JOURNAL OF APPROXIMATION THEORY 12, 283–290 (1974)

Surface Fitting by Separation

E. S. CALL AND F. F. JUDD

Bell Laboratories, Whippany, New Jersey 07981

Communicated by Oved Shisha

1. INTRODUCTION

When fitting a function in two or more independent variables to a surface, the number of basis functions which make up the approximating model can be very large. The resulting system of equations, whose solution is the desired coefficients, is equally large. Solution of such a large system of equations requires sizable amounts of computer memory and time. A method of reducing these requirements is desired.

Standard least-squares curve and surface fitting techniques can be found, for example, in Draper and Smith [1]. There appear to have been two basic approaches to reducing computer memory and computation time requirements. The first involves the use of matrix Kronecker products. The second, in the special case of two independent variables, involves the repeated application of simple curve fitting techniques.

Building on the multivariate interpolation work of Davis [2], Greville [3] has implied that if the grid of base points is a Cartesian product of one-dimensional grids and the basis functions are separable, then the use of Kronecker products results in a decrease in computation time. Clenshaw and Hayes [4] show that a multiple regression (two independent variables) can be accomplished by repeated application of a single regression curve fitting routine. This second approach saves on both computer memory and computation time. In an interesting applications paper, Cornish [5] uses a compact two-sided matrix notation for adjusting an original model to include additional independent variables. He shows that this notation, when applicable, results in savings in computation time over that required to complete a new multiple regression analysis with the full set of independent variables.

In this paper it is shown that the technique suggested by Clenshaw and Hayes [4] can be generalized to k independent variables. With some accompanying restrictions on the data grid, this results in a significant reduction in computer memory and computation time. The new algorithm is not restricted to the use of orthogonal basis functions.

Definitions and some preliminary results are presented in Section 2. In Section 3 the last-squares normal equations are written in a compact matrix notation using Kronecker products. It is assumed in this derivation that the basis functions are separable. The algorithm for surface fitting by separation is then developed in Section 4. This development is made possible by combined use of a two-sided matrix notation similar to that of Cornish [5] and the Kronecker product of matrices. However, the final algorithm does not require the formation of Kronecker products. Finally, an approximate comparison of the regression techniques mentioned above is given in Section 5.

2. Definitions and Preliminary Results

Let x_i be the ith independent variable and x_{ij_i} be the j_ith value of x_i. The x_{ij_i}, $i = 1,..., k$ and $j_i = 1,..., r_i$, are termed *base points*. Let the fitting function, $g(x)$ be a linear combination of n *basis* functions, f_l. Then

$$g(\mathbf{x}) = g(x_1, ..., x_k) = \sum_{l=1}^{n} a_l f_l(x_1, ..., x_k). \tag{1}$$

If the f_l are *separable* (can be written as a product of functions, each in a single variable), we have $f_l(x_1, ..., x_k) = \prod_{i=1}^{k} f_i^{(l_i)}(x_i)$, where $f_i^{(l_i)}$ is the l_ith function in the ith independent variable. The functions $f_i^{(l_i)}$ will be termed *constituent functions*. If n_i is the number of constituent functions in the ith variable (note that $\prod_{i=1}^{k} n_i = n$), and if the coefficients are given additional subscripts to indicate exactly which product of constituent functions they multiply, Eq. (1) can be rewritten

$$g(\mathbf{x}) = \sum_{l_1=1}^{n_1} \cdots \sum_{l_k=1}^{n_k} a_{l_1 \cdots l_k} \prod_{i=1}^{k} f_i^{(l_i)}(x_i). \tag{2}$$

Polynomials provide a good example, in which case $f_i^{(l_i)}(x_i) = x_i^{l_i - 1}$.

Let the *Kronecker, or direct, product* of two matrices A and B be defined

$$A \times B = \{a_{ij} B\}.$$

M. C. Pease [6] proves the following:

THEOREM 1. *If A, B, C, and D are matrices having dimensions which make the products AC, BD, and $(A \times B)(C \times D)$ meaningful, then*

$$(A \times B)(C \times D) = AC \times BD.$$

Corollary 1.1 follows without difficulty.

COROLLARY 1.1. *For matrices of appropriate dimensions*

$$(A_1 \times A_2 \times \cdots \times A_k)(B_1 \times B_2 \times \cdots \times B_k) = A_1B_1 \times A_2B_2 \times \cdots \times A_kB_k .$$

The interested reader may verify the following:

THEOREM 2. *Let M' denote the transpose of matrix M. Then*

$$(A \times B)' = A' \times B'.$$

COROLLARY 2.1. $(A_1 \times A_2 \times \cdots \times A_k)' = A_1' \times A_2' \times \cdots \times A_k'$.

3. THE LEAST-SQUARES NORMAL EQUATIONS

At the jth observation, let f_{lj} be the value of the lth basis function and let y_j be the observed value. Based on the model given by Eq. (1), it is a well-known result [1] that the *least-squares normal equations*, whose solution for the a_i provides the desired coefficients, can be given in matrix form as follows

$$\begin{bmatrix} f_{11} & f_{12} & \cdots & f_{1r} \\ f_{21} & & & \vdots \\ \vdots & & & \\ f_{n1} & & \cdots & f_{nr} \end{bmatrix} \begin{bmatrix} f_{11} & f_{21} & \cdots & f_{n1} \\ f_{12} & & & \vdots \\ \vdots & & & \\ f_{1r} & & \cdots & f_{nr} \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix} = \begin{bmatrix} f_{11} & f_{12} & & f_{1r} \\ f_{21} & & & \vdots \\ \vdots & & & \\ f_{n1} & & \cdots & f_{nr} \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_r \end{bmatrix}, \tag{3}$$

where n is the number of basis functions and r is the number of base points. Defining matrices in the obvious manner, Eq. (3) is more simply written

$$F'FA = F'Y. \tag{4}$$

Now consider the case where the functions f_l are separable and $g(\mathbf{x})$ is given by

$$\begin{aligned} g(\mathbf{x}) &= \sum_{l=1}^{n} a_l f_l(\mathbf{x}) \\ &= \sum_{l_1=1}^{n_1} \cdots \sum_{l_k=1}^{n_k} a_{l_1 \cdots l_k} \prod_{i=1}^{k} f_i^{(l_i)}(x_i), \end{aligned} \tag{5}$$

where n_i is the number of constituent functions in the ith variable and k is the number of independent variables. Let $f_{ij_i}^{(l_i)} = f_i^{(l_i)}(x_{ij_i})$. Then, if r_i is the number of base points in the ith variable and if observed values are available for each combination of base points, Eq. (4) becomes

$$(F_1 \times F_2 \times \cdots \times F_k)' (F_1 \times F_2 \times \cdots \times F_k) A = (F_1 \times F_2 \times \cdots \times F_k)' Y, \quad (6)$$

where

$$F_i = \{f_{ij_i}^{(l_i)}\} = \begin{bmatrix} f_{i1}^{(1)} & f_{i1}^{(2)} & \cdots & f_{i1}^{(n_i)} \\ f_{i2}^{(1)} & & & \\ \vdots & & & \vdots \\ f_{ir_i}^{(1)} & & \cdots & f_{ir_i}^{(n_i)} \end{bmatrix}.$$

If some observed values are missing, those holes can be filled using a procedure described by Cadwell [7]. Using the corollary to Theorem 2, followed by Corollary 1.1 of Theorem 1, Eq. (6) becomes

$$(F_1'F_1 \times F_2'F_2 \times \cdots \times F_k'F_k)A = (F_1' \times F_2' \times \cdots \times F_k')Y. \quad (7)$$

It should be further observed (though it is not obvious) that if we let

$$\mathscr{A} = \begin{bmatrix} a_{1\cdots 111} & a_{1\cdots 112} & \cdots & a_{1\cdots 11n_k} \\ a_{1\cdots 121} & & & \\ \vdots & & & \\ a_{1\cdots 1n_{k-1}1} & & & \vdots \\ a_{1\cdots 211} & & & \\ a_{1\cdots 221} & & & \\ \vdots & & & \\ a_{n_1\cdots n_{k-2}n_{k-1}1} & & \cdots & a_{n_1\cdots n_{k-2}n_{k-1}n_k} \end{bmatrix}_{n_1 n_2 \cdots n_{k-1} \times n_k}$$

and

$$\mathscr{Y} = \begin{bmatrix} y_{1\cdots 111} & y_{1\cdots 112} & \cdots & y_{1\cdots 11r_k} \\ y_{1\cdots 121} & & & \\ \vdots & & & \\ y_{1\cdots 1r_{k-1}1} & & & \vdots \\ y_{1\cdots 211} & & & \\ y_{1\cdots 221} & & & \\ \vdots & & & \\ y_{r_1\cdots r_{k-2}r_{k-1}1} & & \cdots & y_{r_1\cdots r_{k-2}r_{k-1}r_k} \end{bmatrix}_{r_1 r_2 \cdots r_{k-1} \times r_k},$$

then Eq. (7) can be rewritten

$$(F_1'F_1 \times \cdots \times F_{k-1}'F_{k-1}) \mathscr{A} F_k'F_k = (F_1' \times \cdots \times F_{k-1}') \mathscr{Y} F_k. \quad (8)$$

Note that for $k = 2$ Eq. (8) is particularly compact. Eq. (8) will now be used to derive the final results.

4. Surface Fitting by Separation

We first consider the case $k = 2$. Then Eq. (8) is simply

$$F_1'F_1\mathscr{A}F_2'F_2 = F_1'\mathscr{Y}F_2 . \tag{9}$$

Note that the solution of the equations represented by (9) can be obtained by first solving

$$F_1'F_1\mathscr{B} = F_1'\mathscr{Y} \tag{10}$$

for $\mathscr{B}$ and subsequently solving

$$\mathscr{A}F_2'F_2 = \mathscr{B}F_2 \tag{11}$$

for $\mathscr{A}$. Equation (10) is a collection of r_2 normal equations for r_2 curve fits in x_1, one for each of the curves of constant x_2. Each curve fit uses the same approximating model. The matrix $\mathscr{B}$ is made up of the various coefficients. Equation (11) is another collection of normal equations, this time for n_1 curve fits in x_2, one for each coefficient of the model used to approximate the curves of constant x_2. Hence it is seen that the surface fit ($k = 2$) is performed by first fitting a family of curves in x_1 to lines of constant x_2, and subsequently obtaining curves in x_2 for each of the coefficients of the model in x_1.

It will now be shown that given r_k surface fits (one for each base point in the kth variable) in $k - 1$ independent variables, the surface fit in k independent variables is properly obtained by fitting a function in x_k to each coefficient of the surface fit in $k - 1$ independent variables. Simply let the matrix F_1 of the preceding discussion be replaced by the Kronecker product $F_1 \times F_2 \times \cdots \times F_{k-1}$ and let F_2 be replaced by F_k. Then since the r_k surface fits in $k - 1$ independent variables are given, the values of matrix $\mathscr{B}$ in the equation

$$(F_1 \times F_2 \times \cdots \times F_{k-1})'(F_1 \times F_2 \times \cdots \times F_{k-1})\mathscr{B} = (F_1 \times F_2 \times \cdots \times F_{k-1})'\mathscr{Y} \tag{12}$$

are known. The equation analogous to Eq. (11) is then

$$\mathscr{A}F_k'F_k = \mathscr{B}F_k , \tag{13}$$

which is precisely the set of normal equations for the previously mentioned curve fits. By induction on k the result can be carried to any finite number of independent variables.

Assuming that the dependent data is initially arranged in the k-dimensional array $\mathscr{Y}$, whose vectors are of the form:

$$\begin{bmatrix} y_{j_1 j_2} & \cdots & j_{m-1} & 1 & j_{m+1} & \cdots & j_k \\ y_{j_1 j_2} & \cdots & j_{m-1} & 2 & j_{m+1} & \cdots & j_k \\ \vdots & & & & & & \\ y_{j_1 j_2} & \cdots & j_{m-1} & r_m & j_{m+1} & \cdots & j_k \end{bmatrix}.$$

The above results can be summarized mathematically in the following:

Algorithm. A least-squares surface fit to a set of points in k independent variables, where the values of the independent variables are chosen in a rectangular grid, can be obtained by the following:

(1) Let $m = 1$.

(2) For each combination of the subscripts $j_1, j_2, \ldots, j_{m-1}, j_{m+1}, \ldots, j_k$,

$$\left(j_i \text{ ranges from } \begin{matrix} 1 \text{ to } n_i & i < m \\ 1 \text{ to } r_i & i > m \end{matrix}\right).$$

(a) Perform a bivariate least-squares regression analysis using the vector $(x_{m1}, x_{m2}, \ldots, x_{mr_m})'$ as the independent data and the vector

$$\begin{bmatrix} y_{j_1 j_2} & \cdots & j_{m-1} & 1 & j_{m+1} & \cdots & j_k \\ y_{j_1 j_2} & \cdots & j_{m-1} & 2 & j_{m-1} & \cdots & j_k \\ \vdots & & & & & & \\ y_{j_1 j_2} & \cdots & j_{m-1} & r_m & j_{m+1} & \cdots & j_k \end{bmatrix}$$

as the dependent data.

(b) Letting the regression coefficients be b_i, $i = 1, \ldots, n_m$, replace the element $y_{j_1 j_2 \cdots j_{m-1} i j_{m+1} \cdots j_k}$ by b_i, $i = 1, \ldots, n_m$.

(3) Let $m = m + 1$.

(4) If $m \leqslant k$ go to step 2.

(5) The element $y_{j_1 j_2 \cdots j_k}$ of the k-dimensional array $\mathscr{Y}$ is the regression coefficient of $f_1^{(j_1)} f_2^{(j_2)} \cdots f_k^{(j_k)}$.

If desired, weights can be introduced into the above analysis in a straightforward manner.

5. Comparison of Methods

An approximate comparison will now be made between the three multiple regression techniques mentioned in the introduction. These three techniques are:

1. Standard multiple regression (see [1]),
2. Kronecker product method (see Eq. (7)),
3. Surface fit by separation (see Section 4 for algorithm).

These three methods are compared on the basis of computer time and memory. To do this, each method is applied to a multiple least-squares regression using n constituent functions and r basis points in each of k independent variables. All three systems are solved by LL' decomposition and back substitution.

Forsythe and Moler [8] give the approximate number of arithmetic operations for LL' decomposition of an $n \times n$ matrix as $n^3/6$, and for back substitution as n^2. Therefore, the arithmetic operations and memory requirements can be estimated from

$$\begin{pmatrix}\text{Method 1}\\ \text{Standard}\end{pmatrix} \quad \begin{array}{l}\#\text{ of Operations} = \tfrac{1}{6}n^{3k} + n^{2k}\\ \text{Memory} = (n^k + 1)(r^k + 1) - 1\end{array}$$

$$\begin{pmatrix}\text{Method 2}\\ \text{Kronecker}\end{pmatrix} \quad \begin{array}{l}\#\text{ of Operations} = \tfrac{1}{6}kn^3 + n^{2k} + \sum_{i=2}^{k} n^{2i}\\ \text{Memory} = (n^k/2)(n^k + 3) + r^k\end{array}$$

$$\begin{pmatrix}\text{Method 3}\\ \text{Separation}\end{pmatrix} \quad \begin{array}{l}\#\text{ of Operations} = \tfrac{1}{6}kn^3 + k(k-1)\, n^2\\ \text{Memory} = n^k + r^k + nr.\end{array}$$

It is evident that as k increases, the difference between the three approaches

TABLE I

Comparison of Multiple Regression Methods

Conditions/method	Computation time (sec)	Memory (words)
A. $k = 2, n = 5, r = 10$		
1. Standard	3.2×10^{-2}	2625
2. Kronecker	1.3×10^{-2}	450
3. Separation	0.4×10^{-2}	175
B. $k = 4, n = 8, r = 15$		
1. Standard	11,463	207,414,721
2. Kronecker	23.8	8,445,377
3. Separation	0.43	54,841

becomes more significant quite rapidly. Assuming an average rate of 1 μsec/operation, the three methods are compared in Table I at two different sets of values for k, n and r.

References

1. N. R. Draper and H. Smith, "Applied Regression Analysis," John Wiley and Sons, New York, 1966.
2. H. S. Davis, The Kronecker product method of multivariate interpolation, Trans. of the Fifth Conference of Arsenal Mathematicians, Report No. 60-1, Office of Ordnance Research, U. S. Army, Durham, N. C., May 1960, 69–83.
3. T. N. E. Greville, Note on fitting of functions of several independent variables, *J. Soc. Indust. Appl. Math.* **9** (1961), 109–115.
4. C. W. Clenshaw and J. G. Hayes, Curve and surface fitting, *J. Inst. Math. Appl.* **1** (1965), 163–183.
5. E. A. Cornish, An application of the Kronecker product of matrices in multiple regression, *Biometrics* **13** (1957), 19–27.
6. M. C. Pease, "Methods of Matrix Algebra," Academic Press, New York/London, 1965.
7. J. H. Cadwell, A least squares surface-fitting program, *Comput. J.* **3** (1961), 266–269.
8. G. E. Forsythe and C. B. Moler, "Computer Solution of Linear Algebraic Systems," Prentice-Hall, Englewood Cliffs, NJ, 1967.