A formula for the error in multivariate quasi-interpolation which reproduces the linear polynomials is given. From it sharp pointwise $L_\infty$-bounds for the error in linear interpolation (interpolation by linear polynomials) to function values at the vertices of a simplex are obtained. The corresponding 'envelope theorem' giving the optimal recovery of functions is discussed.
1. Introduction

The high applicability of numerical methods based on interpolation from spaces that contain polynomials, such as the finite element method, has lead to a large literature dealing with the error in such schemes. The main contribution of this paper is an error formula for a multivariate quasi-interpolation operator which reproduces $\Pi_1$ (the linear polynomials). By this we mean an operator defined on some space of functions $\Omega \rightarrow \mathbb{R}$, $\Omega \subset \mathbb{R}^n$ (that contains $\Pi_1$) which is of the form

$$Lf := \sum_{v \in \Theta} f(v) p_v,$$  \hspace{1cm} (1.1)

where $\Theta \subset \mathbb{R}^n$ is a finite set of points, and satisfies

$$Lf = f, \hspace{1cm} \forall f \in \Pi_1.$$  \hspace{1cm} (1.2)

This formula covers the much studied case of linear interpolation at the vertices of a triangle (Courant’s ‘original’ finite element [Co43]).

The paper is set out as follows. In the remainder of this section we define the linear functional $f \mapsto \int_{\Theta} f$ and give some of its relevant properties. This linear functional is proving to be the appropriate notation in which to express the error in multivariate polynomial interpolants (see, e.g., de Boor [B95]). In Section 2, the error formula is given. In Section 3, sharp $L_\infty$-error bounds for (multivariate) linear interpolation are obtained from the formula. These bounds are compared with others in the literature, including recent work of Handscomb [H95] and earlier work of Subbotin [Su90_1, Su90_2].

The linear functional $f \mapsto \int_{\Theta} f$

To describe the error in a multivariate quasi-interpolation operator it is often convenient to use the following linear functional called the divided difference functional on $\mathbb{R}^n$ by Micchelli in [M80].

**Definition 1.3.** For $\Theta$ the sequence $[\theta_0, \ldots, \theta_k]$ of $k + 1$ points in $\mathbb{R}^n$, let

$$f \mapsto \int_{\Theta} f := \int_0^1 \int_0^{s_1} \cdots \int_0^{s_k} f(\theta_0 + s_1 (\theta_1 - \theta_0) + \cdots + s_k (\theta_k - \theta_{k-1})) ds_k \cdots ds_2 \ ds_1,$$

with the convention that $\int_{[1]} f := 0$. Note that for one point

$$\int_{[u]} f = f(u).$$  \hspace{1cm} (1.4)

The value of $\int_{\Theta} f$ does not depend on the ordering of the points in $\Theta$. The nature of $\int_{\Theta} f$ becomes more apparent by observing that

$$\int_{\Theta} f = \frac{1}{k!} \int_{\text{conv} \Theta} M(\cdot | \Theta) f,$$  \hspace{1cm} (1.5)
where \( M(\cdot|\Theta) \) is the *simplex spline* with knots \( \Theta \) (which is supported on \( \text{conv} \, \Theta \), the convex hull of the points in \( \Theta \)). The class of functions for which \( \int_\Theta f \) is defined can be determined from (1.5).

Crucial to the arguments of the paper is the following form of the *fundamental theorem of calculus*, that

\[
\int_{[\Theta,v]} f - \int_{[\Theta,w]} f = \int_{[\Theta,v,w]} D_v-w f. \tag{1.6}
\]

This is a form of the *difference identity* for simplex splines (see [M80:Th.6]). Throughout, \( D_y f \) denotes the derivative of \( f \) in the direction \( y \).

Notice that the *Hermite–Genocchi formula* can be written as

\[
[\theta_0,\ldots,\theta_k] f = \int_{[\theta_0,\ldots,\theta_k]} D^k f, \tag{1.7}
\]

where \([\theta_0,\ldots,\theta_k] f \) is the *univariate* divided difference of \( f \) at the points \( \theta_0,\ldots,\theta_k \) in \( \mathbb{R} \).

## 2. The error formula

In this section we give an error formula for a multivariate quasi–interpolation operator which reproduces \( \Pi_1 \) (the linear polynomials). This formula involves only second order derivatives of the function interpolated, and this form permits one to conclude, e.g., by scaling, that a numerical scheme based on such an operator has order of convergence (accuracy) \( h^2 \) in the simplex size \( h \) (for ‘good’ simplices). The idea of the proof below is to use the ‘difference identity’ (1.6) in just the right way so as to introduce these second order derivatives.

**Theorem 2.1.** Suppose \( \Omega \) is starshaped with respect to \( \Theta \). If \( L \) is a multivariate quasi–interpolation operator, as defined by (1.1), which reproduces the linear polynomials, then \( \forall f \in C^2(\Omega) \)

\[
f(x) - Lf(x) = \sum_{\{v,w\} \subseteq \Theta} p_v(x) p_w(x) \int_{[x,v,w]} D_{v-w} D_{w-v} f
\]

\[
= \frac{1}{2} \sum_{v \in \Theta} \sum_{w \in \Theta} p_v(x) p_w(x) \int_{[x,v,w]} D_{v-w} D_{w-v} f, \quad \forall x \in \Omega, \tag{2.2}
\]

where the first sum is taken over all 2-element subsets of \( \Theta \).

**Proof:** Since \( L \) reproduces the constants, it follows from (1.1) that

\[
\sum_{v \in \Theta} p_v = 1. \tag{2.3}
\]

This, together with (1.4) and the ‘difference identity’ (1.6), gives

\[
f(x) - Lf(x) = \sum_{v \in \Theta} \left( \int_{[x]} f - \int_{[v]} f \right) p_v(x) = \sum_{v \in \Theta} \left( \int_{[x,v]} D_{v-x} f \right) p_v(x). \tag{2.4}
\]
Since $L$ reproduces the linear polynomials, and each coordinate of $(\cdot - v)$ is a linear polynomial, it follows from (1.1) that

$$x - v = \sum_{w \in \Theta} (w - v)p_w(x).$$  \hfill (2.5)

Substituting (2.5) into (2.4), and using the linearity of $y \mapsto D_y$ gives

$$f(x) - Lf(x) = \sum_{v \in \Theta} \sum_{w \in \Theta} p_v(x)p_w(x) \int_{[x,v]} D_{w-v}f.$$  \hfill (2.6)

The double summation in (2.6) is over all ordered pairs $(v, w)$ where $v \neq w$ (the terms for $v = w$ are zero). By summing the pairs $(v, w)$ and $(w, v)$ first, we obtain the following sum over the unordered pairs $\{v, w\}$

$$f(x) - Lf(x) = \sum_{\{v, w\} \in \Theta} p_v(x)p_w(x) \left( \int_{[x,v]} D_{w-v}f - \int_{[x,w]} D_{w-v}f \right).$$  \hfill (2.7)

Finally, by the ‘difference identity’ (1.6) again,

$$\int_{[x,v]} D_{w-v}f - \int_{[x,w]} D_{w-v}f = \int_{[x,v,w]} D_v D_{v-w} D_{w-v}f,$$

which gives the result.

This error formula, once known, can be obtained from a general result of [W97*] (which is more involved), by choice of a particular measure. Also obtainable in this way is the multipoint Taylor formula of Ciarlet and Wagschal [CW71] that

$$f(x) - Lf(x) = -\sum_{v \in \Theta} p_v(x) \int_{[x,v,x]} D^2_{v-x}f, \quad \forall x \in \Omega,$$  \hfill (2.8)

which is the multivariate form of Kowalewski’s remainder (see [K32:p21-24]).

The error formula (2.2) reflects the geometry in a particularly appealing way. The error at any point $x$ not lying on a line connecting points in $\Theta$ is the sum over distinct points $v, w \in \Theta$ of 1/2 the average of the second order derivative $D_{v-w} D_{w-v}f$ over the triangle $\text{conv}\{x, v, w\}$ multiplied by the function $p_v p_w$ (which vanishes at all of the points in $\Theta$ if $L$ matches function values at $\Theta$).

Though we will not consider such examples here, it is worth mentioning that (2.2) holds when $\Theta$ is an infinite sequence of points. More generally, for operators of the form

$$Lf(x) := \int f(v)p_v(x) \, d\mu(v),$$

where $\mu$ is a measure (supported on $\Omega$), there is the ‘continuous’ version that

$$f(x) - Lf(x) = \frac{1}{2} \int \int p_v(x)p_w(x) \int_{[x,v,w]} D_{v-w} D_{w-v}f \, d\mu(v) \, d\mu(w).$$  \hfill (2.9)
The special case of (2.2) when \( L \) is the map of Lagrange interpolation from \( \Pi_1 \) is the first in a family of error formulæ for Chung–Yao interpolation from \( \Pi_k \) (the polynomials of degree \( \leq k \)) recently obtained by de Boor [B95]. In Chung–Yao interpolation, see [CY77] for more details, the points of interpolation are the intersections of certain sets of hyperplanes.

There are many examples of univariate quasi–interpolation operators which reproduce the linear polynomials, e.g., the Lagrange interpolation and Bernstein operators, and hence to which Theorem 2.1 applies. The corresponding error formulæ can be expressed in terms of second order divided differences by using the Hermite–Genocchi formula (1.7). Our primary interest here is in multivariate operators, and so we will not elaborate on these.

3. Sharp pointwise \( L_\infty \)–error bounds for linear interpolation

The main result of this section is a sharp pointwise \( L_\infty \)–error bound for linear interpolation. By linear interpolation we mean interpolation by linear polynomials to function values at \( n+1 \) points in \( \mathbb{R}^n \). These \( n+1 \) points are necessarily affinely independent, i.e., the vertices of a (nondegenerate) simplex in \( \mathbb{R}^n \). This simplex will be denoted

\[
T := \text{conv} \Theta,
\]

its diameter by

\[
h := \text{diam} \Theta = \max_{v, w \in \Theta} \|v - w\|,
\]

and the map of linear interpolation by \( L_\Theta \).

To measure the size of the second order derivative of \( f \in C^2(T) \) we define the function \(|D^2 f| \in C(T)\) by the rule

\[
|D^2 f|(x) := \sup_{\|u_1, u_2 \| \leq 1} |D_{u_1} D_{u_2} f(x)| = \sup_{\|u\| = 1} |D^2 \hat{f}(x)|,
\]

where \( \| \cdot \| \) denotes the Euclidean norm, which satisfies

\[
|D_{u_1} D_{u_2} f| \leq |D^2 f| \|u_1\| \|u_2\|, \quad \forall u_1, u_2 \in \mathbb{R}^n.
\]

The \( L_\infty(T) \)–norm of \( |D^2 f| \) gives a seminorm on \( C^2(T) \)

\[
f \mapsto \|f\|_{2, \infty, T} := \| D^2 f \|_{L_\infty(T)}.
\]

**Theorem 3.4.** Suppose that \( L_\Theta \) is the map of linear interpolation at \( \Theta \). Let \( c \) be the centre and \( R \) the radius of the (unique) sphere containing \( \Theta \). Then, for each \( x \in T \), there is the sharp inequality

\[
|f(x) - L_\Theta f(x)| \leq \frac{1}{2} R^2 - \|x - c\|^2 \|f\|_{2, \infty, T}, \quad \forall f \in C^2(T).
\]
Equality in (3.5) occurs when
\[ f \in Q := \text{span}\{\| \cdot \|^2 \} \oplus \Pi_1, \] (3.6)
and for \( x \) in the interior of \( T \) these are the only functions giving equality in (3.5).
In particular, there is the sharp inequality
\[ \| f - L_\Theta f \|_{L_\infty(T)} \leq \frac{1}{2} (R^2 - d^2) \| f \|_{L_2,\infty,T}, \quad \forall f, \] (3.7)
where \( d \) is the distance of \( c \) from \( T \), i.e.,
\[ d := \text{dist}(c, T) = \min_{x \in T} \| x - c \|. \]

Special cases of (3.7) of interest include the following:
(a) If \( c \in T \), then there is the sharp inequality
\[ \| f - L_\Theta f \|_{L_\infty(T)} \leq \frac{1}{2} R^2 \| f \|_{L_2,\infty,T}, \quad \forall f. \] (3.8)
(b) For the bivariate case \((n = 2)\), if \( c \not\in T \), then there is the sharp inequality
\[ \| f - L_\Theta f \|_{L_\infty(T)} \leq \frac{1}{8} h^2 \| f \|_{L_2,\infty,T}, \quad \forall f. \] (3.9)

The inequalities (3.7), (3.8) and (3.9) are sharp, with equality when \( f \in Q \).

**Proof:** The ‘Lagrange form’ of \( L_\Theta \) will be written as
\[ L_\Theta f = \sum_{v \in \Theta} f(v) \lambda_v, \] (3.10)
since the Lagrange polynomials \( \lambda_v \in \Pi_1, \ v \in \Theta \) are the barycentric coordinate functions with respect to the points in \( \Theta \). Notice that each \( \lambda_v \) is nonnegative on \( T \), and so using (3.2) we obtain
\[ \left| \lambda_v(x) \lambda_w(x) \int_{[x,v,w]} D_{v-w} D_{w-v} f \right| \leq \lambda_v(x) \lambda_w(x) \int_{[x,v,w]} \| v - w \|^2 \| f \|_{L_2,\infty,T} \]
\[ = \frac{1}{2} \lambda_v(x) \lambda_w(x) \| v - w \|^2 \| f \|_{L_2,\infty,T}. \] (3.11)
Therefore, from the first inequality (2.2),
\[ |f(x) - L_\Theta f(x)| \leq \frac{1}{2} \sum_{\{v,w\} \subset \Theta} \lambda_v(x) \lambda_w(x) \| v - w \|^2 \| f \|_{L_2,\infty,T}, \quad \forall f \in C^2(T). \] (3.12)
The next part of the proof relies on the fact that

\[ x = \sum_{v \in \Theta} v \lambda_v(x), \quad 1 = \sum_{v \in \Theta} \lambda_v(x), \quad (3.13) \]

which, in view of (3.10), is the statement that \( L_\Theta \) reproduces \( \Pi_1 \). With \( \langle \cdot, \cdot \rangle \) denoting the Euclidean inner product, the quadratic polynomial (of \( x \)) occurring in (3.12) can be expanded and simplified using (3.13) in the following way.

\[
\frac{1}{2} \sum_{v \neq w \in \Theta} \lambda_v(x) \lambda_w(x) \|v - w\|^2
\]

\[
= \frac{1}{4} \sum_v \sum_w \lambda_v(x) \lambda_w(x) \|v - w\|^2
\]

\[
= \frac{1}{4} \sum_v \sum_w \lambda_v(x) \lambda_w(x) (\|v\|^2 - 2\langle v, w \rangle + \|w\|^2)
\]

\[
= \frac{1}{4} \sum_v \lambda_v(x) \|v\|^2 - \frac{1}{2} \sum_v \lambda_v(x) \langle v, \sum_w w \lambda_w(x) \rangle + \frac{1}{4} \sum_w \lambda_w(x) \|w\|^2
\]

\[
= \frac{1}{2} \left( \sum_v \|v\|^2 \lambda_v(x) - \langle \sum_v v \lambda_v(x), x \rangle \right)
\]

\[
= \frac{1}{2} \left( \sum_v \|v\|^2 \lambda_v(x) - \|x\|^2 \right).
\quad (3.14)
\]

Since

\[ x \mapsto \sum_v \|v\|^2 \lambda_v(x) - \|x\|^2 \quad (3.15) \]

is the unique quadratic polynomial which is zero at the points in \( \Theta \) and has the quadratic part of its Taylor series at the origin equal to \(-\| \cdot \|^2\), it must be equal to

\[ R^2 - \| \cdot - c \|^2. \]

This gives (3.5) with equality for \( f \in Q \).

We now show that for \( x \) in the interior of \( T \) these are the only cases of equality. Suppose, without loss of generality, that \( \|f\|_{L_{\infty}, T} = 2 \), and

\[ f(x) - L_\Theta f(x) = R^2 - \|x - c\|^2. \]

Then the function \( \phi \) defined on \( T \) by

\[ \phi(y) = f(y) - L_\Theta f(y) - R^2 + \|y - c\|^2 \]

satisfies

\[ \phi(v) = 0, \quad v \in \Theta, \quad \phi(x) = 0, \]
and is convex, because

\[
D^2_\xi \phi = D^2_\xi f + 2 \geq -\|f\|_{2,\infty,T} + 2 = 0, \quad \|\xi\| = 1.
\]

Since \(x\) is in the interior of \(T\) (the convex hull of \(\Theta\)), this implies that \(\phi = 0\), i.e., \(f \in Q\).

The sharp inequality

\[
\|f - L\Theta f\|_{L_\infty(T)} \leq \frac{1}{2} \max_{x \in T} (R^2 - \|x - c\|^2) \|f\|_{2,\infty,T}, \quad \forall f,
\]

follows immediately from (3.5), and the constant

\[
\max_{x \in T} (R^2 - \|x - c\|^2) = R^2 - \min_{x \in T} \|x - c\|^2 = R^2 - d^2,
\]

giving (3.7). Finally the special cases.

Case (a). If \(c \in T\), then \(R^2 - d^2 = R^2\).

Case (b). If \(c \notin T\), then \(x^*\) the (unique) choice of \(x \in T\) which minimises \(\|x - c\|\) must lie in some facet \(F\) of \(T\), since when \(x\) is in the interior of \(T\) it may be moved closer to \(c\) (thereby reducing \(\|x - c\|\)). In the bivariate case \((n = 2)\), \(T\) is an obtuse angled triangle with \(F\) its largest side and \(x^*\) is the midpoint of \(F\) (see Fig. 4.1). Since the line segment from \(c\) to \(x^*\) is orthogonal to the facet \(F\) which has length \(h\), Pythagoras’s theorem gives

\[
d^2 + (h/2)^2 = R^2,
\]

and so

\[
\frac{1}{2} (R^2 - d^2) = \frac{1}{8} h^2.
\]

\[\square\]

Fig. 4.1. The situation for an obtuse angled triangle: showing the triangle \(T\) (shaded), the facet \(F\) (thick side), and \(x^*\) the closest point to the center \(c\).
Comparison with the sharp $L_\infty$-bounds of Handscomb and Subbotin

The sharp inequality (3.5) is well-known in the univariate case. The inequalities (3.8) and (3.9) were recently proved by Handscomb [H95] for the bivariate case (when $T$ is a triangle). There the condition $c \in T$ (respectively $c \notin T$) is stated in the equivalent way that the triangle $T$ be acute angled (respectively obtuse angled). The inequality (3.9) does not extend to $n > 2$, since in this case for given $h, R$ there is an interval of possible values for $d$ (depending on the geometry of the points $\Theta$). For example, when $n = 3$ the constant $\frac{1}{2}(R^2 - d^2)$ occurring in (3.7) can as small as $\frac{h^2}{8}$ (exactly two of the points are at a distance $h$ from each other), or as large as $\frac{h^2}{6}$ (exactly three of the points are at a distance $h$ from each other).

In the bivariate case (when $T$ is a triangle)$$\sup\{R^2/h^2 : T \text{ acute angled}\} = 1/3,$$with the supremum attained (only) when $T$ is an equilateral triangle. Thus from (3.8) and (3.9) it follows that for all triangles$$\|f - L_\Theta f\|_{L_\infty(T)} \le \frac{1}{6} h^2 \|f\|_{2, \infty,T}, \quad \forall f,$$(3.16)which is sharp if and only if $T$ is an equilateral triangle. The inequality (3.16) was proved in Subbotin [Su90].

More generally, for $n \geq 1$ it can be shown that$$\sup\{R^2/h^2 : c \in T\} = \frac{n}{2(n+1)},$$with the supremum attained (only) when the points in $\Theta$ are equidistant from each other. In this way one obtains the $n > 2$ analogue of (3.16) that$$\|f - L_\Theta f\|_{L_\infty(T)} \le \frac{1}{4} \frac{n}{n+1} h^2 \|f\|_{2, \infty,T}, \quad \forall f,$$(3.17)which is sharp when the points in $\Theta$ are equal distances from each other. This inequality (3.17) was proved by Subbotin [Su90;Th.1] where the sharpness was demonstrated by considering an appropriate quadratic polynomial $f \in Q$, namely$$f = \frac{1}{2} h^2 \sum_{\{v,w\} \in \Theta} \lambda_v \lambda_w,$$which we recognise as the polynomial given by (3.14), with each occurrence of $\|v - w\|$ replaced by $h$.

**Geometric interpretation of the result, the optimal recovery of functions and envelope theorems**

Suppose that the following information about $f \in C^2(T)$ is known:

$$f(v), \ v \in \Theta \quad (\text{its values at the points } \Theta) \quad (3.18)$$

8
and
\[ |D^2 f| \leq K \quad \text{on } T \quad \text{(i.e., } \|f\|_{2, \infty, T} \leq K). \quad (3.19)\]

then it follows from an observation of Golomb and Weinberger [GW59] that there exist functions \( L, U \) for which
\[ L(x) \leq f(x) \leq U(x), \quad \forall x \in T, \]
and these bounds cannot be improved in the sense that there exists an \( f \) taking any value strictly between them. For obvious reasons, some authors refer to these functions \( L \) and \( U \) that enclose \( f \) as \((lower \text{ and } upper) \text{ envelopes for } f\).

Theorem 3.4 provides the solution of this \textit{optimal recovery problem} of determining \( L \) and \( U \) as follows. Since \( f \) satisfies (3.19), inequality (3.5) gives
\[ |f(x) - L_\Theta f(x)| \leq \frac{1}{2} K (R^2 - \|x - c\|^2), \]
which can be rewritten as
\[ L_\Theta f(x) - \frac{1}{2} K (R^2 - \|x - c\|^2) \leq f(x) \leq L_\Theta f(x) + \frac{1}{2} K (R^2 - \|x - c\|^2), \quad (3.20)\]
which is sharp for those \( f \in Q \) with \( \|f\|_{2, \infty, T} = K \). Hence (3.20) provides the envelopes for \( f \), which we now state as a corollary.

**Corollary 3.21 (Envelope Theorem).** Suppose that \( L_\Theta \) is the map of linear interpolation at \( \Theta \), and let \( c \) be the centre and \( R \) the radius of the (unique) sphere containing \( \Theta \). If the value of \( f \in C^2(T) \) is known at the points \( \Theta \), and \( |D^2 f| \leq K \) on \( T \), then
\[ L(x) \leq f(x) \leq U(x), \quad \forall x \in T, \quad (3.22) \]
where
\[ L(x) := L_\Theta f(x) - \frac{1}{2} K (R^2 - \|x - c\|^2), \quad (3.23) \]
\[ U(x) := L_\Theta f(x) + \frac{1}{2} K (R^2 - \|x - c\|^2), \]
and for any of the values allowed by (3.22), there exists a function taking on that value. In particular, the quadratic functions \( L, U \) match \( f \) at \( \Theta \) and satisfy \( |D^3 L|, |D^3 U| = K \) on \( T \).

Notice that the envelope functions \( L, U \) given by (3.23) can be computed from \( \Theta \) and the values given for \( f(v), v \in \Theta \).

Perhaps the best known ‘envelope theorem’ is the result of Gaffney–Powell [GP76] and Micchelli–Rivlin–Winograd [MRW76] which shows that if the values of a \textit{univariate} function \( f \) is known at \( m + k \) points in \([a, b]\) and \( |D^k f| \leq K \) on \([a, b]\), then \( f \) must lie between two \textit{perfect splines} of degree \( k \). Corollary 3.21 is a \textit{multivariate} generalisation of the case \( k = 2 \) (with \( m = n - 1 \)). Though a trivial generalisation in the sense that the envelope functions are such simple multivariate splines (quadratic polynomials), it
is interesting in view of the lack of such results on the optimal recovery of multivariate functions from such more general information.

The corresponding $L_p$-error bounds

It is possible to obtain $L_p$-error bounds from (2.2) by using inequalities such as the multivariate Hardy inequality of [W97]. For example, with $\|D^2f\|_{L_2(T)} := \|D^2f\|_{L_2(T)}$, it can be shown that for linear interpolation on a triangle

$$\|f - L_0f\|_{L_2(T)} \leq \frac{3}{2} h^2 \|D^2f\|_{L_2(T)}, \quad \forall f \in C^2(T),$$

(3.24)

which improves by a factor of $1/2$ the previous best known bound of Arcangeli and Gout [AG76:Ex. 3–1, p. 17].

References


Acknowledgements. Thanks to the editor for the great care he took with this paper, and to S. Konyagin for a useful discussion about extremals.