Computing best l_p approximations by functions nonlinear in one parameter

I. Barrodale, F. D. K. Roberts and C. R. Hunt

Department of Mathematics, University of Victoria, Victoria, B.C., Canada

This paper describes an algorithm for computing best l_1, l_2 and l_∞ approximations to discrete data, by functions of several parameters which depend nonlinearly on just one of these parameters. Such functions (e.g. $a_1 + a_2e^{cx}$, $a_1 + a_2 \sin cx$, $(a_1 + a_2x)/(1 + cx)$) often occur in practice, and a numerical study confirms that it is feasible to compute best approximations in any of the above norms when using these functions. Some remarks on the theory of best l_1 approximations by these functions are included.

(Received October 1969)

1. Introduction

The computation of best l_1, l_2 and l_∞ approximations by linear functions to discrete data has received much attention in the literature of the past few years. However, little has appeared concerning the actual determination of best nonlinear approximations, except for the rational l_∞ case and some l_2 approximations. In this paper we are concerned with the class of approximating functions which are nonlinear in only one of their several parameters (see **Table 1**). We describe an algorithm for obtaining best approximations by such functions, and present a brief summary of some numerical experiments involving this method. Complete details of these computations are contained in Hunt (1970).

The method is the following. For any given value of the nonlinear parameter, the determination of a best l_p approximation can be accomplished by using one of the existing linear approximation algorithms. Considering the error of approximation thus produced as a function of the nonlinear parameter only, the problem of determining a best nonlinear approximation is reduced to that of locating the minimum of a function of one variable. If we assume that this error of approximation is a unimodal function of the nonlinear parameter, then the best nonlinear approximation can be found very efficiently using a Fibonacci search technique.

This algorithm is implemented as a two-stage process. Phase I is a search over a grid of values of the nonlinear parameter to find a small interval containing the minimum. Phase II locates the minimum within this interval by using the Fibonacci method. In practice this scheme has proved to be extremely effective and widely applicable.

In the remainder of this section we make a precise mathematical statement of the problem at hand and the method proposed for its solution. The next section contains a summary of our empirical experience with this algorithm, and the last section consists of some remarks on the theory of best l_1 approximation which have arisen from this study.

The general approximation problem on a set $X = \{x_1, x_2, \ldots, x_N\}$ of real numbers (or points from some Euclidean space E_R) is to choose an approximating function $F(A, x)$ and select a particular form $F(A^*, x)$ which approximates a given function $f(x)$ satisfactorily on X. Here $A = \{a_1, a_2, \ldots, a_n\}$ is a set of free parameters, and $F(A, x)$ is a linear approximating function only if it depends linearly upon these parameters. Thus the functions $a_1 + a_2e^{a_3x}$ and $1/(1 + a_1x)$ are nonlinear. The most general linear function is $F(A, x) = \sum_{j=1}^{n} a_j\phi_j(x)$, where the $\phi_j(x)$'s are given linearly independent functions defined on X. $F(A^*, x)$ is called a best approximation in a norm $||\cdot||$ if for all choices of A, $||f(x) - F(A^*, x)|| \leqslant ||f(x) - F(A, x)||$. The three norms used in practice are:

$$l_1: \quad ||w(x)[f(x) - F(A, x)]||_1 = \sum_{i=1}^{N} w(x_i)|f(x_i) - F(A, x_i)|$$

$$l_2: \quad ||w(x)[f(x) - F(A, x)]||_2 = \left\{\sum_{i=1}^{N} w(x_i)[f(x_i) - F(A, x_i)]^2\right\}^{1/2}$$

$$l_\infty: \quad ||w(x)[f(x) - F(A, x)]||_\infty = \max_{1 \leqslant i \leqslant N} w(x_i)|f(x_i) - F(A, x_i)|$$

where $\{w(x_i)\}$ is a prescribed set of positive weights.

For linear functions $F(A, x)$, best approximations exist in all three norms, but only l_2 approximations are necessarily unique. Best l_1 and l_∞ linear approximations can be determined by linear programming (see Barrodale and Young, 1966), while l_2 approximations are produced by a standard orthonormalising routine. If the discrete data is affected by noise then the distribution of these random errors should dictate the choice of norm. In practice one chooses the l_∞, l_2 or l_1 norm, respectively, according as the errors are very small relative to the error of approximation, normally distributed, or subject to wild points (see Rice and White, 1964, and Barrodale, 1968).

In the discrete nonlinear case best approximations do not necessarily exist. Rice (1969) can be consulted for information on the present state of computational schemes for nonlinear approximation. Generally, progress towards developing algorithms for best nonlinear approximation has been slow; for example, virtually nothing is available in the l_1 case. For our particular problem in the l_2 norm there is an alternate method available which consists of solving the normal equations by reducing them to one nonlinear equation in the single nonlinear variable. In practice however, this technique is not guaranteed to produce a best approximation.

The success of our method depends upon the validity of the assumption that a certain function is unimodal in a given range. Roughly speaking, a function is unimodal if it has only one minimum in the interval to be explored, but it does not have to be smooth or even continuous. For any given tolerance within which values of x are considered indistinguishable, the Fibonacci method is the most efficient search technique known for locating the minimum of a unimodal function. Wilde (1964) contains precise definitions of these terms and a description of the Fibonacci method.

Finally, returning to the problem at hand, suppose that $F(A, x)$ is an approximating function which depends linearly upon parameters $a_1, a_2, \ldots, a_{n-1}$ and nonlinearly upon c, where $A = \{a_1, a_2, \ldots, a_{n-1}, c\}$. Let $A^* = \{a_1^*, \ldots, a_{n-1}^*, c^*\}$ be a set of best parameters for a given norm: we assume here that a best approximation exists. The justification for using a Fibonacci search in Phase II of our method is that the following assumption be true.

Assumption: $G(c) = \min_{a_1,\ldots,a_{n-1}} ||f(x) - F(A, x)||_p$ is unimodal in c, at least for some small interval containing c^*.

Phase I of the algorithm is to calculate $G(c)$ for a discrete set of values of c from an interval large enough to contain c^*, and thus produce a small interval containing c^*. Phase II locates the minimum within this interval by using the Fibonacci method. In many of the test cases described in the next section, the graph of $G(c)$ turns out to be unimodal in the whole range, thus emphasizing that at least in these instances our underlying assumption is very reasonable and Phase I itself is redundant.

The approximating functions $F(A, x)$ used in our numerical study of the algorithm are listed in **Table 1.**

Table 1

Approximating functions $F(A, x)$ considered in the numerical study

1. $a_1 + a_2 e^{cx}$
2. $(a_1 + a_2 x)/(1 + cx)$
3. $(a_1 + a_2 x + a_3 x^2)/(1 + cx)$
4. $a_1 + a_2 \log(1 + cx)$
5. $a_1 + a_2 \sin cx$
6. $a_1 + a_2/(1 + x)^c$
7. $a_1 \sin cx + a_2 \cos cx$
8. $e^{cx}(a_1 + a_2 x)$
9. $e^{cx}(a_1 + a_2 x + a_3 x^2 + a_4 x^3)$
10. $a_1 \sinh cx + a_2 \cosh cx$
11. $a_1 + a_2 x + a_3 (x - c)_+$
12. $a_1 + a_2 x + a_3 x^2 + a_4 x^3 + a_5 (x - c)_+^3$

2. Computational experience with the algorithm

To test the method numerically, we selected twelve approximating functions which are nonlinear in just one of their several parameters (see Table 1). For each one of these functions, discrete best approximations were computed in the l_1, l_2 and l_∞ norm to twelve different sets of data. Each data set consisting of 21 points with abscissae equally spaced on [0, 1], and ordinates recorded to 5D, was generated as follows:

(*a*) One set of exact data produced by assigning the parameters of the approximating function $F(A, x)$ particular numerical values.
(*b*) Three sets of noisy data produced as in (*a*) but with random errors added on. The distributions of these errors are, respectively,
Laplace: $g(\eta) = \exp(-2|\eta|), -\infty < \eta < \infty$
Normal: $g(\eta) = 1/\sqrt{\pi} \exp(-\eta^2), -\infty < \eta < \infty$
Uniform: $g(\eta) = 1, -\frac{1}{2} \leqslant \eta \leqslant \frac{1}{2}$.
(*c*) As (*b*) above but with different numerical values assigned to the parameters of the approximating function $F(A, x)$.
(*d*) $1 + \tan x$, $\sqrt{x}$, $\exp(-x^2)$, $\min(e^x, e^{1/2})$.
(*e*) 21 uniform random numbers in [0, 10].

Table 2 and **Table 3** together summarise the numerical study for the second approximating function considered, $F(A, x) = (a_1 + a_2 x)/(1 + cx)$. The entire study, which was performed in double precision arithmetic on an IBM 360/44, is too extensive to present complete details here: these can be found in Hunt (1970).

In Phase I the grid of values for the nonlinear parameter c was usually the 201 points defined as $c = -4(0{\cdot}04)4$, although in certain cases this was changed (e.g. for $a_1 + a_2 \sin cx$ we used $c = 0(0{\cdot}02)4$). This fine mesh size gives a good indication for the validity of the unimodality assumption over the entire interval. The computer printout from Phase I contains for each c value the error of approximation $G(c)$, and also a table of first differences of $G(c)$. It is thus a simple task to locate on $[-4, 4]$ the minimum of $G(c)$ within 0·04, and also to check for unimodality. In those cases where $G(c)$ was decreasing at one end of the interval, Phase I was repeated with a larger range of c values. In the few cases where a best approximation does not exist (e.g. approximation by $a_1 + a_2 \sin cx$ to our discrete data from $1 + \tan x$), this phenomenon can usually be detected by inspection of the output from Phase I. For all instances where best approximations do exist, the error of approximation appears to be unimodal in a sizeable region enclosing the minimum, and in many cases it is unimodal in the entire range considered. It is important to stress that these remarks are based on the observed behaviour of $G(c)$ for only a discrete set of values of c. However, we feel justified in claiming that the unimodality assumption required for Phase II is most reasonable in practice.

In Phase II the optimum value c^* is located by searching the small interval of unimodality determined in Phase I. The Fibonacci search technique was used here, and in most cases the initial interval was of width 0·08. The tolerance within which c^* was located was 0·000001. The results of Phase II for one of the rational approximating functions are given in Table 2.

Examination of Table 2 reveals that in some cases

Table 2

Complete details of best l_p approximations computed when using $F_2(A, x) = (a_1 + a_2x)/(1 + cx)$

Define $E_i = |f(x_i) - F(A^*, x_i)|$

DATA $x = 0(0{\cdot}05)1$	BEST l_1 PARAMETERS a_1^*	a_2^*	c^*	$\frac{\Sigma E_i}{21}$	BEST l_2 PARAMETERS a_1^*	a_2^*	c^*	$\sqrt{\frac{\Sigma E_i^2}{21}}$	BEST l_∞ PARAMETERS a_1^*	a_2^*	c^*	MAX E_i
f_1: $\frac{1+2x}{1+x}$	1·00000	1·99999	0·99999	0·00000	1·00000	1·99998	0·99999	0·00000	1·00000	2·00004	1·00003	0·00001
f_2: $\frac{3+15x}{1+2x} + \epsilon_L^1$	3·66230	1·15067	—0·31991	0·42187	3·87611	2·20955	—0·09485	0·54768	4·10302	5·57967	0·58859	0·98188
f_3: $\frac{3+15x}{1+2x} + \epsilon_N^1$	2·94022	19·82145	2·77740	0·34060	3·12648	19·02909	2·72056	0·43333	3·70285	5·09068	0·44487	0·76267
f_4: $\frac{3+15x}{1+2x} + \epsilon_U^1$	3·39576	13·05763	1·70736	0·21289	3·22852	13·46381	1·73528	0·26230	3·01563	17·11797	2·31248	0·38253
f_5: $\frac{2+3x}{1-\frac{1}{2}x} + \epsilon_L^2$	1·96249	2·86899	—0·52004	0·58200	2·18411	1·89889	—0·60894	0·71116	2·16127	1·51618	—0·67130	1·12155
f_6: $\frac{2+3x}{1-\frac{1}{2}x} + \epsilon_N^2$	2·17153	2·17032	—0·56465	0·37168	2·21005	2·34352	—0·54836	0·45779	2·45140	1·16545	—0·66382	0·78556
f_7: $\frac{2+3x}{1-\frac{1}{2}x} + \epsilon_U^2$	2·13059	2·20236	—0·58604	0·23153	2·09546	2·29602	—0·57978	0·26827	2·04776	2·56249	—0·56086	0·44058
f_8: $1 + \tan x$	1·02404	0·30111	—0·47446	0·00909	1·02267	0·28974	—0·48255	0·01104	1·01596	0·30265	—0·48116	0·01595
f_9: $\sqrt{x}$	0·17061	1·75785	0·95375	0·01331	0·06815	2·53883	1·68288	0·02469	0·04297	3·18124	2·36898	0·04297
f_{10}: e^{-x^2}	1·04932	—0·83400	—0·36501	0·01327	1·03799	—0·82712	—0·38542	0·01637	1·02426	—0·82747	—0·42732	0·02426
f_{11}: $\min(e^x, e^{1/2})$	0·94572	4·27947	2·06264	0·04524	0·93510	4·64232	2·24602	0·05405	0·91256	5·52770	2·70948	0·08746

where noisy data is being approximated by the data generating function itself, the parameters of best approximation are very different from those used to produce the uncontaminated data. This is because the random errors added to this data are quite large. As might be expected, the last data set (i.e. 21 random numbers) is difficult to approximate closely. The output of Phase I indicates that $G(c)$ does not vary greatly for this case, and appears to contain many local minima. Consequently we cannot guarantee that our results for this function represent best approximations, and we do not display them in Table 2. This random data set was included in the study to test the limitations of the algorithm.

Since our method is guaranteed to produce best approximations only if the unimodality assumption is valid, we have verified that the l_∞ results shown are in fact best approximations, by using the characteristic error equioscillation property. In the l_2 case we checked the results using the normal equations. No such characterisation theorems are available for any of the l_1 results or most of the nonrational l_∞ results.

Finally, we comment briefly on a comparison between our method and that which results from regarding the approximation problem as a pure minimisation problem. The error of approximation is a function of the n variables $a_1, a_2, \ldots, a_{n-1}, c$ and this can be minimised by using any of the known numerical techniques for locating the minimum of a function of several variables. In the l_1 and l_∞ cases the derivative of the error of approximation is not guaranteed to exist everywhere, so we are restricted to optimizing techniques which do not require derivatives. Powell (1964) has developed such a scheme and we used it with success on several of our examples. However, a disadvantage of any of these pure optimisation techniques is that they can converge to a local rather than a global minimum. Phase I of our method ensures that any such failure during Phase II is limited to the small region of c values over which the Fibonacci search is conducted. The major advantage of our method is that it is essentially a one-dimensional procedure rather than an n-dimensional routine. This fact makes the task of choosing upper and lower limits on the activities of the variables much simpler for our method, since we only have to limit c itself. Also, post-computational analysis on our method is far simpler than that which is required to investigate (say) intermediate calculations performed by any n-dimensional optimisation routine. Being able to regard our method as a one-dimensional procedure is both convenient computationally and quite enlightening analytically.

3. Some remarks on nonlinear best l_1 approximation

A theory of best approximation cannot be considered complete until the questions of existence, uniqueness, and characterisation of best approximations are answered, and a satisfactory method is available to determine these approximations in practice. In this section we include some remarks on existence and characterisation for the functions of Table 1, but we do not claim that these remarks constitute complete answers to these difficult questions.

Best linear l_1 approximations are guaranteed to exist. However, for the first ten nonlinear functions of Table 1 we have constructed data sets for which no best approximation exists (see Hunt, 1970). These same examples also demonstrate the possibility of nonexistence for best nonlinear l_p approximations with $0 < p < \infty$. For the last two functions in Table 1 (i.e. spline functions with one free knot) best discrete approximations always exist. Since uniqueness of best linear l_1 approximations is not assured, even in the case of polynomial approximation, any attempt to discover hypotheses which guarantee uniqueness of best nonlinear l_1 approximations would appear to be fruitless.

Rice (1964) contains the proof that the following property is characteristic of best linear l_1 approximations. Recall first that $\{\phi_1(x), \phi_2(x), \ldots, \phi_n(x)\}$ forms a Chebyshev set on $[a, b]$ if $\det|\phi_j(x_i)| \neq 0$ for every

Table 3

Random errors used in defining $f_2, f_3, \ldots, f_7$

x	ϵ_L^1	ϵ_N^1	ϵ_U^1	ϵ_L^2	ϵ_N^2	ϵ_U^2
0·0	0·12119	—0·05978	0·39576	—0·03751	—0·15878	—0·39276
0·05	0·59138	—0·20149	—0·32132	0·57285	0·14117	0·37623
0·10	1·63373	0·79925	0·44350	—0·23326	0·54565	—0·04967
0·15	—0·01027	0·91120	0·22425	0·94078	1·05350	0·14031
0·20	0·92979	0·15312	0·02417	0·03985	—0·57894	0·43535
0·25	—0·22861	—0·13592	0·16845	—0·41057	0·59835	0·26050
0·30	—0·37289	0·30655	0·36961	0·98547	—0·39153	—0·29630
0·35	—0·81332	0·57667	—0·20486	0·36237	0·63852	—0·04687
0·40	—0·27248	—0·09552	0·06392	—0·57615	—0·08338	—0·34234
0·45	—0·32600	—0·37644	0·11942	0·83611	—0·55820	0·21025
0·50	0·05851	—0·39987	—0·23385	—0·74548	—0·09386	—0·18647
0·55	0·06301	0·24062	0·49689	0·73964	—0·32734	—0·40369
0·60	1·00156	—0·88356	0·24025	0·30710	—0·38076	—0·08379
0·65	0·69787	0·00819	—0·10527	0·26385	—0·71750	—0·38699
0·70	—0·11041	0·08635	0·24790	—1·04985	0·42724	—0·42369
0·75	0·57299	0·23558	—0·14875	—0·60061	0·30996	0·49232
0·80	—0·13291	0·72789	—0·28855	—0·94504	0·66792	—0·00517
0·85	0·00631	—0·40999	0·07986	—0·02518	0·10825	0·31742
0·90	—0·15681	0·57207	0·46839	—0·75938	—0·15849	0·33342
0·95	0·99279	0·11019	—0·20942	1·82531	—0·17359	—0·02777
1·00	1·07695	—0·16066	0·41499	0·06650	—0·02672	0·46714

choice of n distinct points x_i from $[a, b]$.

Lemma:

Let $\{\phi_1(x), \phi_2(x), \ldots, \phi_n(x)\}$ be a Chebyshev set on [0, 1] and suppose the discrete point set $X \subset [0, 1]$. Then there exists a best l_1 approximation by $F(A, x) = \sum_{j=1}^{n} a_j \phi_j(x)$ on X which interpolates at least n points of $f(x)$.

This result says that in searching for a best approximation we need only consider those approximating functions which interpolate $f(x)$ at n or more points of X. This raises the question of whether there exists a best nonlinear approximation by, for example, $F(A,x)=a_1+a_2e^{cx}$ which interpolates $f(x)$ in three points. Our empirical study shows that although interpolation at the extra point (due to the presence of the nonlinear parameter c) is most definitely the rule rather than the exception, this phenomenon does not always occur even when the best approximation is unique. This fact clearly has consequences for those who might wish to devise a nonlinear l_1 algorithm based on interpolation at n points of $f(x)$. However, we can use the characteristic interpolatory property of linear approximation to help analyse the situation that prevails with any of the functions in Table 1 (in the case of the spline functions, which do not form Chebyshev sets, the analysis is more difficult but still carries through).

We shall consider l_1 approximation by functions of the form $F(A, x) = a_1 + a_2e^{cx}$, although the arguments can be extended to other nonlinear functions. Existence of a best approximation depends upon the data defining $f(x)$; note that no best approximation exists to the set $\{(0, 0), (1, 1), (2, 2)\}$. For $c \neq 0$ the set $\{1, e^{cx}\}$ forms a Chebyshev set on [0, 1], and hence there exists a best linear l_1 approximation by $a_1 + a_2e^{cx}$ which interpolates at least two points of $f(x)$. If there exists a best nonlinear approximation by $a_1 + a_2e^{cx}$ (c free), then there exists one which interpolates at least two points. Suppose $f(x) = \{(x_1, f_1), (x_2, f_2), \ldots, (x_N, f_N)\}$. For fixed nonzero c and $i \neq j$ let $F_{ij}(x; c) = a_1^{ij}(c)+a_2^{ij}(c)e^{cx}$ be the function which interpolates the ith and jth points of $f(x)$. We are guaranteed the existence of this unique function since $\{1, e^{cx}\}$ forms a Chebyshev set. Let $S_{ij}(c) = \sum_{k=1}^{N} |f_k - F_{ij}(x_k; c)|$, and define $G(c) = \min_{i \neq j} S_{ij}(c)$. If we now allow c to vary, since $F_{ij}(x; c)$ and consequently $S_{ij}(c)$ are continuous functions of c (except for $c = 0$), then $G(c)$ is also continuous. Furthermore, $G(c)$ is precisely that function of c which we calculate in our algorithm. Thus our method is equivalent to that of locating the minimum of a continuous curve (except at $c = 0$) which is the envelope of a finite number of curves, each of which is produced by interpolating a distinct pair of data points.

Defining $H = \min\{\lim_{c \to -\infty} G(c), \lim_{c \to 0} G(c), \lim_{c \to +\infty} G(c)\}$, then $G(c)$ will have a minimum provided there exists an approximation to $f(x)$ which has an error of approximation not exceeding H. We thus have a necessary and sufficient condition for the existence of a best l_1 approximation by $F(A, x) = a_1 + a_2e^{cx}$. The quantity H can be calculated for any data set in practice by considering the limit of $F_{ij}(x; c)$ curves as $c \to -\infty, 0, +\infty$. Provided an approximation can be found which produces a sum of errors not exceeding H, then the existence of a best l_1 approximation is guaranteed.

Although $G(c)$ is a continuous function, for nonzero c, this does not imply that it is necessarily unimodal in some neighbourhood of its minimum. On the other hand a moment's reflection can convince the reader that any continuous function which is not in fact unimodal in a small interval enclosing its minimum is rather special. Such functions (e.g. $f(x) = \left|x \sin \frac{1}{x}\right| + |x|$ for

nonzero $x \epsilon [-1, 1]$ and $f(0) = 0$) do not seem likely to be representable as an envelope of curves S_{ij}.

Concerning the question of interpolation at an extra point, it appears in practice that the minimum of $G(c)$ occurs when two of the $S_{ij}(c)$ curves intersect. That is, the minimum does correspond to interpolation at three points. However we do have counter-examples to show that this is not always the case. Interpolation at some n points of $f(x)$ will therefore not be a general characteristic property of nonlinear best l_1 approximations.

For other approximating functions the analysis is similar. For the rational function $(a_1 + a_2 x)/(1 + cx)$ the function $G(c)$ is also continuous except for poles at the zeros of the denominator. The quantity H in this case is defined to be $H = \min\{\lim_{c \to -\infty} G(c), \lim_{c \to +\infty} G(c)\}$.

Since best linear l_∞ approximations by Chebyshev sets are characterised by the error equioscillation property at $n + 1$ points, it is possible to carry out a similar analysis in this case. For the exponential case we define, for nonzero c, $F_{ijk}(c) = a_1^{ijk}(c) + a_2^{ijk}(c)e^{cx}$ to be that function which has the equioscillation property at the ith, jth and kth points. Defining $S_{ijk}(c)$, $G(c)$ and H as before we obtain necessary and sufficient conditions for the existence of best l_∞ approximations by functions of the type shown in Table 1.

4. Summary

The purpose of this paper is to report that the computation of best l_p approximations by functions nonlinear in only one of their several parameters is feasible in practice.

In particular it has transpired that in any of the three popular l_p norms, the error of approximation $G(c)$ is a well behaved function of the nonlinear parameter c. Indeed it appears that the assumption that this function is unimodal about its minimum is entirely reasonable. For a given discrete function $f(x)$ and approximating function $F(A, x)$, the amount of computation involved in the method depends upon the extensiveness of the preliminary search in Phase I and the tolerance specified for the Fibonacci search of Phase II. In any event the total time required will probably be 50 to 200 times that which is involved in computing a linear best approximation with c fixed.

Finally we note that the method can be extended to deal with functions which have more than one nonlinear parameter. We have computed a few examples where $F(A, x)$ includes two or three nonlinear coefficients. In these few cases we observed that if the error of approximation is regarded as a function of just one of these nonlinear parameters, the other nonlinear parameters are held fixed and the linear ones vary, then this function is still unimodal about the minimum.

Acknowledgements

We wish to acknowledge the financial assistance provided by NRC grants No. A5251 and A7143.

References

BARRODALE, I. (1968). L_1 Approximation and the analysis of data, *Applied Stat.*, Vol. 17, pp. 51–57.

BARRODALE, I., and YOUNG, A. (1966). Algorithms for best L_1 and L_∞ linear approximations on a discrete set, *Numer. Math.*, Vol. 8, pp. 295–306.

HUNT, C. R. (1970). Best l_p approximations by certain nonlinear functions. M.Sc. thesis. Univ. of Victoria, Victoria, B.C., Canada.

POWELL, M. J. D. (1964). An efficient method for finding the minimum of a function of several variables without calculating derivatives, *The Computer Journal*, Vol. 7, pp. 155–162.

RICE, J. R. (1964). *The Approximation of Functions*, Vol. 1. Reading, Mass.: Addison-Wesley.

RICE, J. R. (1969). *The Approximation of Functions*, Vol. 2. Reading, Mass.: Addison-Wesley.

RICE, J. R., and WHITE, J. S. (1964). Norms for smoothing and estimation, *SIAM Rev.*, Vol. 6, pp. 243–256.

WILDE, D. J. *Optimum Seeking Methods*, Englewood Cliffs, N.J.: Prentice-Hall.

Book review

16 Biproportional Matrices and Input-Output Change, by Michael Bacharach, 1970; 170 pages. (*Cambridge University Press*, £3.00)

The main interest of this book for computer programmers and systems analysts is in the mathematical description it contains of the techniques known as the 'Ras techniques'. These techniques are iterative processes for deducing the unknown input/output matrix $\{a_{ij}(1)\}$ at time 1, from the known input/output matrix $\{a_{ij}(0)\}$ at time 0. The new matrix is said to be 'bi-proportional' to the original matrix, where

$$a_{ij}(1) = r_i a_{ij}(0) s_j$$

and r_i and s_j satisfy relations like

$$\sum_i r_i a_{ij}\ (0) s_j = u_i(1)$$

and

$$\sum_j r_i a_{ij}(0) s_j = v_j(1)$$

where $[u_i]$ and $[v_j]$ are known vectors. The name 'Ras technique' follows from this notation.

The problem is, in fact, to deduce from a known matrix a new matrix having known row and column sums. Since the matrices which arise in input/output analysis are usually large in size, say 100×100 and the number of cycles needed to produce a resonable degree of convergence is often 200 or more, the solution of such a problem usually gives full scope to the user of a fairly large and fast computer. I programmed the process first for the Cambridge University computer EDSAC II, *circa* 1960, for matrices up to 50×50, and each cycle took about twenty minutes. Now the current Cambridge computer, Titan, takes about 50 secs, to do 100 cycles, on larger matrices.

Michael Bacharach's book is highly recommended as the first complete and detailed work dealing with these Ras techniques as applied in economics. The book deals with the applications of these techniques to various processes, such as Markov states, and two stage processes in general, and it concludes with the results produced when these researches are applied to data about the British economy.

L. J. SLATER (Cambridge)