Department of Mathematics, University of Utah
Analysis of Numerical Methods I
MTH6610 – Section 001 – Fall 2017

Lecture notes – Polynomial interpolation
Wednesday, November 15, 2017

These notes are not a substitute for class attendance. Their main purpose is to provide a lecture overview summarizing the topics covered.

Reading: Isaacson & Keller, Sections 5.1, 5.2

We will be interested in polynomial approximation in one dimension. Approximating by polynomials is often motivated by the following result, the Weierstrass Approximation Theorem: let $f : [a, b] \to \mathbb{R}$ be any continuous function on the compact interval $[a, b]$. Then there exist a sequence of polynomials p_k, each of degree k, such that

$$\lim_{k\to\infty} \sup_{x\in[a,b]} |f(x) - p_k(x)| = 0.$$

This result is heartening, but it turns out that constructing polynomials having the above approximation property can in fact be quite hard.
Perhaps the conceptually easiest way to construct an approximating polynomial is with interpolation: let $x_1, \ldots, x_n$ be n points in $\mathbb{R}$, and define P_k as the space of polynomials of degree k or less,

$$P_k = \operatorname{span}\left\{1, x, \ldots, x^k\right\}.$$

Suppose now we are given a continuous function $f(x)$. The goal is to find an element of P_k that interpolates f at the sites x_j, $j = 1, \ldots, n$. By a counting argument it seems plausible that we can choose $k = n - 1$ and achieve a unique polynomial from P_{n-1} satisfying the interpolation conditions. There are two fairly straightforward ways to show this.
The first way uses linear algebra: We seek to find $p \in P_{n-1}$ satisfying $p(x_j) = f(x_j)$ for $j = 1, \ldots n$. This means

$$p(x) = \sum_{q=1}^{n} c_q x^{q-1}, \qquad\qquad p(x_j) = \sum_{q=1}^{n} c_q x_j^{q-1} = f(x_j),$$

for $j = 1, \ldots, n$. The conditions above are linear in the coefficients c_q, so we collect the coefficients into a vector $\boldsymbol{c}$, which must satisfy the linear system

$$\boldsymbol{V}\boldsymbol{c} = \boldsymbol{f}, \qquad (V)_{j,q} = x_j^{q-1}, \qquad (f)_j = f(x_j). \tag{1}$$

for $1 \le j, q \le n$. Since $\boldsymbol{V}$ is a square matrix, then a unique solution exists if $\boldsymbol{V}$ is invertible. $\boldsymbol{V}$ is called a *Vandermonde* matrix.
Indeed, $\boldsymbol{V}$ is invertible provided all the x_j are distinct points on $\mathbb{R}$. The standard way to show this is to show that the Vandermonde matrix determinant is nonzero. We will prove the following fact:

$$\det \boldsymbol{V} = \prod_{1\le j<q\le n} (x_q - x_j). \tag{2}$$

We prove the above via induction. For $n = 2$, we have

$$\det \boldsymbol{V} = \det \begin{pmatrix} 1 & x_1 \\ 1 & x_2 \end{pmatrix} = x_2 - x_1 = \prod_{1 \le j < q \le 2} (x_q - x_j),$$

showing the initialization step. Now let $n \ge 2$. The inductive hypothesis assumes (2), and we must show this for $n+1$. In this step, we use $\boldsymbol{V}_m$ to denote the $m \times m$ Vandermonde matrix. The determinant in question has the form

$$\det \boldsymbol{V}_{n+1} = \det \begin{pmatrix} 1 & x_1 & x_1^2 & \cdots & x_1^n \\ 1 & x_2 & x_2^2 & \cdots & x_2^n \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_{n+1} & x_{n+1}^2 & \cdots & x_{n+1}^n \end{pmatrix}.$$

We use the Laplace expansion to expand this determinant along the last row:

$$\det \boldsymbol{V}_{n+1} = (-1)^{n+1} \left[\prod_{j=0}^{n} x_{n+1}^j M_j(x_1, \ldots, x_n) \right],$$

where M_j are the cofactors (signed determinants of the minors) formed from eliminating associated rows and columns from $\boldsymbol{V}_{n+1}$. As notationally suggested, all the cofactors are independent of x_{n+1}. Therefore, as a function of x_{n+1}, the determinant $\det \boldsymbol{V}_{n+1}$ is a polynomial of degree n. The roots of this polynomial are x_j, $j = 1, \ldots, n$, since if $x_{n+1} = x_j$, then two rows of the Vandermonde matrix coincide and so the determinant is zero. Therefore, we have shown

$$\det \boldsymbol{V}_{n+1} = (-1)^{n+1} K(x_1, \ldots, x_n) \prod_{j=1}^{n} (x_{n+1} - x_j).$$

The coefficient $K(x_1, \ldots, x_n)$ is the signed determinant of the upper-left block of $\boldsymbol{V}_{n+1}$; this upper-left block equals $\boldsymbol{V}_n$. Therefore:

$$\begin{aligned} \det \boldsymbol{V}_{n+1} = (-1)^{n+1} K(x_1, \ldots, x_n) \prod_{j=1}^{n} (x_{n+1} - x_j) &= (-1)^{n+1} (-1)^{n+1} \det \boldsymbol{V}_n \prod_{j=1}^{n} (x_{n+1} - x_j) \\ = \left(\prod_{1 \le j < q \le n} (x_q - x_j) \right) \prod_{j=1}^{n} (x_{n+1} - x_j) &= \prod_{1 \le j < q \le n+1} (x_q - x_j), \end{aligned}$$

which completes the proof. We have shown that the linear system in (1) always has a unique solution if the x_j are unique. Therefore, minimal-degree polynomial interpolation is *unisolvent* on distinct nodes in one dimension.

A second, perhaps simpler way to show unisolvence of polynomial interpolation is to explictly construct a solution. Assume the x_j are all distinct. Then by inspection one finds that the n polynomials

$$\ell_j(x) \coloneqq \prod_{\substack{1 \le k \le n \\ k \ne j}} \frac{x - x_k}{x_j - x_k}, \qquad 1 \le j \le n,$$

are all polynomials of degree $n-1$ (i.e., elements of P_{n-1}) and satisfy the condition

$$\ell_j(x_k) = \delta_{j,k}, \qquad\qquad 1 \le k \le n.$$

The polynomials ℓ_j are called the (cardinal) *Lagrange* interpolating polynomials. Due to the above property, we have that

$$p(x) = \sum_{j=1}^{n} f(x_j)\ell_j(x),$$

is a degree-$(n-1)$ polynomial satisfying $p(x_k) = f(x_k)$ for $k = 1, \ldots, n$. It is likewise the only such polynomial: if q is any other polynomial in P_{n-1} interpolating f at the x_k, then $p - q$ is a degree-$(n-1)$ polynomial with n roots at $x_1, \ldots, x_n$. Therefore $p - q = 0$.
Of course, nowhere have we discussed how close an interpolating polynomial is to a polynomial to Weierstrass-like approximation. To make this comparison, we require some notation. Given a compact interval $[a, b]$, $a \neq b$ on the real line, define

$$C\left([a,b];\mathbb{R}\right) = \left\{ f : [a,b] \to \mathbb{R} \;\middle|\; f \text{ is continuous everywhere on } [a,b] \right\},$$

endowed with the norm

$$\|f\|_\infty = \sup_{x \in [a,b]} |f(x)|.$$

This norm makes C a Banach space. Given $x_1, \ldots, x_n$, all distinct nodes, we can define an interpolation operator from the procedure above:

$$I_n : C \to P_{n-1}, \qquad\qquad I_n f(x) = \sum_{j=1}^{n} f(x_j)\ell_j(x).$$

To understand the norm of this operator, assume f satisfies $\|f\|_\infty = 1$, so that

$$|I_n f(x)| = \left| \sum_{j=1}^{n} f(x_j)\ell_j(x) \right| \le \sum_{j=1}^{n} |f(x_j)\ell_j(x)| \le \sum_{j=1}^{n} \ell_j(x).$$

By choosing f such that $f(x_j) = \operatorname{sgn} \ell_j(x)$, and $f(x)$ linearly interpolates point in-between, all the above inequalities can become equalities, so that

$$\sup_{\|f\|_\infty = 1} |I_n f(x)| = \sum_{j=1}^{n} \ell_j(x) =: \lambda(x).$$

The function $\lambda(x)$ is called the *Lebesgue function* (associated to the points $x_1, \ldots, x_n$). The norm of the interpolation operator is now easily computed:

$$\|I_n\|_{C \mapsto C} = \sup_{\|f\|_\infty = 1} \|I_n f\|_\infty = \sup_{x \in [a,b]} \lambda(x) =: \Lambda.$$

The number Λ is the called the Lebesgue constant.

All of this is meant to aid in our understanding of interpolation error. For any continuous function f, note that for any polynomial q in P_{n-1}, we have

$$\begin{aligned}|f(x) - I_n f(x)| &\leq |f(x) - q(x)| + |q(x) - I_n f(x)| \\ &= |f(x) - q(x)| + |I_n (q(x) - f(x))| \\ &\leq |f(x) - q(x)| + \lambda(x) \, |q(x) - f(x)| \\ &= [1 + \lambda(x)] \, |f(x) - q(x)| \\ &\leq [1 + \lambda(x)] \, \|f - q\|_\infty,\end{aligned}$$

where the second line uses the fact that I_n is a projection onto P_{n-1} in C. Infimizing the above result over all $q \in P_{n-1}$, we have proven

$$\begin{aligned}|f(x) - I_n f(x)| &\leq [1 + \lambda(x)] \inf_{q \in P_{n-1}} \|f - q\|_\infty \\ \|f(x) - I_n f(x)\|_\infty &\leq [1 + \Lambda] \inf_{q \in P_{n-1}} \|f - q\|_\infty.\end{aligned}$$

The second equality above is called *Lebesgue's Lemma*, and shows that we can bound interpolation error relative to the best approximating polynomial. Lebesgue's Lemma separates error resulting from the choice of interpolation nodes (Λ) from error resulting from the given function f.

Therefore, the Lebesgue constant gives us a means to understand errors introduced by interpolation. For example, it is known that if x_j are equispaced on $[a, b]$, then Λ grows exponentially with n, yielding a (very) poor approximation. However, consider the following points distributed on $[-1, 1]$:

$$x_j = \cos\theta_j, \qquad \theta_j = \frac{(j-1)\pi}{n-1}, \qquad j = 1, \ldots, n-1.$$

These points are called *Chebyshev* points. If one affinely maps these points to $[a, b]$, then it is known that Λ grows only logarithmically with n. It also turns out that logarthmic growth in n is the best (smallest) possible growth behavior for Λ.