Chapter 1

Preliminaries

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, $\int_a^b |f(x) - p(x)|^2\,dx$?

Exercise 1.1. How do we know that $C[\,a, b\,]$ contains non-polynomial functions? Name one (and explain why it isn't a polynomial)!

Best Approximations in Normed Spaces

Chebyshev's problem is perhaps best understood by rephrasing it in modern terms. What we have here is a problem of *best approximation* in a *normed linear space*. Recall that a *norm* on a (real) *vector space* X is a nonnegative function on X satisfying

$\|x\| \ge 0$, and $\|x\| = 0$ if and only if $x = 0$,

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

$\|x + y\| \le \|x\| + \|y\|$ for any $x,\, y \in X$.

Any norm on X induces a *metric* or distance function by setting $\operatorname{dist}(x, y) = \|x - y\|$. The abstract version of our problem(s) can now be restated:

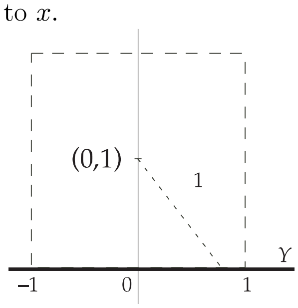
> Given a subset (or even a *subspace*) Y of X and a point $x \in X$, is there an element $y \in Y$ that is *nearest* to x? That is, can we find a vector $y \in Y$ such that $\|x - y\| = \min_{z\in Y} \|x - z\|$? If there is such a *best approximation* to x from elements of Y, is it unique?

It's not hard to see that a satisfactory answer to this question will require that we take Y to be a *closed* set in X, for otherwise points in $\overline{Y} \setminus Y$ (sometimes called the *boundary* of the set Y) will not have nearest points. Indeed, which point in the interval $[0, 1)$ is nearest to 1? Less obvious is that we typically need to impose additional requirements on Y in order to insure the existence (and certainly the uniqueness) of nearest points. For the time being, we will consider the case where Y is a closed *subspace* of a normed linear space X.

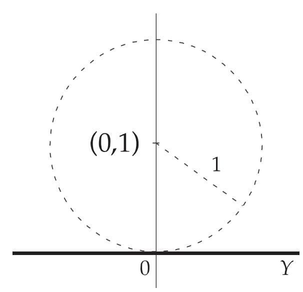
Examples 1.2.

1. As we'll soon see, in $X = \mathbb{R}^n$ with its usual norm $\|(x_k)_{k=1}^n\|_2 = \left(\sum_{k=1}^n |x_k|^2\right)^{1/2}$, the problem has a complete solution for any subspace (or, indeed, any *closed convex set*) Y. This problem is often considered in calculus or linear algebra where it is called "least-squares" approximation. A large part of the current course will be taken up with least-squares approximations, too. For now let's simply note that the problem changes character dramatically if we consider a different norm on $\mathbb{R}^n$, as evidenced by the following example.

2. Consider $X = \mathbb{R}^2$ under the norm $\|(x, y)\| = \max\{|x|, |y|\}$, and consider the subspace $Y = \{(0, y) : y \in \mathbb{R}\}$ (i.e., the y-axis). It's not hard to see that the point $x = (1, 0) \in \mathbb{R}^2$ has infinitely many nearest points in Y; indeed, every point $(0, y)$, $-1 \le y \le 1$, is nearest to x.



sphere of radius 1, max norm



sphere of radius 1, usual norm

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:

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

and

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

It's easy to see that $\|\cdot\|_1$ and $\|\cdot\|_\infty$ define norms. The other cases take a bit more work; for full details see Appendix A.

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

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

In this sense, the usual norm on $\mathbb{R}^n$ is actually induced by the inner product. More generally, any inner product will give rise to a norm in this same way. (But not vice versa. As we'll see, inner product norms satisfy a number of special properties that aren't enjoyed by all norms.)

The presence of an inner product in an abstract space opens the door to geometric arguments that are remarkably similar to those used in $\mathbb{R}^n$. (See Appendix D for more details.) Luckily, inner products are easy to come by in practice. By way of 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

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

We will take full advantage of this fact in later chapters (in particular, Chapters 8–9).

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 $[\,a, b\,]$:

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

(which we will frequently abbreviate by writing $f_n \rightrightarrows f$ on $[\,a, b\,]$). This, by the way, is the norm of choice on $C[\,a, b\,]$ (largely because continuity is preserved by uniform limits). In this particular case we're interested in approximations by elements of $Y = \mathcal{P}_n$, the *subspace* of all polynomials of degree at most n in $C[\,a, b\,]$. It's not hard to see that $\mathcal{P}_n$ is a finite-dimensional subspace of $C[\,a, b\,]$ of dimension exactly $n + 1$. (Why?)

6. If we consider the subspace $Y = \mathcal{P}$ consisting of *all* polynomials in $X = C[\,a, b\,]$, we readily see that the existence of best approximations can be problematic. It follows from the Weierstrass theorem, for example, that each $f \in C[\,a, b\,]$ has distance 0 from $\mathcal{P}$ but, because not every $f \in C[\,a, b\,]$ is a polynomial (why?), we can't hope for a best approximating polynomial to exist in every case. For example, the function

$f(x) = x\sin(1/x)$ is continuous on $[\,0,1\,]$ but can't possibly agree with any polynomial on $[\,0,1\,]$. (Why?) As you may have already surmised, the problem here is that every element of $C[\,a,b\,]$ is the (uniform) limit of a sequence from $\mathcal{P}$; in other words, the closure of Y equals X; in symbols, $\overline{Y} = X$.

Finite-Dimensional Vector Spaces

The key to the problem of polynomial approximation is the fact that each of the spaces $\mathcal{P}_n$, described in Examples 1.2 (5), is *finite-dimensional*. To see how finite-dimensionality comes into play, it will be most efficient to consider the abstract setting of finite-dimensional subspaces of arbitrary normed spaces.

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*

$$A\,\|x\| \le |||x||| \le B\,\|x\|$$

for all vectors $x \in V$.

Proof. Suppose that V is n-dimensional and that $\|\cdot\|$ is a norm on V. Fix a basis $e_1,\ldots,e_n$ for V and consider the norm

$$\left\|\sum_{i=1}^n a_i e_i\right\|_1 = \sum_{i=1}^n |a_i| = \|(a_i)_{i=1}^n\|_1$$

for $x = \sum_{i=1}^n a_i e_i \in V$. Because $e_1,\ldots,e_n$ is a basis for V, it's not hard to see that $\|\cdot\|_1$ is, indeed, a norm on V. [Notice that we've actually set-up a correspondence between $\mathbb{R}^n$ and V; specifically, the map $(a_i)_{i=1}^n \mapsto \sum_{i=1}^n a_i e_i$ is obviously both one-to-one and onto. In fact, this correspondence is an *isometry* between $(\mathbb{R}^n, \|\cdot\|_1)$ and $(V, \|\cdot\|_1)$.]

It now suffices to show that $\|\cdot\|$ and $\|\cdot\|_1$ are equivalent. (Why?)

One inequality is easy to show; indeed, notice that

$$\left\|\sum_{i=1}^n a_i e_i\right\| \le \sum_{i=1}^n |a_i|\,\|e_i\| \le \left(\max_{1\le i\le n}\|e_i\|\right)\sum_{i=1}^n |a_i| = B\left\|\sum_{i=1}^n a_i e_i\right\|_1.$$

The real work comes in establishing the other inequality.

Now the inequality we've just established shows that the function $x \mapsto \|x\|$ is *continuous* on the space $(V, \|\cdot\|_1)$; indeed,

$$\big|\,\|x\| - \|y\|\,\big| \le \|x-y\| \le B\,\|x-y\|_1$$

for any x, $y \in V$. Thus, $\|\cdot\|$ assumes a *minimum* value on the *compact* set

$$S = \{x \in V : \|x\|_1 = 1\}.$$

(Why is S compact?) In particular, there is some $A > 0$ such that $\|x\| \ge A$ whenever $\|x\|_1 = 1$. (Why can we assume that $A > 0$?) The inequality we need now follows from the homogeneity of the norm:

$$\left\|\frac{x}{\|x\|_1}\right\| \ge A \Longrightarrow \|x\| \ge A\,\|x\|_1. \qquad \square$$

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 $\{ y \in Y : \|x - y\| \le M \}$ is compact.*

Proof. Because translation is an isometry, it clearly suffices to show that the set $\{ y \in Y : \|y\| \le M \}$ (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 $\|x - y^*\| \le \|x\| = \|x - 0\|$. Thus, it suffices to look for y^* in the *compact* set

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

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

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

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

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*

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

Example 1.8. Nothing in Corollary 1.7 says that p_n^* will be a polynomial of degree *exactly* n—rather, it's a polynomial of degree *at most* n. For example, the best approximation to $f(x) = x$ by a polynomial of degree at most 3 is, of course, $p(x) = x$. Even examples of nonpolynomial functions are easy to come by; for instance, the best linear approximation to $f(x) = |x|$ on $[-1, 1]$ is actually the constant function $p(x) = 1/2$, and this makes for an entertaining **exercise**.

Before we leave these "soft" arguments behind, let's discuss the problem of *uniqueness* of best approximations. First, let's see why we might like to have unique best approximations:

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 $P_0 = P(x)$. □

Exercise 1.10. Let X be a metric (or normed) space and let $f : X \to X$. Show that f is continuous at $x \in X$ if and only if, whenever $x_n \to x$ in X, some subsequence of $(f(x_n))$ converges to $f(x)$. [Hint: The forward direction is easy; for the backward implication, suppose that $(f(x_n))$ fails to converge to $f(x)$ and work toward a contradiction.]

It should be pointed out that the nearest point map is, in general, *nonlinear* and, as such, can be very difficult to work with. Later we'll see at least one case in which nearest point maps always turn out to be linear.

In spite of any potential difficulties with the nearest point map, we next observe that the set of best approximations has a well-behaved, almost-linear structure.

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 $d = \operatorname{dist}(x, Y) = \inf_{y \in Y} \|x - y\|$.)

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, $y^* \in Y_x$. □

Exercise 1.12. If, in Theorem 1.11, we also assume that Y is finite-dimensional, show that Y_x is *closed* (hence a *compact* convex set).

If Y_x contains more than one point, then, in fact, it contains an entire line segment. Thus, Y_x is either empty, contains exactly one point, or contains infinitely many points. This observation gives us a *sufficient condition* for uniqueness of nearest points: If the normed space X contains no line segments on any sphere $\{x \in X : \|x\| = r\}$, then best approximations (out of any convex subset Y) will necessarily be unique.

A norm $\|\cdot\|$ on a vector space X is said to be *strictly convex* if, for any pair of points $x \ne y \in X$ with $\|x\| = r = \|y\|$, we always have $\|\lambda x + (1-\lambda) y\| < r$ for all $0 < \lambda < 1$. That is, the open line segment between any pair of points on the sphere of radius r lies entirely within the open ball of radius r; in other words, only the endpoints of the line segment can hit the sphere. For simplicity, we often say that the space X is strictly convex, with the understanding that we're actually referring to a property of the norm in X. In any such space, we get an immediate corollary to our last result:

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*

$$x \ne \alpha y,\ \ y \ne \alpha x,\ \ \text{all } \alpha \in \mathbb{R} \implies \|x + y\| < \|x\| + \|y\|.$$

Proof. First suppose that X is strictly convex, and let x and y be nonparallel vectors in X. Then, in particular, the vectors $x/\|x\|$ and $y/\|y\|$ must be different. (Why?) Hence,

$$\left\| \left(\frac{\|x\|}{\|x\| + \|y\|} \right) \frac{x}{\|x\|} + \left(\frac{\|y\|}{\|x\| + \|y\|} \right) \frac{y}{\|y\|} \right\| < 1.$$

That is, $\|x + y\| < \|x\| + \|y\|$.

Next suppose that the triangle inequality is strict on nonparallel vectors, and let $x \neq y \in X$ with $\|x\| = r = \|y\|$. If x and y are parallel, then we must have $y = -x$. (Why?) In this case,

$$\|\lambda x + (1-\lambda)\, y\| = |2\lambda - 1|\, \|x\| < r$$

because $-1 < 2\lambda - 1 < 1$ whenever $0 < \lambda < 1$. Otherwise, x and y are nonparallel. Thus, for any $0 < \lambda < 1$, the vectors λx and $(1-\lambda)\, y$ are likewise nonparallel and we have

$$\|\lambda x + (1-\lambda)\, y\| < \lambda\|x\| + (1-\lambda)\|y\| = r. \qquad \square$$

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?)

Problems

[Problems marked (▷) are essential to a full understanding of the course. Problems marked (∗) are of general interest and are offered as a contribution to your personal growth. Unmarked problems are just for fun.]

The most important collection of functions for our purposes is the space $C[\,a,b\,]$, consisting of all continuous functions $f : [\,a,b\,] \to \mathbb{R}$. It's easy to see that $C[\,a,b\,]$ is a vector space under the usual pointwise operations on functions: $(f+g)(x) = f(x) + g(x)$ and $(\alpha f)(x) = \alpha f(x)$ for $\alpha \in \mathbb{R}$. Actually, we will be most interested in the finite-dimensional subspaces $\mathcal{P}_n$ of $C[\,a,b\,]$, consisting of all *algebraic polynomials* of degree at most n.

▷ 1. The subspace $\mathcal{P}_n$ has dimension exactly $n+1$. Why?

Another useful subset of $C[\,a,b\,]$ is the collection $\text{lip}_K\alpha$, consisting of those f which satisfy a *Lipschitz condition* of order $\alpha > 0$ with constant $0 < K < \infty$; i.e., those f for which $|f(x) - f(y)| \le K\,|x-y|^\alpha$ for all x, y in $[\,a,b\,]$. [Some authors would say that f is *Hölder continuous* with exponent α.]

∗ 2. (a) Show that $\text{lip}_K\alpha$ is, indeed, a subset of $C[\,a,b\,]$.
   (b) If $\alpha > 1$, show that $\text{lip}_K\alpha$ contains only the constant functions.
   (c) Show that $\sqrt{x}$ is in $\text{lip}_1(1/2)$ and that $\sin x$ is in $\text{lip}_1 1$ on $[\,0,1\,]$.
   (d) Show that the collection $\text{lip}\,\alpha$, consisting of all those f which are in $\text{lip}_K\alpha$ for some K, is a subspace of $C[\,a,b\,]$.
   (e) Show that $\text{lip}\,1$ contains all the polynomials.
   (f) If $f \in \text{lip}\,\alpha$ for some $\alpha > 0$, show that $f \in \text{lip}\,\beta$ for all $0 < \beta < \alpha$.
   (g) Given $0 < \alpha < 1$, show that x^α is in $\text{lip}_1\alpha$ on $[\,0,1\,]$ but *not* in $\text{lip}\,\beta$ for any $\beta > \alpha$.

The vector space $C[\,a,b\,]$ is most commonly endowed with the *uniform* or *sup norm*, defined by $\|f\| = \max_{a\le x\le b} |f(x)|$. Some authors use $\|f\|_u$ or $\|f\|_\infty$ here, and some authors refer to this as the *Chebyshev norm*. Whatever the notation used, it is the norm of choice on $C[\,a,b\,]$.

∗ 3. Show that $\mathcal{P}_n$ and $\text{lip}_K\alpha$ are closed subsets of $C[\,a,b\,]$ (under the sup norm). Is $\text{lip}\,\alpha$ closed? A bit harder: Show that $\text{lip}\,1$ is both first category and dense in $C[\,a,b\,]$.

4. Fix n and consider the norm $\|p\|_1 = \sum_{k=0}^n |a_k|$ for $p(x) = a_0 + a_1x + \cdots + a_nx^n \in \mathcal{P}_n$, considered as a subset of $C[\,a,b\,]$. Show that there are constants $0 < A_n$, $B_n < \infty$ such that $A_n\|p\|_1 \le \|p\| \le B_n\|p\|_1$, where $\|p\| = \max_{a\le x\le b}|p(x)|$. Do A_n and B_n really depend on n? Do they depend on the underlying interval $[\,a,b\,]$?

5. Fill-in any missing details from Example 1.8.

We will occasionally consider spaces of real-valued functions defined on finite sets; that is, we will consider $\mathbb{R}^n$ under various norms. (Why is this the same?) We define a scale of norms on $\mathbb{R}^n$ by setting $\|x\|_p = \left(\sum_{i=1}^n |x_i|^p\right)^{1/p}$, where $x = (x_1,\ldots,x_n)$ and $1 \le p < \infty$. (We need $p \ge 1$ in order for this expression to be a legitimate norm, but the expression makes perfect sense for any $p > 0$, and even for $p < 0$ provided no x_i is 0.) Notice, please, that the usual norm on $\mathbb{R}^n$ is given by $\|x\|_2$.

6. Show that $\lim_{p\to\infty} \|x\|_p = \max_{1\le i\le n} |x_i|$. For this reason we define

$$\|x\|_\infty = \max_{1\le i\le n} |x_i|.$$

Thus $\mathbb{R}^n$ under the norm $\|\cdot\|_\infty$ is the same as $C(\{1,2,\dots,n\})$ under its usual norm.

7. Assuming $x_i \neq 0$ for $i = 1,\dots,n$, compute $\lim_{p\to 0^+} \|x\|_p$ and $\lim_{p\to -\infty} \|x\|_p$.

8. Consider $\mathbb{R}^2$ under the norm $\|x\|_p$. Draw the graph of the unit sphere $\{x : \|x\|_p = 1\}$ for various values of p (especially $p = 1,\ 2,\ \infty$).

9. In a normed space normed space $(X, \|\cdot\|)$, prove that the following are equivalent:

   (a) $\|x+y\| = \|x\| + \|y\|$ always implies that x and y lie in the same direction; that is, either $x = \alpha y$ or $y = \alpha x$ for some nonnegative scalar α.

   (b) If $x,\ y \in X$ are nonparallel, then $\left\| \dfrac{x+y}{2} \right\| < \dfrac{\|x\| + \|y\|}{2}$.

   (c) If $x \neq y \in X$ with $\|x\| = 1 = \|y\|$, then $\left\| \dfrac{x+y}{2} \right\| < 1$.

   (d) X is strictly convex (as defined on page 7).

We write ℓ_p^n to denote the vector space of sequences of length n endowed with the p-norm; that is, $\mathbb{R}^n$ supplied with the norm $\|\cdot\|_p$. And we write ℓ_p to denote the vector space of infinite length sequences $x = (x_n)_{n=1}^\infty$ for which $\|x\|_p < \infty$. In each space, the usual algebraic operations are defined pointwise (or coordinatewise) and the norm is understood to be $\|\cdot\|_p$.

10. Show that ℓ_p (and hence ℓ_p^n) is strictly convex for $1 < p < \infty$. Show also that this fails in cases $p = 1$ and $p = \infty$. [Hint: Show that the function $f(t) = |t|^p$ satisfies $f((s+t)/2) < (f(s) + f(t))/2$ whenever $s \neq t$ and $1 < p < \infty$.]

∗ 11. Let X be a normed space and let $B = \{x \in X : \|x\| \le 1\}$. Show that B is a closed convex set.

12. Consider $\mathbb{R}^2$ under the norm $\|\cdot\|_\infty$. Let $B = \{y \in \mathbb{R}^2 : \|y\|_\infty \le 1\}$ and let $x = (2, 0)$. Show that there are infinitely many points in B nearest to x.

13. Let K be a *compact* convex set in a strictly convex space X and let $x \in X$. Show that x has a *unique* nearest point $y_0 \in K$.

14. Let K be a closed subset of a complete normed space X. Prove that K is convex if and only if K is *midpoint convex*; that is, if and only if $(x+y)/2 \in K$ whenever x, $y \in K$. Is this result true in more general settings? For example, can you prove it without assuming completeness? Or, for that matter, is it true for arbitrary sets in any vector space (i.e., without even assuming the presence of a norm)?