الجامعة السعودية الالكترونية
Saudi Electronic University
2011 1432

Elementary Linear Algebra

Name: ________________________
CRN: ________________________
Year: ________________________
Teacher: ________________________

WEEK 2

1.1 Introduction to Systems of Linear Equations.
1.2 Gaussian Elimination (in a separate PPT file).
1.3 Matrices and Matrix Operations (in a separate PPT file).

1.1 Introduction to Systems of Linear Equations

a linear equation in n variables:

$$a_1x_1 + a_2x_2 + a_3x_3 + \cdots + a_nx_n = b$$

$a_1, a_2, a_3, \ldots, a_n$, b: real number

a_1: leading coefficient

x_1: leading variable

Notes:

(1) Linear equations have no products or roots of variables and no variables involved in trigonometric, exponential, or logarithmic functions.

(2) Variables appear only to the first power.

- Ex 1: (Linear or Nonlinear)

Linear $(a)\ 3x + 2y = 7$ — $(b)\ \frac{1}{2}x + y - \pi z = \sqrt{2}$ Linear

Linear $(c)\ x_1 - 2x_2 + 10x_3 + x_4 = 0$ — $(d)\ (\sin\frac{\pi}{2})x_1 - 4x_2 = e^2$ Linear

Nonlinear $(e)\ xy + z = 2$ (not the first power) — $(f)\ e^x - 2y = 4$ (Exponential) Nonlinear

Nonlinear $(g)\ \sin x_1 + 2x_2 - 3x_3 = 0$ (trigonometric functions) — $(h)\ \frac{1}{x} + \frac{1}{y} = 4$ (not the first power) Nonlinear

- a system of m linear equations in n variables:

$$\begin{array}{ccccccccccc} a_{11}x_1 & + & a_{12}x_2 & + & a_{13}x_3 & + & \cdots & + & a_{1n}x_n & = & b_1 \\ a_{21}x_1 & + & a_{22}x_2 & + & a_{23}x_3 & + & \cdots & + & a_{2n}x_n & = & b_2 \\ a_{31}x_1 & + & a_{32}x_2 & + & a_{33}x_3 & + & \cdots & + & a_{3n}x_n & = & b_3 \\ & & & & \vdots & & & & & & \\ a_{m1}x_1 & + & a_{m2}x_2 & + & a_{m3}x_3 & + & \cdots & + & a_{mn}x_n & = & b_m \end{array}$$

- Consistent:

  A system of linear equations has at least one solution.

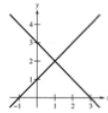
- Inconsistent:

  A system of linear equations has no solution.

Every system of linear equations has either

(1) $\begin{aligned} x + y &= 3 \\ x - y &= -1 \end{aligned}$

two intersecting lines



exactly one solution

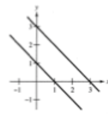
(2) $\begin{aligned} x + y &= 3 \\ 2x + 2y &= 6 \end{aligned}$

two coincident lines

inifinite number

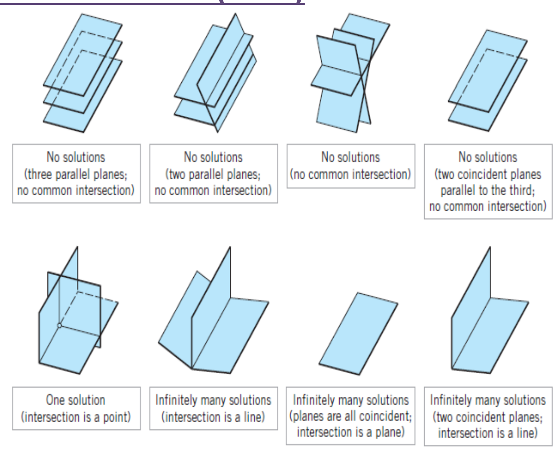
(3) $\begin{aligned} x + y &= 3 \\ x + y &= 1 \end{aligned}$

two parallel lines



no solution

Linear Systems in Three Unknowns (Read)



▲ Figure 1.1.2

Homogeneous Systems: is the system where all equations are set = 0. (b=0)

Theorem 1.2.2 A homogeneous linear system with more unknowns than equations has infinitely many solutions

---

Trace of a matrix

DEFINITION 8 If A is a square matrix, then the ***trace of*** A, denoted by $\text{tr}(A)$, is defined to be the sum of the entries on the main diagonal of A. The trace of A is undefined if A is not a square matrix.

▶ **EXAMPLE 11 Trace of a Matrix**

The following are examples of matrices and their traces.

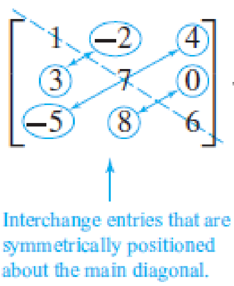
$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}, \quad B = \begin{bmatrix} -1 & 2 & 7 & 0 \\ 3 & 5 & -8 & 4 \\ 1 & 2 & 7 & -3 \\ 4 & -2 & 1 & 0 \end{bmatrix}$$

$$\text{tr}(A) = a_{11} + a_{22} + a_{33} \qquad \text{tr}(B) = -1 + 5 + 7 + 0 = 11$$ ◀

Transpose of a Matrix AT

DEFINITION 7 If A is any $m \times n$ matrix, then the ***transpose of*** A, denoted by A^T, is defined to be the $n \times m$ matrix that results by interchanging the rows and columns of A; that is, the first column of A^T is the first row of A, the second column of A^T is the second row of A, and so forth.

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



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

Transpose Matrix Properties

THEOREM 1.4.8 *If the sizes of the matrices are such that the stated operations can be performed, then:*

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

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

(*c*) $(A - B)^T = A^T - B^T$

(*d*) $(kA)^T = kA^T$

(*e*) $(AB)^T = B^T A^T$

The transpose of a product of any number of matrices is the product of the transposes in the reverse order.

Gaussian Elimination and Gauss-Jordan Elimination

$m \times n$ matrix:

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \cdots & a_{2n} \\ a_{31} & a_{32} & a_{33} & \cdots & a_{3n} \\ & \vdots & & & \\ a_{m1} & a_{m2} & a_{m3} & \cdots & a_{mn} \end{bmatrix} \quad m \text{ rows}$$

$$n \text{ columns}$$

Notes:

(1) Every **entry** a_{ij} in a matrix is a number.

(2) A matrix with <u>m rows</u> and <u>n columns</u> is said to be of **size** $m \times n$.

(3) If $m = n$, then the matrix is called **square of order n**.

(4) For a square matrix, the entries a_{11}, a_{22}, …, a_{nn} are called **the main diagonal entries**.

Elementary row operation:

(1) Interchange two rows. $r_{ij} : R_i \leftrightarrow R_j$

(2) Multiply a row by a nonzero constant. $r_i^{(k)} : (k)R_i \rightarrow R_i$

(3) Add a multiple of a row to another row: $r_{ij}^{(k)} : (k)R_i + R_j \rightarrow R_j$

Ex : (Elementary row operation)

$$\begin{bmatrix} 0 & 1 & 3 & 4 \\ -1 & 2 & 0 & 3 \\ 2 & -3 & 4 & 1 \end{bmatrix} \xrightarrow{r_{12}} \begin{bmatrix} -1 & 2 & 0 & 3 \\ 0 & 1 & 3 & 4 \\ 2 & -3 & 4 & 1 \end{bmatrix}$$

$$\begin{bmatrix} 2 & -4 & 6 & -2 \\ 1 & 3 & -3 & 0 \\ 5 & -2 & 1 & 2 \end{bmatrix} \xrightarrow{r_1^{(\frac{1}{2})}} \begin{bmatrix} 1 & -2 & 3 & -1 \\ 1 & 3 & -3 & 0 \\ 5 & -2 & 1 & 2 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 2 & -4 & 3 \\ 0 & 3 & -2 & -1 \\ 2 & 1 & 5 & -2 \end{bmatrix} \xrightarrow{r_{13}^{(-2)}} \begin{bmatrix} 1 & 2 & -4 & 3 \\ 0 & 3 & -2 & -1 \\ 0 & -3 & 13 & -8 \end{bmatrix}$$

Ex : Using elementary row operations to solve a system

Linear System	Associated Augemented Matrix	Elementary Row Operation
$\begin{aligned} x - 2y + 3z &= 9 \\ -x + 3y &= -4 \\ 2x - 5y + 5z &= 17 \end{aligned}$	$\begin{bmatrix} 1 & -2 & 3 & 9 \\ -1 & 3 & 0 & -4 \\ 2 & -5 & 5 & 17 \end{bmatrix}$	
	$\begin{bmatrix} 1 & -2 & 3 & 9 \\ 0 & 1 & 3 & 5 \\ 2 & -5 & 5 & 17 \end{bmatrix}$	$r_{12}^{(1)} : (1)R_1 + R_2 \to R_2$
	$\begin{bmatrix} 1 & -2 & 3 & 9 \\ 0 & 1 & 3 & 5 \\ 0 & -1 & -1 & -1 \end{bmatrix}$	$r_{13}^{(-2)} : (-2)R_1 + R_3 \to R_3$

Linear System	Associated Augemented Matrix	Elementary Row Operation
	$\begin{bmatrix} 1 & -2 & 3 & 9 \\ 0 & 1 & 3 & 5 \\ 0 & 0 & 2 & 4 \end{bmatrix}$	$r_{23}^{(1)} : (1)R_2 + R_3 \to R_3$
$\begin{aligned} x - 2y + 3z &= 9 \\ y + 3z &= 5 \\ z &= 2 \end{aligned}$	$\begin{bmatrix} 1 & -2 & 3 & 9 \\ 0 & 1 & 3 & 5 \\ 0 & 0 & 1 & 2 \end{bmatrix}$	$r_3^{(\frac{1}{2})} : (\frac{1}{2})R_3 \to R_3$

$$\longrightarrow \quad \begin{aligned} x &= 1 \\ y &= -1 \\ z &= 2 \end{aligned}$$

Row-echelon form:

Reduced row-echelon form:

(1) All row consisting entirely of zeros occur at the bottom of the matrix.

(2) For each row that does not consist entirely of zeros, the first nonzero entry is 1 (called **a leading 1**).

(3) For two successive (nonzero) rows, the leading 1 in the higher row is farther to the left than the leading 1 in the lower row.

(4) Every column that has a leading 1 has zeros in every position above and below its leading 1.

Ex : (Row-echelon form or reduced row-echelon form)

$$\begin{bmatrix} 1 & 2 & -1 & 4 \\ 0 & 1 & 0 & 3 \\ 0 & 0 & 1 & -2 \end{bmatrix} \text{(row-echelon form)} \qquad \begin{bmatrix} 0 & 1 & 0 & 5 \\ 0 & 0 & 1 & 3 \\ 0 & 0 & 0 & 0 \end{bmatrix} \text{(reduced row-echelon form)}$$

$$\begin{bmatrix} 1 & -5 & 2 & -1 & 3 \\ 0 & 0 & 1 & 3 & -2 \\ 0 & 0 & 0 & 1 & 4 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} \text{(row-echelon form)} \qquad \begin{bmatrix} 1 & 0 & 0 & -1 \\ 0 & 1 & 0 & 2 \\ 0 & 0 & 1 & 3 \\ 0 & 0 & 0 & 0 \end{bmatrix} \text{(reduced row-echelon form)}$$

Gaussian elimination:

The procedure for reducing a matrix to a row-echelon form.

Gauss-Jordan elimination:

The procedure for reducing a matrix to a reduced row-echelon form.

Notes:

(1) Every matrix has an unique reduced row echelon form

(2) A row-echelon form of a given matrix is not unique. (Different sequences of row operations can produce different row-echelon forms.)

Ex: (Procedure of Gaussian elimination and Gauss-Jordan elimination)

$$\begin{bmatrix} 0 & 0 & -2 & 0 & 8 & 12 \\ 2 & 8 & -6 & 4 & 12 & 28 \\ 2 & 4 & -5 & 6 & -5 & 4 \end{bmatrix} \xrightarrow{r_{12}} \begin{bmatrix} 2 & 8 & -6 & 4 & 12 & 28 \\ 0 & 0 & -2 & 0 & 8 & 12 \\ 2 & 4 & -5 & 6 & -5 & 4 \end{bmatrix}$$

The first nonzero column; Produce leading 1

$$\xrightarrow{r_1^{(\frac{1}{2})}} \begin{bmatrix} 1 & 4 & -3 & 2 & 6 & 14 \\ 0 & 0 & -2 & 0 & 8 & 12 \\ 2 & 4 & -5 & 6 & -5 & 4 \end{bmatrix} \xrightarrow{r_{13}^{(-2)}} \begin{bmatrix} 1 & 4 & -3 & 2 & 6 & 14 \\ 0 & 0 & -2 & 0 & 8 & 12 \\ 0 & 0 & 5 & 0 & -17 & -24 \end{bmatrix}$$

leading 1; Zeros elements below leading 1; Produce leading 1; The first nonzero column; Submatrix

$$\xrightarrow{r_2^{(-\frac{1}{2})}} \begin{bmatrix} 1 & 4 & -3 & 2 & 6 & 14 \\ 0 & 0 & 1 & 0 & -4 & -6 \\ 0 & 0 & 5 & 0 & -17 & -24 \end{bmatrix} \xrightarrow{r_{23}^{(-5)}} \begin{bmatrix} 1 & 4 & -3 & 2 & 6 & 14 \\ 0 & 0 & 1 & 0 & -4 & -6 \\ 0 & 0 & 0 & 0 & 3 & 6 \end{bmatrix}$$

leading 1

Zeros elements below leading 1

Submatrix

Produce leading 1

Zeros elsewhere

$$\xrightarrow{r_3^{(\frac{1}{3})}} \begin{bmatrix} 1 & 4 & -3 & 2 & 6 & 14 \\ 0 & 0 & 1 & 0 & -4 & -6 \\ 0 & 0 & 0 & 0 & 1 & 2 \end{bmatrix} \xrightarrow{r_{31}^{(-6)}} \begin{bmatrix} 1 & 4 & -3 & 2 & 0 & 2 \\ 0 & 0 & 1 & 0 & -4 & -6 \\ 0 & 0 & 0 & 0 & 1 & 2 \end{bmatrix}$$

leading 1

(row - echelon form) (row - echelon form)

$$\xrightarrow{r_{32}^{(4)}} \begin{bmatrix} 1 & 4 & -3 & 2 & 0 & 2 \\ 0 & 0 & 1 & 0 & 0 & 2 \\ 0 & 0 & 0 & 0 & 1 & 2 \end{bmatrix} \xrightarrow{r_{21}^{(3)}} \begin{bmatrix} 1 & 4 & 0 & 2 & 0 & 8 \\ 0 & 0 & 1 & 0 & 0 & 2 \\ 0 & 0 & 0 & 0 & 1 & 2 \end{bmatrix}$$

(row - echelon form) (reduced row - echelon form)

Ex : Solve a system by Gauss-Jordan elimination method (only one solution)

$$\begin{aligned} x - 2y + 3z &= 9 \\ -x + 3y \quad\quad &= -4 \\ 2x - 5y + 5z &= 17 \end{aligned}$$

Sol:

augmented matrix

$$\begin{bmatrix} 1 & -2 & 3 & 9 \\ -1 & 3 & 0 & -4 \\ 2 & -5 & 5 & 17 \end{bmatrix} \xrightarrow{r_{12}^{(1)},\, r_{13}^{(-2)}} \begin{bmatrix} 1 & -2 & 3 & 9 \\ 0 & 1 & 3 & 5 \\ 0 & -1 & -1 & -1 \end{bmatrix} \xrightarrow{r_{23}^{(1)}} \begin{bmatrix} 1 & -2 & 3 & 9 \\ 0 & 1 & 3 & 5 \\ 0 & 0 & 2 & 4 \end{bmatrix}$$

$$\xrightarrow{r_3^{(\frac{1}{2})}} \begin{bmatrix} 1 & -2 & 3 & 9 \\ 0 & 1 & 3 & 5 \\ 0 & 0 & 1 & 2 \end{bmatrix} \xrightarrow{r_{21}^{(2)},\, r_{32}^{(-3)},\, r_{31}^{(-9)}} \begin{bmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 1 & 2 \end{bmatrix} \longrightarrow \begin{aligned} x &= 1 \\ y &= -1 \\ z &= 2 \end{aligned}$$

(row - echelon form) (reduced row - echelon form)

Ex :Solve a system by Gauss-Jordan elimination method (infinitely many solutions)

$$\begin{aligned} 2x_1 + 4x_1 - 2x_3 &= 0 \\ 3x_1 + 5x_2 \quad\quad &= 1 \end{aligned}$$

Sol: augmented matrix

$$\begin{bmatrix} 2 & 4 & -2 & 0 \\ 3 & 5 & 0 & 1 \end{bmatrix} \xrightarrow{r_1^{(\frac{1}{2})},\, r_{12}^{(-3)},\, r_2^{(-1)},\, r_{21}^{(-2)}} \begin{bmatrix} 1 & 0 & 5 & 2 \\ 0 & 1 & -3 & -1 \end{bmatrix}$$ (reduced row - echelon form)

the corresponding system of equations is

$$\begin{aligned} x_1 \quad + 5x_3 &= 2 \\ x_2 - 3x_3 &= -1 \end{aligned}$$

leading variable : x_1, x_2

free variable : x_3

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

Let $x_3 = t$

$$\begin{aligned} x_1 &= 2 - 5t, \\ x_2 &= -1 + 3t, \qquad t \in R \\ x_3 &= t, \end{aligned}$$

So this system has infinitely many solutions.

Matrices and Matrix Operations

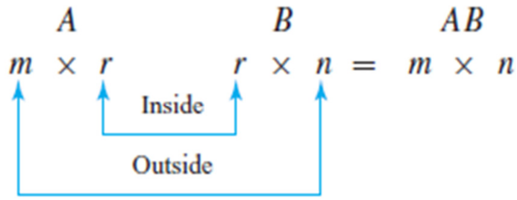
Definition 1 A matrix is a rectangular array of numbers. The numbers in the array are called the entries of the matrix.
The size of a matrix M is written in terms of the number of its rows x the number of its columns. A 2x3 matrix has 2 rows and 3 columns

Arithmetic of Matrices

A + B: add the corresponding entries of A and B
A – B: subtract the corresponding entries of B from those of A
Matrices A and B must be of the same size to be added or subtracted
cA (scalar multiplication): multiply each entry of A by the constant c

Multiplication of Matrices

DEFINITION 5 If A is an $m \times r$ matrix and B is an $r \times n$ matrix, then the ***product*** AB is the $m \times n$ matrix whose entries are determined as follows: To find the entry in row i and column j of AB, single out row i from the matrix A and column j from the matrix B. Multiply the corresponding entries from the row and column together, and then add up the resulting products.



EXAMPLE 5 Multiplying Matrices

Consider the matrices

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

Since A is a 2×3 matrix and B is a 3×4 matrix, the product AB is a 2×4 matrix. To determine, for example, the entry in row 2 and column 3 of AB, we single out row 2 from A and column 3 from B. Then, as illustrated below, we multiply corresponding entries together and add up these products.

$$\begin{bmatrix}1 & 2 & 4\\ 2 & 6 & 0\end{bmatrix}\begin{bmatrix}4 & 1 & 4 & 3\\ 0 & -1 & 3 & 1\\ 2 & 7 & 5 & 2\end{bmatrix}=\begin{bmatrix}\square & \square & \square & \square\\ \square & \square & \boxed{26} & \square\end{bmatrix}$$

$$(2\cdot 4)+(6\cdot 3)+(0\cdot 5)=26$$

The entry in row 1 and column 4 of AB is computed as follows:

$$\begin{bmatrix}1 & 2 & 4\\ 2 & 6 & 0\end{bmatrix}\begin{bmatrix}4 & 1 & 4 & 3\\ 0 & -1 & 3 & 1\\ 2 & 7 & 5 & 2\end{bmatrix}=\begin{bmatrix}\square & \square & \square & \boxed{13}\\ \square & \square & \square & \square\end{bmatrix}$$

$$(1\cdot 3)+(2\cdot 1)+(4\cdot 2)=13$$

The computations for the remaining entries are

$$(1.4)+(2.0)+(4.2)=12$$
$$(1.1)-(2.1)+(4.7)=27$$
$$(1.4)+(2.3)+(4.5)=30$$
$$(2.4)+(6.0)+(0.2)=8$$
$$(2.1)-(6.1)+(0.7)=-4$$
$$(2.3)+(6.1)+(0.2)=12$$

$$AB=\begin{bmatrix}12 & 27 & 30 & 13\\ 8 & -4 & 26 & 12\end{bmatrix}$$

WEEK 3

2.1 Determinants by Cofactor Expansion

DEFINITION 1

If *A* is a square matrix, then the ***minor of entry*** ***aij*** is denoted by **Mij** and is defined to be the determinant of the submatrix that remains after the ***i*th** row and ***j*th** column are deleted from *A*. The number $(-1)^{i+j} M_{ij}$ is denoted by **Cij** and is called the ***cofactor of entry*** ***aij.*** .

EXAMPLE 1 Finding Minors and Cofactors

Let

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

The minor of entry a_{11} is

$$M_{11} = \begin{vmatrix} 3 & 1 & -4 \\ 2 & 5 & 6 \\ 1 & 4 & 8 \end{vmatrix} = \begin{vmatrix} 5 & 6 \\ 4 & 8 \end{vmatrix} = 16$$

The cofactor of a_{11} is

$$C_{11} = (-1)^{1+1} M_{11} = M_{11} = 16$$

Similarly, the minor of entry a_{32} is

$$M_{32} = \begin{vmatrix} 3 & 1 & -4 \\ 2 & 5 & 6 \\ 1 & 4 & 8 \end{vmatrix} = \begin{vmatrix} 3 & -4 \\ 2 & 6 \end{vmatrix} = 26$$

The cofactor of a_{32} is

$$C_{32} = (-1)^{3+2} M_{32} = -M_{32} = -26$$

DEFINITION 2 If A is an $n \times n$ matrix, then the number obtained by multiplying the entries in any row or column of A by the corresponding cofactors and adding the resulting products is called the ***determinant of** A*, and the sums themselves are called ***cofactor expansions of** A*. That is,

$$\det(A) = a_{1j}C_{1j} + a_{2j}C_{2j} + \cdots + a_{nj}C_{nj} \quad (5)$$

[cofactor expansion along the *j*th column]

and

$$\det(A) = a_{i1}C_{i1} + a_{i2}C_{i2} + \cdots + a_{in}C_{in} \quad (6)$$

[cofactor expansion along the *i*th row]

Determinant of 2 x 2 Matrix

The quantity $ad - bc$ in Theorem 1.4.5 is called the ***determinant*** of the 2×2 matrix A and is denoted by

$$\det(A) = ad - bc$$

or alternatively by

$$\begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc$$

EXAMPLE 3 Cofactor Expansion Along the First Row

Find the determinant of the matrix

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

by cofactor expansion along the first row.

Solution

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

$$= 3(-4) - (1)(-11) + 0 = -1$$

EXAMPLE 4 Cofactor Expansion Along the First Column

Let A be the matrix in Example 3, and evaluate $\det(A)$ by cofactor expansion along the first column of A.

Solution

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

$$= 3(-4) - (-2)(-2) + 5(3) = -1$$

This agrees with the result obtained in Example 3.

A technique for determinants of 2x2 and 3x3 matrices *only*

EXAMPLE 7 A Technique for Evaluating 2 × 2 and 3 × 3 Determinants

$$\begin{vmatrix} 3 & 1 \\ 4 & -2 \end{vmatrix} = \begin{vmatrix} 3 & 1 \\ 4 & -2 \end{vmatrix} = (3)(-2) - (1)(4) = -10$$

$$\begin{vmatrix} 1 & 2 & 3 \\ -4 & 5 & 6 \\ 7 & -8 & 9 \end{vmatrix} = \begin{vmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & -8 & 9 \end{vmatrix} \begin{matrix} 1 & 2 \\ -4 & 5 \\ 7 & -8 \end{matrix}$$

$$= [45 + 84 + 96] - [105 - 48 - 72] = 240$$

2.2 Evaluating Determinants by Row Reduction

THEOREM 2.2.1

Let A be a square matrix. If A has a row of zeros or a column of zeros, then $\det(A) = 0$.

THEOREM 2.2.2

Let A be a square matrix. Then $\det(A) = \det(A^T)$.

THEOREM 2.2.3

Let A be an $n \times n$ matrix.

(a) If B is the matrix that results when a single row or single column of A is multiplied by a scalar k, then $\det(B) = k\det(A)$.

(b) If B is the matrix that results when two rows or two columns of A are interchanged, then $\det(B) = -\det(A)$.

(c) If B is the matrix that results when a multiple of one row of A is added to another row or when a multiple of one column is added to another column, then $\det(B) = \det(A)$.

Table 1

Relationship	Operation
$\begin{vmatrix} ka_{11} & ka_{12} & ka_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = k \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}$ $\det(B) = k\det(A)$	The first row of A is multiplied by k.
$\begin{vmatrix} a_{21} & a_{22} & a_{23} \\ a_{11} & a_{12} & a_{13} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = - \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}$ $\det(B) = -\det(A)$	The first and second rows of A are interchanged.
$\begin{vmatrix} a_{11}+ka_{21} & a_{12}+ka_{22} & a_{13}+ka_{23} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}$ $\det(B) = \det(A)$	A multiple of the second row of A is added to the first row.

EXAMPLE 3 Using Row Reduction to Evaluate a Determinant

Evaluate $\det(A)$ where

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

$$\det(A)=\begin{vmatrix}0 & 1 & 5\\ 3 & -6 & 9\\ 2 & 6 & 1\end{vmatrix}=-\begin{vmatrix}3 & -6 & 9\\ 0 & 1 & 5\\ 2 & 6 & 1\end{vmatrix}$$ ← The first and second rows of A where interchanged .

$$=-3\begin{vmatrix}1 & -2 & 3\\ 0 & 1 & 5\\ 2 & 6 & 1\end{vmatrix}$$ ← A common factor of 3 from the first row was taken through the determinant sign .

$$=-3\begin{vmatrix}1 & -2 & 3\\ 0 & 1 & 5\\ 0 & 10 & -5\end{vmatrix}$$ ← -2 times the first row was added to the third row .

$$=-3\begin{vmatrix}1 & -2 & 3\\ 0 & 1 & 5\\ 0 & 0 & -55\end{vmatrix}$$ ← -10 times the second row was added to the third row .

$$=(-3)(-55)\begin{vmatrix}1 & -2 & 3\\ 0 & 1 & 5\\ 0 & 0 & 1\end{vmatrix}$$ ← A common factor of -55 from the last row was taken through the determinant sign .

$$=(-3)(-55)(1)=165$$

EXAMPLE 1 $\det(A+B)\neq\det(A)+\det(B)$

Consider

$$A=\begin{bmatrix}1 & 2\\ 2 & 5\end{bmatrix},\quad B=\begin{bmatrix}3 & 1\\ 1 & 3\end{bmatrix},\quad A+B=\begin{bmatrix}4 & 3\\ 3 & 8\end{bmatrix}$$

We have $\det(A)=1$, $\det(B)=8$, and $\det(A+B)=23$; thus

$$\det(A+B)\neq\det(A)+\det(B)$$

THEOREM 2.3.5

If A is invertible, then

$$\det(A^{-1})=\frac{1}{\det(A)}$$

THEOREM 2.3.4

If A and B are square matrices of the same size, then

$$\det(AB)=\det(A)\det(B)$$

EXAMPLE 4 Verifying That det(AB) = det(A), det(B)

Consider the matrices

$$A = \begin{bmatrix} 3 & 1 \\ 2 & 1 \end{bmatrix}, \quad B = \begin{bmatrix} -1 & 3 \\ 5 & 8 \end{bmatrix}, \quad AB = \begin{bmatrix} 2 & 17 \\ 3 & 14 \end{bmatrix}$$

We leave it for you to verify that

$$\det(A) = 1, \quad \det(B) = -23, \quad \text{and} \quad \det(AB) = -23$$

Thus $\det(AB) = \det(A)\det(B)$, as guaranteed by Theorem 2.3.4.

2.3 Properties of Determinants; Cramer's Rule

THEOREM 2.3.7 Cramer's Rule

If $A\mathbf{x} = \mathbf{b}$ is a system of n linear equations in n unknowns such that $\det(A) \neq 0$, then the system has a unique solution. This solution is

$$x_1 = \frac{\det(A_1)}{\det(A)}, \quad x_2 = \frac{\det(A_2)}{\det(A)}, \ldots, \quad x_n = \frac{\det(A_n)}{\det(A)}$$

where A_j is the matrix obtained by replacing the entries in the jth column of A by the entries in the matrix

$$\mathbf{b} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}$$

EXAMPLE 8 Using Cramer's Rule to Solve a Linear System

Use Cramer's rule to solve

$$\begin{aligned} x_1 + \quad\quad + 2x_3 &= 6 \\ -3x_1 + 4x_2 + 6x_3 &= 30 \\ -x_1 - 2x_2 + 3x_3 &= 8 \end{aligned}$$

Solution

$$A = \begin{bmatrix} 1 & 0 & 2 \\ -3 & 4 & 6 \\ -1 & -2 & 3 \end{bmatrix}, \quad A_1 = \begin{bmatrix} 6 & 0 & 2 \\ 30 & 4 & 6 \\ 8 & -2 & 3 \end{bmatrix},$$

$$A_2 = \begin{bmatrix} 1 & 6 & 2 \\ -3 & 30 & 6 \\ -1 & 8 & 3 \end{bmatrix}, \quad A_3 = \begin{bmatrix} 1 & 0 & 6 \\ -3 & 4 & 30 \\ -1 & -2 & 8 \end{bmatrix}$$

Therefore,

$$x_1 = \frac{\det(A_1)}{\det(A)} = \frac{-40}{44} = \frac{-10}{11}, \quad x_2 = \frac{\det(A_2)}{\det(A)} = \frac{72}{44} = \frac{18}{11},$$

$$x_3 = \frac{\det(A_3)}{\det(A)} = \frac{152}{44} = \frac{38}{11}$$

WEEK 4

1.4 Inverses; Algebraic Properties of Matrices.

THEOREM 1.4.1 Properties of Matrix Arithmetic

Assuming that the sizes of the matrices are such that the indicated operations can be performed, the following rules of matrix arithmetic are valid.

(a) $A + B = B + A$ **(Commutative law for addition)**

(b) $A + (B + C) = (A + B) + C$ **(Associative law for addition)**

(c) $A(BC) = (AB)C$ **(Associative law for multiplication)**

(d) $A(B + C) = AB + AC$ **(Left distributive law)**

(e) $(B + C)A = BA + CA$ **(Right distributive law)**

(f) $A(B - C) = AB - AC$

(g) $(B - C)A = BA - CA$

(h) $a(B + C) = aB + aC$

(i) $a(B - C) = aB - aC$

(j) $(a + b)C = aC + bC$

(k) $(a - b)C = aC - bC$

(l) $a(bC) = (ab)C$

(m) $a(BC) = (aB)C = B(aC)$

Zero Matrices

$$\begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}, \quad \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \quad \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \quad [0]$$

THEOREM 1.4.2 Properties of Zero Matrices

If c is a scalar, and if the sizes of the matrices are such that the operations can be perfomed, then:

(a) $A + 0 = 0 + A = A$

(b) $A - 0 = A$

(c) $A - A = A + (-A) = 0$

(d) $0A = 0$

(e) If $cA = 0$, then $c = 0$ or $A = 0$.

The identity matrix

A square matrix with 1's on the main diagonal and zeros elsewhere is called an ***identity matrix***. Some examples are

$$[1], \quad \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

Inverse Matrices

DEFINITION 1 If A is a square matrix, and if a matrix B of the same size can be found such that $AB = BA = I$, then A is said to be ***invertible*** (or ***nonsingular***) and B is called an ***inverse*** of A. If no such matrix B can be found, then A is said to be ***singular***.

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

The Inverse of a 2x2 matrix

THEOREM 1.4.5 *The matrix*

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

is invertible if and only if $ad - bc \neq 0$, *in which case the inverse is given by the formula*

$$A^{-1} = \frac{1}{ad - bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix} \tag{2}$$

The Determinant

The quantity $ad - bc$ in Theorem 1.4.5 is called the ***determinant*** of the 2×2 matrix A and is denoted by

$$\det(A) = ad - bc$$

or alternatively by

$$\begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc$$

More on Invertible Matrices

THEOREM 1.4.6 *If A and B are invertible matrices with the same size, then AB is invertible and*

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

THEOREM 1.4.7 *If A is invertible and n is a nonnegative integer, then:*

(a) A^{-1} *is invertible and* $(A^{-1})^{-1} = A$.

(b) A^n *is invertible and* $(A^n)^{-1} = A^{-n} = (A^{-1})^n$.

(c) kA *is invertible for any nonzero scalar* k, *and* $(kA)^{-1} = k^{-1}A^{-1}$.

Transpose Matrix Properties

THEOREM 1.4.8 *If the sizes of the matrices are such that the stated operations can be performed, then:*

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

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

(c) $(A - B)^T = A^T - B^T$

(d) $(kA)^T = kA^T$

(e) $(AB)^T = B^TA^T$

THEOREM 1.4.9 *If A is an invertible matrix, then* A^T *is also invertible and*

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

1.5 Elementary Matrices and methods for finding A-1

Using Row Operations to find A-1

Begin with:

$$\left[\begin{array}{ccc|ccc} 1 & 2 & 3 & 1 & 0 & 0 \\ 2 & 5 & 3 & 0 & 1 & 0 \\ 1 & 0 & 8 & 0 & 0 & 1 \end{array}\right]$$

Use successive row operations to produce:

$$\left[\begin{array}{ccc|ccc} 1 & 2 & 3 & 1 & 0 & 0 \\ 2 & 5 & 3 & 0 & 1 & 0 \\ 1 & 0 & 8 & 0 & 0 & 1 \end{array}\right]$$

$$\left[\begin{array}{ccc|ccc} 1 & 2 & 3 & 1 & 0 & 0 \\ 0 & 1 & -3 & -2 & 1 & 0 \\ 0 & -2 & 5 & -1 & 0 & 1 \end{array}\right] \leftarrow$$ We added −2 times the first row to the second and −1 times the first row to the third.

$$\left[\begin{array}{ccc|ccc} 1 & 2 & 3 & 1 & 0 & 0 \\ 0 & 1 & -3 & -2 & 1 & 0 \\ 0 & 0 & -1 & -5 & 2 & 1 \end{array}\right] \leftarrow$$ We added 2 times the second row to the third.

$$\left[\begin{array}{ccc|ccc} 1 & 2 & 3 & 1 & 0 & 0 \\ 0 & 1 & -3 & -2 & 1 & 0 \\ 0 & 0 & 1 & 5 & -2 & -1 \end{array}\right] \leftarrow$$ We multiplied the third row by −1.

$$\left[\begin{array}{ccc|ccc} 1 & 2 & 0 & -14 & 6 & 3 \\ 0 & 1 & 0 & 13 & -5 & -3 \\ 0 & 0 & 1 & 5 & -2 & -1 \end{array}\right] \leftarrow$$ We added 3 times the third row to the second and −3 times the third row to the first.

$$\left[\begin{array}{ccc|ccc} 1 & 0 & 0 & -40 & 16 & 9 \\ 0 & 1 & 0 & 13 & -5 & -3 \\ 0 & 0 & 1 & 5 & -2 & -1 \end{array}\right] \leftarrow$$ We added −2 times the second row to the first.

$$\left[\begin{array}{ccc|ccc} 1 & 0 & 0 & -40 & 16 & 9 \\ 0 & 1 & 0 & 13 & -5 & -3 \\ 0 & 0 & 1 & 5 & -2 & -1 \end{array}\right]$$

Thus,

$$A^{-1} = \begin{bmatrix} -40 & 16 & 9 \\ 13 & -5 & -3 \\ 5 & -2 & -1 \end{bmatrix}$$

Not Invertible Matrix

EXAMPLE 5 Showing That a Matrix Is Not Invertible

Consider the matrix

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

Section 1.6 Linear Systems and Invertible Matrices

THEOREM 1.6.2

If A is an invertible $n \times n$ matrix, then for each $n \times 1$ *matrix* $\mathbf{b}$, the system of equations $A\mathbf{x} = \mathbf{b}$ has exactly one solution,

namely, $\mathbf{x} = A^{-1}\mathbf{b}$

Example

EXAMPLE 1 Solution of a Linear System Using A^{-1}

Consider the system of linear equations

$$\begin{aligned} x_1 + 2x_2 + 3x_3 &= 5 \\ 2x_1 + 5x_2 + 3x_3 &= 3 \\ x_1 \qquad\quad + 8x_3 &= 17 \end{aligned}$$

In matrix form this system can be written as $A\mathbf{x} = \mathbf{b}$, where

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 5 & 3 \\ 1 & 0 & 8 \end{bmatrix}, \quad \mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} 5 \\ 3 \\ 17 \end{bmatrix}$$

In Example 4 of the preceding section, we showed that A is invertible and

$$A^{-1} = \begin{bmatrix} -40 & 16 & 9 \\ 13 & -5 & -3 \\ 5 & -2 & -1 \end{bmatrix}$$

By Theorem 1.6.2, the solution of the system is

$$\mathbf{x} = A^{-1}\mathbf{b} = \begin{bmatrix} -40 & 16 & 9 \\ 13 & -5 & -3 \\ 5 & -2 & -1 \end{bmatrix} \begin{bmatrix} 5 \\ 3 \\ 17 \end{bmatrix} = \begin{bmatrix} 1 \\ -1 \\ 2 \end{bmatrix}$$

or $x_1 = 1$, $x_2 = -1$, $x_3 = 2$. ◀

Section 1.7 Diagonal, Triangular and Symmetric Matrices

Diagonal

A general $n \times n$ diagonal matrix D can be written as

$$D = \begin{bmatrix} d_1 & 0 & \cdots & 0 \\ 0 & d_2 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & d_n \end{bmatrix}$$

Triangular

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ 0 & a_{22} & a_{23} & a_{24} \\ 0 & 0 & a_{33} & a_{34} \\ 0 & 0 & 0 & a_{44} \end{bmatrix} \qquad \begin{bmatrix} a_{11} & 0 & 0 & 0 \\ a_{21} & a_{22} & 0 & 0 \\ a_{31} & a_{32} & a_{33} & 0 \\ a_{41} & a_{42} & a_{43} & a_{44} \end{bmatrix}$$

A general 4 × 4 upper triangular matrix

A general 4 × 4 lower triangular matrix

Symmetric

DEFINITION 1 A square matrix A is said to be *symmetric* if $A = A^T$.

WEEK 5

Euclidean Vector Spaces

Material in this chapter

3.1 Vectors in 2-Space, 3-Space, and n-Space

3.2 Norm, Dot Product, and Distance in Rn

3.3 Orthogonality

3.4 The Geometry of Linear Systems

3.5 Cross Product

Vectors in Coordinate Systems

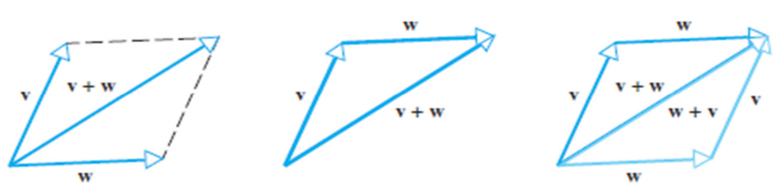
Up until now we have discussed vectors without reference to a coordinate system. However, as we will soon see, computations with vectors are much simpler to perform if a coordinate system is present to work with.

The component forms of the zero vector are $\mathbf{0} = (0, 0)$ in 2-space and $\mathbf{0} = (0, 0, 0)$ in 3-space.

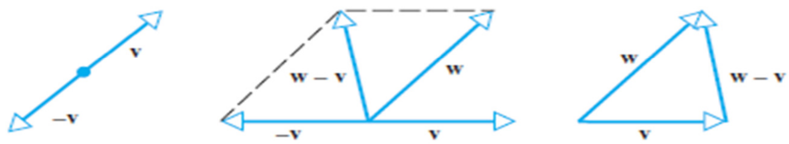
If a vector $\mathbf{v}$ in 2-space or 3-space is positioned with its initial point at the origin of a rectangular coordinate system, then the vector is completely determined by the coordinates of its terminal point (Figure 3.1.10). We call these coordinates the ***components*** of $\mathbf{v}$ relative to the coordinate system. We will write $\mathbf{v} = (v_1, v_2)$ to denote a vector $\mathbf{v}$ in 2-space with components (v_1, v_2), and $\mathbf{v} = (v_1, v_2, v_3)$ to denote a vector $\mathbf{v}$ in 3-space with components (v_1, v_2, v_3).

Vectors

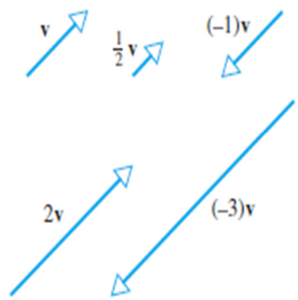
Addition of vectors by the parallelogram or triangle rules



Subtraction:



Scalar Multiplication:



Properties of Vectors

THEOREM 3.1.1 *If* $\mathbf{u}$, $\mathbf{v}$, *and* $\mathbf{w}$ *are vectors in* R^n, *and if* k *and* m *are scalars, then:*

(*a*) $\mathbf{u}+\mathbf{v}=\mathbf{v}+\mathbf{u}$

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

(*c*) $\mathbf{u}+\mathbf{0}=\mathbf{0}+\mathbf{u}=\mathbf{u}$

(*d*) $\mathbf{u}+(-\mathbf{u})=\mathbf{0}$

(*e*) $k(\mathbf{u}+\mathbf{v})=k\mathbf{u}+k\mathbf{v}$

(*f*) $(k+m)\mathbf{u}=k\mathbf{u}+m\mathbf{u}$

(*g*) $k(m\mathbf{u})=(km)\mathbf{u}$

(*h*) $1\mathbf{u}=\mathbf{u}$

Norm, Dot Product, and Distance in Rn

Norm:

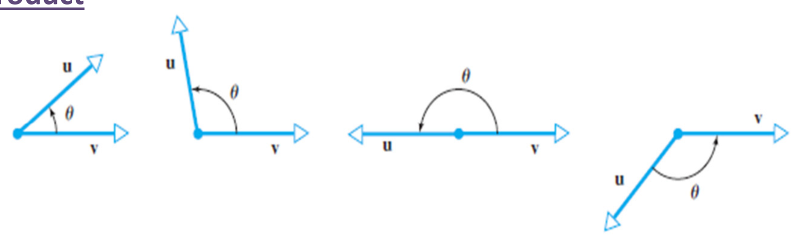
DEFINITION 1 If $\mathbf{v} = (v_1, v_2, \ldots, v_n)$ is a vector in R^n, then the ***norm*** of $\mathbf{v}$ (also called the ***length*** of $\mathbf{v}$ or the ***magnitude*** of $\mathbf{v}$) is denoted by $\|\mathbf{v}\|$, and is defined by the formula

$$\|\mathbf{v}\| = \sqrt{v_1^2 + v_2^2 + v_3^2 + \cdots + v_n^2} \tag{3}$$

Unit Vectors:

$$\mathbf{u} = \frac{1}{\|\mathbf{v}\|}\mathbf{v}$$

The Dot Product



The angle θ between $\mathbf{u}$ and $\mathbf{v}$ satisfies $0 \le \theta \le \pi$.

DEFINITION 3 If $\mathbf{u}$ and $\mathbf{v}$ are nonzero vectors in R^2 or R^3, and if θ is the angle between $\mathbf{u}$ and $\mathbf{v}$, then the ***dot product*** (also called the ***Euclidean inner product***) of $\mathbf{u}$ and $\mathbf{v}$ is denoted by $\mathbf{u} \cdot \mathbf{v}$ and is defined as

$$\mathbf{u} \cdot \mathbf{v} = \|\mathbf{u}\|\|\mathbf{v}\| \cos\theta \tag{12}$$

If $\mathbf{u} = \mathbf{0}$ or $\mathbf{v} = \mathbf{0}$, then we define $\mathbf{u} \cdot \mathbf{v}$ to be 0.

The sign of the dot product reveals information about the angle θ that we can obtain by rewriting Formula (12) as

$$\cos\theta = \frac{\mathbf{u} \cdot \mathbf{v}}{\|\mathbf{u}\|\|\mathbf{v}\|} \tag{13}$$

DEFINITION 4 If $\mathbf{u} = (u_1, u_2, \ldots, u_n)$ and $\mathbf{v} = (v_1, v_2, \ldots, v_n)$ are vectors in R^n, then the ***dot product*** (also called the ***Euclidean inner product***) of $\mathbf{u}$ and $\mathbf{v}$ is denoted by $\mathbf{u} \cdot \mathbf{v}$ and is defined by

$$\mathbf{u} \cdot \mathbf{v} = u_1 v_1 + u_2 v_2 + \cdots + u_n v_n \tag{17}$$

Properties of the Dot Product

THEOREM 3.2.2 *If $\mathbf{u}$, $\mathbf{v}$, and $\mathbf{w}$ are vectors in R^n, and if k is a scalar, then:*

(*a*) $\mathbf{u} \cdot \mathbf{v} = \mathbf{v} \cdot \mathbf{u}$ **[Symmetry property]**

(*b*) $\mathbf{u} \cdot (\mathbf{v} + \mathbf{w}) = \mathbf{u} \cdot \mathbf{v} + \mathbf{u} \cdot \mathbf{w}$ **[Distributive property]**

(*c*) $k(\mathbf{u} \cdot \mathbf{v}) = (k\mathbf{u}) \cdot \mathbf{v}$ **[Homogeneity property]**

(*d*) $\mathbf{v} \cdot \mathbf{v} \geq 0$ *and* $\mathbf{v} \cdot \mathbf{v} = 0$ *if and only if* $\mathbf{v} = \mathbf{0}$ **[Positivity property]**

THEOREM 3.2.3 *If $\mathbf{u}$, $\mathbf{v}$, and $\mathbf{w}$ are vectors in R^n, and if k is a scalar, then:*

(*a*) $\mathbf{0} \cdot \mathbf{v} = \mathbf{v} \cdot \mathbf{0} = 0$

(*b*) $(\mathbf{u} + \mathbf{v}) \cdot \mathbf{w} = \mathbf{u} \cdot \mathbf{w} + \mathbf{v} \cdot \mathbf{w}$

(*c*) $\mathbf{u} \cdot (\mathbf{v} - \mathbf{w}) = \mathbf{u} \cdot \mathbf{v} - \mathbf{u} \cdot \mathbf{w}$

(*d*) $(\mathbf{u} - \mathbf{v}) \cdot \mathbf{w} = \mathbf{u} \cdot \mathbf{w} - \mathbf{v} \cdot \mathbf{w}$

(*e*) $k(\mathbf{u} \cdot \mathbf{v}) = \mathbf{u} \cdot (k\mathbf{v})$

Cauchy-Schwarz Inequality

THEOREM 3.2.4 Cauchy–Schwarz Inequality

If $\mathbf{u} = (u_1, u_2, \ldots, u_n)$ and $\mathbf{v} = (v_1, v_2, \ldots, v_n)$ are vectors in R^n, then

$$|\mathbf{u} \cdot \mathbf{v}| \leq \|\mathbf{u}\| \|\mathbf{v}\| \tag{22}$$

or in terms of components

$$|u_1 v_1 + u_2 v_2 + \cdots + u_n v_n| \leq (u_1^2 + u_2^2 + \cdots + u_n^2)^{1/2} (v_1^2 + v_2^2 + \cdots + v_n^2)^{1/2} \tag{23}$$

Dot Products and Matrices

Table 1

Form	Dot Product	Example	
u a column matrix and **v** a column matrix	$\mathbf{u}\cdot\mathbf{v} = \mathbf{u}^T\mathbf{v} = \mathbf{v}^T\mathbf{u}$	$\mathbf{u} = \begin{bmatrix}1\\-3\\5\end{bmatrix}$ $\mathbf{v} = \begin{bmatrix}5\\4\\0\end{bmatrix}$	$\mathbf{u}^T\mathbf{v} = [1 \quad -3 \quad 5]\begin{bmatrix}5\\4\\0\end{bmatrix} = -7$ $\mathbf{v}^T\mathbf{u} = [5 \quad 4 \quad 0]\begin{bmatrix}1\\-3\\5\end{bmatrix} = -7$
u a row matrix and **v** a column matrix	$\mathbf{u}\cdot\mathbf{v} = \mathbf{u}\mathbf{v} = \mathbf{v}^T\mathbf{u}^T$	$\mathbf{u} = [1 \quad -3 \quad 5]$ $\mathbf{v} = \begin{bmatrix}5\\4\\0\end{bmatrix}$	$\mathbf{u}\mathbf{v} = [1 \quad -3 \quad 5]\begin{bmatrix}5\\4\\0\end{bmatrix} = -7$ $\mathbf{v}^T\mathbf{u}^T = [5 \quad 4 \quad 0]\begin{bmatrix}1\\-3\\5\end{bmatrix} = -7$
u a column matrix and **v** a row matrix	$\mathbf{u}\cdot\mathbf{v} = \mathbf{v}\mathbf{u} = \mathbf{u}^T\mathbf{v}^T$	$\mathbf{u} = \begin{bmatrix}1\\-3\\5\end{bmatrix}$ $\mathbf{v} = [5 \quad 4 \quad 0]$	$\mathbf{v}\mathbf{u} = [5 \quad 4 \quad 0]\begin{bmatrix}1\\-3\\5\end{bmatrix} = -7$ $\mathbf{u}^T\mathbf{v}^T = [1 \quad -3 \quad 5]\begin{bmatrix}5\\4\\0\end{bmatrix} = -7$
u a row matrix and **v** a row matrix	$\mathbf{u}\cdot\mathbf{v} = \mathbf{u}\mathbf{v}^T = \mathbf{v}\mathbf{u}^T$	$\mathbf{u} = [1 \quad -3 \quad 5]$ $\mathbf{v} = [5 \quad 4 \quad 0]$	$\mathbf{u}\mathbf{v}^T = [1 \quad -3 \quad 5]\begin{bmatrix}5\\4\\0\end{bmatrix} = -7$ $\mathbf{v}\mathbf{u}^T = [5 \quad 4 \quad 0]\begin{bmatrix}1\\-3\\5\end{bmatrix} = -7$

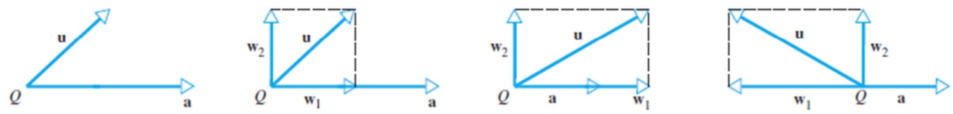
Section 3.3 Orthogonality

DEFINITION 1 Two nonzero vectors **u** and **v** in R^n are said to be ***orthogonal*** (or ***perpendicular***) if $\mathbf{u}\cdot\mathbf{v} = 0$. We will also agree that the zero vector in R^n is orthogonal to *every* vector in R^n. A nonempty set of vectors in R^n is called an ***orthogonal set*** if all pairs of distinct vectors in the set are orthogonal. An orthogonal set of unit vectors is called an ***orthonormal set***.

Orthogonal Projections

THEOREM 3.3.2 Projection Theorem

If **u** *and* **a** *are vectors in* R^n, *and if* $\mathbf{a} \neq \mathbf{0}$, *then* **u** *can be expressed in exactly one way in the form* $\mathbf{u} = \mathbf{w}_1 + \mathbf{w}_2$, *where* $\mathbf{w}_1$ *is a scalar multiple of* **a** *and* $\mathbf{w}_2$ *is orthogonal to* **a**.



Point-line and point-plane Distance formulas

THEOREM 3.3.4

(*a*) *In* R^2 *the distance* D *between the point* $P_0(x_0, y_0)$ *and the line* $ax + by + c = 0$ *is*

$$D = \frac{|ax_0 + by_0 + c|}{\sqrt{a^2 + b^2}} \tag{15}$$

(*b*) *In* R^3 *the distance* D *between the point* $P_0(x_0, y_0, z_0)$ *and the plane* $ax + by + cz + d = 0$ *is*

$$D = \frac{|ax_0 + by_0 + cz_0 + d|}{\sqrt{a^2 + b^2 + c^2}} \tag{16}$$

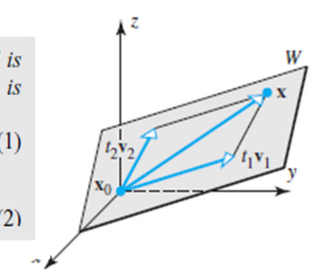
The Geometry of Linear Systems

THEOREM 3.4.1 *Let* L *be the line in* R^2 *or* R^3 *that contains the point* $\mathbf{x}_0$ *and is parallel to the nonzero vector* **v**. *Then the equation of the line through* $\mathbf{x}_0$ *that is parallel to* **v** *is*

$$\mathbf{x} = \mathbf{x}_0 + t\mathbf{v} \tag{1}$$

If $\mathbf{x}_0 = \mathbf{0}$, *then the line passes through the origin and the equation has the form*

$$\mathbf{x} = t\mathbf{v} \tag{2}$$

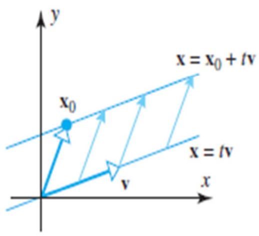


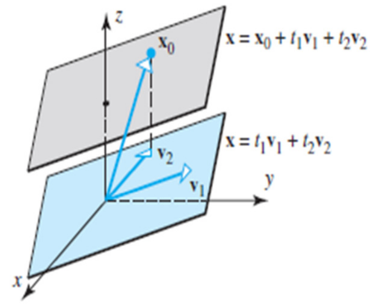
THEOREM 3.4.2 *Let W be the plane in R^3 that contains the point $\mathbf{x}_0$ and is parallel to the noncollinear vectors $\mathbf{v}_1$ and $\mathbf{v}_2$. Then an equation of the plane through $\mathbf{x}_0$ that is parallel to $\mathbf{v}_1$ and $\mathbf{v}_2$ is given by*

$$\mathbf{x} = \mathbf{x}_0 + t_1\mathbf{v}_1 + t_2\mathbf{v}_2 \quad (3)$$

If $\mathbf{x}_0 = \mathbf{0}$, then the plane passes through the origin and the equation has the form

$$\mathbf{x} = t_1\mathbf{v}_1 + t_2\mathbf{v}_2 \quad (4)$$





DEFINITION 1 If $\mathbf{x}_0$ and $\mathbf{v}$ are vectors in R^n, and if $\mathbf{v}$ is nonzero, then the equation

$$\mathbf{x} = \mathbf{x}_0 + t\mathbf{v} \quad (5)$$

defines the ***line through $\mathbf{x}_0$ that is parallel to $\mathbf{v}$***. In the special case where $\mathbf{x}_0 = \mathbf{0}$, the line is said to ***pass through the origin***.

DEFINITION 2 If $\mathbf{x}_0$, $\mathbf{v}_1$, and $\mathbf{v}_2$ are vectors in R^n, and if $\mathbf{v}_1$ and $\mathbf{v}_2$ are not collinear, then the equation

$$\mathbf{x} = \mathbf{x}_0 + t_1\mathbf{v}_1 + t_2\mathbf{v}_2 \quad (6)$$

defines the ***plane through $\mathbf{x}_0$ that is parallel to $\mathbf{v}_1$ and $\mathbf{v}_2$***. In the special case where $\mathbf{x}_0 = \mathbf{0}$, the plane is said to ***pass through the origin***.

Cross Product

DEFINITION 1 If $\mathbf{u} = (u_1, u_2, u_3)$ and $\mathbf{v} = (v_1, v_2, v_3)$ are vectors in 3-space, then the *cross product* $\mathbf{u} \times \mathbf{v}$ is the vector defined by

$$\mathbf{u} \times \mathbf{v} = (u_2v_3 - u_3v_2, u_3v_1 - u_1v_3, u_1v_2 - u_2v_1)$$

or, in determinant notation,

$$\mathbf{u} \times \mathbf{v} = \left(\begin{vmatrix} u_2 & u_3 \\ v_2 & v_3 \end{vmatrix}, - \begin{vmatrix} u_1 & u_3 \\ v_1 & v_3 \end{vmatrix}, \begin{vmatrix} u_1 & u_2 \\ v_1 & v_2 \end{vmatrix} \right) \tag{1}$$

Cross Products and Dot Products

THEOREM 3.5.1 Relationships Involving Cross Product and Dot Product

If $\mathbf{u}$, $\mathbf{v}$, *and* $\mathbf{w}$ *are vectors in 3-space, then*

(*a*) $\mathbf{u} \cdot (\mathbf{u} \times \mathbf{v}) = 0$ ($\mathbf{u} \times \mathbf{v}$ *is orthogonal to* $\mathbf{u}$)

(*b*) $\mathbf{v} \cdot (\mathbf{u} \times \mathbf{v}) = 0$ ($\mathbf{u} \times \mathbf{v}$ *is orthogonal to* $\mathbf{v}$)

(*c*) $\|\mathbf{u} \times \mathbf{v}\|^2 = \|\mathbf{u}\|^2\|\mathbf{v}\|^2 - (\mathbf{u} \cdot \mathbf{v})^2$ (*Lagrange's identity*)

(*d*) $\mathbf{u} \times (\mathbf{v} \times \mathbf{w}) = (\mathbf{u} \cdot \mathbf{w})\mathbf{v} - (\mathbf{u} \cdot \mathbf{v})\mathbf{w}$ (*relationship between cross and dot products*)

(*e*) $(\mathbf{u} \times \mathbf{v}) \times \mathbf{w} = (\mathbf{u} \cdot \mathbf{w})\mathbf{v} - (\mathbf{v} \cdot \mathbf{w})\mathbf{u}$ (*relationship between cross and dot products*)

Properties of Cross Product

THEOREM 3.5.2 Properties of Cross Product

If $\mathbf{u}$, $\mathbf{v}$, *and* $\mathbf{w}$ *are any vectors in 3-space and* k *is any scalar, then:*

(*a*) $\mathbf{u} \times \mathbf{v} = -(\mathbf{v} \times \mathbf{u})$

(*b*) $\mathbf{u} \times (\mathbf{v} + \mathbf{w}) = (\mathbf{u} \times \mathbf{v}) + (\mathbf{u} \times \mathbf{w})$

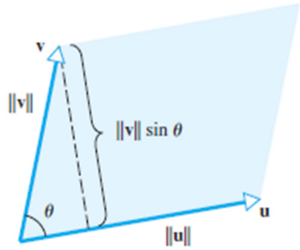
(*c*) $(\mathbf{u} + \mathbf{v}) \times \mathbf{w} = (\mathbf{u} \times \mathbf{w}) + (\mathbf{v} \times \mathbf{w})$

(*d*) $k(\mathbf{u} \times \mathbf{v}) = (k\mathbf{u}) \times \mathbf{v} = \mathbf{u} \times (k\mathbf{v})$

(*e*) $\mathbf{u} \times \mathbf{0} = \mathbf{0} \times \mathbf{u} = \mathbf{0}$

(*f*) $\mathbf{u} \times \mathbf{u} = \mathbf{0}$

Geometry of the Cross Product



$$\|\mathbf{u} \times \mathbf{v}\| = \|\mathbf{u}\| \|\mathbf{v}\| \sin\theta$$

THEOREM 3.5.3 Area of a Parallelogram
If $\mathbf{u}$ *and* $\mathbf{v}$ *are vectors in 3-space, then* $\|\mathbf{u} \times \mathbf{v}\|$ *is equal to the area of the parallelogram determined by* $\mathbf{u}$ *and* $\mathbf{v}$.

WEEK 6

Chapter 4 General Vector Spaces

4.1 Real Vector Spaces
4.2 Subspaces
4.3 Linear Independence
4.4 Coordinates and Basis
4.5 Dimension
4.6 Change of Basis

Section 4.1 Vector Space Axioms

DEFINITION 1 Let V be an arbitrary nonempty set of objects on which two operations are defined: addition, and multiplication by scalars. By ***addition*** we mean a rule for associating with each pair of objects $\mathbf{u}$ and $\mathbf{v}$ in V an object $\mathbf{u} + \mathbf{v}$, called the ***sum*** of $\mathbf{u}$ and $\mathbf{v}$; by ***scalar multiplication*** we mean a rule for associating with each scalar k and each object $\mathbf{u}$ in V an object $k\mathbf{u}$, called the ***scalar multiple*** of $\mathbf{u}$ by k. If the following axioms are satisfied by all objects $\mathbf{u}$, $\mathbf{v}$, $\mathbf{w}$ in V and all scalars k and m, then we call V a ***vector space*** and we call the objects in V ***vectors***.

1. If $\mathbf{u}$ and $\mathbf{v}$ are objects in V, then $\mathbf{u} + \mathbf{v}$ 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 an object $\mathbf{0}$ in V, called a ***zero vector*** for V, such that $\mathbf{0} + \mathbf{u} = \mathbf{u} + \mathbf{0} = \mathbf{u}$ for all $\mathbf{u}$ in V.
5. For each $\mathbf{u}$ in V, there is an object $-\mathbf{u}$ in V, called a ***negative*** of $\mathbf{u}$, such that $\mathbf{u} + (-\mathbf{u}) = (-\mathbf{u}) + \mathbf{u} = \mathbf{0}$.
6. If k is any scalar and $\mathbf{u}$ is any object in V, then $k\mathbf{u}$ is in V.
7. $k(\mathbf{u} + \mathbf{v}) = k\mathbf{u} + k\mathbf{v}$
8. $(k + m)\mathbf{u} = k\mathbf{u} + m\mathbf{u}$
9. $k(m\mathbf{u}) = (km)(\mathbf{u})$
10. $1\mathbf{u} = \mathbf{u}$

Section 4.2 Subspaces

DEFINITION 1 A subset W of a vector space V is called a ***subspace*** of V if W is itself a vector space under the addition and scalar multiplication defined on V.

THEOREM 4.2.1 *If W is a set of one or more vectors in a vector space V, then W is a subspace of V if and only if the following conditions hold.*

(a) If $\mathbf{u}$ and $\mathbf{v}$ are vectors in W, then $\mathbf{u}+\mathbf{v}$ is in W.

(b) If k is any scalar and $\mathbf{u}$ is any vector in W, then $k\mathbf{u}$ is in W.

The 'smallest' subspace of a vector space V

DEFINITION 2 If $\mathbf{w}$ is a vector in a vector space V, then $\mathbf{w}$ is said to be a ***linear combination*** of the vectors $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_r$ in V if $\mathbf{w}$ can be expressed in the form

$$\mathbf{w} = k_1\mathbf{v}_1 + k_2\mathbf{v}_2 + \cdots + k_r\mathbf{v}_r \quad (2)$$

where $k_1, k_2, \dots, k_r$ are scalars. These scalars are called the ***coefficients*** of the linear combination.

THEOREM 4.2.3 *If $S = \{\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_r\}$ is a nonempty set of vectors in a vector space V, then:*

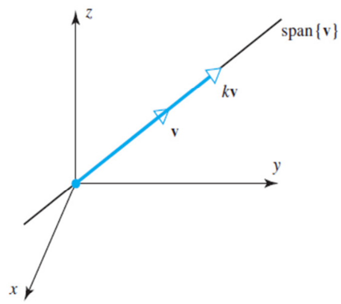
(a) The set W of all possible linear combinations of the vectors in S is a subspace of V.

(b) The set W in part (a) is the "smallest" subspace of V that contains all of the vectors in S in the sense that any other subspace that contains those vectors contains W.

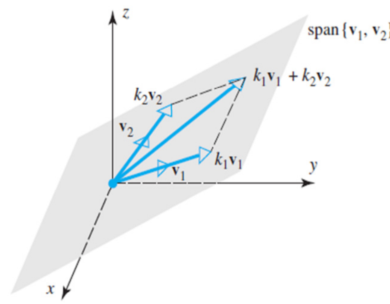
The span of S

DEFINITION 3 The subspace of a vector space V that is formed from all possible linear combinations of the vectors in a nonempty set S is called the ***span of S***, and we say that the vectors in S ***span*** that subspace. If $S = \{\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_r\}$, then we denote the span of S by

$$\text{span}\{\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_r\} \quad \text{or} \quad \text{span}(S)$$



(a) Span{v} is the line through the origin determined by v.



(b) Span{$\mathbf{v}_1, \mathbf{v}_2$} is the plane through the origin determined by $\mathbf{v}_1$ and $\mathbf{v}_2$.

Section 4.3 Linear Independence

DEFINITION 1 If $S = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_r\}$ is a nonempty set of vectors in a vector space V, then the vector equation

$$k_1\mathbf{v}_1 + k_2\mathbf{v}_2 + \cdots + k_r\mathbf{v}_r = \mathbf{0}$$

has at least one solution, namely,

$$k_1 = 0, \quad k_2 = 0, \ldots, \quad k_r = 0$$

We call this the ***trivial solution***. If this is the only solution, then S is said to be a ***linearly independent set***. If there are solutions in addition to the trivial solution, then S is said to be a ***linearly dependent set***.

Linearly independence

THEOREM 4.3.1 *A set S with two or more vectors is*

(*a*) *Linearly dependent if and only if at least one of the vectors in S is expressible as a linear combination of the other vectors in S.*

(*b*) *Linearly independent if and only if no vector in S is expressible as a linear combination of the other vectors in S.*

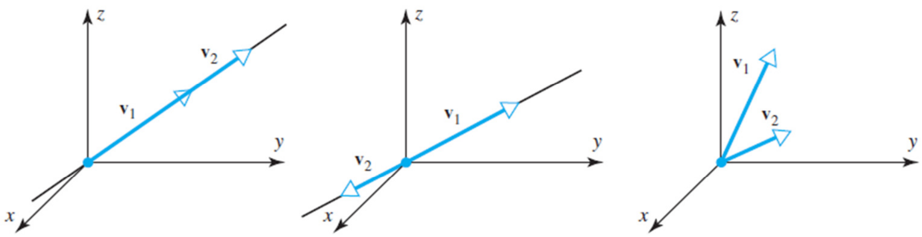
THEOREM 4.3.2

(*a*) *A finite set that contains* **0** *is linearly dependent.*

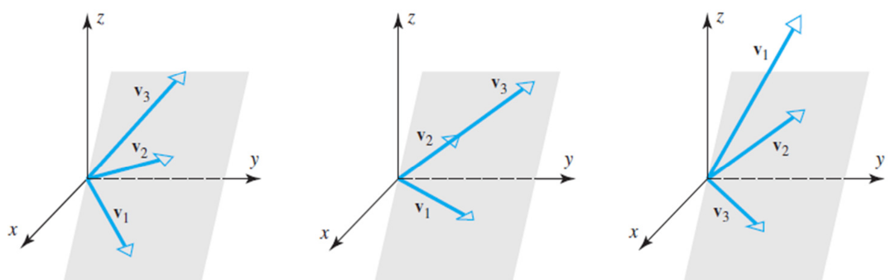
(*b*) *A set with exactly one vector is linearly independent if and only if that vector is not* **0**.

(*c*) *A set with exactly two vectors is linearly independent if and only if neither vector is a scalar multiple of the other.*

Linear Independence in R2 and R3



(*a*) Linearly dependent (*b*) Linearly dependent (*c*) Linearly independent



(*a*) Linearly dependent (*b*) Linearly dependent (*c*) Linearly independent

The Wronskian

DEFINITION 2 If $\mathbf{f}_1 = f_1(x), \mathbf{f}_2 = f_2(x), \ldots, \mathbf{f}_n = f_n(x)$ are functions that are $n-1$ times differentiable on the interval $(-\infty, \infty)$, then the determinant

$$W(x) = \begin{vmatrix} f_1(x) & f_2(x) & \cdots & f_n(x) \\ f_1'(x) & f_2'(x) & \cdots & f_n'(x) \\ \vdots & \vdots & & \vdots \\ f_1^{(n-1)}(x) & f_2^{(n-1)}(x) & \cdots & f_n^{(n-1)}(x) \end{vmatrix}$$

is called the ***Wronskian*** of $f_1, f_2, \ldots, f_n$.

THEOREM 4.3.4 *If the functions* $\mathbf{f}_1, \mathbf{f}_2, \ldots, \mathbf{f}_n$ *have* $n-1$ *continuous derivatives on the interval* $(-\infty, \infty)$, *and if the Wronskian of these functions is not identically zero on* $(-\infty, \infty)$, *then these functions form a linearly independent set of vectors in* $C^{(n-1)}(-\infty, \infty)$.

Section 4.4 Coordinates and Basis

DEFINITION 1 If V is any vector space and $S = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ is a finite set of vectors in V, then S is called a ***basis*** for V if the following two conditions hold:

(a) S is linearly independent.

(b) S spans V.

THEOREM 4.4.1 Uniqueness of Basis Representation

If $S = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ *is a basis for a vector space* V, *then every vector* $\mathbf{v}$ *in* V *can be expressed in the form* $\mathbf{v} = c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_n\mathbf{v}_n$ *in exactly one way.*

The coordinate vector

DEFINITION 2 If $S = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ is a basis for a vector space V, and

$$\mathbf{v} = c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_n\mathbf{v}_n$$

is the expression for a vector $\mathbf{v}$ in terms of the basis S, then the scalars $c_1, c_2, \ldots, c_n$ are called the ***coordinates*** of $\mathbf{v}$ relative to the basis S. The vector $(c_1, c_2, \ldots, c_n)$ in R^n constructed from these coordinates is called the ***coordinate vector of*** $\mathbf{v}$ ***relative to*** S; it is denoted by

$$(\mathbf{v})_S = (c_1, c_2, \ldots, c_n) \tag{6}$$

Section 4.5 Dimension

DEFINITION 1 The ***dimension*** of a finite-dimensional vector space V is denoted by $\dim(V)$ and is defined to be the number of vectors in a basis for V. In addition, the zero vector space is defined to have dimension zero.

Plus / Minus Theorem

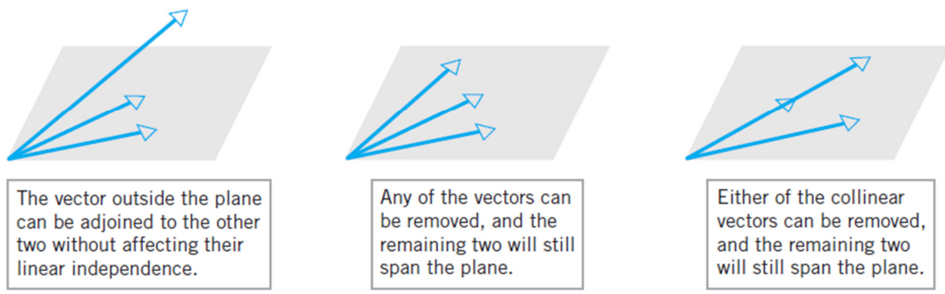
THEOREM 4.5.3 Plus/Minus Theorem

Let S be a nonempty set of vectors in a vector space V.

(a) *If S is a linearly independent set, and if $\mathbf{v}$ is a vector in V that is outside of span(S), then the set $S \cup \{\mathbf{v}\}$ that results by inserting $\mathbf{v}$ into S is still linearly independent.*

(b) *If $\mathbf{v}$ is a vector in S that is expressible as a linear combination of other vectors in S, and if $S - \{\mathbf{v}\}$ denotes the set obtained by removing $\mathbf{v}$ from S, then S and $S - \{\mathbf{v}\}$ span the same space; that is,*

$$\text{span}(S) = \text{span}(S - \{\mathbf{v}\})$$



THEOREM 4.5.4 *Let V be an n-dimensional vector space, and let S be a set in V with exactly n vectors. Then S is a basis for V if and only if S spans V or S is linearly independent.*

THEOREM 4.5.5 *Let S be a finite set of vectors in a finite-dimensional vector space V.*

(a) *If S spans V but is not a basis for V, then S can be reduced to a basis for V by removing appropriate vectors from S.*

(b) *If S is a linearly independent set that is not already a basis for V, then S can be enlarged to a basis for V by inserting appropriate vectors into S.*

THEOREM 4.5.6 *If W is a subspace of a finite-dimensional vector space V, then:*

(a) *W is finite-dimensional.*

(b) $\dim(W) \leq \dim(V)$.

(c) *$W = V$ if and only if* $\dim(W) = \dim(V)$.

Section 4.6 Change of Basis

The Change-of-Basis Problem If $\mathbf{v}$ is a vector in a finite-dimensional vector space V, and if we change the basis for V from a basis B to a basis B', how are the coordinate vectors $[\mathbf{v}]_B$ and $[\mathbf{v}]_{B'}$ related?

Solution of the Change-of-Basis Problem If we change the basis for a vector space V from an old basis $B = \{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n\}$ to a new basis $B' = \{\mathbf{u}'_1, \mathbf{u}'_2, \ldots, \mathbf{u}'_n\}$, then for each vector $\mathbf{v}$ in V, the old coordinate vector $[\mathbf{v}]_B$ is related to the new coordinate vector $[\mathbf{v}]_{B'}$ by the equation

$$[\mathbf{v}]_B = P[\mathbf{v}]_{B'} \tag{7}$$

where the columns of P are the coordinate vectors of the new basis vectors relative to the old basis; that is, the column vectors of P are

$$[\mathbf{u}'_1]_B, \quad [\mathbf{u}'_2]_B, \ldots, \quad [\mathbf{u}'_n]_B \tag{8}$$

Transition Matrices

The columns of the transition matrix from an old basis to a new basis are the coordinate vectors of the old basis relative to the new basis.

THEOREM 4.6.1 *If P is the transition matrix from a basis B' to a basis B for a finite-dimensional vector space V, then P is invertible and P^{-1} is the transition matrix from B to B'.*

Computing the transition matrix

A Procedure for Computing $P_{B \to B'}$

Step 1. Form the matrix $[B' \mid B]$.

Step 2. Use elementary row operations to reduce the matrix in Step 1 to reduced row echelon form.

Step 3. The resulting matrix will be $[I \mid P_{B \to B'}]$.

Step 4. Extract the matrix $P_{B \to B'}$ from the right side of the matrix in Step 3.

This procedure is captured in the following diagram.

$$[\text{new basis} \mid \text{old basis}] \xrightarrow{\text{row operations}} [I \mid \text{transition from old to new}] \quad (14)$$

WEEK 7

Chapter 4 General Vector Spaces

4.7 Row Space, Column Space, and Null Space

4.8 Rank, Nullity, and the Fundamental Matrix Spaces

4.9 Matrix Transformations from RR^n to R^m

4.10 Properties of Matrix Transformations

4.11 Geometry of Matrix Operators on R^2

Row Space, Column Space, and Null Space

DEFINITION 1 For an $m \times n$ matrix

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}$$

the vectors

$$\begin{aligned} \mathbf{r}_1 &= [a_{11} \quad a_{12} \quad \cdots \quad a_{1n}] \\ \mathbf{r}_2 &= [a_{21} \quad a_{22} \quad \cdots \quad a_{2n}] \\ &\vdots \\ \mathbf{r}_m &= [a_{m1} \quad a_{m2} \quad \cdots \quad a_{mn}] \end{aligned}$$

in R^n that are formed from the rows of A are called the ***row vectors*** of A, and the vectors

$$\mathbf{c}_1 = \begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{bmatrix}, \quad \mathbf{c}_2 = \begin{bmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{bmatrix}, \ldots, \quad \mathbf{c}_n = \begin{bmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{bmatrix}$$

in R^m formed from the columns of A are called the ***column vectors*** of A.

Row, column and null spaces

DEFINITION 2 If A is an $m \times n$ matrix, then the subspace of R^n spanned by the row vectors of A is called the ***row space*** of A, and the subspace of R^m spanned by the column vectors of A is called the ***column space*** of A. The solution space of the homogeneous system of equations $A\mathbf{x} = \mathbf{0}$, which is a subspace of R^n, is called the ***null space*** of A.

Systems of linear equations

Question 1. What relationships exist among the solutions of a linear system $A\mathbf{x} = \mathbf{b}$ and the row space, column space, and null space of the coefficient matrix A?

Question 2. What relationships exist among the row space, column space, and null space of a matrix?

THEOREM 4.7.1 *A system of linear equations $A\mathbf{x} = \mathbf{b}$ is consistent if and only if $\mathbf{b}$ is in the column space of A.*

THEOREM 4.7.2 *If $\mathbf{x}_0$ is any solution of a consistent linear system $A\mathbf{x} = \mathbf{b}$, and if $S = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\}$ is a basis for the null space of A, then every solution of $A\mathbf{x} = \mathbf{b}$ can be expressed in the form*

$$\mathbf{x} = \mathbf{x}_0 + c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_k\mathbf{v}_k \qquad (3)$$

Conversely, for all choices of scalars $c_1, c_2, \ldots, c_k$, the vector $\mathbf{x}$ in this formula is a solution of $A\mathbf{x} = \mathbf{b}$.

A basis for span (S)

Problem Given a set of vectors $S = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\}$ in R^n, find a subset of these vectors that forms a basis for span(S), and express those vectors that are not in that basis as a linear combination of the basis vectors.

Basis for Span(S)

Step 1. Form the matrix A having vectors in $S = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\}$ as column vectors.

Step 2. Reduce the matrix A to reduced row echelon form R.

Step 3. Denote the column vectors of R by $\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_k$.

Step 4. Identify the columns of R that contain the leading 1's. The corresponding column vectors of A form a basis for span(S).

Step 5. Obtain a set of dependency equations by expressing each column vector of R that does not contain a leading 1 as a linear combination of preceding column vectors that do contain leading 1's.

Step 6. Replace the column vectors of R that appear in the dependency equations by the corresponding column vectors of A.

This completes the second part of the problem.

Rank and Nullity

DEFINITION 1 The common dimension of the row space and column space of a matrix A is called the ***rank*** of A and is denoted by rank(A); the dimension of the null space of A is called the ***nullity*** of A and is denoted by nullity(A).

THEOREM 4.8.2 Dimension Theorem for Matrices

If A is a matrix with n columns, then

$$\text{rank}(A) + \text{nullity}(A) = n$$

THEOREM 4.8.3 *If A is an $m \times n$ matrix, then*

(a) rank(A) = the number of leading variables in the general solution of $A\mathbf{x} = \mathbf{0}$.

(b) nullity(A) = the number of parameters in the general solution of $A\mathbf{x} = \mathbf{0}$.

THEOREM 4.8.4 Equivalent Statements

If A is an $n \times n$ matrix, then the following statements are equivalent.

(a) A is invertible.

(b) $A\mathbf{x} = \mathbf{0}$ has only the trivial solution.

(c) The reduced row echelon form of A is I_n.

(d) A is expressible as a product of elementary matrices.

(e) $A\mathbf{x} = \mathbf{b}$ is consistent for every $n \times 1$ matrix $\mathbf{b}$.

(f) $A\mathbf{x} = \mathbf{b}$ has exactly one solution for every $n \times 1$ matrix $\mathbf{b}$.

(g) $\det(A) \neq 0$.

(h) The column vectors of A are linearly independent.

(i) The row vectors of A are linearly independent.

(j) The column vectors of A span R^n.

(k) The row vectors of A span R^n.

(l) The column vectors of A form a basis for R^n.

(m) The row vectors of A form a basis for R^n.

(n) A has rank n.

(o) A has nullity 0.

Fundamental Spaces of Matrix A

- Row space of A
- Null space of A
- Column space of A
- Null space of A^T

THEOREM 4.8.10 Equivalent Statements

If A is an $n \times n$ matrix, then the following statements are equivalent.

(a) A *is invertible.*

(b) $A\mathbf{x} = \mathbf{0}$ *has only the trivial solution.*

(c) *The reduced row echelon form of* A *is* I_n.

(d) A *is expressible as a product of elementary matrices.*

(e) $A\mathbf{x} = \mathbf{b}$ *is consistent for every* $n \times 1$ *matrix* $\mathbf{b}$.

(f) $A\mathbf{x} = \mathbf{b}$ *has exactly one solution for every* $n \times 1$ *matrix* $\mathbf{b}$.

(g) $\det(A) \neq 0$.

(h) *The column vectors of* A *are linearly independent.*

(i) *The row vectors of* A *are linearly independent.*

(j) *The column vectors of* A *span* R^n.

(k) *The row vectors of* A *span* R^n.

(l) *The column vectors of* A *form a basis for* R^n.

(m) *The row vectors of* A *form a basis for* R^n.

(n) A *has rank* n.

(o) A *has nullity* 0.

(p) *The orthogonal complement of the null space of* A *is* R^n.

(q) *The orthogonal complement of the row space of* A *is* $\{\mathbf{0}\}$.

Matrix Transformations from R^n to R^m

DEFINITION 1 If V and W are vector spaces, and if f is a function with domain V and codomain W, then we say that f is a ***transformation*** from V to W or that f ***maps*** V to W, which we denote by writing

$$f: V \rightarrow W$$

In the special case where $V = W$, the transformation is also called an ***operator*** on V.

THEOREM 4.9.1 *For every matrix* A *the matrix transformation* $T_A: R^n \rightarrow R^m$ *has the following properties for all vectors* $\mathbf{u}$ *and* $\mathbf{v}$ *in* R^n *and for every scalar* k:

(a) $T_A(\mathbf{0}) = \mathbf{0}$

(b) $T_A(k\mathbf{u}) = kT_A(\mathbf{u})$ **[Homogeneity property]**

(c) $T_A(\mathbf{u} + \mathbf{v}) = T_A(\mathbf{u}) + T_A(\mathbf{v})$ **[Additivity property]**

(d) $T_A(\mathbf{u} - \mathbf{v}) = T_A(\mathbf{u}) - T_A(\mathbf{v})$

Properties of Matrix Transformations

DEFINITION 1 A matrix transformation $T_A: R^n \to R^m$ is said to be ***one-to-one*** if T_A maps distinct vectors (points) in R^n into distinct vectors (points) in R^m.

THEOREM 4.10.1 *If A is an $n \times n$ matrix and $T_A: R^n \to R^n$ is the corresponding matrix operator, then the following statements are equivalent.*

(a) A is invertible.

(b) The range of T_A is R^n.

(c) T_A is one-to-one.

THEOREM 4.10.2 *$T: R^n \to R^m$ is a matrix transformation if and only if the following relationships hold for all vectors $\mathbf{u}$ and $\mathbf{v}$ in R^n and for every scalar k:*

(i) $T(\mathbf{u}+\mathbf{v}) = T(\mathbf{u}) + T(\mathbf{v})$ **[Additivity property]**

(ii) $T(k\mathbf{u}) = kT(\mathbf{u})$ **[Homogeneity property]**

THEOREM 4.10.3 *Every linear transformation from R^n to R^m is a matrix transformation, and conversely, every matrix transformation from R^n to R^m is a linear transformation.*

THEOREM 4.10.4 Equivalent Statements

If A *is an* $n \times n$ *matrix, then the following statements are equivalent.*

(*a*) A *is invertible.*

(*b*) $A\mathbf{x} = \mathbf{0}$ *has only the trivial solution.*

(*c*) *The reduced row echelon form of* A *is* I_n.

(*d*) A *is expressible as a product of elementary matrices.*

(*e*) $A\mathbf{x} = \mathbf{b}$ *is consistent for every* $n \times 1$ *matrix* $\mathbf{b}$.

(*f*) $A\mathbf{x} = \mathbf{b}$ *has exactly one solution for every* $n \times 1$ *matrix* $\mathbf{b}$.

(*g*) $\det(A) \neq 0$.

(*h*) *The column vectors of* A *are linearly independent.*

(*i*) *The row vectors of* A *are linearly independent.*

(*j*) *The column vectors of* A *span* R^n.

(*k*) *The row vectors of* A *span* R^n.

(*l*) *The column vectors of* A *form a basis for* R^n.

(*m*) *The row vectors of* A *form a basis for* R^n.

(*n*) A *has rank* n.

(*o*) A *has nullity* 0.

(*p*) *The orthogonal complement of the null space of* A *is* R^n.

(*q*) *The orthogonal complement of the row space of* A *is* $\{\mathbf{0}\}$.

(*r*) *The range of* T_A *is* R^n.

(*s*) T_A *is one-to-one.*

WEEK 9

Chapter 5 Eigenvalues and Eigenvectors

5.1 Eigenvalues and Eigenvectors

5.2 Diagonalization

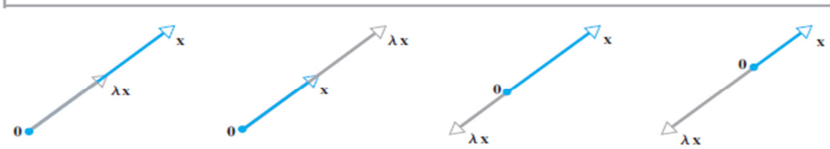
5.3 Complex Vector Spaces

Section 5.1 Eigenvalues and Eigenvectors

DEFINITION 1 If A is an $n \times n$ matrix, then a nonzero vector $\mathbf{x}$ in R^n is called an ***eigenvector*** of A (or of the matrix operator T_A) if $A\mathbf{x}$ is a scalar multiple of $\mathbf{x}$; that is,

$$A\mathbf{x} = \lambda \mathbf{x}$$

for some scalar λ. The scalar λ is called an ***eigenvalue*** of A (or of T_A), and $\mathbf{x}$ is said to be an ***eigenvector corresponding to*** $\boldsymbol{\lambda}$.



(*a*) $0 \le \lambda \le 1$ (*b*) $\lambda \le 1$ (*c*) $-1 \le \lambda \le 0$ (*d*) $\lambda \le -1$

The Characteristic Equation

THEOREM 5.1.1 *If A is an $n \times n$ matrix, then λ is an eigenvalue of A if and only if it satisfies the equation*

$$\det(\lambda I - A) = 0 \tag{1}$$

*This is called the **characteristic equation** of A.*

THEOREM 5.1.3 *If A is an $n \times n$ matrix, the following statements are equivalent.*

(*a*) *λ is an eigenvalue of A.*

(*b*) *The system of equations $(\lambda I - A)\mathbf{x} = \mathbf{0}$ has nontrivial solutions.*

(*c*) *There is a nonzero vector $\mathbf{x}$ such that $A\mathbf{x} = \lambda\mathbf{x}$.*

(*d*) *λ is a solution of the characteristic equation $\det(\lambda I - A) = 0$.*

EXAMPLE 2 Finding Eigenvalues

In Example 1 we observed that $\lambda = 3$ is an eigenvalue of the matrix

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

but we did not explain how we found it. Use the characteristic equation to find all eigenvalues of this matrix.

Solution It follows from Formula 1 that the eigenvalues of A are the solutions of the equation $\det(\lambda I - A) = 0$, which we can write as

$$\begin{vmatrix} \lambda - 3 & 0 \\ -8 & \lambda + 1 \end{vmatrix} = 0$$

from which we obtain

$$(\lambda - 3)(\lambda + 1) = 0 \tag{2}$$

This shows that the eigenvalues of A are $\lambda = 3$ and $\lambda = -1$. Thus, in addition to the eigenvalue $\lambda = 3$ noted in Example 1, we have discovered a second eigenvalue $\lambda = -1$.

$$\begin{bmatrix} \lambda - 3 & 0 \\ -8 & \lambda + 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

If $\lambda = 3$, then this equation becomes

$$\begin{bmatrix} 0 & 0 \\ -8 & 4 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

whose general solution is

$$x_1 = \frac{1}{2}t, \quad x_2 = t$$

(verify) or in matrix form,

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} \frac{1}{2}t \\ t \end{bmatrix} = t \begin{bmatrix} \frac{1}{2} \\ 1 \end{bmatrix}$$

Thus,

$$\begin{bmatrix} \frac{1}{2} \\ 1 \end{bmatrix}$$

is a basis for the eigenspace corresponding to $\lambda = 3$. We leave it as an exercise for you to follow the pattern of these computations and show that

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

is a basis for the eigenspace corresponding to $\lambda = -1$.

THEOREM 5.1.6 Equivalent Statements

If A is an $n \times n$ matrix, then the following statements are equivalent.

(a) *A is invertible.*

(b) *$A\mathbf{x} = \mathbf{0}$ has only the trivial solution.*

(c) *The reduced row echelon form of A is I_n.*

(d) *A is expressible as a product of elementary matrices.*

(e) *$A\mathbf{x} = \mathbf{b}$ is consistent for every $n \times 1$ matrix $\mathbf{b}$.*

(f) *$A\mathbf{x} = \mathbf{b}$ has exactly one solution for every $n \times 1$ matrix $\mathbf{b}$.*

(g) $\det(A) \neq 0$.

(h) *The column vectors of A are linearly independent.*

(i) *The row vectors of A are linearly independent.*

(j) *The column vectors of A span R^n.*

(k) *The row vectors of A span R^n.*

(l) *The column vectors of A form a basis for R^n.*

(m) *The row vectors of A form a basis for R^n.*

(n) *A has rank n.*

(o) *A has nullity 0.*

(p) *The orthogonal complement of the null space of A is R^n.*

(q) *The orthogonal complement of the row space of A is $\{\mathbf{0}\}$.*

(r) *The range of T_A is R^n.*

(s) *T_A is one-to-one.*

(t) *$\lambda = 0$ is not an eigenvalue of A.*

Section 5.2 Diagonalization

DEFINITION 1 If A and B are square matrices, then we say that ***B is similar to A*** if there is an invertible matrix P such that $B = P^{-1}AP$.

DEFINITION 2 A square matrix A is said to be ***diagonalizable*** if it is similar to some diagonal matrix; that is, if there exists an invertible matrix P such that $P^{-1}AP$ is diagonal. In this case the matrix P is said to ***diagonalize*** A.

Similarity Invariants

Table 1 Similarity Invariants

Property	Description
Determinant	A and $P^{-1}AP$ have the same determinant.
Invertibility	A is invertible if and only if $P^{-1}AP$ is invertible.
Rank	A and $P^{-1}AP$ have the same rank.
Nullity	A and $P^{-1}AP$ have the same nullity.
Trace	A and $P^{-1}AP$ have the same trace.
Characteristic polynomial	A and $P^{-1}AP$ have the same characteristic polynomial.
Eigenvalues	A and $P^{-1}AP$ have the same eigenvalues.
Eigenspace dimension	If λ is an eigenvalue of A and hence of $P^{-1}AP$, then the eigenspace of A corresponding to λ and the eigenspace of $P^{-1}AP$ corresponding to λ have the same dimension.

Diagonalizing a Matrix

Procedure for Diagonalizing a Matrix

Step 1. Confirm that the matrix is actually diagonalizable by finding n linearly independent eigenvectors. One way to do this is by finding a basis for each eigenspace and merging these basis vectors into a single set S. If this set has fewer than n vectors, then the matrix is not diagonalizable.

Step 2. Form the matrix $P = [\mathbf{p}_1 \quad \mathbf{p}_2 \quad \cdots \quad \mathbf{p}_n]$ that has the vectors in S as its column vectors.

Step 3. The matrix $P^{-1}AP$ will be diagonal and have the eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$ corresponding to the eigenvectors $\mathbf{p}_1, \mathbf{p}_2, \ldots, \mathbf{p}_n$ as its successive diagonal entries.

Find a matrix P that diagonalizes

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

Solution In Example 7 of the preceding section we found the characteristic equation of A

$$(\lambda - 1)(\lambda - 2)^2 = 0$$

and we found the following bases for the eigenspaces:

$$\lambda = 2: \ \mathbf{p}_1 = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}, \ \mathbf{p}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}; \quad \lambda = 1: \ \mathbf{p}_3 = \begin{bmatrix} -2 \\ 1 \\ 1 \end{bmatrix}$$

There are three basis vectors in total, so the matrix

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

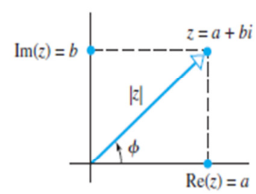
diagonalizes A. As a check, you should verify that

$$P^{-1}AP = \begin{bmatrix} 1 & 0 & 2 \\ 1 & 1 & 1 \\ -1 & 0 & -1 \end{bmatrix} \begin{bmatrix} 0 & 0 & -2 \\ 1 & 2 & 1 \\ 21 & 0 & 3 \end{bmatrix} \begin{bmatrix} -1 & 0 & -2 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Section 5.3 Complex Vector Spaces

Recall that if $z = a + bi$ is a complex number, then:

- $\text{Re}(z) = a$ and $\text{Im}(z) = b$ are called the ***real part*** of z and the ***imaginary part*** of z, respectively,
- $|z| = \sqrt{a^2 + b^2}$ is called the ***modulus*** (or ***absolute value***) of z,
- $\overline{z} = a - bi$ is called the ***complex conjugate*** of z,
- $z\overline{z} = a^2 + b^2 = |z|^2$,
- the angle ϕ in Figure 5.3.1 is called an ***argument*** of z,
- $\text{Re}(z) = |z| \cos \phi$
- $\text{Im}(z) = |z| \sin \phi$
- $z = |z|(\cos \phi + i \sin \phi)$ is called the ***polar form*** of z.



Algebraic Properties of the Complex Conjugate

THEOREM 5.3.1 *If* $\mathbf{u}$ *and* $\mathbf{v}$ *are vectors in* C^n, *and if* k *is a scalar, then*:

(*a*) $\overline{\overline{\mathbf{u}}} = \mathbf{u}$

(*b*) $\overline{k\mathbf{u}} = \overline{k}\overline{\mathbf{u}}$

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

(*d*) $\overline{\mathbf{u} - \mathbf{v}} = \overline{\mathbf{u}} - \overline{\mathbf{v}}$

THEOREM 5.3.2 *If A is an $m \times k$ complex matrix and B is a $k \times n$ complex matrix, then:*

(a) $\overline{\overline{A}} = A$

(b) $\overline{(A^T)} = (\overline{A})^T$

(c) $\overline{AB} = \overline{A}\,\overline{B}$

Dot Product and Norm

DEFINITION 2 If $\mathbf{u} = (u_1, u_2, \ldots, u_n)$ and $\mathbf{v} = (v_1, v_2, \ldots, v_n)$ are vectors in C^n, then the ***complex Euclidean inner product*** of $\mathbf{u}$ and $\mathbf{v}$ (also called the ***complex dot product***) is denoted by $\mathbf{u} \cdot \mathbf{v}$ and is defined as

$$\mathbf{u} \cdot \mathbf{v} = u_1\overline{v}_1 + u_2\overline{v}_2 + \cdots + u_n\overline{v}_n \tag{3}$$

We also define the ***Euclidean norm*** on C^n to be

$$\|\mathbf{v}\| = \sqrt{\mathbf{v} \cdot \mathbf{v}} = \sqrt{|v_1|^2 + |v_2|^2 + \cdots + |v_n|^2} \tag{4}$$

As in the real case, we call $\mathbf{v}$ a ***unit vector*** in C^n if $\|\mathbf{v}\| = 1$, and we say two vectors $\mathbf{u}$ and $\mathbf{v}$ are ***orthogonal*** if $\mathbf{u} \cdot \mathbf{v} = 0$.

Properties of the Dot Product

THEOREM 5.3.3 *If* $\mathbf{u}$, $\mathbf{v}$, *and* $\mathbf{w}$ *are vectors in* C^n, *and if k is a scalar, then the complex Euclidean inner product has the following properties:*

(a) $\mathbf{u} \cdot \mathbf{v} = \overline{\mathbf{v} \cdot \mathbf{u}}$ **[Antisymmetry property]**

(b) $\mathbf{u} \cdot (\mathbf{v} + \mathbf{w}) = \mathbf{u} \cdot \mathbf{v} + \mathbf{u} \cdot \mathbf{w}$ **[Distributive property]**

(c) $k(\mathbf{u} \cdot \mathbf{v}) = (k\mathbf{u}) \cdot \mathbf{v}$ **[Homogeneity property]**

(d) $\mathbf{u} \cdot k\mathbf{v} = \overline{k}(\mathbf{u} \cdot \mathbf{v})$ **[Antihomogeneity property]**

(e) $\mathbf{v} \cdot \mathbf{v} \geq 0$ *and* $\mathbf{v} \cdot \mathbf{v} = 0$ *if and only if* $\mathbf{v} = \mathbf{0}$. **[Positivity property]**

WEEK 10

Chapter 6 Inner Product Spaces

6.1 Inner Products

6.2 Angle and Orthogonality in Inner Product Spaces

6.4 Best Approximation; Least Squares

6.1 Inner Products

DEFINITION 1 An ***inner product*** on a real vector space V is a function that associates a real number $\langle \mathbf{u}, \mathbf{v} \rangle$ with each pair of vectors in V in such a way that the following axioms are satisfied for all vectors $\mathbf{u}$, $\mathbf{v}$, and $\mathbf{w}$ in V and all scalars k.

1. $\langle \mathbf{u}, \mathbf{v} \rangle = \langle \mathbf{v}, \mathbf{u} \rangle$ **[Symmetry axiom]**
2. $\langle \mathbf{u} + \mathbf{v}, \mathbf{w} \rangle = \langle \mathbf{u}, \mathbf{w} \rangle + \langle \mathbf{v}, \mathbf{w} \rangle$ **[Additivity axiom]**
3. $\langle k\mathbf{u}, \mathbf{v} \rangle = k\langle \mathbf{u}, \mathbf{v} \rangle$ **[Homogeneity axiom]**
4. $\langle \mathbf{v}, \mathbf{v} \rangle \geq 0$ and $\langle \mathbf{v}, \mathbf{v} \rangle = 0$ if and only if $\mathbf{v} = \mathbf{0}$ **[Positivity axiom]**

A real vector space with an inner product is called a ***real inner product space***.

$$\langle \mathbf{u}, \mathbf{v} \rangle = \mathbf{u} \cdot \mathbf{v} = u_1 v_1 + u_2 v_2 + \cdots + u_n v_n$$

Algebraic Properties of Inner Products

THEOREM 6.1.2 *If* $\mathbf{u}$, $\mathbf{v}$, *and* $\mathbf{w}$ *are vectors in a real inner product space* V, *and if* k *is a scalar, then*:

(*a*) $\langle \mathbf{0}, \mathbf{v} \rangle = \langle \mathbf{v}, \mathbf{0} \rangle = 0$

(*b*) $\langle \mathbf{u}, \mathbf{v} + \mathbf{w} \rangle = \langle \mathbf{u}, \mathbf{v} \rangle + \langle \mathbf{u}, \mathbf{w} \rangle$

(*c*) $\langle \mathbf{u}, \mathbf{v} - \mathbf{w} \rangle = \langle \mathbf{u}, \mathbf{v} \rangle - \langle \mathbf{u}, \mathbf{w} \rangle$

(*d*) $\langle \mathbf{u} - \mathbf{v}, \mathbf{w} \rangle = \langle \mathbf{u}, \mathbf{w} \rangle - \langle \mathbf{v}, \mathbf{w} \rangle$

(*e*) $k\langle \mathbf{u}, \mathbf{v} \rangle = \langle \mathbf{u}, k\mathbf{v} \rangle$

θ: the angle between u and v

$$\theta = \cos^{-1}\left(\frac{\langle \mathbf{u}, \mathbf{v} \rangle}{\|\mathbf{u}\|\,\|\mathbf{v}\|}\right)$$

DEFINITION 1 Two vectors $\mathbf{u}$ and $\mathbf{v}$ in an inner product space are called ***orthogonal*** if $\langle \mathbf{u}, \mathbf{v} \rangle = 0$.

6.4 Best Approximation: Least Squares

Least Squares Problem Given a linear system $A\mathbf{x} = \mathbf{b}$ of m equations in n unknowns, find a vector $\mathbf{x}$ that minimizes $\|\mathbf{b} - A\mathbf{x}\|$ with respect to the Euclidean inner product on R^m. We call such an $\mathbf{x}$ a ***least squares solution*** of the system, we call $\mathbf{b} - A\mathbf{x}$ the ***least squares error vector***, and we call $\|\mathbf{b} - A\mathbf{x}\|$ the ***least squares error***.

THEOREM 6.4.1 Best Approximation Theorem

*If W is a finite-dimensional subspace of an inner product space V, and if $\mathbf{b}$ is a vector in V, then $\text{proj}_W \mathbf{b}$ is the **best approximation** to $\mathbf{b}$ from W in the sense that*

$$\|\mathbf{b} - \text{proj}_W \mathbf{b}\| < \|\mathbf{b} - \mathbf{w}\|$$

for every vector $\mathbf{w}$ in W that is different from $\text{proj}_W \mathbf{b}$.

Least squares solutions to Ax = b

THEOREM 6.4.2 *For every linear system $A\mathbf{x} = \mathbf{b}$, the associated normal system*

$$A^T A\mathbf{x} = A^T\mathbf{b} \tag{5}$$

is consistent, and all solutions of (5) are least squares solutions of $A\mathbf{x} = \mathbf{b}$. Moreover, if W is the column space of A, and $\mathbf{x}$ is any least squares solution of $A\mathbf{x} = \mathbf{b}$, then the orthogonal projection of $\mathbf{b}$ on W is

$$\text{proj}_W \mathbf{b} = A\mathbf{x} \tag{6}$$

EXAMPLE 1 Least Squares Solution

(a) Find all least squares solutions of the linear system

$$\begin{aligned} x_1 - x_2 &= 4 \\ 3x_1 + 2x_2 &= 1 \\ -2x_1 + 4x_2 &= 3 \end{aligned}$$

(b) Find the error vector and the error.

Solution

(a) It will be convenient to express the system in the matrix form $A\mathbf{x} = \mathbf{b}$, where

$$A = \begin{bmatrix} 1 & -1 \\ 3 & 2 \\ -2 & 4 \end{bmatrix} \quad \text{and} \quad \mathbf{b} = \begin{bmatrix} 4 \\ 1 \\ 3 \end{bmatrix}$$

It follows that

$$A^T A = \begin{bmatrix} 1 & 3 & -2 \\ -1 & 2 & 4 \end{bmatrix} \begin{bmatrix} 1 & -1 \\ 3 & 2 \\ -2 & 4 \end{bmatrix} = \begin{bmatrix} 14 & -3 \\ -3 & 21 \end{bmatrix}$$

$$A^T\mathbf{b} = \begin{bmatrix} 1 & 3 & -2 \\ -1 & 2 & 4 \end{bmatrix} \begin{bmatrix} 4 \\ 1 \\ 3 \end{bmatrix} = \begin{bmatrix} 1 \\ 10 \end{bmatrix}$$

so the normal system $A^T A\mathbf{x} = A^T\mathbf{b}$ is

$$\begin{bmatrix} 14 & -3 \\ -3 & 21 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 1 \\ 10 \end{bmatrix}$$

Solving this system yields a unique least squares solution, namely,

$$x_1 = \frac{17}{95}, \quad x_2 = \frac{143}{285}$$

(b) The error vector is

$$\mathbf{b} - A\mathbf{x} = \begin{bmatrix} 4 \\ 1 \\ 3 \end{bmatrix} - \begin{bmatrix} 1 & -1 \\ 3 & 2 \\ -2 & 4 \end{bmatrix} \begin{bmatrix} \frac{17}{95} \\ \frac{143}{285} \end{bmatrix} = \begin{bmatrix} 4 \\ 1 \\ 3 \end{bmatrix} - \begin{bmatrix} -\frac{92}{285} \\ \frac{439}{285} \\ \frac{95}{57} \end{bmatrix} = \begin{bmatrix} \frac{1232}{285} \\ -\frac{154}{285} \\ \frac{4}{3} \end{bmatrix}$$

and the error is

$$\|\mathbf{b} - A\mathbf{x}\| \approx 4.556$$

WEEK 11

Chapter 7

7.1 Orthogonal Matrices

7.2 Orthogonal Diagonalization

7.3 Quadratic Forms

7.5 Hermitian, Unitary Matrices

7.1 Orthogonal Matrices

DEFINITION 1 A square matrix A is said to be ***orthogonal*** if its transpose is the same as its inverse, that is, if

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

or, equivalently, if

$$AA^T = A^TA = I \quad (1)$$

THEOREM 7.1.1 *The following are equivalent for an $n \times n$ matrix A.*

(*a*) *A is orthogonal.*

(*b*) *The row vectors of A form an orthonormal set in R^n with the Euclidean inner product.*

(*c*) *The column vectors of A form an orthonormal set in R^n with the Euclidean inner product.*

EXAMPLE 1 A 3 × 3 Orthogonal Matrix

The matrix

$$A=\begin{bmatrix} \frac{3}{7} & \frac{2}{7} & \frac{6}{7} \\ -\frac{6}{7} & \frac{3}{7} & \frac{2}{7} \\ \frac{2}{7} & \frac{6}{7} & -\frac{3}{7} \end{bmatrix}$$

is orthogonal since

$$A^TA=\begin{bmatrix} \frac{3}{7} & -\frac{6}{7} & \frac{2}{7} \\ \frac{2}{7} & \frac{3}{7} & \frac{6}{7} \\ \frac{6}{7} & \frac{2}{7} & -\frac{3}{7} \end{bmatrix}\begin{bmatrix} \frac{3}{7} & \frac{2}{7} & \frac{6}{7} \\ -\frac{6}{7} & \frac{3}{7} & \frac{2}{7} \\ \frac{2}{7} & \frac{6}{7} & -\frac{3}{7} \end{bmatrix}=\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

THEOREM 7.1.2

(*a*) *The inverse of an orthogonal matrix is orthogonal.*

(*b*) *A product of orthogonal matrices is orthogonal.*

(*c*) *If A is orthogonal, then* $\det(A)=1$ *or* $\det(A)=-1$.

THEOREM 7.1.3 *If A is an $n\times n$ matrix, then the following are equivalent.*

(*a*) *A is orthogonal.*

(*b*) $\|A\mathbf{x}\|=\|\mathbf{x}\|$ *for all* $\mathbf{x}$ *in* R^n.

(*c*) $A\mathbf{x}\cdot A\mathbf{y}=\mathbf{x}\cdot\mathbf{y}$ *for all* $\mathbf{x}$ *and* $\mathbf{y}$ *in* R^n.

Orthonormal Basis

THEOREM 7.1.4 *If S is an orthonormal basis for an n-dimensional inner product space V, and if*

$$(\mathbf{u})_S=(u_1,u_2,\dots,u_n) \quad \text{and} \quad (\mathbf{v})_S=(v_1,v_2,\dots,v_n)$$

then:

(*a*) $\|\mathbf{u}\|=\sqrt{u_1^2+u_2^2+\cdots+u_n^2}$

(*b*) $d(\mathbf{u},\mathbf{v})=\sqrt{(u_1-v_1)^2+(u_2-v_2)^2+\cdots+(u_n-v_n)^2}$

(*c*) $\langle\mathbf{u},\mathbf{v}\rangle=u_1v_1+u_2v_2+\cdots+u_nv_n$

THEOREM 7.1.5 *Let V be a finite-dimensional inner product space. If P is the transition matrix from one orthonormal basis for V to another orthonormal basis for V, then P is an orthogonal matrix.*

Orthogonal Diagonalization

DEFINITION 1 If A and B are square matrices, then we say that A and B are ***orthogonally similar*** if there is an orthogonal matrix P such that $P^TAP = B$.

If A is orthogonally similar to some diagonal matrix, say

$$P^TAP = D$$

then we say that A is ***orthogonally diagonalizable*** and that P ***orthogonally diagonalizes*** A.

THEOREM 7.2.1 *If A is an $n \times n$ matrix, then the following are equivalent.*

(a) A is orthogonally diagonalizable.

(b) A has an orthonormal set of n eigenvectors.

(c) A is symmetric.

Symmetric Matrices

THEOREM 7.2.2 *If A is a symmetric matrix, then:*

(a) The eigenvalues of A are all real numbers.

(b) Eigenvectors from different eigenspaces are orthogonal.

Orthogonally Diagonalizing an $n \times n$ Symmetric Matrix

Step 1. Find a basis for each eigenspace of A.

Step 2. Apply the Gram–Schmidt process to each of these bases to obtain an orthonormal basis for each eigenspace.

Step 3. Form the matrix P whose columns are the vectors constructed in Step 2. This matrix will orthogonally diagonalize A, and the eigenvalues on the diagonal of $D = P^TAP$ will be in the same order as their corresponding eigenvectors in P.

7.3 Quadratic Forms

$$[x_1 \quad x_2]\begin{bmatrix} a_1 & a_3 \\ a_3 & a_2 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \mathbf{x}^T A\mathbf{x}$$

$$[x_1 \quad x_2 \quad x_3]\begin{bmatrix} a_1 & a_4 & a_5 \\ a_4 & a_2 & a_6 \\ a_5 & a_6 & a_3 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \mathbf{x}^T A\mathbf{x}$$

$$2x^2 + 6xy - 5y^2 = \begin{bmatrix} x & y \end{bmatrix}\begin{bmatrix} 2 & 3 \\ 3 & -5 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix}$$

$$x_1^2 + 7x_2^2 - 3x_3^2 + 4x_1x_2 - 2x_1x_3 + 8x_2x_3 = [x_1 \quad x_2 \quad x_3]\begin{bmatrix} 1 & 2 & -1 \\ 2 & 7 & 4 \\ -1 & 4 & -3 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

Definite quadratic forms

DEFINITION 1 A quadratic form $\mathbf{x}^T A\mathbf{x}$ is said to be

positive definite if $\mathbf{x}^T A\mathbf{x} > 0$ for $\mathbf{x} \neq \mathbf{0}$

negative definite if $\mathbf{x}^T A\mathbf{x} < 0$ for $\mathbf{x} \neq \mathbf{0}$

indefinite if $\mathbf{x}^T A\mathbf{x}$ has both positive and negative values

THEOREM 7.3.2 *If A is a symmetric matrix, then:*

(a) $\mathbf{x}^T A\mathbf{x}$ *is positive definite if and only if all eigenvalues of A are positive.*

(b) $\mathbf{x}^T A\mathbf{x}$ *is negative definite if and only if all eigenvalues of A are negative.*

(c) $\mathbf{x}^T A\mathbf{x}$ *is indefinite if and only if A has at least one positive eigenvalue and at least one negative eigenvalue.*

Section 7.5 Hermitian, Unitary and Normal Matrices

DEFINITION 1 If A is a complex matrix, then the ***conjugate transpose*** of A, denoted by A^*, is defined by

$$A^* = \overline{A}^T \tag{1}$$

THEOREM 7.5.1 *If k is a complex scalar, and if A, B, and C are complex matrices whose sizes are such that the stated operations can be performed, then:*

(*a*) $(A^*)^* = A$

(*b*) $(A + B)^* = A^* + B^*$

(*c*) $(A - B)^* = A^* - B^*$

(*d*) $(kA)^* = \overline{k}A^*$

(*e*) $(AB)^* = B^*A^*$

EXAMPLE 1 Conjugate Transpose

Find the conjugate transpose A^* of the matrix

$$A = \begin{bmatrix} 1+i & -i & 0 \\ 2 & 3-2i & i \end{bmatrix}$$

Solution We have

$$\overline{A} = \begin{bmatrix} 1-i & i & 0 \\ 2 & 3+2i & -i \end{bmatrix} \text{ and hence } A^* = \overline{A}^T = \begin{bmatrix} 1-i & 2 \\ i & 3+2i \\ 0 & -i \end{bmatrix}$$

Hermitian Matrices

DEFINITION 2 A square complex matrix A is said to be ***unitary*** if

$$A^{-1} = A^* \tag{3}$$

and is said to be ***Hermitian**** if

$$A^* = A \tag{4}$$

THEOREM 7.5.2 *The eigenvalues of a Hermitian matrix are real numbers.*

THEOREM 7.5.3 *If A is a Hermitian matrix, then eigenvectors from different eigenspaces are orthogonal.*

EXAMPLE 2 Recognizing Hermitian Matrices

Hermitian matrices are easy to recognize because their diagonal entries are real (why?), and the entries that are symmetrically positioned across the main diagonal are complex conjugates. Thus, for example, we can tell by inspection that

$$A = \begin{bmatrix} 1 & i & 1+i \\ -i & -5 & 2-i \\ 1-i & 2+i & 3 \end{bmatrix}$$

Unitary Matrices

THEOREM 7.5.4 *If A is an $n \times n$ matrix with complex entries, then the following are equivalent.*

(a) *A is unitary.*

(b) $\|A\mathbf{x}\| = \|\mathbf{x}\|$ *for all* $\mathbf{x}$ *in* C^n.

(c) $A\mathbf{x} \cdot A\mathbf{y} = \mathbf{x} \cdot \mathbf{y}$ *for all* $\mathbf{x}$ *and* $\mathbf{y}$ *in* C^n.

(d) *The column vectors of A form an orthonormal set in C^n with respect to the complex Euclidean inner product.*

(e) *The row vectors of A form an orthonormal set in C^n with respect to the complex Euclidean inner product.*

DEFINITION 3 A square complex matrix is said to be ***unitarily diagonalizable*** if there is a unitary matrix P such that $P^*AP = D$ is a complex diagonal matrix. Any such matrix P is said to ***unitarily diagonalize*** A.

WEEK 12

Chapter 8 Linear Transformations

8.1 General Linear Transformations

8.2 Isomorphisms

8.3 Compositions and Inverse Transformations

8.4 Matrices for General Linear Transformations

General Linear Transformations

DEFINITION 1 If $T : V \to W$ is a function from a vector space V to a vector space W, then T is called a ***linear transformation*** from V to W if the following two properties hold for all vectors $\mathbf{u}$ and $\mathbf{v}$ in V and for all scalars k:

(i) $T(k\mathbf{u}) = kT(\mathbf{u})$ [Homogeneity property]

(ii) $T(\mathbf{u} + \mathbf{v}) = T(\mathbf{u}) + T(\mathbf{v})$ [Additivity property]

In the special case where $V = W$, the linear transformation T is called a ***linear operator*** on the vector space V.

THEOREM 8.1.1 *If $T : V \to W$ is a linear transformation, then:*

(a) $T(\mathbf{0}) = \mathbf{0}$.

(b) $T(\mathbf{u} - \mathbf{v}) = T(\mathbf{u}) - T(\mathbf{v})$ *for all* $\mathbf{u}$ *and* $\mathbf{v}$ *in* V.

Image, Kernel and Range

THEOREM 8.1.2 *Let $T : V \to W$ be a linear transformation, where V is finite dimensional. If $S = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ is a basis for V, then the image of any vector $\mathbf{v}$ in V can be expressed as*

$$T(\mathbf{v}) = c_1T(\mathbf{v}_1) + c_2T(\mathbf{v}_2) + \cdots + c_nT(\mathbf{v}_n) \tag{3}$$

where $c_1, c_2, \ldots, c_n$ are the coefficients required to express $\mathbf{v}$ as a linear combination of the vectors in S.

DEFINITION 2 If $T : V \to W$ is a linear transformation, then the set of vectors in V that T maps into $\mathbf{0}$ is called the ***kernel*** of T and is denoted by $\ker(T)$. The set of all vectors in W that are images under T of at least one vector in V is called the ***range*** of T and is denoted by $R(T)$.

THEOREM 8.1.3 *If $T : V \to W$ is a linear transformation, then:*

(a) *The kernel of T is a subspace of V.*

(b) *The range of T is a subspace of W.*

Rank, Nullity and Dimension

DEFINITION 3 Let $T: V \rightarrow W$ be a linear transformation. If the range of T is finite-dimensional, then its dimension is called the ***rank of T***; and if the kernel of T is finite-dimensional, then its dimension is called the ***nullity of T***. The rank of T is denoted by rank(T) and the nullity of T by nullity(T).

THEOREM 8.1.4 Dimension Theorem for Linear Transformations

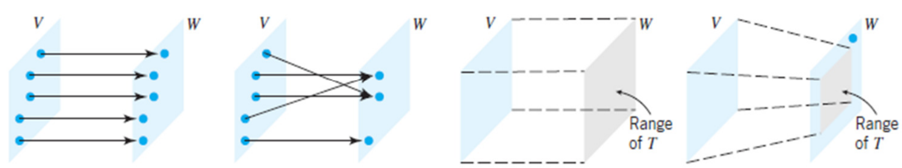
If $T: V \rightarrow W$ is a linear transformation from an n-dimensional vector space V to a vector space W, then

$$\text{rank}(T) + \text{nullity}(T) = n \quad (7)$$

Isomorphism

DEFINITION 1 If $T: V \rightarrow W$ is a linear transformation from a vector space V to a vector space W, then T is said to be ***one-to-one*** if T maps distinct vectors in V into distinct vectors in W.

DEFINITION 2 If $T: V \rightarrow W$ is a linear transformation from a vector space V to a vector space W, then T is said to be ***onto*** (or ***onto W***) if every vector in W is the image of at least one vector in V.



One-to-one. Distinct vectors in V have distinct images in W.

Not one-to-one. There exist distinct vectors in V with the same image.

Onto W. Every vector in W is the image of some vector in V.

Not onto W. Not every vector in W is the image of some vector in V.

THEOREM 8.2.1 *If $T: V \rightarrow W$ is a linear transformation, then the following statements are equivalent.*

(*a*) *T is one-to-one.*

(*b*) $\ker(T) = \{\mathbf{0}\}$.

THEOREM 8.2.2 *If V is a finite-dimensional vector space, and if $T: V \to V$ is a linear operator, then the following statements are equivalent.*

(a) *T is one-to-one.*

(b) $\ker(T) = \{\mathbf{0}\}$.

(c) *T is onto* [*i.e.*, $R(T) = V$].

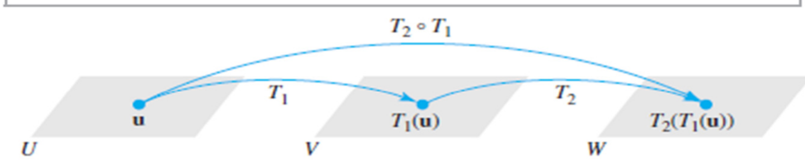
DEFINITION 3 If a linear transformation $T: V \to W$ is both one-to-one and onto, then T is said to be an ***isomorphism***, and the vector spaces V and W are said to be ***isomorphic***.

Compositions and Inverse Transformations

DEFINITION 1 If $T_1: U \to V$ and $T_2: V \to W$ are linear transformations, then the ***composition of T_2 with T_1***, denoted by $T_2 \circ T_1$ (which is read "T_2 circle T_1"), is the function defined by the formula

$$(T_2 \circ T_1)(\mathbf{u}) = T_2(T_1(\mathbf{u})) \tag{1}$$

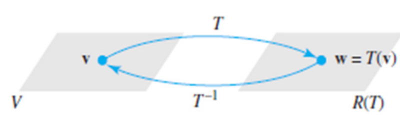
where $\mathbf{u}$ is a vector in U.



Inverses

$$T^{-1}(T(\mathrm{v})) = T^{-1}(\mathrm{w}) = \mathrm{v}$$

$$T(T^{-1}(\mathrm{w})) = T(\mathrm{v}) = \mathrm{w}$$



THEOREM 8.3.2 *If $T_1: U \to V$ and $T_2: V \to W$ are one-to-one linear transformations, then*

(a) *$T_2 \circ T_1$ is one-to-one.*

(b) $(T_2 \circ T_1)^{-1} = T_1^{-1} \circ T_2^{-1}$.

EXAMPLE 4 An Inverse Transformation

Let $T: R^3 \to R^3$ be the linear operator defined by the formula

$$T(x_1, x_2, x_3) = (3x_1 + x_2,\ -2x_1 - 4x_2 + 3x_3, 5x_1 + 4x_2 - 2x_3)$$

Determine whether T is one-to-one; if so, find $T^{-1}(x_1, x_2, x_3)$.

Solution It follows from Formula 12 of Section 4.9 that the standard matrix for T is

$$[T] = \begin{bmatrix} 3 & 1 & 0 \\ -2 & -4 & 3 \\ 5 & 4 & -2 \end{bmatrix}$$

(verify). This matrix is invertible, and from Formula 7 of Section 4.10 the standard matrix for T^{-1} is

$$[T^{-1}] = [T]^{-1} = \begin{bmatrix} 4 & -2 & -3 \\ -11 & 6 & 9 \\ -12 & 7 & 10 \end{bmatrix}$$

Expressing this result in horizontal notation yields

$$T^{-1}(x_1, x_2, x_3) = (4x_1 - 2x_2 - 3x_3,\ -11x_1 + 6x_2 + 9x_3,\ -12x_1 + 7x_2 + 10x_3)$$

A Procedure for Finding Standard Matrices

There is a way of finding the standard matrix for a matrix transformation from R^n to R^m by considering the effect of that transformation on the standard basis vectors for R^n. To explain the idea, suppose that A is unknown and that

$$\mathbf{e}_1,\ \mathbf{e}_2, \ldots,\ \mathbf{e}_n$$

are the standard basis vectors for R^n. Suppose also that the images of these vectors under the transformation T_A are

$$T_A(\mathbf{e}_1) = A\mathbf{e}_1,\ T_A(\mathbf{e}_2) = A\mathbf{e}_2, \ldots,\ T_A(\mathbf{e}_n) = A\mathbf{e}_n$$

It follows from Theorem 1.3.1 that $A\mathbf{e}_j$ is a linear combination of the columns of A in which the successive coefficients are the entries of $\mathbf{e}_j$. But all entries of $\mathbf{e}_j$ are zero except the jth, so the product $A\mathbf{e}_j$ is just the jth column of the matrix A. Thus,

$$A = [T_A(\mathbf{e}_1) | T_A(\mathbf{e}_2) | \cdots | T_A(\mathbf{e}_n)] \tag{12}$$

In summary, we have the following procedure for finding the standard matrix for a matrix transformation:

Finding the Standard Matrix for a Matrix Transformation

Step 1. Find the images of the standard basis vectors $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n$ for R^n in column form.

Step 2. Construct the matrix that has the images obtained in Step 1 as its successive columns. This matrix is the standard matrix for the transformation.

Matrices for General Linear Transformations

EXAMPLE 1 Matrix for a Linear Transformation

Let $T: P_1 \to P_2$ be the linear transformation defined by

$$T(p(x)) = xp(x)$$

Find the matrix for T with respect to the standard bases

$$B = \{\mathbf{u}_1, \mathbf{u}_2\} \quad \text{and} \quad B' = \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$$

where

$$\mathbf{u}_1 = 1, \quad \mathbf{u}_2 = x; \quad \mathbf{v}_1 = 1, \quad \mathbf{v}_2 = x, \quad \mathbf{v}_3 = x^2$$

Solution From the given formula for T we obtain

$$T(\mathbf{u}_1) = T(1) = (x)(1) = x$$
$$T(\mathbf{u}_2) = T(x) = (x)(x) = x^2$$

By inspection, the coordinate vectors for $T(\mathbf{u}_1)$ and $T(\mathbf{u}_2)$ relative to B' are

$$[T(\mathbf{u}_1)]_{B'} = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \quad [T(\mathbf{u}_2)]_{B'} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

Thus, the matrix for T with respect to B and B' is

$$[T]_{B',B} = [[T(\mathbf{u}_1)]_{B'} [T(\mathbf{u}_2)]_{B'}] = \begin{bmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 1 \end{bmatrix}$$

WEEK 13

Chapter 9 Numerical Methods

9.1 LU-Decompositions

9.2 The Power Method

9.5 Singular Value Decomposition

Section 9.1 LU-Decompositions

The Method of LU-Decomposition

Step 1. Rewrite the system $A\mathbf{x} = \mathbf{b}$ as

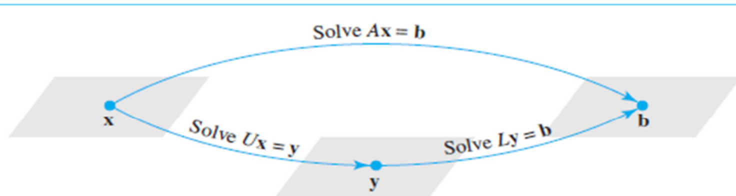
$$LU\mathbf{x} = \mathbf{b} \tag{2}$$

Step 2. Define a new $n \times 1$ matrix $\mathbf{y}$ by

$$U\mathbf{x} = \mathbf{y} \tag{3}$$

Step 3. Use (3) to rewrite (2) as $L\mathbf{y} = \mathbf{b}$ and solve this system for $\mathbf{y}$.

Step 4. Substitute $\mathbf{y}$ in (3) and solve for $\mathbf{x}$.



Finding LU-Decompositions

DEFINITION 1 A factorization of a square matrix A as $A = LU$, where L is lower triangular and U is upper triangular is called an ***LU-decomposition*** (or ***LU-factorization***) of A.

THEOREM 9.1.1 *If A is a square matrix that can be reduced to a row echelon form U by Gaussian elimination without row interchanges, then A can be factored as $A = LU$, where L is a lower triangular matrix.*

Constructing an LU-Decomposition

Procedure for Constructing an LU-Decomposition

Step 1. Reduce A to a row echelon form U by Gaussian elimination without row interchanges, keeping track of the multipliers used to introduce the leading 1's and the multipliers used to introduce the zeros below the leading 1's.

Step 2. In each position along the main diagonal of L, place the reciprocal of the multiplier that introduced the leading 1 in that position in U.

Step 3. In each position below the main diagonal of L, place the negative of the multiplier used to introduce the zero in that position in U.

Step 4. Form the decomposition $A = LU$.

EXAMPLE 2 An *LU*-Decomposition

Find an LU-decomposition of

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

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

Step 1 $\quad \frac{1}{2} \times \text{row } 1 \qquad E_1 = \begin{bmatrix} \frac{1}{2} & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \qquad E_1^{-1} = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$

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

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

Step 2 $\quad (3 \times \text{row } 1) + \text{row } 2 \qquad E_2 = \begin{bmatrix} 1 & 0 & 0 \\ 3 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \qquad E_2^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ -3 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$

$$\begin{bmatrix} 1 & 3 & 1 \\ 0 & 1 & 3 \\ 4 & 9 & 2 \end{bmatrix}$$

Step 3 $\quad (-4 \times \text{row } 1) + \text{row } 3 \qquad E_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -4 & 0 & 1 \end{bmatrix} \qquad E_3^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 4 & 0 & 1 \end{bmatrix}$

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

Step 4 $\quad (3 \times \text{row } 2) + \text{row } 3 \qquad E_4 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 3 & 1 \end{bmatrix} \qquad E_4^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -3 & 1 \end{bmatrix}$

$$\begin{bmatrix} 1 & 3 & 1 \\ 0 & 1 & 3 \\ 0 & 0 & 7 \end{bmatrix}$$

Step 5 $\quad \frac{1}{7} \times \text{row } 3 \qquad E_5 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \frac{1}{7} \end{bmatrix} \qquad E_5^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 7 \end{bmatrix}$

$$\begin{bmatrix} 1 & 3 & 1 \\ 0 & 1 & 3 \\ 0 & 0 & 1 \end{bmatrix} = U$$

$$\begin{aligned} L &= \begin{bmatrix} 2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} 1 & 0 & 0 \\ -3 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 4 & 0 & 1 \end{bmatrix}\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -3 & 1 \end{bmatrix}\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 7 \end{bmatrix} \\ &= \begin{bmatrix} 2 & 0 & 0 \\ -3 & 1 & 0 \\ 4 & -3 & 7 \end{bmatrix} \end{aligned}$$

so

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

is an LU-decomposition of A.

LU decomposition

The main idea of the LU decomposition is to record the steps used in Gaussian elimination on A in the places where the zero is produced. Consider the matrix

$$A = \begin{pmatrix} 1 & -2 & 3 \\ 2 & -5 & 12 \\ 0 & 2 & -10 \end{pmatrix}.$$

The first step of Gaussian elimination is to subtract 2 times the first row from the second row. In order to record what we have done, we will put the multiplier, 2, into the place it was used to make a zero, i.e. the second row, first column. In order to make it clear that it is a record of the step and not an element of A, we will put it in parentheses. This leads to

$$\begin{pmatrix} 1 & -2 & 3 \\ (2) & -1 & 6 \\ 0 & 2 & -10 \end{pmatrix}.$$

There is already a zero in the lower left corner, so we don't need to eliminate anything there. We record this fact with a (0). To eliminate the third row, second column, we need to subtract -2 times the second row from the third row. Recording the -2 in the spot it was used we have

$$\begin{pmatrix} 1 & -2 & 3 \\ (2) & -1 & 6 \\ (0) & (-2) & 2 \end{pmatrix}.$$

Let U be the upper triangular matrix produced, and let L be the lower triangular matrix with the records and ones on the diagonal, i.e.

$$L = \begin{pmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 0 & -2 & 1 \end{pmatrix} \quad \text{and} \quad U = \begin{pmatrix} 1 & -2 & 3 \\ 0 & -1 & 6 \\ 0 & 0 & 2 \end{pmatrix},$$

then we have the following wonderful property:

$$LU = \begin{pmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 0 & -2 & 1 \end{pmatrix} \begin{pmatrix} 1 & -2 & 3 \\ 0 & -1 & 6 \\ 0 & 0 & 2 \end{pmatrix} = \begin{pmatrix} 1 & -2 & 3 \\ 2 & -5 & 12 \\ 0 & 2 & -10 \end{pmatrix} = A.$$

Thus we see that A is actually the product of L and U. Here L is lower triangular and U is upper triangular. When a matrix can be written as a product of simpler matrices, we call that a *decomposition* of A and this one we call the LU decomposition.

Section 9.2 The Power Method

DEFINITION 1 If the *distinct* eigenvalues of a matrix A are $\lambda_1, \lambda_2, \ldots, \lambda_k$, and if $|\lambda_1|$ is larger than $|\lambda_2|, \ldots, |\lambda_k|$, then λ_1 is called a ***dominant eigenvalue*** of A. Any eigenvector corresponding to a dominant eigenvalue is called a ***dominant eigenvector*** of A.

THEOREM 9.2.1 *Let A be a symmetric $n \times n$ matrix with a positive* dominant eigenvalue λ. If $\mathbf{x}_0$ is a unit vector in R^n that is not orthogonal to the eigenspace corresponding to λ, then the normalized power sequence*

$$\mathbf{x}_0, \quad \mathbf{x}_1 = \frac{A\mathbf{x}_0}{\|A\mathbf{x}_0\|}, \quad \mathbf{x}_2 = \frac{A\mathbf{x}_1}{\|A\mathbf{x}_1\|}, \ldots, \quad \mathbf{x}_k = \frac{A\mathbf{x}_{k-1}}{\|A\mathbf{x}_{k-1}\|}, \ldots \tag{1}$$

converges to a unit dominant eigenvector, and the sequence

$$A\mathbf{x}_1 \cdot \mathbf{x}_1, \quad A\mathbf{x}_2 \cdot \mathbf{x}_2, \quad A\mathbf{x}_3 \cdot \mathbf{x}_3, \ldots, \quad A\mathbf{x}_k \cdot \mathbf{x}_k, \ldots \tag{2}$$

converges to the dominant eigenvalue λ.

The Power Method with Euclidean Scaling

The Power Method with Euclidean Scaling

Step 1. Choose an arbitrary nonzero vector and normalize it, if need be, to obtain a unit vector $\mathbf{x}_0$.

Step 2. Compute $A\mathbf{x}_0$ and normalize it to obtain the first approximation $\mathbf{x}_1$ to a dominant unit eigenvector. Compute $A\mathbf{x}_1 \cdot \mathbf{x}_1$ to obtain the first approximation to the dominant eigenvalue.

Step 3. Compute $A\mathbf{x}_1$ and normalize it to obtain the second approximation $\mathbf{x}_2$ to a dominant unit eigenvector. Compute $A\mathbf{x}_2 \cdot \mathbf{x}_2$ to obtain the second approximation to the dominant eigenvalue.

Step 4. Compute $A\mathbf{x}_2$ and normalize it to obtain the third approximation $\mathbf{x}_3$ to a dominant unit eigenvector. Compute $A\mathbf{x}_3 \cdot \mathbf{x}_3$ to obtain the third approximation to the dominant eigenvalue.

Continuing in this way will usually generate a sequence of better and better approximations to the dominant eigenvalue and a corresponding unit eigenvector.*

Positive Dominant Eigenvalue, λ

THEOREM 9.2.2 *Let A be a symmetric $n \times n$ matrix with a positive dominant* eigenvalue λ. If $\mathbf{x}_0$ is a nonzero vector in R^n that is not orthogonal to the eigenspace corresponding to λ, then the sequence*

$$\mathbf{x}_0, \quad \mathbf{x}_1 = \frac{A\mathbf{x}_0}{\max(A\mathbf{x}_0)}, \quad \mathbf{x}_2 = \frac{A\mathbf{x}_1}{\max(A\mathbf{x}_1)}, \ldots, \quad \mathbf{x}_k = \frac{A\mathbf{x}_{k-1}}{\max(A\mathbf{x}_{k-1})}, \ldots \tag{8}$$

converges to an eigenvector corresponding to λ, and the sequence

$$\frac{A\mathbf{x}_1 \cdot \mathbf{x}_1}{\mathbf{x}_1 \cdot \mathbf{x}_1}, \quad \frac{A\mathbf{x}_2 \cdot \mathbf{x}_2}{\mathbf{x}_2 \cdot \mathbf{x}_2}, \quad \frac{A\mathbf{x}_3 \cdot \mathbf{x}_3}{\mathbf{x}_3 \cdot \mathbf{x}_3}, \ldots, \quad \frac{A\mathbf{x}_k \cdot \mathbf{x}_k}{\mathbf{x}_k \cdot \mathbf{x}_k}, \ldots \tag{9}$$

converges to λ.

The Power Method

The Power Method with Maximum Entry Scaling

Step 1. Choose an arbitrary nonzero vector $\mathbf{x}_0$.

Step 2. Compute $A\mathbf{x}_0$ and multiply it by the factor $1/\max(A\mathbf{x}_0)$ to obtain the first approximation $\mathbf{x}_1$ to a dominant eigenvector. Compute the Rayleigh quotient of $\mathbf{x}_1$ to obtain the first approximation to the dominant eigenvalue.

Step 3. Compute $A\mathbf{x}_1$ and scale it by the factor $1/\max(A\mathbf{x}_1)$ to obtain the second approximation $\mathbf{x}_2$ to a dominant eigenvector. Compute the Rayleigh quotient of $\mathbf{x}_2$ to obtain the second approximation to the dominant eigenvalue.

Step 4. Compute $A\mathbf{x}_2$ and scale it by the factor $1/\max(A\mathbf{x}_2)$ to obtain the third approximation $\mathbf{x}_3$ to a dominant eigenvector. Compute the Rayleigh quotient of $\mathbf{x}_3$ to obtain the third approximation to the dominant eigenvalue.

Continuing in this way will generate a sequence of better and better approximations to the dominant eigenvalue and a corresponding eigenvector.

EXAMPLE 2 The Power Method with Euclidean Scaling

Apply the power method with Euclidean scaling to

$$A = \begin{bmatrix} 3 & 2 \\ 2 & 3 \end{bmatrix} \quad \text{with} \quad \mathbf{x}_0 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

Stop at $\mathbf{x}_5$ and compare the resulting approximations to the exact values of the dominant eigenvalue and eigenvector.

Solution We will leave it for you to show that the eigenvalues of A are $\lambda = 1$ and $\lambda = 5$ and that the eigenspace corresponding to the dominant eigenvalue $\lambda = 5$ is the line represented by the parametric equations $x_1 = t, x_2 = t$, which we can write in vector form as

$$\mathbf{x} = t\begin{bmatrix} 1 \\ 1 \end{bmatrix} \tag{6}$$

Setting $t = 1/\sqrt{2}$ yields the normalized dominant eigenvector

$$\mathbf{v}_1 = \begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix} \approx \begin{bmatrix} 0.707106781187\ldots \\ 0.707106781187\ldots \end{bmatrix} \tag{7}$$

Now let us see what happens when we use the power method, starting with the unit vector $\mathbf{x}_0$.

$$A\mathbf{x}_0 = \begin{bmatrix} 3 & 2 \\ 2 & 3 \end{bmatrix}\begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 3 \\ 2 \end{bmatrix} \qquad \mathbf{x}_1 = \frac{A\mathbf{x}_0}{\|A\mathbf{x}_0\|} = \frac{1}{\sqrt{13}}\begin{bmatrix} 3 \\ 2 \end{bmatrix} \approx \frac{1}{3.60555}\begin{bmatrix} 3 \\ 2 \end{bmatrix} \approx \begin{bmatrix} 0.83205 \\ 0.55470 \end{bmatrix}$$

$$A\mathbf{x}_1 \approx \begin{bmatrix} 3 & 2 \\ 2 & 3 \end{bmatrix}\begin{bmatrix} 0.83205 \\ 0.55470 \end{bmatrix} \approx \begin{bmatrix} 3.60555 \\ 3.32820 \end{bmatrix} \qquad \mathbf{x}_2 = \frac{A\mathbf{x}_1}{\|A\mathbf{x}_1\|} \approx \frac{1}{4.90682}\begin{bmatrix} 3.60555 \\ 3.32820 \end{bmatrix} \approx \begin{bmatrix} 0.73480 \\ 0.67828 \end{bmatrix}$$

$$A\mathbf{x}_2 \approx \begin{bmatrix} 3 & 2 \\ 2 & 3 \end{bmatrix}\begin{bmatrix} 0.73480 \\ 0.67828 \end{bmatrix} \approx \begin{bmatrix} 3.56097 \\ 3.50445 \end{bmatrix} \qquad \mathbf{x}_3 = \frac{A\mathbf{x}_2}{\|A\mathbf{x}_2\|} \approx \frac{1}{4.99616}\begin{bmatrix} 3.56097 \\ 3.50445 \end{bmatrix} \approx \begin{bmatrix} 0.71274 \\ 0.70143 \end{bmatrix}$$

$$A\mathbf{x}_3 \approx \begin{bmatrix} 3 & 2 \\ 2 & 3 \end{bmatrix}\begin{bmatrix} 0.71274 \\ 0.70143 \end{bmatrix} \approx \begin{bmatrix} 3.54108 \\ 3.52976 \end{bmatrix} \qquad \mathbf{x}_4 = \frac{A\mathbf{x}_3}{\|A\mathbf{x}_3\|} \approx \frac{1}{4.99985}\begin{bmatrix} 3.54108 \\ 3.52976 \end{bmatrix} \approx \begin{bmatrix} 0.70824 \\ 0.70597 \end{bmatrix}$$

$$A\mathbf{x}_4 \approx \begin{bmatrix} 3 & 2 \\ 2 & 3 \end{bmatrix}\begin{bmatrix} 0.70824 \\ 0.70597 \end{bmatrix} \approx \begin{bmatrix} 3.53666 \\ 3.53440 \end{bmatrix} \qquad \mathbf{x}_5 = \frac{A\mathbf{x}_4}{\|A\mathbf{x}_4\|} \approx \frac{1}{4.99999}\begin{bmatrix} 3.53666 \\ 3.53440 \end{bmatrix} \approx \begin{bmatrix} 0.70733 \\ 0.70688 \end{bmatrix}$$

$$\lambda^{(1)} = (A\mathbf{x}_1)\cdot\mathbf{x}_1 = (A\mathbf{x}_1)^T\mathbf{x}_1 \approx [3.60555 \quad 3.32820]\begin{bmatrix}0.83205\\0.55470\end{bmatrix} \approx 4.84615$$

$$\lambda^{(2)} = (A\mathbf{x}_2)\cdot\mathbf{x}_2 = (A\mathbf{x}_2)^T\mathbf{x}_2 \approx [3.56097 \quad 3.50445]\begin{bmatrix}0.73480\\0.67828\end{bmatrix} \approx 4.99361$$

$$\lambda^{(3)} = (A\mathbf{x}_3)\cdot\mathbf{x}_3 = (A\mathbf{x}_3)^T\mathbf{x}_3 \approx [3.54108 \quad 3.52976]\begin{bmatrix}0.71274\\0.70143\end{bmatrix} \approx 4.99974$$

$$\lambda^{(4)} = (A\mathbf{x}_4)\cdot\mathbf{x}_4 = (A\mathbf{x}_4)^T\mathbf{x}_4 \approx [3.53666 \quad 3.53440]\begin{bmatrix}0.70824\\0.70597\end{bmatrix} \approx 4.99999$$

$$\lambda^{(5)} = (A\mathbf{x}_5)\cdot\mathbf{x}_5 = (A\mathbf{x}_5)^T\mathbf{x}_5 \approx [3.53576 \quad 3.53531]\begin{bmatrix}0.70733\\0.70688\end{bmatrix} \approx 5.00000$$

Thus, $\lambda^{(5)}$ approximates the dominant eigenvalue to five decimal place accuracy and $\mathbf{x}_5$ approximates the dominant eigenvector in 7 correctly to three decimal place accuracy.

Section 9.5 Singular Value Decomposition

THEOREM 9.5.1 *If A is an $m \times n$ matrix, then:*

(*a*) *A and A^TA have the same null space.*

(*b*) *A and A^TA have the same row space.*

(*c*) *A^T and A^TA have the same column space.*

(*d*) *A and A^TA have the same rank.*

THEOREM 9.5.2 *If A is an $m \times n$ matrix, then:*

(*a*) *A^TA is orthogonally diagonalizable.*

(*b*) *The eigenvalues of A^TA are nonnegative.*

Singular Value Decomposition

DEFINITION 1 If A is an $m \times n$ matrix, and if $\lambda_1, \lambda_2, \ldots, \lambda_n$ are the eigenvalues of A^TA, then the numbers

$$\sigma_1 = \sqrt{\lambda_1}, \quad \sigma_2 = \sqrt{\lambda_2}, \ldots, \quad \sigma_n = \sqrt{\lambda_n}$$

are called the ***singular values*** of A.

EXAMPLE 1 Singular Values

Find the singular values of the matrix

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

Solution The first step is to find the eigenvalues of the matrix

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

The characteristic polynomial of $A^T A$ is

$$\lambda^2 - 4\lambda + 3 = (\lambda - 3)(\lambda - 1)$$

so the eigenvalues of $A^T A$ are $\lambda_1 = 3$ and $\lambda_2 = 1$ and the singular values of A in order of decreasing size are

$$\sigma_1 = \sqrt{\lambda_1} = \sqrt{3}, \quad \sigma_2 = \sqrt{\lambda_2} = 1$$