

OPTIMIZATION OF MANAGEMENT SYSTEMS
BY THE CONJUGATE GRADIENT METHODS

by 680

BALAKRISHNAN RADHAKRISHNAN NAIR

B.E., UNIVERSITY OF MADRAS

MADRAS, INDIA, 1959

A MASTER'S REPORT

submitted in partial fulfillment of the
requirements for the degree

MASTER OF SCIENCE

Department of Industrial Engineering

KANSAS STATE UNIVERSITY

Manhattan, Kansas

1969

Approved by


Major Professor

LD
2668
R4
1969
N33

TABLE OF CONTENTS

1. INTRODUCTION	1
2. THE CONJUGATE GRADIENT METHOD OF FLETCHER AND REEVES	4
3. THE CONJUGATE DIRECTION METHOD OF POWELL	9
4. APPLICATIONS	14
An Inventory model with one state variable	14
a. Solution by Fletcher and Reeves' method	16
b. Solution by Powell's method	25
c. Comparison	26
An Inventory and Advertising model with two state variables	29
a. Solution by Fletcher and Reeves' method	32
b. Solution by Powell's method	33
c. Comparison	41
5. DISCUSSION	43
ACKNOWLEDGEMENT	45
REFERENCES	46
APPENDIX A: One dimensional search techniques	48
APPENDIX B: Flowcharts	52
APPENDIX C: Computer programs and results	60

1. INTRODUCTION

The optimization of management systems frequently involves the maximization or minimization of non-linear systems of equations. While efficient algorithms have been developed for handling linear problems, no general algorithms exist for solving non-linear optimization problems. Various methods like the calculus of variations, the maximum principle and dynamic programming have been developed in recent years. However, only dynamic programming has been applied extensively to management problems [1]. The main limitation with this technique is that, owing to the dimensionality difficulties, it cannot be applied to problems with a large number of state variables. While several techniques have been devised to circumvent this limitation, these essentially involve trading computer memory for computing time and are not as efficient as the original dynamic programming algorithm.

The difficulties in obtaining analytical solutions to non-linear optimization problems have led to the development of a large number of direct methods of optimization. Many of them, like Rosenbrock's method [2], the pattern search method of Hooke and Jeeves [3,4], and the simplex method of Nelder and Mead [5], rely on an ad hoc rather than a theoretical approach to the problem [15]. Gradient methods, on the other hand, are classical in origin and based on the fact that if we move in the direction of the gradient of $f(x)$, we will move in the direction of the maximum rate of increase in $f(x)$ [6,7,13]. An iterative procedure is developed where a step is taken along the gradient and the new values of the function and gradient computed. The process is repeated until the

gradient assumes a sufficiently small value, showing that a stationary point has been reached. It has been well established [7] that such a method converges too slowly for practical applications. Hestenes and Stiefel [8] devised an ingenious scheme to accelerate the convergence of the gradient method. They used the accumulated knowledge of the behaviour of the function to generate new directions of search at the end of each iteration. This technique, called the conjugate-gradient method, is used by Fletcher and Reeves [9] to find the unconstrained minimum or maximum of a function of several variables.

All gradient methods require the gradient vector, or the vector of first partial derivatives of the function, to be evaluated. However, it is frequently the case that it is laborious or practically impossible to calculate these derivatives and there is a definite need for optimization techniques which do not require them. Powell [10] devised a method of finding the unconstrained minimum or maximum of a function, by searching along linearly independent directions, called conjugate directions, generated during each iteration cycle. The method does not require the evaluation of derivatives.

Both the methods of Fletcher and Reeves and Powell have the advantage of being able to handle problems with a fairly large number of state variables. They have the property of quadratic convergence, that is, if a quadratic function of n variables is being optimized, the optimum is reached in not more than n iterations. Both the methods can be easily programmed. The disadvantages are: (1) it cannot recognize constraints on the variables and (2) it may converge to a local extremum. The latter disadvantage can be partly overcome by using different starting values.

A comparison is made of the results obtained with each technique. Some conclusions are also drawn regarding the relative merits and demerits of the two methods.

2. THE CONJUGATE GRADIENT METHOD OF FLETCHER AND REEVES

Fletcher and Reeves [9] devised a method of locating the unconstrained minimum of a function of n variables whose value $f(x)$ and the gradient vector $g(x)$ can be calculated at any point x .

Assume that in the neighborhood of the required minimum h , the function may be expanded in the form

$$f(x) = f(h) + \frac{1}{2} (x - h)^T A (x - h) + \text{higher terms} \quad (1)$$

where A , the matrix of second-order partial derivatives, is symmetric and positive-definite.

Virtually all iterative minimization procedures locate h as the limit of a sequence x_0, x_1, x_2, \dots where x_0 is an initial approximation to the position of the minimum. Also, for each $i \geq 0$, x_{i+1} is the position of the minimum of $f(x)$ with respect to variations along the line through x_i in some specified direction ξ_i . Thus, for example, the method of steepest descent uses the direction of the negative gradient of $f(x)$ at x_i , and the method of changing one variable at a time uses cyclically the directions of the n coordinate axes.

In the method of Fletcher and Reeves, the successive directions of search are A -conjugate, satisfying the condition

$$\xi_i^T A \xi_j = 0 \quad \text{for } i \neq j \quad (2)$$

Setting $g(x_i) = g_i$ for each i , the step from x_i to x_{i+1} is determined by the relation

$$g_{i+1}^T \xi_i = 0 \quad (3)$$

$$x_{i+1} = x_i + \lambda_i \xi_i \quad (4)$$

for some scalar parameter λ_i .

The advantage of using A-conjugate directions is that the convergence is quadratic, that is, if a quadratic function of n variables is being minimized, the minimum is reached in not more than n iterations. This can be proved as follows:

Consider the minimization by successive linear searches of the quadratic function

$$f(x) = f(h) + \frac{1}{2} (x - h)^T A (x - h) \quad (5)$$

for which the gradient is

$$g(x) = A (x - h) \quad (6)$$

By repeated use of equation (4), it is seen that

$$x_n = x_{j+1} + \sum_{i=j+1}^{n-1} \lambda_i \xi_i \quad (7)$$

for any j in $0 \leq j \leq n-1$. It then follows from equation (6), that

$$g_n = g_{j+1} + \sum_{i=j+1}^{n-1} \lambda_i A \xi_i \quad (8)$$

and therefore, using equation (3), that

$$g_n^T \xi_j = \sum_{i=j+1}^{n-1} \lambda_i \xi_i^T A \xi_j \quad (9)$$

If the vectors $\xi_0, \xi_1, \xi_2, \dots, \xi_{n-1}$ are A-conjugate, satisfying equation (2), then

$$g_n^T \xi_j = 0 \quad (10)$$

As the conjugate directions $\xi_0, \xi_1, \dots, \xi_{n-1}$ are linearly independent and form a basis,

$$g_n = 0$$

and hence, by equation (6),

$$x_n = h,$$

which is the required minimum.

Thus the minimum is located in the n^{th} iteration, or earlier if any particular λ_i equals zero, if the successive directions of search are A-conjugate.

As $f(x)$ and $g(x)$ are defined numerically, A is not explicitly available. The set of A-conjugate directions is generated in Fletcher and Reeves' method by a different procedure suggested by Hestenes and Stiefel [8]. In this method, successive conjugate directions, satisfying equation (2), are generated such that ξ_{i+1} is a linear combination of $-g_{i+1}$ and $\xi_0, \xi_1, \dots, \xi_i$, as follows:

$$\xi_{i+1} = -g_{i+1} + \lambda_i \xi_i \quad (11)$$

where

$$\lambda_i = \frac{g_{i+1}^2}{g_i^2} \quad (12)$$

A detailed proof that this method yields successive A-conjugate directions is given by Beckman [11].

Using this method of generating conjugate directions, Fletcher and Reeves established the following minimization algorithm:

$$\begin{aligned}
x_0 &= \text{arbitrary} \\
g_0 &= g(x_0); \quad \xi_0 = -g_0 \\
x_{i+1} &= \text{position of minimum of } f(x) \text{ on the line through } x_i \text{ in} \\
&\quad \text{the direction } \xi_i \tag{13} \\
g_{i+1} &= g(x_{i+1}) \\
\lambda_i &= g_{i+1}^2 / g_i^2 \\
\xi_{i+1} &= -g_{i+1} + \lambda_i \xi_i
\end{aligned}$$

This procedure is guaranteed, apart from rounding errors, to locate the minimum of any quadratic function of n variables in at most n iterations. For functions which are not quadratic, the process is iterative rather than n -step, and a test for convergence is required.

During early trials with this procedure, it was noticed that errors in the computation of the conjugate directions could accumulate and result in successive directions being nearly parallel [9]. This could result in slow convergence. Fletcher and Reeves modified the basic procedure so that periodically, all accumulated information about ξ_i was discarded, and the search restarted from the current x , in the direction of steepest descent, $-g_i$. The process remains quadratically convergent so long as such restarts are not more frequent than every n iterations. The modified procedure thus consists of $(n+1)$ iterations.

The linear search problem in the iterative procedure (13) is basically a single variable search along each direction ξ_i . As in any practical application of the algorithm, the time spent to evaluate the function and the gradient may take up most of the computation time, it is preferable

to use a search procedure which reduces the number of these evaluations.

Fletcher and Reeves used a linear search procedure devised by Davidon [12] to find the minimum of the function along each direction. The procedure is described in detail in Appendix A.

The selection of a suitable convergence criterion is very important where the functions being minimized are not quadratic. If any g_i^2 vanishes, the iterations must stop. This is because it is the formal requirement for x_i to be at a minimum, and also to avoid division by zero. However, this condition is unlikely to be realized in practice, due to rounding off errors. A less stringent criterion would be to compare the value of g_i^2 with some tolerable assigned value. A second criterion is to compare the function values at the end of each cycle of $(n+1)$ iterations and to terminate the search if the improvement is less than some small value. In some cases, it might be sufficient to continue the iterations until the change in each argument x_i is less than some assigned value. All these three criteria have been used in solving the problems in this report.

The minimization procedure (13) can be used to handle maximization problems by changing $-g_{i+1}$ into g_{i+1} and modifying the linear search procedure into one of maximization. Another method is to use the same procedure as for minimization and to consider the optimization problem as one of minimizing $-f(x)$. The latter method has been used to solve the inventory and advertising problem in this report.

3. THE CONJUGATE DIRECTION METHOD OF POWELL

There frequently arise optimization problems where it is difficult or practically impossible to calculate derivatives of the function. Powell [10] devised a method of searching along conjugate directions generated without the use of derivatives or the gradient vector. This method, like that of Fletcher and Reeves discussed earlier, has the advantage of quadratic convergence. It can be used for finding only the unconstrained minimum or maximum of a function.

The principle used in generating the conjugate directions is based on the fact that if x_0 is the minimum in a space containing the direction ξ , and x_1 is also the minimum in such a space, then the direction $(x_1 - x_0)$ is conjugate to ξ . The proof is as follows:

Consider the minimization of a quadratic function of n variables

$$f(x) = \frac{1}{2}x^T Ax + Bx + C \quad (1)$$

If x_0 is the minimum along direction ξ ,

$$\frac{\partial}{\partial \lambda} \{f(x_0 + \lambda\xi)\} = 0 \quad \text{at } \lambda = 0. \quad (2)$$

Hence, substituting in equation (1), it is seen that

$$2\lambda\xi^T A\xi + 2Ax_0 + b\xi = 0, \quad \lambda = 0 \quad (3)$$

Also,
$$2\lambda\xi^T A\xi + 2Ax_1 + b\xi = 0, \quad \lambda = 0 \quad (4)$$

Hence
$$\xi^T A (x_1 - x_0) = 0 \quad (5)$$

which is the condition for conjugacy.

The convergence of a quadratic function of n variables to a minimum has already been proved to take place in n iterations, if the successive directions of search are A -conjugate.

Each iteration of Powell's procedure commences with a search down n linearly independent directions $\xi_1, \xi_2, \dots, \xi_n$, starting from the best known approximation to the minimum, x_0 . These directions are chosen to be the coordinate directions initially, so that the start of the first iteration is identical to an iteration of the method of changing one variable at a time. In subsequent iterations, the method is modified to generate conjugate directions, by making each iteration define a new direction, ξ , and choosing the linearly independent directions for the next iteration to be $\xi_2, \xi_3, \dots, \xi_n, \xi$. An iteration of the basic procedure is hence as follows:

- (i) For $r = 1, 2, \dots, n$ calculate λ_r so that $f(x_{r-1} + \lambda_r \xi_r)$ is a minimum, and define $x_r = x_{r-1} + \lambda_r \xi_r$.
- (ii) For $r = 1, 2, \dots, n-1$, replace ξ_r by ξ_{r+1}
- (iii) Replace ξ_n by $(x_n - x_0)$
- (iv) Choose λ so that $f\{x_n + \lambda(x_n - x_0)\}$ is a minimum and replace x_0 by $x_0 + \lambda(x_n - x_0)$

Powell found it necessary to modify the basic procedure in order to ensure that the rate of convergence to the minimum was satisfactory, even when the initial approximation was very poor. The reason was that on some occasions, the basic procedure may choose nearly dependent directions. The resultant directions may not hence span the full parameter space. Considering an instance, if say $\lambda_1 = 0$, then the modified procedure allows a direction other than ξ_1 to be discarded, so that the new set of directions will always be linearly independent. The modified procedure also allows for the search to be continued along the old set of directions in cases where it is unwise to replace any of them.

The procedure recommended by Powell is as follows:

- (i) For $r = 1, 2, \dots, n$ calculate λ_r so that $f(x_{r-1} + \lambda_r \xi_r)$ is a minimum, and define $x_r = x_{r-1} + \lambda_r \xi_r$.
- (ii) Find the integer m , $1 \leq m \leq n$, so that $\{f(x_{m-1}) - f(x_m)\}$ is a maximum, and define $\Delta = f(x_{m-1}) - f(x_m)$.
- (iii) Calculate $f_3 = f(2x_n - x_0)$, and define $f_1 = f(x_0)$ and $f_2 = f(x_n)$.
- (iv) If either $f_3 \geq f_1$ and/or

$$(f_1 - 2f_2 + f_3) \cdot (f_1 - f_2 - \Delta)^2 \geq \frac{1}{2} \Delta (f_1 - f_3)^2 \quad (6)$$
use the old directions $\xi_1, \xi_2, \dots, \xi_n$ for the next iteration and use x_n for the next x_0 , otherwise
- (v) Define $\xi = (x_n - x_0)$ and calculate λ so that $f(x_n + \lambda \xi)$ is a minimum. Use $\xi_1, \xi_2, \dots, \xi_{m-1}, \xi_{m+1}, \dots, \xi_n, \xi$ as the directions and $x_n + \lambda \xi$ as the starting point for the next iteration.

The above modification is based on a very useful result. It is that, if the directions $\xi_1, \xi_2, \dots, \xi_n$ are scaled so that

$$\xi_i^T \xi_i = 1, \quad i = 1, 2, \dots, n$$

the determinant of the matrix whose columns are the vectors ξ_i takes its maximum value if and only if the vectors are mutually conjugate. The proof of this result is given by Powell [10].

This result indicates that $\xi_1, \xi_2, \dots, \xi_n$ should be chosen to make the corresponding determinant as large as possible. The criterion is applied by using the new direction, ξ , defined by an iteration, if it

can cause the determinant to increase, and by rejecting the direction thereplacement of which causes the new determinant to be the largest. Powell has proved that the direction which should be discarded, if any, is ξ_m , $1 \leq m \leq n$, where m is such that $\{f(x_{m-1}) - f(x_m)\}$ is a maximum. The three function values f_1 , f_2 and f_3 are used to predict the stationary value of the function along the new direction and to develop a criterion for the acceptance or rejection of the new direction $(x_n - x_0)$ for the next iteration. The reader is referred to Powell [10] for a detailed derivation of the criterion.

As the Powell's procedure is meant primarily for optimizing functions whose derivatives are not calculated, a linear search procedure for determining λ_r , not requiring the evaluation of derivatives, would be more suitable. The Fibonacci one dimensional search procedure has been chosen for this purpose, to solve the problems in this report. A detailed explanation of the method is given in Appendix A.

One convergence criterion used to terminate the search procedure is by comparing the function values at the end of every cycle and terminating if the improvement is less than an assigned value. This is the only criterion that has been used with this method for solving the problems in this report. However, other criteria could be devised like change in value of the function argument, depending on the requirements of the applications.

The minimization procedure (6) can easily be modified to solve maximization problems by making the following changes:

Modify the Fibonacci search procedure so as to maximize instead of minimize, in steps (i) and (v).

Find integer m , $1 \leq m \leq n$, so that $\{f(x_m) - f(x_{m-1})\}$ is a maximum, and define $\Delta = f(x_m) - f(x_{m-1})$, in step (ii).

Reverse the inequality signs in step (iv).

This modification has been used to solve the Inventory problem in the next section of this report.

4. APPLICATIONS

Two models representing production and inventory situations were considered.

The first model was a simple production planning situation involving one state variable. The second model was more complex and involved two state variables. Solutions were obtained by applying the Fletcher and Reeves' method and Powell's method for both models. A comparison of the two methods was made in each case.

All the problems were solved on an IBM 360 computer using computer programs coded in Fortran.

AN INVENTORY MODEL WITH ONE STATE VARIABLE

This model has been solved by Lee and Shaikh [1], using the functional gradient technique. Consider the case of a manufacturing facility where the rate of sales $Q(t)$ is known with certainty. The rate of change of inventory level $I(t)$ is given by

$$\frac{dI(t)}{dt} = P(t) - Q(t) \quad (1)$$

where $P(t)$ = rate of production at time t .

The problem is to minimize the cost functional:

$$C_T = \int_0^T \left\{ C_I (I_m - I(t))^2 + C_p \exp (P_m - P(t))^2 \right\} dt \quad (2)$$

where C_T is the total cost of inventory and production. C_p is the minimum production cost which occurs when the production rate equals P_m . The quantity P_m can be considered as the capacity of the manufacturing facility. An increase in production over this capacity may entail

additional equipment and manpower and can be very expensive. On the other hand, a decrease in production to a level below capacity can be equally expensive owing to the need for maintaining unused equipment and idle labor which cannot be reduced because of contract agreements. The quantity I_m can be considered as the storage capacity of inventory. C_I is the cost of carrying inventory. In many practical situations, the storage cost is a minimum when the storage capacity is completely utilized. Furthermore, the cost function given by equation (2) has a smoothing capability which is frequently desirable for many manufacturing processes. In this case, I_m and P_m can be considered as the desirable inventory and production levels, respectively. It is assumed that the sales forecast is known and given by the linear relation

$$Q(t) = a + bt \quad (3)$$

and the initial inventory is

$$I(0) = C \quad (4)$$

Equation (1), on substituting for $Q(t)$, becomes

$$\frac{dI(t)}{dt} = P(t) - a - bt \quad (5)$$

In order to apply the conjugate gradient techniques to solve this problem, equations (2) and (5) are approximated by difference equations. From equation (5), the following difference equation is obtained:

$$x(t + \Delta t) = x(t) + (P(t) - a - bt) \Delta t \quad (6)$$

where $x(t) = I(t) \quad (7)$

To obtain the cost, the integral in equation (2) has to be computed over the limits $n\Delta t$ to $(n+1)\Delta t$. If t is small, this integral can be approximated by

$$\int_{n\Delta t}^{(n+1)\Delta t} \left[C_I (I_m - I(t))^2 + C_P \exp (P_m - P(t))^2 \right] dt$$

$$= \left[C_I (I_m - I(t))^2 + C_P \exp (P_m - P(t))^2 \right] \Delta t \quad (8)$$

The following numerical values are assumed:

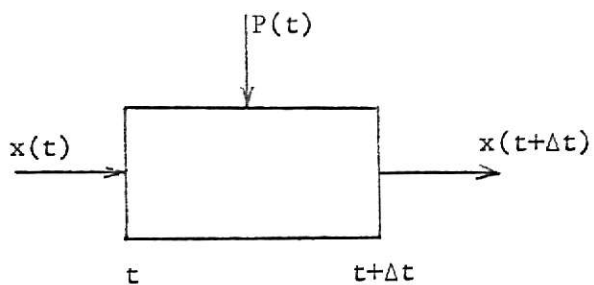
$a = 2$	$b = 1$	$c = 5$
$C_I = 0.1$	$I_m = 10$	$P_m = 5$
$C_P = 0.001$	$t_0 = 0$	$t_f = 1$

a. Solution by Fletcher and Reeves' method

The problem was first solved by applying the conjugate gradient method of Fletcher and Reeves. The number of stages into which the process should be divided is generally influenced by considerations of accuracy and computational cost. A large number of stages will give a better approximation to a continuous process and yield more accurate results. However, this results in increased computational time. For this study, solutions were obtained for a 5 stage and 10 stage process.

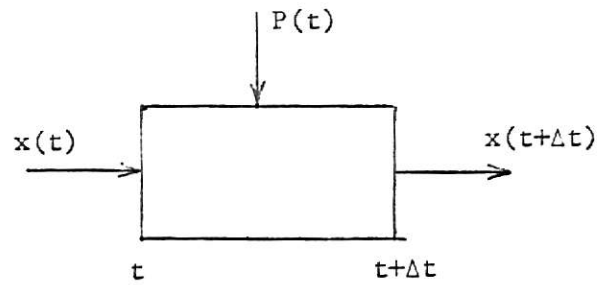
The computer program comprised of two subroutines. Subroutine FUNCT evaluated the inventory, function and gradient values required by the search subroutine FMCG. Subroutine FMCG was a standard subroutine from the IBM Scientific Subroutine Package and incorporated the Fletcher and Reeves' search procedure. This also included the one-dimensional search

Table 1. 5-stage process by Fletcher and Reeves' method



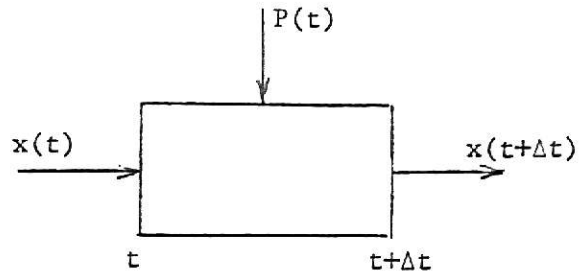
t	$t+\Delta t$	$x(t)$	$P(t)$	$x(t+\Delta t)$	C_T	$Q(t+\Delta t)$
0.0	0.2	5.00	7.15	5.99	0.43	2.20
0.2	0.4	5.99	7.09	6.93	0.69	2.40
0.4	0.6	6.93	6.89	7.79	0.84	2.60
0.6	0.8	7.79	6.64	8.55	0.91	2.80
0.8	1.0	8.55	6.78	9.31	0.94	3.00

Table 2. 10-stage process by Fletcher and Reeves' method



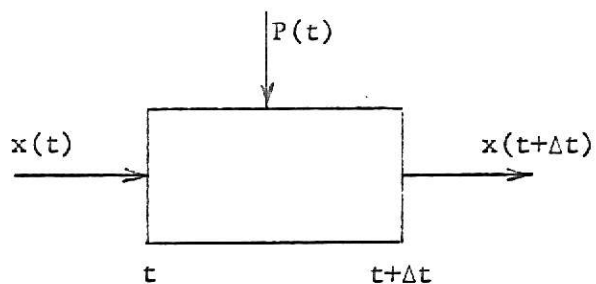
t	$t+\Delta t$	$x(t)$	$P(t)$	$x(t+\Delta t)$	C_T	$Q(t+\Delta t)$
0.0	0.1	5.00	7.17	5.51	0.24	2.10
0.1	0.2	5.51	7.13	6.00	0.43	2.20
0.2	0.3	6.00	7.08	6.48	0.58	2.30
0.3	0.4	6.48	7.03	6.94	0.69	2.40
0.4	0.5	6.94	6.98	7.39	0.77	2.50
0.5	0.6	7.39	6.90	7.82	0.84	2.60
0.6	0.7	7.82	6.77	8.23	0.88	2.70
0.7	0.8	8.23	6.62	8.61	0.90	2.80
0.8	0.9	8.61	6.49	8.97	0.92	2.90
0.9	1.0	8.97	6.63	9.33	0.93	3.00

Table 3. 5-stage process by Powell's method



t	$t+\Delta t$	$x(t)$	$P(t)$	$x(t+\Delta t)$	C_T	$Q(t+\Delta t)$
0.0	0.02	5.00	7.15	5.99	0.43	2.20
0.2	0.4	5.99	7.06	6.92	0.69	2.40
0.4	0.6	6.92	6.94	7.79	0.84	2.60
0.6	0.8	7.79	6.76	8.58	0.91	2.80
0.8	1.0	8.58	6.42	9.27	0.93	3.00

Table 4. 10-stage process by Powell's method



t	$t+\Delta t$	$x(t)$	$P(t)$	$x(t+\Delta t)$	C_T	$Q(t+\Delta t)$
0.0	0.1	5.00	7.17	5.51	0.24	2.10
0.1	0.2	5.51	7.13	6.00	0.43	2.20
0.2	0.3	6.00	7.08	6.48	0.57	2.30
0.3	0.4	6.48	7.03	6.94	0.69	2.40
0.4	0.5	6.94	6.96	7.39	0.77	2.50
0.5	0.6	7.39	6.89	7.82	0.84	2.60
0.6	0.7	7.82	6.80	8.23	0.88	2.70
0.7	0.8	8.23	6.68	8.61	0.90	2.80
0.8	0.9	8.61	6.51	8.97	0.92	2.90
0.9	1.0	8.97	6.15	9.29	0.93	3.00

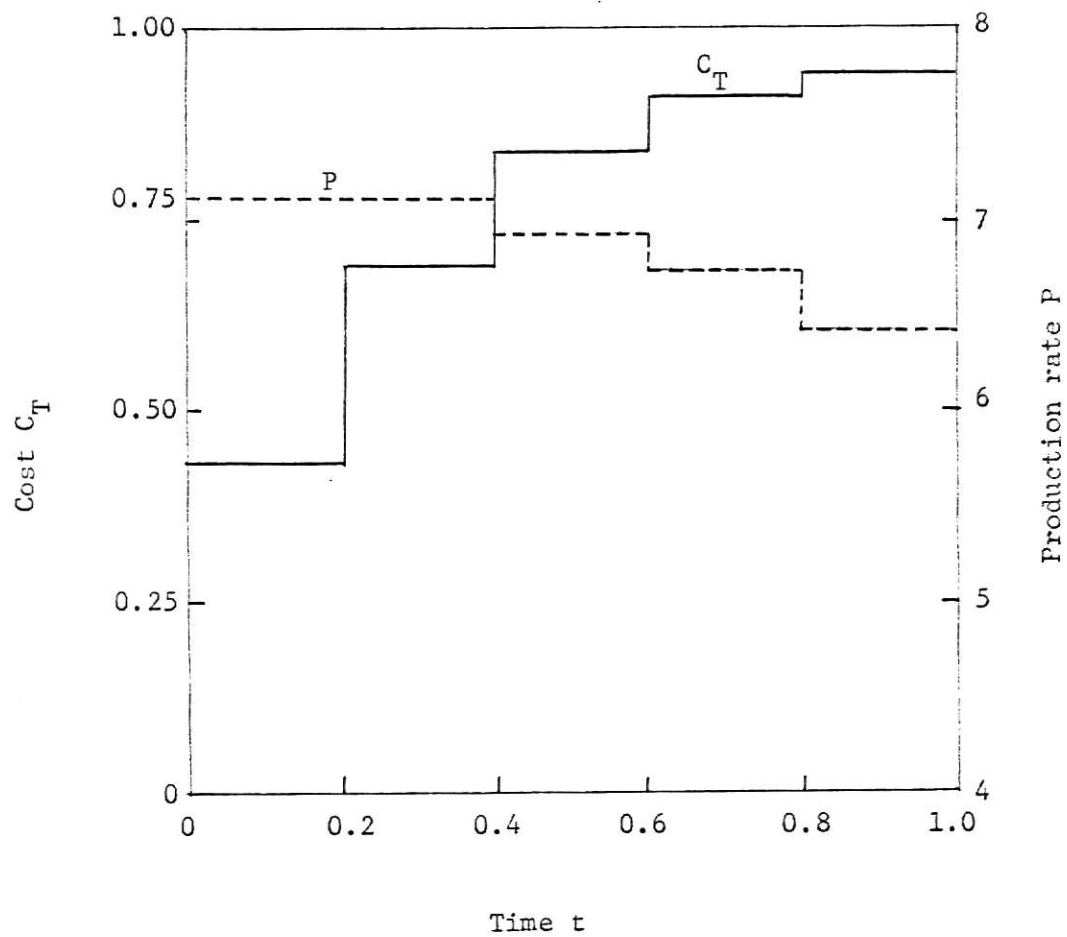


Fig. 1. Cost and production rate for the 5-stage process

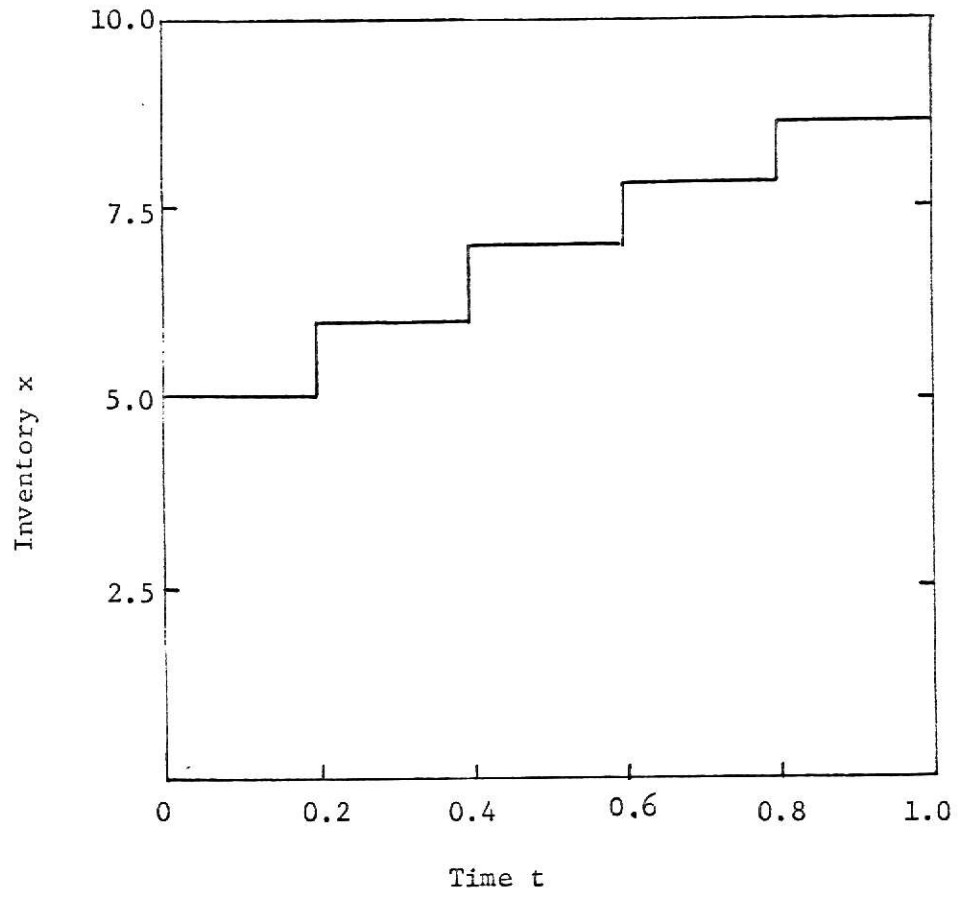


Fig. 2. Inventory for the 5-stage process

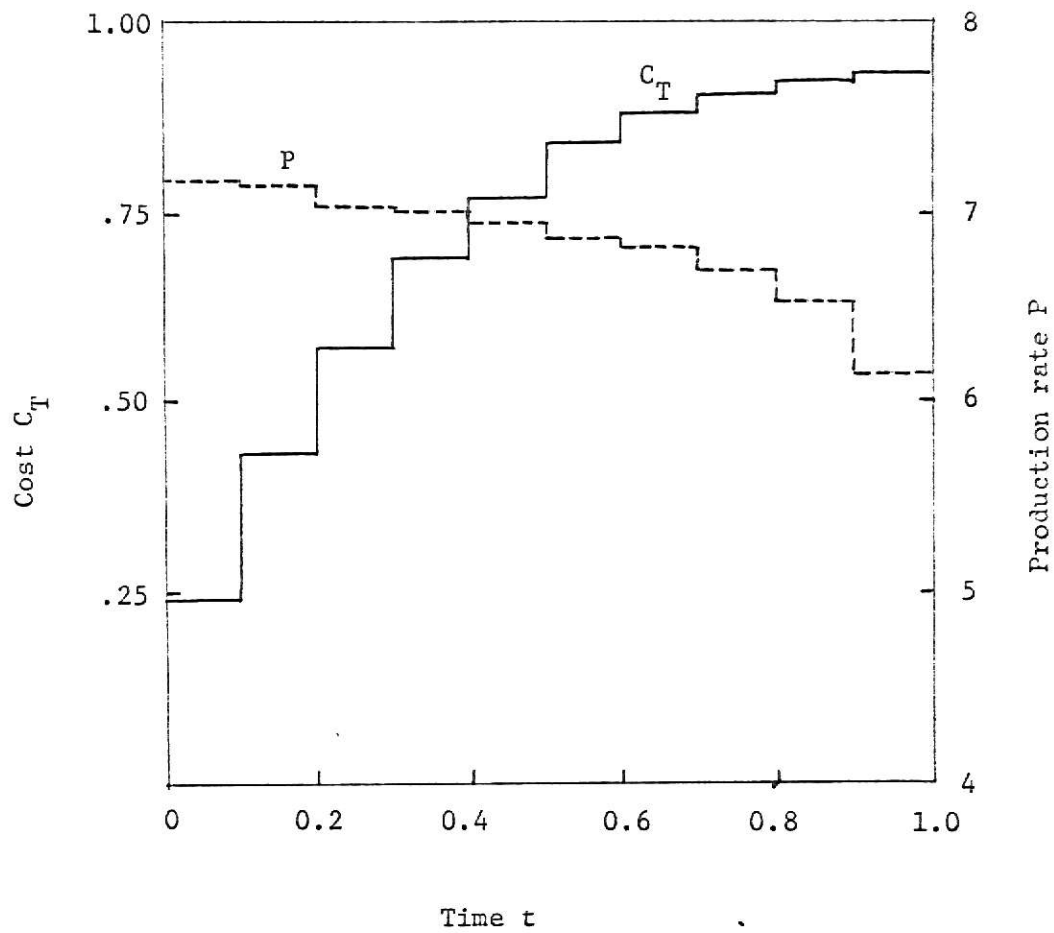


Fig. 3. Cost and production rate for the 10-stage process

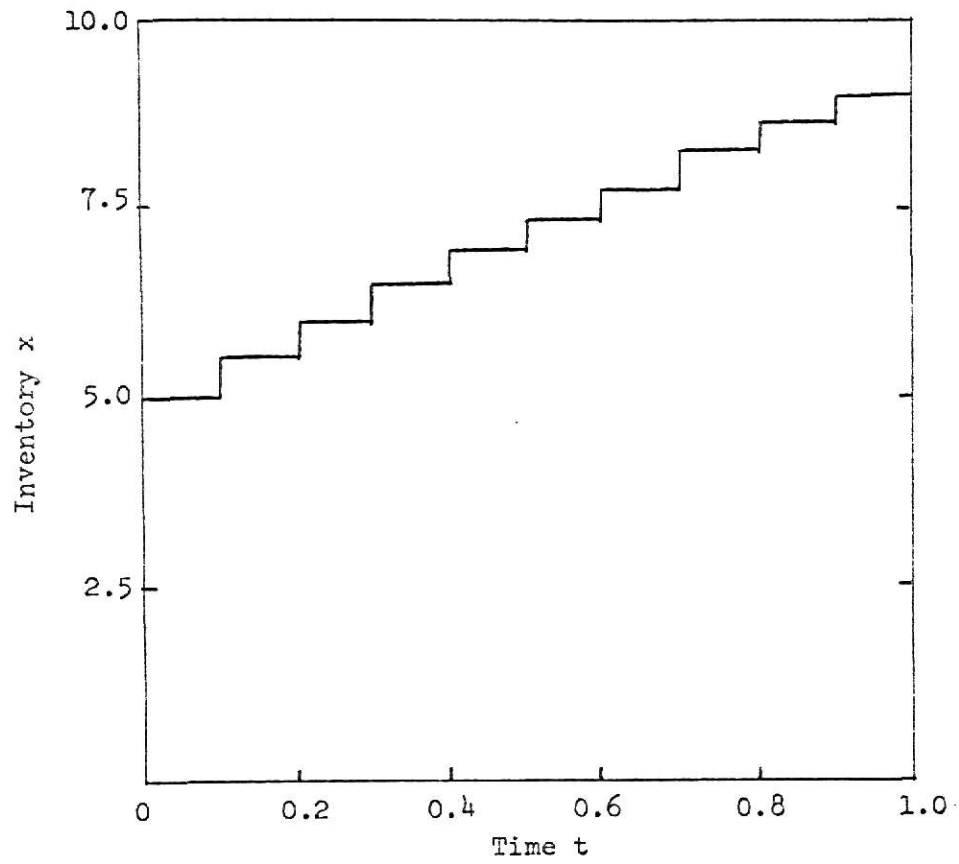


Fig. 4. Inventory for the 10-stage process

procedure of Davidon. The main program communicated the following parameters to the subroutines through the calling sequence:

- (i) the number of variables, n ,
- (ii) initial approximation of the variables P_1, P_2, \dots, P_n
- (iii) the required accuracy of computation, EPS,
- (iv) an integer to limit the total number of iterations, LIMIT and
- (v) an estimate of the value of the function at the unconstrained optimum, EST.

The flowchart for the computer program is given in Fig. 9, Appendix B and the program and sample results given in Appendix C.

Three different sets of values were used for the initial approximation to the variables P_1, P_2, \dots, P_n . It was observed that the optimal values remained the same irrespective of the initial approximation used. A summary of the results obtained is given in Tables 1 and 2. The optimal cost was 0.94 for a 5 stage process and 0.93 for a 10-stage process. There is no significant difference between the two values. It is therefore concluded that, for this particular model, the value of the optimal cost is not improved by increasing the number of stages. This is due to the fact that the inventory and cost functions are smooth functions, as can be seen from Figs. 1 and 2.

This problem has been solved by Lee and Shaikh [1] by using the functional gradient technique and the value of the optimal cost obtained by them was 0.94. There is thus excellent agreement with the results obtained with the Fletcher and Reeves' method.

b. Solution by Powell's method

The same problem was next solved by applying Powell's method. As

before, solutions were obtained for a 5 stage and a 10 stage process.

The computer program comprised of three subroutines. Subroutine POWELL incorporated Powell's search procedure. Subroutine FIBON calculated the unconstrained optimum along each direction computed by Powell's method, using the Fibonacci search procedure. Subroutine FUNCT evaluated the inventory and function values required by the main subroutine POWELL. The main program communicated the following parameters to the subroutines through the calling sequence:

- (i) the number of variables, n ,
- (ii) initial approximation of the variables P_1, P_2, \dots, P_n ,
- (iii) the required accuracy of computation, EPS, and
- (iv) an integer to limit the total number of iterations, LIMIT.

The flowchart for the computer program is given in Fig. 11, Appendix B and the program and sample results given in Appendix C.

As in the case of Fletcher and Reeves' method, the problem was solved using three different sets of values as the initial approximation to the variables P_1, P_2, \dots, P_n . Identical results were obtained in each case. A summary of the results is given in Tables 3 and 4. The value of the optimal cost was seen to be 0.93 for both the 5 and 10 stage processes. This confirms the earlier conclusion that, in this problem, an increase in the number of stages does not yield a more accurate value of the optimal cost.

c. Comparison

A comparison of the results obtained by the two methods gave similar values. However, Powell's method required about 35% to 130% more computing time. This was because Powell's procedure necessitated a large number

Table 5. Convergence rates for the 5 stage process

No. of Iterations	Fletcher & Reeves' method				Powell's method			
	P_1	P_3	P_5	C_T	P_1	P_3	P_5	C_T
	Initial approximation:				All $P_i = 1.0$			
0	1.00	1.00	1.00	8889.39	1.00	1.00	1.00	8889.39
1	6.44	6.44	6.44	1.03	5.96	5.98	5.99	1.14
2	7.21	6.73	6.64	0.94	7.18	6.95	6.42	0.93
3	7.17	6.83	6.72	0.94	7.15	6.94	6.42	0.93
4	7.15	6.84	6.78	0.94				
5	7.15	6.89	6.78	0.94				
6	7.15	6.89	6.78	0.94				
	Initial approximation:				All $P_i = 5.0$			
0	5.00	5.00	5.00	1.43	5.00	5.00	5.00	1.43
1	7.29	6.11	5.66	1.00	7.20	7.10	6.41	0.94
2	7.13	6.45	5.89	0.96	7.15	7.06	6.42	0.93
3	7.20	6.98	6.32	0.94				
4	7.16	7.03	6.42	0.94				
5	7.14	7.02	6.54	0.94				
7	7.17	6.96	6.74	0.94				
10	7.16	6.95	6.78	0.94				
11	7.16	6.95	6.78	0.94				
	Initial approximation:				All $P_i = 10.0$			
0	10.00	10.00	10.00	****	10.00	10.00	10.00	****
1	6.99	6.99	6.99	0.95	5.00	5.00	5.00	1.41
2	7.20	6.95	6.93	0.94	7.20	6.96	6.41	0.94
3	7.11	6.93	6.88	0.94	7.15	6.94	6.42	0.93
5	7.16	6.93	6.86	0.94				
6	7.16	6.93	6.86	0.94				

Table 6. Convergence rates for the 10 stage process

No. of Iterations	Fletcher & Reeves' method				Powell's method			
	P ₁	P ₅	P ₁₀	C _T	P ₁	P ₅	P ₁₀	C _T
Initial approximation: All P _i = 1.0								
0	1.00	1.00	1.00	8889.36	1.00	1.00	1.00	8889.36
1	6.44	6.44	6.44	1.02	5.85	5.90	5.99	1.16
2	7.25	6.76	6.51	0.93	7.21	6.98	6.14	0.93
3	7.15	6.87	6.54	0.93	7.17	6.96	6.15	0.93
4	7.16	6.95	6.57	0.93				
10	7.18	6.98	6.63	0.93				
11	7.17	6.98	6.63	0.93				
12	7.17	6.98	6.63	0.93				
Initial approximation: All P _i = 5.0								
0	5.00	5.00	5.00	1.41	5.00	5.00	5.00	1.41
1	7.62	6.31	5.34	1.04	7.23	7.01	6.13	0.93
2	7.62	6.31	5.34	1.04	7.17	6.97	6.15	0.93
3	7.62	6.33	5.35	1.04				
4	7.52	6.47	5.38	1.01				
10	7.23	7.06	5.65	0.94				
15	7.22	6.92	6.17	0.93				
20	7.18	7.00	6.55	0.93				
21	7.19	6.99	6.59	0.93				
22	7.19	6.99	6.59	0.93				
Initial approximation: All P _i = 10.0								
0	10.00	10.00	10.00	****	10.00	10.00	10.00	****
1	6.99	6.99	6.99	0.94	5.00	5.00	5.00	1.41
2	7.22	6.97	6.87	0.93	7.23	7.01	6.13	0.93
3	7.11	6.96	6.76	0.93	7.17	6.96	6.15	0.93
4	7.20	6.96	6.70	0.93				
10	7.18	6.97	6.69	0.93				
11	7.18	6.97	6.69	0.93				

of function evaluations during every iteration. As function evaluations took up most of the computing time in both the methods, this made a significant difference to their computing times.

In comparing the convergence rates for the two methods, it was found that Powell's method required a fewer number of iterations for convergence. Tables 5 and 6 gives the convergence rates for each method for 5 and 10 stage processes. It is evident that Powell's method is definitely faster in converging to the optimum, though it takes more computational time.

AN INVENTORY AND ADVERTISING MODEL WITH TWO STATE VARIABLES

This model is an extension of an advertising model formulated by Teichroew [14]. Consider a marketing situation where only a certain number of potential customers possess information about a company's product. Assume that the total number of persons in the group under consideration remains constant and that diffusion of information occurs only through personal contact. The number of "contacts" made by an average informed person in an arbitrary unit of time is given by a "contact coefficient". This coefficient is a fixed number which is the same for all members of the group. In a contact, the contactee receives the information if he does not already have it; if he already has it, the contact is wasted so far as increasing the number of people who have the information is concerned.

Let $K(0)$ = number of informed persons at time 0

$K(t)$ = number of informed persons at time t

N = total number of persons in the group

C = contact coefficient i.e the number of contacts made by one informed person per unit time

Then $K(t)/N$ = proportion of informed persons in the group
at time t .

$1-K(t)/N$ = proportion of uninformed persons at time t .

$CK(t) \cdot dt$ = contacts made during a time interval dt .

The increase in total number of informed persons, $dK(t)$, during a short interval of time dt is obtained only from contacts made with uninformed persons.

$$\text{Thus } dK(t) = CK(t) dt \left(1 - \frac{K(t)}{N} \right)$$

This is the differential equation:

$$\frac{dK(t)}{dt} = CK(t) \left(1 - \frac{K(t)}{N} \right) \quad (1)$$

Suppose that the company can increase the number of contacts by spending money on advertising. In particular, it can increase the number of contacts made by the informed persons, by an additional number A per unit of time. Eqn. (1) then becomes

$$\frac{dK(t)}{dt} = K(t)(C + A(t)) \left(1 - \frac{K(t)}{N} \right) \quad (2)$$

If each successful contact results in the sale of n units of the company's product and if $Q(t)$ represents the sale at time t , then

$$Q(t) = nK(t)$$

Assuming $n = 1$ for a particular product and substituting $Q(t)$ for $K(t)$ in eqn (2), we get

$$\frac{dQ(t)}{dt} = Q(t)(C + A(t)) \left(1 - \frac{Q(t)}{N} \right) \quad (3)$$

Next, the rate of change of the company's inventory is given by

$$\frac{dx(t)}{dt} = p(t) - Q(t) \quad (4)$$

where $P(t)$ = production rate at time t

The production rate for the product under consideration is assumed to be a linear function given by

$$P(t) = a + bt \quad (5)$$

where a and b are known constants. This assumption simplifies the model by avoiding a second control variable. The company's objective is to maximize the profit

$$S_T = \int_0^T (FQ(t) - C_I(P_I - x(t))^2 - C_A AQ(t)) dt \quad (6)$$

where S_T is the total net profit, F is the revenue from the sale of one unit of the product, C_I is the inventory carrying cost and C_A is the cost of advertising. P_I may be considered as the available storage capacity for the product.

Equations (3) through (6) completely represent the system. It has two state variables, $x(t)$ and $Q(t)$ and one control variable, $A(t)$. In order to apply the conjugate gradient techniques to solve this model, the differential equations are approximated by difference equations.

Thus, equation (3) is reduced to

$$Q(t + \Delta t) = Q(t) + Q(t) \left[C + A(t) \right] \left[1 - \frac{Q(t)}{N} \right] \Delta t \quad (7)$$

If this equation is used, it is possible for $Q(t + \Delta t)$ to exceed N , which

is not possible in the physical system being considered. To overcome this difficulty, the term $(1 - Q(t)/N)$ is replaced by $(1 - Q(t + \Delta t)/N)$ and this makes it impossible for $Q(t + \Delta t)$ to exceed N . By incorporating this change and rearranging the terms equation (7) becomes

$$Q(t + \Delta t) = \frac{Q(t) (1 + (C + A(t) \Delta t))}{1 + (C + A(t)) Q(t) \Delta t / N} \quad (8)$$

Equation (4) is reduced to

$$x(t + \Delta t) = x(t) + (P(t) - Q(t)) \Delta t \quad (9)$$

To obtain the profit, the integral in equation (6) has to be computed between limits $n\Delta t$ and $(n+1)\Delta t$. If Δt is small, the integral can be approximated by

$$S_T = \{FQ(t) - C_I(P_I - x(t))^2 - C_A A Q(t)\} \Delta t \quad (10)$$

The initial conditions and numerical values for the constants used are:

$P_I = 50$	$b = 100$	$x(0) = 20$
$C_I = 0.15$	$a = 70$	$Q(0) = 20$
$C_A = 1.5$	$c = 2$	$T = 1$
$N = 150$	$F = 10$	

It is assumed that the company has unlimited funds for advertising and thereby no constraint has to be imposed on the control variable $A(t)$.

a) Solution by Fletcher and Reeves' method

An attempt was made to solve this problem for 5 and 10 stage processes. The computer program was similar to the one used for solving the Inventory