1 Linear Equations in Linear Algebra

1.1

SYSTEMS OF LINEAR EQUATIONS

PEARSON

© 2016 Pearson Education, Inc.

LINEAR EQUATION

- A **linear equation** in the variables $x_1, \ldots, x_n$ is an equation that can be written in the form

$$a_1x_1 + a_2x_2 + \cdots + a_nx_n = b$$

where b and the coefficients $a_1, \ldots, a_n$ are real or complex numbers, usually known in advance.

- A **system of linear equations** (or a **linear system**) is a collection of one or more linear equations involving the same variables — say, $x_1, \ldots, x_n$.

© 2016 Pearson Education, Inc.

LINEAR EQUATION

- A **solution** of the system is a list $(s_1, s_2, \ldots, s_n)$ of numbers that makes each equation a true statement when the values $s_1, \ldots, s_n$ are substituted for $x_1, \ldots, x_n$, respectively.

- The set of all possible solutions is called the **solution set** of the linear system.

- Two linear systems are called **equivalent** if they have the same solution set.

© 2016 Pearson Education, Inc.

LINEAR EQUATION

- A system of linear equations has
  1. no solution, or
  2. exactly one solution, or
  3. infinitely many solutions.
- A system of linear equations is said to be **consistent** if it has either one solution or infinitely many solutions.
- A system is **inconsistent** if it has no solution.

© 2016 Pearson Education, Inc.

MATRIX NOTATION

- The essential information of a linear system can be recorded compactly in a rectangular array called a **matrix**. Given the system,

$$\begin{aligned} x_1 - 2x_2 + x_3 &= 0 \\ 2x_2 - 8x_3 &= 8 \\ -4x_1 + 5x_2 + 9x_3 &= -9, \end{aligned}$$

with the coefficients of each variable aligned in columns, the matrix

$$\begin{bmatrix} 1 & -2 & 1 \\ 0 & 2 & -8 \\ -4 & 5 & 9 \end{bmatrix}$$

Is called the **coefficient matrix** (or matrix of **coefficients**) of the system

© 2016 Pearson Education, Inc.

MATRIX NOTATION

- An **augmented matrix** of a system consists of the coefficient matrix with an added column containing the constants from the right sides of the equations.

- For the given system of equations,

$$\begin{bmatrix} 1 & -2 & 1 & 0 \\ 0 & 2 & -8 & 8 \\ -4 & 5 & 9 & -9 \end{bmatrix}$$

is called the **augmented matrix** of the system.

© 2016 Pearson Education, Inc.

MATRIX SIZE

- The size of a matrix tells how many rows and columns it has. If m and n are positive integers, an $\boldsymbol{m \times n}$ **matrix** is a rectangular array of numbers with m rows and n columns. (The number of rows always comes first.)

- The basic strategy for solving a linear system is to *replace one system with an equivalent system (i.e., one with the same solution set) that is easier to solve.*

© 2016 Pearson Education, Inc.

SOLVING SYSTEM OF EQUATIONS

- **Example 1:** Solve the given system of equations.

$$x_1 - 2x_2 + x_3 = 0 \qquad \text{----(1)}$$

$$2x_2 - 8x_3 = 8 \qquad \text{----(2)}$$

$$-4x_1 + 5x_2 + 9x_3 = -9 \qquad \text{----(3)}$$

- **Solution:** The elimination procedure is shown here with and without matrix notation, and the results are placed side by side for comparison.

© 2016 Pearson Education, Inc.

SOLVING SYSTEM OF EQUATIONS

$$\begin{aligned} x_1 - 2x_2 + x_3 &= 0 \\ 2x_2 - 8x_3 &= 8 \\ -4x_1 + 5x_2 + 9x_3 &= -9 \end{aligned} \qquad \begin{bmatrix} 1 & -2 & 1 & 0 \\ 0 & 2 & -8 & 8 \\ -4 & 5 & 9 & -9 \end{bmatrix}$$

- Keep x_1 in the first equation and eliminate it from the other equations. To do so, add 4 times equation 1 to equation 3.

$$\begin{array}{r} 4x_1 - 8x_2 + 4x_3 = 0 \\ -4x_1 + 5x_2 + 9x_3 = -9 \\ \hline -3x_2 + 13x_3 = -9 \end{array}$$

© 2016 Pearson Education, Inc.

SOLVING SYSTEM OF EQUATIONS

- The result of this calculation is written in place of the original third equation:

$$\begin{aligned} x_1 - 2x_2 + x_3 &= 0 \\ 2x_2 - 8x_3 &= 8 \\ -3x_2 + 13x_3 &= -9 \end{aligned} \qquad \begin{bmatrix} 1 & -2 & 1 & 0 \\ 0 & 2 & -8 & 8 \\ 0 & -3 & 13 & -9 \end{bmatrix}$$

- Now, multiply equation 2 by $1/2$ in order to obtain 1 as the coefficient for x_2.

© 2016 Pearson Education, Inc.

SOLVING SYSTEM OF EQUATIONS

$$\begin{aligned} x_1 - 2x_2 + x_3 &= 0 \\ x_2 - 4x_3 &= 4 \\ -3x_2 + 13x_3 &= -9 \end{aligned} \qquad \begin{bmatrix} 1 & -2 & 1 & 0 \\ 0 & 1 & -4 & 4 \\ 0 & -3 & 13 & -9 \end{bmatrix}$$

- Use the x_2 in equation 2 to eliminate the $-3x_2$ in equation 3.

$$\begin{aligned} 3x_2 - 12x_3 &= 12 \\ -3x_2 + 13x_3 &= -9 \\ \hline x_3 &= 3 \end{aligned}$$

© 2016 Pearson Education, Inc.

SOLVING SYSTEM OF EQUATIONS

- The new system has a *triangular* form.

$$\begin{aligned} x_1 - 2x_2 + x_3 &= 0 \\ x_2 - 4x_3 &= 4 \\ \hline x_3 &= 3 \end{aligned} \qquad \begin{bmatrix} 1 & -2 & 1 & 0 \\ 0 & 1 & -4 & 4 \\ 0 & 0 & 1 & 3 \end{bmatrix}$$

- Eventually, you want to eliminate the $-2x_2$ term from equation 1, but it is more efficient to use the x_3 term in equation 3 first to eliminate the $-4x_3$ and x_3 terms in equations 2 and 1.

© 2016 Pearson Education, Inc.

SOLVING SYSTEM OF EQUATIONS

$$\begin{aligned} 4x_3 &= 12 \\ x_2 - 4x_3 &= 4 \\ \hline x_2 &= 16 \end{aligned} \qquad \begin{aligned} -x_3 &= -3 \\ x_1 - 2x_2 + x_3 &= 0 \\ \hline x_1 - 2x_2 &= -3 \end{aligned}$$

- Now, combine the results of these two operations.

$$\begin{aligned} x_1 - 2x_2 &= -3 \\ x_2 &= 16 \\ x_3 &= 3 \end{aligned} \qquad \begin{bmatrix} 1 & -2 & 0 & -3 \\ 0 & 1 & 0 & 16 \\ 0 & 0 & 1 & 3 \end{bmatrix}$$

© 2016 Pearson Education, Inc.

SOLVING SYSTEM OF EQUATIONS

- Move back to the x_2 in equation 2, and use it to eliminate the $-2x_2$ above it. Because of the previous work with x_3, there is now no arithmetic involving x_3 terms. Add 2 times equation 2 to equation 1 and obtain the system:

$$\begin{aligned} x_1 &= 29 \\ x_2 &= 16 \\ x_3 &= 3 \end{aligned} \qquad \begin{bmatrix} 1 & 0 & 0 & 29 \\ 0 & 1 & 0 & 16 \\ 0 & 0 & 1 & 3 \end{bmatrix}$$

© 2016 Pearson Education, Inc.

SOLVING SYSTEM OF EQUATIONS

- Thus, the only solution of the original system is (29,16,3). To verify that (29,16,3) is a solution, substitute these values into the left side of the original system, and compute.

$$(29) - 2(16) + (3) = 29 - 32 + 3 = 0$$

$$2(16) - 8(3) = 32 - 24 = 8$$

$$-4(29) + 5(16) + 9(3) = -116 + 80 + 27 = -9$$

- The results agree with the right side of the original system, so (29,16,3) is a solution of the system.

© 2016 Pearson Education, Inc.

ELEMENTARY ROW OPERATIONS

- Elementary row operations include the following:
  1. (Replacement) Replace one row by the sum of itself and a multiple of another row.
  2. (Interchange) Interchange two rows.
  3. (Scaling) Multiply all entries in a row by a nonzero constant.

- Two matrices are called **row equivalent** if there is a sequence of elementary row operations that transforms one matrix into the other.

© 2016 Pearson Education, Inc.

ELEMENTARY ROW OPERATIONS

- It is important to note that row operations are *reversible*.
- If the augmented matrices of two linear systems are row equivalent, then the two systems have the same solution set.
- Two fundamental questions about a linear system are as follows:
  1. Is the system consistent; that is, does at least one solution *exist*?
  2. If a solution exists, is it the *only* one; that is, is the solution *unique*?

© 2016 Pearson Education, Inc.

EXISTENCE AND UNIQUENESS OF SYSTEM OF EQUATIONS

- **Example 3:** Determine if the following system is consistent:

$$\begin{aligned} x_2 - 4x_3 &= 8 \\ 2x_1 - 3x_2 + 2x_3 &= 1 \\ 5x_1 - 8x_2 + 7x_3 &= 1 \end{aligned} \qquad (5)$$

- **Solution:** The augmented matrix is

$$\begin{bmatrix} 0 & 1 & -4 & 8 \\ 2 & -3 & 2 & 1 \\ 5 & -8 & 7 & 1 \end{bmatrix}$$

© 2016 Pearson Education, Inc.

EXISTENCE AND UNIQUENESS OF SYSTEM OF EQUATIONS

- To obtain an x_1 in in the first equation, interchange rows 1 and 2:

$$\begin{bmatrix} 2 & -3 & 2 & 1 \\ 0 & 1 & -4 & 8 \\ 5 & -8 & 7 & 1 \end{bmatrix}$$

- To eliminate the $5x_1$ term in the third equation, add $-5/2$ times row 1 to row 3.

$$\begin{bmatrix} 2 & -3 & 2 & 1 \\ 0 & 1 & -4 & 8 \\ 0 & -1/2 & 2 & -3/2 \end{bmatrix} \qquad (6)$$

© 2016 Pearson Education, Inc.

EXISTENCE AND UNIQUENESS OF SYSTEM OF EQUATIONS

- Next, use the x_2 term in the second equation to eliminate the $-(1/2)x_2$ term from the third equation. Add $1/2$ times row 2 to row 3.

$$\begin{bmatrix} 2 & -3 & 2 & 1 \\ 0 & 1 & -4 & 8 \\ 0 & 0 & 0 & 5/2 \end{bmatrix} \qquad (7)$$

- The augmented matrix is now in triangular form. To interpret it correctly, go back to equation notation.

$$\begin{aligned} 2x_1 - 3x_2 + 2x_3 &= 1 \\ x_2 - 4x_3 &= 8 \\ 0 &= 5/2 \end{aligned} \qquad (8)$$

© 2016 Pearson Education, Inc.

EXISTENCE AND UNIQUENESS OF SYSTEM OF EQUATIONS

- The equation $0 = 5/2$ is a short form of $0x_1 + 0x_2 + 0x_3 = 5/2.$

- There are no values of x_1, x_2, x_3 that satisfy (8) because the equation $0 = 5/2$ is never true.

- Since (8) and (5) have the same solution set, the original system is inconsistent (*i.e.*, has no solution).

© 2016 Pearson Education, Inc.

1 Linear Equations in Linear Algebra

1.2

Row Reduction and Echelon Forms

PEARSON

© 2016 Pearson Education, Ltd.

ROW REDUCTION AND ECHELON FORMS

- In the definitions that follow, a *nonzero* row or column in a matrix means a row or column that contains at least one nonzero entry; a **leading entry** of a row refers to the leftmost nonzero entry (in a nonzero row).

© 2016 Pearson Education, Ltd.

ECHELON FORM

- A rectangular matrix is in **echelon form** (or **row echelon form, REF**) if it has the following three properties:
  1. All nonzero rows are above any rows of all zeros.
  2. Each leading entry of a row is in a column to the right of the leading entry of the row above it.
  3. All entries in a column below a leading entry are zeros.

© 2016 Pearson Education, Ltd.

ECHELON FORM

- If a matrix in echelon form satisfies the following additional conditions, then it is in **reduced echelon form** (or **reduced row echelon form, RREF**):

  4. The leading entry in each nonzero row is 1.

  5. Each leading 1 is the only nonzero entry in its column.

- An **echelon matrix** (respectively, **reduced echelon matrix**) is one that is in echelon form (respectively, reduced echelon form.)

© 2016 Pearson Education, Ltd.

ECHELON FORM

- Any nonzero matrix may be **row reduced** (i.e., transformed by elementary row operations) into more than one matrix in echelon form, using different sequences of row operations. However, the reduced echelon form one obtains from a matrix is unique.

THEOREM 1

Uniqueness of the Reduced Echelon Form
Each matrix is row equivalent to one and only one reduced echelon matrix.

© 2016 Pearson Education, Ltd.

PIVOT POSITION

- If a matrix A is row equivalent to an echelon matrix U, we call U **an echelon form** (or row echelon form) **of A**; if U is in reduced echelon form, we call U **the reduced echelon form of A**.

- A **pivot position** in a matrix A is a location in A that corresponds to a leading 1 in the reduced echelon form of A. A **pivot column** is a column of A that contains a pivot position.

© 2016 Pearson Education, Ltd.

PIVOT POSITION

- **Example 2:** Row reduce the matrix A below to echelon form, and locate the pivot columns of A.

$$A = \begin{bmatrix} 0 & -3 & -6 & 4 & 9 \\ -1 & -2 & -1 & 3 & 1 \\ -2 & -3 & 0 & 3 & -1 \\ 1 & 4 & 5 & -9 & -7 \end{bmatrix}$$

- **Solution:** The top of the leftmost nonzero column is the first pivot position. A nonzero entry, or *pivot*, must be placed in this position.

© 2016 Pearson Education, Ltd.

PIVOT POSITION

- Now, interchange rows 1 and 4.

Pivot

$$\begin{bmatrix} 1 & 4 & 5 & -9 & -7 \\ -1 & -2 & -1 & 3 & 1 \\ -2 & -3 & 0 & 3 & -1 \\ 0 & -3 & -6 & 4 & 9 \end{bmatrix}$$

Pivot column

- Create zeros below the pivot, 1, by adding multiples of the first row to the rows below, and obtain the next matrix.

© 2016 Pearson Education, Ltd.

PIVOT POSITION

- Choose the 2 in the second row as the next pivot.

Pivot

$$\begin{bmatrix} 1 & 4 & 5 & -9 & -7 \\ 0 & 2 & 4 & -6 & -6 \\ 0 & 5 & 10 & -15 & -15 \\ 0 & -3 & -6 & 4 & 9 \end{bmatrix}$$

Next pivot column

- Add $-5/2$ times row 2 to row 3, and add $3/2$ times row 2 to row 4.

© 2016 Pearson Education, Ltd.

PIVOT POSITION

$$\begin{bmatrix} 1 & 4 & 5 & -9 & -7 \\ 0 & 2 & 4 & -6 & -6 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -5 & 0 \end{bmatrix}$$

- There is no way to create a leading entry in column 3! However, if we interchange rows 3 and 4, we can produce a leading entry in column 4.

© 2016 Pearson Education, Ltd.

PIVOT POSITION

Pivot

$$\begin{bmatrix} 1 & 4 & 5 & -9 & -7 \\ 0 & 2 & 4 & -6 & -6 \\ 0 & 0 & 0 & -5 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Pivot columns

- The matrix is in echelon form and thus reveals that columns 1, 2, and 4 of A are pivot columns.

© 2016 Pearson Education, Ltd.

PIVOT POSITION

Pivot positions

$$A = \begin{bmatrix} \boxed{0} & -3 & -6 & 4 & 9 \\ -1 & \boxed{-2} & -1 & 3 & 1 \\ -2 & -3 & 0 & \boxed{3} & -1 \\ 1 & 4 & 5 & -9 & -7 \end{bmatrix}$$

Pivot columns

- The pivots in the example are 1, 2 and –5.

© 2016 Pearson Education, Ltd.

ROW REDUCTION ALGORITHM

- **Example 3:** Apply elementary row operations to transform the following matrix first into echelon form and then into reduced echelon form:

$$\begin{bmatrix} 0 & 3 & -6 & 6 & 4 & -5 \\ 3 & -7 & 8 & -5 & 8 & 9 \\ 3 & -9 & 12 & -9 & 6 & 15 \end{bmatrix}$$

- **Solution:**
- **STEP 1:** Begin with the *leftmost* nonzero column. This is a pivot column. The pivot position is at the top.

© 2016 Pearson Education, Ltd.

ROW REDUCTION ALGORITHM

$$\begin{bmatrix} 0 & 3 & -6 & 6 & 4 & -5 \\ 3 & -7 & 8 & -5 & 8 & 9 \\ 3 & -9 & 12 & -9 & 6 & 15 \end{bmatrix}$$

Pivot column

- **STEP 2:** Select a nonzero entry in the pivot column as a pivot. If necessary, interchange rows to move this entry into the pivot position.

© 2016 Pearson Education, Ltd.

ROW REDUCTION ALGORITHM

- Interchange rows 1 and 3. (We could have interchanged rows 1 and 2 instead)

Pivot

$$\begin{bmatrix} 3 & -9 & 12 & -9 & 6 & 15 \\ 3 & -7 & 8 & -5 & 8 & 9 \\ 0 & 3 & -6 & 6 & 4 & -5 \end{bmatrix}$$

- **STEP 3:** Use row replacement operations to create zeros in all positions below the pivot.

© 2016 Pearson Education, Ltd.

ROW REDUCTION ALGORITHM

- We could have divided the top row by the pivot, 3, but with two 3s in column 1, it is just as easy to add -1 times row 1 to row 2.

Pivot

$$\begin{bmatrix} 3 & -9 & 12 & -9 & 6 & 15 \\ 0 & 2 & -4 & 4 & 2 & -6 \\ 0 & 3 & -6 & 6 & 4 & -5 \end{bmatrix}$$

- **STEP 4:** Cover (or ignore) the row containing the pivot position and cover all rows, if any, above it. Apply steps 1–3 to the submatrix that remains. Repeat the process until there are no more nonzero rows to modify.

© 2016 Pearson Education, Ltd.

ROW REDUCTION ALGORITHM

- With row 1 covered, step 1 shows that column 2 is the next pivot column; for step 2, select as a pivot the "top" entry in that column.

Pivot

$$\begin{bmatrix} 3 & -9 & 12 & -9 & 6 & 15 \\ 0 & 2 & -4 & 4 & 2 & -6 \\ 0 & 3 & -6 & 6 & 4 & -5 \end{bmatrix}$$

New pivot column

- For step 3, we could insert an optional step of dividing the "top" row of the submatrix by the pivot, 2. Instead, we add $-3/2$ times the "top" row to the row below.

© 2016 Pearson Education, Ltd.

ROW REDUCTION ALGORITHM

- This produces the following matrix.

$$\begin{bmatrix} 3 & -9 & 12 & -9 & 6 & 15 \\ 0 & 2 & -4 & 4 & 2 & -6 \\ 0 & 0 & 0 & 0 & 1 & 4 \end{bmatrix}$$

- When we cover the row containing the second pivot position for step 4, we are left with a new submatrix that has only one row.

$$\begin{bmatrix} 3 & -9 & 12 & -9 & 6 & 15 \\ 0 & 2 & -4 & 4 & 2 & -6 \\ 0 & 0 & 0 & 0 & 1 & 4 \end{bmatrix}$$

© 2016 Pearson Education, Ltd.

ROW REDUCTION ALGORITHM

- Steps 1–3 require no work for this submatrix, and we have reached an echelon form of the full matrix. We perform one more step to obtain the reduced echelon form.
- **STEP 5:** Beginning with the *rightmost* pivot and working upward and to the left, create zeros above each pivot. If a pivot is not 1, make it 1 by a scaling operation.
- The rightmost pivot is in row 3. Create zeros above it, adding suitable multiples of row 3 to rows 2 and 1.

© 2016 Pearson Education, Ltd.

ROW REDUCTION ALGORITHM

$$\begin{bmatrix} 3 & -9 & 12 & -9 & 0 & -9 \\ 0 & 2 & -4 & 4 & 0 & -14 \\ 0 & 0 & 0 & 0 & 1 & 4 \end{bmatrix} \begin{array}{l} \leftarrow \text{Row } 1 + (-6) \times \text{row } 3 \\ \leftarrow \text{Row } 2 + (-2) \times \text{row } 3 \\ \end{array}$$

- The next pivot is in row 2. Scale this row, dividing by the pivot.

$$\begin{bmatrix} 3 & -9 & 12 & -9 & 0 & -9 \\ 0 & 1 & -2 & 2 & 0 & -7 \\ 0 & 0 & 0 & 0 & 1 & 4 \end{bmatrix} \begin{array}{l} \\ \leftarrow \text{Row scaled by } \dfrac{1}{2} \\ \end{array}$$

© 2016 Pearson Education, Ltd.

ROW REDUCTION ALGORITHM

- Create a zero in column 2 by adding 9 times row 2 to row 1.

$$\begin{bmatrix} 3 & 0 & -6 & 9 & 0 & -72 \\ 0 & 1 & -2 & 2 & 0 & -7 \\ 0 & 0 & 0 & 0 & 1 & 4 \end{bmatrix} \leftarrow \text{Row } 1 + (9) \times \text{row } 2$$

- Finally, scale row 1, dividing by the pivot, 3.

© 2016 Pearson Education, Ltd.

ROW REDUCTION ALGORITHM

$$\begin{bmatrix} 1 & 0 & -2 & 3 & 0 & -24 \\ 0 & 1 & -2 & 2 & 0 & -7 \\ 0 & 0 & 0 & 0 & 1 & 4 \end{bmatrix} \leftarrow \text{Row scaled by } \frac{1}{3}$$

- This is the reduced echelon form of the original matrix.

- The combination of steps 1–4 is called the **forward phase** of the row reduction algorithm. Step 5, which produces the unique reduced echelon form, is called the **backward phase**.

© 2016 Pearson Education, Ltd.

SOLUTIONS OF LINEAR SYSTEMS

- The row reduction algorithm leads to an explicit description of the solution set of a linear system when the algorithm is applied to the augmented matrix of the system.
- Suppose, for example, that the augmented matrix of a linear system has been changed into the equivalent *reduced* echelon form.

$$\begin{bmatrix} 1 & 0 & -5 & 1 \\ 0 & 1 & 1 & 4 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

© 2016 Pearson Education, Ltd.

SOLUTIONS OF LINEAR SYSTEMS

- There are three variables because the augmented matrix has four columns. The associated system of equations is

$$\begin{aligned} x_1 - 5x_3 &= 1 \\ x_2 + x_3 &= 4 \\ 0 &= 0 \end{aligned} \qquad (4)$$

- The variables x_1 and x_2 corresponding to pivot columns in the matrix are called **basic variables**. The other variable, x_3, is called a **free variable**.

© 2016 Pearson Education, Ltd.

SOLUTIONS OF LINEAR SYSTEMS

- Whenever a system is consistent, as in (4), the solution set can be described explicitly by solving the *reduced* system of equations for the basic variables in terms of the free variables.

- This operation is possible because the reduced echelon form places each basic variable in one and only one equation.

- In (4), solve the first equation for x_1 and the second for x_2. (Ignore the third equation; it offers no restriction on the variables.)

© 2016 Pearson Education, Ltd.

SOLUTIONS OF LINEAR SYSTEMS

$$\begin{aligned} x_1 &= 1 + 5x_3 \\ x_2 &= 4 - x_3 \qquad\qquad (5) \\ x_3 &\text{ is free} \end{aligned}$$

- The statement "x_3 is free" means that you are free to choose any value for x_3. Once that is done, the formulas in (5) determine the values for x_1 and x_2. For instance, when $x_3 = 0$, the solution is (1,4,0); when $x_3 = 1$, the solution is (6,3,1).

- *Each different choice of x_3 determines a (different) solution of the system, and every solution of the system is determined by a choice of x_3.*

© 2016 Pearson Education, Ltd.

PARAMETRIC DESCRIPTIONS OF SOLUTION SETS

- The descriptions in (5) and (7) are *parametric descriptions* of solutions sets in which the free variables act as parameters.

- *Solving a system* amounts to finding a parametric description of the solution set or determining that the solution set is empty.

- Whenever a system is consistent and has free variables, the solution set has many parametric descriptions.

© 2016 Pearson Education, Ltd.

PARAMETRIC DESCRIPTIONS OF SOLUTION SETS

- For instance, in system (4), we may add 5 times equation 2 to equation 1 and obtain the following equivalent system

$$x_1 + 5x_2 = 21$$

$$x_2 + x_3 = 4$$

- We could treat x_2 as a parameter and solve for x_1 and x_3 in terms of x_2, and we would have an accurate description of the solution set.
- Whenever a system is inconsistent, the solution set is empty, even when the system has free variables. In this case, the solution set has *no* parametric representation.

© 2016 Pearson Education, Ltd.

EXISTENCE AND UNIQUENESS THEOREM

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—i.e., if and only if an echelon form of the augmented matrix has *no* row of the form

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

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 on free variable.

© 2016 Pearson Education, Ltd.

ROW REDUCTION TO SOLVE A LINEAR SYSTEM

Using Row Reduction to Solve a Linear System

1. Write the augmented matrix of the system.
2. Use the row reduction algorithm to obtain an equivalent augmented matrix in echelon form. Decide whether the system is consistent. If there is no solution, stop; otherwise, go to the next step.
3. Continue row reduction to obtain the reduced echelon form.
4. Write the system of equations corresponding to the matrix obtained in step 3.

© 2016 Pearson Education, Ltd.

ROW REDUCTION TO SOLVE A LINEAR SYSTEM

5. Rewrite each nonzero equation from step 4 so that its one basic variable is expressed in terms of any free variables appearing in the equation.

© 2016 Pearson Education, Ltd.

1 Linear Equations in Linear Algebra

1.3

VECTOR EQUATIONS

© 2016 Pearson Education, Ltd.

VECTOR EQUATIONS

Vectors in $\mathbb{R}^2$

- A matrix with only one column is called a **column vector**, or simply a **vector**.
- An example of a vector with two entries is

$$\mathrm{w} = \begin{bmatrix} w_1 \\ w_2 \end{bmatrix},$$

where w_1 and w_2 are any real numbers.

- The set of all vectors with two entries is denoted by $\mathbb{R}^2$ (read "r-two").

© 2016 Pearson Education, Ltd.

VECTOR EQUATIONS

- The $\mathbb{R}$ stands for the real numbers that appear as entries in the vector, and the exponent 2 indicates that each vector contains two entries.
- Two vectors in $\mathbb{R}^2$ are **equal** if and only if their corresponding entries are equal.
- Given two vectors **u** and **v** in $\mathbb{R}^2$, their **sum** is the vector $\mathbf{u} + \mathbf{v}$ obtained by adding corresponding entries of **u** and **v**.
- Given a vector **u** and a real number c, the **scalar multiple** of **u** by c is the vector $c\mathbf{u}$ obtained by multiplying each entry in **u** by c.

© 2016 Pearson Education, Ltd.

VECTOR EQUATIONS

Example 1: Given $\mathrm{u} = \begin{bmatrix} 1 \\ -2 \end{bmatrix}$ and $\mathrm{v} = \begin{bmatrix} 2 \\ -5 \end{bmatrix}$, find

$4\mathbf{u}, (-3)\mathrm{v}$, and $4\mathrm{u} + (-3)\mathrm{v}$.

Solution: $4\mathrm{u} = \begin{bmatrix} 4 \\ -8 \end{bmatrix}, (-3)\mathrm{v} = \begin{bmatrix} -6 \\ 15 \end{bmatrix}$ and

$$4\mathrm{u} + (-3)\mathrm{v} = \begin{bmatrix} 4 \\ -8 \end{bmatrix} + \begin{bmatrix} -6 \\ 15 \end{bmatrix} = \begin{bmatrix} -2 \\ 7 \end{bmatrix}$$

© 2016 Pearson Education, Ltd.

GEOMETRIC DESCRIPTIONS OF $\mathbb{R}^2$

- Consider a rectangular coordinate system in the plane. Because each point in the plane is determined by an ordered pair of numbers, *we can identify a geometric point (a, b) with the column vector* $\begin{bmatrix} a \\ b \end{bmatrix}$.

- So we may regard $\mathbb{R}^2$ as the set of all points in the plane.

FIGURE 1 Vectors as points.

FIGURE 2 Vectors with arrows.

© 2016 Pearson Education, Ltd.

PARALLELOGRAM RULE FOR ADDITION

- If $\mathbf{u}$ and $\mathbf{v}$ in $\mathbb{R}^2$ are represented as points in the plane, then $\mathbf{u} + \mathbf{v}$ corresponds to the fourth vertex of the parallelogram whose other vertices are $\mathbf{u}$, $\mathbf{0}$, and $\mathbf{v}$. See Fig. 3 below.

© 2016 Pearson Education, Ltd.

VECTORS IN $\mathbb{R}^3$ and $\mathbb{R}^n$

- Vectors in $\mathbb{R}^3$ are 3×1 column matrices with three entries.
- They are represented geometrically by points in a three-dimensional coordinate space, with arrows from the origin sometimes included for visual clarity.
- If n is a positive integer, $\mathbb{R}^n$ (read "r-n") denotes the collection of all lists (or *ordered n-tuples*) of n real numbers, usually written as $n\times 1$ column matrices, such as

$$\mathbf{u} = \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix}.$$

FIGURE 6
Scalar multiples.

© 2016 Pearson Education, Ltd.

ALGEBRAIC PROPERTIES OF $\mathbb{R}^n$

- The vector whose entries are all zero is called the **zero vector** and is denoted by **0**.
- For all **u**, **v**, **w** in $\mathbb{R}^n$ and all scalars c and d:

(i) $\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}$

(ii) $(\mathbf{u} + \mathbf{v}) + \mathbf{w} = \mathbf{u} + (\mathbf{v} + \mathbf{w})$

(iii) $\mathbf{u} + 0 = 0 + \mathbf{u} = \mathbf{u}$

(iv) $\mathbf{u} + (-\mathbf{u}) = -\mathbf{u} + \mathbf{u} = 0$,

where $-\mathbf{u}$ denotes $(-1)\mathbf{u}$

(v) $c(\mathbf{u} + \mathbf{v}) = c\mathbf{u} + c\mathbf{v}$

(vi) $(c + d)\mathbf{u} = c\mathbf{u} + d\mathbf{u}$

© 2016 Pearson Education, Ltd.

LINEAR COMBINATIONS

- (vii) $c(d\mathrm{u})=(\mathrm{cd})(\mathrm{u})$

  (viii) $1\mathrm{u} = \mathrm{u}$

- Given vectors $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_p$ in $\mathbb{R}^n$ and given scalars $c_1, c_2, \ldots, c_p$, the vector $\mathbf{y}$ defined by

$$\mathrm{y} = c_1\mathrm{v}_1 + \ldots + c_p\mathrm{v}_p$$

  is called a **linear combination** of $\mathbf{v}_1, \ldots, \mathbf{v}_p$ with **weights** $c_1, \ldots, c_p$.

- The weights in a linear combination can be any real numbers, including zero.

© 2016 Pearson Education, Ltd.

- Example:
  - $u = 3v_1 - 2v_2$
  - $w = (5/2)v_1 - (1/2)v_2$

FIGURE 8 Linear combinations of v_1 and v_2.

LINEAR COMBINATIONS

- **Example 5:** Let $\mathbf{a}_1 = \begin{bmatrix} 1 \\ -2 \\ -5 \end{bmatrix}$, $\mathbf{a}_2 = \begin{bmatrix} 2 \\ 5 \\ 6 \end{bmatrix}$ and $\mathbf{b} = \begin{bmatrix} 7 \\ 4 \\ -3 \end{bmatrix}$.

  Determine whether **b** can be generated (or written) as a linear combination of $\mathbf{a}_1$ and $\mathbf{a}_2$. That is, determine whether weights x_1 and x_2 exist such that

$$x_1\mathbf{a}_1 + x_2\mathbf{a}_2 = \mathbf{b} \qquad (1)$$

  If vector equation (1) has a solution, find it.

© 2016 Pearson Education, Ltd.

LINEAR COMBINATIONS

Solution: Use the definitions of scalar multiplication and vector addition to rewrite the vector equation

$$x_1 \underbrace{\begin{bmatrix} 1 \\ -2 \\ -5 \end{bmatrix}}_{\mathbf{a}_1} + x_2 \underbrace{\begin{bmatrix} 2 \\ 5 \\ 6 \end{bmatrix}}_{\mathbf{a}_2} = \underbrace{\begin{bmatrix} 7 \\ 4 \\ -3 \end{bmatrix}}_{\mathbf{b}},$$

which is same as

$$\begin{bmatrix} x_1 \\ -2x_1 \\ -5x_1 \end{bmatrix} + \begin{bmatrix} 2x_2 \\ 5x_2 \\ 6x_2 \end{bmatrix} = \begin{bmatrix} 7 \\ 4 \\ -3 \end{bmatrix}$$

© 2016 Pearson Education, Ltd.

LINEAR COMBINATIONS

and $$\begin{bmatrix} x_1 + 2x_2 \\ -2x_1 + 5x_2 \\ -5x_1 + 6x_2 \end{bmatrix} = \begin{bmatrix} 7 \\ 4 \\ -3 \end{bmatrix}. \qquad (2)$$

- The vectors on the left and right sides of (2) are equal if and only if their corresponding entries are both equal. That is, x_1 and x_2 make the vector equation (1) true if and only if x_1 and x_2 satisfy the following system.

$$\begin{aligned} x_1 + 2x_2 &= 7 \\ -2x_1 + 5x_2 &= 4 \\ -5x_1 + 6x_2 &= -3 \end{aligned} \qquad (3)$$

© 2016 Pearson Education, Ltd.

LINEAR COMBINATIONS

- To solve this system, row reduce the augmented matrix of the system as follows:

$$\begin{bmatrix} 1 & 2 & 7 \\ -2 & 5 & 4 \\ -5 & 6 & -3 \end{bmatrix} \sim \begin{bmatrix} 1 & 2 & 7 \\ 0 & 9 & 18 \\ 0 & 16 & 32 \end{bmatrix} \sim \begin{bmatrix} 1 & 2 & 7 \\ 0 & 1 & 2 \\ 0 & 16 & 32 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & 3 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \end{bmatrix}$$

- The solution of (3) is $x_1 = 3$ and $x_2 = 2$. Hence $\mathbf{b}$ is a linear combination of $\mathbf{a}_1$ and $\mathbf{a}_2$, with weights $x_1 = 3$ and $x_2 = 2$. That is,

$$3\begin{bmatrix} 1 \\ -2 \\ -5 \end{bmatrix} + 2\begin{bmatrix} 2 \\ 5 \\ 6 \end{bmatrix} = \begin{bmatrix} 7 \\ 4 \\ -3 \end{bmatrix}.$$

© 2016 Pearson Education, Ltd.

LINEAR COMBINATIONS

- Now, observe that the original vectors $\mathbf{a}_1$, $\mathbf{a}_2$, and $\mathbf{b}$ are the columns of the augmented matrix that we row reduced:

$$\begin{array}{c} \begin{bmatrix} 1 & 2 & 7 \\ -2 & 5 & 4 \\ -5 & 6 & -3 \end{bmatrix} \\ \begin{array}{ccc} \uparrow & \uparrow & \uparrow \\ \mathbf{a}_1 & \mathbf{a}_2 & \mathbf{b} \end{array} \end{array}$$

- Write this matrix in a way that identifies its columns.

$$\begin{bmatrix} a_1 & a_2 & b \end{bmatrix} \qquad (4)$$

© 2016 Pearson Education, Ltd.

LINEAR COMBINATIONS

- A vector equation

$$x_1\mathbf{a}_1 + x_2\mathbf{a}_2 + \ldots + x_n\mathbf{a}_n = \mathbf{b}$$

has the same solution set as the linear system whose augmented matrix is

$$\begin{bmatrix} \mathbf{a}_1 & \mathbf{a}_2 & \cdots & \mathbf{a}_n & \mathbf{b} \end{bmatrix} \qquad (5)$$

- In particular, **b** can be generated by a linear combination of $\mathbf{a}_1, \ldots, \mathbf{a}_n$ if and only if there exists a solution to the linear system corresponding to the matrix (5).

© 2016 Pearson Education, Ltd.

LINEAR COMBINATIONS

- **Definition:** If $\mathbf{v}_1, \ldots, \mathbf{v}_p$ are in $\mathbb{R}^n$, then the set of all linear combinations of $\mathbf{v}_1, \ldots, \mathbf{v}_p$ is denoted by Span $\{\mathbf{v}_1, \ldots, \mathbf{v}_p\}$ and is called the **subset of $\mathbb{R}^n$ spanned** (or **generated**) **by** $\mathbf{v}_1, \ldots, \mathbf{v}_p$. That is, Span $\{\mathbf{v}_1, \ldots, \mathbf{v}_p\}$ is the collection of all vectors that can be written in the form

$$c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \ldots + c_p\mathbf{v}_p$$

with $c_1, \ldots, c_p$ scalars.

© 2016 Pearson Education, Ltd.

A GEOMETRIC DESCRIPTION OF SPAN {**V**}

- Let **v** be a nonzero vector in $\mathbb{R}^3$. Then Span {**v**} is the set of all scalar multiples of **v**, which is the set of points on the line in $\mathbb{R}^3$ through **v** and **0**. See Fig. 10 below:

© 2016 Pearson Education, Ltd.

A GEOMETRIC DESCRIPTION OF SPAN {**U, V**}

- If **u** and **v** are nonzero vectors in $\mathbb{R}^3$, with **v** not a multiple of **u**, then Span {**u, v**} is the plane in $\mathbb{R}^3$ that contains **u**, **v**, and **0**.
- In particular, Span {**u**, **v**} contains the line in $\mathbb{R}^3$ through **u** and **0** and the line through **v** and **0**. See Fig. 11 below:

© 2016 Pearson Education, Ltd.

1 Linear Equations in Linear Algebra

1.4

THE MATRIX EQUATION $A\mathbf{x} = \mathbf{b}$

PEARSON

© 2016 Pearson Education, Ltd.

MATRIX EQUATION $A\mathrm{x} = \mathrm{b}$

- **Definition:** If A is an $m \times n$ matrix, with columns $\mathbf{a}_1, \ldots, \mathbf{a}_n$, and if $\mathbf{x}$ is in $\mathbb{R}^n$, then the **product of** A **and** $\mathbf{x}$, denoted by $A\mathbf{x}$, **is the linear combination of the columns of** A **using the corresponding entries in** $\mathbf{x}$ **as weights**; that is,

$$A\mathrm{x} = \begin{bmatrix} \mathrm{a}_1 & \mathrm{a}_2 & \cdots & \mathrm{a}_n \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = x_1\mathrm{a}_1 + x_2\mathrm{a}_2 + \ldots + x_n\mathrm{a}_n.$$

- Note that $A\mathbf{x}$ is defined only if the number of columns of A equals the number of entries in $\mathbf{x}$.

© 2016 Pearson Education, Ltd.

MATRIX EQUATION $A\mathrm{x} = \mathrm{b}$

- **Example 2:** For $\mathbf{v}_1$, $\mathbf{v}_2$, $\mathbf{v}_3$ in $\mathbb{R}^m$, write the linear combination $3\mathrm{v}_1 - 5\mathrm{v}_2 + 7\mathrm{v}_3$ as a matrix times a vector.
- **Solution:** Place $\mathbf{v}_1$, $\mathbf{v}_2$, $\mathbf{v}_3$ into the columns of a matrix A and place the weights $3, -5$, and 7 into a vector $\mathbf{x}$.
- That is,

$$3\mathrm{v}_1 - 5\mathrm{v}_2 + 7\mathrm{v}_3 = \begin{bmatrix} \mathrm{v}_1 & \mathrm{v}_2 & \mathrm{v}_3 \end{bmatrix} \begin{bmatrix} 3 \\ -5 \\ 7 \end{bmatrix} = A\mathrm{x}.$$

© 2016 Pearson Education, Ltd.

MATRIX EQUATION $A\mathbf{x} = \mathbf{b}$

- Now, write the system of linear equations as a vector equation involving a linear combination of vectors.
- For example, the following system

$$\begin{aligned} x_1 + 2x_2 - x_3 &= 4 \\ -5x_2 + 3x_3 &= 1 \end{aligned} \qquad (1)$$

is equivalent to

$$x_1\begin{bmatrix} 1 \\ 0 \end{bmatrix} + x_2\begin{bmatrix} 2 \\ -5 \end{bmatrix} + x_3\begin{bmatrix} -1 \\ 3 \end{bmatrix} = \begin{bmatrix} 4 \\ 1 \end{bmatrix}. \qquad (2)$$

© 2016 Pearson Education, Ltd.

MATRIX EQUATION $A\mathbf{x} = \mathbf{b}$

- As in the example, the linear combination on the left side is a matrix times a vector, so that (2) becomes

$$\begin{bmatrix} 1 & 2 & -1 \\ 0 & -5 & 3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 4 \\ 1 \end{bmatrix}. \qquad (3)$$

- Equation (3) has the form $A\mathbf{x} = \mathbf{b}$. Such an equation is called a **matrix equation**, to distinguish it from a vector equation such as shown in (2).

© 2016 Pearson Education, Ltd.

MATRIX EQUATION $A\mathrm{x} = \mathrm{b}$

THEOREM 3

If A is an $m \times n$ matrix, with columns $\mathrm{a}_1, \ldots, \mathrm{a}_n$, and if b is in $\mathbb{R}^n$, then the matrix equation

$$\mathrm{A}\mathbf{x} = \mathbf{b}$$

has the same solution set as the vector equation

$$x_1 a_1 + x_1 a_2 + \cdots + x_n a_n = b$$

which, in turn, has the same solution set as the system of linear equations whose augmented matrix is

$$[a_1 \quad a_2 \ \ldots \ a_n \quad \mathrm{b}]$$

© 2016 Pearson Education, Ltd.

EXISTENCE OF SOLUTIONS

- The equation $A\mathrm{x} = \mathrm{b}$ has a solution if and only if **b** is a linear combination of the columns of A.

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.

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

© 2016 Pearson Education, Ltd.

COMPUTATION OF $A\mathbf{x}$

- **Example 4:** Compute $A\mathbf{x}$, where $A = \begin{bmatrix} 2 & 3 & 4 \\ -1 & 5 & -3 \\ 6 & -2 & 8 \end{bmatrix}$ and $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$.
- **Solution:** From the definition,

$$\begin{bmatrix} 2 & 3 & 4 \\ -1 & 5 & -3 \\ 6 & -2 & 8 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = x_1\begin{bmatrix} 2 \\ -1 \\ 6 \end{bmatrix} + x_2\begin{bmatrix} 3 \\ 5 \\ -2 \end{bmatrix} + x_3\begin{bmatrix} 4 \\ -3 \\ 8 \end{bmatrix}$$

© 2016 Pearson Education, Ltd.

COMPUTATION OF $A\mathbf{x}$

$$= \begin{bmatrix} 2x_1 \\ -x_1 \\ 6x_1 \end{bmatrix} + \begin{bmatrix} 3x_2 \\ 5x_2 \\ -2x_2 \end{bmatrix} + \begin{bmatrix} 4x_3 \\ -3x_3 \\ 8x_3 \end{bmatrix} \quad (1)$$

$$= \begin{bmatrix} 2x_1 + 3x_2 + 4x_3 \\ -x_1 + 5x_2 - 3x_3 \\ 6x_1 - 2x_2 + 8x_3 \end{bmatrix}.$$

- The first entry in the product $A\mathbf{x}$ is a sum of products (sometimes called *a dot product*), using the first row of A and the entries in $\mathbf{x}$.

© 2016 Pearson Education, Ltd.

COMPUTATION OF $A\mathbf{x}$

- That is, $\begin{bmatrix} 2 & 3 & 4 \\ & & \\ & & \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 2x_1 + 3x_2 + 4x_3 \\ \\ \\ \end{bmatrix}.$

- Similarly, the second entry in $A\mathbf{x}$ can be calculated by multiplying the entries in the second row of A by the corresponding entries in $\mathbf{x}$ and then summing the resulting products.

$$\begin{bmatrix} & & \\ -1 & 5 & -3 \\ & & \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} \\ -x_1 + 5x_2 - 3x_3 \\ \\ \end{bmatrix}$$

© 2016 Pearson Education, Ltd.

ROW-VECTOR RULE FOR COMPUTING $A\mathbf{x}$

- Likewise, the third entry in $A\mathbf{x}$ can be calculated from the third row of A and the entries in $\mathbf{x}$.
- If the product $A\mathbf{x}$ is defined, then the ith entry in $A\mathbf{x}$ is the sum of the products of corresponding entries from row i of A and from the vector $\mathbf{x}$.
- The matrix with 1's on the diagonal and 0's elsewhere is called an **identity matrix** and is denoted by I.
- For example, $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$ is an identity matrix.

© 2016 Pearson Education, Ltd.

PROPERTIES OF THE MATRIX-VECTOR PRODUCT $A\mathbf{x}$

THEOREM 5

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

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

b. $A(c\mathbf{u}) = c(A\mathbf{u})$.

- **Proof:** For simplicity, take $n = 3$, $A = [\mathbf{a}_1 \ \ \mathbf{a}_2 \ \ \mathbf{a}_3]$, and $\mathbf{u}$, $\mathbf{v}$ in $\mathbb{R}^3$.
- For $i = 1, 2, 3$, let u_i and v_i be the ith entries in $\mathbf{u}$ and $\mathbf{v}$, respectively.

© 2016 Pearson Education, Ltd.

PROPERTIES OF THE MATRIX-VECTOR PRODUCT $A\mathbf{x}$

- To prove statement (a), compute $A(\mathbf{u}+\mathbf{v})$ as a linear combination of the columns of A using the entries in $\mathbf{u}+\mathbf{v}$ as weights.

$$A(\mathbf{u}+\mathbf{v}) = \begin{bmatrix} \mathbf{a}_1 & \mathbf{a}_2 & \mathbf{a}_3 \end{bmatrix} \begin{bmatrix} u_1+v_1 \\ u_2+v_2 \\ u_3+v_3 \end{bmatrix}$$

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

$$= (u_1+v_1)\mathbf{a}_1 + (u_2+v_2)\mathbf{a}_2 + (u_3+v_3)\mathbf{a}_3$$

Columns of A

$$= (u_1\mathbf{a}_1 + u_2\mathbf{a}_2 + u_3\mathbf{a}_3) + (v_1\mathbf{a}_1 + v_2\mathbf{a}_2 + v_3\mathbf{a}_3)$$

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

© 2016 Pearson Education, Ltd.

PROPERTIES OF THE MATRIX-VECTOR PRODUCT $A\mathbf{x}$

- To prove statement (b), compute $A(c\mathrm{u})$ as a linear combination of the columns of A using the entries in $c\mathbf{u}$ as weights.

$$A(c\mathrm{u}) = \begin{bmatrix} \mathrm{a}_1 & \mathrm{a}_2 & \mathrm{a}_3 \end{bmatrix} \begin{bmatrix} cu_1 \\ cu_2 \\ cu_3 \end{bmatrix} = (cu_1)\mathrm{a}_1 + (cu_2)\mathrm{a}_2 + (cu_3)\mathrm{a}_3$$

$$= c(u_1\mathrm{a}_1) + c(u_2\mathrm{a}_2) + c(u_3\mathrm{a}_3)$$

$$= c(u_1\mathrm{a}_1 + u_2\mathrm{a}_2 + u_3\mathrm{a}_3)$$

$$= c(A\mathrm{u})$$

© 2016 Pearson Education, Ltd.

Proof of Theorem 4

- (1) statements (a), (b), and (c) are logically equivalent from the definition of Ax and span.
- (2) (a) and (d) are equivalent:
  - Let U be an echelon form of A.
  - Given b in R^m, [A b]~[U d] for some d in R^m.
  - If (d) is true:
    - Each row of U contains a pivot position
    - Ax=b has a solution => (a) is true
  - If (d) is false:
    - The last row of U is all zeros
    - [U d] represents an inconsistent system
    - Since row operations are reversible, [U d]~[A b]
    - Ax=b is also inconsistent => (a) is false

© 2012 Pearson Education, Ltd.

1 Linear Equations in Linear Algebra

1.5

SOLUTION SETS OF LINEAR SYSTEMS

PEARSON

© 2016 Pearson Education, Inc.

HOMOGENEOUS LINEAR SYSTEMS

- A system of linear equations is said to be **homogeneous** if it can be written in the form $A\mathrm{x} = 0$, where A is an $m \times n$ matrix and 0 is the zero vector in $\mathbb{R}^m$.
- Such a system $A\mathrm{x} = 0$ *always* has at least one solution, namely, $\mathrm{x} = 0$ (the zero vector in $\mathbb{R}^n$).
- This zero solution is usually called the **trivial solution**.
- The homogenous equation $A\mathbf{x} = 0$, the important question is whether there exists a **nontrivial solution**, that is, a nonzero vector x that satisfies $A\mathbf{x} = 0$.

© 2016 Pearson Education, Inc.

HOMOGENEOUS LINEAR SYSTEMS

- **Example 1:** Determine if the following homogeneous system has a nontrivial solution. Then describe the solution set.

$$\begin{aligned} 3x_1 + 5x_2 - 4x_3 &= 0 \\ -3x_1 - 2x_2 + 4x_3 &= 0 \\ 6x_1 + x_2 - 8x_3 &= 0 \end{aligned}$$

- **Solution:** Let A be the matrix of coefficients of the system and row reduce the augmented matrix $\begin{bmatrix} A & 0 \end{bmatrix}$ to echelon form:

© 2016 Pearson Education, Inc.

HOMOGENEOUS LINEAR SYSTEMS

$$\begin{bmatrix} 3 & 5 & -4 & 0 \\ -3 & -2 & 4 & 0 \\ 6 & 1 & -8 & 0 \end{bmatrix} \sim \begin{bmatrix} 3 & 5 & -4 & 0 \\ 0 & 3 & 0 & 0 \\ 0 & -9 & 0 & 0 \end{bmatrix} \sim \begin{bmatrix} 3 & 5 & -4 & 0 \\ 0 & 3 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

- Since x_3 is a free variable, $A\mathbf{x} = 0$ has nontrivial solutions (one for each choice of x_3.)
- Continue the row reduction of $\begin{bmatrix} A & 0 \end{bmatrix}$ to *reduced* echelon form:

$$\begin{bmatrix} 1 & 0 & -\frac{4}{3} & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \qquad \begin{aligned} x_1 - \frac{4}{3}x_3 &= 0 \\ x_2 &= 0 \\ 0 &= 0 \end{aligned}$$

© 2016 Pearson Education, Inc.

HOMOGENEOUS LINEAR SYSTEMS

- Solve for the basic variables x_1 and x_2 to obtain $x_1 = \frac{4}{3}x_3$, $x_2 = 0$, with x_3 free.
- As a vector, the general solution of $A\mathrm{x} = 0$ has the form given below.

$$\mathrm{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} \frac{4}{3}x_3 \\ 0 \\ x_3 \end{bmatrix} = x_3 \begin{bmatrix} \frac{4}{3} \\ 0 \\ 1 \end{bmatrix} = x_3\mathrm{v}, \text{ where } \mathrm{v} = \begin{bmatrix} \frac{4}{3} \\ 0 \\ 1 \end{bmatrix}$$

© 2016 Pearson Education, Inc.

HOMOGENEOUS LINEAR SYSTEMS

- Here x_3 is factored out of the expression for the general solution vector.
- This shows that every solution of $A\mathrm{x} = 0$ in this case is a scalar multiple of **v**.
- The trivial solution is obtained by choosing $x_3 = 0$.
- Geometrically, the solution set is a line through 0 in $\mathbb{R}^3$. See Fig.1 below.

© 2016 Pearson Education, Inc.

PARAMETRIC VECTOR FORM

- The equation of the form $\mathbf{x} = s\mathbf{u} + t\mathbf{v}$ (s, t in $\mathbb{R}$) is called a **parametric vector equation** of the plane.

- In Example 1, the equation $\mathbf{x} = x_3\mathbf{v}$ (with x_3 free), or $\mathbf{x} = t\mathbf{v}$ (with t in $\mathbb{R}$), is a parametric vector equation of a line.

- Whenever a solution set is described explicitly with vectors as in Example 1, we say that the solution is in **parametric vector form**.

© 2016 Pearson Education, Inc.

SOLUTIONS OF NONHOMOGENEOUS SYSTEMS

- When a nonhomogeneous linear system has many solutions, the general solution can be written in parametric vector form as one vector plus an arbitrary linear combination of vectors that satisfy the corresponding homogeneous system.
- **Example 3 :** Describe all solutions of $A\mathbf{x} = \mathbf{b}$, where

$$A = \begin{bmatrix} 3 & 5 & -4 \\ -3 & -2 & 4 \\ 6 & 1 & -8 \end{bmatrix} \text{ and } \mathbf{b} = \begin{bmatrix} 7 \\ -1 \\ -4 \end{bmatrix}.$$

© 2016 Pearson Education, Inc.

SOLUTIONS OF NONHOMOGENEOUS SYSTEMS

- **Solution:** Row operations on $\begin{bmatrix} A & 0 \end{bmatrix}$ produce

$$\begin{bmatrix} 3 & 5 & -4 & 7 \\ -3 & -2 & 4 & -1 \\ 6 & 1 & -8 & -4 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & -\frac{4}{3} & -1 \\ 0 & 1 & 0 & 2 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \quad \begin{aligned} x_1 - \frac{4}{3}x_3 &= -1 \\ x_2 &= 2 \\ 0 &= 0 \end{aligned}$$

- Thus $x_1 = -1 + \frac{4}{3}x_3$, $x_2 = 2$, and x_3 is free.

© 2016 Pearson Education, Inc.

SOLUTIONS OF NONHOMOGENEOUS SYSTEMS

- As a vector, the general solution of $A\mathrm{x} = \mathrm{b}$ has the form

$$\mathrm{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} -1 + \frac{4}{3}x_3 \\ 2 \\ x_3 \end{bmatrix} = \begin{bmatrix} -1 \\ 2 \\ 0 \end{bmatrix} = \begin{bmatrix} \frac{4}{3}x_3 \\ 0 \\ x_3 \end{bmatrix} = \underbrace{\begin{bmatrix} -1 \\ 2 \\ 0 \end{bmatrix}}_{\mathrm{p}} + x_3 \underbrace{\begin{bmatrix} \frac{4}{3} \\ 0 \\ 1 \end{bmatrix}}_{\mathrm{v}}$$

© 2016 Pearson Education, Inc.

SOLUTIONS OF NONHOMOGENEOUS SYSTEMS

- The equation $\mathrm{x} = \mathrm{p} + x_3\mathrm{v}$, or, writing t as a general parameter,

$$\mathrm{x} = \mathrm{p} + t\mathrm{v} \quad (t \text{ in } \mathbb{R}) \qquad (3)$$

  describes the solution set of $A\mathrm{x} = \mathrm{b}$ in parametric vector form.

- The solution set of $A\mathrm{x} = 0$ has the parametric vector equation

$$\mathrm{x} = t\mathrm{v} \quad (t \text{ in } \mathbb{R}) \qquad (4)$$

  [with the same **v** that appears in (3)].

- Thus the solutions of $A\mathrm{x} = \mathrm{b}$ are obtained by adding the vector **p** to the solutions of $A\mathrm{x} = 0$.

© 2016 Pearson Education, Inc.

SOLUTIONS OF NONHOMOGENEOUS SYSTEMS

- The vector **p** itself is just one particular solution of $A\mathrm{x} = \mathrm{b}$ [corresponding to $t = 0$ in (3).]

- To describe the solution of $A\mathrm{x} = \mathrm{b}$ geometrically, we can think of vector addition as a *translation*.

- Given **v** and **p** in $\mathbb{R}^2$ or $\mathbb{R}^3$, the effect of adding **p** to **v** is to *move* **v** in a direction parallel to the line through **p** and **0**.

© 2016 Pearson Education, Inc.

SOLUTIONS OF NONHOMOGENEOUS SYSTEMS

- We say that **v** is **translated by p** to $\mathbf{v}+\mathbf{p}$. See the following figure.

- If each point on a line L in $\mathbb{R}^2$ or $\mathbb{R}^3$ is translated by a vector **p**, the result is a line parallel to L. See the following figure.

© 2016 Pearson Education, Inc.

SOLUTIONS OF NONHOMOGENEOUS SYSTEMS

- Suppose L is the line through $\mathbf{0}$ and $\mathbf{v}$, described by equation (4).
- Adding $\mathbf{p}$ to each point on L produces the translated line described by equation (3).
- We call (3) **the equation of the line through p parallel to v**.
- Thus the solution set of $A\mathrm{x} = \mathrm{b}$ *is a line through* $\mathbf{p}$ *parallel to the solution set of* $A\mathrm{x} = 0$. The figure on the next slide illustrates this case.

© 2016 Pearson Education, Inc.

SOLUTIONS OF NONHOMOGENEOUS SYSTEMS

- The relation between the solution sets of $A\mathrm{x} = \mathrm{b}$ and $A\mathrm{x} = 0$ shown in the figure above generalizes to any consistent equation $A\mathrm{x} = \mathrm{b}$, although the solution set will be larger than a line when there are several free variables.

© 2016 Pearson Education, Inc.

SOLUTIONS OF NONHOMOGENEOUS SYSTEMS

THEOREM 6

Suppose the equation $Ax = \mathrm{b}$ is consistent for some given **b**, and let **p** be a solution. Then the solution set of $Ax = \mathrm{b}$ is the set of all vectors of the form $w = p + v_h$, where $\mathbf{v}_h$ is any solution of the homogeneous equation.

- Theorem 6 says that if $Ax = \mathrm{b}$ has a solution, then the solution set is obtained by translating the solution set of $Ax = 0$, using any particular solution **p** of Ax = b for the translation.

© 2016 Pearson Education, Inc.

WRITING A SOLUTION SET (OF A CONSISTENT SYSTEM) IN PARAMETRIC VECTOR FORM

1. Row reduce the augmented matrix to reduced echelon form.
2. Express each basic variable in terms of any free variables appearing in an equation.
3. Write a typical solution **x** as a vector whose entries depend on the free variables, if any.
4. Decompose **x** into a linear combination of vectors (with numeric entries) using the free variables as parameters.

© 2016 Pearson Education, Inc.

1 Linear Equations in Linear Algebra

1.7

LINEAR INDEPENDENCE

PEARSON

© 2016 Pearson Education, Inc.

LINEAR INDEPENDENCE

- **Definition:** A set of vectors $\{\mathbf{v}_1, \ldots, \mathbf{v}_p\}$ is said to be ***linearly independent*** if the vector equation

$$x_1\mathbf{v}_1 + x_2\mathbf{v}_2 + \ldots + x_p\mathbf{v}_p = 0$$

has only the trivial solution. The set $\{\mathbf{v}_1, \ldots, \mathbf{v}_p\}$ is said to be ***linearly dependent*** if there are scalars $c_1, \ldots, c_p$, that are not all zero, such that

$$c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \ldots + c_p\mathbf{v}_p = 0$$

© 2016 Pearson Education, Inc.

LINEAR INDEPENDENCE

$$c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \ldots + c_p\mathbf{v}_p = 0$$

- The above equation is a **linear dependence relation** among $\mathbf{v}_1, \ldots, \mathbf{v}_p$ when the scalars are not all zero.
- An indexed set is linearly dependent if and only if it is not linearly independent.

- **Example 1:** Let $\mathbf{v}_1 = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}, \mathbf{v}_2 = \begin{bmatrix} 4 \\ 5 \\ 6 \end{bmatrix}$, and $\mathbf{v}_3 = \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix}$.

© 2016 Pearson Education, Inc.

LINEAR INDEPENDENCE

a. Determine if the set $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ is linearly independent.

b. If possible, find a linear dependence relation among $\mathbf{v}_1$, $\mathbf{v}_2$, and $\mathbf{v}_3$.

- **Solution:** We must determine if there is a nontrivial solution of the equation (here *p=3* ☺)

$$x_1\mathrm{v}_1 + x_2\mathrm{v}_2 + \ldots + x_p\mathrm{v}_p = 0$$

© 2016 Pearson Education, Inc.

LINEAR INDEPENDENCE

- Row operations on the associated augmented matrix show that

$$\begin{bmatrix} 1 & 4 & 2 & 0 \\ 2 & 5 & 1 & 0 \\ 3 & 6 & 0 & 0 \end{bmatrix} \sim \begin{bmatrix} 1 & 4 & 2 & 0 \\ 0 & -3 & -3 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

- x_1 and x_2 are basic variables, and x_3 is free.
- Each nonzero value of x_3 determines a nontrivial solution of $x_1\mathbf{v}_1 + x_2\mathbf{v}_2 + \ldots + x_p\mathbf{v}_p = 0$
- Hence, $\mathbf{v}_1$, $\mathbf{v}_2$, $\mathbf{v}_3$ are linearly **dependent**.

© 2016 Pearson Education, Inc.

LINEAR INDEPENDENCE

b. To find a linear dependence relation among $\mathbf{v}_1$, $\mathbf{v}_2$, and $\mathbf{v}_3$, completely row reduce the augmented matrix and write the new system:

$$\begin{bmatrix} 1 & 0 & -2 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \qquad \begin{aligned} x_1 - 2x_3 &= 0 \\ x_2 + x_3 &= 0 \\ 0 &= 0 \end{aligned}$$

- Thus, $x_1 = 2x_3$, $x_2 = -x_3$, and x_3 is free.
- Choose any nonzero value for x_3—say, $x_3 = 5$.
- Then $x_1 = 10$ and $x_2 = -5$.

© 2016 Pearson Education, Inc.

LINEAR INDEPENDENCE

- Substitute these values into equation (1) and obtain the equation below.

$$10\mathrm{v}_1 - 5\mathrm{v}_2 + 5\mathrm{v}_3 = 0$$

- This is one (out of infinitely many) possible linear dependence relations among $\mathbf{v}_1$, $\mathbf{v}_2$, and $\mathbf{v}_3$.

© 2016 Pearson Education, Inc.

LINEAR INDEPENDENCE OF MATRIX COLUMNS

- Suppose we begin with a matrix $A = \begin{bmatrix} \mathrm{a}_1 & \cdots & \mathrm{a}_n \end{bmatrix}$ instead of a set of vectors.

- The matrix equation $A\mathrm{x} = 0$ can be written as

$$x_1\mathrm{a}_1 + x_2\mathrm{a}_2 + \ldots + x_n\mathrm{a}_n = 0.$$

- *Each linear dependence relation among the columns of A corresponds to a nontrivial solution of* $A\mathrm{x} = 0$

- The columns of matrix A are linearly independent if and only if the equation $A\mathrm{x} = 0$ has *only* the trivial solution.

© 2016 Pearson Education, Inc.

SETS OF ONE OR TWO VECTORS

- Any set containing the zero vector is linearly dependent because…for any x_1, $x_1 0 = 0$

- A set containing only one vector – say, **v** – is linearly independent if and only if **v** is not the zero vector.

- This is because the vector equation $x_1 \mathrm{v} = 0$ has only the trivial solution when $\mathrm{v} \neq 0$.

© 2016 Pearson Education, Inc.

SETS OF ONE OR TWO VECTORS

- A set of two vectors $\{\mathbf{v}_1, \mathbf{v}_2\}$ is linearly dependent if at least one of the vectors is a multiple of the other.

- The set is linearly independent if and only if neither of the vectors is a multiple of the other.

© 2016 Pearson Education, Inc.

SETS OF TWO OR MORE VECTORS

THEOREM 7

Characterization of Linearly Dependent Sets

An indexed set $S = \{\mathrm{v}_1, \ldots, \mathrm{v}_p\}$ of two or more vectors is linearly dependent if and only if at least one of the vectors in S is a linear combination of the others. In fact, if S is linearly dependent and $\mathrm{v}_1 \neq 0$, then some $\mathbf{v}_j$ (with $j > 1$) is a linear combination of the preceding vectors, $\mathbf{v}_1, \ldots, \mathrm{v}_{j-1}$.

© 2016 Pearson Education, Inc.

SETS OF TWO OR MORE VECTORS

- **Proof:** If some $\mathbf{v}_j$ in S equals a linear combination of the other vectors, then $\mathbf{v}_j$ can be subtracted from both sides of the equation, producing a linear dependence relation with a nonzero weight (-1) on $\mathbf{v}_j$.
- [For instance, if $\mathbf{v}_1 = c_2\mathbf{v}_2 + c_3\mathbf{v}_3$, then $0 = (-1)\mathbf{v}_1 + c_2\mathbf{v}_2 + c_3\mathbf{v}_3 + 0\mathbf{v}_4 + \ldots + 0\mathbf{v}_p$.]
- Thus S is linearly dependent.
- Conversely, suppose S is linearly dependent.
- If $\mathbf{v}_1$ is zero, then it is a (trivial) linear combination of the other vectors in S.

© 2016 Pearson Education, Inc.

SETS OF TWO OR MORE VECTORS

- Otherwise, $\mathbf{v}_1 \neq 0$, and there exist weights $c_1, \ldots, c_p$, not all zero, such that

$$c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \ldots + c_p\mathbf{v}_p = 0.$$

- Let j be the largest subscript for which $c_j \neq 0$.

- If $j = 1$, then $c_1\mathbf{v}_1 = 0$, which is impossible because $\mathbf{v}_1 \neq 0$.

© 2016 Pearson Education, Inc.

SETS OF TWO OR MORE VECTORS

- So $j > 1$, and

$$c_1\mathbf{v}_1 + \ldots + c_j\mathbf{v}_j + 0\mathbf{v}_j + 0\mathbf{v}_{j+1} + \ldots + 0\mathbf{v}_p = 0$$

$$c_j\mathbf{v}_j = -c_1\mathbf{v}_1 - \ldots - c_{j-1}\mathbf{v}_{j-1}$$

$$\mathbf{v}_j = \left(-\frac{c_1}{c_j}\right)\mathbf{v}_1 + \ldots + \left(-\frac{c_{j-1}}{c_j}\right)\mathbf{v}_{j-1}.$$

© 2016 Pearson Education, Inc.

SETS OF TWO OR MORE VECTORS

- Theorem 7 does *not* say that *every* vector in a linearly dependent set is a linear combination of the preceding vectors.
- A vector in a linearly dependent set may fail to be a linear combination of the other vectors.
- **Example 4:** Let $u = \begin{bmatrix} 3 \\ 1 \\ 0 \end{bmatrix}$ and $v = \begin{bmatrix} 1 \\ 6 \\ 0 \end{bmatrix}$. Describe the set spanned by **u** and **v**, and explain why a vector **w** is in Span {**u**, **v**} if and only if {**u**, **v**, **w**} is linearly dependent.

© 2016 Pearson Education, Inc.

SETS OF TWO OR MORE VECTORS

- **Solution:** The vectors **u** and **v** are linearly independent because neither vector is a multiple of the other, and so they span a plane in $\mathbb{R}^3$.
- Span {**u**, **v**} is the x_1x_2-plane (with $x_3 = 0$).
- If **w** is a linear combination of **u** and **v**, then {**u**, **v**, **w**} is linearly dependent, by Theorem 7.
- Conversely, suppose that {**u**, **v**, **w**} is linearly dependent.
- By theorem 7, some vector in {**u**, **v**, **w**} is a linear combination of the preceding vectors (since $\mathrm{u} \neq 0$).
- That vector must be **w**, since **v** is not a multiple of **u**.

© 2016 Pearson Education, Inc.

SETS OF TWO OR MORE VECTORS

- So **w** is in Span {**u**, **v**}. Fig. 2 below

Linearly dependent,
w in Span{**u**, **v**}

Linearly independent,
w *not* in Span{**u**, **v**}

- Example 4 generalizes to any set {**u**, **v**, **w**} in $\mathbb{R}^3$ with **u** and **v** linearly independent.
- The set {**u**, **v**, **w**} will be linearly dependent if and only if **w** is in the plane spanned by **u** and **v**.

© 2016 Pearson Education, Inc.

SETS OF TWO OR MORE VECTORS

THEOREM 8

If a set contains more vectors than there are entries in each vector, then the set is linearly dependent. That is, any set $\{\mathbf{v}_1, \ldots, \mathbf{v}_p\}$ in $\mathbb{R}^n$ is linearly dependent if $p > n$.

- **Proof:** Let $A = \begin{bmatrix} \mathbf{v}_1 & \cdots & \mathbf{v}_p \end{bmatrix}$.
- Then A is $n \times p$, and the equation $A\mathbf{x} = 0$ corresponds to a system of n equations in p unknowns.
- If $p > n$, there are more variables than equations, so there must be a free variable.

© 2016 Pearson Education, Inc.

SETS OF TWO OR MORE VECTORS

- Hence $A\mathrm{x} = 0$ has a nontrivial solution, and the columns of A are linearly dependent.
- See the figure below for a matrix version of this theorem.

$$\begin{array}{cc} & p \\ n & \begin{bmatrix} * & * & * & * & * \\ * & * & * & * & * \\ * & * & * & * & * \end{bmatrix} \end{array}$$

If $p > n$, the columns are linearly dependent.

- Theorem 8 says nothing about the case in which the number of vectors in the set does *not* exceed the number of entries in each vector.

© 2016 Pearson Education, Inc.

SETS OF TWO OR MORE VECTORS

THEOREM 9

If a set $S = \{\mathbf{v}_1, \ldots, \mathbf{v}_p\}$ in $\mathbb{R}^n$ contains the zero vector, then the set is linearly dependent.

- **Proof:** By renumbering the vectors, we may suppose $\mathbf{v}_1 = 0$.

- Then the equation $1\mathbf{v}_1 + 0\mathbf{v}_2 + \ldots + 0\mathbf{v}_p = 0$ shows that S in linearly dependent.

© 2016 Pearson Education, Inc.

1 Linear Equations in Linear Algebra

1.8

INTRODUCTION TO LINEAR TRANSFORMATIONS

PEARSON

© 2016 Pearson Education, Inc.

LINEAR TRANSFORMATIONS

- A **transformation** (or **function** or **mapping**) T from $\mathbb{R}^n$ to $\mathbb{R}^m$ is a rule that assigns to each vector $\mathbf{x}$ in $\mathbb{R}^n$ a vector $T(\mathbf{x})$ in $\mathbb{R}^m$.
- The set $\mathbb{R}^n$ is called **domain** of T, and $\mathbb{R}^m$ is called the **codomain** of T.
- The notation $T: \mathbb{R}^n \to \mathbb{R}^m$ indicates that the domain of T is $\mathbb{R}^n$ and the codomain is $\mathbb{R}^m$.
- For $\mathbf{x}$ in $\mathbb{R}^n$, the vector $T(\mathbf{x})$ in $\mathbb{R}^m$ is called the **image** of $\mathbf{x}$ (under the action of T).
- The set of all images $T(\mathbf{x})$ is called the **range** of T. See Fig. 2 on the next slide

© 2016 Pearson Education, Inc.

MATRIX TRANSFORMATIONS

Domain, codomain, and range of $T: \mathbb{R}^n \to \mathbb{R}^m$.

- For each $\mathbf{x}$ in $\mathbb{R}^n$, $T(\mathbf{x})$ is computed as $A\mathbf{x}$, where A is an $m \times n$ matrix.
- For simplicity, we denote such a *matrix transformation* by $\mathrm{x} \mapsto A\mathrm{x}$.
- Observe that the domain of T is $\mathbb{R}^n$ when A has n columns and the codomain of T is $\mathbb{R}^m$ when each column of A has m entries.

© 2016 Pearson Education, Inc.

MATRIX TRANSFORMATIONS

- The range of T is the set of all linear combinations of the columns of A, because each image $T(\mathbf{x})$ is of the form $A\mathbf{x}$.
- **Example 1:**

Let $A = \begin{bmatrix} 1 & -3 \\ 3 & 5 \\ -1 & 7 \end{bmatrix}$, $\mathrm{u} = \begin{bmatrix} 2 \\ -1 \end{bmatrix}$, $b = \begin{bmatrix} 3 \\ 2 \\ -5 \end{bmatrix}$, $\mathrm{c} = \begin{bmatrix} 3 \\ 2 \\ 5 \end{bmatrix}$.

and define a transformation $T: \mathbb{R}^n \to \mathbb{R}^m$ by $T(\mathrm{x}) = A\mathrm{x}$, so that

$$T(\mathrm{x}) = A\mathrm{x} = \begin{bmatrix} 1 & -3 \\ 3 & 5 \\ -1 & 7 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} x_1 - 3x_2 \\ 3x_1 + 5x_2 \\ -x_1 + 7x_2 \end{bmatrix}.$$

© 2016 Pearson Education, Inc.

MATRIX TRANSFORMATIONS

a. Find $T(\mathbf{u})$, the image of $\mathbf{u}$ under the transformation T.

b. Find an $\mathbf{x}$ in $\mathbb{R}^2$ whose image under T is $\mathbf{b}$.

c. Is there more than one $\mathbf{x}$ whose image under T is $\mathbf{b}$?

d. Determine if $\mathbf{c}$ is in the range of the transformation T.

© 2016 Pearson Education, Inc.

MATRIX TRANSFORMATIONS

Solution:

a. Compute

$$T(\mathrm{u}) = A\mathrm{u} = \begin{bmatrix} 1 & -3 \\ 3 & 5 \\ -1 & 7 \end{bmatrix} \begin{bmatrix} 2 \\ -1 \end{bmatrix} = \begin{bmatrix} 5 \\ 1 \\ -9 \end{bmatrix}.$$

b. Solve $T(\mathrm{x}) = \mathrm{b}$ for **x**. That is, solve $A\mathrm{x} = \mathrm{b}$, or

$$\begin{bmatrix} 1 & -3 \\ 3 & 5 \\ -1 & 7 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 3 \\ 2 \\ -5 \end{bmatrix} \qquad (1)$$

© 2016 Pearson Education, Inc.

MATRIX TRANSFORMATIONS

- Row reduce the augmented matrix:

$$\begin{bmatrix} 1 & -3 & 3 \\ 3 & 5 & 2 \\ -1 & 7 & -5 \end{bmatrix} \sim \begin{bmatrix} 1 & -3 & 3 \\ 0 & 14 & -7 \\ 0 & 4 & -2 \end{bmatrix} \sim \begin{bmatrix} 1 & -3 & 3 \\ 0 & 1 & -.5 \\ 0 & 0 & 0 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & 1.5 \\ 0 & 1 & -.5 \\ 0 & 0 & 0 \end{bmatrix} \quad (2)$$

- Hence $x_1 = 1.5,\ x_2 = -.5,$ and $\mathrm{x} = \begin{bmatrix} 1.5 \\ -.5 \end{bmatrix}$.

- The image of this **x** under T is the given vector **b**.

© 2016 Pearson Education, Inc.

MATRIX TRANSFORMATIONS

c. Any **x** whose image under T is **b** must satisfy equation (1).
   - From (2), it is clear that equation (1) has a unique solution.
   - So there is exactly one **x** whose image is **b**.

d. The vector **c** is in the range of T if **c** is the image of some **x** in $\mathbb{R}^2$, that is, if $\mathrm{c} = T(\mathrm{x})$ for some **x**.
   - This is another way of asking if the system $A\mathrm{x} = \mathrm{c}$ is consistent.

© 2016 Pearson Education, Inc.

MATRIX TRANSFORMATIONS

- To find the answer, row reduce the augmented matrix:

$$\begin{bmatrix} 1 & -3 & 3 \\ 3 & 5 & 2 \\ -1 & 7 & 5 \end{bmatrix} \sim \begin{bmatrix} 1 & -3 & 3 \\ 0 & 14 & -7 \\ 0 & 4 & 8 \end{bmatrix} \sim \begin{bmatrix} 1 & -3 & 3 \\ 0 & 1 & 2 \\ 0 & 14 & -7 \end{bmatrix} \sim \begin{bmatrix} 1 & -3 & 3 \\ 0 & 1 & 2 \\ 0 & 0 & -35 \end{bmatrix}$$

- The third equation, $0 = -35$, shows that the system is inconsistent.
- So **c** is *not* in the range of T.

© 2016 Pearson Education, Inc.

SHEAR TRANSFORMATION

- **Example 3:** Let $A = \begin{bmatrix} 1 & 3 \\ 0 & 1 \end{bmatrix}$. The transformation $T: \mathbb{R}^2 \to \mathbb{R}^2$ defined by $T(\mathrm{x}) = A\mathrm{x}$ is called a **shear transformation**.

- It can be shown that if T acts on each point in the 2×2 square shown in Fig. 4 on the next slide, then the set of images forms the shaded parallelogram.

© 2016 Pearson Education, Inc.

SHEAR TRANSFORMATION

- The key idea is to show that T maps line segments onto line segments and then to check that the corners of the square map onto the vertices of the parallelogram.
- For instance, the image of the point $\mathbf{u} = \begin{bmatrix} 0 \\ 2 \end{bmatrix}$ is

$$T(\mathbf{u}) = \begin{bmatrix} 1 & 3 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 0 \\ 2 \end{bmatrix} = \begin{bmatrix} 6 \\ 2 \end{bmatrix},$$

© 2016 Pearson Education, Inc.

LINEAR TRANSFORMATIONS

and the image of $\begin{bmatrix} 2 \\ 2 \end{bmatrix}$ is $\begin{bmatrix} 1 & 3 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} 2 \\ 2 \end{bmatrix} = \begin{bmatrix} 8 \\ 2 \end{bmatrix}$.

- T deforms the square as if the top of the square were pushed to the right while the base is held fixed.
- **Definition:** A transformation (or mapping) T is **linear** if:
  i. $T(\mathrm{u}+\mathrm{v}) = T(\mathrm{u}) + T(\mathrm{v})$ for all **u**, **v** in the domain of T;
  ii. $T(c\mathrm{u}) = cT(\mathrm{u})$ for all scalars c and all **u** in the domain of T.

© 2016 Pearson Education, Inc.

LINEAR TRANSFORMATIONS

- Linear transformations *preserve the operations of vector addition and scalar multiplication*.
- Property (i) says that the result $T(\mathrm{u}+\mathrm{v})$ of first adding $\mathbf{u}$ and $\mathbf{v}$ in $\mathbb{R}^n$ and then applying T is the same as first applying T to $\mathbf{u}$ and $\mathbf{v}$ and then adding $T(\mathbf{u})$ and $T(\mathbf{v})$ in $\mathbb{R}^m$.
- These two properties lead to the following useful facts.
- If T is a linear transformation, then

$$T(0)=0 \tag{3}$$

© 2016 Pearson Education, Inc.

LINEAR TRANSFORMATIONS

and $T(c\mathrm{u}+d\mathrm{v})=cT(\mathrm{u})+dT(\mathrm{v}).$ (4)

for all vectors **u**, **v** in the domain of T and all scalars c, d.

- Property (3) follows from condition (ii) in the definition, because $T(0)=T(0\mathrm{u})=0T(\mathrm{u})=0.$
- Property (4) requires both (i) and (ii):

$$T(c\mathrm{u}+d\mathrm{v})=T(c\mathrm{u})+T(d\mathrm{v})=cT(\mathrm{u})+dT(\mathrm{v})$$

- *If a transformation satisfies* (4) *for all* **u**, **v** and c, d, *it must be linear.*
- (Set $c=d=1$ for preservation of addition, and set for $d=0$ preservation of scalar multiplication.)

© 2016 Pearson Education, Inc.

LINEAR TRANSFORMATIONS

- Repeated application of (4) produces a useful generalization:

$$T(c_1\mathrm{v}_1 + \ldots + c_p\mathrm{v}_p) = c_1T(\mathrm{v}_1) + \ldots + c_pT(\mathrm{v}_p) \quad (5)$$

- In engineering and physics, (5) is referred to as a *superposition principle*.

- Think of $\mathbf{v}_1, \ldots, \mathbf{v}_p$ as signals that go into a system and $T(\mathbf{v}_1), \ldots, T(\mathbf{v}_p)$ as the responses of that system to the signals.

© 2016 Pearson Education, Inc.

LINEAR TRANSFORMATIONS

- The system satisfies the superposition principle if whenever an input is expressed as a linear combination of such signals, the system's response is the *same* linear combination of the responses to the individual signals.

- Given a scalar r, define $T: \mathbb{R}^2 \to \mathbb{R}^2$ by $T(\mathrm{x}) = r\mathrm{x}$.

- T is called a **contraction** when $0 \le r \le 1$ and a **dilation** when $r > 1$.

© 2016 Pearson Education, Inc.

1 Linear Equations in Linear Algebra

1.9

THE MATRIX OF A LINEAR TRANSFORMATION

PEARSON

© 2016 Pearson Education, Inc.

THE MATRIX OF A LINEAR TRANSFORMATION

- **Theorem 10**: Let T: $\mathbb{R}^n \to \mathbb{R}^m$ be a linear transformation. Then there exists a unique matrix A such that

$$T(x) = Ax \text{ for all x in } \mathbb{R}^n$$

- In fact, A is the m $\times$ n matrix whose j^{th} column is the vector $T(e_j)$, where e_j is the jth column of the identity matrix in $\mathbb{R}^n$

$$A = [T(e_1)...T(e_1)] \qquad (3)$$

© 2016 Pearson Education, Inc.

THE MATRIX OF A LINEAR TRANSFORMATION

- **Proof**: Write x= $I_n x = [e_1 \ \ldots e_2]x = x_1 e_1 + \ \ldots + x_n e_n$, and use the linearity of T to compute

$$T(x) = T(x_1 e_1 + \cdots + x_n e_n) = x_1 T(e_1) + \cdots + x_n T(e_n)$$

$$= [T(e_1) \quad \ldots \quad T(e_1)] \begin{bmatrix} x_1 \\ . \\ . \\ . \\ x_n \end{bmatrix} = Ax$$

© 2016 Pearson Education, Inc.

THE MATRIX OF A LINEAR TRANSFORMATION

- The matrix A in (3) is called the **standard matrix for the linear transformation** T.
- We know now that every linear transformation from $\mathbb{R}^n$ to $\mathbb{R}^m$ can be viewed as a matrix transformation, and vice versa. The term *linear transformation* focuses on a property of a mapping, while *matrix transformation* describes how such a mapping is implemented, as the example on the next slide illustrates.

© 2016 Pearson Education, Inc.

THE MATRIX OF A LINEAR TRANSFORMATION

- **Example 2**: Find the standard matrix A for the dilation transformation T(x)= 3x, for x in $\mathbb{R}^2$.
- **Solution**: Write

$$T(e_1) = e_1 = \begin{bmatrix} 3 \\ 0 \end{bmatrix} \; and \; T(e_1) = e_1 = \begin{bmatrix} 3 \\ 0 \end{bmatrix}$$

$$\mathrm{A} = \begin{bmatrix} 3 & 0 \\ 0 & 3 \end{bmatrix}$$

© 2016 Pearson Education, Inc.

GEOMETRIC LINEAR TRANSFORMATIONS OF $\mathbb{R}^2$

- Tables 1-4 illustrate other common geometric linear transformations of the plane.

TABLE 1 Reflections

Transformation	Image of the Unit Square	Standard Matrix
Reflection through the x_1-axis		$\begin{bmatrix}1 & 0\\0 & -1\end{bmatrix}$

© 2016 Pearson Education, Inc.

EXISTENCE AND UNIQUENESS QUESTIONS

- Table 1 continued:

Reflection through the x_2-axis

$$\begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}$$

Reflection through the line $x_2 = x_1$

$$\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

© 2016 Pearson Education, Inc.

EXISTENCE AND UNIQUENESS QUESTIONS

- Table 1 continued:

Reflection through the line $x_2 = -x_1$		$\begin{bmatrix}0 & -1\\-1 & 0\end{bmatrix}$
Reflection through the origin		$\begin{bmatrix}-1 & 0\\0 & -1\end{bmatrix}$

© 2016 Pearson Education, Inc.

EXISTENCE AND UNIQUENESS QUESTIONS

TABLE 2 Contractions and Expansions

Transformation	Image of the Unit Square	Standard Matrix
Horizontal contraction and expansion		$\begin{bmatrix}k & 0\\0 & 1\end{bmatrix}$
Vertical contraction and expansion		$\begin{bmatrix}1 & 0\\0 & k\end{bmatrix}$

© 2016 Pearson Education, Inc.

EXISTENCE AND UNIQUENESS QUESTIONS

TABLE 3 Shears

Transformation	Image of the Unit Square	Standard Matrix
Horizontal shear		$\begin{bmatrix} 1 & k \\ 0 & 1 \end{bmatrix}$
Vertical shear		$\begin{bmatrix} 1 & 0 \\ k & 1 \end{bmatrix}$

© 2016 Pearson Education, Inc.

EXISTENCE AND UNIQUENESS QUESTIONS

TABLE 4 Projections

Transformation	Image of the Unit Square	Standard Matrix
Projection onto the x_1-axis		$\begin{bmatrix}1 & 0\\0 & 0\end{bmatrix}$
Projection onto the x_2-axis		$\begin{bmatrix}0 & 0\\0 & 1\end{bmatrix}$

© 2016 Pearson Education, Inc.

EXISTENCE AND UNIQUENESS QUESTIONS

- **Definition**: A mapping T: $\mathbb{R}^n \to \mathbb{R}^m$ is said to be **onto** $\mathbb{R}^n$ if each **b** in $\mathbb{R}^m$ is the image of *at least one* **x** in $\mathbb{R}^n$.

- Equivalently, T is onto $\mathbb{R}^m$ when the range of T is all of the codomain $\mathbb{R}^m$. That is, T maps $\mathbb{R}^n$ onto $\mathbb{R}^m$ if, for each **b** in the codomain $\mathbb{R}^m$, there exists at least one solution of T(x)= b . "Does T map $\mathbb{R}^n$ onto $\mathbb{R}^m$?" is an existence question. The mapping T is not onto when there is some **b** in $\mathbb{R}^m$ for which the equation T(x)=b has no solution. See the figure on the next slide.

© 2016 Pearson Education, Inc.

EXISTENCE AND UNIQUENESS QUESTIONS

FIGURE 3 Is the range of T all of $\mathbb{R}^m$?

- **Definition**: A mapping T: $\mathbb{R}^n \to \mathbb{R}^m$ is said to be **one-to-one** if each **b** in $\mathbb{R}^m$ is the image of *at most one* **x** in $\mathbb{R}^n$.

© 2016 Pearson Education, Inc.

EXISTENCE AND UNIQUENESS QUESTIONS

- **Example 4**: Let T be the linear transformation whose standard matrix is

$$A = \begin{bmatrix} 1 & -4 & 8 & 1 \\ 0 & 2 & -1 & 3 \\ 0 & 0 & 0 & 5 \end{bmatrix}$$

- Does T map $\mathbb{R}^4$ onto $\mathbb{R}^3$? Is T a one-to-one mapping?

© 2016 Pearson Education, Inc.

EXISTENCE AND UNIQUENESS QUESTIONS

- **Solution**: Since A happens to be in echelon form, we can see at once that A has a pivot position in each row. By Theorem 4 in Section 1.4, for each **b** in $\mathbb{R}^3$, the equation Ax=b is consistent. In other words, the linear transformation T maps $\mathbb{R}^4$(its domain) onto $\mathbb{R}^3$.
- However, since the equation Ax=b has a free variable (because there are four variables and only three basic variables), each **b** is the image of more than one **x**. This is, T is *not* one-to-one.

© 2016 Pearson Education, Inc.

EXISTENCE AND UNIQUENESS QUESTIONS

- **Theorem 11**: Let T: $\mathbb{R}^n \to \mathbb{R}^m$ be a linear transformation. Then T is one-to-one if and only if the equation T(x)=0 has only the trivial solution.
- **Proof**: Since T is linear, T(0) =0. If T is one-to-one, then the equation T(x)=0 has at most one solution and hence only the trivial solution.
- If T is not one-to-one, then there is a b that is the image of at least two different vectors in $\mathbb{R}^n$--say, **u** and **v**. That is T(u)=b and T(v)=b. But then, since T is linear,

$$T(u - v) = T(u) - T(v) = b - b = 0$$

© 2016 Pearson Education, Inc.

EXISTENCE AND UNIQUENESS QUESTIONS

- The vector **u – v** is not zero, since u ≠ v. Hence the equation T(x)=0 has more than one solution. So, either the two conditions in the theorem are both true or they are both false.
- **Theorem 12**: Let T: $\mathbb{R}^n \to \mathbb{R}^m$ be a linear transformation and let A be the standard matrix for T. Then:

a) T maps $\mathbb{R}^n$ onto $\mathbb{R}^m$ if and only if the columns of A span $\mathbb{R}^m$;

b) T is one-to-one if and only if the columns of A are linearly independent.

© 2016 Pearson Education, Inc.

EXISTENCE AND UNIQUENESS QUESTIONS

- **Proof:**

a) By Theorem 4 in Section 1.4, the columns of A span $\mathbb{R}^m$ if and only if for each b in $\mathbb{R}^m$ the equation Ax=b is consistent—in other words, if and only if for every **b**, the equation T(x)=b has at least one solution. This is true if and only if T maps $\mathbb{R}^n$ onto $\mathbb{R}^m$.

b) The equations T(x)=0 and Ax=0 are the same except for notation. So, by Theorem 11, T is one-to-one if and only if Ax=0 has only the trivial solution. This happens if and only if the columns of A are linearly independent.

© 2016 Pearson Education, Inc.

1 Linear Equations in Linear Algebra

1.10

LINEAR MODELS IN BUSINESS, SCIENCE, AND ENGINEERING

PEARSON

© 2016 Pearson Education, Ltd.

LINEAR EQUATIONS AND ELECTRICAL NETWORKS

- Current flow in a simple electrical network can be described by a system of linear equations.
- A voltage source such as a battery forces a current of electrons to flow through the network. When the current passes through a resistor some of the voltage is "used up"; by Ohm's law, this "voltage drop" across a resistor is given by

$$V = RI$$

- Where the voltage V is measured in *volts*, the resistance R in o*hms* (denoted by Ω), and the current flow I in amperes (*amps*, for short).

© 2016 Pearson Education, Ltd.

LINEAR EQUATIONS AND ELECTRICAL NETWORKS

- **Kirchhoff's Voltage Law:** The algebraic sum of RI voltage drops in one direction around a loop equals the algebraic sum of the voltage sources in the same direction around the loop.

- **Example 2** Determine the loop currents in the network in Fig.1
- **Solution** For loop 1, the current I1 flows through three resistors, and the sum of the RI voltage drops is

$$4I_1 + 4I_1 + 3I_1 = (4 + 4 + 3)I_1 = 11I_1$$

© 2016 Pearson Education, Ltd.

LINEAR EQUATIONS AND ELECTRICAL NETWORKS

FIGURE 1

© 2016 Pearson Education, Ltd.

LINEAR EQUATIONS AND ELECTRICAL NETWORKS

- Current from loop 2 also flows in part of loop 1, through the short *branch* between A and B. The associated RI drop there is $3I_2$ volts. However, the current direction for the branch AB in loop 1 is opposite to that chosen for the flow in loop 2, so the algebraic sum of all RI drops for loop 1 is $11I_1 - 3I_2$.
- Since the voltage in loop 1 is +30 volts, Kirchhoff's voltage law implies that

$$11I_1 - 3I_2 = 30$$

- The equation for loop 2 is

$$-3I_1 + 6I_2 - I_3 = 5$$

© 2016 Pearson Education, Ltd.

LINEAR EQUATIONS AND ELECTRICAL NETWORKS

- The term $-3I_1$ comes from the flow of the loop-1 current through the branch AB (with a negative voltage drop because the current flow there is opposite to the flow in loop 2).
- The term $6I_2$ is the sum of all resistances in loop 2, multiplied by the loop current.
- The term $-I_3 = -1 \cdot I_3$ comes from the loop-3 current flowing through the 1-ohm resistor in branch DC, in the direction opposite to the flow in loop 2. The loop-3 equation is

$$-I_2 + 3I_3 = -25$$

© 2016 Pearson Education, Ltd.

LINEAR EQUATIONS AND ELECTRICAL NETWORKS

- The loop currents are found by solving the system

$$\begin{aligned} 11I_1 - 3I_2 &= 30 \\ -3I_1 + 6I_2 - I_3 &= 5 \\ -I_2 + 3I_3 &= -25 \end{aligned} \qquad (3)$$

- Row operation on the augmented matrix lead to the solution: I_1=3 amps, I_2 =1 amp, and $I_3 = -8$ amps.

© 2016 Pearson Education, Ltd.

LINEAR EQUATIONS AND ELECTRICAL NETWORKS

- It is instructive to look at system (3) as a vector equation:

$$I_1 \underset{\mathbf{r}_1}{\underset{\uparrow}{\begin{bmatrix} 11 \\ -3 \\ 0 \end{bmatrix}}} + I_2 \underset{\mathbf{r}_2}{\underset{\uparrow}{\begin{bmatrix} -3 \\ 6 \\ -1 \end{bmatrix}}} + I_3 \underset{\mathbf{r}_3}{\underset{\uparrow}{\begin{bmatrix} 0 \\ -1 \\ 3 \end{bmatrix}}} = \underset{\mathbf{v}}{\underset{\uparrow}{\begin{bmatrix} 30 \\ 5 \\ -25 \end{bmatrix}}} \qquad (4)$$

- The first entry of each vector concerns the first loop, and similarly for the second and third entries. The first r_1 lists the resistance in the various loops through which current I_1 flows. A resistance is written negatively when I_1 flows against the flow direction in another loop.

© 2016 Pearson Education, Ltd.

LINEAR EQUATIONS AND ELECTRICAL NETWORKS

- Examine Fig. 1 and see how to compute the entries in r_1; then do the same for r_2 and r_3. The matrix form of equation (4),

$$R\mathbf{i} = \mathbf{v}, \quad \text{where} \quad R = [\,\mathbf{r}_1 \quad \mathbf{r}_2 \quad \mathbf{r}_3\,] \quad \text{and} \quad \mathbf{i} = \begin{bmatrix} I_1 \\ I_2 \\ I_3 \end{bmatrix}$$

Provides a matrix version of Ohm's law. If all loop currents are chosen in the same direction (say, counterclockwise), then all entries off the main diagonal of R will be negative.

© 2016 Pearson Education, Ltd.

LINEAR EQUATIONS AND ELECTRICAL NETWORKS

- The matrix equation $R\mathbf{i}=\mathbf{v}$ makes the linearity of this model easy to see at a glance. For instance, if the voltage vector is doubled, then the current vector must double.
- Also, a *superposition principle* holds. That is, the solution of equation (4) is the sum of the solutions of the equations

$$R\mathbf{i} = \begin{bmatrix} 30 \\ 0 \\ 0 \end{bmatrix}, \qquad R\mathbf{i} = \begin{bmatrix} 0 \\ 5 \\ 0 \end{bmatrix}, \quad \text{and} \quad R\mathbf{i} = \begin{bmatrix} 0 \\ 0 \\ -25 \end{bmatrix}$$

Each equation here corresponds to the circuit with only one voltage source (the other sources being replaced by wires that close each loop).

© 2016 Pearson Education, Ltd.

LINEAR EQUATIONS AND ELECTRICAL NETWORKS

- The model for current flow is *linear* precisely because Ohm's law and Kirchhoff's law are linear:
- The voltage drop across a resistor is *proportional* to the current flowing through it (Ohm), and the *sum* of the voltage drops in a loop equals the sum of the voltage sources in the loop (Kirchhoff).

© 2016 Pearson Education, Ltd.

LINEAR EQUATIONS AND ELECTRICAL NETWORKS

- Loop currents in a network can be used to determine the current in any branch of the network.
- If only one loop current passes through a branch, such as from B to D in Fig. 1, the branch current equals the loop current.
- If more than one loop current passes through a branch, such as from A to B, the branch current is the algebraic sum of the loop currents in the branch (Kirchhoff's current law).
- For instance, the current in branch AB is $I_1 - I_2$ =3-1=2 amps, in the direction of I_1. The current in branch CD is $I_2 - I_3 = 9$ amps.

© 2016 Pearson Education, Ltd.

DIFFERENCE EQUATIONS

- In many fields such as ecology, economics, and engineering, a need arises to model mathematically a dynamic system that changes over time.
- Several features of the system are each measured at discrete time intervals, producing a sequence of vectors $x_0, x_1, x_2,$ …..The entries in x_k provide information about the *state* of the system at the time of the *k*th measurement.
- If there is a matrix *A* such that $x_1 = Ax_0, x_2 = Ax_1,$ and, in general

$$x_{k+1} = Ax_k \quad \text{for } k = 0, 1, 2, \ldots$$

- then (5) is called a **linear difference equation** (or **recurrence relation**)

© 2016 Pearson Education, Ltd.

DIFFERENCE EQUATIONS

- Given such an equation, one can compute x_1, x_2, and so on, provided x_0 is known.
- The discussion below illustrates how a difference equation might arise.
- A subject of interest to demographers is the movement of populations or groups of people from one region to another. The simple model here considers the changes in the population of a certain city and its surrounding suburbs over a period of years.

© 2016 Pearson Education, Ltd.

DIFFERENCE EQUATIONS

- Fix an initial year—say, 2000—and denote the populations of the city and suburbs that year by r_0 and s_0, respectively. Let x_0 be the population vector

$$\mathbf{x}_0 = \begin{bmatrix} r_0 \\ s_0 \end{bmatrix} \quad \begin{matrix} \text{City population, 2000} \\ \text{Suburban population, 2000} \end{matrix}$$

- For 2001 and subsequent years, denote the populations of the city and suburbs by the vectors

$$\mathbf{x}_1 = \begin{bmatrix} r_1 \\ s_1 \end{bmatrix}, \qquad \mathbf{x}_2 = \begin{bmatrix} r_2 \\ s_2 \end{bmatrix}, \qquad \mathbf{x}_3 = \begin{bmatrix} r_3 \\ s_3 \end{bmatrix}, \ldots$$

- Our goal is to describe mathematically how these vectors might be related.

© 2016 Pearson Education, Ltd.

DIFFERENCE EQUATIONS

- Suppose demographic studies show that each year about 5% of the city's population moves to the suburbs (and 95% remains in the city), while 3% of the suburban population moves to the city (and 97% remains in the suburbs). See Fig. 2 below:

FIGURE 2 Annual percentage migration between city and suburbs.

© 2016 Pearson Education, Ltd.

DIFFERENCE EQUATIONS

- After 1 year, the original r_0 persons in the city are now distributed between city and suburbs as

$$\begin{bmatrix} .95r_0 \\ .05r_0 \end{bmatrix} = r_0 \begin{bmatrix} .95 \\ .05 \end{bmatrix} \quad \begin{matrix} \text{Remain in city} \\ \text{Move to suburbs} \end{matrix} \qquad (6)$$

- The so persons in the suburbs in 2000 are distributed 1 year later as

$$s_0 \begin{bmatrix} .03 \\ .97 \end{bmatrix} \quad \begin{matrix} \text{Move to city} \\ \text{Remain in suburbs} \end{matrix} \qquad (7)$$

© 2016 Pearson Education, Ltd.

DIFFERENCE EQUATIONS

- The vectors in (6) and (7) account for all of the populations in 2001. Thus

$$\begin{bmatrix} r_1 \\ s_1 \end{bmatrix} = r_0 \begin{bmatrix} .95 \\ .05 \end{bmatrix} + s_0 \begin{bmatrix} .03 \\ .97 \end{bmatrix} = \begin{bmatrix} .95 & .03 \\ .05 & .97 \end{bmatrix} \begin{bmatrix} r_0 \\ s_0 \end{bmatrix}$$

- That is,

$$\mathbf{x}_1 = M\mathbf{x}_0 \tag{8}$$

- Where M is the migration matrix determined by the following table:

From: City	From: Suburbs	To:
.95	.03	City
.05	.97	Suburbs

© 2016 Pearson Education, Ltd.

DIFFERENCE EQUATIONS

- Equation (8) describes how the population changes from 2000 to 2001. If the migration percentages remain constant, then the change from 2001 to 2002 is given by

$$\mathbf{x}_2 = M\mathbf{x}_1$$

- And similarly for 2002 to 2003 and subsequent years. In general,

$$\mathbf{x}_{k+1} = M\mathbf{x}_k \quad \text{for } k = 0, 1, 2, \ldots \tag{9}$$

- The sequence of vectors $\{x_0, x_1, x_2, \ldots\}$ describes the population of the city/suburban region over a period of years.

© 2016 Pearson Education, Ltd.

2 Matrix Algebra

2.1

MATRIX OPERATIONS

PEARSON

© 2016 Pearson Education, Inc.

MATRIX OPERATIONS

- If A is an $m \times n$ matrix—that is, a matrix with m rows and n columns—then the scalar entry in the ith row and jth column of A is denoted by a_{ij} and is called the (i, j)-entry of A. See the Fig. 1 below.
- Each column of A is a list of m real numbers, which identifies a vector in $\mathbb{R}^m$.

$$
\begin{array}{cc}
 & \text{Column } j \\
\text{Row } i & \begin{bmatrix} a_{11} & \cdots & a_{1j} & \cdots & a_{1n} \\ \vdots & & \vdots & & \vdots \\ a_{i1} & \cdots & a_{ij} & \cdots & a_{in} \\ \vdots & & \vdots & & \vdots \\ a_{m1} & \cdots & a_{mj} & \cdots & a_{mn} \end{bmatrix} = A \\
 & \begin{array}{ccccc} \uparrow & & \uparrow & & \uparrow \\ \mathbf{a}_1 & & \mathbf{a}_j & & \mathbf{a}_n \end{array}
\end{array}
$$

Matrix notation.

© 2016 Pearson Education, Inc.

MATRIX OPERATIONS

- The columns are denoted by $\mathbf{a}_1, \ldots, \mathbf{a}_n$, and the matrix A is written as

$$A = [a_1 \quad a_2 \quad \ldots \quad a_n]$$

- The number a_{ij} is the ith entry (from the top) of the jth column vector $\mathbf{a}_j$.
- The **diagonal entries** in an $m \times n$ matrix $A = \left[a_{ij} \right]$ are $a_{11}, a_{22}, a_{33}, \ldots$, and they form the **main diagonal** of A.
- A **diagonal matrix** is a square $n \times m$ matrix whose nondiagonal entries are zero.
- An example is the $n \times n$ identity matrix, I_n.

© 2016 Pearson Education, Inc.

SUMS AND SCALAR MULTIPLES

- An $m \times n$ matrix whose entries are all zero is a **zero matrix** and is written as 0.

- The two matrices are **equal** if they have the same size (*i.e.*, the same number of rows and the same number of columns) and if their corresponding columns are equal, which amounts to saying that their corresponding entries are equal.

- If A and B are $m \times n$ matrices, then the **sum** $A + B$ is the $m \times n$ matrix whose columns are the sums of the corresponding columns in A and B.

© 2016 Pearson Education, Inc.

SUMS AND SCALAR MULTIPLES

- Since vector addition of the columns is done entrywise, each entry in $A+B$ is the sum of the corresponding entries in A and B.

- The sum $A+B$ is defined only when A and B are the same size.

- **Example 1:** Let $A=\begin{bmatrix} 4 & 0 & 5 \\ -1 & 3 & 2 \end{bmatrix}, B=\begin{bmatrix} 1 & 1 & 1 \\ 3 & 5 & 7 \end{bmatrix},$

  and $C=\begin{bmatrix} 2 & -3 \\ 0 & 1 \end{bmatrix}$. Find $A+B$ and $A+C$.

© 2016 Pearson Education, Inc.

SUMS AND SCALAR MULTIPLES

- **Solution:** $A+B=\begin{bmatrix} 5 & 1 & 6 \\ 2 & 8 & 9 \end{bmatrix}$ but $A+C$ is not defined because A and C have different sizes.

- If r is a scalar and A is a matrix, then the **scalar multiple** rA is the matrix whose columns are r times the corresponding columns in A.

- **Theorem 1:** Let A, B, and C be matrices of the same size, and let r and s be scalars.
  a. $A+B=B+A$

© 2016 Pearson Education, Inc.

SUMS AND SCALAR MULTIPLES

b. $(A+B)+C=A+(B+C)$
c. $A+0=A$
d. $r(A+B)=rA+rB$
e. $(r+s)A=rA+sA$
f. $r(sA)=(rs)A$

- Each quantity in Theorem 1 is verified by showing that the matrix on the left side has the same size as the matrix on the right and that corresponding columns are equal.

© 2016 Pearson Education, Inc.

MATRIX MULTIPLICATION

- When a matrix B multiplies a vector $\mathbf{x}$, it transforms $\mathbf{x}$ into the vector $B\mathbf{x}$.
- If this vector is then multiplied in turn by a matrix A, the resulting vector is $A(B\mathbf{x})$. See the Fig. 2 below.

Multiplication by B and then A.

- Thus $A(B\mathbf{x})$ is produced from x by a *composition of mappings*—the linear transformations.

© 2016 Pearson Education, Inc.

MATRIX MULTIPLICATION

- Our goal is to represent this composite mapping as multiplication by a single matrix, denoted by AB, so that $A(B\mathrm{x})=(\mathrm{AB})\mathrm{x}$. See Fig. 3 below

Multiplication by AB.

- If A is $m \times n$, B is $n \times p$, and $\mathbf{x}$ is in $\mathbb{R}^p$, denote the columns of B by $\mathbf{b}_1, \ldots, \mathbf{b}p$ and the entries in $\mathbf{x}$ by $\mathbf{x}_1, \ldots, \mathbf{x}_p$.

© 2016 Pearson Education, Inc.

MATRIX MULTIPLICATION

- Then
$$B\mathrm{x} = x_1\mathrm{b}_1 + \ldots + x_p\mathrm{b}_p$$
- By the linearity of multiplication by A,
$$A(B\mathrm{x}) = A(x_1\mathrm{b}_1) + \ldots + A(x_p\mathrm{b}_p)$$
$$= x_1A\mathrm{b}_1 + \ldots + x_pA\mathrm{b}_p$$
- The vector $A(B\mathbf{x})$ is a linear combination of the vectors $A\mathbf{b}_1, \ldots, A\mathbf{b}_p$, using the entries in $\mathbf{x}$ as weights.
- In matrix notation, this linear combination is written as
$$A(B\mathrm{x}) = \begin{bmatrix} A\mathrm{b}_1 & A\mathrm{b}_2 & \cdots & A\mathrm{b}_p \end{bmatrix}\mathrm{x}$$

© 2016 Pearson Education, Inc.

MATRIX MULTIPLICATION

- Thus multiplication by $\begin{bmatrix} A\mathrm{b}_1 & A\mathrm{b}_2 & \cdots & A\mathrm{b}_p \end{bmatrix}$ transforms $\mathbf{x}$ into $A(B\mathbf{x})$.

- **Definition:** If A is an $m \times n$ matrix, and if B is an $n \times p$ matrix with columns $\mathbf{b}_1, \ldots, \mathbf{b}_p$, then the product AB is the $m \times p$ matrix whose columns are $A\mathbf{b}_1, \ldots, A\mathbf{b}_p$.

- That is,

$$AB = A\begin{bmatrix} \mathrm{b}_1 & \mathrm{b}_2 & \cdots & \mathrm{b}_p \end{bmatrix} = \begin{bmatrix} A\mathrm{b}_1 & A\mathrm{b}_2 & \cdots & A\mathrm{b}_p \end{bmatrix}$$

- *Multiplication of matrices corresponds to composition of linear transformations.*

© 2016 Pearson Education, Inc.

MATRIX MULTIPLICATION

- **Example 3:** Compute AB, where $A = \begin{bmatrix} 2 & 3 \\ 1 & -5 \end{bmatrix}$ and $B = \begin{bmatrix} 4 & 3 & 6 \\ 1 & -2 & 3 \end{bmatrix}$.

- **Solution:** Write $B = \begin{bmatrix} \mathrm{b}_1 & \mathrm{b}_2 & \mathrm{b}_3 \end{bmatrix}$, and compute:

© 2016 Pearson Education, Inc.

MATRIX MULTIPLICATION

$$A\mathbf{b}_1 = \begin{bmatrix} 2 & 3 \\ 1 & -5 \end{bmatrix}\begin{bmatrix} 4 \\ 1 \end{bmatrix}, \quad A\mathbf{b}_2 = \begin{bmatrix} 2 & 3 \\ 1 & -5 \end{bmatrix}\begin{bmatrix} 3 \\ -2 \end{bmatrix}, \quad A\mathbf{b}_3 = \begin{bmatrix} 2 & 3 \\ 1 & -5 \end{bmatrix}\begin{bmatrix} 6 \\ 3 \end{bmatrix}$$

$$= \begin{bmatrix} 11 \\ -1 \end{bmatrix} \qquad = \begin{bmatrix} 0 \\ 13 \end{bmatrix} \qquad = \begin{bmatrix} 21 \\ -9 \end{bmatrix}$$

- Then

$$AB = A\begin{bmatrix} \mathbf{b}_1 & \mathbf{b}_2 & \mathbf{b}_3 \end{bmatrix} = \begin{bmatrix} 11 & 0 & 21 \\ -1 & 13 & -9 \end{bmatrix}$$

$$\begin{array}{ccc} \uparrow & \uparrow & \uparrow \\ A\mathbf{b}_1 & A\mathbf{b}_2 & A\mathbf{b}_3 \end{array}$$

© 2016 Pearson Education, Inc.

MATRIX MULTIPLICATION

- Each column of AB is a linear combination of the columns of A using weights from the corresponding column of B.

Row—column rule for computing AB

- If a product AB is defined, then the entry in row i and column j of AB is the sum of the products of corresponding entries from row i of A and column j of B. If $(AB)_{ij}$ denotes the (i, j)-entry in AB, and if A is an $m \times n$ matrix, then

$$(AB)_{ij} = +a_{i2}b_{2j} + \ldots a_{i1}b_{1j} + a_{in}b_{nj}$$

© 2016 Pearson Education, Inc.

PROPERTIES OF MATRIX MULTIPLICATION

- **Theorem 2:** Let A be an $m \times n$ matrix, and let B and C have sizes for which the indicated sums and products are defined.
  a. $A(BC) = (AB)C$ (associative law of multiplication)
  b. $A(B + C) = AB + AC$ (left distributive law)
  c. $(B + C)A = BA + CA$ (right distributive law)
  d. $r(AB) = (rA)B = A(rB)$ for any scalar r
  e. $I_m A = A = AI_n$ (identity for matrix multiplication)

© 2016 Pearson Education, Inc.

PROPERTIES OF MATRIX MULTIPLICATION

- **Proof:** Property (a) follows from the fact that matrix multiplication corresponds to composition of linear transformations (which are functions), and it is known that the composition of functions is associative. Let

$$C = \begin{bmatrix} \mathbf{c}_1 & \cdots & \mathbf{c}_p \end{bmatrix}$$

- By the definition of matrix multiplication,

$$BC = \begin{bmatrix} B\mathbf{c}_1 & \cdots & B\mathbf{c}_p \end{bmatrix}$$

$$A(BC) = \begin{bmatrix} A(B\mathbf{c}_1) & \cdots & A(B\mathbf{c}_p) \end{bmatrix}$$

© 2016 Pearson Education, Inc.

PROPERTIES OF MATRIX MULTIPLICATION

- The definition of AB makes $A(B\mathrm{x}) = (AB)\mathrm{x}$ for all $\mathbf{x}$, so

$$A(BC) = \begin{bmatrix} (AB)\mathrm{c}_1 & \cdots & (AB)\mathrm{c}_p \end{bmatrix} = (AB)C$$

- The left-to-right order in products is critical because AB and BA are usually not the same.
- Because the columns of AB are linear combinations of the columns of A, whereas the columns of BA are constructed from the columns of B.
- The position of the factors in the product AB is emphasized by saying that A is *right-multiplied* by B or that B is *left-multiplied* by A.

© 2016 Pearson Education, Inc.

PROPERTIES OF MATRIX MULTIPLICATION

- If $AB = BA$, we say that A and B **commute** with one another.

- **Warnings:**
  1. In general, $AB \neq BA$.
  2. The cancellation laws do *not* hold for matrix multiplication. That is, if $AB = AC$, then it is *not* true in general that $B = C$.
  3. If a product AB is the zero matrix, you *cannot* conclude in general that either $A = 0$ or $B = 0$.

© 2016 Pearson Education, Inc.

POWERS OF A MATRIX

- If A is an $n \times n$ matrix and if k is a positive integer, then A^k denotes the product of k copies of A:

$$A^k = \underbrace{A \cdots A}_{k}$$

- If A is nonzero and if $\mathbf{x}$ is in $\mathbb{R}^n$, then $A^k\mathbf{x}$ is the result of left-multiplying $\mathbf{x}$ by A repeatedly k times.

- If $k = 0$, then $A^0\mathbf{x}$ should be $\mathbf{x}$ itself.

- Thus A^0 is interpreted as the identity matrix.

© 2016 Pearson Education, Inc.

THE TRANSPOSE OF A MATRIX

- Given an $m \times n$ matrix A, the **transpose** of A is the $n \times m$ matrix, denoted by A^T, whose columns are formed from the corresponding rows of A.

Theorem 3: Let A and B denote matrices whose sizes are appropriate for the following sums and products.

a. $(A^T)^T = A$

b. $(A+B)^T = A^T + B^T$

c. For any scalar r, $(rA)^T = rA^T$

d. $(AB)^T = B^T A^T$

© 2016 Pearson Education, Inc.

THE TRANSPOSE OF A MATRIX

- The transpose of a product of matrices equals the product of their transposes in the *reverse* order.

© 2016 Pearson Education, Inc.

2 Matrix Algebra

2.2

THE INVERSE OF A MATRIX

© 2016 Pearson Education, Ltd.

MATRIX OPERATIONS

- An $n \times n$ matrix A is said to be **invertible** if there is an $n \times n$ matrix C such that

$$CA = I \quad \text{and} \quad AC = I$$

  where $I = I_n$, the $n \times n$ identity matrix.
- In this case, C is an **inverse** of A.
- In fact, C is uniquely determined by A, because if B were another inverse of A, then

$$B = BI = B(AC) = (BA)C = IC = C.$$

- This unique inverse is denoted by A^{-1}, so that

$$A^{-1}A = I \text{ and } AA^{-1} = I.$$

© 2016 Pearson Education, Ltd.

MATRIX OPERATIONS

- **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 quantity $ad - bc$ is called the **determinant** of A, and we write $\det A = ad - bc$
- This theorem says that a 2×2 matrix A is invertible if and only if $\det A \neq 0$

© 2016 Pearson Education, Ltd.

MATRIX OPERATIONS

- **Theorem 5:** If A is an invertible $n \times n$ matrix, then for each **b** in $\mathbb{R}^n$, the equation $A\mathrm{x} = \mathrm{b}$ has the unique solution $\mathrm{x} = A^{-1}\mathrm{b}$.
- **Proof:** Take any **b** in $\mathbb{R}^n$.
- A solution exists because if $A^{-1}\mathrm{b}$ is substituted for **x**, then $A\mathrm{x} = A(A^{-1}\mathrm{b}) = (AA^{-1})\mathrm{b} = I\mathrm{b} = \mathrm{b}$.
- So $A^{-1}\mathrm{b}$ is a solution.
- To prove that the solution is unique, show that if **u** is any solution, then **u** must be $A^{-1}\mathrm{b}$.
- If $A\mathrm{u} = \mathrm{b}$, we can multiply both sides by A^{-1} and obtain $A^{-1}A\mathrm{u} = A^{-1}\mathrm{b}$, $I\mathrm{u} = A^{-1}\mathrm{b}$, and $\mathrm{u} = A^{-1}\mathrm{b}$.

© 2016 Pearson Education, Ltd.

MATRIX OPERATIONS

- **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,

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

© 2016 Pearson Education, Ltd.

MATRIX OPERATIONS

- **Proof:** To verify statement (a), find a matrix C such that

$$A^{-1}C = I \quad \text{and} \quad CA^{-1} = I$$

- These equations are satisfied with A in place of C. Hence A^{-1} is invertible, and A is its inverse.
- Next, to prove statement (b), compute:

$$(AB)(B^{-1}A^{-1}) = A(BB^{-1})A^{-1} = AIA^{-1} = AA^{-1} = I$$

- A similar calculation shows that $(B^{-1}A^{-1})(AB) = I.$
- For statement (c), use Theorem 3(d), read from right to left, $(A^{-1})^T A^T = (AA^{-1})^T = I^T = I$.
- Similarly, $A^T(A^{-1})^T = I^T = I.$

© 2016 Pearson Education, Ltd.

ELEMENTARY MATRICES

- Hence A^T is invertible, and its inverse is $(A^{-1})T$.
- The generalization of Theorem 6(b) is as follows:

  The product of $n \times n$ invertible matrices is invertible, and the inverse is the product of their inverses in the reverse order.
- An invertible matrix A is row equivalent to an identity matrix, and we can find A^{-1} by *watching the row reduction of A to I.*
- An **elementary matrix** is one that is obtained by performing a single elementary row operation on an identity matrix.

© 2016 Pearson Education, Ltd.

ELEMENTARY MATRICES

- **Example 5:** Let $E_1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -4 & 0 & 1 \end{bmatrix}$, $E_2 = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$,

  $E_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 5 \end{bmatrix}$, $A = \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix}$

  Compute E_1A, E_2A, and E_3A, and describe how these products can be obtained by elementary row operations on A.

© 2016 Pearson Education, Ltd.

ELEMENTARY MATRICES

- **Solution:** Verify that

$$E_1A = \begin{bmatrix} a & b & c \\ d & e & f \\ g-4a & h-4b & i-4c \end{bmatrix}, E_2A = \begin{bmatrix} d & e & f \\ a & b & c \\ g & h & i \end{bmatrix},$$

$$E_3A = \begin{bmatrix} a & b & c \\ d & e & f \\ 5g & 5h & 5i \end{bmatrix}.$$

- Addition of -4 times row 1 of A to row 3 produces E_1A.

© 2016 Pearson Education, Ltd.

ELEMENTARY MATRICES

- An interchange of rows 1 and 2 of A produces E_2A, and multiplication of row 3 of A by 5 produces E_3A.

- Left-multiplication (that is, multiplication on the left) by E_1 in Example 1 has the same effect on any $3 \times n$ matrix.

- Since $E_1 \cdot I = E_1$, we see that E_1 *itself* is produced by this same row operation on the identity.

© 2016 Pearson Education, Ltd.

ELEMENTARY MATRICES

- Example 5 illustrates the following general fact about elementary matrices.

- If an elementary row operation is performed on an $m \times n$ matrix A, the resulting matrix can be written as EA, where the $m \times m$ matrix E is created by performing the same row operation on I_m.

- Each elementary matrix E is invertible. The inverse of E is the elementary matrix of the same type that transforms E back into I.

© 2016 Pearson Education, Ltd.

ELEMENTARY MATRICES

- **Theorem 7:** An $n \times n$ matrix A is invertible if and only if A is row equivalent to I_n, and in this case, any sequence of elementary row operations that reduces A to I_n also transforms I_n into A^{-1}.
- **Proof:** Suppose that A is invertible.
- Then, since the equation $A\mathrm{x} = \mathrm{b}$ has a solution for each **b** (Theorem 5), A has a pivot position in every row.
- Because A is square, the n pivot positions must be on the diagonal, which implies that the reduced echelon form of A is I_n. That is, $A \sim I_n$.

© 2016 Pearson Education, Ltd.

ELEMENTARY MATRICES

- Now suppose, conversely, that $A \sim I_n$.
- Then, since each step of the row reduction of A corresponds to left-multiplication by an elementary matrix, there exist elementary matrices $E_1, \ldots, E_p$ such that $A \sim E_1 A \sim E_2(E_1 A) \sim \ldots \sim E_p(E_{p-1} \ldots E_1 A) = I_n$
- That is,

$$(1) \qquad E_p \ldots E_1 A = I_n$$

- Since the product $E_p \ldots E_1$ of invertible matrices is invertible, (1) leads to

$$(E_p \ldots E_1)^{-1}(E_p \ldots E_1)A = (E_p \ldots E_1)^{-1} I_n$$

$$A = (E_p \ldots E_1)^{-1}$$

© 2016 Pearson Education, Ltd.

ALGORITHM FOR FINDING A^{-1}

- Thus A is invertible, as it is the inverse of an invertible matrix (Theorem 6). Also,

$$A^{-1} = \left[(E_p \ldots E_1)^{-1} \right]^{-1} = E_p \ldots E_1 \cdot$$

- Then $A^{-1} = E_p \ldots E_1 \cdot I_n$, which says that A^{-1} results from applying E_1, ..., E_p successively to I_n.
- This is the same sequence in (1) that reduced A to I_n.
- Row reduce the augmented matrix $\begin{bmatrix} A & I \end{bmatrix}$. If A is row equivalent to I, then $\begin{bmatrix} A & I \end{bmatrix}$ is row equivalent to $\begin{bmatrix} I & A^{-1} \end{bmatrix}$. Otherwise, A does not have an inverse.

© 2016 Pearson Education, Ltd.

ALGORITHM FOR FINDING A^{-1}

- **Example 2:** Find the inverse of the matrix

$$A = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 3 \\ 4 & -3 & 8 \end{bmatrix}, \text{ if it exists.}$$

- **Solution:**

$$\begin{bmatrix} A & I \end{bmatrix} = \begin{bmatrix} 0 & 1 & 2 & 1 & 0 & 0 \\ 1 & 0 & 3 & 0 & 1 & 0 \\ 4 & -3 & 8 & 0 & 0 & 1 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & 3 & 0 & 1 & 0 \\ 0 & 1 & 2 & 1 & 0 & 0 \\ 4 & -3 & 8 & 0 & 0 & 1 \end{bmatrix}$$

© 2016 Pearson Education, Ltd.

ALGORITHM FOR FINDING A^{-1}

$$\sim \begin{bmatrix} 1 & 0 & 3 & 0 & 1 & 0 \\ 0 & 1 & 2 & 1 & 0 & 0 \\ 0 & -3 & -4 & 0 & -4 & 1 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & 3 & 0 & 1 & 0 \\ 0 & 1 & 2 & 1 & 0 & 0 \\ 0 & 0 & 2 & 3 & -4 & 1 \end{bmatrix}$$

$$\sim \begin{bmatrix} 1 & 0 & 3 & 0 & 1 & 0 \\ 0 & 1 & 2 & 1 & 0 & 0 \\ 0 & 0 & 1 & 3/2 & -2 & 1/2 \end{bmatrix}$$

$$\sim \begin{bmatrix} 1 & 0 & 0 & -9/2 & 7 & -3/2 \\ 0 & 1 & 0 & -2 & 4 & -1 \\ 0 & 0 & 1 & 3/2 & -2 & 1/2 \end{bmatrix}$$

© 2016 Pearson Education, Ltd.

ALGORITHM FOR FINDING A^{-1}

- Theorem 7 shows, since $A \sim I$, that A is invertible, and

$$A^{-1} = \begin{bmatrix} -9/2 & 7 & -3/2 \\ -2 & 4 & -1 \\ 3/2 & -2 & 1/2 \end{bmatrix}$$

- Now, check the final answer.

$$AA^{-1} = \begin{bmatrix} 0 & 1 & 2 \\ 1 & 0 & 3 \\ 4 & -3 & 8 \end{bmatrix} \begin{bmatrix} -9/2 & 7 & -3/2 \\ -2 & 4 & -1 \\ 3/2 & -2 & 1/2 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

© 2016 Pearson Education, Ltd.

ANOTHER VIEW OF MATRIX INVERSION

- It is not necessary to check that $A^{-1}A = I$ since A is invertible.

- Denote the columns of I_n by $\mathbf{e}_1, \ldots, \mathbf{e}_n$.

- Then row reduction of $\begin{bmatrix} A & I \end{bmatrix}$ to $\begin{bmatrix} I & A^{-1} \end{bmatrix}$ can be viewed as the simultaneous solution of the n systems

$$A\mathrm{x} = \mathrm{e}_1,\ A\mathrm{x} = \mathrm{e}_2,\ \ldots,\ A\mathrm{x} = \mathrm{e}_n \qquad (2)$$

where the "augmented columns" of these systems have all been placed next to A to form

$$\begin{bmatrix} A & \mathrm{e}_1 & \mathrm{e}_2 & \cdots & \mathrm{e}_n \end{bmatrix} = \begin{bmatrix} A & I \end{bmatrix}.$$

© 2016 Pearson Education, Ltd.

ANOTHER VIEW OF MATRIX INVERSION

- The equation $AA^{-1} = I$ and the definition of matrix multiplication show that the columns of A^{-1} are precisely the solutions of the systems in (2).

© 2016 Pearson Education, Ltd.

2 Matrix Algebra

2.3

CHARACTERIZATIONS OF INVERTIBLE MATRICES

PEARSON

© 2016 Pearson Education, Inc.

THE INVERTIBLE MATRIX THEOREM

- **Theorem 8:** 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. A is an invertible matrix.
  b. A is row equivalent to the $n \times n$ identity matrix.
  c. A has n pivot positions.
  d. The equation $A\mathrm{x} = 0$ has only the trivial solution.
  e. The columns of A form a linearly independent set.

© 2016 Pearson Education, Inc.

THE INVERTIBLE MATRIX THEOREM

f. The linear transformation $\mathbf{x} \mapsto A\mathbf{x}$ is one-to-one.

g. The equation $A\mathbf{x} = \mathbf{b}$ has at least one solution for each b in $\mathbb{R}^n$.

h. The columns of A span $\mathbb{R}^n$.

i. The linear transformation $\mathbf{x} \mapsto A\mathbf{x}$ maps $\mathbb{R}^n$ onto $\mathbb{R}^n$.

j. There is an $n \times n$ matrix C such that $CA = I$.

k. There is an $n \times n$ matrix D such that $AD = I$.

l. A^T is an invertible matrix.

© 2016 Pearson Education, Inc.

THE INVERTIBLE MATRIX THEOREM

- First, we need some notation.
- If the truth of statement (a) always implies that statement (j) is true, we say that (a) *implies* (j) and write $(a) \Rightarrow (j)$.
- The proof will establish the "circle" of implications as shown in the following figure.

$$\begin{tikzcd}
 & (a) \arrow[dr, Rightarrow] & \\
(b) \arrow[ur, Rightarrow] & & (j) \arrow[d, Rightarrow] \\
(c) \arrow[u, Rightarrow] & (d) \arrow[l, Rightarrow] &
\end{tikzcd}$$

- If any one of these five statements is true, then so are the others.

© 2016 Pearson Education, Inc.

THE INVERTIBLE MATRIX THEOREM

- Finally, the proof will link the remaining statements of the theorem to the statements in this circle.
- **Proof:** If statement (a) is true, then A^{-1} works for C in (j), so $(\text{a}) \Rightarrow (\text{j})$.
- Next, $(\text{j}) \Rightarrow (\text{d})$.
- Also, $(\text{d}) \Rightarrow (\text{c})$.
- If A is square and has n pivot positions, then the pivots must lie on the main diagonal, in which case the reduced echelon form of A is I_n.
- Thus $(\text{c}) \Rightarrow (\text{b})$.
- Also, $(\text{b}) \Rightarrow (\text{a})$.

© 2016 Pearson Education, Inc.

THE INVERTIBLE MATRIX THEOREM

- This completes the circle in the previous figure.
- Next, $(\mathrm{a}) \Rightarrow (\mathrm{k})$ because A^{-1} works for D.
- Also, $(\mathrm{k}) \Rightarrow (\mathrm{g})$ and $(\mathrm{g}) \Rightarrow (\mathrm{a})$.
- So (k) and (g) are linked to the circle.
- Further, (g), (h), and (i) are equivalent for any matrix.
- Thus, (h) and (i) are linked through (g) to the circle.
- Since (d) is linked to the circle, so are (e) and (f), because (d), (e), and (f) are all equivalent for *any* matrix A.
- Finally, $(\mathrm{a}) \Rightarrow (\mathrm{l})$ and $(\mathrm{l}) \Rightarrow (\mathrm{a})$.
- This completes the proof.

© 2016 Pearson Education, Inc.

THE INVERTIBLE MATRIX THEOREM

- Theorem 8 could also be written as "The equation $A\mathrm{x} = \mathrm{b}$ has a *unique* solution for each **b** in $\mathbb{R}^n$."
- This statement implies (b) and hence implies that A is invertible.
- The following fact follows from Theorem 8.

  Let A and B be square matrices. If $AB = I$, then A and B are both invertible, with $B = A^{-1}$ and $A = B^{-1}$.
- The Invertible Matrix Theorem divides the set of all $n \times n$ matrices into two disjoint classes: the invertible (nonsingular) matrices, and the noninvertible (singular) matrices.

© 2016 Pearson Education, Inc.

THE INVERTIBLE MATRIX THEOREM

- Each statement in the theorem describes a property of every $n \times n$ invertible matrix.

- The *negation* of a statement in the theorem describes a property of every $n \times n$ singular matrix.

- For instance, an $n \times n$ singular matrix is *not* row equivalent to I_n, does *not* have n pivot position, and has linearly *dependent* columns.

© 2016 Pearson Education, Inc.

THE INVERTIBLE MATRIX THEOREM

- **Example 1:** Use the Invertible Matrix Theorem to decide if A is invertible:

$$A = \begin{bmatrix} 1 & 0 & -2 \\ 3 & 1 & -2 \\ -5 & -1 & 9 \end{bmatrix}$$

- **Solution:**

$$A \sim \begin{bmatrix} 1 & 0 & -2 \\ 0 & 1 & 4 \\ 0 & -1 & -1 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & -2 \\ 0 & 1 & 4 \\ 0 & 0 & 3 \end{bmatrix}$$

© 2016 Pearson Education, Inc.

THE INVERTIBLE MATRIX THEOREM

- So A has three pivot positions and hence is invertible, by the Invertible Matrix Theorem, statement (c).

- The Invertible Matrix Theorem *applies only to square matrices*.

- For example, if the columns of a 4×3 matrix are linearly independent, we cannot use the Invertible Matrix Theorem to conclude anything about the existence or nonexistence of solutions of equation of the form $A\mathrm{x} = \mathrm{b}$.

© 2016 Pearson Education, Inc.

INVERTIBLE LINEAR TRANSFORMATIONS

- Matrix multiplication corresponds to composition of linear transformations.
- When a matrix A is invertible, the equation $A^{-1}A\mathrm{x} = \mathrm{x}$ can be viewed as a statement about linear transformations. See the following figure.

A^{-1} transforms $A\mathbf{x}$ back to $\mathbf{x}$.

© 2016 Pearson Education, Inc.

INVERTIBLE LINEAR TRANSFORMATIONS

- A linear transformation $T: \mathbb{R}^n \to \mathbb{R}^n$ is said to be **invertible** if there exists a function $S: \mathbb{R}^n \to \mathbb{R}^n$ such that

$$S(T(\mathrm{x})) = \mathrm{x} \quad \text{for all } \mathbf{x} \text{ in } \mathbb{R}^n \qquad (1)$$

$$T(S(\mathrm{x})) = \mathrm{x} \quad \text{for all } \mathbf{x} \text{ in } \mathbb{R}^n \qquad (2)$$

- **Theorem 9:** Let $T: \mathbb{R}^n \to \mathbb{R}^n$ be a linear transformation and let A be the standard matrix for T. Then T is invertible if and only if A is an invertible matrix. In that case, the linear transformation S given by $S(\mathrm{x}) = A^{-1}\mathrm{x}$ is the unique function satisfying equation (1) and (2).

© 2016 Pearson Education, Inc.

INVERTIBLE LINEAR TRANSFORMATIONS

- **Proof:** Suppose that T is invertible.
- Then (2) shows that T is onto $\mathbb{R}^n$, for if **b** is in $\mathbb{R}^n$ and $\mathrm{x} = S(\mathrm{b})$, then $T(\mathrm{x}) = T(S(\mathrm{b})) = \mathrm{b}$, so each **b** is in the range of T.
- Thus A is invertible, by the Invertible Matrix Theorem, statement (i).
- Conversely, suppose that A is invertible, and let $S(\mathrm{x}) = A^{-1}\mathrm{x}$. Then, S is a linear transformation, and S satisfies (1) and (2).
- For instance, $S(T(\mathrm{x})) = S(A\mathrm{x}) = A^{-1}(A\mathrm{x}) = \mathrm{x}$.
- Thus, T is invertible.

© 2016 Pearson Education, Inc.

2 Matrix Algebra

2.4

PARTITIONED MATRICES

PEARSON

© 2016 Pearson Education, Ltd.

PARTITIONED MATRICES

- A key feature of our work with matrices has been the ability to regard matrix A as a list of column vectors rather than just a rectangular array of numbers.
- This point of view has been so useful that we wish to consider other **partitions** of A, indicated by horizontal and vertical dividing rules, as in Example 1 on the next slide.

© 2016 Pearson Education, Ltd.

PARTITIONED MATRICES

- **Example 1** The matrix

$$A = \left[\begin{array}{ccc|cc|c} 3 & 0 & -1 & 5 & 9 & -2 \\ -5 & 2 & 4 & 0 & -3 & 1 \\ \hline -8 & -6 & 3 & 1 & 7 & -4 \end{array}\right]$$

- Can also be written as the 2×3 **partitioned** (or **block**) **matrix**

$$A = \begin{bmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \end{bmatrix}$$

- Whose entries are the *blocks* (or *submatrices*)

$$A_{11} = \begin{bmatrix} 3 & 0 & -1 \\ -5 & 2 & 4 \end{bmatrix}, \quad A_{12} = \begin{bmatrix} 5 & 9 \\ 0 & -3 \end{bmatrix}, \quad A_{13} = \begin{bmatrix} -2 \\ 1 \end{bmatrix}$$

$$A_{21} = \begin{bmatrix} -8 & -6 & 3 \end{bmatrix}, \quad A_{22} = \begin{bmatrix} 1 & 7 \end{bmatrix}, \quad A_{23} = \begin{bmatrix} -4 \end{bmatrix}$$

© 2016 Pearson Education, Ltd.

ADDITION AND SCALAR MULTIPLICATION

- If matrices A and B are the same size and are partitioned in exactly the same way, then it is natural to make the same partition of the ordinary matrix sum $A + B$.
- In this case, each block of $A + B$ is the (matrix) sum of the corresponding blocks of A and B.
- Multiplication of a partitioned matrix by a scalar is also computed block by block.

© 2016 Pearson Education, Ltd.

MULTIPLICATION OF PARTITIONED MATRICES

- Partitioned matrices can be multiplied by the usual row—column rule as if the block entries were scalars, provided that for a product AB, the column partition of A matches the row partition of B.
- **Example 3** Let

$$A = \left[\begin{array}{ccc|cc} 2 & -3 & 1 & 0 & -4 \\ 1 & 5 & -2 & 3 & -1 \\ \hline 0 & -4 & -2 & 7 & -1 \end{array}\right] = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}, \quad B = \left[\begin{array}{cc} 6 & 4 \\ -2 & 1 \\ -3 & 7 \\ \hline -1 & 3 \\ 5 & 2 \end{array}\right] = \begin{bmatrix} B_1 \\ B_2 \end{bmatrix}$$

- The 5 columns of A are partitioned into a set of 3 columns and then a set of 2 columns. The 5 rows of B are partitioned in the same way—into a set of 3 rows and then a set of 2 rows.

© 2016 Pearson Education, Ltd.

MULTIPLICATION OF PARTITIONED MATRICES

- We say that the partitions of A and B are conformable for block multiplication. It can be shown that the ordinary product AB can be written as

$$AB = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \begin{bmatrix} B_1 \\ B_2 \end{bmatrix} = \begin{bmatrix} A_{11}B_1 + A_{12}B_2 \\ A_{21}B_1 + A_{22}B_2 \end{bmatrix} = \left[\begin{array}{cc} -5 & 4 \\ -6 & 2 \\ \hline 2 & 1 \end{array}\right]$$

- It is important for each smaller product in the expression for AB to be written with the submatrix from A on the left, since matrix multiplication is not commutative.

© 2016 Pearson Education, Ltd.

MULTIPLICATION OF PARTITIONED MATRICES

- For instance,

$$A_{11}B_1 = \begin{bmatrix} 2 & -3 & 1 \\ 1 & 5 & -2 \end{bmatrix} \begin{bmatrix} 6 & 4 \\ -2 & 1 \\ -3 & 7 \end{bmatrix} = \begin{bmatrix} 15 & 12 \\ 2 & -5 \end{bmatrix}$$

$$A_{12}B_2 = \begin{bmatrix} 0 & -4 \\ 3 & -1 \end{bmatrix} \begin{bmatrix} -1 & 3 \\ 5 & 2 \end{bmatrix} = \begin{bmatrix} -20 & -8 \\ -8 & 7 \end{bmatrix}$$

- Hence the top block in AB is

$$A_{11}B_1 + A_{12}B_2 = \begin{bmatrix} 15 & 12 \\ 2 & -5 \end{bmatrix} + \begin{bmatrix} -20 & -8 \\ -8 & 7 \end{bmatrix} = \begin{bmatrix} -5 & 4 \\ -6 & 2 \end{bmatrix}$$

© 2016 Pearson Education, Ltd.

MULTIPLICATION OF PARTITIONED MATRICES

- **Theorem 10**: Column—Row Expansion of AB

If A is $m \times n$ and B is $n \times p$, then

$$AB = [\,\mathrm{col}_1(A) \quad \mathrm{col}_2(A) \quad \cdots \quad \mathrm{col}_n(A)\,]\begin{bmatrix} \mathrm{row}_1(B) \\ \mathrm{row}_2(B) \\ \vdots \\ \mathrm{row}_n(B) \end{bmatrix} \tag{1}$$

$$= \mathrm{col}_1(A)\,\mathrm{row}_1(B) + \cdots + \mathrm{col}_n(A)\,\mathrm{row}_n(B)$$

- **Proof** For each row index i and column index j, the (i, j)-entry in $\mathrm{col}_k(A)$and b_{kj} from $\mathrm{row}_k(B)$ is the product of a_{ik} from $\mathrm{col}_k(A)$ and b_{kj} from $\mathrm{row}_k(B)$.

© 2016 Pearson Education, Ltd.

MULTIPLICATION OF PARTITIONED MATRICES

- Hence the (i, j)-entry in the sum shown in equation (1) is

$$\underset{(k=1)}{a_{i1}b_{1j}} + \underset{(k=2)}{a_{i2}b_{2j}} + \cdots + \underset{(k=n)}{a_{in}b_{nj}}$$

- This sum is also the (i, j)-entry in AB, by the row—column rule.

© 2016 Pearson Education, Ltd.

INVERSES OF PARTIONED MATRICES

- The next example illustrates calculations involving inverses and partitioned matrices.
- **Example 5** A matrix of the form

$$A = \begin{bmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{bmatrix}$$

- Is said to be *block upper triangular*. Assume that A_{11} is $p \times p$, A_{22} is $q \times q$, and A is invertible. Find a formula for A^{-1}.

© 2016 Pearson Education, Ltd.

INVERSES OF PARTIONED MATRICES

- **Solution** Denote A^{-1} by B and partition B so that

$$\begin{bmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{bmatrix} \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix} = \begin{bmatrix} I_p & 0 \\ 0 & I_q \end{bmatrix} \qquad (2)$$

- This matrix equation provides four equations that will lead to the unknown blocks $B_{11}\ldots,B_{22}$. Compute the product on the left side of equation (2), and equate each entry with the corresponding block in the identity matrix on the right.

© 2016 Pearson Education, Ltd.

INVERSES OF PARTIONED MATRICES

- That is, set

$$A_{11}B_{11} + A_{12}B_{21} = Ip \quad (3)$$

$$A_{11}B_{11} + A_{12}B_{22} = 0 \quad (4)$$

$$A_{22}B_{21} = 0 \quad (5)$$

$$A_{22}B_{22} = Iq \quad (6)$$

- By itself, equation (6) does not show that A_{22} is invertible. However, since A_{22} is square, the Invertible Matrix Theorem and (6) together show that A_{22} is invertible and $B_{22} = A_{22}^{-1}$.

© 2016 Pearson Education, Ltd.

INVERSES OF PARTIONED MATRICES

- Next, left-multiply both sides of (5) by A_{22}^{-1} and obtain
$$B_{21} = A_{22}^{-1}0 = 0$$
- So that (3) simplifies to
$$A_{11}B_{11} + 0 = I_p$$
- Since A_{11} is square, this shows that A_{11} is invertible and $B_{22} = A_{22}^{-1}$. Finally, use these results with (4) to find that

$$A_{11}B_{12} = -A_{12}B_{22} = -A_{12}A_{22}^{-1} \quad \text{and} \quad B_{12} = A_{11}^{-1}A_{12}A_{22}^{-1}$$

© 2016 Pearson Education, Ltd.

INVERSES OF PARTIONED MATRICES

- Thus

$$A^{-1} = \begin{bmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{bmatrix}^{-1} = \begin{bmatrix} A_{11}^{-1} & A_{11}^{-1}A_{12}A_{22}^{-1} \\ 0 & A_{22}^{-1} \end{bmatrix}$$

- A **block diagonal matrix** is a partitioned matrix with zero blocks off the main diagonal (of blocks). Such a matrix is invertible if and only if each block on the diagonal is invertible.

© 2016 Pearson Education, Ltd.

2 Matrix Algebra

2.5

MATRIX FACTORIZATIONS

© 2016 Pearson Education, Ltd.

MATRIX FACTORIZATIONS

- A *factorization* of a matrix A is an equation that expresses A as a product of two or more matrices.
- Whereas matrix multiplication involves a *synthesis* of data (combining the effects of two or more linear transformations into a single matrix), matrix factorization is an *analysis* of data.

© 2016 Pearson Education, Ltd.

THE LU FACTORIZATION

- The LU factorization, described on the next few slides, is motivated by the fairly common industrial and business problem of solving a sequence of equations, all with the same coefficient matrix:

$$Ax = b_1, \quad Ax = b_2, \ldots, \quad Ax = bp \qquad (1)$$

- When A is invertible, one could compute A^{-1} and then compute $A^{-1}b_1$, $A^{-1}b_2$, and so on.
- However, it is more efficient to solve the first equation in the sequence (1) by row reduction and obtain the LU factorization of A at the same time. Thereafter, the remaining equations in sequence (1) are solved with the LU factorization.

© 2016 Pearson Education, Ltd.

THE LU FACTORIZATION

- At first, assume that A is an $m \times n$ matrix that can be row reduced to echelon form, *without row interchanges*.
- Then A can be written in the form $A = LU$, were L is an $m \times m$ lower triangular matrix with 1's on the diagonal and U is an $m \times n$ echelon form of A.
- For instance, see Fig. 1 below. Such a factorization is called an **LU factorization** of A. The matrix L is invertible and is called a unit lower triangular matrix.

$$A = \underbrace{\begin{bmatrix} 1 & 0 & 0 & 0 \\ * & 1 & 0 & 0 \\ * & * & 1 & 0 \\ * & * & * & 1 \end{bmatrix}}_{L} \underbrace{\begin{bmatrix} \blacksquare & * & * & * & * \\ 0 & \blacksquare & * & * & * \\ 0 & 0 & 0 & \blacksquare & * \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}}_{U}$$

© 2016 Pearson Education, Ltd.

THE LU FACTORIZATION

- Before studying how to construct L and U, we should look at why they are so useful. When $A = LU$, the equation $Ax = b$ can be written as $L(Ux) = b$.
- Writing y for Ux, we can find x by solving the pair of equations

$$\boxed{\begin{aligned} Ly &= b \\ Ux &= y \end{aligned}}$$

- First solve $Ly = b$ for y, and then solve $Ux = y$ for x. See Fig. 2 on the next slide. Each equation is easy to solve because L and U are triangular.

© 2016 Pearson Education, Ltd.

THE LU FACTORIZATION

FIGURE 2 Factorization of the mapping $\mathbf{x} \mapsto A\mathbf{x}$.

- **Example 1** It can be verified that

$$A = \begin{bmatrix} 3 & -7 & -2 & 2 \\ -3 & 5 & 1 & 0 \\ 6 & -4 & 0 & -5 \\ -9 & 5 & -5 & 12 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ -1 & 1 & 0 & 0 \\ 2 & -5 & 1 & 0 \\ -3 & 8 & 3 & 1 \end{bmatrix} \begin{bmatrix} 3 & -7 & -2 & 2 \\ 0 & -2 & -1 & 2 \\ 0 & 0 & -1 & 1 \\ 0 & 0 & 0 & -1 \end{bmatrix} = LU$$

- Use this factorization of A to solve $Ax=b$, where $\mathrm{b} = \begin{bmatrix} -9 \\ 5 \\ 7 \\ 11 \end{bmatrix}$

© 2016 Pearson Education, Ltd.

THE LU FACTORIZATION

- **Solution** The solution of $Ly = b$ needs only 6 multiplications and 6 additions, because the arithmetic takes place only in column 5.

$$[L \quad \mathbf{b}] = \begin{bmatrix} 1 & 0 & 0 & 0 & -9 \\ -1 & 1 & 0 & 0 & 5 \\ 2 & -5 & 1 & 0 & 7 \\ -3 & 8 & 3 & 1 & 11 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & 0 & 0 & -9 \\ 0 & 1 & 0 & 0 & -4 \\ 0 & 0 & 1 & 0 & 5 \\ 0 & 0 & 0 & 1 & 1 \end{bmatrix} = [I \quad \mathbf{y}]$$

- Then, for $Ux = y$, the "backward" phase of row reduction requires 4 divisions, 6 multiplications, and 6 additions.

© 2016 Pearson Education, Ltd.

THE LU FACTORIZATION

- For instance, creating the zeros in column 4 of $[U \quad y]$ requires 1 division in row 4 and 3 multiplication-addition pairs to add multiples of row 4 to the rows above.

$$[U \quad \mathbf{y}] = \begin{bmatrix} 3 & -7 & -2 & 2 & -9 \\ 0 & -2 & -1 & 2 & -4 \\ 0 & 0 & -1 & 1 & 5 \\ 0 & 0 & 0 & -1 & 1 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & 0 & 0 & 3 \\ 0 & 1 & 0 & 0 & 4 \\ 0 & 0 & 1 & 0 & -6 \\ 0 & 0 & 0 & 1 & -1 \end{bmatrix}, \quad \mathbf{x} = \begin{bmatrix} 3 \\ 4 \\ -6 \\ -1 \end{bmatrix}$$

- To find x requires 28 arithmetic operations, or "flops" (floating point operations), excluding the cost of finding L and U. In contrast, row reduction of $[A \quad b]$ to $[I \quad x]$ takes 62 operations.

© 2016 Pearson Education, Ltd.

AN LU FACTORIZATION ALGORITHM

- Suppose A can be reduced to an echelon form U using only row replacements that add a multiple of one row to another below it.
- In this case, there exist unit lower triangular elementary matrices $E_1, \ldots, E_p$ such that

$$E_p \ldots E_1 A = U$$

- Then (3)

$$A = (Ep \ldots E_1)^{-1} U = LU$$

- where

$$L = (E_p \ldots E_1)^{-1} \quad (4)$$

- It can be shown that products and inverses of unit lower triangular matrices are also unit lower triangular. Thus L is unit lower triangular.

© 2016 Pearson Education, Ltd.

AN LU FACTORIZATION ALGORITHM

- Note that row operations in equation (3), which reduce A to U, also reduce the L in equation (4) to I, because $E_p \dots E_1 L = (E_p \dots E_1)(E_p \dots E_1)^{-1} = I$. This observation is the key to *constructing* L.

Algorithm for an LU Factorization

1. Reduce A to an echelon form U by a sequence of row replacement operations, if possible.
2. Place entries in L such that the *same sequence of row operations* reduces L to I.

© 2016 Pearson Education, Ltd.

AN LU FACTORIZATION ALGORITHM

- Step 1 is not always possible, but when it is, the argument above shows that an LU factorization exists.
- Example 2 on the followings slides will show how to implement step 2. By construction, L will satisfy

$$(Ep \ldots E_1)L = I$$

- using the same $E_p, \ldots, E_1$ as in equation (3). Thus L will be invertible, by the Invertible Matrix Theorem, with $(E_p \ldots E_1) = \mathrm{L}^{-1}$. From (3), $L^{-1}A = U$, and $A = LU$. So step 2 will produce an acceptable L.

© 2016 Pearson Education, Ltd.

AN LU FACTORIZATION ALGORITHM

- **Example 2** Find an LU factorization of

$$A = \begin{bmatrix} 2 & 4 & -1 & 5 & -2 \\ -4 & -5 & 3 & -8 & 1 \\ 2 & -5 & -4 & 1 & 8 \\ -6 & 0 & 7 & -3 & 1 \end{bmatrix}$$

- **Solution** Since A has four rows, L should be 4×4. The first column of L is the first column of A divided by the top pivot entry:

$$L = \begin{bmatrix} 1 & 0 & 0 & 0 \\ -2 & 1 & 0 & 0 \\ 1 & & 1 & 0 \\ -3 & & & 1 \end{bmatrix}$$

© 2016 Pearson Education, Ltd.

AN LU FACTORIZATION ALGORITHM

- Compare the first columns of A and L. *The row operations that create zeros in the first column of A will also create zeros in the first column of L.*
- To make this same correspondence of row operations on A hold for the rest of L, watch a row reduction of A to an echelon form U. That is, *highlight the entries* in each matrix that are used to determine the sequence of row operations that transform A onto U.

$$A = \begin{bmatrix} 2 & 4 & -1 & 5 & -2 \\ -4 & -5 & 3 & -8 & 1 \\ 2 & -5 & -4 & 1 & 8 \\ -6 & 0 & 7 & -3 & 1 \end{bmatrix} \sim \begin{bmatrix} 2 & 4 & -1 & 5 & -2 \\ 0 & 3 & 1 & 2 & -3 \\ 0 & -9 & -3 & -4 & 10 \\ 0 & 12 & 4 & 12 & -5 \end{bmatrix} = A_1$$

$$\sim A_2 = \begin{bmatrix} 2 & 4 & -1 & 5 & -2 \\ 0 & 3 & 1 & 2 & -3 \\ 0 & 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 4 & 7 \end{bmatrix} \sim \begin{bmatrix} 2 & 4 & -1 & 5 & -2 \\ 0 & 3 & 1 & 2 & -3 \\ 0 & 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 0 & 5 \end{bmatrix} = U \qquad (5)$$

© 2016 Pearson Education, Ltd.

AN LU FACTORIZATION ALGORITHM

- The highlighted entries above determine the row reduction of A to U. At each pivot column, divide the highlighted entries by the pivot and place the result onto L:

$$\begin{array}{cccc} \begin{bmatrix} 2 \\ -4 \\ 2 \\ -6 \end{bmatrix} & \begin{bmatrix} 3 \\ -9 \\ 12 \end{bmatrix} & \begin{bmatrix} 2 \\ 4 \end{bmatrix} & \begin{bmatrix} 5 \end{bmatrix} \\ \div 2 & \div 3 & \div 2 & \div 5 \\ \downarrow & \downarrow & \downarrow & \downarrow \end{array}$$

$$\begin{bmatrix} 1 & & & \\ -2 & 1 & & \\ 1 & -3 & 1 & \\ -3 & 4 & 2 & 1 \end{bmatrix}, \quad \text{and} \quad L = \begin{bmatrix} 1 & 0 & 0 & 0 \\ -2 & 1 & 0 & 0 \\ 1 & -3 & 1 & 0 \\ -3 & 4 & 2 & 1 \end{bmatrix}$$

- An easy calculation verifies that this L and U satisfy $LU = A$.

© 2016 Pearson Education, Ltd.

2 Matrix Algebra

2.8

SUBSPACES OF $\mathbb{R}^n$

© 2016 Pearson Education, Inc.

DEFINITION OF SUBSPACE

- A **subspace** of $\mathbb{R}^n$ is any set H in $\mathbb{R}^n$ that has three properties:
  a) The zero vector is in H.
  b) For each $\mathbf{u}$ and $\mathbf{v}$ in H, the sum $\mathbf{u} + \mathbf{v}$ is in H.
  c) For each $\mathbf{u}$ in H and each scalar c, the vector $c\mathbf{u}$ is in H.

- A subspace is *closed* under addition and scalar multiplication.

© 2016 Pearson Education, Inc.

EXAMPLES OF SUBSPACE

- **Example 1**: If $\mathbf{v}_1$ and $\mathbf{v}_2$ are in $\mathbb{R}^n$ and $H = \text{Span}\{\mathbf{v}_1, \mathbf{v}_2\}$, then H is a subspace of $\mathbb{R}^n$.
- See Example 1 in the book.

FIGURE 1

Span $\{\mathbf{v}_1, \mathbf{v}_2\}$ as a plane through the origin.

© 2016 Pearson Education, Inc.

EXAMPLES OF SUBSPACE

- **Example 2**: A line L not through the origin is not a subspace, because it does not contain the origin, as required. Also, L is not closed under addition or scalar multiplication

$\mathbf{u} + \mathbf{v}$ is not on L

$2\mathbf{w}$ is not on L

© 2016 Pearson Education, Inc.

EXAMPLES OF SUBSPACE

- Note that $\mathbb{R}^n$ is a subspace since it satisfies the three properties for a subspace.
- The set containing only the zero vector in $\mathbb{R}^n$ is called the **zero subspace**, which also satisfies the three properties.

© 2016 Pearson Education, Inc.

COLUMN SPACE OF A MATRIX

- **Definition**: The **column space** of a matrix A is the set of all linear combinations of the columns of A. We call the column space of A as Col A.

- Note that the column space of an $m \times n$ matrix is a subspace of $\mathbb{R}^m$, satisfying the three properties for a subspace.
- Col A equals $\mathbb{R}^m$ only when the columns of A span $\mathbb{R}^m$. Otherwise, Col A is only part of $\mathbb{R}^m$.

- See Example 4 in the book (page 147).

© 2016 Pearson Education, Inc.

NULL SPACE OF A MATRIX

- **Definition**: The **null space** of a matrix A is the set of all solutions of the homogeneous equation $A\mathbf{x} = \mathbf{0}$. We call the null space of A as Nul A.

- **Theorem 12**: The null space of an $m \times n$ matrix is a subspace of $\mathbb{R}^n$. Equivalently, the set of all solutions of a system $A\mathbf{x} = \mathbf{0}$ of m homogeneous linear equations n unknowns is a subspace of $\mathbb{R}^n$.

- See the proof in the book.

© 2016 Pearson Education, Inc.

BASIS FOR A SUBSPACE

- **Definition**: A **basis** for a subspace H of $\mathbb{R}^n$ is a linearly independent set in H that spans H.

- **Example 5**: The columns of an invertible $n \times n$ matrix form a basis for $\mathbb{R}^n$ because they are linearly independent and span $\mathbb{R}^n$, by the Invertible Matrix Theorem.

© 2016 Pearson Education, Inc.

BASIS FOR A SUBSPACE

- One such matrix is the $n \times n$ identity matrix, whose columns are denoted by $\mathbf{e}_1, \ldots, \mathbf{e}_n$:

$$\mathbf{e}_1 = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \mathbf{e}_2 = \begin{bmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{bmatrix}, \ldots, \mathbf{e}_n = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix}$$

- This set $\{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$ is called the standard basis for $\mathbb{R}^n$.

© 2016 Pearson Education, Inc.

BASIS FOR A SUBSPACE

- **Example 6**: Find a basis for the null space of the matrix $A = \begin{bmatrix} -3 & 6 & -1 & 1 & -7 \\ 1 & -2 & 2 & 3 & -1 \\ 2 & -4 & 5 & 8 & -4 \end{bmatrix}$.
- **Solution**: The homogeneous equation $A\mathbf{x} = \mathbf{0}$ can be solved as $[A \quad \mathbf{0}] \sim \begin{bmatrix} 1 & -2 & 0 & -1 & 3 & 0 \\ 0 & 0 & 1 & 2 & -2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$.
- Thus the general solution is

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix} = \begin{bmatrix} 2x_2 + x_4 - 3x_5 \\ x_2 \\ -2x_4 + 2x_5 \\ x_4 \\ x_5 \end{bmatrix} = x_2 \begin{bmatrix} 2 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} + x_4 \begin{bmatrix} 1 \\ 0 \\ -2 \\ 1 \\ 0 \end{bmatrix} + x_5 \begin{bmatrix} -3 \\ 0 \\ 2 \\ 0 \\ 1 \end{bmatrix} = x_2\mathbf{u} + x_4\mathbf{v} + x_5\mathbf{w}.$$

© 2016 Pearson Education, Inc.

BASIS FOR A SUBSPACE

- Nul A is the set of all linear combinations of $\mathbf{u}$, $\mathbf{v}$, and $\mathbf{w}$. Also, these three vectors are linearly independent because $x_2\mathbf{u} + x_4\mathbf{v} + x_5\mathbf{w} = \mathbf{0}$ only if $x_2 = x_4 = x_5 = 0$ (examine entries 2, 4, and 5 in the lefthand-side of the vector equation. Thus, $\{\mathbf{u}, \mathbf{v}, \mathbf{w}\}$ is a basis for Nul A.

© 2016 Pearson Education, Inc.

BASIS FOR A SUBSPACE

- **Example 7**: Find a basis for the column space of the matrix $B = \begin{bmatrix} 1 & 0 & -3 & 5 & 0 \\ 0 & 1 & 2 & -1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$.
- **Solution**: Denote the columns of B by $\mathbf{b}_1, \ldots, \mathbf{b}_5$. Note that $\mathbf{b}_3 = -3\mathbf{b}_1 + 2\mathbf{b}_2$ and $\mathbf{b}_4 = 5\mathbf{b}_1 - \mathbf{b}_2$. Thus, any linear combination $\mathbf{b}_1, \ldots, \mathbf{b}_5$ is actually just a combination of $\mathbf{b}_1, \mathbf{b}_2$, and $\mathbf{b}_5$. Thus, $\{\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_5\}$ is a basis for Col B. In general, the pivot columns of B form a basis for Col B.

© 2016 Pearson Education, Inc.

BASIS FOR A SUBSPACE

- **Example 8**: It can be verified that the matrix Find a basis for the column space of the matrix $A =$ $[\mathbf{a}_1 \quad \mathbf{a}_2 \quad \cdots \quad \mathbf{a}_1] = \begin{bmatrix} 1 & 3 & 3 & 2 & -9 \\ -2 & -2 & 2 & -8 & 2 \\ 2 & 3 & 0 & 7 & 1 \\ 3 & 4 & -1 & 11 & -8 \end{bmatrix}$ is row equivalent to the matrix B in Example 7. Find a basis for Col A.

© 2016 Pearson Education, Inc.

BASIS FOR A SUBSPACE

- **Solution**: A and B are row equivalent, so the solution set is the same for $A\mathbf{x} = \mathbf{0}$ and $B\mathbf{x} = \mathbf{0}$. Thus, the following still holds true: $\mathbf{a}_3 = -3\mathbf{a}_1 + 2\mathbf{a}_2$ and $\mathbf{a}_4 = 5\mathbf{a}_1 - \mathbf{a}_2$. Also $\{\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_5\}$ is linearly independent since any dependence relation among $\mathbf{a}_1, \mathbf{a}_2$, and $\mathbf{a}_5$ would imply the same dependence relation among $\mathbf{b}_1, \mathbf{b}_2$, and $\mathbf{b}_5$. Since $\{\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_5\}$ is linearly independent, $\{\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_5\}$ is also linearly independent and hence is a basis for Col A.

© 2016 Pearson Education, Inc.

BASIS FOR A SUBSPACE

- **Theorem 13**: The pivot columns of a matrix A form a basis for the column space of A.

- **Warning**: Be careful to use pivot columns of A itself for the basis of Col A. The columns of an echelon form B are often not in the column space of A. (For instance, in Example 7 and 8, the columns of B all have zeros in their last entries and cannot generate the columns of A.)

© 2016 Pearson Education, Inc.

2 Matrix Algebra

2.9

DIMENSION AND RANK

PEARSON

© 2016 Pearson Education, Inc.

COORDINATE SYSTEMS

- Suppose $\beta = \{b_1, \ldots, bp\}$ is a basis for H, and suppose a vector $\mathbf{x}$ in H can be generated in two ways, say,

  $x = c_1 b_1 + \cdots + c_p b_p$ and $x = d_1 b_1 + \cdots + d_p b_p$ (1)

- Then, subtracting gives

  $0 = x - x = (c_1 - d_1)b_1 + \cdots + (c_p - dp)b_p$ (2)

- Since β is linearly independent, the weights in (2) must all be zero. That is, $c_j = d_j$ for $1 \leq j \leq p$, which shows that the two representations in (1) are actually the same.

© 2016 Pearson Education, Inc.

COORDINATE SYSTEMS

- **Definition**: Suppose the set $\beta = \{b_1, \dots, bp\}$ is a basis for a subspace *H*. For each x in *H*, the **coordinates of x relative to the basis** β are the weights $c_1, \dots, c_p$ such that $x = c_1 b_1 + \cdots + c_p b_p$, and the vector in $\mathbb{R}^p$

$$[x]_\beta = \begin{bmatrix} c_1 \\ . \\ . \\ . \\ c_p \end{bmatrix}$$

- is called the **coordinate vector of x** (**relative to** β) or the β**-coordinate vector of x**.

© 2016 Pearson Education, Inc.

COORDINATE SYSTEMS

- **Example 1** Let $v_1 = \begin{bmatrix} 3 \\ 6 \\ 2 \end{bmatrix}$, $v_2 = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}$, $x = \begin{bmatrix} 3 \\ 12 \\ 7 \end{bmatrix}$, and $\beta = \{v_1, v_2\}$. Then β is a basis for H = Span $\{v_1, v_2\}$ because v_1 and v_2 are linearly independent. Determine if x is in *H*, and if it is, find the coordinate vector of x relative to β.
- **Solution** If x is in *H*, then the following vector equation is consistent:

$$c_1 \begin{bmatrix} 3 \\ 6 \\ 2 \end{bmatrix} + c_2 = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 3 \\ 12 \\ 7 \end{bmatrix}$$

© 2016 Pearson Education, Inc.

COORDINATE SYSTEMS

- The scalars c_1 and c_2, if they exist, are the β-coordinates of **x**. Row operations show that

$$\begin{bmatrix} 3 & -1 & 3 \\ 6 & 0 & 12 \\ 2 & 1 & 7 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & 2 \\ 0 & 1 & 3 \\ 0 & 0 & 0 \end{bmatrix}$$

- Thus $c_1 = 2$, $c_2 = 3$ and $[x]_\beta = \begin{bmatrix} 2 \\ 3 \end{bmatrix}$. The basis β determines a "coordinate system" on H, which can be visualized by the grid shown in Fig. 1 on the next slide.

© 2016 Pearson Education, Inc.

COORDINATE SYSTEMS

FIGURE 1 A coordinate system on a plane H in $\mathbb{R}^3$.

© 2016 Pearson Education, Inc.

THE DIMESION OF A SUBSPACE

- **Definition**: The **dimension** of a nonzero subspace H, denoted by dim H, is the number of vectors in any basis for H. The dimension of the zero subspace $\{0\}$ is defined to be zero.

- **Definition:** The **rank** of a matrix A, denoted by rank A, is the dimension of the column space of A.

© 2016 Pearson Education, Inc.

THE DIMESION OF A SUBSPACE

- **Example 3** Determine the rank of the matrix

$$A \sim \begin{bmatrix} 2 & 5 & -3 & -4 & 8 \\ 0 & -3 & 2 & 5 & -7 \\ 0 & -6 & 4 & 14 & -20 \\ 0 & -9 & 6 & 5 & -6 \end{bmatrix}$$

- **Solution** Reduce A to echelon form:

$$A \sim \begin{bmatrix} 2 & 5 & -3 & -4 & 8 \\ 0 & -3 & 2 & 5 & -7 \\ 0 & -6 & 4 & 14 & -20 \\ 0 & -9 & 6 & 5 & -6 \end{bmatrix} \sim \cdots \sim \begin{bmatrix} 2 & 5 & -3 & -4 & 8 \\ 0 & -3 & 2 & 5 & -7 \\ 0 & 0 & 0 & 4 & -6 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Pivot columns

- The matrix A has 3 pivot columns, so rank $A = 3$.

© 2016 Pearson Education, Inc.

THE DIMESION OF A SUBSPACE

- **Theorem 14** If a matrix A has n columns, then rank A + dim Nul$A$ = n.

- **Theorem 15** Let H be a p-dimensional subspace of $\mathbb{R}^n$. Any linearly independent set of exactly p elements in H is automatically a basis for H. Also, any set of p elements of H that spans H is automatically a basis for H.

© 2016 Pearson Education, Inc.

RANK AND THE INVERTIBLE MATRIX THEOREM

- **The Invertible Theorem (continued)** Let A be an $n \times n$ matrix. Then the following statements are each equivalent to the statement that A is an invertible matrix.

  m. The columns of A form a basis of $\mathbb{R}^n$.

  n. Col $A = \mathbb{R}^n$

  o. dim Col $A = n$

  p. rank $A = n$

  q. Nul $A = \{0\}$

  r. dim Nul $A = 0$

© 2016 Pearson Education, Inc.

RANK AND THE INVERTIBLE MATRIX THEOREM

- **Proof** Statement (m) is logically equivalent to statements (e) and (h) regarding linear independence and spanning. The other five statements are linked to the earlier ones of the theorem by the following chain of almost trivial implications:

$$(g) \Longrightarrow (n) \Rightarrow (o) \Rightarrow (p) \Rightarrow (r) \Rightarrow (q) \Rightarrow (d)$$

- Statement (g), which says that the equation Ax = b has at least one solution for each **b** in $\mathbb{R}^n$, implies statement (n), because Col A is precisely the set of all **b** such that the equation Ax = b is consistent.

© 2016 Pearson Education, Inc.

RANK AND THE INVERTIBLE MATRIX THEOREM

- The implications $(n) \Rightarrow (o) \Rightarrow (p)$ follow from the definitions of *dimension* and *rank*.
- If the rank of A is n, the number of columns of A, then dim Nul $A = 0$, by the Rank Theorem, and so Nul A $\{0\}$. Thus $(p) \Rightarrow (r) \Rightarrow (q)$.
- Also, statement (q) implies that the equation Ax = 0 has only the trivial solution, which is statement (d).
- Since statements (d) and (g) are already known to be equivalent to the statement that A is invertible, the proof is complete.

© 2016 Pearson Education, Inc.

3 Determinants

3.1

INTRO TO DETERMINANTS

PEARSON

© 2012 Pearson Education, Inc.

DETERMINANTS OF A 2×2 MATRIX

- **Recall**: A 2×2 matrix is **invertible** if and only if its determinant is nonzero.
- If the matrix is invertible, then the inverse can be found as follows:

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

$$\det A = ad - bc.$$

- How can we find the determinant of larger square matrices?

© 2012 Pearson Education, Inc.

DETERMINANTS OF LARGER SQUARE MATRICES

- Consider $A = [a_{ij}]$ with $a_{ij} \neq 0$.

$$A \sim \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{11}a_{21} & a_{11}a_{22} & a_{11}a_{23} \\ a_{11}a_{31} & a_{11}a_{32} & a_{11}a_{33} \end{bmatrix}$$

$$\sim \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ 0 & a_{11}a_{22} - a_{12}a_{21} & a_{11}a_{23} - a_{13}a_{21} \\ 0 & a_{11}a_{32} - a_{12}a_{31} & a_{11}a_{33} - a_{13}a_{31} \end{bmatrix}$$

$$\sim \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ 0 & a_{11}a_{22} - a_{12}a_{21} & a_{11}a_{23} - a_{13}a_{21} \\ 0 & 0 & a_{11}\Delta \end{bmatrix}$$

- where $\Delta = a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32}$
$$-a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33} - a_{13}a_{22}a_{31}$$
- Since A is invertible, Δ must be nonzero.
- We call Δ the determinant of the 3×3 matrix A.

© 2012 Pearson Education, Inc.

DETERMINANTS OF LARGER SQUARE MATRICES

- Recall that $\det A = a_{11}a_{22} - a_{12}a_{21}$ for a 2×2 matrix.
- For a 3×3 matrix from the previous slide,

$$\begin{aligned}\Delta &= a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} \\ &\quad -a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33} - a_{13}a_{22}a_{31} \\ &= (a_{11}a_{22}a_{33} - a_{11}a_{23}a_{32}) - (a_{12}a_{21}a_{33} - a_{12}a_{23}a_{31}) \\ &\quad +(a_{13}a_{21}a_{32} - a_{13}a_{22}a_{31}) \\ &= a_{11}\det\begin{bmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{bmatrix} - a_{12}\det\begin{bmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{bmatrix} + a_{13}\det\begin{bmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{bmatrix} \\ &= a_{11}\det A_{11} - a_{12}\det A_{12} + a_{13}\det A_{13}\end{aligned}$$

- Where A_{11}, A_{12}, and A_{13} are obtained from A by deleting the first row and one of the three columns.

© 2012 Pearson Education, Inc.

DETERMINANTS OF LARGER SQUARE MATRICES

- For any square matrix A, let A_{ij} denote the submatrix formed by deleting the ith row and jth column of A. For instance,

$$A = \begin{bmatrix} 1 & -2 & 5 & 0 \\ 2 & 0 & 4 & -1 \\ 3 & 1 & 0 & 7 \\ 0 & 4 & -2 & 0 \end{bmatrix} \Rightarrow A_{32} = \begin{bmatrix} 1 & 5 & 0 \\ 2 & 4 & -1 \\ 0 & -2 & 0 \end{bmatrix}$$

- We can now give a recursive definition of a determinant. When $n = 3$, $\det A$ is defined using determinants of the 2×2 submatrices A_{ij}. When $n = 4$, $\det A$ uses determinants of the 3×3 submatrices A_{ij}. In general, $n \times n$ determinant is defined by determinants of $(n-1) \times (n-1)$ submatrices.

© 2012 Pearson Education, Inc.

DEFINITION OF DETERMINANTS

- **Definition**: For a 1×1 matrix $A = [a_{11}]$, we define $\det A = a_{11}$. For $n \geq 2$, the determinant of an $n \times n$ matrix $A = [a_{ij}]$ is the sum of n terms of the form $\pm\, a_{1j} \det A_{1j}$, with plus and minus signs alternating, where the entries $a_{11}, a_{12}, \ldots, a_{1n}$ are from the first row of A. In symbols,

$$\begin{aligned}\det A &= a_{11} \det A_{11} - a_{12} \det A_{12} + \cdots + (-1)^{1+n} a_{1n} \det A_{1n} \\ &= \textstyle\sum_{j=1}^{n} (-1)^{1+j} a_{1j} \det A_{1j}\end{aligned}$$

© 2012 Pearson Education, Inc.

DETERMINANTS EXAMPLE

- **Example**: Compute the determinant of

$$A = \begin{bmatrix} 1 & 5 & 0 \\ 2 & 4 & -1 \\ 0 & -2 & 0 \end{bmatrix}$$

- **Solution**:

$$\begin{aligned} \det A &= a_{11} \det A_{11} - a_{12} \det A_{12} + a_{13} \det A_{13} \\ &= 1 \cdot \det \begin{bmatrix} 4 & -1 \\ -2 & 0 \end{bmatrix} - 5 \cdot \det \begin{bmatrix} 2 & -1 \\ 0 & 0 \end{bmatrix} + 0 \cdot \det \begin{bmatrix} 2 & 4 \\ 0 & -2 \end{bmatrix} \\ &= 1 \cdot (0 - 2) - 5 \cdot (0 - 0) + 0 \cdot (-4 - 0) = -2 \end{aligned}$$

- We can rewrite

$$\det A = 1 \cdot \begin{vmatrix} 4 & -1 \\ -2 & 0 \end{vmatrix} - 5 \cdot \begin{vmatrix} 2 & -1 \\ 0 & 0 \end{vmatrix} + 0 \cdot \begin{vmatrix} 2 & 4 \\ 0 & -2 \end{vmatrix}$$

© 2012 Pearson Education, Inc.

DEFINITION OF COFACTOR

- Given $A = [a_{ij}]$, the (i, j)-cofactor of A is the number C_{ij} given by

$$C_{ij} = (-1)^{i+j} \det A_{ij}$$

- Then

$$\det A = a_{11}C_{11} + a_{12}C_{12} + \cdots + a_{1n}C_{1n}$$

- This formula is called a cofactor expansion across the first row of A.

© 2012 Pearson Education, Inc.

COFACTOR EXPANSION FOR DETERMINANTS

- **Theorem 1**: The determinant of an $n \times n$ matrix A can be computed by a cofactor expansion across any row or down any column. The expansion across the ith row using the cofactors is

$$\det A = a_{i1}C_{i1} + a_{i2}C_{i2} + \cdots + a_{in}C_{in}$$

- The cofactor expansion down the jth column is

$$\det A = a_{1j}C_{1j} + a_{2j}C_{2j} + \cdots + a_{nj}C_{nj}$$

- We skip the proof.

© 2012 Pearson Education, Inc.

SIGNS OF COFACTOR

- The plus or minus sign in the (i, j)-cofactor depends on the position of a_{ij} in the matrix regardless of the sign of a_{ij} itself. That is, the factor $(-1)^{i+j}$ determines the following checkerboard pattern of signs:

$$\begin{bmatrix} + & - & + & \cdots \\ - & + & - & \\ + & - & + & \\ \vdots & & & \ddots \end{bmatrix}$$

© 2012 Pearson Education, Inc.

DETERMINANTS EXAMPLE

- **Example**: Use a cofactor expansion across the third row to compute $\det A$, where

$$A = \begin{bmatrix} 1 & 5 & 0 \\ 2 & 4 & -1 \\ 0 & -2 & 0 \end{bmatrix}$$

- **Solution**:

$$\det A = a_{31}C_{31} + a_{32}C_{32} + a_{33}C_{33}$$
$$= (-1)^{3+1}a_{31}\det A_{31} + (-1)^{3+2}a_{32}\det A_{32}$$
$$+(-1)^{3+3}a_{33}\det A_{33}$$

$$= 0 \cdot \begin{vmatrix} 5 & 0 \\ 4 & -1 \end{vmatrix} - (-2) \cdot \begin{vmatrix} 1 & 0 \\ 2 & -1 \end{vmatrix} + 0 \cdot \begin{vmatrix} 1 & 5 \\ 2 & 4 \end{vmatrix}$$

$$= 0 + 2 \cdot (-1) + 0 = -2$$

© 2012 Pearson Education, Inc.

DETERMINANTS EXAMPLE

- **Example**: Compute $\det A$, where

$$A = \begin{bmatrix} 3 & -7 & 8 & 9 & -6 \\ 0 & 2 & -5 & 7 & 3 \\ 0 & 0 & 1 & 5 & 0 \\ 0 & 0 & 2 & 4 & -1 \\ 0 & 0 & 0 & -2 & 0 \end{bmatrix}$$

- **Solution**: The simplest method is to compute the cofactor expansion down the first column of A. Thus,

$$\det A = 3 \cdot \begin{vmatrix} 2 & -5 & 7 & 3 \\ 0 & 1 & 5 & 0 \\ 0 & 2 & 4 & -1 \\ 0 & 0 & -2 & 0 \end{vmatrix} - 0 \cdot C_{21} + 0 \cdot C_{31} - 0 \cdot C_{41} + 0 \cdot C_{51}$$

© 2012 Pearson Education, Inc.

DETERMINANTS EXAMPLE

- Similarly, using the cofactor expansion down the first column of the resulting 4×4 matrix, we obtain

$$\det A = 3 \cdot \begin{vmatrix} 2 & -5 & 7 & 3 \\ 0 & 1 & 5 & 0 \\ 0 & 2 & 4 & -1 \\ 0 & 0 & -2 & 0 \end{vmatrix} = 3 \cdot 2 \cdot \begin{vmatrix} 1 & 5 & 0 \\ 2 & 4 & -1 \\ 0 & -2 & 0 \end{vmatrix}$$

- Similarly, using the cofactor expansion across the third row of the resulting 3×3 matrix, we obtain

$$\det A = 3 \cdot 2 \cdot \begin{vmatrix} 1 & 5 & 0 \\ 2 & 4 & -1 \\ 0 & -2 & 0 \end{vmatrix} = 3 \cdot 2 \cdot 2 \cdot \begin{vmatrix} 1 & 0 \\ 2 & -1 \end{vmatrix} = -12$$

© 2012 Pearson Education, Inc.

DETERMINANTS OF TRIANGULAR MATRICES

- **Theorem 2**: If A is a triangular matrix, then $\det A$ is the product of the entries on the main diagonal of A.

- We skip the proof, but you can use the same strategy from the previous example, to understand this theorem.

© 2012 Pearson Education, Inc.

3 Determinants

3.2

PROPERTIES OF DETERMINANTS

PEARSON

© 2012 Pearson Education, Inc.

ROW OPERATIONS AND DETERMINANTS

- **Theorem 3 (Row Operations)**: Let A be a square matrix.

  a. If a multiple of one row of A is added to another row to produce a matrix B, then $\det B = \det A$.

  b. If two rows of A are interchanged to produce B, then $\det B = -\det A$.

  c. If one row of A is multiplied by k, then $\det B = k \cdot \det A$.

© 2012 Pearson Education, Inc.

ROW OPERATIONS AND DETERMINANTS

- **Example 1**: Compute $\det A$, where $A = \begin{bmatrix} 1 & -4 & 2 \\ -2 & 8 & -9 \\ -1 & 7 & 0 \end{bmatrix}$.
- **Solution**:

$$\det A = \begin{vmatrix} 1 & -4 & 2 \\ -2 & 8 & -9 \\ -1 & 7 & 0 \end{vmatrix} = \begin{vmatrix} 1 & -4 & 2 \\ 0 & 0 & -5 \\ -1 & 7 & 0 \end{vmatrix} = \begin{vmatrix} 1 & -4 & 2 \\ 0 & 0 & -5 \\ 0 & 3 & 2 \end{vmatrix}$$

$$= -\begin{vmatrix} 1 & -4 & 2 \\ 0 & 3 & 2 \\ 0 & 0 & -5 \end{vmatrix} = -1(1)(3)(-5) = 15$$

© 2012 Pearson Education, Inc.

ROW OPERATIONS AND DETERMINANTS

- A common use of Theorem 3(c) is to factor out a common multiple of one row of a matrix. For instance,

$$\begin{vmatrix} * & * & * \\ 5k & -2k & 3k \\ * & * & * \end{vmatrix} = k \begin{vmatrix} * & * & * \\ 5 & -2 & 3 \\ * & * & * \end{vmatrix}$$

© 2012 Pearson Education, Inc.

ROW OPERATIONS AND DETERMINANTS

- **Example 2**: Compute $\det A$, where $A = \begin{bmatrix} 2 & -8 & 6 & 8 \\ 3 & -9 & 5 & 20 \\ -3 & 0 & 1 & -2 \\ 1 & -4 & 0 & 6 \end{bmatrix}$.
- **Solution**:

$$\det A = 2\begin{vmatrix} 1 & -4 & 3 & 4 \\ 3 & -9 & 5 & 20 \\ -3 & 0 & 1 & -2 \\ 1 & -4 & 0 & 6 \end{vmatrix} = 2\begin{vmatrix} 1 & -4 & 3 & 4 \\ 0 & 3 & -4 & -2 \\ 0 & -12 & 10 & 10 \\ 0 & 0 & -3 & 2 \end{vmatrix}$$

$$= 2\begin{vmatrix} 1 & -4 & 3 & 4 \\ 0 & 3 & -4 & -2 \\ 0 & 0 & -6 & 2 \\ 0 & 0 & -3 & 2 \end{vmatrix} = 2\begin{vmatrix} 1 & -4 & 3 & 4 \\ 0 & 3 & -4 & -2 \\ 0 & 0 & -6 & 2 \\ 0 & 0 & 0 & 1 \end{vmatrix}$$

$$= 2(1)(3)(-6)(1) = -36$$

© 2012 Pearson Education, Inc.

DETERMINANTS USING ECHELON FORM

- Suppose a square matrix A has been reduced to an echelon form U by row replacement and row interchanges. (This is always possible since we do not need row scaling to obtain an echelon form.)
- If there are r interchanges, then Theorem 3 shows that

$$\det A = (-1)^r \det U.$$

$$U = \begin{bmatrix} \blacksquare & * & * & * \\ 0 & \blacksquare & * & * \\ 0 & 0 & \blacksquare & * \\ 0 & 0 & 0 & \blacksquare \end{bmatrix} \qquad U = \begin{bmatrix} \blacksquare & * & * & * \\ 0 & \blacksquare & * & * \\ 0 & 0 & 0 & \blacksquare \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

$$\det U \neq 0 \qquad \det U = 0$$

© 2012 Pearson Education, Inc.

DETERMINANTS USING ECHELON FORM

- As seen in the previous slide, since U is triangular, we obtain

$$\det U = \begin{cases} (-1)^r \begin{pmatrix} \text{products of} \\ \text{pivots in } U \end{pmatrix} & \text{when } A \text{ is invertible} \\ 0 & \text{when } A \text{ is not invertible} \end{cases}$$

- **Theorem 4**: A square matrix A is invertible if and only if $\det A \neq 0$.
- This theorem is proved by using the previous slide. This theorem adds the statement "$\det A \neq 0$" to the Invertible Matrix Theorem.

© 2012 Pearson Education, Inc.

DETERMINANTS AND INVERTIBILITY

- A useful corollary is that $\det A = 0$ when the columns of A are linearly dependent. Also, $\det A = 0$ when the rows of A are linearly dependent. (Rows of A are columns of A^T, and linearly dependent columns of A^T make A^T singular. When A^T is singular, so is A, by the Invertible Matrix Theorem.)

© 2012 Pearson Education, Inc.

ROW OPERATIONS AND DETERMINANTS

- **Example 3**: Compute $\det A$, where $A = \begin{bmatrix} 3 & -1 & 2 & -5 \\ 0 & 5 & -3 & -6 \\ -6 & 7 & -7 & 4 \\ -5 & -8 & 0 & 9 \end{bmatrix}$.
- **Solution**: Add 2 times row 1 to row 3 to obtain

$$\det A = \begin{vmatrix} 3 & -1 & 2 & -5 \\ 0 & 5 & -3 & -6 \\ 0 & 5 & -3 & -6 \\ -5 & -8 & 0 & 9 \end{vmatrix} = 0$$

- Because the second and third rows of this matrix are equal, which means the columns are linearly dependent.

© 2012 Pearson Education, Inc.

ROW OPERATIONS AND DETERMINANTS

- **Example 4**: Compute $\det A$, where $A = \begin{bmatrix} 0 & 1 & 2 & -1 \\ 2 & 5 & -7 & 3 \\ 0 & 3 & 6 & 2 \\ -2 & -5 & 4 & 2 \end{bmatrix}$.
- **Solution**:

$$\det A = \begin{vmatrix} 0 & 1 & 2 & -1 \\ 2 & 5 & -7 & 3 \\ 0 & 3 & 6 & 2 \\ 0 & 0 & -3 & 1 \end{vmatrix} = -2 \begin{vmatrix} 1 & 2 & -1 \\ 3 & 6 & 2 \\ 0 & -3 & 1 \end{vmatrix}$$

$$= -2 \begin{vmatrix} 1 & 2 & -1 \\ 0 & 0 & 5 \\ 0 & -3 & 1 \end{vmatrix} = (-2)(1) \begin{vmatrix} 0 & 5 \\ -3 & 1 \end{vmatrix} = (-2)(15) = 30$$

© 2012 Pearson Education, Inc.

DETERMINANTS AND INVERTIBILITY

- We can perform operations on the columns of a matrix in a way that is analogous to the row operations. The next theorem shows that column operations have the same effects on determinants as row operations.
- **Theorem 5**: if A is $n \times n$ matrix, then $\det A^T = \det A$.
- The proof is done via induction.

- Because of Theorem 5, each statement in Theorem 3 is true when the word row is replaced everywhere by column.

© 2012 Pearson Education, Inc.

MULTIPLICATIVE PROPERTY

- **Theorem 6**: if A and B are A is $n \times n$ matrix, then $\det AB = (\det A)(\det B)$.
- We skip the proof, but the proof is done by utilizing Theorem 3.
- **Example 5**: Verify Theorem 6 for $A = \begin{bmatrix} 6 & 1 \\ 3 & 2 \end{bmatrix}$ and $B = \begin{bmatrix} 4 & 3 \\ 1 & 2 \end{bmatrix}$.
- **Solution**: $AB = \begin{bmatrix} 6 & 1 \\ 3 & 2 \end{bmatrix}\begin{bmatrix} 4 & 3 \\ 1 & 2 \end{bmatrix} = \begin{bmatrix} 25 & 20 \\ 14 & 13 \end{bmatrix}$ and $\det AB = 25 \cdot 13 - 20 \cdot 14 = 45$. Since $\det A = 9$ and $\det B = 5$, $\det AB = (\det A)(\det B)$.

© 2012 Pearson Education, Inc.

3 Determinants

3.3

CRAMER'S RULE, VOLUME, AND LINEAR TRANSFORMATIONS

PEARSON
© 2012 Pearson Education, Inc.

Cramer's Rule

- For any $m \times n$ matrix A and any $\mathbf{b}$ in $\mathbb{R}^n$, let $A_i(\mathbf{b})$ be the matrix obtained from A by replacing column i by the vector $\mathbf{b}$.

$$A_i(\mathbf{b}) = [\mathbf{a}_1 \quad \cdots \quad \underset{\substack{\uparrow \\ \text{col } i}}{\mathbf{b}} \quad \cdots \quad \mathbf{a}_n]$$

- **THEOREM 7** **<u>Cramer's Rule</u>**

  Let A be an invertible $n \times n$ matrix. For any $\mathbf{b}$ in $\mathbb{R}^n$, the unique solution $\mathbf{x}$ of $A\mathbf{x} = \mathbf{b}$ has entries given by

$$x_i = \frac{\det A_i(\mathbf{b})}{\det A}, \qquad i = 1, 2, \ldots, n \tag{1}$$

© 2012 Pearson Education, Inc.

Cramer's Rule

- PROOF Denote the columns of A by $\mathbf{a}_1, \ldots, \mathbf{a}_n$ and the columns of the $n \times n$ identity matrix I by $\mathbf{e}_1, \ldots, \mathbf{e}_n$.

  If $A\mathbf{x} = \mathbf{b}$, the definition of matrix multiplication shows that

$$
\begin{aligned}
A \cdot I_i(\mathbf{x}) &= A[\ \mathbf{e}_1 \ \cdots \ \mathbf{x} \ \cdots \ \mathbf{e}_n\ ] \\
&= [\ A\mathbf{e}_1 \ \cdots \ A\mathbf{x} \ \cdots \ A\mathbf{e}_n\ ] \\
&= [\mathbf{a}_1 \ \cdots \ \mathbf{b} \ \cdots \ \mathbf{a}_n\ ] = A_i(\mathbf{b})
\end{aligned}
$$

© 2012 Pearson Education, Inc.

Cramer's Rule

- By the multiplicative property of determinants,

$$(\det A)(\det I_i(\mathbf{x})) = \det A_i(\mathbf{b})$$

- The second determinant on the left is simply x_i . (Make a cofactor expansion along the ith row.)
- Hence $(\det A)\cdot x_i = \det A_i(\mathbf{b})$.
- This proves (1) because A is invertible and $\det A \neq 0$.

© 2012 Pearson Education, Inc.

Cramer's Rule

- **EXAMPLE 1** Use Cramer's rule to solve the system

$$3x_1 - 2x_2 = 6$$
$$-5x_1 + 4x_2 = 8$$

- **Solution** View the system as $A\mathbf{x} = \mathbf{b}$.

Using the notation introduced above,

$$A = \begin{bmatrix} 3 & -2 \\ -5 & 4 \end{bmatrix}, \quad A_1(\mathbf{b}) = \begin{bmatrix} 6 & -2 \\ 8 & 4 \end{bmatrix}, \quad A_2(\mathbf{b}) = \begin{bmatrix} 3 & 6 \\ -5 & 8 \end{bmatrix}$$

© 2012 Pearson Education, Inc.

Cramer's Rule

- Since det $A = 2$, the system has a unique solution.
- By Cramer's rule,

$$x_1 = \frac{\det A_1(\mathrm{b})}{\det A} = \frac{24+16}{2} = 20$$

$$x_2 = \frac{\det A_2(\mathrm{b})}{\det A} = \frac{24+30}{2} = 27$$

© 2012 Pearson Education, Inc.

Application to Engineering

- **EXAMPLE 2** Consider the following system in which s is an unspecified parameter. Determine the values of s for which the system has a unique solution, and use Cramer's rule to describe the solution.

$$3sx_1 - 2x_2 = 4$$
$$-6x_1 + sx_2 = 1$$

© 2012 Pearson Education, Inc.

Application to Engineering

- **Solution** View the system as $A\mathbf{x} = \mathbf{b}$.
- Then

$$A = \begin{bmatrix} 3s & -2 \\ -6 & s \end{bmatrix}, \qquad A_1(\mathbf{b}) = \begin{bmatrix} 4 & -2 \\ 1 & s \end{bmatrix}, \qquad A_2(\mathbf{b}) = \begin{bmatrix} 3s & 4 \\ -6 & 1 \end{bmatrix}$$

- Since

$$\det A = 3s^2 - 12 = 3(s+2)(s-2)$$

- the system has a unique solution precisely when $s \neq \pm 2$.

© 2012 Pearson Education, Inc.

Application to Engineering

- For such an s, the solution is (x_1, x_2), where

$$x_1 = \frac{\det A_1(\mathrm{b})}{\det A} = \frac{4s+2}{3(s+2)(s-2)}$$

$$x_2 = \frac{\det A_2(\mathrm{b})}{\det A} = \frac{3s+24}{3(s+2)(s-2)} = \frac{s+8}{(s+2)(s-2)}$$

© 2012 Pearson Education, Inc.

A Formula for A^{-1}

- Given $A = [a_{ij}]$, the **(i, j)-cofactor** of A is the number C_{ij} given by

$$C_{ij} = (-1)^{i+j} \det A_{ij}$$

- Then

$$A^{-1} = \frac{1}{\det A} \begin{bmatrix} C_{11} & C_{21} & \cdots & C_{n1} \\ C_{12} & C_{22} & \cdots & C_{n2} \\ \vdots & \vdots & & \vdots \\ C_{1n} & C_{2n} & \cdots & C_{nn} \end{bmatrix} \tag{4}$$

- The matrix of cofactors on the right side of (4) is called the **adjugate** (or **classical adjoint**) of A, denoted by adj A.

© 2012 Pearson Education, Inc.

A Formula for A^{-1}

- **THEOREM 8** **<u>An Inverse Formula</u>**

  Let A be an invertible $n \times n$ matrix. Then

$$A^{-1} = \frac{1}{\det A} \operatorname{adj} A$$

- **EXAMPLE 3**

  Find the inverse of the matrix $A = \begin{bmatrix} 2 & 1 & 3 \\ 1 & -1 & 1 \\ 1 & 4 & -2 \end{bmatrix}$.

© 2012 Pearson Education, Inc.

A Formula for A^{-1}

- **Solution** The nine cofactors are

$$C_{11} = +\begin{vmatrix} -1 & 1 \\ 4 & -2 \end{vmatrix} = -2, \quad C_{12} = -\begin{vmatrix} 1 & 1 \\ 1 & -2 \end{vmatrix} = 3, \quad C_{13} = +\begin{vmatrix} 1 & -1 \\ 1 & 4 \end{vmatrix} = 5$$

$$C_{21} = -\begin{vmatrix} 1 & 3 \\ 4 & -2 \end{vmatrix} = 14, \quad C_{22} = +\begin{vmatrix} 2 & 3 \\ 1 & -2 \end{vmatrix} = -7, \quad C_{23} = -\begin{vmatrix} 2 & 1 \\ 1 & 4 \end{vmatrix} = -7$$

$$C_{31} = +\begin{vmatrix} 1 & 3 \\ -1 & 1 \end{vmatrix} = 4, \quad C_{32} = -\begin{vmatrix} 2 & 3 \\ 1 & 1 \end{vmatrix} = 1, \quad C_{33} = +\begin{vmatrix} 2 & 1 \\ 1 & -1 \end{vmatrix} = -3$$

© 2012 Pearson Education, Inc.

A Formula for A^{-1}

- The adjugate matrix is the *transpose* of the matrix of cofactors. [For instance, C_{12} goes in the (2, 1) position.] Thus

$$\operatorname{adj} A = \begin{bmatrix} -2 & 14 & 4 \\ 3 & -7 & 1 \\ 5 & -7 & -3 \end{bmatrix}$$

- We could compute det A directly, but the following computation provides a check on the calculations above *and* produces det A:

© 2012 Pearson Education, Inc.

A Formula for A^{-1}

$$(\text{adj } A)\cdot A = \begin{bmatrix} -2 & 14 & 4 \\ 3 & -7 & 1 \\ 5 & -7 & -3 \end{bmatrix}\begin{bmatrix} 2 & 1 & 3 \\ 1 & -1 & 1 \\ 1 & 4 & -2 \end{bmatrix} = \begin{bmatrix} 14 & 0 & 0 \\ 0 & 14 & 0 \\ 0 & 0 & 14 \end{bmatrix} = 14I$$

- Since $(\text{adj } A)A = 14I$, Theorem 8 shows that det A = 14 and

$$A^{-1} = \frac{1}{14}\begin{bmatrix} -2 & 14 & 4 \\ 3 & -7 & 1 \\ 5 & -7 & -3 \end{bmatrix} = \begin{bmatrix} -1/7 & 1 & 2/7 \\ 3/14 & -1/2 & 1/14 \\ 5/14 & -1/2 & -3/14 \end{bmatrix}$$

© 2012 Pearson Education, Inc.

Determinants as Area or Volume

- **THEOREM 9**
- If A is a 2×2 matrix, the area of the parallelogram determined by the columns of A is $|\det A|$.
- **PROOF**
- The theorem is obviously true for any 2×2 diagonal matrix:

$$\left|\det\begin{bmatrix} a & 0 \\ 0 & d \end{bmatrix}\right| = |ad| = \left\{\begin{matrix} \text{area of} \\ \text{rectangle} \end{matrix}\right\}$$

© 2012 Pearson Education, Inc.

Determinants as Area or Volume

FIGURE 1

Area $= |ad|$.

© 2012 Pearson Education, Inc.

Determinants as Area or Volume

- Let $\mathbf{a}_1$ and $\mathbf{a}_2$ be nonzero vectors. Then for any scalar c, the area of the parallelogram determined by $\mathbf{a}_1$ and $\mathbf{a}_2$ equals the area of the parallelogram determined by $\mathbf{a}_1$ and $\mathbf{a}_2 + c\mathbf{a}_1$.

FIGURE 2 Two parallelograms of equal area.

© 2012 Pearson Education, Inc.

Determinants as Area or Volume

- The proof for $\mathbb{R}^3$ is similar. The theorem is obviously true for a 3×3 diagonal matrix. See Fig 3, and any 3×3 matrix A can be transformed into a diagonal matrix using column operations that do not change $|\det A|$

FIGURE 3
Volume $= |abc|$.

© 2012 Pearson Education, Inc.

Determinants as Area or Volume

A parallelepiped is shown in Fig 4 as a shaded box with two sloping sides. Its volume is the area of the base in the plane Span $\{\mathbf{a}_1, \mathbf{a}_3\}$ times the altitude of $\mathbf{a}_2$ above Span $\{\mathbf{a}_1, \mathbf{a}_3\}$. Any vector $\mathbf{a}_2 + c\mathbf{a}_1$ has the same altitude because $\mathbf{a}_2 + c\mathbf{a}_1$ lies in the plane $\mathbf{a}_2$ + Span $\{\mathbf{a}_1, \mathbf{a}_3\}$, which is parallel to Span $\{\mathbf{a}_1, \mathbf{a}_3\}$. Hence the volume of the parallelepiped is unchanged when $[\mathbf{a}_1 \quad \mathbf{a}_2 \quad \mathbf{a}_3]$ is changed to $[\mathbf{a}_1 \quad \mathbf{a}_2 + c\mathbf{a}_1 \quad \mathbf{a}_3]$.

© 2012 Pearson Education, Inc.

Determinants as Area or Volume

FIGURE 4 Two parallelepipeds of equal volume.

© 2012 Pearson Education, Inc.

Determinants as Area or Volume

EXAMPLE 4

Calculate the area of the parallelogram determined by the points (-2, -2), (0, 3), (4, -1), and (6, 4). See Fig 5.

SOLUTION

First translate the parallelogram to one having the origin as a vertex. For example, subtract the vertex (-2, -2) from each of the four vertices. The new parallelogram has the same area, and its vertices are (0, 0), (2, 5), (6, 1), and (8, 6). See Fig 5(b). This parallelogram is determined by the columns of

$$A = \begin{bmatrix} 2 & 6 \\ 5 & 1 \end{bmatrix}$$

Since $|\det A| = |-28|$, the area of the parallelogram is 28.3

© 2012 Pearson Education, Inc.

Determinants as Area or Volume

FIGURE 5 Translating a parallelogram does not change its area.

© 2012 Pearson Education, Inc.

THEOREN 10

Let $T: \mathbb{R}^2 \to \mathbb{R}^2$ be the linear transformation determined by a 2×2 matrix A. If S is a parallelogram in $\mathbb{R}^2$, then

$$\{\text{ area of } T(S)\} = |\det A| \cdot \{\text{ area of } S\} \tag{5}$$

If T is determined by a 3×3 matrix A, and if S is a parallelepiped in $\mathbb{R}^3$, then

$$\{\text{volume of } T(S)\} = |\det A| \cdot \{\text{volume of } S\} \tag{6}$$

PROOF

Consider the 2×2 case, with $A = [\mathbf{a}_1 \quad \mathbf{a}_2]$.

© 2012 Pearson Education, Inc.

Linear Transformations

A parallelogram at the origin in $\mathbb{R}^2$ determined by vectors $\mathbf{b}_1$ and $\mathbf{b}_2$ has the form

$$S = \{s_1\mathbf{b}_1 + s_2\mathbf{b}_2 : 0 \leqq s_1 \leqq 1,\ 0 \leqq s_2 \leqq 1\}$$

The image of S under T consists of points of the form

$$\begin{aligned} T(s_1\mathbf{b}_1 + s_2\mathbf{b}_2) &= s_1 T(\mathbf{b}_1) + s_2 T(\mathbf{b}_2) \\ &= s_1 A\mathbf{b}_1 + s_2 A\mathbf{b}_2 \end{aligned}$$

Where $0 \leqq s_1 \leqq 1$, $0 \leqq s_2 \leqq 1$. It follows that $T(S)$ is the parallelogram determined by the columns of the matrix $[A\mathbf{b}_1 \quad A\mathbf{b}_2]$. This matrix can be written as AB, where $B = [\mathbf{b}_1 \quad \mathbf{b}_2]$.

© 2012 Pearson Education, Inc.

Linear Transformations

By theorem 9 and the product theorem for determinants,

$$\{\text{area of } T(S)\} = \left|\det AB\right| = \left|\det A\right| \cdot \left|\det B\right|$$
$$= \left|\det A\right| \cdot \{\text{area of } S\}$$

An arbitrary parallelogram has the form $\mathbf{p} + S$, where $\mathbf{p}$ is vector and S is a parallelogram at the origin, as above. It is easy to see that T transforms $\mathbf{p} + S$ into $T(\mathbf{p}) + T(S)$.

$$\begin{aligned}\{\text{area of } T(p + S)\} &= \{\text{area of } T(p) + T(S)\}\\ &= \{\text{area of } T(S)\}\\ &= \left|\det A\right| \cdot \{\text{area of } S\}\\ &= \left|\det A\right| \cdot \{\text{area of } \mathbf{p} + S\}\end{aligned}$$

© 2012 Pearson Education, Inc.

Linear Transformations

This shows that (5) holds for all parallelograms in $\mathbb{R}^2$. The proof of (6) for the 3×3 case is analogous.

FIGURE 6 Approximating a planar region by a union of squares. The approximation improves as the grid becomes finer.

© 2012 Pearson Education, Inc.

Linear Transformations

FIGURE 7 Approximating $T(R)$ by a union of parallelograms.

© 2012 Pearson Education, Inc.

4 Vector Spaces

4.1

VECTOR SPACES AND SUBSPACES

PEARSON

© 2016 Pearson Education, Ltd.

VECTOR SPACES AND SUBSPACES

- **Definition:** A **vector space** is a nonempty set V of objects, called *vectors*, on which are defined two operations, called *addition and multiplication by scalars* (real numbers), subject to the ten axioms (or rules) listed below. The axioms must hold for all vectors **u**, **v**, and **w** in V and for all scalars c and d.

  1. The sum of **u** and **v**, denoted by $\mathbf{u}+\mathbf{v}$, is in V.
  2. $\mathbf{u}+\mathbf{v}=\mathbf{v}+\mathbf{u}$.
  3. $(\mathbf{u}+\mathbf{v})+\mathbf{w}=\mathbf{u}+(\mathbf{v}+\mathbf{w})$.
  4. There is a zero vector 0 in V such that

     $$\mathbf{u}+0=\mathbf{u} \quad .$$

© 2016 Pearson Education, Ltd.

VECTOR SPACES AND SUBSPACES

5. For each **u** in V, there is a vector $-\mathrm{u}$ in V such that $\mathrm{u}+(-\mathrm{u})=0$.
6. The scalar multiple of **u** by c, denoted by c**u**, is in V.
7. $c(\mathrm{u}+\mathrm{v})=c\mathrm{u}+c\mathrm{v}$.
8. $(c+d)\mathrm{u}=c\mathrm{u}+d\mathrm{u}$.
9. $c(d\mathrm{u})=(cd)\mathrm{u}$.
10. $1\mathrm{u}=\mathrm{u}$.

- Using these axioms, we can show that the zero vector in Axiom 4 is unique, and the vector $-\mathrm{u}$, called the **negative** of **u**, in Axiom 5 is unique for each **u** in V.

© 2016 Pearson Education, Ltd.

VECTOR SPACES AND SUBSPACES

- For each **u** in V and scalar c,

$$0\mathrm{u} = 0 \qquad (1)$$
$$c0 = 0 \qquad (2)$$
$$-\mathrm{u} = (-1)\mathrm{u} \qquad (3)$$

- **Example 2:** Let V be the set of all arrows (directed line segments) in three-dimensional space, with two arrows regarded as equal if they have the same length and point in the same direction. Define addition by the parallelogram rule, and for each **v** in V, define c**v** to be the arrow whose length is $|c|$ times the length of **v**, pointing in the same direction as **v** if $c \geq 0$ and otherwise pointing in the opposite direction.

© 2016 Pearson Education, Ltd.

VECTOR SPACES AND SUBSPACES

- See the following figure below. Show that V is a vector space.

- **Solution:** The definition of V is geometric, using concepts of length and direction.
- No $x\,y\,z$-coordinate system is involved.
- An arrow of zero length is a single point and represents the zero vector.
- The negative of **v** is $(-1)\mathbf{v}$.
- So Axioms 1, 4, 5, 6, and 10 are evident. See the figures on the next slide.

© 2016 Pearson Education, Ltd.

SUBSPACES

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

$(\mathbf{u}+\mathbf{v})+\mathbf{w}=\mathbf{u}+(\mathbf{v}+\mathbf{w}).$

- **Definition:** A **subspace** of a vector space V is a subset H of V that has three properties:
  a. The zero vector of V is in H.
  b. H is closed under vector addition. That is, for each **u** and **v** in H, the sum $\mathbf{u}+\mathbf{v}$ is in H.

© 2016 Pearson Education, Ltd.

SUBSPACES

c. H is closed under multiplication by scalars. That is, for each $\mathbf{u}$ in H and each scalar c, the vector $c\mathbf{u}$ is in H.

- Properties (a), (b), and (c) guarantee that a subspace H of V is itself a *vector space*, under the vector space operations already defined in V.

- Every subspace is a vector space.

- Conversely, every vector space is a subspace (of itself and possibly of other larger spaces).

© 2016 Pearson Education, Ltd.

A SUBSPACE SPANNED BY A SET

- The set consisting of only the zero vector in a vector space V is a subspace of V, called the **zero subspace** and written as $\{\mathbf{0}\}$.

- As the term **linear combination** refers to any sum of scalar multiples of vectors, and Span $\{\mathbf{v}_1,\ldots,\mathbf{v}_p\}$ denotes the set of all vectors that can be written as linear combinations of $\mathbf{v}_1,\ldots,\mathbf{v}_p$.

© 2016 Pearson Education, Ltd.

A SUBSPACE SPANNED BY A SET

- **Example 10:** Given $\mathbf{v}_1$ and $\mathbf{v}_2$ in a vector space V, let $H = \text{Span}\{\mathrm{v}_1, \mathrm{v}_2\}$. Show that H is a subspace of V.
- **Solution:** The zero vector is in H, since $0 = 0\mathrm{v}_1 + 0\mathrm{v}_2$.
- To show that H is closed under vector addition, take two arbitrary vectors in H, say,

$$\mathrm{u} = s_1\mathrm{v}_1 + s_2\mathrm{v}_2 \quad \text{and} \quad \mathrm{w} = t_1\mathrm{v}_1 + t_2\mathrm{v}_2.$$

- By Axioms 2, 3, and 8 for the vector space V,

$$\begin{aligned}\mathrm{u} + \mathrm{w} &= (s_1\mathrm{v}_1 + s_2\mathrm{v}_2) + (t_1\mathrm{v}_1 + t_2\mathrm{v}_2) \\ &= (s_1 + t_1)\mathrm{v}_1 + (s_2 + t_2)\mathrm{v}_2\end{aligned}$$

© 2016 Pearson Education, Ltd.

A SUBSPACE SPANNED BY A SET

- So $\mathbf{u} + \mathbf{w}$ is in H.

- Furthermore, if c is any scalar, then by Axioms 7 and 9,

$$c\mathbf{u} = c(s_1\mathbf{v}_1 + s_2\mathbf{v}_2) = (cs_1)\mathbf{v}_1 + (cs_2)\mathbf{v}_2$$

  which shows that $c\mathbf{u}$ is in H and H is closed under scalar multiplication.

- Thus H is a subspace of V.

© 2016 Pearson Education, Ltd.

A SUBSPACE SPANNED BY A SET

- **Theorem 1:** If $\mathbf{v}_1,\ldots,\mathbf{v}_p$ are in a vector space V, then Span $\{\mathbf{v}_1,\ldots,\mathbf{v}_p\}$ is a subspace of V.

- We call Span $\{\mathbf{v}_1,\ldots,\mathbf{v}_p\}$ **the subspace spanned** (or **generated**) by $\{\mathbf{v}_1,\ldots,\mathbf{v}_p\}$.

- Give any subspace H of V, a **spanning** (or **generating**) set for H is a set $\{\mathbf{v}_1,\ldots,\mathbf{v}_p\}$ in H such that

$$H = \text{Span}\{\text{v}_1,\ldots\text{v}_p\}.$$

© 2016 Pearson Education, Ltd.

4 Vector Spaces

4.2

NULL SPACES, COLUMN SPACES, AND LINEAR TRANSFORMATIONS

PEARSON

© 2016 Pearson Education, Inc.

NULL SPACE OF A MATRIX

- **Definition:** The **null space** of an $m \times n$ matrix A, written as Nul A, is the set of all solutions of the homogeneous equation $A\mathrm{x} = 0$. In set notation,

$$\text{Nul } A = \{\mathrm{x} : \mathrm{x} \text{ is in } \mathbb{R}^n \text{ and } A\mathrm{x} = 0\}.$$

- **Theorem 2:** The null space of an $m \times n$ matrix A is a subspace of $\mathbb{R}^n$. Equivalently, the set of all solutions to a system $A\mathrm{x} = 0$ of m homogeneous linear equations in n unknowns is a subspace of $\mathbb{R}^n$.
- **Proof:** Nul A is a subset of $\mathbb{R}^n$ because A has n columns.
- We need to show that Nul A satisfies the three properties of a subspace.

© 2016 Pearson Education, Inc.

NULL SPACE OF A MATRIX

- $\mathbf{0}$ is in Null A.
- Next, let $\mathbf{u}$ and $\mathbf{v}$ represent any two vectors in Nul A.
- Then

$$A\mathrm{u} = 0 \text{ and } A\mathrm{v} = 0$$

- To show that $\mathrm{u} + \mathrm{v}$ is in Nul A, we must show that $A(\mathrm{u} + \mathrm{v}) = 0$.
- Using a property of matrix multiplication, compute

$$A(\mathrm{u} + \mathrm{v}) = A\mathrm{u} + A\mathrm{v} = 0 + 0 = 0$$

- Thus $\mathrm{u} + \mathrm{v}$ is in Nul A, and Nul A is closed under vector addition.

© 2016 Pearson Education, Inc.

NULL SPACE OF A MATRIX

- Finally, if c is any scalar, then

$$A(c\mathrm{u}) = c(A\mathrm{u}) = c(0) = 0$$

  which shows that $c\mathbf{u}$ is in Nul A.
- Thus Nul A is a subspace of $\mathbb{R}^n$.

- **An Explicit Description of Nul A**
- There is no obvious relation between vectors in Nul A and the entries in A.
- We say that Nul A is defined *implicitly*, because it is defined by a condition that must be checked.

© 2016 Pearson Education, Inc.

NULL SPACE OF A MATRIX

- No explicit list or description of the elements in Nul A is given.
- *Solving* the equation $A\mathrm{x} = 0$ amounts to producing an *explicit* description of Nul A.

- **Example 3:** Find a spanning set for the null space of the matrix

$$A = \begin{bmatrix} -3 & 6 & -1 & 1 & -7 \\ 1 & -2 & 2 & 3 & -1 \\ 2 & -4 & 5 & 8 & -4 \end{bmatrix}.$$

© 2016 Pearson Education, Inc.

NULL SPACE OF A MATRIX

- **Solution:** The first step is to find the general solution of $A\mathrm{x} = 0$ in terms of free variables.

- Row reduce the augmented matrix $\begin{bmatrix} A & 0 \end{bmatrix}$ to *reduce* echelon form in order to write the basic variables in terms of the free variables:

$$\begin{bmatrix} 1 & -2 & 0 & -1 & 3 & 0 \\ 0 & 0 & 1 & 2 & -2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}, \qquad \begin{aligned} x_1 - 2x_2 - x_4 + 3x_5 &= 0 \\ x_3 + 2x_4 - 2x_5 &= 0 \\ 0 &= 0 \end{aligned}$$

© 2016 Pearson Education, Inc.

NULL SPACE OF A MATRIX

- The general solution is $x_1 = 2x_2 + x_4 - 3x_5$, $x_3 = -2x_4 + 2x_5$, with x_2, x_4, and x_5 free.
- Next, decompose the vector giving the general solution into a linear combination of *vectors where the weights are the free variables*. That is,

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix} = \begin{bmatrix} 2x_2 + x_4 - 3x_5 \\ x_2 \\ -2x_4 + 2x_5 \\ x_4 \\ x_5 \end{bmatrix} = x_2 \underbrace{\begin{bmatrix} 2 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}}_{\mathbf{u}} + x_4 \underbrace{\begin{bmatrix} 1 \\ 0 \\ -2 \\ 1 \\ 0 \end{bmatrix}}_{\mathbf{v}} + x_5 \underbrace{\begin{bmatrix} -3 \\ 0 \\ 2 \\ 0 \\ 1 \end{bmatrix}}_{\mathbf{w}}$$

© 2016 Pearson Education, Inc.

NULL SPACE OF A MATRIX

$$= x_2\mathbf{u} + x_4\mathbf{v} + x_5\mathbf{w} \qquad (3)$$

- Every linear combination of **u**, **v**, and **w** is an element of Nul A.
- Thus {**u**, **v**, **w**} is a spanning set for Nul A.

1. The spanning set produced by the method in Example (3) is automatically linearly independent because the free variables are the weights on the spanning vectors.
2. When Nul A contains nonzero vectors, the number of vectors in the spanning set for Nul A equals the number of free variables in the equation $A\mathbf{x} = \mathbf{0}$.

© 2016 Pearson Education, Inc.

COLUMN SPACE OF A MATRIX

- **Definition:** The column space of an $m \times n$ matrix A, written as Col A, is the set of all linear combinations of the columns of A. If $A = \begin{bmatrix} \mathrm{a}_1 & \cdots & \mathrm{a}_n \end{bmatrix}$, then

$$\text{Col } A = \text{Span}\{\mathrm{a}_1, \ldots, \mathrm{a}_n\}.$$

- **Theorem 3:** The column space of an $m \times n$ matrix A is a subspace of $\mathbb{R}^m$.
- A typical vector in Col A can be written as $A\mathbf{x}$ for some $\mathbf{x}$ because the notation $A\mathbf{x}$ stands for a linear combination of the columns of A. That is,

$$\text{Col } A = \{\mathrm{b} : \mathrm{b} = A\mathrm{x} \text{ for some } \mathrm{x} \text{ in } \mathbb{R}^n\}.$$

© 2016 Pearson Education, Inc.

COLUMN SPACE OF A MATRIX

- The notation $A\mathbf{x}$ for vectors in Col A also shows that Col A is the *range* of the linear transformation $\mathrm{x} \mapsto A\mathrm{x}$.
- The column space of an $m \times n$ matrix A is all of $\mathbb{R}^m$ if and only if the equation $A\mathrm{x} = \mathrm{b}$ has a solution for each $\mathbf{b}$ in $\mathbb{R}^m$.
- **Example 7:** Let $A = \begin{bmatrix} 2 & 4 & -2 & 1 \\ -2 & -5 & 7 & 3 \\ 3 & 7 & -8 & 6 \end{bmatrix}$, $\mathrm{u} = \begin{bmatrix} 3 \\ -2 \\ -1 \\ 0 \end{bmatrix}$ and $\mathrm{v} = \begin{bmatrix} 3 \\ -1 \\ 3 \end{bmatrix}$.

© 2016 Pearson Education, Inc.

COLUMN SPACE OF A MATRIX

a. Determine if **u** is in Nul A. Could **u** be in Col A?

b. Determine if **v** is in Col A. Could **v** be in Nul A?

- **Solution:**

a. An explicit description of Nul A is not needed here. Simply compute the product $A\mathbf{u}$.

$$A\mathrm{u} = \begin{bmatrix} 2 & 4 & -2 & 1 \\ -2 & -5 & 7 & 3 \\ 3 & 7 & -8 & 6 \end{bmatrix} \begin{bmatrix} 3 \\ -2 \\ -1 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ -3 \\ 3 \end{bmatrix} \neq \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

© 2016 Pearson Education, Inc.

COLUMN SPACE OF A MATRIX

- **u** is *not* a solution of $A\mathrm{x} = 0$, so **u** is not in Nul A.
- Also, with four entries, **u** could not possibly be in Col A, since Col A is a subspace of $\mathbb{R}^3$.

b. Reduce $\begin{bmatrix} A & \mathrm{v} \end{bmatrix}$ to an echelon form.

$$\begin{bmatrix} A & \mathrm{v} \end{bmatrix} = \begin{bmatrix} 2 & 4 & -2 & 1 & 3 \\ -2 & -5 & 7 & 3 & -1 \\ 3 & 7 & -8 & 6 & 3 \end{bmatrix} \sim \begin{bmatrix} 2 & 4 & -2 & 1 & 3 \\ 0 & 1 & -5 & -4 & -2 \\ 0 & 0 & 0 & 17 & 1 \end{bmatrix}$$

- The equation $A\mathrm{x} = \mathrm{v}$ is consistent, so **v** is in Col A.

© 2016 Pearson Education, Inc.

KERNEL AND RANGE OF A LINEAR TRANSFORMATION

- With only three entries, **v** could not possibly be in Nul A, since Nul A is a subspace of $\mathbb{R}^4$.

- Subspaces of vector spaces other than $\mathbb{R}^n$ are often described in terms of a linear transformation instead of a matrix.

- **Definition:** A **linear transformation** T from a vector space V into a vector space W is a rule that assigns to each vector **x** in V a unique vector $T(\mathbf{x})$ in W, such that

  i. $T(\mathrm{u}+\mathrm{v}) = T(\mathrm{u}) + T(\mathrm{v})$ for all **u**, **v** in V, and

  ii. $T(c\mathrm{u}) = cT(\mathrm{u})$ for all **u** in V and all scalars c.

© 2016 Pearson Education, Inc.

KERNEL AND RANGE OF A LINEAR TRANSFORMATION

- The **kernel** (or **null space**) of such a T is the set of all **u** in V such that $T(\mathrm{u}) = 0$ (the zero vector in W).

- The **range** of T is the set of all vectors in W of the form $T(\mathbf{x})$ for some **x** in V.

- The kernel of T is a subspace of V.

- The range of T is a subspace of W.

© 2016 Pearson Education, Inc.

CONTRAST BETWEEN NUL A AND COL A FOR AN $m \times n$ MATRIX A

Nul A	Col A
1. Nul A is a subspace of $\mathbb{R}^n$.	1. Col A is a subspace of $\mathbb{R}^m$.
2. Nul A is implicitly defined; *i.e.*, you are given only a condition $(A\mathrm{x} = 0)$ that vectors in Nul A must satisfy.	2. Col A is explicitly defined; *i.e.*, you are told how to build vectors in Col A.

© 2016 Pearson Education, Inc.

CONTRAST BETWEEN NUL A AND COL A FOR AN $m \times n$ MATRIX A

3. It takes time to find vectors in Nul A. Row operations on $\begin{bmatrix} A & 0 \end{bmatrix}$ are required.	3. It is easy to find vectors in Col A. The columns of a are displayed; others are formed from them.
4. There is no obvious relation between Nul A and the entries in A.	4. There is an obvious relation between Col A and the entries in A, since each column of A is in Col A.

© 2016 Pearson Education, Inc.

CONTRAST BETWEEN NUL A AND COL A FOR AN $m \times n$ MATRIX A

5. A typical vector **v** in Nul A has the property that $A\mathrm{v} = 0$.	5. A typical vector **v** in Col A has the property that the equation $A\mathrm{x} = \mathrm{v}$ is consistent.
6. Given a specific vector **v**, it is easy to tell if **v** is in Nul A. Just compare A**v**.	6. Given a specific vector **v**, it may take time to tell if v is in Col A. Row operations on $\begin{bmatrix} A & \mathrm{v} \end{bmatrix}$ are required.

© 2016 Pearson Education, Inc.

CONTRAST BETWEEN NUL A AND COL A FOR AN $m \times n$ MATRIX A

7. Nul $A = \{0\}$ if and only if the equation $A\mathrm{x} = 0$ has only the trivial solution.	7. Col $A = \mathbb{R}^m$ if and only if the equation $A\mathrm{x} = \mathrm{b}$ has a solution for every $\mathbf{b}$ in $\mathbb{R}^m$.
8. Nul $A = \{0\}$ if and only if the linear transformation $\mathrm{x} \mapsto A\mathrm{x}$ is one-to-one.	8. Col $A = \mathbb{R}^m$ if and only if the linear transformation $\mathrm{x} \mapsto A\mathrm{x}$ maps $\mathbb{R}^n$ onto $\mathbb{R}^m$.

© 2016 Pearson Education, Inc.

4 Vector Spaces

4.3

LINEARLY INDEPENDENT SETS; BASES

PEARSON

© 2012 Pearson Education, Inc.

LINEAR INDEPENDENT SETS; BASES

- An indexed set of vectors $\{\mathbf{v}_1, \ldots, \mathbf{v}_p\}$ in V is said to be **linearly independent** if the vector equation

$$c_1\mathrm{v}_1 + c_2\mathrm{v}_2 + \ldots + c_p\mathrm{v}_p = 0 \qquad \text{----(1)}$$

  has *only* the trivial solution, $c_1 = 0, \ldots, c_p = 0$.
- The set $\{\mathbf{v}_1, \ldots, \mathbf{v}_p\}$ is said to be **linearly dependent** if (1) has a nontrivial solution, *i.e.*, if there are some weights, $c_1, \ldots, c_p$, *not all zero*, such that (1) holds.
- In such a case, (1) is called a **linear dependence relation** among $\mathbf{v}_1, \ldots, \mathbf{v}_p$.

© 2012 Pearson Education, Inc.

LINEAR INDEPENDENT SETS; BASES

- **Theorem 4:** An indexed set $\{\mathbf{v}_1, \ldots, \mathbf{v}_p\}$ of two or more vectors, with $\mathrm{v}_1 \neq 0$, is linearly dependent if and only if some $\mathbf{v}_j$ (with $j > 1$) is a linear combination of the preceding vectors, $\mathrm{v}_1, \ldots, \mathrm{v}_{j-1}$.

- **Definition:** Let H be a subspace of a vector space V. An indexed set of vectors $\mathcal{B} = \{\mathrm{b}_1, \ldots, \mathrm{b}_p\}$ in V is a basis for H if

  (i) $\mathcal{B}$ is a linearly independent set, and

  (ii) The subspace spanned by $\mathcal{B}$ coincides with H; that is, $H = \mathrm{Span}\{\mathrm{b}_1, \ldots, \mathrm{b}_p\}$

© 2012 Pearson Education, Inc.

LINEAR INDEPENDENT SETS; BASES

- The definition of a basis applies to the case when $H = V$, because any vector space is a subspace of itself.

- Thus a basis of V is a linearly independent set that spans V.

- When $H \neq V$, condition (ii) includes the requirement that each of the vectors $\mathbf{b}_1, \ldots, \mathbf{b}_p$ must belong to H, because Span $\{\mathbf{b}_1, \ldots, \mathbf{b}_p\}$ contains $\mathbf{b}_1, \ldots, \mathbf{b}_p$.

© 2012 Pearson Education, Inc.

STANDARD BASIS

- **EXAMPLE 4**
- Let $\mathbf{e}_1, \ldots, \mathbf{e}_n$ be the columns of the $n \times n$ matrix, I_n.
- That is,

$$\mathbf{e}_1 = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \mathbf{e}_2 = \begin{bmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{bmatrix}, \ldots, \mathbf{e}_n = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix}$$

- The set $\{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$ is called the **standard basis** for $\mathbb{R}^n$. See the following figure.

FIGURE 1
The standard basis for $\mathbb{R}^3$.

© 2012 Pearson Education, Inc.

THE SPANNING SET THEOREM

EXAMPLE 6

Let $S = \{1, t, t^2, \ldots t^n\}$. Verify that S is a basis for $\mathbb{P}^n$. This basis is called the **Standard basis** for $\mathbb{P}^n$.

SOLUTION

Certainly S spans $\mathbb{P}^n$. To show that S is linearly independent, suppose that $c_0, \ldots, c_n$ satisfy

$$c_0 \cdot 1 + c_1 t + c_2 t_2 + \ldots + c_n t^n = \mathbf{0}\ (t) \qquad (2)$$

A fundamental theorem in algebra says that the only polynomial in $\mathbb{P}^n$ with more than n zeros is the zero polynomial. Equation (2) holds for all t only if $c_0 = \ldots = c_n = 0$.

© 2012 Pearson Education, Inc.

THE SPANNING SET THEOREM

This proves that S is linearly independent and hence is a basis for $\mathbb{P}^n$. See Fig 2.

FIGURE 2

The standard basis for $\mathbb{P}_2$.

© 2012 Pearson Education, Inc.

THE SPANNING SET THEOREM

- **Theorem 5:** Let $S = \{\mathrm{v}_1, \ldots, \mathrm{v}_p\}$ be a set in V, and let $H = \mathrm{Span}\{\mathrm{v}_1, \ldots, \mathrm{v}_p\}$.
  a. If one of the vectors in S—say, $\mathbf{v}_k$—is a linear combination of the remaining vectors in S, then the set formed from S by removing $\mathbf{v}_k$ still spans H.
  b. If $H \neq \{0\}$, some subset of S is a basis for H.
- **Proof:**
  a. By rearranging the list of vectors in S, if necessary, we may suppose that $\mathbf{v}_p$ is a linear combination of $\mathrm{v}_1, \ldots, \mathrm{v}_{p-1}$—say,

© 2012 Pearson Education, Inc.

THE SPANNING SET THEOREM

$$\mathrm{v}_p = a_1\mathrm{v}_1 + \ldots + a_{p-1}\mathrm{v}_{p-1} \qquad \text{----(2)}$$

- Given any **x** in *H*, we may write

  $$\mathrm{x} = c_1\mathrm{v}_1 + \ldots + c_{p-1}\mathrm{v}_{p-1} + c_p\mathrm{v}_p \qquad \text{----(3)}$$

  for suitable scalars $c_1, \ldots, c_p$.

- Substituting the expression for $\mathbf{v}_p$ from (2) into (3), it is easy to see that **x** is a linear combination of $\mathrm{v}_1, \ldots \mathrm{v}_{p-1}$.

- Thus $\{\mathrm{v}_1, \ldots, \mathrm{v}_{p-1}\}$ spans *H*, because **x** was an arbitrary element of *H*.

© 2012 Pearson Education, Inc.

THE SPANNING SET THEOREM

b. If the original spanning set S is linearly independent, then it is already a basis for H.

- Otherwise, one of the vectors in S depends on the others and can be deleted, by part (a).
- So long as there are two or more vectors in the spanning set, we can repeat this process until the spanning set is linearly independent and hence is a basis for H.
- If the spanning set is eventually reduced to one vector, that vector will be nonzero (and hence linearly independent) because $H \neq \{0\}$.

© 2012 Pearson Education, Inc.

THE SPANNING SET THEOREM

- **Example 7:** Let $\mathbf{v}_1 = \begin{bmatrix} 0 \\ 2 \\ -1 \end{bmatrix}, \mathbf{v}_2 = \begin{bmatrix} 2 \\ 2 \\ 0 \end{bmatrix}, \mathbf{v}_3 = \begin{bmatrix} 6 \\ 16 \\ -5 \end{bmatrix}$

  and $H = \text{Span}\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$.
  Note that $\mathbf{v}_3 = 5\mathbf{v}_1 + 3\mathbf{v}_2$, and show that $\text{Span}\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\} = \text{Span}\{\mathbf{v}_1, \mathbf{v}_2\}$. Then find a basis for the subspace *H*.

- **Solution:** Every vector in Span $\{\mathbf{v}_1, \mathbf{v}_2\}$ belongs to *H* because

$$c_1\mathbf{v}_1 + c_2\mathbf{v}_2 = c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + 0\mathbf{v}_3$$

© 2012 Pearson Education, Inc.

THE SPANNING SET THEOREM

- Now let **x** be any vector in *H*—say,

$$\mathrm{x} = c_1\mathrm{v}_1 + c_2\mathrm{v}_2 + c_3\mathrm{v}_3 .$$

- Since $\mathrm{v}_3 = 5\mathrm{v}_1 + 3\mathrm{v}_2$, we may substitute

$$\begin{aligned}\mathrm{x} &= c_1\mathrm{v}_1 + c_2\mathrm{v}_2 + c_3(5\mathrm{v}_1 + 3\mathrm{v}_2) \\ &= (c_1 + 5c_3)\mathrm{v}_1 + (c_2 + 3c_3)\mathrm{v}_2\end{aligned}$$

- Thus **x** is in Span $\{\mathbf{v}_1, \mathbf{v}_2\}$, so every vector in *H* already belongs to Span $\{\mathbf{v}_1, \mathbf{v}_2\}$.
- We conclude that *H* and Span $\{\mathbf{v}_1, \mathbf{v}_2\}$ are actually the set of vectors.
- It follows that $\{\mathbf{v}_1, \mathbf{v}_2\}$ is a basis of *H* since $\{\mathbf{v}_1, \mathbf{v}_2\}$ is linearly independent.

© 2012 Pearson Education, Inc.

BASIS FOR COL *B*

- **Example 8:** Find a basis for Col *B*, where

$$B = \begin{bmatrix} \mathrm{b}_1 & \mathrm{b}_2 & \cdots & \mathrm{b}_5 \end{bmatrix} = \begin{bmatrix} 1 & 4 & 0 & 2 & 0 \\ 0 & 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

- **Solution:** Each nonpivot column of *B* is a linear combination of the pivot columns.
- In fact, $\mathrm{b}_2 = 4\mathrm{b}_1$ and $\mathrm{b}_4 = 2\mathrm{b}_1 - \mathrm{b}_3$.
- By the Spanning Set Theorem, we may discard $\mathbf{b}_2$ and $\mathbf{b}_4$, and $\{\mathbf{b}_1, \mathbf{b}_3, \mathbf{b}_5\}$ will still span Col *B*.

© 2012 Pearson Education, Inc.

BASIS FOR COL *B*

- Let

$$S = \{\mathbf{b}_1, \mathbf{b}_3, \mathbf{b}_5\} = \left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix} \right\}$$

- Since $\mathbf{b}_1 \neq 0$ and no vector in S is a linear combination of the vectors that precede it, S is linearly independent. (Theorem 4).

- Thus S is a basis for Col B.

© 2012 Pearson Education, Inc.

BASES FOR NUL *A* AND COL *A*

- **Theorem 6:** The pivot columns of a matrix A form a basis for Col A.

- **Proof:** Let B be the reduced echelon form of A.
- The set of pivot columns of B is linearly independent, for no vector in the set is a linear combination of the vectors that precede it.
- Since A is row equivalent to B, the pivot columns of A are linearly independent as well, because any linear dependence relation among the columns of A corresponds to a linear dependence relation among the columns of B.

© 2012 Pearson Education, Inc.

BASES FOR NUL *A* AND COL *A*

- For this reason, every nonpivot column of A is a linear combination of the pivot columns of A.
- Thus the nonpivot columns of a may be discarded from the spanning set for Col A, by the Spanning Set Theorem.
- This leaves the pivot columns of A as a basis for Col A.

- **Warning:** The pivot columns of a matrix A are evident when A has been reduced only to echelon form.
- But, be careful to use the pivot columns of A itself for the basis of Col A.

© 2012 Pearson Education, Inc.

BASES FOR NUL *A* AND COL *A*

- Row operations can change the column space of a matrix.
- The columns of an echelon form B of A are often not in the column space of A.

- **Two Views of a Basis**
- When the Spanning Set Theorem is used, the deletion of vectors from a spanning set must stop when the set becomes linearly independent.
- If an additional vector is deleted, it will not be a linear combination of the remaining vectors, and hence the smaller set will no longer span V.

© 2012 Pearson Education, Inc.

TWO VIEWS OF A BASIS

- Thus a basis is a spanning set that is as small as possible.

- A basis is also a linearly independent set that is as large as possible.

- If S is a basis for V, and if S is enlarged by one vector—say, $\mathbf{w}$—from V, then the new set cannot be linearly independent, because S spans V, and $\mathbf{w}$ is therefore a linear combination of the elements in S.

© 2012 Pearson Education, Inc.

4 Vector Spaces

4.4

COORDINATE SYSTEMS

PEARSON

© 2012 Pearson Education, Inc.

THE UNIQUE REPRESENTATION THEOREM

- **Theorem 7:** Let $\mathcal{B} = \{\mathbf{b}_1, \ldots, \mathbf{b}_n\}$ be a basis for vector space V. Then for each **x** in V, there exists a unique set of scalars $c_1, \ldots, c_n$ such that

$$\mathbf{x} = c_1\mathbf{b}_1 + \ldots + c_n\mathbf{b}_n \qquad \text{----(1)}$$

- **Proof:** Since $\mathcal{B}$ spans V, there exist scalars such that (1) holds.
- Suppose **x** also has the representation

$$\mathbf{x} = d_1\mathbf{b}_1 + \ldots + d_n\mathbf{b}_n$$

for scalars $d_1, \ldots, d_n$.

© 2012 Pearson Education, Inc.

THE UNIQUE REPRESENTATION THEOREM

- Then, subtracting, we have
$$0 = \mathrm{x} - \mathrm{x} = (c_1 - d_1)\mathrm{b}_1 + \ldots + (c_n - d_n)\mathrm{b}_n \quad \text{----(2)}$$

- Since B is linearly independent, the weights in (2) must all be zero. That is, $c_j = d_j$ for $1 \le j \le n$.

- **Definition:** Suppose $\mathrm{B} = \{\mathrm{b}_1, \ldots, \mathrm{b}_n\}$ is a basis for V and **x** is in V. **The coordinates of x relative to the basis B** (or the **B-coordinate of x**) are the weights $c_1, \ldots, c_n$ such that $\mathrm{x} = c_1\mathrm{b}_1 + \ldots + c_n\mathrm{b}_n$.

© 2012 Pearson Education, Inc.

THE UNIQUE REPRESENTATION THEOREM

- If $c_1, \ldots, c_n$ are the **B**-coordinates of **x**, then the vector in $\mathbb{R}^n$.

$$[\mathrm{x}]_{\mathrm{B}} = \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix}$$

  is the **coordinate vector of x** (**relative to B**), or the **B-coordinate vector of x**.

- The mapping $\mathrm{x} \mapsto [\mathrm{x}]_{\mathrm{B}}$ is the **coordinate mapping** (**determined by B**).

© 2012 Pearson Education, Inc.

A Graphical Interpretation of Coordinates

FIGURE 1 Standard graph paper.

FIGURE 2 $\mathcal{B}$-graph paper.

© 2012 Pearson Education, Inc.

A Graphical Interpretation of Coordinates

EXAMPLE 3

In crystallography, the description of a crystal lattice is aided by choosing a basis $\{\mathbf{u}, \mathbf{v}, \mathbf{w}\}$ for $\mathbb{R}^3$ that corresponds to three adjacent edges of one "unit cell" of the crystal. An entire lattice is constructed by stacking together many copies of one cell. There are fourteen basic types of unit cells; three are displayed in Fig 3.

FIGURE 3 Examples of unit cells.

© 2012 Pearson Education, Inc.

COORDINATES IN $\mathbb{R}^n$

- When a basis $\mathcal{B}$ for $\mathbb{R}^n$ is fixed, the $\mathcal{B}$-coordinate vector of a specified **x** is easily found, as in the example below.
- **Example 4:** Let $\mathbf{b}_1 = \begin{bmatrix} 2 \\ 1 \end{bmatrix}, \mathbf{b}_2 = \begin{bmatrix} -1 \\ 1 \end{bmatrix}, \mathbf{x} = \begin{bmatrix} 4 \\ 5 \end{bmatrix}$, and $\mathcal{B} = \{\mathbf{b}_1, \mathbf{b}_2\}$. Find the coordinate vector $[\mathbf{x}]_{\mathcal{B}}$ of **x** relative to $\mathcal{B}$.
- **Solution:** The $\mathcal{B}$-coordinate c_1, c_2 of **x** satisfy

$$c_1 \underbrace{\begin{bmatrix} 2 \\ 1 \end{bmatrix}}_{\mathbf{b}_1} + c_2 \underbrace{\begin{bmatrix} -1 \\ 1 \end{bmatrix}}_{\mathbf{b}_2} = \underbrace{\begin{bmatrix} 4 \\ 5 \end{bmatrix}}_{\mathbf{x}}$$

© 2012 Pearson Education, Inc.

COORDINATES IN $\mathbb{R}^n$

or

$$\begin{array}{cc} \begin{bmatrix} 2 & -1 \\ 1 & 1 \end{bmatrix} & \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} = \begin{bmatrix} 4 \\ 5 \end{bmatrix} \\ \mathbf{b}_1 \quad \mathbf{b}_2 & \mathbf{x} \end{array} \qquad \text{----(3)}$$

- This equation can be solved by row operations on an augmented matrix or by using the inverse of the matrix on the left.
- In any case, the solution is $c_1 = 3, c_2 = 2$.
- Thus $\mathbf{x} = 3\mathbf{b}_1 + 2\mathbf{b}_2$ and

$$[\mathbf{x}]_B = \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} = \begin{bmatrix} 3 \\ 2 \end{bmatrix}.$$

© 2012 Pearson Education, Inc.

COORDINATES IN $\mathbb{R}^n$

- See the following figure.

- **FIGURE 4**

The $\mathcal{B}$-coordinate vector of **x** is (3, 2).

- The matrix in (3) changes the B-coordinates of a vector **x** into the standard coordinates for **x**.
- An analogous change of coordinates can be carried out in $\mathbb{R}^n$ for a basis $\mathsf{B} = \{\mathbf{b}_1, \ldots, \mathbf{b}_n\}$.
- Let $P_{\mathsf{B}} = \begin{bmatrix} \mathbf{b}_1 & \mathbf{b}_2 & \cdots & \mathbf{b}_n \end{bmatrix}$

© 2012 Pearson Education, Inc.

COORDINATES IN $\mathbb{R}^n$

- Then the vector equation

$$\mathbf{x} = c_1\mathbf{b}_1 + c_2\mathbf{b}_2 + \ldots + c_n\mathbf{b}_n$$

is equivalent to

$$\underline{\mathbf{x} = P_{\mathcal{B}}\left[\mathbf{x}\right]_{\mathcal{B}}} \qquad \text{----(4)}$$

- $P_{\mathcal{B}}$ is called the **change-of-coordinates matrix** from $\mathcal{B}$ to the standard basis in $\mathbb{R}^n$.

- Left-multiplication by $P_{\mathcal{B}}$ transforms the coordinate vector $[\mathbf{x}]_{\mathcal{B}}$ into $\mathbf{x}$.

- Since the columns of $P_{\mathcal{B}}$ form a basis for $\mathbb{R}^n$, $P_{\mathcal{B}}$ is invertible (by the Invertible Matrix Theorem).

© 2012 Pearson Education, Inc.

COORDINATES IN $\mathbb{R}^n$

- Left-multiplication by $P_{\mathcal{B}}^{-1}$ converts **x** into its $\mathcal{B}$-coordinate vector:

$$P_{\mathcal{B}}^{-1}\mathbf{x} = [\mathbf{x}]_{\mathcal{B}}$$

- The correspondence $\mathbf{x} \mapsto [\mathbf{x}]_{\mathcal{B}}$, produced by $P_{\mathcal{B}}^{-1}$, is the coordinate mapping.

- Since $P_{\mathcal{B}}^{-1}$ is an invertible matrix, the coordinate mapping is a one-to-one linear transformation from $\mathbb{R}^n$ onto $\mathbb{R}^n$, by the Invertible Matrix Theorem.

© 2012 Pearson Education, Inc.

THE COORDINATE MAPPING

Choosing a basis B = $\{\mathbf{b}_1,\ldots\mathbf{b}_n\}$ for a vector space V introduces a coordinate system in V. The coordinate mapping $\mathbf{x} \longmapsto [\mathbf{x}]_B$ connects the possibly unfamiliar space V to the familiar space $\mathbb{R}^n$. See fig 5.

FIGURE 5 The coordinate mapping from V onto $\mathbb{R}^n$.

© 2012 Pearson Education, Inc.

THE COORDINATE MAPPING

- **Theorem 8:** Let $\mathcal{B} = \{\mathbf{b}_1, \ldots, \mathbf{b}_n\}$ be a basis for a vector space V. Then the coordinate mapping $\mathbf{x} \mapsto [\mathbf{x}]_{\mathcal{B}}$ is a one-to-one linear transformation from V onto $\mathbb{R}^n$.

- **Proof:** Take two typical vectors in V, say,

$$\mathbf{u} = c_1\mathbf{b}_1 + \ldots + c_n\mathbf{b}_n$$

$$\mathbf{w} = d_1\mathbf{b}_1 + \ldots + d_n\mathbf{b}_n$$

- Then, using vector operations,

$$\mathbf{u} + \mathbf{v} = (c_1 + d_1)\mathbf{b}_1 + \ldots + (c_n + d_n)\mathbf{b}_n$$

© 2012 Pearson Education, Inc.

THE COORDINATE MAPPING

- **Theorem 8:** Let $\mathcal{B} = \{\mathbf{b}_1, \ldots, \mathbf{b}_n\}$ be a basis for a vector space V. Then the coordinate mapping $\mathbf{x} \mapsto [\mathbf{x}]_{\mathcal{B}}$ is a one-to-one linear transformation from V onto $\mathbb{R}^n$.

- **Proof:** Take two typical vectors in V, say,

$$\mathbf{u} = c_1\mathbf{b}_1 + \ldots + c_n\mathbf{b}_n$$

$$\mathbf{w} = d_1\mathbf{b}_1 + \ldots + d_n\mathbf{b}_n$$

- Then, using vector operations,

$$\mathbf{u} + \mathbf{v} = (c_1 + d_1)\mathbf{b}_1 + \ldots + (c_n + d_n)\mathbf{b}_n$$

© 2012 Pearson Education, Inc.

THE COORDINATE MAPPING

- It follows that

$$[\mathrm{u}+\mathrm{w}]_{\mathrm{B}} = \begin{bmatrix} c_1 + d_1 \\ \vdots \\ c_n + d_n \end{bmatrix} = \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix} + \begin{bmatrix} d_1 \\ \vdots \\ d_n \end{bmatrix} = [\mathrm{u}]_{\mathrm{B}} + [\mathrm{w}]_{\mathrm{B}}$$

- So the coordinate mapping preserves addition.

- If r is any scalar, then

$$r\mathrm{u} = r(c_1\mathrm{b}_1 + \ldots + c_n\mathrm{b}_n) = (rc_1)\mathrm{b}_1 + \ldots + (rc_n)\mathrm{b}_n$$

© 2012 Pearson Education, Inc.

THE COORDINATE MAPPING

- So

$$[r\mathrm{u}]_{\mathrm{B}} = \begin{bmatrix} rc_1 \\ \vdots \\ rc_n \end{bmatrix} = r\begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix} = r[\mathrm{u}]_{\mathrm{B}}$$

- Thus the coordinate mapping also preserves scalar multiplication and hence is a linear transformation.
- The linearity of the coordinate mapping extends to linear combinations.
- If $\mathbf{u}_1, \ldots, \mathbf{u}_p$ are in V and if $c_1, \ldots, cp$ are scalars, then

$$\left[c_1\mathrm{u}_1 + \ldots + c_p\mathrm{u}_p\right]_{\mathrm{B}} = c_1\left[\mathrm{u}_1\right]_{\mathrm{B}} + \ldots + c_p\left[\mathrm{u}_p\right]_{\mathrm{B}} \text{ ----(5)}$$

© 2012 Pearson Education, Inc.

THE COORDINATE MAPPING

- In words, (5) says that the B-coordinate vector of a linear combination of $\mathbf{u}_1, \ldots, \mathbf{u}_p$ is the *same* linear combination of their coordinate vectors.
- The coordinate mapping in Theorem 8 is an important example of an *isomorphism* from V onto $\mathbb{R}^n$.
- In general, a one-to-one linear transformation from a vector space V onto a vector space W is called an **isomorphism** from V onto W.
- The notation and terminology for V and W may differ, but the two spaces are indistinguishable as vector spaces.

© 2012 Pearson Education, Inc.

THE COORDINATE MAPPING

- *Every vector space calculation in V is accurately reproduced in W, and vice versa.*
- In particular, any real vector space with a basis of n vectors is indistinguishable from $\mathbb{R}^n$.

Example 5

Let B be the standard basis of the space $\mathbb{P}_3$ of polynomials; that is, let B = $\{1, t, t^2, t^3\}$. A typical element **p** of $\mathbb{P}_3$ has the form

$$\mathbf{p}(t) = a_0 + a_1 t + a_2 t^2 + a_3 t^3$$

Since **p** is already displayed as a linear combination of the standard basis vectors, we conclude that

© 2012 Pearson Education, Inc.

THE COORDINATE MAPPING

$$[\mathrm{p}]_B = \begin{bmatrix} a_0 \\ a_1 \\ a_2 \\ a_3 \end{bmatrix}$$

Thus the coordinate mapping $\mathbf{p} \mapsto [\mathbf{p}]_B$ is an isomorphism from $\mathbb{P}_3$ onto $\mathbb{R}_4$. All vector space operations in $\mathbb{P}_3$ correspond to operations in $\mathbb{R}_4$.

If we think of $\mathbb{P}_3$ and $\mathbb{R}^4$ as displays on two computer screens that are connected via the coordinate mapping, then every vector space operation in $\mathbb{P}_3$ on one screen is exactly duplicated by a corresponding vector operation in $\mathbb{R}^4$ on the other screen.

© 2012 Pearson Education, Inc.

THE COORDINATE MAPPING

The vectors on the $\mathbb{P}_3$ screen look different from those on the $\mathbb{R}^4$ screen, but they "act" as vectors in exactly the same way. See Fig 6.

FIGURE 6 The space $\mathbb{P}_3$ is isomorphic to $\mathbb{R}^4$.

© 2012 Pearson Education, Inc.

THE COORDINATE MAPPING

Example 7: Let

$$\mathbf{v}_1 = \begin{bmatrix} 3 \\ 6 \\ 2 \end{bmatrix}, \mathbf{v}_2 = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}, \mathbf{x} = \begin{bmatrix} 3 \\ 12 \\ 7 \end{bmatrix},$$

and $\mathcal{B} = \{\mathbf{v}_1, \mathbf{v}_2\}$. Then $\mathcal{B}$ is a basis for $H = \text{Span}\{\mathbf{v}_1, \mathbf{v}_2\}$. Determine if **x** is in *H*, and if it is, find the coordinate vector of **x** relative to $\mathcal{B}$.

© 2012 Pearson Education, Inc.

THE COORDINATE MAPPING

- **Solution:** If **x** is in *H*, then the following vector equation is consistent:

$$c_1\begin{bmatrix}3\\6\\2\end{bmatrix}+c_2\begin{bmatrix}-1\\0\\1\end{bmatrix}=\begin{bmatrix}3\\12\\7\end{bmatrix}$$

- The scalars c_1 and c_2, if they exist, are the B-coordinates of **x**.

© 2012 Pearson Education, Inc.

THE COORDINATE MAPPING

- Using row operations, we obtain

$$\begin{bmatrix} 3 & -1 & 3 \\ 6 & 0 & 12 \\ 2 & 1 & 7 \end{bmatrix} \square \begin{bmatrix} 1 & 0 & 2 \\ 0 & 1 & 3 \\ 0 & 0 & 0 \end{bmatrix}.$$

- Thus $c_1 = 2$, $c_2 = 3$ and $[\mathbf{x}]_{\mathrm{B}} = \begin{bmatrix} 2 \\ 3 \end{bmatrix}$.

© 2012 Pearson Education, Inc.

THE COORDINATE MAPPING

- The coordinate system on H determined by B is shown in the following figure.

FIGURE 7 A coordinate system on a plane H in $\mathbb{R}^3$.

© 2012 Pearson Education, Inc.

4 Vector Spaces

4.5

THE DIMENSION OF A VECTOR SPACE

PEARSON

© 2016 Pearson Education, Inc.

DIMENSION OF A VECTOR SPACE

- **Theorem 9:** If a vector space V has a basis $\mathcal{B} = \{b_1, \ldots, b_n\}$, then any set in V containing more than n vectors must be linearly dependent.

- **Proof:** Let $\{\mathbf{u}_1, \ldots, \mathbf{u}_p\}$ be a set in V with more than n vectors.

- The coordinate vectors $[\mathbf{u}_1]_{\mathcal{B}}, \ldots, [\mathbf{u}_p]_{\mathcal{B}}$ form a linearly dependent set in $\mathbb{R}^n$, because there are more vectors (p) than entries (n) in each vector.

© 2016 Pearson Education, Inc.

DIMENSION OF A VECTOR SPACE

- So there exist scalars $c_1, \ldots, c_p$, not all zero, such that

$$c_1[\mathbf{u}_1]_{\mathrm{B}} + \ldots + c_p[\mathbf{u}_p]_{\mathrm{B}} = \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix}$$

The zero vector in $\mathbb{R}^n$

- Since the coordinate mapping is a linear transformation,

$$[c_1\mathbf{u}_1 + \ldots + c_p\mathbf{u}_p]_{\mathrm{B}} = \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix}$$

© 2016 Pearson Education, Inc.

DIMENSION OF A VECTOR SPACE

- The zero vector on the right displays the n weights needed to build the vector $c_1\mathrm{u}_1 + \ldots + c_p\mathrm{u}_p$ from the basis vectors in $\mathcal{B}$.

- That is, $c_1\mathrm{u}_1 + \ldots + c_p\mathrm{u}_p = 0 \cdot \mathrm{b}_1 + \ldots + 0 \cdot \mathrm{b}_n = 0.$

- Since the c_i are not all zero, $\{\mathbf{u}_1, \ldots, \mathbf{u}_p\}$ is linearly dependent.
- Theorem 9 implies that if a vector space V has a basis $\mathcal{B} = \{\mathrm{b}_1, \ldots, \mathrm{b}_n\}$, then each linearly independent set in V has no more than n vectors.

© 2016 Pearson Education, Inc.

DIMENSION OF A VECTOR SPACE

- **Theorem 10:** If a vector space V has a basis of n vectors, then every basis of V must consist of exactly n vectors.

- **Proof:** Let $\mathcal{B}_1$ be a basis of n vectors and $\mathcal{B}_2$ be any other basis (of V).
- Since $\mathcal{B}_1$ is a basis and $\mathcal{B}_2$ is linearly independent, $\mathcal{B}_2$ has no more than n vectors, by Theorem 9.
- Also, since $\mathcal{B}_2$ is a basis and $\mathcal{B}_1$ is linearly independent, $\mathcal{B}_2$ has at least n vectors.
- Thus $\mathcal{B}_2$ consists of exactly n vectors.

© 2016 Pearson Education, Inc.

DIMENSION OF A VECTOR SPACE

- **Definition:** If V is spanned by a finite set, then V is said to be **finite-dimensional**, and the **dimension** of V, written as dim V, is the number of vectors in a basis for V. The dimension of the zero vector space **{0}** is defined to be zero. If V is not spanned by a finite set, then V is said to be **infinite-dimensional**.
- **Example 3:** Find the dimension of the subspace

$$H = \left\{ \begin{bmatrix} a - 3b + 6c \\ 5a + 4d \\ b - 2c - d \\ 5d \end{bmatrix} : a, b, c, d \text{ in } \mathbb{R} \right\}$$

© 2016 Pearson Education, Inc.

DIMENSION OF A VECTOR SPACE

- H is the set of all linear combinations of the vectors

$$\mathbf{v}_1 = \begin{bmatrix} 1 \\ 5 \\ 0 \\ 0 \end{bmatrix}, \mathbf{v}_2 = \begin{bmatrix} -3 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \mathbf{v}_3 = \begin{bmatrix} 6 \\ 0 \\ -2 \\ 0 \end{bmatrix}, \mathbf{v}_4 = \begin{bmatrix} 0 \\ 4 \\ -1 \\ 5 \end{bmatrix}$$

- Clearly, $\mathbf{v}_1 \neq 0$, $\mathbf{v}_2$ is not a multiple of $\mathbf{v}_1$, but $\mathbf{v}_3$ is a multiple of $\mathbf{v}_2$.
- By the Spanning Set Theorem, we may discard $\mathbf{v}_3$ and still have a set that spans H.

© 2016 Pearson Education, Inc.

SUBSPACES OF A FINITE-DIMENSIONAL SPACE

- Finally, $\mathbf{v}_4$ is not a linear combination of $\mathbf{v}_1$ and $\mathbf{v}_2$.
- So $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_4\}$ is linearly independent and hence is a basis for H.
- Thus $\dim H = 3$.

- **Theorem 11:** Let H be a subspace of a finite-dimensional vector space V. Any linearly independent set in H can be expanded, if necessary, to a basis for H. Also, H is finite-dimensional and

$$\dim H \leq \dim V$$

© 2016 Pearson Education, Inc.

SUBSPACES OF A FINITE-DIMENSIONAL SPACE

- **Proof:** If $H = \{0\}$, then certainly $\dim H = 0 \leq \dim V$.
- Otherwise, let $S = \{\mathbf{u}_1, \ldots, \mathbf{u}_k\}$ be any linearly independent set in H.
- If S spans H, then S is a basis for H.
- Otherwise, there is some $\mathbf{u}_{k+1}$ in H that is not in Span S.

© 2016 Pearson Education, Inc.

SUBSPACES OF A FINITE-DIMENSIONAL SPACE

- But then $\{\mathrm{u}_1, \ldots, \mathrm{u}_k, \mathrm{u}_{k+1}\}$ will be linearly independent, because no vector in the set can be a linear combination of vectors that precede it (by Theorem 4).

- So long as the new set does not span *H*, we can continue this process of expanding *S* to a larger linearly independent set in *H*.

- But the number of vectors in a linearly independent expansion of *S* can never exceed the dimension of *V*, by Theorem 9.

© 2016 Pearson Education, Inc.

THE BASIS THEOREM

- So eventually the expansion of S will span H and hence will be a basis for H, and $\dim H \leq \dim V$.

- **Theorem 12:** Let V be a p-dimensional vector space, $p \geq 1$. Any linearly independent set of exactly p elements in V is automatically a basis for V. Any set of exactly p elements that spans V is automatically a basis for V.

- **Proof:** By Theorem 11, a linearly independent set S of p elements can be extended to a basis for V.

© 2016 Pearson Education, Inc.

THE BASIS THEOREM

- But that basis must contain exactly p elements, since $\dim V = p$.
- So S must already be a basis for V.
- Now suppose that S has p elements and spans V.
- Since V is nonzero, the Spanning Set Theorem implies that a subset S' of S is a basis of V.
- Since $\dim V = p$, S' must contain p vectors.
- Hence $S = S'$.

© 2016 Pearson Education, Inc.

THE DIMENSIONS OF NUL *A* AND COL *A*

- Let A be an $m \times n$ matrix, and suppose the equation $A\mathbf{x} = \mathbf{0}$ has k free variables.

- A spanning set for Nul A will produce exactly k linearly independent vectors—say, $\mathbf{u}_1, \ldots, \mathbf{u}_k$—one for each free variable.

- So $\{\mathbf{u}_1, \ldots, \mathbf{u}_k\}$ is a basis for Nul A, and the number of free variables determines the size of the basis.

© 2016 Pearson Education, Inc.

DIMENSIONS OF NUL *A* AND COL *A*

- Thus, the dimension of Nul A is the number of free variables in the equation $A\mathrm{x} = 0$, and the dimension of Col A is the number of pivot columns in A.

- **Example 5:** Find the dimensions of the null space and the column space of

$$A = \begin{bmatrix} -3 & 6 & -1 & 1 & -7 \\ 1 & -2 & 2 & 3 & -1 \\ 2 & -4 & 5 & 8 & -4 \end{bmatrix}$$

© 2016 Pearson Education, Inc.

DIMENSIONS OF NUL *A* AND COL *A*

- **Solution:** Row reduce the augmented matrix $\begin{bmatrix} A & 0 \end{bmatrix}$ to echelon form:

$$\begin{bmatrix} 1 & -2 & 2 & 3 & -1 & 0 \\ 0 & 0 & 1 & 2 & -2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

- There are three free variable—x_2, x_4 and x_5.
- Hence the dimension of Nul A is 3.
- Also dim Col $A = 2$ because A has two pivot columns.

© 2016 Pearson Education, Inc.

4 Vector Spaces

4.6

RANK

PEARSON

© 2016 Pearson Education, Inc.

THE ROW SPACE

- If A is an $m \times n$ matrix, each row of A has n entries and thus can be identified with a vector in $\mathbb{R}^n$.
- The set of all linear combinations of the row vectors is called the **row space** of A and is denoted by Row A.
- Each row has n entries, so Row A is a subspace of $\mathbb{R}^n$.
- Since the rows of A are identified with the columns of A^T, we could also write Col A^T in place of Row A.

© 2016 Pearson Education, Inc.

THE ROW SPACE

- **Theorem 13:** If two matrices A and B are row equivalent, then their row spaces are the same. If B is in echelon form, the nonzero rows of B form a basis for the row space of A as well as for that of B.

- **Proof:** If B is obtained from A by row operations, the rows of B are linear combinations of the rows of A.

- It follows that any linear combination of the rows of B is automatically a linear combination of the rows of A.

© 2016 Pearson Education, Inc.

THE ROW SPACE

- Thus the row space of B is contained in the row space of A.
- Since row operations are reversible, the same argument shows that the row space of A is a subset of the row space of B.
- So the two row spaces are the same.

© 2016 Pearson Education, Inc.

THE ROW SPACE

- If B is in echelon form, its nonzero rows are linearly independent because no nonzero row is a linear combination of the nonzero rows below it. (Apply Theorem 4 to the nonzero rows of B in reverse order, with the first row last).

- Thus the nonzero rows of B form a basis of the (common) row space of B and A.

© 2016 Pearson Education, Inc.

THE ROW SPACE

- **Example 2:** Find bases for the row space, the column space, and the null space of the matrix

$$\begin{bmatrix} -2 & -5 & 8 & 0 & -17 \\ 1 & 3 & -5 & 1 & 5 \\ 3 & 11 & -19 & 7 & 1 \\ 1 & 7 & -13 & 5 & -3 \end{bmatrix}$$

- **Solution:** To find bases for the row space and the column space, row reduce A to an echelon form:

© 2016 Pearson Education, Inc.

THE ROW SPACE

$$A \sim B = \begin{bmatrix} 1 & 3 & -5 & 1 & 5 \\ 0 & 1 & -2 & 2 & -7 \\ 0 & 0 & 0 & -4 & 20 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

- By Theorem 13, the first three rows of B form a basis for the row space of A (as well as for the row space of B).

- Thus

  Basis for Row A: $\{(1,3,-5,1,5),(0,1,-2,2,-7),(0,0,0,-4,20)\}$

© 2016 Pearson Education, Inc.

THE ROW SPACE

- For the column space, observe from B that the pivots are in columns 1, 2, and 4.
- Hence columns 1, 2, and 4 of A (not B) form a basis for Col A:

$$\text{Basis for Col } A: \left\{ \begin{bmatrix} -2 \\ 1 \\ 3 \\ 1 \end{bmatrix}, \begin{bmatrix} -5 \\ 3 \\ 11 \\ 7 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 7 \\ 5 \end{bmatrix} \right\}$$

- Notice that any echelon form of A provides (in its nonzero rows) a basis for Row A and also identifies the pivot columns of A for Col A.

© 2016 Pearson Education, Inc.

THE ROW SPACE

- However, for Nul A, we need the *reduced echelon form*.

- Further row operations on B yield

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

© 2016 Pearson Education, Inc.

THE ROW SPACE

- The equation $A\mathrm{x} = 0$ is equivalent to $C\mathrm{x} = 0$, that is,

$$x_1 + x_3 + x_5 = 0$$
$$x_2 - 2x_3 + 3x_5 = 0$$
$$x_4 - 5x_5 = 0$$

- So $x_1 = -x_3 - x_5$, $x_2 = 2x_3 - 3x_5$, $x_4 = 5x_5$, with x_3 and x_5 free variables.

© 2016 Pearson Education, Inc.

THE ROW SPACE

- The calculations show that

$$\text{Basis for Nul } A: \left\{ \begin{bmatrix} -1 \\ 2 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ -3 \\ 0 \\ 5 \\ 1 \end{bmatrix} \right\}$$

- Observe that, unlike the basis for Col A, the bases for Row A and Nul A have no simple connection with the entries in A itself.

© 2016 Pearson Education, Inc.

THE RANK THEOREM

- **Definition:** The **rank** of A is the dimension of the column space of A.
- Since Row A is the same as Col A^T, the dimension of the row space of A is the rank of A^T.
- The dimension of the null space is sometimes called the **nullity** of A.
- **Theorem 14:** The dimensions of the column space and the row space of an $m \times n$ matrix A are equal. This common dimension, the rank of A, also equals the number of pivot positions in A and satisfies the equation

$$\text{rank } A + \dim \text{Nul } A = n$$

© 2016 Pearson Education, Inc.

THE RANK THEOREM

- **Proof:** By Theorem 6, rank A is the number of pivot columns in A.
- Equivalently, rank A is the number of pivot positions in an echelon form B of A.
- Since B has a nonzero row for each pivot, and since these rows form a basis for the row space of A, the rank of A is also the dimension of the row space.
- The dimension of Nul A equals the number of free variables in the equation $A\mathrm{x} = 0$.
- Expressed another way, the dimension of Nul A is the number of columns of A that are *not* pivot columns.

© 2016 Pearson Education, Inc.

THE RANK THEOREM

- (It is the number of these columns, not the columns themselves, that is related to Nul A).

- Obviously,

$$\left\{\begin{matrix}\text{number of}\\ \text{pivot columns}\end{matrix}\right\}+\left\{\begin{matrix}\text{number of}\\ \text{nonpivot columns}\end{matrix}\right\}=\left\{\begin{matrix}\text{number of}\\ \text{columns}\end{matrix}\right\}$$

- This proves the theorem.

© 2016 Pearson Education, Inc.

THE RANK THEOREM

- **Example 3:**
  a. If A is a 7×9 matrix with a two-dimensional null space, what is the rank of A?
  b. Could a 6×9 matrix have a two-dimensional null space?
- **Solution:**
  a. Since A has 9 columns, $(\text{rank } A) + 2 = 9$, and hence rank $A = 7$.
  b. No. If a 6×9 matrix, call it B, has a two-dimensional null space, it would have to have rank 7, by the Rank Theorem.

© 2016 Pearson Education, Inc.

THE INVERTIBLE MATRIX THEOREM (CONTINUED)

- But the columns of B are vectors in $\mathbb{R}^6$, and so the dimension of Col B cannot exceed 6; that is, rank B cannot exceed 6.

- **Theorem:** Let A be an $n \times n$ matrix. Then the following statements are each equivalent to the statement that A is an invertible matrix.

  m. The columns of A form a basis of $\mathbb{R}^n$.

  n. Col $A = \mathbb{R}^n$

  o. Dim Col $A = n$

  p. rank $A = n$

© 2016 Pearson Education, Inc.

RANK AND THE INVERTIBLE MATRIX THEOREM

q. $\text{Nul } A = \{0\}$

r. $\text{Dim Nul } A = 0$

- **Proof:** Statement (m) is logically equivalent to statements (e) and (h) regarding linear independence and spanning.

- The other five statements are linked to the earlier ones of the theorem by the following chain of almost trivial implications:

$$(\text{g}) \Rightarrow (\text{n}) \Rightarrow (\text{o}) \Rightarrow (\text{p}) \Rightarrow (\text{r}) \Rightarrow (\text{q}) \Rightarrow (\text{d})$$

© 2016 Pearson Education, Inc.

RANK AND THE INVERTIBLE MATRIX THEOREM

- Statement (g), which says that the equation $A\mathrm{x} = \mathrm{b}$ has at least one solution for each **b** in $\mathbb{R}^n$, implies (n), because Col A is precisely the set of all **b** such that the equation $A\mathrm{x} = \mathrm{b}$ is consistent.

- The implications $(\mathrm{n}) \Rightarrow (\mathrm{o}) \Rightarrow (\mathrm{p})$ follow from the definitions of dimension and rank.

- If the rank of A is n, the number of columns of A, then dim Nul $A = 0$, by the Rank Theorem, and so Nul $A = \{0\}$.

© 2016 Pearson Education, Inc.

RANK AND THE INVERTIBLE MATRIX THEOREM

- Thus $(\mathrm{p}) \Rightarrow (\mathrm{r}) \Rightarrow (\mathrm{q})$.

- Also, (q) implies that the equation $A\mathrm{x} = 0$ has only the trivial solution, which is statement (d).

- Since statements (d) and (g) are already known to be equivalent to the statement that A is invertible, the proof is complete.

© 2016 Pearson Education, Inc.

4 Vector Spaces

4.7

CHANGE OF BASIS

PEARSON

CHANGE OF BASIS

- When a basis B is chosen for an n-dimensional vector space V, the associated coordinate mapping onto $\square^n$ provides a coordinate system for V. Each $\mathbf{x}$ in V is identified uniquely by is B-coordinate vector $[\mathbf{x}]_{\mathrm{B}}$.

- In some applications, a problem is described initially using a basis B, But the problem's solution is aided by changing B to a new basis X. Each vector is assigned a new X-coordinate vector. In this section, we study how $[\mathbf{x}]_{\mathrm{X}}$ and $[\mathbf{x}]_{\mathrm{B}}$ are related for each $\mathbf{x}$ in V

© 2012 Pearson Education, Inc.

CHANGE OF BASIS

- To visualize the problem, consider the two coordinate systems in Fig.1. In Fig. 1(a), $\mathbf{x} = 3\mathbf{b}_1 + \mathbf{b}_2$,while in Fig. 1(b), the same $\mathbf{x}$ is shown as $\mathbf{x} = 6\mathbf{c}_1 + 4\mathbf{c}_2$.

- That is ,

$$[\mathbf{x}]_B = \begin{bmatrix} 3 \\ 1 \end{bmatrix} \text{ and } [\mathbf{x}]_X = \begin{bmatrix} 6 \\ 4 \end{bmatrix}$$

© 2012 Pearson Education, Inc.

CHANGE OF BASIS

- Our problem is to find the connection between the two coordinate vectors. Example 1 shows how to do this, provided we know how $\mathbf{b}_1$ and $\mathbf{b}_2$ are formed from $\mathbf{c}_1$ and $\mathbf{c}_2$.

- Figure 1 two systems for the same vector space.

© 2012 Pearson Education, Inc.

CHANGE OF BASIS

- **Example 1**: Consider two base $\mathrm{B} = \{\mathbf{b}_1, \mathbf{b}_2\}$ and $\mathrm{X} = \{\mathbf{c}_1, \mathbf{c}_2\}$ for a vector space V, such that

$$\mathbf{b}_1 = 4\mathbf{c}_1 + \mathbf{c}_2 \text{ and } \mathbf{b}_2 = -6\mathbf{c}_1 + \mathbf{c}_2 \quad (1)$$

suppose

$$\mathbf{x} = 3\mathbf{b}_1 + \mathbf{b}_2 \qquad (2)$$

That is ,suppose $[\mathbf{x}]_{\mathrm{B}} = \begin{bmatrix} 3 \\ 1 \end{bmatrix}$. Find $[\mathbf{x}]_{\mathrm{X}}$.

© 2012 Pearson Education, Inc.

CHANGE OF BASIS

- **Solution:** Apply the coordinate mapping determined by X to **x** in (2). Since the coordinate mapping is a linear transformation,

$$[\mathbf{x}]_X = [3\mathbf{b}_1 + \mathbf{b}_2]_X$$
$$= 3[\mathbf{b}_1]_X + [\mathbf{b}_2]_X$$

We can write this vector equation as a matrix equation , using the vectors in the linear combination as the column of a matrix :

$$[\mathbf{x}]_X = \begin{bmatrix} [\mathbf{b}_1]_X & [\mathbf{b}_2]_X \end{bmatrix} \begin{bmatrix} 3 \\ 1 \end{bmatrix} \qquad (3)$$

© 2012 Pearson Education, Inc.

CHANGE OF BASIS

- This formula gives $[\mathbf{x}]_X$, once we know the columns of the matrix. From (1),

$$[\mathbf{b}_1]_X = \begin{bmatrix} 4 \\ 1 \end{bmatrix} \text{ and } [\mathbf{b}_2]_X = \begin{bmatrix} -6 \\ 1 \end{bmatrix}$$

Thus (3) provides the solution:

$$[\mathbf{x}]_X = \begin{bmatrix} 4 & -6 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 3 \\ 1 \end{bmatrix} = \begin{bmatrix} 6 \\ 4 \end{bmatrix}$$

The X-coordinates of $\mathbf{x}$ match those of $\mathbf{x}$ in Fig. 1.

The argument used to derive formula (3) can be generalized to yield the following result. (See Exercises 15 and 16.)

© 2012 Pearson Education, Inc.

CHANGE OF BASIS

- **THEOREM 15**

Let B=$\{\mathbf{b}_1, \ldots, \mathbf{b}_n\}$ and X =$\{\mathbf{c}_1, \ldots, \mathbf{c}_n\}$ be bases of a vector space V. The there is a unique $n \times n$ matrix $\underset{\mathrm{X}\leftarrow\mathrm{B}}{P}$ such that

$$[\mathbf{x}]_{\mathrm{X}} = \underset{\mathrm{X}\leftarrow\mathrm{B}}{P}[\mathbf{x}]_{\mathrm{B}} \qquad (4)$$

The column of $\underset{\mathrm{X}\leftarrow\mathrm{B}}{P}$ are the X-coordinate vectors of the vectors in the basis B.

That is

$$\underset{\mathrm{X}\leftarrow\mathrm{B}}{P} = \left[[\mathbf{b}_1]_{\mathrm{X}}\ [\mathbf{b}_2]_{\mathrm{X}} \cdots [\mathbf{b}_n]_{\mathrm{X}}\right] \qquad (5)$$

© 2012 Pearson Education, Inc.

THEOREM 15

- The matrix $\underset{X \leftarrow B}{P}$ in the Theorem 15 is called **change-of-coordinate matrix from B to X**. Multiplication by $\underset{X \leftarrow B}{P}$ converts B-coordinates into X-coordinates. Figure 2 illustrates the change-of coordinates equation (4).
-

- Figure 2 two coordinate systems for V

© 2012 Pearson Education, Inc.

THEOREM 15

- The columns of $\underset{\mathcal{C}\leftarrow\mathcal{B}}{\boldsymbol{P}}$ are linearly independent because they are the coordinate vectors of the linearly independent set B. Since $\underset{\mathcal{C}\leftarrow\mathcal{B}}{\boldsymbol{P}}$ is square, it must be invertible, by the Invertible Matrix Theorem. Left-multiplying both side of equation (4) by $(\underset{\mathcal{C}\leftarrow\mathcal{B}}{\boldsymbol{P}})^{-1}$ yeilds

$$(\underset{\mathcal{C}\leftarrow\mathcal{B}}{\boldsymbol{P}})^{-1}[\mathbf{x}]_{\mathrm{X}} = [\mathbf{x}]_{\mathrm{B}}$$

© 2012 Pearson Education, Inc.

THEOREM 15

- Thus $\left(\underset{X\leftarrow B}{P}\right)^{-1}$ is the matrix that converts X-coordinates into B-coordinates.

  That is

$$\left(\underset{X\leftarrow B}{P}\right)^{-1} = \underset{B\leftarrow X}{P} \qquad (6)$$

© 2012 Pearson Education, Inc.

CHANGE OF BASIS in $\mathbb{R}^n$

- If $\mathrm{B} = \{\mathbf{b}_1, \dots, \mathbf{b}_n\}$ and ε is the standard basis $\{\mathbf{e}_1, \dots, \mathbf{e}_\mathrm{n}\}$ in $\mathbb{R}^n$, then $[\mathbf{b}_1]_\varepsilon = \mathbf{b}_1$, and likewise for the vectors in B. In this case, $\underset{\varepsilon \leftarrow \mathrm{B}}{P}$ is the same as the change-of-coordinate matrix P_B introduced in section 4.4, namely, $P_\mathrm{B} = [\mathbf{b}_1 \quad \mathbf{b}_2 \dots \ \mathbf{b}_n]$
- To change coordinates between two nonstandard bases in $\mathbb{R}^n$, we need Theorem 15 . The theorem shows that to solve the change-of-basis problem .we need coordinate vectors of the old basis relative to the new basis.

© 2012 Pearson Education, Inc.

CHANGE OF BASIS in $\mathbb{R}^n$

- **Example 2**

  let $\mathbf{b}_1 = \begin{bmatrix} -9 \\ 1 \end{bmatrix}, \mathbf{b}_2 = \begin{bmatrix} -5 \\ 1 \end{bmatrix}, \boldsymbol{c}_1 = \begin{bmatrix} 1 \\ -4 \end{bmatrix}, \quad \boldsymbol{c}_2 = \begin{bmatrix} 3 \\ -5 \end{bmatrix}$, and consider the bases for $\mathbb{R}^2$ given by $\mathrm{B} = \{\mathbf{b}_1, \mathbf{b}_2\}$ and $\mathrm{X} = \{\boldsymbol{c}_1, \boldsymbol{c}_2\}$. Find the change-of-coordinate matrix from B *to* X.

- Solution: the matrix $\underset{\mathrm{X} \leftarrow \mathrm{B}}{P}$ involves the X-coordinate vectors of $\mathbf{b}_1$ and $\mathbf{b}_2$ let $\mathbf{b}_1 = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}, \mathbf{b}_2 = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$ .Then,by definition,

$$[\,\mathbf{c}_1 \quad \mathbf{c}_2]\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \mathbf{b}_1 \text{ and } [\,\boldsymbol{c}_1 \quad \boldsymbol{c}_2]\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \boldsymbol{b}_2$$

© 2012 Pearson Education, Inc.

CHANGE OF BASIS in $\mathbb{R}^n$

- To solve both systems simultaneously . augment the coefficient matrix with $\mathbf{b}_1$ and $\mathbf{b}_2$,and row reduce:

$$[\mathbf{c}_1 \ \ \mathbf{c}_2 \ \vdots \ \mathbf{b}_1 \ \ \mathbf{b}_2]$$

$$= \left[\begin{array}{cc:cc} 1 & 3 & -9 & -5 \\ -4 & -5 & 1 & -1 \end{array}\right] \sim \left[\begin{array}{cc:cc} 1 & 0 & 6 & 4 \\ 0 & 1 & -5 & -3 \end{array}\right] \quad (7)$$

© 2012 Pearson Education, Inc.

CHANGE OF BASIS in $\mathbb{R}^n$

- The desired change-of-coordinate matrix is therefore

$$\underset{X \leftarrow B}{P} = [[\mathbf{b}_1]_{\mathcal{C}} \quad [\mathbf{b}_2]_{\mathcal{C}}] = \begin{bmatrix} 6 & 4 \\ -5 & -3 \end{bmatrix}$$

- Observe that the matrix $\underset{X \leftarrow B}{P}$ in Example 2 already appeared in (7). This is not surprising because the first column of $\underset{X \leftarrow B}{P}$ results from row reducing $[\mathbf{c}_1 \;\; \mathbf{c}_2 \;\vdots\; \mathbf{b}_1]$ to $[I \;\vdots\; [\mathbf{b}_1]_{\mathcal{C}}]$, and similarly for the second column of $\underset{\mathcal{C} \leftarrow \mathcal{B}}{P}$ .Thus

$$\boxed{[\mathbf{c}_1 \;\; \mathbf{c}_2 \;\vdots\; \mathbf{b}_1 \;\; \mathbf{b}_2] \sim \left[I \;\vdots\; \underset{X \leftarrow B}{P} \right]}$$

- An analogous procedure work for finding the change-of-coordinates matrix between any basis in $\mathbb{R}^n$.

© 2012 Pearson Education, Inc.

CHANGE OF BASIS in $\mathbb{R}^n$
Example 3 solution *P*

- **Example 3** let $\mathbf{b}_1 = \begin{bmatrix} 1 \\ -3 \end{bmatrix}$, $\mathbf{b}_2 = \begin{bmatrix} -2 \\ 4 \end{bmatrix}$, $\mathbf{c}_1 = \begin{bmatrix} -7 \\ 9 \end{bmatrix}$, $\mathbf{c}_2 = \begin{bmatrix} -5 \\ 7 \end{bmatrix}$, and consider the bases for R^2 given by $\mathrm{B} = \{\mathbf{b}_1, \mathbf{b}_2\}$ and $\mathrm{X} = \{\mathbf{c}_1, \mathbf{c}_2\}$.

a. Find the change-of-coordinates matrix from X to B.

b. Find the change-of-coordinates matrix from B to X.

© 2012 Pearson Education, Inc.

CHANGE OF BASIS in $\mathbb{R}^n$
Example 3 solution P

- a. Notice that $\underset{\mathrm{B}\leftarrow\mathrm{X}}{P}$ is needed rather than $\underset{\mathrm{X}\leftarrow\mathrm{B}}{P}$ and compute $\begin{bmatrix}\mathbf{b}_1 & \mathbf{b}_2 \vdots \mathbf{c}_1 & \mathbf{c}_2\end{bmatrix} = \begin{bmatrix} 1 & -2 \vdots -7 & -5 \\ -3 & 4 \vdots 9 & 7 \end{bmatrix}$

$$\sim \begin{bmatrix} 1 & 0 & \vdots & 5 & 3 \\ 0 & 1 & \vdots & 6 & 4 \end{bmatrix}$$

- So $\underset{\mathrm{B}\leftarrow\mathrm{X}}{P} = \begin{bmatrix} 5 & 3 \\ 6 & 4 \end{bmatrix}$.
- b. By part (a) and property (6) above (with B and X interchanged),

$$\underset{\mathrm{X}\leftarrow\mathrm{B}}{P} = \left(\underset{\mathrm{B}\leftarrow\mathrm{X}}{P}\right)^{-1} = \frac{1}{2}\begin{bmatrix} 4 & -3 \\ -6 & 5 \end{bmatrix} = \begin{bmatrix} 2 & -3/2 \\ -3 & 5/2 \end{bmatrix}$$

© 2012 Pearson Education, Inc.

CHANGE OF BASIS in $\mathbb{R}^n$

- Another description of the change-of-coordinates matrix $\underset{\mathrm{X}\leftarrow\mathrm{B}}{P}$ uses the change-of-coordinate matrices P_{B} and P_{X} that convert B-coordinates and X-coordinates, respectively, into standard coordinates. Recall that for each $\mathbf{x}$ in R^n,

$$P_{\mathrm{B}}[\mathbf{x}]_{\mathrm{B}} = \mathbf{x}, P_{\mathrm{X}}[\mathbf{x}]_{\mathrm{X}} = \mathbf{x}, and\ [\mathbf{x}]_{\mathrm{X}} = P_{\mathrm{X}}{}^{-1}\mathbf{x}$$

Thus

$$[\mathbf{x}]_{\mathrm{X}}=P_{\mathrm{X}}{}^{-1}\mathbf{x}=P_{\mathrm{X}}{}^{-1}P_{\mathrm{B}}[\mathbf{x}]_{\mathrm{B}}$$

© 2012 Pearson Education, Inc.

CHANGE OF BASIS in $\mathbb{R}^n$

- In R^n, the change-of-coordinates matrix $\underset{\mathrm{X}\leftarrow\mathrm{B}}{P}$ may be computed as $P_{\mathrm{X}}{}^{-1}P_{\mathrm{B}}$. Actually, for matrices larger than 2×2, an algorithm analogous to the one in Example 3 is faster than computing $P_{\mathrm{X}}{}^{-1}$ and then $P_{\mathrm{X}}{}^{-1}P_{\mathrm{B}}$. See Exercise 12 in Section 2.2.

© 2012 Pearson Education, Inc.

5 Eigenvalues and Eigenvectors

5.1

EIGENVECTORS AND EIGENVALUES

© 2016 Pearson Education, Ltd.

EIGENVECTORS AND EIGENVALUES

- **Definition:** An **eigenvector** of an $n \times n$ matrix A is a nonzero vector **x** such that $A\mathrm{x} = \lambda\mathrm{x}$ for some scalar λ. A scalar λ is called an **eigenvalue** of A if there is a nontrivial solution **x** of $A\mathrm{x} = \lambda\mathrm{x}$; such an **x** is called an *eigenvector corresponding to λ*.
- λ is an eigenvalue of an $n \times n$ matrix A if and only if the equation

$$(A - \lambda I)\mathrm{x} = 0 \qquad (3)$$

  has a nontrivial solution.
- The set of *all* solutions of (3) is just the null space of the matrix $A - \lambda I$.

© 2016 Pearson Education, Ltd.

EIGENVECTORS AND EIGENVALUES

- So this set is a *subspace* of $\mathbb{R}^n$ and is called the **eigenspace** of A corresponding to λ.

- The eigenspace consists of the zero vector and all the eigenvectors corresponding to λ.

- **Example 3:** Show that 7 is an eigenvalue of matrix

$A = \begin{bmatrix} 1 & 6 \\ 5 & 2 \end{bmatrix}$ and find the corresponding eigenvectors.

© 2016 Pearson Education, Ltd.

EIGENVECTORS AND EIGENVALUES

- **Solution:** The scalar 7 is an eigenvalue of A if and only if the equation

$$A\mathrm{x} = 7\mathrm{x} \qquad (1)$$

has a nontrivial solution.

- But (1) is equivalent to $A\mathrm{x} - 7\mathrm{x} = 0$, or

$$(A - 7I)\mathrm{x} = 0 \qquad (2)$$

- To solve this homogeneous equation, form the matrix

$$A - 7I = \begin{bmatrix} 1 & 6 \\ 5 & 2 \end{bmatrix} - \begin{bmatrix} 7 & 0 \\ 0 & 7 \end{bmatrix} = \begin{bmatrix} -6 & 6 \\ 5 & -5 \end{bmatrix}$$

© 2016 Pearson Education, Ltd.

EIGENVECTORS AND EIGENVALUES

- The columns of $A - 7I$ are obviously linearly dependent, so (2) has nontrivial solutions.
- To find the corresponding eigenvectors, use row operations:

$$\begin{bmatrix} -6 & 6 & 0 \\ 5 & -5 & 0 \end{bmatrix} \sim \begin{bmatrix} 1 & -1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

- The general solution has the form $x_2 \begin{bmatrix} 1 \\ 1 \end{bmatrix}$.
- Each vector of this form with $x_2 \neq 0$ is an eigenvector corresponding to $\lambda = 7$.

© 2016 Pearson Education, Ltd.

EIGENVECTORS AND EIGENVALUES

- **Example 4:** Let $A = \begin{bmatrix} 4 & -1 & 6 \\ 2 & 1 & 6 \\ 2 & -1 & 8 \end{bmatrix}$. An eigenvalue of A is 2. Find a basis for the corresponding eigenspace.
- **Solution:** Form

$$A - 2I = \begin{bmatrix} 4 & -1 & 6 \\ 2 & 1 & 6 \\ 2 & -1 & 8 \end{bmatrix} - \begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix} = \begin{bmatrix} 2 & -1 & 6 \\ 2 & -1 & 6 \\ 2 & -1 & 6 \end{bmatrix}$$

and row reduce the augmented matrix for $(A - 2I)\mathrm{x} = 0$.

© 2016 Pearson Education, Ltd.

EIGENVECTORS AND EIGENVALUES

$$\begin{bmatrix} 2 & -1 & 6 & 0 \\ 2 & -1 & 6 & 0 \\ 2 & -1 & 6 & 0 \end{bmatrix} \sim \begin{bmatrix} 2 & -1 & 6 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

- At this point, it is clear that 2 is indeed an eigenvalue of A because the equation $(A-2I)\mathrm{x}=0$ has free variables.
- The general solution is

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = x_2 \begin{bmatrix} 1/2 \\ 1 \\ 0 \end{bmatrix} + x_3 \begin{bmatrix} -3 \\ 0 \\ 1 \end{bmatrix}, \; x_2 \text{ and } x_3 \text{ free.}$$

© 2016 Pearson Education, Ltd.

EIGENVECTORS AND EIGENVALUES

- The eigenspace, shown in the following figure, is a two-dimensional subspace of $\mathbb{R}^3$.A basis is

$$\left\{ \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix}, \begin{bmatrix} -3 \\ 0 \\ 1 \end{bmatrix} \right\}$$

A acts as a dilation on the eigenspace.

© 2016 Pearson Education, Ltd.

EIGENVECTORS AND EIGENVALUES

- **Theorem 1:** The eigenvalues of a triangular matrix are the entries on its main diagonal.
- **Proof:** For simplicity, consider the 3×3 case.
- If A is upper triangular, the $A - \lambda I$ has the form

$$A - \lambda I = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ 0 & a_{22} & a_{23} \\ 0 & 0 & a_{33} \end{bmatrix} - \begin{bmatrix} \lambda & 0 & 0 \\ 0 & \lambda & 0 \\ 0 & 0 & \lambda \end{bmatrix}$$

$$= \begin{bmatrix} a_{11} - \lambda & a_{12} & a_{13} \\ 0 & a_{22} - \lambda & a_{23} \\ 0 & 0 & a_{33} - \lambda \end{bmatrix}$$

© 2016 Pearson Education, Ltd.

EIGENVECTORS AND EIGENVALUES

- The scalar λ is an eigenvalue of A if and only if the equation $(A-\lambda I)\mathrm{x}=0$ has a nontrivial solution, that is, if and only if the equation has a free variable.

- Because of the zero entries in $A-\lambda I$, it is easy to see that $(A-\lambda I)\mathrm{x}=0$ has a free variable if and only if at least one of the entries on the diagonal of $A-\lambda I$ is zero.

- This happens if and only if λ equals one of the entries a_{11}, a_{22}, a_{33} in A.

© 2016 Pearson Education, Ltd.

EIGENVECTORS AND EIGENVALUES

- **Theorem 2:** If $\mathbf{v}_1, \ldots, \mathbf{v}_r$ are eigenvectors that correspond to distinct eigenvalues $\lambda_1, \ldots, \lambda_r$ of an $n \times n$ matrix A, then the set $\{\mathbf{v}_1, \ldots, \mathbf{v}_r\}$ is linearly independent.
- **Proof:** Suppose $\{\mathbf{v}_1, \ldots, \mathbf{v}_r\}$ is linearly dependent.
- Since $\mathbf{v}_1$ is nonzero, Theorem 7 in Section 1.7 says that one of the vectors in the set is a linear combination of the preceding vectors.
- Let p be the least index such that v_{p+1} is a linear combination of the preceding (linearly independent) vectors.

© 2016 Pearson Education, Ltd.

EIGENVECTORS AND EIGENVALUES

- Then there exist scalars $c_1, \ldots, c_p$ such that

$$c_1\mathbf{v}_1 + \cdots + c_p\mathbf{v}_p = \mathbf{v}_{p+1} \tag{5}$$

- Multiplying both sides of (5) by A and using the fact that $Av_k = \lambda_k v_k$ for each k, we obtain

$$c_1 A\mathbf{v}_1 + \cdots + c_p A\mathbf{v}_p = A\mathbf{v}_{p+1}$$

$$c_1\lambda_1\mathbf{v}_1 + \cdots + c_p\lambda_p\mathbf{v}_p = \lambda_{p+1}\mathbf{v}_{p+1} \tag{6}$$

- Multiplying both sides of (5) by λ_{p+1} and subtracting the result from (6), we have

$$c_1(\lambda_1 - \lambda_{p+1})\mathbf{v}_1 + \cdots + c_p(\lambda_p - \lambda_{p+1})\mathbf{v}_p = 0 \tag{7}$$

© 2016 Pearson Education, Ltd.

EIGENVECTORS AND EIGENVALUES

- Since $\{\mathbf{v}_1, \ldots, \mathbf{v}_p\}$ is linearly independent, the weights in (7) are all zero.

- But none of the factors $\lambda_i - \lambda_{p+1}$ are zero, because the eigenvalues are distinct.

- Hence $c_i = 0$ for $i = 1, \ldots, p$.

- But then (5) says that $\mathrm{v}_{p+1} = 0$, which is impossible.

© 2016 Pearson Education, Ltd.

EIGENVECTORS AND DIFFERENCE EQUATIONS

- Hence $\{\mathbf{v}_1, \ldots, \mathbf{v}_r\}$ cannot be linearly dependent and therefore must be linearly independent.

- If A is an $n \times n$ matrix, then (8) is a *recursive* description of a sequence $\{x_k\}$ in $\mathbb{R}^n$.

$$\mathrm{x}_{k+1} = A\mathrm{x}_k \qquad (k = 0, 1, 2 \ldots) \tag{8}$$

- A **solution** of (8) is an explicit description of $\{x_k\}$ whose formula for each x_k does not depend directly on A or on the preceding terms in the sequence other than the initial term $\mathbf{x}_0$.

© 2016 Pearson Education, Ltd.

EIGENVECTORS AND DIFFERENCE EQUATIONS

- The simplest way to build a solution of (8) is to take an eigenvector $\mathbf{x}_0$ and its corresponding eigenvalue λ and let

$$\mathrm{x}_k = \lambda^k \mathrm{x}_0 \quad (k = 1, 2, \ldots) \qquad (9)$$

- This sequence is a solution because

$$A\mathrm{x}_k = A(\lambda^k \mathrm{x}_0) = \lambda^k (A\mathrm{x}_0) = \lambda^k (\lambda \mathrm{x}_0) = \lambda^{k+1} \mathrm{x}_0 = \mathrm{x}_{k+1}$$

© 2016 Pearson Education, Ltd.

5 Eigenvalues and Eigenvectors

5.2

THE CHARACTERISTIC EQUATION

PEARSON

© 2016 Pearson Education, Inc.

DETERMINANATS

- Let A be an $n \times n$ matrix, let U be any echelon form obtained from A by row replacements and row interchanges (without scaling), and let r be the number of such row interchanges.

- Then the determinant of A, written as det A, is $(-1)^r$ times the product of the diagonal entries $u_{11}, \ldots, u_{nn}$ in U.

- If A is invertible, then $u_{11}, \ldots, u_{nn}$ are all pivots (because $A \sim I_n$ and the u_{ii} have not been scaled to 1's).

© 2016 Pearson Education, Inc.

DETERMINANATS

- Otherwise, at least u_{nn} is zero, and the product $u_{11} \ldots u_{nn}$ is zero.

- Thus

$$\det A = \begin{cases} (-1)^r \cdot \begin{pmatrix} \text{product of} \\ \text{pivots in } U \end{pmatrix}, & \text{when A is invertible} \\ 0, & \text{when A is not invertible} \end{cases}$$

© 2016 Pearson Education, Inc.

DETERMINANATS

- **Example 1:** Compute det A for $A = \begin{bmatrix} 1 & 5 & 0 \\ 2 & 4 & -1 \\ 0 & -2 & 0 \end{bmatrix}$.

- **Solution:** The following row reduction uses one row interchange:

$$A \sim \begin{bmatrix} 1 & 5 & 0 \\ 0 & -6 & -1 \\ 0 & -2 & 0 \end{bmatrix} \sim \begin{bmatrix} 1 & 5 & 0 \\ 0 & -2 & 0 \\ 0 & -6 & -1 \end{bmatrix} \sim \begin{bmatrix} 1 & 5 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & -1 \end{bmatrix} = U_1$$

© 2016 Pearson Education, Inc.

DETERMINANATS

- So det A equals $(-1)^1(1)(-2)(-1) = -2$.
- The following alternative row reduction avoids the row interchange and produces a different echelon form.
- The last step adds $-1/3$ times row 2 to row 3:

$$A \sim \begin{bmatrix} 1 & 5 & 0 \\ 0 & -6 & -1 \\ 0 & -2 & 0 \end{bmatrix} \sim \begin{bmatrix} 1 & 5 & 0 \\ 0 & -6 & -1 \\ 0 & 0 & 1/3 \end{bmatrix} = U_2$$

- This time det A is $(-1)^0(1)(-6)(1/3) = -2$, the same as before.

© 2016 Pearson Education, Inc.

THE INVERTIBLE MATRIX THEOREM (CONTINUED)

- **Theorem:** Let A be an $n \times n$ matrix. Then A is invertible if and only if:
  s. The number 0 is *not* an eigenvalue of A.
  t. The determinant of A is *not* zero.

- **Theorem 3: Properties of Determinants**
- Let A and B be $n \times n$ matrices.
  a. A is invertible if and only if $\det A \neq 0$.
  b. $\det AB = (\det A)(\det B)$.
  c. $\det A^T = \det A$.

© 2016 Pearson Education, Inc.

PROPERTIES OF DETERMINANTS

d. If A is triangular, then det A is the product of the entries on the main diagonal of A.

e. 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.

© 2016 Pearson Education, Inc.

THE CHARACTERISTIC EQUATION

- Theorem 3(a) shows how to determine when a matrix of the form $A - \lambda I$ is *not* invertible.

- The scalar equation $\det(A - \lambda I) = 0$ is called the **characteristic equation** of A.

- A scalar λ is an eigenvalue of an $n \times n$ matrix A if and only if λ satisfies the characteristic equation

$$\det(A - \lambda I) = 0$$

© 2016 Pearson Education, Inc.

THE CHARACTERISTIC EQUATION

- **Example 2:** Find the characteristic equation of

$$A = \begin{bmatrix} 5 & -2 & 6 & -1 \\ 0 & 3 & -8 & 0 \\ 0 & 0 & 5 & 4 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

- **Solution:** Form $A - \lambda I$, and use Theorem 3(d):

© 2016 Pearson Education, Inc.

THE CHARACTERISTIC EQUATION

$$\det(A-\lambda I)=\det\begin{bmatrix}5-\lambda & -2 & 6 & -1\\ 0 & 3-\lambda & -8 & 0\\ 0 & 0 & 5-\lambda & 4\\ 0 & 0 & 0 & 1-\lambda\end{bmatrix}$$

$$=(5-\lambda)(3-\lambda)(5-\lambda)(1-\lambda)$$

- The characteristic equation is

$$(5-\lambda)^2(3-\lambda)(1-\lambda)=0$$

or

$$(\lambda-5)^2(\lambda-3)(\lambda-1)=0$$

© 2016 Pearson Education, Inc.

THE CHARACTERISTIC EQUATION

- Expanding the product, we can also write

$$\lambda^4 - 14\lambda^3 + 68\lambda^2 - 130\lambda + 75 = 0$$

- If A is an $n \times n$ matrix, then $\det(A - \lambda I)$ is a polynomial of degree n called the **characteristic polynomial** of A.
- The eigenvalue 5 in Example 2 is said to have *multiplicity* 2 because $(\lambda - 5)$ occurs two times as a factor of the characteristic polynomial.
- In general, the (**algebraic**) **multiplicity** of an eigenvalue λ is its multiplicity as a root of the characteristic equation.

© 2016 Pearson Education, Inc.

SIMILARITY

- If A and B are $n \times n$ matrices, then A **is similar to** B if there is an invertible matrix P such that $P^{-1}AP = B$, or, equivalently, $A = PBP^{-1}$.

- Writing Q for P^{-1}, we have $Q^{-1}BQ = A$.

- So B is also similar to A, and we say simply that A and B **are similar**.

- Changing A into $P^{-1}AP$ is called a **similarity transformation**.

© 2016 Pearson Education, Inc.

SIMILARITY

- **Theorem 4:** If $n \times n$ matrices A and B are similar, then they have the same characteristic polynomial and hence the same eigenvalues (with the same multiplicities).

- **Proof:** If $B = P^{-1}AP$ then,

$$B - \lambda I = P^{-1}AP - \lambda P^{-1}P = P^{-1}(AP - \lambda P) = P^{-1}(A - \lambda I)P$$

- Using the multiplicative property (b) in Theorem (3), we compute

$$\det(B - \lambda I) = \det\left[P^{-1}(A - \lambda I)P\right]$$

$$= \det(P^{-1}) \cdot \det(A - \lambda I) \cdot \det(P) \quad \text{----(1)}$$

© 2016 Pearson Education, Inc.

SIMILARITY

- Since $\det(P^{-1}) \cdot \det(P) = \det(P^{-1}P) = \det I = 1$, we see from equation (1) that $\det(B - \lambda I) = \det(A - \lambda I)$.

- **Warnings:**
  1. The matrices

$$\begin{bmatrix} 2 & 1 \\ 0 & 2 \end{bmatrix} \text{ and } \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}$$

     are not similar even though they have the same eigenvalues.

© 2016 Pearson Education, Inc.

SIMILARITY

2. Similarity is not the same as row equivalence. (If A is row equivalent to B, then $B = EA$ for some invertible matrix E). Row operations on a matrix usually change its eigenvalues.

© 2016 Pearson Education, Inc.

5 Eigenvalues and Eigenvectors

5.3

DIAGONALIZATION

PEARSON

© 2016 Pearson Education, Inc.

DIAGONALIZATION

- **Example 2:** Let $A = \begin{bmatrix} 7 & 2 \\ -4 & 1 \end{bmatrix}$. Find a formula for A^k, given that $A = PDP^{-1}$, where

$$P = \begin{bmatrix} 1 & 1 \\ -1 & -2 \end{bmatrix} \quad \text{and} \quad D = \begin{bmatrix} 5 & 0 \\ 0 & 3 \end{bmatrix}$$

- **Solution:** The standard formula for the inverse of a 2×2 matrix yields

$$P^{-1} = \begin{bmatrix} 2 & 1 \\ -1 & -1 \end{bmatrix}$$

© 2016 Pearson Education, Inc.

DIAGONALIZATION

- Then, by associativity of matrix multiplication,

$$A^2 = (PDP^{-1})(PDP^{-1}) = PD\underbrace{(P^{-1}P)}_{I}DP^{-1} = PDDP^{-1}$$

$$= PD^2P^{-1} = \begin{bmatrix} 1 & 1 \\ -1 & -2 \end{bmatrix}\begin{bmatrix} 5^2 & 0 \\ 0 & 3^2 \end{bmatrix}\begin{bmatrix} 2 & 1 \\ -1 & -1 \end{bmatrix}$$

- Again,

$$A^3 = (PDP^{-1})A^2 = (PD\underbrace{P^{-1})P}_{I}D^2P^{-1} = PDD^2P^{-1} = PD^3P^{-1}$$

© 2016 Pearson Education, Inc.

DIAGONALIZATION

- In general, for $k \geq 1$,

$$A^k = PD^kP^{-1} = \begin{bmatrix} 1 & 1 \\ -1 & -2 \end{bmatrix} \begin{bmatrix} 5^k & 0 \\ 0 & 3^k \end{bmatrix} \begin{bmatrix} 2 & 1 \\ -1 & -1 \end{bmatrix}$$

$$= \begin{bmatrix} 2 \cdot 5^k - 3^k & 5^k - 3^k \\ 2 \cdot 3^k - 2 \cdot 5^k & 2 \cdot 3^k - 5^k \end{bmatrix}$$

- A square matrix A is said to be **diagonalizable** if A is similar to a diagonal matrix, that is, if $A = PDP^{-1}$ for some invertible matrix P and some diagonal, matrix D.

© 2016 Pearson Education, Inc.

THE DIAGONALIZATION THEOREM

- **Theorem 5:** 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 and 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.

  In other words, A is diagonalizable if and only if there are enough eigenvectors to form a basis of $\mathbb{R}^n$. We call such a basis an **eigenvector basis** of. $\mathbb{R}^n$

© 2016 Pearson Education, Inc.

THE DIAGONALIZATION THEOREM

- **Proof:** First, observe that if P is any $n \times n$ matrix with columns $\mathbf{v}_1, \ldots, \mathbf{v}_n$, and if D is any diagonal matrix with diagonal entries $\lambda_1, \ldots, \lambda_n$, then

$$AP = A\begin{bmatrix} \mathbf{v}_1 & \mathbf{v}_2 & \cdots & \mathbf{v}_n \end{bmatrix} = \begin{bmatrix} A\mathbf{v}_1 & A\mathbf{v}_2 & \cdots & A\mathbf{v}_n \end{bmatrix} \quad (1)$$

while

$$PD = P\begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix} = \begin{bmatrix} \lambda_1\mathbf{v}_1 & \lambda_2\mathbf{v}_2 & \cdots & \lambda_n\mathbf{v}_n \end{bmatrix} \quad (2)$$

© 2016 Pearson Education, Inc.

THE DIAGONALIZATION THEOREM

- Now suppose A is diagonalizable and $A = PDP^{-1}$. Then right-multiplying this relation by P, we have $AP = PD$.
- In this case, equations (1) and (2) imply that

$$\begin{bmatrix} A\mathbf{v}_1 & A\mathbf{v}_2 & \cdots & A\mathbf{v}_n \end{bmatrix} = \begin{bmatrix} \lambda_1\mathbf{v}_1 & \lambda_2\mathbf{v}_2 & \cdots & \lambda_n\mathbf{v}_n \end{bmatrix} \tag{3}$$

- Equating columns, we find that

$$A\mathbf{v}_1 = \lambda_1\mathbf{v}_1, A\mathbf{v}_2 = \lambda_2\mathbf{v}_2, \ldots, A\mathbf{v}_n = \lambda_n\mathbf{v}_n \tag{4}$$

- Since P is invertible, its columns $\mathbf{v}_1, \ldots, \mathbf{v}_n$ must be linearly independent.

© 2016 Pearson Education, Inc.

THE DIAGONALIZATION THEOREM

- Also, since these columns are nonzero, the equations in (4) show that $\lambda_1, \ldots, \lambda_n$ are eigenvalues and $\mathbf{v}_1, \ldots, \mathbf{v}_n$ are corresponding eigenvectors.

- This argument proves the "only if " parts of the first and second statements, along with the third statement, of the theorem.

- Finally, given any n eigenvectors $\mathbf{v}_1, \ldots, \mathbf{v}_n$, use them to construct the columns of P and use corresponding eigenvalues $\lambda_1, \ldots, \lambda_n$ to construct D.

© 2016 Pearson Education, Inc.

THE DIAGONALIZATION THEOREM

- By equations (1)–(3), $AP = PD$.
- This is true without any condition on the eigenvectors.
- If, in fact, the eigenvectors are linearly independent, then P is invertible (by the Invertible Matrix Theorem), and $AP = PD$ implies that $A = PDP^{-1}$.

© 2016 Pearson Education, Inc.

DIAGONALIZING MATRICES

- **Example 3:** Diagonalize the following matrix, if possible.

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

  That is, find an invertible matrix P and a diagonal matrix D such that $A = PDP^{-1}$.

- **Solution:** There are four steps to implement the description in Theorem 5.
- *Step 1.* ***Find the eigenvalues of A.***
- Here, the characteristic equation turns out to involve a cubic polynomial that can be factored:

© 2016 Pearson Education, Inc.

DIAGONALIZING MATRICES

$$0 = \det(A - \lambda I) = -\lambda^3 - 3\lambda^2 + 4$$
$$= -(\lambda - 1)(\lambda + 2)^2$$

- The eigenvalues are $\lambda = 1$ and $\lambda = -2$.
- *Step 2.* ***Find three linearly independent eigenvectors of A.***
- *Three* vectors are needed because A is a 3×3 matrix.
- This is a critical step.
- If it fails, then Theorem 5 says that A cannot be diagonalized.

© 2016 Pearson Education, Inc.

DIAGONALIZING MATRICES

- Basis for $\lambda = 1 : \mathrm{v}_1 = \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}$

- Basis for $\lambda = -2 : \mathrm{v}_2 = \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}$ and $\mathrm{v}_3 = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}$

- You can check that $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ is a linearly independent set.

© 2016 Pearson Education, Inc.

DIAGONALIZING MATRICES

- *Step 3.* ***Construct P from the vectors in step 2.***
- The order of the vectors is unimportant.
- Using the order chosen in step 2, form

$$P = \begin{bmatrix} \mathrm{v}_1 & \mathrm{v}_2 & \mathrm{v}_3 \end{bmatrix} = \begin{bmatrix} 1 & -1 & -1 \\ -1 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix}$$

- *Step 4.* ***Construct D from the corresponding eigenvalues.***
- In this step, it is essential that the order of the eigenvalues matches the order chosen for the columns of *P*.

© 2016 Pearson Education, Inc.

DIAGONALIZING MATRICES

- Use the eigenvalue $\lambda = -2$ twice, once for each of the eigenvectors corresponding to $\lambda = -2$:

$$D = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & -2 \end{bmatrix}$$

- To avoid computing P^{-1}, simply verify that $AD = PD$.
- Compute

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

© 2016 Pearson Education, Inc.

DIAGONALIZING MATRICES

$$PD = \begin{bmatrix} 1 & -1 & -1 \\ -1 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & -2 \end{bmatrix} = \begin{bmatrix} 1 & 2 & 2 \\ -1 & -2 & 0 \\ 1 & 0 & -2 \end{bmatrix}$$

- **Theorem 6:** An $n \times n$ matrix with n distinct eigenvalues is diagonalizable.
- **Proof:** Let $\mathbf{v}_1, \ldots, \mathbf{v}_n$ be eigenvectors corresponding to the n distinct eigenvalues of a matrix A.
- Then $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ is linearly independent, by Theorem 2 in Section 5.1.
- Hence A is diagonalizable, by Theorem 5.

© 2016 Pearson Education, Inc.

MATRICES WHOSE EIGENVALUES ARE NOT DISTINCT

- It is not *necessary* for an $n \times n$ matrix to have n distinct eigenvalues in order to be diagonalizable.
- The 3 x 3 matrix in Example 3 is diagonalizable even though it has only two distinct eigenvalues.

- If an $n \times n$ matrix A has n distinct eigenvalues, with corresponding eigenvectors $\mathbf{v}_1, \ldots, \mathbf{v}_n$, and if $P = \begin{bmatrix} \mathrm{v}_1 & \cdots & \mathrm{v}_2 \end{bmatrix}$, then P is automatically invertible because its columns are linearly independent, by Theorem 2.

© 2016 Pearson Education, Inc.

MATRICES WHOSE EIGENVALUES ARE NOT DISTINCT

- When A is diagonalizable but has fewer than n distinct eigenvalues, it is still possible to build P in a way that makes P automatically invertible, as the next theorem shows.

- **Theorem 7:** Let A be an $n \times n$ matrix whose distinct eigenvalues are $\lambda_1, \ldots, \lambda_p$.
  a. For $1 \le k \le p$, the dimension of the eigenspace for λ_k is less than or equal to the multiplicity of the eigenvalue λ_k.

© 2016 Pearson Education, Inc.

MATRICES WHOSE EIGENVALUES ARE NOT DISTINCT

b. The matrix A is diagonalizable if and only if the sum of the dimensions of the eigenspaces equals n, and this happens if and only if (*i*) the characteristic polynomial factors completely into linear factors and (*ii*) the dimension of the eigenspace for each λ_k equals the multiplicity of λ_k.

c. If A is diagonalizable and $\mathcal{B}_k$ is a basis for the eigenspace corresponding to $\mathcal{B}_k$ for each k, then the total collection of vectors in the sets $\mathcal{B}_1, \ldots, \mathcal{B}_p$ forms an eigenvector basis for $\mathbb{R}^n$.

© 2016 Pearson Education, Inc.

5 Eigenvalues and Eigenvectors

5.4

Eigenvectors and Linear Transformations

PEARSON

© 2012 Pearson Education, Inc.

THE MATRIX OF A LINEAR TRANSFORMATION

- Given any $\mathbf{x}$ in V, the coordinate vector $[\mathbf{x}]_B$ is in $\mathbb{R}^n$ and the coordinate vector of its image, $[T(\mathbf{x})]_C$, as shown in Fig 1.

FIGURE 1 A linear transformation from V to W.

© 2012 Pearson Education, Inc.

THE MATRIX OF A LINEAR TRANSFORMATION

- The connection between $[\mathbf{x}]_B$ and $[T(\mathbf{x})]_C$ is easy to find. Let $\{\mathbf{b}_1,\ldots\mathbf{b}_n\}$ be the basis B for V. If $\mathbf{x} = r_1\mathbf{b}_1 + \ldots + r_n\mathbf{b}_n$, then

$$[\mathbf{x}]_B = \begin{bmatrix} r_1 \\ \vdots \\ r_n \end{bmatrix}$$

- and
- $T(\mathbf{x}) = T(r_1\mathbf{b}_1 + \ldots + r_n\mathbf{b}_n) = r_1 T(\mathbf{b}_1) + \ldots + r_n T(\mathbf{b}_n)$
- (1)
- Because T is linear. Now, since the coordinate mapping from W to $\mathbb{R}^m$ is linear.

© 2012 Pearson Education, Inc.

THE MATRIX OF A LINEAR TRANSFORMATION

- $$[T(\mathbf{x})]_C = r_1[T(\mathbf{b}_1)]_C + \ldots + r_n[T(\mathbf{b}_n)]_C \qquad (2)$$

Since C-coordinate vectors are in $\mathbb{R}^m$, the vector equation (2) can be written as a matrix equation, namely,

$$[T(\mathbf{x})]_C = M[\mathbf{x}]_B \qquad (3)$$

where

$$M = \begin{bmatrix} [T(\mathbf{b}_1)]_C & [T(\mathbf{b}_2)]_C & \cdots & [T(\mathbf{b}_n)]_C \end{bmatrix} \qquad (4)$$

The matrix M is a matrix representation of T, called the **matrix for T relative to the bases B and C**. See Fig 2.

© 2012 Pearson Education, Inc.

THE MATRIX OF A LINEAR TRANSFORMATION

$$\begin{tikzcd} \mathbf{x} \arrow[r, "T"] \arrow[d] & T(\mathbf{x}) \arrow[d] \\ {[\mathbf{x}]_{\mathcal{B}}} \arrow[r, "\text{Multiplication by } M"'] & {[T(\mathbf{x})]_{\mathcal{C}}} \end{tikzcd}$$

FIGURE 2

- **EXAMPLE 1**
- Suppose $B = \{\mathbf{b}_1, \mathbf{b}_2\}$ is a basis for V and $C = \{c_1, c_2, c_3\}$ is a basis for W. Let $T: V \to W$ be a linear transformation with the property that
- $T(\mathbf{b}_1) = 3\mathbf{c}_1 - 2\mathbf{c}_2 + 5\mathbf{c}_3$ and $T(\mathbf{b}_2) = 4\mathbf{c}_1 + 7\mathbf{c}_2 - \mathbf{c}_3$
- Find the matrix M for T relative to B and C.

© 2012 Pearson Education, Inc.

THE MATRIX OF A LINEAR TRANSFORMATION

- **SOLUTION** The C-coordinate vectors of the images of $\mathbf{b}_1$ and $\mathbf{b}_2$ are

$$[T(\mathbf{b}_1)]_C = \begin{bmatrix} 3 \\ -2 \\ 5 \end{bmatrix} \quad \textbf{and} \quad [T(\mathbf{b}_2)]_C = \begin{bmatrix} 4 \\ 7 \\ -1 \end{bmatrix}$$

- Hence

$$M = \begin{bmatrix} 3 & 4 \\ -2 & 7 \\ 5 & -1 \end{bmatrix}$$

- If B and C are bases for the same space V and if T is the identity transformation $T(\mathbf{x}) = \mathbf{x}$ for $\mathbf{x}$ in V, then matrix M in (4) is just a change-of-coordinates matrix.

© 2012 Pearson Education, Inc.

Linear Transformations from V into V

- In the common case where W is the same as V and the basis C is the same as B, the matrix M in (4) is called the **matrix for T relative to B**, or simple the **B-matrix for T**, and is denoted by $[T]_B$. See Fig 3.

$$\begin{tikzcd}
\mathbf{x} \arrow[r, "T"] \arrow[d] & T(\mathbf{x}) \arrow[d] \\
{[\mathbf{x}]_{\mathcal{B}}} \arrow[r, "\text{Multiplication by } [T]_{\mathcal{B}}"'] & {[T(\mathbf{x})]_{\mathcal{B}}}
\end{tikzcd}$$

FIGURE 3

The B-matrix for $T: V \to V$ satisfies

- $$[T(\mathbf{x})]_B = [T]_B[\mathbf{x}]_B, \qquad \text{for all } \mathbf{x} \text{ in } V$$

(5)

© 2012 Pearson Education, Inc.

Linear Transformations from V into V

EXAMPLE 2

The mapping $T: \mathbb{P}_2 \to \mathbb{P}_2$ defined by

- $$T(a_0 + a_1 t + a_2 t_2) = a_1 + 2a_2 t$$

Is a linear transformation.

A. Find the B-matrix for T, when B is the basis $\{1, t, t_2\}$.

B. Verify that $[T(\mathbf{p})]_B = [T]_B[\mathbf{p}]_B$ for each $\mathbf{p}$ in $\mathbb{P}_2$.

Solution

a. Compute the images of the basis vectors:

$T(1) = 0$ The zero polynomial

$T(t) = 1$ The polynomial whose value is always 1

$T(t^2) = 2t$

© 2012 Pearson Education, Inc.

Linear Transformations from V into V

The write the B-coordinate vectors of $T(1)$, $T(t)$, and $T(t^2)$ and place them together as the B-matrix for T:

$$[T(1)]_B = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \quad [T(t)]_B = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \quad [T(t^2)]_B = \begin{bmatrix} 0 \\ 2 \\ 0 \end{bmatrix}$$

$$[T]_B = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{bmatrix}$$

© 2012 Pearson Education, Inc.

Linear Transformations from V into V

b. For a general $\mathbf{p}(t) = a_0 + a_1 t + a_2 t^2$,

$$[T(\mathbf{p})]_B = [a_1 + 2a_2 t]_B = \begin{bmatrix} a_1 \\ 2a_2 \\ 0 \end{bmatrix}$$

$$= \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ a_2 \end{bmatrix} = [T]_B [\mathbf{p}]_B$$

See Fig 4.

© 2012 Pearson Education, Inc.

Linear Transformations from V into V

FIGURE 4 Matrix representation of a linear transformation.

© 2012 Pearson Education, Inc.

Linear Transformations on $\mathbb{R}^n$

THEOREM 8 Diagonal Matrix Representation

Suppose $A = PDP^{-1}$, where D is diagonal $n \times n$ matrix. If B is the basis for $\mathbb{R}^n$ formed from the columns of P, then D is the B-matrix for the transformation $\mathbf{x} \mapsto A\mathbf{x}$.

PROOF

Denote the columns of P by $\mathbf{b}_1, \ldots \mathbf{b}_n$, so that $B = \{\mathbf{b}_1, \ldots \mathbf{b}_n\}$ and $P = [\mathbf{b}_1 \ldots \mathbf{b}_n]$. In this case, P is the change of coordinates matrix P_B.

$$P[\mathbf{x}]_B = \mathbf{x} \text{ and } [\mathbf{x}]_B = P^{-1}\mathbf{x}$$

If $T(\mathbf{x}) = A\mathbf{x}$ for $\mathbf{x}$ in $\mathbb{R}^n$, then

© 2012 Pearson Education, Inc.

Linear Transformations on $\mathbb{R}^n$

$$[T]_B = \begin{bmatrix} [T(b_1)]_B & \cdots & [T(\mathbf{b}_n)]_B \end{bmatrix}$$ Definition of $[T]_B$

$$= \begin{bmatrix} [A\mathbf{b}_1]_B & \cdots & [A\mathbf{b}_n]_B \end{bmatrix}$$ Since $T(\mathbf{x}) = A\mathbf{x}$

$$= \begin{bmatrix} P^{-1}A\mathbf{b}_1 & \cdots P^{-1}A\mathbf{b}_n \end{bmatrix}$$ Change of coordinates

$$= P^{-1}A\begin{bmatrix} \mathbf{b}_1 & \cdots \mathbf{b}_n \end{bmatrix}$$ Matrix multiplication

$$= P^{-1}AP \qquad (6)$$

Since $A = PDP^{-1}$, we have $[T]_B = P^{-1}AP = D$.

© 2012 Pearson Education, Inc.

Similarity of Matrix Representations

The proof of Theorem 8 did not use the information that D was diagonal. Hence, if A is similar to a matrix C, with $A = PCP^{-1}$, then C is the $\mathcal{B}$-matrix for the transformation $\mathbf{x} \mapsto A\mathbf{x}$ when the basis $\mathcal{B}$ is formed from the columns of P. The factorization $A = PCP^{-1}$ is shown in Fig 5.

$$\begin{tikzcd}
\mathbf{x} \arrow[r, "\text{Multiplication by } A"'] \arrow[d, "\text{Multiplication by } P^{-1}"'] & A\mathbf{x} \\
{[\mathbf{x}]_{\mathcal{B}}} \arrow[r, "\text{Multiplication by } C"'] & {[A\mathbf{x}]_{\mathcal{B}}} \arrow[u, "\text{Multiplication by } P"']
\end{tikzcd}$$

FIGURE 5 Similarity of two matrix representations: $A = PCP^{-1}$.

© 2012 Pearson Education, Inc.

Similarity of Matrix Representations

If $T: \mathbb{R}^n \to \mathbb{R}^n$ is defined by $T(\mathbf{x}) = A\mathbf{x}$, and if B is any basis for $\mathbb{R}^n$, then the B-matrix for T is similar to A. In fact, the calculations in the proof of Theorem 8 show that if P is the matrix whose columns come from the vectors in B, then $[T]_B = P^{-1}AP$.

EXAMPLE 4

Let $A = \begin{bmatrix} 4 & -9 \\ 4 & -8 \end{bmatrix}$, $\mathrm{b}_1 = \begin{bmatrix} 3 \\ 2 \end{bmatrix}$, and $\mathrm{b}_2 = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$ The characteristic polynomial of A is $(\lambda+2)^2$, but the eigenspace for the eigenvalue -2 is only one-dimensional; so A is not diagonalizable./

© 2012 Pearson Education, Inc.

Similarity of Matrix Representations

However, the basis $B = \{\mathbf{b}_1, \mathbf{b}_2\}$ has the property that the B-matrix for the transformation $\mathbf{x} \mapsto A\mathbf{x}$ is a triangular matrix called the *Jordan form* of $\boldsymbol{A}$**.** Find this B-matrix.

SOLUTION

If $P = [\mathbf{b}_1 \quad \mathbf{b}_2]$, then the B-matrix is $P^{-1}AP$, Compute

$$AP = \begin{bmatrix} 4 & -9 \\ 4 & -8 \end{bmatrix} \begin{bmatrix} 3 & 2 \\ 2 & 1 \end{bmatrix} = \begin{bmatrix} -6 & -1 \\ -4 & 0 \end{bmatrix}$$

$$P^{-1}AP = \begin{bmatrix} -1 & 2 \\ 2 & -3 \end{bmatrix} \begin{bmatrix} -6 & -1 \\ -4 & 0 \end{bmatrix} = \begin{bmatrix} -2 & 1 \\ 0 & -2 \end{bmatrix}$$

© 2012 Pearson Education, Inc.

Similarity of Matrix Representations

NUMERICAL NOTE

An efficient way to compute a B-matrix $P^{-1}AP$ is to compute AP and then to row reduce the augmented matrix $[P \quad AP]$ to $[I \quad P^{-1}AP]$. A separate computation of P^{-1} is unnecessary.

© 2012 Pearson Education, Inc.

5 Eigenvalues and Eigenvectors

5.5

COMPLEX EIGENVALUES

PEARSON

© 2012 Pearson Education, Inc.

COMPLEX EIGENVALUES

- The matrix eigenvalue–eigenvector theory already developed for $\mathbb{R}^n$ applies equally well to $\mathbb{C}^n$
- So a complex scalar λ satisfies $\det(A - \lambda I) = 0$ if and only if there is a nonzero vector $\mathbf{x}$ in $\mathbb{C}^n$ such that $A\mathbf{x} = \lambda\mathbf{x}$.
- We call λ a (**complex**) **eigenvalue** and $\mathbf{x}$ a (**complex**) **eigenvector** corresponding to λ.
- **EXAMPLE 1** If $A = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$, then the linear transformation $\mathbf{x} \to A\mathbf{x}$ on $\mathbb{R}^2$ rotates the plane counterclockwise through a quarter-turn.

© 2012 Pearson Education, Inc.

COMPLEX EIGENVALUES

- The action of A is periodic, since after four quarter-turns, a vector is back where it started.
- Obviously, no nonzero vector is mapped into a multiple of itself, so A has no eigenvectors in $\mathbb{R}^2$ and hence no real eigenvalues.
- In fact, the characteristic equation of A is

$$\lambda^2 + 1 = 0$$

The only roots are complex: $\lambda = i$ and $\lambda = -i$.

© 2012 Pearson Education, Inc.

COMPLEX EIGENVALUES

$$\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} 1 \\ -i \end{bmatrix} = \begin{bmatrix} i \\ 1 \end{bmatrix} = i\begin{bmatrix} 1 \\ -i \end{bmatrix}$$

$$\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} 1 \\ i \end{bmatrix} = \begin{bmatrix} -i \\ 1 \end{bmatrix} = -i\begin{bmatrix} 1 \\ i \end{bmatrix}$$

However, if we permit A to act on C2, then

Thus i and −i are eigenvalues, with $\begin{bmatrix} 1 \\ -i \end{bmatrix}$ and $\begin{bmatrix} 1 \\ i \end{bmatrix}$ as corresponding eigenvectors.
(A method for finding complex eigenvectors is discussed in Example 2.)

© 2012 Pearson Education, Inc.

COMPLEX EIGENVALUES

- **EXAMPLE 2** Let $A = \begin{bmatrix} .5 & -.6 \\ .75 & 1.1 \end{bmatrix}$. Find the eigenvalues of A, and find a basis for each eigenspace.
- **Solution** The characteristic equation of A is

$$0 = \det \begin{bmatrix} .5-\lambda & -.6 \\ .75 & 1.1-\lambda \end{bmatrix} = (.5-\lambda)(1.1-\lambda) - (-.6)(.75)$$

$$= \lambda^2 - 1.6\lambda + 1$$

- From the quadratic formula, $\lambda = \frac{1}{2}\left[1.6 \pm \sqrt{(-1.6)^2 - 4}\right] = .8 \pm .6i$

© 2012 Pearson Education, Inc.

COMPLEX EIGENVALUES

- A basis for the eigenspace corresponding to $\lambda = .8 - .6i$ is

$$\mathbf{v}_1 = \begin{bmatrix} -2-4i \\ 5 \end{bmatrix}$$

- Analogous calculations for $\lambda = .8 + .6i$ produce the eigenvector

$$\mathbf{v}_2 = \begin{bmatrix} -2+4i \\ 5 \end{bmatrix}$$

- As a check on the work, compute

$$A\mathbf{v}_2 = \begin{bmatrix} .5 & -.6 \\ .75 & 1.1 \end{bmatrix}\begin{bmatrix} -2+4i \\ 5 \end{bmatrix} = \begin{bmatrix} -4+2i \\ 4+3i \end{bmatrix} = (.8 + .6\text{i})\mathbf{v}_2$$

© 2012 Pearson Education, Inc.

COMPLEX EIGENVALUES

- **EXAMPLE 3** One way to see how multiplication by the A in Example 2 affects points is to plot an arbitrary initial point—say, $\mathbf{x}_0 = (2, 0)$—and then to plot successive images of this point under repeated multiplications by A.
- That is, plot

$$x_1 = Ax_0 = \begin{bmatrix} .5 & -.6 \\ .75 & 1.1 \end{bmatrix}\begin{bmatrix} 2 \\ 0 \end{bmatrix} = \begin{bmatrix} 1.0 \\ 1.5 \end{bmatrix}$$

$$x_2 = Ax_1 = \begin{bmatrix} .5 & -.6 \\ .75 & 1.1 \end{bmatrix}\begin{bmatrix} 1.0 \\ 1.5 \end{bmatrix} = \begin{bmatrix} -.4 \\ 2.4 \end{bmatrix}$$

$$x_3 = Ax_2, \ldots$$

- Figure 1 shows $\mathbf{x}_0, \ldots, \mathbf{x}_8$ as larger dots. The smaller dots are the locations of $\mathbf{x}_9, \ldots, \mathbf{x}_{100}$. The sequence lies along an elliptical orbit.

© 2012 Pearson Education, Inc.

Real and Imaginary Parts of Vectors

- The complex conjugate of a complex vector **x** in $\mathbb{C}^n$ is the vector $\bar{x}$ in $\mathbb{C}^n$ whose entries are the complex conjugates of the entries in **x**.
- The **real** and **imaginary parts** of a complex vector **x** are the vectors Re **x** and Im **x** formed from the real and imaginary parts of the entries of **x**.

FIGURE 1 Iterates of a point $\mathbf{x}_0$ under the action of a matrix with a complex eigenvalue.

© 2012 Pearson Education, Inc.

Real and Imaginary Parts of Vectors

- **EXAMPLE 4** If $\mathrm{x} = \begin{bmatrix} 3-i \\ i \\ 2+5i \end{bmatrix} = \begin{bmatrix} 3 \\ 0 \\ 2 \end{bmatrix} + i \begin{bmatrix} -1 \\ 1 \\ 5 \end{bmatrix}$,

  then Re $\mathbf{x} = \begin{bmatrix} 3 \\ 0 \\ 2 \end{bmatrix}$, Im $\mathbf{x} = \begin{bmatrix} -1 \\ 1 \\ 5 \end{bmatrix}$

  and $\bar{x} = \begin{bmatrix} 3 \\ 0 \\ 2 \end{bmatrix} - i \begin{bmatrix} -1 \\ 1 \\ 5 \end{bmatrix} = \begin{bmatrix} 3+i \\ -i \\ 2-5i \end{bmatrix}$

© 2012 Pearson Education, Inc.

Eigenvalues and Eigenvectors of a Real Matrix That Acts on $\mathbb{C}^n$

- Let A be an $m \times n$ matrix whose entries are real.
- Then $\overline{Ax} = \overline{A}\overline{x} = A\overline{x}$.
- If λ is an eigenvalue of A and **x** is a corresponding eigenvector in $\mathbb{C}^n$, then

$$A\overline{x} = \overline{Ax} = \overline{\lambda x} = \overline{\lambda}\overline{x}$$

- Hence $\overline{\lambda}$ is also an eigenvalue of A, with $\overline{x}$ a corresponding eigenvector.

© 2012 Pearson Education, Inc.

Eigenvalues and Eigenvectors of a Real Matrix That Acts on $\mathbb{C}^n$

- **EXAMPLE 5** The eigenvalues of the real matrix in Example 2 are complex conjugates, namely, $.8 - .6i$ and $.8 + .6i$.
- The corresponding eigenvectors found in Example 2 are also conjugates:

$$\mathbf{v}_1 = \begin{bmatrix} -2-4i \\ 5 \end{bmatrix} \quad \text{and} \quad \mathbf{v}_2 = \begin{bmatrix} -2+4i \\ 5 \end{bmatrix} = \overline{v}_1$$

© 2012 Pearson Education, Inc.

Eigenvalues and Eigenvectors of a Real Matrix That Acts on $\mathbb{C}^n$

- **EXAMPLE 6** If $C = \begin{bmatrix} a & -b \\ b & a \end{bmatrix}$, where a and b are real and not both zero, then the eigenvalues of C are $\lambda = a \pm bi$.
- Also, if $r = |\lambda| = \sqrt{a^2 + b^2}$, then

$$C = r\begin{bmatrix} a/r & -b/r \\ b/r & a/r \end{bmatrix} = \begin{bmatrix} r & 0 \\ 0 & r \end{bmatrix}\begin{bmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{bmatrix}$$

- where ϕ is the angle between the positive x-axis and the ray from (0, 0) through (a, b).See Fig. 2

© 2012 Pearson Education, Inc.

Eigenvalues and Eigenvectors of a Real Matrix That Acts on $\mathbb{C}^n$

- Thus the transformation $\mathbf{x} \to C\mathbf{x}$ may be viewed as the composition of a rotation through the angle ϕ and a scaling by $|\lambda|$ (see Fig. 3).

FIGURE 2

FIGURE 3 A rotation followed by a scaling.

© 2012 Pearson Education, Inc.

Eigenvalues and Eigenvectors of a Real Matrix That Acts on $\mathbb{C}^n$

- **EXAMPLE 7** Let $A = \begin{bmatrix} .5 & -.6 \\ .75 & 1.1 \end{bmatrix}$, $\lambda = .8 - .6i$, and $\mathbf{v}_1 = \begin{bmatrix} -2-4i \\ 5 \end{bmatrix}$, as in Example 2.
- Also, let P be the 2×2 real matrix

$$P = [\text{Re } \mathbf{v}_1 \quad \text{Im } \mathbf{v}_1] = \begin{bmatrix} -2 & -4 \\ 5 & 0 \end{bmatrix}$$

- and let

$$C = P^{-1}AP = \frac{1}{20}\begin{bmatrix} 0 & 4 \\ -5 & -2 \end{bmatrix}\begin{bmatrix} .5 & -.6 \\ .75 & 1.1 \end{bmatrix}\begin{bmatrix} -2 & -4 \\ 5 & 0 \end{bmatrix} = \begin{bmatrix} .8 & -.6 \\ .6 & .8 \end{bmatrix}$$

© 2012 Pearson Education, Inc.

Eigenvalues and Eigenvectors of a Real Matrix That Acts on $\mathbb{C}^n$

- By Example 6, C is a pure rotation because $|\lambda|^2 = (.8)^2 + (.6)^2 = 1$.
- From $C = P^{-1}AP$, we obtain

$$A = PCP^{-1} = P\begin{bmatrix} .8 & -.6 \\ .6 & .8 \end{bmatrix} P^{-1}$$

- Here is the rotation "inside" A!
- The matrix P provides a change of variable, say, $\mathbf{x} = P\mathbf{u}$.

© 2012 Pearson Education, Inc.

Eigenvalues and Eigenvectors of a Real Matrix That Acts on $\mathbb{C}^n$

- The action of A amounts to a change of variable from $\mathbf{x}$ to $\mathbf{u}$, followed by a rotation, and then a return to the original variable. See Fig. 4.
- The rotation produces an ellipse, as in Fig. 1, instead of a circle, because the coordinate system determined by the columns of P is not rectangular and does not have equal unit lengths on the two axes.

FIGURE 4 Rotation due to a complex eigenvalue.

© 2012 Pearson Education, Inc.

Eigenvalues and Eigenvectors of a Real Matrix That Acts on $\mathbb{C}^n$

THEOREM 9

- Let A be a real 2×2 matrix with a complex eigenvalue $\lambda = a - bi$ ($b \neq 0$) and an associated eigenvector $\mathbf{v}$ in $\mathbb{C}^2$.
- Then

  $A = PCP^{-1}$, where $P = [\text{Re } \mathbf{v} \quad \text{Im } \mathbf{v}]$ and $C = \begin{bmatrix} a & -b \\ b & a \end{bmatrix}$

© 2012 Pearson Education, Inc.

5 Eigenvalues and Eigenvectors

5.6

DISCRETE DYNAMICAL SYSTEMS

PEARSON
© 2012 Pearson Education, Inc.

DISCRETE DYNAMICAL SYSTEMS

- Eigenvalues and eigenvectors provide the key to understanding the long-term behavior, or *evolution*, of a dynamical system described by a difference equation

$$\mathbf{x}_{k+1} = A\mathbf{x}_k$$

© 2012 Pearson Education, Inc.

DISCRETE DYNAMICAL SYSTEMS

- We assume that A is diagonalizable, with n linearly independent eigenvectors, $\mathbf{v}_1, \ldots, \mathbf{v}_n$, and corresponding eigenvalues, $\lambda_1, \ldots, \lambda_n$.
- For convenience, assume the eigenvectors are arranged so that $|\lambda_1| \geq |\lambda_2| \geq \cdots \geq |\lambda_n|$.
- Since $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ is a basis for $\mathbb{R}^n$, any initial vector $\mathbf{x}_0$ can be written uniquely as

$$\mathbf{x}_0 = c_1\mathbf{v}_1 + \cdots + c_n\mathbf{v}_n$$

- This *eigenvector decomposition* of $\mathbf{x}_0$ determines what happens to the sequence $\{\mathbf{x}_k\}$.

© 2012 Pearson Education, Inc.

DISCRETE DYNAMICAL SYSTEMS

- Since the $\mathbf{v}i$ are eigenvectors,

$$\mathbf{x}_1 = A\mathbf{x}_0 = c_1 A\mathbf{v}_1 + \cdots + c_n A\mathbf{v}_n$$
$$= c_1\lambda_1\mathbf{v}_1 + \cdots + c_n\lambda_n\mathbf{v}_n$$

- In general,

$$\mathbf{x}_k = c_1(\lambda_1)^k\mathbf{v}_1 + \cdots + c_n(\lambda_n)^k\mathbf{v}_n \qquad (k = 0, 1, 2, \ldots)$$

© 2012 Pearson Education, Inc.

Graphical Description of Solutions

- When A is 2×2, algebraic calculations can be supplemented by a geometric description of a system's evolution.
- We can view the equation $\mathbf{x}_{k+1} = A\mathbf{x}_k$ as a description of what happens to an initial point $\mathbf{x}_0$ in $\mathbb{R}^2$ as it is transformed repeatedly by the mapping $\mathbf{x} \rightarrow A\mathbf{x}$.
- The graph of $\mathbf{x}_0, \mathbf{x}_1, \ldots$ is called a **trajectory** of the dynamical system.

© 2012 Pearson Education, Inc.

Graphical Description of Solutions

- **EXAMPLE 2** Plot several trajectories of the dynamical system $\mathbf{x}_{k+1} = A\mathbf{x}_k$, when

$$A = \begin{bmatrix} .80 & 0 \\ 0 & .64 \end{bmatrix}$$

- **Solution** The eigenvalues of A are .8 and .64, with eigenvectors $\mathbf{v}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ and $\mathbf{v}_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$.

© 2012 Pearson Education, Inc.

Graphical Description of Solutions

- If $\mathbf{x}_0 = c_1\mathbf{v}_1 + c_2\mathbf{v}_2$, then

$$x_k = c_1(.8)^k \begin{bmatrix} 1 \\ 0 \end{bmatrix} + c_2(.64)^k \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

- Of course, $\mathbf{x}_k$ tends to $\mathbf{0}$ because $(.8)^k$ and $(.64)^k$ both approach 0 as $k \to \infty$. But *the way* $\mathbf{x}_k$ goes toward $\mathbf{0}$ is interesting.
- Figure 1 shows the first few terms of several trajectories that begin at points on the boundary of the box with corners at $(\pm 3, \pm 3)$.

© 2012 Pearson Education, Inc.

Graphical Description of Solutions

- The points on each trajectory are connected by a thin curve, to make the trajectory easier to see.

FIGURE 1 The origin as an attractor.

© 2012 Pearson Education, Inc.

Graphical Description of Solutions

- **EXAMPLE 3** Plot several typical solutions of the equation $\mathbf{x}_{k+1} = A\mathbf{x}_k$, where

$$A = \begin{bmatrix} 1.44 & 0 \\ 0 & 1.2 \end{bmatrix}$$

- **Solution** The eigenvalues of A are 1.44 and 1.2.
- If $\mathbf{x}_0 = \begin{bmatrix} c_1 \\ c_2 \end{bmatrix}$, then

$$x_1 = c_1(1.44)^k \begin{bmatrix} 1 \\ 0 \end{bmatrix} + c_2(1.2)^k \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

© 2012 Pearson Education, Inc.

Graphical Description of Solutions

- Both terms grow in size, but the first term grows faster.
- So the direction of greatest repulsion is the line through **0** and the eigenvector for the eigenvalue of larger magnitude.
- Figure 2 shows several trajectories that begin at points quite close to **0**.

© 2012 Pearson Education, Inc.

Graphical Description of Solutions

FIGURE 2 The origin as a repellor.

© 2012 Pearson Education, Inc.

Graphical Description of Solutions

- **EXAMPLE 4** Plot several typical solutions of the equation $\mathbf{y}_{k+1} = D\mathbf{y}_k$, where

$$D = \begin{bmatrix} 2.0 & 0 \\ 0 & 0.5 \end{bmatrix}$$

Show that a solution $\{\mathbf{y}_k\}$ is unbounded if its initial point is not on the x_2-axis.

- **Solution** The eigenvalues of D are 2 and .5.

If $\mathbf{y}_0 = \begin{bmatrix} c_1 \\ c_2 \end{bmatrix}$, then

$$y_k = c_1 2^k \begin{bmatrix} 1 \\ 0 \end{bmatrix} + c_2 (.5)^k \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

© 2012 Pearson Education, Inc.

Graphical Description of Solutions

- If $\mathbf{y}_0$ is on the x_2-axis, then $c_1 = 0$ and $\mathbf{y}_k \to \mathbf{0}$ as $k \to \infty$.
- But if $\mathbf{y}_0$ is not on the x_2-axis, then the first term in the sum for $\mathbf{y}_k$ becomes arbitrarily large, and so $\{\mathbf{y}_k\}$ is unbounded.
- Figure 3 shows ten trajectories that begin near or on the x_2-axis.

© 2012 Pearson Education, Inc.

Graphical Description of Solutions

FIGURE 3 The origin as a saddle point.

© 2012 Pearson Education, Inc.

Complex Eigenvalues

- When a 2×2 matrix A has complex eigenvalues, A is not diagonalizable (when acting on $\mathbb{R}^n$), but the dynamical system $\mathbf{x}_{k+1} = A\mathbf{x}_k$ is easy to describe.

© 2012 Pearson Education, Inc.

Complex Eigenvalues

- **EXAMPLE 6** It can be verified that the matrix

$$A = \begin{bmatrix} .8 & .5 \\ -.1 & 1.0 \end{bmatrix}$$

has eigenvalues $.9 \pm .2i$, with eigenvectors $\begin{bmatrix} 1 \mp 2i \\ 1 \end{bmatrix}$

- Figure 5 (on page 350) shows three trajectories of the system $\mathbf{x}_{k+1} = A\mathbf{x}_k$, with initial vectors $\begin{bmatrix} 0 \\ 2.5 \end{bmatrix}$, $\begin{bmatrix} 3 \\ 0 \end{bmatrix}$, and $\begin{bmatrix} 0 \\ -2.5 \end{bmatrix}$.

© 2012 Pearson Education, Inc.

Complex Eigenvalues

FIGURE 5 Rotation associated with complex eigenvalues.

© 2012 Pearson Education, Inc.

5 Eigenvalues and Eigenvectors

5.7

APPLICATIONS TO DIFFERENTIAL EQUATIONS

PEARSON
© 2012 Pearson Education, Inc.

APPLICATIONS TO DIFFERENTIAL EQUATIONS

- In many applied problems, several quantities are varying continuously in time, and they are related by a system of differential equations:
- $$\begin{aligned} x'_1 &= a_{11}x_1 + \cdots + a_{1n}x_n \\ x'_2 &= a_{21}x_1 + \cdots + a_{2n}x_n \\ &\vdots \\ x'_n &= a_{n1}x_1 + \cdots + a_{nn}x_n \end{aligned}$$
- Here $x_1, \ldots, x_n$ are differentiable functions of t, with derivatives $x'_1, \ldots, x'_n$, and the a_{ij} are constants.

© 2012 Pearson Education, Inc.

APPLICATIONS TO DIFFERENTIAL EQUATIONS

- We can write the system as a matrix differential equation

$$\mathbf{x}' = A\mathbf{x} \qquad (1)$$

- where

$$\mathbf{x}(t) = \begin{bmatrix} x_1(t) \\ \vdots \\ x_n(t) \end{bmatrix}, \mathbf{x}'(t) = \begin{bmatrix} x_1'(t) \\ \vdots \\ x_n'(t) \end{bmatrix}, \text{ and } A = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{bmatrix}$$

- A **solution** of (1) is a vector-valued function that satisfies (1) for all t in some interval of real numbers, such as $t \geq 0$.

© 2012 Pearson Education, Inc.

APPLICATIONS TO DIFFERENTIAL EQUATIONS

- There always exists what is called a **fundamental set of solutions** to (1).
- If A is $n \times n$, then there are n linearly independent functions in a fundamental set, and each solution of (1) is a unique linear combination of these n functions.
- That is, a fundamental set of solutions is a *basis* for the set of all solutions of (1), and the solution set is an n-dimensional vector space of functions.
- If a vector $\mathbf{x}_0$ is specified, then the **initial value problem** is to construct the (unique) function $\mathbf{x}$ such that $\mathbf{x}' = A\mathbf{x}$ and $x(0) = \mathbf{x}_0$.

© 2012 Pearson Education, Inc.

APPLICATIONS TO DIFFERENTIAL EQUATIONS

- When A is a diagonal matrix, the solutions of (1) can be produced by elementary calculus.
- For instance, consider

$$\begin{bmatrix} x_1'(t) \\ x_2'(t) \end{bmatrix} = \begin{bmatrix} 3 & 0 \\ 0 & -5 \end{bmatrix} \begin{bmatrix} x_1(t) \\ x_1(t) \end{bmatrix} \qquad (2)$$

- that is,

$$\begin{aligned} x_1'(t) &= 3x_1(t) \\ x_2'(t) &= -5x_2(t) \end{aligned} \qquad (3)$$

© 2012 Pearson Education, Inc.

APPLICATIONS TO DIFFERENTIAL EQUATIONS

- the solutions of (3) are $x_1(t) = c_1 e^{3t}$ and $x_2(t) = c_2 e^{-5t}$, for any constants c_1 and c_2.
- Each solution of (2) can be written in the form

$$\begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix} = \begin{bmatrix} c_1 e^{3t} \\ c_2 e^{-5t} \end{bmatrix} = c_1 \begin{bmatrix} 1 \\ 0 \end{bmatrix} e^{3t} + c_2 \begin{bmatrix} 0 \\ 1 \end{bmatrix} e^{-5t}$$

- This example suggests that for the general equation $\mathbf{x}' = A\mathbf{x}$, a solution might be a linear combination of functions of the form

$$\mathbf{x}(\mathrm{t}) = \mathbf{v} e^{\lambda t} \qquad (4)$$

© 2012 Pearson Education, Inc.

APPLICATIONS TO DIFFERENTIAL EQUATIONS

- Observe that

$$\mathbf{x}'(t) = \lambda \mathbf{v} e^{\lambda t}$$ By calculus, since v is a constant vector

$$A\mathbf{x}(t) = A\mathbf{v} e^{\lambda t}$$ Multiplying both sides of (4) by A

- Since $e^{\lambda t}$ is never zero, $\mathbf{x}'(t)$ will equal $A\mathbf{x}(t)$ if and only if $\lambda\mathbf{v} = A\mathbf{v}$, that is, if and onlyif λ is an eigenvalue of A and $\mathbf{v}$ is a corresponding eigenvector.
- Thus each eigenvalue–eigenvector pair provides a solution (4) of $\mathbf{x}' = A\mathbf{x}$.
- Such solutions are sometimes called *eigenfunctions* of the differential equation.

© 2012 Pearson Education, Inc.

APPLICATIONS TO DIFFERENTIAL EQUATIONS

- **EXAMPLE 1** The circuit in Fig. 1 can be described by the differential equation

$$\begin{bmatrix} v'1(t) \\ v'2(t) \end{bmatrix} = \begin{bmatrix} -(1/R_1 + 1/R_2)/C_1 & 1/(R_2C_1) \\ 1/(R_2C_2) & -1/(R_2C_2) \end{bmatrix} \begin{bmatrix} v_1(t) \\ v_2(t) \end{bmatrix}$$

FIGURE 1

- where $v_1(t)$ and $v_2(t)$ are the voltages across the two capacitors at time t.

© 2012 Pearson Education, Inc.

APPLICATIONS TO DIFFERENTIAL EQUATIONS

- Suppose resistor R_1 is 1 ohm, R_2 is 2 ohms,
- capacitor C_1 is 1 farad, and C_2 is .5 farad,
- and suppose there is an initial charge of 5 volts on capacitor C_1 and 4 volts on capacitor C_2.

- Find formulas for $v_1(t)$ and $v_2(t)$ that describe how the voltages change over time.

© 2012 Pearson Education, Inc.

APPLICATIONS TO DIFFERENTIAL EQUATIONS

- **Solution** For the data given, set $A = \begin{bmatrix} -1.5 & .5 \\ 1 & -1 \end{bmatrix}$, $\mathbf{x} = \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}$, and $\mathbf{x}_0 = \begin{bmatrix} 5 \\ 4 \end{bmatrix}$

- The vector $\mathbf{x}_0$ lists the initial values of $\mathbf{x}$.

- From A, we obtain eigenvalues $\lambda_1 = -.5$ and $\lambda_2 = -2$, with corresponding eigenvectors

$$\mathbf{v}_1 = \begin{bmatrix} 1 \\ 2 \end{bmatrix} \quad \text{and} \quad \mathbf{v}_2 = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$$

© 2012 Pearson Education, Inc.

APPLICATIONS TO DIFFERENTIAL EQUATIONS

- The eigenfunctions $\mathbf{x}_1(t) = \mathbf{v}_1 e^{\lambda 1 t}$ and $\mathbf{x}_2(t) = \mathbf{v}_2 e^{\lambda 2 t}$ both satisfy $\mathbf{x}' = A\mathbf{x}$, and so does any linear combination of $\mathbf{x}_1$ and $\mathbf{x}_2$.
- Set

$$\mathbf{x}(t) = c_1\mathbf{v}_1 e^{\lambda 1 t} + c_2\mathbf{v}_2 e^{\lambda 2 t} = c_1 \begin{bmatrix} 1 \\ 2 \end{bmatrix} e^{-.5t} + c_2 \begin{bmatrix} -1 \\ 1 \end{bmatrix} e^{-2t}$$

and note that $\mathbf{x}(0) = c_1\mathbf{v}_1 + c_2\mathbf{v}_2$.

© 2012 Pearson Education, Inc.

APPLICATIONS TO DIFFERENTIAL EQUATIONS

- Since $\mathbf{v}_1$ and $\mathbf{v}_2$ are obviously linearly independent and hence span $\mathbb{R}^2$, c_1 and c_2 can be found to make $\mathbf{x}(0)$ equal to $\mathbf{x}_0$.
- In fact, the equation

$$c_1 \underbrace{\begin{bmatrix} 1 \\ 2 \end{bmatrix}}_{\mathbf{v}_1} + c_2 \underbrace{\begin{bmatrix} -1 \\ 1 \end{bmatrix}}_{\mathbf{v}_2} = \underbrace{\begin{bmatrix} 5 \\ 4 \end{bmatrix}}_{\mathbf{x}_0}$$

leads easily to $c_1 = 3$ and $c_2 = -2$.

© 2012 Pearson Education, Inc.

APPLICATIONS TO DIFFERENTIAL EQUATIONS

- Thus the desired solution of the differential equation $\mathbf{x}' = A\mathbf{x}$ is

$$\mathbf{x}(\mathrm{t}) = 3\begin{bmatrix}1\\2\end{bmatrix}e^{-5t} - 2\begin{bmatrix}-1\\1\end{bmatrix}e^{-2t}$$

or

$$\begin{bmatrix}v_1(t)\\v_2(t)\end{bmatrix} = \begin{bmatrix}3e^{-5t}+2e^{-2t}\\6e^{-5t}-2e^{-2t}\end{bmatrix}$$

- Figure 2 shows the graph, or *trajectory*, of $\mathbf{x}(t)$, for $t \geq 0$, along with trajectories for some other initial points.
- The trajectories of the two eigenfunctions $\mathbf{x}_1$ and $\mathbf{x}_2$ lie in the eigenspaces of A.

© 2012 Pearson Education, Inc.

APPLICATIONS TO DIFFERENTIAL EQUATIONS

- The functions $\mathbf{x}_1$ and $\mathbf{x}_2$ both decay to zero as $t \to \infty$, but the values of $\mathbf{x}_2$ decay faster because its exponent is more negative.
- The entries in the corresponding eigenvector $\mathbf{v}_2$ show that the voltages across the capacitors will decay to zero as rapidly as possible if the initial voltages are equal in magnitude but opposite in sign.

FIGURE 2 The origin as an attractor.

© 2012 Pearson Education, Inc.

5 Eigenvalues and Eigenvectors

5.8

ITERATIVE ESTIMATES FOR EIGENVALUES

PEARSON
© 2012 Pearson Education, Inc.

The Power Method

- In scientific applications of linear algebra, eigenvalues are seldom known precisely. Fortunately, a close numerical approximation is usually quite satisfactory.
- The power method applies to an $n \times n$ matrix A with a **strictly dominant eigenvalue** λ_1, which means that λ_1 must be larger in absolute value than all the other eigenvalues.
- In this case, the power method produces a scalar sequence that approaches λ_1 and a vector sequence that approaches a corresponding eigenvector.

© 2012 Pearson Education, Inc.

The Power Method

- **THE POWER METHOD FOR ESTIMATING A STRICTLY DOMINANT EIGENVALUE**
- **1.** Select an initial vector $\mathbf{x}_0$ whose largest entry is 1.
- **2.** For $k = 0, 1, \ldots,$
  a. Compute $A\mathbf{x}_k$.
  b. Let μ_k be an entry in $A\mathbf{x}_k$ whose absolute value is as large as possible.
  c. Compute $\mathbf{x}_{k+1} = (1/\mu_k)A\mathbf{x}_k$.
- **3.** For almost all choices of $\mathbf{x}_0$, the sequence $\{\mu_k\}$ approaches the dominant eigenvalue, and the sequence $\{\mathbf{x}_k\}$ approaches a corresponding eigenvector.

© 2012 Pearson Education, Inc.

The Power Method

- **EXAMPLE 2** Apply the power method to $A = \begin{bmatrix} 6 & 5 \\ 1 & 2 \end{bmatrix}$ with $\mathbf{x}_0 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$. Stop when $k = 5$, and estimate the dominant eigenvalue and a corresponding eigenvector of A.
- **Solution** Calculations in this example and the next were made with MATLAB, which computes with 16-digit accuracy, although we show only a few significant figures here.

© 2012 Pearson Education, Inc.

The Power Method

- To begin, compute $A\mathbf{x}_0$ and identify the largest entry μ_0 in $A\mathbf{x}_0$:

$$A\mathbf{x}_0 = \begin{bmatrix} 6 & 5 \\ 1 & 2 \end{bmatrix}\begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 5 \\ 2 \end{bmatrix}, \mu_0 = 5$$

- Scale $A\mathbf{x}_0$ by $1/\mu_0$ to get $\mathbf{x}_1$, compute $A\mathbf{x}_1$, and identify the largest entry in $A\mathbf{x}_1$:

$$\mathbf{x}_1 = \frac{1}{\mu_0}A\mathbf{x}_0 = \frac{1}{5}\begin{bmatrix} 5 \\ 2 \end{bmatrix} = \begin{bmatrix} 1 \\ .4 \end{bmatrix}$$

$$A\mathbf{x}_1 = \begin{bmatrix} 6 & 5 \\ 1 & 2 \end{bmatrix}\begin{bmatrix} 1 \\ 4 \end{bmatrix} = \begin{bmatrix} 8 \\ 1.8 \end{bmatrix}, \mu_1 = 8$$

© 2012 Pearson Education, Inc.

The Power Method

- Scale $A\mathbf{x}_1$ by $1/\mu_1$ to get $\mathbf{x}_2$, compute $A\mathbf{x}_2$, and identify the largest entry in $A\mathbf{x}_2$:

$$\mathbf{x}_2 = \frac{1}{\mu_1} A\mathbf{x}_1 = \frac{1}{8}\begin{bmatrix} 8 \\ 1.8 \end{bmatrix} = \begin{bmatrix} 1 \\ .225 \end{bmatrix}$$

$$A\mathbf{x}_2 = \begin{bmatrix} 6 & 5 \\ 1 & 2 \end{bmatrix}\begin{bmatrix} 1 \\ .225 \end{bmatrix} = \begin{bmatrix} 7.125 \\ 1.450 \end{bmatrix}, \ \mu_2 = 7.125$$

- Scale $A\mathbf{x}_2$ by $1/\mu_2$ to get $\mathbf{x}_3$, and so on.

© 2012 Pearson Education, Inc.

The Power Method

- The results of MATLAB calculations for the first five iterations are arranged in Table 2.

- **TABLE 2** The Power Method for Example 2

k	0	1	2	3	4	5
$\mathbf{x}_k$	$\begin{bmatrix} 0 \\ 1 \end{bmatrix}$	$\begin{bmatrix} 1 \\ .4 \end{bmatrix}$	$\begin{bmatrix} 1 \\ .225 \end{bmatrix}$	$\begin{bmatrix} 1 \\ .2035 \end{bmatrix}$	$\begin{bmatrix} 1 \\ .2005 \end{bmatrix}$	$\begin{bmatrix} 1 \\ .20007 \end{bmatrix}$
$A\mathbf{x}_k$	$\begin{bmatrix} 5 \\ 2 \end{bmatrix}$	$\begin{bmatrix} 8 \\ 1.8 \end{bmatrix}$	$\begin{bmatrix} 7.125 \\ 1.450 \end{bmatrix}$	$\begin{bmatrix} 7.0175 \\ 1.4070 \end{bmatrix}$	$\begin{bmatrix} 7.0025 \\ 1.4010 \end{bmatrix}$	$\begin{bmatrix} 7.00036 \\ 1.40014 \end{bmatrix}$
μ_k	5	8	7.125	7.0175	7.0025	7.00036

© 2012 Pearson Education, Inc.

The Power Method

- The evidence from Table 2 strongly suggests that $\{\mathbf{x}_k\}$ approaches (1, .2) and $\{\mu_k\}$ approaches 7.
- If so, then (1, .2) is an eigenvector and 7 is the dominant eigenvalue.
- Thisis easily verified by computing

$$A\begin{bmatrix}1\\.2\end{bmatrix}=\begin{bmatrix}6 & 5\\1 & 2\end{bmatrix}\begin{bmatrix}1\\.2\end{bmatrix}=\begin{bmatrix}7\\1.4\end{bmatrix}=7\begin{bmatrix}1\\.2\end{bmatrix}$$

© 2012 Pearson Education, Inc.

6 Orthogonality and Least Squares

6.1

INNER PRODUCT, LENGTH, AND ORTHOGONALITY

© 2016 Pearson Education, Ltd.

- In order to find an approximate solution to an inconsistent system of equations that has no actual solution, a well-defined notion of nearness is needed
  - Distance, orthogonality (Sec. 6.1)
  - To identify the point within a subspace W that is nearest to a point lying outside of W (Sec. 6.2-6.3)
  - Matrix factorization widely used in numerical linear algebra (Sec. 6.4)
  - Approximate solutions (Sec.6.5)

© 2012 Pearson Education, Ltd.

INNER PRODUCT

- If $\mathbf{u}$ and $\mathbf{v}$ are vectors in $\mathbb{R}^n$, then we regard $\mathbf{u}$ and $\mathbf{v}$ as $n \times 1$ matrices.

- The transpose $\mathbf{u}^T$ is a $1 \times n$ matrix, and the matrix product $\mathbf{u}^T\mathbf{v}$ is a 1×1 matrix, which we write as a single real number (a scalar) without brackets.

- The number $\mathbf{u}^T\mathbf{v}$ is called the **inner product** of $\mathbf{u}$ and $\mathbf{v}$, and it is written as $u \cdot v$.

- This inner product is also referred to as a **dot product**.

© 2016 Pearson Education, Ltd.

INNER PRODUCT

- If $\mathrm{u}=\begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix}$ and $\mathrm{v}=\begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix}$,

  then the inner product of **u** and **v** is

$$\begin{bmatrix} u_1 & u_2 & \cdots & u_n \end{bmatrix}\begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix} = u_1v_1 + u_2v_2 + \cdots + u_nv_n.$$

© 2016 Pearson Education, Ltd.

INNER PRODUCT

- **Theorem 1:** Let **u**, **v**, and **w** be vectors in $\mathbb{R}^n$, and let c be a scalar. Then

  a. $u \cdot v = v \cdot u$

  b. $(u + v) \cdot w = u \cdot w + v \cdot w$

  c. $(cu) \cdot v = c(u \cdot v) = u \cdot (cv)$

  d. $u \cdot u \geq 0$, and $u \cdot u = 0$ if and only if $\mathrm{u} = 0$

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

$$(c_1 u_1 + \ldots + c_p u_p) \cdot w = c_1(u_1 \cdot w) + \cdots + c_p(u_p \cdot w)$$

© 2016 Pearson Education, Ltd.

THE LENGTH OF A VECTOR

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

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

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

- Suppose **v** is in $\mathbb{R}^2$, say, $\mathrm{v} = \begin{bmatrix} a \\ b \end{bmatrix}$

© 2016 Pearson Education, Ltd.

THE LENGTH OF A VECTOR

- If we identify **v** with a geometric point in the plane, as usual, then $\|\mathrm{v}\|$ coincides with the standard notion of the length of the line segment from the origin to **v**.
- This follows from the Pythagorean Theorem applied to a triangle such as the one shown in the following figure.

Interpretation of $\|\mathbf{v}\|$ as length.

- For any scalar c, the length $c\mathbf{v}$ is $|c|$ times the length of v. That is,

$$\|c\mathrm{v}\| = |c|\|\mathrm{v}\|$$

© 2016 Pearson Education, Ltd.

THE LENGTH OF A VECTOR

- A vector whose length is 1 is called a **unit vector**.

- If we *divide* a nonzero vector **v** by its length—that is, multiply by $1/\|\mathbf{v}\|$—we obtain a unit vector **u** because the length of **u** is $(1/\|\mathbf{v}\|)\|\mathbf{v}\|$.

- The process of creating **u** from **v** is sometimes called **normalizing** **v**, and we say that **u** is *in the same direction* as **v**.

© 2016 Pearson Education, Ltd.

THE LENGTH OF A VECTOR

- **Example 2:** Let $\mathrm{v} = (1, -2, 2, 0)$. Find a unit vector **u** in the same direction as **v**.
- **Solution:** First, compute the length of **v**:

$$\|\mathrm{v}\|^2 = v \cdot v = (1)^2 + (-2)^2 + (2)^2 + (0)^2 = 9$$

$$\|\mathrm{v}\| = \sqrt{9} = 3$$

- Then, multiply **v** by $1/\|\mathrm{v}\|$ to obtain

$$\mathrm{u} = \frac{1}{\|\mathrm{v}\|}\mathrm{v} = \frac{1}{3}\mathrm{v} = \frac{1}{3}\begin{bmatrix} 1 \\ -2 \\ 2 \\ 0 \end{bmatrix} = \begin{bmatrix} 1/3 \\ -2/3 \\ 2/3 \\ 0 \end{bmatrix}$$

© 2016 Pearson Education, Ltd.

DISTANCE IN $\mathbb{R}^n$

- To check that $\|\mathrm{u}\| = 1$, it suffices to show that $\|\mathrm{u}\|^2 = 1$.

$$\|\mathrm{u}\|^2 = u \cdot u = \left(\frac{1}{3}\right)^2 + \left(-\frac{2}{3}\right) + \left(\frac{2}{3}\right)^2 + (0)^2$$

$$= \frac{1}{9} + \frac{4}{9} + \frac{4}{9} + 0 = 1$$

- **Definition:** For **u** and **v**in $\mathbb{R}^n$, the **distance between u and v**, written as dist (**u**, **v**), is the length of the vector $\mathrm{u} - \mathrm{v}$. That is,

$$\text{dist}\,(\mathrm{u},\mathrm{v}) = \|\mathrm{u} - \mathrm{v}\|$$

© 2016 Pearson Education, Ltd.

DISTANCE IN $\mathbb{R}^n$

- **Example 4:** Compute the distance between the vectors $\mathbf{u} = (7,1)$ and $\mathbf{v} = (3,2)$.
- **Solution:** Calculate

$$\mathbf{u} - \mathbf{v} = \begin{bmatrix} 7 \\ 1 \end{bmatrix} - \begin{bmatrix} 3 \\ 2 \end{bmatrix} = \begin{bmatrix} 4 \\ -1 \end{bmatrix}$$

$$\|\mathbf{u} - \mathbf{v}\| = \sqrt{4^2 + (-1)^2} = \sqrt{17}$$

- The vectors **u**, **v**, and $\mathbf{u} - \mathbf{v}$ are shown in the figure on the next slide.
- When the vector $\mathbf{u} - \mathbf{v}$ is added to **v**, the result is **u**.

© 2016 Pearson Education, Ltd.

DISTANCE IN $\mathbb{R}^n$

The distance between **u** and **v** is the length of $\mathbf{u} - \mathbf{v}$.

- Notice that the parallelogram in the above figure shows that the distance from **u** to **v** is the same as the distance from $\mathrm{u} - \mathrm{v}$ to **0**.

© 2016 Pearson Education, Ltd.

ORTHOGONAL VECTORS

- Consider $\mathbb{R}^2$or $\mathbb{R}^3$and two lines through the origin determined by vectors **u** and **v**.
- See the figure below. The two lines shown in the figure are geometrically perpendicular if and only if the distance from **u** to **v** is the same as the distance from **u** to $-$v.

- This is the same as requiring the squares of the distances to be the same.

© 2016 Pearson Education, Ltd.

ORTHOGONAL VECTORS

- Now

$$[dist(u,-v)]^2 = \|u-(-v)\|^2 = \|u+v\|^2$$
$$= (u+v)\cdot(u+v)$$
$$= u\cdot(u+v)+v\cdot(u+v) \quad \text{Theorem 1(b)}$$
$$= u\cdot u+u\cdot v+\mathrm{v}\cdot u+v\cdot v \quad \text{Theorem 1(a), (b)}$$
$$= \|u\|^2+\|v\|^2+2\mathrm{u}\cdot v \quad \text{Theorem 1(a)}$$

- The same calculations with **v** and –v interchanged show that

$$\left[\mathrm{dist}(\mathrm{u},\mathrm{v})\right]^2 = \|\mathrm{u}\|^2+\|-\mathrm{v}\|^2+2u\cdot(-v)$$
$$= \|\mathrm{u}\|^2+\|\mathrm{v}\|^2-2u\cdot v$$

© 2016 Pearson Education, Ltd.

ORTHOGONAL VECTORS

- The two squared distances are equal if and only if $2u \cdot v = -2u \cdot v$, which happens if and only if $u \cdot v = 0$.
- This calculation shows that when vectors **u** and **v** are identified with geometric points, the corresponding lines through the points and the origin are perpendicular if and only if $u \cdot v = 0$.
- **Definition:** Two vectors **u** and **v** in $\mathbb{R}^n$ are **orthogonal** (to each other) if $u \cdot v = 0$.
- The zero vector is orthogonal to every vector in $\mathbb{R}^n$ because $0^T \mathrm{v} = 0$ for all **v**.

© 2016 Pearson Education, Ltd.

THE PYTHOGOREAN THEOREM

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

FIGURE 6

© 2016 Pearson Education, Ltd.

ORTHOGONAL COMPLEMENTS

Orthogonal Complements

- If a vector **z** is orthogonal to every vector in a subspace W of $\mathbb{R}^n$, then **z** is said to be **orthogonal to** W.
- The set of all vectors **z** that are orthogonal to W is called the **orthogonal complement** of W and is denoted by $W^{\perp}$ (and read as "W perpendicular" or simply "W perp").

FIGURE 7

A plane and line through **0** as orthogonal complements.

© 2016 Pearson Education, Ltd.

ORTHOGONAL COMPLEMENTS

1. A vector **x** is in $W^{\perp}$ if and only if **x** is orthogonal to every vector in a set that spans W.
2. $W^{\perp}$ is a subspace of $\mathbb{R}^n$.

- **Theorem 3:** Let A be an $m \times n$ matrix. The orthogonal complement of the row space of A is the null space of A, and the orthogonal complement of the column space of A is the null space of A^T:

$$(\text{Row } A)^{\perp} = \text{Nul } A \quad \text{and} \quad (\text{Col } A)^{\perp} = \text{Nul } A^T.$$

© 2016 Pearson Education, Ltd.

FIGURE 8 The fundamental subspaces determined by an $m \times n$ matrix A.

© 2012 Pearson Education, Ltd.

ORTHOGONAL COMPLEMENTS

- **Proof:** The row-column rule for computing $A\mathbf{x}$ shows that if $\mathbf{x}$ is in Nul A, then $\mathbf{x}$ is orthogonal to each row of A (with the rows treated as vectors in $\mathbb{R}^n$).

- Since the rows of A span the row space, $\mathbf{x}$ is orthogonal to Row A.

- Conversely, if $\mathbf{x}$ is orthogonal to Row A, then $\mathbf{x}$ is certainly orthogonal to each row of A, and hence $A\mathrm{x} = 0$.

- This proves the first statement of the theorem.

© 2016 Pearson Education, Ltd.

ORTHOGONAL COMPLEMENTS

- Since this statement is true for any matrix, it is true for A^T.

- That is, the orthogonal complement of the row space of A^T is the null space of A^T.

- This proves the second statement, because $\text{Row } A^T = \text{Col } A$.

© 2016 Pearson Education, Ltd.

ANGLES IN $\mathbb{R}^2$ AND $\mathbb{R}^3$ (OPTIONAL)

- If **u** and **v** are nonzero vectors in either $\mathbb{R}^2$ or $\mathbb{R}^3$, then there is a nice connection between their inner product and the angle ϑ between the two line segments from the origin to the points identified with **u** and **v**.

- The formula is

$$u \cdot v = \|\mathrm{u}\|\|\mathrm{v}\|\cos\vartheta \qquad (2)$$

- To verify this formula for vectors in $\mathbb{R}^2$, consider the triangle shown in the figure on the next slide with sides of lengths, $\|\mathrm{u}\|$, $\|\mathrm{v}\|$, and $\|\mathrm{u}-\mathrm{v}\|$.

© 2016 Pearson Education, Ltd.

ANGLES IN $\mathbb{R}^2$ AND $\mathbb{R}^3$ (OPTIONAL)

The angle between two vectors.

- By the law of cosines,

$$\|\mathrm{u} - \mathrm{v}\|^2 = \|\mathrm{u}\|^2 + \|\mathrm{v}\|^2 - 2\|\mathrm{u}\|\|\mathrm{v}\|\cos\vartheta$$

which can be rearranged to produce the equations on the next slide.

© 2016 Pearson Education, Ltd.

ANGLES IN $\mathbb{R}^2$ AND $\mathbb{R}^3$ (OPTIONAL)

$$\|\mathrm{u}\|\|\mathrm{v}\|\cos\vartheta = \frac{1}{2}\left[\|\mathrm{u}\|^2 + \|\mathrm{v}\|^2 - \|\mathrm{u}-\mathrm{v}\|^2\right]$$

$$= \frac{1}{2}\left[u_1^2 + u_2^2 + v_1^2 + v_2^2 - (u_1 - v_1)^2 - (u_2 - v_2)^2\right]$$

$$= u_1 v_1 + u_2 v_2$$

$$= u \cdot v$$

- The verification for $\mathbb{R}^3$ is similar.
- When $n > 3$, formula (2) may be used to *define* the angle between two vectors in $\mathbb{R}^n$.
- In statistics, the value of $\cos\vartheta$ defined by (2) for suitable vectors **u** and **v** is called a *correlation coefficient*.

© 2016 Pearson Education, Ltd.

6 Orthogonality and Least Squares

6.2

ORTHOGONAL SETS

PEARSON

© 2016 Pearson Education, Inc.

ORTHOGONAL SETS

- A set of vectors $\{\mathbf{u}_1,\ldots,\mathbf{u}_p\}$ in $\mathbb{R}^n$ is said to be an **orthogonal set** if each pair of distinct vectors from the set is orthogonal, that is, if $\mathbf{u}_i \bullet \mathbf{u}_j = 0$ whenever $i \neq j$.

- **Theorem 4:** If $S = \{\mathbf{u}_1,\ldots,\mathbf{u}_p\}$ is an orthogonal set of nonzero vectors in $\mathbb{R}^n$, then S is linearly independent and hence is a basis for the subspace spanned by S.

© 2016 Pearson Education, Inc.

ORTHOGONAL SETS

- **Proof:** If $0 = c_1\mathbf{u}_1 + \cdots + c_p\mathbf{u}_p$ for some scalars $c_1,\ldots,c_p$, then

$$\begin{aligned} 0 = \mathbf{0}\bullet\mathbf{u}_1 &= (c_1\mathbf{u}_1 + c_2\mathbf{u}_2 + \cdots + c_p\mathbf{u}_p)\bullet\mathbf{u}_1 \\ &= (c_1\mathbf{u}_1)\bullet\mathbf{u}_1 + (c_2\mathbf{u}_2)\bullet\mathbf{u}_1 + \cdots + (c_p\mathbf{u}_p)\bullet\mathbf{u}_1 \\ &= c_1(\mathbf{u}_1\bullet\mathbf{u}_1) + c_2(\mathbf{u}_2\bullet\mathbf{u}_1) + \cdots + c_p(\mathbf{u}_p\bullet\mathbf{u}_1) \\ &= c_1(\mathbf{u}_1\bullet\mathbf{u}_1) \end{aligned}$$

  because $\mathbf{u}_1$ is orthogonal to $\mathbf{u}_2,\ldots,\mathbf{u}_p$.
- Since $\mathbf{u}_1$ is nonzero, $\mathbf{u}_1\bullet\mathbf{u}_1$ is not zero and so $c_1 = 0$.
- Similarly, $c_2,\ldots,c_p$ must be zero.

© 2016 Pearson Education, Inc.

ORTHOGONAL SETS

- Thus S is linearly independent.
- **Definition:** An **orthogonal basis** for a subspace W of $\mathbb{R}^n$ is a basis for W that is also an orthogonal set.

- **Theorem 5:** Let $\{\mathbf{u}_1,\ldots,\mathbf{u}_p\}$ be an orthogonal basis for a subspace W of $\mathbb{R}^n$. For each $\mathbf{y}$ in W, the weights in the linear combination

$$\mathbf{y} = c_1\mathbf{u}_1 + \cdots + c_p\mathbf{u}_p$$

are given by

$$c_j = \frac{\mathbf{y}\bullet\mathbf{u}_j}{\mathbf{u}_j\bullet\mathbf{u}_j} \qquad (j = 1,\ldots,p)$$

© 2016 Pearson Education, Inc.

ORTHOGONAL SETS

- **Proof:** The orthogonality of $\{\mathbf{u}_1,\ldots,\mathbf{u}_p\}$ shows that

$$\mathbf{y}\bullet\mathbf{u}_1 = (c_1\mathbf{u}_1 + c_2\mathbf{u}_2 + \cdots + c_p\mathbf{u}_p)\bullet\mathbf{u}_1 = c_1(\mathbf{u}_1\bullet\mathbf{u}_1)$$

- Since $\mathbf{u}_1\bullet\mathbf{u}_1$ is not zero, the equation above can be solved for c_1.

- To find c_j for $j = 2,\ldots, p$, compute $\mathbf{y}\bullet\mathbf{u}_j$ and solve for c_j.

© 2016 Pearson Education, Inc.

AN ORTHOGONAL PROJECTION

- Given a nonzero vector **u** in $\mathbb{R}^n$, consider the problem of decomposing a vector **y** in $\mathbb{R}^n$ into the sum of two vectors, one a multiple of **u** and the other orthogonal to **u**.
- We wish to write

$$\mathbf{y} = \hat{\mathbf{y}} + \mathbf{z} \qquad \text{----(1)}$$

where $\hat{\mathbf{y}} = \alpha\mathbf{u}$ for some scalar α and **z** is some vector orthogonal to **u**. See the following figure.

Finding α to make $\mathbf{y} - \hat{\mathbf{y}}$ orthogonal to $\mathbf{u}$.

© 2016 Pearson Education, Inc.

AN ORTHOGONAL PROJECTION

- Given any scalar α, let $\mathbf{z} = \mathbf{y} - \alpha\mathbf{u}$, so that (1) is satisfied.
- Then $\mathbf{y} - \hat{\mathbf{y}}$ is orthogonal to **u** if an only if

$$0 = (\mathbf{y} - \alpha\mathbf{u})\bullet\mathbf{u} = \mathbf{y}\bullet\mathbf{u} - (\alpha\mathbf{u})\bullet\mathbf{u} = \mathbf{y}\bullet\mathbf{u} - \alpha(\mathbf{u}\bullet\mathbf{u})$$

- That is, (1) is satisfied with **z** orthogonal to **u** if and

$$\text{only if } \alpha = \frac{\mathbf{y}\bullet\mathbf{u}}{\mathbf{u}\bullet\mathbf{u}} \text{ and } \hat{\mathbf{y}} = \frac{\mathbf{y}\bullet\mathbf{u}}{\mathbf{u}\bullet\mathbf{u}}\mathbf{u}.$$

- The vector $\hat{\mathbf{y}}$ is called the **orthogonal projection of y onto u**, and the vector **z** is called the **component of y orthogonal to u**.

© 2016 Pearson Education, Inc.

AN ORTHOGONAL PROJECTION

- If c is any nonzero scalar and if $\mathbf{u}$ is replaced by $c\mathbf{u}$ in the definition of $\hat{\mathbf{y}}$, then the orthogonal projection of $\mathbf{y}$ onto $c\mathbf{u}$ is exactly the same as the orthogonal projection of $\mathbf{y}$ onto $\mathbf{u}$.
- Hence this projection is determined by the *subspace* L spanned by $\mathbf{u}$ (the line through $\mathbf{u}$ and $\mathbf{0}$).
- Sometimes $\hat{\mathbf{y}}$ is denoted by $\text{proj}_L\mathbf{y}$ and is called the **orthogonal projection of y onto** L.
- That is,

$$\hat{\mathbf{y}} = \text{proj}_L\mathbf{y} = \frac{\mathbf{y}\bullet\mathbf{u}}{\mathbf{u}\bullet\mathbf{u}}\mathbf{u} \qquad \text{----(2)}$$

© 2016 Pearson Education, Inc.

AN ORTHOGONAL PROJECTION

- **Example 1:** Let $\mathbf{y} = \begin{bmatrix} 7 \\ 6 \end{bmatrix}$ and $\mathbf{u} = \begin{bmatrix} 4 \\ 2 \end{bmatrix}$. Find the orthogonal projection of **y** onto **u**. Then write **y** as the sum of two orthogonal vectors, one in Span {**u**} and one orthogonal to **u**.
- **Solution:** Compute

$$\mathbf{y} \bullet \mathbf{u} = \begin{bmatrix} 7 \\ 6 \end{bmatrix} \bullet \begin{bmatrix} 4 \\ 2 \end{bmatrix} = 40$$

$$\mathbf{u} \bullet \mathbf{u} = \begin{bmatrix} 4 \\ 2 \end{bmatrix} \bullet \begin{bmatrix} 4 \\ 2 \end{bmatrix} = 20$$

© 2016 Pearson Education, Inc.

AN ORTHOGONAL PROJECTION

- The orthogonal projection of **y** onto **u** is

$$\hat{\mathbf{y}} = \frac{\mathbf{y}\bullet\mathbf{u}}{\mathbf{u}\bullet\mathbf{u}}\mathbf{u} = \frac{40}{20}\mathbf{u} = 2\begin{bmatrix}4\\2\end{bmatrix} = \begin{bmatrix}8\\4\end{bmatrix}$$

and the component of **y** orthogonal to **u** is

$$\mathbf{y} - \hat{\mathbf{y}} = \begin{bmatrix}7\\6\end{bmatrix} - \begin{bmatrix}8\\4\end{bmatrix} = \begin{bmatrix}-1\\2\end{bmatrix}$$

- The sum of these two vectors is **y**.

© 2016 Pearson Education, Inc.

AN ORTHOGONAL PROJECTION

- That is,

$$\begin{bmatrix} 7 \\ 6 \end{bmatrix} = \begin{bmatrix} 8 \\ 4 \end{bmatrix} + \begin{bmatrix} -1 \\ 2 \end{bmatrix}$$

$$\mathbf{y} \qquad \hat{\mathbf{y}} \qquad (\mathbf{y} - \hat{\mathbf{y}})$$

- The decomposition of **y** is illustrated in the following figure.

The orthogonal projection of **y** onto a line L through the origin.

© 2016 Pearson Education, Inc.

AN ORTHOGONAL PROJECTION

- *Note:* If the calculations above are correct, then $\{\hat{\mathbf{y}}, \mathbf{y} - \hat{\mathbf{y}}\}$ will be an orthogonal set.

- As a check, compute

$$\hat{\mathbf{y}} \bullet (\mathbf{y} - \hat{\mathbf{y}}) = \begin{bmatrix} 8 \\ 4 \end{bmatrix} \bullet \begin{bmatrix} -1 \\ 2 \end{bmatrix} = -8 + 8 = 0$$

- Since the line segment in the figure on the previous slide between $\mathbf{y}$ and $\hat{\mathbf{y}}$ is perpendicular to L, by construction of $\hat{\mathbf{y}}$, the point identified with $\hat{\mathbf{y}}$ is the closest point of L to $\mathbf{y}$.

© 2016 Pearson Education, Inc.

ORTHONORMAL SETS

- A set $\{\mathbf{u}_1,\ldots,\mathbf{u}_p\}$ is an **orthonormal set** if it is an orthogonal set of unit vectors.

- If W is the subspace spanned by such a set, then $\{\mathbf{u}_1,\ldots,\mathbf{u}_p\}$ is an **orthonormal basis** for W, since the set is automatically linearly independent, by Theorem 4.

- The simplest example of an orthonormal set is the standard basis $\{\mathbf{e}_1,\ldots,\mathbf{e}_n\}$ for $\mathbb{R}^n$.

- Any nonempty subset of $\{\mathbf{e}_1,\ldots,\mathbf{e}_n\}$ is orthonormal, too.

© 2016 Pearson Education, Inc.

ORTHONORMAL SETS

- **Example 2:** Show that $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ is an orthonormal basis of $\mathbb{R}^3$, where

$$\mathbf{v}_1 = \begin{bmatrix} 3/\sqrt{11} \\ 1/\sqrt{11} \\ 1/\sqrt{11} \end{bmatrix}, \ \mathbf{v}_2 = \begin{bmatrix} -1/\sqrt{6} \\ 2/\sqrt{6} \\ 1/\sqrt{6} \end{bmatrix}, \ \mathbf{v}_3 = \begin{bmatrix} -1/\sqrt{66} \\ -4/\sqrt{66} \\ 7/\sqrt{66} \end{bmatrix}$$

- **Solution:** Compute

$$\mathbf{v}_1 \bullet \mathbf{v}_2 = -3/\sqrt{66} + 2/\sqrt{66} + 1/\sqrt{66} = 0$$

$$\mathbf{v}_1 \bullet \mathbf{v}_3 = -3/\sqrt{726} - 4/\sqrt{726} + 7/\sqrt{726} = 0$$

© 2016 Pearson Education, Inc.

ORTHONORMAL SETS

$$\mathbf{v}_2 \bullet \mathbf{v}_3 = 1/\sqrt{396} - 8/\sqrt{396} + 7/\sqrt{396} = 0$$

- Thus $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ is an orthogonal set.
- Also, $\mathbf{v}_1 \bullet \mathbf{v}_1 = 9/11 + 1/11 + 1/11 = 1$

$$\mathbf{v}_2 \bullet \mathbf{v}_2 = 1/6 + 4/6 + 1/6 = 1$$

$$\mathbf{v}_3 \bullet \mathbf{v}_3 = 1/66 + 16/66 + 49/66 = 1$$

  which shows that $\mathbf{v}_1$, $\mathbf{v}_2$, and $\mathbf{v}_3$ are unit vectors.
- Thus $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ is an orthonormal set.
- Since the set is linearly independent, its three vectors form a basis for $\mathbb{R}^3$. See the figure on the next slide.

© 2016 Pearson Education, Inc.

ORTHONORMAL SETS

- When the vectors in an orthogonal set of nonzero vectors are *normalized* to have unit length, the new vectors will still be orthogonal, and hence the new set will be an orthonormal set.

© 2016 Pearson Education, Inc.

ORTHONORMAL SETS

- **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 $\mathbb{R}^m$.
- Let $U = \begin{bmatrix} \mathbf{u}_1 & \mathbf{u}_2 & \mathbf{u}_3 \end{bmatrix}$ and compute

$$U^T U = \begin{bmatrix} \mathbf{u}_1^T \\ \mathbf{u}_2^T \\ \mathbf{u}_3^T \end{bmatrix} \begin{bmatrix} \mathbf{u}_1 & \mathbf{u}_2 & \mathbf{u}_3 \end{bmatrix} = \begin{bmatrix} \mathbf{u}_1^T\mathbf{u}_1 & \mathbf{u}_1^T\mathbf{u}_2 & \mathbf{u}_1^T\mathbf{u}_3 \\ \mathbf{u}_2^T\mathbf{u}_1 & \mathbf{u}_2^T\mathbf{u}_2 & \mathbf{u}_2^T\mathbf{u}_3 \\ \mathbf{u}_3^T\mathbf{u}_1 & \mathbf{u}_3^T\mathbf{u}_2 & \mathbf{u}_3^T\mathbf{u}_3 \end{bmatrix} \quad \text{----(3)}$$

© 2016 Pearson Education, Inc.

ORTHONORMAL SETS

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

$$\mathbf{u}_1^T\mathbf{u}_2 = \mathbf{u}_2^T\mathbf{u}_1 = 0, \mathbf{u}_1^T\mathbf{u}_3 = \mathbf{u}_3^T\mathbf{u}_1 = 0, \mathbf{u}_2^T\mathbf{u}_3 = \mathbf{u}_3^T\mathbf{u}_2 = 0 \qquad \text{----(4)}$$

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

$$\mathbf{u}_1^T\mathbf{u}_1 = 1,\ \mathbf{u}_2^T\mathbf{u}_2 = 1,\ \mathbf{u}_3^T\mathbf{u}_3 = 1 \qquad \text{----(5)}$$

- The theorem follows immediately from (3)–(5).

© 2016 Pearson Education, Inc.

ORTHONORMAL SETS

- **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

  a. $\|U\mathbf{x}\| = \|\mathbf{x}\|$

  b. $(U\mathbf{x})\bullet(U\mathbf{y}) = \mathbf{x}\bullet\mathbf{y}$

  c. $(U\mathbf{x})\bullet(U\mathbf{y}) = 0$ if and only if $\mathbf{x}\bullet\mathbf{y} = 0$

- Properties (a) and (c) say that the linear mapping $\mathbf{x} \mapsto U\mathbf{x}$ preserves lengths and orthogonality.

© 2016 Pearson Education, Inc.

6 Orthogonality and Least Squares

6.3

ORTHOGONAL PROJECTIONS

PEARSON

© 2016 Pearson Education, Inc.

ORTHOGONAL PROJECTIONS

- The orthogonal projection of a point in $\mathbb{R}^2$ onto a line through the origin has an important analogue in $\mathbb{R}^n$.

- Given a vector $\mathbf{y}$ and a subspace W in $\mathbb{R}^n$, there is a vector $\hat{\mathbf{y}}$ in W such that (1) $\hat{\mathbf{y}}$ is the unique vector in W for which $\mathbf{y} - \hat{\mathbf{y}}$ is orthogonal to W, and (2) $\hat{\mathbf{y}}$ is the unique vector in W closest to $\mathbf{y}$. See the following figure.

© 2016 Pearson Education, Inc.

THE ORTHOGONAL DECOMPOSITION THEOREM

- These two properties of $\hat{\mathrm{y}}$ provide the key to finding the least-squares solutions of linear systems.
- **Theorem 8:** Let W be a subspace of $\mathbb{R}^n$. Then each $\mathbf{y}$ in $\mathbb{R}^n$ can be written uniquely in the form

$$\mathrm{y} = \hat{\mathrm{y}} + \mathrm{z} \qquad (1)$$

where $\hat{\mathrm{y}}$ is in W and $\mathbf{z}$ is in $W^{\perp}$.

- In fact, if $\{\mathbf{u}_1,\ldots,\mathbf{u}_p\}$ is any orthogonal basis of W, then

$$\hat{\mathrm{y}} = \frac{y \cdot u_1}{u_1 \cdot u_1}\mathrm{u}_1 + \cdots + \frac{y \cdot u_p}{u_p \cdot u_p}\mathrm{u}_p \qquad (2)$$

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

© 2016 Pearson Education, Inc.

THE ORTHOGONAL DECOMPOSITION THEOREM

- The vector $\hat{\mathbf{y}}$ in (1) is called the **orthogonal projection of y onto** W and often is written as $\text{proj}_W\mathbf{y}$. See the following figure:

The orthogonal projection of $\mathbf{y}$ onto W.

- **Proof:** Let $\{\mathbf{u}_1,\ldots,\mathbf{u}_p\}$ be any orthogonal basis for W, and define $\hat{\mathbf{y}}$ by (2).
- Then $\hat{\mathbf{y}}$ is in W because $\hat{\mathbf{y}}$ is a linear combination of the basis $\mathbf{u}_1,\ldots,\mathbf{u}_p$.

© 2016 Pearson Education, Inc.

THE ORTHOGONAL DECOMPOSITION THEOREM

- Let $z = y - \hat{y}$.
- Since $\mathbf{u}_1$ is orthogonal to $\mathbf{u}_2,\ldots,\mathbf{u}_p$, it follows from (2) that

$$z \cdot u_1 = (y - \hat{y}) \cdot u_1 = y \cdot u_1 - \left(\frac{y \cdot u_1}{u_1 \cdot u_1}\right) u_1 \cdot u_1 - 0 \ldots 0 - 0$$

$$= y \cdot u_1 - y \cdot u_1 = 0$$

- Thus $\mathbf{z}$ is orthogonal to $\mathbf{u}_1$.
- Similarly, $\mathbf{z}$ is orthogonal to each $\mathbf{u}_j$ in the basis for W.
- Hence $\mathbf{z}$ is orthogonal to every vector in W.
- That is, $\mathbf{z}$ is in $W^{\perp}$.

© 2016 Pearson Education, Inc.

THE ORTHOGONAL DECOMPOSITION THEOREM

- To show that the decomposition in (1) is unique, suppose **y** can also be written as $y = \hat{y}_1 + z_1$, with $\hat{y}_1$ in W and $\mathbf{z}_1$ in $W^{\perp}$.
- Then $\hat{y} + z = \hat{y}_1 + z_1$ (since both sides equal **y**), and so

$$\hat{y} - \hat{y}_1 = z_1 - z$$

- This equality shows that the vector $v = \hat{y} - \hat{y}_1$ is in W and in $W^{\perp}$(because $\mathbf{z}_1$ and **z** are both in $W^{\perp}$, and $W^{\perp}$ is a subspace).
- Hence $v \cdot v = 0$, which shows that $v = 0$.
- This proves that $\hat{y} = \hat{y}_1$ and also $z_1 = z$.

© 2016 Pearson Education, Inc.

THE ORTHOGONAL DECOMPOSITION THEOREM

- The uniqueness of the decomposition (1) shows that the orthogonal projection $\hat{\mathrm{y}}$ depends only on W and not on the particular basis used in (2).

- **Example 1:** Let $\mathrm{u}_1 = \begin{bmatrix} 2 \\ 5 \\ -1 \end{bmatrix}, \mathrm{u}_2 = \begin{bmatrix} -2 \\ 1 \\ 1 \end{bmatrix}$, and $\mathrm{y} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$.

  Observe that $\{\mathbf{u}_1, \mathbf{u}_2\}$ is an orthogonal basis for $W = \mathrm{Span}\{\mathrm{u}_1, \mathrm{u}_2\}$. Write $\mathbf{y}$ as the sum of a vector in W and a vector orthogonal to W.

© 2016 Pearson Education, Inc.

THE ORTHOGONAL DECOMPOSITION THEOREM

- **Solution:** The orthogonal projection of **y** onto W is

$$\hat{\mathrm{y}} = \frac{y \cdot u_1}{u_1 \cdot u_1}\mathrm{u}_1 + \frac{y \cdot u_2}{u_2 \cdot u_2}\mathrm{u}_2$$

$$= \frac{9}{30}\begin{bmatrix} 2 \\ 5 \\ -1 \end{bmatrix} + \frac{3}{6}\begin{bmatrix} -2 \\ 1 \\ 1 \end{bmatrix} = \frac{9}{30}\begin{bmatrix} 2 \\ 5 \\ -1 \end{bmatrix} + \frac{15}{30}\begin{bmatrix} -2 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} -2/5 \\ 2 \\ 1/5 \end{bmatrix}$$

- Also

$$\mathrm{y} - \hat{\mathrm{y}} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} - \begin{bmatrix} -2/5 \\ 2 \\ 1/5 \end{bmatrix} = \begin{bmatrix} 7/5 \\ 0 \\ 14/5 \end{bmatrix}$$

© 2016 Pearson Education, Inc.

THE ORTHOGONAL DECOMPOSITION THEOREM

- Theorem 8 ensures that $\mathrm{y} - \hat{\mathrm{y}}$ is in $W^{\perp}$.

- To check the calculations, verify that $\mathrm{y} - \hat{\mathrm{y}}$ is orthogonal to both $\mathbf{u}_1$ and $\mathbf{u}_2$ and hence to all of W.

- The desired decomposition of $\mathbf{y}$ is

$$\mathrm{y} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} = \begin{bmatrix} -2/5 \\ 2 \\ 1/5 \end{bmatrix} + \begin{bmatrix} 7/5 \\ 0 \\ 14/5 \end{bmatrix}$$

© 2016 Pearson Education, Inc.

PROPERTIES OF ORTHOGONAL PROJECTIONS

- If $\{\mathbf{u}_1,\ldots,\mathbf{u}_p\}$ is an orthogonal basis for W and if $\mathbf{y}$ happens to be in W, then the formula for $\text{proj}_W\mathbf{y}$ is exactly the same as the representation of $\mathbf{y}$ given in Theorem 5 in Section 6.2.

- In this case, $\text{proj}_W \mathrm{y} = \mathrm{y}$.

- If $\mathbf{y}$ is in $W = \text{Span}\{\mathrm{u}_1,\ldots,\mathrm{u}_p\}$, then $\text{proj}_W \mathrm{y} = \mathrm{y}$.

© 2016 Pearson Education, Inc.

THE BEST APPROXIMATION THEOREM

- **Theorem 9:** Let W be a subspace of $\mathbb{R}^n$, let **y** be any vector in $\mathbb{R}^n$, and let $\hat{\mathbf{y}}$ be the orthogonal projection of **y** onto W. Then $\hat{\mathbf{y}}$ is the closest point in W to **y**, in the sense that

$$\|\mathbf{y} - \hat{\mathbf{y}}\| < \|\mathbf{y} - \mathbf{v}\| \qquad (3)$$

  for all **v** in W distinct from $\hat{\mathbf{y}}$.
- The vector $\hat{\mathbf{y}}$ in Theorem 9 is called **the best approximation to y by elements of** W.
- The distance from **y** to **v**, given by $\|\mathbf{y} - \mathbf{v}\|$, can be regarded as the "error" of using **v** in place of **y**.
- Theorem 9 says that this error is minimized when $\mathbf{v} = \hat{\mathbf{y}}$.

© 2016 Pearson Education, Inc.

THE BEST APPROXIMATION THEOREM

- Inequality (3) leads to a new proof that $\hat{\mathbf{y}}$ does not depend on the particular orthogonal basis used to compute it.

- If a different orthogonal basis for W were used to construct an orthogonal projection of $\mathbf{y}$, then this projection would also be the closest point in W to $\mathbf{y}$, namely, $\hat{\mathbf{y}}$.

© 2016 Pearson Education, Inc.

THE BEST APPROXIMATION THEOREM

- **Proof:** Take **v** in W distinct from $\hat{y}$. See the following figure:

The orthogonal projection of **y** onto W is the closest point in W to **y**.

- Then $\hat{y} - v$ is in W.
- By the Orthogonal Decomposition Theorem, $y - \hat{y}$ is orthogonal to W.
- In particular, $y - \hat{y}$ is orthogonal to $\hat{y} - v$ (which is in W).

© 2016 Pearson Education, Inc.

THE BEST APPROXIMATION THEOREM

- Since
$$y - v = (y - \hat{y}) + (\hat{y} - v)$$
the Pythagorean Theorem gives
$$\|y - v\|^2 = \|y - \hat{y}\|^2 + \|\hat{y} - v\|^2$$
- (See the colored right triangle in the figure on the previous slide. The length of each side is labeled.)
- Now $\|\hat{y} - v\|^2 > 0$ because $\hat{y} - v \neq 0$, and so inequality (3) follows immediately.

© 2016 Pearson Education, Inc.

PROPERTIES OF ORTHOGONAL PROJECTIONS

- **Example 4:** The distance from a point **y** in $\mathbb{R}^n$ to a subspace W is defined as the distance from **y** to the nearest point in W. Find the distance from **y** to $W = \text{Span}\{\mathbf{u}_1, \mathbf{u}_2\}$, where

$$\mathbf{y} = \begin{bmatrix} -1 \\ -5 \\ 10 \end{bmatrix}, \mathbf{u}_1 = \begin{bmatrix} 5 \\ -2 \\ 1 \end{bmatrix}, \mathbf{u}_2 = \begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix}$$

- **Solution:** By the Best Approximation Theorem, the distance from **y** to W is $\|\mathbf{y} - \hat{\mathbf{y}}\|$, where $\hat{\mathbf{y}} = \text{proj}_W \mathbf{y}$.

© 2016 Pearson Education, Inc.

PROPERTIES OF ORTHOGONAL PROJECTIONS

- Since $\{\mathbf{u}_1, \mathbf{u}_2\}$ is an orthogonal basis for W,

$$\hat{y} = \frac{15}{30}u_1 + \frac{-21}{6}u_2 = \frac{1}{2}\begin{bmatrix} 5 \\ -2 \\ 1 \end{bmatrix} - \frac{7}{2}\begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix} = \begin{bmatrix} -1 \\ -8 \\ 4 \end{bmatrix}$$

$$y - \hat{y} = \begin{bmatrix} -1 \\ -5 \\ 10 \end{bmatrix} - \begin{bmatrix} -1 \\ -8 \\ 4 \end{bmatrix} = \begin{bmatrix} 0 \\ 3 \\ 6 \end{bmatrix}$$

$$\|y - \hat{y}\|^2 = 3^2 + 6^2 = 45$$

- The distance from $\mathbf{y}$ to W is $\sqrt{45} = 3\sqrt{5}$.

© 2016 Pearson Education, Inc.

PROPERTIES OF ORTHOGONAL PROJECTIONS

- **Theorem 10:** If $\{\mathbf{u}_1,\ldots,\mathbf{u}_p\}$ is an orthogonal basis for a subspace W of $\mathbb{R}^n$, then

$$proj_W y = (y \cdot u_1)u_1 + (y \cdot u_2)u_2 + \cdots + (y \cdot u_p)u_p \quad (4)$$

If $U = \begin{bmatrix} \mathrm{u}_1 & \mathrm{u}_2 & \cdots & \mathrm{u}_p \end{bmatrix}$, then

$$\mathrm{proj}_W \mathrm{y} = UU^T \mathrm{y} \quad \text{for all } \mathbf{y} \text{ in } \mathbb{R}^n \qquad (5)$$

- **Proof:** Formula (4) follows immediately from (2) in Theorem 8.

© 2016 Pearson Education, Inc.

PROPERTIES OF ORTHOGONAL PROJECTIONS

- Also, (4) shows that $\text{proj}_W\mathbf{y}$ is a linear combination of the columns of U using the weights $y \cdot u_1, y \cdot u_2, \ldots y \cdot u_p$.

- The weights can be written as $\mathrm{u}_1^T\mathrm{y}, \mathrm{u}_2^T\mathrm{y}, \ldots, \mathrm{u}_p^T\mathrm{y}$, showing that they are the entries in $U^T\mathbf{y}$ and justifying (5).

© 2016 Pearson Education, Inc.

6 Orthogonality and Least Squares

6.4

THE GRAM-SCHMIDT PROCESS

PEARSON

© 2016 Pearson Education, Inc.

THE GRAM-SCHMIDT PROCESS

- **Theorem 11: The Gram-Schmidt Process**
- Given a basis $\{x_1, \ldots, x_p\}$ for a nonzero subspace W of $\mathbb{R}^n$, define

$$v_1 = x_1$$

$$v_2 = x_2 - \frac{x_2 \cdot v_1}{v_1 \cdot v_1} v_1$$

$$v_3 = x_3 - \frac{x_3 \cdot v_1}{v_1 \cdot v_1} v_1 - \frac{x_3 \cdot v_2}{v_2 \cdot v_2} v_2$$

$$\vdots$$

$$v_p = xp - \frac{x_p \cdot v_1}{v_1 \cdot v_1} v_1 - \frac{x_p \cdot v_2}{v_2 \cdot v_2} v_2 - \ldots - \frac{x_p \cdot v_{p_1}}{v_{p_1} \cdot v_{p_1}} v_{p_1}$$

- Then $\{v_1, \ldots, v_p\}$ is an orthogonal basis for W. In addition

$$\text{Span}\{v_1, \ldots, v_k\} = \text{Span}\{x_1, \ldots, x_k\} \quad \text{for } 1 \le k \le p \qquad (1)$$

© 2016 Pearson Education, Inc.

THE GRAM-SCHMIDT PROCESS

- **Proof** For $1 \leq k \leq p$, let $W_k = \text{Span}\{x_1, \ldots, x_k\}$. Set $v_1 = x_1$, so that $\text{Span}\{v_1\} = \text{Span}\{x_1\}$. Suppose, for some $k < p$, we have constructed $v_1, \ldots, v_k$ so that $\{v_1, \ldots, v_k\}$ is an orthogonal basis for W_k. Define

$$v_k+1 = x_{k+1} - \text{proj}_{W_k} x_{k+1} \qquad (2)$$

- By the Orthogonal Decomposition Theorem, v_{k+1} is orthogonal to W_k. Furthermore, $v_{k+1} \neq 0$ because x_{k+1} is not in $W_k = \text{Span}\{x_1, \ldots, x_k\}$
- Hence $\{v_1, \ldots, v_k\}$ is an orthogonal set of nonzero vectors in the (k + 1)-dimensional space W_{k+1}. By the Basis Theorem in Section 4.5, this set is an orthogonal basis for W_{k+1}. Hence $W_{k+1} = \text{Span}\{v_1, \ldots, v_{k+1}\}$. When $k + 1 = p$, the process stops.

© 2016 Pearson Education, Inc.

ORTHONORMAL BASES

- **Example 3** Example 1 constructed the orthogonal basis

$$v_1 = \begin{bmatrix} 3 \\ 6 \\ 0 \end{bmatrix}, \; v_2 = \begin{bmatrix} 0 \\ 0 \\ 2 \end{bmatrix}$$

- An orthonormal basis is

$$u_1 = \frac{1}{\|v_1\|}\mathrm{v}_1 = \frac{1}{\sqrt{45}}\begin{bmatrix} 3 \\ 6 \\ 0 \end{bmatrix} = \begin{bmatrix} 1/\sqrt{5} \\ 2/\sqrt{5} \\ 0 \end{bmatrix}$$

$$u_2 = \frac{1}{\|v_2\|}\mathrm{v}_2 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

© 2016 Pearson Education, Inc.

QR FACTORIZATION OF MATRICES

- **Theorem 12: The QR Factorization**
- If A is an $m \times n$ matrix with linearly independent columns, then A can be factored as $A = QR$, where Q is an $m \times n$ matrix whose columns form an orthonormal basis for Col A and R is an $n \times n$ upper triangular invertible matrix with positive entries on its diagonal.

- **Proof** The columns of A form a basis $\{x_1, \ldots, x_n\}$ for Col A. Construct an orthonormal basis $\{u_1, \ldots, u_n\}$ for W = Col A with property (1) in Theorem 11. This basis may be constructed by the Gram-Schmidt process or some other means.

© 2016 Pearson Education, Inc.

QR FACTORIZATION OF MATRICES

- Let

$$Q = [\mathrm{u}_1 \quad \mathrm{u}_2 \quad \ldots \quad \mathrm{un}]$$

- For $k = 1, \ldots, \mathrm{x}_k$ is in Span$\{\mathrm{x}_1, \ldots, \mathrm{x}_k\}$ = Span$\{\mathrm{u}_1, \ldots, \mathrm{u}_k\}$. So there are constants, $\mathrm{r}_{1k}, \ldots, \mathrm{r}_{kk}$, such that

$$x_k = r_{1ku1} {}_{+ \cdots +} rkkuk + 0 \cdot uk {}_{+1} + \cdots + 0 \cdot un$$

- We may assume that $\mathrm{r}_{kk} \geq 0$. This shows that x_k is a linear combination of the columns of Q using as weights the entries in the vector

© 2016 Pearson Education, Inc.

QR FACTORIZATION OF MATRICES

$$rk = \begin{bmatrix} r1k \\ . \\ . \\ . \\ rkk \\ 0 \\ . \\ . \\ . \\ 0 \end{bmatrix}$$

- That is, $x_k = Qr_k$ for $k = 1, \ldots, n$. Let $R = [r_1 \ldots r_n]$. Then
$$A = [x_1 \ldots x_n] = [Qr_1 \ldots Qr_n] = QR$$
- The fact that R is invertible follows easily from the fact that the columns of A are linearly independent. Since R is clearly upper triangular, its nonnegative diagonal entries must be positive.

© 2016 Pearson Education, Inc.

QR FACTORIZATION OF MATRICES

- **Example 4** Find a QR factorization of $A = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$.
- **Solution** The columns of A are the vectors x_1, x_2, and x_3 in Example 2. An orthogonal basis for Col A = Span$\{x_1, x_2, x_3\}$ was found in that example:

$$v_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}, \quad v_2 = \begin{bmatrix} -3 \\ 1 \\ 1 \\ 1 \end{bmatrix}, \quad v_3 = \begin{bmatrix} 0 \\ -2/3 \\ 1/3 \\ 1/3 \end{bmatrix}$$

© 2016 Pearson Education, Inc.

QR FACTORIZATION OF MATRICES

- To simplify the arithmetic that follows, scale v_3 by letting $v_3 = 3v_3$. Then normalize the three vectors to obtain u_1, u_2, and u_3, and use these vectors as the columns of Q:

$$Q = \begin{bmatrix} 1/2 & -3/\sqrt{12} & 0 \\ 1/2 & 1/\sqrt{12} & -2/\sqrt{6} \\ 1/2 & 1/\sqrt{12} & 1/\sqrt{6} \\ 1/2 & 1/\sqrt{12} & 1/\sqrt{6} \end{bmatrix}.$$

- By construction, the first k columns of Q are an orthonormal basis of Span$\{x_1, \ldots, x_k\}$.

© 2016 Pearson Education, Inc.

QR FACTORIZATION OF MATRICES

- From the proof of Theorem 12, $A = QR$ for some R. To find R, observe that $Q^TQ = I$, because the columns of Q are orthonormal. Hence

$$Q^TA = QT(QR) = IR = R$$

- and

$$\mathrm{R} = \begin{bmatrix} 1/2 & 1/2 & 1/2 & 1/2 \\ -3/\sqrt{12} & 1/\sqrt{12} & 1/\sqrt{12} & 1/\sqrt{12} \\ 0 & -2/\sqrt{6} & 1/\sqrt{12} & 1/\sqrt{6} \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$$

$$= \begin{bmatrix} 2 & 3/2 & 1 \\ 0 & 3/\sqrt{12} & 2/\sqrt{12} \\ 0 & 0 & 3/\sqrt{12} \end{bmatrix}$$

© 2016 Pearson Education, Inc.

6 Orthogonality and Least Squares

6.5

LEAST-SQUARES PROBLEMS

PEARSON

© 2016 Pearson Education, Inc.

LEAST-SQUARES PROBLEMS

- **Definition:** If A is $m \times n$ and **b** is in $\mathbb{R}^m$, a **least-squares solution** of $A\mathrm{x} = \mathrm{b}$ is an $\hat{\mathrm{x}}$ in $\mathbb{R}^n$ such that

$$\|\mathrm{b} - A\hat{\mathrm{x}}\| \le \|\mathrm{b} - A\mathrm{x}\|$$

  for all **x** in $\mathbb{R}^n$.

- The most important aspect of the least-squares problem is that no matter what **x** we select, the vector A**x** will necessarily be in the column space, Col A.

- So we seek an **x** that makes A**x** the closest point in Col A to **b**. See the figure on the next slide.

© 2016 Pearson Education, Inc.

LEAST-SQUARES PROBLEMS

The vector **b** is closer to $A\hat{\mathbf{x}}$ than to $A\mathbf{x}$ for other **x**.

- **Solution of the General Least-Squares Problem**

- Given A and **b**, apply the Best Approximation Theorem to the subspace Col A.

- Let $$\hat{\mathrm{b}} = \mathrm{proj}_{\mathrm{Col}\,A}\mathrm{b}$$

© 2016 Pearson Education, Inc.

SOLUTION OF THE GENREAL LEAST-SQUARES PROBLEM

- Because $\hat{\mathrm{b}}$ is in the column space A, the equation $A\mathrm{x} = \hat{\mathrm{b}}$ *is* consistent, and there is an $\hat{\mathrm{x}}$ in $\mathbb{R}^n$ such that

$$A\hat{\mathrm{x}} = \hat{\mathrm{b}} \qquad (1)$$

- Since $\hat{\mathrm{b}}$ is the closest point in Col A to **b**, a vector $\hat{\mathrm{x}}$ is a least-squares solution of $A\mathrm{x} = \mathrm{b}$ if and only if $\hat{\mathrm{x}}$ satisfies (1).

- Such an $\hat{\mathrm{x}}$ in $\mathbb{R}^n$ is a list of weights that will build $\hat{\mathrm{b}}$ out of the columns of A. See the figure on the next slide.

© 2016 Pearson Education, Inc.

SOLUTION OF THE GENREAL LEAST-SQUARES PROBLEM

The least-squares solution $\hat{\mathbf{x}}$ is in $\mathbb{R}^n$.

- Suppose $\hat{\mathrm{x}}$ satisfies $A\hat{\mathrm{x}} = \hat{\mathrm{b}}$.
- By the Orthogonal Decomposition Theorem, the projection $\hat{\mathrm{b}}$ has the property that $\mathrm{b} - \hat{\mathrm{b}}$ is orthogonal to Col A, so $\mathrm{b} - A\hat{\mathrm{x}}$ is orthogonal to each column of A.
- If $\mathbf{a}_j$ is any column of A, then $a_j \cdot (b - A\hat{x}) = 0$, and $a_j^T(b - A\hat{x}) = 0$.

© 2016 Pearson Education, Inc.

SOLUTION OF THE GENREAL LEAST-SQUARES PROBLEM

- Since each $\mathbf{a}_j^T$ is a row of A^T,

$$A^T(\mathbf{b} - A\hat{\mathbf{x}}) = 0 \qquad (2)$$

- Thus

$$A^T\mathbf{b} - A^T A\hat{\mathbf{x}} = 0$$

$$A^T A\hat{\mathbf{x}} = A^T\mathbf{b}$$

- These calculations show that each least-squares solution of $A\mathbf{x} = \mathbf{b}$ satisfies the equation

$$A^T A\mathbf{x} = A^T\mathbf{b} \qquad (3)$$

- The matrix equation (3) represents a system of equations called the **normal equations** for $A\mathbf{x} = \mathbf{b}$.
- A solution of (3) is often denoted by $\hat{\mathbf{x}}$.

© 2016 Pearson Education, Inc.

SOLUTION OF THE GENREAL LEAST-SQUARES PROBLEM

- **Theorem 13:** The set of least-squares solutions of $A\mathrm{x} = \mathrm{b}$ coincides with the nonempty set of solutions of the normal equation $A^T A\mathrm{x} = A^T\mathrm{b}$.
- **Proof:** The set of least-squares solutions is nonempty and each least-squares solution $\hat{\mathrm{x}}$ satisfies the normal equations.
- Conversely, suppose $\hat{\mathrm{x}}$ satisfies $A^T A\hat{\mathrm{x}} = A^T\mathrm{b}$.
- Then $\hat{\mathrm{x}}$ satisfies (2), which shows that $\mathrm{b} - A\hat{\mathrm{x}}$ is orthogonal to the rows of A^T and hence is orthogonal to the columns of A.
- Since the columns of A span Col A, the vector $\mathrm{b} - A\hat{\mathrm{x}}$ is orthogonal to all of Col A.

© 2016 Pearson Education, Inc.

SOLUTION OF THE GENREAL LEAST-SQUARES PROBLEM

- Hence the equation

$$\mathrm{b} = A\hat{\mathrm{x}} + (\mathrm{b} - A\hat{\mathrm{x}})$$

  is a decomposition of **b** into the sum of a vector in Col A and a vector orthogonal to Col A.

- By the uniqueness of the orthogonal decomposition, $A\hat{\mathrm{x}}$ must be the orthogonal projection of **b** onto Col A.

- That is, $A\hat{\mathrm{x}} = \hat{\mathrm{b}}$ and $\hat{\mathrm{x}}$ is a least-squares solution.

© 2016 Pearson Education, Inc.

SOLUTION OF THE GENREAL LEAST-SQUARES PROBLEM

- **Example 1:** Find a least-squares solution of the inconsistent system $A\mathrm{x} = \mathrm{b}$ for

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

- **Solution:** To use normal equations (3), compute:

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

© 2016 Pearson Education, Inc.

SOLUTION OF THE GENREAL LEAST-SQUARES PROBLEM

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

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

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

© 2016 Pearson Education, Inc.

SOLUTION OF THE GENREAL LEAST-SQUARES PROBLEM

- Row operations can be used to solve the system on the previous slide, 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 solve $A^T A\mathrm{x} = A^T \mathrm{b}$ as

$$\hat{\mathrm{x}} = (A^T A)^{-1} A^T \mathrm{b}$$

$$= \frac{1}{84}\begin{bmatrix} 5 & -1 \\ -1 & 17 \end{bmatrix}\begin{bmatrix} 19 \\ 11 \end{bmatrix} = \frac{1}{84}\begin{bmatrix} 84 \\ 168 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$$

© 2016 Pearson Education, Inc.

SOLUTION OF THE GENREAL LEAST-SQUARES PROBLEM

- **Theorem 14:** Let A be an $m \times n$ matrix. The following statements are logically equivalent:
  a. The equation $A\mathrm{x} = \mathrm{b}$ has a unique least-squares solution for each **b** in $\mathbb{R}^m$.
  b. The columns of A are linearly independent.
  c. The matrix $A^T A$ is invertible.

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

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

- When a least-squares solution $\hat{\mathrm{x}}$ is used to produce $A\hat{\mathrm{x}}$ as an approximation to **b**, the distance from **b** to $A\hat{\mathrm{x}}$ is called the **least-squares error** of this approximation.

© 2016 Pearson Education, Inc.

ALTERNATIVE CALCULATIONS OF LEAST-SQUARES SOLUTIONS

- **Example 4:** Find a least-squares solution of $A\mathrm{x} = \mathrm{b}$ for

$$A = \begin{bmatrix} 1 & -6 \\ 1 & -2 \\ 1 & 1 \\ 1 & 7 \end{bmatrix}, \mathrm{b} = \begin{bmatrix} -1 \\ 2 \\ 1 \\ 6 \end{bmatrix}$$

- **Solution:** Because the columns $\mathbf{a}_1$ and $\mathbf{a}_2$ of A are orthogonal, the orthogonal projection of $\mathbf{b}$ onto Col A is given by

$$\hat{b} = \frac{b \cdot a_1}{a_1 \cdot a_1} a_1 + \frac{b \cdot a_2}{a_2 \cdot a_2} a_2 = \frac{8}{4} a_1 + \frac{45}{90} a_2 \qquad (5)$$

© 2016 Pearson Education, Inc.

ALTERNATIVE CALCULATIONS OF LEAST-SQUARES SOLUTIONS

$$= \begin{bmatrix} 2 \\ 2 \\ 2 \\ 2 \end{bmatrix} + \begin{bmatrix} -3 \\ -1 \\ 1/2 \\ 7/2 \end{bmatrix} = \begin{bmatrix} -1 \\ 1 \\ 5/2 \\ 11/2 \end{bmatrix}$$

- Now that $\hat{\mathrm{b}}$ is known, we can solve $A\hat{\mathrm{x}} = \hat{\mathrm{b}}$.
- But this is trivial, since we already know weights to place on the columns of A to produce $\hat{\mathrm{b}}$.
- It is clear from (5) that

$$\hat{x} = \begin{bmatrix} 8/4 \\ 45/90 \end{bmatrix} = \begin{bmatrix} 2 \\ 1/2 \end{bmatrix}$$

© 2016 Pearson Education, Inc.

ALTERNATIVE CALCULATIONS OF LEAST-SQUARES SOLUTIONS

- **Theorem 15:** Given an $m \times n$ matrix A with linearly independent columns, let $A = QR$ be a QR factorization of A. Then, for each **b** in $\mathbb{R}^m$, the equation $A\mathrm{x} = \mathrm{b}$ has a unique least-squares solution, given by

$$\hat{\mathrm{x}} = R^{-1}Q^T\mathrm{b} \qquad (6)$$

- **Proof:** Let $\hat{\mathrm{x}} = R^{-1}Q^T\mathrm{b}$.
- Then

$$A\hat{\mathrm{x}} = QR\hat{\mathrm{x}} = QRR^{-1}Q^T\mathrm{b} = QQ^T\mathrm{b}$$

© 2016 Pearson Education, Inc.

ALTERNATIVE CALCULATIONS OF LEAST-SQUARES SOLUTIONS

- The columns of Q form an orthonormal basis for Col A.
- Hence, by Theorem 10, $QQ^T\mathbf{b}$ is the orthogonal projection $\hat{\mathbf{b}}$ of $\mathbf{b}$ onto Col A.
- Then $A\hat{\mathbf{x}} = \hat{\mathbf{b}}$, which shows that $\hat{\mathbf{x}}$ is a least-squares solution of $A\mathbf{x} = \mathbf{b}$.
- The uniqueness of $\hat{\mathbf{x}}$ follows from Theorem 14.

© 2016 Pearson Education, Inc.

7 Symmetric Matrices and Quadratic Forms

7.1

DIAGONALIZATION OF SYMMETRIC MATRICES

PEARSON

© 2016 Pearson Education, Inc.

SYMMETRIC MATRIX

- A **symmetric matrix** is a matrix A such that $A^T = A$.
- Such a matrix is necessarily square.
- Its main diagonal entries are arbitrary, but its other entries occur in pairs—on opposite sides of the main diagonal.

© 2016 Pearson Education, Inc.

SYMMETRIC MATRIX

- **Theorem 1:** If A is symmetric, then any two eigenvectors from different eigenspaces are orthogonal.
- **Proof:** Let $\mathbf{v}_1$ and $\mathbf{v}_2$ be eigenvectors that correspond to distinct eigenvalues, say, λ_1 and λ_2.
- To show that $v_1.v_2 = 0$, compute

$$\lambda_1 v_1.v_2 = (\lambda_1 v_1)^T v_2 = (Av_1)^T v_2 \quad \text{Since } \mathbf{v}_1 \text{ is an eigenvector}$$

$$= (v_1^T A^T)\, v_2 = v_1^T (Av_2) \quad \text{Since } A^T = A$$

$$= v_1^T (\lambda_2 v_2) \quad \text{Since } \mathbf{v}_2 \text{ is an eigenvector}$$

$$= \lambda_2 v_1^T v_2 = \lambda_2 v_1.v_2$$

© 2016 Pearson Education, Inc.

SYMMETRIC MATRIX

- Hence $(\lambda_1 - \lambda_2)\, v_1.v_2 = 0$
- But $\lambda_1 - \lambda_2 \neq 0$, so $v_1.v_2 = 0$.
- An $n \times n$ matrix A is said to be **orthogonally diagonalizable** if there are an orthogonal matrix P (with $P^{-1} = P^T$) and a diagonal matrix D such that

$$A = PDP^T = PDP^{-1} \qquad (1)$$

- Such a diagonalization requires n linearly independent and orthonormal eigenvectors.
- When is this possible?
- If A is orthogonally diagonalizable as in (1), then

$$A^T = (PDP^T)^T = P^{TT} D^T P^T = PDP^T = A$$

© 2016 Pearson Education, Inc.

SYMMETRIC MATRIX

- Thus A is symmetric!
- **Theorem 2:** An $n \times n$ matrix A is orthogonally diagonalizable if and only if A is symmetric matrix.

- **Example 3:** Orthogonally diagonalize the matrix

$$A = \begin{bmatrix} 3 & -2 & 4 \\ -2 & 6 & 2 \\ 4 & 2 & 3 \end{bmatrix}, \text{ whose characteristic equation is}$$

$$0 = -\lambda^3 + 12\lambda^2 - 21\lambda - 98 = -(\lambda - 7)^2(\lambda + 2)$$

© 2016 Pearson Education, Inc.

SYMMETRIC MATRIX

- **Solution:** The usual calculations produce bases for the eigenspaces:

$$\lambda = 7 : \mathrm{v}_1 = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \mathrm{v}_2 = \begin{bmatrix} -1/2 \\ 1 \\ 0 \end{bmatrix}; \lambda = -2 : \mathrm{v}_3 = \begin{bmatrix} -1 \\ -1/2 \\ 1 \end{bmatrix}$$

- Although $\mathbf{v}_1$ and $\mathbf{v}_2$ are linearly independent, they are not orthogonal.
- The projection of $\mathbf{v}_2$ onto $\mathbf{v}_1$ is $\frac{v_2 \cdot v_1}{v_1 \cdot v_1} v_1$.

© 2016 Pearson Education, Inc.

SYMMETRIC MATRIX

- The component of $\mathbf{v}_2$ orthogonal to $\mathbf{v}_1$ is

$$\mathbf{z}_2 = \mathbf{v}_2 - \frac{v_2 \cdot v_1}{v_1 \cdot v_1} v_1 = \begin{bmatrix} -1/2 \\ 1 \\ 0 \end{bmatrix} - \frac{-1/2}{2} \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} -1/4 \\ 1 \\ 1/4 \end{bmatrix}$$

- Then $\{\mathbf{v}_1, \mathbf{z}_2\}$ is an orthogonal set in the eigenspace for $\lambda = 7$.

- (Note that $\mathbf{z}_2$ is linear combination of the eigenvectors $\mathbf{v}_1$ and $\mathbf{v}_2$, so $\mathbf{z}_2$ is in the eigenspace).

© 2016 Pearson Education, Inc.

SYMMETRIC MATRIX

- Since the eigenspace is two-dimensional (with basis $\mathbf{v}_1$, $\mathbf{v}_2$), the orthogonal set $\{\mathbf{v}_1, \mathbf{z}_2\}$ is an *orthogonal basis* for the eigenspace, by the Basis Theorem.

- Normalize $\mathbf{v}_1$ and $\mathbf{z}_2$ to obtain the following orthonormal basis for the eigenspace for $\lambda = 7$:

$$\mathrm{u}_1 = \begin{bmatrix} 1/\sqrt{2} \\ 0 \\ 1/\sqrt{2} \end{bmatrix}, \mathrm{u}_2 = \begin{bmatrix} -1/\sqrt{18} \\ 4/\sqrt{18} \\ 1/\sqrt{18} \end{bmatrix}$$

© 2016 Pearson Education, Inc.

SYMMETRIC MATRIX

- An orthonormal basis for the eigenspace for $\lambda = -2$ is

$$\mathbf{u}_3 = \frac{1}{\|2\mathbf{v}_3\|} 2\mathbf{v}_3 = \frac{1}{3}\begin{bmatrix} -2 \\ -1 \\ 2 \end{bmatrix} = \begin{bmatrix} -2/3 \\ -1/3 \\ 2/3 \end{bmatrix}$$

- By Theorem 1, $\mathbf{u}_3$ is orthogonal to the other eigenvectors $\mathbf{u}_1$ and $\mathbf{u}_2$.

- Hence $\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\}$ is an orthonormal set.

© 2016 Pearson Education, Inc.

SYMMETRIC MATRIX

- Let

$$P = \begin{bmatrix} \mathrm{u}_1 & \mathrm{u}_2 & \mathrm{u}_3 \end{bmatrix} = \begin{bmatrix} 1/\sqrt{2} & -1/\sqrt{18} & -2/3 \\ 0 & 4/\sqrt{18} & -1/3 \\ 1/\sqrt{2} & 1/\sqrt{18} & 2/3 \end{bmatrix}, D = \begin{bmatrix} 7 & 0 & 0 \\ 0 & 7 & 0 \\ 0 & 0 & -2 \end{bmatrix}$$

- Then P orthogonally diagonalizes A, and $A = PDP^{-1}$.

© 2016 Pearson Education, Inc.

THE SPECTRAL THEOREM

- The set if eigenvalues of a matrix A is sometimes called the *spectrum* of A, and the following description of the eigenvalues is called a *spectral theorem*.

- **Theorem 3:** The Spectral Theorem for Symmetric Matrices
- An $n \times n$ symmetric matrix A has the following properties:
  a. A has n real eigenvalues, counting multiplicities.

© 2016 Pearson Education, Inc.

THE SPECTRAL THEOREM

b. The dimension of the eigenspace for each eigenvalue λ equals the multiplicity of λ as a root of the characteristic equation.

c. The eigenspaces are mutually orthogonal, in the sense that eigenvectors corresponding to different eigenvalues are orthogonal.

d. A is orthogonally diagonalizable.

© 2016 Pearson Education, Inc.

SPECTRAL DECOMPOSITION

- Suppose $A = PDP^{-1}$, where the columns of P are orthonormal eigenvectors $\mathbf{u}_1,\ldots,\mathbf{u}_n$ of A and the corresponding eigenvalues $\lambda_1,\ldots,\lambda_n$ are in the diagonal matrix D.
- Then, since $P^{-1} = P^T$,

$$A = PDP^T = \begin{bmatrix} \mathrm{u}_1 & \cdots & \mathrm{u}_n \end{bmatrix} \begin{bmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{bmatrix} \begin{bmatrix} \mathrm{u}_1^T \\ \vdots \\ \mathrm{u}_n^T \end{bmatrix}$$

$$= \begin{bmatrix} \lambda_1 \mathrm{u}_1 & \cdots & \lambda_n \mathrm{u}_n \end{bmatrix} \begin{bmatrix} \mathrm{u}_1^T \\ \vdots \\ \mathrm{u}_n^T \end{bmatrix}$$

© 2016 Pearson Education, Inc.

SPECTRAL DECOMPOSITION

- Using the column-row expansion of a product, we can write

$$A = \lambda_1 \mathbf{u}_1 \mathbf{u}_1^T + \lambda_2 \mathbf{u}_2 \mathbf{u}_2^T + \cdots + \lambda_n \mathbf{u}_n \mathbf{u}_n^T \qquad (2)$$

- This representation of A is called a **spectral decomposition** of A because it breaks up A into pieces determined by the spectrum (eigenvalues) of A.
- Each term in (2) is an $n \times n$ matrix of rank 1.
- For example, every column of $\lambda_1 \mathbf{u}_1 \mathbf{u}_1^T$ is a multiple of $\mathbf{u}_1$.
- Each matrix $\mathbf{u}_j \mathbf{u}_j^T$ is a **projection matrix** in the sense that for each $\mathbf{x}$ in $\mathbb{R}^n$, the vector $(\mathbf{u}_j \mathbf{u}_j^T)\mathbf{x}$ is the orthogonal projection of $\mathbf{x}$ onto the subspace spanned by $\mathbf{u}_j$.

© 2016 Pearson Education, Inc.

SPECTRAL DECOMPOSITION

- **Example 4:** Construct a spectral decomposition of the matrix A that has the orthogonal diagonalization

$$A = \begin{bmatrix} 7 & 2 \\ 2 & 4 \end{bmatrix} = \begin{bmatrix} 2/\sqrt{5} & -1/\sqrt{5} \\ 1/\sqrt{5} & 2/\sqrt{5} \end{bmatrix} \begin{bmatrix} 8 & 0 \\ 0 & 3 \end{bmatrix} \begin{bmatrix} 2/\sqrt{5} & 1/\sqrt{5} \\ -1/\sqrt{5} & 2/\sqrt{5} \end{bmatrix}$$

- **Solution:** Denote the columns of P by $\mathbf{u}_1$ and $\mathbf{u}_2$.

- Then $$A = 8\mathrm{u}_1\mathrm{u}_1^T + 3\mathrm{u}_2\mathrm{u}_2^T$$

© 2016 Pearson Education, Inc.

SPECTRAL DECOMPOSITION

- To verify the decomposition of A, compute

$$\mathbf{u}_1\mathbf{u}_1^T = \begin{bmatrix} 2/\sqrt{5} \\ 1/\sqrt{5} \end{bmatrix} \begin{bmatrix} 2/\sqrt{5} & 1/\sqrt{5} \end{bmatrix} = \begin{bmatrix} 4/5 & 2/5 \\ 2/5 & 1/5 \end{bmatrix}$$

$$\mathbf{u}_2\mathbf{u}_2^T = \begin{bmatrix} -1/\sqrt{5} \\ 2/\sqrt{5} \end{bmatrix} \begin{bmatrix} -1/\sqrt{5} & 2/\sqrt{5} \end{bmatrix} = \begin{bmatrix} 1/5 & -2/5 \\ -2/5 & 4/5 \end{bmatrix}$$

and

$$8\mathbf{u}_1\mathbf{u}_1^T + 3\mathbf{u}_2\mathbf{u}_2^T = \begin{bmatrix} 32/5 & 16/5 \\ 16/5 & 8/5 \end{bmatrix} + \begin{bmatrix} 3/5 & -6/5 \\ -6/5 & 12/5 \end{bmatrix} = \begin{bmatrix} 7 & 2 \\ 2 & 4 \end{bmatrix} = A$$

© 2016 Pearson Education, Inc.

7 Symmetric Matrices and Quadratic Forms

7.2

QUADRATIC FORMS

© 2016 Pearson Education, Ltd.

QUADRATIC FORMS

- A **quadratic form** on $\mathbb{R}^n$ is a function Q defined on $\mathbb{R}^n$ whose value at a vector $\mathbf{x}$ in $\mathbb{R}^n$ can be computed by an expression of the form $Q(\mathrm{x}) = \mathrm{x}^T A\mathrm{x}$, where A is an $n \times n$ symmetric matrix.

- The matrix A is called the **matrix of the quadratic form**.

© 2016 Pearson Education, Ltd.

QUADRATIC FORMS

- **Example 1:** Let $\mathrm{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$. Compute $\mathbf{x}^T A\mathbf{x}$ for the following matrices.

  a. $A = \begin{bmatrix} 4 & 0 \\ 0 & 3 \end{bmatrix}$

  b. $A = \begin{bmatrix} 3 & -2 \\ -2 & 7 \end{bmatrix}$

© 2016 Pearson Education, Ltd.

QUADRATIC FORMS

- **Solution:**

a. $\mathrm{x}^T A\mathrm{x} = \begin{bmatrix} x_1 & x_2 \end{bmatrix} \begin{bmatrix} 4 & 0 \\ 0 & 3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} x_1 & x_2 \end{bmatrix} \begin{bmatrix} 4x_1 \\ 3x_2 \end{bmatrix} = 4x_1^2 + 3x_2^2.$

b. There are two -2 entries in A.

$$\mathrm{x}^T A\mathrm{x} = \begin{bmatrix} x_1 & x_2 \end{bmatrix} \begin{bmatrix} 3 & -2 \\ -2 & 7 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} x_1 & x_2 \end{bmatrix} \begin{bmatrix} 3x_1 - 2x_2 \\ -2x_1 + 7x_2 \end{bmatrix}$$

$$= x_1(3x_1 - 2x_2) + x_2(-2x_1 + 7x_2)$$

$$= 3x_1^2 - 2x_1x_2 - 2x_2x_1 + 7x_2^2$$

$$= 3x_1^2 - 4x_1x_2 + 7x_2^2$$

© 2016 Pearson Education, Ltd.

QUADRATIC FORMS

- The presence of $-4x_1x_2$ in the quadratic form in Example 1(b) is due to the -2 entries off the diagonal in the matrix A.

- In contrast, the quadratic form associated with the diagonal matrix A in Example 1(a) has no x_1x_2 *cross-product* term.

© 2016 Pearson Education, Ltd.

CHANGE OF VARIBALE IN A QUADRATIC FORM

- If **x** represents a variable vector in $\mathbb{R}^n$, then a **change of variable** is an equation of the form

$$\mathrm{x} = P\mathrm{y}, \text{ or equivalently, } \mathrm{y} = P^{-1}\mathrm{x} \qquad (1)$$

  where P is an invertible matrix and **y** is a new variable vector in $\mathbb{R}^n$.

- Here **y** is the coordinate vector of **x** relative to the basis of $\mathbb{R}^n$ determined by the columns of P.

- If the change of variable (1) is made in a quadratic form $\mathbf{x}^T A\mathbf{x}$, then

$$(2) \quad \mathrm{x}^T A\mathrm{x} = (P\mathrm{y})^T A(P\mathrm{y}) = \mathrm{y}^T P^T AP\mathrm{y} = \mathrm{y}^T (P^T AP)\mathrm{y}$$

  and the new matrix of the quadratic form is $P^T AP$.

© 2016 Pearson Education, Ltd.

CHANGE OF VARIBALE IN A QUADRATIC FORM

- Since A is symmetric, Theorem 2 guarantees that there is an *orthogonal* matrix P such that P^TAP is a diagonal matrix D, and the quadratic form in (2) becomes $\mathbf{y}^TD\mathbf{y}$.

- **Example 4:** Make a change of variable that transforms the quadratic form $Q(\mathrm{x}) = x_1^2 - 8x_1x_2 - 5x_2^2$ into a quadratic form with no cross-product term.

- **Solution:** The matrix of the given quadratic form is

$$A = \begin{bmatrix} 1 & -4 \\ -4 & -5 \end{bmatrix}$$

© 2016 Pearson Education, Ltd.

CHANGE OF VARIBALE IN A QUADRATIC FORM

- The first step is to orthogonally diagonalize A.
- Its eigenvalues turn out to be $\lambda = 3$ and $\lambda = -7$.
- Associated unit eigenvectors are

$$\lambda = 3 : \begin{bmatrix} 2/\sqrt{5} \\ -1/\sqrt{5} \end{bmatrix}; \lambda = -7 : \begin{bmatrix} 1/\sqrt{5} \\ 2/\sqrt{5} \end{bmatrix}$$

- These vectors are automatically orthogonal (because they correspond to distinct eigenvalues) and so provide an orthonormal basis for $\mathbb{R}^2$.

© 2016 Pearson Education, Ltd.

CHANGE OF VARIBALE IN A QUADRATIC FORM

- Let

$$P = \begin{bmatrix} 2/\sqrt{5} & 1/\sqrt{5} \\ -1/\sqrt{5} & 2/\sqrt{5} \end{bmatrix}, D = \begin{bmatrix} 3 & 0 \\ 0 & -7 \end{bmatrix}$$

- Then $A = PDP^{-1}$ and $D = P^{-1}AP = P^{T}AP.$

- A suitable change of variable is

$$\mathrm{x} = P\mathrm{y}, \text{where } \mathrm{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \text{ and } \mathrm{y} = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}.$$

© 2016 Pearson Education, Ltd.

CHANGE OF VARIBALE IN A QUADRATIC FORM

- Then

$$x_1^2 - 8x_1x_2 - 5x_2^2 = \mathrm{x}^T A\mathrm{x} = (P\mathrm{y})^T A(P\mathrm{y})$$
$$= \mathrm{y}^T P^T AP\mathrm{y} = \mathrm{y}^T D\mathrm{y}$$
$$= 3\mathrm{y}_1^2 - 7\mathrm{y}_2^2$$

- To illustrate the meaning of the equality of quadratic forms in Example 4, we can compute $Q(\mathbf{x})$ for $\mathrm{x} = (2, -2)$ using the new quadratic form.

© 2016 Pearson Education, Ltd.

CHANGE OF VARIBALE IN A QUADRATIC FORM

- First, since $\mathrm{x} = P\mathrm{y}$,

$$\mathrm{y} = P^{-1}\mathrm{x} = P^{T}\mathrm{x}$$

so

$$y = \begin{bmatrix} 2/\sqrt{5} & -1/\sqrt{5} \\ 1/\sqrt{5} & 2/\sqrt{5} \end{bmatrix} \begin{bmatrix} 2 \\ -2 \end{bmatrix} = \begin{bmatrix} 6/\sqrt{5} \\ -2/\sqrt{5} \end{bmatrix}$$

- Hence

$$3\mathrm{y}_1^2 - 7\mathrm{y}_2^2 = 3(6/\sqrt{5})^2 - 7(-2/\sqrt{5})^2 = 3(36/5) - 7(4/5)$$
$$= 80/5 = 16$$

- This is the value of $Q(\mathbf{x})$ when $\mathrm{x} = (2, -2)$.

© 2016 Pearson Education, Ltd.

THE PRINCIPAL AXIS THEOREM

- See the figure below.

Change of variable in $\mathbf{x}^T A\mathbf{x}$.

- **Theorem 4:** Let A be an $n \times n$ symmetric matrix. Then there is an orthogonal change of variable, $\mathrm{x} = P\mathrm{y}$, that transforms the quadratic form $\mathbf{x}^T A\mathbf{x}$ into a quadratic form $\mathbf{y}^T D\mathbf{y}$ with no cross-product term.

© 2016 Pearson Education, Ltd.

THE PRINCIPAL AXIS THEOREM

- The columns of P in theorem 4 are called the **principal axes** of the quadratic form $\mathbf{x}^T A\mathbf{x}$.

- The vector $\mathbf{y}$ is the coordinate vector of $\mathbf{x}$ relative to the orthonormal basis of $\mathbb{R}^n$ given by these principal axes.

A Geometric View of Principal Axes

- Suppose $Q(\mathrm{x}) = \mathrm{x}^T A\mathrm{x}$, where A is an invertible 2×2 symmetric matrix, and let c be a constant.

© 2016 Pearson Education, Ltd.

A GEOMETRIC VIEW OF PRINCIPAL AXES

- It can be shown that the set of all **x** in $\mathbb{R}^2$ that satisfy

(3) $$\mathrm{x}^T A\mathrm{x} = c$$

either corresponds to an ellipse (or circle), a hyperbola, two intersecting lines, or a single point, or contains no points at all.

- If A is a diagonal matrix, the graph is in *standard position*, such as in the figure below.

An ellipse and a hyperbola in standard position.

© 2016 Pearson Education, Ltd.

A GEOMETRIC VIEW OF PRINCIPAL AXES

- If A is not a diagonal matrix, the graph of equation (3) is rotated out of standard position, as in the figure below.

(a) $5x_1^2 - 4x_1x_2 + 5x_2^2 = 48$ (b) $x_1^2 - 8x_1x_2 - 5x_2^2 = 16$

An ellipse and a hyperbola *not* in standard position.

- Finding the *principal axes* (determined by the eigenvectors of A) amounts to finding a new coordinate system with respect to which the graph is in standard position.

© 2016 Pearson Education, Ltd.

CLASSIFYING QUADRATIC FORMS

- **Definition:** A quadratic form Q is:
  a. **positive definite** if $Q(\mathrm{x}) > 0$ for all $\mathrm{x} \neq 0$,
  b. **negative definite** if $Q(\mathrm{x}) < 0$ for all $\mathrm{x} \neq 0$,
  c. **indefinite** if $Q(\mathbf{x})$ assumes both positive and negative values.

- Also, Q is said to be **positive semidefinite** if $Q(\mathrm{x}) \geq 0$ for all **x**, and **negative semidefinite** if $Q(\mathrm{x}) \leq 0$ for all **x**.

© 2016 Pearson Education, Ltd.

QUADRATIC FORMS AND EIGENVALUES

- **Theorem 5:** Let A be an $n \times n$ symmetric matrix. Then a quadratic form $\mathbf{x}^T A\mathbf{x}$ is:
  a. positive definite if and only if the eigenvalues of A are all positive,
  b. negative definite if and only if the eigenvalues of A are all negative, or
  c. indefinite if and only if A has both positive and negative eigenvalues.

© 2016 Pearson Education, Ltd.

QUADRATIC FORMS AND EIGENVALUES

- **Proof:** By the Principal Axes Theorem, there exists an orthogonal change of variable $\mathrm{x} = P\mathrm{y}$ such that

$$Q(\mathrm{x}) = \mathrm{x}^T A\mathrm{x} = \mathrm{y}^T D\mathrm{y} = \lambda_1 y_1^2 + \lambda_2 y_2^2 + \cdots + \lambda_n y_n^2 \quad (4)$$

where $\lambda_1,\ldots,\lambda_n$ are the eigenvalues of A.

- Since P is invertible, there is a one-to-one correspondence between all nonzero **x** and all nonzero **y**.

© 2016 Pearson Education, Ltd.

QUADRATIC FORMS AND EIGENVALUES

- Thus the values of $Q(\mathbf{x})$ for $\mathrm{x} \neq 0$ coincide with the values of the expression on the right side of (4), which is controlled by the signs of the eigenvalues $\lambda_1,\ldots,\lambda_n$, in three ways described in the theorem.

© 2016 Pearson Education, Ltd.