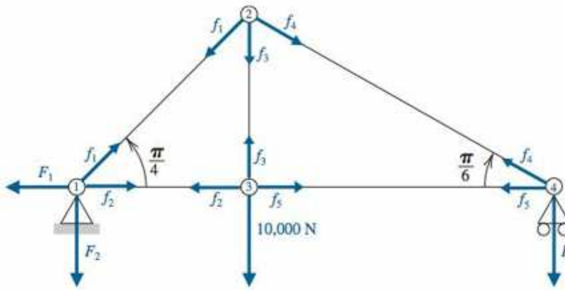
b. Solve this system using $n = 10, 50$, and 100.

c. Change the probabilities to α and $1 - \alpha$ for movement to the left and right, respectively, and derive the linear system similar to the one in (a).

d. Repeat (b) using the system in (c) with $\alpha = \frac{1}{3}$.

5. The forces on the bridge truss shown here satisfy the equations in the following table:



Joint	Horizontal Component	Vertical Component
①	$-F_1 + \frac{\sqrt{2}}{2} f_1 + f_2 = 0$	$\frac{\sqrt{2}}{2} f_1 - F_2 = 0$
②	$-\frac{\sqrt{2}}{2} f_1 + \frac{\sqrt{3}}{2} f_4 = 0$	$-\frac{\sqrt{2}}{2} f_1 - f_3 - \frac{1}{2} f_4 = 0$
③	$-f_2 + f_5 = 0$	$f_3 - 10{,}000 = 0$
④	$-\frac{\sqrt{3}}{2} f_4 - f_5 = 0$	$\frac{1}{2} f_4 - F_3 = 0$

This linear system can be placed in the matrix form

$$\begin{bmatrix} -1 & 0 & 0 & \frac{\sqrt{2}}{2} & 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & \frac{\sqrt{2}}{2} & 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 & 0 & 0 & \frac{1}{2} & 0 \\ 0 & 0 & 0 & -\frac{\sqrt{2}}{2} & 0 & -1 & \frac{1}{2} & 0 \\ 0 & 0 & 0 & 0 & -1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & -\frac{\sqrt{2}}{2} & 0 & 0 & \frac{\sqrt{3}}{2} & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & -\frac{\sqrt{3}}{2} & -1 \end{bmatrix} \begin{bmatrix} F_1 \\ F_2 \\ F_3 \\ f_1 \\ f_2 \\ f_3 \\ f_4 \\ f_5 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 10{,}000 \\ 0 \\ 0 \end{bmatrix}.$$

Approximate the solution of the resulting linear system to within 10^{-2} in the l_∞ norm using as initial approximation the vector all of whose entries are 1s and (i) the Gauss-Seidel method, (ii) the Jacobi method, and (iii) the SOR method with $\omega = 1.25$.

7.6 Error Bounds and Iterative Refinement

This section considers the errors in approximation that are likely to occur when solving linear systems by both direct and iterative methods. There is no universally superior technique for approximating the solution to linear systems, but some methods will give better results than others when the matrix satisfies certain conditions.

It seems intuitively reasonable that if $\tilde{\mathbf{x}}$ is an approximation to the solution $\mathbf{x}$ of $A\mathbf{x} = \mathbf{b}$ and the **residual vector**, defined by

$$\mathbf{r} = \mathbf{b} - A\tilde{\mathbf{x}},$$

has the property that if $\|\mathbf{r}\|$ is small, then $\|\mathbf{x} - \tilde{\mathbf{x}}\|$ should be small as well. This is often the case, but certain systems, which occur quite often in practice, fail to have this property.

Example 1 The linear system $A\mathbf{x} = \mathbf{b}$ given by

$$\begin{bmatrix} 1 & 2 \\ 1.0001 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 3 \\ 3.0001 \end{bmatrix}$$

has the unique solution $\mathbf{x} = (1, 1)^t$. Determine the residual vector for the poor approximation $\tilde{\mathbf{x}} = (3, -0.0001)^t$.

Solution We have

$$\mathbf{r} = \mathbf{b} - A\tilde{\mathbf{x}} = \begin{bmatrix} 3 \\ 3.0001 \end{bmatrix} - \begin{bmatrix} 1 & 2 \\ 1.0001 & 2 \end{bmatrix} \begin{bmatrix} 3 \\ -0.0001 \end{bmatrix} = \begin{bmatrix} 0.0002 \\ 0 \end{bmatrix},$$

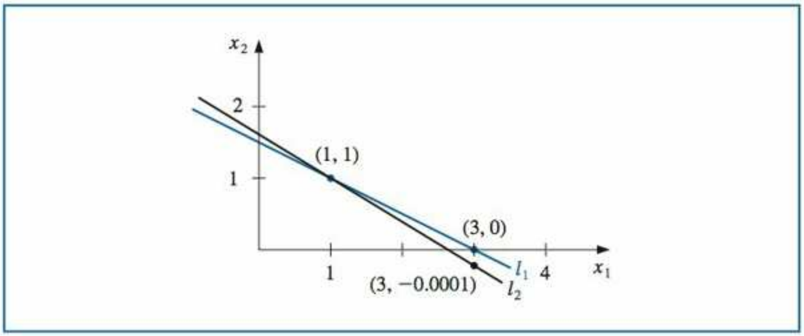
so $\|\mathbf{r}\|_\infty = 0.0002$. Although the norm of the residual vector is small, the approximation $\tilde{\mathbf{x}} = (3, -0.0001)^t$ is obviously quite poor; in fact, $\|\mathbf{x} - \tilde{\mathbf{x}}\|_\infty = 2$. ■

The difficulty in Example 1 is explained quite simply by noting that the solution to the system represents the intersection of the lines

$$l_1: \quad x_1 + 2x_2 = 3 \quad \text{and} \quad l_2: \quad 1.0001x_1 + 2x_2 = 3.0001.$$

The point $(3, -0.0001)$ lies on l_2, and the lines are nearly the same. This means that $(3, -0.0001)$ also lies close to the line l_1, even though it differs significantly from the solution of the system, which is the intersection point $(1, 1)$. (See Figure 7.6.)

Figure 7.6



Example 1 was clearly constructed to show the difficulties that might—and, in fact, do—arise. Had the lines not been nearly coincident, we would expect a small residual vector to imply an accurate approximation. In the general situation, we cannot rely on the geometry of the system to give an indication of when problems might occur. We can, however, obtain this information by considering the norms of the matrix and its inverse.

Residual Vector Error Bounds

If $\tilde{\mathbf{x}}$ is an approximation to the solution of $A\mathbf{x} = \mathbf{b}$ and A is a nonsingular matrix, then for any natural norm,

$$\|\mathbf{x} - \tilde{\mathbf{x}}\| \leq \|\mathbf{b} - A\tilde{\mathbf{x}}\| \cdot \|A^{-1}\|$$

and

$$\frac{\|\mathbf{x} - \tilde{\mathbf{x}}\|}{\|\mathbf{x}\|} \leq \|A\| \cdot \|A^{-1}\| \frac{\|\mathbf{b} - A\tilde{\mathbf{x}}\|}{\|\mathbf{b}\|}, \quad \text{provided } \mathbf{x} \neq \mathbf{0} \text{ and } \mathbf{b} \neq \mathbf{0}.$$

This result implies that $\|A^{-1}\|$ and $\|A\| \cdot \|A^{-1}\|$ provide an indication of the connection between the residual vector and the accuracy of the approximation. In general, the relative error $\|\mathbf{x} - \tilde{\mathbf{x}}\|/\|\mathbf{x}\|$ is of most interest. Any convenient norm can be used for this approximation; the only requirement is that it be used consistently throughout.

Condition Numbers

The **condition number**, $K(A)$, of the nonsingular matrix A relative to a norm $\|\cdot\|$ is

$$K(A) = \|A\| \cdot \|A^{-1}\|.$$

Note that for any nonsingular matrix A and natural norm $\|\cdot\|$,

$$1 = \|I\| = \|A \cdot A^{-1}\| \leq \|A\| \cdot \|A^{-1}\| = K(A).$$

With this notation, we can reexpress the inequalities in the previous result as

$$\|\mathbf{x} - \tilde{\mathbf{x}}\| \leq K(A) \frac{\|\mathbf{b} - A\tilde{\mathbf{x}}\|}{\|A\|} \quad \text{and} \quad \frac{\|\mathbf{x} - \tilde{\mathbf{x}}\|}{\|\mathbf{x}\|} \leq K(A) \frac{\|\mathbf{b} - A\tilde{\mathbf{x}}\|}{\|\mathbf{b}\|}.$$

A matrix A is well-behaved (called **well-conditioned**) if $K(A)$ is close to 1, and A is not well-behaved (called **ill-conditioned**) when $K(A)$ is significantly greater than 1. Conditioning in this instance refers to the relative security that a small residual vector implies a correspondingly accurate approximate solution.

Example 2 Determine the condition number for the matrix

$$A = \begin{bmatrix} 1 & 2 \\ 1.0001 & 2 \end{bmatrix}.$$

Solution We saw in Example 1 that the very poor approximation $(3, -0.0001)^t$ to the exact solution $(1, 1)^t$ had a residual vector with small norm, so we should expect the condition number of A to be large. We have $\|A\|_\infty = \max\{|1| + |2|, |1.001| + |2|\} = 3.0001$, which would not be considered large. However,

$$A^{-1} = \begin{bmatrix} -10000 & 10000 \\ 5000.5 & -5000 \end{bmatrix}, \quad \text{so} \quad \|A^{-1}\|_\infty = 20000,$$

and for the infinity norm,

$$K_\infty(A) = (20000)(3.0001) = 60002.$$

The size of the condition number for this example should certainly keep us from making hasty accuracy decisions based on the residual of an approximation. ∎

In MATLAB, the condition number $K_\infty(A)$ for the matrix in Example 2 can be computed using the command cond. To obtain the l_∞ condition number, use the command

```
cond(A,Inf)
```

MATLAB responds with

$$ans = 6.000199999999003e + 004$$

The default for cond is the l_2 condition number, and either of the commands `cond(A)` or `cond(A,2)` gives

$$K_2(A) = 5.000100002987370e + 004.$$

Iterative Refinement

The residual of an approximation can also be used to improve the accuracy of the approximation. Suppose that $\tilde{\mathbf{x}}$ is an approximation to the solution of the linear system $A\mathbf{x} = \mathbf{b}$ and that $\mathbf{r} = \mathbf{b} - \tilde{A}\mathbf{x}$ is the residual vector associated with $\tilde{\mathbf{x}}$. Consider $\tilde{\mathbf{y}}$, the approximate solution to the system $A\mathbf{y} = \mathbf{r}$. Then

$$\tilde{\mathbf{y}} \approx A^{-1}\mathbf{r} = A^{-1}(\mathbf{b} - A\tilde{\mathbf{x}}) = A^{-1}\mathbf{b} - A^{-1}A\tilde{\mathbf{x}} = \mathbf{x} - \tilde{\mathbf{x}}.$$

So

$$\mathbf{x} \approx \tilde{\mathbf{x}} + \tilde{\mathbf{y}}.$$

The program ITREF74 implements the Iterative Refinement method.

This new approximation $\tilde{\mathbf{x}} + \tilde{\mathbf{y}}$ is often much closer to the solution of $A\mathbf{x} = \mathbf{b}$ than is $\tilde{\mathbf{x}}$, and $\tilde{\mathbf{y}}$ is easy to determine because it involves the same matrix, A, as the original system. This technique is called **iterative refinement**, or *iterative improvement*, and is shown in the following Illustration. To increase accuracy, the residual vector is computed using double-digit arithmetic.

Illustration The linear system given by

$$\begin{bmatrix} 3.3330 & 15920 & -10.333 \\ 2.2220 & 16.710 & 9.6120 \\ 1.5611 & 5.1791 & 1.6852 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 15913 \\ 28.544 \\ 8.4254 \end{bmatrix}$$

has the exact solution $\mathbf{x} = (1, 1, 1)^t$.

Using Gaussian elimination and five-digit rounding arithmetic leads successively to the augmented matrices

$$\left[\begin{array}{ccc:c} 3.3330 & 15920 & -10.333 & 15913 \\ 0 & -10596 & 16.501 & 10580 \\ 0 & -7451.4 & 6.5250 & -7444.9 \end{array}\right]$$

and

$$\left[\begin{array}{ccc:c} 3.3330 & 15920 & -10.333 & 15913 \\ 0 & -10596 & 16.501 & -10580 \\ 0 & 0 & -5.0790 & -4.7000 \end{array}\right].$$

The approximate solution to this system is

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

The residual vector corresponding to $\tilde{\mathbf{x}}$ is computed in double precision to be

$$\begin{aligned}
\mathbf{r} &= \mathbf{b} - A\tilde{\mathbf{x}} \\
&= \begin{bmatrix} 15913 \\ 28.544 \\ 8.4254 \end{bmatrix} - \begin{bmatrix} 3.3330 & 15920 & -10.333 \\ 2.2220 & 16.710 & 9.6120 \\ 1.5611 & 5.1791 & 1.6852 \end{bmatrix} \begin{bmatrix} 1.2001 \\ 0.99991 \\ 0.92538 \end{bmatrix} \\
&= \begin{bmatrix} 15913 \\ 28.544 \\ 8.4254 \end{bmatrix} - \begin{bmatrix} 15913.00518 \\ 28.26987086 \\ 8.611560367 \end{bmatrix} = \begin{bmatrix} -0.00518 \\ 0.27412914 \\ -0.186160367 \end{bmatrix},
\end{aligned}$$

so

$$\|\mathbf{r}\|_\infty = \|\mathbf{b} - A\tilde{\mathbf{x}}\|_\infty = 0.27413.$$

To use iterative refinement to improve this approximation, we now solve the system $A\mathbf{y} = \mathbf{r}$ for $\tilde{\mathbf{y}}$. Using five-digit arithmetic and Gaussian elimination, the approximate solution $\tilde{\mathbf{y}}$ to the equation $A\mathbf{y} = \mathbf{r}$ is

$$\tilde{\mathbf{y}} = (-0.20008, 8.9987 \times 10^{-5}, 0.074607)^t$$

and we have the improved approximation to the system $A\mathbf{x} = \mathbf{b}$ given by

$$\begin{aligned}
\tilde{\mathbf{x}} + \tilde{\mathbf{y}} &= (1.2001, 0.99991, 0.92538)^t + (-0.20008, 8.9987 \times 10^{-5}, 0.074607)^t \\
&= (1.0000, 1.0000, 0.99999)^t.
\end{aligned}$$

This approximation has the residual vector

$$\|\tilde{\mathbf{r}}\|_\infty = \|\mathbf{b} - A(\tilde{\mathbf{x}} + \tilde{\mathbf{y}})\|_\infty = 0.0001.$$

If we were continuing the iteration processes, we would, of course, use $\tilde{\mathbf{x}} + \tilde{\mathbf{y}}$ as our starting values rather than $\tilde{\mathbf{x}}$. □

EXERCISE SET 7.6

1. Compute the l_∞ condition numbers of the following matrices.

   a. $\begin{bmatrix} \frac{1}{2} & \frac{1}{3} \\ \frac{1}{3} & \frac{1}{4} \end{bmatrix}$ b. $\begin{bmatrix} 3.9 & 1.6 \\ 6.8 & 2.9 \end{bmatrix}$

   c. $\begin{bmatrix} 1 & 2 \\ 1.0001 & 2 \end{bmatrix}$ d. $\begin{bmatrix} 1.003 & 58.09 \\ 5.550 & 321.8 \end{bmatrix}$

   e. $\begin{bmatrix} 1 & -1 & -1 \\ 0 & 1 & -1 \\ 0 & 0 & -1 \end{bmatrix}$ f. $\begin{bmatrix} 0.04 & 0.01 & -0.01 \\ 0.2 & 0.5 & -0.2 \\ 1 & 2 & 4 \end{bmatrix}$

2. The following linear systems $A\mathbf{x} = \mathbf{b}$ have $\mathbf{x}$ as the actual solution and $\tilde{\mathbf{x}}$ as an approximate solution. Using the results of Exercise 1, compute $\|\mathbf{x} - \tilde{\mathbf{x}}\|_\infty$ and

$$K_\infty(A)\frac{\|\mathbf{b} - A\tilde{\mathbf{x}}\|_\infty}{\|A\|_\infty}.$$

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. $3.9x_1 + 1.6x_2 = 5.5$,
$6.8x_1 + 2.9x_2 = 9.7$,
$\mathbf{x} = (1, 1)^t$, $\tilde{\mathbf{x}} = (0.98, 1.1)^t$.

c. $x_1 + 2x_2 = 3$,
$1.0001x_1 + 2x_2 = 3.0001$,
$\mathbf{x} = (1, 1)^t$, $\tilde{\mathbf{x}} = (0.96, 1.02)^t$.

d. $1.003x_1 + 58.09x_2 = 68.12$,
$5.550x_1 + 321.8x_2 = 377.3$,
$\mathbf{x} = (10, 1)^t$, $\tilde{\mathbf{x}} = (-10, 1)^t$.

e. $x_1 - x_2 - x_3 = 2\pi$,
$x_2 - x_3 = 0$,
$-x_3 = \pi$,
$\mathbf{x} = (0, -\pi, -\pi)^t$, $\tilde{\mathbf{x}} = (-0.1, -3.15, -3.14)^t$.

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

3. The linear system

$$\begin{bmatrix} 1 & 2 \\ 1.0001 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 3 \\ 3.0001 \end{bmatrix}$$

has the solution $(1, 1)^t$. Change A slightly to

$$\begin{bmatrix} 1 & 2 \\ 0.9999 & 2 \end{bmatrix}$$

and consider the linear system

$$\begin{bmatrix} 1 & 2 \\ 0.9999 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 3 \\ 3.0001 \end{bmatrix}.$$

Compute the new solution using five-digit rounding arithmetic, and compare the change in A to the change in $\mathbf{x}$.

4. The linear system $A\mathbf{x} = \mathbf{b}$ given by

$$\begin{bmatrix} 1 & 2 \\ 1.00001 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 3 \\ 3.00001 \end{bmatrix}$$

has the solution $(1, 1)^t$. Use seven-digit rounding arithmetic to find the solution of the perturbed system

$$\begin{bmatrix} 1 & 2 \\ 1.000011 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 3.00001 \\ 3.00003 \end{bmatrix},$$

and compare the change in A and $\mathbf{b}$ to the change in $\mathbf{x}$.

5. (i) Use Gaussian elimination and three-digit rounding arithmetic to approximate the solutions to the following linear systems. (ii) Then use one iteration of iterative refinement to improve the approximation, and compare the approximations to the actual solutions.

a. $0.03x_1 + 58.9x_2 = 59.2$
$5.31x_1 - 6.10x_2 = 47.0$
Actual solution $(10, 1)^t$.

b. $3.3330x_1 + 15920x_2 + 10.333x_3 = 7953$
$2.2220x_1 + 16.710x_2 + 9.6120x_3 = 0.965$
$-1.5611x_1 + 5.1792x_2 - 1.6855x_3 = 2.714$
Actual solution $(1, 0.5, -1)^t$.

c. $1.19x_1 + 2.11x_2 - 100x_3 + x_4 = 1.12$
$14.2x_1 - 0.122x_2 + 12.2x_3 - x_4 = 3.44$
$100x_2 - 99.9x_3 + x_4 = 2.15$
$15.3x_1 + 0.110x_2 - 13.1x_3 - x_4 = 4.16$
Actual solution $(0.17682530, 0.01269269, -0.02065405, -1.18260870)^t$.

d. $\pi x_1 - ex_2 + \sqrt{2}x_3 - \sqrt{3}x_4 = \sqrt{11}$
$\pi^2 x_1 + ex_2 - e^2 x_3 + \frac{3}{7}x_4 = 0$
$\sqrt{5}x_1 - \sqrt{6}x_2 + x_3 - \sqrt{2}x_4 = \pi$
$\pi^3 x_1 + e^2 x_2 - \sqrt{7}x_3 + \frac{1}{9}x_4 = \sqrt{2}$
Actual solution $(0.78839378, -3.12541367, 0.16759660, 4.55700252)^t$.

6. Repeat Exercise 5 using four-digit rounding arithmetic.

7. The $n \times n$ *Hilbert* matrix, $H^{(n)}$, defined by

$$H_{ij}^{(n)} = \frac{1}{i+j-1}, \quad 1 \leq i, j \leq n$$

is an ill-conditioned matrix that arises when solving for the coefficients of least squares polynomials (see Section 8.3, page 331).

a. Show that

$$[H^{(4)}]^{-1} = \begin{bmatrix} 16 & -120 & 240 & -140 \\ -120 & 1200 & -2700 & 1680 \\ 240 & -2700 & 6480 & -4200 \\ -140 & 1680 & -4200 & 2800 \end{bmatrix},$$

and compute $K_\infty(H^{(4)})$.

b. Show that

$$[H^{(5)}]^{-1} = \begin{bmatrix} 25 & -300 & 1050 & -1400 & 630 \\ -300 & 4800 & -18900 & 26880 & -12600 \\ 1050 & -18900 & 79380 & -117600 & 56700 \\ -1400 & 26880 & -117600 & 179200 & -88200 \\ 630 & -12600 & 56700 & -88200 & 44100 \end{bmatrix}$$

and compute $K_\infty(H^{(5)})$.

c. Solve the linear system

$$H^{(4)} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \end{bmatrix}$$

using three-digit rounding arithmetic, and compare the actual error to the residual vector error bound.

8. **a.** Use four-digit rounding arithmetic to compute the inverse H^{-1} of the 3×3 Hilbert matrix H.
b. Use four-digit rounding arithmetic to compute $\hat{H} = (H^{-1})^{-1}$.
c. Determine $\|H - \hat{H}\|_\infty$.

7.7 The Conjugate Gradient Method

Magnus Hestenes (1906–1991) and Eduard Stiefel (1907–1998) published the original paper on the conjugate gradient method in 1952 while working at the Institute for Numerical Analysis on the campus of UCLA.

The conjugate gradient method of Hestenes and Stiefel [HS] was originally developed as a direct method designed to solve an $n \times n$ positive definite linear system. As a direct method it is generally inferior to Gaussian elimination with pivoting because both methods require n steps to determine a solution, and the steps of the conjugate gradient method are more computationally expensive than those in Gaussian elimination.

However, the conjugate gradient method is useful when employed as an iterative approximation method for solving large sparse systems with nonzero entries occurring in predictable patterns. These problems frequently arise in the solution of boundary-value problems, and too much computation is required for direct methods in these situations. When the matrix has been preconditioned to make the calculations more effective, good results are obtained in only about $\sqrt{n}$ steps. Employed in this way, the method is preferred over Gaussian elimination and the previously discussed iterative methods.

Throughout this section we assume that the matrix A is positive definite. We will use the *inner product* notation

$$\langle \mathbf{x}, \mathbf{y} \rangle = \mathbf{x}^t \mathbf{y}, \tag{7.4}$$

where $\mathbf{x}$ and $\mathbf{y}$ are n-dimensional vectors. We will also need some additional standard results from linear algebra. A review of this material is found in Section 9.2.

The next result follows easily from the properties of transposes (see Exercise 12).

Inner Product Properties

For any vectors $\mathbf{x}$, $\mathbf{y}$, and $\mathbf{z}$ and any real number α, we have

(i) $\langle \mathbf{x}, \mathbf{y} \rangle = \langle \mathbf{y}, \mathbf{x} \rangle$;

(ii) $\langle \alpha\mathbf{x}, \mathbf{y} \rangle = \langle \mathbf{x}, \alpha\mathbf{y} \rangle = \alpha\langle \mathbf{x}, \mathbf{y} \rangle$;

(iii) $\langle \mathbf{x} + \mathbf{z}, \mathbf{y} \rangle = \langle \mathbf{x}, \mathbf{y} \rangle + \langle \mathbf{z}, \mathbf{y} \rangle$;

(iv) $\langle \mathbf{x}, \mathbf{x} \rangle \geq 0$;

(v) $\langle \mathbf{x}, \mathbf{x} \rangle = 0$ if and only if $\mathbf{x} = \mathbf{0}$.

When A is positive definite, $\langle \mathbf{x}, A\mathbf{x} \rangle = \mathbf{x}^t A\mathbf{x} > 0$ unless $\mathbf{x} = \mathbf{0}$. Also, because A is symmetric, we have

$$\langle \mathbf{x}, A\mathbf{y} \rangle = \mathbf{x}^t A\mathbf{y} = \mathbf{x}^t A^t \mathbf{y} = (A\mathbf{x})^t \mathbf{y} = \langle A\mathbf{x}, \mathbf{y} \rangle. \tag{7.5}$$

The following result is a basic tool in the development of the conjugate gradient method.

Minimization Condition for Positive Definite Matrices

The vector $\mathbf{x}$ is a solution to the positive definite linear system $A\mathbf{x} = \mathbf{b}$ if and only if $\mathbf{x}$ minimizes

$$g(\mathbf{x}) = \langle \mathbf{x}, A\mathbf{x} \rangle - 2\langle \mathbf{x}, \mathbf{b} \rangle.$$

To show this result we fix the vectors $\mathbf{x}$ and $\mathbf{v}$ and consider the single-variable function

$$h(t) = g(\mathbf{x} + t\mathbf{v}).$$