Orthogonal Polynomials

(Com S 477/577 Notes)

Yan-Bin Jia

Nov 17, 2016

1 Introduction

We have seen the importance of orthogonal projection and orthogonal decomposition, particularly in the solution of systems of linear equations and in the least-squares data fitting. In fact, these ideas can generalize from vectors to functions. For instance, we might let V be the "vector space" of continuous integrable functions defined on some interval $[a, b]$. Such a space V has infinite dimensions.

Many inner products can be defined on V. Which one to use is generally determined by your application. Here we define the inner product of two functions g and h in one of two ways:

$$\begin{aligned} \langle g, h \rangle &= \int_a^b g(x)h(x)w(x)\,dx, \qquad \text{or} && (1) \\ \langle g, h \rangle &= \sum_{i=1}^{n} g(x_i)h(x_i)w(x_i), && (2) \end{aligned}$$

where $w(x)$ is a positive function, called a *weighting function.* The form (2) is used often when we do not have information about g and h at all points. Instead, we may have information only at n discrete points $x_1, \dots, x_n$.

With the inner product defined, we say that the two functions $g(x)$ and $h(x)$ are *orthogonal* if

$$\langle g, h \rangle = 0.$$

Example 1. It is easy to verify, for example, that the functions $g(x) = 1$, $h(x) = x$ are orthogonal if the inner product is

$$\langle g, h \rangle = \int_{-1}^{1} g(x)h(x)\,dx,$$

or if it is

$$\langle g, h \rangle = \sum_{i=-10}^{10} g(i)h(i),$$

or if it is

$$\langle g, h \rangle = \int_{-1}^{1} \frac{g(x)h(x)}{\sqrt{1-x^2}}\,dx.$$

The functions $g(x) = \sin nx$, $h(x) = \sin mx$, integers n and m, are orthogonal if

$$\langle g, h \rangle = \int_0^{2\pi} g(x)h(x)\,dx,$$

and $n \neq m$, as are the functions $g(x) = \sin nx$ and $h(x) = \cos mx$.

Why are we concerned with orthogonal functions? Because there are two parts to good data fitting: a) selecting a good basis of functions $\mathcal{B}$ and b) approximating a given function or a set of observed data using the basis functions. We have seen how to carry out part b) but we have yet to discuss part a). The point here is that if we find an orthogonal basis $\mathcal{B}$, we would be able to approximate or decompose a function f by the rule

$$f \cong \sum_{g \in \mathcal{B}} \frac{\langle f, g \rangle}{\langle g, g \rangle}\, g.$$

The above is an equality if $f \in \mathrm{span}(\mathcal{B})$, that is, f is a linear combination of some functions in $\mathcal{B}$. Otherwise, it is an orthogonal projection of f onto $\mathrm{span}(\mathcal{B})$.

2 Orthogonal Polynomials

A sequence of *orthogonal polynomials* consists of $p_0(x), p_1(x), p_2(x), \ldots$ (finite or infinite) such that

a) $p_i(x)$ is a polynomial of degree i;

b) $\langle p_i, p_j \rangle = 0$ whenever $i \neq j$.

Example 2. The polynomials

$$p_0(x) = 1, \qquad p_1(x) = x, \qquad \text{and} \qquad p_2(x) = 3x^2 - 1$$

constitute a sequence of orthogonal polynomials under the inner product

$$\langle g, h \rangle = \int_{-1}^{1} g(x)h(x)\,dx.$$

We know from Example 1 that $\langle p_0, p_1 \rangle = 0$. Also we obtain that

$$\begin{aligned} \langle p_0, p_2 \rangle &= \int_{-1}^{1} 1 \cdot (3x^2 - 1)\,dx &= x^3 - x \Big|_{-1}^{1} = 0, \\ \langle p_1, p_2 \rangle &= \int_{-1}^{1} x \cdot (3x^2 - 1)\,dx &= \frac{3}{4}x^4 - \frac{1}{2}x^2 \Big|_{-1}^{1} = 0. \end{aligned}$$

Example 3. Chebyshev polynomials

$$\begin{aligned} T_0(x) &= 1, \\ T_1(x) &= x, \\ T_{k+1}(x) &= 2xT_k(x) - T_{k-1}(x), \qquad k = 1, 2, \ldots, \end{aligned}$$

are orthogonal with respect to two inner products:

a)

$$\langle g, h\rangle = \int_{-1}^{1} \frac{g(x)h(x)}{\sqrt{1-x^2}}\, dx.$$

In this case,

$$\langle T_i, T_j\rangle = \int_{-1}^{1} \frac{T_i(x)T_j(x)}{\sqrt{1-x^2}}\, dx = \begin{cases} 0, & i \neq j \\ \frac{\pi}{2}, & i = j \neq 0 \\ \pi, & i = j = 0 \end{cases}$$

b)

$$\langle g, h\rangle = \sum_{k=0}^{m-1} g(\xi_{k,m})h(\xi_{k,m}),$$

where $\xi_{k,m} = \cos\frac{2k+1}{2m}\pi$, $k = 0, \ldots, m-1$, are the m zeros of $T_m(x)$. In this case,

$$\langle T_i, T_j\rangle = \sum_{k=0}^{m-1} T_i(\xi_{k,m})T_j(\xi_{k,m}) = \begin{cases} 0, & i \neq j \\ \frac{m}{2}, & i = j \neq 0 \\ m, & i = j = 0 \end{cases}$$

The following facts can be proved about a finite sequence of orthogonal polynomials $p_0(x), p_1(x), \ldots, p_k(x)$:

i) If $p(x)$ is any polynomial of degree at most k, then one can write

$$p(x) = d_0p_0(x) + d_1p_1(x) + \cdots + d_kp_k(x)$$

with the coefficients $d_0, \ldots, d_k$ uniquely determined when $\langle p_i, p_i\rangle \neq 0$ for all i. Let us take the inner product of $p_i(x)$ with both sides of the above equation:

$$\begin{aligned} \langle p, p_i\rangle &= d_0\langle p_0, p_i\rangle + \cdots + d_k\langle p_k, p_i\rangle \\ &= 0 + \cdots + 0 + d_i\langle p_i, p_i\rangle + 0 + \cdots + 0, \qquad \langle p_j, p_i\rangle = 0 \text{ for all } j \neq i. \end{aligned}$$

Hence we have

$$d_i = \frac{\langle p, p_i\rangle}{\langle p_i, p_i\rangle}.$$

ii) If $p(x)$ is any polynomial of degree less than k, then

$$\langle p, p_k\rangle = 0.$$

By Property i), $p(x) = d_0p_0(x) + d_1p_1(x) + \cdots + d_lp_l(x)$, where $l \leq k$ is the degree of p. Taking the inner product of p_k with both sides of this equation, we obtain that $\langle p, p_k\rangle = 0$.

iii) If the inner product is given by (1), then $p_k(x)$ has k simple real zeros in the interval (a, b).

iv) The orthogonal polynomials satisfy a three-term recurrence relation:

$$p_{i+1}(x) = A_i(x - B_i)p_i(x) - C_ip_{i-1}(x), \qquad i = 0, 1, \ldots, k-1, \tag{3}$$

where $p_{-1}(x) = 0$ and

$$\begin{aligned} A_i &= \frac{\text{leading coefficient of } p_{i+1}}{\text{leading coefficient of } p_i}, \\ B_i &= \frac{\langle xp_i, p_i \rangle}{\langle p_i, p_i \rangle}, \\ C_i &= \begin{cases} \text{arbitrary} & \text{if } i = 0, \\ \dfrac{A_i \cdot \langle p_i, p_i \rangle}{A_{i-1} \cdot \langle p_{i-1}, p_{i-1} \rangle} & \text{if } i > 0. \end{cases} \end{aligned}$$

In the case where the polynomials are *monic* (with leading coefficient 1), the following recurrence holds:

$$\begin{aligned} p_{-1}(x) &= 0, \\ p_0(x) &= 1, \\ p_1(x) &= \left(x - \frac{\langle xp_0, p_0 \rangle}{\langle p_0, p_0 \rangle} \right) p_0(x), \\ p_{i+1}(x) &= \left(x - \frac{\langle xp_i, p_i \rangle}{\langle p_i, p_i \rangle} \right) p_i(x) - \frac{\langle p_i, p_i \rangle}{\langle p_{i-1}, p_{i-1} \rangle} p_{i-1}(x), \qquad i = 1, 2, \ldots \end{aligned}$$

This property allows us to generate an orthogonal polynomial sequence provided $\langle p_i, p_i \rangle \neq 0$ for all i.

EXAMPLE 4. *Legendre polynomials* The inner product is given by

$$\langle g, h \rangle = \int_{-1}^{1} g(x)h(x)\,dx.$$

Starting with $p_0(x) = 1$, we get

$$\begin{aligned} \langle p_0, p_0 \rangle &= \int_{-1}^{1} 1\,dx = 2, \\ \langle xp_0, p_0 \rangle &= \int_{-1}^{1} x\,dx = 0. \end{aligned}$$

Hence

$$p_1(x) = (x - 0)p_0(x) = x,$$

and

$$\begin{aligned} \langle p_1, p_1 \rangle &= \int_{-1}^{1} x^2\,dx = \frac{2}{3}, \\ \langle xp_1, p_1 \rangle &= \int_{-1}^{1} x^3\,dx = 0. \end{aligned}$$

So now we have

$$p_2(x) = (x - 0)p_1(x) - \frac{\frac{2}{3}}{2} p_0(x) = x^2 - \frac{1}{3}.$$

Continuing this process we would get

$$\begin{aligned} p_0(x) &= 1, \\ p_1(x) &= x, \\ p_2(x) &= x^2 - \frac{1}{3}, \\ p_3(x) &= x^3 - \frac{3}{5}x, \\ p_4(x) &= x^4 - \frac{6}{7}x^2 + \frac{3}{35}, \\ &\vdots \end{aligned}$$

Legendre polynomials can also be "normalized" in the sense that

$$p_k(1) = 1, \qquad \text{for all } k.$$

The coefficients in the recurrence (3) then become

$$\begin{aligned} A_k &= \frac{2k+1}{k+1}, \\ B_k &= 0, & k = 0, 1, \ldots \\ C_k &= \frac{k}{k+1}. \end{aligned}$$

And the recurrence subsequently reduces to

$$p_{k+1}(x) = \frac{(2k+1)xp_k(x) - kp_{k-1}(x)}{k+1}.$$

3 Least-Squares Approximation by Polynomials

Given a function $f(x)$ defined on some interval (a, b), we want to approximate it by a polynomial of degree at most k. Here we measure the difference between $f(x)$ and a polynomial $p(x)$ by

$$\langle f(x) - p(x), f(x) - p(x) \rangle,$$

where the inner product is defined by either (1) or (2). And we would like to seek a polynomial of degree at most k to minimize the above inner product. Such a polynomial is a least-squares approximation to $f(x)$ by polynomials of degrees not exceeding k.

We proceed by finding an orthogonal sequence of polynomials $p_0(x), \ldots, p_k(x)$ for the chosen inner product such that $\langle p_i, p_j \rangle = 0$ whenever $i \neq j$. Then every polynomial of degree at most k can be written uniquely as

$$p(x) = d_0 p_0(x) + \cdots + d_k p_k(x)$$

where

$$d_i = \frac{\langle p, p_i \rangle}{\langle p_i, p_i \rangle}$$

So now we try to minimize

$$\langle f(x) - p(x), f(x) - p(x) \rangle = \langle f(x) - d_0 p_0(x) - \cdots - d_k p_k(x), f(x) - d_0 p_0(x) - \cdots - d_k p_k(x) \rangle$$

over all possible choices of $d_0, \ldots, d_k$, or equivalently, over all polynomials of degree at most k. The partial derivatives of the above inner product with respect to $d_0, \ldots, d_k$ must all vanish; in other words, the "best" coefficients must satisfy the *normal equations*

$$d_0\langle p_0, p_i\rangle + d_1\langle p_1, p_i\rangle + \cdots + d_k\langle p_k, p_i\rangle = \langle f, p_i\rangle, \qquad i = 0, \ldots, k.$$

Due to the orthogonality of $p_j(x)$, the normal equations reduce to

$$d_i\langle p_i, p_i\rangle = \langle f, p_i\rangle, \qquad i = 0, \ldots, k.$$

Hence the best coefficients are given by

$$d_i = \frac{\langle f, p_i\rangle}{\langle p_i, p_i\rangle}, \qquad i = 0, \ldots, k. \tag{4}$$

Here is the analogy to the case of the least-squares technique over a vector space. In the space of all functions, the orthogonal polynomials $p_0, \ldots p_k$ constitute an "orthogonal basis" for the subspace of polynomial functions of degree no more than k. The least-squares approximation of a function f by polynomials in this subspace is then its orthogonal projection onto the subspace. The coordinates of this projection along the axes $p_0, \ldots, p_k$ are thus $\langle f, p_0\rangle/\langle p_0, p_0\rangle, \ldots, \langle f, p_k\rangle/\langle p_k, p_k\rangle$.

Below we illustrate the use of orthogonal polynomials for obtaining least-squares approximations with respect to both continuous and discrete versions of inner products.

Example 5. Calculate the polynomial at degree at most 3 that best approximates e^x over the interval $[-1, 1]$ in the least-squares sense.

Here we obtain a best approximation by orthogonally projecting e^x onto the subspace of functions spanned by Legendre polynomials $p_0, \ldots, p_3$. In other words,

$$p(x) = \sum_{i=0}^{3} d_i p_i(x),$$

where

$$d_i = \frac{\langle e^x, p_i\rangle}{\langle p_i, p_i\rangle}.$$

We compute the following inner products:

$$\begin{aligned}
\langle p_0, p_0\rangle &= \int_{-1}^{1} 1\,dx = 2,\\
\langle p_1, p_1\rangle &= \int_{-1}^{1} x^2\,dx = \frac{2}{3},\\
\langle p_2, p_2\rangle &= \int_{-1}^{1} \left(x^4 - \frac{2}{3}x^2 + \frac{1}{9}\right) dx = \frac{8}{45},\\
\langle p_3, p_3\rangle &= \int_{-1}^{1} \left(x^6 - \frac{6}{5}x^4 + \frac{9}{25}x^2\right) dx = \frac{8}{175},\\
\langle e^x, p_0\rangle &= \int_{-1}^{1} e^x\,dx = e - \frac{1}{e},\\
\langle e^x, p_1\rangle &= \int_{-1}^{1} e^x x\,dx = \frac{2}{e},
\end{aligned}$$

$$\begin{aligned}
\langle e^x, p_2\rangle &= \int_{-1}^{1} e^x\left(x^2-\frac{1}{3}\right)dx = \frac{2}{3}e-\frac{14}{3e},\\
\langle e^x, p_3\rangle &= \int_{-1}^{1} e^x\left(x^3-\frac{3}{5}x\right)dx = -2e+\frac{74}{5e}.
\end{aligned}$$

Then

$$\begin{aligned}
d_0 &= \frac{1}{2}\left(e-\frac{1}{e}\right) \approx 1.1752012,\\
d_1 &= \frac{3}{2}\cdot\frac{2}{e} \approx 1.1036383,\\
d_2 &= \frac{45}{8}\left(\frac{2}{3}e-\frac{14}{3e}\right) \approx 0.53672153,\\
d_3 &= \frac{175}{8}\left(-2e+\frac{74}{5e}\right) \approx 0.17613908.
\end{aligned}$$

So the least-squares approximation to e^x on $(-1,1)$ is

$$\begin{aligned}
p(x) &= 1.1752012\cdot p_0(x)+1.1036383\cdot p_1(x)+0.53672153\cdot p_2(x)+0.17613908\cdot p_3(x)\\
&= 0.99629402+0.99795487x+0.53672153x^2+0.17613908x^3.
\end{aligned}$$

Example 6. Find the least-squares approximation of $f(x)=10-2x+x^2/10$ by a quadratic polynomial over supporting points

$$x_i = 10+\frac{i-1}{5} \qquad \text{and} \qquad f_i=f(x_i), \qquad i=1,\ldots,6.$$

In this case, we seek the polynomial of degree at most 2 which minimizes

$$\sum_{i=1}^{6}\Big(f_i-p(x_i)\Big)^2.$$

So the inner product (2) is used with $w(x)\equiv 1$. We start by calculating the following

$$\begin{aligned}
p_0(x) &= 1,\\
\langle p_0,p_0\rangle &= \sum_{i=1}^{6} 1\cdot 1 = 6,\\
\langle xp_0,p_0\rangle &= \sum_{i=1}^{6}\left(10+\frac{i-1}{5}\right)\cdot 1 = 63.
\end{aligned}$$

Therefore

$$\begin{aligned}
p_1(x) &= \left(x-\frac{\langle xp_0,p_0\rangle}{\langle p_0,p_0\rangle}\right)\cdot 1-0 = x-10.5,\\
\langle p_1,p_1\rangle &= \sum_{i=1}^{6}\left(\frac{i-1}{5}-0.5\right)^2 = 0.7,\\
\langle xp_1,p_1\rangle &= \sum_{i=1}^{6}\left(10+\frac{i-1}{5}\right)\left(\frac{i-1}{5}-0.5\right)^2 = 7.35.
\end{aligned}$$

We can go on to calculate $p_2(x)$, obtaining

$$\begin{aligned}
p_2(x) &= \left(x - \frac{\langle xp_1, p_1\rangle}{\langle p_1, p_1\rangle}\right) p_1(x) - \frac{\langle p_1, p_1\rangle}{\langle p_0, p_0\rangle} p_0(x) \\
&= \left(x - \frac{7.35}{0.7}\right)(x - 10.5) - \frac{0.7}{6} \\
&= (x - 10.5)^2 - 0.1166667, \\
\langle p_2, p_2\rangle &= 0.05973332.
\end{aligned}$$

Next, we calculate the coefficients for the least-squares approximation:

$$\begin{aligned}
\frac{\langle f, p_0\rangle}{\langle p_0, p_0\rangle} &= \sum_{i=1}^{6} \frac{f_i}{6} &= 0.0366667, \\
\frac{\langle f, p_1\rangle}{\langle p_1, p_1\rangle} &= \sum_{i=1}^{6} \frac{f_i \cdot p_1(x_i)}{0.7} &= 0.1, \\
\frac{\langle f, p_2\rangle}{\langle p_2, p_2\rangle} &= \sum_{i=1}^{6} \frac{f_i \cdot p_2(x_i)}{0.05973332} &= 0.0999999.
\end{aligned}$$

So the least-squares approximation for $f(x)$ is

$$\begin{aligned}
p(x) &= 0.03666667 + 0.1(x - 10.5) + 0.0999999\Big((x - 10.5)^2 - 0.1166667\Big) \\
&= 9.99998 - 2x + 0.0999999x^2.
\end{aligned}$$

References

[1] S. D. Conte and Carl de Boor. *Elementary Numerical Analysis.* McGraw-Hill, Inc., 2nd edition, 1972.

[2] M. Erdmann. Lecture notes for *16-811 Mathematical Fundamentals for Robotics.* The Robotics Institute, Carnegie Mellon University, 1998.