Research Report Number 71 - 3

Department of Computer Science

YALE UNIVERSITY

New Haven, Connecticut 06520

April 1971

DISCRETE TCHEBYCHEFF APPROXIMATION

FOR MULTIVARIATE SPLINES

by

#4 Martin H. Schultz 71-3

The preparation of this manuscript was supported in part by the Office of Naval Research (NR 044-401).

In this paper we give the theoretical analysis for the combination of two ideas in numerical analysis. The first is to approximate the Tchebycheff approximation to a function over a continuum, X, in R^M by Tchebycheff approximations over finite, discrete subsets of X, cf. [4], [5], [7], and [8], and the second is the use of multivariate spline functions as approximators. Experimental results for this combination have previously been reported in [5].

To be precise, let X be a compact subset of R^M. If Y is any closed subset of X and g is a real-valued, continuous function on Y, let

$$||g||_Y \equiv \max \{|g(y)| \; |y \in Y\}.$$

Given a real-valued, continuous function f and n linearly independent, real-valued, continuous basis functions $\{B_j(x)\}_{j=1}^n$, a common problem in numerical analysis is to solve the optimization problem

$$(1) \qquad \inf \{||f - \sum_{j=1}^{n} \beta_j B_j ||_X \mid \underline{\beta} \in R^n\} .$$

The standard difficulties are that (i) f is usually given only on a finite discrete point set, (ii) the basis functions $\{B_j\}_{j=1}^n$ don't satisfy the Haar condition in general so that Remez type algorithms don't work, and (iii) interpolation type schemes are impossible to define for general domains in R^M, $M \geq 2$.

The approach studied in this paper is to replace X by an appropriate discrete subset Y and to consider the approximate optimization problem

$$(2) \qquad \inf \{ || f - \sum_{j=1}^{n} \beta_j B_j ||_Y \mid \underline{\beta} \in R^n \} ,$$

which, following [4], [5], and [7], is solved by being reformulated as a linear programming problem, which in turn is solved by either the simplex or dual simplex method.

We now consider a reformulation of problem (2). Let $R^{n+1} \cap K \equiv \{ \underline{\alpha} \in R^{n+1} \mid \alpha_i \geq 0, 1 \leq i \leq n+1 \}$ and consider

$$(3) \qquad \inf \{ || f - \sum_{j=1}^{n+1} \alpha_j B_j ||_Y \mid \underline{\alpha} \in R^{n+1} \cap K \}$$

where $B_{n+1} \equiv - \sum_{j=1}^{n} B_j$. The following standard equivalence result is easy to prove.

<u>Theorem 1.</u> The two formulations (2) and (3) of the optimization problem are equivalent.

<u>Proof.</u> It suffices to show that

$$\{ \sum_{j=1}^{n} \beta_j B_j \mid \underline{\beta} \in R^n \} \equiv \{ \sum_{j=1}^{n+1} \alpha_j B_j \mid \underline{\alpha} \in R^{n+1} \cap K \}.$$

Clearly the right-hand side is a subset of the left-hand side and hence it suffices to show the converse. Given $\underline{\beta} \in R^n$, let $\alpha_{n+1} \equiv \max(0, -\min_{1\leq j\leq n} \beta_j)$ and $\alpha_j \equiv \alpha_{n+1} + \beta_j$, $1 \leq j \leq n$. Then

$$\sum_{j=1}^{n} \beta_j B_j = \sum_{j=1}^{n} \beta_j B_j + \alpha_{n+1} B_{n+1} + \alpha_{n+1} \sum_{j=1}^{n} B_j = \sum_{j=1}^{n} \alpha_j B_j + \alpha_{n+1} B_{n+1} = \sum_{j=1}^{n+1} \alpha_j B_j .$$

QED.

Let $Y \equiv \{y_i\}_{i=1}^{N}$, $f_i \equiv f(y_i)$, and $B_{ij} \equiv B_j(y_i)$, for all $1 \leq j \leq n+1$, $1 \leq i \leq N$. Then, if $\varepsilon(\underline{\alpha}) \equiv ||f - \sum_{j=1}^{n+1} \alpha_j B_j ||_Y$, we wish to minimize ε with respect to all $(\underline{\alpha}, \varepsilon) \in R^{n+2} \cap K$ subject to the constraints

$$(4) \qquad -\varepsilon \leq f_i - \sum_{j=1}^{n+1} \alpha_j B_{ij} \leq \varepsilon, \quad 1 \leq i \leq N,$$

i.e., there are n+2 unknowns and 2N constraints. Rewriting (4) we have

$$(5) \qquad \varepsilon - \sum_{j=1}^{n+1} \alpha_j B_{ij} \geq -f_i, \quad 1\leq i\leq N, \text{ and}$$

$$(6) \qquad \varepsilon + \sum_{j=1}^{n+1} \alpha_j B_{ij} \geq f_i, \quad 1\leq i\leq N.$$

But this is the form of a standard linear programming problem, i.e., given $\underline{b} \in R^{n+2}$, A a real $2N \times (n+2)$ matrix, and $\underline{c} \in R^{2N}$, minimize $(y, \underline{b})$ with respect to $y \in R^{n+2} \cap K$ subject to the constraint that $A\,\underline{y} \geq \underline{c}$. This problem has the dual problem of maximizing $(\underline{x},\underline{c})$ with respect to $\underline{x} \in R^{2N} \cap K$ subject to the constraint that $\underline{x}^T A \leq \underline{b}$, cf. [6].

In this case, $\underline{b} \equiv (0, \ldots, 0, 1)$, $\underline{y} \equiv (\varepsilon, \alpha_1, \ldots, \alpha_{n+1})$, $\underline{c} \equiv (-f_1, \ldots, -f_N, f_1, \ldots, f_N)$, and

$$A \equiv \begin{bmatrix} \vdots & \boxed{-B} \\ \vdots & \boxed{B} \end{bmatrix}$$ where $B \equiv [B_{ij}]$. Since, in general

we use the simplex method to solve a linear program the number of arithmetic operations involved is directly proportional to the number of constraints and in general $2N > (n+2)$. Hence, we expect that the dual program, solved by the simplex method, will be <u>more</u> efficient, cf. [6]. Furthermore, we remark that in general we expect to obtain a "degenerate" programming problem. However, such problems present no difficulties for the simplex method, cf. [1], [3], [4], and [6].

Hence, in general we seek to maximize

$$\sum_{i=1}^{N} \{ s_i f_i + t_i (-f_i) \} = \sum_{i=1}^{N} f_i(s_i - t_i) \text{ with respect to}$$

$(\underline{s},\underline{t}) \in R^{2N} \cap K$ subject to the constraints $\sum_{i=1}^{N} B_{ij}(s_i - t_i) \leq 0,$

$1 \leq i \leq n+1$ and $\sum_{i=1}^{N} (s_i + t_i) \leq 1.$

We turn now to the choice of the basis functions, $\{B_j\}_{j=1}^{n}$. We first examine the one dimensional case of $X \equiv [0,1]$. The classical choice for basis functions are the algebraic polynomials, cf. [8]. However, polynomials are numerically unstable and give rise to unwanted oscillations in the approximation. Moreover, the matrices A are dense and many function evaluations are needed. To remedy these we consider polynomial spline basis functions.

In particular, let P denote the set of all partitions, Δ, of [0,1] of the form, $\Delta : 0 = x_0 < \cdots < x_N < x_{N+1} = 1$ and for each $\Delta \in P$ and each positive integer d, S (Δ, d) denote the set of functions s(x) which are a polynomial of degree d on each subinterval $[x_i, x_{i+1}]$ defined by Δ and which are in $C^{d-1}[0,1]$. We remark that all the results of this paper may easily be extended to the case in which s(x) is assumed to be in C^{z_i}, $0 \leq z_i \leq d-1$, at each interior knot x_i, $1 \leq i \leq N$.

To define suitable basis functions for $S(\Delta,d)$, we follow [2] and augment the partition $\Delta : 0 = x_0 < \cdots < x_{N+1} = 1$ with the points $x_{-d} < x_{-d+1} < \cdots < x_{-1} < x_0$ and $x_{N+1} < x_{N+1+1} < \cdots < x_{N+1+d}$ to form a new partition $\tilde{\Delta} : x_{-d} < \cdots < x_0 < \cdots < x_{N+1} < x < \cdots < x_{N+1+d}$.

Letting $x_+^d \equiv \begin{cases} x^d , & \text{if } x \geq 0 , \\ 0 , & \text{if } x < 0 , \end{cases}$ and $W_i(x) \equiv \prod_{k=0}^{d+1} (x - x^{i+k})$

for $-d \leq i \leq N$, we define $M_{d,i}(x;\tilde{\Delta}) \equiv \sum_{k=0}^{d+1} (d+1) \dfrac{(x^{i+k}-x)_+^d}{W_i(x^{i+k})}$

for $-d \leq i \leq M$. As a basis for $S(\Delta,d)$ we take the restriction of the functions $\{ M_{d,i}(x; \tilde{\Delta}) \}_{i=-d}^{N}$ to the interval [0,1].

If Y is a finite subset of [0,1] and $| Y | \equiv \max_{x\varepsilon[0,1]} \min_{y\varepsilon Y} | x-y |$, then we obtain the following new error bound which relates the error in approximating f by a solution, s_Y, of the discrete optimization problem to the error in approximating f by a solution s_X of the continuous optimization problem. The proof uses a technique developed in [8] for the case of polynomial basis functions.

Theorem 2. If $\Delta \in P$ and $2d^2 \underline{\Delta}^{-1} |Y| < 1$, where $\underline{\Delta} \equiv \min_{0 \le i \le N} (x_{i+1} - x_i)$, then

(7) $$\|f - s_Y\|_X \le [2(1 - 2d^2 \underline{\Delta}^{-1} |Y|)^{-1} + 1] \, \|f - s_X\|_X .$$

Proof. By the triangle inequality

(8) $$\|f - s_Y\|_X \le \|f - s_X\|_X + \|s_X - s_Y\|_X .$$

Let $t \in [0,1]$ be such that $|(s_X - s_Y)(t)| = \|s_X - s_Y\|_X$.

Then there exists $y \in Y$ such that $|t - y| \le |Y|$ and

$$|(s_X - s_Y)(t)| \le |(s_X - s_Y)(y)| + |Y| \, \|D(s_X - s_Y)\|_X .$$

Hence, using the Markov inequality for polynomial splines, cf. [9],

(9) $$\|s_X - s_Y\|_X \le \|s_X - s_Y\|_Y + |Y| \, 2d^2 \underline{\Delta}^{-1} \|s_X - s_Y\|_X$$

and $$\|s_X - s_Y\|_X \le (1 - |Y| \, 2d^2 \underline{\Delta}^{-1})^{-1} \|s_X - s_Y\|_Y$$

$$\le (1 - |Y| \, 2d^2 \underline{\Delta}^{-1})^{-1} (\|f - s_X\|_Y + \|f - s_Y\|_Y)$$

$$\le (1 - |Y| \, 2d^2 \underline{\Delta}^{-1})^{-1} (2 \|f - s_X\|_X).$$ The required result

now follows from the triangle inequality and (7) and (8). QED.

If we assume a certain regularity of the function f, then we can bound the right hand side of (7). Using results of deBoor, cf. [2], we obtain

<u>Corollary 1</u>. Let $2d^2 \underline{\Delta}^{-1} |Y| < 1$ and $f \varepsilon W^{t,\infty} [0,1]$, $0 \le t \le d+1$,

i.e., $D^{t-1}f$ is absolutely continuous and $D^t f \varepsilon L^\infty [0,1]$.

There exists a positive constant, $K_{d,t}$, such that if $\Delta \varepsilon P$ and

$2d^2 \underline{\Delta}^{-1} |Y| < 1$ then

(10) $$||f - s_Y||_X \perp [2 (1 - 2d^2 \underline{\Delta}^{-1} |Y|)^{-1} + 1] K_{d,t} \bar{\Delta}^t ||D^t f||_X,$$

where $\bar{\Delta} \equiv \max\limits_{0 \le i \le N} (x_{i+1} = x_i)$.

We remark that for $S(\Delta,d)$, $|Y|$ need only be of order $\underline{\Delta}$, for Theorem 2 to hold. While for polynomials of degree n, $|Y|$ need be of order n^{-2}, for the corresponding result to hold, cf. [8].

We may obtain still a further Corrollary about computing the maximum absolute value of a polynomial spline function, s(x). The idea is that by sampling the size of a spline at a sufficiently large number of points we may give a rigorous estimate of it everywhere.

<u>Corollary 2</u>. If $\Delta \varepsilon P$, $s(x) \varepsilon S(\Delta,d)$, and $2d^2 \underline{\Delta}^{-1} |Y| < 1$, then

(11) $$||s||_Y \leq ||s||_X \leq (1 - 2d^2 \underline{\Delta}^{-1} |Y|)^{-1} ||s||_Y ,$$

and

(12) $$0 \leq ||s||_Y - ||s||_Y \leq [(1 - 2d^2 \underline{\Delta}^{-1} |Y|)^{-1} - 1] ||s||_Y$$
$$\leq (2d^2 \underline{\Delta}^{-1} |Y|) (1 - 2d^2 \underline{\Delta}^{-1} |Y|)^{-1} ||s||_Y .$$

We now turn to the multivariate case. Let $\Omega \varepsilon R^M$ be a closed set contined in the unit cube $\chi_{i=1}^{M} [0,1]_i$ in R^M and for each $1 \leq i \leq N$ let $\Delta_i : 0=x_1 < x_2 < \cdots < x_{N_i} < x_{N_i+1} = 1$ be a partition of $[0,1]_i$. Let P_M denote the set of all partitions, P, of the cube of the form $P \equiv \chi_{i=1}^{M} \Delta_i$, $\overline{P} \equiv \max_{1 \leq i \leq M} \{\overline{\Delta}_i\}$, and $\underline{P} \equiv \min_{1 \leq i \leq M} \{\underline{\Delta}_i\}$, i.e., $\underline{P}$ is the minimum distance between two partition points. Furthermore, let $S(d,P) \equiv \chi_{i=1}^{N} S(d,\Delta_i)$, i.e., S(d,P) is the space of multivariate polynomial spline functions of degree d with respect to P, $\Omega_P \equiv \{x \varepsilon \Omega \mid$ the "N" - cell" of P containing x is contained in $\Omega\}$, and $Y_P \equiv \{ y \varepsilon Y | y \varepsilon \Omega_P \}$.

Finally, let $|Y_P| \equiv \max_{x\in\Omega_P} \min_{y\in Y_P} \inf \{ \int_{\alpha \in \Gamma(x,y)} ||d\underline{\alpha}||_{\ell_1} \mid \Gamma(x,y)$ is a piecewise smooth curve all of whose points lie in Ω_P and which connect y to x , i.e., given $x \in \Omega_P$ there exists $y \in Y_P$ such that the ℓ_1 - distance in Ω_P between x and y is no more than $|Y_P|$.

The following result is a multivariate analogue of Theorem 2.

Theorem 3. If $\Delta \in P$ and, $2d^2\ \underline{P}^{-1}\ |Y_P| < 1$, then

(13) $$|| f - s_{Y_P} ||_{\Omega_P} \leq [2 (1 - 2d^2\ \underline{P}^{-1}\ |Y_P|)^{-1} +1]||f - s_{\Omega_P}||_{\Omega_P}$$

Proof. $$|| f - s_{Y_P} ||_{\Omega_P} \leq ||f - s_{\Omega_P} ||_{\Omega_P} + ||s_{Y_P} - s_{\Omega_P} ||_{\Omega_P}.$$

Let $t \in \Omega_P$ be such that $|s(t)| \equiv | s_{Y_P} (t) - s_{\Omega_P} (t) |$

$= ||s_{Y_P} - s_{\Omega_P} ||_{\Omega_P}$. There exists a point $y \in Y_P$ such that

$$| s(t)| \leq | s(y)| + \sum_{i=1}^{N} | D_i\ s (\xi_i)|\ | y_i -t_i|$$

$$\leq || s ||_{Y_P} + \sum_{i=1}^{N} || D_i\ s ||_{\Omega_P} | y_i - t_i |$$

$$\leq || s ||_{Y_P} + 2d^2 \sum_{i=1}^{N} \underline{\Delta}_i^{-1} || s ||_{\Omega_P} | y_i - t_i |$$

$$\leq \|s\|_{Y_P} + 2d^2 \underline{P}^{-1} |Y_P| .$$

Thus,

$$\| s_{Y_P} - s_{\Omega_P} \|_{\Omega_P} \leq (1 - | Y_P| \, 2d^2\underline{P}^{-1})^{-1} \|s_{Y_P} - s_{\Omega_P}\|_{Y_P},$$

and the result follows as in Theorem 2. QED.

Let $W^{t,\infty}(\Omega)$ denote the closure of the set of real-valued, infinitely differentiable functions on Ω with respect to the norm

$$\| \phi \|_{W^{t,\infty}(\Omega)} \equiv \max_{|\alpha| \leq t} \|D^\alpha \phi\|_{L^\infty(\Omega)} .$$

Using the results of [9] we obtain the following multivariate analogue of Corollary 1 of Theorem 2.

Corollary 1. Let $f \in W^{t,\infty}(\Omega)$, $0 \leq t \leq d+1$.

There exists a positive constant, $C_{d,t}$, such that if $P \in P_M$ and $2d^2 \underline{P}^{-1} |Y_P| < 1$, then

(14) $$\|f - s_{Y_P} \|_{\Omega_P} \leq [2 (1 - 2d^2 \underline{P}^{-1} |Y_P|)^{-1} +1] C_{d,t}\overline{P}^t\|f\|_{W^{t,\infty}(\Omega)}$$

Similiarly, we can prove the following multivariate analogue of Corollary 2 of Theorem 2.

<u>Corollary 2.</u> If $P \varepsilon P_M$, $s \varepsilon S(P,d)$ and $2d^2 \underline{P}^{-1} |Y_P| < 1$, then

(15) $|| s ||_{Y_P} \leq ||s||_{\Omega_P} \leq (1 - |Y| 2d^2 \underline{P}^{-1})^{-1} || s ||_{Y_P}$, and

(16) $0 \leq || s ||_{\Omega_P} - || s||_{Y_P} \leq [(1 - |Y_P| 2d^2 \underline{P}^{-1})^{-1} - 1] || s ||_{Y_P}$

$$\leq (2d^2 |Y_P| \underline{P}^{-1}) (1 - |Y_P| 2d^2 \underline{P}^{-1})^{-1} || s ||_{Y_P} .$$

References

[1] Beale, E.M.L., Cycling in the Dual Simplex Algorithm. Naval Res. Logist. Quart. 2, 269-275 (1955).

[2] deBoor, C., On Uniform Approximation by Splines. J. Approx Theory 1, 219-235 (1968).

[3] Charnes, A., Optimality and Degeneracy in Linear Programming. Econometrica 20, 160-170 (1952).

[4] Descloux, J., Dégénérescence dans les approximations de Tschebyscheff lineaires et discretes. Numer. Math. 3, 180-187 (1961).

[5] Each, R.E. and W.L. Eastman, Computational Methods for Best Spline Function Approximation. J. Approx. Theory 2, 85-96 (1969).

[6] Hadley, G., Linear Programming. Addison-Wesley, Reading. (1962).

[7] Kelley, J.E., Jr., An Application of Linear Programming to Curve Fitting. SIAM Journal on Applied Math 6, 15-22 (1958).

[8] Rivlin. T.J., and Cheney E.W., A Comparison of Uniform Approximations on an Interval and a Finite Subset Thereof. SIAM J. Numer. Anal. 3, 311-320 (1966).

[9] Schultz, M.H., L^∞ - Multivariate Approximation Theory. SIAM J. Numer. Anal. 6, 161-183 (1969).