TRANSFORMATION METHODS

IN

NONLINEAR PROGRAMMING

BY

D.M. RYAN

A thesis submitted to the
Australian National University
for the degree of Doctor of Philosophy.

August, 1971.

TABLE OF CONTENTS

Acknowledgement	(i)
Preface	(ii)
Abstract	(iii)

Chapter 1 : Introduction

1.1 Problem Description	1
1.2 Transformation Methods for Constrained Problems	7
1.3 Scope of the Thesis	11
1.4 Definitions and Classical Results	14

Chapter 2 : Barrier Function Methods

2.1 Introduction	26
2.2 Trajectory Analysis and Error Estimates	31
2.3 A New Family of Barrier Functions	40
2.4 Numerical Results	46
2.5 Hessian Conditioning and Computational Problems	57

Chapter 3 : Penalty Function Methods

3.1 Introduction	62
3.2 The Schmit-Fox Method and Morrison's Method	63
3.3 The Tangent Parameter	71
3.4 An Algorithm	76
3.5 Numerical Results	80
3.6 Powell's Method	86
3.7 Properties of the Hessian of $T(\tilde{x}, \tilde{\theta}, \tilde{\sigma})$	92

Chapter 4 : A Hybrid Method

4.1	Introduction	95
4.2	Equality Constrained Barrier Function Transformations	96
4.3	Inequality Constraint Classification	104
4.4	Active Constraint Transfers	107
4.5	Hessian Corrections	111
4.6	The Hybrid Algorithm	116
4.7	Numerical Results Using the Hybrid Algorithm	118

Chapter 5 : Unconstrained and Linearly Constrained Methods

5.1	Introduction	128
5.2	Unconstrained Methods	129
5.3	Projection Methods for Linearly Constrained Problems	133
5.4	A Hybrid Projection Algorithm for Minimizing $T(\underline{x}, \underline{\theta}, \underline{\sigma}, r)$	139
5.5	Numerical Results	146

Bibliography	151
--------------	-----

ACKNOWLEDGEMENTS

I am particularly indebted to my wife, Ruth, for her support, encouragement and longsuffering - attributes that have received a great deal of exercise especially during the preparation of this thesis.

I wish to thank Dr. M.R. Osborne for his encouraging and stimulating supervision during my three years of study at the Computer Centre, Australian National University. I am particularly grateful for his interest and the frequent discussions that have led to much of the work presented in this thesis. I would also like to thank Dr. R.S. Anderssen, first for his supervision during the past nine months and secondly, for his constructive criticism of the early version of the thesis.

I also wish to thank Mrs. Stephanie Larkham, whose patience and undoubted skill are clearly evident in her preparation of the typescript. I am also grateful to Mr. L.S. Jennings for his careful reading and checking of the typescript.

Finally, I would like to thank the Australian National University for the financial support of a Postgraduate Scholarship.

PREFACE

Many of the results described in this thesis have been established during or as the result of collaboration and discussion with M.R. Osborne and have subsequently been published in the following joint papers: Kowalik, Osborne and Ryan (1969) and Osborne and Ryan (1970a, 1970b). When discussing these results in this thesis the text of the relevant papers has been closely followed.

In presenting the basic theory of barrier function methods in Chapter 2, repeated reference is made to Fiacco and McCormick (1968) which can be taken as a standard reference for transformation methods. The reader is referred in Chapter 5 to the original source references for the theoretical properties of the unconstrained and linearly constrained methods discussed. Emphasis in this chapter is on the computational properties of the methods with particular reference to their use as a computational tool for transformation methods.

Elsewhere in this thesis, unless another source is acknowledged, the work described is my own and the numerical results presented in support of my argument I believe to be original.

A handwritten signature in black ink, appearing to read 'L. J. Ryan', is located at the bottom right of the page.

ABSTRACT

This thesis investigates a number of difficulties associated with the design and implementation of transformation methods for the solution of the continuous mathematical programming problem,

$$\begin{aligned} &\text{minimize} && f(\underline{x}) , && \underline{x} \in E^n, \\ &\text{subject to} && g_i(\underline{x}) \geq 0 , && i=\overline{1, \ell}, \\ &\text{and} && h_i(\underline{x}) = 0 , && i=\overline{1, m}, \end{aligned}$$

where the problem functions $f(\underline{x})$, $g_i(\underline{x})$, $i=\overline{1, \ell}$, $h_i(\underline{x})$, $i=\overline{1, m}$, are at least continuous. The aim is to improve convergence properties of current transformation methods by designing modified methods which avoid the computational difficulties.

In Chapter 2, barrier function methods (interior-point methods) are investigated and a new family of barrier functions, exhibiting improved convergence properties, is proposed. The theoretical and computational aspects of two penalty function methods (exterior-point methods) are then considered in Chapter 3. Based on the properties of the barrier and penalty function approaches discussed in the second and third chapters, a new hybrid method is

constructed in Chapter 4. The theoretical validity of the new approach is established and it is shown that under mild conditions on the problem (e.g. the Kuhn-Tucker conditions) the computational difficulties of the parent methods are avoided and improved convergence properties are evident. For problems not satisfying such conditions, the hybrid method behaves as a barrier function method. In Chapter 5, unconstrained methods and projection methods for linearly constrained problems are considered and numerical evidence is presented to suggest that if linear constraints are active at the solution then improved convergence can be obtained by treating such constraints using projection instead of including them in barrier or penalty function methods.

CHAPTER 1

INTRODUCTION

1.1 Problem Description.

The general optimization problem to be considered in this thesis is the mathematical programming problem:

$$\text{Problem 1.1 : minimize } f(\tilde{x}) \quad (1.1)$$

$$\text{subject to } g_i(\tilde{x}) \geq 0, \quad i=\overline{1,\ell}, \quad (1.2)$$

$$\text{and } h_i(\tilde{x}) = 0, \quad i=\overline{1,m}, \quad (1.3)$$

where $f(\tilde{x})$, the objective function, and $g_i(\tilde{x})$, $i=\overline{1,\ell}$, and $h_i(\tilde{x})$, $i=\overline{1,m}$, the constraint functions, are real-valued functions of a vector $\tilde{x}$ in the n -dimensional vector space E^n . It will be assumed throughout this thesis that the problem functions, $f(\tilde{x})$, $g_i(\tilde{x})$, $i=\overline{1,\ell}$, and $h_i(\tilde{x})$, $i=\overline{1,m}$, are at least continuous and further restrictions, in the form of differentiability conditions, will be imposed when necessary. It should be noted that the convention of adopting the minimization form of optimization results in no loss of generality since it is equivalent to maximizing $[-f(\tilde{x})]$. Much of the interest in Problem 1.1 has developed since the basic results of Dantzig (1951) and Kuhn and Tucker (1951) were published and has been stimulated by the

advent of modern computers and advances in many fields of application including operations research, economics, applied mathematics and industrial design and control. In the development of theory and solution methods for Problem 1.1, some classes of the general problem have proved more tractable than others. A number of these classes will now be described and the development of theory and solution methods, applicable to each form, will be briefly surveyed.

The linear programming problem, where all problem functions are linear, has been widely treated since Dantzig (1951) developed the Simplex Method and an extensive literature now exists. A report of the basic developments can be found in Dantzig (1963). Much of the recent interest has centred on adaptations of the basic methods to linear programming problems with special structures.

The linearly constrained problem, Problem 1.2, with all constraint functions linear in $\tilde{x}$ has been studied in two forms:

(i) Quadratic Objective Functions.

Contributions to quadratic programming have been made by Beale (1959), Wolfe (1959), Dorn (1960), Lemke (1962), Murray (1969a) and Fletcher (1970).

(ii) Convex and General Nonlinear Objective Functions.

Results and solution methods relating to convex and general nonlinear forms have been developed by Beale (1955), Rosen (1960), Goldfarb and Lapidus (1968), Fletcher (1968), Goldfarb (1969b) and Murtagh and Sargent (1969).

In the last two decades, the convex programming problem with $f(\underline{x})$ convex, $g_i(\underline{x})$, $i=\overline{1, \ell}$, concave and $h_i(\underline{x})$, $i=\overline{1, m}$, linear has received considerable attention. Using the convexity/concavity and the implied "smoothness" of the problem functions, powerful existence and uniqueness results concerning problem solutions have been derived. The original contributions are due to John (1948), Kuhn and Tucker (1951), Arrow and Hurwicz (1956) and Arrow, Hurwicz and Uzawa (1958), but many more recent results and methods have appeared, among them Charnes and Lemke (1954), Frisch (1955), Cheney and Goldstein (1959), Kelley (1960), Zoutendijk (1960), Wolfe (1961), Pietrzykowski (1962), Fiacco and McCormick (1963, 1964a, 1967, 1968), Hartley and Hocking (1963), Hadley (1964), Geoffrion (1966, 1967), Fiacco (1967), Mangasarian and Fromowitz (1967) and Zangwill (1969).

Results relating directly to the general nonlinear programming problems are not as abundant since, typically, both the theory and solution methods

require only local knowledge about the problem. However, convexity, and consequently the results and methods of convex programming, can often be applied in a local sense. Three important forms of the nonlinear problem require independent analysis.

Problem 1.3 $\equiv$ Problem 1.1 with $\ell=0$, $m \neq 0$.

Problem 1.4 $\equiv$ Problem 1.1 with $\ell \neq 0$, $m=0$.

Problem 1.5 $\equiv$ Problem 1.1 with $\ell=0$, $m=0$.

In this thesis, the central investigation will be a detailed examination of transformation methods for Problem 1.3 and Problem 1.4. A brief survey of some typical transformation methods is given in §1.2. In §1.3, the structure and motivation of the thesis are outlined.

Before investigating transformation methods, we shall consider results and methods related to Problem 1.5. The unconstrained problem arises in many applications (Kowalik and Osborne (1968)), and in particular, plays a vital role in transformation methods for constrained problems. Classes of methods for Problem 1.5 are now surveyed briefly. They fall naturally into two classes.

(i) Direct Search Methods (Kowalik and Osborne (1968)).

Direct search methods do not involve the explicit calculation of the objective function gradient. Experience indicates that they only

compete favourably with more sophisticated methods when the dimension n is small.

(ii) Gradient Methods (Kowalik and Osborne (1968)).

These methods require first order partial derivatives of the objective function. The simplest of them, the method of steepest descent (Curry (1944)), has a number of important disadvantages (Forsythe and Motzkin (1951), Akaike (1959), Forsythe (1967), Kowalik and Osborne (1968)) mainly reflected in the poor rate of convergence as the minimum is approached. In recent years, an extensive literature has appeared on more effective gradient methods which can be broadly classified into two groups. The older methods developed and discussed by Davidon (1959), Zoutendijk (1960), Fletcher and Powell (1963), Fletcher and Reeves (1964), Shah, Buehler and Kempthorne (1964), Daniel (1967a, 1967b), Bard (1968), Fiacco and McCormick (1968), Meyers (1968), Goldfarb (1969a, 1970), McCormick (1969), McCormick and Pearson (1969), Pearson (1969), Polak and Ribière (1969), Powell (1969b), Broyden (1970a) and Greenstadt (1970) incorporate, at least computationally and often theoretically, a full linear search (i.e. one dimensional minimization) in each iteration. The methods are generally based on a quasi-Newton (or variable metric) approach and sometimes possess the additional feature of conjugate search directions when applied to quadratic functions. The second and more recent group of methods, based purely on a quasi-Newton approach, have been proposed by Broydon (1967, 1970b), Davidon

(1968), Murtagh and Sargent (1969,1970), Powell (1969a,1970) and Fletcher (1969c). The distinguishing feature of these methods is the elimination of the linear search requirement.

Recently, Huang (1969) has developed a unified treatment of gradient methods in which all presently known methods appear as particular cases. Surveys, comparisons and reviews of direct search and gradient methods can be found in Spang (1962), Fletcher (1965, 1969b), Box (1966), Rosen (1966), Greenstadt (1967), Kowalik and Osborne (1968), Powell (1968), Box, Davies and Swann (1969), McCormick and Pearson (1969), Huang and Levy (1969) and Bard (1970).

Crockett and Chernoff (1955), Goldstein (1962) and Fiacco and McCormick (1968) have investigated the use of Newton's method for function minimization. In particular, Fiacco and McCormick have considered the special case of minimizing transformation functions arising from constrained minimization. Although convergence properties are exceptionally good for general problems, the method suffers from the disadvantage of requiring explicit calculation and inversion of the second partial derivative matrix during each iteration. It will be shown in Chapter 2 that, due to possible ill-conditioning, such inversions can lead to unreliable solutions and a possible loss in efficiency.

1.2 Transformation Methods for Constrained Problems.

Since the basic aim of this thesis is to investigate certain properties of transformation methods for nonlinear mathematical programming, we shall now discuss the background and an intuitive basis for these methods. The methods are often called penalty function methods, but we prefer to follow Murray (1969c) and reserve this term for a particular class of transformations. The transformation approach reduces the computational process of constrained minimization to that of a sequence of unconstrained minimizations of a transformation function involving the original problem functions and controlling parameters. Much of the historical and theoretical development of transformation methods is described by Fiacco and McCormick (1968), while practical aspects of problem solving by transformation methods can be found in Bracken and McCormick (1968).

Two basic classes of transformations exist.

(i) Penalty Function Methods.

Penalty function methods, sometimes called "exterior-point methods", are designed to impose an increasing penalty on the transformation function as any constraint is increasingly violated. The solution of Problem 1.1 is, in general, approached from the exterior of the feasible region, R , defined by

$$R = \{\underline{x} : g_i(\underline{x}) \geq 0, i=\overline{1, \ell}; h_i(\underline{x})=0, i=\overline{1, m}\}. \quad (1.4)$$

Penalty function methods can be designed to handle both equality and inequality constraints. The first use of these functions was made by Courant (1943) in the context of analysing constrained motion problems. Much later this approach was generalized to multiple inequality constraints and further developed by Ablow and Brigham (1955), Camp (1955), Butler and Martin (1962), Pietrzykowski (1962) and Fiacco and McCormick (1967). This method is now known as the quadratic penalty function method. A more general development of penalty function methods, based on a transformation

$$T(\underline{x}, r) = f(\underline{x}) + r \Phi(g(\underline{x}), h(\underline{x})) \quad (1.5)$$

where Φ is zero if $g_i(\underline{x}) \geq 0, i=\overline{1, \ell}$, and $h_i(\underline{x})=0, i=\overline{1, m}$, and positive otherwise, has been given by Zangwill (1967b). The function, $T(\underline{x}, r)$, is minimized for a positive increasing sequence of r -values.

Two independent contributions by Schmit and Fox (1965) and Morrison (1968) have investigated the properties of the penalty function transformation

$$T(\underline{x}, X) = (f(\underline{x}) - X)^2 + \sum_{i=1}^m h_i(\underline{x})^2 \quad (1.6)$$

for the equality constrained Problem 1.3. In Chapter 3, this transformation will be further developed by deriving improved convergence results and a new computational algorithm.

Another penalty function method for Problem 1.3 has been developed by Powell (1967) and is based on the transformation

$$T(\tilde{x}, \tilde{\theta}, \tilde{\sigma}) = f(\tilde{x}) + \sum_{i=1}^m \sigma_i [\theta_i + h_i(\tilde{x})]^2 \quad (1.7)$$

where the components of $\tilde{\theta}$ and $\tilde{\sigma} > 0$ are controlling parameters. This method is also examined in some detail in Chapter 3. Further contributions involving penalty function transformations have been made by Edelbaum (1962), Kelley (1962), Beltrami and McGill (1966), Bellmore, Greenburg and Jarvis (1970) and Haarhoff and Buys (1970).

(ii) Barrier Function Methods.

The barrier function methods, in contrast to the penalty function methods, are designed so that a "barrier" is constructed at the boundary of the feasible region given by (1.4) with $m=0$. The solution of Problem 1.4 is approached from R_0 , the interior of R , which is assumed to be non-empty. It is clear then that this form of transformation is not applicable to equality constraints. Barrier function methods are also referred to as "interior-point methods". The

earliest reference to these methods was made by Frisch (1955) with a logarithmic transformation based on

$$T(\tilde{x}, r) = f(\tilde{x}) - r \sum_{i=1}^{\ell} \log [g_i(\tilde{x})], \quad (1.8)$$

where $T(\tilde{x}, r)$ is minimized for a monotonic decreasing null sequence of r -values. This approach was developed for linear constraints by Parisot (1961) and extended to nonlinear constraints by Lootsma (1967, 1968a). A second barrier function approach, based on an inverse transformation

$$T(\tilde{x}, r) = f(\tilde{x}) + r \sum_{i=1}^{\ell} g_i(\tilde{x})^{-1}, \quad (1.9)$$

again with $T(\tilde{x}, r)$ minimized for a monotonic decreasing null sequence of r -values, was originally suggested by Carrol (1961) and later developed by Fiacco and McCormick (1963, 1964a, 1964b, 1966, 1968), Stong (1965) and Pomentale (1965). Other results concerning barrier function transformations have been discussed by Rosenbrock (1960), Huard (1964, 1967, 1968), Faure and Huard (1966), Bui Trong Lieu and Huard (1966), Kowalik (1966), Tremolières (1968), Box, Davies and Swann (1969), Fletcher and McCann (1969), Allran and Johnsen (1970) and Osborne and Ryan (1970a).

Combinations of penalty and barrier function methods have lead to the development by Fiacco (1967), Fiacco and McCormick (1968) and Lootsma (1968b, 1970) of mixed transformation methods. Such methods are considered further in Chapter 4.

Recently, Murray (1969b, 1969c), Lootsma (1969, 1970) and Fletcher and McCann (1969) have considered the numerical conditioning of certain barrier and penalty function transformations as functions of the controlling parameters. These results and their implications are discussed in Chapter 2 and provide motivation for developments later in the thesis.

An interesting and important extension of the transformation approach has been reported by Fiacco (1970) in which the continuity condition for the problem functions is relaxed and a more general description of the constraints is permitted.

1.3 Scope of the Thesis.

In this thesis, a number of difficulties associated with the design and implementation of transformation methods for the solution of continuous mathematical programming problems are investigated. The aim is to improve convergence properties of the current transformation methods by designing modified methods which avoid the computational difficulties.

The implications of these modifications are evaluated, where possible, by numerical comparison with results yielded by the original transformations. All calculations reported in this thesis were programmed in Fortran and performed in double precision arithmetic on an IBM 360/50 computer at the Computer Centre, A.N.U.

Chapter 1 is concluded with definitions and the development of standard results that will play a basic role in the subsequent chapters. The results include characterizations of the optimal solution in terms of the problem functions.

Chapter 2, based on the work of Osborne and Ryan (1970a), considers, in some detail, the barrier function methods for Problem 1.4. Convergence results are examined and on the basis of improving convergence rates, a new family of barrier functions is proposed. Finally, a number of computational difficulties associated with barrier function methods are discussed and the new barrier function family is examined numerically.

The third chapter describes an algorithm, based on the penalty function transformation (1.6), for solving Problem 1.3. Morrison's method is modified to provide an automatic starting procedure using the Schmit and Fox method and more rapid convergence is obtained by the development of a new correction term for the controlling parameter.

Kelley's device (Kelley (1962)) for transforming inequality constraints into equality constraints is used and numerical results of the method are presented. The results related to transformation (1.6) are taken from Kowalik, Osborne and Ryan (1969). The third chapter concludes with a discussion of Powell's method (Powell (1967)) for equality constraints. Attention is drawn to conditions under which favourable convergence rates apply. Results relating to Lagrange multipliers are also given.

Chapter 4 is devoted to the development of a new hybrid algorithm for nonlinear programming. The new algorithm is based on the methods and results of Chapter 2 and §3.5 and §3.6 and is motivated by

- (a) the computational problems associated with barrier function methods, and
- (b) the excellent convergence properties of Powell's method.

It will be shown that the hybrid algorithm avoids the computational difficulties of barrier function methods and utilizes Powell's method in circumstances which ensure its rapid convergence. These two features provide a highly competitive method which is illustrated by numerical comparisons with a standard barrier function method.

In Chapter 5, a number of quasi-Newton methods for unconstrained minimization and two projection algorithms for linearly constrained minimization (Problem 1.2) will be considered. A modified projection algorithm, based on these methods, will be developed to minimize linearly constrained problems on a subset of E^n . Such problems arise in Chapter 4 when using transformation methods to handle nonlinear inequality constraints.

1.4 Definitions and Classical Results.

This section will discuss the classical results of mathematical programming theory developed originally by John (1948) and Kuhn and Tucker (1951) and later extended by Arrow, Hurwicz and Uzawa (1961). The results involve necessary and sufficient conditions for the existence of local minima in terms of the original problem functions $f(\underline{x})$, $g_i(\underline{x})$, $i=\overline{1, \ell}$, and $h_i(\underline{x})$, $i=\overline{1, m}$. Concepts used in this derivation are due to Hestenes (1966), but alternative approaches are described by Fiacco and McCormick (1968) and Lootsma (1970). Attention will be restricted to optimization over Euclidean n -space assuming that all problem functions are at least continuous. Further restrictions, in the form of differentiability conditions, will be imposed when required. In establishing the necessary

and sufficient conditions, only major results will be proved but references or outlines of proofs will be provided in other cases.

Definition 1.1. The feasible region, R , for Problem 1.1 is defined by (1.4).

Any $\underline{x} \in R$ is said to be feasible. It is assumed throughout this thesis that $R \neq \emptyset$ (the empty set). The following definitions provide a characterization of minima on R .

Definition 1.2. (i) The function $f(\underline{x})$ has a local minimum, $\underline{x}^*$, on R (i.e. $\underline{x}^* \in R$) iff there exists $\delta > 0$ s.t. $f(\underline{x}) \geq f(\underline{x}^*)$ for all $\underline{x} \in N(\underline{x}^*, \delta) \cap R$ where $N(\underline{x}^*, \delta) = \{\underline{x} : ||\underline{x} - \underline{x}^*|| \leq \delta\}$.

(ii) The local minimum is isolated iff $f(\underline{x}) > f(\underline{x}^*)$ for all $\underline{x} \in N(\underline{x}^*, \delta) \cap R$, $\underline{x} \neq \underline{x}^*$.

(iii) The local minimum is global on R iff $f(\underline{x}) \geq f(\underline{x}^*)$ for all $\underline{x} \in R$.

(iv) The local minimum is unconstrained iff $f(\underline{x}) \geq f(\underline{x}^*)$ for all $\underline{x} \in N(\underline{x}^*, \delta)$ for some $\delta > 0$.

Definition 1.3. A sequence of points $\{\underline{x}^k\}$, $\underline{x}^k \in E^n$, converges to $\underline{x}^*$ in the direction $\underline{d}$ (i.e. $\underline{x}^k \xrightarrow{\underline{d}} \underline{x}^*$) where

$||\underline{d}|| = 1$ and $\underline{x}^k \neq \underline{x}^*$ for all k iff

$$\lim_{k \rightarrow \infty} ||\underline{x}^k - \underline{x}^*|| = 0 \text{ and } \lim_{k \rightarrow \infty} \frac{\underline{x}^k - \underline{x}^*}{||\underline{x}^k - \underline{x}^*||} = \underline{d}.$$

Definition 1.4. The tangent cone, $C(\underline{x}^*)$, of R at $\underline{x}^* \in R$ is

$$C(\underline{x}^*) = \{\lambda \underline{d} : \lambda \geq 0, ||\underline{d}|| = 1, \text{ there exists } \{\underline{x}^k\} \subset R \text{ s.t. } \underline{x}^k \xrightarrow{\underline{d}} \underline{x}^*\}.$$

Definition 1.5. (i) A function, $f(\underline{x})$, has a first order differential at $\underline{x}^*$ iff $f(\underline{x})$ is defined in a neighbourhood of $\underline{x}^*$ and there exists a function $f'(\underline{x}^*, \underline{d})$, linear in $\underline{d}$, which, for any sequence $\{\underline{x}^k\}$ s.t. $\underline{x}^k \xrightarrow{\underline{d}} \underline{x}^*$, satisfies

$$\lim_{k \rightarrow \infty} \frac{f(\underline{x}^k) - f(\underline{x}^*)}{||\underline{x}^k - \underline{x}^*||} = f'(\underline{x}^*, \underline{d}).$$

(ii) A function, $f(\underline{x})$, has a second order differential at $\underline{x}^*$ iff $f(\underline{x})$ has a first order differential at $\underline{x}^*$ and there exists a function $f''(\underline{x}^*, \underline{d})$, quadratic in $\underline{d}$, which satisfies

$$\lim_{k \rightarrow \infty} \frac{f(\underline{x}^k) - f(\underline{x}^*) - f'(\underline{x}^*, \underline{x}^k - \underline{x}^*)}{||\underline{x}^k - \underline{x}^*||^2} = \frac{1}{2} f''(\underline{x}^*, \underline{d}).$$

Remarks 1.1. 1. We may write $f'(\underline{x}^*, \underline{d}) = \underline{d}^T \nabla f(\underline{x}^*)$ and $f''(\underline{x}^*, \underline{d}) = \underline{d}^T \nabla^2 f(\underline{x}^*) \underline{d}$ where $\nabla f(\underline{x}^*)$ and $\nabla^2 f(\underline{x}^*)$ are called the gradient and Hessian respectively, of $f(\underline{x})$ at $\underline{x}^*$.

2. If $f(\underline{x}) \in C^{(1)}$ (i.e. once continuously differentiable) in a neighbourhood of $\underline{x}^*$, then $f(\underline{x})$ has a first order differential at $\underline{x}^*$ and the vector, $\nabla f(\underline{x}^*)$, has as components, the first partial derivatives of $f(\underline{x})$ at $\underline{x}^*$.

3. If $f(\underline{x}) \in C^{(2)}$ in a neighbourhood of $\underline{x}^*$, then $f(\underline{x})$ has a second order differential at $\underline{x}^*$ and the matrix, $\nabla^2 f(\underline{x}^*)$, has as elements, the second partial derivatives of $f(\underline{x})$ at $\underline{x}^*$.

Using Definition 1.2, Definition 1.5 and Remarks 1.1, Lemma 1.1 can be established and describes first and second order necessary conditions for local minima on R .

Lemma 1.1.

(i) If $f(\underline{x})$ has a first order differential at $\underline{x}^* \in R$ where $\underline{x}^*$ is a local minimum on R , then $f'(\underline{x}^*, \underline{d}) = \underline{d}^T \nabla f(\underline{x}^*) \geq 0$ for all $\underline{d} \in C(\underline{x}^*)$.

(ii) If $f(\underline{x})$ has a second order differential at $\underline{x}^* \in R$ where $\underline{x}^*$ is a local minimum on R and $\nabla f(\underline{x}^*) = 0$, then $f''(\underline{x}^*, \underline{d}) = \underline{d}^T \nabla^2 f(\underline{x}^*) \underline{d} \geq 0$ for all $\underline{d} \in C(\underline{x}^*)$.

Sufficient conditions for $\tilde{x}^*$ to be a local minimum on R are now stated in Lemma 1.2. Again the proof of this lemma follows from Definition 1.2, Definition 1.5 and Remarks 1.1.

Lemma 1.2.

If $f(\tilde{x})$ has a second order differential at $\tilde{x}^* \in R$, $\nabla f(\tilde{x}^*) = 0$ and $f''(\tilde{x}^*, \tilde{d}) = \tilde{d}^T \nabla^2 f(\tilde{x}^*) \tilde{d} > 0$ for all $\tilde{d} \in C(\tilde{x}^*)$, $\tilde{d} \neq 0$, then $\tilde{x}^*$ is an isolated local minimum on R (i.e. there exist $\mu, \delta > 0$ s.t. $f(\tilde{x}) \geq f(\tilde{x}^*) + \mu ||\tilde{x} - \tilde{x}^*||^2$ for all $\tilde{x} \in N(\tilde{x}^*, \delta) \cap R$).

The results of Lemma 1.1 and Lemma 1.2 are now extended by developing a description of $C(\tilde{x}^*)$ in terms of the constraint functions, $g_i(\tilde{x})$, $i=\overline{1, \ell}$, and $h_i(\tilde{x})$, $i=\overline{1, m}$. If $\tilde{x}^* \in R$ and $g_i(\tilde{x}^*) > 0$ for some i , then by the continuity of $g_i(\tilde{x})$, it follows that $g_i(\tilde{x}) > 0$ on a sufficiently small neighbourhood of $\tilde{x}^*$. This observation implies that the local character of R at $\tilde{x}^*$ and hence $C(\tilde{x}^*)$, will be defined by the active constraints at $\tilde{x}^*$.

Definition 1.6. The set of active inequality constraints at $\tilde{x}$ is $B(\tilde{x}) = \{i : g_i(\tilde{x}) = 0, i=\overline{1, \ell}\}$.

Lemma 1.3.

If $\tilde{d} \in C(\tilde{x}^*)$ and the constraint functions, $g_i(\tilde{x})$, $i \in B(\tilde{x}^*)$, and $h_i(\tilde{x})$, $i = \overline{1, m}$, have first order differentials at $\tilde{x}^*$, then $\tilde{d}$ satisfies

$$g'_i(\tilde{x}^*, \tilde{d}) = \tilde{d}^T \nabla g_i(\tilde{x}^*) \geq 0, \quad i \in B(\tilde{x}^*), \quad (1.10)$$

and

$$h'_i(\tilde{x}^*, \tilde{d}) = \tilde{d}^T \nabla h_i(\tilde{x}^*) = 0, \quad i = \overline{1, m} \quad . \quad (1.11)$$

This lemma states that all directions in the tangent cone of R at $\tilde{x}^*$ are into the feasible region. Kuhn and Tucker (1951) have constructed a simple example showing that the converse is not necessarily true.

Definition 1.7. The point, $\tilde{x}^*$, is a regular point of R iff for all $\tilde{d}$ satisfying (1.10) and (1.11), $\tilde{d} \in C(\tilde{x}^*)$.

A characterization of a regular point, in terms of the constraint functions, is given by Lemma 1.4. Extensive discussion and proofs of these conditions, known as constraint qualifications, have been given by many authors including John (1948), Kuhn and Tucker (1951), Arrow, Hurwicz and Uzawa (1961),

Hestenes (1966), Mangasarian and Fromowitz (1967), Fiacco and McCormick (1968) and Lootsma (1970).

Lemma 1.4.

Any one of the following conditions is sufficient for $\tilde{x}^*$ to be a regular point of R .

- (i) The vectors, $\nabla g_i(\tilde{x}^*)$, $i \in B(\tilde{x}^*)$, $\nabla h_i(\tilde{x}^*)$, $i=\overline{1,m}$, are linearly independent.
- (ii) There exists $\tilde{d}$ s.t. $\tilde{d}^T \nabla g_i(\tilde{x}^*) > 0$, $i \in B(\tilde{x}^*)$, $\tilde{d}^T \nabla h_i(\tilde{x}^*) = 0$, $i=\overline{1,m}$, and $\nabla h_i(\tilde{x}^*)$, $i=\overline{1,m}$, are linearly independent.
- (iii) The Kuhn-Tucker constraint qualification (Kuhn and Tucker (1951), Fiacco and McCormick (1968)) holds at $\tilde{x}^*$.

This lemma is usually proved by showing

(i) $\Rightarrow$ (iii) $\Rightarrow \tilde{x}^*$ is regular and (ii) $\Rightarrow \tilde{x}^*$ is regular.

The following result concerning linear functionals will also be useful in establishing the required necessary and sufficient conditions for constrained optimality.

Lemma 1.5. (Hestenes (1966), p.13).

Let $F(\tilde{x})$, $G_i(\tilde{x})$, $i=\overline{1,\ell}$, $H_i(\tilde{x})$, $i=\overline{1,m}$, be linear functionals on a real linear space X . If $F(\tilde{x}) \geq 0$ for all $\tilde{x} \in X$ s.t. $G_i(\tilde{x}) \geq 0$, $i=\overline{1,\ell}$, and $H_i(\tilde{x}) = 0$, $i=\overline{1,m}$, then there exist multipliers u_i , $i=\overline{1,\ell}$, and v_i , $i=\overline{1,m}$, s.t. $u_i \geq 0$, $i=\overline{1,\ell}$, and

$$F(\tilde{x}) - \sum_{i=1}^{\ell} u_i G_i(\tilde{x}) + \sum_{i=1}^m v_i H_i(\tilde{x}) = 0$$

for all $\tilde{x} \in X$. Furthermore, if $G_i(\tilde{x})$, $i=\overline{1,\ell}$, $H_i(\tilde{x})$, $i=\overline{1,m}$, are linearly independent, the multipliers are unique.

We now introduce a function which plays an important role in both the classical theory of mathematical programming and the results of this thesis.

Definition 1.8. The Lagrangian function, $L(\tilde{x}, \tilde{u}, \tilde{v})$, associated with Problem 1.1, is

$$L(\tilde{x}, \tilde{u}, \tilde{v}) = f(\tilde{x}) - \sum_{i=1}^{\ell} u_i g_i(\tilde{x}) + \sum_{i=1}^m v_i h_i(\tilde{x}) \quad (1.12)$$

where $\tilde{u} \in E^{\ell}$ and $\tilde{v} \in E^m$.

Assuming the problem functions have first order differentials, the well known Kuhn-Tucker first order necessary conditions can now be established.

Theorem 1.1.

If $\tilde{x}^*$ is a regular point of R and a local minimum of $f(\tilde{x})$ on R , then there exists $\tilde{u}^* \in E^{\ell}$ and $\tilde{v}^* \in E^m$ satisfying

$$g_i(\tilde{x}^*) \geq 0, \quad i=\overline{1, \ell}, \quad (1.13)$$

$$h_i(\tilde{x}^*) = 0, \quad i=\overline{1, m}, \quad (1.14)$$

$$u_i^* g_i(\tilde{x}^*) = 0, \quad i=\overline{1, \ell}, \quad (1.15)$$

$$u_i^* \geq 0, \quad i=\overline{1, \ell}, \quad (1.16)$$

$$\text{and } \nabla_{\tilde{x}} L(\tilde{x}^*, \tilde{u}^*, \tilde{v}^*) = 0. \quad (1.17)$$

Proof. Since $\tilde{x}^*$ is regular, $\tilde{d} \in C(\tilde{x}^*)$ for all $\tilde{d}$ satisfying (1.10) and (1.11). For $\tilde{d} \in C(\tilde{x}^*)$, using Lemma 1.1(i), $f'(\tilde{x}^*, \tilde{d}) \geq 0$ since $\tilde{x}^*$ is a local minimum on R . Using Lemma 1.5, we deduce the existence of multipliers u_i , $i \in B(\tilde{x}^*)$, v_i^* , $i=\overline{1, m}$, s.t. $u_i \geq 0$, $i \in B(\tilde{x}^*)$, and

$$f'(\tilde{x}^*, \tilde{d}) - \sum_{i \in B(\tilde{x}^*)} u_i g'_i(\tilde{x}^*, \tilde{d}) + \sum_{i=1}^m v_i^* h'_i(\tilde{x}^*, \tilde{d}) = 0$$

for all $\tilde{d} \in C(\tilde{x}^*)$. Defining $u_i^* = u_i$, $i \in B(\tilde{x}^*)$, and $u_i^* = 0$, $i \notin B(\tilde{x}^*)$, it follows that $\nabla_{\tilde{x}} L(\tilde{x}^*, \tilde{u}^*, \tilde{v}^*) = 0$ and $\tilde{u}^*$ satisfies (1.15) and (1.16).

Remarks 1.2. 1. The vectors $\tilde{u}^*$, $\tilde{v}^*$ are called generalized Lagrange multipliers (GLM) or dual variables. The latter term arises from the theory of convex programming and duality. For this special problem, the conditions of Theorem 1.1 are also sufficient for $\tilde{x}^*$ to be a global minimum on R .

2. If $\tilde{x}^*$, a local minimum on R , is not regular, then the GLM are not all finite.

3. If condition (i) of Lemma 1.4 holds, then the GLM are unique. This result follows from the uniqueness result of Lemma 1.5.

Assuming the problem functions have second order differentials, then the second order necessary and sufficient conditions may be developed.

Definition 1.9. Let (i) $B_1(\tilde{x}^*) = \{i: i \in B(\tilde{x}^*), u_i^* > 0\}$,
 (ii) $R_1 = \{\tilde{x}: \tilde{x} \in R, g_i(\tilde{x}) = 0, i \in B_1(\tilde{x}^*)\}$,
 and (iii) $C_1(\tilde{x}^*)$ be the tangent cone of R_1 at $\tilde{x}^*$.

Theorem 1.2.

If $\tilde{x}^*$ is a regular point of R and a local minimum of $f(\tilde{x})$ on R , then the first order necessary conditions of Theorem 1.1 are satisfied and, for all $\tilde{d} \in C_1(\tilde{x}^*)$,

$$\tilde{d}^T \nabla^2 L(\tilde{x}^*, \tilde{u}^*, \tilde{v}^*) \tilde{d} \geq 0. \quad (1.18)$$

Proof. Since $L(\tilde{x}, \tilde{u}^*, \tilde{v}^*) = f(\tilde{x})$ on R_1 and $\tilde{x}^*$ is a local minimum of $f(\tilde{x})$ on R_1 , the result follows immediately from Lemma 1.1(ii).

Remarks 1.3. 1. If $\tilde{d} \in C_1(\tilde{x}^*)$, then $\tilde{d} \in C(\tilde{x}^*)$ and $\tilde{d}^T \nabla g_{\tilde{i}}(\tilde{x}^*) = 0$, $\tilde{i} \in B_1(\tilde{x}^*)$.

2. $\tilde{x}^*$ is a regular point of R_1 if $\tilde{d} \in C_1(\tilde{x}^*)$ for all $\tilde{d} \in C(\tilde{x}^*)$ satisfying $\tilde{d}^T \nabla g_{\tilde{i}}(\tilde{x}^*) = 0$, $\tilde{i} \in B_1(\tilde{x}^*)$.

3. Sufficient conditions for $\tilde{x}^*$ to be a regular point of R_1 are provided by Condition (i) of Lemma 1.4.

Theorem 1.3.

If, at $\tilde{x}^*$, there exist vectors $\tilde{u}^*$ and $\tilde{v}^*$ satisfying the conditions of Theorem 1.1 and, for all $\tilde{d} \in C_1(\tilde{x}^*)$,

$$\tilde{d}^T \nabla^2 L(\tilde{x}^*, \tilde{u}^*, \tilde{v}^*) \tilde{d} > 0, \quad (1.19)$$

then $\tilde{x}^*$ is an isolated local minimum on R .

Proof. Assume that $\tilde{x}^*$ is not an isolated local minimum on R . Then for every integer, k , there exists $\tilde{x}^k \in R$, $\tilde{x}^k \neq \tilde{x}^*$ satisfying

$$||\tilde{x}^k - \tilde{x}^*|| < \frac{1}{k} \quad \text{and} \quad f(\tilde{x}^k) - f(\tilde{x}^*) < \frac{1}{k} ||\tilde{x}^k - \tilde{x}^*||^2. \quad (1.20)$$

From the sequence $\{\tilde{x}^k\}$, a subsequence, also denoted by $\{\tilde{x}^k\}$, can be chosen s.t. $\tilde{x}^k \xrightarrow{\tilde{d}} \tilde{x}^*$ where $\tilde{d} \in C(\tilde{x}^*)$.

Defining the function

$$G(\tilde{x}) = \sum_{i \in B(\tilde{x}^*)} u_i^* g_i(\tilde{x}) - \sum_{i=1}^m v_i^* h_i(\tilde{x})$$

we note that $G(\tilde{x}^*)=0$ and $G(\tilde{x}) \geq 0$ for all $\tilde{x} \in R$. For all $\tilde{x} \in R$ it follows that

$$f(\tilde{x}) = L(\tilde{x}, \tilde{u}^*, \tilde{v}^*) + G(\tilde{x})$$

and from (1.20),

$$\frac{L(\tilde{x}^k, \tilde{u}^*, \tilde{v}^*) - L(\tilde{x}^*, \tilde{u}^*, \tilde{v}^*)}{\|\tilde{x}^k - \tilde{x}^*\|^2} + \frac{G(\tilde{x}^k)}{\|\tilde{x}^k - \tilde{x}^*\|^2} < \frac{1}{k}.$$

Using (1.17), Definition 1.5(ii) and Remark 1.1(1) we have

$$\frac{1}{2} \tilde{d}^T \nabla^2 L(\tilde{x}^*, \tilde{u}^*, \tilde{v}^*) \tilde{d} + \limsup_{k \rightarrow \infty} \frac{G(\tilde{x}^k)}{\|\tilde{x}^k - \tilde{x}^*\|^2} \leq 0 \quad (1.21)$$

and since $G(\tilde{x}^k) \geq 0$, the second term is bounded. Therefore

$$\tilde{d}^T \nabla^2 L(\tilde{x}^*, \tilde{u}^*, \tilde{v}^*) \tilde{d} \leq 0. \quad (1.22)$$

It also follows from (1.21) and (1.11) that

$$\lim_{k \rightarrow \infty} \frac{G(\tilde{x}^k)}{\|\tilde{x}^k - \tilde{x}^*\|} = G'(\tilde{x}^*, \tilde{d}) = \sum_{i \in B_1(\tilde{x}^*)} u_i^* g'_i(\tilde{x}^*, \tilde{d}) = 0,$$

and since $u_i^* > 0$, $i \in B(\tilde{x}^*)$, then $g'_i(\tilde{x}^*, \tilde{d}) = 0$, $i \in B_1(\tilde{x}^*)$. But this implies $\tilde{d} \in C_1(\tilde{x}^*)$ and thus, (1.22) provides the contradiction.

CHAPTER 2

BARRIER FUNCTION METHODS2.1 Introduction.

In this chapter, barrier function transformations for the solution of Problem 1.4 will be studied. For the nonlinear form of Problem 1.4, the solutions generated are local solutions which, for the particular case of convex problems, can be shown to be global. The transformation is written

$$T(\tilde{x}, r) = f(\tilde{x}) + r \Phi(g(\tilde{x})) , \quad (2.1)$$

where r is a positive controlling parameter and Φ is a function of the inequality constraint vector $g(\tilde{x})$ and thus $\tilde{x}$. Convergence properties, including rates of convergence, will be discussed. The basic theory of barrier function transformations can be found in Fiacco and McCormick (1968).

Here, attention will be restricted to separable transformations of the form

$$\Phi(g(\tilde{x})) = \sum_{i=1}^{\ell} \phi_i(g_i(\tilde{x})) \quad (2.2)$$

with the ϕ_i defined by Definition 2.1.

Definition 2.1. The function ϕ_i , is a barrier function for $g_i(\underline{x})$ if

(i) ϕ_i is a continuous function of $g_i(\underline{x})$ for $\underline{x} \in R_0$ where R_0 is the nonempty interior of R , R given by (1.4) with $m=0$.

(ii) $\phi_i(g_i(\underline{x})) \rightarrow \infty$ as $g_i(\underline{x}) \rightarrow 0+$.

(iii) ϕ_i is a twice continuously differentiable function of $g_i(\underline{x})$ for $g_i(\underline{x}) > 0$ (i.e. $\underline{x} \in R_0$).

(iv) $\phi_i(g_i(\underline{x}))$ satisfies:

$$(a) \quad \frac{d\phi_i(g_i(\underline{x}))}{dg_i} < 0 \quad \text{for all } \underline{x} \in R_0,$$

$$(b) \quad \frac{d^2\phi_i(g_i(\underline{x}))}{dg_i^2} > 0 \quad \text{for all } \underline{x} \in R_0,$$

$$(c) \quad \frac{d^2\phi_i(g_i(\underline{x}))}{dg_i^2} \text{ is a monotonic decreasing}$$

function of $g_i(\underline{x})$, $\underline{x} \in R_0$.

Remarks 2.1. 1. Conditions (i) and (ii) of Definition 2.1 are the essential defining properties of a barrier function, while conditions (iii) and (iv) are only required to establish uniqueness properties for an isolated trajectory of minimizing points, $\underline{x}(r)$, of (2.1) as r tends to zero. (See Fiacco and McCormick (1968)).

2. The logarithmic barrier function (1.8) and the inverse barrier function (1.9) both satisfy (2.2) and Definition 2.1.

A basis, given by Fiacco and McCormick, for the barrier function methods is provided by the following lemmas and theorem. Proofs of the lemmas can be found in the references cited.

Lemma 2.1. (Rudin (1953), p.67).

A continuous function on a nonempty compact set attains its minimum on the set.

Lemma 2.2. (Fiacco and McCormick (1968), p.46).

Let R be a closed set, S a compact set and $R_0 \cap S \neq \emptyset$. If $T(\underline{x})$ is continuous on $R_0 \cap S$ with the property that for every sequence $\{\underline{x}^k\} \subset R_0 \cap S$ with $\underline{x}^k \rightarrow \hat{\underline{x}} \in (R - R_0) \cap S$, $\lim_{k \rightarrow \infty} T(\underline{x}^k) = \infty$, then $T(\underline{x})$ attains a finite minimum value on $R_0 \cap S$.

Lemma 2.3. (Fiacco and McCormick (1968), p.47).

If a set of local minima, A^* , corresponding to a local minimum value, f^* , of Problem 1.4 is a nonempty, isolated, compact set of $A = \{\underline{x} : f(\underline{x}) = f^*\}$, then there exists a compact set S such that $A^* \subset S_0$ and for any point $\underline{x} \in R \cap S$ such that $\underline{x} \notin A^*$, $f(\underline{x}) > f^*$ (i.e. the local minima in A^* are global minima on $R \cap S$).

Theorem 2.1. (Fiacco and McCormick (1968), p.47).

- If
- (i) the problem functions are continuous,
 - (ii) each ϕ_i satisfies Definition 2.1,
 - (iii) $\{r^k\}$ is a strictly decreasing null sequence,
- and
- (iv) there exists a nonempty isolated compact set, A^* , of local minima corresponding to a local minimum value, f^* , such that $A^* \cap \text{cl}(R_0) \neq \emptyset$, where $\text{cl}(\cdot)$ denotes closure,

then (a) there exists a compact set, S , such that $A^* \subset S_0$ and for k sufficiently large, the unconstrained local minima, $\tilde{x}^k$, of $T(\tilde{x}, r^k)$ given by (2.1) exist in $R_0 \cap S_0$ and every limit point of any subsequence, $\{\tilde{x}^k\}$, is in A^* ,

$$(b) \quad \lim_{k \rightarrow \infty} r^k \sum_{i=1}^{\ell} \phi_i (g_i(\tilde{x}^k)) = 0,$$

$$(c) \quad \lim_{k \rightarrow \infty} f(\tilde{x}^k) = f(\tilde{x}^*),$$

$$(d) \quad \lim_{k \rightarrow \infty} T(\tilde{x}^k, r^k) = f(\tilde{x}^*),$$

$$(e) \quad \{f(\tilde{x}^k)\} \text{ is a monotonic decreasing sequence,}$$

and

$$(f) \quad \left\{ \sum_{i=1}^{\ell} \phi_i (g_i(\tilde{x}^k)) \right\} \text{ is a monotonic increasing}$$

sequence.

Proof. (Fiacco and McCormick (1968), p.48 and p.60).

The proof is based directly on Definition 2.1 and Lemmas 2.1, 2.2 and 2.3.

A more general result is available for the transformation

$$\hat{T}(\tilde{x}, \tilde{r}^k) = f(\tilde{x}) + \sum_{i=1}^{\ell} r_i^k \phi_i(g_i(\tilde{x})),$$

where $\{\tilde{r}^k\}$ is a strictly decreasing null sequence with the ordering determined componentwise.

Theorem 2.2.

If conditions (i), (ii) and (iv) of Theorem 2.1 are satisfied and $\{\tilde{r}^k\}$ is a strictly decreasing null sequence, then conclusions (a), (b), (c) and (d) of Theorem 2.1 remain valid.

Proof. The proof is similar to the proof of Theorem 2.1.

Remark 2.2. Given a sequence of positive vectors tending to zero, it is possible to select a subsequence which is strictly decreasing. Theorem 2.2 may be applied using such a subsequence.

2.2 Trajectory Analysis and Error Estimates.

An important feature of the barrier function transformation is that a GLM vector, $\underline{u}^*$, satisfying the Kuhn-Tucker necessary conditions (1.15), (1.16) and (1.17) at $\underline{x}^*$, can be derived from the sequence of unconstrained minimizations if either condition (i) or condition (ii) of Lemma 1.4 is satisfied. For condition (i), this result is given by Fiacco and McCormick (1968), p.73. A constructive proof of the result for condition (ii) has been given by Beltrami (1969) and an alternative approach has been discussed by Osborne and Ryan (1970).

Let $K(\underline{x}^*)$ be the matrix with columns formed from the vectors $\nabla g_i(\underline{x}^*)$, $i \in B(\underline{x}^*)$, and assume $K(\underline{x}^*)$ has rank $s \leq n$. After reordering the columns, if necessary, let $K(\underline{x}^*)$ be partitioned into matrices, $K_1(\underline{x}^*)$ with rank s and $K_2(\underline{x}^*)$, such that $K_2(\underline{x}^*) = K_1(\underline{x}^*)U$, where U expresses the linear dependence of columns of $K_2(\underline{x}^*)$ on columns of $K_1(\underline{x}^*)$. Denoting the matrices formed from the columns of $K_1(\underline{x}^*)$ and $K_2(\underline{x}^*)$, evaluated at $\underline{x}$, by $K_1(\underline{x})$ and $K_2(\underline{x})$ respectively, it follows that the rank of $K_1(\underline{x})$ is s if $\underline{x}$ is sufficiently close to $\underline{x}^*$. Finally, define

$$\underline{u}_i^k = -r^k \frac{d\phi_i(g_i(\underline{x}^k))}{dg_i}, \quad i=\overline{1, \ell}, \quad (2.3)$$

and let the components of α^k and β^k be selected from the components of u^k corresponding to the constraints of $K_1(\tilde{x}^k)$ and $K_2(\tilde{x}^k)$ respectively. It should be noted, using Definition 2.1, that u_i^k , $i=\overline{1, \ell}$, are bounded below and $u_i^k \rightarrow 0$ as $r^k \rightarrow 0$ for all $i \notin B(\tilde{x}^*)$.

Theorem 2.3. (Osborne and Ryan (1970)).

If $\{r^k\}$ is a strictly decreasing null sequence s.t. $\tilde{x}^k \rightarrow \tilde{x}^*$ and the components of β^k are bounded for every $k \geq \hat{k}$, where $\hat{k}$ is sufficiently large, then the sequence of vectors $\alpha^k + U\beta^k$, $k = \hat{k}, \hat{k}+1, \dots$, converges and

$$\lim_{k \rightarrow \infty} (\alpha^k + U\beta^k) = [K_1(\tilde{x}^*)^T K_1(\tilde{x}^*)]^{-1} K_1(\tilde{x}^*)^T \nabla f(\tilde{x}^*). \quad (2.4)$$

Proof. Since $\tilde{x}^k$ minimizes $T(\tilde{x}, r^k)$,

$$\begin{aligned} \nabla f(\tilde{x}^k) &= \sum_{i \in B(\tilde{x}^*)} -r^k \frac{d\phi_i(g_i(\tilde{x}^k))}{dg_i} \nabla g_i(\tilde{x}^k) + o(r^k) \\ &= K_1(\tilde{x}^k) \alpha^k + K_2(\tilde{x}^k) \beta^k + o(r^k) \\ &= K_1(\tilde{x}^k) (\alpha^k + U\beta^k) + (K_2(\tilde{x}^k) - K_1(\tilde{x}^k)U) \beta^k + o(r^k). \end{aligned} \quad (2.5)$$

But (2.5) is compatible by construction and therefore

$$\begin{aligned} (\alpha^k + U\beta^k) &= [K_1(\tilde{x}^k)^T K_1(\tilde{x}^k)]^{-1} K_1(\tilde{x}^k)^T \{ \nabla f(\tilde{x}^k) - \\ &\quad (K_2(\tilde{x}^k) - K_1(\tilde{x}^k)U) \beta^k + o(r^k) \}. \end{aligned} \quad (2.6)$$

The result follows from (2.6) using the boundedness of β^k .

Corollary 1. If α^k and β^k are bounded for every $k \geq \hat{k}$, then there exists a vector, u^* , satisfying the Kuhn-Tucker conditions (1.15), (1.16) and (1.17) at x^* .

Proof. It does not follow from Theorem 2.3 that α^k and β^k converge but, using the boundedness of α^k and β^k , it is possible to select a convergent subsequence, $\{(\alpha^k, \beta^k)\}$, with limit (α, β) satisfying

$$\alpha + U\beta = [K_1(x^*)^T K_1(x^*)]^{-1} K_1(x^*)^T \nabla f(x^*) .$$

When padded out with zeros corresponding to the inactive inequality constraints (i.e. $i \notin B(x^*)$), the limit yields the required GLM vector, u^* .

Corollary 2. If the Kuhn-Tucker necessary conditions are not satisfied at x^* , then the sequence, $\{\max_{i \in B(x^*)} [u_i^k]\}$, is unbounded as $k \rightarrow \infty$.

Corollary 3. If condition (i) of Lemma 1.4 is satisfied, then there exists a constant, γ , s.t.

$$\max_{i \in B(x^*)} [u_i^k] < \gamma \text{ for all } k \geq \hat{k} .$$

Proof. By condition (i) of Lemma 1.4, $K(x^*)$ has full rank (i.e. $K(x^*) = K_1(x^*)$) and for $\hat{k}$ sufficiently large, $K(x^k)$, $k \geq \hat{k}$, also has full rank. It follows from (2.6) that the elements of α^k are bounded for $\hat{k}$ sufficiently large.

It is now shown that the components of α^k and β^k are bounded if condition (ii) of Lemma 1.4 is satisfied, and thus, the existence of a GLM vector, u^* , is established by Corollary 1 of Theorem 2.3.

Theorem 2.4. (Osborne and Ryan (1970)).

If condition (ii) of Lemma 1.4 is satisfied, then there exists a constant, γ , s.t.

$$\max_{i \in B(\tilde{x}^*)} [u_i^k] < \gamma \quad \text{for all } k \geq \hat{k}.$$

Proof. For $\hat{k}$ sufficiently large and using the continuity of the constraint functions,

$$\tilde{d}^T \nabla g_i(\tilde{x}^k) > 0, \quad i \in B(\tilde{x}^*), \quad k \geq \hat{k}, \quad (2.7)$$

where the existence of $\tilde{d}$ is guaranteed by Lemma 1.4(ii). At $\tilde{x}^k$, a local minimum of $T(\tilde{x}, r^k)$,

$$\nabla f(\tilde{x}^k) = \sum_{i \in B(\tilde{x}^*)} u_i^k \nabla g_i(\tilde{x}^k) + o(r^k), \quad (2.8)$$

whence

$$\sum_{i \in B(\tilde{x}^*)} u_i^k \leq \frac{\tilde{d}^T \nabla f(\tilde{x}^k) - o(r^k)}{\min_{i \in B(\tilde{x}^*)} [\tilde{d}^T \nabla g_i(\tilde{x}^k)]}. \quad (2.9)$$

The desired result follows from (2.9) as each term in the sum on the left is positive and the denominator on the right is bounded away from zero by (2.7) for $k \geq \hat{k}$.

Using the convergence properties of the GLM estimates, $\underline{u}^k$, an error expression for $[f(\underline{x}^k) - f(\underline{x}^*)]$ is now developed in terms of the problem functions and the GLM estimates at $\underline{x}^k$. A second application for the GLM estimates is discussed in Chapter 4.

Theorem 2.5. (Osborne and Ryan (1970)).

- If (i) the problem functions are twice continuously differentiable,
 (ii) $\{r^k\}$ is a strictly decreasing null sequence
 s.t. $\underline{x}^k \rightarrow \underline{x}^*$,
 (iii) the components of $\underline{u}^k$ are bounded for all
 $k \geq \hat{k}$, and
 (iv) $(\underline{x}^k - \underline{x}^*)^T \nabla f(\underline{x}^k)$ satisfies a strict order relation

$$(\underline{x}^k - \underline{x}^*)^T \nabla f(\underline{x}^k) = S0(||\underline{x}^k - \underline{x}^*||), \quad (2.10)$$

where the S0 notation indicates that

$$\min_{k > \hat{k}} \frac{|(\underline{x}^k - \underline{x}^*)^T \nabla f(\underline{x}^k)|}{||\underline{x}^k - \underline{x}^*||} > 0,$$

then

$$f(\underline{x}^k) - f(\underline{x}^*) = E_k + O(\max[r^k ||\underline{x}^k - \underline{x}^*||, ||\underline{x}^k - \underline{x}^*||^2]), \quad (2.11)$$

where

$$E_k = \sum_{i \in B(\tilde{x}^*)} u_i^k g_i(\tilde{x}^k) \quad (2.12)$$

and the left hand side of (2.11) is $S_0(\|\tilde{x}^k - \tilde{x}^*\|)$.

Proof. The result follows by taking the scalar product of both sides of (2.8) with $(\tilde{x}^k - \tilde{x}^*)$ and using the first order Taylor expansions about $\tilde{x}^*$ for the functions. It should be noted that (2.11) is valid without (2.10) which only ensures that E_k dominates the right hand side.

Corollary 1. If ϕ_i is the inverse barrier function given by (1.9), then $E_k = O([r^k]^{\frac{1}{2}})$. If ϕ_i is the log barrier function given by (1.8), then $E_k = O(r^k)$.

Proof. For the inverse barrier function, under the conditions of the Theorem, (2.3) implies $g_i(\tilde{x}^k) = O([r^k]^{\frac{1}{2}})$, $i \in B(\tilde{x}^*)$, and the result follows from (2.12). For the log barrier function, (2.3) implies $g_i(\tilde{x}^k) = O(r^k)$, and again the result follows from (2.12).

Remark 2.3. Corollary 1 of Theorem 2.5 implies that, for a given strictly decreasing null sequence $\{r^k\}$, the log barrier function can be expected to provide a faster convergence to $\tilde{x}^*$ than the inverse barrier function. This observation motivates, in part, the results of §2.3.

The strict order condition (2.10) is related to the condition of strict complementarity,

$$u_i^* > 0 \text{ for all } i \in B(\tilde{x}^*), \quad (2.13)$$

which is used by Fiacco and McCormick (1968) (Chapter 6) in a similar context.

Theorem 2.6. (Osborne and Ryan (1970)).

If conditions (i), (ii) and (iii) of Theorem 2.5 hold, then

(a) for the inverse and log barrier functions, the strict order condition (2.10), implies the strict complementarity condition (2.13), and

(b) for the log barrier function, assuming Lemma 1.4(ii) is satisfied, condition (2.13) implies (2.10).

Proof. (a) Using Theorem 2.5(i),

$$g_i(\tilde{x}^k) = (\tilde{x}^k - \tilde{x}^*)^T \nabla g_i(\tilde{x}^*) + o(\|\tilde{x}^k - \tilde{x}^*\|^2)$$

for all $i \in B(\tilde{x}^*)$ so that at most

$$g_i(\tilde{x}^k) = o(\|\tilde{x}^k - \tilde{x}^*\|), \quad i \in B(\tilde{x}^*). \quad (2.14)$$

If (2.13) is not satisfied, then for at least one $i \in B(\tilde{x}^*)$,

$$\lim_{k \rightarrow \infty} u_i^k = \lim_{k \rightarrow \infty} -r^k \frac{d\phi_i(g_i(\tilde{x}^k))}{dg_i} = 0. \quad (2.15)$$

For the inverse barrier function, (2.15) implies $r^k = o(g_i(\tilde{x}^k)^2)$. Using (2.14), $r^k = o(\|\tilde{x}^k - \tilde{x}^*\|^2)$. Thus (2.11) becomes

$$(\tilde{x}^k - \tilde{x}^*)^T \nabla f(\tilde{x}^k) = o(\|\tilde{x}^k - \tilde{x}^*\|) \quad (2.16)$$

showing that (2.10) is not satisfied.

For the log barrier function, (2.15) implies $r^k = o(g_i(\tilde{x}^k))$. Using (2.14), $r^k = o(\|\tilde{x}^k - \tilde{x}^*\|)$. Thus (2.11) takes the form

$$(\tilde{x}^k - \tilde{x}^*)^T \nabla f(\tilde{x}^k) = \sum_{i \in B(\tilde{x}^*)} r^k + o(\|\tilde{x}^k - \tilde{x}^*\|^2) \quad (2.17)$$

again showing that (2.10) is not satisfied.

(b) For the log barrier function, the required result follows from Theorem 15 of Fiacco and McCormick (1968), p.81, which guarantees the existence of derivatives of the minimizing trajectory, $\tilde{x} = \tilde{x}(r)$, in a neighbourhood of $r = 0$, to an order consequent on the smoothness of the problem functions. Let the α -order derivative be the first non-vanishing derivative at $r = 0$. Then

$$\tilde{x}^k - \tilde{x}^* = \frac{[r^k]^\alpha}{\alpha!} \frac{d^\alpha \tilde{x}(0)}{dr^\alpha} + o([r^k]^{\alpha+1}),$$

and thus, $r^k = S_0(\|x^k - x^*\|^{\frac{1}{\alpha}})$, $\alpha \geq 1$. But by (2.17), $\alpha = 1$ is the only case for which (2.13) holds and thus (2.10) is satisfied.

The error expression (2.12) can be used in a number of ways. First, if the active constraints at x^* (i.e. $i \in B(x^*)$) can be determined, then (2.12) provides an asymptotically correct error estimate of $[f(x^k) - f(x^*)]$. Secondly, the sum in (2.12) can be extended over all constraints and in this case, a bound for the error, valid for r^k sufficiently small, should be obtained. It is known that this bound is strict when Problem 1.4 is convex. For this problem, it can be derived using properties of the dual problem (Fiacco and McCormick (1968), p.98). As the inactive constraints do not influence the solution, the sum over active constraints also gives a strict bound.

Osborne and Ryan (1970) have treated a number of simple numerical examples to demonstrate the effect of violating the Kuhn-Tucker conditions or conditions (2.10) and (2.13). They show that rates of convergence of barrier function methods are not necessarily affected adversely if the Kuhn-Tucker conditions do not hold; but they can be adversely affected if either (2.10) or (2.13) is violated.

2.3 A New Family of Barrier Functions.

Motivated by the observations of Remark 2.3 concerning the inverse and log barrier functions, a new family of barrier functions is constructed with improved convergence properties for a given strictly decreasing null sequence, $\{r^k\}$. This is achieved by choosing ϕ to "reduce" the error, E_k , given by (2.12).

If R is bounded, then for $\delta > 0$, let G_i , $i=\overline{1, \ell}$, be chosen so that

$$G_i > 1 + \delta + \max_{\tilde{x} \in R_0} \log(g_i(\tilde{x})), \quad i=\overline{1, \ell},$$

and if R is unbounded, choose G_i sufficiently large so that when $\log(g_i(\tilde{x})) \geq G_i - 1$, $g_i(\tilde{x})$ can safely be ignored in any reasonably scaled problem. Let

$$\hat{R} = \{\tilde{x} : \tilde{x} \in R_0; 0 < g_i(\tilde{x}) < e^{G_i - 1 - \delta}, i=\overline{1, \ell}, \delta > 0\}.$$

Remarks 2.4. 1. If $G_i = 10$, then $g_i(\tilde{x})$ can take values of order 10^4 on $\hat{R}$ and, in particular, near the minimum of the transformation function. Such a constraint is unlikely to be violated under these conditions if the problem is reasonably scaled.

2. If R is bounded, then $\hat{R} = R_0$.

Define, for $\underline{x} \in \hat{R}$,

$$\left. \begin{aligned} \phi_i^{(1)}(g_i(\underline{x})) &= \log[G_i - \log(g_i(\underline{x}))] \\ \phi_i^{(j+1)}(g_i(\underline{x})) &= \log[\sigma_i^{(j)} + \phi_i^{(j)}(g_i(\underline{x}))], \quad j=1,2,\dots \end{aligned} \right\} i=\overline{1,\ell}, \quad (2.18)$$

where $\sigma_i^{(j)} \geq 1$, $i=\overline{1,\ell}$, $j=1,2,\dots$, and write

$$\Phi^{(j)}(\underline{g}(\underline{x})) = \sum_{i=1}^{\ell} \phi_i^{(j)}(g_i(\underline{x})). \quad (2.19)$$

Then the family of functions, $\Phi^{(j)}(\underline{g}(\underline{x}))$, $j=1,2,\dots$, are barrier functions in the following restricted sense.

Lemma 2.1.

Writing $\phi_i^{(j)}$ for $\phi_i^{(j)}(g_i(\underline{x}))$ and g_i for $g_i(\underline{x})$, then

- (a) the function, $\phi_i^{(1)}$, satisfies Definition 2.1 on $\hat{R}$,
- (b) the functions, $\phi_i^{(j)}$, $j=2,3,\dots$, satisfy properties (i), (ii), (iii) and (iva) of Definition 2.1 on $\hat{R}$ and
- (c) $\sigma_i^{(j)}$ can be chosen so that $\phi_i^{(j+1)}$, $j=1,2,\dots$, is convex on $\hat{R}$ (i.e. satisfies Definition 2.1 (ivb)).

Proof. Definition 2.1 (i), (ii) and (iii) are obviously satisfied by $\phi_i^{(j)}$, $j=1,2,\dots$, on $\hat{R}$.

Differentiating (2.18) w.r.t. g_i gives

$$\frac{d\phi_i^{(j)}}{dg_i} = \frac{-1}{g_i [G_i - \log(g_i)]} \prod_{s=1}^{j-1} \frac{1}{\sigma_i^{(s)} + \phi_i^{(s)}}, \quad j=1, 2, \dots, \quad (2.20)$$

and for $\tilde{x} \in \hat{R}$ and $\sigma_i^{(j)} \geq 1$, $\frac{d\phi_i^{(j)}}{dg_i} < 0$, $j=1, 2, \dots$, thus

satisfying Definition 2.1(iva) on $\hat{R}$ and establishing part (b) of the Lemma. Part (a) of the Lemma is verified by examining the properties of

$$\frac{d^2\phi_i^{(1)}}{dg_i^2} \quad \text{and} \quad \frac{d^3\phi_i^{(1)}}{dg_i^3} \quad \text{on } \hat{R}. \quad \text{Part (c) of the Lemma is}$$

proved inductively using the convexity of $\phi_i^{(1)}$ established in (a). From (2.18),

$$\frac{d^2\phi_i^{(j+1)}}{dg_i^2} = \frac{1}{\sigma_i^{(j)} + \phi_i^{(j)}} \left[\frac{d^2\phi_i^{(j)}}{dg_i^2} - \frac{1}{\sigma_i^{(j)} + \phi_i^{(j)}} \left[\frac{d\phi_i^{(j)}}{dg_i} \right]^2 \right].$$

If $\frac{d^2\phi_i^{(j)}}{dg_i^2} > 0$ for $\tilde{x} \in \hat{R}$, then $\frac{d^2\phi_i^{(j+1)}}{dg_i^2} > 0$ for

$\tilde{x} \in \hat{R}$ iff

$$\sigma_i^{(j)} > \max_{\tilde{x} \in \hat{R}} \left\{ \left[\frac{d\phi_i^{(j)}}{dg_i} \right]^2 / \frac{d^2\phi_i^{(j)}}{dg_i^2} - \phi_i^{(j)} \right\} = S_i^{(j)},$$

and thus $\sigma_i^{(j)} = \max[1, S_i^{(j)} + \varepsilon]$, $\varepsilon > 0$, is chosen to ensure the convexity of $\phi_i^{(j+1)}$ assuming $\phi_i^{(j)}$ is convex. It should be noted that $S_i^{(j)}$ is bounded above on $\hat{R}$.

Remarks 2.5. 1. As $g_i \rightarrow 0$, $\phi_i^{(j)} \rightarrow \infty$, $j=1,2,\dots$, and therefore, on a small enough neighbourhood of $\tilde{x}^*$, $\sigma_i^{(j)} = 1$, $j=1,2,\dots$, $i \in B(\tilde{x}^*)$, can be expected to preserve convexity.

2. The functions, $\phi_i^{(j)}$, $j=2,3,\dots$, need not satisfy Definition 2.1(ivc). This is not a serious disadvantage, since this property is not required for the basic convergence results of the barrier function methods.

3. For $\phi_i^{(1)}$, if the strict order condition (2.10) is satisfied under conditions (i), (ii) and (iii) of Theorem 2.5, then the strict complementarity condition (2.13) is satisfied. Using (2.18), (2.15) becomes

$$\lim_{k \rightarrow \infty} \frac{r^k}{g_i(\tilde{x}^k) [G_i - \log(g_i(\tilde{x}^k))]} = 0,$$

for at least one $i \in B(\tilde{x}^*)$. Since $\lim_{t \rightarrow 0} t \log(t) = 0$, then at least, $r^k = o(g_i(\tilde{x}^k))$. It follows from (2.11), using (2.14), that

$$(\tilde{x}^k - \tilde{x}^*)^T \nabla f(\tilde{x}^k) = o \left(\frac{\|\tilde{x}^k - \tilde{x}^*\|}{[G_i - \log(g_i(\tilde{x}^k))]} \right),$$

which shows that (2.10) is not satisfied.

For the barrier function family (2.18), the error estimate (2.12) becomes

$$E_k^{(j)} = \sum_{i \in B(\tilde{x}^*)} \frac{r^k}{[G_i - \log(g_i(\tilde{x}^k))]} \prod_{s=1}^{j-1} \frac{1}{\sigma_i^{(s)} + \phi_i^{(s)}(g_i(\tilde{x}^k))},$$

$$j=1, 2, \dots, \quad (2.21)$$

which would appear to provide a faster rate of convergence for each successive barrier function, $\phi^{(j)}$, and fixed r^k -sequence. In §2.4, this observation is examined numerically.

A further application of (2.21) shows that in certain circumstances, it is possible to choose a barrier function from the family (2.18) with the property that the solution of Problem 1.4 is approximated arbitrarily closely by the result of a single unconstrained minimization of the corresponding transformation function. Let

$$T^{(j)}(\tilde{x}, \gamma) = f(\tilde{x}) + \gamma \sum_{i=1}^{\ell} \phi_i^{(j)}(g_i(\tilde{x})), \quad \gamma > 0,$$

and

$$Q(\tilde{x}, r) = f(\tilde{x}) + \sum_{i=1}^{\ell} [-r_i \log(g_i(\tilde{x}))].$$

Theorem 2.6. (Osborne and Ryan (1970)).

Let $Q(\underline{x}, \underline{r})$ have a unique minimum stationary value on R_0 for $\underline{r} > 0$ and, choosing G_i , $i=\overline{1, \ell}$, sufficiently large, let $\underline{x}^{(j)} \in \hat{R}$ minimize $T^{(j)}(\underline{x}, \gamma)$ for $j=1, 2, \dots$, and γ fixed. Then the limit points of the sequence $\{\underline{x}^{(j)}\}$ are local minima of Problem 1.4.

Proof. If $\underline{x}^{(j)}$ minimizes $T^{(j)}(\underline{x}, \gamma)$, then

$$\nabla f(\underline{x}^{(j)}) - \gamma \sum_{i=1}^{\ell} \frac{1}{[G_i - \log(g_i(\underline{x}^{(j)}))]} \prod_{s=1}^{j-1} \frac{1}{\sigma_i^{(s)} + \phi_i^{(s)}(g_i(\underline{x}^{(j)}))} \\ \frac{1}{g_i(\underline{x}^{(j)})} \nabla g_i(\underline{x}^{(j)}) = 0$$

which is in the form $\nabla Q(\underline{x}^{(j)}, \underline{r}^{(j)}) = 0$, where

$$r_i^{(j)} = \frac{\gamma}{[G_i - \log(g_i(\underline{x}^{(j)}))]} \prod_{s=1}^{j-1} \frac{1}{\sigma_i^{(s)} + \phi_i^{(s)}(g_i(\underline{x}^{(j)}))}, i=\overline{1, \ell}.$$

The assumed uniqueness of these stationary values implies that the sequence $\{\underline{x}^{(j)}\}$ also corresponds to the sequence minimizing $Q(\underline{x}, \underline{r}^{(j)})$. Now, since $\sigma_i^{(s)} \geq 1$ and $\phi_i^{(s)} > 0$, $i=\overline{1, \ell}$, $s=1, 2, \dots$, $r_i^{(j)}$ can be made as small as desired for each i by taking j large enough. The required result is then a consequence of Theorem 2.2 and Remark 2.2 noting that $Q(\underline{x}, \underline{r})$ is the logarithmic transformation function.

Remark 2.6. If Problem 1.4 is convex, then $Q(\underline{x}, \underline{r})$ is convex for $\underline{r} > 0$, and further, if $f(\underline{x})$ is strictly convex or any $g_i(\underline{x})$, $i=\overline{1, \ell}$, is strictly concave, then $Q(\underline{x}, \underline{r})$ is strictly convex. In this latter case, the conditions of Theorem 2.6 are satisfied and the local minimum of Problem 1.4 is global.

2.4 Numerical Results.

Numerical results from several test problems are now presented to illustrate the results of §2.2 and §2.3. The following four problems are used.

(i) Kuhn-Tucker. (Fiacco and McCormick (1968), p.22).

$$\begin{aligned} \text{Minimize} \quad & f(\underline{x}) = -x_1 \\ \text{subject to} \quad & g_1(\underline{x}) = (1-x_1)^3 - x_2 \geq 0, \\ & g_2(\underline{x}) = x_1 \geq 0, \\ & g_3(\underline{x}) = x_2 \geq 0. \end{aligned}$$

The constrained minimum is at $\underline{x}^* = (1, 0)^T$ with $f(\underline{x}^*) = -1$ and the first and third constraints active. An initial feasible point is $\underline{x}^0 = (0.25, 0.25)^T$ with $f(\underline{x}^0) = -0.25$. We note that the Kuhn-Tucker conditions do not hold at $\underline{x}^*$.

(ii) Rosen-Suzuki. (Rosen and Suzuki (1965)).

$$\begin{aligned} \text{Minimize} \quad & f(\underline{x}) = x_1^2 + x_2^2 + 2x_3^2 + x_4^2 - 5x_1 - 5x_2 - 21x_3 + 7x_4 \\ \text{subject to} \quad & g_1(\underline{x}) = -x_1^2 - x_2^2 - x_3^2 - x_1 + x_2 - x_3 + x_4 + 8 \geq 0, \\ & g_2(\underline{x}) = -x_1^2 - 2x_2^2 - x_3^2 - 2x_4^2 + x_1 + x_4 + 10 \geq 0, \\ & g_3(\underline{x}) = -2x_1^2 - x_2^2 - x_3^2 - 2x_1 + x_2 + x_4 + 5 \geq 0, \end{aligned}$$

The constrained minimum is at $\underline{x}^* = (0, 1, 2, -1)^T$ with $f(\underline{x}^*) = -44$ and the first and third constraints active. An initial feasible point is $\underline{x}^0 = (0, 0, 0, 0)^T$ with $f(\underline{x}^0) = 0$.

(iii) Post Office parcel problem. (Rosenbrock (1960)).

$$\begin{aligned} \text{Minimize} \quad & f(\underline{x}) = -x_1 x_2 x_3 \\ \text{subject to} \quad & 0 \leq x_1 \leq 20, \quad 0 \leq x_2 \leq 11, \quad 0 \leq x_3 \leq 42 \text{ and} \\ & 0 \leq x_1 + 2x_2 + 2x_3 \leq 72. \end{aligned}$$

The constrained minimum is at $\underline{x}^* = (20, 11, 15)^T$ with $f(\underline{x}^*) = -3300$ and the upper bounds on the first and second variables as well as the last constraint active. An initial point is $\underline{x}^0 = (15, 3, 25)^T$ with $f(\underline{x}^0) = -1125$.

(iv) Colville: Problem 1. (Colville (1968)).

$$\begin{aligned} \text{Minimize} \quad & f(\underline{x}) = \sum_{j=1}^5 e_j x_j + \sum_{j=1}^5 \sum_{i=1}^5 c_{ij} x_i x_j + \sum_{j=1}^5 d_j x_j^3 \\ \text{subject to} \quad & g_i(\underline{x}) = \sum_{j=1}^5 a_{ij} x_j + b_i \geq 0, \quad i = \overline{1, 10}, \\ \text{and} \quad & g_{i+10}(\underline{x}) = x_i \geq 0, \quad i = \overline{1, 5}, \end{aligned}$$

where a_{ij} , c_{ij} , b_i , d_j , e_j are given in Table 2.1.

$i \backslash j$		1	2	3	4	5		
e_j		-15	-27	-36	-18	-12		
c_{ij}	1	30	-20	-10	32	-10		
	2	-20	39	-6	-31	32		
	3	-10	-6	10	-6	-10		
	4	32	-31	-6	39	-20		
	5	-10	32	-10	-20	30		
d_j		4	8	10	6	2	b_i	
a_{ij}	1	-16	2	0	1	0	40	
	2	0	-2	0	0.4	2	2	
	3	-3.5	0	2	0	0	0.25	
	4	0	-2	0	-4	-1	4	
	5	0	-9	-2	1	-2.8	4	
	6	2	0	-4	0	0	1	
	7	-1	-1	-1	-1	-1	40	
	8	-1	-2	-3	-2	-1	60	
	9	1	2	3	4	5	-5	
	10	1	1	1	1	1	-1	

Table 2.1. Coefficients for Colville: Problem 1

The constrained minimum is at $\tilde{x}^* = (0.3, 0.3334676, 0.4, 0.4283101, 0.2239649)^T$ with $f(\tilde{x}^*) = -32.34868$ and the third, fifth, sixth and ninth constraints active. An initial feasible point is $\tilde{x}^0 = (0.125, 0.0625, 0.125, 0.125, 1.0)^T$ with $f(\tilde{x}^0) = 4.724609$.

Each problem has been solved using the three barrier function transformations (1.8), (1.9) and

$$T(\tilde{x}, r^k) = f(\tilde{x}) + r^k \sum_{i=1}^{\ell} \log[10 - \log(g_i(\tilde{x}))]. \quad (2.22)$$

The transformation (2.22) is the first member of the family (2.18) developed in §2.3. The Davidon-Fletcher-Powell method, implementing a rather inefficient golden-section algorithm for linear searches, was used for the unconstrained minimizations.

B.F.	r^k	x_1^k	x_2^k	$f(\tilde{x}^k)$
(1.9)	10^0	0.191936	2.6382D-1	-0.191936
	10^{-1}	0.215089	2.4179D-1	-0.215089
	10^{-2}	0.419509	9.7803D-2	-0.419509
	10^{-3}	0.669210	1.8098D-2	-0.669210
(1.8)	10^0	0.162278	2.9395D-1	-0.162278
	10^{-1}	0.500000	6.2500D-2	-0.500000
	10^{-2}	0.940631	1.0463D-4	-0.940631
	10^{-3}	0.994006	1.0767D-7	-0.994006
(2.22)	10^0	0.617140	2.8060D-2	-0.617140
	10^{-1}	0.972317	1.0607D-5	-0.972317
	10^{-2}	0.997952	4.2958D-9	-0.997952
	10^{-3}	0.999837	2.1553D-12	-0.999837

Table 2.2(a). Kuhn-Tucker

The results, generated by the r^k -sequence, $r^k = 10^{1-k}$, $k=1,2,\dots$, are given in Tables 2.2 - 2.5. Each table is subdivided, part (a) showing the convergence of $\tilde{x}^k$ and $f(\tilde{x}^k)$ and part (b) showing the behaviour of u_i^k , $i \in B(\tilde{x}^*)$, E_k given by (2.12) and the predicted solution minimum, $\hat{f}(\tilde{x}^*) = f(\tilde{x}^k) - E_k$.

It can be seen that (2.22) gives the best rate of convergence and the tables are continued until the solutions for this transformation are correct to seven figures. Table 2.2 is truncated at an earlier stage since the rapid convergence of x_2 to zero causes difficulties with the tolerances of the linear search procedure. The results obtained are in excellent agreement with the predictions of §2.2 and §2.3. We note, in particular, the divergence of the GLM estimates in Table 2.2(b), the convergence rates associated with the three barrier functions in Tables 2.3 - 2.5 and the accuracy of the error estimate (2.12).

B.F.	r^k	$-r^k \frac{d\phi_1}{dg_1}$	$-r^k \frac{d\phi_3}{dg_3}$	E_k	$f(\tilde{x}^k) - E_k$
(1.9)	10^0	1.4368D 1	1.4368D 1	7.5809D 0	-7.772872
	10^{-1}	1.7106D 0	1.7106D 0	8.2718D-1	-1.042266
	10^{-2}	1.0454D 0	1.0452D 0	2.0449D-1	-0.624011
	10^{-3}	3.0531D 0	3.0531D 0	1.1051D-1	-0.779719
(1.8)	10^0	3.4020D 0	3.4020D 0	2.0000D 0	-2.162277
	10^{-1}	1.6000D 0	1.6000D 0	2.0000D-1	-0.700001
	10^{-2}	9.5577D 1	9.5577D 1	2.0000D-2	-0.960631
	10^{-3}	9.2872D 3	9.2872D 3	2.0000D-3	-0.996006
(2.22)	10^0	2.6256D 0	2.6256D 0	1.4735D-1	-0.764487
	10^{-1}	4.3943D 2	4.3943D 2	9.3223D-3	-0.981639
	10^{-2}	7.9542D 4	7.9542D 4	6.8339D-4	-0.998635
	10^{-3}	1.2586D 7	1.2586D 7	5.4255D-5	-0.999892

Table 2.2(b). Kuhn-Tucker

B.F.	r^k	x_1^k	x_2^k	x_3^k	x_4^k	$f(x^k)$
(1.9)	10^0	-1.5642D-2	0.912479	1.899167	-0.835025	-41.46807
	10^{-1}	-8.7436D-3	0.970932	1.967363	-0.957171	-43.22576
	10^{-2}	-3.6955D-3	0.993318	1.988775	-0.988594	-43.75809
	10^{-3}	-1.2667D-3	0.998503	1.996281	-0.996571	-43.92364
	10^{-4}	-4.0913D-4	0.999606	1.998804	-0.998928	-43.97586
	10^{-5}	-1.3014D-4	0.999884	1.999620	-0.999662	-43.99236
(1.8)	10^0	4.1762D-2	0.867958	1.938858	-0.736371	-41.63005
	10^{-1}	5.7665D-3	0.968053	2.000055	-0.970539	-43.78449
	10^{-2}	4.5151D-4	0.995907	2.000163	-0.997545	-43.97980
	10^{-3}	4.3409D-5	0.999578	2.000018	-0.999761	-43.99800
	10^{-4}	4.2939D-6	0.999957	2.000002	-0.999976	-43.99980
	10^{-5}	4.3817D-7	0.999996	2.000000	-0.999998	-43.99998
(2.22)	10^0	6.0530D-3	0.967234	2.002343	-0.737576	-43.82941
	10^{-1}	4.4360D-4	0.995937	2.000457	-0.981749	-43.98682
	10^{-2}	4.0860D-5	0.999585	2.000054	-0.999844	-43.99888
	10^{-3}	3.9152D-6	0.999958	2.000005	-0.999986	-43.99990
	10^{-4}	4.0358D-7	0.999996	2.000001	-0.999999	-43.99999
	10^{-5}	3.6478D-8	1.000000	2.000000	-1.000000	-44.00000

Table 2.3(a). Rosen-Suzuki

B.F.	r^k	$-r^k \frac{d\phi_1}{dg_1}$	$-r^k \frac{d\phi_2}{dg_2}$	E_k	$f(\underline{x}^k) - E_k$
(1.9)	10^0	2.235698	0.895115	2.4413D 0	-43.90940
	10^{-1}	2.106723	0.942892	7.6606D-1	-43.99182
	10^{-2}	2.030369	0.991962	2.4209D-1	-44.00017
	10^{-3}	2.007888	1.000397	7.6438D-2	-44.00008
	10^{-4}	2.002246	1.000520	2.4153D-2	-44.00001
	10^{-5}	2.000682	1.000208	7.6355D-3	-44.00000
(1.8)	10^0	1.921990	0.918628	2.0000D 0	-43.63005
	10^{-1}	2.059076	0.889972	2.0000D-1	-43.98449
	10^{-2}	2.010331	0.983628	2.0000D-2	-43.99980
	10^{-3}	2.001092	0.998292	2.0000D-3	-44.00000
	10^{-4}	2.000101	0.999826	2.0000D-4	-44.00000
	10^{-5}	1.999979	1.000001	2.0000D-5	-44.00000
(2.22)	10^0	2.064759	0.878275	1.5584D-1	-43.98525
	10^{-1}	2.010999	0.982454	1.3015D-2	-43.99984
	10^{-2}	2.001177	0.998149	1.1221D-3	-44.00000
	10^{-3}	2.000122	0.999821	9.8734D-5	-44.00000
	10^{-4}	1.999868	1.000016	8.8213D-6	-44.00000
	10^{-5}	1.999539	0.999981	7.9757D-7	-44.00000

Table 2.3(b). Rosen-Suzuki

B.F.	r^k	x_1^k	x_2^k	x_3^k	$f(\underline{x}^k)$
(1.9)	10^0	19.86704	10.89100	15.12741	-3273.147
	10^{-1}	19.95755	10.96494	15.04117	-3291.509
	10^{-2}	19.98653	10.98885	15.01311	-3297.315
	10^{-3}	19.99573	10.99647	15.00415	-3299.151
(1.8)	10^0	19.98190	10.98579	15.01690	-3297.006
	10^{-1}	19.99818	10.99875	15.00170	-3299.700
	10^{-2}	19.99982	10.99987	15.00017	-3299.970
	10^{-3}	19.99998	10.99999	15.00002	-3299.997
(2.22)	10^0	19.99892	10.99927	15.00101	-3299.826
	10^{-1}	19.99990	10.99993	15.00008	-3299.985
	10^{-2}	19.99999	10.99999	15.00000	-3299.999
	10^{-3}	20.00000	11.00000	15.00000	-3300.000

Table 2.4(a). Post Office parcel

B.F.	r^k	$-r^k \frac{d\phi_2}{dg_2}$	$-r^k \frac{d\phi_4}{dg_4}$	$-r^k \frac{d\phi_8}{dg_8}$	E_k	$f(x^k) - E_k$
(1.9)	10^0	56.56766	84.17029	108.1877	2.7097D 1	-3300.244
	10^{-1}	55.50888	81.35206	109.4169	8.5161D 0	-3300.026
	10^{-2}	55.16232	80.43111	109.8145	2.6875D 0	-3300.003
	10^{-3}	55.05148	80.13667	109.9413	8.4929D-1	-3300.000
	10^{-4}					
(1.8)	10^0	55.25840	80.57489	109.8051	3.0000D 0	-3300.006
	10^{-1}	55.02603	80.05791	109.9804	3.0000D-1	-3300.000
	10^{-2}	55.00262	80.00575	109.9979	3.0000D-2	-3300.000
	10^{-3}	55.00021	80.00055	109.9998	3.0000D-3	-3300.000
	10^{-4}					
(2.22)	10^0	55.01869	80.03869	109.9910	1.7438D-1	-3300.000
	10^{-1}	55.00175	80.00362	109.9993	1.5276D-2	-3300.000
	10^{-2}	54.99999	80.00084	109.9993	1.3600D-3	-3300.000
	10^{-3}	55.00069	79.99893	109.9998	1.2262D-4	-3300.000
	10^{-4}					

Table 2.4(b). Post Office parcel

B.F.	r^k	x_1^k	x_2^k	x_3^k	x_4^k	x_5^k	$f(\tilde{x}^k)$
(1.9)	10^0	0.176354	0.257550	0.288939	0.569868	0.427768	-26.26803
	10^{-1}	0.234806	0.307222	0.346302	0.503774	0.309355	-29.88951
	10^{-2}	0.277309	0.324216	0.381580	0.457650	0.256897	-31.56492
	10^{-3}	0.292670	0.330016	0.394059	0.439024	0.236670	-32.10196
	10^{-4}	0.297667	0.332229	0.398109	0.432030	0.228584	-32.27078
	10^{-5}	0.299261	0.333050	0.399401	0.429542	0.225531	-32.32405
(1.8)	10^0	0.242829	0.250176	0.356302	0.523538	0.443040	-29.18271
	10^{-1}	0.291231	0.317514	0.393635	0.455342	0.277745	-31.97916
	10^{-2}	0.299067	0.330521	0.399324	0.433629	0.234649	-32.31008
	10^{-3}	0.299906	0.333075	0.399932	0.429028	0.225414	-32.34471
	10^{-4}	0.299991	0.333426	0.399993	0.428386	0.224118	-32.34828
	10^{-5}	0.299999	0.333463	0.399999	0.428318	0.223980	-32.34864
(2.22)	10^0	0.293805	0.318774	0.395646	0.452151	0.274285	-32.07532
	10^{-1}	0.299448	0.331051	0.399605	0.432516	0.232796	-32.32401
	10^{-2}	0.299951	0.333201	0.399965	0.428778	0.224952	-32.34647
	10^{-3}	0.299996	0.333444	0.399997	0.428352	0.244053	-32.34848
	10^{-4}	0.300000	0.333465	0.400000	0.428314	0.223973	-32.34866
	10^{-5}	0.300000	0.333467	0.400000	0.428310	0.223966	-32.34868

Table 2.5(a). Colville: Problem 1

B.F.	r^k	$-r^k \frac{d\phi_3}{dg_3}$	$-r^k \frac{d\phi_5}{dg_5}$	$-r^k \frac{d\phi_6}{dg_6}$	$-r^k \frac{d\phi_9}{dg_9}$	E_k	$f(\tilde{x}^k) - E_k$
(1.9)	10^0	22.53835	4.372517	25.78012	1.055988	1.2943D 1	-39.21157
	10^{-1}	6.854993	3.087339	14.03623	0.493765	2.7905D 0	-32.68005
	10^{-2}	5.515711	3.052425	12.48845	0.290406	8.1685D-1	-32.38177
	10^{-3}	5.271903	3.058298	12.06361	0.181026	2.5120D-1	-32.35316
	10^{-4}	5.204573	3.060346	11.91632	0.132259	7.8464D-2	-32.34924
	10^{-5}	5.183713	3.060891	11.86489	0.113512	2.4690D-2	-32.34874
(1.8)	10^0	8.872962	3.136410	16.54220	0.891711	4.0000D 0	-33.18271
	10^{-1}	5.567500	3.052257	12.62271	0.315199	4.0000D-1	-32.37916
	10^{-2}	5.219729	3.060335	11.97170	0.151871	4.0000D-2	-32.35008
	10^{-3}	5.179054	3.061075	11.85667	0.110697	4.0000D-3	-32.34871
	10^{-4}	5.174563	3.061073	11.84135	0.104619	4.0000D-4	-32.34868
	10^{-5}	5.174136	3.061098	11.83975	0.103968	4.0000D-5	-32.34868
(2.22)	10^0	5.511443	3.047380	12.54530	0.298947	2.9716D-1	-32.37248
	10^{-1}	5.211455	3.059749	11.94775	0.142352	2.5542D-2	-32.34955
	10^{-2}	5.177846	3.060978	11.85135	0.108261	2.2174D-3	-32.34869
	10^{-3}	5.174394	3.061022	11.84060	0.104277	1.9526D-4	-32.34868
	10^{-4}	5.174121	3.061295	11.83711	0.103925	1.7451D-5	-32.34868
	10^{-5}	5.174265	3.061019	11.83969	0.103901	1.5784D-6	-32.34868

Table 2.5(b). Colville: Problem 1

The convergence property of Theorem 2.6 has also been tested numerically using the Kuhn-Tucker problem and Colville: Problem 1. The results, obtained with parameter values of $\gamma = 1$, $G_i = 10$ and $\sigma_i^{(j)} = 2$ for all i and j , are displayed in Tables 2.6 and 2.7.

j	$x_1^{(j)}$	$x_2^{(j)}$	$f(\tilde{x}^{(j)})$
1	0.61714	0.0281	-0.61714
2	0.93776	1.2E-4	-0.93776
3	0.98658	1.2E-6	-0.98658
4	0.99666	1.9E-8	-0.99666
5	0.99912	3.4E-10	-0.99912
6	0.99976	7.1E-12	-0.99976

Table 2.6. Kuhn-Tucker (using (2.18)).

j	$x_1^{(j)}$	$x_2^{(j)}$	$x_3^{(j)}$	$x_4^{(j)}$	$x_5^{(j)}$	$f(\tilde{x}^{(j)})$
1	0.29345	0.31772	0.39533	0.45359	0.27766	-32.0521
2	0.29869	0.32836	0.39907	0.43693	0.24236	-32.2893
3	0.29967	0.33173	0.39976	0.43130	0.23036	-32.3332
4	0.29991	0.33292	0.39993	0.42926	0.22600	-32.3443
5	0.29997	0.33330	0.39998	0.42859	0.22457	-32.3474
6	0.29999	0.33342	0.39999	0.42839	0.22414	-32.3483
7	0.30000	0.33345	0.40000	0.42833	0.22402	-32.3486

Table 2.7. Colville: Problem 1 (using (2.18)).

It should be noted that the unconstrained minimizations were started from the minimum of the previous member of the family (2.18). In general, the minimizations become progressively more difficult for higher order members of the family.

2.5 Hessian Conditioning and Computational Problems.

A number of serious computational difficulties arise in the application of barrier function transformations. These difficulties are most evident for small values of the controlling parameter, r^k . It should be noted that the use of small r^k -values is necessary when high accuracy is sought since it effectively allows $\tilde{x}^k$ to approach closer to the active constraints, and thus closer to the required solution x^* .

The overall efficiency of transformation methods for Problem 1.4 depends, to a great extent, on the efficiency of the methods for unconstrained minimization. The best methods currently available are Newton's method and the quasi-Newton methods (see Chapter 5), the latter being based, theoretically, on properties of the quadratic form with symmetric positive definite Hessian matrix. For general functions, if the Hessian matrix is ill-conditioned, obvious difficulties arise in the use of Newton's method. For quasi-Newton methods, where an approximation of the inverse Hessian at the minimum is constructed, it can be seen that ill-conditioning of the Hessian implies that at points close to the minimum, the approximation can be far from its true value. With both methods, Hessian ill-conditioning could lead to computational inefficiencies in

unconstrained minimization. It has been shown by Murray (1969b,1969c) (and also Fletcher and McCann (1969) and Lootsma (1969,1970)) that ill-conditioning of the Hessian matrix of transformation function (2.1) can occur for small r^k -values if the solution, $\tilde{x}^*$, is constrained (i.e. $B(\tilde{x}^*) \neq \emptyset$). From (2.1) and (2.2)

$$\nabla^2 T(\tilde{x}^k, r^k) = \nabla^2 L(\tilde{x}^k, \tilde{u}^k) + \sum_{i=1}^{\ell} r^k \frac{d^2 \phi_i(g_i(\tilde{x}^k))}{dg_i^2} \nabla g_i(\tilde{x}^k) \nabla g_i(\tilde{x}^k)^T, \quad (2.23)$$

where $\nabla^2 L(\tilde{x}^k, \tilde{u}^k)$ is given by (1.12) (noting that $m=0$) and $\tilde{u}^k$ is given by (2.3).

Theorem 2.7. Fiacco and McCormick (1968).

If $B_1(\tilde{x}^*) = B(\tilde{x}^*)$ (i.e. the strict complementarity condition (2.13) is satisfied) and the sufficiency conditions of Theorem 1.3 are satisfied then $\nabla^2 T(\tilde{x}^k, r^k)$ is positive definite for k sufficiently large.

Proof. Fiacco and McCormick (1968), p. 75.

Corollary 1. The minimum eigenvalue of $\nabla^2 T(\tilde{x}^k, r^k)$ is positive and bounded away from zero for k sufficiently large.

Remark 2.7. The conditions of Theorem 2.7 are sufficient but not necessary for the positive definiteness

of $\nabla^2 T(\tilde{x}^k, r^k)$ and the positivity of the sequence of its minimum eigenvalues.

Theorem 2.8.

If $u_i^* > 0$, then for the barrier functions (1.9), (1.8) and (2.22), $r^k \frac{d^2 \phi_i(g_i(\tilde{x}^k))}{dg_i^2}$ can be approximated, for k sufficiently large, by $\frac{\gamma_i(r^k)}{||\tilde{x}^k - \tilde{x}^*||}$ where $\gamma_i(r^k) > 0$ is bounded.

Proof. (i) If $\phi_i(g_i(\tilde{x})) = \frac{1}{g_i(\tilde{x})}$,

$$r^k \frac{d^2 \phi_i(g_i(\tilde{x}^k))}{dg_i^2} = \frac{2 r^k}{g_i(\tilde{x}^k)^3} = \frac{2 u_i^k}{g_i(\tilde{x}^k)}.$$

(ii) If $\phi_i(g_i(\tilde{x})) = -\log(g_i(\tilde{x}))$,

$$r^k \frac{d^2 \phi_i(g_i(\tilde{x}^k))}{dg_i^2} = \frac{r^k}{g_i(\tilde{x}^k)^2} = \frac{u_i^k}{g_i(\tilde{x}^k)}.$$

(iii) If $\phi_i(g_i(\tilde{x})) = \log(G - \log(g_i(\tilde{x})))$,

$$\begin{aligned} r^k \frac{d^2 \phi_i(g_i(\tilde{x}^k))}{dg_i^2} &= \frac{r^k [G - 1 - \log(g_i(\tilde{x}^k))]}{g_i(\tilde{x}^k)^2 [G - \log(g_i(\tilde{x}^k))]^2} \\ &= \frac{u_i^k}{g_i(\tilde{x}^k)} \frac{[G - 1 - \log(g_i(\tilde{x}^k))]}{[G - \log(g_i(\tilde{x}^k))]} . \end{aligned}$$

In each case, the result follows using (2.14) since $u_i^* > 0$. Note that the u_i^* need not be unique if $\nabla g_i(\tilde{x}^*)$, $i \in B(\tilde{x}^*)$ are linearly dependent.

Corollary 1. If $B_1(\tilde{x}^*) \neq \emptyset$ then the maximum eigenvalue of $\nabla^2 T(\tilde{x}^k, r^k)$, for k sufficiently large, is given by $\frac{\hat{\gamma}(r^k)}{||\tilde{x}^k - \tilde{x}^*||}$ where $\hat{\gamma}(r^k) > 0$ is bounded.

Proof. The result follows by observing that

$r^k \frac{d^2 \phi_i(g_i(\tilde{x}^k))}{dg_i^2}$ is the scalar multiplier of a rank-one matrix in the expansion of $\nabla^2 T(\tilde{x}^k, r^k)$ given by (2.23).

Using the corollaries of Theorems 2.7 and 2.8, the condition number of $\nabla^2 T(\tilde{x}^k, r^k)$, given by the ratio of maximum to minimum eigenvalues, can be written as $\frac{\gamma(r^k)}{||\tilde{x}^k - \tilde{x}^*||}$ where $\gamma(r^k) > 0$ is bounded. These results imply that each barrier function will generate similar ill-conditioning problems as the active constraints are approached.

A second computational difficulty arises from the behaviour of $\Phi(g(\tilde{x}))$ as the active constraints are approached. In this case, the transformation function can be badly scaled near $\tilde{x}^*$. Powell (1967) has commented on this situation and has suggested that the unconstrained minimization methods may have difficulty converging. This has been experienced in computations and, in particular, when r^k is small, very small steps are taken along search directions, but gradient magnitudes remain relatively large compared with $||\tilde{x} - \tilde{x}^k||$.

It has been shown by Bard (1968) that, under these conditions, a numerical instability can arise in the popular Davidon-Fletcher-Powell method for unconstrained minimization. The relatively large gradient magnitudes also provide difficulties in choosing convergence criteria for the unconstrained minimization methods which are valid for all r^k -values.

It is important to note, in conclusion, that the computational difficulties discussed here are all due to the barrier functions corresponding to active inequality constraints. The effects of inactive constraints can be ignored as r^k decreases. This point will be considered again in Chapter 4.

CHAPTER 3

PENALTY FUNCTION METHODS3.1 Introduction.

Two different penalty function transformations for the solution of Problem 1.3 are examined in this chapter. A solution of the problem is assumed to exist and is denoted by $\tilde{x}^*$.

Initially, we consider the transformation

$$T(\tilde{x}, X^k) = (f(\tilde{x}) - X^k)^2 + \sum_{i=1}^m h_i(\tilde{x})^2, \quad (3.1)$$

where X^k is a controlling parameter which is iteratively adjusted so that $\tilde{x}^*$ is given by the unconstrained minimum of $T(\tilde{x}, X^k)$ for some X^k . The minimum of $T(\tilde{x}, X^k)$ will be denoted by $\tilde{x}^k$ and is assumed to be well defined. Two methods based on (3.1) and proposed by Schmit and Fox (1965) and Morrison (1968), will be examined in §3.2. Then an improved generating relationship for the X^k -sequence is developed in §3.3. A new computational algorithm based on (3.1) and the results of §3.3 is described and evaluated numerically in §3.4 and §3.5.

The remainder of the chapter is devoted to an examination of Powell's method (Powell (1967)) which, under certain circumstances, exhibits excellent convergence properties. It is based on the transformation

$$T(\underline{x}, \underline{\theta}^K, \underline{\sigma}^K) = f(\underline{x}) + (\underline{\theta}^K + \underline{h}(\underline{x}))^T S_K (\underline{\theta}^K + \underline{h}(\underline{x})), \quad (3.2)$$

where $S_K = \text{diag}(\sigma_i^K)$ and the components of the m -dimensional vectors $\underline{\theta}^K$ and $\underline{\sigma}^K > 0$ are controlling parameters which are iteratively adjusted during the procedure.

Remark 3.1. It can readily be seen that, for $\underline{\theta}^K = 0$ and the components of $\underline{\sigma}^K$ given by $[r^k]^{-1}$, (3.2) reduces to the penalty function transformation for equality constraints discussed by Fiacco and McCormick (1968).

3.2 The Schmit-Fox Method and Morrison's Method.

The Schmit-Fox method is based on the behaviour of (3.1) when X^k satisfies

$$f(\underline{x}^*) \leq X^k \leq Y, \quad (3.3)$$

where $Y = \max_{\underline{x} \in R} f(\underline{x})$. Starting with X^k satisfying (3.3),

$T(\underline{x}, X^k)$ is minimized for a decreasing sequence X^k , $k = 1, 2, \dots$, commencing the k th unconstrained minimization from $\underline{x}^{k-1}$ with $\underline{x}^0$ chosen as some initial estimate of $\underline{x}^*$. Initiating the procedure presents little difficulty, in theory, if $f(\underline{x})$ is unbounded on R . If this is not the case, a suitable starting value can be obtained by setting $X^1 = f(\underline{x}^0)$, where $\underline{x}^0 \in R$. The sequence of minimizations is continued until, for some k , $T(\underline{x}^k, X^k) > 0$ when the following theorem implies that $X^k < f(\underline{x}^*)$.

Theorem 3.1.

If R is connected, then $T(\underline{x}^k, X^k) > 0$ for some k iff $X^k < f(\underline{x}^*)$ or $X^k > Y$.

Proof. If $X^k < f(\underline{x}^*)$, then $T(\underline{x}^k, X^k) > 0$ since otherwise $T(\underline{x}^k, X^k) = 0$ implies $\underline{x}^k \in R$ and $f(\underline{x}^k) = X^k < f(\underline{x}^*)$, which contradicts the definition of $\underline{x}^*$. If $X^k > Y$ and $T(\underline{x}^k, X^k) = 0$, then a similar contradiction of the definition of Y can be deduced.

If $T(\underline{x}^k, X^k) > 0$, assume $f(\underline{x}^*) \leq X^k \leq Y$.

Then, by the continuity of the problem functions and the connectedness of R , there exists $\hat{\underline{x}} \in R$ s.t. $f(\hat{\underline{x}}) = X^k$ and $0 = T(\hat{\underline{x}}, X^k) < T(\underline{x}^k, X^k)$. This contradicts the definition of $\underline{x}^k$ and thus the theorem is proved.

Since $\tilde{x}^*$ is assumed to exist, the Schmit-Fox method will generate, for k sufficiently large, an $\tilde{x}^{k-1}$ and $\tilde{x}^k$ satisfying

$$\tilde{x}^k < f(\tilde{x}^*) \leq \tilde{x}^{k-1} . \quad (3.4)$$

Using Theorem 3.1 to ensure that (3.4) is satisfied, the interval $(\tilde{x}^k, \tilde{x}^{k-1})$ can be refined by a further minimization with $\tilde{x}^{k+1} \in (\tilde{x}^{k-1}, \tilde{x}^k)$. This refinement can be continued to any desired precision.

In contrast to the Schmit-Fox method, Morrison's method depends on the behaviour of (3.1) for $\tilde{x}^k \leq f(\tilde{x}^*)$. The following theorems describe the main results used in establishing the method. Connectedness is not required.

Theorem 3.2. (Morrison (1968)).

If $\tilde{x}^k \leq f(\tilde{x}^*)$, then $f(\tilde{x}^k) \leq f(\tilde{x}^*)$.

Proof. By definition,

$$T(\tilde{x}^*, \tilde{x}^k) \geq T(\tilde{x}^k, \tilde{x}^k) ,$$

and thus

$$(f(\tilde{x}^*) - \tilde{x}^k)^2 \geq (f(\tilde{x}^k) - \tilde{x}^k)^2 . \quad (3.5)$$

If $f(\tilde{x}^k) \geq \tilde{x}^k$, then from (3.5), $f(\tilde{x}^*) \geq f(\tilde{x}^k)$.

If $f(\tilde{x}^k) < \tilde{x}^k$, then $f(\tilde{x}^k) < \tilde{x}^k \leq f(\tilde{x}^*)$.

Theorem 3.3. (Morrison (1968)).

The function, $f(\tilde{x}^k)$, is a monotonic nondecreasing function of X^k .

Proof. Suppose $X^k \leq X^{k+1}$. Then

$$T(\tilde{x}^k, X^k) \leq T(\tilde{x}^{k+1}, X^k) \quad (3.6)$$

$$T(\tilde{x}^{k+1}, X^{k+1}) \leq T(\tilde{x}^k, X^{k+1}). \quad (3.7)$$

Adding (3.6) and (3.7) gives

$$f(\tilde{x}^k)(X^{k+1} - X^k) \leq f(\tilde{x}^{k+1})(X^{k+1} - X^k),$$

and consequently, the result.

Theorems 3.2 and 3.3 suggest that under suitable conditions, $f(\tilde{x}^k)$ may converge to $f(\tilde{x}^*)$ if an increasing sequence X^k , $k = 1, 2, \dots$, is chosen so that $X^k \leq f(\tilde{x}^*)$ for all k . It is now shown that with additional mild assumptions about the problem functions, this is true. Morrison's method for generating the X^k sequence is given by

$$X_M^{k+1} = X^k + T(\tilde{x}^k, X^k)^{\frac{1}{2}}, \quad (3.8)$$

where X_M^{k+1} is called the Morrison parameter.

Theorem 3.4. (Morrison (1968)).

If $x^k \leq f(\tilde{x}^*)$ then $x_M^{k+1} \leq f(\tilde{x}^*)$.

Proof. $T(\tilde{x}^k, x^k) \leq T(\tilde{x}^*, x^k)$
 $= (f(\tilde{x}^*) - x^k)^2$.

But $f(\tilde{x}^*) - x^k \geq 0$ and it follows that

$$x_M^{k+1} = x^k + T(\tilde{x}^k, x^k)^{\frac{1}{2}} \leq f(\tilde{x}^*)$$

To establish convergence of Morrison's method, the notion of constraint perturbations is introduced. Consider the problem:

$$\begin{aligned} &\text{minimize } f(\tilde{x}) \\ &\text{subject to } h_i(\tilde{x}) - \theta_i = 0, \quad i = \overline{1, m}, \end{aligned}$$

where θ is a perturbation of the constraints of Problem 1.3. A solution of the perturbed problem is denoted by $\tilde{x}(\theta)$, and therefore, $\tilde{x}^* = \tilde{x}(0)$.

Definition 3.1. The function, f , is a continuous function of constraint perturbations, θ , at $\theta = 0$ if for every $\epsilon > 0$, there exists $\delta > 0$ s.t. $\|\theta\| < \delta$ implies $|f(\tilde{x}(\theta)) - f(\tilde{x}^*)| < \epsilon$.

Theorem 3.5. (Morrison (1968)).

The parameter sequence $\tilde{x}^k$, $k=1,2,\dots$, generated by (3.8) with $\tilde{x}^0 \leq f(\tilde{x}^*)$, converges. If f is a continuous function of the constraint perturbations, $\tilde{\theta}$, at $\tilde{\theta} = 0$, then for any $\epsilon_1 > 0$ and $\epsilon_2 > 0$, k can be chosen sufficiently large to satisfy

$$f(\tilde{x}^k) - f(\tilde{x}^*) < \epsilon_1 \text{ and } ||\tilde{h}(\tilde{x}^k)|| = \left[\sum_{i=1}^m h_i(\tilde{x}^k)^2 \right]^{\frac{1}{2}} < \epsilon_2.$$

Proof. By (3.8) and Theorem 3.4, $\{\tilde{x}^k\}$ is non-decreasing, bounded above and therefore, convergent. Using the continuity assumptions, δ is chosen so that for all $\tilde{\theta}$ satisfying $||\tilde{\theta}|| < \delta$,

$$|f(\tilde{x}(\tilde{\theta})) - f(\tilde{x}^*)| < \epsilon_1.$$

Now, choose k sufficiently large so that $\tilde{x}^{k+1} - \tilde{x}^k < \min(\epsilon_2, \delta)$. Then, by (3.8),

$$T(\tilde{x}^k, \tilde{x}^k) < \min(\epsilon_2^2, \delta^2),$$

and thus,

$$||\tilde{h}(\tilde{x}^k)||^2 = \sum_{i=1}^m h_i(\tilde{x}^k)^2 < \min(\epsilon_2^2, \delta^2). \quad (3.9)$$

Since by (3.9), $||\tilde{h}(\tilde{x}^k)|| < \delta$, it follows that

$$|f(\tilde{x}(\tilde{h}(\tilde{x}^k))) - f(\tilde{x}^*)| < \epsilon_1. \quad (3.10)$$

The definition of $\tilde{x}(h(\tilde{x}^k))$ and Theorem 3.2 imply

$$f(\tilde{x}(h(\tilde{x}^k))) \leq f(\tilde{x}^k) \leq f(\tilde{x}^*)$$

and, from (3.10), it follows that

$$0 \leq f(\tilde{x}^*) - f(\tilde{x}^k) < \varepsilon_1. \quad (3.11)$$

The theorem now follows from (3.9) and (3.11).

In terms of the problem functions, sufficient conditions for f to be a continuous function of the constraint perturbations, $\tilde{\theta}$, at $\tilde{\theta} = 0$ are given by Theorem 3.6. A rigorous discussion of constraint perturbation and its effect on the problem solution in the case of inequality constraints can be found in Evans and Gould (1970).

Theorem 3.6.

If the sufficiency conditions of Theorem 1.3 are satisfied at $\tilde{x}^*$ and $\nabla h_i(\tilde{x}^*)$, $i=\overline{1,m}$, are linearly independent, then f is a continuous function of constraint perturbations, $\tilde{\theta}$, at $\tilde{\theta} = 0$.

Proof. Using Theorem 1.1 and Remark 1.2(2), the assumption of linear independence implies that there exists a unique GLM vector, $\tilde{v}^*$, satisfying the

Kuhn-Tucker necessary conditions (1.14) and (1.17) at $\underline{x}^*$. These $(n+m)$ equations in $(n+2m)$ variables may be written

$$\nabla f(\underline{x}) + \sum_{i=1}^m v_i \nabla h_i(\underline{x}) = 0, \quad (3.12)$$

$$h_i(\underline{x}) - \theta_i = 0, \quad i=\overline{1, m},$$

and are satisfied at $(\underline{x}, \underline{v}, \underline{\theta}) = (\underline{x}^*, \underline{v}^*, 0)$. With respect to $(\underline{x}, \underline{v})$, the Jacobian of this system, $J(\underline{x}, \underline{v})$, is given by

$$J(\underline{x}, \underline{v}) = \begin{bmatrix} \nabla^2 L(\underline{x}, \underline{v}) & \Lambda(\underline{x}) \\ \Lambda(\underline{x})^T & 0 \end{bmatrix},$$

where $\Lambda(\underline{x}) = [\nabla h_1(\underline{x}), \dots, \nabla h_m(\underline{x})]$. It is now shown that there is no nonzero solution to the equations

$$J(\underline{x}^*, \underline{v}^*) \underline{z} = \begin{bmatrix} \nabla^2 L(\underline{x}^*, \underline{v}^*) \underline{\alpha} + \Lambda(\underline{x}^*) \underline{\beta} \\ \Lambda(\underline{x}^*)^T \underline{\alpha} \end{bmatrix} = 0, \quad (3.13)$$

where $\underline{z} = \begin{bmatrix} \underline{\alpha} \\ \underline{\beta} \end{bmatrix}$ and $\underline{\alpha}$ and $\underline{\beta}$ are $(n \times 1)$ and $(m \times 1)$ vectors, respectively. If $\underline{\alpha} = 0$, then for any $\underline{\beta} \neq 0$, $\Lambda(\underline{x}^*) \underline{\beta} = 0$. This contradicts the linear independence assumption and thus, $\underline{\beta} = 0$. If $\underline{\alpha} \neq 0$, then by (3.13),

$$\Lambda(\underline{x}^*)^T \underline{\alpha} = 0 \quad (3.14)$$

and premultiplying (3.13) by $\underline{z}^T$ yields

$$\underline{\alpha}^T \nabla^2 L(\underline{x}^*, \underline{v}^*) \underline{\alpha} = 0. \quad (3.15)$$

The contradiction follows from (1.19) since (3.14) is just the condition that ensures $\alpha \in C_1(\tilde{x}^*) = C(\tilde{x}^*)$. Since the inverse of $J(\tilde{x}^*, \tilde{v}^*)$ exists, the implicit function theorem (Hestenes (1966), p.22) is now applied to (3.12) at $(\tilde{x}^*, \tilde{v}^*, 0)$ and the existence of a unique differentiable function $(\tilde{x}, \tilde{v}) = (\tilde{X}(\tilde{\theta}), \tilde{V}(\tilde{\theta}))$ is deduced in a neighbourhood of $\tilde{\theta} = 0$. The required result follows using the continuity of $f(\tilde{x})$.

3.3 The Tangent Parameter.

The rate of convergence of Morrison's method, discussed in §3.2, depends on the efficiency of the Morrison parameter defined by (3.8). Numerical observations have suggested that, in some cases, convergence of X^k to $f(\tilde{x}^*)$ is rather slow. Motivated by this evidence and also by a suggestion by P. Wolfe, a referee of Morrison (1968), a new method for generating the X^k -sequence has been developed by Kowalik, Osborne and Ryan (1969).

Definition 3.2. The tangent parameter, X_T^{k+1} , is given by

$$X_T^{k+1} = X^k + \frac{T(\tilde{x}^k, X^k)}{[T(\tilde{x}^k, X^k) - \sum_{i=1}^m h_i(\tilde{x}^k)^2]^{\frac{1}{2}}} . \quad (3.16)$$

Remark 3.2. The term "tangent parameter" results from the geometric interpretation of (3.16) (Morrison (1968)).

Theorem 3.7.

The inequality, $x_T^k \geq x_M^k$, holds for all k .

Proof. Since $T(\tilde{x}^k, x^k) \geq 0$,

$$\begin{aligned} x_T^{k+1} &= x^k + T(\tilde{x}^k, x^k)^{\frac{1}{2}} \left[\frac{T(\tilde{x}^k, x^k)}{T(\tilde{x}^k, x^k) - \sum_{i=1}^m h_i (\tilde{x}^k)^2} \right]^{\frac{1}{2}} \\ &\geq x^k + T(\tilde{x}^k, x^k)^{\frac{1}{2}} \\ &= x_M^{k+1} . \end{aligned}$$

Theorem 3.7 suggests that the parameter sequence $x^k = x_T^k$, $k=1,2,\dots$, with $x^0 \leq f(\tilde{x}^*)$ could possibly provide more rapid convergence than the Morrison parameter sequence provided

$$x^k \leq f(\tilde{x}^*) \Rightarrow x_T^{k+1} \leq f(\tilde{x}^*) . \quad (3.17)$$

In order to establish (3.17), which is equivalent to a similar result for the Morrison parameter (see Theorem 3.4), a number of additional conditions are imposed on the problem functions.

Lemma 3.1. (Kowalik, Osborne and Ryan (1969)).

- If (i) $f(\underline{x}), h_i(\underline{x}), i=\overline{1,m}$, are continuously differentiable functions,
- (ii) $\underline{x}^k, k=1,2,\dots$, is an increasing sequence converging to $f(\underline{x}^*)$ from below such that $\underline{x}^k$ converges to $\underline{x}^*$, and
- (iii) $\nabla h_i(\underline{x}^k), i=\overline{1,m}$, are linearly independent for k sufficiently large,

then $\lim_{k \rightarrow \infty} v_i^k = v_i^*, i=\overline{1,m}$, where

$$v_i^k = \frac{h_i(\underline{x}^k)}{f(\underline{x}^k) - \underline{x}^k}, \quad i=\overline{1,m}, \quad (3.18)$$

and $\underline{v}^*$ is the GLM vector, existence and uniqueness of which is guaranteed by the continuity and linear independence assumptions.

Proof. From (1.17) (noting that $\ell=0$) it follows that

$$\nabla L(\underline{x}^*, \underline{v}^*) = \nabla f(\underline{x}^*) + \sum_{i=1}^m v_i^* \nabla h_i(\underline{x}^*) = 0, \quad (3.19)$$

and at $\underline{x}^k$, (3.1) yields

$$\frac{1}{2} \nabla T(\underline{x}^k, \underline{x}^k) = (f(\underline{x}^k) - \underline{x}^k) \nabla f(\underline{x}^k) + \sum_{i=1}^m h_i(\underline{x}^k) \nabla h_i(\underline{x}^k) = 0. \quad (3.20)$$

Using (3.19) and (3.20)

$$\sum_{i=1}^m \left\{ v_i^* - \frac{h_i(\tilde{x}^k)}{f(\tilde{x}^k) - x^k} \right\} \nabla h_i(\tilde{x}^k) = \nabla f(\tilde{x}^k) - \nabla f(x^*) + \sum_{i=1}^m v_i^* (\nabla h_i(\tilde{x}^k) - \nabla h_i(x^*)). \quad (3.21)$$

Since the right hand side of (3.21) tends to zero as $\tilde{x}^k \rightarrow x^*$, the result follows from the linear independence of $\nabla h_i(\tilde{x}^k)$, $i = \overline{1, m}$.

Remark 3.3. Under the conditions of Lemma 3.1, (3.21) ensures that $f(\tilde{x}^k) - x^k \neq 0$ for $x^k < f(x^*)$.

Theorem 3.8. (Kowalik, Osborne and Ryan (1969)).

If the conditions of Lemma 3.1 hold and if there exists a set $\Omega \subset E^n \times E^m$ where $(\tilde{x}^*, \tilde{v}^*) \in \Omega_0$ and s.t:

$$(i) \quad L(\tilde{x}^*, \tilde{v}) \leq L(\tilde{x}^*, \tilde{v}^*) \leq L(\tilde{x}, \tilde{v}^*) \quad (3.22)$$

for $(\tilde{x}, \tilde{v}) \in \Omega$ (actually $L(\tilde{x}^*, \tilde{v}) = L(\tilde{x}^*, \tilde{v}^*)$

for Problem 1.3), and

(ii) for $\hat{v}$ s.t. $(\tilde{x}, \hat{v}) \in \Omega$, there exists $\hat{x}$ s.t.

$(\hat{x}, \hat{v}) \in \Omega$, $L(\hat{x}, \hat{v}) \leq L(\tilde{x}, \hat{v})$ for all $\tilde{x}$ s.t.

$(\tilde{x}, \hat{v}) \in \Omega$ and $(\hat{x}, \hat{v})$ is the unique solution

in Ω to the system of equations

$$\nabla_x L(\tilde{x}, \hat{v}) = 0, \quad (3.23)$$

then $x_T^{k+1} \leq f(\tilde{x}^*)$ provided x^k is sufficiently close to $f(\tilde{x}^*)$.

Proof. By Lemma 3.1, it follows that if x^k is sufficiently close to $f(x^*)$, then $(\tilde{x}^k, \tilde{y}^k) \in \Omega$ where $\tilde{y}^k$ is given by (3.18). From (3.20) we have $\nabla L(\tilde{x}^k, \tilde{y}^k) = 0$, and therefore, by the second condition,

$$L(\tilde{x}^k, \tilde{y}^k) \leq L(\tilde{x}^*, \tilde{y}^k) . \quad (3.24)$$

By the first condition, using (3.24),

$$L(\tilde{x}^k, \tilde{y}^k) \leq L(\tilde{x}^*, \tilde{y}^*) = f(\tilde{x}^*)$$

and the required result follows since

$$\begin{aligned} L(\tilde{x}^k, \tilde{y}^k) &= f(\tilde{x}^k) + \sum_{i=1}^m \frac{h_i(\tilde{x}^k)}{f(\tilde{x}^k) - x^k} h_i(\tilde{x}^k) \\ &= x^k + \frac{(f(\tilde{x}^k) - x^k)^2 + \sum_{i=1}^m h_i(\tilde{x}^k)^2}{f(\tilde{x}^k) - x^k} \\ &= x_T^{k+1} . \end{aligned}$$

Remarks 3.4. 1. If the Lagrangian satisfies condition (i) of Theorem 3.8, it is said to have the saddle point property at (x^*, y^*) .

2. The second condition of the theorem is a regularity condition which ensures $L(\tilde{x}, \tilde{y})$ is not fluted in a sufficiently small neighbourhood of the

saddle point, $(\tilde{x}^*, \tilde{y}^*)$. This should not be a difficult condition to satisfy and therefore the tangent parameter should have reasonably general application, at least in a small neighbourhood of the solution. A sufficient condition for condition (ii) is provided when $L(\tilde{x}, \tilde{y})$ is convex in $\tilde{x}$ for $(\tilde{x}, \tilde{y}) \in \Omega$.

3. If Problem 1.3 is convex, then $L(\tilde{x}, \tilde{y})$ is convex in $\tilde{x}$ for all $\tilde{y}$. Convexity also ensures that $(\tilde{x}^*, \tilde{y}^*)$ is a saddle point of $L(\tilde{x}, \tilde{y})$.

3.4 An Algorithm.

An algorithm, based on the results of §3.2 and §3.3 is now developed and its convergence is established. The Schmit-Fox method is used in an initial phase to bracket $f(\tilde{x}^*)$ and the solution is finally obtained using the Morrison and tangent parameter sequences. A particular advantage of this procedure is that it is not necessary for X^0 in (3.1) to be a lower bound for $f(\tilde{x}^*)$. This allows more flexibility in starting the calculation. The use of the Schmit-Fox method in this manner also allows the more rapid convergence of the tangent parameter to be used whenever possible and provides a simple means for detecting its failure. A description of the algorithm follows.

Algorithm 3.1.1. Initial phase

- (i) Choose $\tilde{x}^0 \in R$; $X^0 = f(\tilde{x}^0)$; $k=1$.
- (ii) $X^k = X^{k-1} - \delta X$ where δX is an increasing increment by which X^k is decreased.
- (iii) Minimize $T(\tilde{x}, X^k)$ to find $\tilde{x}^k$.
- (iv) If $T(\tilde{x}^k, X^k) < \epsilon$, (ϵ -machine test on zero) then $\delta X = 2\delta X$, $k=k+1$ and go to 1(ii); else $BU = X^{k-1}$ and go to 2(i). (BU , BL - upper and lower bounds respectively for $f(x^*)$).

2. Iteration phase

- (i) Compute X_M^{k+1} , X_T^{k+1} ; $BL = X_M^{k+1}$; $k=k+1$.
- (ii) If $X_T^k < BU$, then $X^k = X_T^k$; else $X^k = X_M^k$.
- (iii) Minimize $T(\tilde{x}, X^k)$ to find $\tilde{x}^k$.
- (iv) If $T(\tilde{x}^k, X^k) > \epsilon$, then go to 2(i); else $BU = X^k$. If $BU - BL < \epsilon$, then stop; else $X^k = BL$ and go to 2(iii).

Remark 3.5. On return to 2(i), $X^k \leq f(\tilde{x}^*)$.

A proof of convergence of Algorithm 3.1 is based on Theorem 3.5. The tests on zero are considered to be exact.

Theorem 3.9.

- If (i) f is a continuous function of the constraint perturbations, θ , at $\theta=0$,
 (ii) $T(\tilde{x}^k, x^k)$ can be found by unconstrained minimization methods, and
 (iii) $T(\tilde{x}^k, x^k) > 0 \Rightarrow f(\tilde{x}^k) < f(x^*)$, (3.25)

then Algorithm 3.1 converges in the sense that for any $\varepsilon_1, \varepsilon_2 > 0$, k can be chosen sufficiently large to satisfy

$$|f(\tilde{x}^k) - f(x^*)| < \varepsilon_1 \text{ and } ||h(\tilde{x}^k)|| < \varepsilon_2.$$

Proof. By (3.25), the Initial phase ensures that an upper and lower bound for $f(x^*)$ can be found. The Iteration phase produces an increasing sequence $x^k, k=\hat{k}, \hat{k}+1, \dots$, ($\hat{k}$ the iteration count when the lower bound is found), which is bounded above by $f(\tilde{x}^*)$. Thus $\{x^k\}$ converges whence $\{x_M^k\}$ converges since $x^k \leq x_M^{k+1} \leq x^{k+1}$. The result follows by applying Theorem 3.5.

Convergence of $\tilde{x}^k, k=1,2,\dots$, can be established under stronger conditions.

Theorem 3.10. (Kowalik, Osborne and Ryan (1969)).

If there exists a closed bounded region, Γ , in which $T(\tilde{x}^k, x^k)$ can be found by unconstrained minimization methods with $\tilde{x}^k \in \Gamma$ and (3.25) holds,

then Algorithm 3.1 converges and limit points of $\{\tilde{x}^k\}$ are solutions of Problem 1.3.

Proof. By the argument used in Theorem 3.9, the sequence $\{X_M^k\}$ converges so that $\lim_{k \rightarrow \infty} T(\tilde{x}^k, X^k) = 0$ giving

$$\lim_{k \rightarrow \infty} ||h(\tilde{x}^k)||^2 = 0 \quad (3.26)$$

and

$$\lim_{k \rightarrow \infty} |f(\tilde{x}^k) - X^k| = 0.$$

Let $\hat{x}$ be a limit point of $\{\tilde{x}^k\}$; then there is a subsequence $\tilde{x}^{\sigma_j}$, $j=0,1,\dots$, converging to $\hat{x}$. From (3.26) and the continuity of $h(\tilde{x})$, it follows that $h(\hat{x}) = 0$ whence $f(\hat{x}) \geq f(x^*)$. However, by Theorem 3.2, $f(\tilde{x}^{\sigma_j}) \leq f(x^*)$ so that, by the continuity of f , we have $f(\hat{x}) \leq f(x^*)$ and the result follows.

Remarks 3.6. 1. Theorem 3.10 assumes that Algorithm 2.1 does not terminate in a finite number of iterations. The argument is readily adapted should this occur.

2. The conditions of Theorem 3.10, although difficult to verify, can be expected to hold in most cases.

3. Keeping an upper bound for $f(\tilde{x}^*)$ ensures that the tangent parameter is not accepted in the case $X_T^k > Y$ which also gives $T(\tilde{x}^k, X^k) > 0$. If

such a value were accepted, then the subsequent values of x^k would diverge to $+\infty$. In such circumstances,

$$-\nabla f(\tilde{x}) + \frac{1}{2X} \nabla T(\tilde{x}, 0) = 0$$

holds at a stationary point of $T(\tilde{x}, X)$. It is clear that either

- (a) $\tilde{x}$ tends to an unconstrained local maximum of $f(\tilde{x})$ and $x_M^{k+1} = 2x_M^k$, or
- (b) $f(\tilde{x}) \rightarrow +\infty$ as $X \rightarrow \infty$.

3.5 Numerical Results.

To demonstrate the procedures that form the basis of Algorithm 3.1, its application to three simple test problems is considered. Although the problems chosen are inequality constrained, use is made of a suggestion, originally due to Kelley (1962) and also Schmit and Fox (1965), that inequality constraints be transformed into equality constraints using the Heaviside function, $H(t)$, defined as

$$\begin{aligned} H(t) &= 1, & t > 0, \\ &= 0, & t \leq 0. \end{aligned}$$

Then, the inequality constraint, $g(\tilde{x}) \geq 0$, is equivalent to the equality constraint

$$h(\tilde{x}) = g(\tilde{x}) H(-g(\tilde{x})) = 0.$$

This device, although preserving continuity, introduces discontinuities in the gradients of the constraint functions at the boundary of the feasible region, and therefore, the use of unconstrained methods of minimization, based on first and second derivative information, could produce convergence difficulties at a constrained solution. For this reason, the direct search simplex method of Nelder and Mead (1965), as discussed by Kowalik and Osborne (1968), is used for the unconstrained minimizations. In the examples considered, a feasible point could be found by inspection, but use has been made of the simplex method to find a feasible point by minimizing $\sum_{i=1}^m h_i(\tilde{x})^2$. The simplex method, in this case, can be terminated as soon as a sufficiently small value has been found, often before the simplex collapses. This saves function evaluations. The following three test problems are used.

(i) Rosen-Suzuki. (Rosen and Suzuki (1965)).

See §2.4.

(ii) Beale. (Beale (1967)).

Minimize $f(\underline{x}) = 9 - 8x_1 - 6x_2 - 4x_3 + 2x_1^2 + x_3^2 + 2x_1x_2 + 2x_1x_3 + 2x_2^2$
 subject to $x_1 \geq 0, x_2 \geq 0, x_3 \geq 0$ and $x_1 + x_2 + 2x_3 \leq 3$.

The constrained minimum is at $\underline{x}^* = (\frac{4}{3}, \frac{7}{9}, \frac{4}{9})^T$ with
 $f(\underline{x}^*) = \frac{1}{9}$ and the fourth constraint active. An
 initial feasible point is $\underline{x}^0 = (\frac{1}{2}, \frac{1}{2}, \frac{1}{2})^T$ with
 $f(\underline{x}^0) = 2.25$.

(iii) Post Office parcel problem. (Rosenbrock (1960)).

See §2.4.

In applying the algorithm, the main interest centres on the performance of the tangent parameters, which are expected to control at least the final stages of the process. In fact, for each of the problems considered, the tangent parameters are used from the lower bound found by the Schmit-Fox procedure. This is illustrated in Tables 3.1(a)-3.3(a). For comparison, the Morrison parameters have also been used to determine the solution starting from the same lower bound and these results are given in Tables 3.1(b)-3.3(b). It will be seen from Table 3.3 that the advantage of the tangent parameters over the Morrison parameters can be marked.

k	$f(\tilde{x}^k)$	x_1^k	x_2^k	x_3^k	x_4^k
0	-44.810	0.05034	1.0236	2.0555	-0.9555
1	-44.029	0.00200	1.0010	2.0023	-0.9971
2	-44.000	0.00010	1.0000	1.9999	-1.0001
3	-44.010	0.00110	1.0005	2.0011	-0.9985
4	-44.009	0.00063	1.0003	2.0007	-0.9991
5	-44.005	0.00039	1.0002	2.0004	-0.9995
6	-44.003	0.00023	1.0001	2.0002	-0.9997
8	-44.001	0.00003	0.9999	2.0002	-0.9998
10	-44.000	0.00000	1.0000	2.0000	-0.9999

Table 3.1(a). Rosen-Suzuki (Tangent parameter).

k	$f(\tilde{x}^k)$	x_1^k	x_2^k	x_3^k	x_4^k
0	-44.810	0.050340	1.0236	2.0555	-0.9555
1	-44.460	0.031718	1.0147	2.0337	-0.9663
2	-44.273	0.019428	1.0090	2.0203	-0.9772
3	-44.160	0.011706	1.0054	2.0122	-0.9855
4	-44.094	0.006993	1.0033	2.0073	-0.9911
5	-44.056	0.004161	1.0019	2.0043	-0.9946
10	-44.004	0.000301	1.0001	2.0003	-0.9996
15	-44.000	0.000166	1.0001	1.9999	-1.0001
20	-44.000	0.000024	1.0002	1.9999	-1.0001
21	-44.000	0.000000	1.0001	1.9999	-1.0000

Table 3.1(b). Rosen-Suzuki (Morrison parameter).

k	$f(\tilde{x}^k)$	x_1^k	x_2^k	x_3^k	x_4^k
0	0.0423	1.2057	0.8629	0.6572	3.3829
1	0.1104	1.3322	0.7785	0.4464	3.0034
2	0.1111	1.3336	0.7779	0.4442	3.0000
3	0.1111	1.3334	0.7778	0.4444	3.0000

Table 3.2(a). Beale (Tangent parameter).

k	$f(\tilde{x}^k)$	x_1^k	x_2^k	x_3^k	x_4^k
0	0.0423	1.2057	0.8629	0.6572	3.3829
1	0.1091	1.3303	0.7798	0.4494	3.0089
2	0.1110	1.3333	0.7778	0.4446	3.0002
3	0.1111	1.3333	0.7779	0.4444	3.0001
4	0.1111	1.3333	0.7778	0.4445	3.0000

Table 3.2(b). Beale (Morrison parameter).

k	$f(\tilde{x}^k)$	x_1^k	x_2^k	x_3^k	x_4^k
0	-3305.0	20.013	11.018	14.988	72.026
1	-3300.0	20.000	11.000	15.000	72.000
2	-3300.0	20.000	11.000	15.000	72.000

Table 3.3(a). Post Office parcel (Tangent parameter).

k	$f(\tilde{x}^k)$	x_1^k	x_2^k	x_3^k	x_4^k
0	-3305.0	20.013	11.018	14.988	72.026
1	-3305.0	20.013	11.018	14.988	72.025
2	-3304.9	20.013	11.018	14.988	72.025
3	-3304.9	20.013	11.018	14.988	72.025
4	-3304.9	20.012	11.018	14.988	72.025
5	-3304.8	20.012	11.018	14.988	72.025
10	-3304.7	20.012	11.017	14.989	72.024
15	-3304.5	20.012	11.017	14.989	72.023
20	-3304.4	20.011	11.016	14.989	72.022
30	-3304.1	20.010	11.015	14.990	72.021
40	-3303.8	20.010	11.014	14.991	72.019
50	-3303.6	20.009	11.013	14.991	72.018
100	-3302.5	20.006	11.009	14.994	72.013
200	-3301.8	20.004	11.007	14.996	72.010

Table 3.3(b). Post Office parcel (Morrison parameter).

Remarks 3.6. 1. Since the condition, $T(\tilde{x}^k, X^k) < \epsilon$, is used to test for $X^k \geq f(\tilde{x}^*)$, the size of the set $\{\tilde{x} : X^k < f(\tilde{x}^*), T(\tilde{x}, X^k) < \epsilon\}$ is critical to the accuracy attained by Algorithm 3.1. It also reflects the conditioning of the problem.

2. For each k , $T(\tilde{x}^k, X^k)$ must be computed to high accuracy. This is an expensive requirement especially when using direct search unconstrained minimization methods.

3. Comparative results, obtained using an inverse barrier function transformation to solve the same problems, are reported by Kowalik, Osborne and Ryan (1970).

The general use of Algorithm 3.1 is limited by a number of factors. As mentioned in Remarks 3.6(1) and (2), the most important is the high accuracy required for $T(\tilde{x}^k, X^k)$. The use of direct search methods of minimization to avoid gradient discontinuities introduced by the inequality constraints will also restrict the application of Algorithm 3.1 to problems of small dimension.

3.6 Powell's Method.

Powell's aim in developing a method based on (3.2) was to avoid computational difficulties due to active constraints and small r^k -values as discussed in Chapter 2. The method avoids the computational difficulties by minimizing (3.2) for moderate values of the parameters, $\underline{\theta}$ and $\underline{\sigma}$, which do not cause ill-conditioning and scaling problems.

The method is based on the observation that if $\underline{x}^k$ minimizes (3.2), then $\underline{x}^k$ is also a solution of the problem

$$\begin{aligned} &\text{minimize} && f(\underline{x}) \\ &\text{subject to} && h_i(\underline{x}) = h_i(\underline{x}^k), \quad i=\overline{1,m}. \end{aligned}$$

In order to solve Problem 1.3, it is sufficient to find $\underline{\theta}^k$ and $\underline{\sigma}^k$ such that $\underline{x}^k = \underline{x}(\underline{\theta}^k, \underline{\sigma}^k)$ solves the system of equations

$$h(\underline{x}(\underline{\theta}^k, \underline{\sigma}^k)) = 0. \quad (3.27)$$

If at $\underline{x}^*$, a solution of Problem 1.3, $\nabla h_i(\underline{x}^*)$, $i=\overline{1,m}$, are linearly independent, then by Theorem 1.1 and Remark 1.2(3), there exists a unique GLM vector, $\underline{v}^*$, satisfying the Kuhn-Tucker conditions

(1.14) and (1.17) at $\tilde{x}^*$. For points, $\tilde{x}^k$, sufficiently close to $\tilde{x}^*$, it follows, using the continuity of $\tilde{h}(\tilde{x})$ and the consequent linear independence of $\nabla \tilde{h}_i(\tilde{x}^k)$, $i=\overline{1,m}$, that estimates

$$\tilde{y}^k = 2 S_k(\tilde{\theta}^k + \tilde{h}(\tilde{x}^k)) , \quad (3.28)$$

derived from $\nabla T(\tilde{x}^k, \tilde{\theta}^k, \tilde{\sigma}^k) = 0$, exist and have as limit points, when $\tilde{x}^k \rightarrow \tilde{x}^*$, the values $\tilde{y}^* = 2 S_* \tilde{\theta}^*$, where $\tilde{\theta}^*$ and $\tilde{\sigma}^*$ correspond to $\tilde{x}^*$. It is clear then that the controlling parameters, $\tilde{\theta}_i$ and $\tilde{\sigma}_i$, $i=\overline{1,m}$, are only determined by the unique GLM values to within multiplicative constants. Therefore, one of the vector parameters can be fixed and (3.27) can be regarded as a system of m equations in the m variables of the other parameter. Setting $\tilde{\theta} = 0$, (3.2) reduces to the quadratic penalty transformation described by Fiacco and McCormick (1968) in which the solution is obtained in the limit as $\tilde{\sigma} \rightarrow \infty$. It can be shown in this case that ill-conditioning and scaling difficulties, similar to those for the barrier function transformation (see §2.5), arise and therefore Powell's method treats $\tilde{\sigma}$ as a "constant" and varies $\tilde{\theta}$ to satisfy (3.27). A particular advantage of Powell's approach lies in the simplicity of an iterative scheme for solving (3.27).

Since $\tilde{x}^K$ minimizes $T(\tilde{x}, \tilde{\theta}^K, \tilde{\sigma}^K)$,

$$\nabla f(\tilde{x}^K) + 2 \Lambda(\tilde{x}^K) S_K(\tilde{\theta}^K + \tilde{h}(\tilde{x}^K)) = 0 \quad (3.29)$$

where $\Lambda(\tilde{x}) = [\nabla h_1(\tilde{x}), \dots, \nabla h_m(\tilde{x})]$. Assuming the columns of $\Lambda(\tilde{x}^K)$ are linearly independent, it follows that

$$\tilde{\theta}^K + \tilde{h}(\tilde{x}^K) = -\frac{1}{2} S_K^{-1} [\Lambda(\tilde{x}^K)^T \Lambda(\tilde{x}^K)]^{-1} \Lambda(\tilde{x}^K)^T \nabla f(\tilde{x}^K). \quad (3.30)$$

But, if $\tilde{\theta}^{K+1} = \tilde{\theta}^K + \delta \tilde{\theta}$ and $\tilde{\sigma}^{K+1} = \tilde{\sigma}^K$, then $\tilde{x}^{K+1} = \tilde{x}^K + \delta \tilde{x}$ satisfies

$$\nabla f(\tilde{x}^K + \delta \tilde{x}) + 2 \Lambda(\tilde{x}^K + \delta \tilde{x}) S_K(\tilde{\theta}^K + \delta \tilde{\theta}) = 0 \quad (3.31)$$

assuming $\tilde{x}^{K+1}$ satisfies (3.27) (i.e. $\tilde{h}(\tilde{x}^{K+1}) = 0$). Again using the assumption of linear independence, it follows that

$$\tilde{\theta}^K + \delta \tilde{\theta} = -\frac{1}{2} S_K^{-1} [\Lambda(\tilde{x}^K + \delta \tilde{x})^T \Lambda(\tilde{x}^K + \delta \tilde{x})]^{-1} \Lambda(\tilde{x}^K + \delta \tilde{x})^T \nabla f(\tilde{x}^K + \delta \tilde{x}). \quad (3.32)$$

Subtracting (3.32) from (3.30) and applying the mean value theorem gives

$$\tilde{h}(\tilde{x}^K) - \delta \tilde{\theta} = \frac{1}{2} S_K^{-1} K(\bar{\tilde{x}}) \delta \tilde{x}, \quad (3.33)$$

where $K(\bar{\tilde{x}}) = \nabla ([\Lambda(\bar{\tilde{x}})^T \Lambda(\bar{\tilde{x}})]^{-1} \Lambda(\bar{\tilde{x}})^T \nabla f(\bar{\tilde{x}}))$ and $\bar{\tilde{x}}$ is a

mean value. Since $\tilde{x}^K$ is the minimum of $T(\tilde{x}, \tilde{\theta}^K, \tilde{\sigma}^K)$, it is reasonable to assume that the Hessian of this function is positive definite at $\tilde{x}^K$. The application of the implicit function theorem to (3.29) then implies that $\tilde{x} = \tilde{x}(\tilde{\theta})$ in a neighbourhood of $\tilde{\theta}^K$ and further

$$\delta \tilde{x} = M(\tilde{\theta}^K) \delta \tilde{\theta}, \quad (3.34)$$

where $M_{ij}(\tilde{\theta}) = \frac{\partial x_i(\tilde{\theta})}{\partial \theta_j}$. From (3.33), it follows that

$$\tilde{h}(\tilde{x}^K) - \delta \tilde{\theta} = \frac{1}{2} S_K^{-1} K(\tilde{x}) M(\tilde{\theta}^K) \delta \tilde{\theta}. \quad (3.35)$$

It is now shown that the right hand side of (3.35) involves terms of order $S_K^{-1} \delta \tilde{\theta}$. The total derivative of (3.29) with respect to $\tilde{\theta}$, evaluated at $\tilde{\theta}^K$, is given by

$$\begin{aligned} & [\nabla^2 f(\tilde{x}^K) + 2\nabla_x [\Lambda(\tilde{x}^K)] S_K (\tilde{\theta}^K + \tilde{h}(\tilde{x}^K)) + 2\Lambda(\tilde{x}^K) S_K \Lambda(\tilde{x}^K)^T] M(\tilde{\theta}^K) \\ & + 2\Lambda(\tilde{x}^K) S_K = 0. \end{aligned} \quad (3.36)$$

The assumption of positive definiteness and (3.28) imply that

$$A = \nabla^2 f(\tilde{x}^K) + 2\nabla_x [\Lambda(\tilde{x}^K)] S_K (\tilde{\theta}^K + \tilde{h}(\tilde{x}^K)) + 2\Lambda(\tilde{x}^K) \bar{S} \Lambda(\tilde{x}^K)^T \quad (3.37)$$

is positive definite for $\bar{S} = \text{diag}(\bar{\sigma}_i)$ and $||\bar{\sigma}|| = 0(1)$.

It follows from (3.36) and (3.37), writing $\hat{S} = S_K - \bar{S}$, that

$$[A + 2\Lambda(\tilde{x}^K)\hat{S}\Lambda(\tilde{x}^K)^T]M(\tilde{\theta}^K) = -2\Lambda(\tilde{x}^K)S_K,$$

and hence

$$[I + 2 \cdot B \hat{S} B^T]C = -2 B S_K, \quad (3.38)$$

where $B = A^{-\frac{1}{2}} \Lambda(\tilde{x}^K)$, $C = A^{\frac{1}{2}} M(\tilde{\theta}^K)$ and $A^{\frac{1}{2}}$ exists since A is assumed positive definite. Solving (3.38) for C , the expansion

$$C = -B\{(B^TB)^{-1} - \frac{1}{2}(B^TB)^{-2}\hat{S}^{-1} + \frac{1}{4}(B^TB)^{-3}\hat{S}^{-2} - \dots\}\hat{S}^{-1}S_K \quad (3.39)$$

is generated. But since $||\hat{S}^{-1}S_K|| = 0(1)$, it follows that $||C|| = 0(1)$, and thus $||M(\tilde{\theta}^K)|| = 0(1)$. It now follows from (3.35) that

$$h(\tilde{x}^K) - \delta\tilde{\theta} = 0(||S_K^{-1}|| ||\delta\tilde{\theta}||).$$

This establishes the correction

$$\tilde{\theta}^{K+1} = \tilde{\theta}^K + h(\tilde{x}^K) \quad (3.40)$$

to within terms of order $S_K^{-1} \delta\tilde{\theta}$. The accuracy of (3.40) is therefore dependent on the magnitude of the σ parameter.

The σ parameter is also related to the convergence rate of Powell's method as it attempts to solve (3.27). Expansion of $\tilde{h}(\tilde{x})$ in a Taylor series about $\tilde{x}^K$ gives

$$\begin{aligned}\tilde{h}(\tilde{x}^{K+1}) &= \tilde{h}(\tilde{x}^K) + \Lambda(\tilde{x}^K)^T \delta \tilde{x} + O(\|\delta \tilde{x}\|^2) \\ &= \tilde{h}(\tilde{x}^K) + \Lambda(\tilde{x}^K)^T M(\tilde{\theta}^K) \delta \tilde{\theta} + O(\|M(\tilde{\theta}^K) \delta \tilde{\theta}\|^2).\end{aligned}\tag{3.41}$$

If $\tilde{h}(\tilde{x}^{K+1}) = 0$, then the correction given by Newton's method in solving (3.27) is

$$\delta \tilde{\theta} = - [\Lambda(\tilde{x}^K)^T M(\tilde{\theta}^K)]^{-1} \tilde{h}(\tilde{x}^K).\tag{3.42}$$

By using (3.42) instead of (3.40), second order convergence of $\tilde{h}(\tilde{x})$ to zero could be attained and Osborne (1971) has derived conditions, based on the Kantorovich convergence proof of Newton's method (Luenburger (1969)), under which (3.27) can be solved as a system of equations using Newton's method and starting from an initial estimate, $\tilde{\theta}^0 = 0$. The conditions state essentially that Newton's method converges if σ^0 is chosen sufficiently large. In claiming that (3.40) is in fact an approximation of (3.42), Powell (1967) has shown that by choosing σ sufficiently large, the difference between I and

$[-\Lambda(\tilde{x}^*)^T M(\tilde{\theta}^*)]$ can be made arbitrarily small. From (3.40) and (3.41),

$$\tilde{h}(\tilde{x}^{K+1}) = [I + \Lambda(\tilde{x}^K)^T M(\tilde{\theta}^K)] \tilde{h}(\tilde{x}^K) + O(\|M(\tilde{\theta}^K) \tilde{h}(\tilde{x}^K)\|^2). \quad (3.43)$$

Using Powell's argument again, (3.43) suggests that arbitrarily fast linear convergence of $\tilde{h}(\tilde{x})$ to zero can be attained by choosing $\tilde{\sigma}$ sufficiently large.

Powell's method essentially proceeds by minimizing $T(\tilde{x}, \tilde{\theta}^K, \tilde{\sigma}^K)$ starting from $\tilde{x}^{K-1}$. If $\tilde{h}(\tilde{x}^K)$ is converging to zero sufficiently rapidly, then $\tilde{\theta}^K$ is updated according to (3.40), but otherwise, $\tilde{\sigma}^K$ is increased keeping $S_K \tilde{\theta}^K$ constant, thus improving the accuracy of (3.40) and the convergence rate of $\tilde{h}(\tilde{x}^K)$ to zero.

3.7 Properties of the Hessian of $T(\tilde{x}, \tilde{\theta}, \tilde{\sigma})$.

A feature of Powell's method that will be used in a different context in Chapter 4, is found in the modifications which can be made to the inverse Hessian of $T(\tilde{x}, \tilde{\theta}, \tilde{\sigma})$ when changes are made to $\tilde{\theta}$ and $\tilde{\sigma}$. These modifications enable an estimate of the inverse Hessian matrix, generated by quasi-Newton methods of

unconstrained minimization, to be corrected and used as an initial approximation in the following minimization. Further discussion of this procedure is given in §4.4.

Powell has shown that changes in the Hessian can be ignored when small corrections are made to $\tilde{\theta}^K$ using (3.40), and that a sum of rank-one corrections can be computed when $\tilde{\sigma}^K$ is increased while $S_K \tilde{\theta}^K$ is held constant. From (3.2), it follows that

$$\begin{aligned} \nabla^2 T(\tilde{x}, \tilde{\theta}, \tilde{\sigma}) = \nabla^2 f(\tilde{x}) + 2 \sum_{i=1}^m \sigma_i \{ (\theta_i + h_i(\tilde{x})) \nabla^2 h_i(\tilde{x}) \\ + \nabla h_i(\tilde{x}) \nabla h_i(\tilde{x})^T \}. \end{aligned} \quad (3.44)$$

When $\tilde{\theta}^K$ is corrected using (3.40), the correction to $\nabla^2 T(\tilde{x}, \tilde{\theta}, \tilde{\sigma})$ is $\nabla^2 T(\tilde{x}^K, \tilde{\theta}^{K+1}, \tilde{\sigma}^K) - \nabla^2 T(\tilde{x}^K, \tilde{\theta}^K, \tilde{\sigma}^K) = 2 \sum_{i=1}^m \sigma_i h_i(\tilde{x}^K) \nabla^2 h_i(\tilde{x}^K)$. This may be neglected if $\|h(\tilde{x}^K)\|$ is small. The correction for $\nabla^2 T(\tilde{x}^K, \tilde{\theta}^K, \tilde{\sigma}^K)$ when σ_i is increased and θ_i is decreased by a factor $\tau_i \geq 1$, can be written

$$\begin{aligned} \nabla^2 T(\tilde{x}^K, \hat{\tilde{\theta}}^K, \hat{\tilde{\sigma}}^K) - \nabla^2 T(\tilde{x}^K, \tilde{\theta}^K, \tilde{\sigma}^K) = 2 \sum_{i=1}^m (\tau_i - 1) \sigma_i \{ h_i(\tilde{x}^K) \nabla^2 h_i(\tilde{x}^K) + \\ \nabla h_i(\tilde{x}^K) \nabla h_i(\tilde{x}^K)^T \}, \end{aligned}$$

where $\hat{\tilde{\theta}}^K = (\theta_1^K / \tau_1, \dots, \theta_m^K / \tau_m)^T$ and $\hat{\tilde{\sigma}}^K = (\tau_1 \sigma_1^K, \dots, \tau_m \sigma_m^K)^T$.

Assuming $||h(\tilde{x}^K)||$ is small, the correction can be simplified to a sum of rank-one corrections of the form

$$2 \sum_{i=1}^m (\tau_i - 1) \sigma_i \nabla h_i(\tilde{x}^K) \nabla h_i(\tilde{x}^K)^T.$$

The inverse Hessian estimate is then corrected by repeated application of the formula

$$(H + \lambda \tilde{a} \tilde{a}^T)^{-1} = H^{-1} - \frac{(H^{-1} \tilde{a}) (H^{-1} \tilde{a})^T}{\lambda^{-1} + \tilde{a}^T (H^{-1} \tilde{a})}. \quad (3.45)$$

Experience has shown that Powell's method is both reliable and efficient, especially when $\tilde{\theta}^K$ is close to $\tilde{\theta}^*$. In Chapter 4, Powell's method will be used in circumstances where such an assumption can be made.

CHAPTER 4

A HYBRID METHOD4.1 Introduction.

Motivation for the development of hybrid methods for constrained optimization is provided by

- (a) the computational problems, discussed in §2.5, associated with the active constraints treated by barrier function techniques, and
- (b) the excellent convergence properties of Powell's method when accurate estimates of the θ parameter are known.

The new approach is based on the following proposition.

Proposition 4.1. Solving the general mathematical programming problem (i.e. Problem 1.1) is equivalent to solving the sequence of subproblems:

$$\begin{aligned} &\text{minimize} && \{f(\tilde{x}) + r^k \sum_{i=1}^{\ell} \phi_i(g_i(\tilde{x}))\} \\ &\text{subject to} && h_i(\tilde{x}) = 0, \quad i = \overline{1, m}, \end{aligned} \tag{4.1}$$

where each subproblem, generated by an r^k -value from a strictly decreasing null r^k -sequence, is equality constrained.

This proposition will now be established.

4.2 Equality Constrained Barrier Function Transformations.

Writing $V = \{\tilde{x} : g_i(\tilde{x}) \geq 0, i=\overline{1,\ell}\}$ and $W = \{\tilde{x} : h_i(\tilde{x}) = 0, i=\overline{1,m}\}$, it follows that $R = V \cap W$. In seeking a minimum of $f(\tilde{x})$ on R , the barrier function transformation

$$T(\tilde{x}, r^k) = f(\tilde{x}) + r^k \sum_{i=1}^{\ell} \phi_i(g_i(\tilde{x}))$$

is defined on V_0 and is minimized, not on V_0 , but on $V_0 \cap W$ for a strictly decreasing null sequence $\{r^k\}$. Proposition 4.1 can now be established immediately since Lemmas 2.1 and 2.2 and Theorem 2.1, describing convergence properties of the standard barrier function methods, can be generalized in an obvious manner by replacing R by $V \cap W$, R_0 by $V_0 \cap W$ and $\text{cl}(R_0)$ by $\text{cl}(V_0) \cap W$. The proofs, given by Fiacco and McCormick (1968), can be extended without difficulty. Consequently, the results of Lemmas 2.1 and 2.2 and Theorem 2.1 can be assumed for the sequence of subproblems (4.1). In particular, the existence of local minima on $V_0 \cap W$ for subproblems generated by r^k

sufficiently small, and the basic convergence properties of barrier function transformations on $V_0 \cap W$ then follow.

Each subproblem of (4.1) is solved by Powell's method with a transformation function of the form

$$T(\underline{x}, \underline{\theta}, \underline{\sigma}, r) = f(\underline{x}) + r \sum_{i=1}^{\ell} \phi_i(g_i(\underline{x})) + (\underline{\theta} + \underline{h}(\underline{x}))^T S(\underline{\theta} + \underline{h}(\underline{x})), \quad (4.2)$$

where $\underline{\theta}$, $\underline{\sigma}$ and S are defined in (3.2) and r is the controlling parameter for the barrier term. The solution of the subproblem at r^k is denoted by $\underline{x}^k$ and the corresponding Powell parameters are denoted by $\underline{\theta}^k$ and $\underline{\sigma}^k$. Convergence results for the GLM estimates $\underline{u}^k$, given by (2.3), and $\underline{v}^k$, given by (3.28), both estimates being evaluated at $(\underline{x}^k, \underline{\theta}^k, \underline{\sigma}^k)$, are now considered.

The development follows a similar pattern to that presented for barrier functions in §2.2.

Let $K(\underline{x}^*)$ be the matrix with columns formed from the vectors $\nabla g_i(\underline{x}^*)$, $i \in B(\underline{x}^*)$, and $\nabla h_i(\underline{x}^*)$, $i = \overline{1, m}$, and assume $K(\underline{x}^*)$ has rank $s \leq n$. As in §2.2, let $K(\underline{x}^*)$ be partitioned into $K_1(\underline{x}^*)$ with rank s and $K_2(\underline{x}^*)$ such that $K_2(\underline{x}^*) = K_1(\underline{x}^*)U$ and denote by $K_1(\underline{x}^k)$ and $K_2(\underline{x}^k)$, the matrices $K_1(\underline{x}^*)$ and $K_2(\underline{x}^*)$ evaluated at $\underline{x}^k$.

Also, let the components of α^k and β^k be selected from u^k and $(-v^k)$ corresponding to the constraints in $K_1(\tilde{x}^*)$ and $K_2(\tilde{x}^*)$, respectively. The analogue of Theorem 2.3 can now be established.

Theorem 4.1.

If $\{r^k\}$ is a strictly decreasing null sequence s.t. $\tilde{x}^k \rightarrow \tilde{x}^*$, $h(\tilde{x}^k) = 0$ for all k and the components of β^k are bounded for $k \geq \hat{k}$, where $\hat{k}$ is sufficiently large, then the sequence of vectors $(\alpha^k + U\beta^k)$, $k = \hat{k}, \hat{k}+1, \dots$ converges and

$$\lim_{k \rightarrow \infty} (\alpha^k + U\beta^k) = [K_1(\tilde{x}^*)^T K_1(\tilde{x}^*)]^{-1} K_1(\tilde{x}^*)^T \nabla f(\tilde{x}^*). \quad (4.3)$$

Furthermore, if α^k is bounded then a convergent subsequence of $\{(\alpha^k, \beta^k)\}$ exists with limit (α, β) satisfying (4.3). When padded out with zeros corresponding to the inactive inequality constraints, the limit also satisfies the Kuhn-Tucker necessary conditions (1.15), (1.16) and (1.17).

Proof. The Theorem is proved using the arguments of Theorem 2.3 and its Corollary 1.

Theorem 4.2.

If $\{r^k\}$ is a strictly decreasing null sequence s.t. $\tilde{x}^k \rightarrow \tilde{x}^*$, $h(\tilde{x}^k) = 0$ for all k , then

for k sufficiently large, the GLM estimates α^k and β^k are bounded if either condition (i) or condition (ii) of Lemma 1.4 is satisfied.

Proof. For condition (i), the result follows using the arguments of Corollary 3 of Theorem 2.3.

If condition (ii) of Lemma 1.4 is satisfied, there exists $\tilde{d}$ s.t. $\tilde{d}^T \nabla g_i(\tilde{x}^*) > 0$, $i \in B(\tilde{x}^*)$ and $\tilde{d}^T \nabla h_i(\tilde{x}^*) = 0$, $i = \overline{1, m}$, where $\nabla h_i(\tilde{x}^*)$, $i = \overline{1, m}$, are linearly independent. By the continuity of the constraints and for k sufficiently large,

$$\begin{aligned} \tilde{d}^T \nabla g_i(\tilde{x}^k) &> 0, \quad i \in B(\tilde{x}^*), \\ |\tilde{d}^T \nabla h_i(\tilde{x}^k)| &< \delta, \quad i = \overline{1, m}, \quad \delta > 0, \end{aligned} \tag{4.4}$$

and $\nabla h_i(\tilde{x}^k)$, $i = \overline{1, m}$, are linearly independent.

Applying Theorem 1.1 to (4.1) at $\tilde{x}^k$, using the linear independence assumption at $\tilde{x}^k$, the existence of a unique bounded GLM vector, $\tilde{v}^k$, satisfying

$$\{\nabla f(\tilde{x}^k) + r^k \sum_{i=1}^{\ell} \frac{d\phi_i(g_i(\tilde{x}^k))}{dg_i} \nabla g_i(\tilde{x}^k)\} + \sum_{i=1}^m \tilde{v}_i^k \nabla h_i(\tilde{x}^k) = 0$$

may be deduced. It follows that

$$\tilde{d}^T \nabla f(\tilde{x}^k) - \sum_{i \in B(\tilde{x}^*)} u_i^k \tilde{d}^T \nabla g_i(\tilde{x}^k) + \sum_{i=1}^m \tilde{v}_i^k \tilde{d}^T \nabla h_i(\tilde{x}^k) + 0(r^k) = 0,$$

where $\tilde{d}$ satisfies condition (ii) of Lemma 1.4. Then

$$\sum_{i \in B(\tilde{x}^*)} u_i^k \leq \frac{\tilde{d}^T \nabla f(\tilde{x}^k) + \sum_{i=1}^m v_i^k \tilde{d}^T \nabla h_i(\tilde{x}^k) + 0(r^k)}{\min_{i \in B(\tilde{x}^*)} [\tilde{d}^T \nabla g_i(\tilde{x}^k)]}$$

and the boundedness of u_i^k follows since each term in the sum on the left is positive by Definition 2.1 and the right side is bounded by (4.4).

The application of Powell's method will only be computationally efficient if it converges rapidly for each value of r^k (i.e. if the equality constrained problem is easily solved using $(\tilde{x}^{k-1}, \tilde{\theta}^{k-1}, \tilde{\sigma}^{k-1})$ as the initial estimate of $(\tilde{x}^k, \tilde{\theta}^k, \tilde{\sigma}^k)$). The following lemma and theorem show that for k sufficiently large, $\tilde{x}$ and $\tilde{\theta}$ are differentiable functions of r in a sufficiently small neighbourhood of r^k .

Lemma 4.1.

If $\{r^k\}$ is a strictly decreasing null sequence s.t. $\tilde{x}^k \rightarrow \tilde{x}^*$, $h(\tilde{x}^k) = 0$ for all k , $B_1(\tilde{x}^*) = B(\tilde{x}^*)$ (i.e. the strictly complementarity condition (2.13) is satisfied) and the sufficiency conditions of Theorem 1.3 for $\tilde{x}^*$ to be an isolated local minimum of Problem 1.1 are satisfied, then for k sufficiently large and for all $\tilde{d}$ satisfying $\|\tilde{d}\| = 1$ and

$$\tilde{d}^T \nabla h_i(\tilde{x}^*) = 0, \quad i = \overline{1, m}, \quad (4.5)$$

it follows that

$$\tilde{d}^T \nabla_{\tilde{x}}^2 T(\tilde{x}^k, \tilde{\theta}^k, \tilde{\sigma}^k, r^k) \tilde{d} > 0. \quad (4.6)$$

Proof. From (4.2), using (2.3) and (3.28),

$$\begin{aligned} \nabla_{\tilde{x}}^2 T(\tilde{x}^k, \tilde{\theta}^k, \tilde{\sigma}^k, r^k) &= \nabla^2 f(\tilde{x}^k) - \sum_{i=1}^{\ell} u_i^k \nabla^2 g_i(\tilde{x}^k) \\ &\quad + \sum_{i=1}^{\ell} r^k \frac{d^2 \phi_i(g_i(\tilde{x}^k))}{dg_i^2} \nabla g_i(\tilde{x}^k) \nabla g_i(\tilde{x}^k)^T \\ &\quad + \sum_{i=1}^m v_i^k \nabla^2 h_i(\tilde{x}^k) + \sum_{i=1}^m 2\sigma_i^k \nabla h_i(\tilde{x}^k) \nabla h_i(\tilde{x}^k)^T. \end{aligned} \quad (4.7)$$

But for k sufficiently large, $u_i^k \nabla^2 g_i(\tilde{x}^k)$, $i \notin B_1(\tilde{x}^*)$ (i.e. all i s.t. $u_i^* = 0$), can be neglected. Then for any unit vector, $\tilde{d}$, satisfying (4.5) and $\tilde{d}^T \nabla g_i(\tilde{x}^*) \neq 0$ for some $i \in B_1(\tilde{x}^*)$, the result follows for sufficiently large k using Definition 2.1 and the consequent result (Fiacco and McCormick (1968), p.74) that

$$\lim_{k \rightarrow \infty} r^k \frac{d^2 \phi_i(g_i(\tilde{x}^k))}{dg_i^2} = \infty, \quad i \in B_1(\tilde{x}^*).$$

If $\tilde{d}$ satisfies (4.5) and

$$\tilde{d}^T \nabla g_i(\tilde{x}^*) = 0, \quad i \in B_1(\tilde{x}^*), \quad (4.8)$$

it follows from Theorem 1.3 that, for k sufficiently large,

$$\begin{aligned} \tilde{d}^T \nabla^2 T(\tilde{x}^k, \tilde{\theta}^k, \tilde{\sigma}^k, \tilde{r}^k) \tilde{d} &= \tilde{d}^T \{ \nabla^2 f(\tilde{x}^k) - \sum_{i \in B_1(\tilde{x}^*)} u_i^k \nabla^2 g_i(\tilde{x}^k) \\ &\quad + \sum_{i=1}^m v_i^k \nabla^2 h_i(\tilde{x}^k) \} \tilde{d} + O(\tilde{r}^k) > 0. \end{aligned}$$

The conditions of Theorem 1.3 require that $\tilde{d} \in C_1(\tilde{x}^*)$. This is an immediate consequence of (4.5) and (4.8).

Theorem 4.3.

If $\{\tilde{r}^k\}$ is a strictly decreasing null sequence s.t. $\tilde{x}^k \rightarrow \tilde{x}^*$, $\tilde{h}(\tilde{x}^k) = 0$ for all k , $B_1(\tilde{x}^*) = B(\tilde{x}^*)$ (i.e. the strictly complementarity condition (2.13) is satisfied) and the sufficiency conditions of Theorem 1.3 for $\tilde{x}^*$ to be an isolated local minimum of Problem 1.1 are satisfied, then for k sufficiently large $(\tilde{x}, \tilde{\theta})$ is given by

$$(\tilde{x}, \tilde{\theta}) = (\tilde{X}(r), \tilde{\Theta}(r))$$

where $\tilde{X}_i(r)$, $i = \overline{1, n}$, and $\tilde{\Theta}_i(r)$, $i = \overline{1, m}$, are differentiable functions of r in a sufficiently small neighbourhood of $\tilde{r}^k$, excluding $r=0$.

Proof. Since $\tilde{x}^k$ minimizes $T(\tilde{x}, \tilde{\theta}^k, \tilde{\sigma}^k, r^k)$,

$$\nabla T(\tilde{x}^k, \tilde{\theta}^k, \tilde{\sigma}^k, r^k) = \nabla f(\tilde{x}^k) - \sum_{i=1}^{\ell} u_i^k \nabla g_i(\tilde{x}^k) + \sum_{i=1}^m v_i^k \nabla h_i(\tilde{x}^k) = 0 \quad (4.9)$$

and also

$$\tilde{h}(\tilde{x}^k) = 0. \quad (4.10)$$

Treating $\tilde{\sigma}$ as a constant, (4.9) and (4.10) form a system of $(n+m)$ equations in $(n+m+1)$ variables $(\tilde{x}, \tilde{\theta}, r)$ which is satisfied by $(\tilde{x}^k, \tilde{\theta}^k, r^k)$. Using Lemma 4.1 and the argument of Theorem 3.6, it can be shown that the Jacobian of (4.9) and (4.10) with respect to $(\tilde{x}, \tilde{\theta})$, evaluated at $(\tilde{x}^k, \tilde{\theta}^k, r^k)$, is nonsingular. The result then follows by applying the implicit function theorem at $(\tilde{x}^k, \tilde{\theta}^k, r^k)$.

Corollary 1. If the r^k -sequence is generated by

$$r^{k+1} = \gamma r^k, \quad k=1, 2, \dots, \quad 0 < \gamma < 1,$$

then for k sufficiently large and γ sufficiently close to 1, $(\tilde{x}^k, \tilde{\theta}^k)$ provides a good initial estimate of $(\tilde{x}^{k+1}, \tilde{\theta}^{k+1})$.

Proof. The result follows using the differentiability of $\tilde{x}(r)$ and $\tilde{\theta}(r)$ in a sufficiently small neighbourhood of r^k , noting that $r^{k+1} = r^k + \delta r$, where $\delta r = (\gamma - 1)r^k$.

Remark 4.1. The linear independence of vectors $\nabla h_i(\tilde{x}^k)$, $i=\overline{1,m}$, is a basic requirement for the equality constraints. It is also required by Powell's method in solving the subproblem of (4.1) generated by r^k .

The basic approach described in this section has no advantage over the mixed transformation methods of Fiacco and McCormick (1968) and Lootsma (1968b). In fact, this approach still suffers from the computational difficulties discussed in §2.5 as well as compounding these difficulties by requiring that the equality constraints be strictly satisfied at the minimum, $\tilde{x}^k$, for each r^k . In the following sections, a number of important extensions of the basic approach are developed. These extensions provide improved convergence properties and lead to the construction of an efficient computational algorithm.

4.3 Inequality Constraint Classification.

An important aspect of the new approach involves classification of active and inactive inequality constraints treated by the barrier functions of (4.2). This is accomplished by observing the behaviour of the constraint values and the GLM estimates, $\tilde{u}^k$, generated by the barrier

function term at $\tilde{x}^k$. The following results can be deduced.

Results 4.1. 1. If $\tilde{x}^*$ is a regular point of R , then

$$(a) \quad g_i(\tilde{x}^*) > 0 \Rightarrow u_i^k \rightarrow 0, \quad g_i(\tilde{x}^k) + g_i(\tilde{x}^*) > 0 \text{ as } r^k \rightarrow 0, \text{ and}$$

$$(b) \quad g_i(\tilde{x}^*) = 0 \Rightarrow u_i^k \rightarrow u_i^* \geq 0, \quad g_i(\tilde{x}^k) \rightarrow 0 \text{ as } r^k \rightarrow 0.$$

2. If $\tilde{x}^*$ is not a regular point of R , then

$$(a) \quad g_i(\tilde{x}^*) > 0 \Rightarrow u_i^k \rightarrow 0, \quad g_i(\tilde{x}^k) + g_i(\tilde{x}^*) > 0 \text{ as } r^k \rightarrow 0, \text{ and}$$

$$(b) \quad g_i(\tilde{x}^*) = 0 \Rightarrow u_i^k \rightarrow \infty, \quad g_i(\tilde{x}^k) \rightarrow 0 \text{ as } r^k \rightarrow 0.$$

From these results, Table 4.1 is constructed.

Conditions	Classification
$\left. \begin{array}{l} g_i(\tilde{x}^k) \rightarrow 0 \\ -r^k \frac{d\phi_i(g_i(\tilde{x}^k))}{dg_i} \rightarrow L \geq 0 \end{array} \right\} r^k \rightarrow 0$	Constraint, $g_i(\tilde{x})$, is active at $\tilde{x}^*$.
$\left. \begin{array}{l} g_i(\tilde{x}^k) \rightarrow K > 0 \\ -r^k \frac{d\phi_i(g_i(\tilde{x}^k))}{dg_i} \rightarrow 0 \end{array} \right\} r^k \rightarrow 0$	Constraint, $g_i(\tilde{x})$, is not active at $\tilde{x}^*$.
$\left. \begin{array}{l} g_i(\tilde{x}^k) \rightarrow 0 \\ -r^k \frac{d\phi_i(g_i(\tilde{x}^k))}{dg_i} \rightarrow \infty \end{array} \right\} r^k \rightarrow 0$	Kuhn-Tucker conditions are not satisfied at $\tilde{x}^*$.

Table 4.1. Classification of Active and Inactive Constraints

The testing for convergence of constraint values and GLM estimates is computed using the solutions of (4.1) at successive values of r^k , even though the convergence results relating to the GLM estimates were derived for a subsequence of such values (see Theorem 4.1). It can be expected that for most "well behaved" problems in which $\tilde{x}^*$ is a regular point of R , the sequence $\tilde{u}^k$, $k=1,2,\dots$ will be convergent.

The following conditions, deduced from Table 4.1, will be used as a basis for attempting to classify each inequality constraint as active or inactive. Constraints are considered active if the conditions

$$g_i(\tilde{x}^k)/g_i(\tilde{x}^{k-1}) < \gamma r^k/r^{k-1}, \gamma > 1, \quad (4.11)$$

and

$$|u_i^k - u_i^{k-1}|/(1 + u_i^k) < \epsilon, \quad 0 < \epsilon \ll 1,$$

are both satisfied. Constraints are considered inactive if the conditions

$$u_i^k/u_i^{k-1} < \gamma r^k/r^{k-1}, \gamma > 1, \quad (4.12)$$

and

$$|\phi_i(g_i(\tilde{x}^k)) - \phi_i(g_i(\tilde{x}^{k-1}))|/[1 + |\phi_i(g_i(\tilde{x}^k))|] < \epsilon, \quad 0 < \epsilon \ll 1,$$

are both satisfied. The choice of γ and ϵ in (4.11)

and (4.12) will depend on the severity of the test required. If a constraint does not satisfy either (4.11) or (4.12), it remains unclassified and a further classification attempt is made after the next minimization at r^{k+1} . If the Kuhn-Tucker conditions do not hold, inactive constraints will be recognised, but for active constraints, the GLM test of (4.11) will be difficult to satisfy. These constraints will then remain unclassified in the barrier function. If the strict complementarity condition (2.13) is not satisfied, inactive constraints will be recognized, but for active constraints with $u_{\ell}^* = 0$, the GLM test of (4.11) will again be difficult to satisfy if ϵ is too small. In practice, values of $\gamma = 1.1$ and $\epsilon = 0.01$ have been found to be satisfactory when r^k is generated by $r^k = 10^{1-k}$, $k=1,2,\dots$.

The implications of this constraint classification are now discussed.

4.4 Active Constraint Transfers.

After solving (4.1) at r^k , an attempt is made to classify inequality constraints treated by the barrier function term. Constraints classified as inactive are omitted from the barrier function

term of $T(\underline{x}, \underline{\theta}, \underline{\sigma}, r)$ (given by (4.2)) in following minimizations. To ensure that these constraints remain inactive in the subsequent minimizations, they may be evaluated at each point where $T(\underline{x}, \underline{\theta}, \underline{\sigma}, r)$ is evaluated. Violated constraints may be returned to the barrier function term as unclassified and the current minimization can then be restarted from the previous feasible point. In practice, such action has not been necessary and only constraints inactive at $\underline{x}^*$ have been classified as inactive.

Constraints classified as active are treated as equality constraints by transferring them to the Powell term of (4.2). The transfer is effected using the barrier function estimates of the GLM, $\underline{u}^k$, associated with the active inequality constraints, to provide an initial estimate of the θ parameter occurring in the Powell term. Using (2.3) and (3.28), denoting the classified active constraint as $g(\underline{x}) \equiv h(\underline{x})$, and noting that a sign change is required to preserve the form of (1.12), it follows that

$$2\sigma(\theta + h(\underline{x}^k)) = r^k \frac{d\phi(g(\underline{x}^k))}{dg} \quad (4.13)$$

Choosing σ and θ to satisfy (4.13) ensures that $\underline{x}^k$ will still minimize (4.2). The first application of (3.40) for correcting θ then implies that θ should be

chosen according to

$$\theta = \frac{r^k}{2\sigma} \frac{d\phi(g(\tilde{x}^k))}{dg} . \quad (4.14)$$

A suitable choice for σ is made by considering the influence of σ on the accuracy of (3.40) and on the rate of convergence of Powell's method. If the barrier function GLM estimate is accurate to three or four decimal places and θ is chosen to be $0(10g(\tilde{x}^k))$, then by (3.40), $h(\tilde{x})$ will be decreased in the following minimization by a factor of between 10^{-2} and 10^{-3} . It should also be noted that since the GLM estimates are accurate, $||\delta\theta||$ is small and thus σ does not have to be very large to ensure the accuracy of (3.40). In practice, σ is seldom required to exceed 100.

In transferring active constraints, a constraint which is inactive at $\tilde{x}^*$ may be classified as active and transferred in error. Such an event can be recognised and rectified in the following way. From (2.3) and by the definition of $\phi_i(g_i(\tilde{x}))$, the GLM estimates corresponding to the inequality constraints are essentially nonnegative. When the classified active constraints are transferred to the Powell term, the corresponding GLM estimates become essentially nonpositive. It can be seen from (1.17)

that if ever one of the GLM estimates for the transferred inequalities becomes positive, the corresponding constraint is no longer active since $f(\tilde{x})$ can be decreased by treating that constraint as a strict inequality. A check on the sign of the relevant GLM estimates during each iteration of Powell's method provides a means of recognising such occurrences. This use of the GLM estimates was suggested by Powell (1967) in a similar context. It can also be seen from (3.40) that $h(\tilde{x}^K)$ must be positive for such a sign change to occur since, by assumption,

$$2\sigma^{K-1}(\theta^{K-1} + h(\tilde{x}^{K-1})) = 2\sigma^K\theta^K \leq 0, \quad \sigma^{K-1} = \sigma^K > 0,$$

and

$$2\sigma^K(\theta^K + h(\tilde{x}^K)) > 0$$

from which it can be deduced that $h(\tilde{x}^K) > 0$. This observation implies that $\tilde{x}^K$ is feasible with respect to $h(\tilde{x})$ and the constraint may therefore be transferred back to the barrier function term of (4.2). The minimizations can then be restarted from $\tilde{x}^K$ and further attempts to reclassify this constraint may be made in the normal manner.

4.5 Hessian Corrections.

A further important aspect of the hybrid approach of this chapter is provided by an extension of the ideas of §3.7 concerning corrections for the inverse Hessian estimates of the transformation function (3.2). When inequality constraints are classified as inactive, their contributions to the Hessian may be ignored since their GLM estimates will be tending to zero and $r^k \frac{d^2 \phi(g(\tilde{x}^k))}{dg^2}$ will be small.

When active inequality constraints are transferred to the Powell term of (4.2) a simple rank-one correction can be developed for the inverse Hessian estimate. The Hessian of $T(\tilde{x}, \tilde{\theta}, \tilde{\sigma}, r)$ is given by (4.7) and, without loss of generality, the ℓ th constraint, $g_\ell(\tilde{x})$, may be transferred to the Powell term as $h_v(\tilde{x})$, where $v = m+1$. The change in the Hessian at $\tilde{x}^k$, δH^k , is given by

$$\begin{aligned} \delta H^k = & 2\sigma_v [(\theta_v + h_v(\tilde{x}^k)) \nabla^2 h_v(\tilde{x}^k) + \nabla h_v(\tilde{x}^k) \nabla h_v(\tilde{x}^k)^T] \\ & - r^k \left[\frac{d\phi_\ell(g_\ell(\tilde{x}^k))}{dg_\ell} \nabla^2 g_\ell(\tilde{x}^k) + \frac{d^2 \phi_\ell(g_\ell(\tilde{x}^k))}{dg_\ell^2} \nabla g_\ell(\tilde{x}^k) \nabla g_\ell(\tilde{x}^k)^T \right]. \end{aligned}$$

But, $h_v(\tilde{x}) \equiv g_\ell(\tilde{x})$ and $h_v(\tilde{x}^k)$ is small. Therefore, using (4.14), it follows that

$$\delta H^k \approx [2\sigma_v - r^k \frac{d^2 \phi_\ell(g_\ell(\tilde{x}^k))}{dg_\ell^2}] \nabla g_\ell(\tilde{x}^k) \nabla g_\ell(\tilde{x}^k)^T. \quad (4.15)$$

Since both $2\sigma_v$ and $r^k \frac{d^2 \phi_\ell(g_\ell(\tilde{x}^k))}{dg_\ell^2}$ can be large,

δH^k can represent a substantial modification of $\nabla^2 T(\tilde{x}^k, \tilde{\theta}^k, \tilde{\sigma}^k, r^k)$. The inverse of the Hessian is corrected by a single application of (3.45).

The correction procedures of §3.7 and this section provide useful initial approximations of the new inverse Hessian which appear to improve the convergence properties of the unconstrained methods of minimization. The importance of using such an approximation is now demonstrated numerically. It is noted that similar conclusions have been reached by Huang and Levy (1969), p.29, in a slightly different context. Table 4.2 contains the results of applying the hybrid method described by Algorithm 4.1 (see §4.6), and Algorithm 5.1 (see §5.4) based on the Davidon-Fletcher-Powell unconstrained minimization method, to the following test problem.

(i) Colville: Problem 3. (Colville (1968)).

$$\begin{aligned} \text{Minimize } f(\tilde{x}) = & 5.3578547x_3^2 + 0.8356891x_1x_5 + 37.293239x_1 \\ & - 40792.141 \end{aligned}$$

subject to

$$g_1(\underline{x}) = 85.334407 + 0.0056858x_2x_5 + 0.0006262x_1x_4 - 0.0022053x_3x_5 \geq 0,$$

$$g_2(\underline{x}) = 92 - g_1(\underline{x}) \geq 0,$$

$$g_3(\underline{x}) = -9.48751 + 0.0071317x_2x_5 + 0.0029955x_1x_2 + 0.0021813x_3^2 \geq 0,$$

$$g_4(\underline{x}) = 20 - g_3(\underline{x}) \geq 0,$$

$$g_5(\underline{x}) = -10.699039 + 0.0047026x_3x_5 + 0.0012547x_1x_3 + 0.0019085x_3x_4 \geq 0,$$

$$g_6(\underline{x}) = 5 - g_5(\underline{x}) \geq 0,$$

$$g_7(\underline{x}) = x_1 - 78 \geq 0,$$

$$g_8(\underline{x}) = 102 - x_1 \geq 0,$$

$$g_9(\underline{x}) = x_2 - 33 \geq 0,$$

$$g_{10}(\underline{x}) = 45 - x_2 \geq 0,$$

$$\left. \begin{aligned} g_{5+2i}(\underline{x}) &= x_i - 29 \geq 0 \\ g_{6+2i}(\underline{x}) &= 45 - x_i \geq 0 \end{aligned} \right\} i = \overline{3,5}.$$

The constrained minimum is at $\underline{x}^* = (78.0, 33.0, 29.9952560, 45.0, 36.7758129)$ with $f(\underline{x}^*) = -30665.5387$ and the second, fifth, seventh, ninth and fourteenth constraints active. An initial feasible point is $\underline{x}^0 = (78.62, 33.44, 31.07, 44.18, 35.32)$. Although results are only presented for this problem, similar results have been obtained for all other problems tested.

In Table 4.2 the columns headed "Reset" give the number of iterations, "it", and function

Iteration		Classification			Reset		Corrected	
k	κ	un	in	ac	it	ev	it	ev
1	1	16	0	0	14	34	14	34
2	1	16	0	0	28	72	12	26
3	1	7	7	2	18	41	11	25
	2				7	14	3	5
	3				1	1		
	4				2	3		
4	1	0	11	5	7	13	4	7
	2				6	11	2	3
	3				2	3	1	1
	4				4	7	2	3
	5						2	3
Totals					89	199	51	107

Table 4.2. Colville: Problem 3. Comparison of Resetting or Correcting Inverse Hessian Estimates

evaluations, "ev", required to minimize $T(\underline{x}, \underline{\theta}, \underline{\sigma}, r)$ when the inverse Hessian was reset to the identity before each minimization. The columns headed "Corrected" give the corresponding number of iterations and function evaluations required when the previous estimates of the inverse Hessian were retained and corrected where possible. The columns headed " k " and " κ " indicate the κ th Powell iteration within the solution of (4.1) at r^k . An indication

is also given of the constraint classification at each stage with the columns headed "un", "in" and "ac" giving the number of constraints at the beginning of each unconstrained minimization which are unclassified, inactive or active, respectively.

Remarks 4.2. 1. Convergence tests used in the computations are discussed in §4.6. The same tests were used for both the "Corrected" and "Reset" results.

2. A function evaluation is defined as the evaluation of $T(\underline{x}, \underline{\theta}, \underline{\sigma}, r)$ and its gradient vector. It should be noted that this requires evaluations of problem functions and their gradients. As the inequality constraints are classified as inactive and are removed, a function evaluation will require considerably less computation.

3. The apparent discrepancies in Table 4.2 at iterations $k=3$, $\kappa=3$ and $\kappa=4$ and $k=4$, $\kappa=5$ are due to the convergence tests being satisfied at slightly different points. This does not affect the implications of the results which show the advantages of retaining an estimate of the inverse Hessian whenever possible.

4.6 The Hybrid Algorithm.

The hybrid approach of this chapter can now be written in algorithm form.

Algorithm 4.1.

- (i) Set $k=\kappa=1$. If $m > 0$ then $\theta_{\tilde{i}}^1 = 0$, $\sigma_{\tilde{i}}^1 = 1$,
 $\tilde{i}=\overline{1,m}$; else continue.
- (ii) Minimize $T(\tilde{x}, \tilde{\theta}^k, \tilde{\sigma}^k, r^k)$ to find $\tilde{x}^{k,\kappa}$.
- (iii) If $m = 0$ then define $\tilde{x}^k = \tilde{x}^{k,\kappa}$ and go to (iv);
 else if $\max_{\tilde{i}=\overline{1,m}} |h_{\tilde{i}}(\tilde{x}^{k,\kappa})| \leq \varepsilon_1$ then define
 $\tilde{x}^k = \tilde{x}^{k,\kappa}$, $\tilde{\theta}^k = \tilde{\theta}^{\kappa}$, $\tilde{\sigma}^k = \tilde{\sigma}^{\kappa}$ and
 go to (iv);
 else: (a) Check GLM sign of transferred inequality constraints
 (see §4.3).
 (b) Correct $\tilde{\theta}^{\kappa}$ by (3.40) or increase $\tilde{\sigma}^{\kappa}$ and correct the inverse Hessian as required by Powell's method.
 (c) Go to (ii) with $\kappa=\kappa+1$.
- (iv) If $\ell = 0$ then stop.
 else attempt to classify inequality constraints.
 If inactive, remove from the barrier term (i.e. $\ell=\ell-1$).

If active:

- (a) Transfer to the Powell term (i.e. $\ell=\ell-1$, $m=m+1$).
 - (b) Predict θ_m^k , σ_m^k (see §4.3).
 - (c) Correct the inverse Hessian estimate (see §4.4).
- (v) If $r^k < \epsilon_2$ then stop;
 else if $m=0$ then go to (ii) with $k=k+1$;
 else define $\theta_i^1 = \theta_i^k$, $\sigma_i^1 = \sigma_i^k$, $i = \overline{1, m}$ and
 go to (ii) with $\kappa=1$ and $k=k+1$.

The use of (2.22) as the barrier function transformation in (4.1) has two related advantages. The convergence properties of this transformation, discussed by Osborne and Ryan (1970a) and developed in Chapter 2, enable early classification of both active and inactive constraints and consequently only a small number of moderate r -values are required in the minimizing sequence. This early classification effectively reduces the number of subproblems given by (4.1) and also reduces the numerical difficulties of §2.5 associated with active constraints treated by barrier function techniques. The transformation (2.22) also provides accurate estimates for θ when constraints are transferred to the Powell term. It is under just these conditions that Powell's method has been found

to be so efficient, and the estimates seldom require more than two or three corrections by (3.40). The minimizations required during these corrections are also extremely simple, supporting the implications of Theorem 4.3 and Corollary 1.

4.7 Numerical Results Using Algorithm 4.1.

The performance of Algorithm 4.1 is now numerically evaluated using the following standard test problems.

(i) Colville: Problem 1. (Colville (1968)).

See §2.4.

(ii) Colville: Problem 2. (Colville (1968)).

$$\text{Minimize } f(\tilde{x}) = \sum_{i=1}^{10} b_i x_{i+5} - \sum_{i=1}^5 \sum_{j=1}^5 c_{ij} x_i x_j - 2 \sum_{i=1}^5 d_i x_i^3$$

subject to

$$g_j(\tilde{x}) = e_j + 2 \sum_{i=1}^5 c_{ij} x_i + 3d_j x_j^2 - \sum_{i=1}^{10} a_{ij} x_{i+5} \geq 0, \quad j = \overline{1, 5},$$

$$g_{i+5}(\tilde{x}) = x_i \geq 0, \quad i = \overline{1, 15},$$

where a_{ij} , c_{ij} , b_i , d_j , e_j are given in Table 2.1.

The constrained minimum is at $\tilde{x}^* = (0.3, 0.3334676, 0.4, 0.4283101, 0.2239649, 0, 0, 5.174041, 0, 3.061109, 11.83955, 0, 0, 0.1038962, 0)$ with $f(\tilde{x}^*) = 32.34868$ and constraints 1, 2, 3, 4, 5, 11, 12, 14, 17, 18 and 20 active. An initial feasible point is $\tilde{x}^0 = (0.123066, 0.153779, 0.418090, 0.546784, 0.454095, 0.023147, 0.170608, 6.50313, 0.076608, 3.31380, 12.5012, 0.006822, 0.010191, 0.255809, 0.063178)$ with $f(\tilde{x}^0) = 36.7547$. This problem is the dual formulation of Colville: Problem 1.

(iii) Colville: Problem 3. (Colville (1968)).

See §4.4.

(iv) Colville: Problem 7. (Colville (1968)).

$$\text{Minimize } f(\tilde{x}) = \sum_{i=1}^{16} \sum_{j=i}^{16} A_{ij} (x_i^2 + x_i + 1.0) (x_j^2 + x_j + 1.0)$$

$$\left. \begin{array}{l} \text{subject to } g_i(\tilde{x}) = x_i \geq 0 \\ g_{i+16}(\tilde{x}) = 5 - x_i \geq 0 \end{array} \right\} i = \overline{1, 16},$$

$$h_1(\tilde{x}) = 0.22x_1 + 0.2x_2 + 0.19x_3 + 0.25x_4 + 0.15x_5 + 0.11x_6 + 0.12x_7 + 0.13x_8 + x_9 - 2.5 = 0,$$

$$h_2(\tilde{x}) = -1.46x_1 - 1.3x_3 + 1.82x_4 - 1.15x_5 + 0.8x_7 + x_{10} - 1.1 = 0,$$

$$h_3(\tilde{x}) = 1.29x_1 - 0.89x_2 - 1.16x_5 - 0.96x_6 - 0.49x_8 + x_{11} + 3.1 = 0,$$

$$h_4(\tilde{x}) = -1.1x_1 - 1.06x_2 + 0.95x_3 - 0.54x_4 - 1.78x_6 - 0.41x_7 + x_{12} + 3.5 = 0,$$

$$h_5(\tilde{x}) = -1.43x_4 + 1.51x_5 + 0.59x_6 - 0.33x_7 - 0.43x_8 + x_{13} - 1.3 = 0,$$

$$h_6(\tilde{x}) = -1.72x_2 - 0.33x_3 + 1.62x_5 + 1.24x_6 + 0.21x_7 - 0.26x_8 + x_{14} - 2.1 = 0,$$

$$h_7(\tilde{x}) = 1.12x_1 + 0.31x_4 + 1.12x_7 - 0.36x_9 + x_{15} - 2.3 = 0,$$

$$h_8(\tilde{x}) = 0.45x_2 + 0.26x_3 - 1.10x_4 + 0.58x_5 - 1.03x_7 + 0.1x_8 + x_{16} + 1.5 = 0,$$

where A_{ij} , $i=\overline{1,16}$, $j=\overline{1,16}$, is given in Table 4.3.

The constrained minimum is at $\tilde{x}^* = (0.0398473, 0.7919832, 0.2038704, 0.8443579, 1.269906, 0.9347387, 1.681962, 0.1553008, 1.567870, 0, 0, 0, 0.6602040, 0, 0.6742560, 0)$ with $f(\tilde{x}^*) = 244.899670$ and, besides the eight equality constraints, constraints 10, 11, 12, 14 and 16 active. An initial feasible point for the inequality constraints was taken as $\tilde{x}^0 = \tilde{e}$ with $f(\tilde{x}^0) = 414$.

$i \backslash j$	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
1	1			1			1	1								1
2		1	1				1			1						
3			1				1		1	1				1		
4				1			1				1				1	
5					1	1				1		1				1
6						1		1							1	
7							1				1		1			
8								1		1					1	
9									1			1				1
10										1				1		
11											1		1			
12												1		1		
13													1	1		
14														1		
15															1	
16																1

Table 4.3. A_{ij} values for Colville: Problem 7.

(v) Hexagon Problem. (Murray (1969c)).

$$\begin{aligned}
&\text{Minimize} && f(\underline{x}) = -\frac{1}{2}(x_2 x_6 + x_7(x_3 - x_1) + x_8(x_5 - x_3) - x_4 x_9) \\
&\text{subject to} && g_i(\underline{x}) = 1 - (x_i^2 + x_{i+5}^2) \geq 0, \quad i=1,2, \\
&&& g_3(\underline{x}) = 1 - x_3^2 \geq 0, \\
&&& g_i(\underline{x}) = 1 - (x_i^2 + x_{i+4}^2) \geq 0, \quad i=4,5, \\
&&& g_6(\underline{x}) = 1 - [(x_2 - x_1)^2 + (x_7 - x_6)^2] \geq 0, \\
&&& g_7(\underline{x}) = 1 - [(x_3 - x_1)^2 + x_6^2] \geq 0, \\
&&& g_{i+4}(\underline{x}) = 1 - [(x_i - x_1)^2 + (x_{i+4} - x_6)^2] \geq 0, \\
&&& \quad \quad \quad i=4,5, \\
&&& g_{10}(\underline{x}) = 1 - [(x_3 - x_2)^2 + x_1^2] \geq 0, \\
&&& g_{i+7}(\underline{x}) = 1 - [(x_i - x_2)^2 + (x_{i+4} - x_7)^2] \geq 0, \\
&&& \quad \quad \quad i=4,5, \\
&&& g_{i+9}(\underline{x}) = 1 - [(x_i - x_3)^2 + x_{i+4}^2] \geq 0, \quad i=4,5, \\
&&& g_{15}(\underline{x}) = 1 - [(x_5 - x_4)^2 + (x_9 - x_8)^2] \geq 0, \\
&&& g_{i+15}(\underline{x}) = x_i \geq 0, \quad i=\overline{1,7}, \\
&&& g_{i+15}(\underline{x}) = -x_i \geq 0, \quad i=8,9 \\
&&& g_{25}(\underline{x}) = x_3 - x_2 \geq 0, \\
&&& g_{26}(\underline{x}) = x_3 - x_4 \geq 0, \\
&&& g_{27}(\underline{x}) = x_2 - x_1 + 0.1 \geq 0, \\
&&& g_{28}(\underline{x}) = x_4 - x_5 + 0.1 \geq 0.
\end{aligned}$$

A constrained minimum (not unique) is at $\underline{x}^* =$
 $(0.4023507, 0.9390534, 1.0, 0.9390534, 0.4023507, 0.5,$
 $0.3437714, -0.3437714, -0.5)$ with $f(\underline{x}^*) = -0.6749814$
and constraints 2,3,4,8,9 and 12 active. An initial
feasible point is $\underline{x}^0 = (0.3, 0.6, 0.9, 0.6, 0.3, 0.4,$

0.4, -0.4, -0.4) with $f(\tilde{x}^0) = -0.48$. This problem maximizes the area of a hexagon (actually minimizes the negative of the area) in which no two vertices are separated by more than one unit. The first fifteen constraints describe this restriction while the remaining constraints preserve the topology of the hexagon.

The performance of Algorithm 4.1 is evaluated by comparing it with an extrapolated barrier function method and a standard barrier function method on the basis of numbers of iterations and function evaluations (see Remarks 4.2(ii)) required by each minimization. The extrapolated barrier function methods have been developed extensively by Fiacco and McCormick (1966, 1968) and Lootsma (1968a, 1970) for the barrier functions (1.8) and (1.9). The methods predict the minimum $\tilde{x}^{\hat{k}+1}$ at $r^{\hat{k}+1}$ (and also the problem solution, $\tilde{x}^*$, if required) by treating $\tilde{x}$ as a function of r and extrapolating from the $\tilde{x}^k$, $k=1, 2, \dots, \hat{k}$, (generated by r^k , $k=1, 2, \dots, \hat{k}$) using a polynomial expansion for $\tilde{x}(r)$. Since extrapolation cannot be generalized immediately to the family of barrier functions (2.18), the logarithmic transformation (1.8) is used with a quadratic extrapolation from $\tilde{x}^{k-2}$, $\tilde{x}^{k-1}$ and $\tilde{x}^k$ to predict an estimate of $\tilde{x}^{k+1}$, $k > 2$.

Linear extrapolation is used to predict an estimate of $\tilde{x}^3$ from $\tilde{x}^1$ and $\tilde{x}^2$. Fletcher and McCann (1969) have suggested that quadratic extrapolation is sufficient since little improvement in the estimate of $\tilde{x}^{k+1}$ is obtained from higher order approximations. The log-log barrier function (2.22) is used as the standard barrier function transformation and in all cases the r^k -sequence is generated by $r^k = 10^{1-k}$, $k=1,2,\dots$. Algorithm 5.1, based on the Davidon-Fletcher-Powell method and developed in Chapter 5, is used for all unconstrained minimizations.

Convergence of Powell's method, when solving each subproblem of (4.1) (i.e. step (iii) of Algorithm 4.1), is determined with $\epsilon_1 = 10^{-8}$ while convergence of Algorithm 5.1 is determined with tests of the form

$$||\nabla T(\tilde{x}^\alpha, \tilde{\theta}^k, \tilde{\sigma}^k, r^k)|| \leq n^2 10^{-13} \quad \text{or} \quad \max_{i=1, n} \frac{|\tilde{x}_i^{\alpha+1} - \tilde{x}_i^\alpha|}{1 + |\tilde{x}_i^{\alpha+1}|} \leq 10^{-8}, \quad (4.16)$$

where α is an iteration count for Algorithm 5.1. The second condition is required for small values of r^k when it becomes impossible to satisfy the first condition (see §2.5).

The results are presented in Tables 4.4-4.8 in the same format as Table 4.2 except that the "Reset" column is replaced by columns headed "B.F." and "E.B.F." giving the comparative results of the barrier function and extrapolated barrier function methods, respectively. Extrapolated results are not given for Colville: Problem 7 since the linear equality constraints, treated by Powell's method, prevent the use of extrapolation. The barrier function results are continued until the accuracy of Algorithm 4.1 has been attained.

Algorithm 4.1							B.F.			E.B.F.	
<i>k</i>	κ	un	in	ac	it	ev	<i>k</i>	it	ev	it	ev
1	1	15	0	0	25	66	1	25	66	16	40
2	1	15	0	0	13	27	2	13	27	12	29
3	1	4	10	1	10	22	3	11	32	14	30
	2				3	5	4	10	22	12	26
4	1	1	11	3	5	19	5	7	15	9	21
	2				2	12	6	19	41	7	13
5	1	0	11	4	4	10	7	6	13	3	6
	2				2	3	8	4	9	2	4
	3				2	3	9			2	4
							10			2	4
Totals							95	225		79	177

Table 4.4. Colville Problem 1.

Algorithm 4.1							B.F.			E.B.F.	
k	κ	un	in	ac	it	ev	k	it	ev	it	ev
1	1	20	0	0	34	74	1	34	74	31	61
2	1	20	0	0	26	55	2	26	55	22	48
3	1	10	3	7	20	41	3	21	47	21	44
	2				7	13	4	19	42	16	33
	3				2	3	5	18	39	13	29
4	1	2	7	11	7	13	6	21	48	7	15
	2				4	7	7	18	39	2	3
	3				2	3	8	24	53	2	4
5	1	1	8	11	4	7	9	21	43	2	4
	2				3	5	10			2	4
6	1	0	9	11	3	5					
	2				2	3					
Totals						114	229	202	440	118	245

Table 4.5. Colville Problem 2.

Algorithm 4.1							B.F.			E.B.F.	
k	κ	un	in	ac	it	ev	k	it	ev	it	ev
1	1	16	0	0	14	34	1	14	34	13	27
2	1	16	0	0	12	26	2	12	26	12	25
3	1	7	7	2	11	25	3	11	23	16	31
	2				3	5	4	8	17	6	12
4	1	0	11	5	4	7	5	13	31	2	3
	2				2	3	6	9	21	3	7
	3				1	1	7	4	8	2	4
	4				2	3	8	3	8	2	4
	5				2	3	9	2	5	2	4
							10			2	6
Totals						51	107	76	173	60	123

Table 4.6. Colville Problem 3.

Algorithm 4.1							B.F.			
k	κ	un	in	ac	it	ev	k	κ	it	ev
1	1	32	0	8	30	65	Same as for Algorithm 4.1 until k = 3.			
	2				9	18				
	3				6	11				
	4				6	11				
	5				4	7				
	6				3	5				
	7				3	5				
	8				2	3				
	9				2	3				
	10				2	3				
	11				2	3				
2	1	32	0	8	14	30	3	1	12	26
	2				6	11				
	3				3	5				
	4				3	5				
	5				2	3				
	6				2	3				
	7				2	3				
	8				2	3				
3	1	8	23	9	11	23	4	1	8	17
	2				4	7				
	3				3	5				
	4				2	3				
	5				2	3				
	6				2	3				
4	1	2	25	13	4	7	5	1	6	13
	2				3	5				
	3				2	3				
	4				2	3				
	5				2	3	6	1	8	17
	6				2	3				
5	1	0	27	13	3	5	7	1	6	14
	2				3	5				
	3				2	3				
	4				1	1				
	5				2	3				
Totals					151	279			187	367

Table 4.7. Colville Problem 7.

Algorithm 4.1							B.F.			E.B.F.	
k	κ	un	in	ac	it	ev	k	it	ev	it	ev
1	1	28	0	0	10	20	1	10	20	12	27
2	1	28	0	0	23	50	2	23	50	24	51
3	1	16	12	0	15	33	3	15	36	16*	35*
4	1	6	20	2	13	29	4	14	30	13*	28*
	2				4	7	5	11	27	7	14
5	1	0	22	6	4	7	6	9	20	2	3
	2				3	5	7	11	27	2	4
							8	7	16	2	7
							9	16	36	2	4
							10			2	3
Totals						72	151	116	262	82	176

Table 4.8. Hexagon Problem.

* Extrapolated point infeasible.

Two observations can be made concerning the results. For each problem, Algorithm 4.1 finds the solution in fewer unconstrained minimization iterations and fewer function evaluations than both the standard barrier function method and the extrapolated barrier function method. The extrapolated barrier function method, although improving on the performance of the standard barrier function method, still suffers from the conditioning and scaling problems of §2.5. The slower logarithmic barrier function convergence is also evident in that smaller r^k -values are required to attain the same accuracy for $\tilde{x}^*$. As the inequality constraints are classified, the unconstrained minimizations appear to become simpler reflecting the convergence properties of Powell's method and supporting the suggestion that $(\tilde{x}^k, \tilde{\theta}^k)$ are good initial estimates of $(\tilde{x}^{k+1}, \tilde{\theta}^{k+1})$.

UNCONSTRAINED AND LINEARLY CONSTRAINED METHODS5.1 Introduction.

In this chapter a number of methods for unconstrained and linearly constrained minimization will be considered. The emphasis will be on the computational properties of the methods with particular reference to their use as a computational tool for implementing transformation methods. Barrier function transformations, discussed in Chapter 2, which are defined only on a subset of the vector space (i.e. on $\{\underline{x}: g_i(\underline{x}) > 0, i=\overline{1, \ell}\}$) generate a number of special problems for unconstrained methods. Aspects of these problems have been discussed by Murray (1969c), Fiacco and McCormick (1968) and Fletcher and McCann (1969).

Brief details of unconstrained gradient minimization methods on E^n will be discussed in §5.2. In §5.3, three projection methods for linearly constrained optimization will be examined. In §5.4, a projection algorithm, based on the methods of §5.2 and §5.3 will be constructed with computational properties particularly attractive in minimizing $T(\underline{x}, \underline{\theta}, \underline{\sigma}, r)$ given by (4.2). Convergence properties of the algorithm can be deduced from the methods on which it is based. Numerical results, generated by this algorithm, will be presented in §5.5.

5.2 Unconstrained Methods.

Since many theoretical results for transformation methods derived in previous chapters have required the existence of continuous derivatives of problem functions, it is natural to make use of gradient methods for minimizing these functions. This is also supported by the efficiency of gradient methods when compared with direct search methods. Fiacco and McCormick (1968) have discussed the use of Newton's method requiring second partial derivations of the problem functions. This will not be pursued here. Attention will be restricted to first order methods for minimizing a function $f(\underline{x})$, $\underline{x} \in E^n$, which take the form at the $(\alpha+1)$ th iteration

$$\underline{d}^\alpha = -H^\alpha \nabla f(\underline{x}^\alpha) , \quad (5.1)$$

$$\underline{x}^{\alpha+1} = \underline{x}^\alpha + \lambda^\alpha \underline{d}^\alpha , \quad \lambda^\alpha > 0 , \quad (5.2)$$

$$H^{\alpha+1} = H^\alpha + \Gamma^\alpha , \quad (5.3)$$

where H^α is an $(n \times n)$ matrix operator, $\underline{d}^\alpha$ is the direction of linear search, λ^α is the step-length in direction $\underline{d}^\alpha$ and Γ^α is the correction matrix for H^α based on information at $\underline{x}^\alpha$ and $\underline{x}^{\alpha+1}$. It is usual to require that H^α be symmetric, positive definite (to ensure $\underline{d}^\alpha$ is "downhill" for $f(\underline{x})$ at $\underline{x}^\alpha$) and to approximate, in some sense, the inverse Hessian of $f(\underline{x})$ at $\underline{x}^\alpha$ - hence the name quasi-Newton methods. The correction, Γ^α , must be chosen to preserve these properties. A choice of λ^α and Γ^α

characterize a particular method. The derivation and convergence of these methods is usually based on their application to quadratic functions and the justification for applying the methods to general functions is provided by the observation that in a sufficiently small neighbourhood of the minimum, a general function can be approximated arbitrarily closely by a quadratic.

Two classes of methods based on (5.1), (5.2) and (5.3) are evident.

1. Full linear search methods.

The first class, described comprehensively by Huang (1969) and Huang and Levy (1969) (see §1.1 for further references) choose λ^α by a full linear search in the direction $\underline{d}^\alpha$ so that

$$f(\underline{x}^\alpha + \lambda^\alpha \underline{d}^\alpha) = \min_{\lambda} f(\underline{x}^\alpha + \lambda \underline{d}^\alpha) . \quad (5.4)$$

One of the oldest and most effective of these methods is due to Davidon (1959) and Fletcher and Powell (1963). Here Γ^α is given by

$$\Gamma^\alpha = \frac{\underline{p}^\alpha (\underline{p}^\alpha)^T}{(\underline{p}^\alpha)^T \underline{q}^\alpha} - \frac{H^\alpha \underline{q}^\alpha (\underline{q}^\alpha)^T H^\alpha}{(\underline{q}^\alpha)^T H^\alpha \underline{q}^\alpha} , \quad (5.5)$$

where $\underline{p}^\alpha = \underline{x}^{\alpha+1} - \underline{x}^\alpha = \lambda^\alpha \underline{d}^\alpha$ and $\underline{q}^\alpha = \nabla f(\underline{x}^{\alpha+1}) - \nabla f(\underline{x}^\alpha)$.

It can be shown that (5.5) preserves symmetry and positive definiteness of successive H^α matrices if H^0 is symmetric and positive definite. For quadratic functions, (5.5) also ensures convergence to the minimum in not more than n

iterations. Powell (1969b) has also proved convergence of this method for functions with strictly positive definite Hessian matrices at all points, $\tilde{x} \in E^n$.

The most serious disadvantage of these methods is the requirement of a full linear search to determine λ^α . This is an expensive procedure in terms of function evaluations. For this reason, the full linear search is often replaced by a partial linear search in which λ^α is chosen to minimize an approximating (e.g. extrapolating or interpolating) function and in addition to satisfy the descent property

$$f(\tilde{x}^\alpha + \lambda^\alpha \tilde{d}^\alpha) < f(\tilde{x}^\alpha) . \quad (5.6)$$

Although the theoretical properties depending on (5.4) (e.g. the generation of conjugate search directions and convergence in n iterations) are no longer valid, the partial linear search does not appear to seriously affect the computational convergence of the methods especially when care is taken in choosing the approximating functions. Greenstadt (1970) has reported that the use of a partial linear search can render the methods less reliable. In our applications, using $\tilde{x}^{k-1}$ as a reasonable initial estimate of the minimum at r^k , the failure of the methods appears to be less likely. An important feature of the partial linear search is the significant reduction in the number of function evaluations per iteration.

2. Predicted step-length methods.

The second class of gradient methods (see §1.1 for references), all of the quasi-Newton type, do not choose λ^α by a linear search but instead predict, using previous information, λ^α to satisfy the descent property (5.6). The correction matrix, Γ^α , is chosen so that $H^{\alpha+1}$ converges, in some sense, to the inverse Hessian of the function at the minimum. Such a property provides some guarantee of a fast ultimate convergence of the methods. The methods have the obvious advantage of requiring few function evaluations per iteration but they usually require an increased number of iterations for convergence. In predicting λ^α and choosing Γ^α , the methods incorporate a number of computational devices designed to ensure the "stability" of the method (i.e. $\tilde{d}^\alpha$ is downhill for $f(\tilde{x})$ at $\tilde{x}^\alpha$ and (5.6) is satisfied). A method of particular interest in §5.4, due to Murtagh and Sargent (1968), chooses Γ^α according to the rank-one correction

$$\Gamma^\alpha = \frac{\tilde{z}^\alpha (\tilde{z}^\alpha)^T}{c^\alpha} \quad (5.7)$$

where $\tilde{z}^\alpha = \tilde{p}^\alpha - H^\alpha \tilde{q}^\alpha$ and $c^\alpha = (\tilde{q}^\alpha)^T \tilde{z}^\alpha$ with $\tilde{p}^\alpha$ and $\tilde{q}^\alpha$ defined in (5.5). To ensure stability, (5.7) is only applied when the criterion

$$(\tilde{z}^\alpha)^T \nabla f(\tilde{x}^\alpha) / c^\alpha < 0 \quad (5.8)$$

is satisfied. If (5.8) is violated, Murtagh and Sargent suggest correcting H^α by (5.7) with $c^\alpha = (\tilde{z}^\alpha)^T \tilde{z}^\alpha$, which will preserve positive definiteness properties.

Other methods based on (5.7) have been discussed by Davidon (1968), Powell (1969a) and Fletcher (1969c).

5.3 Projection Methods for Linearly Constrained Problems.

A particular feature of Problem 1.1 which can influence the choice of an unconstrained method is found in the occurrence of linear constraints. Such constraints often appear as simple upper and lower variable bounds of the form $\beta_{\ell} \leq x \leq \beta_u$. The linearly constrained problem (Problem 1.4) may, in general, be written as

$$\begin{aligned} & \text{minimize } f(x), \quad x \in E^n, \\ & \text{subject to } C_s^T x \geq b, \end{aligned} \quad (5.9)$$

where C_s is an $(n \times s)$ matrix with columns, c_i , $i=\overline{1,s}$, (the normalised vectors of constraint coefficients) and b is the vector of right-hand sides corresponding to the linear constraints. Constraints (5.9) can also be written as $c_i^T x - b_i \geq 0$, $i=\overline{1,s}$. In solving Problem 1.4, the constraints could be included in barrier function transformations. This appears to be a retrograde step since no account is taken of the important linearity property and further nonlinearity is introduced by the transformation. Three related projection methods for solving Problem 1.4 will now be briefly described and details relevant to the linearly constrained minimization of transformation functions will be discussed.

At x^α , suppose that a basis, $L_t = \{i: c_i^T x^\alpha - b_i = 0, i=\overline{1,s}\}$, of t of the s linear constraints are active. Let

C_t be the $(n \times t)$ matrix formed by the t vectors $\underline{c}_i$, $i \in L_t$. For simplicity, it is assumed that C_t has rank $t \leq n$, although the methods can be modified to handle linearly dependent constraints. The projection methods, based on (5.2) and (5.3), modify (5.1) to choose $\underline{d}^\alpha$, and consequently by (5.2), $\underline{x}^{\alpha+1}$, in the $(n-t)$ dimensional constraint manifold of active constraints until one of the following conditions is satisfied:

- (i) the solution is found, or
- (ii) it is determined that a currently active constraint should be removed from the basis, or
- (iii) it is determined that a new constraint should be added to the basis.

The projection methods replace (5.1) by

$$\hat{\underline{d}}^\alpha = P_t^\alpha \underline{d}^\alpha = -P_t^\alpha H^\alpha \nabla f(\underline{x}^\alpha) \quad , \quad (5.10)$$

where P_t^α is a non-Euclidean projection operator of the form

$$P_t^\alpha = \{I - H^\alpha C_t (C_t^T H^\alpha C_t)^{-1} C_t^T\} \quad , \quad (5.11)$$

which can easily be shown to project E^n onto the constraint manifold defined by the constraints of the basis L_t . Derivations and discussions of this operator, in the context of constrained optimization, are given by Goldfarb (1969b), Murtagh and Sargent (1969) and Beltrami (1970).

The oldest method, due to Rosen (1960), defines $H^\alpha \equiv I$ in (5.10) and (5.11) and $\Gamma^\alpha \equiv 0$ in (5.3). P_t^α then

becomes a Euclidean projection operator and (5.10) describes the projection of the steepest descent direction. The algorithm is plagued by problems associated with the steepest descent directions and its practical use is therefore limited. It does, however, provide a basis for the following more efficient gradient methods based on unconstrained conjugate gradient and quasi-Newton techniques.

1. The Goldfarb-Lapidus method.

The Goldfarb-Lapidus method (Goldfarb and Lapidus (1968) and Goldfarb (1969b)), based on the Davidon-Fletcher-Powell method, selects Γ^α according to (5.5) noting that symmetry and positive definiteness are preserved. The method defines $\hat{H}_t^\alpha$ by

$$\hat{H}_t^\alpha = P_t^\alpha H^\alpha,$$

where $\hat{H}_t^\alpha$ can be interpreted as the orthogonal projection of H^α onto the constraint manifold defined by L_t and therefore

$$\hat{d}_t^\alpha = -\hat{H}_t^\alpha \nabla f(\tilde{x}^\alpha).$$

Computationally, $\hat{H}_t^\alpha$ is stored and no record is kept of H^α . A full linear search is used to determine λ^α in the interval $(0, \beta_j^\alpha]$, where

$$\beta_j^\alpha = \frac{b_j - \tilde{c}_j^T \tilde{x}^\alpha}{\tilde{c}_j^T \tilde{d}_t^\alpha} = \min_{i \notin L_t} \left\{ \frac{b_i - \tilde{c}_i^T \tilde{x}^\alpha}{\tilde{c}_i^T \tilde{d}_t^\alpha} : \tilde{c}_i^T \tilde{d}_t^\alpha < 0 \right\} \quad (5.12)$$

is the distance to the nearest linear constraint. Two possibilities arise in the linear search:

- (i) If $\lambda^\alpha \in (0, \beta_j^\alpha]$ satisfies (5.4), then $\hat{H}_t^\alpha$ is corrected according to (5.5). It can be shown that this correction preserves the projection properties of $\hat{H}_t^\alpha$ (Goldfarb (1969b)).
- (ii) If $\lambda^\alpha = \beta_j^\alpha$, then $\underline{c}_j^T \underline{x}^{\alpha+1} - b_j = 0$ and this constraint must be added to the basis. A rank-one correction can be derived for obtaining $\hat{H}_{t+1}^\alpha$ (Goldfarb (1969b)) by effectively removing curvature information orthogonal to the j th constraint.

Constraints currently in the basis can be recognised as inactive by considering the behaviour of the corresponding Lagrange multipliers in a manner similar to that used in §4.4. A full discussion of the procedures for removing an active constraint can be found in Goldfarb (1969b).

Goldfarb (1969b) has shown that when applied to a quadratic program and while the constraint basis remains unchanged, the method generates conjugate search directions and the solution is found in not more than $(n-t)$ iterations. He has also shown that if $(n-t)$ steps are taken without a basis change then $\hat{H}_t^\alpha$ will be the projection of the inverse Hessian within the $(n-t)$ iterations.

Rounding errors, which accumulate during the repeated application of the recurrence formulae, cause $\hat{\underline{d}}^\alpha$ to have small components that do not lie in the constraint manifold, and therefore, yield an $\underline{x}^{\alpha+1}$ which violates the active constraints. Two correction procedures have been proposed:

- (i) Rosen's error - squaring procedure (Rosen (1960))

can be used to generate $\hat{x}^{\alpha+1}$ satisfying

$$||C_t^T \hat{x}^{\alpha+1} - b_t|| = O(\delta^2) \text{ where } \max_{i \in L_t} \{c_i^T \hat{x}^{\alpha+1} - b_i\} = \delta.$$

- (ii) When rounding errors become large, $\hat{H}_t^\alpha$ can be reconstructed by setting $\hat{H}_0^\alpha = I$ and reforming the basis by adding the active constraints one at a time using the relevant recursion relation. The use of this procedure, referred to as basis reinversion, destroys all curvature information accumulated in previous iterations.

Two computational disadvantages are evident in the use of the Goldfarb-Lapidus method:

- (i) The Davidon-Fletcher-Powell formula (5.5) is not theoretically applicable unless (5.4) is satisfied. In general, when $\lambda^\alpha = \beta_j^\alpha$ and a new active constraint is added to the basis, $\hat{H}_t^\alpha$ can not be corrected by (5.5) and curvature information available at this step is ignored.
- (ii) The projection of H^α destroys curvature information orthogonal to the constraint manifold. Effects similar to those evident in Table 4.2 can be expected if constraints are repeatedly added and then removed from the basis.

2. The Murtagh-Sargent method.

A third projection method, developed by Murtagh and Sargent (1969), preserves the distinction between the projection operator, P_t^α , and the inverse Hessian estimate,

H^α , by storing the matrices H^α and $(C_t^T H^\alpha C_t)^{-1}$, the latter being used to compute P_t^α by (5.11). The method is based on the rank-one correction (5.7) which can be applied (provided the stability condition (5.8) is satisfied) for any arbitrary step-length. This, therefore, overcomes the first computational disadvantage of the Goldfarb-Lapidus method.

The Murtagh-Sargent method uses a predicted step-length method for determining λ^α satisfying (5.6) and $\lambda^\alpha \in (0, \beta_j^\alpha]$, where β_j^α is given by (5.12).

The following recursion formulae can be easily derived for correcting $(C_t^T H^\alpha C_t)^{-1}$ when:

(i) H^α is updated by (5.7);

$$(C_t^T H^{\alpha+1} C_t)^{-1} = \left\{ (C_t^T H^\alpha C_t) + \frac{1}{c^\alpha} C_t^T z^\alpha (C_t^T z^\alpha)^T \right\}^{-1} \quad (5.13)$$

and the right side can be computed by applying (3.45).

(ii) $\lambda^\alpha = \beta_j^\alpha$ and the j th constraint is added to the basis. Let $C_{t+1} = [C_t | c_j]$. Then

$$C_{t+1}^T H^\alpha C_{t+1} = \begin{bmatrix} A_{11} & A_{12} \\ A_{12}^T & A_{22} \end{bmatrix}$$

where $A_{11} = C_t^T H^\alpha C_t$, $A_{12} = C_t^T H^\alpha c_j$,

$A_{22} = c_j^T H^\alpha c_j$. Define $A_0 = A_{22} - A_{12}^T A_{11}^{-1} A_{12}$

noting that $A_{11}^{-1} = (C_t^T H^\alpha C_t)^{-1}$ and then

$$(C_{t+1}^T H^\alpha C_{t+1})^{-1} = \begin{bmatrix} B_{11} & B_{12} \\ B_{12}^T & B_{22} \end{bmatrix} \quad (5.14)$$

where $B_{11} = A_{11}^{-1} + (A_{11}^{-1} A_{12}) A_0^{-1} (A_{11}^{-1} A_{12})^T$,

$B_{12} = -A_0^{-1} (A_{11}^{-1} A_{12})$ and $B_{22} = A_0^{-1}$.

(5.15)

(iii) The t th constraint is removed from the basis.

Let $C_t = [C_{t-1} | c_t]$. Then

$$(C_t^T H^\alpha C_t)^{-1} = A^{-1} = B = \begin{bmatrix} B_{11} & B_{12} \\ B_{12}^T & B_{22} \end{bmatrix}$$

and solving (5.15) for A_{11}^{-1} gives

$$(C_{t-1}^T H^\alpha C_{t-1})^{-1} = B_{11} - B_{12} B_{22}^{-1} B_{12}^T. \quad (5.16)$$

As with the unconstrained methods based on (5.7), care must be taken to preserve positive definiteness of H^α when the correction is applied. Murtagh and Sargent have shown that for quadratic programming and a constant constraint basis, L_t , their method finds a stationary point in at most $(n-t+1)$ iterations.

Rounding error correction formulae similar to those used by the Goldfarb-Lapidus method can be derived for correcting the effects of rounding error. It should be noted that basis reinversion in this case involves no loss of curvature information since H^α is retained at all times and $(C_t^T H^\alpha C_t)^{-1}$ can be recomputed by starting with H^α and adding the active constraints one at a time using (5.14).

5.4 A Hybrid Projection Algorithm for Minimizing $T(\tilde{x}, \tilde{\theta}, \tilde{\sigma}, r)$.

Interest in projection methods is generated by the specific problem of minimizing the successive members

of (4.1) subject to the linear constraints of Problem 1.1. Before describing the proposed hybrid projection method, Algorithm 5.1, the special properties of (4.2) which influence its construction will be discussed.

1. Preservation of Hessian information.

The importance of retaining the inverse Hessian estimates when solving (4.1) has been demonstrated by the results of Table 4.2. Algorithm 5.1, like the Murtagh-Sargent approach, must therefore preserve the distinction between the projection operator and the inverse Hessian estimate. Basis reinversion in this case need not reset the inverse Hessian to the identity. It should be noted that Rosen's error-squaring procedure is inapplicable while inequality constraints remain unclassified since the corrected point, although in the constraint manifold, may not be in the feasible region with respect to the unclassified nonlinear inequality constraints.

2. Linear searches.

Murray (1969c) and Fletcher and McCann (1969) have discussed the problem of linear searches associated with minimizing barrier function transformations. They conclude that the interpolating and extrapolating functions should be chosen to reflect the behaviour of transformation functions as the boundary of the feasible region is approached. In the present context, this conclusion will not be as important since the barrier function terms of $T(\underline{x}, \underline{\theta}, \underline{g}, r)$ are gradually eliminated by classifying constraints.

In minimizing $T(\underline{x}, \underline{\theta}, \underline{\sigma}, r)$, the use of at least a partial linear search appears desirable. Conditioning problems discussed in §2.5 and the general difficulty of minimizing barrier function transformations prevent accurate predictions of step-length and methods of this type can therefore be expected to have difficulty satisfying the descent condition (5.6) in one function evaluation. Furthermore, it seems unwise to aggravate Hessian ill-conditioning associated with barrier function transformations by using a predicted step-length method involving further stability problems when correcting the inverse Hessian estimate. Corrections of the form (5.7) are, however, useful in allowing the Hessian to be corrected when active linear constraints are added to the basis and (5.5) is not applicable.

Motivated by these considerations, Algorithm 5.1 uses the Davidon-Fletcher-Powell correction (5.5) and incorporates a cubic extrapolation and interpolation procedure for partial linear searches. The projection facilities are taken, basically, from the Murtagh-Sargent approach. For simplicity, considerations of linear dependence of active linear constraints and reinversion of the constraint basis are omitted from the following algorithm description. They can be based directly on the Murtagh-Sargent approach. Explicit treatment of linear equality constraints is also omitted.

Algorithm 5.1.

At the $(\alpha+1)$ th iteration, begin at $\tilde{x}^\alpha$ with a positive definite, symmetric inverse Hessian estimate H^α , an active constraint basis L_t and the matrix $(C_t^T H^\alpha C_t)^{-1}$.

(i) If $t=0$ then set $\tilde{d}^\alpha = -H^\alpha \nabla f(\tilde{x}^\alpha)$ and go to (iii);

else compute

$$(a) \quad \tilde{\delta}^\alpha = -(C_t^T H^\alpha C_t)^{-1} C_t^T (-H^\alpha \nabla f(\tilde{x}^\alpha))$$

($\tilde{\delta}^\alpha$ is an estimate of the Lagrange multipliers corresponding to the active linear constraints).

$$(b) \quad \tilde{d}^\alpha = -H^\alpha \nabla f(\tilde{x}^\alpha) + H^\alpha C_t \tilde{\delta}^\alpha.$$

(ii) Compute $\gamma = -\frac{1}{2} \delta_j^\alpha / |b_{jj}|$ where $\delta_j^\alpha = \min_{i=1, \dots, t} \delta_i^\alpha$ and b_{jj} is the j th diagonal element of $(C_t^T H^\alpha C_t)^{-1}$.

If $||\tilde{d}^\alpha|| \leq \varepsilon_1$ and $\gamma \leq \varepsilon_1$ then stop;

else if $||\tilde{d}^\alpha|| > \gamma$ then go to (iv);

else remove the j th constraint from

L_t and C_t .

(a) Swop j th and t th rows and columns of $(C_t^T H^\alpha C_t)^{-1}$.

(b) Compute $(C_{t-1}^T H^\alpha C_{t-1})^{-1}$ using (5.16).

(c) Go to (i) with $t = t-1$.

(iii) If $||\nabla f(\tilde{x}^\alpha)|| < \varepsilon_2$ or $\max_{i=1, \dots, n} \frac{|x_i^\alpha - x_i^{\alpha-1}|}{1 + |x_i^\alpha|} < \varepsilon_3$

then stop; else go to (iv).

(iv) If $0 \leq t < s$ then compute $\lambda_M = \beta_j^\alpha$ using (5.12); else set $\lambda_M = \infty$. By a partial linear search determine λ^α satisfying

$$f(\tilde{x}^\alpha + \lambda^\alpha \tilde{d}^\alpha) = \min_{0 < \lambda \leq \lambda_M} f(\tilde{x}^\alpha + \lambda \tilde{d}^\alpha).$$

Set $\tilde{x}^{\alpha+1} = \tilde{x}^\alpha + \lambda^\alpha \tilde{d}^\alpha$, $\tilde{p}^\alpha = \lambda^\alpha \tilde{d}^\alpha$, $\tilde{q}^\alpha = \nabla f(\tilde{x}^{\alpha+1}) - \nabla f(\tilde{x}^\alpha)$. If $\lambda^\alpha = \lambda_M (\neq \infty)$ then go to (v); else go to (vii).

(v) Add the j th constraint to the basis by applying (5.14). Go to (vi) with $t = t+1$.

(vi) Murtagh-Sargent rank-one correction.

If (5.8) is satisfied compute $H^{\alpha+1}$ using (5.7); else compute $H^{\alpha+1}$ using (5.7) with $c^\alpha = (\tilde{z}^\alpha)^T \tilde{z}^\alpha$. Compute $(C_t^T H^{\alpha+1} C_t)^{-1}$ using (5.13) and go to (i) with $\alpha = \alpha+1$.

(vii) Davidon rank-two correction.

If $\kappa_1 = (\tilde{q}^\alpha)^T H^\alpha \tilde{q}^\alpha = 0$ or $\kappa_2 = (\tilde{p}^\alpha)^T \tilde{p}^\alpha = 0$ then go to (vi);

else compute the Davidon correction as two rank-one corrections, viz.

$$(a) \text{ Compute } \bar{H}^\alpha = H^\alpha - \frac{H^\alpha \tilde{q}^\alpha (H^\alpha \tilde{q}^\alpha)^T}{\kappa_1}$$

and $(C_t^T \bar{H}^\alpha C_t)^{-1}$ using (5.13) with

$$\tilde{z}^\alpha = H^\alpha \tilde{q}^\alpha \text{ and } c^\alpha = -\kappa_1.$$

$$(b) \text{ Compute } H^{\alpha+1} = \bar{H}^\alpha + \frac{\tilde{p}^\alpha (\tilde{p}^\alpha)^T}{\kappa_2} \text{ and}$$

$(C_t^T H^{\alpha+1} C_t)^{-1}$ using (5.13) with

$$\tilde{z}^\alpha = \tilde{p}^\alpha \text{ and } c^\alpha = \kappa_2.$$

Go to (i) with $\alpha = \alpha+1$.

The repeated minimization of $T(\tilde{x}, \tilde{\theta}, \tilde{\sigma}, r)$ by

Algorithm 5.1 has three beneficial byproducts:

- (i) The linear constraint basis is usually formed early in the sequence of minimizations and remains constant in following minimizations. The effect of a large constraint basis is evident in the reduced number of iterations for convergence in each minimization although more computation per iteration is required.
- (ii) It is possible that at $\tilde{x}^k$, the solution of (4.1) at r^k , the basis may incorporate linear constraints which, for smaller values of r^k , are inactive. Such constraints are said to be pseudo-active. This situation is demonstrated by the problem; minimize x subject to $x \geq 0$ and $x \leq \frac{1}{2}$ where the first constraint is included in a log barrier function and the second is treated by projection. The unconstrained minimum of the transformation function is $x^k = r^k$ and for $r^k \geq \frac{1}{2}$, the second constraint can be seen to be pseudo-active. In general, for any linear constraint $\tilde{c}_i^T \tilde{x} - b_i \geq 0$, active at $\tilde{x}^k$, it is necessary that $\tilde{c}_i^T \nabla T(\tilde{x}^k, \tilde{\theta}^k, \tilde{\sigma}^k, r^k) \geq 0$. For such a constraint, $\rho_i < r^k$ can be determined to satisfy

$$\tilde{c}_i^T \nabla T(\tilde{x}^k, \tilde{\theta}^k, \tilde{\sigma}^k, \rho_i) = 0 \quad (5.17)$$

provided

$$\tilde{c}_i^T \sum_{j=1}^{\ell} \frac{d\phi_j(g_j(\tilde{x}^k))}{dg_j} \nabla g_j(\tilde{x}^k) > 0 \quad (5.18)$$

since otherwise

$$0 < \tilde{c}_i^T \nabla T(\tilde{x}^k, \tilde{\theta}^k, \tilde{\sigma}^k, r^k) < \tilde{c}_i^T \nabla T(\tilde{x}^k, \tilde{\theta}^k, \tilde{\sigma}^k, \rho_i)$$

for $\rho_i < r^k$. Solving (5.17) for ρ_i gives

$$\rho_i = - \frac{c_i^T \{ \nabla f(\tilde{x}^k) + 2\Lambda(\tilde{x}^k) S_k \tilde{\theta}^k \}}{c_i^T \sum_{j=1}^{\ell} \frac{d\phi_j(g_j(\tilde{x}^k))}{dg_j} \nabla g_j(\tilde{x}^k)}$$

and, using (5.18), it follows that ρ_i is positive only if

$$c_i^T \{ \nabla f(\tilde{x}^k) + 2\Lambda(\tilde{x}^k) S_k \tilde{\theta}^k \} < 0. \quad (5.19)$$

The inequalities (5.18) and (5.19) are now used as tests in an attempt to recognise pseudo-active constraints in the basis at $\tilde{x}^k$. The next r^k value is determined by

$$r^{k+1} = \gamma \min_i (\rho_i), \quad 0 < \gamma < 1,$$

for c_i , $i \in L_k$, satisfying (5.18) and (5.19).

Constraints satisfying (5.18) and (5.19) are removed from the basis since

$$c_i^T \nabla T(\tilde{x}^k, \tilde{\theta}^k, \tilde{\sigma}^k, r^{k+1}) < 0.$$

It is clear that because of the nonlinearity of $T(\tilde{x}, \tilde{\theta}, \tilde{\sigma}, r)$, these constraints may become active again and satisfy $c_i^T \nabla T(\tilde{x}^{k+1}, \tilde{\theta}^{k+1}, \tilde{\sigma}^{k+1}, r^{k+1}) \geq 0$.

(iii) Although Rosen's error-squaring procedure is inapplicable while $\ell \neq 0$, it can be applied after all nonlinear inequalities have been classified thus ensuring an ultimate solution free from excessive rounding error.

It is interesting to observe that in solving Problem 1.1, Powell's method and the projection method are used to ensure that active constraints are strictly satisfied at $\tilde{x}^k$. Convergence of the GLM estimates for the linear constraints may be deduced by applying Theorem 4.1 since there is no restriction on the method of satisfying equality constraints be they linear or nonlinear.

5.5 Numerical Results.

Numerical results of using Algorithm 5.1 to perform the linearly constrained minimizations of $T(\tilde{x}, \tilde{\theta}, \tilde{g}, r)$ are now given. The problems described in §4.7 are used as test problems.

Some difficulty has been experienced in implementing Algorithm 5.1 due to the use of convergence criteria based on $||P_t^\alpha H^\alpha \nabla f(\tilde{x}^\alpha)||$. Because of the scaling effect of H^α , it has been found necessary to use, in step (ii) of the algorithm description, a test of the form

$$||P_t \nabla f(\tilde{x}^\alpha)|| \leq n^2 10^{-13} \quad \text{and} \quad \gamma \leq 10^{-8},$$

where $P_t = \{I - C_t(C_t^T C_t)^{-1} C_t^T\}$ is a Euclidean projection operator projecting E^n onto the constraint manifold. This is still not entirely satisfactory and it is believed that improved convergence tests will provide a more reliable algorithm.

Tables 5.1 - 5.5 show the results of treating linear inequality constraints by Algorithm 5.1. The column headings of §4.7 are used and a further column headed "Linear" gives the numbers of linear inequality constraints remaining inactive or active (i.e. members of the constraint basis) after each application of Algorithm 5.1.

k	κ	Nonlinear			Linear			
		un	in	ac	in	ac	it	ev
1	1	0	0	0	11	4	7	15

Table 5.1. Colville: Problem 1. (Projection).

Remark 5.1. Nine figure accuracy for $f(\underline{x}^*)$ is obtained by Algorithm 5.1 in solving Colville: Problem 1. Results of an apparent eight figure accuracy for $f(\underline{x}^*)$, provided by the Goldfarb-Lapidus and Murtagh-Sargent methods have been published by Murtagh and Sargent (1969), p. 243.

Goldfarb-Lapidus: 10 iterations; 13 function evaluations.

Murtagh-Sargent : 7 iterations; 7 function evaluations.

k	κ	Nonlinear			Linear			
		un	in	ac	in	ac	it	ev
1	1	5	0	0	8	7	52*	151*
2	1	5	0	0	8	7	14	29
3	1	2	0	3	9	6	18	42
	2				9	6	3	5
4	1	0	0	5	9	6	5	9
	2				9	6	31*	61*
Totals							123	297

Table 5.2. Colville: Problem 2. (Projection).

* Unsatisfactory convergence tests.

Remarks 5.2. The comparatively poor results (see Table 4.5) are due to convergence difficulties associated with Algorithm 5.1 at $k = \kappa = 1$ and $k = 4$ and $\kappa = 2$. Apart from these two iterations the results are encouraging.

k	κ	Nonlinear			Linear		it	ev
		un	in	ac	in	ac		
1	1	6	0	0	7	3	11	22
2	1	6	0	0	7	3	7	16
3	1	2	4	0	7	3	6	13
4	1	0	4	2	7	3	3	6
	2				7	3	2	3
	3				7	3	1	1
	4				7	3	2	4
Totals							32	65

Table 5.3. Colville: Problem 3. (Projection).

Remark 5.3. The linear constraint basis is formed during the first unconstrained minimization (i.e. $k = \kappa = 1$) and remains unchanged during succeeding minimizations.

k	κ	Nonlinear			Linear		it	ev
		un	in	ac	in	ac		
1	1	0	0	8	29	5	19	32
	2				29	5	5	9
	3				29	5	4	7
	4				29	5	4	7
	5				29	5	3	5
	6				29	5	3	5
	7				29	5	2	3
	8				29	5	2	3
	9				29	5	2	3
	10				29	5	2	3
Totals							46	77

Table 5.4. Colville: Problem 7. (Projection).

Remarks 5.4. 1. The constraint basis is formed during the first unconstrained minimization (i.e. $k = \kappa = 1$) and remains unchanged during succeeding minimizations.

2. Because the present implementation of Algorithm 5.1 does not handle linear equality constraints, these constraints are handled by Powell's method. Further improvements can be expected by treating them also by projection.

k	κ	Nonlinear			Linear		it	ev
		un	in	ac	in	ac		
1	1	15	0	0	13	0	10	23
2	1	15	0	0	13	0	27	61
3	1	12	3	0	13	0	13	31
4	1	2	9	4	13	0	9	19
	2				13	0	3	5
5	1	0	9	6	13	0	5	9
	2				13	0	2	3
Totals							69	151

Table 5.5. Hexagon Problem. (Projection).

Remark 5.5. The results of Table 5.5 are similar to those of Table 4.8 since none of the linear constraints are active at the solution. They therefore have little influence on the barrier function transformation near the solution.

The results of Tables 5.1 - 5.5 suggest, in conflict with the comments of Lootsma (1970), p. 80, that if linear constraints are active at the solution, then improvements can be realised by treating these constraints using projection methods instead of barrier function transformations. This observation must be qualified by repeating that unsuitable convergence tests can nullify the advantages.

It can be seen from Tables 4.4 - 4.8 and Tables 5.1 - 5.5 that the use of a partial linear search provides an average of a little more than two function evaluations per iteration. This can be regarded as near optimal under the strategy of one evaluation to bound the minimum and a second evaluation at the interpolated minimum.

Bibliography

- Ablow, C.M. and Brigham, G. (1955). An analog solution of programming problems. *Op. Res.*, 3, pp 388-394.
- Akaike, H. (1959). On a successive transformation of probability distribution and its application to the analysis of the optimal gradient method. *Ann. Inst. Stat. Math.*, Tokyo, 11, p 1.
- Allran, R.R. and Johnsen, S.E.J. (1970). An algorithm for solving nonlinear programming problems subject to nonlinear inequality constraints. *Comp. J.*, 13(2), pp 171-177.
- Arrow, K.J. and Hurwicz, L. (1956). Reduction of constrained maxima to saddle point problems; in "Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability", ed. J. Neyman. University of California Press, pp 1-20.
- Arrow, K.J., Hurwicz, L. and Uzawa, H. (1958). "Studies in Linear and Nonlinear Programming". Stanford University Press.
- Arrow, K.J., Hurwicz, L. and Uzawa, H. (1961). Constraint qualifications in maximization problems. *Nav. Res. Log. Q.*, 8, pp 175-191.
- Bard, Y. (1968). On a numerical instability of Davidon-like methods. *Maths. Comp.*, 22 (103), pp 665-666.

- Bard, Y. (1970). Comparison of gradient methods for the solution of nonlinear parameter estimation problems. SIAM J. Numer. Anal., 7 (1), pp 157-186.
- Beale, E.M.L. (1955). On minimizing a convex function subject to linear inequalities. J. Roy. Stat. Soc., Ser. B, 17 (2), pp 173-184.
- Beale, E.M.L. (1959). On quadratic programming. Nav. Res. Log. Q., 6, pp 227-243.
- Beale, E.M.L. (1967). Numerical methods; in "Nonlinear Programming", ed. J. Abadie. North-Holland Publishing Co., p 150.
- Bellmore, M., Greenberg, H.J. and Jarvis, J.J. (1970). Generalized penalty function concepts in mathematical optimization. Op. Res., 18 (2), pp 229-252.
- Beltrami, E.J. (1969). A constructive proof of the Kuhn-Tucker multiplier rule. J. Math. Anal. Appl., 26, pp 297-306.
- Beltrami, E.J. (1970). "An Algorithmic Approach to Nonlinear Analysis and Optimization". Academic Press.
- Beltrami, E.J. and McGill, R. (1966). A class of variational problems in search theory and the maximum principle. Op. Res., 14 (2), pp 267-278.

- Box, M.J. (1966). A comparison of several current optimization methods and the use of transformations in constrained problems. *Comp. J.*, 9, pp 67-77.
- Box, M.J., Davies, D. and Swann, W.H. (1969). "Nonlinear Optimization Techniques". ICI Monograph, Oliver and Boyd.
- Bracken, J. and McCormick, G.P. (1968). "Selected Applications of Nonlinear Programming". Wiley.
- Broyden, C.G. (1965). A class of methods for solving nonlinear simultaneous equations. *Math. Comp.*, 19 (92), pp 577-593.
- Broyden, C.G. (1967). Quasi-Newton methods and their application to function minimization. *Math. Comp.*, 21 (19), pp 368-381.
- Broyden, C.G. (1970a). The convergence of a class of double-rank minimization algorithms; 1. general considerations. *J. Inst. Math. Appl.*, 6 (1) pp 76-90.
- Broyden, C.G. (1970b). The convergence of single-rank quasi-Newton methods. *Math. Comp.*, 24 (110), pp 365-382.
- Bui Trong Lieu and Huard, P. (1966). La méthode des centres dans un espace topologique. *Numér. Math.*, 8 (1), pp 56-67.

- Butler, T. and Martin, A.V. (1962). On a method of Courant for minimizing functionals. J. Math. Phys., 41, pp 291-299.
- Camp, G.D. (1955). Inequality-constrained stationary value problems. J. Op. Res. Soc. Am., 3, pp 548-550.
- Carrol, C.W. (1961). The created response surface technique for optimizing nonlinear restrained systems. Op. Res., 9 (2), pp 169-184.
- Charnes, A. and Lemke, C.E. (1954). Minimization of nonlinear separable convex functionals. Nav. Res. Log. Q., 1, pp 301-302.
- Cheney, E.W. and Goldstein, A.A. (1959). Newton's method for convex programming and Tchebycheff approximation. Numer. Math., 1, pp 253-268.
- Colville, A.R. (1968). A comparative study on nonlinear programming codes. IBM New York Scientific Centre Tech. Rep. No. 320-2949.
- Courant, R. (1943). Variational methods for the solution of problems of equilibrium and vibrations. Bull. Am. Math. Soc., 49, pp 1-23.
- Crockett, J.B. and Chernoff, H. (1955). Gradient methods of maximization. Pac. J. Math., 5, pp 33-50.

- Curry, H.B. (1944). The method of steepest descent for non-linear minimization problems. *Quart. Appl. Math.*, 2 (3), pp 250-261.
- Daniel, J.W. (1967a). The conjugate gradient method for linear and nonlinear operator equations. *SIAM J. Numer. Anal.*, 4, pp 10-26.
- Daniel, J.W. (1967b). Convergence of the conjugate gradient method with computationally convenient modifications. *Numer. Math.*, 10, pp 125-131.
- Dantzig, G.B. (1951). Maximization of a linear function of variables subject to linear inequalities; in "Activity Analysis of Production and Allocation", ed. T.C. Koopmans. Cowles Commission Monograph No. 13, Wiley.
- Dantzig, G.B. (1963). "Linear Programming and Extensions". Princeton University Press, Princeton, N.J.
- Davidon, W.C. (1959). Variable metric method for minimization. U.S.A.E.C. Research and Development Rep. ANL-5990, Argonne National Laboratory.
- Davidon, W.C. (1968). Variance algorithm for minimization. *Comp. J.*, 10, pp 406-410.
- Dorn, W.S. (1960). Duality in quadratic programming. *Quart. Appl. Math.*, 18, pp 155-162.

- Edelbaum, T.N. (1962). Theory of maxima and minima; in "Optimization Techniques", ed. G. Leitmann. Academic Press.
- Evans, J.P. and Gould, F.J. (1970). Stability in nonlinear programming. *Op. Res.*, 18 (1), pp 107-118.
- Faure, P. and Huard, P. (1966). Résultats nouveaux relatifs à la méthode des centres. Quatrième Conférence de Recherche Opérationnelle, Cambridge, Mass.
- Fiacco, A.V. (1967). Sequential unconstrained minimization methods for nonlinear programming. Ph.D. Dissertation, Northwestern University, Ill.
- Fiacco, A.V. (1970). Penalty methods for mathematical programming in E^n with general constraint sets. *J. Opt. Th. Appl.*, 6 (3), pp 252-268.
- Fiacco, A.V. and McCormick, G.P. (1963). Programming under nonlinear constraints by unconstrained minimization: a primal-dual method. Tech. Paper RAC-TP-96, Research Analysis Corporation.
- Fiacco, A.V. and McCormick, G.P. (1964a). The sequential unconstrained minimization technique for nonlinear programming: a primal-dual method. *Man. Sci.*, 10 (2), pp 360-366.
- Fiacco, A.V. and McCormick, G.P. (1964b). Computational algorithm for the sequential unconstrained minimization technique for nonlinear programming. *Man. Sci.*, 10 (2), pp 601-617.

- Fiacco, A.V. and McCormick, G.P. (1966). Extensions of SUMT for nonlinear programming: equality constraints and extrapolation. *Man. Sci.*, 12 (11), pp 816-829.
- Fiacco, A.V. and McCormick, G.P. (1967). The slacked unconstrained minimization technique for convex programming. *SIAM J. Appl. Math.*, 15 (3), pp 505-515.
- Fiacco, A.V. and McCormick, G.P. (1968). "Nonlinear Programming: Sequential Unconstrained Minimization Techniques". Wiley.
- Fletcher, R. (1965). Function minimization without evaluating derivatives - a review. *Comp. J.*, 8, pp 33-41.
- Fletcher, R. (1968). Programming under linear equality and inequality constraints. ICI Rep. No. MSDH/68/19.
- Fletcher, R. (Ed.) (1969a). "Optimization". Academic Press.
- Fletcher, R. (1969b). A review of methods for unconstrained optimization; in "Optimization", ed. R. Fletcher. Academic Press.
- Fletcher, R. (1969c). A new approach to variable metric algorithms. AERE Tech. Rep. No. 383.
- Fletcher, R. (1970). A general quadratic programming algorithm. AERE Tech. Rep. No. 401.

- Fletcher, R. and McCann, A.P. (1969). Acceleration techniques for nonlinear programming; in "Optimization", ed. R. Fletcher. Academic Press.
- Fletcher, R. and Powell, M.J.D. (1963). A rapidly convergent descent method for minimization. *Comp. J.*, 6, pp 163-168.
- Fletcher, R. and Reeves, C.M. (1964). Function minimization by conjugate gradients. *Comp. J.*, 7, pp 149-154.
- Forsythe, G.E. (1967). On the asymptotic directions of the s -dimensional optimum gradient method. Tech. Rep. No. CS61, Stanford University.
- Forsythe, G.E. and Motzkin, T.S. (1951). Asymptotic properties of the optimum gradient method (abstract). *Bull. Am. Math. Soc.*, 57, p 183.
- Frisch, K.R. (1955). The logarithmic potential method of convex programming. Memorandum May 13, 1955, University Institute of Economics, Oslo.
- Geoffrion, A.M. (1966). Strictly concave parametric programming, part 1: basic theory. *Man. Sci.*, 13, pp 244-253.
- Geoffrion, A.M. (1967). Strictly concave parametric programming, part 2: additional theory and computational considerations. *Man. Sci.*, 13 (5), pp 359-370.

- Goldfarb, D. (1969a). Sufficient conditions for the convergence of a variable metric algorithm; in "Optimization", ed. R. Fletcher. Academic Press.
- Goldfarb, D. (1969b). Extension of Davidon's variable metric method to maximization under linear inequality and equality constraints. SIAM J. Appl. Math., 17, pp 739-764.
- Goldfarb, D. (1970). A family of variable-metric methods derived by variational means. Math. Comp., 24, pp 23-26.
- Goldfarb, D. and Lapidus, L. (1968). Conjugate gradient method for nonlinear programming problems with linear constraints. I. & E.C. Fundamentals, 7 (1), pp 142-151.
- Goldstein, A.A. (1962). Cauchy's method of minimization. Numer. Math., 4, pp 146-150.
- Greenstadt, J. (1967). On the relative efficiencies of gradient methods. Math. Comp., 21, pp 360-367.
- Greenstadt, J. (1970). Variations on variable-metric methods. Math. Comp., 24, pp 1-25.
- Haarhoff, P.C. and Buys, J.D. (1970). A new method for the optimization of a nonlinear function subject to nonlinear constraints. Comp. J., 13 (2), pp 178-184.
- Hadley, G. (1964). "Nonlinear and Dynamic Programming". Addison-Wesley.

- Hartley, H.O. and Hocking, R.R. (1963). Convex programming by tangential approximation. *Man.Sci.*, 9, pp 600-612.
- Hestenes, M.R. (1966). "Calculus of Variations and Optimal Control Theory". Wiley.
- Huang, H.Y. (1969). Unified approach to quadratically convergent algorithms for function minimization. Aero-Astronautics Rep. No. 64, Rice University.
- Huang, H.Y. and Levy, A.V. (1969). Numerical experiments on quadratically convergent algorithms for function minimization. Aero-Astronautics Rep. No. 66, Rice University.
- Huard, P. (1964). Résolution de programmes mathématiques à contraintes non-linéaires par la méthode des centres. Note E.D.F., HR5690/3/317.
- Huard, P. (1967). Resolution of mathematical programming with nonlinear constraints by the method of centres; in "Nonlinear Programming", ed. J. Abadie. North-Holland Pub. Co., Amsterdam.
- Huard, P. (1968). Programmation mathématique convexe. R.I.R.O., 7, pp 43-59.
- John, F. (1948). Extremum problems with inequalities as subsidiary conditions: in "Studies and Essays". Courant Anniversary Volume, Interscience, pp 187-204.

- Kelley, H.J. (1962). Methods of gradients; in "Optimization Techniques", ed. G. Leitmann. Academic Press.
- Kelley, J.E. Jr. (1960). The cutting plane method for solving convex programs. J. Soc. Ind. Appl. Math., 8 (4), pp 703-712.
- Kowalik, J. (1966). Nonlinear programming procedures and design optimization. Acta Polytechnica Scandinavica, 13, Trondheim.
- Kowalik, J. and Osborne, M.R. (1968). "Methods for Unconstrained Problems". American Elsevier.
- Kowalik, J., Osborne, M.R. and Ryan, D.M. (1969). A new method for constrained optimization problems. Op. Res., 17 (6), pp 973-983.
- Kuhn, H.W. and Tucker, A.W. (1951). Non-linear Programming; in "Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability", ed. J. Neyman. University of California Press, pp 481-493.
- Lemke, C.E. (1962). A method for solution of quadratic programming problems. Man. Sci., 8, pp 442-453.
- Lootsma, F.A. (1967). Logarithmic programming: a method of solving nonlinear-programming problems. Philips Res. Rep., 22 (3), pp 329-344.

- Lootsma, F.A. (1968a). Extrapolation in logarithmic programming. Philips Res. Rep., 23, pp 108-116.
- Lootsma, F.A. (1968b). Constrained optimization via penalty functions. Philips Res. Rep., 23, pp 408-423.
- Lootsma, F.A. (1969). Hessian matrices of penalty functions for solving constrained-optimization problems. Philips Res. Rep., 24, pp 322-331.
- Lootsma, F.A. (1970). Boundary properties of penalty functions for constrained minimization. Philips Res. Rep. Supp. No. 3.
- Luenburger, D.G. (1969). "Optimization by Vector Space Methods". Wiley.
- Mangasarian, O.L. and Fromowitz, S. (1967). The Fritz John necessary optimality conditions in the presence of equality and inequality constraints. J. Math. Anal. Appl., 17 (1), pp 37-47.
- McCormick, G.P. (1969). The rate of convergence of the reset Davidon variable metric method. MRC Tech. Rep. No. 1012, University of Wisconsin.
- McCormick, G.P. and Pearson, J.D. (1969). Variable metric methods and unconstrained optimization; in "Optimization", ed. R. Fletcher. Academic Press.
- Meyers, G.E. (1968). Properties of the conjugate-gradient and Davidon methods. J. Opt. Th. Appl., 2 (4), pp 209-219.

Morrison, D.D. (1968). Optimization by least squares.

SIAM J. Numer. Anal., 5 (1), pp 83-88.

Murray, W. (1969a). Indefinite quadratic programming.

Nat. Phys. Lab. Rep. Ma 76.

Murray, W. (1969b). Behaviour of Hessian matrices of barrier and penalty functions arising in optimization.

Nat. Phys. Lab. Rep. Ma 77.

Murray, W. (1969c). Constrained Optimization.

Nat. Phys. Lab. Rep. Ma 79.

Murtagh, B.A. and Sargent, R.W.H. (1969). A constrained minimization method with quadratic convergence; in "Optimization", ed. R. Fletcher. Academic Press.

Murtagh, B.A. and Sargent, R.W.H. (1970). Computational experience with quadratically convergent minimization methods. Comp. J., 13 (2), pp 185-194.

Nelder, J.A. and Mead, R. (1965). A simplex method for function minimization. Comp. J., 7, pp 308-313.

Osborne, M.R. (1971). Private communication.

Osborne, M.R. and Ryan, D.M. (1970a). On penalty function methods for nonlinear programming problems. J. Math. Anal. Appl., 31 (3), pp 559-578.

Osborne, M.R. and Ryan, D.M. (1970b). An algorithm for nonlinear programming. Tech. Rep. No. 35, Computer Centre, A.N.U.

- Parisot, G.R. (1961). Résolution numérique approchée du problème de programmation linéaire par application de la programmation logarithmique. *Rev. Fr. Recherche Opérationnelle*, 20, pp 227-259.
- Pearson, J.D. (1969). On variable metric methods of minimization. *Comp. J.*, 12, pp 171-178.
- Pietrzykowski, T. (1962). Application of the steepest descent method to concave programming; in "Proceedings of IFIPS Congress", Munich, North-Holland Pub. Co., pp 185-189.
- Polak, E. and Ribière, G. (1969). Note sur la convergence de méthodes de directions conjuguées. *R.I.R.O.*, 3, pp 35-43.
- Pomentale, T. (1965). A new method for solving conditioned maxima problems. *J. Math. Anal. Appl.*, 10, pp 216-220.
- Powell, M.J.D. (1967). A method for nonlinear constraints in minimization problems. *AERE Tech. Rep. No. 310*.
- Powell, M.J.D. (1968). A survey of numerical methods for unconstrained optimization. *AERE Tech. Rep. No. 340*.
- Powell, M.J.D. (1969a). Rank one methods for unconstrained optimization. *AERE Tech. Rep. No. 372*.
- Powell, M.J.D. (1969b). On the convergence of the variable metric algorithm. *AERE Tech. Rep. No. 382*.

- Powell, M.J.D. (1970). A new algorithm for unconstrained optimization. AERE Tech. Rep. No. 393.
- Rosen, E.M. (1966). A review of quasi-Newton methods in nonlinear equation solving and unconstrained optimization; in "Proceedings 21st National Conference of the A.C.M.". Thompson Book Co., pp 37-41.
- Rosen, J.B. (1960). The gradient projection method for nonlinear programming, part 1: linear constraints. J. Soc. Ind. Appl. Math., 8 (1), pp 181-217.
- Rosen, J.B. and Suzuki, S. (1965). Construction of nonlinear programming test problems. Comm. A.C.M., 8, p 113.
- Rosenbrock, H.H. (1960). An automatic method for finding the greatest or least value of a function. Comp. J., 3, pp 175-184.
- Rudin, W. (1953). "Principles of Mathematical Analysis". McGraw-Hill, N.Y.
- Schmit, L.A. and Fox, R.L. (1965). Integrated approach to structural synthesis. AIAA J., 3 (6), pp 1104-1112.
- Shah, B.V., Buehler, R.J. and Kempthorne, O. (1964). Some algorithms for minimizing a function of several variables. J. Soc. Ind. Appl. Math., 12, pp 74-92.
- Spang, H.A. (1962). A review of minimization techniques for nonlinear functions. SIAM Review, 4 (4), pp 343-365.

- Stong, R.E. (1965). A note on the sequential unconstrained minimization technique for non-linear programming. *Man. Sci.*, 12 (1), pp 142-144.
- Tremolières, R. (1968). La méthode des centre à troncature variable. Thèse, Paris.
- Wolfe, P. (1959). The simplex method for quadratic programming. *Econometrica*, 27 (3), pp 382-398.
- Wolfe, P. (1961). A duality theorem for nonlinear programming. *Quart. Appl. Math.*, 19 (3), pp 239-244.
- Zangwill, W.I. (1967). Nonlinear programming via penalty functions. *Man. Sci.*, 13 (5), pp 344-358.
- Zangwill, W.I. (1969). "Nonlinear Programming: A Unified Approach". Prentice-Hall.
- Zoutendijk, G. (1960). "Methods of Feasible Directions". Elsevier.