

ALGORITHMS FOR CONSTRAINED OPTIMIZATION

J F A de O PANTOJA

A Thesis submitted for the degree of Doctor of
Philosophy

Control Section
Department of Electrical Engineering
Imperial College of Science and Technology
University of London

ABSTRACT

This thesis is divided into two parts.

The first part is concerned with the development of efficient algorithms for solving the general mathematical programming problems. Such problems frequently arise in optimization based design of control system or circuits. The thesis concentrates on a specific class of algorithms which obtain a search direction by solving a quadratic approximation to the original problem and then utilizes an exact penalty function to compute the step length. One of the defects of existing algorithms in this category is that the quadratic approximation to the problem is such that no feasible solution exists. A new quadratic approximation is therefore introduced with the following properties: a feasible solution always exists and the solution of the programme yields a search direction which is a descent direction for the exact penalty function used to compute step length. Using this a new algorithm for constrained optimization is developed and shown to be globally convergent. A simplified version of the algorithm for the convex case is also developed. Finally the special structure of the quadratic approximation is exploited to obtain a simple procedure for updating the penalty parameter. Global convergence of this version of the algorithm is also established.

It is then shown that if a secant procedure is used to generate the approximation to the Hessian of the Lagrangian (employed in the quadratic approximation) the resultant algorithm is globally convergent and has a Q-superlinear rate of convergence. A new version of the B.F.G.S. procedure for approximating the Hessian is proposed

and the resultant algorithm is shown to be globally convergent and have an R-superlinear rate of convergence. Numerical examples illustrate the results.

In the second part, a modified version of the differential dynamic programming (DDP) method, for discrete unconstrained control problems, is introduced and is proven to be a 'step by step' implementation of Newton's method for solving systems of nonlinear equations. Finally, an extension of this method is developed to cope with control constraints. The resultant algorithm is shown, under some assumptions, to be globally convergent and have a quadratic rate of convergence.

List of Contents

ACKNOWLEDGEMENTS	
GLOSSARY OF SYMBOLS	
INTRODUCTION	1
PART I - MATHEMATICAL PROGRAMMING PROBLEMS	12
I.1 Basic Algorithm	12
I.1.1 Introduction	12
I.1.2 The Algorithm	19
I.1.3 Convergence Results	23
I.1.4 Remarks	41
APPENDIX	50
I.2 Convex Inequalities Case	63
I.2.2 Introduction	63
I.2.2 The Algorithm	65
I.2.3 Convergence Results	66
I.2.4 Results	75
APPENDIX	78
I.3 Secant Update Case	79
I.3.1 Introduction	79
I.3.2 Algorithm	83
I.3.3 Global Convergence	85
I.3.4 Rate of Convergence	91
I.3.5 Remarks	96
APPENDIX	116
I.4 Modification of BFGS Update Case	120
I.4.1 Introduction	120
I.4.2 The Algorithms	124
I.4.3 Global Convergence	127
I.4.4 Rate of Convergence	146
I.4.5 Remarks	162
APPENDIX	175

	page
I.5 Numerical Results	200
I.5.1 Introduction	200
I.5.2 Results	205
I.5.3 Remarks	256
I.6 Suggestions for Further Research	258
 PART II - DISCRETE OPTIMAL CONTROL PROBLEMS	 262
II.1 Unconstrained Control Case	262
II.1.1 Introduction	262
II.1.2 Modified D.D.P./Newton's method	267
II.1.3 Results	272
II.1.4 Remarks	284
 II.2 Constrained Control Case	 299
II.2.1 Introduction	299
II.2.2 The Algorithm	307
II.2.3 Global Convergence	310
II.2.4 Rate of Convergence	321
II.2.5 Remarks	334
 APPENDIX	 339
II.3 Suggestions for Further Research	346
 REFERENCES	 349

Aos meus pais,

Doralice e José

e à minha amiga,

Marisia

ACKNOWLEDGEMENTS

I would like to express my most sincere gratitude to my supervisor, Professor D Q Mayne, for his invaluable help and advice during the course of the research that led to this thesis.

I would also like to thank Dr José de Oliveira Pantoja, my father, for his continuous incentive throughout my studies and Professor Miguel Arcangelo Bernardes Filho, my attorney, for keeping stable the 'home front'.

I am greatly indebted to the Fundacao Universidade Federal do Maranhao (UFMA) and Coordenacao do Aperfeicoamento do Pessoal do Ensino Superior (CAPES) for financial support during my stay in England.

I acknowledge the help of Mr A Heunis when implementing the computational routines.

I thank Mrs Doris Abeysekera and Ms Elaine Gascoyne for the painstaking work of typing this thesis

Glossary of Symbols

Besides the symbols defined along the text, we followed the notation below:

Standard Symbols:

ϕ ; ϵ ; ∇ ; $\subset$; $\supset$; $\cup$; $\cap$; $\{\cdot\}$; $=$; $\neq$; $<$; $\leq$; $>$; $\geq$; $\exists$; $\forall$; $+$; $-$; $\times$;
 $/$; ∞ ; $(a,b]$; $[a,b)$; (a,b) ; $[a,b]$; $\|\cdot\|$; $|\cdot|$; $;$; $\rightarrow$; $\Rightarrow$;
 $\Leftrightarrow$; $\triangle$; $\sum_{i=1}^n$; $\sum_{j \in S}$; $<\cdot, \cdot>$;
 p^T ; $\lim_{i \rightarrow \infty}$; $\lim_{r \rightarrow 0_+}$; $(\text{mod}.n)$; $\int_0^1$; $\mathbb{R}$; $\mathbb{R}^n$; $L(\mathbb{R}^n)$; ∇ ; ∇^2 ;
 ∇_x ; ∇_{xx} ; f_x ; f_{xx} ; f_{xu} .

Special Symbols

$f : \mathbb{R}^n \rightarrow \mathbb{R}^m$ - the map f from $\mathbb{R}^n$ into $\mathbb{R}^m$
 $x \mapsto p'(\cdot, B^*, c)$ - the map $p'(x, B^*, c)$
 $\{x_i\}_{i \in I}$ or $\{x_i\}_I$ - the set $\{x_i \in \{x_i\} | i \in I\}$
 $x_i \xrightarrow{I} x^*$ - $\{x_i\}_I$ converging to x^*
 $I - I'$ - the set $\{i \in I | i \notin I'\}$
K-T - short for Kuhn-Tucker
 $\frac{d\bar{z}(0)}{d\Delta_i}$ - derivative of $\bar{z}(\cdot)$ with respect to Δ_i computed at the origin
 $\|\cdot\|_p$ - the ℓ_p -norm on $\mathbb{R}^n$, $1 \leq p \leq \infty$
 $A \times B$ - cartesian product (A, B sets)

INTRODUCTION

This thesis is divided into two parts.

In the first part we deal with iterative methods to solve the general mathematical programming problem (GMP) below;

GMP

$$\begin{array}{ll} \min & f(x) \\ (x \in R^n) & \\ \text{subject to:} & g(x) \leq 0 \\ & h(x) = 0 \end{array}$$

where:

$$f : R^n \rightarrow R, \quad g : R^n \rightarrow R^m, \quad h : R^n \rightarrow R^\ell$$

There is an extremely diverse range of practical applications which require the determination of the 'best' solution to certain mathematically defined problems, either in the general form above or in some special forms as : linear programming, linearly constrained programming, quadratic programming, convex programming, etc. A number of such applications can be found in the proceedings of a conference on 'Optimization in Action' [Dixon (1976)] and many more, of course, in the specialized technical literature of almost every activity in which numerical information is processed (Science, Engineering, Mathematics, Economics, Commerce, etc.)

Although most of these applications are finite dimensional problems, the solution of GMP is also of interest for some infinite dimensional problems. In particular, continuous time optimal control may be put into the GMP form through the use of an appropriate discretization [see Cannon, Cullum and Polak (1970)].

Also, many problems of synthesizing a system from components whose values are only known within a certain tolerance, for example circuit design problems with tolerances and multivariable control design problems, can be cast into the form of a mathematical programming problem where the constraints are infinite dimensional. Then, either by the selection of a finite number of 'worst cases' [Bandler, Liu and Tromp (1976)] or through outer approximations algorithms [Mayne, Polak and Trahan (1979)] with automatic constraint dropping schemes [Gonzaga, Polak (1979)], these problems can be solved as the limit of a sequence of finite dimensional GMP's.

Most of the methods proposed to solve the GMP are iterative; given a start point x_0 they build a sequence $\{x_i\}$ through the general iteration

$$x_{i+1} = x_i + \alpha_i p_i$$

where p_i is a 'search direction' previously obtained by solving an auxiliary sub-problem and α_i is the 'step length' computed by some procedure in order to assure that x_{i+1} satisfies a certain improvement condition.

In fact, the nature of the sub-problem used to determine p_i and the way α_i is computed may be used to characterize diverse types of methods to solve GMP's.

In the so called 'methods of feasible directions' [Zoutendijk (1960)] the sequence $\{x_i\}$ remains feasible, that is : x_i satisfies the constraints of the GMP for all i . The search directions are in general obtained by minimizing, over the feasible set of GMP, a function which approximates the objective function $f(\cdot)$ about the current

point x_i , either linearly, as in the conditional gradient method, or quadratically, as in the Newton's method for constrained optimization [Levitin and Polyak (1966)].

Since in these methods x_i remains feasible (for $\alpha_i \leq \bar{\alpha} < \infty$, say) the step length is chosen to ensure a strictly monotonic decrease of the objective function so that under reasonable assumptions a convergent sequence, whose limit point solves GMP, is obtained [see Ortega and Rheinboldt (1970) for a discussion of convergence of 'descent methods']

On the other hand, nonlinear equality constraints can make the solution of the sub-problems used to determine p_i quite cumbersome and unless the equality constraint functions of GMP are affine such methods do not have much appeal.

Simpler sub-problems can be achieved if the constraints are linearized about the current point, although the objective function must then be modified to include information about the curvature of the constraints [Beale (1967)] if superlinear convergence is required.

In particular, the use of second order approximation to the objective function of GMP and linearization of the constraints [see Wilson (1963)] lead to a quadratic problem which, because it can be solved by many well known algorithms in a finite number of iterations [see Fletcher (1971)], is a suitable search direction sub-problem.

Algorithms of this class have been shown to be locally convergent with a quadratic rate of convergence [see Robinson (1974)] but their global stabilization posed a new question. Because they generate sequences $\{x_i\}$ which are not necessarily feasible, it may be difficult to decide whether, for a certain value of α , the iteration

$$x_{i+1} = x_i + \alpha p_i$$

leads to a improvement.

Alternatively, the search direction can be found by the unconstrained minimization of an appropriate function, as in the multiplier method of Glad and Polak (1969) and exact penalty function methods [see Pietrzykowski (1969)].

Although these methods have an advantage over the sequential unconstrained minimization of the classical exterior penalty function [see Fiacco and McCormick (1968)] in the sense that the penalty parameter can be kept bounded from above rather than going to infinity, their success depends greatly on the particular unconstrained optimization algorithm eventually chosen for minimizing the seldom simple and often nondifferentiable augmented objective function.

Moreover, despite the fact that the penalty parameter required is finite it has to be estimated. Repetitive unconstrained minimization with an increasing penalty parameter may lead to unnecessary waste of time; moreover, as the penalty parameter increases, the unconstrained minimization becomes more difficult.

Nevertheless, exact penalty functions are a useful mechanism to order unfeasible points, thereby permitting a suitable choice of step length to be made in algorithms which generate a search direction by solving an approximation to GMP. This idea was employed by Han (1977) and later, but independently, by Mayne and Maratos (1979).

Recently the use of Quasi-Newton methods to determine the search direction and an exact penalty function to determine step length has been a common feature to many algorithms [see Powell (1978), Chamberlain (1980), Yamashita (1982), Mayne and Polak (1982)].

It is interesting to note in these methods the double role of the penalty parameter. First, it is necessary that the final value of this

parameter should be large enough (although finite) to enforce equivalence with the original problem (that is, to ensure that a strict local minimum of the penalty function is a strict local minimum of GMP). Second, at each iteration the parameter must be chosen so that the search direction is a descent direction of the exact penalty function; in this respect these algorithms differ from the pioneering works of Conn (1973); Conn and Pietrzykowski (1977) where the search direction is chosen to be a descent direction for the penalty function with the penalty parameter being selected a priori.

Much of the literature on these methods either assumes that a suitable value of the penalty parameter is known a priori (in practice many computations with different choices of the penalty parameter may be necessary) or employs heuristic procedures for iteratively updating the parameter.

Mayne and Polak (1982) suggested an automatic rule for the choice of the parameter which ensures ultimate equivalence with the original problem but, except in the neighbourhood of a solution point of GMP, does not guarantee that the direction obtained from the quadratic problem is a descent direction of the penalty function; in this case an alternative first order search direction is generated. This same alternative is employed by the last two authors to cope with inconsistent constraints in the quadratic problem, while Powell (1982) suggests that, if the linearized constraints are inconsistent, the objective function of the quadratic problem should be minimized over an extended feasible set. Besides requiring the addition of further routines to the main algorithm, both remedies suffer from the drawback of being activated only after an unsuccessful attempt to solve the quadratic problem.

The rate of convergence of this class of algorithms (indeed, the rate of convergence of Quasi-Newton methods to unconstrained minimization) depends on the choice of the matrices of the quadratic problem and also on the achievement of an asymptotic step length of unity. Much research has been done on the problem of choosing the sequence of matrices so that, if the step length is of unity for all i sufficiently large, then the rate of the convergence is superlinear.

The question of unitary step length is also important. Maratos (1978) has pointed out that the use of a line search strategy to guarantee a decrease of the penalty function may lead to a truncation of the step length no matter how close the current point is to a solution. This difficulty occurs because of the linearization of the constraints and may prevent the algorithm from achieving a superlinear rate of convergence.

To avoid this difficulty, Chamberlain et al (1981) suggested a 'watchdog technique' where in some iterations the criterion of reducing the penalty function is relaxed so that unitary step length can be used. The technique requires some heuristic choice of parameters and does not ensure the decrease of the penalty function at all iterations. As a result, many iterations may be required before the algorithm converges. A mechanism to be preferred is employed by Mayne and Polak (1982). They avoid the eventual occurrence of the so called 'Maratos effect' by replacing the linear search strategy.

$$x_{i+1} = x_i + \alpha_i p_i$$

by the search arc

$$x_{i+1} = x_i + \alpha_i p_i + \alpha_i^2 \tilde{p}_i$$

where $\tilde{p}_i$ is chosen so that the above arc satisfies the constraints to second order; the calculation of second order derivatives is avoided.

All the algorithms presented in Part I of this thesis solve a quadratic problem to find the search direction and compute the step length using an exact penalty function.

In the first chapter a special quadratic problem is introduced which always has a solution and generates a descent direction of the penalty function independently on the value of the penalty parameter. An algorithm is then developed and it is proved to be, under reasonable assumptions, globally convergent.

In the second chapter a simplified version of this Basic Algorithm is suggested to solve problems with convex inequalities. Here the special structure of the quadratic problem is also used to develop a simpler procedure to update the penalty parameter.

In the third chapter the Basic Algorithm is modified so that a 'secant update' can be used to construct the sequence of matrices. Under fairly mild assumptions the resultant algorithm is proved to be globally convergent. Furthermore, for all i large enough a step length of unity is obtained; it can then be shown that the algorithm generates sequences with a superlinear rate of convergence.

In the fourth chapter a new modification of the classical BFGS update is introduced to deal with constrained problems. A modified version of the algorithm presented in the first chapter is also developed and, under reasonable assumptions, proved to be globally convergent. Also, the step length is shown to be of unity for i sufficiently large and a superlinear rate of convergence is shown to be achieved.

Some numerical results are given in the fifth chapter and in the last chapter of Part I suggestions are given for further research.

The second part of this thesis deals with the N-stages discrete optimal control problem (DCP) below :

DCP

$$\min_{u \in R^{m \times N}} \sum_{i=0}^{N-1} T_i(x_i, u_i) + \phi(x_N)$$

with (x_i, u_i) connected by:

$$x_{i+1} = f_i(x_i, u_i), \quad i = 0 \dots N-1$$

and satisfying :

$$g_i^s(u_i) \leq 0, \quad i = 1, \dots, N-1$$

$$s = 1, \dots, \ell_i$$

$$T_i : R^n \times R^m \rightarrow R$$

$$\phi : R^n \rightarrow R$$

$$f_i : R^n \times R^m \rightarrow R^n \quad ; \quad g_i^s : R^m \rightarrow R, \quad \forall i, s$$

$$x_0 = x^0, \text{ given}$$

$$x_i \in R^n \quad (\text{the state vector})$$

$$u_i \in R^m \quad (\text{the control vector}).$$

The solution of this fixed-time optimal control problem is relevant to problems with free-end-time and problems with constrained terminal state (that is $x_N \in X_N \subset R^n$) as they can be solved as the limit of a sequence of fixed-time problems with additional penalty terms [see Canon, Cullum and Polak (1970) and Mayne and Polak (1980)].

Also note that when using digital computers to solve continuous optimal control problems we may encounter a DCP through the use of some form of discretization [see Canon, Cullum and Polak (1970)].

Now, the DCP can always be treated as a GMP in $R^{(m+n)N}$ so that any method to solve mathematical programming problems could be applied to solve the DCP. However, this may not be a wise choice because the dimension of the GMP would be dependent on the number of stages N and consequently it could be forbiddingly high.

On the other hand, specific necessary conditions of optimality have been directly developed for control problems, either through the calculus of variations, as in Pontryagin's maximum principle [see Halkin (1966)] or through dynamic programming [see Bellman (1957)].

The use of necessary conditions expressed in the form of a maximum principle to construct a method for solving discrete control problems requires the solution of a system of coupled equations involving state variable, control variable and costates vectors; this may be rather difficult to solve. In particular, the presence of control constraints and/or severe nonlinear dynamic equations makes it impossible to express, a priori, the optimal control in terms of optimal state and costate making this approach highly unattractive.

Dynamic programming provides a very general approach with a broad range of applications [see Bellman and Dreyfus (1962)], but the backward procedure implied by the 'principle of optimality' requires that, at each stage, a large number of minimizations must be performed (with different state vectors). Such procedure also requires a great deal of storage and, consequently, the application of conventional dynamic programming to DCP remains a formidable task, unless the dimension of the state vector is smaller than, say, 3.

A mechanism to avoid this difficulty was provided by Mayne (1966) with the introduction of the Differential Dynamic Programming (DDP) method to solve optimal control problems. The idea behind this approach is

the following: given state and control nominal sequences the DPP method compares them with neighbouring sequences and selects that neighbouring sequence which yields an optimal decrease in cost. The new control and state sequences become the nominal sequences for the next iteration. Since at each iteration attention is confined to a suitable small neighbourhood of the current nominal state sequence, it is possible to approximate the improved cost function by a truncated Taylor series expansion. This results in a tremendous reduction of the storage required, when compared with the conventional dynamic programming method. Among other applications of this technique, Jacobson and Mayne (1970) developed a second order algorithm for unconstrained discrete control problems.

The relative simplicity of this algorithm, which at each stage solves a sub-problem of the same dimension as the control space, makes it quite appealing.

Moreover, although not proven, practical applications of the algorithm suggested a quadratic rate of convergence. Actually, the fact that it provides one step convergence for problems with quadratic objective function and linear dynamics (LQP) hints at some link between the DDP method and Newton's method to solve nonlinear equations.

Another problem is the extension of the DDP technique to constrained discrete control problems.

In the first chapter of Part II a modified version of the DDP method for unconstrained discrete control problems is introduced and is proven to be identical to Newton's method. That is, given a nominal control sequence both methods will compute the same successor sequence. It is afterwards shown that the new version coincides with the original whenever the dynamics are linear. Hence, the results for LQP may be thought of as a special case of the fact that the original version coincides with Newton's method if the dynamical equations are linear. The last result

of this chapter shows that the second-order algorithm suggested by Jacobson and Mayne (1970) for unconstrained discrete control problems indeed has a quadratic rate of convergence.

In the second chapter of Part II an algorithm to solve the constrained discrete control problem is introduced and, under some assumptions, is proven to be globally convergent with a quadratic rate.

As in Part I, in the last chapter of Part II further research is suggested.

The references come next, listed in alphabetical order by author and then in chronological order.

It is useful to indicate now the general numbering scheme adopted throughout the thesis.

The two parts (numbered I and II) are each divided into chapters (numbered from I.1. to I.6 and from II.1 to II.4) which are divided into the several sections (denoted by numbers of the type I.2.1, I.2.2, etc.) and, for some chapters, an appendix (AP.I.2 denoting the appendix of Chapter 2 in Part I, etc.).

Lemmas, propositions and theorems are indicated by capital letters within each chapter regardless the section (e.g., I.3.A) and, when required, expressions obtained during the proof of a particular proposition (or lemma or theorem) will be indicated by low-key roman numbers (i, ii, etc.). That is, II.2.A.xi stands for expression (xi), Proposition A, Chapter 2 of Part II.

Expressions outside the proofs are indicated by capital roman numbers but now only within chapters. That is, II.1.V stands for expression V in the first chapter of Part II.

Finally, the results and definitions of the literature which are given in the appendix are numbered consecutively inside each appendix which, unless expressly stated, is assumed to

correspond to the chapter where the result or definition was referred. That is, AP.1 denotes the result (or definition) 1 in the appendix at the end of the current chapter.

I.1 BASIC ALGORITHM

I.1.1 Introduction

The use of exact penalty functions to compute step length along a direction obtained by solving a quadratic programming sub-problem, is a common feature of many algorithms lately suggested for solving the general mathematical programming problem (GMP) below :

GMP

$$\begin{aligned} \min_{x \in \mathbb{R}^n} \quad & f(x) \\ \text{s.t.} \quad & \\ & g(x) \leq 0 \\ & h(x) = 0 \end{aligned}$$

where:

$$f: \mathbb{R}^n \rightarrow \mathbb{R} \quad g: \mathbb{R}^n \rightarrow \mathbb{R}^m \quad h: \mathbb{R}^n \rightarrow \mathbb{R}^\ell .$$

The quadratic problem that have been generally used is of the form:

$$\begin{aligned} \text{QP}_i^0 \\ \min_{p \in \mathbb{R}^n} \quad & f_x(x_i)p + p^T B_i p \\ \text{s.t.} \quad & \\ & g^j(x_i) + g_x^j(x_i)p \leq 0, \quad j \in \underline{m} \\ & h^j(x_i) + h_x^j(x_i)p = 0, \quad j \in \underline{\ell} \end{aligned}$$

where:

$x_i \in R^n$ denotes the current point at iteration i

B_i is a real $n \times n$ matrix

$\underline{m}, \underline{\ell}$ denote, respectively, the sets $\{1, 2, \dots, m\}$
and $\{1, 2, \dots, \ell\}$

Calling p_i^0 the solution of QP_i^0 , the next point x_{i+1} of the iteration process is set as :

$$x_{i+1} = x_i + \alpha_i p_i^0$$

where $\alpha_i \in (0, 1]$ is chosen by some procedure which ensures the satisfaction of the following condition :

I.1.I

$$\gamma(x_i + \alpha_i p_i^0, c_i) < \bar{\gamma}(x_i, c_i)$$

where $\gamma : R^n \times R \rightarrow R$ belongs to the following class of penalty functions studied in Han and Mangasarian (1979) :

I.1.II

$$\gamma(x, c) = f(x) + c Q(\|g(x)_+, h(x)\|)$$

where:

c is a nonnegative real number

$$g^j(x)_+ \triangleq \max\{0, g^j(x)\}$$

$\|\cdot\|$ is any fixed vector norm in $R^{m+\ell}$

and

Q is some function from the nonnegative real line into itself with the following properties:

I.1.III

$$Q(0) = 0$$

$$Q(r) \geq 0 \quad \text{for } r > 0$$

$$\infty > Q'(0_+) \triangleq \lim_{r \rightarrow 0_+} \frac{Q(r) - Q(0)}{r} > 0$$

A condition to assure that $\gamma(x, c)$ has a minimum at a minimum of GMP for all sufficiently large but finite values of c is due to Pietrzykowski (1969) and requires the linear independence of the gradients of all equality constraints and of those inequalities satisfied as equalities at the point being considered.

Note that the classical exact penalty function

$$\gamma_1(x, c) \triangleq f(x) + c \left\{ \sum_{j=1}^m g^j(x)_+ + \sum_{j=1}^{\ell} |h^j(x)| \right\}$$

can be obtained from I.1.II by setting $Q(r) = r$ and using the 1-norm.

On the other hand, the classical exterior penalty function

$$\gamma_2(x, c) \triangleq f(x) + c \left\{ \sum_{j=1}^m [g^j(x)_+]^2 + \sum_{j=1}^{\ell} [h^j(x)]^2 \right\}$$

corresponds to setting $Q(r) = r^2$ and using the 2-norm, but does not satisfy the third part of I.1.III as in this case

$$Q'(0_+) = 0$$

which is not a surprise because it is well known that for the classical exterior penalty function the penalty parameter goes to infinity.

In fact, if the second order sufficiency conditions (see AP.1) are satisfied for GMP at some $(\bar{x}, \bar{\lambda}, \bar{\mu}) \in \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^L$ then a result due to Han and Mangasarian (1979) shows that $\bar{x}$ will be a strict local minimum of $\gamma(x, c)$ for all $c > \bar{c}$ with $\bar{c}$ being defined by:

I.1.IV

$$\bar{c} = \frac{\|(\bar{\lambda}, \bar{\mu})\|'}{Q'(0)}$$

where:

$\|\cdot\|'$ is the dual norm of the vector norm used in II.1.II.

Now, it is readily seen that two problems may occur when generating the search direction through the quadratic problem QP_i^0 :

I.1.V

The solution p_i^0 may not be a descent direction to the penalty function $\gamma(x_i, c_i)$

I.1.VI

QP_i^0 may have inconsistent constraints.

The occurrence of I.1.V. depends basically on two factors:

- The nature of the matrix B_i in QP_i^0
- The current value c_i of the penalty parameter.

Actually, the definition of what is a descent direction may vary with the type of exact penalty function being used (some are not differentiable).

The above factors are combined in the following sufficient condition :

"If,

$$(p_i^0)^T B_i p_i^0 > 0$$

and

$$c_i > \|(\lambda_i^0, \mu_i^0)\|' .$$

Then,

$$p_i^0 \text{ is a descent direction to } \gamma(x_i, c_i)$$

where

(λ_i^0, μ_i^0) are the multipliers associated with p_i^0 at the solution of QP_i^0 .

A convenient way of ensuring $(p_i^0)^T B_i p_i^0 > 0$ is to assume positive definiteness of the sequence $\{B_i\}$, as in Han (1977), or to construct the sequence $\{B_i\}$ on such a way that positive definiteness is indeed enforced, as in Powell (1978).

The condition $c_i > \|(\lambda_i^0, \mu_i^0)\|'$ presents an interesting question. Large parameter values cause the penalty function to have steep sides at the boundary of the feasible region so that the satisfaction of the condition

$$\gamma(x_{i+1}, c_i) - \gamma(x_i, c_i) < 0$$

may lead to excessively small step lengths.

Also, if one tries to satisfy $c_i > \|(\lambda_i^0, \mu_i^0)\|'$ by setting $c_i = c$ large enough or by increasing c_i whenever this condition is violated, one may end working with far larger values of c_i than required to ensure the compatibility of $\gamma(\cdot, c_i)$ with GMP (see I.1.D). Moreover,

the multipliers (λ_i^0, μ_i^0) may vary discontinuously and even be unbounded so that setting $c_i > \|(\lambda_i^0, \mu_i^0)\|$ may bring convergence problems [see Chamberlain (1979)] .

Inconsistent constraints in QP_i^0 have been dealt with in different ways.

Han (1977) simply rules out I.1.VI by assuming that the quadratic problem always has a Kuhn-Tucker point [see Theorem 3.2 in Han (1977)] or makes some assumptions, for instance $g(\cdot)$ is convex [see Theorem 3.2 in Han (1977)], which are sufficient to assure that I.1.VI never occurs but that also clearly reduce the range of application of his algorithm.

Powell (1982) suggests that, if I.1.VI occurs, the objective function should be minimized satisfying the following conditions:

$$\begin{aligned}\xi h^j(x_i) + h_x^j(x_i)p &= 0, & j &= 1, 2, \dots, \ell \\ \xi g^j(x_i) + g_x^j(x_i)p &\leq 0, & j &\in V_i \\ g^j(x_i) + g_x^j(x_i)p &\geq 0, & j &\in S_i\end{aligned}$$

where:

$$\bar{m} = V_i \cup S_i$$

V_i is the set of indices of the inequalities that are violated at x_i

ξ is a constant chosen from the interval $[0, \bar{\xi}]$, with $\bar{\xi}$ being the largest value in $[0, 1]$ such that the above conditions are satisfied.

This procedure may require many computations not only to find $\bar{\xi}$ but also in solving a quadratic problem with inconsistent constraints.

This last drawback is shared by the remedy proposed by Mayne and Polak (1982) which consists of generating an alternative search direction, if I.1.VI occurs, through the following unconstrained minimization:

I.1.VII

$$\min_{p \in \mathbb{R}^n} \{1/2 p^T \bar{H} p + \theta(x, c, p)\}$$

where

$\bar{H}$ is any positive definite matrix (e.g. ηI , $\eta > 0$)

$$\theta(x, c, p) \triangleq f_x(x)p + c(\hat{\psi}(x, p) - \psi(x))$$

with:

$$\psi(x) \triangleq \max\{g^j(x), j \in \underline{m}; |h^j(x)|, j \in \underline{\ell}; 0\}$$

and

$$\hat{\psi}(x, p) \triangleq \max\{g^j(x) + g_x^j(x)p, j \in \underline{m}; |h^j(x) + h_x^j(x)p|, j \in \underline{\ell}; 0\}.$$

In this chapter we present an algorithm which generates the search direction through a quadratic problem which always has a solution which is a descent direction to $\gamma(\cdot, c_i)$ (regardless of the value of c_i). This allows us to update the penalty parameter with the sole objective of ultimately satisfying I.1.IV. With fairly mild assumptions it is proved that any accumulation point of a bounded infinite sequence constructed by the algorithm satisfies the Kuhn-Tucker conditions for GMP.

If second order sufficiency conditions are satisfied at any Kuhn-Tucker point of GMP we prove a much stronger result, that is:

$$x_i \rightarrow x^* \text{ (strict local minimum of GMP).}$$

I.1.2 The Algorithm

Consider the general mathematical programming problem (GMP) defined in the previous section.

Let $\psi: R^n \rightarrow R^m$ be defined as :

$$\psi(x) = \max\{g^j(x), j \in \underline{m}; |h^j(x)|, j \in \underline{\ell}\}.$$

Let $\gamma: R^n \times R \rightarrow R$ be defined as :

$$\gamma(x, c) \triangleq f(x) + \psi(x)c$$

where $c > 0$ is a penalty parameter.

For each iteration i let $(p_i, w_i, \lambda_i, \mu_i, v_i)$ denote the Kuhn-Tucker point for the quadratic programming problem below:

QP_i^w

$$\min_{(p, w) \in R^{n+1}} f_x(x_i)p + 1/2 p^T B_i p + c_i w$$

subject:

$$\begin{aligned} g^j(x_i) + g_x^j(x_i)p &\leq w, \quad j \in \underline{m} \\ |h^j(x_i) + h_x^j(x_i)p| &\leq w, \quad j \in \underline{\ell} \end{aligned}$$

$$w \geq 0$$

where,

x_i is the current point

c_i is the current value of the penalty parameter

$\lambda_i \in R^m$ is the multiplier associated with the first m inequalities

$\mu_i \in R^\ell$ is the multiplier associated with the next
 ℓ inequalities

$v_i \in R$ is the multiplier of the last inequality
 $p \in R^n$ and $w \in R$.

Let $\bar{\lambda}: R^n \rightarrow R$ and $\bar{\mu}: R^n \rightarrow R^m$ be defined by:

$$(\bar{\lambda}(x), \bar{\mu}(x)) \triangleq \arg \min_{(\lambda, \mu) \in R^m \times R^\ell} \{ \|\nabla f(x) + g_x(x)^T \lambda + h_x(x)^T \mu\|^2 + \sum_{j \in \underline{m}} (\psi(x) - g^j(x))^2 (\lambda^j)^2 + \sum_{j \in \underline{\ell}} (\psi(x) - |h^j(x)|^2) (\mu^j)^2 \}$$

Let $\bar{c}: R^n \rightarrow R$ be defined by:

$$\bar{c}(x) \triangleq \{ \sum_{j \in \underline{m}} |\bar{\lambda}^j(x)| + \sum_{j \in \underline{\ell}} |\bar{\mu}^j(x)| + b \}, b > 0, \text{ "small" }.$$

Let $\hat{\psi}: R^n \times R^n \rightarrow R$ be defined by:

$$\hat{\psi}(x, p) \triangleq \max\{g^j(x) + g_x^j(x)p, j \in \underline{m}; |h^j(x) + h_x^j(x)p|, j \in \underline{\ell}\}.$$

Let $\hat{\gamma}: R^n \times R^n \times R \rightarrow R$ be defined by:

$$\hat{\gamma}(x, p, c) \triangleq f(x) + f_x(x)p + c\hat{\psi}(x, p).$$

Let $T: R^n \rightarrow R$ be defined by:

$$T(x) \triangleq [\psi(x) + \|\nabla f(x) + g_x(x)^T \bar{\lambda}_+(x) + h_x(x)^T \bar{\mu}(x)\|^2]^2$$

where,

$\bar{\lambda}_+(x)$ is the vector $\in R^m$ whose j^{th} component is:
 $\bar{\lambda}_+^j(x) = \max\{0; \bar{\lambda}^j(x)\}.$

Let $\theta : \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R} \rightarrow \mathbb{R}$ be defined by :

$$\theta(x,p,c) \triangleq \hat{\gamma}(x,p,c) - \gamma(x,c).$$

We say that $p' \in \mathbb{R}^n$ is a 'descent direction' to the exact penalty function $\gamma(x,c)$ if:

$$\theta(x,p',c) < 0.$$

Finally, for each i , let $\tilde{p}_i$ denote the minimum norm solution, if a solution with norm smaller than $\|p_i\|$ exists, of the following system :

$$g^j(x_i + p_i) + g_x^j(x_i)p = 0, \quad j \in \hat{I}_i$$

$$h^j(x_i + p_i) + h_x^j(x_i)p = 0, \quad j \in \underline{\ell}$$

where,

$$\hat{I}_i \triangleq \{j \in \underline{m} \mid \lambda_i^j > 0\}.$$

If the minimum norm solution of this system has norm greater than $\|p_i\|$, then $\tilde{p}_i$ is set equal to zero.

We are now able to state the Basic Algorithm.

Basic Algorithm

Data : $x_1 \in \mathbb{R}^n$; B_1 ($n \times n$ positive definite matrix); $c_0 \in (0, \infty)$

$\beta, \varepsilon \in (0, 1)$; $\delta \in (1, 2)$

Step 0 : Set $i = 1$

Step 1 : If $c_{i-1} \geq \bar{c}(x_i)$, set $c_i = c_{i-1}$

Else, set $c_i = \max\{\delta c_{i-1}; \bar{c}(x_i)\}$

Step 2 : Find (p_i, w_i) , the solution of QP_i^w

Step 3 : If :

$$T(x_i) \leq \varepsilon$$

$$w_i = 0$$

Then:

compute $\tilde{p}_i$

Else:

$$\text{set } \tilde{p}_i = 0$$

Step 4: Compute k_i , the first integer in the sequence $\{0, 1, \dots\}$

satisfying:

$$\gamma(x_i + \beta^{k_i} p_i + \beta^{2k_i} \tilde{p}_i, c_i) - \gamma(x_i, c_i) \leq \beta^{k_i} \theta(x_i, p_i, c_i) / 8$$

Step 5: Set $x_{i+1} = x_i + \beta^{k_i} p_i + \beta^{2k_i} \tilde{p}_i$

Step 6 : Update B_i to B_{i+1}

Step : 7 Set $i = i+1$; Go to Step 1.

Before proceeding to establish convergence of the above algorithm it is perhaps interesting to point out that the problem solved at Step 2 is indeed a quadratic programming problem because the ℓ nonlinear inequalities of QP_i^w ,

$$|h^j(x_i) + h_x^j(x_i)p| \leq w, \quad j \in \underline{\ell}$$

are equivalent to the 2ℓ linear inequalities below :

$$h^j(x_i) + h_x^j(x_i)p \leq w, \quad j \in \underline{\ell}$$

and

$$-h^j(x_i) - h_x^j(x_i)p \leq w, \quad j \in \underline{\ell}.$$

Note also that, in contrast to GPM and QP_i^0 , there exist no equality constraints in QP_i^w . Therefore all the multipliers $\lambda_i^j, \mu_i^j, v_i$, associated with (p_i, w_i) at the solution of QP_i^w , are non-negative, for all i and j .

I.1.3 Convergence Results

In the analysis of the global convergence properties of the algorithm we will initially assume that the three following conditions are satisfied:

H1: The functions f, g, h are all continuously differentiable.

H2: For all $x \in R^n$ the vectors $\{\nabla g^j(x), j \in I(x); \nabla h^j(x), j \in I_\ell(x)\}$ are linearly independent, where :

$$I(x) \triangleq \{j \in \underline{m} \mid g^j(x) = \psi(x)\} \text{ and } I_\ell(x) \triangleq \{j \in \underline{\ell} \mid |h^j(x)| = \psi(x)\}$$

H3: For all i , $B_i \in L^+ \subset L(R^n)$ where :

$$L^+ \triangleq \{B \in L(R^n) \mid \text{there exist } m', M \text{ positive s.t. :}$$

$$m' \|y\|^2 \leq y^T B y \leq M \|y\|^2, \quad \forall y \in R^n \}.$$

Proposition I.1.A

For all $i = 1, 2, 3 \dots$

$$\theta(x_i, p_i, c_i) \leq -p_i^T B_i p_i - v_i \psi(x_i)$$

Proof:

Consider the ℓ nonlinear inequalities of QP_i^W as 2ℓ linear inequalities and let t_i^j and s_i^j denote the multipliers respectively associated with,

$$h^j(x_i) + h_x^j(x_i)p \leq w, \quad j \in \underline{\ell}$$

and

$$-h^j(x_i) - h_x^j(x_i)p \leq w, \quad j \in \underline{\ell}.$$

At the solution of QP_i^W the following Kuhn-Tucker conditions are satisfied, in addition to the constraints :

$$\begin{aligned} n+1 \text{ equations} \quad & \left(\begin{array}{c} \nabla f(x_i) + B_i p_i \\ c_i \end{array} \right) + \sum_{j \in \underline{m}} \lambda_i^j \left(\begin{array}{c} \nabla g^j(x_i) \\ -1 \end{array} \right) + \sum_{j \in \underline{\ell}} t_i^j \left(\begin{array}{c} \nabla h^j(x_i) \\ -1 \end{array} \right) \\ & + \sum_{j \in \underline{\ell}} s_i^j \left(\begin{array}{c} -\nabla h^j(x_i) \\ -1 \end{array} \right) + v_i \left(\begin{array}{c} 0 \\ \vdots \\ 0 \\ -1 \end{array} \right) = \left(\begin{array}{c} 0 \\ \vdots \\ 0 \\ 0 \end{array} \right) \end{aligned}$$

$$\lambda_i^j (g^j(x_i) + g_x^j(x_i)p_i - w_i) = 0, \quad j \in \underline{m}$$

$$t_i^j (h^j(x_i) + h_x^j(x_i)p_i - w_i) = 0, \quad j \in \underline{\ell}$$

$$s_i^j (-h^j(x_i) - h_x^j(x_i)p_i - w_i) = 0, \quad j \in \underline{\ell}$$

$$v_i w_i = 0$$

$$\lambda_i^j \geq 0, \quad t_i^j \geq 0, \quad s_i^j \geq 0, \quad v_i \geq 0, \quad \forall i, j$$

Now,

$$\hat{\gamma}(x_i, p_i, c_i) - \gamma(x_i, c_i) = f_x(x_i)p_i + c_i(\hat{\psi}(x_i, p_i) - \psi(x_i))$$

But,

$$0 \leq \hat{\psi}(x_i, p_i) \leq w_i$$

and

$$\begin{aligned} f_x(x_i)p_i &= -p_i^T B_i p_i - \sum_{j \in \underline{m}} \langle \lambda_i^j g_x^j(x_i), p_i \rangle - \sum_{j \in \underline{\ell}} \langle t_i^j h_x^j(x_i), p_i \rangle \\ &\quad - \sum_{j \in \underline{\ell}} \langle s_i^j [-h_x^j(x_i)], p_i \rangle. \end{aligned}$$

As the constraints are satisfied we have :

$$g_x^j(x_i)p_i \leq w_i - g^j(x_i), \quad j \in \underline{m}$$

$$h_x^j(x_i)p_i \leq w_i - h^j(x_i), \quad j \in \underline{\ell}$$

and

$$-h_x^j(x_i)p_i \leq w_i + h^j(x_i), \quad j \in \underline{\ell}.$$

Then,

$$\begin{aligned} \hat{\gamma}(x_i, p_i, c_i) - \gamma(x_i, c_i) &\leq -p_i^T B_i p_i + \sum_{j \in \underline{m}} \lambda_i^j g^j(x_i) + \sum_{j \in \underline{\ell}} t_i^j h^j(x_i) \\ &\quad + \sum_{j \in \underline{\ell}} s_i^j [-h^j(x_i)] \\ &\quad - w_i \left[\sum_{j \in \underline{m}} \lambda_i^j + \sum_{j \in \underline{\ell}} t_i^j + \sum_{j \in \underline{\ell}} s_i^j \right] + c_i (w_i - \psi(x_i)). \end{aligned}$$

But, from the $(n+1)^{\text{th}}$ equation,

$$(i) \quad c_i = \sum_{j \in \underline{m}} \lambda_i^j + \sum_{j \in \underline{\ell}} t_i^j + \sum_{j \in \underline{\ell}} s_i^j + v_i$$

then,

$$\begin{aligned} \hat{\gamma}(x_i, p_i, c_i) - \gamma(x_i, c_i) &\leq -p_i^T B_i p_i + \sum_{j \in \underline{m}} \lambda_i^j g^j(x_i) + \sum_{j \in \underline{\ell}} t_i^j h^j(x_i) + \\ &\quad + \sum_{j \in \underline{\ell}} s_i^j [-h^j(x_i)] \\ &\quad - w_i (c_i - v_i) + c_i (w_i - \psi(x_i)). \end{aligned}$$

But, by the complementarity condition, $w_i v_i = 0$.

Also,

$$\begin{aligned} g^j(x_i) &\leq \psi(x_i) \quad , \quad \forall j \in \underline{m} \\ h^j(x_i) &\leq \psi(x_i) \quad , \quad \forall j \in \underline{\ell} \\ -h^j(x_i) &\leq \psi(x_i) \quad , \quad \forall j \in \underline{\ell} . \end{aligned}$$

Then,

$$\begin{aligned} \hat{\gamma}(x_i, p_i, c_i) - \gamma(x_i, c_i) &\leq -p_i^T B_i p_i + \psi(x_i) \left[\sum_{j \in \underline{m}} \lambda_i^j + \sum_{j \in \underline{\ell}} t_i^j + \sum_{j \in \underline{\ell}} s_i^j \right] \\ &\quad - c_i \psi(x_i) \end{aligned}$$

or,

$$\hat{\gamma}(x_i, p_i, c_i) - \gamma(x_i, c_i) \leq -p_i^T B_i p_i - v_i \psi(x_i) \quad . \quad \square$$

Now, denote by $p'(x, B, c)$ the map that leads to p_i and by $\tilde{p}(x, p)$ the map that leads to $\tilde{p}_i$, i.e.

$$p_i \triangleq p'(x_i, B_i, c_i)$$

and

$$\tilde{p}_i \triangleq \tilde{p}(x_i, p_i) \quad .$$

The next proposition is a result due to Mayne and Polak (1978)

Proposition I.1.B

Let $\bar{x}$ be any point in R^n and let N be a compact neighbourhood of $\bar{x}$.

Then, there exists a function

$$(\alpha, x, p', p) \rightarrow \emptyset(\alpha, x, p', p)$$

such that

$$|\gamma(x + \alpha p' + \alpha^2 \tilde{p}, c) - \hat{\gamma}(x, \alpha p' + \alpha^2 \tilde{p}, c)| \leq \alpha \emptyset(\alpha, x, p', \tilde{p})$$

and

$$\emptyset(\alpha, x, p', \tilde{p}) \rightarrow 0 \text{ uniformly in}$$

$$(x, p', \tilde{p}) \in N \times \beta_k(0) \times \beta_k(0), \text{ as } \alpha \rightarrow 0$$

where, k and c are both arbitrary positive constants and $\beta_k(0)$ is the compact ball $\subset R^n$ with radius k centered at the origin.

Proof:

See Proposition A.3 in Mayne and Polak (1978).

Proposition I.1.C

For all $c > 0$ and all $B \in L^+$
 the map $x \rightarrow p'(x, B, c)$ is continuous.

Proof:

See appendix at the end of this chapter.

Proposition I.1.D

Consider the problem below:

$$\tilde{\theta}(x, m, c) = \min_{p \in \mathbb{R}^n} \{ (m/2) \|p\|^2 + \theta(x, p, c) \}$$

If $\tilde{\theta}(x, m_1, c_1) = 0$, for some $m_1 > 0$ and $c_1 > 0$

Then,

$$\tilde{\theta}(x, m, c) = 0, \text{ for all } m > 0 \text{ and } c \geq c_1.$$

Proof:

See appendix at the end of this chapter.

Corollary

Consider the problem below:

$$\theta_1(x, B, c) = \min_{p \in \mathbb{R}^n} \{ 1/2 p^T B p + \theta(x, p, c) \}.$$

If, for some $\bar{x} \in \mathbb{R}^n$, $\bar{c} > 0$ and $B_1 \in L^+$

$$\theta_1(\bar{x}, B_1, \bar{c}) = 0.$$

Then,

$$\theta_1(\bar{x}, B, \bar{c}) = 0, \text{ for all } B \in L^+$$

Proof:

For all $B \in L^+$ we know that,

$$m' \|p\|^2 \leq p^T B p \leq M \|p\|^2, \text{ for all } p \in \mathbb{R}^n.$$

Then, for an arbitrary $B \in L^+$ we have :

$$(i) \quad \tilde{\theta}(\bar{x}, m', \bar{c}) \leq \theta_1(\bar{x}, B, \bar{c}) \leq \tilde{\theta}(\bar{x}, M, \bar{c}).$$

Now, by hypothesis $B_1 \in L^+$ and $\theta_1(\bar{x}, B_1, \bar{c}) = 0$.

Hence, from (i) and the fact that $\tilde{\theta}(x, m, c) \leq 0$, for all $x \in \mathbb{R}^n$ and all positive m and c , we have :

$$\tilde{\theta}(\bar{x}, M, \bar{c}) = 0.$$

But this implies, by Proposition I.1.D, that :

$$\tilde{\theta}(\bar{x}, m', \bar{c}) = 0.$$

Then, from (i), $\theta_1(\bar{x}, B, \bar{c}) = 0$, for all $B \in L^+$.

Lemma I.1.E

For any $x^* \in \mathbb{R}^n$, $c > 0$ and B_1, B_2 both belonging to L^+ let,

$$Q_1(p, w, c) \triangleq 1/2 p^T B_1 p + \nabla f(x^*) p + cw$$

$$Q_2(p, w, c) \triangleq 1/2 p^T B_2 p + \nabla f(x^*) p + cw$$

$$F \triangleq \{(p, w) \in \mathbb{R}^{n+1} \mid g^j(x^*) + g^j(x^*)p \leq w, j \in \underline{m}; |h^j(x^*) + h_x^j(x^*)p| \leq w, j \in \underline{\ell}\}$$

$$(p_1, w_1) \triangleq \arg \min\{Q_1(p, w, c) \mid (p, w) \in F\}$$

$$(p_2, w_2) \triangleq \arg \min\{Q_2(p, w, c) \mid (p, w) \in F\}.$$

Then,

$$\|p_2 - p_1\| \leq (1/m') \|p_1\| \|B_2 - B_1\|$$

Proof:

See appendix at the end of this chapter.

Now define

$$D_c \triangleq \{x \in R^n \mid \tilde{\theta}(x, 1, c) = 0\}$$

and,

$$D \triangleq \left\{ x \in R^n \mid \exists (\lambda, u) \text{ s.t. } (x, \lambda, u) \text{ is a Kuhn-Tucker point of GMP} \right\}.$$

Proposition I.1.F.

Let $c \geq \bar{c}(x)$

Then,

$$x \in D_c \iff x \in D.$$

Proof:

$\Rightarrow$ This part of the proof is identical to the proof of part (i) of proposition 5 in Mayne and Polak (1978).

$\Leftarrow$ For any $m > 0$ consider the problem,

$$\theta(x, m, c) \triangleq \min_p \{ (m/2) \|p\|^2 + \theta(x, p, c) \mid p \in R^n \}$$

As shown in Proposition A.2 of Mayne and Polak (1978), this problem has the following dual :

$$\tilde{\theta}(x, m, c) = \max_{(y, k)} c [\langle y, g(x) \rangle + \langle k, h(x) \rangle - \psi(x)] - (1/2)m \|\nabla L\|^2$$

subject to:

$$y \geq 0; k = k_1 - k_2; k_1 \geq 0; k_2 \geq 0$$

and

$$\sum_{j \in m} y^j + \sum_{j \in \ell_1} k_1^j + \sum_{j \in \ell_2} k_2^j \leq 1$$

where:

$$\nabla L \triangleq \nabla f(x) + c(g_x^T(x) y + h_x^T(x) k).$$

By hypothesis $x \in D$. Then,

$$\begin{aligned} h(x) &= 0; \quad g(x) \leq 0; \quad \psi(x) = 0; \quad \exists (\lambda, \mu) \text{ such that :} \\ \lambda &> 0; \quad \langle \lambda, g(x) \rangle = 0; \quad \nabla f(x) + g_x^T(x) \lambda + h_x^T(x) \mu = 0. \end{aligned}$$

But, as shown in Proposition 2 of Mayne and Polak (1978), if $x \in D$ we have :

$$\bar{\lambda}(x) = \lambda \quad \text{and} \quad \bar{\mu}(x) = \mu.$$

Let $c \geq \bar{c}(x)$ be arbitrary.

$$\text{By construction, } \bar{c}(x) \geq \sum_{j \in \underline{m}} \lambda^j + \sum_{j \in \underline{\ell}} \mu^j.$$

Then, setting $y = (1/c)\lambda$ and $k = (1/c)\mu$ yields

$$\nabla L = 0; \quad y \geq 0; \quad \sum_{j \in \underline{m}} y^j + \sum_{j \in \underline{\ell}} k^j \leq 1.$$

Substituting these values into the last dual problem we have :

$$\tilde{\theta}(x, m, c) = 0.$$

The result then follows from Proposition I.1.D.

Let $\{A_i^!(x_i, B_i, c_i)\}$ denote the sequence of maps, in the algorithm, which find the next point x_{i+1} with $\tilde{p}_i = 0$. That is :

$$A_i^!(x_i, B_i, c_i) = x_{i+1} = x_i + \beta^{k_i} p_i.$$

Proposition I.1.G.

Let $c > 0$ be arbitrary. Then, any accumulation point x^* of an

infinite bounded sequence $\{x_i\}$, where

$$x_{i+1} = A_i^1(x_i, B_i, c) \text{ for all } i, \text{ satisfies,}$$

$$x^* \in D_c.$$

Proof:

Let I be a sequence such that

$$x_i \xrightarrow{I} x^*.$$

As $\{B_i\}_{i \in I} \subset L^+$, $\exists$ at least one subsequence $I' \subset I$ such that:

$$x_i \xrightarrow{I'} x^* \text{ and } B_i \xrightarrow{I'} B^* \in L^+.$$

Now, there exists a unique solution p^* to the problem,

$$\theta_1(x^*, B^*, c) \triangleq \min_{p \in R} \{1/2 p^T B^* p + \theta(x^*, p, c)\}$$

Suppose $x^* \notin D_c$.

Then, from the Corollary of Proposition I.1.D,

$$\theta_1(x^*, B^*, c) < 0.$$

Hence, as $B^* \in L^+$,

$$(i) \quad \theta(x^*, p^*, c) < 0.$$

Now, as shown in the proof of Proposition I.1.C, there exists a continuous map $S(x)$ such that, for all $B \in L^+$ and all $x \in R^n$,

$$\begin{aligned} \|p'(x, B, c)\| \leq S(x) \triangleq (1/m') \|\nabla f(x)\| + [(1/m')^2 \|\nabla f(x)\|^2 + \\ + 2(c/m')\psi(x)]^{1/2}. \end{aligned}$$

By H1 and the boundedness of $\{x_i\}$ the above inequality implies the existence of a $k < \infty$ such that

$$p_i \in \beta_k(0) \quad \text{for all } i.$$

From Proposition I.1.B it can be readily seen that:

$$(ii) \quad |\gamma(x + \alpha p, c) - \hat{\gamma}(x, \alpha p, c)| \leq \alpha \vartheta(\alpha, x, p)$$

where,

$$\vartheta(\alpha, x, p) \rightarrow 0 \quad \text{uniformly in } (x, p) \in N \times \beta_k(0) \text{ as}$$

$$\alpha \rightarrow 0, \text{ with } N \text{ being any compact neighbourhood of } x^*.$$

From (i), there exists an integer k' such that :

$$(iii) \quad \vartheta(\beta^{k'}, x, p) \leq -\theta(x^*, p^*, c)/8$$

$$\text{for all } (x, p) \in N \times \beta_k(0).$$

Now, since $\hat{\gamma}(x, \alpha p, c)$ is convex in α ,

$$(iv) \quad \hat{\gamma}(x, \alpha p, c) - \gamma(x, c) \leq \alpha \theta(x, p, c), \text{ for all } \alpha \in [0, 1]$$

Hence, from (ii), (iii) and (iv),

$$(v) \quad \gamma(x+\beta^k p, c) - \gamma(x, c) \leq \beta^k (\theta(x, p, c) - \theta(x^*, p^*, c)/8)$$

for all $k \geq k'$, all $(x, p) \in N \times \beta_k(0)$.

Now, the map $\theta(\cdot, \cdot, c)$ is continuous with respect to both arguments.

Then, there exists $\varepsilon_1 > 0$ such that

$$(vi) \quad \theta(x, p, c) \leq \theta(x^*, p^*, c)/2$$

for all $(x, p) \in \beta_{\varepsilon_1}(x^*) \times \beta_{\varepsilon_1}(p^*)$.

On the other hand, by Proposition I.1.C the map $x \mapsto p'(x, B^*, c)$ is continuous.

Then, there exists $\varepsilon_2 > 0$ such that , for all $x \in \beta_{\varepsilon_2}(x^*)$,

$$(vii) \quad p'(x, B^*, c) \in \beta_{\varepsilon_1}(p^*) .$$

If we now define $\varepsilon_3 \triangleq \min\{\varepsilon_1, \varepsilon_2\}$, from (vi) and (vii),

$$(viii) \quad \theta(x, p'(x, B^*, c), c) \leq \theta(x^*, p^*, c)/2$$

for all $x \in \beta_{\varepsilon_3}(x^*)$.

But, by construction,

$$\begin{aligned} \theta(x, p_i, c) - \theta(x, p'(x, B^*, c), c) &= \nabla f(x)^T (p_i - p'(x, B^*, c)) \\ &\quad + c(\hat{\psi}(x, p_i) - \hat{\psi}(x, p'(x, B^*, c))) \end{aligned}$$

or,

$$\begin{aligned}
 (ix) \quad \theta(x, p_i, c) - \theta(x, p'(x, B^*, c), c) &\leq \nabla f(x)^T (p_i - p'(x, B^*, c)) \\
 &+ c \max\{ |g_X^j(x)^T (p_i - p'(x, B^*, c))|, j \in \underline{m}; \\
 &|h_X^j(x)^T (p_i - p'(x, B^*, c))|, j \in \underline{\ell} \}
 \end{aligned}$$

But, $B_i \xrightarrow{I'} B^*$.

Then, from (ix) and Proposition I.1.D,

$$\begin{aligned}
 (x) \quad |\theta(x_i, p_i, c) - \theta(x_i, p'(x_i, B^*, c))| &\leq -\theta(x^*, p^*, c)/4 \\
 \text{for all } i \in I', i \geq i_0 \text{ say.}
 \end{aligned}$$

From (viii), (x) and the convergence $x_i \xrightarrow{I'} x^*$, there exists an $\epsilon_4 > 0$ such that :

$$\begin{aligned}
 (xi) \quad \theta(x_i, p_i, c) &\leq \theta(x^*, p^*, c)/4 \\
 \text{for all } x_i &\in \beta_{\epsilon_4}(x^*), i \in I'.
 \end{aligned}$$

Hence, from (v) and (xi),

$$\begin{aligned}
 \gamma(x_i + \beta^k p_i, c) - \gamma(x_i, c) &\leq \beta^k \theta(x^*, p^*, c)/8 \\
 \text{for all } x_i &\in \beta_{\epsilon_4}(x^*), i \in I' \text{ and all } k \geq k' \\
 \text{But, from the computation of Step 4, } k_i &\leq k'.
 \end{aligned}$$

Therefore,

$$\begin{aligned}
 \gamma(x_{i+1}, c) - \gamma(x_i, c) &\leq \beta^{k'} \theta(x^*, p^*, c)/8 \triangleq -\delta_1 < 0 \\
 \text{for all } x_i &\in \beta_{\epsilon_4}(x^*), i \in I'.
 \end{aligned}$$

This last result contradicts the fact that $x_i \xrightarrow{I'} x^*$.

The contradiction arose from the supposition that :

$$\theta_1(x^*, B^*, c) < 0 .$$

Hence, $\theta_1(x^*, B^*, c) = 0$ and the proposition follows from the Corollary of Proposition I.1.D.

Let $\{\bar{A}_i(x_i, B_i, c_i)\}$ denote the sequence of maps which find the next point x_{i+1} with $\tilde{p}_i \neq 0$. That is :

$$\bar{A}_i(x_i, B_i, c_i) = x_{i+1} = x_i + \beta^{k_i} p_i + \beta^{2k_i} \tilde{p}_i .$$

Proposition I.1.H

Let $c > 0$ be arbitrary. Then, any accumulation point x^* of an infinite bounded sequence $\{x_i\}$, where

$$x_{i+1} = \bar{A}(x_i, B_i, c) \text{ for all } i,$$

satisfies,

$$x^* \in D_c .$$

Proof:

The proof is similar to the last proposition and is given in the appendix at the end of this chapter.

Proposition I.1.I

Let $c > 0$ be arbitrary. Then, any accumulation point x^* of an infinite bounded sequence $\{x_i\}$, where for all i x_{i+1} is either equal

to $A'(x_i, B_i, c_i)$ or equal to $\bar{A}(x_i, B_i, c_i)$, satisfies

$$x^* \in D_c$$

Proof:

Let I' be a subsequence such that :

$$x_i \xrightarrow{I'} x^* \quad \text{and} \quad B_i \xrightarrow{I'} B^* \in L^+.$$

Let $I'' \subset I'$ and $I''' \subset I'$ be the subsequences defined as:

$$I'' \triangleq \{i \in I' \mid x_{i+1} = A'(x_i, B_i, c)\}$$

and

$$I''' \triangleq \{i \in I' \mid x_{i+1} = \bar{A}(x_i, B_i, c)\}.$$

Suppose,

$$\theta_1(x^*, p^*, c) < 0.$$

Then,

(i) as shown in the proof of Proposition I.1.G there exists $\delta_1 > 0$ such that :

$$\gamma(x_{i+1}, c) - \gamma(x_i, c) \leq -\delta_1 < 0$$

for all $i \in I''$, $i \geq i_1$ say.

(ii) as shown in the proof of Proposition I.1.H there exists $\delta_2 > 0$ such that :

$$\gamma(x_{i+1}, c) - \gamma(x_i, c) \leq -\delta_2 < 0$$

for all $i \in I'''$, $i \geq i_2$ say.

But, $I' = I'' \cup I'''$

Hence, from (i) and (ii), there exists $\delta_3 \triangleq \min\{\delta_1; \delta_2\}$ and $i_3 \triangleq \max\{i_1; i_2\}$ such that :

$$\gamma(x_{i+1}, c) - \gamma(x_i, c) \leq -\delta_3 < 0$$

for all $i \in I'$, $i \geq i_3$.

This last fact contradicts $x_i \xrightarrow{I'} x^*$.

Then, $\theta_1(x^*, p^*, c) = 0$ and the proposition follows.

Theorem I.1.J

Any accumulation point x^* of an infinite bounded sequence $\{x_i\}$ constructed by the algorithm satisfies

$$x^* \in D.$$

Proof:

Again let I' be the subsequence where

$$x_i \xrightarrow{I'} x^* \text{ and } B_i \xrightarrow{I'} B^*.$$

Now, from Proposition I.1.F, Proposition I.1.I and the continuity of $\bar{c}(\cdot)$, all the hypothesis of Theorem 1 in Mayne and Polak (1978) are satisfied.

Hence, as x_i is assumed to be bounded, c_{i-1} is increased only finitely often.

Then, there exists an i^* and a $c > 0$ such that

$$c_i = c \quad \text{for all } i \geq i^*.$$

The theorem then follows from Proposition I.1.I and Proposition I.1.F.

To proceed further we need to strengthen our hypothesis. We replace H1 by:

H1': f, g and h are three times continuously differentiable and add a further assumption:

H4: At each Kuhn-Tucker triple $(\hat{x}, \hat{\lambda}, \hat{u})$ of GMP the second-order sufficiency conditions (see AP.1) hold with strict complementary slackness, i.e. $(\hat{\lambda}^j) > 0$ for $j \in I(\hat{x}) \triangleq \{j \in \underline{m} \mid g^j(\hat{x}) = 0\}$.

Lemma I.1.K

Let $\bar{x}$ be any point in $\mathbb{R}^n$ and let N be any compact neighbourhood of $\bar{x}$. Let $k > 0$ and $c > 0$ be arbitrary finite numbers. Then, there exists a constant $d_1 > 0$ such that :

$$\hat{\gamma}(x, \alpha p' + \alpha^2 \tilde{p}, c) - \gamma(x, c) \leq \alpha [\theta(x, p', c) + d_1 \alpha]$$

for all $(x, p', \tilde{p}) \in N \times \beta_k(0) \times \beta_k(0)$, $\alpha \in (0, 1]$

Proof:

See appendix at the end of this chapter.

Proposition I.1.L

Let $\{x_i\}$ be an infinite sequence constructed by the algorithm.

Suppose I is a subsequence such that :

$$c_i = c > 0 \text{ for all } i \in I$$

and

$$x_i \xrightarrow{I} \bar{x} \in \mathbb{R}^n.$$

Then,

$$p_i \xrightarrow{I} 0.$$

Proof:

See appendix at the end of this chapter.

Theorem I.1.M

Let $\{x_i\}$ be an infinite bounded sequence constructed by the algorithm. Then,

$$x_i \rightarrow x^* \in D, \text{ as } i \rightarrow \infty.$$

Proof:

Consider x^* any accumulation point of $\{x_i\}$.

By Theorem I.1.J, $x^* \in D$ and $c_i = \bar{c}$ (constant) for $i \geq i_0$, say.

Let C be a compact ball in R^n containing $\{x_i\}$.

Because of H4 we can use Theorem 4 of Fiacco and McCormick (1968) to see that any Kuhn-Tucker point of GPM is isolated.

Then, as shown in the proof of Theorem 3 of Mayne and Polak (1978), the set $S \triangleq \{x_1^*, \dots, x_s^*\}$ of accumulation points of $\{x_i\}$ in C is finite and each element of this set is a K-T point of GPM.

Now, as $\{x_i\}$ is bounded by hypothesis any infinite sequence of points contains at least one convergent sub-sequence.

That is, given $\varepsilon > 0$ there exists an i_1 such that, for $i \geq i_1$, x_i belongs to at least one of the balls with radius ε and centered at each point $x_k^* \in S$. In other words, for $i \geq i_1$, say, x_i belongs to at least one subsequence converging to some point $x_k^* \in S$.

Then from Proposition I.1.L, we have :

$$(i) \quad \|x_{i+1} - x_i\| \rightarrow 0 \text{ as } i \rightarrow \infty.$$

Let,

$$d \triangleq \min\{\|x_{s_1}^* - x_{s_2}^*\| \mid x_{s_1}^*, x_{s_2}^* \in S\}.$$

As the points are isolated, $d > 0$.

Let $\varepsilon = d/4$.

From (i), for $i \geq i_2$, say, $\|x_{i+1} - x_i\| \leq \varepsilon$.

Hence, if $x_{i_2} \in \beta_\varepsilon(x^*)_\ell$ for some $\ell \in S$, $x_i \in \beta_\varepsilon(x^*)_\ell$ for all

$i \geq i_2$, say.

That is,

$$x_i \rightarrow x_\ell^* \triangleq x^*, \text{ as } i \rightarrow \infty.$$

I.1.4 Remarks

The remedy suggested by Mayne and Polak (1982) when the quadratic sub-problem QP_i^0 has inconsistent constraints or when its solution is not a descent direction of the exact penalty function is to generate an alternative search direction by solving the unconstrained problem below:

I.1.VIII

$$\min_{p \in \mathbb{R}^n} \{ (1/2)p^T \bar{H} p \} + \theta(x, p, c)$$

where :

$\bar{H}$ is any positive definite matrix

and

$\theta(x, p, c)$ is defined in section I.1.2.

As pointed out by the two authors, I.1.VIII is equivalent to the following quadratic problem :

I.1.IX

$$\min_{(p,w) \in \mathbb{R}^n \times \mathbb{R}} \left[f_x(x)p + 1/2 p^T \bar{H} p + c(w - \psi(x)) \right]$$

subject to:

$$\begin{aligned} g^j(x) + g_x^j(x)p &\leq w, & j = 1, \dots, m \\ h^j(x) + h_x^j(x)p &\leq w, & j = 1, \dots, \ell \end{aligned}$$

where:

$\psi(x)$ is defined in section I.1.2.

Now, I.1.IX always has a solution and its solution $\bar{p}_i$ is a descent direction of the exact penalty function $\gamma(x,c)$ (defined in Section I.1.2) for any value of c [see Mayne and Polak (1978)].

Hence, the sub-problem QP_i^w can be thought of as an extension of the robust first order method given by I.1.IX which allows the matrices to vary, a necessary feature to achieve a superlinear rate of convergence.

On the other hand, QP_i^w can be put into the form I.1.VIII with B_i replacing $\bar{H}$. This fact makes the search direction p_i generated by QP_i^w similar to the direction used in the non-differentiable optimization method suggested by Fletcher (1981b). The differences are that here a quadratic sub-problem is solved and a L_∞ exact penalty function is used to generate global convergence, while Fletcher (1981b) minimizes a L_1 exact penalty function subject to the additional constraint below:

I.1.X

$$\|p\| \leq h_i.$$

This step restriction is used to define a 'trust region' with h_i being automatically changed in order to force the convergence [see Fletcher (1981a)]. For different reasons, a constraint of the type I.1.X is also used in the algorithm suggested by Mayne and Polak (1982). Actually, the matrices used by this algorithm in the quadratic sub-problem are not required to be positive definite and a restriction on the norm of its solution is necessary.

Here, assumption H3 ensures that the sequence of matrices $\{B_i\}$ remains positive definite with eigenvalues bounded away from zero, and this provides an upper bound to $\|p_i\|$. [see inequality (i) in the proof of Proposition I.1.C in AP.I.1]

It is interesting to compare Theorem I.1.J with the global convergence result given by Theorem 3.2 of Han (1977), for problems where all the constraints are inequalities.

In the latter, besides H1 and H3, two other assumptions are required to show that any accumulation point x^* is a Kuhn-Tucker point of the mathematical programming problem to be solved. Namely:

I.1.X1

for each i , there exists a Kuhn-Tucker point of QP_i^0 with a Lagrange multiplier vector bounded from above by the penalty parameter, in ∞ - norm.

I.1.XII

$$S^0(x^*) \triangleq \{p \in R^n \mid g(x^*) + \nabla g(x^*)^T p < 0\} \neq \emptyset.$$

Here, in the absence of equalities, the extra assumptions are :

H2: For all $x \in R^n$ the vectors $\{\nabla g^i(x), j \in I(x)\}$ are linearly independent, where :

$$I(x) = \{j \in \underline{m} \mid g^j(x) = \psi(x)\}.$$

I.1.XIII

$\{x_i\}$ remains bounded.

The consistency assumption of QP_i^0 required in I.1.XI, implies that:

I.1.XIV

$$S(x_i) \triangleq \{p \in R^n \mid g(x_i) + \nabla g(x_i)^T p \leq 0\} \neq \emptyset.$$

Now, if the points $\{x_i\}$ are considered to be feasible, I.1.XI and I.1.XIV would be trivially satisfied with $p = 0$. Otherwise, suppose these points to be infeasible. Then, a necessary condition to I.1.XI and I.1.XIV is :

I.1.XV

For all $x \in X \triangleq \{x_i\} \cup \bar{X}$ there exists $p \in R^n$ such that :

$$\nabla g^j(x)p < 0, \quad j \in I_1(x)$$

where:

$\bar{X}$ is the set of all accumulation points of $\{x_i\}$.

and

$$I_1(x) = \{j \in \underline{m} \mid g^j(x) > 0\}.$$

In fact, as the set X cannot be specified a priori, I.1.XV should, for practical purposes, be assumed to be satisfied for all $x \in R^n$.

In this case, it can be shown using theorems of alternative [see McCormick (1967)] that I.1.XV would be equivalent to :

I.1.VI

For all $x \in R^n$ there exists no $a^j > 0$, $j \in I_1(x)$ such that:

$$\sum_{j \in I_1(x)} a^j \nabla^j g(x) = 0.$$

Now, the restriction of the a^j to the positive side of the real line makes the above assumption milder than requiring the linear independence of the gradients.

On the other hand, the cardinality of $I(x)$ is generally expected to be small and, moreover, the definition of $\psi(x)$ clearly implies that, for any infeasible x we have :

$$I(x) \subset I_1(x) \quad .$$

The requirement of H2 is not linked with the existence of solution to the quadratic sub-problem (QP_i^W is always feasible). The purpose of H2 is to make sure that the estimates $\bar{\lambda}(\cdot)$, $\bar{\mu}(\cdot)$ are well defined

in I.1.2 so that a continuous function $\bar{c}(\cdot)$ can be used to update the penalty parameter c_i . If an appropriate value of the penalty parameter is supposed to be a priori known, as in I.1.XI, the assumption H2 could be dropped.

Note yet that I.1.XI also implies that for each i the multipliers of QP_i^0 must be finite, something that may not be always true unless further assumptions are made with respect to the constraints of the quadratic sub-problem [see Theorem 1 of Fiacco and McCormick (1968), for instance].

Furthermore, there is a real possibility of the sequence of multipliers of QP_i^0 becoming unbounded so that I.1.XI is not satisfied, as can be seen in the example given by Chamberlain (1979) where the multipliers of QP_i^0 go to infinity even if the sequence $\{x_i\}$ is bounded.

Here, the boundedness of $\{x_i\}$ implies that the penalty parameter c_i is increased only finitely often and this in turn is sufficient to guarantee that the sequence of multipliers of QP_i^w will remain bounded.

As pointed out by Mayne and Polak (1982) two conditions mild enough to be found in practice are sufficient to assure the boundedness of $\{x_i\}$. Namely:

- The function $f(\cdot)$ in the GMP is bounded from below by a number d .
- Any of the $m + 2\ell$ sets $\{x | g^j(x) \leq e, j \in \underline{m}\}, \{x | h^j(x) \leq e, j \in \underline{\ell}\}$ and $\{x | h^j(x) \geq -e, j \in \underline{\ell}\}$ is compact, where $e = [\gamma(x_1, c_1) - d]/c_1$.

Further assumptions, such as the boundedness of $\{x_i\}$, to avoid the

possibility that $c_i \rightarrow \infty$, is often found in exact penalty algorithms.

Note, however, that :

$$\psi(x) = \| g(x)_+ , h(x) \|_\infty . .$$

Therefore, the exact penalty function $\gamma(x,c)$ belongs to the class of functions defined by I.1.II and the derivative indicated in I.1.III is equal to one. Then, as H2 implies linear independence of the gradients of the active constraints at any feasible point of GMP, we satisfy all the conditions of Theorem 4.4 of Han and Mangasarian (1979), so that :

there exists a $0 \leq \bar{c} < \infty$ such that for all $c \geq \bar{c}$

any strict local minimum of the GMP, is a strict local minimum of $\gamma(x,c)$.

Furthermore, as $(\bar{\lambda}(x_i), \bar{\mu}(x_i)) \rightarrow (\bar{\lambda}(x^*), \bar{\mu}(x^*)) = (\lambda^*, \mu^*)$ and because of the way c_i is updated, at any Kuhn-Tucker point (x^*, λ^*, μ^*) of GMP we have :

$$c_i \geq \|(\lambda^*, \mu^*)\|_1 .$$

Hence, with the additional assumption H4, it is also possible to apply Theorem 4.6 of Han and Mangasarian (1979) to see that the x^* of Theorem I.1.M is indeed a strict local minimum of $\gamma(x, \bar{c})$, with $c_i = \bar{c}$ for i sufficiently large.

Note that the proof of the strong result given by Theorem I.1.M depended on two facts :

- Because of H4 the number of Kuhn-Tucker points of GMP inside any compact set is finite
- $p_i \rightarrow 0$ as $i \rightarrow \infty$.

This observation makes Theorem I.1.M closely related to Theorem 14.1.5 of Ortega and Rheinboldt (1969).

The positive definiteness of $\{B_i\}$ makes possible to apply successfully a number of methods to the solution of the quadratic sub-problem. However, when implementing the routine to solve QP_i^W we must bear in mind that such solution has to be found for all iterations. Therefore, information about the solution of previous sub-problems may be helpful to find the solution of QP_i^W at the current iteration. In particular the convergence of p_i to zero, the linear independence of H2 and the strict complementarity of H4 will enable us to show (see Theorem I.4.0) that, for $i \geq i^*$ say, QP_i^W will predict correctly the active constraints at x^* .

This fact makes quite attractive the use of the 'active set method' [see Fletcher (1981)] in the routine to solve the quadratic sub-problem.

The reason for the test $w_i = 0$ in Step 3 is that the addition of $\tilde{p}_i$ forces the vector $(p_i + \tilde{p}_i)$ to satisfy the constraints up to second order. This is only meaningful if the constraints are already satisfied to the first order, which implies $w_i = 0$.

On the other hand, the test $T(x_i) \leq \varepsilon$ is intended to delay the activation of the procedure to compute $\tilde{p}_i$ until x_i is 'close' to x^* .

Now, although some numerical results suggest that the use of the second order component $\tilde{p}_i$ relatively far away from the solution may

reduce the total number of iterations (see Chapter I-5), the search arc is not necessary to prove the global convergence results of Theorem I.1.J and Theorem I.1.M.

However, $\tilde{p}_i$ is necessary for obtaining a superlinear rate of convergence of the algorithms of Chapter I.3 and Chapter I.4. The use of the search arc in the algorithm of this chapter was intended to show how it does not spoil the global convergence eventually obtained through the use of the line search. [see Proposition I.1.G and Proposition I.1.H]

APPENDIX (Chapter I,1)

AP.1 Second-Order Sufficiency Conditions

A Kuhn-Tucker point $(\bar{x}, \bar{\lambda}, \bar{u})$ of GMP satisfies the second-order sufficiency conditions if for every nonzero vector $y \in \mathbb{R}^n$ satisfying

$$y^T \nabla g^j(\bar{x}) = 0, \quad \text{for all } j \in F_1 \triangleq \{j \in \underline{m} \mid \bar{\lambda}^j > 0\}$$

$$y^T \nabla g^j(\bar{x}) \leq 0, \quad \text{for all } j \in F_2 \triangleq I(\bar{x}) - F_1$$

and

$$y^T \nabla h^j(\bar{x}) = 0, \quad \text{for all } j \in \underline{\ell}$$

it follows that,

$$y^T \nabla^2 L(\bar{x}, \bar{\lambda}, \bar{u}) y \geq 0$$

where,

$$\nabla^2 L(\bar{x}, \bar{\lambda}, \bar{u}) \triangleq \nabla^2 f(\bar{x}) + \sum_{j \in \underline{m}} \bar{\lambda}^j \nabla^2 g^j(\bar{x}) + \sum_{j \in \underline{\ell}} \bar{u}^j \nabla^2 h^j(\bar{x}).$$

Proof of Proposition I.1.C

As $\psi(x)$ does not depend on p or w , the quadratic problem QP_i^w can be written as :

$$\min_{(p, w)} \{f_x(x)p + \frac{1}{2} p^T B p + c(w - \psi(x))\}$$

subject to;

$$g^j(x) + g_x^j(x)p \leq w, \quad j \in \underline{m}$$

$$|h^j(x) + h_x^j(x)p| \leq w, \quad j \in \underline{\ell}.$$

From the definition we have :

$$\theta(x, p, c) \triangleq f_x(x)p + c(\hat{\psi}(x, p) - \psi(x))$$

$$\hat{\psi}(x, p) \triangleq \max\{g^j(x) + g_x^j(x)p, \quad j \in \underline{m}; |h^j(x) + h_x^j(x)p|, \quad j \in \underline{\ell}; 0\}$$

Then the above problem is equivalent to :

$$\theta_1(x, B, c) \triangleq \min_{p \in \mathbb{R}^n} \{ \frac{1}{2} p^T B p + \theta(x, p, c) \}.$$

Setting $p = 0$ we can see that $\theta_1(x, B, c) \leq 0$.

Then,

$$p'(x, B, c) \in T(x, B, c) \triangleq \{p \in \mathbb{R}^n \mid \frac{1}{2} p^T B p + \theta(x, p, c) \leq 0\}.$$

As $\hat{\psi}(x, p) \geq 0$, we have:

$$T(x, B, c) \subset T_1(x, B, c) \triangleq \{p \in \mathbb{R}^n \mid \frac{1}{2} p^T B p + \langle \nabla f(x), p \rangle - c\psi(x) \leq 0\}.$$

Then, from H3

$$T(x, B, c) \subset T_2(x, B, c) \triangleq \{p \in \mathbb{R}^n \mid (\bar{m}'/2) \|p\|^2 + \langle \nabla f(x), p \rangle - c\psi(x) \leq 0\}$$

$$(i) \quad \|p\| \leq (1/m') \|\nabla f(x)\| + [(1/m')^2 \|\nabla f(x)\|^2 + 2(c/m') \psi(x)]^{1/2} \triangleq S(x).$$

Now, the map $x \mapsto S(x)$ is continuous and,

$$T_2(x, B, c) \subset \{p \in \mathbb{R}^n \mid \|p\| \leq S(x)\}.$$

Hence, for any $x \in \mathbb{R}^n$ and $\bar{\varepsilon} > 0$, there exists a compact set F such that,

$$T(x, B, c) \subset T_1(x, B, c) \subset T_2(x, B, c) \subset F$$

for all $x \in \beta_{\bar{\varepsilon}}(\bar{x})$.

Then, for $x \in \beta_{\bar{\varepsilon}}(\bar{x})$ we can always say :

$$\theta_1(x, B, c) = \min\{\frac{1}{2} p^T B p + \theta(x, p, c) \mid p \in F\}.$$

But,

$z(p) \triangleq \frac{1}{2} p^T B p + \theta(x, p, c)$ is a continuous numerical functional.

Then, by the Maximum Theorem [see Berge (1962) pg. 116] we have:

$\theta_1(\cdot, B, c)$ is continuous at $\bar{x}$ and the solution set is upper-semicontinuous at $\bar{x}$.

As $Z(\cdot)$ is strictly convex and $\bar{x}$ is arbitrary, the proposition is proved.

Proof of Proposition I.1,D

It is shown in Proposition A.2 of Mayne and Polak (1978) that $\tilde{\theta}(x,m,c)$ is dual to the problem below:

$$\theta'_m(x,c) = \max_{t,s} \left[c(\langle t, g(x) \rangle + \langle s, h(x) \rangle - \psi(x)) - m/2 \cdot \|\nabla L\|^2 \right]$$

$$\begin{aligned} & t > 0, \quad s = s_1 - s_2, \quad s_1 > 0, \quad s_2 > 0 \\ & \sum_{j \in \underline{m}} t^j + \sum_{j \in \underline{\ell}} s_1^j + \sum_{j \in \underline{\ell}} s_2^j \leq 1 \end{aligned}$$

where

$$\nabla L \triangleq \nabla f(x) + c(g_x^T(x)t + h_x^T(x)s).$$

Now,

$$g^j(x) \leq \psi(x), \quad |h^j(x)| \leq \psi(x), \quad \psi(x) \geq 0 \quad \text{for all } j \text{ and } x \in \mathbb{R}^n.$$

For $m = m_1$, $c = c_1$, and taking into account that the multipliers sum to less than unity, at the solution $(\hat{t}, \hat{s}_1, \hat{s}_2)$ of the above problem the following equations hold :

$$c_1(\langle \hat{t}, g(x) \rangle + \langle \hat{s}_1, h(x) \rangle + \langle \hat{s}_2, -h(x) \rangle - \psi(x)) = 0$$

$$\nabla L = \nabla f(x) + c_1(g_x^T(x)t + h_x^T(x)(\hat{s}_1 - \hat{s}_2)) = 0.$$

Clearly,

$$(i) \quad \tilde{\theta}(x, m, c) \leq 0, \quad \forall x \in \mathbb{R}^n, \forall m, \quad c > 0.$$

Hence, if for any $m > 0$ and $c \geq c_1$ we let :

$$t' = \hat{t}(c_1/c); \quad s'_1 = \hat{s}_1(c_1/c); \quad s'_2 = \hat{s}_2(c_1/c)$$

we would have :

$$c(\langle t', g(x) \rangle + \langle s'_1, h(x) \rangle + \langle s'_2, -h(x) \rangle - \psi(x)) = 0$$

and

$$\nabla L = \nabla f(x) + c(g_x^T(x)t' + h_x^T(x)(s'_1 - s'_2)) = 0.$$

Moreover, as $c \geq c_1$, t' , s'_1 , s'_2 satisfy the constraints of $\theta'_m(x, c)$.

Then, from (i) we have $\tilde{\theta}(x, m, c) = 0$ and the proposition is proved.

Proof of Lemma I.1.E

Let,

$$z_1 \triangleq (p_1, w_1) \quad ; \quad z_2 \triangleq (p_2, w_2).$$

As the above vectors minimize differentiable functionals over the same convex set, we know that :

$$(i) \quad (z - z_1)^T \nabla Q_1(z_1, c) \geq 0, \quad \text{for all } z \in F$$

$$(ii) \quad (z - z_2)^T \nabla Q_2(z_2, c) \geq 0, \quad \text{for all } z \in F.$$

Substituting $z = z_2$ in (i) and $z = z_1$ in (ii), after a rearrangement we have:

$$(z_2 - z_1)^T [\nabla Q_2(z_2, c) - \nabla Q_2(z_1, c)] \leq (z_2 - z_1)^T [\nabla Q_1(z_1, c) - \nabla Q_2(z_1, c)] .$$

But,

$$\nabla Q_2(z_2, c) = \begin{pmatrix} B_2 p_2 + \nabla f(x^*) \\ c \end{pmatrix}; \quad \nabla Q_1(z_1, c) = \begin{pmatrix} B_1 p_1 + \nabla f(x^*) \\ c \end{pmatrix};$$

$$\nabla Q_2(z_1, c) = \begin{pmatrix} B_2 p_1 + \nabla f(x^*) \\ c \end{pmatrix} .$$

Then,

$$(p_2 - p_1, w_2 - w_1)^T \begin{pmatrix} B_2(p_2 - p_1) \\ 0 \end{pmatrix} \leq (p_2 - p_1, w_2 - w_1)^T \begin{pmatrix} (B_1 - B_2)p_1 \\ 0 \end{pmatrix}$$

or,

$$(p_2 - p_1)^T B_2 (p_2 - p_1) \leq (p_2 - p_1)^T (B_1 - B_2) p_1 .$$

By hypothesis, $B_2 \in L^+$.

Then,

$$m' \|p_2 - p_1\|^2 \leq \|p_2 - p_1\| \|B_2 - B_1\| \|p_1\|$$

or,

$$\|p_2 - p_1\| \leq (1/m') \|p_1\| \|B_2 - B_1\| .$$

Proof of Proposition I.1.H

Let I' be the subsequence defined in Proposition I.1.G.

That is :

$$x_1 \xrightarrow{I'} x^* \quad ; \quad B_i \xrightarrow{I'} B^*.$$

By construction,

$$\begin{aligned} \hat{\gamma}(x, \alpha p_i, c) - \hat{\gamma}(x, \alpha p_i + \alpha^2 \tilde{p}_i, c) &= -\alpha^2 \langle \nabla f(x), \tilde{p}_i \rangle + c \max\{g^j(x) + \alpha g_x^j(x)^T p_i, j \in \underline{m}; \\ &\quad |h^j(x) + \alpha h_x^j(x)^T p_i|, j \in \underline{\ell}; 0\} \\ &\quad - c \max\{g^j(x) + \alpha g_x^j(x) p_i + \alpha^2 g_x^j(x) \tilde{p}_i, j \in \underline{m}; \\ &\quad |h^j(x) + \alpha h_x^j(x) p_i + \alpha^2 h_x^j(x) \tilde{p}_i|, j \in \underline{\ell}; 0\} \end{aligned}$$

Then,

$$\begin{aligned} |\hat{\gamma}(x, \alpha p_i, c) - \hat{\gamma}(x, \alpha p_i + \alpha^2 \tilde{p}_i, c)| &\leq \alpha [\alpha \|\nabla f(x)\| + c \max\{\alpha \|g_x^j(x)\|, j \in \underline{m}; \\ &\quad \alpha \|h_x^j(x)\|, j \in \underline{\ell}; 0\} \|\tilde{p}_i\| \end{aligned}$$

or

$$|\hat{\gamma}(x, \alpha p_i, c) - \hat{\gamma}(x, \alpha p_i + \alpha^2 \tilde{p}_i, c)| \leq \alpha t(\alpha, x, \tilde{p}_i).$$

The map $t(\cdot, x, \tilde{p}_i)$ is continuous for any $(x, \tilde{p}_i) \in \mathbb{R}^n \times \mathbb{R}^n$.

Then,

(i) $t(\cdot, x, \tilde{p}_i) \rightarrow 0$ uniformly in $(x, \tilde{p}_i) \in N \times \beta_k(0)$ as $\alpha \rightarrow 0$ where N is any compact neighbourhood of x^* and $\beta_k(0)$ any compact ball.

Now, as shown in the proof of Proposition I.1.G, we can always define a compact ball $\beta_k(0)$ such that : $p_i \in \beta_k(0)$ for all i .

By construction,

$$p_i \in \beta_k(0) \Rightarrow \tilde{p}_i \in \beta_k(0).$$

Suppose $x^* \notin D_c$.

Then, as shown in the proof of Proposition I.1.G., there exists a $\delta_1 > 0$ such that :

$$(ii) \quad \theta(x^*, p^*, c) \leq -\delta_1 < 0$$

where

$$p^* = \arg \min_{p \in \mathbb{R}^n} \{ \frac{1}{2} p^T B^* p + \theta(x^*, p, c) \} .$$

From (i) and (ii), there exists an $\bar{\alpha} \in (0, 1]$ such that :

$$t(\alpha, x, \tilde{p}_i) \leq -\theta(x^*, p^*, c)/32$$

$$\forall x \in \mathbb{R}^n, i \in I' (i \geq i_0, \text{ say}) \text{ and } \alpha \in (0, \bar{\alpha}].$$

Consequently,

$$(iii) \quad |\hat{\gamma}(x, \alpha p_i, c) - \hat{\gamma}(x, \alpha p_i + \alpha^2 \tilde{p}_i, c)| \leq -\alpha \theta(x^*, p^*, c)/32$$

for all $x \in N$, $i \in I' (i \geq i_0, \text{ say})$ and $\alpha \in (0, \bar{\alpha}]$.

Now, from Proposition I.1.B, there exists $\bar{\bar{\alpha}} \in (0, 1]$ such that:

$$(iv) \quad |\gamma(x + \alpha p_i + \alpha^2 \tilde{p}_i, c) - \hat{\gamma}(x, \alpha p_i + \alpha^2 \tilde{p}_i, c)| \leq -\alpha \theta(x^*, p^*, c)/32$$

for all $x \in N$, $i \in I' (i \geq i_0, \text{ say})$ and $\alpha \in (0, \bar{\bar{\alpha}}]$.

Let $\alpha' \triangleq \min\{\bar{\alpha}, \bar{\bar{\alpha}}\}$. From (iii) and (iv) :

$$(v) \quad |\gamma(x + \alpha p_i + \alpha^2 \tilde{p}_i, c) - \hat{\gamma}(x, \alpha p_i, c)| \leq -\alpha \theta(x^*, p^*, c)/16$$

for all $x \in N$, $i \in I' (i \geq i_0, \text{ say})$ and $\alpha \in (0, \alpha']$

Again by Proposition I.1.B, there exists $\bar{\alpha} \in (0,1]$ such that :

$$(vi) \quad |\hat{\gamma}(x, \alpha p_i, c) - \gamma(x + \alpha p_i, c)| \leq -\alpha \theta(x^*, p^*, c)/16$$

for all $x \in N$, $i \in I'$ ($i \geq i_0$, say) and $\alpha \in (0, \bar{\alpha}]$.

Let $\alpha^0 \triangleq \min\{\alpha'; \bar{\alpha}\}$.

From (v) and (vi),

$$(vii) \quad |\gamma(x + \alpha p_i + \alpha^2 \tilde{p}_i, c) - \gamma(x + \alpha p_i, c)| \leq -\alpha \theta(x^*, p^*, c)/8$$

for all $x \in N$, $i \in I'$ ($i \geq i_0$, say), and $\alpha \in (0, \alpha^0]$.

Now, as $x_i \xrightarrow{I'} x^*$, there exists an $i_1 \in I'$ such that, for all $i \in I'$, $i \geq i_1$, we have : $x_i \in N$.

Furthermore, because $\theta(x^*, p^*, c) < 0$, similar reasoning to the proof of Proposition I.1.G will show that :

$$(viii) \quad \text{there exists an } i_2 \text{ and } \alpha^1 \in (0,1) \text{ such that}$$

$$\gamma(x_i + \alpha p_i, c) - \gamma(x_i, c) \leq -\alpha \theta(x^*, p^*, c)/4$$

for all $i \in I'$ ($i \geq i_2$) and for all $\alpha \in (0, \alpha^1]$.

From (vii) and (viii), there exists an integer $\tilde{k} \in \{0, 1, \dots\}$, with $\beta^{\tilde{k}} \leq \min\{\alpha^0, \alpha^1\}$, and $i^* \triangleq \max\{i_1, i_2\}$ such that :

$$\gamma(x_i + \beta^k p_i + \beta^{2k} \tilde{p}_i, c) - \gamma(x_i, c) \leq -\beta^k \theta(x^*, p^*, c)/8$$

for all $i \in I'$, $i \geq i^*$, and all $k \geq \tilde{k}$.

But, from the computation of the step length, $k_i \leq \tilde{k}$. Hence,

$$\gamma(x_{i+1}, c) - \gamma(x_i, c) \leq \beta^{\tilde{k}} \theta(x^*, p^*, c)/8 \stackrel{\Delta}{=} -\delta_2 < 0$$

for all $i \in I', i \geq i^*$,

This contradicts the fact that $x_i \xrightarrow{I'} x^*$.

The contradiction arose from the assumption that $x^* \notin D_c$ and, therefore, the proposition is proved.

Proof of Lemma I.1.K.

By construction,

$$\begin{aligned} \hat{\gamma}(x, \alpha p' + \alpha^2 \tilde{p}, c) - \gamma(x, c) &= \alpha \langle \nabla f(x), p' \rangle + \alpha^2 \langle \nabla f(x), \tilde{p} \rangle + \\ &+ c \cdot \max \left(\begin{array}{l} g^j(x) + \alpha g_x^j(x) T_{p'} + \alpha^2 g_x^j(x) T_{\tilde{p}}, \quad j \in \underline{m} \\ |h^j(x) + \alpha h_x^j(x) T_{p'} + \alpha^2 h_x^j(x) T_{\tilde{p}}|, \quad j \in \underline{\ell} \end{array} \right) \\ &- c\psi(x) \end{aligned}$$

or,

$$(i) \quad \begin{aligned} \hat{\gamma}(x, \alpha p' + \alpha^2 \tilde{p}, c) - \gamma(x, c) &\leq \alpha \langle \nabla f(x), p' \rangle + c(\hat{\psi}(x, \alpha p') - \psi(x)) + \\ &+ \alpha^2 \left(\langle \nabla f(x), \tilde{p} \rangle + c \cdot \max \left(\begin{array}{l} |g_x^j(x) \tilde{p}|, j \in \underline{m} \\ |h_x^j(x) \tilde{p}|, j \in \underline{\ell} \end{array} \right) \right) \end{aligned}$$

But, for any compact neighbourhood N of $\bar{x}$ and any compact ball $\beta_k(0)$, there exists $d_1 > 0$, finite, such that :

$$(ii) \quad |\langle \nabla f(x), \tilde{p} \rangle + c \max\{|g_x^j(x) \tilde{p}|, j \in \underline{m}; |h_x^j(x) \tilde{p}|, j \in \underline{\ell}\}| \leq d_1$$

for all $(x, \tilde{p}) \in N \times \beta_k(0)$.

The lemma follows from (i), (ii) and the definition of

$\theta(\cdot, \cdot, \cdot)$.

Proof of Proposition I.1.L

As shown in the proof Proposition I.1.A

$$\theta(x_i, p_i, c) \leq -p_i^T B_i p_i.$$

Then, by H3,

$$(i) \quad \theta(x_i, p_i, c) \leq -m' \|p_i\|^2 \quad \text{for all } i.$$

Now, as shown in the proof of Proposition I.1.C, there exists a continuous function $S(\cdot)$ such that :

$$\|p_i\| \leq S(x_i), \quad \text{for all } i \in I.$$

But, by hypothesis, $x_i \xrightarrow{I} x^*$.

Hence, there exists a $k < \infty$ and an i_0 such that :

$$(ii) \quad \|\tilde{p}_i\| \leq \|p_i\| \leq k \quad \text{for all } i \in I, i \geq i_0.$$

From Lemma I.1.K, there exists a constant $d_1 > 0$ such that :

$$(iii) \quad \hat{\gamma}(x_i, \alpha p_i + \alpha^2 \tilde{p}_i, c) - \gamma(x_i, c) \leq \alpha[\theta(x_i, p_i, c) + d_1 \alpha] \\ \text{for all } (x_i, p_i, \tilde{p}_i) \in N \times \beta_k(0) \times \beta_k(0).$$

From Proposition I.1,B,

$$(iv) \quad |\gamma(x_i + \alpha p_i + \alpha^2 \tilde{p}_i, c) - \hat{\gamma}(x_i, \alpha p_i + \alpha^2 \tilde{p}_i, c)| \leq \alpha \vartheta(\alpha, x_i, p_i, \tilde{p}_i)$$

where,

$$\begin{aligned} \vartheta(\alpha, x_i, p_i, \tilde{p}_i) &\rightarrow 0 \text{ uniformly as } \alpha \rightarrow 0 \\ \text{for all } (x_i, p_i, \tilde{p}_i) &\in N \times \beta_k(0) \times \beta_k(0). \end{aligned}$$

Suppose that for some infinite subsequence $I' \subset I$ there exists a $\delta_1 > 0$ such that :

$$(v) \quad \theta(x_i, p_i, c) \leq -\delta_1, \quad \text{for all } i \in I'.$$

By (iv) there exists an integer k' such that :

$$\begin{aligned} (vi) \quad \vartheta(\beta^{k'}, x_i, p_i, \tilde{p}_i) + d_1 \beta^{k'} &\leq \delta_1/2 \\ \text{for all } (x_i, p_i, \tilde{p}_i) &\in N \times \beta_k(0) \times \beta_k(0). \end{aligned}$$

Hence, from (ii), (iii), (iv), (v) and (vi) we have :

$$\begin{aligned} \gamma(x_i + \beta^{k'} p_i + \beta^{2k'} \tilde{p}_i, c) - \gamma(x_i, c) &\leq -\beta^{k'} \delta_1/2 \\ \text{for all } i \in I', i &\geq i_0. \end{aligned}$$

But, from the computation of the step length, $k_i \geq k'$ for all $i \in I', i \geq i_0$.

Hence,

$$\begin{aligned} \gamma(x_{i+1}, c) - \gamma(x_i, c) &\leq -\beta^{k'} \delta_1/2 < 0 \\ \text{for all } i \in I'', i &\geq i_0, \end{aligned}$$

This last result contradicts the hypothesis that :

$$x_i \not\rightarrow \bar{x}.$$

The contradiction arose from (v).

Hence, there not exist any infinite subsequence $I' \subset I$ where $\theta(x_i, p_i, c)$ remains bounded away from zero.

That is,

$$(vii) \quad \theta(x_i, p_i, c) \not\rightarrow 0.$$

The proposition follows from (vii) and (i).

I.2 CONVEX INEQUALITIES CASE

I.2.1 Introduction

Consider the problem CIP below, which is a special case of the GMP, having no equalities and convex inequalities.

CIP

$$\min_{x \in \mathbb{R}^n} \{f(x) \mid g(x) \leq 0\}$$

where:

$$f : \mathbb{R}^n \rightarrow \mathbb{R}, \quad g : \mathbb{R}^n \rightarrow \mathbb{R}^m$$

and

$$g^j(\cdot) \text{ is convex for all } j \in \underline{m}.$$

In this case the sub-problem QP_i^0 would be :

$$QP_i^0$$

$$\min_{p \in \mathbb{R}^n} f_x(x_i)p + p^T B_i p$$

s.t.

$$g^j(x_i) + g_x^j(x_i)p \leq 0, \quad \forall j \in \underline{m}.$$

Because of the convexity assumption, it is readily seen that :
if there exists a point x' satisfying the constraints of CIP then the constraints of QP_i^0 are consistent. In fact, if x' is any point satisfying

$$g^j(x') \leq 0 \quad \text{for all } j \in \underline{m}$$

then $p' \triangleq (x' - x_i)$ satisfies:

$$g^j(x_i) + g_x^j(x_i)p' \leq 0, \quad \forall j \in \underline{m}.$$

In this case the use of QP_i^w instead of QP_i^0 is only useful to ensure that the solution of the quadratic problem is a descent direction for the exact penalty function (with the current value of c_i). On the other hand, it is possible to show [see Lemma I.2A] that : if there exists a point p satisfying the constraints of QP_i^0 then the solution w_i of QP_i^w satisfies $w_i = 0$, whenever c_i satisfies :

I.2.I

$$c_i \geq \|\lambda_i^0\|_1$$

where

(p_i^0, λ_i^0) is the Kuhn-Tucker couple of QP_i^0 .

Actually, the absence of equalities enable us to strengthen this result a little and say that w_i will be equal to zero if and only if I.2.I is satisfied [see Proposition I.2.B].

Although p_i^w is a descent direction for the penalty function regardless of the value of the penalty parameter c_i , so that I.2.I does not need to be satisfied at all iterations, in the convex inequalities case the occurrence of $w_i > 0$ at the solution of QP_i^w can be used in a procedure to update the penalty parameter ensuring that, for i sufficiently large,

$$c_i = \bar{c} > \|\bar{\lambda}\|_1$$

where

$(\bar{x}, \bar{\lambda})$ is a Kuhn-Tucker couple of CIP.

This idea is used to construct the next algorithm which, under some assumptions, will generate a sequence of points $\{x_i\}$ such that : any accumulation point $\{x_i\}$ satisfies the Kuhn-Tucker conditions to CIP. As in the algorithm of I.1, if the second order sufficient conditions are satisfied at all Kuhn-Tucker points of CIP we can prove a stronger result :

As $i \rightarrow \infty$, $x_i \rightarrow x^*$ (strict local minimizer of CIP).

I.2.2 The Algorithm

Consider the convex constrained problem (CIP) defined in the last section. Let the maps ψ , γ , $\hat{\psi}$, $\hat{\gamma}$ and θ be as previously defined. Let $(p_i, w_i, \lambda_i, v_i)$ be the solution and multipliers of :

$$\begin{aligned} & \overline{QP}_i^w \\ & \min_{(p,w) \in R^n \times R} f_x(x_i)p + \frac{1}{2}p^T B_i p + c_i w \end{aligned}$$

subject to:

$$g^j(x_i) + g_x^j(x_i)p \leq w, \quad j \in \underline{m} \quad w \geq 0.$$

λ_i^j are the multipliers of the first m inequalities and v_i of the last.

Algorithm

Data: $x_1 \in R^n$, B_1 ($n \times n$ positive definite); b , c_1 , $\delta \in (0, \infty)$; $\beta \in (0, 1)$

Step 0 : Set $i = 1$

Step 1 : Find (p_i, w_i) to solve $\overline{QP}_i^w$

Step 2 : If $w_i > 0$, set $c_{i+1} = c_i + bw_i + \delta$

Else,

set $c_{i+1} = c_i$

Step 3:

a) Compute k_i , the first integer in $\{0,1,\dots\}$ satisfying :

$$\gamma(x_i + \beta^{k_i} p_i, c_i) - \gamma(x_i, c_i) \leq \beta^{k_i} \theta(x_i, p_i, c_i)/8$$

b) Set $x_{i+1} = x_i + \beta^{k_i} p_i$

Step 4: Update B_i to B_{i+1}

Step 5 : Set $i = i+1$; Go to Step 1.

Because we are only concerned in this chapter with the eventual effects that the use of the simple procedure to update c_i might cause to the global convergence of the algorithm, we use, for the sake of simplicity, a line search at step 3. Note, however, that, as shown in the previous chapter, the addition of $\tilde{p}_i$ would not destroy the convergence.

I.2.3 Convergence Results

To study the convergence of the algorithm we initially assume that the following conditions are satisfied:

C1 : f, g are continuously differentiable and $g^j(\cdot)$ is convex for all j

C2 : f is bounded from below and

$$V^0 \triangleq \{x \mid g(x) < 0\} \neq \emptyset$$

C3 : For all i , $B_i \in L^+ \triangleq \{B \in L(R^n) \mid m' \|y\|^2 \leq y^T B y \leq M \|y\|^2 \text{ for all } y \in R^n\}$

where

$$m' > 0 \text{ and } M < \infty$$

C4 : There exists at least one $j \in \underline{m}$ such that :

$$V^j \triangleq \{x \mid g^j(x) \leq 0\} \text{ is compact}$$

C5 : For all i , the gradients of the active constraints of $\overline{QP}_i^w$ are linearly independent.

Lemma I.2.A

Let $(\hat{p}_i, \hat{\lambda}_i, \hat{\mu}_i)$ be the K-T point of QP_i^0 (see section I.1.1)

Let $(\bar{p}_i, \bar{w}_i, \bar{\lambda}_i, \bar{\mu}_i)$ be the K-T point of QP_i^w (see section I.1.2)

If QP_i^0 is feasible and

$$c_i \geq \sum_{j \in \underline{m}} \hat{\lambda}_i^j + \sum_{j \in \underline{\ell}} |\bar{\mu}_i^j| \quad .$$

Then,

$$\bar{p}_i = \hat{p}_i \text{ and } \bar{w}_i = 0.$$

Proof:

For any $w > 0$ let p_i^w denote the solution of :

$$Q2_i^w$$

$$\min_{p \in R^n} f_x(x_i)p + \frac{1}{2} p^T B_i p$$

subject to:

$$g^j(x_i) + g_x^j(x_i)p \leq w, \quad j \in \underline{m}$$

$$h^j(x_i) + h_x^j(x_i)p \leq w, \quad j \in \underline{\ell}$$

For $w = \bar{w}_i, p_i^{\bar{w}_i}$ denotes the solution of $Q2_i^{\bar{w}_i}$.

Hence, as $\bar{p}_i$ is a feasible point for $Q2_i^{\bar{w}_i}$,

$$(i) \quad f_x(x_i) p_i^{\bar{w}_i} + \frac{1}{2} (p_i^{\bar{w}_i})^T B_i p_i^{\bar{w}_i} \leq f_x(x_i) \bar{p}_i + \frac{1}{2} \bar{p}_i^T B_i \bar{p}_i.$$

Now define :

$$y_i \triangleq [\bar{w}_i, \bar{w}_i, \dots, \bar{w}_i] \in R^{m+\ell}$$

and

$$r_i \triangleq [\bar{\lambda}_i^1, \dots, \bar{\lambda}_i^m, \bar{u}_i^m, \dots, \bar{u}_i^\ell] \in R^{m+\ell}.$$

As the objective function and the constraints of QP_i^0 and $Q2_i^w$ are convex for all i and w , we can apply the sensitivity

result of AP.1 which is proved by Luenberger (1969). Then,

$$(ii) \quad [f_x(x_i) \hat{p}_i + \frac{1}{2} \hat{p}_i^T B_i \hat{p}_i] - [f_x(x_i) p_i^{\bar{w}_i} + \frac{1}{2} (p_i^{\bar{w}_i})^T B_i p_i^{\bar{w}_i}] \leq \langle y_i, r_i \rangle.$$

From (i) and (ii)

$$(iii) \quad [f_x(x_i) \hat{p}_i + \frac{1}{2} \hat{p}_i^T B_i \hat{p}_i] - [f_x(x_i) \bar{p}_i + \frac{1}{2} \bar{p}_i^T B_i \bar{p}_i] \leq \langle y_i, r_i \rangle.$$

Let us compare the values of the objective function of QP_i^w for the two pairs : $(\bar{p}_i, \bar{w}_i)$; $(\hat{p}_i, 0)$.

Call,

$$V_1 \triangleq f_x(x_i) \bar{p}_i + \frac{1}{2} \bar{p}_i^T B_i \bar{p}_i + c_i \bar{w}_i$$

and

$$V_2 \triangleq f_x(x_i) \hat{p}_i + \frac{1}{2} \hat{p}_i^T B_i \hat{p}_i.$$

That is,

$$V_2 - V_1 = [f_x(x_i) \hat{p}_i + \frac{1}{2} \hat{p}_i^T B_i \hat{p}_i] - [f_x(x_i) \bar{p}_i + \frac{1}{2} \bar{p}_i^T B_i \bar{p}_i] - c_i \bar{w}_i.$$

From (iii),

$$V_2 - V_1 \leq \langle y_i, r_i \rangle - c_i \bar{w}_i.$$

$$\text{But, } \langle y_i, r_i \rangle \leq \left(\sum_{j \in \underline{m}} |\bar{\lambda}_i^j| + \sum_{j \in \underline{\ell}} |\bar{\mu}_i^j| \right) \bar{w}_i$$

Hence,

$$V_2 - V_1 \leq \left(\sum_{j \in \underline{m}} |\bar{\lambda}_i^j| + \sum_{j \in \underline{\ell}} |\bar{\mu}_i^j| - c_i \right) \bar{w}_i.$$

$$\text{But } \bar{w}_i \geq 0 \text{ and, by hypothesis, } c_i \geq \sum_{j \in \underline{m}} |\bar{\lambda}_i^j| + \sum_{j \in \underline{\ell}} |\bar{\mu}_i^j|.$$

Then,

$$(iv) \quad V_2 - V_1 \leq 0.$$

But, $(\bar{p}_i, \bar{w}_i)$ solves QP_i^w and $(\bar{p}_i, 0)$ is a feasible point to this problem. Then

$$(v) \quad V_1 - V_2 \leq 0.$$

The lemma follows from (iv), (v) and the uniqueness of the solution to QP_i^w .

Proposition I.2.B

Let $(\hat{p}_i, \hat{\lambda}_i)$ be the K-T point of $\overline{QP}_i^0$ defined below :

$$\overline{QP}_i^0$$

$$\min_{p \in \mathbb{R}^n} f_x(x_i)p + \frac{1}{2}p^T B_i p$$

subject to :

$$g^j(x_i) + g_x^j(x_i)p \leq 0, \quad j \in \underline{m}.$$

Then, the solution of $\overline{QP}_i^w$ will be

$$(p_i, w_i) = (\hat{p}_i, 0)$$

if and only if,

$$c_i \geq \sum_{j \in \underline{m}} \hat{\lambda}_i^j.$$

Proof:

The first part comes directly from Lemma I.1.A.

Now, suppose $(p_i, w_i) = (\hat{p}_i, 0)$.

At the solution of $\overline{QP}_i^w$ the following K-T conditions

are satisfied :

$$\begin{matrix} n+1 \\ \text{equations} \end{matrix} \left\{ \begin{pmatrix} \nabla f(x_i) + B_i p_i \\ c_i \end{pmatrix} + \sum_{j=1}^m \lambda_i^j \begin{pmatrix} \nabla g^j(x_i) \\ -1 \end{pmatrix} + v_i \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 0 \end{pmatrix} \right.$$

$$\lambda_i^j \geq 0, \quad j \in \underline{m}; \quad v_i \geq 0.$$

Then,

$$(i) \quad \nabla f(x_i) + B_i \hat{p}_i + \sum_{j \in \underline{m}} \lambda_i^j \nabla g^j(x_i) = 0$$

and

$$(ii) \quad c_i = \sum_{j \in \underline{m}} \lambda_i^j + v_i.$$

But, at the solution of $\overline{QP}_i^0$,

$$(iii) \quad f(x_i) + B_i \hat{p}_i + \sum_{j \in \underline{m}} \hat{\lambda}_i^j \nabla g^j(x_i) = 0$$

Now, by the complementarity condition which is satisfied at the solution of $\overline{QP}_i^w$ and $\overline{QP}_i^0$ and because of assumption C5, we have:

$$\hat{\lambda}_i^j = \lambda_i^j, \quad \text{for all } j \in \underline{m}.$$

The proposition then follows from (iii) and the fact that $v_i \geq 0$.

Proposition 1.2.C

There exists $\bar{c}$ and i_0 , both finite, such that :

$$c_i = \bar{c} ; \quad w_i = 0, \quad \text{for all } i \geq i_0.$$

Proof:

From C4, there exists $j' \in \underline{m}$ such that:

$$V^{j'} \triangleq \{x \in \mathbb{R}^n \mid g^{j'}(x) \leq 0\} \text{ is compact.}$$

But, from C1, $g^{j'}(\cdot)$ is convex and continuously differentiable.

Then, by Lemma 4.1.14 pg 139 of Stoer and Witzgall (1970),

$$(i) \quad V_s^{j'} \triangleq \{x \in \mathbb{R}^n \mid g^{j'}(x) \leq s\} \text{ is compact for all finite } s.$$

By C2, there exists a $d > -\infty$ such that :

$$f(x) \geq d, \quad \forall x \in \mathbb{R}^n.$$

Since $c_i \geq c_1$, it can be shown that

$$\psi(x_i) \leq s_1 \triangleq [\gamma(x_1, c_1) - d]/c_1 < \infty, \quad \text{for all } i.$$

Hence, x_i is bounded and

$$x_i \in V_{s_1}^{j'} \triangleq \{x \in \mathbb{R}^n \mid g^{j'}(x) \leq s_1\}, \quad \text{for all } i.$$

But, from (i), $V_{s_1}^{j'}$ is compact.

Then we can apply Lemma 3.3 in Han (1977) [see AP.2] to show that :

(ii) there exists $r < \infty$ such that :

$$\sum_{j \in \underline{m}} \hat{\lambda}^j \leq r, \quad \text{for all } i.$$

Now, suppose there exists an infinite subsequence I_1 such that

(iii) $w_i > 0$ for all $i \in I_1$.

As at each $i \in I_1$ w_i is increased by an amount greater than $\delta > 0$ and I_1 is supposed to be an infinite subsequence, there exists a finite i_0 such that :

$$c_i \geq r \quad \text{for all } i \geq i_0.$$

Hence, from (ii) and Proposition I.2.B,

$$w_i = 0 \quad \text{for all } i \geq i_0.$$

This contradicts (iii) and therefore no such subsequence I_1 exists, that is, $w_i > 0$ only finitely often.

The proposition then follows from the fact that c_i is only increased if $w_i > 0$.

Theorem I.2.D

Any accumulation point x^* of an infinite sequence $\{x_i\}$ constructed by the algorithm satisfies,

$$x^* \in D.$$

Proof:

Let I be such that:

$$x_i \xrightarrow{I} x^*.$$

Because of C3, there exists $B^* \in L^+$ and $I' \subset I$ such that :

$$(i) \quad B_i \xrightarrow{I'} B^*.$$

From the last proposition, for all i sufficiently large

$$(ii) \quad c_i = \bar{c} > 0 \quad \text{and} \quad w_i = 0.$$

Hence, from Lemma I.1.E and (i),

$$(iii) \quad p_i \xrightarrow{I'} p^*$$

where p^* is the unique solution of :

$$\min_{p \in \mathbb{R}^n} f_x(x^*)p + \frac{1}{2}p^T B^*p$$

subject to:

$$g_x^j(x^*) + g_x^j(x^*)p \leq 0, \quad j \in \bar{m}.$$

Clearly,

$$(iv) \quad p^* = 0 \Rightarrow x^* \in D.$$

Now, from (ii) and Proposition I.2.B,

$$c_i = \bar{c} \geq \sum_{j \in \underline{m}} \hat{\lambda}_i^j, \quad \text{for all } i \text{ sufficiently large.}$$

Then, as shown in Proposition 1 of Mayne and Polak (1978),
for $i \geq i_0$ say, we have:

$$(v) \quad \theta(x_i, p_i, \bar{c}) \leq -p_i^T B_i p_i \leq -m' \|p_i\|^2.$$

Suppose $p^* \neq 0$.

Then, from (iii) and (v), there exists a $\delta_2 > 0$ such that :

$$(vi) \quad \theta(\dot{x}^*, p^*, \bar{c}) \leq -\delta_2 < 0.$$

But then, following the reasoning of the proof of Proposition I.1.G, there exists a finite i_1 such that :

$$\gamma(x_{i+1}, \bar{c}) - \gamma(x_i, \bar{c}) \leq -\delta_3 < 0, \quad \text{for all } i \in I, i \geq i_1.$$

This last inequality contradicts the fact that:

$$x_i \xrightarrow{I'} x^*.$$

Hence, $p^* = 0$ and, from (iv), the theorem is shown.

As in the previous chapter if C1 is replaced by C'_1

C'_1 : f, g are three times continuously differentiable

with the further assumption below:

C6: at each Kuhn-Tucker couple (x^*, λ^*) of CIP the second-order sufficiency conditions hold with strict slackness
then, a stronger result can be shown. Namely :

Theorem I.2.E

Let x_i be an infinite sequence constructed by the algorithm.

Then,

$$x_i \rightarrow x^* \in D, \text{ as } i \rightarrow \infty.$$

Proof:

As shown in the last theorem, for any convergent subsequence

I we have :

$$x_i \xrightarrow{I} x^* \in D.$$

By Proposition I.2.C,

$$c_i = \bar{c} > 0 \text{ for all } i \text{ sufficiently large.}$$

Then, a detailed examination shows that Proposition I.1.L remains valid.

Therefore,

$$p_i \xrightarrow{I} 0, \text{ for all convergent subsequence } I.$$

The rest of the proof is identical to the proof of Theorem I.1.M.

I.2.4. Remarks

The variable w in the quadratic sub-problem QP_i^w has in the algorithm of this chapter a different role from that of the previous algorithm.

While in the latter the addition of w to QP_i^0 served to overcome the possibility of this quadratic sub-problem having inconsistent constraints, in the former that possibility is avoided through the assumptions made. That is:

$$\left. \begin{array}{l} g^j \text{ convex for all } j \\ V = \{x \mid g^+(x) < 0\} \neq \emptyset \end{array} \right\} \Rightarrow QP_i^0 \text{ always has feasible points.}$$

On the other hand, the variable w is used here in the procedure which updates the penalty parameter c_i without computing the estimates $\bar{\lambda}(\cdot)$, $\bar{u}(\cdot)$ and the function $\bar{c}(\cdot)$.

Now, in the first algorithm the continuity of $\bar{c}(\cdot)$ is sufficient to the existence of a bound to $\{c_i\}$ whenever the sequence $\{x_i\}$ remains bounded. As no such continuous function is employed in the simpler procedure to update c_i introduced in this chapter, it is interesting to emphasize how the penalty parameter is proved to be finite. [see Proposition I.2.C]. The finiteness of c_i results directly from the way c_i is updated and the existence of a $r < \infty$ such that :

$$\|\hat{\lambda}_i\|_1 \leq r \quad \text{for all } i$$

where $\hat{\lambda}_i \in \mathbb{R}^m$ is the multiplier vector associated with $\hat{p}_i$ at the solution of $\overline{QP}_i^0$.

Nevertheless, the boundedness of $\{x_i\}$ is still linked with c_i being increased only finitely often because the existence of the bound on $\|\hat{\lambda}_i\|_1$ could only be established to bounded sequences $\{x_i\}$.

In order to explain this remark better, we point out the role played by the several assumptions.

First recall that :

$$\left. \begin{array}{l} f \text{ bounded from below} \\ V^{j'} \triangleq \{x \mid g^{j'}(x) \leq 0 \text{ compact} \\ \text{for some } j \in \underline{m} \end{array} \right\} \Rightarrow \{x_i\} \text{ is bounded.}$$

Clearly,

$$\{x_i\} \text{ bounded} \Rightarrow \text{there exists a compact set } U \text{ such that} \\ \{x_i\} \subset U.$$

But, as shown in Lemma 3.3 of Han (1977), [see AP.2] we have :

$$\left. \begin{array}{l} g^j \text{ convex for all } j \\ V^0 \triangleq \{ \underset{+}{g(x)} < 0 \} \neq \emptyset \\ C_3 \\ \{x_i\} \subset U \text{ compact} \end{array} \right\} \Rightarrow \text{there exists a } r < \infty \text{ such that} \\ \text{for all } i \quad \|\hat{\lambda}_i\|_1 \leq r.$$

Now, the set $V \triangleq \{x | g(x) \leq 0\}$ is the intersection of all the sets,

$$V^j \triangleq \{\dot{x} | g^j(\dot{x}) \leq 0\}, \quad j \in \underline{m}.$$

Then,

$$\left. \begin{array}{l} V^j = \{x | g^j(x) \leq 0\} \text{ compact} \\ \text{for some } j \in \underline{m} \end{array} \right\} \Rightarrow V \text{ compact}.$$

The conditions assumed by Theorem I.2.D are identical to the conditions of Theorem 3.3 of Han (1977), see AP.3, with the addition of C5. In fact C5 was required in order to make sure that the procedure to update c_i would not lead to $c_i \rightarrow \infty$. [see proof of Proposition I.2.C]. Therefore, the further requirement of C5 is linked with the fact that here the algorithm updates the penalty parameter automatically while Han (1977) assumes that an appropriate constant value for the penalty parameters is chosen from the start.

Another difference between the algorithms is that our step

length procedure does not require exact minimization along the search direction, therefore involving less computation.

APPENDIX (Chapter I.2)

AP.1 Sensitivity Result [see Luenberger (1969)].

Let f and G be convex and suppose x_0, x_1 are the respective solution of :

$$\begin{array}{ll} \min f(x) & \min f(x) \\ \text{s.t.} & ; \quad \text{s.t.} \\ G(x) \leq z_0 \in \mathbb{R}^m & G(x) \leq z_1 \in \mathbb{R}^m . \end{array}$$

Then,

$$\langle z_1 - z_0, z_1^* \rangle \leq f(x_0) - f(x_1) \leq \langle z_1 - z_0, z_0^* \rangle$$

where,

z_0^* and z_1^* are the multipliers associated with x_0 and x_1 respectively.

AP.2 Paraphrase of Lemma 3.3 in Han (1977).

Let f and $g^j, j \in \underline{m}$, be continuously differentiable and let the following conditions be satisfied:

- (i) $g^j(\cdot)$ are convex for all $j \in \underline{m}$
- (ii) $V^0 \triangleq \{x \in \mathbb{R}^n \mid g(x) < 0\} \neq \emptyset$.
- (iii) $B_i \in L^+$ for all i

then

for any compact set $U \subset \mathbb{R}^n$ and any vector norm $\|\cdot\|$, there exists $r > 0$ such that, for all $x_i \in U$ we have :

$$\|\hat{\lambda}_i\| \leq r \quad \text{where } \hat{\lambda}_i \text{ is the multiplier vector of } \overline{QP}_i^0 .$$

AP.3 Paraphrase of Theorem 3.3 given in Han (1977) to the algorithm suggested there.

Let f and g^j , $j \in \underline{m}$, be continuously differentiable and let the following conditions be satisfied:

- (i) $g^j(\cdot)$ are convex for all $j \in \underline{m}$
- (ii) $V^0 \triangleq \{x \in \mathbb{R}^n \mid g(x) < 0\} \neq \emptyset$
- (iii) $V \triangleq \{x \in \mathbb{R}^n \mid g(x) \leq 0\}$ is compact
- (iv) f is bounded from below
- (v) $B_i \in L^+$ for all i .

Then, if $\{x_i\}$ is an infinite sequence constructed by the algorithm, for any starting point x^0 there exists a $\bar{r} > 0$ such that :

if $\|\hat{\lambda}_i\|_\infty \leq \bar{r}$ for all i , then :

any accumulation point x^* of $\{x_i\}$ belongs to D .

I.3 SECANT UPDATE CASE

I.3.1 Introduction

The satisfaction of assumption H3 by the sequence of matrices $\{B_i\}$ can, of course, be achieved if B_i is set equal to a fixed positive definite matrix, for instance the unity matrix. Although the resultant method would be globally convergent its rate of convergence would in general be at most Q-linear (see AP.1). This makes the use of a fixed positive definite matrix in the quadratic sub-problem unattractive because to be competitive the algorithm should be superlinearly convergent (see AP.1).

In the unconstrained minimization of a function $f(x)$ through quasi-Newton methods, a standard technique for proving that the method is Q-superlinearly convergent (see AP.1) is to show that, as $\{x_i\}$ converges to x^* , $\{B_i\}$ converges to the Hessian matrix $\nabla_{xx} f(x^*)$. Although not necessary, the last convergence is indeed a sufficient condition for the achievement of a superlinear rate of convergence [see Dennis and More (1974)]. In the constrained case, the convergence of $\{B_i\}$ to $\nabla_{xx} L(x^*)$, the Hessian matrix of the Lagrangian function at x^* , is a sufficient condition to Q-superlinear convergence [see Han (1976)].

The first method to use, in the quadratic sub-problems, matrices converging to $\nabla_{xx} L(x^*)$ is due to Wilson (1963), but as in Newton's method to unconstrained problems, it requires the computation of second order derivatives which is undesirable.

The most direct approach for avoiding the computation of derivatives is to approximate them by difference quotients. In particular, the use of $n+1$ auxiliary points to interpolate the function $F : \mathbb{R}^n \rightarrow \mathbb{R}^m$ in a neighbourhood of x_i characterizes the 'general secant method' for solving the system of equations

$$F(x) = 0.$$

The method [see Ortega and Rheinboldt (1970)] is defined by:

I.3.1

$$x_{i+1} = x_i - J(x_i, H_i)^{-1} F(x_i)$$

where

$$J(x, H) = [F(x+He^1) - F(x), \dots, F(x+He^n) - F(x)]H^{-1}$$

and

with given auxiliary points x_i^k , $k = 1, \dots, n$

$$H_i = (x_i^1 - x_i, \dots, x_i^n - x_i).$$

In general, the secant method requires $n+1$ evaluations of the function F at each iteration - namely at the points x_i , x_i^1 , x_i^2 , ..., x_i^n - but it is well known that particular definitions of the auxiliary points can save the number of function evaluations. The greatest saving is obtained through the 'sequential $n+1$ - point secant method' suggested by Wolfe (1959) which uses the function evaluations computed in the last n previous iteration so that only one new evaluation is required-namely, at x_i .

Local convergence results for secant methods rely upon the operator J of I.2.I being a 'consistent approximation' [see definition in Ortega and Rheinboldt (1970)] to the derivative whose computation is being avoided. The operator J of Wolfe's method is not guaranteed to satisfy this condition; hence local convergence is not guaranteed for this method. In fact, examples have been found [see Ortega and Rheinboldt (1970)] showing that it does not always converge.

For unconstrained minimization a secant method, which is locally convergent at a superlinear rate yet requires few function evaluations - namely, two new evaluations by iteration - is suggested by Polak (1974). In this method a sequence of matrices $\{H_i\}$ is constructed such that H_i differs from H_{i-1} by only one column j which is set as the vector below:

I.3.II

$$(1/\epsilon_i)[F(x_i + \epsilon_i e_j) - F(x_i)]$$

where

$$\epsilon_i = \|x_i - x_{i-1}\|.$$

In the next section an extension of the above procedure to

constrained problems, suggested by Mayne and Polak (1982), will be used to build an algorithm derived from the Basic Algorithm of the first chapter.

Now, when using a secant update, assumption H3 becomes too strong for practical purposes. On the other hand, the positive definiteness of the matrices $\{B_i\}$, although sufficient, is not a necessary condition to assure that :

I.3.III

$$p_i^T B_i p_i < 0.$$

If the second order sufficiency conditions are satisfied at x^* , B_i is close enough to $\nabla_{xx} L(x^*)$ and x_i is close enough to x^* then I.3.II will be satisfied even if B_i is not positive definite.

Hence, besides constructing the sequence of matrices $\{B_i\}$ such that

$$B_i \rightarrow \nabla_{xx} L(x^*) \text{ as } x_i \rightarrow x^*,$$

to ensure global convergence we need to guarantee that I.3.III will be satisfied also when x_i is not close to x^* . This will be done adding a multiple of the unity matrix I to the matrix H_i obtained through the secant update yielding :

I.3.IV

$$B_i = H_i + \gamma_i I$$

where γ_i is a nonnegative number which is set equal to zero at the start of each iteration and is then increased if some tests are not satisfied. The addition of a multiple of the unity matrix to the matrix $\nabla_{xx} f(x_i)$, in unconstrained optimization, is due to Goldfeld, Quandt and Trotter (1966). But, as H_i is intended to be an approxi-

mation to $\nabla_{xx} L(x_i)$ rather than this matrix itself, I.3.IV is perhaps more closely related to the works of Levenber (1944) and Marquardt (1963) which, in the context of nonlinear least squares, also added a multiple of the unity matrix to an approximation of the Hessian matrix.

The resultant algorithm is shown to generate sequences $\{x_i\}$ with all the accumulation points x^* lying on the set D of Kuhn-Tucker points for GMP. Furthermore, the rate of convergence of x_i to x^* in D is Q-superlinear.

I.3.2 The Algorithm

Consider the general mathematical programming problem (GMP) defined in the first chapter.

Let the maps $\psi, \gamma, \hat{\psi}, \hat{\gamma}, \bar{\lambda}, \bar{u}, T$ and θ be as defined in section I.1.2.

Again let $(p_i, w_i, \lambda_i, t_i, s_i, v_i)$ denote the solution and multipliers of the quadratic problems QP_i^w (see I.1.2) and $\tilde{p}_i$ the minimum norm solution, if a solution with norm smaller than $\|p_i\|$ exists, of the system:

$$g^j(x_i + p_i) + g_x^j(x_i)p = 0, \quad j \in \hat{I}_1 = \underline{\Delta}\{j \in \underline{m} \mid \lambda_i^j > 0\}$$

$$h^j(x_i + p_i) + h_x^j(x_i)p = 0, \quad j \in \underline{\ell}.$$

If such a solution does not exist $\tilde{p}_i$ is set equal to zero.

Now, suppose $\{H_i\}$ is a sequence of matrices generated by the secant update where H_{i+1} is obtained from H_i by replacing the column $i \pmod n$ by the vector :

I.3.V

$$(1/\Delta_i) [\nabla_x L(x_{i+1} + \Delta_i e_i, \bar{\lambda}(\bar{x}_{i+1}), \bar{\mu}(\bar{x}_{i+1})) - \nabla_x L(x_{i+1}, \bar{\lambda}(x_{i+1}), \bar{\mu}(x_{i+1}))]$$

with,

$$\Delta_i \triangleq \min\{\|x_{i+1} - x_i\|; \varepsilon_2\}$$

$$\varepsilon_2 \triangleq \text{small positive number}$$

$$e_i \triangleq i^{\text{th}} \text{ orthogonal basis vector of } \mathbb{R}^n$$

$L(\cdot, \cdot, \cdot)$ is the Lagrangian function.

The matrix B_i in $Q^P_i^W$ is obtained from H_i with the addition of a multiple of the unity matrix (see I.3.IV).

Algorithm

Data: $x_1 \in \mathbb{R}^n$; $H_1 \in L(\mathbb{R}^n)$; $c_0, b, K \in (0, \infty)$; $\delta > 1$; $\beta, \varepsilon \in (0, 1)$

Step 1: Set $i = 1$

Step 2: Set $B_i = H_i$; $\gamma_i = 0$

Step 3: If $c_{i-1} \geq \bar{c}(x_i)$, set $c_i = c_{i-1}$

Else,

$$\text{set } c_i = \max\{\delta c_{i-1}; \bar{c}(x_i)\}$$

Step 4: Find $(\bar{p}_i, w_i)$, the minimum norm solution of $Q^P_i^W$

Step 5: If,

$$(a) \quad \|\bar{p}_i\| > K$$

or

$$(b) \quad \theta(x_i, p_i, c_i) > \max\{-T(x_i); -\varepsilon\}.$$

Then,

set $\gamma_i = \gamma_i + b$; set $B_i = H_i + \gamma_i I$; Go to Step 4.

Else,

$$\text{set } p_i = \bar{p}_i$$

Step 6: If $w_i = 0$, compute $\tilde{p}_i$.

Else,

$$\text{set } \tilde{p}_i = 0$$

Step 7:

a) Compute k_i , the first integer in $\{0,1,\dots\}$ satisfying :

$$\gamma(x_i + \beta^{k_i} p_i + \beta^{2k_i} \tilde{p}_i, c_i) - \gamma(x_i, c_i) \leq \beta^{k_i} \theta(x_i, p_i, c_i) / 8$$

b) Set $x_{i+1} = x_i + \beta^{k_i} p_i + \beta^{2k_i} \tilde{p}_i$

Step 8: Update H_i to H_{i+1} by the secant update defined in the previous page.

Step 9: Set $i = i+1$, Go to Step 2

I.3.3 Global Convergence

In order to prove that any accumulation point of a bounded sequence constructed by the algorithm is a Kuhn-Tucker point of GMP, we will assume that the assumptions $H1'$ and $H2$ of the first chapter are satisfied.

That is:

$H1'$: f, g, h are three times continuously differentiable.

$H2$: for all $x \in R^n$ the vectors $\{\nabla g^j(x), j \in I(x); \nabla h^j(x), j \in I_\ell(x)\}$ are linearly independent, where :

$$I(x) \triangleq \{j \in \underline{m} \mid g^j(x) = \psi(x)\} \text{ and } I_\ell(x) \triangleq \{j \in \underline{\ell} \mid h^j(x) = \psi(x)\}.$$

Also,

$$D_c \triangleq \{x \in R^n \mid \tilde{\theta}(x, n, c) = 0\}, \quad n > 0 \text{ arbitrary}$$

where, $\tilde{\theta}(x, n, c) \triangleq \min_{p \in R^n} \{ (n/2) \|p\|^2 + \theta(x, p, c) \}$

$D \triangleq \{\hat{x} \in \mathbb{R}^n \mid \text{there exists } (\hat{\lambda}, \hat{\mu}) \text{ such that: } (\hat{x}, \hat{\lambda}, \hat{\mu}) \text{ is a K-T triple to GMP}\}$

Our first result was given (without proof) in Mayne and Polak (1978).

Proposition I.3.A

$$T(\bar{x}) = 0 \iff \bar{x} \in D.$$

Proof:

Given in the appendix at the end of the chapter.

Proposition I.3.B

Let $\{x_i\}$ be a bounded sequence constructed by the algorithm .
Then there exists a finite positive number M s.t. , for all i ,

$$-M \|y\|^2 \leq y^T H_i y \leq M \|y\|^2, \text{ for all } y \in \mathbb{R}^n .$$

Proof:

Given in the appendix at the end of the chapter.

Proposition I.3.C

Let $\{x_i\}$ be a bounded infinite sequence constructed by the algorithm.

Then, for all i , the loop between Step 5 and Step 4 is performed a finite number of times.

Proof:

For some i let k denote the number of times the loop is performed.

Because of the last proposition, the fact that $b > 0$ and the way B_i is constructed at Step 5, we have :

(i) for any $m' > 0$ there exists $k^0 < \infty$ such that :

$$y^T B_i y \geq m' \|y\|^2, \forall y \in \mathbb{R}^n \text{ and } k \geq k^0 .$$

Now, as shown in the proof of Proposition I.1.C [see AP.I.1], if $y^T B_i y \geq m' \|y\|^2$, for all $y \in R^n$ then,

$$\|\bar{p}_i\| \leq (1/m') \|\nabla f(x_i)\| + [(1/m')^2 \|\nabla f(x_i)\|^2 + 2(c_i/m') \psi(x_i)]^{1/2}.$$

Clearly, the right hand side of the above inequality diminishes with the increase of m' .

Then, from (i), there exists a $k^* < \infty$ such that, for all i ,

$$k \geq k^* \Rightarrow \|\bar{p}_i\| \leq K \quad (\text{constant of Step 5}).$$

That is, at each iteration test (a) of Step 5 will never be satisfied for k sufficiently large (but finite).

Now, as shown in the proof of Proposition I.1.A, for all i we have :

$$\theta(x_i, \bar{p}_i, c_i) \leq -\bar{p}_i^T B_i \bar{p}_i.$$

Then, from (i), for k sufficiently large but finite

$$\theta(x_i, \bar{p}_i, c_i) \leq -\bar{p}_i^T B_i \bar{p}_i \leq -\epsilon$$

and the proposition is proved, as both tests of Step 5 will never be satisfied for $k \geq \bar{k}$, say ($\bar{k} < \infty$).

Proposition I.3.D

Let $c \geq \bar{c}(x)$

Then,

$$x \in D_c \Leftrightarrow x \in D.$$

Proof:

Because D_c , D and $\bar{c}(\cdot)$ are defined exactly in the same way as in Mayne and Polak (1982), the proposition is identical to Proposition 5 proved in this paper.

Proposition I.3.E

Let $\{x_i\}$ be a bounded infinite sequence constructed by the algorithm.

Let I be a subsequence such that :

$$x_i \xrightarrow{I} x^*$$

with, for all $i \in I$,

$$c_i = c, \text{ for some } c > 0.$$

Then,

$$x^* \in D_c$$

Proof:

From Lemma I.1.K, there exists a $d_1 > 0$ such that, for any compact neighbourhood N of x^* and any $k > 0$ we have,

$$(i) \quad \hat{\gamma}(x_i, \alpha p_i + \alpha^2 \tilde{p}_i, c) - \gamma(x_i, c) \leq \alpha [\theta(x_i, p_i, c) + d_1 \alpha]$$

for all $(x_i, p_i, \tilde{p}_i) \in N \times \beta_k(0) \times \beta_k(0)$.

Also, from Proposition I.1.B,

$$(ii) \quad |\gamma(x_i + \alpha p_i + \alpha^2 \tilde{p}_i, c) - \hat{\gamma}(x_i, \alpha p_i + \alpha^2 \tilde{p}_i, c)| \leq \alpha \vartheta(\alpha, x_i, p_i, \tilde{p}_i)$$

where,

$$\vartheta(\alpha, x_i, p_i, \tilde{p}_i) \rightarrow 0 \text{ as } \alpha \rightarrow 0$$

for all $(x_i, p_i, \tilde{p}_i) \in N \times \beta_k(0) \times \beta_k(0)$.

Now, as $x_i \xrightarrow{I} x^*$ and $\|p_i\| \leq K$ by construction, (i) and (ii) will always be satisfied for all $i \in I$.

If $x^* \in D$ the proposition would follow from Proposition I.3.D.

Otherwise, suppose $x^* \notin D$.

Then, from Proposition I.3.A, $T(x^*) > 0$.

As $T(\cdot)$ is continuous and $x_i \rightarrow x^*$ we can then say that:

there exists a $\delta_1 > 0$ such that :

$$T(x_i) \geq \delta_1 \quad \text{for all } i \in I \text{ (} i \geq i_0, \text{ say).}$$

But then, as the test (b) at each iteration is ultimately not satisfied, we have :

$$\begin{aligned} \text{(iii)} \quad \theta(x_i, p_i, c) &\leq -d_2 \triangleq \max\{-\delta_1; -\epsilon\} < 0 \\ &\text{for all } i \in I \text{ (} i \geq i_0, \text{ say).} \end{aligned}$$

Now, from (ii), there exists an integer k' such that:

$$\text{(iv)} \quad \theta(\beta^{k'}, x_i, p_i, \tilde{p}_i) + d_1 \beta^{k'} \leq d_2/2 \quad \text{for all } i \in I.$$

Then, from (i), (ii), (iii) and (iv), for all $i \in I$ ($i \geq i_0$, say),

$$\gamma(x_i + \beta^{k'} p_i + \beta^{2k'} \tilde{p}_i, c) - \gamma(x_i, c_i) \leq \beta^{k'} \theta(x_i, p_i, c)/2 \leq -\beta^{k'} d_2/2.$$

But,

$$\beta^{k'} \theta(x_i, p_i, c)/2 < \beta^{k'} \theta(x_i, p_i, c)/8.$$

Then, at Step 7 a step length greater than or equal to $\beta^{k'}$ is found and:

$$\gamma(x_{i+1}, c) - \gamma(x_i, c) \leq -\beta^{k'} d_2/8 < 0$$

for all $i \in I$ ($i \geq i_0$, say).

Now, this result contradicts the fact that

$$x_i \xrightarrow{I} x^*.$$

The contradiction arose from the supposition that $x^* \notin D$.

Hence $x^* \in D$ and the proposition follows from Proposition I.3.D.

Theorem I.3.F

Any accumulation point x^* of a bounded infinite sequence constructed by the algorithm satisfies :

$$x^* \in D.$$

Proof:

Let I be the subsequence where

$$x_i \xrightarrow{I} x^*.$$

Because of Proposition I.3.D, Proposition I.3.E and the continuity of $\bar{c}(\cdot)$, all the assumptions of Theorem 1 in Mayne and Polak (1978) are satisfied. Hence, as $\{x_i\}$ is by hypothesis bounded, c_{i-1} is increased only finitely often. That is, for $i \geq i_0$, say, $c_i = c > 0$ (constant).

The theorem then follows Proposition I.3.E and I.3.D.

Observation:

Before proceeding with the analysis of the rate of convergence of the algorithm of this chapter, it is interesting to point out that the stronger result of Theorem I.1.M cannot be shown because we have not proved that $p_i \rightarrow 0$ in all convergent subsequences.

I.3.4 Rate of Convergence

In this section we will assume that the following conditions, defined in the first chapter are satisfied :

H1' (replacing H1); H2; H4.

In addition, we assume that :

$$x_i \rightarrow x^* \in D, \text{ as } i \rightarrow \infty$$

Let $(\bar{p}_i^0, w_i^0)$ denote the minimum norm solution of QP_i^W at the first performance of Step 4 at iteration i ($\gamma_i = 0$; $B_i = H_i$).

Considering the ℓ nonlinear inequalities of QP_i^W as 2ℓ linear inequalities (see Proposition I.1.A) let $(\lambda_i^0, t_i^0, s_i^0, v_i^0)$ be the multiplier vector associated with $(\bar{p}_i^0, w_i^0)$.

Proposition I.3.G

$$[\bar{p}_i^0, w_i^0, \lambda_i^0, (t_i^0 - s_i^0)] \rightarrow [0, 0, \lambda^*, \mu^*]$$

where

(λ^*, μ^*) is the multiplier associated with x^* .

Proof:

As shown in the proof of Theorem I.3.F, for i sufficiently large $c_i = c$.

As shown in Proposition 7 of Mayne and Polak (1978)

$$H_i \rightarrow \nabla_{xx} L(x^*, \lambda^*, u^*) \text{ as } x_i \rightarrow x^* \in D.$$

Then, for $\gamma_i = 0$, the quadratic problem QP_i^W can be considered as a perturbation of the following 'unperturbed' problem

$$\min_{(p,w)} \nabla f(x^*)p + \frac{1}{2}p^T \nabla_{xx} L(x^*, \lambda^*, u^*)p + cw$$

subject to,

$$g^j(x^*) + g_x^j(x^*)p \leq w \quad , \quad j \in \underline{m}$$

$$h^j(x^*) + h_x^j(x^*)p \leq w \quad , \quad j \in \underline{\ell}$$

$$-h^j(x^*) - h_x^j(x^*)p \leq w \quad , \quad j \in \underline{\ell}$$

$$w \geq 0$$

Let (p', w') be the minimum norm solution of this problem and (λ', t', s', v') the associated multiplier vector.

As $x^* \in D$, the minimum norm solution is the pair :

$$(i) \quad (p', w') = (0, 0).$$

Clearly, from H4, at this solution the 'unperturbed' problem satisfies the second-order sufficiency condition with strict complementarity slackness. Hence we can use Theorem 2.1 of Robinson (1974) to show that :

$$(ii) \quad (\bar{p}_i^0, w_i^0, \lambda_i^0, t_i^0, s_i^0, v_i^0) \rightarrow (p', w', \lambda', s', v').$$

As there exists only one multiplier vector associated with x^* at the solution of GMP, we have :

$$(iii) \quad \lambda^* = \lambda' \quad \text{and} \quad \mu^* = (t' - s').$$

The proposition follows from (i), (ii) and (iii).

Proposition 1.3.H

For $\gamma_i = 0$, for i large enough $w_i^0 = 0$ at the solution of QP_i^w .

Proof:

At the solution of QP_i^w we have:

$$(i) \quad c_i = \sum_{j \in \underline{m}} (\lambda_i^0)^j + \sum_{j \in \underline{\ell}} (t_i^0)^j + \sum_{j \in \underline{\ell}} (s_i^0)^j + v_i^0.$$

Also, by construction

$$(ii) \quad c_i \geq \sum_{j \in \underline{m}} |\bar{\lambda}(x_i)^j| + \sum_{j \in \underline{\ell}} |\bar{\mu}(x_i)^j| + b.$$

Note that for i sufficiently large, $c_i = c$.

Suppose there exists an infinite set $I^0 \subset \{1, 2, \dots\}$ such that:

$$(iii) \quad w_i^0 > 0, \text{ for all } i \in I^0.$$

Clearly, if $w_i^0 > 0$ only one of the multipliers $(t_i^0)^j, (s_i^0)^j$ can be different from zero, for all $j \in \underline{\ell}$

That is :

$$(iv) \quad \text{if } w_i > 0, (s_j^0)^j \neq 0 \Rightarrow (t_i^0)^j = 0 \text{ and } (t_i^0)^j \neq 0 \Rightarrow (s_j^0)^j = 0.$$

Let b_i^0 be the vector: $b_i^0 \triangleq (t_i^0 - s_i^0)$.

From (iv) we have:

$$(v) \quad (b_i^0)^j = \begin{cases} (t_i^0)^j, & \text{if } (t_i^0)^j \neq 0 \\ -(s_i^0)^j, & \text{if } (s_i^0)^j \neq 0 \\ 0, & \text{if } (t_i^0)^j = (s_i^0)^j = 0. \end{cases}$$

From the last proposition we have :

$$(vi) \quad b_i^0 \rightarrow \mu^* ; \quad \lambda_i^0 \rightarrow \lambda^* .$$

Also, as shown in Proposition 2 of Mayne and Polak (1978)

$$(vii) \quad \bar{\lambda}(x_i) \rightarrow \lambda^* ; \quad \bar{\mu}(x_i) \rightarrow \mu^* .$$

From (vi) and (vii), there exists an i_1 such that :

$$(viii) \quad \sum_{j \in \underline{m}} | |\bar{\lambda}_i^j(x_i)| - (\lambda_i^0)^j | \leq b/4, \quad \text{for all } i \geq i_1$$

and

$$(ix) \quad \sum_{j \in \underline{\ell}} | |\bar{u}_i^j(x_i)| - |(b_i^0)^j| | \leq b/4, \quad \text{for all } i \geq i_1.$$

But, from (v),

$$| (b_i^0)^j | = (s_i^0)^j + (t_i^0)^j.$$

Hence, from (ii), (viii) and (ix), we have :

$$(x) \quad c \geq \sum_{j \in \underline{m}} (\lambda_i^0)^j + \sum_{j \in \underline{\ell}} (t_i^0)^j + \sum_{j \in \underline{\ell}} (s_i^0)^j + b/2$$

for all i sufficiently large, $i \in I^0$.

From (x) and (i), we obtain:

$$v_i^0 \geq b/2 > 0, \quad \text{for all } i \text{ sufficiently large, } i \in I^0.$$

This contradicts (iii) because of the complementarity condition.

Hence, no such subsequence I^0 exists and the proposition is proved.

Proposition I.3.I

For $i \geq i_0$, say, the tests of Step 5 will never be satisfied with $\gamma_i = 0$.

Proof:

As shown in Proposition I.3.G.,

$$\bar{p}_i^0 \rightarrow 0 \quad \text{as } i \rightarrow \infty.$$

Hence, for $i \geq i_0$, say, $\|\bar{p}_i^0\| \leq K$ and test (a) is never satisfied, for i sufficiently large.

Now, as shown in the proof of Proposition 8 in Mayne and Polak (1978), for x_i close enough to x^* there exists a unique Kuhn-Tucker point for the problem $QP(x_i, H_i)$ below:

$$\min_{p \in \mathbb{R}^n} \nabla f(x_i)p + \frac{1}{2}p^T H_i p$$

subject to:

$$g_i^j(x_i) + g_x^j(x_i)p \leq 0, \quad j \in \underline{m}$$

$$h_i^j(x_i) + h_x^j(x_i)p = 0, \quad j \in \underline{\ell}$$

Hence, from the last proposition, $\bar{p}_i^0$ is the unique solution for the above problem and, therefore, the proposition follows from Proposition 11 in Mayne and Polak (1978) and the fact that $T(x_i) \rightarrow 0$ as $x_i \rightarrow x^*$.

Theorem I.3.J

$$x_i \rightarrow x^* \in D, \quad \text{as } i \rightarrow \infty,$$

Q - superlinearly.

Proof:

Because of Proposition I.3.H and Proposition I.3.I, for $i \geq i_0$, say, the algorithm of section I.3.2 coincides with the algorithm of Mayne and Polak (1978). Hence, the result follows directly from Theorem 4 of this paper.

I.3.5 Remarks

The use of the secant update given by I.3.V and the test of Step 5 make the algorithm of this chapter very similar to the algorithm of Mayne and Polak (1982). There exists two features where the algorithms are different, namely :

- The structure of the quadratic sub-problem where the matrix H_i obtained by the secant update is tentatively used.
- The nature of the alternative procedure adopted in case the solution of the quadratic sub-problem above is considered to be not satisfactory.

The difference between the quadratic sub-problems is that here we use QP_i^w instead of QP_i^0 .

As QP_i^w is always feasible, it is not necessary to test the existence of solution to the quadratic sub-problem and, consequently, the algorithm of page 84 has one test less than the algorithm of Mayne and Polak (1982).

Another effect of using QP_i^w is given in Proposition I.1.A, that is:

for any value of c_i we have :

$$p_i^T B_i p_i > 0 \Rightarrow \theta(x_i, p_i, c_i) < 0 .$$

The implication above is not necessarily true when using QP_i^0 , unless c_i satisfies the further condition

$$c_i \geq \|(\lambda_i^0, \mu_i^0)\|_1$$

where (λ_i^0, μ_i^0) is the multiplier vector associated with the solution p_i^0 of QP_i^0 . As a result, the solution p_i of QP_i^w with $B_i = H_i$ is more likely to satisfy the test

$$\theta(x_i, p_i, c_i) < -T(x_i)$$

than the correspondent solution of QP_i^0 .

Therefore, because of this last conclusion and the consistency of the constraints in QP_i^w , the algorithm presented here will probably activate the alternative procedure less often than the algorithm of Mayne and Polak (1982).

We now compare the alternative procedure of both algorithms. In the algorithm of Mayne and Polak (1982) the search direction used when the solution of QP_i^0 fails to satisfy some tests is obtained by solving :

$$\min_{p \in R} \{ p^T \tilde{H} p + \theta(x_i, p, c_i) \}$$

where:

$$\tilde{H} = \eta I, \quad \eta \text{ being a positive constant.}$$

Here, the alternative direction is the solution of :

$$\min_{p \in R} \{ p^T [H_i + \gamma_i I] p + \theta(x_i, p, c_i) \}$$

where :

H_i is obtained by the secant up date

and

$\gamma_i > 0$ is automatically set.

That is, while the strategy of Mayne and Polak (1982) to stabilize the locally superlinear convergent secant method is to revert to the 'steepest [with respect to the norm $\|p\| = (p^T \bar{H} p)^{1/2}$] descent direction', here we add a multiple of the unity matrix to the H_i obtained by the secant update.

Similar differences occur, in the context of unconstrained minimization of a function $f(x)$, between the method of Goldstein and Price (1967) and the method of Goldfeld, Quandt and Totter (1966), both suggested to stabilize Newton's method. Indeed, whenever the Hessian matrix $\nabla_{xx} f(x_i)$ is not positive definite or the direction obtained in the minimization of the quadratic approximation is nearly orthogonal to $\nabla f(x_i)$, Goldstein and Price (1967) switch to the steepest descent direction $-\nabla f(x_i)$, while Goldfeld, et al. (1966) prefer to add a multiple of the unity matrix to $\nabla_{xx} f(x_i)$ and use the resultant matrix, say

$$A_i = \nabla_{xx} f(x_i) + \gamma_i I, \gamma_i > 0,$$

in the quadratic form to be minimized.

Now, as pointed out by Fletcher (1980), the latter method (and by extension the method suggested here) has the advantage of using the quadratic information given by $\nabla_{xx} f(x_i)$ (H_i in our case), even when this matrix is not positive definite. In particular, if the Hessian (or H_i) is close to being positive definite then it may only be necessary to add a small multiple of I so as to give a good search direction. Because the curvature information is ignored by the alternative procedure of Goldstein and Price (1967), as it is also ignored by the first order method suggested by Mayne and Polak (1982), if the alternative iteration operates successively for a number of iterations, it is likely that slow oscillatory behaviour would be observed.

On the other hand, the computation of γ_i through the loop between Step 5 and Step 4 is clearly a disadvantage of the algorithm, although some devices can be used to reduce the number of times this loop is repeated. First note that for $i \geq i_0$, say, the tests of Step 5 are never satisfied so that the alternative procedure is not activated. (see Proposition I.2.J). In any case the loop is repeated only a finite number of times because we increase the eigenvalues of H_i by adding the term $\gamma_i I$ so that for γ_i sufficiently large the matrix $B_i = H_i + \gamma_i I$ will be positive definite (see Proposition I.3.C). Hence, it would be advantageous if the matrix H_i given by the secant update remains positive definite during the initial iterations.

Accordingly, unless x_1 is close to x^* and a good approximation to $\nabla_{xx} L(x^*, \lambda^*, \mu^*)$ is known, an adequate choice for H_1 is :

$$H_1 = \eta I, \quad \eta > 0.$$

Actually, the idea of giving an initial bias towards the steepest descent direction which would be gradually reduced as x_i approaches x^* can be exploited with the use of the function $T(\cdot)$. Recall that $T(x)$ is a continuous function which assumes positive values for all $x \in \mathbb{R}^n$ except at Kuhn-Tucker points for GMP where it is equal to zero. Hence, a possibly good choice to η would be $\eta = T(x_1)$.

This would take into account eventual bad scaling and would make the positive eigenvalues of H_1 depend on how far x_1 is from the solution.

Similar reasoning suggests the following modification to Step 5.

in case any test is satisfied set,

$$\gamma_i = \gamma_i + b \quad ; \quad B_i = H_i + \gamma_i T(x_i) I.$$

Now because of Proposition I.1.A, it is easily seen that, if test (b) is satisfied,

$$p_i^T B_i p_i < \varepsilon_2 \triangleq \min\{T(x_i) ; \varepsilon\}.$$

Hence, another modification of Step 5 which can make the increase on γ_i in a more efficient way is :

in case test (b) is satisfied set,

$$\gamma_i = \gamma_i + (\varepsilon_2 - p_i^T B_i p_i)/T(x_i) + b; \quad B_i = H_i + \gamma_i T(x_i)I.$$

Note that with b being chosen small, the above expression for γ_i avoids excessive change on the matrix H_i obtained through the secant update. This feature is particularly useful when H_i is close to being positive definite.

These modifications are likely to keep the number of times the loop is repeated low.

In addition, the tests of Step 5 can be made more difficult to satisfy with ε small (say 10^{-8}) and K large (say, $10 \|\bar{p}_1\|$), as suggested by Mayne and Polak (1982).

Hidden in Theorem I.3.J is the fact that the computation of Step 7 will lead to step length of unity for i sufficiently large.

This is proved in Proposition 12 of Mayne and Polak (1978) and the proof relies upon the convergence,

$$H_i \rightarrow \nabla_{xx} L(x^*, \lambda^*, u^*) \text{ as } i \rightarrow \infty$$

and the use of the search arc $\alpha p_i + \alpha^2 \tilde{p}_i$ in the computation of the

step length α .

Our next theorem shows that the above convergence, although sufficient, is not necessary for the achievement of step length of unity when using the search arc. We start with some further definitions.

Let N_i be the matrix whose columns are the gradients $g_x^j(x_i)$, $j \in I(x^*) \triangleq \{j \in \mathbb{M} \mid g^j(x^*) = 0\}$. Let $r(x_i)$ denote the vector whose components are $g^j(x_i)$, $j \in I(x^*)$, ordered in the same way as the columns of N_i . Let Y_i denote a matrix whose columns are orthogonal basis vector for the null space of N_i . That is :

$$N_i^T Y_i = 0 \quad \text{for all } i.$$

$$\text{Let } P_i \triangleq Y_i Y_i^T.$$

Let $a_i \sim b_i$ mean that there exists an integer i_1 and positive constants d_1, d_2 such that :

$$d_1 \leq (a_i/b_i) \leq d_2, \quad \text{for all } i \geq i_1.$$

Let $a_i \approx b_i$ mean that there exists an integer i_2 and a positive constant d_3 such that:

$$a_i \leq d_3 b_i, \quad \text{for all } i \geq i_2.$$

Let $\mathcal{O}(\cdot)$ stands for an expression whose ratio to the term inside the brackets tends to zero as $i \rightarrow \infty$.

Consider the problem QP_i^0 , defined in I.1.1, with B_i being given by:

$$B_i = L_i + A_i, \quad \text{for all } i,$$

where:

$L_i \in L(R^n)$ satisfies : $L_i \rightarrow \nabla_{xx} L(x^*, \lambda^*, \mu^*)$, as $i \rightarrow \infty$,

and

$\{A_i\}$ is any sequence of positive definite matrices with

$A_i \rightarrow A^* \in L(R^n)$ as $i \rightarrow \infty$.

Lemma I.3.K

Let the assumptions of the last section be satisfied.

Let $\{B_i\}$ be constructed as above.

Then,

for i sufficiently large QP_i^0 will correctly predict the active constraints to GMP at x^* .

Proof:

Given in the appendix at the end of the chapter.

Theorem I.3.L

Let the assumptions of the last section be satisfied.

Let p_i be obtained by solving QP_i^0 with $\{B_i\}$ being constructed as above and c_i being updated as in Step 3.

Then,

the step length obtained through the search arc of Step 7 will be unity for all i sufficiently large.

Proof:

Because of the last lemma, two results obtained in the proof of Proposition 12 of Mayne and Polak (1978) are easily seen to remain valid. Namely :

$$(i) \quad \|\tilde{p}_i\| \approx \|p_i\|^2$$

and

$$(ii) \quad \gamma(x_i + p_i + \tilde{p}_i, c_i) - \gamma(x_i, c_i) = \phi(x_i, c_i) + \hat{\epsilon}_i$$

where,

$$\phi(x_i, c_i) = \theta(x_i, p_i, c_i) + \langle \nabla f(x_i), \tilde{p}_i \rangle + \frac{1}{2} \langle p_i, \nabla_{xx} f(x_i) p_i \rangle$$

and

$$\hat{\epsilon}_i \approx \|p_i\|^3.$$

As shown in the proof of Lemma 1.3.K,

$$(iii) \quad \lambda_i^* \rightarrow \lambda^*, \quad \mu_i \rightarrow \mu^*, \quad \text{as } x_i \rightarrow x^*$$

where,

(λ_i, μ_i) is the multiplier vector associated with p_i .

Then,

$$\nabla_{xx} L(x_i, \lambda_i, \mu_i) \rightarrow \nabla_{xx} L(x^*, \lambda^*, \mu^*)$$

and,

$$(iv) \quad (L_i - \nabla_{xx} L(x_i, \lambda_i, \mu_i)) \rightarrow 0.$$

Now, the step length is of unity if,

$$T_i \triangleq \gamma(x_i + p_i + \tilde{p}_i, c_i) - \gamma(x_i, c_i) - \theta(x_i, p_i, c_i)/8 \leq 0.$$

By (ii),

$$(v) \quad T_i = (7/8)\theta(x_i, p_i, c_i) + \langle \nabla f(x_i), \tilde{p}_i \rangle + \frac{1}{2} \langle p_i, \nabla_{xx} f(x_i) p_i \rangle + \hat{\epsilon}_i.$$

But,

$$\nabla f(x_i) = -B_i p_i - \lambda_i^T g_x(x_i) - \mu_i^T h_x(x_i).$$

Then,

$$(vi) \quad \langle \nabla f(x_i), \tilde{p}_i \rangle = -\langle p_i, B_i \tilde{p}_i \rangle - \lambda_i^T g_x(x_i) \tilde{p}_i - \mu_i^T h_x(x_i) \tilde{p}_i.$$

By hypothesis, $L_i \rightarrow \nabla_{xx} L(x^*, \lambda^*, \mu^*)$ and $A_i \rightarrow A^*$.

Hence, $\{B_i\}$ is bounded and, from (i)

$$(vii) \quad |\langle p_i, B_i \tilde{p}_i \rangle| \approx \|\tilde{p}_i\|^3.$$

From (vi), (vii) and taking into account the way $\tilde{p}_i$ is computed, we have :

$$(viii) \quad \langle \nabla f(x_i), \tilde{p}_i \rangle = \lambda_i^T g(x_i + p_i) + \mu_i^T h(x_i + p_i) + \tilde{\varepsilon}_i$$

where,

$$\tilde{\varepsilon}_i \approx \|p_i\|^3.$$

Now, from Proposition 1 of Mayne and Polak (1978), (iii), the convergences

$\bar{\lambda}(x_i) \rightarrow \lambda^*$; $\bar{\mu}(x_i) \rightarrow \mu^*$ and the way c_i is updated, we have:

$$\theta(x_i, p_i, c_i) \leq -p_i^T B_i p_i.$$

Then, from (viii), (v) and expanding $g(x_i + p_i)$ and $h(x_i + p_i)$ to third order we have :

$$(ix) \quad T_i \leq (3/8)\theta(x_i, p_i, c_i) + \frac{1}{2}\langle p_i, (\nabla_{xx} L(x_i, \lambda_i, \mu_i) - B_i)p_i \rangle + \varepsilon_i'$$

where,

$$\varepsilon_i^1 \approx \|p_i\|^3,$$

Now,

$$(x) \quad \langle p_i, (\nabla_{xx} L(x_i, \lambda_i, u_i) - B_i) p_i \rangle = \langle p_i, (\nabla_{xx} L(x_i, \lambda_i, u_i) - L_i - A_i) p_i \rangle.$$

Hence, from (ix) and (x) we have :

$$T_i \leq (3/8)\theta(x_i, p_i, c_i) - p_i^T A_i p_i + o(\|p_i\|^2).$$

But, by hypothesis, A_i is positive definite. Then,

$$(xi) \quad T_i \leq (3/8)\theta(x_i, p_i, c_i) + o(\|p_i\|^2).$$

Now, again by the positive definiteness of A_i , we have :

$$Y_i^T L_i Y_i \text{ positive definite} \Rightarrow Y_i^T B_i Y_i \text{ positive definite.}$$

But, by H4 and the convergence $L_i \rightarrow \nabla_{xx} L(x^*, \lambda^*, u^*)$, for i sufficiently large $Y_i^T L_i Y_i$ is indeed positive definite.

Hence, a detailed examination shows that Proposition 10 of Mayne and Polak (1978) remains true. That is :

$$(xii) \quad \theta(x_i, p_i, c_i) \approx [\|r(x_i)\| + \|p_i^T \nabla f(x_i)\|^2].$$

From (xii) and (xi) we have :

$$(xiii) \quad T_i \approx -[\| r(x_i) \| + \| P_i \nabla f(x_i) \|^2] + o(\| p_i \|^2).$$

But, as proved by Powell (1978a),

$$\| p_i \| \sim \| r(x_i) \| + \| P_i \nabla f(x_i) \|.$$

Hence

$$(xiv) \quad \| p_i \|^2 \approx \| r(x_i) \|^2 + \| P_i \nabla f(x_i) \|^2.$$

But,

$$\| r(x_i) \|^2 + \| P_i \nabla f(x_i) \|^2 \rightarrow 0.$$

It is obvious from (xiv) and (xiii) that, for i sufficiently large, $T_i \leq 0$ and the theorem is proved.

We now compare Theorem I.3, L with previous results in the literature of quasi-Newton methods.

In the unconstrained case the conditions which ensure that the steps length along a descent direction p_i will be set equal to unity, for i sufficiently large, are often also sufficient to guarantee a superlinear rate of convergence. Among these conditions is the convergence, in direction and length, of the vectors p_i to the Newton iteration [see Theorem 6.4 of Dennis and Moré (1977)].

It has also been shown [see Theorem 8.9 of Dennis and Moré (1977)] for algorithms generating a sequence $\{B_i\}$ of positive definite matrices, that the convergence of B_i to the Hessian matrix is not necessary for asymptotic step length of unity and superlinear rate of convergence.

In the presence of constraints similar conditions were found to guarantee superlinear rate of convergence for variable metric algorithms [see Han (1976)] but these conditions are no longer sufficient for an asymptotic step length of unity. Because the constraints are linearized in the quadratic sub-problem, examples by Maratos (1978) show that no such result can be proved when using line search procedures. Hence, it is interesting to know whether the use of the search arc $(\alpha p_i + \alpha^2 \tilde{p}_i)$ makes sufficient conditions for superlinear convergence also sufficient for an asymptotic step length of unity.

Under assumptions similar to those of Theorem I.3.L, Han (1976) shows that the following condition is sufficient to ensure a Q-superlinear rate of convergence for variable metric algorithms.

I.3.VI

$$\lim_{i \rightarrow \infty} \frac{\| B_i - \nabla_{xx} L(z^*) (x_{i+1} - x_i) \|}{\| z_{i+1} - z_i \|} = 0$$

where,

$$z_i = (x_i, \lambda_i, \mu_i) \text{ and } z^* = (x^*, \lambda^*, \mu^*).$$

The next theorem shows that the above condition is also sufficient to ensure unity step length if the search arc is used. As the sequence of matrices $\{B_i\}$ is supposed to remain positive definite with eigenvalues bounded away from zero, the theorem is stated for the Basic Algorithm of Chapter I.1.

Theorem I.3. M

Consider the algorithm of page 21 and let assumptions H1, H2, H3 and H4 be satisfied.

If,

condition I.3.VI is satisfied

then,

for $i \geq i_0$, say, the computation of Step 4 will set a step length of unity.

Proof:

Note that the assumptions of Theorem I.1.M. are satisfied, so that:

$$z_i \rightarrow z^*, \quad \text{as } i \rightarrow \infty.$$

Clearly, this convergence implies:

$$(i) \quad \nabla_{xx} L(z_i) \rightarrow \nabla_{xx} L(z^*), \quad \text{as } i \rightarrow \infty.$$

Now, as in the proof of Theorem I.3.L, we can show that the step length is set equal to unity at Step 4 whenever $T_i \leq 0$, where :

$$(ii) \quad T_i \leq (7/16)\theta(x_i, p_i, c_i) + \frac{1}{2}\langle p_i, (\nabla_{xx} L(z_i) - B_i)p_i \rangle + o(\|p_i\|^2).$$

But, because of (i) and condition I.3.VI, we have :

$$(iii) \quad \langle p_i, (\nabla_{xx} L(z_i) - B_i)p_i \rangle = o(\|p_i\|^2).$$

Hence, from (ii), (iii), H3 and Proposition I.1.A there exists a $m_1 > 0$ such that :

$$T_i \leq -m_1 \|p_i\|^2 + o(\|p_i\|^2).$$

The theorem follows from the convergence : $\|p_i\| \rightarrow 0$.

Now, recall that in Theorem I.3.L we did not require B_i to be positive definite but assumed :

I.3.VII

$$B_i = H_i + A_i$$

with:

$$H_i \rightarrow \nabla_{xx} L(z^*) \quad \text{as } i \rightarrow \infty$$

and

A_i positive definite for all i .

Suppose A_i is of the form:

I.3.VIII

$$A_i = v_i (T(x_i) I + N_i C_i N_i^T)$$

where:

N_i is the matrix whose columns are the active gradients of the quadratic sub-problem.

C_i is any positive definite matrix

and

$$v_i > 0$$

In this case, if assumption H4 is satisfied, there exists a $v^* < \infty$ s.t. for all $v_i \geq v^*$ the matrix B_i given by I.3.VII is positive definite. Note however that, while a step length of unity can be shown (Theorem I.3.L) to be ultimately achieved, this result could not be proven using Theorem I.3.M because condition I.3.VI is not satisfied.

On the other hand, since $T(x_i) \rightarrow 0$ as $i \rightarrow \infty$ and $P_i N_i = 0$, the following limit is satisfied.

I.3.IX

$$\lim_{i \rightarrow \infty} \frac{\| P_i (B_i - \nabla_{xx} L(z^*)) P_i P_i \|}{\| P_i \|} = 0$$

where,

P_i is the projection matrix onto the null space of N_i .

Now, I.3.IX is shown by Powell (1978a) to be a sufficient condition for variable metric algorithms to achieve a two-step superlinear rate of convergence.

Hence, the use of an update of the form I.3.VII, with A_i given by I.3.VIII and v_i sufficiently large, ensures the global convergence of the Basic Algorithm at a two-step superlinear rate. A related

idea will be employed to update B_i in the next chapter.

A result similar to Theorem I.3.M can be proved if the condition I.3.IX is satisfied and the further assumption below is made.

I.3.X

there exists a $r_1 > 0$ such that:

$$\|p_i\| > r_1 \|p_{i-1}\|^2 \quad \text{for all } i.$$

It is interesting to note that this condition is not intended to induce a quicker convergence of p_i to zero. On the contrary, it avoids the case where $\|p_i\| / \|p_{i-1}\|^2 \rightarrow 0$. As will be discussed in I.4.5, condition I.3.X is also useful to establish Q-superlinear rate of convergence of variable metric algorithms which, otherwise, could only be shown to achieve a R-superlinear rate.

Theorem I.3.N

Consider the algorithm of page 21 and let assumptions H1', H2, H3 and H4 be satisfied.

If,

conditions I.3.IX and I.3.X are satisfied,

then,

for $i \geq i_0$, say, the computation of Step 4 will set the step length equal to unity.

Proof:

Again by Theorem I.1.M, we have :

$$(i) \quad z_i \rightarrow z^* \text{ and } \nabla_{xx} L(z_i) \rightarrow \nabla_{xx} L(z^*), \text{ as } i \rightarrow \infty,$$

Also, the step length is set equal to unity if $T_i \leq 0$, where:

$$(ii) \quad T_i \leq (7/16)\theta(x_i, p_i, c_i) + \frac{1}{2}\langle p_i, (\nabla_{xx} L(z_i) - B_i)p_i \rangle \\ + o(\|p_i\|^2).$$

But, by Proposition I.1.A and assumption H3 there exists a $m_1 > 0$ such that

$$(iii) \quad (7/16)\theta(x_i, p_i, c_i) \leq -m_1 \|p_i\|^2.$$

Now, as shown in the proof of Proposition 10 of Mayne and Polak (1978), $p_i = p_i^1 + p_i^2$ where :

p_i^1 lies on the null space of N_i^T

and

p_i^2 is the minimum norm solution of :

$$N_i^T p + r(x_i) = 0.$$

That is:

$$(iv) \quad p_i^1 = P_i p_i$$

and

$$(v) \quad \|p_i^2\| \sim \|r(x_i)\|.$$

Now, recall that $I(x^*)$ was defined as ;

$$I(x^*) \triangleq \{j \in \underline{m} \mid g^j(x^*) = 0\}.$$

Because $\{B_i\}$ is bounded, $\|p_i\| \rightarrow 0$ and H4 gives a strict complementarity slackness, it is possible to show [Powell (1978a)] that for i sufficiently large we have :

$$g^j(x_i) + g_x^j(x_i)p_i = 0, \quad \forall j \in I(x^*).$$

Also,

$$h^j(x_j) + h_x^j(x_i)p_i = 0, \quad \forall j \in \underline{\ell}.$$

Hence, as shown in the proof of Proposition 12 of Mayne and Polak (1978) we have :

$$|g^j(x_{i-1} + p_{i-1} + \tilde{p}_{i-1})| \approx \|p_{i-1}\|^3$$

and

$$|h^j(x_{i-1} + p_{i-1} + \tilde{p}_{i-1})| \approx \|p_{i-1}\|^3.$$

Then, by the definition of $r(x_i)$ (see page 101) and (v), we have :

$$(vi) \quad \|p_i^2\| \approx \|p_{i-1}\|^3.$$

But, from (iv) we have :

$$\begin{aligned}
\text{(vii)} \quad \langle p_i, (\nabla_{xx} L(z_i) - B_i) p_i \rangle &= \langle (p_i^1 + p_i^2), (\nabla_{xx} L(z_i) - B_i) (p_i^1 + p_i^2) \rangle \\
&= \langle p_i p_i, (\nabla_{xx} L(z_i) - B_i) p_i p_i \rangle \\
&\quad + 2 \langle p_i p_i, (\nabla_{xx} L(z_i) - B_i) p_i^2 \rangle \\
&\quad + \langle p_i^2, (\nabla_{xx} L(z_i) - B_i) p_i^2 \rangle .
\end{aligned}$$

From the boundedness of P_i and condition I.3.IX we have :

$$\text{(viii)} \quad |\langle p_i p_i, (\nabla_{xx} L(z_i) - B_i) p_i p_i \rangle| = o(\|p_i\|^2).$$

From the boundedness of the matrices involved and (vi), there exists a $M_1 < \infty$ such that :

$$|\langle p_i p_i, (\nabla_{xx} L(z_i) - B_i) p_i^2 \rangle| \leq M_1 \|p_i\| \|p_{i-1}\|^3 .$$

Hence, from the convergence $p_{i-1} \rightarrow 0$ and condition I.3.X, we have :

$$\text{(x)} \quad |\langle p_i p_i, (\nabla_{xx} L(z_i) - B_i) p_i^2 \rangle| = o(\|p_i\|^2) .$$

Similarly,

$$\text{(x)} \quad |\langle p_i^2, (\nabla_{xx} L(z_i) - B_i) p_i^2 \rangle| = o(\|p_i\|^2)$$

Now, (x), (iv) and (viii) into (vii) lead to:

$$\text{(xi)} \quad |\langle p_i, (\nabla_{xx} L(z_i) - B_i) p_i \rangle| = o(\|p_i\|^2) .$$

The theorem then follows from (xi), (iii), (ii) and the convergence $p_i \rightarrow 0$.

APPENDIX (Chapter I.3)

AP.1 Quotient Convergence Factors [From Ortega and Rheinboldt (1970)]

Definition:

Let $\{x_i\} \subset \mathbb{R}^n$ be any convergent sequence with limit x^* .

Then the quantities

$$Q_p\{x_i\} \triangleq \begin{cases} 0 & \text{if } x_i = x^*, \text{ for all but finitely many } i \\ \limsup_{i \rightarrow \infty} \frac{\|x_{i+1} - x^*\|}{\|x_i - x^*\|^p} & \text{if } x_i \neq x^*, \text{ for all but finitely many } i \\ +\infty & \text{otherwise} \end{cases}$$

defined for all $p \in [1, \infty)$, are the quotient convergence factors, or Q-factors, for short, of $\{x_i\}$ with respect to the norm $\|\cdot\|$ on $\mathbb{R}^n$.

Definition:

Let $C(A, x^*)$ denote the set of all sequences with limit x^* generated by an iterative process A . Then,

$$Q_p(A, x^*) = \sup\{Q_p\{x_i\} \mid \{x_i\} \in C(A, x^*)\}$$

are the Q-factors of A at x^* with respect to the norm in which the $Q_p\{x_i\}$ are computed.

Definition:

Whenever $Q_1(A, x^*) = 0$, we say that the process has Q-super-

linear convergence at x^* , while, if $0 < Q_1(A, x^*) < 1$ in some norm, the convergence is called Q-linear.

Proof of Proposition I.3.A

$$(a) \quad \Rightarrow T(\bar{x}) = 0 \Rightarrow \begin{cases} \psi(\bar{x}) = 0 \Rightarrow \bar{x} \text{ is a feasible point to GMP} \\ \nabla f(\bar{x}) + g_X^T(\bar{x}) \bar{\lambda}_+(\bar{x}) + h_X^T(\bar{x}) \bar{\mu}(\bar{x}) = 0. \end{cases}$$

Then, it only remains to show that $(\bar{x}, \bar{\lambda}_+(\bar{x}))$ satisfies the complementarity condition, i.e.,

$$\bar{\lambda}_+^j(\bar{x}) g^j(\bar{x}) = 0, \quad j \in \underline{m}.$$

Now, consider the definition of $(\bar{\lambda}(\bar{x}), \bar{\mu}(\bar{x}))$. That is :

$$\begin{aligned} (\bar{\lambda}(\bar{x}), \bar{\mu}(\bar{x})) &\triangleq \arg \min_{(\lambda, \mu)} \{ \|\nabla f(\bar{x}) + g_X^T(\bar{x}) \lambda + h_X^T(\bar{x}) \mu\|^2 + \\ &+ \sum_{j \in \underline{m}} [(\psi(\bar{x}) - g^j(\bar{x}))^2 (\lambda^j)^2]^2 + \sum_{j \in \underline{\ell}} [(\psi(\bar{x}) - |h^j(\bar{x})|)^2 (\mu^j)^2]^2 \}. \end{aligned}$$

As $(\bar{\lambda}(\bar{x}), \bar{\mu}(\bar{x}))$ is the solution of the above problem and taking into account that $\nabla f(\bar{x}) + g_X^T(\bar{x}) \bar{\lambda}_+(\bar{x}) + h_X^T(\bar{x}) \bar{\mu}(\bar{x}) = 0$,

$$\begin{aligned} \|\nabla f(\bar{x}) + g_X^T(\bar{x}) \bar{\lambda}(\bar{x}) + h_X^T(\bar{x}) \bar{\mu}(\bar{x})\|^2 + \sum_{j \in \underline{m}} (-g^j(\bar{x}) \bar{\lambda}^j(\bar{x}))^2 \\ \leq \sum_{j \in \underline{m}} (-g^j(\bar{x}) \bar{\lambda}_+^j(\bar{x}))^2. \end{aligned}$$

But, by definition of $\bar{\lambda}_+(\cdot)$ we have :

$$(g^j(\bar{x}) \bar{\lambda}_+^j(\bar{x}))^2 \leq (g^j(\bar{x}) \bar{\lambda}^j(\bar{x}))^2, \quad \text{for all } j \in \underline{m}.$$

Then,

$$(\bar{\lambda}_+(\bar{x}), \bar{\mu}(\bar{x})) = (\bar{\lambda}(\bar{x}), \bar{\mu}(\bar{x})).$$

Now at the solution of the last problem, we must have :

$$\| \nabla f(\bar{x}) + g_X^T(\bar{x}) \bar{\lambda}(x) + h_X^T(\bar{x}) \bar{\mu}(x) \|^2 + \sum_{j \in \underline{m}} (-g^j(\bar{x}) \bar{\lambda}^j(\bar{x}))^2 = 0.$$

Hence,

$$\bar{\lambda}_+^j(\bar{x}) g^j(\bar{x}) = \bar{\lambda}^j(\bar{x}) g^j(\bar{x}) = 0.$$

(b) $\Leftarrow$

$$\bar{x} \in D \Rightarrow \begin{cases} \psi(\bar{x}) = 0. \\ \nabla f(\bar{x}) + g_X^T(\bar{x}) \bar{\lambda} + h_X^T(\bar{x}) \bar{\mu} = 0, \text{ for some } (\bar{\lambda}, \bar{\mu}). \\ \bar{\lambda}^j g^j(\bar{x}) = 0, \quad j \in \underline{m}. \\ \bar{\lambda} \geq 0. \end{cases}$$

But then $(\bar{\lambda}, \bar{\mu})$ would be the unique solution of the unconstrained minimization which defines $(\bar{\lambda}(\bar{x}), \bar{\mu}(\bar{x}))$ and, $T(\bar{x}) = 0$.

Proof Proposition I.3.B.

Consider the following sequence of matrices $\{G_i\}$

$$G_i \triangleq \nabla_{xx} L(x_{i+1}, \bar{\lambda}(x_{i+1}), \bar{\mu}(x_{i+1})).$$

As all the functions involved are continuous and $\{x_i\}$ is supposed to be bounded, there exists a $M_1 < \infty$, such that :

$$(i) \quad \|G_i\| \leq M_1, \text{ for all } i.$$

Now, let $i > n+1$.

Suppose, without loss of generality, that the n^{th} column of H_i is:

$$(1/\Delta_i) [\nabla_x L(x_{i+1} + \Delta_i e_n, \bar{\lambda}(\bar{x}_{i+1}), \bar{\mu}(\bar{x}_{i+1})) - \nabla_x L(x_{i+1}, \bar{\lambda}(\bar{x}_{i+1}), \bar{\mu}(x_{i+1}))].$$

Then, the norm of the difference between the n^{th} columns of H_i and G_i is :

$$\left\| \int_0^1 [\nabla_{xx} L(x_{i+1}, \bar{\lambda}(x_{i+1}), \bar{\mu}(x_{i+1})) - \nabla_{xx} L(x_{i+1} + t \Delta_i e_n, \bar{\lambda}(x_{i+1}), \bar{\mu}(x_{i+1}))] e_n dt \right\|.$$

Call $\bar{\Delta}$ the above norm.

As f, g, h are all three times continuously differentiable, there exists a $M_2 < \infty$ such that :

$$\bar{\Delta} \leq M_2 \int_0^1 \| t \Delta_i \| dt.$$

As $\Delta_i \leq \epsilon_2$ for all i , there exists a $M_3 < \infty$ such that :

$$\bar{\Delta} \leq M_3, \text{ for all } i.$$

Moreover, the same bound M_3 applies to the norm of the difference between the $(n-k)^{\text{th}}$ columns of H_i and G_{i-k} , $k = 0, \dots, n-1$.

The proposition follows from (i).

Proof of Lemma I.3.K

As $L_i \rightarrow \nabla_{xx} L(x^*, \lambda^*, \mu^*)$ and $A_i \rightarrow A^*$, we can consider QP_i^0 as a perturbation of the problem below:

$$\min_{p \in \mathbb{R}^n} \nabla f(x^*)p + \frac{1}{2}p^T [\nabla_{xx} L(x^*, \lambda^*, \mu^*) + A^*]p$$

subject to:

$$\begin{aligned} g^j(x^*) + g_x^j(x^*)p &\leq 0, \quad j \in \underline{m} \\ h^j(x^*) + h_x^j(x^*)p &= 0, \quad j \in \underline{\ell}. \end{aligned}$$

By H4, and as A^* is at least positive semi-definite, $(0, \lambda^*, u^*)$ is the unique Kuhn-Tucker triple for the above problem.

Moreover, at this point the second-order sufficiency conditions are satisfied.

Then, we can use Theorem 2.1 of Robinson (1974) to see that:

$$(i) \quad p_i \rightarrow 0, \quad \text{as } i \rightarrow \infty$$

and

$$(ii) \quad \lambda_i \rightarrow \lambda^*, \quad \text{as } i \rightarrow \infty.$$

Recall that $I_2(x^*) \triangleq \{j \in \underline{m} \mid g^j(x^*) < 0\}$.

From (i) we can say that, for i sufficiently large,

$$g^j(x_i) + g_x^j(x_i)p_i < 0 \quad \text{for all } j \in I_2(x^*).$$

That is, for $i \geq i_1$, say, none of the non-active constraints at x^* is predicted to be active.

The lemma then follows from (ii) and the strict complementarity slackness given by H4.

I.4 MODIFICATION OF BFGS UPDATE CASE

I.4.1 Introduction

Although the loop between Step 5 and Step 6 of the last algorithm is not executed for all i sufficiently large [see Proposition I.3.C] and a suitable choice of parameters makes the tests of Step 6 more difficult to satisfy [see discussion in I.3.5], the existence of this loop is an undesirable feature.

The tests are used to enforce global convergence and, as in the Basic Algorithm, are not required if the matrices $\{B_i\}$ are bounded and have positive eigenvalues bounded away from zero [see assumption H3, pg 23]. Furthermore, the use of $\{B_i\}$ satisfying H3 have also the advantages of making the solution of the quadratic sub-problem easier-many routines only perform well if the matrix is at least positive semidefinite - while providing a bound to $\|p_i\|$ [see inequality I.1.C.11].

In the unconstrained minimization of a function $f(x)$, the use of symmetric positive definite matrices for the determination of the search direction characterizes the 'variable metric methods, the most successful being [see Fletcher (1980)] the one which generates the sequence of matrices through the BFGS update below:

I.4.1

$$H_{i+1} = H_i - \frac{H_i \delta_i \delta_i^T H_i}{\delta_i^T H_i \delta_i} + \frac{\gamma_i \gamma_i^T}{\delta_i^T \gamma_i}$$

where:

$$\delta_i = x_{i+1} - x_i$$

and

$$\gamma_i = f_x(x_{i+1}) - f_x(x_i).$$

From the Kuhn-Tucker first order optimality conditions, a natural extension of the BFGS update to constrained problems would be obtained by setting :

I.4.II

$$\gamma_i = \nabla_x L(x_{i+1}, \tilde{\lambda}_i, \tilde{\mu}_i) - \nabla_x L(x_i, \tilde{\lambda}_i, \tilde{\mu}_i)$$

with:

$(\tilde{\lambda}_i, \tilde{\mu}_i)$ being estimates of the multipliers (λ^*, μ^*)

associated with x^* at the solution of GMP.

It is well known [see Fletcher and Powell (1963)] that: if H_i is positive definite and $\delta_i^T \gamma_i > 0$ then H_{i+1} will also be positive definite, when I.4.I is used. This property, which works well in unconstrained problems, is not so readily applied in the constrained case because I.4.II may lead to $\delta_i^T \gamma_i \leq 0$. Hence, Powell (1973) suggested that, instead of using γ_i in the update I.4.I, we use the vector η_i defined by :

I.4.III

$$\eta_i = \theta_i \gamma_i + (1 - \theta_i) H_i \delta_i$$

where:

$\theta_i \in (0, 1]$ is chosen so that:

$$\delta_i^T \eta_i \geq r' \delta_i^T H_i \delta_i, \text{ for some } r' > 0.$$

This definition and the subsequent use of η_i in I.4.I. has two consequences which, for reasons that will be made clearer later in this chapter and the next, we prefer to avoid.

Namely:

- The vector γ_i is multiplied by a coefficient which may be different from unity no matter how close x_i , is to x^* .
- The use of H_i in the definition of η_i makes the last term of I.4.1 also dependent on H_i .

Instead we define the vector η_i as :

I.4.IV

$$\eta_i = \gamma_i + A_i \delta_i$$

where,

A_i is a $(n \times n)$ positive definite matrix which does not depend on $\{H_i\}$, but is chosen to ensure,

$$\delta_i^T \eta_i \geq \bar{r} \|\delta_i\|^2, \text{ for some } \bar{r} > 0.$$

As we shall see, the definition of A_i is motivated by the second-order sufficiency conditions.

The matrix B_i to be used in QP_i^W is then computed as :

I.4.V

$$B_i = H_i + T(x_i)I.$$

In the next section we will present two algorithms.

The first one is a direct application of the above procedure to generate $\{B_i\}$ in the Basic Algorithm of the first chapter. With milder assumptions than those of Theorem I.1.M, - namely : the eigenvalues of $\{B_i\}$ are not assumed to remain bounded away from zero - any bounded infinite sequence generated by the algorithm will converge to $x^* \in D$. (see Theorem I.4.I)

The second algorithm is a slight modification of the previous one. A further test is performed before setting $\tilde{p}_i$. Although no result of the kind $x_i \rightarrow x^* \in D$ is proved to this algorithm, under reasonable conditions the algorithm will generate sequences with accumulation points in D . [see Theorem I.4.N] .

An interesting point to be observed in the proof of Theorem I.4.N is that it does not require H3 and, in this sense, it resembles the convergence result of Powell (1971) for the variable metric method for unconstrained optimization .

From the definition I.4.V given above and Proposition I.3.A, the sequence of matrices $\{B_i\}$ will have eigenvalues bounded away from zero, except possibly in subsequences such that :

$$x_i \rightarrow x^* \text{ (a Kuhn-Tucker point for GMP) .}$$

The assumption that $\{B_i\}$ is bounded can be replaced by the condition :

- The term $\frac{\eta_i \eta_i^T}{\delta_i \eta_i}$ in I.4.I is bounded from above.

That this condition is satisfied is established in Proposition I.4.L. The result is possible because of the choice of A_i . In general boundedness of this term is not guaranteed if it depends on H_i , unless, of course, we assume $\{H_i\}$ is bounded which is tantamount to assuming the boundedness of $\{B_i\}$ because $T(\cdot)$ is a continuous function and the sequence $\{x_i\}$ is assumed bounded.

The results on the rate of convergence are valid for both algorithms. Under the assumptions of section I.4.4, both algorithms generate sequences converging to $x^* \in D$ and, furthermore, for i sufficiently large they are equivalent.

In contrast to the secant update case, here $\{H_i\}$ is not expected to converge to $\nabla_{xx} L(x^*)$. Hence, we can not prove from I.4.V and Theorem I.3.L that unity of the step length will ultimately be achieved. Nevertheless, under assumptions mild enough to be satisfied whenever $\{H_i\}$ converges to $H^* \in L(R^n)$ as $x_i \rightarrow x^*$, a result similar to Theorem I.3.L is shown for constrained problem versions of the update I.4.I [see Theorem I.4.S]. Through this result we prove that the algorithm of page 126 will set the step length equal to unity for all i sufficiently large and a R-superlinear rate of convergence [see AP.1] is established.

I.4.2 The Algorithms

Consider the general mathematical programming problem (GMP), defined in the first chapter. The maps $\psi, \gamma, \hat{\psi}, \hat{\gamma}, \bar{\lambda}, \bar{\mu}, T, \bar{c}, \theta$ are as defined in section I.1.2. Let $(p_i, w_i, \lambda_i, t_i, s_i, v_i)$ denote the solution and multipliers of QP_i^W (see I.1.2) and let $\tilde{p}_i$ be defined as before. Let $L(\cdot, \cdot, \cdot)$ be the estimate of the Lagrange function defined by:

$$L(x, \bar{\lambda}(x), \bar{\mu}(x)) \triangleq f(x) + \sum_{j \in \underline{m}} \bar{\lambda}^j(x) g^j(x) + \sum_{j \in \underline{\ell}} \bar{\mu}^j(x) h^j(x) .$$

Also,

$$\nabla_x L(x, \bar{\lambda}(x), \bar{\mu}(x)) \triangleq \nabla f(x) + \sum_{j \in \underline{m}} \bar{\lambda}^j(x) \nabla_g^j(x) + \sum_{j \in \underline{\ell}} \bar{\mu}^j(x) \nabla_h^j(x)$$

and,

$$\begin{aligned} \nabla_{xx} L(x, \bar{\lambda}(x), \bar{\mu}(x)) \triangleq \nabla_{xx} f(x) + \sum_{j \in \underline{m}} \bar{\lambda}^j(x) \nabla_{xx} g^j(x) + \\ + \sum_{j \in \underline{\ell}} \bar{\mu}^j(x) \nabla_{xx} h^j(x) . \end{aligned}$$

Let $I'(\cdot)$ and $\tilde{I}(\cdot, \cdot)$ be the maps defined by :

$$I'(x) \triangleq \{j \in \underline{m} \mid \bar{\lambda}^j(x) > 0\}$$

and

$$\tilde{I}(x, p) \triangleq \{j \in \underline{m} \mid g^j(x) + g_x^j(x)p = 0\} .$$

Let $\tilde{N}(x, p)$ be the matrix whose columns are the gradients $g_x^j(x)$, $j \in I'(x) \cap \tilde{I}(x, p)$.

Let $N(x, p)$ be the matrix defined below :

$$N(x, p) \triangleq [\tilde{N}(x, p) \quad \nabla h^1(x) \quad \dots \quad \nabla h^\ell(x)] .$$

Let $\{H_i\}$ be the sequence of matrices constructed as follows:

I.4.VI

$$H_{i+1} = H_i - \frac{H_i \delta_i \delta_i^T H_i}{\delta_i^T H_i \delta_i} + \frac{\eta_i \eta_i^T}{\delta_i^T \eta_i}$$

where,

$$\delta_i = x_{i+1} - x_i$$

$$\eta_i = \gamma_i + a_i(T(x_i)\delta_i + N(x_i, p_i)\delta_i), \text{ where:}$$

$$\gamma_i \triangleq \nabla_x L(x_{i+1}, \bar{\lambda}(x_i), \bar{\mu}(x_i)) - \nabla_x L(x_i, \bar{\lambda}(x_i), \bar{\mu}(x_i))$$

and

$$a_i \triangleq \begin{cases} 0 & , \text{ if } \delta_i^T \gamma_i \geq \varepsilon_1 \|\delta_i\|^2, \varepsilon_1 \in (0,1) \\ 1 & , \text{ if } 0 \leq \delta_i^T \gamma_i < [\min\{T(x_i); \varepsilon_1\}] \|\delta_i\|^2 \\ 1 + \frac{(T(x_i) \|\delta_i\|^2 - \delta_i^T \gamma_i)}{(T(x_i) \|\delta_i\|^2 + \delta_i^T [N(x_i, p_i)]^T \delta_i)} & , \text{ otherwise} \end{cases}$$

Algorithm A

The algorithm is the Basic Algorithm (repeated below for convenience) with $\{B_i\}$ constructed as follows :

$$B_i = T(x_i)I + H_i$$

where

$\{H_i\}$ is generated by the update I.4.VI .

Data: $x_1 \in \mathbb{R}^n$; H_1 ($n \times n$ positive definite); $c_0 \in (0, \infty)$, $\delta \in (1, 2)$,
 $\beta, \varepsilon \in (0, 1)$.

Step : Set $i = 1$.

Step 1 : If $c_{i-1} \geq \bar{c}(x_i)$, set $c_i = c_{i-1}$.

Else, set $c_i = \max\{\delta c_{i-1}; \bar{c}(x_i)\}$.

Step 2 : Find (p_i, w_i) , the solution of QP_i^w .

Step 3 : If

$$T(x_i) \leq \varepsilon \quad \text{and} \quad w_i = 0.$$

Then,

compute $\tilde{p}_i$.

Else, set $\tilde{p}_i = 0$.

Step 4: Compute k_i , the first integer in $\{0, 1, \dots\}$ satisfying:

$$\gamma(x_i + \beta^{k_i} p_i + \beta^{2k_i} \tilde{p}_i, c_i) - \gamma(x_i, c_i) \leq \beta^{k_i} \theta(x_i, p_i, c_i) / 8 .$$

Step 5 : Set $x_{i+1} = x_i + \beta^{k_i} p_i + \beta^{2k_i} \tilde{p}_i$.

Step 6 : Update H_i to H_{i+1} by I.4.VI ; set $B_i = T(x_i)I + H_i$.

Step 7 : Set $i = i+1$; Go to Step 1.

Algorithm B

The algorithm is identical to Algorithm A with the addition

of one further test after Step 3. Namely:

Step 3' : If $\|B_i \tilde{p}_i\| > M \|p_i\|$, set $\tilde{p}_i = 0$.

Else,

proceed to Step 4.

$M < \infty$ is a positive constant which should be added to the data.

I.4.3 Global Convergence

Algorithm A

In the analysis of the global convergence properties of the first algorithm given in this chapter, we will assume that the two

following conditions, defined in I.1.3, are satisfied :

H1 ; H2.

In addition, we will require :

H3b: there exist $M < \infty$, positive, such that

$$\| y^T H_i y \| \leq M \| y \|^2 , \forall y \in R^n, \forall i$$

where $\{H_i\}$ is defined by I.4.VI.

Note that H3b is milder than H3 (defined in the first chapter) because the eigenvalues of $\{B_i\}$ are not assumed to be bounded away from zero.

Since Algorithm A is identical to the Basic Algorithm, with $\{B_i\}$ being generated by the method described in the previous section, if H3 were satisfied all the global convergence results of the first chapter would hold.

As it happens, the replacement of H3 by H3b will require some changes in the proofs although most of the propositions of I.1.3 (those which did not require the sequence $\{B_i\}$ to have eigenvalues bounded away from zero) clearly remain valid. Thus, Proposition I.1.A; Proposition I.1.B; Proposition I.1.C; Proposition I.1.D and Corollary; Lemma I.1.E; Proposition I.1.F; Lemma I.1.K are all valid.

We start by showing that the update given in the previous section is well defined and constructs a sequence $\{H_i\}$ of positive definite matrices.

Proposition I.4.A

For any finite i^*

$$p_{i^*} = 0 \Rightarrow x_{i^*} \in D.$$

Proof: Because of H1 and the way $\{B_i\}$ is constructed it is readily seen that, for any finite i^* , there exists a $M_{i^*} < \infty$ such that

$$p^T B_{i^*} p \leq M_{i^*} \|p\|^2, \text{ for all } p \in R^n.$$

Then,

$$(i) \quad \min_{p \in R^n} \{p^T B_{i^*} p + \theta(x_{i^*}, p, c_{i^*})\} \leq \min_{p \in R^n} \{M_{i^*} \|p\|^2 + \theta(x_{i^*}, p, c_{i^*})\}.$$

Now, as shown in the proof of Proposition I.1.C,

$$(ii) \quad QP_{i^*}^w \text{ is equivalent to : } \min_{p \in R^n} \{p^T B_{i^*} p + \theta(x_{i^*}, p, c_{i^*})\}.$$

But, $p_{i^*} = 0$ is the solution of $QP_{i^*}^w$.

Then, from (i) and (ii) we have :

$$\tilde{\theta}(x_{i^*}, M_{i^*}, c_{i^*}) \triangleq \min_{p \in R^n} \{M_{i^*} \|p\|^2 + \theta(x_{i^*}, p, c_{i^*})\} = 0.$$

Hence, from Proposition I.1.D we have :

$$\tilde{\theta}(x_{i^*}, 1, c_{i^*}) = 0 \Rightarrow x_{i^*} \in D_{c_{i^*}}.$$

But, by construction, $c_{i^*} \geq \bar{c}(x_{i^*})$.

Then, by the first part of Proposition I.1.F, which clearly remains valid, $x_{i^*} \in D$ and the proposition is proved.

Proposition I.4.B

For any finite i^* , there exists an integer $k' \in \{0, 1, \dots\}$ such that, at Step 7 of the algorithm,

$$\beta^{k_{i^*}} \geq \beta^{k'} > 0$$

Proof:

Given in the appendix at the end of the chapter.

Proposition I.4.C

For any finite i^* , H_{i^*} is positive definite.

Proof:

Let $K_1 \subset \{1, 2, \dots\}$ be the subset where $a_i = 0$. By construction,

$$(i) \quad \delta_i^T \eta_i \geq \varepsilon_1 \|\delta_i\|^2, \quad \forall i \in K_1.$$

Let $K_2 \subset \{1, 2, \dots\}$ be the subset where $a_i \geq 0$. By construction,

$$(ii) \quad \delta_i^T \eta_i \geq T(x_i) \|\delta_i\|^2, \quad \forall i \in K_2.$$

Now, from Proposition I.3.A, which clearly remains valid:

$$(iii) \quad T(x_i) = 0 \iff x_i \in D.$$

As we are interested in infinite sequences $\{x_i\}$, we can assume, without loss of generality that :

$$x_{i^*} \notin D.$$

Then, from (i), (ii), (iii), there exists a $r_1 > 0$ such that :

$$(iv) \quad \delta_{i*}^T \eta_{i*} \geq r_1 \|\delta_{i*}\|^2.$$

But, from Proposition I.4.A and Proposition I.4.B,

$$(v) \quad \|\delta_{i*}\| > 0.$$

From (iv) and (v), there exists $r_2 > 0$ such that

$$\delta_{i*}^T \eta_{i*} \geq r_2 > 0.$$

It is well known that, if H_1 is positive definite and $\delta_i^T \eta_i > 0$, H_{i+1} will be positive definite when computed through I.4.VI. The proposition is then proved because H_1 is positive definite.

Again let $\{A'(x_i, B_i, c_i)\}$ denote the sequence of maps which find the next point x_{i+1} with $\tilde{p}_i = 0$. That is:

$$A'(x_i, B_i, c_i) = x_{i+1} = \sum_{i=0}^k p_i + x_i.$$

Proposition I.4.D

Let $c > 0$ be arbitrary. If x^* is any accumulation point of a bounded infinite sequence $\{x_i\}$ where

$$x_{i+1} = A'_i(x_i, B_i, c) \text{ for all } i,$$

Then, $x^* \in D_c$.

Proof:

Let I be a sequence such that :

$$x_i \xrightarrow{I} x^*.$$

Because of H3b and the boundedness of $\{x_i\}$ it is clear that $\{B_i\}$ is bounded.

Hence, there exists at least one subsequence $I' \subset I$ such that:

$$x_i \xrightarrow{I'} x^* \quad \text{and} \quad B_i \xrightarrow{I'} B^*.$$

Note that, because of Proposition I.4.C, B^* is at least positive semi-definite.

Suppose $x^* \notin D_c$. From Proposition I.1.F, $x^* \notin D$. Then, $T(x^*) > 0$ and, by construction, there exists a $m' > 0$ such that :

$$(i) \quad p^T B_i p > m' \|p\|^2, \quad \forall p \in R^n, \quad \forall i \in I'$$

Also, if $x^* \notin D_c$, $\tilde{\theta}(x^*, 1, c) < 0$. Then, from the Corollary of Proposition I.1.D,

$$(ii) \quad \theta_1(x^*, B^*, c) < 0.$$

Now, from (i), B^* is positive definite. Hence, (ii) implies:

$$\theta(x^*, B^*, c) < 0.$$

The rest of the proof is identical to the proof of Proposition

I.1.G where it is shown that expression (ii) leads to a contradiction so that x^* indeed belongs to D_c .

Theorem I.4.E

Any accumulation point x^* of an infinite bounded sequence $\{x_i\}$ built by the algorithm satisfies

$$x^* \in D.$$

Proof:

Because of the last proposition, an examination of the proof of Proposition I.1.H and Proposition I.1.I show that they remain valid. The proof of the theorem is then identical to the proof of Theorem I.1.J.

As in the first chapter, we will now strengthen our hypothesis by assuming that the following conditions are satisfied:

H1' (replacing H1) ; H4.

Proposition I.4.F

Suppose $\{x_i\}$ is an infinite sequence constructed by the algorithm :

Let (x^*, λ^*, μ^*) be any K-T triple for GMP. Suppose $\{x_i\}_I$

is a subsequence such that:

$$x_i \xrightarrow{I} x^* \in D.$$

Then,

there exists an i_0 such that:

$$c_i > \sum_{j \in \underline{m}} \hat{\lambda}_i^j + \sum_{j \in \underline{\ell}} |\hat{\mu}_i^j|, \quad \forall i \in I, i \geq i_0.$$

where,

$(\hat{p}_i, \hat{\lambda}_i, \hat{\mu}_i)$ is the K-T triple for QP_i^0 .

Proof:

For simplicity of notation let : $(\bar{\lambda}_i, \bar{\mu}_i) \triangleq (\bar{\lambda}(x_i), \bar{\mu}(x_i))$.

By construction, $c_i \geq \sum_{j \in \underline{m}} |\bar{\lambda}_i^j| + \sum_{j \in \underline{\ell}} |\bar{\mu}_i^j| + b$.

We already know that :

$$(\bar{\lambda}_i, \bar{\mu}_i) \rightarrow (\lambda^*, \mu^*) \text{ as } x_i \rightarrow x^*.$$

Hence,

(i) there exists an i_1 such that

$$c_i \geq \sum_{j \in \underline{m}} (\lambda^*)^j + \sum_{j \in \underline{\ell}} |(\mu^*)^j| + b/2$$

for all $i \in I, i \geq i_1$.

Now, let $(p_i^!, \lambda_i^!, \mu_i^!)$ be the K-T point of the sub-problem P_i^* below :

$$\begin{aligned} P_i^* : & \min_{p \in \mathbb{R}^n} f_x(x^*)p + 1/2 p^T B_i p \\ \text{s.t.} \quad & g_x^j(x^*) + g_x^j(x^*)p \leq 0, j \in \underline{m} \\ & |h_x^j(x^*) + h_x^j(x^*)p| \leq 0, j \in \underline{\ell} . \end{aligned}$$

As $x^* \in D$ and B_i is positive definite we have :

(ii) $p_i^! = 0$, for all i .

Then, by H2, we have :

$$(iii) \quad (\lambda_i^!, \mu_i^!) = (\lambda^*, \mu^*), \text{ for all } i.$$

Clearly, for all i the sub-problem P_i^* itself satisfies the second-order sufficiency conditions with strict complementarity slackness.

Moreover as $B_i p^! = 0$ for all i , the map f_i^* defined below, does not depend on B_i .

$$f_i^* \triangleq [\nabla L_i^!, (\lambda_i^!)^T [g_i^!(x^*)], (\mu_i^!)^T [h_i^!(x^*)]]$$

where,

$$\nabla L_i^! \triangleq f(z^*) + B_i p_i^! + \sum_{j \in \underline{m}} (\lambda_i^!)^j g_x^j(x^*) + \sum_{j \in \underline{\ell}} (\mu_i^!)^j h_x^j(x^*)$$

$$g_i^!(x^*)^T = [g^1(x^*) + g_x^1(x^*) p_i^!, \dots, g^m(x^*) + g_x^m(x^*) p_i^!]$$

and

$$h_i^!(x^*)^T = [h^1(x^*) + h_x^1(x^*) p_i^!, \dots, h^\ell(x^*) + h_x^\ell(x^*) p_i^!].$$

Then, considering (for all i) sub-problem QP_i^0 as a perturbation of P_i^* , we can apply Theorem 2.1 of Robinson (1974) to see that :

(iv) There exists an open neighbourhood N_0 of x^* such that, for all $x_i \in N_0$, $(\hat{p}_i, \hat{\lambda}_i, \hat{\mu}_i)$ is the unique K-T triple of QP_i^0 , satisfies the second order sufficiency conditions with strict complementarity slackness and,

$$(\hat{p}_i, \hat{\lambda}_i, \hat{\mu}_i) \rightarrow (p_i^!, \lambda_i^!, \mu_i^!).$$

From (ii), (iii), and (iv) we have :

$$(\hat{p}_i, \hat{\lambda}_i, \hat{\mu}_i) \rightarrow (0, \lambda^*, \mu^*).$$

Then there exists an i_2 such that, for all $i \in I$, $i \geq i_2$

we have:

$$(v) \quad \left| \left(\sum_{j \in \underline{m}} \hat{\lambda}_i^j + \sum_{j \in \underline{\ell}} |\hat{\mu}_i^j| \right) - \left(\sum_{j \in \underline{m}} (\lambda^*)^j + \sum_{j \in \underline{\ell}} |(\mu^*)^j| \right) \right| \leq b/4.$$

Hence, from (i) and (v), for all $i \in I$ sufficiently large

$$c_i \geq \sum_{j \in \underline{m}} \hat{\lambda}_i^j + \sum_{j \in \underline{\ell}} |\hat{\mu}_i^j| + b/4$$

and the proposition is proved.

Proposition I.4.G

Suppose $\{x_i\}$ is an infinite sequence constructed by the algorithm.

Let (x^*, λ^*, μ^*) be any K-T triple for GMP. Suppose $\{x_i\}_{i \in I}$ is a sub-sequence such that :

$$x_i \xrightarrow{I} x^* \in D.$$

Then,

there exists an i' such that, for all $i \in I$, $i \geq i'$

$$p_i = \hat{p}_i; w_i = 0; \lambda_i = \hat{\lambda}_i; (t_i - s_i) = \hat{\mu}_i.$$

Proof:

By hypothesis, $x_i \xrightarrow{I} x^* \in D$.

Then, by the continuity properties of g , g_x , h and h_x , we can say that, for all $i \in I$, i sufficiently large, sub-problem QP_i^0 is feasible.

Now, as B_i is positive definite, Lemma I.2.A remains valid.

Hence, the proposition follows from Proposition I.4.F and the fact, shown in the proof of this proposition that, for all $i \in I$, $i \geq i_0$, say, $(\hat{p}_i, \hat{\lambda}_i, \hat{\mu}_i)$ is the unique K-T triple for QP_i^0 .

Proposition I.4.H

Let $\{x_i\}$ be an infinite sequence constructed by the algorithm.

Suppose $\{x_i\}_I$ is an infinite subsequence such that

$$c_i = c > 0, \quad \text{for all } i \in I$$

and

$$x_i \xrightarrow{I} \bar{x} \in \mathbb{R}^n.$$

Then,

$$p_i \xrightarrow{I} 0.$$

Proof:

If $\bar{x} \in D$, from Proposition I.4.G

$$p_i = \hat{p}_i, \quad \text{for all } i \in I, \text{ } i \text{ sufficiently large.}$$

But, as shown in the proof of Proposition I.4.F,

$$\hat{p}_i \xrightarrow{I} 0.$$

Then, if $\bar{x} \in D$ the proposition is proved.

Otherwise, suppose $\bar{x} \notin D$.

In this case, there exists a $m^* > 0$ such that :

$$T(x_i) \geq m^*, \quad \text{for all } i \in I.$$

Hence, from Proposition I.1.A and the way B_i is defined,

$$\theta(x_i, p_i, c) \leq -p_i^T B_i p_i \leq -m^* \|p_i\|^2$$

$$\text{for all } i \in I.$$

The rest of the proof of the proposition is identical to the proof of Proposition I.1.L.

Theorem I.4.I

Let x_i be an infinite bounded sequence constructed by the algorithm.

Then,

$$x_i \rightarrow x^* \in D.$$

Proof:

Because of Theorem I.4.E, Proposition I.4.I and assumption H4, the proof of this theorem is identical to the proof of Theorem I.1.M.

Algorithm B

In the analysis of the global convergence properties of the second algorithm given in this chapter, we will assume that the three following conditions, defined in the first chapter, are satisfied:

$$H1' ; H2 ; H4$$

Note that assumptions H3 or H3b are not required. Actually, no assumption are made about the sequence $\{B_i\}$. This fact make the previous approaches, to show the global convergence of the Basic Algorithm (see I.1.2) and Algorithm A, inadequate. Nevertheless, a careful analysis shows that some previous results, those which did not depend on H3, remain true. Namely: Proposition I.1.A ; Proposition I.1.B; Proposition I.1.F; Lemma I.2.A ; Propostion I.3.A ; Proposition I.4.A ; Proposition I.4.B ; Proposition I.4.C ; Proposition I.4.F ; Proposition I.4.G ; Proposition I.4.H.

Proposition I.4.J

Let $\{x_i\}$ be an infinite sequence constructed by the algorithm .

Then,

there exists a constant $r' > 0$

such that :

$$\delta_i^T \eta_i \geq r' \|\delta_i\|^2 \text{ for all } i.$$

Proof:

Given in the appendix at the end of the chapter.

Proposition I.4.K

Let $\{x_i\}$ be a bounded sequence constructed by the algorithm.

Then,

there exists a $L' < \infty$ such that :

$$a_i \leq L' , \text{ for all } i.$$

Proof:

Given in the appendix at the end of the chapter.

Proposition I.4.L

Let $\{x_i\}$ be a bounded infinite sequence constructed by the algorithm.

Then,

there exists a $\tilde{K} < \infty$ such that :

$$\frac{\|\eta_i\|^2}{\alpha_i^T \eta_i} \leq \tilde{K} , \text{ for all } i.$$

Proof:

Given in the appendix at the end of the chapter.

Proposition I.4.M

Let $\{x_i\}$ be a bounded sequence constructed by the algorithm.

Let $\{x_i\}_I$ be any infinite subsequence such that:

$$c_i = c > 0, \text{ for all } i \in I$$

and

$$x_i \xrightarrow{I} \bar{x} \notin D.$$

Then,

there exists an integer $k' \in \{0, 1, \dots\}$

such that :

$$\beta_i^{k'} \geq \beta^{k'} > 0, \text{ for all } i \in I.$$

Proof:

Given in the appendix at the end of the chapter.

Theorem I.4.N

Any accumulation point x^* of a bounded infinite sequence constructed by the algorithm, under the further assumptions that :

there exists a compact neighbourhood

N_1 of x^* where,

QP_i^0 is feasible for all $x_i \in N_1$

and

$$c_i \geq \sum_{j \in \underline{m}} \hat{\lambda}_i^j + \sum_{j \in \underline{\ell}} |\hat{\mu}_i^j| ,$$

satisfies :

$$x^* \in D.$$

Proof:

Let $I \subset \{1, 2, \dots\}$ be such that :

$$x_i \xrightarrow{I} x^*.$$

Hence, for i sufficiently large, $i \in I$, $x_i \in N_1$.

Because of the assumptions on N_1 we can see from Lemma I.2.A

that,

for all $i \in I$, i sufficiently large,

$$QP_i^w \text{ is equivalent to } QP_i^0.$$

Hence, for all $i \in I$, i sufficiently large, the following constraints are satisfied :

$$(i) \quad g^j(x_i) + g_x^j(x_i)p_i \leq 0, \quad j \in \underline{m}$$

and

$$h^j(x_i) + h_x^j(x_i)p_i = 0, \quad j \in \underline{\ell}.$$

But, from Proposition I.4.H,

$$(ii) \quad p_i \xrightarrow{I} 0,$$

Then, by the continuity of the scalar product and $H1'$, we have from (i) and (ii)

$$\begin{aligned} g^j(x_i) &\xrightarrow{I} g^j(x^*) \leq 0, \quad \text{for all } j \in \underline{m} \\ \text{and} \\ h^j(x_i) &\xrightarrow{I} h^j(x^*) = 0, \quad \text{for all } j \in \underline{\ell} \end{aligned}$$

That is :

$$(iii) \quad x^* \text{ is a feasible point for GMP.}$$

Now, consider the trace of the matrix H_{i+1}

$$\text{tr}(H_{i+1}) = \text{tr}(H_i) - \frac{\|H_i \delta_i\|^2}{\delta_i^T H_i \delta_i} + \frac{\|\eta_i\|^2}{\delta_i^T \eta_i}.$$

Hence, for any $j \in \{0, 1, \dots\}$ we have :

$$(iv) \quad \text{tr}(H_{j+1}) = \text{tr}(H_1) - \sum_{i=1}^j \frac{\|H_i \delta_i\|^2}{\delta_i^T H_i \delta_i} + \sum_{i=1}^j \frac{\|\eta_i\|^2}{\delta_i^T \eta_i}.$$

Now, as H_{j+1} is positive definite for any j , the right hand side of (iv) is positive. Then,

$$(v) \quad \sum_{i=1}^j \frac{\|H_i \delta_i\|^2}{\delta_i^T H_i \delta_i} \leq \text{tr}(H_1) + \sum_{i=1}^j \frac{\|\eta_i\|^2}{\delta_i^T \eta_i}.$$

From (v) and Proposition I.4.L, there exists a number $M_2 < \infty$ and independent of j such that :

$$\sum_{i=1}^j \frac{\|H_i \delta_i\|^2}{\delta_i^T H_i \delta_i} < j M_2, \quad \text{for any } j.$$

Hence,

$$(vi) \quad \sum_{\substack{i \in I \\ i \leq j}} \frac{\|H_i \delta_i\|^2}{\delta_i^T H_i \delta_i} < j M_2, \quad \text{for any } j.$$

Now, as $\{x_i\}$ is assumed bounded, we can apply Theorem 1 in Mayne and Polak (1978) to see that, c_{i-1} is increased only finitely often.

That is, for i sufficiently large we have :

$$c_i = c > 0.$$

Then, for $i \in I$, sufficiently large,

$$(vii) \quad c \geq \sum_{j \in \underline{m}} \lambda_i^j + \sum_{i \in \underline{\ell}} |\mu_i^j|, \quad \text{where } \mu_i \triangleq t_i - s_i.$$

Now, at the solution of QP_i^w , the following equation is satisfied

$$(viii) \quad \nabla f(x_i) + B_i p_i + \sum_{j \in \underline{m}} \lambda_i^j \nabla g^j(x_i) + \sum_{j \in \underline{\ell}} \mu_i^j \nabla h^j(x_i) = 0.$$

But, from (vii), the multipliers λ_i^j and μ_i^j are bounded.

Hence, from H1' and the boundedness of $\{x_i\}$, we have a $M_3 < \infty$ such that:

$$(ix) \quad \left\| \nabla f(x_i) + \sum_{j \in \underline{m}} \lambda_i^j \nabla g^j(x_i) + \sum_{j \in \underline{\ell}} \mu_i^j \nabla h^j(x_i) \right\| \leq M_3$$

for all $i \in I$.

Then, from (ix) and (viii),

$$(x) \quad \| B_i p_i \| \leq M_3, \quad \text{for all } i \in I.$$

Now, from (ii) and the fact that $\| \tilde{p}_i \| \leq \| p_i \|$ by construction, we have :

$$(xi) \quad \delta_i \xrightarrow{I} 0.$$

Also,

$$(xii) \quad \| B_i \delta_i \| \leq \| B_i p_i \| + \| B_i \tilde{p}_i \|, \quad \text{for all } i.$$

But, from the test at Step 3',

$$(xiii) \quad \| B_i \tilde{p}_i \| \leq M \| p_i \|, \quad \text{for all } i.$$

Then, from (xiii), (xii), (x) and (ii), we have a $M_4 < \infty$ such that

$$(xiv) \quad \| B_i \delta_i \| \leq M_4, \quad \text{for all } i \in I.$$

From (xiv) and (xi), we have:

$$\delta_i^T B_i \delta_i \xrightarrow{I} 0.$$

But,

$$\delta_i^T H_i \delta_i \leq \delta_i^T B_i \delta_i, \quad \text{for all } i.$$

Hence,

$$(xv) \quad \delta_i^T H_i \delta_i \xrightarrow{I} 0.$$

The only way of (xv) and (vi) being simultaneously satisfied is:

$$(xvi) \quad H_i \delta_i \xrightarrow{I} 0.$$

Now,

$$(xvii) \quad H_i \delta_i = \beta^{k_i} H_i p_i + \beta^{2k_i} H_i \tilde{p}_i$$

From the definition of B_i , the continuity of $T(\cdot)$ and the boundedness of $\{x_i\}$, there exists a $M_5 < \infty$ such that :

$$\|H_i \tilde{p}_i\| \leq \|B_i \tilde{p}_i\| + M_5 \|\tilde{p}_i\| \quad \text{for all } i.$$

Hence, from the test at Step 3', and the fact that $\|\tilde{p}_i\| \leq \|p_i\|$, we have :

$$(xviii) \quad \|H_i \tilde{p}_i\| \leq (M + M_5) \|p_i\|.$$

From (xviii) and (ii) we have :

$$(xix) \quad H_i \tilde{p}_i \xrightarrow{I} 0.$$

Suppose $x^* \notin D$.

Then, from Proposition I.4.M, there exists k' such that:

$$(xx) \quad \beta^{k_i} \geq \beta^{k'} > 0, \quad \text{for all } i \in I.$$

From (xx), (xix) and (xvi) into (xvii), we have :

$$(xxi) \quad H_i p_i \xrightarrow{I} 0.$$

But,

$$(xxii) \quad \|B_i p_i\| \leq \|H_i p_i\| + M_5 \|p_i\| \text{ for some } M_5 < \infty.$$

Then, from (xxii), (xxi) and (ii),

$$(xxiii) \quad B_i p_i \xrightarrow{I} 0.$$

From (xxiii) and (viii),

$$(xxiv) \quad \nabla f(x_i) + \sum_{j \in \underline{m}} \lambda_i^j \nabla g^j(x_i) + \sum_{j \in \underline{\ell}} \mu_i^j \nabla h^j(x_i) \xrightarrow{I} 0.$$

But, as the multipliers are bounded, there exists at least one subsequence $I_2 \subset I$ such that :

$$(xxv) \quad \lambda_i \xrightarrow{I_2} \lambda^* ; \quad \mu_i \xrightarrow{I_2} \mu^*$$

where

$$(\lambda^*, \mu^*) \in \mathbb{R}^m \times \mathbb{R}^\ell.$$

Clearly,

$$x_i \xrightarrow{I_2} x^*.$$

Then, from (xxiv) and (xxv),

$$(xxvi) \quad \nabla f(x^*) + \sum_{j \in \underline{m}} (\lambda^*)^j \nabla g^j(x^*) + \sum_{j \in \underline{\ell}} (\mu^*)^j \nabla h^j(x^*) = 0.$$

Recall that, from (iii),

$$(xxvii) \quad \psi(x^*) = 0.$$

Let I^* denote the set : $I^* \triangleq \{j \in m \mid g^j(x^*) < 0\}$.

From (ii), the convergence $x_i \xrightarrow{I} x^*$ and the complementarity condition at the solution of QP_i^W , we have :

$$(xxviii) \quad \lambda_i^j = 0, \quad j \in I^*$$

for all $i \in I$, i sufficiently large.

From (xxviii) and (xxv),

$$(xxix) \quad (\lambda^*)^j = 0, \quad j \in I^*.$$

Then, from (xxvi), (xxvii), (xxix), the definition of $(\bar{\lambda}(x), \bar{\mu}(x))$ and H2, we have :

$$(\bar{\lambda}(x^*), \bar{\mu}(x^*)) = (\lambda^*, \mu^*).$$

Hence, from (xxvi) and (xxvii) we have :

$$T(x^*) = 0.$$

The theorem then follows from Proposition I.3.A.

I.4.4 Rate of Convergence

Throughout this section we will assume that the following conditions, defined in the first chapter, are satisfied.

$$H1' ; H2 ; H3 ; H4.$$

Note that because of H3, both Algorithm A and Algorithm B can be shown to satisfy Theorem I.1.M that is ,

if $\{x_i\}$ is a bounded infinite sequence constructed by the algorithm, then,

$$x_i \rightarrow x^* \in D.$$

As we shall show (see Proposition I.4.Q),

$$\|\tilde{p}_i\| \approx \|p_i\|^2.$$

Then, by the boundedness of $\{B_i\}$ (ensured by H3) and the convergence of $p_i \rightarrow 0$, for i large enough the test of Step 3' in Algorithm B will be always satisfied. That is,

for all i sufficiently large Algorithm A and Algorithm B are equivalent.

Hence, the results of this section are valid for both algorithms.

In addition to the four conditions above, we introduce a new condition which is assumed to be satisfied throughout this section.

Namely:

$$H5: \quad \|H_i \delta_i - \eta_i\| \rightarrow 0, \text{ as } i \rightarrow \infty.$$

Now, H5 can be better understood if we note that :

$$H_{i+1} \delta_i = \eta_i, \text{ for all } i.$$

Therefore, a sufficient condition to H5 being satisfied is :

$$H_i \rightarrow H^* \in L(R^n), \text{ as } i \rightarrow \infty$$

Clearly, such H^* would be at least, positive semi-definite.

So, H3 does not seem to be excessively strong.

Proposition I.4.0

Let $\{x_i\}$ be a bounded infinite sequence constructed by the algorithm. Let $x^* \in D$ be such that :

$$x_i \rightarrow x^* \text{ as } i \rightarrow \infty.$$

Then

there exists an i^* such that, for all $i \geq i^*$, QP_i^W does predict correctly the constraints to GMP^1 which are active at x^* .

Proof:

We already know that, as $\{x_i\}$ is bounded,

$$c_i = c > 0, \text{ for all } i \text{ sufficiently large.}$$

Now, as shown in the proof of Proposition I.4.F,

$$\hat{\lambda}_i \rightarrow \lambda^* ; \hat{p}_i \rightarrow 0, \text{ as } i \rightarrow \infty$$

But, from Proposition I.4.G, for i sufficiently large,

$$\lambda_i = \hat{\lambda}_i ; p_i = \hat{p}_i ; w_i = 0.$$

Hence,

$$\lambda_i \rightarrow \lambda^* ; p_i \rightarrow 0, \text{ as } i \rightarrow \infty.$$

From the strict complementarity slackness assured by H4, we can then see that :

there exists an i_1 such that :

$$\lambda_i^j > 0, \text{ for all } j \text{ s.t. } g^j(x^*) = 0, \text{ } i \geq i_1.$$

Now, as $w_i = 0$ for i sufficiently large, the last ℓ constraints are ultimately active at the solution of QP_i^W .

Then, all the active constraints at x^* are active at the solution of QP_i^W , for all i sufficiently large.

But, as $x_i \rightarrow x^* \in D$ and $p_i \rightarrow 0$, we can see that, for i sufficiently large,

$$g^j(x_i) + g_x^j(x_i)p_i < 0, \quad \text{for all } j \in I_2(x^*)$$

where,

$$I_2(x^*) \triangleq \{j \in \underline{m} \mid g^j(x^*) < 0\}.$$

Hence, for i sufficiently large, none of the non-active constraints of GMP will be active at the solution of QP_i^W and the proposition is proved.

Proposition I.4.P

There exists an integer $k' \in \{0, 1, \dots\}$ such that,

$$\beta_i^{k_i} \geq \beta^{k'} > 0, \quad \text{for all } i.$$

Proof:

Note in the proof of Proposition I.4.M that the hypothesis that $x \notin D$ was only necessary to make sure that $T(x_i)$ remains bounded away from zero, so that :

$$-p_i^T B_i p_i < -m \|p_i\|^2 \quad \text{for some } m > 0.$$

This condition is now assured by H3. The proof of this proposition is then identical to the proof of Proposition I.4.M.

Proposition I.4.Q

$$\| \tilde{p}_i \| \approx \| p_i \|^2$$

Proof:

Given in the appendix at the end of the chapter.

Proposition I.4.R

$$\gamma(x_i + p_i + \tilde{p}_i, c_i) - \gamma(x_i, c_i) = \phi(x_i, c_i) + \bar{e}_i$$

where,

$$\phi(x_i, c_i) = \theta(x_i, p_i, c_i) + \langle \nabla f(x_i, \tilde{p}_i) \rangle + \frac{1}{2} \langle p_i, \nabla_{xx} f(x_i) p_i \rangle$$

and

$$\bar{e}_i \approx \| p_i \|^3$$

Proof:

Given in the appendix at the end of the chapter.

Theorem I.4.S

Consider the Basic Algorithm (see I.1.2) with $\{B_i\}$ being updated through I.4.VI with η_i given by :

$$\eta_i = L_i \delta_i + \bar{A}_i \delta_i, \text{ where}$$

$L_i \in L(R^n)$ satisfies : $L_i \rightarrow \nabla_{xx} L(x^*, \lambda^*, u^*)$ as $x_i \rightarrow x^* \in D$ and $\{\bar{A}_i\}$ is a sequence of positive-definite matrices.

If H1', H2, H3, H4, H5 are satisfied, then the step length obtained through the search arc of Step 4 will be unity for all i sufficiently large.

Proof:

First note that, because H1, H2, H3 and H4 are assumed to be satisfied, Theorem I.1.M is valid so that :

$$x_i \rightarrow x^* \in D.$$

Then, for i sufficiently large, $T(x_i) < \varepsilon$.

As shown in Proposition I.4.G, $w_i = 0$ for all i sufficiently large. Also, from Proposition I.4.Q and the fact that $p_i \rightarrow 0$, for all i sufficiently large, $\|\tilde{p}_i\| \leq \|p_i\|$. Hence, for all i sufficiently large $\tilde{p}_i$ is not set equal to zero.

Now, the step length is unity if :

$$T_i \triangleq \gamma(x_i + p_i + \tilde{p}_i, c_i) - \gamma(x_i, c_i) - \theta(x_i, p_i, c_i)/8 \leq 0.$$

Because of the last two propositions, we can follow the same reasoning of the proof of Theorem I.3.L to see that :

$$(i) \quad T_i \leq (3/8)\theta(x_i, p_i, c_i) + \frac{1}{2} \langle p_i, (\nabla_{xx} L(x_i, \lambda_i, \mu_i) - B_i)p_i \rangle + o(\|p_i\|^2).$$

By hypothesis, condition H5 is satisfied. Then,

$$(ii) \quad \langle \delta_i, (\nabla_{xx} L(x_i, \lambda_i, \mu_i) - H_i)\delta_i \rangle \leq \langle \delta_i, (\nabla_{xx} L(x_i, \lambda_i, \mu_i)\delta_i - \eta_i) \rangle + o(\|\delta_i\|^2).$$

But, by definition,

$$(iii) \quad \eta_i = L_i \delta_i + \bar{A}_i \delta_i.$$

As shown by Proposition I.4.F and Proposition I.4.G, $(x_i, \lambda_i, \mu_i) \rightarrow (x^*, \lambda^*, \mu^*)$ as $i \rightarrow \infty$. Then,

$$(iv) \quad \nabla_{xx} L(x_i, \lambda_i, \mu_i) \rightarrow \nabla_{xx} L(x^*, \lambda^*, \mu^*) \text{ as } i \rightarrow \infty.$$

Also by hypothesis,

$$(v) \quad L_i \rightarrow \nabla_{xx} L(x^*, \lambda^*, \mu^*) \text{ as } i \rightarrow \infty.$$

From (iii), (iv), (v) into (ii), we have :

$$\langle \delta_i, (\nabla_{xx} L(x_i, \lambda_i, \mu_i) - H_i) \delta_i \rangle \leq - \langle \delta_i, \bar{A}_i \delta_i \rangle + o(\|\delta_i\|^2).$$

But, by hypothesis, $\bar{A}_i$ is positive definite. Then,

$$(vi) \quad \langle \delta_i, (\nabla_{xx} L(x_i, \lambda_i, \mu_i) - H_i) \delta_i \rangle \leq o(\|\delta_i\|^2).$$

$$\text{Now, } \delta_i = \beta^{k_i} p_i + \beta^{2k_i} \tilde{p}_i.$$

Then,

$$\begin{aligned} (vii) \quad & \langle \beta^{k_i} p_i, (\nabla_{xx} L(x_i, \lambda_i, \mu_i) - H_i) \beta^{k_i} p_i \rangle = \langle \delta_i, (\nabla_{xx} L(x_i, \lambda_i, \mu_i) - H_i) \delta_i \rangle \\ & - 2 \langle \delta_i, (\nabla_{xx} L(x_i, \lambda_i, \mu_i) - H_i) \beta^{2k_i} \tilde{p}_i \rangle \\ & + \langle \beta^{2k_i} \tilde{p}_i, (\nabla_{xx} L(x_i, \lambda_i, \mu_i) - H_i) \beta^{2k_i} \tilde{p}_i \rangle. \end{aligned}$$

From Proposition I.4.P, $\beta^{k_i} \geq \beta^{k_{i'}} > 0$, for all i . Hence, dividing both sides of (vii) by β^{2k_i} , taking into account (vi) and Proposition I.4.Q, we have :

$$(viii) \quad \langle p_i, (\nabla_{xx} L(x_i, \lambda_i, \mu_i) - H_i) p_i \rangle \leq o(\|p_i\|^2)$$

Recall that $B_i = H_i + T(x_i)I$. By the continuity of $T(\cdot)$ and Proposition I.3.A, we have : $T(x_i) \rightarrow 0$ as $i \rightarrow \infty$. Then,

$$(ix) \quad \langle p_i, (\nabla_{xx} L(x_i, \lambda_i, \mu_i) - B_i) p_i \rangle \leq \langle p_i, (\nabla_{xx} (Lx_i, \lambda_i, \mu_i) - H_i) p_i \rangle \\ + o(\| \delta_i \|^2) .$$

From Proposition I.1.A and H3, there exists a $r_1 > 0$ such that :

$$(x) \quad 3/8 \theta(x_i, p_i, c_i) \leq -r_1 \| p_i \|^2 , \text{ for all } i.$$

From (x), (ix) , (viii) and (i), we obtain:

$$T_i \leq -r_1 \| p_i \|^2 + o(\| p_i \|^2).$$

The theorem then follows from the convergence

$$\| p_i \| \rightarrow 0 \quad \text{as} \quad i \rightarrow \infty.$$

Proposition I.4.T

For i sufficiently large the algorithms of this chapter will compute a step length of unity.

Proof:

We already know that $x_i \rightarrow x^* \in D$ as $i \rightarrow \infty$.

Now, by construction, $\eta_i = \gamma_i + a_i A_i \delta_i$ where :

$$A_i \triangleq T(x_i)I + [N(x_i, p_i)][N(x_i, p_i)]^T.$$

By definition, $\gamma_i = \nabla_x L(x_{i+1}, \bar{\lambda}(x_i), \bar{\mu}(x_i)) - \nabla_x L(x_i, \bar{\lambda}(x_i), \bar{\mu}(x_i))$

then,

$$\gamma_i = G_i \delta_i, \text{ for all } i$$

where,

$$G_i \triangleq \int_{t=0}^1 \nabla_{xx} L(x_i + t\delta_i, \bar{\lambda}(x_i), \bar{\mu}(x_i)) dt.$$

But, $\delta_i \rightarrow 0$ and $(\bar{\lambda}(x_i), \bar{\mu}(x_i)) \rightarrow (\lambda^*, \mu^*)$

as $x_i \rightarrow x^* \in -D$.

Then it is clear that

$$G_i \rightarrow \nabla_{xx} L(x^*, \lambda^*, \mu^*) \text{ as } i \rightarrow \infty.$$

But, the matrix $a_i A_i$ is positive definite for all i . The proposition then follows from Theorem I.4.S.

Proposition I.4.U

There exists a $v_3 > 0$

such that : $\|p_{i+1}\| \leq v_3 \|p_i\|$

for all i sufficiently large.

Proof:

Given in the appendix at the end of the chapter.

Theorem I.4.V

Let $\{x_i\}$ be a bounded infinite sequence constructed by the algorithms.

Then,

$$x_i \rightarrow x^* \in D$$

with a rate which is, at least, R-superlinear.

Proof:

We already know that $x_i \rightarrow x^* \in D$, from Theorem I.1.M.

Let $\tilde{K}_1 \subset \{1, 2, \dots\}$ be the set of iterations where

(i) there exists a $r_1 > 0$ such that :

$$\|p_i\| \geq r_1 \|p_{i-1}\|^2, \text{ for all } i \in \tilde{K}_1.$$

Let $\tilde{K}_2$ be the complement of $\tilde{K}_1$.

As no $r_1 > 0$ satisfying (i) can be found for $\tilde{K}_2$, we can say that

$$(ii) \quad \frac{\|p_i\|}{\|p_{i-1}\|^2} \xrightarrow{I} 0 \text{ as } i \rightarrow \infty$$

for any infinite subsequence $I \subset \tilde{K}_2$.

As shown in the proof of Proposition I.4.U,

$$(iii) \quad \|p_i\| \sim \|x_i - x^*\|.$$

Also,

$$(iv) \quad \|p_{i+1}\| \leq v_3 \|p_i\|, \text{ for all } i \text{ sufficiently large.}$$

From (ii), (iii) and (iv) we have :

$$(v) \quad \frac{\|x_{i+1} - x^*\|}{\|x_i - x^*\|^2} \xrightarrow{I} 0, \text{ as } i \rightarrow \infty$$

for any infinite subsequence $I \subset \tilde{K}_2$.

Let $x_i^!$ denote $x_i + p_i$

Let z_i denote $(x_i, \lambda_{i-1}, u_{i-1})$

Let $z_i^!$ denote $(x_i^!, \lambda_i, u_i)$

We already know that : $z_i \rightarrow z^* \triangleq (x^*, \lambda^*, u^*)$.

Also from Proposition I.4.Q, for $i \geq i_0$, say, $\tilde{p}_i$ is not set equal to zero. That is, $\|\tilde{p}_i\| \leq \|p_i\|$ and $\|B_i \tilde{p}_i\| \leq M \|p_i\|$, for all i sufficiently large.

Then, for all i sufficiently large, we have :

$$x_{i+1} = x_i^! + \tilde{p}_i.$$

Now, for all z let $q(z)$ be defined by :

$$q(z) \triangleq \begin{pmatrix} \nabla L(z) \\ \lambda_i^1 g^1(z) \\ \vdots \\ \lambda_i^m g^m(z) \\ h_i^1(z) \\ \vdots \\ h_i^\ell(z) \end{pmatrix}$$

Under our assumptions [see pg 7 of Garcia and Mangasarian (1976)], we have M_1, M_2 such that :

$$(vi) \quad \|q(z_i^!)\| \leq M_1 \|z_i^! - z_i\|^2 + \|(\nabla_{xx} L(z_i) - B_i)(x_i^! - x_i)\|$$

and [see Lemma 3.7 in Garcia and Mangasarian (1976)]

$$(vii) \quad \|z'_{i+1} - z_{i+1}\| \leq M_2 \|q(z_{i+1})\|.$$

Now, by H5, Proposition I.4.T and Proposition I.4.Q we can show, as in the proof of Theorem I.4.S, that :

$$(viii) \quad \|(\nabla_{xx} L(z_i) - H_i)(x'_i - x_i)\| \leq \|(\nabla_{xx} L(z_i)(x'_i - x_i) - Q_i(x'_i - x_i)\| + o(\|x'_i - x_i\|).$$

where

$$Q_i \triangleq G_i + a_i(T(x_i)I + [N(x_i, p_i)] [N(x_i, p_i)]^T)$$

But, as shown by Mayne and Polak (1978),

$$p_i = r_i + q_i$$

where,

r_i lies on the null space of $[N(x_i, p_i)]^T$
 q_i is orthogonal to null space of $[N(x_i, p_i)]^T$.

Then,

$$(ix) \quad \|(\nabla_{xx} L(z_i) - Q_i)(x'_i - x_i)\| = \|(\nabla_{xx} L(z_i) - Q_i)(q_i + r_i)\|.$$

We already know that :

$$(x) \quad \|(\nabla_{xx} L(z_i) - G_i)\| \rightarrow 0, \text{ as } i \rightarrow \infty$$

$$(xi) \quad T(x_i) \rightarrow 0, \text{ as } i \rightarrow \infty$$

$$(xii) \quad a_i \leq L' \text{ for all } i$$

$$(xiii) \quad [N(x_i, p_i)]^T r_i = 0, \text{ for all } i.$$

Hence, from (ix), (x), (xi), (xii) and (xiii), there exists a $M_4 < \infty$ such that :

$$\| (\nabla_{xx} L(z_i) - Q_i)(x'_i - x_i) \| \leq M_4 \| q_i \| + o(\| r_i \|) .$$

But, $q_i \perp r_i \Rightarrow \| r_i \| \leq \| p_i \|$. Then,

$$(xiv) \quad \| (\nabla_{xx} L(z_i) - Q_i)(x'_i - x_i) \| \leq M_4 \| q_i \| + o(\| p_i \|) .$$

Now, as shown in the proof of Proposition I.4.R,

$$|g^j(x_{i-1} + p_{i-1} + \tilde{p}_{i-1})| \approx \| p_{i-1} \|^3$$

for all $j \in I(x^*) \triangleq \{j \in \underline{m} \mid g^j(x^*) = 0\}$.

Hence, from Proposition I.4.T,

$$(xv) \quad |g^j(x_i)| \approx \| p_{i-1} \|^3 , \text{ for all } j \in I(x^*) .$$

On the other hand, q_i is the minimum norm solution of,

$$(xvi) \quad [N(x_i, p_i)]^T p + r(x_i) = 0 .$$

Then, from (xvi), (xv), Proposition I.4.0 and (i), we obtain,

$$(xvii) \quad \| q_i \| = o(\| x'_i - x_i \|) \quad \text{for all } i \in \tilde{K}_1 .$$

From (viii), (xi), (xiv) and (xvii) we have :

$$(xviii) \quad \left\| (\nabla_{xx} L(z_i) - B_i)(x_i' - x_i) \right\| = o(\|x_i' - x_i\|) \\ \text{for all } i \in \tilde{K}_1.$$

$$\text{But, } \|x_i' - x_i\| \leq \|z_i' - z_i\|.$$

Then, from (xviii) and (vi) we have :

$$(xix) \quad \|q(z_i')\| \leq \tilde{\alpha}_i \|z_i' - z_i\|, \text{ for all } i \in \tilde{K}_1 \\ \text{with } \tilde{\alpha}_i \rightarrow 0 \text{ as } i \rightarrow \infty.$$

Also, $(z_i' - z^*) \rightarrow 0$ and $\tilde{p}_i \rightarrow 0$, so, there exists $M_5 < \infty$ such that:

$$(xx) \quad \|q(z_{i+1})\| \leq \|q(z_i')\| + M_5 \|\tilde{p}_i\| \\ \text{for all } i \in \tilde{K}_1.$$

But,

$$(xxi) \quad \|\tilde{p}_i\| \approx \|p_i\|^2 \leq \|z_i' - z_i\|^2.$$

Hence, from (xix), (xx) and (xxi) we obtain :

$$(xxii) \quad \|q(z_{i+1})\| \leq \alpha_i' \|z_i' - z_i\|, \text{ for all } i \in \tilde{K}_1 \\ \text{with } \alpha_i' \rightarrow 0 \text{ as } i \rightarrow \infty.$$

Then, for (xxii) and (vii) we have :

$$(xxiii) \quad \|p_{i+1}\| \leq \|z_{i+1}' - z_{i+1}\| \leq \alpha_i'' \|z_i' - z_i\|, \text{ for all } i \in \tilde{K}_1 \\ \text{with } \alpha_i'' \rightarrow 0 \text{ as } i \rightarrow \infty.$$

Now, as shown by Powell (1978a),

$$(xxiv) \quad \| p_i \| \sim \| r(x_i) \| + \| P_i \nabla f(x_i) \| .$$

Define the vector $\tilde{w}_i = [\tilde{\lambda}_i^j, j \in I(x^*) : \tilde{\mu}_i^1, \dots, \tilde{\mu}_i^\ell]$

as follows:

$$\tilde{w}_i \triangleq \arg \min_w \{ \| N_i^T w + \nabla f(x_i) \| \}$$

where,

$$N_i \triangleq N(x_i, p_i) , \quad \text{for all } i.$$

Hence,

$$(xxv) \quad N_i [N_i^T \tilde{w}_i + \nabla f(x_i)] = 0.$$

As $\psi(x_i) \rightarrow 0$, we can see that :

$$(xxvi) \quad \tilde{\lambda}_i \rightarrow \bar{\lambda}(x_i) \quad \text{and} \quad \tilde{\mu}_i \rightarrow \bar{\mu}(x_i)$$

as $i \rightarrow \infty$.

Now, because $\tilde{w}_i$ is the least square solution of (xxv), the projection of $\nabla f(x_i)$ onto the null space of N_i^T is :

$$(xxvii) \quad P_i \nabla f(x_i) = \nabla f(x_i) + N_i^T \tilde{w}_i .$$

From (xxvi) and (xxvii) we have :

$$(xxviii) \quad \| P_i \nabla f(x_i) \| \sim \| \nabla f(x_i) + g_X^T(x_i) \bar{\lambda}(x_i) + h_X^T(x_i) \bar{\mu}(x_i) \|$$

But, $\lambda_{i-1} \rightarrow \lambda^*$, $\lambda_i \rightarrow \lambda^*$, $\mu_{i-1} \rightarrow \mu^*$, $\mu_i \rightarrow \mu^*$, $\dot{\lambda}(x_i) \rightarrow \lambda^*$
and $\bar{\mu}(x_i) \rightarrow \mu^*$, as $i \rightarrow \infty$

Hence, from (xxviii) we obtain :

$$(xxix) \quad \| P_i \nabla f(x_i) \| \sim \| \nabla_{x^L}(z_i) \| .$$

Recall that :

$$(xxx) \quad r(x_i) \triangleq [g^j(x_i), j \in I(x^*), h^j(x_i), j \in \underline{\ell}] .$$

Also,

$$(xxxi) \quad \lambda_i^j \leq c, \text{ for all } j \text{ and } i \text{ sufficiently large.}$$

and

$$(xxxii) \quad \left. \begin{array}{l} \lambda_i^j = 0, j \notin I(x^*) \\ \text{and} \\ \lambda_i^j > 0, j \in I(x^*) \end{array} \right\} , \text{ for } i \text{ sufficiently large.}$$

From (xxix), (xxx), (xxxi), (xxxii) and the definition of $q(\cdot)$, we have :

$$(xxxiii) \quad \| q(z_i) \| \sim \| r(x_i) \| + \| P_i \nabla f(x_i) \| .$$

But, from (vii) and (xxii), we have :

$$(xxxiv) \quad \| p_{i+1} \| \leq \alpha_i'' M_2 \| q(z_i) \| , \text{ for all } i \in \tilde{K}_1$$

with $\alpha_i'' \rightarrow 0$ as $i \rightarrow \infty$.

As shown in the proof of Proposition I.4.U,

$$(xxxv) \quad \|p_{i+1}\| \sim \|x_{i+1} - x^*\| ,$$

Hence, from (xxxv), (xxxiv), (xxxiii), (xxiv) and (iii) we have :

$$(xxxvi) \quad \frac{\|x_{i+1} - x^*\|}{\|x_i - x^*\|} \xrightarrow{I'} 0, \quad \text{as } i \rightarrow \infty$$

for any infinite subsequence $I' \subset \tilde{K}_1$.

As $\{1, 2, \dots\} = K_1 \cup K_2$, the theorem follows from (v), (xxxvi) and Proposition I.4.U.

I.4.5 Remarks

First of all, note the similarity between the algorithms of this chapter and the Basic Algorithm of the first chapter. Besides the introduction of an update procedure to B_i , the only difference is the test after the computation of $\tilde{p}_i$ in Algorithm B. Namely :

$$(i) \quad \|B_i \tilde{p}_i\| > M_1 \|p_i\| \Rightarrow \tilde{p}_i = 0.$$

Now, by definition, $\tilde{p}_i$ is only set different from zero, if,

$$\|\tilde{p}_i\| \leq \|p_i\| .$$

Hence, as the constant M_1 can be chosen arbitrarily large, the reason for the inclusion of test (i) is that (in contrast to the first chapter where H3 was assumed) we do not not assume the boundedness of $\{B_i\}$ when proving Theorem I.4.N. It is interesting to point out

how the absence of H3 was overcome in the proof of Theorem I.4.N, because this can cast some light in other features of the update given by I.4.VI. Recall that H3 implies existence of two constants, $m' > 0$ and $M < \infty$, such that :

$$(H3a) \quad p^T B_i p \geq m' \|p\|^2, \text{ for all } i \text{ and all } p \in \mathbb{R}^n.$$

and

$$(H3b) \quad p^T B_i p \leq M \|p\|^2, \text{ for all } i \text{ and all } p \in \mathbb{R}^n.$$

Assumption H3a was avoided by the addition of $T(x_i)I$ to the positive definite matrix H_i obtained in Step 8. The sequence $\{B_i\}$ constructed is then shown to have positive eigenvalues which are bounded away from zero, except perhaps, in subsequences such that

$$x_i \rightarrow x^* \in D.$$

Now, test (i) is related to the way in which the necessity for H3b was avoided. Although H3b is not enforced, we can still prove the existence of the following bounds :

$$(ii) \quad \frac{\|\eta_i\|}{\delta_i^T \eta_i} \leq \tilde{K} < \infty, \text{ for all } i.$$

$$(iii) \quad \|B_i p_i\| \leq M_2 < \infty, \text{ for all } i.$$

Note that (ii) can only be proved because (in contrast to the suggestion of Powell (1978)) the update given by I.4.VI does not include H_i in the definition of η_i . The bound of (iii) comes from the satisfaction of the Kuhn-Tucker conditions at the solution of QP_i^W with multiplier vectors which remain bounded.

An examination of the proof of Theorem I.4.N shows that (iii), (ii), test (i) and the convergence of $p_i \rightarrow 0$ in all convergent subsequences (see Proposition I.4.H) are the means for avoiding H3b.

However, the main interest of Theorem I.4.N is theoretical. The possibility of showing the result of this theorem suggests that indeed the sequence of matrices will remain bounded whenever $\{x_i\}$ is itself bounded.

This point will be briefly discussed in the next chapter.

We now discuss the rate of convergence of the algorithms and compare the results of section I.4.4 with other results in the literature.

Powell (1978a) proved that a variable metric algorithm to solve GMP, which uses in the quadratic sub-problem a matrix B_i obtained through another modification (see I.4.III) of the BFGS update given by I.4.I, will have a R-superlinear rate of convergence, provided that the following assumptions are satisfied :

- P1 - f, g, h are twice continuously differentiable and the second order derivatives are Lipschitz continuous on R^n .
- P2 - $x_i \rightarrow x^*$ as $i \rightarrow \infty$.
- P3 - equivalent to H4.
- P4 - there exists a positive constant v_1 such that, for i sufficiently large,

$$y^T B_i y \geq v_1 \|y\|^2.$$

for all $y \in R^n$ satisfying :

$$y^T (N^*) (N^*)^T y = 0$$

where, $N^* \triangleq N(x^*, 0)$.

P5 - The step length is of unity.

P6 - $\{B_i\}$ remains bounded.

Note that P1 is slightly milder than H1' in the sense that P1 implies H1' almost everywhere. Also, P4 is milder than H3(a), although P4 together with P3 indicate some relation between the accumulation points of $\{B_i\}$ and $\nabla_{xx} L(z^*)$.

The next proposition shows that H3(a) can be substituted by P4 in our rate of convergence result.

Proposition I.4.X

Theorem I.4.V will remain valid if assumption H3a is replaced by assumption P4 above.

Proof:

Given in the appendix.

A less clear correspondence is the one between H5 and P5.

This is the object of the next proposition.

Proposition I.4.W

Theorem I.4.V will remain valid if assumption H5 is replaced by P5.

Proof:

Because assumptions H1, H2, H3a, H3b, H4, Theorem I.1.M and the step length of unity ensure that all conditions listed above are satisfied, to prove a R-superlinear rate of convergence to the algorithm it is convenient to follow the method of Powell (1978a). We start showing that the key result of Lemma 7 in this last paper has a corresponding result, whose proof is given in the appendix. That is :

Lemma I.4.Y

Let $\bar{v}_3$ be a positive constant such that :

$$\| r(x_i) \| \leq \bar{v}_3 \| r_i \| \quad \text{for all } i.$$

Then,

$$\begin{aligned} \| D_i^{\frac{1}{2}} H_{i+1} D_i^{\frac{1}{2}} - I \|_F &\leq \| D_i^{\frac{1}{2}} H_i D_i^{\frac{1}{2}} - I \|_F - \frac{(D_i H_i r_i - r_i)^T H_i (D_i H_i r_i - r_i)}{r_i^T H_i r_i} \\ &\quad + \bar{\varepsilon} + o(\| p_i \|^3) \end{aligned}$$

where:

$$\bar{\varepsilon} \approx (\| r(x_i) \| / \| r_i \|) \text{ and}$$

$$D_i = (P_i G_i P_i)^+, \text{ with } + \text{ denoting the generalized inverse.}$$

If we follow carefully the subsequent results of Powell (1978a) we shall see that they are not affected by the additional term $o(\| p_i \|^3)$. Actually, the results become easier to prove because the above lemma does not involve the update parameter θ_i . Then, if δ_i were equal to the solution of the quadratic sub-problem with H_i replacing B_i , the proposition would be proved. That is, if $p_i^!$ denotes this solution :

$$(i) \quad \lim_{i \rightarrow \infty} \frac{\| p_{i+1}^! \|}{\| p_{i-1}^! \|} = 0.$$

Actually, as the step length is of unity we have :

$$(ii) \quad \delta_i = p_i + \tilde{p}_i.$$

But, as shown in Proposition I.4.Q,

$$(iii) \quad \| \tilde{p}_i \| \approx \| p_i \|^2 ,$$

From (i) we have :

$$(iv) \quad \| p_i \| - \| \tilde{p}_i \| \leq \| \delta_i \| \leq \| p_i \| + \| \tilde{p}_i \|$$

Hence, from (iii) and (iv), there exist positive constants m_0 and M_0 such that, for all i sufficiently large :

$$0 < m_0 \leq (\| \delta_i \| / \| p_i \|) \leq M_0 < \infty.$$

Then,

$$(v) \quad \lim_{i \rightarrow \infty} \frac{\| \delta_{i+1} \|}{\| \delta_{i-1} \|} \leq \frac{M_0}{m_0} \lim_{i \rightarrow \infty} \frac{\| p_{i+1} \|}{\| p_{i-1} \|} .$$

Now define :

$$d_i = p_i - p'_i .$$

Similarly to (iv), we have :

$$(vi) \quad \| p'_i \| - \| d_i \| \leq \| p_i \| \leq \| p'_i \| + \| d_i \| .$$

But,

$$\| B_i - H_i \| = T(x_i) .$$

Hence, from Lemma I.1.E, there exists a constant $M' < \infty$ such that :

$$(vii) \quad \|d_i\| \leq T(x_i)M_1 \|p'_i\|.$$

But, $T(x_i) \rightarrow 0$ as $i \rightarrow \infty$.

Then, from (vii) and (vi), there exist positive constants m_1 and M_2 such that, for all i sufficiently large :

$$0 < m_1 \leq (\|p_i\| / \|p'_i\|) \leq M_1 < \infty.$$

Hence, the proposition is proved because from (i) and (v),

$$\lim_{i \rightarrow \infty} \frac{\|\delta_{i+1}\|}{\|\delta_{i+1}\|} \leq \frac{M_0}{m_0} \quad \lim_{i \rightarrow \infty} \frac{\|p_{i+1}\|}{\|p_{i-1}\|} \leq \frac{M_0 M_1}{m_0 m_1} \quad \lim_{i \rightarrow \infty} \frac{\|p'_{i+1}\|}{\|p'_{i-1}\|} = 0.$$

Because the use of the search arc $(\alpha p_i + \alpha^2 \tilde{p}_i)$ still permits a monotone decrease on the exact penalty function, it is possible to show (see Theorem I.1.M) that, under the assumptions H1, H2, H3, H4, we indeed have :

$$x_i \rightarrow x^* \in D \quad \text{as } i \rightarrow \infty.$$

But, if the alternative approach of the 'watchdog technique' [see Chamberlain et al (1982)] is used to achieve a step length of unity in variable metric algorithms, as done by Powell (1982b), only the weaker result of Theorem 1 of Chamberlain et al (1982) could be proved. That is :

there exists a subsequence converging to a point belonging to D .

Hence, when using the 'watchdog technique', the simultaneous requirement of P2 and P5 seems to be a bit too strong. It would thus be interesting to know whether H5 and the use of the search arc would also ensure a step length of unity when using the update of Powell (1978).

Recall that in H5 the vector η_i was defined as :

$$\eta_i = G_i \delta_i + a_i (T(x_i)I + N_i N_i^T) \delta_i$$

where,

a_i is given by I.4.VI.

The corresponding vector η'_i in the update of Powell (1978) is:

I.4.VII

$$\eta'_i = \theta_i G_i \delta_i + (1-\theta_i) H_i \delta_i$$

where :

$$\theta_i = \begin{cases} 1, & \text{if } \delta_i^T \gamma_i \geq 0.2 \delta_i^T H_i \delta_i \\ \frac{0.8 \delta_i^T H_i \delta_i}{\delta_i^T H_i \delta_i - \delta_i^T \gamma_i}, & \text{if } \delta_i^T \gamma_i < 0.2 \delta_i^T H_i \delta_i. \end{cases}$$

Now, when proving Proposition I.4.T directly through the reasoning of Theorem I.4.S, we would arrive at the following inequality :

$$\begin{aligned} \langle \delta_i, (\nabla_{xx} L(x_i, \lambda_i, \mu_i) - H_i) \delta_i \rangle &= \langle \delta_i, (\nabla_{xx} L(x_i, \lambda_i, \mu_i) \delta_i - \eta_i) \rangle \\ &\quad + o(\|\delta_i\|^2) \end{aligned}$$

As $a_i (T(x_i)I + N_i N_i^T)$ is positive definite, we deduce that:

$$\begin{aligned} \langle \delta_i, (\nabla_{xx} L(x_i, \lambda_i, \mu_i) - H_i) \delta_i \rangle &\leq \langle \delta_i, (\nabla_{xx} L(x_i, \lambda_i, \mu_i) - G_i) \delta_i \rangle \\ &+ o(\|\delta_i\|^2) . \end{aligned}$$

The inequality above and the convergence

$$(G_i - \nabla_{xx} L(x_i, \lambda_i, \mu_i)) \rightarrow 0 \quad \text{as } i \rightarrow \infty$$

is enough to ensure a step length of unity.

But, if we had η'_i instead of η_i , after similar reasoning we would arrive at the following inequality :

$$\begin{aligned} \langle \delta_i, (\nabla_{xx} L(x_i), \lambda_i, u_i) - H_i) \delta_i \rangle &\leq \langle \delta_i, (\nabla_{xx} L(x_i, \lambda_i, \mu_i) - \theta_i G_i) \delta_i \rangle \\ &+ o(\|\delta_i\|^2) . \end{aligned}$$

Hence, the convergence of $(G_i - \nabla_{xx} L(x_i, \lambda_i, \mu_i))$ to zero would no longer ensure a step length of unity, unless $\theta_i = 1$ for i large enough, something that may not happen. Another important difference between η_i and η'_i is that the latter depends on H_i while in the former this matrix does not appear. As already pointed out, this feature enabled us to prove the result of the Theorem I.4.N, without requiring H3(b).

It would thus be interesting to know whether the sequence of matrices $\{H_i\}$ will indeed remain bounded whenever the sequence $\{x_i\}$ is itself bounded. This would be a useful result, particularly if the convergence $H_i \rightarrow H^*$ could be proved as then H5 is shown to be satisfied. On the other hand, H5 is not sufficient to establish a bound for the sequence $\{B_i\}$, although a slightly weaker result can be shown.

Proposition I.4.2

Let H1, H2, H3a, H4 and H5 be satisfied.

Let $\{x_i\}$ be a bounded infinite sequence built by the algorithm.

Then ,

there exists a $L^* < \infty$ such that :

$$\|B_i p_i\| \leq L^* \|p_i\|, \text{ for all } i.$$

Proof:

By the continuity of $f_{xx}(\cdot)$, $g_{xx}(\cdot)$, $h_{xx}(\cdot)$, $\lambda(\cdot)$, $\bar{\mu}(\cdot)$, $T(\cdot)$ and the convergences,

$$x_i \rightarrow x^* \in D; \quad p_i \rightarrow 0, \quad \text{as } i \rightarrow \infty,$$

it is clear that :

(i) there exists $L_0 < \infty$ such that :

$$\|Q_i\| \leq L_0, \quad \text{for all } i.$$

where,

$$Q_i \triangleq G_i + a_i (T(x_i)I + [N(x_i, p_i)][N(x_i, p_i)]^T)$$

Now,

$$\eta_i = Q_i \delta_i, \quad \text{for all } i.$$

Hence, from H5 we have :

$$(ii) \quad \|H_i \delta_i\| \leq \|Q_i \delta_i\| + k_i'' \|\delta_i\|$$

where,

$$k_i'' \rightarrow 0 \quad \text{and } i \rightarrow \infty.$$

Then, there exists $L_1 < \infty$ such that :

$$\| H_i \delta_i \| \leq L_1 \| \delta_i \| .$$

Now,

$$\delta_i = \beta^{k_i} p_i + \beta^{2k_i} \tilde{p}_i .$$

From Proposition I.4.P, we have a $r' > 0$ such that :

$$\beta^{k_i} > r' > 0, \quad \text{for all } i.$$

Hence, from (ii), we have :

$$r' \| H_i p_i \| \leq \| \beta^{k_i} H_i p_i \| \leq \| Q_i \delta_i \| + k_i'' \| \delta_i \| + \| \beta^{2k_i} H_i \tilde{p}_i \| .$$

But, by construction, $\beta^{k_i} \leq 1$ for all i . Then,

$$\begin{aligned} \text{(iii)} \quad \| H_i p_i \| &\leq (1/r') [\| Q_i \| (\| p_i \| + \| \tilde{p}_i \|) + \\ &\quad + k_i'' (\| p_i \| + \| \tilde{p}_i \|) + \| H_i \tilde{p}_i \|] . \end{aligned}$$

Now, $B_i = H_i + T(x_i)I$ and $T(x_i) \rightarrow 0$ as $x_i \rightarrow x^* \in D$.

Then, from the test of Step 3', we have :

(iv) there exists a $L_2 < \infty$ such that :

$$\| H_i \tilde{p}_i \| \leq L_2 \| p_i \| , \quad \text{for all } i.$$

By construction, $\| \tilde{p}_i \| \leq \| p_i \|$. Then, from (iv), (iii) and (i), we obtain :

there exists a $L_3 < \infty$, such that :

$$\| H_i p_i \| \leq L_3 \| p_i \| .$$

The proposition follows from the boundedness of $T(x_i)$.

Note that, in the definition of γ_i , we used the estimates $\bar{\lambda}(\cdot)$, $\bar{\mu}(\cdot)$ while Powell (1978) suggests the use of the multipliers obtained when solving the quadratic sub-problem. When using QP_i^0 the multipliers may become unbounded even if the sequence $\{x_i\}$ remains bounded. Consequently [see Chamberlain (1979)] the update given by I.4.VI, with η_i given by I.4.VII [as suggested by Powell (1978)] may make the sequence of matrices $\{H_i\}$ unbounded and the algorithm may cycle between non-desirable points.

Because we use QP_i^W instead of QP_i^0 , it is possible to show (From I.1.A.i and the fact that : $\{x_i\}$ bounded $\Rightarrow c_i = c > 0$ for i sufficiently large) that the multipliers (λ_i, μ_i) will remain bounded whenever the algorithm generates a bounded sequence $\{x_i\}$. Therefore, the vector γ_i could also be defined as below without bringing further complications.

$$\gamma_i = \nabla_x L(x_{i+1}, \lambda_i, \mu_i) - \nabla_x L(x_i, \lambda_i, \mu_i)$$

where,

$$(\lambda_i, \mu_i) \text{ is the multiplier vector of } QP_i^W.$$

Finally, we comment on the two-step superlinear rate of convergence given by Theorem I.4.V. It was shown during the proof of this theorem, that a Q-superlinear rate of convergence can be achieved if,

I.4.VIII

there exists a $r_1 > 0$ such that $\|p_i\| \geq r_1 \|p_{i-1}\|^2$
for all i .

An interesting consequence of this condition is that it implies the limit

$$\lim_{i \rightarrow \infty} \frac{\|r(x_i)\|}{\|x_i - x^*\|} = 0$$

which is a condition reckoned by Powell (1978a) as sufficient for establishing a Q-superlinear rate of convergence for an algorithm which, otherwise, could only be proved to be R-superlinearly convergent.

In fact, as shown in the proof of Theorem I.4.V,

$\|r(x_i)\| \approx \|p_{i-1}\|^3$. But, as shown in the proof of Proposition I.4.V,

$$\|x_i - x^*\| \sim \|p_i\|.$$

Hence,

$$\frac{\|r(x_i)\|}{\|x_i - x^*\|} \approx \frac{\|p_{i-1}\|^3}{\|p_i\|}$$

and, from I.4.VIII,

$$\lim_{i \rightarrow \infty} \frac{\|r(x_i)\|}{\|x_i - x^*\|} = 0.$$

The conclusion was made possible because the search arc satisfies the constraints to second order. Indeed, if the

search line is used we can only guarantee that :

$$\| r(x_i) \| \approx \| p_{i-1} \|^2 .$$

In this case, the corresponding sufficient condition to the above limit being zero is that :

there exists a $r_2 > 0$ such that

$$\| p_i \| \geq r_2 \| p_{i-1} \| .$$

As $p_i \rightarrow 0$, this last condition is stronger than I.4.VIII.

In fact, the use of the search arc also made it possible to show (see proof of Theorem I.4.V) that :

for any convergent subsequence I' either

$$\lim_{i \in I'} \frac{\| x_{i+1} - x^* \|}{\| x_i - x^* \|} = 0$$

or

$$\lim_{i \in I'} \frac{\| x_{i+1} - x^* \|}{\| x_{i-1} - x^* \|^2} = 0 .$$

In other words, for the subsequences for which we could not prove one step superlinear rate of convergence we were able to show super-quadratic two step rate of convergence. Hence, proving a R-superlinear rate of convergence instead of a Q-superlinear rate does not seem to be important.

APPENDIX (chapter I.4)

AP.1 - Root-Convergence Factors [From Ortega and Rheinboldt (1970)]

Definition:

Let $\{x_i\} \subset \mathbb{R}^n$ be any sequence converging to x^* . Then the numbers

$$R_p\{x_i\} = \begin{cases} \limsup_{i \rightarrow \infty} \|x_i - x^*\|^{1/i} & , \text{ if } p = 1 \\ \limsup_{i \rightarrow \infty} \|x_i - x^*\|^{1/p^i} & , \text{ if } p > 1 \end{cases}$$

are the root convergence factors, or R-factors for short, of $\{x_i\}$.

If A is an iterative process with limit point x^* , and $C(A, x^*)$ is the set of all sequences generated by A which converge to x^* , then

$$R_p(A, x^*) = \sup_{1 \leq p < \infty} \{R_p\{x_i\} \mid \{x_i\} \in C(A, x^*)\},$$

are the R-factors of A at x^* .

Definition:

If $0 < R_1(A, x^*) < 1$, we shall say that the convergence of A at x^* is R-linear while for $R_1 = 0$ and $R_1 = 1$ we call the convergence R-superlinear or R-sublinear, respectively. Similarly, if

$$0 < R_2(A, x^*) < 1$$

we say that the convergence is quadratic, while for $R_2 = 0$ we call the convergence R-superquadratic.

Proof of Proposition I.4.B

Let $T'_i : R \rightarrow R$ be the map defined by :

$$T'_i(\alpha) \triangleq \gamma(x_i + \alpha p_i + \alpha^2 \tilde{p}_i, c_i) - \gamma(x_i, c_i) - \alpha \theta(x_i, p_i, c_i) / 8 .$$

Then,

$$(i) \quad T'_i(\alpha) = f(x_i + \alpha p_i + \alpha^2 \tilde{p}_i) - f(x_i) + c_i [\psi(x_i + \alpha p_i + \alpha^2 \tilde{p}_i) - \psi(x_i)] - \alpha \theta(x_i, p_i, c_i) / 8 .$$

Because of H1' there exists $k'' < \infty$ such that for $i = i^*$ we have:

$$(ii) \quad f(x_i + \alpha p_i + \alpha^2 \tilde{p}_i) \leq f(x_i) + \alpha f_x(x_i)(p_i + \alpha \tilde{p}_i) + \alpha^2 k'' \|p_i + \alpha \tilde{p}_i\|^2$$

and, for all $j \in \underline{m}$,

$$(iii) \quad g^j(x_i + \alpha p_i + \alpha^2 \tilde{p}_i) \leq g^j(x_i) + \alpha g_x^j(x_i)(p_i + \alpha \tilde{p}_i) + \alpha^2 k'' \|p_i + \alpha \tilde{p}_i\|^2$$

and, for all $j \in \underline{\ell}$

$$(iv) \quad h^j(x_i + \alpha p_i + \alpha^2 \tilde{p}_i) \leq h^j(x_i) + \alpha h_x^j(x_i)(p_i + \alpha \tilde{p}_i) + \alpha^2 k'' \|p_i + \alpha \tilde{p}_i\|^2 .$$

From (i), (ii), (iii) and (iv), we have :

$$(v) \quad T'_i(\alpha) \leq \alpha \theta(x_i, p_i, c_i) + \alpha^2 f_x(x_i) \tilde{p}_i + \alpha^2 [\max\{g^j(x_i) \tilde{p}_i, j \in \underline{m}\}] + \alpha^2 [\max\{|h^j(x_i) \tilde{p}_i|, j \in \underline{\ell}\}] + \alpha^2 k'' \|p_i + \alpha \tilde{p}_i\|^2 - \alpha \theta(x_i, p_i, c_i) / 8$$

for $i = i^*$.

Also, there exists a $K''' < \infty$ such that :

$$(vi) \quad f_x(x_i) \tilde{p}_i \leq K''' \|\tilde{p}_i\|$$

$$(vii) \quad g_x^j(x_i) \tilde{p}_i \leq K''' \|\tilde{p}_i\|, \text{ for all } j \in \underline{m}$$

$$(viii) \quad |h^j(x_i) \tilde{p}_i| \leq K''' \|\tilde{p}_i\|, \text{ for all } j \in \underline{\ell}.$$

From (v), (vi), (vii) and (viii) we obtain :

$$(ix) \quad T_i'(\alpha) \leq \alpha(7/8)\theta(x_i, p_i, c_i) + 3K''' \|\tilde{p}_i\| \alpha^2 + 3K'' \|p_i + \alpha \tilde{p}_i\|^2 \alpha^2. \\ \text{for } i = i^*.$$

Now, for any finite i^* , there exists a m_{i^*} such that,

$$T(x_{i^*}) > m_{i^*} > 0$$

unless $x_{i^*} \in D$, hypothesis that we did not consider because we are interested in infinite sequences.

Hence, from Proposition I.1.A and the way $\{B_i\}$ is defined, we have:

$$(x) \quad \theta(x_{i^*}, p_{i^*}, c_{i^*}) \leq -p_{i^*}^T B_{i^*} p_{i^*} \leq -m_{i^*} \|p_{i^*}\|^2.$$

From (x), (ix) and Proposition C, we obtain :

$$(xi) \quad T_{i^*}'(\alpha) \leq (7/8) m_{i^*} \|p_{i^*}\|^2 \alpha + 3K''' \|\tilde{p}_{i^*}\| \alpha^2 + 3K'' \|p_{i^*} + \alpha \tilde{p}_{i^*}\|^2 \alpha^2.$$

Suppose $\tilde{p}_{i*} = 0$.

Then, from (x_i) , we have :

$$(xii) \quad T_{i*}(\alpha) \leq 0, \text{ for all } \alpha \leq (7/8) m_{i*}/K''$$

Otherwise let $\tilde{p}_{i*} \neq 0$.

From the definition of $\tilde{p}_i$ we then have :

$$(xiii) \quad w_{i*} = 0.$$

By the mean value theorem, there exist $\xi_{i*}^j \in [x_{i*}, x_{i*} + p_{i*}]$ such that, for all $j \in (\underline{m}\underline{\ell})$, we have :

$$(xiv) \quad g^j(x_{i*} + p_{i*}) = g^j(x_{i*}) + g_x^j(x_{i*})p_{i*} + \frac{1}{2}p_{i*}^T [\nabla_{xx} g^j(\xi_{i*}^j)] p_{i*}$$

and

$$(xv) \quad h^j(x_{i*} + p_{i*}) = h^j(x_{i*}) + h_x^j(x_{i*})p_{i*} + \frac{1}{2}p_{i*}^T [\nabla_{xx} h^j(\xi_{i*}^j)] p_{i*}$$

From (xiii), at the solution of QP_{i*}^w , we have :

$$(xvi) \quad g^j(x_{i*}) + g_x^j(x_{i*})p_{i*} = 0, \quad j \in \{j \in \underline{m} \mid \lambda_i^j > 0\}$$

and

$$(xvii) \quad h^j(x_{i*}) + (h_x^j(x_{i*})p_{i*}) = 0, \quad j \in \underline{\ell}.$$

Hence, from (xvii), (xvxi), (xv), (xiv) and H1', there exists a $\hat{K} < \infty$, such that :

$$\begin{aligned}
 \text{(xix)} \quad T_{i*}(\alpha) \leq & -(7/8) m_{i*} \|p_{i*}\|^2 \alpha + 3K''' \hat{K} \|p_{i*}\|^2 \alpha^2 + 3K'' \|p_{i*}\|^2 \alpha^2 \\
 & + 3K''' \hat{K} 2\alpha \|p_{i*}\|^3 \alpha^3 + 3K''' \hat{K}^2 \|p_{i*}\|^2 \alpha^4.
 \end{aligned}$$

Now, the last two terms on the right hand side of the above inequality can be rewritten as :

$$\begin{aligned}
 \text{(xx)} \quad 3K'' \hat{K} 2\alpha \|p_{i*}\|^3 \alpha^3 & \equiv (3K''' \hat{K} 2\alpha^2 \|p_{i*}\|) \|p_{i*}\|^2 \alpha^2 \\
 & \triangleq d_{i*} \|p_{i*}\|^2 \alpha^2
 \end{aligned}$$

and

$$\begin{aligned}
 3K''' \hat{K}^2 \|p_{i*}\|^4 \alpha^4 & \equiv (3K''' \hat{K}^2 \|p_{i*}\|^2 \alpha^2) \|p_{i*}\|^2 \alpha^2 \\
 & \triangleq b_{i*} \|p_{i*}\|^2 \alpha^2.
 \end{aligned}$$

But, as shown in the proof of Proposition I.1.C there exists a continuous function $S(\cdot)$ such that :

$$\begin{aligned}
 \text{(xxi)} \quad \|p_{i*}\| \leq S(x_{i*}) & \triangleq (1/m_{i*}) \|\nabla f(x_{i*})\| + [(1/m_{i*})^2 \|\nabla f(x_{i*})\| + \\
 & + 2(c_i/m_{i*})^2 \psi(x_{i*})]^{1/2}.
 \end{aligned}$$

Hence, from (xx), (xxi) and H_1' there exists a $K_0 < \infty$ such that :

$$\text{(xxii)} \quad d_{i*} \leq K_0; \quad b_{i*} \leq K_0.$$

From (xxii), (xx) and (xix), we obtain :

$$T_{i*}(\alpha) \leq -(7/8) m_{i*} \|p_{i*}\|^2 \alpha + [3K''' + 3K'' + 2K_0] \|p_{i*}\|^2 \alpha^2.$$

Hence,

$$(xxiii) \quad T_{i*}(\alpha) \leq 0, \quad \text{for all } \alpha \leq (7/8) \frac{m_{i*}}{3K''' + 3K'' + 2K_0}.$$

The proposition then follows from (xxiii) and (xii).

Proof of Proposition I.4.J

Let $K_3 \subset \{1, 2, \dots\}$ be the subset where $a_i = 0$.

By construction,

$$(i) \quad \delta_i^T \eta_i \geq \varepsilon_1 \|\delta_i\|^2, \quad \text{for all } i \in K_3.$$

Let $K_4 \subset \{1, 2, \dots\}$ be the subset where $a_i \geq 1$.

By construction,

$$(ii) \quad \delta_i^T \eta_i \geq T(x_i) \|\delta_i\|^2 + \delta_i^T [N(x_i, p_i)] [N(x_i, p_i)]^T \delta_i$$

for all $i \in K_4$.

Because of Proposition I.4.A, Proposition I.4.B, Proposition I.3.A and as we are interested in infinite sequences $\{x_i\}$, we can say, without loss of generality, that :

$$(iii) \quad T(x_i) > 0, \quad \text{for all finite } i.$$

Hence, by (i), (ii) and (iii), it only remains to prove the existence of such $r^* > 0$ in the subsequences $\{x_i\}_I \subset K_4$ such that

$$(iv) \quad T(x_i) \xrightarrow{I} 0.$$

As $T(\cdot)$ is continuous, (iv) and Proposition I.3.A imply,

$$(v) \quad x_i \xrightarrow{I} x^* \in D.$$

By Proposition I.4.H,

$$(vi) \quad p_i \xrightarrow{I} 0.$$

Consequently, $\tilde{p}_i \xrightarrow{I} 0$; $\delta_i \rightarrow 0$.

Now, by construction,

$$\begin{aligned} (vii) \quad \gamma_i &\triangleq \nabla_x L(x_{i+1}, \bar{\lambda}(x_i), \bar{\mu}(x_i)) - \nabla_x L(x_i, \bar{\lambda}(x_i), \bar{\mu}(x_i)) \\ &= G_i \delta_i \end{aligned}$$

where,

$$G_i = \int_{t=0}^1 \nabla_{xx} L(x_i + t\delta_i, \bar{\lambda}(x_i), \bar{\mu}(x_i)) dt.$$

Let (λ^*, μ^*) be the multiplier associated with $x^* \in D$.

By Proposition 2 of Mayne and Polak (1978) we have :

$$(viii) \quad \bar{\lambda}(x_i) \xrightarrow{I} \lambda^* ; \quad \bar{\mu}(x_i) \xrightarrow{I} \mu^* .$$

From (v), (vi) and (viii), we obtain :

$$(ix) \quad G_i \xrightarrow{I} G^* \triangleq \nabla_{xx} L(x^*, \lambda^*, \mu^*) .$$

Let I^* denote the set of active constraints to GMP at x^* .

By (viii) and the strict complementarity slackness given by H4, for all $i \in I$ sufficiently large,

$$(x) \quad \bar{\lambda}^j(x_i) > 0 \quad \text{for all } j \in I^*.$$

Also, by H1' and (vi), for all $i \in I$ sufficiently large,

$$(xi) \quad g^j(x_i) + g_x^j(x_i)p_i < 0, \quad \text{for all } j \notin I^*.$$

Hence, by H1', (xi), (x) and the definition of $N(\cdot, \cdot)$, we have :

$$(xii) \quad N(x_i, p_i) \xrightarrow{I} N(x^*, 0).$$

By H4, there exists a $\bar{r}_2 > 0$ such that :

$$(xiii) \quad p^T G^* p \geq \bar{r}_2^2 \|p\|^2 \quad \text{for all } p \in R^n \text{ such that :}$$

$$p^T [N(x^*, 0)] [N(x^*, 0)]^T p = 0.$$

Now, $p^T [N(x^*, 0)] [N(x^*, 0)]^T p \geq 0$, for all $p \in R^n$.

From (ix) and (xii) there exists $\bar{r}_3 > 0$ such that:

$$(xiv) \quad p^T G_i p \geq \bar{r}_3 \|p\|^2$$

for all $p \in R^n$ such that, $p^T [N(x^*, 0)] [N(x^*, 0)]^T p = 0$.

By definition,

$$\delta_i^T \eta_i = \delta_i^T \gamma_i + a_i (T(x_i) \|\delta_i\|^2 + \delta_i^T [N(x_i, p_i)] [N(x_i, p_i)]^T \delta_i).$$

Let $K_5 \subset K_4$ be the subset where $a_i = 1$.

Then, from (vii)

$$(xv) \quad \delta_i^T \eta_i = \delta_i^T G_i \delta_i + \delta_i^T [N(x_i, p_i)] [N(x_i, p_i)]^T \delta_i + T(x_i) \|\delta_i\|^2$$

for all $i \in K_5$.

But, from (ix), (xii), (xiii), (xiv) and the continuity of the scalar product, there exist $\bar{r}_4 > 0$ and i_0 such that :

$$(xvi) \quad \delta_i^T G_i \delta_i + \delta_i^T [N(x_i, p_i)] [N(x_i, p_i)]^T \delta_i \geq \bar{r}_4 \|\delta_i\|^2$$

for all $i \in I$, $i \geq i_0$.

But, from Proposition I.3.A and the continuity of $T(\cdot)$, there exists $\bar{r}_5 > 0$ such that

$$(xvii) \quad T(x_i) \geq \bar{r}_5, \text{ for all } i \in I, i \leq i_0.$$

Hence, from (xvii), (xvi) and (xv), there exists $\bar{r}_6 \triangleq \min\{\bar{r}_5; \bar{r}_4\}$ such that :

$$(xviii) \quad \delta_i^T \eta_i \geq \bar{r}_6 \|\delta_i\|^2, \text{ for all } i \in I \text{ with } a_i = 1.$$

If $a_i > 1$, then $\delta_i^T \gamma_i < 0$. Let $K_6 \subset K_4$ be the subset where,

$$a_i = 1 + \frac{(T(x_i) \parallel \delta_i \parallel^2 - \delta_i^T \gamma_i)}{T(x_i) \parallel \delta_i \parallel^2 + \delta_i^T [N(x_i, p_i)] [N(x_i, p_i)]^T \delta_i}$$

Hence,

$$(xix) \quad \delta_i^T \eta_i = 2T(x_i) \parallel \delta_i \parallel^2 + \delta_i^T [N(x_i, p_i)] [N(x_i, p_i)]^T \delta_i$$

for all $i \in K_6$.

Again by (ix), (xii), (xiii), (xiv) and the continuity of the scalar product, there exists $\bar{r}_7 > 0$ and i_1 such that,

$$(xx) \quad \delta_i^T [N(x_i, p_i)] [N(x_i, p_i)]^T \delta_i \geq \bar{r}_7 \parallel \delta_i \parallel^2$$

for all $i \in I \cap K_6$, $i \geq i_1$.

Also, by Proposition I.3.A and the continuity of $T(\cdot)$, there exists $\bar{r}_8 > 0$ such that,

$$(xxi) \quad T(x_i) > \bar{r}_8, \text{ for all } i \in I, i \leq i_1.$$

From (xix), (xx) and (xxi), there exists $\bar{r}_9 \triangleq \min \{\bar{r}_7, \bar{r}_8\}$ such that :

$$(xxii) \quad \delta_i^T \eta_i \geq \bar{r}_9 \parallel \delta_i \parallel^2, \quad \text{for all } i \in I \text{ where } a_i > 1.$$

The proposition follows from (xxii), (xviii) and (i).

Proof of Proposition I.4.K

If $\delta_i^T \gamma_i \geq 0$, $a_i \leq 1$. Otherwise let $\delta_i^T \gamma_i < 0$.

Then,

$$(i) \quad a_i = 1 + \frac{(T(x_i) \parallel \delta_i \parallel^2 - \delta_i^T \gamma_i)}{T(x_i) \parallel \delta_i \parallel^2 + \delta_i^T [N(x_i, p_i)] [N(x_i, p_i)]^T \delta_i}.$$

Now, by H1', the continuity of $\bar{\lambda}(\cdot)$, $\bar{\mu}(\cdot)$ and the boundedness of $\{x_i\}$, there exists a $M' < \infty$ such that :

$$\parallel G_i \parallel \leq M' \text{ for all } i.$$

Then,

$$(ii) \quad -\delta_i^T \gamma_i \leq M' \parallel \delta_i \parallel^2, \text{ for all } i.$$

Also, by the continuity of $T(\cdot)$ and the boundedness of $\{x_i\}$, there exists a $M_1 < \infty$ such that :

$$(iii) \quad T(x_i) \leq M_1, \text{ for all } i.$$

Hence, from (i), (ii), and (iii) we obtain :

$$(iv) \quad a_i \leq 1 + \frac{(M' + M_1) \parallel \delta_i \parallel^2}{T(x_i) \parallel \delta_i \parallel^2 + \delta_i^T [N(x_i, p_i)] [N(x_i, p_i)]^T \delta_i}.$$

But, as shown in the proof of Proposition I.4.J, there exists $r_0 > 0$ such that :

$$(v) \quad T(x_i) \|\delta_i\|^2 + \delta_i^T [N(x_i, p_i)] [N(x_i, p_i)]^T \delta_i \geq r_0 \|\delta_i\|^2$$

for all i where $a_i > 1$.

From (iv) and (v) we obtain :

$$a_i \leq 1 + \frac{(M' + M_1)}{r_0} \triangleq L' < \infty, \text{ for all } i.$$

Proof of Proposition I.4.L

By Proposition I.4.J, there exists $r' > 0$, such that :

$$\delta_i^T \eta_i \geq r' \|\delta_i\|^2 \quad \text{for all } i$$

Then,

$$(i) \quad \frac{\|\eta_i\|^2}{\delta_i^T \eta_i} \leq \frac{\|\eta_i\|^2}{r' \|\delta_i\|^2}.$$

Now, as shown in the proof of Proposition I.4.K,

$$(iii) \quad T(x_i) \leq M_1 < \infty, \text{ for all } i$$

$$(iv) \quad \|\gamma_i\| \leq M' \|\delta_i\|, \quad M' < \infty$$

$$(v) \quad a_i \leq L' < \infty, \text{ for all } i.$$

Also, by H1' and the boundedness of $\{x_i\}$, there exists $K' < \infty$ such that :

$$(vi) \quad \| [N(x_i, p_i)] [N(x_i, p_i)]^T \| \leq K', \quad \text{for all } i.$$

From (i), (ii), (iii), (iv), (v), (vi), we obtain :

$$\frac{\| \eta_i \|^2}{\delta_i^T \eta_i} \leq \frac{[M' + L'(M_1 + K')]}{r'} \triangleq \tilde{K} < \infty, \quad \text{for all } i.$$

Proof of Proposition I.4.M

An examination of the proof of Proposition I.4.B shows that the sole purpose of the assumption that i^* is finite is to ensure the condition

$$(i) \quad T(x_{i^*}) > 0, \text{ so that } p_{i^*}^T B_{i^*} p_{i^*} > 0.$$

Now, as $\bar{x}$ is supposed $\notin D$, condition (i) will be satisfied for all $i \in I$. The rest of the proof is identical to the proof of Proposition I.4.B.

Proof of Proposition I.4.Q

Let $\tilde{r}(x_i)$ be a vector whose components are $g^j(x_i + p_i)$, $j \in I(x^*)$ and $h^j(x_i + p_i)$, $j \in \underline{\ell}$, ordered in the same way as $r(x_i)$.

By Proposition I.4.0, for i sufficiently large, we have :

$$\begin{aligned} g^j(x_i) + g_x^j(x_i) p_i &= 0, & j \in I(x^*) &= \{j \in \underline{m} \mid g^j(x^*) = 0\} \\ h^j(x_i) + h_x^j(x_i) p_i &= 0, & j \in \underline{\ell}. \end{aligned}$$

Hence, by H1' we have : $\| \tilde{r}(x_i) \| \approx \| p_i \|^2$.

Now, from the definition of $\tilde{p}_i$, the definition of $N(x_i, p_i)$,

the fact that $\bar{\lambda}(x_i) \rightarrow \lambda^*$ and Proposition I.4.0, for i sufficiently large,

$$\tilde{p}_i = \arg \min_{p \in \mathbb{R}^n} \{ \|p\| \mid N_i^T p + \tilde{r}(x_i) = 0 \} .$$

Hence,

$$\| \tilde{p}_i \| \approx \| r(\bar{x}_i) \| \approx \| p_i \|^2 .$$

Proposition I.4.R

From H1' and the mean value theorem, for all i and all j , we have :

$$g^j(x_i + p_i + \tilde{p}_i) = g^j(x_i + p_i) + g_x^j(x_i + p_i) \tilde{p}_i + \frac{1}{2} \tilde{p}_i^T \nabla_{xx} g^j(\xi_i^j) \tilde{p}_i$$

where,

$$\xi_i^j \in [x_i + p_i, x_i + p_i + \tilde{p}_i]$$

Now for i sufficiently large, we have :

$$g^j(x_i + p_i) + g_x^j(x_i) \tilde{p}_i = 0, \text{ for all } j \in I(x^*) .$$

$$\text{Since } \| g_x^j(x_i + p_i) - g_x^j(x_i) \| \approx \| p_i \| \text{ and } \| \tilde{p}_i \| \approx \| p_i \|^2 ,$$

we have :

$$| g^j(x_i + p_i + \tilde{p}_i) | \approx \| p_i \|^3$$

for all $j \in I(x^*)$.

A similar result holds for $|h^j(x_i + p_i + \tilde{p}_i)|$, $j \in \underline{\ell}$

so that :

$$(i) \quad \psi(x_i + p_i + \tilde{p}_i) \approx \|p_i\|^3$$

and

$$(ii) \quad |f(x_i + p_i + \tilde{p}_i) - f(x_i) - \langle \nabla f(x_i), p_i + \tilde{p}_i \rangle - \frac{1}{2} \langle p_i, \nabla_{xx} f(x_i) p_i \rangle| \approx \|p_i\|^3.$$

From (i) and (ii) we have :

$$\gamma(x_i + p_i + \tilde{p}_i, c_i) - \gamma(x_i, c_i) = \emptyset(x_i, c_i) + \bar{e}_i$$

where,

$$\begin{aligned} \emptyset(x_i, c_i) &\triangleq \langle \nabla f(x_i), p_i + \tilde{p}_i \rangle + \frac{1}{2} \langle p_i, \nabla_{xx} f(x_i) p_i \rangle - c_i \psi(x_i) \\ &= \emptyset(x_i, p_i, c_i) + \langle \nabla f(x_i), \tilde{p}_i \rangle + \frac{1}{2} \langle p_i, \nabla_{xx} f(x_i) p_i \rangle \end{aligned}$$

and,

$$\bar{e}_i \approx \|p_i\|^3.$$

Proof of Proposition I.4.U

Because H3 is supposed to be satisfied, we can use the same arguments used in the proofs of Lemma 4 and Lemma 5 in Powell (1978a) to see that the following relations are satisfied.

$$(i) \quad \|p_i\| \sim \|r(x_i)\| + \|P_i \nabla f(x_i)\|$$

and

$$(ii) \quad \|x_i - x^*\| \sim \|r(x_i)\| + \|P_i \nabla f(x_i)\|$$

where $x_i \rightarrow x^* \in D$.

From (i) and (ii) we obtain :

$$(iii) \quad \| p_{i+1} \| \sim \| x_{i+1} - x^* \|$$

and

$$(iv) \quad \| x_i - x^* \| \sim \| p_i \| ,$$

Now,

$$(v) \quad \| x_{i+1} - x^* \| \leq \| x_i - x^* \| + \| x_{i+1} - x_i \| .$$

From Proposition I.4.T and Proposition I.4.Q we have:

$$(vi) \quad \| x_{i+1} - x_i \| \approx \| p_i \| .$$

From (iv), (v) and (vi) we obtain :

$$(vii) \quad \| x_{i+1} - x^* \| \approx \| p_i \| .$$

The proposition follows from (vii) and (iii).

Proof of Proposition I.4.X

P4 is sufficient to ensure that the following condition (as shown in the proof of Lemma 4 in Powell (1978a)) is satisfied :

$$(i) \quad \| p_i \| \sim \| r(x_i) \| + \| P_i \nabla f(x_i) \| .$$

Also, as shown in the proof of Proposition 10 of Mayne and Polak (1978), (i) and P4 are sufficient to guarantee the satisfaction of the condition below:

$$(ii) \quad \theta(x_i, p_i, c_i) \approx -(\| r(x_i) \| + \| P_i \nabla f(x_i) \|^2) .$$

Now, an examination of the proofs of Proposition I.4.0, Proposition I.4.Q and Proposition I.4.R shows that they do not rely upon

H3, so these results remain true.

Following the proof of Proposition I.4.T, we arrive at the following inequality:

$$(iii) \quad T_1 \leq (3/8)\theta(x_i, p_i, c_i) + o(\|p_i\|^2)$$

From (ii) and (iii)

$$(iv) \quad T_i \approx -[\|r(x_i)\| + \|P_i \nabla f(x_i)\|^2] + o(\|p_i\|^2) .$$

But from (i),

$$(v) \quad \|p_i\|^2 \approx [\|r(x_i)\| + \|P_i \nabla f(x_i)\|^2] .$$

Since $[\|r(x_i)\| + \|P_i \nabla f(x_i)\|^2] \rightarrow 0$, it is obvious from (iv) and (v) that,

$$T_i \leq 0, \text{ for all } i \text{ sufficiently large.}$$

Hence, Proposition I.4.T remains true. That is, for i large enough the step length is unity. But, P4 and this last condition are sufficient [see Lemma 5 in Powell (1978a)] to show that

$$(vi) \quad \|x_i - x^*\| \sim \|r(x_i)\| + \|P_i \nabla f(x_i)\| .$$

From (i), (vi) and the step length of unity, an examination of the proof of Proposition I.4.U will show that its result remains valid. The proposition follows because Theorem I.4.V itself does not depend on H3.

Proof of Lemma I.4.Y

Let $\bar{H}_{i+1}$ be the matrix below :

$$\bar{H}_{i+1} = H_i - \frac{H_i r_i r_i^T H_i}{r_i^T H_i r_i} + \frac{\beta_i \beta_i^T}{r_i^T \beta_i}$$

where,

$$(i) \quad r_i \text{ lies on the null space of } [N(x_i, p_i)]^T$$

and

$$(ii) \quad \beta_i = G_i r_i + a_i (T(x_i) I r_i + [N(x_i, p_i)] [N(x_i, p_i)]^T r_i)$$

From (i) and (ii),

$$(iii) \quad \beta_i = G_i r_i + a_i T(x_i) r_i.$$

As shown in the proof of Lemma 6 of Powell (1978a), the fact that the Frobenius norm of a matrix A is the square root of the trace of $A^T A$ gives the equation :

$$(iv) \quad \|D_i^{\frac{1}{2}} \bar{H}_{i+1} D_i^{\frac{1}{2}} - I\|_F = \|D_i^{\frac{1}{2}} H_i D_i^{\frac{1}{2}} - I\|_F = u_0 + u_1$$

where,

$$u_0 \triangleq \left\{ 1 - \frac{\beta_i^T D_i \beta_i}{r_i^T \beta_i} \right\}^2 - \left\{ 1 - \frac{r_i^T H_i D_i H_i r_i}{r_i^T H_i r_i} \right\}^2$$

and

$$u_1 \triangleq 2 \left(\frac{(\beta_i^T D_i H_i D_i \beta_i)(r_i^T H_i r_i) - (\beta_i^T D_i H_i r_i)^2}{(r_i^T \beta_i)(r_i^T H_i r_i)} \right) - 2 \left(\frac{(r_i^T H_i D_i H_i D_i H_i r_i)(r_i^T H_i r_i) - (r_i^T H_i D_i H_i r_i)^2}{(r_i^T H_i r_i)^2} \right).$$

Let u_0^1 and u_1^1 denote the first terms of u_0 and u_1 respectively.

Now, as shown in the proof of Lemma 7 in Powell (1978a),

$$(v) \quad D_i G_i r_i = r_i.$$

Then, from (iii), (iv) and (v) we have :

$$u_0^1 = \left(1 - \frac{\beta_i^T r_i + a_i T(x_i) \beta_i^T D_i r_i}{r_i^T \beta_i} \right)^2 = \left(a_i T(x_i) \frac{\beta_i^T D_i r_i}{r_i^T \beta_i} \right)^2.$$

But,

$$\beta_i^T D_i r_i = (r_i^T G_i + a_i T(x_i) r_i^T) D_i r_i = r_i^T r_i + a_i T(x_i) r_i^T D_i r_i.$$

Hence

$$(vi) \quad u_0^1 = \left(\frac{a_i T(x_i) (r_i^T r_i + a_i T(x_i) r_i^T D_i r_i)}{r_i^T \beta_i} \right).$$

Now, as $[N(x_i, p_i)]^T r_i = 0$ and because of H3a (or P4), there exists a $r^* > 0$ such that

$$(vii) \quad r_i^T \beta_i \geq r^* \|r_i\|^2$$

and

$$(viii) \quad r_i^T H_i r_i \geq r^* \|r_i\|^2.$$

But D_i is a bounded semi-definite matrix. Then, from (vi) and (vii) we obtain :

$$(ix) \quad u_0^1 \leq \left(\frac{a_i T(x_i)}{r^*} \right)^2 \bar{M} ; \quad \bar{M} < \infty.$$

Now, from (iii) and (v) we obtain :

$$(x) \quad \beta_i^T D_i H_i D_i \beta_i = (r_i^T + a_i T(x_i) r_i^T D_i) H_i (r_i + a_i T(x_i) D_i r_i)$$

and

$$(xi) \quad \beta_i^T D_i H_i r_i = (r_i^T + a_i T(x_i) r_i^T D_i) H_i r_i.$$

From (iv), (viii), (x) and (xi) we have :

$$(xii) \quad u_1^1 \approx T(x_i)^2.$$

We already know that :

$$(xiii) \quad \|p_i\| \sim \|r(x_i)\| + \|P_i \nabla f(x_i)\|.$$

Define the vector $\tilde{w}_i = [\tilde{\lambda}_i^j, j \in I(x^*); \tilde{\mu}_i^j, j \in \bar{J}]$

as follows :

$$\tilde{w}_i \triangleq \arg \min_w \{ \| [N(x_i, p_i)]^T w + \nabla f(x_i) \| \}.$$

Hence,

$$(xiv) \quad N(x_i, p_i) ([N(x_i, p_i)]^T \tilde{w}_i + \nabla f(x_i)) = 0.$$

As $\psi(x_i) \rightarrow 0$, it is clear that :

$$(xv) \quad \tilde{\lambda}_i \rightarrow \bar{\lambda}(x_i) \quad \text{and} \quad \tilde{\mu}_i \rightarrow \bar{\mu}(x_i).$$

Now, $\tilde{w}_i$ is the least squares solution of (xiv). Then, the projection of $\nabla f(x_i)$ onto the null space of $[N(x_i, p_i)]$ is:

$$(xvi) \quad P_i \nabla f(x_i) = \nabla f(x_i) + [N(x_i, p_i)]^T \tilde{w}_i.$$

From (xv) and (xvi) we obtain :

$$(xvii) \quad \|P_i \nabla f(x_i)\| \sim \|\nabla f(x_i) + g_x^T(x_i) \bar{\lambda}(x_i) + h_x^T(x_i) \bar{\mu}(x_i)\|$$

$$\text{But, } \psi(x_i) \leq \|r(x_i)\|_\infty.$$

so,

$$(xviii) \quad \psi(x_i) \approx \|r(x_i)\|.$$

From (xviii), (xvii), (xiii) and the definition of $T(\cdot)$, we have :

$$(xix) \quad T(x_i) \approx \|p_i\|^2.$$

From (ix), (xii) and (xix) we obtain:

$$(xx) \quad u_0^1 = o(\|p_i\|^3); \quad u_1^1 = o(\|p_i\|^3).$$

Hence, from (iv), (xii) and (xx), we have :

$$(xxi) \quad \left\| D_i^{\frac{1}{2}} \bar{H}_{i+1} D_i^{\frac{1}{2}} - I \right\|_F^2 \leq \left\| D_i^{\frac{1}{2}} H_i D_i^{\frac{1}{2}} - I \right\|_F^2 - \frac{(D_i H_i r_i - r_i)^T H_i (D_i H_i r_i - r_i)}{r_i^T H_i r_i} + o(\|p_i\|^3).$$

Then, to prove the lemma it only remains to show that :

$$(xxii) \quad \|H_{i+1} - \bar{H}_{i+1}\| \approx [\|r(x_i)\| / \|r_i\|].$$

Now, by Proposition I.4.J there exists a $\bar{r}_1 > 0$ such that:

$$(xxiii) \quad \delta_i^T \eta_i \geq \bar{r}_1 \|\delta_i\|^2.$$

As the step length is unity, we have :

$$(xxiv) \quad \delta_i = p_i + \tilde{p}_i.$$

We can also see in the proof of Proposition I.4.Q that it does not depend on H5 so that it remains true, that is,

$$(xxv) \quad \|\tilde{p}_i\| \approx \|p_i\|^2.$$

From (xxiv) we have : $\|\delta_i\| \geq \|p_i\| - \|\tilde{p}_i\|$.

Then, from (xiii),

$$\delta_i^T \eta_i \geq \bar{r}_1 \|p_i\|^2 - 2 \bar{r}_1 \|p_i\| \|\tilde{p}_i\| + \|\tilde{p}_i\|^2.$$

As $p_i \rightarrow 0$, we can use (xxv) to see that :

$$\delta_i^T \eta_i \geq \bar{r}_1 \|p_i\|^2 - o(\|p_i\|^2).$$

Then, there exists a $r'' > 0$ such that :

$$(xxvi) \quad \delta_i^T \eta_i \geq r'' \|p_i\|^2.$$

Let $N_i \triangleq [N(x_i, p_i)]$, for simplicity of notation.

Now,

$$(xxvii) \quad p_i = q_i + r_i$$

where,

$$q_i = -B_i^{-1} N_i (N_i^T B_i^{-1} N_i)^{-1} r(x_i)$$

and

$$r_i = \{B_i^{-1} N_i (N_i^T B_i^{-1} N_i)^{-1} N_i^T B_i^{-1} - B_i^{-1}\} \nabla f(x_i).$$

$$\text{As } q_i \perp r_i, \quad \|p_i\| \geq \|r_i\|$$

then, from (xxvi) we have :

$$(xxviii) \quad \delta_i^T \eta_i \geq r'' \|r_i\|^2, \quad r'' > 0.$$

$$\text{Now, } y^T B_i y \geq y^T H_i y, \quad \forall y \in \mathbb{R}^n.$$

Then, from (xxvii) and (viii) we obtain :

$$(xxix) \quad p_i^T B_i p_i = r_i^T B_i r_i + r(x_i) (N_i^T B_i^{-1} N_i)^{-1} r(x_i) \geq r^* \|r_i\|^2.$$

But,

$$\delta_i^T B_i \delta_i = (p_i^T + \tilde{p}_i^T) B_i (p_i + \tilde{p}_i).$$

or

$$\delta_i^T B_i \delta_i = p_i^T B_i p_i + \tilde{p}_i^T B_i p_i + p_i^T B_i \tilde{p}_i + \tilde{p}_i^T B_i \tilde{p}_i.$$

Then, by H3a (or P4), (xxv) and the convergence $p_i \rightarrow 0$, there exists a $\bar{r}_2 > 0$ such that :

$$(xxx) \quad \delta_i^T B_i \delta_i \geq \bar{r}_2 \|p_i\|^2.$$

$$\text{Now, } \delta_i^T H_i \delta_i = \delta_i^T B_i \delta_i - T(x_i) \|\delta_i\|^2.$$

Then, from (xix), (xxx) and the convergence of $\|\delta_i\| \rightarrow 0$, there exists a $\tilde{r} > 0$ such that :

$$(xxxi) \quad \delta_i^T H_i \delta_i \geq \tilde{r} \|p_i\|^2 \geq \tilde{r} \|r_i\|^2.$$

From (xxxi), (xxviii), (viii) and (vii) we can see that all the denominators that occur in the definitions of $\bar{H}_{i+1}$ and H_{i+1} , are bounded from below by a multiple of $\|r_i\|^2$.

Now, as shown in the proof of Theorem I.4.V, q_i is the minimum norm solution of :

$$N_i^T p + r(x_i) = 0.$$

then,

$$p_i = r_i + e_i^!$$

where,

$$\| e_i^! \| \approx \| r(x_i) \| .$$

Also, by hypothesis, $\| r(x_i) \| \leq \bar{v}_3 \| r_i \|$.

Hence,

$$(xxxii) \quad \left\| \frac{H_i r_i r_i^T H_i}{r_i^T H_i r_i} - \frac{H_i \delta_i \delta_i^T H_i}{\delta_i^T H_i \delta_i} \right\| \approx [\| r(x_i) \| / \| r_i \|] .$$

Now,

$$\beta_i = \eta_i + \bar{e}_i$$

where,

$$\| \bar{e}_i \| \approx \| r(x_i) \| .$$

Hence, from (vii), (xxviii) and the condition $\| r(x_i) \| \leq \bar{v}_3 \| r_i \|$, we have :

$$(xxxiii) \quad \left\| \frac{\beta_i \beta_i^T}{\beta_i^T r_i} - \frac{\eta_i \eta_i^T}{\eta_i^T \delta_i} \right\| \approx [\| r(x_i) \| / \| r_i \|] .$$

From (xxxiii) and (xxxii) it is clear that (xxii) is indeed true and the lemma is proved.

I.5 Numerical Results

I.5.1 Introduction

In this chapter practical versions of the Basic Algorithm are applied to solve the following problems:

Problem A

$$\min_{x \in \mathbb{R}^7} (1/2) x_1 x_3$$

subject to:

$$(x_4 - x_6)^2 + (x_5 - x_7)^2 \geq 4$$

$$x_5 \geq 1$$

$$(x_3 x_4 - x_2 x_5) / (x_2^2 + x_3^2)^{1/2} \geq 1$$

$$[x_1 x_3 - x_3 x_4 + (x_2 - x_1) x_5] / [x_3^2 + (x_2 - x_1)^2]^{1/2} \geq 1$$

$$x_7 \geq 1$$

$$(x_3 x_6 - x_2 x_7) / (x_2^2 + x_3^2)^{1/2} \geq 1$$

$$[x_1 x_3 - x_3 x_6 + (x_2 - x_1) x_7] / [x_3^2 + (x_2 - x_1)^2]^{1/2} \geq 1$$

$$x_1 \geq 0$$

$$x_3 \geq 0$$

Problem B

$$\min_{x \in \mathbb{R}^5} e^{x_1 x_2 x_3 x_4 x_5}$$

subject to:

$$x_1^2 + x_2^2 + x_3^2 + x_4^2 + x_5^2 - 10 = 0$$

$$x_2 x_3 - 5 x_4 x_5 = 0$$

$$x_1^3 + x_2^3 + 1 = 0$$

where,

e^x is the exponential function.

Problem A is due to Powell (1977) and consists in finding the triangle of smallest area that can contain two non-overlapping discs of unit radius and Problem B is given in Powell (1969) as an illustration of a linear convergent method for solving problems with nonlinear equality constraints.

As pointed out by Mayne and Polak (1982) it is possible to employ any heuristic for choosing c_i finitely often without destroying the results obtained. In the application of the Basic Algorithm to the above problems we incorporate in the algorithm the following procedure, suggested in this last referred paper, which has the advantage of permitting high initial values and low final values for c_i . That is:

If $\hat{c}_{i-1} \geq \bar{c}(x_i)$, set $\hat{c}_i = \hat{c}_{i-1}$.

If $\hat{c}_{i-1} \leq \bar{c}(x_i)$, set $\hat{c}_i = \max \{ \hat{c}_{i-1} + \Delta ; \bar{c}(x_i) \}$.

Set $\tilde{c}_i = (1/2)\tilde{c}_{i-1}$.

Set $c_i = \max \{ \hat{c}_i, \tilde{c}_i \}$

Note in this procedure that the problem dependent rule

$$\hat{c}_i = \max \{ c_{i-1} + \Delta ; c(x_i) \} , \quad \Delta > 0$$

replaces

$$\hat{c}_i = \max \{ \delta c_{i-1} ; \bar{c}(x_i) \} , \quad \delta > 1$$

something that clearly does not invalid any of our results.

In the solution of Problem A, two versions of the BFGS update for constrained problems were used. Namely : a slight modification of the method suggested by Powell - the estimates $\bar{\lambda}(x_i), \bar{\mu}(x_i)$ replacing the multipliers of the quadratic sub-problem in the definition

of γ_i - and the method suggested in the last chapter with the additional practical feature of only using the matrix $N_i N_i^T$ in the computation of η_i if $T(x_i)$ is small. In order to emphasize the features of the new update given by I.4.VI, this last method was applied to Problem B.

Two strategies were adopted in relation to the use of the search arc $\alpha p_i + \alpha^2 \tilde{p}_i$. First we delay the computation of $\tilde{p}_i$ until $T(x_i) \leq \varepsilon$ (given small number). Afterwards we compute $\tilde{p}_i$ from the start. Because the test $w_i = 0$ does not affect the rate of convergence (see Proposition I.4.G) and as we are interested in the effects of using $QP_i^{w_i}$ instead of QP_i^0 , this test was not used.

In the results of the next section, the value of w_i is given as the last element of the vector PI and it is underlined whenever different from zero.

The data used was the following :

Problem A

$$x_0 = 3.0; 0.0; 2.0; -1.5; 1.5; 5.0; 0.0; 0.0.$$

$$H_0 = \begin{pmatrix} I & 0 \\ \vdots & \vdots \\ 0.. & 0 \end{pmatrix}, \text{ where } I \text{ is the } (7 \times 7) \text{ identity matrix.}$$

$$c_0 = 100 ; \Delta = 1.0; b = 1.0 \text{ (definition of } \bar{c}(\cdot)).$$

$$\varepsilon = .5 \text{ (test of step 3); } \varepsilon_1 = .1 \text{ (in the update of I.4.VI).}$$

$$\beta = .9 \text{ (step length procedure).}$$

Problem' B

$$x_0 = 1.0; 1.0; 1.0; 1.0; 1.0; 0.0.$$

$$H_0 = \begin{pmatrix} I...0 \\ \vdots \\ 0...0 \end{pmatrix}, \text{ where } I \text{ is the } (5 \times 5) \text{ identity matrix.}$$

c_0 , Δ , b , ϵ , ϵ_1 , β , are the same as in Proposition A.

The variables involved in the numerical results given in the next section are defined below:

ICT = Iteration number; COST = $f(x_i)$;

$X = x_i$; $XI = \psi(x_i)$; $C = c_i$;

GPM(X) = $\bar{\lambda}(x_i)$; (MU(X) = $\mu(x_i)$);

TX = $\min\{T(x_i) ; \epsilon_2\}$

where,

$$\epsilon_2 = \begin{cases} .1 & \text{for Problem A} \\ 1.0 & \text{for Problem B} \end{cases}$$

ATX = $T(x_i)^{\frac{1}{2}}$; PI = p_i ; L11 = $\tilde{p}_i$;

NQACTIV = Number of active constraints at the solution of QP_i .

IQACTIV = Set of active constraints at the solution of QP_i^w

QPLMD = Multipliers at the solution of QP_i^w

THETA = $\theta(x_i, p_i, c_i)$; STEP LENGTH = α_i ;

NEWX = x_{i+1} ; $C_1 = \gamma_i^T \delta_i$; $C_2 = \delta_i^T H_i \delta_i$.

$$C_3 = \begin{cases} \eta_i^T \delta_i, & \text{in Powell's update} \\ \|\delta_i\|^2, & \text{in the update of I.4.VI.} \end{cases}$$

$$C_4 = \delta_i^T N_i N_i^T \delta_i ; D_1 = TX \|\delta_i\|^2$$

$$C_5 = \eta_i^T \delta_i, \text{ in the update of I.4.VI.}$$

DLGR2 = η_i

CONSTR = $h_i^j(x_i)$, $j = 1, \dots, \ell$ (in Problem B).

Note in the data how x_i and H_i both have additional elements, always set equal to zero, as consequence of the inclusion of the variable w in the quadratic sub-problem. They are left out of the update procedure.

The algorithm is stopped when

$$\| p_i \| \leq 10^{-5}$$

at the solution of the quadratic sub-problem. The reason for this test is that (see Proposition I.4.A) :

$$\| p_i \| = 0 \Rightarrow x_i \in D,$$

but better stop rules may be found.

I.5.2 Results

A.1 Basic Algorithm, with the update given by I.4.VI and $\tilde{p}_i$ computed after $T(x_i) \leq \varepsilon$, applied to Problem A

ICT=1

THE 1 BY 9 MATRIX X

3.000000	0.	2.000000	-1.500000	1.500000
5.000000	0.	0.	0.	0.

COST=3.
XI=2.5
C=50.

THE 1 BY 9 MATRIX GPM(X)

.6359913E-03	.1294196E-01	.9184579E-01	.1578301	.4812452E-01
.2190287E-02	.1956819	.2793716E-01	.5609402E-01	0.

TX=.1

NQACTIV=5

ARRAY IJACTV

I=1IJACTV=5

I=2IJACTV=3

I=3IJACTV=4

I=4IJACTV=7

I=5IJACTV=2

I=6IJACTV=1

THE 1 BY 9 MATRIX GPLMD

.1040583	3.237211	5.173400	4.865792	0.814070
0.	5.410520	0.	0.	0.

THE 1 BY 8 MATRIX PI

2.659575	-1.338798	1.561678	1.120902	-1.500000
-1.648329	1.000000	0.	0.	0.

THEIA=-119.9976082422

STEP LENGTH=1.

THE 1 BY 8 MATRIX NEWX

5.659575	-1.338798	3.561878	-3.790983	1.000000
3.351671	1.000000	0.	0.	0.

C1=4.365935958963

C2=13.1173871435

C3=13.1173871435

C4=0.

D1=1.31173871435

IN THE NEW UPDATE SCHEME

PARAM=0.

THE 1 BY 8 MATRIX PLGR2

.8326329	-.5510573E-01	1.284337	-.1311759E-01	-.2998773E-01
-.1956049E-01	.1031159E-01	0.	0.	0.

C5=4.365935958963

THE 8 BY 8 MATRIX H

.7583753	.2592296	.1565875E-01	-.1670465	.0787108E-01
.2382391	-.1448303	0.	0.	0.

.2592296	.8140949	.1420138	.1139329	-.5037432E-01
-.1670434	.1013610	0.	0.	0.

.1565875E-01	.1420138	1.243169	-.1004904	.3457848E-01
.1363463	-.8317540E-01	0.	0.	0.

-.1670465	.1139329	-.1004904	.9306905	.3107150E-01
.1020390	-.6189982E-01	0.	0.	0.

.6787108E-01	-.5037432E-01	.3457848E-01	.3102150E-01	.9863335
-.4536040E-01	.2752935E-01	0.	0.	0.

.2382391	-.1670434	.1363463	.1020390	-.4536040E-01
.8501218	.9093431E-01	0.	0.	0.

-.1448303	.1013610	-.8317540E-01	-.6189982E-01	.2752935E-01
.9093431E-01	.4448288	0.	0.	0.

0.	0.	0.	0.	0.
0.	0.	0.	0.	0.

0.	0.	0.	0.	0.
----	----	----	----	----

ICT=2

THE 1 BY 8 MATRIX X

5.659575	-1.338798	3.561878	-3.790983	1.000000
3.351671	1.000000	0.	0.	0.

COST=10.07935740166

XI=.9130143249954

C=25.

THE 1 BY 9 MATRIX GPM(X)

.2417262E-01	.1521983	.2307496	.4215331	.8456124
.7798805E-01	1.758139	.2711741E-01	.9431471E-01	0.

TX=.1

NQACTIV=5

ARRAY IJACTV

I=1IJACTV=3

I=2IJACTV=5

I=3IJACTV=7

I=4IJACTV=4

I=5IJACTV=1

THE 1 BY 9 MATRIX GPLMD

.5102988	0.	5.107661	1.516135	7.921467
0.	5.209125	0.	0.	0.

```

THE 1 BY 8 MATRIX PI
1.491300 -2.2535105 .1451352 .5342575 .7251122
-1.7750445 .5133571E-11
THEIA=-19.0462170445
STEP LENGTH=1.

THE 1 BY 8 MATRIX NEAX
7.550475 -2.537403 3.767056 .1551593 1.785112
2.555525 1.000000 0.

C1=.513380554734
C2=3.003155153175
C3=5.193701794272
C4=0.
D1=.5193701794272
IN THE NEW UPDATE SCHEME
PARAM=0.

THE 1 BY 8 MATRIX DLG92
.2259235 -3.3510257E-01 .6268726 -.9044141E-01 -.3372433E-01
-3.3111382E-01 .5442247E-01 0.
C5=.513380554734

THE 3 BY 3 MATRIX H
.4527355 .5714433E-01 .2938157 -.2405584 -.5241257
.2555053 .4754333E-01 0. .4754333E-01
.5714433E-01 .7474255 0. .1003534 .1004791 -.1155474
-.1427403 .1551153 0. 1.882317 -.1443556 -.1235710E-01
.2938157 .1217576E-01 0. 1.882317 -.1443556 -.1235710E-01
-.2405584 .1003534 -.1443556 .4420110 .4420110E-01
.1144831 -.6441023E-01 0. .1235710E-01 .4493733E-02 .6737206
-.3211550 -.1354708 0. .7883456E-01 .1144831 .4231470E-02
.4231470E-02 -.1427403 0. 1.882317 -.1443556 -.1235710E-01
.2655053 .5359317E-01 0. .1217576E-01 .4420110 0.
.347389 .1551153 0. 0. 0. 0.
.4954333E-01 .6054215 0. 0. 0. 0.
.6369317E-01 0. 0. 0. 0. 0.
0. 0. 0. 0. 0. 0.
*****
ICT=3

THE 1 BY 3 MATRIX X
7.550475 -2.537403 3.767056 .1551593 1.785112
2.555525 1.000000 0.
COST=13.99577564592
XI=.1422431401388
C=16.84466584422

THE 1 BY 9 MATRIX GPM(X)
.3100918 .1944531 2.904380 2.441242 4.967722
.151974E-01 4.935051 .1174374E-01 .2325190E-01

TX=.1
NOACTIV=5
APRAY IOACTV
I=1 IOACTV=3
I=2 IOACTV=7
I=3 IOACTV=5
I=4 IOACTV=1
I=5 IOACTV=4

THE 1 BY 9 MATRIX QPLMD
.2939040 0. 2.740232 1.820473 5.936994
0. 5.853063 0. 0. 0.

THE 1 BY 8 MATRIX PI
-1.064560 .5453514 -.4803155E-01 .1850431 .5367391E-01
-1.3517456 -.4437857E-01 .4437857E-01
THEIA=-3.794542581023
STEP LENGTH=1.

THE 1 BY 9 MATRIX NEAX
6.496214 -1.251057 3.659064 .3412024 1.839714
2.204551 .7551213 0.

C1=-.07411638444211
C2=1.514421155932
C3=2.015341096654
C4=0.
D1=.2015341096654
THE VECTOR NNT*PI WAS SET TO ZERO
IN THE NEW UPDATE SCHEME
PARAM=2.367760993735

THE 1 BY 3 MATRIX DLG92
-.4548583 .1592511 -.6044832 -.3417321 .3707970
.3854021 -.8425450E-01 0.
C5=.4030682193307

THE 8 BY 8 MATRIX H
.7304124 .1255041 .3246933 .3401209 -.8591307
-.6654693 .1496423 0. .1496423
.1255041 .5655180 0. .1411511E-01 -.2276417 -.1324477
.3670673 .1177326 0. 2.692323 .4432883 -.5145779
.9246933 .1411511E-01 0. .4432883 1.070270 -.3725130
-.3401209 -.2276417 0. .5145078 -.5725130 .4543578
-.1572498 .3409495E-02 0. .9731626 -.1572496 .5373677
-.5541807 -.1326677 0. .1419820 .3409495E-02 .7144213E-01
-.6654693 .3670673 0. 0. 0.
1.385771 -.7373223E-01 0. 0. 0.
.1406425 .1177326 0. 0. 0.
-.5373223E-01 .7251064 0. 0. 0.
0. 0. 0. 0. 0.
0. 0. 0. 0. 0.

```

ICT=4

THE 1 BY 8 MATRIX X			
6.436214	-1.271057	3.657084	.3412024
2.204881	.5551213		1.332714

COST=11.86573779609
X1=.09228719125116
C=18.36921977094

THE 1 BY 9 MATRIX GPM(X)			
.5112221	-.9941221E-01	3.601017	2.300742
-.2795574E-01	5.063049	.1114990E-01	-.6370237E-02

TX=.1

NQACTIV=5
ARRAY IQACTV
I=1IQACTV=1
I=2IQACTV=3
I=3IQACTV=5
I=4IQACTV=7
I=5IQACTV=4

THE 1 BY 9 MATRIX CPLMD			
.8014471	0.	4.136616	2.396470
0.	5.049473	0.	5.985798

THE 1 BY 8 MATRIX PI			
-1.025501	.6335692	.3812145	.3086120
.1398239	.2124439E-01	.2303378E-01	.2201273

THE TA=-1.717234866255
STEP LENGTH=1.

THE 1 BY 8 MATRIX NEWX			
5.460713	-.5174877	4.040279	.6498144
2.344705	.9763662	0.	2.059943

C1=.5419022320374
C2=.8623694394397
C3=1.762086074521
C4=0.

O1=.1762086074521
IN THE NEW UPDATE SCHEME
PARAM=0.

THE 1 BY 8 MATRIX DLGR2			
-.4453434	-.4286006E-01	-.7936865E-01	-.5863152E-01
.8457520	-.1222265	0.	.2147073

C5=.5419022320374

THE 8 BY 8 MATRIX H			
.8258303	.2679645	.9350490	.3417825
-.8955771	.2326392	0.	-.5057710
.2679645	.5245037	.2350286E-02	-.2784625
.1186395	.1264936	0.	-.2784625
.9350490	.2850286E-02	2.696515	-.5004573
-1.173957	.1594997	0.	.4624069
.3417825	-.2008165	.4624069	1.067903
-1.1705875	.1935772E-01	0.	-.3330124
-.5050010	-.2734625	-.6004573	.5677359
.4052895	.2035201E-01	0.	.3330124
-.3965771	.1186395	-1.173957	-.1705875
1.910709	-.2484658	0.	.4067925
.2526392	.1254936	.1594997	.1935772E-01
-.2484658	.9506549	0.	.2035201E-01
0.	0.	0.	0.
0.	0.	0.	0.

ICT=5

THE 1 BY 8 MATRIX X			
5.460713	-.5174877	4.040279	.6498144
2.344705	.9763662	0.	2.059943

COST=11.03140235551
X1=.08317207257105
C=18.36921977094

THE 1 BY 9 MATRIX GPM(X)			
.7314849	-.2902227E-01	4.007791	2.347918
.1698214E-03	4.426019	.5500959E-02	-.3450443E-02

TX=.03135270345615

NQACTIV=5
ARRAY IQACTV
I=1IQACTV=1
I=2IQACTV=3
I=3IQACTV=5
I=4IQACTV=7
I=5IQACTV=4

THE 1 BY 9 MATRIX CPLMD			
.8547437	0.	4.662884	3.027962
0.	4.385636	0.	5.437794

THE 1 BY 8 MATRIX PI			
-.5751720	.5368723	.3855739	.2928591
.9877103E-01	-.2345793E-01	.4709172E-01	.2331709

THETA=-1.065801421177
PASSED TEST OF STEP 3

THE 1 BY 3 MATRIX L11			
.1951409	-.2325263E-01	.2688520	.0234435E-01
.1530080E-01	.4709172E-01	0.	.2349473E-01

STEP LENGTH=1.

THE 1 BY 8 MATRIX NEWX			
4.980582	-.3953333E-02	4.694705	1.005013
2.458777	1.000000	0.	2.397422

C1=.3996887402345
 C2=.2963419041042
 C3=1.152926711785
 C4=0.
 D1=.03645095904122
 IN THE NEW UPDATE SCHEME
 PARAM=0.

THE 1 BY 9 MATRIX DLGR2
 -.4770255 -.6965043E-C2 .9663934E-02 -.6611004E-02 -.1454282E-01
 .8455481 -.1489188E-C1 0.
 C5=.3996887502345

THE 8 BY 8 MATRIX H
 1.328019 .2978679 1.057939 .3666955 -.4719139
 -1.682206 .2932568 0.3592885E-01 -.2055423 -.2824400
 .2978879 .5232054 0.4281002 .4281002 -.6325691
 .3567120F-G1 .1366740 0.4281002 1.063696 -.3367856
 1.057939 -.3592885E-01 0.6325681 -.3367850 .6645374
 -1.601428 .2194280 0.6325681 -.3367850 .6645374
 .3666955 -.2055423 0.6325681 -.3367850 .6645374
 -.2412482 .2817364E-01 0.6325681 -.3367850 .6645374
 -.4718138 -.2826690 0.6325681 -.3367850 .6645374
 .3228188 .2940934E-01 0.6325681 -.3367850 .6645374
 -1.682206 .3567139E-01 0.6325681 -.3367850 .6645374
 2.955280 -.2724142 0.6325681 -.3367850 .6645374
 .2932568 .1366740 0.6325681 -.3367850 .6645374
 -.2724142 .9454419 0.6325681 -.3367850 .6645374
 0. 0. 0. 0. 0.
 0. 0. 0. 0. 0.
 0. 0. 0. 0. 0.

 ICT=6

THE 1 BY 9 MATRIX X
 4.980682 -.3568039E-02 4.694705 1.005016 2.397400
 2.456777 1.000000 0.
 COST=11.69141622586
 X1=.01436001304852
 C=16.36921577094

THE 1 BY 9 MATRIX GPM(X)
 .8451201 -.3481750E-02 4.061753 3.219658 4.975925
 .5174036E-02 3.606311 .1477066E-02 -.1146835E-02

TX=.0002830611347779

NOACTIV=5
 ARRAY IQACTV
 I=1IQACTV=1
 I=2IQACTV=5
 I=3IQACTV=3
 I=4IQACTV=4
 I=5IQACTV=7

THE 1 BY 9 MATRIX OPLMD
 .8581472 0. 4.762480 3.303591 5.026342
 0. 3.628007 0. 0. 0.

THE 1 BY 8 MATRIX PI
 -.7375261E-01 .3767929E-01 .5566070E-01 .1221084E-01 .5638197E-02
 -.9313288E-02 -.8146195E-12 0.
 THETA=-.2982914806918
 PASSED TEST OF STEP 3

THE 1 BY 8 MATRIX L11
 .7145896E-03 .1254278E-03 .7331419E-03 -.3445119E-04 -.4669394E-04
 -.2100733E-03 .1002155E-11 0.
 STEP LENGTH=1.

THE 1 BY 8 MATRIX NEWX
 4.907644 .3343668E-01 4.751099 1.017144 2.395091
 2.449253 1.000000 0. 0. 0.

C1=.002937236202734
 C2=.004334850752607
 C3=.01022491995494
 C4=0.
 D1=.00000289427745459
 IN THE NEW UPDATE SCHEME
 PARAM=0.

THE 1 BY 8 MATRIX DLGR2
 -.4748187E-01 -.2391246E-01 .1990504E-01 .3676474E-02 -.1079937E-01
 .8114340E-01 -.3332019E-01 0.
 C5=.002937236202734

THE 8 BY 8 MATRIX H
 2.070872 .6636910 .9087176 .3129357 -.3327104
 -2.975721 .8290565 0.8290565 .3129357 -.3327104
 .6636910 .7004530 -.5308227E-01 -.2307100 -.2239447
 -.6096392 .4055557 0.1356824 .4134126 -.4647842
 .9087176 -.5308227E-01 0.1356824 .4134126 -.4647842
 -1.178658 .1393160E-01 0.4134126 1.066996 -.3423530
 .3129357 -.2307100 0.4134126 1.066996 -.3423530
 -.1438616 .1288128E-01 0.4649862 -.3423689 .6567327
 -.3320104 -.2239447 0.4649862 -.3423689 .6567327
 .5039464E-01 .1481445 0.1481445 .5039464E-01 .1481445
 -2.975721 -.6096392 0.1481445 .5039464E-01 .1481445
 5.183515 -1.190819 0.1481445 .5039464E-01 .1481445
 .8290565 .4055557 0.1393160E-01 -.1288128E-01 .1481445
 -1.190819 1.323102 0. 0. 0.
 0. 0. 0. 0. 0.
 0. 0. 0. 0. 0.

ICT=7

```

THE 1 BY 3 MATRIX X
  4.907544      .3343553E-01      4.751099      1.017144      2.394771
  2.449253      1.000000      0.
COST=11.65835070046
XI=.000002044711713722
C=18.36921977034

```

```

THE 1 BY 3 MATRIX GPM(X)
  .8553315      -.1470003E-02      4.736964      3.281922      4.917756
  .2438549E-02      3.522764      .3745774E-03      -.5840043E-03

```

TX=3.071775218827E-7

NDACTIV=5

ARRAY IQACTV

I=1 IQACTV=1

I=2 IQACTV=5

I=3 IQACTV=3

I=4 IQACTV=4

I=5 IQACTV=7

```

THE 1 BY 9 MATRIX QPLMD
  0.8604436      0.      4.774489      3.306349      4.893572
  0.      3.501132      0.      0.

```

```

THE 1 BY 3 MATRIX PI
  -.3865522E-01      .5255179E-02      .3502855E-01      .2609553E-02      .1147739E-01
  -.9532138E-02      .7904433E-10      0.
THETA=-.001150014384485
PASSED TEST OF STEP 3

```

```

THE 1 BY 3 MATRIX L11
  .1841972E-03      .2134044E-04      .1900920E-03      .3955889E-05      -.3095170E-04
  -.3480355E-04      -.7504788E-10      0.
STEP LENGTH=1.

```

```

THE 1 BY 8 MATRIX NEWX
  4.871177      .3521343E-01      4.786317      1.019758      2.407497
  2.440666      1.000000      0.

```

C1=.0005444743509653

C2=.001090541052104

C3=.00280841415127

C4=.6.477704205416E-11

O1=.8.626816994076E-10

IN THE NEW UPDATE SCHEME

PARAM=0.

```

THE 1 BY 8 MATRIX DLGR2
  -.1959873E-01      -.6479498E-02      .1638011E-01      -.3187428E-03      -.3297402E-01
  -.3913517E-01      -.1765055E-03      0.
C5=.3005444743509653

```

```

THE 8 BY 8 MATRIX H
  2.473352      .5701757      .6484888      .3580381      .3595554
  -4.041729      .5595423      0.      .5595423      .3595554
  .5701757      .4033201      .1108117      -.1910745      .2097793
  -.7182380      .1044059      0.      .1044059      .1044059
  -.6484888      .1103117      0.      .1103117      .1103117
  -.3738482      .2976311      0.      .2976311      .2976311
  -.3560381      -.1910745      0.      .3694392      1.063381      -.3245951
  -.2025452      .1497500E-01      0.      .1497500E-01      .1497500E-01
  .3595554      .2092203      -.1472915      -.3245951      2.657271
  -.2336252      .1716918      0.      .1716918      .1716918
  -4.041729      -.7182380      -.3738482      -.2025452      -2.336752
  7.608842      -.9028022      0.      .2976511      .1497600E-01      .1716918
  .5695423      .1044059      0.      0.      0.
  -.9028022      1.0689465      0.      0.      0.
  0.      0.      0.      0.

```

ICT=3

```

THE 1 BY 3 MATRIX X
  4.871177      .3521343E-01      4.746317      1.019758      2.407497
  2.440666      1.000000      0.
COST=11.659749960434
XI=.0059151353331E-7
C=18.36921977034

```

```

THE 1 BY 9 MATRIX GPM(X)
  .8558325      -.3721097E-03      4.774043      3.334380      4.871759
  .1701566E-02      3.466228      .3078639E-03      -.3087198E-03

```

TX=6.626632616615E-8

NDACTIV=5

ARRAY IQACTV

I=1 IQACTV=1

I=2 IQACTV=5

I=3 IQACTV=3

I=4 IQACTV=4

I=5 IQACTV=7

```

THE 1 BY 9 MATRIX QPLMD
  .8553956      0.      4.798975      3.372183      4.861570
  0.      3.447421      0.      0.

```

```

THE 1 BY 3 MATRIX PI
  -.4006327E-01      -.1699733E-01      .3898427E-01      -.8054911E-02      .9174743E-02
  -.1775202E-01      -.11671334E-03      0.

```

THEIA=-.0002274055574513
PASSED TEST OF STEP 3

THE 1 BY 8 MATRIX L11
.22559321E-03 -.16277954E-04 .2493944E-03 .3350721E-04 -.4307302E-04
.174441E-04 .1571339E-06
STEP LENGTH=1.

THE 1 BY 8 MATRIX NEWX
4.831339 .2219777E-01 4.825551 1.011127 2.414471
2.422929 1.000000 0.

C1=.0007267853276313
C2=.0009261265683846
C3=.003888253100323
C4=.0968043920424E-13
D1=2.575622361523E-10
IN THE NEW UPDATE SCHEME
PARAM=0.

THE 1 BY 8 MATRIX DLGR2
-.1567150E-01 -.9750736E-02 .2382568E-01 -.4421832E-03 -.4801089E-01
.3139352E-01 -.795447E-04 0.

C5=.0007287853276313

THE 8 BY 8 MATRIX H
2.759774 .7136323 .2854030 .3270197 1.701744
-4.681388 .5038891 0. .1160515E-01 -.2360801 .5631540
.7136323 .4461759 0. .1160515E-01 -.2360801 .5631540
-1.092810 .1526468 0. 1.829958 .4686876 -2.450113
.2854030 -.1160515E-01 0. .4686876 1.034799 -.4484500
.5481135 .1987313 0. .4686876 1.034799 -.4484500
.3270197 -.2360801 0. .4686876 1.034799 -.4484500
-1.746919 .3988748E-01 0. .4686876 1.034799 -.4484500
1.701344 .5381549 -2.450113 -.4484509 5.019949
-4.264245 .3065323 0. .5481135 -.1946919 -4.264245
-4.581388 -1.072910 0. .1987313 .3988748E-01 .3065323
8.935084 -.9290134 0. 0. 0. 0.
.6038891 .1526848 0. 0. 0. 0.
-.9290134 1.0568812 0. 0. 0. 0.
0. 0. 0. 0. 0. 0.
0. 0. 0. 0. 0. 0.

ICT=9

THE 1 BY 8 MATRIX X
4.831339 .2219777E-01 4.825551 1.011127 2.415521
2.422929 1.000000 0.

COST=11.65693726844
XI=7.984996157609E-7
C=18.36921977094

THE 1 BY 9 MATRIX GPM(X)
.8548791 .3483796E-03 4.820268 3.401258 4.830724
.4534390E-03 3.411445 .1641323E-04 -.1505553E-04

TX=1.925611311433E-9

NQACTIV=5
ARRAY IQACTV
I=1IQACTV=1
I=2IQACTV=5
I=3IQACTV=3
I=4IQACTV=4
I=5IQACTV=7

THE 1 BY 9 MATRIX QPLMD
.8541994 0. 4.817065 3.398399 4.839509
0. 3.425297 0.

THE 1 BY 8 MATRIX PI
-.5320046E-02 -.1583988E-01 .5261936E-02 -.7962710E-02 -.7747635E-03
-.7184926E-02 -.1211831E-03 0.

THEIA=-.0001346413127135
PASSED TEST OF STEP 3

THE 1 BY 8 MATRIX L11
.1163992E-04 -.9849183E-05 .1493100E-04 .1163155E-04 .8208702E-07
.1112233E-04 .1211931E-03 0.

STEP LENGTH=1.

THE 1 BY 8 MATRIX NEWX
4.826031 .6348057E-02 4.830828 1.003175 2.415946
2.415755 1.000000 0.

C1=.0001002195471709
C2=.00013084678819
C3=.0004225212385889
C4=1.34108295408E-12
D1=8.136115763477E-13
IN THE NEW UPDATE SCHEME
PARAM=0.

THE 1 BY 8 MATRIX DLGR2
.1323490E-02 -.2720028E-02 .6588643E-02 -.3216600E-04 -.1318530E-01
-.2648931E-02 -.3608839E-05 0.

C5=.0001002195471709

THE 8 BY 8 MATRIX H
2.568304 .7436156 .2715840 .4076158 1.564799
-4.146390 .5425329 0. .1586235 -.2607617 .9341537
.7436156 .4492128 0. .214454 .5056698 -3.298914
-1.200535 .1721200 0. .5056698 1.003393 -.4587864
.2715840 -.1586235 0. .5056698 1.003393 -.4587864
.6487708 .1689094 0. .5056698 1.003393 -.4587864
.4076158 -.2607617 0. .5056698 1.003393 -.4587864
-1.4146516 .6366155E-01 0. .5056698 1.003393 -.4587864
1.564799 .9341537 -3.298814 -.4587864 6.746306
-4.018132 .3180304 0. .6487708 -.4146616 -4.018132
-4.146890 -1.200535 0. .1639094 .6366155E-01 .3180304
7.453999 -.7618229 0. 0. 0. 0.
.5425329 .1721200 0. 0. 0. 0.
-.7618229 1.0508823 0. 0. 0. 0.
-.7618229 1.0508823 0. 0. 0. 0.
0. 0. 0. 0. 0. 0.
0. 0. 0. 0. 0. 0.

ICT=10

THE 1 BY 8 MATRIX X
4.326031 .6348357E-02 4.330824 1.003175 2.415944
2.415755 1.000000 0.
COST=11.65686243836
XI=3.787466340555E-8
C=18.36921977094

THE 1 BY 9 MATRIX GPM(X)
.9539306 .1847852E-03 4.829547 3.415421 4.825449
.4878529E-04 3.408518 -.1987564E-04 .1992781E-04

TX=1.626742620913E-11

NOACTIV=5
ARRAY IQACTV
I=1IQACTV=1
I=2IQACTV=5
I=3IQACTV=3
I=4IQACTV=4
I=5IQACTV=7

THE 1 BY 9 MATRIX QPLMD
.8538495 0. 4.626536 3.411613 4.830009
0. 3.415772 0.

THE 1 BY 8 MATRIX PI
.1752226E-02 -.6115911E-02 -.1760523E-02 -.3060824E-02 -.1410377E-02
-.1647185E-02 .1367795E-11 0.

THETA=-.00001502592885
PASSED TEST OF STEP 3

THE 1 BY 8 MATRIX L11
.4547473E-12 .1354242E-11 .1364242E-11 .3520477E-06 -.3523591E-06
-.3520436E-06 .3528522E-05 0.

STEP LENGTH=.59049
THE 1 BY 8 MATRIX NEWX
4.827066 .2736673E-02 4.329788 1.001368 2.415013
2.414782 1.000000 0.

C1=.00001445340109676
C2=.00001545224467106
C3=.00005764441813587
C4=1.63557920797E-15
D1=9.378416072299E-16
IN THE NEW UPDATE SCHEME
PARAM=0.

THE 1 BY 8 MATRIX DLGR2
.2405112E-02 -.2574279E-03 .5483333E-03 .1926399E-05 -.1294103E-02
-.4817539E-02 -.2296251E-05 0.

C5=.00001446340109675
THE 8 BY 8 MATRIX H
2.442450 .7015072 .4561303 .4673771 1.101974
-3.745993 .5030262 0. .5030262
-.7015072 .5031469 -.1709559 -.2610644 .9597744
-1.116767 .1723383 0. .1723383
.4561303 -.1709559 2.232317 .4970814 -3.327671
.2603127 .1745163 0. .1745163
.4673771 -.2610644 0. .4970814 .9966731 -.4309548
-.5483285 .5030262E-01 0. .4970814 .9966731 -.4309548
1.101974 .4592044 -3.320671 -.4309548 6.745397
-3.030196 .2398031 0. .2603127 -.5484285 -3.030196
-3.745993 -1.116767 0. .1745163 .5808421E-01 .2999731
6.401505 -.6731047 0. .1745163
.5030262 1.047912 0. .1745163
-.6731047 0. 0. 0.
0. C. 0.
0. C. 0.

ICT=11

THE 1 BY 8 MATRIX X
4.827066 .2736673E-02 4.329788 1.001368 2.415013
2.414782 1.000000 0.
COST=11.6568625284636
XI=4.768551562506E-7
C=18.36921977094

THE 1 BY 9 MATRIX GPM(X)
.8537157 .8849061E-04 4.829259 3.415230 4.824799
.1291900E-04 3.411261 -.1107663E-04 .1108488E-04

TX=1.646039730525E-12

NOACTIV=5
ARRAY IQACTV
I=1IQACTV=1
I=2IQACTV=5
I=3IQACTV=3
I=4IQACTV=4
I=5IQACTV=7

THE 1 BY 9 MATRIX QPLMD
.8535630 0. 4.828228 3.413935 4.829603
0. 3.414394 0.

THE 1 BY 8 MATRIX PI
.1307561E-02 -.2758981E-02 -.1308521E-02 -.1374552E-02 -.7567709E-03
-.5917935E-03 -.1230285E-05 0.

THETA=-.00000020534-2442201
PASSED TEST OF STEP 3

THE 1 BY 8 MATRIX L11

.4547473E-12 .1324242E-11
-.1095134E-06 .1096422E-06
STEP LENGTH=.9

.1364242E-11

.1095179E-06

-.1095391E-05

THE 1 BY 8 MATRIX NEWX

4.828243 .2535900E-03
2.414249 1.000000

4.828611

1.000127

2.414375

C1=.000003394450785914
C2=.000003405877412119
C3=.00001390632153047
C4=.0033044559102E-13
D1=2.292134536844E-17
IN THE NEW UPDATE SCHEME
PARAM=0.

THE 1 BY 8 MATRIX DLGR2

.1343255E-02 -.1421063E-04
-.2687873E-02 -.1291593E-06
C5=.000003394450785914

.3558718E-04

.7504745E-07

-.7012794E-04

THE 8 BY 8 MATRIX H

2.416173 .6893116
-3.634028 .4932533
.6893116 .5041290
-1.092019 .1722325
.5023395 -.1707203
.1461613 .1750345
.4795310 -.2609219
-.5738856 .8923031E-01
.9857139 .9384567
-2.793118 .2483762
-3.684028 -1.042019
6.177134 -.6344756
.4939533 .1722325
-.6544756 1.047765
0. 0.
0. 0.

.5023395

.4795310

.9857139

0.

0.

0.

-.1707263

-.2609219

.9584567

0.

0.

0.

2.230840

.4963838

-3.31431E

0.

.4963838

-.4290357

0.

-.316315

-.4290352

0.

.1661613

-.5738856

0.

.1750346

.6828031E-01

0.

0.

0.

ICT=12

THE 1 BY 3 MATRIX X

4.828243 .2535900E-03
2.414249 1.000000

4.828611

1.000127

2.414375

COST=11.65535188747

XI=4.682115459780E-7

C=18.36921977034

THE 1 BY 9 MATRIX GRV(X)

.9535683 .9770591E-05
-.2755131E-06 3.413852

4.828566

3.414395

4.829213

-.1471218E-05

.1471256E-05

TX=2.271293136467E-13

NQACTIV=5

ARRAY IQACTV

I=1 IQACTV=1

I=2 IQACTV=5

I=3 IQACTV=3

I=4 IQACTV=4

I=5 IQACTV=7

THE 1 BY 9 MATRIX QPLMD

.8535543 0.
0. 3.414235

4.828407

3.414185

4.829447

THE 1 BY 3 MATRIX PI

.1793522E-03 -.2565835E-03
-.3842067E-04 -.1011579E-06

-.1783955E-03

-.1291993E-03

-.3984924E-04

THETA=-.000000254665062737

PASSED TEST OF STEP 3

THE 1 BY 8 MATRIX L11

0. .9094947E-12
.1818989E-11 .1818989E-11
STEP LENGTH=1.

.9094947E-12

.1364242E-11

.1364242E-11

THE 1 BY 8 MATRIX NEWX

4.828422 -.3053444E-05
2.414211 1.000000

4.828432

.9999985

2.414215

C1=3.67176932954E-8

C2=3.683222496379E-8

C3=1.558661001556E-7

C4=2.781942395803E-13

D1=3.54018403263E-20

IN THE NEW UPDATE SCHEME

PARAM=0.

THE 1 BY 8 MATRIX DLGR2

.1538266E-03 .1035573E-04
-.3068550E-03 .3559402E-06
C5=3.67176932954E-8

-.2501023E-04

-.3573200E-06

.5012857E-04

THE 8 BY 8 MATRIX H

2.380470 .6728169
-3.583035 .4841533
.6728169 .5017748
-1.056470 .1713382
.5524713 -.1441338
.5918715E-01 .1773654
.4896914 -.2590957
-.5946784 .6547049E-01
.8541991 .4225042
-2.515332 .2931854
-3.583035 -1.056470
5.913279 -.6344163
.4841533 .1713382
-.6344183 1.047582
0. 0.
0. 0.

.5524713

.4896914

.8541991

0.

-.1641338

-.2549957

0.

0.

.9425047

0.

2.212572

.4939705

0.

0.

-.3.272612

0.

.4939705

-.4234677

0.

-.272612

0.

0.

-.4230622

0.

0.

.5916715E-01

-.5946784

0.

0.

-.2.515332

0.

.1773654

.6847049E-01

0.

0.

0.

ICT=13

THE 1 BY 3 MATRIX
 4.824422 -3093444E-05 4.824422 4.444445 2.414214
 2.414211 1.000000 0.
 COS1=11.0555554214E-9
 X1=7.80021710000E-9
 C=18.36921077094

THE 1 BY 9 MATRIX GPM(X)
 .8535532 .7955138E-07 4.828433 3.414223 4.829422
 -.1865299E-06 3.414207 -.3901063E-07 .3901060E-07

TX=6.089315231524E-17

QUACTIV=5
 ARRAY IQACTV
 I=1 IQACTV=4
 I=2 IQACTV=1
 I=3 IQACTV=5
 I=4 IQACTV=3
 I=5 IQACTV=7

THE 1 BY 9 MATRIX JPLMD
 .8535535 0. 4.828425 3.414211 4.829429
 0. 3.414215 0. 0.

THE 1 BY 8 MATRIX PI
 .4757429E-05 .2592788E-05 -.4742507E-05 .1348159E-05 -.7970977E-05
 .2339581E-05 .4705033E-10 0.
 THETA=-1.065757714479E-7
 PASSED TEST OF STEP 3

THE 1 BY 8 MATRIX L11
 0. .9094947E-12 .1364242E-11 .1364242E-11
 .1814929E-11 .1814989E-11 0.
 STEP LENGTH=1.

THE 1 BY 8 MATRIX NEJA
 4.828427 -.4506551E-06 4.828428 .9994998 2.414214
 2.414213 1.000000 0.

C1=1.191177570625E-11
 C2=1.325376985525E-11
 C3=6.066091171322E-11
 C4=3.4192909973E-16
 D1=3.653334134899E-27
 IN THE NEW UPDATE SCHEME

PARAM=0.

THE 1 BY 8 MATRIX DLGRZ
 .1705965E-05 .1258212E-05 -.3044898E-05 -.9292933E-08 .5095779E-05
 -.3394280E-05 -.4353495E-09 0.
 C5=1.191177570625E-11

THE 8 BY 8 MATRIX H
 2.340116 .0230014 .7380129 .5330595 .3995524
 -3.390160 .4490018 0. .1577782E-01 -.2206790 .5151911
 .5230014 .4490951 0. .1635479 .3379779 -1.935423
 -.8671258 .1429817 0. .2539876 .9875553 -.1969720
 .7380129 .1577782E-01 0. .3879779 .4875553 -.1969720
 -.5536901 .2539876 0. .7458795E-01 .1295322
 .5380598 -.2206790 0. -1.935823 -.1964349 3.564220
 -.7103499 .7458795E-01 0. .5536901 -.7103499 -1.097220
 .3995526 .5151911 -1.935823 0. .2539876 .7458795E-01 .1295322
 -1.090220 .1295322 0. 0. 0. 0.
 -3.390160 -.4490951 -.5536901 0. 0. 0.
 5.263320 -.5507221 0. 0. 0. 0.
 .4490018 .1429817 0. 0. 0. 0.
 -.5507221 1.043264 0. 0. 0. 0.
 0. 0. 0. 0. 0. 0.
 0. 0. 0. 0. 0. 0.

A.2 Basic Algorithm, with the slightly modified Powell's update and $\tilde{p}_i$ computed after $T(x_i) \leq \epsilon$, applied to Problem A.

```

*****
ICT=1
THE 1 BY 8 MATRIX X
  3.000000  0.  2.000000 -1.500000  1.500000
  5.000000  0.  0.  0.  0.
COST=3.
XI=2.5
C=50.

THE 1 BY 9 MATRIX GPM(X)
  .6359913E-03  .1294196E-01  .7184579E-01  .1575361  .4812452E-01
  .2190287E-02  .1856819  .2793716E-01  .6609982E-01

TX=1.

NQACTION=6
ARRAY IQACTV
I=1 IQACTV=5
I=2 IQACTV=3
I=3 IQACTV=4
I=4 IQACTV=7
I=5 IQACTV=2
I=6 IQACTV=1

THE 1 BY 9 MATRIX OPLMD
  .1040683  3.237211  5.173400  4.866792  5.314022
  0.  5.410520  0.  0.

THE 1 BY 8 MATRIX PI
  2.659575 -1.838798  1.561878  1.120902 -1.500000
  -1.648329  1.000000  0.
THETA=-119.9976082422
STEP LENGTH=1.

THE 1 BY 8 MATRIX NEWX
  5.659575 -1.838798  3.561878 -3.790983  1.000000
  3.351671  1.000000  0.

C1=4.365935958963
C2=18.1173871435
IN POWELL UPDATE SCHEME
PARAM=1.

THE 1 BY 8 MATRIX DLGR2
  .8326329 -5.510578E-01  1.284337 -1.1311759E-01 -2.898273E-01
  -1.956049E-01  .1031159E-01  0.
C3=4.365935958963

THE 8 BY 8 MATRIX H
  .7683753  .2592296  .1565875E-01 -1.670465  .6787108E-01
  .2382391 -1.1448303  0.  .1420158  .1139329 -5.037432E-01
  .2592296  .8140949  0.  .1420158  .1139329 -5.037432E-01
  -.1670434  .1013610  0.  1.243169 -1.1004904  .3457848E-01
  .1565875E-01  .1420158  0.  1.243169 -1.1004904  .3457848E-01
  .1363463 -8.317540E-01  0.  -1.004904  .9306905  .3102150E-01
  -.1670465  .1139329  0.  -1.004904  .9306905  .3102150E-01
  .1020390 -6.189982E-01  0.  .3457848E-01  .3102150E-01  .9863035
  .6787108E-01 -5.037432E-01  0.  .3457848E-01  .3102150E-01  .9863035
  -.4536040E-01  .2752935E-01  0.  .1363463  .1020390 -4.536040E-01
  .2382391 -1.670434  0.  .1363463  .1020390 -4.536040E-01
  .8501218  .9093431E-01  0.  -.8317540E-01 -6.189982E-01  .2752935E-01
  -.1448303  .1013610  0.  -.8317540E-01 -6.189982E-01  .2752935E-01
  .9093431E-01  .9448288  0.  0.  0.
  0.  0.  0.  0.
*****
ICT=2
THE 1 BY 8 MATRIX X
  5.659575 -1.838798  3.561878 -3.790983  1.000000
  3.351671  1.000000  0.
COST=10.07935740166
XI=.9130143249954
C=25.

THE 1 BY 9 MATRIX GPM(X)
  .2417262E-01  .1521983  .2807496  .4515331  .8466126
  .7798805E-01  1.758189  .2711741E-01  .9931971E-01

TX=1.

NQACTION=5
ARRAY IQACTV
I=1 IQACTV=3
I=2 IQACTV=5
I=3 IQACTV=7
I=4 IQACTV=4
I=5 IQACTV=1

THE 1 BY 9 MATRIX OPLMD
  .5102888  0.  5.107661  1.516185  7.921467
  0.  9.208125  0.  0.

```

```

THE 1 BY 8 MATRIX PI
1.891300 -2.2586100 .1451682 .5342576 .7851122
-1.7950448 .6295409E-11 0.
THETA=-19.04621704489
STEP LENGTH=1.

THE 1 BY 8 MATRIX NEWX
7.550875 -2.097408 3.707066 .1551593 1.785112
2.556626 1.000000 0.

C1=.6138805587488
C2=3.009095168212
IN POWELL UPDATE SCHEME
PARAM=1.

THE 1 BY 8 MATRIX DLGR2
.2869435 -.3510297E-01 .6268726 -.3094198E-01 -.3379458E-01
-.5811382E-01 .8842247E-01 0.
C3=.6138805587488

THE 8 BY 8 MATRIX H
.4529355 .6714845E-01 .2938167 -.2400584 -.3241950
.2665053 .4954935E-01 0. .1003639 .1069791 -.1954766
.6714845E-01 .7474556 0. 1.882817 -.1943850 -.1235719E-01
.2938167 .1561106 0. 1.882817 -.1943850 -.1235719E-01
.7883456E-01 .1217676E-01 0. .1943850 .9420110 .9993733E-02
-.2406584 .1069791 -.1943850 .9420110 .9993733E-02
.1144831 -.6441023E-01 0. .1235719E-01 .9993733E-02 .6733776
-.3241950 -.1954766 0. .7883456E-01 .1144831 .4231490E-02
.4231490E-02 .1507610 0. .1217676E-01 -.6441023E-01 .1507610
.2665053 -.1420603 0. .7883456E-01 .1144831 .4231490E-02
.8487889 .6359317E-01 0. .1217676E-01 -.6441023E-01 .1507610
.4954935E-01 .1561106

.6369317E-01 .9054615 0. 0. 0.
0. 0. 0. 0. 0.
0. 0. 0. 0. 0.
*****
ICT=3

THE 1 BY 8 MATRIX X
7.550875 -2.097408 3.707066 .1551593 1.785112
2.556626 1.000000 0.
COST=13.99579564597
XI=.1422431401356
C=16.84466584428

THE 1 BY 9 MATRIX GPM(X)
.3100918 .1944631 2.904380 2.441242 4.967222
.6151974E-01 4.935051 .1174379E-01 .2325190E-01

TX=1.
NOACTIV=5
ARRAY IQACTV
I=1 IQACTV=3
I=2 IQACTV=7
I=3 IQACTV=5
I=4 IQACTV=1
I=5 IQACTV=4

THE 1 BY 9 MATRIX QPLMD
.2939040 0. 2.940232 1.820473 5.935394
0. 5.853063 0. 0.

THE 1 BY 8 MATRIX PI
-1.064660 .8463514 -.4800166E-01 .1860431 .5360381E-01
-.3517456 -.4487867E-01 .4487867E-01
THETA=-3.794682581025
STEP LENGTH=1.

THE 1 BY 8 MATRIX NEWX
6.486214 -1.251057 3.659064 .3412024 1.839714
2.204881 .9551213 0.

C1=-.07411638443296
C2=1.514421155979
IN POWELL UPDATE SCHEME
PARAM=.7626744058368

THE 1 BY 8 MATRIX DLGR2
-.2964777 .1138403 -.5433350 -.1768680 .3225630
.3467947 -.5357203E-01 0.
C3=.3028842311958

THE 8 BY 8 MATRIX H
.5073162 .1959134 .6743828 .1276068 -.5570438
-.3443085 .1066626 0. .5022967E-01 -.1582531 -.1584675
.1959134 .5455930 0. 2.760450 .2480695 -.5378335
.2646685 .1312144 0. .2480695 .8838218 -.2469391
.6743828 .5022967E-01 0. .5378335 -.2469391 .9877192
-.7173371 .1112765 0. .7173371 .1365107 .4689144
.1276068 -.1582531 0. .1112765 -.3699491E-01 .920573E-01
.1365107 -.3699491E-01 0. 0. 0.
-.5570638 -.1586475 0. 0. 0.
.4689144 .920573E-01 0. 0. 0.
-.3443085 .2646685 0. 0. 0.
.9336244 .7733104E-02 0. 0. 0.
.1066626 .1312144 0. 0. 0.
.7733104E-02 .9148443 0. 0. 0.
0. 0. 0. 0. 0.
0. 0. 0. 0. 0.

```

ICT=4

THE 1 BY 9 MATRIX X
6.486214 -1.251057 3.659064 .3412024 1.439714
2.204881 .4551213 0.
COST=11.8667779607
X1=.08228719125265
C=18.36921977093

THE 1 BY 9 MATRIX GPM(X)
-.6112221 -.8941221E-01 3.601017 2.300742 5.659302
-.2795574E-01 5.063049 .1114990E-01 -.5370237E-02

TX=.3242121864552

NQACTIV=5
ARRAY IQACTV
I=1 IQACTV=1
I=2 IQACTV=3
I=3 IQACTV=5
I=4 IQACTV=7
I=5 IQACTV=4

THE 1 BY 9 MATRIX QPLMD
.7955867 0. 4.194125 2.232478 6.031408
0. 5.114922 0. 0.

THE 1 BY 8 MATRIX PI
-1.544858 .8614935 .4184449 .3653067 .2671053
1.541217 -.5861787E-02 .5074046E-01
THETA=-2.0487942239
PASSED TEST OF STEP 3

THE 1 BY 8 MATRIX L11
.2008367 .4808073E-02 .4316625 .1020185 .1874566
1.885585E-01 .5074046E-01 0.
STEP LENGTH=1.

THE 1 BY 8 MATRIX NEWX
5.142193 -.3847554 4.509172 .8085276 2.293278
2.377859 1.000000 0.

C1=1.11903978347
C2=1.900622869964
IN POWELL UPDATE SCHEME
PARAM=1.

THE 1 BY 8 MATRIX DLGR2
-.6676656 -.9644810E-01 .8890344E-01 -.6962154E-01 .6085327E-01
1.377317 -.1598248 0.
C3=1.11903978347

THE 8 BY 8 MATRIX H
.8872385 .2689832 .7430834 .1913071 -.5369518
-1.126694 .2150076 0.
-.2689832 .5408314 -.5996011E-01 -.1709157 -.2114064
.2010232 .1340524 0.
.7430834 -.5996011E-01 1.963506 .9618432E-01 -.9055980
-.9451957 .1281244E-01 0.
.1913071 -.1709157 0.
-.1210882E-01 -.4266351E-01 .9618432E-01 .8615124 -.3185496
-.5369518 -.2114064 0.
.3878585 .4362783E-01 -.9055980 -.3185496 .8183555
-1.126694 .1010232 -.9451957 -.1210882E-01 .3878585
2.535594 -.2279848 0.
-.2150076 .1340524 0.
-.2279848 .9285219 0.
0. 0. 0.
0. 0. 0.
0. 0. 0.

ICT=5

THE 1 BY 8 MATRIX X
5.142193 -.3847554 4.509172 .8085276 2.293278
2.377859 1.000000 0.
COST=11.59351649428
X1=.03737459140851
C=19.36921977093

THE 1 BY 9 MATRIX GPM(X)
.7924523 -.1215683E-01 4.497769 3.095386 5.089312
.4781940E-02 3.932037 .3510898E-02 -.1983515E-02

TX=.003596952001138

NQACTIV=5
ARRAY IQACTV
I=1 IQACTV=1
I=2 IQACTV=3
I=3 IQACTV=5
I=4 IQACTV=7
I=5 IQACTV=4

THE 1 BY 9 MATRIX QPLMD
.8541355 0. 4.956991 3.549135 5.114922
0. 3.798319 0. 0.

THE 1 BY 8 MATRIX PI
-.3729530 .3772396 .3335720 .1868062 .1324262
-.3454458E-01 .3431033E-09 0.
THETA=-.7071255623483
PASSED TEST OF STEP 3

THE 1 BY 8 MATRIX L11
.2375516E-01 .3179723E-02 .2569959E-01 .3441591E-02 .1120076E-02
-.1045522E-01 -.3425526E-09 0.
STEP LENGTH=1.

THE 1 BY 8 MATRIX NEWX
.792995 -.4336044E-02 4.868443 .9987753 2.424824
2.401949 1.000000 0.

C1=.09977284340987
C2=.1081777365939
IN POWELL UPDATE SCHEME
PARAM=1.

THE 1 BY 8 MATPIX DLGR2
-.2727913 -.6399045E-01 .8023163E-01 -.3158258E-01 -.5909730E-01
-.5784176 -.7480438E-01 0.
C3=.09977284340987

THE 8 BY 8 MATRIX H
1.633002 .4448214 .5319222 .2783134 -.3797968
-2.702570 .4187810 0. .1993706 -.1576908 -.1261643
.4448214 .5724319 0. .1208606 .5282752E-02 -.5120824
-.3298797 .1900830 0. .5282752E-02 .8662732 -.2645958
.5319222 -.1993706 0. .5120824 -.2645858 .6159791
-1.038388 -.1296628E-01 0. .2769353E-01 -.1298628E-01 .4754970E-01
.2783134 -.1576908 0. .5282752E-02 .8662732 -.2645958
-.2398365 -.1296628E-01 0. .5120824 -.2645858 .6159791
-.3797868 -.1261643 0. .2769353E-01 -.1298628E-01 .4754970E-01
.3457589 -.4754970E-01 0. .5282752E-02 .8662732 -.2645958
-2.702570 -.3298797 0. .5319222 .2783134 -.3797968
5.508458 -.6105256 0. .2769353E-01 -.1298628E-01 .4754970E-01
.4187810 .1900830 0. .5120824 -.2645858 .6159791
-.6105256 .9777353 0. .5282752E-02 .8662732 -.2645958
0. 0. 0. 0. 0.
0. 0. 0. 0. 0.
0. 0. 0. 0. 0.

ICT=6

THE 1 BY 8 MATRIX X
4.792995 -.4336044E-02 4.868443 .9987753 2.424824
2.401949 1.000000 0.
COST=11.66721334661
XI=.0007687487873156
C=19.36921977093

THE 1 BY 9 MATRIX GPM(X)
-.8530271 .7895031E-03 4.871395 3.477742 4.790311
-.9242305E-03 3.356778 -.2503792E-03 .2505658E-03

TX=7.405393128506E-7

NOACTIV=5
ARRAY IQACTV
I=1 IQACTV=1
I=2 IQACTV=5
I=3 IQACTV=3
I=4 IQACTV=7
I=5 IQACTV=4

THE 1 BY 9 MATRIX QPLMD
0.8504182 0. 3.375697 4.853597 3.460133 4.799316
0. 0. 0. 0. 0.

THE 1 BY 8 MATRIX PI
.1675605E-01 -.1399110E-01 -.2131815E-01 -.7908834E-02 -.9811651E-02
.3852302E-03 .2489955E-08 0.
THEIA=-.02519097795758
PASSED TEST OF STEP 3

THE 1 BY 8 MATRIX L11
.5501588E-04 .1081429E-04 .4891350E-04 .7256226E-05 -.1730098E-04
-.3396145E-04 -.2489948E-08 0.
STEP LENGTH=1.

THE 1 BY 8 MATRIX NEWX
4.809806 -.1831632E-01 4.847174 .9908738 2.416995
2.402300 1.000000 0. 0. 0.

C1=.000146613233485
C2=.0002407811915024
IN POWELL UPDATE SCHEME
PARAM=1.

THE 1 BY 8 MATRIX DLGR2
.1163040E-01 .1507213E-02 -.6876712E-02 .9250946E-03 .1601085E-01
-.2680874E-01 -.3917908E-02 0.
C3=.000146613233485

THE 8 BY 8 MATRIX H
2.097320 .2988772 .3943259 .2451445 .7643597
-4.030683 -.3225487E-01 0. .3374136E-01 -.2099124 -.2294786E-01
.2988772 .4341056 -.3374136E-01 0. .168074 .5673420E-01 -1.169749
-.1428432 .6855870E-01 0. .5673420E-01 .8473360 -.1881945
.3943259 -.3374136E-01 0. .168074 .5673420E-01 -1.169749
-.4917303 .3362830 0. .5673420E-01 .8473360 -.1881945
.2451445 -.2099124 0. .168074 .5673420E-01 -1.169749
-.2233272 -.2294786E-01 0. .168074 .5673420E-01 -1.169749
-.7843597 -.4127249 0. .168074 .5673420E-01 -1.169749
-2.397271 -.1428432 0. .168074 .5673420E-01 -1.169749
-4.030683 -.3502396 0. .168074 .5673420E-01 -1.169749
9.019093 .6855870E-01 0. .168074 .5673420E-01 -1.169749
-.3225487E-01 .1039518 0. .168074 .5673420E-01 -1.169749
.3502396 0. 0. 0. 0.
0. 0. 0. 0. 0.
0. 0. 0. 0. 0.

ICT=7

THE 1 BY 8 MATRIX X
4.809806 -.1831632E-01 4.847174 .9908738 2.416995
2.402300 1.000000 0. 0. 0.
COST=11.65698477189
XI=3.421007903626E-7
C=19.36921977093

```

THE 1 BY 9 MATRIX GPM(X)
  .8524201      .1421750E-03      4.852616      3.449723      4.807272
  -.8022805E-03      3.391631      -.1373731E-03      .1372614E-03
TX=3.753932225448E-9
NQACTIV=5
ARRAY IQACTV
I=1IQACTV=1
I=2IQACTV=3
I=3IQACTV=5
I=4IQACTV=7
I=5IQACTV=4

THE 1 BY 9 MATRIX QPLMD
  .8523306      0.      4.846467      3.441559      4.807990
  0.      3.392030      0.      0.

THE 1 BY 8 MATRIX PI
  .1074242E-01      .1835109E-02      -.1086521E-01      .9041997E-03      -.2883530E-02
  .3799237E-02      -.2153833E-10      0.
THETA=-.0001012135014269
PASSED TEST OF STEP 3

THE 1 BY 8 MATRIX L11
  .1702406E-04      .3653520E-05      .1748063E-04      .1233514E-05      -.3442524E-05
  -.1224965E-05      .2153655E-10      0.
STEP LENGTH=1.

THE 1 BY 8 MATRIX NEWX
  4.820566      -.1648085E-01      4.836326      .9917792      2.414109
  2.406098      1.000000      0.

C1=.00004981684087487
C2=.00009550610280919
IN POWELL UPDATE SCHEME
PARAM=1.

THE 1 BY 8 MATRIX DLGR2
  .4929148E-02      .2420474E-02      -.5856226E-02      -.5520223E-04      .1171190E-01
  -.9827637E-02      -.1660001E-04      0.
C3=.00004981684087487

THE 8 BY 8 MATRIX H
  2.561916      .4806131      -.7726176E-01      .2070207      1.864796
  -5.044740      -.1194297E-01      0.
  .4806131      .4074193      -.4883169E-01      -.2941869      .3507155

  -.7244101      .1225977      0.
  -.7726176E-01      -.4883169E-01      1.353335      .2155928      -2.177372
  .8578933      .2358159      0.
  .2070207      -.2941869      .2155928      .8012567      -.3120562
  -.2712854      -.3928101E-01      0.
  1.864796      .3500155      -.2.179372      -.3120562      4.824796
  -4.849163      -.3420945      0.
  -5.044740      -.7244101      .8578933      -.2712854      -4.849163
  10.88279      .3930706      0.
  -.1194297E-01      .1225977      .2358159      -.3928101E-01      -.3420945
  .3930706      1.018677      0.
  0.      0.      0.
  0.      0.      0.
*****
ICT=8

THE 1 BY 8 MATRIX X
  4.820566      -.1648085E-01      4.836326      .9917792      2.414109
  2.406098      1.000000      0.
COST=11.65691512442
XI=8.352124325484E-9
C=19.36921977093

THE 1 BY 9 MATRIX GPM(X)
  .8525543      -.1078857E-03      4.840633      3.432578      4.820842
  -.4896239E-03      3.407481      -.5521306E-04      .5552059E-04

TX=9.317900028927E-10
NQACTIV=5
ARRAY IQACTV
I=1IQACTV=7
I=2IQACTV=1
I=3IQACTV=3
I=4IQACTV=5
I=5IQACTV=4

THE 1 BY 9 MATRIX QPLMD
  .8529872      0.      4.837284      3.429401      4.818519
  0.      3.400493      0.      0.

THE 1 BY 3 MATRIX PI
  .8939749E-02      .8886702E-02      -.9002598E-02      .4418854E-02      -.1324983E-02
  .5742739E-02      .5403482E-09      0.
THETA=-.00008119418978863
PASSED TEST OF STEP 3

THE 1 BY 8 MATRIX L11
  .1352299E-04      -.4468628E-05      .1589124E-04      .5292952E-05      -.2153575E-05
  .6206266E-05      -.5403464E-09      0.
STEP LENGTH=1.

THE 1 BY 8 MATRIX NEWX
  4.829519      -.7598617E-02      4.827340      .9962034      2.412782
  2.411847      1.000000      0.

```

C1=.00006135226174522
C2=.00008084142004401
IN POWELL UPDATE SCHEME
PARAM=1.

THE 1 BY 8 MATRIX DLGR2
-.2267618E-02 .2785824E-02 -.6704629E-02 -.3941053E-04 .1341102E-01
-.4518506E-02 -.5145043E-05 0.
C3=.00006136226174522

THE 8 BY 8 MATRIX H
2.558302 .6634794 -.4733749 .1957691 2.482949
-4.932722 .3398392E-01 0. -.2175938 -.2970230 .8467442
.6634794 .4608303 0. 1.834149 .2032759 -3.435567
-1.184622 .8020163E-01 0. 1.825078 .2032759 .8001844 -.3069266
-.4733749 -.2175938 0. 1.825078 -.2371201 -6.229149
1.825078 .3146420 0. 1.825078 -.2371201 -6.229149
-.1957691 -.2870230 0. 1.825078 -.2371201 -6.229149
-2.371201 -.3411002E-01 0. 1.825078 -.2371201 -6.229149
2.482949 .8467442 0. 1.825078 -.2371201 -6.229149
-6.228149 -.4079230 0. 1.825078 -.2371201 -6.229149
-4.932722 -1.184622 0. 1.825078 -.2371201 -6.229149
10.32504 .2462571 0. 1.825078 -.2371201 -6.229149
.3398392E-01 .8020163E-01 0. 1.825078 -.2371201 -6.229149
.2462571 .9943467 0. 1.825078 -.2371201 -6.229149
0. 0. 0. 0. 0.
0. 0. 0. 0. 0.

ICT=9

THE 1 BY 9 MATRIX X
4.829519 -.7598617E-02 4.827340 .9902034 2.412792
2.411847 1.000000 0.
COST=11.65686486155
X1=0.
C=19.36921977093

THE 1 BY 9 MATRIX GPM(X)
.8530988 -.1755185E-03 4.828999 3.415493 4.830037
-.1045737E-03 3.418306 .1048076E-04 -.1039245E-04
TX=3.203389593899E-11

NQACTIV=5
ARRAY IQACTV
I=1 IQACTV=7
I=2 IQACTV=1
I=3 IQACTV=3
I=4 IQACTV=5
I=5 IQACTV=4

THE 1 BY 9 MATRIX QPLMD
.8533478 0. 4.831542 3.419573 4.824891
0. 3.409337 0. 0. 0.

THE 1 BY 8 MATRIX PI
.3552810E-03 .5881019E-02 -.3616611E-03 .2936013E-02 .7555616E-03
.2181977E-02 .2698428E-09 0.
THETA=-.00001579237099072
PASSED TEST OF STEP 3

THE 1 BY 8 MATRIX L11
.1183424E-05 -.1004682E-05 .1386534E-05 .1381473E-05 .1421998E-06
.8368459E-06 -.2698428E-09 0.
STEP LENGTH=1.

THE 1 BY 8 MATRIX NEWX
4.829876 -.1718603E-02 4.826979 .9991407 2.413538
2.414030 1.000000 0.
C1=.00001228406779662
C2=.00001564507466039
IN POWELL UPDATE SCHEME
PARAM=1.

THE 1 BY 8 MATRIX DLGR2
-.1290407E-02 .6854847E-03 -.1650479E-02 -.2598024E-05 .3303896E-02
.2579071E-02 -.4312901E-06 0.
C3=.00001228406779662

THE 8 BY 8 MATRIX H
1.984212 .6416417 -.3266430 .1768260 1.781981
-3.552438 .1403324 0. 0. 1.056131
.6416417 .4955352 -.3078112 -.2358094
-1.157440 .7265788E-01 0. 2.054905 .2029034 -3.893765
-.3266430 -.3078112 0. 2.029034 .7996646 -.3172085
1.540556 .3186438 0. 2.029034 .7996646 -.3172085
-.1768260 -.2858094 0. 2.029034 .7996646 -.3172085
-1.929533 -.3122978E-01 0. 2.029034 .7996646 -.3172085
1.781981 1.056131 -3.893765 -.3172085 8.298017
-4.711022 -.3549951 0. 2.029034 .7996646 -.3172085
-3.552438 -1.157440 1.540556 -.1929533 -4.711022
7.024463 -.1320721E-02 0. 1.320721E-02 -4.711022
-1.403924 .7265788E-01 .3186438 -.3122978E-01 -.3549951
-.1320721E-02 .9784048 0. 0. 0.
0. 0. 0. 0. 0.
0. 0. 0. 0. 0.

ICT=10

THE 1 BY 8 MATRIX X
4.829876 -.1718603E-02 4.826979 .9991407 2.413538
2.414030 1.000000 0.
COST=11.65685520825
X1=9.665797051639E-10
C=19.36921977093

THE 1 BY 9 MATRIX GPM(X)
.8534516 -.7149446E-04 4.827270 3.412672 4.830089
.6883741E-05 3.416970 .1147707E-04 -.1147634E-04

TX=2.370705388511E-13

NQACTIV=5
 ARRAY IOACTV
 I=1IOACTV=7
 I=2IOACTV=1
 I=3IOACTV=3
 I=4IOACTV=4
 I=5IOACTV=5

THE 1 BY 9 MATRIX OPLMD

0.8535140	0.	4.029010	3.415216	4.827742
	3.413299	0.	0.	

THE 1 BY 8 MATRIX PI

-0.1161978E-02	-0.1695375E-02	0.1160565E-02	.8475759E-03	.5834550E-03
-0.2596713E-03	-0.4791242E-09			

THETA=-0.000001736415902027
 PASSED TEST OF STEP 3

THE 1 BY 8 MATRIX L11

-0.4547474E-12	-0.1364242E-11	0.1364242E-11	.6118821E-07	-0.6112413E-07
-0.6118412E-07	-0.6113022E-07			

STEP LENGTH=.2541865828329

THE 1 BY 8 MATRIX NEWX

4.829580	-0.1287661E-02	4.827274	.9993562	2.413697
2.414096	1.000000	0.		

C1=1.01895175714E-7

C2=1.116365863865E-7

IN POWELL UPDATE SCHEME

PARAM=1.

THE 1 BY 8 MATRIX DLGR2

-0.2553824E-03	-0.1644840E-04	0.3968332E-04	.9768613E-08	-0.7927525E-04
-0.5106430E-03	-0.3102036E-07			

C3=1.01895175714E-7

THE 8 BY 8 MATRIX H

1.614625	.5422165	-0.2610867E-01	.1664710	.9042091
-2.660023	.2124439	0.		
.5422165	.4785972	-0.2584961	-0.2872501	.9187716
-0.9372655	.8269613E-01	0.		
-0.2610867E-01	-0.2584961	1.911896	.2069998	-3.499190
.8788530	.2901483	0.		
.1664710	-0.2872501	.2069998	.7995589	-0.3282301
-0.1706787	-0.3049286E-01	0.		
.9042081	.9189716	-3.498180	-0.3282301	7.212007
-2.792328	-0.2781818	0.		
-2.660023	-0.9372655	.6788530	-0.1706787	-2.792328
4.909991	-0.1564331	0.		
.2124439	.8269613E-01	.2901483	-0.3049286E-01	-0.2781818
-0.1564331	.9732670	0.		
0.	0.	0.	0.	0.
0.	0.	0.	0.	0.

ICT=11

THE 1 BY 8 MATRIX X

4.829580	-0.1287661E-02	4.827274	.9993562	2.413697
2.414096	1.000000	0.		

COST=11.65685472506

XI=2.159131184953E-8

C=19.36921977093

THE 1 BY 9 MATRIX GPM(X)

.8534772	-0.5534636E-04	4.827487	3.412955	4.827746
.6888547E-05	3.416383	.9113134E-05	-0.9113010E-05	

IX=9.503845021976E-14

NQACTIV=5

ARRAY IOACTV

I=1IOACTV=4

I=2IOACTV=1

I=3IOACTV=3

I=4IOACTV=5

I=5IOACTV=7

THE 1 BY 9 MATRIX OPLMD

0.8535492	0.	4.828489	3.414320	4.828357
	3.414117	0.	0.	

THE 1 BY 8 MATRIX PI

-0.1134201E-02	-0.1301776E-02	0.1133229E-02	.6508407E-03	.5227544E-03
.1284711E-03	-0.3829882E-08			

THETA=-0.000001459453583741
 PASSED TEST OF STEP 3

THE 1 BY 8 MATRIX L11

-0.4547473E-12	-0.1364242E-11	0.1364242E-11	.4829285E-07	-0.4825320E-07
-0.4828276E-07	-0.4825738E-07			

STEP LENGTH=.2541865828329

THE 1 BY 8 MATRIX NEWX

4.829292	-0.9567669E-03	4.827563	.9995216	2.413820
2.414128	1.000000	0.		

C1=7.319095104304E-8

C2=7.330537891076E-8

IN POWELL UPDATE SCHEME

PARAM=1.

THE 1 BY 8 MATRIX DLGR2

-0.2268559E-03	-0.2538430E-04	0.6124988E-04	.1201259E-07	-0.1224410E-03
.4536346E-03	-0.2135242E-07			

C3=7.319095104304E-8


```

THE 8 BY 8 MATRIX H
  1.578652      .5306079      .8994601E-02      .1654017      .8029246
 -2.570619      .2207077      0.      .2522602      -.2873803      .9024977
  .5306079      .4763719      -.2522602      0.      .2073240      -3.454297
 -2.9119181      .8369382E-01      0.      1.894690      .7995575      -.3289217
  .8994601E-02      -.2522602      0.      .2073240      .7995575      -.3289217
  .8033413      .2876507      0.      -.3454287      -.3289217      7.103957
  .1654017      -.2873803      0.      .8033413      -.1654160      -2.579279
 -1.1685160      -.3048144E-01      0.      0.      .2876507      -.3048144E-01      -.2729459
  .8028246      .9026932      0.      0.      0.      0.
 -2.578229      -.2728450      0.      0.      0.      0.
 -2.570619      -.9119181      0.      0.      0.      0.
  4.695866      -.1730505      0.      0.      0.      0.
  .2207077      .8369382E-01      0.      0.      0.      0.
 -1.1730505      .9731772      0.      0.      0.      0.
  0.      0.      0.      0.      0.      0.
  0.      0.      0.      0.      0.      0.
*****
ICT=12

THE 1 BY 8 MATRIX X
  4.829292      -.9567669E-03      4.827563      .9995216      2.413920
  2.414128      1.000000      0.      .
COST=11.65685441882
XI=3.405396853395E-8
C=19.36921977093

THE 1 BY 9 MATRIX GPM(X)
  .8534968      -.4132341E-04      4.827720      3.413266      4.829416
  .5329919E-05      3.415838      .6831665E-05      -.6831622E-05

TX=3.74641882628E-14
NQACTION=5
ARRAY IQACTV
I=1 IQACTV=4
I=2 IQACTV=1
I=3 IQACTV=3
I=4 IQACTV=5

I=5 IQACTV=7

THE 1 BY 9 MATRIX OPLMD
  .8535530      0.      4.828433      3.414224      4.829420
  0.      3.414204      0.      0.

THE 1 BY 8 MATRIX PI
  -.8628836E-03      .9581181E-03      .8623734E-03      .4790330E-03      .3937758E-03
  .8623954E-04      -.6021047E-08      0.      .
THETA=-.000001144507904511
PASSED TEST OF STEP 3

THE 1 BY 8 MATRIX L11
  .4547473E-12      .1364242E-11      .1364242E-11      .2729978E-07      -.2728794E-07
  -.2729568E-07      .2728505E-07      0.      .
STEP LENGTH=.3486784401

THE 1 BY 8 MATRIX NEWX
  4.828991      -.6226918E-03      4.827863      .9996887      2.413957
  2.414159      1.000000      0.      .
C1=7.699402587551E-8
C2=7.69885740343E-8
IN POWELL UPDATE SCHEME
PARAM=1.

THE 1 BY 8 MATRIX DLGR2
  -.2339451E-03      -.2767484E-04      .6679119E-04      .5791591E-08      -.1335331E-03
  .4678342E-03      -.2145529E-07      0.      .
C3=7.699402587551E-8

THE 8 BY 8 MATRIX H
  1.574303      .5291250      .1385912E-01      .1653307      .7889947
 -2.559396      .2219024      0.      .2514033      -.2873888      .9004970
  .5291250      .4760806      0.      .2514033      .1892249      .2073445      -3.449713
 -1.9086503      .8384168E-01      0.      1.892249      .2073445      .7995575      -.3289631
  .1385912E-01      -.2514033      0.      .2073445      .7995575      -.3289631      7.089406
  .7928791      .2873039      0.      -.3448213      -.3289631      7.089406
  .1653307      -.2873888      0.      .7928791      -.1683738      -2.549090
 -1.1683738      -.3048141E-01      0.      .2873039      -.3048141E-01      -.2721451
  .7889947      .9004890      0.      0.      0.      0.
 -2.549090      -.2721451      0.      0.      0.      0.
 -2.559396      -.9086503      0.      0.      0.      0.
  4.668359      -.1754437      0.      0.      0.      0.
  .2219024      .8384168E-01      0.      0.      0.      0.
 -1.1754437      .9731754      0.      0.      0.      0.
  0.      0.      0.      0.      0.      0.
  0.      0.      0.      0.      0.      0.
*****
ICT=13

THE 1 BY 8 MATRIX X
  4.828991      -.6226918E-03      4.827863      .9996887      2.413957
  2.414159      1.000000      0.      .
COST=11.65685420418
XI=4.144579435206E-8
C=19.36921977093

THE 1 BY 9 MATRIX GPM(X)
  .8535166      -.2691475E-04      4.827966      3.413595      4.829072
  .3502632E-05      3.415272      .4454284E-05      -.4454269E-05

TX=1.190307308384E-14

```

NOACTIV=5
 ARRAY IQACTV
 I=1IQACTV=4
 I=2IQACTV=1
 I=3IQACTV=3
 I=4IQACTV=5
 I=5IQACTV=7

THE 1 BY 9 MATRIX QPLMD
 0.8535533 0. 3.414213

0.4.828428 0.3.414214 4.828427

THE 1 BY 8 MATRIX PI
 -.5636981E-03 .6226879E-03
 .5513000E-04 -.7199432E-08
 THETA=-8.912185161227E-7
 PASSED TEST OF STEP 3

0.5635295E-03 .3113307E-03 .2567957E-03

THE 1 BY 8 MATRIX L11
 .4547474E-12 .1364242E-11
 .1160897E-07 .1160931E-07
 STEP LENGTH=.6561

0.1364242E-11 .1161306E-07 -.1160521E-07

THE 1 BY 8 MATRIX NEWX
 4.828621 .2141462E-03
 2.414195 1.000000

0.4.828233 .9998929 2.414125

C1=1.157064058728E-7
 C2=1.15690684916E-7
 IN POWELL UPDATE SCHEME
 PARAM=1.

THE 1 BY 8 MATRIX DLGR2
 -.2870629E-03 -.3426180E-04
 .5740868E-03 -.3108194E-07
 C3=1.157064058728E-7

0.8271289E-04 -.5303121E-08 -.1653751E-03

THE 8 BY 8 MATRIX H
 1.574153 .5290193
 -2.558555 .2221748

0.1465777E-01 .1654092 .7869530

.5290193 .4760575
 -.9083743 .8387422E-01
 .1465777E-01 -.2512826
 .7911254 .2872252
 .1654092 .2873794
 -.1685308 .3048142E-01
 .7869530 .9001945
 -2.544693 .2719877
 -2.558555 .9083743
 4.665594 .1759886
 .2221748 .8387422E-01
 -.1759886 .9731754
 0. 0.
 0. 0.
 0. 0.

-.2512826 -.2873794 .9001945
 0.1.891800 .2073218 -3.447187
 0.2073218 .7995575 -.3289177
 0.-3.447187 -.3289177 7.087078
 0.7911254 -.1685308 -2.544693
 0.2872252 -.3048142E-01 -.2719877
 0. 0.
 0. 0.
 0. 0.

 ICT=14

THE 1 BY 8 MATRIX X
 4.828621 .2141462E-03
 2.414195 1.000000
 COST=11.6568540773
 XI=4.328671820986E-8
 C=19.36921977093

0.4.828233 .9998929 2.414125

THE 1 BY 9 MATRIX GPM(X)
 .8535407 .9256256E-05
 .1212034E-05 3.414578

4.828268 3.414000 4.828640
 .1533017E-05 -.1533016E-05

TX=2.630671360932E-15

NOACTIV=5
 ARRAY IQACTV
 I=1IQACTV=4
 I=2IQACTV=1
 I=3IQACTV=3
 I=4IQACTV=5
 I=5IQACTV=7

THE 1 BY 9 MATRIX QPLMD
 0.8535534 0. 3.414214

0.4.828427 0.3.414214 4.828427

THE 1 BY 8 MATRIX PI
 -.1940981E-03 .2141315E-03
 .1884309E-04 -.7502223E-08
 THETA=-6.814031876765E-7
 PASSED TEST OF STEP 3

0.1941471E-03 .1070606E-03 .8822100E-04

THE 1 BY 8 MATRIX L11
 0. .9094947E-12
 .1818989E-11 .1318989E-11
 STEP LENGTH=1.

0.9094947E-12 .1364242E-11 .1364242E-11

THE 1 BY 8 MATRIX NEWX
 4.828427 .1473591E-07
 2.414214 1.000000

0.4.828427 1.000000 2.414214

C1=3.181101110427E-8
C2=3.180964739421E-8
IN POWELL UPDATE SCHEME
PARAM=1.

THE 1 BY 8 MATRIX DLGR2
-1.505923E-03 -1.1804189E-04
-3.012061E-03 -1.108437E-07
C3=3.181101110427E-8
0.4353713E-04 -1.131330E-07 -1.8705492E-04

THE 8 BY 8 MATRIX H
1.574970 .5292801 .1494458E-01 .1655500 .7866042
-2.559981 .2222098 0.2512955 -2.873626 .9002471
-5.292801 .4761082 0.1891566 .2072811 -3.444794
-9.088712 .8387839E-01 0.1891566 .2072811 -3.444794
1.494458E-01 -2.512955 0.2072811 .7495575 -3.288362
7.904917 .2872150 0.2072811 .7495575 -3.288362
1.655500 -2.873626 0.2072811 .7495575 -3.288362
-1.688125 -3.048144E-01 0.2072811 .7495575 -3.288362
7.866042 .9002471 -3.446784 -3.288362 7.086421
-2.543876 -2.719674 0.7904917 -1.688125 -2.543876
-2.559981 -9.088712 0.7904917 -1.688125 -2.543876
4.668032 -1.760587 0.2872150 -3.048144E-01 -2.719674
2.222098 .8387839E-01 0.2872150 -3.048144E-01 -2.719674
-1.760587 .9731754 0.2872150 -3.048144E-01 -2.719674
0. 0. 0. 0. 0.
0. 0. 0. 0. 0.

ICT=15

THE 1 BY 8 MATRIX X
4.828427 -1.473591E-07 4.828427 1.000000 2.414214
2.414214 1.000000 0.
COST=11.65685421449
XI=7.965674342358E-9
C=19.36921977093

THE 1 BY 9 MATRIX GPM(X)
.8535534 -8.153501E-09 4.828427 3.414213 4.828427
.8674761E-09 3.414214 .2464417E-09 -2.2464306E-09

TX=6.345196909843E-17

NQACTIV=5
ARRAY IQACTV
I=1IQACTV=4
I=2IQACTV=1
I=3IQACTV=3
I=4IQACTV=5
I=5IQACTV=7

THE 1 BY 9 MATRIX QPLMD
0.8535534 0. 3.414214 4.828427 3.414214 4.828427
0. 3.414214 0. 0. 0.

THE 1 BY 8 MATRIX PI
-3.408769E-07 .1563026E-07 .4825337E-07 .8347058E-09 .1371201E-07
-1.086501E-07 .6039613E-13 0.
THETA=-1.161325941108E-7
PASSED TEST OF STEP 3

THE 1 BY 8 MATRIX L11
0.9094947E-12 .1364242E-11 .1364242E-11
.1818989E-11 .1818989E-11 0.
STEP LENGTH=1.

THE 1 BY 8 MATRIX NEWX
4.828427 .8952643E-09 4.828427 1.000000 2.414214
2.414214 1.000000 0.

C1=3.048848686567E-16
C2=6.817311419973E-16
IN POWELL UPDATE SCHEME
PARAM=1.

THE 1 BY 8 MATRIX DLGR2
-2.016025E-07 -1.187563E-07 .1942796E-07 -8.302962E-08 -3.226910E-07
.5728954E-07 -1.086264E-08 0.
C3=3.048848686567E-16

THE 8 BY 8 MATRIX H
2.873458 1.294453 -1.043512 .7413102 2.524280
-6.180898 .3322471 0.8685782 .5253010E-01 1.914203
1.294453 .9262901 0.1559589 -5.073530 -2.767872
-3.037219 .1497467 0.1559589 -5.073530 -2.767872
-1.043512 -8.685782 0.1559589 -5.073530 -2.767872
3.279918 -4.719859E-01 0.1559589 -5.073530 -2.767872
7.413102 -5.253010E-01 0.1559589 -5.073530 -2.767872
-1.866223 -3.224167E-01 0.1559589 -5.073530 -2.767872
2.526289 1.914203 0.1559589 -5.073530 -2.767872
-6.584407 .3050194 0.1559589 -5.073530 -2.767872
-6.180898 -3.037219 0.1559589 -5.073530 -2.767872
14.57420 -5.763187 0.1559589 -5.073530 -2.767872
3.322471 .1497467 0.1559589 -5.073530 -2.767872
-5.763187 .9322498 0.1559589 -5.073530 -2.767872
0. 0. 0. 0. 0.
0. 0. 0. 0. 0.

A.3 Basic Algorithm, with the update given by I.4.VI and $\tilde{p}_i$ computed from the start, applied to Problem A.

```

*****
ICT=1

THE 1 BY 8 MATRIX X
3.000000 0. 2.000000 -1.500000 1.500000
5.000000 0. 0. 0. 0.
COST=3.
XI=2.5
C=50.

THE 1 BY 2 MATRIX GPM(X)
.6354913E-03 .1294196E-01 .9184579E-01 .1578361
.2190287E-02 .1856819 .2793716E-01 .6609982E-01 .4812452E-01

TX=.1
NQACTIV=6
ARRAY IQACTV
I=1IQACTV=5
I=2IQACTV=3
I=3IQACTV=4
I=4IQACTV=7
I=5IQACTV=2
I=6IQACTV=1

THE 1 BY 9 MATRIX QPLMO
.1040683 3.237211 5.173400 4.866792 5.814029
0. 5.410520 0. 0. 0.

THE 1 BY 8 MATRIX PI
2.659575 -1.838798 1.561878 1.120902 -.5000000
-1.648329 1.000000 0. 0. 0.
THETA=-119.9976082422
PASSED TEST OF STEP 3

THE 1 BY 8 MATRIX L11
.7407258 -.6057378 .7995119E-01 .4238319 -.2828404E-11
-.3391403 .8633094E-12 0. 0. 0.
STEP LENGTH=1.

THE 1 BY 8 MATRIX NEWX
6.400301 -2.444535 3.641829 .4473366E-01 1.000000
3.012531 1.000000 0. 0. 0.

C1=4.365935958963
C2=18.1173871435
C3=18.1173871435
C4=0.
D1=1.81173871435
IN THE NEW UPDATE SCHEME
PARAM=0.

THE 1 BY 8 MATRIX DLGR2
.8326329 -.5610578E-01 1.284337 -.1311759E-01 -.2898273E-01
-.1958049E-01 .1031159E-01 0. 0. 0.
C5=4.365935958963

THE 8 BY 8 MATRIX H
.7683753 .2592296 .1565875E-01 -.1670465 .6787108E-01
.2382391 -.1448303 0. -.1448303 0.
.2592296 .8140949 0.1420158 .1139329 -.5037432E-01
-.1670434 .1013610 0.1013610 0.1013610 0.
.1565875E-01 .1420158 0.1243169 -.1004904 .3457848E-01
.1363463 -.8317540E-01 0. -.1004904 .9306905 .3102150E-01
-.1670465 .1139329 0. .3457848E-01 .3102150E-01 .9863935
.1020390 -.6189982E-01 0. .3457848E-01 .3102150E-01 .9863935
.6787108E-01 -.5037432E-01 0. .1363463 .1020390 -.4536040E-01
-.4536040E-01 .2752935E-01 0. .8501218 .9093431E-01
.2382391 -.1670434 0. .1448303 .1013610
.8501218 .9093431E-01 .9448288
.9093431E-01 .9448288
0. 0. 0. 0. 0.
0. 0. 0. 0. 0.
0. 0. 0. 0. 0.
*****
ICT=2

THE 1 BY 8 MATRIX X
6.400301 -2.444535 3.641829 .4473366E-01 1.000000
3.012531 1.000000 0. 0. 0.
COST=11.65440060996
XI=.6348404389385
C=25.

THE 1 BY 9 MATRIX GPM(X)
.7078324E-01 .2251396 .7147024 .7723693 1.675365
.1225813 2.615912 .2156885E-01 .9872368E-01 0.

TX=.1
NQACTIV=5
ARRAY IQACTV
I=1IQACTV=3
I=2IQACTV=5
I=3IQACTV=7
I=4IQACTV=4
I=5IQACTV=1

THE 1 BY 9 MATRIX QPLMO
.4734375 0. 3.915526 1.272844 8.470421
0. 9.219083 0. 0. 0.

```

THE 1 BY 8 MATRIX PI
-.5383493 .9151717 .1905429 .2096030E-01 .7817404
-.7890377 .1598721E-11 0.
THETA=-14.18991131472

PASSED TEST OF STEP 3

THE 1 BY 8 MATRIX L11
-.7902512E-01 -.4099904E-01 -.3038035E-02 .3185842 -.1195176
-.4588947E-01 -.2505218E-11 0.
STEP LENGTH=1.

THE 1 BY 8 MATRIX NEWX
6.909625 -1.570363 3.829334 .3842781 1.862223
2.269162 1.000000 0.
C1=-.05867469093131
C2=2.853026946147
C3=2.806834594718
C4=0.
D1=.2806834594718
THE VECTOR NNT*PI WAS SET TO ZERO
IN THE NEW UPDATE SCHEME
PARAM=2.209042210901

THE 1 BY 8 MATRIX DLGR2
.3551515 .1778869 -.2284814E-02 -.1575892 .2014777

C5=.5631618E-02 .1285567 0.
C5=.5613669189436

THE 8 BY 8 MATRIX H
.8802023 .1710177 -.4551771E-01 -.2580522 -.4598907E-01
.3810066 -.5269484E-01 0. .7946003E-01 -.3421673
.1710177 .5133730 .3504480E-01
.8235269E-01 .1613150 0. .9524785E-01 -.7205603E-01
-.4551771E-01 .3504480E-01 1.211566 -.9524785E-01 -.1013678E-01
.2099964 -.7798100E-01 0. .9742599
-.2580522 .7946003E-01 -.9524785E-01 .7045150
-.8973491E-01 -.9882097E-01 -.1013678E-01 .7045150
-.4598907E-02 -.3421673 -.7205603E-01 -.9882097E-01 .9280756E-01
.2032645 .9280756E-01 0. .2099964 .8973491E-01 .2032645
.3810066 .8235269E-01 0. .7798100E-01 -.9882097E-01 .9280756E-01
.6784813 .7889886E-01 0. 0. 0.
-.5269484E-01 .1613150 0. 0. 0.
.7889886E-01 .9732349 0. 0. 0.
0. 0. 0. 0. 0.
0. 0. 0. 0. 0.

ICT=3

THE 1 BY 8 MATRIX X
6.909625 -1.570363 3.829334 .3842781 1.862223
2.269162 1.000000 0.
COST=13.22963047621
X1=.01165113083177
C=19.40176651691

THE 1 BY 9 MATRIX GPM(X)
.6166670 -.1211852 3.786967 2.508866 5.959600
-.3314612E-01 5.359347 .1066547E-01 -.5322962E-02

TX=.1
NQACTION=5
ARRAY IQACTV
I=1IQACTV=3
I=2IQACTV=5
I=3IQACTV=7
I=4IQACTV=4
I=5IQACTV=1

THE 1 BY 9 MATRIX QPLMD
.5182833 0. 4.115408 2.979173 5.545501
0. 4.862857 0. 0.

THE 1 BY 8 MATRIX PI
-.8490120 1.133437 .3638672E-01 .3157533 .1334241
.1761427 -.2444267E-11 0.
THETA=-1.72591842377
PASSED TEST OF STEP 3

THE 1 BY 8 MATRIX L11
.2713175E-01 -.2816190E-01 .6431258E-01 .7092323E-01 .3693754E-01
.4413382E-01 .2595923E-11 0.
STEP LENGTH=1.

THE 1 BY 8 MATRIX NEWX
6.087745 -.4650879 3.930033 .7709546 2.032584
2.489439 1.000000 0.
C1=.5298822042943
C2=1.133074757979
C3=2.155352662461
C4=0.
D1=.2155352662461
IN THE NEW UPDATE SCHEME
PARAM=0.

THE 1 BY 8 MATRIX DLGR2
-.3455678 .8022277E-01 -.4873593 -.1811553 .7165695
.7090257 -.1228928 0.
C5=.5298822042943

THE 8 BY 8 MATRIX H

.8187215	.3360075	.3325711	.1759409	-.6027049
-.1053073	.1380521	0.		
.3363076	.3608911	-.8438614E-01	-.1872537	-.1345984
.2078154	.5992049E-01	0.		
.3325711	-.8438614E-01	1.647159	.5022981E-02	-.7036487
-.4371067	.1181764E-01	0.		
.1759489	-.1872537	.5022981E-02	.6883871	-.1110999

-.1263309	-.1785943	0.		
-.6027049	-.1345984	-.7036487	-.1110999	1.613912
1.151189	-.2295374E-01	0.		
-.1053073	.2078154	-.4371067	-.1263309	1.151189
1.625222	-.7632032E-01	0.		
.1380521	.5992049E-01	.1181764E-01	-.1785943	-.2295374E-01
-.7632032E-01	.9590916	0.		
0.	0.	0.	0.	0.
0.	0.	0.	0.	0.

ICT=4

THE 1 BY 8 MATRIX X

6.087745	-.4650879	3.930033	.7709546	2.032584
2.489439	1.000000	0.		

CDST=11.9625193574
XI=.008508883413974
C=19.40176651691

THE 1 BY 9 MATRIX GPM(X)

.7628100	-.7323941E-01	3.957683	2.187611	5.885488
-.6483730E-03	5.026307	.7810430E-02	-.8627329E-02	

TX=.04259658418871

NOACTIV=5
ARRAY IQACTV
I=1 IQACTV=5
I=2 IQACTV=7
I=3 IQACTV=3
I=4 IQACTV=1
I=5 IQACTV=4

THE 1 BY 9 MATRIX QPLMD

.9276963	0.	4.173570	2.011086	6.023884
0.	4.950167	0.	0.	

THE 1 BY 8 MATRIX PI

-.7358202	.7424481	.3276605	.3488080	.2302870
.2047856	-.3057110E-11	0.		

THEIA=-.6136297404617
PASSED TEST OF STEP 3

THE 1 BY 8 MATRIX L11

.5542504E-01	-.3865991E-02	.8743318E-01	.2848699E-01	.3872207E-01
-.1624492E-01	.3074541E-11	0.		

STEP LENGTH=1.

THE 1 BY 8 MATRIX NEWX

5.407349	.2734942	4.345127	1.148250	2.301593
2.677980	1.000000	0.		

C1=.5587624615129
C2=.4628440773017
C3=1.416658168515
C4=0.
D1=.06034479894176
IN THE NEW UPDATE SCHEME
PARAM=0.

THE 1 BY 8 MATRIX DLGR2

-.5026077	-.2333356E-01	-.6389588E-01	.4263492E-01	.1525986
.8652015	-.1087809	0.		

C5=.5587624615129

THE 8 BY 8 MATRIX H

1.016658	.3120138	.3771084	.7965925E-01	-.2348448
-.4075057	.1342504	0.		
.3120138	.3539043	-.8400757E-01	-.1992885	-.5157169E-01
.2559389	.4547255E-01	0.		
.3771084	-.8400757E-01	1.653807	-.2801644E-02	-.6953871
-.5118128	.1908284E-01	0.		
.7965925E-01	-.1992885	-.2801644E-02	.6784321	.1569405E-01
.4820885E-01	-.2100673	0.		
-.2348448	-.5157169E-01	-.6953871	.1569405E-01	.6516935
.4413615	.1493603	0.		
-.4075067	.2559389	-.5118128	.4820885E-01	.4413615
2.073260	-.5436449E-01	0.		
.1342504	.4547255E-01	.1908284E-01	-.2100673	.1493603
-.5436449E-01	.9396145	0.		
0.	0.	0.	0.	0.
0.	0.	0.	0.	0.

ICT=5

THE 1 BY 8 MATRIX X

5.407349	.2734942	4.345127	1.148250	2.301593
2.677980	1.000000	0.		

CDST=11.74780965811
XI=.005283188767557
C=19.40176651691

THE 1 BY 9 MATRIX GPM(X)

.8519126	-.1070126E-01	4.209586	2.520968	5.371540
.1253737E-01	4.130998	.3479847E-02	-.3791979E-02	

TX=.0009418603033447

NOACTIV=5
 ARRAY IQACTV
 I=1IQACTV=7
 I=2IQACTV=1
 I=3IQACTV=5
 I=4IQACTV=3
 I=5IQACTV=4

THE 1 BY 9 MATRIX QPLMD
 0.8502535 0. 4.268947 2.590177 5.337669
 0. 4.107563 0. 0. 0.

THE 1 BY 8 MATRIX PI
 -.1461889 .6131778E-02 .9996451E-01 .4136691E-03 .3023302E-01
 -.3049536E-01 .3316458E-11 0.
 THETA=-.1498363985099
 PASSED TEST OF STEP 3

THE 1 BY 8 MATRIX L11
 .1939021E-02 .1966966E-03 .2440191E-02 .2036235E-04 -.6305590E-04
 -.6700096E-03 -.3319244E-11 0.
 STEP LENGTH=1.

THE 1 BY 8 MATRIX NEWX
 5.263099 .2798226 4.447532 1.148684 2.331763
 2.640814 1.000000 0. 0. 0.

C1=.007133359851819

C2=.02607901235427

C3=.03364781894243

C4=0.

D1=.000031691544956

IN THE NEW UPDATE SCHEME

PARAM=0.

THE 1 BY 8 MATRIX DLGR2
 -.7048719E-01 -.3395942E-01 .5187791E-01 -.1309948E-01 -.9493980E-01
 .1444859 -.1940500E-01 0.
 C5=.007133359851819

THE 8 BY 8 MATRIX H
 1.320342 .4037778 .2810226 .1541794 .5795007
 -2.038273 .2832049 0.
 .4037778 .3642412 -.7242754E-01 -.1710161 .3235718
 -.5579573 .1112915 0.
 .2810226 -.7242754E-01 1.589354 -.3982560E-01 -1.254567
 .7543083 -.7665953E-01 0.
 .1541794 -.1710161 -.3982560E-01 .6948083 .1727297
 -.2455124 -.1804161 0.
 .5795007 .3235718 -1.254567 .1727297 1.876260
 -1.545634 .3941400 0.
 -2.038273 -.5579573 .7543083 -.2455124 -1.545634
 4.894844 -.4695344 0.
 .2832049 .1112915 -.7665953E-01 -.1804161 .3941400
 -.4695344 .9877400 0.
 0. 0. 0.
 0. 0. 0.
 0. 0. 0.

 ICT=6

THE 1 BY 8 MATRIX X
 5.263099 .2798226 4.447532 1.148684 2.331763
 2.640814 1.000000 0. 0. 0.

COST=11.70390064815

XI=.00002512969418689

C=19.40176651691

THE 1 BY 9 MATRIX GPM(X)
 .4617611 -.6251562E-02 4.335643 2.688395 5.239919
 .1107243E-01 3.945859 .2667746E-02 -.2839148E-02 0.

TX=.0002871301311358

NOACTIV=5
 ARRAY IQACTV
 I=1IQACTV=1
 I=2IQACTV=3
 I=3IQACTV=5
 I=4IQACTV=4
 I=5IQACTV=7

THE 1 BY 9 MATRIX QPLMD
 0.8719921 0. 4.632313 3.125399 4.950502
 0. 3.560703 0. 0. 0.

THE 1 BY 8 MATRIX PI
 -.4561289 -.1772693 .3542149 -.1011190 .1004817
 -.1908173 -.8316903E-11 0.
 THETA=-.08267718518833
 PASSED TEST OF STEP 3

THE 1 BY 8 MATRIX L11
 .2022451E-01 -.1577701E-02 .2602690E-01 .2739759E-02 -.2341833E-02
 -.1249441E-02 .8321593E-11 0.
 STEP LENGTH=1.

THE 1 BY 8 MATRIX NEWX
 4.827195 .1009757 4.827773 1.050304 2.429903
 2.448748 1.000000 0. 0. 0.

C1=.05275372885474
C2=.08218352879094
C3=.4216784798509
C4=0.

D1=.0001210767407818
IN THE NEW UPDATE SCHEME
PARAM=0.

THE 1 BY 8 MATRIX DLGR2
-.1755847 -.8404070E-01
.3457355 -.9744297E-02
C5=.06275372885474

0.2214968 -.3192466E-01 -.4771687
0.

THE 8 BY 8 MATRIX H
1.563804 .4337775
-2.603929 .2807640
.4337775 .3070063
-.6885058 .8320810E-01
-.2323615E-01 -.1079538
1.463333 -.4779558E-01
1.390065 -.2147470
-.2520456 -.1964115
1.393478 .5312965
-3.329975 .3637420
-2.603929 -.6885058
6.148615 -.4426738
-.2607690 .8320810E-01
-.4426738 .9792878
0. 0.
0. 0.

-.2323615E-01 .1390065 1.393478
0. -.1079538 .5312965
0. 1.969606 -.1950454E-01 -2.275496
0. -.1950454E-01 .6664952 .1957788
0. -2.275496 .1957788 4.408907
0. 1.463333 -.2520456 -3.329975
0. -.4779558E-01 -.1964115 .3637420
0. 0.
0. 0.

ICT=7

THE 1 BY 8 MATRIX X
4.827195 .1009757
2.448748 1.000000
COST=11.55230214582
XI=.0009410457052965
C=19.40176651691

4.827773 1.050304 2.429903
0.

THE 1 BY 9 MATRIX GPM(X)
.8592011 .1764041E-02
.1799025E-02 .3.378228

4.804133 3.376544 4.820425
.2427235E-04 -.7591410E-05

TX=.000003359560307014

NOACTIV=5
ARRAY IQACTV
I=1IQACTV=4
I=2IQACTV=3
I=3IQACTV=1
I=4IQACTV=5
I=5IQACTV=7

THE 1 BY 9 MATRIX QPLMD
.8492625 0.
0. 3.415014

4.777388 3.368839 4.827855
0.

THE 1 BY 8 MATRIX PI
-.7755141E-02 -.9473331E-01
-.3522034E-01 -.1771454E-09
THEIA=-.01533556907706
PASSED TEST OF STEP 3

.8966867E-02 -.4770639E-01 -.1230451E-01
0.

THE 1 BY 8 MATRIX L11
.3139187E-03 -.2858654E-03
.2381980E-03 .1771454E-09
STEP LENGTH=1.

.3807240E-03 .3745034E-03 .2585101E-04
0.

THE 1 BY 8 MATRIX NEWX
4.819754 .5956492E-02
2.413765 1.000000

4.837121 1.002973 2.417625
0.

C1=.003282556684255
C2=.00338492851082
C3=.01279272015427
C4=0.

D1=4.294431924597E-8
IN THE NEW UPDATE SCHEME
PARAM=0.

THE 1 BY 8 MATRIX DLGR2
.1952341E-01 -.1188449E-01
-.4089077E-01 -.1489421E-02
C5=.003282556684255

.2870810E-01 .9849785E-03 -.5343474E-01
0.

THE 8 BY 8 MATRIX H
1.617769 .3864847
-2.568523 .2084082
.3864847 .3412306
-.6453163 .1049728
.1241189 -.2030885
1.210565 -.7719261E-01
.1716935 -.2284102
-.3845781 -.1780805
1.195002 .6798454
-3.199264 .4715111
-2.568523 -.6453163
5.409095 -.2291161
-.2084082 .1049728
-.2291161 .9495154
0. 0.
0. 0.

.1241189 .1716935 1.195002
0. -.2030885 -.2284102 .6798454
0. 2.211875 -.7939189E-03 -2.697910
0. -.7939189E-03 .6557101 .1282343
0. -2.697910 .1282343 5.049621
0. 1.210565 -.3845781 -3.199264
0. -.7719261E-01 -.1780805 .4715111
0. 0.
0. 0.

ICT=8

THE 1 BY 8 MATRIX X
4.819754 .5956492E-02 4.837121 1.002973 2.417625
2.413765 1.000000 0.
COST=11.6568664222
XI=.000005341872309828
C=19.40176651691

THE 1 BY 9 MATRIX GPM(X)
.8538983 .3411895E-03 4.836371 3.425149 4.818709
-.1185704E-03 3.399077 -.6743200E-04 .6734725E-04

TX=2.059177614596E-10

NOACTIV=2
ARRAY IQACTV
I=1IQACTV=3
I=2IQACTV=4
I=3IQACTV=1
I=4IQACTV=5
I=5IQACTV=7

THE 1 BY 9 MATRIX QPLMD
.8531645 0. 4.824456 3.409828 4.826665
0. 3.412594 0. 0.

THE 1 BY 8 MATRIX PI
.6552650E-02 -.5866704E-02 -.6586951E-02 -.2927685E-02 -.2748450E-02
-.6450027E-04 -.2701483E-10 0. 0. 0.
THETA=-.0001293988659856
PASSED TEST OF STEP 3

THE 1 BY 8 MATRIX L11
.7874954E-05 .1148147E-05 .7057245E-05 .6157722E-06 -.1057909E-05
-.3704370E-05 .2701484E-10 0. 0. 0.
STEP LENGTH=1.

THE 1 BY 8 MATRIX NEWX
4.826314 .2909355E-03 4.830541 1.000145 2.414875
2.413697 1.000000 0. 0. 0.

C1=.00002891069918569
C2=.00003540421943497
C3=.0001339906182824
C4=.6499805724891E-11
D1=2.759104817329E-14
IN THE NEW UPDATE SCHEME
PARAM=0.

THE 1 BY 8 MATRIX DLGR2
.4687994E-02 .7845185E-03 -.1876184E-02 .1926868E-05 .3755356E-02
-.9396452E-02 -.2648841E-05 0. 0. 0.
C5=.00002891069918569

THE 8 BY 8 MATRIX H
1.928554 .4285740 .4131908 .1446479 .9348516
-2.783475 .1514609 0. 0. 0.
4.285740 .3463951 -.1416173 -.2335401 .6171265
-.6524002 .9419534E-01 0. 0. 0.
4.131908 .1416173 1.550333 .3520001E-01 -1.794205
.9253611E-01 -.2404040E-02 0. 0. 0.
.1446479 -.2335401 .3520001E-01 .6540447 .7557588E-01
-.3055323 -.1815213 0. 0. 0.
9348516 .6171265 -.1.794205 .7557588E-01 3.856643
-1.888829 .3618652 0. 0. 0.
-2.783475 -.6524002 .9253611E-01 -.3055323 -1.888829
4.651831 -.6366394E-01 0. 0. 0.
.1514609 .9419534E-01 -.2404040E-02 -.1815213 .3618652
-.6366394E-01 .9424077 0. 0. 0.
0. 0. 0. 0. 0.
0. 0. 0. 0. 0.

ICT=9

THE 1 BY 8 MATRIX X
4.826314 .2909355E-03 4.830541 1.000145 2.414875
2.413697 1.000000 0. 0. 0.

COST=11.65686513108
XI=1.804320959309E-8
C=19.40176651691

THE 1 BY 9 MATRIX GPM(X)
.8535687 .6143639E-04 4.830624 3.417320 4.826129
-.4864253E-04 3.410902 -.1612593E-04 .1611782E-04

TX=1.488123470635E-13

NOACTIV=5
ARRAY IQACTV
I=1IQACTV=3
I=2IQACTV=4
I=3IQACTV=1
I=4IQACTV=5
I=5IQACTV=7

THE 1 BY 9 MATRIX QPLMD
.8535271 0. 4.828258 3.414086 4.828253
0. 3.414004 0. 0. 0.

THE 1 BY 8 MATRIX PI
.2076281E-02 -.3457449E-03 -.2078844E-02 -.1728203E-03 -.6588595E-03
.4366624E-03 .5794476E-11 0. 0. 0.

THETA=-.000002145377015348
PASSED TEST OF STEP 3

THE 1 BY 8 MATRIX L11
.4547473E-12 .1364242E-11 .1364242E-11 .7677715E-07 -.7684538E-07
-.7677306E-07 .7684947E-07 0.
STEP LENGTH=.2541865828329

THE 1 BY 8 MATRIX NEWX
4.826842 .2030518E-03 4.830013 1.000102 2.414708
2.413821 1.000000 0.

C1=.000001878951998139
C2=.000001854036331212
C3=.000009452879504518
C4=4.504029732403E-16
D1=1.406705185575E-18
IN THE NEW UPDATE SCHEME
PARAM=0.

THE 1 BY 8 MATRIX DLGR2
.1124441E-02 .3950219E-03 -.9530158E-03 -.1917127E-06 .1905529E-02
-.2249723E-02 -.3365996E-06 0.
C5=.000001828951998139

THE 8 BY 8 MATRIX H
2.078837 .4657614 .4187639 .1464842 1.034493
-2.986319 .1248520 0.
.4657614 .3535260 -.1225960 .2328386 .6212092
-.6896109 .8408587E-01 0.
.4187639 -.1225960 1.400264 .3316351E-01 -1.615266
-.2557364E-01 -.2663587E-01 0.
.1464842 -.2328386 0.3316351E-01 .6540377 .7924765E-01
-.3095595 -.1814259 0.
1.034493 .6212092 -1.615266 .7924765E-01 3.718346
-1.894358 .3092070 0.
-2.986319 -.6896109 -.2557364E-01 -.3095595 -1.894358
4.844238 -.5652032E-02 0.
.1248520 .8408587E-01 0.2663587E-01 -.1814259 .3092070
-.5652032E-02 .9411194 0.
0. 0. 0.
0. 0. 0.

ICT=10

THE 1 BY 8 MATRIX X
4.826842 .2030518E-03 4.830013 1.000102 2.414708
2.413821 1.000000 0.
COST=11.65685453521
X1=4.504734363309E-8
C=19.40176651691

THE 1 BY 9 MATRIX GPM(X)
.8535640 .4581454E-04 4.830078 3.416550 4.826704
-.3673465E-04 3.411734 -.1209197E-04 .1208740E-04

IX=6.326348281739E-14

NQACTIV=5
ARRAY IQACTV
I=1IQACTV=3
I=2IQACTV=4
I=3IQACTV=1
I=4IQACTV=5
I=5IQACTV=7

THE 1 BY 9 MATRIX QPLMD
.8535522 0. 4.828419 3.414208 4.828419
0. 3.414205 0. 0.

THE 1 BY 8 MATRIX PI

.1584648E-02 -.2069140E-03 -.1586026E-02 -.1034361E-03 -.4944785E-03
.3913877E-03 -.5022271E-08 0.

THETA=-.000001687000197165
PASSED TEST OF STEP 3

THE 1 BY 8 MATRIX L11
.4547473E-12 .1364242E-11 .1364242E-11 .4324035E-07 -.4326693E-07
-.4323625E-07 .4327102E-07 0.
STEP LENGTH=.282429536481

THE 1 BY 8 MATRIX NEWX
4.827290 .1446132E-03 4.829565 1.000072 2.414568
2.413931 1.000000 0.

C1=.000001055803310212
C2=.000001055832152262
C3=.000005477791398777
C4=3.348584969061E-15
D1=3.465441620338E-19
IN THE NEW UPDATE SCHEME
PARAM=0.

THE 1 BY 8 MATRIX DLGR2
.8439827E-03 .3072017E-03 -.7413136E-03 -.1214648E-06 .1482351E-02
-.1688382E-02 -.1795171E-06 0.
C5=.000001055803310212

```

THE 8 BY 8 MATRIX H
  2.088042      .4681212      .4188904      .1466166      1.041074
-2.998868      .1232354      0.      .1216513      -.2327901      .6218180
.4681212      .3540162      -.1216513      .3304435E-01      -1.606427
-.6921514      .8349518E-01      0.      .3304435E-01      .6540376      .7948358E-01
.4188904      -.1216513      0.      1.392800      .3304435E-01      .6540376
-3.114009E-01      -.2807415E-01      0.      -.1.606427      .7948358E-01      3.712774
.1466166      -.2327901      0.      -.3114009E-01      -.3098265      -1.896959
-3.098265      -.1814254      0.      .2807415E-01      -.1814254      .3063461
1.041074      .6218180      0.      0.      0.      0.
-1.096959      .3063461      0.      0.      0.      0.
-2.998868      -.6921514      0.      0.      0.      0.
.4.857355      -.2404686E-02      0.      0.      0.      0.
.1232354      .8349518E-01      0.      0.      0.      0.
-.2404686E-02      .9411162      0.      0.      0.      0.
0.      0.      0.      0.      0.      0.
0.      0.      0.      0.      0.      0.
*****
ICT=11

THE 1 BY 8 MATRIX X
  4.827290      .1446132E-03      4.829565      1.000072      2.414568
  2.413931      1.000000      0.      0.      0.
COST=11.65685420505
XI=5.930813529176E-8
C=19.40176651691

THE 1 BY 9 MATRIX GPM(X)
  .8535610      .3286680E-04      4.829612      3.415890      4.827191
-2.636825E-04      3.412435      -.8676276E-05      .8673922E-05      0.
TX=2.742492269706E-14
NQACTION=5
ARRAY IQACTV
I=1 IQACTV=3
I=2 IQACTV=4
I=3 IQACTV=1
I=4 IQACTV=5
I=5 IQACTV=7

THE 1 BY 9 MATRIX QPLMD
  .8535533      0.      4.828426      3.414213      4.828427
0.      3.414214      0.      0.      0.

THE 1 BY 8 MATRIX PI
  .1137419E-02      -.1453882E-03      -.1138050E-02      -.7268844E-04      -.3544466E-03
  .2819322E-03      -.7067500E-08      0.      0.      0.
THETA=-.000001377150174527
PASSED TEST OF STEP 3

THE 1 BY 8 MATRIX L11
  .4547473E-12      .1364242E-11      0.1364242E-11      .2221626E-07      -.2222376E-07
-2.221217E-07      .2222785E-07      0.      0.      0.
STEP LENGTH=.43046721

THE 1 BY 8 MATRIX NEWX
  4.827779      .8202835E-04      4.829075      1.000041      2.414415
  2.414053      1.000000      0.      0.      0.
C1=5.435323629669E-7
C2=5.43532648623E-7
C3=.000002820420094491
C4=.6.167239834422E-15
D1=7.734980306464E-20
IN THE NEW UPDATE SCHEME
PARAM=0.

THE 1 BY 8 MATRIX DLGR2
  .6050228E-03      .2207058E-03      -.5326948E-03      -.6627562E-07      .1065278E-02
-1.210213E-02      -.8696735E-07      0.      0.      0.
C5=5.435323629669E-7

THE 8 BY 8 MATRIX H
  2.087881      .4679722      .4190152      .1466086      1.040805
-2.998266      .1231376      0.      .1215424      -.2327930      .6215932
.4679722      .3539355      0.      0.      0.      0.
-.6917879      .8345951E-01      0.      1.392743      .3305139E-01      -1.606297
.4190152      -.1215424      0.      .3305139E-01      .6540376      .7946950E-01
-3.154801E-01      -.2816026E-01      0.      -.3305139E-01      .6540376      .7946950E-01
.1466086      -.2327930      0.      0.      0.      0.
-3.098105      -.1814254      0.      -.1.606297      .7946950E-01      3.712481
1.040805      .6215932      0.      -.3154801E-01      -.3098105      -1.896105
-1.896105      .3061739      0.      .2816026E-01      -.1814254      .3061739
-2.998266      -.6917879      0.      0.      0.      0.
.4.855792      -.2209075E-02      0.      0.      0.      0.
.1231376      .8345951E-01      0.      0.      0.      0.
-.2209075E-02      .9411162      0.      0.      0.      0.
0.      0.      0.      0.      0.      0.
0.      0.      0.      0.      0.      0.
*****
ICT=12

THE 1 BY 8 MATRIX X
  4.827779      .8202835E-04      4.829075      1.000041      2.414415
  2.414053      1.000000      0.      0.      0.
COST=11.65685398707
XI=6.601413105045E-8
C=19.40176651691

THE 1 BY 9 MATRIX GPM(X)
  .8535577      .1871904E-04      4.829102      3.415169      4.827723
-1.1501477E-04      3.413201      -.4940689E-05      .4939926E-05      0.
TX=1.009706788114E-14

```

NOACTIV=5
 ARRAY IQACTV
 I=1 IQACTV=3
 I=2 IQACTV=4
 I=3 IQACTV=1
 I=4 IQACTV=5
 I=5 IQACTV=7

THE 1 BY 9 MATRIX QPLMD
 .8535534 0.
 0. 3.414214

4.828427 3.414213 4.828427
 0. 0.

THE 1 BY 8 MATRIX PI
 .6478233E-03 -.8219761E-04
 .1607202E-03 -.7922353E-08
 THETA=-.000001104589989946
 PASSED TEST OF STEP 3

-.6479246E-03 -.4110373E-04 -.2017723E-03
 0.

THE 1 BY 8 MATRIX L11
 0. .9094947E-12
 .1818989E-11 .1818989E-11
 STEP LENGTH=1.

0. .9094947E-12 .1364242E-11 .1364242E-11

THE 1 BY 8 MATRIX NEWX
 4.828427 -.1692640E-06
 2.414214 1.000000

4.828427 .9999999 2.414214
 0.

C1=1.762429780316E-7
 C2=1.762432967975E-7
 C3=9.144702948522E-7
 C4=8.034923693911E-15
 D1=9.233468642412E-21
 IN THE NEW UPDATE SCHEME
 PARAM=0.

THE 1 BY 8 MATRIX DLGR2
 .3444519E-03 .1256786E-03
 -.6888992E-03 -.2316114E-07
 C5=1.762429780316E-7

-.3034202E-03 -.2509960E-07 .6068472E-03
 0.

THE 8 BY 8 MATRIX H
 .2.088885 .4675840
 -.2.996482 .1231258
 .4675840 .3537848
 -.6910870 .8345522E-01
 .4191730 .1214628
 -.3168151E-01 .2817063E-01
 .1465846 .2328017
 -.3097625 .1814254
 1.039950 .6212370
 -.1.894759 .3061532
 -.2.996482 .6910870
 4.852639 .2185583E-02
 .1231258 .8345522E-01
 -.2185583E-02 .9411162
 0. 0.
 0. 0.
 0. 0.

.4191730 .1465846 1.039950
 0. .1214628 -.2328017 .6212370
 0. 1.393237 .3307256E-01 -1.606810
 0. .3307256E-01 .6540376 .7942718E-01
 0. -1.606810 .7942718E-01 3.712555
 0. -.3168151E-01 -.3097625 -1.894759
 0. .2817063E-01 -.1814254 .3061532
 0. 0.
 0. 0.
 0. 0.

 ICT=13

THE 1 BY 8 MATRIX X
 4.828427 -.1692640E-06
 2.414214 1.000000
 CDS1=11.65685395253
 XI=5.64465594266E-8
 C=19.40176651691

4.828427 .9999999 2.414214
 0.

THE 1 BY 9 MATRIX GPM(X)
 .8535534 -.2038887E-08
 .5748720E-09 3.414214

4.828427 3.414213 4.828427
 .3827694E-09 -.3827790E-09

TX=3.186214112397E-15

NOACTIV=5
 ARRAY IQACTV
 I=1 IQACTV=3
 I=2 IQACTV=4
 I=3 IQACTV=1
 I=4 IQACTV=5
 I=5 IQACTV=7

THE 1 BY 9 MATRIX QPLMD
 .8535534 0.
 0. 3.414214

4.828427 3.414214 4.828427
 0. 0.

THE 1 BY 8 MATRIX PI
 .3491643E-08 .2039578E-06
 .3699299E-07 -.1488800E-09
 THETA=-7.778103499371E-7
 PASSED TEST OF STEP 3

0. .1196344E-06 .1000456E-06 .3410929E-07

THE 1 BY 8 MATRIX L11
 0. .9094947E-12
 .1818989E-11 .1818989E-11
 STEP LENGTH=1.

0. .9094947E-12 .1364242E-11 .1364242E-11

THE 1 BY 8 MATRIX NEWX
 4.828427 .3469473E-07
 2.414214 1.000000

4.828427 1.000000 2.414214
 0.

C1=-2.982203387472E-16
C2=1.643007439316E-14
C3=6.846445420821E-14
C4=1.273049532654E-14
D1=2.181424104507E-28
IN THE NEW UPOATE SCHEME
PARAM=1.023425666567

THE 1 BY 8 MATRIX DLGR2
-.7260772E-08 .4420792E-08 .4671557E-07 .1671349E-06 -.2093269E-06
-.8893390E-07 .1732113E-06 0.
C5=1.273049532654E-14

THE 8 BY 8 MATRIX H
1.575072 .2880930 -.1157107 -.2549826E-01 1.145059
-2.346145 -.4791095E-01 0.
-.2380930 .2946203 -.2795639 -.2010901 .5436485
-.5849045 .1188244 0.
-.1157107 -.2795639 1.064023 .5707770 -2.389017
-.3560767E-01 .5926171 0.
-.2549826E-01 -.2010901 .5707770 2.836884 -2.670888
-1.417899 2.081867 0.
1.145059 .5436485 -2.389017 -2.670888 7.154110
-.4213639 -2.543951 0.
-2.546145 -.5849045 .3560767E-01 -1.417899 -.4213639
5.164414 -1.156266 0.
-.4791095E-01 .1188244 .5926171 2.081867 -2.543951
-1.156266 3.287716 0.
0. 0. 0. 0. 0.
0. 0. 0. 0. 0.

A.4 Basic Algorithm, with the slightly modified Powell's update and $\tilde{p}_i$ computed from the start, applied to Problem A.

```

*****
ICT=1
THE 1 BY 4 MATRIX X
3.000000 0. 2.000000 -1.500000 1.500000
5.000000 0.
COSI=3.
XI=2.5
C=50.

THE 1 BY 9 MATRIX GPM(X)
.4359913E-02 .1244194E-01 .9184579E-01 .1576351
.2190287E-02 .1356919 .2793716E-01 .680944E-01 .4812452E-01
TX=.1
NNOACTIV=5
ARRAY IQACTV
I=1IQACTV=5
I=2IQACTV=3
I=3IQACTV=4
I=4IQACTV=7
I=5IQACTV=2
I=6IQACTV=1

THE 1 BY 9 MATRIX CPLMD
.1040683 3.237211 5.173400 4.460792 5.414020
0. 5.410520

THE 1 BY 9 MATRIX PI
2.659575 -1.838798 1.561878 1.120902 -.5000000
-1.648329 1.000000
THEIA=-119.9976082422
PASSED TEST OF STEP 3

THE 1 BY 9 MATRIX L11
.7407258 -.6057378 .7995119E-01 .4238319 -.2829404E-11
-.3391403 .8533094E-12

THE 1 BY 8 MATRIX NEWX
6.400301 -2.444535 3.641829 .4473366E-01 1.000000
3.012531 1.000000
IN POWELL UPDATE SCHEME
PARAM=1.

THE 8 BY 8 MATRIX H
.7162714 .2902775 .4539468E-01 -.1912537 .5593240E-01
.2383203 -.1201232 0. .1283931 .1353449 -.4360117E-01
.2902775 .7857456 0. .1362172 -.9572393E-01 .1984204E-01
-.1743760 .8773515E-01 0. .1362172 -.9572393E-01 .1984204E-01
.4539468E-01 .1283931 0. .1362172 -.9572393E-01 .1984204E-01
.1046427 -.5509257E-01 0. .1362172 -.9572393E-01 .1984204E-01
.1104337 -.5556575E-01 0. .1362172 -.9572393E-01 .1984204E-01
.5593240E-01 -.4360117E-01 0. .1362172 -.9572393E-01 .1984204E-01
-.3391403 .8533094E-12 0. .1362172 -.9572393E-01 .1984204E-01
.2383203 -.1743760 0. .1362172 -.9572393E-01 .1984204E-01
.8581403 .7137583E-01 0. .1362172 -.9572393E-01 .1984204E-01
-.1201232 .8773515E-01 0. .1362172 -.9572393E-01 .1984204E-01
.7137583E-01 .4040677 0. .1362172 -.9572393E-01 .1984204E-01
0. 0. 0. 0. 0.
0. 0. 0. 0. 0.
*****
ICT=2
THE 1 BY 8 MATRIX X
6.400361 -2.444535 3.641829 .4473366E-01 1.000000
3.012531 1.000000
COSI=11.55440060996
XI=.6344404389385
C=25.

THE 1 BY 9 MATRIX GPM(X)
.7078324E-01 .2251346 .7147024 .7723643 1.675745
.1225813 2.615912 .2156485E-01 .4872308E-01

TX=.1
NNOACTIV=5
ARRAY IQACTV
I=1IQACTV=3
I=2IQACTV=5
I=3IQACTV=7
I=4IQACTV=4
I=5IQACTV=1

THE 1 BY 9 MATRIX CPLMD
.4744780 0. 3.425380 1.292468 4.494224
0. 9.235859 0. 0.

```

THE 1 BY 3 MATRIX PI
 .7131487 .118259
 -.7464879 -.9999979E-11
 THETA=-14.18780852538
 PASSED TEST OF STEP 3

THE 1 BY 4 MATRIX L11
 -.7574240E-01 -.3431601E-01
 .4329303E-01 .4171774E-11

THE 1 BY 4 MATRIX NEWX
 6.217757 -1.372055
 2.269336 1.000000

IN POWELL UPDATE SCHEME
 PARA=7490215425893

THE 3 BY 8 MATRIX H
 .7718444 .2234000
 .3361433 -.7354604E-01
 .2234000 .4778481
 .4371729E-01 .1403899
 .1745257E-01 .3456168E-01
 .1819216 -.4275395E-01
 -.2429784 -.2931357E-02
 .1727952 -.0214320E-01
 .1007403E-01 -.3032784
 .1914506 .6760174E-01
 .3501453 .8371734E-01
 .5953217 .5785330E-01
 -.7364504E-01 .1403899
 .5785330E-01 .9725935
 0. 0.
 0. 0.

 ICT=3

THE 1 BY 9 MATRIX X
 6.437757 -1.572055
 2.269336 1.000000
 COSI=13.24748746571
 XI=.010313947114006
 C=14.41843556267

THE 1 BY 9 MATRIX GPM(X)
 .6157775 -.1236594
 -.3426459E-01 5.378195

TX=.1

NOACTIV=5

APRAY IQACTV
 I=110ACTIV=3
 I=210ACTIV=5
 I=310ACTIV=7
 I=410ACTIV=4
 I=510ACTIV=1

THE 1 BY 9 MATRIX UPLMO
 .5225231 0.
 0. 4.783441

THE 1 BY 8 MATRIX PI
 -1.134622 1.428581
 .2486150 -.4334311E-12
 THETA=-2.044157433543
 PASSED TEST OF STEP 3

THE 1 BY 8 MATRIX L11
 .5143397E-01 -.4075422E-01
 .5384947E-01 .1414202E-11

THE 1 BY 8 MATRIX NEWX
 5.854575 -.1842425
 2.581401 1.000000

IN POWELL UPDATE SCHEME
 PARA=1.

THE 8 BY 8 MATRIX H
 .8418944 .3101270
 -.1342133 .5324939E-01
 .3101270 .3969461
 .1447504 .6402758E-01
 .3371598 -.7005755E-01
 -.2809517 -.4035182E-01
 -.7526426E-01 -.2934163
 -.7528408E-01 -.1741025
 -.4506334 -.1834186
 .3667783 .2062120E-01
 -.1342133 .1447504
 .1757176 -.8727362E-01
 .3324939E-01 .6602758E-01
 -.3727362E-01 .9407907
 0. 0.
 0. 0.

.1771934 .2301908E-01 .9745605
 0.

.8037324E-04 .3130434 -.1172729
 0.

3.819069 .3820876 1.857226
 0.

.1748257E-01 -.2820784 .1007403E-01
 0. .3436168E-01 -.2931357E-02 -.3032784
 0. .1334255 -.1316472 -.6062755E-01
 0. .1316472 .4041726 -.9842566E-01
 0. -.6062755E-01 -.9842566E-01 .7735074
 0. .1819216 .1737957 .1914500
 0. -.4275395E-01 -.6214320E-01 .0760178E-01
 0. 0.
 0. 0.

3.819069 .3820876 1.857226
 0.

3.779344 2.435538 5.974238
 .1067909E-01 -.5478903E-02

4.257335 3.082163 5.541727
 0. 0.

.9188251E-01 .4141682 .1914245
 0.

.1189187 .1133353 .7159172E-01
 0.

4.029871 .9095912 2.120307
 0.

.3371698 .8706426E-01 -.4504334
 0. -.7005755E-01 -.2039163 -.1834186
 0. .1464302 -.1451498 -.3297341
 0. -.1751498 .5497912 -.1135165
 0. -.3290861 -.1135165 1.272227
 0. .2809617 -.7528408E-01 .9667783
 0. -.4035182E-01 -.1741025 .2062120E-01
 0. 0.
 0. 0.

ICJ-4

```

THE 1 4Y 0 0AIP1Y 0
      5.454575 -1.1342126 4.029471 .9095912 2.120102
      2.581461 1.000000 0.
COS1=11.74634355414
X1=.51471475833748
C=19.41823556267

```

```

THE 1 BY 2 MATRIX SP4(X)
.7905057 -4.474341E-01 3.419976 2.227077 5.745024
.7771200E-02 4.701219 .6321404E-02 -.7551215E-02

```

TX=,01242978743071

```
NOACTIV=5  
ARRAY IOACTV  
I=1IOACTV=2  
I=2IOACTV=5  
I=3IOACTV=7  
I=4IOACTV=1  
I=5IOACTV=4
```

THE 1 BY 9 MATP1X QPLMD
0.8959417 0.4643060 0.146222 0.233080 5.841412

```
THE 1 BY 8 MATRIX PI
-.5039587      .4745474      .2919348      .2378446      .1702277
.1079770      -.3407985E-13      0.
THEIA=-.525907422308
PASSED TEST OF STEP 3
```

```

THE 1 BY 3 MATRIX L11
      .2945052E-01      .8512752E-03      .4117408E-01      .9475813E-02      .1437711E-01
      -.1390318E-01      .1345273E-12      0.

```

THE 1 RY 8 MATRIX NEWY				
5.380067	2.2911650	4.362984	1.156907	2.305547
2.675475	1.000000	0.		

IN POWELL UPDATE SCHEME
PARAM=1.

```

THE      3 BY 8 MATRIX H
1.181172      .3617213      .4705348      .2699713E-01      -.8591057E-01
-.7500167      .1011304      0.      .3523810E-01      -.2210265      -.9529072E-01
.3617213      .4035967      0.      .3523810E-01      -.1217615      -.5828664
.1168157E-01      .6529679E-01      0.      .8655095E-02      .5072128      .3979270E-01
.4705348      -.3523810E-01      0.      .1.334908      -.1217615      .5072128      .3979270E-01
-.4920777      .8655095E-02      0.      .1.334908      -.1217615      .5072128      .3979270E-01
.2639713E-01      -.2210265      0.      .1.334908      -.1217615      .5072128      .3979270E-01
.7043588E-02      -.2002969      0.      .1.334908      -.1217615      .5072128      .3979270E-01
-.4591057E-01      -.9529072E-01      0.      .1.334908      -.1217615      .5072128      .3979270E-01
.3358933      .1341335      0.      .1.334908      -.1217615      .5072128      .3979270E-01
-.7500167      .1158159E-01      0.      .1.334908      -.1217615      .5072128      .3979270E-01
.3.650409      -.1583005      0.      .1.334908      -.1217615      .5072128      .3979270E-01
.1011394      .6529679E-01      0.      .1.334908      -.1217615      .5072128      .3979270E-01
-.1563005      .9317694      0.      .1.334908      -.1217615      .5072128      .3979270E-01
0.      0.      0.      .      .      .      .
0.      0.      0.      .      .      .      .
0.      0.      0.      .      .      .      .

```

0.

ICT=5

```

THE 1 BY 3 MATRIX X
5.380067      .2911660      4.362984      1.106907      2.305547
2.675875      1.000000      0.
COST=11.72657383994
X1=.001535605908089
C=19.41893556267

```

```
THE 1 BY 9 MATRIX GPM(X)
.8569829      -.4722030E-02      4.231843      2.541624      5.347255
.1225835E-01      4.044954      .3269777E-02      -.3603890E-02
```

TX= .0006384537678234

```

NQACTIV=5
ARRAY IQACTV
I=1 IQACTV=3
I=2 IQACTV=5
I=3 IQACTV=7
I=4 IQACTV=1
I=5 IQACTV=4

```

THE 1.0RY 9 MATPIX QPLMD
0.8519458 0.4013459 0.274762 0.2507920 0.254934

```

THE 1-BY 4 MATRIX PI
  -1.864132E-01 .1241034E-01 .1350140 .4288124E-02 .4884740E-01
  -1.4155359E-01 -1.1950617E-10 0.
THETA=-.074377284488597
PASSED TEST OF STEP 3

```

```

THE 1 BY 5 MATRIX L11
      .3377085E-02      .4111371E-03      .4086226E-02      .5422340E-04      -.4084044E-04
      -.1387841E-02      .1949411E-10      0.

```

```

THE 1 HY R MATRIX NEWX
5.196531      1.3039875      4.502085      1.161249      2.354754
2.632434      1.000000      0.

```


IN POWELL UPDATE SCHEME
PARAM=1.

THE 8 BY 4 MATRIX H

1.538219	.4010231	.2110649	.6423421E-01	.2000000
-2.607559	.1730021	0.	0.	0.
.4210231	.3691125	-.6920485E-01	-.2208726	.2559999
-.7154892	.9642648E-01	0.	0.	0.
.2110231	-.6720442E-01	1.527521	-.1905111	-1.349571
.9143609	-.5484791E-01	0.	.5107905	.1874111
-.6328051E-01	-.2208726	0.	0.	0.
-.2522705	-.1481211	-.1505111	.1774111	2.460313
.9000834	-.2351343	1.349631	0.	0.
-2.070827	.2753435	0.	.8143609	-.2522705
-2.607559	-.7154892	0.	-.5484791E-01	-.1905111
6.799599	-.3691125	0.	0.	0.
.1830621	.4545648E-01	0.	0.	0.
-.3606211	.4437513	0.	0.	0.
0.	0.	0.	0.	0.
0.	0.	0.	0.	0.

ICT=6

THE 1 BY 8 MATRIX X

5.195531	.3039875	4.502085	1.151249	2.354354
2.632934	1.000000	0.	0.	0.
COST=11.59761050159				
XI=.0000504999330151				
C=19.418435556267				

THE 1 BY 9 MATRIX GPM(X)

.8652392	-.3573638E-02	4.392774	2.764029	5.172364
.1062324E-01	3.843078	.2259385E-02	-.2355202E-02	0.

TX=.0002323650932088

NDACTIV=5
ARRAY IQACTV

I=1 IQACTV=5
I=2 IQACTV=3
I=3 IQACTV=7
I=4 IQACTV=1
I=5 IQACTV=4

THE 1 BY 9 MATRIX QPLMD

.8550435	0.	4.595655	3.092736	4.997745
0.	3.544260	0.	0.	0.

THE 1 BY 3 MATRIX PI

-.3861333	-.1743679	.3089594	-.7948338E-01	.4197557E-01
-.1743679	.2497503E-09	0.	0.	0.
THETA=-.26743229750095				
PASSED TEST OF STEP 3				

THE 1 BY 8 MATRIX L11

.1521634E-01	-.1844437E-02	.1941126E-01	.2402751E-02	-.2151631E-02
.1638075E-03	-.2497377E-09	0.	0.	0.

THE 1 BY 8 MATRIX NEWX

4.825614	.1277751	4.830455	1.064168	2.434179
2.458129	1.000000	0.	0.	0.

IN POWELL UPDATE SCHEME
PARAM=1.

THE 8 BY 8 MATRIX H

1.835299	.5818683	-.1742786	.1319100	1.762372
-3.304783	.2108346	0.	0.	0.
.5818683	.4301071	-.2032881	-.1950532	.6740937
-1.145429	.1144162	0.	0.	0.
-.1742786	-.2032881	1.814164	-.2044097	-2.244426
1.835299	-.6531594E-01	0.	0.	0.
-.1319100	-.1950532	-.2084097	.5217595	.3452104
-.4184433	-.1812197	0.	0.	0.
1.762372	.6740937	-2.244426	.3459104	4.717247
-4.241186	.3652247	0.	0.	0.
-3.304783	-1.145429	1.866816	-.4184433	-4.241186
8.336993	-.4184433	0.	0.	0.
.2108346	.1144162	-.9931594E-01	-.1812197	.3652247
-.4184433	.9459747	0.	0.	0.
0.	0.	0.	0.	0.
0.	0.	0.	0.	0.

ICT=7

THE 1 BY 8 MATRIX X

4.825614	.1277751	4.830455	1.064168	2.434179
2.458129	1.000000	0.	0.	0.
COST=11.55445582391				
XI=.0000024294548795				
C=19.418435556267				

THE 1 BY 9 MATRIX GPM(X)

.9608372	.2326838E-02	4.801100	3.354707	4.516210
.2107493E-02	3.366430	-.2917150E-05	.2852545E-04	0.

TX=.000004046329328411

NDACTIV=5
ARRAY IQACTV

I=1 IQACTV=5
I=2 IQACTV=3
I=3 IQACTV=4
I=4 IQACTV=1
I=5 IQACTV=7

```

THE 1 BY 9 MATRIX SPLMD
.9477146 0. 2.454019 4.749301 2.371623 4.451172
0. 0. 0. 0. 0. 0.

THE 1 BY 8 MATRIX PI
-.1214355E-02 -.1248514
-.4434330E-01 -.3478244E-09 0. .6229039E-03 -.5353302E-01 -.1461522E-01
THETA=-.01257352227859
PASSED TEST OF STEP 3

THE 1 BY 8 MATRIX L11
.5317087E-03 -.4316252E-03 0. .6105502E-03 .5094240E-03 .4567192E-04
.2934045E-03 -.3478533E-09 0. 0. 0. 0.

THE 1 BY 8 MATRIX NEWX
4.824317 .4585332E-01 4.831015 1.022481 2.421252
2.429022 1.000000 0. 0. 0. 0.

IN POWELL UPDATE SCHEME
PARAM=1.

THE 8 BY 8 MATRIX H
1.839903 .5326082 .8089016E-02 .1167372 1.479244
-3.030479 .1576079 0. -.2777149 -.1934335 .8071058
.5326082 .4566081 0. 0. 0. 0.
-1.040585 .1228545 0. 0. 0. 0.
.8089016E-02 -.2777149 2.001106 -.2068579 -2.592422
1.484131 -.1041827 0. .2068579 .5201732 .3482312
.1167372 -.1934335 0. -.2068579 .3489312 5.371356
-.3500700 -.1859065 0. 0. 0. 0.
-1.478264 .8071058 0. 0. 0. 0.
-3.747759 .3418076 0. 0. 0. 0.
-3.030479 -.1040583 0. 1.484131 -.3540700 -3.747759
6.674010 -.2279034 0. 0. 0. 0.
.1576079 .1228545 0. 0. 0. 0.
-.2279034 .9326023 0. 0. 0. 0.
0. 0. 0. 0. 0. 0.
0. 0. 0. 0. 0. 0.
*****
ICT=8

THE 1 BY 8 MATRIX X
4.824317 .4585332E-01 4.831015 1.022481 2.421252
2.429022 1.000000 0. 0. 0. 0.
COST=11.65434146379
X1=.0005517824990271
C=19.41893556267

THE 1 BY 9 MATRIX GPM(X)
.8561297 .9026212E-03 4.820605 3.400914 4.821777
.7405463E-03 3.394190 -.2095426E-04 .2447013E-04 0.

TX=5.408464557319E-7
NOACTIV=5
ARRAY IOACTIV
I=1 IOACTIV=3
I=2 IOACTIV=4
I=3 IOACTIV=1
I=4 IOACTIV=5
I=5 IOACTIV=7

THE 1 BY 9 MATRIX SPLMD
.8533808 0. 4.826725 3.411142 4.829661
0. 3.416799 0. 0. 0. 0.

THE 1 BY 8 MATRIX PI
.4316451E-02 -.4630524E-01 -.3455309E-02 -.2281342E-01 -.5051471E-02
-.1478231E-01 -.2322439E-09 0. 0. 0. 0.
THETA=-.006624219263455
PASSED TEST OF STEP 3

THE 1 BY 8 MATRIX L11
.7797590E-04 -.5133524E-04 .9290956E-04 .8024704E-04 .8953494E-05
.2522161E-04 .4057149E-11 0. 0. 0. 0.

THE 1 BY 8 MATRIX NEWX
4.829211 -.5032503E-03 4.827643 .9997479 2.412910
2.414265 1.000000 0. 0. 0. 0.

IN POWELL UPDATE SCHEME
PARAM=1.

THE 8 BY 8 MATRIX H
1.836120 .5339977 .1231248E-01 .1004791 1.465690
-3.001394 .1573730 0. -.2787528 -.1885891 .8104477
.5339977 .4561074 0. 0. 0. 0.
-1.089490 .1225196 0. 0. 0. 0.
.1231248E-01 -.2787528 2.009415 -.2137369 -2.519333
1.491075 -.1044703 0. 0. 0. 0.
.1004791 -.1885891 0. .2187368 .5192375 .3705952
-.3250640 -.1856495 0. 0. 0. 0.
-1.465690 .6104477 0. -.2618333 .3705953 5.412538
-3.750580 .3424154 0. 0. 0. 0.
-3.001394 -.1040592 0. 1.491075 -.3250680 -3.750580
6.572115 -.2277543 0. 0. 0. 0.
.1573730 .1224195 0. 0. 0. 0.
-.2277543 .9326110 0. 0. 0. 0.
0. 0. 0. 0. 0. 0.
0. 0. 0. 0. 0. 0.
0. 0. 0. 0. 0. 0.

```

ICT=9

THE 1 BY 8 MATRIX X
4.522211 -1.503250E-03 4.827643 .4997479 2.413210
2.414240 1.000000 0.
COST=11.4148532203
XI=.000001624957-46752
C=19.41892556267

THE 1 BY 9 MATRIX GRM(X)
.4535240 -1.3046540E-04 4.827702 3.413016 4.827702
.1113320E-04 3.413016 .2091204E-05 -1.3046540E-05

TX=1.222009672057E-12

NOACTIV=5
ARRAY IQACTIV
I=1 IQACTIV=3
I=2 IQACTIV=4
I=3 IQACTIV=1
I=4 IQACTIV=5
I=5 IQACTIV=7

THE 1 BY 9 MATRIX QPLMD
.4535242 0. 4.828787 3.414664 4.828787
0. 3.414030 0. 0.

THE 1 BY 3 MATRIX PI
-1.6931144E-03 .5029315E-03 .6932563E-03 .2519500E-03 .2772213E-03
-1.2472387E-04 .5222469E-12 0.
THEIA=-.00001902542652363
PASSED TEST OF STEP 3

THE 1 BY 3 MATRIX L11
.4547474E-12 .1364242E-11 .1364242E-11 .1364242E-11
-1.1356463E-07 .1356335E-07 0. .1356373E-07 -1.1356326E-07

THE 1 BY 8 MATRIX NEWX
4.828518 -1.3187916E-06 4.828336 .9999999 2.414137
2.414240 1.000000 0.

IN POTELL UPDATE SCHEME
PARAM=1.

THE 3 BY 8 MATRIX H
1.849045 .4810088 .3137076 -1.1490544E-01 1.044121
-2.555342 .7524083E-01 0. -1.725647 -2.192204 .6460950
4.310088 .4312448 0. -1.180366 -1.959204
-1.8589454 .1012314 0. -1.180366 .5010029 .1966834
3.137076 -1.1725647 0. -1.959204 .1966834 4.365714
4.783670 -1.3343233E-01 0. 4.783670 -2.263571E-01 -2.192467
-1.1490544E-01 -1.2192204 0. -1.180366 .5010029 .1966834
-1.2263621E-01 -1.197511 0. -1.180366 .5010029 .1966834
1.044121 .6460950 -1.959204 .1966834 4.365714
-2.192467 .2695265 0. 4.783670 -2.263571E-01 -2.192467
-2.555342 -1.8589454 0. 4.783670 -2.263571E-01 -2.192467
4.447003 -1.1347288E-01 0. -1.180366 .5010029 .1966834
7.524083E-01 .1012314 0. -1.180366 .5010029 .1966834
-1.1347288E-01 .9240188 0. 0. 0. 0.
0. 0. 0. 0. 0. 0.

ICT=10

THE 1 BY 8 MATRIX X
4.828518 -1.3187916E-06 4.828336 .9999999 2.414197
2.414240 1.000000 0.
COST=11.65685386558
XI=8.697706022076E-8
C=19.41893556267

THE 1 BY 9 MATRIX GRM(X)
.4535244 -1.2416220E-05 4.828324 3.414075 4.828324
.2308795E-05 3.414075 .6919545E-06 -1.3046540E-05

TX=7.603950412311E-15

NOACTIV=5
ARRAY IQACTIV
I=1 IQACTIV=3
I=2 IQACTIV=4
I=3 IQACTIV=1
I=4 IQACTIV=5
I=5 IQACTIV=7

THE 1 BY 9 MATRIX QPLMD
.4535551 0. 4.828439 3.414228 4.828439
0. 3.414210 0. 0.

THE 1 BY 3 MATRIX PI
-1.0318506E-04 .1843137E-05 .9034359E-04 .9062505E-05 .2672625E-04
-1.2581594E-04 -1.1360749E-07 0.
THEIA=-.000001305859942754
PASSED TEST OF STEP 3

THE 1 BY 4 MATRIX L11
0. .9094947E-12 .1364242E-11 .1364242E-11
.1818989E-11 .1818989E-11 0.

THE 1 BY 8 MATRIX NEWX
4.828424 .1324346E-05 4.828426 1.000001 2.414214
2.414214 1.000000 0.

IN PJWELL UPDATE SCHEME
PARAM=1.

THE 8 BY 8 MATRIX H
1.405366 .4891444 .3655650 -.5845828E-01 1.017870
-2.526621 .4592955E-01 0. -.1535250 -.2351736 .6017255
-.4491444 .2794500 0. 1.422847 -.5575294E-01 -1.594263
-.4124706 .8450617E-01 0. 0. 0. 0.
.3055050 -.1337750 0. 1.422847 -.5575294E-01 -1.594263
.2056422 -.4623713E-03 0. 0. 0. 0.
-.2447425E-01 -.2241536 -.5578249E-01 .4381518 .1003126 .4007270
-.7546611E-01 -.2002413 0. 0. 0. 0.
1.017870 .6017255 -1.694053 .1003126 4.007270
-1.422847 .2056422 0. 0. 0. 0.
-2.526621 -.8128756 .2056422 .7546611E-01 -1.622350
4.265190 .3289664E-01 0. .9553713E-03 -.2002413 .2047954
.4592955E-01 .8450617E-01 0. 0. 0. 0.
.5269564E-01 .225124 0. 0. 0. 0.
0. 0. 0. 0. 0. 0.
0. 0. 0. 0. 0. 0.

ICT=11

THE 1 BY 8 MATRIX X
4.828427 .1524344E-05 4.828426 1.000001 2.414214
2.414214 1.000000 0. 0. 0. 0.
COST=11.45645424381
X1=1.03536221267E-9
C=19.41493654267

THE 1 BY 8 MATRIX GPM(X)
.3535535 .4000413E-03 4.828426 3.414213 4.829427
.5075907E-07 3.414213 .6846457E-06 -.5546451E-08 0.

TX=1.074504420172E-18

NOACTIV=5
ARRPAY IDACTIV
I=1 IDACTIV=3
I=2 IDACTIV=4
I=3 IDACTIV=1
I=4 IDACTIV=5
I=5 IDACTIV=7

THE 1 BY 9 MATRIX OPLMD
.9535534 0. 4.828427 3.414214 4.829427
0. 3.414213 0. 0. 0. 0. 0. 0. 0.

THE 1 BY 8 MATRIX PI
-.9561491E-06 -.1488373E-05 .9585989E-06 -.7442020E-06 .6222253E-07
-.8049162E-06 .1297300E-10 0. 0. 0. 0.
THEIA=-1.280772212123E-8
PASSED TEST OF STEP 3

THE 1 BY 8 MATRIX L11
0. .9094947E-12 .1364242E-11 .1364242E-11
.1818489E-11 .1818989E-11 0. 0. 0. 0. 0.

THE 1 BY 8 MATRIX NEWX
4.828427 .3597418E-07 4.828427 1.000000 2.414214
2.414214 1.000000 0. 0. 0. 0.

IN PJWELL UPDATE SCHEME
PARAM=1.

THE 8 BY 8 MATRIX H
1.912562 .5048318 .3332517 -.6202037E-01 1.089920
-2.5229145 .4344061E-01 0. -.8861670E-01 -.2622714 .5544230
-.5048318 .4243802 0. 0. 0. 0.
-.7638410 .7147795E-01 0. 1.215138 -.5482819E-03 -1.395520
.3332517 -.8861670E-01 0. 0. 0. 0.
.5011290E-01 .4708662E-01 0. 0. 0. 0.
-.5202237E-01 -.2522718 -.6982319E-03 .4428351 -.2241715E-01
.1007296 -.2039559 0. 0. 0. 0.
1.089920 .5544230 -1.395520 -.2241715E-01 3.627677
-1.562297 .1191740 0. 0. 0. 0.
-2.5229145 -.7638410 .5011290E-01 .1007296 -1.562297
3.987314 .7005297E-01 0. .4708662E-01 -.2039559 .1191740
.4344061E-01 .7147795E-01 0. 0. 0. 0.
.7805297E-01 .4199175 0. 0. 0. 0.
0. 0. 0. 0. 0. 0.
0. 0. 0. 0. 0. 0.

B.1 Basic Algorithm, with the update given by I.4.VI and $\tilde{p}_i$ computed after $T(x_i) \leq \epsilon$, applied to Problem B.

ICT=1

THE 1 BY 6 MATRIX X
1.000000 1.000000 1.000000 1.000000 1.000000

COST=2.718281828459
XI=5.
C=50.

THE 1 BY 3 MATRIX MU(X)
-1.359141 0. 0.

ATX=5.
TX=1.

NQACTIV=3
ARRAY IQACTV
I=1IQACTV=4
I=2IQACTV=5
I=3IQACTV=3

THE 1 BY 6 MATRIX QPLMD
0. 0. 1.269841 2.860842 .6122449
0.

THE 1 BY 6 MATRIX PI
-.8061224 -.1938776 3.615646 -.5782313E-01 -.5782313E-01
0.
THETA=-243.2042954253
STEP LENGTH=.6561

THE 1 BY 6 MATRIX NEWX
.4711031 .8727969 3.372226 .9620622 .9620622
0.

C1=-23.54969161054
C2=5.926245008488
C3=5.926245008488
C4=0.
D1=5.926245008488
THE VECTOR NNT*PI WAS SET TO ZERO
IN THE NEW UPDATE SCHEME
PARAM=5.973796489481

THE 1 BY 6 MATRIX DLGR2
5.390869 2.174003 6.377931 1.972251 1.972251
0.
C5=11.85249001698

THE 6 BY 6 MATRIX H
3.404728 .9774495 3.112588 .8936535 .8936535
0. .9774495 1.396029 1.220769 .3609392 .3609392
0. 3.112588 1.220769 3.482440 1.076472 1.076472
0. .8936535 .3609392 1.076472 1.327939 .3279393
0. .8936535 .3609392 1.076472 .3279393 1.327939
0. 0. 0. 0. 0.
0.

THE 1 BY 3 MATRIX C7NSTR
4.206745 -1.684551 1.769430

ICT=2

THE 1 BY 6 MATRIX X
.4711031 .8727969 3.372226 .9620622 .9620622
0.
COST=3.608783088924
XI=4.206745008534
C=25.

THE 1 BY 3 MATRIX MU(X)
-.6065259 .5644547 -1.742332

ATX=81.19917005324
TX=1.

NQACTIV=3
ARRAY IQACTV
I=1IQACTV=2
I=2IQACTV=6
I=3IQACTV=1

THE 1 BY 6 MATRIX QPLMD
.9010495 .8215580 0. 0. 0.
3.035294

THE 1 BY 6 MATRIX PI
-2.254218 -.1175063 -.1473681 -.2296557 -.2296557
0.
THETA=-130.3668746116
STEP LENGTH=.729

THE 1 BY 6 MATRIX NEWX
-1.172222 .7871349 3.264794 .7946432 .7946432
0.

```

C1=24.44984928022
C2=11.8160790467
C3=2.775453905567
C4=0.
D1=2.775453905567
IN THE NEW UPDATE SCHEME
PARAM=0.

THE 1 BY 6 MATRIX DLGR2
-13.61774 -4.680533 -1.378383 -4.495695 -4.495695
0.
C5=24.44984928022

THE 6 BY 6 MATRIX H
7.617102 2.639123 .6994650 2.386664 2.386664
0. 2.639123 2.039184 .4992684 .9415875 .9415875
0. .6994650 .4992684 .5598541 .3763622 .3763622
0. 2.386664 .9415875 .3763622 1.851519 .9515194
0. 2.386664 .9415875 .3763622 .8515194 1.851519
0. 0. 0. 0. 0. 0.
0.
0.

THE 1 BY 3 MATRIX CONSTR
3.915482 -.5874561 -.1230600
*****
ICT=3

THE 1 BY 6 MATRIX X
-1.172222 .7371349 3.264794 .7946432 .7946432
0.
COST=.1492380148984
XI=3.915481802851
C=12.5

THE 1 BY 3 MATRIX MU(X)
.5045301E-01 -.2792634E-01 .5210355E-02
ATX=4.171978681292
TX=1.

NDACTIV=3
ARRAY IQACTV
I=1 IQACTV=5
I=2 IQACTV=1
I=3 IQACTV=3

THE 1 BY 6 MATRIX QPLMD
.6718755E-01 0. .8302392E-01 0. .8260925E-01
0.

THE 1 BY 6 MATRIX PI
-.2610329 .6451231 -.8986198 .1021091 .1021091
0.
THETA=-49.2342220596
STEP LENGTH=1.

THE 1 BY 6 MATRIX NEWX
-1.433255 1.432258 2.366174 .8967523 .8967523
0.
C1=.4323492023122
C2=.5920836700038
C3=1.312692097748
C4=0.
D1=1.312692097748
IN THE NEW UPDATE SCHEME
PARAM=0.

THE 1 BY 6 MATRIX DLGR2
-.2030505 .4183539 -.5495330E-01 .2941681 .2941681
0.
C5=.4323492023122

THE 6 BY 6 MATRIX H
7.404652 2.709615 .5185339 2.192443 2.192443
0. 2.709615 2.212449 .6254023 1.274861 1.274861
0. .5185339 .6254023 .4279836 .3013155 .3013155
0. 2.192443 1.274861 .3013155 2.041457 1.041457
0. 2.192443 1.274861 .3013155 1.041457 2.041457
0. 0. 0. 0. 0.
0. 0.

THE 1 BY 3 MATRIX CONSTR
1.312692 -.6318517 .9938628
*****
ICT=4

THE 1 BY 6 MATRIX X
-1.433255 1.432258 2.366174 .8967523 .8967523
0.
COST=.02012007772103
XI=1.312692097706
C=6.25

THE 1 BY 3 MATRIX MU(X)
.1667471E-01 -.1244056E-01 .2390092E-02
ATX=1.314427332775
TX=1.

```

NOACTIV=3
 ARRAY IQACTV
 I=1 IQACTV=3
 I=2 IQACTV=5
 I=3 IQACTV=1

THE 1 BY 6 MATRIX QPLMD
 .7558908E-01 0. .1152526 0. .1042036
 0.

THE 1 BY 6 MATRIX PI
 .9611360E-01 -.2577437 .3728697E-01 -.1325130 -.1325130
 0.
 THETA=-8.162924720537
 STEP LENGTH=1.

THE 1 BY 6 MATRIX NEWX
 -1.337141 1.174514 2.403461 .7642393 .7642393
 0.

C1=.1172094031451
 C2=.2382107739954
 C3=.1121793839413
 C4=0.
 D1=.1121793839413
 IN THE NEW UPDATE SCHEME
 PARAM=0.

THE 1 BY 6 MATRIX DLGR2
 .1283096 -.1660314 -.6350464E-01 -.2431891 -.2431891
 0.
 C5=.1172094031451

THE 6 BY 6 MATRIX H
 6.282522 1.090420 .4553915E-01 .7402118 .7402118
 0. 1.090420 .8111347 .2560079 .2690920 .2690920
 0. .4553915E-01 .2560079 .3334551 .5407272E-01 .5407273E-01
 0. .7402118 .2690920 .5407272E-01 1.431957 .4319571
 0. .7402118 .2690920 .5407273E-01 .4319571 1.431957
 0. 0. 0. 0. 0. 0.

THE 1 BY 3 MATRIX CNSTR
 .1121794 -.9740901E-01 .2294876

 ICT=5

THE 1 BY 6 MATRIX X
 -1.337141 1.174514 2.403461 .7642393 .7642393
 0.
 COST=.1102937661536
 XI=.2294875849982
 C=3.913312694022

THE 1 BY 3 MATRIX MU(X)
 .6131193E-01 -.5338734E-01 .1516078E-01

ATX=.2738175418159
 TX=.07497604620612

NOACTIV=3
 ARRAY IQACTV
 I=1 IQACTV=5
 I=2 IQACTV=3
 I=3 IQACTV=1

THE 1 BY 6 MATRIX QPLMD
 .4625000E-01 0. .9168365E-01 0. .8359672E-01
 0.

THE 1 BY 6 MATRIX PI
 -.9737872E-01 .7075982E-01 -.1076194 -.7031960E-02 -.7031959E-02
 0.
 THETA=-.9150514601138
 PASSED TEST OF STEP 3

THE 1 BY 6 MATRIX L11
 .3612416E-02 .3840251E-03 -.2773709E-02 -.1334286E-02 -.1334286E-02
 0.
 STEP LENGTH=1.

THE 1 BY 6 MATRIX NEWX
 -1.430907 1.245658 2.293068 .7558731 .7558731
 0.

C1=.006988559449801
 C2=.04792197621523
 C3=.02618016554244
 C4=0.
 D1=.001962866800915
 IN THE NEW UPDATE SCHEME
 PARAM=0.

THE 1 BY 6 MATRIX DLGR2
 -.2357975E-01 .4802324E-01 -.1473576E-01 .1587990E-01 .1587990E-01
 0.
 C5=.006988559449801

```

THE 6 BY 6 MATRIX H
.5242234 .7519423E-01 -.1671245 -.1061255 -.1061255
0. .7519423E-01 1.016443 .1164015 .2623537 .2623537
0. -.1671245 .1164015 .3527335 -.1504145E-01 -.1504145E-01
0. -.1061255 .2623537 -.1504145E-01 1.360387 .3603872
0. -.1061255 .2623537 -.1504145E-01 .3603872 1.360387
0. 0. 0. 0. 0.
0. 0. 0. 0. 0.

THE 1 BY 3 MATRIX CONSTR
.9793407E-05 -.3413889E-03 .3066683E-02
*****
ICT=6

THE 1 BY 6 MATRIX X
.1430907 1.245658 2.293068 .7558731 .7558731
0.
COST=.09679061804432
XI=.003066682971323
C=3.913312694022

THE 1 BY 3 MATRIX MU(X)
.5848395E-01 -.5265001E-01 .1323609E-01
0. 0. 0.

ATX=.02833403159683
TX=.0008028173465301
NQACTIV=3
ARRAY IQACTV
I=1IQACTV=6
I=2IQACTV=1
I=3IQACTV=5

THE 1 BY 6 MATRIX QPLMD
.3717853E-01 0. 0. .3389272E-01
.2960121E-02 0. 0.

THE 1 BY 6 MATRIX PI
.1530972 .2013603 -.2211381 .2459793E-01 .2459793E-01
0.
THETA=-.06563459393974
PASSED TEST OF STEP 3

THE 1 BY 6 MATRIX L11
.8184179E-03 -.1083827E-01 -.1189419E-01 -.1153934E-01 -.1153934E-01
0.
STEP LENGTH=.81

THE 1 BY 6 MATRIX NEWX
.1555453 1.401649 2.106143 .7682264 .7682264
0.
C1=.0252916906506
C2=.03103768790226
C3=.07509117949924
C4=0.
D1=.00006028450147339
IN THE NEW UPDATE SCHEME
PARAM=0.

THE 1 BY 6 MATRIX DLGR2
.4183130E-01 .9728814E-01 -.1712406E-01 .6900131E-01 .6900131E-01
0.
C5=.0252916906506

THE 6 BY 6 MATRIX H
.5733664 .2189939E-01 -.1607684 -.1573989 -.1573989
0. .2189939E-01 .8128939 .1684672 .1903295 .1903295
0. -.1607684 .1684672 .3402547 .7119633E-02 .7119633E-02
0. -.1573989 .1903295 .7119633E-02 1.351555 .3515551
0. -.1573989 .1903295 .7119633E-02 .3515551 1.351555
0. 0. 0. 0. 0.
0. 0. 0. 0. 0.

THE 1 BY 3 MATRIX CONSTR
.2340679E-03 .1213417E-02 -.9608132E-02
*****
ICT=7

THE 1 BY 6 MATRIX X
.1555453 1.401649 2.106143 .7682264 .7682264
0.
COST=.06653967003095
XI=.009608131961073
C=3.913312694022

THE 1 BY 3 MATRIX M'(X)
.4633309E-01 -.4218606E-01 .8249053E-02
0. 0. 0.

ATX=.01467807342949
TX=.0002154458396016
NQACTIV=3
ARRAY IQACTV
I=1IQACTV=5
I=2IQACTV=1
I=3IQACTV=3

```



```

THE 1 BY 6 MATRIX QPLMD
.3596096E-01 0. .3797177E-02 0. .3400761E-01

THE 1 BY 6 MATRIX PI
-.1040857 .1298120 -.1671469 .5250365E-02 .5250365E-02
0.
THETA=-.0545203881321
PASSED TEST OF STEP 3

THE 1 BY 6 MATRIX L11
.1247207E-02 -.5160639E-02 -.5080650E-02 -.5184128E-02 -.5184128E-02
0.
STEP LENGTH=1.

THE 1 BY 6 MATRIX NEWX
-1.658291 1.526300 1.933915 .7682927 .7682927
0.
C1=.01170253963444
C2=.01530767984239
C3=.05577608104092
C4=0.
D1=.00001201672460955
IN THE NEW UPDATE SCHEME
PARAM=0.

THE 1 BY 6 MATRIX DLGR2
-.2037758E-01 .5121644E-01 -.1869242E-01 .2559069E-01 .2559069E-01
0.
C5=.01170253963444

THE 6 BY 6 MATRIX H
.5555404 .6350907E-01 -.1675344 -.1295566 -.1295566
0.
.6350907E-01 .7161492 .1831175 .1246892 .1246892
0.
.1675344 .1831175 .3411176 .1964027E-01 .1964027E-01
0.
.1295566 .1246892 .1964027E-01 1.309180 .3091800
0.
.1295566 .1246892 .1964027E-01 .3091800 1.309180
0.
0. 0. 0. 0. 0.
0.

THE 1 BY 3 MATRIX CONSTR
.9784827E-04 .3670848E-03 -.4528230E-02
*****
ICT=8

THE 1 BY 6 MATRIX X
-1.658291 1.526300 1.933915 .7682927 .7682927
0.
COST=.05561550504315
XI=.004528230229283
C=3.913312694022

THE 1 BY 3 MATRIX MU(X)
.4107500E-01 -.3814729E-01 .5983851E-02
0.
ATX=.005078118214501
TX=.00002578728460045
NQACTIV=3
ARRAY IQACTV
I=1IQACTV=5
I=2IQACTV=1
I=3IQACTV=3

THE 1 BY 6 MATRIX QPLMD
.3883123E-01 0. .4930399E-02 0. .3671893E-01
0.

THE 1 BY 6 MATRIX PI
-.5105201E-01 .6091146E-01 -.8990371E-01 -.2480226E-02 -.2480226E-02
0.
THETA=-.02057252725024
PASSED TEST OF STEP 3

THE 1 BY 6 MATRIX L11
.6091714E-03 -.1307966E-02 -.1157538E-02 -.1275968E-02 -.1275968E-02
0.
STEP LENGTH=1.

THE 1 BY 6 MATRIX NEWX
-1.708734 1.585904 1.842954 .7645365 .7645365
0.
C1=.002434678725396
C2=.00274026440445
C3=.01441742393818
C4=0.
D1=3.717862142991E-7
IN THE NEW UPDATE SCHEME
PARAM=0.

THE 1 BY 6 MATRIX DLGR2
.5771325E-02 .1761339E-01 -.1217648E-01 .2005755E-02 .2005755E-02
0.
C5=.002434678725396

```

```

THE 6 BY 6 MATRIX H
.5453161 .8567221E-01 -.1732872 -.1164844 -.1164444
0. .8567221E-01 .6590303 .1895603 .9051785E-01 .9051785E-01
0. -.1732872 .1895603 .3508164 .3597519E-01 .3597519E-01
0. -.1164844 .9051785E-01 .3597519E-01 1.297254 .2972545
0. -.1164844 .9051785E-01 .3597519E-01 .2972545 1.297254
0. 0. 0. 0. 0.
0.

THE 1 BY 3 MATRIX CONSTR
.5931448E-05 .8810362E-05 -.4237261E-03
*****
ICT=9

THE 1 BY 6 MATRIX X
-1.708734 1.585904 1.842854 .7645365 .7645365
0.
COST=.05398598646872
XI=.0004237260842217
C=3.913312644022

THE 1 BY 3 MATRIX MU(X)
.4019375E-01 -.3784002E-01 .5311445E-02
ATX=.0004348224457852
TX=1.890705593586E-7
NQACTION=3
ARRAY IOACTV
I=1IOACTV=5
I=2IOACTV=1
I=3IOACTV=3

THE 1 BY 6 MATRIX QPLMD
.4009502E-01 0. .5210994E-02 0. .3789650E-01
0.

THE 1 BY 6 MATRIX PI
-.8300226E-02 .9691903E-02 -.1533780E-01 -.8442648E-03 -.8442648E-03
0.
THETA=-.001727099082655
PASSED TEST OF STEP 3

THE 1 BY 6 MATRIX L11
.2024663E-04 -.3597323E-04 -.2981532E-04 -.3476537E-04 -.3476540E-04
0.
STEP LENGTH=1.

THE 1 BY 6 MATRIX NEWX
-1.717014 1.595560 1.827486 .7636574 .7636574
0.
C1=.00006560198733754
C2=.00006653641715276
C3=.0003995041732633
C4=2.936061608981E-7
D1=7.553447750497E-11
IN THE NEW UPDATE SCHEME
PARAM=0.

THE 1 BY 6 MATRIX DLGR2
-.6987428E-03 .2492645E-02 -.2282278E-02 -.3836194E-03 -.3836194E-03
0.
C5=.00006560193733754

THE 6 BY 6 MATRIX H
.5430465 .9230593E-01 -.1761003 -.1138349 -.1138349
0. .9230593E-01 .6559144 .1909737 .8060946E-01 .8060946E-01
0. -.1761003 .1909737 .3581824 .4550604E-01 .4550604E-01
0. -.1138349 .8060946E-01 .4550604E-01 1.299296 .2992957
0. -.1138349 .8060946E-01 .4550604E-01 .2992957 1.299296
0. 0. 0. 0. 0.
0.

THE 1 BY 3 MATRIX CONSTR
.3611603E-08 -.3567685E-07 -.1596467E-05
*****
ICT=10

THE 1 BY 6 MATRIX X
-1.717014 1.595560 1.827486 .7636574 .7636574
0.
COST=.05394986260146
XI=.000001590466830733
C=3.913312694022

THE 1 BY 3 MATRIX MU(X)
.4016295E-01 -.3795629E-01 .5223953E-02
ATX=.000001599039426848
TX=2.557086495057E-12
NQACTION=3
ARRAY IOACTV
I=1IOACTV=5
I=2IOACTV=1
I=3IOACTV=3

```

```

THE 1 BY 6 MATRIX QPLMD
0.4016261E-01 0. .5222640E-02 0. .3795766E-01

THE 1 BY 6 MATRIX PI
0.1293490E-03 .1499997E-03 -.2404952E-03 -.1435693E-04 -.1435695E-04
0.
THETA=-.000000270182204335
PASSED TEST OF STEP 3

THE 1 BY 6 MATRIX L11
0.5110905E-08 -.8734105E-08 -.7190465E-08 -.8455231E-09 -.8434199E-08
STEP LENGTH=1.

THE 1 BY 6 MATRIX NEWX
-1.717144 1.595710 1.827246 .7636431 .7636431
0.
C1=1.59973509351E-8
C2=1.599131029118E-8
C3=9.748123374066E-8
C4=2.496005844793E-12
D1=2.492679950603E-19
IN THE NEW UPDATE SCHEME
PARAM=0.

THE 1 BY 6 MATRIX DLGR2
-1.035526E-04 .3797563E-04 -.3634089E-04 -.7723137E-05 -.7723137E-05
0.
C5=1.59973509351E-8

THE 6 BY 6 MATRIX H
0.5424939 .9345254E-01 -.1768454 -.1136142 -.1136142
0.9345254E-01 .6548288 .1907638 .7922066E-01 .7922066E-01
0.1768454 .1907638 .3595612 .4706699E-01 .4706699E-01
0.1136142 .7922066E-01 .4706699E-01 1.299877 .2998771
0.1136142 .7922066E-01 .4706699E-01 .2998771 1.299877
0.
0. 0. 0. 0. 0.
0.

THE 1 BY 3 MATRIX CONSTR
0.2444267E-11 -.1959677E-10 .3382183E-11
*****
ICT=11

THE 1 BY 6 MATRIX X
-1.717144 1.595710 1.827246 .7636431 .7636431
0.
COST=.05394984776941
XI=1.959676865226E-11
C=3.913312694022

THE 1 BY 3 MATRIX MU(X)
.4016274E-01 -.3795777E-01 .5222643E-02

ATX=1.959714225517E-11
TX=3.840479845694E-22
NQACTIV=3
ARRAY IQACTV
I=1IQACTV=5
I=2IQACTV=1
I=3IQACTV=3

THE 1 BY 6 MATRIX QPLMD
0.4016274E-01 0. .5222643E-02 0. .3795777E-01

THE 1 BY 6 MATRIX PI
0.4380924E-07 -.5652181E-07 .9069032E-07 .5454783E-08 .5389033E-08
0.
THETA=-1.279488522341E-12
PASSED TEST OF STEP 3

THE 1 BY 6 MATRIX L11
0.4547474E-12 .2728484E-11 -.1818989E-11 -.2160050E-11 .9549694E-11
STEP LENGTH=1.

THE 1 BY 6 MATRIX NEWX
-1.717144 1.595710 1.827246 .7636431 .7636431
0.
C1=2.273927455947E-15
C2=2.273902245906E-15
C3=1.385998345139E-14
C4=1.44427805434E-22
D1=5.32289871067E-36
IN THE NEW UPDATE SCHEME
PARAM=0.

THE 1 BY 6 MATRIX DLGR2
0.3902201E-08 -.1430526E-07 .1370793E-07 .2934311E-08 .2925630E-08
0.
C5=2.273927455947E-15

```

THE 6 BY 6 MATRIX H
0.5424122 .9357449E-01 -.1769843 -.1136758 -.1135974
0.9357449E-01 .6550286 .1906526 .7931290E-01 .7902792E-01
0.1769843 .1906526 .3595909 .4696204E-01 .4723542E-01
0.1136758 .7931290E-01 .4696204E-01 1.299831 .2998896
0.1135974 .7902792E-01 .4723542E-01 .2993846 1.299947
0. 0. 0. 0. 0.

0.
THE 1 BY 3 MATRIX CONSTR
-.3865352E-11 -.1353558E-10 .1168132E-10

B.2 Basic Algorithm, with the update given by I.4.VI and $\tilde{p}_i$ computed from the start, applied to Problem B.

```

*****
ICT=1
THE 1 BY 6 MATRIX X
0. 1.000000 1.000000 1.000000 1.000000 1.000000
COST=2.718231928459
XI=5.
C=50.
THE 1 BY 3 MATRIX MU(X)
-1.359141 0. 0.
ATX=5.
TX=1.
NFACTIV=3
ARRAY IQACTV
I=1IQACTV=4
I=2IQACTV=5
I=3IQACTV=3
THE 1 BY 6 MATRIX OPLMD
0. 0. 1.269841 2.360342 .6122449
0.
THE 1 BY 6 MATRIX PI
-1.8061224 -1.1938776 3.615646 -.5782313E-01 -.5782313E-01
0.
THETA=-243.2042954253
PASSED TEST OF STEP 3
THE 1 BY 6 MATRIX L11
0. .1151475 -.5255251 -5.087066 -.6430301 -.6430301
STEP LENGTH=.6561
THE 1 BY 6 MATRIX NEWX
0. .4711031 .8727969 3.372226 .9620622 .9620622
C1=-23.54963161054
C2=5.926245008488
C3=5.926245008488
C4=0.
D1=5.926245008488
THE VECTOR NNT*PI WAS SET TO ZERO
IN THE NEW UPDATE SCHEME
PARAM=5.973796489481
THE 1 BY 6 MATRIX DLGR2
0. 5.390869 2.174003 6.377931 1.972251 1.972251
C5=11.85249001698
THE 6 BY 6 MATRIX H
3.404728 .9774495 3.112588 .8936535 .8936535
0. .9774495 1.396029 1.220769 .3609392 .3609392
0. 3.112588 1.220769 3.482440 1.076472 1.076472
0. .8936535 .3609392 1.076472 1.327939 .3279393
0. .8936535 .3609392 1.076472 .3279393 1.327939
0. 0. -0. 0. 0.
0. 0.
THE 1 BY 3 MATRIX CONSTR
4.206745 -1.684551 1.769430
*****
ICT=2
THE 1 BY 6 MATRIX X
0. .4711031 .8727969 3.372226 .9620622 .9620622
COST=3.608793088924
XI=4.206745008534
C=25.
THE 1 BY 3 MATRIX MU(X)
-1.6065259 .5644547 -1.742332
ATX=81.19917005324
TX=1.
NFACTIV=3
ARRAY IQACTV
I=1IQACTV=2
I=2IQACTV=6
I=3IQACTV=1

```

THE 1 BY 5 MATRIX OPLMD
 .9010495 .215590
 3.035294

0.

0.

0.

THE 1 BY 5 MATRIX PI
 -2.254218 -.1175063
 0.
 THETA=-130.3568746116
 PASSED TEST OF STEP 3

-.1473681

-.2246557

-.2246557

THE 1 BY 5 MATRIX L11
 .5707393 1.588393
 0.

-1.534689

.4264757

.4269702

STEP LENGTH=.729

THE 1 BY 5 MATRIX NEWX
 -1.172222 .7371349
 0.

3.264794

.7946432

.7946432

C1=24.44984928022
 C2=11.8160790467
 C3=2.775453905567
 C4=0.
 D1=2.775453905567
 IN THE NEW UPDATE SCHEME
 PARAM=0.

THE 1 BY 6 MATRIX DLGR2
 -13.61774 -4.880533
 0.
 C5=24.44984928022

-1.376383

-4.495695

-4.495695

THE 6 BY 6 MATRIX H
 7.617102 2.639123
 0. 2.639123 2.039184
 0. .6994650 .4992684
 0. 2.386664 .9415875
 0. 2.386664 .9415875
 0. 0.
 0. 0.

.6994650

2.386664

2.386664

.4992684

.9415875

.9415875

.5598541

.3763622

.3763622

.3763622

1.851519

.8515194

.3763622

.8515194

1.851519

0.

0.

0.

THE 1 BY 3 MATRIX CONSTR
 3.913482 -.5874561

 ICT=3

-.1230600

THE 1 BY 5 MATRIX X
 -1.172222 .7871349
 0.
 COST=.1492380148984
 XI=3.615481802851
 C=12.5

3.264794

.7946432

.7946432

THE 1 BY 3 MATRIX MU(X)
 .5045301E-01 -.2792634E-01 .5210355E-02

ATX=4.171978681292
 TX=1.

NOACTIV=3
 ARRAY IQACTV
 I=1IQACTV=5
 I=2IQACTV=1
 I=3IQACTV=3

THE 1 BY 6 MATRIX QPLMD
 .6718755E-01 0.
 0.

.8302392E-01

0.

.8260925E-01

THE 1 BY 6 MATRIX PI
 -.2610329 .6451231
 0.
 THETA=-49.23422220596
 PASSED TEST OF STEP 3

-.8986198

.1021091

.1021091

THE 1 BY 6 MATRIX L11
 -.1975419 -.9658870E-01
 0.
 STEP LENGTH=1.

-.1818824

-.1372138

-.1372138

THE 1 BY 6 MATRIX NEWX
 -1.630796 1.335669
 0.

2.184292

.7595386

.7595386

C1=.3752251806689
 C2=1.78754121102
 C3=1.68113053765
 C4=0.
 D1=1.68113053765
 IN THE NEW UPDATE SCHEME
 PARAM=0.

THE 1 BY 6 MATRIX DLGR2
-1.1601967 .3323215 -.1181466 .1165744 .1165744
0.
C5=.3752251806669

THE 6 BY 6 MATRIX H
2.755157 1.339283 -.3764950 .5440559 .5440559
0. 1.339283 2.061544 .1300793 .6237591 .6237591
0. -.3764950 .1300793 .2397137 -.6994101E-01 -.6994101E-01
0. .5440559 .6237591 -.6994101E-01 1.235800 .2357999
0. .5440559 .6237591 -.6994101E-01 .2357999 1.235600
0.
0.
0.

THE 1 BY 3 MATRIX CONSTR
***.3684384 .3299742E-01 -.9542477

ICT=4

THE 1 BY 6 MATRIX X
-1.530796 1.335669 2.164292 .7595386 .7595386
0.
COST=.06426182672335
XI=.9542476977343
C=6.25

THE 1 BY 3 MATRIX MU(X)
.4488451E-01 -.4085192E-01 .9184123E-02

ATX=.9620547876309
TX=.9255494144035

NOACTIV=3
AKRAY IQACTV
I=1 IQACTV=6
I=2 IQACTV=1
I=3 IQACTV=5

THE 1 BY 6 MATRIX QPLMD
.3391632E-01 0. 0. 0. .1442040E-01
.4184730E-01

THE 1 BY 6 MATRIX PI
.4083812E-01 .1174172 -.1355722 .1427066E-01 .1427066E-01
0.
THETA=-5.970917303685

PASSED TEST OF STEP 3

THE 1 BY 6 MATRIX L11
-2.2614583E-02 -.5215021E-02 -.3570807E-02 -.4357555E-02 -.4357555E-02
0.
STEP LENGTH=1.

THE 1 BY 6 MATRIX NEWX
-1.592573 1.447871 2.045149 .7694517 .7694517
0.

C1=.007298471420569
C2=.0522617436759
C3=.03360768374089
C4=0.
D1=.03110557200584
IN THE NEW UPDATE SCHEME
PARAM=0.

THE 1 BY 6 MATRIX DLGR2
-6.6582910E-03 .3821513E-01 -.2000580E-01 .1271778E-01 .1271778E-01
0.
C5=.007298471420569

THE 6 BY 6 MATRIX H
.8110583 -.3522367 -.7917610E-01 -.1591277 -.1591277
0. .3522367 .7959205 .2819172 .8078598E-01 .8078598E-01
0. .7917610E-01 .2819172 .3496329 .1908756E-02 .1908756E-02
0. .1591277 .8078598E-01 .1908756E-02 1.004455 .4455230E-02
0. .1591277 .8078598E-01 .1908756E-02 .4455230E-02 1.004455
0.
0.
0.

THE 1 BY 3 MATRIX CONSTR
-6.3339921E-03 .8332702E-03 -.4005468E-02

ICT=5

THE 1 BY 6 MATRIX X
-1.592573 1.447871 2.045149 .7694517 .7694517
0.
COST=.06129733491748
XI=.004005457550233
C=3.913312694022

THE 1 BY 3 MATRIX MU(X)
.4388478E-01 -.4018735E-01 .7254625E-02

ATX=.006712247649571
TX=.00004505426850917

NOACTIV=3
ARRAY IQACTV
I=1IQACTV=5
I=2IQACTV=1
I=3IQACTV=3

THE 1 BY 5 MATRIX QPLMD
-.3839339E-01 0. .6262692E-02 0. .3781163E-01
0.

THE 1 BY 5 MATRIX PI
-.4889454E-01 .5979293E-01 -.8084356E-01 .7885657E-03 .7885657E-03
0.
THETA=-.02158237475559
PASSED TEST OF STEP 3

THE 1 BY 5 MATRIX L11
.3891213E-03 -.1139282E-02 -.1091677E-02 -.1136660E-02 -.1136660E-02
0.
STEP LENGTH=1.

THE 1 BY 5 MATRIX NEWX
-1.641078 1.506525 1.963214 .7691034 .7691034
0.

C1=.002470339947282
C2=.00564195626111
C3=.01250665214867
C4=0.
D1=.634780440564E-7
IN THE NEW UPDATE SCHEME
PARAM=0.

THE 1 BY 5 MATRIX DLGR2
-.9218305E-02 .2247702E-01 -.8695641E-02 .1089091E-01 .1089091E-01
0.
C5=.002470389947282

THE 6 BY 6 MATRIX H
.3399883 -.5169033E-01 -.1250308 -.8666556E-01 -.8666556E-01
0. .5169033E-01 .7060696 .2623502 .9386135E-01 .9386135E-01
0. -.1250308 .2623502 .3681112 -.1890596E-01 -.1890596E-01
0. -.8666556E-01 .9386135E-01 -.1890596E-01 1.027161 .2716143E-01
0. -.8666556E-01 .9386135E-01 -.1890596E-01 .2716143E-01 1.027161
0. 0. 0. 0. 0.
0. 0. 0. 0. 0.

THE 1 BY 3 MATRIX CONSTR
.3856276E-05 .3057546E-04 -.4142871E-03

ICT=6

THE 1 BY 6 MATRIX X
-1.641078 1.506525 1.963214 .7691034 .7691034
0.
COST=.05563869985161
XI=.0004142870535146
C=3.913312694022

THE 1 BY 3 MATRIX MU(X)
.4159183E-01 -.3846059E-01 .6253509E-02
ATX=.001335960009204
TX=.000001784789146194

NOACTIV=3
ARRAY IQACTV
I=1IQACTV=5
I=2IQACTV=1
I=3IQACTV=3

THE 1 BY 5 MATRIX QPLMD
.3839331E-01 0. .4872954E-02 0. .3612570E-01
0.

THE 1 BY 6 MATRIX PI
-.6900967E-01 .8194801E-01 -.1187410 -.2337052E-02 -.2337052E-02
0.
THETA=-.006481169075554
PASSED TEST OF STEP 3

THE 1 BY 5 MATRIX L11
.1050521E-02 -.2293231E-02 -.2103443E-02 -.2266131E-02 -.2266131E-02
0.
STEP LENGTH=.9

THE 1 BY 6 MATRIX NEWX
-1.702336 1.578421 1.854643 .7651645 .7651645
0.

C1=.003544395959743
C2=.00378187853525
C3=.02074013925117
C4=0.
D1=3.701677542604E-8
IN THE NEW UPDATE SCHEME
PARAM=0.


```

THE 1 BY 5 MATRIX DLGR2
-.7689601E-02 .2232974E-01 -.1401046E-01 .4213017E-02 .4213017E-02
0.
C5=.003564395959743

THE 6 BY 6 MATRIX H
.3286023 -.3227520E-01 -.1309661 -.6957399E-01 -.6867399E-01
0. .3227520E-01 .6946653 .2619484 .5432789E-01 .5432789E-01
0. .1309661 .2619484 .3764352 -.4597362E-03 -.4597360E-03
0. .6867399E-01 .5482789E-01 -.4597362E-03 1.005927 .5926554E-02
0. .6867399E-01 .5482789E-01 -.4597360E-03 .5926554E-02 1.005927
0.
0. 0. 0. 0. 0.

THE 1 BY 3 MATRIX CONSTR
.1410726E-04 .2361484E-04 -.7651585E-03
*****
ICT=7

THE 1 BY 6 MATRIX X
-1.702336 1.578421 1.954643 .7651645 .7651645
0.
COST=.05405883507457
XI=.000785158465378
C=3.913312694022

THE 1 BY 3 MATRIX MU(X)
.4023977E-01 -.3784487E-01 .5383160E-02
ATX=.0006194770201331
TX=6.715425865263E-7

NOACTIV=3
ARRAY IQACTV
I=1IQACTV=5
I=2IQACTV=1
I=3IQACTV=3

THE 1 BY 6 MATRIX QPLMD
.3997863E-01 0. .5187072E-02 0. .3774545E-01
0.

THE 1 BY 5 MATRIX PI
-.1447739E-01 .1694478E-01 -.2658054E-01 -.1372925E-02 -.1372925E-02
0.
THETA=-.003250644220135
PASSED TEST OF STEP 3.

THE 1 BY 6 MATRIX L11
.6000106E-04 -.1087314E-03 -.9097477E-04 -.1052167E-03 -.1052167E-03
0.
STEP LENGTH=1.

THE 1 BY 6 MATRIX NEWX
-1.716753 1.595257 1.827971 .7635863 .7636863
0.
C1=.0001983595520441
C2=.0002028887631236
C3=.001207053124747
C4=.000001922569556169
D1=8.105875774674E-10
IN THE NEW UPDATE SCHEME
PARAM=0.

THE 1 BY 6 MATRIX DLGR2
-.1264015E-02 .4412800E-02 -.3910249E-02 -.5240984E-03 -.5240984E-03
0.
C5=.0001983595520441

THE 6 BY 6 MATRIX H
.3242773 -.2255888E-01 -.1352678 -.5197705E-01 -.6197705E-01
0. .2255888E-01 .6671957 .2642556 .3240793E-01 .3240794E-01
0. .1352678 .2642556 .3845628 .1779505E-01 .1779505E-01
0. .6197705E-01 .3290793E-01 .1779505E-01 1.006401 .6400915E-02
0. .6197705E-01 .3290794E-01 .1779505E-01 .6400915E-02 1.006401
0.
0. G. 0. 0. 0.

THE 1 BY 3 MATRIX CONSTR
.3779166E-07 -.1414138E-06 -.8594165E-05
*****
ICT=8

THE 1 BY 5 MATRIX X
-1.716753 1.595257 1.827971 .7636863 .7636863
0.
COST=.05394235861607
XI=.000008594165137765
C=3.913312694022

THE 1 BY 3 MATRIX MU(X)
.4016340E-01 -.3795333E-01 .5226606E-02
ATX=.000008618067287061
TX=7.42710837643E-11

```

NOACTIV=3
 ARRAY IOACTV
 I=1IOACTV=5
 I=2IOACTV=1
 I=3IOACTV=3

THE 1 BY 6 MATRIX OPLMD
 .4016187E-01 0. .5222521E-02 0. .3795668E-01
 0.

THE 1 BY 6 MATRIX PI
 -.3902231E-03 .4530520E-03 -.7257840E-03 -.4318365E-04 -.4318357E-04
 0.

THETA=-.00003381530404334
 PASSED TEST OF STEP 3

THE 1 BY 6 MATRIX L11
 .4638628E-07 -.7956427E-07 -.6545906E-07 -.7703363E-07 -.7699646E-07
 0.
 STEP LENGTH=1.

THE 1 BY 6 MATRIX NEWX
 -1.717144 1.595710 1.827246 .7636431 .7636431
 0.

C1=1.457570768824E-7
 C2=1.457236551349E-7
 C3=8.880221416985E-7
 C4=7.158161117092E-11
 D1=6.595435687065E-17
 IN THE NEW UPDATE SCHEME
 PARAM=0.

THE 1 BY 6 MATRIX DLGR2
 -.3130587E-04 .1147610E-03 -.1096107E-03 -.2309044E-04 -.2309042E-04
 0.
 C5=1.457570768824E-7

THE 6 BY 6 MATRIX H
 .3234384 -.2069782E-01 -.1363737 -.6103686E-01 -.6103684E-01
 0.
 .2069782E-01 .6646293 .2643526 .2881615E-01 .2881607E-01
 0.
 -.1363737 .2643526 .3866591 .2206013E-01 .2206021E-01
 0.
 -.6103686E-01 .2881615E-01 .2206013E-01 1.007923 .7922870E-02
 0.
 -.6103684E-01 .2881607E-01 .2206021E-01 .7922870E-02 1.007923
 0.
 0. 0. 0. 0. 0.

THE 1 BY 3 MATRIX CONSTR
 -.0151790E-11 -.2289369E-10 -.1324167E-09

 ICT=9

THE 1 BY 6 MATRIX X
 -1.717144 1.595710 1.827246 .7636431 .7636431
 0.

COST=-.05394984777046
 XI=1.324167442363E-10
 C=3.913312694022

THE 1 BY 3 MATRIX MU(X)
 .4016274E-01 -.3795778E-01 .5222642E-02

ATX=1.324179246412E-10
 TX=1.753450676627E-20

NOACTIV=3
 ARRAY IOACTV
 I=1IOACTV=5
 I=2IOACTV=1
 I=3IOACTV=3

THE 1 BY 6 MATRIX OPLMD
 .4016274E-01 0. .5222643E-02 0. .3795777E-01
 0.

THE 1 BY 6 MATRIX PI
 .9583909E-07 -.1005429E-06 .1613416E-06 .9698688E-08 .9602874E-08
 0.

THETA=-4.490652083967E-10
 PASSED TEST OF STEP 3

THE 1 BY 6 MATRIX L11
 .4547474E-12 .1918989E-11 -.3183231E-11 .5911716E-11 .9094947E-12
 0.
 STEP LENGTH=1.

THE 1 BY 6 MATRIX NEWX
 -1.717144 1.595710 1.827246 .7636431 .7636431
 0.

C1=7.195631614261E-15
C2=7.196383750086E-15
C3=4.386510491965E-14
C4=1.991401559337E-20
D1=7.691701135237E-34
IN THE NEW UPDATE SCHEME
PARAM=0.

THE 1 BY 5 MATRIX DLGR2
0.6941510E-08 - .2544678E-07 .2436766E-07 .5224090E-08 .5207895E-08
C5=7.195631614261E-15

THE 6 BY 5 MATRIX H

0.	.3233496	-.2056764E-01	-.1365300	-.5106542E-01	-.6598315E-01
0.	-.2056764E-01	.6648628	.2642433	.2877488E-01	.2847612E-01
0.	-.1365300	.2642433	.3866541	.2207619E-01	.2236296E-01
0.	-.5106542E-01	.2877488E-01	.2207618E-01	1.007930	.7991318E-02
0.	-.6098315E-01	.2847612E-01	.2236296E-01	.7991318E-02	1.009051
0.	0.	0.	0.	0.	0.
0.	0.	0.	0.	0.	0.

THE 1 BY 3 MATRIX CONSTR
-.1051603E-10 -.1043077E-10 .8185452E-11

I.5.3 Remarks

The apparent advantage of computing $\tilde{p}_i$ from the start, suggested by the lower number of iterations taken by the algorithm to converge, does not have a theoretical explanation, although it can be argued that as by definition $\tilde{p}_i \neq 0$ only if $\|\tilde{p}_i\| \leq \|p_i\|$ this test is perhaps a better indicator of when to start the computation of $\tilde{p}_i$ than the test $T(x_i) \leq \varepsilon$.

The table below summarizes the results obtained here for Problem A when using the new modification of the BFGS update, case A.3, and the results given by Powell (1982b) for a variable metric algorithm which incorporates the 'watchdog technique' [see Chamberlain et al (1982)] and updates the matrices by the scheme suggested by Powell (1977). Both algorithms started from the same initial point.

Algorithm	No. of iterations	Last point (x_f)	$\ x_f - x^*\ $
A.3	13	(4.828472; .3469473 x 10 ⁻⁷ ; 4.828427; 1.0; 2.414214; 2.414214; 1.0)	6.44704 x 10 ⁻⁷
Powell's	14	(4.830647; .004184; 4.826209; 1.002093; 2.414177; 2.416344; 1.0)	6.02268 x 10 ⁻³

The vector $x^* = (2 + 2\sqrt{2}; 0; 2 + 2\sqrt{2}; 1.0, 1 + \sqrt{2}; 1.0)$ is the exact solution of Problem A. The graph on the next page illustrates the behaviour of both algorithms.

We do not have results of variable metric algorithms applied to Problem B. In Maratos (1978) a quadratically convergent algorithm is used, starting from $\bar{x}_0 = (-2; 2; 2; -1; -1)$, and after 6 iterations it achieves a point x_6 such that $\|x_6 - x_M\| \leq 4 \times 10^{-14}$, with $x_M = (-1.71714; 1.59571; 1.82725; -.76364; -.76364)$ being an approximation of the exact solution x^* .

Here, case B2, the point $x_f = (-1.717144; 1.595710; 1.827246; .7636431; .7636431)$ is obtained after 8 iterations starting from $x'_0 = (1.1; 1; 1; 1)$. Clearly the signs of the last two components of x_f and x_M indicate the existence of two exact solutions x^* to Problem B. Note however that our starting point is farther from any x^* than

I.5.3 Remarks

In relation to the number of iterations, both versions of the Basic Algorithm compare well with reported applications of variable metric algorithms for solving Problem A.

For instance, the point achieved after 13 iterations when using the new update procedure and $\tilde{p}_i$ computed from the start, namely:

$$x_{13} = (4.828427; .3469473 \times 10^{-7}; 4.828427; 1.000000; 2.414214; 2.414214; 1.000000)$$

compares favourably with the point obtained by Powell (1982b) after 14 iterations, namely :

$$\bar{x}_{14} = (4.830647; .004184; 4.826209; 1.002093; 2.414177; 2.416344; 1.000000)$$

in the sense that some conditions that prevail at the solution x^* , namely :

$$x_1^* = x_3^* ; x_5^* = x_6^*$$

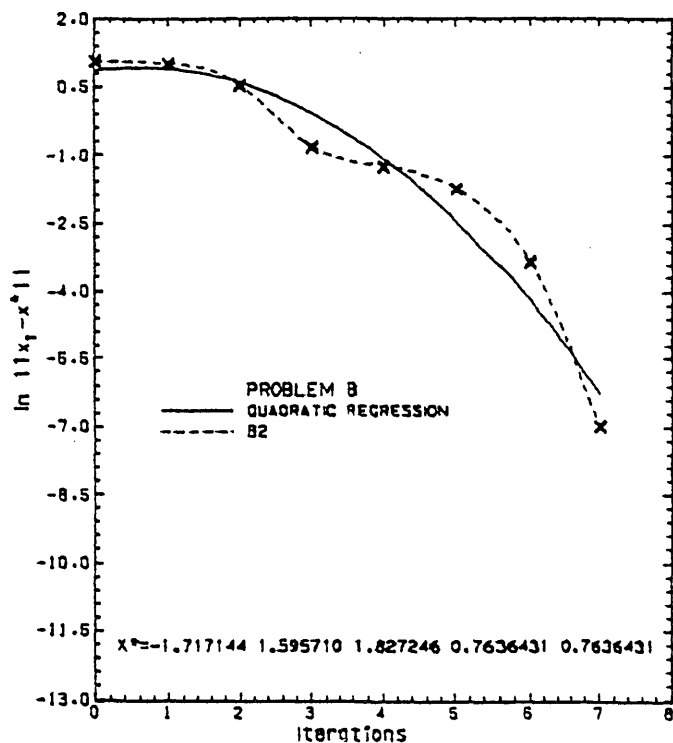
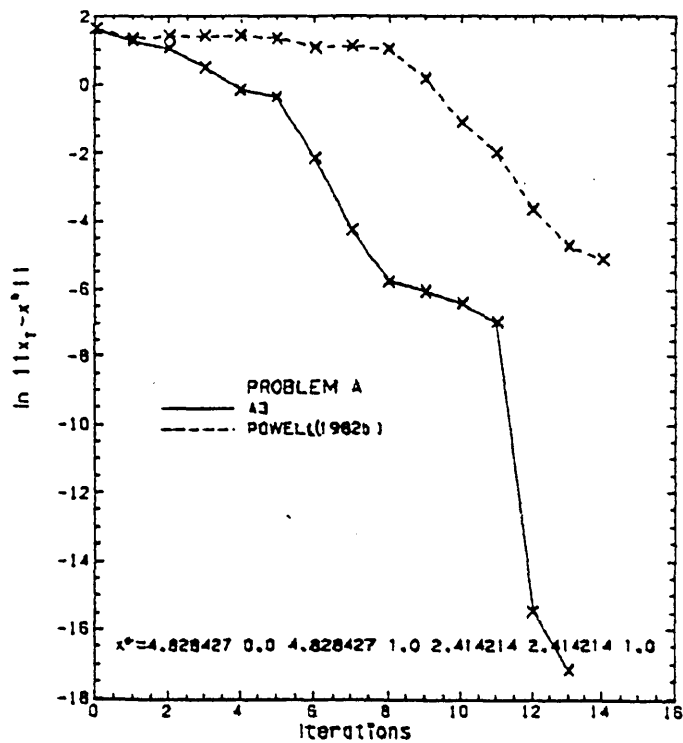
are here satisfied up to the sixth decimal digit.

Also, the number of iterations (11) required when using the slightly modified Powell's update - recall $\bar{\lambda}(\cdot)$ replacing λ_i in the definition of γ_i - is the same number required by the algorithm of Mayne and Polak (1982) when using the original version of Powell's update.

We do not have results of other variable metric algorithms applied to Problem B but the results here obtained seem to be satisfactory.

As to the apparent advantage of computing $\tilde{p}_i$ from the start,

in the Maratos case, that is: $\|x'_0 - x^*\| = 2.921265464$ and $\|\bar{x}_0 - x^*\| = .620511759$. A comparison of the performance of both algorithms is made more difficult by the fact that x^* is not exactly determined. Actually, if we call $v(x)$ the maximum violation of the constraints and $f(x)$ the value of the objective function, we have: $v(x_M) = 3.4 \times 10^{-5}$; $f(x_M) = .053951049$ and $v(x_F) = -1.324167 \times 10^{-10}$; $f(x_F) = .053949748$. That is, the point obtained here after 8 iterations is clearly a better approximation to x^* than the point adopted by Maratos (1978) as the optimum point. The graph below illustrates the performance of the algorithm presented here.



it does not have a theoretical explanation, although it can be argued that as by definition $\tilde{p}_i \neq 0$ only if

$$\|\tilde{p}_i\| \leq \|p_i\| \quad ,$$

this test is perhaps a better indication of when starting the computation of p_i than the test $T(x_i) \leq \epsilon$.

I.6 SUGGESTIONS FOR FURTHER RESEARCH

As pointed out in the introduction, the linearization of the constraints makes necessary a modification of the objective function of GMP to include information about the curvature of the constraints. An example given in Beale (1967) shows that the use, in the sub-problem which generates the search direction, of a linear approximation to a quadratic constraint leads to oscillatory behaviour. The question posed is whether the satisfaction of the constraints to second-order is sufficient to avoid such modification of the objective function $f(\cdot)$. In other words, whether the matrices H_i generated by the update could then be computed to approximate $\nabla_{xx} f(x^*)$ rather than $\nabla_{xx} L(x^*)$. In order to understand the problem better, consider, without loss of generality, that GMP has no equality constraints and define the two sub-problems below :

Q1ⁱ

$$\min_{p \in \mathbb{R}^n} f_x(x_i)p + \frac{1}{2}p^T [\nabla_{xx} f(x_i) + \sum_{j \in \underline{m}} (\bar{\lambda}(x_i))^j \nabla_{xx} g^j(x_i)] p$$

subject to:

$$g^j(x_i) + g_x^j(x_i)p \leq 0, \quad j \in \underline{m}.$$

Q2ⁱ

$$\min_{p \in \mathbb{R}^n} f_x(x_i)p + p^T \nabla_{xx} f(x_i)p$$

subject to :

$$g^j(x_i) + g_x^j(x_i)p + p^T \nabla_{xx} g^j(x_i)p \leq 0, \quad j \in \underline{m}.$$

Let (p_1, λ_1) and (p_2, λ_2) be Kuhn-Tucker couples of Q1ⁱ and

$Q2^i$ respectively.

At the solution of $Q1^i$ the following condition is satisfied :

I.6.I

$$f_x(x_i) + p_1^T [\nabla_{xx} f(x_i) + \sum_{j \in \underline{m}} (\bar{\lambda}(x_i))^j \nabla_{xx} g^j(x_i)] + \\ + \sum_{j \in \underline{m}} \lambda_1^j g_x^j(x_i) = 0.$$

At the solution of $Q2^i$ the following condition is satisfied :

$$f_x(x_i) + p_2^T \nabla_{xx} f(x_i) + \sum_{j \in \underline{m}} \lambda_2^j [g_x^j(x_i) + p_2^T \nabla_{xx} g^j(x_i)] = 0.$$

or, rewriting,

I.6.II

$$f_x(x_i) + p_2^T [\nabla_{xx} f(x_i) + \sum_{j \in \underline{m}} \lambda_2^j \nabla_{xx} g^j(x_i)] + \\ + \sum_{j \in \underline{m}} \lambda_2^j g_x^j(x_i) = 0.$$

If the second-order sufficiency conditions are satisfied at $x^* \in D$, with $x_i \rightarrow x^*$ as $i \rightarrow \infty$, for i sufficiently large both $Q1^i$ and $Q2^i$ have unique couples satisfying I.6.I and I.6.II, respectively. Moreover, as $\bar{\lambda}(x_i)$ and λ_2 both converge to λ^* , as $x_i \rightarrow x^*$, near a solution (x^*, λ^*) for GMP the vectors p_1 and p_2 must be very close. Thus, the use of $Q1^i$ and $Q2^i$ to generate the search direction are likely to give similar rates of convergence.

In practice, the use of $Q2^i$ has not much appeal because the quadratic constraints makes its solution far more complicated than the solution of a quadratic sub-problem. On the other hand, although

not requiring second-order derivatives, the search arc $\alpha p_i + \alpha^2 p_i$ defined in the first chapter satisfies the constraints to second-order. It is then a topic for further research to ask whether the use of the search arc ensures superlinear rate of convergence even if the objective function of the quadratic sub-problem does not involve the constraints.

The advantages of this procedure are emphasized when applying the algorithm of the second chapter. In contrast to the other algorithms of Part I, the function $\bar{c}(\cdot)$ is not used there to update the penalty parameter. Thus, the estimates $\bar{\lambda}(\cdot)$ and $\bar{\mu}(\cdot)$ are only necessary to compute the matrix B_i , either through the secant update or through the modification of the BFGS. Hence, if the simplified objective function is used in the quadratic sub-problem, the sometimes cumbersome computation of $\bar{\lambda}(\cdot)$ and $\bar{\mu}(\cdot)$ can be avoided.

Another aspect is the use of the simple procedure to update the penalty parameter, introduced in I.2, in the presence of non-convex inequalities. In theory, an infinite subsequence I may be generated where the constraints of the quadratic sub-problems QP_i^0 are inconsistent, for all $i \in I$. It may also happen that, for some infinite subsequence I' , $\|\lambda_i^0\| \xrightarrow{I'} \infty$. Consequently (see Proposition I.2.B), w_i may be greater than zero infinitely often and the penalty parameter may become excessively high.

Nevertheless, convergence can still be proven if the following conditions are included in the assumptions:

- The sequence of multipliers $\{\lambda_i^0\}$ remains bounded.
- for all i sufficiently large the constraints of QP_i^0 are consistent.

As these conditions are likely to be found in practice, it is perhaps advantageous apply the algorithm of I.2 even if the constraints of GMP are not convex. Further research is necessary on this topic.

The asymptotic properties of the sequence of matrices $\{H_i\}$ generated by the update of I.4.VI is another topic where further research is required. For the unconstrained minimization of convex functions, Powell (1971) shows that the variable metric method with the BFGS update is globally convergent and *afterwards* it is established that the positive eigenvalues of the sequence $\{H_i\}$ remain bounded from above and bounded away from zero. Recently, Ren-Pu and Powell (1982) show that the sequence $\{H_i\}$ generated by the BFGS update in unconstrained minimization actually converges to some $H^* \in L(R^n)$.

Now in the same way as Theorem 1 of Powell (1971), Theorem I.4.N is proven without requiring a priori either the boundedness of $\{H_i\}$ or that its positive eigenvalues remain bounded away from zero. It is then useful to investigate whether results similar to those of Powell (1971) can be proven for the update given by I.4.VI and to extend the results of Ren-Pu and Powell (1982) to constrained problems.

Finally, although the numerical examples of the last chapter are encouraging, more applications of the algorithms suggested here are required before asserting their competitiveness.

II.1 UNCONSTRAINED CONTROL CASE

II.1.1 Introduction

We deal in this chapter with the special case of the discrete optimal control problem (DCP) where there exists no constraints. That is :

$$\min_{u \in R^{n \times N}} \left\{ \sum_{i=0}^{N-1} T_i(x_i, u_i) + \phi(x_N) \right\}$$

with the dynamic system:

$$\begin{aligned} x_{i+1} &= f_i(x_i, u_i), \quad i = 0, N-1 \\ x_0 &= x^0, \text{ given} \end{aligned}$$

where:

$$T_i : R^n \times R^m \rightarrow R ; \quad \Phi : R^n \rightarrow R ; \quad f_i : R^n \times R^m \rightarrow R^n$$

$x_i \in R^n$ being the state vector

$u_i \in R^m$ being the control vector , $i = 0, \dots, N-1$.

Now, defining an augmented state vector with the introduction of

$$x_{i+1}^{n+1} = f_i^{n+1}(x_i, u_i), \quad i = 1, \dots, N-1; \quad x_0^{n+1} = 0$$

with

$$f_i^{n+1}(x_i, u_i) = x_i^{n+1} + T_i(x_i, u_i)$$

and a objective function $F(x_N)$ given by:

$$F(x_N) = x_N^{n+1} + \phi(x_N)$$

it is readily seen that we can, without loss of generality, consider the unconstrained discrete control problem in the following form.

UDCP

$$\min_{u \in R^{m \times N}} F(x_N)$$

with the dynamic system : $x_{i+1} = f_i(x_i, u_i)$, $i = 0, \dots, N-1$
 $x_0 = x^0$, given

where:

$$F : R^n \rightarrow R \quad ; \quad f_i : R^n \times R^m \rightarrow R^n.$$

Let the cost due to the initial condition $x_0 = x^0$ and control sequence $\{u_i\}$ be expressed by :

$$V_0(x_0, \{u_i\}) = F(x_N)$$

with:

$$x_0 = x^0 \quad ; \quad x_{i+1} = f_i(x_i, u_i), \quad i = 0, \dots, N-1.$$

Now, to employ dynamic programming, it is necessary to define the cost of a trajectory with initial condition x_i (at time i) and control schedule $\{u_k\} = \{u_i, u_{i+1}, \dots, u_{N-1}\}$. This cost is expressed by :

$$V_i(x_i, \{u_k\}) = F(x_N)$$

with:

$$x_{k+1} = f_k(x_k, u_k) \quad , \quad k = i, \dots, N-1.$$

Let π denote any policy (i.e., sequence of control laws $u_k = h_k(x_k)$, $k = 0, \dots, N-1$) and ${}^0\pi$ the optimal policy.

Also, let ${}^0V_i(x_i)$ denotes the cost due to initial condition x_i (at time i) if the optimal policy is employed.

Then, by the 'principle of optimality' [see Bellman (1957)] we have :

II.1.1

$${}^0V_i(x_i) = \min_{u_i \in R^m} \{ {}^0V_{i+1}(f_i(x_i, u_i)) \} ,$$

Performing the minimization indicated above for all $x_i \in R^n$, with the boundary condition ${}^0V_N(x_N) = F(x_N)$, yields the optimal cost function ${}^0V_i(\cdot)$ and the optimal control law ${}^0h_i(\cdot)$, at time i . Iteration of this procedure for $i = N-1, \dots, 0$ yields the optimal control policy ${}^0\pi = \{ {}^0h_0(\cdot), \dots, {}^0h_{N-1}(\cdot) \}$ and, given the initial condition $x_0 = x^0$, the problem is solved. On the other hand, the performance of the minimization II.1.1 and the storage of ${}^0V_i(x_i)$, ${}^0h_i(x_i)$ for all $x_i \in R^n$ is usually impossible.

Instead of finding a solution to the UDCP in one iteration, as in the dynamic programming, we are interested in iterative methods. That is :

given a 'nominal control sequence' $\{\bar{u}_i\}$ and a correspondent 'nominal trajectory' $\{\bar{x}_i\} = 0, \dots, N-1$, find a 'successor control sequence' $\{u_i\}$ and a correspondent 'successor trajectory' $\{x_i\}$, $i = 0, \dots, N-1$, such that an improvement is obtained towards the solution of UDCP. The new control and state sequences become the nominal sequences for the next iteration.

In the Differential Dynamic Programming method [see Jacobson and Mayne (1970)] the measurement of such improvement is done in a neighbourhood of $\{\bar{x}_i\}$, where the policy $\pi = \{h_0(\cdot), \dots, h_{N-1}(\cdot)\}$ is defined by :

$$\begin{aligned} u_i &= \bar{u}_i + \delta u_i \\ \delta u_i &= \alpha_i + \beta_i \delta x_i \\ \delta x_i &= x_i - \bar{x}_i . \end{aligned}$$

for $i = 0, \dots, N-1$.

In the second-order algorithm of Jacobson and Mayne (1970), , the function $V_i(x_i, \pi)$, which denotes the cost due to initial condition x_i (at time i) and policy π , is approximated to second order by Taylor series and this approximation is minimized, for $i = N-1, \dots, 0$, to generate the sequences $\{\alpha_i\}$ and $\{\beta_i\}$. Omitting the argument π for notational simplicity, the Taylor series coefficients $V_x^i(\bar{x}_i)$ and $V_{xx}(\bar{x}_i)$ are estimated through the following recursive equations :

II.1.II

$$V_x^i(\bar{x}_i) = H_x^i + \beta_i^T H_u^i$$

II.1.III

$$V_{xx}^i(\bar{x}_i) = A_i - \beta_i^T C_i \beta_i$$

where:

$$H_i(x_i, u_i, V_x^{i+1}(\bar{x}_{i+1})) = [V_x^{i+1}(\bar{x}_{i+1})]^T f_i(x_i, u_i)$$

so that :

$$H_x^i = [f_x^i]^T V_x^{i+1}(\bar{x}_{i+1}) ; H_u^i = [f_u^i]^T V_x^{i+1}(\bar{x}_{i+1}),$$

the unspecified arguments being $\bar{x}_i, \bar{u}_i, V_x^{i+1}(\bar{x}_{i+1})$.

Also,

$$\beta_i = C_i^{-1} B_i$$

and,

II.1.IV

$$A_i = [f_{xx}^i]^T V_x^{i+1}(\bar{x}_{i+1}) + [f_x^i]^T V_{xx}^{i+1}(\bar{x}_{i+1}) [f_x^i]$$

$$B_i = [f_{ux}^i]^T V_x^{i+1}(\bar{x}_{i+1}) + [f_u^i]^T V_{xx}^{i+1}(\bar{x}_{i+1}) [f_x^i]$$

$$C_i = [f_{uu}^i]^T V_x^{i+1}(\bar{x}_{i+1}) + [f_u^i]^T V_{xx}^{i+1}(\bar{x}_{i+1}) [f_u^i] .$$

The boundary conditions are ;

$$V_x^N(\bar{x}_N) = F_x(\bar{x}_N) : V_{xx}^N(\bar{x}_N) = F_{xx}(\bar{x}_N).$$

The sequence $\{\alpha_i\}$ is computed backwards as :

II.1.V

$$\alpha_i = \arg \min_{\alpha \in R} \{ [H_u^i]^T \alpha + \frac{1}{2} \alpha^T C_i \alpha \} \Rightarrow \alpha_i = -C_i^{-1} H_u^i.$$

In order to enforce global convergence, the algorithm suggested by Jacobson and Mayne (1970) incorporates a step-length procedure using the estimated change in cost of the new policy given by:

II.1.VI

$$a_0 = - \sum_{i=0}^{N-1} [H_u^i]^T C_i^{-1} [H_u^i] .$$

The complete algorithm is given below.

- 1 - Given initial condition x^0 and schedule $\{\bar{u}_i\}$, calculate $\{\bar{x}_i\}$. Set $\epsilon = 1$.
- 2 - Calculate $\{V_x^i(\bar{x}_i)\}$ and $\{V_{xx}^{i+1}(\bar{x}_i)\}$ using II.1.II, II.1.III and the boundary conditions. Store $\alpha_i = -C_i^{-1} H_u^i$, $\beta_i = -C_i^{-1} B_i$, $i = N-1, \dots, 0$.
- 3 - Compute : $x_{i+1} = f_i(\bar{x}_i, u_i)$, $x_0 = x^0$.

$$u_i = \bar{u}_i + \epsilon \alpha_i + \beta_i (x_i - \bar{x}_i)$$
- 4 - Calculate a_0 using II.1.VI and the actual change in cost :

$$\Delta V_0 = F(x_N) - F(\bar{x}_N)$$

$$5 - \quad \text{If } \Delta V_0 > 0 \quad \text{or} \quad \frac{|\Delta V_0|}{|\epsilon(1-\epsilon/2)a_0|} < c, \quad 0 < c < 1,$$

set $\epsilon = \epsilon/2$ and repeat Step 3.

$$6 - \quad \text{Set } \bar{u}_i = u_i, \quad i = 0, \dots, N-1; \quad \bar{x}_i = x_i, \quad i = 0, \dots, N$$

and repeat Step 1.

Jacobson and Mayne (1970) do not discuss the global convergence of this algorithm but indicated how it could be proved. Although a formal proof can be obtained as a special case of the results of the next chapter, here we will discuss the rate of convergence of the algorithm and its relationship with Newton's method when applied to the UDCP.

A slightly modified version of the DDP method is presented in the next section and the successor sequences $(\{u_i\}, \{x_i\})$ it computes is later proved to be identical to the successor obtained through Newton's method. As a result, it is shown that the original version of the DDP method does not in general coincide with Newton's method, except when the dynamic system is linear.

Finally, it is proved that a quadratic rate of convergence is also achieved by the original version of the DDP method.

II.1.2 Modified DDP/Newton's method

Modified DDP

Consider the algorithm of Jacobson and Mayne (1970) which was stated in the previous section. The addition of the term $\beta_i^T H_u^i$ to H_x^i in expression II.1.II, and of $[-\beta_i^T C_i \beta_i]$ to A_i in II.1.III, is meant to send information backwards and is clearly important

for obtaining a superlinear rate of convergence. But, since the corrected estimate $V_x^i(\bar{x}_i)$ is also used to compute equations II.1.IV the term $\beta_i^T H_u^i$ is also affecting the computation of the matrices A_i, B_i, C_i . As shown on pages 108-109 of Jacobson and Mayne (1970), the error on the estimate $V_{xx}^i(\bar{x}_i)$ is already of order α_i . Hence, as H_u^i and α_i are of the same order (recall that $\alpha_i = -C_i^{-1} H_u^i$), no advantage is obtained using the correction $\beta_i^T H_u^i$ in the equations for estimating the second-order derivatives. Another way of seeing this is to observe that $H_u^i \rightarrow 0$ as $\bar{u}_i \rightarrow u^*$ (desirable), so that with or without the term $\beta_i^T H_u^i$ the sequences of matrices $\{A_i\}, \{B_i\}$ and $\{C_i\}$ will converge to the same matrices. Therefore, the rate of convergence to the algorithm is not expected to be affected if, instead of using II.1.IV, we define the matrices A_i, B_i, C_i as follows:

II.1.VII

$$\begin{aligned} A_i &= [f_{xx}^i]^T \bar{V}_x^{i+1}(\bar{x}_{i+1}) + [f_x^i]^T V_{xx}^{i+1}(\bar{x}_{i+1}) [f_x^i] \\ B_i &= [f_{ux}^i]^T \bar{V}_x^{i+1}(\bar{x}_{i+1}) + [f_u^i]^T V_{xx}^{i+1}(\bar{x}_{i+1}) [f_x^i] \\ C_i &= [f_u^i]^T \bar{V}_x^{i+1}(\bar{x}_{i+1}) + [f_u^i]^T V_{xx}^{i+1}(\bar{x}_{i+1}) [f_u^i] \end{aligned}$$

where: $\bar{V}_x^i = f_x^i \bar{V}_x^{i+1}$

with

$$\bar{V}_x^N = F_x(x_N).$$

Another point to be discussed is the use of δx_i in Step 3, when constructing onwards the new control sequence. The method of Jacobson and Mayne (1970) uses a linear control policy of the form:

$$\delta u_i = \alpha_i + \beta_i \delta x_i.$$

This is then justified (see pg 19) as follows. Because the equation

$$V_i^0(x_i) = \min_{u_i} [V_{i+1}^0(x_{i+1})]$$

is expressed to second order only when generating the recursive equations II.1.II and II.1.III, there is no point in finding a relationship between δu_i and δx_i of order higher than linear. Suppose otherwise that δu_i is computed as :

$$\delta u_i = \alpha_i + \beta_i \delta x_i + \varepsilon^2(\delta x_i) + \text{terms of higher order.}$$

where,

$$\varepsilon^2(x_i) \text{ indicated a } 2^{\text{nd}} \text{ order term in } \delta x_i.$$

Because of the differentiability assumptions and the mean value theorem, when neglecting $\varepsilon^2(\delta x_i)$ in the above equation, the DDP method is neglecting second order terms in $(u_j - \bar{u}_j)$, $j < i$.

Therefore, after deciding to use β_i as the coefficient in a linear policy, there is no gain in accuracy from multiplying β_i by δx_i instead of the first order estimate Δ_i given below:

II.1.VIII

$$\Delta_i \triangleq f_x^{i-1} f_x^{i-2} \dots f_x^1 f_u^0(u_0 - \bar{u}_0) + f_x^{i-1} f_x^{i-2} f_x^2 f_u^1(u_1 - \bar{u}_1) + \dots + f_u^{i-1}(u_{i-1} - \bar{u}_{i-1}).$$

The modified version of the DDP method consists in the same algorithm given in the previous section, with two modifications:

- Compute A_i, B_i, C_i through II.1.V II, at Step 2
- Set,

$$u_i = \bar{u}_i + \epsilon \alpha_i + \beta_i^T \Delta_i, \text{ at Step 3.}$$

Newton's method

Using the dynamic equations and initial condition, we can transform the UDCP, given in II.1.1, into the equivalent problem below:

P2

$$\min_{(u \in \mathbb{R}^{m \times n})} \tilde{F}(u) \triangleq F(f_{N-1}(f_{N-2}(\dots f_0(x_0, u_0), u_1), \dots, u_{N-1})).$$

Given a nominal sequence $\bar{u}$, the new control obtained through Newton's method is :

$$u^N = \arg \min_{(u \in \mathbb{R}^{m \times N})} \phi^N(u)$$

where $\phi^N(\cdot)$ is the following quadratic function :

$$\phi^N(u) = \tilde{F}_u(\bar{u})^T (u - \bar{u}) + \frac{1}{2} (u - \bar{u})^T \tilde{F}_{uu}(\bar{u}) (u - \bar{u})$$

with

$$\tilde{F}_u(\bar{u}) = \begin{pmatrix} \tilde{F}_{u_0}(\bar{u}) \\ \vdots \\ \tilde{F}_{u_{N-1}}(\bar{u}) \end{pmatrix} \quad \text{and} \quad \tilde{F}_{uu}(\bar{u}) = \begin{pmatrix} \tilde{F}_{u_0 u_0}(\bar{u}) \dots \tilde{F}_{u_0 u_{N-1}}(\bar{u}) \\ \vdots \\ \tilde{F}_{u_{N-1} u_0}(\bar{u}) \dots \tilde{F}_{u_{N-1} u_{N-1}}(\bar{u}) \end{pmatrix}$$

Omitting the arguments $(\bar{x}_i, \bar{u}_i)$, from the definition of $F(\cdot)$ we obtain :

II.1.IX

$$\tilde{F}_{u_i} = (f_u^i)^T (f_x^i)^T \dots (f_x^{N-1})^T F_x,$$

and, for $j > i$.

II.1.X

$$\begin{aligned} \tilde{F}_{u_i u_j} &= (f_u^i)^T (f_x^{i+1})^T \dots (f_x^{j-1})^T (f_{ux}^j)^T (f_x^{j+1})^T \dots (f_x^{N-1})^T F_x \\ &+ (f_u^i)^T (f_x^{i+1})^T \dots (f_x^{j-1})^T \{ (f_u^j)^T [(f_{xx}^{j+1})^T (f_x^{j+2})^T (f_x^{j+3})^T \\ &\dots (f_x^{N-1})^T F_x \\ &+ (f_x^{j+1})^T (f_{xx}^{j+2})^T (f_x^{j+3})^T \dots (f_x^{N-1})^T F_x + (f_x^{j+1})^T \\ &(f_x^{j+2})^T (f_{xx}^{j+3})^T (f_x^{j+4})^T \dots (f_x^{N-1})^T F_x \\ &+ \dots + (f_x^{j+1})^T (f_x^{j+2})^T \dots (f_x^{N-1})^T F_{xx} (f_x^{N-1})^T \dots \\ &(f_x^{j+1})^T] f_x^j \}. \end{aligned}$$

Also,

II.i.XI

$$\begin{aligned} \tilde{F}_{u_i u_i} &= (f_{uu}^i)^T (f_x^{i+1})^T \dots (f_x^{N-1})^T F_x + (f_u^i)^T (f_{xx}^{i+1})^T (f_x^{i+2})^T \dots \\ &(f_x^{N-1})^T F_x \\ &+ (f_x^{i+1})^T (f_{xx}^{i+2})^T (f_x^{i+3})^T \dots (f_x^{N-1})^T F_x + \dots + \\ &(f_x^{i+1})^T (f_x^{i+2})^T \dots F_{xx} f_x^{N-1} \dots f_x^{i+1}] f_u^i. \end{aligned}$$

Observation

$$(f_x^i)^T (f_x^{i+1})^T \dots (f_x^{N-1})^T F_x = \sum_{j=1}^n \{ [f_j^i]^T [(f_x^{i+1})^T \dots (f_x^{N-1})^T F_x]_j \}$$

and

$$(f_{uu}^i)^T (f_x^{i+1})^T \dots (f_x^{N-1})^T F_x = \sum_{j=1}^n \{ [f_j^i]^T_{uu} [(f_x^{i+1})^T \dots (f_x^{N-1})^T F_x]_j \}.$$

where the subscript j indicates j^{th} component of the n -vectors.

II.1.3 Results

Assume the functions F , f all have the required continuous derivatives and let :

- u^D denotes the successor obtained through the original version of the DDP method ($\epsilon = 1$)
- u' denotes the successor obtained through the modified version of the DDP method ($\epsilon = 1$)
- u^N denotes the successor obtained through Newton's method.

Theorem II.1.A

Let $\tilde{F}_{uu}(\bar{u})$ be positive definite.

Then,

the two successors u' , u^N coincide.

Proof:

From II.1.II, II.1.III and II.1.VII we have :

$$(i) \quad H_u^i = (f_u^i)^T (f_x^i)^T \dots (f_x^{N-1})^T F_x + (f_u^i)^T (f_x^{i+1})^T \dots (f_x^{N-2})^T \beta_{N-1}^T H_u^{N-1} \\ + (f_u^i)^T (f_x^{i+1})^T \dots (f_x^{N-3})^T \beta_{N-2}^T H_u^{N-2} + \dots + (f_u^i)^T \beta_{i+1}^T H_u^{i+1}.$$

$$\begin{aligned}
 \text{(ii)} \quad B_i &= (f_{ux}^i)^T (f_x^{i+1})^T \dots (f_x^{N-1})^T F_x + (f_u^i)^T [(f_{xx}^{i+1})^T (f_x^{i+2})^T \dots \\
 &\quad (f_x^{N-1})^T F_x \\
 &\quad + (f_x^{i+1})^T (f_{xx}^{i+2})^T (f_x^{i+3})^T (f_x^{i+4})^T \dots (f_x^{N-1})^T F_x (f_x^{i+1}) \\
 &\quad \vdots \\
 &\quad + (f_x^{i+1})^T (f_x^{i+2})^T \dots (f_x^{N-1})^T F_{xx} (f_x^{N-1}) \dots \\
 &\quad (f_x^{i+1})] (f_u^i) \\
 &\quad - (f_u^i)^T (f_x^{i+1})^T \dots (f_x^{N-2})^T [\beta_{N-1}^T C_{N-1} \beta_{N-1}] (f_x^{N-2}) \dots (f_x^i) \\
 &\quad - (f_u^i)^T (f_x^{i+1})^T \dots (f_x^{N-3})^T [\beta_{N-2}^T C_{N-2} \beta_{N-2}] (f_x^{N-3}) \dots (f_x^i) \\
 &\quad - (f_u^i)^T (f_x^{i+1})^T \dots (f_x^{N-4})^T [\beta_{N-3}^T C_{N-3} \beta_{N-3}] (f_x^{N-4}) \dots (f_x^i) \\
 &\quad \vdots \\
 &\quad - (f_u^i)^T [\beta_{i+1}^T C_{i+1} \beta_{i+1}] (f_x^i).
 \end{aligned}$$

$$\begin{aligned}
 \text{(iii)} \quad C_i &= (f_{uu}^i)^T (f_x^{i+1})^T \dots (f_x^{N-1})^T F_x + (f_u^i)^T [(f_{xx}^{i+1})^T (f_x^{i+2})^T \dots \\
 &\quad (f_x^{N-1})^T F_x \\
 &\quad + (f_x^{i+1})^T (f_{xx}^{i+2})^T (f_x^{i+3})^T \dots (f_x^{N-1})^T F_x (f_x^{i+1}) \\
 &\quad + (f_x^{i+1})^T (f_x^{i+2})^T (f_{xx}^{i+3})^T (f_x^{i+4})^T \dots (f_x^{N-1})^T F_x (f_x^{i+2}) (f_x^{i+1}) \\
 &\quad \vdots \\
 &\quad + (f_x^{i+1})^T (f_x^{i+2})^T \dots (f_x^{N-1})^T F_{xx} (f_x^{N-1}) \dots (f_x^{i+1})] (f_u^i) \\
 &\quad - (f_u^i)^T (f_x^{i+1})^T \dots (f_x^{N-2})^T [\beta_{N-1}^T C_{N-1} \beta_{N-1}] (f_x^{N-2}) \dots \\
 &\quad (f_x^{i+1}) (f_u^i)
 \end{aligned}$$

$$\begin{aligned}
& - (f_u^i)^T (f_x^{i+1})^T \dots (f_x^{N-3})^T [\beta_{N-2}^T \ C_{N-2} \ \beta_{N-2}] (f_x^{N-3}) \dots \\
& \qquad \qquad \qquad (f_x^{i+1}) (f_u^i) \\
& - (f_u^i)^T (f_x^{i+1})^T \dots (f_x^{N-4})^T [\beta_{N-3}^T \ C_{N-3} \ \beta_{N-3}] (f_x^{N-4}) \dots \\
& \qquad \qquad \qquad \vdots \qquad \qquad \qquad (f_x^{i+1}) (f_u^i) \\
& - (f_u^i)^T [\beta_{i+1}^T \ C_{i+1} \ \beta_{i+1}] (f_u^i).
\end{aligned}$$

From II.1.IX and (i) we obtain :

$$\begin{aligned}
\text{(iv)} \quad \tilde{F}_{u_i} &= H_u^i - (f_u^i)^T (f_x^{i+1})^T \dots (f_x^{N-2})^T \beta_{N-1}^T H_u^{N-1} - (f_u^i)^T (f_x^{i+1})^T \dots \\
& \qquad \qquad \qquad (f_x^{N-3})^T \beta_{N-2}^T H_u^{N-2} \\
& - (f_u^i)^T (f_x^{i+1})^T \dots (f_x^{N-4})^T \beta_{N-3}^T H_u^{N-3} - \dots - (f_u^i)^T \beta_{i+1}^T H_u^{i+1} .
\end{aligned}$$

From II.1.X and (ii) we obtain :

$$\begin{aligned}
\text{(v)} \quad \tilde{F}_{u_i u_j} &= (f_u^i)^T (f_x^{i+1})^T \dots (f_x^{j-1})^T \beta_j^T + (f_u^i)^T (f_x^{i+1})^T \dots \\
& \qquad \qquad \qquad (f_x^{N-2})^T [\beta_{N-1}^T \ C_{N-1} \ \beta_{N-1}] (f_x^{N-2}) \dots \\
& \qquad \qquad \qquad (f_x^{j+1}) (f_u^j) \\
& + (f_u^i)^T (f_x^{i+1})^T \dots (f_x^{N-3})^T [\beta_{N-2}^T \ C_{N-2} \ \beta_{N-2}] (f_x^{N-3}) \dots \\
& \qquad \qquad \qquad (f_x^{j+1}) (f_u^j) \\
& \qquad \qquad \qquad \vdots \\
& + (f_u^i)^T (f_x^{i+1})^T \dots (f_x^j)^T [\beta_{j+1}^T \ C_{j+1} \ \beta_{j+1}] (f_u^j) .
\end{aligned}$$

From II,1,XI and (iii) we obtain ;

$$\begin{aligned}
 \text{(vi)} \quad \tilde{F}_{u_i u_i} &= C_i + (f_u^i)^T (f_x^{i+1})^T \dots (f_x^{N-2})^T [\beta_{N-1}^T C_{N-1} \beta_{N-1}] (f_x^{N-2}) \dots \\
 &\quad (f_x^{i+1}) (f_u^i) \\
 &+ (f_u^i)^T (f_x^{i+1})^T \dots (f_x^{N-3})^T [\beta_{N-2}^T C_{N-2} \beta_{N-2}] (f_x^{N-3}) \dots \\
 &\quad (f_x^{i+1}) (f_u^i) \\
 &+ (f_u^i)^T (f_x^{i+1})^T \dots (f_x^{N-4})^T [\beta_{N-3}^T C_{N-3} \beta_{N-3}] (f_x^{N-4}) \dots \\
 &\quad (f_x^{i+1}) (f_u^i) \\
 &\quad \vdots \\
 &+ (f_u^i)^T [\beta_{i+1}^T C_{i+1} \beta_{i+1}] (f_u^i)
 \end{aligned}$$

Now,

$$u'_i = \bar{u}_i + \alpha_i + \beta_i^T \Delta_i.$$

Also,

$$\alpha_i = C_i^{-1} H_u^i; \quad \beta_i = -C_i^{-1} B_i.$$

Hence,

$$\text{(vii)} \quad H_u^i = -C_i (u'_i - \bar{u}_i) - B_i \Delta_i.$$

From (iv) and (vii) we obtain :

$$\begin{aligned}
 \text{(viii)} \quad \tilde{F}_{u_i} &= H_u^i - (f_u^i)^T (f_x^{i+1})^T \dots (f_x^{N-2})^T B_{N-1}^T (u'_{N-1} - \bar{u}_{N-1}) \\
 &\quad - (f_u^i)^T (f_x^{i+1})^T \dots (f_x^{N-2})^T [\beta_{N-1}^T C_{N-1} \beta_{N-1}] \Delta_{N-1}.
 \end{aligned}$$

$$\begin{aligned}
& - (f_u^i)^T (f_x^{i+1})^T \dots (f_x^{N-3})^T B_{N-2}^T (u_{N-2}' - \bar{u}_{N-2}) \\
& - (f_u^i)^T (f_x^{i+1})^T \dots (f_x^{N-3})^T [\beta_{N-2}^T C_{N-2} \beta_{N-2}] \Delta_{N-2} \\
& - (f_u^i)^T (f_x^{i+1})^T \dots (f_x^{N-2})^T B_{N-3}^T (u_{N-3}' - \bar{u}_{N-3}) \\
& - (f_u^i)^T (f_x^{i+1})^T \dots (f_x^{N-4})^T [\beta_{N-3}^T C_{N-3} \beta_{N-3}] \Delta_{N-3} \\
& \quad \vdots \\
& - (f_u^i)^T B_{i+1}^T (u_{i+1}' - \bar{u}_{i+1}) - (f_u^i)^T [\beta_{i+1}^T C_{i+1} \beta_{i+1}] \Delta_{i+1} .
\end{aligned}$$

Recall that :

$$\begin{aligned}
\Delta_i = & f_x^{i-1} f_x^{i-2} \dots f_x^1 f_u^0 (u_0' - \bar{u}_0) + f_x^{i-1} \dots f_x^2 f_u^1 (u_1' - \bar{u}_1) + \\
& + \dots + f_u^{i-1} (u_{i-1}' - \bar{u}_{i-1}) .
\end{aligned}$$

Then, from (v) and (viii) we have :

$$\begin{aligned}
(ix) \quad H_u^i = & \tilde{F}_{u_i} + \tilde{F}_{u_i u_{i+1}} (u_{i+1}' - \bar{u}_{i+1}) + \tilde{F}_{u_i u_{i+2}} (u_{i+2}' - \bar{u}_{i+2}) + \\
& + \dots + \tilde{F}_{u_i - u_{N-1}} (u_{N-1}' - \bar{u}_{N-1}) \\
& + (f_u^i)^T (f_x^{i+1})^T \dots (f_x^{N-2})^T [\beta_{N-1}^T C_{N-1} \beta_{N-1}] [f_x^{N-2} \dots f_x^{i+1} \\
& \quad f_u^i (u_i' - \bar{u}_i) \\
& + f_x^{N-2} f_x^{N-3} \dots f_x^i f_u^{i-1} (u_{i-1}' - \bar{u}_{i-1}) + \dots + f_x^{N-2} \dots \\
& \quad f_x^1 f_u^0 (u_0' - \bar{u}_0)]
\end{aligned}$$

$$\begin{aligned}
& + (f_u^i)^T (f_x^{i+1})^T \dots (f_x^{N-3})^T [\beta_{N-2}^T C_{N-2} \beta_{N-2}] [f_x^{N-3} \dots \\
& \qquad \qquad \qquad f_x^{i+1} f_u^i (u_i' - \bar{u}_i)] \\
& + f_x^{N-3} \dots f_x^i f_u^{i-1} (u_{i-1}' - \bar{u}_{i-1}) + \dots + f_x^{N-3} \dots \\
& \qquad \qquad \qquad f_x^1 f_u^0 (u_0' - \bar{u}_0)] \\
& \qquad \qquad \qquad \vdots \\
& + (f_u^i)^T [\beta_{i+1}^T C_{i+1} \beta_{i+1}] [f_u^i (u_i' - \bar{u}_i) + f_x^i f_u^{i-1} (u_{i-1}' - \bar{u}_{i-1}) \\
& + f_x^i \dots f_x^1 f_u^0 (u_0' - \bar{u}_0)]
\end{aligned}$$

From (vi) and (ix) we have :

$$\begin{aligned}
(x) \quad H_u^i + C_i (u_i' - \bar{u}_i) &= \tilde{F}_{u_i} + \tilde{F}_{u_i u_i} (u_i' - \bar{u}_i) + \tilde{F}_{u_i u_{i+1}} (u_{i+1}' - \bar{u}_{i+1}) \\
&+ \dots + \tilde{F}_{u_i u_{N-1}} (u_{N-1}' - \bar{u}_{N-1}) \\
&+ (f_u^i)^T (f_x^{i+1})^T \dots (f_x^{N-2})^T [\beta_{N-1}^T C_{N-1} \beta_{N-1}] [f_x^{N-2} \dots \\
& \qquad \qquad \qquad f_x^i f_u^{i-1} (u_{i-1}' - \bar{u}_{i-1}) \\
& + f_x^{N-2} \dots f_x^{i-1} f_u^{i-2} (u_{i-2}' - \bar{u}_{i-2}) + \dots + f_x^{N-2} \dots f_x^1 \\
& \qquad \qquad \qquad f_u^0 (u_0' - \bar{u}_0)] \\
& + (f_u^i)^T (f_x^{i+1})^T \dots (f_x^{N-3})^T [\beta_{N-2}^T C_{N-2} \beta_{N-2}] [f_x^{N-3} \dots f_x^i f_u^{i-1} (u_{i-1}' - \bar{u}_{i-1}) \\
& + f_x^{N-3} \dots f_x^{i-1} f_u^{i-2} (u_{i-2}' - \bar{u}_{i-2}) + \dots + f_x^{N-3} \dots f_x^1 f_u^0 (u_0' - \bar{u}_0)] \\
& \qquad \qquad \qquad \vdots \\
& (f_u^i)^T [\beta_{i+1}^T C_{i+1} \beta_{i+1}] [f_x^i f_u^{i-1} (u_{i-1}' - \bar{u}_{i-1}) + \dots + \\
& \qquad \qquad \qquad f_x^i \dots f_x^1 f_u^0 (u_0' - \bar{u}_0)] .
\end{aligned}$$

Now,

$$\tilde{F}_{u_i u_j} = (\tilde{F}_{u_j u_i})^T,$$

Then, from (v) and the definition of Δ_i , we obtain :

$$\begin{aligned} \text{(xi)} \quad B_i \Delta_i &= \sum_{j=0}^{i-1} \tilde{F}_{u_i u_j} (u'_j - \bar{u}_j) - (f_u^i)^T (f_x^{i+1})^T \dots (f_x^{N-2})^T \\ &\quad [\beta_{N-1}^T \ C_{N-1} \ \beta_{N-1}] [f_x^{N-2} \dots f_x^i f_u^{i-1} (u'_{i-1} - \bar{u}_{i-1}) \\ &\quad + f_x^{N-2} \dots f_x^{i-1} f_u^{i-2} (u'_{i-2} - \bar{u}_{i-2}) + \dots + \\ &\quad f_x^{N-2} \dots f_x^1 f_u^0 (u'_0 - \bar{u}_0)] \\ &\quad - (f_u^i)^T (f_x^{i+1})^T \dots (f_x^{N-3})^T [\beta_{N-2}^T \ C_{N-2} \ \beta_{N-2}] [f_x^{N-3} \dots f_x^i \\ &\quad f_u^{i-1} (u'_{i-1} - \bar{u}_{i-1}) \\ &\quad + f_x^{N-3} \dots f_x^{i-1} f_u^{i-2} (u'_{i-2} - \bar{u}_{i-2}) + \dots + f_x^{N-3} \dots f_x^1 f_u^0 \\ &\quad (u'_0 - \bar{u}_0)] \\ &\quad - (f_u^i)^T [\beta_{i+1}^T \ C_{i+1} \ \beta_{i+1}] [f_x^i f_u^{i-1} (u'_{i-1} - \bar{u}_{i-1}) + \dots + \\ &\quad f_x^i \dots f_x^1 f_u^0 (u'_0 - \bar{u}_0)] \end{aligned}$$

Hence, from (x) and (xi), we have :

$$\tilde{F}_{u_i} + \sum_{j=0}^{N-1} \tilde{F}_{u_i u_j} (u'_j - \bar{u}_j) = H_u^i + B_i \Delta_i + C_i (u'_i - \bar{u}_i)$$

Pre-multiplying the right hand side of the above equality by C_i^{-1} we obtain :

$$\tilde{F}_{u_i} + \sum_{j=0}^{N-1} \tilde{F}_{u_i u_j} (u_j^i - \bar{u}_j) = -C_i^{-1} H_u^i - (C_i^{-1} B_i) \Delta_i + (u_i^i - \bar{u}_i) = 0.$$

This equality holds for all i . Then,

$$F(\bar{u}) + F_{uu}(\bar{u})(u' - \bar{u}) = 0$$

where,

$$u' \triangleq \{u_0', u_1', \dots, u_{N-1}'\}.$$

That is, u' satisfies the first order condition to minimize $\Phi^N(u)$.

The theorem follows from the positive definiteness of $\tilde{F}_{uu}(\bar{u})$.

Theorem II.1.B

Let $\tilde{F}_{uu}(\bar{u})$ be positive definite.

Let the dynamic equations of the unconstrained control problem be linear.

Then,

the two successors u^N, u^D coincide.

Proof :

As the dynamic equations are linear by hypothesis

$$f_{ux}^i = f_{xx}^i = f_{uu}^i = 0, \text{ for all } i.$$

Hence the matrices A_i , B_i and C_i are identically computed by both versions of the DDP method. Also, by the linearity of the dynamic equations

$$\delta x_i = \Delta_i, \text{ for all } i.$$

Therefore $u^D = u'$ and the result follows from Theorem II.1.A.

Let D be the set of all schedules $\{u_k\}$ satisfying the first order necessary condition of optimality below :

$$H_u^i = 0, \quad i = 0, \dots, N-1.$$

Theorem II.1.C

Let C_i be invertible for all i

Let $\tilde{F}_{uu}(u^*)$ be positive definite for all $u^* \triangleq \{u_0^*, u^*, \dots, u_{N-1}^*\} \in D$.

Then,

the convergence $u^D \rightarrow u^*$ has a quadratic rate.

Proof:

By hypothesis, $u^D \rightarrow u^* \in D$ and $F_{uu}(u^*)$ is positive definite for all $u^* \in D$.

Hence, for all $\bar{u}$ near enough to u^* we have :

$\tilde{F}_{uu}(\bar{u})$ is positive definite.

As we are interested in asymptotic properties, we can, without loss of generality, apply Theorem I.1.A and say that :

$$u' = \bar{u}^N.$$

Now, let $C_i^!$, $B_i^!$ denote the matrices obtained by the modified version of the DDP method and C_i , B_i the matrices obtained by the original version.

From II.1.IV and II.1.VII, we have :

$$\begin{aligned} (i) \quad C_i^! - C_i &= (f_{uu}^i)^T [(f_x^{i+1})^T (f_x^{i+2})^T \dots (f_x^{N-2})^T \beta_{N-1}^T H_u^{N-1} + (f_x^{i+1})^T \\ &\quad \dots (f_x^{N-3})^T \beta_{N-2}^T H_u^{N-2} \\ &\quad + \dots + (f_x^{i+1})^T \beta_{i+2}^T H_u^{i+2} + \beta_{i+1}^T H_u^{i+1}] \end{aligned}$$

and

$$(ii) \quad B'_i - B_i = (f_{ux}^i)^T [(f_x^{i+1})^T (f_x^{i+2})^T \dots (f_x^{N-2})^T \beta_{N-1}^T H_u^{N-1} + \\ (f_x^{i+1})^T \beta_{N-2}^T H_u^{N-2} + \dots + (f_x^{i+1})^T \beta_{i+2}^T H_u^{i+2} + \\ \beta_{i+1}^T H_u^{i+1}].$$

Now,

$$C_0(u_0^D - \bar{u}_0) = -H_u^0$$

and

$$C'_0(u'_0 - \bar{u}) = -H_u^0.$$

Then, from (i), we have :

$$(iii) \quad C_0(u_0^D - u'_0) = - (f_{uu}^0)^T [(f_x^1)^T \dots (f_x^{N-2})^T \beta_{N-1}^T H_u^{N-1} + \dots + \\ \beta_1^T H_u^1] (u'_0 - \bar{u}_0)$$

But, $H_u^i(u^*) = 0$ and $\bar{u} \rightarrow u^*$.

Then, by the differentiability assumptions we have :

(iv) there exists a $\bar{c} < \infty$ s.t.

$$\| H_u^i(\bar{u}) \| \leq \bar{c} \| \bar{u} - u^* \|, \text{ for all } i$$

From the existence of C_i^{-1} for all i , (iv) and (iii), we have :

(v) there exists a $\tilde{c} < \infty$

such that:

$$\| u_0^D - u'_0 \| \leq \tilde{c} \| \bar{u} - u^* \| \| u' - \bar{u} \|.$$

But, from the quadratic convergence of Newton's method and Theorem II.1.A, we have :

(vi) there exists a $c' < \infty$

such that :

$$\|u' - \bar{u}\| \leq c' \|\bar{u} - u^*\|.$$

From (v) and (vi) we obtain :

(vii) there exists a $c^0 < \infty$

such that :

$$\|u_0^D - u_0'\| \leq c^0 \|\bar{u} - u^*\|^2.$$

Now,

$$(viii) \quad C_1(u_1^D - \bar{u}_1) = -H_u^1 - B_1 \delta x_1.$$

From (ii) and (iv) we can say that :

$$(ix) \quad B_1 \delta x_1 = B_1' \delta x_1 - \bar{e}_1 \delta x_1$$

where,

$$\|\bar{e}_1\| \approx \|\bar{u} - u^*\|.$$

By the mean value theorem, $\exists u_\alpha \in [\bar{u}_0, u_0^D]$ such that :

$$(x) \quad \delta x_1 = f_u^0(\bar{u}_0)(u_0^D - \bar{u}_0) + \frac{1}{2} (u_0^D - \bar{u}_0)^T f_{uu}^0(u_\alpha)(u_0^D - \bar{u}_0)$$

Also,

$$(xi) \quad \|u_0^D - \bar{u}_0\| \leq \|u_0^D - u_0'\| + \|u_0' - \bar{u}_0\|.$$

From (xi), (x), (ix), (viii) and (vii) we obtain :

$$(xii) \quad C_1(u_1^D - \bar{u}_1) = -H_u^1 - B_1' \delta x_1 + \bar{e}_2$$

where,

$$\|\bar{e}_2\| \approx \|\bar{u} - u^*\|^2.$$

But,

$$(xiii) \quad C_1'(u_1' - \bar{u}_1) = -H_u^1 - B_1' \Delta_1.$$

Now,

$$(xiv) \quad \Delta_1 = f_u^0(\bar{u}_0)(u_0' - \bar{u}_0).$$

Then, from (xiv), (x), (xi), (vii) and (vi), we have :

(xv) there exists a $\hat{c} < \infty$ such that :

$$\| \Delta_1 - \delta x_1 \| \leq \hat{c} \| \bar{u} - u^* \|^2.$$

From (xv), (xiii) and (xii) we have :

$$(xvi) \quad C_1(u_1^D - \bar{u}_1) - C_1'(u_1' - \bar{u}_1) = \bar{e}_3$$

where,

$$\| \bar{e}_3 \| \approx \| \bar{u} - u^* \|^2.$$

But, from (i),

$$(xvii) \quad C_1'(u_1' - \bar{u}_1) = C_1(u_1' - \bar{u}_1) + (f_{uu}^i)^T [f_x^{i+1}]^T \dots (f_x^{N-2})^T \beta_{N-1}^T H_u^{N-1} \\ + \dots + \beta_2^T H_u^2 (u_1' - \bar{u}_1).$$

Then, from (iv), (vi), (xvi) and (xvii), we obtain :

$$C_1(u_1^D - u_1') = \bar{e}_4$$

where

$$\| \bar{e}_4 \| \approx \| \bar{u} - u^* \|^2.$$

Hence, by the existence of C_1^{-1} , we have :

(xviii) there exists a $c^1 < \infty$

such that

$$\| u_1^D - u_1' \| \leq c^1 \| \bar{u} - u^* \|^2.$$

Repeating the reasoning which led to (vii) and (xviii) for $i = 2, \dots, N-1$, similar results can be obtained. That is

there exists $c^i < \infty$, $i = 0, \dots, N-1$,

such that:

$$\|u_i^D - u_i'\| \leq c^i \|\bar{u} - u^*\|^2$$

Hence,

(xix) there exists a $M_1 < \infty$

such that:

$$\|u^D - u'\| \leq M_1 \|\bar{u} - u^*\|^2.$$

But,

(xx) $\|u^D - u^*\| \leq \|u^D - u'\| + \|u' - u^*\|$.

By the quadratic convergence of Newton's method and Theorem I.1.1.A, we can see that :

(xxi) there exists a $M_2 < \infty$

such that:

$$\|u' - u^*\| \leq M_2 \|\bar{u} - u^*\|^2.$$

From (xix), (xx) and (xxi) we obtain

there exists a $M < \infty$

such that :

$$\|u^D - u^*\| \leq M \|\bar{u} - u^*\|^2$$

which is the desired result.

II.1.4 Remarks

First note that a direct application of Newton's method to solve the unconstrained discrete control problem has in general no

practical interest because of the high dimension of the resultant Hessian matrix. In other words, the DDP method has the advantage of performing N inversions of $(m \times m)$ matrices instead of the cumbersome task of inverting a $(mN \times mN)$ matrix. This reduction in the dimension of the matrices is also relevant to the solution of some systems of equations which can be transformed into an equivalent discrete control problem.

For instance, solving the system below,

II.1.XII

$$z^2 + 2y^2 = 1$$

$$-3z^2 + y^3 = -1$$

is equivalent to solving the following control problem :

II.1.XIII

$$\min_{(u_0, u_1)} \{ \|x_2\|^2 \} = F(x_2)$$

with the dynamics given by :

$$\left. \begin{aligned} 1x_1 &= 1x_0 + u_0^2 \\ 2x_1 &= 2x_0 - 3u_0 \end{aligned} \right\} x_1 = f_0(x_0, u_0)$$

$$\left. \begin{aligned} 1x_2 &= 1x_1 + 2u_1^2 - 1 \\ 2x_2 &= 2x_1 + u_1^3 + 1 \end{aligned} \right\} x_2 = f_1(x_1, u_1)$$

with,

$$x_i \in R^2, i = 0, 1, 2 \dots N; x_0 = (0, 0)$$

jx_i denoting the j^{th} component of x_i

$$u_i \in R, i = 0, 1.$$

Clearly, if $x_2^* = (0,0)$ is a reachable point through some control sequence $\{u_0^*, u_1^*\}$ this sequence is the solution of II.1.XIII and $z = u_0^*$, $y = u_1^*$ will be the solution of I.1.XII. If at the solution of II.1.XIII $x_2^* \neq (0,0)$ then the system II.1.XII has no solution. Now, if we apply Newton's method to solve II.1.XII a (2×2) matrix has to be inverted at each iteration. The application of the modified version of the DDP method, described in the section II.1.2, to solve II.1.XIII would generate the same sequence of points as Newton's method but only requiring, at each iteration, the trivial inversion of two real numbers.

So, besides providing an easy way of proving the results of Theorem II.1.B and Theorem II.1.C to the original version of the DDP method, the modifications suggested here have the interest of being a 'step by step' implementation of the well known and highly successful Newton's method. Indeed, although we were mainly concerned in this chapter with the rate of convergence of the DDP method, it is clear that Theorem II.1.A makes possible the direct application of a great number of results in the literature related to Newton's method [e.g., the Global Damped Newton Theorem of Ortega and Rheinboldt (1970)]. In particular, the one step convergence of the DDP method for LQP problems (linear systems, quadratic performance index), which was already shown [see Jacobson and Mayne (1970)] by noting that in these cases the method reduces to the conventional dynamic programming can now also be seen, through Theorem II.1.B, as a consequence of the well known one step convergence of Newton's method when applied to quadratic functions. This last approach seems more natural because both the DDP and Newton's methods converge in one iteration only in the quadratic case, while the conventional dynamic programming is an one-step convergent method.

As an illustration of the results shown in this chapter, we will apply the different methods to the solution of the unconstrained discrete control problem given by II.1.XIII. For both the versions of the DDP method and for Newton's method we start by detailing their first iterations. Afterwards another iteration of Newton's method and the corresponding one of the modified DDP method are again detailed. Hopefully this will help to clarify the proof of Theorem II.1.A. Finally the complete solution of the problem is given. As we shall see, the original version of the DDP method does differ from Newton's method, but ultimately both methods perform equally well, something to be expected from Theorem II.1.C.

The nominal sequences from which all the methods started is :

$$\{\bar{u}_i\} = \{1;1\} \Rightarrow \{\bar{x}_i\} = \left[\begin{pmatrix} 0 \\ 0 \end{pmatrix} ; \begin{pmatrix} 1 \\ -3 \end{pmatrix} ; \begin{pmatrix} 2 \\ -1 \end{pmatrix} \right] .$$

A similar step length procedure to the one suggested by Jacobson and Mayne (1970) is applied to the direction originated by Newton's method but with the estimate of the change in cost a_0 being computed as :

$$a_0 = -[\tilde{F}_{u_0}(u_0^N - \bar{u}_0) + \tilde{F}_{u_1}(u_1^N - \bar{u}_1)]$$

where (u_0^N, u_1^N) is obtained by Newton's method. Because of Theorem II.1.A, Theorem II.1.C and the quadratic convergence of Newton's method, we expected a step length of unity to be ultimately achieved by all methods. As it happened, the step length was set, by all methods, equal to unity from start to finish.

A 10-digit calculator was used and differences smaller than 9×10^{-8} are negligible.

1st ITERATIONDDP method (original)

$$\text{Date : } \{\bar{u}_1\} = \{1 ; 1\} \Rightarrow \{\bar{x}_1\} = \left\{ \begin{pmatrix} 0 \\ 0 \end{pmatrix} ; \begin{pmatrix} 1 \\ -3 \end{pmatrix} ; \begin{pmatrix} 2 \\ -1 \end{pmatrix} \right\}$$

backward:

$$A_1 = H_{xx}^1 + (f_x^1)^T F_{xx}(\bar{x}_2) f_x^1 = F_{xx}(\bar{x}_2) = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}$$

$$B_1 = H_{ux}^1 + (f_u^1)^T F_{xx}(\bar{x}_2) f_x^1 = [4\bar{u}_1 \quad 3\bar{u}_1^2] \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix} = [8 \quad 6].$$

$$C_1 = H_{uu}^1 + (f_u^1)^T F_{xx}(\bar{x}_2) f_u^1 = [4 \quad 6\bar{u}_1] \begin{pmatrix} 2_1 \bar{x}_2 \\ 2_2 \bar{x}_2 \end{pmatrix} \\ + [4\bar{u}_1 \quad 3\bar{u}_1^2] \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix} \begin{pmatrix} 4\bar{u}_1 \\ 3\bar{u}_1^2 \end{pmatrix} \\ = 54.$$

$$B_1 = -C_1^{-1} B_1 = [-8/54 \quad -6/54]$$

$$V_x^1 = H_x^1 + \beta_1^T H_u^1 = \begin{pmatrix} 1.518518519 \\ -3.111111111 \end{pmatrix}$$

$$V_{xx}^1 = A_1 - \beta_1^T C_1 \beta_1 = \begin{pmatrix} .814814815 & -.888888888 \\ -.888888888 & 1.333333334 \end{pmatrix}$$

$$\alpha_1 = -C_1^{-1} H_u^1 = -.185185185$$

$$a_1 = 0 - \frac{1}{2}(H_u^1)^T C_1^{-1} H_u^1 = -50/54,$$

$$A_0 = H_{xx}^0 + (f_x^0)^T V_{xx}^1 f_x^0 = f_{xx}^0 V_x^1 + (f_x^0)^T V_{xx}^1 f_x^0 = V_{xx}^1.$$

$$C_0 = f_{uu}^0 V_x^1 + (f_u^0)^T V_{xx}^1 f_u^0 = [2 \quad 0] V_x^1 + [2\bar{u}_0 \quad -3] V_{xx}^1 \begin{pmatrix} 2\bar{u}_0 \\ -3 \end{pmatrix}$$

$$= 30.96296295$$

$$\alpha_0 = -C_0^{-1} H_u^0 = -.464114832.$$

$$a_0 = a_1 = \frac{1}{2}(H_u^0)^T C_0^{-1} H_u^0 = -4.26027942$$

onward:

$$\varepsilon = 1$$

$$u_0^D = \bar{u}_0 + \alpha_0 = .535885168$$

$$1x_1 = 1x_0 + u_0^2 = .287172913$$

$$2x_1 = 2x_0 - 3u_0 = -1.60765504$$

$$\delta x_1 = [x_1 \quad -\bar{x}_1] = \begin{pmatrix} -.712827087 \\ 1.392344496 \end{pmatrix}$$

$$u_1^D = \bar{u}_1 + \alpha_1 + \beta_1 \Delta x_1 = .765713883$$

$$1x_2 = 1x_1 + 2u_1^2 = .459808414$$

$$2x_2 = 2x_1 + u_1^3 = -.158703864$$

$$F(x_2) = 1x_2 + 2x_2^2 = .236610693$$

step length

procedure:

$$\Delta V_0 = F(x_2) - F(\bar{x}_2) = -4.763389307 < 0$$

$$\frac{|\Delta V_0|}{a_0^{1/2}} = 2.23 \dots > \frac{1}{2},$$

Then,

$\varepsilon = 1$ is accepted.

1st ITERATION

Newton's method

$$\text{Data : } \{\bar{u}_1\} = \{1 ; 1\} \Rightarrow \{\bar{x}_1\} = \left\{ \begin{pmatrix} 0 \\ 0 \end{pmatrix} ; \begin{pmatrix} 1 \\ -3 \end{pmatrix} ; \begin{pmatrix} 2 \\ -1 \end{pmatrix} \right\}.$$

$$\tilde{F}_{u_1} = f_u^1(\bar{x}_1, \bar{u}_1)^T F_x(\bar{x}_2) = [4\bar{u}_1 \quad 3\bar{u}_1^2] \begin{pmatrix} 2 & 1\bar{x}_2 \\ 2 & 2\bar{x}_2 \end{pmatrix} = 10$$

$$\tilde{F}_{u_0} = f_u^0(\bar{x}_0, \bar{u}_0)^T f_x^1(\bar{x}_1, \bar{u}_1)^T F_x(\bar{x}_2) = [2\bar{u}_0 \quad -3] \begin{pmatrix} 2 & 1\bar{x}_2 \\ 2 & 2\bar{x}_2 \end{pmatrix} = 14$$

$$\begin{aligned} \tilde{F}_{u_0 u_1} &= F_{u_1 u_0}^T = f_u^0(\bar{x}_0, \bar{u}_0)^T [f_{ux}^1(\bar{x}_1, \bar{u}_1)^T F_x(\bar{x}_2) + f_u^1(\bar{x}_1, \bar{u}_1)^T \\ &\quad F_{xx}(\bar{x}_2) f_x^1(\bar{x}_1, \bar{u}_1)]^T \\ &= [2\bar{u}_0 \quad -3] \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix} \begin{pmatrix} 4\bar{u}_1 \\ 3\bar{u}_1^2 \end{pmatrix} = -2. \end{aligned}$$

$$\begin{aligned} \tilde{F}_{u_0 u_0} &= f_{uu}^0(\bar{x}_0, \bar{u}_0)^T f_x^1(\bar{x}_1, \bar{u}_1)^T F_x(\bar{x}_2) + f_u^0(\bar{x}_0, \bar{u}_0)^T [f_{xx}^1(\bar{x}_1, \bar{u}_1)^T \\ &\quad F_x(\bar{x}_2) + f_x^1(\bar{x}_1, \bar{u}_1)^T F_{xx}(\bar{x}_2) f_x^1(\bar{x}_1, \bar{u}_1)] f_u^0(\bar{x}_0, \bar{u}_0) \\ &= [2 \quad 0] \begin{pmatrix} 2 & 1\bar{x}_2 \\ 2 & 2\bar{x}_2 \end{pmatrix} + [2\bar{u}_0 \quad -3] \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix} \begin{pmatrix} 2\bar{u}_0 \\ -3 \end{pmatrix} = 34 \end{aligned}$$

$$\begin{aligned}\tilde{F}_{u_1 u_1} &= f_{uu}^1(\bar{x}_1, \bar{u}_1)^T F_x(\bar{x}_2) + f_u^1(\bar{x}_1, \bar{u}_1)^T F_{xx}(\bar{x}_2) f_u^1(\bar{x}_1, \bar{u}_1) \\ &= [4 \quad 6\bar{u}_1] \begin{bmatrix} 2 & \bar{x}_2 \\ 2 & 2\bar{x}_2 \end{bmatrix} + [4\bar{u}_1 \quad 3\bar{u}_1^2] \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} 4\bar{u}_1 \\ 3\bar{u}_1^2 \end{bmatrix}\end{aligned}$$

That is :

$$\tilde{F}_u = \begin{bmatrix} 1 & 4 \\ 1 & 0 \end{bmatrix} ; \quad \tilde{F}_{uu} = \begin{bmatrix} 34 & -2 \\ -2 & 54 \end{bmatrix}$$

$$\text{Now: } (\tilde{F}_{uu})^{-1} = \begin{bmatrix} .029475982 & 1.091703 \times 10^{-3} \\ 1.091703 \times 10^{-3} & .018558951 \end{bmatrix}$$

and

$$u^N = \bar{u} - (\tilde{F}_{uu})^{-1} \tilde{F}_u = \begin{bmatrix} .576419222 \\ .799126648 \end{bmatrix}$$

New trajectory :

$${}_1x_1 = {}_1x_0 + u_0^2 = .332259119$$

$${}_2x_1 = {}_2x_0 - 3u_0 = -1.729257665$$

$${}_1x_2 = {}_1x_1 + 2u_1^2 - 1 = .609465919$$

$${}_2x_2 = {}_2x_1 + u_1^3 + 1 = .218932675$$

$$F(x_2) = {}_1x_2^2 + {}_2x_2 = .419380226$$

step length

$$\text{procedure : } \Delta V_0 = F(x_2) - F(\bar{x}_2) = -4.580619774 < 0$$

$$a_0 = -F_{u_0}(u_0^N - \bar{u}_0) - F_{u_1}(u_1^N - \bar{u}_1) = -7.938864412$$

$$\frac{|\Delta V_0|}{|a_0/2|} = 1.153 \dots > \frac{1}{2}$$

$\epsilon = 1$ is accepted.

1st ITERATION

DDP method (modified)

$$\text{Data : } \{\bar{u}_i\} = \{1 ; 1\} \Rightarrow \{\bar{x}_i\} = \left(\begin{pmatrix} 0 \\ 0 \end{pmatrix} ; \begin{pmatrix} 1 \\ -3 \end{pmatrix} ; \begin{pmatrix} 2 \\ -1 \end{pmatrix} \right)$$

backward:

$$A_1(\text{unchanged}) = F_{xx}(\bar{x}_2) = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}$$

$$B_1(\text{unchanged}) = [4_{u_1} \quad 3\bar{u}_1^2] \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix} = [8 \quad 6]$$

$$C_1(\text{unchanged}) = [4 \quad 6\bar{u}_2] \begin{pmatrix} 2 & 1\bar{x}_2 \\ 2 & 2\bar{x}_2 \end{pmatrix} + [4\bar{u}_1 \quad 3\bar{u}_1^2] \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix} \begin{pmatrix} 4\bar{u}_1 \\ 3\bar{u}_1^2 \end{pmatrix} = 54$$

$$\beta_1(\text{unchanged}) = -C_1^{-1} B_1 = [-8/54 \quad -6/54]$$

$$V_x^1(\text{unchanged}) = H_x^1 + \beta_1^T H_u^1 = \begin{pmatrix} 2.518518519 \\ -3.111111111 \end{pmatrix}$$

$$\bar{V}_x^1(\text{new}) = H_x^1 = \begin{pmatrix} 2 & 1\bar{x}_2 \\ 2 & 2\bar{x}_2 \end{pmatrix} = \begin{pmatrix} 4 \\ -2 \end{pmatrix}$$

$$V_{xx}^1 \text{ (unchanged)} = A_1 - \beta_1^T C_1 \beta_1 = \begin{bmatrix} .814814815 & -.888888888 \\ -.888888888 & 1.333333334 \end{bmatrix}$$

$$\alpha_1 \text{ (unchanged)} = - C_1^{-1} H_u^1 = -.185185185$$

$$A_0 \text{ (modified)} = f_{xx}^0 \bar{V}_x^1 + (f_x^0)^T V_{xx}^1 f_x^0 = V_{xx}^1$$

$$C_0 \text{ (modified)} = \begin{bmatrix} 2 & 0 \end{bmatrix} \bar{V}_x^1 + \begin{bmatrix} 2\bar{u}_0 & -3 \end{bmatrix} V_{xx}^1 \begin{bmatrix} 2\bar{u}_0 \\ -3 \end{bmatrix}$$

$$= 33.92592592$$

$$\alpha_0 = - C_0^{-1} H_u^0 = -.423580786$$

onward :

$$u'_0 = \bar{u}_0 + \alpha_0 = .576419214$$

$$1x_1 = 1x_0 + u_0^2 = .332259111$$

$$2x_1 = 2x_0 - 3\bar{u}_0 = - 1.729257642$$

$$\delta x_1 \text{ (Just to compare)} = \begin{bmatrix} -.667740889 \\ 1.270742358 \end{bmatrix}$$

$$\Delta_1 \text{ (new)} = f_u^0 (u'_0 - \bar{u}_0) = \alpha_0 \begin{bmatrix} 2\bar{u}_0 \\ -3 \end{bmatrix} = \begin{bmatrix} -.847161572 \\ 1.270742358 \end{bmatrix}$$

$$u'_1 \text{ (modified)} = \bar{u}_1 + \alpha_1 + \beta_1 \Delta_1 = .799126638.$$

Observation:

Note that :

$$|u_0^N - u'_0| = 8 \times 10^{-9} < 9 \times 10^{-8}$$

and

$$|u_1^N - N'_1| = 1 \times 10^{-8} < 9 \times 10^{-8},$$

These differences are negligible and are due to lack of precision.

In order to avoid the accumulation of these errors disturbing our comparisons, we will use in the next iteration the same nominal sequence both to Newton's method and the modified version of the DDP method.

2nd ITERATION

Newton's method

Current sequences:

$$\{\bar{u}_1\} = \begin{pmatrix} .576419222 \\ .799126648 \end{pmatrix} \Rightarrow \{\bar{x}_1\} = \left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}; \begin{pmatrix} .332259119 \\ -1.729257665 \end{pmatrix} \right);$$

$$\begin{pmatrix} .609465919 \\ -.218932675 \end{pmatrix} \Bigg) .$$

$$F_{u_1} = [4\bar{u}_1 \quad 3\bar{u}_1^2] \begin{pmatrix} 2 & 1\bar{x}_2 \\ 2 & 2\bar{x}_2 \end{pmatrix} = 3.057456752$$

$$F_{u_0} = [2\bar{u}_0 \quad -3] \begin{pmatrix} 2 & 1\bar{x}_2 \\ 2 & 2\bar{x}_2 \end{pmatrix} = 2.718827533$$

$$F_{u_0 u_1} = [2\bar{u}_0 \quad -3] \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix} \begin{pmatrix} 4\bar{u}_1 \\ 3\bar{u}_1^2 \end{pmatrix} = -4.12474983$$

$$\tilde{F}_{u_0 u_0} = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix} \begin{pmatrix} 2 \\ 1 \end{pmatrix} \bar{x}_2 + \begin{bmatrix} 2\bar{u}_0 & -3 \\ 0 & 2 \end{bmatrix} \begin{pmatrix} 2 \\ 0 \end{pmatrix} \begin{pmatrix} 2\bar{u}_0 \\ -3 \end{pmatrix}$$

$$= 23.09593663$$

$$\tilde{F}_{u_1 u_1} = \begin{bmatrix} 4 & 6\bar{u}_1 \\ 6\bar{u}_1 & 3\bar{u}_1^2 \end{bmatrix} \begin{pmatrix} 2 \\ 1 \end{pmatrix} \bar{x}_2 + \begin{bmatrix} 4\bar{u}_1 & 3\bar{u}_1^2 \\ 0 & 2 \end{bmatrix} \begin{pmatrix} 2 \\ 0 \end{pmatrix} \begin{pmatrix} 4\bar{u}_1 \\ 3\bar{u}_1^2 \end{pmatrix}$$

$$= 30.55223434.$$

That is :

$$\tilde{F}_u = \begin{pmatrix} 2.718827533 \\ 3.057456752 \end{pmatrix} ; \quad \tilde{F}_{uu} = \begin{pmatrix} 23.09593663 & -4.12474983 \\ -4.12474983 & 30.555223434 \end{pmatrix}$$

Now:

$$(\tilde{F}_{uu})^{-1} = \begin{pmatrix} .044367405 & 5.9898875 \times 10^{-3} \\ 5.9898875 \times 10^{-3} & .033539504 \end{pmatrix}$$

and :

$$u^N = \bar{u} - (\tilde{F}_{uu})^{-1} \tilde{F}_u = \begin{pmatrix} .437478077 \\ .680295596 \end{pmatrix}$$

2nd ITERATION

DDP method (modified)

Current sequences:

$$\{\bar{u}_1\} = \begin{pmatrix} .576419222 \\ .799126648 \end{pmatrix} \Rightarrow \{\bar{x}_1\} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} ; \begin{pmatrix} .332259119 \\ -1.729357665 \end{pmatrix} ; \begin{pmatrix} .609465919 \\ -.218932675 \end{pmatrix}$$

backward :

$$A_1 = F_{xx}(\bar{x}_2) = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}$$

$$B_1 = [4\bar{u}_1 \quad 3\bar{u}_1^2] \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix} = [6.393013184 \quad 3.831620397]$$

$$C_1 = [4 \quad 6\bar{u}_1] \begin{pmatrix} 2 & \bar{x}_2 \\ 2 & \bar{x}_2 \end{pmatrix} + [4\bar{u}_1 \quad 3\bar{u}_1^2] \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix} \begin{pmatrix} 4\bar{u}_1 \\ 3\bar{u}_1^2 \end{pmatrix}$$

$$= 30.55223434$$

$$\beta_1 = -C_1^{-1} B_1 = [-.20924863 \quad -.125412117]$$

$$V_x^1 = H_x^1 + \beta_1^T H_u^1 = \begin{pmatrix} .579163202 \\ -.821307473 \end{pmatrix}$$

$$\bar{V}_x^1 = H_x^1 = \begin{pmatrix} 2 & \bar{x}_2 \\ 2 & \bar{x}_2 \end{pmatrix} = \begin{pmatrix} 1.218931838 \\ -.43786535 \end{pmatrix}$$

$$V_{xx}^1 = A_1 - \beta_1^T C_1 \beta_1 = \begin{pmatrix} .662270756 & -.801761316 \\ -.801761316 & 1.519468379 \end{pmatrix}$$

$$\alpha_1 = -C_1^{-1} H_u^1 = -.100073098.$$

$$A_0 = V_{xx}^1$$

$$C_0 = f_{xx}^0 \bar{V}_x^1 + [2\bar{u}_0 \quad -3] V_{xx}^1 \begin{pmatrix} 2\bar{u}_0 \\ -3 \end{pmatrix} = 22.53906866$$

$$\alpha_0 = -C_0^{-1} H_u^0 = -.138941145$$

onward:

$$u'_0 = \bar{u}_0 + \alpha_0 = .437478077.$$

$$\Delta_1 = f_u^0(u'_0 - \bar{u}_0) = \alpha_0 \begin{pmatrix} 2\bar{u}_0 \\ -3 \end{pmatrix} = \begin{pmatrix} -.160176693 \\ .416823435 \end{pmatrix}$$

$$u'_1 = \bar{u}_1 + \alpha_1 + \beta_1 \Delta_1 = .680295594.$$

Comparing with Newton's we have :

$$|u_0^N - u'_0| < 1. \times 10^{-9}$$

and

$$|u_0^N - u'_0| < 3. \times 10^{-9}.$$

Again the difference is negligible and are due to lack of precision.

Finally, in the next page , we have the complete solution of II.1.XIII by both Newton's and the original version of the DDP method.

Method →	NEWTON's	DDP
Iteration ↓		(original)
1	$\bar{u} = \{1; 1\}$ $u^N = \{.576419222; .799126648\}$ $\ x_2\ ^2 = .419380226$	$\bar{u} = \{1; 1\}$ $u^D = \{.535885168; .765713883\}$ $\ x_2\ ^2 = .236610693$
2	$\bar{u} = \{.576419222; .799126648\}$ $u^N = \{.437478077; .680295596\}$ $\ x_2\ ^2 = .013840913$	$\bar{u} = \{.535885168; .765713883\}$ $u^D = \{.425362292; .663617459\}$ $\ x_2\ ^2 = .0040692655$
3	$\bar{u} = \{.437478077; .680295596\}$ $u^N = \{.424464777; .643873171\}$ $\ x_2\ ^2 = 1.2854101 \times 10^{-4}$	$\bar{u} = \{.425362292; .663617459\}$ $u^D = \{.421254214; .642502274\}$ $\ x_2\ ^2 = 1.160171099 \times 10^{-5}$
4	$\bar{u} = \{.424464777; .643873171\}$ $u^N = \{.421257407; .641320785\}$ $\ x_2\ ^2 = 1.8099364 \times 10^{-8}$	$\bar{u} = \{.421254214; .642502274\}$ $u^D = \{.421250877; .641310327\}$ $\ x_2\ ^2 = 1.26540677 \times 10^{-10}$
5	$\bar{u} = \{.421257407; .641320785\}$ $u^N = \{.42125084; .641306374\}$ $\ x_2\ ^2 = 8.84 \times 10^{-16}$	$\bar{u} = \{.421250877; .641310327\}$ $u^D = \{.421250849; .641306369\}$ $\ x_2\ ^2 = 1.8 \times 10^{-17}$

II.2 CONSTRAINED CONTROL CASE

II.2.1 Introduction

In this chapter we will be concerned with the solution of the discrete control problem (DCP) below:

$$\begin{array}{l} \text{DCP} \\ \min_{u \in R^{m \times N}} \left\{ \sum_{i=0}^{N-1} T_i(x_i, u_i) + \phi(x_N) \right\} \end{array}$$

with the dynamic system:

$$x_{i+1} = f_i(x_i, u_i), \quad i = 0, \dots, N-1$$

$$x_0 = x^0, \text{ given}$$

and satisfying:

$$g_i^s(u_i) \leq 0, \quad i = 0, \dots, N-1$$

$$s = 1, \dots, \ell_i.$$

For the same reasons given in the introduction to the previous chapter we can, without loss of generality, consider the above problem in the form

$$\begin{array}{l} \text{DCP} \\ \min_{u \in R^{m \times N}} F(x_N) \end{array}$$

with the dynamic system:

$$x_{i+1} = f_i(x_i, u_i), \quad i = 0, \dots, N-1$$

$$x_0 = x^0 \text{ (given)}$$

and satisfying:

$$g_i(u_i) \leq 0, \quad i = 0, \dots, N-1$$

where for all i we have:

$$f_i: R^n \times R^m \rightarrow R^n; \quad g_i: R^m \rightarrow R^{\ell_i}$$

and

$$x_i \in R^n \text{ being the state vector, } \quad i = 0, \dots, N$$

$$u_i \in R^m \text{ being the control vector, } i = 0, \dots, N-1$$

$$F: R^n \rightarrow R.$$

Clearly, with the definition of the vector

$$z_i = (x_i, u_i) \in R^{(n+m)N}$$

this problem can be solved as a GMP in $R^{(m+n)N}$. However, the dependence of the dimension on the number of stages makes this approach unattractive, although the similarities between the DCP and the GMP are exploited during the analysis of the algorithm which is proposed in the next section for solving DCP. The algorithm is an extension of the DDP method, given in the second section of the previous chapter, for problems with control constraints.

The definition of α_i by II.1.V suggests that in the constrained case we should have:

$$\alpha_i = \arg \min_{\alpha \in R^m} \left\{ \begin{aligned} & \left[H_u^i \right]^T \alpha + \frac{1}{2} \alpha^T C_i \alpha \Big| g_i^s(\bar{u}_i + \alpha) \leq 0 \\ & s = 1, \dots, \ell_i \end{aligned} \right\}$$

where $(\{\bar{x}_i\}, \{\bar{u}_i\})$ are given nominal sequences.

But, as in Newton's method for constrained optimization [see Levitin and Polyak(1966)] the non-linearity of the constraints $g(\cdot)$ may make this sub-problem rather difficult. Hence, as already

done in Part I of this thesis, a linearization of the constraints is performed and the objective function is modified to take into account the curvature of $g(\cdot)$. That is, the algorithm computes, backwards, (α_i, λ_i) , the Kuhn-Tucker couple of the following quadratic sub-problem:

$$\begin{aligned} &\underline{\text{QP}_i} \\ &\min_{\alpha \in \mathbb{R}^m} [H_u^i]^T \alpha + \frac{1}{2} \alpha^T D_i \alpha \end{aligned}$$

subject to:

$$g_i^s(\bar{u}_i) + {}_i g_u^s(\bar{u}_i) \alpha \leq 0, \quad s = 1, \dots, \ell_i$$

where

$$D_i = C_i + \sum_{s=1}^{\ell_i} \bar{\lambda}_i^s(\bar{u}_i) {}_i g_{uu}^s(\bar{u}_i)$$

with $\bar{\lambda}_i^s(\bar{u}_i)$ being an estimate to λ_i^s .

In the unconstrained case the new sequence is obtained forwards by adding the term $\beta_i \Delta_i$ to the α_i computed backwards. That is:

II.2.1

$$\delta u_i = \alpha_i + \beta_i \Delta_i$$

with Δ_i given by II.1.VIII.

In the constrained case the addition of a similar term may lead to the new sequence not satisfying the linearized constraints.

On the other hand, with the α_i , β_i being defined by

$$\alpha_i = -C_i^{-1} H_u^i \quad \text{and} \quad \beta_i = -C_i^{-1} B_i,$$

II.2.1 is clearly identical to:

$$\delta u_i = \arg \min_{\delta u \in \mathbb{R}^m} \left\{ [H_u^i + B_i \Delta_i]^T \delta u + \frac{1}{2} \delta u^T C_i \delta u \right\}$$

Hence, in the presence of constraints the new sequence $\{u_i\}$ can be constructed as follows:

II.2.II

$$u_i = \bar{u}_i + \delta u_i$$

where, for all i , δu_i solves the quadratic sub-problem below:

DQP_i

$$\min_{\delta u \in \mathbb{R}^m} [H_u^i + B_i \Delta_i]^T \delta u + \frac{1}{2} \delta u^T D_i \delta u$$

subject to:

$$g_i^s(\bar{u}_i) + g_{u_i}^s(\bar{u}_i) \delta u \leq 0, \quad s = 1, \dots, \ell_i$$

Note that, considering Δ_i as a perturbation, the sub-problem above may be thought of as a perturbation of QP_i . Calling $\bar{z}: \mathbb{R}^n \rightarrow \mathbb{R}^m$ the map which for each Δ_i associates the solution of the perturbed sub-problem DQP_i , we can use results from Fiacco and McCormick(1968) on sensitivity to show that:

$$\frac{d\bar{z}(0)}{d\Delta_i} = -D_i^{-1} B_i$$

The expression above is the appropriate extension to the definition of β_i in the unconstrained case and

II.2.III

$$\delta u_i' = \alpha_i + \beta_i \Delta_i$$

with:

$$\alpha_i \text{ solving } QP_i \text{ and } \beta_i = -D_i^{-1} B_i$$

is a first order approximation to the solution of DQP_i .

In order to understand better the modifications that the recursive equations II.1.II, II.1.III (and consequently the matrices A_i , B_i , C_i) undergo because of the constraints $g(\cdot)$, recall that $V_i(\bar{x}_i, \pi)$ denotes the cost due to initial condition $\bar{x}_i$ (at time i) and policy $\pi = \{h_i(\cdot), \dots, h_{N-1}(\cdot)\}$. That is:

$$V_i(\bar{x}_i, \pi) = F(x_N)$$

with

$$x_{k+1} = f_i(x_k, u_k) \quad , \quad k = i, \dots, N-1$$

and

$$u_k = h_k(x_k) \quad , \quad k = i, \dots, N-1$$

Now, let π' be the policy defined by:

$$\pi' = \{\bar{u}_i, h_{i+1}(\cdot), \dots, h_{N-1}(\cdot)\}$$

That is, π' differs from π only at time i .

After computing α_i , a first order approximation to $V_x(\bar{x}_i, \pi)$ can be obtained from $V_x^i(\bar{x}_i, \pi')$ as:

II.2.IV

$$V_x^i(\bar{x}_i, \pi) = V_x^i(\bar{x}_i, \pi') + V_{xu}^i(\bar{x}_i, \pi') \alpha_i$$

But,

$$V_i(\bar{x}_i, \pi') = V_{i+1}(f(\bar{x}_i, \bar{u}_i), \pi)$$

So that:

II.2.V

$$V_x^i(\bar{x}_i, \pi') = f_x^i(\bar{x}_i, \bar{u}_i) V_x^{i+1}(\bar{x}_{i+1}, \pi)$$

Hence, from the definition of H_x^i in II.1.II, we have:

II.2.VI

$$V_x^i(\bar{x}_i, \pi') = H_x^i$$

Also, from II.2.V, we have:

$$V_x^i(\bar{x}_i, \pi') = [f_{ux}^i(\bar{x}_i, \bar{u}_i)]^T V_x^{i+1}(\bar{x}_{i+1}, \pi) + [f_u^i(\bar{x}_i, \bar{u}_i)]^T V_{xx}^{i+1}(\bar{x}_{i+1}, \pi) f_x^i(\bar{x}_i, \bar{u}_i)$$

Then, from II.1.IV, we have:

II.2.VII

$$V_{xu}^i(\bar{x}_i, \pi') = B_i$$

Hence, II.2.VI and II.2.VII into II.2.IV lead to:

II.2.VIII

$$V_x^i(\bar{x}_i) = H_x^i + B_i \alpha_i$$

where the argument π is omitted.

Now the minimization indicated in II.1.V implies that:

$$C_i \alpha_i + H_u^i = 0 \Rightarrow \alpha_i = C_i^{-1} H_u^i$$

Hence, II.1.II comes from the above equation, II.2.VIII and the definition: $\beta_i = -C_i^{-1} B_i$.

In the constrained case α_i is the solution of QP_i and satisfies the following equation:

II.2.IX

$$H_u^i + \sum_{s=1}^{\ell_i} \lambda_i^s g_u^i + D_i \alpha_i = 0 \Rightarrow \alpha_i = -D_i^{-1} [H_u^i + \sum_{s=1}^{\ell_i} \lambda_i^s g_u^i]$$

Hence, II.2.IX into II.2.VIII leads to the expression below which corresponds to II.1.II in the presence of constraints:

II.2.X

$$V_x^i(\bar{x}_i) = H_x^i + \beta_i^T [H_u^i + \sum_{s=1}^{\ell_i} \lambda_i^s g_u^i]$$

$$\beta_i = -D_i^{-1} B_i$$

and

$$\lambda_i \in R^{\ell_i}$$

being the multiplier associated with α_i at the solution of QP_i with the boundary condition:

$$V_x^N(\bar{x}_N) = F_x(\bar{x}_N).$$

Now, as can be seen in the proof of Theorem II.1.A, the matrix with the correction term $\beta_i^T C_i \beta_i$ in II.1.III must be the same matrix used to define α_i so that the step by step computation does not lose contact with the whole problem. Hence, the use of D_i in QP_i makes it necessary to replace II.1.III by the expression given below:

II.2.XI

$$V_{xx}^i(\bar{x}_i) = A_i - \beta_i^T D_i \beta_i$$

where

A_i is given by II.1.VII

with the boundary condition:

$$V_{xx}^N(\bar{x}_N) = F_{xx}(\bar{x}_N)$$

As in the unconstrained case, a step length procedure is incorporated in the algorithm in order to guarantee its global convergence. As the solution of the sub-problem DQP_i may not satisfy the inequalities of DCP, we use an exact penalty function to determine step length.

The definition of Δ_i involves the new controls $\{u_k\}$, $k = 0, \dots, i-1$. Hence, whenever the step length of unity is not accepted by the step length procedure we have to solve a new sequence of quadratic sub-problems DQP_i . In order to avoid this drawback, we make use of the approximation given by II.2.III. That is, if the sequence $\{\delta u_i\}$ obtained by solving $\{DQP_i\}$ is such that $\{\bar{u}_i + \delta u_i\}$ fails to make a 'substantial' reduction on the exact penalty function, then the new sequence $\{u_i\}$ is set as:

$$u_i = \bar{u}_i + \varepsilon \alpha_i + \beta_i \Delta_i$$

where $\varepsilon \in (0, 1)$ is the automatically adjusted step length. The resultant algorithm is shown to be, under some assumptions, globally convergent. Moreover, if the $\{\delta u_i\}$ obtained by solving the sub-problems DQP_i ultimately satisfies the test in the step length procedure, then the rate of convergence is quadratic.

II.2.2 The Algorithm

Let $\psi_i: R^m \rightarrow R$ be defined as:

$$\psi_i(\bar{u}_i) \triangleq \max\{0; g_i^s(\bar{u}_i), s = 1, 2, \dots, l_i\}$$

Let S_i denote the set $\{1, 2, \dots, l_i\}$

Let $\bar{\lambda}_i: \mathbb{R}^m \rightarrow \mathbb{R}^{\ell_i}$ be the map below:

$$\bar{\lambda}_i(\bar{u}_i) \triangleq \arg \min_{\lambda \in \mathbb{R}^{\ell_i}} \left\{ \left\| H_u^i + \sum_{s \in S_i} \lambda^s g_u^s(\bar{u}_i) \right\| + \sum_{s \in S_i} [(\psi_i(\bar{u}_i) - g_i^s(\bar{u}_i))^2 (\lambda^s)^2] \right\}$$

Recall that QP_i and DQP_i are defined as:

$$\begin{aligned} & \underline{QP_i} \\ & \min_{\delta u \in \mathbb{R}^m} (H_u^i)^T \delta u + \frac{1}{2} \delta u^T D_i \delta u \end{aligned}$$

subject to:

$$g_i^s(\bar{u}_i) + g_u^s(\bar{u}_i) \delta u \leq 0, \quad s \in S_i$$

and

$$\begin{aligned} & \underline{DQP_i} \\ & \min_{\delta u \in \mathbb{R}^m} [H_u^i + B_i \Delta_i]^T \delta u + \frac{1}{2} \delta u^T D_i \delta u \end{aligned}$$

subject to:

$$g_i^s(\bar{u}_i) + g_u^s(\bar{u}_i) \delta u \leq 0, \quad s \in S_i$$

Let $\tilde{\psi}: \mathbb{R}^{m \times N} \rightarrow \mathbb{R}$ be defined as:

$$\tilde{\psi}(\{\bar{u}_i\}) \triangleq \sum_{i=0}^{N-1} \psi_i(\bar{u}_i)$$

Let $\gamma: \mathbb{R}^{m \times N} \times \mathbb{R} \rightarrow \mathbb{R}$ be the penalty function below:

$$\gamma(\{\bar{u}_i\}, r) \triangleq F(\bar{x}_N) + r \tilde{\psi}(\{\bar{u}_i\})$$

where $r > 0$ is the penalty parameter.

Note that the above definitions are valid for any pair

$(\{x_i\}, \{u_i\})$ satisfying the dynamic equations of DCP.

Also, for any $\{\delta u_i\} \in R^{m \times N}$ define:

$$\hat{\psi}(\{\bar{u}_i\}, \{\delta u_i\}) \triangleq \sum_{i=0}^{N-1} \max\{0; g_i^s(\bar{u}_i) + g_u^s(\bar{u}_i) \delta u_i, s \in S_i\}$$

$$a_0 \triangleq \sum_{i=0}^{N-1} (H_u^i)^T \delta u_i$$

$$\hat{\gamma}(\{\bar{u}_i\}, \{\delta u_i\}, r) \triangleq F(\bar{x}_N) + a_0 + r \hat{\psi}(\{\bar{u}_i\}, \{\delta u_i\})$$

and

$$\theta(\{\bar{u}_i\}, \{\delta u_i\}, r) \triangleq \hat{\gamma}(\{\bar{u}_i\}, \{\delta u_i\}, r) - \gamma(\{\bar{u}_i\}, r)$$

Finally, recall that Δ_i was defined as:

$$\begin{aligned} \Delta_i = & f_x^{i-1} f_x^{i-2} \dots f_x^1 f_u^0 (u_0 - \bar{u}_0) + f_x^{i-1} f_x^{i-2} \dots f_x^2 f_u^1 (u_1 - \bar{u}_1) + \dots \\ & + f_u^{i-1} (u_{i-1} - \bar{u}_{i-1}) \dots \end{aligned}$$

ALGORITHM

Data: $\{\bar{u}_i\} \in R^{m \times N} (\Rightarrow \{\bar{x}_i\} \in R^{n \times N})$; $r > 0$; $\delta \in (0, 1)$.

Step 1: Set $i = N-1$; set $\epsilon = 1$; set $\ell = 0$.

Step 2: Compute A_i , B_i , C_i through II.1.VII, with

$$V_x^i, V_{xx}^i \text{ given by II.2.X and II.2.XI.}$$

Step 3: Compute $\bar{\lambda}_i \triangleq \bar{\lambda}_i(\bar{u}_i)$

Step 4: Set, $D_i = C_i + \sum_{s \in S_i} \bar{\lambda}_i^s g_{uu}^s(\bar{u}_i)$

Step 5: Find α_i , the solution of QP_i .

Step 6: Compute $\beta_i = -D_i^{-1} B_i$.

Step 7: If $i > 0$, set $i = i-1$; Go to Step 2.

Step 8: Compute the new control and state sequences by:

$$x_{i+1} = f_i(\bar{x}_i, u_i) \quad ; \quad x_0 = x^0$$

$$u_i = \bar{u}_i + \delta u_i, \text{ where:}$$

$$\delta u_i = \text{solves DQP}_i.$$

Step 9: Compute $\tilde{\theta} \triangleq \theta(\{\bar{u}_i\}, \{\alpha_i\}, r)$

Step 10: If $\gamma(\{u_i\}, r) - \gamma(\{\bar{u}_i\}, r) \leq \varepsilon \tilde{\theta}_i / 16$

$$\text{a) set } \bar{x}_i = x_i$$

$$\text{b) set } \ell = \ell+1, \text{ go to step 1.}$$

Step 11: Set $\varepsilon = \delta \varepsilon$; set $i = 0$.

Step 12: Compute:

$$x_{i+1} = f_i(\bar{x}_i, u_i), \quad x_0 = \bar{x}_0$$

and

$$u_i = \bar{u}_i + \varepsilon \alpha_i + \beta_i \Delta_i$$

Step 13: If $i < N-1$, set $i = i+1$; go to step 12.

Else,

go to step 10.

II.2.3 - Global Convergence

We will require in our convergence results, to the algorithm of the previous section, the three following assumptions:

G1: All the functions F, f_i, g_i^s are three times continuously differentiable in all their arguments.

G2: There exist m and M , both finite positive numbers, such that:

$$m \|p\|^2 \leq p^T D_i p \leq M \|p\|^2$$

for all $p \in \mathbb{R}^m$, $\forall i$ and all iterations.

G3: $r \geq \sum_{s \in S_i} \lambda_i^s$, for all i , for all iterations

where,

λ_i^s denotes the multiplier associated with

$$g_i^s(\bar{u}_i) + {}_i g_u^s(\bar{u}_i)^T (u_i - \bar{u}_i) \leq 0$$

at the solution of QP_i .

Proposition II.2.A

$$\tilde{\theta} \leq \sum_{i=0}^{N-1} [-\alpha_i^T D_i \alpha_i] < 0$$

Proof:

At the solution of QP_i , the following equation is satisfied

$$H_u^i = -D_i \alpha_i - \sum_{s \in S_i} \lambda_i^s {}_i g_u^s.$$

Then,

$$(i) \quad (H_u^i)^T \alpha_i = -\alpha_i^T D_i \alpha_i - \sum_{s \in S_i} \lambda_i^s ({}_i g_u^s)^T \alpha_i$$

But,

$$(ii) \quad ({}_i g_u^s)^T \alpha_i = -g_i^s(\bar{u}_i)$$

for all $s \in S_i$ such that: $\lambda_i^s \neq 0$.

From (i), (ii) and the definition of $\tilde{\theta}$ we have:

$$(iii) \quad \tilde{\theta} = \sum_{i=0}^{N-1} [-\alpha_i^T D_i \alpha_i + \sum_{s \in S_i} \lambda_i^s g_i^s(\bar{u}_i)] + r[\hat{\psi}(\{\bar{u}_i\}, \{\alpha_i\}) - \tilde{\psi}(\{\bar{u}_i\})].$$

But,

$$(iv) \quad \hat{\psi}(\{\bar{u}_i\}, \{\alpha_i\}) = 0$$

and

$$(v) \quad \tilde{\psi}(\{\bar{u}_i\}) \triangleq \sum_{i=0}^{N-1} \psi_i(\bar{u}_i).$$

Also,

$$(vi) \quad \psi_i(\bar{u}_i) \geq g_i^s(\bar{u}_i), \quad \forall s \in S_i.$$

Then, (iv), (v), (vi) into (iii) lead to:

$$(vii) \quad \tilde{\theta} \leq \sum_{i=0}^{N-1} [-\alpha_i^T D_i \alpha_i + \sum_{s \in S_i} \lambda_i^s \psi_i(\bar{u}_i)] - r[\sum_{i=0}^{N-1} \psi_i(\bar{u}_i)].$$

But, by G3,

$$(viii) \quad r \geq \sum_{s \in S_i} \lambda_i^s, \quad \text{for all } i$$

Then, from (viii) and (vii), we obtain:

$$\tilde{\theta} \leq \sum_{i=0}^{N-1} [-\alpha_i^T D_i \alpha_i]$$

The proposition then follows from G2

Proposition II.2.B

Let $(\{\bar{u}_i\}, \{\bar{x}_i\})$ be such that

$$\alpha_i = 0, \quad \forall i$$

Then,

$\{\bar{u}_i\}$ satisfies the first order optimality conditions to DCP.

Proof:

Clearly, DCP is equivalent to the mathematical programming problem below:

MP1

$$\min_{z \triangleq (x,u) \in \mathbb{R}^{(n+m) \times N}} \{F(z)\}$$

subject to:

$$\tilde{f}_i(z) = 0, \quad i = 0, \dots, N-1$$

$$g_i(z) \leq 0, \quad i = 0, \dots, N-1$$

$$x_0 = x^0; \quad \tilde{f}_i(z) \in \mathbb{R}^n; \quad g_i(z) \in \mathbb{R}^m$$

where,

$$\tilde{f}_i(z) \triangleq x_{i+1} - f_i(x_i, u_i)$$

and

$$g_i(z) \triangleq g_i(u_i)$$

Now, as $(\{\bar{x}_i\}, \{\bar{u}_i\})$ follows the dynamic equations and, by hypothesis, $\alpha_i = 0$ at the solution of QP_i for all i , we have:

(i) $(\{\bar{x}_i\}, \{\bar{u}_i\})$ is feasible for MP1.

At a Kuhn-Tucker point z^* of MP1, the following equations must be satisfied, for some multipliers

$$a \triangleq (a_0, \dots, a_{N-1}); \quad \tilde{\lambda} \triangleq (\tilde{\lambda}_0, \dots, \tilde{\lambda}_{N-1})$$

where,

$$a_i \in \mathbb{R}^n; \tilde{\lambda}_i \in \mathbb{R}^m.$$

$$(ii) \quad F_z(z^*) + \sum_{i=0}^{N-1} \tilde{f}_z^i(z^*)^T a_i + \sum_{i=0}^{N-1} g_z^i(z^*)^T \tilde{\lambda}_i = 0.$$

$$(iii) \quad \tilde{\lambda}_i^s \geq 0, \forall i, s.$$

$$(iv) \quad \tilde{\lambda}_i^s g_i^s(z^*) = 0, \forall i, s.$$

Let,

$$z^* \triangleq (x_1^*, \dots, x_N^*, u_0^*, \dots, u_{N-1}^*)$$

From (ii) and the definitions of $\tilde{f}_i(z)$ and $g_i(z)$, $(a, \tilde{\lambda})$ must satisfy:

$$\left. \begin{aligned} (v) \quad & [f_x^i(x_i^*, u_i^*)]^T a_i - a_{i-1} = 0 \\ (vi) \quad & -f_u^i(x_i^*, u_i^*)^T a_i + [g_u^i(u_i^*)]^T \tilde{\lambda}_i = 0 \end{aligned} \right\} i = 0, \dots, N-2$$

and

$$(vii) \quad a_{N-1} = -F_x(x_N^*)$$

$$(viii) \quad -[f_u^{N-1}(x_{N-1}^*, u_{N-1}^*)]^T a_{N-1} + [g_u^{N-1}(u_{N-1}^*)]^T \tilde{\lambda}_{N-1} = 0$$

Clearly, any $(a, \tilde{\lambda})$ satisfying (v) to (viii) would also satisfy (ii).

Now, at the solution of QP_i the following equations are satisfied:

$$H_u^i + D_i \alpha_i + g_u^i(\bar{u}_i)^T \lambda_i = 0, \forall i$$

$$\lambda_i^s \geq 0, \forall i, s$$

$$\lambda_i^s [g_i^s(\bar{u}_i) + g_u^s(\bar{u}_i) \alpha_i] = 0, \forall i, s$$

or, as $\alpha_i = 0$ by hypothesis:

$$(ix) \quad H_u^i + g_u^i (\bar{u}_i)^T \lambda_i + D_i \alpha_i = 0, \text{ for all } i$$

$$(x) \quad \lambda_i^s \geq 0, \text{ for all } i, s$$

$$(xi) \quad \lambda_i^s g_i^s (\bar{u}_i) = 0, \text{ for all } i, s.$$

Now, as $\alpha_i = 0$ for all i , from II.2.X and the definition of H_x we have:

$$(xii) \quad V_x^i (\bar{x}_i) = f_x^i V_x^{i+1} (\bar{x}_{i+1}), \quad i = 0, \dots, N-2$$

with the following boundary condition:

$$(viii) \quad V_x^{N-1} (\bar{x}_i) = f_x^{N-1} F_x (\bar{x}_N)$$

Also, as $\alpha_i = 0$ for all i , from (ix) and the definition of H_u^i , we have:

$$(xiv) \quad f_u^i V_x^{i+1} (\bar{x}_i) + g_u^i (\bar{u}_i)^T \lambda_i = 0, \quad i = 0, \dots, N-2$$

and

$$(xv) \quad f_u^{N-1} F_x (\bar{x}_N) + g_u (\bar{u}_i)^T \lambda_{N-1} = 0.$$

Hence,

$$(xvi) \quad \text{with } a_{N-1} = -F(\bar{x}_N) \text{ and } \tilde{\lambda}_{N-1} = \lambda_{N-1}$$

we have:

$$(xv) \cup (xiii) \Rightarrow (viii) \cup (vii)$$

Similarly, for $i = 0, \dots, N-2$, we have:

(xvii) with,

$$a_i = (f_x^{i+1})^T (f_x^{i+2})^T \dots (f_x^{N-1})^T F_x (\bar{x}_N)$$

and

$$\tilde{\lambda}_i = \lambda_i$$

we have:

$$(xiv) \cup (xii) \Rightarrow (v) \cup (vi)$$

In other words, with $(a, \tilde{\lambda})$ given by (xvi) and (xvii), the condition (ii) is satisfied at $(\{\bar{x}_i\}, \{\bar{u}_i\})$. But, for $\tilde{\lambda}_i = \lambda_i$, we have: $(x) \Rightarrow (iii)$ and $(xi) \Rightarrow (iv)$. Hence, the Kuhn-Tucker conditions (ii) to (iv) are satisfied at $(\{\bar{x}_i\}, \{\bar{u}_i\})$ and the proposition is proved.

Proposition II.2.C

$$\text{If } \theta(\{\bar{u}_i\}, \{\alpha_i\}, r) = 0$$

then,

$\{\bar{u}_i\}$ satisfies the first order optimality conditions to DCP

Proof:

From Proposition II.1.A, we have:

$$\tilde{\theta} \leq - \sum_{i=0}^{N-1} [\alpha_i^T D_i \alpha_i]$$

Hence, from G2, $\tilde{\theta} = 0 \Rightarrow \alpha_i = 0$ for all i .

The Proposition then follows from Proposition II.2.B.

Proposition II.2.D

Let $\{u_i'\}$ be any control sequence.

Let $U \subset \underbrace{R^m \times R^m \times \dots \times R^m}_{N \text{ times}}$ be any compact neighbourhood of $\{u_i'\}$.

Let $r, k > 0$ be arbitrary.

Then, there exists a function

$$(\varepsilon, \{u_i\}, \{\delta u_i'\}) \rightarrow \phi(\varepsilon, \{u_i\}, \{\delta u_i'\})$$

with

$$\phi(\varepsilon, \{u_i\}, \{\delta u_i'\}) \rightarrow 0 \text{ as } \varepsilon \rightarrow 0$$

uniformly in $(\{u_i\}, \{\delta u_i'\}) \in U \times \underbrace{\beta_k(0) \times \beta_k(0) \dots \times \beta_k(0)}_{N \text{ times}}$

such that

$$|\gamma(\{u_i + \varepsilon \delta u_i'\}, r) - \hat{\gamma}(\{u_i\}, \{\varepsilon \delta u_i'\}, r)| \leq \varepsilon \phi(\varepsilon, \{u_i\}, \{\delta u_i'\})$$

Proof:

Given in the appendix at the end of the chapter.

Theorem II.2.E

Let $\{u_i^*\}$ be any accumulation point of the sequence of control sequences $\{u_i\}^\ell$ constructed by the algorithm. Then,

$\{u_i^*\}$ satisfies the first order optimality conditions for DCP.

Proof:

Let J be a subset of iterations such that:

$$(i) \quad \{\bar{u}_i\}^\ell \xrightarrow{J} \{u_i^*\}.$$

As $x_0 = x^0$, there exists a $\{x_i^*\}$ such that:

$$(ii) \quad \{\bar{x}_i\}^{\ell} \xrightarrow{J} \{x_i^*\} .$$

Also, by G2, there exists $J' \subset J$ and $\{D_i^*\}$ such that:

$$(iii) \quad \{D_i\}^{\ell} \xrightarrow{J'} \{D_i^*\}$$

From (i), (ii), (iii) and Lemma 3.2 in Han (1977), we have:

$$(iv) \quad \text{there exists } \{\alpha_i^*\} \text{ with } \alpha_i^* \text{ solving } QP_i \\ \text{to } (\{\bar{x}_i\}, \{\bar{u}_i\}) = (\{x_i^*\}, \{u_i^*\}) \text{ and such that:}$$

$$\{\alpha_i\}^{\ell} \xrightarrow{J'} \{\alpha_i^*\} .$$

Suppose there exists an infinite subset $J'' \subset J'$ where u_i is set at step 8. Then, also from (i), (ii), (iii) and Lemma 3.2 in Han (1977), we have:

$$(v) \quad \{\delta u_i\}^{\ell} \xrightarrow{J''} \{\delta u_i^*\} ,$$

$$\delta u_i^* \text{ solving } DQP_i \text{ to } (\{\bar{x}_i\}, \{\bar{u}_i\}) = (\{x_i^*\}, \{u_i^*\}) .$$

But, if u_i is set at step 8, $\epsilon = 1$. Then from (i) we have:

$$(vi) \quad \{\delta u_i\}^{\ell} \xrightarrow{J''} 0 .$$

In this case, $QP_i \equiv DQP_i$ at $(\{x_i^*\}, \{u_i^*\})$. But then, from (iv), (v) and (vi), $\alpha_i^* = 0$ for all i and the theorem is proved by Proposition II.2.B. Otherwise suppose that no such J'' exists. That is, from one point onwards, u_i is set at step 12. If $\{\alpha_i^*\} \neq 0$ then, by Proposition II.2.B,

$$(vii) \quad \theta(\{u_i^*\}, \{\alpha_i^*\}, r) < 0$$

Now, from (i), (iv) and G1, we have:

$$(viii) \quad \theta(\{u_i\}^\ell, \{\alpha_i\}^\ell, r) \xrightarrow{J'} \theta(\{u_i\}^*, \{\alpha_i^*\}, r)$$

Hence, from (vii) and (viii), there exists $b > 0$ and $\ell^* < \infty$ such that:

$$(ix) \quad \theta(\{u_i\}^\ell, \{\alpha_i\}^\ell, r) \leq -b < 0$$

for all $\ell \in J'$, $\ell \geq \ell^*$.

But, since $\hat{\gamma}(\{u_i'\}, \{\varepsilon' \delta u_i'\}, r)$ is convex in ε' ,

$$(x) \quad \hat{\gamma}(\{u_i'\}, \{\varepsilon' \delta u_i'\}, r) - \gamma(\{u_i'\}, r) \leq \varepsilon' \theta(\{u_i'\}, \{\delta u_i'\}, r)$$

for any sequences $\{u_i'\}, \{\delta u_i'\}$ and all $\varepsilon' \in [0, 1]$.

Also, because of (i), (ii), (iv), G1 and G2, it is possible to show [see proof of Proposition I.1.C] that:

(xi) there exists a $K < \infty$ such that:

$$\|\alpha_i^\ell\| \leq K, \quad \forall i, \ell \in J'$$

But, from (i), there exists a compact neighbourhood U of $\{u_i^*\}$ such that:

$$(xii) \quad \{\bar{u}_i\}^\ell \in U, \text{ for all } \ell \in J', \ell \geq \ell', \text{ say.}$$

Then, from (xi), (xii) and Proposition II.2.D, there exists an integer $k' \in \{0, 1, \dots\}$ such that

$$(xiii) \quad |\gamma(\{\bar{u}_i + \delta^{k'} \alpha_i\}^\ell, r) - \hat{\gamma}(\{\bar{u}_i\}^\ell, \{\delta^{k'} \alpha_i\}^\ell, r)| \leq \delta^{k'} b / 4$$

for all $\ell \in J'$, $\ell \geq \ell'$, say.

Now, from (i), $(u_{i-1} - \bar{u}_{i-1}) \xrightarrow{J} 0$, for all i , hence

$$\Delta_i \xrightarrow{J} 0, \text{ for all } i.$$

Then, there exists a $\bar{\ell} < \infty$ such that, for all $\ell \in J$
 $\ell \geq \bar{\ell}$, we have :

$$(xiv) \quad |\gamma(\{\bar{u}_i + \delta^{k'} \alpha_i + \beta_i \Delta_i\}^\ell, r) - \gamma(\{\bar{u}_i + \delta^{k'} \alpha_i\}^\ell, r)| \leq \delta^{k'} b/4 .$$

From (xiii) and (xiv), there exists a $\hat{\ell} < \infty$ such that:

$$(xv) \quad |\gamma(\{\bar{u}_i + \delta^{k'} \alpha_i + \beta_i \Delta_i\}^\ell, r) - \hat{\gamma}(\{\bar{u}_i\}^\ell, \{\delta^{k'} \alpha_i\}, r)| \leq \delta^{k'} b/2$$

for all $\ell \in J'$, $\ell \geq \hat{\ell}$.

From (ix), (x) and (xv) we obtain:

(xvi) there exists $\tilde{\ell} < \infty$ such that:

$$\gamma(\{\bar{u}_i + \delta^{k'} \alpha_i + \beta_i \Delta_i\}^\ell, r) - \gamma(\{\bar{u}_i\}^\ell, r) \leq -\delta^{k'} b/2$$

for all $\ell \in J'$, $\ell \geq \tilde{\ell}$.

Hence, from the way ε is set at step 11 and step 10 with u_i computed at step 12, we have:

$$(xvii) \quad \varepsilon \geq \delta^{k'} , \text{ for all } \ell \in J', \ell \geq \tilde{\ell} .$$

From (xvi) and (xvii) we have:

$$(xviii) \quad \gamma(\{\bar{u}_i\}^{\ell+1}, r) - \gamma(\{\bar{u}_i\}^\ell, r) \leq -\delta^{k'} b/2 < 0$$

for all $\ell \in J'$, ℓ large enough .

But (xviii) contradicts (i), the contradiction coming from (vii) through (ix).

So, $\theta(\{u_i^*\}, \{\alpha_i^*\}, r) = 0$ and the theorem follows from Proposition II.2.C.

II.2.4 - Rate of Convergence

If we eliminate the variable x in DCP, through the dynamic equations and initial condition $x_0 = x^0$, we would arrive at the following mathematical programming problem which is equivalent to DCP.

MP2

$$\min_{(u \in \mathbb{R}^{m \times N})} \tilde{F}(u) \triangleq F(f_{N-1}(f_{N-2}(\dots f_0(x_0, u_0), u_1), u_2) \dots, u_{N-1})$$

subject to:

$$g_i^s(u_i) \leq 0, \quad i = 0, \dots, N-1$$

$$s \in S_i.$$

In the analysis of the rate of convergence of the algorithm introduced in this chapter, we will require G1 and the further assumptions below:

G4: $\{u_i\}^k \rightarrow u^*$ (Kuhn-Tucker point for MP2).

G5: At each K-T couple for MP2, the second-order sufficiency conditions are satisfied with strict complementarity slackness and at these points the same second-order sufficiency conditions also hold for each DQP_i, $i = 0, \dots, N-1$.

G6: For all $u \in \mathbb{R}^{m \times N}$ the vectors ${}_i g_u^s \in I_i(u)$ are linearly independent for all $i = 0, \dots, N-1$, with

$$I_i(u) \triangleq \{ {}_i g_u^s(u) \mid g_i^s(u) = \psi_i(u) \}.$$

G7: The step length ε is unity.

Suppose we apply to MP2 a slightly modified version of Wilson's method [see Wilson (1963)] where the estimate of the multipliers $\bar{\lambda}_i(\bar{u}_i)$ is used to compute the estimate of the Hessian matrix. That is:

at each iteration we find u^w to solve
the quadratic problem QP2 and iteratively
make:

$$\bar{u}^{\ell+1} = u^w$$

with QP2 being the following quadratic problem:

QP2

$$\min_{(u \in \mathbb{R}^{m \times N})} \tilde{F}_u(\bar{u})^T(u - \bar{u}) + \frac{1}{2}(u - \bar{u})^T P(\bar{u})(u - \bar{u})$$

subject to:

$$g_i^s(u_i) + {}_i g_u^s(\bar{u}_i)(u - \bar{u}_i) \leq 0$$

for all $i = 0, \dots, N-1$

for all $s = 1, \dots, \ell_i$

where,

$$P(\bar{u}) \triangleq \tilde{F}_{uu}(\bar{u}) + \sum_{i=0}^{N-1} \left[\sum_{s \in S_i} \bar{\lambda}_i^s(\bar{u}_i) {}_i g_{uu}^s(\bar{u}_i) \right].$$

Proposition II.2.F

Let the sequence of points u^w be such that:

$$u^w \rightarrow u^* \text{ (K-T point to MP2)}$$

Let H5 and H6 be satisfied. Then,

$$u^w \rightarrow u^*$$

at least R-quadratically.

Proof:

Given in the appendix at the end of the chapter.

Proposition II.2.G

Let $u^w \rightarrow u^*$ (K-T point to MP2)

Let H4, H5, H6, H7 be satisfied. Then,

for $\bar{u}$ near enough to u^* both the
quadratic problem QP2 and the sequence
 $\{DQP_i\}$ will predict correctly the
active constraints to MP2 at u^* .

Proof:

Given in the appendix at the end of the chapter.

Observation:

As shown in the proof of the last proposition, the sequence $\{\lambda_i^D\}$ of multipliers to $\{DQP_i\}$ will converge to λ^* , the multiplier associated with u^* at the solution of MP2. Hence, the strict complementary slackness holds for $\{DQP_i\}$ as a consequence of being assumed for MP2. The sole reason for including $\{DQP_i\}$ in G5 was to guarantee that:

II.2.XII

If $p \in R^m$ is such that: ${}_i g_u^T p = 0$,

then:

$$p^T D_i p > 0, \text{ for all } i$$

with D_i being computed at $\bar{u} = u^*$.

Note, however, that II.2.XII is milder than assuming the positive definiteness of $\{D_i\}$ at $\bar{u} = u^*$.

Theorem II.2.H

Let $\{u_i\}^\ell$ be constructed by the algorithm of this chapter.

Let G1, G4, G5, G6 and G7 be satisfied. Then,

$$\{u_i\}^\ell \rightarrow u^*, \text{ as } \ell \rightarrow \infty.$$

at least R-quadratically.

Proof:

If we consider $\{\bar{u}_i\}$ as a perturbation of u^* we can regard subproblem QP2 as a perturbation of MP2. Hence, calling $p \triangleq \{\bar{u}_i\}$ and $\bar{p} \triangleq u^*$, we can, as suggested by Robinson (1974), define the following maps:

$$L(u, \lambda, p) = \theta(u, p) + \lambda^T G(u, p)$$

where,

$$\theta(u, p) \triangleq \tilde{F}_u(\bar{u})^T (u - \bar{u}) + \frac{1}{2} (u - \bar{u})^T P(\bar{u}) (u - \bar{u})$$

and

$$\lambda^T G(u, p) \triangleq \begin{bmatrix} (\lambda_0)^T [g_0(\bar{u}) + {}_0g_u(\bar{u}_0)(u_0 - \bar{u}_0)] \\ \vdots \\ (\lambda_{N-1})^T [g_{N-1}(\bar{u}_{N-1}) + {}_{N-1}g_u(\bar{u}_{N-1})(u_{N-1} - \bar{u}_{N-1})] \end{bmatrix}$$

Also define:

$$\begin{aligned}
 h(u, \lambda, p) \triangleq & \{L_u(u, \lambda, p), \lambda_0^1 [g_0^1(\bar{u}_0) + {}_0g_u^1(\bar{u}_0)(u_0 - \bar{u}_0)], \dots \\
 & \dots, \lambda_0^{\ell_0} [g_0^{\ell_0}(\bar{u}_0) + {}_0g_u^{\ell_0}(\bar{u}_0)(u_0 - \bar{u}_0)], \dots \\
 & \dots, \lambda_{N-1}^1 [g_{N-1}^1(\bar{u}_{N-1}) + {}_{N-1}g_u^1(\bar{u}_{N-1})(u_{N-1} - \bar{u}_{N-1})], \dots \\
 & \dots, \lambda_{N-1}^{\ell_{N-1}} [g_{N-1}^{\ell_{N-1}}(\bar{u}_{N-1}) + {}_{N-1}g_u^{\ell_{N-1}}(\bar{u}_{N-1})(u_{N-1} - \bar{u}_{N-1})]\}^T.
 \end{aligned}$$

Now, because of G5 we can use Theorem 2.1 of Robinson (1974) to see that:

there exists a continuous map $Z(\cdot)$

which for any p near enough to $\bar{p}$

associates the K-T point to QP2 and

such that:

$$Z(\bar{p}) = (u^*, \lambda^*) = \text{K-T couple to MP2.}$$

As the conditions to Theorem 2.2 in the last paper referred to are also satisfied, we have:

(i) for any $\varepsilon \in (0, 1]$, $\exists$ open neighbourhoods

$M_\varepsilon(\bar{p})$ and $N_\varepsilon(u^*, \lambda^*)$ such that:

for any $\tilde{p} \in M_\varepsilon(\bar{p})$ and any $\tilde{z} \in N_\varepsilon(z^*)$

we have:

$$\|\tilde{z} - Z(\tilde{p})\| \leq (1-\varepsilon)^{-1} \|h_u(z^*, p^*)^{-1}\| \|h(\tilde{z}, \tilde{p})\|$$

with

$$z^* \triangleq (u^*, \lambda^*).$$

Now, G5 and G6 ensure the existence of the inverse above.

Therefore, from (i), we obtain:

(ii) there exists a $c_1 < \infty$ such that, for

$\bar{u}$ near enough to u^* , we have:

$$\|z^D - Z(\bar{u})\| \leq c_1 \|h(z^D, \bar{u})\|$$

where,

$$z^D \triangleq (u^D, \lambda^D) \triangleq \underbrace{[(u_0^D, \lambda_0^D), (u_1^D, \lambda_1^D), \dots, (u_{N-1}^D, \lambda_{N-1}^D)]}_{\text{K-T couples to DQP}_i}$$

Also let,

$$z^W \triangleq (u^W, \lambda^W) = Z(\bar{u}).$$

Now,

$$g_{u_i, u_j}^s = g_{u_j, u_i}^s = 0, \text{ for all } s, j, i.$$

Hence, recalling the definition of $P(\cdot)$ we have:

$$\begin{aligned} & \left[\begin{array}{l} \tilde{F}_{u_0} + \tilde{F}_{u_0 u_1} (u_1^D - \bar{u}_1) + \tilde{F}_{u_0 u_2} (u_2^D - \bar{u}_2) + \dots + \tilde{F}_{u_0 u_{N-1}} (u_{N-1}^D - \bar{u}_{N-1}) + \bar{A}_0 (u_0^D - \bar{u}_0) + [\lambda_0^D]^T {}_0 g_u \\ \tilde{F}_{u_1} + \tilde{F}_{u_1 u_0} (u_0^D - \bar{u}_0) + \tilde{F}_{u_1 u_2} (u_2^D - \bar{u}_2) + \dots + \tilde{F}_{u_1 u_{N-1}} (u_{N-1}^D - \bar{u}_{N-1}) + \bar{A}_1 (u_1^D - \bar{u}_1) + [\lambda_1^D]^T {}_1 g_u \\ \vdots \\ \tilde{F}_{u_{N-1}} + \tilde{F}_{u_{N-1} u_0} (u_0^D - \bar{u}_0) + \tilde{F}_{u_{N-1} u_1} (u_1^D - \bar{u}_1) + \dots + \tilde{F}_{u_{N-1} u_{N-2}} (u_{N-2}^D - \bar{u}_{N-2}) + \bar{A}_{N-1} (u_{N-1}^D - \bar{u}_{N-1}) + [\lambda_{N-1}^D]^T g_u \end{array} \right] \\ & h(z^D, \bar{u}) = \left[\begin{array}{l} {}^1 \lambda_0^D [{}_0 g_0^1(\bar{u}_0) + {}_0 g_u^1(\bar{u}_0) (u_0^D - \bar{u}_0)] \\ \vdots \\ {}^\ell \lambda_0^D [{}_0 g_0^\ell(\bar{u}_0) + {}_0 g_u^\ell(\bar{u}_0) (u_0^D - \bar{u}_0)] \\ \vdots \\ {}^1 \lambda_{N-1}^D [{}_{N-1} g_{N-1}^1(\bar{u}_{N-1}) + {}_{N-1} g_u^1(\bar{u}_{N-1}) (u_{N-1}^D - \bar{u}_{N-1})] \\ \vdots \\ {}^\ell \lambda_{N-1}^D [{}_{N-1} g_{N-1}^\ell(\bar{u}_{N-1}) + {}_{N-1} g_u^\ell(\bar{u}_{N-1}) (u_{N-1}^D - \bar{u}_{N-1})] \end{array} \right] \end{aligned}$$

where,

$$\bar{A}_i \triangleq \tilde{F}_{u_i u_i} + [\bar{\lambda}(\bar{u}_i)]^T i g_{uu}$$

From the recursive equations we have:

$$\begin{aligned} \text{(iv)} \quad H_u^i &= (f_u^i)^T (f_x^{i+1})^T \dots (f_x^{N-1})^T F_x + (f_u^i)^T (f_x^{i+1})^T \dots (f_x^{N-2})^T \beta_{N-1}^T \\ &\quad [H_u^{N-1} + \lambda_{N-1}^T g_u] \\ &\quad + (f_u^i)^T (f_x^{i+1})^T \dots (f_x^{N-3})^T \beta_{N-2}^T [H_u^{N-2} + \lambda_{N-2}^T g_u] + \dots \\ &\quad \dots + f_u^i \beta_{i+1}^T [H_u^{i+1} + \lambda_{i+1}^T g_u] \end{aligned}$$

where

λ_i is the multiplier associated with α_i at QP_i

$$\begin{aligned} \text{(v)} \quad B_i &= (f_{ux}^i)^T (f_x^{i+1})^T \dots (f_x^{N-1})^T F_x + (f_u^i)^T [(f_{xx}^{i+1})^T (f_x^{i+2})^T \dots \\ &\quad \dots (f_x^{N-1})^T F_x + (f_x^{i+1})^T (f_{xx}^{i+2})^T \dots F_x f_x^{i+1} \\ &\quad + (f_x^{i+1})^T (f_x^{i+2})^T (f_{xx}^{i+3})^T (f_x^{i+4})^T \dots F_x f_x^{i+1} f_x^{i+2} + \dots \\ &\quad \dots + (f_x^{i+1})^T \dots (f_x^{N-1})^T F_{xx} f_x^{N-1} \dots f_x^{i+1}] f_x^i \\ &\quad - (f_u^i)^T (f_x^{i+1})^T \dots (f_x^{N-2})^T [\beta_{N-1}^T D_{N-1} \beta_{N-1}] f_x^{N-2} \dots f_x^i - (f_u^i)^T (f_x^{i+1})^T \dots \\ &\quad \dots (f_x^{N-3})^T [\beta_{N-2}^T D_{N-2} \beta_{N-2}] f_x^{N-3} \dots f_x^i \\ &\quad - \dots - (f_u^i)^T [\beta_{i+1}^T D_{i+1} \beta_{i+1}] f_x^i. \end{aligned}$$

Observation:

$$(f_x^i)^T (f_x^{i+1})^T \dots F_x = \sum_{j=1}^n \{ (f_{x_j}^i)^T (f_{x_j}^{i+1})^T \dots F_{x_j} \} .$$

Also,

$$\begin{aligned} \text{(vi)} \quad C_i &= (f_{uu}^i)^T (f_x^{i+1})^T \dots (f_x^{N-1})^T F_x + (f_u^i)^T [(f_{xx}^{i+1})^T (f_x^{i+2})^T \dots (f_x^{N-1})^T F_x \\ &\quad + (f_x^{i+1})^T (f_{xx}^{i+2})^T (f_x^{i+3})^T \dots (f_x^{N-1})^T F_x f_x^{i+1} + \dots + (f_x^{i+1})^T \dots \\ &\quad \dots (f_x^{N-1})^T F_{xx} f_x^{N-1} \dots f_x^{i+1}] f_u^i \\ &\quad - (f_u^i)^T (f_x^{i+1})^T \dots (f_x^{N-2})^T [\beta_{N-1}^T D_{N-1} \beta_{N-1}] f_x^{N-2} \dots f_x^{i+1} f_u^i \\ &\quad - (f_u^i)^T (f_x^{i+1})^T \dots (f_x^{N-3})^T [\beta_{N-2}^T D_{N-2} \beta_{N-2}] f_x^{N-3} \dots f_x^{i+1} f_u^i - \dots \\ &\quad \dots - (f_u^i)^T [\beta_{i+1}^T D_{i+1} \beta_{i+1}] f_u^i . \end{aligned}$$

Now, from the definition of $\tilde{F}(\cdot)$, we have:

$$\text{(vii)} \quad \tilde{F}_{u_i} = (f_u^i)^T (f_x^{i+1})^T \dots (f_x^{N-1})^T F_x$$

Hence, for $j > i$

$$\begin{aligned} \text{(viii)} \quad \tilde{F}_{u_i u_j} &= (f_u^i)^T (f_x^{i+1})^T \dots (f_x^{j-1})^T (f_{ux}^j)^T (f_x^{j+1})^T \dots (f_x^{N-1})^T F_x \\ &\quad + (f_u^i)^T (f_x^{i+1})^T \dots (f_x^{j-1})^T \{ (f_u^j)^T [(f_{xx}^{j+1})^T (f_x^{j+2})^T \dots (f_x^{N-1})^T F_x \\ &\quad + (f_x^{j+1})^T (f_{xx}^{j+2})^T (f_x^{j+3})^T \dots (f_x^{N-1})^T F_x f_x^{j+1} + (f_x^{j+1})^T \\ &\quad (f_x^{j+2})^T (f_{xx}^{j+3})^T \dots F_x f_x^{j+1} f_x^{j+2} \\ &\quad + \dots + (f_x^{j+1})^T (f_x^{j+2})^T \dots (f_x^{N-1})^T F_{xx} f_x^{N-1} \dots f_x^{j+1}] f_x^j \} \end{aligned}$$

while,

$$\begin{aligned}
 \text{(ix)} \quad \tilde{F}_{u_i u_i} = & (f_{uu}^i)^T (f_x^{i+1})^T \dots (f_x^{N-1})^T F_x + (f_u^i)^T [(f_{xx}^{i+1})^T (f_x^{i+2})^T \dots (f_x^{N-1})^T F_x \\
 & + (f_x^{i+1})^T (f_{xx}^{i+2})^T (f_x^{i+3})^T \dots F_x f_x^{i+1} + \dots + (f_x^{i+1})^T \dots \\
 & \dots (f_x^{N-1})^T F_{xx} f_x^{N-1} \dots f_x^{i+1}] f_u^i .
 \end{aligned}$$

From (iv) and (vii) we obtain

$$\begin{aligned}
 \text{(x)} \quad \tilde{F}_{u_i} = & H_u^i - (f_u^i)^T (f_x^{i+1})^T \dots (f_x^{N-2})^T \beta_{N-1}^T [H_u^{N-1} + \lambda_{N-1}^T \beta_{N-1} g_u] \\
 & - (f_u^i)^T (f_x^{i+1})^T \dots (f_x^{N-3})^T \beta_{N-2}^T [H_u^{N-2} + \lambda_{N-2}^T \beta_{N-2} g_u] - \dots \\
 & - (f_u^i)^T \beta_{i+1}^T [H_u^i + \lambda_{i+1}^T \beta_{i+1} g_u] .
 \end{aligned}$$

From (v) and (viii), we obtain for $j > 1$

$$\begin{aligned}
 \text{(xi)} \quad \tilde{F}_{u_i u_j} = & (f_u^i)^T (f_x^{i+1})^T \dots (f_x^{j-1})^T \beta_j^T + (f_u^i)^T (f_x^{i+1})^T \dots (f_x^{N-2})^T \\
 & [\beta_{N-1}^T D_{N-1} \beta_{N-1}] f_x^{N-2} \dots f_x^{j+1} f_u^j \\
 & + (f_u^i)^T (f_x^{i+1})^T \dots (f_x^{N-3})^T [\beta_{N-2}^T D_{N-2} \beta_{N-2}] f_x^{N-3} \dots \\
 & \dots f_x^{j+1} f_u^j + \dots \\
 & + (f_u^i)^T (f_x^{i+1})^T \dots (f_x^j)^T [\beta_{j+1}^T D_{j+1} \beta_{j+1}] f_u^j .
 \end{aligned}$$

From (vi) and (ix) we have:

$$\begin{aligned}
 \text{(xii)} \quad \tilde{F}_{u_i, u_i} &= C_i + (f_u^i)^T (f_x^{i+1})^T (f_x^{i+2})^T \dots (f_x^{N-2})^T [\beta_{N-1}^T D_{N-1} \beta_{N-1}] \\
 &\quad f_x^{N-2} \dots f_x^{i+1} f_u^i \\
 &\quad + (f_u^i)^T (f_x^{i+1})^T \dots (f_x^{N-3})^T [\beta_{N-2}^T D_{N-2} \beta_{N-2}] f_x^{N-3} \dots f_x^{i+1} f_u^i + \dots \\
 &\quad + (f_u^i)^T [\beta_{i+1}^T D_{i+1} \beta_{i+1}] f_u^i
 \end{aligned}$$

Now, recalling that (α_i, λ_i) is the K-T couple of QP_i we have:

$$H_u^i + \lambda_i^T g_u = -D_i(\alpha_i - \bar{u}_i)$$

But, from H5 the sequence of sub-problems $\{DQP_i\}$ also satisfies the second-order sufficiency conditions with strict complementarity slackness, at the point $\{\bar{u}_i\} = u^*$.

Hence, from theorem 2.1 of Robinson (1974), for all $\{\bar{u}_i\}$ near enough to u^* , DQP_i will also have a unique K-T couple $(\delta u_i, \lambda_i^D)$ where the second order sufficiency conditions are satisfied with strict complementary slackness. Then, considering DQP_i as a perturbation of QP_i , Theorem 6 of Fiacco and McCormick (1968) shows that a first order approximation to $(\delta u_i - \alpha_i)$ is:

$$(\alpha_i - \delta u_i) = D_i^{-1} B_i \Delta_i$$

Moreover, as the functions F , f_i and g_i^s are all supposed to be three times continuously differentiable, we can, similarly to Theorem 15 of Fiacco and McCormick (1968), see that:

$$\alpha_i = \delta u_i + D_i^{-1} B_i \Delta_i + \bar{e}_1$$

where

$$\|\bar{e}_1\| \approx \|\Delta_i\|^2$$

Hence,

$$(xiii) \quad H_u^i + \lambda_i^T g_u = -D_i(u_i^D - \bar{u}_i) - B_i \Delta_i + \bar{e}_i$$

where,

$$\{u_i^D\} = \{\bar{u}_i + \delta u_i\}$$

with

δu_i being the solution of DQP_i .

From (x), (xiii) and recalling that $B_i = -D_i \beta_i$, we obtain:

$$\begin{aligned} (xiv) \quad \tilde{F}_{u_i} &= H_u^i - (f_u^i)^T (f_x^{i+1})^T \dots (f_x^{N-2})^T B_{N-1}^T (u_{N-1}^D - \bar{u}_{N-1}) \\ &\quad - (f_u^i)^T (f_x^{i+1})^T \dots (f_x^{N-2})^T [\beta_{N-1}^T D_{N-1} \beta_{N-1}] \Delta_{N-1} \\ &\quad - (f_u^i)^T (f_x^{i+1})^T \dots (f_x^{N-3})^T B_{N-2}^T (u_{N-2}^D - \bar{u}_{N-2}) \\ &\quad - (f_u^i)^T (f_x^{i+1})^T \dots (f_x^{N-3})^T [\beta_{N-2}^T D_{N-2} \beta_{N-2}] \Delta_{N-2} \\ &\quad \vdots \\ &\quad - (f_u^i)^T B_{i+1}^T (u_{i+1}^D - \bar{u}_{i+1}) - (f_u^i)^T [\beta_{i+1}^T D_{i+1} \beta_{i+1}] \Delta_{i+1} \\ &\quad + \bar{e}_1 \end{aligned}$$

where,

$$\|\bar{e}_1\| \approx \|\Delta_i\|^2$$

From (xi), (xiv) and the definition of Δ_i , we have:

$$\begin{aligned}
(xv) \quad H_u^i &= \tilde{F}_{u_i} + \tilde{F}_{u_i u_{i+1}} (u_{i+1}^D - \bar{u}_{i+1}) + \tilde{F}_{u_i u_{i+2}} (u_{i+2}^D - \bar{u}_{i+2}) + \dots + \tilde{F}_{u_i u_{N-1}} (u_{N-1}^D - \bar{u}_{N-1}) \\
&+ (f_u^i)^T (f_x^{i+1})^T \dots (f_x^{N-2})^T [\beta_{N-1}^T D_{N-1} \beta_{N-1}] [f_x^{N-2} f_x^{N-3} \dots f_x^{i+1} f_u^i (u_{i-1}^D - \bar{u}_{i-1}) \\
&+ f_x^{N-2} f_x^{N-3} \dots f_x^i f_u^{i-1} (u_{i-1}^D - \bar{u}_{i-1}) + \dots + f_x^{N-2} \dots f_x^1 f_u^0 (u_0^D - \bar{u}_0)] \\
&+ (f_u^i)^T (f_x^{i+1})^T \dots (f_x^{N-3})^T [\beta_{N-2}^T D_{N-2} \beta_{N-2}] [f_x^{N-3} \dots f_x^{i+1} f_u^i (u_{i-1}^D - \bar{u}_{i-1}) \\
&+ f_x^{N-3} f_x^{N-4} \dots f_x^i f_u^{i-1} (u_{i-1}^D - \bar{u}_{i-1}) + \dots + f_x^{N-3} \dots f_x^1 f_u^0 (u_0^D - \bar{u}_0)] \\
&\vdots \\
&(f_u^i)^T [\beta_{i+1}^T D_{i+1} \beta_{i+1}] [f_u^i (u_i^D - \bar{u}_i) + f_x^i f_u^{i-1} (u_{i-1}^D - \bar{u}_{i-1}) + \dots \\
&+ f_x^i f_x^{i-2} \dots f_x^1 f_u^0 (u_0^D - \bar{u}_0)] - \bar{e}_i.
\end{aligned}$$

From (xv) and (xii) we obtain

$$\begin{aligned}
(xvi) \quad H_u^i + C_i (u_i^D - \bar{u}_i) &= \tilde{F}_{u_i} + \tilde{F}_{u_i u_i} (u_i^D - \bar{u}_i) + \tilde{F}_{u_i u_{i+1}} (u_{i+1}^D - \bar{u}_{i+1}) + \dots + \tilde{F}_{u_i u_{N-1}} (u_{N-1}^D - \bar{u}_{N-1}) \\
&+ (f_u^i)^T (f_x^{i+1})^T \dots (f_x^{N-2})^T [\beta_{N-1}^T D_{N-1} \beta_{N-1}] [f_x^{N-2} \dots f_x^i f_u^{i-1} (u_{i-1}^D - \bar{u}_{i-1}) \\
&+ f_x^{N-2} \dots f_x^{i-1} f_u^{i-2} (u_{i-2}^D - \bar{u}_{i-2}) + \dots + f_x^{N-2} \dots f_x^1 f_u^0 (u_0^D - \bar{u}_0)] \\
&\vdots \\
&(f_u^i)^T [\beta_{i+1}^T D_{i+1} \beta_{i+1}] [f_x^i f_u^{i-1} (u_{i-1}^D - \bar{u}_{i-1}) + \dots + f_x^i \dots f_x^1 f_u^0 (u_0^D - \bar{u}_0)] \\
&- \bar{e}_i.
\end{aligned}$$

As $\tilde{F}_{u_i u_j} = (\tilde{F}_{u_j u_i})^T$, from (xi) and the definition of Δ_i we

have :

$$\begin{aligned}
 \text{(xvii)} \quad B_i \Delta_i &= \sum_{j=0}^{i-1} \tilde{F}_{u_i u_j} (u_j^D - \bar{u}_j) \\
 &- (f_u^i)^T (f_x^{i+1})^T \dots (f_x^{N-2})^T [\beta_{N-1}^T D_{N-1} \beta_{N-1}] [f_x^{N-2} \dots f_x^i f_u^{i-1} (u_{i-1}^D - \bar{u}_{i-1}) \\
 &+ f_x^{N-2} \dots f_x^{i-1} f_u^{i-2} (u_{i-2}^D - \bar{u}_{i-2}) + \dots + f_x^{N-2} \dots f_x^1 f_u^0 (u_0^D - \bar{u}_0)] \\
 &- (f_u^i)^T [\beta_{i+1}^T D_{i+1} \beta_{i+1}] [f_x^i f_u^{i-1} (u_{i-1}^D - \bar{u}_{i-1}) + \dots + f_x^i \dots f_x^1 f_u^0 (u_0^D - \bar{u}_0)].
 \end{aligned}$$

From (xvi) and (xvii) we have:

$$\text{(xviii)} \quad H_u^i + B_i \Delta_i + C_i (u_i^D - \bar{u}_i) = \tilde{F}_{u_i} + \sum_{j=0}^{N-1} \tilde{F}_{u_i u_j} (u_j^D - \bar{u}_j) - \bar{e}_i.$$

By definition

$$\text{(xix)} \quad D_i = C_i + [\bar{\lambda}(\bar{u}_i)]^T {}_i g_{uu}$$

and

$$\text{(xx)} \quad \bar{A}_i = \tilde{F}_{u_i u_i} + [\bar{\lambda}(\bar{u}_i)]^T {}_i g_{uu}.$$

Now, at the solution of DQP_i we have:

$$\text{(xxi)} \quad H_u^i + B_i \Delta_i + D_i (u_i^D - \bar{u}_i) + [\lambda_i^D]^T {}_i g_u = 0$$

and

$$\begin{aligned}
 \text{(xxii)} \quad {}_s \lambda_i^D [{}_i g_i^s(\bar{u}_i) + {}_i g_u^s(\bar{u}_i) (u_i^D - \bar{u}_i)] &= 0, \text{ for all } i = 0, \dots, N-1 \\
 &\text{for all } s = 1, \dots, \ell_i.
 \end{aligned}$$

Hence, (xxii), (xxi), (xx) and (xix) into (iii) lead to:

$$(xxiii) \quad \|h(Z^D, \bar{u})\| \approx \|\Delta_i\|^2$$

But, because of H5, it is possible to show [see proofs of Lemma 4 and Lemma 5 in Powell (1978a)] that:

(xxiv) there exists a constant $c < \infty$
such that:

$$\|u^D - \bar{u}\| \leq c \|\bar{u} - u^*\|$$

Then, from (xxiii), (xxiv) and the definition of Δ_i , we obtain:

(xxv) there exists $c^* < \infty$ such that:

$$\|h(Z^D, \bar{u})\| \leq c^* \|\bar{u} - u^*\|^2$$

Then, from (ii) and (xxv), we can see that:

(xxvi) there exists a $c^0 < \infty$ such that:

$$\|u^D - u^W\| \leq c^0 \|\bar{u} - u^*\|^2$$

where,

u^W is the successor obtained by the modified Wilson method.

Now, because of Proposition II.2.F, we already know that:

(xxvii) there exists a $c' < \infty$ such that:

$$\|u^W - \bar{u}\| \leq c' \|\bar{u} - u^*\|^2$$

But,

$$(xxviii) \quad \|u^D - \bar{u}\| \leq \|u^D - u^W\| + \|u^W - \bar{u}\|$$

Hence, from (xxviii), (xxvii) and (xvii), we obtain:

there exists a $\bar{c} < \infty$ such that:

$$\|u^D - \bar{u}\| \leq \bar{c} \|\bar{u} - u^*\|^2$$

which is the desired result.

II.2.5 - Remarks

Note that the algorithm presented in this chapter is not a discrete time version of the algorithm suggested by Jacobson and Mayne (1970) for continuous time control problems with control inequality constraints. While here the sequence $\{\alpha_i, i = N-1, \dots, 0\}$ is found backwards by solving the sequence of quadratic programming problems $\{QP_i\}$, the discretization of the corresponding computation in the algorithm of Jacobson and Mayne (1970) would require the solution, for $i = N-1, \dots, 0$, of the following non-linear programming problem:

II.2.XIII

$$\min_{\delta u \in \mathbb{R}^m} H_i(\bar{x}_i, \bar{u}_i + \delta u, v_x^i)$$

subject to

$$\bar{u}_i + \delta u \in U = \text{feasible controls set.}$$

Now, in contrast to the solution of the quadratic subproblem, which can be found in a finite number of steps, the above minimization, in general, involves an infinite iterative process, where the solution is only reached as a limit. If the method suggested by Wilson (1963) is applied to solve II.2.XIII a sequence of quadratic programming problems must be solved.

In this sense, the solution of QP_i may be thought of as the first iteration of a process intended to solve the nonlinear programming problem given by II.2. XIII. That is, instead of pursuing the complete minimization indicated by II.2. XIII we only take, at each iteration, one step towards its solution. This feature may greatly reduce the number of computations. On the other hand, as pointed out by Jacobson and Mayne (1970), the solution to II.2. XIII can be found even when the matrix $H_{uu}^i(\bar{x}_i, \bar{u}_i, v_x^i)$ is not positive definite, although their algorithm would require the positive definiteness of the matrix $[H_{uu}^i + \lambda_i g_{uu}^i]$ when evaluated at $\bar{u}_i + \delta u_i^*$ (λ_i being the multiplier associated with δu_i^* at the solution of II.2. XIII).

Here, the matrices $D_i = H_{uu}^i + \bar{\lambda}_i(\bar{u}_i)g_{uu}$ are assumed to be positive definite for all the nominal sequences, when proving the global convergence. Given the way D_i is constructed, assumption G2 is perhaps excessively strong. Some remedies can be suggested to overcome this drawback.

As in the changes made to the Basic Algorithm of the first chapter, to cope with the secant update (see algorithm of section I.3.2), a multiple of the identity matrix can be added to D_i whenever the solution of QP_i fails to satisfy some tests. This device may enforce global convergence without requiring the rather restrictive assumption G2. Moreover, if γ_i remains equal to zero for all i large enough, or $\gamma_i \rightarrow 0$ as $\{\bar{u}_i\} \rightarrow \{u_i^*\}$, a quadratic rate of convergence could still be achieved (see discussion in section I.3.5).

Another approach would be the replacement of QP_i by some variable metric method. That is, instead of computing D_i directly from the recursive equations, an update could be used keeping the matrices $\{D_i\}$ positive definite yet allowing the achievement of a superlinear rate of convergence. This idea will be briefly discussed in the next section.

The role played by the positive definiteness of D_i is to ensure that the solution α_i of QP_i is a descent direction for the exact penalty function used to compute the step length. Now, because x_0 is fixed, a reduction on step length ε not only affects the movement along the direction α_i but also reduces the norm of Δ_i , therefore making the solution of QP_i closer to the solution of DQP_i . Hence, if the eigenvalues of the matrices $\{D_i\}$ are positive and bounded away from zero, we could see through Proposition II.2.A that for ε small enough (but bounded away from zero) we would have a reduction on the exact penalty function even if no future correction is made.

In other words, if the new control sequence $\{u_i\}$ is constructed ignoring the future change on the trajectory, that is according to:

II.2.XIV

$$u_i = \bar{u}_i + \varepsilon \alpha_i$$

then we can still prove global convergence, provided that assumption G2 holds. Note that II.2.XIV differs from the first order method below, due to Jacobson and Mayne (1970),

II.2.XV

$$u_i = \bar{u}_i - \varepsilon H_u^i$$

because $(-H_u^i)$ would be equal to α_i only if $D_i = I$. Therefore, iteration II.2.XIV can be expected to work better than II.2.XV. However, its rate of convergence is not likely to be faster than linear. Actually, a superlinear rate of convergence will depend not only on the way the matrices $\{D_i\}$ are defined but also on the forward correction of u_i due to the change of trajectory. As shown in the proof of Theorem II.2.H, the computations of step 8 and step 12 only differ by a second order term in Δ_i . But

$$\Delta_i \rightarrow 0 \text{ as } \{u_i\} \rightarrow \{u^*\}$$

Therefore, although the quadratic convergence result of Theorem II.2 was proved assuming that the forward correction is ultimately performed by step 8, the computation of step 12 may also lead to a superlinear rate of convergence provided that $\varepsilon = 1$. Indeed, it is the achievement of a step length of unity which motivated the introduction of the quadratic problem DQP_i , instead of making the correction always through step 12. As we have seen in Part I, the degree of accuracy with which the constraints are satisfied may play a major role in the establishment of superlinear rate of convergence for algorithms which approximate the constraints around the current point.

While the solution of the quadratic problem DQP_i satisfies the constraints to first order, this may not be so with the computation of step 12.

Furthermore, as in Part I, a second order term $\tilde{p}_i$ may be added to the solution of DQP_i and the resultant search arc be used to compute the step length. As this arc satisfies the constraints to second order, a step length of unity is then more likely to be accepted by the test of step 10. In particular, this could avoid trouble eventually arising from the occurrence of the Maratos "effect" [see Maratos (1978)].

The structure of the method described in this chapter makes it possible to apply most of the ideas employed in Part I. For instance, with the additional variable w_i new sub-problems DQP_i^w and QP_i^w can be used replacing DQP_i and QP_i respectively, thus preventing any inconsistent constraints in the quadratic optimization problems.

A procedure, similar to the one suggested by Mayne and Polak(1982), which was used in the Basic Algorithm to update the penalty parameter c_i , could also be incorporated into the algorithm. Alternatively, when the constraints $g(\cdot)$ are convex, something often found in practice, the simpler procedure described in chapter I.2 may also be useful.

APPENDIX (Chapter II.2)

Proof of Proposition II.2.D

Let $d(\epsilon, \{u_i\}, \{\delta u_i^1\}, r) \triangleq |\gamma(\{u_i + \epsilon \delta u_i^1\}, r) - \hat{\gamma}(\{u_i\}, \{\epsilon \delta u_i^1\}, r)|$

Let

$\{x_i\}$ denote the sequence obtained with $\{u_i\}$

and

$\{x_i^1\}$ denote the sequence obtained with $\{u_i + \epsilon \delta u_i^1\}$

From the definitions of $\gamma(\cdot, r)$ and $\hat{\gamma}(\cdot, \cdot, r)$ we have:

$$(i) \quad d(\epsilon, \{u_i\}, \{\delta u_i^1\}, r) \leq |F(x_N^1) - F(x_N) - \sum_{i=0}^{N-1} (H_u^i)^T \epsilon \delta u_i^1| \\ + r |\tilde{\psi}(\{u_i + \epsilon \delta u_i^1\}) - \hat{\psi}(\{u_i\}, \{\epsilon \delta u_i^1\})|.$$

Now, the expression (for the discrete case) of the change in cost due to a change in the control sequence, corresponding to the continuous time expression in Mayne (1973), is:

$$(ii) \quad F(x_N^1) - F(x_N) = \sum_{i=0}^{N-1} [H_i(x_i, u_i + \epsilon \delta u_i^1, v_x^{i+1}) - H_i(x_i, u_i, v_x^{i+1})]$$

where

$$H_i(x, u, w) \triangleq w^T f_i(x, u)$$

Because of G1 and the mean value theorem, there exists an $\tilde{u}_i$ lying on the segment $[u_i, u_i + \epsilon \delta u_i^1]$ such that:

$$(iii) \quad H_i(x_i, u_i + \epsilon \delta u_i^1, v_x^{i+1}) - H_i(x_i, u_i, v_x^{i+1}) = [H_u^i(x_i, \tilde{u}_i, v_x^{i+1})]^T \epsilon \delta u_i^1$$

From (ii) and (iii) we obtain:

$$(iv) \quad F(x_N') - F(x_N) = \sum_{i=0}^{N-1} [H_u^i(x_i, \tilde{u}_i, v_x^{i+1})]^T \varepsilon \delta u_i'$$

where,

for all i , $\tilde{u}_i$ lies on the segment $[u_i, u_i + \varepsilon \delta u_i']$

(iv) into (i) leads to:

$$(v) \quad d(\varepsilon, \{u_i\}, \{\delta u_i'\}, r) \leq \varepsilon \sum_{i=0}^{N-1} [\| H_u^i(x_i, \tilde{u}_i, v_x^{i+1}) - H_u^i(x_i, u_i, v_x^{i+1}) \|] \| \delta u_i' \| \\ + r | \psi(\{u_i + \varepsilon \delta u_i'\}) - \hat{\psi}(\{u_i\}, \{\varepsilon \delta u_i'\}) | .$$

Now, again by G1 and the mean value theorem, for all i , all $s \in S_i$, there exists an $\hat{u}_i^s$, lying on the segment $[u_i, u_i + \varepsilon \delta u_i']$ such that:

$$(vi) \quad g_i^s(u_i + \varepsilon \delta u_i') = g_i^s(u_i) + {}_i g_u^s(\hat{u}_i^s) \varepsilon \delta u_i' .$$

Hence, from (vi) and the definition of $\psi(\cdot)$, $\hat{\psi}(\cdot, \cdot)$ we have:

$$(vii) \quad | \psi(\{u_i + \varepsilon \delta u_i'\}) - \hat{\psi}(\{u_i\}, \{\varepsilon \delta u_i'\}) | \leq \sum_{i=0}^{N-1} [\max_{s \in S_i} \{ \| {}_i g_u^s(\hat{u}_i^s) - {}_i g_u^s(u_i) \| \} \varepsilon \| \delta u_i' \|] .$$

Now, there exists a compact set U' such that, for $\varepsilon \in (0, 1]$, $u_i \in U$ and $\delta u_i' \in \beta_K(0)$, we have:

$$[u_i, u_i + \varepsilon \delta u_i'] \subset U', \quad \text{for all } i .$$

Also, the functions H_u^i , ${}_i g_u^s$ are all continuous and, therefore, uniformly continuous in U' .

Hence,

$$(viii) \quad \| H_u^i(x_i, \tilde{u}_i, v_x^{i+1}) - H_u^i(x_i, u_i, v_x^{i+1}) \| \rightarrow 0,$$

uniformly in $U \times \beta_K(0)$, as $\varepsilon \rightarrow 0$

and

$$(ix) \quad \| {}_i g_u^s(\hat{u}_i^s) - {}_i g_u^s(u_i) \| \rightarrow 0,$$

uniformly in $U \times \beta_K(0)$, as $\varepsilon \rightarrow 0$.

The proposition follows from (ix), (viii) and (v).

Proof of Proposition II.2.F

Recall that in the first chapter the map $\bar{\lambda}(\cdot)$ was defined for all $x \in \mathbb{R}^n$. If we apply these maps to estimate the multipliers of the mathematical programming problem MP2, we have:

$$(i) \quad \bar{\lambda}(\bar{u}) \triangleq \min_{\lambda \in \mathbb{R}^{\bar{\ell}}} \{ \| \tilde{F}_u(\bar{u}) + g_x^T(\bar{u})\lambda \|^2 + \sum_{s=1}^{\bar{\ell}} [(\psi'(\bar{u}) - g_i^s(\bar{u}))^2(\lambda)^2] \}$$

where, $\bar{\ell} \triangleq \sum_{i=0}^{N-1} \ell_i$ and $\psi'(\bar{u}) = \max\{0; g_i^s(\bar{u})\}$

$$i = 0, \dots, N-1$$

$$s = 1, \dots, \ell_i.$$

We already know that:

$$\bar{\lambda}(\bar{u}) \rightarrow \lambda^* \text{ as } \bar{u} \rightarrow u^* \text{ (K-T point to MP2)}$$

Now, for all i the maps $\bar{\lambda}_i(\cdot)$ were defined as:

$$(ii) \quad \bar{\lambda}_i(\bar{u}) \triangleq \arg \min_{\lambda \in \mathbb{R}^{\ell_i}} \{ \| H_u^i + \sum_{s \in S_i} \lambda^s g_u^s(\bar{u}_i) \|^2 + \sum_{s \in S_i} (\psi_i(\bar{u}_i) - g_i^s(\bar{u}_i))^2(\lambda)^2 \}$$

where,

$$\psi_i(\bar{u}_i) \triangleq \max_{s \in \ell_i} \{0; g_i^s(\bar{u}_i)\}$$

Since $\bar{u} \rightarrow u^*$ (K-T point to MP2), we have:

$$(iii) \psi_i(\bar{u}_i) \rightarrow 0, \text{ for all } i$$

and

$$(iv) \psi'(\bar{u}) \rightarrow 0$$

From the recursive equations and the definition of $\tilde{F}(\cdot)$ it is possible to show (see Theorem II.2.H) that:

$$\begin{aligned} (vi) \quad \tilde{F}_{u_i} &= H_u^i - (f_u^i)^T (f_x^{i+1})^T \dots (f_x^{N-2})^T \beta_{N-1}^T [H_u^{N-1} + \lambda_{N-1}^T g_u] \\ &\quad - (f_u^i)^T (f_x^{i+1})^T \dots (f_x^{N-3})^T \beta_{N-2}^T [H_u^{N-2} + \lambda_{N-2}^T g_u] \\ &\quad \vdots \\ &\quad - (f_u^i)^T \beta_{i+1}^T [H_u^i + \lambda_{i+1}^T g_u] \end{aligned}$$

Now, at the solution of QP_i we have:

$$(vii) D_i \alpha_i + [H_u^i + \lambda_i^T g_u] = 0$$

As shown in the proof of Theorem II.2.E,

$$(viii) \{\alpha_i\} \rightarrow \{\alpha_i^*\} \text{ as } \bar{u} \rightarrow u^*$$

where,

α_i^* solves the problem below:

$$\min_{\alpha \in \mathbb{R}^m} [H_u^i(u^*)]^T \alpha + \frac{1}{2} \alpha^T D_i(u^*) \alpha$$

subject to:

$$g_i^s(u_i^*) + g_u^s(u_i^*) \alpha \leq 0, \quad s \in S_i$$

where,

$D_i(u^*)$ means that the sequence $\{u_i^*\}$
is used to compute D_i .

Clearly, from G5, $\alpha_i^* = 0$ for all i .

Hence, from (viii) and (vii) we have:

$$(ix) \quad [H_u^i + \lambda_i^T g_u] \rightarrow 0, \text{ for all } i.$$

Then, from (ix) and (vi) we have:

$$(x) \quad (\tilde{F}_{u_i} - H_u^i) \rightarrow 0, \text{ for all } i.$$

From (x), (iv), (iii), (ii) and (i) we then have:

$$(\bar{\lambda}(\bar{u}) - \{\bar{\lambda}_i(\bar{u}_i)\}) \rightarrow 0$$

Then,

$$\{\bar{\lambda}_i(\bar{u}_i)\} \rightarrow \lambda^*.$$

From the above convergence, G1 and the definition of $P(\cdot)$, we have:

$$P(u^w) \rightarrow \tilde{F}_{uu}(u^*) + \sum_{i,s}^s \lambda_i g_{uu}^s(u_i^*)$$

Then, the proposition follows from Theorem 3.1 in Robinson(1974).

Proof of Proposition II.2.G

By G4 we have: $\{u_i\} \rightarrow u^*$ (K-T point to MP2).

Then, by G7,

$$\{\delta u_i\} \rightarrow 0$$

Also, as $u^w \rightarrow u^*$, $\delta u^w \triangleq (u^w - \bar{u}) \rightarrow 0$

Hence, for $\bar{u}$ near enough to u^* , none of the non-active inequalities at u^* are predicted to be active, either by QP2 or the sequence $\{DQP_i\}$.

Now, let λ_i^D denote the multiplier vector of DQP_i and λ^W the multiplier of QP2. Because of G1 and the boundedness of $P(\bar{u})$, it is possible to show [as in the proof of Proposition I.4.F] that:

$$\lambda^W \rightarrow \lambda^* \text{ (multiplier to MP2).}$$

Then, the first part of the proposition (concerning QP2) comes directly from the strict complementary slackness assumed in G5.

Now, let N_i be the matrix whose columns are the gradients of the active $g_i(\cdot)$ at u_i^* , calculated at $\bar{u}_i$. From the fact that none of the non-active $g_i(\cdot)$ at u_i^* are predicted to be active by DQP_i , at the solution of this problem we have:

$$H_u^i + B_i \Delta_i + D_i \delta u_i = N_i Z_i$$

with

$$Z_i \triangleq \{s \lambda_i^D | g_i^s(u_i^*) = 0\}$$

Then, as $\delta u_i \rightarrow 0$ and B_i and D_i are bounded, we have:

$$(i) \lim_{i \rightarrow \infty} N^i Z_i = N^i Z$$

where,

N^i is the matrix whose columns are the gradients of the active $g_i(\cdot)$ at u_i^* , calculated at u_i^*

and

$$Z = \{\hat{\lambda}_i^s \mid g_i^s(u^*) = 0\}$$

with

$\hat{\lambda}_i$ being the multiplier associated with the solution of DQP_i , for $\bar{u} = u^*$.

From (i) and the independence of the columns of N^i assured by G6, we have:

$$(ii) \quad \lambda_i^D \rightarrow \hat{\lambda}_i.$$

But, for $\bar{u} = u^*$ we have seen in the proof of the last proposition that $\alpha_i = 0$, for all i . That is

$$DQP_i \equiv QP_i.$$

Then, (see Proposition II.2.B), the K-T conditions to MP2 are satisfied at $\bar{u} = u^*$ with:

$$(iii) \quad \lambda^* = \{\hat{\lambda}_i\}.$$

But, by G6, the multiplier λ^* is unique, so that, from (ii) and (iii), we have:

$$\{\lambda_i^D\} \rightarrow \lambda^*$$

And again by the strict complementary slackness assumed in G5, the last part of the proposition is proved.

II.3 SUGGESTIONS FOR FURTHER RESEARCH

The modified version of the DDP method, presented in II.1, is shown (see Theorem II.1A) to give the same successor as Newton's method, without requiring the cumbersome task of inverting a $(mN) \times (mN)$ matrix at each iteration. In practice, when applying Newton's method, it is a common feature to replace the inversion of the matrix by the solution of a system of linear equations. The Gauss elimination is among the most commonly used methods to solve linear systems. It is a point for further research to compare the computation involved in the method of II.1 with that involved in the triangularization procedure of Gauss's method.

The requirement of positive definiteness of the matrices $\{D_i\}$ to establish convergence of the algorithm of II.2 reduces the applicability of this algorithm, in particular if the matrices D_i are computed as suggested. That is:

$$D_i = C_i + \sum_{s=1}^{\ell_i} \bar{\lambda}_i^s (\bar{u}_i) g_{uu}^s(\bar{u}_i).$$

Of course if, at all iterations, D_i is set equal to a positive definite matrix (say $\bar{\alpha}I$, $\bar{\alpha} > 0$) for all i , assumption G2 would be satisfied and any accumulation point of the sequence of control sequence $\{u_i\}$ constructed by the algorithm satisfies the first-order optimality conditions for the discrete control problem (see Theorem II.2.E). But the result of Theorem II.2.H would no longer be valid as only a linear rate of convergence can be expected.

A method which keeps the sequence $\{D_i\}$ positive definite while updating $\{D_i\}^{\ell+1}$ through an extension of the update given by

I.4.VI is given below.

Let $T_i \triangleq \min\{m'; [\psi_i(\bar{u}_i) + \|H_u^i + \sum_{s=1}^{\ell_i} \lambda_i^s g_{uu}^s(\bar{u}_i)\|]\}$ and for each iteration ℓ define: $T^\ell \triangleq \max\{T_i | i = 0, \dots, N-1\}$.

For $i = 0, \dots, N-1$ let N_i be the matrix whose columns are the gradients of the active constraints at the solution of QP_i .

The sequence $\{D_i\}^\ell$ can then be updated to $\{D_i\}^{\ell+1}$ as follows:

$$D_i^{\ell+1} = D_i^\ell - \frac{D_i^\ell \delta_i^\ell (\delta_i^\ell)^T D_i^\ell}{(\delta_i^\ell)^T D_i^\ell \delta_i^\ell} + \frac{\eta_i^\ell (\eta_i^\ell)^T}{(\delta_i^\ell)^T \eta_i^\ell}$$

for $i = 0, \dots, N-1$, where:

$$\delta_i^\ell \triangleq \bar{u}_i^{\ell+1} - \bar{u}_i^\ell$$

$$\eta_i^\ell \triangleq \gamma_i^\ell + a_i^\ell [T^\ell + N_i^\ell (N_i^\ell)^T] \delta_i^\ell$$

with

$$\gamma_i \triangleq [H_u^i + \sum_{s=1}^{\ell_i} \lambda_i^s g_{uu}^s(\bar{u}_i)]^{\ell+1} - [H_u^i + \sum_{s=1}^{\ell_i} \lambda_i^s g_{uu}^s(\bar{u}_i)]^\ell$$

and

$$a_i^\ell \triangleq \begin{cases} 0, & \text{if } (\delta_i^T \gamma_i)^\ell \geq \varepsilon_1 (\|\delta_i\|^2)^\ell, \varepsilon_1 \in (0, 1) \\ 1, & \text{if } 0 \leq (\delta_i^T \gamma_i)^\ell \leq [\min\{\varepsilon_1; T^\ell\}] (\|\delta_i\|^2)^\ell \\ 1 + \frac{T^\ell (\|\delta_i\|^2)^\ell - (\delta_i^T \gamma_i)^\ell}{T^\ell (\|\delta_i\|^2)^\ell + (\delta_i^T N_i N_i^T \delta_i)^\ell}, & \text{otherwise} \end{cases}$$

Proposition II.2.B, the definition of T^ℓ and assumption G5 are sufficient to show that the above update is well defined and generates a sequence $\{D_i\}^\ell$ of positive definite matrices for all i . It is then a point for further research to know whether a super-linear rate of convergence can be established for the algorithm

of II.2 when using the above update procedure. Note that, at each stage i , the constraints g_i^j , $j = 1, \dots, \ell_i$, only depend on u_i . Hence, the matrices D_i are not obtained recursively but are constructed from C_i using estimates of the multipliers λ_i . If the constraints are of the form $g_i^s(u_0, u_1, \dots, u_{N-1})$ the recursive equations must be modified in a similar way to that suggested by Jacobson and Mayne (1970) for the continuous time case. Further research is required on this problem as well as on the question of how to cope with state constraints.

Finally, application of the algorithm of II.2 to practical problems is required before asserting its competitiveness.

REFERENCES

- Bandler, J.W., Liu, P.C. and Tromp, H. (1976) - 'A Nonlinear Programming Approach to Optimal Design Centering, Tolerancing and Tuning' - IEEE Transactions on Circuits and Systems, No. 3, March, pp 155-165.
- Beale, E.M.L. (1967) - 'Numerical Methods' (in 'Nonlinear Programming', Ed. J. Abadie), North Holland, Amsterdam, pp 189-196.
- Bellman, R. (1957) - 'Dynamic Programming', Princeton University Press, New Jersey.
- Bellman, R. and Dreyfus, S.E. (1962) - 'Applied Dynamic Programming', Princeton University Press, New Jersey.
- Berge, C. (1962) - 'Topological Spaces', The MacMillan Company, New York
- Cannon, M.D., Cullum, C.D. and Polak, E. (1970) - 'Theory of Optimal Control and Mathematical Programming', McGraw Hill, USA.
- Chamberlain, R.M (1979) - 'Some Examples of Cycling in Variable Metric Methods for Constrained Optimization', Mathematical Programming, 16, pp 378-383.
- Chamberlain, R.M. (1980) - 'The Theory and Application of Variable Metric Methods to Constrained Optimization Problems', PhD Dissertation, University of Cambridge.
- Chamberlain, R.M., Lemaréchal, C., Pedersen, H.C. and Powell, M.J.D (1982) - 'The Watchdog Technique for Forcing Convergence in Algorithms for Constrained Optimization' - Math. Program. Study 16, pp 1-17.
- Conn, A.R. (1973) - 'Constrained Optimization Using a Non-differentiable Penalty Function' - SIAM J. of Numerical Analysis, pp 760-784.
- Conn, A.R. and Pietrzykowski, T. - 'A Penalty Function Method Converging Directly to a Constrained Optimum' - SIAM J. of Numerical Analysis, pp 348-378.
- Dennis, J.E. and Moré, J.J. (1974) - 'A Characterization of Superlinear Convergence and its Application to Quasi-Newton Methods', Mathematics of Computation, pp 549-560.
- Dennis, J.E. and Moré, J.J. (1977) - 'Quasi-Newton Methods, Motivation and Theory', SIAM Rev. 19, pp 46-89.
- Dixon, L.C.W. (Ed.) (1976) - 'Optimization in Action', Academic Press, London

Fiacco, A.V. and McCormick, G.P. (1968) - 'Nonlinear Programming: Sequential Unconstrained Minimization Techniques' - John Wiley and Sons, USA.

Fletcher, R. and Powell, M.J.D. (1963) - 'A Rapidly Convergent Descent Method for Minimization' - Computer J., 6, pp 163-168.

Fletcher, R. (1971) - 'A General Quadratic Programming Algorithm', J. Inst. Maths. Applics. 7, pp 76-91.

Fletcher, R. (1980) - 'Practical Methods of Optimization' Vol. 1, John Wiley and Sons, USA.

Fletcher, R. (1981a) - 'Practical Methods of Optimization' Vol 2, John Wiley and Sons, USA.

Fletcher, R. (1981b) - 'Numerical Analysis with an Exact L_1 Penalty Function Method' (in 'Nonlinear Programming 4', Eds. O.L. Mangasarian, R.R. Meger and S.M. Robinson), Academic Press, New York, pp 99-129.

Garcia-Palomares, U.M. and Mangasarian, O.L. (1976) - 'Superlinearly Convergent Quasi-Newton Algorithms for Nonlinearly Constrained Optimization Problems', Math. Prog. 11, pp 1-13.

Glad, T. and Polak, E. (1979) - 'A Multiplier Method with Automatic Limitation of Penalty Growth' - Math. Prog. 17, pp 140-144.

Goldfeld, S.M., Quandt, R.E. and Trotter, H.F. (1966) - 'Maximization by Quadratic Hill-Climbing', Econometrica, 34, pp 541-551.

Goldstein, A.A. and Price, J.F. (1967) - 'An Effective Algorithm for Minimization' - Numer. Math. 10, pp 184-189.

Gonzaga, C. and Polak, E. (1979) - 'On Constraint Dropping Schemes and Optimality Functions for a Class of Outer Approximations Algorithms', SIAM J. Control and Optimization, Vol 17, pp 477-493.

Halkin, H. (1966), SIAM Rev. 8, October, pp 547.

Han, S.D. and Mangasarian, O.L. (1974) - 'Exact Penalty Functions in Nonlinear Programming', Math. Prog., 17, pp 251-269.

Han, S.P. (1976) - 'Superlinearly Convergent Variable Metric Algorithms for General Nonlinear Mathematical Programming Problems' - Math. Prog., 11, pp 263-282.

Han, S.P. (1977) - 'A Globally Convergent Method for Nonlinear Programming' - JOTA, Vol 22, No. 3, pp 297-309.

Jacobson, D.H. and Mayne, D.Q. (1970) - 'Differential Dynamic Programming', Elsevier, New York.

- Levenberger, K. (1944) - 'A Method for the Solution of Certain Nonlinear Problems in Least Squares', Quant. Appl. Math., 2, pp 164-168.
- Levitin, E.S. and Polyak, B.T. (1966) - 'Constrained Minimization Methods', USSR Computational Mathematics and Mathematical Physics, pp 1-49.
- Luenberger, D.G. (1969) - 'Optimization by Vector Space Methods', John Wiley and Sons, New York.
- Maratos, N. (1978) - 'Exact Penalty Function Algorithms for Finite Dimensional and Control Optimization Problems', PhD Thesis, Imperial College.
- Marquardt, D.W. (1963) - 'An Algorithm for Least Squares Estimation of Nonlinear Parameters', SIAM J., 11, pp 431-441.
- Mayne, D.Q. (1966), International J. Control, 3, pp 85.
- Mayne, D.Q. (1973) - 'Differential Dynamic Programming' (in 'Control and Dynamic Systems', Ed. C.T. Leondes), Academic Press, New York, pp 180-252.
- Mayne, D.Q. and Polak, E. (1980) - 'An Exact Penalty Function Algorithm for Optimal Control Problems with Control and Terminal Equality Constraints' - JOTA, 2, pp 211-246.
- Mayne, D.Q. and Polak, E. (1978) - 'A Superlinearly Convergent Algorithm for Constrained Optimization Problems', Research Report, CCD/Imperial College.
- Mayne, D.Q., Polak, E. and Trahan, H. (1979a) - 'An Outer Approximations Algorithm for Computer Aided Design Problems', JOTA, Vol 28, pp 331-352.
- Mayne, D.Q. and Maratos, N. (1979b) - 'A First Order Exact Penalty Function Algorithm for Equality Constrained Optimization Problems', Math. Prog., 16, pp 303-324.
- Mayne, D.Q. and Polak, E. (1982) - 'A Superlinearly Convergent Algorithm for Constrained Optimization Problems', Math. Prog. Study 16, pp 45-61.
- McCormick, G.P. (1967) - 'Second Order Conditions for Constrained Minima' - SIAM J. Appl. Maths., 15, pp 641-652.
- Ortega, J.M. and Rheinboldt, E.C. (1970) - 'Iterative Solution of Nonlinear Equations in Several Variables', Academic Press, New York.
- Pietrzykowski, T (1969) - 'An Exact Potential Method for Constrained Maxima' - SIAM J. Numerical Analysis 6, pp 299-304.

- Polak, E. (1974) - 'A Globally Converging Secant Method with Applications to Boundary Value Problems', SIAM J. Numerical Analysis, 11, pp 529-537.
- Powell, M.J.D. (1969) - 'A Method for Nonlinear Constraints in Minimization Problems' (in 'Optimization', Ed. R. Fletcher), Academic Press, London, pp 283-298.
- Powell, M.J.D. (1971) - 'On the Convergence of the Variable Metric Algorithm', J. Inst. Maths. Appls., 7, pp 21-36.
- Powell, M.J.D. (1977) - 'Variable Metric Methods for Constrained Optimization', 3rd International Symposium on Computing Methods in Applied Sciences and Engineering, Paris.
- Powell, M.J.D. (1978a) - 'The Convergence of the Variable Metric Methods for Nonlinearly Constrained Optimization Problems', in Nonlinear Programming 3, Ed. O.L. Mangasarian, R.R. Meyer, S.M. Robinson, pp 27-61.
- Powell, M.J.D. (1978b) - 'A Fast Algorithm for Nonlinearly Constrained Optimization Problems', Lecture Notes in Mathematics 630 (Ed. G.A. Watson), Springer-Verlag, Berlin, pp 144-157.
- Powell, M.J.D. (1982a) - 'State-of-the-Art Tutorial on Variable Metric Methods for Constrained Optimization', D.A.M.T.P., University of Cambridge.
- Powell, M.J.D. (1982b) - 'V.M.C.W.D.: A FORTRAN Subroutine for Constrained Optimization', D.A.M.T.P., University of Cambridge.
- Ren-Pu, G. and Powell, M.J.D. (1983) - 'The Convergence of Variable Metric Matrices in Unconstrained Optimization', Math. Prog., 27, pp 123-143.
- Robinson, S.M. (1974) - 'Perturbed Kuhn-Tucker Points and Rates of Convergence for a Class of Nonlinear Programming Algorithms', Math. Prog., 7, pp 1-16.
- Stoer, J. and Witzgall, C. (1970) - 'Convexity and Optimization in Finite Dimension', Springer-Verlag, Berlin.
- Wilson, R.B. (1963) - 'A Simplicial Algorithm for Concave Programming' - PhD Dissertation, Harvard University.
- Wolfe, P. (1959) - 'The Secant Method for Simultaneous Nonlinear Equations', Comm. A.C.M., 2, pp 12-14.
- Yamashita, H. (1982) - 'A Globally Convergent Constrained Quasi-Newton Method with an Augmented Lagrangian-type Penalty Function', Math. Prog., 23, pp 75-86.
- Zoutendijk, G. (1969) - 'Methods of Feasible Directions', Elsevier, Amsterdam.