First Recitation Class
Linear Algebra

YAO Shaoxiong

UM-SJTU Joint Institute

March 5, 2019

Table of contents

From Linear Equations to Matrices

For a system of linear equations of ***real*** numbers,

$$
\begin{aligned}
a_{11}x_1 + a_{12}x_2 + \ldots + a_{1m}x_m &= b_1 \\
a_{21}x_1 + a_{22}x_2 + \ldots + a_{2m}x_m &= b_2 \\
&\vdots \\
a_{n1}x_1 + a_{n2}x_2 + \ldots + a_{nm}x_m &= b_n
\end{aligned}
$$

here a_{ij}, b_i are coefficients and x_i are unknowns.
We write it as in the matrix form.

$$
\begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nm} \end{bmatrix}
\begin{bmatrix} x_1 \\ \vdots \\ x_m \end{bmatrix}
=
\begin{bmatrix} b_1 \\ \vdots \\ b_n \end{bmatrix}
$$

Two ways to understand linear system

Method 1. Intersection of n "planes" in $\mathbb{R}^m$.
For $n = 2$ and $m = 2$

$$a_{11}x_1 + a_{12}x_2 = b_1$$
$$a_{21}x_1 + a_{22}x_2 = b_2$$

The solution is the intersection of two lines defined by two equations.
Method 2. Span m vectors of in $\mathbb{R}^n$.
We rewrite the equation,

$$\begin{pmatrix} a_{11} \\ a_{21} \end{pmatrix} x_1 + \begin{pmatrix} a_{21} \\ a_{22} \end{pmatrix} x_2 = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix}$$

Solution is the appropriate coefficients to make two vectors span the third vector.

Matrices

Definition

We call

$$A := \begin{bmatrix} a_{11} & \cdots & a_{1m} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nm} \end{bmatrix}$$

a $n \times m$ matrix. We can also write $A = (a_{ij}), i = \overline{1, n}, j = \overline{1, m}$.

Notice:

There are several types of important matrices.

- If $n = m$, A is a square matrix.
- If $a_{ij} = 0$ for $i \neq j$, A is diagonal.
- If $a_{ij} = 0$ for $i \geq j$, A is upper triangle.

Note:

A $n \times m$ matrix has n ***rows*** and m ***columns***. It come from n ***equations*** and m ***unknowns***.

Vectors

Definition

We define vectors as $n \times 1$ matrices,

$$x = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}.$$

We denote the space of all vectors

$$\mathbb{R}^n := \{\overline{a} = (a_1, ..., a_n) : a_i \in \mathbb{R}, \forall i \in \overline{1, n}\}.$$

A $1 \times m$ matrix is called a ***row*** vector.

Vectors and Matrices

Notice:

We here define the product between appropriate matrix and vectors.

$$\begin{bmatrix} a_{11} & \cdots & a_{1m} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nm} \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_m \end{bmatrix} = \begin{bmatrix} a_{11}x_1 + \cdots + a_{1m}x_m \\ \vdots \\ a_{n1}x_1 + \cdots + a_{nm}x_m \end{bmatrix}$$

A special type of $n \times n$ matrices is called identity matrix:

$$I = \begin{bmatrix} 1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 1 \end{bmatrix}.$$

When they operate on vectors $x \in \mathbb{R}^n$, we have

$$Ix = x.$$

Solution set of a system of linear equations

Definition.

For a system of linear equations, the set of all solutions is called the solution set,

$$S = \{x \in \mathbb{R}^n : x \text{ is a solution}\}$$

Claim.

There are only three possible cases for the solution set,

- $S = \emptyset$,
- S contains only one element,
- S is infinite.

Question.

Why there are only three possible cases?

Gaussian-Jordan Elimination

Restrictions.

To solve a system, we can only apply following types of operations.

- Times an equation with a constant.
- Add an equation to another.
- Exchange two rows.

Motivation.

For $n \times n$ case, the identity form can be solved.
We want to transform a system to this form.

$$\begin{aligned} x + 0 + 0 &= b_1 \\ 0 + y + 0 &= b_2 \\ 0 + 0 + z &= b_3 \end{aligned}$$

Gaussian-Jordan Elimination

Solution.

Our solution contains two steps:

- Gaussian: eliminate the lower triangle,

$$\begin{aligned} a_{11}x + a_{12}y + a_{13}z &= \diamond \\ a_{21}x + a_{22}y + a_{23}z &= \diamond \\ a_{31}x + a_{32}y + a_{33}z &= \diamond \,. \end{aligned} \quad \Rightarrow \quad \begin{aligned} x + \star + \star &= \diamond \\ 0 + y + \star &= \diamond \\ 0 + 0 + z &= \diamond \,. \end{aligned}$$

- Jordan: eliminate the upper triangle,

$$\begin{aligned} x + \star + \star &= \diamond \\ 0 + y + \star &= \diamond \\ 0 + 0 + z &= \diamond \,. \end{aligned} \quad \Rightarrow \quad \begin{aligned} x + 0 + 0 &= \diamond \\ 0 + y + 0 &= \diamond \\ 0 + 0 + z &= \diamond \,. \end{aligned}$$

Gaussian-Jordan Elimination

Example.

$$\begin{bmatrix} 2 & 4 & 1 & 2 \\ 2 & 5 & 4 & 4 \\ 4 & 9 & 5 & 11 \\ 2 & 4 & 2 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}$$

Additional operations and failure

If we have zero at the leading position, we need to exchange rows.
If all lower elements are zeros, we fail.

Gaussian-Jordan Algorithm

We apply this method to $n \neq m$.

Example.

$$\begin{bmatrix} 1 & 2 & 2 & 2 \\ 2 & 4 & 6 & 8 \\ 3 & 6 & 8 & 10 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 2 \\ 3 \\ 4 \end{bmatrix}$$

Notice:
Although we cannot get to the identity form, the result system is easy to solve.

Reduced Row Echelon Form

Definiotion

(i) If a row is non-zero and the first non-zero element is 1, this element is called ***leading 1*** in this row.

(ii) If a column contains a ***leading 1***, all other entries are 0.

(iii) If a row contains a ***leading 1***, all ***leading 1***s of the above rows are on the left.

We will write $rref(A)$ to refer the reduced row echelon form of A.

Question:

Why this form is easy to solve?

Rank

Definition

The number of leading 1's in rref(A) is called the ***rank*** of the matrix A.

Notice:

We can exchange columns (rename variables in the equations) to write $rref(A)$ as follows,

$$rref(A) = \begin{bmatrix} I & F \\ 0 & 0 \end{bmatrix}.$$

We denote $\text{rank}A = r$. I is the $r \times r$ identity matrix and F is a $r \times (m - r)$ matrix.

Rank

We summarize all situations,

(i) $\text{rank}A = n = m$, $rref(A) = I$, the system will have one solution.

(ii) $\text{rank}A = n < m$, $rref(A) = [I \;\; F]$, the system will have infinitely many solutions.

(iii) $\text{rank}A = m < n$, $rref(A) = \left[\begin{smallmatrix} I \\ 0 \end{smallmatrix}\right]$, the system will either be inconsistent or have one solution.

(iv) $\text{rank}A < m, \text{rank}A < n, rref(A) = \left[\begin{smallmatrix} I & F \\ 0 & 0 \end{smallmatrix}\right]$, the system will either be inconsistent or have infinitely many solutions.

Review of Vectors in $\mathbb{R}^n$

$$\mathbb{R}^n = \{\overline{x} = (x_1, ..., x_n) : x_i \in \mathbb{R}^n, i = \overline{1, n}\}$$

Stanard Representation

For $x \in \mathbb{R}^n$, we can write it as

$$x = x_1 e_1 + \cdots + x_n e_n,$$

where

$$e_i = \begin{bmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{bmatrix}, 1 \text{ at ith entry.}$$

Properties of Vectors

Definition

For $\overline{a}, \overline{b}, \overline{c} \in \mathbb{R}^n$ and $k_1, k_2 \in \mathbb{R}$,

1. ***Addition between vectors***

   (i) $\overline{a} + \overline{b} = \overline{b} + \overline{a}$,

   (ii) $(\overline{a} + \overline{b}) + \overline{c} = \overline{a} + (\overline{b} + \overline{c})$.

2. ***Scalar product***

   (i) $k_1(k_2\overline{a}) = (k_1 k_2)\overline{a}$,

   (ii) $k(1)(\overline{a} + \overline{b}) = k_1\overline{a} + k_1\overline{b}$,

   (iii) $(k(1) + k(2))\overline{a} = k_1\overline{a} + k_2\overline{a}$.

Comment:

Vector space is a more general concept, functions and sequences with appropriate definition can also be vectors.

An important question is to find basis for these spaces. Can you find a basis in $C^\infty[a, b]$?

Projection and Dot Product

Deifnition

Geometrically, we define the projection of ***a vector on a line*** as a vector,

$$proj_L \overline{AB} = \overline{O_A O_B}.$$

The projection of ***a vector on a vectro*** as a number,

$$proj_{\overline{l}} = \pm|\overline{O_A O_B}|.$$

Definition

We define the dot product as

$$\overline{a} \cdot \overline{b} = |\overline{a}||\overline{b}| \cos(\overline{a}, \overline{b}).$$

Note that $\overline{a} \cdot \overline{b} = |\overline{a}| proj_{\overline{a}} \overline{b} = |\overline{b}| proj_{\overline{b}} \overline{a}$.

Properties of Dot Product

Geometrically, we have projection is linear, then we can prove

$$\overline{a} \cdot (\overline{b} + \overline{c}) = |\overline{a}| proj_{\overline{a}}(\overline{b} + \overline{c}) = |\overline{a}| proj_{\overline{a}}\overline{b} + |\overline{a}| proj_{\overline{a}}\overline{c}.$$

For standard basis, we have

$$e_i \cdot e_j = \begin{cases} 0, i \neq j \\ 1, i = j \end{cases}.$$

With standard representation, we have

$$\overline{a} \cdot \overline{b} = a_1 b_1 + a_2 b_2.$$

Properties of Dot Product

Perpendicular

Vectors $\overline{a}, \overline{b}$ are perpendicular if $\overline{a} \cdot \overline{b} = 0$.

Euclidean Norm

The length of a vector is

$$\|\overline{a}\| = \sqrt{\overline{a} \cdot \overline{a}}.$$

Definition

A vector $\overline{w} \in \mathbb{R}^n$ is called a unit vector is $\|\overline{a}\| = 1$. For $\overline{a} \in \mathbb{R}^n$ with $\overline{a} \neq 0$, vector $\frac{\overline{a}}{\|\overline{a}\|}$ is called the normalized vector of $\overline{a}$.

Cross Product

Definition

The cross product $\overline{a} \times \overline{b}$ of two vectors $\overline{a}, \overline{b} \in \mathbb{R}^3$ is a vector $\overline{c} \in \mathbb{R}^3$ that satisfies:

(i) $\overline{c} \perp \overline{a}$, $\overline{c} \perp \overline{b}$,

(ii) $\|\overline{c}\| = \|\overline{a}\| \cdot \|\overline{b}\| \cdot \sin(\overline{a}, \overline{b})$,

(iii) the ordered tuple $(\overline{a}, \overline{b}, \overline{c})$ is right-handed.

Notice:

We notice that $\overline{a} \times \overline{a} = 0$ and $\overline{a} \times \overline{b} = \|\overline{a}\| \cdot \|\overline{b}\|$ for $\overline{a} \perp \overline{b}$.
We can then use basis representation to calculate cross product.