Discrete Tchebycheff Approximation for Multivariate Splines

MARTIN H. SCHULTZ

Department of Computer Science, Yale University, New Haven, Connecticut 06520*

Received July 22, 1971

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

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

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

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

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

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

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.

* The preparation of this manuscript was supported in part by the Office of Naval Research (NR 044-401).
We now consider a reformulation of problem (2). Let
\[ R^{n+1} \cap K = \{ \mathbf{a} \in R^{n+1} \mid \alpha_i \geq 0, 1 \leq i \leq n+1 \} \]
and consider
\[ \inf \| f - \sum_{j=1}^{n+1} \alpha_j B_j \|_\gamma \mid \mathbf{a} \in R^{n+1} \cap K, \]
(3)
where \( B_{n+1} = -\sum_{j=1}^{n} B_j \). The following standard equivalence result is easy to prove.

**Theorem 1.** The two formulations (2) and (3) of the optimization problem are equivalent.

**Proof.** It suffices to show that
\[ \sum_{j=1}^{n} \beta_j B_j = \sum_{j=1}^{n+1} \alpha_j B_j \mid \mathbf{b} \in R^n \]
\[ \mathbf{a} \in R^{n+1} \cap K. \]
Clearly the right-hand side is a subset of the left-hand side and hence it suffices to show the converse. Given \( \mathbf{b} \in R^n \), let \( \alpha_{n+1} = \max(0, -\min_{1 \leq j \leq n} \beta_j) \) and \( \alpha_j = \alpha_{n+1} + \beta_j \), \( 1 \leq j \leq n \). Then
\[ \sum_{j=1}^{n} \beta_j B_j = \sum_{j=1}^{n} \beta_j B_j + \alpha_{n+1} B_{n+1} + \alpha_{n+1} \sum_{j=1}^{n} B_j = \sum_{j=1}^{n} \alpha_j B_j + \alpha_{n+1} B_{n+1} = \sum_{j=1}^{n+1} \alpha_j B_j. \]
Q.E.D.

Let \( Y = \{ y_i \}_{i=1}^{N}, f_i = f(y_i), \) and \( B_{ij} = B_j(y_i) \), for all \( 1 \leq j \leq n + 1, 1 \leq i \leq N \). Then, if \( \epsilon(\mathbf{a}) = \| f - \sum_{j=1}^{n+1} \alpha_j B_j \|_\gamma \), we wish to minimize \( \epsilon \) with respect to all \( (\mathbf{a}, \epsilon) \in R^{n+2} \cap K \) subject to the constraints
\[ -\epsilon \leq f_i - \sum_{j=1}^{n+1} \alpha_j B_{ij} \leq \epsilon, \quad 1 \leq i \leq N, \]
(4)
i.e., there are \( n + 2 \) unknowns and \( 2N \) constraints. Rewriting (4) we have
\[ \epsilon - \sum_{j=1}^{n+1} \alpha_j B_{ij} \geq -f_i, \quad 1 \leq i \leq N, \]
(5)
and
\[ \epsilon + \sum_{j=1}^{n+1} \alpha_j B_{ij} \geq f_i, \quad 1 \leq i \leq N. \]
(6)
But this is the form of a standard linear programming problem, i.e., given \( \mathbf{b} \in R^{n+2} \), \( A \) a real \( 2N \times (n + 2) \) matrix, and \( \mathbf{c} \in R^2N \), minimize \( (y, \mathbf{b}) \) with respect to
$y \in R^{n+2} \cap K$ subject to the constraint that $Ay \geq c$. This problem has the dual problem of maximizing $(x, e)$ with respect to $x \in R^{2N} \cap K$ subject to the constraint that $x^T A \leq b$, cf. [6].

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

$$A = \begin{bmatrix} -B \\ -B \end{bmatrix}, \quad \text{where} \quad B = [B_{ij}].$$

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 more 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

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

with respect to $(s, t) \in R^{2N} \cap K$ subject to the constraints

$$\sum_{i=1}^{N} B_{ij} (s_i - t_i) \leq 0, \quad 1 \leq i \leq n + 1 \quad \text{and} \quad \sum_{i=1}^{N} (t_i + t_i) \leq 1.$$

We turn now to the choice of the basis functions, $\{B_{ij}\}_{j=1}^{n}$. We first examine the one dimensional case of $X = [0, 1]$. 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 $A$ of $[0, 1]$ of the form, $A : 0 = x_0 < \cdots < x_N < x_{N+1} = 1$ and for each $A \in P$ and each positive integer $d$, $S(A, d)$ denote the set of functions $s(x)$ which are a polynomial of degree $d$ on each subinterval $[x_i, x_{i+1}]$ defined by $A$ 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 $S(A, d)$, we follow [2] and [9] and augment the partition $A : 0 = x_0 < \cdots < x_{N+1} = 1$ with the points $x_{-d} < x_{-d+1} < \cdots < x_0$ and $x_{N+1} < x_{N+1+1} < \cdots < x_{N+1+d}$ to form a new partition

$$\tilde{A} : x_{-d} < \cdots < x_0 < \cdots < x_{N+1} < x < \cdots < x_{N+1+d}.$$
Letting
\[ x_i^d := \begin{cases} x^d, & \text{if } x \geq 0, \\ 0, & \text{if } x < 0, \end{cases} \quad \text{and} \quad W_i(x) = \prod_{k=0}^{d+1} (x - x^{i+k}) \]
for \(-d \leq i \leq N\), we define the “B-splines” of Schoenberg and Curry
\[ M_{d,i}(x; \Delta) = \sum_{k=0}^{d+1} (d + 1) \frac{(x^{i+k} - x)^d}{W_i(x^{i+k})} \]
for \(-d \leq i \leq M\). As a basis for \( S(\Delta, d) \) we take the restriction of the functions \( \{M_{d,i}(x; \Delta)\}_{i=-d}^{N} \) to the interval \([0, 1]\).

If \( Y \) is a finite subset of \([0, 1]\) and \( |Y| \equiv \max_{x \in [0,1]} \min_{y \in Y} |x - y| \), 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 \( \Delta \in \mathbb{P} \) and \( 2d^2 \Delta^{-2} |Y|^2 < 1 \), where \( \Delta \equiv \min_{0 \leq i \leq N} (x_{i+1} - x_i) \), then
\[ \|f - s_Y\|_X \leq [2(1 - 2d^2 \Delta^{-1} |Y|^{-1}) + 1] \|f - s_X\|_X. \]  \hspace{1cm} (7)

**Proof.** By the triangle inequality
\[ \|f - s_Y\|_X \leq \|f - s_X\|_X + \|s_X - s_Y\|_X. \]  \hspace{1cm} (8)
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 \(|t - y| \leq |Y|\) and \(|(s_X - s_Y)(t)| \leq |(s_X - s_Y)(y)| + |Y| \|D(s_X - s_Y)\|_X\).

Hence, using the Markov inequality for polynomial splines, cf. [9],
\[ \|s_X - s_Y\|_X \leq \|s_X - s_Y\|_Y + |Y| \|D(s_X - s_Y)\|_X \]  \hspace{1cm} (9)
and
\[ \|s_X - s_Y\|_X \leq (1 - |Y| \|2d^2 \Delta^{-1}\| s_X - s_Y\|_Y \leq (1 - |Y| \|2d^2 \Delta^{-1}\| \|f - s_X\|_Y + \|f - s_Y\|_Y) \leq (1 - |Y| \|2d^2 \Delta^{-1}\| (2 \|f - s_X\|_X). \]

The required result now follows from the triangle inequality and (7) and (8). \hspace{1cm} \text{Q.E.D.}

If we assume a certain regularity of the function \( f \), then we can bound the right hand side of (7). Using results of deBoor [2], we obtain
COROLLARY 1. Let $2d^2 \Delta^{-1} | Y | < 1$ and $f \in W^{t, \infty}[0, 1]$, $0 \leq t \leq d + 1$, i.e., $D^{t-1}f$ is absolutely continuous and $D^t f \in L^{\infty}[0, 1]$. There exists a positive constant, $K_{d,t}$, such that if $A \in P$ and $2d^2 \Delta^{-1} | Y | < 1$, then

$$
\|f - s_y\|_X \leq [2(1 - 2d^2 \Delta^{-1} | Y |)^{-1} + 1] K_{d,t} \Delta \|D^t f\|_X,
$$

(10)

where $\Delta = \max_{0 \leq i \leq N(x_{i+1} - x_i)}$.

We remark that for $S(\Delta, d)$, $| Y |$ need only be of order $\Delta$, 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 corollary 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.

COROLLARY 2. If $A \in P$, $s(x) \in S(\Delta, d)$, and $2d^2 \Delta^{-1} | Y | < 1$, then

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

(11)

and

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

(12)

We now turn to the multivariate case. Let $\Omega \in \mathbb{R}^M$ be a closed set contained in the unit cube $\times_{i=1}^M [0, 1]_i$ in $\mathbb{R}^M$ 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 $P \equiv \times_{i=1}^M \Delta_i$, $\overline{P} = \max_{1 \leq i \leq M} \{\Delta_i\}$, i.e., $P$ is the minimum distance between two partition points. Furthermore, let $S(d, P) \equiv \times_{i=1}^M S(d, \Delta_i)$, i.e., $S(d, P)$ is the space of multivariate polynomial spline functions of degree $d$ with respect to $P$, $\Omega_p \equiv \{ x \in \Omega \mid \text{the "N-cell" of } P \text{ containing } x \text{ is contained in } \Omega \}$, and $Y_p \equiv \{ y \in Y \mid y \in \Omega_p \}$. Finally, let

$$
| Y | \equiv \max_{x \in \Omega_p, y \in Y_p} \min_{\Gamma(x,y)} \left\{ \int_{\alpha \in \Gamma(x,y)} \|d\alpha\|_{\ell_1} \left| \Gamma(x,y) \right| \right\}
$$

is a piecewise smooth curve all of whose points lie in $\Omega_p$ 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 $\Omega_p$ between $x$ and $y$ is no more than $| Y_p |$.

The following result is a multivariate analog of Theorem 2.
THEOREM 3. If $P \in P_M$ and $2d^2P^{-1} | Y_P | < 1$, then

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

(13)

Proof. $\| f - s_{Y_P} \|_{\alpha_P} \leq \| f - s_{\alpha_P} \|_{\alpha_P} + \| s_{Y_P} - s_{\alpha_P} \|_{\alpha_P}$. Let $t \in \Omega_P$ be such that $| s(t) | = | s_{Y_P}(t) - s_{\alpha_P}(t) | = \| s_{Y_P} - s_{\alpha_P} \|_{\alpha_P}$. There exists a point $y \in Y_P$ such that

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

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

$$
\leq \| s \|_{Y_P} + 2d^2 \sum_{i=1}^{N} \Delta_i^{-1} | s(\alpha_P) | | y_i - t_i |
$$

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

Thus,

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

and the result follows as in Theorem 2. Q.E.D.

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

$$
\| f \|_{W^{t,\infty}(\Omega)} = \max_{|t| \leq t} \| D^t f \|_{L^\infty(\Omega)}
$$

Using the results of [10] 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 $2d^2P^{-1} | Y_P | < 1$, then

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

(14)

Similarly, we can prove the following multivariate analog of Corollary 2 of Theorem 2.

COROLLARY 2. If $P \in P_M$, $s \in S(P, d)$ and $2d^2P^{-1} | Y_P | < 1$, then

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

(15)

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

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

(16)
REFERENCES