MATH240 Summer 2022

Haoran Li

Contents

1 Lectures

1.1 Lecture 1 - 05/31/2022

Example 1.1.1 (Example in the ancient Chinese mathematics book titled *The Nine Chapters on the Mathematical Art*)**.** In a cage there are chicken and rabbits. The total number of heads is 10 and the total number of legs is 26. Question: how many rabbits and chicken are there?

Definition 1.1.2. A linear equation in the variables $x_1, x_2, x_3, \cdots, x_n$ is an equation that can be written in the form

$$a_1x_1 + a_2x_2 + a_3x_3 + \cdots + a_nx_n = b \tag{1.1.1}$$

where b and the coefficients $a_1, a_2, a_3, \cdots, a_n$ are real or complex numbers, usually known in advance.

Remark. We adopt the following notations

- Natural numbers $\mathbb{N} = \{0, 1, 2, 3, \cdots\}$.
- Integers $\mathbb{Z} = \{\cdots, -2, -1, 0, 1, 2, \cdots\}$.
- Rational numbers $\mathbb{Q}$ is the set of fractions.
- Real numbers $\mathbb{R}$ contains irrational numbers (like $\sqrt{2}$, $\sqrt[3]{3}$) and transcendental numbers (like π, e)
- Complex numbers $\mathbb{C} = \{a+bi | a, b \in \mathbb{R}\}$, here $\in$ means belong to, $i = \sqrt{-1}$ is the imaginary number such that $i^2 = -1$.
- $\mathbb{N} \subset \mathbb{Z} \subset \mathbb{Q} \subset \mathbb{R} \subset \mathbb{C}$.
- $\mathbb{R}^n = \{(r_1, r_2, \cdots, r_n) | r_1, r_2, \cdots, r_n \in \mathbb{R}\}$ is the set of all n-tuples of real numbers. Geometrically, it is the n-dimensional Euclidean space. For example
  - $\mathbb{R}^1 = \mathbb{R}$ is a line
  - $\mathbb{R}^2$ is a plane
  - $\mathbb{R}^3$ is our usual three dimensional space
  - $\vdots$

Example 1.1.3 (Examples and non-examples of linear equations)**.**

- $x_1 + \frac{1}{2}x_2 = 2$, ✓
- $\pi(x_1 + x_2) - 9.9x_3 = e$ ✓. Because we can expand to get $\pi x_1 + \pi x_2 - 9.9x_3 = e$ in which case $a_1 = \pi, a_2 = \pi, a_3 = -9.9, b = e$ as in the form of (1.1.1)
- $|x_2| - 1 = 0$, ✗
- $x_1 + x_2^2 = 9$, ✗
- $\sqrt{x_1} + \sqrt{x_2} = 1$, ✗

Question. Why do we use the word “linear”? What do they mean geometrically?

Answer. A linear equation uniquely characterizes a hyperplane. A hyperplane is a one-dimension less subspace, for example

- In $\mathbb{R}^1$, it is a point.
- In $\mathbb{R}^2$, it is a line.
- In $\mathbb{R}^3$, it is a plane.
- In $\mathbb{R}^4$, it is a hyperplane.

- $\vdots$

Definition 1.1.4. A system of linear equations (or a linear system) is a collection of one or more linear equations involving the same variables, say $x_1, x_2, \cdots, x_n$.

Example 1.1.5. In Example 1.1.1, if we assume the number of chickens and rabbits are x_1 and x_2, then we can use the information we know about heads and legs to get the linear system

$$\begin{cases} x_1 + x_2 = 10 \\ 2x_1 + 4x_2 = 26 \end{cases} \tag{1.1.2}$$

Definition 1.1.6. A solution of the system is $\begin{cases} x_1 = s_1 \\ x_2 = s_2 \\ x_3 = s_3 \\ \vdots \\ x_n = s_n \end{cases}$ which makes each equation true.

The set of all possible solutions is called the solution set of the linear system. To solve a linear system is to find all its solutions.

Question. How many solutions could a linear system have?

Example 1.1.7. There are three different possibilities

- In Example (1.1.2), there is a unique solution $\begin{cases} x_1 = 7 \\ x_2 = 3 \end{cases}$
- $\begin{cases} x_1 + 2x_2 = 3 \\ 2x_1 + 4x_2 = 6 \end{cases}$ has uncountably infinitely many solutions since the two equations describes the same line. (the second equation is just twice of the first one)
- $\begin{cases} x_1 + 2x_2 = 3 \\ -2x_1 - 4x_2 = 6 \end{cases}$ has no solutions since these two equations defines two parallel, non-intersecting lines.

Answer. A linear system consists of two equations with two variables has

- infinitely many solutions if these two lines are overlapping.
- a unique solution if these two lines intersect.
- no solutions if these two lines are parallel and non-intersecting.

Question. How should we solve a linear system?

Answer. We use Gaussian elimination which is illustrated in the following example

Example 1.1.8. Let's solve the linear system $\begin{cases} 2x_1 + 2x_2 = 20 & ① \\ 2x_1 + 4x_2 = 26 & ② \end{cases}$ in Example 1.1.5

Step 1. Multiply ① by 2, we get $\begin{cases} 2x_1 + 2x_2 = 20 & ③ \\ 2x_1 + 4x_2 = 26 & ② \end{cases}$

Step 2. Replace ① by ③ − ②, we get $\begin{cases} -2x_2 = -6 & ④ \\ 2x_1 + 4x_2 = 26 & ② \end{cases}$

Step 3. Divide ④ by −2, we get $\begin{cases} x_2 = 3 & ⑤ \\ 2x_1 + 4x_2 = 26 & ② \end{cases}$

Step 4. Replace ② by ② − 4⑤, we get $\begin{cases} \quad\quad x_2 = 3 & ⑤ \\ 2x_1 \quad\quad = 14 & ⑥ \end{cases}$

Step 5. Divide ⑥ by 2, we get $\begin{cases} \quad\quad x_2 = 3 & ⑤ \\ x_1 \quad\quad = 7 & ⑦ \end{cases}$

Step 6. Interchange ⑤ and ⑦, we finally have the solution $\begin{cases} x_1 \quad\quad = 7 & ⑦ \\ \quad\quad x_2 = 7 & ⑤ \end{cases}$

Definition 1.1.9. We say a linear system is consistent it has solution(s), and inconsistent if it has none

1.2 Lecture 2 - 06/01/2022

Definition 1.2.1. A m by n (or $m \times n$) matrix is a rectangular array of numbers of m rows and n columns, we use the (i, j)-th entry to mean the entry on the i-th row and j-column. A matrix is

- a zero matrix is a matrix with all entries zeros.
- a square matrix is a matrix with the same number of rows and columns, i.e. $m = n$.
- a vector if it only has one column, i.e. $n = 1$
- the identity matrix if it is a square matrix with diagonal elements 1's, and 0's otherwise. Here the diagonal are the (i, i)-th entries

Example 1.2.2.

- A general m by n matrix looks like $\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}$, here a_{ij}'s are numbers
- A general $n \times n$ square matrix looks like $\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix}$, here a_{ij}'s are numbers
- A general vector looks like $\begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix}$, here a_i's are numbers
- A zero matrix looks like $\begin{bmatrix} 0 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & 0 \end{bmatrix}$
- The n by n identity matrix looks like $\begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix}$

Definition 1.2.3. Soon we will be getting tired of writing all these equations in a linear system, instead we write down its augmented matrix, obtained by omitting x_i's, pluses, and equal signs. If we delete the last column, we will get the coefficient matrix

Example 1.2.4. For (1.1.2), its augmented matrix and coefficient matrix are

$$\begin{bmatrix} 1 & 1 & 10 \\ 2 & 4 & 26 \end{bmatrix}, \begin{bmatrix} 1 & 1 \\ 2 & 4 \end{bmatrix}$$

For $\begin{cases} x_1 - x_2 + x_3 = 1 \\ 2x_1 \qquad - x_3 = 1 \\ x_1 + x_2 + x_3 = 3 \end{cases}$, its augmented matrix and coefficient matrix are

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

In general, a linear system of m equations in n variables has a m by $(n+1)$ augmented matrix and a m by n coefficient matrix.

Definition 1.2.5. Look more carefully at Example 1.1.8, we define the following three elementary row operations

- Replacement: Replace one row by the sum of itself and a multiple of another row.
- Interchange: Interchange two rows.
- Scaling: Multiply all entries in a row by a nonzero constant.

We say matrices A, B are row equivalent ($A \sim B$) if B can obtained by applying a sequence of elementary row operations to A (or vise versa)

Example 1.2.6. Let's rewrite the process in Example 1.1.8

$$\begin{bmatrix} 1 & 1 & 10 \\ 2 & 4 & 26 \end{bmatrix} \overset{2R1}{\sim} \begin{bmatrix} 2 & 2 & 20 \\ 2 & 4 & 26 \end{bmatrix} \xrightarrow{R1 \to R1 - R2} \begin{bmatrix} 0 & -2 & -6 \\ 2 & 4 & 26 \end{bmatrix} \overset{\frac{R1}{-2}}{\sim} \begin{bmatrix} 0 & 1 & 3 \\ 2 & 4 & 26 \end{bmatrix}$$
$$\xrightarrow{R2 \to R2 - 4R1} \begin{bmatrix} 0 & 1 & 3 \\ 2 & 0 & 14 \end{bmatrix} \overset{\frac{1}{2}R2}{\sim} \begin{bmatrix} 0 & 1 & 3 \\ 1 & 0 & 7 \end{bmatrix} \xrightarrow{R1 \leftrightarrow R2} \begin{bmatrix} 1 & 0 & 7 \\ 0 & 1 & 3 \end{bmatrix}$$

Example 1.2.7. $\begin{cases} x_1 - x_2 + x_3 = 1 \\ 2x_1 \qquad - x_3 = 1 \\ x_1 + x_2 + x_3 = 3 \end{cases}$, first we write out its augmented matrix

$$\begin{bmatrix} 1 & -1 & 1 & 1 \\ 2 & 0 & -1 & 1 \\ 1 & 1 & 1 & 3 \end{bmatrix} \xrightarrow{R2 \to R2 - 2R1} \begin{bmatrix} 1 & -1 & 1 & 1 \\ 0 & 2 & -3 & -1 \\ 1 & 1 & 1 & 3 \end{bmatrix} \xrightarrow{R3 \to R3 - R1} \begin{bmatrix} 1 & -1 & 1 & 1 \\ 0 & 2 & -3 & -1 \\ 0 & 2 & 0 & 2 \end{bmatrix}$$
$$\xrightarrow{R3 \to R3 - R2} \begin{bmatrix} 1 & -1 & 1 & 1 \\ 0 & 2 & -3 & -1 \\ 0 & 0 & 3 & 3 \end{bmatrix} \overset{\frac{R3}{3}}{\sim} \begin{bmatrix} 1 & -1 & 1 & 1 \\ 0 & 2 & -3 & -1 \\ 0 & 0 & 1 & 1 \end{bmatrix} \xrightarrow[R2 \to R2 + 3R3]{R1 \to R1 - R3} \begin{bmatrix} 1 & -1 & 0 & 0 \\ 0 & 2 & 0 & 2 \\ 0 & 0 & 1 & 1 \end{bmatrix}$$
$$\overset{\frac{1}{2}R2}{\sim} \begin{bmatrix} 1 & -1 & 0 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \end{bmatrix} \xrightarrow{R1 \to R1 + R2} \begin{bmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \end{bmatrix}$$

This gives the unique solution $\begin{cases} x_1 = 1 \\ x_2 = 1 \\ x_3 = 1 \end{cases}$

Definition 1.2.8. A leading entry of a row refers to the leftmost nonzero entry (in a nonzero row). A matrix is of row echelon form, REF if it is of a "staircase shape". The leading entries of an REF matrix are called pivots, the position of pivots are called pivot positions, and the column pivots are in are called pivot columns. An REF of reduced row echelon form (RREF) if all its pivots are 1's and in each column, every entry except the pivot are 0's

$$\overset{\text{REF}}{\begin{bmatrix} \blacksquare & * & * & * & * & * & * & * \\ 0 & 0 & \blacksquare & * & * & * & * & * \\ 0 & 0 & 0 & 0 & \blacksquare & * & * & * \\ 0 & 0 & 0 & 0 & 0 & \blacksquare & * & * \end{bmatrix}}, \quad \overset{\text{RREF}}{\begin{bmatrix} 1 & * & 0 & * & 0 & 0 & * & * \\ 0 & 0 & 1 & * & 0 & 0 & * & * \\ 0 & 0 & 0 & 0 & 1 & 0 & * & * \\ 0 & 0 & 0 & 0 & 0 & 1 & * & * \end{bmatrix}}$$

Example 1.2.9. In Example 1.2.7, $\begin{bmatrix} 1 & -1 & 1 & 1 \\ 0 & 2 & -3 & -1 \\ 0 & 2 & 0 & 2 \end{bmatrix}$ is not an REF. $\begin{bmatrix} 1 & -1 & 1 & 1 \\ 0 & 2 & -3 & -1 \\ 0 & 0 & 3 & 3 \end{bmatrix}$ is an REF, but not an RREF. $\begin{bmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \end{bmatrix}$ is an RREF

Theorem 1.2.10. Every matrix is row equivalent to some REF matrix (which is not in general unique), and it is row equivalent to some unique RREF matrix.

Remark. This ensures that the pivot positions are well-defined, i.e. you won't get different pivot positions if you applied different row operations

Example 1.2.11. In Example 1.2.7, $\begin{bmatrix} 1 & -1 & 1 & 1 \\ 0 & 2 & -3 & -1 \\ 0 & 0 & 3 & 3 \end{bmatrix}$ is an REF row equivalent to the original matrix $\begin{bmatrix} 1 & -1 & 1 & 1 \\ 2 & 0 & -1 & 1 \\ 1 & 1 & 1 & 3 \end{bmatrix}$, and $\begin{bmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \end{bmatrix}$ is the unique row equivalent RREF matrix.

Remark. A linear system has a unique solution if and only if its RREF deleting the last column gives the identity matrix.

Example 1.2.12. $\begin{cases} x_1 - x_2 + x_3 = 1 \\ 2x_1 \phantom{{}-x_2} - x_3 = 1 \\ x_1 + x_2 - 2x_3 = 1 \end{cases}$, we write down its augmented matrix

$$\begin{bmatrix} 1 & -1 & 1 & 1 \\ 2 & 0 & -1 & 1 \\ 1 & 1 & -2 & 1 \end{bmatrix} \underset{\sim}{\overset{R2\to R2-2R1}{R3\to R3-R1}} \begin{bmatrix} 1 & -1 & 1 & 1 \\ 0 & 2 & -3 & -1 \\ 0 & 2 & -3 & 0 \end{bmatrix} \overset{R3\to R3-R2}{\sim} \begin{bmatrix} 1 & -1 & 1 & 1 \\ 0 & 2 & -3 & -1 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

You might notice that the last row is a bit weird which indeed is true since it represents $0x_1 + 0x_2 + 0x_3 = 0 = 1$, this is a contradiction, hence the linear system is inconsistent, i.e. it has no solutions

Remark. This only happens if and only if the last pivot column is the last column

Example 1.2.13. $\begin{cases} x_1 - x_2 + x_3 = 1 \\ 2x_1 \phantom{{}-x_2} - x_3 = 1 \end{cases}$, we write down its augmented matrix

$$\begin{aligned} \begin{bmatrix} 1 & -1 & 1 & 1 \\ 2 & 0 & -1 & 1 \end{bmatrix} &\overset{R2\to R2-2R1}{\sim} \begin{bmatrix} 1 & -1 & 1 & 1 \\ 0 & 2 & -3 & -1 \end{bmatrix} \\ \overset{\frac{R2}{2}}{\sim} \begin{bmatrix} 1 & -1 & 1 & 1 \\ 0 & 1 & -\frac{3}{2} & -\frac{1}{2} \end{bmatrix} &\overset{R1\to R1+R2}{\sim} \begin{bmatrix} 1 & 0 & -\frac{1}{2} & \frac{1}{2} \\ 0 & 1 & -\frac{3}{2} & -\frac{1}{2} \end{bmatrix} \end{aligned}$$

This gives the solution set

$$\begin{cases} x_1 \phantom{{}+x_2} - \frac{1}{2}x_3 = \frac{1}{2} \\ \phantom{x_1 +{}} x_2 - \frac{3}{2}x_3 = \frac{1}{2} \end{cases} \Rightarrow \begin{cases} x_1 = \frac{1}{2}x_3 + \frac{1}{2} \\ x_2 = \frac{3}{2}x_3 - \frac{1}{2} \end{cases} \tag{1.2.1}$$

Let's formalize the row reduction algorithm

Step 1. Begin with the leftmost nonzero column. This is a pivot column. The pivot position is at the top.

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

Step 3. Use row replacement operations to create zeros in all positions below the pivot.

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

Step 5. Beginning with the rightmost pivot and working upward and to the left, create zeros above each pivot. If a pivot is not 1, make it 1 by a scaling operation.

Here steps 1-4 are call Forward phase, and step 5 is called backward phase.

Definition 1.2.14. The variables corresponding to pivot columns in a matrix are called basic variables, the other variables are called free variables. In a solution set, basic variables are expressed in terms of free variables, and a free variable can take any value.

Example 1.2.15. In Example 1.2.13, x_1, x_2 are basic variables and x_3 is a free variable. And we formally write our solution set as $\begin{cases} x_1 = \frac{1}{2}x_3 + \frac{1}{2} \\ x_2 = \frac{3}{2}x_3 - \frac{1}{2} \\ x_3 \text{ is free} \end{cases}$

Exercise 1.2.16. Find the general solution of the system $\begin{cases} x_1 - 2x_2 - x_3 + 3x_4 = 0 \\ -2x_1 + 4x_2 + 5x_3 - 5x_4 = 3 \\ 3x_1 - 6x_2 - 4x_3 + 8x_4 = 2 \end{cases}$

Solution.

$$\begin{bmatrix} 1 & -2 & -1 & 3 & 0 \\ -2 & 4 & 5 & -5 & 3 \\ 3 & -6 & -4 & 8 & 2 \end{bmatrix} \xrightarrow[R3 \to R3 - 3R1]{R2 \to R2 + 2R1} \begin{bmatrix} 1 & -2 & -1 & 3 & 0 \\ 0 & 0 & 3 & 1 & 3 \\ 0 & 0 & -1 & -1 & 2 \end{bmatrix} \xrightarrow{(-1)\cdot R3} \begin{bmatrix} 1 & -2 & -1 & 3 & 0 \\ 0 & 0 & 3 & 1 & 3 \\ 0 & 0 & 1 & 1 & -2 \end{bmatrix}$$

$$\xrightarrow{R2 \to R2 - 3R3} \begin{bmatrix} 1 & -2 & -1 & 3 & 0 \\ 0 & 0 & 0 & -2 & 9 \\ 0 & 0 & 1 & 1 & -2 \end{bmatrix} \xrightarrow{R2 \leftrightarrow R3} \begin{bmatrix} 1 & -2 & -1 & 3 & 0 \\ 0 & 0 & 1 & 1 & -2 \\ 0 & 0 & 0 & -2 & 9 \end{bmatrix}$$

$$\overset{\frac{R3}{-2}}{\sim} \begin{bmatrix} 1 & -2 & -1 & 3 & 0 \\ 0 & 0 & 1 & 1 & -2 \\ 0 & 0 & 0 & 1 & -\frac{9}{2} \end{bmatrix} \xrightarrow[R1 \to R1 - 3R3]{R2 \to R2 - R3} \begin{bmatrix} 1 & -2 & -1 & 0 & \frac{27}{2} \\ 0 & 0 & 1 & 0 & \frac{5}{2} \\ 0 & 0 & 0 & 1 & -\frac{9}{2} \end{bmatrix} \xrightarrow{R1 \to R1 + R2} \begin{bmatrix} 1 & -2 & 0 & 0 & 16 \\ 0 & 0 & 1 & 0 & \frac{5}{2} \\ 0 & 0 & 0 & 1 & -\frac{9}{2} \end{bmatrix}$$

Write this as solution set, we get

$$\begin{cases} x_1 - 2x_2 = 16 \\ x_3 = \frac{5}{2} \\ x_4 = -\frac{9}{2} \end{cases} \Rightarrow \begin{cases} x_1 = 2x_2 + 16 \\ x_2 \text{ is free} \\ x_3 = \frac{5}{2} \\ x_4 = -\frac{9}{2} \end{cases}$$

□

Theorem 1.2.17. Consider the RREF matrix $\begin{bmatrix} A & \mathbf{b} \end{bmatrix}$ equivalent to the augmented matrix of a linear system, then the linear system has

- no solutions $\iff$ the last column of M is a pivot column, i.e. has a pivot
- infinitely many solutions $\iff$ the last column is not a pivot column and there are free variables
- a unique solution $\iff$ M deleting the last column is an identity matrix

1.3 Lecture 3 - 06/02/2022

Definition 1.3.1. Recall a vector is a matrix with one column, the zero vector is a vector with all entries zero. For scalar (i.e. a number) c, and vectors $\mathbf{a} = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix}$, $\mathbf{b} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}$, we could define

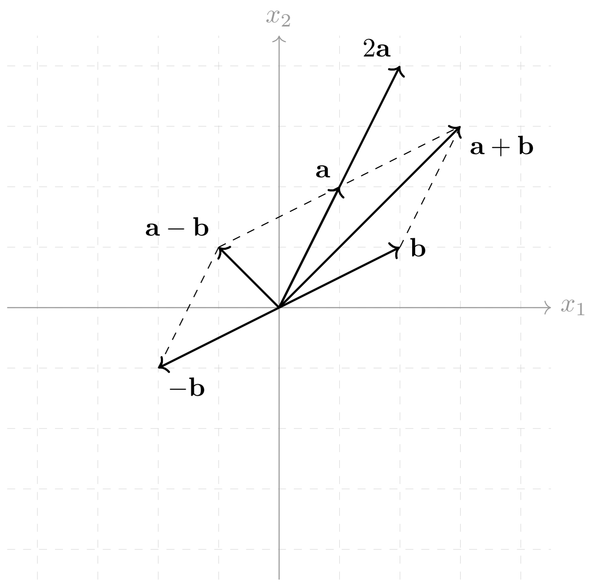
- Addition $\mathbf{a} + \mathbf{b} = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix} + \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix} = \begin{bmatrix} a_1 + b_1 \\ a_2 + b_2 \\ \vdots \\ a_n + b_n \end{bmatrix}$
- Scalar multiplication $c\mathbf{a} = c \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix} = \begin{bmatrix} ca_1 \\ ca_2 \\ \vdots \\ ca_n \end{bmatrix}$
- Subtraction $\mathbf{a} - \mathbf{b} = \mathbf{a} + (-1)\mathbf{b} = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix} - \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix} = \begin{bmatrix} a_1 - b_1 \\ a_2 - b_2 \\ \vdots \\ a_n - b_n \end{bmatrix}$

A linear combination of vectors $\mathbf{a}_1, \mathbf{a}_2, \cdots, \mathbf{a}_n$ means a sum $c_1\mathbf{a}_1 + c_2\mathbf{a}_2 + \cdots + c_n\mathbf{a}_n$ for some scalars $c_1, \cdots, c_n$

Remark. In handwritings, we use $\vec{v}$ or $\overrightarrow{v}$ to denote a vector, while in printing materials we often use $\mathbf{v}$

Example 1.3.2. Let $\mathbf{a} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$, $\mathbf{b} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$, then

- $\mathbf{a} + \mathbf{b} = \begin{bmatrix} 3 \\ 3 \end{bmatrix}$
- $2\mathbf{a} = \begin{bmatrix} 2 \\ 4 \end{bmatrix}$
- $-\mathbf{b} = (-1)\mathbf{b} = \begin{bmatrix} -2 \\ -1 \end{bmatrix}$
- $\mathbf{a} - \mathbf{b} = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$



We can now rewrite a linear system as a vector equation

Example 1.3.3. (1.1.2) can be written as $x_1\mathbf{a}_1 + x_2\mathbf{a}_2 = \mathbf{b}$

$$x_1\begin{bmatrix}1\\2\end{bmatrix} + x_2\begin{bmatrix}1\\4\end{bmatrix} = \begin{bmatrix}x_1\\2x_1\end{bmatrix} + \begin{bmatrix}x_2\\4x_2\end{bmatrix} = \begin{bmatrix}x_1+x_2\\2x_1+4x_2\end{bmatrix} = \begin{bmatrix}10\\26\end{bmatrix}$$

Here $\mathbf{a}_1 = \begin{bmatrix}1\\2\end{bmatrix}$, $\mathbf{a}_2 = \begin{bmatrix}1\\4\end{bmatrix}$ and $\mathbf{b} = \begin{bmatrix}10\\26\end{bmatrix}$

Remark. In general, vector equation

$$x_1\mathbf{a}_1 + x_2\mathbf{a}_2 + \cdots + x_n\mathbf{a}_n = \mathbf{b} \tag{1.3.1}$$

can be thought of as a generalization of (1.1.1)

Example 1.3.4. $\begin{cases} x_1 + x_3 = 1 \\ 2x_1 + x_3 = 2 \end{cases}$ can be written as

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

Definition 1.3.5. The span of $\{\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_n\}$ is the set of all its linear combinations, which we denote $\mathrm{Span}\{\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_n\}$. We see that (1.3.1) has solution(s) $\iff$ $\mathbf{b}$ is in $\mathrm{Span}\{\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_n\}$ (i.e. $\mathbf{b}$ is a linear combination of $\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_n$) $\iff$ the linear system is consistent

Exercise 1.3.6. Let $\mathbf{a}_1 = \begin{bmatrix}1\\2\\1\end{bmatrix}$, $\mathbf{a}_2 = \begin{bmatrix}-1\\0\\1\end{bmatrix}$, $\mathbf{a}_3 = \begin{bmatrix}1\\-1\\-2\end{bmatrix}$, $\mathbf{b} = \begin{bmatrix}1\\1\\1\end{bmatrix}$. Is $\mathbf{b}$ in $\mathrm{Span}\{\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3\}$

Solution. This is equivalent of asking if the following linear system is consistent $\begin{cases} x_1 - x_2 + x_3 = 1 \\ 2x_1 \quad\quad - x_3 = 1 \\ x_1 + x_2 - 2x_3 = 1 \end{cases}$, we find its RREF

$$\begin{bmatrix}1 & -1 & 1 & 1\\2 & 0 & -1 & 1\\1 & 1 & -2 & 1\end{bmatrix} \xrightarrow[R3\to R3-R1]{R2\to R2-2R1} \begin{bmatrix}1 & -1 & 1 & 1\\0 & 2 & -3 & -1\\0 & 2 & -3 & 0\end{bmatrix} \xrightarrow{R3\to R3-R2} \begin{bmatrix}1 & -1 & 1 & 1\\0 & 2 & -3 & -1\\0 & 0 & 0 & 1\end{bmatrix}$$

Since there is a pivot in the last column, by Theorem 1.2.17, the linear system is inconsistent, hence $\mathbf{b}$ is not in $\mathrm{Span}\{\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3\}$ □

1.4 Lecture 4 - 06/03/2022

Definition 1.4.1. Let's use $M_{m\times n}(\mathbb{R})$ to denote the set of all (real-valued) matrices of dimension m by n

Definition 1.4.2. Suppose A, B are $m \times n$ matrices, c is a scalar (i.e. a number), then we can define

- Addition $\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} + \begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1n} \\ b_{21} & b_{22} & \cdots & b_{2n} \\ \vdots & \vdots & & \vdots \\ b_{m1} & b_{m2} & \cdots & b_{mn} \end{bmatrix} = \begin{bmatrix} a_{11}+b_{11} & a_{12}+b_{12} & \cdots & a_{1n}+b_{1n} \\ a_{21}+b_{21} & a_{22}+b_{22} & \cdots & a_{2n}+b_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1}+b_{m1} & a_{m2}+b_{m2} & \cdots & a_{mn}+b_{mn} \end{bmatrix}$
- Scalar multiplication $c\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} = \begin{bmatrix} ca_{11} & ca_{12} & \cdots & ca_{1n} \\ ca_{21} & ca_{22} & \cdots & ca_{2n} \\ \vdots & \vdots & & \vdots \\ ca_{m1} & ca_{m2} & \cdots & ca_{mn} \end{bmatrix}$

Remark. Note that in the case of vectors (i.e. $n = 1$), these recovers addition and scalar multiplication of vectors.

Example 1.4.3. $\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} + \begin{bmatrix} 4 & 3 \\ 2 & 1 \end{bmatrix} = \begin{bmatrix} 5 & 5 \\ 5 & 5 \end{bmatrix}, 2\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} = \begin{bmatrix} 2 & 4 \\ 6 & 8 \end{bmatrix}$

Definition 1.4.4. Suppose A is a $m \times n$ matrix, and B is a $n \times p$ matrix, we can define matrix multiplication AB to be the $m \times p$ matrix, computed via the rule-column rule: The (i, j)-entry is to multiply the i-row and j-th column

$$\begin{bmatrix} \\ \\ a_{i1} & a_{i2} & \cdots & a_{in} \\ \\ \\ \end{bmatrix} \begin{bmatrix} & b_{1j} & \\ & b_{2j} & \\ & \vdots & \\ & b_{nj} & \end{bmatrix} = \begin{bmatrix} \\ & \blacksquare & \\ \\ \end{bmatrix}$$

Where $\blacksquare = a_{i1}b_{1j} + a_{i2}b_{2j} + \cdots + a_{in}b_{nj}$.

If A is a square matrix, then we could define matrix power A^k to be simply $\overbrace{AA\cdots A}^{k \text{ times}}$

Example 1.4.5.

- $\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{bmatrix} \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \\ b_{31} & b_{32} \end{bmatrix} = \begin{bmatrix} a_{11}b_{11} + a_{12}b_{21} + a_{13}b_{31} & a_{11}b_{12} + a_{12}b_{22} + a_{13}b_{32} \\ a_{21}b_{11} + a_{22}b_{21} + a_{23}b_{31} & a_{21}b_{12} + a_{22}b_{22} + a_{23}b_{32} \end{bmatrix}$
- $\begin{bmatrix} 1 & 2 & 2 \\ 2 & 1 & 1 \end{bmatrix} \begin{bmatrix} 3 & 1 \\ 1 & 2 \\ 2 & 3 \end{bmatrix} = \begin{bmatrix} 9 & 11 \\ 9 & 7 \end{bmatrix}$
-
$$\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 \end{bmatrix} = x_1\begin{bmatrix} a_{11} \\ a_{21} \end{bmatrix} + x_1\begin{bmatrix} a_{12} \\ a_{22} \end{bmatrix} + x_1\begin{bmatrix} a_{13} \\ a_{23} \end{bmatrix} \tag{1.4.1}$$

- $$\begin{bmatrix}0&1\\0&0\end{bmatrix}\begin{bmatrix}1&0\\0&0\end{bmatrix}=\begin{bmatrix}0&0\\0&0\end{bmatrix} \tag{1.4.2}$$

Fact 1.4.6. Suppose A, B, C, D are matrices, c is a scalar, 0 is the zero matrix, I is the identity matrix. we have the following facts

a) Matrix multiplication is not commutative, i.e. $AB \neq BA$

b) Matrix multiplication is associative, i.e. the order of multiplication doesn't matter, in other words $(AB)C = A(BC)$, so there is no confusion in writing successive multiplication $A_1A_2A_3 \cdots A_n$

c) Scalar multiplication and matrix multiplication commutes, $A(cB) = c(AB) = (cA)B$.

d) Matrix multiplication is distributive over addition, i.e. $A(B+C) = AB+AC$, $(A+B)C = AC + BC$

e) $A + 0 = 0 + A = A$, $A0 = 0A = 0$

f) $IA = AI = A$

g) Even if $A \neq 0$, $B \neq 0$, AB could still be 0, take (1.4.2) for an example

h) $AB = AC$ does not imply $B = C$

Remark. Some of the properties of matrices are really similar to that of numbers, so we dub this the name of matrix algebra

Definition 1.4.7. A is a partitioned (or block) matrix if is divided into smaller submatrix by some horizontal and vertical lines. And the submatrices are the blocks

$$\left[\begin{array}{c|c|c}A_{11}&A_{12}&A_{13}\\\hline A_{21}&A_{22}&A_{23}\\\hline A_{31}&A_{32}&A_{33}\end{array}\right]=\left[\begin{array}{cc|cccc|c}*&*&*&*&*&*&*\\ *&*&*&*&*&*&*\\\hline *&*&*&*&*&*&*\\\hline *&*&*&*&*&*&*\\ *&*&*&*&*&*&*\\ *&*&*&*&*&*&*\end{array}\right]$$

Here the blocks are $A_{11} = \begin{bmatrix}*&*\\ *&*\end{bmatrix}$, $A_{12} = \begin{bmatrix}*&*&*&*\\ *&*&*&*\end{bmatrix}$, $A_{13} = \begin{bmatrix}*\\ *\end{bmatrix}$, $A_{21} = \begin{bmatrix}*&*\end{bmatrix}$, $A_{22} = \begin{bmatrix}*&*&*&*\end{bmatrix}$, $A_{23} = \begin{bmatrix}*\end{bmatrix}$, $A_{31} = \begin{bmatrix}*&*\\ *&*\\ &*\end{bmatrix}$, $A_{32} = \begin{bmatrix}*&*&*&*\\ *&*&*&*\\ *&*&*&*\end{bmatrix}$, $A_{33} = \begin{bmatrix}*\\ *\\ *\end{bmatrix}$

Fact 1.4.8. Suppose $A = \begin{bmatrix}A_{11}&A_{12}&\cdots&A_{1q}\\A_{21}&A_{22}&\cdots&A_{2q}\\\vdots&\vdots&&\vdots\\A_{p1}&A_{p2}&\cdots&A_{pq}\end{bmatrix}$, $B = \begin{bmatrix}B_{11}&B_{12}&\cdots&B_{1r}\\B_{21}&B_{22}&\cdots&B_{2r}\\\vdots&\vdots&&\vdots\\B_{q1}&B_{q2}&\cdots&B_{qr}\end{bmatrix}$ are partitioned matrices, the number of columns of submatrix A_{1k} is equal to the number of rows of submatrix B_{k1} (i.e. all submatrices multiplications make sense). Then the usual matrix multiplication is the same as the first treating these submatrix as numbers, do the row-column rule and then multiply them out

Example 1.4.9. Consider $\left[\begin{array}{c|c}A_{11}&A_{12}\\\hline A_{21}&A_{22}\end{array}\right] = \left[\begin{array}{cc|c}1&1&1\\2&2&1\\\hline 2&1&1\end{array}\right]$, $\left[\begin{array}{c|c}B_{11}&B_{12}\\\hline B_{21}&B_{22}\end{array}\right] = \left[\begin{array}{c|cc}1&1&1\\2&2&1\\\hline 2&1&1\end{array}\right]$, then we have

$$\left[\begin{array}{c|cc}5&3&7\\8&5&11\\\hline 7&4&10\end{array}\right]=AB=\left[\begin{array}{c|c}A_{11}B_{11}+A_{12}B_{21}&A_{11}B_{12}+A_{12}B_{22}\\\hline A_{21}B_{11}+A_{22}B_{21}&A_{21}B_{12}+A_{22}B_{22}\end{array}\right]$$

Now we are able to write the vector equation (1.3.1) as a matrix equation

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

Here $A = \begin{bmatrix}\mathbf{a}_1 & \mathbf{a}_2 & \cdots & \mathbf{a}_n\end{bmatrix}$ as a partitioned matrix, $\mathbf{x} = \begin{bmatrix}x_1\\x_2\\\vdots\\x_n\end{bmatrix}$

Example 1.4.10. Follow Example 1.3.4, we can rewrite the vector equation as a matrix equation as in (1.4.1)

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

1.5 Lecture 5 - 06/06/2022

Question. Suppose A is $m\times n$, when does the matrix equation $A\mathbf{x} = \mathbf{b}$ always has a solution for any $\mathbf{b}$ in $\mathbb{R}^n$

Theorem 1.5.1. Let A be an $m\times n$ matrix. Then the following statements are logically equivalent. That is, for a particular A, either they are all true statements or they are all false.

a. For each $\mathbf{b}$ in $\mathbb{R}^m$, the equation $A\mathbf{x} = \mathbf{b}$ has a solution.

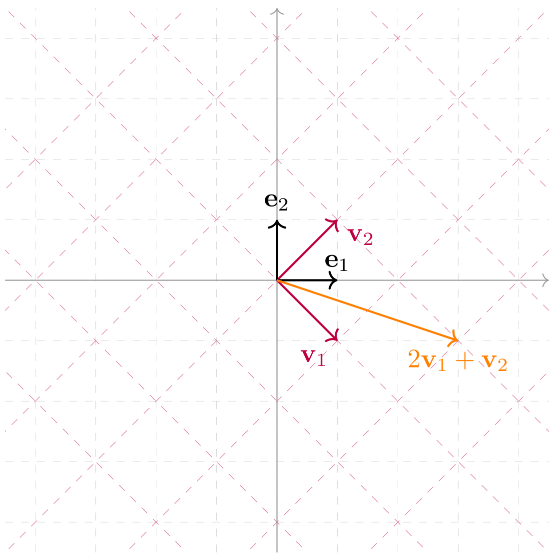
b. Each $\mathbf{b}$ in $\mathbb{R}^m$ is a linear combination of the columns of A.

c. The columns of A span $\mathbb{R}^m$.

d. A has a pivot position in every row. (Equivalentely, in the last row)

Question. What is the geometric meaning of spans

Example 1.5.2. Consider $\mathbf{v}_1 = \begin{bmatrix}1\\-1\end{bmatrix}$, $\mathbf{v}_2 = \begin{bmatrix}1\\1\end{bmatrix}$, $\mathbf{e}_1 = \begin{bmatrix}1\\0\end{bmatrix}$



What is Span$\{\mathbf{v}_1, \mathbf{v}_2\}$, find c_1, c_2 such that $\mathbf{e}_1 = c_1\mathbf{v}_1 + c_2\mathbf{v}_2$

Answer. Apply Theorem 1.5.1 to $A = \begin{bmatrix} \mathbf{v}_1 & \mathbf{v}_2 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix} \sim \begin{bmatrix} 1 & 1 \\ 0 & 2 \end{bmatrix}$ which have a pivot on each row, so the column vectors $\{\mathbf{v}_1, \mathbf{v}_2\}$ of A span the whole plane $\mathbb{R}^2$, i.e. $\mathrm{Span}\{\mathbf{v}_1, \mathbf{v}_2\} = \mathbb{R}^2$. To solve c_1, c_2, let's consider

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

Hence $\mathbf{e}_1 = \frac{1}{2}\mathbf{v}_1 + \frac{1}{2}\mathbf{v}_2$

Definition 1.5.3. $\{\mathbf{v}_1, \cdots, \mathbf{v}_n\}$ is linearly dependent if some $\mathbf{v}_i$ can be written as a linear combination of the others, or equivalently, if there is a non-trivial solution $c_1, \cdots, c_n$ (i.e. not all c_i's are 0) to the vector equation

$$c_1\mathbf{v}_1 + \cdots + c_n\mathbf{v}_n = \mathbf{0} \tag{1.5.1}$$

(1.5.1) is refered to as a linear dependence between $\{\mathbf{v}_1, \cdots, \mathbf{v}_n\}$. If (1.5.1) has only the trivial solution (i.e. $c_1, \cdots, c_n$ are all 0, which is of course always a solution), $\{\mathbf{v}_1, \cdots, \mathbf{v}_n\}$ is said to be linearly independent

Remark. Equivalence between two different definitions of linear dependence

- If $\mathbf{v}_i = c_1\mathbf{v}_1 + \cdots + c_n\mathbf{v}_n$, then $c_1\mathbf{v}_1 + \cdots + (-1)\mathbf{v}_i + \cdots + c_n\mathbf{v}_n = \mathbf{0}$
- If $c_1\mathbf{v}_1 + \cdots + c_i\mathbf{v}_i + \cdots + c_n\mathbf{v}_n = \mathbf{0}$ and $c_i \neq 0$ (since not all c_i's are zero, we may assume c_i is nonzero), then $\mathbf{v}_i = -\frac{c_1}{c_i}\mathbf{v}_1 - \cdots - \frac{c_n}{c_i}\mathbf{v}_n$

Theorem 1.5.4. To determine the linear dependence of $\{\mathbf{v}_1, \cdots, \mathbf{v}_n\}$, we may consider $A = \begin{bmatrix} \mathbf{v}_1 & \cdots & \mathbf{v}_n \end{bmatrix}$, then the solution to $x_1\mathbf{v}_1 + \cdots + x_n\mathbf{v}_n = \mathbf{0}$ is the solution to the RREF form of $\begin{bmatrix} A & \mathbf{0} \end{bmatrix}$. In other words, $\{\mathbf{v}_1, \cdots, \mathbf{v}_n\}$ is linearly independent $\iff$ each column of the RREF of A is a pivot column.

Example 1.5.5. In Example 1.5.2, $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{e}_1\}$ is linearly dependent since

$$\begin{bmatrix} \mathbf{v}_1 & \mathbf{v}_2 & \mathbf{e}_1 & \mathbf{0} \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 & 0 \\ -1 & 1 & 0 & 0 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & \frac{1}{2} & 0 \\ 0 & 1 & \frac{1}{2} & 0 \end{bmatrix}$$

Note that this computation is essentially the same as in Example 1.5.2, the $\mathbf{0}$ in the augmented matrix doesn't change when performing row reductions. We get

$$\begin{cases} x_1 = -\frac{1}{2}x_3 \\ x_2 = -\frac{1}{2}x_3 \\ x_3 \text{ is free} \end{cases}$$

Choose $x_3 = -1$ we get the linear dependence $\frac{1}{2}\mathbf{v}_1 + \frac{1}{2}\mathbf{v}_2 - \mathbf{e}_1 = \mathbf{0}$. On the other hand, $\mathbf{v}_1, \mathbf{v}_2$ are linearly independent since

$$\begin{bmatrix} \mathbf{v}_1 & \mathbf{v}_2 & \mathbf{0} \end{bmatrix} = \begin{bmatrix} 1 & 1 & 0 \\ -1 & 1 & 0 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$$

Which says that $\begin{cases} x_1 = 0 \\ x_2 = 0 \end{cases}$, the trivial solution.

Definition 1.5.6. Suppose $\{\mathbf{v}_1, \cdots, \mathbf{v}_n\} \subseteq \mathbb{R}^m$ (meaning a subset of vectors in $\mathbb{R}^m$) is linearly independent and $\mathrm{Span}\{\mathbf{v}_1, \cdots, \mathbf{v}_n\} = \mathbb{R}^m$, then n necessarily equals to m. We call such a set $\{\mathbf{v}_1, \cdots, \mathbf{v}_n\}$ basis for $\mathbb{R}^m$, and the maximal number of independent vectors (in this case m) the dimension of $\mathbb{R}^m$. (This only has to do with the space $\mathbb{R}^m$ it self, not the choice of basis.)

Example 1.5.7. In Example 1.5.2, $\{\mathbf{v}_1, \mathbf{v}_2\}$ is a basis for $\mathbb{R}^2$, if we let $\mathbf{e}_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$, then $\{\mathbf{e}_1, \mathbf{e}_2\}$ is another basis for $\mathbb{R}^2$. In fact, any two vectors in $\mathbb{R}^2$ that are not parallel form a basis.

Definition 1.5.8. The standard basis for $\mathbb{R}^m$ is the set of vectors $\{\mathbf{e}_1, \cdots, \mathbf{e}_m\}$, where

$$\mathbf{e}_j = \begin{bmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{bmatrix} \leftarrow j\text{-th entry}$$

Example 1.5.9. $\left\{ \mathbf{e}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \mathbf{e}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \mathbf{e}_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \right\}$ is the standard basis for $\mathbb{R}^3$, and

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} x_1 \\ 0 \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ x_2 \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ x_3 \end{bmatrix} = x_1 \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + x_2 \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} + x_3 \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = x_1\mathbf{e}_1 + x_2\mathbf{e}_2 + x_3\mathbf{e}_3$$

Theorem 1.5.10. $\{\mathbf{v}_1, \cdots, \mathbf{v}_m\}$ is a basis $\iff$ the RREF of $\begin{bmatrix} \mathbf{v}_1 & \cdots & \mathbf{v}_m \end{bmatrix}$ is the identity matrix

Theorem 1.5.11. If $\{\mathbf{v}_1, \cdots, \mathbf{v}_m\}$ is a basis for $\mathbb{R}^m$, then $A = \begin{bmatrix} \mathbf{v}_1 & \cdots & \mathbf{v}_n \end{bmatrix} \sim I$. For any vector $\mathbf{b}$ in $\mathbb{R}^m$, $\mathbf{v}$ can be uniquely written as some linear combination of the basis vectors (since $A\mathbf{x} = \mathbf{b}$ always has a unique solution)

1.6 Lecture 6 - 06/07/2022

A linear system is homogeneous if it has matrix equation $A\mathbf{x} = \mathbf{0}$, note that this always have the zero solution, called the trivial solution. $A\mathbf{x} = \mathbf{0}$ has a non-trivial solution $\iff$ it has at least one free variable $\iff$ columns of A are linearly dependent.
We can express the solution set to a linear system in parametric vector form

Example 1.6.1. In Example 1.2.13, the solution set can be written as

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

In Example 1.2.16, the solution set can be written as

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 2x_2 + 16 \\ x_2 \\ \frac{5}{2} \\ -\frac{9}{2} \end{bmatrix} = \begin{bmatrix} 2x_2 \\ x_2 \\ 0 \\ 0 \end{bmatrix} + \begin{bmatrix} 16 \\ 0 \\ \frac{5}{2} \\ -\frac{9}{2} \end{bmatrix} = x_2 \begin{bmatrix} 2 \\ 1 \\ 0 \\ 0 \end{bmatrix} + \begin{bmatrix} 16 \\ 0 \\ \frac{5}{2} \\ -\frac{9}{2} \end{bmatrix}$$

Exercise 1.6.2. Suppose the augmented matrix of a linear system is equivalent to the following matrix

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

Write down the solution set in parametric vector form

Solution.

$$\begin{cases} x_1 + x_2 + 2x_4 = 3 \\ x_3 - 2x_4 = 2 \\ x_5 = 1 \end{cases} \Rightarrow \begin{cases} x_1 = 3 - x_2 - 2x_4 \\ x_2 \text{ is free} \\ x_3 = 2 + 2x_4 \\ x_4 \text{ is free} \\ x_5 = 1 \end{cases}$$

So the solution in parametric vector form would be

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

□

Question. What is the relation of the RREF of A and the RREF of $\begin{bmatrix} A & \mathbf{b} \end{bmatrix}$

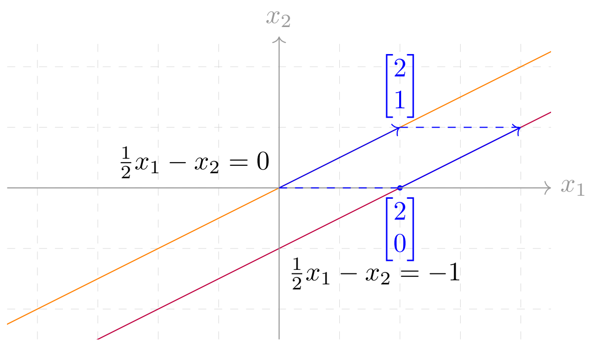
Answer. Suppose $\begin{bmatrix} A & \mathbf{b} \end{bmatrix} \sim \begin{bmatrix} U & \mathbf{d} \end{bmatrix}$ is the RREF, then U will be the RREF of A. Note that the solutions of $A\mathbf{x} = \mathbf{0}$ and $A\mathbf{x} = \mathbf{b}$ differs by $\mathbf{d}$, i.e.

$$\mathbf{d} + \{\text{solutions of } A\mathbf{x} = \mathbf{0}\} = \{\text{solutions of } A\mathbf{x} = \mathbf{b}\}$$

Question. What is the geometric meaning of a solution set in parametric vector form

Answer. It is the hyperplane (solutions of $A\mathbf{x} = \mathbf{b}$) translated from the hyperplane through the origin (solutions of $A\mathbf{x} = \mathbf{0}$) by $\mathbf{d}$

Example 1.6.3. $\frac{1}{2}x_1 - x_2 = -1$ has solution set $\begin{cases} x_1 = -2 + 2x_2 \\ x_2 \text{ is free} \end{cases}$ which is $\begin{bmatrix} -2 \\ 0 \end{bmatrix} + x_2 \begin{bmatrix} 2 \\ 1 \end{bmatrix}$ in parametric vector form, and $\frac{1}{2}x_1 - x_2 = 0$ has solution set $\begin{cases} x_1 = 2x_2 \\ x_2 \text{ is free} \end{cases}$ which is $x_2 \begin{bmatrix} 2 \\ 1 \end{bmatrix}$ in parametric vector form.



Definition 1.6.4. A Linear transformation (or mapping) T is a mapping from $\mathbb{R}^n$ to $\mathbb{R}^m$ satisfying

- $T(\mathbf{u} + \mathbf{v}) = T(\mathbf{u}) + T(\mathbf{v})$ for any $\mathbf{u}, \mathbf{v}$ in $\mathbb{R}^n$
- $T(c\mathbf{u}) = cT(\mathbf{u})$ for any scalar c and any $\mathbf{u}$ in $\mathbb{R}^n$

Definition 1.6.5. A matrix transformation is a linear transformation defined via matrix multiplication, i.e. $T(\mathbf{x}) = A\mathbf{x}$ for some $m \times n$ matrix. It is linear thanks to Fact 1.4.6 c),d) since

- $T(\mathbf{u} + \mathbf{v}) = A(\mathbf{u} + \mathbf{v}) = A\mathbf{u} + A\mathbf{v} = T(\mathbf{u}) + T(\mathbf{v})$
- $T(c\mathbf{u}) = A(c\mathbf{u}) = c(A\mathbf{u}) = cT(\mathbf{u})$

Question. How to know a linear transformation T?

Answer. Just need to know its effect on any basis! Suppose $\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_n$ is a basis of $\mathbb{R}^n$, then any vector $\mathbf{v}$ can be written as $c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_n\mathbf{v}_n$, and by linearality, we have

$$T(\mathbf{v}) = T(c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_n\mathbf{v}_n) = c_1T(\mathbf{v}_1) + c_2T(\mathbf{v}_2) + \cdots + c_nT(\mathbf{v}_n) \tag{1.6.1}$$

Let $\{\mathbf{e}_1, \cdots, \mathbf{e}_n\}$ as the standard basis as in Definition 1.5.8, if $T : \mathbb{R}^n \to \mathbb{R}^m$ is a linear transformation, then any $\mathbf{x} = x_1\mathbf{e}_1 + \cdots + x_n\mathbf{e}_n$, by (1.6.1)

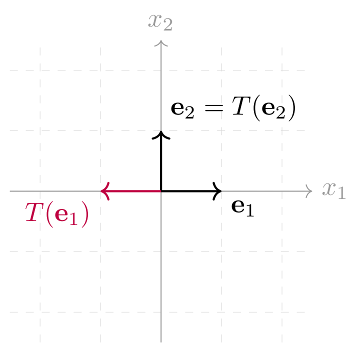
$$T(\mathbf{x}) = x_1T(\mathbf{e}_1) + \cdots + x_nT(\mathbf{e}_n) = \begin{bmatrix} T(\mathbf{e}_1) & \cdots & T(\mathbf{e}_n) \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \tag{1.6.2}$$

So $A = \begin{bmatrix} T(\mathbf{e}_1) & \cdots & T(\mathbf{e}_n) \end{bmatrix}$ (which is called the standard matrix for the linear transformation T) is the unique matrix satisfies $T(\mathbf{x}) = A\mathbf{x}$

Example 1.6.6. Suppose $T : \mathbb{R}^2 \to \mathbb{R}^2$ is a linear transformation

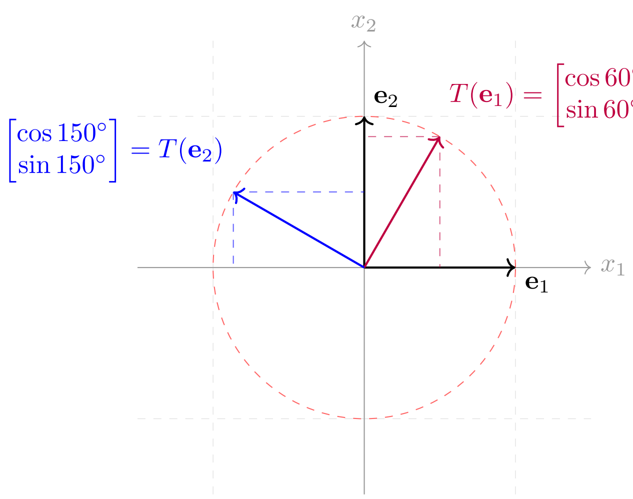
a) Assume T is the reflection over x_2-axis, then the standard matrix for T is

$$A = \begin{bmatrix} T(\mathbf{e}_1) & T(\mathbf{e}_2) \end{bmatrix} = \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}$$



b) Assume T is the rotation by $60°$ counter-clockwise, then the standard matrix for T is

$$A = \begin{bmatrix} T(\mathbf{e}_1) & T(\mathbf{e}_2) \end{bmatrix} = \begin{bmatrix} \frac{1}{2} & -\frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & \frac{1}{2} \end{bmatrix}$$



1.7 Lecture 7 - 06/08/2022

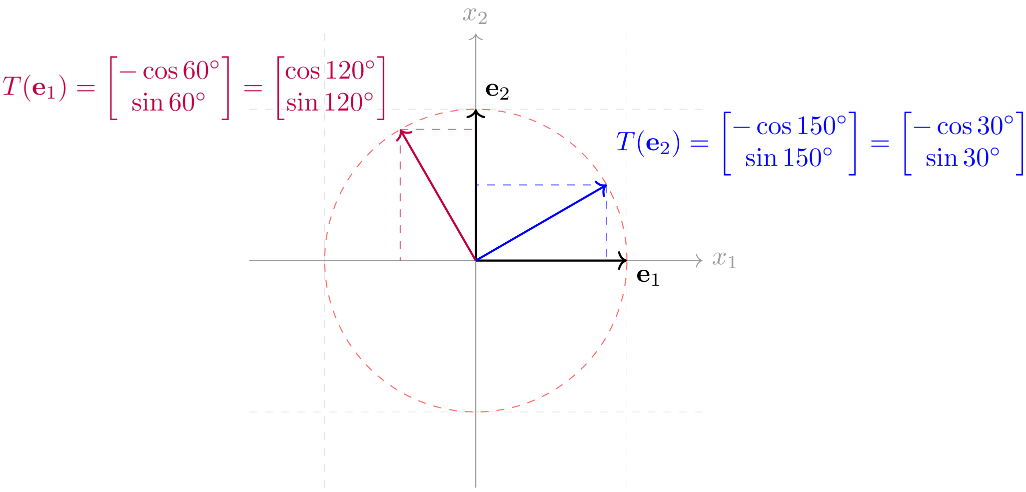
Exercise 1.7.1. Suppose $T : \mathbb{R}^2 \to \mathbb{R}^2$ is defined by $T\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right) = \begin{bmatrix} x_1 + x_2 + 1 \\ x_2 \end{bmatrix}$. Is T is a linear mapping?

Solution. T is not a linear transformation since $T(\mathbf{x}) = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} 1 \\ 0 \end{bmatrix} = A\mathbf{x} + \mathbf{p}$ is not of the matrix transformation form (1.6.2) ($\mathbf{p} \neq 0$) □

Exercise 1.7.2. $T : \mathbb{R}^2 \to \mathbb{R}^2$ is the linear transformation that rotate $60°$ counter-clockwise and then reflects over x_2-axis, what is its standard matrix?

Solution. The standard matrix for T is

$$A = \begin{bmatrix} T(\mathbf{e}_1) & T(\mathbf{e}_2) \end{bmatrix} = \begin{bmatrix} -\frac{1}{2} & \frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & \frac{1}{2} \end{bmatrix}$$



□

Definition 1.7.3. Suppose $T : \mathbb{R}^n \to \mathbb{R}^m$ is a mapping. We call

- $\mathbb{R}^n$ the domain of T
- $\mathbb{R}^m$ the codomain of T
- $T(\mathbf{x})$ the image of $\mathbf{x}$ under T
- $T^{-1}(\mathbf{b}) = \{\mathbf{x} | T(\mathbf{x}) = \mathbf{b}\}$ the preimage of $\mathbf{b}$ under T
- the set of images $\{T\mathbf{x} | \mathbf{x} \in \mathbb{R}^n\}$ the range of T

Exercise 1.7.4. Suppose the linear transformation $T : \mathbb{R}^3 \to \mathbb{R}^3$ is defined by $T\left(\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}\right) = \begin{bmatrix} x_1 - x_2 + x_3 \\ 2x_1 - x_3 \\ x_1 + x_2 + x_3 \end{bmatrix}$, what is the standard matrix of T? What is the image $\begin{bmatrix} 2 \\ 0 \\ 1 \end{bmatrix}$, what is the preimage of $\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$, what is the range?

Solution. The standard matrix is $A = \begin{bmatrix} 1 & -1 & 1 \\ 2 & 0 & -1 \\ 1 & 1 & 1 \end{bmatrix}$, the image $\begin{bmatrix} 2 \\ 0 \\ 1 \end{bmatrix}$ under T is

$$T(\mathbf{x}) = \begin{bmatrix} 1 & -1 & 1 \\ 2 & 0 & -1 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 2 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 2 - 0 + 1 \\ 2 \cdot 2 - 1 \\ 2 + 0 + 1 \end{bmatrix} = \begin{bmatrix} 3 \\ 3 \\ 3 \end{bmatrix}$$

The preimage of $\mathbf{b} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$ under T is the solution set to $T(\mathbf{x}) = A\mathbf{x} = \mathbf{b}$ (this is Example 1.2.7), which is $\left\{ \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \right\}$. And since there is a pivot in each row, by Theorem 1.5.1, the range of T is $\mathbb{R}^3$ □

Definition 1.7.5. A mapping T is said to be onto $\mathbb{R}^m$ if each $b \in \mathbb{R}^m$ is the image of at least one $x \in \mathbb{R}^n$.

Codomain is larger than the range if T is not onto

Definition 1.7.6. A mapping T is said to be one-to-one if each $b \in \mathbb{R}^m$ is the image of at most one $x \in \mathbb{R}^n$.

Theorem 1.7.7. Suppose A is the standard matrix for linear transformation T (i.e. $T(\mathbf{x}) = A\mathbf{x}$), then

- T is one-to-one $\iff$ $A\mathbf{x} = \mathbf{0}$ has a unique solution $\iff$ $A\mathbf{x} = \mathbf{b}$ has at most a unique solution $\iff$ RREF of A has a pivot in each column $\iff$ columns of A are linearly independent.
- T is onto $\iff$ the columns of A span $\mathbb{R}^m$ $\iff$ RREF of A has a pivot in each row.

Exercise 1.7.8. Suppose the linear transformation $T : \mathbb{R}^3 \to \mathbb{R}^2$ is defined by $T\left(\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}\right) = \begin{bmatrix} x_1 - x_2 + x_3 \\ 2x_1 - x_3 \end{bmatrix}$, Is T onto? Is T one-to-one?

Solution. This is the Example 1.2.13. The standard matrix for T is $A = \begin{bmatrix} 1 & -1 & 1 \\ 2 & 0 & -1 \end{bmatrix} \xrightarrow{R2 \to R2 - 2R1} \begin{bmatrix} 1 & -1 & 1 \\ 0 & 2 & -3 \end{bmatrix}$, since there is a pivot in each row but not in each column, it is onto but not one-to-one □

Question.

- If $T : \mathbb{R}^3 \to \mathbb{R}^2$ is a linear transformation, could it be one-to-one?
- If $T : \mathbb{R}^2 \to \mathbb{R}^3$ is a linear transformation, could it be onto?

Answer. Both no! Due to Theorem 1.7.7

- Since A is a 2×3 matrix, there will be at most 2 pivots (only 2 rows), so there won't be enough pivots to fill all columns.
- Since A is a 3×2 matrix, there will be at most 2 pivots (only 2 columns), so there won't be enough pivots to fill all rows.

1.8 Lecture 8 - 06/09/2022

Question. What happens if we compose two linear transformation (say $T_1 : \mathbb{R}^p \to \mathbb{R}^n$, $T_2 : \mathbb{R}^n \to \mathbb{R}^m$)?

Definition 1.8.1. We have
$$\begin{tikzcd} \mathbb{R}^p \arrow[r, "T_1"] \arrow[rr, bend right, "T_2 \circ T_1"'] & \mathbb{R}^n \arrow[r, "T_2"] & \mathbb{R}^m \end{tikzcd}$$
, Here $\circ$ means composition which you compose from the left. suppose the standard matrices for T_1, T_2 are A_1, A_2 respectively, then A_1 is $n \times p$, A_2 is $m \times n$, and for $\mathbf{x} \in \mathbb{R}^p$, $(T_2 \circ T_1)(\mathbf{x}) = T_2(T_1(\mathbf{x})) = A_2(T_1(\mathbf{x})) = A_2(A_1\mathbf{x}) = (A_2A_1)\mathbf{x}$. So we have concluded that the standard matrix for $T_2 \circ T_1$ is the $m \times p$ matrix A_2A_1

Example 1.8.2. Consider Example 1.6.6., If we let $T_1 : \mathbb{R}^2 \to \mathbb{R}^2$ to denote the rotation by 60° counter-clockwise, $T_2 : \mathbb{R}^2 \to \mathbb{R}^2$ to denote reflection over x_2-axis, and their standard matrices are

$$\begin{bmatrix} \frac{1}{2} & -\frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & \frac{1}{2} \end{bmatrix}, \quad \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}$$

Then look at Exercise 1.7.2, this is the composition $T_2 \circ T_1$, which has the standard matrix

$$A_2A_1 = \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \frac{1}{2} & -\frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & \frac{1}{2} \end{bmatrix} = \begin{bmatrix} -\frac{1}{2} & \frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & \frac{1}{2} \end{bmatrix}$$

Question. Suppose $A = \begin{bmatrix} \frac{1}{2} & -\frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & \frac{1}{2} \end{bmatrix}$ is the standard matrix for the linear transformation of rotating $60°$ counter-clockwise (Example 1.6.6). What is A^7?

Answer. $A^7 = AAAAAAA$ is the standard matrix for composition of linear transformations $T \circ T \circ T \circ T \circ T \circ T \circ T$ which is rotate $7 \times 60° = 420°$, but that is the same as rotating $420° - 360° = 60°$ which is the same linear transformation as T, so $A^7 = A$

$$\begin{tikzcd} A^7 = AAAAAAA \arrow[r, equal] & A \\ T^{\circ 7} = T \circ T \circ T \circ T \circ T \circ T \circ T \arrow[u, "\text{Standard matrix}"] \arrow[r, equal, "\text{same effect}"'] & T \arrow[u, "\text{Standard matrix}"'] \end{tikzcd}$$

Definition 1.8.3. Suppose $A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}$ is a $m \times n$ matrix, we define its transpose by flipping it over the diagonal, and this is the $n \times m$ matrix

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

Example 1.8.4. Suppose $A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix}$, then $A^T = \begin{bmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{bmatrix}$

Theorem 1.8.5. Here are some properties of matrix transpose

- $(A^T)^T = A$
- $(A + B)^T = A^T + B^T$
- $(cA)^T = cA^T$
- $(AB)^T = B^T A^T$

Definition 1.8.6. For any $\mathbf{v}, \mathbf{w} \in \mathbb{R}^n$, we can define the dot product to be $\mathbf{v} \cdot \mathbf{w} = \mathbf{v}^T \mathbf{w} = v_1 w_1 + \cdots + v_n w_n$. $\|\mathbf{v}\| = \sqrt{\mathbf{v}^T \mathbf{v}} = \sqrt{v_1^2 + \cdots + v_n^2}$ is the length $\mathbf{v}$

Elementary row operations on matrix A can be realized as elementary matrices E multiplication on A.

Exercise 1.8.7. Let A be a 3 by 3 matrix, write down the elementary matrix E such that

1. EA acts as subtracting 3 times row 3 from row 1
2. AE acts as adding 2 times column 3 to column 1
3. EA acts as scaling the third row by 2
4. AE acts as scaling the third column by 3
5. EA acts as interchanging row 1 and row 3
6. AE acts as interchanging column 1 and column 2

Solution.

1. $E = \begin{bmatrix} 1 & 0 & -3 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$, $EA = \begin{bmatrix} 1 & 0 & -3 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} R1 \\ R2 \\ R3 \end{bmatrix} = \begin{bmatrix} R1 - 3R3 \\ R2 \\ R3 \end{bmatrix}$

2. $E = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 2 & 0 & 1 \end{bmatrix}$, $AE = \begin{bmatrix} \mathbf{a}_1 & \mathbf{a}_2 & \mathbf{a}_3 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 2 & 0 & 1 \end{bmatrix} = \begin{bmatrix} \mathbf{a}_1 + 2\mathbf{a}_3 & \mathbf{a}_2 & \mathbf{a}_3 \end{bmatrix}$

3. $E = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{bmatrix}$, $EA = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} R1 \\ R2 \\ R3 \end{bmatrix} = \begin{bmatrix} R1 \\ 2R2 \\ R3 \end{bmatrix}$

4. $E = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 1 \end{bmatrix}$, $AE = \begin{bmatrix} \mathbf{a}_1 & \mathbf{a}_2 & \mathbf{a}_3 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} \mathbf{a}_1 & 3\mathbf{a}_2 & \mathbf{a}_3 \end{bmatrix}$

5. $E = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}$, $EA = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} R1 \\ R2 \\ R3 \end{bmatrix} = \begin{bmatrix} R3 \\ R2 \\ R1 \end{bmatrix}$

6. $E = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$, $AE = \begin{bmatrix} \mathbf{a}_1 & \mathbf{a}_2 & \mathbf{a}_3 \end{bmatrix} \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} \mathbf{a}_2 & \mathbf{a}_1 & \mathbf{a}_3 \end{bmatrix}$

□

1.9 Lecture 9 - 06/10/2022

Definition 1.9.1. Suppose linear transformation $T : \mathbb{R}^n \to \mathbb{R}^m$ is both onto and one-to-one (i.e. every vector $\mathbf{b}$ in the codomain $\mathbb{R}^m$ is a unique image, which we denote as $T^{-1}(\mathbf{b})$), and A is the standard matrix for T, then m necessary equal n as showned in Question 1.7, so A must be a square matrix. We know $T(\mathbf{x}) = \mathbf{b}$ always has a unique solution which is $T^{-1}(\mathbf{b})$, it can be shown that $T^{-1} : \mathbb{R}^n \to \mathbb{R}^n$ as mapping is actually also a linear transformation (Why? See if you can figure this out). Then the standard matrix of T^{-1} is defined to be A^{-1} the inverse matrix of A, since

$$\begin{aligned}(T \circ T^{-1})(\mathbf{b}) &= T(T^{-1}(\mathbf{b})) = T(\mathbf{x}) = \mathbf{b} \\ (T^{-1} \circ T)(\mathbf{x}) &= T^{-1}(T(\mathbf{x})) = T^{-1}(\mathbf{b}) = \mathbf{x}\end{aligned}$$

Note that $T \circ T^{-1}$, $T^{-1} \circ T$ work like the identity map, so $AA^{-1} = A^{-1}A = I$. In this case, we see that A is equivalent to the identity matrix (because of Theorem 1.5.2, A has a pivot in each row and column).

Remark. Because we can write elementary row operations as left elementary matrix multiplications, so we know there are elementary matrices $E_1, E_2, \cdots, E_k$ such that $E_k E_{k-1} \cdots E_2 E_1 A = I$. If we multiply A^{-1} on the right on both sides, we get $E_k E_{k-1} \cdots E_2 E_1 = A^{-1}$

Using the remark above, we can deduce an algorithm for computing matrix inverses. Let's consider the RREF of the following partitioned matrix

$$\left[\begin{array}{c|c} A & I \end{array}\right] \sim \left[\begin{array}{c|c} E_k E_{k-1} \cdots E_2 E_1 A & E_k E_{k-1} \cdots E_2 E_1 I \end{array}\right] = \left[\begin{array}{c|c} I & A^{-1} \end{array}\right]$$

1.10 Lecture 10 - 06/13/2022

Exercise 1.10.1. Find the inverse of the following matrices.

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

Solution.

a)

$$\left[\begin{array}{cc|cc} -1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{array}\right] \underset{\smile}{\overset{(-1)R1}{}} \left[\begin{array}{cc|cc} 1 & 0 & -1 & 0 \\ 0 & 1 & 0 & 1 \end{array}\right]$$

Hence $A^{-1} = \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}$

b)

$$\left[\begin{array}{cc|cc} 1 & 2 & 1 & 0 \\ 3 & 5 & 0 & 1 \end{array}\right] \underset{\smile}{\overset{R2\to R2-3R1}{}} \left[\begin{array}{cc|cc} 1 & 2 & 1 & 0 \\ 0 & -1 & -3 & 1 \end{array}\right]$$
$$\underset{\smile}{\overset{R1\to R1+2R2}{}} \left[\begin{array}{cc|cc} 1 & 0 & -5 & 2 \\ 0 & -1 & -3 & 1 \end{array}\right] \underset{\smile}{\overset{(-1)R2}{}} \left[\begin{array}{cc|cc} 1 & 0 & -5 & 2 \\ 0 & 1 & 3 & -1 \end{array}\right]$$

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

c)

$$\left[\begin{array}{ccc|ccc} 1 & -1 & 1 & 1 & 0 & 0 \\ 2 & 0 & -1 & 0 & 1 & 0 \\ 1 & 1 & 1 & 0 & 0 & 1 \end{array}\right] \underset{\smile}{\overset{\substack{R2\to R2-2R1 \\ R3\to R3-R1}}{}} \left[\begin{array}{ccc|ccc} 1 & -1 & 1 & 1 & 0 & 0 \\ 0 & 2 & -3 & -2 & 1 & 0 \\ 0 & 2 & 0 & -1 & 0 & 1 \end{array}\right]$$
$$\underset{\smile}{\overset{R3\to R3-R2}{}} \left[\begin{array}{ccc|ccc} 1 & -1 & 1 & 1 & 0 & 0 \\ 0 & 2 & -3 & -2 & 1 & 0 \\ 0 & 0 & 3 & 1 & -1 & 1 \end{array}\right] \underset{\smile}{\overset{R2\to R2+R3}{}} \left[\begin{array}{ccc|ccc} 1 & -1 & 1 & 1 & 0 & 0 \\ 0 & 2 & 0 & -1 & 0 & 1 \\ 0 & 0 & 3 & 1 & -1 & 1 \end{array}\right]$$
$$\underset{\smile}{\overset{\substack{R2/2 \\ R3/3}}{}} \left[\begin{array}{ccc|ccc} 1 & -1 & 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & -\frac{1}{2} & 0 & \frac{1}{2} \\ 0 & 0 & 1 & \frac{1}{3} & -\frac{1}{3} & \frac{1}{3} \end{array}\right] \underset{\smile}{\overset{R1\to R1+R2-R3}{}} \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & \frac{1}{6} & \frac{1}{3} & \frac{1}{6} \\ 0 & 1 & 0 & -\frac{1}{2} & 0 & \frac{1}{2} \\ 0 & 0 & 1 & \frac{1}{3} & -\frac{1}{3} & \frac{1}{3} \end{array}\right]$$

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

□

Exercise 1.10.2. What is $(A^T)^{-1}$ in Exercise 1.10.1, c)?

Solution. Use Theorem 1.10.9, we know

$$(A^T)^{-1} = (A^{-1})^T = \begin{bmatrix} \frac{1}{6} & \frac{1}{3} & \frac{1}{6} \\ -\frac{1}{2} & 0 & \frac{1}{2} \\ \frac{1}{3} & -\frac{1}{3} & \frac{1}{3} \end{bmatrix}^T = \begin{bmatrix} \frac{1}{6} & -\frac{1}{2} & \frac{1}{3} \\ \frac{1}{3} & 0 & -\frac{1}{3} \\ \frac{1}{6} & \frac{1}{2} & \frac{1}{3} \end{bmatrix}$$

□

Definition 1.10.3. A square matrix A is invertible (or non-singular) if it has an inverse A^{-1} such that $AA^{-1} = A^{-1}A = I$. A is called singular if A is not invertible.

Theorem 1.10.4. Suppose T is a linear transformation with standard matrix A, then

$$T \text{ is invertible with inverse } T^{-1} \iff A \text{ is invertible with inverse } A^{-1} \iff A \sim I$$

Theorem 1.10.5. $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$, $A^{-1} = \dfrac{1}{ad-bc}\begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$, here $\det A = ad - bc$

Example 1.10.6. If $A = \begin{bmatrix} \frac{1}{2} & \frac{\sqrt{3}}{2} \\ -\frac{\sqrt{3}}{2} & \frac{1}{2} \end{bmatrix}$, then

$$A^{-1} = \frac{1}{\frac{1}{2}\frac{1}{2} - \frac{\sqrt{3}}{2}\left(-\frac{\sqrt{3}}{2}\right)} \begin{bmatrix} \frac{1}{2} & -\frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & \frac{1}{2} \end{bmatrix} = \begin{bmatrix} \frac{1}{2} & -\frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & \frac{1}{2} \end{bmatrix}$$

Theorem 1.10.7. If A is invertible, then the linear system $A\mathbf{x} = \mathbf{b}$ has a unique solution $\mathbf{x} = A^{-1}\mathbf{b}$

Example 1.10.8. Let's consider (1.1.2), in which case $A = \begin{bmatrix} 1 & 1 \\ 2 & 4 \end{bmatrix}$, $\mathbf{b} = \begin{bmatrix} 10 \\ 26 \end{bmatrix}$, then $A^{-1} = \dfrac{1}{1 \cdot 4 - 1 \cdot 2}\begin{bmatrix} 4 & -1 \\ -2 & 1 \end{bmatrix} = \begin{bmatrix} 2 & -\frac{1}{2} \\ -1 & \frac{1}{2} \end{bmatrix}$, and

$$\mathbf{x} = A^{-1}\mathbf{b} = \begin{bmatrix} 2 & -\frac{1}{2} \\ -1 & \frac{1}{2} \end{bmatrix}\begin{bmatrix} 10 \\ 26 \end{bmatrix} = \begin{bmatrix} 2 \cdot 10 - \frac{1}{2} \cdot 26 \\ -10 + \frac{1}{2} \cdot 26 \end{bmatrix} = \begin{bmatrix} 7 \\ 3 \end{bmatrix}$$

Theorem 1.10.9. Here are some properties of matrix inverse

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

Exercise 1.10.10. Let $T : \mathbb{R}^n \to \mathbb{R}^n$ be a linear transformation with standard matrix A.

- If A is invertible, then A has n pivots. ✓
- If T is one-to-one, then A is invertible. ✓
- If columns of A span $\mathbb{R}^n$, then A is invertible. ✓
- If A is invertible, $A\mathbf{x} = \mathbf{0}$ only has the trivial solution. ✓
- If T is onto, then T is one-to-one. ✓
- If T is one-to-one, then T is onto. ✓

1.11 Lecture 11 - 06/14/2022

Exercise 1.11.1. Suppose $T : \mathbb{R}^3 \to \mathbb{R}^3$ is a linear transformation with standard matrix $A = \begin{bmatrix} 1 & 2 & 3 \\ 0 & -1 & -2 \\ 0 & 0 & 1 \end{bmatrix}$. Is A^{-1} invertible? Is T one-to-one? Is A^T invertible? If so, what is $(A^T)^{-1}$? If so, what is $(A^{-1})^{-1}$. Is T invertible (i.e. does T^{-1} exist)? What is the standard matrix of T^{-1}? Is T onto?

Solution.

$$\left[\begin{array}{c|c} A & I \end{array}\right] = \left[\begin{array}{ccc|ccc} 1 & 2 & 3 & 1 & 0 & 0 \\ 0 & -1 & -2 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 \end{array}\right] \xrightarrow[R2 \to R2 + 2R3]{R1 \to R1 - 3R3} \left[\begin{array}{ccc|ccc} 1 & 2 & 0 & 1 & 0 & -3 \\ 0 & -1 & 0 & 0 & 1 & 2 \\ 0 & 0 & 1 & 0 & 0 & 1 \end{array}\right]$$

$$\xrightarrow{(-1)R2} \left[\begin{array}{ccc|ccc} 1 & 2 & 0 & 1 & 0 & -3 \\ 0 & 1 & 0 & 0 & -1 & -2 \\ 0 & 0 & 1 & 0 & 0 & 1 \end{array}\right] \xrightarrow{R1 \to R1 - 2R2} \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & 2 & 1 \\ 0 & 1 & 0 & 0 & -1 & -2 \\ 0 & 0 & 1 & 0 & 0 & 1 \end{array}\right] = \left[\begin{array}{c|c} I & A^{-1} \end{array}\right]$$

So $A^{-1} = \begin{bmatrix} 1 & 2 & 1 \\ 0 & -1 & -2 \\ 0 & 0 & 1 \end{bmatrix}$. By Theorem 1.10.9, we know

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

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

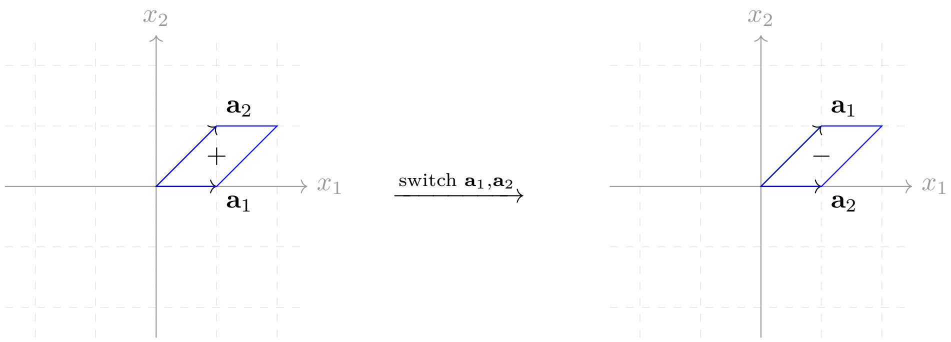
Therefore we know A^{-1} and A^T are invertible.
In general, if T is invertible, then A is invertible, so A^{-1} will be the standard matrix for T^{-1} as $T^{-1}(\mathbf{x}) = A^{-1}\mathbf{x}$, in more explict terms, we have

$$T^{-1}\left(\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}\right) = \begin{bmatrix} 1 & 2 & 1 \\ 0 & -1 & -2 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} x_1 + 2x_2 + x_3 \\ -x_2 - 2x_2 \\ x_3 \end{bmatrix}$$

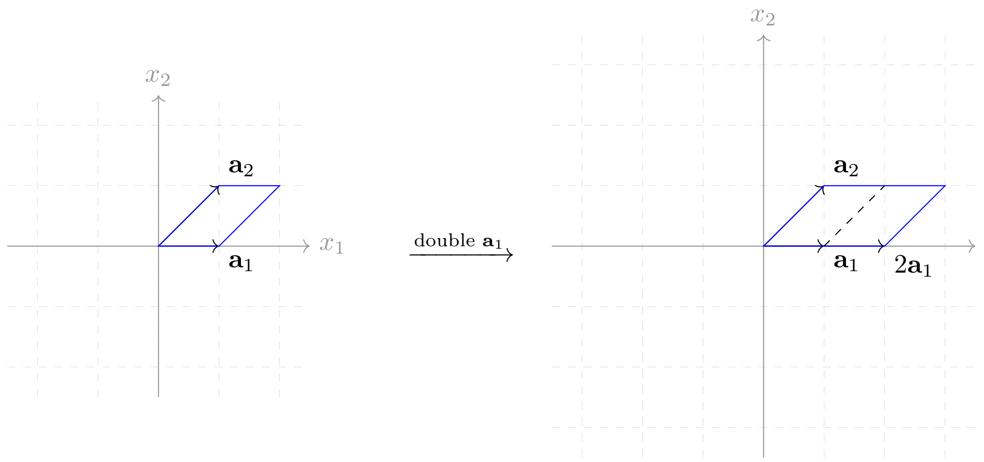
$\square$

Now let's talk about determinants (ONLY for square matrices!!!): Consider the parallelepiped P with edges $\mathbf{a}_1, \cdots, \mathbf{a}_n$ in $\mathbb{R}^n$. We want to think of determinant of $A = \begin{bmatrix} \mathbf{a}_1 & \cdots & \mathbf{a}_n \end{bmatrix}$ (Usually denoted $\det A$ or $|A| = \begin{vmatrix} \mathbf{a}_1 & \cdots & \mathbf{a}_n \end{vmatrix}$, replacing brackets with vertical lines) as signed volumes of P . Therefore we have $\mathrm{Vol}(P) = |\det A|$, i.e. actual volume is the absolute value of the determinant. Note that the determinant/signed volume has these following three properties (Take $n = 2$ for an example, in this case, volume is really area):

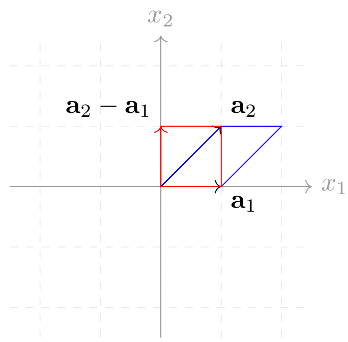
1. The following corresponds to interchange in elementary row operations, this changes the sign of the determinant/signed volume



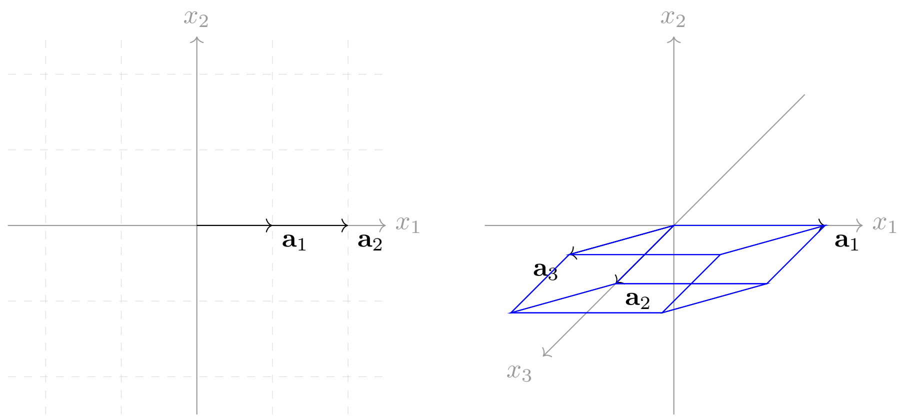
2. The following corresponds to scaling in elementary row operations, this scales the determinant/signed volume



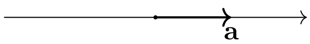
3. The following corresponds to replacement in elementary row operations, this doesn't change the determinant/signed volume



Remark. If $\{\mathbf{a}_1, \cdots, \mathbf{a}_n\}$ is linearly dependent, then A is singular, i.e. not invertible, then the determinant will be zero, since the parallelepiped will be constraint in a hyperplane which has zero volume. Take $n = 2$ and 3 for examples

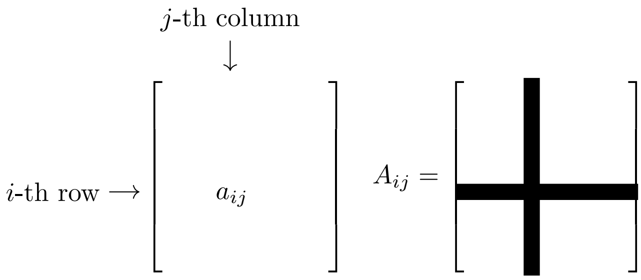


With geometric intepretation of determinants, the case where $n = 1$ is rather straightforward. Suppose $A = [a]$ is a 1 by 1 matrix, then it corresponds to the signed length of the number a on the real line $\mathbb{R}^1$, which is really the number a itself! In other words, we know $\det A = a$.



For $n \geq 2$, we use the following cofactor expansion as an inductive definition.

Definition 1.11.2. We use a_{ij} to be denote the (i, j)-th entry of the matrix A, and A_{ij} to denote the submatrix of A by deleting the i-th row and the j-th column



We define the (i, j)-cofactor to be $C_{ij} = (-1)^{i+j} \det A_{ij}$. The cofactor expansion across the i-th row is

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

The cofactor expansion across the j-th column is

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

$$\begin{bmatrix} & & & \\ & & & \\ a_{i1} & a_{i2} & \cdots & a_{in} \\ & & & \\ & & & \end{bmatrix} \qquad \begin{bmatrix} & a_{1j} & \\ & a_{2j} & \\ & \vdots & \\ & a_{nj} & \end{bmatrix}$$

Exercise 1.11.3. Consider $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$. What is $A_{11}, A_{12}, A_{21}, A_{22}$? What is $C_{11}, C_{12}, C_{21}, C_{22}$. Write down the cofactor expansion of A across the

- 1st row
- 2nd row
- 1st column
- 2nd column

Solution. $A_{11} = [d], A_{12} = [c], A_{21} = [b], A_{22} = [a]$ are all 1 by 1 matrices. $C_{11} = (-1)^{1+1} \det A_{11} = d, C_{12} = (-1)^{1+2} \det A_{21} = -c, C_{21} = (-1)^{2+1} \det A_{21} = -b, C_{22} = (-1)^{2+2} \det A_{22} = a$. So the cofactor expansions are

- $\det A = aC_{11} + bC_{12} = ad - bc$
- $\det A = cC_{21} + dC_{22} = -bc + ad$
- $\det A = aC_{11} + cC_{21} = ad - bc$
- $\det A = bC_{12} + dC_{22} = -bc + ad$

Note that all of the above calculations show that $\det A = \det \begin{bmatrix} a & b \\ c & d \end{bmatrix} = ad - bc$. ☐

1.12 Lecture 12 - 06/15/2022

Exercise 1.12.1. Suppose $A = \begin{bmatrix} 1 & -1 & 1 \\ 2 & 0 & -1 \\ 1 & 1 & 1 \end{bmatrix}$. Please find the cofactor expansion of A across the

a) 1st row

b) 2nd column

And evaluate determinant of A.

Solution.

$$\begin{aligned} \det A &= a_{11}C_{11} + a_{12}C_{12} + a_{13}C_{13} \\ &= a_{11}(-1)^{1+1} \det A_{11} + a_{12}(-1)^{1+2} \det A_{12} + a_{13}(-1)^{1+3} \det A_{13} \\ &= 1 \cdot (-1)^{1+1} \begin{vmatrix} 0 & -1 \\ 1 & 1 \end{vmatrix} + (-1) \cdot (-1)^{1+2} \begin{vmatrix} 2 & -1 \\ 1 & 1 \end{vmatrix} + 1 \cdot (-1)^{1+3} \begin{vmatrix} 2 & 0 \\ 1 & 1 \end{vmatrix} \\ &= 1 \cdot (0 \cdot 1 - (-1) \cdot 1) + (-1) \cdot (2 \cdot 1 - (-1) \cdot 1) + 1 \cdot (-1)(2 \cdot 1 - (-1) \cdot 1) \\ &= 6 \end{aligned}$$

$$\begin{aligned} \det A &= a_{12}C_{12} + a_{22}C_{22} + a_{32}C_{32} \\ &= a_{11}(-1)^{1+1} \det A_{11} + a_{12}(-1)^{1+2} \det A_{12} + a_{13}(-1)^{1+3} \det A_{13} \\ &= (-1) \cdot (-1)^{1+2} \begin{vmatrix} 2 & -1 \\ 1 & 1 \end{vmatrix} + 0 \cdot (-1)^{2+2} \begin{vmatrix} 1 & 1 \\ 1 & 1 \end{vmatrix} + 1 \cdot (-1)^{3+2} \begin{vmatrix} 1 & 1 \\ 2 & -1 \end{vmatrix} \\ &= (-1) \cdot (-1)(2 \cdot 1 - (-1) \cdot 1) + 0 \cdot (1 \cdot 1 - 1 \cdot 1) + 1 \cdot (-1)(1 \cdot (-1) - 1 \cdot 2) \\ &= 6 \end{aligned}$$

☐

Exercise 1.12.2 (Quiz 1)**.** Suppose $A = \begin{bmatrix} 1 & 2 & 1 \\ 1 & 0 & -1 \\ -1 & 2 & 1 \end{bmatrix}$. Write out the cofactor expansion of A across the second row, and evaluate the determinant $\det A$.

Solution.

$$\begin{aligned}
\det A &= a_{21}C_{21} + a_{22}C_{22} + a_{23}C_{23} \\
&= a_{21}(-1)^{2+1}\det A_{21} + a_{22}(-1)^{2+2}\det A_{22} + a_{23}(-1)^{2+3}\det A_{23} \\
&= 1\cdot(-1)^{2+1}\begin{vmatrix} 2 & 1 \\ 2 & 1 \end{vmatrix} + 0\cdot(-1)^{2+2}\begin{vmatrix} 1 & 1 \\ -1 & 1 \end{vmatrix} + (-1)\cdot(-1)^{2+3}\begin{vmatrix} 2 & 1 \\ 2 & 1 \end{vmatrix} \\
&= 1\cdot(-1)(2\cdot 1 - 1\cdot 2) + 0\cdot(1\cdot 1 - 1\cdot(-1)) + (-1)\cdot(-1)(1\cdot 2 - 2\cdot(-1)) \\
&= 4
\end{aligned}$$

□

Remark. When use the cofactor expansion, we want to apply it to rows/columns with more 0's

1.13 Lecture 13 - 06/16/2022

Theorem 1.13.1. Cofactor expansion across a row in A is the same as cofactor expansion across a column in A^T. From this obeservation we conclude inductively $\det(A) = \det(A^T)$

Definition 1.13.2. We say a square matrix A is upper triangular if it only has zeros to the left of the diagonal

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

A is lower triangular if it only has zeros to the right of the diagonal

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

A is diagonal if A only has nonzero entries on the diagonal

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

A diagonal matrix is both upper triangular and lower triangular

Exercise 1.13.3. Suppose $A = \begin{bmatrix} 2 & -1 & 3 & 1 \\ 0 & -2 & 0 & -1 \\ 0 & 0 & 3 & 1 \\ 0 & 0 & 0 & 1 \end{bmatrix}$. Please find use cofactor expansion to find the $\det A$

Solution. Note that A is upper triangular, so we could do cofactor expansions across first columns multiple times

$$\begin{vmatrix} 2 & -1 & 3 & 1 \\ 0 & -2 & 0 & -1 \\ 0 & 0 & 3 & 1 \\ 0 & 0 & 0 & 1 \end{vmatrix} = 2(-1)^{1+1} \begin{vmatrix} -2 & 0 & -1 \\ 0 & 3 & 1 \\ 0 & 0 & 1 \end{vmatrix} = 2 \cdot (-2)(-1)^{1+1} \begin{vmatrix} 3 & 1 \\ 0 & 1 \end{vmatrix} = 2 \cdot (-2) \cdot 3 \cdot (-1)^{1+1} 1 = -12$$

□

We could summerize this into the following

Theorem 1.13.4. If A is triangular, then $\det A$ is the product of the diagonal entries.

Exercise 1.13.5. Suppose I is the $n \times n$ identity matrix, what is $\det I$, $\det(-I)$, $\det(2I)$ and $\det(aI)$?

Solution. Note that I is a diagonal matrix. $\det I = 1$, $\det(-I) = (-1)^n$, $\det(2I) = 2^n$, and in general $\det(aI) = a^n$ □

As discussed in lecture 1.11, thinking in terms of signed volume, we know that

1. For column interchangement, determinant pick up a negtive sign

$$\det \begin{bmatrix} \mathbf{a}_1 & \mathbf{a}_3 & \mathbf{a}_2 \end{bmatrix} = -\det \begin{bmatrix} \mathbf{a}_1 & \mathbf{a}_2 & \mathbf{a}_3 \end{bmatrix}$$

2. For column scaling, determinant scales correspondingly

$$\det \begin{bmatrix} \mathbf{a}_1 & c\mathbf{a}_2 & \mathbf{a}_3 \end{bmatrix} = c \det \begin{bmatrix} \mathbf{a}_1 & \mathbf{a}_2 & \mathbf{a}_3 \end{bmatrix}$$

3. For column replacement, determinant doesn't change

$$\det \begin{bmatrix} \mathbf{a}_1 & \mathbf{a}_2 + c\mathbf{a}_1 & \mathbf{a}_3 \end{bmatrix} = \det \begin{bmatrix} \mathbf{a}_1 & \mathbf{a}_2 & \mathbf{a}_3 \end{bmatrix}$$

With the help of Theorem 1.13.1 we may conclude an algorithm for computing determinants using elementary row operations (instead elementary column operations discussed above). 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 $\det B = \det A$.

$$\det \begin{bmatrix} R1 + cR3 \\ R2 \\ R3 \end{bmatrix} = \det \begin{bmatrix} R1 \\ R2 \\ R3 \end{bmatrix}$$

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

$$\det \begin{bmatrix} R1 \\ R3 \\ R2 \end{bmatrix} = -\det \begin{bmatrix} R1 \\ R2 \\ R3 \end{bmatrix}$$

3. If one row of A is multiplied by k to produce B , then $\det B = k \det A$.

$$\det \begin{bmatrix} R1 \\ kR2 \\ R3 \end{bmatrix} = k \det \begin{bmatrix} R1 \\ R2 \\ R3 \end{bmatrix}$$

Example 1.13.6. Use elementary row operations to evaluate the following i.

$$\begin{vmatrix}1 & -1 & 1\\ 2 & 0 & -1\\ 1 & 1 & 1\end{vmatrix} \xlongequal[R3\to R3-R1]{R2\to R2-2R1} \begin{vmatrix}1 & -1 & 1\\ 0 & 2 & -3\\ 0 & 2 & 0\end{vmatrix} \xlongequal{\text{factor } R3} 2\begin{vmatrix}1 & -1 & 1\\ 0 & 2 & -3\\ 0 & 1 & 0\end{vmatrix} \xlongequal{R2\to R2-2R3} 2\begin{vmatrix}1 & -1 & 1\\ 0 & 0 & -3\\ 0 & 1 & 0\end{vmatrix}$$

$$\xlongequal{R2\leftrightarrow R3} (-1)\cdot 2\begin{vmatrix}1 & -1 & 1\\ 0 & 1 & 0\\ 0 & 0 & -3\end{vmatrix} = (-1)\cdot 2\cdot 1\cdot 1\cdot(-3) = 6$$

ii.

$$\begin{vmatrix}1 & 2 & 1\\ 2 & 0 & -1\\ -1 & 2 & 2\end{vmatrix} \xlongequal[R3\to R3+R1]{R2\to R2-2R1} \begin{vmatrix}1 & 2 & 1\\ 0 & -4 & -3\\ 0 & 4 & 3\end{vmatrix} \xlongequal{R3\to R3+R2} \begin{vmatrix}1 & 2 & 1\\ 0 & -4 & -3\\ 0 & 0 & 0\end{vmatrix} = 1\cdot(-4)\cdot 0 = 0$$

iii.

$$\begin{vmatrix}1 & 2 & 3 & 0\\ 0 & 3 & -1 & 0\\ -1 & 2 & 1 & 2\\ 2 & -3 & 1 & 0\end{vmatrix} \xlongequal{\text{cofactor expansion across last column}} 2(-1)^{3+4}\begin{vmatrix}1 & 2 & 3\\ 0 & 3 & -1\\ 2 & -3 & 1\end{vmatrix}$$

$$\xlongequal{R3\to R3-2R1} (-2)\begin{vmatrix}1 & 2 & 3\\ 0 & 3 & -1\\ 0 & -7 & -5\end{vmatrix} \xlongequal{\text{cofactor expansion across first column}} (-2)\cdot 1(-1)^{1+1}\begin{vmatrix}3 & -1\\ -7 & -5\end{vmatrix}$$

$$= (-2)(3(-5)-(-1)(-7)) = 44$$

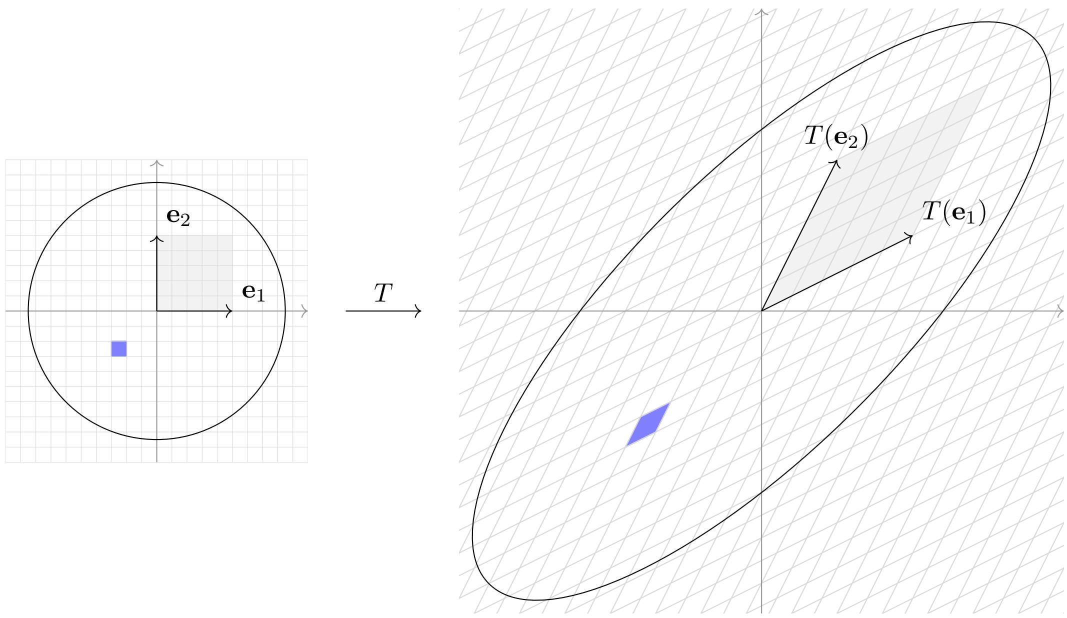
Remark. The REF of a square matrix A is upper triangular, and $\det A = 0$ if A has less then n pivots.

Theorem 1.13.7. Suppose A, B are $n\times n$ matrices, then $\det(AB) = (\det A)(\det B)$

If A is invertible, then A^{-1} is well-defined, then $1 = \det I = \det(AA^{-1}) = (\det A)(\det(A^{-1}) \Rightarrow \det(A^{-1}) = \dfrac{1}{\det A}$, so $\det A \neq 0$. Conversely, if $\det A \neq 0$, A would have n pivots, so a pivot in each row and column, thus A will be invertible. Therefore we have the following theorem

Theorem 1.13.8. *A is invertible $\iff$ $\det A \neq 0$. In addition, $\det(A^{-1}) = \dfrac{1}{\det A}$.*

Geometric meaning of the determinant of a linear transformation T (with standard matrix $A = \begin{bmatrix}T(\mathbf{e}_1) & T(\mathbf{e}_2)\end{bmatrix}$)



Since every small blue squares has been deformed into small parallelograms which are congruent to the gray square, parallelogram respectively and are of the same (signed) ratio $\det A$, so any shape under T gets scaled by $\det A$.

1.14 Lecture 14 - 06/22/2022

To motivate the definition of a vector space, let's consider the following example

Example 1.14.1. Let $\mathbb{P}_n$ denote the set of (real) polynomials of degree less or equal to n. For example $\mathbb{P}_0 = \mathbb{R}$ is just the set of real numbers, and

$$\begin{aligned}
\mathbb{P}_1 &= \{a_0 + a_1 t | a_0, a_1 \in \mathbb{R}\} \\
\mathbb{P}_2 &= \{a_0 + a_1 t + a_2 t^2 | a_0, a_1, a_2 \in \mathbb{R}\} \\
\mathbb{P}_3 &= \{a_0 + a_1 t + a_2 t^2 + a_3 t^3 | a_0, a_1, a_2, a_3 \in \mathbb{R}\} \\
&\vdots \\
\mathbb{P}_n &= \{a_0 + a_1 t + a_2 t^2 + \cdots + a_n t^n | a_0, a_1, a_2, \cdots, a_n \in \mathbb{R}\}.
\end{aligned}$$

You may soon realize that $\mathbb{P}_n$ can be identified with $\mathbb{R}^{n+1}$.

$$\begin{tikzcd}
\mathbb{P}_1 \arrow[r, equal, "\sim"] \arrow[d, equal] & \mathbb{R}^2 \arrow[d, equal] \\
\{a_0 + a_1 t\} \arrow[r, leftrightsquigarrow] & \left\{\begin{bmatrix} a_0 \\ a_1 \end{bmatrix}\right\}
\end{tikzcd}
\qquad
\begin{tikzcd}
\mathbb{P}_2 \arrow[r, equal, "\sim"] \arrow[d, equal] & \mathbb{R}^3 \arrow[d, equal] \\
\{a_0 + a_1 t + a_2 t^2\} \arrow[r, leftrightsquigarrow] & \left\{\begin{bmatrix} a_0 \\ a_1 \\ a_2 \end{bmatrix}\right\}
\end{tikzcd}$$

$$\begin{tikzcd}
\mathbb{P}_n \arrow[r, equal, "\sim"] \arrow[d, equal] & \mathbb{R}^{n+1} \arrow[d, equal] \\
\{a_0 + a_1 t + a_2 t^2 + \cdots + a_n t^n\} \arrow[r, leftrightsquigarrow] & \left\{\begin{bmatrix} a_0 \\ a_1 \\ \vdots \\ a_n \end{bmatrix}\right\}
\end{tikzcd}$$

More concrete examples could be

1. For $\mathbb{P}_1 \cong \mathbb{R}^2$, $1 + 2t \leftrightsquigarrow \begin{bmatrix} 1 \\ 2 \end{bmatrix}$
2. For $\mathbb{P}_2 \cong \mathbb{R}^3$, $3t^2 - 1 \leftrightsquigarrow \begin{bmatrix} -1 \\ 0 \\ 3 \end{bmatrix}$

If we consider addition and scalar multiplication, we have

$$\begin{bmatrix} 1 \\ 0 \\ 2 \end{bmatrix} + \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 3 \\ 1 \\ 2 \end{bmatrix} \qquad\qquad 2 \cdot \begin{bmatrix} 1 \\ 0 \\ 2 \end{bmatrix} = \begin{bmatrix} 2 \\ 0 \\ 4 \end{bmatrix}$$

$$(1 + 2t^2) + (2 + t) = 3 + t + 2t^2 \qquad\qquad 2 \cdot (1 + 2t^2) = 2 + 4t^2$$

So we may conclude that addition and scalar multiplication in $\mathbb{P}_n$ can be identically translated to addition and scalar multiplication in $\mathbb{R}^{n+1}$

Remark. We call $\{1, t, t^2, \cdots, t^n\}$ the *standard basis* of $\mathbb{P}_n$, corresponding to the standard basis for $\mathbb{R}^{n+1}$

Example 1.14.2. $\{1, t, t^2\}$ is the standard basis for $\mathbb{P}_2$, and

$$p(t) = a_0 + a_1 t + a_2 t^2 = a_0 \cdot 1 + a_1 \cdot t + a_2 \cdot t^2$$

Example 1.14.3. Let's denote $M_{m\times n}(\mathbb{R})$ the set of $m \times n$ matrices. For example

$$\begin{aligned}
M_{2\times 2}(\mathbb{R}) &= \left\{ \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \middle| a_{11}, a_{12}, a_{21}, a_{22} \in \mathbb{R} \right\} \\
M_{3\times 2}(\mathbb{R}) &= \left\{ \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \\ a_{31} & a_{32} \end{bmatrix} \middle| a_{11}, a_{12}, a_{21}, a_{22}, a_{31}, a_{32} \in \mathbb{R} \right\} \\
M_{2\times 3}(\mathbb{R}) &= \left\{ \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{bmatrix} \middle| a_{11}, a_{12}, a_{13}, a_{21}, a_{22}, a_{23} \in \mathbb{R} \right\} \\
&\vdots \\
M_{m\times n}(\mathbb{R}) &= \left\{ \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \middle| a_{11}, \cdots, a_{1n}, \cdots, a_{m1}, \cdots, a_{mn} \in \mathbb{R} \right\}.
\end{aligned}$$

You may realize that $M_{m\times n}(\mathbb{R})$ can be identified with $\mathbb{R}^{mn}$

$$\begin{tikzcd}
M_{2\times 2}(\mathbb{R}) \arrow[r, equal, "\sim"] \arrow[d, equal] & \mathbb{R}^4 \arrow[d, equal] \\
\left\{ \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \right\} \arrow[r, leftrightsquigarrow] & \left\{ \begin{bmatrix} a_{11} \\ a_{12} \\ a_{21} \\ a_{22} \end{bmatrix} \right\}
\end{tikzcd}
\qquad
\begin{tikzcd}
M_{3\times 2}(\mathbb{R}) \arrow[r, equal, "\sim"] \arrow[d, equal] & \mathbb{R}^6 \arrow[d, equal] \\
\left\{ \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \\ a_{31} & a_{32} \end{bmatrix} \right\} \arrow[r, leftrightsquigarrow] & \left\{ \begin{bmatrix} a_{11} \\ a_{12} \\ a_{21} \\ a_{22} \\ a_{31} \\ a_{32} \end{bmatrix} \right\}
\end{tikzcd}$$

$$\begin{tikzcd}
M_{2\times 3}(\mathbb{R}) \arrow[r, equal, "\sim"] \arrow[d, equal] & \mathbb{R}^6 \arrow[d, equal] \\
\left\{ \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{bmatrix} \right\} \arrow[r, leftrightsquigarrow] & \left\{ \begin{bmatrix} a_{11} \\ a_{12} \\ a_{13} \\ a_{21} \\ a_{22} \\ a_{23} \end{bmatrix} \right\}
\end{tikzcd}$$

$$\begin{tikzcd}
M_{m\times n}(\mathbb{R}) \arrow[r, equal, "\sim"] \arrow[d, equal] & \mathbb{R}^{mn} \arrow[d, equal] \\
\left\{ \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \right\} \arrow[r, leftrightsquigarrow] & \left\{ \begin{bmatrix} a_{11} \\ \vdots \\ a_{1n} \\ \vdots \\ a_{m1} \\ \vdots \\ a_{mn} \end{bmatrix} \right\}
\end{tikzcd}$$

In more concrete terms, addition and scalar multiplication can be identified as the following

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

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

So we may conclude that addition and scalar multiplication in $M_{m\times n}(\mathbb{R})$ can be identically translated to addition and scalar multiplication in $\mathbb{R}^{mn}$

Remark. We call $\{E_{ij}\}$ the *standard basis* of $M_{m\times n}(\mathbb{R})$, corresponding to the standard basis for $\mathbb{R}^{mn}$. Here E_{ij} is the $m \times n$ matrix that only has a single 1 in the (i, j)-th spot, but 0's elsewhere.

Example 1.14.4. $\left\{E_{11} = \begin{bmatrix}1 & 0\\ 0 & 0\end{bmatrix}, E_{12} = \begin{bmatrix}0 & 1\\ 0 & 0\end{bmatrix}, E_{21} = \begin{bmatrix}0 & 0\\ 1 & 0\end{bmatrix}, E_{22} = \begin{bmatrix}0 & 0\\ 0 & 1\end{bmatrix}\right\}$ is the standard basis for $M_{2\times 2}(\mathbb{R})$, and

$$\begin{aligned}\begin{bmatrix}a_{11} & a_{12}\\ a_{21} & a_{22}\end{bmatrix} &= \begin{bmatrix}a_{11} & 0\\ 0 & 0\end{bmatrix} + \begin{bmatrix}0 & a_{12}\\ 0 & 0\end{bmatrix} + \begin{bmatrix}0 & 0\\ a_{21} & 0\end{bmatrix} + \begin{bmatrix}0 & 0\\ 0 & a_{22}\end{bmatrix}\\ &= a_{11}\begin{bmatrix}1 & 0\\ 0 & 0\end{bmatrix} + a_{12}\begin{bmatrix}0 & 1\\ 0 & 0\end{bmatrix} + a_{21}\begin{bmatrix}0 & 0\\ 1 & 0\end{bmatrix} + a_{22}\begin{bmatrix}0 & 0\\ 0 & 1\end{bmatrix}\\ &= a_{11}E_{11} + a_{12}E_{12} + a_{21}E_{21} + a_{22}E_{22}\end{aligned}$$

Definition 1.14.5. A (real) vector space is a set V of objects, called *vectors*, on which are defined two operations, called *addition* $\boldsymbol{+}$ and *(left) scalar multiplication* $\bullet$, subject to axioms

0. $\boldsymbol{u} \boldsymbol{+} \boldsymbol{v}$ and $c \bullet \boldsymbol{v}$ are still in V
1. $\boldsymbol{u} \boldsymbol{+} \boldsymbol{v} = \boldsymbol{v} \boldsymbol{+} \boldsymbol{u}$
2. $(\boldsymbol{u} \boldsymbol{+} \boldsymbol{v}) \boldsymbol{+} \boldsymbol{w} = \boldsymbol{u} \boldsymbol{+} (\boldsymbol{v} \boldsymbol{+} \boldsymbol{w})$
3. There is a *zero vector* $\mathbf{0}$ such that $\boldsymbol{u} \boldsymbol{+} \mathbf{0} = \boldsymbol{u}$
4. For each $\boldsymbol{u}$ in V, there is a vector $\boldsymbol{-u}$ in V such that $\boldsymbol{u} \boldsymbol{+} (\boldsymbol{-u}) = \mathbf{0}$
5. $c \bullet (\boldsymbol{u} \boldsymbol{+} \boldsymbol{v}) = c \bullet \boldsymbol{u} \boldsymbol{+} c \bullet \boldsymbol{v}$
6. $(c + d) \bullet \boldsymbol{u} = c \bullet \boldsymbol{u} \boldsymbol{+} d \bullet \boldsymbol{u}$
7. $c \bullet (d \bullet \boldsymbol{u}) = (cd) \bullet \boldsymbol{u}$
8. $1 \bullet \boldsymbol{u} = \boldsymbol{u}$

Example 1.14.6. Set V to be $\mathbb{R}^n$, $\boldsymbol{+}$ to be addition $+$ for vectors, $\bullet$ to be scalar multiplication $\cdot$ for vectors, then this is a vector space

Example 1.14.7 (non-example)**.** Suppose $V = \mathbb{R}$, $a \boldsymbol{+} b = a + b + 1$, $c \bullet a = c \cdot a = ca$, we can check

0. $a \boldsymbol{+} b = a + b + 1 \in \mathbb{R}$, $c \bullet a = ca \in \mathbb{R}$
1. $a \boldsymbol{+} b = a + b + 1 = b + a + 1 = b \boldsymbol{+} a$
2. $(a \boldsymbol{+} b) \boldsymbol{+} c = (a + b + 1) + c + 1 = a + (b + c + 1) + 1 = a \boldsymbol{+} (b \boldsymbol{+} c)$
3. There is a *zero vector* $\mathbf{0} = -1$ such that $a \boldsymbol{+} \mathbf{0} = a + (-1) + 1 = a$
4. For each a, we have $\boldsymbol{-}a = -a - 2$ such that $a \boldsymbol{+} (\boldsymbol{-}a) = a + (-a - 2) + 1 = -1 = \mathbf{0}$

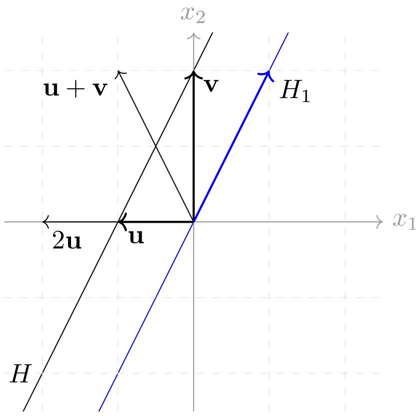
However $2 \bullet (a \boldsymbol{+} b) = 2(a + b + 1) \neq 2a + 2b + 1 = 2 \bullet a \boldsymbol{+} 2 \bullet b$. Therefore, this is not a vector space

1.15 Lecture 15 - 06/23/2022

Definition 1.15.1. Suppose V is a vector space with addition $\boldsymbol{+}$ and scalar multiplication $\boldsymbol{\cdot}$. A subspace H is a non-empty subset which closed under addition and scalar multiplication, i.e. for any $\boldsymbol{u}, \boldsymbol{v} \in H$, $c \in \mathbb{R}$, $\boldsymbol{u} \boldsymbol{+} \boldsymbol{v}, c \boldsymbol{\cdot} \boldsymbol{u} \in H$

Remark. It is easy to check that a subspace H is again a vector space.

Example 1.15.2. Consider the vector space $V = \mathbb{R}^2$, and H is the set of solutions to the linear equation $2x_1 - x_2 + 2 = 0$, then H is not a subspace. For example, if we choose $\mathbf{u} = \begin{bmatrix} -1 \\ 0 \end{bmatrix}$, $\mathbf{v} = \begin{bmatrix} 0 \\ 2 \end{bmatrix}$, then $\mathbf{u} + \mathbf{v} = \begin{bmatrix} -1 \\ 2 \end{bmatrix}$ is not in H, nor is $2\mathbf{u} = \begin{bmatrix} -2 \\ 0 \end{bmatrix}$



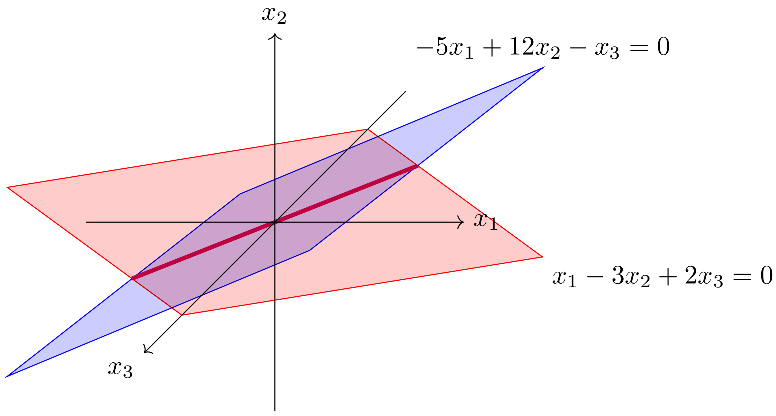
The reason is that H is not homogeneous. If we consider H_1 to be solution set of the homogeneous equation $2x_1 - x_2 = 0$, we see that H_1 is a subspace as it is the span of a single vector $\begin{bmatrix} 1 \\ 2 \end{bmatrix}$.

Definition 1.15.3. Suppose A is a $m \times n$ matrix, we define the null space of A to be $\operatorname{Nul} A = \{\mathbf{x} \in \mathbb{R}^n | A\mathbf{x} = \mathbf{0}\}$. Note that the solution set to linear system $A\mathbf{x} = \mathbf{0}$ is the intersection of m hyperplanes (one for each homogeneous equation) that pass through the origin.

Example 1.15.4. $A = \begin{bmatrix} 1 & -3 & 2 \\ -5 & 12 & -1 \end{bmatrix}$, the to find the $\operatorname{Nul} A$ is equivalent to solve $A\mathbf{x} = \mathbf{0}$

$$[A \quad \mathbf{0}] \xrightarrow{R2 \to R2+5R1} \begin{bmatrix} 1 & -3 & 2 & 0 \\ 0 & -3 & 9 & 0 \end{bmatrix} \xrightarrow{R2/(-3)} \begin{bmatrix} 1 & -3 & 2 & 0 \\ 0 & 1 & -3 & 0 \end{bmatrix} \xrightarrow{R1 \to R1+3R2} \begin{bmatrix} 1 & 0 & -7 & 0 \\ 0 & 1 & -3 & 0 \end{bmatrix}$$

Hence the solution set is $\begin{cases} x_1 = 7x_3 \\ x_2 = 3x_3 \\ x_3 \text{ is free} \end{cases}$, in parametric form, $\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = x_3 \begin{bmatrix} 7 \\ 3 \\ 1 \end{bmatrix}$, which describes a line in $\mathbb{R}^3$ that passes through the origin, and this line is the intersection of planes $x_1 - 3x_2 + 2x_3 = 0$ and $-5x_1 + 12x_2 - x_3 = 0$



Remark. As discussed in Example 1.15.2, in general, the solution set of $A\mathbf{x} = \mathbf{b}$ is not a subspace of $\mathbb{R}^n$ unless $\mathbf{b} = \mathbf{0}$. And in fact, any subspace of $\mathbb{R}^n$ is the null space for some $m \times n$ matrix A, i.e. the intersection of hyperplanes passing through the origin

1.16 Lecture 16 - 06/24/2022

Definition 1.16.1. Suppose $A = \begin{bmatrix} \mathbf{a}_1 & \mathbf{a}_2 & \cdots & \mathbf{a}_n \end{bmatrix}$ is an $m \times n$ matrix, then the column space (denote as $\operatorname{Col} A$) is the subspace $\operatorname{Span}\{\mathbf{a}_1, \mathbf{a}_2, \cdots, \mathbf{a}_n\}$ in $\mathbb{R}^m$. Suppose $A = \begin{bmatrix} R1 \\ R2 \\ \vdots \\ Rm \end{bmatrix}$, then the row space (denote as $\operatorname{Row} A$) is the subspace spaned by row vectors $\operatorname{Span}\{R1, R2, \cdots, Rm\}$ in $\mathbb{R}^n$ written horizontally.

Remark. Suppose column vectors $\mathbf{a}_1, \mathbf{a}_2, \cdots, \mathbf{a}_n$ in A has some linear dependence $A\mathbf{x} = \mathbf{0}$, then, after elementary row reduction $A \sim EA$, $(EA)\mathbf{x} = E(A\mathbf{x}) = E\mathbf{0} = \mathbf{0}$ again has the same linear dependence. In other words, the linear dependence of columns of A is preserved by row equivalence.

Remark. Since row elementary operations can be reversed, so $\operatorname{Row} A$ preserved under row equivalence

We may conclude the following algorithm for finding basis for $\operatorname{Nul} A$, $\operatorname{Col} A$, $\operatorname{Row} A$ by simply use elementary row reductions!!!

Theorem 1.16.2. Suppose A is a $m \times n$ matrix, $A \sim U$ is of RREF form

- The solution set of $\begin{bmatrix} U & \mathbf{0} \end{bmatrix}$ in parametric vector form gives a basis for $\operatorname{Nul} A$. Note that $\dim \operatorname{Nul} A =$ the number of free variables.
- A basis for $\operatorname{Col} A$ could be the set of pivot columns in A. Note that $\dim \operatorname{Col} A =$ the number of pivots
- A basis for $\operatorname{Row} A$ could be the set of non-zero row vectors in U (Or any REF of A actually). Note that $\dim \operatorname{Row} A =$ the number of pivots

Definition 1.16.3. $\dim \operatorname{Nul} A$ is also name the nullity of A. The number of pivots of A (which is equal to both $\dim \operatorname{Col} A$ and $\dim \operatorname{Row} A$) is called the rank of A

Theorem 1.16.4 (Rank-Nullity theorem)**.** Notice that the number of columns in A (say a $m \times n$ matrix) is equal to the number of free variables and the number of pivot columns, thus we have

$$n = \text{nullity} + \text{rank}$$

Example 1.16.5. $A = \begin{bmatrix} 1 & -1 & 1 \\ 2 & 0 & -1 \\ 1 & 1 & 2 \end{bmatrix} \sim \begin{bmatrix} 1 & -1 & 1 \\ 0 & 2 & -3 \\ 0 & 0 & 0 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & -\frac{1}{2} \\ 0 & 1 & -\frac{3}{2} \\ 0 & 0 & 0 \end{bmatrix}$, which is an REF and an RREF respectively. There is only one free variable x_3, so the nullity is 1, and the 1st, 2nd columns are pivot columns, so the rank is 2. We see that Theorem 1.16.4 holds as $3 = 1 + 2$, and

$$\operatorname{Nul} A = \operatorname{Span}\left\{ \begin{bmatrix} \frac{1}{2} \\ \frac{3}{2} \\ 1 \end{bmatrix} \right\}$$

$$\operatorname{Col} A = \operatorname{Span}\left\{ \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}, \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} \right\}$$

$$\operatorname{Row} A = \operatorname{Span}\left\{ \begin{bmatrix} 1 & 0 & -\frac{1}{2} \end{bmatrix}, \begin{bmatrix} 0 & 1 & -\frac{3}{2} \end{bmatrix} \right\} \text{ or } \operatorname{Span}\left\{ \begin{bmatrix} 1 & -1 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 2 & -3 \end{bmatrix} \right\}$$

Exercise 1.16.6. $A = \begin{bmatrix} 1 & 2 & 0 & 4 & 5 \\ 0 & 0 & 5 & -7 & 8 \\ 0 & 0 & 0 & 0 & -9 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \sim \begin{bmatrix} 1 & 2 & 0 & 4 & 0 \\ 0 & 0 & 1 & -\frac{7}{5} & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$ Note that here we have 2 free variables x_2, x_4, so the nullity is 2, and the 1st, 3rd, 5th columns are pivot columns, so the rank is 3. We see that Theorem 1.16.4 holds as $5 = 2 + 3$, and

$$\operatorname{Nul} A = \operatorname{Span}\left\{ \begin{bmatrix} -2 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} -4 \\ 0 \\ \frac{7}{5} \\ 1 \end{bmatrix} \right\}$$

$$\operatorname{Col} A = \operatorname{Span}\left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 5 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 5 \\ 8 \\ -9 \\ 0 \end{bmatrix} \right\}$$

$$\operatorname{Row} A = \operatorname{Span}\left\{ \begin{bmatrix} 1 & 2 & 0 & 4 & 5 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 5 & -7 & 8 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 0 & 0 & -9 \end{bmatrix} \right\}$$

Question. If you have a set $\mathcal{S}$ of vectors in $\mathbb{R}_m$, how do you find a subset of $\mathcal{S}$ that is a basis for $\operatorname{Span}\{\mathcal{S}\}$ (i.e. remove linear dependences)?

Answer. Collect these vectors as the column vectors of a matrix, and then find its columns space.

Exercise 1.16.7. Recall that $M_{2\times 2}(\mathbb{R})$ is the set of 2 by 2 matrices, and that a square matrix A is *symmetric* if $A^T = A$. Consider a subset V consists of 2 by 2 symmetric matrices, i.e. $V = \{A \in M_{2\times 2}(\mathbb{R}) | A^T = A\}$

1. Show that V is a vector space.

2. Find a basis for V.

Solution.

1. For any $A, B \in V$, $c \in \mathbb{R}$, by definition we know that $A^T = A$, $B^T = B$, we want to show that $A + B \in V$, $cA \in V$ (condition for subspace), i.e. $(A+B)^T = A^T + B^T$, $(cA)^T = cA$. This is true because

$$(A+B)^T = A^T + B^T = A + B, \qquad (cA)^T = cA^T = cA$$

Therefore V is a subspace of $M_{2\times 2}(\mathbb{R})$, and thus a vector space

2. Suppose $A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \in M_{2\times 2}(\mathbb{R})$, then $a_{12} = a_{21}$, so we may conclude that

$$V = \left\{ \begin{bmatrix} a & b \\ b & c \end{bmatrix} \in M_{2\times 2}(\mathbb{R}) \middle| a, b, c \in \mathbb{R} \right\}$$

Note that

$$\begin{bmatrix} a & b \\ b & c \end{bmatrix} = \begin{bmatrix} a & 0 \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} 0 & b \\ b & 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 0 & c \end{bmatrix} = a \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} + b \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} + c \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \tag{1.16.1}$$

And that linear combination (1.16.1) is the zero matrix $\iff$ $a = b = c = 0$, thus $\mathcal{B} = \left\{ \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \right\}$ is a basis for V

□

1.17 Lecture 17 - 06/27/2022

Definition 1.17.1. Suppose V, W are vector spaces, a *linear transformation* $T : V \to W$ is such that

- $T(\boldsymbol{u} + \boldsymbol{v}) = T(\boldsymbol{u}) + T(\boldsymbol{v})$
- $T(c \cdot u) = c \cdot T(\boldsymbol{u})$

Just as before, we call V the *domain* of T, W the *codomain* of V, the *image* of $\boldsymbol{u}$ under T is $T(\boldsymbol{u})$, the set of images $\{T(\boldsymbol{u})|\boldsymbol{u} \in V\}$ the *range* (denoted as $\operatorname{Range} T$), and the set $\{\boldsymbol{u}|T(\boldsymbol{w}) = \boldsymbol{u}\}$ the preimage of $\boldsymbol{w}$ under T. We still say that T is *one-to-one* if any $\boldsymbol{w} \in W$, there is at most one $\boldsymbol{u} \in V$ such that $T(\boldsymbol{u}) = \boldsymbol{w}$. T is *onto* the range is the codomain. T is said to be *invertible* if T has a inverse (this happens if and only if T is both one-to-one and onto), in this case we also call T an isomorphism.

Definition 1.17.2. We call $\{\boldsymbol{u}|T(\boldsymbol{u}) = \boldsymbol{0}\}$ the kernel (or null space) of T

Example 1.17.3. The identification

$$T : \mathbb{P}_2 \to \mathbb{R}^3, \qquad T(a_0 + a_1 t + a_2 t^2) = \begin{bmatrix} a_0 \\ a_1 \\ a_2 \end{bmatrix}$$

in Eaxmple 1.14.1 is an invertible linear transformation with inverse linear transformation

$$T^{-1} : \mathbb{R}^3 \to \mathbb{P}_2, \qquad T^{-1}\left(\begin{bmatrix} a_0 \\ a_1 \\ a_2 \end{bmatrix}\right) = a_0 + a_1 t + a_2 t^2$$

Example 1.17.4. The identification

$$T : M_{2\times 2}(\mathbb{R}) \to \mathbb{R}^4, \qquad T\left(\begin{bmatrix} a & b \\ c & d \end{bmatrix}\right) = \begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix}$$

in Eaxmple 1.14.3 is an invertible linear transformation with inverse linear transformation

$$T^{-1} : \mathbb{R}^4 \to M_{2\times 2}(\mathbb{R}), \qquad T^{-1}\left(\begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix}\right) = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

Theorem 1.17.5. Suppose $T : V \to W$ is a linear transformation between vector spaces, then

- $\ker T$ is a subspace of V.
- $\operatorname{Range} T$ is a subspace of W.

Proof.

- Suppose $\boldsymbol{u}, \boldsymbol{v} \in \ker T$, then by definition $T(\boldsymbol{u}) = \boldsymbol{0}$, $T(\boldsymbol{v}) = \boldsymbol{0}$, so $T(\boldsymbol{u} + \boldsymbol{v}) = T(\boldsymbol{u}) + T(\boldsymbol{v}) = \boldsymbol{0}$. And for any $c \in \mathbb{R}$, $T(c \cdot \boldsymbol{u}) = c \cdot T(\boldsymbol{u}) = \boldsymbol{0}$. In other words, we have shown that $\boldsymbol{u} + \boldsymbol{v}, c \cdot \boldsymbol{u} \in \ker T$, so $\ker T$ is a subspace.
- For any $T(\boldsymbol{u}), T(\boldsymbol{v}) \in \operatorname{Range} T$, $T(\boldsymbol{u}) + T(\boldsymbol{v}) = T(\boldsymbol{u} + \boldsymbol{v}) \in \operatorname{Range} T$, and for any $c \in \mathbb{R}$, $c \cdot T(\boldsymbol{u}) = T(c \cdot \boldsymbol{u}) \in \operatorname{Range} T$, so $\operatorname{Range} T$ is a subspace

$\square$

Remark. Suppose $T : \mathbb{R}^n \to \mathbb{R}^m$, $T(\mathbf{x}) = A\mathbf{x}$ is a matrix transformation, then $\ker T = \operatorname{Nul} A$, $\operatorname{Range} T = \operatorname{Col} A$

Example 1.17.6. Suppose $T : \mathbb{P}_2 \to \mathbb{R}$ takes the sum of coefficients, i.e. $T(a_0+a_1t+a_2t^2) = a_0+a_1+a_2$. T is a linear transformation, since for any $p(t) = a_0+a_1+a_2, q(t) = b_0+b_1t+b_2t^2 \in \mathbb{P}_2$, $c \in \mathbb{R}$, we have

$$\begin{aligned} T(p+q) &= T((a_0+b_0)+(a_1+b_1)t+(a_2+b_2)t^2) = (a_0+b_0)+(a_1+b_1)+(a_2+b_2) \\ &= (a_0+a_1+a_2)+(b_0+b_1+b_2) = T(p)+T(q) \end{aligned}$$

$$T(cp) = T((ca_0)+(ca_1)t+(ca_2)t^2) = (ca_0)+(ca_1)+(ca_2) = c(a_0+a_1+a_2) = cT(p)$$

So $V = \ker T$ is a subspace of V by Theorem 1.17.5

1.18 Lecture 18 - 06/28/2022

Theorem 1.18.1 (Unique representation theorem)**.** Suppose $\mathcal{B} = \{\boldsymbol{b}_1, \cdots, \boldsymbol{b}_n\}$ is a basis of a vector sapce V, then any vector $\boldsymbol{v} \in V$ can be uniquely represented as a linear combination $x_1 \bullet \boldsymbol{b}_1 \boldsymbol{+} \cdots \boldsymbol{+} x_n \bullet \boldsymbol{b}_n$

Remark. Here $[\boldsymbol{v}]_{\mathcal{B}} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$ is called the $\mathcal{B}$-coordinate vector (or coordinate vector relative to the basis $\mathcal{B}$) of $\boldsymbol{v}$

Proof. Suppose $\boldsymbol{v} = c_1 \bullet \boldsymbol{b}_1 \boldsymbol{+} \cdots \boldsymbol{+} c_n \bullet \boldsymbol{b}_n = d_1 \bullet \boldsymbol{b}_1 \boldsymbol{+} \cdots \boldsymbol{+} d_n \bullet \boldsymbol{b}_n$ are two linear combinations that express $\boldsymbol{v}$, then we have $(c_1 - d_1) \bullet \boldsymbol{b}_1 \boldsymbol{+} \cdots \boldsymbol{+} (c_n - d_n) \bullet \boldsymbol{b}_n = \mathbf{0}$, since $\{\boldsymbol{b}_1, \cdots, \boldsymbol{b}_n\}$ is linearly independent, we have $c_1 = d_1, \cdots, c_n = d_n$. Therefore the expression is unique. □

Definition 1.18.2. We call $[\quad]_{\mathcal{B}} : V \to \mathbb{R}^n$ the coordinate mapping

Theorem 1.18.3. The coordinate mapping $[\quad]_{\mathcal{B}} : V \to \mathbb{R}^n$ is a linear transformation

Example 1.18.4. The identifications in Example 1.14.1 and Example 1.14.3 are coordinate mappings $[\quad]_{\mathcal{E}} : \mathbb{P}_2 \to \mathbb{R}^3$ with standard basis $\mathcal{E} = \{1, t, t^2\}$ and $[\quad]_{\mathcal{E}} : M_{2\times 2}(\mathbb{R}) \to \mathbb{R}^4$ with standard basis $\mathcal{E} = \left\{ \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \right\}$, respectively

Example 1.18.5. $\mathcal{B} = \left\{ \mathbf{b}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \mathbf{b}_2 = \begin{bmatrix} 1 \\ 2 \end{bmatrix} \right\}$ is a basis for $\mathbb{R}^2$, $\mathbf{x} = \begin{bmatrix} 1 \\ 6 \end{bmatrix}$. Solve linear system $\mathbf{x} = x_1\mathbf{b}_1 + x_2\mathbf{b}_2 = \begin{bmatrix} \mathbf{b}_1 & \mathbf{b}_2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$ we get $[\mathbf{x}]_{\mathcal{B}} = \begin{bmatrix} -2 \\ 3 \end{bmatrix}$

Question. How should we study linear transformations via matrices in general?

Assume $T : V \to W$ is a linear transformation between vector spaces, $\{\boldsymbol{b}_1, \cdots, \boldsymbol{b}_n\}$ is a basis for V and $\{\boldsymbol{c}_1, \cdots, \boldsymbol{c}_m\}$ is a basis for W, then for any

$$\boldsymbol{v} = x_1 \bullet \boldsymbol{b}_1 \boldsymbol{+} \cdots \boldsymbol{+} x_n \bullet \boldsymbol{b}_n = \begin{bmatrix} \boldsymbol{b}_1 & \cdots & \boldsymbol{b}_n \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} \boldsymbol{b}_1 & \cdots & \boldsymbol{b}_n \end{bmatrix} [\boldsymbol{v}]_{\mathcal{B}}$$

We have

$$\begin{aligned} T(\boldsymbol{v}) &= T(x_1 \bullet \boldsymbol{b}_1 \boldsymbol{+} \cdots \boldsymbol{+} x_n \bullet \boldsymbol{b}_n) = x_1 \bullet T(\boldsymbol{b}_1) \boldsymbol{+} \cdots \boldsymbol{+} x_n \bullet T(\boldsymbol{b}_n) = x_1 \bullet T(\boldsymbol{b}_1) \boldsymbol{+} \cdots \boldsymbol{+} x_n \bullet T(\boldsymbol{b}_n) \\ &= \begin{bmatrix} T(\boldsymbol{b}_1) & \cdots & T(\boldsymbol{b}_n) \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} T(\boldsymbol{b}_1) & \cdots & T(\boldsymbol{b}_n) \end{bmatrix} [\boldsymbol{v}]_{\mathcal{B}} \end{aligned}$$

By Theorem 1.18.1, we can write

$$\begin{aligned} T(\boldsymbol{b}_1) &= a_{11} \bullet \boldsymbol{c}_1 + a_{21} \bullet \boldsymbol{c}_2 + \cdots + a_{m1} \bullet \boldsymbol{c}_m \\ T(\boldsymbol{b}_2) &= a_{12} \bullet \boldsymbol{c}_1 + a_{22} \bullet \boldsymbol{c}_2 + \cdots + a_{m2} \bullet \boldsymbol{c}_m \\ &\vdots \\ T(\boldsymbol{b}_n) &= a_{1n} \bullet \boldsymbol{c}_1 + a_{2n} \bullet \boldsymbol{c}_2 + \cdots + a_{mn} \bullet \boldsymbol{c}_m \end{aligned}$$

Therefore we have

$$\begin{bmatrix} T(\boldsymbol{b}_1) & \cdots & T(\boldsymbol{b}_n) \end{bmatrix} = \begin{bmatrix} \boldsymbol{c}_1 & \cdots & \boldsymbol{c}_m \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} = \begin{bmatrix} \boldsymbol{c}_1 & \cdots & \boldsymbol{c}_m \end{bmatrix} A \tag{1.18.1}$$

Where

$$A = \begin{bmatrix} [T(\boldsymbol{b}_1)]_{\mathcal{C}} & \cdots & [T(\boldsymbol{b}_n)]_{\mathcal{C}} \end{bmatrix} \tag{1.18.2}$$

is called the matrix of T relative to bases $\mathcal{B}$ and $\mathcal{C}$, thus

$$T(\boldsymbol{v}) = \begin{bmatrix} \boldsymbol{c}_1 & \cdots & \boldsymbol{c}_m \end{bmatrix} A[\boldsymbol{v}]_{\mathcal{B}}$$

On the other hand, we should have

$$T(\boldsymbol{v}) = \begin{bmatrix} \boldsymbol{c}_1 & \cdots & \boldsymbol{c}_m \end{bmatrix} [T(\boldsymbol{v})]_{\mathcal{C}}$$

So we may also conclude that

$$[T(\boldsymbol{v})]_{\mathcal{C}} = A[\boldsymbol{v}]_{\mathcal{B}} \tag{1.18.3}$$

The above discussion can be summerized by the following commutative diagram

$$\begin{tikzcd} V \arrow[r, "T"] \arrow[d, "{[\ \ ]_{\mathcal{B}}}"'] & W \arrow[d, "{[\ \ ]_{\mathcal{C}}}"] \\ \mathbb{R}^n \arrow[r, "A"] & \mathbb{R}^m \end{tikzcd} \qquad \begin{tikzcd} \boldsymbol{v} \arrow[r, mapsto] \arrow[d, mapsto] & T(\boldsymbol{v}) \arrow[d, mapsto] \\ {[\boldsymbol{v}]_{\mathcal{B}}} \arrow[r, mapsto] & A[\boldsymbol{v}]_{\mathcal{B}} = [T(\boldsymbol{v})]_{\mathcal{C}} \end{tikzcd} \tag{1.18.4}$$

Remark. The philosophy here is that any statement about a general linear transformation can be converted to a corresponding statement about matrix transformation.

Example 1.18.6. Suppose $V = \mathbb{R}^n$, $W = \mathbb{R}^m$ with bases $\{\mathbf{e}_1, \cdots, \mathbf{e}_n\}$, $\{\mathbf{e}_1, \cdots, \mathbf{e}_m\}$, then (1.18.1) reads $A = \begin{bmatrix} T(\mathbf{e}_1) & \cdots & T(\mathbf{e}_n) \end{bmatrix}$ is the standard matrix

Example 1.18.7. Consider linear transformation

$$T : \mathbb{P}_2 \to \mathbb{R}^3, \qquad T(a_0 + a_1 t + a_2 t^2) = \begin{bmatrix} a_0 \\ a_1 \\ a_2 \end{bmatrix}$$

and $\mathcal{B} = \{1+t, t+t^2, 1+t^2\}$ is a basis for $\mathbb{P}_2$, $\mathcal{E}$ is the standard basis for $\mathbb{R}^3$, then matrix of T relative to bases $\mathcal{B}$ and $\mathcal{E}$ can be read from (1.18.2) as

$$A = \begin{bmatrix} T(1+t) & T(t+t^2) & T(1+t^2) \end{bmatrix} = \begin{bmatrix} 1 & 0 & 1 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix}$$

Example 1.18.8. Consider linear transformation $T : \mathbb{R}^2 \to \mathbb{R}^2$ is defined by $T\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right) = \begin{bmatrix} x_1 + 2x_2 \\ x_2 \end{bmatrix}$, $\mathcal{B} = \left\{\begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ -1 \end{bmatrix}\right\}$, $\mathcal{C} = \left\{\begin{bmatrix} -1 \\ 1 \end{bmatrix}, \begin{bmatrix} -1 \\ -1 \end{bmatrix}\right\}$ are bases for $\mathbb{R}^2$. Then the matrix of T relative to bases $\mathcal{B}$ and $\mathcal{C}$ can be read from (1.18.1)

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

Remark. Suppose $T : V \to W$ and $S : W \to U$ are linear transformations between vector spaces, $\mathcal{B} = \{\boldsymbol{b}_1, \cdots, \boldsymbol{b}_n\}$ is a basis for V, $\mathcal{C} = \{\boldsymbol{c}_1, \cdots, \boldsymbol{c}_n\}$ is a basis for W and $\mathcal{D} = \{\boldsymbol{d}_1, \cdots, \boldsymbol{d}_n\}$ is a basis for U, then the matrix of $S \circ T$ relative to bases $\mathcal{B}$, $\mathcal{D}$ is BA.

$$\begin{tikzcd} V \arrow[r, "T"] \arrow[d, "{[\ \ ]_{\mathcal{B}}}"'] \arrow[rr, bend left, "S\circ T"] & W \arrow[r, "S"] \arrow[d, "{[\ \ ]_{\mathcal{C}}}"] & U \arrow[d, "{[\ \ ]_{\mathcal{D}}}"] \\ \mathbb{R}^n \arrow[r, "A"] \arrow[rr, bend right, "BA"'] & \mathbb{R}^m \arrow[r, "B"] & \mathbb{R}^p \end{tikzcd} \qquad \begin{tikzcd} \boldsymbol{v} \arrow[r, mapsto] \arrow[d, mapsto] & T(\boldsymbol{v}) \arrow[r] \arrow[d, mapsto] & S(T(\boldsymbol{v})) = (S\circ T)(\boldsymbol{v}) \arrow[d, mapsto] \\ {[\boldsymbol{v}]_{\mathcal{B}}} \arrow[r, mapsto] & A[\boldsymbol{v}]_{\mathcal{B}} = [T(\boldsymbol{v})]_{\mathcal{C}} \arrow[r, mapsto] & BA[\boldsymbol{v}]_{\mathcal{B}} = [S(T(\boldsymbol{v}))]_{\mathcal{D}} \end{tikzcd}$$

If T is invertible, then the matrix of T^{-1} relative to bases $\mathcal{C}$, $\mathcal{B}$ is A^{-1}

$$\begin{tikzcd} V \arrow[r, "T"] \arrow[d, "{[\ \ ]_{\mathcal{B}}}"'] & W \arrow[l, bend left, "T_{-1}"] \arrow[d, "{[\ \ ]_{\mathcal{C}}}"] \\ \mathbb{R}^n \arrow[r, "A"] & \mathbb{R}^m \arrow[l, bend left, "A^{-1}"] \end{tikzcd}$$

1.19 Lecture 19 - 06/29/2022

Now let's talk about change of basis. Suppose V is a vector space with two basis $\mathcal{B} = \{\boldsymbol{b}_1, \cdots, \boldsymbol{b}_n\}$ and $\mathcal{C} = \{\boldsymbol{c}_1, \cdots, \boldsymbol{c}_n\}$, and $\mathrm{id}_V : V \to V$, $\mathrm{id}_V(\boldsymbol{v}) = \boldsymbol{v}$ is the identity mapping. Diagram (1.18.4) becomes

$$\begin{tikzcd} V \arrow[r, equal, "\mathrm{id}_V"] \arrow[d, "{[\ \ ]_{\mathcal{B}}}"'] & V \arrow[d, "{[\ \ ]_{\mathcal{C}}}"] \\ \mathbb{R}^n \arrow[r, "\underset{\mathcal{C}\leftarrow\mathcal{B}}{P}"] & \mathbb{R}^n \end{tikzcd} \tag{1.19.1}$$

Where equation (1.18.2) and equation (1.18.3) become

$$\underset{\mathcal{C}\leftarrow\mathcal{B}}{P} = \begin{bmatrix} [\boldsymbol{b}_1]_{\mathcal{C}} & \cdots & [\boldsymbol{b}_n]_{\mathcal{C}} \end{bmatrix}$$

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

which is the matrix of id_V relative to basis $\mathcal{B}$ and $\mathcal{C}$, and we call this the change of coordinates matrix from $\mathcal{B}$ to $\mathcal{C}$. Also remark 1.18 gives us

$$\underset{\mathcal{D}\leftarrow\mathcal{B}}{P} = \left(\underset{\mathcal{D}\leftarrow\mathcal{C}}{P}\right)\left(\underset{\mathcal{C}\leftarrow\mathcal{B}}{P}\right), \quad \left(\underset{\mathcal{C}\leftarrow\mathcal{B}}{P}\right)^{-1} = \underset{\mathcal{B}\leftarrow\mathcal{C}}{P}$$

Example 1.19.1. Continue Example 1.16.7, we have shown that $\mathcal{B} = \left\{\begin{bmatrix}1&0\\0&0\end{bmatrix}, \begin{bmatrix}0&1\\1&0\end{bmatrix}, \begin{bmatrix}0&0\\0&1\end{bmatrix}\right\} = \{B_1, B_2, B_3\}$. Let's show that $\mathcal{C} = \left\{\begin{bmatrix}1&1\\1&0\end{bmatrix}, \begin{bmatrix}1&0\\0&1\end{bmatrix}, \begin{bmatrix}0&1\\1&1\end{bmatrix}\right\} = \{C_1, C_2, C_3\}$ is another basis for V. First note that

$$C_1 = \begin{bmatrix}1&1\\1&0\end{bmatrix} = \begin{bmatrix}1&0\\0&0\end{bmatrix} + \begin{bmatrix}0&1\\1&0\end{bmatrix} = B_1 + B_2$$

$$C_2 = \begin{bmatrix}1&0\\0&1\end{bmatrix} = \begin{bmatrix}1&0\\0&0\end{bmatrix} + \begin{bmatrix}0&0\\0&1\end{bmatrix} = B_1 + B_3$$

$$C_3 = \begin{bmatrix}0&1\\1&1\end{bmatrix} = \begin{bmatrix}0&1\\1&0\end{bmatrix} + \begin{bmatrix}0&0\\0&1\end{bmatrix} = B_2 + B_3$$

So $\{[C_1]_{\mathcal{B}}, [C_2]_{\mathcal{B}}, [C_3]_{\mathcal{B}}\} = \left\{\begin{bmatrix}1\\1\\0\end{bmatrix}, \begin{bmatrix}1\\0\\1\end{bmatrix}, \begin{bmatrix}0\\1\\1\end{bmatrix}\right\}$, and it is easy to show that this is a basis for $\mathbb{R}^3$, hence $\mathcal{C}$ is a basis for V. Also we know that

$$\underset{\mathcal{B}\leftarrow\mathcal{C}}{P} = \begin{bmatrix} [C_1]_{\mathcal{B}} & [C_2]_{\mathcal{B}} & [C_3]_{\mathcal{B}} \end{bmatrix} = \begin{bmatrix}1&1&0\\1&0&1\\0&1&1\end{bmatrix}$$

And

$$\underset{\mathcal{C}\leftarrow\mathcal{B}}{P} = \left(\underset{\mathcal{B}\leftarrow\mathcal{C}}{P}\right)^{-1} = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} & -\frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} & \frac{1}{2} \\ -\frac{1}{2} & \frac{1}{2} & \frac{1}{2} \end{bmatrix}$$

Note that here the diagram (1.19.1) becomes

$$\begin{tikzcd} V \arrow[r, equal, "\mathrm{id}_V"] \arrow[d, "{[\ ]_{\mathcal{B}}}"'] & V \arrow[d, "{[\ ]_{\mathcal{C}}}"] \\ \mathbb{R}^3 \arrow[r, "\underset{\mathcal{C}\leftarrow\mathcal{B}}{P}"] & \mathbb{R}^3 \end{tikzcd}$$

Example 1.19.2. Continue Example 1.17.6, let's find a basis for $V = \ker T$ which is supposed to correspond to $\operatorname{Nul} A$, where A is the matrix relative to both standard bases $\mathcal{E} = \{1, t, t^2\}$ for $\mathbb{P}_2$ and $\mathcal{E} = \left\{ \mathbf{e}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \mathbf{e}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \mathbf{e}_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \right\}$ which can be read from (1.18.3)

$$A = \begin{bmatrix} [T(1)]_{\mathcal{E}} & [T(t)]_{\mathcal{E}} & [T(t^2)]_{\mathcal{E}} \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 \end{bmatrix}$$

So we get

$$\operatorname{Nul} A = \operatorname{Span}\left\{ \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} \right\}$$

Which gives a basis $\{-1+t, -1+t^2\}$ for V. Note that here the diagram (1.18.4) becomes

$$\begin{tikzcd} \mathbb{P}_2 \arrow[r, "T"] \arrow[d, "{[\ ]_{\mathcal{E}}}"'] & \mathbb{R} \arrow[d, equal, "{[\ ]_{\mathcal{E}}}"] \\ \mathbb{R}^3 \arrow[r, "A"] & \mathbb{R} \end{tikzcd}$$

1.20 Lecture 20 - 06/30/2022

An algorithm for computing $A^{-1}B$

$$\left[\begin{array}{c|c} A & B \end{array}\right] \sim \left[\begin{array}{c|c} I & A^{-1}B \end{array}\right]$$

Exercise 1.20.1. Suppose $\mathcal{B} = \{2t^2 - 1, 3t + 1 - t^2, 3 - t\}$ and $\mathcal{C} = \{1 + t, t^2, -t\}$ are both bases for $\mathbb{P}_2$. Please find the change of basis matrix form $\mathcal{B}$ to $\mathcal{C}$

Solution. First let's find the change of basis matrices from $\mathcal{B}$ and $\mathcal{C}$ to the standard basis $\mathcal{E} = \{1, t, t^2\}$. We have

$$\underset{\mathcal{E}\leftarrow\mathcal{B}}{P} = \begin{bmatrix} [-1+2t^2]_{\mathcal{E}} & [1+3t-t^2]_{\mathcal{E}} & [3-t]_{\mathcal{E}} \end{bmatrix} = \begin{bmatrix} -1 & 1 & 3 \\ 0 & 3 & -1 \\ 2 & -1 & 0 \end{bmatrix}$$

And

$$\underset{\mathcal{E}\leftarrow\mathcal{C}}{P} = \begin{bmatrix} [1-t]_{\mathcal{E}} & [t]_{\mathcal{E}} & [t^2]_{\mathcal{E}} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 0 & -1 \\ 0 & 1 & 0 \end{bmatrix}$$

Hence we have

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

Which can computed via the above algorithm

$$\left[\begin{array}{ccc|ccc} 1 & 0 & 0 & -1 & 1 & 3 \\ 1 & 0 & -1 & 0 & 3 & -1 \\ 0 & 1 & 0 & 2 & -1 & 0 \end{array}\right] \sim \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & -1 & 1 & 3 \\ 0 & 1 & 0 & 2 & -1 & 0 \\ 0 & 0 & 1 & -1 & -2 & 4 \end{array}\right]$$

$$
\begin{tikzcd}
\mathbb{P}_2 \arrow[r, equal, "\mathrm{id}_V"] \arrow[d, "{[\ \ ]_{\mathcal{B}}}"'] & \mathbb{P}_2 \arrow[r, equal, "\mathrm{id}_V"] \arrow[d, "{[\ \ ]_{\mathcal{E}}}"'] & \mathbb{P}_2 \arrow[d, "{[\ \ ]_{\mathcal{C}}}"] \\
\mathbb{R}^3 \arrow[r, "\underset{\mathcal{E}\leftarrow\mathcal{B}}{P}"'] \arrow[rr, bend right, "\underset{\mathcal{C}\leftarrow\mathcal{B}}{P}"'] & \mathbb{R}^3 \arrow[r, bend left, "\underset{\mathcal{C}\leftarrow\mathcal{E}}{P}"] & \mathbb{R}^3 \arrow[l, "\underset{\mathcal{E}\leftarrow\mathcal{C}}{P}"]
\end{tikzcd}
$$

$\square$

Example 1.20.2. Suppose $\mathbb{R}^2$ has bases $\mathcal{B} = \left\{\mathbf{b}_1 = \begin{bmatrix} 7 \\ 5 \end{bmatrix}, \mathbf{b}_2 = \begin{bmatrix} -3 \\ -1 \end{bmatrix}\right\}$, $\mathcal{C} = \left\{\mathbf{c}_1 = \begin{bmatrix} 1 \\ -5 \end{bmatrix}, \mathbf{c}_2 = \begin{bmatrix} -2 \\ 2 \end{bmatrix}\right\}$. $\underset{\mathcal{C}\leftarrow\mathcal{B}}{P} = \left(\underset{\mathcal{E}\leftarrow\mathcal{C}}{P}\right)^{-1}\left(\underset{\mathcal{E}\leftarrow\mathcal{B}}{P}\right)$, which can be computed via

$$
\left[\begin{array}{cc|cc} 1 & -2 & 7 & -3 \\ -5 & 2 & 5 & -1 \end{array}\right] \sim \left[\begin{array}{cc|cc} 1 & 0 & -3 & 1 \\ 0 & 1 & -5 & 2 \end{array}\right]
$$

So $\underset{\mathcal{C}\leftarrow\mathcal{B}}{P} = \begin{bmatrix} -3 & 1 \\ -5 & 2 \end{bmatrix}$

$$
\begin{tikzcd}
\mathbb{R}^2 \arrow[r, equal] \arrow[d, "{[\ \ ]_{\mathcal{B}}}"'] & \mathbb{R}^2 \arrow[r, equal] \arrow[d, "{[\ \ ]_{\mathcal{E}}}"'] & \mathbb{R}^2 \arrow[d, "{[\ \ ]_{\mathcal{C}}}"] \\
\mathbb{R}^2 \arrow[r, "\underset{\mathcal{E}\leftarrow\mathcal{B}}{P}"'] \arrow[rr, bend right, "\underset{\mathcal{C}\leftarrow\mathcal{B}}{P}"'] & \mathbb{R}^2 \arrow[r, bend left, "\underset{\mathcal{C}\leftarrow\mathcal{E}}{P}"] & \mathbb{R}^2 \arrow[l, "\underset{\mathcal{E}\leftarrow\mathcal{C}}{P}"]
\end{tikzcd}
$$

Theorem 1.20.3. Let's generalize the rank-nullity theorem 1.16.4, with the remark 1.17, we have

$$\dim V = \dim \operatorname{Range} T + \dim \ker T$$

Definition 1.20.4. Suppose A is a square $n \times n$ matrix. A λ-eigenvector of an $n \times n$ matrix A is a nonzero vector $\mathbf{x}$ such that $A\mathbf{x} = \lambda\mathbf{x}$ for some scalar λ. A scalar λ is called an eigenvalue of A has λ-eigenvectors.

1.21 Lecture 21 - 07/01/2022

Definition 1.21.1. We say that a vector space V is trivial if $V = \{0\}$ is the zero vector space.

Question. How to decide whether a nonzero vector $\mathbf{x}$ is an eigenvector?

Answer. We can evaluate $A\mathbf{x}$ and see if it is a multiple of $\mathbf{x}$

Example 1.21.2. $A = \begin{bmatrix} 1 & 6 \\ 5 & 2 \end{bmatrix}$, determine whether $\mathbf{u} = \begin{bmatrix} 6 \\ -5 \end{bmatrix}$, $\mathbf{v} = \begin{bmatrix} 3 \\ -2 \end{bmatrix}$ are eigenvectors of A.

$$
A\mathbf{u} = \begin{bmatrix} 1 & 6 \\ 5 & 2 \end{bmatrix}\begin{bmatrix} 6 \\ -5 \end{bmatrix} = \begin{bmatrix} -24 \\ 20 \end{bmatrix} = (-4)\mathbf{u}
$$

Hence $\mathbf{u}$ is a (-4)-eigenvector of A.

$$
A\mathbf{v} = \begin{bmatrix} 1 & 6 \\ 5 & 2 \end{bmatrix}\begin{bmatrix} 3 \\ -2 \end{bmatrix} = \begin{bmatrix} -9 \\ 11 \end{bmatrix}
$$

Which is not a multiple of $\mathbf{v}$, so $\mathbf{v}$ is not an eigenvector.

Question. How to decide if λ is an eigenvalue of A?

Answer. By definition, we know that λ is an eigenvalue of $A \iff A\mathbf{x} = \lambda\mathbf{x}$ has a nontrivial solution $\iff \lambda\mathbf{x} - A\mathbf{x} = \lambda I\mathbf{x} - A\mathbf{x} = (\lambda I - A)\mathbf{x} = \mathbf{0}$ has a nontrivial solution $\iff \mathrm{Nul}(\lambda I - A)$ is nontrivial $\iff \lambda I - A$ is invertible $\iff \det(\lambda I - A) = 0$

Example 1.21.3. $A = \begin{bmatrix} 4 & -1 & 6 \\ 2 & 1 & 6 \\ 2 & -1 & 8 \end{bmatrix}$. Determine whether $\lambda = 2$ is an eigenvalue of A

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

Definition 1.21.4. From above criterion, we see that the set of λ-vectors is $\mathrm{Nul}(\lambda I - A)$ which is a subspace of $\mathbb{R}^n$, we call this the λ-eigenspace.

Example 1.21.5. Continue Example 1.21.3. Let's find a basis for the 2-eigenspace of A, which is equivalent of finding a basis for $\mathrm{Nul}(2I - A)$, since $\left[\begin{array}{c|c} 2I - A & \mathbf{0} \end{array}\right] \sim \left[\begin{array}{ccc|c} 1 & -\frac{1}{2} & 3 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right]$ Hence a basis could be $\left\{ \begin{bmatrix} \frac{1}{2} \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -3 \\ 0 \\ 1 \end{bmatrix} \right\}$

Question. How could we find all the eigenvalues of A?

Definition 1.21.6. From above discussion, we see that λ is an eigenvalue of $A \iff \det(\lambda I - A) = 0$, this motivates the following definition. We call $\det(tI - A)$ the characteristic polynomial of A, and $\det(tI - A) = 0$ the characteristic equation. And the roots of the characteristic polynomial are the eigenvalues of A.

Definition 1.21.7. We call the dimension of the λ-eigenspace ($\dim \mathrm{Nul}(\lambda I - A)$) the geometric multiplicity of λ, and the multiplicity of λ as a root in the charateristic polynomial $\det(tI - A)$ the algebraic multiplicity of λ.

Example 1.21.8. Suppose $A = \begin{bmatrix} 1 & -4 \\ 4 & 2 \end{bmatrix}$, then the characteristic polynomial of A would be

$$\det(tI - A) = \det\left(t\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} - \begin{bmatrix} 1 & -4 \\ 4 & 2 \end{bmatrix}\right) = \begin{vmatrix} t-1 & 4 \\ -4 & t-2 \end{vmatrix} = (t-1)(t-2) - 4\cdot(-4) = t^2 - 3t + 18$$

And the characteristic equation is $t^2 - 3t + 18 = 0$. Since $\Delta = (-3)^2 - 4\cdot 1 \cdot 18 = -63 < 0$, this equation has no (real) solutions, A doesn't have (real) eigenvalues

Note. Recall that the quadratic equation $ax^2 + bx + c = 0$ has no (real) solutions $\iff$ the discriminant $\Delta = b^2 - 4ac < 0$.

Example 1.21.9. Suppose $A = \begin{bmatrix} 3 & -2 \\ 2 & -1 \end{bmatrix}$, then the characteristic polynomial of A would be

$$\det(tI - A) = \det\left(t\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} - \begin{bmatrix} 3 & -2 \\ 2 & -1 \end{bmatrix}\right) = \begin{vmatrix} t-3 & 2 \\ -2 & t+1 \end{vmatrix} = (t-3)(t+1) - 2\cdot(-2) = t^2 - 2t + 1$$

And the characteristic equation is $t^2 - 2t + 1 = (t-1)^2 = 0$. Hence the eigenvalues of A is 2 with algebraic multiplicity 2

Example 1.21.10. Suppose $A = \begin{bmatrix} 3 & -2 \\ 2 & -1 \end{bmatrix}$, then the characteristic polynomial of A would be

$$\det(tI - A) = \begin{vmatrix} t-3 & 2 & -3 \\ 0 & t+1 & 0 \\ -6 & -7 & t+4 \end{vmatrix} \xlongequal{\text{cofactor expansion across the 2nd row}} (t+1)(-1)^{2+2}\begin{vmatrix} t-3 & -3 \\ -6 & t+4 \end{vmatrix}$$
$$= (t+1)((t-3)(t+4) - (-3)\cdot(-6)) = (t+1)(t^2 + t - 30) = (t+1)(t+6)(t-5)$$

So we see that the eigenvalues of A are $-1, -6, -5$, all with algebraic multiplicities 1,1,1.

Example 1.21.11. Suppose $A = \begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix}$, then characteristic polynomial is

$$\begin{aligned}
\det(tI - A) = \begin{vmatrix} t & -1 & 1 \\ -1 & t & -1 \\ -1 & -1 & t \end{vmatrix} &\xlongequal[R2\to R2-R3]{R2\to R1+t\cdot R2} \begin{vmatrix} 0 & -1-t & 1+t^2 \\ 0 & t+1 & -1-t \\ -1 & -1 & t \end{vmatrix} \\
&\xlongequal{\text{cofactor expansion across the 1st column}} (-1)(-1)^{3+1} \begin{vmatrix} -1-t & 1+t^2 \\ t+1 & -1-t \end{vmatrix} \\
&= (-1)((-1-t)^2 - (1+t^2)(t+1)) = (-1)(t-t^3) = t(t+1)(t-1)
\end{aligned}$$

Thus the eigenvalues are $0, 1, -1$, with algebraic multiplicities $1, 1, 1$.

Example 1.21.12. Suppose

$$A = \begin{bmatrix} \lambda_1 & * & \cdots & * \\ 0 & \lambda_2 & \cdots & * \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix}$$

is an $n \times n$ triangular matrix, then the characteristic polynomial of A is

$$\det(tI - A) = \begin{vmatrix} t-\lambda_1 & * & \cdots & * \\ 0 & t-\lambda_2 & \cdots & * \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & t-\lambda_n \end{vmatrix} = (t-\lambda_1)(t-\lambda_2)\cdots(t-\lambda_n)$$

so the eigenvalues are the diagonal elements $\lambda_1, \lambda_2, \cdots, \lambda_n$. Note that λ_i's might not be distinct, so there might be some multiplicities.

Example 1.21.13. Suppose $A = \begin{bmatrix} 1 & 2 & 0 & -1 \\ 0 & 2 & 1 & 3 \\ 0 & 0 & 3 & 4 \\ 0 & 0 & 0 & 1 \end{bmatrix}$, then the characteristic polynomial is

$$\det(tI - A) = \begin{vmatrix} t-1 & -2 & 0 & 1 \\ 0 & t-2 & -1 & -3 \\ 0 & 0 & t-3 & -4 \\ 0 & 0 & 0 & t-1 \end{vmatrix} = (t-1)^2(t-2)(t-3)$$

And the eigenvalues are $1, 2, 3$, with multiplicities $2, 1, 1$ respectively.

Definition 1.21.14. A and B are said to be similar if there exists an invertible matrix P such that $A = PBP^{-1}$. Note that this definition is *symmetric* in the sense that we also have $B = P^{-1}AP = (P^{-1})A(P^{-1})^{-1}$. Similarity is *transitive* in the sense that if A, B are similar, B, C are similar, then so do A, C. The reason is that suppose $A = PBP^{-1}$, $B = RCR^{-1}$, we would have $A = PBP^{-1} = PRCR^{-1}P^{-1} = (PR)C(PR)^{-1}$.

Definition 1.21.15. We can the mapping $A \mapsto PAP^{-1}$ a similar transformation

Theorem 1.21.16. Similar matrices have the same

- determinant
- characteristic polynomial

Proof. Suppose A, B are similar, $A = PBP^{-1}$, then

- $\det(A) = \det(PBP^{-1}) = \det(P)\det(B)\det(P^{-1}) = \det(P)\det(B)\det(P)^{-1} = \det(B)$
- Note that $tI - A = tPIP^{-1} - PBP^{-1} = P(tI - A)P^{-1}$, so $tI - A$ and $tI - B$ are similar, so they have the same determinant which is the characteristic polynomial.

□

1.22 Lecture 22 - 07/05/2022

Theorem 1.22.1 (diagonalization theorem)**.** Suppose $\mathbf{v}_1, \cdots, \mathbf{v}_n$ are eigenvectors of A with eigenvalues $\lambda_1, \cdots, \lambda_n$, then we have

$$A\begin{bmatrix}\mathbf{v}_1 & \cdots & \mathbf{v}_n\end{bmatrix} = \begin{bmatrix}A\mathbf{v}_1 & \cdots & A\mathbf{v}_n\end{bmatrix} = \begin{bmatrix}\lambda_1\mathbf{v}_1 & \cdots & \lambda_n\mathbf{v}_n\end{bmatrix} = \begin{bmatrix}\mathbf{v}_1 & \cdots & \mathbf{v}_n\end{bmatrix}\begin{bmatrix}\lambda_1 & \cdots & 0\\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_n\end{bmatrix}$$

Therefore we get a diagonalization of A, which reads $A = PDP^{-1}$, where $P = \begin{bmatrix}\mathbf{v}_1 & \cdots & \mathbf{v}_n\end{bmatrix}$ and $D = \begin{bmatrix}\lambda_1 & \cdots & 0\\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_n\end{bmatrix}$

Theorem 1.22.2. If $\mathbf{v}_1, \cdots, \mathbf{v}_r$ are eigenvectors that correspond to distinct eigenvalues $\lambda_1, \cdots, \lambda_r$ of an $n \times n$ matrix A, then the set $\{\mathbf{v}_1, \cdots, \mathbf{v}_r\}$ is linearly independent.

Proof. Suppose $\{\mathbf{v}_1, \cdots, \mathbf{v}_n\}$ is linearly independent, up to reordering the indices, we may assume $\mathbf{v}_r = c_1\mathbf{v}_1 + \cdots + c_{r-1}\mathbf{v}_{r-1}$ □

Exercise 1.22.3. Diagonalize $A = \begin{bmatrix}1 & 3 & 3\\ -3 & -5 & -3\\ 3 & 3 & 1\end{bmatrix}$

Solution. First we want to find all eigenvalues of A

$$\begin{aligned}
\det(tI - A) = \begin{vmatrix}t-1 & -3 & -3\\ 3 & t+5 & 3\\ -3 & -3 & t-1\end{vmatrix} &\xlongequal{R3 \to R3+R2} \begin{vmatrix}t-1 & -3 & -3\\ 3 & t+5 & 3\\ 0 & t+2 & t+2\end{vmatrix}\\
&\xlongequal{\text{factor out } (t+2) \text{ from row 3}} (t+2)\begin{vmatrix}t-1 & -3 & -3\\ 3 & t+5 & 3\\ 0 & 1 & 1\end{vmatrix}\\
&\xlongequal[R2\to R2-3R3]{R1\to R1+3R3} (t+2)\begin{vmatrix}t-1 & 0 & 0\\ 3 & t+2 & 0\\ 0 & 1 & 1\end{vmatrix} = (t+2)(t-1)(t+2)(1) = (t+2)^2(t-1)
\end{aligned}$$

So the eigenvalues of A are $\lambda_1 = 1$ with algebraic multiplicity 1 and $\lambda_2 = \lambda_3 = -2$ which is of algebraic multiplicity 2. For a basis of the 1-eigenspace $\mathrm{Nul}(I - A)$, consider

$$\left[\begin{array}{c|c} I - A & \mathbf{0}\end{array}\right] = \left[\begin{array}{ccc|c}0 & -3 & -3 & 0\\ 3 & 6 & 3 & 0\\ -3 & -3 & 0 & 0\end{array}\right] \sim \left[\begin{array}{ccc|c}1 & 0 & -1 & 0\\ 0 & 1 & 1 & 0\\ 0 & 0 & 0 & 0\end{array}\right]$$

So it could be $\left\{\mathbf{v}_1 = \begin{bmatrix}1\\-1\\0\end{bmatrix}\right\}$ For a basis of the (-2)-eigenspace $\mathrm{Nul}(-2I - A)$, consider

$$\left[\begin{array}{c|c} -2I - A & \mathbf{0}\end{array}\right] = \left[\begin{array}{ccc|c}-3 & -3 & -3 & 0\\ 3 & 3 & 3 & 0\\ -3 & -3 & -3 & 0\end{array}\right] \sim \left[\begin{array}{ccc|c}1 & 1 & 1 & 0\\ 0 & 0 & 0 & 0\\ 0 & 0 & 0 & 0\end{array}\right]$$

So it could be $\left\{\mathbf{v}_2 = \begin{bmatrix}-1\\1\\0\end{bmatrix}, \mathbf{v}_3 = \begin{bmatrix}-1\\0\\1\end{bmatrix}\right\}$. By Theorem 1.22.2, we know that $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ are linearly independent, thus forming a basis for $\mathbb{R}^3$. And we have a diagonalization

$$A = PDP^{-1}, \quad P = \begin{bmatrix}\mathbf{v}_1 & \mathbf{v}_2 & \mathbf{v}_3\end{bmatrix} = \begin{bmatrix}1 & -1 & -1\\ -1 & 1 & 0\\ 0 & 0 & 1\end{bmatrix}, \quad D = \begin{bmatrix}\lambda_1 & 0 & 0\\ 0 & \lambda_2 & 0\\ 0 & 0 & \lambda_3\end{bmatrix} = \begin{bmatrix}1 & 0 & 0\\ 0 & -2 & 0\\ 0 & 0 & -2\end{bmatrix}$$

□

1.23 Lecture 23 - 07/06/2022

Example 1.23.1. $A = \begin{bmatrix} 2 & 4 & 3 \\ -4 & -6 & -3 \\ 3 & 3 & 1 \end{bmatrix}$, the characteristic polynomial of A is

$$\det(tI - A) = \begin{vmatrix} t-2 & -4 & -3 \\ 4 & t+6 & 3 \\ -3 & -3 & t-1 \end{vmatrix} \xlongequal{R1 \to R1+R2} \begin{vmatrix} t+2 & t+2 & 0 \\ 4 & t+6 & 3 \\ -3 & -3 & t-1 \end{vmatrix}$$

$$\xlongequal{\text{factor } (t+2) \text{ from row 1}} (t+2) \begin{vmatrix} 1 & 1 & 0 \\ 4 & t+6 & 3 \\ -3 & -3 & t-1 \end{vmatrix} \xlongequal[R3 \to R3+3R1]{R2 \to R2-4R1} (t+2) \begin{vmatrix} 1 & 1 & 0 \\ 0 & t+2 & 3 \\ 0 & 0 & t-1 \end{vmatrix}$$

$$= (t-1)(t+2)^2$$

So the eigenvalues of A are $\lambda_1 = 1$ with algebraic multiplicity 1 and $\lambda_2 = \lambda_3 = -2$ which is of algebraic multiplicity 2. To find a basis for the 1-eigenspace, consider

$$\left[\begin{array}{c|c} I - A & \mathbf{0} \end{array}\right] = \left[\begin{array}{ccc|c} -1 & -4 & -3 & 0 \\ 4 & 7 & 3 & 0 \\ -3 & -3 & 0 & 0 \end{array}\right] \sim \left[\begin{array}{ccc|c} -1 & -4 & -3 & 0 \\ 0 & -9 & -9 & 0 \\ 0 & 9 & 9 & 0 \end{array}\right] \sim \left[\begin{array}{ccc|c} -1 & -4 & -3 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right]$$

$$\sim \left[\begin{array}{ccc|c} 1 & 0 & -1 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right]$$

It could be $\left\{ \mathbf{v}_1 = \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix} \right\}$. To find a basis for the (-2)-eigenspace, consider

$$\left[\begin{array}{c|c} -2I - A & \mathbf{0} \end{array}\right] = \left[\begin{array}{ccc|c} -4 & -4 & -3 & 0 \\ 4 & 4 & 3 & 0 \\ -3 & -3 & -3 & 0 \end{array}\right] \sim \left[\begin{array}{ccc|c} -4 & -4 & -3 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 \end{array}\right]$$

$$\sim \left[\begin{array}{ccc|c} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 \end{array}\right] \sim \left[\begin{array}{ccc|c} 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right]$$

It could be $\left\{ \mathbf{v}_2 = \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} \right\}$. Note that there is not enough eigenvectors $(2 < 3!!!)$ for a basis for $\mathbb{R}^3$. The real reason that this failed to be a basis was that the geometric multiplicity of -2 (which is equal to 1) is less that the algebraic multiplicity of -2 (which is equal to 2)

Theorem 1.23.2. Supoose A is an $n \times n$ matrix. Then the geometric multiplicity of an eigenvalue λ is less or equal to the algebraic multiplicity of λ. A is diagonalizable $\iff$ A has n linearly independent eigenvectors $\iff$ geometric multiplicities are always the same as algebraic multiplicities.

Theorem 1.23.3. Combining Theorem 1.22.2, we know that an $n \times n$ matrix with n distinct eigenvalues is diagonalizable.

Definition 1.23.4. Suppose $T : V \to V$ is a linear transformation, and $\mathcal{B} = \{\boldsymbol{b}_1, \cdots, \boldsymbol{b}_n\}$ is a basis for V, we can the matrix relative to $\mathcal{B}$ and $\mathcal{B}$ the $\mathcal{B}$-matrix of T (denoted $[T]_{\mathcal{B}}$), i.e. (1.18.4) becomes

$$\begin{tikzcd} V \arrow[r, "T"] \arrow[d, "{[\ ]_{\mathcal{B}}}"'] & V \arrow[d, "{[\ ]_{\mathcal{B}}}"] \\ \mathbb{R}^n \arrow[r, "{[T]_{\mathcal{B}}}"] & \mathbb{R}^n \end{tikzcd} \qquad \begin{tikzcd} \boldsymbol{v} \arrow[r, mapsto] \arrow[d, mapsto] & T(\boldsymbol{v}) \arrow[d, mapsto] \\ {[\boldsymbol{v}]_{\mathcal{B}}} \arrow[r, mapsto] & {[T]_{\mathcal{B}}[\boldsymbol{v}]_{\mathcal{B}} = [T(\boldsymbol{v})]_{\mathcal{B}}} \end{tikzcd}$$

And (1.18.2) gives $[T]_{\mathcal{B}} = \begin{bmatrix} [T(\boldsymbol{b}_1)]_{\mathcal{B}} & \cdots & [T(\boldsymbol{b}_n)]_{\mathcal{B}} \end{bmatrix}$ Suppose both $\mathcal{B} = \{\boldsymbol{b}_1, \cdots, \boldsymbol{b}_n\}$ and $\mathcal{C} = \{\boldsymbol{c}_1, \cdots, \boldsymbol{c}_n\}$ form bases for V, then

$$[T]_{\mathcal{B}} = \underset{\mathcal{B}\leftarrow\mathcal{C}}{P}[T]_{\mathcal{C}}\underset{\mathcal{C}\leftarrow\mathcal{B}}{P} = \underset{\mathcal{B}\leftarrow\mathcal{C}}{P}[T]_{\mathcal{C}}\left(\underset{\mathcal{B}\leftarrow\mathcal{C}}{P}\right)^{-1}$$

is similar via

$$\begin{tikzcd}
V \arrow[r, "T"] \arrow[d, "{[\ \ ]_{\mathcal{B}}}"'] \arrow[dd, bend right=60, "{[\ \ ]_{\mathcal{C}}}"'] & V \arrow[d, "{[\ \ ]_{\mathcal{B}}}"] \arrow[dd, bend left=60, "{[\ \ ]_{\mathcal{C}}}"] \\
\mathbb{R}^n \arrow[r, "{[T]_{\mathcal{B}}}"] & \mathbb{R}^n \arrow[d, "\underset{\mathcal{C}\leftarrow\mathcal{B}}{P}"'] \\
\mathbb{R}^n \arrow[r, "{[T]_{\mathcal{C}}}"] \arrow[u, "\underset{\mathcal{B}\leftarrow\mathcal{C}}{P}"'] & \mathbb{R}^n
\end{tikzcd}$$

If $T : \mathbb{R}^n \to \mathbb{R}^n$, $T(\mathbf{x}) = A\mathbf{x}$ is a matrix transformation, $\mathcal{E}$ is the standard basis, and $\mathcal{C} = \{\mathbf{v}_1, \cdots, \mathbf{v}_n\}$ is basis consists of eigenvectors, then $[T]_{\mathcal{C}} = D$ is diagonal with eigenvalues, as shown in Theorem 1.22.1.

Remark. In fact, any two similar matrix can be realized as matrices of the same linear transformation in different basis.

Definition 1.23.5. By Theorem 1.21.16 we know that similar matrices have the same characteristic polynomials (hence the same eigenvalues) and determinants. So we may define notions like *eigenvalues*, *eigenvectors*, *characteristic polynomials* and *determinants* for a linear transformation $T : V \to V$.

Example 1.23.6. Suppose $T : \mathbb{P}_2 \to \mathbb{P}_2$, $T(a_0 + a_1 t + a_2 t^2) = a_0 + 2a_1 t$ is a linear transformation (verify this!), $\mathcal{E} = \{1, t, t^2\}$ is the standard basis, we have

$$D = [T]_{\mathcal{B}} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

so the characteristic polynomial of T is $\det(tI - [T]_{\mathcal{E}}) = t(t-1)(t-2)$, so the eigenvalues for T will be $\lambda_1 = 0, \lambda_2 = 1, \lambda_3 = 2$, and the corresponding eigenvectors for $[T]_{\mathcal{E}}$ could be $\left\{ \mathbf{v}_1 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}, \mathbf{v}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \mathbf{v}_1 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \right\}$, which in turn corresponds to eigenvectors $\{t^2, 1, t\}$ for T. However if we choose basis $\mathcal{C} = \{1 + t, t, t^2\}$ is the standard basis, we have

$$P = \underset{\mathcal{E}\leftarrow\mathcal{C}}{P} = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad P^{-1} = \underset{\mathcal{C}\leftarrow\mathcal{E}}{P} = \begin{bmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

and thus

$$A = [T]_{\mathcal{C}} = \underset{\mathcal{B}\leftarrow\mathcal{C}}{P}[T]_{\mathcal{C}}\left(\underset{\mathcal{B}\leftarrow\mathcal{C}}{P}\right)^{-1} = PDP^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 0 \end{bmatrix}\begin{bmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ -1 & 2 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

The characteristic polynomial of T is $\det(tI - [T]_{\mathcal{E}}) = \det(tI - [T]_{\mathcal{C}}) = t(t-1)(t-2)$, so the eigenvalues for T will be $\lambda_1 = 0, \lambda_2 = 1, \lambda_3 = 2$, and the corresponding eigenvectors for $[T]_{\mathcal{E}}$ (or $[T]_{\mathcal{C}}$) could be $\left\{ \mathbf{v}_1 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}, \mathbf{v}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \mathbf{v}_1 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \right\}$, which in turn corresponds to eigenvectors $\{t^2, 1, t\}$ for T. The determinant of T is equal to either $\det([T]_{\mathcal{B}})$ or $\det([T]_{\mathcal{C}})$, which is zero.

Example 1.23.7. $A = \begin{bmatrix} 2 & 1 \\ 2 & 3 \end{bmatrix}$, let's compute A^{50}. First we find the eigenvalues $\lambda_1 = 1$, $\lambda_2 = 4$. and we get corresponding eigenvectors $\mathbf{v}_1 = \begin{bmatrix} -1 \\ 1 \end{bmatrix}, \mathbf{v}_2 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$, so we have $D = \begin{bmatrix} 1 & 0 \\ 0 & 4 \end{bmatrix}$ and

$P = \begin{bmatrix} -1 & 1 \\ 1 & 2 \end{bmatrix}$, so $P^{-1} = \frac{1}{3}\begin{bmatrix} -2 & 1 \\ 1 & 1 \end{bmatrix}$. So we have

$$\begin{aligned} A^{50} &= \overbrace{(PDP^{-1})(PDP^{-1})(PDP^{-1})\cdots(PDP^{-1})}^{50} \\ &= \overbrace{(PD\cancel{P^{-1}})(\cancel{P}D\cancel{P^{-1}})(\cancel{P}D\cancel{P^{-1}})\cdots(\cancel{P}DP^{-1})}^{50} \\ &= PD^{50}P^{-1} = \frac{1}{3}\begin{bmatrix} -1 & 1 \\ 1 & 2 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 0 & 4^{50} \end{bmatrix}\begin{bmatrix} -2 & 1 \\ 1 & 1 \end{bmatrix} \end{aligned}$$

1.24 Lecture 24 - 07/07/2022

Recall for $\mathbf{u} = \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix}, \mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix} \in \mathbb{R}^n$, we can define the *dot product* $\mathbf{u}\cdot\mathbf{v}$ to be $\mathbf{u}^T\mathbf{v} = \begin{bmatrix} u_1 & u_2 & \cdots & u_n \end{bmatrix}, \mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix} = u_1v_1 + u_2v_2 + \cdots u_nv_n$. The length(or norm) of a vector $\mathbf{u}$ can be expressed as $\|\mathbf{u}\| = \sqrt{\mathbf{u}\cdot\mathbf{u}} = \sqrt{u_1^2 + \cdots + u_n^2}$. Geometrically, the dot product can interpreted as $\mathbf{u}\cdot\mathbf{v} = \|\mathbf{u}\|\|\mathbf{v}\|\cos\theta$, here θ is the angle between $\mathbf{u}$ and $\mathbf{v}$ (which could be between 0 and π)

- If $\mathbf{u}\cdot\mathbf{v} < 0$, then $\cos\theta < 0$, θ is obtuse, if in addition $\mathbf{u}\cdot\mathbf{v} = \|\mathbf{u}\|\|\mathbf{v}\|$, then $\mathbf{u}, \mathbf{v}$ are of the opposite direction.
- If $\mathbf{u}\cdot\mathbf{v} = 0$, then $\cos\theta = 0$, $\theta = \frac{\pi}{2}$, or $\mathbf{u} = \mathbf{0}$, or $\mathbf{v} = \mathbf{0}$. In this case, we say that $\mathbf{u}, \mathbf{v}$ are orthogonal (denoted $\mathbf{u} \perp \mathbf{v}$, $\perp$ stands for perpendicular).
- If $\mathbf{u}\cdot\mathbf{v} > 0$, then $\cos\theta > 0$, θ is acute, if in addition $\mathbf{u}\cdot\mathbf{v} = \|\mathbf{u}\|\|\mathbf{v}\|$, then $\mathbf{u}, \mathbf{v}$ are of the same direction.

Definition 1.24.1. We say a vector $\mathbf{u}$ is of unit length (or a unit vector, or a normalized vector) if $\|\mathbf{u}\| = 1$

Definition 1.24.2. $\mathcal{B} = \{\mathbf{v}_1, \cdots, \mathbf{v}_n\}$ is called an orthogonal set if $\mathbf{v}_i, \mathbf{v}_j$ $(i \neq j)$ are orthogonal. $\mathcal{B}$ is called an orthogonal basis if $\mathcal{B}$ is in addition a basis. $\mathcal{B}$ is called an orthonormal basis if the basis vectors are in addition normalized.

Remark. Suppose $\mathcal{B} = \{\mathbf{v}_1, \cdots, \mathbf{v}_n\}$, to test if $\mathcal{B}$ is an orthogonal (or orthonormal) set, we just need to write $A = \begin{bmatrix} \mathbf{v}_1 & \cdots & \mathbf{v}_n \end{bmatrix}$, and test if

$$A^TA = \begin{bmatrix} \mathbf{v}_1^T \\ \mathbf{v}_2^T \\ \vdots \\ \mathbf{v}_n^T \end{bmatrix} \begin{bmatrix} \mathbf{v}_1 & \mathbf{v}_2 & \cdots & \mathbf{v}_n \end{bmatrix} = \begin{bmatrix} \mathbf{v}_1^T\mathbf{v}_1 & \mathbf{v}_1^T\mathbf{v}_2 & \cdots & \mathbf{v}_1^T\mathbf{v}_n \\ \mathbf{v}_2^T\mathbf{v}_1 & \mathbf{v}_2^T\mathbf{v}_2 & \cdots & \mathbf{v}_2^T\mathbf{v}_n \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{v}_n^T\mathbf{v}_1 & \mathbf{v}_n^T\mathbf{v}_2 & \cdots & \mathbf{v}_n^T\mathbf{v}_n \end{bmatrix}$$

is diagonal (or if $A^TA = I$)

Theorem 1.24.3. Suppose $\mathcal{B} = \{\mathbf{v}_1, \cdots, \mathbf{v}_p\}$ is an orthogonal set of non-zero vectors, then $\mathcal{B}$ is linearly independent

Proof. Suppose not, we may assume $\mathbf{0} \neq \mathbf{v}_p = c_1\mathbf{v}_1 + \cdots + r_{p-1}\mathbf{v}_{p-1}$, but then

$$0 < \|\mathbf{v}_p\|^2 = \mathbf{v}_p \cdot (c_1\mathbf{v}_1 + \cdots + r_{p-1}\mathbf{v}_{p-1}) = 0$$

Which is a contradiction. ☐

1.25 Lecture 25 - 07/11/2022

Question. Suppose you have two vectors $\mathbf{u}, \mathbf{v}$ (here $\mathbf{v} \neq \mathbf{0}$), what is the orthogonal projection of $\mathbf{u}$ onto $\mathbf{v}$ (Which we denote as $\operatorname{Proj}_{\mathbf{v}} \mathbf{u}$)?

Answer. First you realize that $\operatorname{Proj}_{\mathbf{v}} \mathbf{u}$ is parallel to $\mathbf{v}$, so we write it as $\lambda \mathbf{v}$, and we know $\|\operatorname{Proj}_{\mathbf{v}} \mathbf{u}\| = \lambda \|\mathbf{v}\| = \|\mathbf{u}\| \cos\theta$, so we may conclude that

$$\lambda = \frac{\|\mathbf{u}\| \cos\theta}{\|\mathbf{v}\|} = \frac{\|\mathbf{u}\|\|\mathbf{v}\| \cos\theta}{\|\mathbf{v}\|^2} = \frac{\mathbf{u} \cdot \mathbf{v}}{\mathbf{v} \cdot \mathbf{v}}$$

Therefore we have derived the equation

$$\operatorname{Proj}_{\mathbf{v}} \mathbf{u} = \frac{\mathbf{u} \cdot \mathbf{v}}{\mathbf{v} \cdot \mathbf{v}} \mathbf{v} \tag{1.25.1}$$

Question. If W is a subspace of $\mathbb{R}^n$ with an orthogonal basis $\mathcal{B} = \{\mathbf{u}_1, \cdots, \mathbf{u}_p\}$, how could you express the orthogonal projection of a vector $\mathbf{y} \in \mathbb{R}^n$ onto W (denoted by $\operatorname{Proj}_W \mathbf{y}$).

Suppose $\operatorname{Proj}_W \mathbf{y} = \lambda_1 \mathbf{u}_1 + \cdots + \lambda_p \mathbf{u}_p$, then we should have for any i

$$0 = (\mathbf{y} - \operatorname{Proj}_W \mathbf{y}) \cdot \mathbf{u}_i = (\mathbf{y} - \lambda_1 \mathbf{u}_1 + \cdots + \lambda_p \mathbf{u}_p) \cdot \mathbf{u}_i = \mathbf{y} \cdot \mathbf{u} - \lambda_i \mathbf{u}_1 \cdot \mathbf{u}_1 \Rightarrow \lambda_i = \frac{\mathbf{y} \cdot \mathbf{u}_i}{\mathbf{u}_i \cdot \mathbf{u}_i}$$

There we have the equation

$$\begin{aligned} \operatorname{Proj}_W \mathbf{y} &= \operatorname{Proj}_{\mathbf{u}_1} \mathbf{y} + \cdots + \operatorname{Proj}_{\mathbf{u}_p} \mathbf{y} \\ &= \frac{\mathbf{y} \cdot \mathbf{u}_1}{\mathbf{u}_1 \cdot \mathbf{u}_1} \mathbf{u}_1 + \cdots + \frac{\mathbf{y} \cdot \mathbf{u}_p}{\mathbf{u}_p \cdot \mathbf{u}_p} \mathbf{u}_p \end{aligned} \tag{1.25.2}$$

If we further assume that $\mathcal{B}$ is an orthonormal basis, then $\mathbf{u}_1 \cdot \mathbf{u}_1 = \|\mathbf{u}_i\|^2 = 1$. Let's write $u = \begin{bmatrix} \mathbf{u}_1 & \cdots & \mathbf{u}_p \end{bmatrix}$. (1.25.2) becomes

$$\operatorname{Proj}_W \mathbf{y} = (\mathbf{y} \cdot \mathbf{u}_1)\mathbf{u}_1 + \cdots + (\mathbf{y} \cdot \mathbf{u}_p)\mathbf{u}_p = UU^T \mathbf{y} \tag{1.25.3}$$

Example 1.25.1. Consider $\mathcal{B} = \left\{ \mathbf{u}_1 = \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix}, \mathbf{u}_2 = \begin{bmatrix} -1 \\ 2 \\ 0 \end{bmatrix} \right\}$, $\mathbf{y} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$, $W = \operatorname{Span}\{\mathbf{u}_1, \mathbf{u}_2\}$. Let $U = \begin{bmatrix} \mathbf{u}_1 & \mathbf{u}_2 \end{bmatrix} = \begin{bmatrix} 2 & -1 \\ 1 & 2 \\ 0 & 0 \end{bmatrix}$, then $A^T A = \begin{bmatrix} 5 & 0 \\ 0 & 5 \end{bmatrix}$ is diagonal, hence $\mathcal{B}$ is an orthogonal set, the orthogonal projection of $\mathbf{y}$ onto W is

$$\begin{aligned} \operatorname{Proj}_W \mathbf{y} &= \frac{\mathbf{y} \cdot \mathbf{u}_1}{\mathbf{u}_1 \cdot \mathbf{u}_1} \mathbf{u}_1 + \frac{\mathbf{y} \cdot \mathbf{u}_2}{\mathbf{u}_2 \cdot \mathbf{u}_2} \mathbf{u}_2 \\ &= \frac{1 \cdot 2 + 1 \cdot 1 + 1 \cdot 0}{2^2 + 1^2 + 0^2} \mathbf{u}_1 + \frac{1 \cdot (-1) + 1 \cdot 2 + 1 \cdot 0}{(-1)^2 + 2^2 + 0^2} \mathbf{u}_2 \\ &= \frac{3}{5} \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix} + \frac{1}{5} \begin{bmatrix} -1 \\ 2 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} \end{aligned}$$

1.26 Lecture 26 - 07/12/2022

Question. Suppose we are given arbitrary basis $\mathcal{B} = \{\mathbf{v}_1, \cdots, \mathbf{v}_p\}$ for a subspace W of $\mathbb{R}^n$, how could we get a orthogonal (or orthonormal) basis $\mathcal{U} = \{\mathbf{u}_1, \cdots, \mathbf{u}_p\}$ from it

Answer. We apply the Gram-Schmidt process

- $\mathbf{u}_1 = \mathbf{v}_1$
- $\begin{aligned} \mathbf{u}_2 &= \mathbf{v}_2 - \operatorname{Proj}_{\mathbf{u}_1} \mathbf{v}_2 \\ &= \mathbf{v}_2 - \frac{\mathbf{v}_2 \cdot \mathbf{u}_1}{\mathbf{u}_1 \cdot \mathbf{u}_1} \mathbf{u}_1 \end{aligned}$

- $\mathbf{u}_3 = \mathbf{v}_3 - \text{Proj}_{\mathbf{u}_1} \mathbf{v}_3 - \text{Proj}_{\mathbf{u}_2} \mathbf{v}_3$
  $= \mathbf{v}_3 - \dfrac{\mathbf{v}_3 \cdot \mathbf{u}_1}{\mathbf{u}_1 \cdot \mathbf{u}_1}\mathbf{u}_1 - \dfrac{\mathbf{v}_3 \cdot \mathbf{u}_2}{\mathbf{u}_2 \cdot \mathbf{u}_2}\mathbf{u}_2$
- $\vdots$
- $\mathbf{u}_p = \mathbf{v}_p - \text{Proj}_{\mathbf{u}_1} \mathbf{v}_p - \cdots - \text{Proj}_{\mathbf{u}_{p-1}} \mathbf{v}_p$
  $= \mathbf{v}_p - \dfrac{\mathbf{v}_p \cdot \mathbf{u}_1}{\mathbf{u}_1 \cdot \mathbf{u}_1}\mathbf{u}_1 - \cdots - \dfrac{\mathbf{v}_p \cdot \mathbf{u}_{p-1}}{\mathbf{u}_{p-1} \cdot \mathbf{u}_{p-1}}\mathbf{u}_{p-1}$

Then $\{\mathbf{u}_1, \cdots, \mathbf{u}_p\}$ would be an orthogonal basis. To get an orthonormal basis, just normalize these vectors.

Example 1.26.1. Consider $\left\{\mathbf{v}_1 = \begin{bmatrix}1\\1\\1\end{bmatrix}, \mathbf{v}_2 = \begin{bmatrix}1\\2\\2\end{bmatrix}, \mathbf{v}_3 = \begin{bmatrix}3\\-1\\1\end{bmatrix}\right\}$. Let's use Gram-Schmidt process to find an orthogonal (and an orthonormal) basis from it.

- First we choose $\mathbf{u}_1 = \mathbf{v}_1 = \begin{bmatrix}1\\1\\1\end{bmatrix}$
- $\mathbf{v}_2 - \dfrac{\mathbf{v}_2 \cdot \mathbf{u}_1}{\mathbf{u}_1 \cdot \mathbf{u}_1}\mathbf{u}_1 = \begin{bmatrix}1\\2\\2\end{bmatrix} - \frac{5}{3}\begin{bmatrix}1\\1\\1\end{bmatrix} = \begin{bmatrix}-\frac{2}{3}\\\frac{1}{3}\\\frac{1}{3}\end{bmatrix}$. Let's instead take a multiple of this to be our $\mathbf{u}_2$, namely we set $\mathbf{u}_2 = \begin{bmatrix}-2\\1\\1\end{bmatrix}$
- $\mathbf{u}_3 = \mathbf{v}_3 - \dfrac{\mathbf{v}_3 \cdot \mathbf{u}_1}{\mathbf{u}_1 \cdot \mathbf{u}_1}\mathbf{u}_1 - \dfrac{\mathbf{v}_3 \cdot \mathbf{u}_2}{\mathbf{u}_2 \cdot \mathbf{u}_2}\mathbf{u}_2 = \begin{bmatrix}3\\-1\\1\end{bmatrix} - \begin{bmatrix}1\\1\\1\end{bmatrix} - \begin{bmatrix}-2\\1\\1\end{bmatrix} = \begin{bmatrix}0\\-1\\1\end{bmatrix}$

Thus $\left\{\mathbf{u}_1 = \begin{bmatrix}1\\1\\1\end{bmatrix}, \mathbf{u}_2 = \begin{bmatrix}-2\\1\\1\end{bmatrix}, \mathbf{u}_3 = \begin{bmatrix}0\\-1\\1\end{bmatrix}\right\}$ is an orthogonal basis for $\mathbb{R}^3$. If we further normalize it, we have

$$\mathbf{w}_1 = \frac{\mathbf{u}_1}{\|\mathbf{u}_1\|} = \frac{1}{\sqrt{3}}\begin{bmatrix}1\\1\\1\end{bmatrix} = \begin{bmatrix}\frac{1}{\sqrt{3}}\\\frac{1}{\sqrt{3}}\\\frac{1}{\sqrt{3}}\end{bmatrix}, \mathbf{w}_2 = \frac{\mathbf{u}_2}{\|\mathbf{u}_2\|} = \frac{1}{\sqrt{6}}\begin{bmatrix}-2\\1\\1\end{bmatrix} = \begin{bmatrix}-\frac{1}{\sqrt{3}}\\\frac{1}{\sqrt{6}}\\\frac{1}{\sqrt{6}}\end{bmatrix}, \mathbf{w}_3 = \frac{\mathbf{u}_3}{\|\mathbf{u}_3\|} = \frac{1}{\sqrt{2}}\begin{bmatrix}0\\-1\\1\end{bmatrix} = \begin{bmatrix}0\\-\frac{1}{\sqrt{2}}\\\frac{1}{\sqrt{2}}\end{bmatrix}$$

is an orthonormal basis.

1.27 Lecture 27 - 07/13/2022

Definition 1.27.1. A square matrix U is called an orthogonal matrix if the columns of U form an orthonormal basis $\iff U^TU = UU^T = I$. Note that in this particular case, $U^T = U^{-1}$

Remark. The rows of U also form an orthonormal basis.

Note. The name *orthogonal matrix* is unfortunate since it is made out of orthonormal basis. It should really be called the orthonormal matrix. But since mathematicians has been using this for a long time. So we will stick to this usage.

The Gram-Schmidt process gives us the so-called QR-factorization. First we can rewrite the

Gram-Schmidt process as

$$
\begin{aligned}
\mathbf{v}_1 &= \mathbf{u}_1 \\
\mathbf{v}_2 &= \frac{\mathbf{v}_2 \cdot \mathbf{u}_1}{\mathbf{u}_1 \cdot \mathbf{u}_1}\mathbf{u}_1 + \mathbf{u}_2 \\
\mathbf{v}_3 &= \frac{\mathbf{v}_3 \cdot \mathbf{u}_1}{\mathbf{u}_1 \cdot \mathbf{u}_1}\mathbf{u}_1 + \frac{\mathbf{v}_3 \cdot \mathbf{u}_2}{\mathbf{u}_2 \cdot \mathbf{u}_2}\mathbf{u}_2 + \mathbf{u}_3 \\
&\vdots \\
\mathbf{v}_n &= \frac{\mathbf{v}_n \cdot \mathbf{u}_1}{\mathbf{u}_1 \cdot \mathbf{u}_1}\mathbf{u}_1 + \cdots + \frac{\mathbf{v}_n \cdot \mathbf{u}_{n-1}}{\mathbf{u}_{n-1} \cdot \mathbf{u}_{n-1}}\mathbf{u}_{n-1} + \mathbf{u}_n
\end{aligned}
$$

After normalization (let's set $\mathbf{u}_i := \frac{\mathbf{u}_i}{\|\mathbf{u}_i\|}$ to be unit vectors), we may write the above as the following set of equations with coefficients

$$
\begin{aligned}
\mathbf{v}_1 &= r_{11}\mathbf{u}_1 \\
\mathbf{v}_2 &= r_{12}\frac{\mathbf{v}_2 \cdot \mathbf{u}_1}{\mathbf{u}_1 \cdot \mathbf{u}_1}\mathbf{u}_1 + r_{22}\mathbf{u}_2 \\
\mathbf{v}_3 &= r_{13}\frac{\mathbf{v}_3 \cdot \mathbf{u}_1}{\mathbf{u}_1 \cdot \mathbf{u}_1}\mathbf{u}_1 + r_{23}\frac{\mathbf{v}_3 \cdot \mathbf{u}_2}{\mathbf{u}_2 \cdot \mathbf{u}_2}\mathbf{u}_2 + r_{33}\mathbf{u}_3 \\
&\vdots \\
\mathbf{v}_n &= r_{1n}\frac{\mathbf{v}_n \cdot \mathbf{u}_1}{\mathbf{u}_1 \cdot \mathbf{u}_1}\mathbf{u}_1 + \cdots + r_{n-1,n}\frac{\mathbf{v}_n \cdot \mathbf{u}_{n-1}}{\mathbf{u}_{n-1} \cdot \mathbf{u}_{n-1}}\mathbf{u}_{n-1} + r_{nn}\mathbf{u}_n
\end{aligned}
$$

If we write $A = \begin{bmatrix}\mathbf{v}_1 & \cdots & \mathbf{v}_n\end{bmatrix}$, $Q = \begin{bmatrix}\mathbf{u}_1 & \cdots & \mathbf{u}_n\end{bmatrix}$ (this is an orthogonal matrix), then have

$$A = \begin{bmatrix}\mathbf{v}_1 & \cdots & \mathbf{v}_n\end{bmatrix} \mp [\begin{bmatrix}\mathbf{u}_1 & \cdots & \mathbf{u}_n\end{bmatrix} = RQ$$

It can be shown that QR-factorization always exists, even A is not invertible (columns does not form a basis)

Example 1.27.2. Continuing Example 1.26.1, we consider the QR-factorization of $A = \begin{bmatrix}\mathbf{v}_1 & \mathbf{v}_2 & \mathbf{v}_3\end{bmatrix} = \begin{bmatrix}1 & 1 & 31 & 2 & -11 & 2 & 1\end{bmatrix}$, we know that $Q = \begin{bmatrix}\mathbf{w}_1 & \mathbf{w}_2 & \mathbf{w}_3\end{bmatrix} = \begin{bmatrix}\frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{3}} & 0 \\ \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{6}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{2}}\end{bmatrix}$, so we know that

$$R = Q^{-1}A = Q^TA = \begin{bmatrix}\frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} \\ -\frac{1}{\sqrt{3}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{6}} \\ 0 & -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}}\end{bmatrix}\begin{bmatrix}1 & 1 & 31 & 2 & -11 & 2 & 1\end{bmatrix} = \begin{bmatrix}\sqrt{3} & \frac{5}{\sqrt{3}} & \sqrt{3} \\ 0 & \frac{2}{\sqrt{6}} & -\sqrt{6} \\ 0 & 0 & \sqrt{2}\end{bmatrix}$$

Note that here we didn't actually use Gram-Schmidt process to evaluate R, instead we only used the existence of QR-factorization.

Definition 1.27.3. Suppose W is a subspace of $\mathbb{R}^n$, the orthogonal complement $W^\perp = \{\mathbf{v} | \mathbf{v} \perp W\}$ is the set of vectors that are orthogonal to W

Remark. It is easy to show that W is a subspace of $\mathbb{R}^n$ and $(W^\perp)^\perp = W$. This is kind of a *duality*, take $\mathbb{R}^3$ for an example, the orthogonal complement of a line would be the perpendicular plane, the orthogonal complement of a plane would be the perpendicular line, the orthogonal complement of the origin (a point) would be the whole space ($\mathbb{R}^3$) and the orthogonal complement of $\mathbb{R}^3$ would be the origin $\{\mathbf{0}\}$. From this we also see that $\dim W + \dim W^\perp = n$.

We make the following crucial observation: $\mathbf{x}$ is in $\operatorname{Nul} A \iff \mathbf{x}$ is orthogonal all row vectors of $A \iff \mathbf{x} \perp \operatorname{Row} A$, so we may conclude that $(\operatorname{Row} A)^\perp = \operatorname{Nul} A$, and similarly we know that $(\operatorname{Col} A)^\perp = (\operatorname{Row}(A^T))^\perp = \operatorname{Nul}(A^T)$

1.28 Lecture 28 - 07/14/2022

Question. linear system $A\mathbf{x} = \mathbf{b}$ may not always be solvable, nonetheless, we still want to find the *best possible* solution $\hat{\mathbf{x}}$

Answer. According to Gauss, by best possible we mean the least-square solutions. In concrete terms, such $\hat{\mathbf{x}}$ stisfies that $\|A\hat{\mathbf{x}} - \mathbf{b}\| \leq \|A\mathbf{x} - \mathbf{b}\|$ for any other choice of $\mathbf{x}$.

By a simple argument we can show that $A\hat{\mathbf{x}}$ must be the projection of $\mathbf{b}$ onto Col A, so we know that

$$(A\hat{\mathbf{x}} - \mathbf{b}) \perp \operatorname{Col} A \iff (A\hat{\mathbf{x}} - \mathbf{b}) \in (\operatorname{Col} A)^\perp = \operatorname{Nul}(A^T) \iff A^T(A\hat{\mathbf{x}} - \mathbf{b}) = \mathbf{0} \iff A^TA\hat{\mathbf{x}} = A^T\mathbf{b}$$

We call equation

$$A^TA\mathbf{x} = A^T\mathbf{b} \tag{1.28.1}$$

the normal equation of $A\mathbf{x} = \mathbf{b}$. And we have shown that the least square solutions of the linear system $A\mathbf{x} = \mathbf{b}$ is the solution set to the normal equation (1.28.1). The solution to the least square problem might not be unique.

Theorem 1.28.1. $A^TA\mathbf{x} = A^T\mathbf{b}$ has a unique solution iff the columns of A is linearly independent $\iff$ A^TA is invertible. And the unique solution would of course be $(A^TA)^{-1}A^T\mathbf{b}$

Example 1.28.2. Suppoes $A = \begin{bmatrix} 4 & 0 \\ 0 & 2 \\ 1 & 1 \end{bmatrix}$, $\mathbf{b} = \begin{bmatrix} 2 \\ 0 \\ 11 \end{bmatrix}$. It is easy to verify that the linear system $A\mathbf{x} = \mathbf{b}$ is not linearly consistent. To solve for the least-square solution, we look at its normal equation

$$\begin{bmatrix} 17 & 1 \\ 1 & 5 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 4 & 0 & 1 \\ 0 & 2 & 1 \end{bmatrix}\begin{bmatrix} 4 & 0 \\ 0 & 2 \\ 1 & 1 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = A^TA\mathbf{x} = A^T\mathbf{b} = \begin{bmatrix} 4 & 0 & 1 \\ 0 & 2 & 1 \end{bmatrix}\begin{bmatrix} 2 \\ 0 \\ 11 \end{bmatrix} = \begin{bmatrix} 19 \\ 11 \end{bmatrix}$$

And solve for $\mathbf{x}(A^TA)^{-1}A^T\mathbf{b} = (A^TA)^{-1}A^T\mathbf{b} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$

1.29 Lecture 29 - 07/15/2022

Question. Suppose you know a curve has the form

$$y = \beta_1 f_1(x_1, \cdots, x_k) + \beta_2 f_2(x_1, \cdots, x_k) + \cdots + \beta_n f_n(x_1, \cdots, x_k)$$

And you know what functions $f_1, f_2, \cdots, f_n$ should be (they could be linear functions, quadratic functions, polynomial functions, exponential functions, logarithmic function, trigonometry functions, etc.). You also have collected a bunch of data points $(x_1^{(1)}, \cdots, x_k^{(1)}, y^{(1)}), (x_1^{(2)}, \cdots, x_k^{(2)}, y^{(2)}), \cdots, (x_1^{(m)}, \cdots, x_k^{(m)}, y^{(m)})$. How do you find best fitting coefficients $\beta_1, \beta_2, \cdots, \beta_n$? First let's define the residual

$$\epsilon^{(i)} = y^{(i)} - \beta_1 f_1(x_1^{(i)}, \cdots, x_k^{(i)}) - \beta_2 f_2(x_1^{(i)}, \cdots, x_k^{(i)}) - \cdots - \beta_n f_n(x_1^{(i)}, \cdots, x_k^{(i)})$$

Our goal to minimize the residual. One might want to minimize the quantity $\sum_{i=1}^m |\epsilon^{(i)}|$, however, this is cumbersome, difficult to work with, and mathematically unsatisfactory. So Gauss instead considered the square sum $\sum_{i=1}^m |\epsilon^{(i)}|^2$, posed and solved the least square problem!!!

Answer. We use normal equations.

Let's write

$$\mathbf{y} = \begin{bmatrix} y^{(1)} \\ \vdots \\ y^{(m)} \end{bmatrix}, \beta = \begin{bmatrix} \beta_1 \\ \vdots \\ \beta_n \end{bmatrix}, \epsilon = \begin{bmatrix} \epsilon^{(1)} \\ \vdots \\ \epsilon^{(m)} \end{bmatrix}, X = \begin{bmatrix} f_1(x_1^{(1)}, \cdots, x_k^{(1)}) & \cdots & f_n(x_1^{(1)}, \cdots, x_k^{(1)}) \\ \vdots & \ddots & \vdots \\ f_1(x_1^{(m)}, \cdots, x_k^{(m)}) & \cdots & f_n(x_1^{(m)}, \cdots, x_k^{(m)}) \end{bmatrix}$$

Then have $\mathbf{y} = X\beta = \epsilon$. To minimize $\|\epsilon\|^2$ will be the same as solving the normal equation $X^TX\beta = X^T\mathbf{y}$ of $X\beta = \mathbf{y}$

1.30 Lecture 30 - 07/18/2022

Theorem 1.30.1. Consider quadratic equation $at^2 + bt + c = 0$, the quadratic formula reads $t = \frac{-b\pm\sqrt{\Delta}}{2a}$, where $\Delta = b^2 - 4ac$ is the discriminant.

Let's review some general stuff about complex numbers.

Definition 1.30.2. The set of complex numbers $\mathbb{C}$ is the defined to be $\{z = a + bi | a, b \in \mathbb{R}\}$, with $i = \sqrt{-1}$ so that $i^2 = -1$. We call $\operatorname{Re} z = a$ the real part of z and $\operatorname{Im} z = b$ to be the imaginary part of z. We call $r = |z| = \sqrt{a^2 + b^2}$ the modulus (or absolute value) of z, and the angle φ between z and the real axis the argument of z.

Definition 1.30.3. There is a natural identification between the complex numbers and the plane $\mathbb{R}^2$ via $z = a + bi \leftrightsquigarrow (a, b)$ (this is why the set of complex numbers is often called the complex plane). We can define

- Addition via $(a + bi) + (c + di) = (a + c) + (b + d)i$
- Multiplication via $(a + bi)(c + di) = ac + adi + bci + bdi^2 = (ac - bd) + (ad + bc)i$.
- Conjugate as $\overline{z} = a - bi$.

Remark. Through this identification it is easy to see that $a = r\cos\varphi$, $b = r\sin\varphi$, and we have Euler's formula $z = re^{i\varphi} = r\cos\varphi = i\sin\varphi$.

Remark. If we have a complex-valued matrix A, then the conjugation is defined entrywise (Note that column vectors are matrices). If we write a complex valued matrix $A = \operatorname{Re} A + i \operatorname{Im} A$, then the conjugation would be $\overline{A} = \operatorname{Re} A - i \operatorname{Im} A$. For example, if $A = \begin{bmatrix} 1+2i & 3 \\ -1-i & i \end{bmatrix} = \begin{bmatrix} 1 & 3 \\ -1 & 0 \end{bmatrix} + i\begin{bmatrix} 2 & 0 \\ -1 & 1 \end{bmatrix}$, then $\overline{A} = \begin{bmatrix} 1-2i & 3 \\ -1+i & -i \end{bmatrix} = \begin{bmatrix} 1 & 3 \\ -1 & 0 \end{bmatrix} - i\begin{bmatrix} 2 & 0 \\ -1 & 1 \end{bmatrix}$. It is easy to check that $\overline{A+B} = \overline{A} + \overline{B}$, $\overline{cA} = \overline{c}\overline{A}$, $\overline{AB} = \overline{AB}$.

Question. $A = \begin{bmatrix} 1 & -2 \\ 1 & 3 \end{bmatrix}$, can we diagonalize it?

Answer. First we compute its characteristic polynomial $t^2 - 4t + 5$, an then we can solve the quadratic equation as follows

$$\Delta = (-4)^2 - 4 \cdot 1 \cdot 5 = -4, t = \frac{-(-4) \pm \sqrt{-4}}{2} = \frac{-(-4) \pm 2i}{2} = 2 \pm i$$

Hence the eigenvalues are $\lambda_1 = 2 + i, \lambda_2 = 2 - i$. The eigenvector for λ_1 can be computed via

$$\left[\begin{array}{c|c} \lambda_1 I - A & \mathbf{0} \end{array}\right] \sim \left[\begin{array}{cc|c} 1 & 1-i & 0 \\ 0 & 0 & 0 \end{array}\right]$$

So we get $\mathbf{v}_1 = \begin{bmatrix} -1+i \\ 1 \end{bmatrix} = \begin{bmatrix} -1 \\ 1 \end{bmatrix} + \begin{bmatrix} i \\ 0 \end{bmatrix} = \begin{bmatrix} -1 \\ 1 \end{bmatrix} + \begin{bmatrix} 1 \\ 0 \end{bmatrix} i = \operatorname{Re} \mathbf{v}_1 + i \operatorname{Im} \mathbf{v}_1$. The eigenvector for λ_2 can be computed via

$$\left[\begin{array}{c|c} \lambda_2 I - A & \mathbf{0} \end{array}\right] \sim \left[\begin{array}{cc|c} 1 & 1+i & 0 \\ 0 & 0 & 0 \end{array}\right]$$

So we get $\mathbf{v}_2 = \begin{bmatrix} -1-i \\ 1 \end{bmatrix} = \begin{bmatrix} -1 \\ 1 \end{bmatrix} + \begin{bmatrix} -i \\ 0 \end{bmatrix} = \begin{bmatrix} -1 \\ 1 \end{bmatrix} + \begin{bmatrix} -1 \\ 0 \end{bmatrix} i = \operatorname{Re} \mathbf{v}_2 + i \operatorname{Im} \mathbf{v}_2$. Now we may realize that we have λ_1 and λ_2 are conjugate of each other. $\mathbf{v}_1$ and $\mathbf{v}_2$ are conjugate of each other.

Theorem 1.30.4. In general, if A is a 2 by 2 real-valued matrix with complex eigenvalues (characteristic polynomial have complex roots, no real roots), then they are conjugate of each other, we can write them as $\lambda = a - bi$ and $\overline{\lambda} = a + bi$, suppose the eigenvector for λ is $\mathbf{v} = \operatorname{Re} \mathbf{v} + i \operatorname{Im} \mathbf{v}$, then the eigenvector for $\overline{\lambda}$ will be $\overline{\mathbf{v}} = \operatorname{Re} \mathbf{v} - i \operatorname{Im} \mathbf{v}$. If we write $P = \begin{bmatrix} \operatorname{Re} \mathbf{v} & \operatorname{Im} \mathbf{v} \end{bmatrix}$, and $C = \begin{bmatrix} a & -b \\ b & a \end{bmatrix}$, then $AP = PC$, therefore we get decomposition $A = PCP^{-1}$.

Proof. We have $A\mathbf{v} = \lambda\mathbf{v}$ (and hence by conjugation we have $A\overline{\mathbf{v}} = \overline{A}\overline{\mathbf{v}} = \overline{A}\overline{\mathbf{v}} = \overline{A\mathbf{v}} = \overline{\lambda}\overline{\mathbf{v}}$, note that here A is real-valued, so $\overline{A} = A$), we can rewrite this as

$$(A\operatorname{Re}\mathbf{v}) + i(A\operatorname{Im}\mathbf{v}) = A(\operatorname{Re}\mathbf{v} + i\operatorname{Im}\mathbf{v}) = A\mathbf{v} = \lambda\mathbf{v} = (a - bi)(\operatorname{Re}\mathbf{v} + i\operatorname{Im}\mathbf{v}) = (a\operatorname{Re}\mathbf{v} + b\operatorname{Im}\mathbf{v}) + i(a\operatorname{Im}\mathbf{v} - b\operatorname{Re}\mathbf{v})$$

Looking at its real and imaginary parts we conclude that

$$A\operatorname{Re}\mathbf{v} = a\operatorname{Re}\mathbf{v} + b\operatorname{Im}\mathbf{v}, \qquad A\operatorname{Im}\mathbf{v} = a\operatorname{Im}\mathbf{v} - b\operatorname{Re}\mathbf{v}$$

This is precisely $AP = PC$ □

Remark. Matrix C is special in the following sense (it can be decomposed as a composition of rotation and scaling)

$$C = \begin{bmatrix} a & -b \\ b & a \end{bmatrix} = \begin{bmatrix} r\cos\varphi & -r\sin\varphi \\ r\sin\varphi & r\cos\varphi \end{bmatrix} = r\begin{bmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{bmatrix}$$

Here we suppose $\overline{\lambda} = a + bi = re^{i\varphi} = r\cos\varphi + ir\sin\varphi$ so that $a = r\cos\varphi$, $b = r\sin\varphi$.

Example 1.30.5. In the previous example, we could take $\lambda = 2 - i$ so that $a = 2, b = 1$, $\mathbf{v} = \begin{bmatrix} -1 \\ 1 \end{bmatrix} + i\begin{bmatrix} -1 \\ 0 \end{bmatrix}$ so that $\operatorname{Re}\mathbf{v} = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$, $\operatorname{Im}\mathbf{v} = \begin{bmatrix} -1 \\ 0 \end{bmatrix}$, and we have the decomposition

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

1.31 Lecture 31 - 07/19/2022

Theorem 1.31.1. Suppose A is a symmetric ($A^T = A$) real-valued matrix, and $\mathbf{v}_1$, $\mathbf{v}_2$ are λ_1-eigenvector, λ_2-eigenvectors respectively. Then $\mathbf{v}_1 \cdot \mathbf{v}_2 = 0$

Proof. Consider

$$\lambda_1(\mathbf{v}_1 \cdot \mathbf{v}_2) = (\lambda_1\mathbf{v}_1)^T\mathbf{v}_2 = (A\mathbf{v}_1)^T\mathbf{v}_2 = \mathbf{v}_1^TA^T\mathbf{v}_2 = \mathbf{v}_1^TA\mathbf{v}_2 = \mathbf{v}_1 \cdot (\lambda_2\mathbf{v}_2) = \lambda_2(\mathbf{v}_1 \cdot \mathbf{v}_2)$$

We get $(\lambda_1 - \lambda_2)(\mathbf{v}_1 \cdot \mathbf{v}_2) = 0$, since $\lambda_1 - \lambda_2 \neq 0$, $\mathbf{v}_1 \cdot \mathbf{v}_2 = 0$ □

Theorem 1.31.2. Suppose A is a symmetric real-valued matrix, then the eigenvalues are also real.

Proof. Suppose λ is an eigenvalue of A, then there exists some λ-eigenvector such that $A\mathbf{v} = \lambda\mathbf{v}$, and that the length of the complex vector $\mathbf{v}$ is $\|\mathbf{v}\|^2 = \overline{\mathbf{v}} \cdot \mathbf{v}$ is a positive real number (Note that for a complex number $z = a + bi$, $|z|^2 = a^2 + b^2 = (a + bi)(a - bi) = z\overline{z}$). Since A is symmetric and real-valued, $\overline{A}^T = A$. We have

$$\overline{\lambda}\|\mathbf{v}\|^2 = \overline{(A\mathbf{v})}^T\mathbf{v} = \overline{\mathbf{v}}^T\overline{A}^T\mathbf{v} = \overline{\mathbf{v}}^TA\mathbf{v} = \lambda\|\mathbf{v}\|^2$$

Which implies that $(\lambda - \overline{\lambda})\|\mathbf{v}\|^2 = 0$, so $\lambda = \overline{\lambda}$, i.e. λ is real-valued. □

Fact 1.31.3. A symmetric real-valued matrix is diagonalizable.

Theorem 1.31.4. Suppose A is a symmetric real-valued matrix with eigenvalues $\lambda_1, \cdots, \lambda_n$ (maybe repeated). A can be orthogonal diagonalized as $A = PDP^T$, where D is the diagonal matrix $\begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix}$, and P is an orthonormal matrix.

Proof. To get an orthonormal basis for each eigenspace $\operatorname{Nul}(\lambda I - A)$, you just need to find an arbitrary basis, and then apply Gram-Schmidt process to get an orthonormal basis, then by Theorem 1.31.1, the set of eigenvectors form an orthonormal basis for $\mathbb{R}^n$, assume they are $\mathbf{u}_1, \cdots, \mathbf{u}_n$, in corresponds to eigenvalues $\lambda_1, \cdots, \lambda_n$, then we get orthogonal diagonalization

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

Here $P = \begin{bmatrix} \mathbf{u}_1 & \cdots & \mathbf{u}_n \end{bmatrix}$ is an orthogonal basis and thus $P^{-1} = P^T$. □

Example 1.31.5. Suppose $A = \begin{bmatrix} 7 & 2 \\ 2 & 4 \end{bmatrix}$, the characteristic polynomial is $(t-3)(t-8)$, so we have eigenvalues $\lambda_1 = 3$ and $\lambda_2 = 8$, we can then find the eigenvectors $\mathbf{v}_1 = \begin{bmatrix} -1 \\ 2 \end{bmatrix}$ and $\mathbf{v}_2 = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$, we may realized that $\mathbf{v}_1 \cdot \mathbf{v}_2 = 0$, i.e. $\mathbf{v}_1$ is orthogonal to $\mathbf{v}_2$, we can further normalize them into $\mathbf{u}_1 = \frac{\mathbf{v}_1}{\|\mathbf{v}_1\|} = \begin{bmatrix} -\frac{1}{\sqrt{5}} \\ \frac{2}{\sqrt{5}} \end{bmatrix}$ and $\mathbf{u}_2 = \frac{\mathbf{v}_2}{\|\mathbf{v}_2\|} = \begin{bmatrix} \frac{2}{\sqrt{5}} \\ \frac{1}{\sqrt{5}} \end{bmatrix}$. So we have orthogonal diagonalization

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

Theorem 1.31.6. Suppose A is a symmetric real-valued matrix, and $A = PDP^T$ is its orthogonal diagonalization, then we have the so-called spectral decomposition

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

Proof.

$$A = PDP^T = \begin{bmatrix} \mathbf{u}_1 & \cdots & \mathbf{u}_n \end{bmatrix} \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix} \begin{bmatrix} \mathbf{u}_1^T \\ \vdots \\ \mathbf{u}_n^T \end{bmatrix} = \begin{bmatrix} \lambda_1 \mathbf{u}_1 & & \\ & \ddots & \\ & & \lambda_n \mathbf{u}_n \end{bmatrix} \begin{bmatrix} \mathbf{u}_1^T \\ \vdots \\ \mathbf{u}_n^T \end{bmatrix} = \lambda_1 \mathbf{u}_1 \mathbf{u}_1^T + \cdots + \lambda_n \mathbf{u}_n \mathbf{u}_n^T$$

□

Remark. Recall that if $\mathbf{u}$ is of unit length, then $\operatorname{Proj}_{\mathbf{u}} \mathbf{x} = \mathbf{u}\mathbf{u}^T \mathbf{x}$, so

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

You can think of this as decompose the matrix transformation $\mathbf{x} \mapsto A\mathbf{x}$ into the sum of scaled orthogonal projections.

1.32 Lecture 32 - 07/20/2022

Let's briefly talk about quadratic forms: $ax_1 + 2bx_1x_2 + cx_2^2 = \begin{bmatrix} x_1 & x_2 \end{bmatrix} \begin{bmatrix} a & b \\ b & c \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \mathbf{x}^T A \mathbf{x}$. In Example 1.31.5, $\mathbf{x}^T A \mathbf{x}$ is the quadratic form $7x_1^2 + 4x_1x_2 + 4x_2^2$, note that $\mathbf{y} = P^T \mathbf{x}$ gives change of (orthonormal) coordinates (actually differ by a rotation) $\begin{cases} y_1 &= -\frac{1}{\sqrt{5}}x_1 + \frac{1}{\sqrt{5}}x_2 \\ y_2 &= \frac{2}{\sqrt{5}}x_1 + \frac{1}{\sqrt{5}}x_2 \end{cases}$, then the quadratic form becomes $\mathbf{x}^T A \mathbf{x} = \mathbf{x}^T PDP^T \mathbf{x} = (P^T\mathbf{x})^T D (P^T \mathbf{x}) = \mathbf{y}^T D \mathbf{y} = 3y_1^2 + 8y_2^2$, note this is without cross term $y_1 y_2$.

Example 1.32.1. Suppose $A = \begin{bmatrix} 3 & -2 & 4 \\ -2 & 6 & 2 \\ 4 & 2 & 3 \end{bmatrix}$, then characteristic polynomial is

$$\begin{aligned} \det(tI - A) &= \begin{vmatrix} t-3 & 2 & -4 \\ 2 & t-6 & -2 \\ -4 & -2 & t-3 \end{vmatrix} \xlongequal{R3 \to R3 + 2R2} \begin{vmatrix} t-3 & 2 & -4 \\ 2 & t-6 & -2 \\ 0 & 2t-14 & t-7 \end{vmatrix} \\ &\xlongequal{\text{factor } (t-7) \text{ from 3rd row}} (t-7) \begin{vmatrix} t-3 & 2 & -4 \\ 2 & t-6 & -2 \\ 0 & 2 & 1 \end{vmatrix} \xlongequal{C2 \to C2 - C3} (t-7) \begin{vmatrix} t-3 & 10 & -4 \\ 2 & t-2 & -2 \\ 0 & 0 & 1 \end{vmatrix} \\ &\xlongequal{\text{cofactor expansion on the 3rd row}} (t-7) \cdot 1 \cdot (-1)^{3+3} \begin{vmatrix} t-3 & 10 \\ 2 & t-2 \end{vmatrix} \\ &= (t-7)(t^2 - 5t - 14) = (t-7)^2(t+2) \end{aligned}$$

Hence the eigenvalues are $\lambda_1 = \lambda_2 = 7$, $\lambda_3 = -2$

$$\left[\begin{array}{c|c} 7I - A & \mathbf{0} \end{array}\right] \sim \left[\begin{array}{ccc|c} 1 & \frac{1}{2} & -1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right]$$

so we may choose

$$\mathbf{v}_1 = 2\begin{bmatrix} -\frac{1}{2} \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} -1 \\ 2 \\ 0 \end{bmatrix}, \quad \mathbf{v}_2 = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$$

as eigenvectors for $\lambda_1 = \lambda_2 = 7$. We can use Gram-Schmidt process to get an orthogonal set: $\mathbf{w}_1 = \mathbf{v}_1 = \begin{bmatrix} -1 \\ 2 \\ 0 \end{bmatrix}$, $\mathbf{v}_2 - \dfrac{\mathbf{v}_2 \cdot \mathbf{w}_1}{\mathbf{w}_1 \cdot \mathbf{w}_1}\mathbf{w}_1 = \begin{bmatrix} \frac{4}{5} \\ \frac{2}{5} \\ 1 \end{bmatrix}$, we can choose $\mathbf{w}_2 = \begin{bmatrix} 4 \\ 2 \\ 5 \end{bmatrix}$. We can normalize them into

$$\mathbf{u}_1 = \frac{1}{\sqrt{5}}\begin{bmatrix} -1 \\ 2 \\ 0 \end{bmatrix}, \quad \mathbf{u}_2 = \frac{1}{3\sqrt{5}}\begin{bmatrix} 4 \\ 2 \\ 5 \end{bmatrix}$$

$$\left[\begin{array}{c|c} -2I - A & \mathbf{0} \end{array}\right] \sim \left[\begin{array}{ccc|c} 1 & 0 & 1 & 0 \\ 0 & 1 & \frac{1}{2} & 0 \\ 0 & 0 & 0 & 0 \end{array}\right]$$

so we may choose

$$\mathbf{v}_3 = 2\begin{bmatrix} -1 \\ -\frac{1}{2} \\ 1 \end{bmatrix} = \begin{bmatrix} -2 \\ -1 \\ 2 \end{bmatrix}$$

as the eigenvector for $\lambda_3 = -2$. We can normalize it into

$$\mathbf{u}_3 = \frac{\mathbf{v}_3}{\|\mathbf{v}_3\|} = \frac{1}{3}\begin{bmatrix} -2 \\ -1 \\ 2 \end{bmatrix}$$

Now we get the orthogonal decomposition

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

2 Online Assignments

2.1 Online Assignment 1

Problem 2.1.1. Rewrite the following linear systems as augmented matrices and then solve them, show all your work

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

2. $\begin{cases} x_1 + x_2 + x_3 = 6 \\ x_1 - x_2 + x_3 = 2 \\ -x_1 + x_2 + x_3 = 4 \end{cases}$

Solution.

1. The augmented matrix is

$$\begin{bmatrix} 5 & 1 & 2 \\ 3 & -1 & 6 \end{bmatrix} \underset{\sim}{5R2} \begin{bmatrix} 5 & 1 & 2 \\ 15 & -5 & 30 \end{bmatrix} \underset{\sim}{R2 \to R2 - 3R1} \begin{bmatrix} 5 & 1 & 2 \\ 0 & -8 & 24 \end{bmatrix} \underset{\sim}{R2/(-8)} \begin{bmatrix} 5 & 1 & 2 \\ 0 & 1 & -3 \end{bmatrix}$$
$$\underset{\sim}{R1 \to R1 - R2} \begin{bmatrix} 5 & 0 & 5 \\ 0 & 1 & -3 \end{bmatrix} \underset{\sim}{R1/5} \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & -3 \end{bmatrix}$$

Thus the solution to this linear system is $\begin{cases} x_1 = 1 \\ x_2 = -3 \end{cases}$

2. The augmented matrix is

$$\begin{bmatrix} 1 & 1 & 1 & 6 \\ 1 & -1 & 1 & 2 \\ -1 & 1 & 1 & 4 \end{bmatrix} \underset{\sim}{\substack{R2 \to R2 - R1 \\ R3 \to R3 + R1}} \begin{bmatrix} 1 & 1 & 1 & 6 \\ 0 & -2 & 0 & -4 \\ 0 & 2 & 2 & 10 \end{bmatrix} \underset{\sim}{R3 \to R3 + R2} \begin{bmatrix} 1 & 1 & 1 & 6 \\ 0 & -2 & 0 & -4 \\ 0 & 0 & 2 & 6 \end{bmatrix}$$
$$\underset{\sim}{\substack{R2/(-2) \\ R3/3}} \begin{bmatrix} 1 & 1 & 1 & 6 \\ 0 & 1 & 0 & 2 \\ 0 & 0 & 1 & 3 \end{bmatrix} \underset{\sim}{R1 \to R1 - R2 - R3} \begin{bmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 2 \\ 0 & 0 & 1 & 3 \end{bmatrix}$$

Thus the solution to this linear system is $\begin{cases} x_1 = 1 \\ x_2 = 2 \\ x_3 = 3 \end{cases}$

□

Problem 2.1.2. How many solutions does the following linear systems of equations have

1. $\begin{cases} 5x_1 + 7x_2 = 3 \\ -10x_1 - 14x_2 = -3 \end{cases}$

2. $\begin{cases} 2x_1 - x_2 = 4 \\ x_1 - \frac{1}{2}x_2 = 2 \end{cases}$

Solution.

1. The augmented matrix is

$$\begin{bmatrix} 5 & 7 & 3 \\ -10 & -14 & -3 \end{bmatrix} \underset{\sim}{R2 \to R2 + 2R1} \begin{bmatrix} 5 & 7 & 3 \\ 0 & 0 & 3 \end{bmatrix}$$

Since the last column has pivot, by Theorem 1.2.17, this linear system has no solutions

2.

$$\begin{bmatrix} 2 & -1 & 4 \\ 1 & -\frac{1}{2} & 2 \end{bmatrix} \xrightarrow{R2 \to R2 - \frac{1}{2}R1} \begin{bmatrix} 2 & -1 & 4 \\ 0 & 0 & 0 \end{bmatrix}$$

Since the last column has no pivot and there is a free variable x_2, by Theorem 1.2.17, this linear system has infinitely solutions

□

Problem 2.1.3. Consider the following matrix

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

1. Which columns are the pivot columns of A?
2. Write down the RREF of the this matrix

Solution.

1. The pivot columns of A are columns 1,3,4.
2. $$A \xrightarrow[R3/2]{R2/(-1)} \begin{bmatrix} 1 & 2 & 2 & 3 & 1 \\ 0 & 0 & 1 & -2 & -1 \\ 0 & 0 & 0 & 1 & 2 \end{bmatrix} \xrightarrow[R2 \to R2+2R3]{R1 \to R1-3R3} \begin{bmatrix} 1 & 2 & 2 & 0 & -5 \\ 0 & 0 & 1 & 0 & 3 \\ 0 & 0 & 0 & 1 & 2 \end{bmatrix} \xrightarrow{R1 \to R1-2R2} \begin{bmatrix} 1 & 2 & 0 & 0 & -11 \\ 0 & 0 & 1 & 0 & 3 \\ 0 & 0 & 0 & 1 & 2 \end{bmatrix}$$

□

Problem 2.1.4. Determine which of the following statements are true

1. The following matrix is of row reduced echelon form
$$\begin{bmatrix} 1 & 2 & 2 & 3 & 1 \\ 0 & 0 & 0 & 2 & 4 \\ 0 & 0 & -1 & 2 & 1 \end{bmatrix}$$
2. The following two matrices are equivalent
$$\begin{bmatrix} 1 & -1 & 1 & 3 & 2 \\ 2 & 4 & 1 & 2 & 4 \\ 1 & 1 & -3 & 2 & 1 \end{bmatrix} \sim \begin{bmatrix} 1 & 2 & 2 & 3 & 1 \\ 0 & 0 & 0 & 2 & 4 \\ 0 & 0 & -1 & 2 & 1 \end{bmatrix}$$

Solution.

1. False. $(3,3)$, $(2,4)$-th entries are pivots which breaks the "staircase shape"
$$\begin{bmatrix} 1 & 2 & 2 & 3 & 1 \\ 0 & 0 & 0 & 2 & 4 \\ 0 & 0 & -1 & 2 & 1 \end{bmatrix}$$
2. False. Because
$$\begin{bmatrix} 1 & -1 & 1 & 3 & 2 \\ 2 & 4 & 1 & 2 & 4 \\ 1 & 1 & -3 & 2 & 1 \end{bmatrix} \xrightarrow[R3-R1]{R2 \to R2-2R1} \begin{bmatrix} 1 & -1 & 1 & 3 & 2 \\ 0 & 6 & -1 & -4 & 0 \\ 0 & 2 & -4 & -1 & -1 \end{bmatrix}$$
$$\xrightarrow{R2 \to R2-3R3} \begin{bmatrix} 1 & -1 & 1 & 3 & 2 \\ 0 & 0 & 11 & 14 & 3 \\ 0 & 2 & -4 & -1 & -1 \end{bmatrix} \xrightarrow{R2 \leftrightarrow R3} \begin{bmatrix} 1 & -1 & 1 & 3 & 2 \\ 0 & 2 & -4 & -1 & -1 \\ 0 & 0 & 11 & 14 & 3 \end{bmatrix}$$

has pivot columns 1,2,3, and

$$\begin{bmatrix}1&2&2&3&1\\0&0&0&2&4\\0&0&-1&2&1\end{bmatrix}\xrightarrow{R2\leftrightarrow R3}\begin{bmatrix}1&2&2&3&1\\0&0&-1&2&1\\0&0&0&2&4\end{bmatrix}$$

has pivot columns 1,3,4. By Theorem 1.2.10, they are not equivalent, otherwise they would have the same RREF, which implies same pivot columns.

□

Problem 2.1.5. Determine the value(s) of h such that the matrix is the augmented matrix of a consistent linear system $\begin{bmatrix}1&h&1\\2&4&4\end{bmatrix}$

Solution. First consider

$$\begin{bmatrix}1&h&1\\2&4&4\end{bmatrix}\xrightarrow{R2\to R2-2R1}\begin{bmatrix}1&h&1\\0&4-2h&2\end{bmatrix}$$

By Theorem 1.2.17, the linear system has solutions $\iff$ the last column is not a pivot column $\iff 4-2h\neq 0 \iff h\neq 2$ □

Problem 2.1.6. Do the three lines $x_1-4x_2=1$, $2x_1-x_2=-3$, and $-x_1-3x_2=4$ have a common point of intersection? Explain.

Solution. Note that a common point of intersection would be a solution to the linear system $\begin{cases}x_1-4x_2=1\\2x_1-\ \ x_2=-3\\-x_1-3x_2=4\end{cases}$, consider its augmented matrix

$$\begin{bmatrix}1&-4&1\\2&-1&-3\\-1&-3&4\end{bmatrix}\xrightarrow[R3\to R3+R1]{R2\to R2-2R1}\begin{bmatrix}1&-4&1\\0&7&-5\\0&-7&5\end{bmatrix}\xrightarrow{R3\to R3+R2}\begin{bmatrix}1&-4&1\\0&7&-5\\0&0&0\end{bmatrix}$$

By Theorem 1.2.17, since the last column is not a pivot column, the linear system is consistent, i.e. these three lines has comon point(s) of intersection. □

2.2 Online Assignment 2

Problem 2.2.1. Consider the following statements

1. For any four distinct vectors $\mathbf{v}_1,\mathbf{v}_2,\mathbf{v}_3,\mathbf{v}_4$ in $\mathbb{R}^3$, $\text{Span}\{\mathbf{v}_1,\mathbf{v}_2,\mathbf{v}_3,\mathbf{v}_4\}=\mathbb{R}^3$
2. Suppose we know that the augmented matrix of the linear system

$$\begin{cases}a_1x_1+a_2x_2+a_3x_3=d_1\\b_1x_1+b_2x_2+b_3x_3=d_2\\c_1x_1+c_2x_2+c_3x_3=d_3\end{cases}$$

   has two pivot columns, then how many solutions could the linear system have?
3. Consider matrix equation $A\mathbf{x}=\mathbf{0}$ where A is a 3 by 4 matrix, then it always has more than one solution (obviously $\mathbf{x}=\mathbf{0}$ will be a solution)

Solution.

1. False. A counter-example would be $\mathbf{v}_1 = \begin{bmatrix}1\\0\\0\end{bmatrix}, \mathbf{v}_2 = \begin{bmatrix}0\\1\\0\end{bmatrix}, \mathbf{v}_3 = \begin{bmatrix}1\\1\\0\end{bmatrix}, \mathbf{v}_4 = \begin{bmatrix}1\\-1\\0\end{bmatrix}$, then $\mathbf{b} = \begin{bmatrix}0\\0\\1\end{bmatrix}$ is not in Span$\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \mathbf{v}_4\}$ since

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

has pivot in the last row, and apply Theorem 1.2.17

2. The number of solutions could be none or infinitely many. There are in total 6 possible cases (apply Theorem 1.2.17)

Case i: The pivot columns are 1,2, then augmented matrix is equivalent to the RREF matrix $\begin{bmatrix}1 & 0 & * & *\\0 & 1 & * & *\\0 & 0 & 0 & 0\end{bmatrix}$, so the linear system have infinitely many solutions and x_3 is a free variable.

Case ii: The pivot columns are 1,3, then augmented matrix is equivalent to the RREF matrix $\begin{bmatrix}1 & * & 0 & *\\0 & 0 & 1 & *\\0 & 0 & 0 & 0\end{bmatrix}$, so the linear system have infinitely many solutions. and x_2 is a free variable.

Case iii: The pivot columns are 1,4, then augmented matrix is equivalent to the RREF matrix $\begin{bmatrix}1 & * & * & 0\\0 & 0 & 0 & 1\\0 & 0 & 0 & 0\end{bmatrix}$, so the linear system have no solutions.

Case iv: The pivot columns are 2,3, then augmented matrix is equivalent to the RREF matrix $\begin{bmatrix}0 & 1 & 0 & *\\0 & 0 & 1 & *\\0 & 0 & 0 & 0\end{bmatrix}$, so the linear system have infinitely many solutions. and x_1 is a free variable.

Case v: The pivot columns are 2,4, then augmented matrix is equivalent to the RREF matrix $\begin{bmatrix}0 & 1 & * & 0\\0 & 0 & 0 & 1\\0 & 0 & 0 & 0\end{bmatrix}$, so the linear system have no solutions.

Case vi: The pivot columns are 3,4, then augmented matrix is equivalent to the RREF matrix $\begin{bmatrix}0 & 0 & 1 & 0\\0 & 0 & 0 & 1\\0 & 0 & 0 & 0\end{bmatrix}$, so the linear system have no solutions.

3.

$$\begin{bmatrix}* & * & * & * & 0\\* & * & * & * & 0\\* & * & * & * & 0\end{bmatrix} = \begin{bmatrix}A & \mathbf{0}\end{bmatrix} \sim \begin{bmatrix}U & \mathbf{0}\end{bmatrix}$$

Since it doesn't have a pivot in the last column, it must have a solution (indeed at least the trivial solution) by Theorem 1.2.17, and since A has more columns that rows, there must be a free variable, therefore it has infinitely many solutions

□

Problem 2.2.2. Answer the following questions

1. Determine whether $\mathbf{b} = \begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix}$ is in the span of $\left\{ \mathbf{v}_1 = \begin{bmatrix} 1 \\ 0 \\ 2 \end{bmatrix}, \mathbf{v}_2 = \begin{bmatrix} 2 \\ 2 \\ 6 \end{bmatrix}, \mathbf{v}_3 = \begin{bmatrix} -1 \\ -2 \\ -4 \end{bmatrix} \right\}$, why?

2. Assume $\mathbf{w} = \begin{bmatrix} 1 \\ 0 \\ 2 \end{bmatrix}$, $\mathbf{v}_1 = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$, $\mathbf{v}_2 = \begin{bmatrix} 1 \\ 4 \\ -4 \end{bmatrix}$, $\mathbf{v}_3 = \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix}$. Find constants c_1, c_2 such that $\mathbf{w} = c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + 2\mathbf{v}_3$. Show your work on how you found c_1, c_2.

3. Suppose
$$A = \begin{bmatrix} 1 & 3 & 0 & 3 \\ -1 & -1 & -1 & 1 \\ 0 & -4 & 2 & -8 \\ 2 & 0 & 3 & -1 \end{bmatrix}$$
Is it true that for any vector $\mathbf{b}$ in $\mathbb{R}^4$, matrix equation $A\mathbf{x} = \mathbf{b}$ always has solution(s)? If it is, please give your reason. If it is not, please find one such $\mathbf{b}$ and justify your answer.

Solution.

1. Consider
$$\begin{bmatrix} \mathbf{v}_1 & \mathbf{v}_2 & \mathbf{v}_3 & \mathbf{b} \end{bmatrix} = \begin{bmatrix} A & \mathbf{b} \end{bmatrix} = \begin{bmatrix} 1 & 2 & -1 & 1 \\ 0 & 2 & -2 & 1 \\ 2 & 6 & -4 & 2 \end{bmatrix} \xrightarrow{R3 \to R3 - 2R1} \begin{bmatrix} 1 & 2 & -1 & 1 \\ 0 & 2 & -2 & 1 \\ 0 & 2 & -2 & 0 \end{bmatrix}$$
$$\xrightarrow{R3 \to R3 - R2} \begin{bmatrix} 1 & 2 & -1 & 1 \\ 0 & 2 & -2 & 1 \\ 0 & 0 & 0 & -1 \end{bmatrix}$$
By Theorem 1.2.17, $A\mathbf{x} = \mathbf{b}$ has no solution, i.e. $\mathbf{b}$ is not in Span$\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$. So we should consider augmented matrix
$$\begin{bmatrix} 1 & 1 & 1 \\ 1 & 4 & -2 \\ 0 & -4 & 4 \end{bmatrix} \xrightarrow{R2 \to R2 - R1} \begin{bmatrix} 1 & 1 & 1 \\ 0 & 3 & -3 \\ 0 & -4 & 4 \end{bmatrix} \xrightarrow{R2/3} \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & -1 \\ 0 & -4 & 4 \end{bmatrix} \xrightarrow[R1 \to R1 - R2]{R3 \to R3 + 4R2} \begin{bmatrix} 1 & 0 & 2 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{bmatrix}$$
Hence $c_1 = 2$, $c_2 = -1$.

2. It is equivalent to solving $c_1\mathbf{v}_1 + c_2\mathbf{v}_2 = \mathbf{w} - 2\mathbf{v}_3 = \begin{bmatrix} 1 \\ -2 \\ 4 \end{bmatrix}$

3.
$$A \xrightarrow[R4 \to R4 - 2R1]{R2 \to R2 + R1} \begin{bmatrix} 1 & 3 & 0 & 3 \\ 0 & 2 & -1 & 4 \\ 0 & -4 & 2 & -8 \\ 0 & -6 & 3 & -7 \end{bmatrix} \xrightarrow[R4 \to R4 + 3R2]{R3 \to R3 + 2R2} \begin{bmatrix} 1 & 3 & 0 & 3 \\ 0 & 2 & -1 & 4 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 5 \end{bmatrix} \xrightarrow{R3 \leftrightarrow R4} \begin{bmatrix} 1 & 3 & 0 & 3 \\ 0 & 2 & -1 & 4 \\ 0 & 0 & 0 & 5 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$
Which doesn't have pivot in the last row, so $A\mathbf{x} = \mathbf{b}$ doesn't always have a solution by Theorem 1.5.1. To find one such $\mathbf{b}$, you can just try (Since the range of the linear transformation $\mathbf{x} \mapsto A\mathbf{x}$ is the span of the columns of A which is a hyperplane in $\mathbb{R}^4$, so choosing an arbitrary point is most likely not on that hyperplane, i.e. $A\mathbf{x} = \mathbf{b}$ not

solvable!). Here we just try $\mathbf{b} = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}$, then

$$[A \quad \mathbf{b}] = \begin{bmatrix} 1 & 3 & 0 & 3 & 0 \\ -1 & -1 & -1 & 1 & 0 \\ 0 & -4 & 2 & -8 & 1 \\ 2 & 0 & 3 & -1 & 0 \end{bmatrix} \underbrace{}_{\substack{R2 \to R2+R1 \\ R4 \to R4-2R1}} \begin{bmatrix} 1 & 3 & 0 & 3 & 0 \\ 0 & 2 & -1 & 4 & 0 \\ 0 & -4 & 2 & -8 & 1 \\ 0 & -6 & 3 & -7 & 0 \end{bmatrix}$$

$$\underbrace{}_{\substack{R3 \to R3+2R2 \\ R4 \to R4+3R2}} \begin{bmatrix} 1 & 3 & 0 & 3 & 0 \\ 0 & 2 & -1 & 4 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 5 & 0 \end{bmatrix} \underbrace{}_{R3 \leftrightarrow R4} \begin{bmatrix} 1 & 3 & 0 & 3 & 0 \\ 0 & 2 & -1 & 4 & 0 \\ 0 & 0 & 0 & 5 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

□

2.3 Online Assignment 3

Problem 2.3.1. Suppose $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$ are vectors in $\mathbb{R}^4$, if $\{\mathbf{v}_1, \mathbf{v}_2\}$ is linearly independent, $\{\mathbf{v}_2, \mathbf{v}_3\}$ is linearly independent, and $\{\mathbf{v}_1, \mathbf{v}_3\}$ is linearly independent, then $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ is linearly independent

Solution. False. For example $\left\{ \mathbf{v}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \mathbf{v}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \mathbf{v}_3 = \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \end{bmatrix} \right\}$ □

Problem 2.3.2. Suppose A is a m by n matrix, and the matrix equation $A\mathbf{x} = \mathbf{b}$ always has solution for any vector $\mathbf{b}$ in $\mathbb{R}^m$, the columns of A are linearly independent

Solution. False. There could be free variables □

Problem 2.3.3. Solve the linear system $\begin{cases} x_1 + x_2 + x_3 + x_4 = 1 \\ 2x_1 - x_2 + x_3 - x_4 = -1 \end{cases}$ and express its solution set in parametric vector form

Solution. Write the augmented matrix

$$\begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 2 & -1 & 1 & -1 & -1 \end{bmatrix} \underbrace{}_{R2 \to R2-2R1} \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 0 & -3 & -1 & -3 & -3 \end{bmatrix} \underbrace{}_{R2/(-3)} \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 0 & 1 & \frac{1}{3} & 1 & 1 \end{bmatrix}$$

$$\underbrace{}_{R1 \to R1-R2} \begin{bmatrix} 1 & 0 & \frac{2}{3} & 0 & 0 \\ 0 & 1 & \frac{1}{3} & 1 & 1 \end{bmatrix}$$

So the solution set is

$$\begin{cases} x_1 + \frac{2}{3}x_3 = 0 \\ x_2 + \frac{1}{3}x_3 + x_4 = 1 \end{cases} \Rightarrow \begin{cases} x_1 = -\frac{2}{3}x_3 \\ x_2 = 1 - \frac{1}{3}x_3 - x_4 \\ x_3 \text{ is free} \\ x_4 \text{ is free} \end{cases}$$

Its parametric vector form is

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} -\frac{2}{3}x_3 \\ 1 - \frac{1}{3}x_3 - x_4 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix} + \begin{bmatrix} -\frac{2}{3}x_3 \\ -\frac{1}{3}x_3 \\ x_3 \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ -x_4 \\ 0 \\ x_4 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix} + x_3 \begin{bmatrix} -\frac{2}{3} \\ -\frac{1}{3} \\ 1 \\ 0 \end{bmatrix} + x_4 \begin{bmatrix} 0 \\ -1 \\ 0 \\ 1 \end{bmatrix}$$

□

Problem 2.3.4. Suppose $\mathbf{v}_1 = \begin{bmatrix}1\\0\\1\end{bmatrix}$, $\mathbf{v}_2 = \begin{bmatrix}0\\-1\\1\end{bmatrix}$, $\mathbf{v}_3 = \begin{bmatrix}-2\\3\\-11\end{bmatrix}$, $\mathbf{v}_4 = \begin{bmatrix}0\\0\\-3\end{bmatrix}$. Is $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \mathbf{v}_4\}$ linearly independent? If so, please give your reason, if not, please find a linear dependence (i.e. some linear combination $c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + c_2\mathbf{v}_3 + c_4\mathbf{v}_4 = 0$, c_1, c_2, c_3, c_4 not all zero)

Solution. Consider

$$\begin{bmatrix}\mathbf{v}_1 & \mathbf{v}_2 & \mathbf{v}_3 & \mathbf{v}_4 & \mathbf{0}\end{bmatrix} = \begin{bmatrix}1 & 0 & -2 & 0 & 0\\ 0 & -1 & 3 & 0 & 0\\ 1 & 1 & -11 & -3 & 0\end{bmatrix} \xrightarrow{R3\to R3-R1} \begin{bmatrix}1 & 0 & -2 & 0 & 0\\ 0 & -1 & 3 & 0 & 0\\ 0 & 1 & -9 & -3 & 0\end{bmatrix}$$

$$\xrightarrow{R3\to R3+R2} \begin{bmatrix}1 & 0 & -2 & 0 & 0\\ 0 & -1 & 3 & 0 & 0\\ 0 & 0 & -6 & -3 & 0\end{bmatrix} \xrightarrow[R3/(-6)]{(-1)R2} \begin{bmatrix}1 & 0 & -2 & 0 & 0\\ 0 & 1 & -3 & 0 & 0\\ 0 & 0 & 1 & \frac{1}{2} & 0\end{bmatrix} \xrightarrow[R2\to R2+3R3]{R1\to R1+2R3} \begin{bmatrix}1 & 0 & 0 & 1 & 0\\ 0 & 1 & 0 & \frac{3}{2} & 0\\ 0 & 0 & 1 & \frac{1}{2} & 0\end{bmatrix}$$

So the solution set is

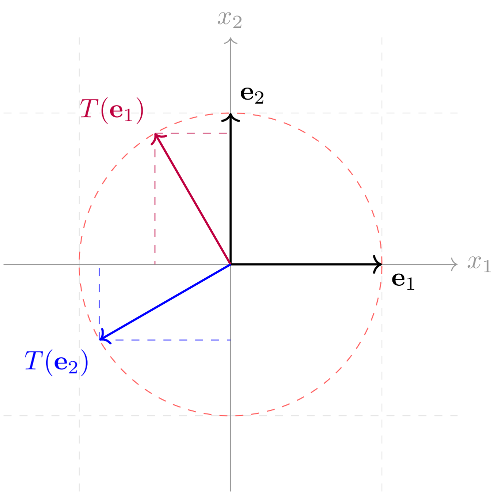
$$\begin{cases} x_1 = -x_4 \\ x_2 = -\frac{3}{2}x_4 \\ x_3 = -\frac{1}{2}x_4 \\ x_4 \text{ is free} \end{cases}$$

We can choose $x_4 = -2$, then $x_1 = 2, x_2 = 3, x_3 = 1$ so that we know $2\mathbf{v}_1 + 3\mathbf{v}_2 + \mathbf{v}_3 - 2\mathbf{v}_4 = \mathbf{0}$ is a linear dependence. □

Problem 2.3.5. Suppose linear transformation $T : \mathbb{R}^2 \to \mathbb{R}^2$ rotates the plane $\mathbb{R}^2$ counter-clockwise by $120°$, what is the standard matrix for the this linear transformation?

Solution. the standard matrix for T is

$$A = \begin{bmatrix}T(\mathbf{e}_1) & T(\mathbf{e}_2)\end{bmatrix} = \begin{bmatrix}-\frac{1}{2} & -\frac{\sqrt{3}}{2}\\ \frac{\sqrt{3}}{2} & -\frac{1}{2}\end{bmatrix}$$



□

Problem 2.3.6. Suppose $T : \mathbb{R}^n \to \mathbb{R}^m$ is a linear transformation and $T(\mathbf{x}) = A\mathbf{x}$, where

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

What is n and m? Find the preimage of $\begin{bmatrix}1\\2\end{bmatrix}$ under T. Is T one-to-one? Is T onto?

Solution. $m = 2$, $n = 3$, the preimage of $\mathbf{b} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$ under T will be the solution to $A\mathbf{x} = \mathbf{b}$, so we have

$$\begin{bmatrix} A & \mathbf{b} \end{bmatrix} = \begin{bmatrix} 1 & 2 & 3 & 1 \\ 0 & 2 & 1 & 2 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & 2 & -1 \\ 0 & 1 & \frac{1}{2} & 1 \end{bmatrix}$$

The solution set is given in parametric vector form

$$\begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} + x_3 \begin{bmatrix} -2 \\ -\frac{1}{2} \\ 1 \end{bmatrix}$$

Therefore T is not one-to-one but onto. $\square$

2.4 Online Assignment 4

Problem 2.4.1. Suppose $A = \begin{bmatrix} 1 & 2 & 1 \\ 2 & 1 & 2 \\ 0 & 2 & 1 \end{bmatrix}$.

- Determine if A is invertible, if not, explain why, if it is, find the inverse matrix A^{-1}
- Find A^T, the transpose of A. Is A^T invertible? If yes, please evaluate $(A^T)^{-1}$

Please show all your work.

Solution.

- A is invertible

$$\left[\begin{array}{ccc|ccc} 1 & 2 & 1 & 1 & 0 & 0 \\ 2 & 1 & 2 & 0 & 1 & 0 \\ 0 & 2 & 1 & 0 & 0 & 1 \end{array}\right] \xrightarrow{R2 \to R2 - 2R1} \left[\begin{array}{ccc|ccc} 1 & 2 & 1 & 1 & 0 & 0 \\ 0 & -3 & 0 & -2 & 1 & 0 \\ 0 & 2 & 1 & 0 & 0 & 1 \end{array}\right]$$

$$\xrightarrow{R2/(-3)} \left[\begin{array}{ccc|ccc} 1 & 2 & 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & \frac{2}{3} & -\frac{1}{3} & 0 \\ 0 & 2 & 1 & 0 & 0 & 1 \end{array}\right] \xrightarrow{R3 \to R3 - 2R2} \left[\begin{array}{ccc|ccc} 1 & 2 & 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & \frac{2}{3} & -\frac{1}{3} & 0 \\ 0 & 0 & 1 & -\frac{4}{3} & \frac{2}{3} & 1 \end{array}\right]$$

$$\xrightarrow{R3 \to R3 - 2R2 - R3} \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & 0 & -1 \\ 0 & 1 & 0 & \frac{2}{3} & -\frac{1}{3} & 0 \\ 0 & 0 & 1 & -\frac{4}{3} & \frac{2}{3} & 1 \end{array}\right]$$

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

- $A^T = \begin{bmatrix} 1 & 2 & 0 \\ 2 & 1 & 2 \\ 1 & 2 & 1 \end{bmatrix}$, and $(A^T)^{-1} = (A^{-1})^T = \begin{bmatrix} 1 & \frac{2}{3} & -\frac{4}{3} \\ 0 & -\frac{1}{3} & \frac{2}{3} \\ -1 & 0 & 1 \end{bmatrix}$

$\square$

Problem 2.4.2. Suppose $A = \begin{bmatrix} -\frac{1}{2} & -\frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & -\frac{1}{2} \end{bmatrix}$ is the standard matrix for the linear transformation of rotating $120°$ counter-clockwise. Evaluate A^{24} and explain why.

Solution. Note that A^{24} is the standard matrix for the composition of linear transformations $\overbrace{T \circ T \circ \cdots \circ T}^{24}$ which is rotate $24 \times 120° = 8 \times 360°$, which is the same as rotate $0°$, so we should have $A^{24} = I$, the identity matrix. $\square$

Problem 2.4.3. We say n is the order of a square matrix A if n is the smallest positive integer such that $A^n = I$, where I is the identity matrix. Suppose $T : \mathbb{R}^2 \to \mathbb{R}^2$ is the linear transformation of reflecting over x_1-axis, and A is the standard matrix of T, the order of A is

Solution. Note that $T \circ T$ is reflecting over x_2-axis twice, which amounts to doing nothing, so A^2 as the standard matrix of $T \circ T$ is the identity matrix, i.e. the order of A is 2. $\square$

Problem 2.4.4. If A, B are both invertible, then $A + B$ is also invertible

Solution. False. For example we can take $B = -A = -I$, then $A + B = 0$ which is not invertible. $\square$

Problem 2.4.5. We call $q(x_1, x_2) = ax_1^2 + 2bx_1x_2 + cx_2^2$ a quadratic form, a, b, c are constants. Suppose $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$, try to rewrite $q(x_1, x_2)$ as the form of a matrix multiplication $\mathbf{x}^T A\mathbf{x}$, where A is a symmetric matrix (i.e. $A^T = A$). Please find A.

Solution. We should choose $A = \begin{bmatrix} a & b \\ b & c \end{bmatrix}$, then

$$\mathbf{x}^T A\mathbf{x} = \begin{bmatrix} x_1 & x_2 \end{bmatrix} \begin{bmatrix} a & b \\ b & c \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} ax_1 + bx_2 & bx_1 + cx_2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = ax_1^2 + 2bx_1x_2 + cx_2^2 = q(x_1, x_2)$$

$\square$

Problem 2.4.6. Suppose $T : \mathbb{R}^3 \to \mathbb{R}^3$ is a linear transformation defined by $T\left(\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}\right) = \begin{bmatrix} x_1 - 2x_2 + x_3 \\ 2x_2 + 3x_3 \\ x_1 - x_3 \end{bmatrix}$. Please find T^{-1}, then standard matrix for T^{-1}. Is T^{-1} onto? Is T^{-1} is one-to-one? Show all your work.

Solution. Note that $T(\mathbf{x}) = A\mathbf{x}$, where $A = \begin{bmatrix} 1 & -2 & 1 \\ 0 & 2 & 3 \\ 1 & 0 & -1 \end{bmatrix}$, use the algorithm to find A^{-1}

$$\left[\begin{array}{ccc|ccc} 1 & -2 & 1 & 1 & 0 & 0 \\ 0 & 2 & 3 & 0 & 1 & 0 \\ 1 & 0 & -1 & 0 & 0 & 1 \end{array}\right] \xrightarrow{R3 \to R3 - R1} \left[\begin{array}{ccc|ccc} 1 & -2 & 1 & 1 & 0 & 0 \\ 0 & 2 & 3 & 0 & 1 & 0 \\ 0 & 2 & -2 & -1 & 0 & 1 \end{array}\right]$$

$$\xrightarrow{R3 \to R3 - R2} \left[\begin{array}{ccc|ccc} 1 & -2 & 1 & 1 & 0 & 0 \\ 0 & 2 & 3 & 0 & 1 & 0 \\ 0 & 0 & -5 & -1 & -1 & 1 \end{array}\right] \xrightarrow[R3/(-5)]{R2/2} \left[\begin{array}{ccc|ccc} 1 & -2 & 1 & 1 & 0 & 0 \\ 0 & 1 & \frac{3}{2} & 0 & \frac{1}{2} & 0 \\ 0 & 0 & 1 & \frac{1}{5} & \frac{1}{5} & -\frac{1}{5} \end{array}\right]$$

$$\xrightarrow{R2 \to R2 - \frac{3}{2}R3} \left[\begin{array}{ccc|ccc} 1 & -2 & 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & -\frac{3}{10} & \frac{1}{5} & \frac{3}{10} \\ 0 & 0 & 1 & \frac{1}{5} & \frac{1}{5} & -\frac{1}{5} \end{array}\right] \xrightarrow{R1 \to R1 + 2R2 - R3} \left[\begin{array}{ccc|ccc} 1 & -2 & 1 & \frac{1}{5} & \frac{1}{5} & \frac{4}{5} \\ 0 & 1 & 0 & -\frac{3}{10} & \frac{1}{5} & \frac{3}{10} \\ 0 & 0 & 1 & \frac{1}{5} & \frac{1}{5} & -\frac{1}{5} \end{array}\right]$$

Hence $A^{-1} = \begin{bmatrix} \frac{1}{5} & \frac{1}{5} & \frac{4}{5} \\ -\frac{3}{10} & \frac{1}{5} & \frac{3}{10} \\ \frac{1}{5} & \frac{1}{5} & -\frac{1}{5} \end{bmatrix}$, so we have $T^{-1}\left(\begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix}\right) = \begin{bmatrix} \frac{1}{5}y_1 + \frac{1}{5}y_2 + \frac{4}{5}y_3 \\ -\frac{3}{10}y_1 + \frac{1}{5}y_2 + \frac{3}{10}y_3 \\ \frac{1}{5}y_1 + \frac{1}{5}y_2 - \frac{1}{5}y_3 \end{bmatrix}$. T^{-1} is both onto and one-to-one. $\square$

Problem 2.4.7. An affine transformation is a mapping $T : \mathbb{R}^2 \to \mathbb{R}^2$ defined by $T(\mathbf{x}) = A\mathbf{x} + \mathbf{b}$, Using partitioned matrix, note that $A\mathbf{x} + \mathbf{b} = \begin{bmatrix} A & \mathbf{b} \end{bmatrix} \begin{bmatrix} \mathbf{x} \\ 1 \end{bmatrix}$. To make thing even nicer, we can use the trick of "adding" 1 at the end of the coordinates, which gives us $\begin{bmatrix} A\mathbf{x} + \mathbf{b} \\ 1 \end{bmatrix} = \begin{bmatrix} A & \mathbf{b} \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \mathbf{x} \\ 1 \end{bmatrix}$.

- Suppose $T_1(\mathbf{x}) = A_1\mathbf{x} + \mathbf{b}_1$ and $T_2(\mathbf{x}) = A_2\mathbf{x} + \mathbf{b}_2$ are both affine transformations. What is $T_2 \circ T_1$? What happens if you try the trick?

- Suppose $T(\mathbf{x}) = A\mathbf{x} + \mathbf{b}$ is an affine transformation and A is invertible. What is T^{-1}? What happens if you try the trick?

Solution.

- $(T_2 \circ T_1)(\mathbf{x}) = T_2(T_1(\mathbf{x})) = A_2(T_1(\mathbf{x})) + \mathbf{b}_2 = A_2(A_1\mathbf{x} + \mathbf{b}_1) + \mathbf{b}_2 = A_2A_1\mathbf{x} + (A_2\mathbf{b}_1 + \mathbf{b}_2)$, this can be explained use the trick as follows

$$\begin{bmatrix} A_2 & \mathbf{b}_2 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} A_1 & \mathbf{b}_1 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} A_2A_1 & A_2\mathbf{b}_1 + \mathbf{b}_2 \\ 0 & 1 \end{bmatrix}$$

- The inverse should be $T^{-1}(\mathbf{y}) = A^{-1}\mathbf{y} - A^{-1}\mathbf{b}$, via the trick we can interpreted it as

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

Since
$$\begin{bmatrix} A^{-1} & -A^{-1}\mathbf{b} \\ 0 & 1 \end{bmatrix} \begin{bmatrix} A & \mathbf{b} \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} I & \mathbf{0} \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} A & \mathbf{b} \\ 0 & 1 \end{bmatrix} \begin{bmatrix} A^{-1} & -A^{-1}\mathbf{b} \\ 0 & 1 \end{bmatrix}$$

□

2.5 Online Assignment 5

Problem 2.5.1. Suppose $A = \begin{bmatrix} 1 & 2 & 2 \\ 2 & 1 & 3 \\ -1 & 2 & 0 \end{bmatrix}$. Use cofactor expansion of A across the last row to evaluate the determinant of A

Solution.

$$\begin{aligned} \begin{vmatrix} 1 & 2 & 2 \\ 2 & 1 & 3 \\ -1 & 2 & 0 \end{vmatrix} &= (-1) \cdot (-1)^{3+1} \begin{vmatrix} 2 & 2 \\ 1 & 3 \end{vmatrix} + 2 \cdot (-1)^{3+2} \begin{vmatrix} 1 & 2 \\ 2 & 3 \end{vmatrix} + 0 \\ &= (-1)(2 \cdot 3 - 2 \cdot 1) + (-2)(1 \cdot 3 - 2 \cdot 2) = -2 \end{aligned}$$

□

Problem 2.5.2. Compute the determinant of $A = \begin{bmatrix} 1 & 2 & 2 & -1 \\ 0 & 2 & 0 & 0 \\ 3 & 5 & 2 & 3 \\ 2 & 1 & 0 & -1 \end{bmatrix}$. Is A invertible?

Solution.

$$\begin{vmatrix} 1 & 2 & 2 & -1 \\ 0 & 2 & 0 & 0 \\ 3 & 5 & 2 & 3 \\ 2 & 1 & 0 & -1 \end{vmatrix} \xlongequal{\text{cofactor expansion second row}} 2(-1)^{2+2} \begin{vmatrix} 1 & 2 & -1 \\ 3 & 2 & 3 \\ 2 & 0 & -1 \end{vmatrix} \xlongequal[R3 \to R3 - 2R1]{R2 \to R3 - 3R1} 2 \begin{vmatrix} 1 & 2 & -1 \\ 0 & -4 & 6 \\ 0 & -4 & 1 \end{vmatrix}$$

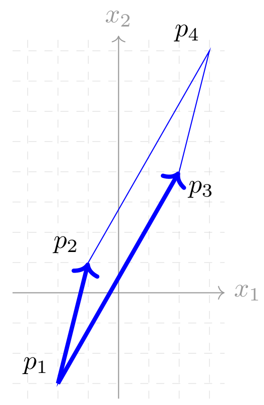
$$\xlongequal{R3 \to R3 - R2} 2 \begin{vmatrix} 1 & 2 & -1 \\ 0 & -4 & 6 \\ 0 & 0 & -5 \end{vmatrix} = 2 \cdot 1 \cdot (-4) \cdot (-5) = 40$$

A is invertible since $\det A \neq 0$ □

Problem 2.5.3. Consider the parallelgram P with vertices $(-2, -3), (-1, 1), (2, 4), (3, 8)$, use determinants to evaluate the area of P

Solution. Name these four points p_1, p_2, p_3, p_4, and the vectors $\mathbf{a}_1 = \begin{bmatrix} 1 \\ 4 \end{bmatrix}$ and $\mathbf{a}_2 = \begin{bmatrix} 4 \\ 7 \end{bmatrix}$. And we consider $A = \begin{bmatrix} \mathbf{a}_1 & \mathbf{a}_2 \end{bmatrix} = \begin{bmatrix} 1 & 4 \\ 4 & 7 \end{bmatrix}$, then area of P would be

$$|\det A| = |(1 \cdot 7 - 4 \cdot 4))| = 9$$



□

Problem 2.5.4. $\det(A - B) = \det A - \det B$

Solution. This is false, for example, we could just take $A = I = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$, $B = -I$, then $4 = \det(2I) = \det(A - B) \neq \det A - \det B = 1 - 1 = 0$ □

Problem 2.5.5. If A is a 3 by 3 matrix, then $\det(2A) = 8(\det A)$

Solution. This is true. □

Problem 2.5.6. Suppose A, B are both 3 by 3 matrices, and $\det A = 2$, $\det B = \frac{1}{3}$, then the determinant of $A^T B^{-1}$ is

Solution. $\det(A^T B^{-1}) = \det(A^T)\det(B^{-1}) = (\det A)(\det B)^{-1} = 2 \cdot 3 = 6.$ □

Problem 2.5.7. Suppose A is a 3 by 3 matrix with entries integers, and $A^3 = I$ is the identity matrix. Then the determinant of A has to be

Solution. $\det A$ must be some real number. $1 = \det I = \det(A^3) = (\det A)^3 \Rightarrow \det A = \sqrt[3]{1} = 1$. □

2.6 Online Assignment 6

Problem 2.6.1. Suppose $A = \begin{bmatrix} 1 & 2 & 1 & 2 & -1 & 3 \\ 0 & 0 & 0 & 0 & 1 & -1 \\ 0 & 0 & 1 & 3 & 1 & -2 \end{bmatrix}$

a) Find a basis for the null space of A

b) Find a basis for the column space of A

c) Find a basis for the row space of A

Solution. First realize

$$A \xrightarrow{R2 \leftrightarrow R3} \begin{bmatrix} 1 & 2 & 1 & 2 & -1 & 3 \\ 0 & 0 & 1 & 3 & 1 & -2 \\ 0 & 0 & 0 & 0 & 1 & -1 \end{bmatrix} \sim \begin{bmatrix} 1 & 2 & 0 & -1 & 0 & 3 \\ 0 & 0 & 1 & 3 & 0 & -1 \\ 0 & 0 & 0 & 0 & 1 & -1 \end{bmatrix}$$

Hence the solution in parametric form is $x_2 \begin{bmatrix} -2 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} + x_4 \begin{bmatrix} 1 \\ 0 \\ -3 \\ 1 \\ 0 \\ 0 \end{bmatrix} + x_6 \begin{bmatrix} -3 \\ 0 \\ 1 \\ 0 \\ 1 \\ 1 \end{bmatrix}$. And the pivot columns are the 1st, 3rd and 5th columns

a) A basis for Nul A could be $\left\{ \begin{bmatrix} -2 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ -3 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} -3 \\ 0 \\ 1 \\ 0 \\ 1 \\ 1 \end{bmatrix} \right\}$

b) A basis for Col A could be $\left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} -1 \\ 1 \\ 1 \end{bmatrix} \right\}$

c) A basis for Row A could be

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

□

Problem 2.6.2. Suppose $A = \begin{bmatrix} 3 & -1 & -5 \\ 1 & 1 & -1 \\ -2 & 2 & 4 \end{bmatrix}$

a) Determine whether $\mathbf{u} = \begin{bmatrix} -3 \\ 1 \\ -2 \end{bmatrix}$ is in the null space of A. Explain your reasoning.

b) Determine whether $\mathbf{b} = \begin{bmatrix} 1 \\ -3 \\ -4 \end{bmatrix}$ is in the column space of A. Explain your reasoning.

Solution.

a) $\mathbf{u} \in \operatorname{Nul} A$ since $A\mathbf{u} = \mathbf{0}$

b) $\mathbf{b} \in \operatorname{Col} A$ since linear system $A\mathbf{x} = \mathbf{b}$ is consistent

□

Problem 2.6.3. Recall $\mathbb{P}_2 = \{a_0 + a_1 t + a_2 t^2 | a_0, a_1, a_2 \in \mathbb{R}\}$ is the set of polynomials of degree less or equal to 2. Let V be the subset of $\mathbb{P}_2$ consists of polynomials that evalute to 0 at $t = 1$ (i.e. polynomial $p(t)$ is in V if $p(1) = 0$ and of degree less or equal to 2). With the usual addition and scalar multiplication for polynomials

a) Show that V is a subspace.

b) Find a basis of V.

c) Cosider $T : V \to \mathbb{R}^2$ that maps polynomial $p(t)$ to $\begin{bmatrix} p(-1) \\ p(2) \end{bmatrix}$, show T is a linear transformation

Solution.

a) Realize that $V = \ker S$ for the linear transformation $S : \mathbb{P}_2 \to \mathbb{R}$, $S(a_0 + a_1 t + a_2 t^2) = a_0 + a_1 + a_2$, thus V is a subspace of $\mathbb{P}_2$

b) This is precisely Example 1.19.2

c) For any $p(t), q(t) \in V$ and $c \in \mathbb{R}$, we have

$$T(p+q) = \begin{bmatrix} (p+q)(-1) \\ (p+q)(2) \end{bmatrix} = \begin{bmatrix} p(-1) + q(-1) \\ p(2) + q(2) \end{bmatrix} = \begin{bmatrix} p(-1) \\ p(2) \end{bmatrix} + \begin{bmatrix} q(-1) \\ q(2) \end{bmatrix} = T(p) + T(q)$$

$$T(cp) = \begin{bmatrix} (cp)(-1) \\ (cp)(2) \end{bmatrix} = \begin{bmatrix} c \cdot p(-1) \\ c \cdot p(2) \end{bmatrix} = c \begin{bmatrix} p(-1) \\ p(2) \end{bmatrix} = cT(p)$$

Therefore $T : V \to \mathbb{R}^2$ is a linear transformation

$\square$

Problem 2.6.4. We say a square matrix A is anti-symmetric if $A^T = -A$. Denote the set of 3×3 anti-symmetric matrices V.

a) Show that V is a vector space.

b) What is the dimension of V?

c) Find a basis of V.

d) Show that

$$\mathcal{B} = \left\{ B_1 = \begin{bmatrix} 0 & -1 & -1 \\ 1 & 0 & -1 \\ 1 & 1 & 0 \end{bmatrix}, B_2 = \begin{bmatrix} 0 & 2 & 0 \\ -2 & 0 & -1 \\ 0 & 1 & 0 \end{bmatrix}, B_3 = \begin{bmatrix} 0 & -1 & -2 \\ 1 & 0 & -3 \\ 2 & 3 & 0 \end{bmatrix} \right\}$$

form a basis for V.

Solution.

a) For any $A, B \in V$ and $c \in \mathbb{R}$, by definition $A^T = -A, B^T = -B$, so

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

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

hence $A+B, cA \in V$, i.e. V is a closed under addition and scalar multiplication. Therefore V is a subspace of $M_{3\times 2}(\mathbb{R})$, and consequently a vector space.

b) It is not hard to realize

$$V = \left\{ \begin{bmatrix} 0 & -a & -b \\ a & 0 & -c \\ b & c & 0 \end{bmatrix} \in M_{3\times 3}(\mathbb{R}) \middle| a, b, c \in \mathbb{R} \right\} \cong \mathbb{R}^3$$

Thus $\dim V = 3$

c) Note that

$$\begin{aligned} \begin{bmatrix} 0 & -a & -b \\ a & 0 & -c \\ b & c & 0 \end{bmatrix} &= \begin{bmatrix} 0 & -a & 0 \\ a & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 & -b \\ 0 & 0 & 0 \\ b & 0 & 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -c \\ 0 & c & 0 \end{bmatrix} \\ &= a \begin{bmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} + b \begin{bmatrix} 0 & 0 & -1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix} + c \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{bmatrix} \end{aligned} \tag{2.6.1}$$

So we know that

$$\mathcal{E} = \{E_1, E_2, E_3\} = \left\{ \begin{bmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 & -1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{bmatrix} \right\}$$

is a basis for V, this is linearly independent since the linear combination (2.6.1) is equal to zero $\iff a = b = c = 0$

d) The coordinate vectors $\{[B_1]_{\mathcal{E}}, [B_2]_{\mathcal{E}}, [B_3]_{\mathcal{E}}\} = \left\{ \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} \right\}$ form a basis for $\mathbb{R}^3$, so $\mathcal{B}$ form a basis for V

$\square$

2.7 Online Assignment 7

Problem 2.7.1. Suppose $\mathbf{v}$ is an eigenvector for matrices A and B, then $\mathbf{v}$ is an eigenvector for $A+B$ and AB.

Solution. This is true. Since $\mathbf{v}$ is an eigenvector for both A and B, there exist eigenvalues λ_1, λ_2 such that $A\mathbf{v} = \lambda_1\mathbf{v}$, $B\mathbf{v} = \lambda_2\mathbf{v}$, so we have $(A+B)\mathbf{v} = A\mathbf{v} + B\mathbf{v} = \lambda_1\mathbf{v} + \lambda_2\mathbf{v} = (\lambda_1+\lambda_2)\mathbf{v}$. In other words, $\mathbf{v}$ is an eigenvector for $A+B$ with eigenvalue $\lambda_1+\lambda_2$. Similarly, we also have $AB\mathbf{v} = A(\lambda_1\mathbf{v}_1) = \lambda_2 A\mathbf{v} = \lambda_1\lambda_2\mathbf{v}$. In other words, $\mathbf{v}$ is an eigenvector for AB with eigenvector $\lambda_1\lambda_2$ □

Problem 2.7.2. Suppose $t+3t^2-2t^3$ is the characteristic polynomial of a 3 by 3 matrix, then A is not invertible.

Solution. This is true. Note that $t=0$ is a root of the characteristic polynomial, so the null space of A is not trivial, A is not invertible. □

Problem 2.7.3. Suppose $\mathcal{B} = \{1+t, 1+2t^2, 1-t+t^2\}$ and $\mathcal{C} = \{1-t, t, t^2\}$ are two bases for $\mathbb{P}_2$, What is the change of basis matrix $\underset{\mathcal{C}\leftarrow\mathcal{B}}{P}$ from $\mathcal{B}$ to $\mathcal{C}$. Please show all your work.

Solution. First let's find the change of basis matrices from $\mathcal{B}$ and $\mathcal{C}$ to the standard basis $\mathcal{E} = \{1, t, t^2\}$. We have

$$\underset{\mathcal{E}\leftarrow\mathcal{B}}{P} = \begin{bmatrix}[1+t]_{\mathcal{E}} & [1+2t^2]_{\mathcal{E}} & [1-t+t^2]_{\mathcal{E}}\end{bmatrix} = \begin{bmatrix}1 & 1 & 1\\ 1 & 0 & -1\\ 0 & 2 & 1\end{bmatrix}$$

And

$$\underset{\mathcal{E}\leftarrow\mathcal{C}}{P} = \begin{bmatrix}[1-t]_{\mathcal{E}} & [t]_{\mathcal{E}} & [t^2]_{\mathcal{E}}\end{bmatrix} = \begin{bmatrix}1 & 0 & 0\\ -1 & 1 & 0\\ 0 & 0 & 1\end{bmatrix}$$

Hence we have

$$\underset{\mathcal{C}\leftarrow\mathcal{B}}{P} = \left(\underset{\mathcal{C}\leftarrow\mathcal{E}}{P}\right)\left(\underset{\mathcal{E}\leftarrow\mathcal{B}}{P}\right) = \left(\underset{\mathcal{E}\leftarrow\mathcal{C}}{P}\right)^{-1}\left(\underset{\mathcal{E}\leftarrow\mathcal{B}}{P}\right) = \begin{bmatrix}1 & 0 & 0\\ -1 & 1 & 0\\ 0 & 0 & 1\end{bmatrix}^{-1}\begin{bmatrix}1 & 1 & 1\\ 1 & 0 & -1\\ 0 & 2 & 1\end{bmatrix} = \begin{bmatrix}1 & 1 & 1\\ 2 & 1 & 0\\ 0 & 2 & 1\end{bmatrix}$$

□

Problem 2.7.4. If $NulA$ is 2-dimensional, then 0 is an eigenvalue of A.

Solution. This is true. Since if $\operatorname{Nul} A$ is 2-dimensional, then $\operatorname{Nul} A$ is non-trivial, thus 0 is an eigenvalue of A □

Problem 2.7.5. Is $\begin{bmatrix}0\\0\\1\end{bmatrix}$ an eigenvector of $\begin{bmatrix}2 & 0 & 0\\ 3 & 2 & 0\\ 0 & 4 & 3\end{bmatrix}$? If so, find the corresponding eigenvalue, if not, please explain why.

Solution. Note that $\begin{bmatrix}2 & 0 & 0\\ 3 & 2 & 0\\ 0 & 4 & 3\end{bmatrix}\begin{bmatrix}0\\0\\1\end{bmatrix} = \begin{bmatrix}0\\0\\3\end{bmatrix}$. so $\begin{bmatrix}0\\0\\1\end{bmatrix}$ is an eigenvector with eigenvalue 3 □

Problem 2.7.6. Find all eigenvalues of $A = \begin{bmatrix}6 & -2 & 0\\ -2 & 9 & 0\\ 5 & 8 & 3\end{bmatrix}$. Please show all your work.

Solution.

$$|tI-A| = \begin{vmatrix}t-6 & 2 & 0\\ 2 & t-9 & 0\\ -5 & -8 & t-3\end{vmatrix} \overset{\text{cofator expansion across 3rd column}}{=\!=\!=\!=\!=} (t-3)(-1)^{3+3}\begin{vmatrix}t-6 & 2\\ 2 & t-9\end{vmatrix}$$

$$= (t-3)((t-6)(t-9)-2\cdot 2) = (t-3)(t^2-5t+50) = (t-3)(t-5)(t-10)$$

Therefore eigenvalues for A are 3,5,10. □

Problem 2.7.7. Assume that A is similar to an upper triangular matrix U, then $\det A$ is the product of all its eigenvalues (counting multiplicity). Please explain why.

Solution. Suppose the diagonal elements of A are $\lambda_1, \cdots, \lambda_n$, then the charateristic polynomial of A is the same the charateristic polynomial U (Since they are similar) which would be $(t - \lambda_1) \cdots (t - \lambda_n)$, so $\lambda_1, \cdots, \lambda_n$ are the eigenvalues for A. And the determinant of A is the same as the determinant of U (Since they are similar) which is $\lambda_1 \cdots \lambda_n$ □

2.8 Online Assignment 8

Problem 2.8.1. Suppose $A = \begin{bmatrix} 2 & 2 & 1 \\ 1 & 3 & 1 \\ 1 & 2 & 2 \end{bmatrix}$, determine whether A is diagonalizable. If it is, please find a diagonalization. If not, please explain why.

Solution. First we evaluate the characteristic polynomial of A

$$\det(tI - A) = \begin{vmatrix} t-2 & -2 & -1 \\ -1 & t-3 & -1 \\ -1 & -2 & t-2 \end{vmatrix} \xlongequal[R2 \to R2 - R3]{R1 \to R1 + (t-2)R3} \begin{vmatrix} 0 & -2(t-1) & (t-1)(t-3) \\ 0 & t-1 & -(t-1) \\ -1 & -2 & t-2 \end{vmatrix}$$

$$\xlongequal{\text{factor out } (t-1) \text{ on row 1 and row 2}} (t-1)^2 \begin{vmatrix} 0 & -2 & (t-3) \\ 0 & 1 & -1 \\ -1 & -2 & t-2 \end{vmatrix} = (t-1)^2(t-5)$$

So the eigenvalues are $\lambda_1 = \lambda_2 = 1$, $\lambda_3 = 5$. For the 1-eigenspace, we consider

$$\left[\begin{array}{c|c} I - A & \mathbf{0} \end{array}\right] = \left[\begin{array}{ccc|c} -1 & -2 & -1 & 0 \\ -1 & -2 & -1 & 0 \\ -1 & -2 & -1 & 0 \end{array}\right] \sim \left[\begin{array}{ccc|c} 1 & 2 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right]$$

So we get two basis vectors $\left\{ \mathbf{v}_1 = \begin{bmatrix} -2 \\ 1 \\ 0 \end{bmatrix}, \mathbf{v}_2 = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} \right\}$. For the 5-eigenspace, we consider

$$\left[\begin{array}{c|c} 5I - A & \mathbf{0} \end{array}\right] = \left[\begin{array}{ccc|c} 3 & -2 & -1 & 0 \\ -1 & 2 & -1 & 0 \\ -1 & -2 & 3 & 0 \end{array}\right] \sim \left[\begin{array}{ccc|c} 1 & 0 & -1 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right]$$

So we get a basis vector $\left\{ \mathbf{v}_3 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \right\}$

From above, we get the diagonalization

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

□

Problem 2.8.2. Suppose $A = \begin{bmatrix} -6 & 8 \\ -4 & 6 \end{bmatrix}$, Please evaluate A^{101}, show all your work.

Solution. First we can diagonalize A as

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

$$A^{101} = PD^{101}P^{-1} = \begin{bmatrix} 1 & 2 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 2^{101} & 0 \\ 0 & -2^{101} \end{bmatrix} \begin{bmatrix} -1 & 2 \\ 1 & -1 \end{bmatrix}$$

□

Problem 2.8.3. Suppose $T : \mathbb{R}^3 \to \mathbb{R}^3$ is a linear transformation, $\mathcal{B} = \{\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3\}$, $\mathcal{C} = \{\mathbf{c}_1, \mathbf{c}_2, \mathbf{c}_3\}$ are two different bases for $\mathbb{R}^3$. Determine whether the following is possible.

a) $[T]_{\mathcal{B}} = \begin{bmatrix} 1 & 2 & 4 \\ 3 & -1 & -2 \\ 2 & -1 & 3 \end{bmatrix}$ and $[T]_{\mathcal{C}} = \begin{bmatrix} 1 & -3 & 1 \\ 2 & 1 & 6 \\ 0 & 3 & 8 \end{bmatrix}$

b) $[T]_{\mathcal{B}} = \begin{bmatrix} 3 & 0 & 0 \\ 2 & 2 & 0 \\ 2 & 3 & 4 \end{bmatrix}$ and $[T]_{\mathcal{C}} = \begin{bmatrix} 1 & -1 & 0 \\ 2 & 4 & 0 \\ 3 & -2 & 4 \end{bmatrix}$

Solution.

a) This is not possible since these two matrices have different determinant.

b) This is possible since these two matrix has the same the characteristic polynomial $(t-2)(t-3)(t-4)$, hence they are similar to same diagonal matrix, and thus similar.

□

Problem 2.8.4. Suppose A is similar to B.

a) Could you conclude that $3A$ is similar to $3B$. If you can, please give your reasons, if not, please find a counter-example.

b) Could you conclude that A^{-1} is similar to B^{-1}. If you can, please give your reasons, if not, please find a counter-example.

Solution. Since A is similar to B, we may assume $A = PBP^{-1}$.

a) $3A = 3PBP^{-1} = P(3B)P^{-1}$ is similar.

b) $A^{-1} = (PBP^{-1})^{-1} = PB^{-1}P^{-1}$ is similar.

□

2.9 Online Assignment 9

Problem 2.9.1. Determine whether the following statements are true

a) $\|\mathbf{u}\|^2 + \|\mathbf{v}\|^2 = \|\mathbf{u}+\mathbf{v}\|^2$ if and only if $\mathbf{u}, \mathbf{v}$ are orthogonal.

b) If W is a subspace of $\mathbb{R}^n$, and vector $\mathbf{v}$ is orthogonal to both W and $W^\perp$, then $\mathbf{v} = \mathbf{0}$.

c) If $W = \operatorname{Span}\{\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3\}$ with $\{\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3\}$ linearly independent, and if $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ is an orthogonal set in W, then $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ is a basis for W.

Solution.

a) This is true. $\|\mathbf{u}+\mathbf{v}\|^2 = (\mathbf{u}+\mathbf{v}) \cdot (\mathbf{u}+\mathbf{v}) = \mathbf{u} \cdot \mathbf{u} + 2\mathbf{u} \cdot \mathbf{v} + \mathbf{v} \cdot \mathbf{v} = \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2 + 2\mathbf{u} \cdot \mathbf{v}$. Thus the equality holds $\iff$ $\mathbf{u} \cdot \mathbf{v} = 0$, i.e. $\mathbf{u}, \mathbf{v}$ are orthogonal.

b) Consider $\mathbf{v} \cdot \mathbf{v} = \|\mathbf{v}\|^2$, it should be 0 since $\mathbf{v}$ is in both W and $W^\perp$, so $\mathbf{v} = \mathbf{0}$

c) This is true by Theorem 1.24.3

□

Problem 2.9.2. Suppose $A = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 1 & -1 \\ 0 & 2 & 4 \\ 3 & 5 & -1 \end{bmatrix}$, please find a basis for $(\operatorname{Col} A)^\perp$.

Solution. Recall that $(\operatorname{Col} A)^{\perp} = \operatorname{Nul}(A^T)$. And we can find a basis for $\operatorname{Nul}(A^T)$ through

$$\left[\begin{array}{c|c} A^T & \mathbf{0} \end{array}\right] = \left[\begin{array}{cccc|c} 1 & 2 & 0 & 3 & 0 \\ 2 & 1 & 2 & 5 & 0 \\ 3 & -1 & 4 & 1 & 0 \end{array}\right] \sim \left[\begin{array}{cccc|c} 1 & 0 & 0 & -13 & 0 \\ 0 & 1 & 0 & 8 & 0 \\ 0 & 0 & 1 & \frac{23}{2} & 0 \end{array}\right]$$

Thus

$$(\operatorname{Col} A)^{\perp} = \operatorname{Span}\left\{\begin{bmatrix} 13 \\ -8 \\ -\frac{23}{2} \\ 1 \end{bmatrix}\right\}$$

□

Problem 2.9.3. Suppose we have $\mathcal{B} = \left\{\mathbf{u}_1 = \begin{bmatrix} 1 \\ 1 \\ 0 \\ -1 \end{bmatrix}, \mathbf{u}_2 = \begin{bmatrix} 1 \\ 0 \\ 1 \\ 1 \end{bmatrix}, \mathbf{u}_3 = \begin{bmatrix} 0 \\ -1 \\ 1 \\ -1 \end{bmatrix}\right\}$, please justify that $\mathcal{B}$ is an orthogonal set, suppose $\mathbf{y} = \begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \end{bmatrix}$, L is the subspace spanned by $\mathcal{B}$, compute the projection $\operatorname{Proj}_L \mathbf{y}$ of $\mathbf{y}$ onto L.

Solution. Let $A = \begin{bmatrix} \mathbf{u}_1 & \mathbf{u}_2 & \mathbf{u}_3 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & -1 \\ 0 & 1 & 1 \\ -1 & 1 & -1 \end{bmatrix}$, we can check that $A^T A = I$ is the 3 by 3 identity matrix, so $\mathcal{B}$ is an orthogonal set. And therefore

$$\begin{aligned} \operatorname{Proj}_L \mathbf{y} &= \frac{\mathbf{y} \cdot \mathbf{u}_1}{\mathbf{u}_1 \cdot \mathbf{u}_1}\mathbf{u}_1 + \frac{\mathbf{y} \cdot \mathbf{u}_2}{\mathbf{u}_2 \cdot \mathbf{u}_2}\mathbf{u}_2 + \frac{\mathbf{y} \cdot \mathbf{u}_3}{\mathbf{u}_3 \cdot \mathbf{u}_3}\mathbf{u}_3 \\ &= \frac{-1}{3}\mathbf{u}_1 + \frac{8}{3}\mathbf{u}_2 + \frac{-3}{3}\mathbf{u}_3 \\ &= -\frac{1}{3}\begin{bmatrix} 1 \\ 1 \\ 0 \\ -1 \end{bmatrix} + \frac{8}{3}\begin{bmatrix} 1 \\ 0 \\ 1 \\ 1 \end{bmatrix} - \begin{bmatrix} 0 \\ -1 \\ 1 \\ -1 \end{bmatrix} = \begin{bmatrix} \frac{7}{3} \\ \frac{2}{3} \\ \frac{5}{3} \\ 4 \end{bmatrix} \end{aligned}$$

□

Problem 2.9.4. Suppose $A = \begin{bmatrix} -1 & 6 & 6 \\ 3 & -8 & 3 \\ 1 & -2 & 6 \\ 1 & -4 & -3 \end{bmatrix}$, please find an orthogonal basis for $\operatorname{Col} A$.

Solution. We apply the Gram-Schmidt process here

- $\mathbf{u}_1 = \mathbf{v}_1 = \begin{bmatrix} -1 \\ 3 \\ 1 \\ 1 \end{bmatrix}$
- $\mathbf{u}_2 = \mathbf{v}_2 - \dfrac{\mathbf{v}_2 \cdot \mathbf{u}_1}{\mathbf{u}_1 \cdot \mathbf{u}_1}\mathbf{u}_1 = \begin{bmatrix} 6 \\ -8 \\ -2 \\ -4 \end{bmatrix} - \frac{-36}{12}\begin{bmatrix} -1 \\ 3 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 3 \\ 1 \\ 1 \\ -1 \end{bmatrix}$
- $\mathbf{u}_3 = \mathbf{v}_3 - \dfrac{\mathbf{v}_3 \cdot \mathbf{u}_1}{\mathbf{u}_1 \cdot \mathbf{u}_1}\mathbf{u}_1 - \dfrac{\mathbf{v}_3 \cdot \mathbf{u}_2}{\mathbf{u}_2 \cdot \mathbf{u}_2}\mathbf{u}_2 = \begin{bmatrix} 6 \\ 3 \\ 6 \\ -3 \end{bmatrix} - \frac{6}{12}\begin{bmatrix} -1 \\ 3 \\ 1 \\ 1 \end{bmatrix} - \frac{30}{12}\begin{bmatrix} 3 \\ 1 \\ 1 \\ -1 \end{bmatrix} = \begin{bmatrix} -1 \\ -1 \\ 3 \\ -1 \end{bmatrix}$

So we have an orthogonal basis for $\operatorname{Col} A$

$$\operatorname{Col} A = \operatorname{Span}\left\{\begin{bmatrix}-1\\3\\1\\1\end{bmatrix}, \begin{bmatrix}3\\1\\1\\-1\end{bmatrix}, \begin{bmatrix}-1\\-1\\3\\-1\end{bmatrix}\right\}$$

□

2.10 Online Assignment 10

Problem 2.10.1. Determine whether the following statements are correct

a) If A is symmetric and if vectors $\mathbf{u}$ and $\mathbf{v}$ such that $A\mathbf{u} = \mathbf{u}$ and $\mathbf{v}$ is in $\operatorname{Nul} A$, then $\mathbf{u}\cdot\mathbf{v} = \mathbf{0}$.

b) $\operatorname{Nul} A = \operatorname{Nul} A^T A$.

Solution.

a) The eigenvalues for eigenvectors $\mathbf{u}$ and $\mathbf{v}$ are 1 and 0, and A is symmetric, by Theorem 1.31.1, $\mathbf{u} \cdot \mathbf{v} = 0$.

b) If $\mathbf{x} \in \operatorname{Nul} A$, then $A\mathbf{x} = \mathbf{0}$, so $A^T A\mathbf{x} = \mathbf{0}$. If $\mathbf{x} \in \operatorname{Nul}(A^T A)$, then $A^T A\mathbf{x} = \mathbf{0}$, so $0 = \mathbf{x}^T A^T A\mathbf{x} = \|A\mathbf{x}\|^2$, so $A\mathbf{x} = 0$.

□

Problem 2.10.2. Find the least-squares solution(s) to $\begin{bmatrix}1 & 5\\3 & 1\\-2 & 4\end{bmatrix}\begin{bmatrix}x_1\\x_2\end{bmatrix} = \begin{bmatrix}4\\-2\\-3\end{bmatrix}$.

Solution. Let's denote $A = \begin{bmatrix}1 & 5\\3 & 1\\-2 & \end{bmatrix}$, $\mathbf{b} = \begin{bmatrix}4\\-2\\-3\end{bmatrix}$, then $A^T A = \begin{bmatrix}14 & 0\\0 & 42\end{bmatrix}$, and $A^T\mathbf{b} = \begin{bmatrix}4\\6\end{bmatrix}$, then we get the least -square solution $\hat{\mathbf{x}} = (A^T A)^{-1} A^T \mathbf{b} = \begin{bmatrix}\frac{2}{7}\\\frac{1}{7}\end{bmatrix}$ □

Problem 2.10.3. Orthogonal diagonalize the matrix $\begin{bmatrix}3 & 4\\4 & 9\end{bmatrix}$.

Solution. First note that $A = \begin{bmatrix}3 & 4\\4 & 9\end{bmatrix}$ is symmetric and real-valued, we can get its eigenvalues $\lambda_1 = 1$, $\lambda_2 = 11$, and normalized eigenvectors $\mathbf{u}_1 = \begin{bmatrix}-\frac{2}{\sqrt{5}}\\\frac{1}{5}\end{bmatrix}$, $\mathbf{u}_2 = \begin{bmatrix}\frac{1}{5}\\\frac{2}{5}\end{bmatrix}$. Hence we have the orthogonal diganolization

$$A = \begin{bmatrix}3 & 4\\4 & 9\end{bmatrix} = \begin{bmatrix}-\frac{2}{\sqrt{5}} & \frac{1}{\sqrt{5}}\\\frac{1}{5} & \frac{2}{\sqrt{5}}\end{bmatrix}\begin{bmatrix}1 & 0\\0 & 11\end{bmatrix}\begin{bmatrix}-\frac{2}{\sqrt{5}} & \frac{1}{\sqrt{5}}\\\frac{1}{5} & \frac{2}{\sqrt{5}}\end{bmatrix} = PDP^T$$

□

Problem 2.10.4. Suppose $A = \begin{bmatrix}1 & 5\\-2 & 3\end{bmatrix}$.

a) Please find the eigenvalues of A.

b) Please find the eigenvectors of A.

c) Please write A as matrix multiplication PCP^{-1}, where C is of the form $\begin{bmatrix}a & -b\\b & c\end{bmatrix}$.

Solution.

a) The characteristic polynomial is $t^2 - 4t13$, and so the eigenvalues are $\lambda = 2-3i$, $\overline{\lambda} = 2+3i$, so we have $a = 2$, $b = 3$ and $C = \begin{bmatrix} 2 & -3 \\ 3 & 2 \end{bmatrix}$

b) The eigenvalues for $\lambda, \overline{\lambda}$ are $\mathbf{v} = \begin{bmatrix} 1+3i \\ 2 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \end{bmatrix} + i \begin{bmatrix} 3 \\ 0 \end{bmatrix}$, $\overline{\mathbf{v}} = \begin{bmatrix} 1-3i \\ 2 \end{bmatrix}$. so $P = \begin{bmatrix} 1 & 3 \\ 2 & 0 \end{bmatrix}$

c) From above computation we get decomposition

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

□

3 Exams

3.1 Exam 1

Problem 3.1.1.

a) **(15 pt)** Write linear system

$$\begin{cases} 2x_1 + 4x_2 + 2x_3 = 2 \\ -x_1 + 3x_2 + x_3 + 2x_4 = 1 \\ 3x_1 + 2x_2 - x_3 + 2x_4 = -1 \end{cases}$$

as a matrix equation, and then solve it, write your solution in parametric vector form.

b) **(5 pt)** Is $\left\{ \mathbf{a}_1 = \begin{bmatrix} 2 \\ -1 \\ 3 \end{bmatrix}, \mathbf{a}_2 = \begin{bmatrix} 4 \\ 3 \\ 2 \end{bmatrix}, \mathbf{a}_3 = \begin{bmatrix} 2 \\ 1 \\ -1 \end{bmatrix}, \mathbf{a}_4 = \begin{bmatrix} 0 \\ 2 \\ 2 \end{bmatrix} \right\}$ linearly independent? Why or why not.

c) **(4 pt)** What is the span of $\{\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3, \mathbf{a}_4\}$

Solution.

a) We could rewrite the linear system as

$$A\mathbf{x} = \begin{bmatrix} 2 & 4 & 2 & 0 \\ -1 & 3 & 1 & 2 \\ 3 & 2 & -1 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 2 \\ 1 \\ -1 \end{bmatrix} = \mathbf{b}$$

$$\begin{bmatrix} A & \mathbf{b} \end{bmatrix} \xrightarrow[R3 \to R3+3R2]{R1 \to R1+2R2} \begin{bmatrix} 0 & 10 & 4 & 4 & 4 \\ -1 & 3 & 1 & 2 & 1 \\ 0 & 11 & 2 & 8 & 2 \end{bmatrix} \xrightarrow{R3 \to R3-R1} \begin{bmatrix} 0 & 10 & 4 & 4 & 4 \\ -1 & 3 & 1 & 2 & 1 \\ 0 & 1 & -2 & 4 & -2 \end{bmatrix}$$

$$\xrightarrow[(-1)R2]{R1 \to R1-10R3} \begin{bmatrix} 0 & 0 & 24 & -36 & 24 \\ 1 & -3 & -1 & -2 & -1 \\ 0 & 1 & -2 & 4 & -2 \end{bmatrix} \xrightarrow[\text{move } R3 \text{ to the third row}]{R1/24} \begin{bmatrix} 1 & -3 & -1 & -2 & -1 \\ 0 & 1 & -2 & 4 & -2 \\ 0 & 0 & 1 & -\frac{3}{2} & 1 \end{bmatrix}$$

$$\xrightarrow{R2 \to R2+2R3} \begin{bmatrix} 1 & -3 & -1 & -2 & -1 \\ 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & -\frac{3}{2} & 1 \end{bmatrix} \xrightarrow{R1 \to R1+3R2+R3} \begin{bmatrix} 1 & 0 & 0 & -\frac{1}{2} & 0 \\ 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & -\frac{3}{2} & 1 \end{bmatrix}$$

So the solution is

$$\begin{cases} x_1 = \frac{1}{2}x_4 \\ x_2 = -x_4 \\ x_3 = 1 + \frac{3}{2}x_4 \\ x_4 \text{ is free} \end{cases} \Rightarrow \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix} + x_4 \begin{bmatrix} \frac{1}{2} \\ -1 \\ \frac{3}{2} \\ 1 \end{bmatrix}$$

b) Note the RREF of A is $\begin{bmatrix} 1 & 0 & 0 & -\frac{1}{2} \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & -\frac{3}{2} \end{bmatrix}$ Since the RREF of A doesn't have pivots in each column, $\{\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3, \mathbf{a}_4\}$ is linearly dependent.

c) Since the RREF A has a pivot in each row, the columns of A (i.e. $\{\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3, \mathbf{a}_4\}$) span $\mathbb{R}^3$

□

Problem 3.1.2. Suppose $T : \mathbb{R}^3 \to \mathbb{R}^3$ is a linear transformation defined by $T\left(\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}\right) = \begin{bmatrix} x_1 + 2x_2 + 2x_3 \\ 2x_1 + 3x_3 \\ 3x_1 + 5x_2 + 6x_3 \end{bmatrix}$. Denote the standard matrix for T as A.

a) (3 pt) Evaluate A.

b) (4 pt) Is T is onto? Is T one-to-one?

c) (17 pt) Is T is invertible? If so, what is the standard matrix for T^{-1}?

d) (15 pt) Find A^T ,$(A^T)^{-1}$.

Solution.

a) The standard matrix for T is $A = \begin{bmatrix} 1 & 2 & 2 \\ 2 & 0 & 3 \\ 3 & 5 & 6 \end{bmatrix}$

b) T is onto and one-to-to.

c) T is invertible. And the standard matrix of T^{-1} is A^{-1}

$$\left[\begin{array}{c|c} A & I \end{array}\right] = \left[\begin{array}{ccc|ccc} 1 & 2 & 2 & 1 & 0 & 0 \\ 2 & 0 & 3 & 0 & 1 & 0 \\ 3 & 5 & 6 & 0 & 0 & 1 \end{array}\right] \underbrace{\substack{R2\to R2-2R1 \\ R3\to R3-3R1}} \left[\begin{array}{ccc|ccc} 1 & 2 & 2 & 1 & 0 & 0 \\ 0 & -4 & -1 & -2 & 1 & 0 \\ 0 & -1 & 0 & -3 & 0 & 1 \end{array}\right]$$

$$\underbrace{\substack{R2\to R2-4R3 \\ R3\to R1+2R3}} \left[\begin{array}{ccc|ccc} 1 & 0 & 2 & -5 & 0 & 2 \\ 0 & 0 & -1 & 10 & 1 & -4 \\ 0 & -1 & 0 & -3 & 0 & 1 \end{array}\right] \underbrace{\substack{(-1)R2 \\ (-1)R3}} \left[\begin{array}{ccc|ccc} 1 & 0 & 2 & -5 & 0 & 2 \\ 0 & 0 & 1 & -10 & -1 & 4 \\ 0 & 1 & 0 & 3 & 0 & -1 \end{array}\right]$$

$$\underbrace{R2\leftrightarrow R3} \left[\begin{array}{ccc|ccc} 1 & 0 & 2 & -5 & 0 & 2 \\ 0 & 1 & 0 & 3 & 0 & -1 \\ 0 & 0 & 1 & -10 & -1 & 4 \end{array}\right] \underbrace{R1\to R1-2R3} \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 15 & 2 & -6 \\ 0 & 1 & 0 & 3 & 0 & -1 \\ 0 & 0 & 1 & -10 & -1 & 4 \end{array}\right]$$

$$= \left[\begin{array}{c|c} I & A^{-1} \end{array}\right]$$

Hence $A^{-1} = \begin{bmatrix} 15 & 2 & -6 \\ 3 & 0 & -1 \\ -10 & -1 & 4 \end{bmatrix}$

d) $A^T = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 0 & 5 \\ 2 & 3 & 6 \end{bmatrix}$, $(A^T)^{-1} = (A^{-1})^T = \begin{bmatrix} 15 & 3 & -10 \\ 2 & 0 & -1 \\ -6 & -1 & 4 \end{bmatrix}$

□

Problem 3.1.3.

a) (5 pt) Suppose $A = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & -2 \\ 0 & 0 & h & 0 & k & 3 \\ 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix}$ is of reduced row echelon form (RREF), could we know what h, k are? If not, please explain why, if so, please give their values.

b) (3 pt) `TRUE` or `FALSE`. If $T : \mathbb{R}^3 \to \mathbb{R}^3$ is a one-to-one linear transformation, then it is also onto.

c) (3 pt) `TRUE` or `FALSE`. If A is a 3 by 3 matrix and A^3 is invertible, then so is A.

d) (3 pt) `TRUE` or `FALSE`. If A is a 4 by 3 matrix, then $A\mathbf{x} = \mathbf{b}$ cannot have non-trivial solution(s)

Solution.

a) $h = 1$, $k = 0$. Since h has be in a pivot position, and k is in a pivot column.

b) `TRUE`. Since its standard matrix would have a pivot in each column, that is three pivots, so there must be a pivot in each row also, hence T is also onto.

c) `TRUE`. Since A^3 is invertible, $0 \neq \det(A^3) = (\det A)^3 \Rightarrow \det A \neq 0$, hence A is invertible.

d) `FALSE`. Take A to be $\begin{bmatrix}1&1&1\\0&0&0\\0&0&0\\0&0&0\end{bmatrix}$, and $\mathbf{b}=\mathbf{0}$, and the linear system have two free variables.

$\square$

Problem 3.1.4. Suppose $A=\begin{bmatrix}1&0&3&0\\2&0&6&1\\2&1&2&-4\\3&0&2&1\end{bmatrix}$.

a) (16 pt) Evaluate $\det A$.

b) (3 pt) Is A invertible?

c) (4 pt) What is $\det(-2A)$?

Solution.

a)

$$\begin{vmatrix}1&0&3&0\\2&0&6&1\\2&1&2&-4\\3&0&2&1\end{vmatrix} \xlongequal{\text{cofactor expansion across second column}} 1(-1)^{3+2}\begin{vmatrix}1&3&0\\2&6&1\\3&2&1\end{vmatrix}$$

$$\xlongequal[R3\to R3-3R1]{R2\to R2-2R1}(-1)\begin{vmatrix}1&3&0\\0&0&1\\0&-7&1\end{vmatrix}\xlongequal{R2\leftrightarrow R3}(-1)(-1)\begin{vmatrix}1&3&0\\0&-7&1\\0&0&1\end{vmatrix}=(-1)(-1)1(-7)1=-7$$

b) A is invertible since $\det A=-7\neq 0$

c) Since A is a 4 by 4 matrix, $\det(-2A)=(-2)^4\det A=16\cdot(-7)=-112$

$\square$

3.2 Exam 2

Problem 3.2.1. Suppose $A=\begin{bmatrix}1&-1&5&1&6&0\\2&0&3&5&3&6\\0&1&-4&2&-1&2\\3&-2&12&4&10&4\end{bmatrix}$.

a) (6 pt) Please find a basis for $\operatorname{Col}A$.

b) (6 pt) Please find a basis for $\operatorname{Row}A$.

c) (6 pt) Please find a basis for $\operatorname{Nul}A$.

d) (6 pt) What is $\dim\operatorname{Nul}A^T$?

Solution. First note that

$$A\sim\begin{bmatrix}1&-1&5&1&6&0\\0&1&-4&2&-1&2\\0&0&1&-1&-7&2\\0&0&0&0&0&0\end{bmatrix}\sim\begin{bmatrix}1&0&0&4&12&0\\0&1&0&-2&-29&10\\0&0&1&-1&-7&2\\0&0&0&0&0&0\end{bmatrix}$$

So we may conclude that the pivot columns are the 1st, 2nd, 3rd column

a) A basis for $\operatorname{Col}A$ could be $\left\{\begin{bmatrix}1\\2\\0\\3\end{bmatrix},\begin{bmatrix}-1\\0\\1\\-2\end{bmatrix},\begin{bmatrix}5\\3\\-4\\12\end{bmatrix}\right\}$

b) A basis for Row A could be

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

c) A basis for Nul A could be $\left\{\begin{bmatrix}-4\\2\\1\\1\\0\\0\end{bmatrix}, \begin{bmatrix}-12\\29\\7\\0\\1\\0\end{bmatrix}, \begin{bmatrix}0\\-10\\-2\\0\\0\\1\end{bmatrix}\right\}$

d) By the rank nullity theorem. we have $\dim \operatorname{Nul} A^T = 4 - \operatorname{rank} A^T = 4 - \operatorname{rank} A = 4 - 3 = 1$

□

Problem 3.2.2. Suppose $A = \begin{bmatrix}4 & 2 & -2\\4 & 2 & 4\\-1 & 1 & 5\end{bmatrix}$.

a) (15 pt) Find all eigenvalues and corresponding eigenvectors

b) (8 pt) Diagonalize A.

c) (8 pt) Evaluate A^{99}.

Solution.

a) First we evaluate the characteristic polynomial of A

$$\det(tI - A) = \begin{vmatrix}t-4 & -2 & 2\\-4 & t-2 & -4\\1 & -1 & t-5\end{vmatrix} \xlongequal[R2\to R2+4R3]{R1\to R1-(t-4)R3} \begin{vmatrix}0 & t-6 & -t^2+9t-18\\0 & t-6 & 4t-24\\1 & -1 & t-5\end{vmatrix}$$

$$\xlongequal{R1\to R1-R2} \begin{vmatrix}0 & 0 & -t^2+5t+6\\0 & t-6 & 4t-24\\1 & -1 & t-5\end{vmatrix} = (t-6)^2(t+1)$$

So the eigenvalues are $\lambda_1 = \lambda_2 = 6$, $\lambda_3 = -1$. For the 6-eigenspace, we consider

$$\left[\begin{array}{c|c}6I - A & \mathbf{0}\end{array}\right] = \left[\begin{array}{ccc|c}2 & -2 & 2 & 0\\-4 & 4 & -4 & 0\\1 & -1 & 1 & 0\end{array}\right] \sim \left[\begin{array}{ccc|c}1 & -1 & 1 & 0\\0 & 0 & 0 & 0\\0 & 0 & 0 & 0\end{array}\right]$$

So we get two basis vectors $\left\{\mathbf{v}_1 = \begin{bmatrix}1\\1\\0\end{bmatrix}, \mathbf{v}_2 = \begin{bmatrix}-1\\0\\1\end{bmatrix}\right\}$. For the (-1)-eigenspace, we consider

$$\left[\begin{array}{c|c}-I - A & \mathbf{0}\end{array}\right] = \left[\begin{array}{ccc|c}-5 & -2 & 2 & 0\\-4 & -3 & -4 & 0\\1 & -1 & -6 & 0\end{array}\right] \sim \left[\begin{array}{ccc|c}1 & 0 & -2 & 0\\0 & 1 & 4 & 0\\0 & 0 & 0 & 0\end{array}\right]$$

So we get a basis vector $\left\{\mathbf{v}_3 = \begin{bmatrix}2\\-4\\1\end{bmatrix}\right\}$

b) From above, we get the diagonalization

$$A = PDP^{-1} = \begin{bmatrix}1 & -1 & 2\\1 & 0 & -4\\0 & 1 & 1\end{bmatrix}\begin{bmatrix}6 & 0 & 0\\0 & 6 & 0\\0 & 0 & -1\end{bmatrix}\begin{bmatrix}\frac{4}{7} & \frac{3}{7} & \frac{4}{7}\\-\frac{1}{7} & \frac{1}{7} & \frac{6}{7}\\\frac{1}{7} & -\frac{1}{7} & \frac{1}{7}\end{bmatrix}$$

c)

$$A^{99} = PD^{99}P^{-1} = \begin{bmatrix} 1 & -1 & 2 \\ 1 & 0 & -4 \\ 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} 6^{99} & 0 & 0 \\ 0 & 6^{99} & 0 \\ 0 & 0 & -1 \end{bmatrix} \begin{bmatrix} \frac{4}{7} & \frac{3}{7} & \frac{4}{7} \\ -\frac{1}{7} & \frac{1}{7} & \frac{6}{7} \\ \frac{1}{7} & -\frac{1}{7} & \frac{1}{7} \end{bmatrix}$$

□

Problem 3.2.3. Suppose V is the set of 2 by 2 matrices that diagonal elements sum to 0. for example $\begin{bmatrix} 3 & 7 \\ 2 & -3 \end{bmatrix}$ is in V since $3 + (-3) = 0$ while $\begin{bmatrix} 3 & 4 \\ 2 & -2 \end{bmatrix}$ is not because $3 + (-2) = 1 \neq 0$.

a) (10 pt) Show that V is a vector space.

b) (10 pt) Find a basis $\mathcal{B}$ (this cannot simply be $\mathcal{C}$) of V and justify your choice.

c) (8 pt) Explain why $\mathcal{C} = \left\{ C_1 = \begin{bmatrix} 2 & 1 \\ 1 & -2 \end{bmatrix}, C_2 = \begin{bmatrix} 1 & 2 \\ -1 & 0 \end{bmatrix}, C_3 = \begin{bmatrix} 3 & -3 \\ 2 & -3 \end{bmatrix} \right\}$ also form a basis for V.

d) (8 pt) Find the change-of-coordinates matrix $\underset{\mathcal{C}\leftarrow\mathcal{B}}{P}$.

Solution.

a) Consider linear transformation $T : M_{2\times 2}(\mathbb{R}) \to \mathbb{R}$, $\begin{bmatrix} a & b \\ c & d \end{bmatrix} \mapsto a + d$, we see that $V = \ker T$ which is a subspace of $M_{2\times 2}(\mathbb{R})$

b) Note that $V = \left\{ \begin{bmatrix} a & b \\ c & -a \end{bmatrix} \right\}$ A basis for V could be $\left\{ B_1 = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}, B_2 = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, B_3 = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} \right\}$, this is spanning since

$$\begin{bmatrix} a & b \\ c & -a \end{bmatrix} = a \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} + b = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} + c \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}$$

This is linearly independent because if a linear combination

$$a \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} + b = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} + c \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} a & b \\ c & -a \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$

Then $a = b = c = 0$.

c) The $\mathcal{B}$ coordinates of $\mathcal{C}$ is $\left\{ [C_1]_{\mathcal{B}} = \begin{bmatrix} 2 \\ 1 \\ 1 \end{bmatrix}, [C_2]_{\mathcal{B}} = \begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix}, [C_3]_{\mathcal{B}} = \begin{bmatrix} 3 \\ -3 \\ 2 \end{bmatrix} \right\}$, can we can show that they are linearly independent

d) From above we know

$$\underset{\mathcal{B}\leftarrow\mathcal{C}}{P} = \begin{bmatrix} [C_1]_{\mathcal{B}} & [C_2]_{\mathcal{B}} & [C_3]_{\mathcal{B}} \end{bmatrix} = \begin{bmatrix} 2 & 1 & 3 \\ 1 & 2 & -3 \\ 1 & -1 & 2 \end{bmatrix}$$

So we know

$$\underset{\mathcal{C}\leftarrow\mathcal{B}}{P} = \left(\underset{\mathcal{B}\leftarrow\mathcal{C}}{P} \right)^{-1} = \begin{bmatrix} 2 & 1 & 3 \\ 1 & 2 & -3 \\ 1 & -1 & 2 \end{bmatrix}^{-1} = \begin{bmatrix} -\frac{1}{12} & \frac{5}{12} & \frac{3}{4} \\ \frac{5}{12} & -\frac{1}{12} & -\frac{3}{4} \\ \frac{1}{4} & -\frac{1}{4} & -\frac{1}{4} \end{bmatrix}$$

□

Problem 3.2.4. Determine whether the following is true.

a) **(3 pt)** `TRUE` or `FALSE`. If the eigenvalues for a 3 by 3 matrix A are $1, -1, 3$, then A is diagonalizable.

b) **(3 pt)** `TRUE` or `FALSE`. If both $\mathbf{v}_1, \mathbf{v}_2$ are eigenvectors of a matrix A, then $\mathbf{v}_1 + \mathbf{v}_2$ is again an eigenvector for A.

c) **(3 pt)** `TRUE` or `FALSE`. If A is similar to $\begin{bmatrix} 2 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}$, then A is not diagonalizable.

Solution.

a) `TRUE`. By Theorem 1.23.3

b) `FALSE`. Consider $A = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$, $\mathbf{v}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ is eigenvector for $\lambda_1 = 1$ and $\mathbf{v}_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$ is eigenvector for $\lambda_2 = -1$, however $\mathbf{v}_1 + \mathbf{v}_2 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ is not an eigenvector.

c) `TRUE`. Since $\begin{bmatrix} 2 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}$ is not diagonalizable, the geometric multiplicity for eigenvalue 1 is only 1 which is less than its algebraic multiplicity, which is 2.

□

3.3 Final

Problem 3.3.1. Suppose $A = \begin{bmatrix} 1 & 0 & b & 0 & 2 \\ 0 & 0 & a & 0 & -3 \\ 0 & 0 & 0 & 1 & 1 \end{bmatrix}$ is of reduced row echelon form (RREF).

a) **(4 pt)** Evaluate a, b.

b) **(6 pt)** Find a basis for $\operatorname{Col} A$.

c) **(8 pt)** Find a basis for $(\operatorname{Row} A)^\perp$.

Solution.

a) $a = 1, b = 0$.

b) A basis for $\operatorname{Col} A$ could be $\left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \right\}$

c) A basis for $(\operatorname{Row} A)^\perp = \operatorname{Nul} A$ is $\left\{ \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} -2 \\ 0 \\ 3 \\ -1 \\ 1 \end{bmatrix} \right\}$

□

Problem 3.3.2. Recall that $\mathcal{E} = \{1, t, t^2\}$ is the *standard basis* for $\mathbb{P}_2$. Given that $\mathcal{B} = \{1 + t, 2 - t^2, 3 - t + 2t^2\}$, $\mathcal{C} = \{2, 1 + t, -t^2\}$ are bases for $\mathbb{P}_2$.

a) **(15 pt)** Find the change-of-coordinate matrix $\underset{\mathcal{C} \leftarrow \mathcal{B}}{P}$.

b) **(15 pt)** Suppose $T : \mathbb{P}_2 \to \mathbb{P}_2$ is a linear transformation with $\mathcal{B}$ matrix $[T]_{\mathcal{B}} = \begin{bmatrix} 1 & 0 & 1 \\ 0 & -1 & 0 \\ 1 & 0 & 1 \end{bmatrix}$, please find the $\mathcal{C}$ matrix $[T]_{\mathcal{C}}$ for T.

c) (15 pt) Suppose $p(t)$ is a polynomial in $\mathbb{P}_2$ such that $[p(t)]_{\mathcal{B}} = \begin{bmatrix} 0 \\ 2 \\ -1 \end{bmatrix}$. Evalue polynomial $T(p(t))$ in $\mathbb{P}_2$.

Solution.

a) First we evaluate matrices

$$\underset{\mathcal{E}\leftarrow\mathcal{B}}{P} = \begin{bmatrix} 1 & 2 & 3 \\ 1 & 0 & -1 \\ 0 & -1 & 2 \end{bmatrix}$$

$$\underset{\mathcal{E}\leftarrow\mathcal{C}}{P} = \begin{bmatrix} 2 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{bmatrix}$$

$$\underset{\mathcal{C}\leftarrow\mathcal{B}}{P} = \left(\underset{\mathcal{E}\leftarrow\mathcal{C}}{P}\right)^{-1} \underset{\mathcal{E}\leftarrow\mathcal{B}}{P} = \begin{bmatrix} 0 & 1 & 2 \\ 1 & 0 & -1 \\ 0 & 1 & -2 \end{bmatrix}$$

b)

$$\underset{\mathcal{B}\leftarrow\mathcal{C}}{P} = \left(\underset{\mathcal{C}\leftarrow\mathcal{B}}{P}\right)^{-1} = \begin{bmatrix} \frac{1}{4} & 1 & -\frac{1}{4} \\ \frac{1}{2} & 0 & \frac{1}{2} \\ \frac{1}{4} & 0 & -\frac{1}{4} \end{bmatrix}$$

$$[T]_{\mathcal{C}} = \underset{\mathcal{C}\leftarrow\mathcal{B}}{P}[T]_{\mathcal{B}}\underset{\mathcal{B}\leftarrow\mathcal{C}}{P} = \begin{bmatrix} \frac{1}{2} & 2 & -\frac{3}{2} \\ 0 & 0 & 0 \\ -\frac{3}{2} & -2 & \frac{1}{2} \end{bmatrix}$$

c) First note that

$$[T(p(t))]_{\mathcal{B}} = [T]_{\mathcal{B}}[p(t)]_{\mathcal{B}} = \begin{bmatrix} -1 \\ -2 \\ -1 \end{bmatrix}$$

So $T(p(t)) = (-1)(1+t) + (-2)(2-t^2) + (-1)(3-t+2t^2) = -8$.

$\square$

Problem 3.3.3. Consider linear transformation $T : \mathbb{R}^3 \to \mathbb{R}^3$ which is defined by $T\left(\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}\right) = \begin{bmatrix} x_1 + x_2 - 2x_3 \\ -x_1 + 2x_2 \\ x_1 + 4x_2 - 4x_3 \end{bmatrix}$.

a) (5 pt) Is T onto? Explain why.

b) (5 pt) Is T one-to-one? Explain why.

c) (5 pt) Is T is invertible? If so, please find T^{-1}.

Solution. The standard matrix for T is $A = \begin{bmatrix} 1 & 1 & -2 \\ -1 & 2 & 0 \\ 1 & 4 & -4 \end{bmatrix} \sim \begin{bmatrix} 1 & 1 & -2 \\ 0 & 3 & -2 \\ 0 & 0 & 0 \end{bmatrix}$

a) T is not onto since there is not a pivot on each row of A.

b) T is not one-to-one since there is not a pivot on each column of A.

c) T is not invertible, since T is neither one-to-one nor onto.

$\square$

Problem 3.3.4. Suppose $A = \begin{bmatrix} 3 & -4 \\ 2 & -1 \end{bmatrix}$.

a) (8 pt) Please find the eigenvalues of A (they may be complex).

b) (12 pt) Please find the corresponding eigenvectors.

c) (6 pt) Write down a factorization of the form $A = PCP^{-1}$, where C is of the form $\begin{bmatrix} a & -b \\ b & a \end{bmatrix}$.

Solution.

a) The eigenvalues of A are $\lambda = 1 - 2i$ and $\overline{\lambda} = 1 + 2i$, so we have $a = 1$, $b = 2$.

b) The eigenvectors of A are $\mathbf{v} = \begin{bmatrix} 1-i \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix} + i \begin{bmatrix} -1 \\ 0 \end{bmatrix}$ and $\overline{\mathbf{v}} = \begin{bmatrix} 1 \\ 1 \end{bmatrix} - i \begin{bmatrix} -1 \\ 0 \end{bmatrix}$.

c)

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

□

Problem 3.3.5. Suppose $A = \begin{bmatrix} 1 & 3 \\ 2 & 2 \\ 3 & 1 \end{bmatrix}$, $\mathbf{b} = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$.

a) (6 pt) Show that the linear system is linearly inconsistent.

b) (10 pt) Find the least-square solution(s).

Solution.

a)

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

Which has a pivot in the last column.

b) $A^TA = \begin{bmatrix} 14 & 10 \\ 10 & 14 \end{bmatrix}$, $A^T\mathbf{b} = \begin{bmatrix} 4 \\ 4 \end{bmatrix}$, then $\hat{\mathbf{x}} = (A^TA)^{-1}(A^T\mathbf{b}) = \begin{bmatrix} \frac{1}{6} \\ \frac{1}{6} \end{bmatrix}$ is the least-square solution.

□

Problem 3.3.6. Suppose $A = \begin{bmatrix} 4 & 1 & 0 \\ 1 & 4 & 0 \\ 0 & 0 & 5 \end{bmatrix}$.

a) (6 pt) Please find the characteristic polynomial of A.

b) (4 pt) Please find the eigenvalues of A.

c) (12 pt) Please find an orthonormal basis for $\mathbb{R}^3$ consists of eigenvectors.

d) (8 pt) Please orthogonal diagonalize A.

Solution.

a) The characteristic polynomial is $(t-3)(t-5)^2$.

b) The eigenvalues of A are $\lambda_1 = 3$, $\lambda_2 = \lambda_3 = 5$.

c) An orthonormal basis consists of eigenvectors could be $\left\{ \begin{bmatrix} \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \\ 0 \end{bmatrix}, \begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \right\}$.

d)

$$A = \begin{bmatrix} 4 & 1 & 0 \\ 1 & 4 & 0 \\ 0 & 0 & 5 \end{bmatrix} = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 3 & 0 & 0 \\ 0 & 5 & 0 \\ 0 & 0 & 5 \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} & 0 \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ 0 & 0 & 1 \end{bmatrix} = PDP^T$$

$\square$

Index