Approximation Theory

Department of Mathematics

Fourth Class

Introduction

In 1853, the great Russian mathematician, P. L. Chebyshev (Čebyšev), while working on a problem of *linkages*, devices which translate the linear motion of a steam engine into the circular motion of a wheel, considered the following problem:

> Given a continuous function f defined on a closed interval $[a, b]$ and a positive integer n, can we "represent" f by a polynomial $p(x) = \sum_{k=0}^{n} a_k x^k$, of degree at most n, in such a way that the maximum error at any point x in $[a, b]$ is controlled? In particular, is it possible to construct p so that the error $\max_{a \le x \le b} |f(x) - p(x)|$ is minimized?

This problem raises several questions, the first of which Chebyshev himself ignored:

– Why should such a polynomial even *exist*?

– If it does, can we hope to *construct* it?

– If it exists, is it also *unique*?

– What happens if we change the measure of error to, say, $\int_a^b |f(x) - p(x)|^2 \, dx$?

Best Approximations in Normed Spaces

Recall that a norm on a vector space X is a nonnegative function on X satisfying:

$$\|x\| \geq 0, \text{ and } \|x\| = 0 \text{ if and only if } x = 0,$$

$$\|\alpha x\| = |\alpha|\|x\| \text{ for any } x \in X \text{ and } \alpha \in \mathbb{R},$$

$$\|x + y\| \leq \|x\| + \|y\| \text{ for any } x,\, y \in X.$$

Examples

1. As we'll soon see, in $X = \mathbb{R}^n$ with its usual norm $\|(x_k)_{k=1}^n\|_2 = \left(\sum_{k=1}^n |x_k|^2\right)^{1/2}$,

2. Consider $X = \mathbb{R}^2$ under the norm $\|(x, y)\| = \max\{|x|, |y|\}$,

3. There are many norms we might consider on $\mathbb{R}^n$. Of particular interest are the ℓ_p-norms; that is, the scale of norms:

$$\|(x_i)_{i=1}^n\|_p = \left(\sum_{k=1}^n |x_k|^p \right)^{1/p}, \quad 1 \leq p < \infty,$$

and

$$\|(x_i)_{i=1}^n\|_\infty = \max_{1 \leq i \leq n} |x_i|.$$

It's easy to see that $\|\cdot\|_1$ and $\|\cdot\|_\infty$ define norms. The other cases take a bit more

4. The ℓ_2-norm is an example of a norm induced by an *inner product* (or "dot" product). You will recall that the expression

$$\langle\, x, y \,\rangle = \sum_{i=1}^n x_i y_i,$$

where $x = (x_i)_{i=1}^n$ and $y = (y_i)_{i=1}^n$, defines an inner product on $\mathbb{R}^n$ and that the norm in $\mathbb{R}^n$ satisfies

$$\|x\|_2 = \sqrt{\langle\, x, x \,\rangle}\,.$$

one example, consider this: Given a positive Riemann integrable *weight function* $w(x)$ defined on some interval $[a, b]$, it's not hard to check that the expression

$$\langle f, g \rangle = \int_a^b f(t)\, g(t)\, w(t)\, dt$$

defines an inner product on $C[a, b]$, the space of all continuous, real-valued functions $f : [a, b] \to \mathbb{R}$, with associated norm

$$\| f \|_2 = \left(\int_a^b |f(t)|^2 w(t)\, dt \right)^{1/2}.$$

5. Our original problem concerns the space $X = C[a, b]$ under the *uniform norm* $\| f \| = \max_{a \le x \le b} |f(x)|$. The adjective "uniform" is used here because convergence in this norm is the same as uniform convergence on $[a, b]$:

$$\| f_n - f \| \to 0 \iff f_n \to f \text{ uniformly on } [a, b]$$

Lemma 1.3. *Let V be a finite-dimensional vector space. Then, all norms on V are equivalent. That is, if $\|\cdot\|$ and $|||\cdot|||$ are norms on V, then there exist constants $0 < A, B < \infty$ such that*

$$A\|x\| \leq |||x||| \leq B\|x\|$$

for all vectors $x \in V$.

Corollary 1.4. *Every finite-dimensional normed space is complete (that is, every Cauchy seqeuence converges). In particular, if Y is a finite-dimensional subspace of a normed linear space X, then Y is a closed subset of X.*

Corollary 1.5. *Let Y be a finite-dimensional normed space, let $x \in Y$, and let $M > 0$. Then, any closed ball $\{ y \in Y : \|x - y\| \leq M \}$ is compact.*

Proof. Because translation is an isometry, it clearly suffices to show that the set $\{ y \in Y : \|y\| \leq M \}$ (i.e., the ball about 0) is compact.

Suppose now that Y is n-dimensional and that $e_1, \ldots, e_n$ is a basis for Y. From Lemma 1.3 we know that there is some constant $A > 0$ such that

$$A \sum_{i=1}^{n} |a_i| \le \left\| \sum_{i=1}^{n} a_i e_i \right\|$$

for all $x = \sum_{i=1}^{n} a_i e_i \in Y$. In particular,

$$A\,|a_i| \le \left\| \sum_{i=1}^{n} a_i e_i \right\| \le M \implies |a_i| \le M/A \quad \text{for} \quad i = 1, \ldots, n.$$

Thus, $\{\, y \in Y : \|y\| \le M \,\}$ is a *closed* subset (why?) of the *compact* set

$$\left\{ x = \sum_{i=1}^{n} a_i e_i : |a_i| \le M/A,\ i = 1, \ldots, n \right\} = [\,-M/A,\ M/A\,]^n. \qquad \square$$

Theorem 1.6. *Let Y be a finite-dimensional subspace of a normed linear space X, and let $x \in X$. Then, there exists a (not necessarily unique) vector $y^* \in Y$ such that*

$$\|x - y^*\| = \min_{y \in Y} \|x - y\|$$

for all $y \in Y$. That is, there is a best approximation to x by elements from Y.

Proof. First notice that because $0 \in Y$, we know that any nearest point y^* will satisfy $\|x - y^*\| \leq \|x\| = \|x - 0\|$. Thus, it suffices to look for y^* in the *compact* set

$$K = \{\, y \in Y : \|x - y\| \leq \|x\,\|\}.$$

To finish the proof, we need only note that the function $f(y) = \|x - y\|$ is *continuous*:

$$|f(y) - f(z)| = \Big|\, \|x - y\| - \|x - z\| \,\Big| \leq \|y - z\|,$$

and hence attains a minimum value at some point $y^* \in K$. □

Corollary 1.7. *For each $f \in C[a,b]$ and each positive integer n, there is a (not necessarily unique) polynomial $p_n^* \in \mathcal{P}_n$ such that*

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

Lemma 1.9. *Let Y be a finite-dimensional subspace of a normed linear space X, and suppose that each $x \in X$ has a unique nearest point $y_x \in Y$. Then the nearest point map $x \mapsto y_x$ is continuous.*

Proof. Let's write $P(x) = y_x$ for the nearest point map, and let's suppose that $x_n \to x$ in X. We want to show that $P(x_n) \to P(x)$, and for this it's enough to show that there is a subsequence of $(P(x_n))$ that converges to $P(x)$. (Why?)

Because the sequence (x_n) is bounded in X, say $\|x_n\| \le M$ for all n, we have

$$\|P(x_n)\| \le \|P(x_n) - x_n\| + \|x_n\| \le 2\|x_n\| \le 2M.$$

Thus, $(P(x_n))$ is a bounded sequence in Y, a finite-dimensional space. As such, by passing to a subsequence, we may suppose that $(P(x_n))$ converges to some element $P_0 \in Y$. (How?) Now we need to show that $P_0 = P(x)$. But

$$\|P(x_n) - x_n\| \le \|P(x) - x_n\|$$

for any n. (Why?) Hence, letting $n \to \infty$, we get

$$\|P_0 - x\| \le \|P(x) - x\|.$$

Because nearest points in Y are unique, we must have $P_0 = P(x)$. □

Theorem 1.11. *Let Y be a subspace of a normed linear space X, and let $x \in X$. The set Y_x, consisting of all best approximations to x out of Y, is a bounded* convex *set.*

Proof. As we've seen, the set Y_x is a subset of the ball $\{ y \in X : \|x-y\| \le \|x\| \}$ and, as such, is bounded. (More generally, the set Y_x is a subset of the sphere $\{ y \in X : \|x-y\| = d \}$, where $d = \text{dist}(x, Y) = \inf_{y \in Y} \|x-y\|$.)

Next recall that a subset K of a vector space V is said to be *convex* if K contains the line segment joining any pair of its points. Specifically, K is convex if

$$x, y \in K,\ 0 \le \lambda \le 1 \implies \lambda x + (1-\lambda)y \in K.$$

Thus, given y_1, $y_2 \in Y_x$ and $0 \le \lambda \le 1$, we want to show that the vector $y^* = \lambda y_1 + (1-\lambda) y_2 \in Y_x$. But y_1, $y_2 \in Y_x$ means that

$$\|x - y_1\| = \|x - y_2\| = \min_{y \in Y} \|x - y\|.$$

Hence,

$$\begin{aligned} \|x - y^*\| &= \|x - (\lambda y_1 + (1-\lambda) y_2)\| \\ &= \|\lambda(x - y_1) + (1-\lambda)(x - y_2)\| \\ &\le \lambda\|x - y_1\| + (1-\lambda)\|x - y_2\| \\ &= \min_{y \in Y} \|x - y\|. \end{aligned}$$

Consequently, $\|x - y^*\| = \min_{y \in Y} \|x - y\|$; that is, $y^* \in Y_x$. □

Corollary 1.13. *If X has a strictly convex norm, then, for any subspace Y of X and any point $x \in X$, there can be at most one best approximation to x out of Y. That is, Y_x is either empty or consists of a single point.*

In order to arrive at a condition that's somewhat easier to check, let's translate our original definition into a statement about the triangle inequality in X.

Lemma 1.14. *A normed space X has a strictly convex norm if and only if the triangle inequality is strict on nonparallel vectors; that is, if and only if*

$$x \neq \alpha y, \; y \neq \alpha x, \; \textit{all } \alpha \in \mathbb{R} \implies \|x+y\| < \|x\| + \|y\|.$$

Examples 1.15.

1. The usual norm on $C[\,a,b\,]$ is *not* strictly convex (and so the problem of uniqueness of best approximations is all the more interesting to tackle). For example, if $f(x) = x$ and $g(x) = x^2$ in $C[\,0,1\,]$, then $f \neq g$ and $\|f\| = 1 = \|g\|$, while $\|f+g\| = 2$. (Why?)

2. The usual norm on $\mathbb{R}^n$ is strictly convex, as is any one of the norms $\|\cdot\|_p$ for $1 < p < \infty$. (See Problem 10.) The norms $\|\cdot\|_1$ and $\|\cdot\|_\infty$, on the other hand, are *not* strictly convex. (Why?)