Interpolating Polynomials

Handout March 7, 2012

Again we work over our favorite field F (such as $\mathbb{R}$, $\mathbb{Q}$, $\mathbb{C}$ or $\mathbb{F}_p$). We wish to find a polynomial $y = f(x)$ passing through n specified data points (x_1, y_1), (x_2, y_2), $\ldots$, (x_n, y_n) in the plane. For this, we must assume that the x-coordinates $x_1, x_2, \ldots, x_n$ are distinct. (Otherwise if $x_i = x_j$, it is clearly impossible to fit a function exactly to the data unless also $y_i = y_j$.) It is always possible to fit a polynomial function to the data exactly if one allows polynomials of sufficiently high degree. The good news is that there is a polynomial $f(x)$ of degree $< n$ which fits the data exactly. It is not hard to see that such a polynomial must be unique. For suppose $y = f(x)$ and $y = g(x)$ are two polynomials of degree $< n$, both of which fit the n data points. Then the difference polynomial $f(x) - g(x)$ has degree $< n$ and vanishes at $x_1, x_2, \ldots, x_n$. But this implies that $f(x) - g(x) = 0$, i.e. $f(x) = g(x)$.

Theorem 1. Let (x_1, y_1), (x_2, y_2), $\ldots$, (x_n, y_n) be ordered pairs of elements of F, and suppose that the values $x_1, x_2, \ldots, x_n$ are distinct. Then there exists a unique polynomial $p(x) \in F[x]$ of degree $< n$ such that $p(x_i) = y_i$ for $i = 1, 2, \ldots, n$.

We have seen why $f(x)$ must be unique, but not yet why it exists. We will in fact give two different proofs of this fact. Let us write

$$p(x) = a_0 + a_1 x + a_2 x^2 + \cdots + a_{n-1} x^{n-1}$$

We determine the n unknown coefficients $a_0, \ldots, a_{n-1} \in F$ using the n linear equations

$$\begin{aligned}
a_0 + a_1 x_1 + a_2 x_1^2 + \cdots + a_{n-1} x_1^{n-1} &= y_1, \\
a_0 + a_1 x_2 + a_2 x_2^2 + \cdots + a_{n-1} x_2^{n-1} &= y_2, \\
a_0 + a_1 x_3 + a_2 x_3^2 + \cdots + a_{n-1} x_3^{n-1} &= y_3, \\
\vdots \qquad\qquad\qquad & \quad \vdots \\
a_0 + a_1 x_n + a_2 x_n^2 + \cdots + a_{n-1} x_n^{n-1} &= y_n.
\end{aligned}$$

In matrix form, this linear system is expressed as $AX = B$, where

$$A = \begin{bmatrix} 1 & x_1 & x_1^2 & \cdots & x_1^{n-1} \\ 1 & x_2 & x_2^2 & \cdots & x_2^{n-1} \\ 1 & x_3 & x_3^2 & \cdots & x_3^{n-1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_n & x_n^2 & \cdots & x_n^{n-1} \end{bmatrix}, \quad X = \begin{bmatrix} a_0 \\ a_1 \\ a_2 \\ \vdots \\ a_{n-1} \end{bmatrix}, \quad B = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ \vdots \\ y_n \end{bmatrix}.$$

The matrix A shown is a very special $n \times n$ matrix, known as a *Vandermonde matrix.* The linear system $AX = B$ has a unique solution, namely $X = A^{-1}B$, if and only if $\det(A) \neq 0$. So to prove Theorem 2, what we really need to show is that $\det(A) \neq 0$. It is not hard to compute the case $n = 1$:

$$A = \begin{bmatrix} 1 & x_1 \\ 1 & x_2 \end{bmatrix}, \qquad \det(A) = x_2 - x_1$$

and the case $n = 2$:

$$A = \begin{bmatrix} 1 & x_1 & x_1^2 \\ 1 & x_2 & x_2^2 \\ 1 & x_3 & x_3^2 \end{bmatrix}, \qquad \det(A) = (x_2 - x_1)(x_3 - x_1)(x_3 - x_2)\,.$$

The general expression for $\det(A)$ is given by the following:

Lemma 2. The determinant of the $n \times n$ Vandermonde matrix A as above, is given by

$$\det(A) = \prod_{1 \le i < j \le n} (x_j - x_i)\,.$$

Here the notation '$\prod$' means 'product' (similar to the notation '$\sum$' meaning 'sum', with which you should be familiar). Note that the cases $n = 1, 2$ of Lemma 2 give exactly the determinants we computed above. Rather than give a complete proof of Lemma 2, we give the ideas of the proof. Note that $\det(A)$ is a polynomial $h(x_1, x_2, \ldots, x_n)$. Also, expanding $\det(A)$ using the definition of determinant, each of the $n!$ terms is a monomial of degree $0 + 1 + 2 + \cdots + (n-1) = n(n-1)/2$, so $\deg(h) = n(n-1)/2$. Also, this polynomial vanishes whenever $x_i = x_j$ where $i < j$. This implies that the product of all of the linear factors $(x_j - x_i)$ divides $h(x_1, x_2, \ldots, x_n)$, so

$$h(x_1, x_2, \ldots, x_n) = c(x_1, x_2, \ldots, x_n) \prod_{1 \le i < j \le n} (x_j - x_i)$$

for some polynomial $c(x_1, x_2, \ldots, x_n) \in F[x_1, x_2, \ldots, x_n]$. However, the number of such linear factors $x_j - x_i$ with $1 \le i < j \le n$ equals the binomial coefficient

$$\binom{n}{2} = \frac{n!}{2!(n-2)!} = \frac{n(n-1)}{2}\,.$$

By comparing degrees, we must have $\deg c(x_1, \ldots, x_n) = 0$, i.e. we have a constant polynomial $c(x_1, \ldots, x_n) = c \in F$. To determine the constant c, observe that the main diagonal of A gives a term $x_2 x_3^2 x_4^3 \cdots x_n^{n-1}$ in $h(x_1, \ldots, x_n)$. A similar term in the expansion of the product $\prod \cdots$ shows that $c = 1$. This proves Lemma 2. $\square$

Now it is easy to prove Theorem 1 from Lemma 2. Since by the hypothesis of Theorem 1, we assume the x_i's to be distinct, Lemma 2 gives $\det(A) \neq 0$. Thus A is invertible and the system $AX = B$ has a unique solution $X = A^{-1}B$. The entries of the column vector X give the coefficients of the polynomial $p(x)$. $\square$

Now let us give an alternative proof of Theorem 1. Let V be the vector space of all polynomials in x of degree $< n$ with coefficients in F, i.e.

$$V = \{a_0 + a_1 x + a_2 x^2 + \cdots + a_{n-1} x^{n-1} \ : \ a_0, a_1, \ldots, a_{n-1} \in F\}.$$

Every vector $p(x) \in V$ can be uniquely expressed as a linear combination of the vectors $1, x, x^2, \ldots, x^{n-1}$, viz.

$$p(x) = a_0 \cdot 1 + a_1 \cdot x + a_2 \cdot x^2 + \cdots + a_{n-1} \cdot x^{n-1}$$

where $a_0, a_1, \ldots, a_{n-1} \in F$. This means that $\{1, x, x^2, \ldots, x^{n-1}\}$ is a basis for V. It is not the only basis (soon we will exhibit others) but all bases have the same size. This size (i.e. n) is by definition the dimension of V.

Consider another vector space W consisting of all column vectors of length n over F, i.e.

$$W = \left\{ \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix} \ : \ b_1, b_2, \ldots, b_n \in F \right\}.$$

Every vector $w \in W$ can be uniquely expressed as a linear combination of the n vectors

$$\begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{bmatrix}, \ldots, \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix};$$

namely,

$$w = b_1 \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} + b_2 \begin{bmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{bmatrix} + \cdots + b_n \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix}.$$

So we have a set of n column vectors forming a basis for W. This means that W also has dimension n.

Now suppose $x_1, x_2, \ldots, x_n \in F$ are distinct, and consider the map

$$T : V \to W, \qquad T(p(x)) = \begin{bmatrix} p(x_1) \\ p(x_2) \\ \vdots \\ p(x_n) \end{bmatrix}.$$

It is easy to see that T is a linear transformation, i.e. $T(ap(x)+bq(x)) = aT(p(x))+bT(q(x))$ for all $a, b \in F$ and all $p(x), q(x) \in V$. The Fundamental Theorem of Linear Algebra says that

$$\dim\big(\text{null space of } T\big) + \dim\big(\text{image of } T\big) = \dim V = n.$$

Now the null space of T consists of all polynomials $p(x) \in V$ such that

$$T(p(x)) = \begin{bmatrix} p(x_1) \\ p(x_2) \\ \vdots \\ p(x_n) \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}.$$

The only polynomial satisfying this condition is $p(x) = 0$, since a nonzero polynomial of degree less than n cannot have n distinct roots. So the null space of T is zero (the subspace consisting of just the zero vector). This subspace has dimension 0. Now the image of T is a subspace of W of dimension n. But W itself has dimension n. So the image of T must consist of all of W, i.e. the linear transformation $T : V \to W$ is onto. It is also one-to-one since its null space is zero. (These are among the most basic facts from Linear Algebra; if they are not familiar, you should review them.)

Now the (second) proof of Theorem 1 is clear: Given

$$w = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} \in W,$$

there exists a unique $p(x) \in V$ such that $T(p(x)) = w$. The existence of $p(x) \in V$ comes from the fact that T is onto; and the uniqueness we have seen before (and is equivalent to the fact that T is one-to-one, since its null space is zero). $\square$

As an example, consider the problem of finding the unique polynomial $p(x) = a + bx + cx^2$ of degree < 3 through the three points $(0, 7)$, $(1, 6)$, $(2, 9)$. This gives a linear system

$$\begin{aligned} a \qquad\qquad\qquad &= 7, \\ a \quad + b \quad + c \quad &= 6, \\ a \quad + 2b \quad + 4c \quad &= 9. \end{aligned}$$

Solving this system by elementary row operations on the corresponding augmented matrix (which is faster than inverting the 3×3 coefficient matrix) gives the unique solution $a = 7$, $b = -3$, $c = 2$, and so the unique polynomial of degree < 3 passing through the three points given is $p(x) = 7 - 3x + 2x^2$. (You should check that this polynomial does indeed pass through the three points given.)

Working with $n \times n$ matrices, while possible in principle, becomes impractical for large values of n. We now give two other approaches to determining the interpolating polynomial $p(x)$, using more direct formulas which minimize the requirements on computational resources.

Newton's Interpolation Formula

Our first practical method for determining the interpolating polynomial $p(x)$ applies only in cases where the x-values $x_1, x_2, \ldots, x_n$ are equally spaced, separated by a common distance $h > 0$. This means that $x_k = x_1 + (k-1)h$ for $k = 1, 2, \ldots, n$. Define the *first-order differences* $\Delta y_k = y_{k+1} - y_k$, the *second-order differences* $\Delta^2 y_k = \Delta y_{k+1} - \Delta y_k$, etc. These differences may be conveniently tabulated by listing columns for the values of k, x_k and y_k, then listing the differences of the y_k's in a new column Δy_k, then the differences of the Δy_k's in a new column $\Delta^2 y_k$, then additional columns for $\Delta^3 y_k$, $\Delta^4 y_k$, etc. as far as possible, i.e. until you run out of values to take differences with. Also recall that the *factorial* of n is

$$n! = 1 \times 2 \times 3 \times \cdots \times n.$$

Theorem 3 (Newton's Interpolation Formula). If $x_k = x_1 + (k-1)h$ for $k = 1, 2, \ldots, n$, then the polynomial $p(x)$ of Theorem 1 is given by

$$p(x) = y_1 + \Delta y_1 \frac{x - x_1}{h} + \frac{\Delta^2 y_1}{2!} \frac{(x-x_1)(x-x_2)}{h^2} + \frac{\Delta^3 y_1}{3!} \frac{(x-x_1)(x-x_2)(x-x_3)}{h^3}$$
$$+ \cdots + \frac{\Delta^{n-1} y_1}{(n-1)!} \frac{(x-x_1)(x-x_2)(x-x_3)\cdots(x-x_{n-1})}{h^{n-1}}.$$

For the above example, we obtain the table of values

k	x_k	y_k	Δy_k	$\Delta^2 y_k$
1	0	7	−1	4
2	1	6	3	—
3	2	9	—	—

In our case $n = 2$, $h = 1$ and the interpolating polynomial is

$$p(x) \; = \; 7 + (-1)\frac{x}{1} + \frac{4}{2}\frac{x(x-1)}{1} \; = \; 7 - 3x + 2x^2 ,$$

which agrees with the answer we obtained from our linear system.

Using Theorem 3, it is possible to deduce the following:

Proposition 4. Given data points (x_i, y_i) with equally spaced values x_i, there exists an interpolating polynomial $p(x)$ of degree $< n$ iff $\Delta^k y_j = 0$ for all $k \geq n$.

As an illustration of Proposition 4, consider the following table of values obtained from the polynomial $7 - 3x + 2x^2$; note that the differences $\Delta^k y_\ell$ are zero for $k > 2$. This should remind you of the fact that the second-order derivative has constant value 4, and all higher order derivatives are zero.

x_k	y_k	Δy_k	$\Delta^2 y_k$	$\Delta^3 y_k$	$\Delta^4 y_k$	$\Delta^5 y_k$
−2	21	−9	4	0	0	0
−1	12	−5	4	0	0	—
0	7	−1	4	0	—	—
1	6	3	4	—	—	—
2	9	7	—	—	—	—
3	16	—	—	—	—	—

Also note that for the simple task of evaluating $p(x)$ at other *integer* values of x beyond the original data (i.e. *extrapolating*), it is *not necessary* to know the polynomial $p(x)$ explicitly; we only need to extend the evident pattern of tabulated differences as far as required. In the example above, starting with the three rows of the table for $x = 0, 1, 2$, we extend the table as far as desired starting with the rightmost columns (where the pattern is most obvious) and working toward the left. This allows us to determine $p(x)$ for $x = -2, -1, 3$ without having to compute a, b, c and substituting x-values into the resulting polynomial.

Lagrange's Interpolation Formula

We give the corresponding interpolation formula for the more general case when the x_k's are not necessarily equally spaced. Again, we are looking for the polynomial $p(x)$ of degree $< n$ which passes through the points (x_1, y_1), (x_2, y_2), $\ldots$, (x_n, y_n). First we define the polynomials

$$\phi_j(x) = \prod_{\substack{1 \le i \le n \\ i \ne j}} \frac{x - x_i}{x_j - x_i} = \frac{(x - x_1)(x - x_2) \cdots (x - x_{j-1})(x - x_{j+1}) \cdots (x - x_n)}{(x_j - x_1)(x_j - x_2) \cdots (x_j - x_{j-1})(x_j - x_{j+1}) \cdots (x_j - x_n)}$$

for $j = 1, 2, \ldots, n$. Observe that the product '$\prod$' has $n - 1$ factors in this case (for all values of $i = 1, 2, \ldots, n$ *excluding* j). Also, it is easy to see that

$$\phi_j(x_i) = \begin{cases} 1, & \text{for } i = j; \\ 0, & \text{for } i \ne j. \end{cases}$$

Theorem 5 (Lagrange's Interpolation Formula). The interpolating polynomial $p(x)$ of Theorem 1 is given by

$$p(x) = y_1\phi_1(x) + y_2\phi_2(x) + \cdots + y_n\phi_n(x).$$

Proof. In the sum $p(x_i) = \sum_{j=1}^{n} y_j\phi_j(x_i)$, all terms are zero except the term for $j = i$, and we get $p(x_i) = y_i \cdot 1 = y_i$. So the polynomial $p(x) = \sum_{j=1}^{n} y_j\phi_j(x)$ passes through the n data points (x_1, y_1), (x_2, y_2), $\ldots$, (x_n, y_n). Each $\phi_j(x)$ is a polynomial in x of degree $< n$, so $\sum_{j=1}^{n} y_j\phi_j(x)$ has degree $< n$. So $p(x) = \sum_{j=1}^{n} y_j\phi_j(x)$ must be *the* unique polynomial given in Theorem 1. $\square$

As an example, we determine the unique polynomial of degree < 3 through the three points $(0,-5)$, $(1,-2)$, $(3,10)$. Notice that in this case the x_k's are not equally spaced, so that Newton's Interpolation Formula does not apply. We compute

$$\begin{aligned}
\phi_1(x) &= \frac{(x-1)(x-3)}{(0-1)(0-3)} = \tfrac{1}{3}(x^2-4x+3)\,, \\
\phi_2(x) &= \frac{(x-0)(x-3)}{(1-0)(1-3)} = -\tfrac{1}{2}(x^2-3x)\,, \\
\phi_3(x) &= \frac{(x-0)(x-1)}{(3-0)(3-1)} = \tfrac{1}{6}(x^2-x)\,, \\
p(x) &= -5\phi_1(x) - 2\phi_2(x) + 10\phi_3(x) \\
&= -\tfrac{5}{3}(x^2-4x+3) + (x^2-3x) + \tfrac{10}{6}(x^2-x) \\
&= x^2+2x-5\,.
\end{aligned}$$

Again, you should check that this polynomial does in fact pass through the three points given.

HOMEWORK #4 Due Wed March 28

In each of the following, use the field $F = \mathbb{Q}$ of rational numbers.

1. Let n be a positive integer, and consider the unique polynomial $P(x)$ of degree $< n$ such that $P(k) = 2^k$ for $k = 0, 1, 2, \ldots, n-1$. Determine $P(n)$ and prove that your answer is correct. (*Hint:* It is not necessary to explicitly determine the polynomial $P(x)$. Use the pattern evident in a table of differences, and consider the remarks following Proposition 4.)

2. Consider the five data points $(0,-5)$, $(2,-15)$, $(4,-25)$, $(6,13)$, $(8,147)$. Observe that the x-values are equally spaced.
   (a) What is h?
   (b) Construct a table of values for k, x_k, y_k, Δy_k, $\Delta^2 y_k$, $\Delta^3 y_k$, $\Delta^4 y_k$.
   (c) Use Newton's Interpolation Formula to find a polynomial $p(x)$ of degree < 5 passing through the five data points.
   (d) Check that your answer to (c) is correct by evaluating $p(x_k)$ for $k = 1, 2, 3, 4, 5$.
   (e) What is the degree of $p(x)$?
   (f) Explain the relationship between your answers to (b) and (e), using Proposition 4.

3. Use Lagrange's Interpolation Formula to find a polynomial of degree < 4 passing through the four data points $(-1, 16)$, $(0, 19)$, $(1, 22)$, $(3, -20)$. Check that your polynomial is correct by evaluating $p(x_k)$ for $k = 1, 2, 3, 4$.

4. Find the unique polynomial function $y = p(x)$ of degree < 3 passing through the three points $(-3, -30)$, $(0, 12)$, $(3, 0)$

   (a) by solving the appropriate system of three linear equations in three unknowns;

   (b) by Newton's Interpolation Formula;

   (c) by Lagrange's Interpolation Formula.

   Compare your answers in (a), (b) and (c).