Part III - Lent Term 2003

Approximation Theory – Lecture 21

21 Orthogonal spline projector

21.1 Least squares approximation

Definition 21.1 (Inner product space) A linear space $\mathbb{X}$ is called an *inner product space* if, for every f and g in $\mathbb{X}$, there is a scalar (f, g), called the *scalar product* that has the followig properties:

$$1)\ (f,g) = (g,f), \quad 2)\ (f,f) \ge 0 \text{ with equality iff } f = g, \quad 3)\ (f,g) \text{ is linear in both } f \text{ and } g.$$

One can deduce the well-known Cauchy-Schwarz and triangle inequalities

$$|(f,g)| \le (f,f)^{1/2}(g,g)^{1/2}, \qquad (f+g, f+g)^{1/2} \le (f,f)^{1/2} + (g,g)^{1/2}.$$

Thus, with the choice $\|f\| = (f,f)^{1/2}$, $\mathbb{X}$ becomes a normed linear space.

Theorem 21.2 *Let $\mathbb{X}$ be an inner product space, $\mathcal{U}_n$ be a subspace. Then $u^* \in \mathcal{U}_n$ is a best approximation to $f \in \mathbb{X}$ if and only if*

$$(f - u^*, v) = 0 \quad \forall v \in \mathcal{U}_n. \tag{21.1}$$

Proof. If (21.1) holds, then, for any $u \in \mathcal{U}_n$, letting $v := u^* - u$, we find that

$$\|f - u\|^2 = \|(f - u^*) + v\|^2 = \|f - u^*\|^2 + \|v\|^2 > \|f - u^*\|^2,$$

i.e., u^* is a b.a. Conversely, if $(f - u^*, v) \ne 0$ for some $v \in \mathcal{U}_n$, then with $\lambda = -\frac{(f-u^*,v)}{\|v\|^2}$ we obtain

$$\|(f - u^*) + \lambda v\|^2 = \|f - u^*\|^2 + 2\lambda(f - u^*, v) + \lambda^2\|v\|^2 = \|f - u^*\|^2 - \lambda^2\|v\|^2 < \|f - u^*\|^2,$$

i.e., u^* is not optimal. ∎

Corollary 21.3 *If $u^* \in \mathcal{U}_n$ is a best approximation to $f \in \mathbb{X}$, then*

$$\|f - u^*\|^2 + \|u^*\|^2 = \|f\|^2, \quad \textit{in particular} \quad \|u^*\| \le \|f\|,$$

the latter inequality being strict for $x \in \mathbb{X} \setminus \mathcal{U}_n$

Method 21.4 If u_i is a basis for $\mathcal{U}_n$ and if we write $u^* = \sum a_i u_i$, then running v in (21.1) through the basis functions u_i we obtain a linear system of equations for determining the coefficients a,

$$Ga = b, \quad G = [(u_i, u_j)]_{i,j=1}^n, \quad b = [(f, u_i)]_{i=1}^n.$$

These equations are called the *normal equations*. The matrix G is called the *Gram matrix*. Since the system is uniquely solvable, the Gram matrix G is invertible.

21.2 Smooth interpolation

Discussion 21.5 Let $\mathbb{X}$ be a Hilbert space (a complete inner product space), let $f \in \mathbb{X}$, and let the available information about f consists of interpolating data $\mu_i(f)$ with respect to some sequence of functionals $(\mu_i) \in \mathbb{X}^*$. We seek a (quasi-) interpolation u to f that satisfies the same interpolating conditions, but among many candidates we choose the "smoothest" one in the sense that it has the least norm

$$\widehat{u} = \arg\min \{\|u\| : \mu_i(u) = \mu_i(f)\}.$$

By Riesz theorem, any linear functional on the Hilbert space $\mathbb{X}$ can be represented as the scalar product,

$$\mu_i(f) = (f, u_i), \quad u_i \in \mathbb{X}.$$

Lemma 21.6 *Given $(\mu_i(f))$, the smoothest interpolant $\widehat{u}$ coincides with u^*, the best approximation to f from $\mathcal{U}_n := \mathrm{span}(u_i)$.*

Proof. The best approximation $u^* = \sum a_i u_i$ to f is a solution of the normal equations

$$Ga = b, \quad b = (f, u_i) = \mu_i(f), \tag{21.2}$$

and it satisfies the inequality $\|u^*\| \le \|f\|$. For any other interpolant g, with $\mu_i(g) = \mu_i(f)$, the right-hand side of (21.2) remains the same, therefore u^* is the best approximation to such a g, too, and thus $\|u^*\| \le \|g\|$ for any such g. Hence, u^* is the smoothest interpolant. ∎

Theorem 21.7 *Let $\Delta = (t_i)_{i=1}^{n+k}$ be a knot-sequence in $[a, b]$, and let the values (y_i) be given. Then a solution to the minimization problem*

$$s := \arg\min \{\|f^{(k)}\|_{L_2[a,b]} : f(t_i) = y_i\} \tag{21.3}$$

is a spline of degree $2k-1$ on the knot sequence Δ. In other words, among all interpolants to a given data the intepolating spline of degree $2k-1$ has the least L_2-norm of the k-th derivative.

Proof. From the sequence $(y_i)_{i=1}^{n+k}$ on Δ, let us form the sequence

$$z_i := y[t_i ... t_{i+k}], \quad i = 1...n,$$

of divided differences of order k, and let

$$\widetilde{s} := \arg\min \{\|f^{(k)}\|_{L_2[a,b]} : f[t_i ... t_{i+k}] = z_i\}.$$

Since both the kth derivative and a divided difference of order k vanish on polynomials of degree $k-1$, this $\widetilde{s}$ is determined up to such a polynomial. We claim that s, a solution to the original problem (21.3), is given by $s = \widetilde{s} - \ell_\delta(s) + \ell_\delta(y)$, where ℓ_δ are Lagrange interpolating polynomials of degree $k-1$ on $(t_1, ..., t_k)$. Indeed, by definition,

$$s(t_i) = y_i, \quad i = 1...k$$

and, because divided differences of order k vanish on polynomials, we also have

$$s[t_i ... t_{i+k}] = \widetilde{s}[t_i ... t_{i+k}] = z_i = y[t_i ... t_{i+k}].$$

For $[t_1 ... t_{k+1}]$, the latter equality combined with the previous ones gives $s(t_{k+1}) = y_{k+1}$, and increasing i one by one we obtain successively that $s(t_i) = y_i$ for all i. Since also $s^{(k)} = \widetilde{s}^{(k)}$, the value of $\|s^{(k)}\|$ is minimal.

2) Now recall that B-splines are Peano kernels for the divided diffrences

$$k!\, f[t_i ... t_{i+k}] = \int_a^b M_i(t) f^{(k)}(t)\, dt = (M_i, f^{(k)}),$$

i.e., with $g = f^{(k)}$ and $\sigma = \widetilde{s}^{(k)}$, we are looking for the solution

$$\sigma := \arg\min \{\|g\|_{L_2[a,b]} : (M_i, g) = k! z_i\}.$$

By the previous lemma, the minimal solution is a linear combination $\sum_i a_i M_i$. Hence,

$$\sigma = s^{(k)} \in \mathcal{S}_k(\Delta),$$

and respectively $s \in \mathcal{S}_{2k-1}(\Delta)$. ∎

21.3 Exercises

21.1. Prove that, for any basis (u_i) of $\mathcal{U}_n$, the Gram matrix $G = [(u_i, u_j)]$ is positive definite, i.e., $(Gx, x)_{\ell_2} > 0$ for any nonzero vector $x \in \mathbb{R}^n$.

Prove the converse: for any positive definite matrix $G \in \mathbb{R}^{n \times n}$, there exists a basis (u_i) of $\mathcal{U}_n$ such that $G = [(u_i, u_j)]$.

21.2. Prove that $\widehat{u}$ that interpolates $(\mu_i(f)$ is the smoothest innterpolant if and only if $\widehat{u}$ is orthogonal to $\ker \mu$, i.e.,

$$(\widehat{u}, g) = 0 \text{ for all } g \text{ such that } \mu_i(g) = 0.$$

21.3. (*Optimal recovery.*) Condiser the problem:

$$\text{given functionals } (\mu_i)_{i=1}^n \text{ and } \lambda\text{, and the values } \mu_i(f) = y_i \quad \to \quad \text{evaluate } \lambda(f)$$

(e.g., a quadrature rule $\lambda(f) = \int_a^b f(t)\,dt$ or a formula for numerical differentiation $\lambda(f) = f'(x_0)$). We expect that the value $\lambda(f)$ is bounded, so we need to put some restriction on f, say $\|f\| \le M$. Set

$$\mathcal{F} := \{f \in \mathbb{X} : \ \mu_i(f) = y_i, \ \|f\| \le M\}$$

Then the values of the functional λ on $\mathcal{F}$ lie inside a certain finite interval

$$a \le \inf_{f \in \mathcal{F}} \lambda(f) \le \sup_{f \in \mathcal{F}} \lambda(f) \le b.$$

Clearly, the optimal approximation (or, *recovery*) of λ on $\mathcal{F}$ is the value $\frac{a+b}{2}$.

Prove: the smoothest interpolant $\widehat{u}$ to (y_i) gives the optimal recovery to any λ, i.e.,

$$\lambda(\widehat{u}) = \tfrac{a+b}{2}.$$