Mathematical Tripos Part IB: Easter Term 2003

Numerical Analysis – Lecture 8

8 Linear least squares (cont.)

8.1 Orthogonal polynomials

Consider $\mathbb{X} = C[a,b]$, the space of all continuous real-valued functions $f : [a,b] \to \mathbb{R}$, and define a scalar product on $C[a,b]$ by

$$(f,g) := (f,g)_w := \int_a^b f(x)g(x)\,w(x)\,dx. \tag{8.1}$$

Here w, the so-called *weight function*, is a fixed positive function, such that the integral $\int h(x)w(x)\,dx$ exists for all $h \in C[a,b]$.

We denote by $\mathcal{P}_n$ the space of all algebraic polynomials of degree (at most) n, i.e., $p \in \mathcal{P}_n$ if $p(x) = \sum_{k=0}^n a_k x^k$. If the *leading* coefficent a_n equals 1, then p is called a *monic* polynomial. Given a scalar product (8.1), we say that $Q_n \in \mathcal{P}_n$ is the n-th *orthogonal polynomial* if

$$(Q_n, p) = 0 \quad \forall p \in \mathcal{P}_{n-1}.$$

(Note: different weights lead to different orthogonal polynomials!)

Lemma 8.1 *For every n there exists a unique monic orthogonal polynomial Q_n. Any $p \in \mathcal{P}_n$ is uniquely expressible as a linear combination*

$$p = \textstyle\sum_{k=0}^n c_k Q_k, \quad c_k = (p,Q_k)/\|Q_k\|^2. \tag{8.2}$$

Proof. We just repeat the Gram-Schmidt algorithm. Starting with $Q_0 \equiv 1$ we set

$$Q_n(x) := x^n - \textstyle\sum_{k=1}^{n-1} \frac{(x^n,Q_k)}{(Q_k,Q_k)} Q_k(x), \quad n = 1,2,\ldots.$$

Then Q_n is orthogonal to each previous Q_k, and $x^n \in \operatorname{span}(Q_k)_{k=0}^n$, hence $\mathcal{P}_n = \operatorname{span}(Q_k)_{k=0}^n$. Therefore $Q_n \perp \mathcal{P}_{n-1}$, and any $p \in \mathcal{P}_n$ has an expansion (8.2), where each coefficient c_ℓ is uniquely determined by multiplying both sides with Q_ℓ. If $\widetilde{Q}_n$ is another n-th orthogonal polynomial, then in its expansion (8.2) all $c_k = 0$ for $k < n$, and if it is monic then $c_n = 1$, i.e., $\widetilde{Q}_n \equiv Q_n$. ∎

Theorem 8.2 (Three-term recurrence relation) *Monic orthogonal polynomials satisfy the relation*

$$Q_{n+1}(x) = (x - a_n)Q_n(x) - b_n Q_{n-1}(x), \qquad n = 0,1,\ldots \tag{8.3}$$

where $Q_{-1}(x) \equiv 0$, $Q_0(x) \equiv 1$, and

$$a_n := \frac{(xQ_n,Q_n)}{(Q_n,Q_n)}, \qquad b_n := \frac{(Q_n,Q_n)}{(Q_{n-1},Q_{n-1})} > 0.$$

Proof. Based on (8.2), let us look at the coefficients of the expansion

$$xQ_n(x) = \sum_{k=0}^{n+1} c_k Q_k(x), \qquad c_k = \frac{(xQ_n,Q_k)}{(Q_k,Q_k)} = \frac{(Q_n,xQ_k)}{(Q_k,Q_k)}$$

- $k = n+1 \;\to\; c_{n+1} = 1$, since both $xQ_n(x)$ and $Q_{n+1}(x)$ are monic polynomials of degree $n+1$.
- $k = n \;\to$ By definition, $c_n = a_n$.
- $k = n-1 \;\to$ Because of monicity, we have the equality $xQ_{n-1} = Q_n + p_{n-1}$ where $p_{n-1} \in \mathcal{P}_{n-1}$, so that $(Q_n, xQ_{n-1}) = (Q_n, Q_n + p_{n-1}) = (Q_n,Q_n)$, hence $c_{n-1} = b_n$.
- $k < n-1 \;\to$ Since $xQ_k \in \mathcal{P}_{n-1}$, we obtain $(Q_n, xQ_k) = 0$, thus $c_k = 0$.

It follows that $xQ_n(x) = Q_{n+1} + a_n Q_n + b_n Q_{n-1}$, and that is equivalent to (8.3). ∎

Remark 8.3 If (Q_k) have leading coefficients (α_k), then the recurrence takes the form

$$Q_{n+1}(x) = \frac{\alpha_{n+1}}{\alpha_n}(x - a_n)Q_n(x) - \frac{\alpha_{n+1}\alpha_{n-1}}{\alpha_n^2} b_n Q_{n-1}(x), \qquad n = 0, 1, \ldots$$

and, with an appropriate choice of (α_k), may become very simple.

8.2 Examples

Example 8.4 Classical examples of orthogonal (non-monic) polynomials include

Name	Notation	Interval	Weight function	Recurrence
Legendre	P_n	$[-1, 1]$	$w(x) \equiv 1$	$(n+1)P_{n+1}(x) = (2n+1)xP_n(x) - nP_{n-1}(x)$
Chebyshev	T_n	$[-1, 1]$	$w(x) = (1-x^2)^{-1/2}$	$T_{n+1}(x) = 2xT_n(x) - T_{n-1}(x)$
Laguerre	L_n	$[0, \infty)$	$w(x) = \mathrm{e}^{-x}$	$(n+1)L_{n+1}(x) = (2n+1-x)L_n(x) - nL_{n-1}(x)$
Hermite	H_n	$(-\infty, \infty)$	$w(x) = \mathrm{e}^{-x^2}$	$H_{n+1}(x) = 2xH_n(x) - 2nH_{n-1}(x)$

Example 8.5 (Chebyshev polynomials) Chebyshev polynomials T_n of degree n are defined by

$$T_n(x) = \cos n \arccos x, \quad x \in [-1, 1],$$

or, in a more instructive form,

$$T_n(x) = \cos n\theta, \quad x = \cos\theta, \quad \theta \in [0, \pi]\,.$$

They satisfy the relations

$$T_0(x) \equiv 1, \quad T_1(x) = x, \quad T_{n+1}(x) = 2xT_n(x) - T_{n-1}(x), \quad n \geq 1,$$

where the first two ones are straightforward and the recurrence follows from the equality

$$\cos(n+1)\theta + \cos(n-1)\theta = 2\cos\theta\cos n\theta\,.$$

In particular, T_n are indeed algebraic polynomials with leading coefficients 2^{n-1}. Further,

$$\begin{aligned}(T_n, T_m) &:= \int_{-1}^{1} T_n(x)T_m(x)\frac{dx}{\sqrt{1-x^2}} = \int_0^{\pi} \cos n\theta \cos m\theta \, d\theta \\ &= \int_0^{\pi} \tfrac{1}{2}[\cos(n+m)\theta + \cos(n-m)\theta]\, d\theta = 0, \quad n \neq m,\end{aligned}$$

i.e., (T_n) are orthogonal with respect to the scalar product

$$(f, g) := \int_{-1}^{1} f(x)g(x)\frac{dx}{\sqrt{1-x^2}}\,.$$

8.3 Least squares polynomial fitting

Let f be a continuous function defined on some interval $[a, b]$, and suppose we wish to approximate f by a polynomial of degree n. This is a continuous analogue of the discrete problem considered in Lecture 7. If we equip $C[a, b]$ with the distance $\|f - g\| := (f - g, f - g)^{1/2}$ induced by an inner product $(f, g) = \int_a^b f(x)g(x)\, w(x)dx$, then it is natural to seek a polynomial

$$p^* := \arg\min_{p \in \mathcal{P}_n} \|f - p\| .$$

for which the distance $\|f - p\|$ is as small as possible. Such a polynomial is called a (*weighted*) *least squares approximant*.

Theorem 8.6 *Let $(Q_k)_{k=0}^n$ be polynomials orthogonal with respect to a given inner product. Then the least squares approximant to any $f \in C[a, b]$ from $\mathcal{P}_n$ is given by the formula*

$$p^* = p^*(f) = \textstyle\sum_{k=0}^n c_k^* Q_k, \quad c_k^* = \frac{(f,Q_k)}{\|Q_k\|^2}, \tag{8.4}$$

and the value of the least squares approximation is

$$\|f - p^*\|^2 = \|f\|^2 - \textstyle\sum_{k=1}^n \frac{(f,Q_k)^2}{\|Q_k\|^2} . \tag{8.5}$$

Remark 8.7 By orthogonality of Q_k, the norm of the extremal p^* is equal to

$$\|p^*\|^2 = \|\textstyle\sum_{k=0}^n c_k^* Q_k\|^2 = \sum_{k=0}^n |c_k^*|^2 \|Q_k\|^2 = \sum_{k=1}^n \frac{(f,Q_k)^2}{\|Q_k\|^2},$$

hence formula (8.5) takes the form

$$\|f - p^*\|^2 = \|f\|^2 - \|p^*\|^2$$

what is a reminiscent of the Pythagoras theorem.

Proof. Let $p = \sum_{k=0}^n c_k Q_k$. Then

$$F(c) := (f - p, f - p) = (f - \sum_{k=0}^n c_k Q_k, f - \sum_{k=0}^n c_k Q_k) = \|f\|^2 - 2\sum_{k=0}^n c_k (f, Q_k) + \sum_{k=0}^n c_k^2 \|Q_k\|^2 \tag{8.6}$$

The right-hand side is a quadratic polynomial in each c_k, therefore it attains its minimum when

$$\tfrac{\partial F}{\partial c_k}\big|_{c_k = c_k^*} = -2(f, Q_k) + 2c_k^* \|Q_k\|^2 = 0, \quad k = \overline{0, n},$$

hence the conclusion (8.4). Putting optimal c_k^* into (8.6) we obtain (8.5).

Method 8.8 Suppose we want to approximate $f \in C[a, b]$ with a prescribed accuracy ϵ, i.e., to find $n = n(\epsilon)$ such that

$$\|f - p^*\| \le \epsilon, \qquad p^* \in \mathcal{P}_n.$$

From (8.5), it follows that the required value n can be calculated by summing the terms of $\frac{(f,Q_k)^2}{\|Q_k\|^2}$ until we reach the bound

$$\textstyle\sum_{k=1}^n \frac{(f,Q_k)^2}{\|Q_k\|^2} \ge \|f\|^2 - \epsilon^2.$$

The next theorem assures that this bound can be reached.

Theorem 8.9 (Parceval identity) *If $[a, b]$ is finite, then $\sum_{k=0}^\infty \frac{(f,Q_k)^2}{\|Q_k\|^2} = \|f\|^2$.*

Proof (Incomplete). Let

$$\sigma_n^2 := \|p^*\|^2 = \sum_{k=0}^{n} \frac{(f,Q_k)^2}{\|Q_k\|^2},$$

hence

$$\inf_{p\in\mathcal{P}_n} \|f-p\|^2 = \|f-p^*\|^2 = \|f\|^2 - \sigma_n^2.$$

According to the *Weierstrass theorem*, any function in $C[a,b]$ can be approximated arbitrarily close by a polynomial, hence $\lim_{n\to\infty} \inf_{p\in\mathcal{P}_n} \|f-p\|^2 = 0$ and we deduce that $\sigma_n^2 \to \|f\|^2$ as $n\to\infty$. ∎

Remark 8.10 Analogous identity is valid for the trigonometric system.

8.4 Exercises

8.1. Set $f(x) = 2x-1$, $x\in[0,1]$. We require a function of form

$$p(x) = \sum_{k=0}^{n} a_k \cos(k\pi x), \quad 0\le x\le 1,$$

that satisfies the condition

$$\int_0^1 [f(x)-p(x)]^2\,\mathrm{d}x < 10^{-4}.$$

Explain why it is sufficient if the value of $a_0^2 + \frac{1}{2}\sum_{k=1}^{n} a_k^2$ exceeds $\frac{1}{3} - 10^{-4}$, where the coefficients $\{a_k\}_{k=0}^n$ are calculated to minimize this integral. Hence find the smallest acceptable value of n.

8.2. The polynomials $\{p_n\}_{n\in\mathbb{Z}^+}$ are defined by the three-term recurrence formula

$$\begin{aligned} p_0(x) &\equiv 1, \\ p_1(x) &= 2x, \\ p_{n+1}(x) &= 2xp_n(x) - p_{n-1}(x), \qquad n=1,2,\ldots. \end{aligned}$$

Prove that they are orthogonal with respect to the inner product

$$(f,g) = \int_{-1}^{1} f(x)g(x)\sqrt{1-x^2}\,dx$$

and evaluate (p_n,p_n) for $n\in\mathbb{Z}^+$. [*Hint*: Prove that $p_n(x) = \sin(n+1)\theta/\sin\theta$, where $x=\cos\theta$.]

8.3. The functions $p_0, p_1, p_2, \ldots$ are generated by the Rodrigues formula

$$p_n(x) = \mathrm{e}^x \frac{\mathrm{d}^n}{\mathrm{d}x^n}(x^n\mathrm{e}^{-x}), \qquad 0\le x<\infty.$$

Show that these functions are polynomials and prove by integration by parts that for every $p\in\mathcal{P}_{n-1}$ we have the orthogonality condition $(p_n,p)=0$ with respect to the scalar product

$$(f,g) := \int_0^\infty f(x)g(x)\,\mathrm{e}^{-x}\,\mathrm{d}x.$$

Derive the coefficients of p_3, p_4 and p_5 from the Rodrigues formula. Verify that these coefficients are compatible with a three term recurrence relation of the form

$$p_5(x) = (\gamma x-\alpha)p_4(x) - \beta p_3(x), \qquad x\in\mathbb{R},$$

where α,β and γ are constants.