MATH 247 Introductory Linear Algebra

Textbook *"Linear Algebra and Its Applications"* by David Lay, Pearson Education, Inc., 2006 (3rd edition update).

1.1. Systems of Linear Equations.

Definition. A linear equation in the variables $x_1, \dots, x_n$ is an equation of the form

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

where $a_1, \dots, a_n$, and b are real or complex numbers.

Example. $x_1 - 2x_2 = 3$, $ix_2 - \sqrt{2}x_1 = \pi$ are linear equations in x_1 and x_2. Equations $x_1^2 = 5$ and $x_1 \ln(x_2) = -7$ are not linear.

Definition. A system of linear equations is a collection of linear equations in the same variables $x_1, \dots, x_n$.

Example.

$$\begin{cases} x_1 - x_2 = 1 \\ x_1 + x_2 = 5 \end{cases}$$

is a system of two linear equations in two unknowns x_1 and x_2.

$$\begin{cases} x_1 + x_3 = 8 \\ x_3 - x_2 = 4 \end{cases}$$

is a system of two linear equations in three unknowns x_1, x_2, and x_3.

Definition. A solution of a system of linear equations is a set of numbers $s_1, \dots, s_n$ that make the equations true statements if $x_1 = s_1, \dots, x_n = s_n$.

Example. The system

$$\begin{cases} x_1 - x_2 = 1 \\ x_1 + x_2 = 5 \end{cases}$$

has solution $x_1 = 3$ and $x_2 = 2$.

Definition. A system of linear equations has either
(i) no solution (inconsistent system),
(ii) exactly one solution (consistent system),
or
(iii) infinitely many solutions (consistent system).

Example. A graph of a linear equation is a straight line. Therefore, finding a solution of a system of two linear equations is equivalent to finding an intersection of two lines. Lines can be either (i) parallel (no solution), (ii) intersect at a point (exactly one solution), or (iii) coincide (infinitely many solutions). For example, system $\begin{cases} x_1 + x_2 = 3 \\ 2x_1 + 2x_2 = 4 \end{cases}$ has no solution, system $\begin{cases} x_1 + x_2 = 4 \\ x_1 - x_2 = 2 \end{cases}$ has exactly one solution, and system $\begin{cases} x_1 + x_2 = 4 \\ 2x_1 + 2x_2 = 8 \end{cases}$ has infinitely many solutions.

Definition. Two systems of equations are called equivalent if they have the same solution.

Example. Systems $\begin{cases} x_1 + 3x_2 = 7 \\ 2x_1 - x_2 = 0 \end{cases}$ and $\begin{cases} x_1 + 3x_2 = 7 \\ 3x_1 + 2x_2 = 7 \end{cases}$ are equivalent since they have the same solution $x_1 = 1$ and $x_2 = 2$.

Definition. Given the system

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2 \\ \dots \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n = b_n \end{cases},$$

the matrix

$$\begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{bmatrix}$$

is called the coefficient matrix or matrix of coefficients. The matrix

$$\begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} & b_1 \\ a_{21} & a_{22} & \dots & a_{2n} & b_2 \\ \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} & b_n \end{bmatrix}$$

is called the augmented matrix.

Example. Consider the system

$$\begin{cases} -x_1 + x_3 = -3 \\ 2x_1 + x_2 - x_3 = 4 \\ x_1 + 2x_2 = 1 \end{cases}.$$

Its augmented matrix is

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

Definition. A matrix of size $m \times n$ has m rows and n columns.

Example. In the above example, we have 3×4 matrix.

1.1. Solving Linear System.

To solve a system of linear equations we replace it with an equivalent one that is easier to solve. Three basic operations are used to simplify a linear system (i) replace an equation by the sum of itself and a multiple of another equation, (ii) interchange two equations, and (iii) multiply an equation by a nonzero constant.

Example.

$$\begin{cases} x_1 - x_2 = 1 \\ x_1 + x_2 = 5 \end{cases} \overset{e_1 \to e_1 + e_2}{\Longrightarrow} \begin{cases} 2x_1 = 6 \\ x_1 + x_2 = 5 \end{cases}$$

$$\overset{e_1 \to e_1/2}{\Longrightarrow} \begin{cases} x_1 = 3 \\ x_1 + x_2 = 5 \end{cases} \overset{e_2 \to e_2 - e_1}{\Longrightarrow} \begin{cases} x_1 = 3 \\ x_2 = 2. \end{cases}$$

This solution can be written in terms of augmented matrices as follows:

$$\begin{bmatrix} 1 & -1 & 1 \\ 1 & 1 & 5 \end{bmatrix} \overset{r_1 \to r_1 + r_2}{\Longrightarrow} \begin{bmatrix} 2 & 0 & 6 \\ 1 & 1 & 5 \end{bmatrix}$$

$$\overset{r_1 \to r_1/2}{\Longrightarrow} \begin{bmatrix} 1 & 0 & 3 \\ 1 & 1 & 5 \end{bmatrix} \overset{r_2 \to r_2 - r_1}{\Longrightarrow} \begin{bmatrix} 1 & 0 & 3 \\ 0 & 1 & 2 \end{bmatrix}.$$

To solve a system of three or more equations, it is advisable to work with augmented matrices. Three elementary row operations that lead to a row equivalent augmented matrix are
(i) replace one row by the sum of itself and a multiple of another row,
(ii) interchange two rows,
and (iii) multiply a row by a nonzero constant.

1.2. Row Reduction and Echelon Forms.

Example. Consider two matrices

$$\begin{bmatrix} 2 & -2 & 0 & 1 \\ 0 & 1 & 2 & 3 \\ 0 & 0 & 0 & 5 \\ 0 & 0 & 0 & 0 \end{bmatrix} \text{ and } \begin{bmatrix} 1 & 0 & 0 & -1 \\ 0 & 1 & 0 & 3 \\ 0 & 0 & 1 & 2 \end{bmatrix}.$$

Definition. In a matrix, a leading entry of a nonzero row is the leftmost nonzero entry of the row.

In our example, the leading entries are 2, 1 and 5 in the first matrix, and 1, 1, and 1 in the second matrix.

Definition. A matrix is in <u>row echelon form</u> if:
(i) all nonzero rows are above any rows of all zeros,
(ii) each leading entry of a row is to the right of the leading entry of the row above it,
and (iii) all entries in a column below a leading entry are zeros.

In our example, both matrices are in row echelon form.

Definition. A matrix is in <u>reduced row echelon form</u> if:
(i) the leading entry in each nonzero row is 1,
and (ii) each leading 1 is the only nonzero entry in its column.

In our example, the second matrix is in reduced row echelon form.

Definition. A <u>pivot position</u> in a matrix is a location in this matrix that corresponds to a leading 1 in its reduced row echelon form. A <u>pivot column</u> is a column that contains a pivot position.

Algorithm for Solving a Linear System.

To solve a system of equations, we reduce the augmented matrix to its row echelon form and then to its reduced row echelon form. It is done in several steps.

<u>Step 1.</u> Take the leftmost nonzero column. This will be your pivot column. Make sure the top entry (pivot) is not zero. Interchange rows if necessary.

Example.

$$\begin{cases} x_1 + x_2 + 2x_3 = 0 \\ 2x_2 + x_3 = 4 \\ x_1 + 2x_3 = -3 \end{cases} \quad \leftrightarrow \begin{bmatrix} 1 & 1 & 2 & 0 \\ 0 & 2 & 1 & 4 \\ 1 & 0 & 2 & -3 \end{bmatrix}.$$

The first column is a pivot column.

<u>Step 2.</u> Use the elementary row operations to create zeros in all positions below the pivot.

In our example,

$$\overset{r_3 \to r_3 - r_1}{\Longrightarrow} \begin{bmatrix} 1 & 1 & 2 & 0 \\ 0 & 2 & 1 & 4 \\ 0 & -1 & 0 & -3 \end{bmatrix}.$$

<u>Step 3.</u> Select a pivot column in the matrix with the first row ignored. Repeat the previous steps.

In our example,

$$\xrightarrow{r_3 \leftrightarrow 2r_3 + r_2} \begin{bmatrix} 1 & 1 & 2 & 0 \\ 0 & 2 & 1 & 4 \\ 0 & 0 & 1 & -2 \end{bmatrix}.$$

This matrix is in row echelon form.

Step 4. Starting with the rightmost pivot, create zeros above each pivot. Make pivots equal 1 by rescaling rows if necessary.

In our example,

$$\xrightarrow{r_2 \to r_2 - r_3} \begin{bmatrix} 1 & 1 & 2 & 0 \\ 0 & 2 & 0 & 6 \\ 0 & 0 & 1 & -2 \end{bmatrix} \xrightarrow{r_1 \to r_1 - 2r_3} \begin{bmatrix} 1 & 1 & 0 & 4 \\ 0 & 2 & 0 & 6 \\ 0 & 0 & 1 & -2 \end{bmatrix} \xrightarrow{r_2 \to r_2/2} \begin{bmatrix} 1 & 1 & 0 & 4 \\ 0 & 1 & 0 & 3 \\ 0 & 0 & 1 & -2 \end{bmatrix}$$

$$\xrightarrow{r_1 \to r_1 - r_2} \begin{bmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 3 \\ 0 & 0 & 1 & -2 \end{bmatrix} \leftrightarrow \begin{cases} x_1 = 1 \\ x_2 = 3 \\ x_3 = -2. \end{cases}$$

1.6. Applications of Linear Systems.

Example. (Applications in Economics: Example 1, page 58). Consider coal, electric and steel sectors of economy. Suppose coal sells 60% of its output to electric and 40% to steel; electric sells 40% to coal, 10% to electric and 50% to steel; steel sells 60% to coal, 20% to electric and 20% to steel. It can be summarized by a table

Coal	Electric	Steel	Purchased by
0	.4	.6	coal
.6	.1	.2	electric
.4	.5	.2	steel

Let p_c, p_e and p_s denote the prices of the total outputs of coal, electric and steel, respectively. Find prices that balance each sector's income and expenditures.

SOLUTION: The prices must satisfy the system

$$\begin{cases} p_c = .4p_e + .6p_s \\ p_e = .6p_c + .1p_e + .2p_s \\ p_s = .4p_c + .5p_e + .2p_s \end{cases}.$$

The reduced row echelon form of the augmented matrix of this system is

$$\begin{bmatrix} 1 & 0 & -.94 & 0 \\ 0 & 1 & -.85 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

Thus, the solution is $p_c = .94p_s, p_e = .85p_s$ and p_s is free.

Example. (Applications in Chemistry: Example on page 59).
When propane burns, the propane C_3H_8 combines with oxygen O_2 and forms carbon dioxide CO_2 and water H_2O. Write a balanced chemical equation for this reaction.

SOLUTION: The number of atoms is preserved in a chemical reaction. Hence, we have to find integers x_1, x_2, x_3, and x_4 such that $x_1C_3H_8 + x_2O_2 \rightarrow x_3CO_2 + x_4H_2O$. That is, we have to solve the system

$$\begin{cases} 3x_1 = x_3 \\ 8x_1 = 2x_4 \\ 2x_2 = 2x_3 + x_4. \end{cases}$$

The solution is

$$\begin{cases} x_1 = .25x_4 \\ x_2 = 1.25x_4 \\ x_3 = .75x_4. \end{cases}$$

Since all x's must be integers, $x_4 = 4, x_1 = 1, x_2 = 5$, and $x_3 = 3$.

Example. (Network Flow: Example 2 on page 61).
In a network, the in-flow and out-flow of every node are equal. Consider intersections A, B, C, and D. Picture. Determine the flows at these intersections.

SOLUTION: The unknown flows satisfy the linear system

$$\begin{cases} x_1 + x_2 = 800 \\ x_2 + x_4 = x_3 + 300 \\ x_4 + x_5 = 500 \\ x_1 + x_5 = 600. \end{cases}$$

The solution is

$$\begin{cases} x_1 = 600 - x_5 \\ x_2 = 200 + x_5 \\ x_3 = 400 \\ x_4 = 500 - x_5. \end{cases}$$

The flow x_5 is free, though it must satisfy $x_5 \leq 500$ since $x_4 \geq 0$.

1.3. Vector Equations.

Definition. A column vector or a vector is a list of numbers arranged in a column.

Example. $\begin{bmatrix} 1 \\ -3 \\ 7 \end{bmatrix}$.

Notation. Vectors are usually denoted by $\mathbf{u}, \mathbf{v}, \mathbf{w}, \mathbf{x}, \mathbf{y}$.

Notation. The set of real numbers is denoted by $\mathbb{R}$.

Notation. The set of all vectors with n entries is denoted by $\mathbb{R}^n$. We write $\mathbf{u} \in \mathbb{R}^n$.

Definition. Two vectors are equal iff their corresponding entries are equal. That is, equality of vectors is defined entry-wise.

Example. Let $\mathbf{u} = \begin{bmatrix} 1 \\ -1 \end{bmatrix}, \mathbf{v} = \begin{bmatrix} 1 \\ i^2 \end{bmatrix}$, and $\mathbf{w} = \begin{bmatrix} e^{\pi i} \\ -1 \end{bmatrix}$. Entry-wise, $\mathbf{u} = \mathbf{v} \neq \mathbf{w}$.

Definition. Vector addition and scalar multiplication are defined entry-wise.

Example. In our example, $\mathbf{u} + \mathbf{w} = \begin{bmatrix} 1 + (-1) \\ -1 + (-1) \end{bmatrix} = \begin{bmatrix} 0 \\ -2 \end{bmatrix}$, and $2\mathbf{u} = \begin{bmatrix} 2(1) \\ 2(-1) \end{bmatrix} = \begin{bmatrix} 2 \\ -2 \end{bmatrix}$.

Geometric Presentation of a Vector.

A vector $\mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \\ \cdots \\ v_n \end{bmatrix}$ is plotted in n-dimensional coordinate system as a pointed line from the origin $\mathbf{0}$ to point $(v_1, v_2, \ldots, v_n)$. Picture in $\mathbb{R}^2$. Example $\mathbf{u} = \begin{bmatrix} 2 \\ 3 \end{bmatrix}$, $\mathbf{v} = \begin{bmatrix} -2 \\ 1 \end{bmatrix}$. How to picture the sum of two vectors and a vector multiplied by a scalar?

Parallelogram Rule for Addition.
The vector $\mathbf{u}+\mathbf{v}$ points from the origin and lies on the diagonal of the parallelogram formed by vectors $\mathbf{u}$ and $\mathbf{v}$. Picture. Example. Where is $\mathbf{u}-\mathbf{v}$? $\mathbf{v}-\mathbf{u}$?

Scalar Multiplication.
Vector $c\mathbf{v}$ lies on the same line as $\mathbf{v}$, its length is $|c|$ times the length of $\mathbf{v}$, and it points in the direction of $\mathbf{v}$ if $c > 0$ and in the opposite direction if $c < 0$. Picture. Example.

Definition. For any vectors $\mathbf{v_1}, \dots, \mathbf{v_p}$ in $\mathbb{R}^n$ and any $c_1, \dots, c_p \in \mathbb{R}$, the vector $c_1\mathbf{v_1} + \cdots + c_p\mathbf{v_p}$ is called a linear combination of $\mathbf{v_1}, \dots, \mathbf{v_p}$ with weights $c_1, \dots, c_p$.

Give an example.

Example. Let $\mathbf{u} = \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}$, $\mathbf{v} = \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix}$, and $\mathbf{w} = \begin{bmatrix} 3 \\ 0 \\ 1 \end{bmatrix}$. Is vector $\mathbf{y} = \begin{bmatrix} 0 \\ 2 \\ -1 \end{bmatrix}$ a linear combination of $\mathbf{u}, \mathbf{v}$, and $\mathbf{w}$?

SOLUTION: The goal is to find, if possible, weights c_u, c_v, and c_w such that $\mathbf{y} = c_u\mathbf{u}+c_v\mathbf{v}+c_w\mathbf{w}$ or $\begin{bmatrix} 0 \\ 2 \\ -1 \end{bmatrix} = c_u \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix} + c_v \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix} + c_w \begin{bmatrix} 3 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} c_u + 2c_v + 3c_w \\ c_v \\ -c_u + c_w \end{bmatrix}$.
The answer is $\mathbf{y} = -(1/4)\mathbf{u} + 2\mathbf{v} - (5/4)\mathbf{w}$.

Definition. The set of all linear combinations of vectors $\mathbf{v_1}, \mathbf{v_2}, \dots, \mathbf{v_p} \in \mathbb{R}^n$ is called the subset of $\mathbb{R}^n$ spanned (or generated) by $\mathbf{v_1}, \mathbf{v_2}, \dots, \mathbf{v_p}$ and is denoted by $Span\{\mathbf{v_1}, \mathbf{v_2}, \dots, \mathbf{v_p}\}$.

Example. $Span\{\mathbf{u}\}$ where $\mathbf{u} \in \mathbb{R}^3$. Picture.

Example. $Span\{\mathbf{u}, \mathbf{v}\}$ where the vectors are in $\mathbb{R}^3$. Picture.

Definition. A vector equation is a linear equation of the form $x_1\mathbf{v_1}+x_2\mathbf{v_2}+\cdots+x_p\mathbf{v_p} = \mathbf{b}$. It is equivalent to a linear system with the augmented matrix $\begin{bmatrix} \mathbf{v_1} & \mathbf{v_2} & \dots & \mathbf{v_p} & \mathbf{b} \end{bmatrix}$.

1.4. The Matrix Equation $\mathbf{Ax} = \mathbf{b}$.

Definition. A system of linear equations with augmented matrix $\begin{bmatrix} \mathbf{A} | \mathbf{b} \end{bmatrix}$ can be written as a matrix equation $\mathbf{Ax} = \mathbf{b}$.

1.5. Solution Sets of Linear Systems.

Homogeneous Linear Systems.

Definition. A system $\mathbf{Ax} = \mathbf{0}$ is called homogeneous. It has a trivial solution $\mathbf{x} = \mathbf{0}$. A non-trivial solution exists iff the system has at least one free variable.

Example (Example 1 on page 50). Describe the solution set of $\mathbf{Ax} = \mathbf{0}$ where

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

SOLUTION: The reduced row echelon form of $\begin{bmatrix}\mathbf{A}|\mathbf{0}\end{bmatrix}$ is

$$\begin{bmatrix} 3 & 5 & -4 & 0 \\ 0 & 3 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix},$$

meaning $x_1 = -4/3x_3, x_2 = 0$, and x_3 is free. Thus, any solution

$$\mathbf{x} = x_3 \begin{bmatrix} -4/3 \\ 0 \\ 1 \end{bmatrix} = x_3\mathbf{v},$$

that is, a non-trivial solution of this homogeneous system is in $Span\{v\}$.

Example (Example 2 on page 51). Describe the solution set of $10x_1 - 3x_2 - 2x_3 = 0$.

SOLUTION: From this equation, $x_1 = .3x_2 + .2x_3$ where x_2 and x_3 are free. Any solution

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} .3x_2 + .2x_3 \\ x_2 \\ x_3 \end{bmatrix} = x_2 \begin{bmatrix} .3 \\ 1 \\ 0 \end{bmatrix} + x_3 \begin{bmatrix} .2 \\ 0 \\ 1 \end{bmatrix} = x_2\mathbf{u} + x_3\mathbf{v},$$

that is, a non-trivial solution of this equation is in $Span\{\mathbf{u}, \mathbf{v}\}$.

Generally speaking, for a homogeneous system with k free variables, a non-trivial solution is in the span of k vectors.

Nonhomogeneous Linear Systems.

Definition. A system $\mathbf{Ax} = \mathbf{b}$ with nontrivial right-side ($\mathbf{b} \neq \mathbf{0}$) is called nonhomogeneous linear system.

Proposition. The solution set of a nonhomogeneous system $\mathbf{Ax} = \mathbf{b}$ is the set of all vectors of the form $\mathbf{w} = \mathbf{p} + \mathbf{v_h}$ where $\mathbf{p}$ is some particular solution of the nonhomogeneous system, and $\mathbf{v_h}$ is any solution of the homogeneous system $\mathbf{Ax} = \mathbf{0}$.
In other words, in order to solve a nonhomogeneous system, one has to find all solutions of the homogeneous system and add a particular solution of the nonhomogeneous system.

Example (Example 3 on page 52).

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

The reduced row echelon form of $\begin{bmatrix} \mathbf{A} | \mathbf{b} \end{bmatrix}$ is

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

meaning

$$\begin{cases} x_1 = -1 + 4/3x_3 \\ x_2 = 2 \\ x_3 \text{ free.} \end{cases}$$

Thus,

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} -1 + 4/3x_3 \\ 2 \\ x_3 \end{bmatrix} = \begin{bmatrix} -1 \\ 2 \\ 0 \end{bmatrix} + \begin{bmatrix} 4/3x_3 \\ 0 \\ x_3 \end{bmatrix} = \begin{bmatrix} -1 \\ 2 \\ 0 \end{bmatrix} + x_3 \begin{bmatrix} 4/3 \\ 0 \\ 1 \end{bmatrix} = \mathbf{p} + x_3\mathbf{v}.$$

Picture of $Span\{\mathbf{v}\}$, translated by a vector.

1.7. Linear Independence.

Definition. Vectors $\mathbf{v_1}, \dots, \mathbf{v_p} \in \mathbb{R}^n$ are linearly independent if the vector equation $x_1\mathbf{v_1} + \cdots + x_p\mathbf{v_p} = \mathbf{0}$ has only the trivial solution (there are no free variables). Otherwise, the vectors are linearly dependent and the vector equation with a non-trivial solution is called a linear dependence relation.

Example (Example 1 on page 65). Decide whether vectors $\mathbf{v_1} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$, $\mathbf{v_2} =$

$\begin{bmatrix}4\\5\\6\end{bmatrix}$, and $\mathbf{v_3} = \begin{bmatrix}2\\1\\0\end{bmatrix}$ are linearly independent.
SOLUTION: A row echelon form of the augmented matrix is

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

Thus, a non-trivial solution is possible. The vectors are not linear independent.

Find a linear dependence relation.
SOLUTION: The reduced row echelon form is

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

Thus, for example, $10\mathbf{v_1} - 5\mathbf{v_2} + 5\mathbf{v_3} = \mathbf{0}$.

Proposition. The columns of a matrix $\mathbf{A}$ are linearly independent iff the equation $\mathbf{Ax} = \mathbf{0}$ has only the trivial solution.

Example (Example 2 on pages 66 – 67). Determine if the columns of the matrix $\mathbf{A} = \begin{bmatrix}0 & 1 & 4\\1 & 2 & -1\\5 & 8 & 0\end{bmatrix}$ are independent.
SOLUTION: A row echelon form of the augmented matrix is

$$\begin{bmatrix}1 & 2 & -1 & 0\\0 & 1 & 4 & 0\\0 & 0 & 13 & 0\end{bmatrix}.$$

There are no free variables, therefore only a trivial solution exists and the columns are independent.

Proposition. (i) One vector is linearly independent iff it is a non-zero vector. (ii) Two vectors are linearly dependent iff one is a multiple of the other (that is, if they lie on the same line through the origin).
PROOF: (i) The equation $x_1\mathbf{v_1} = \mathbf{0}$ has a non-trivial solution iff $\mathbf{v_1} \neq \mathbf{0}$.
(ii) The equation $x_1\mathbf{v_1} + x_2\mathbf{v_2} = \mathbf{0}$ has a non-trivial solution iff $\mathbf{v_1} = -x_2/x_1\mathbf{v_2}$ (assumed $x_1 \neq 0$). $\square$

Example. Vector $\begin{bmatrix}1\\3\end{bmatrix}$ is linearly independent. Vectors $\begin{bmatrix}1\\-3\end{bmatrix}$ and $\begin{bmatrix}-3\\9\end{bmatrix}$ are linearly independent, while $\begin{bmatrix}1\\-3\end{bmatrix}$ and $\begin{bmatrix}2\\6\end{bmatrix}$ are linearly dependent.

Theorem 7 on page 68. Vectors $\mathbf{v_1}, \ldots, \mathbf{v_p}, p \geq 2$ are linearly dependent iff at least one of the vectors is a linear combination of the others (that is, at least one vector is in the set spanned by the others).
PROOF: The p vectors are linearly dependent iff the equation $x_1\mathbf{v_1} + \cdots + x_p\mathbf{v_p} = \mathbf{0}$ has a non-trivial solution iff (say, $x_1 \neq 0$) $\mathbf{v_1} = -x_2/x_1\mathbf{v_2} - \cdots - x_p/x_1\mathbf{v_p}$. $\square$

Example. (Example 4 on page 68). Consider vectors $\mathbf{v_1} = \begin{bmatrix} 3 \\ 0 \\ 1 \end{bmatrix}$ and $\mathbf{v_2} = \begin{bmatrix} 1 \\ 6 \\ 0 \end{bmatrix}$. They are not multiples of each other, therefore, $Span\{\mathbf{v_1}, \mathbf{v_2}\}$ is a plane through the origin. A vector in this plane is linearly dependent with $\mathbf{v_1}$ and $\mathbf{v_2}$. A vector not in this plane is linearly independent. Picture.

Theorem 8 on page 69. Vectors $\mathbf{v_1}, \ldots, \mathbf{v_p} \in \mathbb{R}^n$ are linearly dependent if $p > n$. In other words, if a set contains more vectors than there are entries in each vector, then the set is linearly dependent.
PROOF: Let $\mathbf{A} = \begin{bmatrix} \mathbf{v_1} & \ldots & \mathbf{v_p} \end{bmatrix}$. Then the equation $\mathbf{Ax} = \mathbf{0}$ is a system of n equations in p unknowns. If $p > n$, there must be free variables, and, thus, a non-trivial solution. $\square$

Example 5 on page 69. Three vectors $\begin{bmatrix} 2 \\ 1 \end{bmatrix}$, $\begin{bmatrix} 4 \\ -1 \end{bmatrix}$, and $\begin{bmatrix} -2 \\ 2 \end{bmatrix}$ are linearly dependent, since $3 = p > n = 2$. The dependence relation is $-\begin{bmatrix} 2 \\ 1 \end{bmatrix} + \begin{bmatrix} 4 \\ -1 \end{bmatrix} + \begin{bmatrix} -2 \\ 2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$.

Theorem 9 on page 69. If a set of vectors contains the zero vector, the set is linearly dependent.
PROOF: Suppose $\mathbf{v_1} = \mathbf{0}$. Then $1 \cdot \mathbf{v_1} + 0 \cdot \mathbf{v_2} + \cdots + 0 \cdot \mathbf{v_p} = \mathbf{0}$ solves the equation nontrivially. $\square$

Example 6(b) on page 69. The set of vectors $\begin{bmatrix} 2 \\ 3 \\ 5 \end{bmatrix}$, $\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$, and $\begin{bmatrix} 1 \\ 1 \\ 8 \end{bmatrix}$ is linearly dependent since it contains the zero vector.

1.8. Linear Transformations.

Definition. A <u>transformation</u> (or <u>function</u> or <u>mapping</u>) T from $\mathbb{R}^n$ to $\mathbb{R}^m$ is a rule that assigns to each vector $\mathbf{x}$ in $\mathbb{R}^n$ a vector $T(\mathbf{x})$ in $\mathbb{R}^m$. Picture.

Definition. The set $\mathbb{R}^n$ is called the domain of T, and $\mathbb{R}^m$ is called the codomain of T. The notation $T : \mathbb{R}^n \to \mathbb{R}^m$ means that T maps $\mathbb{R}^n$ into $\mathbb{R}^m$. The vector $T(\mathbf{x})$ is called the image of $\mathbf{x}$. The set of all images is called the range of T. Picture.

A matrix equation $\mathbf{A}\mathbf{x} = \mathbf{b}$ can be considered as a transformation.

Definition. Let $\mathbf{x} \in \mathbb{R}^n$. The transformation $T(\mathbf{x}) = \mathbf{A}\mathbf{x}$ where $\mathbf{A}$ is a $m \times n$ matrix is called a matrix transformation and is denoted by $\mathbf{x} \mapsto \mathbf{A}\mathbf{x}$.

Example 1 on page 74. Define a matrix transformation $T : \mathbb{R}^2 \to \mathbb{R}^3$ by $T(\mathbf{x}) = \mathbf{A}\mathbf{x}$ where $\mathbf{A} = \begin{bmatrix} 1 & -3 \\ 3 & 5 \\ -1 & 7 \end{bmatrix}$.

(i) Find $T(\mathbf{u})$ where $\mathbf{u} = \begin{bmatrix} 2 \\ -1 \end{bmatrix}$.

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

(ii) Find $\mathbf{x}$ which image is $\begin{bmatrix} 3 \\ 2 \\ -5 \end{bmatrix}$.

SOLUTION: $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$ solves $\begin{bmatrix} 1 & -3 \\ 3 & 5 \\ -1 & 7 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 3 \\ 2 \\ -5 \end{bmatrix}$. Solution is $x_1 = 1.5, x_2 = -.5$ and is unique.

(iii) Determine if $\begin{bmatrix} 3 \\ 2 \\ 5 \end{bmatrix}$ is in the range. Answer: no.

Example 2 on page 76. The matrix transformation with

$$\mathbf{A} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

is a projection on the x_1x_2-plane since

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \mapsto \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} x_1 \\ x_2 \\ 0 \end{bmatrix}.$$

Define projections on x_1x_3- and x_2x_3-planes.

Definition. A transformation T is called linear if

(i) $T(\mathbf{u}+\mathbf{v}) = T(\mathbf{u}) + T(\mathbf{v})$ for any vectors $\mathbf{u}$ and $\mathbf{v}$ in the domain of T,
(ii) $T(c\mathbf{u}) = cT(\mathbf{u})$ for any vector $\mathbf{u}$ and any scalar c.

Example. A matrix transformation is linear.

Exercise. Show the superposition principle $T(c_1\mathbf{v_1}+\cdots+c_p\mathbf{v_p}) = c_1T(\mathbf{v_1})+\cdots+c_pT(\mathbf{v_p})$.

Exercise. Define $T : \mathbb{R}^2 \to \mathbb{R}^2$ by $T(\mathbf{u}) = rT(\mathbf{u})$. If $0 \le r \le 1$, T is called a contraction; if $r > 1$, T is called a dilation. Show that T is a linear transformation.

Exercise. Show that if T is a linear transformation, then $T(\mathbf{0}) = \mathbf{0}$.

1.9. The Matrix of a Linear Transformation.

It turns out that any linear transformation is in fact a matrix transformation.

Proposition. Let $T : \mathbb{R}^n \to \mathbb{R}^m$ be a linear transformation. Then there exists a unique matrix $\mathbf{A}$ such that $T(\mathbf{u}) = \mathbf{A}\mathbf{u}$ for any vector $\mathbf{u} \in \mathbb{R}^n$. The matrix $\underset{m\times n}{\mathbf{A}} = \begin{bmatrix} T(\mathbf{e_1}) & \dots & T(\mathbf{e_n}) \end{bmatrix}$ where $\underset{m\times 1}{\mathbf{e_1}} = \begin{bmatrix} 1 \\ 0 \\ \dots \\ 0 \end{bmatrix}$, etc.

Proof: $\mathbf{u} = \begin{bmatrix} u_1 \\ \dots \\ u_n \end{bmatrix} = u_1\mathbf{e_1} + \cdots + u_n\mathbf{e_n}$. Thus, $T(\mathbf{u}) = u_1T(\mathbf{e_1}) + \cdots + u_nT(\mathbf{e_n}) = \begin{bmatrix} T(\mathbf{e_1}) & \dots & T(\mathbf{e_n}) \end{bmatrix} \begin{bmatrix} u_1 \\ \dots \\ u_n \end{bmatrix}$. $\square$

Example 1 on page 82. Let $T : \mathbb{R}^2 \to \mathbb{R}^3$ be a linear transformation such that $T(\mathbf{e_1}) = \begin{bmatrix} 5 \\ -7 \\ 2 \end{bmatrix}$ and $T(\mathbf{e_2}) = \begin{bmatrix} -3 \\ 8 \\ 0 \end{bmatrix}$. Find $\mathbf{A}$, the standard matrix for T.

2.1. Matrix Operations.

Definition. An $m \times n$ matrix $\mathbf{A}$ has the form

$$\mathbf{A} = \begin{bmatrix} a_{11} & \dots & a_{1j} & \dots & a_{1n} \\ & & \dots & & \\ a_{i1} & \dots & a_{ij} & \dots & a_{in} \\ & & \dots & & \\ a_{m1} & \dots & a_{mj} & \dots & a_{mn} \end{bmatrix}$$

where a_{ij} is called the (i, j)-entry.

Definition. The diagonal entries of $\mathbf{A}$ are $a_{11}, a_{22}, \ldots$, and they form the main diagonal of $\mathbf{A}$.

Definition. A diagonal matrix is a square matrix with the zero entries off the main diagonal.

Definition. The identity matrix $\mathbf{I_n}$ is a diagonal matrix with ones on the main diagonal.

Definition. Two matrices are equal if they have the same dimensions and their corresponding entries are equal.

Definition. The addition of matrices of equal dimensions and the scalar multiplication of a matrix are defined entrywise.

Definition. If $\mathbf{A}$ is a $m \times n$ matrix and $\mathbf{B}$ is a $n \times p$ matrix, then the product $\mathbf{AB}$ is a $m \times p$ matrix with (i, j) entry

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

Example.

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

Remark. (i) The product of matrices is not commutative, that is, in general, $\mathbf{AB} \neq \mathbf{BA}$. Example.
(ii) The cancellation law doesn't hold for matrix multiplication, that is, if $\mathbf{AB} = \mathbf{AC}$, then, in general, $\mathbf{B} \neq \mathbf{C}$. Give example.
(iii) If $\mathbf{AB} = \mathbf{0}$, then, in general, it is not true that $\mathbf{A} = \mathbf{0}$ or $\mathbf{B} = \mathbf{0}$. Give example.

Definition. The kth power of a matrix $\mathbf{A}$ is defined as the product of k copies of $\mathbf{A}$, that is, $\mathbf{A}^k = \underbrace{\mathbf{A} \ldots \mathbf{A}}_{k\,times}$.

Definition. The transpose of a $n \times m$ matrix $\mathbf{A}$ is the $m \times n$ matrix $\mathbf{A}^T$ with a_{ji} as the (i, j) entry. Example.

Theorem 3 on page 115. (i) $(\mathbf{A}^T)^T = \mathbf{A}$, (ii) $(\mathbf{A} + \mathbf{B})^T = \mathbf{A}^T + \mathbf{B}^T$, (iii) $(r\mathbf{A})^T = r\mathbf{A}^T$ for any scalar r, and (iv) $(\mathbf{AB})^T = \mathbf{B}^T\mathbf{A}^T$.

Examples.

2.2. The Inverse of a Matrix.

Definition. The inverse of a square matrix $\mathbf{A}$ is the (unique) matrix $\mathbf{A}^{-1}$ such that $\mathbf{AA}^{-1} = \mathbf{A}^{-1}\mathbf{A} = \mathbf{I}$.

Theorem 4 on page 119. The inverse of a 2×2 matrix $\mathbf{A} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ is $\mathbf{A}^{-1} = \frac{1}{ad-bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$.
If $ad - bc = 0$, the inverse does not exist, and $\mathbf{A}$ is called <u>singular</u> or <u>not invertible</u>.
PROOF:

Theorem 5 on page 120. If $\mathbf{A}$ is invertible, then the matrix equation $\mathbf{Ax} = \mathbf{b}$ has the unique solution $\mathbf{x} = \mathbf{A}^{-1}\mathbf{b}$.

Example. Solve

$$\begin{cases} 3x_1 + 4x_2 = 3 \\ 5x_1 + 6x_2 = 7. \end{cases}$$

Theorem 6 on page 121. (i) $(\mathbf{A}^{-1})^{-1} = \mathbf{A}$, (ii) $(\mathbf{AB})^{-1} = \mathbf{B}^{-1}\mathbf{A}^{-1}$, and (iii) $(\mathbf{A}^T)^{-1} = (\mathbf{A}^{-1})^T$.
PROOF:

We know how to find the inverse of a 2×2 matrix. For larger matrices there is no explicit formula. However, one can find $\mathbf{A}^{-1}$ by applying the following algorithm based on the fact that $\mathbf{AA}^{-1} = \mathbf{I}$.

Algorithm for Finding $\mathbf{A}^{-1}$: Take the augmented matrix $\begin{bmatrix} \mathbf{A} & \mathbf{I} \end{bmatrix}$. Applying elementary operations reduce the matrix to $\begin{bmatrix} \mathbf{I} & \mathbf{A}^{-1} \end{bmatrix}$. If it is impossible, then $\mathbf{A}$ is not invertible.

Example 7 on page 124. Find the inverse of the matrix

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

SOLUTION:

$$\begin{bmatrix} \mathbf{A} & \mathbf{I} \end{bmatrix} = \begin{bmatrix} 0 & 1 & 2 & 1 & 0 & 0 \\ 1 & 0 & 3 & 0 & 1 & 0 \\ 4 & -3 & 8 & 0 & 0 & 1 \end{bmatrix} \Longrightarrow \begin{bmatrix} 1 & 0 & 0 & -9/2 & 7 & -3/2 \\ 0 & 1 & 0 & -2 & 4 & -1 \\ 0 & 0 & 1 & 3/2 & -2 & 1/2 \end{bmatrix}.$$

It is useful to check the answer.

3.1. Determinants.

Notation. Denote by $\mathbf{A}_{ij}$ the matrix obtained from matrix $\mathbf{A}$ by removing the ith row and the jth column.

Example. Find $\mathbf{A}_{11}$.

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

Definition. The determinant of a $n \times n$ matrix $\mathbf{A} = [a_{ij}]$ is a number given recursively by

$$det\mathbf{A} = \sum_{j=1}^{n} (-1)^{j+1} a_{1j} det\mathbf{A}_{1j}.$$

Definition. The quantity $(-1)^{i+j} det\mathbf{A}_{ij}$ is called the (i,j)-cofactor of $\mathbf{A}$.

Definition. The determinant is defined as a cofactor expansion across the first row.

Example. Find

$$det \begin{bmatrix} 1 & 5 & 0 \\ 2 & 4 & -1 \\ 0 & -2 & 0 \end{bmatrix}.$$

Theorem 1 on page 188. The determinant of a matrix can be computed by a cofactor expansion across any row or any column.

Example. In above example, it is simpler to expand the determinant along the last row or last column.

Exercise. Show that if one row or one column of a square matrix is zero, its determinant is zero (singular matrix).

Definition. A matrix is called triangular if all entries below or above the main diagonal are zeros. Example. Special case: diagonal matrix.

Theorem 2 on page 189. The determinant of a triangular matrix is the product of the entries on the main diagonal.

3.2. Properties of Determinants.

Theorem 3 on page 192. (i) If a multiple of one row a matrix $\mathbf{A}$ is added to another row to produce a matrix $\mathbf{B}$, then $det\mathbf{B} = det\mathbf{A}$.
(ii) If two rows of $\mathbf{A}$ are interchanged to produce $\mathbf{B}$, then $det\mathbf{B} = -det\mathbf{A}$.
(iii) If one row of $\mathbf{A}$ is multiplied by c to produce $\mathbf{B}$, then $det\mathbf{B} = cdet\mathbf{A}$.

Example 1 on page 193.

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

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

Exercise. Compute $det \begin{bmatrix} 0 & 2 & 2 \\ 1 & 0 & 3 \\ 2 & 1 & 1 \end{bmatrix}$. Answer: 12.

Exercise. Show that if two rows of a square matrix are equal, its determinant is zero.

Exercise. Show that if one column of a square matrix is a linear combination of the others, then its determinant is zero. The converse is also true.

Theorem 4 on page 194. A square matrix is invertible iff its determinant is nonzero.

Remark. It follows from the above Exercise and Theorem 4 that a matrix is invertible iff its determinant is nonzero iff its columns are linearly independent.

Theorem 5 on page 196. $det\mathbf{A}^T = det\mathbf{A}$.

Theorem 6 on page 196. $det\mathbf{AB} = (det\mathbf{A})(det\mathbf{B})$.

Exercise. $det\mathbf{A} = 5$. Find $det\mathbf{A}^{-1}$.

Exercise. A matrix $\mathbf{A}$ is such that $\mathbf{A} = \mathbf{A}^{-1}$. Find $det\mathbf{A}$.

3.3. Cramer's Rule.

Notation. Let $\mathbf{A}_i(\mathbf{b})$ denote the matrix obtained from a $n \times n$ matrix $\mathbf{A}$ by replacing its ith column by vector $\mathbf{b} \in \mathbb{R}^n$.

Theorem 7 on page 201 (Cramer's Rule). The unique solution of a matrix equation $\mathbf{Ax} = \mathbf{b}$ is

$$x_i = det\mathbf{A}_i(\mathbf{b})/det\mathbf{A},\ i = 1, \ldots, n.$$

PROOF: Let $\mathbf{A} = \begin{bmatrix} \mathbf{a_1} & \ldots & \mathbf{a_n} \end{bmatrix}$ and $\mathbf{I} = \begin{bmatrix} \mathbf{e_1} & \ldots & \mathbf{e_n} \end{bmatrix}$. If $\mathbf{Ax} = \mathbf{b}$, then $\mathbf{A}\,\mathbf{I}_i(\mathbf{x}) = \begin{bmatrix} \mathbf{Ae_1} & \ldots & \mathbf{Ax} & \ldots & \mathbf{Ae_n} \end{bmatrix} = \begin{bmatrix} \mathbf{a_1} & \ldots & \mathbf{b} & \ldots & \mathbf{a_n} \end{bmatrix} = \mathbf{A}_i(\mathbf{b})$. Thus, $det\big(\mathbf{A}\,\mathbf{I}_i(\mathbf{x})\big) = det\mathbf{A}\,det\mathbf{I}_i(\mathbf{x}) = det\mathbf{A}_i(\mathbf{b})$, but $det\mathbf{I}_i(\mathbf{x}) = x_i$. $\square$

Example 1 on page 202. Use Cramer's rule to solve

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

Answer: $x_1 = 20, x_2 = 27$.

Exercise 6 on page 209. Use Cramer's rule to solve

$$\begin{cases} 2x_1 + x_2 + x_3 = 4 \\ -x_1 + 2x_3 = 2 \\ 3x_1 + x_2 + 3x_3 = -2. \end{cases}$$

A Formula for $\mathbf{A}^{-1}$.

Notation. Denote by $C_{ij} = (-1)^{i+j} det\mathbf{A}_{ij}$, the (i, j)-cofactor of $\mathbf{A}$.

Theorem 8 on page 203. The inverse

$$\mathbf{A}^{-1} = \frac{1}{det\mathbf{A}} \begin{bmatrix} C_{11} & C_{21} & \dots & C_{n1} \\ C_{12} & C_{22} & \dots & C_{n2} \\ & & \dots & \\ C_{1n} & C_{2n} & \dots & C_{nn} \end{bmatrix}.$$

The matrix of cofactors is called the <u>adjugate</u> or <u>adjoint</u> of $\mathbf{A}$ and is denoted by $adj\mathbf{A}$. Thus, $\mathbf{A}^{-1} = \frac{1}{det\mathbf{A}} adj\mathbf{A}$.

Example 3 on pages 203 – 204. Find the inverse of

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

Answer:

$$\mathbf{A}^{-1} = (1/14) \begin{bmatrix} -2 & 14 & 4 \\ 3 & -7 & 1 \\ 5 & -7 & -3 \end{bmatrix}.$$

Exercise 12 on page 210. Find the inverse of

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

Determinants as Area or Volume.

Theorem 9 on page 205 (Geometric Interpretation of Determinants). The absolute value of the determinant of a 2×2 matrix is the area of the parallelogram determined by the columns of the matrix. The $|\cdot|$ of the determinant of a 3×3 matrix is the volume of the parallelepiped determined by the columns of the matrix.

Example. Draw the picture for $\left| det \begin{bmatrix} a & 0 \\ 0 & d \end{bmatrix} \right|$ and $\left| det \begin{bmatrix} a & 0 & 0 \\ 0 & b & 0 \\ 0 & 0 & c \end{bmatrix} \right|$.

Example 4 on page 206. Find the area of the parallelogram with vertices $(-2,-2)$, $(0,3)$, $(4,-1)$, and $(6,4)$.
SOLUTION: First move the parallelogram to the origin by subtracting $(-2,-2)$, for example. Then the area of the parallelogram with vertices $(0,0)$, $(2,5)$, $(6,1)$, and $(8,6)$ is $|det \begin{bmatrix} 2 & 6 \\ 5 & 1 \end{bmatrix}| = |-28| = 28$.

4.1. Vector Spaces and Subspaces.

Definition. A vector space V is a collection of objects, called vectors, for which two operations are defined: addition and multiplication by a scalar. For any $\mathbf{u}, \mathbf{v}, \mathbf{w} \in V$ and any $c, d \in \mathbb{R}$, the following ten axioms hold:
1. $\mathbf{u} + \mathbf{v} \in V$.
2. $\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}$ (commutative law for addition).
3. $(\mathbf{u} + \mathbf{v}) + \mathbf{w} = \mathbf{u} + (\mathbf{v} + \mathbf{w})$ (associative law for addition).
4. There exists a zero vector $\mathbf{0} \in V$ such that $\mathbf{u} + \mathbf{0} = \mathbf{u}$.
5. For every $\mathbf{u}$, there exists a vector $-\mathbf{u} \in V$, called the negative of $\mathbf{u}$ such that $\mathbf{u} + (-\mathbf{u}) = \mathbf{0}$.
6. Vector $c\mathbf{u} \in V$.
7. $c(\mathbf{u} + \mathbf{v}) = c\mathbf{u} + c\mathbf{v}$ (distributive law).
8. $(c + d)\mathbf{u} = c\mathbf{u} + d\mathbf{u}$ (distributive law).
9. $c(d\mathbf{u}) = (cd)\mathbf{u}$.
10. $1 \cdot \mathbf{u} = \mathbf{u}$.

Example 1 on page 217. The space $\mathbb{R}^n$ is a vector space. Each element is represented by a column vector.

Example 2 on pages 217 – 218. The space of all arrows is a vector space. Here two vectors are equal if they have the same length and point in the same direction. Addition is defined by the parallelogram rule, and multiplication by a scalar as contracting or dilating a vector by $|c|$ and reversing the direction if $c < 0$.

Example 3 on page 218. The space of all doubly infinite sequences of numbers $(\dots, x_{-2}, x_{-1}, x_0, x_1, x_2, \dots)$ is a vector space.

Example 4 on pages 218 – 219. The space of all polynomials of degree at most n is a vector space.

Example 5 on page 219. The space of all real-valued functions is a vector space.

Example. The space of $n \times n$ matrices is a vector space.

Example. The space of natural numbers $\mathbb{N} = \{1, 2, 3, \dots\}$ is not a vector space. No negative vector, no zero.

Example. The space of integer numbers $\mathbb{Z} = \{0, 1, -1, 2, -2, \dots\}$ is not a vector space. Multiplication by a real scalar doesn't work.

PROPOSITION 1. The negative of $\mathbf{u}$ is unique.

PROOF: Let $\mathbf{a}$ and $\mathbf{b}$ be two negatives of $\mathbf{u}$. Then, by axiom 5, $\mathbf{u} + \mathbf{a} = \mathbf{0}$. Since vectors commute by axiom 2, we can write $\mathbf{a} + \mathbf{u} = \mathbf{0}$, which, by axiom 5 again, means that $\mathbf{u} = -\mathbf{a}$. Similar, $\mathbf{u} = -\mathbf{b}$. So, $-\mathbf{a} = -\mathbf{b}$. Therefore, by axiom 5 once again, $\mathbf{b} + (-\mathbf{a}) = \mathbf{0}$. Adding $\mathbf{a}$ to both sides of this identity, we get that the left-hand side equals $\big(\mathbf{b} + (-\mathbf{a})\big) + \mathbf{a} = \{$by axiom 3 $\} = \mathbf{b} + \big((-\mathbf{a}) + \mathbf{a}\big) = \{$ by axiom 2$\} = \mathbf{b} + \big(\mathbf{a} + (-\mathbf{a})\big) = \{$by axiom 5$\} = \mathbf{b} + \mathbf{0} = \{$by axiom 4$\} = \mathbf{b}$, while the right-hand side equals $\mathbf{0} + \mathbf{a} = \{$by axiom 2$\} = \mathbf{a} + \mathbf{0} = \{$by axiom 4$\} = \mathbf{a}$. So, $\mathbf{a} = \mathbf{b}$. $\square$

PROPOSITION 2. $0 \cdot \mathbf{u} = \mathbf{0}$.

PROOF: By axiom 4, it suffices to show that $\mathbf{u} + 0 \cdot \mathbf{u} = \mathbf{u}$. Indeed, $\mathbf{u} + 0 \cdot \mathbf{u} = \{$by axiom 10$\} = 1 \cdot \mathbf{u} + 0 \cdot \mathbf{u} = \{$by axiom 8$\} = (1 + 0)\mathbf{u} = 1 \cdot \mathbf{u} = \{$by axiom 10 again$\} = \mathbf{u}$. $\square$

PROPOSITION 3. $-\mathbf{u} = (-1)\mathbf{u}$.

PROOF: By axiom 5, it suffices to show that $\mathbf{u} + (-1)\mathbf{u} = \mathbf{0}$. Indeed, $\mathbf{u} + (-1)\mathbf{u} = \{$by axiom 10$\} = 1 \cdot \mathbf{u} + (-1)\mathbf{u} = \{$by axiom 8$\} = (1 - 1)\mathbf{u} = 0 \cdot \mathbf{u} = \{$by proposition 2$\} = \mathbf{0}$. $\square$

PROPOSITION 4. $c\,\mathbf{0} = \mathbf{0}$.

PROOF: By axiom 5, $c\,\mathbf{0} = c\,\big(\mathbf{u} + (-\mathbf{u})\big) = \{$by axiom 7$\} = c\mathbf{u} + c(-\mathbf{u}) = \{$by proposition 3$\} = c\mathbf{u} + c\big((-1)\mathbf{u}\big) = \{$by axiom 9$\} = c\mathbf{u} + (c(-1))\mathbf{u} = c\mathbf{u} + ((-1)c)\mathbf{u} = \{$by axiom 9 again$\} = c\mathbf{u} + (-1)(c\mathbf{u}) = \{$by proposition 3$\} = c\mathbf{u} + \big(-(c\mathbf{u})\big) = \{$by axiom 5$\} = \mathbf{0}$. $\square$

Subspaces.

Definition. A <u>subspace</u> of a vector space V is a subset H with three properties: (i) has the zero vector, (ii) closed under vector addition, (iii) closed under scalar multiplication. These properties guarantee that a subspace is itself a vector space.

Example 6 on page 220. The set consisting only of the zero vector is a subspace. It is denoted by $\{\mathbf{0}\}$.

Example 7 on page 220. The set of all polynomials with real coefficients, $\mathbb{P}$ is a subspace of the space of real-valued functions. The set of all polynomials of degree at most n is a subspace of $\mathbb{P}$.

Example 8 on page 220. The set $\left\{ \begin{bmatrix} x_1 \\ x_2 \\ 0 \end{bmatrix} \right\}$ is a subspace of $\mathbb{R}^3$. It looks and acts like $\mathbb{R}^2$ but is not $\mathbb{R}^2$.

Example 10 on page 220. If $\mathbf{v_1}, \dots, \mathbf{v_p} \in V$, then $Span\{\mathbf{v_1}, \dots, \mathbf{v_p}\}$ is a subspace of V.

4.2. Null Spaces, Column Spaces, and Linear Transformations.

Definition. The null space of an $m \times n$ matrix $\mathbf{A}$ is the set of all solutions of the homogeneous equation $\mathbf{Ax} = \mathbf{0}$. In set notation,

$$Nul\mathbf{A} = \{\mathbf{x} : \, \mathbf{x} \in \mathbb{R}^n, \; \mathbf{Ax} = \mathbf{0}\}.$$

Geometrically, the null space of $\mathbf{A}$ is the set that is mapped into zero by the linear transformation $\mathbf{x} \mapsto \mathbf{Ax}$. Picture.

Theorem 2 on page 227. The null space of an $m \times n$ matrix is a subspace of $\mathbb{R}^n$.
PROOF: (i) $\mathbf{0} \in \mathbb{R}^n$ is in the null space of $\mathbf{A}$ since $\mathbf{A0} = \mathbf{0}$.
(ii) If $\mathbf{Au} = \mathbf{0}$ and $\mathbf{Av} = \mathbf{0}$, then $\mathbf{A}(\mathbf{u} + \mathbf{v}) = \mathbf{0}$.
(iii) If $\mathbf{Au} = \mathbf{0}$, then $\mathbf{A}(c\mathbf{u}) = c(\mathbf{Au}) = c(\mathbf{0}) = \mathbf{0}$. $\square$

Example. Find the null space of $\mathbf{A} = \begin{bmatrix} 1 & -3 & -2 \\ -5 & 9 & 1 \end{bmatrix}$.

SOLUTION: $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$ is in $Nul\mathbf{A}$ if it solves $\begin{cases} x_1 - 3x_2 - 2x_3 = 0 \\ -5x_1 + 9x_2 + x_3 = 0. \end{cases}$
Therefore, $Nul\mathbf{A} = \{\mathbf{x} \in \mathbb{R}^3 : \, x_1 = -(5/2)x_3, \; x_2 = -(3/2)x_3\}$. It is a subspace of $\mathbb{R}^3$.

An Explicit Description of $Nul\mathbf{A}$.

Example 3 on page 228. Find a spanning set for the null space of the matrix

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

SOLUTION: The general solution of the equation $\mathbf{Ax} = \mathbf{0}$ is $x_1 = 2x_2 + x_4 - 3x_5$, $x_3 = -2x_4 + 2x_5$, and x_2, x_4, and x_5 are free variables. We can now decompose any vector in $\mathbb{R}^5$ into a linear combination of vectors where weights are the free variables. That is,

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

Notice that $\mathbf{u}, \mathbf{v}$, and $\mathbf{w}$ are linearly independent since the weights are free variables. Thus, $Nul\mathbf{A} = Span\{\mathbf{u}, \mathbf{v}, \mathbf{w}\}$.

The Column Space of a Matrix.

Definition. The <u>column space</u> of an $m \times n$ matrix $\mathbf{a} = [\mathbf{a_1}, \mathbf{a_2}, \dots, \mathbf{a_n}]$ is the set of all linear combinations of its columns, that is,

$$Col\mathbf{A} = Span\{\mathbf{a_1}, \mathbf{a_2}, \dots, \mathbf{a_n}\} = \{\mathbf{b} : \mathbf{b} = \mathbf{Ax} \text{ for some } \mathbf{x} \in \mathbb{R}^n\}.$$

Theorem 3 on page 229. $Col\mathbf{A}$ is a subspace of $\mathbb{R}^m$.
PROOF: exercise.

Example 4 on pages 229 − 230. Find a matrix $\mathbf{A}$ such that $\mathbf{W} = Col\mathbf{A}$ where $\mathbf{W} = \left\{ \begin{bmatrix} 6a - b \\ a + b \\ -7a \end{bmatrix} : a, b \in \mathbb{R} \right\}$.
SOLUTION:

$$\mathbf{W} = \left\{ a \begin{bmatrix} 6 \\ 1 \\ -7 \end{bmatrix} + b \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} : a, b \in \mathbb{R} \right\} = Span\left\{ \begin{bmatrix} 6 \\ 1 \\ -7 \end{bmatrix}, \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} \right\}.$$

Thus, the matrix is

$$\mathbf{A} = \begin{bmatrix} 6 & -1 \\ 1 & 1 \\ -7 & 0 \end{bmatrix}.$$

Kernel and Range of a Linear Transformation.

Recall that a linear transformation preserves vector addition and multiplication by a scalar, that is, $T(\mathbf{u} + \mathbf{v}) = T(\mathbf{u}) + T(\mathbf{v})$ and $T(c\mathbf{u}) = cT(\mathbf{u})$.

Definition. The <u>kernel</u> of a linear transformation is the set of vectors that are mapped into zero. That is, for $T : V \to W$ the kernel is the set of all $\mathbf{v} \in V$ such that $T(\mathbf{v}) = \mathbf{0}$.

Definition. The <u>range</u> of a linear transformation is the image under the mapping, that is, the range of T is the set of all vectors $\mathbf{w} \in W$ such that $\mathbf{w} = T(\mathbf{v})$ for some $\mathbf{v} \in V$.

Picture.

Remark. Recall that any linear transformation is a matrix transformation, that is, $T(\mathbf{x}) = \mathbf{Ax}$ for some $\mathbf{A}$. Therefore, by definition, the kernel of T is the null space of $\mathbf{A}$, and the range of T is the column space of $\mathbf{A}$.

4.3. Bases.

Definition. A set of vectors $\mathfrak{B} = \{\mathbf{b_1}, \ldots, \mathbf{b_p}\}$ is a basis for a vector space V if (i) $\mathfrak{B}$ is a linearly independent set, and (ii) the vector space is spanned by $\mathfrak{B}$, that is, $V = Span\{\mathbf{b_1}, \ldots, \mathbf{b_p}\}$.

Example 3 on page 238. Let $\mathbf{A}$ be an invertible $n \times n$ matrix. Then the columns of $\mathbf{A}$ form a basis for $\mathbb{R}^n$ because they are linearly independent and they span $\mathbf{R}^n$. Explain more.

Example 4 on page 238. The set

$$\left\{\mathbf{e_1} = \begin{bmatrix} 1 \\ 0 \\ \cdots \\ 0 \end{bmatrix}, \mathbf{e_2} = \begin{bmatrix} 0 \\ 1 \\ \cdots \\ 0 \end{bmatrix}, \ldots, \mathbf{e_n} = \begin{bmatrix} 0 \\ 0 \\ \cdots \\ 1 \end{bmatrix}\right\}$$

is called the standard basis for $\mathbb{R}^n$. Picture.

Example 5 on page 238. Determine whether $\left\{\mathbf{u} = \begin{bmatrix} 3 \\ 0 \\ -6 \end{bmatrix}, \mathbf{v} = \begin{bmatrix} -4 \\ 1 \\ 7 \end{bmatrix}, \mathbf{w} = \begin{bmatrix} -2 \\ 1 \\ 5 \end{bmatrix}\right\}$ is a basis for $\mathbb{R}^3$.

Solution: These vectors are linearly independent iff $det \begin{bmatrix} 3 & -4 & -2 \\ 0 & 1 & 1 \\ -6 & 7 & 5 \end{bmatrix} \neq 0$, which is true. Thus, by Example 3, this set is a basis.

Example 6 on pages 238 – 239. Show that $\{1, t, t^2, \ldots, t^n\}$ for a basis, called standard basis, for $\mathbb{P}_n$, the set of all polynomial of degree at most n.

Theorem 5 on page 239 (The Spanning Set Theorem).

Suppose a vector space $V = Span\{\mathbf{v_1}, \ldots, \mathbf{v_p}\}$. If one of the vectors, say $\mathbf{v_1}$, is a linear combination of the others, then the set with this vector removed still spans V, that is, $V = Span\{\mathbf{v_2}, \ldots, \mathbf{v_p}\}$. The set with all such vectors removed is a basis for V since it contains only linearly independent vectors spanning V.

Example 7 on page 239. Let $\mathbf{v_1} = \begin{bmatrix} 0 \\ 2 \\ -1 \end{bmatrix}$, $\mathbf{v_2} = \begin{bmatrix} 2 \\ 2 \\ 0 \end{bmatrix}$, and $\mathbf{v_3} = \begin{bmatrix} 6 \\ 16 \\ -5 \end{bmatrix}$, and suppose $V = Span\{\mathbf{v_1}, \mathbf{v_2}, \mathbf{v_3}\}$. Find a basis for V.
SOLUTION: Note that $\mathbf{v_3} = 5\mathbf{v_1} + 3\mathbf{v_2}$. By the theorem, $\{\mathbf{v_1}, \mathbf{v_2}\}$ is a basis for V.

Bases for $Nul\mathbf{A}$.

Recall Example 3 on page 228 discussed on the last lecture. We learned how to produce linearly independent set of vectors that spans the null space of a matrix. That is, we learned how to find a basis.

Bases for $Col\mathbf{A}$.

Example 8 on page 240. Find a basis for $Col\mathbf{B}$ where

$$\mathbf{B} = \begin{bmatrix} 1 & 4 & 0 & 2 & 0 \\ 0 & 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} \mathbf{b_1} & \dots & \mathbf{b_5} \end{bmatrix}.$$

SOLUTION: Each nonpivot column of $\mathbf{B}$ is a linear combination of the pivot columns. In fact, the pivot columns are $\mathbf{b_1}, \mathbf{b_3}$, and $\mathbf{b_5}$, and the nonpivot columns are $\mathbf{b_2} = 4\mathbf{b_1}$, and $\mathbf{p_4} = 2\mathbf{b_1} - \mathbf{b_3}$. By the Spanning Set Theorem, we may discard the nonpivot columns. The pivot columns are linearly independent and form a basis for $Col\mathbf{B}$.

Notice that the matrix $\mathbf{B}$ is in the reduced row echelon form. Suppose a matrix is not in this form. How to find a basis for its column space?

Theorem 6 on page 241. (i) Elementary row operations on a matrix do not affect the linear dependence relations among the columns of the matrix. (ii) The pivot columns of a matrix form a basis for its column space.

Example 9 on page 241. Find a basis for $Col\mathbf{A}$ where

$$\mathbf{A} = \begin{bmatrix} 1 & 4 & 0 & 2 & -1 \\ 3 & 12 & 1 & 5 & 5 \\ 2 & 8 & 1 & 3 & 2 \\ 5 & 20 & 2 & 8 & 8 \end{bmatrix}.$$

SOLUTION: It can be shown that $\mathbf{A}$ is row equivalent to the matrix $\mathbf{B}$ in Example 8. Thus, the columns $\mathbf{a_1}, \mathbf{a_3}$, and $\mathbf{a_5}$ are the pivot columns and form a basis.

4.4. Coordinate Systems.

Theorem 7 on page 246 (The Unique Representation Theorem). Let $\mathfrak{B} = \{\mathbf{b_1}, \dots, \mathbf{b_n}\}$ be a basis for a vector space V. Then for any $\mathbf{x} \in V$, there exists a unique set of scalars $c_1, \dots, c_n$ such that $\mathbf{x} = c_1\mathbf{b_1} + \cdots + c_n\mathbf{b_n}$.

PROOF: by contradiction, use linear independence.

Definition. The coordinates of $\mathbf{x}$ relative to a basis $\mathfrak{B} = \{\mathbf{b_1}, \dots, \mathbf{b_n}\}$ (or $\mathfrak{B}$-coordinates of $\mathbf{x}$) are the weights $c_1, \dots, c_n$ such that $\mathbf{x} = c_1\mathbf{b_1} + \cdots + c_n\mathbf{b_n}$.

Definition. The vector in $\mathbb{R}^n$

$$[\mathbf{x}]_{\mathfrak{B}} = \begin{bmatrix} c_1 \\ \dots \\ c_n \end{bmatrix}$$

is the coordinate vector of $\mathbf{x}$ (relative to $\mathfrak{B}$) or the $\mathfrak{B}$-coordinate vector of $\mathbf{x}$. The mapping $\mathbf{x} \mapsto [\mathbf{x}]_{\mathfrak{B}}$ is the coordinate mapping determined by $\mathfrak{B}$.

Examples 1 and 2 on page 247. Consider $\mathfrak{B} = \{\mathbf{b_1} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \mathbf{b_2} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}\}$. It is a basis for $\mathbb{R}^2$. Suppose a vector $\mathbf{x} \in \mathbb{R}^2$ has the coordinate vector $[\mathbf{x}]_{\mathfrak{B}} = \begin{bmatrix} -2 \\ 3 \end{bmatrix}$. Find $\mathbf{x}$.

SOLUTION: $\mathbf{x} = c_1\mathbf{b_1} + c_2\mathbf{b_2} = \begin{bmatrix} 1 \\ 6 \end{bmatrix}$. Note that the entries of this vector are the coordinates of $\mathbf{x}$ relative to the standard basis $\left\{\mathbf{e_1} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \mathbf{e_2} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}\right\}$.
Indeed, $\mathbf{x} = \begin{bmatrix} 1 \\ 6 \end{bmatrix} = 1 \cdot \begin{bmatrix} 1 \\ 0 \end{bmatrix} + 6 \cdot \begin{bmatrix} 0 \\ 1 \end{bmatrix} = 1 \cdot \mathbf{e_1} + 6 \cdot \mathbf{e_2}$.
Graphically, the coordinates can be interpreted in the following way. Plot the standard basis and $\mathbf{x}$, plot $\mathfrak{B}$. In the $\mathfrak{B}$-graph paper, $\mathbf{x}$ has coordinates $\begin{bmatrix} -2 \\ 3 \end{bmatrix}$.

Example 4 on page 249. Let $\mathbf{b_1} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}, \mathbf{b_2} = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$, and $\mathbf{x} = \begin{bmatrix} 4 \\ 5 \end{bmatrix}$. Find the $\mathfrak{B}$-coordinates of $\mathbf{x}$.
SOLUTION: $[\mathbf{x}]_{\mathfrak{B}} = \begin{bmatrix} c_1 \\ c_2 \end{bmatrix}$ where $\mathbf{x} = \begin{bmatrix} 4 \\ 5 \end{bmatrix} = c_1\mathbf{b_1} + c_2\mathbf{b_2} = c_1 \begin{bmatrix} 2 \\ 1 \end{bmatrix} + c_2 \begin{bmatrix} -1 \\ 1 \end{bmatrix}$.
Solving the system, obtain $c_1 = 3, c_2 = 2$.
Note that the augmented matrix is $\begin{bmatrix} \mathbf{b_1} & \mathbf{b_2} & \mathbf{x} \end{bmatrix}$. This can be generalized for $\mathbb{R}^n$.

Definition. Consider a basis $\mathfrak{B} = \{\mathbf{b_1}, \dots, \mathbf{b_n}\}$ in $\mathbb{R}^n$. Let $P_{\mathfrak{B}} = \begin{bmatrix} \mathbf{b_1} & \dots & \mathbf{b_n} \end{bmatrix}$. The change-of-coordinates equation $\mathbf{x} = c_1\mathbf{b_1} + \cdots + c_n\mathbf{b_n}$ is equivalent to $\mathbf{x} = P_{\mathfrak{B}}\,[\mathbf{x}]_{\mathfrak{B}}$. The matrix $P_{\mathfrak{B}}$ is called the change-of-coordinates matrix from

$\mathfrak{B}$ to the standard basis in $\mathbb{R}^n$. The coordinate mapping from the standard basis to $\mathfrak{B}$ is given by $[\mathbf{x}]_{\mathfrak{B}} = P_{\mathfrak{B}}^{-1}\mathbf{x}$. Note that $P_{\mathfrak{B}}$ is invertible since its columns are linearly independent.

The Coordinate Mapping.

Definition (on page 87). A mapping $T : U \to V$ is said to be <u>onto V</u> if each vector in V is the image of at least one vector in U.
Picture.

Definition (on page 87). A mapping $T : U \to V$ is said to be <u>one-to-one</u> if each vector in V is the image of at most one vector in U.
Picture.

Theorem 8 on page 250. Let $\mathfrak{B} = \{\mathbf{b_1}, \ldots, \mathbf{b_n}\}$ be a basis for a vector space V. Then the coordinate mapping $\mathbf{x} \mapsto [\mathbf{x}]_{\mathfrak{B}}$ is a <u>one-to-one</u> linear transformation from V <u>onto</u> $\mathbb{R}^n$.
PROOF: The coordinate mapping preserves vector addition and scalar multiplication (show!). Hence it is a linear transformation. Show the rest.

Definition. A one-to-one linear transformation from a vector space V onto a vector space W is called an <u>isomorphism</u> between V and W. The vector spaces are called <u>isomorphic</u>. The word "isomorphism" is from Greek "iso"="the same" and "morph"="structure". The name reflects the fact that two isomorphic spaces are indistinguishable as vector spaces. Every vector space calculation in V is accurately reproduced in W and vice versa.

The above theorem states that the possibly unfamiliar space V is isomorphic to the familiar $\mathbb{R}^n$.

Example 5 on page 251. Let $\mathfrak{B} = \{1, t, t^2 t^3\}$ be the standard basis for $\mathbb{P}_3$. A typical vector in $\mathbb{P}_3$ has the form $\mathbf{p}(t) = a_0 + a_1 t + a_2 t^2 + a_3 t^3$ which corresponds to the coordinate vector $[\mathbf{p}]_{\mathfrak{B}} = \begin{bmatrix} a_0 \\ a_1 \\ a_2 \\ a_3 \end{bmatrix}$ in $\mathbb{R}^4$. By the theorem, the coordinate mapping $\mathbf{p} \mapsto [\mathbf{p}]_{\mathfrak{B}}$ is an isomorphism between $\mathbb{P}_3$ and $\mathbb{R}^4$.

4.5. The Dimension of a Vector Space.

Theorem 9 on page 256. If a vector space V has a basis $\mathfrak{B} = \{\mathbf{b_1}, \ldots, \mathbf{b_n}\}$, then any set in V containing more than n vectors must be linearly dependent. This implies that any linearly independent set in V has no more than n vectors.
PROOF: Suppose $\{\mathbf{u_1}, \ldots, \mathbf{u_p}\}$ is a set in V and $p > n$. Then the set $\{[\mathbf{u_1}]_{\mathfrak{B}}, \ldots, [\mathbf{u_p}]_{\mathfrak{B}}\}$ is a linearly dependent set in $\mathbb{R}^n$ since there are more

vectors (p) than entries (n) in each vector. Therefore, there exist scalars $c_1, \ldots, c_p$, not all zeros, such that

$$\begin{bmatrix} 0 \\ \ldots \\ 0 \end{bmatrix} = c_1[\mathbf{u_1}]_{\mathfrak{B}} + \cdots + c_p[\mathbf{u_p}]_{\mathfrak{B}}$$

$=$ {the coordinate mapping is a linear transformation} $= [c_1\mathbf{u_1} + \cdots + c_p\mathbf{u_p}]_{\mathfrak{B}}$.

Hence, $c_1\mathbf{u_1} + \cdots + c_p\mathbf{u_p} = \mathbf{0}$. The scalars are not all zero, therefore, the set $\{\mathbf{u_1}, \ldots, \mathbf{u_p}\}$ is linearly dependent. □

Theorem 10 on page 257. If a vector space V has a basis consisting of n vectors, then every basis in V must have exactly n vectors.
PROOF: Let $\mathfrak{B}_1$ be a basis consisting of n vectors, and $\mathfrak{B}_2$ be any other basis. Since $\mathfrak{B}_1$ is a basis and $\mathfrak{B}_2$ is linearly independent, $\mathfrak{B}_2$ has no more than n vectors by Theorem 9. Also, since $\mathfrak{B}_2$ is a basis and $\mathfrak{B}_1$ is linearly independent, $\mathfrak{B}_2$ has at least n vectors. Thus, $\mathfrak{B}_2$ consist of exactly n vectors. □

Definition. The dimension of a vector space V, $dimV$, is the number of vectors in a basis of V. The dimension of the zero vector space $\{\mathbf{0}\}$ is defined to be zero. If V is not spanned by a finite number of vectors, then $dimV = \infty$.

Remark. This is a consistent definition since by Theorem 10, all bases in V have the same number of vectors.

Example 1 on page 257. $dim\mathbb{R}^n = n$ since the standard basis consists of n vectors. $dim\mathbb{P}_n = n+1$ since the set $\{1, t, t^2, \ldots, t^n\}$ is a basis. $dim\mathbb{P} = \infty$.

Example 3 on page 258. Find the dimension of the subspace of $\mathbb{R}^4$

$$H = \left\{ \begin{bmatrix} a - 3b + 6c \\ 5a + 4d \\ b - 2c - d \\ 5d \end{bmatrix} : a, b, c, d \in \mathbb{R} \right\}.$$

SOLUTION: Note that

$$H = Span\left\{ \mathbf{v_1} = \begin{bmatrix} 1 \\ 5 \\ 0 \\ 0 \end{bmatrix}, \mathbf{v_2} = \begin{bmatrix} -3 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \mathbf{v_3} = \begin{bmatrix} 6 \\ 0 \\ -2 \\ 0 \end{bmatrix}, \mathbf{v_4} = \begin{bmatrix} 0 \\ 4 \\ -1 \\ 5 \end{bmatrix} \right\}.$$

Since $\mathbf{v_3} = -2\mathbf{v_2}$ and the other vectors are linearly independent, $H = Span\{\mathbf{v_1}, \mathbf{v_2}, \mathbf{v_4}\}$. Thus, $dimH = 3$.

Theorem 12 on page 259 (The Basis Theorem). Let V be a p-dimensional vector space, $p \geq 1$. Any linearly independent set of exactly p vectors in V is a basis in V. Any set of exactly p vectors that spans V is a basis in V.

Proposition. The dimension of $Nul\mathbf{A}$ is the number of free variables in the equation $\mathbf{Ax} = \mathbf{0}$. The dimension of $Col\mathbf{A}$ is the number of pivot columns in $\mathbf{A}$.

Example 5 on page 260. Find the dimensions of the null space and the column space of

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

SOLUTION: Row reduce the augmented matrix $\begin{bmatrix} \mathbf{A} & \mathbf{0} \end{bmatrix}$ to echelon form

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

There are three free variables – x_2, x_4, and x_5. Hence $dimNul\mathbf{A} = 3$. Also, $dimCol\mathbf{A} = 2$ because $\mathbf{A}$ has two pivot columns.

4.6. Rank.

Definition. The rank of a matrix $\mathbf{A}$ is the dimension of the column space of $\mathbf{A}$. That is , $rank\mathbf{A} = dimCol\mathbf{A}$.

Theorem 14 on page 265 (The Rank Theorem).

$$rank\mathbf{A} + dimNul\mathbf{A} = n.$$

"PROOF:" The rank of $\mathbf{A}$ is the number of pivot columns. They correspond to the non-free variables. The other variables are free. □

Note that in Example 5, $rank\mathbf{A} + dimNul\mathbf{A} = 2 + 3 = 5$.

5.1. Eigenvectors and Eigenvalues.

Definition. An eigenvector of an $n \times n$ matrix $\mathbf{A}$ is a nonzero vector $\mathbf{x}$ such that $\mathbf{Ax} = \lambda\mathbf{x}$ where the scalar λ is called the eigenvalue of $\mathbf{A}$ corresponding to the eigenvector $\mathbf{x}$.
In other words, the matrix transformation $\mathbf{x} \mapsto \mathbf{Ax}$ stretches or shrinks $\mathbf{x}$.

Definition. The eigenspace of a matrix corresponding to eigenvalue λ is the space consisting of the zero vector and all the eigenvectors corresponding to λ.

Example 2 on page 303. Let $\mathbf{A} = \begin{bmatrix} 1 & 6 \\ 5 & 2 \end{bmatrix}$, $\mathbf{u} = \begin{bmatrix} 6 \\ -5 \end{bmatrix}$, and $\mathbf{v} = \begin{bmatrix} 3 \\ -2 \end{bmatrix}$.

Are $\mathbf{u}$ and $\mathbf{v}$ eigenvectors of $\mathbf{A}$?
SOLUTION: $\mathbf{Au} = -4\mathbf{u}$ and $\mathbf{Av} \neq \lambda\mathbf{v}$.

5.2. The Characteristic Equation.

To find eigenvalues of a matrix $\mathbf{A}$, one has to find all λ's such that the equation $\mathbf{Ax} = \lambda\mathbf{x}$ or, equivalently, $(\mathbf{A} - \lambda\mathbf{I})\mathbf{x} = \mathbf{0}$ has a nontrivial solution. This problem is equivalent to finding all λ's such that the matrix $\mathbf{A} - \lambda\mathbf{I}$ is not invertible, which is equivalent to solving the characteristic equation of $\mathbf{A}$, $det(\mathbf{A} - \lambda\mathbf{I}) = 0$.

Definition. If $\mathbf{A}$ is an $n \times n$ matrix, then $det(\mathbf{A} - \lambda\mathbf{I})$ is a polynomial of degree n called the characteristic polynomial of $\mathbf{A}$.

Example 3 on pages 313 – 314. Find the characteristic equation of

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

ANSWER: The characteristic equation is $(\lambda - 1)(\lambda - 3)(\lambda - 5)^2 = 0$.

Definition. The multiplicity of an eigenvalue λ_1 is its multiplicity as a root of the characteristic equation, that is, if the characteristic polynomial of $\mathbf{A}$ has factor $(\lambda - \lambda_1)^k$, then the multiplicity of λ_1 is k.

In our example, $\lambda_1 = 1$ and $\lambda_2 = 3$ have multiplicities 1, while $\lambda_3 = 5$ has multiplicity 2.

Example. Find eigenvalues and eigenvectors of $\mathbf{A} = \begin{bmatrix} 1 & 5 & 0 \\ 0 & 4 & -1 \\ 0 & 0 & -2 \end{bmatrix}$.

SOLUTION: Solve

$$0 = det \begin{bmatrix} 1-\lambda & 5 & 0 \\ 0 & 4-\lambda & -1 \\ 0 & 0 & -2-\lambda \end{bmatrix} = (1-\lambda)(4-\lambda)(-2-\lambda).$$

The solution is $\lambda_1 = -2, \lambda_2 = 1$, and $\lambda_3 = 4$. Note that the eigenvalues of a diagonal matrix are the diagonal entries.
The eigenvector $\mathbf{x_1}$ corresponding to $\lambda_1 = -2$ solves $(\mathbf{A} + 2\mathbf{I})\mathbf{x} = \mathbf{0}$. The augmented matrix of this matrix equation is

$$\begin{bmatrix} 3 & 5 & 0 & 0 \\ 0 & 6 & -1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \Longrightarrow \begin{bmatrix} 1 & 0 & 5/18 & 0 \\ 0 & 1 & -1/6 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

The solution is $\mathbf{x_1} = \begin{bmatrix} -(5/18)a \\ (1/6)a \\ a \end{bmatrix}$ where a is free.

Eigenvector $\mathbf{x_2}$ solves $(\mathbf{A} - \mathbf{I})\mathbf{x} = \mathbf{0}$. The augmented matrix is

$$\begin{bmatrix} 0 & 5 & 0 \\ 0 & 3 & -1 \\ 0 & 0 & -3 \end{bmatrix} \Longrightarrow \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}.$$

The solution is $\mathbf{x_2} = \begin{bmatrix} b \\ 0 \\ 0 \end{bmatrix}$ where b is free.

Eigenvector $\mathbf{x_3}$ solves $(\mathbf{A} - 4\mathbf{I})\mathbf{x} = \mathbf{0}$. The augmented matrix is

$$\begin{bmatrix} -3 & 5 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & -6 & 0 \end{bmatrix} \Longrightarrow \begin{bmatrix} 0 & -5/3 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}.$$

The solution is $\mathbf{x_3} = \begin{bmatrix} (5/3)c \\ c \\ 0 \end{bmatrix}$ where c is free.

For example, we may take $\mathbf{x_1} = \begin{bmatrix} -5 \\ 3 \\ 18 \end{bmatrix}$, $\mathbf{x_2} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$, and $\mathbf{x_3} = \begin{bmatrix} 5 \\ 3 \\ 0 \end{bmatrix}$.

Theorem 2 on page 307. If $\mathbf{v_1}, \ldots, \mathbf{v_r}$ are eigenvectors corresponding to distinct eigenvalues $\lambda_1, \ldots, \lambda_r$ of a matrix $\mathbf{A}$, then the vectors are linearly independent.

Indeed, in our example, consider the linear combination $\alpha\mathbf{x_1} + \beta\mathbf{x_2} + \gamma\mathbf{x_3} =$
$\alpha \begin{bmatrix} -5 \\ 3 \\ 18 \end{bmatrix} + \beta \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + \gamma \begin{bmatrix} -5 \\ 3 \\ 0 \end{bmatrix} = 0 \Longleftrightarrow \alpha = \beta = \gamma = 0.$

Exercise 14 on page 317. Find eigenvalues and eigenvectors of

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

SOLUTION: $det(\mathbf{A} - \lambda\mathbf{I}) = -(\lambda + 4)(\lambda - 1)(\lambda - 7) = 0$, thus, the eigenvalues are $\lambda_1 = -4$, $\lambda_2 = 1$, and $\lambda_3 = 7$. The eigenvector $\mathbf{x_1}$ solves

$$\begin{bmatrix} 9 & -2 & 3 & 0 \\ 0 & 5 & 0 & 0 \\ 6 & 7 & 2 & 0 \end{bmatrix} \Longrightarrow \mathbf{x_1} = \begin{bmatrix} -(1/3)a \\ 0 \\ a \end{bmatrix}.$$

The eigenvector $\mathbf{x_2}$ solves

$$\begin{bmatrix} 4 & -2 & 3 & 0 \\ 0 & 0 & 0 & 0 \\ 6 & 7 & -3 & 0 \end{bmatrix} \Longrightarrow \mathbf{x_2} = \begin{bmatrix} -(3/8)b \\ 3b/4 \\ b \end{bmatrix}.$$

The eigenvector $\mathbf{x_3}$ solves

$$\begin{bmatrix} -2 & -2 & 3 & 0 \\ 0 & -6 & 0 & 0 \\ 6 & 7 & -9 & 0 \end{bmatrix} \Longrightarrow \mathbf{x_3} = \begin{bmatrix} (3/2)c \\ 0 \\ c \end{bmatrix}.$$

For example, we can take $\mathbf{x_1} = \begin{bmatrix} -1 \\ 0 \\ 3 \end{bmatrix}$, $\mathbf{x_2} = \begin{bmatrix} -3 \\ 6 \\ 8 \end{bmatrix}$, and $\mathbf{x_3} = \begin{bmatrix} 3 \\ 0 \\ 2 \end{bmatrix}$. Again, they are linearly independent.

5.3. Diagonalization.

Definition. A square matrix $\mathbf{A}$ is diagonalizable if there exist an invertible matrix $\mathbf{P}$ and a diagonal matrix $\mathbf{D}$ such that $\mathbf{A} = \mathbf{PDP}^{-1}$. The matrix $\mathbf{A}$ is said to be similar to $\mathbf{D}$.

Example 2 on page 320. Let $\mathbf{A} = \begin{bmatrix} 7 & 2 \\ -4 & 1 \end{bmatrix}$, $\mathbf{P} = \begin{bmatrix} 1 & 1 \\ -1 & -2 \end{bmatrix}$, and $\mathbf{D} = \begin{bmatrix} 5 & 0 \\ 0 & 3 \end{bmatrix}$. Check that $\mathbf{A} = \mathbf{PDP}^{-1}$ and find $\mathbf{A}^3$.

SOLUTION: Check $\mathbf{AP} = \mathbf{PD}$.

$\mathbf{A}^3 = \mathbf{PD}^3\mathbf{P}^{-1} = \begin{bmatrix} 1 & 1 \\ -1 & -2 \end{bmatrix} \begin{bmatrix} 125 & 0 \\ 0 & 27 \end{bmatrix} \begin{bmatrix} 2 & 1 \\ -1 & -1 \end{bmatrix} = \begin{bmatrix} 223 & 98 \\ -196 & -71 \end{bmatrix}$.

Theorem 5 on page 320 (The Diagonalization Theorem). An $n \times n$ matrix $\mathbf{A}$ is diagalizable iff it has n linearly independent eigenvectors. In fact, $\mathbf{A} = \mathbf{PDP}^{-1}$ iff the columns of $\mathbf{P}$ are n linearly independent eigenvectors of $\mathbf{A}$, and the diagonal entries of $\mathbf{D}$ are the corresponding eigenvalues of $\mathbf{A}$.

In our example, $det(\mathbf{A} - \lambda\mathbf{I}) = (7-\lambda)(1-\lambda) + 8 = (\lambda - 5)(\lambda - 3) = 0$, etc.

Example 3 on pages 321 – 322. If possible, diagonalize the following matrix

$$\mathbf{A} = \begin{bmatrix} 1 & 3 & 3 \\ -3 & -5 & -3 \\ 3 & 3 & 1 \end{bmatrix}.$$

SOLUTION: The characteristic equation is $-(\lambda - 1)(\lambda + 2)^2 = 0$. Hence, the eigenvalues are 1 and -2. Three linearly independent eigenvectors are $\mathbf{v_1} = \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}$, $\mathbf{v_2} = \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}$, and $\mathbf{v_3} = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}$.

Example 4 on page 323. Diagonalize $\mathbf{A} = \begin{bmatrix} 2 & 4 & 3 \\ -4 & -6 & -3 \\ 3 & 3 & 1 \end{bmatrix}$.

SOLUTION: The characteristic equation is $-(\lambda - 1)(\lambda + 2)^2 = 0$. Hence, the eigenvalues are 1 and -2. We can find only two linearly independent vectors $\mathbf{v_1} = \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}$ and $\mathbf{v_2} = \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}$, so $\mathbf{A}$ is not diagonalizable.

Theorem 6 on page 323. An $n \times n$ matrix with n distinct eigenvalues is diagonalizable.

5.4. Eigenvectors and Linear Transformations.

Suppose $T : V \to W$ where $dimV = n$, and $dimW = m$. Let $\mathfrak{B} = \{\mathbf{b_1}, \dots, \mathbf{b_n}\}$ and $\mathfrak{C}$ be two bases in V and W, respectively. Given $x \in V$, the coordinate vector $[\mathbf{x}]_{\mathfrak{B}} \in \mathbb{R}^n$, and the coordinate vector of $T(\mathbf{x}) \in W$, $[T(\mathbf{x})]_{\mathfrak{C}}$ is in $\mathbb{R}^m$. The connection between $[\mathbf{x}]_{\mathfrak{B}}$ and $[T(\mathbf{x})]_{\mathfrak{C}}$ is as follows. If $\mathbf{x} = r_1\mathbf{b_1} + \cdots + r_n\mathbf{b_n}$, then $[\mathbf{x}]_{\mathfrak{B}} = \begin{bmatrix} r_1 \\ \dots \\ r_n \end{bmatrix}$ and $T(\mathbf{x}) = r_1 T(\mathbf{b_1}) + \cdots + r_n T(\mathbf{b_n})$ because T is linear.
In terms of the $\mathfrak{C}$-coordinate vectors $[T(\mathbf{x})]_{\mathfrak{C}} = r_1[T(\mathbf{b_1})]_{\mathfrak{C}} + \cdots + r_n[T(\mathbf{b_n})]_{\mathfrak{C}}$.
Hence,

$$[T(\mathbf{x})]_{\mathfrak{C}} = \mathbf{M}\,[\mathbf{x}]_{\mathfrak{B}}$$

where $\mathbf{M} = \begin{bmatrix}[T(\mathbf{b_1})]_{\mathfrak{C}} & \dots & [T(\mathbf{b_n})]_{\mathfrak{C}}\end{bmatrix}$ is the matrix representation of T called the matrix for T relative to the bases $\mathfrak{B}$ and $\mathfrak{C}$.

Example 1 on page 329. Let $\mathfrak{B} = \{\mathbf{b_1}, \mathbf{b_2}\}$ and $\mathfrak{C} = \{\mathbf{c_1}, \mathbf{c_2}, \mathbf{c_3}\}$, and suppose $T(\mathbf{b_1}) = 3\mathbf{c_1} - 2\mathbf{c_2} + 5\mathbf{c_3}$ and $T(\mathbf{b_2}) = 4\mathbf{c_1} + 7\mathbf{c_2} - \mathbf{c_3}$. Then the matrix for T relative to $\mathfrak{B}$ and $\mathfrak{C}$ is $\mathbf{M} = \begin{bmatrix}[T(\mathbf{b_1})]_{\mathfrak{C}} & [T(\mathbf{b_2})]_{\mathfrak{C}}\end{bmatrix} = \begin{bmatrix} 3 & 4 \\ -2 & 7 \\ 5 & -1 \end{bmatrix}$.

Definition. If $T : V \to V$ and the basis $\mathfrak{C}$ is the same as $\mathfrak{B}$, then the matrix $\mathbf{M}$ is called the matrix for T relative to $\mathfrak{B}$ (or the $\mathfrak{B}$-matrix for T), and is denoted by $[T]_{\mathfrak{B}}$.

Example 2 on page 329. Consider differentiation in $\mathbb{P}^2$. It is defined as $T : \mathbb{P}^2 \to \mathbb{P}^2$ such that $T(a_0 + a_1 t + a_2 t^2) = a_1 + 2a_2 t$. Suppose $\mathfrak{B} = \{1, t\, t^2\}$. Since $T(1) = 0$, $T(t) = 1$, and $T(t^2) = 2t$, the $\mathfrak{B}$-matrix for T is $[T]_{\mathfrak{B}} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{bmatrix}$.

For a general $\mathbf{p}(t) = a_0 + a_1 t + a_2 t^2$, the coordinate vector is $[\mathbf{p}]_{\mathfrak{B}} = \begin{bmatrix} a_0 \\ a_1 \\ a_2 \end{bmatrix}$, and the image $[T(\mathbf{p})]_{\mathfrak{B}} = [a_1 + 2a_2 t]_{\mathfrak{B}} = \begin{bmatrix} a_1 \\ 2a_2 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ a_2 \end{bmatrix} = [T]_{\mathfrak{B}}[\mathbf{p}]_{\mathfrak{B}}$.

Theorem 8 on page 331. Suppose $\mathbf{A} = \mathbf{PDP}^{-1}$ where D is a diagonal $n \times n$ matrix. If $\mathfrak{B}$ is the basis for $\mathbb{R}^n$ formed from the columns of $\mathbf{P}$, then $\mathbf{D}$ is the $\mathfrak{B}$- matrix for the transformation $T(\mathbf{x}) = \mathbf{Ax}$.

Example 3 on page 331. Let $\mathbf{A} = \begin{bmatrix} 7 & 2 \\ -4 & 1 \end{bmatrix}$. Then $\mathbf{P} = \begin{bmatrix} 1 & 1 \\ -1 & -2 \end{bmatrix}$ and $\mathbf{D} = \begin{bmatrix} 5 & 0 \\ 0 & 3 \end{bmatrix}$. $\mathbf{D}$ is the $\mathfrak{B}$-matrix for T when $\mathfrak{B} = \left\{ \begin{bmatrix} 1 \\ -1 \end{bmatrix}, \begin{bmatrix} 1 \\ -2 \end{bmatrix} \right\}$.

6.1. Inner Product, Length, and Orthogonality.

Definition. If $\mathbf{u} = \begin{bmatrix} u_1 \\ \dots \\ u_n \end{bmatrix}$ and $\mathbf{v} = \begin{bmatrix} v_1 \\ \dots \\ v_n \end{bmatrix}$ are two vectors in $\mathbb{R}^n$, then the inner product or dot product is the number

$$\mathbf{u} \cdot \mathbf{v} = \mathbf{u}^T \mathbf{v} = \begin{bmatrix} u_1 & \dots & u_n \end{bmatrix} \begin{bmatrix} v_1 \\ \dots \\ v_n \end{bmatrix} = u_1 v_1 + \cdots + u_n v_n.$$

Example. Compute the dot product of $\mathbf{u} = \begin{bmatrix} 1 \\ -3 \\ 2 \end{bmatrix}$ and $\mathbf{v} = \begin{bmatrix} 5 \\ 0 \\ -2 \end{bmatrix}$.

Theorem 1 on page 376. The dot product has the following properties:
(a) $\mathbf{u} \cdot \mathbf{v} = \mathbf{v} \cdot \mathbf{u}$ (commutative law)
(b) $(\mathbf{u} + \mathbf{v}) \cdot \mathbf{w} = \mathbf{u} \cdot \mathbf{w} + \mathbf{v} \cdot \mathbf{w}$ (distributive law)
(c) $(c\mathbf{u}) \cdot \mathbf{v} = c(\mathbf{u} \cdot \mathbf{v}) = \mathbf{u} \cdot (c\mathbf{v})$ (associative law)
(d) $\mathbf{u} \cdot \mathbf{u} \geq 0, \forall \mathbf{u}$, and $\mathbf{u} \cdot \mathbf{u} = 0$ iff $\mathbf{u} = \mathbf{0}$

Definition. The length (or norm) of a vector $\mathbf{v}$ is $\|\mathbf{v}\| = \sqrt{\mathbf{v} \cdot \mathbf{v}} = \sqrt{v_1^2 + \cdots + v_n^2}$.
Example.

Definition. The distance between $\mathbf{u}$ and $\mathbf{v}$ is $\text{dist}(\mathbf{u}, \mathbf{v}) = \|\mathbf{u} - \mathbf{v}\|$.
Example. Picture.

Definition. Vectors $\mathbf{u}$ and $\mathbf{v}$ are orthogonal iff $\mathbf{u} \cdot \mathbf{v} = 0$.

Theorem 2 on page 380 (The Pythagorean Theorem). Vectors $\mathbf{u}$ and $\mathbf{v}$ are orthogonal iff $\|\mathbf{u} + \mathbf{v}\|^2 = \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2$.
PROOF:

Definition. Let $\mathbf{W}$ be a vector space. The orthogonal complement of $\mathbf{W}$, $\mathbf{W}^{\perp}$, is the set of all vectors that are orthogonal to every vector in $\mathbf{W}$.

Example. Let $\mathbf{W} = \mathbb{R}^2$ Then $\mathbf{W}^{\perp}$ is the line through the origin.

6.2. Orthogonal Sets.

Definition. A set of vectors $\{\mathbf{u_1}, \ldots, \mathbf{u_p}\}$ is an orthogonal set if $\mathbf{u_i} \cdot \mathbf{u_j} = 0$ for any $i \neq j$.

Example 1 on page 384. The set $\left\{ \begin{bmatrix} 3 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} -1 \\ 2 \\ 1 \end{bmatrix}, \begin{bmatrix} -1/2 \\ -2 \\ 7/2 \end{bmatrix} \right\}$ is an orthogonal set.

Theorem 4 on page 384. If $S = \{\mathbf{u_1}, \ldots, \mathbf{u_p}\}$ is an orthogonal set, then S is linearly independent and therefore is a basis for the subspace spanned by S.
PROOF: The vector equation $\mathbf{0} = c_1\mathbf{u_1} + \cdots + c_p\mathbf{u_p}$ has only the trivial solution. Indeed, $0 = \mathbf{0} \cdot \mathbf{u_1} = c_1\mathbf{u_1} \cdot \mathbf{u_1}$ hence $c_1 = 0$, etc.

Theorem 5 on page 385. Let $\{\mathbf{u_1}, \ldots, \mathbf{u_p}\}$ be an orthogonal basis for a vector space W. Then for any $\mathbf{y} \in W$,

$$\mathbf{y} = \sum_{i=1}^{p} \frac{\mathbf{y} \cdot \mathbf{u_i}}{\mathbf{u_i} \cdot \mathbf{u_i}} \mathbf{u_i}.$$

PROOF: $\mathbf{y} \cdot \mathbf{u_1} = c_1\mathbf{u_1} \cdot \mathbf{u_1}$, etc.

Example 2 on pages 385 – 386. Let $\mathbf{y} = \begin{bmatrix} 6 \\ 1 \\ -8 \end{bmatrix}$. Write $\mathbf{y}$ as a linear combination of the vectors in S in Example 1. Answer: $\mathbf{y} = \mathbf{u_1} - 2\mathbf{u_2} - 2\mathbf{u_3}$.

Definition. Let $L = Span\{\mathbf{u}\}$ for some vector $\mathbf{u} \in \mathbb{R}^n$. Take any $\mathbf{y} \notin L$. The orthogonal projection of y onto L is

$$\hat{\mathbf{y}} = \text{proj}_L\, \mathbf{y} = \frac{\mathbf{y} \cdot \mathbf{u}}{\mathbf{u} \cdot \mathbf{u}} \mathbf{u}.$$

Example. $\mathbf{u} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$.

Definition. A set of vectors is orthonormal if it is an orthogonal set of unit vectors.
Example.

Theorem 6 on page 390. Matrix $\mathbf{U}$, called the unitarian matrix, has orthonormal columns iff $\mathbf{U}^T\mathbf{U} = \mathbf{I}$.

Theorem 7 on page 390. Unitarian matrices have the property that$(\mathbf{Ux}) \cdot (\mathbf{Uy}) = \mathbf{x} \cdot \mathbf{y}$. In particular, $\|\mathbf{Ux}\| = \|\mathbf{x}\|$.

6.3. Orthogonal Projections.

Theorem 8 on page 395 (The Orthogonal Decomposition Theorem). Let W be a subspace of $\mathbb{R}^n$. Then any $\mathbf{y} \in \mathbb{R}^n$ can be written uniquely as

$$\mathbf{y} = \hat{\mathbf{y}} + \mathbf{z}$$

where $\hat{\mathbf{y}} \in W$ and $\mathbf{z} \in W^{\perp}$.
In fact, if $\{\mathbf{u_1}, \ldots, \mathbf{u_p}\}$ is an orthogonal basis for W, then

$$\hat{\mathbf{y}} = \sum_{i=1}^{p} \frac{\mathbf{y} \cdot \mathbf{u_i}}{\mathbf{u_i} \cdot \mathbf{u_i}} \mathbf{u_i}$$

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

Picture.

Example 2 on pages 396 – 397. Let $\mathbf{u_1} = \begin{bmatrix} 2 \\ 5 \\ -1 \end{bmatrix}, \mathbf{u_2} = \begin{bmatrix} -2 \\ 1 \\ 1 \end{bmatrix}$, and $\mathbf{y} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$. Then the decomposition of $\mathbf{y}$ is

$$\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} = \begin{bmatrix} -2/5 \\ 2 \\ 1/5 \end{bmatrix} + \begin{bmatrix} 7/5 \\ 0 \\ 14/5 \end{bmatrix}.$$

Theorem 9 on page 398. An orthogonal projection $\hat{\mathbf{y}}$ of a vector $\mathbf{y}$ onto W is the closest vector in W to $\mathbf{y}$ in the sense that $\text{dist}(\mathbf{y}, \hat{\mathbf{y}}) = \|\mathbf{y} - \hat{\mathbf{y}}\| = \min_{\mathbf{v} \in W} \|\mathbf{y} - \mathbf{v}\|$.

Picture.

Definition. The distance between $\mathbf{y}$ and W is the distance between $\mathbf{y}$ and the closest vector in W, that is, $\text{dist}(\mathbf{y}, W) = \|\mathbf{y} - \text{proj}_W \mathbf{y}\|$.

Example 4 on page 399. Find the distance between $\mathbf{y} = \begin{bmatrix} = -1 \\ -5 \\ 10 \end{bmatrix}$ and $W = Span\left\{\mathbf{u_1} = \begin{bmatrix} 5 \\ -2 \\ 1 \end{bmatrix}, \mathbf{u_2} = \begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix}\right\}$.

SOLUTION:

$$\hat{\mathbf{y}} = (1/2)\mathbf{u_1} - (7/2)\mathbf{u_2} = \begin{bmatrix} -1 \\ -8 \\ 4 \end{bmatrix}, \ \mathbf{y} - \hat{\mathbf{y}} = \begin{bmatrix} 0 \\ 3 \\ 6 \end{bmatrix}, \ \|\mathbf{y} - \hat{\mathbf{y}}\| = 3\sqrt{5}.$$

Theorem 10 on page 399. If $\{\mathbf{u_1}, \ldots, \mathbf{u_p}\}$ is an orthonormal basis in W, then $\text{proj}_W \mathbf{y} = \sum_{i=1}^{p} (\mathbf{y} \cdot \mathbf{u_i})\mathbf{u_i}$. If $\mathbf{U} = \begin{bmatrix} \mathbf{u_1} & \cdots & \mathbf{u_p} \end{bmatrix}$, then $\text{proj}_W \mathbf{y} = \mathbf{U}\mathbf{U}^T\mathbf{y}$.

6.4. The Gram – Schmidt Orthogonalization Process.

Theorem 11 on page 404 (The Gram – Schmidt Process). Suppose $\{\mathbf{x_1}, \ldots, \mathbf{x_p}\}$ is a basis for W. The following algorithm produces an orthogonal basis for W. Take

$$\mathbf{v_1} = \mathbf{x_1}, \quad \mathbf{v_2} = \mathbf{x_2} - \frac{\mathbf{x_2} \cdot \mathbf{v_1}}{\mathbf{v_1} \cdot \mathbf{v_1}}\mathbf{v_1}, \quad \ldots,$$

$$\mathbf{v_p} = \mathbf{x_p} - \frac{\mathbf{x_p} \cdot \mathbf{v_1}}{\mathbf{v_1} \cdot \mathbf{v_1}}\mathbf{v_1} - \frac{\mathbf{x_p} \cdot \mathbf{v_2}}{\mathbf{v_2} \cdot \mathbf{v_2}}\mathbf{v_2} - \cdots - \frac{\mathbf{x_p} \cdot \mathbf{v_{p-1}}}{\mathbf{v_{p-1}} \cdot \mathbf{v_{p-1}}}\mathbf{v_{p-1}}.$$

The set $\{\mathbf{v_1}, \ldots, \mathbf{v_p}\}$ is an orthogonal basis for W.

Example 2 on pages 402 – 403. Suppose

$$\mathbf{x_1} = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}, \mathbf{x_2} = \begin{bmatrix} 0 \\ 1 \\ 1 \\ 1 \end{bmatrix}, \mathbf{x_3} = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 1 \end{bmatrix}.$$

Then

$$\mathbf{v_1} = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}, \mathbf{v_2} = \begin{bmatrix} -3/4 \\ 1/4 \\ 1/4 \\ 1/4 \end{bmatrix}, \mathbf{v_3} = \begin{bmatrix} 0 \\ -2/3 \\ 1/3 \\ 1/3 \end{bmatrix}.$$

6.7. Inner Product Spaces.

Definition. An <u>inner product</u> on a vector space V is a function that for any $\mathbf{u}$ and $\mathbf{v} \in V$, assigns a real number $\langle \mathbf{u}, \mathbf{v} \rangle$ satisfying the following properties:
1. $\langle \mathbf{u}, \mathbf{v} \rangle = \langle \mathbf{v}, \mathbf{u} \rangle$ (commutative law)
2. $\langle \mathbf{u} + \mathbf{v}, \mathbf{w} \rangle = \langle \mathbf{u}, \mathbf{w} \rangle + \langle \mathbf{v}, \mathbf{w} \rangle$ (distributive law)
3. $\langle c\mathbf{u}, \mathbf{v} \rangle = c\langle \mathbf{u}, \mathbf{v} \rangle$
4. $\langle \mathbf{u}, \mathbf{u} \rangle \geq 0$ and $\langle \mathbf{u}, \mathbf{u} \rangle = 0$ iff $\mathbf{u} = \mathbf{0}$.

A vector space with an inner product is called an <u>inner product space</u>.

Example 1 on page 428. Let $\mathbf{u} = \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}$ and $\mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}$. Show that $\langle \mathbf{u}, \mathbf{v} \rangle = 4u_1v_1 + 5u_2v_2$ is an inner product.

Generally, $\langle \mathbf{u}, \mathbf{v} \rangle = \sum_{i=1}^{n} a_i u_i v_i$ defines an inner product in $\mathbb{R}^n$.

Example 2 on page 429. Fix $t_0, \ldots, t_n \in \mathbb{R}$. For any $\mathbf{p}$ and $\mathbf{q} \in \mathbb{P}_n$ define $\langle \mathbf{p}, \mathbf{q} \rangle = \sum_{i=0}^{n} \mathbf{p}(t_i)\mathbf{q}(t_i)$. Show that it is an inner product.
SOLUTION: $\langle \mathbf{p}, \mathbf{p} \rangle = 0$ iff the polynomial vanishes at $n+1$ points iff it is a zero polynomial since its degree is at most n.

Example 3 on page 429. Let $t_0 = 0, t_1 = 1/2, t_2 = 1$ and let $\mathbf{p}(t) = 12t^2$ and $\mathbf{q} = 2t - 1$. Then $\langle \mathbf{p}, \mathbf{q} \rangle = \mathbf{p}(0)\mathbf{q}(0) + \mathbf{p}(1/2)\mathbf{q}(1/2) + \mathbf{p}(1)\mathbf{q}(1) = 12$.

Definition. The length (or norm) of a vector $\mathbf{v}$ in an inner product space V is $\|\mathbf{v}\| = \sqrt{\langle \mathbf{v}, \mathbf{v} \rangle}$.

Example 4 on page 429. In Example 3, $\|\mathbf{p}\| = \sqrt{153}$ and $\|\mathbf{q}\| = \sqrt{2}$.

The Gram – Schmidt Orthogonalization process is applicable to an inner product space.

Example 5 on pages 430 – 431. Consider $\mathbb{P}_2$ with the inner product $\langle \mathbf{p}, \mathbf{q} \rangle = \mathbf{p}(-1)\mathbf{q}(-1) + \mathbf{p}(0)\mathbf{q}(0) + \mathbf{p}(1)\mathbf{q}(1)$. Take $\{1, t, t^2\}$, the standard basis in $\mathbb{P}^2$. Orthogonalize this basis.

SOLUTION: For any polynomial $\mathbf{p} \in \mathbb{P}_2$, consider $\begin{bmatrix} \mathbf{p}(-1) \\ \mathbf{p}(0) \\ \mathbf{p}(1) \end{bmatrix} \in \mathbb{R}^3$. The standard basis corresponds to the set $\left\{ \mathbf{v_1} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \mathbf{v_2} = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}, \mathbf{v_3} = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} \right\}$. Notice that $\mathbf{x_1} = \mathbf{v_1}$ and $\mathbf{x_2} = \mathbf{v_2}$ are orthogonal. Apply the Gram – Schmidt process to find $\mathbf{x_3} = \mathbf{v_3} - \frac{\langle \mathbf{x_3}, \mathbf{v_1} \rangle}{\langle \mathbf{v_1}, \mathbf{v_1} \rangle} \mathbf{v_1} - \frac{\langle \mathbf{x_3}, \mathbf{v_2} \rangle}{\langle \mathbf{v_2}, \mathbf{v_2} \rangle} \mathbf{v_2} = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} - (2/3) \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 1/3 \\ -2/3 \\ 1/3 \end{bmatrix}$.
This vector corresponds to a polynomial $a + bt + ct^2 \in \mathbb{P}_2$ where the coefficients satisfy

$$\begin{cases} a - b + c = 1/3 \\ a = -2/3 \\ a + b + c = 1/3. \end{cases}$$

Thus, $\{1, t, t^2 - 2/3\}$ is an orthogonal basis for $\mathbb{P}_2$. An orthonormal basis is $\{1/3, t/2, (3/2)t^2 - 1\}$.

Theorem 16 on page 432 (The Cauchy – Schwarz Inequality). For any $\mathbf{u}$ and $\mathbf{v}$,

$$|\langle \mathbf{u}, \mathbf{v} \rangle| \leq \|\mathbf{u}\| \|\mathbf{v}\|.$$

Three applications of the Cauchy – Schwarz Inequality.

1. **Theorem 17 on page 433 (The Triangle Inequality).** For any $\mathbf{u}$ and $\mathbf{v}$,

$$\|\mathbf{u}+\mathbf{v}\| \leq \|\mathbf{u}\| + \|\mathbf{v}\|.$$

Picture.

PROOF: $\|\mathbf{u}+\mathbf{v}\|^2 = \langle \mathbf{u}+\mathbf{v}, \mathbf{u}+\mathbf{v}\rangle = \langle \mathbf{u},\mathbf{u}\rangle + 2\langle \mathbf{u},\mathbf{v}\rangle + \langle \mathbf{v},\mathbf{v}\rangle \leq \|\mathbf{u}\|^2 + 2|\langle \mathbf{u},\mathbf{v}\rangle| + \|\mathbf{v}\|^2 \overset{C-S.\neq}{\leq} \|\mathbf{u}\|^2 + 2\|\mathbf{u}\|\|\mathbf{v}\| + \|\mathbf{v}\|^2 = (\mathbf{u}+\mathbf{v})^2.$ □

2. **Exercise 19 on page 436.** Show that $\sqrt{ab} \leq (a+b)/2$, that is, the geometric mean is less than or equal to the arithmetic mean.

SOLUTION: Consider $\mathbf{u} = \begin{bmatrix} \sqrt{a} \\ \sqrt{b} \end{bmatrix}$ and $\mathbf{v} = \begin{bmatrix} \sqrt{b} \\ \sqrt{a} \end{bmatrix}$. The inner product is $\langle \mathbf{u},\mathbf{v}\rangle = 2\sqrt{ab}$. The norms are $\|\mathbf{u}\| = \|\mathbf{v}\| = \sqrt{a+b}$. By the Cauchy – Schwarz inequality, $2\sqrt{ab} \leq a+b$.

3. **Exercise 20 on page 436.** Show that $(a+b)^2 \leq 2a^2 + 2b^2$.

SOLUTION: Consider $\mathbf{u} = \begin{bmatrix} a \\ b \end{bmatrix}$ and $\mathbf{v} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$. The inner product is $\langle \mathbf{u},\mathbf{v}\rangle = a+b$. The norms are $\|\mathbf{u}\| = \sqrt{a^2+b^2}$ and $\|\mathbf{v}\| = \sqrt{2}$. By the Cauchy – Schwarz inequality, $a+b \leq \sqrt{2(a^2+b^2)}$.

7.1. Diagonalization of Symmetric Matrices.

Definition. A square matrix $\mathbf{A}$ is symmetric iff $\mathbf{A} = \mathbf{A}^T$.

Example.

Theorem 1 on page 450. If $\mathbf{A}$ is symmetric, then any two eigenvectors from different eigenspaces are orthogonal.

PROOF: Let $\mathbf{v_1}$ and $\mathbf{v_2}$ be two eigenvectors corresponding to two distinct eigenvalues λ_1 and λ_2, respectively. We want to show that $\mathbf{v_1} \cdot \mathbf{v_2} = 0$. We have $\lambda_1 \mathbf{v_1} \cdot \mathbf{v_2} = (\lambda_1 \mathbf{v_1}) \cdot \mathbf{v_2} = (\lambda_1 \mathbf{v_1})^T \mathbf{v_2} = (\mathbf{A}\mathbf{v_1})^T \mathbf{v_2} = \mathbf{v_1}^T \mathbf{A}^T \mathbf{v_2} = \mathbf{v_1}^T(\mathbf{A}\mathbf{v_2}) = \mathbf{v_1}^T(\lambda_2 \mathbf{v_2}) = \lambda_2 \mathbf{v_1} \cdot \mathbf{v_2}$. Since $\lambda_1 \neq \lambda_2$, $\mathbf{v_1} \cdot \mathbf{v_2} = 0$. □

Example 2 on pages 449 – 450. Consider $\mathbf{A} = \begin{bmatrix} 6 & -2 & -1 \\ -2 & 6 & -1 \\ -1 & -1 & 5 \end{bmatrix}$. The characteristic equation of $\mathbf{A}$ is $-(\lambda-3)(\lambda-6)(\lambda-8) = 0$. The eigenvectors are

$$\lambda_1 = 3 : \mathbf{v_1} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \quad \lambda_2 = 6 : \mathbf{v_2} = \begin{bmatrix} -1 \\ -1 \\ 2 \end{bmatrix}, \quad \lambda_3 = 8 : \mathbf{v_3} = \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}.$$

The are orthogonal.

Since a nonzero multiple of an eigenvector is still an eigenvector, we can normalize the orthogonal eigenvectors $\{\mathbf{v_1}, \mathbf{v_2}, \mathbf{v_3}\}$ to produce the unit eigenvectors (orthonormal eigenvectors) $\{\mathbf{u_1}, \mathbf{u_2}, \mathbf{u_3}\}$. In our example,

$$\mathbf{u_1} = \begin{bmatrix} 1/\sqrt{3} \\ 1/\sqrt{3} \\ 1/\sqrt{3} \end{bmatrix}, \quad \mathbf{u_2} = \begin{bmatrix} -1/\sqrt{6} \\ -1/\sqrt{6} \\ 2/\sqrt{6} \end{bmatrix}, \quad \mathbf{u_3} = \begin{bmatrix} -1/\sqrt{2} \\ 1/\sqrt{2} \\ 0 \end{bmatrix}.$$

Definition. Let $\mathbf{P} = \begin{bmatrix} \mathbf{u_1} & \dots & \mathbf{u_n} \end{bmatrix}$ where $\{\mathbf{u_1}, \dots, \mathbf{u_n}\}$ are the orthonormal eigenvectors of an $n \times n$ matrix $\mathbf{A}$ that correspond to not necessarily distinct eigenvalues $\lambda_1, \dots, \lambda_n$. The matrix $\mathbf{P}$ is called an <u>orthonormal</u> matrix. It has the property that $\mathbf{P}^{-1} = \mathbf{P}^T$. Let $\mathbf{D} = \begin{bmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ & & \dots & \\ 0 & 0 & \dots & \lambda_n \end{bmatrix}$ be the diagonal matrix with the eigenvalues on the main diagonal. Then $\mathbf{A} = \mathbf{PDP^{-1}} = \mathbf{PDP}^T$. The matrix $\mathbf{A}$ is said to be <u>orthogonally diagonalizable</u>.

In our example, $\mathbf{P} = \dots$, $\mathbf{D} = \dots$.

Theorem 2 on page 451. An $n \times n$ matrix $\mathbf{A}$ is orthogonally diagonalizable iff $\mathbf{A}$ is symmetric.

Example 3 on pages 451 – 452. Orthogonally diagonalize the matrix $\mathbf{A} = \begin{bmatrix} 3 & -2 & 4 \\ -2 & 6 & 2 \\ 4 & 2 & 3 \end{bmatrix}$.

SOLUTION: The characteristic polynomial is $-(\lambda + 2)(\lambda - 7)^2$. The eigenvectors are

$$\lambda_1 = -2: \ \mathbf{v_1} = \begin{bmatrix} -1 \\ -1/2 \\ 1 \end{bmatrix}, \quad \lambda_2 = 7: \ \mathbf{v_2} = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \ \mathbf{v_3} = \begin{bmatrix} -1/2 \\ 1 \\ 0 \end{bmatrix}.$$

The basis $\{\mathbf{v_2}, \mathbf{v_3}\}$ can be orthogonalized by using the Gram – Schmidt process. It produces $\mathbf{u_1} = \begin{bmatrix} -2/3 \\ -1/3 \\ 2/3 \end{bmatrix}$, $\mathbf{u_2} = \begin{bmatrix} 1/\sqrt{2} \\ 0 \\ 1/\sqrt{2} \end{bmatrix}$, $\mathbf{u_3} = \begin{bmatrix} -1/\sqrt{18} \\ 4/\sqrt{18} \\ 1/\sqrt{18} \end{bmatrix}$. Then $\mathbf{A} = \mathbf{PDP}^T$.

Definition. Write $\mathbf{A} = \mathbf{PDP}^T = \begin{bmatrix} \mathbf{u_1} & \dots & \mathbf{u_n} \end{bmatrix} \begin{bmatrix} \lambda_1 & 0 & \dots & 0 \\ & & \dots & \\ 0 & 0 & \dots & \lambda_n \end{bmatrix} \begin{bmatrix} \mathbf{u_1}^T \\ \dots \\ \mathbf{u_n}^T \end{bmatrix} =$ $\begin{bmatrix} \lambda_1 \mathbf{u_1} & \dots & \lambda_n \mathbf{u_n} \end{bmatrix} \begin{bmatrix} \mathbf{u_1}^T \\ \dots \\ \mathbf{u_n}^T \end{bmatrix}$. Thus, we can write $\mathbf{A} = \lambda_1 \mathbf{u_1} \mathbf{u_1}^T + \dots + \lambda_n \mathbf{u_n} \mathbf{u_n}^T$. This representation of $\mathbf{A}$ is called a <u>spectral decomposition</u> of $\mathbf{A}$. The set of eigenvalues of $\mathbf{A}$ is called the spectrum.

Example 4 on page 453. Construct a spectral decomposition of $\mathbf{A} = \begin{bmatrix} 7 & 2 \\ 2 & 4 \end{bmatrix} = \begin{bmatrix} 2/\sqrt{5} & -1/\sqrt{5} \\ 1/\sqrt{5} & 2/\sqrt{5} \end{bmatrix} \begin{bmatrix} 8 & 0 \\ 0 & 3 \end{bmatrix} \begin{bmatrix} 2/\sqrt{5} & 1/\sqrt{5} \\ -1/\sqrt{5} & 2/\sqrt{5} \end{bmatrix}$.

SOLUTION: $\mathbf{A} = 8\mathbf{u_1}\mathbf{u_1}^T + 3\mathbf{u_2}\mathbf{u_2}^T = \begin{bmatrix} 32/5 & 16/5 \\ 16/5 & 8/5 \end{bmatrix} + \begin{bmatrix} 3/5 & -6/5 \\ -6/5 & 12/5 \end{bmatrix}$.

7.2. Quadratic Forms.

Definition. A quadratic form on $\mathbb{R}^n$ is $Q(\mathbf{x}) = \mathbf{x}^T\mathbf{A}\mathbf{x}$ where $\mathbf{A}$ is an $n \times n$ symmetric matrix called the matrix of quadratic form.

Example 1 on page 456. Let $\mathbf{A} = \begin{bmatrix} 3 & -2 \\ -2 & 7 \end{bmatrix}$. The quadratic form $Q(\mathbf{x}) = \mathbf{x}^T\mathbf{A}\mathbf{x} = 3x_1^2 - 4x_1x_2 + 7x_2^2$.

Example 2 on page 456. Let $Q(\mathbf{x}) = 5x_1^2 + 3x_2^2 + 2x_3^2 - x_1x_2 + 8x_2x_3$. The matrix of the quadratic form is $\mathbf{A} = \begin{bmatrix} 5 & -1/2 & 0 \\ -1/2 & 3 & 4 \\ 0 & 4 & 2 \end{bmatrix}$.

Definition. If $\mathbf{x}$ represents a variable vector in $\mathbb{R}^n$, then the change of variable is an equation of the form $\mathbf{x} = \mathbf{P}\mathbf{y}$ or $\mathbf{y} = \mathbf{P}^{-1}\mathbf{x}$.

Definition. If the change of variable is made in the quadratic form, then $\mathbf{x}^T\mathbf{A}\mathbf{x} = (\mathbf{P}\mathbf{y})^T\mathbf{A}(\mathbf{P}\mathbf{y}) = \mathbf{y}^T(\mathbf{P}^T\mathbf{A}\mathbf{P})\mathbf{y}$, and the new matrix of the quadratic form is $\mathbf{P}^T\mathbf{A}\mathbf{P}$. If $\mathbf{P}$ orthogonally diagonalizes $\mathbf{A}$, then $\mathbf{P}^T\mathbf{A}\mathbf{P} = \mathbf{D}$. Thus, the matrix of the new quadratic form is diagonal.

Example 4 on pages 457 – 458. Let $\mathbf{A} = \begin{bmatrix} 1 & -4 \\ -4 & -5 \end{bmatrix}$. Then $\mathbf{P} = \begin{bmatrix} 2/\sqrt{5} & 1/\sqrt{5} \\ -1/\sqrt{5} & 2/\sqrt{5} \end{bmatrix}$ and $\mathbf{D} = \begin{bmatrix} 3 & 0 \\ 0 & -7 \end{bmatrix}$. The orthogonal change of variable is $\mathbf{x} = \mathbf{P}\mathbf{y}$ where $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$ and $\mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$. The quadratic form is $\mathbf{x}^T\mathbf{A}\mathbf{x} = x_1^2 - 8x_1x_2 - 5x_2^2 = \mathbf{y}^T\mathbf{D}\mathbf{y} = 3y_1^2 - 7y_2^2$.

Theorem 4 on page 458. If $\mathbf{A}$ is an $n \times n$ symmetric matrix, then there exists an orthogonal change of variable $\mathbf{x} = \mathbf{P}\mathbf{y}$ that transforms the quadratic form $\mathbf{x}^T\mathbf{A}\mathbf{x}$ into a quadratic form $\mathbf{y}^T\mathbf{D}\mathbf{y}$ with no cross-product terms.

Definition. The columns of $\mathbf{P}$ are called the principal axes of the quadratic form $\mathbf{x}^T\mathbf{A}\mathbf{x}$. The vector $\mathbf{y}$ is actually vector $\mathbf{x}$ relative to the orthonormal basis of $\mathbb{R}^n$ given by the principal axes.

Geometric Interpretation of Quadratic Forms in $\mathbb{R}^2$.

Consider $Q(\mathbf{y}) = \mathbf{y}^T\mathbf{D}\mathbf{y}$, and let c be a constant. Then the equation $\mathbf{y}^T\mathbf{D}\mathbf{y} = c$ can be written in either of the following six forms:
(i) $x_1^2/a^2 + x_2^2/b^2 = 1$, $a \geq b > 0$ (an ellipse if $a > b$, a circle if $a = b$),
(ii) $x_1^2/a^2 + x_2^2/b^2 = 0$ (a single point $(0, 0)$),
(iii) $x_1^2/a^2 + x_2^2/b^2 = -1$ (empty set of points),
(iv) $x_1^2/a^2 - x_2^2/b^2 = 1, a \geq b > 0$ (a hyperbola),
(v) $x_1^2/a^2 - x_2^2/b^2 = 0$ (two intersecting lines $x_2 = \pm(b/a)x_1$),
and
(vi) $x_1^2/a^2 - x_2^2/b^2 = -1$ (a hyperbola $x_2^2/b^2 - x_1^2/a^2 = 1$).

Consider $Q(\mathbf{x}) = \mathbf{x}^T\mathbf{A}\mathbf{x}$ where $\mathbf{A}$ is a 2×2 symmetric matrix. If $\mathbf{A}$ is diagonal, the graph is in standard position. If $\mathbf{A}$ is not diagonal, the graph is rotated out of standard position. Finding the principal axes amounts to finding the new coordinate system with respect to which that graph is in standard position.

Pictures of an ellipse, hyperbola, rotation of axes, etc.

Definition. A quadratic form Q is:
a. positive definite if $Q(\mathbf{x}) > 0$ for all $\mathbf{x} \neq \mathbf{0}$,
b. negative definite if $Q(\mathbf{x}) < 0$ for all $\mathbf{x} \neq \mathbf{0}$,
and
c. indefinite if $Q(\mathbf{x})$ assumes both positive and negative values.

Example. $Q(\mathbf{x}) = x_1^2 + 2x_2^2$, $Q(\mathbf{x}) = -x_1^2 - 2x_2^2$, $Q(\mathbf{x}) = x_1^2 - 2x_2^2$.

Theorem 5 on page 461. Let $\mathbf{A}$ be an $n \times n$ symmetric matrix. Then the quadratic form $\mathbf{x}^T\mathbf{A}\mathbf{x}$ is
a. positive definite iff the eigenvalues of $\mathbf{A}$ are all positive,
b. negative definite iff the eigenvalues of $\mathbf{A}$ are all negative,
and
c. indefinite iff some eigenvalues of $\mathbf{A}$ are positive and some are negative.
PROOF: $\mathbf{x}^T\mathbf{A}\mathbf{x} = \mathbf{y}^T\mathbf{D}\mathbf{y} = \lambda_1 y_1^2 + \cdots + \lambda_n y_n^2$. $\square$

Definition. An $n \times n$ symmetric matrix $\mathbf{A}$ is called a positive definite matrix (negative definite or indefinite) if the corresponding quadratic form $\mathbf{x}^T\mathbf{A}\mathbf{x}$ is positive definite (negative definite or indefinite).

Example 5 on page 461. Is $Q(\mathbf{x}) = 3x_1^2 + 2x_2^2 + x_3^2 + 4x_1x_2 + 4x_2x_3$ positive definite?

SOLUTION: $\mathbf{A} = \begin{bmatrix} 3 & 2 & 0 \\ 2 & 2 & 2 \\ 0 & 2 & 1 \end{bmatrix}$. The eigenvalues are -1, 2, and 5. Hence, Q is indefinite.