

THEOREM 2

Existence and Uniqueness Theorem

A linear system is consistent if and only if the rightmost column of the augmented matrix is *not* a pivot column—that is, if and only if an echelon form of the augmented matrix has *no* row of the form

$$[0 \quad \cdots \quad 0 \quad b] \quad \text{with } b \text{ nonzero}$$

If a linear system is consistent, then the solution set contains either (i) a unique solution, when there are no free variables, or (ii) infinitely many solutions, when there is at least one free variable.

THEOREM 4

Let A be an $m \times n$ matrix. Then the following statements are logically equivalent. That is, for a particular A , either they are all true statements or they are all false.

- For each \mathbf{b} in \mathbb{R}^m , the equation $A\mathbf{x} = \mathbf{b}$ has a solution.
- Each \mathbf{b} in \mathbb{R}^m is a linear combination of the columns of A .
- The columns of A span \mathbb{R}^m .
- A has a pivot position in every row.

Properties of the Matrix–Vector Product $A\mathbf{x}$

The facts in the next theorem are important and will be used throughout the text. The proof relies on the definition of $A\mathbf{x}$ and the algebraic properties of \mathbb{R}^n .

THEOREM 5

If A is an $m \times n$ matrix, \mathbf{u} and \mathbf{v} are vectors in \mathbb{R}^n , and c is a scalar, then:

- $A(\mathbf{u} + \mathbf{v}) = A\mathbf{u} + A\mathbf{v}$;
- $A(c\mathbf{u}) = c(A\mathbf{u})$.

THEOREM 3

Let A and B denote matrices whose sizes are appropriate for the following sums and products.

- $(A^T)^T = A$
- $(A + B)^T = A^T + B^T$
- For any scalar r , $(rA)^T = rA^T$
- $(AB)^T = B^T A^T$

THEOREM 4

Let $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$. If $ad - bc \neq 0$, then A is invertible and

$$A^{-1} = \frac{1}{ad - bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$$

If $ad - bc = 0$, then A is not invertible.

The simple proof of Theorem 4 is outlined in Exercises 25 and 26. The quantity $ad - bc$ is called the **determinant** of A , and we write

$$\det A = ad - bc$$

THEOREM 5

If A is an invertible $n \times n$ matrix, then for each \mathbf{b} in \mathbb{R}^n , the equation $A\mathbf{x} = \mathbf{b}$ has the unique solution $\mathbf{x} = A^{-1}\mathbf{b}$.

THEOREM 6

- a. If A is an invertible matrix, then A^{-1} is invertible and

$$(A^{-1})^{-1} = A$$

- b. If A and B are $n \times n$ invertible matrices, then so is AB , and the inverse of AB is the product of the inverses of A and B in the reverse order. That is,

$$(AB)^{-1} = B^{-1}A^{-1}$$

- c. If A is an invertible matrix, then so is A^T , and the inverse of A^T is the transpose of A^{-1} . That is,

$$(A^T)^{-1} = (A^{-1})^T$$

THEOREM 8**The Invertible Matrix Theorem**

Let A be a square $n \times n$ matrix. Then the following statements are equivalent. That is, for a given A , the statements are either all true or all false.

- A is an invertible matrix.
- A is row equivalent to the $n \times n$ identity matrix.
- A has n pivot positions.
- The equation $A\mathbf{x} = \mathbf{0}$ has only the trivial solution.
- The columns of A form a linearly independent set.
- The linear transformation $\mathbf{x} \mapsto A\mathbf{x}$ is one-to-one.
- The equation $A\mathbf{x} = \mathbf{b}$ has at least one solution for each \mathbf{b} in \mathbb{R}^n .
- The columns of A span \mathbb{R}^n .
- The linear transformation $\mathbf{x} \mapsto A\mathbf{x}$ maps \mathbb{R}^n onto \mathbb{R}^n .
- There is an $n \times n$ matrix C such that $CA = I$.
- There is an $n \times n$ matrix D such that $AD = I$.
- A^T is an invertible matrix.

THEOREM 13

The pivot columns of a matrix A form a basis for the column space of A .

THEOREM 2

If A is a triangular matrix, then $\det A$ is the product of the entries on the main diagonal of A .

THEOREM 5

If A is an $n \times n$ matrix, then $\det A^T = \det A$.

THEOREM 6**Multiplicative Property**

If A and B are $n \times n$ matrices, then $\det AB = (\det A)(\det B)$.

THEOREM**The Invertible Matrix Theorem (continued)**

Let A be an $n \times n$ matrix. Then A is invertible if and only if:

- The number 0 is *not* an eigenvalue of A .
- The determinant of A is *not* zero.

THEOREM 3

Properties of Determinants

Let A and B be $n \times n$ matrices.

- A is invertible if and only if $\det A \neq 0$.
- $\det AB = (\det A)(\det B)$.
- $\det A^T = \det A$.
- If A is triangular, then $\det A$ is the product of the entries on the main diagonal of A .
- A row replacement operation on A does not change the determinant. A row interchange changes the sign of the determinant. A row scaling also scales the determinant by the same scalar factor.

THEOREM 5

The Diagonalization Theorem

An $n \times n$ matrix A is diagonalizable if and only if A has n linearly independent eigenvectors.

In fact, $A = PDP^{-1}$, with D a diagonal matrix, if and only if the columns of P are n linearly independent eigenvectors of A . In this case, the diagonal entries of D are eigenvalues of A that correspond, respectively, to the eigenvectors in P .

THEOREM 6

An $n \times n$ matrix with n distinct eigenvalues is diagonalizable.

THEOREM 1

Let \mathbf{u} , \mathbf{v} , and \mathbf{w} be vectors in \mathbb{R}^n , and let c be a scalar. Then

- $\mathbf{u} \cdot \mathbf{v} = \mathbf{v} \cdot \mathbf{u}$
- $(\mathbf{u} + \mathbf{v}) \cdot \mathbf{w} = \mathbf{u} \cdot \mathbf{w} + \mathbf{v} \cdot \mathbf{w}$
- $(c\mathbf{u}) \cdot \mathbf{v} = c(\mathbf{u} \cdot \mathbf{v}) = \mathbf{u} \cdot (c\mathbf{v})$
- $\mathbf{u} \cdot \mathbf{u} \geq 0$, and $\mathbf{u} \cdot \mathbf{u} = 0$ if and only if $\mathbf{u} = \mathbf{0}$

Properties (b) and (c) can be combined several times to produce the following useful rule:

$$(c_1\mathbf{u}_1 + \cdots + c_p\mathbf{u}_p) \cdot \mathbf{w} = c_1(\mathbf{u}_1 \cdot \mathbf{w}) + \cdots + c_p(\mathbf{u}_p \cdot \mathbf{w})$$

The Length of a Vector

If \mathbf{v} is in \mathbb{R}^n , with entries v_1, \dots, v_n , then the square root of $\mathbf{v} \cdot \mathbf{v}$ is defined because $\mathbf{v} \cdot \mathbf{v}$ is nonnegative.

DEFINITION

The **length** (or **norm**) of \mathbf{v} is the nonnegative scalar $\|\mathbf{v}\|$ defined by

$$\|\mathbf{v}\| = \sqrt{\mathbf{v} \cdot \mathbf{v}} = \sqrt{v_1^2 + v_2^2 + \cdots + v_n^2}, \quad \text{and} \quad \|\mathbf{v}\|^2 = \mathbf{v} \cdot \mathbf{v}$$

DEFINITION

Two vectors \mathbf{u} and \mathbf{v} in \mathbb{R}^n are **orthogonal** (to each other) if $\mathbf{u} \cdot \mathbf{v} = 0$.

Observe that the zero vector is orthogonal to every vector in \mathbb{R}^n because $\mathbf{0}^T \mathbf{v} = 0$ for all \mathbf{v} .

The next theorem provides a useful fact about orthogonal vectors. The proof follows immediately from the calculation in (1) above and the definition of orthogonality. The right triangle shown in Figure 6 provides a visualization of the lengths that appear in the theorem.

THEOREM 2

The Pythagorean Theorem

Two vectors \mathbf{u} and \mathbf{v} are orthogonal if and only if $\|\mathbf{u} + \mathbf{v}\|^2 = \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2$.

$$\mathbf{u} + \mathbf{v}$$

THEOREM 6

An $m \times n$ matrix U has orthonormal columns if and only if $U^T U = I$.

PROOF To simplify notation, we suppose that U has only three columns, each a vector in \mathbb{R}^m . The proof of the general case is essentially the same. Let $U = [\mathbf{u}_1 \ \mathbf{u}_2 \ \mathbf{u}_3]$ and compute

$$U^T U = \begin{bmatrix} \mathbf{u}_1^T \\ \mathbf{u}_2^T \\ \mathbf{u}_3^T \end{bmatrix} [\mathbf{u}_1 \ \mathbf{u}_2 \ \mathbf{u}_3] = \begin{bmatrix} \mathbf{u}_1^T \mathbf{u}_1 & \mathbf{u}_1^T \mathbf{u}_2 & \mathbf{u}_1^T \mathbf{u}_3 \\ \mathbf{u}_2^T \mathbf{u}_1 & \mathbf{u}_2^T \mathbf{u}_2 & \mathbf{u}_2^T \mathbf{u}_3 \\ \mathbf{u}_3^T \mathbf{u}_1 & \mathbf{u}_3^T \mathbf{u}_2 & \mathbf{u}_3^T \mathbf{u}_3 \end{bmatrix} \quad (4)$$

The entries in the matrix at the right are inner products, using transpose notation. The columns of U are orthogonal if and only if

$$\mathbf{u}_1^T \mathbf{u}_2 = \mathbf{u}_2^T \mathbf{u}_1 = 0, \quad \mathbf{u}_1^T \mathbf{u}_3 = \mathbf{u}_3^T \mathbf{u}_1 = 0, \quad \mathbf{u}_2^T \mathbf{u}_3 = \mathbf{u}_3^T \mathbf{u}_2 = 0 \quad (5)$$

The columns of U all have unit length if and only if

$$\mathbf{u}_1^T \mathbf{u}_1 = 1, \quad \mathbf{u}_2^T \mathbf{u}_2 = 1, \quad \mathbf{u}_3^T \mathbf{u}_3 = 1 \quad (6)$$

The theorem follows immediately from (4)–(6). ■

THEOREM 7

Let U be an $m \times n$ matrix with orthonormal columns, and let \mathbf{x} and \mathbf{y} be in \mathbb{R}^n . Then

- $\|U\mathbf{x}\| = \|\mathbf{x}\|$
- $(U\mathbf{x}) \cdot (U\mathbf{y}) = \mathbf{x} \cdot \mathbf{y}$
- $(U\mathbf{x}) \cdot (U\mathbf{y}) = 0$ if and only if $\mathbf{x} \cdot \mathbf{y} = 0$

THEOREM 8

The Orthogonal Decomposition Theorem

Let W be a subspace of \mathbb{R}^n . Then each \mathbf{y} in \mathbb{R}^n can be written uniquely in the form

$$\mathbf{y} = \hat{\mathbf{y}} + \mathbf{z} \quad (1)$$

where $\hat{\mathbf{y}}$ is in W and \mathbf{z} is in W^\perp . In fact, if $\{\mathbf{u}_1, \dots, \mathbf{u}_p\}$ is any orthogonal basis of W , then

$$\hat{\mathbf{y}} = \frac{\mathbf{y} \cdot \mathbf{u}_1}{\mathbf{u}_1 \cdot \mathbf{u}_1} \mathbf{u}_1 + \dots + \frac{\mathbf{y} \cdot \mathbf{u}_p}{\mathbf{u}_p \cdot \mathbf{u}_p} \mathbf{u}_p \quad (2)$$

and $\mathbf{z} = \mathbf{y} - \hat{\mathbf{y}}$.

THEOREM 13

The set of least-squares solutions of $A\mathbf{x} = \mathbf{b}$ coincides with the nonempty set of solutions of the normal equations $A^T A\mathbf{x} = A^T \mathbf{b}$.

EXAMPLE 1 Find a least-squares solution of the inconsistent system $A\mathbf{x} = \mathbf{b}$ for

$$A = \begin{bmatrix} 4 & 0 \\ 0 & 2 \\ 1 & 1 \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} 2 \\ 0 \\ 11 \end{bmatrix}$$

SOLUTION To use normal equations (3), compute:

$$A^T A = \begin{bmatrix} 4 & 0 & 1 \\ 0 & 2 & 1 \end{bmatrix} \begin{bmatrix} 4 & 0 \\ 0 & 2 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 17 & 1 \\ 1 & 5 \end{bmatrix}$$

$$A^T \mathbf{b} = \begin{bmatrix} 4 & 0 & 1 \\ 0 & 2 & 1 \end{bmatrix} \begin{bmatrix} 2 \\ 0 \\ 11 \end{bmatrix} = \begin{bmatrix} 19 \\ 11 \end{bmatrix}$$

Then the equation $A^T A \mathbf{x} = A^T \mathbf{b}$ becomes

$$\begin{bmatrix} 17 & 1 \\ 1 & 5 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 19 \\ 11 \end{bmatrix}$$

Row operations can be used to solve this system, but since $A^T A$ is invertible and 2×2 , it is probably faster to compute

$$(A^T A)^{-1} = \frac{1}{84} \begin{bmatrix} 5 & -1 \\ -1 & 17 \end{bmatrix}$$

and then to solve $A^T A \mathbf{x} = A^T \mathbf{b}$ as

$$\begin{aligned} \hat{\mathbf{x}} &= (A^T A)^{-1} A^T \mathbf{b} \\ &= \frac{1}{84} \begin{bmatrix} 5 & -1 \\ -1 & 17 \end{bmatrix} \begin{bmatrix} 19 \\ 11 \end{bmatrix} = \frac{1}{84} \begin{bmatrix} 84 \\ 168 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \end{bmatrix} \end{aligned} \quad \blacksquare$$

In many calculations, $A^T A$ is invertible, but this is not always the case. The next example involves a matrix of the sort that appears in what are called *analysis of variance* problems in statistics.

THEOREM 14

Let A be an $m \times n$ matrix. The following statements are logically equivalent:

- The equation $A\mathbf{x} = \mathbf{b}$ has a unique least-squares solution for each \mathbf{b} in \mathbb{R}^m .
- The columns of A are linearly independent.
- The matrix $A^T A$ is invertible.

When these statements are true, the least-squares solution $\hat{\mathbf{x}}$ is given by

$$\hat{\mathbf{x}} = (A^T A)^{-1} A^T \mathbf{b} \quad (4)$$

1	$A^T A = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 4 & 2 & 2 \\ 2 & 2 & 0 \\ 2 & 0 & 2 \end{bmatrix}$ $A^T b = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 3 \\ 8 \\ 2 \end{bmatrix} = \begin{bmatrix} 14 \\ 4 \\ 10 \end{bmatrix}$	<p>We know that each least square solution of $Ax = b$ satisfies the equation $A^T Ax = A^T b$. Therefore first of all find $A^T A, A^T b$.</p>
2	$\begin{bmatrix} 4 & 2 & 2 & 4 \\ 2 & 2 & 0 & 4 \\ 2 & 0 & 2 & 10 \end{bmatrix}$	<p>The augmented matrix for $A^T Ax = A^T b$ to solve this system.</p>
3	$\begin{bmatrix} 4 & 2 & 2 & 4 \\ 2 & 2 & 0 & 4 \\ 2 & 0 & 2 & 10 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & 1 & 5 \\ 0 & 1 & -1 & -3 \\ 0 & 0 & 0 & 0 \end{bmatrix}$	<p>Use row operation to solve this system.</p>
4	$x_1 = 5 - x_3, x_2 = -3 + x_3, x_3 = 0$	<p>The solution of the augmented matrix. $x_3 = 0$ because it is free as we can see.</p>
5	$\hat{x} = \begin{bmatrix} 5 \\ -3 \\ 0 \end{bmatrix} + x_3 \begin{bmatrix} -1 \\ 1 \\ 1 \end{bmatrix}$	<p>The general least square solution for $Ax=b$.</p>
RESULT		$\hat{x} = \begin{bmatrix} 5 \\ -3 \\ 0 \end{bmatrix} + x_3 \begin{bmatrix} -1 \\ 1 \\ 1 \end{bmatrix}$