Part III - Lent Term 2005

Approximation Theory – Lecture 1

1 Basic concepts

1.1 Approximation and best approximation

Let $\mathbb{X}$ be a *metric space*, which means that, for every f and g in $\mathbb{X}$, there exists a real number $d(f,g)$ that has the following properties:

$$1)\ d(f,g) = d(g,f), \quad 2)\ d(f,g) \ge 0 \text{ with equality iff } f = g, \quad 3)\ d(f,h) \le d(f,g) + d(g,h).$$

We call d the *distance function*.

Definition 1.1 (*Approximation scheme*). An approximation method requires a set of approximating functions, $\mathcal{U}$ say, which is a subset of $\mathbb{X}$. Specifically, the method is just a mapping from $\mathbb{X}$ to $\mathcal{U}$. In other words, given any $f \in \mathbb{X}$, the method picks the element u_f, say, from $\mathcal{U}$, which is regarded as an approximation to f. To find whether the method is good or bad, it should be compared with the best approximation.

Definition 1.2 (*Best approximation*). A *best approximation* to $f \in \mathbb{X}$ from $\mathcal{U}$ is an element $u^* \in \mathcal{U}$ s.t.

$$d(f,u^*) = \inf\{d(f,u) : u \in \mathcal{U}\} =: \operatorname{dist}(f,\mathcal{U}).$$

Question 1.3 Regarding the element of best approximation, basic questions are (i) existence, (ii) uniqueness, (iii) characterization: how one can recognize u^* other than comparing it with all other elements of $\mathcal{U}$, this is important for (iv) construction.

Often too much work would be needed to calculate a best approximation, so we are interested in "good" methods that choose u_f so that $d(f,u_f)$ is close to $\operatorname{dist}(f,\mathcal{U})$, and which are "cheap".

Usually $\mathbb{X}$ is a *normed linear space*, $\mathcal{U}$ its linear subspace, and $d(f,g) = \|f - g\|$. In this case linear approximation methods are "good" in a sense.

1.2 Linear approximation and projection

Definition 1.4 Let M be the mapping from $\mathbb{X}$ to $\mathcal{U}$ that is mentioned in §1.1. It is a *linear approximation method* if the following conditions hold: (i) $\mathbb{X}$ is a linear space and $\mathcal{U}$ is a linear subspace of $\mathbb{X}$, and (ii) the mapping enjoys the property

$$M(\alpha f + \beta g) = \alpha M(f) + \beta M(g),$$

for every f and g in $\mathbb{X}$ and every scalars α and β. The mapping is a *projection method* if it satisfies the condition $M(u) = u$, $\forall u \in \mathcal{U}$.

Example 1.5 (*Polynomial interpolation on* $[a,b]$). Let $\mathbb{X}$ be $C[a,b]$ and let $\mathcal{U}$ be $\mathcal{P}_n$, which are the linear spaces of continuous functions from $[a,b]$ to $\mathbb{R}$ and of polynomials of degree at most n, respectively. Then the following approximation method is useful. We pick points x_i, $i = 0..n$, that satisfy $a \le x_0 < x_1 < \cdots < x_n \le b$, and, given any $f \in C[a,b]$, we let $p = M(f)$ be the polynomial in $\mathcal{P}_n$ that satisfies the conditions

$$p(x_i) = f(x_i), \qquad i = 0..n.$$

Thus M is a linear approximation method that is also a projection, the solution of the interpolation problem being unique. It is often written in the *Lagrange* form

$$p(x) = \sum_{i=0}^{n} \ell_i(x)\, f(x_i), \qquad x \in [a,b],$$

where ℓ_i, $i = 0..n$, is the polynomial $\ell_i(x) = \prod_{j=0, j\ne i}^{n} \{(x - x_j)/(x_i - x_j)\}$, so that $\ell_i(x_j) = \delta_{ij}$.

Definition 1.6 (*Norms of mappings*). Let $\mathbb{X}$ be a normed linear space (where the norm provides the distance function). Then, if M is a linear mapping from $\mathbb{X}$ to $\mathcal{U}$, its *norm* is defined to be the number

$$\|M\| := \sup\{\, \|M(f)\| : f \in \mathbb{X},\ \|f\| = 1\}.$$

It follows from linearity of the norm that the inequality $\|M(f)\| \le \|M\|\,\|f\|$ holds for every $f \in \mathbb{X}$.

Example 1.7 (*A norm for polynomial interpolation*). In Example 1.5, we can pick the function norm

$$\|f\|_\infty := \sup_{x\in[a,b]} |f(x)|\,,$$

and it gives the mapping norm

$$\begin{aligned} \|M\| &:= \sup\{\, |\textstyle\sum_{i=0}^n \ell_i(x)\, f(x_i)| : x \in [a,b],\ \|f\|_\infty = 1\} \\ &= \sup\{\, \textstyle\sum_{i=0}^n |\ell_i(x)| : x \in [a,b]\}. \end{aligned}$$

This assertion is proved by choosing, for any x, any f that satisfies $\|f\|_\infty = 1$ and $\ell_i(x)f(x_i) = |\ell_i(x)|$, $i = 0..n$.

Theorem 1.8 (Lebesgue[1]inequality) *Let $\mathbb{X}$ be a normed linear space, and let $M : \mathbb{X} \to \mathcal{U}$ be a linear projection. Then, for every $f \in \mathbb{X}$, the approximation $M(f)$ has the property*

$$\|f - M(f)\| \le (\|M\| + 1)\,\mathrm{dist}(f, \mathcal{U}).$$

Remark 1.9 The number $\|M\|$ is called the *Lebesgue constant* of the approximation method.

Proof. For f given, let u be any element of $\mathcal{U}$. Then, because M is a linear and $M(u) = u$, we have

$$\begin{aligned} \|f - M(f)\| &= \|(f-u) - M(f-u)\| = \|(I - M)(f-u)\| \\ &\le \|I - M\|\,\|f - u\| \le (\|M\| + 1)\|f - u\|. \end{aligned}$$

Now take the infimum over $u \in \mathcal{U}$. ■

Remark 1.10 With respect to the Lebesgue inequality, it is suggestive to look for the *best* linear projection from $\mathbb{X}$ to $\mathcal{U}$, i.e. the projection M with minimal norm.

Example 1.11 (*Lebesgue constant of polynomial interpolation*). Of course it depends on the positions of the interpolation points. For example, if $n = 20$, then the equally spaced points $x_i = i/10$, $i = -10..10$, provide $\|M\| = 10986.71$, but the choice $x_i = \cos(\pi i/10)$, gives $\|M\| = 2.48$. Thus a polynomial interpolation method, when $n = 20$, can calculate an approximation $p_f \approx f$ with the property that $\|f - p_f\|_\infty$ is guaranteed to be within a factor of 3.48 of the least maximum error, namely $\mathrm{dist}(f, \mathcal{P}_{20})$.

1.3 Degree of approximation

Suppose we are given a sequence $(\mathcal{U}_n)$ of subsets of $\mathbb{X}$. Degree of approximation considers the behaviour of $E_n(f) := \mathrm{dist}(f, \mathcal{U}_n)$ as a function of the parameter n as $n \to \infty$. The first question is whether

$$E_n(f) \overset{?}{\to} 0 \quad (n \to \infty),$$

i.e. whether our choice of approx. sets $(\mathcal{U}_n)$ is reasonable at all. The next question of interest is just how 'fast' or 'slow' this convergence is for particular f's. These terms are made precise by comparing $E_n(f)$ with certain standard sequences, e.g., $(n^{-\alpha})$. With $\mathbb{Y} \subset \mathbb{X}$, a class of functions sharing with particular f certain characteristics (e.g., $\|f''\| \le 1$), one distinguishes

(i) Jackson-type (or, direct) theorems: $f \in \mathbb{Y} \;\Rightarrow\; E_n(f) = \mathcal{O}(n^{-\alpha})$,

(ii) Bernstein-type (or, inverse) theorems: $E_n(f) = \mathcal{O}(n^{-\alpha}) \;\Rightarrow\; f \in \mathbb{Y}$.

[1]Henri Lebesgue, 1875-1941, his measure theory and the Lebesgue integral, which form a basis of the modern analysis, appeared in his PhD thesis in 1902. It is interesting that Hilbert, while formulating the needs of mathematics in the 20th century in his famous problems of 1900, did not mention theory of integral.

1.4 Exercises

1.1. Prove that the norm of a linear projector M satisfies $\|M\| \ge 1$. By considering the linear interpolation $M : C[a,b] \to \mathcal{P}_1$ at the end-points of $[a,b]$ show that, for the linear projectors, both inequalities

$$\|M\|_\infty \ge 1, \qquad \|f - M(f)\|_\infty \le (\|M\|_\infty + 1)\,\mathrm{dist}(f, \mathcal{U})$$

may become equalities (i.e., they are *sharp*).

1.2. An approximation $M : \mathbb{X} \to \mathcal{U}$ is called *near-best*, if there is a constant $\gamma > 0$ such that $\|f - M(f)\| \le \gamma\,\mathrm{dist}(f, \mathcal{U})$, $\forall f \in \mathbb{X}$. Prove that the near-best approximation M is a projection.

1.3. Let $M : C[a,b] \to \mathcal{P}_n$ be the interpolation at some given points $a \le x_0 < \cdots < x_n \le b$, and let $\ell_i(x) = \prod_{j \ne i} \frac{x - x_j}{x_i - x_j}$ be the corresponding fundamental polynomials. The function

$$L(x) := \textstyle\sum_{i=0}^n |\ell_i(x)|$$

is called the *Lebesgue* function (so that, from Example 1.7, $\|L\|_\infty = \|M\|_\infty$). Prove that, on $[x_k, x_{k+1}]$, the Lebesgue function L coincides with the polynomial $q \in \mathcal{P}_n$ such that

$$|q(x_i)| = 1 \quad \text{all } i, \qquad q(x_k) = q(x_{k+1}) = 1, \qquad q(x_i) = -q(x_{i+1}) \quad \text{for} \quad i \ne k,$$

while on $[a, x_0]$ and on $[x_n, b]$ it is the polynomial $q \in \mathcal{P}_n$ that satisfies

$$|q(x_i)| = 1 \quad \text{all } i, \qquad q(x_i) = -q(x_{i+1}) \quad \text{all } i.$$

1.4. Use the previous result to prove that, for the quadratic interpolation M, the minimal norm of M is not smaller than $5/4$. Further, find all positions of the interpolation points that satisfy $-1 \le x_0 < x_1 < x_2 \le 1$ and that provide this minimal value $\|M\|_\infty = 5/4$.

Hint. Prove that, for any x_i, the Lebesgue function on $[x_0, x_2]$ satisfies $\max L(x) \ge 5/4$, and for those x_i for which there is equality sign investigate $L(x)$ on the rest of the interval.