INTRODUCTION TO LINEAR ALGEBRA

FARID ALINIAEIFARD

Contents

1. Week 1: Systems of Linear equations

Goals of the first week:

(1) What is a system of linear equations.
(2) Echolen form (or row echelon form) and reduced echelon form (or row reduced echelOn form).
(3) The number of solutions of a linear system.
(4) Solutions of linear systems.

1.1. **Week 1, Lecture 1, Aug. 28, 2017, What is a system of linear equations.** A **linear equation** in the variables $x_1, x_2, \ldots, x_n$ is an equation that can be written in the form

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

where b and the **coefficients** $a_1, \ldots, a_n$ are real (complex) numbers.

Example 1.1. $4x_1 - 5x_2 + 2 = x_1$ *is a linear equation because it can be rearranged as* $4x_1 - x_1 - 5x_2 = -2$ *and it is the same as* $3x_1 - 5x_2 = -2$.

A **system of linear equations** (or a **linear system**) is a collection of one or more linear equations.

Example 1.2.

$$\begin{aligned} 2x_1 - x_2 + 1.5x_3 &= 8 \\ x_1 \qquad\quad - 4x_3 &= -7 \end{aligned}$$

A **solution** of a system of linear equations is a list $\begin{bmatrix} s_1 \\ s_2 \\ \vdots \\ s_n \end{bmatrix}$ of numbers that makes each equation a true statement when the values $s_1, s_2, \ldots, s_n$ are substituted for $x_1, x_2, \ldots, x_n$, respectively. Also, the set of all possible solutions is called **solution set** of the linear system.

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

Example 1.3. *The system*

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

has $\{\begin{bmatrix} 1 \\ 1 \end{bmatrix}\}$ *as the solution set.*

1.1.1. *Coefficient Matrix and Augmented Matrix of a Linear System.* Given the system

$$\begin{array}{cccc} x_1 & -2x_2 & +x_3 & = 0 \\ & 2x_2 & -8x_3 & = 8 \\ 5x_1 & & -5x_3 & = 10 \end{array}$$

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

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

is called the coefficient matrix of the system above, and

$$\begin{bmatrix} 1 & -2 & 1 & 0 \\ 0 & 2 & -8 & 8 \\ 5 & 0 & -5 & 10 \end{bmatrix}$$

is called the augmented matrix of the system above.

1.2. **Week 1, Lecture 2, Aug. 30, 2017, Echolen form (or row echelon form) and reduced echelon form (or row reduced echelon form).**

1.2.1. *Elementary Row Operations.* The following three are called elementary row operations:

- (Replacement) replace one row by the sum of itself and a multiple of another row.
- (Interchanging) interchanging two rows.
- (Scaling) multiply all entries in a row by a nonzero constant.

Example 1.4.

$$\begin{bmatrix} 1 & -2 & 1 & 0 \\ 0 & 2 & -8 & 8 \\ 5 & 0 & -5 & 10 \end{bmatrix} \overset{\text{Intechanging}}{R1 \longleftrightarrow R2} \begin{bmatrix} 0 & 2 & -8 & 8 \\ 1 & -2 & 1 & 0 \\ 5 & 0 & -5 & 10 \end{bmatrix} \overset{\text{ScalingR3 by } 1/5}{\longleftrightarrow} \begin{bmatrix} 0 & 2 & -8 & 8 \\ 1 & -2 & 1 & 0 \\ 1 & 0 & -1 & 2 \end{bmatrix}$$

$$\overset{\text{Replacing R2 by R2+}(-2)\text{R1}}{\longleftrightarrow} \begin{bmatrix} 0 & 2 & -8 & 8 \\ 1 & -6 & 17 & -16 \\ 1 & 0 & -1 & 2 \end{bmatrix}$$

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

1.2.2. *Echolen form (or row echelon form) and reduced echelon form (or row reduced echelon form).* A matrix is in echelon form (or row echelon form) if it has the following three properties:

(1) All nonzero rows are above any rows of all zeros.
(2) Each leading entry (the leftmost nonzero entry in a nonzero row) is in a column to the right of the leading entry of the row above it.
(3) All entries in a column below a leading entry are zeros.

Example 1.5.

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

is in echelon form.

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

(4) The leading entry in each nonzero row is 1.
(5) Each leading 1 is the only nonzero entry in its column.

Example 1.6.

$$\begin{bmatrix} 1 & -3 & 0 & 0 & 21 \\ 0 & 0 & 1 & 0 & 8 \\ 0 & 0 & 0 & 1 & 3 \end{bmatrix}$$

is in reduced row echelon form.

Theorem 1.7. *Each matrix is row equivalent to one and only one reduced echelon matrix.*

Note that the theorem above is not true for echelon form.

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

The row reduction algorithm: The algorithm that follows consists of four steps, and by using elementary row operations,it produces a matrix in echelon form. The steps 5-7 produces a matrix in reduced echelon form.

Example 1.8. *Transfer the following matrix*

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

first into echelon form and then into reduced echelon form.

Solution:

STEP 1: Begin with the leftmost nonzero column.

STEP 2: Select a nonzero entry in the column in Step 1, call this entry c. By interchanging the rows, move the nonzero entry to the first row.

$$\begin{bmatrix} 0 & 3 & -6 & 6 & 4 & -5 \\ 3 & -7 & 8 & -5 & 8 & 9 \\ 3 & -9 & 12 & -9 & 6 & 15 \end{bmatrix} \overset{\text{Intechanging}}{R1 \longleftrightarrow R3} \begin{bmatrix} 3 & -9 & 12 & -9 & 6 & 15 \\ 3 & -7 & 8 & -5 & 8 & 9 \\ 0 & 3 & -6 & 6 & 4 & -5 \end{bmatrix}$$

STEP 3: Use row replacement operation to create zeros in all positions below the entry c.

$$\begin{bmatrix} 3 & -9 & 12 & -9 & 6 & 15 \\ 3 & -7 & 8 & -5 & 8 & 9 \\ 0 & 3 & -6 & 6 & 4 & -5 \end{bmatrix} \overset{\text{Replacing R2 by R2+(}-1\text{)R1}}{\longleftrightarrow} \begin{bmatrix} 3 & -9 & 12 & -9 & 6 & 15 \\ 0 & 2 & -4 & 4 & 2 & -6 \\ 0 & 3 & -6 & 6 & 4 & -5 \end{bmatrix}$$

STEP 4: Cover (or ignore) the row containing c. Apply steps 1-3 to the submatrix that remains. Repeat the process until there are no more nonzero rows to modify.

$$\begin{bmatrix} 3 & -9 & 12 & -9 & 6 & 15 \\ 0 & 2 & -4 & 4 & 2 & -6 \\ 0 & 3 & -6 & 6 & 4 & -5 \end{bmatrix} \overset{\text{Replacing R3 by R3+}(-3/2)\text{R2}}{\longleftrightarrow} \begin{bmatrix} 3 & -9 & 12 & -9 & 6 & 15 \\ 0 & 2 & -4 & 4 & 2 & -6 \\ 0 & 0 & 0 & 0 & 1 & 4 \end{bmatrix}$$

Until now, we have the echelon form of the matrix. By the following steps 5-7, we will produce the reduced echelon form.

STEP 5: Begin with last nonzero row. By a scaling operation, make the leading entry 1.

STEP 6: Use row replacement operation to create zeros in all positions above the entry 1 in Step 5.

$$\begin{bmatrix} 3 & -9 & 12 & -9 & 6 & 15 \\ 0 & 2 & -4 & 4 & 2 & -6 \\ 0 & 0 & 0 & 0 & 1 & 4 \end{bmatrix} \begin{matrix} \overset{\text{Replacing R1 by R1+}(-6)\text{R3}}{\longleftrightarrow} \\ \overset{\text{Replacing R2 by R2+}(-2)\text{R3}}{\longleftrightarrow} \end{matrix} \begin{bmatrix} 3 & -9 & 12 & -9 & 0 & -9 \\ 0 & 2 & -4 & 4 & 0 & -14 \\ 0 & 0 & 0 & 0 & 1 & 4 \end{bmatrix}$$

STEP 7: Cover (or ignore) the row containing the entry 1 and, if any, the rows below it. Apply Steps 5 and 6 to the submatrix that remains. Repeat the process until there is no nonzero row to modify.

$$\begin{bmatrix} 3 & -9 & 12 & -9 & 0 & -9 \\ 0 & 2 & -4 & 4 & 0 & -14 \\ 0 & 0 & 0 & 0 & 1 & 4 \end{bmatrix} \overset{\text{Scaling R2 by }1/2}{\longleftrightarrow} \begin{bmatrix} 3 & -9 & 12 & -9 & 0 & -9 \\ 0 & 1 & -2 & 2 & 0 & -7 \\ 0 & 0 & 0 & 0 & 1 & 4 \end{bmatrix}$$

$$\overset{\text{Replacing R1 by R1+9R2}}{\longleftrightarrow} \begin{bmatrix} 3 & 0 & -6 & 9 & 0 & -72 \\ 0 & 1 & -2 & 2 & 0 & -7 \\ 0 & 0 & 0 & 0 & 1 & 4 \end{bmatrix} \overset{\text{Scaling R1 by }1/3}{\longleftrightarrow} \begin{bmatrix} 1 & 0 & -2 & 3 & 0 & -24 \\ 0 & 1 & -2 & 2 & 0 & -7 \\ 0 & 0 & 0 & 0 & 1 & 4 \end{bmatrix}$$

1.3. **Week 1, Lecture 3, Sept. 1, 2017 The number of solutions of a linear system.** A system of linear equations has

(1) no solution, or
(2) exactly one solution, or
(3) infinitely many solutions.

A system of linear equations is said to be **consistent** if it has wither one solution or infinitely many solutions, a system is **inconsistent** if it has no solution.

Theorem 1.9. (1) *A linear system is inconsistent if and only if an echelon form of the augmented matrix has a row of the form*

$$[0 \;\; 0 \;\; \cdots \;\; 0 \;\; b].$$

(2) *If a echelon form of the augmented matrix does not have a row of the form* $[0 \;\; 0 \;\; \cdots \;\; 0 \;\; b]$ *and the coefficient matrix of the system correspond to echelon form has the same number of nonzero rows and columns, then the linear system has only one solution.*

(3) *if none of the above happened, then the linear system has infinitely many solutions.*

Example 1.10. *Determine the existence and uniqueness of the solutions of the system*

$$\begin{array}{cccccc} & 3x_2 & -6x_3 & +6x_4 & +4x_5 & = -5 \\ 3x_1 & -7x_2 & +8x_3 & -5x_4 & +8x_5 & = 9 \\ 3x_1 & -9x_2 & +12x_3 & -9x_4 & +6x_5 & = 15 \end{array}$$

Answer: *The augmented matrix is*

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

Look at the Example 1.8, the echelon form of augmented matrix of this linear system is

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

Since it does not have a row of the form $[0 \;\; 0 \;\; \ldots \;\; 0 \;\; b]$ *it is consistent.*
Also, the coefficient matrix of the system correspond to ehcolen form is

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

does not have equal number of nonzero rows and columns, so the system has infinitely many solutions.

1.3.1. *Solutions of linear systems.* Suppose, for example, that the augmented matrix of a linear system has been changed into the equivalent reduced echelon form

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

There are three variables because augmented matrix has four columns. The associate system of equation is

$$\begin{array}{cccc} x_1 & & -5x_3 & = 1 \\ & x_2 & +x_3 & = 4 \\ & & 0 & = 0 \end{array}$$

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

Solve the equations for basic variables.

$$\begin{cases} x_1 = 1 + 5x_3 \\ x_2 = 4 - x_3 \\ x_3 \ \ \textbf{is free} \end{cases}$$

Let $x_3 = t$. Then all solutions are

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 1+5t \\ 4-t \\ t \end{bmatrix} = \begin{bmatrix} 1 \\ 4 \\ 0 \end{bmatrix} + t \begin{bmatrix} 5 \\ -1 \\ 1 \end{bmatrix}.$$

So the solution set is

$$\left\{ \begin{bmatrix} 1 \\ 4 \\ 0 \end{bmatrix} + t \begin{bmatrix} 5 \\ -1 \\ 1 \end{bmatrix} : t \in \mathbb{R} \right\}$$

Using Row Reduction to Solve a Linear System:

(1) Write the augmented matrix of the system.
(2) By row reduction algorithm, Find an echelon form of the augmented matrix, then check the number of solutions by Theorem 1.9. If it does not have solution stop.
(3) Continue to obtain the reduced echelon form.
(4) Write the equations corresponding to the reduced echelon form in Step 3.
(5) Solve equations in a way that each basic variable is expressed in terms of free variables. Then write the set of all solutions.

Remark. For a consistent system if we do not have free variables it means the system has only one solution.

Example 1.11. *Previously in Example 1.8, the row reduced form of the augmented matrix was*

$$\begin{bmatrix} 1 & 0 & -2 & 3 & 0 & -24 \\ 0 & 1 & -2 & 2 & 0 & -7 \\ 0 & 0 & 0 & 0 & 1 & 4 \end{bmatrix}$$

The equation associated to the reduced echelon form are

$$\begin{array}{ccccccc} x_1 & & -2x_3 & +3x_4 & & = -24 \\ & x_2 & -2x_3 & +2x_4 & & = -7 \\ & & & & x_5 & = 4 \end{array}$$

The pivot columns are column 1, column 2 and column 5. Therefore, the basic variables are x_1, x_2, x_5, and the free variables are x_3 and x_4. We solve the equations in terms of free variables, we have

$$\begin{gathered} x_1 = -24 + 2x_3 - 3x_4 \\ x_2 = -7 + 2x_3 - 2x_4 \\ x_5 = 4 \end{gathered}$$

Let $x_3 = t$ and $x_4 = s$. Then

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix} = \begin{bmatrix} -24 & +2t & -3s \\ -7 & +2t & -2s \\ & t & \\ & & s \\ 4 & & \end{bmatrix} = \begin{bmatrix} -24 \\ -7 \\ 0 \\ 0 \\ 4 \end{bmatrix} + t \begin{bmatrix} 2 \\ 2 \\ 1 \\ 0 \\ 0 \end{bmatrix} + s \begin{bmatrix} -3 \\ -2 \\ 0 \\ 1 \\ 0 \end{bmatrix}$$

So the solution set is

$$\left\{ \begin{bmatrix} -24 \\ -7 \\ 0 \\ 0 \\ 4 \end{bmatrix} + t \begin{bmatrix} 2 \\ 2 \\ 1 \\ 0 \\ 0 \end{bmatrix} + s \begin{bmatrix} -3 \\ -2 \\ 0 \\ 1 \\ 0 \end{bmatrix} : t, s \in \mathbb{R} \right\}$$

2. Week 2: Vector Equations and The Matrix Equation $Ax = b$

Goals of the second week:

(1) Linear combination.
(2) Span of a set of vectors.
(3) Solutions of $Ax = b$, when A is an $m \times n$ matrix and $b \in \mathbb{R}^n$.
(4) When a set of vectors spans $\mathbb{R}^m$.

2.1. **Week 2, Lecture 1, Sept. 4, 2017, Vector Equations.** A matrix with only one column is called a **vector**. A vector in $\mathbb{R}^2$ is of the form $\begin{bmatrix} a \\ b \end{bmatrix}$. For example, $\begin{bmatrix} -1 \\ 2 \end{bmatrix}$ is a vector in $\mathbb{R}^2$. A vector in $\mathbb{R}^3$ is of the form $\begin{bmatrix} a \\ b \\ c \end{bmatrix}$. For example, $\begin{bmatrix} 5 \\ \sqrt{2} \\ -10 \end{bmatrix}$ is a vector in $\mathbb{R}^3$.

Given a vector u and a real number c, the **scalar multiplication** of u by c is the vector obtained by multiplying each entry in u by c. The number c, in cu is called a **scalar**.

Example 2.1. *Let* $u = \begin{bmatrix} 1 \\ -2 \\ 3 \end{bmatrix}$ *and* $v = \begin{bmatrix} 4 \\ 1 \\ -5 \end{bmatrix}$ *are both in* $\mathbb{R}^3$.

$$2u = \begin{bmatrix} 2 \\ -4 \\ 6 \end{bmatrix}$$

$$2u + (-3)v = \begin{bmatrix} 2 \\ -4 \\ 6 \end{bmatrix} + \begin{bmatrix} -12 \\ -3 \\ 15 \end{bmatrix} = \begin{bmatrix} -10 \\ -7 \\ 21 \end{bmatrix}$$

If n is a positive integer, $\mathbb{R}^n$ denotes the set of all vectors of the form $\begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix}$, where $a_1, a_2, \dots, a_n \in \mathbb{R}$.

2.2. **Week 2, Lecture 2, Sept 6, 2017, Linear Combination.** Given vectors $v_1, v_2, \ldots, v_p$ in $\mathbb{R}^n$ and given scalers $c_1, c_2, \ldots, c_p$, the vector y defined by

$$y = c_1v_1 + c_2v_2 + \cdots + c_pv_p$$

is called a linear combination of $v_1, v_2, \ldots, v_p$ with weights $c_1, c_2, \ldots, c_p$. For example, $y = \sqrt{3}v_1 + v_2$ is a linear combination of v_1 and v_2, and also $z = 4\begin{bmatrix} 2 \\ -4 \\ 6 \end{bmatrix} - 5\begin{bmatrix} -12 \\ -3 \\ 15 \end{bmatrix}$ is a linear combination of $\begin{bmatrix} 2 \\ -4 \\ 6 \end{bmatrix}$ and $\begin{bmatrix} -12 \\ -3 \\ 15 \end{bmatrix}$.

Example 2.2. *Let* $a_1 = \begin{bmatrix} 1 \\ -2 \\ -5 \end{bmatrix}$, $a_2 = \begin{bmatrix} 2 \\ 5 \\ 6 \end{bmatrix}$ *and* $b = \begin{bmatrix} 7 \\ 4 \\ 3 \end{bmatrix}$. *Determine, weather b can be generated (written) as a linear combination of a_1 and a_2.*

Solution. *If b is a linear combination of a_1 and a_2, then there are scalars x_1 and x_2 in $\mathbb{R}$ such that*

$$x_1\begin{bmatrix} 1 \\ -2 \\ -5 \end{bmatrix} + x_2\begin{bmatrix} 2 \\ 5 \\ 6 \end{bmatrix} = \begin{bmatrix} 7 \\ 4 \\ 3 \end{bmatrix}$$

Thus

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

which means

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

, and so we have the following linear system

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

This system has the augmented matrix $\begin{bmatrix} 1 & 2 & 7 \\ -2 & 5 & 4 \\ -5 & 6 & 3 \end{bmatrix}$ *and this matrix has the row reduced form* $\begin{bmatrix} 1 & 0 & 3 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \end{bmatrix}$. *Therefore,* $\begin{cases} x_1 = 3 \\ x_2 = 2 \end{cases}$. *So we have*

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

Thus, b is a linear combination of a_1 and a_2.

Remark. *As we see in the previous Example, a vector equation*

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

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

$$[\mathbf{a}_1 \quad \mathbf{a}_2 \quad \cdots \quad \mathbf{a}_n \quad \mathbf{b}] \qquad\qquad (*)$$

In particular, $\mathbf{b}$ can be generated by a linear combination of $\mathbf{a}_1, \mathbf{a}_2, \cdots, \mathbf{a}_n$ if and only if there exists a solution to the linear system corresponding to the matrix $(*)$.

Definition. *If $v_1, \ldots, v_p$ are in $\mathbb{R}^n$, the Span$\{v_1, \ldots, v_p\} = \{c_1v_1 + \ldots + c_pv_p : c_i \in \mathbb{R}\}$ is called the subset of $\mathbb{R}^n$* **spanned (generated)** *by $v_1, \ldots, v_p$) and it is the set of all linear combinations of $v_1, \ldots, v_p$.*

2.3. **Week 2, Lecture 3, Sept 8, 2017, The solutions of $AX = b$.**

Definition. *We denote by $M_{m,n}(\mathbb{R})$ the set of all $m \times n$ matrices with entries in $\mathbb{R}$. If A is in $M_{m,n}(\mathbb{R})$, with columns $A_1, A_2, \dots, A_n$, and x is in $\mathbb{R}^n$, then the product of A by x, denoted by Ax, is*

$$Ax = [A_1|A_2|\cdots|A_n]\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = x_1A_1 + x_2A_2 + \cdots + x_nA_n.$$

Example 2.3.

$$Ax = \begin{bmatrix} 1 & 2 & -1 \\ 0 & -5 & 3 \end{bmatrix}\begin{bmatrix} 4 \\ 3 \\ 7 \end{bmatrix} = 4\begin{bmatrix} 1 \\ 0 \end{bmatrix} + 3\begin{bmatrix} 2 \\ -5 \end{bmatrix} + 7\begin{bmatrix} -1 \\ 3 \end{bmatrix} = \begin{bmatrix} 3 \\ 6 \end{bmatrix}.$$

Theorem 2.4. *If $A = [A_1|A_2|\cdots|A_n] \in M_{m,n}(\mathbb{R})$, and $b \in \mathbb{R}^n$, then the matrix equation*

$$Ax = b$$

has the same solution set as the vector equations

$$x_1A_1 + x_2A_2 + \ldots + x_nA_n = b$$

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

$$[A_1|A_2|\cdots|A_n|b].$$

Example 2.5. *Let* $A = \begin{bmatrix} 1 & 3 & 4 \\ -4 & 2 & -6 \\ -3 & -2 & -7 \end{bmatrix}$ *and* $b = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}$. *Is the equation $Ax = b$ consistent for all possible b_1, b_2, and b_3.*

Solution. *By the Theorem above, $Ax = b$ has the same solution set as the system of linear equations whose augmented matrix is*

$$\begin{bmatrix} 1 & 3 & 4 & b_1 \\ -4 & 2 & -6 & b_2 \\ -3 & -2 & -7 & b_3 \end{bmatrix}.$$

An echelon form of this matrix is

$$\begin{bmatrix} 1 & 3 & 4 & b_1 \\ 0 & 14 & 10 & b_2 + 4b_1 \\ 0 & 0 & 0 & b_3 + 3b_1 - 1/2(b_2 + 4b_1) \end{bmatrix}.$$

Since we have a row of the form $[0 \quad 0 \quad 0 \quad b_3 + 3b_1 - 1/2(b_2 + 4b_1)]$, and for $b_1 = b_2 = b_3 = 2$, $b_3 + 3b_1 - 1/2(b_2 + 4b_1) = 3$, we conclude the the system is not consistent for all possible b_1, b_2, and b_3.

Definition. *A set of vectors* $\{v_1, v_2, \dots, v_p\}$ *in* $\mathbb{R}^m$ **spans (or generate)** $\mathbb{R}^m$ *if every vector in* $\mathbb{R}^n$ *is a linear combination of* $v_1, v_2, \dots, v_p$*, that is, if* $\text{Span}\{v_1, v_2, \dots, v_p\} = \mathbb{R}^m$*.*

Theorem 2.6. *Let A be an $m \times n$ matrix. Then the following are equivalent.*

(1) *For each $b \in \mathbb{R}^m$, the equation $Ax = b$ has a solution.*
(2) *Each $b \in \mathbb{R}^m$ is a linear combination of the columns of A.*
(3) *The columns of A span $\mathbb{R}^m$.*
(4) *A has a pivot position in every row.*

Example 2.7. *Does the set of vectors* $v_1 = \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}, v_2 = \begin{bmatrix} -7 \\ 0 \\ 7 \end{bmatrix}, v_3 = \begin{bmatrix} 0 \\ 1 \\ -4 \end{bmatrix}$, *and* $v_4 = \begin{bmatrix} 6 \\ -2 \\ 0 \end{bmatrix}$ *span* $\mathbb{R}^3$*.*

Lemma 2.8. *Let A be a matrix. The leftmost nonzero entry in any row of an echelon form of A corresponds to a pivot.*

Solution. By Theorem 2.6, this set of vectors span $\mathbb{R}^3$ if a matrix with columns v_1, v_2, v_3, and v_4 has a pivot position in every row. Let

$$A = [v_1|v_2|v_3|v_4] = \begin{bmatrix} 1 & -7 & 0 & 6 \\ 0 & 0 & 1 & -2 \\ -1 & 7 & -4 & 0 \end{bmatrix}.$$

Then an echelon form of this matrix is

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

and it has a pivot position in every row. So the set of vectors $\{v_1, v_2, v_3, v_4\}$ spans $\mathbb{R}^3$.

A system of linear equations is said to be **homogeneous** if it can be written in the form $Ax = \mathbf{0}$, where $A \in M_{m,n}(\mathbb{R})$ and $\mathbf{0}$ is the zero matrix in $\mathbb{R}^m$. Note that homogeneous systems always are consistent since at least $x = \mathbf{0}$ is a solution of the homogeneous systems.

Theorem 2.9. *Suppose the equation $Ax = b$ is consistent for some given vector b, and let $\mathbf{p}$ be a solution. Then the solution set of $Ax = b$ is the set of all vectors of the form $\mathbf{w} = \mathbf{p} + \mathbf{v}$, where v is any solution of the homogeneous equation $Ax = b$.*

Example 2.10. *Determine if the following homogeneous system has a nontrivial solution.*

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

The augmented matrix is

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

The echelon form is

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

(Just to see it has nontrivial solution) The reduced echelon form is

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

Now you can find the solutions.

3. Week 3: Linear independent sets and linear transformations

Goals of the third week:

(1) Linear independent and related theorems
(2) Linear transformations
(3) Matrix of a linear transformation
(4) When a linear transformation is onto and when it is one-to-one.

3.1. Week 3, Lecture 1, Sept 11, Linear independent sets.

An indexed set of vectors $\{v_1, v_2, \ldots, v_p\}$ in $\mathbb{R}^n$ is said to be **linearly independent** if the vector equation

$$x_1v_1 + x_2v_2 + \ldots + x_pv_p = 0$$

has only trivial solution (i.e., the only solution is $x_1 = x_2 = \ldots = x_p = 0$). The set $\{v_1, v_2, \ldots, v_p\}$ is said to be **linearly dependent** if there exist weights $c_1, c_2, \ldots, c_p$, not all zero, such that

$$c_1v_1 + c_2v_2 + \ldots + c_pv_p = 0 \qquad (2).$$

The expression (2) is called a **linear dependence relation** among $v_1, \ldots, v_p$ when the weights are not all zero.

Example 3.1. *Let* $v_1 = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$, $v_2 = \begin{bmatrix} 4 \\ 5 \\ 6 \end{bmatrix}$, *and* $v_3 = \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix}$.

(1) *Determine if the set* $\{v_1, v_2, v_3\}$ *is linearly independent.*
(2) *If possible, find a linear dependence relation among* v_1, v_2, *and* v_3.

Solution. The vectors v_1, v_2, and v_3 are linearly dependent if the equation

$$x_1v_1 + x_2v_2 + x_3v_3 = 0,$$

i.e.,

$$x_1 \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} + x_2 \begin{bmatrix} 4 \\ 5 \\ 6 \end{bmatrix} + x_3 \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

has a nontrivial solution. This equation turns to

$$\begin{array}{cccc} x_1 & +4x_2 & +2x_3 & = 0 \\ 2x_1 & +5x_2 & +x_3 & = 0 \\ 3x_1 & +6x_2 & & = 0 \end{array}$$

So the augmented matrix is

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

To solve the linear system above we find the reduced form of the augmented matrix which is the following matrix

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

Since we have a free variable then we have infinitely many solution so the vector equation has a non-trivial solution, therefore, v_1, v_2 and v_3 are linearly dependent.

Now we want to find a linear dependency relation. Note that we can write the linear system correspond to reduced echelon form as

$$\begin{array}{cccc} x_1 & & -2x_3 & = 0 \\ & x_2 & +x_3 & = 0 \\ & & 0 & = 0 \end{array}$$

Let $x_3 = t$. Then the set of solution is

$$\left\{ \begin{bmatrix} 2t \\ -t \\ t \end{bmatrix} : t \in \mathbb{R} \right\}.$$

Take $t = 1$, then $x_1 = 2, x_2 = -1$, and $x_3 = 1$, and so

$$2\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} - \begin{bmatrix} 4 \\ 5 \\ 6 \end{bmatrix} + \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

is a linear dependence relation among v_1, v_2, and v_3.

Remark. The columns of a matrix A are linearly independent if and only if the equation $Ax = 0$ has only the trivial solution.

Example 3.2. *Determine if the columns of the matrix*

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

are linearly independent.

Solution. We only need to solve the equation $Ax = 0$, then if the equation has only one solution $\mathbf{0}$, then the columns of A are linearly independent. The augmented matrix correspond to the equation $Ax = 0$ is $[A|0]$, i.e.,

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

and the echelon form is

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

Therefore, by Theorem 1.9, since the coefficient matrix correspond to the linear system of the echelon form has the equal number of rows and columns then $Ax = 0$ has only one solution and so the columns of the matrix are linearly independent.

Theorem 3.3. *A set of vectors $S = \{v_1, \ldots, v_p\}$ of two or more vectoes is linearly dependent if and only if at least one of the vectors is S is a linear combination of the others.*

Theorem 3.4. *If a set contains more vectors than there are entries in each vector, then the set is linearly dependent. That is, $v_1, v_2, \ldots, v_p$ in $\mathbb{R}^n$ are linearly if $p > n$.*

Theorem 3.5. *If a set $S = \{v_1, v_2, \ldots, v_p\}$ in $\mathbb{R}^n$ contains the zero vector, then the set S is linearly dependent.*

Proof. Let $v_i = 0$. Then for every $c \neq 0$,

$$0v_1 + \ldots + c_i v_i + \ldots + 0v_p = 0.$$

Therefore, $\{v_1, \ldots, v_p\}$ is linearly dependent. $\square$

Example 3.6. *The vectors $v_1 = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$, $v_2 = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$, and $v_3 = \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix}$ are linearly dependent.*

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

Definition. *A transformation $T : \mathbb{R}^n \to \mathbb{R}^m$ is called a* **linear** *if*

(1) *$T(u + v) = T(u) + T(v)$ for all $u, v \in \mathbb{R}^n$.*
(2) *$T(cu) = cT(u)$ for all scalars c and $u \in \mathbb{R}^n$.*

3.2. **Week 3, Lecture 2, Sept. 13, Linear Transformations.** A **transformation (or function or mapping)** T from $\mathbb{R}^n$ to $\mathbb{R}^m$ is a rule that assigns to each vector x in $\mathbb{R}^n$ a vector $T(x)$ in $\mathbb{R}^m$. We use the notation

$$T : \mathbb{R}^n \longrightarrow \mathbb{R}^m.$$

The set $\mathbb{R}^n$ is called the **domain** of T and $\mathbb{R}^m$ is called the **codomain** of T. For $x \in \mathbb{R}^n$, the vector $T(x) \in \mathbb{R}^m$ is called the **image** of x under the action of T. The set of all $T(x)$ is called the **range or image** of T.

Example 3.7. *Let* $A = \begin{bmatrix} 1 & -3 \\ 3 & 5 \\ -1 & 7 \end{bmatrix}$, $u = \begin{bmatrix} 2 \\ -1 \end{bmatrix}$, $b = \begin{bmatrix} 3 \\ 2 \\ -5 \end{bmatrix}$, $c = \begin{bmatrix} 3 \\ 2 \\ 5 \end{bmatrix}$ *and define transformation* $T : \mathbb{R}^2 \longrightarrow \mathbb{R}^3$ *by* $T(x) = Ax$.

(1) *Find $T(u)$, the image of u under the transformation T.*
(2) *Find an $x \in \mathbb{R}^2$ whose image under T is b.*
(3) *Is there more than one x whose image under T is b.*
(4) *Determine if c is in the range of the transformation T.*

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

(2) We want to find a vector $x \in \mathbb{R}^2$ such that $T(x) = \begin{bmatrix} 3 \\ 2 \\ 05 \end{bmatrix}$. So we want to solve the equation $T(x) = Ax = \begin{bmatrix} 3 \\ 2 \\ -5 \end{bmatrix}$, i.e.,

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

the augmented matrix corresponding to this equation is

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

and its reduced echelon form is

$$\begin{bmatrix} 1 & 0 & 1.5 \\ 0 & 1 & -0.5 \\ 0 & 0 & 0 \end{bmatrix} (**)$$

so $x_1 = 1.5$ and $x_2 = -0.5$. Therefore, $x = \begin{bmatrix} 1.5 \\ -0.5 \end{bmatrix}$. So, $T(\begin{bmatrix} 1.5 \\ -0.5 \end{bmatrix}) = \begin{bmatrix} 3 \\ 2 \\ -5 \end{bmatrix}$.

(3) No. Any x whose image under T is b must satisfy $(*)$. Since the reduced echelon form of the equation in $(*)$ is $(**)$, We can see that $(*)$ has only one solution.

(4) The vector c is in the range of T if there exists $x \in \mathbb{R}^2$ such that $T(x) = c$, i.e.,

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

The augmented matrix of this equation is

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

and the reduced echelon form is

$$\begin{bmatrix} 1 & -3 & 3 \\ 0 & 1 & 2 \\ 0 & 0 & -35 \end{bmatrix}.$$

Since we have a row of the form

$$[0 \quad 0 \cdots 0 \quad 35]$$

this equation does not have a solution. So $\begin{bmatrix} 3 \\ 2 \\ -5 \end{bmatrix}$ *is not in the range of T.*

A transformation (or mapping) T is **linear** if

(1) $T(u + v) = T(u) + T(v)$ for all u and v in the domain of T.
(2) $T(cu) = cT(u)$ for all scalar c and all u in the domain of T.

Theorem 3.8. *If T is a linear transformation,*

(1) $T(0) = 0$.
(2) $T(cu + dv) = cT(u) + dT(v)$ *for all vectors u, v in the domain of T and all scalars c and d.*
(3) $T(c_1v_1 + \ldots + c_pv_p) = c_1T(v_1) + \ldots + c_pT(v_p)$ *for all vectors $v_1, \ldots, v_p$ in the domain of T and scalars $c_1, \ldots, c_p$ in $\mathbb{R}$.*

Example 3.9. *Given a scalar r, define $T : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ by $T(x) = rx$. Show that T is a linear transformation.*

proof. We must show that $T(u+v) = T(u)+T(v)$ and $T(cu) = cT(u)$ for any vector $u, v \in \mathbb{R}^2$ and $c \in \mathbb{R}$. Note that $T(u + v) = r(u + v) = ru + rv = T(u) + T(v)$, and also $T(cu) = r(cu) = (rc)u = (cr)u = c(ru) = cT(u)$. Therefore, T is linear.

3.3. **Week 3, Lecture 3, Sept. 15, The matrix of a linear transformation.** We start with a theorem.

Theorem 3.10. *Let $T : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ be a linear transformation. Then there exists a unique matrix A such that*

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

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

$$A = [T(e_1)| \dots |T(e_n)]$$

The matrix A in the above theorem is called the **standard matrix for the linear transformation** T.
proof. Write

$$x = I_n x = [e_1| \dots |e_n] \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = x_1 e_1 + \dots + x_n e_n,$$

and use the linearity of T to compute

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

$$[T(e_1)| \dots |T(e_n)] \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = Ax.$$

Example 3.11. *Find the standard matrix A for the dilation transformation $T(x) = 3x$ for $x \in \mathbb{R}^2$.*

solution. Write

$$T(e_1) = 3e_1 = 3 \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 3 \\ 0 \end{bmatrix} \textit{and}$$

$$T(e_2) = 3e_2 = 3 \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 3 \end{bmatrix}$$

Therefore

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

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

Example 3.12. *Let T be the linear transformation whose standard matrix is*

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

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

Solution. First note that $T(x) = Ax$, we can also write that

$$\begin{array}{rccc} T: & \mathbb{R}^4 & \longrightarrow & \mathbb{R}^3 \\ & x & \mapsto & Ax \end{array}$$

Onto. T is onto if each $b \in \mathbb{R}^3$ is in the range of T, i.e., for every $b \in \mathbb{R}^3$, there is $x \in \mathbb{R}^4$ such that $T(x) = b$, that is, $T(x) = Ax = b$.

Therefore, if $Ax = b$ has a solution for every $b \in \mathbb{R}^3$, then T is onto. Since A has a pivot position in each row, thus by Theorem 2.6, $Ax = b$ has a solution for every $b \in \mathbb{R}^3$, therefore T is onto.

One-to-one. The mapping T is one-to-one if each $b \in \mathbb{R}^3$ is the image of at most one $x \in \mathbb{R}^4$, that is, the equation $Ax = b$ for every $b \in \mathbb{R}^3$ has at most one solution. However we can see that $Ax = b$ has a free variable. Therefore, $Ax = b$ has infinitely many solutions and so T is not one-to-one.

Notation: Let $T : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ be a linear map. Denote

$$Img(T) = \{\ T(x) : \quad x \in \mathbb{R}^n\}$$

and the kernel of T

$$Ker(T) = \{\ x \in \mathbb{R}^n : \quad T(x) = 0\}\,.$$

Theorem 3.13. *Let $T : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ be a linear transformation. Then T is one-to-one if and only if $T(x) = 0$ has only one solution. Therefore, T is one-to-one if $Ker(T) = 0$, and T is onto if $Img(T) = \mathbb{R}^m$.*

proof. Since T is linear $T(0) = 0$. If T is one-to-one, then from the definition of one-to-one $T(x) = 0$ has only one solution. For converse, we proof by contradiction. Assume that $T(x) = 0$ only have one solution, and for some $b \in \mathbb{R}^n$ there are different b_1 and b_2 such that $T(b_1) = b$ and $T(b_2) = b$. Then

$$T(b_2) - T(b_1) = b - b = 0.$$

Note that $T(b_2) - T(b_1) = T(b_2 - b_1) = 0$ Since $T(x) = 0$ has only one solution which is zero. Therefore, $b_2 - b_1 = 0$, and so $b_2 = b_1$, a contradiction since b_2 and b_1 are different.

Theorem 3.14. *Let $T : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ be a linear transformation, and let A be the standard matrix for T. Then:*

(1) *T maps $\mathbb{R}^n$ onto $\mathbb{R}^m$ if and only if the columns of A span $\mathbb{R}^m$.*
(2) *T is one-to-one if the columns of A are linearly independent.*

Example 3.15. *Let*

$$T(x_1, x_2) = (3x_1 + x_2, 5x_1 + 7x_2, x_1 + 3x_2).$$

Show that T is a one-to-one linear transformation. Does T map $\mathbb{R}^2$ onto $\mathbb{R}^3$?

proof. Let rewrite T as

$$\begin{array}{cccc} T: & \mathbb{R}^2 & \longrightarrow & \mathbb{R}^3 \\ & \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} & \mapsto & \begin{bmatrix} 3x_1 + x_2 \\ 5x_1 + 7x_2 \\ x_1 + 3x_2 \end{bmatrix} \end{array}$$

i.e.,

$$T(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}) = \begin{bmatrix} 3x_1 + x_2 \\ 5x_1 + 7x_2 \\ x_1 + 3x_2 \end{bmatrix}.$$

First find the standard matrix of T. Note that

$$T(\begin{bmatrix} 1 \\ 0 \end{bmatrix}) = \begin{bmatrix} 3 \\ 5 \\ 1 \end{bmatrix}$$

and

$$T(\begin{bmatrix} 0 \\ 1 \end{bmatrix}) = \begin{bmatrix} 1 \\ 7 \\ 3 \end{bmatrix}$$

Thus $A = \begin{bmatrix} 3 & 1 \\ 5 & 7 \\ 1 & 3 \end{bmatrix}$. So by previous theorem T is one-to-one if the columns of A are linearly independent. And T is onto if the columns of A span $\mathbb{R}^3$. Therefore T is onto and T is not one-to-one (why?).

4. Week 4: Matrix Operations

Goals of the forth week:

(1) Matrix operations, sum, multiplication, transpose
(2) Inverse of a matrix
(3) Elementary matrices
(4) Algorithm for finding A^{-1}

4.1. Week 4, Lecture 1, Sept. 18, 2017,Matrix operations, sum, product, transpose.

• Sums and scalar multiple:

Let $A = \begin{bmatrix} 3 & 2 \\ -1 & 1 \\ 0 & 1 \end{bmatrix}$ and $B = \begin{bmatrix} 6 & 1 \\ 7 & -10 \\ -1 & 5 \end{bmatrix}$, then

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

$$A - 3B = \begin{bmatrix} 6 & 1 \\ 7 & -10 \\ -1 & 5 \end{bmatrix} - \begin{bmatrix} 18 & 3 \\ 21 & -30 \\ -3 & 15 \end{bmatrix} = \begin{bmatrix} -12 & -2 \\ -14 & 20 \\ 4 & -10 \end{bmatrix}.$$

Theorem 4.1. *Let A, B, and C be matrices of the same size, and let r and s be scalars. Then,*

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

Notation. Let A be an $m \times n$ matrix. We denote by a_{ij} or $(A)_{ij}$ the entry in the row i and column j.

Let $A = \begin{bmatrix} 2 & -1 & 0 \\ 3 & 5 & 1 \\ 6 & 1 & -2 \end{bmatrix}$, Then $a_{12} = (A)_{12} = -1$

And $row_i(A)$ is the ith row of A, $col_j(A)$ is the jth column of A.

$row_2(A) = [3 \;\; 5 \;\; 1]$

$col_3(A) = \begin{bmatrix} 0 \\ 1 \\ -2 \end{bmatrix}$

• when A is $m \times n$ matrix and $B = [b_1| \dots |b_p]$ is an $n \times p$ matrix, then the product AB is an $m \times p$ matrix with

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

also the columns of AB are $Ab_1, \ldots, Ab_p$. That is $col_j(AB) = A.col_j(B)$.

$$AB = A\,[b_1|\ldots|b_p] = [Ab_1|\ldots|Ab_p]$$

and

$$row_i(AB) = row_i(A).B$$

Example 4.2. $B = \begin{bmatrix} -1 & 3 \\ 2 & 1 \\ 7 & 8 \end{bmatrix}$.

Then,

$$AB = \begin{bmatrix} 2 & -1 & 0 \\ 3 & 5 & 1 \\ 6 & 1 & -2 \end{bmatrix} \begin{bmatrix} -1 & 3 \\ 2 & 1 \\ 7 & 8 \end{bmatrix} = \begin{bmatrix} -4 & 5 \\ 14 & 22 \\ -10 & 3 \end{bmatrix}$$

$$row_1(A) = [2 \quad -1 \quad 0]$$

$$row_1(AB) = row_1(A)B = [2 \quad -1 \quad 0] \begin{bmatrix} -1 & 3 \\ 2 & 1 \\ 7 & 8 \end{bmatrix} = [-5 \quad 4]$$

$$col_2(B) = \begin{bmatrix} 3 \\ 1 \\ 8 \end{bmatrix}$$

$$col_2(AB) = A.col_2(B) = \begin{bmatrix} 2 & -1 & 0 \\ 3 & 5 & 1 \\ 6 & 1 & -2 \end{bmatrix} \begin{bmatrix} 3 \\ 1 \\ 8 \end{bmatrix} = \begin{bmatrix} 5 \\ 22 \\ 3 \end{bmatrix}$$

The identity matrix I_n is the matrix

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

Theorem 4.3. *Let A be an $m \times n$ matrix, B an $n \times p$ matrix, C an $n \times k$ matrix, D an $p \times s$ matrix. Then*

(1) *$A(BC) = (AB)C$ (associativity)*
(2) *$A(B + C) = AB + AC$ (left distributive)*
(3) *$(B + C)A = BA + CA$ (right distributive)*
(4) *$r(AB) = (rA)B = A(rB)$ for any scalar r*
(5) *$I_nA = A = AI_n$.*

Example 4.4. *Let*

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

Then

$$row_1(AB) = row_1(A)B = \begin{bmatrix} 1 & 0 \end{bmatrix} \begin{bmatrix} 0 & 0 \\ 1 & 1 \end{bmatrix} = [0 \;\; 0],$$

$$row_2(AB) = row_2(A)B = [-2 \;\; 0] \begin{bmatrix} 0 & 0 \\ 1 & 1 \end{bmatrix} = [0 \;\; 0].$$

Therefore,

$$AB = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$

$$col_1(BA) = Bcol_1(A) = \begin{bmatrix} 0 & 0 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ -2 \end{bmatrix} = \begin{bmatrix} 0 \\ -1 \end{bmatrix}$$

$$col_2(BA) = Bcol_2(A) = \begin{bmatrix} 0 & 0 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

Therefore,

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

Lemma 4.5. (1) *We do not have necessarily $AB = BA$.*

(2) *The cancellation laws do not hold for matrix multiplication. That is, if $AB = AC$, then not necessarily $B = C$.*

(3) *It is possible that $AB = 0$ but $A \neq 0$ and $B \neq 0$. (0 here is a matrix whose all entries are zero).*

• Powers of a matrix: Let A be an $n \times n$ matrix. Then $A^k = AA \cdots A$ k times.
• **Transpose** of a matrix: Let A be an $m \times n$ matrix. Then the transpose of A, denoted by A^T, is an $n \times m$ matrix whose column j is the row j of A.

Example 4.6. *Let*

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

Then

$$A^2 = AA = \begin{bmatrix} 1 & 3 \\ 4 & -1 \end{bmatrix} \begin{bmatrix} 1 & 3 \\ 4 & -1 \end{bmatrix} = \begin{bmatrix} 13 & 0 \\ 0 & 13 \end{bmatrix}$$

$$B^T = \begin{bmatrix} 0 & 1 & 4 & -3 \\ 2 & 1 & 5 & 6 \end{bmatrix}$$

$$(B^T)^T = \begin{bmatrix} 0 & 2 \\ 1 & 1 \\ 4 & 5 \\ -3 & 6 \end{bmatrix}.$$

Theorem 4.7. *Let A be an $m \times n$ matrix, B an $m \times n$ matrix, and C an $n \times p$ matrix. Then*

(1) $(A^T)^T = A$

(2) $(A+B)^T = A^T + B^T$
(3) *for any scalar* r, $(rA)^T = rA^T$,
(4) $(AB)^T = B^T A^T$.

4.2. **Week 4, Lecture 2, Sept. 20, 2017, The inverse of a matrix.**

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

$$CA = I_n \qquad and \qquad AC = I_n$$

The matrix C is called the inverse of matrix A.
• The inverse of a matrix A is unique and is denoted by A^{-1}. So that

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

Example 4.8. *If* $A = \begin{bmatrix} 2 & 7 \\ -1 & -3 \end{bmatrix}$ *and* $C = \begin{bmatrix} -3 & -7 \\ 1 & 2 \end{bmatrix}$

$$AC = \begin{bmatrix} 2 & 7 \\ -1 & -3 \end{bmatrix} \begin{bmatrix} -3 & -7 \\ 1 & 2 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

$$CA = \begin{bmatrix} -3 & -7 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} 2 & 7 \\ -1 & -3 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

Theorem 4.9. *Let* $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$. *If* $\ ad - bc \neq 0$, *then* A *is invertible and*

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

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

• For a 2×2 matrix the determination of A is

$$detA = ad - bc.$$

Example 4.10. *Let* $A = \begin{bmatrix} 1 & 4 \\ 3 & -1 \end{bmatrix}$. *Then*

$$det(A) = -1 - 12 = -13 = ad - bc$$

and

$$A^{-1} = 1/(ad - bc) \begin{bmatrix} d & -b \\ -c & a \end{bmatrix} = 1/-13 \begin{bmatrix} -1 & -4 \\ -3 & 1 \end{bmatrix} = \begin{bmatrix} 1/13 & 4/13 \\ 3/13 & -1/13 \end{bmatrix}.$$

Theorem 4.11. *If* A *is an* $n \times n$ *matrix , then for every* $b \in \mathbb{R}^n$, *the equation* $Ax = b$ *has the unique solution* $x = A^{-1}b$.

Example 4.12. *Solve the following equation*

$$x_1 + 4x_2 = 1$$

$$3x_1 - x_2 = 3$$

Solution. This equation is equivalent to the equation

$$\begin{bmatrix} 1 & 4 \\ 3 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 1 \\ 3 \end{bmatrix}.$$

By provious theorem
$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 1 & 4 \\ 3 & -1 \end{bmatrix}^{-1} \begin{bmatrix} 1 \\ 3 \end{bmatrix}$$
is the unique solution. So
$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 1/13 & 4/13 \\ 3/13 & -1/13 \end{bmatrix} \begin{bmatrix} 1 \\ 3 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}.$$

Theorem 4.13. (1) *the inverse of the inverse of a matrix A is A, i.e.,*
$$(A^{-1})^{-1} = A$$
(2) $(AB)^{-1} = B^{-1}A^{-1}$, *if A and B are invertible AB is so.*
(3) $(A^T)^{-1} = (A^{-1})^T$, *if A is invertible, A^T is so.*

• **Elementary Matrices:** The result of operating a single elementary row operation on identity matrix is called an elementary matrix.

Example 4.14.
$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \xleftrightarrow{\mathrm{R}_3 \leftrightarrow \mathrm{R}_3 + 2\mathrm{R}_2} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 2 & 1 \end{bmatrix}$$
$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \xleftrightarrow{\textit{Scaling R1 by 2}} \begin{bmatrix} 2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

• When we operate an elementary row operation on matrix A , the result is the same as EA where E is an $m \times m$ matrix created by performing the same row operation on I.

Example 4.15. *Let* $A = \begin{bmatrix} 2 & -1 & 3 \\ 0 & 1 & 5 \\ -1 & 2 & 6 \end{bmatrix}$.

Then
$$\begin{bmatrix} 2 & -1 & 3 \\ 0 & 1 & 5 \\ -1 & 2 & 6 \end{bmatrix} \xleftrightarrow{\mathrm{R}_3 \leftrightarrow \mathrm{R}_3 + (-1)\mathrm{R}_1} \begin{bmatrix} 2 & -1 & 3 \\ 0 & 1 & 5 \\ -3 & 3 & 3 \end{bmatrix}.$$
$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \xleftrightarrow{\mathrm{R}_3 \leftrightarrow \mathrm{R}_3 + (-1)\mathrm{R}_1} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -1 & 0 & 1 \end{bmatrix} = E$$
$$\textit{Then } EA = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -1 & 0 & 1 \end{bmatrix} \begin{bmatrix} 2 & -1 & 3 \\ 0 & 1 & 5 \\ -1 & 2 & 6 \end{bmatrix} = \begin{bmatrix} 2 & -1 & 3 \\ 0 & 1 & 5 \\ -3 & 3 & 3 \end{bmatrix}.$$

• Every elementary matrix E is invertible. The inverse of E is the elementary matrix correspond to the elementary operation that transforms E back into I_m.

Example 4.16. *Let* $E = \begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$*, then*

$$\begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \xleftrightarrow{\mathrm{R}_2 \leftrightarrow \mathrm{R}_2 + (-2)\mathrm{R}_1} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Therefore,

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \xleftrightarrow{\mathrm{R}_2 \leftrightarrow \mathrm{R}_2 + (-2)\mathrm{R}_1} \begin{bmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

So we have $\begin{bmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$ *is the* E^{-1}.

4.3. **Week 4, Lecture 3, Sept. 22, 2017, Algorithm for finding A^{-1}.** Let A be an $n \times n$ matrix. Then the following process will give us A^{-1}.

(1) Start with an $n \times 2n$ matrix whose the first n columns are columns of A and the last n columns are columns of I_n, i.e., $[A|I_n]$.
(2) If the reduced echelon form of A is I_n, A is invertible, otherwise A is not invertible.
(3) By elementary row operations on $[A|I_n]$, change $[A|I_n]$ to a matrix of the form $[I_n|B]$.
(4) B is the inverse of A, that is $A^{-1} = B$.

Example 4.17. *Let*

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

Then

$$\left[\begin{array}{ccc|ccc} 0 & 4 & -1 & 1 & 0 & 0 \\ 1 & -2 & 1 & 0 & 1 & 0 \\ 3 & 0 & 1 & 0 & 0 & 1 \end{array}\right] \xleftrightarrow{\textit{Interchanging } R1 \textit{ and } R2} \left[\begin{array}{ccc|ccc} 1 & -2 & 1 & 0 & 1 & 0 \\ 0 & 4 & -1 & 1 & 0 & 0 \\ 3 & 0 & 1 & 0 & 0 & 1 \end{array}\right]$$

$$\xleftrightarrow{\textit{Replacing } R3 \textit{ and } R3+(-3)R1} \left[\begin{array}{ccc|ccc} 1 & -2 & 1 & 0 & 1 & 0 \\ 0 & 4 & -1 & 1 & 0 & 0 \\ 0 & 6 & -2 & 0 & -3 & 1 \end{array}\right]$$

$$\xleftrightarrow{\textit{Replacing } R3 \textit{ and } R3+(-3/2)R2} \left[\begin{array}{ccc|ccc} 1 & -2 & 1 & 0 & 1 & 0 \\ 0 & 4 & -1 & 1 & 0 & 0 \\ 0 & 0 & -1/2 & -3/2 & -3 & 1 \end{array}\right]$$

$$\xleftrightarrow{\textit{Scaling } R3 \textit{ by } -2} \left[\begin{array}{ccc|ccc} 1 & -2 & 1 & 0 & 1 & 0 \\ 0 & 4 & -1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 3 & 6 & -2 \end{array}\right]$$

$$\begin{array}{c} \xleftrightarrow{\textit{Replacing } R2 \textit{ by } R2+R3} \\ \xleftrightarrow{\textit{Replacing } R1 \textit{ by } R1-R3} \end{array} \left[\begin{array}{ccc|ccc} 1 & -2 & 0 & -3 & -5 & 2 \\ 0 & 4 & 0 & 4 & 6 & -2 \\ 0 & 0 & 1 & 3 & 6 & -2 \end{array}\right]$$

$$\xleftrightarrow{\textit{Scaling } R2 \textit{ by } 1/4} \left[\begin{array}{ccc|ccc} 1 & -2 & 0 & -3 & -5 & 2 \\ 0 & 1 & 0 & 1 & 3/2 & -1/2 \\ 0 & 0 & 1 & 3 & 6 & -2 \end{array}\right]$$

$$\xleftrightarrow{\textit{Replacing } R1 \textit{ by } R1+2R2} \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & -1 & -2 & 1 \\ 0 & 1 & 0 & 1 & 3/2 & -1/2 \\ 0 & 0 & 1 & 3 & 6 & -2 \end{array}\right]$$

Therefore,

$$A^{-1} = \begin{bmatrix} -1 & -2 & 1 \\ 1 & 3/2 & -1/2 \\ 3 & 6 & -2 \end{bmatrix}$$

Theorem 4.18. *An $n \times n$ matrix is invertible if and only if A is row equivalent to I_n; moreover, any sequence of elementary row operations that reduces A to I_n also transforms I_n into A^{-1}.*

The theorem above means, if

$$E_1 E_2 \dots E_k A = I_n$$

then

$$A^{-1} = E_k^{-1} \dots E_2^{-1} E_1^{-1} I_n.$$

5. WEEK 5, SUBSPACES OF $\mathbb{R}^n$

Goals of the week:

- A subspace of $\mathbb{R}^n$, a basis for the subspace, and the dimension of subspace
- Column space, null space, and rank of a matrix.
- Three theorems: Rank theorem, basis theorem, and the invertible matrix theorem.

5.1. Week 5, Lecture 1, Sept. 25, Subspaces of $\mathbb{R}^n$.

Review. Let $\{v_1, \ldots, v_q\}$ be a set of vectors in $\mathbb{R}^n$.

(1) We say that the vectors $v_1, \ldots, v_q$ are linearly independent if the equation

$$c_1v_1 + \ldots + c_qv_q = 0$$

has only trivial solution.

(2) The span of $\{v_1, \ldots, v_q\}$, denoted by $\mathrm{Span}\{v_1, \ldots, v_q\}$, is the set of all linear combinations of $\{v_1, \ldots, v_q\}$, i.e.,

$$\mathrm{Span}\{v_1, \ldots, v_q\} = \{c_1v_1 + \ldots + c_qv_q : c_1, \ldots, c_q \in \mathbb{R}^n\}.$$

Definition. *A* **subspace** *H of $\mathbb{R}^n$ is any subset in $\mathbb{R}^n$ with three properties:*

(1) The zero vector is in H.

(2) For every u and v in H, $u + v \in H$.

(3) For every scalar c and $v \in H$, $cv \in H$.

Example 5.1. *Let v_1 and v_2 be in $\mathbb{R}^n$. Then $H = \mathrm{Span}\{v_1, v_2\}$ is a subspace of $\mathbb{R}^n$.*

Solution: Note that

$$span\{v_1, v_2\} = \{c_1v_1 + c_2v_2 : c_1, c_2 \in \mathbb{R}\}$$

is a subset of $\mathbb{R}^n$, we only need to check the three properties.

(1) Let $c_1 = 0$, $c_2 = 0$. then $0v_1 + 0v_2 = 0 \in span\{v_1, v_2\}$

(2) Let $u, v \in span\{v_1, v_2\}$. Then there are scalars, c_1 and c_2, d_1 and d_2 such that $v = c_1v_1 + c_2v_2$ and $u = d_1v_1 + d_2v_2$. Therefore,

$$v + u = c_1v_1 + c_2v_2 + d_1v_1 + d_2v_2 = (c_1 + d_1)v_1 + (c_2 + d_2)v_2$$

is in $span\{v_1, v_2\}$.

(3) Let c be an scalar , and $v \in span\{v_1, v_2\}$. Then there are $c_1, c_2 \in \mathbb{R}$ such that $v = c_1v_1 + c_2v_2$. so,

$$c(c_1v_1 + c_2v_2) = (cc_1)v_1 + (cc_2)v_2 \in span\{v_1, v_2\}.$$

Therefore $span\{v_1, v_2\}$ is a subspace of $\mathbb{R}^n$.

- The **column space** of a matrix $A = [A_1|\ldots|A_n]$, denoted by $ColA$, is

$$ColA = span\{A_1, \ldots, A_n\} = \{c_1A_1 + \ldots + c_nA_n : c_1, \ldots, c_n \in \mathbb{R}\},$$

which is the set of all linear combination of columns of A.

Example 5.2. *Is* $b = \begin{bmatrix} -1 \\ 1 \\ 2 \end{bmatrix}$ *in the column space of* $A = \begin{bmatrix} 0 & -1 & 3 \\ 3 & -2 & 1 \\ -2 & 4 & 1 \end{bmatrix}$.

Solution. *Note that the column space of A is*

$$span\left\{\begin{bmatrix} 0 \\ 3 \\ -2 \end{bmatrix}, \begin{bmatrix} -1 \\ -2 \\ 4 \end{bmatrix}, \begin{bmatrix} 3 \\ 1 \\ 1 \end{bmatrix}\right\} =$$

$$\left\{x_1 \begin{bmatrix} 0 \\ 3 \\ -2 \end{bmatrix} + x_2 \begin{bmatrix} -1 \\ -2 \\ 4 \end{bmatrix} + x_3 \begin{bmatrix} 3 \\ 1 \\ 1 \end{bmatrix} : x_1, x_2, x_3 \in \mathbb{R}\right\}$$

Therefore b is in column space of A if the system

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

has a solution. The augmented matrix of the system is

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

and an echelon form is

$$\begin{bmatrix} 3 & -2 & 1 & 1 \\ 0 & 1 & -3 & 1 \\ 0 & 0 & 29 & 0 \end{bmatrix}.$$

Since the system is consistent, therefore b is in column space of A.

• The **null space** of a matrix A, denoted by $NulA$, is the set of all solutions of the homogeneous system

$$Ax = 0.$$

Theorem 5.3. *The null space of a matrix A is a subspace of $\mathbb{R}^n$. That is, the solution set of $Ax = 0$ is a subspace of $\mathbb{R}^n$. (Do it as an exercise.)*

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

Example 5.4. $e_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$, $e_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$, $e_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$ *is a basis for* $\mathbb{R}^3$. *Since any arbitrary element* $\begin{bmatrix} a \\ b \\ c \end{bmatrix}$ *in* $\mathbb{R}^n$ *can be written as*

$$a \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + b \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} + c \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}.$$

And also

$$x_1 \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + x_2 \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} + x_3 \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = 0$$

has only one solution.

Theorem 5.5. *The pivot columns of a matrix A form a basis for column space of A.*

Example 5.6. *Find a basis for column space of*

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

Solution.

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

$$\begin{array}{c} \text{Replacing} \quad \underset{\longleftrightarrow}{\text{R2 by}} \quad \text{R2+(−2)R1} \\ \text{Replacing} \quad \underset{\longleftrightarrow}{\text{R3}} \quad \text{by} \quad \text{R3+3R1} \end{array} \begin{bmatrix} 1 & -1 & 5 & 1 \\ 0 & 2 & -3 & -1 \\ 0 & -8 & 12 & 5 \end{bmatrix}$$

$$\text{Replacing} \quad \underset{\longleftrightarrow}{\text{R3}} \quad \text{by} \quad \text{R3+4R2} \begin{bmatrix} 1 & -1 & 5 & 1 \\ 0 & 2 & -3 & -1 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

Therefore the set of columns 1, 2 and 4 of A is a basis for column space of A.

5.2. **Week 5, Lecture 2, Basis and Basis Coordinates.** • A **basis** for a subspace H of $\mathbb{R}^n$ is a linearly independent set in H that spans H.

Example 5.7. $e_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$, $e_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$, $e_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$ *is a basis for* $\mathbb{R}^3$. *Since any arbitrary element* $\begin{bmatrix} a \\ b \\ c \end{bmatrix}$ *in* $\mathbb{R}^n$ *can be written as*

$$a\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + b\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} + c\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}.$$

And also

$$x_1\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + x_2\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} + x_3\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = 0$$

has only one solution.

Theorem 5.8. *The pivot columns of a matrix A form a basis for column space of A.*

Example 5.9. *Find a basis for column space of*

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

Solution.

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

$$\begin{matrix} \text{Replacing} \underset{\longleftrightarrow}{\text{R2 by}} \text{R2+}(-2)\text{R1} \\ \text{Replacing} \underset{\longleftrightarrow}{\text{R3}} \text{ by } \text{R3+3R1} \end{matrix} \begin{bmatrix} 1 & -1 & 5 & 1 \\ 0 & 2 & -3 & -1 \\ 0 & -8 & 12 & 5 \end{bmatrix}$$

$$\text{Replacing} \underset{\longleftrightarrow}{\text{R3}} \text{ by } \text{R3+4R2} \begin{bmatrix} 1 & -1 & 5 & 1 \\ 0 & 2 & -3 & -1 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

Therefore the set of columns 1, 2 and 4 of A is a basis for column space of A.

Definition. *Suppose the set $\beta = \{b_1, \ldots, b_p\}$ is a basis for subspace H. For each x in H, the* **coordinate of x relative to basis** β *are the weights $c_1, \ldots, c_p$ such that*

$$x = c_1b_1 + c_2b_2 + \ldots + c_pb_p,$$

and the vector

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

is called the β**-coordinate vector** *of x.*

Example 5.10. *Let* $v_1 = \begin{bmatrix} 1 \\ 2 \\ -3 \end{bmatrix}$, $v_2 = \begin{bmatrix} -1 \\ -6 \\ 5 \end{bmatrix}$, $x = \begin{bmatrix} 1 \\ -2 \\ -1 \end{bmatrix}$.

(1) *Find a basis* B *for* $span\{v_1, v_2\}$.
(2) *Is* x *in* $span\{v_1, v_2\}$.
(3) *Write* $[x]_B$ *the* $B-$ *coordinate vector of* x.

Solution. *(Whenever a set of vectors* $\{v_1, \ldots, v_p\}$ *is given and you want to find a basis for* $span\{v_1, \ldots, v_p\}$ *do this process:*
(1) Construct a matrix $[v_1|\ldots|v_p]$.
(2) Find the pivot positions.
(3) The columns correspond to pivot positions give a basis.)
Look at the following matrix

$$\begin{bmatrix} 1 & -1 \\ 2 & -6 \\ -3 & 5 \end{bmatrix},$$

then an echelon form of this matrix is

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

Since both columns have pivot positions therefore, v_1 *and* v_2 *are a basis for* $span\{v_1, v_2\}$.

(2) $x = \begin{bmatrix} 1 \\ -2 \\ -1 \end{bmatrix}$ *is in* $span\{v_1, v_2\}$ *if the following system is consistent.*

$$x_1 \begin{bmatrix} 1 \\ 2 \\ -3 \end{bmatrix} + x_2 \begin{bmatrix} -1 \\ -6 \\ 5 \end{bmatrix} = \begin{bmatrix} 1 \\ -2 \\ -1 \end{bmatrix}.$$

The augmented matrix is

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

Then an echelon form of this matrix is

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

Thus, the system is consistent and a solution is $x_2 = 1$ *and* $x_1 = 2$. *Thus,*

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

(3) Therefore $[x]_\beta = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$.

5.3. Week 5, Lecture 3, Dimension and Theorems.

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

Example 5.11. *Find a basis and the dimension for*

$$span \left\{ \begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix}, \begin{bmatrix} 3 \\ 2 \\ 1 \end{bmatrix}, \begin{bmatrix} 4 \\ 4 \\ 0 \end{bmatrix}, \begin{bmatrix} 2 \\ 0 \\ 2 \end{bmatrix} \right\}.$$

Solution. (Put the columns on a matrix, the number of pivot position will be the dimension.)

So look at

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

An echelon form of the matrix is

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

Therefore, the pivot position are in row 1 and row 2, therefore, $\{ \begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix}, \begin{bmatrix} 3 \\ 2 \\ 1 \end{bmatrix} \}$ is a basis and the dimension is 2.

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

Theorem 5.12 (The Rank Theorem)**.** *If a matrix A has n columns, then*

$$rank A + dim\ Nul A = n$$

Theorem 5.13 (The basis theorem)**.** *Let H be a subspace of dimension p of $\mathbb{R}^n$.*

(1) *Any linearly independent set of p element automatically span H.*
(2) *If p vectors span H, then they are linearly independent.*

Theorem 5.14 (The inverse matrix theorem)**.** *Let A be an $n \times n$ matrix. Then the following statements are equivalent.*

(1) *A is invertible.*
(2) *The columns of A form a basis of $\mathbb{R}^n$.*
(3) *$Col\ A = \mathbb{R}^n$.*
(4) *$dim\ Col\ A = n$*
(5) *$rank\ A = n$*
(6) *$Nul\ A = \{0\}$*
(7) *$dim\ Nul\ A = 0$.*

Example 5.15. *Let*

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

Find a basis for $NulA$*. What is the dimension of* $NulA$*.*

Solution. The reduced echelon form of the augmented matrix is

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

Therefore, x_1 and x_3 are basic variables and x_2, x_4, and x_5 are free variables. Then

$$\begin{array}{ccccc} x_1 & -2x_2 & -x_4 & +3x_4 & = 0 \\ & x_3 & +2x_4 & -2x_5 & = 0 \end{array}$$

Now, let $x_2 = t, x_4 = r, x_5 = y$. Then

$$x_1 = 2t + r - 3y$$

$$x_3 = -2r + 2y$$

Therefore,

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix} = \begin{bmatrix} 2t + r - 3y \\ t \\ -2r + 2y \\ r \\ y \end{bmatrix} = t \begin{bmatrix} 2 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} + r \begin{bmatrix} 1 \\ 0 \\ -2 \\ 1 \\ 0 \end{bmatrix} + y \begin{bmatrix} -3 \\ 0 \\ 2 \\ 0 \\ 1 \end{bmatrix}$$

Therefore,

$$NulA = span \left\{ \begin{bmatrix} 2 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ -2 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -3 \\ 0 \\ 2 \\ 0 \\ 1 \end{bmatrix} \right\}$$

Moreover, they are linearly independent and so the dimension is 3.

6. Week 6, Determinants

Goals of the week:

- The Determinant of an $n \times n$ matrix and its properties
- First midterm

6.1. Week 6, Lecture 1, Oct.2, Determinant.

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

Example 6.1. *Let*

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

Then

$$A_{12} = \begin{bmatrix} 0 & 10 & 5 \\ -2 & 7 & 0 \\ -5 & 1 & 2 \end{bmatrix}.$$

$$A_{34} = \begin{bmatrix} -1 & 3 & 4 \\ 0 & 1 & 10 \\ -5 & 3 & 1 \end{bmatrix}.$$

Definition. *For $n \geq 2$, the determinant of an $n \times n$ matrix $A = [a_{ij}]$ is*

$$detA = a_{11}detA_{11} - a_{12}detA_{12} + \ldots + (-1)^{n+1}a_{1n}detA_{1n}$$
$$= \Sigma_{j=1}^{n}(-1)^{1+j}a_{1j}detA_{1j}.$$

Example 6.2. *Let*

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

$$detA = -1det\begin{bmatrix} -1 & 0 \\ -2 & 1 \end{bmatrix} - 3det\begin{bmatrix} 7 & 0 \\ 5 & 1 \end{bmatrix}$$
$$+2det\begin{bmatrix} 7 & -1 \\ 5 & -2 \end{bmatrix} = (-1)(-1) - 3(7) + 2(-14+5)$$
$$= 1 - 21 - 18 = -38$$

Definition. *Given $A = [a_{ij}]$, the (i,j)-cofactor of A is*

$$C_{ij} = (-1)^{i+j}detA_{ij}$$

So

$$detA = a_{11}C_{11} + a_{12}C_{12} + \ldots + a_{1n}C_{1n}.$$

Example 6.3. *Let*

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

Then

$$C_{23} = (-1)^5 detA_{23}$$
$$= (-1)^5 \begin{vmatrix} -1 & 3 \\ 5 & -2 \end{vmatrix} = (-1)(2-15) = 13.$$

Theorem 6.4. *The determinant of an $n \times n$ matrix A can be computed by cofactor expression across any row or down any columns. The determinant is the* **cofactor expression across the *i*th row***,*

$$detA = a_{i1}C_{i1} + a_{i2}C_{i2} + \ldots + a_{in}C_{in}$$

The **cofactor expression down the** *jth* **column** *is*

$$detA = a_{1j}C_{1j} + a_{2j}C_{2j} + \ldots + a_{nj}C_{nj}$$

Example 6.5. *Let*

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

(1) Find the determinant of A by cofactor expression of row 2.
(2) Find the determinant of A by cofactor expression down to 3th column.

Solution. (1) The cofactor expression of row 2 is

$$detA = a_{21}C_{21} + a_{22}C_{22} + a_{23}C_{23}$$
$$= 0C_{21} + 1 \times C_{22} + (-1)C_{23}$$

$$C_{22} = (-1)^{2+2} \begin{vmatrix} 3 & -1 \\ 5 & 3 \end{vmatrix} = 1 \times (9+5) = 14$$
$$C_{23} = (-1)^{2+3} \begin{vmatrix} 3 & 2 \\ 5 & -1 \end{vmatrix} = (-1)(-3-10) = 13.$$

Therefore, $detA = 14 + (-1)(13) = 1$.

(2) The cofactor expression of column 3 is

$$detA = a_{13}C_{13} + a_{23}C_{23} + a_{33}C_{33}$$
$$= (-1)C_{13} + (-1)C_{23} + 3C_{33}$$
$$C_{13} = (-1)^{1+3} \begin{vmatrix} 0 & 1 \\ 5 & -1 \end{vmatrix} = -5$$
$$C_{23} = (-1)^{2+3} \begin{vmatrix} 3 & 2 \\ 5 & -1 \end{vmatrix} = 13$$
$$C_{33} = (-1)^{3+3} \begin{vmatrix} 3 & 2 \\ 0 & 1 \end{vmatrix} = 3$$

So

$$\begin{aligned} detA &= (-1)(-5) + (-1)(13) + 3 \times 3 \\ &= 5 - 13 + 9 = 1 \end{aligned}$$

Definition. *A triangular matrix is a matrix with all entries below main diagonal zero.*

Example 6.6.

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

Theorem 6.7. *If A is a triangular matrix, then $detA$ is the product of entries on the main diagonal of A.*

6.2. **Week 6, Lecture 2, Oct.4, Inquiries.**

6.3. **Week 6, Lecture 3, Oct.6, Midterm 1.**

7. Week 7, Determinants, Cramer's Rule, and Inverse of a Matrix

Goals of the week:

- Properties of Determinant
- Cramer's Rule
- Inverse of a Matrix
- linear Transformation

7.1. Week 7, Lecture 1, Oct.9, Properties of Determinants.

• We study the effect of row operations on the determinant.

Theorem 7.1. *(Row operations) Let A be a square matrix.*

(1) *If a multiple of one row of A is added to another row to produce a matrix B, then*
$$detB = detA.$$
(2) *If two rows of A are interchanged to produce B, then $detB = -detA$*
(3) *If one row of A is multiplied by the scalar k to produce B, then $detB = k.detA$.*

Theorem 7.2. *If A is an $n \times n$ matrix, then $detA^T = detA$.*

Example 7.3. *Let*
$$A = \begin{bmatrix} 2 & 0 & 3 \\ 4 & -1 & 0 \\ 6 & 1 & 3 \end{bmatrix}.$$
Compute $detA$.

Solution.
$$detA = det \begin{bmatrix} 2 & 0 & 3 \\ 0 & -1 & -6 \\ 0 & 1 & 6 \end{bmatrix} = 2 \begin{vmatrix} -1 & -6 \\ 1 & 6 \end{vmatrix} = 0$$

Example 7.4. *If we have $detA = -2$ where*
$$A = \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix}$$
Compute the determinant of the following matrices.

(1)
$$\begin{bmatrix} 2a & d & -g \\ 2b & e & -h \\ 2c & f & -i \end{bmatrix}$$
(2)
$$\begin{bmatrix} a & b & c \\ -2a+d & -2b+e & -2c+f \\ g & h & i \end{bmatrix}$$

Solution.

(1)
$$det\begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix} = 1/2det\begin{bmatrix} 2a & 2b & 2c \\ d & e & f \\ g & h & i \end{bmatrix} =$$
$$-1/2\begin{bmatrix} 2a & 2b & 2c \\ d & e & f \\ -g & -h & -i \end{bmatrix} = -1/2det\begin{bmatrix} 2a & d & -g \\ 2b & e & -h \\ 2c & f & -i \end{bmatrix}$$

Therefore,
$$-2 = -1/2det\begin{bmatrix} 2a & d & -g \\ 2b & e & -h \\ 2c & f & -i \end{bmatrix}$$
and so
$$det\begin{bmatrix} 2a & d & -g \\ 2b & e & -h \\ 2c & f & -i \end{bmatrix} = 4.$$

(2) The determinant of the matrix in (2) is the same as the determinant of A since we replace the second row of A by the sum of the second row and a multiple of the first row.

Theorem 7.5. *A square matrix is invertible if and only if* $detA \neq 0$.

Theorem 7.6. *If* A *and* B *are* $n \times n$ *matrix, then*
$$det(AB) = det(BA) = det(A)det(B).$$

• Determinant of elementary matrix E
$$detE = \begin{cases} 1 & \text{if } E \text{ is a row replacement} \\ -1 & \text{if } E \text{ is an interchanger} \\ r & \text{if } E \text{ is a scalar by } r \end{cases}$$

7.2. **Week 7, Lecture 2, Oct. 11, Cramer's Rule and Inverse of a Matrix.**

• **Cramer's Rule:** It gives you the solution of the equation $Ax = b$ for an invertible matrix A.
For any invertible $n \times n$ matrix A and $b \in \mathbb{R}^n$, let $A_i(b)$ be the matrix obtained form A by replacing column i by the vector b. So if $A = [A_1|\dots|A_i|\dots|A_n]$, then

$$A_i(b) = [A_1|\dots|b|\dots|A_n].$$

Theorem 7.7 (Cramer's Rule)**.** *Let A be an invertible $n\times n$ matrix. Then the unique solution of the equation* $A\begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = b$ *has entries,*

$$x_i = \frac{detA_i(b)}{detA}, i = 1, 2, \dots, n.$$

Example 7.8. *Use the Cramer's rule to solve the system*

$$\begin{array}{rrl} x_1 & -2x_2 & = 6 \\ -2x_1 & +3x_2 & = 1 \end{array}$$

Solution. We can write the equation as

$$\begin{bmatrix} 1 & -2 \\ -2 & 3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 6 \\ 1 \end{bmatrix}$$

So by Cramer's rule we have

$$x_1 = \frac{detA_1(b)}{detA} \qquad x_2 = \frac{detA_2(b)}{detA}$$

Therefore, we should first find the determinant of A, which is $detA = 1 \times 3 - (-2) \times (-2) = -1$. Also,

$$detA_1(b) = det\begin{bmatrix} 6 & -2 \\ 1 & 3 \end{bmatrix} = 20 \qquad and \qquad detA_2(b) = det\begin{bmatrix} 1 & 6 \\ -2 & 1 \end{bmatrix} = 13$$

Therefore, $x_1 = \frac{20}{-1} = -20$ and $x_2 = \frac{13}{-1} = -13$. So $\begin{bmatrix} -20 \\ -13 \end{bmatrix}$ is the unique solution of the system.

• **A formula for A^{-1}**

Remember that the (i, j)-cofactor is

$$C_{ij} = (-1)^{i+j} detA_{ij}$$

Let A be an $n \times n$ matrix. The following matrix is denoted by $adj\ A$

$$adjA = \begin{bmatrix} C_{11} & C_{12} & \dots & C_{1n} \\ C_{21} & C_{22} & \dots & C_{2n} \\ \vdots & & & \\ C_{n1} & C_{n2} & \dots & C_{nn} \end{bmatrix}^T$$

Theorem 7.9. *(An Inverse formula) Let A be an invertible $n \times n$ matrix. Then*

$$A^{-1} = \frac{1}{detA} adjA.$$

Example 7.10. *Find the inverse of the matrix*

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

solution. What we should find are $adjA$ and the determinant of A. The cofactors are

$$C_{11} = (-1)^{1+1} \begin{vmatrix} 0 & 5 \\ -1 & 0 \end{vmatrix} = 5 \qquad C_{12} = (-1)^{1+2} \begin{vmatrix} 6 & 5 \\ -3 & 0 \end{vmatrix} = -15$$

$$C_{13} = (-1)^{1+3} \begin{vmatrix} 6 & 0 \\ -3 & -1 \end{vmatrix} = -6 \qquad C_{21} = (-1)^{2+1} \begin{vmatrix} 2 & 1 \\ -1 & 0 \end{vmatrix} = -1$$

$$C_{22} = (-1)^{2+2} \begin{vmatrix} 3 & 1 \\ -3 & 0 \end{vmatrix} = 3 \qquad C_{23} = (-1)^{2+3} \begin{vmatrix} 3 & 2 \\ -3 & -1 \end{vmatrix} = -3$$

$$C_{31} = (-1)^{3+1} \begin{vmatrix} 2 & 1 \\ 0 & 5 \end{vmatrix} = 10 \qquad C_{32} = (-1)^{3+2} \begin{vmatrix} 3 & 1 \\ 6 & 5 \end{vmatrix} = -9$$

$$C_{33} = (-1)^{3+3} \begin{vmatrix} 3 & 2 \\ 6 & 0 \end{vmatrix} = -6$$

We can compute the determinant by cofactor expression of the first row, so

$$\begin{aligned} detA &= a_{11}C_{11} + a_{12}C_{12} + a_{13}C_{13} \\ &= 3 \times 5 + 2 \times (-15) + 1 \times (-6) = -21 \end{aligned}$$

Also

$$\begin{aligned} adjA &= \begin{bmatrix} C_{11} & C_{12} & C_{13} \\ C_{21} & C_{22} & C_{23} \\ C_{31} & C_{32} & C_{33} \end{bmatrix}^T \\ &= \begin{bmatrix} 5 & -15 & -6 \\ -1 & 3 & -3 \\ 10 & -9 & -6 \end{bmatrix}^T = \begin{bmatrix} 5 & -1 & 10 \\ -15 & 3 & -9 \\ -6 & -3 & -6 \end{bmatrix} \end{aligned}$$

Therefore,

$$\begin{aligned} A^{-1} &= \frac{1}{detA} adjA = \frac{1}{-21} \begin{bmatrix} 5 & -1 & 10 \\ -15 & 3 & -9 \\ -6 & -3 & -6 \end{bmatrix} \\ &= \begin{bmatrix} 5/-21 & -1/-21 & 10/-21 \\ -15/-21 & 3/-21 & -9/-21 \\ -6/-21 & -3/-21 & -6/-21 \end{bmatrix}. \end{aligned}$$

• Parallelogram is a simple quadrilateral with two pairs of parallel sides.

Example 7.11. *draw a picture*

- The volume of the parallelepiped
  draw a picture

Theorem 7.12. *If A is a 2×2 matrix, the area of the parallelogram determined by the columns of A is $|detA|$. If A is a 3×3 matrix, the volume of the parallelepiped determined by the columns of A is $|detA|$.*

Example 7.13. *(1) Find the area of the parallelogram determined by $\begin{bmatrix} 1 \\ 2 \end{bmatrix}$ and $\begin{bmatrix} -1 \\ 0 \end{bmatrix}$.*

draw a picture

$$area = \left| det \begin{bmatrix} 1 & -1 \\ 2 & 0 \end{bmatrix} \right| = 2$$

(2) Find the volume of the parallelepiped determined by $\begin{bmatrix} 3 \\ 6 \\ -3 \end{bmatrix}, \begin{bmatrix} 2 \\ 0 \\ -1 \end{bmatrix}, \begin{bmatrix} 1 \\ 5 \\ 0 \end{bmatrix}$.

draw a picture

7.3. Week 7, Lecture 3, Oct. 13, Vector Space.

Definition: A **vector space** consist of the following:

(1) a Field of scalars; In this course we only work with $\mathbb{R}$, the field of real numbers;
(2) a set V of objects, called vectors;
(3) **addition:** a rule(or operation), called vector addition, which associate with each pair of vectors u and v in V a vector $u+v$ in V, called the sum of u and v, in such a way that
  (a) $u+v=v+u$; (commutativity of addition)
  (b) u+ (v+ w)= (u+ v)+ w (associativity of addition) w is also in V
  (c) there is a zero vector 0 in V such that $u+0=u$.
  (d) for each vector u in V, there is a vector $-u$ in V such that $u+(-u)=0$.
(4) **scalar multiplication:** a rule (or operation) called scalar multiplication which associate each scalar c in $\mathbb{R}$ and vector $u \in V$ a vector cu in V, called product of c and u, in such a way that
  (a) $1u=u$ for every $u \in V$;
  (b) $c(du)=(cd)u$;
  (c) $(c+d)u=cu+du$;
  (d) $c(u+v)=cu+cv$.

Example 7.14. *Show that $M_{2\times 3}(\mathbb{R})$, the set of all 2×3 matrices is a vector space.*

Solution.

(1) $\mathbb{R}$ is the field of scalars
(2) The set of objects is the set of 2×3 matrices
(3) addition operator is the sum (or addition of matrices): Let $A=\begin{bmatrix} a & b & c \\ d & e & f \end{bmatrix}$, $B=\begin{bmatrix} g & h & i \\ k & l & m \end{bmatrix}$, $C=\begin{bmatrix} w & y & z \\ r & p & q \end{bmatrix}$ be three arbitrary matrix in $M_{2\times 3}(\mathbb{R})$. Then
$$\begin{bmatrix} a & b & c \\ d & e & f \end{bmatrix}+\begin{bmatrix} g & h & i \\ k & l & m \end{bmatrix}=\begin{bmatrix} a+g & b+h & c+i \\ d+k & e+l & f+m \end{bmatrix}$$
is in $M_{2\times 3}(\mathbb{R})$.
  (a)
$$\begin{bmatrix} a & b & c \\ d & e & f \end{bmatrix}+\begin{bmatrix} g & h & i \\ k & l & m \end{bmatrix}=$$
$$\begin{bmatrix} a+g & b+h & c+i \\ d+k & e+l & f+m \end{bmatrix}=\begin{bmatrix} g+1 & h+b & i+c \\ k+d & l+e & m+f \end{bmatrix}$$
$$=\begin{bmatrix} g & h & i \\ k & l & m \end{bmatrix}+\begin{bmatrix} a & b & c \\ d & e & f \end{bmatrix}$$
  (b) For any $A,B,C \in M_{2\times 3}(\mathbb{R})$,
$$A+(B+C)=(A+B)+C$$

(c) There is the zero matrix $0 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$ such that $A + 0 = A$

(d) For each matrix $\begin{bmatrix} a & b & c \\ d & e & f \end{bmatrix} \in M_{2\times 3}(\mathbb{R})$ there is a matrix $\begin{bmatrix} -a & -b & -c \\ -d & -e & -f \end{bmatrix}$ such that

$$\begin{bmatrix} a & b & c \\ d & e & f \end{bmatrix} + \begin{bmatrix} -a & -b & -c \\ -d & -e & -f \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

(4) (Scalar multiplication) For any $c \in \mathbb{R}$ and $A \in M_{2\times 3}(\mathbb{R})$, we have that $cA \in M_{2\times 3}(\mathbb{R})$

(a) The scalar $1.A = A$ for any $A \in M_{2\times 3}(\mathbb{R})$ since if we multiply the entries of A by 1 we have A.

(b) Let c and d be in $\mathbb{R}$, then $c(dA) = (cd)A$.

(c) $(c + d)A = cA + dA$

(d) $c(A + B) = cA + cB$

8. Week 8, Vector Space

Goals of the week:

- More examples for vector space, and subspace
- Review of the null space and column space
- Linear Transformation
- Linear dependent sets and bases

8.1. **Week 8, Lecture 1, Oct.16, some example for vector space.**

Definition. Polynomial: *The set of polynomials of degree at most n, is denoted by $\mathbb{P}_n$, consists of all polynomials of the form*

$$p(t) = a_0 + a_1 t + \ldots + a_m t^m$$

where $m \leq n$.

Example 8.1.

$$1 + 2t + 3t^2 \in \mathbb{P}_3$$
$$1 - 3t + 4t^4 \in \mathbb{P}_4$$

Example 8.2. *Show that the set of all polynomials of degree n, is a vector space.*

(1) Field of scalars is $\mathbb{R}$
(2) objects are polynomials
(3) (addition) is the sum of polynomials.
Let $k \leq m \leq n$, and $f(t) = a_0 + a_1 t + \ldots + a_m t^m \in \mathbb{P}_n$, $g(t) = b_0 + b_1 t + \ldots + b_k t^k \in \mathbb{P}_n$, then

$$f(t) + g(t) = (a_0 + b_0) + (a_1 + b_1)t + \ldots + (a_k + b_k)t^k + \ldots + (a_m)t^m \in \mathbb{P}_n.$$

(a) $f(t) + s(t) = s(t) + f(t)$

$$f(t) + s(t) = (a_0 + b_0) + (a_1 + b_1)t + \ldots + (a_k + b_k)t^k + \ldots + (a_m)t^m$$

and

$$s(t) + f(t) = (b_0 + a_0) + (b_1 + a_1)t + \ldots + (b_k + a_k)t^k + \ldots + (a_m)t^m$$

so

$$f(t) + s(t) = s(t) + f(t).$$

(b) $f(t) + (s(t) + g(t)) = (f(t) + s(t)) + g(t)$
(c) The zero polynomial is $0 + 0t + \ldots + 0t^n$, and so

$$f(t) + 0 = a_0 + a_1 t + \ldots + a_n t^n +$$
$$0 + 0t + \ldots + 0t^n =$$
$$a_0 + a_1 t + \ldots + a_n t^n = f(t).$$

(d) For each $f(t) = a_0 + a_1 t + \ldots + a_n t^n$,

$$-f(t) = -a_0 - a_1 t + \ldots - a_n t^n$$

and so $f(t) + (-f(t)) = 0$

(4) we have the scalar multiplication as follows
$$c.f(t) = ca_0 + \ldots + ca_n t^n$$
It is clear that $c.f(t)$ is a polynomial in P_n.
(a) $1.f(t) = f(t)$
(b) $c(df(t)) = (cd)f(t)$
(c) $(c + d)f(t) = cf(t) + df(t)$
(d) $c(f(t) + g(t)) = cf(t) + cg(t)$.

8.2. **Week 8, Lecture 2, Oct.18, Subspace.**

Example 8.3. *Let K be the set of all functions on the interval $[a,b]$. Then K is a vector space over $\mathbb{R}$.*

1. *(sum) Let f and g be in K. Then $(f+g)(c) = f(c)+g(c)$ for every $c \in [a,b]$ (Example: $f = sinx + 1$ and $g = 2x + x^2$, then $(f+g)(c) = f(c) + g(c) = sin\ c + 1 + 2c + c^2$.*
2. *(zero) The zero element of K is zero function, i.e., $\mathbf{0}$: $[a,b] \to \mathbb{R}$ such that $\mathbf{0}(c)=0$.*
3. *The negative of any function f in K is $(-f)$ where $(-f)(c) = -f(c)$.*
4. *(scalar multiplication) for every $f \in K$ and scalar d, (df) is a function such that $(df)(c) = df(c)$.*

Definition. *A* **subspace** *of a vector space V is a non-empty subset H of V such that*

1. *the zero vector is in H.*
2. *for any two vectors $v, u \in H$, $u + v \in H$.*
3. *for any vector $v \in V$ and scalar c, $cv \in H$.*

• Every vector space V has two trivial subspace V and $\{0\}$.

Example 8.4. *Let V be the set of all continuous functions on the interval $[a,b]$. By drawing the graph we showed in class what continuous function means. Then V is a subspace of K, the set of all functions.*

1. *It is clear that 0 is a continuous function.*
2. *If f and g are continuous functions, then $f+g$ is a continuous function.*
3. *If c is a scalar, the cf is continuous for any continuous function f.*

Example 8.5. *Let W be the set of all differentiable functions on the interval $[a,b]$. By drawing the graph we showed what differentiable function means. With the same proof, W is a subspace of K, the set of all functions.*

Example 8.6. *Show that $H = \left\{ \begin{bmatrix} a \\ b \\ 0 \end{bmatrix} : a, b \in \mathbb{R} \right\}$ is a subspace of $\mathbb{R}^3$.*

Solution. *We only need to show the three properties.*

1. *The zero vector is in H, since when $a = b = 0$, then $\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$ is in H.*
2. *Let $u = \begin{bmatrix} a_1 \\ b_1 \\ 0 \end{bmatrix}$ and $\begin{bmatrix} a_2 \\ b_2 \\ 0 \end{bmatrix} \in H$, then $u + v = \begin{bmatrix} a_1 + a_2 \\ b_1 + b_2 \\ 0 \end{bmatrix}$ is in H.*
3. *Let c be a scalar and $u = \begin{bmatrix} a_1 \\ b_1 \\ 0 \end{bmatrix} \in H$, then $cu = \begin{bmatrix} ca_1 \\ cb_1 \\ 0 \end{bmatrix}$ is in H.*

Lemma 8.7. *Let V be a vector space. Then for any v_1 and v_2 in V, $Span\{v_1, v_2\} = \{c_1v_1 + c_2v_2 : c_1, c_2 \in \mathbb{R}\}$ is a subspace of V.*

• *A geometrical view of the span of two vectors in $\mathbb{R}^3$ is shown the following video.* *CLICK*

Solution: Note that

$$span\{v_1, v_2\} = \{c_1v_1 + c_2v_2 : c_1, c_2 \in \mathbb{R}\}$$

is a subset of $\mathbb{R}^n$, we only need to check the three properties.

(1) Let $c_1 = 0$, $c_2 = 0$. then $0v_1 + 0v_2 = 0 \in span\{v_1, v_2\}$

(2) Let $u, v \in span\{v_1, v_2\}$. Then there are scalars, c_1 and c_2, d_1 and d_2 such that $v = c_1v_1 + c_2v_2$ and $u = d_1v_1 + d_2v_2$. Therefore,

$$v + u = c_1v_1 + c_2v_2 + d_1v_1 + d_2v_2 = (c_1 + d_1)v_1 + (c_2 + d_2)v_2$$

is in $span\{v_1, v_2\}$.

(3) Let c be an scalar , and $v \in span\{v_1, v_2\}$. Then there are $c_1, c_2 \in \mathbb{R}$ such that $v = c_1v_1 + c_2v_2$. so,

$$c(c_1v_1 + c_2v_2) = (cc_1)v_1 + (cc_2)v_2 \in span\{v_1, v_2\}.$$

Therefore $span\{v_1, v_2\}$ is a subspace of $\mathbb{R}^n$.

Theorem 8.8. *For any set of vectors $v_1, \dots, v_p$ in a vector space V, $Span\{v_1, \dots, v_p\}$ is a subspace of V.*

Example 8.9. *Let $H = \{(a, b - a, 3a - b) : a, b \in \mathbb{R}\}$. Show that H is a subspace of $\mathbb{R}^3$ (To solve this example see Example 11 page 196 of the textbook).*

Example 8.10. *Read Example 12 page 197 of the textbook.*

Theorem 8.11. *Any subspace of a vector space is a vector space.*

YOU ARE RESPONSIBLE TO REVIEW WEEK 5 BEFORE NEXT LECTURE

8.3. Week 8, Lecture 3, Oct.20, Linear Transformation, Linear dependent sets and bases.

A **linear transformation** T from a vector space V into a vector space W is a rule that assigns to each vector $x \in V$ a unique vector $T(x)$ in W, such that

(1) $T(u+v)$=$T(u)+T(v)$ for all $u, v \in V$
(2) $T(cu)$= $cT(u)$ for all $c \in \mathbb{R}$ and $u \in V$.

Moreover, the **range** of T is denoted by $range(T) = \{T(x) : x \in V\}$, and **kernel** of T is denoted by $ker(T) = \{x \in V : T(x) = 0\}$.

Example 8.12. *Let V be the vector space of continuous functions, and W be the set of all differentiable functions. Define*

$$D : W \longrightarrow V \quad by \quad D(f) = f' \quad the \quad derivation \quad of \quad f.$$

Show that D is a linear transformation. Also find $ker(T)$.

Solution. Let f and g be in W. Then

$$D(f+g) = (f+g)' = f' + g'$$
$$D(f) = f'$$
$$D(g) = g'$$

Therefore, $D(f+g) = f' + g' = D(f) + D(g)$.

Also $D(cf) = (cf)' = cf' = cD(f)$ for every $c \in \mathbb{R}$. Therefore D is a linear transformation.

Note that $ker(T) = \{f \in W : T(f) = 0\} = \{f \in W : f^{'} = 0\} = \{f \in W : f = c \text{ where c is a scalar}\}$.

Definition. • *An index set of vectors $\{v_1, \ldots, v_p\}$ of a vector space V are* **linearly independent** *if the equation*

$$c_1v_1 + \ldots + c_pv_p = 0$$

has only one trivial solution $c_1 = \ldots = c_p = 0$.
• *If there are some weights $c_1, \ldots, c_p$ not all zero such that*

$$c_1v_1 + \ldots + c_pv_p = 0 \qquad (1)$$

Then equation (1) is said to be a **linear dependence relation among** $v_1, \ldots, v_p$.

Theorem 8.13. *An index set of vectors $\{v_1, \ldots, v_p\}$ of two or more vectors, with $v_1 \neq 0$, is linearly dependent if and only if some v_j(with $j > 1$) is a linear combination of the preceding vectors $v_1, \ldots, v_{j-1}$.*

Example 8.14. *(See also Example 1 and 2 page 211 of the textbook) Let*

$$p_1(t) = 3t^2 + 1 \quad p_2(t) = t^2 \quad p_3(t) = 2.$$

Then $\{p_1, p_2, p_3\}$ is linearly dependent because

$$p_3(t) = 2 = 2(3t^2 + 1) + (-6)t^2.$$

Definition. *Let H be a subspace of a vector space V. Then an indexed set $\{v_1, \dots, v_p\}$ is a basis for H, if*

(1) *$\{v_1, \dots, v_p\}$ is linearly independent.*
(2) *$H = span\{v_1, \dots, v_p\}$.*

• *See Examples 3, 4, 5 page 211 of the textbook.*

Definition. *Let $f(t) = a_0 + a_1t + \dots + a_nt^n = 0$ be a non-zero polynomial. A root for f is a number c such that $f(c) = a_0 + a_1c + \dots + a_nc^n = 0$; for example $f(t) = t^2 - 1$ has roots 1 and -1.*

Theorem 8.15. *Every polynomial in $\mathbb{P}_n$ has at most n roots.*

Example 8.16. *$S = \{1, t, t^2, \dots, t^n\}$ is a basis for $\mathbb{P}_n$.*

Solution. Any polynomial is of the form

$$f(t) = a_0 + a_1t + \dots + a_mt^m$$

where $m \leq n$ so $f(t) \in span\{1, t, \dots, t^n\}$.

Now, we should show that $\{1, t, \dots, t^n\}$ are linearly independent.

Let

$$c_0 + c_1t + \dots + c_nt^n = 0,$$

then it means the polynomial $c_0 + c_1t + \dots + c_nt^n$ has infinitely many roots which is not possible because every polynomial of degree at most n has at most n roots.

• See Example 7 pages 212 of the textbook.

Theorem 8.17. *Let $S = \{v_1, \dots, v_p\}$ be a set in V, and Let $H = span\{v_1, \dots, v_p\}$.*

(1) *If $v_i \in \{v_1, \dots, v_p\}$ is a linear combination of $\{v_1, \dots, v_{i-1}, v_{i+1}, \dots, v_p\}$ then*
$$H = span\{v_1, \dots, v_{i-1}, v_{i+1}, \dots, v_p\}$$
.
(2) *If $H \neq 0$ some subset of S is a basis for H.*

9. Week 9, More on vector spaces

Goals of the week:

- Coordinate system
- Isomorphism
- Review of Rank
- Change of basis

9.1. **Week 9, Lecture 1, Oct.23, coordinate basis.**

Theorem 9.1. *Let $\mathcal{B}$ be a basis for a vector space V. Then for each x in V, there exists unique set of scalars $\{c_1, \ldots, c_n\}$ such that*

$$x = c_1 b_1 + \ldots + c_n b_n.$$

Proof. Since $\mathcal{B} = \{b_1, \ldots, b_n\}$ is a basis there are scalars $c_1, \ldots, c_n$ such that $x = c_1 b_1 + \ldots + c_n b_n$. Suppose also x has the representation

$$x = d_1 b_1 + \ldots + d_n b_n.$$

Then

$$0 = x - x = (c_1 - d_1) b_1 + \ldots + (c_n - d_n) b_n.$$

Note that $\{b_1, \ldots, b_n\}$ is linearly independent, so

$$c_1 - d_1 = 0, \ldots, c_n - d_n = 0 \Rightarrow c_1 = d_1, \ldots, c_n = d_n.$$

$\square$

Definition. *Suppose $\mathcal{B} = \{b_1, \ldots, b_n\}$ is a basis for V and x is in V. Let*

$$x = c_1 b_1 + \ldots + c_n b_n.$$

The **coordinate vector for x relative to the basis $\mathcal{B}$** *is*

$$[x]_{\mathcal{B}} = \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix}.$$

Note that $[x]_{\mathcal{B}} \in \mathbb{R}^n$ for any basis $\mathcal{B}$ of V.

- **Coordinates in $\mathbb{R}^n$**

Example 9.2. *Let $\mathcal{B} = \{b_1, b_2\}$ be a basis for $\mathbb{R}^2$ where $b_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ and $b_2 = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$. If $[x]_{\mathcal{B}} = \begin{bmatrix} 3 \\ 4 \end{bmatrix}$. Find x.*

Solution. $[x]_{\mathcal{B}} = 3 \begin{bmatrix} 1 \\ 0 \end{bmatrix} + 4 \begin{bmatrix} 2 \\ 1 \end{bmatrix} = \begin{bmatrix} 11 \\ 4 \end{bmatrix}$.

Example 9.3. *Let $\mathcal{B}$ be the standard basis for $\mathbb{R}^2$, i.e., $\mathcal{B} = \{e_1, e_2\}$, where $e_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ and $e_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$. Let $x = \begin{bmatrix} 3 \\ 1 \end{bmatrix}$ what is $[x]_{\mathcal{B}}$?*

Solution. Since $\begin{bmatrix} 3 \\ 1 \end{bmatrix} = 3\begin{bmatrix} 1 \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \end{bmatrix} = 3e_1 + e_2$, we have $[x]_{\mathcal{B}} = \begin{bmatrix} 3 \\ 1 \end{bmatrix}$.

• If $\mathcal{B}$ is the standard basis for $\mathbb{R}^n$, then $[x]_{\mathcal{B}} = x$.

Example 9.4. *Let* $b_1 = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$, $b_2 = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$, *and* $x = \begin{bmatrix} 4 \\ 5 \end{bmatrix}$, *and* $\mathcal{B} = \{b_1, b_2\}$. *find the coordinate vector* $[x]_{\mathcal{B}}$.

Solution. We have that $[x]_{\mathcal{B}} = \begin{bmatrix} c_1 \\ c_2 \end{bmatrix}$ where

$$c_1\begin{bmatrix} 2 \\ 1 \end{bmatrix} + c_2\begin{bmatrix} -1 \\ 1 \end{bmatrix} = \begin{bmatrix} 4 \\ 5 \end{bmatrix},$$

i.e.,

$$\begin{bmatrix} 2c_1 - c_2 \\ c_1 + c_2 \end{bmatrix} = \begin{bmatrix} 4 \\ 5 \end{bmatrix},$$

we can write it as

$$\begin{bmatrix} 2 & -1 \\ 1 & 1 \end{bmatrix}\begin{bmatrix} c_1 \\ c_2 \end{bmatrix} = \begin{bmatrix} 4 \\ 5 \end{bmatrix}.$$

Then you can solve this equation and find $c_1 = 3$ and $c_2 = 2$.

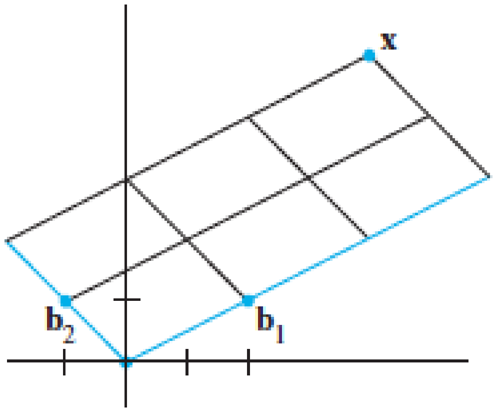


FIGURE 4
The $\mathcal{B}$-coordinate vector of **x** is $(3, 2)$.

□

In the above example the matrix

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

has a especial name.

Definition. *Let $\mathcal{B} = \{b_1, \ldots, b_n\}$ be a basis for $\mathbb{R}^n$. The matrix*

$$P_{\mathcal{B}} = [b_1| \ldots |b_n]$$

is called the **change-of-coordinates matrix from $\mathcal{B}$ to the standard basis of** $\mathbb{R}^n$. *Also when $x = c_1 b_1 + \ldots + c_n b_n$, we have*

$$x = P_{\mathcal{B}}[x]_{\mathcal{B}} = P_{\mathcal{B}} \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix}.$$

Remark.

(1) The matrix $P_{\mathcal{B}}$ is an $n \times n$ matrix.
(2) The columns of $P_{\mathcal{B}}$ form a basis for $\mathbb{R}^n$, so $P_{\mathcal{B}}$ is invertible.
(3) We can also write $P_{\mathcal{B}}^{-1} x = [x]_{\mathcal{B}}$.

• The coordinate mapping

Theorem 9.5. *Let $\mathcal{B} = \{b_1, \ldots, b_n\}$ be a basis for a vector space V. Then the coordinate mapping*

$$\begin{array}{rccl} T: & V & \to & \mathbb{R}^n \\ & x & \mapsto & [x]_{\mathcal{B}} \end{array}$$

is a one-to-one linear transformation form V onto $\mathbb{R}^n$.

Proof. Let $u = c_1 b_1 + \ldots + c_n b_n$ and $w = d_1 b_1 + \ldots + d_n b_n$. Then

$$u + w = (c_1 + d_1) b_1 + \ldots + (c_n + d_n) b_n.$$

It follows that

$$[u+w]_{\mathcal{B}} = \begin{bmatrix} c_1 + d_1 \\ \vdots \\ c_d + d_d \end{bmatrix} = \begin{bmatrix} c_1 \\ \vdots \\ c_d \end{bmatrix} + \begin{bmatrix} d_1 \\ \vdots \\ d_n \end{bmatrix} = [u]_{\mathcal{B}} + [w]_{\mathcal{B}}.$$

Now let $r \in \mathbb{R}$,

$$ru = r(c_1 b_1 + \ldots + c_n d_n) = (rc_1) b_1 + \ldots + (rc_n) d_n.$$

Therefore,

$$[ru]_{\mathcal{B}} = \begin{bmatrix} rc_1 \\ \vdots \\ rc_n \end{bmatrix} = r \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix} = r[u]_{\mathcal{B}}.$$

□

Theorem 9.6. *Let V and W be vector spaces, and $T : V \to W$ be a linear transformation. Then*

(1) *T is one-to-one if ker $(T) = \{v \in V : T(v) = 0\} = \{0\}$.*
(2) *T is onto if range$(T) = \{T(v) : v \in V\} = W$.*

Definition. *A linear transformation T from a vector space V to a vector space W is an* **isomorphism** *if T is one-to-one and onto. Moreover, we say V and W are* **isomorphic**.

Theorem 9.7. *Let $\mathcal{B} = \{b_1, \ldots, b_n\}$ be a basis for a vector space V. Then the coordinate mapping*

$$\begin{array}{rccc} T: & V & \to & \mathbb{R}^n \\ & x & \mapsto & [x]_{\mathcal{B}} \end{array}$$

is a one-to-one linear transformation form V onto $\mathbb{R}^n$.

Solution. Previously we showed that T is a linear transformation. Now, we will show that it is one-to-one and onto.

one-to-one: $ker(T) = \{x \in V : [x]_{\mathcal{B}} = 0\}$. Note that if $[x]_{\mathcal{B}} = \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix}$, then $x = 0b_1 + \ldots + 0b_n = 0$. Therefore, $ker(T) = 0$ and so T is one-to-one.

onto: For any $y = \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix} \in \mathbb{R}^n$, there is a vector $x = y_1 b_1 + \ldots + y_n b_n \in V$ such that $[x]_{\mathcal{B}} = y$.

9.2. Week 9, Lecture 2, Oct.25, Linearly independent sets, basis, and dimension.

Definition. *Let $f(t) = a_0 + a_1t + \ldots + a_nt^n = 0$ be a non-zero polynomial. A root for f is a number c such that $f(c) = a_0 + a_1c + \ldots + a_nc^n = 0$; for example $f(t) = t^2 - 1$ has roots 1 and -1.*

Theorem 9.8. *Every polynomial in $\mathbb{P}_n$ has at most n roots.*

Example 9.9. *$S = \{1, t, t^2, \ldots, t^n\}$ is a basis for $\mathbb{P}_n$.*

Solution. Any polynomial is of the form

$$f(t) = a_0 + a_1t + \ldots + a_mt^m$$

where $m \leq n$ so $f(t) \in span\{1, t, \ldots, t^n\}$.

Now, we should show that $\{1, t, \ldots, t^n\}$ are linearly independent.

Let

$$c_0 + c_1t + \ldots + c_nt^n = 0,$$

then it means the polynomial $c_0 + c_1t + \ldots + c_nt^n$ has infinitely many roots which is not possible because every polynomial of degree at most n has at most n roots.

Example 9.10. *Let $B = \{1, t, t^2, t^3\}$ be the standard basis for $\mathbb{P}_3$. Show that $\mathbb{P}_3$ is isomorphic to $\mathbb{R}^4$.*

Solution. By the previous theorem we have

$$T : \mathbb{P}_3 \longrightarrow \mathbb{R}^4$$

$$p = a_0 + a_1t + a_2t^2 + a_3t^3 \mapsto [p]_B = \begin{bmatrix} a_0 \\ a_1 \\ a_2 \\ a_3 \end{bmatrix}$$

is a isomorphism.

Example 9.11. *Let*

$$v_1 = \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix} \quad v_2 = \begin{bmatrix} -1 \\ 0 \\ -3 \end{bmatrix} \quad x = \begin{bmatrix} 1 \\ 4 \\ -1 \end{bmatrix}$$

and $B = \{v_1v_2\}$. Then $\mathcal{B}$ is a basis for $H = span\{v_1, v_2\}$. Determine if x is in H. Find $[x]_{\mathcal{B}}$.

Solution. If the following system is consistent

$$c_1 \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix} + c_2 \begin{bmatrix} -1 \\ 0 \\ -3 \end{bmatrix} = \begin{bmatrix} 1 \\ 4 \\ -1 \end{bmatrix}$$

Then $\begin{bmatrix} 1 \\ 4 \\ -1 \end{bmatrix}$ is in $span\{v_1, v_2\}$. The augmented matrix is

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

An echelon form is

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

so the system is consistent and if you solve it, you have $c_1 = 2$ and $c_2 = 1$. Therefore $[x]_{\mathcal{B}} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$.

9.3. **Week, Lecture 3, Oct. 27, The dimension of a vector space.**

Theorem 9.12. *Let $T : V \longrightarrow W$ be an isomorphism.*

(1) $v_1, \dots, v_n$ *are linearly independent (dependent) in* V *if and only if* $T(v_1), \dots, T(v_n)$ *are linearly independent (dependent) in* W*.*

(2) *A vector* x *is in span*$\{v_1, \dots, v_n\}$ *if and only if* $T(x)$ *is in span*$\{T(v_1), \dots, T(v_n)\}$*.*

Example 9.13. (1) *Verify that the polynomials* $1 + 2t^2$*,* $4 + t + 5t^2$*, and* $3 + 2t$ *are linearly independent.*

(2) *Is* $g(t) = t - 3t^2$ *in span*$\{1 + 2t^2, 4 + t + 5t^2, 3 + 2t\}$*?*

Solution. (1) Let $\mathcal{B} = \{1, t, t^2, t^3\}$ be the standard basis for $\mathbb{P}_3$. We have by Theorem 9.7 $T : \mathbb{P}_3 \longrightarrow \mathbb{R}^4$ where

$$p \mapsto [p]_{\mathcal{B}}$$

is an isomorphism. Therefore by theorem above $1 + 2t^2$, $4 + t + 5t^2$ and $3 + 2t$ are linearly independent if and only if

$$\left[1 + 2t^2\right]_{\mathcal{B}}, \left[4 + t + 5t^2\right]_{\mathcal{B}}, [3 + 2t]_{\mathcal{B}}$$

are linearly independent. We have

$$\left[1 + 2t^2\right]_{\mathcal{B}} = \begin{bmatrix} 1 \\ 0 \\ 2 \\ 0 \end{bmatrix}, \left[4 + t + 5t^2\right]_{\mathcal{B}} = \begin{bmatrix} 4 \\ 1 \\ 5 \\ 0 \end{bmatrix}, [3 + 2t]_{\mathcal{B}} = \begin{bmatrix} 3 \\ 2 \\ 0 \\ 0 \end{bmatrix}$$

Therefore, we only need to show that

$$\left\{ \begin{bmatrix} 1 \\ 0 \\ 2 \\ 0 \end{bmatrix}, \begin{bmatrix} 4 \\ 1 \\ 5 \\ 0 \end{bmatrix}, \begin{bmatrix} 3 \\ 2 \\ 0 \\ 0 \end{bmatrix} \right\}$$

are linearly independent. (Do it as an Exercise).

(2) By the above theorem we only need to show that

$$[g(t)]_{\mathcal{B}} \in span \left\{ \begin{bmatrix} 1 \\ 0 \\ 2 \\ 0 \end{bmatrix}, \begin{bmatrix} 4 \\ 1 \\ 5 \\ 0 \end{bmatrix}, \begin{bmatrix} 3 \\ 2 \\ 0 \\ 0 \end{bmatrix} \right\},$$

i.e.,

$$\begin{bmatrix} 0 \\ 1 \\ -3 \\ 0 \end{bmatrix} \in span \left\{ \begin{bmatrix} 1 \\ 0 \\ 2 \\ 0 \end{bmatrix}, \begin{bmatrix} 4 \\ 1 \\ 5 \\ 0 \end{bmatrix}, \begin{bmatrix} 3 \\ 2 \\ 0 \\ 0 \end{bmatrix} \right\}$$

$\square$

Theorem 9.14. *If a vector space* V *has a basis* $\mathcal{B} = \{b_1, \dots, b_n\}$ *then any set containing more than* n *vectors must be linearly dependent.*

Theorem 9.15. *If V is a vector space and V has a basis of n vectors, then every basis of V must consist of exactly n vectors.*

Definition.
(1) *A vector space is said to be* **finite-dimensional** *if it is spanned by a finite set of vectors in V*
(2) **Dimension** *of V, dim V, is the number of vectors in a basis of V. Also dimension of zero space $\{0\}$ is 0.*
(3) *If V is not spanned by a finite set, then V is said to be infinite-dimensional.*

Example 9.16. *Find dimension of the subspace*

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

Solution. We have

$$\begin{bmatrix} a-3b+c \\ 2a+2d \\ b-3c-d \\ 2d-b \end{bmatrix} = a \begin{bmatrix} 1 \\ 2 \\ 0 \\ 0 \end{bmatrix} + b \begin{bmatrix} -3 \\ 0 \\ 1 \\ -1 \end{bmatrix} + c \begin{bmatrix} 1 \\ 0 \\ -3 \\ 0 \end{bmatrix} + d \begin{bmatrix} 0 \\ 2 \\ -1 \\ 2 \end{bmatrix}$$

Now we have

$$H = span \left\{ \begin{bmatrix} 1 \\ 2 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} -3 \\ 0 \\ 1 \\ -1 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ -3 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 2 \\ -1 \\ 2 \end{bmatrix} \right\}$$

Now, we want to find a basis for H, we had a process for finding the basis.(Do it as an exercise.)

Theorem 9.17. *Let H be a subspace of a finite dimensional vector space V. Any linearly independent set in H can be expanded to a basis for H. Also*

$$dim\ H \leq dim\ V$$

Theorem 9.18. *(***The Basis Theorem***) Let V be a p-dimensional vector space $p \geq 1$.*
(1) *Any linearly independent set of exactly p elements in V is automatically a basis for V.*
(2) *Any set of exactly p elements that spans V is automatically a basis for V.*

Remember: The dimension of $Nul\ A$ is the number of free variables in the equation $Ax = 0$, and the dimension of $Col\ A$ is the number of pivot columns in A, and the pivot columns of A gives a basis for column space of A.

10. Week 10, change of basis, eigenvalue and eigenvector, characteristic equations

10.1. **Week 10, Lecture 1, change of basis.** Goals:

(1) Change of Basis
(2) Eigenvalues and eigenvectors
(3) Characteristic equations

Example 10.1. *Let*

$$b_1 = \begin{bmatrix} 2 \\ 0 \end{bmatrix}, b_2 = \begin{bmatrix} -1 \\ 1 \end{bmatrix}, c_1 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, c_2 = \begin{bmatrix} 2 \\ 1 \end{bmatrix}.$$

Then $\mathcal{B} = \{b_1, b_2\}$ *and* $\mathcal{C} = \{c_1, c_2\}$ *are two basis for* $\mathbb{R}^2$. *Let* $x = \begin{bmatrix} 0 \\ 2 \end{bmatrix}$. *Then*

$$x = \begin{bmatrix} 0 \\ 2 \end{bmatrix} = \begin{bmatrix} 2 \\ 0 \end{bmatrix} + 2\begin{bmatrix} -1 \\ 1 \end{bmatrix} = b_1 + 2b_2$$

Therefore, $[x]_{\mathcal{B}} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$. *Also*

$$x = \begin{bmatrix} 0 \\ 2 \end{bmatrix} = 2\begin{bmatrix} 0 \\ 1 \end{bmatrix} + 0\begin{bmatrix} 2 \\ 1 \end{bmatrix} = 2c_1 + 0c_2$$

so $[x]_{\mathcal{C}} = \begin{bmatrix} 2 \\ 0 \end{bmatrix}$. *Then there is a matrix* $\underset{\mathcal{C}\leftarrow\mathcal{B}}{P}$ *such that*

$$[x]_{\mathcal{C}} = \underset{\mathcal{C}\leftarrow\mathcal{B}}{P}[x]_{\mathcal{B}} = [[b_1]_{\mathcal{C}} \quad [b_2]_{\mathcal{C}}][x]_{\mathcal{B}}.$$

Since

$$b_1 = \begin{bmatrix} 2 \\ 0 \end{bmatrix} = (-1)\begin{bmatrix} 0 \\ 1 \end{bmatrix} + \begin{bmatrix} 2 \\ 1 \end{bmatrix} = (-1)c_1 + c_2$$

we have

$$[b_1]_{\mathcal{C}} = \begin{bmatrix} -1 \\ 1 \end{bmatrix}.$$

Also

$$b_2 = \begin{bmatrix} -1 \\ 1 \end{bmatrix} = 3/2\begin{bmatrix} 0 \\ 1 \end{bmatrix} + (-1/2)\begin{bmatrix} 2 \\ 1 \end{bmatrix} = 3/2c_1 - 1/2c_2$$

Therefore,

$$[x]_{\mathcal{C}} = \begin{bmatrix} -1 & 3/2 \\ 1 & -1/2 \end{bmatrix}\begin{bmatrix} 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 2 \\ 0 \end{bmatrix}$$

Theorem 10.2. *Let* $\mathcal{B} = \{b_1, \dots, b_n\}$ *and* $\mathcal{C} = \{c_1, \dots, c_n\}$ *be bases of a vector space* V. *Then there is a unique matrix* $\underset{\mathcal{C}\leftarrow\mathcal{B}}{P}$ *such that*

$$[x]_{\mathcal{C}} = \underset{\mathcal{C}\leftarrow\mathcal{B}}{P}[x]_{\mathcal{B}}$$

The columns of $\underset{\mathcal{C}\leftarrow\mathcal{B}}{P}$ *are the* $\mathcal{C}$*-coordinate vectors of the vectors in the basis* $\mathcal{B}$*. That is,*

$$\underset{\mathcal{C}\leftarrow\mathcal{B}}{P} = [[b_1]_{\mathcal{C}} \quad [b_2]_{\mathcal{C}} \quad \ldots \quad [b_n]_{\mathcal{C}}].$$

Definition. *The matrix* $\underset{\mathcal{C}\leftarrow\mathcal{B}}{P}$ *in the above theorem is called* **change-of-coordinates matrix from** $\mathcal{B}$ **to** $\mathcal{C}$*.*

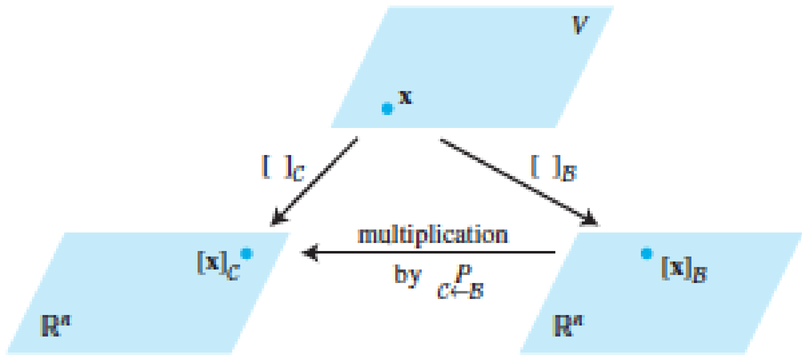


FIGURE 2 Two coordinate systems for V.

Remark. We have

$$[x]_{\mathcal{C}} = \underset{\mathcal{C}\leftarrow\mathcal{B}}{P}[x]_{\mathcal{B}}$$

so

$$\underset{\mathcal{C}\leftarrow\mathcal{B}}{P}^{-1}[x]_{\mathcal{C}} = [x]_{\mathcal{B}}$$

Therefore,

$$\underset{\mathcal{B}\leftarrow\mathcal{C}}{P} = (\underset{\mathcal{C}\leftarrow\mathcal{B}}{P})^{-1}$$

• Change of Basis in $\mathbb{R}^n$

Remark. Let $\mathcal{B} = \{b_1, \ldots, b_n\}$ a basis for $\mathbb{R}^n$. Let $\mathcal{E} = \{e_1, \ldots, e_n\}$ be the standard basis for $\mathbb{R}^n$. Then $P_{\mathcal{B}} = [b_1|\ldots|b_n]$ is the same as $\underset{\mathcal{E}\leftarrow\mathcal{B}}{P}$.

• Let $\mathcal{B} = \{b_1, \ldots, b_n\}$ and $\mathcal{C} = \{c_1, \ldots, c_n\}$ be bases for $\mathbb{R}^n$. Then by row operation we can reduce the matrix

$$[c_1 \quad \ldots \quad c_n | b_1 \quad \ldots \quad b_n]$$

to

$$[I|\underset{\mathcal{C}\leftarrow\mathcal{B}}{P}].$$

Example 10.3. *Let* $b_1 = \begin{bmatrix} -9 \\ 1 \end{bmatrix}$, $b_2 = \begin{bmatrix} -5 \\ -1 \end{bmatrix}$, $c_1 = \begin{bmatrix} 1 \\ -4 \end{bmatrix}$, *and* $c_2 = \begin{bmatrix} 3 \\ -5 \end{bmatrix}$, *and consider the bases for* $\mathbb{R}^2$ *given by* $\mathcal{B} = \{b_1, b_2\}$ *and* $\mathcal{C} = \{c_1, c_2\}$*. Find the change-of-coordinate matrix from* $\mathcal{B}$ *to* $\mathcal{C}$*.*

Solution. We can reduce the matrix $[c_1\ c_2|b_1\ b_2]$ to $[I|\underset{\mathcal{C}\leftarrow\mathcal{B}}{P}]$, and so we can find $\underset{\mathcal{C}\leftarrow\mathcal{B}}{P}$. Therefore, we have

$$\left[\begin{array}{cc|cc} 1 & 3 & -9 & -5 \\ -4 & -5 & 1 & -1 \end{array}\right] \underset{\longleftrightarrow}{\text{Replace R2 by R2+4R1}}$$

$$\left[\begin{array}{cc|cc} 1 & 3 & -9 & -5 \\ 0 & 7 & -35 & -21 \end{array}\right] \underset{\longleftrightarrow}{\text{Scaling R2 by 1/7}}$$

$$\left[\begin{array}{cc|cc} 1 & 3 & -9 & -5 \\ 0 & 1 & -5 & -3 \end{array}\right] \underset{\longleftrightarrow}{\text{Replace R1 by R1−3R2}} \left[\begin{array}{cc|cc} 1 & 0 & 6 & 4 \\ 0 & 1 & -5 & -3 \end{array}\right]$$

Therefore,

$$\underset{\mathcal{C}\leftarrow\mathcal{B}}{P} = \begin{bmatrix} 6 & 4 \\ -5 & -3 \end{bmatrix}.$$

Example 10.4. *Let* $b_1 = \begin{bmatrix} 1 \\ -3 \end{bmatrix}$, $b_2 = \begin{bmatrix} -2 \\ 4 \end{bmatrix}$, $c_1 = \begin{bmatrix} -7 \\ 9 \end{bmatrix}$, $c_2 = \begin{bmatrix} -5 \\ 7 \end{bmatrix}$, *and consider the bases for* $\mathbb{R}^2$ *given by* $\mathcal{B} = \{b_1, b_2\}$ *and* $\mathcal{C} = \{c_1, c_2\}$.

(1) *Find the change-of-coordinates matrix from* $\mathcal{C}$ *to* $\mathcal{B}$.
(2) *Find the change-of-coordinates matrix from* $\mathcal{B}$ *to* $\mathcal{C}$.

Solution.

(1) Note that we need to find $\underset{\mathcal{B}\leftarrow\mathcal{C}}{P}$, so compute

$$[b_1\ \ b_2|c_1\ \ c_2] = \left[\begin{array}{cc|cc} 1 & -2 & -7 & -5 \\ -3 & 4 & 9 & 7 \end{array}\right] \leftrightarrow \left[\begin{array}{cc|cc} 1 & 0 & 5 & 3 \\ 0 & 1 & 6 & 4 \end{array}\right].$$

Therefore,

$$\underset{\mathcal{B}\leftarrow\mathcal{C}}{P} = \begin{bmatrix} 5 & 3 \\ 6 & 4 \end{bmatrix}.$$

(2) We now want to compute $\underset{\mathcal{C}\leftarrow\mathcal{B}}{P}$. Note that

$$\underset{\mathcal{C}\leftarrow\mathcal{B}}{P} = (\underset{\mathcal{B}\leftarrow\mathcal{C}}{P})^{-1} = \begin{bmatrix} 5 & 3 \\ 6 & 4 \end{bmatrix}^{-1} = \begin{bmatrix} 2 & -3/2 \\ -3 & 5/2 \end{bmatrix}.$$

Remark. Let $\mathcal{B} = \{b_1, b_2, \dots, b_n\}$ and $\{c_1, \dots, c_n\}$ be bases for $\mathbb{R}^n$. We denoted by $P_{\mathcal{B}}$ the following matrix

$$P_{\mathcal{B}} = [b_1|b_2|\dots|b_n],$$

Also we had

$$P_{\mathcal{C}} = [c_1|c_2|\dots|c_n].$$

It was shown that

$$x = P_{\mathcal{B}}[x]_{\mathcal{B}} \qquad x = P_{\mathcal{C}}[x]_{\mathcal{C}}.$$

So we have

$$P_{\mathcal{C}}[x]_{\mathcal{C}} = P_{\mathcal{B}}[x]_{\mathcal{B}}.$$

Therefore,

$$[x]_{\mathcal{C}} = P_{\mathcal{C}}^{-1} P_{\mathcal{B}}[x]_{\mathcal{B}}.$$

We also have

$$[x]_{\mathcal{C}} = \underset{\mathcal{C}\leftarrow\mathcal{B}}{P}[x]_{\mathcal{B}}.$$

So,

$$P_{\mathcal{C}}^{-1}P_{\mathcal{B}} = \underset{\mathcal{C}\leftarrow\mathcal{B}}{P}.$$

• Change of basis for polynomials

Example 10.5. *Let* $\mathcal{B} = \{1+t, 1+t^2, 1+t+t^2\}$ *and* $\mathcal{C} = \{2-t, -t^2, 1+t^2\}$ *be bases for* $\mathbb{P}_2$. *Find* $\underset{\mathcal{C}\leftarrow\mathcal{B}}{P}$.

Solution. Let $\mathcal{E} = \{1, t, t^2\}$ be the standard basis for $\mathbb{P}_2$. Then

$$\begin{array}{rccc} T: & \mathbb{P}_2 & \to & \mathbb{R}^3 \\ & f & \mapsto & [f]_{\mathcal{E}} \end{array}$$

is an isomorphism. We have

$$[1+t]_{\mathcal{E}} = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, [1+t^2]_{\mathcal{E}} = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, [1+t+t^2]_{\mathcal{E}} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix},$$

and also

$$[2-t]_{\mathcal{E}} = \begin{bmatrix} 2 \\ -1 \\ 0 \end{bmatrix}, [-t^2]_{\mathcal{E}} = \begin{bmatrix} 0 \\ 0 \\ -1 \end{bmatrix}, [1+t^2]_{\mathcal{E}} = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}.$$

Now we have

$$\mathcal{B} = \left\{ \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \right\}$$

and

$$\mathcal{C} = \left\{ \begin{bmatrix} 2 \\ -1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ -1 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} \right\}$$

be bases for $\mathbb{R}^3$. We are looking for the matrix $\underset{\mathcal{C}\leftarrow\mathcal{B}}{P}$.

10.2. **Week 10, Lecture 2, Nov. 1, Eigenvalues and eigenvectors.**

Example 10.6. *Let* $A = \begin{bmatrix} 3 & -2 \\ 1 & 0 \end{bmatrix}$, $u = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$, $v = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$. *Then*

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

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

Precisely we have $Av = 2v$.

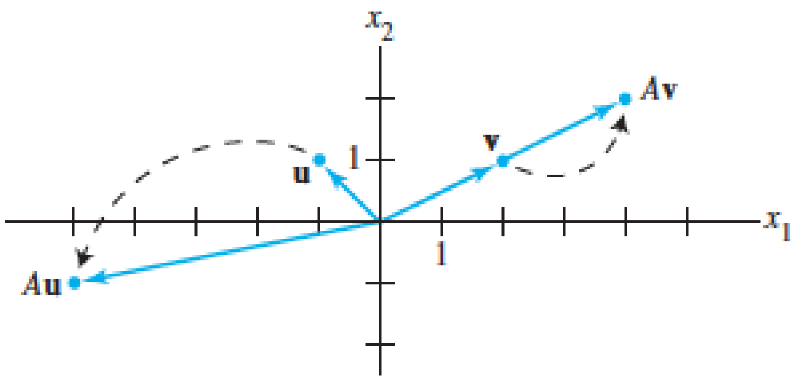


FIGURE 1 Effects of multiplication by A.

Definition. *An* **eigenvector** *of an* $n \times n$ *matrix* A *is a nonzero vector* x *such that* $Ax = \lambda x$ *for some scalar* λ. *A scalar* λ *is called an* **eigenvalue** *of* A *if there is a nonzero vector* x *such that* $Ax = \lambda x$*; such* x *is called an* **eigenvector corresponding to** λ.

Example 10.7. *Let* $A = \begin{bmatrix} 2 & -4 \\ -1 & -1 \end{bmatrix}$, $v = \begin{bmatrix} -4 \\ 1 \end{bmatrix}$, $u = \begin{bmatrix} 3 \\ 2 \end{bmatrix}$.

$$Av = \begin{bmatrix} 2 & -4 \\ -1 & -1 \end{bmatrix} \begin{bmatrix} -4 \\ 1 \end{bmatrix} = \begin{bmatrix} -12 \\ 3 \end{bmatrix} = 3 \begin{bmatrix} -4 \\ 1 \end{bmatrix}$$

so $\begin{bmatrix} -4 \\ 1 \end{bmatrix}$ *is an eigenvector and* 3 *is an eigenvalue.*

$$Au = \begin{bmatrix} 2 & -4 \\ -1 & -1 \end{bmatrix} \begin{bmatrix} 3 \\ 2 \end{bmatrix} = \begin{bmatrix} -2 \\ -5 \end{bmatrix} \neq \lambda \begin{bmatrix} 3 \\ 2 \end{bmatrix}$$

for some λ.

Example 10.8. *Show that* 7 *is an eigenvalue of* $A = \begin{bmatrix} 1 & 5 \\ 6 & 2 \end{bmatrix}$.

Solution. The number 7 is an eigenvalue. For some vector x we have

$$Ax = 7x$$

so

$$Ax - 7x = 0$$

we can write the above equation as

$$(A - 7I)x = 0$$

so if $(A - 7I)x = 0$ has a nonzero solution say x', then

$$(A - 7I)x' = 0 \Rightarrow Ax' - 7x' = 0$$
$$\Rightarrow Ax' = 7x'$$

and so 7 is an eigenvalue. Therefore, we only need to solve

$$(A - 7I)x = 0, \quad i.e.,$$
$$\left(\begin{bmatrix} 1 & 6 \\ 5 & 2 \end{bmatrix} - 7\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}\right)\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$
$$\Rightarrow \begin{bmatrix} -6 & 6 \\ 5 & -5 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = 0$$

when we solve the equation we have at least a nonzero solution $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$. Therefore 7 is an eigenvalue.

• **How to find all eigenvalues of a matrix A.**

λ is an eigenvalue for A if and only if

$$Ax = \lambda x \quad at\ least\ for\ a\ nonzero\ vector\ x.$$

So we can say λ is an eigenvalue of a matrix A if and only if

$$(A - \lambda I)x = 0 \quad at\ least\ for\ some\ nonzero\ x.$$

Which means the equation $(A - \lambda I)x = 0$ does not have only trivial solution if and only if

$$det(A - \lambda I) = 0$$

λ is an eigenvalue of A if and only if

$$det(A - \lambda I) = 0.$$

Definition. *The equation $det(A - \lambda I) = 0$ is called the* **characteristic equation**.

Definition. *Let λ be an eigenvalue of $n \times n$ matrix A. Then the* **eigenspace of A corresponding to λ** *is the solution set of*

$$(A - \lambda I)x = 0$$

Remark. Note that we already have the solution set of $(A - \lambda I)x = 0$ is a subspace.

Example 10.9. *let* $A = \begin{bmatrix} 4 & -1 & 6 \\ 2 & 1 & 6 \\ 2 & -1 & 8 \end{bmatrix}$.

(a) Find all eigenvalues of A.

(b) For each eigenvalue λ of A, find a basis for the eigenspace of A corresponding to λ.

Solution. (a) To find all eigenvalues of A we must find all λ such that

$$det(A - \lambda I) = 0.$$

Note that

$$det(A - \lambda I) = det(\begin{bmatrix} 4 & -1 & 6 \\ 2 & 1 & 6 \\ 2 & -1 & 8 \end{bmatrix} - \begin{bmatrix} \lambda & 0 & 0 \\ 0 & \lambda & 0 \\ 0 & 0 & \lambda \end{bmatrix}) = 0$$

$$\Rightarrow det(\begin{bmatrix} 4-\lambda & -1 & 6 \\ 2 & 1-\lambda & 6 \\ 2 & -1 & 8-\lambda \end{bmatrix}) = 0$$

you already know how to compute the determinant. We have

$$det\left(\begin{bmatrix} 4-\lambda & -1 & 6 \\ 2 & 1-\lambda & 6 \\ 2 & -1 & 8-\lambda \end{bmatrix}\right) = -(\lambda - 9)(\lambda - 2)^2$$

so $\lambda = 9$ and $\lambda = 2$, are the eigenvalues of A.

(b) We first find the basis for eigenspace of A corresponding to $\lambda = 2$, which is the same as the finding the basis of the solution set of $(A - 2I)x = 0$ which means we should find the basis for null space of $A - 2I$ (you know how to do it). The null space of $A - 2I$ contains all vectors $\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$ such that $(A - 2I)\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = 0$. i.e.,

$$\begin{bmatrix} 2 & -1 & 6 \\ 2 & -1 & 6 \\ 2 & -1 & 6 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = 0$$

The augmented matrix is

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

and the reduced echelon form is

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

So x_1 is basic and x_2 and x_3 are free. We have $x_1 - 1/2x_2 + 3x_3 = 0$

$$\Rightarrow x_1 = 1/2x_2 - 3x_3$$

Let $x_2 = t$ and $x_3 = s$. Then

$$x_1 = 1/2t - 3s.$$

So

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 1/2t - 3s \\ t \\ s \end{bmatrix} = t\begin{bmatrix} 1/2 \\ 1 \\ 0 \end{bmatrix} + s\begin{bmatrix} -3 \\ 0 \\ 1 \end{bmatrix}$$

so the eigenspace of A corresponding to 2 is

$$\left\{ t \begin{bmatrix} 1/2 \\ 1 \\ 0 \end{bmatrix} + s \begin{bmatrix} -3 \\ 0 \\ 1 \end{bmatrix} : s, t \in \mathbb{R} \right\}$$

and the basis for the eigenspace of A corresponding to 2 is

$$\left\{ \begin{bmatrix} 1/2 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -3 \\ 0 \\ 1 \end{bmatrix} \right\}.$$

Now you will find the eigenspace and the basis of it for $\lambda = 9$ (Do it as an exercise).

10.3. **Week 10, Lecture 3, Nov. 3, Characteristic polynomial and diagonalization.**

Theorem 10.10. *The eigenvalues of a triangular matrix are the entries on its main diagonal.*

Example 10.11. *Let* $A = \begin{bmatrix} a & b & c \\ 0 & d & e \\ 0 & 0 & f \end{bmatrix}$. *Then eigenvalues of A are a, d, and f.*

Why? because

$$det(A - \lambda I) = det(\begin{bmatrix} a & b & c \\ 0 & d & e \\ 0 & 0 & f \end{bmatrix} - \begin{bmatrix} \lambda & 0 & 0 \\ 0 & \lambda & 0 \\ 0 & 0 & \lambda \end{bmatrix}) =$$

$$det(\begin{bmatrix} a-\lambda & b & c \\ 0 & d-\lambda & e \\ 0 & 0 & f-\lambda \end{bmatrix}) =$$

$$(a-\lambda)(d-\lambda)(f-\lambda)$$

Theorem 10.12. *If $v_1, \ldots, v_r$ are eigenvectors that correspond to distinct eigenvalues $\lambda_1, \ldots, \lambda_r$ of an $n \times n$ matrix A, then the set $\{v_1, \ldots, v_r\}$ is linearly independent.*

Example 10.13. *let* $A = \begin{bmatrix} 4 & -1 & 6 \\ 2 & 1 & 6 \\ 2 & -1 & 8 \end{bmatrix}$. *Then 2 and 9 are eigenvalues of A.*

The eigenspace corresponding to 2 has a basis

$$\left\{ \begin{bmatrix} 1/2 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -3 \\ 0 \\ 1 \end{bmatrix} \right\}.$$

Also, the eigenspace corresponding to 9 has a basis

$$\left\{ \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \right\}.$$

Then

$$\left\{ \begin{bmatrix} 1/2 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \right\} \quad \textit{and} \quad \left\{ \begin{bmatrix} -3 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \right\}$$

are linearly independent.

• **When 0 is an eigenvalue of an $n \times n$ matrix A:**

If 0 is an eigenvalue, then there is a nonzero vector x such that $Ax = 0x$

$$\Rightarrow \quad Ax = 0$$

which means that $Ax = 0$ has a nonzero solution, which also means A is not invertible and $det\ A = 0$.

Theorem 10.14. *Let A be an $n \times n$ matrix. Then A is invertible if and only if one of the following holds:*

(1) *The number 0 is not eigenvalue of A.*
(2) *The determinant of A is not zero.*

• Similarity:

Definition. *Two $n \times n$ matrices A and B are said to be similar if there exists an invertible matrix P such that $A = PBP^{-1}$.*

Definition. *The expression $det(A - \lambda I) = 0$ is called the* **characteristic polynomial**.

Let A and B are similar. Then there exists an invertible matrix P such that

$$A = PBP^{-1} \qquad \Leftrightarrow \qquad A - \lambda I = PBP^{-1} - \lambda I$$

Note that $PP^{-1} = I$, so

$$A - \lambda I = PBP^{-1} - \lambda PP^{-1} = P(B - \lambda I)P^{-1}$$

Now

$$\begin{aligned} det(A - \lambda I) &= det(P(B - \lambda I)P^{-1}) \\ &= det(P)det(B - \lambda I)det(P^{-1}) \\ &= det(P)det(P^{-1})det(B - \lambda I) \\ &= det(B - \lambda I) \end{aligned}$$

Therefore, A and B have the same characteristic polynomial and so they have the same eigenvalues.

Proposition 10.15. *Similar matrices have the same characteristic polynomial and so they have the same eigenvalues.*

• Diagonalization

Example 10.16. *If* $D = \begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix}$, *Then*

$$D^2 = \begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix} \begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix} = \begin{bmatrix} 2^2 & 0 \\ 0 & 3^2 \end{bmatrix}$$

$$D^3 = \begin{bmatrix} 2^3 & 0 \\ 0 & 3^3 \end{bmatrix}$$

and for k we have

$$D^k = \begin{bmatrix} 2^k & 0 \\ 0 & 3^k \end{bmatrix}$$

Definition. *A matrix D is a* **diagonal matrix** *if it is of the form*

$$\begin{bmatrix} d_1 & 0 & 0 & \dots & 0 \\ 0 & d_2 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & d_n \end{bmatrix}.$$

Definition. *A matrix is called* **diagonalizable** *if A is similar to a diagonal matrix, i.e., there is an invertible matrix P and a diagonal matrix D such that*

$$A = PDP^{-1}.$$

Theorem 10.17. *An $n \times n$ matrix A is diagonalizable if and only if it has n linearly independent eigenvectors.*

• How to diagonalize a matrix:

(1) First check that if the matrix has n linearly dependent eigenvectors, if so, the matrix is diagonalizable.

(2) Find a basis for the set of all eigenvectors, say $\{v_1, \dots, v_n\}$.

(3) Let $P = [v_1| \dots |v_n]$, then $D = P^{-1}AP$ is an diagonal matrix with eigenvalues on its diagonal.

Example 10.18. *Find if $A = \begin{bmatrix} 1 & 2 \\ 0 & -3 \end{bmatrix}$ is diagonalizable, if so find an invertible matrix P and a diagonal matrix D such that $D = P^{-1}AP$.*

Solution. First we should find basis for eigenspaces. Note that $det(A - \lambda I) = (1 - \lambda)(-3 - \lambda)$. So, A has two eigenvalues 1 and -3. The eigenspace corresponding to 1 has the basis $\left\{ \begin{bmatrix} 1 \\ 0 \end{bmatrix} \right\}$ and the eigenspace corresponding to -3 has the basis $\left\{ \begin{bmatrix} -1/2 \\ 1 \end{bmatrix} \right\}$. Then we have $P = \begin{bmatrix} 1 & -1/2 \\ 0 & 1 \end{bmatrix}$, and $D = \begin{bmatrix} 1 & 0 \\ 0 & -3 \end{bmatrix}$. Check that $D = P^{-1}AP$.

11. Week 11, Diagonalization, Linear transformation and eigenvalues, and complex eigenvalues

11.1. **Week 11, Lecture 1, Nov. 6, Diagonalization.**

Goals of the week:
(1) Diagonalization
(2) Eigenvectors and linear transformation
(3) Complex Eigenvalues

Example 11.1. *If* $D = \begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix}$ *, Then*

$$D^2 = \begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix} \begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix} = \begin{bmatrix} 2^2 & 0 \\ 0 & 3^2 \end{bmatrix}$$

$$D^3 = \begin{bmatrix} 2^3 & 0 \\ 0 & 3^3 \end{bmatrix}$$

and for k *we have*

$$D^k = \begin{bmatrix} 2^k & 0 \\ 0 & 3^k \end{bmatrix}$$

Definition. *A matrix* D *is a* **diagonal matrix** *if it is of the form*

$$\begin{bmatrix} d_1 & 0 & 0 & \dots & 0 \\ 0 & d_2 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & d_n \end{bmatrix}.$$

Definition. *A matrix is called* **diagonalizable** *if* A *is similar to a diagonal matrix, i.e., there is an invertible matrix* P *and a diagonal matrix* D *such that*

$$A = PDP^{-1}.$$

Example 11.2. *Let* $A = \begin{bmatrix} 7 & 2 \\ -4 & 1 \end{bmatrix}$. *Find a formula for* A^k*, given that* $A = PDP^{-1}$. *Where* $P = \begin{bmatrix} 1 & 1 \\ -1 & -2 \end{bmatrix}$ *and* $D = \begin{bmatrix} 5 & 0 \\ 0 & 3 \end{bmatrix}$.

Solution. We can find the inverse of P which is

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

Then

$$A^2 = (PDP^{-1})(PDP^{-1}) = PD(P^{-1}P)DP^{-1} =$$

$$PD^2P^{-1} = \begin{bmatrix} 1 & 1 \\ -1 & -2 \end{bmatrix} \begin{bmatrix} 5 & 0 \\ 0 & 3 \end{bmatrix}^2 \begin{bmatrix} 2 & 1 \\ -1 & -2 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ -1 & -2 \end{bmatrix} \begin{bmatrix} 5^2 & 0 \\ 0 & 3^2 \end{bmatrix} \begin{bmatrix} 2 & 1 \\ -1 & -2 \end{bmatrix}$$

Again,

$$A^3 = AA^2 = (PDP^{-1})(PD^2P^{-1}) = PD(P^{-1}P)D^2P^{-1} = PD^3P^{-1}.$$

In general, for $k >= 1$,

$$A^k = PD^kP^{-1} = \begin{bmatrix} 1 & 1 \\ -1 & -2 \end{bmatrix} \begin{bmatrix} 5^k & 0 \\ 0 & 3^k \end{bmatrix} \begin{bmatrix} 2 & 1 \\ -1 & -2 \end{bmatrix}$$
$$= \begin{bmatrix} 2.5^k - 3^k & 5^k - 3^k \\ 2.3^k - 2.5^k & 2.3^k - 5^k \end{bmatrix}.$$

Theorem 11.3. *(The diagonal theorem) An $n \times n$ matrix A is diagonalizable if and only if A has n linearly independent eigenvectors.*

Definition. *An eigenvector basis of $\mathbb{R}^n$ corresponding to A is a basis $\{v_1, \dots, v_n\}$ of $\mathbb{R}^n$ such that $v_1, \dots, v_n$ are eigenvectors of A.*

• An $n \times n$ matrix A is diagonalizable if and only if there are eigenvectors $v_1, \dots, v_n$ such that $\{v_1, \dots, v_n\}$ are a basis for $\mathbb{R}^n$, i.e., $\{v_1, \dots, v_n\}$ is an eigenvector basis for $\mathbb{R}^n$ corresponding to A.

• **How to diagonalize an $n \times n$ matrix A.**

Step 1. First find the eigenvalues of A.

Step 2. Find a basis for each eigenspace. That is, if

$$det(A - \lambda I) = (x - \lambda_1)^{k_1}(x - \lambda_2)^{k_2} \dots (x - \lambda_p)^{k_p},$$

we should find the basis of eigenspace corresponding to each λ_i.

Step 3. If the number of all vectors in bases in Step 2 is n, then A is diagonalizable, otherwise it is not and we stop.

Step 4. Let $v_1, v_2, \dots, v_n$ be all vectors we find in Step 2, then

$$P = [v_1|v_2|\dots|v_n].$$

Step 5. Constructing D form eigenvalues. If the multiplicity of an eigenvalue λ_i is k_i, we repeat λ_i, k_i times, on the diagonal of D.

11.2. **Week 11, Lecture 2, Nov. 8, Diagonalization, Eigenvectors and linear transformations.**

Example 11.4. *Diagonalize the following matrix, if possible.*

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

That is, find an invertible matrix P and a diagonal matrix D such that $A = PDP^{-1}$.

Solution. Step 1. Find eigenvalues of A.

$$0 = det(A - \lambda I) = -\lambda^3 - 3\lambda^2 + 4 = -(\lambda - 1)(\lambda + 2)^2.$$

Therefore, $\lambda = 1$ and $\lambda = -2$ are the eigenvalues.

Step 2. Find a basis for each eigenspace. The eigenspace corresponding to $\lambda = 1$ is the solution set of

$$(A - I)x = 0.$$

A basis for this space is

$$\left\{ \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \right\}.$$

The eigenspace corresponding to $\lambda = -2$ is the solution set of

$$(A - (-2)I)x = 0.$$

A basis for this space is

$$\left\{ \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} \right\}.$$

Step 3. Since we find three vectors

$$\left\{ \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} \right\}.$$

So A is diagonalizable.

Step 4.

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

Step 5.

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

It is a good idea to check that P and D work, i.e.,

$$A = PDP^{-1} \quad \text{or} \quad AP = PD.$$

If we compute we have

$$AP = \begin{bmatrix} 1 & 2 & 2 \\ -1 & -2 & 0 \\ 1 & 0 & -2 \end{bmatrix} \qquad PD = \begin{bmatrix} 1 & 2 & 2 \\ -1 & -2 & 0 \\ 1 & 0 & -2 \end{bmatrix}.$$

Example 11.5. *Diagonalize the following matrix, if possible.*

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

Solution. First we find the eigenvalues, which are the roots of characteristic polynomial $det(A - \lambda I)$.

$$0 = det(A - \lambda I) = -\lambda^3 - 3\lambda^2 + 4 = -(\lambda - 1)(\lambda + 2)^2$$

So $\lambda = 1$ and $\lambda = -2$ are eigenvalues.

A basis for eigenspace corresponding to $\lambda = 1$ is

$$\left\{ \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix} \right\}$$

and a basis for eigenspace corresponding to $\lambda = -2$ is

$$\left\{ \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} \right\}.$$

Since we can not find 3 eigenvectors that are linearly independent, so A is not diagonalizable.

Theorem 11.6. *An $n \times n$ matrix with n distinct eigenvalues is diagonalizable.*

Theorem 11.7. *Let characteristic polynomial of A is*

$$(x - \lambda_1)^{k_1}(x - \lambda_2)^{k_2} \ldots (x - \lambda_p)^{k_p}.$$

(1) *For each $1 \leq i \leq p$ The dimension of eigenspace corresponding to λ_i is at most k_i.*
(2) *The matrix A is diagonalizable if and only if the sum of the dimensions of the eigenspaces equals n, and this happens if and only if*
   (a) *the characteristic polynomial factors completely into linear factors and*
   (b) *the dimension of the eigenspace for each λ_i equals the multiplicity of λ_i.*
(3) *If A is diagonalizable and $\mathcal{B}_i$ is a basis for the eigenspace corresponding to λ_i for each i, then the total collection of vectors in the sets $\mathcal{B}_1, \ldots, \mathcal{B}_p$ forms an eigenvector basis for $\mathbb{R}^n$.*

(1) Let $a, b_1, b_2, \ldots, b_n$ be vectors in $\mathbb{R}^m$. Complete the following definitions:
   (a) The set $\{b_1, b_2, \ldots, b_n\}$ is said to be linearly independent if and only if $c_1b_1 + c_2b_2 + \ldots + c_nb_n = 0$, for scalars $c_1, \ldots, c_n$, we have that $c_1 = c_2 = \ldots = c_n = 0$. ...

$$a = c_1b_1 + c_2b_2 + \ldots + c_nb_n$$

   for some scalars $c_1, c_2, \cdots, c_n$.
   (b) The set $\mathcal{B} = \{b_1, b_2, \ldots, b_n\}$ is a basis for the subspace H if it is linearly independent and every $a \in H$ is a linear combination of elements in $\mathcal{B}$.

(2) What does it mean $T : V \to W$ is a linear transformation? It means T is a function such that for every $x, y \in V$ and $c \in \mathbb{R}$, $T(cx) = cT(x)$ and $T(x + y) = T(x) + T(y)$.

(3) Let $\mathcal{B} = \{v_1, v_2, \ldots, v_n\}$ and $\mathcal{C} = \{w_1, w_2, \ldots, w_n\}$ be two bases for vector space V. Write a formula for $\underset{\mathcal{C}\leftarrow\mathcal{B}}{P}$.
   We have $\underset{\mathcal{C}\leftarrow\mathcal{B}}{P} = [[v_1]_{\mathcal{C}}\ [v_2]_{\mathcal{C}}\ \ldots\ [v_n]_{\mathcal{C}}]$.

11.3. **Week 11, Lecture 3, Nov. 10, Eigenvectors and linear transformations.**

When A is diagonalizable there exist an invertible matrix P and a diagonal matrix D such that $A = PDP^{-1}$. Our goal is to show that the following two linear transformations are essentially the same.

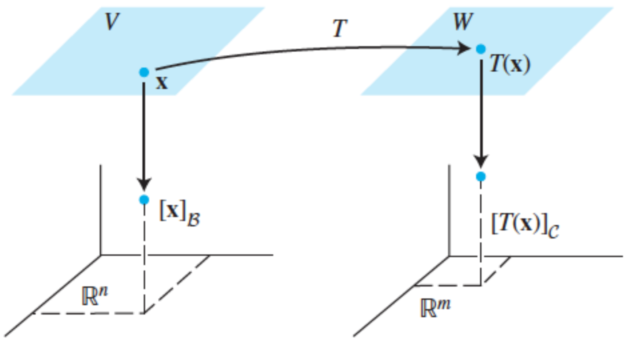
$$\begin{array}{ccc} \mathbb{R}^n & \to & \mathbb{R}^n \\ x & \mapsto & Ax \end{array} \qquad\qquad \begin{array}{ccc} \mathbb{R}^n & \to & \mathbb{R}^n \\ u & \mapsto & Du \end{array}$$

Remark. Let $\mathcal{B} = \{b_1, \ldots, b_n\}$ be a basis for a vector space V. Then the coordinate mapping

$$\begin{array}{cccc} T: & V & \to & \mathbb{R}^n \\ & x & \mapsto & [x]_{\mathcal{B}} \end{array}$$

is a one-to-one linear transformation form V onto $\mathbb{R}^n$.

• The matrix of a linear transformation: Let V be an n-dimensional vector space and W be an m-dimensional vector space.



Let $\mathcal{B}$ and $\mathcal{C}$ be bases for V and W, respectively. The connection between $[x]_{\mathcal{B}}$ and $[T(x)]_{\mathcal{C}}$ is easy to find. Let $\mathcal{B} = \{b_1, b_2, \ldots, b_n\}$ be the basis of V. If $x = r_1b_1 + r_2b_2 + \ldots + r_nb_n$, then

$$x_{\mathcal{B}} = \begin{bmatrix} r_1 \\ r_2 \\ \vdots \\ r_n \end{bmatrix}.$$

Note that

$$T(x) = T(r_1b_1 + r_2b_2 + \ldots + r_nb_n) = r_1T(b_1) + r_2T(b_2) + \ldots + r_nT(b_n).$$

Since the coordinate mapping from W to $\mathbb{R}^m$ is a linear transformation, we have

$$[T(x)]_{\mathcal{C}} = [r_1T(b_1) + r_2T(b_2) + \ldots + r_nT(b_n)]_{\mathcal{C}} =$$

$$r_1[T(b_1)]_{\mathcal{C}} + r_2[T(b_2)]_{\mathcal{C}} + \ldots + r_n[T(b_n)]_{\mathcal{C}} =$$

$$[\ [T(b_1)]_{\mathcal{C}} \quad [T(b_2)]_{\mathcal{C}} \quad \dots \quad [T(b_n)]_{\mathcal{C}}\ ] \begin{bmatrix} r_1 \\ r_2 \\ \vdots \\ r_n \end{bmatrix} =$$

$$[\ [T(b_1)]_{\mathcal{C}} \quad [T(b_2)]_{\mathcal{C}} \quad \dots \quad [T(b_n)]_{\mathcal{C}}\ ]\, [x]_{\mathcal{B}}.$$

So

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

where

$$M = [\ [T(b_1)]_{\mathcal{C}} \quad [T(b_2)]_{\mathcal{C}} \quad \dots \quad [T(b_n)]_{\mathcal{C}}\ ].$$

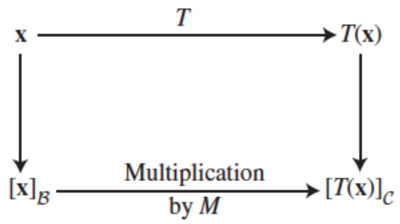
So

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

where

$$M = [\ [T(b_1)]_{\mathcal{C}} \quad [T(b_2)]_{\mathcal{C}} \quad \dots \quad [T(b_n)]_{\mathcal{C}}\ ].$$

The matrix M is called **matrix for T relative to the bases $\mathcal{B}$ and $\mathcal{C}$.**



Example 11.8. *Let $\mathcal{B} = \{b_1, b_2\}$ be a basis for V and $\mathcal{C} = \{c_1, c_2, c_3\}$ be a basis for W. Let $T : V \to W$ be a linear transformation such that*

$$T(b_1) = 3c_1 - 2c_2 + 5c_3 \qquad\qquad T(b_2) = 4c_1 + 7c_2 - c_3$$

Find matrix M for T relative to $\mathcal{B}$ and $\mathcal{C}$.

Solution. We have that

$$M = [[T(b_1)]_{\mathcal{C}} \quad [T(b_2)]_{\mathcal{C}}].$$

We have

$$[T(b_1)] = \begin{bmatrix} 3 \\ -2 \\ 5 \end{bmatrix} \qquad [T(b_2)] = \begin{bmatrix} 4 \\ 7 \\ -1 \end{bmatrix}.$$

So

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

• **Linear transformation from V into V**

Now, we want to find the matrix M when V and W are the same, and the basis $\mathcal{C}$ is the same as $\mathcal{B}$. The matrix M in this case called **Matrix for T relative to $\mathcal{B}$**, or simply $\mathcal{B}$-matrix for T.

The $\mathcal{B}$-matrix for T satisfies $$[T(x)]_{\mathcal{B}} = [T]_{\mathcal{B}}[x]_{\mathcal{B}} \quad \text{for all } x \text{ in } V.$$

Example 11.9. *The linear transformation $T : \mathbb{P}_2 \to \mathbb{P}_2$ defined by*

$$T(a_0 + a_1 t + a_2 t^2) = a_1 + 2a_2 t$$

is a linear transformation.

(1) *Find the $\mathcal{B}$-matrix for T, when $\mathcal{B}$ is the basis $\{1, t, t^2\}$.*
(2) *Verify that $[T(p)]_{\mathcal{B}} = [T]_{\mathcal{B}}[p]_{\mathcal{B}}$ for each $p \in \mathbb{P}_2$.*

Solution. (1) We have that

$$[T]_{\mathcal{B}} = [[T(1)]_{\mathcal{B}} \;\; [T(t)]_{\mathcal{B}} \;\; [T(t^2)]_{\mathcal{B}}].$$

Note that

$$T(1) = 0 \qquad T(t) = 1 \qquad T(t^2) = 2t$$

Therefore,

$$[T(1)]_{\mathcal{B}} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \qquad [T(t)]_{\mathcal{B}} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \qquad [T(t^2)]_{\mathcal{B}} = \begin{bmatrix} 0 \\ 2 \\ 0 \end{bmatrix}.$$

So

$$[T]_{\mathcal{B}} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{bmatrix}.$$

(2) Any polynomial $p(t) \in \mathbb{P}_2$ is of the form $p(t) = a_0 + a_1 t + a_2 t^2$ for some scalars a_0, a_1 and a_2. Thus,

$$[T(p)]_{\mathcal{B}} = [a_1 + 2a_2 t]_{\mathcal{B}} = \begin{bmatrix} a_1 \\ 2a_2 \\ 0 \end{bmatrix}$$

and

$$[T(p)]_{\mathcal{B}} = [T]_{\mathcal{B}}[p]_{\mathcal{B}} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ a_2 \end{bmatrix} = \begin{bmatrix} a_1 \\ 2a_2 \\ 0 \end{bmatrix}.$$

• Linear transformation on $\mathbb{R}^n$

Theorem 11.10. *Diagonal matrix representation Suppose that $A = PDP^{-1}$ where P is an invertible matrix and D is a diagonal matrix. Assume that*

$$P = [v_1|v_2|\ldots|v_n].$$

Let $\mathcal{B} = \{v_1, v_2, \ldots, v_n\}$. Let

$$\begin{array}{rccc} T: & \mathbb{R}^n & \to & \mathbb{R}^n \\ & x & \mapsto & Ax \end{array}$$

Then $D = [T]_{\mathcal{B}}$*, i.e.,*

$$[T(x)]_{\mathcal{B}} = D[x]_{\mathcal{B}}.$$

Example 11.11. *Define* $T : \mathbb{R}^2 \to \mathbb{R}^2$ *by* $T(x) = Ax$*, where* $A = \begin{bmatrix} 7 & 2 \\ -4 & 1 \end{bmatrix}$*. Find a basis for* $\mathbb{R}^2$ *with the property that the* $\mathcal{B}$*-matrix for* T *is a diagonal matrix.*

Solution. By the previous Theorem if we find an invertible matrix P and a diagonal matrix D such that $A = PDP^{-1}$, then the columns of P produce the basis $\mathcal{B}$. We can find $P = \begin{bmatrix} 1 & 1 \\ -1 & -2 \end{bmatrix}$ and $D = \begin{bmatrix} 5 & 0 \\ 0 & 3 \end{bmatrix}$ such that $A = PDP^{-1}$. So $\mathcal{B} = \{\begin{bmatrix} 1 \\ -1 \end{bmatrix}, \begin{bmatrix} 1 \\ -2 \end{bmatrix}\}$.

Theorem 11.12. *Suppose that* $A = PCP^{-1}$ *where* P *is an invertible matrix. Assume that*

$$P = [v_1|v_2|\dots|v_n].$$

Let $\mathcal{B} = \{v_1, v_2, \dots, v_n\}$*. Let*

$$\begin{array}{rccc} T: & \mathbb{R}^n & \to & \mathbb{R}^n \\ & x & \mapsto & Ax \end{array}$$

Then $C = [T]_{\mathcal{B}}$*, i.e.,*

$$[T(x)]_{\mathcal{B}} = C[x]_{\mathcal{B}}.$$

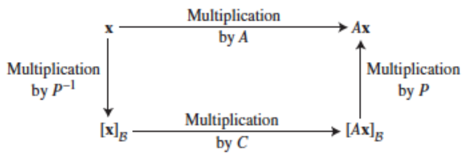


FIGURE 5 Similarity of two matrix representations: $A = PCP^{-1}$.

12. Week 12, Inner Product and Orthogonality

12.1. Week 12, Lecture 1, Nov. 13, Inner Product, length and orthogonality.

Goals:

(1) Complex eigenvalues
(2) Inner product and length
(3) Orthogonal sets

Definition. *A complex eigenvalue for a matrix A is a complex scalar λ such that there is a non-zero vector x in $\mathbb{C}^n$ s.t $Ax = \lambda x$. Moreover, x is called a complex eigenvector corresponding to λ.*

Remark. The complex eigenvalues are the roots of $det(A - \lambda I)$. Also, the set of all eigenvectors corresponding to λ are the non-zero vectors $x \in \mathbb{C}^n$ such that

$$(A - \lambda I)x = 0.$$

Example 12.1. *If* $A = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$ *, find eigenvalues.*

Solution. To find the eigenvalues, we should find the roots of $det(A - \lambda I)$.

$$det(A - \lambda I) = det \begin{bmatrix} 0-\lambda & -1 \\ 1 & 0-\lambda \end{bmatrix} = \lambda^2 + 1$$

The roots of $\lambda^2 + 1$ are i and $-i$. So eigenvalues are i and $-i$. And also we have

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

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

So $\begin{bmatrix} 1 \\ i \end{bmatrix}$ and $\begin{bmatrix} 1 \\ -i \end{bmatrix}$ are eigenvectors corresponding to $-i$ and i respectively.

• The inner product

Let $u = \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix} \in \mathbb{R}^n$, then $u^T = [u_1 u_2 \dots u_n]$. The **inner product(or dot product)** of two vectors $u, v \in \mathbb{R}^n$ is the number $u^T v$, and often it is written as $u.v$.

Example 12.2. *Compute $u.v$ and $v.u$ for* $u = \begin{bmatrix} 2 \\ -5 \\ -1 \end{bmatrix}$ *and* $v = \begin{bmatrix} 3 \\ 2 \\ -3 \end{bmatrix}$.

Solution.

$$u.v = u^T v = \begin{bmatrix} 2 & -5 & -1 \end{bmatrix} \begin{bmatrix} 3 \\ 2 \\ -3 \end{bmatrix} = 2 \times 3 + (-5) \times 2 + (-1) \times (-3) = -1$$

$$v.u = v^T u = \begin{bmatrix} 3 & 2 & -3 \end{bmatrix} \begin{bmatrix} 2 \\ -5 \\ -1 \end{bmatrix} = 3 \times 2 + 2 \times (-5) + (-3) \times (-1) = -1$$

Theorem 12.3. *Let u, v and w be vectors in $\mathbb{R}^n$, and let c be a scalar. Then*

a. $u.v = v.u$
b. $(u + v).w = u.w + v.w$
c. $(cu).v = c(u.v) = u.(cv)$
d. $u.u \geq 0$ *and* $u.u = 0$ *if and only if* $u = 0$.

Combining (b) and (c) we have

$$(c_1 u_1 + \ldots + c_p u_p).w = c_1(u_1.w) + \ldots + c_p(u_p.w).$$

• **The length of a vector:**

Definition. *The* **length (or norm)** *of* $v = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix}$ *is the nonnegative scalar* $\|v\|$ *defined by*

$$\|v\| = \sqrt{v.v} = \sqrt{v_1^2 + v_2^2 + \ldots + v_n^2}$$

and $\|v\|^2 = v.v$.

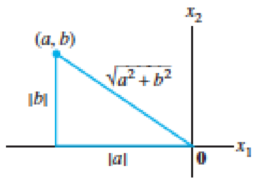


FIGURE 1
Interpretation of $\|\mathbf{v}\|$ as length.

• For any scalar c, the length of cv is $|c|$ times the length of v, that is

$$\|cv\| = |c|\|v\|.$$

Definition. *A vector v with $\|v\| = 1$ is called a* **unit vector**.

Normalizing a vector: Let u be a vector, then $(1/\|u\|)u$ is a unit vector. The process of dividing a vector to its length is called **normalizing**. Moreover, u and $(1/\|u\|)u$ have the same direction.

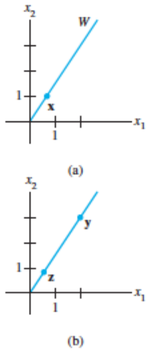


FIGURE 2
Normalizing a vector to produce a unit vector.

Example 12.4. *Let $v = (1, -2, 2, 4)$. Find a unit vector u in the same direction as v.*

Solution. First compute the length of v:

$$\|v\| = \sqrt{v.v} = \sqrt{1^2 + (-2)^2 + 2^2 + 4^2} = \sqrt{25} = 5$$

Then we multiply v by $1/\|v\|$ to obtain u.

$$u = 1/\|v\|v = 1/5v = 1/5 \begin{bmatrix} 1 \\ -2 \\ 2 \\ 4 \end{bmatrix} = \begin{bmatrix} 1/5 \\ -2/5 \\ 2/5 \\ 4/5 \end{bmatrix}.$$

To check $\|u\| = 1$,

$$\|u\| = \sqrt{u.u} = \sqrt{(1/5)^2 + (-2/5)^2 + (2/5)^2 + (4/5)^2} =$$

$$\sqrt{1/25 + 4/25 + 4/25 + 16/25} = \sqrt{25/25} = 1$$

Example 12.5. *Let W be a subspace of $\mathbb{R}^2$ spanned by $x = \begin{bmatrix} 3/2 \\ 1 \end{bmatrix}$. Find a unit vector z that is a basis for W.*

Solution. Note that $W = \{c \begin{bmatrix} 3/2 \\ 1 \end{bmatrix} : c \in \mathbb{R}\}$. We have that $1/\|x\| \in \mathbb{R}$ so $(1/\|x\|)x$ is a vector in W, and spanning it. It is enough to compute $1/\|x\|x$.

$$\|x\| = \sqrt{x.x} = \sqrt{(3/2)^2 + 1^2} = \sqrt{9/4 + 1} = \sqrt{13/4} = \sqrt{13}/2$$

so $(1/\|x\|)x = \frac{1}{\sqrt{13}/2} \begin{bmatrix} 3/2 \\ 1 \end{bmatrix} = 2/\sqrt{13} \begin{bmatrix} 3/2 \\ 1 \end{bmatrix} = \begin{bmatrix} 6/2\sqrt{13} \\ 2/\sqrt{13} \end{bmatrix}$.

12.2. **Week 12, Lecture 2, Nov. 15, Distance in $\mathbb{R}^n$ and Orthogonality.**

Definition. *For u and v in $\mathbb{R}^n$, the* **distance** *between u and v, written as $dist(u, v)$, is the length of vector $u - v$. That is $dist(u, v) = ||u - v||$.*

Example 12.6. *Compute the distance between the vectors $u = (7, 1)$ and $v = (3, 2)$.*

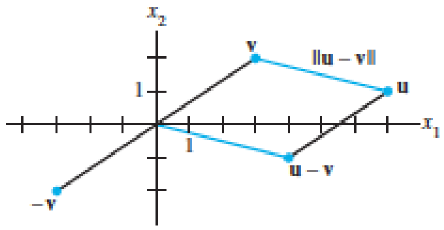


FIGURE 4 The distance between **u** and **v** is the length of **u** − **v**.

Solution.

$$u - v = \begin{bmatrix} 7 \\ 1 \end{bmatrix} - \begin{bmatrix} 3 \\ 2 \end{bmatrix} = \begin{bmatrix} 4 \\ -1 \end{bmatrix}$$

$$||u - v|| = \sqrt{4^2 + (-1)^2} = \sqrt{17}$$

Example 12.7. *If $u = (u_1, u_2, u_3)$ and $v = (v_1, v_2, v_3)$, then*

$$dist(u, v) = ||u - v|| = \sqrt{(u - v).(u - v)} = \sqrt{(u_1 - v_1)^2 + (u_2 - v_2)^2 + (u_3 - v_3)^2}$$

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

Lemma 12.8. *If vectors u and v are orthogonal, then*

$$dist(u, v) = \|u - v\| = \|u - (-v)\| = \|u + v\| = dist(u, -v).$$

Theorem 12.9. *(The pythagorean Theorem) Two vectors u and v are orthogonal if and only if*

$$\|u + v\|^2 = \|u\|^2 + \|v\|^2.$$

Orthogonal Complement

Definition.
- *If a vector z is orthogonal to every vector in a subspace W of $\mathbb{R}^n$, then z is said to be* **orthogonal to** W.
- *The set of all vectors z that are orthogonal to W is said* **orthogonal complement of** W *and is denoted by $W^\perp$ (W perp)*

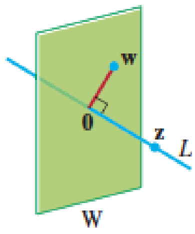


FIGURE 7
A plane and line through **0** as orthogonal complements.

Theorem 12.10. (1) *A vector x is in $W^{\perp}$ if and only if x is orthogonal to every vector in a set that spans W.*
(2) *$W^{\perp}$ is a subspace of $\mathbb{R}^n$.*

Definition. *Let $A = [A_1|A_2|\dots|A_n]$ be an $m \times n$ matrix. Also A has m rows, denote them by $A_1^{'}, \dots, A_m^{'}$.*

$$Col\ A = span\{A_1, \cdots, A_n\} \qquad Row\ A = span\{A_1^{'}, \dots, A_m^{'}\}.$$

Theorem 12.11. *Let A be an $m \times n$ matrix.*
(1) *$(Row\ A)^{\perp} = Nul\ A$, that is the orthogonal complement of the row space of A is the null space of A.*
(2) *$(Col\ A)^{\perp} = Nul\ A^T$, that is the orthogonal complement of the column space of A is the null space of A^T.*

- Let u and v be in $\mathbb{R}^2$ or $\mathbb{R}^3$, then
  (1)
$$u.v = \|u\|\|v\|cos\theta,$$
  where θ is the angle between the two line segments from the origin to the points identified with u and v.
  (2) We also have
$$\|u - v\|^2 = \|u\|^2 + \|v\|^2 - 2\|u\|\|v\|cos\theta$$

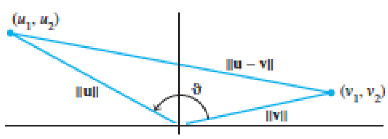


FIGURE 9 The angle between two vectors.

Example 12.12. *Find the angle between $u = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ and $v = \begin{bmatrix} -1 \\ 0 \end{bmatrix}$*

Solution. We have
$$u.v = \|u\|\|v\|cos\theta.$$

Note that $\|u\| = \sqrt{1+1^2} = \sqrt{2}$ and $\|v\| = \sqrt{(-1)^2 + 0^2} = 1$ and $u.v = u^T.v = -1$. So $-1 = \sqrt{2}.cos\theta$. Therefore, $\theta = \frac{3\pi}{4}$.

• **Orthogonal Sets:**

Definition. *A set of vectors $\{u_1, u_2, \ldots, u_p\}$ in $\mathbb{R}^n$ is said to be orthogonal set if each pair of distinct vectors from the set are orthogonal, that is, $u_i.u_j = 0$ if $i \neq j$.*

Example 12.13. *Show that $\{u_1, u_2, u_3\}$ is an orthogonal set where*

$$u_1 = \begin{bmatrix} 3 \\ 1 \\ 1 \end{bmatrix}, u_2 = \begin{bmatrix} -1 \\ 2 \\ 1 \end{bmatrix}, \text{ and } u_3 = \begin{bmatrix} -1/2 \\ -2 \\ 7/2 \end{bmatrix}.$$

Solution. We must show that $u_1.u_2 = 0$, $u_1.u_3 = 0$, and $u_2.u_3 = 0$.

$$u_1.u_2 = 3(-1) + 1(2) + 1(1) = 0 \qquad u_1.u_3 = 3(-1/2) + 1(-2) + 1(7/2) = 0$$
$$u_2.u_3 = -1(-1/2) + 2(-2) + 1(7/2) = 0.$$

Theorem 12.14. *If $S = \{u_1, u_2, u_3\}$ is an orthogonal set of non-zero vectors in $\mathbb{R}^n$, then S is linearly independent and hence is a basis for the subspace spanned by S.*

Definition. *An orthogonal basis for a subspace W of $\mathbb{R}^n$ is a basis for W that is also orthogonal set.*

Theorem 12.15. *Let $\{u_1, \ldots, u_p\}$ be an orthogonal basis for a subspace W of $\mathbb{R}^n$. For each $y \in W$, the weights in the linear combination*

$$y = c_1u_1 + \cdots + c_pu_p$$

are given by

$$c_j = \frac{y.u_j}{u_j.u_j} \quad (j = 1, 2, \ldots, p)$$

Example 12.16. *The set $S = \{u_1, u_2, u_3\}$, where*

$$u_1 = \begin{bmatrix} 3 \\ 1 \\ 1 \end{bmatrix}, u_2 = \begin{bmatrix} -1 \\ 2 \\ 1 \end{bmatrix}, \text{ and } u_3 = \begin{bmatrix} -1/2 \\ -2 \\ 7/2 \end{bmatrix}$$

is an orthogonal basis for $\mathbb{R}^3$. Express the vector $y = \begin{bmatrix} 6 \\ 1 \\ -8 \end{bmatrix}$ as a linear combination of the vectors in S.

Solution. If we write $y = c_1u_1 + c_2u_2 + c_3u_3$, then

$$c_1 = \frac{y.u_1}{u_1.u_1} = \frac{11}{11} = 1 \qquad c_2 = \frac{y.u_2}{u_2.u_2} = \frac{-12}{6} = -2$$
$$c_3 = \frac{y.u_3}{u_3.u_3} = \frac{-33}{33/2} = -2$$

Therefore, $y = 1u_1 - 2u_2 - 2u_3$.

12.3. **Week 12, Lecture 3, Nov. 17, Orthogonal projection and orthonormal sets.** Assume that u is in $\mathbb{R}^n$. then $L = span\{u\} = \{cu : c \in \mathbb{R}\}$ is a line.

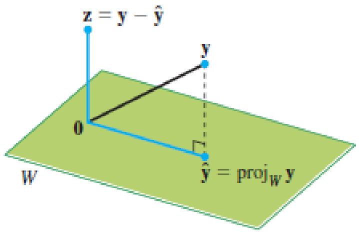


FIGURE 2
Finding α to make $\mathbf{y} - \hat{\mathbf{y}}$ orthogonal to $\mathbf{u}$.

We want to write a vector y as a sum of a vector in $L = span\{u\}$ and a vector orthogonal to u. Then $y = \hat{y} + (y - \hat{y})$, where

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

$\hat{y} = \mathbf{proj}_L y$ is called orthogonal projection of y onto L. Also $y - \hat{y}$ is called the **complement of y orthogonal to u.**

Example 12.17. *Let* $y = \begin{bmatrix} 7 \\ 6 \end{bmatrix}$, *and* $u = \begin{bmatrix} 4 \\ 2 \end{bmatrix}$. *Find the orthogonal projection of y onto u. Then write y as the sum of two orthogonal vectors, one in $span\{u\}$ and one orthogonal to u.*

Solution. Compute

$$y.u = \begin{bmatrix} 7 \\ 6 \end{bmatrix} \begin{bmatrix} 4 \\ 2 \end{bmatrix} = 40$$

$$u.u = \begin{bmatrix} 4 \\ 2 \end{bmatrix} \begin{bmatrix} 4 \\ 2 \end{bmatrix} = 20$$

$$\Rightarrow \hat{y} = \frac{y.u}{u.u} u = (40/20)u = 2 \begin{bmatrix} 4 \\ 2 \end{bmatrix} = \begin{bmatrix} 8 \\ 4 \end{bmatrix}$$

and the complement of y orthogonal to u.

$$y - \hat{y} = \begin{bmatrix} 7 \\ 6 \end{bmatrix} - \begin{bmatrix} 8 \\ 4 \end{bmatrix} = \begin{bmatrix} -1 \\ 2 \end{bmatrix}.$$

so $y = \hat{y} + (y - \hat{y}) = \begin{bmatrix} 8 \\ 4 \end{bmatrix} + \begin{bmatrix} -1 \\ 2 \end{bmatrix}$.

• It is easy to visualize the case in which $w = \mathbb{R}^2 = span\{u_1, u_2\}$ with u_1 and u_2 orthogonal. Any $y \in \mathbb{R}^2$ can be written in the form

$$y = \frac{y.u_1}{u_1.u_1} u_1 + \frac{y.u_2}{u_2.u_2} u_2$$

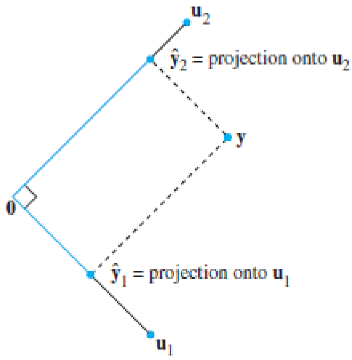


FIGURE 4 A vector decomposed into the sum of two projections.

Definition. *A set* $\{u_1, \dots, u_p\}$ *is an* **orthonormal set** *if it is an orthogonal of unit vectors.*

Example 12.18. *Show that* $\{v_1, v_2, v_3\}$ *is an orthonormal basis of* $\mathbb{R}^3$*. Where*

$$v_1 = \begin{bmatrix} 3/\sqrt{11} \\ 1/\sqrt{11} \\ 1/\sqrt{11} \end{bmatrix}, v_2 = \begin{bmatrix} -1/\sqrt{6} \\ 2/\sqrt{6} \\ 1/\sqrt{6} \end{bmatrix}, and \quad v_3 = \begin{bmatrix} -1/\sqrt{66} \\ -4/\sqrt{66} \\ 7/\sqrt{66} \end{bmatrix}$$

Solution. Compute

$$v_1.v_2 = -3/\sqrt{66} + 2/\sqrt{66} + 1/\sqrt{66} = 0$$

$$v_1.v_3 = -3/\sqrt{726} + -4/\sqrt{726} + 7/\sqrt{726} = 0$$

$$v_2.v_3 = 1/\sqrt{396} + -8/\sqrt{396} + 7/\sqrt{396} = 0$$

so $\{v_1, v_2, v_3\}$ is an orthogonal set.

Now we show that v_1, v_2, v_3 are unit vector.

$$\|u_1\| = \sqrt{v_1.v_1} = \sqrt{9/11 + 1/11 + 1/11} = 1$$

$$\|u_2\| = \sqrt{v_2.v_2} = \sqrt{1/6 + 4/6 + 1/6} = 1$$

$$\|u_3\| = \sqrt{v_3.v_3} = \sqrt{1/66 + 16/66 + 49/66} = 1$$

So $\{v_1, v_2, v_3\}$ is orthonormal basis for $\mathbb{R}^3$.

Theorem 12.19. *An* $m \times n$ *matrix* U *has orthonormal columns if and only if* $U^T U = I$.

Theorem 12.20. *Let* U *be an* $m \times n$ *matrix with orthonormal columns and let* x *and* y *be in* $\mathbb{R}^n$*. Then*

(1) $\|Ux\| = \|x\|$.
(2) $(Ux).(Uy) = x.y$.
(3) $(Ux).(Uy) = 0 \quad if\ and\ only\ if \quad x.y = 0$

Example 12.21. *Let* $U = \begin{bmatrix} 1/\sqrt{2} & 2/3 \\ 1/\sqrt{2} & -2/3 \\ 0 & 1/3 \end{bmatrix}$ *and* $x = \begin{bmatrix} \sqrt{2} \\ 3 \end{bmatrix}$*. Notice that* U *has orthonormal columns and*

$$U^T U = \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} & 0 \\ 2/3 & -2/3 & 1/3 \end{bmatrix} \begin{bmatrix} 1/\sqrt{2} & 2/3 \\ 1/\sqrt{2} & -2/3 \\ 0 & 1/3 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

verify that $||Ux|| = ||x||$*.*

$$Ux = \begin{bmatrix} 1/\sqrt{2} & 2/3 \\ 1/\sqrt{2} & -2/3 \\ 0 & 1/3 \end{bmatrix} \begin{bmatrix} \sqrt{2} \\ 3 \end{bmatrix} = \begin{bmatrix} 3 \\ -1 \\ 1 \end{bmatrix}.$$

$$\|Ux\| = \sqrt{9+1+1} = \sqrt{111}$$

$$\|Ux\| = \sqrt{2+9} = \sqrt{11}$$

Definition. *An* **orthonormal matrix** *is a square invertible matrix* U *such that*

$$U^{-1} = U^T.$$

Example 12.22. *The matrix*

$$U = \begin{bmatrix} 3/\sqrt{11} & -1/\sqrt{6} & -1/\sqrt{66} \\ 1/\sqrt{11} & 2/\sqrt{6} & -4/\sqrt{66} \\ 1/\sqrt{11} & 1/\sqrt{6} & 7/\sqrt{66} \end{bmatrix}$$

is an orthonormal matrix.

13. Week 13, Reading Week

14. Week 14, Orthogonal projection, the Gram-Schmidt process, and least-squares problems

14.1. Week 14, Lecture 1, Nov. 27, Orthogonal Projection.

Goals:

(1) Orthogonal projection
(2) The Gram-Schmidt process
(3) Least-squares problems

Example 14.1. *Let* $\{u_1, \ldots, u_5\}$ *be an orthogonal basis for* $\mathbb{R}^5$ *and let*

$$y = c_1u_1 + \ldots + c_5u_5.$$

Consider the subspace $W = span\{u_1, u_2\}$*, and write* y *as the sum of a vector* z_1 *in* W *and a vector* z_2 *in* $W^{\perp}$*.*

Solution. Write

$$y = \underbrace{c_1u_1 + c_2u_2}_{z_1} + \underbrace{c_3u_3 + c_4u_4 + c_5u_5}_{z_2}$$

where $z_1 = c_1u_1 + c_2u_2$ is in $span\{u_1, u_2\} = W$ and $z_2 = c_3u_3 + c_4u_4 + c_5u_5$ is in $span\{u_3, u_4, u_5\}$.

To show that z_2 is in $W^{\perp}$ it is enough to show that $z_2.u_i = 0$, for $i = 1$ *and* $i = 2$.

$$z_2.u_1 = (c_3u_3 + c_4u_4 + c_5u_5).u_1$$

$$= c_3u_3.u_1 + c_4u_4.u_1 + c_5u_5.u_1 = 0$$

because $\{u_1, \ldots, u_5\}$ is an orthogonal set. Similarly $z_2.u_2 = 0$. Therefore $z_2 \in W^{\perp}$.

Theorem 14.2. *(The Orthogonal Decomposition Theorem)*

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

$$y = \widehat{y} + z \qquad (1)$$

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

$$\widehat{y} = \frac{y.u_1}{u_1.u_1}u_1 + \ldots + \frac{y.u_p}{u_p.u_p}u_p$$

and $z = y - \widehat{y}$*.*

Definition. *The vector* $\widehat{y}$ *in (1) is called* **the orthogonal projection of** y **onto** W*, and it sometimes denoted by* $\mathbf{proj}_W y$*.*

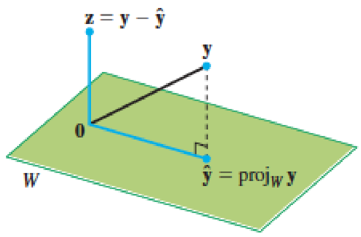


FIGURE 2 The orthogonal projection of **y** onto W.

Example 14.3. *Let* $u_1 = \begin{bmatrix} 2 \\ 5 \\ -1 \end{bmatrix}$, $u_2 = \begin{bmatrix} -2 \\ 1 \\ 1 \end{bmatrix}$, *and* $y = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$. *Observe that* $\{u_1, u_2\}$ *is an orthogonal basis for* $W = span\{u_1, u_2\}$. *Write* y *as the sum of a vector in* W *and a vector orthogonal to* W.

Solution. The orthogonal projection of y onto W is

$$\widehat{y} = \frac{y.u_1}{u_1.u_1}u_1 + \frac{y.u_2}{u_2.u_2}u_2$$

$$= 9/30 \begin{bmatrix} 2 \\ 5 \\ -1 \end{bmatrix} + 3/6 \begin{bmatrix} -2 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} -2/5 \\ 2 \\ 1/5 \end{bmatrix}$$

Also

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

By previous theorem $y - \widehat{y}$ is in $W^\perp$. And

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

• A Geometric Interpretation of the Orthogonal Projection

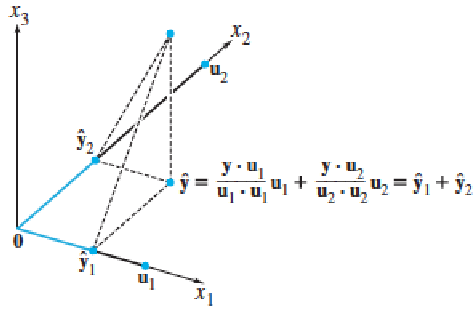


FIGURE 3 The orthogonal projection of **y** is the sum of its projections onto one-dimensional subspaces that are mutually orthogonal.

• Properties of Orthogonal Projections

Proposition 14.4. *If y is in $W = span\{u_1, \ldots, u_p\}$, then $proj_W y = y$.*

Theorem 14.5. *(The Best Approximation Theorem) Let W be a subspace of $\mathbb{R}^n$, let y be any vector in $\mathbb{R}^n$, and let $\widehat{y}$ be the orthogonal projection of y onto W. Then $\widehat{y}$ is the closest point in W to y, in the sense that*

$$\|y - \widehat{y}\| \leq \|y - v\|$$

for all v in W distinct from $\widehat{y}$.

Definition. *The vector $\widehat{y}$ is called* **the best approximation to** y **by elements of** W**.**

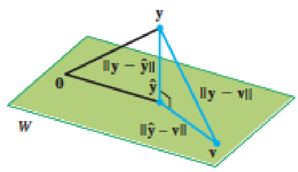


FIGURE 4 The orthogonal projection of **y** onto W is the closest point in W to **y**.

Example 14.6. *If* $u_1 = \begin{bmatrix} 2 \\ 5 \\ -1 \end{bmatrix}$, $u_2 = \begin{bmatrix} -2 \\ 1 \\ 1 \end{bmatrix}$, $y = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$ *and* $W = span\{u_1, u_2\}$. *Find the closest point in W to y.*

Solution. By the theorem the point is

$$\widehat{y} = \frac{y.u_1}{u_1.u_1}u_1 + \frac{y.u_2}{u_2.u_2}u_2 = \begin{bmatrix} -2/5 \\ 2 \\ 1/5 \end{bmatrix}$$

(we already computed $\widehat{y}$ in one of the examples.)

Example 14.7. *The* **distance from a point** $y \in \mathbb{R}^n$ **to a subspace** W *is defined as the distance from* y *to the nearest point in* W. *Find the distance from* y *to* $W = span\{u_1, u_2\}$, *where*

$$y = \begin{bmatrix} -1 \\ -5 \\ 10 \end{bmatrix}, \quad u_1 = \begin{bmatrix} 5 \\ -2 \\ 1 \end{bmatrix}, \quad u_2 = \begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix}.$$

Solution. By the best approximation theorem, the distance from y to W is $\|y - \widehat{y}\|$, where $\widehat{y} = \mathbf{proj}_W y$. Since $\{u_1, u_2\}$ is an orthogonal basis for W,

$$\widehat{y} = 15/30u_1 + -21/6u_2 = 1/2 \begin{bmatrix} 5 \\ -2 \\ 1 \end{bmatrix} - 7/2 \begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix} = \begin{bmatrix} -1 \\ -8 \\ 4 \end{bmatrix}$$

$$y - \widehat{y} = \begin{bmatrix} -1 \\ -5 \\ 10 \end{bmatrix} - \begin{bmatrix} -1 \\ -8 \\ 4 \end{bmatrix} = \begin{bmatrix} 0 \\ 3 \\ 6 \end{bmatrix}$$

$$\|y - \widehat{y}\| = \sqrt{3^2 + 6^2} = \sqrt{45}.$$

Therefore, the distance from y to W is $\sqrt{45} = 3\sqrt{5}$.

Theorem 14.8. *If* $\{u_1, \ldots, u_5\}$ *is an orthogonal basis for a subspace* W *of* $\mathbb{R}^n$, *then*

$$\mathbf{proj}_W y = (y.u_1)u_1 + (y.u_2)u_2 + \ldots + (y.u_p)u_p$$

if $U = [u_1 u_2 \ldots u_p]$, *then*

$$\mathbf{proj}_W y = UU^T y \quad \textit{for all } y \textit{ in } \mathbb{R}^n.$$

14.2. **Week 14, Lecture 2, Nov. 29, The Gram-Schmidt process.**

Example 14.9. *Let* $W = span\{x_1, x_2\}$, *where* $x_1 = \begin{bmatrix} 3 \\ 6 \\ 0 \end{bmatrix}$ *and* $x_2 = \begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix}$. *Construct an orthogonal basis* $\{v_1, v_2\}$ *for* W.

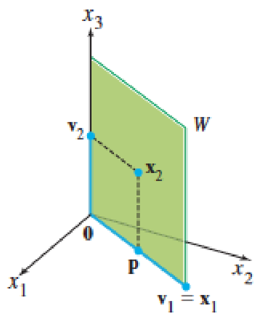


FIGURE 1
Construction of an orthogonal basis $\{\mathbf{v}_1, \mathbf{v}_2\}$.

Solution. Let $v_1 = x_1$. Let p be orthogonal projection of x_2 onto x_1 , i.e., $p = \frac{x_1 . x_2}{x_1 . x_1} x_1$. We have that

$$v_2 = x_2 - \frac{x_1 . x_2}{x_1 . x_1} x_1 = \begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix} - 15/45 \begin{bmatrix} 3 \\ 6 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 2 \end{bmatrix}.$$

Then $\{v_1, v_2\}$ is an orthogonal set of non-zero vectors in W. Since $dim\ W = 2$, then set $\{v_1, v_2\}$ is a basis for W.

Example 14.10. *Let* $x_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}$, $x_2 = \begin{bmatrix} 0 \\ 1 \\ 1 \\ 1 \end{bmatrix}$, *and* $x_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 1 \end{bmatrix}$. *Then* $\{x_1, x_2, x_3\}$ *is clearly linearly independent and thus is a basis for* W. *Construct an orthogonal basis for* W.

Step1. *Let* $v_1 = x_1$ *and* $W_1 = span\{x_1\} = span\{v_1\}$.

Step2. $v_2 = x_2 - \mathbf{proj}_{w_1} x_2$

$$= x_2 - \frac{x_2 . v_1}{v_1 . v_1} x_1$$

$$= \begin{bmatrix} 0 \\ 1 \\ 1 \\ 1 \end{bmatrix} - 3/4 \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} -3/4 \\ 1/4 \\ 1/4 \\ 1/4 \end{bmatrix}.$$

Let $W_2 = span\{v_1, v_2\}$*. Then* $\{v_1, v_2\}$ *is an orthogonal basis for* $W_2 = span\{v_1, v_2\} = span\{x_1, x_2\}$*.*

Step3. $v_3 = x_3 - \mathbf{proj}_{W_2} x_3$

$$\mathbf{proj}_{W_2} x_3 = \frac{x_3.v_1}{v_1.v_1} v_1 + \frac{x_3.v_2}{v_2.v_2} v_2$$

$$= 1/2 \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} + 2/3 \begin{bmatrix} -3/4 \\ 1/4 \\ 1/4 \\ 1/4 \end{bmatrix} = \begin{bmatrix} 0 \\ 2/3 \\ 2/3 \\ 2/3 \end{bmatrix}$$

Then $v_3 = x_3 - \mathbf{proj}_{w_2} x_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 1 \end{bmatrix} - \begin{bmatrix} 0 \\ 2/3 \\ 2/3 \\ 2/3 \end{bmatrix} = \begin{bmatrix} 0 \\ -2/3 \\ 1/3 \\ 1/3 \end{bmatrix}.$

So $\{v_1, v_2, v_3\}$ *is an orthogonal basis for* W*.*

Theorem 14.11. *(The Gram-Schmidt process) Given a basis* $\{x_1, \dots, x_p\}$ *for non-zero subspace* W *of* $\mathbb{R}^n$*, define*

$v_1 = x_1$

$v_2 = x_2 - \frac{x_2.v_1}{v_1.v_1} v_1$

$v_3 = x_3 - \frac{x_3.v_1}{v_1.v_1} v_1 - \frac{x_3.v_2}{v_2.v_2} v_2$

$\vdots$

$v_p = x_p - \frac{x_p.v_1}{v_1.v_1} v_1 - \frac{x_p.v_2}{v_2.v_2} v_2 - \dots - \frac{x_p.v_{p-1}}{v_{p-1}.v_{p-1}} v_{p-1}$

Then $\{v_1, \dots, v_p\}$ *is an orthogonal basis for* W*. In addition* $span\{v_1, \dots, v_k\} = span\{x_1, \dots, x_k\}$ *for* $1 \leq k \leq p$*.*

Theorem 14.12. *(The QR factorization) If* A *is an* $m \times n$ *matrix with linearly independent columns, then* A *can be factored as* $A = QR$*, where* Q *is an* $m \times n$ *matrix whose columns from an orthogonal basis for* $Col\ A$ *and* R *is an* $n \times n$ *upper triangular invertible matrix with positive entries on its diagonal.*

14.3. **Week 14, Lecture 3, Dec. 1, Least squares problems.**

Sometimes $Ax = b$ does not have a solution. However, we can find the vector $\widehat{x}$ such that $A\widehat{x}$ is the best approximation to b.

Definition. *If A is $m \times n$ and b is in $\mathbb{R}^m$, a* **least-squares solution** *of $Ax = b$ is an $\widehat{x}$ in $\mathbb{R}^n$ such that*

$$\|b - A\widehat{x}\| \leq \|b - Ax\|$$

for all x in $\mathbb{R}^n$.

• Goal: Finding the set of least-squares solution of $Ax = b$.

Theorem 14.13. *(Best Approximation Theorem): Let W be a subspace of $\mathbb{R}^n$, let y be any vector in $\mathbb{R}^n$, and let $\widehat{y}$ be the orthogonal projection of y onto W. Then $\widehat{y}$ is the closest point in W to y, in the sense that*

$$\|y - \widehat{y}\|\langle\|y - v\|$$

for all v in W distinct from $\widehat{y}$.

• **Solution of the general least-squares problem:**
We apply the theorem above to find the set of least-squares solution of $Ax = b$. Consider $Col\ A$. Let

$$\widehat{b} = \mathbf{proj}_{Col\ A} b$$

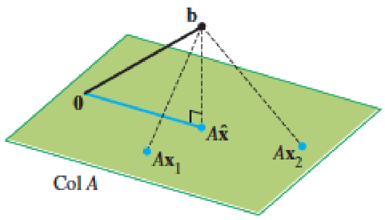


FIGURE 1 The vector **b** is closer to $A\hat{\mathbf{x}}$ than to $A\mathbf{x}$ for other **x**.

Since $\widehat{b} \in Col\ A$, there is $\widehat{x}$ such that

$$A\widehat{x} = \widehat{b} \qquad (1)$$

Note that $\widehat{b}$ is the closest point in $Col\ A$ to b. Therefore, a vector $\widehat{x}$ is a least- squares solution if and only if $\widehat{x}$ satisfies $A\widehat{x} = \widehat{b}$. We have by the Orthogonal Decomposition Theorem that $b - \widehat{b}$ is orthogonal to $Col\ A$. So $b - \widehat{b}$ is orthogonal to each column A_j of A. Therefore,

$$\begin{aligned} 0 = A_j.(b - \widehat{b}) &= A_j.(b - A\widehat{x}) \\ &= A_j^T(b - A\widehat{x}) = 0 \\ &\Rightarrow A^T(b - A\widehat{x}) = 0 \\ &\Rightarrow A^T b = A^T A\widehat{x}. \end{aligned}$$

So the set of least squares solutions of $Ax = b$ is the same as all $\widehat{x}$ such that $A^T b = A^T A\widehat{x}$. So we have the following theorem.

Theorem 14.14. *The set of least-squares solutions of $Ax = b$ coincides with the nonempty set of solution of the* **normal equations** $A^T Ax = A^T b$.

Theorem 14.15. *Let A be an $m \times n$ matrix. The following statements are logically equivalent:*

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

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

$$\widehat{x} = (A^T A)^{-1} A^T b.$$

Example 14.16. *Find a least-squares solution of the inconsistent system $Ax = b$ for*

$$A = \begin{bmatrix} 4 & 0 \\ 0 & 2 \\ 1 & 1 \end{bmatrix} \quad \textit{and} \quad b = \begin{bmatrix} 2 \\ 0 \\ 11 \end{bmatrix}.$$

Solution. Example 1 page 364 of the textbook.

Example 14.17. *Find a least-squares solution of $Ax = b$ for*

$$A = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 \end{bmatrix} \quad \textit{and} \quad b = \begin{bmatrix} -3 \\ -1 \\ 0 \\ 2 \\ 5 \\ 1 \end{bmatrix}.$$

Solution. Example 2 page 364 of the textbook.

15. Week 15, Inner Product Space

15.1. **Week 15, Lecture 1, Dec. 4, Inner product space.**

Goals:

(1) Inner product space
(2) Second test preparation
(3) Second test

Definition. *An inner product on a vector space V is a function*

$$\langle ., . \rangle : V \times V \longrightarrow \mathbb{R}$$

satisfying the following axioms:
1. $\langle u, v\rangle = \langle v, u\rangle$
2. $\langle u+v, w\rangle = \langle u, w\rangle + \langle v, w\rangle$
3. $\langle cu, v\rangle = c\langle u, v\rangle$
4. $\langle u, u\rangle \geq 0$ *and* $\langle u, u\rangle = 0$ *if and only if* $u = 0$.
A vector space with an inner product is called an inner product space.

Example 15.1. *Show that $\mathbb{R}^2$ with the following function*

$$\langle \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}, \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} \rangle = 4u_1v_1 + 5u_2v_2$$

is an inner product space.

Solution. We know that $\mathbb{R}^2$ is a vector space, so we only need to show that the function is an inner product, i.e., checking that the axioms satisfy.

(1) $\langle \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}, \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} \rangle = 4u_1v_1 + 5u_2v_2 = 4v_1u_1 + 5v_2u_2 = \langle \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}, \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} \rangle$

(2) Let $w = \begin{bmatrix} w_1 \\ w_2 \end{bmatrix}$ be another element in $\mathbb{R}^2$. Then

$$\langle \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} + \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}, \begin{bmatrix} w_1 \\ w_2 \end{bmatrix} \rangle = \langle \begin{bmatrix} u_1+v_1 \\ u_2+v_2 \end{bmatrix}, \begin{bmatrix} w_1 \\ w_2 \end{bmatrix} \rangle =$$

$$\begin{aligned} 4(u_1+v_1)w_1 + 5(u_2+v_2)w_2 &= 4u_1w_1 + 4v_1w_1 + 5u_2w_2 + 5v_2w_2 \\ &= (4u_1w_1 + 5u_2w_2) + (4v_1w_1 + 5v_2w_2) \\ &= \langle \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}, \begin{bmatrix} w_1 \\ w_2 \end{bmatrix} \rangle + \langle \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}, \begin{bmatrix} w_1 \\ w_2 \end{bmatrix} \rangle \end{aligned}$$

(3) $\langle c\begin{bmatrix} u_1 \\ u_2 \end{bmatrix}, \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} \rangle = \langle \begin{bmatrix} cu_1 \\ cu_2 \end{bmatrix}, \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} \rangle$

$$= 4cu_1v_1 + 5cu_2v_2 = c(4u_1v_1 + 5u_2v_2) = c\langle \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}, \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} \rangle.$$

(4) $\langle \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}, \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} \rangle = 4u_1^2 + 5u_2^2 \geq 0$

and also note that if $\langle\begin{bmatrix} u_1 \\ u_2 \end{bmatrix}, \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}\rangle = 4u_1^2 + 5u_2^2 = 0$ then $u_1 = 0$ and $u_2 = 0$. Therefore, $\begin{bmatrix} u_1 \\ u_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$.

Example 15.2. *Let $t_0, \dots, t_n$ be distinct real numbers. For p and q in P_n, define*

$$\langle p, q\rangle = p(t_0)q(t_0) + p(t_1)q(t_1) + \dots + p(t_n)q(t_n).$$

Solution. Axioms 1-3 are readily checked. For axiom 4,

$$\langle p, p\rangle = [p(t_0)]^2 + \dots + [p(t_n)]^2 = 0.$$

So if $[p(t_0)]^2 + \dots + [p(t_n)]^2 = 0$ we must have $p(t_0) = 0, \dots, p(t_n) = 0$. It means $t_0, \dots, t_n$ are roots for p. Therefore, p has $n+1$ roots, which is impossible if $p \neq 0$ since any non-zero polynomial of degree n has at most n roots.

• Length, Distance, and Orthogonality

Definition. *Let V be an inner product space and u and $v \in V$. Then we define*

(1) *the* **length or norm** *of a vector to be the scalar*

$$\|v\| = \sqrt{\langle v.v\rangle}$$

(2) *A* **unit** *vector is one whose length is* 1.
(3) *The* **distance** *between u and v is $\|u - v\| = \sqrt{\langle u - v, u - v\rangle}$.*
(4) *Two vectors u and v are said to be* **orthogonal** *if and only if $\langle u, v\rangle = 0$.*

Example 15.3. *Let $\mathbb{P}_2$ have the inner product*

$$\langle p, q\rangle = p(0)q(0) + p(1/2)q(1/2) + p(1)q(1).$$

Compute the length of the following vectors $p(t) = 12t^2$ and $q(t) = 2t - 1$.

Solution. Note that $\|p\| = \sqrt{\langle p, p\rangle}$. We have

$$\langle p, p\rangle = [p(0)]^2 + [p(1/2)]^2 + [p(1)]^2 = 0 + 3^2 + 12^2 = 153.$$

Therefore, $\|p\| = \sqrt{153}$. Also, $\|q\| = \sqrt{2}$ (check it).

• The Gram-Schmidt Process:

Theorem 15.4. *(The Gram-Schmidt process) Given a basis $\{x_1, \dots, x_p\}$ for non-zero subspace W of $\mathbb{R}^n$, define*

$v_1 = x_1$

$v_2 = x_2 - \frac{x_2.v_1}{v_1.v_1}v_1$

$v_3 = x_3 - \frac{x_3.v_1}{v_1.v_1}v_1 - \frac{x_3.v_2}{v_2.v_2}v_2$

$\vdots$

$v_p = x_p - \frac{x_p.v_1}{v_1.v_1}v_1 - \frac{x_p.v_2}{v_2.v_2}v_2 - \ldots - \frac{x_p.v_{p-1}}{v_{p-1}.v_{p-1}}v_{p-1}$
Then $\{v_1, \ldots, v_p\}$ *is an orthogonal basis for* W*. In addition span*$\{v_1, \ldots, v_k\}$ $=$ *span*$\{x_1, \ldots, x_k\}$ *for* $1 \leq k \leq p$*.*

Theorem 15.5. *(The Gram-Schmidt process for an inner product space) Given a basis* $\{x_1, \ldots, x_p\}$ *for non-zero subspace* W *of an inner product space* V*, define*

$v_1 = x_1$

$v_2 = x_2 - \frac{\langle x_2, v_1\rangle}{\langle v_1, v_1\rangle}v_1$

$v_3 = x_3 - \frac{\langle x_3, v_1\rangle}{\langle v_1, v_1\rangle}v_1 - \frac{\langle x_3, v_2\rangle}{\langle v_2, v_2\rangle}v_2$

$\vdots$

$v_p = x_p - \frac{\langle x_p, v_1\rangle}{\langle v_1, v_1\rangle}v_1 - \frac{\langle x_p, v_2\rangle}{\langle v_2, v_2\rangle}v_2 - \ldots - \frac{\langle x_p, v_{p-1}\rangle}{\langle v_{p-1}, v_{p-1}\rangle}v_{p-1}$
Then $\{v_1, \ldots, v_p\}$ *is an orthogonal basis for* W*. In addition span*$\{v_1, \ldots, v_k\}$ $=$ *span*$\{x_1, \ldots, x_k\}$ *for* $1 \leq k \leq p$*.*

Example 15.6. *Define the following inner product for* $\mathbb{P}_4$*,*

$$\langle p, q\rangle = p(-2)q(-2) + p(-1)q(-1) + p(0)q(0) + p(1)q(1) + p(2)q(2).$$

Let $\mathbb{P}_2$ *be the subspace of* $\mathbb{P}_4$ *with the basis* $\{p_1, p_2, p_3\}$*, where* $p_1 = 1, p_2 = t, p_3 = t^2$*. Produce an orthogonal basis for* $\mathbb{P}_2$ *by applying the Gram-Schmidt Process.*

Solution.

$f_1 = P - 1 = 1$

$f_2 = p_2 - \frac{\langle p_2, f_1\rangle}{\langle f_1, f_1\rangle}f_1$

$f_3 = p_3 - \frac{\langle p_3, f_1\rangle}{\langle f_1, f_1\rangle}f_1 - \frac{\langle p_3, f_2\rangle}{\langle f_2, f_2\rangle}f_2$

$$\langle t, 1\rangle = (-2) \times 1 + (-1) \times 1 + 0 \times 1 + 1 \times 1 + 2 \times 1 = 0.$$
$$\langle f_1, f_1\rangle = \langle 1, 1\rangle = 1 \times 1 + 1 \times 1 + 1 \times 1 + 1 \times 1 + 1 \times 1 = 5$$

Therefore, $f_2 = t - \frac{0}{5} = t$.

$$\langle p_3, f_1\rangle = \langle t^2, 1\rangle = (-2)^2 \times 1 + (-1)^2 \times 1 + 0^2 \times 1 + 1^2 \times 1 + 2^2 \times 1 = 10.$$
$$\langle p_3, f_2\rangle = \langle t^2, t\rangle = (-2)^2 \times -2 + (-1)^2 \times (-1) + 0^2 \times 0 + 1^2 \times 1 + 2^2 \times 2 = 0.$$
$$\langle f_2, f_2\rangle = \langle t, t\rangle = (-2)^2 + (-1)^2 + 0^2 + 1^2 + 2^2 = 10.$$

Therefore, $f_3 = t^2 - \frac{10}{5}1 - \frac{0}{10}t = t^2 - 2$. Therefore,

$$\{1, t, t^2 - 2\}$$

is an orthogonal basis for $\mathbb{P}_2$ (check orthogonality).

References

[1] Linear Algebra and Its Applications, 5th Edition, by Lay, Lay, McDonald.