JOURNAL OF APPROXIMATION THEORY **10**, 69–73 (1974)

The Continuity and Differentiability of the Parameters of Best Linear L_p Approximations

R. FLETCHER

Mathematics Branch, A.E.R.E. Harwell, Didcot, Berks, United Kingdom

J. A. GRANT

Department of Computational Science, University of Leeds, Leeds, England

AND

M. D. HEBDEN

Mathematics Branch, A.E.R.E. Harwell, Didcot, Berks, United Kingdom

Communicated by Oved Shisha

Received December 17, 1971

It is shown that the parameters of a best linear L_p approximation are both continuous and differentiable when considered as functions of p. An expression for each derivative is obtained. The result assumes only very mild conditions on the problem.

The Polya algorithm [4] (see also [5] and [1], for example) has been an interesting theoretical method for the determination of Chebyshev (L_∞) approximations by taking the limit of a sequence of L_p approximations as $p \to \infty$. Although the theoretical basis of the algorithm has been investigated (see [5, 2], for example) it is by no means complete, and the purpose of this paper is to present a new result concerning the continuity and differentiability of the parameters of the best linear L_p approximations as p is varied. Without any such result, the practical implementation of the Polya algorithm would be fraught with difficulty. The result is valid for a wide class of problems with minimal restrictions on the functions involved; in particular there is no need to assume the fulfillment of Haar's condition.

It will be assumed that the function $f(x)$ is to be approximated by a linear combination

$$l(\mathbf{a}, x) = \sum_{i=1}^{n} a_i \phi_i(x) \tag{1}$$

of linearly independent functions $\phi_1(x)$, $\phi_2(x),\ldots,$ $\phi_n(x)$, the region of approximation to be either a discrete point set or a closed continuum B of one or many variables. It is assumed that $f(x)$ and the $\phi_i(x)$ are continuous over the region. The discussion is given for the continuum, the extension to the discrete case being immediate. The vector of parameters of the best L_p approximation is denoted by $\mathbf{a}_p$ and it minimizes the L_p norm of the error curve

$$r(\mathbf{a}, x) = f(x) - l(\mathbf{a}, x), \tag{2}$$

namely

$$\| r(\mathbf{a}_p , x)\|_p = \min_{\mathbf{a}} \| r(\mathbf{a}, x)\|_p = \min_{\mathbf{a}} \left[\int_B | r(\mathbf{a}, x)|^p \, dB\right]^{1/p} \tag{3}$$

For convenience $r(\mathbf{a}_p , x)$ will usually be denoted by $r_p(x)$ or by r_p.

An equivalent formulation to (3) is to define $\mathbf{a}_p$ as the vector of parameters which solves the problem

$$\min_{\mathbf{a}} M(\mathbf{a}, p) = \min_{\mathbf{a}} \int_B | r(\mathbf{a}, x)|^p \, dB \tag{4}$$

where $M(\mathbf{a}, p)$ is now continuous, and twice continuously differentiable if $p \geqslant 2$ with respect to either $\mathbf{a}$ or p. The derivatives with respect to $\mathbf{a}$ can be denoted by

$$\frac{\partial}{\partial a_i} M(\mathbf{a}, p) = -p \int_B | r(\mathbf{a}, x)|^{p-2} \, r(\mathbf{a}, x) \, \phi_i(x) \, dB = -p\psi_i(\mathbf{a}, p) \tag{5}$$

and

$$\frac{\partial^2}{\partial a_i \, \partial a_j} M(\mathbf{a}, p) = p(p-1) \int_B | r(\mathbf{a}, x)|^{p-2} \, \phi_i(x) \, \phi_j(x) \, dB$$

$$= p(p-1) \, C_{ij}(\mathbf{a}, p). \tag{6}$$

Thus Eqs. (5) and (6) implicitly define a vector $\boldsymbol{\psi}(\mathbf{a}, p)$ and a matrix $C(\mathbf{a}, p)$ related to the derivatives of $M(\mathbf{a}, p)$, and such that

$$C_{ij}(\mathbf{a}, p) = (-1/(p-1))(\partial/\partial a_j) \, \psi_i(\mathbf{a}, p). \tag{7}$$

The equations defining $\mathbf{a}_p$ can now be obtained by equating the first derivatives of $M(\mathbf{a}, p)$ to zero, so that

$$\psi_i(\mathbf{a}, p) = 0 \qquad i = 1, 2,\ldots, n. \tag{8}$$

The basic result to be established in this paper is the following:

THEOREM. *For $p \geqslant 2$, $\mathbf{a}_p$ considered as a function of p is continuous and differentiable, its derivative being given by*

$$\frac{d}{dp}(\mathbf{a}_p)_i = \frac{1}{(p-1)} \sum_{j=1}^{n} [C(\mathbf{a}_p, p)^{-1}]_{ij} \int_B | r_p |^{p-2} r_p \log | r_p | \phi_j \, dB. \quad (9)$$

In proving the theorem two simple lemmas will be required.

LEMMA 1. *The matrix $C(\mathbf{a}, p)$ is strictly positive definite for any $\mathbf{a}$ and $p \geqslant 2$.*

Proof. See Fletcher *et al.* [3].

COROLLARY. *For any closed bounded region of $R^{n+1} = (\mathbf{a} \otimes p)$, $p \geqslant 2$, the smallest eigenvalue of $C(\mathbf{a}, p)$ is bounded below by a strictly positive number.*

LEMMA 2. *For any function $g(\mathbf{x})$, $\mathbf{x} \in R^k$, with continuous first derivatives,*

$$g(\boldsymbol{\xi} + \mathbf{h}) - g(\boldsymbol{\xi}) = \sum_{i=1}^{k} h_i \int_0^1 \frac{\partial g}{\partial x_i} (\boldsymbol{\xi} + \theta \mathbf{h}) \, d\theta. \quad (10)$$

Proof. Consider $g^*(\theta) = g(\boldsymbol{\xi} + \theta \mathbf{h})$ as a function of the single variable θ, and let $\mathbf{x} = \boldsymbol{\xi} + \theta \mathbf{h}$. Then

$$\begin{aligned} g(\boldsymbol{\xi} + \mathbf{h}) - g(\boldsymbol{\xi}) &= g^*(1) - g^*(0) \\ &= \int_0^1 \frac{dg^*}{d\theta} \, d\theta \\ &= \int_0^1 \sum_{i=1}^{k} \frac{\partial g}{\partial x_i} \frac{\partial x_i}{\partial \theta} \, d\theta \\ &= \int_0^1 \sum_{i=1}^{k} h_i \frac{\partial g}{\partial x_i} \, d\theta. \end{aligned}$$

Equation (10) follows on interchanging the summation and the integration.

Proof of the Main Theorem

Let $\mathbf{a}_p$ and $\mathbf{a}_{p+\delta p}$ be vectors of parameters satisfying (8) corresponding to p and $p + \delta p$ respectively where δp is an increment in p such that

$p + \delta p \geqslant 2$, and let $\boldsymbol{\delta}\mathbf{a} = \mathbf{a}_{p+\delta p} - \mathbf{a}_p$. Using (8) and then (10) with $\mathbf{x} = (\mathbf{a}, p)$

$$\begin{aligned} 0 &= \psi_i(\mathbf{a}_p + \boldsymbol{\delta}\mathbf{a}, p + \delta p) - \psi_i(\mathbf{a}_p, p) \qquad i = 1, 2,\ldots, n \\ &= \sum_{j=1}^{n} \delta a_j \int_0^1 \frac{\partial \psi_i}{\partial a_j} (\mathbf{a}_p + \theta \boldsymbol{\delta}\mathbf{a}, p + \theta \delta p)\, d\theta \\ &\quad + \delta p \int_0^1 \frac{\partial \psi_i}{\partial p} (\mathbf{a}_p + \theta \boldsymbol{\delta}\mathbf{a}, p + \theta \delta p)\, d\theta. \end{aligned}$$

Thus, (7) gives

$$\begin{aligned} &\sum_{j=1}^{n} \delta a_j \int_0^1 C_{ij}(\mathbf{a}_p + \theta \boldsymbol{\delta}\mathbf{a}, p + \theta \delta p)\, d\theta \\ &\qquad = \frac{\delta p}{p-1} \int_0^1 \frac{\partial \psi_i}{\partial p} (\mathbf{a}_p + \theta \boldsymbol{\delta}\mathbf{a}, p + \theta \delta p)\, d\theta. \end{aligned} \tag{11}$$

Now by differentiating $\psi_i(\mathbf{a}, p)$, as defined in (5), with respect to p,

$$\frac{\partial \psi_i}{\partial p} (\mathbf{a}, p) = \int_B | r(\mathbf{a}, x)|^{p-2} r(\mathbf{a}, x) \log | r(\mathbf{a}, x)| \, \phi_i(x)\, dB \tag{12}$$

which is bounded if $\mathbf{a}$ and p are bounded. Furthermore a matrix $\bar{C}$ can be defined by

$$\bar{C} = \int_0^1 C(\mathbf{a}_p + \theta \boldsymbol{\delta}\mathbf{a}, p + \theta \delta p)\, d\theta \tag{13}$$

and is positive definite by Lemma 1. From the corollary to Lemma 1, $\bar{C}^{-1}$ is bounded above. Now from (11)

$$\delta a_i = \frac{\delta p}{p-1} \sum_{j=1}^{n} [\bar{C}^{-1}]_{ij} \int_0^1 \frac{\partial \psi_j}{\partial p} (\mathbf{a}_p + \theta \boldsymbol{\delta}\mathbf{a}, p + \theta \delta p)\, d\theta, \tag{14}$$

and because of the bounds on $\bar{C}^{-1}$ and $\partial\psi_j/\partial p$ it follows that as $\delta p \to 0$ then $\delta a_i \to 0$ for all i. Therefore $\mathbf{a}_p$ is continuous. Furthermore as $\delta p \to 0$ both $\bar{C} \to C(\mathbf{a}_p, p)$ and

$$\int_0^1 \frac{\partial \psi_i}{\partial p} (\mathbf{a}_p + \theta \boldsymbol{\delta}\mathbf{a}, p + \theta \delta p)\, d\theta \to \frac{\partial \psi_i}{\partial p} (\mathbf{a}_p, p). \tag{15}$$

Thus from (14) the limit $\delta a_i/\delta p$ exists and is given, using (14), (15) and (12), by Eq. (9). Q.E.D.

Note in Proof. M. J. D. Powell has pointed out that Lemma 1 does not hold for the discrete problem. Hence our statement in the second paragraph, that the Theorem extends to the discrete problem, does not follow.

References

1. E. W. Cheney, "Introduction to Approximation Theory," McGraw–Hill, New York, 1966.
2. J. Descloux, Approximations in L^p and Chebyshev Approximations, *SIAM J.* **11** (1963), 1017–1026.
3. R. Fletcher, J. A. Grant, and M. D. Hebden, The calculation of linear best L_p approximations, *Comput. J.* **14** (1971), 276–279.
4. G. Polya, Sur une algorithm toujours convergent pour obtenir les polynomes de meilleure approximation de Tchebysheff pour une fonction continue quelconque, *Comptes Rendus*, **157** (1913), 840–843.
5. J. R. Rice, "The Approximation of Functions," Vol. 1, Addison–Wesley, Reading, Mass., 1964.