Mathematical Tripos Part IB: Easter Term 2002

Numerical Analysis – Lecture 8

8 Linear least squares (cont.)

8.1 Orthogonal systems

Recall that a set (q_j) in an inner product space $\mathbb{X}$ is orthogonal if $(q_i, q_j) = 0$ for $i \neq j$, and it is orthonormal if $(q_i, q_j) = \delta_{ij}$ The least squares approximation by orthogonal systems is especially simple and effective.

Theorem 8.1 *Let $\mathbb{X}$ be an inner product space, $\mathcal{A}_n \subset \mathbb{X}$ be a subspace, and let $(q_k)_{k=1}^n$ be an orthogonal set that spans $\mathcal{A}_n$. Then the least squares approximant to any $f \in \mathbb{X}$ from $\mathcal{A}_n$ is given by the formula*

$$a^* := a^*(f) = \textstyle\sum_{k=1}^n c_k q_k, \quad c_k = c_k(f) = (f, q_k)/\|q_k\|^2$$

and the value of the least squares approximation is

$$\|f - a^*\|^2 = \|f\|^2 - \textstyle\sum_{k=1}^n c_k^2 \, \|q_k\|^2$$

Proof. By Theorem 7.4, the best $a^* \in \mathcal{A}_n$ is determined by the conditions $(f, q_\ell) = (a^*, q_\ell)$, for all ℓ, hence, since (q_k) are orthogonal, we obtain

$$(f, q_\ell) = (a^*, q_\ell) = \textstyle\sum_{k=1}^n c_k (q_k, q_\ell) = c_\ell (q_\ell, q_\ell) = c_\ell \, \|q_\ell\|^2.$$

Further, by Corollary 7.5, we have $\|f - a^*\|^2 = \|f\|^2 - \|a^*\|^2$, and, by orthogonality of (q_k), $\|a^*\|^2 = \sum_{k=1}^n c_k^2 \, \|q_k\|^2$. ∎

Method 8.2 Suppose we are given an infinite orthogonal sequence $(q_k)_{k=1}^\infty$ and we want to approximate $f \in \mathbb{X}$ with a prescribed accuracy ε by a linear combination of the first terms q_k, i.e., find n and q_n^* such that

$$\|f - q_n^*\| \leq \varepsilon, \qquad q_n^* \in \mathcal{A}_n = \operatorname{span}(q_k)_{k=1}^n.$$

Firstly, it follows that, for any n, the optimal q_n^* equals $a^* = \sum_{k=1}^n \frac{(f,q_k)}{\|q_k\|} q_k$ and, secondly, 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

$$\sum_{k=1}^n \frac{(f, q_k)^2}{\|q_k\|^2} \geq \|f\|^2 - \varepsilon^2. \tag{8.1}$$

Example 8.3 For $\mathbb{X} = C(\mathbb{T})$, the space of continuous 2π-periodic functions, equipped with the scalar product $(f, g) := \int_{-\pi}^{\pi} f(t) g(t) \, dt$, the trigonometric functions $(\cos kx, \sin kx)_{k=0}^\infty$ form an orthogonal system, so that for $f \in C(\mathbb{T})$ its *n-th partial Fourier series* given by

$$s_n(f, x) := \tfrac{a_0}{2} + \textstyle\sum_{k=1}^n a_k \cos kx + b_k \sin kx, \quad \left(\begin{smallmatrix} a_k \\ b_k \end{smallmatrix}\right) := \tfrac{1}{\pi} \int_{\mathbb{T}} f(t) \left(\begin{smallmatrix} \cos kt \\ \sin kt \end{smallmatrix}\right) dt$$

is the best approximation to f from $\mathcal{T}_n$, the space of all trigonometric polynomial of degree n.

If we are given $\mathcal{A}_n = \operatorname{span}(a_k)$ with non-orthogonal (a_k), we can get an orthogonal basis (q_k) by the Gram-Schmidt algorithm. This procedure is not stable numerically, but generally it is the only one available. However, in some cases there exist better algorithms. For $\mathcal{A}_n \subset \mathbb{R}^m$, this was the QR factorization by means of Givens or Householder transfomations. Another case is construction of orthogonal algebraic polynomials to be considered next.

8.2 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

$$(f,g) := (f,g)_w := \int_a^b f(x)g(x)\,w(x)\,dx. \tag{8.2}$$

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 the *monic* polynomial. Given a scalar product (8.2), 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 8.4 *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*

$$p = \textstyle\sum_{k=0}^n c_k Q_k, \quad c_k = (p,Q_k)/\|Q_k\|^2. \tag{8.3}$$

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 (8.3) where each coefficient c_ℓ is uniquely determined by multiplying both sides with Q_ℓ. If $\widetilde{Q}_n$ is n-th orthogonal polynomial, then in its expansion (8.3) 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 8.5 (Three-term recurrence relation) *Monic orthogonal polynomials satisfy the relation*

$$Q_{n+1}(x) = (x - a_n)Q_n(x) - b_n Q_{n-1}(x), \qquad n = 0,1,\ldots \tag{8.4}$$

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 (8.3), 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 (8.4). ∎

Remark 8.6 If (Q_k) have leading coefficients (α_k), then the recurrence takes the form

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

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

8.3 Examples

Example 8.7 Classical examples of orthogonal 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 8.8 (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

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

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

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

8.4 Least squares polynomial fitting

Next theorem justifies Method 8.2 in the case of polynomial least squares approximation.

Theorem 8.9 (Parceval identity) *If $[a,b]$ is finite, then*

$$\textstyle\sum_{k=0}^{\infty} \frac{(f,Q_k)^2}{\|Q_k\|^2} = \|f\|^2.$$

Proof. Let

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

hence

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

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

Remark 8.10 Analogous indentity is valid for the trigonometric system.