The methods presented in Chapter 6 used direct techniques to solve a system of $n \times n$ linear equations of the form $A\mathbf{x} = \mathbf{b}$. In this chapter, we present iterative methods to solve a system of this type.

7.1 Norms of Vectors and Matrices

In Chapter 2 we described iterative techniques for finding roots of equations of the form $f(x) = 0$. An initial approximation (or approximations) was found, and new approximations are then determined based on how well the previous approximations satisfied the equation. The objective is to find a way to minimize the difference between the approximations and the exact solution.

To discuss iterative methods for solving linear systems, we first need to determine a way to measure the distance between n-dimensional column vectors. This will permit us to determine whether a sequence of vectors converges to a solution of the system.

A scalar is a real (or complex) number generally denoted using italic or Greek letters. Vectors are denoted using boldface letters.

In actuality, this measure is also needed when the solution is obtained by the direct methods presented in Chapter 6. Those methods required a large number of arithmetic operations, and using finite-digit arithmetic leads only to an approximation to an actual solution of the system.

Vector Norms

Let $\mathbb{R}^n$ denote the set of all n-dimensional column vectors with real-number components. To define a distance in $\mathbb{R}^n$ we use the notion of a norm, which is the generalization of the absolute value on $\mathbb{R}$, the set of real numbers.

Definition 7.1 A **vector norm** on $\mathbb{R}^n$ is a function, $\|\cdot\|$, from $\mathbb{R}^n$ into $\mathbb{R}$ with the following properties:

(i) $\|\mathbf{x}\| \geq 0$ for all $\mathbf{x} \in \mathbb{R}^n$,

(ii) $\|\mathbf{x}\| = 0$ if and only if $\mathbf{x} = \mathbf{0}$,

(iii) $\|\alpha\mathbf{x}\| = |\alpha|\|\mathbf{x}\|$ for all $\alpha \in \mathbb{R}$ and $\mathbf{x} \in \mathbb{R}^n$,

(iv) $\|\mathbf{x} + \mathbf{y}\| \leq \|\mathbf{x}\| + \|\mathbf{y}\|$ for all $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$. ■

Vectors in $\mathbb{R}^n$ are column vectors, and it is convenient to use the transpose notation presented in Section 6.3 when a vector is represented in terms of its components. For example, the vector

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

will be written $\mathbf{x} = (x_1, x_2, \ldots, x_n)^t$.

We will need only two specific norms on $\mathbb{R}^n$, although a third norm on $\mathbb{R}^n$ is presented in Exercise 2.

Definition 7.2 The l_2 and l_∞ norms for the vector $\mathbf{x} = (x_1, x_2, \ldots, x_n)^t$ are defined by

$$\|\mathbf{x}\|_2 = \left\{\sum_{i=1}^{n} x_i^2\right\}^{1/2} \quad \text{and} \quad \|\mathbf{x}\|_\infty = \max_{1\leq i\leq n} |x_i|. \qquad ■$$

Note that each of these norms reduces to the absolute value in the case $n = 1$.

The l_2 norm is called the **Euclidean norm** of the vector $\mathbf{x}$ because it represents the usual notion of distance from the origin in case $\mathbf{x}$ is in $\mathbb{R}^1 \equiv \mathbb{R}$, $\mathbb{R}^2$, or $\mathbb{R}^3$. For example, the l_2 norm of the vector $\mathbf{x} = (x_1, x_2, x_3)^t$ gives the length of the straight line joining the points $(0, 0, 0)$ and (x_1, x_2, x_3). Figure 7.1 shows the boundary of those vectors in $\mathbb{R}^2$ and $\mathbb{R}^3$ that have l_2 norm less than 1. Figure 7.2 is a similar illustration for the l_∞ norm.

Figure 7.1

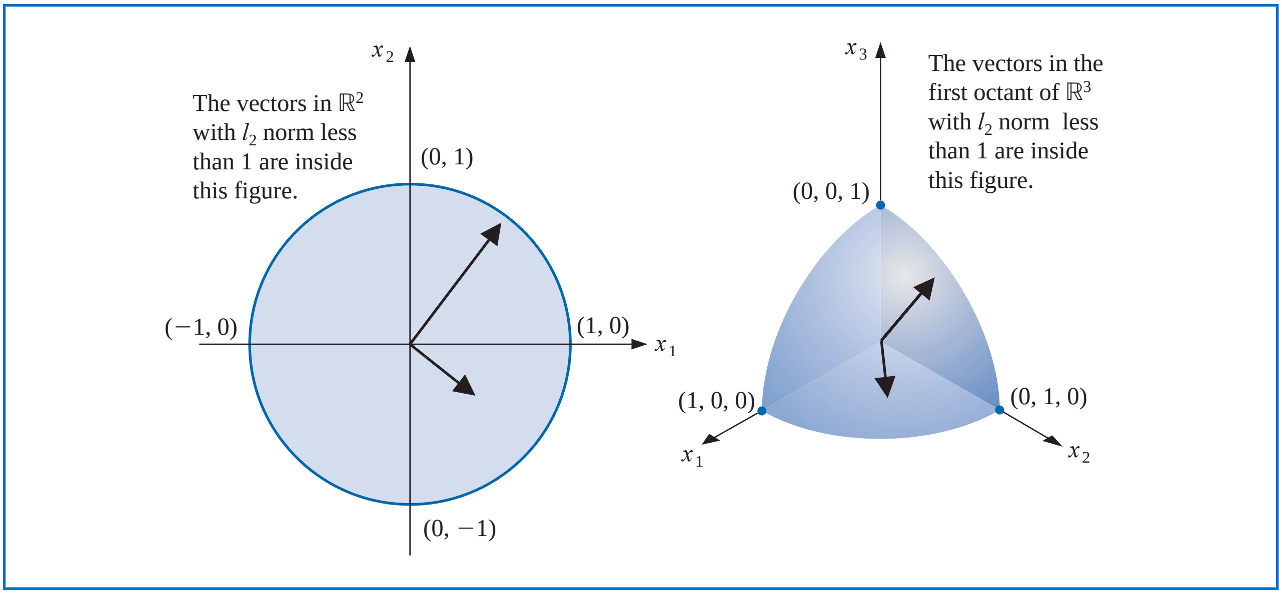
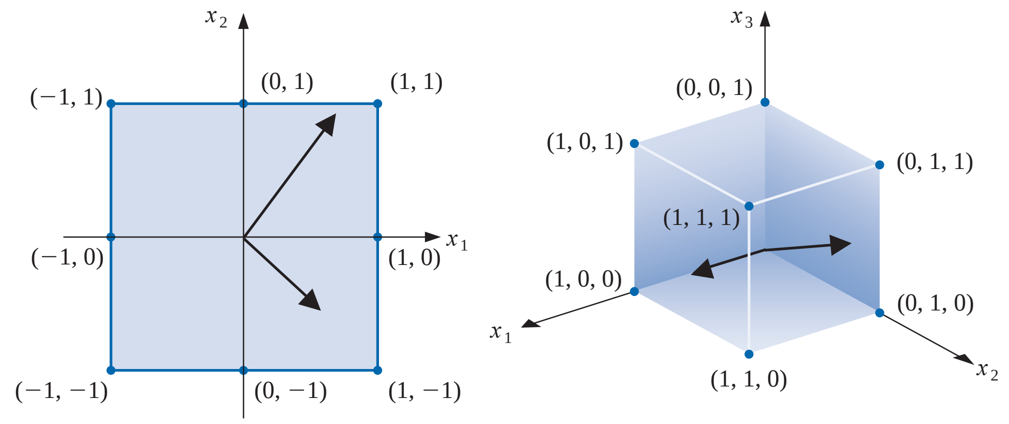


Figure 7.2



The vectors in $\mathbb{R}^2$ with l_∞ norm less than 1 are inside this figure.

The vectors in the first octant of $\mathbb{R}^3$ with l_∞ norm less than 1 are inside this figure.

Example 1 Determine the l_2 norm and the l_∞ norm of the vector $\mathbf{x} = (-1, 1, -2)^t$.

Solution The vector $\mathbf{x} = (-1, 1, -2)^t$ in $\mathbb{R}^3$ has norms

$$\|\mathbf{x}\|_2 = \sqrt{(-1)^2 + (1)^2 + (-2)^2} = \sqrt{6}$$

and

$$\|\mathbf{x}\|_\infty = \max\{|-1|, |1|, |-2|\} = 2.$$

■

It is easy to show that the properties in Definition 7.1 hold for the l_∞ norm because they follow from similar results for absolute values. The only property that requires much demonstration is (iv), and in this case if $\mathbf{x} = (x_1, x_2, \ldots, x_n)^t$ and $\mathbf{y} = (y_1, y_2, \ldots, y_n)^t$, then

$$\|\mathbf{x}+\mathbf{y}\|_\infty = \max_{1\le i\le n} |x_i + y_i| \le \max_{1\le i\le n} (|x_i| + |y_i|) \le \max_{1\le i\le n} |x_i| + \max_{1\le i\le n} |y_i| = \|\mathbf{x}\|_\infty + \|\mathbf{y}\|_\infty.$$

The first three conditions also are easy to show for the l_2 norm. But to show that

$$\|\mathbf{x}+\mathbf{y}\|_2 \le \|\mathbf{x}\|_2 + \|\mathbf{y}\|_2, \quad \text{for each } \mathbf{x}, \mathbf{y} \in \mathbb{R}_n,$$

we need a famous inequality.

Theorem 7.3 **(Cauchy-Bunyakovsky-Schwarz Inequality for Sums)**

For each $\mathbf{x} = (x_1, x_2, \ldots, x_n)^t$ and $\mathbf{y} = (y_1, y_2, \ldots, y_n)^t$ in $\mathbb{R}^n$,

$$\mathbf{x}^t\mathbf{y} = \sum_{i=1}^n x_i y_i \le \left\{\sum_{i=1}^n x_i^2\right\}^{1/2} \left\{\sum_{i=1}^n y_i^2\right\}^{1/2} = \|\mathbf{x}\|_2 \cdot \|\mathbf{y}\|_2. \tag{7.1}$$

■

There are many forms of this inequality, hence many discoverers. Augustin Louis Cauchy (1789–1857) describes this inequality in 1821 in *Cours d'Analyse Algébrique*, the first rigorous calculus book. An integral form of the equality appears in the work of Viktor Yakovlevich Bunyakovsky (1804–1889) in 1859, and Hermann Amandus Schwarz (1843–1921) used a double integral form of this inequality in 1885. More details on the history can be found in [Stee].

Proof If $\mathbf{y} = \mathbf{0}$ or $\mathbf{x} = \mathbf{0}$, the result is immediate because both sides of the inequality are zero.

Suppose $\mathbf{y} \ne \mathbf{0}$ and $\mathbf{x} \ne \mathbf{0}$. Note that for each $\lambda \in \mathbb{R}$ we have

$$0 \le ||\mathbf{x} - \lambda\mathbf{y}||_2^2 = \sum_{i=1}^n (x_i - \lambda y_i)^2 = \sum_{i=1}^n x_i^2 - 2\lambda \sum_{i=1}^n x_i y_i + \lambda^2 \sum_{i=1}^n y_i^2,$$

so that

$$2\lambda \sum_{i=1}^n x_i y_i \le \sum_{i=1}^n x_i^2 + \lambda^2 \sum_{i=1}^n y_i^2 = \|\mathbf{x}\|_2^2 + \lambda^2 \|\mathbf{y}\|_2^2.$$

However $\|\mathbf{x}\|_2 > 0$ and $\|\mathbf{y}\|_2 > 0$, so we can let $\lambda = \|\mathbf{x}\|_2 / \|\mathbf{y}\|_2$ to give

$$\left(2\frac{\|\mathbf{x}\|_2}{\|\mathbf{y}\|_2}\right)\left(\sum_{i=1}^n x_i y_i\right) \le \|\mathbf{x}\|_2^2 + \frac{\|\mathbf{x}\|_2^2}{\|\mathbf{y}\|_2^2}\|\mathbf{y}\|_2^2 = 2\|\mathbf{x}\|_2^2.$$

Hence

$$2\sum_{i=1}^n x_i y_i \le 2\|\mathbf{x}\|_2^2 \frac{\|\mathbf{y}\|_2}{\|\mathbf{x}\|_2} = 2\|\mathbf{x}\|_2\|\mathbf{y}\|_2,$$

and

$$\mathbf{x}^t\mathbf{y} = \sum_{i=1}^{n} x_i y_i \le \|\mathbf{x}\|_2\|\mathbf{y}\|_2 = \left\{\sum_{i=1}^{n} x_i^2\right\}^{1/2}\left\{\sum_{i=1}^{n} y_i^2\right\}^{1/2}.$$

■ ■ ■

With this result we see that for each $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$,

$$\|\mathbf{x}+\mathbf{y}\|_2^2 = \sum_{i=1}^{n}(x_i+y_i)^2 = \sum_{i=1}^{n} x_i^2 + 2\sum_{i=1}^{n} x_i y_i + \sum_{i=1}^{n} y_i^2 \le \|\mathbf{x}\|_2^2 + 2\|\mathbf{x}\|_2\|\mathbf{y}\|_2 + \|\mathbf{y}\|_2^2,$$

which gives norm property **(iv)**:

$$\|\mathbf{x}+\mathbf{y}\|_2 \le \left(\|\mathbf{x}\|_2^2 + 2\|\mathbf{x}\|_2\|\mathbf{y}\|_2 + \|\mathbf{y}\|_2^2\right)^{1/2} = \|\mathbf{x}\|_2 + \|\mathbf{y}\|_2.$$

Distance between Vectors in $\mathbb{R}^n$

The norm of a vector gives a measure for the distance between an arbitrary vector and the zero vector, just as the absolute value of a real number describes its distance from 0. Similarly, the **distance between two vectors** is defined as the norm of the difference of the vectors just as distance between two real numbers is the absolute value of their difference.

Definition 7.4 If $\mathbf{x} = (x_1, x_2, \ldots, x_n)^t$ and $\mathbf{y} = (y_1, y_2, \ldots, y_n)^t$ are vectors in $\mathbb{R}^n$, the l_2 and l_∞ distances between $\mathbf{x}$ and $\mathbf{y}$ are defined by

$$\|\mathbf{x}-\mathbf{y}\|_2 = \left\{\sum_{i=1}^{n}(x_i-y_i)^2\right\}^{1/2} \quad \text{and} \quad \|\mathbf{x}-\mathbf{y}\|_\infty = \max_{1\le i\le n}|x_i-y_i|.$$

■

Example 2 The linear system

$$3.3330x_1 + 15920x_2 - 10.333x_3 = 15913,$$
$$2.2220x_1 + 16.710x_2 + 9.6120x_3 = 28.544,$$
$$1.5611x_1 + 5.1791x_2 + 1.6852x_3 = 8.4254$$

has the exact solution $\mathbf{x} = (x_1, x_2, x_3)^t = (1, 1, 1)^t$, and Gaussian elimination performed using five-digit rounding arithmetic and partial pivoting (Algorithm 6.2), produces the approximate solution

$$\tilde{\mathbf{x}} = (\tilde{x}_1, \tilde{x}_2, \tilde{x}_3)^t = (1.2001, 0.99991, 0.92538)^t.$$

Determine the l_2 and l_∞ distances between the exact and approximate solutions.

Solution Measurements of $\mathbf{x} - \tilde{\mathbf{x}}$ are given by

$$\|\mathbf{x}-\tilde{\mathbf{x}}\|_\infty = \max\{|1-1.2001|, |1-0.99991|, |1-0.92538|\}$$
$$= \max\{0.2001, 0.00009, 0.07462\} = 0.2001$$

and

$$\|\mathbf{x}-\tilde{\mathbf{x}}\|_2 = \left[(1-1.2001)^2 + (1-0.99991)^2 + (1-0.92538)^2\right]^{1/2}$$
$$= [(0.2001)^2 + (0.00009)^2 + (0.07462)^2]^{1/2} = 0.21356.$$

Although the components $\tilde{x}_2$ and $\tilde{x}_3$ are good approximations to x_2 and x_3, the component $\tilde{x}_1$ is a poor approximation to x_1, and $|x_1 - \tilde{x}_1|$ dominates both norms. ■

The concept of distance in $\mathbb{R}^n$ is also used to define a limit of a sequence of vectors in this space.

Definition 7.5 A sequence $\{\mathbf{x}^{(k)}\}_{k=1}^{\infty}$ of vectors in $\mathbb{R}^n$ is said to **converge** to $\mathbf{x}$ with respect to the norm $\|\cdot\|$ if, given any $\varepsilon > 0$, there exists an integer $N(\varepsilon)$ such that

$$\|\mathbf{x}^{(k)} - \mathbf{x}\| < \varepsilon, \quad \text{for all } k \geq N(\varepsilon).$$ ■

Theorem 7.6 The sequence of vectors $\{\mathbf{x}^{(k)}\}$ converges to $\mathbf{x}$ in $\mathbb{R}^n$ with respect to the l_∞ norm if and only if $\lim_{k\to\infty} x_i^{(k)} = x_i$, for each $i = 1, 2, \ldots, n$. ■

Proof Suppose $\{\mathbf{x}^{(k)}\}$ converges to $\mathbf{x}$ with respect to the l_∞ norm. Given any $\varepsilon > 0$, there exists an integer $N(\varepsilon)$ such that for all $k \geq N(\varepsilon)$,

$$\max_{i=1,2,\ldots,n} |x_i^{(k)} - x_i| = \|\mathbf{x}^{(k)} - \mathbf{x}\|_\infty < \varepsilon.$$

This result implies that $|x_i^{(k)} - x_i| < \varepsilon$, for each $i = 1, 2, \ldots, n$, so $\lim_{k\to\infty} x_i^{(k)} = x_i$ for each i.

Conversely, suppose that $\lim_{k\to\infty} x_i^{(k)} = x_i$, for every $i = 1, 2, \ldots, n$. For a given $\varepsilon > 0$, let $N_i(\varepsilon)$ for each i represent an integer with the property that

$$|x_i^{(k)} - x_i| < \varepsilon,$$

whenever $k \geq N_i(\varepsilon)$.

Define $N(\varepsilon) = \max_{i=1,2,\ldots,n} N_i(\varepsilon)$. If $k \geq N(\varepsilon)$, then

$$\max_{i=1,2,\ldots,n} |x_i^{(k)} - x_i| = \|\mathbf{x}^{(k)} - \mathbf{x}\|_\infty < \varepsilon.$$

This implies that $\{\mathbf{x}^{(k)}\}$ converges to $\mathbf{x}$ with respect to the l_∞ norm. ▪ ▪ ▪

Example 3 Show that

$$\mathbf{x}^{(k)} = (x_1^{(k)}, x_2^{(k)}, x_3^{(k)}, x_4^{(k)})^t = \left(1, 2 + \frac{1}{k}, \frac{3}{k^2}, e^{-k} \sin k\right)^t.$$

converges to $\mathbf{x} = (1, 2, 0, 0)^t$ with respect to the l_∞ norm.

Solution Because

$$\lim_{k\to\infty} 1 = 1, \quad \lim_{k\to\infty} (2 + 1/k) = 2, \quad \lim_{k\to\infty} 3/k^2 = 0 \quad \text{and} \quad \lim_{k\to\infty} e^{-k} \sin k = 0,$$

Theorem 7.6 implies that the sequence $\{\mathbf{x}^{(k)}\}$ converges to $(1, 2, 0, 0)^t$ with respect to the l_∞ norm. ■

To show directly that the sequence in Example 3 converges to $(1, 2, 0, 0)^t$ with respect to the l_2 norm is quite complicated. It is better to prove the next result and apply it to this special case.

Theorem 7.7 For each $\mathbf{x} \in \mathbb{R}^n$,

$$\|\mathbf{x}\|_\infty \leq \|\mathbf{x}\|_2 \leq \sqrt{n}\|\mathbf{x}\|_\infty.$$ ■

Proof Let x_j be a coordinate of $\mathbf{x}$ such that $\|\mathbf{x}\|_\infty = \max_{1\le i\le n} |x_i| = |x_j|$. Then

$$\|\mathbf{x}\|_\infty^2 = |x_j|^2 = x_j^2 \le \sum_{i=1}^{n} x_i^2 = \|\mathbf{x}\|_2^2,$$

and

$$\|\mathbf{x}\|_\infty \le \|\mathbf{x}\|_2.$$

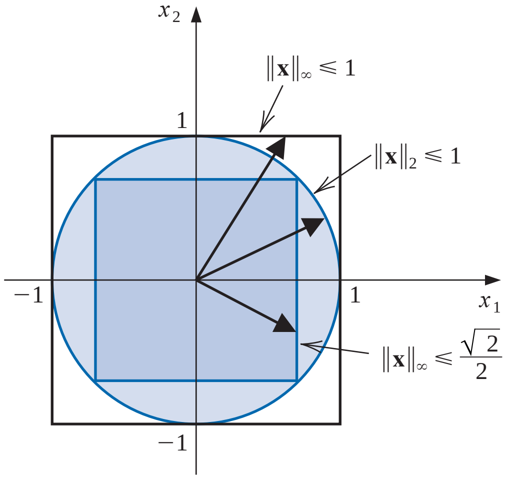
So

$$\|\mathbf{x}\|_2^2 = \sum_{i=1}^{n} x_i^2 \le \sum_{i=1}^{n} x_j^2 = nx_j^2 = n||\mathbf{x}||_\infty^2,$$

and $\|\mathbf{x}\|_2 \le \sqrt{n}\|\mathbf{x}\|_\infty$. ■ ■ ■

Figure 7.3 illustrates this result when $n = 2$.

Figure 7.3



Example 4 In Example 3, we found that the sequence $\{\mathbf{x}^{(k)}\}$, defined by

$$\mathbf{x}^{(k)} = \left(1, 2 + \frac{1}{k}, \frac{3}{k^2}, e^{-k}\sin k\right)^t,$$

converges to $\mathbf{x} = (1, 2, 0, 0)^t$ with respect to the l_∞ norm. Show that this sequence also converges to $\mathbf{x}$ with respect to the l_2 norm.

Solution Given any $\varepsilon > 0$, there exists an integer $N(\varepsilon/2)$ with the property that

$$\|\mathbf{x}^{(k)} - \mathbf{x}\|_\infty < \frac{\varepsilon}{2},$$

whenever $k \ge N(\varepsilon/2)$. By Theorem 7.7, this implies that

$$\|\mathbf{x}^{(k)} - \mathbf{x}\|_2 \le \sqrt{4}\|\mathbf{x}^{(k)} - \mathbf{x}\|_\infty \le 2(\varepsilon/2) = \varepsilon,$$

when $k \ge N(\varepsilon/2)$. So $\{\mathbf{x}^{(k)}\}$ also converges to $\mathbf{x}$ with respect to the l_2 norm. ■

It can be shown that all norms on $\mathbb{R}^n$ are equivalent with respect to convergence; that is, if $\|\cdot\|$ and $\|\cdot\|'$ are any two norms on $\mathbb{R}^n$ and $\{\mathbf{x}^{(k)}\}_{k=1}^{\infty}$ has the limit $\mathbf{x}$ with respect to $\|\cdot\|$, then $\{\mathbf{x}^{(k)}\}_{k=1}^{\infty}$ also has the limit $\mathbf{x}$ with respect to $\|\cdot\|'$. The proof of this fact for the general case can be found in [Or2], p. 8. The case for the l_2 and l_∞ norms follows from Theorem 7.7.

Matrix Norms and Distances

In the subsequent sections of this and later chapters, we will need methods for determining the distance between $n \times n$ matrices. This again requires the use of a norm.

Definition 7.8 A **matrix norm** on the set of all $n \times n$ matrices is a real-valued function, $\|\cdot\|$, defined on this set, satisfying for all $n \times n$ matrices A and B and all real numbers α:

(i) $\|A\| \geq 0$;

(ii) $\|A\| = 0$, if and only if A is O, the matrix with all 0 entries;

(iii) $\|\alpha A\| = |\alpha|\|A\|$;

(iv) $\|A + B\| \leq \|A\| + \|B\|$;

(v) $\|AB\| \leq \|A\|\|B\|$. ■

The **distance between $n \times n$ matrices** A and B with respect to this matrix norm is $\|A - B\|$.

Although matrix norms can be obtained in various ways, the norms considered most frequently are those that are natural consequences of the vector norms l_2 and l_∞.

These norms are defined using the following theorem, whose proof is considered in Exercise 13.

Theorem 7.9 If $||\cdot||$ is a vector norm on $\mathbb{R}^n$, then

$$\|A\| = \max_{\|\mathbf{x}\|=1} \|A\mathbf{x}\| \tag{7.2}$$

is a matrix norm. ■

Every vector norm produces an associated natural matrix norm.

Matrix norms defined by vector norms are called the **natural**, or *induced*, **matrix norm** associated with the vector norm. In this text, all matrix norms will be assumed to be natural matrix norms unless specified otherwise.

For any $\mathbf{z} \neq \mathbf{0}$, the vector $\mathbf{x} = \mathbf{z}/\|\mathbf{z}\|$ is a unit vector. Hence

$$\max_{\|\mathbf{x}\|=1} \|A\mathbf{x}\| = \max_{\mathbf{z}\neq\mathbf{0}} \left\| A\left(\frac{\mathbf{z}}{\|\mathbf{z}\|}\right) \right\| = \max_{\mathbf{z}\neq\mathbf{0}} \frac{\|A\mathbf{z}\|}{\|\mathbf{z}\|},$$

and we can alternatively write

$$\|A\| = \max_{\mathbf{z}\neq\mathbf{0}} \frac{\|A\mathbf{z}\|}{\|\mathbf{z}\|}. \tag{7.3}$$

The following corollary to Theorem 7.9 follows from this representation of $\|A\|$.

Corollary 7.10 For any vector $\mathbf{z} \neq \mathbf{0}$, matrix A, and any natural norm $\|\cdot\|$, we have

$$\|A\mathbf{z}\| \leq \|A\| \cdot \|\mathbf{z}\|. \quad ■$$

The measure given to a matrix under a natural norm describes how the matrix stretches unit vectors relative to that norm. The maximum stretch is the norm of the matrix. The matrix norms we will consider have the forms

$$\|A\|_\infty = \max_{\|\mathbf{x}\|_\infty=1} \|A\mathbf{x}\|_\infty, \quad \text{the } l_\infty \text{ norm,}$$

and

$$\|A\|_2 = \max_{\|\mathbf{x}\|_2=1} \|A\mathbf{x}\|_2, \quad \text{the } l_2 \text{ norm.}$$

An illustration of these norms when $n = 2$ is shown in Figures 7.4 and 7.5 for the matrix

$$A = \begin{bmatrix} 0 & -2 \\ 2 & 0 \end{bmatrix}$$

Figure 7.4

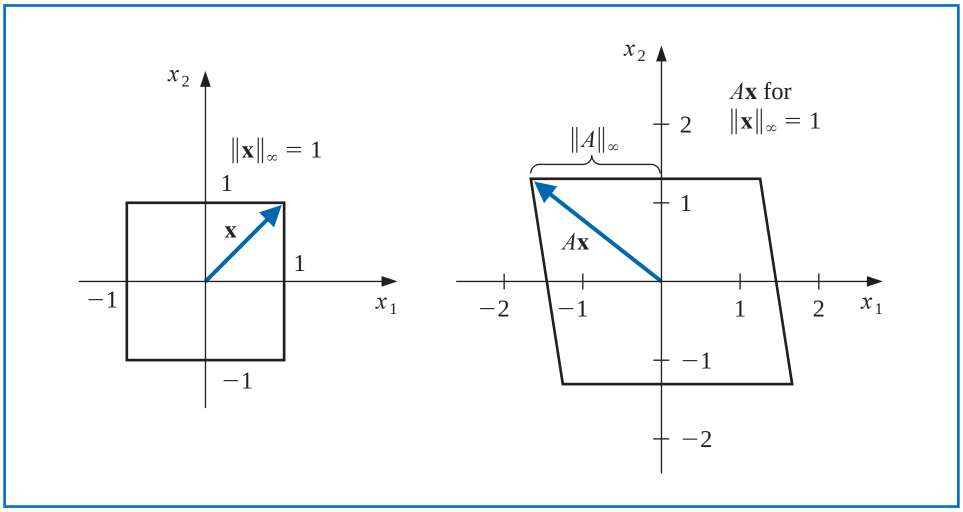
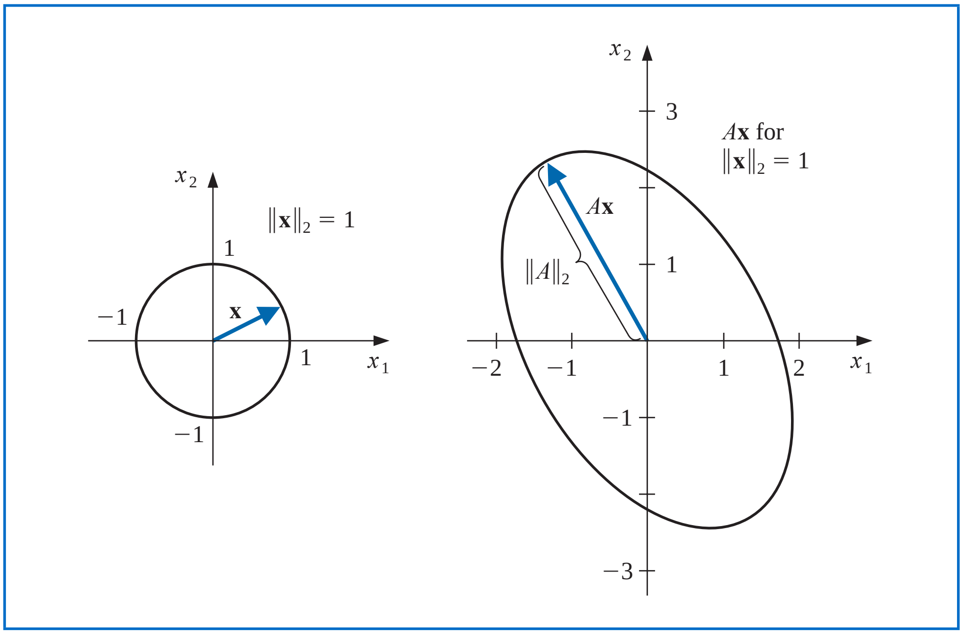


Figure 7.5



The l_∞ norm of a matrix can be easily computed from the entries of the matrix.

Theorem 7.11 If $A = (a_{ij})$ is an $n \times n$ matrix, then

$$\|A\|_\infty = \max_{1\le i\le n} \sum_{j=1}^{n} |a_{ij}|. \qquad \blacksquare$$

Proof First we show that $\|A\|_\infty \le \max_{1\le i\le n} \sum_{j=1}^{n} |a_{ij}|$.

Let $\mathbf{x}$ be an n-dimensional vector with $1 = \|\mathbf{x}\|_\infty = \max_{1\le i\le n} |x_i|$. Since $A\mathbf{x}$ is also an n-dimensional vector,

$$\|A\mathbf{x}\|_\infty = \max_{1\le i\le n} |(A\mathbf{x})_i| = \max_{1\le i\le n} \left| \sum_{j=1}^{n} a_{ij}x_j \right| \le \max_{1\le i\le n} \sum_{j=1}^{n} |a_{ij}| \max_{1\le j\le n} |x_j|.$$

But $\max_{1\le j\le n} |x_j| = \|\mathbf{x}\|_\infty = 1$, so

$$\|A\mathbf{x}\|_\infty \le \max_{1\le i\le n} \sum_{j=1}^{n} |a_{ij}|,$$

and consequently,

$$\|A\|_\infty = \max_{\|\mathbf{x}\|_\infty=1} \|A\mathbf{x}\|_\infty \le \max_{1\le i\le n} \sum_{j=1}^{n} |a_{ij}|. \tag{7.4}$$

Now we will show the opposite inequality. Let p be an integer with

$$\sum_{j=1}^{n} |a_{pj}| = \max_{1\le i\le n} \sum_{j=1}^{n} |a_{ij}|,$$

and $\mathbf{x}$ be the vector with components

$$x_j = \begin{cases} 1, & \text{if } a_{pj} \ge 0, \\ -1, & \text{if } a_{pj} < 0. \end{cases}$$

Then $\|\mathbf{x}\|_\infty = 1$ and $a_{pj}x_j = |a_{pj}|$, for all $j = 1, 2, \ldots, n$, so

$$\|A\mathbf{x}\|_\infty = \max_{1\le i\le n} \left| \sum_{j=1}^{n} a_{ij}x_j \right| \ge \left| \sum_{j=1}^{n} a_{pj}x_j \right| = \left| \sum_{j=1}^{n} |a_{pj}| \right| = \max_{1\le i\le n} \sum_{j=1}^{n} |a_{ij}|.$$

This result implies that

$$\|A\|_\infty = \max_{\|\mathbf{x}\|_\infty=1} \|A\mathbf{x}\|_\infty \ge \max_{1\le i\le n} \sum_{j=1}^{n} |a_{ij}|.$$

Putting this together with Inequality (7.4) gives $\|A\|_\infty = \max_{1\le i\le n} \sum_{j=1}^{n} |a_{ij}|$. ▪ ▪ ▪

Example 5 Determine $\|A\|_\infty$ for the matrix

$$A = \begin{bmatrix} 1 & 2 & -1 \\ 0 & 3 & -1 \\ 5 & -1 & 1 \end{bmatrix}.$$

Solution We have

$$\sum_{j=1}^{3} |a_{1j}| = |1| + |2| + |-1| = 4, \quad \sum_{j=1}^{3} |a_{2j}| = |0| + |3| + |-1| = 4,$$

and

$$\sum_{j=1}^{3} |a_{3j}| = |5| + |-1| + |1| = 7.$$

So Theorem 7.11 implies that $\|A\|_\infty = \max\{4, 4, 7\} = 7$. ■

In the next section, we will discover an alternative method for finding the l_2 norm of a matrix.

EXERCISE SET 7.1

1. Find l_∞ and l_2 norms of the vectors.

a. $\mathbf{x} = (3, -4, 0, \frac{3}{2})^t$

b. $\mathbf{x} = (2, 1, -3, 4)^t$

c. $\mathbf{x} = (\sin k, \cos k, 2^k)^t$ for a fixed positive integer k

d. $\mathbf{x} = (4/(k+1), 2/k^2, k^2 e^{-k})^t$ for a fixed positive integer k

2. **a.** Verify that the function $\|\cdot\|_1$, defined on $\mathbb{R}^n$ by

$$\|\mathbf{x}\|_1 = \sum_{i=1}^{n} |x_i|,$$

is a norm on $\mathbb{R}^n$.

b. Find $\|\mathbf{x}\|_1$ for the vectors given in Exercise 1.

c. Prove that for all $\mathbf{x} \in \mathbb{R}^n$, $\|\mathbf{x}\|_1 \geq \|\mathbf{x}\|_2$.

3. Prove that the following sequences are convergent, and find their limits.

a. $\mathbf{x}^{(k)} = (1/k, e^{1-k}, -2/k^2)^t$

b. $\mathbf{x}^{(k)} = \left(e^{-k}\cos k, k\sin(1/k), 3 + k^{-2}\right)^t$

c. $\mathbf{x}^{(k)} = (ke^{-k^2}, (\cos k)/k, \sqrt{k^2+k} - k)^t$

d. $\mathbf{x}^{(k)} = (e^{1/k}, (k^2+1)/(1-k^2), (1/k^2)(1+3+5+\cdots+(2k-1)))^t$

4. Find the l_∞ norm of the matrices.

a. $\begin{bmatrix} 10 & 15 \\ 0 & 1 \end{bmatrix}$

b. $\begin{bmatrix} 10 & 0 \\ 15 & 1 \end{bmatrix}$

c. $\begin{bmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{bmatrix}$

d. $\begin{bmatrix} 4 & -1 & 7 \\ -1 & 4 & 0 \\ -7 & 0 & 4 \end{bmatrix}$

5. The following linear systems $A\mathbf{x} = \mathbf{b}$ have $\mathbf{x}$ as the actual solution and $\tilde{\mathbf{x}}$ as an approximate solution. Compute $\|\mathbf{x} - \tilde{\mathbf{x}}\|_\infty$ and $\|A\tilde{\mathbf{x}} - \mathbf{b}\|_\infty$.

a. $\frac{1}{2}x_1 + \frac{1}{3}x_2 = \frac{1}{63}$,
$\frac{1}{3}x_1 + \frac{1}{4}x_2 = \frac{1}{168}$,
$\mathbf{x} = \left(\frac{1}{7}, -\frac{1}{6}\right)^t$,
$\tilde{\mathbf{x}} = (0.142, -0.166)^t$.

b. $x_1 + 2x_2 + 3x_3 = 1$,
$2x_1 + 3x_2 + 4x_3 = -1$,
$3x_1 + 4x_2 + 6x_3 = 2$,
$\mathbf{x} = (0, -7, 5)^t$,
$\tilde{\mathbf{x}} = (-0.33, -7.9, 5.8)^t$.

c. $x_1 + 2x_2 + 3x_3 = 1$,
$2x_1 + 3x_2 + 4x_3 = -1$,
$3x_1 + 4x_2 + 6x_3 = 2$,
$\mathbf{x} = (0, -7, 5)^t$,
$\tilde{\mathbf{x}} = (-0.2, -7.5, 5.4)^t$.

d. $0.04x_1 + 0.01x_2 - 0.01x_3 = 0.06$,
$0.2x_1 + 0.5x_2 - 0.2x_3 = 0.3$,
$x_1 + 2x_2 + 4x_3 = 11$,
$\mathbf{x} = (1.827586, 0.6551724, 1.965517)^t$,
$\tilde{\mathbf{x}} = (1.8, 0.64, 1.9)^t$.

6. The matrix norm $\|\cdot\|_1$, defined by $\|A\|_1 = \max_{\|\mathbf{x}\|_1=1} \|A\mathbf{x}\|_1$, can be computed using the formula

$$\|A\|_1 = \max_{1\le j\le n} \sum_{i=1}^{n} |a_{ij}|,$$

where the vector norm $\|\cdot\|_1$ is defined in Exercise 2. Find $\|\cdot\|_1$ for the matrices in Exercise 4.

7. Show by example that $\|\cdot\|_{\otimes}$, defined by $\|A\|_{\otimes} = \max_{1\le i,j\le n} |a_{ij}|$, does not define a matrix norm.

8. Show that $\|\cdot\|_{①}$, defined by

$$\|A\|_{①} = \sum_{i=1}^{n}\sum_{j=1}^{n} |a_{ij}|,$$

is a matrix norm. Find $\|\cdot\|_{①}$ for the matrices in Exercise 4.

9. **a.** The Frobenius norm (which is not a natural norm) is defined for an $n \times n$ matrix A by

$$\|A\|_F = \left(\sum_{i=1}^{n}\sum_{j=1}^{n} |a_{ij}|^2\right)^{1/2}.$$

Show that $\|\cdot\|_F$ is a matrix norm.

b. Find $\|\cdot\|_F$ for the matrices in Exercise 4.

c. For any matrix A, show that $\|A\|_2 \le \|A\|_F \le n^{1/2}\|A\|_2$.

10. In Exercise 9 the Frobenius norm of a matrix was defined. Show that for any $n \times n$ matrix A and vector $\mathbf{x}$ in $\mathbb{R}^n$, $\|A\mathbf{x}\|_2 \le \|A\|_F\|\mathbf{x}\|_2$.

11. Let S be a positive definite $n \times n$ matrix. For any $\mathbf{x}$ in $\mathbb{R}^n$ define $\|\mathbf{x}\| = (\mathbf{x}^t S\mathbf{x})^{1/2}$. Show that this defines a norm on $\mathbb{R}^n$. [*Hint:* Use the Cholesky factorization of S to show that $\mathbf{x}^t S\mathbf{y} = \mathbf{y}^t S\mathbf{x} \le (\mathbf{x}^t S\mathbf{x})^{1/2}(\mathbf{y}^t S\mathbf{y})^{1/2}$.]

12. Let S be a real and nonsingular matrix, and let $\|\cdot\|$ be any norm on $\mathbb{R}^n$. Define $\|\cdot\|'$ by $\|\mathbf{x}\|' = \|S\mathbf{x}\|$. Show that $\|\cdot\|'$ is also a norm on $\mathbb{R}^n$.

13. Prove that if $\|\cdot\|$ is a vector norm on $\mathbb{R}^n$, then $\|A\| = \max_{\|\mathbf{x}\|=1} \|A\mathbf{x}\|$ is a matrix norm.

14. The following excerpt from the *Mathematics Magazine* [Sz] gives an alternative way to prove the Cauchy-Buniakowsky-Schwarz Inequality.

a. Show that when $\mathbf{x} \ne \mathbf{0}$ and $\mathbf{y} \ne \mathbf{0}$, we have

$$\frac{\sum_{i=1}^{n} x_i y_i}{\left(\sum_{i=1}^{n} x_i^2\right)^{1/2}\left(\sum_{i=1}^{n} y_i^2\right)^{1/2}} = 1 - \frac{1}{2}\sum_{i=1}^{n}\left(\frac{x_i}{\left(\sum_{j=1}^{n} x_j^2\right)^{1/2}} - \frac{y_i}{\left(\sum_{j=1}^{n} y_j^2\right)^{1/2}}\right)^2.$$