On the Invariance of the Interpolation Points of the Discrete ℓ_1-approximation

by

Uri Ascher

Technical Report 77-1

February 1977

Department of Computer Science
University of British Columbia
Vancouver, B. C.

On the invariance of the interpolation points of the discrete ℓ_1-approximation

Uri Ascher
Department of Computer Science
University of British Columbia
Vancouver, B. C., Canada

Abstract

Consider discrete ℓ_1-approximations to a data function f, on some finite set of points X, by functions from a linear space of dimension $m < \infty$. It is known that there always exists a best approximation which interpolates f on a subset of m points of X. This does not generally hold for the "continuous" L_1-approximation on an interval, as we show by means of an example.

We investigate the invariance of the interpolation points of the discrete ℓ_1-approximation under a change in the approximated function. Conditions are given, under which the interpolant to a function g on a set of "best ℓ_1 points" of a function f is a best ℓ_1-approximant to g. Additional results are then obtained for the particular case of spline ℓ_1-approximation.

1. Introduction

One of the properties of the discrete linear ℓ_1-approximation to a function f over some finite set of points X, is that there always exists a best approximation which could be determined as an interpolant to f on some subset of X.

Specifically, let $X = \{x_1, \ldots, x_n\}$ and let W be a set of associated positive weights, $W = \{w_1, \ldots, w_n\}$. Let the functions $\phi_1, \ldots, \phi_m$ be linearly independent over X, $m < n$, and consider approximations to f(x) of the form

$$(1.1) \qquad v(\alpha;x) := \sum_{i=1}^{m} \alpha_i \phi_i(x) \; ; \quad \alpha = (\alpha_1, \ldots, \alpha_m)^T.$$

The ℓ_1-approximation problem is to find an $\alpha = \alpha^*$ which solves the minimization problem

$$(1.2) \qquad \min_{\alpha}\{ \frac{1}{n} \sum_{j=1}^{n} w_j \; |v(\alpha;x_j) - f(x_j)|\} = \frac{1}{n} \sum_{j=1}^{n} w_j \; |v(\alpha^*;x_j) - f(x_j)|.$$

Let

$$(1.3) \qquad Z(g;\beta) := \{x \in X \mid v(\beta;x) = g(x)\}; \qquad \beta = (\beta_1, \ldots, \beta_m)^T.$$

The following theorem may be found in Barrodale and Roberts [2]. It can also be obtained constructively from the dual linear programming formulation of (1.2) (see, e.g. [6]).

Theorem 1.1 [2]

For a given data function f, there exists a best ℓ_1-approximation $v(\alpha^*;\cdot)$ to $f(\cdot)$ on X, such that there exist m points in X

$$\xi_1, \ldots, \xi_m \in X$$

which satisfy

$$(1.4) \qquad \xi_1, \ldots, \xi_m \in Z(f;\alpha^*)$$

and

$$(1.5) \qquad \det \begin{bmatrix} \phi_1, & \ldots, & \phi_m \\ \xi_1, & \ldots, & \xi_m \end{bmatrix} := \det(\phi_i(\xi_j)) \neq 0.$$

The subset of interpolation points $\xi_1, \ldots, \xi_m$ depends generally on f and X and is not usually known in advance. Our purpose in this note is to determine criteria for the interpolation points to remain invariant under a change in f; i.e., to define a class of functions for which the interpolation points $\xi_1, \ldots, \xi_m$ are "ℓ_1-best". Thus, once the points are known for a specific class, (say, by carrying out the linear programming computation of (1.2) for one function in the class) the problem of ℓ_1-approximation for other functions in that class is reduced to that of interpolation of order m.

Our general theorem appears in section 2. It gives conditions for the case where a set of "ℓ_1-best" interpolation points for one function is also "ℓ_1- best" for another function. In section 3 we recall corresponding results for the "continuous" L_1 -approximation on an interval I. It is well known that in the polynomial case, $\phi_i(x) := x^{i-1}$, i=1, ..., m, interpolation at the zeros of the m-th order Chebyshev polynomial of the second kind (transformed from $[-1,1]$ to I) will provide the unique best L_1 -approximation for any function in C^m whose m-th derivative does not vanish on the interval. This was generalized in Micchelli [4] to weak Chebyshev systems. By comparison, in the corresponding ℓ_1 -approximation there is no uniqueness and the Hobby-Rice theorem [3] does not hold; on the other hand, theorem 1.1 does not extend to the continuous L_1 -approximation in such generality. An example is given to prove this last point.

In section 4 we consider the case where $X \subset I$ and arrive at a discrete analogue to Micchelli's result. Finally, we treat the case of spline ℓ_1 -approximation and show, that the unique set of "ℓ_1-best" interpolation points, obtained from the ℓ_1-approximation of a certain perfect spline, provide a best ℓ_1-approximation for every function in the corresponding convexity cone.

The conditions given in section 2 for the invariance of the "best ℓ_1 points" under a change in the approximated function may at times prove to be quite restrictive, especially when X represents a discretization of some connected domain in R^k, $k > 1$. Nevertheless, it has been noted in practical calculations with cross products of B-splines that the interpolation points $\xi_1, \ldots, \xi_m$, determined by best ℓ_1-approximation to a function f, were also "good", though not "best", for other functions tested which did not satisfy the invariance conditions. That is, for another function g, the error when using $\xi_1, \ldots, \xi_m$ to determine the approximant by interpolation, was of the same order of magnitude as the error obtained for the best ℓ_1-approximation to g. This observation has instigated motivation to use the ℓ_1 -points as collocation points in the numerical solution of partial differential equations [8], [1].

2. Invariance of ℓ_1 -interpolation points

Before stating and proving our theorem we recall the following characterization theorem for best ℓ_1- approximations (see, e.g.,[7]). With the notation

(2.1) $$\operatorname{sgn}\{x\} := \begin{cases} 1 & x>0 \\ 0 & x=0 \\ -1 & x<0 \end{cases}$$

we have

Theorem 2.1

$v(\alpha^*;\cdot)$ is a best ℓ_1-approximation to $f(\cdot)$ if and only if

(2.2) $$\left| \sum_{j=1}^{n} w_j v(\alpha;x_j) \operatorname{sgn}\{v(\alpha^*;x_j)-f(x_j)\} \right| \le \sum_{x_k \varepsilon Z(f;\alpha^*)} w_k \left| v(\alpha;x_k) \right|$$

for all $\alpha \varepsilon R^m$.

Our theorem follows.

Theorem 2.2

Let f and g be two given data functions on X. Let α^* and $\xi_1, \ldots, \xi_m$ be so constructed so that $v(\alpha^*;\cdot)$ is a best ℓ_1-approximation to f on X and (1.4) - (1.5) hold. Let $\tilde{\alpha}$ be determined so that $v(\tilde{\alpha};\cdot)$ interpolates $g(\cdot)$ at $\xi_1, \ldots, \xi_m$. If

(i) $Z(g;\tilde{\alpha}) \supset Z(f;\alpha^*)$

(ii) $\exists\, \sigma \varepsilon \{-1,1\}$ such that for any j, $1 \le j \le n$, either

$$\operatorname{sgn}\left\{\det \begin{bmatrix} \phi_1, & \ldots, & \phi_m, & g \\ \xi_1, & \ldots, & \xi_m, & x_j \end{bmatrix}\right\} = \sigma \operatorname{sgn}\left\{\det \begin{bmatrix} \phi_1, \ldots, \phi_m, f \\ \xi_1, \ldots, \xi_m, x_j \end{bmatrix}\right\}$$

or

$$\det\begin{bmatrix}\phi_1, & \dots, & \phi_m, & g\\ \xi_1, & \dots, & \xi_m, & x_j\end{bmatrix} = 0,$$

then $v(\tilde{\alpha};\cdot)$ is a best ℓ_1-approximation to g.

Proof

The characterization (2.2) holds for f with α^*. We want to show that it holds for g with $\tilde{\alpha}$.

Define a function $\hat{f}$ on X by

$$\text{(2.3)} \qquad \hat{f}(x_j) := \begin{cases} v(\alpha^*;x_j) & x_j \in Z(g;\tilde{\alpha}) \\ f(x_j) & \text{otherwise} \end{cases}$$

Then, by assumption (i),

$$Z(\hat{f};\alpha^*) = Z(g;\tilde{\alpha}).$$

We claim that (2.2) holds with $\hat{f}$ replacing f. To show this we need consider only x_j's which satisfy

$$x_j \in Z(g;\tilde{\alpha}) - Z(f;\alpha^*).$$

For each such x_j and any $\alpha \in R^m$, the term $w_j|v(\alpha;x_j)|$ is added to the right hand side of (2.2) and the term $w_j v(\alpha;x_j)$ or $-w_j v(\alpha;x_j)$ is eliminated from the left hand sum. Thus, since the inequality (2.2) holds for f, it must also hold for $\hat{f}$:

$$(2.4)\qquad \left|\sum_{j=1}^{n} w_j v(\alpha;x_j)\,\mathrm{sgn}\{v(\alpha^*;x_j)-\hat{f}(x_j)\}\right| \le \sum_{x_k \varepsilon Z(g;\tilde{\alpha})} w_k |v(\alpha;x_k)|$$

for all $\alpha \ \varepsilon R^m$.

Now, $v(\tilde{\alpha};\cdot)$ interpolates $g(\cdot)$ at exactly the same points as $v(\alpha^*;\cdot)$ interpolates $\hat{f}(\cdot)$, and

$$(2.5)\qquad \mathrm{sgn}\{\det\begin{bmatrix}\phi_1, & \ldots, & \phi_m, & \hat{f}\\ \xi_1, & \ldots, & \xi_m, & x_j\end{bmatrix}\} = \sigma\,\mathrm{sgn}\{\det\begin{bmatrix}\phi_1, & \ldots, & \phi_m, & g\\ \xi_1, & \ldots, & \xi_m, & x_j\end{bmatrix}\}$$

$j = 1, \ldots, n.$

But the errors of interpolation can be written as

$$v(\tilde{\alpha};x_j)-g(x_j) = \frac{-\det\begin{bmatrix}\phi_1, & \ldots, & \phi_m, & g\\ \xi_1, & \ldots, & \xi_m, & x_j\end{bmatrix}}{\det\begin{bmatrix}\phi_1, & \ldots, & \phi_m\\ \xi_1, & \ldots, & \xi_m\end{bmatrix}} \qquad 1 \le j \le n$$

$$v(\alpha^*;x_j)-\hat{f}(x_j) = \frac{-\det\begin{bmatrix}\phi_1, & \ldots, & \phi_m, & \hat{f}\\ \xi_1, & \ldots, & \xi_m, & x_j\end{bmatrix}}{\det\begin{bmatrix}\phi_1, & \ldots, & \phi_m\\ \xi_1, & \ldots, & \xi_m\end{bmatrix}}$$

The determinant in the two denominators is the same (and is nonzero), and (2.5) now yields that

$$(2.6)\qquad \mathrm{sgn}\{v(\tilde{\alpha};x_j)-g(x_j)\} = \sigma\,\mathrm{sgn}\{v(\alpha^*;x_j)-\hat{f}(x_j)\} \quad j = 1,\ldots,n.$$

Thus we obtain, inserting (2.6) into (2.4),

$$\left|\sum_{j=1}^{n} w_j v(\alpha;x_j)\,\mathrm{sgn}\{v(\tilde{\alpha};x_j)-g(x_j)\}\right| \le \sum_{x_k \varepsilon Z(g;\tilde{\alpha})} w_k |v(\alpha;x_k)|$$

for all $\alpha \varepsilon R^m$

and by theorem 2.1, this proves the desired conclusion.

Q. E. D.

3. The continuous L_1-approximation

For purpose of comparison we now consider the case for L_1-approximation on an interval I := [0,1], say. Let $\phi_1, \ldots, \phi_m$ and f be continuous on I. With a uniform weight function, the problem is to find an $\alpha = \alpha^*$ which solves the minimization problem

$$\text{(3.1)} \qquad \min_{\alpha} \left\{\int_0^1 |v(\alpha;x)-f(x)|dx\right\} = \int_0^1 |v(\alpha^*;x)-f(x)|dx.$$

A characterization for α^* is given by (see, e.g. [7])

$$\text{(3.2)} \qquad \left|\int_0^1 v(\alpha;x)\,\mathrm{sgn}\{v(\alpha^*;x)-f(x)\}dx\right| \le \int_{Z(f;\alpha^*)} |v(\alpha;x)|dx$$

for all $\alpha \varepsilon R^m$

with $Z(f;\alpha^*)$ defined as in (1.3), I replacing X.

A general theorem, relevant here, is due to Hobby and Rice [3]

Theorem 3.1 [3]

For any set of functions $\phi_1, \ldots, \phi_m$, linearly independent in $L_1[0,1]$, there exist points

$$0 = \xi_0 < \xi_1 < \ldots < \xi_r < \xi_{r+1} = 1 \quad , \quad r \le m$$

such that

$$\sum_{j=1}^{r+1} (-1)^j \int_{\xi_{j-1}}^{\xi_j} \phi_i(x)dx = 0 \qquad i = 1, \ldots, m. \tag{3.3}$$

Now, if $\phi_1, \ldots, \phi_m$ and f are such that (i) r = m, (ii) interpolation to f on $\{\xi_i\}_{i=1}^m$ is possible and (iii) the error of interpolation changes sign on $\{\xi_i\}_{i=1}^m$ and only there, then by (3.2) we have a best L_1-approximation. Such a result is proved in [4] for weak Chebyshev systems, and we state it below.

Recall that the set of linearly independent continuous functions $\{\phi_1, \ldots, \phi_m\}$ is called a weak Chebyshev system on (0,1) provided that for any $0 < x_1 < \ldots < x_m < 1$,

$$\det \begin{bmatrix} \phi_1, & \ldots, & \phi_m \\ x_1, & \ldots, & x_m \end{bmatrix} \ge 0. \tag{3.4}$$

The subspace $S = \text{span}\{\phi_1, \ldots, \phi_m\}$ is then called a weak Chebyshev subspace of C[0,1], dimS = m. If the determinants in (3.4) are all strictly positive, then the set is called a Chebyshev system. Also, denote by K_c the class of all continuous functions in the convexity cone of $\{\phi_1, \ldots, \phi_m\}$, i.e., all continuous functions f for which, either with h := f or with h := -f

(3.5) $$\det \begin{bmatrix} \phi_1, & \ldots, & \phi_m, & h \\ x_1, & \ldots, & x_m, & x_{m+1} \end{bmatrix} \geq 0$$

for all $0 < x_1 < \ldots < x_{m+1} < 1$. Finally, let

$$F[x_1, \ldots, x_m] := \{(f(x_1), \ldots, f(x_m)); f\varepsilon K_c\}$$

for every $0 < x_1 < \ldots < x_m < 1$ and let $d[x_1, \ldots, x_m]$ be the dimension of the smallest linear subspace of R^m containing $F[x_1, \ldots, x_m]$.

Theorem 3.2 [4]

Suppose $S = \text{span}\{\phi_1, \ldots, \phi_m\}$ is a weak Chebyshev subspace of dimension m of C[0,1], and for every $0 < x_1 < \ldots < x_m < 1$, $d[x_1, \ldots, x_m] = m$. Then every $f \varepsilon K_c$ has a unique best L_1-approximation by elements of S. Furthermore, we have r = m in (3.3) and the best L_1-approximation $v(\alpha^*;\cdot)$ to $f(\cdot)$ is determined by the condition that it interpolates f at $\xi_1, \ldots, \xi_m$.

Note that $\xi_1, \ldots, \xi_m$ do not depend on f. When passing to the discrete ℓ_1-approximation we do not have uniqueness, and the corresponding version of (3.3) does not hold any more (i.e., the left hand side of (2.2) cannot usually be made equal to 0). Nevertheless we obtain, in the next section, corresponding results about invariance of the interpolation points, using theorem 2.2. On the other hand, we show now by means of an example, that theorem 1.1 cannot be stated in such generality for the continuous L_1-approximation.

Example

Let $\phi_i(x) := x^{2i}$, $i = 1, \ldots, m$ and $f(x) := x^{2m+1}$ be defined on $I := [-1,1]$. Then $\phi_1, \ldots, \phi_m$ are linearly independent over I. It is clearly seen from (3.2) that a best L_1-approximation is provided here by $\alpha^* \equiv 0$. Now, let $\beta = (\beta_1, \ldots, \beta_m)^T$ provide another best L_1-approximation to f. Then, for each $x \in I$ (see [7]),

$$[v(\beta;x)-f(x)][v(\alpha^*;x)-f(x)] \geq 0.$$

Therefore, we must have

$$(3.6) \qquad \begin{aligned} v(\beta;x) &\leq f(x) \qquad & x \in (0,1] \\ v(\beta;x) &\geq f(x) \qquad & x \in [-1,0). \end{aligned}$$

Assume, without loss of generality, that $v(\beta;x) \geq 0$ for x in some neighborhood of 0 (note that $v(\beta;x)$ is symmetric around $x = 0$). Then, if $\beta \not\equiv 0$, we get that there exists $\varepsilon > 0$ such that

$$v(\beta;x) > 0 \qquad x \in (-\varepsilon,\varepsilon) - \{0\} .$$

But, by the choice of f we then have that there exists $\delta > 0$ such that

$$v(\beta;x) > f(x) \qquad x \in (-\delta,\delta) - \{0\} .$$

This contradicts (3.6); hence $\alpha^* \equiv 0$ provides the unique best L_1-approximation here. Now, $v(\alpha^*;\cdot) \equiv 0$ interpolates $f(\cdot)$ at only one point, $\xi_1 = 0$, for any positive integer m.

4. Discrete ℓ_1-approximation in one dimension

We restrict ourselves here to $X \subset I$ and use theorem 2.2 to obtain results analogous to part of theorem 3.2 for the discrete ℓ_1-approximation.

Let

(4.1)
$$A = \begin{pmatrix} \phi_1(x_1) & \cdots\cdots & \phi_1(x_n) \\ \vdots & & \vdots \\ \phi_m(x_1) & \cdots\cdots & \phi_m(x_n) \end{pmatrix} .$$

We say that the set $\{\phi_1, \ldots, \phi_m\}$ forms a weak Chebyshev system on X if rank(A)=m and every m by m submatrix of A has a nonnegative determinant. If all m by m determinants are strictly positive then we have a Chebyshev system. A function f, defined on X, is said to belong to the convexity cone of $\{\phi_1, \ldots, \phi_m\}$ if either for $h := f$ or for $h := -f$ we have that for all $x_1 < \ldots < x_{m+1}$, $\{x_i\}_{i=1}^{m+1} \subset X$,

(4.2)
$$\det\begin{bmatrix} \phi_1, & \ldots, & \phi_m, & h \\ x_1, & \ldots, & x_m, & x_{m+1} \end{bmatrix} \geq 0 .$$

We have the following consequence of theorem 2.2.

Corollary 4.1

Let f and g both belong to the convexity cone of the set of m linearly independent functions $\phi_1, \ldots, \phi_m$ on X. With α^* and $\tilde{\alpha}$ defined as in theorem 2.2, assume

$$(4.3) \qquad Z(g;\tilde{\alpha}) \supset Z(f;\alpha^*).$$

Then $v(\tilde{\alpha};\cdot)$ is a best ℓ_1-approximation to g.

Proof

Condition (i) of theorem 2.2 is assumed here. Condition (ii) follows from the definition of the convexity cone. Thus theorem 2.2 is applicable and the conclusion follows.

Q. E. D.

Note that we do not assume above that the functions $\phi_1, \ldots, \phi_m$ form a weak Chebyshev system; only that they are linearly independent on X.

From corollary 4.1 it is clear that if we want to find a set of points $\{\xi_1, \ldots, \xi_m\} \subset X$ which would be invariant for all functions in the convexity cone on X, we have to find a function f in the convexity cone with a minimal set of interpolation points (which always includes $\xi_1, \ldots, \xi_m$). If $\{\phi_1, \ldots, \phi_m, f\}$ is a Chebyshev system on X, then f is such a desired function, since then

$$Z(f;\alpha^*) = \{\xi_1, \ldots, \xi_m\} .$$

But even the requirement that $\phi_1, \ldots, \phi_m$ form a weak Chebyshev system on X does not guarantee the existence of such an f. In particular, for spline functions of order k:

(4.4) $$\phi_i(x) := x^{i-1} \quad i=1,\ldots,k; \quad \phi_{k+i}(x) := (x-\tau_i)_+^{k-1} \quad i=1,\ldots,\nu$$

with $m=k+\nu$ and $0<\tau_1< \ldots <\tau_\nu <1$, where $(x)_+ := \frac{1}{2}(x+|x|)$ and $X \subset I :=[0,1]$, there is no function f such that $\{\phi_1, \ldots, \phi_m, f\}$ is a Chebyshev system if X is dense enough in I. Nevertheless we have for splines

Corollary 4.2

Let f be the perfect spline

(4.5) $$f(x) := x^k + 2 \sum_{i=1}^{\nu} (-1)^i (x-\tau_i)_+^k$$

and let $\xi_1, \ldots, \xi_m$ be obtained as interpolation points of the best discrete ℓ_1-approximation to f by spline functions defined in (4.4), which satisfies (1.4) - (1.5). Then for any function in the convexity cone of $\{\phi_1,\ldots, \phi_m\}$ on X, interpolation on $\xi_1, \ldots, \xi_m$ provides a best spline ℓ_1-approximation.

Proof

Since $f^{(k)}$ changes sign exactly at $\tau_1, \ldots, \tau_\nu$ we have that f belongs to the convexity cone of $\{\phi_1, \ldots, \phi_m\}$ defined by (4.4) (see [4]). Also, since there cannot be more than m interpolation points to this f by any spline $v(\alpha;x) = \sum_{i=1}^{m} \alpha_i\phi_i(x)$ [5], we have that

$$Z(f;\alpha^*) = \{\xi_1, \ldots, \xi_m\} \subset Z(g;\tilde{\alpha})$$

for any g in the convexity cone, and corresponding $\tilde{\alpha}$ which is determined by interpolation on $\xi_1, \ldots, \xi_m$. Hence corollary 4.1 applies here.

Q. E. D.

References

1. U. Ascher and J.B. Rosen,"A Collocation Method for Parabolic Quasilinear Problems on General Domains," in preparation.

2. I. Barrodale and F.D.K. Roberts, "Application of Mathematical Programming to ℓ_p Approximations," Nonlinear Programming, J.B. Rosen, O.L. Mangasarian and K. Ritter (eds.), Academic Press, 447-464 (1970).

3. C.R. Hobby and J.R. Rice, "A Moment Problem in L_1 Approximation," Proc. Amer. Math. Soc., 65; 665-670 (1965).

4. C.A. Micchelli, "Best L^1 Approximation by Weak Chebyshev Systems and the Uniqueness of Interpolating Perfect Splines," to appear in J. Approx. Th.

5. A. Pinkus, private communication.

6. P. Rabinowitz, "Applications of Linear Programming to Numerical Analysis," SIAM Rev. 10 (1968), 121-159.

7. J.R. Rice, The Approximation of Functions, Vol. 1, Addison-Wesley MA (1964).

8. J.B.Rosen, "Solution of Boundary Value Problems by a Weighted L_1 Minimization," Proc. IFIP Congress 1974, North Holland Publ. 456-460 (1974).