Multivariate polynomial interpolation on lower sets

Nira Dyn* and Michael S. Floater†

Abstract

In this paper we study multivariate polynomial interpolation on lower sets of points. A lower set can be expressed as the union of blocks of points. We show that a natural interpolant on a lower set can be expressed as a linear combination of tensor-product interpolants over various intersections of the blocks that define it.

Math Subject Classification: 41A05, 41A10, 65D05
Keywords: multivariate polynomial interpolation, lower sets, tensor-product interpolants

1 Introduction

Lower sets are subsets of Cartesian grids of points in $\mathbb{R}^d$ that admit unique polynomial interpolation from a natural, associated space of polynomials, and include rectangular and triangular grids as special cases. Such interpolation has been studied, for example, in [1, 3, 4, 5, 6, 8, 9, 10, 11, 12, 13]. This natural polynomial space is the span of a collection of monomials with powers taken from a similar lower set of indices. The resulting interpolant is the "least interpolant", introduced in [3], and studied further in [2].

A lower set of points can be expressed as the union of blocks (rectangular grids) of points. The purpose of this paper is (1) to make the observation that

*School of Mathematical Sciences Tel Aviv University, Ramat Aviv, Tel Aviv 69978, Israel, *email: niradyn@math.tau.ac.il*

†Department of Mathematics, University of Oslo, PO Box 1053, Blindern, 0316 Oslo, Norway, *email: michaelf@ifi.uio.no*

the interpolant can be expressed as a linear combination of tensor-product interpolants, each corresponding to an intersection of some of these blocks, and (2) to give an explicit formula for the linear combination, and to show how the formula simplifies in various important cases, including the 2D case, and the total degree case. It was brought to our attention, after submitting the paper, that similar relations were obtained for piecewise polynomial interpolants on sparse grids, for the numerical solution of PDE's. In particular, our formula for the total degree polynomial interpolant is similar to the 'combination technique' for sparse grids (see e.g. [7]).

Here is an outline of the paper. In Section 2 we introduce some notation, and discuss tensor-product polynomial interpolation. In Section 3 we present the interpolation problem on lower sets, and give an explicit expression for the interpolants in terms of a Newton-type formula. An incremental double sum formula for these interpolants, based on blocks, is derived in Section 4, while its simplifications are obtained in Section 5 for the 2D case. For the arbitrary dimension case, a general formula in terms of tensor-product interpolants is derived in Section 6, and again simplified. The simplified formula is applied in Section 7 to interpolation by polynomials of a fixed total degree.

2 Tensor-product interpolation

We consider a Cartesian grid of points $\mathcal{X} \subset \mathbb{R}^d$. For each $j \in \{1, \ldots, d\}$, let $x_{j,k}$, $k \in \mathbb{N}_0$, be a sequence of distinct real points. Then the points of $\mathcal{X}$ are

$$x_\alpha = (x_{1,\alpha_1}, x_{2,\alpha_2}, \ldots, x_{d,\alpha_d}), \qquad \alpha \in \mathbb{N}_0^d,$$

where we use the multi-index notation

$$\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_d) \in \mathbb{N}_0^d,$$

with $|\alpha| := \alpha_1 + \cdots + \alpha_d$, and the convention that $\alpha \leq \beta$ means that $\alpha_j \leq \beta_j$ for $j = 1, \ldots, d$. Given any $\beta \in \mathbb{N}_0^d$, we call the set of indices

$$B_\beta = \{\alpha \in \mathbb{N}_0^d : 0 \leq \alpha \leq \beta\},$$

a *block*, and with it we can associate the set of grid points

$$X_\beta = \{x_\alpha : \alpha \in B_\beta\}.$$

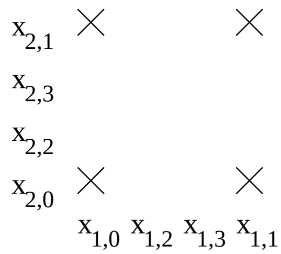

Figure 1: An example of the set $X_{(1,1)}$.

Note that unlike B_β, the points of X_β are not necessarily consecutive in the grid $\mathcal{X}$, as illustrated in Figure 1, showing an example of the set $X_{(1,1)}$ corresponding to the block $B_{(1,1)}$.

If we denote a monomial in $\mathbb{R}^d$ by

$$x^\alpha = x_1^{\alpha_1} \cdots x_d^{\alpha_d},$$

for $\alpha \in \mathbb{N}_0^d$, then

$$P_\beta = \operatorname{span}\{x^\alpha : \alpha \in B_\beta\}$$

is an interpolation space for X_β. In other words, for every function $f : \mathbb{R}^d \to \mathbb{R}$ there is a unique tensor-product polynomial $p \in P_\beta$ such that $p(x_\alpha) = f(x_\alpha)$ for all $\alpha \in B_\beta$.

One way of expressing p is the Newton form; see [8], Chap. 5. We define, for each $j = 1, \ldots, d$, the univariate polynomials $\omega_{j,0}(y) = 1$ and

$$\omega_{j,k}(y) = \prod_{i=0}^{k-1} (y - x_{j,i}), \qquad k \geq 1, \quad y \in \mathbb{R},$$

and the d-variate polynomial

$$\omega_\alpha(x) = \omega_{1,\alpha_1}(x_1) \cdots \omega_{d,\alpha_d}(x_d), \qquad \alpha \in \mathbb{N}_0^d.$$

We further define the divided difference $\Delta_{\alpha,\beta} f$, for $0 \leq \alpha \leq \beta$, by the recursion

$$\Delta_{\alpha,\beta} f := \frac{\Delta_{\alpha+e^j,\beta} f - \Delta_{\alpha,\beta-e^j} f}{x_{j,\beta_j} - x_{j,\alpha_j}}$$

if $\alpha_j < \beta_j$, and otherwise

$$\Delta_{\alpha,\beta} f := f(x_\alpha).$$

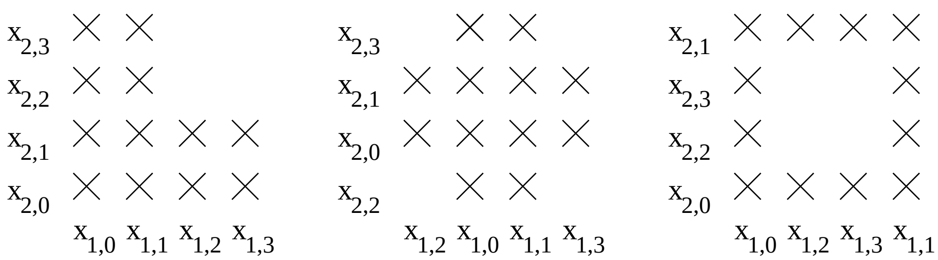

Figure 2: Three configurations of the set Y_L.

Here, $e^j \in \mathbb{N}_0^d$ is the multi-index with $e_i^j = \delta_{ij}$. This kind of divided difference is the usual tensor-product one, i.e., it is applied to the function coordinate-wise. Then we can express p as

$$p(x) = \sum_{\alpha \in B_\beta} \omega_\alpha(x) \Delta_{0,\alpha} f, \qquad x \in \mathbb{R}^d. \tag{1}$$

In later sections it will be convenient to denote the interpolant p by $p(B_\beta)$.

3 Interpolation on lower sets

We call a finite set $L \subset \mathbb{N}_0^d$ a *lower set* if whenever $\mu \in L$ and $0 \le \alpha \le \mu$ then $\alpha \in L$. We call a point $\beta \in L$ a *maximal* point if there is no $\mu \in L$ such that $\mu \ne \beta$ and $\beta \le \mu$. There is a natural space of polynomials associated with L,

$$P_L = \operatorname{span}\{x^\alpha : \alpha \in L\}.$$

The monomials x^α, $\alpha \in L$, form a basis for P_L, and so the dimension of P_L is $|L|$, the cardinality of L.

Corresponding to a lower set L, let Y_L denote the set of points

$$Y_L = \{x_\alpha : \alpha \in L\}.$$

Due to the arbitrary ordering of the coordinates $x_{j,k}$ for each $j = 1, \dots, d$, the set Y_L can take on different configurations. This is illustrated in Figure 2 for the lower set $L = B_{(1,3)} \cup B_{(3,1)}$ where three different configurations of Y_L are shown.

The following theorem has been established in various ways in various papers in various cases. The earliest reference appears to be Kuntzmann [9] for $d = 2$. Here, for the sake of completeness, we prove it in the general case.

Theorem 1 *There is a unique polynomial $p \in P_L$ that interpolates f on Y_L. This polynomial can be expressed as*

$$p(x) = \sum_{\alpha \in L} \omega_\alpha(x) \Delta_{0,\alpha} f, \qquad x \in \mathbb{R}^d. \tag{2}$$

Proof. To see that p interpolates f, suppose $x = x_\mu$ in (2) for some $\mu \in L$. Then, since $\omega_\alpha(x_\mu) = 0$ if $\alpha \not\leq \mu$,

$$p(x_\mu) = \sum_{0 \leq \alpha \leq \mu} \omega_\alpha(x_\mu) \Delta_{0,\alpha} f.$$

But the polynomial

$$q(x) = \sum_{0 \leq \alpha \leq \mu} \omega_\alpha(x) \Delta_{0,\alpha} f,$$

in view of (1), is the tensor-product interpolant to f in P_μ at the points X_μ, and so $p(x_\mu) = q(x_\mu) = f(x_\mu)$.

Having shown that p interpolates f, and since f is arbitrary, it follows from elementary linear algebra that p is also unique. $\square$

In later sections we will sometimes refer to the interpolant p as $p(L)$.

4 Block interpolation

A lower set L can, by definition, be expressed as the union of the blocks associated with its maximal points:

$$L = \bigcup_{\alpha \in V} B_\alpha, \tag{3}$$

where $V \subset L$ is the set of maximal points of L. Conversely, every finite union of blocks is a lower set. The intersection of blocks, on the other hand, is itself a block. In fact, the intersection of the two blocks B_α and B_β is the block B_γ where $\gamma = \min(\alpha, \beta)$, namely $\gamma_j = \min(\alpha_j, \beta_j)$ for $j = 1, \ldots, d$.

In what follows we will show that interpolation on L in the space P_L can be expressed as linear combinations of tensor-product interpolants over the intersections of the blocks B_α, $\alpha \in V$. The main ingredient in this is the following identity.

Lemma 1 *If L and M are lower sets, then*

$$p(L \cup M) = p(L) + p(M) - p(L \cap M). \tag{4}$$

Proof. From the Newton form in Theorem 1,

$$p(L \cup M)(x) = \sum_{\alpha \in L \cup M} \omega_\alpha(x) \Delta_{0,\alpha} f, \qquad x \in \mathbb{R}^d,$$

and so the result follows from the fact that

$$\sum_{\alpha \in L \cup M} = \sum_{\alpha \in L} + \sum_{\alpha \in M} - \sum_{\alpha \in L \cap M}.$$

□

In order to interpolate on a lower set, we can express it as the union of several blocks, and apply recursion, based on the lemma, with M a block, to build up the interpolant incrementally. From now on, when referring to a block we will sometimes write B_i when it is the i-th block in a sequence of blocks. Given blocks B_i, $i = 1, 2, \ldots$, let $L_r = \cup_{i=1}^r B_i$. Then Lemma 1 gives

$$p(L_n) = p(L_{n-1}) + p(B_n) - p(L_{n-1} \cap B_n).$$

This formula leads to a double sum formula

$$p(L_n) = \sum_{i=1}^n p(B_i) - \sum_{i=2}^n p(L_{i-1} \cap B_i). \tag{5}$$

5 Two dimensions

Blocks in $\mathbb{R}^2$ are rectangles. Suppose $B_1, \ldots, B_n$ are n rectangles, such that $B_i = B_{\beta^i}$ with $\beta^i \in \mathbb{N}_0^2$ the multi-index $\beta^i = (\beta_1^i, \beta_2^i)$. Let us further assume that $\beta^i \not\leq \beta^k$ for any $i \neq k$. Then by reordering the rectangles, if necessary, we can assume that

$$0 \leq \beta_1^1 < \beta_1^2 < \cdots < \beta_1^n, \qquad \beta_2^1 > \beta_2^2 > \cdots > \beta_2^n \geq 0. \tag{6}$$

Thus the configuration of the rectangles can be thought of as a kind of staircase shape; see Figure 3, in which $n = 5$ and $\beta^1 = (2, 10)$, $\beta^2 = (3, 8)$, $\beta^3 = (5, 5)$, $\beta^4 = (8, 4)$, $\beta^5 = (9, 2)$.

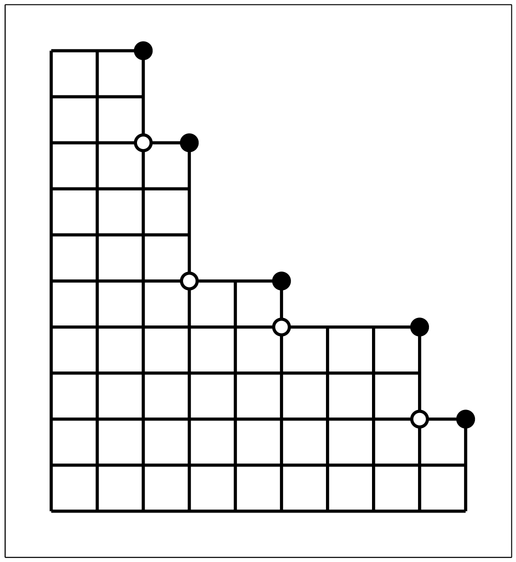

Figure 3: A lower set in $\mathbb{N}_0^2$. The points β^i, $i = 1, \dots, 5$, are shown as black circles, the points $(\beta_1^{i-1}, \beta_2^i)$, $i = 2, \dots, 5$, as white circles.

Theorem 2 *In the staircase configuration the interpolant reduces to*

$$p(L_n) = \sum_{i=1}^{n} p(B_i) - \sum_{i=2}^{n} p(B_{i-1} \cap B_i). \tag{7}$$

Proof. This follows from the double sum formula (5), since by (6),

$$L_{i-1} \cap B_i = \{\alpha : 0 \leq \alpha \leq (\beta_1^{i-1}, \beta_2^i)\} = B_{i-1} \cap B_i.$$

□

6 Arbitrary dimension

Consider now general dimension d. Repeated use of the double sum formula (5) leads to the following general formula.

Theorem 3 *Let L_n be the union of the n blocks $B_1, B_2, \dots, B_n$. Then*

$$p(L_n) = \sum_{k=1}^{n} (-1)^{k-1} \sum_{1 \leq i_1 < i_2 < \cdots < i_k \leq n} p(B_{i_1} \cap B_{i_2} \cap \cdots \cap B_{i_k}). \tag{8}$$

Proof. The proof is by induction on n. The formula trivially holds when $n = 1$. For $n > 1$, since

$$L_{i-1} \cap B_i = (B_1 \cap B_i) \cup \cdots \cup (B_{i-1} \cap B_i), \qquad i = 2, \ldots, n,$$

we conclude that $L_{i-1} \cap B_i$ is the union of $i - 1$ blocks, and $i - 1 \leq n - 1$. Thus we can apply the induction hypothesis to the double sum formula (5) to deduce that

$$p(L_n) = \sum_{i=1}^{n} p(B_i) - \sum_{i=2}^{n} \sum_{k=1}^{i-1} (-1)^{k-1} \sum_{1 \leq i_1 < \cdots < i_k \leq i-1} p(B_{i_1} \cap \cdots \cap B_{i_k} \cap B_i).$$

The second sum above can be rewritten as

$$-\sum_{k=2}^{n} (-1)^{k-1} \sum_{1 \leq i_1 < \cdots < i_k \leq n} p(B_{i_1} \cap \cdots \cap B_{i_k}),$$

which when subtracted from the first sum gives (8). □

With the shorthand $p = p(L_n)$ and

$$p_{i_1,\ldots,i_k} := p(B_{i_1} \cap \cdots \cap B_{i_k}),$$

the formulas in the cases $n = 2, 3, 4$ are

$$\begin{aligned} p &= (p_1 + p_2) - p_{12}, \\ p &= (p_1 + p_2 + p_3) - (p_{12} + p_{13} + p_{23}) + p_{123}, \\ p &= (p_1 + p_2 + p_3 + p_4) - (p_{12} + p_{13} + p_{14} + p_{23} + p_{24} + p_{34}) \\ &\quad + (p_{123} + p_{124} + p_{134} + p_{234}) - p_{1234}. \end{aligned}$$

If $n > d$ some of the block intersections in the sum (8) are over more than d blocks, in which case they reduce to intersections of at most d of the blocks. Thus, in general, many of the terms in the sum (8) are repeated. It follows from Theorem 3 that there must be integer coefficients c_α, $\alpha \in L$, such that

$$p(L) = \sum_{\alpha \in L} c_\alpha p(B_\alpha). \tag{9}$$

We now use Lemma 1 to derive a general formula for the coefficients. For any set $L \subset \mathbb{N}_0^d$, we denote by $\chi(L) : \mathbb{N}_0^d \to \{0, 1\}$ its characteristic function defined by

$$\chi(L)(\alpha) = \begin{cases} 1 & \text{if } \alpha \in L; \\ 0 & \text{otherwise.} \end{cases}$$

Theorem 4 *The coefficients in (9) are given by*

$$c_\alpha = \sum_{\epsilon \in \{0,1\}^d} (-1)^{|\epsilon|} \chi(L)(\alpha + \epsilon), \qquad \alpha \in L. \tag{10}$$

We note that the sum in (9) can be extended to all $\alpha \in \mathbb{N}_0^d$ if necessary because if $\alpha \notin L$ then $c_\alpha = 0$. This follows from the fact that if $\alpha \notin L$ then $\alpha + \epsilon \notin L$ for all $\epsilon \in \{0,1\}^d$ in the sum in (10). This fact will facilitate the proof of the theorem by induction.

Proof. The proof of the formula is by induction on $n \geq 1$, the number of blocks defining L. Suppose first that $n = 1$ and that $L = B_\beta$ for some $\beta \in \mathbb{N}_0^d$. For any $\alpha \in B_\beta$ define $\gamma \in \{0,1\}^d$ by $\gamma_j = 1$ if $\alpha_j < \beta_j$ and $\gamma_j = 0$ if $\alpha_j = \beta_j$, $j = 1, \ldots, d$. Then for $\epsilon \in \{0,1\}^d$, $\alpha + \epsilon \in L$ if and only if $0 \leq \epsilon \leq \gamma$, and so

$$c_\alpha = \sum_{0 \leq \epsilon \leq \gamma} (-1)^{|\epsilon|}.$$

Therefore, if $\alpha = \beta$, $\gamma = 0$, and

$$c_\alpha = (-1)^0 = 1.$$

On the other hand, if $\alpha \neq \beta$, $\gamma \neq 0$, and, letting $k = |\gamma|$,

$$c_\alpha = \sum_{i=0}^{k} (-1)^i \binom{k}{i} = (1-1)^k = 0,$$

which proves (10) for $n = 1$.

Otherwise, suppose that $L = L_n \cup B$, where L_n is the union of n blocks and B is a block. We can assume by the induction hypothesis that the formula holds for the union of at most n blocks. By Lemma 1,

$$p(L) = p(L_n) + p(B) - p(L_n \cap B),$$

and so, by applying (9–10) to the three terms on the right hand side, it follows that (9) holds for L with

$$c_\alpha = \sum_{\epsilon \in \{0,1\}^d} (-1)^{|\epsilon|} (\chi(L_n) + \chi(B) - \chi(L_n \cap B))(\alpha + \epsilon).$$

This reduces to (10) because

$$\chi(L_n) + \chi(B) - \chi(L_n \cap B) = \chi(L).$$

□

Let

$$1_d := (1, 1, \ldots, 1) \in \mathbb{N}_0^d.$$

We call a point $\alpha \in L$ a *boundary* point if $\alpha + 1_d \notin L$, and an *interior* point otherwise.

Corollary 1 *For $\alpha \in L$ an interior point, $c_\alpha = 0$.*

Proof. Because L is a lower set, if $\alpha + 1_d$ belongs to L, so does $\alpha + \epsilon$ for any $\epsilon \in \{0, 1\}^d$ and so (10) implies

$$c_\alpha = \sum_{\epsilon \in \{0,1\}^d} (-1)^{|\epsilon|} = \sum_{i=0}^{d} \binom{d}{i} (-1)^i = 0.$$

□

It follows that a point $\alpha \in L$ with a non-zero coefficient c_α must be both an intersection of some of the blocks defining L and a boundary point of L. This means that in the staircase configuration (6), where $d = 2$, α must belong to either $\{\beta^i : i = 1, \ldots, n\}$ or $\{(\beta_1^{i-1}, \beta_2^i) : i = 2, \ldots, n\}$. Since Theorem 4 gives $c_\alpha = 1$ in the first case, and $c_\alpha = -1$ in the second, it thus recovers Theorem 2.

7 Interpolation of total degree

As an example of interpolation on a lower set, consider polynomial interpolation of total degree $m \geq 0$, so that

$$L = \{\alpha \in \mathbb{N}_0^d : |\alpha| \leq m\}.$$

Then the set of maximal points is

$$V = \{\alpha \in \mathbb{N}_0^d : |\alpha| = m\},$$

and L is the union of the blocks B_α with $|\alpha| = m$.

In the special case $d = 2$, the interpolant is a staircase interpolant (7),

$$p(L) = \sum_{|\alpha|=m} p(B_\alpha) - \sum_{|\alpha|=m-1} p(B_\alpha).$$

For general d, we apply Theorem 4 and Corollary 1. Suppose $\alpha \in L$ with $|\alpha| \le m-d$. Then $|\alpha+1_d| \le m$, and so $\alpha+1_d \in L$ and α is an interior point, and by Corollary 1, $c_\alpha = 0$. Suppose otherwise that $|\alpha| = m-k$ for some $k \in \{0, 1, \ldots, d-1\}$, in which case it is a boundary point. Then Theorem 4 gives

$$c_\alpha = \sum_{\epsilon \in \{0,1\}^d : |\epsilon| \le k} (-1)^{|\epsilon|} = \sum_{i=0}^{k} (-1)^i \binom{d}{i},$$

and so

$$c_\alpha = 1 + \sum_{i=1}^{k} (-1)^i \left(\binom{d-1}{i-1} + \binom{d-1}{i} \right) = (-1)^k \binom{d-1}{k}.$$

So, for example, the interpolant for $d = 3$ is

$$p(L) = \sum_{|\alpha|=m} p(B_\alpha) - 2 \sum_{|\alpha|=m-1} p(B_\alpha) + \sum_{|\alpha|=m-2} p(B_\alpha),$$

where, as usual, an empty sum has value 0.

References

[1] V. Barthelmann, E. Novak, and K. Ritter, High dimensional polynomial interpolation on sparse grids, Adv. Comput. Math. **12** (2000), 273–288.

[2] C. de Boor, On the error in multivariate polynomial interpolation, Appl. Num. Math. **10** (1992), 297–305.

[3] C. de Boor and A. Ron, On multivariate interpolation, Const. Approx. **6** (1990), 287–302.

[4] A. Chkifa, A. Cohen, and C. Schwab, High-dimensional adaptive sparse polynomial interpolation and applications to parametric PDEs, to appear in Found. Comput. Math.

[5] M. Gasca and T. Sauer, On the history of multivariate polynomial interpolation, J. Comp. Appl. Math. **122** (2000), 23–35.

[6] M. Gasca and T. Sauer, Polynomial interpolation in several variables, Adv. Comput. Math. **12** (2000), 377-410.

[7] M. Hegland, The combination technique and some generalisations, Linear Alg. Appl. **420** (2007), 249–275.

[8] E. Isaacson and H. B. Keller, *Analysis of numerical methods*, Dover, 1994.

[9] J. Kuntzmann, *Méthodes numériques: interpolation, dérivées*, Dunod, Paris, 1959.

[10] G. G. Lorentz and R. A. Lorentz, Solvability problems of bivariate interpolation I, Const. Approx. **2** (1986), 153–169.

[11] G. Mühlbach, On multivariate interpolation by generalized polynomials on subsets of grids, Computing **40** (1988), 201–215.

[12] T. Sauer, Lagrange interpolation on subgrids of tensor product grids, Math. Comp. **73**, (2004), 181–190.

[13] H. Werner, Remarks on Newton type multivariate interpolation for subsets of grids, Computing **25** (1980), 181-191.