An application of separable programming

By A. J. Akeroyd*

A method of solving a general nonlinear function of several variables within a constrained region using an interpolation technique with separable programming is described, with reference to a problem formulated by Box (1965).

Box (1965) describes a nonlinear programming problem that arose in his work, for which he described two new computational procedures. The problem is in fact linear except for the presence of four product terms. It may be of interest to point out that problems of this type can be solved by an extension of linear programming known as *separable programming*, and Box's problem has been used to illustrate this. The paper outlines the technique of separable programming and the interpolation procedure associated with it in the C–E–I–R mathematical programming code LP/90/94 for the IBM 7094 computer. The application of the technique to Box's problem is then described.

Separable programming and interpolation

The method of separable programming was first formulated by Miller (1963). It provides a simple technique for handling arbitrary nonlinear functions of single arguments in otherwise linear programming problems—and can readily be adapted to handle product terms.

Assume we have a variable z and a function $f(z)$ whose graph appears as in **Fig. 1**.

This curve can be replaced by a set of straight lines defined by a finite number of points (in this example 8 points have been used).

Allowing the coordinates of P_i to be (a_i, b_i) we can introduce 8 new non-negative variables $u_1 \ldots u_8$ such that

$$u_1 + \ldots + u_8 = 1 \tag{2.1}$$

$$a_1u_1 + \ldots + a_8u_8 = z \tag{2.2}$$

$$b_1u_1 + \ldots + b_8u_8 = f(z). \tag{2.3}$$

The variables $u_1 \ldots u_8$ are called a group of special variables, and in a particular problem each nonlinear constraint has a group associated with it.

If $u_1 = 1$ then $u_2 \ldots u_8 = 0$ and $z = a_1$, $f(z) = b_1$, the point P_1. If $u_1 = \frac{1}{2}$, $u_2 = \frac{1}{2}$, then $u_3 \ldots u_8 = 0$ and $z = \frac{1}{2}(a_1 + a_2)$, $f(z) = \frac{1}{2}(b_1 + b_2)$, a point halfway along the line P_1P_2. In this way, allowing two neighbouring special variables to take non-zero values, we can move along the set of lines P_1P_2, $P_2P_3 \ldots P_7P_8$. However, if two non-neighbouring special variables take non-zero values we have an invalid point (e.g. if $u_1 = \frac{1}{2}$, $u_3 = \frac{1}{2}$ we have a point midway along the line P_1P_3).

* *C–E–I–R Ltd.*, 30–31 *Newman Street, London, W.*1.

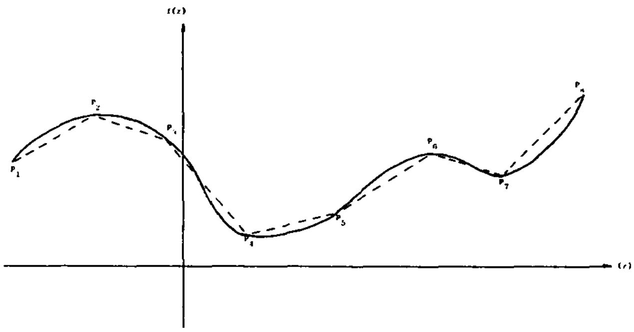

Fig. 1

Separable programming is a modification to the normal simplex method that restricts the choice of variable to enter the basis, and thereby avoids having any illegal combinations of the special variables.

In order to handle product terms by separable programming we use the relationship

$$z_1z_2 = \left(\frac{z_1 + z_2}{2}\right)^2 - \left(\frac{z_1 - z_2}{2}\right)^2 \tag{2.4}$$

which converts a product of two linear variables to the difference of two nonlinear functions of linear variables.

If we replace the right-hand side of equation (2.4) by $(u^2 - v^2)$ we have

$$u = \left(\frac{z_1 + z_2}{2}\right) \quad \text{and} \quad v = \left(\frac{z_1 - z_2}{2}\right).$$

If we can restrict each z_i to lie between 0 and 1, it follows that $0 \leqslant u \leqslant 1$ and $-\frac{1}{2} \leqslant v \leqslant +\frac{1}{2}$.

Thus we can handle a product term by introducing two groups of special variables (u and v), each group being defined by two equations, the first corresponding to equation (2.1) and the second to equation (2.2).

For example, taking five equally spaced variables in each group, we have

$$u_1 + u_2 + u_3 + u_4 + u_5 = 1 \tag{2.5}$$

$$\tfrac{1}{4}(u_2) + \tfrac{1}{2}(u_3) + \tfrac{3}{4}(u_4) + (u_5) = \tfrac{1}{2}z_1 + \tfrac{1}{2}z_2 \tag{2.6}$$

$$v_1 + v_2 + v_3 + v_4 + v_5 = 1 \tag{2.7}$$

$$-\tfrac{1}{2}(v_1) - \tfrac{1}{4}(v_2) \quad + \tfrac{1}{4}(v_4) + \tfrac{1}{2}(v_5) = \tfrac{1}{2}z_1 - \tfrac{1}{2}z_2. \tag{2.8}$$

Interpolation

In order to get an accurate solution from a separable programming problem it is necessary to specify a large number of special variables, and this can considerably slow down the speed with which the problem is solved. This difficulty can be overcome, as suggested by Miller (Miller, 1963), by starting with a few variables and, when a solution is reached, generating further variables in the region of this solution.

The interpolation agendum written by Mrs. P. Griffiths for the C–E–I–R LP/90/94 system interpolates between points P_i and P_{i+1} by fitting a cubic to the points P_{i-1}, P_i, P_{i+1} and P_{i+2} or by fitting a quadratic to three of these points if one of the end points does not exist. Rows corresponding to equations of the type (2.6) are given special names and the program looks at each of these rows. If one variable is in the basis (say the ith), it interpolates n equally spaced new variables between the $(i-1)$th and the ith and another n new variables between the ith and the $(i+1)$th, where n is a numerical parameter specified on the agendum call card. If there are two adjacent variables in the basis, it interpolates n equally spaced new variables between them. If there are two non-adjacent variables in the basis, it interpolates n variables in the interval spanned by them, although this situation rarely occurs in practice. If there are no variables in the basis, it ignores this group. The entries in other rows are calculated as cubic functions of the entries in the reference rows.

The use of interpolation to solve a problem

The example the author used is that described by Box (Box, 1965), except that the function $-f$ was minimized rather than f maximized.

The problem was to minimize the function f subject to the 8 constraints given below:

$$b = x_2 + 0{\cdot}01x_3 \tag{3.1}$$

$$x_6 = (k_1 + k_2x_2 + k_3x_3 + k_4x_4 + k_5x_5)x_1 \tag{3.2}$$

$$y_1 = k_6 + k_7x_2 + k_8x_3 + k_9x_4 + k_{10}x_5 \tag{3.3}$$

$$y_2 = k_{11} + k_{12}x_2 + k_{13}x_3 + k_{14}x_4 + k_{15}x_5 \tag{3.4}$$

$$y_3 = k_{16} + k_{17}x_2 + k_{18}x_3 + k_{19}x_4 + k_{20}x_5 \tag{3.5}$$

$$y_4 = k_{21} + k_{22}x_2 + k_{23}x_3 + k_{24}x_4 + k_{25}x_5 \tag{3.6}$$

$$x_7 = (y_1 + y_2 + y_3)x_1 \tag{3.7}$$

$$x_8 = (k_{26} + k_{27}x_2 + k_{28}x_3 + k_{29}x_4 + k_{30}x_5)x_1 + x_6 + x_7 \tag{3.8}$$

$$f = -(a_2y_1 + a_3y_2 + a_4y_3 + a_5y_4 + 7840a_6 - 100000a_0 - 50800ba_7 + k_{31} + k_{32}x_2 + k_{33}x_3 + k_{34}x_4 + k_{35}x_5)x_1 + 24345 - a_1x_6 \tag{3.9}$$

where x_1, x_2, x_3, x_4 and x_5 are the independent variables and the following constraints must be satisfied:

$$0 \leqslant x_1$$
$$1{\cdot}2 \leqslant x_2 \leqslant 2{\cdot}4$$
$$20 \leqslant x_3 \leqslant 60$$
$$9 \leqslant x_4 \leqslant 9{\cdot}3$$
$$6{\cdot}5 \leqslant x_5 \leqslant 7{\cdot}0$$
$$0 \leqslant x_6 \leqslant 294000$$
$$0 \leqslant x_7 \leqslant 294000$$
$$0 \leqslant x_8 \leqslant 277200.$$

The numerical values of the constants are specified in the appendix to Box (1965). It is convenient to start by reducing all product terms to products of variables lying between 0 and 1. So we write;

$$k_1 + k_2x_2 + k_3x_3 + k_4x_4 + k_5x_5 = c_1 + d_1z_1 \tag{3.10}$$

$$x_1 = d_2z_2 \tag{3.11}$$

$$y_1 + y_2 + y_3 = c_3 + d_3z_3 \tag{3.12}$$

$$k_{26} + k_{27}x_2 + k_{28}x_3 + k_{29}x_4 + k_{30}x_5 = c_4 + d_4z_4 \tag{3.13}$$

$$a_2y_1 + a_3y_2 + a_4y_3 + a_5y_4 + 7840a_6 - 50800a_7x_2 - 508a_7x_3 + k_{31} + k_{32}x_2 + k_{33}x_3 + k_{34}x_4 + k_{35}x_5 = c_5 + d_5z_5 \tag{3.14}$$

where the c_i and d_i are chosen so that all z_i lie between 0 and 1. We then have;

$$x_6 = c_1d_2z_2 + d_1d_2z_1z_2 \tag{3.15}$$

$$x_7 = c_3d_2z_2 + d_3d_2z_3z_2 \tag{3.16}$$

$$x_8 = c_4d_2z_2 + d_4d_2z_4z_2 + x_6 + x_7 \tag{3.17}$$

$$f = -c_5d_2z_2 - d_5d_2z_5z_2 + 24345 - a_1x_6. \tag{3.18}$$

For each of the product terms (say z_1z_2) we have four rows corresponding to equations (2.5), (2.6), (2.7) and (2.8):

u_1	u_2	u_3	u_4	u_5	v_1	v_2	v_3	v_4	v_5	z_1	z_2		
1	1	1	1	1								=	1
0	$+\frac{1}{4}$	$+\frac{1}{2}$	$+\frac{3}{4}$	1						$-\frac{1}{2}$	$-\frac{1}{2}$	=	0
					1	1	1	1	1			=	1
					$-\frac{1}{2}$	$-\frac{1}{4}$	0	$+\frac{1}{4}$	$+\frac{1}{2}$	$-\frac{1}{2}$	$+\frac{1}{2}$	=	0.

Equation (3.15) then becomes:

$$d_1d_2((1/16)u_2 + (1/4)u_3 + (9/16)u_4 + u_5 - (1/4)v_1 - (1/16)v_2 - (1/16)v_4 - (1/4)v_5) + c_1d_2z_2 - x_6 = 0.$$

Similarly a further six sets of special variables and their related equations are used to represent the three remaining product terms.

Since the z_i are confined to the range $0 \leqslant z_i \leqslant 1$ the values of the c_i and d_i can be determined by calculating

Table 1

	f	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	TIME	NUMBER OF ITERATIONS
	5,007,517	7·30902	1·2	60·0	9·3	6·52841	38,444	240,848	277,200	0·37	50
1st Interp.	5,326,545	5·70859	1·2	60·0	9·3	6·94147	64,999	209,110	277,200	0·72	64
2nd Interp.	5,226,302	5·30990	1·2	60·0	9·3	7·0	69,887	204,103	277,200	1·07	69
3rd Interp.	5,251,280	5·03160	1·60338	60·0	9·3	7·0	71,783	201,993	277,200	1·46	85
4th Interp.	5,293,116	4·57073	2·4	60·0	9·3	7·0	75,118	198,612	277,200	1·89	109
5th Interp.	5,284,811	4·53681	2·4	60·0	9·3	7·0	75,831	197,896	277,200	2·26	115
6th Interp.	5,280,189	4·53710	2·4	60·0	9·3	7·0	75,580	198,148	277,200	2·59	115
7th Interp.	5,280,342	4·53743	2·4	60·0	9·3	7·0	75,570	198,157	277,200	2·92	115
8th Interp.	5,280,335	4·53743	2·4	60·0	9·3	7·0	75,570	198,157	277,200	3·25	115
9th Interp.	5,280,336	4·53743	2·4	60·0	9·3	7·0	75,570	198,157	277,200	3·85	115

the maximum and minimum of equations (3.10) to (3.14). Because many of the coefficients are so large the rows (3.3) to (3.6) and (3.10) to (3.14) were scaled by a factor of 1000, and rows (3.15) to (3.18) by 10,000.

The problem has 40 rows and 68 vectors, the rows being the 11 constraints for $x_1 \ldots x_8$, the 16 defining the product terms, the 4 (including the functional) containing the product terms, the 5 defining the z_i and the 4 defining $y_1 \ldots y_4$. Many of the variables can be eliminated by substitution but it is not worth the effort in a problem as small as this.

Table 1 show the results obtained. In all cases the parameter n (the number of new variables interpolated) was 3. The two end columns show the time in minutes and the number of simplex iterations from the start of the run.

The work took four minutes on an IBM 7094 Mark II computer, and it cannot be claimed that this is particularly fast. Dr. McCormick reports in a private communication that he has solved the same problem in about 30 seconds on an IBM 7040 computer using the method described by Fiacco and McCormick (1964). However, it should be pointed out that separable programming is a convenient general method of dealing with nonlinearities in an otherwise linear programming problem, and has the important advantage over other apparently more general methods of constrained optimization that there is no real limit to the number of linear variables that can be represented: typical real separable programming problems have a few hundred such variables. Also the LP/90/94 system is most efficient on larger problems—with small problems a large proportion of the total time is spent in noniterative routines.

Acknowledgements

The author wishes to acknowledge his sincere thanks to Mr. E. M. L. Beale and Dr. P. F. Windley of C–E–I–R Ltd. for their valuable assistance in solving the problem, and to Mr. Beale for his constant help and advice in preparing this paper.

References

Box, M. J. (1965). "A new method of constrained optimization and a comparison with other methods," *The Computer Journal*, Vol. 8, p. 42.

Fiacco, A. V., and McCormick, G. P. (1964). "Computational Algorithm for the Sequential Unconstrained Minimization Technique for Nonlinear Programming," *Management Science*, Vol. 10, p. 601.

Miller, C. E. (1963). "The Simplex Method for Local Separable Programming" from *Recent Advances in Mathematical Programming*, p. 89. Edited by R. L. Groves and P. Wolfe. McGraw-Hill.