

On best p-norm approximation of discrete data by polynomials

Michael S. Floater*

September 28, 2023

Abstract

In this note we derive a solution to the problem of finding a polynomial of degree at most n that best approximates data at $n + 2$ points in the l_p norm. Analogous to a result of de la Vallée Poussin, one can express the solution as a convex combination of the Lagrange interpolants over subsets of $n + 1$ points, and the error oscillates in sign.

Keywords: Best approximation, equioscillation, least squares approximation, polynomial interpolation.

1 Introduction

A fundamental theorem in approximation theory, due to Chebyshev, is that the best approximation by a polynomial $P(x)$ of degree at most n to a continuous real function f on a closed interval $[a, b]$ has the equioscillation property: there must exist $n + 2$ points in $[a, b]$ at which the error $f - P$ alternates in sign and equals $\max_{x \in [a, b]} |f(x) - P(x)|$ in magnitude. A similar property holds for approximation to a finite data set: the error in the best approximation of degree at most n to a real function f on a set of m real points, $m \geq n + 2$, is attained at a subset of $n + 2$ of the points, with alternating sign;

*Department of Mathematics, University of Oslo, Moltke Moes vei 45, 0851 Oslo, Norway, email: michaelf@math.uio.no

see for example Powell [11, Chap. 7], Cheney [4, Chap. 3], and Dahlquist and Björck [5, Subsec. 4.5.7].

This equioscillation property led de la Vallée Poussin in [6] to make a detailed study of the special case $m = n + 2$, i.e, the problem of finding the unique polynomial P of degree at most n that minimizes

$$E = \max_{i=0,1,\dots,n+1} |f(x_i) - P(x_i)|, \quad (1)$$

where $x_0 < x_1 < \dots < x_{n+1}$ and f is some real function defined at these points. There must be a constant λ such that

$$f(x_i) - P(x_i) = (-1)^i \lambda, \quad i = 0, 1, \dots, n + 1, \quad (2)$$

and then $E = |\lambda|$, and de la Vallée Poussin looked for the solution in the monomial form

$$P(x) = \sum_{j=0}^n c_j x^j \quad (3)$$

by solving the linear system of $n + 2$ equations

$$\sum_{j=0}^n c_j x_i^j + (-1)^i \lambda = f(x_i), \quad i = 0, 1, \dots, n + 1, \quad (4)$$

in the $n + 2$ unknowns $c_0, c_1, \dots, c_n$ and λ . He showed, using Cramer's rule, that

$$\lambda = (-1)^{n+1} \delta / \sum_{i=0}^{n+1} \frac{1}{|\sigma_i|},$$

where

$$\sigma_i = \prod_{\substack{j=0 \\ j \neq i}}^{n+1} (x_i - x_j)$$

and δ is the divided difference of f ,

$$\delta = [x_0, \dots, x_{n+1}]f = \sum_{i=0}^{n+1} \frac{f(x_i)}{\sigma_i}.$$

He also used the solution to the coefficients c_j to derive a formula for P as a convex combination of Lagrange polynomial interpolants. For $i =$

$0, 1, \dots, n+1$, let P_i denote the polynomial of degree at most n that interpolates f at the points $x_0, \dots, x_{i-1}, x_{i+1}, \dots, x_{n+1}$. Then P is the convex combination

$$P(x) = \sum_{i=0}^{n+1} \frac{P_i(x)}{|\sigma_i|} \bigg/ \sum_{i=0}^{n+1} \frac{1}{|\sigma_i|}. \quad (5)$$

Recently, in [3], it was discovered that the polynomial solution to the weighted least squares problem at an arbitrary number of data ($m \geq n+1$) can also be expressed as a convex combination of interpolants. In the special case of $n+2$ data, the problem is to minimize

$$E_2 = \sum_{i=0}^{n+1} w_i |f(x_i) - P(x_i)|^2,$$

with positive weights w_i , and the theorem of [3] shows that the solution can be expressed as

$$P(x) = \sum_{i=0}^{n+1} \frac{P_i(x)}{w_i |\sigma_i|^2} \bigg/ \sum_{i=0}^{n+1} \frac{1}{w_i |\sigma_i|^2}. \quad (6)$$

The formulas (5) and (6) suggest there might be a similar formula for minimization in other p norms, $p > 1$. The purpose of this paper is to report that there is indeed both a convex combination formula and an oscillation property for the minimization of

$$E_p = \sum_{i=0}^{n+1} w_i |f(x_i) - P(x_i)|^p, \quad (7)$$

for any p , $1 < p < \infty$. The convex combination is as follows:

Theorem 1 *For any p with $1 < p < \infty$, there is a unique polynomial P of degree at most n that minimizes E_p and it is given by the formula*

$$P(x) = \sum_{i=0}^{n+1} \frac{P_i(x)}{w_i^{q-1} |\sigma_i|^q} \bigg/ \sum_{i=0}^{n+1} \frac{1}{w_i^{q-1} |\sigma_i|^q}, \quad (8)$$

where q is such that $1/p + 1/q = 1$.

The oscillation property also has its analogue in the l_p case although the oscillation is no longer equioscillation: the error alternates in sign, but varies in magnitude.

Theorem 2 *The polynomial P of Theorem 1 satisfies the equations*

$$f(x_i) - P(x_i) = (-1)^i \frac{\lambda}{w_i^{q-1} |\sigma_i|^{q-1}}, \quad i = 0, 1, \dots, n+1, \quad (9)$$

where

$$\lambda = (-1)^{n+1} \delta / \sum_{j=0}^{n+1} \frac{1}{w_j^{q-1} |\sigma_j|^q}. \quad (10)$$

Later we discuss the case $p = 1$, where the solution is not always unique, and present some numerical examples.

We note that Motzkin and Walsh in [8] and [9] also studied the minimization of E_p for $p \geq 1$ and found a characterization of P (unique for $p > 1$) in terms of the polynomial $T(x) = x^{n+1} + \dots$ that minimizes $\sum_{i=0}^{n+1} w_i |T(x_i)|^p$, but one would need to compute the zeros of T to find P .

2 Affine combinations of the P_i

We start by considering the error between f and any affine combination of the polynomials $P_0, P_1, \dots, P_{n+1}$.

Lemma 1 *Let P be a polynomial of the form*

$$P(x) = \sum_{i=0}^{n+1} \mu_i P_i(x), \quad (11)$$

where $\mu_i \in \mathbb{R}$ and $\sum_{i=0}^{n+1} \mu_i = 1$. Then

$$f(x_k) - P(x_k) = \delta \mu_k \sigma_k, \quad k = 0, 1, \dots, n+1.$$

Proof. For any $k = 0, 1, \dots, n+1$,

$$\begin{aligned} f(x_k) - P(x_k) &= \sum_{i=0}^{n+1} \mu_i (f(x_k) - P_i(x_k)) \\ &= \mu_k (f(x_k) - P_k(x_k)) + \sum_{\substack{i=0 \\ i \neq k}}^{n+1} \mu_i (f(x_k) - P_i(x_k)) \\ &= \mu_k (f(x_k) - P_k(x_k)). \end{aligned}$$

By the Newton error formula for the interpolation of f by P_k ,

$$f(x) - P_k(x) = (x - x_0) \cdots (x - x_{k-1})(x - x_{k+1}) \cdots (x - x_{n+1}) \\ \times [x_0, \dots, x_{k-1}, x_{k+1}, \dots, x_{n+1}, x]f, \quad x \in \mathbb{R},$$

and letting $x = x_k$ gives $f(x_k) - P_k(x_k) = \sigma_k \delta$, which gives the result. $\square$

Next we show that for a non-zero divided difference δ , any polynomial of degree at most n can be represented as an affine combination of the P_i .

Lemma 2 *If $\delta \neq 0$ then any polynomial P of degree at most n can be represented in the form of (11), where*

$$\mu_i = \frac{f(x_i) - P(x_i)}{\delta \sigma_i}, \quad i = 0, 1, \dots, n+1. \quad (12)$$

Proof. First we show $\sum_i \mu_i = 1$. This follows from the fact that

$$\sum_{i=0}^{n+1} \frac{f(x_i) - P(x_i)}{\sigma_i} = \sum_{i=0}^{n+1} \frac{f(x_i)}{\sigma_i} - \sum_{i=0}^{n+1} \frac{P(x_i)}{\sigma_i} \\ = [x_0, \dots, x_{n+1}]f - [x_0, \dots, x_{n+1}]P = \delta - 0 = \delta.$$

Next let

$$Q(x) := \sum_{i=0}^{n+1} \mu_i P_i(x).$$

By Lemma 1 and (12),

$$f(x_k) - Q(x_k) = \delta \mu_k \sigma_k = f(x_k) - P(x_k), \quad k = 0, 1, \dots, n+1,$$

and so $Q(x_k) = P(x_k)$, $k = 0, 1, \dots, n+1$, and hence $Q = P$. $\square$

3 Proofs of theorems

We now prove Theorem 1.

Proof. Suppose $\delta = 0$. Then the unique interpolant P to f of degree at most $n+1$ at the points $x_0, x_1, \dots, x_{n+1}$ has degree at most n . With this P we have $E_p = 0$ and $P_0 = P_1 = \dots = P_{n+1} = P$ and so the theorem holds in this case.

So we assume from now on that $\delta \neq 0$. Then by Lemma 2, we can look for P in the form of (11) for some μ_i that sum to 1. Then, by Lemma 1,

$$E_p = |\delta|^p \sum_{i=0}^{n+1} w_i |\mu_i|^p |\sigma_i|^p,$$

and the task is to minimize this over $\mu_i \in \mathbb{R}$ subject to the constraint that $\sum_{i=0}^{n+1} \mu_i = 1$. Suppose that one or more of the μ_i are negative. Then

$$1 = \sum_{i=0}^{n+1} \mu_i < \sum_{i=0}^{n+1} |\mu_i|,$$

and defining

$$\hat{\mu}_i := |\mu_i| / \sum_{j=0}^{n+1} |\mu_j|,$$

it follows that $\sum_{i=0}^{n+1} \hat{\mu}_i = 1$, and

$$\hat{E}_p := |\delta|^p \sum_{i=0}^{n+1} w_i |\hat{\mu}_i|^p |\sigma_i|^p = E_p / \left(\sum_{j=0}^{n+1} |\mu_j| \right)^p < E_p.$$

Therefore the minimum of E_p can only be attained by non-negative μ_i , which we assume from now on.

Next recall Hölder's inequality,

$$\sum_{i=0}^{n+1} |a_i b_i| \leq \left(\sum_{i=0}^{n+1} |a_i|^p \right)^{1/p} \left(\sum_{i=0}^{n+1} |b_i|^q \right)^{1/q},$$

with equality if and only if

$$|a_i|^p = c |b_i|^q, \quad i = 0, 1, \dots, n+1,$$

for some constant $c > 0$. Let

$$a_i = w_i^{1/p} \mu_i |\sigma_i|, \quad \text{and} \quad b_i = \frac{1}{w_i^{1/p} |\sigma_i|}.$$

Then

$$\sum_i |a_i b_i| = \sum_i \mu_i = 1,$$

and so

$$E_p = |\delta|^p \sum_i |a_i|^p \geq |\delta|^p / \left(\sum_i |b_i|^q \right)^{p-1},$$

(since $p/q = p - 1$), with equality when

$$w_i \mu_i^p |\sigma_i|^p = \frac{c}{w_i^{q-1} |\sigma_i|^q},$$

or equivalently when

$$\mu_i = \frac{c^{1/p}}{w_i^{q-1} |\sigma_i|^q}.$$

The condition that the μ_i sum to 1 then implies that

$$\mu_i = \frac{1}{w_i^{q-1} |\sigma_i|^q} / \sum_{j=0}^{n+1} \frac{1}{w_j^{q-1} |\sigma_j|^q}. \quad (13)$$

□

We note also from the proof that the minimum value of E_p is

$$E_p = |\delta|^p / \left(\sum_{i=0}^{n+1} \frac{1}{w_i^{q-1} |\sigma_i|^q} \right)^{p-1}.$$

We next prove Theorem 2.

Proof. By Lemma 2 and since $|\sigma_i| = (-1)^{n+1-i} \sigma_i$,

$$f(x_i) - P(x_i) = \delta (-1)^{n+1} (-1)^i \mu_i |\sigma_i|, \quad k = 0, 1, \dots, n+1.$$

Substituting the expression for μ_i in (13) into this, we obtain equation (9) with λ as in (10). □

4 The case $p = 1$

In the case that $p = 1$, the problem is to find a polynomial P of degree $\leq n$ that minimizes E_1 in (7). In this case there may or may not be a unique minimizer, but we can express the set of minimizers precisely in terms of the Lagrange interpolants P_i .

Theorem 3 *Let $J \subset \{0, 1, \dots, n+1\}$ be the set of indices such that*

$$w_j |\sigma_j| = \min_{i=0,1,\dots,n+1} w_i |\sigma_i|, \quad j \in J.$$

Then E_1 is minimized by $P(x) = \sum_{j \in J} \mu_j P_j(x)$ for any $\mu_j \geq 0$, $j \in J$, such that $\sum_{j \in J} \mu_j = 1$.

Proof. As in the proof of Theorem 1, we can look for P in the form of (11) for some μ_i that sum to 1, and by Lemma 1,

$$E_1 = |\delta| \sum_{i=0}^{n+1} w_i |\mu_i| |\sigma_i|.$$

By the same argument as previously, the minimization of E_1 requires non-negative μ_i , and therefore

$$E_1 \geq |\delta| \min_{i=0,1,\dots,n+1} w_i |\sigma_i|,$$

with equality if and only if $\mu_k = 0$ for all $k \notin J$. $\square$

Thus the minimum value of E_1 is

$$E_1 = |\delta| \min_{i=0,1,\dots,n+1} w_i |\sigma_i|,$$

and $P(x_k) = f(x_k)$ for all $k \notin J$, for any minimizer P . However, P is unique only in the case that $|J| = 1$.

5 Numerical examples

Figure 1 shows examples of the minimizing polynomial P for $n = 1, 2, 3, 4$. In each figure there are three curves corresponding to the values $p = 1.1, 2, 11$: solid for $p = 1.1$, dashed for $p = 2$; and dotted for $p = 11$. All the weights are equal to 1. The solid curve corresponding to $p = 1.1$ is close to that of $p = 1$, and as expected, is close to interpolating all but one point, the point x_i with the minimum value of $|\sigma_i|$. The dotted curve corresponding to $p = 11$ is close to that of the case $p = \infty$ in which the $n + 2$ errors are equal.

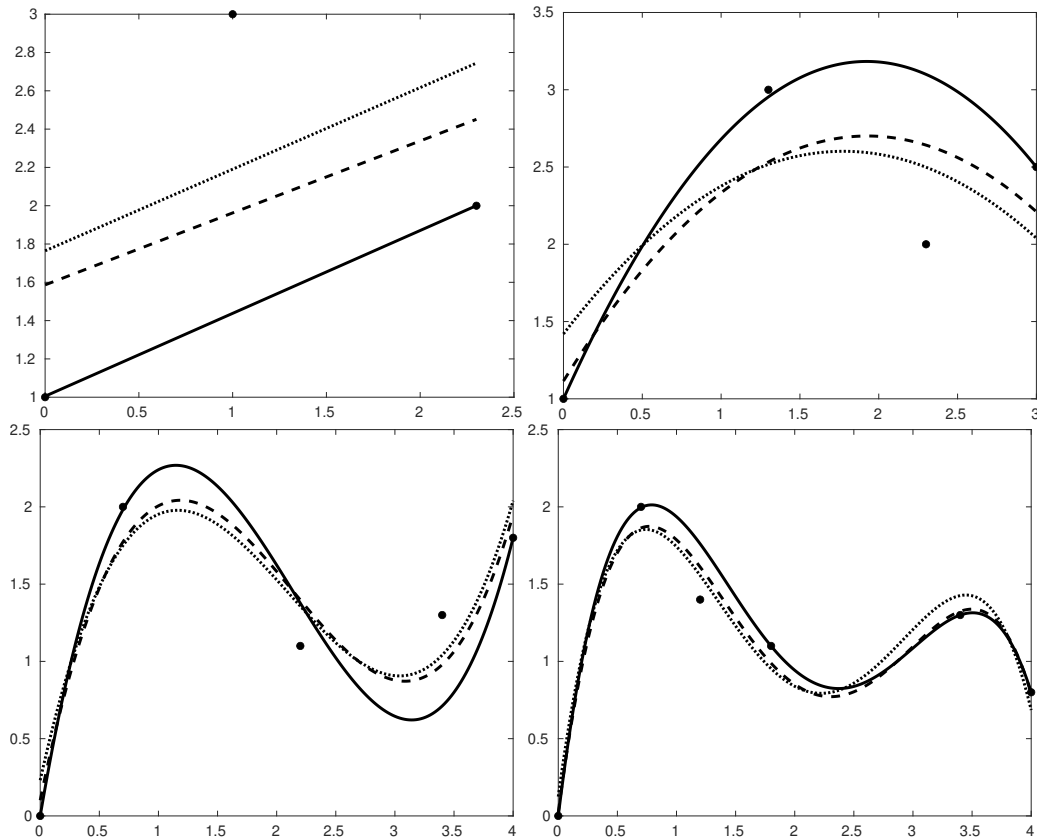


Figure 1: Examples with $n = 1, 2, 3, 4$ and $p = 1.1, 2, 11$.

An efficient way to evaluate P is to recognize it as the unique solution to the $n + 1$ interpolation conditions

$$P(x_i) = f(x_i) - (-1)^i \frac{\lambda}{w_i^{q-1} |\sigma_i|^{q-1}}, \quad i = 0, 1, \dots, n. \quad (14)$$

We can first compute the σ_i and λ in $O(n^2)$ operations, and then use any one of several known methods of representing such an interpolant, computing its representation in $O(n^2)$ operations, and evaluating it at a point x in $O(n)$ operations. For example, if we represent P in the monomial form (3) then the matrix of the system (14) is the Vandermonde matrix $A = [x_i^j]_{i,j=0,1,\dots,n}$. One can then solve this system through the ‘Newton’ LU factorization of A discussed in [10] and [2]. In particular, it is shown in [2] how the entries of L^{-1} and U^{-1} can be computed directly and recursively.

Alternatively, following the approach of [5, Subsec. 4.5.7], one could find the Newton coefficients of P with respect to the nodes $x_0, x_1, \dots, x_n$, also in $O(n^2)$ operations. Another option is to compute the weights in the barycentric form [12, 7] of P as discussed in [1].

Acknowledgement. I wish to thank the referees for their careful reading of the manuscript and their valuable comments which helped to improve it.

References

- [1] J.-P. Berrut and L. N. Trefethen, *Barycentric Lagrange interpolation*, SIAM Review **46** (2004), 501–517.
- [2] J. Carnicer, Y. Khia, and J. M. Peña, *Factorization of Vandermonde matrices and the Newton formula*, Monografías Matemáticas, “García de Galdeano”, Universidad de Zaragoza **40** (2015), 53–60.
- [3] Q. Chang, C. Deng, and M. S. Floater, *An interpolatory view of polynomial least squares approximation*, J. Approx. Theory **252** (2020), Paper No. 105360 (11 pages).
- [4] E. W. Cheney, *Introduction to Approximation Theory*, second edition, AMS Chelsea, 1982.
- [5] G. Dahlquist and Å. Björck, *Numerical Methods in Scientific Computing, Volume 1*, SIAM, 2008.
- [6] Ch.-J. de la Vallée Poussin, *Sur les polynômes d’approximation et la représentation approchée d’un angle*, Académie Royale de Belgique, Bulletins de la Classe des Sciences **12** (1910), 808–844.
- [7] M. Dupuy, *Le calcul numérique des fonctions par l’interpolation barycentrique*, Comptes Rendus Acad. Sci. Paris **226** (1948), 158–159.
- [8] T. S. Motzkin and J. L. Walsh, *Least p th power polynomials on a real finite point set*, Trans. Amer. Math. Soc. **78** (1955), 67–81.
- [9] ———, *Least p th power polynomials on a finite point set*, Trans. Amer. Math. Soc. **83** (1956), 371–396.

- [10] P. J. Olver, *On multivariate interpolation*, Stud. Appl. Math. **2** (2006), 201–240.
- [11] M. J. D. Powell, *Approximation Theory and Methods*, Cambridge University Press, 1981.
- [12] W. J. Taylor, *Method of Lagrangian curvilinear interpolation*, J. Research Nat. Bur. Standards **35** (1945), 151–155.