Linear Algebra

Exercises:

Chapter 1 – System of Linear Equations and Matrices

Exercise 1.1 – Introduction to System of Linear Equations

1. Multiply a row through by a nonzero constant.
2. Interchange two rows.
3. Add a constant times one row to another.

These are called ***elementary row operations*** on a matrix.

Exercise 1.2 – Gaussian Elimination

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

from which the solution $x = 1, y = 2, z = 3$ became evident. This is an example of a matrix that is in ***reduced row echelon form***. To be of this form, a matrix must have the following properties:

1. If a row does not consist entirely of zeros, then the first nonzero number in the row is a 1. We call this a ***leading 1***.
2. If there are any rows that consist entirely of zeros, then they are grouped together at the bottom of the matrix.
3. In any two successive rows that do not consist entirely of zeros, the leading 1 in the lower row occurs farther to the right than the leading 1 in the higher row.
4. Each column that contains a leading 1 has zeros everywhere else in that column.

A matrix that has the first three properties is said to be in ***row echelon form***. (Thus, a matrix in reduced row echelon form is of necessity in row echelon form, but not conversely.)

Theorem 1.2.1

Free Variable Theorem for Homogeneous Systems

If a homogeneous linear system has n unknowns, and if the reduced row echelon form of its augmented matrix has r nonzero rows, then the system has $n - r$ free variables.

Theorem 1.2.2

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

Although row echelon forms are not unique, the reduced row echelon form and all row echelon forms of a matrix A have the same number of zero rows, and the leading 1's always occur in the same positions. Those are called the ***pivot positions*** of A. The columns containing the leading 1's in a row echelon or reduced row echelon form of A are called the ***pivot columns*** of A, and the rows containing the leading 1's are called the ***pivot rows*** of A. A *nonzero* entry in a pivot position of A is called a ***pivot*** of A.

Exercise 1.3 – 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.

Definition 2

Two matrices are defined to be ***equal*** if they have the same size and their corresponding entries are equal.

Definition 3

If A and B are matrices of the same size, then the ***sum*** $A + B$ is the matrix obtained by adding the entries of B to the corresponding entries of A, and the ***difference*** $A - B$ is the matrix obtained by subtracting the entries of B from the corresponding entries of A. Matrices of different sizes cannot be added or subtracted.

Definition 4

If A is any matrix and c is any scalar, then the ***product*** cA is the matrix obtained by multiplying each entry of the matrix A by c. The matrix cA is said to be a ***scalar multiple*** of A.

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 the resulting products.

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

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

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

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

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

$$\boxed{j\text{th column vector of } AB = A[\,j\text{th column vector of } B]}$$

$$\boxed{i\text{th row vector of } AB = [i\text{th row vector of } A]B}$$

Definition 6

If $A_1, A_2, \dots, A_r$ are matrices of the same size, and if $c_1, c_2, \dots, c_r$ are scalars, then an expression of the form

$$c_1A_1 + c_2A_2 + \cdots + c_rA_r$$

is called a ***linear combination*** of $A_1, A_2, \dots, A_r$ with ***coefficients*** $c_1, c_2, \dots, c_r$.

Theorem 1.3.1

If A is an $m \times n$ matrix, and if $\mathbf{x}$ is an $n \times 1$ column vector, then the product $A\mathbf{x}$ can be expressed as a linear combination of the column vectors of A in which the coefficients are the entries of $\mathbf{x}$.

The matrix product

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

can be written as the following linear combination of column vectors:

$$2\begin{bmatrix} -1 \\ 1 \\ 2 \end{bmatrix} - 1\begin{bmatrix} 3 \\ 2 \\ 1 \end{bmatrix} + 3\begin{bmatrix} 2 \\ -3 \\ -2 \end{bmatrix} = \begin{bmatrix} 1 \\ -9 \\ -3 \end{bmatrix}$$

$$\boxed{AB = \mathbf{c}_1\mathbf{r}_1 + \mathbf{c}_2\mathbf{r}_2 + \cdots + \mathbf{c}_r\mathbf{r}_r} \tag{11}$$

We call (11) the ***column-row expansion*** of AB.

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.

Definition 8

If A is a square matrix, then the ***trace of A***, denoted by $\mathrm{tr}(A)$, 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.

Exercise 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 matrix addition]**

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

(c) $A(BC) = (AB)C$ **[Associative law for matrix 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) $a(BC) = (aB)C = B(aC)$

Theorem 1.4.3

If R is the reduced row echelon form of an $n \times n$ matrix A, then either R has at least one row of zeros or R is the identity matrix I_n.

Definition 1

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

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 $(kA)^{-1} = k^{-1}A^{-1}$.

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

Theorem 1.4.9

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

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

Exercise 1.5 - Elementary Matrices and a Method for Finding A^{-1}

Definition 1

Matrices A and B are said to be ***row equivalent*** if either (hence each) can be obtained from the other by a sequence of elementary row operations.

Definition 2

A matrix E is called an ***elementary matrix*** if it can be obtained from an identity matrix by performing a *single* elementary row operation.

Theorem 1.5.1

Row Operations by Matrix Multiplication

If the elementary matrix E results from performing a certain row operation on I_m and if A is an $m \times n$ matrix, then the product EA is the matrix that results when this same row operation is performed on A.

TABLE 1

Row Operation on I That Produces E	Row Operation on E That Reproduces I
Multiply row i by $c \neq 0$	Multiply row i by $1/c$
Interchange rows i and j	Interchange rows i and j
Add c times row i to row j	Add $-c$ times row i to row j

Theorem 1.5.2

Every elementary matrix is invertible, and the inverse is also an elementary matrix.

Theorem 1.5.3

Equivalent Statements

If A is an $n \times n$ matrix, then the following statements are equivalent, that is, all true or all false.

(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.

Inversion Algorithm To find the inverse of an invertible matrix A, find a sequence of elementary row operations that reduces A to the identity and then perform that same sequence of operations on I_n to obtain A^{-1}.

Exercise 1.6 - More on Linear Systems and Invertible Matrices

Theorem 1.6.1

A system of linear equations has zero, one, or infinitely many solutions. There are no other possibilities.

Theorem 1.6.2

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

Frequently, one is concerned with solving a sequence of systems

$$A\mathbf{x} = \mathbf{b}_1, \quad A\mathbf{x} = \mathbf{b}_2, \quad A\mathbf{x} = \mathbf{b}_3, \ldots, \quad A\mathbf{x} = \mathbf{b}_k$$

each of which has the same square coefficient matrix A. If A is invertible, then the solutions

$$\mathbf{x}_1 = A^{-1}\mathbf{b}_1, \quad \mathbf{x}_2 = A^{-1}\mathbf{b}_2, \quad \mathbf{x}_3 = A^{-1}\mathbf{b}_3, \ldots, \quad \mathbf{x}_k = A^{-1}\mathbf{b}_k$$

can be obtained with one matrix inversion and k matrix multiplications. An efficient way to do this is to form the partitioned matrix

$$[A \mid \mathbf{b}_1 \mid \mathbf{b}_2 \mid \cdots \mid \mathbf{b}_k] \tag{1}$$

Theorem 1.6.3

Let A be a square matrix.

(*a*) If B is a square matrix satisfying $BA = I$, then $B = A^{-1}$.

(*b*) If B is a square matrix satisfying $AB = I$, then $B = A^{-1}$.

Theorem 1.6.4

Equivalent Statements

If A is an $n \times n$ matrix, then the following 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 $\mathbf{b}$.

Theorem 1.6.5

Let A and B be square matrices of the same size. If AB is invertible, then A and B must also be invertible.

A Fundamental Problem Let A be a fixed $m \times n$ matrix. Find all $m \times 1$ matrices $\mathbf{b}$ such that the system of equations $A\mathbf{x} = \mathbf{b}$ is consistent.

Exercise 1.7 - Diagonal, Triangular, and Symmetric Matrices

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

A general 4×4 upper triangular matrix

A general 4×4 lower triangular matrix

$$\begin{bmatrix} d_1 & 0 & 0 \\ 0 & d_2 & 0 \\ 0 & 0 & d_3 \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \end{bmatrix} = \begin{bmatrix} d_1a_{11} & d_1a_{12} & d_1a_{13} & d_1a_{14} \\ d_2a_{21} & d_2a_{22} & d_2a_{23} & d_2a_{24} \\ d_3a_{31} & d_3a_{32} & d_3a_{33} & d_3a_{34} \end{bmatrix}$$

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \\ a_{41} & a_{42} & a_{43} \end{bmatrix} \begin{bmatrix} d_1 & 0 & 0 \\ 0 & d_2 & 0 \\ 0 & 0 & d_3 \end{bmatrix} = \begin{bmatrix} d_1a_{11} & d_2a_{12} & d_3a_{13} \\ d_1a_{21} & d_2a_{22} & d_3a_{23} \\ d_1a_{31} & d_2a_{32} & d_3a_{33} \\ d_1a_{41} & d_2a_{42} & d_3a_{43} \end{bmatrix}$$

In words, *to multiply a matrix A on the left by a diagonal matrix D, multiply successive rows of A by the successive diagonal entries of D, and to multiply A on the right by D, multiply successive columns of A by the successive diagonal entries of D.*

Theorem 1.7.1

(*a*) The transpose of a lower triangular matrix is upper triangular, and the transpose of an upper triangular matrix is lower triangular.

(*b*) The product of lower triangular matrices is lower triangular, and the product of upper triangular matrices is upper triangular.

(*c*) A triangular matrix is invertible if and only if its diagonal entries are all nonzero.

(*d*) The inverse of an invertible lower triangular matrix is lower triangular, and the inverse of an invertible upper triangular matrix is upper triangular.

Definition 1

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

Theorem 1.7.2

If A and B are symmetric matrices with the same size, and if k is any scalar, then:

(*a*) A^T is symmetric.

(*b*) $A + B$ and $A - B$ are symmetric.

(*c*) kA is symmetric.

Theorem 1.7.3

The product of two symmetric matrices is symmetric if and only if the matrices commute.

Theorem 1.7.4

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

Theorem 1.7.5

If A is an invertible matrix, then AA^T and A^TA are also invertible.

Exercise 1.8 - Introduction to Linear Transformations

Definition 1

If T is a function with domain R^n and codomain R^m, then we say that T is a ***transformation*** from R^n to R^m or that T ***maps*** from R^n to R^m, which we denote by writing

$$T : R^n \to R^m$$

In the special case where $m = n$, a transformation is sometimes called an ***operator*** on R^n.

Theorem 1.8.1

For every matrix A the matrix transformation $T_A : R^n \to R^m$ has the following properties for all vectors $\mathbf{u}$ and $\mathbf{v}$ 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) $T_A(\mathbf{u}-\mathbf{v}) = T_A(\mathbf{u}) - T_A(\mathbf{v})$

Theorem 1.8.2

$T : R^n \to R^m$ 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]**

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.

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.

TABLE 1

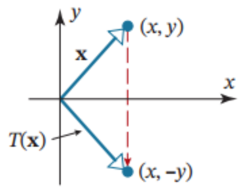
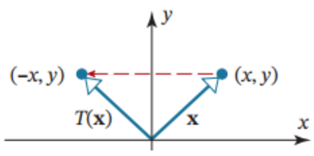
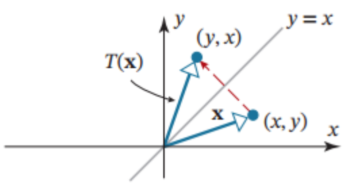
Operator	Illustration	Images of $\mathbf{e}_1$ and $\mathbf{e}_2$	Standard Matrix
Reflection about the x-axis $T(x,y) = (x, -y)$		$T(\mathbf{e}_1) = T(1,0) = (1,0)$ $T(\mathbf{e}_2) = T(0,1) = (0,-1)$	$\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$
Reflection about the y-axis $T(x,y) = (-x,y)$		$T(\mathbf{e}_1) = T(1,0) = (-1,0)$ $T(\mathbf{e}_2) = T(0,1) = (0,1)$	$\begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}$
Reflection about the line $y = x$ $T(x,y) = (y,x)$		$T(\mathbf{e}_1) = T(1,0) = (0,1)$ $T(\mathbf{e}_2) = T(0,1) = (1,0)$	$\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$

TABLE 2

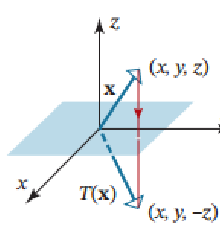
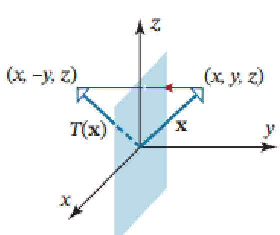
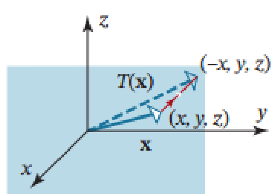
Operator	Illustration	Images of $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$	Standard Matrix
Reflection about the xy-plane $T(x,y,z) = (x,y,-z)$		$T(\mathbf{e}_1) = T(1,0,0) = (1,0,0)$ $T(\mathbf{e}_2) = T(0,1,0) = (0,1,0)$ $T(\mathbf{e}_3) = T(0,0,1) = (0,0,-1)$	$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{bmatrix}$
Reflection about the xz-plane $T(x,y,z) = (x,-y,z)$		$T(\mathbf{e}_1) = T(1,0,0) = (1,0,0)$ $T(\mathbf{e}_2) = T(0,1,0) = (0,-1,0)$ $T(\mathbf{e}_3) = T(0,0,1) = (0,0,1)$	$\begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$
Reflection about the yz-plane $T(x,y,z) = (-x,y,z)$		$T(\mathbf{e}_1) = T(1,0,0) = (-1,0,0)$ $T(\mathbf{e}_2) = T(0,1,0) = (0,1,0)$ $T(\mathbf{e}_3) = T(0,0,1) = (0,0,1)$	$\begin{bmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$

TABLE 3

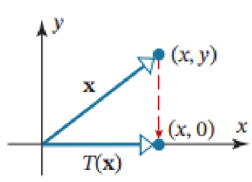
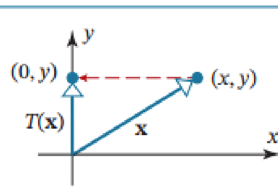
Operator	Illustration	Images of $\mathbf{e}_1$ and $\mathbf{e}_2$	Standard Matrix
Orthogonal projection onto the x-axis $T(x,y) = (x,0)$		$T(\mathbf{e}_1) = T(1,0) = (1,0)$ $T(\mathbf{e}_2) = T(0,1) = (0,0)$	$\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$
Orthogonal projection onto the y-axis $T(x,y) = (0,y)$		$T(\mathbf{e}_1) = T(1,0) = (0,0)$ $T(\mathbf{e}_2) = T(0,1) = (0,1)$	$\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$

TABLE 4

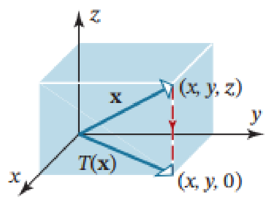
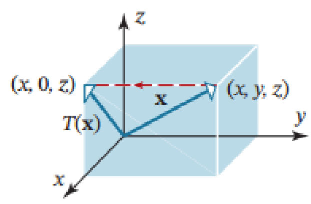
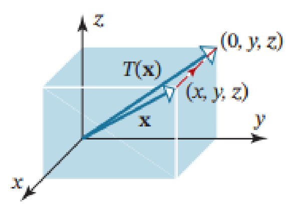
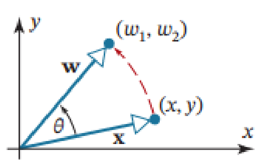
Operator	Illustration	Images of $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$	Standard Matrix
Orthogonal projection onto the xy-plane $T(x, y, z) = (x, y, 0)$		$T(\mathbf{e}_1) = T(1,0,0) = (1,0,0)$ $T(\mathbf{e}_2) = T(0,1,0) = (0,1,0)$ $T(\mathbf{e}_3) = T(0,0,1) = (0,0,0)$	$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$
Orthogonal projection onto the xz-plane $T(x, y, z) = (x, 0, z)$		$T(\mathbf{e}_1) = T(1,0,0) = (1,0,0)$ $T(\mathbf{e}_2) = T(0,1,0) = (0,0,0)$ $T(\mathbf{e}_3) = T(0,0,1) = (0,0,1)$	$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$
Orthogonal projection onto the yz-plane $T(x, y, z) = (0, y, z)$		$T(\mathbf{e}_1) = T(1,0,0) = (0,0,0)$ $T(\mathbf{e}_2) = T(0,1,0) = (0,1,0)$ $T(\mathbf{e}_3) = T(0,0,1) = (0,0,1)$	$\begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$

TABLE 5

Operator	Illustration	Images of $\mathbf{e}_1$ and $\mathbf{e}_2$	Standard Matrix
Counterclockwise rotation about the origin through an angle θ		$T(\mathbf{e}_1) = T(1,0) = (\cos\theta, \sin\theta)$ $T(\mathbf{e}_2) = T(0,1) = (-\sin\theta, \cos\theta)$	$\begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$

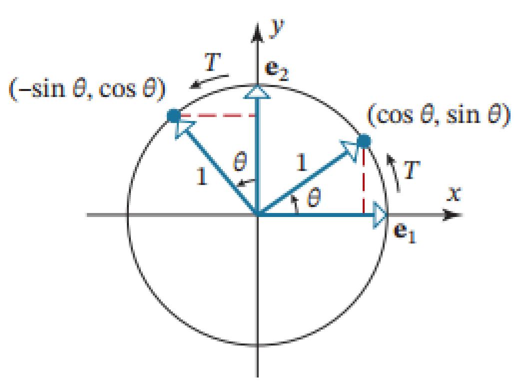


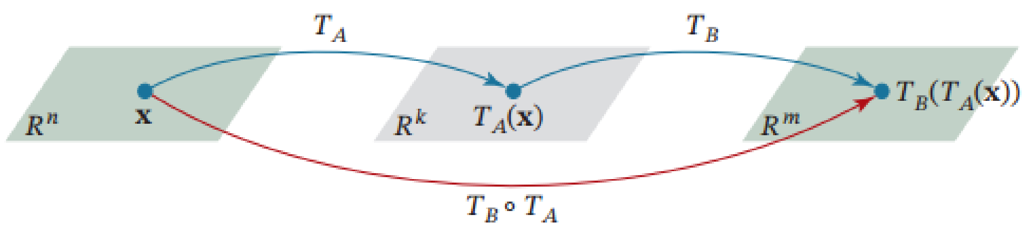
FIGURE 1.8.5

Exercise 1.9 - Compositions of Matrix Transformations

$$T_B \circ T_A$$

which is read "T_B circle T_A." As illustrated in **Figure 1.9.1**, the transformation T_A in the formula is performed first; that is,

$$\boxed{(T_B \circ T_A)(\mathbf{x}) = T_B(T_A(\mathbf{x}))} \tag{1}$$



Theorem 1.9.1

If $T_A: R^n \to R^k$ and $T_B: R^k \to R^m$ are matrix transformations, then $T_B \circ T_A$ is also a matrix transformation and

$$T_B \circ T_A = T_{BA} \tag{2}$$

Since it is *not* generally true that $AB = BA$, it is also *not* generally true that $T_{AB} = T_{BA}$, so in general

$$T_A \circ T_B \neq T_B \circ T_A$$

Thus, *composition of matrix transformations is not commutative.* In those special cases where equality holds, we say that T_A and T_B ***commute***. Note, for example, that the linear transformations in Example 1 do not commute, since $AB \neq BA$.

Exercise 1.10 - Applications of Linear Systems

Network Analysis

The concept of a *network* appears in a variety of applications. Loosely stated, a ***network*** is a set of ***branches*** through which something "flows." For example, the branches might be electrical wires through which electricity flows, pipes through which water or oil flows, traffic lanes through which vehicular traffic flows, or economic linkages through which money flows, to name a few possibilities.

In most networks, the branches meet at points, called ***nodes*** or ***junctions***, where the flow divides. For example, in an electrical network, nodes occur where three or more wires join, in a traffic network they occur at street intersections, and in a financial network they occur at banking centers where incoming money is distributed to individuals or other institutions.

In the study of networks, there is generally some numerical measure of the rate at which the medium flows through a branch. For example, the flow rate of electricity is often measured in amperes, the flow rate of water or oil in gallons per minute, the flow rate of traffic in vehicles per hour, and the flow rate of European currency in millions of Euros per day. We will restrict our attention to networks in which there is ***flow conservation*** at each node, by which we mean that *the rate of flow into any node is equal to the rate of flow out of that node.* This ensures that the flow medium does not build up at the nodes and block the free movement of the medium through the network.

Chapter 2 – Determinants

Exercise 2.1 – Determinants by Cofactor Expansion

Definition 1

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

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,

$$\det(A) = a_{1j}C_{1j} + a_{2j}C_{2j} + \cdots + a_{nj}C_{nj} \tag{7}$$

[cofactor expansion along the *j*th column]

and

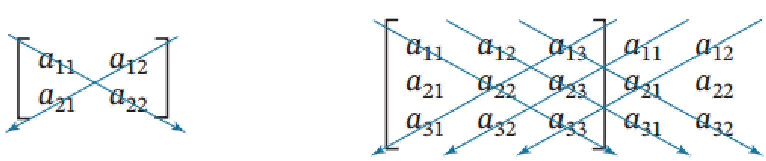
$$\det(A) = a_{i1}C_{i1} + a_{i2}C_{i2} + \cdots + a_{in}C_{in} \tag{8}$$

[cofactor expansion along the *i*th row]

Theorem 2.1.2

If A is an $n \times n$ triangular matrix (upper triangular, lower triangular, or diagonal), then $\det(A)$ is the product of the entries on the main diagonal of the matrix; that is, $\det(A) = a_{11}a_{22}\cdots a_{nn}$.

Determinants of 2×2 and 3×3 matrices can be evaluated very efficiently using the pattern suggested in **Figure 2.1.1**.



Exercise 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 or when a multiple of one column is added to another, then $\det(B) = \det(A)$.

Theorem 2.2.4

Let E be an $n \times n$ elementary matrix.

(a) If E results from multiplying a row of I_n by a nonzero number k, then $\det(E) = k$.

(b) If E results from interchanging two rows of I_n, then $\det(E) = -1$.

(c) If E results from adding a multiple of one row of I_n to another, then $\det(E) = 1$.

Theorem 2.2.5

If A is a square matrix with two proportional rows or two proportional columns, then $\det(A) = 0$.

$$\boxed{\det(kA) = k^n \det(A)} \tag{1}$$

For example,

$$\begin{vmatrix} ka_{11} & ka_{12} & ka_{13} \\ ka_{21} & ka_{22} & ka_{23} \\ ka_{31} & ka_{32} & ka_{33} \end{vmatrix} = k^3 \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}$$

Unfortunately, no simple relationship exists among $\det(A)$, $\det(B)$, and $\det(A+B)$. In particular, $\det(A+B)$ will usually *not* be equal to $\det(A) + \det(B)$. The following example illustrates this fact.

Exercise 2.3 – Properties of Determinants; Cramer's Rule

Theorem 2.3.1

Let A, B, and C be $n \times n$ matrices that differ only in a single row, say the rth, and assume that the rth row of C can be obtained by adding corresponding entries in the rth rows of A and B. Then

$$\det(C) = \det(A) + \det(B)$$

The same result holds for columns.

Lemma 2.3.2

If B is an $n \times n$ matrix and E is an $n \times n$ elementary matrix, then

$$\det(EB) = \det(E)\det(B)$$

Theorem 2.3.3

A square matrix A is invertible if and only if $\det(A) \neq 0$.

Theorem 2.3.4

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

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

Theorem 2.3.5

If A is invertible, then

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

Definition 1

If A is any $n \times n$ matrix and C_{ij} is the cofactor of a_{ij}, then the matrix

$$\begin{bmatrix} C_{11} & C_{12} & \cdots & C_{1n} \\ C_{21} & C_{22} & \cdots & C_{2n} \\ \vdots & \vdots & & \vdots \\ C_{n1} & C_{n2} & \cdots & C_{nn} \end{bmatrix}$$

is called the ***matrix of cofactors from A***. The transpose of this matrix is called the ***adjoint of A*** and is denoted by $\operatorname{adj}(A)$.

Theorem 2.3.6

Inverse of a Matrix Using Its Adjoint

If A is an invertible matrix, then

$$A^{-1} = \frac{1}{\det(A)}\operatorname{adj}(A) \tag{6}$$

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

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

Chapter 4 – General Vector Spaces

Exercise 4.1 – Real Vector Spaces

Definition 1

Let V be an arbitrary nonempty set of objects for which two operations are defined: addition and multiplication by numbers called ***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 exists an object in V, called the ***zero vector***, that is denoted by $\mathbf{0}$ and has the property 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. $1\mathbf{u}=\mathbf{u}$

Exercise 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

Subspace Test

If W is a nonempty set of vectors in a vector space V, then W is a subspace of V if and only if the following conditions are satisfied.

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

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

Theorem 4.2.2

If $W_1, W_2, \ldots, W_r$ are subspaces of a vector space V, then the intersection of these subspaces is also a subspace of V.

Theorem 4.2.3

The solution set of a homogeneous system $A\mathbf{x} = \mathbf{0}$ of m equations in n unknowns is a subspace of R^n.

Because the solution set of a homogeneous system in n unknowns is actually a subspace of R^n, we will generally refer to it as the ***solution space*** of the system.

Remark Whereas the solution set of every *homogeneous* system of m equations in n unknowns is a subspace of R^n, it is *never* true that the solution set of a *nonhomogeneous* system of m equations in n unknowns is a subspace of R^n. There are two possible scenarios: first, the system may not have any solutions at all, and second, if there are solutions, then the solution set will not be closed under either addition or scalar multiplication (Exercise 22).

Theorem 4.2.4

If A is an $m \times n$ matrix, then the kernel of the matrix transformation $T_A : R^n \rightarrow R^m$ is a subspace of R^n.

Exercise 4.3 – Spanning Sets

Definition 1

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

$$\mathbf{w} = k_1\mathbf{v}_1 + k_2\mathbf{v}_2 + \cdots + k_r\mathbf{v}_r \tag{1}$$

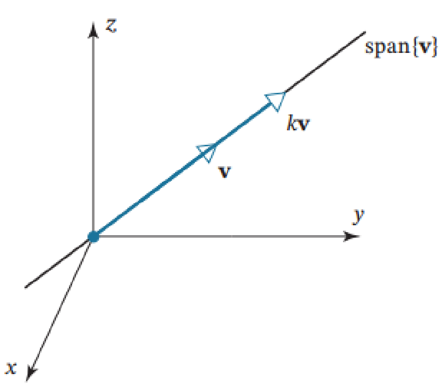
where $k_1, k_2, \ldots, k_r$ are scalars. These scalars are called the ***coefficients*** of the linear combination.

Theorem 4.3.1

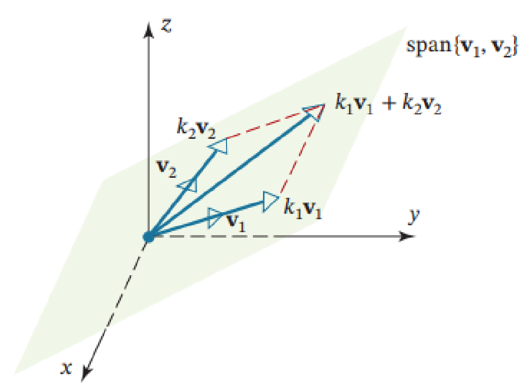
If $S = \{\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_r\}$ is a nonempty set of vectors in a vector space V, then:

(*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.



(*a*) span{$\mathbf{v}$} is the line through the origin determined by $\mathbf{v}$



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

The polynomials $1, x, x^2, \ldots, x^n$ span the vector space P_n defined in Example 10 since each polynomial $\mathbf{p}$ in P_n can be written as

$$\mathbf{p} = a_0 + a_1 x + \cdots + a_n x^n$$

which is a linear combination of $1, x, x^2, \ldots, x^n$. We can denote this by writing

$$P_n = \text{span}\{1, x, x^2, \ldots, x^n\}$$

A Procedure for Identifying Spanning Sets

Step 1. Let $S = \{\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_r\}$ be a given set of vectors in V, and let $\mathbf{x}$ be an arbitrary vector in V.

Step 2. Set up the augmented matrix for the linear system that results by equating corresponding components on the two sides of the vector equation

$$k_1\mathbf{w}_1 + k_2\mathbf{w}_2 + \cdots + k_r\mathbf{w}_r = \mathbf{x} \tag{2}$$

Step 3. Use the techniques developed in Chapters 1 and 2 to investigate the consistency or inconsistency of that system. If it is consistent for *all* choices of $\mathbf{x}$, the vectors in S span V, and if it is inconsistent for *some* vector $\mathbf{x}$, they do not.

Theorem 4.3.2

If $S = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_r\}$ and $S' = \{\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_k\}$ are nonempty sets of vectors in a vector space V, then

$$\text{span}\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_r\} = \text{span}\{\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_k\}$$

if and only if each vector in S is a linear combination of those in S', and each vector in S' is a linear combination of those in S.

Exercise 4.4 – Linear Independence

Definition 1

If $S = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_r\}$ is a set of two or more vectors in a vector space V, then S is said to be a ***linearly independent set*** if no vector in S can be expressed as a linear combination of the others. A set that is not linearly independent is said to be ***linearly dependent***. If S has only one vector, we will agree that it is linearly independent if and only if that vector is nonzero.

Theorem 4.4.1

A nonempty set $S = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_r\}$ in a vector space V is linearly independent if and only if the only coefficients satisfying the vector equation

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

are $k_1 = 0, k_2 = 0, \ldots, k_r = 0$.

Theorem 4.4.2

(a) A set with finitely many vectors that contains $\mathbf{0}$ is linearly dependent.

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

Theorem 4.4.3

Let $S = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_r\}$ be a set of vectors in R^n. If $r > n$, then S is linearly dependent.

Exercise 4.5 – Coordinate and Basis

Definition 1

If $S = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ is a set of vectors in a finite-dimensional vector space V, then S is called a ***basis*** for V if:

(a) S spans V.

(b) S is linearly independent.

fall into two categories: A vector space V is said to be ***finite-dimensional*** if there is a finite set of vectors in V that spans V and is said to be ***infinite-dimensional*** if no such set exists.

Theorem 4.5.1

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.

Exercise 4.6 – Dimensions

Theorem 4.6.1

All bases for a finite-dimensional vector space have the same number of vectors.

Theorem 4.6.2

Let V be a finite-dimensional vector space, and let $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ be any basis for V.

(a) If a set in V has more than n vectors, then it is linearly dependent.

(b) If a set in V has fewer than n vectors, then it does not span V.

We can now see rather easily why Theorem 4.6.1 is true; for if

$$S = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$$

is an *arbitrary* basis for V, then the linear independence of S implies that any set in V with more than n vectors is linearly dependent and any set in V with fewer than n vectors does not span V. Thus, unless a set in V has exactly n vectors it cannot be a basis.

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.

EXAMPLE 1 | Dimensions of Some Familiar Vector Spaces

$\dim(R^n) = n$ **[The standard basis has n vectors.]**

$\dim(P_n) = n + 1$ **[The standard basis has $n + 1$ vectors.]**

$\dim(M_{mn}) = mn$ **[The standard basis has mn vectors.]**

$$\dim\left[\operatorname{span}\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_r\}\right] = r$$

In words, the dimension of the space spanned by a linearly independent set of vectors is equal to the number of vectors in that set.

Theorem 4.6.3

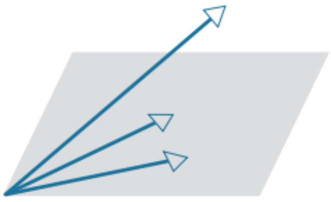
Plus/Minus Theorem

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 $\operatorname{span}(S)$, 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,

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



The vector outside the plane can be adjoined to the other two without affecting their linear independence.

Any of the vectors can be removed, and the remaining two will still span the plane.

Either of the collinear vectors can be removed, and the remaining two will still span the plane.

Theorem 4.6.4

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 4.6.5

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 4.6.6

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 $\dim(W) = \dim(V)$.

***Proof* (b)** Part (*a*) tells us that W is finite-dimensional, so it has a basis

$$S = \{\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_m\}$$

Either S is also a basis for V or it is not. If it is a basis, then $\dim(V) = m$, which means that $\dim(V) = \dim(W)$. If not, then because S is a linearly independent set it can be enlarged to a basis for V by part (*b*) of Theorem 4.6.5. But this implies that $\dim(W) < \dim(V)$, so we have shown that $\dim(W) \leq \dim(V)$ in all cases.

Figure 4.6.2 illustrates the geometric relationship between the subspaces of R^3 in order of increasing dimension.

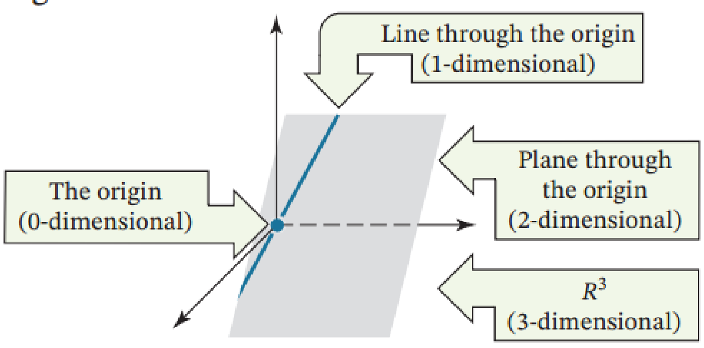


FIGURE 4.6.2

Exercise 4.8 – 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

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

in R^n 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.

Definition 2

If A is an $m \times n$ matrix, then the subspace of R^n spanned by the row vectors of A is denoted by row(A) and is called the ***row space*** of A, and the subspace of R^m spanned by the column vectors of A is denoted by col(A) and 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 denoted by null(A) and is called the ***null space*** of A.

Thus, a linear system, $A\mathbf{x} = \mathbf{b}$, of m equations in n unknowns can be written as

$$x_1\mathbf{c}_1 + x_2\mathbf{c}_2 + \cdots + x_n\mathbf{c}_n = \mathbf{b} \tag{2}$$

from which we conclude that $A\mathbf{x} = \mathbf{b}$ is consistent if and only if $\mathbf{b}$ is expressible as a linear combination of the column vectors of A. This yields the following theorem.

Theorem 4.8.1

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 4.8.2

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

$$\mathbf{x} = \mathbf{x}_0 + c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_k\mathbf{v}_k \tag{6}$$

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

Theorem 4.8.3

(*a*) *Row equivalent matrices have the same row space.*

(*b*) *Row equivalent matrices have the same null space.*

Theorem 4.8.4

If a matrix R is in row echelon form, then the row vectors with the leading 1's (the nonzero row vectors) form a basis for the row space of R, and the column vectors with the leading 1's of the row vectors form a basis for the column space of R.

Theorem 4.8.5

If A and B are row equivalent matrices, then:

(*a*) A given set of column vectors of A is linearly independent if and only if the corresponding column vectors of B are linearly independent.

(*b*) A given set of column vectors of A forms a basis for the column space of A if and only if the corresponding column vectors of B form a basis for the column space of B.

Find a basis for the row space of

$$A = \begin{bmatrix} 1 & -2 & 0 & 0 & 3 \\ 2 & -5 & -3 & -2 & 6 \\ 0 & 5 & 15 & 10 & 0 \\ 2 & 6 & 18 & 8 & 6 \end{bmatrix}$$

consisting entirely of row vectors from A.

Solution We will transpose A, thereby converting the row space of A into the column space of A^T; then we will use the method of Example 5 to find a basis for the column space of A^T; and then we will transpose again to convert column vectors back to row vectors.

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 each vector that is not in that basis as a linear combination of the basis vectors.

Basis for the Space Spanned by a Set of Vectors

Step 1. Form the matrix A whose columns are the vectors in the set $S = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\}$.

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).

This completes the first part of the problem.

Step 5. Obtain a set of dependency equations for the column vectors $\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_k$ of R by successively expressing each $\mathbf{w}_i$ that does not contain a leading 1 of R as a linear combination of predecessors that do.

Step 6. In each dependency equation obtained in Step 5, replace the vector $\mathbf{w}_i$ by the vector $\mathbf{v}_i$ for $i = 1, 2, \ldots, k$.

This completes the second part of the problem.

Exercise 4.9 – Rank, Nullity, and the Fundamental Matrix Spaces

Theorem 4.9.1

The row space and the column space of a matrix A have the same dimension.

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 $\text{rank}(A)$; the dimension of the null space of A is called the ***nullity*** of A and is denoted by $\text{nullity}(A)$.

EXAMPLE 2 | Maximum Value for Rank

What is the maximum possible rank of an $m \times n$ matrix A that is not square?

Solution Since the row vectors of A lie in R^n and the column vectors in R^m, the row space of A is at most n-dimensional and the column space is at most m-dimensional. Since the rank of A is the common dimension of its row and column space, it follows that the rank is at most the smaller of m and n. We denote this by writing

$$\text{rank}(A) \leq \min(m, n)$$

in which $\min(m, n)$ is the minimum of m and n.

Theorem 4.9.2

Dimension Theorem for Matrices

If A is a matrix with n columns, then

$$\text{rank}(A) + \text{nullity}(A) = n \tag{4}$$

Theorem 4.9.3

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

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

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

Theorem 4.9.4

If $A\mathbf{x} = \mathbf{b}$ is a consistent linear system of m equations in n unknowns, and if A has rank r, then the general solution of the system contains $n - r$ parameters.

row space of A	column space of A
null space of A	null space of A^T

These are called the ***fundamental spaces*** of the matrix A. The row space and null space of A are subspaces of R^n, whereas the column space of A and the null space of A^T are subspaces of R^m. The null space of A^T is also called the ***left null space of A*** because transposing both sides of the equation $A^T\mathbf{x} = \mathbf{0}$ produces the equation $\mathbf{x}^T A = \mathbf{0}^T$ in which the unknown is on the left. The dimension of the left null space of A is called the ***left nullity of A***. We will now consider how the four fundamental spaces are related.

Theorem 4.9.5

If A is any matrix, then $\text{rank}(A) = \text{rank}(A^T)$.

$$\boxed{\text{rank}(A) + \text{nullity}(A^T) = m} \tag{5}$$

This alternative form of Formula (4) makes it possible to express the dimensions of all four fundamental spaces in terms of the size and rank of A. Specifically, if $\text{rank}(A) = r$, then

$$\boxed{\begin{array}{ll} \dim[\text{row}(A)] = r & \dim[\text{col}(A)] = r \\ \dim[\text{null}(A)] = n - r & \dim[\text{null}(A^T)] = m - r \end{array}} \tag{6}$$

Bases for the Fundamental Spaces

An efficient way to obtain bases for the four fundamental spaces of an $m \times n$ matrix A is to adjoin the $m \times m$ identity matrix to A to obtain an augmented matrix $[A \mid I]$ and apply elementary row operations to this matrix to put A in reduced row echelon form R, thereby putting the augmented matrix in the form $[R \mid E]$. In the case where A is invertible the matrix E will be A^{-1}, but in general it will not. The rank r of A can then be obtained by counting the number of pivots (leading l's) in R, and the nullity of A^T can be obtained from the relationship

$$\text{nullity}(A^T) = m - r \tag{7}$$

that follows from Formula (5). Bases for three of the fundamental spaces can be obtained directly from $[R \mid E]$ as follows:

- A basis for $\text{row}(A)$ will be the r rows of R that contain the leading 1's (the pivot rows).
- A basis for $\text{col}(A)$ will be the r columns of A that contain the leading 1's of R (the pivot columns).
- A basis for $\text{null}(A^T)$ will be the bottom $m - r$ rows of E (see the proof at the end of this section)

Definition 2

If W is a subspace of R^n, then the set of all vectors in R^n that are orthogonal to every vector in W is called the ***orthogonal complement*** of W and is denoted by the symbol $W^\perp$.

Theorem 4.9.6

If W is a subspace of R^n, then:

(*a*) $W^\perp$ is a subspace of R^n.

(*b*) The only vector common to W and $W^\perp$ is $\mathbf{0}$.

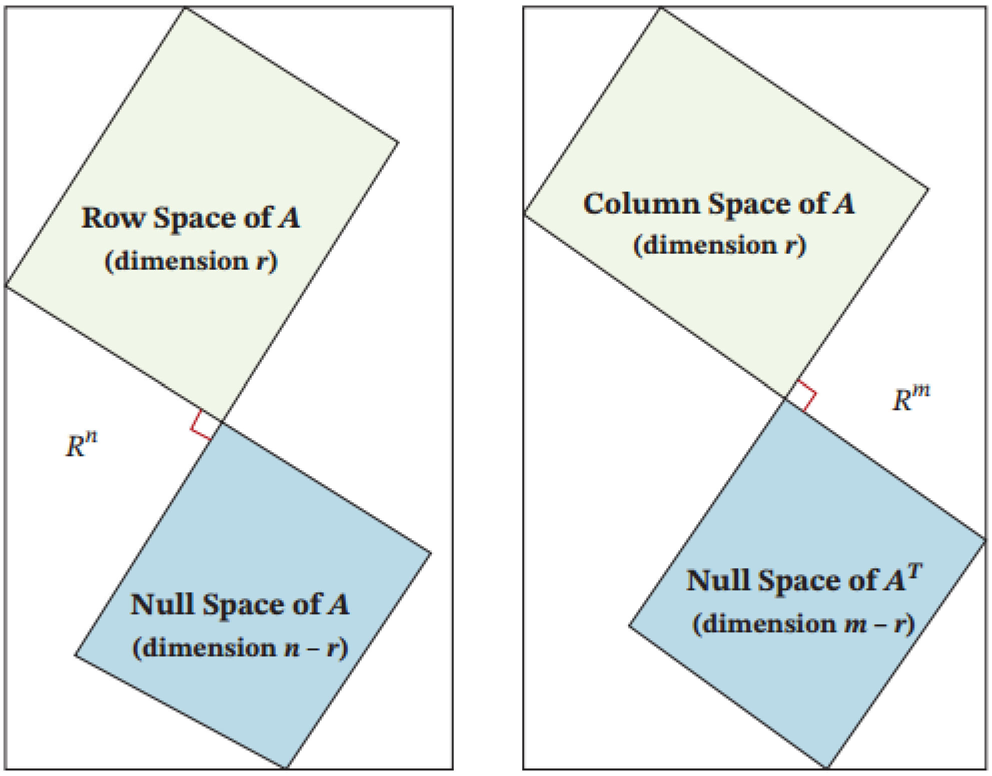
(*c*) The orthogonal complement of $W^\perp$ is W.

Theorem 4.9.7

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

(*a*) The null space of A and the row space of A are orthogonal complements in R^n.

(*b*) The null space of A^T and the column space of A are orthogonal complements in R^m.



Theorem 4.9.8

Equivalent Statements

If A is an $n \times n$ matrix in which there are no duplicate rows and no duplicate columns, 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 $\{\mathbf{0}\}$.

Theorem 4.9.9

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

(a) (***Overdetermined Case***). If $m > n$, then the linear system $A\mathbf{x} = \mathbf{b}$ is inconsistent for at least one vector $\mathbf{b}$ in R^n.
(b) (***Underdetermined Case***). If $m < n$, then for each vector $\mathbf{b}$ in R^m the linear system $A\mathbf{x} = \mathbf{b}$ is either inconsistent or has infinitely many solutions.

Chapter 5 – Eigenvalues and Eigenvectors

Exercise 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 λ***.

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 5.1.2

If A is an $n \times n$ triangular matrix (upper triangular, lower triangular, or diagonal), then the eigenvalues of A are the entries on the main diagonal 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) λ is a solution of the characteristic equation $\det(\lambda I - A) = 0$.

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

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

$$(\lambda I - A)\mathbf{x} = \mathbf{0}$$

Thus, we can find the eigenvectors of A corresponding to λ by finding the nonzero vectors in the solution space of this linear system. This solution space, which is called the ***eigenspace*** of A corresponding to λ, can also be viewed as:

1. the null space of the matrix $\lambda I - A$
2. the kernel of the matrix operator $T_{\lambda I - A}: R^n \to R^n$
3. the set of vectors for which $A\mathbf{x} = \lambda\mathbf{x}$

Exercise 5.2 – Diagonalization

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.

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.

Theorem 5.2.1

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

(a) A is diagonalizable.

(b) A has n linearly independent eigenvectors.

Theorem 5.2.2

(a) If $\lambda_1, \lambda_2, \ldots, \lambda_k$ are distinct eigenvalues of a matrix A, and if $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k$ are corresponding eigenvectors, then $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\}$ is a linearly independent set.

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

A Procedure for Diagonalizing an $n \times n$ Matrix

Step 1. Determine first whether the matrix is actually diagonalizable by searching for n linearly independent eigenvectors. One way to do this is to find a basis for each eigenspace and count the total number of vectors obtained. If there is a total of n vectors, then the matrix is diagonalizable, and if the total is less than n, then it is not.

Step 2. If you ascertained that the matrix is diagonalizable, then form the matrix $P = [\mathbf{p}_1 \quad \mathbf{p}_2 \quad \cdots \quad \mathbf{p}_n]$ whose column vectors are the n basis vectors you obtained in Step 1.

Step 3. $P^{-1}AP$ will be a diagonal matrix whose successive diagonal entries are the eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$ that correspond to the successive columns of P.

In general, there is no preferred order for the columns of P. Since the ith diagonal entry of $P^{-1}AP$ is an eigenvalue for the ith column vector of P, changing the order of the columns of P just changes the order of the eigenvalues on the diagonal of $P^{-1}AP$. Thus,

Theorem 5.2.3

If k is a positive integer, λ is an eigenvalue of a matrix A, and $\mathbf{x}$ is a corresponding eigenvector, then λ^k is an eigenvalue of A^k and $\mathbf{x}$ is a corresponding eigenvector.

$$A^k = PD^kP^{-1} = P \begin{bmatrix} \lambda_1^k & 0 & \cdots & 0 \\ 0 & \lambda_2^k & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & \lambda_n^k \end{bmatrix} P^{-1}$$

There is some terminology that is related to these ideas. If λ_0 is an eigenvalue of an $n \times n$ matrix A, then the dimension of the eigenspace corresponding to λ_0 is called the ***geometric multiplicity*** of λ_0, and the number of times that $\lambda - \lambda_0$ appears as a factor in the characteristic polynomial of A is called the ***algebraic multiplicity*** of λ_0. The following theorem, which we state without proof, summarizes the preceding discussion.

Theorem 5.2.4

Geometric and Algebraic Multiplicity

If A is a square matrix, then:

(*a*) For every eigenvalue of A, the geometric multiplicity is less than or equal to the algebraic multiplicity.

(*b*) A is diagonalizable if and only if its characteristic polynomial can be expressed as a product of linear factors, and the geometric multiplicity of every eigenvalue is equal to the algebraic multiplicity.

Chapter 6 – Inner Product Spaces

Exercise 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***.

Because the axioms for a real inner product space are based on properties of the dot product, these inner product space axioms will be satisfied automatically if we define the inner product of two vectors $\mathbf{u}$ and $\mathbf{v}$ in R^n to be

$$\boxed{\langle \mathbf{u}, \mathbf{v} \rangle = \mathbf{u} \cdot \mathbf{v} = u_1v_1 + u_2v_2 + \cdots + u_nv_n} \tag{1}$$

This inner product is commonly called the ***Euclidean inner product*** (or the ***standard inner product***) on R^n to distinguish it from other possible inner products that might be defined on R^n. We call R^n with the Euclidean inner product ***Euclidean n-space***.

Definition 2

If V is a real inner product space, then the ***norm*** (or ***length***) of a vector $\mathbf{v}$ in V is denoted by $\|\mathbf{v}\|$ and is defined by

$$\|\mathbf{v}\| = \sqrt{\langle \mathbf{v}, \mathbf{v} \rangle}$$

and the ***distance*** between two vectors is denoted by $d(\mathbf{u}, \mathbf{v})$ and is defined by

$$d(\mathbf{u}, \mathbf{v}) = \|\mathbf{u} - \mathbf{v}\| = \sqrt{\langle \mathbf{u} - \mathbf{v}, \mathbf{u} - \mathbf{v} \rangle}$$

A vector of norm 1 is called a ***unit vector***.

Theorem 6.1.1

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

(*a*) $\|\mathbf{v}\| \geq 0$ with equality if and only if $\mathbf{v} = \mathbf{0}$.
(*b*) $\|k\mathbf{v}\| = |k|\|\mathbf{v}\|$.
(*c*) $d(\mathbf{u}, \mathbf{v}) = d(\mathbf{v}, \mathbf{u})$.
(*d*) $d(\mathbf{u}, \mathbf{v}) \geq 0$ with equality if and only if $\mathbf{u} = \mathbf{v}$.

Although the Euclidean inner product is the most important inner product on R^n, there are various applications in which it is desirable to modify it by *weighting* each term differently. More precisely, if

$$w_1, w_2, \ldots, w_n$$

are *positive* real numbers, called ***weights***, and 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 it can be shown that the formula

$$\boxed{\langle \mathbf{u}, \mathbf{v} \rangle = w_1u_1v_1 + w_2u_2v_2 + \cdots + w_nu_nv_n} \tag{2}$$

defines an inner product on R^n that we call the ***weighted Euclidean inner product with weights*** $\boldsymbol{w_1, w_2, \ldots, w_n}$.

Definition 3

If V is an inner product space, then the set of points in V that satisfy

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

is called the ***unit sphere*** in V (or the ***unit circle*** in the case where $V = R^2$).

and let A be an *invertible* $n \times n$ matrix. It can be shown (Exercise 47) that if $\mathbf{u} \cdot \mathbf{v}$ is the Euclidean inner product on R^n, then the formula

$$\boxed{\langle \mathbf{u}, \mathbf{v} \rangle = A\mathbf{u} \cdot A\mathbf{v}} \tag{5}$$

also defines an inner product; it is called the ***inner product on R^n generated by A***.

and the weighted Euclidean inner product

$$\langle \mathbf{u}, \mathbf{v} \rangle = w_1 u_1 v_1 + w_2 u_2 v_2 + \cdots + w_n u_n v_n \tag{7}$$

is generated by the matrix

$$A = \begin{bmatrix} \sqrt{w_1} & 0 & 0 & \cdots & 0 \\ 0 & \sqrt{w_2} & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots & \sqrt{w_n} \end{bmatrix}$$

If $\mathbf{u} = U$ and $\mathbf{v} = V$ are matrices in the vector space M_{nn}, then the formula

$$\boxed{\langle \mathbf{u}, \mathbf{v} \rangle = \operatorname{tr}(U^T V)}$$

defines an inner product on M_{nn} called the ***standard inner product*** on that space

$$\langle \mathbf{u}, \mathbf{v} \rangle = \operatorname{tr}(U^T V) = u_1 v_1 + u_2 v_2 + u_3 v_3 + u_4 v_4$$

which is just the dot product of the corresponding entries in the two matrices. And it follows from this that

$$\|\mathbf{u}\| = \sqrt{\langle \mathbf{u}, \mathbf{u} \rangle} = \sqrt{\operatorname{tr}\langle U^T U \rangle} = \sqrt{u_1^2 + u_2^2 + u_3^2 + u_4^2}$$

If

$$\mathbf{p} = a_0 + a_1 x + \cdots + a_n x^n \quad \text{and} \quad \mathbf{q} = b_0 + b_1 x + \cdots + b_n x^n$$

are polynomials in P_n, then the following formula defines an inner product on P_n (verify) that we will call the ***standard inner product*** on this space:

$$\langle \mathbf{p}, \mathbf{q} \rangle = a_0 b_0 + a_1 b_1 + \cdots + a_n b_n \tag{9}$$

The norm of a polynomial $\mathbf{p}$ relative to this inner product is

$$\|\mathbf{p}\| = \sqrt{\langle \mathbf{p}, \mathbf{p} \rangle} = \sqrt{a_0^2 + a_1^2 + \cdots + a_n^2}$$

If

$$\mathbf{p} = p(x) = a_0 + a_1 x + \cdots + a_n x^n \quad \text{and} \quad \mathbf{q} = q(x) = b_0 + b_1 x + \cdots + b_n x^n$$

are polynomials in P_n, and if $x_0, x_1, \ldots, x_n$ are distinct real numbers (called ***sample points***), then the formula

$$\langle \mathbf{p}, \mathbf{q} \rangle = p(x_0)q(x_0) + p(x_1)q(x_1) + \cdots + p(x_n)q(x_n) \tag{10}$$

defines an inner product on P_n called the ***evaluation inner product*** at $x_0, x_1, \ldots, x_n$. Algebraically, this can be viewed as the dot product in R^n of the n-tuples

$$(p(x_0), p(x_1), \ldots, p(x_n)) \quad \text{and} \quad (q(x_0), q(x_1), \ldots, q(x_n))$$

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

Exercise 6.2 – Angle and Orthogonality in Inner Product Spaces

Theorem 6.2.1

Cauchy–Schwarz Inequality
If $\mathbf{u}$ and $\mathbf{v}$ are vectors in a real inner product space V, then

$$|\langle \mathbf{u}, \mathbf{v}\rangle| \le \|\mathbf{u}\|\|\mathbf{v}\| \tag{3}$$

The following two alternative forms of the Cauchy–Schwarz inequality are useful to know:

$$\boxed{\langle \mathbf{u}, \mathbf{v}\rangle^2 \le \langle \mathbf{u}, \mathbf{u}\rangle\langle \mathbf{v}, \mathbf{v}\rangle} \tag{4}$$

$$\boxed{\langle \mathbf{u}, \mathbf{v}\rangle^2 \le \|\mathbf{u}\|^2\|\mathbf{v}\|^2} \tag{5}$$

The first of these formulas was obtained in the proof of Theorem 6.2.1, and the second is a variation of the first.

This enables us to *define* the ***angle*** θ ***between*** $\mathbf{u}$ ***and*** $\mathbf{v}$ to be

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

Theorem 6.2.2

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

(*a*) $\|\mathbf{u}+\mathbf{v}\| \le \|\mathbf{u}\| + \|\mathbf{v}\|$ **[Triangle inequality for vectors]**
(*b*) $d(\mathbf{u}, \mathbf{v}) \le d(\mathbf{u}, \mathbf{w}) + d(\mathbf{w}, \mathbf{v})$ **[Triangle inequality for distances]**

Definition 1

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

Theorem 6.2.3

Generalized Theorem of Pythagoras
If $\mathbf{u}$ and $\mathbf{v}$ are orthogonal vectors in a real inner product space, then

$$\|\mathbf{u}+\mathbf{v}\|^2 = \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2$$

Definition 2

If W is a subspace of a real inner product space V, then the set of all vectors in V that are orthogonal to every vector in W is called the ***orthogonal complement*** of W and is denoted by the symbol $W^{\perp}$.

Theorem 6.2.4

If W is a subspace of a real inner product space V, then:

(*a*) $W^{\perp}$ is a subspace of V.

(*b*) $W \cap W^{\perp} = \{\mathbf{0}\}$.

Theorem 6.2.5

If W is a subspace of a real finite-dimensional inner product space V, then the orthogonal complement of $W^{\perp}$ is W; that is,

$$(W^{\perp})^{\perp} = W$$

Exercise 6.3 – Gram–Schmidt Process QR-Decomposition

Definition 1

A set of two or more vectors in a real inner product space is said to be ***orthogonal*** if all pairs of distinct vectors in the set are orthogonal. An orthogonal set in which each vector has norm 1 is said to be ***orthonormal***.

Theorem 6.3.1

If $S = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ is an orthogonal set of nonzero vectors in an inner product space, then S is linearly independent.

Theorem 6.3.2

(*a*) If $S = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ is an orthogonal basis for an inner product space V, and if $\mathbf{u}$ is any vector in V, then

$$\mathbf{u} = \frac{\langle \mathbf{u}, \mathbf{v}_1 \rangle}{\|\mathbf{v}_1\|^2}\mathbf{v}_1 + \frac{\langle \mathbf{u}, \mathbf{v}_2 \rangle}{\|\mathbf{v}_2\|^2}\mathbf{v}_2 + \cdots + \frac{\langle \mathbf{u}, \mathbf{v}_n \rangle}{\|\mathbf{v}_n\|^2}\mathbf{v}_n \tag{3}$$

(*b*) If $S = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ is an orthonormal basis for an inner product space V, and if $\mathbf{u}$ is any vector in V, then

$$\mathbf{u} = \langle \mathbf{u}, \mathbf{v}_1 \rangle \mathbf{v}_1 + \langle \mathbf{u}, \mathbf{v}_2 \rangle \mathbf{v}_2 + \cdots + \langle \mathbf{u}, \mathbf{v}_n \rangle \mathbf{v}_n \tag{4}$$

Theorem 6.3.3

Projection Theorem

If W is a finite-dimensional subspace of an inner product space V, then every vector $\mathbf{u}$ in V can be expressed in exactly one way as

$$\mathbf{u} = \mathbf{w}_1 + \mathbf{w}_2 \tag{8}$$

where $\mathbf{w}_1$ is in W and $\mathbf{w}_2$ is in $W^{\perp}$.

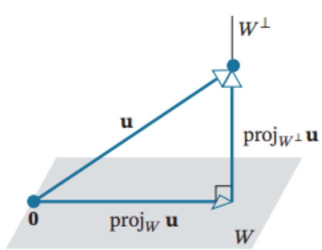


FIGURE 6.3.1

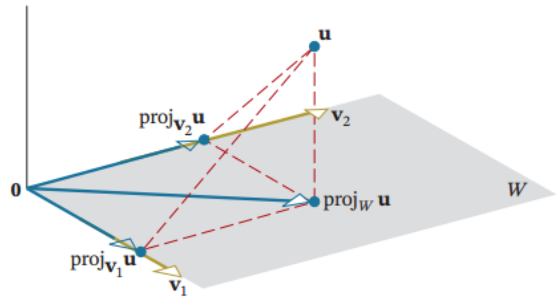


FIGURE 6.3.2

The vectors $\mathbf{w}_1$ and $\mathbf{w}_2$ in Formula (8) are commonly denoted by

$$\mathbf{w}_1 = \text{proj}_W \mathbf{u} \quad \text{and} \quad \mathbf{w}_2 = \text{proj}_{W^\perp} \mathbf{u} \tag{9}$$

These are called the ***orthogonal projection of* $\mathbf{u}$ *on* W** and the ***orthogonal projection of* $\mathbf{u}$ *on* $W^\perp$**, respectively. The vector $\mathbf{w}_2$ is also called the ***component of* $\mathbf{u}$ *orthogonal to* W**. Using the notation in (9), Formula (8) can be expressed as

$$\boxed{\mathbf{u} = \text{proj}_W \mathbf{u} + \text{proj}_{W^\perp} \mathbf{u}} \tag{10}$$

(Figure 6.3.1). Moreover, since $\text{proj}_{W^\perp}\mathbf{u} = \mathbf{u} - \text{proj}_W \mathbf{u}$, we can also express Formula (10) as

$$\boxed{\mathbf{u} = \text{proj}_W \mathbf{u} + (\mathbf{u} - \text{proj}_W \mathbf{u})} \tag{11}$$

Theorem 6.3.4

Let W be a finite-dimensional subspace of an inner product space V.

(*a*) If $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_r\}$ is an orthogonal basis for W, and $\mathbf{u}$ is any vector in V, then

$$\text{proj}_W \mathbf{u} = \frac{\langle \mathbf{u}, \mathbf{v}_1\rangle}{\|\mathbf{v}_1\|^2}\mathbf{v}_1 + \frac{\langle \mathbf{u}, \mathbf{v}_2\rangle}{\|\mathbf{v}_2\|^2}\mathbf{v}_2 + \cdots + \frac{\langle \mathbf{u}, \mathbf{v}_r\rangle}{\|\mathbf{v}_r\|^2}\mathbf{v}_r \tag{12}$$

(*b*) If $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_r\}$ is an orthonormal basis for W, and $\mathbf{u}$ is any vector in V, then

$$\text{proj}_W \mathbf{u} = \langle \mathbf{u}, \mathbf{v}_1\rangle \mathbf{v}_1 + \langle \mathbf{u}, \mathbf{v}_2\rangle \mathbf{v}_2 + \cdots + \langle \mathbf{u}, \mathbf{v}_r\rangle \mathbf{v}_r \tag{13}$$

Theorem 6.3.5

Every nonzero finite-dimensional inner product space has an orthonormal basis.

The Gram–Schmidt Process

To convert a basis $\{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_r\}$ into an orthogonal basis $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_r\}$, perform the following computations:

Step 1. $\mathbf{v}_1 = \mathbf{u}_1$

Step 2. $\mathbf{v}_2 = \mathbf{u}_2 - \dfrac{\langle \mathbf{u}_2, \mathbf{v}_1\rangle}{\|\mathbf{v}_1\|^2}\mathbf{v}_1$

Step 3. $\mathbf{v}_3 = \mathbf{u}_3 - \dfrac{\langle \mathbf{u}_3, \mathbf{v}_1\rangle}{\|\mathbf{v}_1\|^2}\mathbf{v}_1 - \dfrac{\langle \mathbf{u}_3, \mathbf{v}_2\rangle}{\|\mathbf{v}_2\|^2}\mathbf{v}_2$

Step 4. $\mathbf{v}_4 = \mathbf{u}_4 - \dfrac{\langle \mathbf{u}_4, \mathbf{v}_1\rangle}{\|\mathbf{v}_1\|^2}\mathbf{v}_1 - \dfrac{\langle \mathbf{u}_4, \mathbf{v}_2\rangle}{\|\mathbf{v}_2\|^2}\mathbf{v}_2 - \dfrac{\langle \mathbf{u}_4, \mathbf{v}_3\rangle}{\|\mathbf{v}_3\|^2}\mathbf{v}_3$

$\vdots$

(continue for r steps)

Optional Step. To convert the orthogonal basis into an orthonormal basis $\{\mathbf{q}_1, \mathbf{q}_2, \ldots, \mathbf{q}_r\}$, normalize the orthogonal basis vectors.

Theorem 6.3.6

If W is a finite-dimensional inner product space, then:

(*a*) Every orthogonal set of nonzero vectors in W can be enlarged to an orthogonal basis for W.

(*b*) Every orthonormal set in W can be enlarged to an orthonormal basis for W.

Theorem 6.3.7

***QR*-Decomposition**

If A is an $m \times n$ matrix with linearly independent column vectors, then A can be factored as

$$A = QR$$

where Q is an $m \times n$ matrix with orthonormal column vectors, and R is an $n \times n$ invertible upper triangular matrix.

Chapter 7 – Diagonalization and Quadratic Forms

Exercise 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

$$AA^T = A^TA = I \tag{1}$$

A matrix transformation $T_A : R^n \to R^n$ is said to be an ***orthogonal transformation*** or an ***orthogonal operator*** if A is an orthogonal matrix.

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.

$$AA^T = \begin{bmatrix} \mathbf{r}_1\mathbf{c}_1^T & \mathbf{r}_1\mathbf{c}_2^T & \cdots & \mathbf{r}_1\mathbf{c}_n^T \\ \mathbf{r}_2\mathbf{c}_1^T & \mathbf{r}_2\mathbf{c}_2^T & \cdots & \mathbf{r}_2\mathbf{c}_n^T \\ \vdots & \vdots & & \vdots \\ \mathbf{r}_n\mathbf{c}_1^T & \mathbf{r}_n\mathbf{c}_2^T & \cdots & \mathbf{r}_n\mathbf{c}_n^T \end{bmatrix} = \begin{bmatrix} \mathbf{r}_1 \cdot \mathbf{r}_1 & \mathbf{r}_1 \cdot \mathbf{r}_2 & \cdots & \mathbf{r}_1 \cdot \mathbf{r}_n \\ \mathbf{r}_2 \cdot \mathbf{r}_1 & \mathbf{r}_2 \cdot \mathbf{r}_2 & \cdots & \mathbf{r}_2 \cdot \mathbf{r}_n \\ \vdots & \vdots & & \vdots \\ \mathbf{r}_n \cdot \mathbf{r}_1 & \mathbf{r}_n \cdot \mathbf{r}_2 & \cdots & \mathbf{r}_n \cdot r_n \end{bmatrix}$$

Theorem 7.1.2

(a) The transpose of an orthogonal matrix is orthogonal.

(b) The inverse of an orthogonal matrix is orthogonal.

(c) A product of orthogonal matrices is orthogonal.

(d) 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.

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.

Thus, to find P we must find the coordinates of the old basis vectors with respect to the new basis. We leave it for you to deduce the following results Figure 7.1.2*b*.

$$[\mathbf{u}_1]_{B'} = \begin{bmatrix} \cos\theta \\ -\sin\theta \end{bmatrix} \quad \text{and} \quad [\mathbf{u}_2]_{B'} = \begin{bmatrix} \sin\theta \\ \cos\theta \end{bmatrix} \tag{3}$$

Thus

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} \tag{4}$$

or equivalently

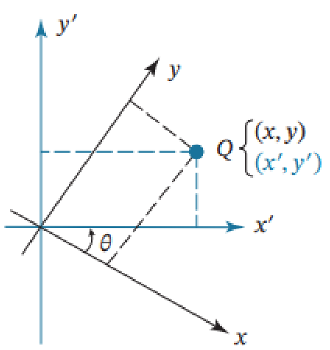
$$\boxed{\begin{aligned} x' &= x\cos\theta + y\sin\theta \\ y' &= -x\sin\theta + y\cos\theta \end{aligned}} \tag{5}$$

These are sometimes called the ***rotation equations*** for R^2.

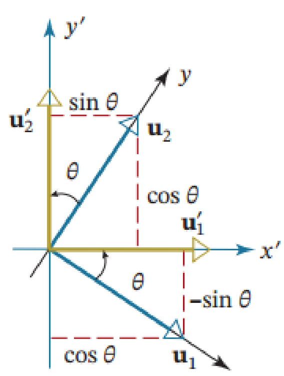
$$P^{-1} = \begin{bmatrix} \cos\theta & \sin\theta & 0 \\ -\sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

(verify). Thus, the new coordinates (x', y', z') of a point Q can be computed from its old coordinates (x, y, z) by

$$\begin{bmatrix} x' \\ y' \\ z' \end{bmatrix} = \begin{bmatrix} \cos\theta & \sin\theta & 0 \\ -\sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix}$$



(*a*)



(*b*)

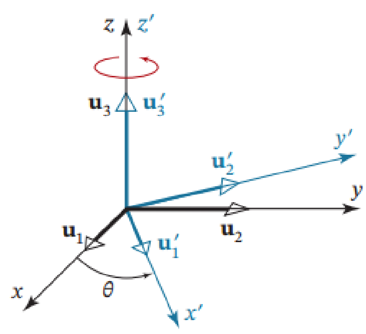


FIGURE 7.1.3

Exercise 7.2 – Orthogonal Diagonalization

Definition 1

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

Note that if B is orthogonally similar to A, then it is also true that A is orthogonally similar to B since we can express A as $A = PBP^T = Q^TBQ$, where $Q = P^T$. This being the case we will say that A and B are ***orthogonally similar matrices*** if either is orthogonally similar to the other.

If A is orthogonally similar to some diagonal matrix, say

$$P^TAP = D$$

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

Theorem 7.2.1

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

(a) A is orthogonally diagonalizable.

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

(c) A is symmetric.

Theorem 7.2.2

If A is a symmetric matrix with real entries, 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.

Remark The justification of this procedure should be clear: Theorem 7.2.2 ensures that eigenvectors from *different* eigenspaces are orthogonal, and applying the Gram–Schmidt process ensures that the eigenvectors within the *same* eigenspace are orthonormal. Thus the *entire* set of eigenvectors obtained by this procedure will be orthonormal.

Exercise 7.3 – Quadratic Forms

If A is a symmetric matrix with real entries that is orthogonally diagonalized by

$$P = [\mathbf{u}_1 \quad \mathbf{u}_2 \quad \cdots \quad \mathbf{u}_n]$$

$$\boxed{A = \lambda_1 \mathbf{u}_1 \mathbf{u}_1^T + \lambda_2 \mathbf{u}_2 \mathbf{u}_2^T + \cdots + \lambda_n \mathbf{u}_n \mathbf{u}_n^T}$$

which is called a ***spectral decomposition of A***.*

Problem 1	If $\mathbf{x}^T A\mathbf{x}$ is a quadratic form on R^2 or R^3, what kind of curve or surface is represented by the equation $\mathbf{x}^T A\mathbf{x} = k$?
Problem 2	If $\mathbf{x}^T A\mathbf{x}$ is a quadratic form on R^n, what conditions must A satisfy for $\mathbf{x}^T A\mathbf{x}$ to have positive values for $\mathbf{x} \neq \mathbf{0}$?
Problem 3	If $\mathbf{x}^T A\mathbf{x}$ is a quadratic form on R^n, what are its maximum and minimum values if $\mathbf{x}$ is constrained to satisfy $\|\mathbf{x}\| = 1$?