Research Report Number 71 - 3
Department of Computer Science
YALE UNIVERSITY
New Haven, Connecticut 06520
April 1971

DISCRETE TCHEBYCHEFF APPROXIMATION

FOR MULTIVARIATE SPLINES

by

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 $\mathbb{R}^M$ 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 $\mathbb{R}^M$. If $Y$ is any closed subset of $X$ and $g$ is a real-valued, continuous function on $Y$, let

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

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

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

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 Neur condition in general so that Reineck type algorithms don't work, and (iii) interpolation type schemes are impossible to define for general domains in $\mathbb{R}^M$, $M \geq 2$. 
The approach studied in this paper is to replace $X$ by an appropriate discrete subset $Y$ and to consider the approximate optimization problem

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

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

$$\mathbb{R}^{n+1} \cap K = \{ \alpha \in \mathbb{R}^{n+1} \mid \alpha_i \geq 0, 1 \leq i \leq n+1 \}$$

and consider

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

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

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

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

QED.

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

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

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

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

and

$$\varepsilon + \sum_{j=1}^{n+1} \alpha_j B_{ij} \geq f_i, \quad 1 \leq i \leq N.$$
But this is the form of a standard linear programming problem, i.e.,
given $b \in \mathbb{R}^{n+2}$, $A$ a real $2N \times (n+2)$ matrix, and $c \in \mathbb{R}^{2N}$, minimize
$(y, b)$ with respect to $y \in \mathbb{R}^{n+2} \cap K$ subject to the constraint
that $A y \geq c$. This problem has the dual problem of maximizing $(x, c)$
with respect to $x \in \mathbb{R}^{2N} \cap K$ subject to the constraint that
\[
x^T A \leq b, \text{ cf. [6].}
\]

In this case, $b = (0, \ldots, 0, 1), c = (c, a_1, \ldots, a_{n+1})$,
\[
c = (-f_1, \ldots, -f_N, f_1, \ldots, f_N),
\]
and
\[
A = \begin{bmatrix}
\vdots & \begin{array}{c}
- B \\
\end{array} \\
\vdots & \begin{array}{c}
\end{array} \\
\vdots & \begin{array}{c}
B \\
\end{array}
\end{bmatrix}
\]
where $B = [B_{ij}]$. 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
\[ \sum_{i=1}^{N} \left( s_i f_i + t_i (-f_i) \right) = \sum_{i=1}^{N} f_i (s_i - t_i) \] with respect to
\[ (s, t) \in \mathbb{R}^{2N} \cap \mathcal{K} \] 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 \( \mathcal{X} = [0, 1] \). 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, \( \Delta \), 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 \( d \), \( S(\Delta, d) \) denote the set of functions \( s(x) \) which are a polynomial of degree \( d \) on each sub-interval \( [x_i, x_{i+1}] \) defined by \( \Delta \) 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 \( C^{d} \), \( 0 \leq z_i \leq d-1 \), 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 $\widetilde{\Delta} : x_{-d} < \cdots < x_0 < \cdots < x_{N+1} < x_{N+1+d}$.

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

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

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 $s_\Delta$ of the continuous optimization problem. The proof uses a technique developed in [5] for the case of polynomial basis functions.
Theorem 2. If \( \Delta \in \mathcal{P} \) and \( 2d \Delta^{-1} |Y| < 1 \), where \( \Delta = \min_{0 \leq i < N} (x_{i+1} - x_i) \), then
\[
||f - s_Y||_X \leq \left[ 2 (1 - 2d \Delta^{-1} |Y|)^{-1} + 1 \right] ||f - s_X||_X.
\]

Proof. By the triangle inequality
\[
||f - s_Y||_X \leq ||f - s_X||_X + ||s_X - s_Y||_X.
\]

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 \( |t - y| \leq |Y| \) and
\[
|(s_X - s_Y)(t)| \leq |(s_X - s_Y)(y)| + |Y| ||s_X - s_Y||_X.
\]

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

and
\[
||s_X - s_Y||_X \leq (1 - |Y| 2d \Delta^{-1})^{-1} ||s_X - s_Y||_Y
\]
\[
\leq (1 - |Y| 2d \Delta^{-1})^{-1} (||f - s_X||_Y + ||f - s_Y||_Y)
\]
\[
\leq (1 - |Y| 2d \Delta^{-1})^{-1} (2 ||f - s_X||_X). \text{ The required result}
\]

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

**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 \( \Delta \in \mathcal{P} \) and \( 2d^2 \Delta^{-1} |Y| < 1 \) then

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

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

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

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

and

$$0 \leq \|s\|_Y - \|s\|_Y \leq \left(1 - 2d^2 \Delta^{-1} |Y| \right)^{-1} \|s\|_Y \leq (2d^2 \Delta^{-1} |Y|) \left(1 - 2d^2 \Delta^{-1} |Y| \right)^{-1} \|s\|_Y.$$

We now turn to the multivariate case. Let $\Omega \in \mathbb{R}^M$ be a closed set contained in the unit cube $\prod_{i=1}^M [0,1]_i$ in $\mathbb{R}^M$ and for each $1 \leq i \leq N$ let $\Delta_i : 0 = x_0 < x_1 < x_2 < \cdots < x_{N_i} < x_{N_i + 1} = 1$

be a partition of $[0,1]_i$. Let $P_M$ denote the set of all partitions, $P$, of the cube of the form $P = \prod_{i=1}^M \Delta_i$, $P = \max_{1 \leq i \leq M} \{\Delta_i\}$, and $P = \min_{1 \leq i \leq M} \{\Delta_i\}$, i.e., $P$ is the minimum distance between two partition points. Furthermore, let $S(d,P) = \prod_{i=1}^N S(d,\Delta_i)$, i.e., $S(d,P)$ is the space of multivariate polynomial spline functions of degree $d$ with respect to $P$, $\Omega_P = \{x \in \Omega \mid$ the "N" - cell of $P$ containing $x$ is contained in $\Omega\}$, and $Y_P = \{y \in Y \mid y \in \Omega_P\}$. 
Finally, let \(|Y_p| \equiv \max_{x \in \Omega_p} \min_{y \in Y_p} \inf_{\alpha \in \Gamma(x,y)} \|\alpha\|_{\ell_1} \|\Gamma(x,y)\|_{\ell_1}\) 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 \(\ell_1\)-distance in \(\Omega_p\) 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, \(2d^2 \frac{P^{-1}}{|Y_p|} < 1\), then

\[
||f - s_{Y_p}||_{\Omega_p} \leq \left[ 2 \left(1 - 2d^2 \frac{P^{-1}}{|Y_p|} \right)^{-1} + 1 \right] ||f - s_{\hat{\Omega}_p}||_{\Omega_p}.
\]

**Proof.**

\[
||f - s_{Y_p}||_{\Omega_p} \leq ||f - s_{\hat{\Omega}_p}||_{\Omega_p} + ||s_{Y_p} - s_{\hat{\Omega}_p}||_{\Omega_p}.
\]

Let \(t \in \Omega_p\) be such that \(|s(t)| = |s_{Y_p}(t) - s_{\hat{\Omega}_p}(t)|\)

\[
= ||s_{Y_p} - s_{\hat{\Omega}_p}||_{\Omega_p}.
\]

There exists a point \(y \in Y_p\) such that

\[
|s(t)| \leq |s(y)| + \sum_{i=1}^{N} D_1 s(x_i) \left| y_i - t_i \right|
\]

\[
\leq \|s\|_{\hat{\Omega}_p} + \sum_{i=1}^{N} \|D_1 s\|_{\Omega_p} \left| y_i - t_i \right|
\]

\[
\leq \|s\|_{\hat{\Omega}_p} + \frac{N}{\Delta^{-1}} \|s\|_{\Omega_p} \left| y_i - t_i \right|
\]
\[ \left\| s_{Y_P} - s_{\Omega_P} \right\|_{\Omega_P} \leq (1 - |Y_P| 2d^2 P^{-1})^{-1} \left\| s_{Y_P} - s_{\Omega_P} \right\|_{Y_P}, \]

Thus, 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 \( \Omega \) with respect to the norm

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

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 \( 2d^2 P^{-1} |Y_P| < 1 \), then

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

Similarly, we can prove the following multivariate analogue of

Corollary 2 of Theorem 2.
Corollary 2. If $P \in P_{M'}$, $s \in S(P, d)$ and $2d^2 \frac{1}{P} < 1$, then

\begin{align}
(15) \quad \| s \|_{Y_P} & \leq \| s \|_{\Omega_P} \leq (1 - |Y| \cdot 2d^2 \frac{1}{P})^{-1} \| s \|_{Y_P}, \quad \text{and} \\
(16) \quad 0 & \leq \| s \|_{\Omega_P} - \| s \|_{Y_P} \leq [(1 - |Y| \cdot 2d^2 \frac{1}{P})^{-1} - 1] \| s \|_{Y_P}
\end{align}

\begin{align*}
\leq (2d^2 \cdot \frac{1}{P}) \cdot (1 - |Y| \cdot 2d^2 \frac{1}{P})^{-1} \| s \|_{Y_P}.
\end{align*}
References


