Part III - Lent Term 2005
Approximation Theory – Lecture 8

8 Approximation in L_1 and L_2

8.1 Approximation of continuous functions in L_1-norm

Lemma 8.1 *Let $f, h \in C[a,b]$. If f has no more than a finite number of roots and if $\int h \,\mathrm{sgn}\, f > 0$, then, for some γ,*

$$\int |f - \gamma h| < \int |f|.$$

Proof. For $(x_i)_1^m$ being the roots of f, define the sets $B = \cup(x_i - \varepsilon, x_i + \varepsilon)$ and $A = [a,b] \setminus B$, and then select $\varepsilon > 0$ small enough to ensure that

$$\int_B |h| < \int_A h \,\mathrm{sgn}\, f. \tag{8.1}$$

Since $f \neq 0$ on A, and the set A is closed, the number $\delta := \min_{x \in A} |f(x)|$ is positive, hence, if we choose γ to satisfy $0 < \gamma \|h\| < \delta$, it will follow that $\mathrm{sgn}\,(f - \gamma h) = \mathrm{sgn}\, f$ on A. With that we obtain

$$\int_A |f - \gamma h| = \int_A (f - \gamma h)\,\mathrm{sgn}\, f = \int_A |f| - \int_A \gamma h \,\mathrm{sgn}\, f, \quad \text{while} \quad \int_B |f - \gamma h| \le \int_B |f| + \int_B \gamma |h|,$$

and summation of integrals on both sides of these relations gives

$$\int |f - \gamma h| \le \int |f| - \int_A \gamma h \,\mathrm{sgn}\, f + \int_B \gamma |h| \overset{(8.1)}{<} \int |f|. \qquad \square$$

Theorem 8.2 *Let U be a subspace and f an element of $C[a,b]$, and let $p^* \in U$ coincides with f in no more than a finite number of points. Then*

$$p^* \text{ is a b.a. to } f \text{ in } L_1\text{-norm} \quad \Leftrightarrow \quad \mathrm{sgn}\,(f - p^*) \perp U \tag{8.2}$$

Proof. If the condition fails, then $\int q \,\mathrm{sgn}\,(f - p^*) > 0$ for some $q \in \mathcal{U}$, and by the previous lemma there is a γ such that

$$\int |f - p^* - \gamma q| < \int |f - p^*|.$$

If the condition is fulfilled, then (no matter in how many points f and p^* coincide, and whether f is continuous or not), for any $p \in \mathcal{U}$,

$$\int |f - p| \ge \int (f - p)\,\mathrm{sgn}\,(f - p^*) = \int (f - p^*)\,\mathrm{sgn}\,(f - p^*) = \int |f - p^*|. \qquad \square$$

Theorem 8.3 *On the interval $[-1,1]$, among all polynomials of degree n with leading coefficient $a_n = 1$, the polynomial $\frac{1}{2^n(n+1)} T'_{n+1}$ deviates least from zero, i.e.,*

$$\inf_{(a_i)} \|x^n + a_{n-1}x^{n-1} + \cdots + a_0\|_1 = \tfrac{1}{2^n(n+1)} \|T'_{n+1}\|_1 = \tfrac{1}{2^{n-1}}.$$

Proof. Let p^* be the polynomial that interpolates $f(x) = x^n$ at the points $(x_k)_1^n$, where $x_k = \cos \frac{\pi k}{n+1}$ are zeros of T'_{n+1}. Comparing the leading coefficients we see that $f - p^* = \frac{1}{2^n(n+1)} T'_{n+1}$, therefore

$$\mathrm{sgn}\,[f(x) - p^*(x)] = \mathrm{sgn}\,\cos(n+1)\theta, \quad x = \cos\theta.$$

The latter function is orthogonal to $\mathcal{P}_{n-1}$ on $[-1,1]$ (exercise), hence by Theorem 8.2 the polynomial p^* is a best appoximant to $f(x) = x^n$. As to the value of best approximation, it follows from the relations $\|T'_{n+1}\|_1 = \mathrm{Var}\,[T_{n+1}] = 2(n+1)$. $\square$

Theorem 8.4 (Jackson [1921]) *Let $\mathcal{U}_n$ be a Chebyshev subspace of $C[a,b]$. Then each $f \in C[a,b]$ possesses a unique polynomial of best approximation in L_1-norm.*

Proof. 1) First of all, for a b.a. p^* to f from $\mathcal{U}_n$, the difference $f - p^*$ has at least $n+1$ zeros on (a,b). Otherwise, by Lemma 6.2, there exists a polynomial $q \in \mathcal{U}_n$ such that $\operatorname{sgn} q = \operatorname{sgn}(f - p^*)$, hence $\int q \operatorname{sgn}(f-p^*) > 0$, a contradiction to Theorem 8.2.

2) Assume now that, for a function f, there are two polynomials of best approximation, p and q. By the triangle inequality, the polynomial $r := \frac{1}{2}(p+q)$ is a best approximation, too, whence

$$\int \left(|f-r| - \tfrac{1}{2}|f-p| - \tfrac{1}{2}|f-q|\right) = 0.$$

Since the integrand is continious and ≤ 0, it must vanish identically on $[a,b]$, so, at the points where $f(x) = r(x)$, we have $f(x) = p(x) = q(x)$. But as we showed above, the number of these points is at least $n+1$, hence, by the Haar condition, p and q are identical. □

Remark 8.5 The analogue of the Haar unicity theorem for approximation in L_1-norm is not true, i.e., there are non-Chebyshev subspaces of $C[a,b]$ which provide uniqueness of best approximation to any $f \in C[a,b]$.

8.2 Least squares approximation

Definition 8.6 (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) $(f,g) = (g,f)$, 2) $(f,f) \ge 0$ with equality iff $f = g$, 3) (f,g) is linear in both f 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 8.7 *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*

$$(f - u^*, v) = 0 \quad \forall v \in \mathcal{U}_n. \tag{8.3}$$

Proof. If (8.3) 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 8.8 *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 8.9 If u_i is a basis for $\mathcal{U}_n$ and if we write $u^* = \sum a_i u_i$, then running v in (8.3) 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.

Theorem 8.10 *On the interval $[-1,1]$, among all polynomials of degree n with leading coefficient $a_n = 1$, the Legendre polynomial $c_n P_n$, where*

$$P_n(x) = \frac{1}{2^n n!}\frac{d}{dx^n}(x^2-1)^n,$$

deviates least from zero, i.e.,

$$\inf_{(a_i)} \|x^n + a_{n-1}x^{n-1} + \cdots + a_0\|_1 = c_n\,\|P_n\|_2 = \frac{2^n (n!)^2}{(2n)!}\frac{1}{\sqrt{n+1/2}}.$$

8.3 Exercises

8.1. Prove the orthogonality condition used in the proof of Theorem 8.3:

$$\int_{-1}^{1} p(x)\,\mathrm{sgn}\,T'_{n+1}(x)\,dx = 0 \quad \forall\, p \in \mathcal{P}_{n-1}\,.$$

Hint. Make a substitution $x = \cos\theta$ and use Lemma 4.10.

8.2. Prove the following generalization of Theorem 8.3: if $f^{(n)} > 0$ on $[-1,1]$, then p^*, the polynomial of b.a. to f from $\mathcal{P}_{n-1}$ in L_1-norm, is the Lagrange polynomial that interpolates f at n points $(\cos\frac{\pi k}{n+1})_{k=1}^{n}$.

8.3. Show that orthogonality condition (8.2) is not necessary, i.e., construct an $f \in C[a,b]$ and a subspace U such that, for the best (unique) approximation p^*, condition (8.2) is not fulfilled.

8.4. 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 $G = [(u_i, u_j)]$.

8.5. Prove that the value of the least squares approximation to $f \in \mathbb{X}$ from $\mathcal{U}_n = \mathrm{span}(u_i)$ is given by

$$E(f)^2 = \frac{\det G(f, u_1, ..., u_n)}{\det G(u_1, ..., u_n)}\,,$$

where G-s the corresponding Gram-matrices.

Hint. We have $E(f)^2 = (f,f) - (f,u^*)$ and, if $u^* = \sum a_i u_i$, this is equivalent to

$$\sum_{i=1}^{n} a_i (u_i, f) = (f,f) - E(f)^2.$$

Join this equation to the normal equations and consider the resulting singular system.

8.6. Deduce from the previous exercise that

$$\det G(u_1, ..., u_n) \le (u_1, u_1)\cdots(u_n, u_n)\,,$$

and, more generally, that the determinant of any positive definite matrix is not greater that the product of its diagonal elements. Hence, derive the Hadamard inequality: for any matrix $A = (a_{ij})$

$$|\det A| \le \prod_{i=1}^{n} \Big(\sum_{j=1}^{n} |a_{ij}|^2\Big)^{1/2}.$$

8.7. Let G be the class of functions g in $W_2^k[a,b]$ such that

$$g^{(r)}(a) = g^{(r)}(b) = 0, \quad r = 0...k-1\,.$$

Prove that $f \in L_2[a,b]$ is orthogonal to all polynomials of degree $k-1$ if and only if $f = g^{(k)}$ for some $g \in G$. Hence, derive Theorem 8.10.