EXTRAPOLATION METHOD IN CONSTRAINED OPTIMIZATION.

BY

PATRICK NOAH OKOLO
(M.TECH/SSSE/557/2000/2001)

A PROJECT SUBMITTED TO THE DEPARTMENT OF
MATHEMATICS/COMPUTER SCIENCE
FEDERAL UNIVERSITY OF TECHNOLOGY, MINNA.

IN PARTIAL FULFILMENT FOR THE AWARD OF MASTER OF TECHNOLOGY (M. TECH) IN MATHEMATICS.

JULY, 2003.

CERTIFICATION

This thesis titled "EXTRAPOLATION METHOD IN CONSTRAINED OPTIMIZATION" by Patrick Noah Okolo, meets the regulations governing the award of the degree of Masters of Technology in Mathematics. Federal University of Minna and is approved for its contribution to knowledge and literary presentation.

Dr. S. A. Reju
Supervisor

Date: 16-12-2003

Mr. L. N. Ezeako
Head of Department

Date: 26-1-2004

Prof. (Mrs) H.O. Akanya
Dean, School of Science
and Science Education

Date: 4th-02-2004

Prof. J. A. Abalaka 23/8/04
Dean, Postgraduate School

Date:..............................

Prof. A. U. Afuwape
External Examiner

Date: 31/7/2003

AKNOWLEDGEMENT

Glory and Honour be the name of God, almighty and Jesus Christ my Saviour for the grace and empowerment from above. I am conscious of the fact that without you, Lord, I can do nothing.

I am immensely indebted to my project Supervisor, Dr Reju, S. A. for been there despite his tight schedules. You are indeed peculiar and may God bring to pass all his good counsel for your life.

I am also grateful to the H.O.D. Maths/computer Science and all the lecturers in the department. You have further opened my eyes and equipped me to be effective and meet the challenges ahead.

To Dr and Mrs. Raymond Jatau, thank God for making you to be father to some of us. Your expectations will not be cut off and to Mummy (Mrs.) Dinatu Ajegana, your concern and care can not be condensed in these lines. But I want to say thank you.

I wish to acknowledge, Mr. Gabriel Omotosho, who has been my room mate, course mate and colleague and Mr. Sunday Abbah, you are not just a course mate but a brother.

To Mrs. Dr Reju, thanks for all contribution to the success of this work. I also wish to appreciate all my course mates for all their care and concern.

I am grateful to Pastor Caleb Agosaah and Victor Churchill for making great sacrifices during my quest for reference materials.

I wish to acknowledge Mr. Daniel Barau, Mr. Michael Ibrahim, Mr. John Saleh and Tahir Adejo for all your labour of love. I will not forget to mention my nephews, Solomon and Iko.

Finally to my wife, Carol, I cannot thank my self. God is always on the throne.

ABSTRACT

The idea of extrapolation is to use the solution of previous unconstrained minimization or the barrier sub problem to fit a polynomial to the barrier trajectory, and then use this polynomial to predict the solution of the constrained minimization problem.

This study designed extrapolation algorithm. BASIC computer program was developed and used to test the performance of the algorithm.

The extrapolation technique helps interior penalty function by accelerating the rate of convergence to the optimum solution.

TABLE OF CONTENT

CHAPTER THREE

EXTRAPOLATION METHOD IN INTERIOR PENALTY METHOD

CHAPTER FOUR

COMPUTATIONAL EXPERIMENTATION OF EXTRAPOLATED ALGORITHM

CHAPTER FIVE

PERFORMANCE OF EXTRAPOLATION ALGORITHM AND CONCLUSION

CHAPTER ONE

THEORY OF CONSTRAINED OPTIMIZATION.

1.1: INTRODUCTION.

The optimist proclaims that we live in the best of all possible worlds; and the pessimist fears this is true. J.B. Cabell.

All of us must make decisions in the course of our day-to- day events in order to accomplish certain tasks. Usually there are several, perhaps many, possible ways to accomplish these tasks, although some choices will generally be better than others. Consciously or unconsciously, we must therefore decide upon the best – or optimal- way to realize our objectives.

Optimization is applied in virtually all areas of human endeavour, including engineering, system design, economics, power systems, water and land use, transportation systems, scheduling systems, resource allocation, personnel planning, portfolio selection, mining operations, blending of raw materials, structural design, and control systems. Optimizers or decision makers use optimization in the design of systems and processes, in the production of products and in the operation of systems.

Techniques of optimization assume such varied forms that no one general description is possible. Thus this study intends to examine the extrapolation method in interior penalty function for solving constrained optimization problems.

With the advent of modern technology move and more emphasis has been placed on optimization of various types and special thinking has developed to the extent that it is meaningful to speak of a mathematical theory of optimization. Computer technology has been critically important in practical application, hence computational experimentation of extrapolation algorithm will feature in the course of this study.

1.2: OPTIMIZATION AND ITS FRAMEWORK.

The fundamental problem of optimization is to arrive at the best possible decision in any given set of circumstances. Of course, many situations arise where the 'best' is unattainable for one reason or the another; sometimes what is 'best' for one person is 'worse' for another; more often we are not at all sure what is meant by 'best'. The first step, therefore in a mathematical optimization problem is to choose some quantity, typically a function of several variables, to be maximized or minimized, subject possibly to one or more constraints. The commonest types of constraint are equalities and inequalities which must be satisfied by the variables of the problem, but many other types of constraint are possible; for example a solution in integers may be required. The next step is to choose a mathematical method to solve the optimization problem, such methods are usually called optimization techniques, or algorithms.

The optimization problems that have been posed and solved in recent years have tended to become more and more elaborate, not to say abstract.

The simply- stated problem of maximizing or minimizing a given function f several variables has attracted the attention of many mathematicians over the past three decades or so. The direct search methods of solution which involve function evaluations and comparisons only, are usually simpler, though less accurate for the same computational effort, than the indirect or gradient methods, which requires values of the function and its derivatives. Both types of method are still undergoing development, with the emphasis being on the search for efficient and reliable algorithms to deal with general non linear functions.

1.3: CONSTRAINED OPTIMIZATION TECHNIQUES.

Different techniques abound for the solution of a constrained optimization problem which can be stated in the standard form as:

Find X such that

$F(X) \rightarrow \text{Minimum}$ 1.3.1

and $g_j(X) \leq 0 \quad j = 1,2,\ldots,m$ 1.3.2

These methods can be classified as classical and non-classical methods of solution.

If the expression for the objective function and the constraints are fairly simple in terms of design variables, the classical methods of optimization can be used to solve the problem. To this, the Lagrange multiplier method is considered.

(i) LAGRANGE MULTIPLIER METHOD.

This classical technique was developed to derive the equations of constrained motion of particles, and has applied to solve optimization problem with only equality constraints.

The derivation of the theory of this method is best illustrated by considering a two variable problem with a single constraint as shown below:

$$\left.\begin{aligned} &\text{Min} \quad f(X) = f(x_1,x_2) \\ &\text{subject to } g(x_1,x_2) = b \end{aligned}\right\} \qquad 1.3.3$$

Assume that f, g $\in$ c′ i.e. their first partial derivatives exist, and further assume that

$$\frac{\partial g(x^*)}{\partial x_2} \neq 0$$

i.e., this partial derivative(the determinant of the Jacobian incase of more than one constraint) does not vanish at the optimum solution. Then by implicit function theorem; there exists an eps - neighbourhood of x_1^* where it is possible to solve $g(x_1,x_2) - b = 0$, explicitly for x_2 to obtain the relationship

$$x_2 = \phi(x_1) \qquad 1.3.4$$

substituting for x_2 in $f(x_1,x_2)$ it follows that

$$\left.\begin{aligned} &f(X) = h(x_1) = f(x_1,\phi(x_1)), \\ &\text{for} \\ &|x_1 - x_1^*| < eps \end{aligned}\right\} \qquad 1.3.5$$

If f takes on a relative minimum at x* for x satisfying $g(x_1,x_2) = b$, it must be true that there exists a δ such that $0 < \delta < eps$ and $h(x_1) < h(x_1^*)$ for all x_1 in his neighbourhood; in other words the function $h(x_1)$ has an unconstrained relative minimum at x_1^*. For the function $\phi(x_1)$ the differentiation is obtained by applying the implicit function theorem and for the function $h(x_1)$ the rules for differentiating compound functions are applied. Hence the conditions for the minimum of $h(x_1)$ may be deduced as,

$$\frac{dh}{dx_1} = \frac{\partial f}{\partial x_1} + \frac{\partial f}{\partial x_2}\frac{\partial \Phi}{\partial x_1} = \frac{\partial f}{\partial x_1} - \frac{\partial f}{\partial x_2}\frac{\partial g/\partial x_1}{\partial g/\partial x_2} = 0 \qquad 1.3.6$$

where

$$\frac{\partial \phi}{\partial x_1} = \left(\frac{\partial g/\partial x_1}{\partial g/\partial x_2}\right)$$

If a new variable λ is introduced to express the ratio,

$$\lambda = \left(\frac{\partial f/\partial x_2}{\partial g/\partial x_2}\right) \qquad 1.3.7$$

then (1.3.6) reduces to

$$\frac{\partial f}{\partial x_1} - \lambda \frac{\partial g}{\partial x_1} = 0 \qquad 1.3.8$$

and (1.3.7) may be re written

$$\frac{\partial f}{\partial x_2} - \lambda \frac{\partial g}{\partial x_2} = 0 \qquad 1.3.9$$

To these, append equality constraint of (1.3.3)

$$g(x_1, x_2) = b \qquad 1.3.10$$

then the relationships (1.3.8), (1.3.9), (1.3.10) express the necessary (but not sufficient) conditions for obtaining the constrained minimum of the problem stated in (1.3.3). Note there are three equations (1.3.8), (1.3.9), (1.3.10) and there are three variables x_1, x_2, λ which take on the values x_1^*, x_2^*, λ^* at the optimum solution and also satisfy these equations. A less rigorous approach tates that to obtain the optimum solution of (1.3.1) and (1.3.2), it suffices to ntroduce m language multipliers, $\lambda = (\lambda_1, \lambda_2, ..., \lambda_m)$ and investigate the stationary points of the Lagrangean function,

$$F(x, \lambda) = f(x_1, x_2, ..., x_n) - \sum_{i=1}^{m} \lambda_i\left[g_i(x_1, x_2, ..., x_n) - b_i\right] \qquad 1.3.11$$

It is then a necessary condition that the optimum solution of (1.3.1) and (1.3.2) is constrained in one of the stationary points of (1.3.11), and hence must satisfy the n + m equations obtained by taking n partial derivatives in λ and setting these expression equal to zero,

$$\frac{\partial F}{\partial x_j} = \frac{\partial f}{\partial x_j} - \sum_{i=1}^{m} \lambda_i \frac{\partial g_i}{\partial x_j} = 0 \qquad j= 1, 2, ..., m$$

and

$$\frac{\partial F}{\partial \lambda_i} = g_i(x_1, x_2, ..., x_n) - b_i = 0 \quad i = 1,2,...,m$$

1.3.12

Provided the rank of the Jacobian associated with the optimum solution is m, i.e.,

$r(J_x^*) = m$, it is possible to determine uniquely the optimum solution x^* and the associated multiplier λ^*.

Example 1.3.1

(a) Minimize $f(x_1,x_2) = 6 - 6x_1 + 2x_1^2 - 2x_1x_2 + 2x_2^2$

subject to $x_1 + x_2 = 2$

Solution:

To solve the problem set up the Lagrangean, (for a maximizing function)

$$F(x, \lambda) = -(6 - 6x_1 + 2x_1^2 - 2x_1x_2 + 2x_2^2) - \lambda(x^1 + x_2 - 2) \qquad (E_1)$$

and take partial derivatives with respect to x_1, x_2, λ and set these to zero

$$\left.\begin{aligned} 6 - 4x_1 + 2x_2 - \lambda = 0 \\ 2x_1 - 4x_2 - \lambda = 0 \\ x_1 + x_2 - 2 = 0 \end{aligned}\right\} \qquad E_2$$

This set of equations may be solved to obtain the optimum solution, $x_1 = 3/2$, $x_2 = 1/2$,

$\lambda = 1$ and $f(x) = ½$ the optimum value of the objective. The Lagrangean $F(x, \lambda)$ has only one stationary point the necessary condition in (E_2) for this convex problem is also sufficient to guarantee global optimality.

(b) suppose a nuclear reactor is to have the shape of a cylinder of radius R and height H. Neutron diffusion theory tells that such reactor must have the following constraint

$$g(R,H) = \left(\frac{2.4048}{R}\right)^2 + \left(\frac{\pi}{H}\right)^2 = \text{const}^2$$

We would like to minimize the volume of the reactor

$$F(R,H) = \pi R^2 H$$

Solution:

By using the equations above, the

$$\frac{\partial f}{\partial R} + \lambda \frac{\partial g}{\partial R} = 2\pi RH - 2\lambda\left(\frac{2.4048}{R^3}\right)^2 = 0 \qquad E_1$$

$$\frac{\partial f}{\partial H} + \lambda \frac{\partial g}{\partial H} = \pi R^2 - 2\lambda \frac{\pi}{H^3}^2 = 0 \qquad E_2$$

By multiplying (E_1) by R/2 and (E_2) by H we should obtain

$$H = \frac{\sqrt{2\pi R}}{2.4048}$$

The non-classical methods can be classified into two broad categories, namely, the direct methods and indirect methods.

The direct methods, in which the constraints are handled in an explicit manner include:

(i) Heuristic Search methods – the complex method.

The heuristic search methods are mostly intuitive and do not have much theoretical support.

(ii) Constraint Approximation Methods.

In these methods, the non linear objective function and the constraints are linearlized about some point and the approximating linear programming problem is solved by using linear programming techniques.

The resulting optimum solution is then used to construct a new linear approximation which will again be solved by using LP techniques. This procedure is continued until the specified convergence criteria are satisfied.

For example,

Let the given problem be:

$$\left.\begin{array}{l}\text{Find } (x_1, x_2, \ldots, x_n) \text{ which} \\ \text{Minimize } f(x_1, x_2, \ldots, x_n) \text{ subject to} \\ \text{The constraints } g_j(x_1, x_2, \ldots, x_n) \le 0 \quad j = 1,2,\ldots,m\end{array}\right\} \quad 1.3.13$$

Introduce a new variable, say x_{n+1}, and transform this problem into an equivalent form as:

Find $(x_1,x_2,\ldots,x_n, x_{n+1})$ which minimize x_{n+1}

subject to the constraints

$g_j(x_1,x_2,\ldots,x_n) \le 0 \quad j = 1,2,\ldots,m$ 1.3.14

and $g_{m+1}(x_1,x_2,\ldots,x_n, x_{n+1}) = f(x_1,x_2,\ldots,x_n) - x_{n+1} \le 0$

In effect, by the addition of one variable and one constraint, we have converted the original problem with a non linear objective function, Eq(1.3.13), into a problem with a linear objective function, Eq(1.3.14). Thus, without loss of generality, we can assume that the given problem is:

$$\left.\begin{array}{l}\text{Find } X \text{ which minimizes} \\ f(X) = C^T X = \sum_{i=1}^{n} C_i x_i \\ \text{Subject to the constraints} \\ g_j(X) \le 0 \quad j = 1,2,\ldots,m\end{array}\right\} \quad 1.3.15$$

The cutting plane algorithm can be stated by the following steps:

(i) start with an initial point X_1 and set the iteration number as

$i = 1$. The point X_1 need not be feasible.

(ii) Linearize the constraint functions $\nabla g_j(X)$ about the point X_1 as

$$g_j(X) \approx g_j(X_i) + \nabla g_i(X_i)^T (X - X_i), \; j = 1,2,\ldots,m \quad 1.3.16$$

(iii) Formulate the approximating linear programming problem as

$$\left.\begin{array}{l}\text{Minimize } C^T X \text{ subject to} \\ g_j(X_i) + \nabla g_i(X_i)^T (X - X_i) \le 0 \quad j = 1,2,\ldots,m\end{array}\right\} \quad 1.3.17$$

(iv) Solve the corresponding L P problem to obtain the solution vector X_{i+1}

(v) Evaluate the original constraints at X_{i+1}, i.e. find $g_j(X_{i+1})$, $j = 1,2,\ldots,m$.

if $g_j(X_{i+1}) \leq \varepsilon$ for $j = 1,2,\ldots,m$ where ε is a prescribed small positive tolerance, all original constraints can be assumed to have been satisfied.

Hence stop procedure by taking

$$X_{opt} \simeq X_{i+1}$$

If $g_j(X_{i+1}) > \varepsilon$ for some values of j, find the most violated constraint as

$$g_k(X_{i+1}) = \max_j \left(g_j(X_{i+1}) \right)$$

Relinearize the constraint $g_k(X) \leq 0$ about the point X_{i+1} as

$$g_k(X) \simeq g_k(X_{i+1}) + \nabla g_k(g_k(X_{i+1}))^T(X - X_{i+1}) \leq 0 \qquad 1.3.18$$

and add this as the (m+1)th constraint to the previous L P problem.

(vi) Set the new iteration number as $i = i+1$, the total number of constraints in the new approximating L P problem as $m = m+1$, and go to step (iv).

Example 1.3.2:

Minimize $f(x_1,x_2) = x_1 - x_2$ subject to

$$g_1(x_1,x_2) = 3x_1^2 - 2x_1x_2 + x_2^2 - 1 \leq 0$$

using the cutting plane method. Take the convergence limit in step (v) as $\varepsilon = 0.02$.

Solution:

STEPS 1,2,3,: Although we can start the solution from the initial point X_1 to avoid the possible unbound on x_1 and x_2 as $-2 \leq x_1 \leq 2$ and $-2 \leq x_2 \leq 2$ and solve the following LP problem:

Minimize $f = x_1 - x_2$ subject to

$$\left.\begin{array}{l} -2 \le x_1 \le 2 \\ -2 \le x_2 \le 2 \end{array}\right\} \quad E_1$$

The solution of this problem can be obtained as

$$X = \begin{Bmatrix} -2 \\ 2 \end{Bmatrix} \quad \text{with } f(X) = -4$$

STEP 4: Since we have solved

$$X_{i+1} = X_2 = \begin{Bmatrix} -2 \\ 2 \end{Bmatrix}$$

STEP 5: Since $g_1(X_2) = 23 > \varepsilon$, we linearize

$g_1(X)$ about the point X_2 as

$$g_1(X) \approx g_1(X_2) + \nabla g_1(X_2)^T(X - X_2) \le 0 \qquad E_2$$

As

$$g_1(X) = 23, \ \frac{dg_1}{dx_1}\bigg|_{X_2} = 6x_1 - 2x_2\bigg|_{X_2} = -16$$

and

$$\frac{\partial g_1}{\partial x_1}\bigg|_{X_2} = (-2x_1 + 2x_2)\bigg|_{X_2} = 8, \quad (E_2) \text{ becomes}$$

$$g_1(X) \approx -16x_1 + 8x_2 - 25 \le 0$$

By adding this constraint to the previous LP problem, the new LP problem becomes:

Minimize $f = x_1 - x_2$ subject to

$$\left.\begin{array}{l} -2 \le x_1 \le 2 \\ -2 \le x_2 \le 2 \\ -16x_1 + 8x_2 - 25 \le 0 \end{array}\right\} \quad E_3$$

STEP 6: Set the iteration number as $i = 2$ and go to step 4.

STEP 4: Solve the approximating LP problem in (E_3) and obtain the solution

$$X_3 = \begin{Bmatrix} -0.5625 \\ 2.0 \end{Bmatrix} \text{ with } f3 = f(X_3) = -2.6526$$

STEP 5: As $g_1(X_3) \approx 6.19922 > \varepsilon$, we linearize

$g_1(X)$ about X_3 as

$$g_1(X_3) \approx g_1(X_3) + \nabla g_1(X_3)^T(X - X_3) \leq 0 \qquad E_4$$

As $g_1(X_3) = 6.19922$, $\left.\frac{\partial g_1}{\partial x_1}\right|_{X_3} = -7.375$

and $\left.\frac{\partial g_1}{\partial x_1}\right|_{X_3} = 5.125$

(E_4) becomes

$$g_1(X) \approx -7.375x_1 + 5.125x_2 - 8.19922 \leq 0$$

This gives the new LP problem as :

$$\left.\begin{aligned} &\text{Minimize} \quad f = x_1 - x_2 \text{ subject to} \\ &-2 \leq x_1 \leq 2 \\ &-2 \leq x_2 \leq 2 \\ &-16x_1 + 8x_2 - 25 \leq 0 \\ &-7.375x_1 + 5.125x_2 - 8.19922 \leq 0 \end{aligned}\right\} \qquad E_5$$

STEP 6: Set the iteration number as i = 3 and go to step 4.

STEP 4: Solve the approximating LP problem of (E_5) to obtain the solution

$$X_4 = \begin{Bmatrix} 0.27870 \\ 2.00000 \end{Bmatrix} \text{ and } f_4 = f(X_4) = -1.72193$$

This procedure is continued until the specified convergence criterion, $g_1(X) \leq \varepsilon$ in step 5 is satisfied.

The cutting plane method – constraint Approximation method is an efficient technique for solving convex programming problems with nearly linear objective and constraint functions.

Where as in most of the indirect method, the constrained minimization problem is solved as a sequence of unconstrained minimization problems. The following are the two basic types of indirect optimization methods:

(i) Transformation of Variables.

Some of the constrained optimization problems have their constraints expressed as simple and explicit function of the decision variables. In such cases, it may be possible to make a change of variable such that the constraints are automatically satisfied. Transformation of variables will be discussed in sec. 1.5

(ii) Penalty Function Methods.

There are two types of penalty function methods – the interior penalty function method and the exterior penalty function method. In both types of methods, the constrained minimization problem is transformed into a sequence of unconstrained minimization problem. Penalty function method will be the main focus of chapter two of this work.

1.4: CHARACTERISTICS OF CONSTRAINED PROBLEM.

The presence of constraints in a nonlinear programming problem creates more problems while finding the minimum. Several situations can be identified depending on the effect of constraints on the objective function. The simplest situation is when the constraints do not have any influence on the minimum point.

If the constraints are also nonlinear, the minimum point X* can also nonlinear, the minimum point X* can be found by making use of the necessary and sufficient conditions:

$$\nabla f \big|_{X^*} = 0 \qquad 1.4.1$$

and

$$Jx^* = \left(\frac{\partial^2 f}{\partial x_i \partial x_j} \right) x^* \text{ is positive definite} \qquad 1.4.2$$

Thus in order to use these conditions, one must know for sure, that the constraints are not going to have any effect on the minimum. Thus one has to passed with the general assumption that the constraints will have same influence on the optimum point.

Another situation is one where the minimization problem has two or more local minima. If the objective function has two or more unconstrained local minima and if at least one of them is contained in the feasible region, the constrained problem would have at least two local minima.

It can be sum up that, the minimum of a nonlinear programming problem will not be, in general, an extreme point of the feasible region and may have local minima even if the corresponding unconstrained problem is not having local minima. Furthermore, none of the local minima may correspond to the global minimum of the unconstrained problem. All these characteristics are direct consequences of the introduction of constraints.

1.5: TRANSFORMATION TECHNIQUES.

(i) Change of variables

If the constraints gj(X) are explicit functions of the variables xi and have certain simple forms, it may be possible to make a transformation of the independent variables such that the constraints are automatically satisfied. Thus it may be possible to convert a constrained optimization problem into an unconstrained one by making change of variables. One of the frequently encountered constraints, which can be satisfied in this way, is that when the variable is bounded below and above by certain constraints:

$$l_i \leq x_i \leq u_i \qquad 1.5.1$$

where l_i and u_i are respectively, the lower and the upper limits on x_i . These constraints can be satisfied by transforming the variable x_i as

$$x_i = l_i + (u_i - l_i) \sin^2 y_i \qquad 1.5.2$$

where y_i is the new variable which can take any value.

In the particular case when the variable x_i is restricted to lie in the interval (0,1), we can use any of the following transformations:

$$\left.\begin{aligned} x_i &= \sin^2 y_i \\ x_i &= \cos 2yi \\ x_i &= \frac{e^{y_i}}{(e^{y_i} + e\text{-}yi)} \\ x_i &= \frac{e^{y_i^2}}{(1 + y_i^2)} \end{aligned}\right\} \qquad 1.5.3$$

If the variable x_i is contained to take only positive values, the transformation has to be as follows:

$$\left.\begin{aligned} X_i &= \text{abs}\,(y_i) \\ x_i &= y_i^2 \\ x_i &= e^{y_i} \end{aligned}\right\} \qquad 1.5.4$$

On the other hand, if the variable is restricted to take values lying only in between –1 and 1, the transformation is given by

$$\left.\begin{aligned} x_i &= \sin y_i \\ x_i &= \cos y_i \\ x_i &= \frac{2y_i}{1 + y_i^2} \end{aligned}\right\} \qquad 1.5.5$$

Applying these transformation, then the unconstrained minimum of the objective function is sought with respect to the new variables y_i.

The following points are to be noted in applying this transformation technique:

(a) the constraints $g_j(X)$ have to be very simple function of x_i

(b) for certain constraints it may not be possible to find the necessary transformation.

(c) If it is not possible to eliminate all the constraints by making change of variables, it may be better not to use the transformation at all. The partial transformation may, sometimes produce a distorted objective function which might be more difficult to minimize than the original function.

Example 1.5.1

Find the dimensions of a rectangular prism of parcel which has the largest volume when its length plus girth are limited to take the maximum values of 42 cm and 72 cm respectively.

Solution:

Let x_1, x_2 and x_3 denote the length, breadth and depth of the parcel respectively.

The problem can be stated as follows:

Maximize $f(x_1, x_2 x_3) = x_1 x_2 x_3$ $\quad E_1$

Subject to $x_1 + 2x_2 + 2x_3 \leq 72$ $\quad E_2$

$x_1 \leq 42$ $\quad E_3$

$x_1 \geq 0$ $\quad E_4$

$x_2 \geq 0$ $\quad E_5$

$x_3 \geq 0$ $\quad E_6$

by introducing the new variables y_1 , y_2 and y_3 as

$$\left.\begin{array}{l} y_1 = x_1 \\ y_1 = x_2 \\ y_1 = x_1 + 2x_1 + 2x_3 \end{array}\right\}, \text{ i.e. } \left.\begin{array}{l} x_1 = y_1 \\ x_2 = y_2 \\ x_3 = \frac{1}{2}(y_3 - y_1 - 2y_2) \end{array}\right\} \quad E_7$$

the constraints (E_1) to (E_1) can be expressed as

$$\left.\begin{array}{l} 0 \leq y_1 \leq 42 \\ 0 \leq y_2 \leq 36 \\ 0 \leq y_3 \leq 72 \end{array}\right\} \quad E_8$$

where the upper bound on y_2 is obtained by setting $x_1 = x_3 = 0$ in (E_3). As x_1 and x_2 are constrained to be positive, intuitively we feel that a negative value of x_3 (negative volume) may not correspond to the maximum of f. hence we do dot consider (E_6) in the subsequent computations. The constraints (E_8) will automatically satisfied if we define the new variables z_1, z_2 and z_3 as

$$\left.\begin{array}{l} y_1 = 42 \sin^2 z_1 \\ y_2 = 36 \sin^2 z_2 \\ y_3 = 72 \sin^2 z_3 \end{array}\right\} \quad (E_9)$$

If we use (E_7) and (E_9), the original constrained optimization problem can be transformed into an equivalent unconstrained problem as:

Maximize $f(z_1, z_2, z_3)$

$$= \tfrac{1}{2} (42 \sin^2 z_1)(36 \sin^2 z_2)(72 \sin^2 z_3 - 42 \sin^2 z_1 - 72 \sin^2 z_2)$$

$$= 4536 \sin^2 z_1 \sin^2 z_2 (12 \sin^2 z_3 - 7 \sin^2 z_1 - 12 \sin^2 z_2) \quad E_{10}$$

To find the maximum of f, we use the necessary conditions:

$$\frac{\partial f}{\partial z_1} = (4536)(4) \sin z_1 \cos z_1 \sin^2 z_2 (6 \sin^2 z_3 - 7 \sin^2 z_1 - 6 \sin^2 z_2) = 0 \quad E_{11}$$

$$\frac{\partial f}{\partial z_1} = (4536)(2) \sin^2 z_1 \sin z_2 \cos z_2 (12 \sin^2 z_3 - 7 \sin^2 z_1 - 24 \sin^2 z_2) = 0 \quad E_{12}$$

$$\frac{\partial f}{\partial z_1} = (4536)(24) \sin^2 z_1 \sin^2 z_2 \sin z_3 \cos z_3 = 0 \quad E_{13}$$

Equation (E_{12}) shows that $\sin z_1 = 0$ or $\sin z_2 = 0$ or $\sin z_3 = 0$

or cos z_3 = 0. However if sin z_1 = 0 or sin z_2 = 0 is taken, we do not get any information about z_2 and z_3 from the other equations. Similarly, sin z_3 = 0 will lead to the trivial solution $z_1 = z_2 = 0$ from (E_{11}) and (E_{12}). Hence

cos z_3 must be zero to make (E_{13}) valid. This lead to $\sin^2 z_3 = 1$,

$\sin^2 z_1 = 4/7$ and $\sin^2 z_2 = 1/3$. This solution can be verified to correspond to a relative maximum of f. using (E_9) and (E_7), the solution of the problem, in terms of the original variables, can be obtained as:

x_1^* = length = 24cm, x_2^* = breadth = 12cm, x_3^* = depth = 12cm and

f_{max} = maximum volume = 3456 cm^3

(ii) Elimination of variables.

If an optimization problem in inequality constraints, all of them may not be active (i.e. a constraint which is satisfied with equality sign). If it is known, in advance, which constraints are going to be active at the optimum point, we use those constraint equations to eliminate the variables from the problem. Thus if r (< n) specific constraints are known to be active at the optimum point, can eliminate any r variables from the problem and obtain a new problem involving only n - r variables and m – r constraints. This problem will be, in general, much easier to solve compared to the original problem.

The major drawback of this method is that it will be very difficult to know beforehand, which of the constraints are going to be active at the optimum point. Thus, in a general problem with m constraints, we need to check

(a) the minimum of f(x) with no constraints (assuming that no constraint is active at the optimum point).

(b) The minimum of f(x) by taking one constraint at a time with equality sign (assuming that one constraint is active at the optimum point).

(c) The minimum of f(x) by taking all possible combinations of constraints taken two at a time (assuming that two constraints are active at the optimum point) e.t.c. if any of these solutions satisfies the Kuhn – Tucker necessary conditions, it is likely to be a local minimum of the original optimization problem.

It can be seen that in the absence of a prior knowledge about which constraints are going to be active at the optimum point, the number of problems to be solved is given by

$$1+m+\frac{m(m-1)}{2!}+\frac{m(m-1)(m-2}{3!}+\ldots+\frac{m}{n!(m-n)!}=\sum_{k=0}^{n}\frac{m}{k!(m-k)!} \qquad 1.5.6$$

For example, if the original optimization problem has 5 variables and 10 constraints, the number of problems to be solved will be 638, which can be seen to be very large.

However, in LP problems, it is known that exactly n – m variables will be zero at the optimum point. In such cases, we need to solve only

$\left(\frac{n}{(n-m)!m!}\right)$ problems to identify the optimum solution.

For example if m = 5 and n = 10, the number of problems to be solved will be 252, which is still a large number in terms of practical computations. Hence this approach is not feasible even for solving LP problems.

Example 1.5.2

Find the dimension of the parcel described in example 1.5.1 if the constraint given is known to be active at the optimum point.

Solution:

The problem can be treated as follows:

Maximize $f(x_1, x_2, x_3) = x_1 x_2 x_3$ (E_1)

$x_1 + 2x_2 + 2x_3 = 72$ (E_2)

$x_1 \leq 42$ (E_3)

$x_1 \geq 0$ (E_4)

$x_2 \geq 0$ (E_5)

$x_3 \geq 0$ (E_6)

With the help of (E_2), x_1 can be expanded as

$= 2(36 - x_2 - x_3)$ (E_7)

and the objective function as

$f(x_2, x_3) = 2(36x_2x_3 - x_2^2x_3 - x_2x_3^2)$ (E_8)

Now we maximize $f(x_2, x_3)$ as an unconstrained problem and accept the solution if it satisfies the constraints (E_3) to (E_6)

By solving the equations

$$\frac{\partial f}{\partial x_2} = 2(36x_3 - 2x_2x_3 - x_3^2) = 0 \quad (E_9)$$

and

$$\frac{\partial f}{\partial x_3} = 2(36x_2 - x_2^2 - 2x_2x_3) = 0 \quad (E_{10})$$

at all times. It also has the reassuring feature that, should the algorithm be terminated prematurely, a feasible solution is always returned.

The second class of penalty function methods, the exterior penalty function methods, are characterized by their use of infeasible points.

Before developing on the two classes of methods for the penalty function method, it is pertinent to consider briefly some properties of problem (2.1.1) and then introduce the Kuhn-Tucker first order necessary conditions for its solution. The feasible region of (2.1.1) will be written as

$$X = \{ x : g(x) \leq 0 \} \qquad 2.1.3$$

The interior of the feasible region will be written as

$$X_1 = \{ x : g(x) < 0 \} \qquad 2.1.4$$

We shall assume throughout that the problem functions are continuous and also, mainly for computational convenience, that they are differentiable. We shall also assume a local solution, X*, exists and that at least one constraint is active at X* (i.e. J, the active constraint set at X* is non-empty). For problem (2.1.1), we write the Lagrangrian as

$$L(X,\lambda) = F(X) - \sum_{i=1}^{m} \lambda_i g_i(x) \qquad 2.1.5$$

and following similar arguments, we can establish the Kuhn-Tucker first – order necessary condition for X* to be a local minimum of (2.1.1)

<u>KUHN – TUCKER FIRST – ORDER CONDITIONS</u>

If X* is a local solution of (2.1.1) and the constraints $g_i(x)$, $i \in J$, satisfy "constraint qualification" then there exists $\lambda^* \in E^n$ such that

$$G(X^*) - \sum_{i=1}^{m} \lambda_i^* a_i(X^*) = 0 \qquad 2.1.6$$

CHAPTER TWO

BASIC APPROACH IN THE PENALTY FUNCTION METHOD.

2.1 INTRODUCTION:

The penalty function method has been widely used to transform constrained optimization problems into a sequence of unconstrained optimization problem. Let the basic optimization problem be of the form:

$$\left.\begin{array}{l} \text{Find } X \text{ which minimize } F(X) \\ \text{Subject to } G_i(x) \leq 0 \quad i = 1, 2, \ldots, m \end{array}\right\} \qquad 2.1.1$$

This problem is converted into an unconstrained minimization problem by constructing a function of the form

$$\phi(X, r_k) = F(X) + r_k \sum_{i=1}^{m} G_i \left[g_i(x) \right] \qquad 2.1.2$$

Where G_i is some function of the constraints g_i and r_k is a positive constant known as the penalty parameter. The second term on the right side of (2.1.2) is called the penalty term. If the unconstrained minimization of the ϕ – function is repeated for a sequence of values of the penalty parameter $r_k(k = 1,2,\ldots)$,the solution may be brought to converge to that of the original problem stated in (2.1.1). this is the reason why the penalty function method have also been referred to as Sequential Unconstrained Minimization Techniques commonly abbreviated to SUMT. Two major classes of such methods can be identified. The first, the interior penalty function method (also known as the barrier function methods), are characterized by their property of preserving constraint feasibility

$\lambda_i^* \geq 0 \quad i \in J$ 2.1.7

$\lambda_i^* = 0 \quad i \notin J$ 2.1.8

where g(x) and $a_i(x)$, i = 1,2,...,m are the gradients of the problem functions.

The importance of the Kuhn – Tucker conditions in relation to penalty function methods lies in the fact that they or at least expressions of the same form, appear naturally in the foundation of penalty function methods. The conditions under which the Kuhn – Tucker conditions holds also enable us to establish some very attractive and useful results concerning the methods.

2.2: INTERIOR PENALTY FUNCTION METHOD.

As indicated in the previous section, in the interior penalty function method, a new function $\phi(X, r_k)$ is defined so that a barrier is constructed at the boundary of the feasible region X, and the solution X*, is approached from the interior of X(i.e. X_1 is non-empty) by modifying the barrier using the controlling parameter.

Let us define $G(g(x), r_k)$ as

$$G(g(x), r_k) = r_k \sum_{i=1}^{m} G_i(g_i(x)) \qquad 2.2.1$$

Where r_k is a positive scalar and Gi(t) is defined continuously on the interval t > 0. Let $G_i(t) \rightarrow \infty$ as $t \rightarrow 0_+$. For computational convenience we also assume G_i is differentiable. It is interesting to note that from a practical point of view the conditions imposed on $G_i(t)$ are not of great importance provided one such function can be found. The conditions that must be considered important in establishing the penalty function method are those that the problem function must satisfy. Te interior penalty function method is then

$$\phi (X, r_k) = F(X) + r_k \sum_{i=1}^{m} G_i (g_i(x)) , \quad r_k > 0 \qquad 2.2.2$$

$\phi (X, r_k)$ is defined as X_1 and $\phi(X, r_k) \to \infty$ as $g_i(x) \to 0$ for any i. If $g_i(x^*) = 0$ then

$X \to X^*$, the growth of $G_i (g_i(x))$ can be controlled or "cancelled" by decreasing r_k

Example 2.2.1

(i) Inverse interior penalty function (Carroll, 1961)

$$G_i (g_i(x)) = - g_i(x)^{-1} \qquad 2.2.3$$

(ii) Log Interior penalty function (Frish, 1955)

$$G_i (g_i(x)) = \log[- g_i(x)] \qquad 2.2.4$$

The behaviour of 2.2.2 can be interpreted in the following way. Assume $G_\alpha(X^*) = 0$ for some α. If r_k is decreased, then $G_\alpha(g_\alpha(x))$ can be increased without increasing

$\phi(X, r_k)$. This implies that $g_\alpha(x)$ can be decreased, so permitting $X(r_k)$ to approach X^*. As a by-product of the convergence $X(r_k) \to X^*$,we would also expect $F(X(r_k))$ to decrease towards $F(X^*)$. This observation forms the basis of the computational procedure.

Using (2.2.3), the ϕ – function defined originally by Carroll is

$$\phi (X, r_k) = F(X) - r_k \sum_{i=1}^{m} \frac{1}{g_i(x)} \qquad 2.2.5$$

It can be seen that the value of the function ϕ will always be greater than F since $g_i(x)$ is negative for all feasible points X. If any constraint $g_i(x)$ is satisfied critically (with equality sign), the value of Φ tends to infinity. It is to be noted that the penalty term in (2.2.5) is not defined if X is infeasible. This introduces serious shortcoming while using (2.2.5). Since this equation does not allow any constraint to be violated, it requires a feasible starting point for the search towards the optimum point. Since the initial point as well as each of the subsequent points generated in this method is classified as an "interior penalty function" formulation.

2.2.6 BASIC ALGORITHM

The iteration procedure of this method can be summarized as follows.

(i) start with an initial feasible point X_1 satisfying all the constraints with strict inequality sign, that is $g_i(x) < o$ for $i = 1,2,\ldots,m$ and initial value of $r_i > 0$.

Set $k = 1$.

(ii) Minimize $\phi\ (X,\ r_k)$ by using any of the unconstrained minimization methods and obtain the solution X_k^*.

(iii) Test whether X_k^* is the optimum solution of the original problem. If X_k^* is found to be optimum, terminate the process. Otherwise, go to the next step.

(iv) Find the value of the next penalty parameter, r_{k+1} as

$r_{k+1} = c\ r_k$ where $c < 1$

(v) Set the new value of $k = k+1$, take the new starting point as $X_1 = X_k^*$ and go to step (ii)

These steps are shown in the form of a flowchart in fig 2.2.1

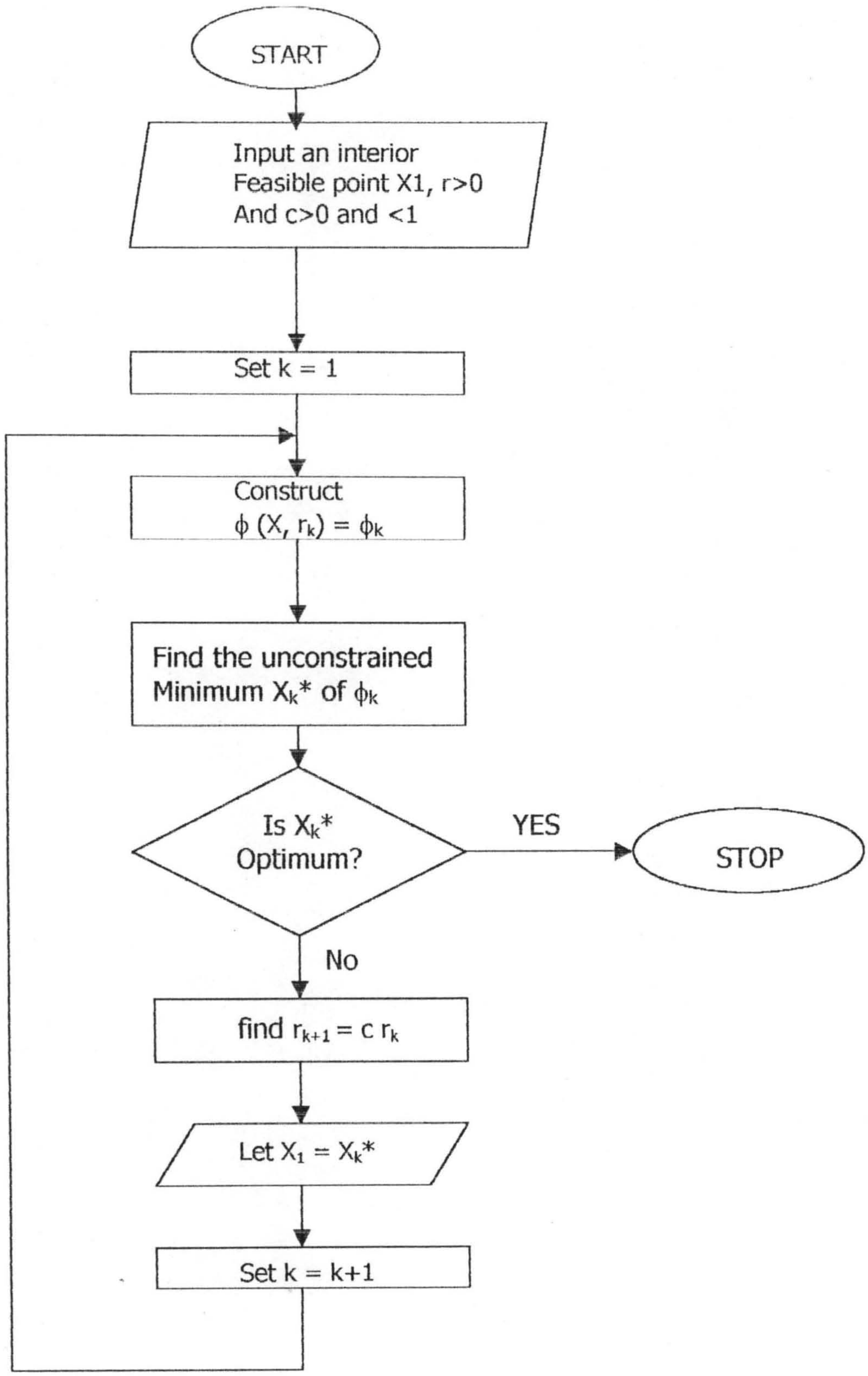


Fig 2.2.1: flowchart for the interior penalty function method.

(ii) Identify the constraint which is violated most at the point X_1, that is find the integer k such that

$g_k(X_1) = \max\left[g_i\,(X_1)\right] \geq 0$ $\quad$ $i = m - r+1,\ m - r+2, \ldots,\ m$

2.3.2

(iii) Now formulate a new optimization problem as :

Find X which minimizes $g_k(X)$ subject to

$$\left.\begin{array}{ll} g_i(X) \leq 0, & i = 1,2,..,m\text{-}r \\ \text{and} & \\ g_i(X) - g_k(X) \leq 0, & i = m\text{-}r+1,\ m\text{-}r+2, \ldots,\ k\text{-}1,\ k\text{-}2, \ldots,\ m \end{array}\right\} \quad 2.3.3$$

(iv) Solve the optimization problem formulated in step (iii) by taking the point X_1a, as the feasible starting point using the interior penalty function method. Note that this optimization method can be terminated whenever the value of the objective function $G_k(X)$ drop below zero. Thus the solution obtained X_M will satisfy at least one more constraint than did the original point X_1.

(v) If all the constraints are satisfied at the point X_M, set the new starting point as $X_1 = X_M$, and renumber the constraints such that the last constraints will be the unsatisfied ones (this value of r will be different from the previous value), and go to step (ii)

This procedure is repeated until all the constraints are satisfied, and a point $X_1 = \dot{X}_M$ is obtained for which $g_i(X_1) < 0$ $I = 1,2,..,m$

If the constraints are consistent, it should be possible to obtain by applying the above procedure, a point X_1 that satisfies all the constraints.

2.3 INTERIOR PENALTY FUNCTION IMPLEMENTATION

Although the algorithm is straight forward, there are a number of points to be considered in implementing this method. These are:

(i) The starting feasible point X_1 may not be readily available in some cases.

(ii) A suitable value of the initial penalty parameter (r_1) has to be found.

(iii) A proper value has to be selected for the multiplication factor, c.

(iv) Suitable convergence criteria have to be chosen to identify the optimum point.

(vi) The constraints have to be normalized so that each one of them vary between −1 and 0 only.

All these aspects are discussed in the following sub headings.

(a) STARTING FEASIBLE POINT X_1

There may be some situations where the feasible design points could not be found easily. In such cases, the required feasible starting points can be found by using the interior penalty function method itself as follows:

(i) Choose an arbitrary point X_1 and evaluate the constraints $g_i(x)$ at the point X_1. Since the point X_1 is arbitrary, it may not satisfy all the constraints with strict inequality sign. If r out of a total of m constraints are violated, renumber the constraints such that the last r constraints will become the violated ones, that is,

$$\left.\begin{array}{ll} g_i(x) < 0 & i = 1,2,\ldots, m-r \\ \text{and} & \\ g_i(x) \geq 0 & i = m-r+1, m-r+2,\ldots, m \end{array}\right\} \quad 2.3.1$$

However, there may exist situations in which the solution of the problem formulated in step (iii) gives the unconstrained or constrained local minimum of $g_k(X)$ that is positive. In such cases, one has to start afresh with a new point X_1 from step(i) onwards.

(b) INITIAL VALUE OF THE PENALTY PARAMETER (r_1)

Since the unconstrained minimization of ϕ (X, r_k) is to be carried out for a decreasing sequence of r_k, it might appear that by choosing a very small value of r_1, we can avoid excessive number of minimizations of the function Φ. But from computational point of view, it will be easier to minimize the unconstrained function ϕ ϕ (X, r_k) if r_k is large. However, the minimum of ϕ_k, X_k^*, will be farther away from the desired minimum X^* if r_k is large. Thus it requires an excessive number of unconstrained minimizations of ϕ (X, r_k) (for several values of r_k) to reach the point X^* if $r_1$is selected to be very large. Thus a "moderate" value has to be chosen for the initial penalty parameter (r_1). In practice, a value of r_1 , which gives the value of

ϕ (X_1, r_1) approximately equal to 1.1 to 2.0 times the value of $f(X_1)$ has been found to be quite satisfactory in achieving quick convergence of the process. Thus, for any initial feasible starting point X_1, the value of r_1 can be taken as

$$r_1 \approx 0.1 \text{ to } 1.0 \quad \frac{f(X_1)}{\left(-\sum_{i=1}^{m} \frac{1}{ci(X_1)}\right)} \qquad 2.3.4$$

(c) SUBSEQUENT VALUES OF THE PENALTY PARAMETER.

Once the initial value of r_k is chosen, the subsequent values r_k of have to be chosen such that

$$r_{k+1} < r_k \qquad 2.3.5$$

For convenience, the value of r_k are chosen according to the relation

$$r_{k+1} = c\, r_k \qquad 2.3.6$$

where $c < 1$. The value of c can be taken as 0.1 or 0.2 or 0.5 etc.

(c) CONVERGENCE CRITERIA.

Since the unconstrained minimization of ϕ (X, r_k) has to be carried out for a decreasing sequence of values of , it is necessary to use proper convergence criteria to identify the optimum point and to avoid unnecessarily large number of unconstrained minimizations. The process can be terminated whenever the following conditions are satisfied.

(i) The relative difference between the values of the objective function obtained at the end of any two consecutive unconstrained minimization falls below a small number ε_1 i. e.

$$\left| \frac{f(X_k^*) - f(X_{k-1}^*)}{f(X_k^*)} \right| \leq \varepsilon_1 \qquad 2.3.7$$

(ii) The difference between the optimum points X^*_k and X^*_{k-1} become very small. This can be judged in several ways. Some of them are given below:

$$\left| (\Delta X)_j \right| \leq \varepsilon_2 \qquad 2.3.8$$

where $\Delta X = X^*_k - X^*_{k-1}$ and $(\Delta X)_j$ the jth component of the vector ΔX.

$$\text{Max} \left| (\Delta X)_j \right| \leq \varepsilon_3 \qquad 2.3.9$$

$$\Delta X = [\, (\Delta X)^2_1 + (\Delta X)^2_2 + \ldots + (\Delta X)^2_n]^{1/n} \leq \varepsilon_4 \qquad 2.3.10$$

Note that the values of ε_1 to ε_4 have to be chosen depending on the characteristics of the problem at hand.

(e) NORMALIZATION OF CONSTRAINTS:

It is advisable to normalize the constraints so that they vary between –1 and 0 as far as possible. For example, a structural optimization problem might be having constraints on deflection(δ) and the stress(σ) as

$$C_1(x) = \delta(x) - \delta_{max} \leq 0 \qquad 2.3.11$$

$$C_2(x) = \sigma(x) - \sigma_{max} \leq 0 \qquad 2.3.12$$

The normalization can be done as

$$C'_1(x) = \frac{C_1(x)}{\delta_{max}} = \frac{\delta(x)}{\delta_{max}} - 1 \leq 0 \qquad 2.3.13$$

And

$$C'_2(x) = \frac{C_2(x)}{\sigma_{max}} = \frac{\sigma(x)}{\sigma_{max}} - 1 \leq 0 \qquad 2.3.14$$

The problem can still be solved effectively without normalizing the constraints by defining different penalty parameters for different constraints as:

$$\Phi(X,r_k) = f(X) - r_k \sum_{i=1}^{m} \frac{R_i}{C_i(x)} \qquad 2.3.15$$

Where R_1, R2, ...,R_m are selected such that the contributions of different $C_i(X)$ to the

ϕ – function will be approximately same as the initial point X_1. when the unconstrained minimization of ϕ (X,r_k) is carried for a decreasing sequence of values of r_k, the values of R_1, R2, ...,R_m will not be altered; however, they are expected to be effective in reducing the disparities between the contributions of the various constraints to the

ϕ – function.

Example 2.3.1

Minimize $f(x_1, x_2) = 1/3\,(x_1 + 1)^3 + x_2$

Subject to $C_1(x_1, x_2) = -x_1 + 1 \leq 0$

$C_1(x_1, x_2) = -x_2 \leq 0$

Solution:

To illustrate the interior penalty function method, we use calculus method for solving the unconstrained minimization problem in this case. Hence there is no need to have an initial feasible point X_1. the ϕ – function is

$$\Phi(X, r) = 1/3\,(x_1 + 1)^3 + x_2 - r\left(\frac{1}{-x_1+1} - \frac{1}{x_2}\right)$$

To find the unconstrained minimum of Φ, we use the necessary conditions:

$$\frac{\partial \Phi}{\partial x_1} = (x_1 + 1)^2 - \frac{r}{(1-x_1)^2} = 0 \quad \text{i e} \quad (x_1^2 - 1)^2 = r$$

$$\frac{\partial \Phi}{\partial x_1} = 1 - \frac{r}{x_2^2} = 0 \quad \text{i e} \quad x_2^2 = r$$

This equations gives

$$X^*_1(r) = (r^{1/2} + 1); \quad X^*_2(r) = r^{1/2}$$

and

$$\phi_{min}(r) = \left\{1/3\,[\,(r^{1/2} + 1)^{1/2} + 1]^3 + \frac{1}{2r^{1/2} - 1/r - (1/r^{3/2} + 1/r^2)^{1/2}}\right\}$$

To obtain the solution of the original problem, we know that

$$f_{min} = \lim_{r\to 0} \phi_{min}(r)$$

$$X_1^* = \lim_{r\to 0} X_1^*(r)$$

$$X_2^* = \lim_{r\to 0} X_2^*(r)$$

The values of f, X_1^* and X_2^* corresponding to a decreasing sequence of values of r are shown in the following table.

Values of r	$X_1^*(r) = (r^{1/2} + 1)^{1/2}$	$X_2^*(r) = r^{1/2}$	$\phi_{min}(r)$
1000	5.71164	31.62278	376.2636
100	3.31662	10.0000	89.9772
10	2.04017	3.16228	25.3048
1	1.41421	1.00000	9.1046
0.1	1.14727	0.31623	4.6117
0.01	1.04881	0.10000	3.2716
0.001	1.01569	0.03162	2.8569
0.0001	1.00499	0.01000	2.7267
0.00001	1.00158	0.00316	2.6856
0.000001	1.00050	0.00100	2.6727
Exact solution 0	1	0	8/3

The interior penalty function method illustrated in the above table gives convergence as the parameter r_k is decreased sequentially. The sequential process permits a graded approximation to be used in the analysis in the system. It can also be seen that k (iteration) did not exceed 10 before computer rounding error becomes significant. Furthermore, the value of the function ϕ is greater than f since $C_i(x)$ is negative for all feasible points X.

2.4: INTERIOR PENALTY FUNCTION STRENGHTS.

This section examines some properties of the basis algorithm which can be properties of any penalty function method. In general, the results can be considered as strengths of the methods.

(a) INTERIOR PENALTY FUNCTION CONVERGENCE.

It has been shown under mild topological conditions and for r_k sufficiently small, that a sequence of minimizing points $\{X_k\}$, produced by the basis algorithm and corresponding to $\{r_k\}$, exist and $X_k \rightarrow X^*$ as $k \rightarrow \infty$. Furthermore, it is easy to establish the following results which confirm the earlier observation concerning the behaviour of the interior penalty formulation (2.2.2)

(i) $\lim_{k \to \infty} r_k \sum_{i=1}^{m} G_i(c_i(x_k)) = 0$

(ii) $\lim_{k \to \infty} F(X_k) = F(X^*)$ and $\{F(X_k)\}$ is monotonic decreasing.

(iii) $\lim_{k \to \infty} \phi(X_k, r_k) = F(X^*)$

(iv) $\left\{ \sum_{i=1}^{m} G_i(c_i(x_k)) \right\}$ is monotonic increasing.

To give some idea of the generality of the basic algorithm, topological conditions required amount to continuity of the problem functions and an assumption that X^* is the closure of X_1. The condition implies that the interior penalty function approach will not converge to a local minima which are isolated points of the feasible region. In establishing convergence of the basic algorithm, we do not require assumptions as strong as Kuhn – Tucker constraint qualifications. Indeed, the interior penalty function method will converge to a local minimum at which the Kuhn – Tucker first order conditions fails to hold. The generality of these convergence results suggest that the basic algorithm has wide applicability.

(b) AN ERROR ESTIMATE OR BOUND.

Under stronger conditions on the problem functions it is possible to derive an estimate of

$(F(X(r)) - F(X^*))$. At $X(r)$ we have

$$b(X(r)) \doteq \sum_{i \in J} \lambda_i(r)\, a_i(x(r)) + 0(r) \qquad 2.4.1$$

where the $0(r)$ term represents the contribution from inactive constraints. Taking the scalar product of Eq. (2.4.1) with $(X(r) - X^*)$ and using the first – order Taylor expansions about X^*, we obtain

$$F(X(r)) - F(X^*) = \sum_{i \in J} \lambda_i(r)\, c_i(X(r)) + 0(\max[r \,||\, X(r) - X^* \,||^2]) \qquad 2.4.2$$

If the condition of strict complementary $\lambda_i^* > 0$, $i \in J$, holds, it is possible to show for the log interior penalty function that the first term on the right – hand side of (2.4.2) dominates and the order term can be ignored. When the active constraint set is unknown, the sum in (2.4.2) can be extended over all constraints. In this case we would expect

$$\sum_{i=1}^{m} \lambda_i(r)\, C_i(X(r)$$

to be bounded for the error $F(X(r)) - F(X^*)$.

(c) ROBUSTNESS.

Mifflin (1972) has established under much stronger conditions on the problem functions that the basic algorithm, using log interior penalty function, will converge even when $\phi(X, r_k)$ is "minimized" only to within a predetermined tolerance of $\phi(X(r), r_k)$. the rate of convergence with respect to r_k is, however, adversely affected by the magnitude of the tolerance. The importance of this result lies in the observation that in exact minimizations are an unavoidable consequence of any numerical implementation of the basic algorithm exhibits a certain robustness in overcoming this form of systematic error.

2.5 INTERIOR PENALTY FUNCTION WEAKNESS.

In contrast to the desirable properties of interior function method discussed in section (2.4), it is also possible to identify a number of undesirable properties of this method. These weaknesses of the basic algorithm are of a computational in nature and are most serious when the controlling parameter is small. In this section, we shall consider some of such weakness.

(a) REPEATED UNCONSTRAINED MINIMIZATIONS.

The penalty function method requires repeated unconstrained minimization of $\phi(X,r_k)$ with no strong indication of how r_k should be chosen. The choice of r1 and the rate at which r_k tends to zero can seriously affect the computational effort required to find the solution X*. we must therefore seek a compromise between the use of I few very difficult minimizations when r_k is chosen to converge rapidly and a large number of less difficult minimization when r_k is chosen to converge slowly.

(b) UNCONSTRAINED MINIMIZATION DIFFICULTIES.

The direct application of a typical unconstrained minimization algorithm to minimize the penalty function is seldom satisfactory, especially when the controlling parameter becomes small and the minimizing point of the penalty function approaches the boundary of the feasible region.

Three associated difficulties can be distinguished.

(i) The unconstrained minimization algorithm, in general, assumes the objective function is defined on E^n which contains the minimum as an interior point. Although this is true of the interior penalty function, it is also true that

$X(r_k)$ approaches the boundary of X as r decreases. This implies that linear search procedures within the algorithm will need to be modified to accommodate attempts to evaluate $\phi(X, r_k)$ at infeasible points. The simple expedient of setting the function value to "infinity" is unlikely to be sufficient of this value is then to determine an interpolating function to perform the linear search or if the function value is used to compute or update derivative information about $\phi(X, r_k)$

(ii) the linear search procedures incorporated in many unconstrained algorithms are usually performed by fitting quadratic or cubic – extrapolation and interpolation functions to function values and sometimes gradient values. It is impractical to expect such functions to model accurately the behaviour of the interior penalty function with its singularity at the boundary of the feasible region.

(iii) A third difficulty in minimization $\phi(X, r_k)$ using a standard unconstrained minimization algorithm arises in the specification of convergence criteria which are satisfactory for all values of the controlling parameter. As r_k decreases the penalty function becomes extremely steep valleyed and in a very small neighbourhood of X(r), the magnitude of the gradient $\phi(X, r_k)$ can take large values. In these circumstances, it is practically impossible to reduce the gradient magnitude numerically and this implies that the convergence criterion based solely on gradient magnitudes is unsuitable. Care must also be taken to avoid premature convergence resulting from the use of tests based on objective function improvement or step lengths in successive iterations of the unconstrained algorithm.

(c) ILL – CONDITIONING OF THE HESSIAN MATRIX.

A third weakness of the interior penalty function method is related to the ill – conditioned nature of its Hessian matrix. An analysis of the form of its Hessian (matrix of second partial derivatives) of the penalty function shows clearly that the matrix becomes increasingly ill – conditioned as $r_k \to \infty$. This would lead us to expect that interior penalty methods would perform poorly – yield in accurate result – as the solution is approached.

2.6: EXTERIOR PENALTY FUNCTION METHOD.

The exterior penalty function formulation, in contrast to the interior function formulation discussed in section 2.2 to 2.5, is defined to impose an increasing penalty on the objective function as constraint violation increases. The controlling parameter is effectively to increase the magnitude of the penalty. In the exterior penalty function method, the ϕ - function is generally taken as

$$\phi(X, r_k) = f(X) + rk \sum_{i=1}^{m} \langle gi(X) \rangle^q \qquad 2.6.1$$

where r_k is a positive penalty parameter, the exponent q is a non negative constant, and the bracket function $\langle g_i(X) \rangle$ is defined as

$$\langle gi(X) \rangle = \max\langle gi(x), 0 \rangle = \begin{cases} g_i(x) \text{ if } g_i(x) > 0 & \text{constrained violated} \\ 0 \text{ if } g_i(x) \leq 0 & \text{constrained is satisfied} \end{cases} \qquad 2.6.2$$

Usually, the function $\phi(X, r_k)$ possesses a minimum as a function of X in the infeasible region. The unconstrained minima X^*_k converge to the solution of the original problem as $k \to \infty$ and $r_k \to \infty$. Thus the unconstrained minima approach the feasible domain gradually and as $k \to \infty$, the X^*_k eventually lies in the feasible region.

Example 2.6.1

$G_i = \max [0, g_i(x)]$ 2.6.3

and

$G_i = \{\max [0, g_i(x)] \}^2$ 2.6.4

2.6.5 ALGORITHM

The exterior penalty function method can be stated by using the following steps. (i) Start from any design X_1 and a suitable value of r_1 . set $k = 1$

(ii) find the vector X^*_k that minimizes the function

$$\phi(X, r_k) = f(X) + r_k \sum_{i=1}^{m} \langle gi(X) \rangle^q$$

(iii) Test whether the point X^*_k satisfies all the constraints. If X^*_k is feasible, it is the desired optimum and hence terminate the procedure otherwise go to step (iv)

(iv) Choose the next value of the penalty parameter which satisfies the relation

$$r_{k+1} > r_k$$

and set the new value of k as original k plus one and go to step (ii).

Example 2.6.2:

Minimize

$$F(x_1,x_2) = 1/3\ (x_1 + 1) + x_2$$

Subject to $C_1(x_1 , x_2) = 1 - x_1 \leq 0$

$C_1(x_1 , x_2) = -x_2 \leq 0$

Solution:

To illustrate the exterior penalty function method, we solve the unconstrained minimization problem by using differential calculus method. As such, it is not necessary to have an initial trial point X_1. The ϕ - function is

$\phi(X, r_k) = 1/3\,(x_1 + 1)^3 + x_2 + r\,[\max(0, 1 - x_1)]^2 + r[\max(0, -x_2)]^2$

The necessary conditions for the unconstrained minimum of $\phi(X, r)$ are

$$\frac{\partial\phi}{\partial x_1} = (x_1 + 1)^2 - 2\,r[\max(0,1 - x_1)] = 0$$

$$\frac{\partial\phi}{\partial x_1} = 1 - 2\,r[\max(0, -x_2)] = 0$$

These equations can be written as

$\text{Min}\,[(x_1 + 1)^2, (x_1 + 1)^2 - 2\,r(1 - x_1)] = 0$ E_1

And

$\text{Min}\,[1, 1 + 2r\,x_2] = 0$ E_2

In Eq. (E_1), if $(x_1 + 1)^2 = 0$, $x_1 = -1$ (This violates the first constraint) and if

$(x_1 + 1) - 2r(1 - x_1) = 0$, $x_1 = -1 - r + \sqrt{r^2 + 4r}$

In Eq.(E_2), the only possibility is that

$1 + 2r\,x_2 = 0$ and hence

$x_2 = -\frac{1}{2}r$

Thus the solution of the unconstrained minimization problem is giving by

$X^*_1(r) = -1 - r + r(1 + 4/r)^{1/2}$ E_3

And

$X^*_2(r) = -1/2r$ E_4

From this, the solution of the original constrained problem can be obtained as

$$X_1^* = \lim_{r\to\infty} X_1^*(r) = 1, \quad X_2^* = \lim_{r\to\infty} X_2^*(r) = 0$$

and

$$f_{min} = \lim_{r\to\infty} \phi_{min}(r) = 8/3$$

The converge of the method, as r increases gradually, can be seen from the following table.

Value of r	X_1^*	X_2^*	$\phi_{min}(r)$	$f_{min}(r)$
0.001	-0.93775	-500.00000	-249.9962	-500.0000
0.01	-0.80970	-50.00000	-24.9650	-49.9977
0.1	-0.45969	-5.0000	-2.2344	-4.9474
1	0.23607	-0.5000	0.9631	0.1295
10	0.83216	-0.05000	2.3068	2.0001
100	0.98039	-0.00500	2.6249	2.5840
1000	0.99800	-0.00050	2.6624	2.6582
10000	0.99963	-0.00005	2.6655	2.6652
∞	1	0	8/3	8/3

It can be seen from the table above that the unconstrained minima of ϕ - converge to the optimum X* as the parameter r_k is increased sequentially. We can also observe that the sequential nature of the method allows a gradual or sequential approach to critically of the constraint.

REMARK 1.6.6

The interpretation of the behaviours of Exterior penalty function method as r_k decreases is similar to that given for the barrier (Interior penalty) function. In particular, the basic algorithm can again be applied and equivalent theoretical result can be established to ensure convergence. The only difference in the basic algorithm for exterior penalty function is that we no longer require a feasible initial point X_1 in step (i). It is also possible to identify computational difficulties similar to those discussed for interior penalty function method in section 2.5.

Again these difficulties result from $\phi(X, r_k)$ forming an increasing steep- sided valley as the controlling parameter decreases. In particular, the Hessian matrix of the exterior penalty function (2.6.2) can be shown to be ill- conditioned as r_k decreases.

There are recent contributions in the field of penalty function methods for non linear programming which can be interpreted as attempts to overcome the computational difficulties associated with the basic penalty function methods as implemented in the basic algorithm. This study intends to examine the extrapolation procedures and show it attempts to improve the performance of the basic algorithm in the subsequent chapters.

CHAPTER THREE

EXTRAPOLATION METHOD IN INTERIOR PENALTY METHOD

3.1 INTRODUCTION

In the interior penalty function method, the ϕ - function is minimized sequentially for a decreasing sequence of values $r_1 > r_2 > \ldots > r_k$ to find the unconstrained minima X^*_1, $X^*, \ldots, X^*_k$ respectively. Let the values of the objective function corresponding to $X^*_1, X^*_2, \ldots, X^*_k$ be f^*_1, $f^*_2, \ldots, f_k$ respectively. It has been proved that the sequence X^*_1, $X^*_1, \ldots, X^*_k$ converges to the minimum value f^* of the original constrained problem stated in (2.1.1) as $r_k \rightarrow 0$. After carrying out a certain number of unconstrained minimizations of ϕ, the results obtained thus far can be used to estimate the minimum of the original constrained problem by a method known as the extrapolation technique. Looking at the examples in chapter two, one might expect that when a set of local minima consists of one point a unique trajectory of unconstrained local minima exists converging to that point. Furthermore, one would hope that an examination of a few points on such a trajectory would give information about the final point, the local minimum to which it is converging.

Thus the idea of extrapolation is to use the solutions of the previous interior penalty sub problems to fit a polynomial to the interior penalty trajectory.

It is the intention of this chapter to consider the extrapolation of the design vector and the objective function. Furthermore, this chapter intends to explore the feasibility of extrapolation point and how extrapolation formulas can be used to accelerate convergence.

Let X* be a local solution of (2.1.1); we define

τ = 0(r) if there exist some positive constant k so that

$\|\tau\| \leq$ kr $\forall$ sufficiently small r $>$ 0.

We also make use of the assumption that the functions f and g have three continuous derivatives.

3.2 EXTRAPOLATION OF THE DESIGN VECTOR X

Since different vectors X^*_i, i = 2,...,k are obtained as unconstrained minima of ϕ(X, r) for any value of r, X*(r), can be approximated by a polynomial in r as:

$$X^*(r) = \sum_{i=0}^{k-1} A_j (r)^j = Ao+rA_1\, r^2A_2+\; +r^{k-1}A_{k-1} \qquad 3.2.1$$

where A_j are n-component vectors. By substituting the known conditions

$$X^*(r=r_i) = x^*_i \quad i=1,2,...,k \qquad 3.2.2$$

in (3.2.1), we can determine the vectors A_j j = 0,1,2,...,k-1 uniquely. Then x*(r), given by (3.2.1) will be a good approximation for the unconstrained minimum of ϕ(x, r) in the interval (0,r_1). By setting r = 0 in (3.2.1) we can obtain an estimate to the true minimum X*, as

$$X^* = x^*(r=0) = A_0 \qquad 3.2.3$$

It is to be noted that it is not necessary to approximate X*(r) by a (k-1)st order polynomial in r. Infact, any polynomial of order 1$\leq$ P $\leq$ K-1 can be used to approximate X*(r). In such a case, we need only P+1 points out of

$X^*_1, X^*_2, \ldots, X^*_K$ to completely define the polynomial.

As a simplest case, let us consider approximating $X^*(r)$ by a first order polynomial (linear equation) in r as

$$X^*(r) = A_0 + r A_1 \qquad 3.2.4$$

To evaluate the vectors A_0 and A_1, we need the data of two unconstrained minima. If the extrapolation is being done at the end of Kth unconstrained minimization, we generally use the latest information to find the constant vectors A_0 and A_1. Let X^*_{k-1} and x_k^* be the unconstrained minima corresponding to r_{k-1} respectively. Since

$r_k = cr_{k-1}$ $(c < 1)$, (3.2.4) gives

$$\left.\begin{aligned} x^*(r=r_{k-1}) &= A_0 + r_{k-1}A_1 = X^*_{k-1} \\ x^*(r=r_k) &= A_0 + cr_{k-1}A_1 = X^*_k \end{aligned}\right\} \qquad 3.2.5$$

These equations give

$$\left.\begin{aligned} A_0 &= \frac{X^*_k - cX^*_{k-1}}{1 - c} \\ &\text{and} \\ A_1 &= \frac{X^*_{k-1} - X^*_k}{R_{k-1}(1 - c)} \end{aligned}\right\} \qquad 3.2.6$$

From (3.2.4) and (3.2.6), the extrapolated value of the true minimum can be obtained as

$$X^*(r=0) = A_0 = \frac{X^*_k - cX^*_{k-1}}{1 - c} \qquad 3.2.7$$

The extrapolation technique (3.2.1) has several advantages:

a. It can be used to find a good estimate to the optimum of the original problem with the help of (3.2.1)

b. It can be used to provide an additional convergence criterion to terminate the minimization process. The point obtained at the end of Kth iteration, X^*_k, can be taken as the true minimum if the relation

$$|X^*_k - X^*(r=0)| \leq \epsilon \qquad 3.2.8$$

is satisfied, where ϵ is the vector of prescribed small quantities.

c. This method, can be used as to estimate the next minimum of the ϕ - function after a number of minimizations have been completed. The estimate obtained form X^* can also be used as a starting point for the (k+1)st minimization of the ϕ - function. The estimate of the (k+1)st minimum, based on the information collected from the previous k minima, is given by (3.2.1) as:

$$X^*_{k-1} \approx X^* (r= r_{k+1}= r_1c^k) = A_0 + (r_1c^k)\, A_1 + (r_1c^k)^2\, A_2 + \dots + A_{K-1}(r_1c^k)^{k-1} \qquad 3.2.9$$

If (3.2.4) and (3.2.6) are used, the estimate becomes

$$X_{k+1} \approx X^* (r=c^2=c^2r_{k-1}) = A_0=c^2r_{k-1}\, A_1 = (1+c)\, X^*_k - cX^*_{k-1} \qquad 3.2.10$$

Sometimes the extrapolation made with the help of a (k-1)st order polynomial in $r^{1/2}$ rater than in r are found to be better. Here the approximation is taken as

$$X^* (r) = \sum_{j=0}^{k-1} A_j\, (r^{1/2})^j = A_o+r^{1/2} A_1\; r A_2+\dots+ r^{(k-1)/2} A_{k-1} \qquad 3.2.11$$

The vectors $A_0, A_1, \dots, A_{k-1}$ are evaluated with the help of the unconstrained minima $X^*_1, X^*_2, \dots, X^*_k$ corresponding to $r_1, r_2, \dots, r_k$ respectively.

By setting r= 0 (3.2.11) gives the estimate of X* as

$$X^* \cong X^*(r=0) = A_0 \qquad 3.2.12$$

If only two terms are considered in (3.2.11), we have

$$X^*(r) = A_o + r^{1/2} A_1 \qquad 3.2.13$$

If X^*_{k-1} and X^*_k are the unconstrained minima of ϕ corresponding to r_{k-1} and $r_k = cr_{k-1}$, we obtain, using (3.2.13),

$$\left.\begin{aligned} A_0 &= \frac{X^*_k - c^{1/2} X^*_{k-1}}{(1 - c^{1/2})} \\ A_1 &= \frac{X^*_{k-1} - X^*_k}{r^{1/2}_{k-1}(1 - c^{1/2})} \end{aligned}\right\} \qquad 3.2.14$$

Thus the estimates of X* and X^*_{k+1} can be found from (3.2.13) as

$$X^* \approx x^*(r=0) = A_0 = \frac{X^*_k - c^{1/2} X^*_{k-1}}{1 - c^{1/2}} \qquad 3.2.15$$

and

$$\begin{aligned} X^*_{k+1} \approx X^*(r = r_{k+1} = c^2 r_{k-1}) &= A_0 + C r^{1/2}_{k-1} A_1 \\ &= (1-c^{1/2}) X^*_k - c^{1/2} X^*_{k-1} \qquad 3.2.16 \end{aligned}$$

By using the point obtained in either (3.2.15) and (3.2.16) a starting point for the (k+1)st minimization of ϕ - function, the overall convergence of the process can be speeded up.

3.3 EXTRAPOLATION OF THE OBJECTIVE FUNCTION, F

It is possible to use extrapolation technique to estimate the optimum value of the original objective function, f*. For this, let $f^*_1, f^*_2, \ldots, f^*_k$. be the values of the objective function corresponding to the vectors X^{*1}, $X^*_2, \ldots, X^*_k$. Since the points $X^*_1, X^*_2, \ldots, X^*_k$ have been found to be the unconstrained minima of the ϕ - function corresponding to $r_1, r_2, \ldots, r_k$, respectively, the objective function, f^*, can be assumed to be a function of r. By approximating f^* by a (k-1)st order polynomial in k, we have

$$f^*(r) = \sum_{j=0}^{k-1} a_j (r)^j = a_o + a_1 r + a_2 r^2 + \ldots + a_{k-1} r_{k-1} \qquad 3.3.4$$

where the k constants a_j, $j = 0,1,2,\ldots, k-1$ can be evaluated by substituting the known conditions:

$$F^*(r=r_i)=f^*_i=a_0 + a_1 r_i + a_2 r_i^2 + \ldots + a_{k-1} r_i^{k-1}, \quad i = 1,2,\ldots, k-1 \qquad 3.3.2$$

Since (3.3.1) is a good approximation for the true f^* in the interval $(0,r_1)$, we can obtain an estimate for the constrained minimum of f as

$$f^* \approx f^*(r=0) = a_0 \qquad 3.3.3$$

As a particular case, a linear approximation can be made for f^* by using the last two data points. Thus if f^*_{k-1} and f^*_k are the function values corresponding to r_{k-1} and $r_k = c\, r_{k-1}$ we have

$$\left.\begin{aligned} f^*_{k-1} &= a_0 + r_{k-1} a_1 \\ &\text{and} \\ f^*_k &= a_0 + c\, r_{k-1} a_1 \end{aligned}\right\} \qquad 3.3.\ 4$$

These equation yield

$$a_0 = \frac{f^*_k - cf^*_{k-1}}{1 - c} \qquad 3.3.5$$

$$a_1 = \frac{f^*_{k-1} - f^*_k}{r_{k-1}(1 - c)} \qquad 3.3.6$$

and

$$f^*(r) = \frac{f^*_k - cf^*_{k-1}}{1 - c} + \frac{r \ (f^*_{k-1} - f^*_k)}{r_{k-1}\ (1 - c)} \qquad 3.3.7$$

(3.3.7) gives an estimate of f* as

$$f^* \approx f^*\ (r=0) = a_0 = \frac{f^*_k - cf^*_{k-1}}{1 - c} \qquad 3.3.8$$

Instead of (3.3.1), extrapolation are sometimes made by assuming a (k-1)st order polynomial in $r^{1/2}$ for f* (r) as

$$f^*\ (r) = a_0 + a_1 r^{1/2} + a_2\ r + \ldots + a_{k-1}\ r^{\ (k-1)/2} \qquad 3.3.9$$

If we retain only two terms in (3.3.9), we obtain, as in the case of (3.2.13),

$$f^*(r) = a_0 + r^{\ 1/2} a_1 \qquad 3.3.10$$

with

$$a_0 = \frac{f^*_k - c^{1/2} f^*_{k-1}}{1 - c^{\ 1/2}} \qquad 3.3.11$$

and

$$a_1 = \frac{f^*_{k-1} - f^*_k}{r^{1/2}_{k-1}(1 - c^{\ 1/2})} \qquad 3.3.12$$

(3.3.10) gives the estimate of f* as

$$f^* \approx f^*\ (r=0) = \frac{f^*_k - c^{1/2} f^*_{k-1}}{1 - c^{\ 1/2}} \qquad 3.3.13$$

The extrapolated value as can be used to provide an additional convergence criterion for terminating the interior penalty function method. The criterion is that whenever the value of f^*_k obtained at the end of kth unconstrained minimization of ϕ is sufficiently close to the extrapolated valve as i.e when

$$\left| \frac{f^*_k - a_0}{f^*_k} \right| \leq \in \qquad 3.3.14$$

where $\in$ is a specified small quantity, the process can be terminated.

Example 3.3.1

The following results are obtained during the minimization of

$f(x) = x^3_1 - 6x^2_1 + 11x_1 + x_3$ subject to

$$X^2_1 + x^2_2 - x^2_3 \le 0$$

$$4 - x^2_1 + x^2_2 - x^2_3 \le 0$$

$$-x_i \le 0, \ i = 1,2,3$$

using the interior penalty function method

Value of r_k	Startingpoint for minimization of ϕ_k	Unconstrained minimum of $\phi(X_1, rk) = X^*_k$	f^*_k
1.0×10^0	$\begin{pmatrix} 0.1 \\ 0.1 \\ 3.0 \end{pmatrix}$	$\begin{pmatrix} 0.37898 \\ 1.67965 \\ 2.34617 \end{pmatrix}$	5.70766
1.0×10^{-1}	$\begin{pmatrix} 0.37898 \\ 1.67965 \\ 2.34617 \end{pmatrix}$	$\begin{pmatrix} 0.10088 \\ 1.41945 \\ 1.68302 \end{pmatrix}$	2.73267
1.0×10^{-2}	$\begin{pmatrix} 0.10088 \\ 1.41945 \\ 1.68302 \end{pmatrix}$	$\begin{pmatrix} 0.03066 \\ 1.41411 \\ 1.49842 \end{pmatrix}$	1.83012
1.0×10^{-3}	$\begin{pmatrix} 0.03066 \\ 1.41411 \\ 1.49842 \end{pmatrix}$	$\begin{pmatrix} 0.009576 \\ 1.41419 \\ 1.44081 \end{pmatrix}$	1.54560

Find the extrapolated values of X and f using the result of minimization of

$\phi(x, r_1)$ and $\phi(x, r_2)$.

Solution:

From the result above we have for $r_1 = 1.0$

$$X^*_1 = \begin{Bmatrix} 0.37898 \\ 1.67965 \\ 2.34617 \end{Bmatrix} \qquad f^*_1 = 5.70766$$

and for $r_2 = 0.1 \quad c = 0.1$

$$X^*_2 = \begin{Bmatrix} 0.10088 \\ 1.41945 \\ 1.68302 \end{Bmatrix}, \quad f^*_2 = 2.73267$$

If (3.2.13) is used to approximate X*(r), the extrapolated value of X* can be obtained from (3.2.15) as

$$X^* \approx A_0 = \frac{X^*_k - c^{1/2} X^*_{k-1}}{1 - c^{1/2}} = 1/0.6838 \left(\begin{Bmatrix} 0.10088 \\ 1.41945 \\ 1.68302 \end{Bmatrix} - 0.3162 \begin{Bmatrix} 0.37898 \\ 1.67965 \\ 2.34617 \end{Bmatrix} \right)$$

$$= \begin{Bmatrix} -0.0278 \\ 1.2990 \\ 1.3770 \end{Bmatrix} \qquad E_1$$

Similarly the use of a linear relation for f(r) in terms of $r^{1/2}$ gives

$$f^* \approx \frac{f^*_k - c^{1/2} f^*_{k-1}}{1 - c^{1/2}} = 1/0.6838 \left(2.73267 - 0.3162(5.70766) \right)$$

$$= 1.3570 \qquad E_2$$

If (3.2.4) is used for approximating x* (r), the extrapolated vector X* is given by (3.2.7) as

$$X^* \approx A_0 = \frac{X^*_k - cX^*_{k-1}}{1 - c} = 1/0.9 \left(\begin{Bmatrix} 0.10088 \\ 1.41945 \\ 1.60933 \end{Bmatrix} - 0.1 \begin{Bmatrix} 0.37898 \\ 1.67865 \\ 2.34617 \end{Bmatrix} \right)$$

$$= \begin{Bmatrix} 0.06998 \\ 1.39053 \\ 1.60933 \end{Bmatrix} \qquad E_3$$

Similarly the linear relationship $f^*(r) = a_o + a_1 r$ leads to

$$f^* \approx \frac{f^*_k - cf^*_{k-1}}{1 - c} = 1/0.9 \left[(2.73267 - 0.1(5.70766) \right]$$

$$= 2.40211 \qquad E_4.$$

By comparing the present values with the true optimum results given in the above table, it can be found that the linear approximation in terms of r gave better result compared to the linear equation in $r^{1/2}$. Further, the design vector X* predicted by the linear equation in $r^{1/2}$ can be seen to be infeasible whereas the one predicted by the linear equation in r can be seen to be feasible.

3.4: FEASIBILITY OF EXTRAPOLATION POINT

In spite of the above mentioned advantages, the extrapolation technique has a serious limitation. This is that the extrapolated points given by any of the Eq. (3.2.3), (3.2.5), (3.2.10), (3.2.15) and (3.2.16) may sometimes violate the constraints. Hence, we have to check any extrapolated point for feasibility before using it as a starting point for the next minimization of ϕ.

Let X^*_{k-1} and X^*_k be the unconstrained minima corresponding to r_{k-1} and r_k respectively. In the following lemma, we show that for a binding constraint (if we ignore all but the dominant terms)

$$g_j(X_{k+1}) \approx g_j(x)\frac{r_{k+1}}{r_k} < 0 \qquad 3.4.1$$

where X_{k+1} is the result of linear extrapolation at r_{k+1} i.e. X_{k+1} is the initial guess for the next unconstrained minimization.

For a non bonding constraints

$$g_j(X_{k+1}) = g_j(X) + 0(r) = 0(1) + 0(r) \qquad 3.4.2$$

and so again, if we ignore all but dormant terms

$$g_j(X_{k+1}) < 0. \qquad 3.4.3$$

Thus, we can expect the extrapolated point to be feasible.

Lemma 3.4.1

Let $g_j(X) \leq 0$ be a constraint that is binding at x^*, then

$$g_j(X_{k+1}) = g_j(X) \frac{r_{k+1}}{r_k} + 0\,(r^2_{k-1}) \qquad 3.4.3$$

Proof

Suppose that $g_j(X) \leq 0$ is binding at X^*.

Then since

$$X_{k+1} = X + \frac{r_{k+1} - r_k}{r_k - r_{k-1}} (X_k - X_{k-1})$$

We obtain

$$gj(X_{k+1}) = gj(X_k) + \frac{r_{k+1} - r_k}{r_k - r_{k-1}} \nabla gj(X_k)^T(X_k - X_{k-1}) + 0(\| X_k - X_{k-1} \|) \qquad 3.4.4$$

A Taylor series gives

$$gj(X_{k-1}) = gj(X_k) + \nabla gj(X_k)^T(X_{k-1} - X_k) + 0(\| X_{k-1} - X_k \|^2) \qquad 3.4.5$$

combining these formulas gives

$$g_j(X_{k+1}) = g_j(X_k) + \frac{r_{k+1} - r_k}{r_k - r_{k-1}} \left[g_j(X_k) - g_j(X_{k-1}) \right] + 0(\| X_{k-1} - X_k \|^2) \qquad 3.4.6$$

From Langrange multiplier method,

$$\left.\begin{aligned} \frac{r_{k-1}}{g_j(X_{k-1})} &= (\lambda^*)_j + 0(r_{k-1}) \\ &\text{and} \\ \frac{r_k}{g_j(X_k)} &= (\lambda^*)_j + 0(r_k) \end{aligned}\right\} \qquad 3.4.7$$

where $\lambda^* > 0$ is the Langrange multiplier for $g_j(X^*)$

Then

$$g_j(X_k) - g_j(X_{k-1}) = g_j(X_k) - \frac{r_{k-1}}{(\lambda^*)_j + 0(r_{k-1})}$$

$$= g_j(X_k) - \cfrac{g_{k\,1}}{\cfrac{r}{g_i(X_k)} - 0(r_k) + 0(r_{k-1})}$$

$$= g_i(X_k)\left(1 - \frac{r_{k-1}}{r_k + g_j(X_k)0(r_{k-1})}\right)$$

$$= g_i(X_k)\left(1 - \frac{r_{k-1}}{r_k + 0(r_k\, r_{k-1})}\right)$$

$$= g_i(X_k)\left(1 - \frac{r_{k-1}}{r_k} + 0(r_{k-1})\right) \qquad 3.4.8$$

combining this with the earlier result, we obtain

$$g_i(X_{k+1}) = g_i(X_k) + \frac{r_{k+1} - r_k}{r_k - r_{k-1}}\, g_i(X_k)\left(1 - \frac{r_{k-1}}{r_k} + 0(r_{k-1})\right) + (\|X_k - X_{k-1}\|) \qquad 3.4.9$$

since $r_{k+1} < r_k < r_{k-1}$ and $g_i(X_k) \le 0$

we obtain

$$g_i(X_{k+1}) = g_i(X_k)\left(1 - \frac{r_{k+1} - r_k}{r_k - r_{k-1}}\left(1 - \frac{r_{k-1}}{r_k}\right)\right) + 0(r^2_k - 1)$$

$$= g_i(X_k)\frac{r_{k+1}}{r_k} + 0(r^2_k - 1) \qquad 3.4.10$$

3.5: ACCELERATION BY EXTRAPOLATION

In this section, let the basic optimization problem be of the form

$$\left.\begin{array}{l}\text{Find X which minimizes f(X) subject to} \\ g_i(X) \ge 0 \quad j = 1,2, \ldots, m\end{array}\right\} \qquad 3.5.1$$

In the following discussion' the existence of $D^i\left(x(0)\right)$ and other appropriate conditions are assumed to hold.

Suppose the ϕ - function has been uniquely minimized for $r_1 > ... > r_p > 0$ at $x^*_1, ... x^*_p$. A polynomial in r that yields $x^*_1,..., x^*_p$ is given by a set of equations of the form

$$X^*(r_k) = x^*_k = \sum_{j=0}^{p-1} a_j (r_k)^j , \quad k = 1, ..., p \qquad 3.5.2$$

where the a_j are n – component vectors.

The determinant of the matrix

$$R = \begin{vmatrix} r^0_1, ..., r^0_p \\ . \\ . \\ . \\ r^{p-1}_1, ..., r^{p-1}_p \end{vmatrix}$$

(called the Vandermonde determinant) is equal to

$$\prod_{i<j} (r_j - r_i)$$

and since $r_j \neq r_i$ $(i \neq j)$, R is non singular. Thus the vectors a_i are uniquely determined by 3.5.2. Then

$$\sum_{j=0}^{p-1} a_j (r)^j$$ is an approximation of $X^*(r)$ in the interval $[0, r_1]$ and

$X^*(0) = X^*$(a solution) is approximated by a_0. That this approximation converges to a solution, and infact that the estimate improves with each minimum that is determined, is seen as follows.

USING TAYLOR SERIES:

The exact Taylor series expansion of X^*_k in r_k about 0 is

$$X^*_k = \sum_{j=0}^{p-1} (r_k)^j \frac{D^j x(0)}{j!} + \varepsilon^k \qquad k = 1,..., p \qquad 3.5.3$$

where

$$\varepsilon^k = \left(\frac{r^p{}_k}{p!}\right)\left(\frac{d^p x_1(\eta_1 k)}{dr^p}, \ldots, \frac{d^p x_n(\eta_n k)}{dr^p}\right) \quad 0 \le \eta_j k \le r_k \quad j = 1, \ldots, n$$

Setting (3.5.2) and (3.5.3) equal, subtracting, and combining yields

$$[\varepsilon^1, \ldots, \varepsilon^P]\, R^{-1} = A - \left(x(0), \frac{D^1 x(0)}{1!}, \ldots, \frac{D^{P-1} x(0)}{(P-1)!}\right),$$

where $A = (a_0, \ldots, a_{p-1})$. Clearly, then the difference between a_0 and $x^*(0)$ is of the order of $r_1{}^P$. thus as $r_1 \to 0$, $a_0 \to x^*(0)$. More important, the estimate using p minima are better than using (p-1) minima. When $r_{k+1} = r_k/c$ $(c>1)$, the particular structure of these equations makes it possible to develop a simple iterative scheme for computing a series of estimates based on using a giving number of terms in the polynomial. Notice that the a_j need not be computed to obtain these estimates.

Let $x_{i,j}$, $i = 1,\ldots,p$, $j = 0,\ldots,i-1$ signify the jth order estimate of x(0) after i minima have been achieved, with the understanding that r_1 is the initial value of r. note that the order corresponds to the index of a_j in (3.5.2).

Then it follows that

$$\left.\begin{aligned} X_{i,0} &= X\,\frac{r_1}{C^{i-1}} \qquad i = 1, \ldots, p \\ &\text{and} \\ X_{i,j} &= \frac{(C^j)\, X_{i,j-1} - X_{i-1,j-1}}{C^j - 1}, \quad i = 2, \ldots, p-1 \end{aligned}\right\} \qquad 3.5.4$$

Then a_0, the "best" estimate of X(0), is given by

$$X(0) = X_{p,p-1} = a_0 \qquad 3.5.5$$

The extrapolation formulas (3.5.4) can also be used to estimate the next minimum of the ϕ - function after a number of minima have been completed. For example, the (p+1)st minimum, based on the information collected from the previous p minima, is estimated by

$$X_{p+1,0} = x\left(\frac{r_1}{C^p}\right) = a_0 + \ldots + a_{j-1}\left(\frac{r_1}{C^p}\right)^{j-1} + \ldots + a_{p-1}\left(\frac{r_1}{C^p}\right)^{p-1} \qquad 3.5.6$$

Although the $\{a_j\}$ have not been computed explicitly, it is possible to make use of relations (3.5.4) and (3.5.5) and work backward to $X_{p+1,0}$. This is accomplished by setting i= p+1 in (3.5.4) and solving for $X_{p+1,j-1}$. This gives the recursive relation

$$X_{p+1,j-1} = \frac{(C^j - 1)X_{p+1,j} + X_{p,j-1}}{C^j} \qquad 3.5.7$$

Noting that $a_0 = X_{p,p-1} = X_{p+1,p-1}$ from (3.5.5) and using the values previously obtained from (3.5.4), we can evaluate (3.5.7) for j = p-1, p −2 ,...,1. The last computation will yield the required estimate $X_{p+1,0}$.

The following example shows how estimates can be used to accelerate convergence.

Examples 3.5.1

Minimize $\ln x_1 - x_2$

Subject to

$X_1 - 1 \leq 0$

$X_1^2 + X_2^2 - 4 = 0$

Solution

To illustrate the extrapolation technique, we use calculus method for solving the unconstrained minimization problem. In table 3.5.1 and the data for the iterations.

Table 3.5.1: Use of extrapolation of Accelerate convergence.

r_k	Estimates X^k_1 (1) (2) (3)	Estimates X^k_2 (1) (2) (3)	In X_1-X_2 f(x)
1.0	1.5527902	1.3328244	-0.8927710
0.25	1.1593310	1.6412662	-1.493523
	1.021779	1.7442134	-1.7164252
6.25×10^{-2}	1.0398432	1.7111091	-1.6720392
	1.0000139	1.7343567	-1.7343428
	0.9981363	1.7336995	-1.7355649
1.5625×10^{-2}	1.0099207	1.7269415	-1.7170697
	0.9999465	1.322189	-1.7322724
	0.9999420	1.7320763	-1.7321343
	0.999706	1.7320505	-1.7320799
3.960625×10^{-3}	1.0024774	1.730819	-1.7283076
	0.99999630	1.7320620	-1.7320657
	0.99999960	1.7320515	-1.7320519
	1.0000005	1.7320511	-1.7320506
Theoretical soln	1.0	$\sqrt{3}$ $\cong$1.7320505	$-\sqrt{3}$ $\cong$ -1.7320505

The convergence of the estimates to the theoretical solution can be seen by reading down the columns. The power of the extrapolation formulas can be summed up by noting that the third – order estimates using the last four minima (last line before theoretical solution) agree with the theoretical solution to seven places; whereas the minimum for r= 3.960625×10^{-3} agrees to only three places.

REMARK 3.5.8

This estimate can be used as the starting point for the (P+1)st minimization of the ϕ - function.

As more minima are achieved, the estimate eventually improves. This accelerates the entire process by substantially reducing the effort required to minimize the successive ϕ - functions.

In practice, computer storage requirements and accuracy considerations such as round – off error (which becomes critical for higher order estimates) limit the number of estimates possible. However, it has been found that considerable computational advantage is gained even when only first and second – order estimates are made of a next ϕ minimum and the optimum.

CHAPTER FOUR

COMPUTATIONAL EXPERIMENTATION OF EXTRAPOLATION ALGORITHM.

4.1: INTRODUCTION:

A vital test and justification of any body of theory of how to solve problems is the feasibility of computational implementation and practical application. In this chapter the computational questions implicit in the theoretical development of extrapolation algorithms are discussed.

The algorithm described will be described will be applicable to penalty function algorithm as discussed in chapter two. For definiteness it is assumed that the logarithmic penalty function is applied to the constraints where interior feasibility is required, and the quadratic less function as the exterior point penalty term.

Te problem to be solved is

Minimize $f(X)$ 4.1.1

Subject to $g_i(X) \leq 0$ $\quad i = 1, 2, \ldots,m$ 4.1.2

The unconstrained function has the form

$$\phi_k = \phi(X, r_k) = f(X) + \sum_{i=1}^{m} G_i(g_i(x)) \qquad 4.1.3$$

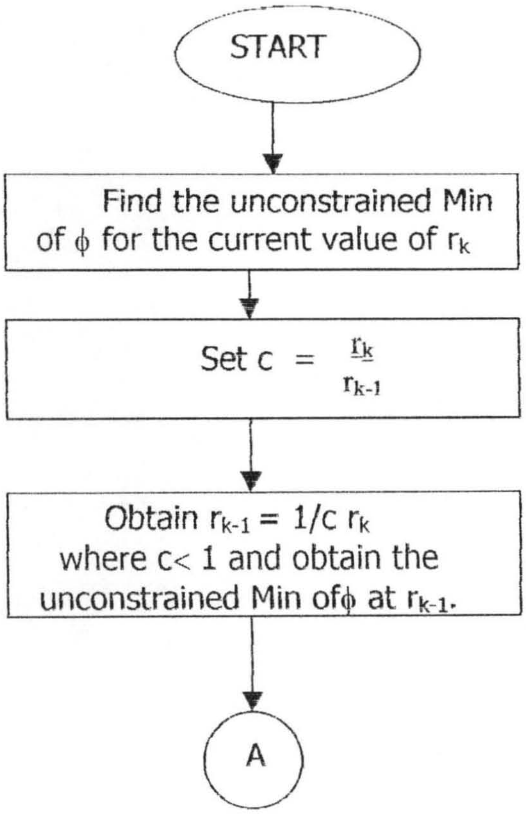
it is also the intention of this chapter to examine computational experiments and analysis of error due to extrapolation.

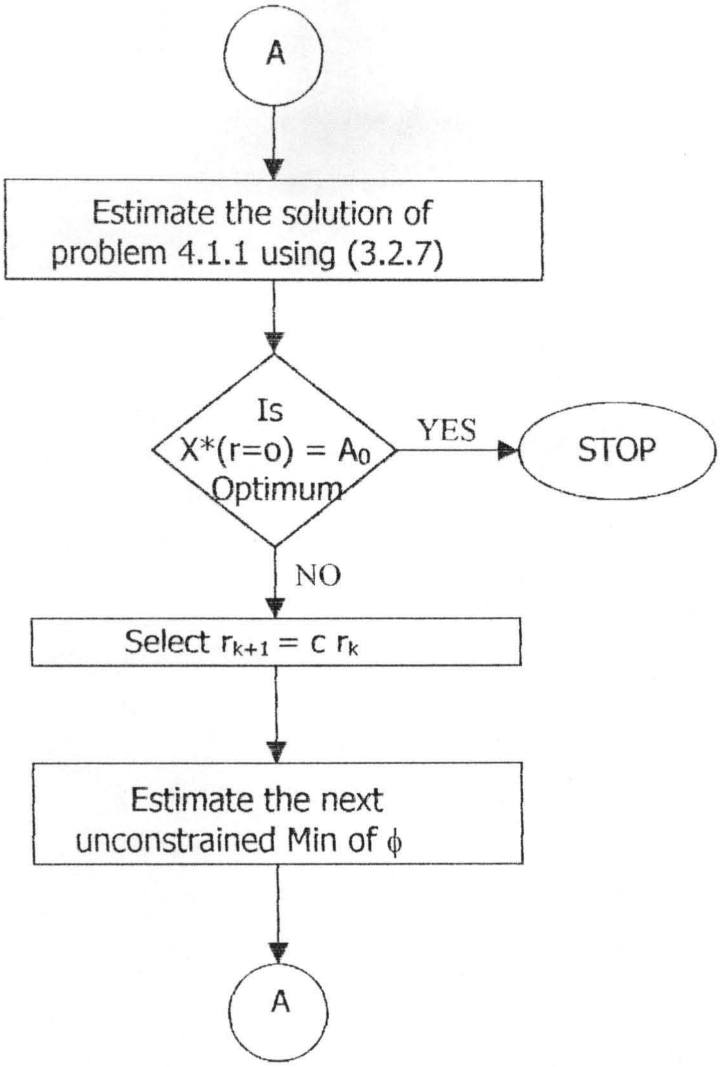
4.2: COMPUTATIONAL ALGORITHM.

The computational algorithm for extrapolation techniques is built on penalty function algorithm. The algorithm is summarized as follows:

1. Determine the unconstrained Min of ϕ for the current value of rk and use it to obtain r_{k-1} and $c< 1$ where $r_k = c\, r_{k-1}$.
2. Estimate the solution of (4.1.1) using the extrapolation formula in (3.2.7)
3. Terminate the computation if the estimated solution is acceptable. Otherwise
4. Select $r_{k+1} = c\, r_k$ where $c< 1$
5. Estimate the next unconstrained Min of the ϕ - function using the extrapolation formula in (3.5.5) and go to step 2.

These steps are shown in the form of a flowchart as in fig 4.2.1.





4.3: COMPUTATIONAL EXPERIMENTS.

Example 4.3.1

The following results are obtained during the minimization of

$$F(x) = 9 - 8x_1 - 6x_2 - 4x_3 + 2x_1^2 + 2x_2^2 + x_3^2 + 2x_1x_2 + 2x_1x_3$$

Subject to

$$x_1 + x_2 + 2x_3 \leq 3$$

and $x_i \geq 0, \quad i = 1,2,3$

using the exterior penalty function method

Value of r_i	Starting point for minimization of $\phi(X, r_i)$	Unconstrained Min of $\phi(X, r_i) = X_i^*$	$F(X_i^*) = f_i^*$
1	$\begin{Bmatrix} 0.1 \\ 0.1 \\ 0.1 \end{Bmatrix}$	$\begin{Bmatrix} 0.8884 \\ 0.7188 \\ 0.2760 \end{Bmatrix}$	0.7072
0.01	$\begin{Bmatrix} 0.8884 \\ 0.7188 \\ 0.2760 \end{Bmatrix}$	$\begin{Bmatrix} 1.3313 \\ 0.7539 \\ 0.3710 \end{Bmatrix}$	0.1564
0.0001	$\begin{Bmatrix} 1.3313 \\ 0.7539 \\ 0.3710 \end{Bmatrix}$	$\begin{Bmatrix} 1.3478 \\ 0.7720 \\ 0.4293 \end{Bmatrix}$	0.1158

Use extrapolation technique to predict the optimum solution using the following relations.

(i) $X(r) = A_0 + r A_1; \quad f(r) = a_0 + r a_1$

(ii) $X(r) = A_0 + r^{1/2} A_1, \quad f(r) = a_0 + r^{1/2} a_1$

Compare your result with the exact solution.

$$X^* = \begin{Bmatrix} 12/9 \\ 7/9 \\ 4/9 \end{Bmatrix} \qquad \text{and } f_{min} = 1/9$$

Solution:

From the result above

We have r1 = 1.0

$$X_1^* = \begin{Bmatrix} 0.8884 \\ 0.7188 \\ 0.2760 \end{Bmatrix} \qquad f_1^* = 0.7072$$

And for r2 = 0.01, c= 0.01

$$X_2^* = \begin{Bmatrix} 1.3313 \\ 0.7543 \\ 0.3710 \end{Bmatrix}, \qquad f_2^* = 0.1564$$

(i) using the relation $X(r) = A_0 + r A_1$ leads to

$$X^* \approx A_0 = \frac{X_2^* - cX_1^*}{1 - c}$$

$$= 1/0.99 \left(\begin{Bmatrix} 1.3313 \\ 0.7543 \\ 0.3710 \end{Bmatrix} - (0.01) \begin{Bmatrix} 0.8884 \\ 0.7188 \\ 0.2760 \end{Bmatrix} \right)$$

$$= \begin{Bmatrix} 1.3358 \\ 0.7543 \\ 0.3720 \end{Bmatrix}$$

Similarly the use of a linear relationship

$f^*(r) = a_0 + a_1 r$ leads to

$$f^* = \frac{f_2^* - cf_1^*}{1 - c} = 1/0.99 \left[(0.1564) - (0.01)(0.7072) \right]$$

$$= 0.1508$$

(ii) The linear relationship $X(r) = A_0 + r^{1/2} A_1$ leads to

$$X^* \approx A_0 = \frac{X_2^* - c^{1/2} X_1^*}{1 - c^{1/2}}$$

$$= 1/0.9 \left(\begin{Bmatrix} 1.3313 \\ 0.7543 \\ 0.3710 \end{Bmatrix} - 0.1 \begin{Bmatrix} 0.8884 \\ 0.7188 \\ 0.2760 \end{Bmatrix} \right)$$

$$= \begin{Bmatrix} 1.38051 \\ 0.7578 \\ 0.3816 \end{Bmatrix}$$

Similarly the use of a linear relation for $f(r)$ in terms of $r^{1/2}$ gives

$$f^* = \frac{f_2^* - c^{1/2} f_1^*}{1 - c^{1/2}}$$

$$= 1/0.9 \ \ 0.1564 - 0.1 \ (0.7072)$$

$$= 0.0952$$

By comparing the present values with the true (exact) solution in Example 4.3.1, it can be found that the linear approximation in terms $r^{1/2}$ gave a better results compared to the linear equation in r. further, both the design vector X* predicted by the linear equations in r and $r^{1/2}$ are seen to be feasible.

Example 4.3.2

Find the extrapolated values of X and f in Example 2.3.1, using the results of minimization of $\phi(X,r_6)$ and $\phi(X,r_7)$

Solution:

From the results of Example 2.3.1, we have for $r_6 = 0.01$

$$X_6^* = \begin{Bmatrix} 1.04881 \\ 0.10000 \end{Bmatrix} ; \quad f_6^* = 2.9667$$

And for $r_7 = 0.001$, $c = 0.1$

$$X_7^* = \begin{Bmatrix} 1.01569 \\ 0.03162 \end{Bmatrix} ; \quad f_7^* = 2.7615$$

If (3.2.4) is used to approximate X*(r), the extrapolated value of X* can be obtained from (3.2.7) as

$$X^* \approx A_0 = \frac{X_7^* - cX_6^*}{1 - c}$$

$$= 1/0.9\left(\begin{Bmatrix} 1.01569 \\ 0.03162 \end{Bmatrix} - 0.1\begin{Bmatrix} 1.04881 \\ 0.10000 \end{Bmatrix} \right)$$

$$= \begin{Bmatrix} 1.01201 \\ 0.02402 \end{Bmatrix}$$

Similarly the linear relationship $f^*(r) = a_0 + a_1 r$ leads to

$$f^* = \frac{F_7^* - cf_6^*}{1 - c} = \left[1/0.9 \ (2.7615) - 0.1(2.9667)\right]$$

$$= 2.7387$$

If (3.2.13) is used for approximating $X^*(r)$, the extrapolated value of X^* can be obtain from (3.2.15) as

$$X^* \approx A_0 = \frac{X_2^* - c^{1/2} X_1^*}{1 - c^{1/2}}$$

$$= 1/0.8638\left(\begin{Bmatrix}1.01569\\0.03162\end{Bmatrix} - 0.31623\begin{Bmatrix}1.04881\\0.10000\end{Bmatrix}\right)$$

$$= \begin{Bmatrix}1.00033\\-0.000004\end{Bmatrix}$$

Similarly the use of a linear relation for $f(r)$ in terms of $r^{1/2}$ gives

$$F^* \approx \frac{f_2^* - c^{1/2} f_1^*}{1 - c^{1/2}}$$

$$= \left[1/0.6838\ (2.7615) - 0.31623(2.9667)\right]$$

$$= 2.6665$$

Comparing the present values with the exact solution given in Example 2.3.1, it can be found that the extrapolation using the linear relationship

$X^*(r) = A_0 + rA_1$ agree with the theoretical solution two places, whereas the minimum for $r^{1/2}$ gave a better results compared to the linear equation in r.

<u>Example 4.3.3</u>

Minimize

$$f(X) = -15x_1 - 27x_2 - 36x_3 - 18x_4 - 12x_5$$

$$+ 30x_1^2 + 39x_2^2 + 10x_3^2 + 39x_4^2 + 30x_5^2$$

$$- 40x_1x_2 - 20x_1x_3 + 64x_1x_4 - 20x_1x_5$$

$$- 12x_2x_3 - 62x_2x_4 + 64x_2x_5 - 12x_3x_4 - 20x_3x_5 - 40x_4x_5$$

$$+ 4x_1^3 + 8x_2^3 + 10x_3^3 + 6x_4^3 + 2x_5^3$$

subject to the constraints

$16\, x_1 - 2\, x_2 - x_4 \leq 40$

$2\, x_2 - 0.4\, x_4 - 2\, x_5 \leq 2$

$7/2\, x_1 - 2\, x_3 \leq 1/4$

$2\, x_2 + 4\, x_4 + x_5 \leq 4$

$9\, x_2 + 2\, x_3 - x_4 + 2.8\, x_5 \leq 4$

$-\, 2\, x_1 + 4\, x_3 \leq 1$

$x_1 + x_2 + x_3 + x_4 + x_5 \leq 40$

$x_1 + 2\, x_2 + 3\, x_3 + 2\, x_4 + x_5 \leq 60$

$-\, x_1 - 2\, x_2 - 3\, x_3 - 4\, x_4 - 5\, x_5 \leq -5$

$-\, x_1 - x_2 - x_3 - x_4 - x_5 \leq -1$

and $-\, x_i \leq 0 \quad i = 1$ to 5

The optimum solution of this problem is given by

$$X^* = \begin{Bmatrix} 0.3 \\ 0.33347 \\ 0.4 \\ 0.42831 \\ 0.2239 \end{Bmatrix} \quad \text{and } f_{min} = -32.34868$$

Using extrapolation technique.

Solution

The starting point is taken as

$$X_1 = \begin{Bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 1 \end{Bmatrix} \quad \text{with } f(X_1) = 20.0$$

We use Newton's method of unconstrained minimization with $c = r_k/r_{k+1} = 50$.

The results are summarize in the following table.

Value of r_k	x_1	x_2	x_3	x_4	x_5	f(x)
1.0	0.1762	0.2575	0.2888	0.5698	0.4272	-26
2×10^{-2}	0.2684	0.3208	0.3743	0.4675	0.2677	-31.23
4×10^{-4}	0.2954	0.3311	0.3961	0.4352	0.2327	-32.19

Using the extrapolated technique at $\phi(X,r_2)$ and $\phi(X,r_3)$

From the result above we have $r_2 = 0.02$

$$X_2^* = \begin{Bmatrix} 0.2684 \\ 0.3208 \\ 0.3743 \\ 0.4675 \\ 0.2677 \end{Bmatrix}, \quad f_2^* = -31.23$$

And for $r_3 = 0.0004$ $\quad c = 0.02$

$$X_3^* = \begin{Bmatrix} 0.2954 \\ 0.3311 \\ 0.3961 \\ 0.4352 \\ 0.2327 \end{Bmatrix}, \quad , \quad f_3^* = -32.19$$

The linear relationship $X(r) = A_0 + rA_1$ leads to

$$X^* \approx A_0 = \frac{X_7^* - cX_6^*}{1 - c}$$

$$= 1/0.98 \left(\begin{Bmatrix} 0.2954 \\ 0.3311 \\ 0.3961 \\ 0.4352 \\ 0.2327 \end{Bmatrix} - 0.02 \begin{Bmatrix} 0.2684 \\ 0.3208 \\ 0.3743 \\ 0.4675 \\ 0.2677 \end{Bmatrix} \right)$$

$$= \begin{Bmatrix} 0.2960 \\ 0.3313 \\ 0.3965 \\ 0.4245 \\ 0.2320 \end{Bmatrix}$$

Similarly the linear relationship $f^*(r) = a_0 + a_1 r$ leads to

$$f^* \approx \frac{F_7^* - cf_6^*}{1 - c} = 1/0.98\left[(-32.19 - 0.12(-31.23)\right]$$

$$= -32.20959$$

the relationship $X(r) = A_0 + r^{1/2} A_1$ leads to

$$X^* \approx A_0 = \frac{X_2^* - c^{1/2} X_1^*}{1 - c^{1/2}}$$

$$= 1/0.8586\left(\begin{Bmatrix} 0.2954 \\ 0.3311 \\ 0.3961 \\ 0.4352 \\ 0.2327 \end{Bmatrix} - 0.1414\begin{Bmatrix} 0.2684 \\ 0.3208 \\ 0.3743 \\ 0.4675 \\ 0.2677 \end{Bmatrix}\right)$$

$$= \begin{Bmatrix} 0.2999 \\ 0.3328 \\ 0.3997 \\ 0.4299 \\ 0.2269 \end{Bmatrix}$$

similarly the use of a linear relation for $f(r)$ in terms of $r^{1/2}$ gives

$$f^* = \frac{f_2^* - c^{1/2} f_1^*}{1 - c^{1/2}} = 1/0.8586\left[(-32.19) - 0.1414(-31.23)\right]$$

$$= -32.3481$$

Looking at the approximated results and the optimum value, though the prediction using the linear relation in r and $r^{1/2}$ are feasible, the linear relationship in terms of $r^{1/2}$ give a better approximation. The f(r) in terms of $r^{1/2}$ agrees with the theoretical result to five places while that in terms of r agree to only two places.

From the above results, the convergence of the estimates to the theoretical solution can be seen. Thus the extrapolation technique accelerates the entire process of minimization by substantially reducing the effort required to minimize the successive ϕ - functions.

4.4: ANALYSIS OF ERROR DUE TO EXTRAPOLATION

From the previous sections, the power of extrapolation came to play where, it was seen that the prediction (approximation) agrees with the optimum or theoretical solution to a certain number of places. In this section, an attempt is made to examine the error(s) due to the use of extrapolation techniques.

Let define the error due to extrapolation as the difference between the true minimum X^* and the estimate (extrapolated value) $X^*(r=0) = A_0$. in terms of the above notation this is given by

$$X^* - X^*(r=0) = E \qquad 4.4.1$$

Where E is the error due to extrapolation. If the objective function is used instead of the design vector, then we define the error due to extrapolation as the difference between optimum value of the original objective function f^* and the estimate, $f^*(r=0)$ a_0.

i.e. $$f^* - f^*(r=0) = E \qquad 4.4.2$$

attempt is made to analyse the error using the results of the examples in chapter three and section 4.3. The error is summarize in the table below.

Table 4.4.1: Error due to extrapolation.

Example	Error on the Design vector	Error on the objective function
3.5.1	$\left\{ \begin{matrix} 2.5 \times 10^{-3} \\ 1.2 \times 10^{-3} \end{matrix} \right\}$	3.7×10^{-3}
3.5.1 (3-order)	$\left\{ \begin{matrix} 2.5 \times 10^{-7} \\ 1.2 \times 10^{-7} \end{matrix} \right\}$	1.0×10^{-7}
4.3.1	$\left\{ \begin{matrix} 4.7 \times 10^{-2} \\ 2.0 \times 10^{-2} \\ 6.2 \times 10^{-2} \end{matrix} \right\}$	1.5×10^{-2}
4.3.2	$\left\{ \begin{matrix} 3.3 \times 10^{-4} \\ 4 \times 10^{-5} \end{matrix} \right\}$	$2. \times 10^{-4}$
4.3.3	$\left\{ \begin{matrix} 1 \times 10^{-4} \\ 7 \times 10^{-4} \\ 3 \times 10^{-4} \\ 1.5 \times 10^{-3} \\ 2.9 \times 10^{-3} \end{matrix} \right\}$	6×10^{-4}

From Example 3.5.1, the third –order estimate agree with the optimum value to seven places and the error due to extrapolation on the objective function is 1.0×10^{-7}. This is acceptable to converge to the true minimum.

The other examples shows that the error is acceptable, although it will be minimize if 2nd or 3rd order is use in the estimate. Thus we can put it that extrapolation techniques accelerate rate of convergence to the optimum solution.

CHAPTER FIVE

PERFORMANCE OF ALGORITHM AND CONCLUSION.

5.1 INTRODUCTION:

This study deals with the use of extrapolation method in solving a general nonlinear constrained optimization. As already stated the problem of interest is to find the n- vector X that minimize the scalar function

$$f(X) = f(X_1,\ldots,X_n) \qquad 5.1.1$$

called the performance index, subject to the inequality constraints

$$g_j(x) \leq 0 \qquad j = 1,2,\ldots,m \qquad 5.1.2$$

In this chapter we report the performance of the algorithm described in chapter four, compare the result and sum up the findings so far.

5.2: PERFORMANCE OF EXTRAPOLATION ALGORITHM.

Several Computer programs are available to solve constrained programming problem.

The algorithm described in section 4.2 is coded using BASIC computer program and tested using problems of examples 4.2.1, 4.3.2, and 4.3.3. The computer program is in the Appendix. The output (results) are summarized in the table below

Problem	X*(r)	X*(r ½)	F*(r)	F*(r ½)
4.3.1	{1.3805, 0.7578, 0.3816}	{1.3358, 0.7543, 0.3720}	0.1508	0.0952
4.3.2	{1.01201, 0.02402}	{1.0000373, -0.000004}	2.7387	2.6666
4.3.3	{0.2960, 0.3313, 0.3965, 0.4345, 0.2320}	{0.2998, 0.3328, 0.3997, 0.4299, 0.2269}	-32.20959	-32.3481

From the table above, Problem 4.3.1 results agree with those obtained in section 4.3. by comparing the present value with the true optimum(solution), it can be found that the linear approximation in $r^{1/2}$ gives a better result compared to the linear equation in r.

Comparing the present values of problem 4.3.2 with the exact solution, extrapolation using the linear relationship r, agrees with the theoretical solution to two places, whereas the minimum for $r^{1/2}$ agrees to four.

Looking at the approximated results and the optimum value, though the prediction in r and $r^{1/2}$ are feasible, the linear relationship in terms of $r^{1/2}$ gives a better approximation.

Thus the extrapolation technique is a very powerful computational tool that accelerates the entire process of minimization by substituting reducing the effort required to minimize the successive ϕ - functions.

5.3: CONCLUSION

From the foregoing study on the extrapolation method in constrained optimization, it has been proved that, under certain conditions, the difference between the true minimum X* and the estimate X*(r=0) = A_0 will be of the order r_1^k. Thus as $r_1 \to 0$ $A_0 \to X^*$. moreover, if $r_1 < 1$, the estimates of X^* obtained by using K minima, will be better than those using (k – 1) minima and so on. Hence, as more minima are achieved, the estimates can be used as the starting point for the (k + 1) st minimization of the ϕ - function. This accelerates the entire process by substantially reducing the effort needed to minimize the successive ϕ - functions. Sometimes, the extrapolations made with the help of (k–1) st order polynomial in $r^{1/2}$ rather than in r are found to be better. Furthermore, the extrapolated value a_0 can be used to provide an additional convergence criterion for terminating the interior penalty function method.

The analysis of error due to extrapolation was also examined and although there was some slight deviation from the true optimum value, the use of extrapolation formulas accelerates convergence.

From this study, one can conclude that the idea of extrapolation is to use the solution of previous unconstrained minimization or the barrier sob problem to fit to a polynomial to the barrier trajectory, and then use this polynomial to predict the solution of the constrained minimization problem. This extrapolation method accelerates the entire process by substantially reducing the effort needed to minimize the successive ϕ - function. Thus extrapolation method helps interior penalty function method.

5.4 RECOMMENDATION

This study uses a first order polynomial in r to approximate the design vector and objective function, while Rao[(10)] observed that extrapolation with the help of even quadratic and cubic equations in r, generally yield good estimates. There is therefore need to examine the use of second or third order polynomial to ascertain rate of convergence.

It was stated in section (2.5c) that one of the weaknesses of the interior penalty function method is related to the ill- conditioned nature of its Hessian matrix, but if Newton method of unconstrained minimization is used, and initial guesses are obtained by extrapolation, then the potential deficiencies of interior penalty method will go away.

The extrapolation method is a very powerful tool that can be used to estimate the minimum of the original constrained problem. Thus instead of the effort needed to minimize the successive ϕ - function as in the interior penalty function method, extrapolation can be applied.

The algorithm described in the previous chapters cannot be directly applied to solve problems with mixed equality and inequality constraints. Hence there is need to examine extrapolation method with mixed equality and inequality constraints in further research effort.

Furthermore, the use of extrapolation method in solving constrained parameter optimization problem remains one of intrinsic interest.

Fiacco and McCormick(3) demonstrated that there is a trajectory of optimal points varying continuously as a function of r. thus it is worth considering trajectory analysis and extrapolation method in barrier function method.

REFERENCES:

1. Bertsekas, D. P. Constrained optimization and Lagrange Multiplier Method. Academic Press, London, 1982.

2. Betts, J. T. An Improved Penalty Function Method for solving constrained Parameter Optimization problems. Journal of optimization and Application. 16(1/2) 1975, 1 – 23.

3. Fiacco, A. V. and McCormick, G. P. Nonlinear programming, Sequential Unconstrained Minimization Techniques. SIAM, Philadelphia, 1970

4. Foulds, L. R. Optimization Techniques. (An introduction). Springer-Verlag, NewYork 1981

5. Mitra, G. Theory and Application of Mathematical programming. Academic Press, London, 1976

6. Nash, G. N. and Ariela, S. Why Extrapolation Helps Barrier Methods, Operation Research and Engineering Dept. George Mason University(sent for publication).

7. Osborne, M. R. and Ryan, D.M. On Penalty Function Methods for nonlinear programming problem. Journal of Mathematical Analysis and Application; 31 (1970), 559 - 578

8. Peressini, A. L. et al. The Mathematics of Nonlinear programming. Springer – Verlag. NewYork, 1998.

9. Polak, E. Methods in Optimization. Academic Press, London, 1976

10 Rao, S. S. Optimization theory and Application, Wiley Eastern, New Delhi, 1978

11. Walsh, G. R. Methods of Optimization. John Wiley and sons, NewYork,1975

APPENDIX

```
REM =====================================================================
DECLARE SUB prob1 ()
DECLARE SUB prob2 ()
DECLARE SUB prob3 ()
DECLARE SUB bgrnd ()
DECLARE SUB scrn ()
DIM X(5), Xx(5), X1(5), X2(5), X3(5), X4(5), X5(5), X6(5), X7(5)
GOTO menu
'======================================== SIGNING OFF
q:
scrn
LOCATE 12, 10: PRINT "Program terminated by the user, Press any
key..."
END

REM ==================================================== Program menu
menu:
scrn
bgrnd
COLOR 2
LOCATE 4, 27: PRINT " EXTRAPOLATION TECHNIQUE PROGRAM"
LOCATE 6, 27: PRINT "(Using Interior Penalty Function)"
LOCATE 8, 37: PRINT "Program Menu"
COLOR 10
LOCATE 12, 30: PRINT "1"
COLOR 15
LOCATE 12, 31: PRINT ".  Problem 1"
COLOR 10
LOCATE 14, 30: PRINT "2"
COLOR 15
LOCATE 14, 31: PRINT ".  Problem 2"
COLOR 10
LOCATE 16, 30: PRINT "3"
COLOR 15
LOCATE 16, 31: PRINT ".  Problem 3"
COLOR 10
LOCATE 18, 30: PRINT "4"
COLOR 15
LOCATE 18, 31: PRINT ".  Quit"
COLOR 2
COLOR 11
LOCATE 27, 20: PRINT "Developed By: (                           MTECH/
/  )"
COLOR 14
LOCATE 20, 28: INPUT "Enter your choice..."; no%
COLOR 15
IF no% < 1 OR no% > 4 THEN
        LOCATE 21, 25: PRINT "Invalid choice [Enter 1, 2, 3 or 4]."
        FOR Y = 1 TO 1000 STEP .01: NEXT: GOTO menu
END IF
IF no% = 1 THEN prob1
IF no% = 2 THEN prob2
IF no% = 3 THEN prob3
IF no% = 4 THEN GOTO q
GOTO menu
END
REM ------------------------- End of Program ----------------------
```

```
SUB probl
REM PROGRAM TO PREDICT THE OPTIMUM SOLUTION USING EXTRAPOLATION
TECHNIQUE
REM (INTERIOR PENALTY METHOD)
DIM X(5), Xx(5), X1(5), CX1(5), X2(5)
REM Xx = X*,  X2 = X2*, X1 = X1*, F1 = F1*, F2 = F2*, fx = f*
LET DATA1$ = "PROB1.DAT"
LET RES1$ = "RESUL1.OUT"
CLS
OPEN DATA1$ FOR INPUT AS #1
OPEN RES1$ FOR OUTPUT AS #2
IF EOF(1) THEN GOTO 5
        INPUT #1, r1
        FOR i% = 1 TO 3
                INPUT #1, X1(i%)
        NEXT i%
        INPUT #1, f1
        INPUT #1, r2
        FOR i% = 1 TO 3
                INPUT #1, X2(i%)
        NEXT i%
        INPUT #1, f2
5 CLOSE #1
LET Fmin = 1 / 9: C = .01
REM USING THE RELATION X(r) = Ao + rA1
REM X* ~ Ao = (X2* - CX1*)/(1-C)
FOR i% = 1 TO 3
        CX1(i%) = -C * X1(i%)
NEXT i%
FOR i% = 1 TO 3
        Xx(i%) = CX1(i%) + X2(i%)
NEXT i%
FOR i% = 1 TO 3
        Xx(i%) = Xx(i%) / (1 - C)
NEXT i%
        PRINT #2, "RESULTS FOR PROBLEM 1"
        PRINT #2, "---------------------"
        PRINT #2, "(i)  X(r) = Ao + rA1; f(r) = ao + ra1"
        PRINT #2, "X*"
FOR i% = 1 TO 3
        PRINT USING "##.####"; Xx(i%)
        PRINT #2, USING "##.####"; Xx(i%)
NEXT i%
        PRINT #2, "f*"

REM --- FOR LINEAR RELATIONSHIP
REM f*(r) = ao + a1r  leads to
REM f* = (f2* - Cf1*)/(1-C)
        fx = (f2 - C * f1) / (1 - C)
PRINT USING "##.####"; fx
PRINT #2, USING "##.####"; fx: PRINT #2,
        PRINT #2, "(ii)  X(r) = Ao + r(1/2)A1; f(r) = ao + r(1/2)a1"
        PRINT #2, "X*"
REM  ---------------------------------  (ii)
REM FROM X(r)=Ao + r(1/2) A1
REM X* ~ Ao = (X2* - C(1/2) X1*)/(1-C(1/2))
FOR i% = 1 TO 3
        CX1(i%) = -C ^ (.5) * X1(i%)
NEXT i%
FOR i% = 1 TO 3
        Xx(i%) = CX1(i%) + X2(i%)
```

```
NEXT i%
FOR i% = 1 TO 3
        Xx(i%) = Xx(i%) / (1 - C ^ (.5))
NEXT i%
FOR i% = 1 TO 3
        PRINT USING "##.####"; Xx(i%)
        PRINT #2, USING "##.####"; Xx(i%)
NEXT i%
        PRINT #2, "f*"
REM --- FOR LINEAR RELATIONSHIP
REM f*(r) in terms of r(1/2)
REM f* = (f2* - C1/2f1*)/(1-C1/2)
        fx = (f2 - C ^ .5 * f1) / (1 - C ^ .5)
PRINT USING "##.####"; fx
PRINT #2, USING "##.####"; fx
CLOSE #2
LOCATE 10, 20: PRINT "COMPUTATION COMPLETED SUCCESSFULLY  !!!"
   LOCATE 12, 20: PRINT "Check file 'RESUL1.OUT' for details."
   FOR Y = 1 TO 4000 STEP .01: NEXT Y
END SUB

SUB prob2
REM PROGRAM TO PREDICT THE OPTIMUM SOLUTION USING EXTRAPOLATION
TECHNIQUE
REM (INTERIOR PENALTY METHOD)
REM Xx = X*,  X6 = X6*, X7 = X7*, F6 = F6*, F7 = F7*, fx = f*
DIM X(5), Xx(5), CX6(5), X6(5), X7(5)
LET DATA2$ = "PROB2.DAT"
LET RES2$ = "RESUL2.OUT"
CLS
OPEN DATA2$ FOR INPUT AS #1
OPEN RES2$ FOR OUTPUT AS #2
IF EOF(1) THEN GOTO 15
        INPUT #1, r6
        FOR i% = 1 TO 2
                INPUT #1, X6(i%)
        NEXT i%
        INPUT #1, f6
        INPUT #1, r7
        FOR i% = 1 TO 2
                INPUT #1, X7(i%)
        NEXT i%
        INPUT #1, f7
15 CLOSE #1
LET Fmin = 1 / 9: C = .1
REM  ***********************USING THE RELATION X(r) = Ao + rA1
REM X* ~ Ao = (X7* - CX6*)/(1-C)
FOR i% = 1 TO 2
        CX6(i%) = -C * X6(i%)
NEXT i%
FOR i% = 1 TO 2
        Xx(i%) = CX6(i%) + X7(i%)
NEXT i%
FOR i% = 1 TO 2
        Xx(i%) = Xx(i%) / (1 - C)
NEXT i%
        PRINT #2, "RESULTS FOR PROBLEM 2"
        PRINT #2, "---------------------"
        PRINT #2, "(i)  X(r) = Ao + rA1; f(r) = ao + a1r"
        PRINT #2, "X*"
FOR i% = 1 TO 2
```

```
        PRINT USING "##.#####"; Xx(i%)
        PRINT #2, USING "##.#####"; Xx(i%)
NEXT i%
        PRINT #2, "f*"

REM --- FOR LINEAR RELATIONSHIP
REM f*(r) = ao + a1r  leads to
REM f* = (f2* - Cf1*)/(1-C)
        fx = (f7 - C * f6) / (1 - C)
PRINT USING "##.####"; fx
PRINT #2, USING "##.####"; fx: PRINT #2,
        PRINT #2, "(ii)  X(r) = Ao + r(1/2)A1; f(r) = ao +a1r(1/2)"
        PRINT #2, "X*"
REM  ------------------------------------ (ii)
REM FROM X(r)=Ao + r(1/2) A1
REM X* ~ Ao = (X7* - C(1/2) X6*)/(1-C(1/2))
FOR i% = 1 TO 2
        CX6(i%) = -C ^ (.5) * X6(i%)
NEXT i%
FOR i% = 1 TO 2
        Xx(i%) = CX6(i%) + X7(i%)
NEXT i%
FOR i% = 1 TO 2
        Xx(i%) = Xx(i%) / (1 - C ^ (.5))
NEXT i%
FOR i% = 1 TO 2
        PRINT USING "##.######"; Xx(i%)
        PRINT #2, USING "##.######"; Xx(i%)
NEXT i%
        PRINT #2, "f*"
REM --- FOR LINEAR RELATIONSHIP
REM f*(r) in terms of r(1/2)
REM f* = (f2* - C6/2f6*)/(1-C1/2)
        fx = (f7 - C ^ .5 * f6) / (1 - C ^ .5)
PRINT USING "##.####"; fx
PRINT #2, USING "##.####"; fx
CLOSE #2
LOCATE 10, 20: PRINT "COMPUTATION COMPLETED SUCCESSFULLY  !!!"
   LOCATE 12, 20: PRINT "Check file 'RESUL2.OUT' for details."
   FOR Y = 1 TO 4000 STEP .01: NEXT Y
END SUB

SUB prob3
REM PROGRAM TO PREDICT THE OPTIMUM SOLUTION USING EXTRAPOLATION
TECHNIQUE
REM (INTERIOR PENALTY METHOD)
REM Xx = X*,  X2 = X2*, X1 = X1*, X3=X3*
REM F1 = F1*, F2 = F2*, F3= F3*, fx = f*
DIM X(5), Xx(5), X1(5), CX2(5), X2(5), X3(5)
LET DATA3$ = "PROB3.DAT"
LET RES3$ = "RESUL3.OUT"
CLS
OPEN DATA3$ FOR INPUT AS #1
OPEN RES3$ FOR OUTPUT AS #2
IF EOF(1) THEN GOTO 25
        FOR i% = 1 TO 5
                INPUT #1, X1(i%)
        NEXT i%
        INPUT #1, f1

        INPUT #1, r2
```

```
        FOR i% = 1 TO 5
                INPUT #1, X2(i%)
        NEXT i%
        INPUT #1, f2

        INPUT #1, r3
        FOR i% = 1 TO 5
                INPUT #1, X3(i%)
        NEXT i%
        INPUT #1, f3
25 CLOSE #1
LET Fmin = -32.34868: C = .02
REM ***********************USING THE RELATION X(r) = Ao + rA1
REM X* ~ Ao = (X3* - CX2*)/(1-C)
FOR i% = 1 TO 5
        CX2(i%) = -C * X2(i%)
NEXT i%
FOR i% = 1 TO 5
        Xx(i%) = CX2(i%) + X3(i%)
NEXT i%
FOR i% = 1 TO 5
        Xx(i%) = Xx(i%) / (1 - C)
NEXT i%
        PRINT #2, "RESULTS FOR PROBLEM 3"
        PRINT #2, "---------------------"
        PRINT #2, "(i)  The starting point is taken as:"
        PRINT #2, "X1"
FOR i% = 1 TO 5
        PRINT USING "#"; X1(i%)
        PRINT #2, USING "#"; X1(i%)
NEXT i%
        PRINT #2, : PRINT #2, "f(x1)"
        PRINT USING "###.#"; f1
        PRINT #2, USING "###.#"; f1: PRINT #2,

        PRINT #2, "(ii)  X(r) = Ao + rA1; f(r) = ao + a1r"
        PRINT #2, "X*"
FOR i% = 1 TO 5
        PRINT USING "##.####"; Xx(i%)
        PRINT #2, USING "##.####"; Xx(i%)
NEXT i%
        PRINT #2, "f*"

REM --- FOR LINEAR RELATIONSHIP
REM f*(r) = ao + a1r  leads to
REM f* = (f3* - Cf2*)/(1-C)
        fx = (f3 - C * f2) / (1 - C)
PRINT USING "##.#####"; fx
PRINT #2, USING "##.#####"; fx: PRINT #2,
        PRINT #2, "(iii)  X(r) = Ao + r(1/2)A1; f(r) = ao + r(1/2)a1"
        PRINT #2, "X*"
REM  ---------------------------------  (ii)
REM FROM X(r)=Ao + r(1/2) A1
REM X* ~ Ao = (X3* - C(1/2) X2*)/(1-C(1/2))
FOR i% = 1 TO 5
        CX2(i%) = -C ^ (.5) * X2(i%)
NEXT i%
FOR i% = 1 TO 5
        Xx(i%) = CX2(i%) + X3(i%)
NEXT i%
FOR i% = 1 TO 5
```

```
        Xx(i%) = Xx(i%) / (1 - C ^ (.5))
NEXT i%
FOR i% = 1 TO 5
        PRINT USING "##.####"; Xx(i%)
        PRINT #2, USING "##.####"; Xx(i%)
NEXT i%
        PRINT #2, "f*"
REM --- FOR LINEAR RELATIONSHIP
REM f*(r) in terms of r(1/2)
REM f* = (f3* - C1/2f2*)/(1-C(1/2))
        fx = (f3 - C ^ .5 * f2) / (1 - C ^ .5)
PRINT USING "##.####"; fx
PRINT #2, USING "##.####"; fx
CLOSE #2
LOCATE 10, 20: PRINT "COMPUTATION COMPLETED SUCCESSFULLY  !!!"
   LOCATE 12, 20: PRINT "Check file 'RESUL3.OUT' for details."
   FOR Y = 1 TO 4000 STEP .01: NEXT Y
END SUB
```

RESULTS FOR PROBLEM 1

(i) X(r) = Ao + rA1; f(r) = ao + ra1

X*

1.3358

0.7543

0.3720

f*

0.1508

(ii) X(r) = Ao + r(1/2)A1; f(r) = ao + r(1/2)a1

X*

1.3805

0.7578

0.3816

f*

0.0952

RESULTS FOR PROBLEM 2

(i) X(r) = Ao + rA1; f(r) = ao + ra1

X*

1.01201

0.02402

f*

2.7387

(ii) X(r) = Ao + r(1/2)A1; f(r) = ao + r(1/2)a1

X*

1.000373

-0.000004

f*

2.6666

RESULTS FOR PROBLEM 3

(i) The starting point is taken as:

X1
0
0
0
0
1

f(x1)
20.0

(ii) X(r) = Ao + rA1; f(r) = ao + a1r

X*
0.2960
0.3313
0.3965
0.4345
0.2320

f*
-32.20959

(iii) X(r) = Ao + r(1/2)A1; f(r) = ao + r(1/2)a1

X*
0.2998
0.3328
0.3997
0.4299
0.2269

f*
-32.3481