On the convergence of exchange algorithms for calculating minimax approximations

By A. R. Curtis and M. J. D. Powell*

Given a function $f(x)$ and a range of the variable, S, the general minimax approximation problem is to determine that function of a class, C, which is the best approximation to $f(x)$ in the sense that the maximum error of the approximation as x ranges over S is minimized. We specialize to the usual case in which the functions of C are determined by n real parameters, $\lambda_1, \lambda_2, \ldots \lambda_n$, and we use the notation $\phi(x, \lambda_1, \lambda_2, \ldots, \lambda_n) \varepsilon C$. Most algorithms for calculating the required function $\phi(x, \boldsymbol{\lambda})$ depend on the maximum error of the minimax approximation occurring for $(n+1)$ distinct values of the variable x. In particular exchange algorithms seek these values iteratively, usually calculating on each iteration a best approximation over $(n+1)$ distinct points of S, $x_0, x_1, \ldots, x_n$ say. The value of the minimax error over the point set, $\eta(x_0, x_1, \ldots, x_n)$, is regarded as a function of the points; so are $\mu_t(x_0, x_1, \ldots, x_n)$ which are the values of λ_t, $t = 1, 2, \ldots, n$, yielding η. In this paper we present theorems on the first and second derivatives of η and μ_t. They provide much insight into the convergence of exchange algorithms.

1. Introduction

The theorems presented are stated in Section 5 and discussed in Section 6. The preceding sections necessarily include a synopsis of the basis of exchange algorithms.

2. Notation

$f(x)$ is the function to be approximated, S is the range of the variable x, and C is the class of approximating functions. It is assumed that the functions of C are determined by n real parameters, $\lambda_1, \lambda_2, \ldots, \lambda_n$, which are denoted by the vector $\boldsymbol{\lambda}$. Therefore the functions of C are called $\phi(x, \boldsymbol{\lambda})$. For example if the class of approximating functions is the class of polynomials of degree $(n-1)$, we may define

$$\phi(x, \boldsymbol{\lambda}) = \sum_{i=1}^{n} \lambda_i x^{i-1}. \tag{1}$$

To avoid consideration of limits, a number of assumptions will be made. They are

(i) S is closed and compact,
(ii) $f(x)$ is a continuous function of x,
(iii) For all real values of the parameters, $\phi(x, \boldsymbol{\lambda}) \varepsilon C$,
(iv) The parameters are such that those which define a minimax approximating function are necessarily finite, and
(v) $\phi(x, \boldsymbol{\lambda})$ is a continuous function of x and is also continuous in the parameters $\lambda_1, \lambda_2, \ldots, \lambda_n$.

Following the notation of Curtis and Powell (1966), we reserve $h(\boldsymbol{\lambda})$ for the maximum absolute error in approximating $f(x)$ by $\phi(x, \boldsymbol{\lambda})$. The assumptions guarantee the existence of a best approximating function, $\phi(x, \boldsymbol{\lambda}^*)$ say, and we abbreviate $h(\boldsymbol{\lambda}^*)$ by h^*. Therefore

$$h^* = \min_{\boldsymbol{\lambda}} h(\boldsymbol{\lambda}), \tag{2}$$

where

$$h(\boldsymbol{\lambda}) = \max_{x \varepsilon S} |f(x) - \phi(x, \boldsymbol{\lambda})|. \tag{3}$$

In exchange algorithms we do not calculate the required minimax approximation over the full range of x directly. Ratheı we iterate, calculating on each iteration an approximation over just $(n+1)$ distinct points of S. It is convenient to employ Stiefel's (1959) nomenclature and call a set of $(n+1)$ points, $(x_0, x_1, \ldots, x_n)$ a *reference*. The approximation required by an iteration is usually the minimax approximation over the reference, but it may not be because we may require an approximation subject to a condition known to hold for $\phi(x, \boldsymbol{\lambda}^*)$ being satisfied. For example, minimax rational approximations are necessarily pole-free over S (Maehly, 1963). In any event for each reference we define $\mu_1, \mu_2, \ldots, \mu_n$ to be the values of the parameters $\lambda_1, \lambda_2, \ldots, \lambda_n$ yielding the required approximation, and we regard the numbers μ_i as functions of $x_0, x_1, \ldots, x_n$. Further we define

$$\eta(x_0, x_1, \ldots, x_n) = \max_{\substack{x = x_i \\ i = 0, 1, \ldots, n}} |f(x) - \phi(x, \boldsymbol{\mu})|. \tag{4}$$

3. The basis of exchange algorithms

The basis of exchange algorithms is that

$$\eta(x_0, x_1, \ldots, x_n) \leqslant h^*, \tag{5}$$

because the minimax error over a part of S cannot exceed the minimax error over the whole. Further it is assumed that for some choice of $x_0, x_1, \ldots, x_n$, η attains the value h^*. Hence the minimization problem defined by (2) is replaced by the maximization problem defined by (5). It is of interest that in the event that $\phi(x, \boldsymbol{\lambda})$ depends linearly on the parameters and S is composed of a finite number of discrete points this reformulation is

* *Applied Mathematics Group, Theoretical Physics Division, A.E.R.E., Harwell, Didcot, Berks.*

equivalent to a duality theorem in linear programming (Stiefel, 1960).

Of course there are many classes C for which equality in (5) is not attained (Rice, 1960), so the applications of exchange algorithms are limited. However, they have a great advantage over minimizing $h(\boldsymbol{\lambda})$ directly. It is that for a given reference $x_0, x_1, \ldots, x_n$ the dependence of η on $f(x)$ is contained in just the $(n+1)$ function values $f(x_0), f(x_1), \ldots, f(x_n)$ while, for given $\lambda_1, \lambda_2, \ldots, \lambda_n$, the calculation of $h(\boldsymbol{\lambda})$ requires consideration of $f(x)$ over the whole of the range of the approximation. Usually this would be balanced by the greater difficulty of adjusting the variables of (5), but we shall prove that the derivatives of $\eta(x_0, x_1, \ldots, x_n)$ have properties that may be exploited.

In common with all exchange algorithms known to the authors, we make the restrictive assumption that for each choice of reference the error of the required approximation $\phi(x, \boldsymbol{\mu})$ is η at each of the points $x_0, x_1, \ldots, x_n$. Thus we have

$$|f(x_i) - \phi(x_i, \boldsymbol{\mu})| = \eta, \quad i = 0, 1, \ldots, n. \tag{6}$$

(6) is used to calculate the required numbers, $\mu_1, \mu_2, \ldots, \mu_n$ and η, there being $(n+1)$ equations in $(n+1)$ unknowns. Generally (6) has many solutions, but usually the equations may be written without modulus signs, known multipliers $s_i = \pm 1$ being introduced in the right-hand sides. Hence (6) becomes

$$f(x_i) - \phi(x_i, \boldsymbol{\mu}) = s_i\eta; \; i = 0, 1, \ldots, n. \tag{7}$$

Even (7) may have many solutions, and in the case of rational approximation this reintroduces the requirement that the approximation shall be pole-free (Maehly, 1963).

4. Assumptions made in proving the theorems

It is assumed that $f(x)$ and $\phi(x, \boldsymbol{\lambda})$ are differentiable functions of x and $\boldsymbol{\lambda}$.

It is assumed that $\eta(x_0, x_1, \ldots, x_n)$ and $\mu_i(x_0, x_1, \ldots, x_n)$, $i = 1, 2, \ldots, n$, are differentiable functions of $x_0, x_1, \ldots, x_n$.

It is assumed that $x_0, x_1, \ldots, x_n$ are such that the $(n+1) \times (n+1)$ matrix

$$D_{ij} = \begin{cases} s_i, j = 0 \\ \left[\dfrac{\partial}{\partial \lambda_j}\phi(x_i, \boldsymbol{\lambda})\right]_{\boldsymbol{\lambda}=\boldsymbol{\mu}}, j = 1, 2, \ldots, n, \end{cases} \tag{8}$$

is non-singular.

We require the inverse of D, and use the notation E_{ij} for its elements, which are, of course, functions of the reference. Further we write $\sigma_i = E_{0i}$, and the final assumption is that the numbers σ_i are non-zero. This is a sufficient condition for (6) to hold (Curtis and Powell, 1966).

There is no doubt that in many applications of algorithms based on the theorems it will not be possible to guarantee the assumptions. In particular for some choices of reference one σ_i may be zero. This should not deter one from trying an exchange algorithm, as such a reference may not occur.

5. The theorems

For the sake of brevity we use primes to denote differentiation with respect to the variable x, for example

$$\phi'(x_j, \boldsymbol{\mu}) = \left[\frac{\partial}{\partial x}\phi(x, \boldsymbol{\mu})\right]_{x=x_j}. \tag{9}$$

Theorem 1

$$\frac{\partial \eta}{\partial x_j} = \sigma_j\{f'(x_j) - \phi'(x_j, \boldsymbol{\mu})\} \tag{10}$$

and

$$\frac{\partial \mu_t}{\partial x_j} = E_{tj}\{f'(x_j) - \phi'(x_j, \boldsymbol{\mu})\}, \; t = 1, 2, \ldots, n. \tag{11}$$

Proof. Differentiating (7) with respect to x_j,

$$s_i\frac{\partial \eta}{\partial x_j} = \delta_{ij}\{f'(x_j) - \phi'(x_j, \boldsymbol{\mu})\} - \sum_{p=1}^{n} \frac{\partial \mu_p}{\partial x_j}D_{ip}. \tag{12}$$

Multiplying by σ_i and summing over i, equation (10) results; multiplying by E_{ti} and summing over i gives (11).

Theorem 2

If the reference is such that $\dfrac{\partial \eta}{\partial x_j} = 0$, then

$$\frac{\partial \mu_t}{\partial x_j} = 0, \; t = 1, 2, \ldots, n. \tag{13}$$

Proof. This is a simple consequence of (10), (11) and the assumption that $\sigma_j \neq 0$.

Theorem 3

If the reference is such that $\dfrac{\partial \eta}{\partial x_j} = 0$, then

$$\frac{\partial^2 \eta}{\partial x_j^2} = \sigma_j\{f''(x_j) - \phi''(x_j, \boldsymbol{\mu})\}, \tag{14}$$

and

$$\frac{\partial^2 \mu_t}{\partial x_j^2} = E_{tj}\{f''(x_j) - \phi''(x_j, \boldsymbol{\mu})\}, \; t = 1, 2, \ldots, n. \tag{15}$$

Proof. Differentiating (12) with respect to x_j,

$$s_i\frac{\partial^2 \eta;}{\partial x_j^2} = \delta_{ij}\{f''(x_j) - \phi''(x_j, \boldsymbol{\mu})\} - \delta_{ij}\sum_{p=1}^{n} \frac{\partial \mu_p}{\partial x_j}\frac{\partial D_{jp}}{\partial x_j}$$
$$- \sum_{p=1}^{n} \frac{\partial^2 \mu_p}{\partial x_j^2}D_{ip} - \sum_{p=1}^{n} \frac{\partial \mu_p}{\partial x_j}\frac{\partial D_{ip}}{\partial x_j}. \tag{16}$$

If $\dfrac{\partial \eta}{\partial x_j} = 0$, the second and fourth terms on the right-hand side of (16) are zero because of Theorem 2. Multiplying (16) by σ_i and summing over i gives (14), while multiplying by E_{ti} and summing over i gives (15).

Theorem 4

If the reference is such that $\frac{\partial\eta}{\partial x_j} = \frac{\partial\eta}{\partial x_k} = 0, j \neq k$, then

$$\frac{\partial^2\eta}{\partial x_k \partial x_j} = 0 \qquad (17)$$

and

$$\frac{\partial^2\mu_t}{\partial x_k \partial x_j} = 0,\ t = 1, 2, \ldots, n. \qquad (18)$$

Proof. Differentiating (12) with respect to x_k and using $\frac{\partial\mu_p}{\partial x_j} = \frac{\partial\mu_p}{\partial x_k} = 0$,

$$s_i \frac{\partial^2\eta}{\partial x_k \partial x_j} = \delta_{ij}\frac{\partial}{\partial x_k}\{f'(x_j) - \phi'(x_j, \boldsymbol{\mu})\} - \sum_{p=1}^{n} \frac{\partial^2\mu_p}{\partial x_k \partial x_j} D_{ip}. \qquad (19)$$

The first term on the right-hand side of (19) is zero if $k \neq j$ because there is no explicit dependence on x_k. (17) and (18) then follow as in Theorem 3.

6. Discussion of the theorems

Theorem 1 relates the derivatives, with respect to x_j, of both $\eta(x_0, x_1, \ldots, x_n)$ and μ_t, to the derivative of the error function of the approximation $\phi(x, \boldsymbol{\mu})$ at $x = x_j$. The obvious deduction may be made that if $x = x_j$ is a turning point of the error function these derivatives are zero.

Theorems 2, 3 and 4 provide insight into the ultimate convergence of exchange algorithms because, in the reference that maximizes $\eta(x_0, x_1, \ldots, x_n)$, the first partial derivatives of η are zero with respect to those x_i that are interior points of S. Those points of the reference that are boundary points, for instance the end-points of an interval, are usually recognized early in the calculation.

A special case of Theorem 2 has been stated by Murnaghan and Wrench (1959). It is easy to visualize why relatively inaccurate values of the required $x_0, x_1, \ldots, x_n$ can furnish a good approximation to the required function.

Theorem 3 provides the most significant term in the Taylor series expansion of $\eta(x_0, x_1, \ldots, x_n)$ and $\boldsymbol{\mu}(x_0, x_1, \ldots, x_n)$ about the required solution.

Theorem 4 is the most important since a corollary of it is that if $\eta(x_0, x_1, \ldots, x_n)$ were maximized by the simplest of methods, namely that of varying the points of the reference one at a time, the ultimate convergence rate would be quadratic. Curtis and Frank (1959) have experienced this fast convergence in calculating best polynomial approximations, and it was their paper that motivated us to postulate and prove the general theorems.

References

Curtis, A. R., and Powell, M. J. D. (1966). "Necessary conditions for a minimax approximation," *The Computer Journal*, Vol. 8, p. 358.

Curtis, P. C., and Frank, W. L. (1959). "An Algorithm for the Determination of the polynomial of best minimax approximation to a function defined on a finite point set," *J. Assoc. Comp. Mach.*, Vol. 6, p. 395.

Maehly, H. J. (1963). "Methods for fitting rational approximations, parts II and III," *J. Assoc. Comp. Mach.*, Vol. 10, p. 257.

Murnaghan, F. D., and Wrench, J. W. (1959). "The determination of the Chebyshev approximating polynomial for a differentiable function," *M.T.A.C.*, Vol. 13, p. 185.

Rice, J. R. (1960). "The characterization of best non-linear Tchebycheff approximations," *Trans. Amer. Math. Soc.*, Vol. 96, p. 322.

Stiefel, E. L. (1959). "Numerical methods of Tchebycheff approximation" from *On numerical approximation*, edited by R. E. Langer, Univ. of Wisconsin Press, p. 217.

Stiefel, E. L. (1960). "Note on Jordan elimination, linear programming and Tchebycheff approximation," *Numer. Math.*, Vol. 2, p. 1.

Notice: Future publication of Algorithms Supplement

Considerable interest has been shown in the Algorithms Supplement, which has been published in recent issues of *The Computer Bulletin*. As a result it has been decided to transfer the Supplement to this *Journal*, and the next Supplement will appear in the August issue.

A. S. Radford, who has been the editor of the Supplement since its foundation, has given up the work upon accepting a position in Canada. The new editor is

P. Hammersley
Northampton College of Advanced Technology
St. John Street, London E C 1.

All correspondence concerning the Supplement should be addressed to him.