

MAT

2041



MAT 2041 Linear Algebra and Applications

Dr. Cosme Louart

I. Introduction to Vectors

Vectors

Def. Column vectors

$$\text{Def. } \vec{v} = [v_1, v_2, \dots, v_i, v_n]^T = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix} \equiv (v_1, v_2, \dots, v_n)$$

e.g. Zero vector $\vec{0} = (0, 0, \dots, 0)$.

A list data structure in computer algorithms

A time series (a sequence of data points) in data science problems.

Example:
Color RGB Vectors

(1, 0, 0)	(0, 1, 0)	(0, 0, 1)
(1, 1, 0)	(1, 0.5, 0.5)	(0.5, 0.5, 0.5)

Operations

$$\vec{v} = (v_1, v_2, \dots, v_n), \vec{w} = (w_1, w_2, \dots, w_n)$$

$$\textcircled{1} \quad \vec{v} + \vec{w} = (v_1 + w_1, v_2 + w_2, \dots, v_n + w_n)$$

Vector Addition.

$$\textcircled{2} \quad c\vec{v} = (cv_1, cv_2, \dots, cv_n)$$

Vector Multiplication

$$\textcircled{3} \quad c\vec{v} + d\vec{w} = (cv_1 + dw_1, cv_2 + dw_2, \dots, cv_n + dw_n) \quad \text{Linear Combination}$$

Vector Norm ("Geometric length")

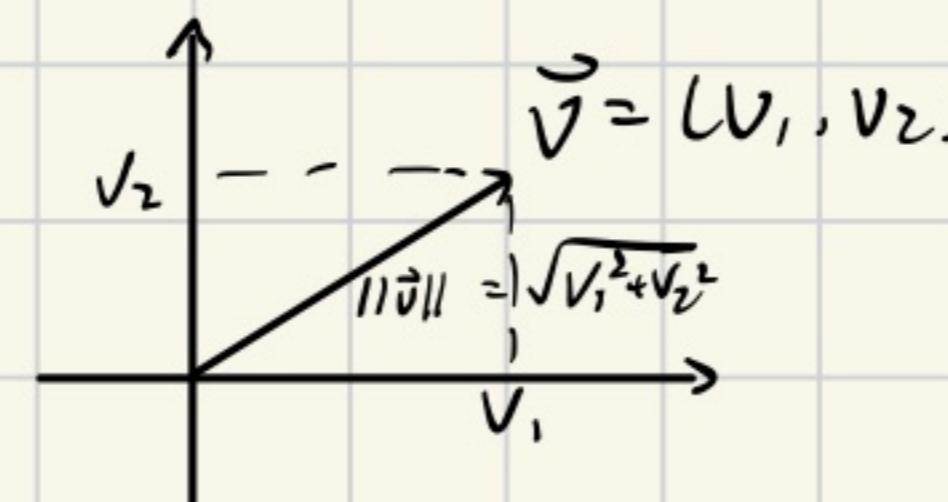
Consider a n -dimensional vector

Def. L_2 -norm

Let $\vec{v} = (v_1, v_2, \dots, v_n)$ be a n -length vector. The l_2 -norm of \vec{v} , denoted by $\|\vec{v}\|_2$, is defined as

$$\|\vec{v}\|_2 = (v_1^2 + v_2^2 + \dots + v_n^2)^{\frac{1}{2}}$$

(We often abbreviate $\|\vec{v}\|_2$ as $\|\vec{v}\|$).



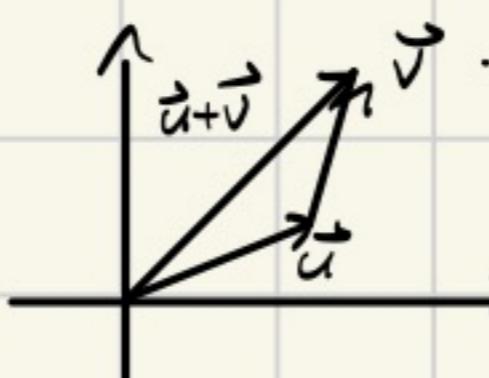
Properties:

if and only if

$$\textcircled{1} \quad \|\vec{v}\| \geq 0, \quad \|\vec{v}\| \text{ iff } \vec{v} = 0$$

$$\textcircled{2} \quad \|c\vec{v}\| = |c| \cdot \|\vec{v}\|$$

$$\textcircled{3} \quad \|\vec{u} + \vec{v}\| \leq \|\vec{u}\| + \|\vec{v}\|$$



Triangle inequality.

Unit vector

Def. A vector \vec{v} is called a unit vector if $\|\vec{v}\|=1$.

For any non-zero vector \vec{v} , $\frac{\vec{v}}{\|\vec{v}\|}$ is a unit vector.

$$\text{e.g. } \vec{e}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \quad \vec{e}_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \quad \vec{v} = \begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix}.$$

Inner Product (Dot Product)

Def. A dot product between 2 vectors (of the same size) $\vec{v} = (v_1, v_2, \dots, v_n), \vec{w} = (w_1, w_2, \dots, w_n)$ is defined as $\vec{v} \cdot \vec{w} = \vec{v}^T \vec{w} = \langle \vec{v}, \vec{w} \rangle = \sum_{i=1}^n v_i w_i$.

Recommend writing as $\langle \vec{v}, \vec{w} \rangle$ to avoid confusion.

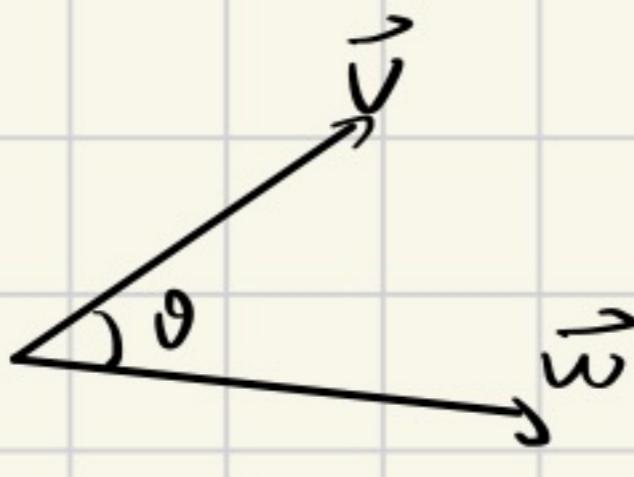
Properties of Inner Products.

- ① $\langle a\vec{v} + b\vec{u}, \vec{w} \rangle = a\langle \vec{v}, \vec{w} \rangle + b\langle \vec{u}, \vec{w} \rangle$ Linearity
- ② $\langle \vec{v}, \vec{w} \rangle = \langle \vec{w}, \vec{v} \rangle$ Symmetry
- ③ $\langle \vec{v}, \vec{v} \rangle = \|\vec{v}\|^2 \geq 0$ Positivity
- $\langle \vec{v}, \vec{v} \rangle = 0 \text{ iff } \vec{v} = \vec{0}$

Properties of Inner Product and Norm

1. Cosine Similarity

$$\cos \theta = \frac{\langle \vec{v}, \vec{w} \rangle}{\|\vec{v}\| \cdot \|\vec{w}\|}$$



2. Cauchy-Schwarz Inequality

$$|\langle \vec{v}, \vec{w} \rangle| \leq \|\vec{v}\| \cdot \|\vec{w}\|$$

3. Pythagoras Law

$$\|\vec{v}\|^2 + \|\vec{w}\|^2 = \|\vec{v} - \vec{w}\|^2 \text{ iff } \langle \vec{v}, \vec{w} \rangle = 0$$

II. System of Linear Equations

Linear Equations

Def. A linear equation is the equation of the form

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

where $a_i, b \in \mathbb{R}$. x_i are variables.

Vector form: $\vec{a} = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix}, \vec{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}, \langle \vec{a}, \vec{x} \rangle = b$.

$$\text{e.g. } -x_1 + 4x_2 = 2x_2 + 3x_3$$

System of LE

Def. An $m \times n$ system of linear equations is a collection of m linear equations and n variables.

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n = b_2 \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n = b_m \end{cases}$$

where $a_{ij}, b_i \in \mathbb{R}$. x_i are variables.

Row-vector Form (Unknown vector satisfies n linear equations simultaneously).

$$(x_1, x_2) \cdot (a_{11}, a_{12}) = b_1$$

$$(x_1, x_2) \cdot (a_{21}, a_{22}) = b_2$$

Column-vector form (Unknown combination of columns produces vector \vec{b}).

$$x_1 \cdot \begin{pmatrix} a_{11} \\ a_{21} \end{pmatrix} + x_2 \cdot \begin{pmatrix} a_{12} \\ a_{22} \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix}$$

Matrix form (Given matrix times unknown vector produces \vec{b}).

$$\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}$$

III. Matrix

• Definition

An $m \times n$ matrix A is a rectangular array of numbers with m rows and n columns in the form:

$$\Delta A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} =: (a_{ij})_{m \times n} \in \mathbb{R}^{m \times n}$$

where all a_{ij} are scalars. P.S. Matrix = row vector of column vectors
= column vector of row vectors

Dimension:

For an $m \times n$ matrix, the dimension of A is $m \times n$.

For an $m \times 1$ vector, the dimension of the vector is m , or $m \times 1$.

• Notation

- ① For a matrix A , a_{ij} is called the (i,j) -th entry (element) of A .
- ② Matrices are denoted by A, B, C, \dots
- ③ When $m = n$, A is called a square matrix. Otherwise, it's called a rectangular matrix.
- ④ When all entries are zeros, A is called a zero matrix. (Similar to zero vector).
- ⑤ When $m = 1$, A is a row vector.

When $n = 1$, A is a column vector.

$$\Delta \text{⑥ Column of a matrix: } \vec{a}_j = \begin{bmatrix} a_{1j} \\ \vdots \\ a_{mj} \end{bmatrix}$$

Row of a matrix: $\vec{a}^{(i)} = [a_{1i}, \dots, a_{ni}]$

• Matrix - vector Product

Def. Let $\vec{u} = \begin{bmatrix} u_1 \\ \vdots \\ u_n \end{bmatrix}$ and A be an $m \times n$ matrix $A = (a_{ij})$.

The matrix - vector product $A\vec{u}$ is an $m \times 1$ column vector:

$$A\vec{u} = \begin{bmatrix} \sum_{j=1}^n a_{1j} \cdot u_j \\ \sum_{j=1}^n a_{2j} \cdot u_j \\ \vdots \\ \sum_{j=1}^n a_{mj} \cdot u_j \end{bmatrix}$$

Remark: Dimension checking $\underset{m \times n}{A\vec{u}} = \vec{v}$

Properties:

① Row-form of matrix - vector product

Let A be an $m \times n$ matrix and \vec{b} an $n \times 1$ column vector. Suppose $A = \begin{bmatrix} \vec{a}^{(1)} \\ \vec{a}^{(2)} \\ \vdots \\ \vec{a}^{(m)} \end{bmatrix}$ then

$$A\vec{b} = \begin{bmatrix} \vec{a}^{(1)} \cdot \vec{b} \\ \vec{a}^{(2)} \cdot \vec{b} \\ \vdots \\ \vec{a}^{(m)} \cdot \vec{b} \end{bmatrix}$$

② Column-form of matrix - vector product

Let A be an $m \times n$ matrix and \vec{w} be an $n \times 1$ column vector. Suppose $A = [\vec{a}_1, \vec{a}_2, \dots, \vec{a}_m]$ then

$$A\vec{w} = w_1\vec{a}_1 + w_2\vec{a}_2 + \dots + w_m\vec{a}_m$$

IV. Elimination

Motivation

How to solve n by n system?

1. Eliminate one variable by subtracting one equation from others
2. Solve the remaining $(n-1)$ by $(n-1)$ system.
3. Continue the process until getting 1 variable and 1 equation.

Coefficient Matrix & Augmented Matrix

Def. Given a linear equation,

$$a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n = b_1$$

$$a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n = b_2$$

:

$$a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n = b_m$$

The coefficient matrix of the system is an $m \times n$ matrix:

$$A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix} = (a_{ij})_{m \times n}$$

The corresponding augmented matrix is:

$$[A|b] = \left[\begin{array}{cccc|c} a_{11} & a_{12} & \dots & a_{1n} & b_1 \\ a_{21} & a_{22} & \dots & a_{2n} & b_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} & b_m \end{array} \right]$$

Overdetermined, Underdetermined and Square

An $m \times n$ system of linear equations is

- 1^o overdetermined system if $m > n$. (tall)
- 2^o underdetermined system if $m < n$. (wide)
- 3^o square system if $m = n$. (square)

Special Matrices

Lower Triangular Matrix

Def. A square matrix $(L_{ij})_{n \times n}$ where
 $L_{ij} = 0$, for any $1 \leq i < j \leq n$.

$$\Rightarrow \begin{bmatrix} l_{11} & 0 & \dots & \dots & 0 \\ l_{21} & l_{22} & \dots & \dots & 0 \\ l_{31} & l_{32} & l_{33} & \dots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ l_{n1} & l_{n2} & l_{n3} & \dots & l_{nn} \end{bmatrix}$$

Upper Triangular Matrix

Def. A square matrix $(U_{ij})_{n \times n}$ where
 $U_{ij} = 0$, for any $1 \leq j < i \leq n$.

$$\Rightarrow \begin{bmatrix} u_{11} & u_{12} & u_{13} & \dots & u_{1n} \\ 0 & u_{22} & u_{23} & \dots & u_{2n} \\ 0 & 0 & u_{33} & \dots & u_{3n} \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & u_{nn} \end{bmatrix}$$

Diagonal Matrix

Def. A square matrix $(D_{ij})_{n \times n}$ where
 $D_{ij} = 0$, $\forall i \neq j$.

$$\Rightarrow \begin{bmatrix} d_{11} & 0 & 0 & \dots & 0 \\ 0 & d_{22} & 0 & \dots & 0 \\ 0 & 0 & d_{33} & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & d_{nn} \end{bmatrix} \quad (\text{d}_{ii} \text{ can be } 0!)$$

▷ Gaussian Elimination

Step 1: Forward Elimination → triangular matrix

Perform elementary row operations and try to get an upper triangular matrix.

Step 2: Backward Substitution → diagonal matrix

Perform elementary row operations and try to get a diagonal matrix.

Assumption 1: At each iteration of the forward elimination, the pivot is nonzero. (may not hold for some problems)

↪ Claim 1: Under A1, we can get a diagonal matrix at the end of step 2.

Corollary 1: Under A1, the system has a unique solution.

• Allowable Operations on Rows

$$\text{Given } A = \begin{bmatrix} a^{(1)} \\ a^{(2)} \\ \vdots \\ a^{(n)} \end{bmatrix}$$

① [Multiplication] Multiply a row by a non-zero scalar.

$$a^{(i)} \rightarrow \lambda a^{(i)}$$

② [Addition] Add to one row a scalar multiple of another.

$$a^{(i)} \rightarrow a^{(i)} + \lambda a^{(j)}$$

③ [Interchange] Swap the positions of two rows.

$$\begin{bmatrix} a^{(i)} \\ a^{(j)} \\ a^{(k)} \\ \vdots \\ a^{(n)} \end{bmatrix} \rightarrow \begin{bmatrix} a^{(j)} \\ a^{(i)} \\ a^{(k)} \\ \vdots \\ a^{(n)} \end{bmatrix} \quad (\text{swap } i\text{-th row \& } j\text{-th row}).$$

V. Matrix Multiplication

• Definition

Let $A \in \mathbb{R}^{m \times n}$ and $B \in \mathbb{R}^{n \times k}$. The product $C = AB$ is an $m \times k$ matrix where :

$$C_{ij} = \sum_{k=1}^n a_{ik} b_{kj}$$

AB exists iff A 's column # = B 's row # Dimension Match.

Q: How many multiplications?

A: $m \times n \times k$.

$$\xrightarrow{\text{Zero matrix}} \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \end{bmatrix}$$

Let A, B, C be 3 matrices, $I \in \mathbb{R}^{n \times n}$ is an identity matrix, $\alpha \in \mathbb{R}$, we have :

- 0. $AI = A$. $A0 = 0$
- 1. $A(B+C) = AB + AC$, $A \in \mathbb{R}^{m \times n}$, $B, C \in \mathbb{R}^{n \times k}$
- 2. $(B+C)A = BA + CA$, $B, C \in \mathbb{R}^{m \times n}$, $A \in \mathbb{R}^{n \times k}$
- 3. $\alpha(AB) = (\alpha A)B = A(\alpha B)$. $A \in \mathbb{R}^{m \times n}$, $B \in \mathbb{R}^{n \times k}$.
- 4. $(AB)C = A(BC)$, $A \in \mathbb{R}^{m \times n}$, $B \in \mathbb{R}^{n \times k}$, $C \in \mathbb{R}^{k \times p}$.

Left Distributive

Right Distributive

Scalar Associate

Associative

Warning: $AB \neq BA$?

1. AB exists doesn't imply that BA exists.

2. Even if both AB & BA exists, they are generally not equal. ($AB \neq BA$).

Transpose

Def. let $A = (a_{ij})_{m \times n}$, then the transpose of A is the matrix $B = (b_{ji})_{n \times m}$. where
 $b_{ji} = a_{ij}$ ($i = 1, 2, \dots, m$, $j = 1, 2, \dots, n$).

Notation: $\underline{B = A^T}$

e.g. $A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$, $A^T = \begin{bmatrix} 1 & 3 \\ 2 & 4 \end{bmatrix}$
 $A = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$, $A^T = [1, 2, 3]$

$$A = \begin{bmatrix} 1 & 0 & 3 \\ 5 & -1 & 0 \end{bmatrix}, A^T = \begin{bmatrix} 1 & 5 \\ 0 & -1 \\ 3 & 0 \end{bmatrix}$$

Properties:

Let $A, B \in R^{m \times n}$, $\alpha \in R$, then

- 1. $(A+B)^T = A^T + B^T$
- 2. $(\alpha A)^T = \alpha A^T$
- 3. $(A^T)^T = A$
- 4. $(AB)^T = B^T A^T \neq A^T B^T$

Symmetric Matrix

Def. If a matrix $A \in R^{n \times n}$ satisfies $A = A^T$, we call A is symmetric.

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

Vector Outer Product

Def. let $\vec{u} \in R^{m \times 1}$ and $\vec{v} \in R^{n \times 1}$ are two column vectors. The outer product of \vec{u} and \vec{v} is:

$$\vec{u} \vec{v}^T = \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_m \end{bmatrix} [v_1, v_2, \dots, v_n] = \begin{bmatrix} u_1 v_1 & u_1 v_2 & \dots & u_1 v_n \\ u_2 v_1 & u_2 v_2 & \dots & u_2 v_n \\ \vdots & \ddots & \ddots & \vdots \\ u_m v_1 & u_m v_2 & \dots & u_m v_n \end{bmatrix}$$

OPS: The outer product of \vec{u} and \vec{v} is $\vec{u} \vec{v}^T \Rightarrow$ Matrix

The inner product of \vec{u} and \vec{v} is $\vec{u}^T \vec{v} \Rightarrow$ Scalar

Block Partition

Def. The matrix $A = \begin{bmatrix} A_{11} & \cdots & A_{1t} \\ \vdots & \ddots & \vdots \\ A_{s1} & \cdots & A_{st} \end{bmatrix}$ is a partition of matrix with $s \times t$ blocks if A_{ij} :

① For each fixed i , the number of rows of all A_{ij} are equal.

② For each fixed j , the number of columns of all A_{ij} are equal.

The matrix A_{ij} is called the (i, j) -block of A .

We say A has s block-rows and t block-columns.

Multiplication of Partitioned Matrices

Rule 1: Block structure match.

The number of block-columns in A must equal the number of block-rows in B .

Then can treat each block as an "entry", for standard matrix multiplication.

Rule 2: Block size match

When multiplying two blocks, the dimensions of the blocks must match.

VI. Breakdown Cases & REF, RREF

Zeros in GJ-E

Type 1: No solution

$$\text{e.g. } \begin{cases} x_1 + x_2 = 2 \\ 2x_1 + 2x_2 = 5 \end{cases} \Rightarrow \left[\begin{array}{cc|c} 1 & 1 & 2 \\ 2 & 2 & 5 \end{array} \right] \Rightarrow \left[\begin{array}{cc|c} 1 & 1 & 2 \\ 0 & 0 & 1 \end{array} \right] ?$$

Type 2: Infinite solutions

$$\text{e.g. } \begin{cases} x_1 + x_2 = 2 \\ 2x_1 + 2x_2 = 4 \end{cases} \Rightarrow \left[\begin{array}{cc|c} 1 & 1 & 2 \\ 2 & 2 & 4 \end{array} \right] \Rightarrow \left[\begin{array}{cc|c} 1 & 1 & 2 \\ 0 & 0 & 0 \end{array} \right] ?$$

Type 3: 1 solution (0 in diagonal)

$$\text{e.g. } \begin{cases} x_2 = 2 \\ 2x_1 + 2x_2 = 4 \end{cases} \Rightarrow \left[\begin{array}{cc|c} 0 & 1 & 2 \\ 2 & 2 & 4 \end{array} \right] \Rightarrow \left[\begin{array}{cc|c} 1 & 0 & 0 \\ 0 & 1 & 2 \end{array} \right] \Rightarrow \begin{cases} x_1 = 0 \\ x_2 = 2 \end{cases}.$$

Solution Test

Def. A solution of an $m \times n$ linear system with variables (x_1, x_2, \dots, x_n) is a vector such that if we let $x_i = s_i$ for all $i = 1, 2, \dots, n$, the m equations hold simultaneously. The solution set of an $m \times n$ linear system is a set that contains all solutions.

$$\text{e.g. } \begin{cases} 2x_1 + 3x_2 = 3 \\ x_1 - x_2 = 4 \end{cases} \Rightarrow \text{The solution set is } \{(3, -1)\}.$$

$$\begin{cases} 2x_1 + 3x_2 = 3 \\ 4x_1 + 6x_2 = 6 \end{cases} \Rightarrow \text{The solution set is } \left\{ \left(t, \frac{3-2t}{3} \right) \mid t \in \mathbb{R} \right\}.$$

REF & RREF

Def. A matrix is in row echelon form (REF) if

1. Leading 1s: The first entry of each non-zero row is 1. (called "pivot").

2. Leading 1s Moving Right: If a row has a leading 1, the leading 1 in the next row appears in a later column.

3. Zero Rows at the Bottom: If there is a row whose entries are all zero, then all rows below it have all zero entries.

And is reduced row echelon form (RREF) if

4. Zeros in the columns of the leading 1s: Each leading 1 is the only nonzero entry in its column (the corresponding column is called pivot column)

$$\text{e.g. REF: } \left[\begin{array}{ccccc} 1 & 2 & 2 & 3 & 1 \\ 0 & 1 & 5 & 4 & 5 \\ 0 & 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array} \right]$$

$$\text{RREF: } \left[\begin{array}{ccccc} 1 & 0 & 2 & 0 & 1 \\ 0 & 1 & 5 & 0 & 5 \\ 0 & 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array} \right]$$

Property of RREF:

If a matrix A is in RREF, then exchanging columns of A can lead to the form

$$\left[\begin{array}{cc} I_k & F \\ 0 & 0 \end{array} \right]$$

where I_k is the $k \times k$ identity matrix.

Number of Solutions for Square Systems

Problem: Solve $n \times n$ linear system of equations.

Preparation: Write the linear system as augmented matrix form.

Step 1:

Perform elementary row operations to get an upper triangular matrix.

Step 2:

Case 1: All diagonal entries are nonzero (i.e. n pivots). Unique solution.

Then perform "row multiplication and addition" to get an identity matrix.

Case 2: Some diagonal entries are zero (i.e. less than n pivots):

At the end of elimination, if:

① exists $0 = *$ (nonzero number). No solution.

② besides $x_{it} = *$, all are $0 = 0$. ∞ many solutions

VII. Matrix Inverse

Matrix Inverse

Def. Suppose $A \in \mathbb{R}^{n \times n}$. If a matrix $B \in \mathbb{R}^{n \times n}$ satisfies $\underline{AB = BA = I_n}$, then we say B is the inverse of A . nxn unit matrix.
Denoted as $\underline{B = A^{-1}}$.

Suppose $A \in \mathbb{R}^{n \times n}$. If there exists B such that $B = A^{-1}$, then we say A is invertible.

If A is invertible, then we say A is nonsingular.

If A is not invertible, then we say A is singular.

e.g. let $A = \begin{bmatrix} 3 & 0 \\ 0 & 4 \end{bmatrix}$, $B = \begin{bmatrix} \frac{1}{3} & 0 \\ 0 & \frac{1}{4} \end{bmatrix}$.

$$\text{then } AB = \begin{bmatrix} 3 & 0 \\ 0 & 4 \end{bmatrix} \begin{bmatrix} \frac{1}{3} & 0 \\ 0 & \frac{1}{4} \end{bmatrix} = \begin{bmatrix} \frac{1}{3} & 0 \\ 0 & \frac{1}{4} \end{bmatrix} \begin{bmatrix} 3 & 0 \\ 0 & 4 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

$$\text{Thus, } A^{-1} = B = \begin{bmatrix} \frac{1}{3} & 0 \\ 0 & \frac{1}{4} \end{bmatrix}.$$

Theorem: Matrix inverse is unique. Uniqueness

Suppose the square matrix A has an inverse. Then A^{-1} is unique.

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

Matrix Inverse of Special Matrices

1. Diagonal Matrix

A diagonal matrix $\begin{bmatrix} a_{11} & 0 & \cdots & 0 \\ 0 & a_{22} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_{nn} \end{bmatrix}$ is invertible if $a_{ii} \neq 0$ ($i=1, 2, \dots, n$).

and the inverse is $\begin{bmatrix} \frac{1}{a_{11}} & 0 & \cdots & 0 \\ 0 & \frac{1}{a_{22}} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \frac{1}{a_{nn}} \end{bmatrix}$

2. Triangular Matrix

- ① A triangular matrix is invertible iff the diagonal entries are invertible.
- ② The inverse of an upper triangular matrix (if exists) is upper triangular matrix.
lower

Inverse of a 2×2 Matrix

Let $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$. we have $A^{-1} = \frac{1}{\det(A)} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$, where $\det(A) = \begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc$.

And A^{-1} exists iff $\det(A) \neq 0$.

Inverse of Products

Theorem: Matrix Inverse of Matrices Product

Suppose A and B are invertible matrices of size n , then AB is invertible and

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

e.g. A_1, A_2, \dots, A_k are real invertible matrices.

Then $(A_1 A_2 \dots A_k)^{-1} = A_k^{-1} A_{k-1}^{-1} \dots A_1^{-1}$.

VIII. Elementary Matrix

Elementary Matrix

Def. The matrices corresponding to a single elementary row operation are called elementary matrices.

- For a given matrix A , performing elementary row operation for A is equivalent to premultiplying A by the corresponding elementary matrix.

Inverse of Elementary Matrix

Proposition:

1. $E_{RiRj}^{-1} = E_{RjRi}$, Corresponding to the reverse row operation 1: $R_i \leftrightarrow R_j$.
2. $E_{\alpha R_i}^{-1} = E_{\frac{1}{\alpha}R_i}$ ($\alpha \neq 0$), corresponding to the reverse row operation 2: $R_i \rightarrow \frac{1}{\alpha}R_i$.
3. $E_{\beta R_i + R_j}^{-1} = E_{-\beta R_j + R_i}$. Corresponding to the reverse row operation 3: $R_j \rightarrow -\beta R_i + R_j$.

Remark: The inverse of the elementary matrices corresponding to the reverse row operations and belong to the same type of elementary matrices.

IX. LU Decomposition

n by n Good Case

Given a coefficient matrix A (square matrix).

We have Gaussian elimination:

$$A \rightarrow A_1 \rightarrow A_2 \dots \rightarrow U$$

$$\Rightarrow E_k \cdot E_{k-1} \cdots E_1 \cdot A = U$$

$$\Rightarrow A = (E_k \cdot E_{k-1} \cdots E_1)^{-1} \cdot U \quad \Rightarrow \quad A = LU$$

$L = [I | E_i]^{-1}$

Algorithm of Computation

Step 1: Forward elimination.

Run forward elimination, till get upper triangular matrix U .

IF row exchange needed:

STOP and report: the algorithm fails to generate LU decomposition.

ELSE (no row exchange) Go to Step 2.

Step 2: Record elementary matrices

Record E_1, E_2, \dots, E_k in Step 1.

Step 3: Compute $L = E_1^{-1} E_2^{-1} \dots E_k^{-1}$.

Step 4: Conclusion

A has an LU decomposition $A = LU$.

PLU Decomposition

Theorem: Any square matrix A can be written as

$$PA = LU$$

where L is lower triangular, U is upper triangular, P is a certain permutation matrix.

Permutation Matrix

Def. Let e_i be the i^{th} row of the identity matrix I .

$$\text{e.g. if } P = \begin{bmatrix} 1 & & \\ 3 & 1 & \\ 2 & & 1 \end{bmatrix}, \text{ then } Pe_1 = \begin{bmatrix} e_1 \\ e_3 \\ e_2 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$$

A permutation matrix is a square matrix that has exactly one entry of 1 in each row and each column and 0s elsewhere.

The product of row exchange matrices

$$P = E_{i_1, j_1} E_{i_2, j_2} \dots E_{i_k, j_k}$$

P is a permutation matrix

X. Matrix Inverse Continue

When is A invertible?

Ans: A is invertible iff A has n pivots (Assuming A is n by n matrix).

Left Inverse and Right Inverse

AB

Def. Let A and B be matrices. If $\underline{BA} = I$, where I is the identity matrix, then B is called the left inverse of A .

right

not necessary

Lemma: If $MA = I_n$, (and M B invertible), then $A^{-1} = M$.

How to express A^{-1} ?

We have $\underline{C_E} - \bar{E}$ (case of n pivots). $A \rightarrow A_1 \rightarrow A_2 \rightarrow \dots \rightarrow U \rightarrow \dots B_1 \rightarrow \dots \rightarrow I_n$.

Noted $U \rightarrow \bar{E}_p \rightarrow \bar{E}_{p_1} \rightarrow \dots \rightarrow \bar{E}_1 \rightarrow I_n$. And we can get $A^{-1} = \underline{\bar{E}_p \bar{E}_{p_1} \dots \bar{E}_1}$.

Claim: A matrix is invertible if it can be written as the product of elementary matrices.

• How to compute A^{-1} ?

Algorithm 1:

Step 1: Forward elimination (with column exchange)

Run forward elimination (with column exchange), till get upper triangular matrix U .

If U contains zero diagonal entry:

STOP and report: No inverse.

ELSE Go to Step 2.

Step 2: Run backward substitution, till get identity matrix I_n .

Step 3: Record elementary matrices E_1, E_2, \dots, E_k in Step 1.

Record elementary matrices $\bar{E}_{k+1}, \bar{E}_{k+2}, \dots, \bar{E}_p$ in Step 2.

Step 4: Compute inverse

$$A^{-1} = E_p \cdots \bar{E}_2 \bar{E}_1.$$

Algorithm 2:

Apply GE to $[A | I_n]$

$$[A | I_n] \xrightarrow{\text{op}} D \xrightarrow{\text{op}} \cdots \xrightarrow{\text{op}} [I_n | A^{-1}]$$

XI. Linear Space

Linear Space

Def. Suppose V is a set associated with two operations:

of (i) Addition "+": $u + v \in V$, $\forall u \in V, v \in V$.

(ii) Scalar multiplication: $\alpha u \in V$, $\forall \alpha \in R, u \in V$.

V is called a linear space over R if the 8 axioms hold ($u, v, w \in V, \alpha, \beta \in R$):

1. $u + v = v + u$

2. $u + (v + w) = (u + v) + w = u + v + w$

3. There exists an element 0 s.t. $u + 0 = u$.

4. $-u = (-1) \cdot u$, s.t. $u + (-u) = 0$

5. $\alpha(u + v) = \alpha u + \alpha v$

6. $(\alpha + \beta)u = \alpha u + \beta u$

7. $\alpha(\beta u) = (\alpha \beta)u$

8. $1u = u$.

Euclidean Space

Euclidean space is a linear space, when equipped with addition and scalar-vector product.

e.g. R^n is called n-dimensional Euclidean space.

Polynomial Space

Set of polynomials with degree no more than k is a linear space.

• Subspace

Def. Suppose V is a linear space. We say W is a subspace of V if two conditions hold:

1. W is a subset of V
2. W is a linear space

In words: A subspace of V is a subset that is itself a linear space.

Property: A subspace W must contain the zero element: $0 \in W$

XII. Linear Independence

• Linear Dependence and Independence

Def. Suppose V is a linear space over \mathbb{R} . $u_1, u_2, \dots, u_k \in V$.

We say u_1, u_2, \dots, u_k are

1. linearly dependent if

$\exists c_1, c_2, \dots, c_k \in \mathbb{R}$ s.t. $(c_1, c_2, \dots, c_k) \neq (0, 0, \dots, 0)$ and $c_1 u_1 + c_2 u_2 + \dots + c_k u_k = 0$.

2. linearly independent if

$\nexists c_1, c_2, \dots, c_k \in \mathbb{R}$ s.t. $(c_1, c_2, \dots, c_k) \neq (0, 0, \dots, 0)$ and $c_1 u_1 + c_2 u_2 + \dots + c_k u_k \neq 0$.

In short, $c_1 u_1 + c_2 u_2 + \dots + c_k u_k = 0$ only happens when $c_1 = c_2 = \dots = c_k = 0$

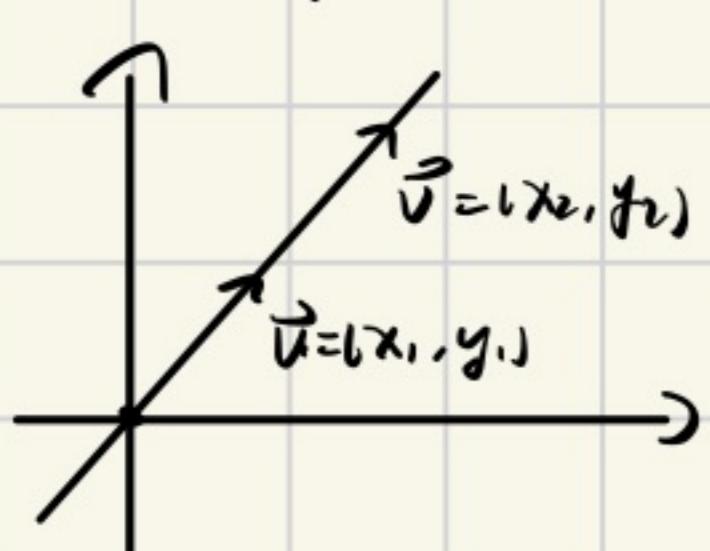
• Checking Linear Independence

If the space is \mathbb{R}^n , then just need to solve a linear system!

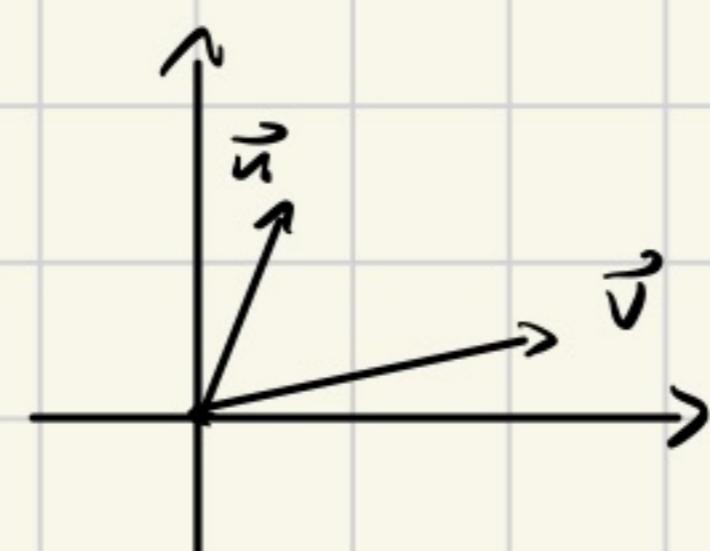
If the space is not \mathbb{R}^n , need extra tools. (not discuss here).

• Geometry

In the plane:

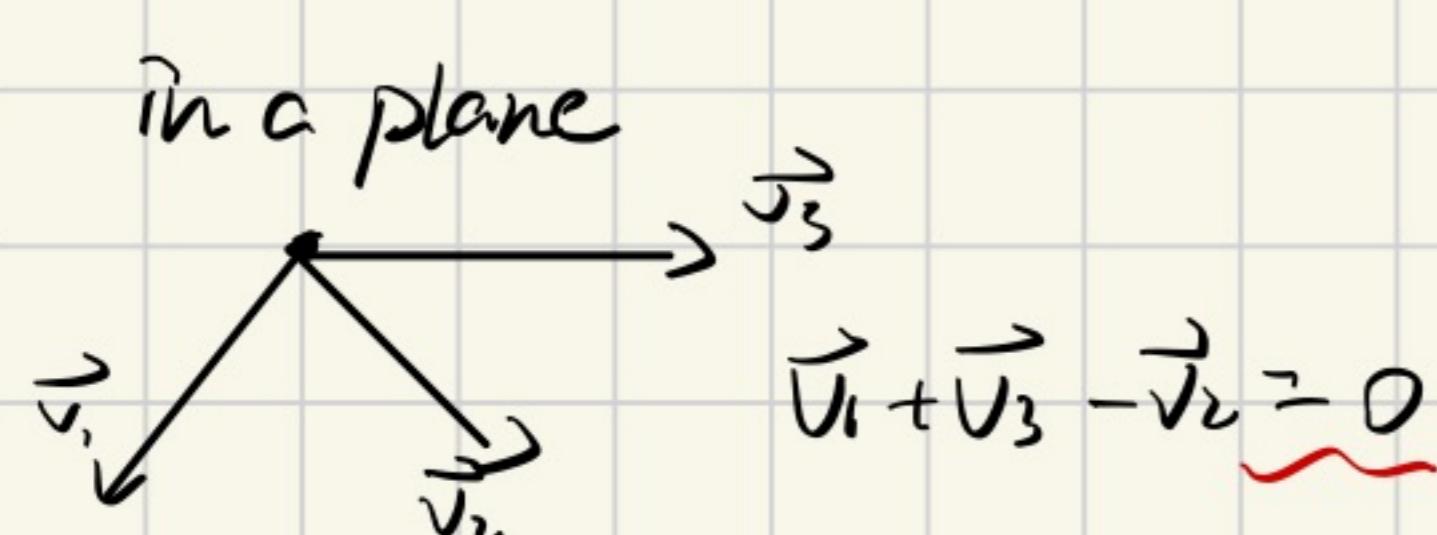


\vec{u} & \vec{v} linearly dependent

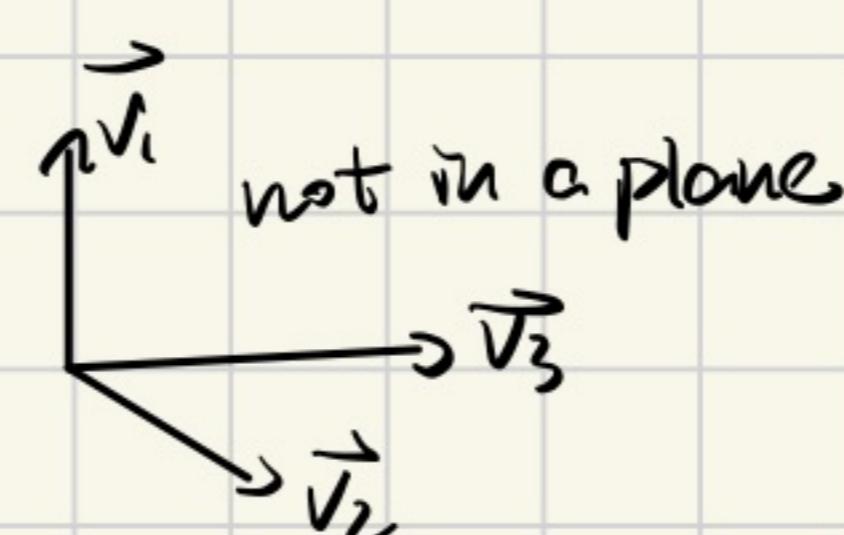


\vec{u} & \vec{v} linearly independent

In a 3D space:



linear dependent vectors



linear independent vectors

XIII. Null Space

• Homogeneous Linear System

Def. A homogeneous linear system is

$$Ax = 0$$

where $A \in \mathbb{R}^{m \times n}$ are given and $x \in \mathbb{R}^n$ is the variable.

Theorem:

The solution set of a homogeneous linear system $Ax=0$ is a linear space (also a subspace of \mathbb{R}^n).

Def. The solution set of $Ax=0$ is called the null space of A , denoted as $N(A)$.

XIV. Span

Two Ways to Generate Subspaces

1. Solution set of linear equation + Taking intersection
 \Rightarrow null space

2. Linear combination \Rightarrow span

Span

Def. Suppose V is a linear space, $u = \{u_1, u_2, \dots, u_k\}$ is a subset of V .

The span of u is defined as

$$\Delta \text{span}(u) \triangleq \{a_1u_1 + a_2u_2 + \dots + a_ku_k \mid a_1, a_2, \dots, a_k \in \mathbb{R}\}.$$

In words: the span of elements of a linear space is the set of all linear combinations.

Fact: The span of any finite subset of V is a subspace of V .

e.g. RGB as a spanning set.

The set $\{R, G, B\}$ spans the entire color space.

Span of Unit Vectors

Def. Consider two points $e_1 = (0, 1)$, $e_2 = (1, 0)$ on the plane.

$$\text{span}\{(1, 0), (0, 1)\} = \mathbb{R}^2.$$

Consider n vectors e_1, e_2, \dots, e_n are linearly independent.

$$\text{Then } \text{span}\{e_1, e_2, \dots, e_n\} = \mathbb{R}^n.$$

Spanning Set

Def. Suppose V is a linear space, $u = \{u_1, u_2, \dots, u_k\}$ is a subset of V .

If $\text{span}(u) = V$, then we say u is a spanning set of V , or u spans V .

e.g. $\{e_1, e_2, \dots, e_n\}$ is a spanning set of \mathbb{R}^n .

XV. Solving $Ax=b$

Column Space

Def. Suppose $A = [a_1, a_2, \dots, a_n] \in \mathbb{R}^{m \times n}$ is a matrix.

The span ($\{a_1, a_2, \dots, a_n\}$) is called the column space of A , denoted as $C(A)$.

In words: A 's column space is the span of A 's column vectors.

Δ Solvability: $Ax=b$ has a solution iff $b \in C(A)$.

Solution Set of $Ax = b$

Def. Suppose V is a linear space. For any element $v \in V$, and a subspace U , define $v+U \triangleq \{v+u \mid u \in U\}$.

Proposition:

The solution set of a linear system $\underline{Ax=b}$ is:

- o either (i) an empty set
- o or (ii) $x_p + N(A)$, where x_p is any solution of $Ax=b$.

XVI. Basis

Redundancy of Set

Theorem: Dependent Sets is Redundant

strictly included

Suppose $M = \{u_1, u_2, \dots, u_r\}$ is linear dependent, then $\exists H \subsetneq M$ s.t. $\text{span}(M) = \text{span}(H)$.

In words, it can be reduced without changing its span.

By contrast, if M is independent, then there's no set $H \subsetneq M$ s.t. $\text{span}(M) = \text{span}(H)$.

In words, removing any elements changes the span.

Basis

Def. Suppose V is a linear space over R , $u \triangleq \{u_1, u_2, \dots, u_k\} \subseteq V$.

We say u is a basis if:

- o 1. u is linearly independent.
- o 2. $\text{span}(u) = V$.

In words, a basis of V is linearly independent & spanning set of V .

A basis of V is a minimal spanning set of V in the sense that we can't further shrink the set while being a spanning set of V .

Dimension

Def. Dimension = Size of Bases.

Theorem: Bases have same size

If $\{u_1, u_2, \dots, u_m\}$ and $\{v_1, v_2, \dots, v_n\}$ are bases of a linear space V , then $m=n$.

Def. Suppose V is a linear space

If V has a basis u with n elements, then we say the dimension of V is n .

Denoted as $\dim(V)=n$. or V is n -dimensional.

Quick way to check basis.

Proposition:

Consider a linear space V with $\dim(V)=n$.

n vector in R^n { Independence
Span } \Rightarrow Basis

Invertibility Conditions

Theorem: Equivalent Conditions for Invertibility

Let $A \in R^{n \times n}$. The following statements are equivalent.

1. A is invertible.
2. The linear system $Ax=0$ has a unique solution $x=0$.
3. A is a product of elementary matrices.
4. A has n pivots.
5. The columns of A are linearly independent.
6. The columns of A span R^n .
7. The columns of A form a basis.
8. $\dim(C(A)) = n$, $\text{rank}(A) = n$.
9. $Ax = b$ is solvable for any b .

XVII. Rank

Row Space

Def. Consider a matrix $A = \begin{bmatrix} A_{11}^T \\ A_{12}^T \\ \vdots \\ A_{m1}^T \end{bmatrix} \in R^{m \times n}$

$\text{span}(\{A_{11}, A_{12}, \dots, A_{m1}\}) \subseteq R^n$ is called the row space of A , denoted as $\text{Row}(A)$.

A 's row space is the span of A 's row vectors: $\text{Row}(A) = C(A^T)$.

Rank

Def. Row-rank of A is defined as $\dim(\text{Row}(A))$, denoted as $r_R(A)$.

Column-rank of A is defined as $\dim(C(A))$, denoted as $r_C(A)$.

Theorem: Row-rank, column-rank and rank of a matrix are the same.

$$\Delta r_R(A) = r_C(A) = \text{rank}(A).$$

$$\Leftrightarrow \dim(\text{Row}(A)) = \dim(C(A)).$$

Def. The rank of a matrix A is $r_R(A)$, which is also $r_C(A)$, denoted as $\text{rank}(A)$.

$$\text{For } A \in R^{n \times n}, \text{rank}(A) \leq \min\{m, n\}.$$

Rank and Pivots

Theorem: Suppose M is a RREF.

$$\text{Then } r_R(M) = r_C(M) = \text{Number of pivots} = \text{rank}(A).$$

Full Rank Matrix

Def. Let $A \in R^{m \times n}$.

If $\text{rank}(A) = m$, then we say A has full row rank.

If $\text{rank}(A) = n$, then we say A has full column rank.

If $\text{rank}(A) = \min\{m, n\}$, then we say A has full rank.

Rank and Solutions

Def. Consider a matrix $A \in \mathbb{R}^{m \times n}$.

1. If $\text{rank}(A) = n$ (full column rank), then for any b , $Ax = b$ has at most one solution.

(# of solutions must be 0 or 1. Eliminate possibility of ∞).

This means: No free columns in RREF \Rightarrow No free variables

2. If $\text{rank}(A) = m$ (full row rank), then $Ax = b$ has at least one solution.

(# of solutions must be 1 or ∞ , eliminate possibility of 0).

This means: No "zero rows" in RREF.

XVII. Orthogonality

Orthogonality

Def. Two vectors $u, v \in \mathbb{R}^n$ are orthogonal if $u^T v = 0$. Denote $u \perp v$.

$$\text{And } \cos \theta = \frac{u^T v}{\|u\| \|v\|} = 0$$

We say a vector $u \in \mathbb{R}^n$ is orthogonal to a subspace $V \subseteq \mathbb{R}^n$ if $u \perp v, \forall v \in V$. Denote $u \perp V$.

We say two subspaces $U, V \subseteq \mathbb{R}^n$ are orthogonal if $u \perp v, \forall u \in U, v \in V$. Denote $U \perp V$.

Orthogonal Complement

Def. For any subspace $V \subseteq \mathbb{R}^n$, the set of vectors that are orthogonal to V ($u \in \mathbb{R}^n, u \perp v, \forall v \in V$)

is called the orthogonal complement of V , denoted as $U = V^\perp$

Remark: A subspace V has a unique orthogonal complement.

Theorem: Dimension of V prep.

Suppose S is a subspace of \mathbb{R}^n . Then the following hold:

i) S^\perp is a subspace

ii) $\dim(S) + \dim(S^\perp) = n$

iii) If $\{u_1, u_2, \dots, u_r\}$ is a basis of S , and $\{u_{r+1}, \dots, u_n\}$ is a basis of S^\perp , then $\{u_1, \dots, u_r, u_{r+1}, \dots, u_n\}$ is a basis of \mathbb{R}^n .

XIX. Four Fundamental Subspaces

Left Null Space

Def. The left null space of a matrix A is defined as $N(A^T)$.

Fundamental Theorem

Theorem: Suppose $A \in \mathbb{R}^{m \times n}$. Then:

1. $N(A) = \text{Row}(A)^\perp$, $\dim(N(A)) = n - \text{rank}(A)$.

2. $N(A^T) = C(A)^\perp$, $\dim(N(A^T)) = m - \text{rank}(A)$.

XX. Least Square

Linear Regression

Minimization Problem

Method: Make error as small as possible

Find x s.t. $\|Ax - b\|$ is the smallest among all.

More precisely, find x^* s.t. $\|Ax^* - b\| \leq \|Ax - b\|, \forall x$.

Least Squares

Def. Given $A \in \mathbb{R}^{m \times n}$, $b \in \mathbb{R}^m$, the least square problem is

$$\min_{x \in \mathbb{R}^n} \|Ax - b\|$$

y is a solution of the problem if

$$\|Ay - b\| \leq \|Ax - b\|, \forall x$$

And y is called the regression vector

The method to find y called linear regression.

Suppose a_1^T, \dots, a_m^T are rows of A .

$$\|Ax - b\|^2 = (a_1^T x - b_1)^2 + \dots + (a_m^T x - b_m)^2$$

Solving LS is Finding Projection

Lemma: Orthogonal Projection

Suppose S is a subspace of \mathbb{R}^m . Suppose $p \in S$, then

$$\|b - p\| \leq \|b - z\|, \forall z \in S \Leftrightarrow b - p \perp S$$

Def. Suppose S is a subspace of \mathbb{R}^m .

Suppose $p \in S$ and $b - p \perp S$, then we say p is the projection of b onto S .

Finding Projection is Solving Equation

Lemma:

Suppose $S = C(A)$ is the column space of matrix A .

Then: $b - Ay \perp C(A) \Leftrightarrow A^T A y = A^T b$.

Theorem: LS Solution and normal equation

Given $A \in \mathbb{R}^{m \times n}$, $b \in \mathbb{R}^m$. The following statements are equivalent:

- of
1. y is a solution of the problem $\min_{x \in \mathbb{R}^n} \|Ax - b\|$.
2. $A^T A y = A^T b$

Analyzing Normal Equation

Proposition (Existence):

The linear system $A^T A y = A^T b$ has at least one solution.

Corollary (Characterization)

Suppose A has linearly independent columns (i.e. has full column rank),

then $A^T A$ is invertible solution of the least square problem $\min_{x \in \mathbb{R}^n} \|Ax - b\|$ is $y = (A^T A)^{-1} A^T b$.

XXI. Orthonormal Basis

• Orthonormal Set of Vectors

Def. Let $\{v_1, v_2, \dots, v_k\}$ be a set of nonzero vectors in \mathbb{R}^n . If $\langle v_i, v_j \rangle = 0$ for any $i \neq j$, then this set is called an orthogonal set.

An orthonormal set $\{v_1, v_2, \dots, v_k\}$ is an orthogonal set of unit-norm vectors.
i.e. $\|v_i\| = 1$ for all $i=1, \dots, n$.

Proposition: If $S = \{v_1, v_2, \dots, v_k\}$ is an orthogonal set, then v_1, \dots, v_k are linearly independent.

• Orthonormal Basis

Def. A set of vectors $S = \{v_1, \dots, v_n\}$ is called an orthonormal basis of a linear space V if

1. S is an orthonormal set.
2. Vector in S form a basis of V .

• Orthogonal Matrix

Def. An orthogonal matrix $Q \in \mathbb{R}^{n \times n}$ is a real square matrix whose columns form an orthonormal basis in \mathbb{R}^n .

A matrix $Q \in \mathbb{R}^{n \times n}$ is an orthogonal matrix iff $Q^T Q = I_n$.

Corollary:

If $Q \in \mathbb{R}^{n \times n}$ is orthogonal then $Q^{-1} = Q^T$.

Properties:

1. $\langle Qx, Qy \rangle = \langle x, y \rangle$
2. $\|Qx\| = \|x\|$.

• Representation Elements Using Basis

Proposition (Unique Representation via Basis)

Let $u = \{u_1, \dots, u_n\}$ be a basis of a linear space V .

Any element $v \in V$ can be uniquely represented as a linear combination of u_1, \dots, u_n .

Representation: $v = a_1 u_1 + \dots + a_n u_n$.

Theorem: Representation via Orthonormal Basis

Let $u = \{u_1, \dots, u_n\}$ be an orthonormal basis of \mathbb{R}^n .

Any vector $v \in \mathbb{R}^n$ can be represented as a linear combination

$$v = \sum_{i=1}^n u_i u_i^T v = \sum_{i=1}^n \langle v, u_i \rangle u_i$$

e.g. $\{u_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, u_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, u_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}\}$ is an orthonormal basis of \mathbb{R}^3 .

For any $v = [x, y, z]^T$, one has

$$v = \langle v, u_1 \rangle u_1 + \langle v, u_2 \rangle u_2 + \langle v, u_3 \rangle u_3.$$

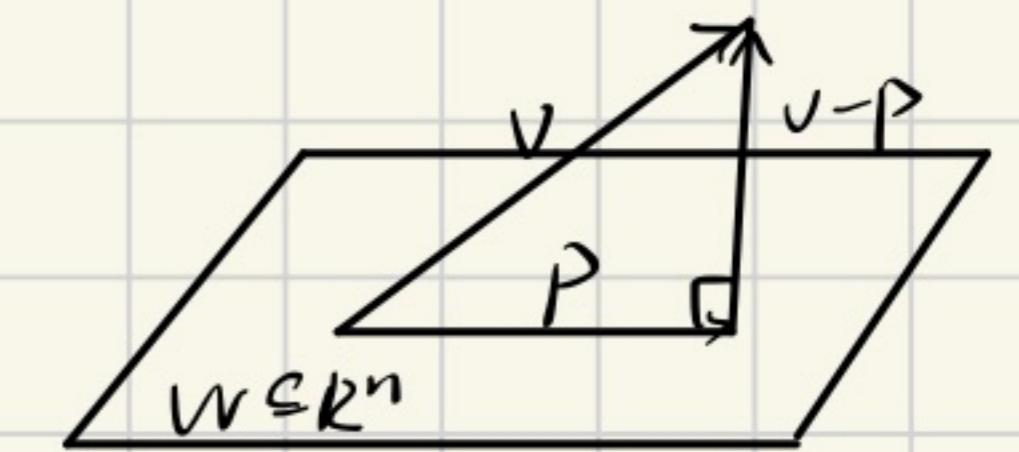
$$= \begin{bmatrix} x \\ 0 \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ y \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ z \end{bmatrix}$$

$$= [x, y, z]^T.$$

Projection

Def. Suppose S is a subspace of \mathbb{R}^n .

Suppose $p \in S$ and $b-p \perp S$, then we say p is the projection of b onto S .



Proposition:

For any subspace $W \subset \mathbb{R}^n$, any vector $v \in \mathbb{R}^n$, there is a unique vector $p \in W$ s.t. $(v-p) \perp W$.

Theorem: Projection when Ortho-Basis Exists

Let W be a subspace of \mathbb{R}^n and suppose $S = \{w_1, \dots, w_k\}$ is an orthonormal basis of W .

The projection of v onto W can be represented as

$$\Delta P = \sum_{i=1}^k \langle w_i, v \rangle w_i.$$

Gram-Schmidt Process

Def. Suppose $U = \{u_1, \dots, u_k\}$ is a set of linearly independent elements.

Input $U = \{u_1, \dots, u_k\}$.

For $i = 1, \dots, k$

If $i=1$:

$$v_1 = \frac{u_1}{\|u_1\|}.$$

Else:

$$\Delta \text{ let } p_i = u_i - \sum_{j=1}^{i-1} \text{proj}_{v_j} u_i = u_i - \sum_{j=1}^{i-1} \langle u_i, v_j \rangle v_j$$

$$v_i = \frac{p_i}{\|p_i\|}$$

Return $V = \{v_1, \dots, v_k\}$.

Proposition:

The set $V = \{v_1, \dots, v_k\}$ returned by the Gram-Schmidt process is an orthonormal basis of $\text{span}(U)$.

XXII. Determinant

Geometry of Determinant

Def. Consider a matrix $A \in \mathbb{R}^{n \times n}$, then

$\det(A) = \text{Volume of the polytope formed by } a_1, a_2, \dots, a_n \in \mathbb{R}^n$. ($A = [a_1, a_2, \dots, a_n]$).

And we have

$\Delta \det(A) \neq 0$ iff A is invertible

Notation: $|A| = \det(A)$.

Decomposition of $\det(A)$

Properties:

$$\textcircled{1} \cdot \begin{vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{vmatrix} = \begin{vmatrix} a_1 & 0 & 0 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{vmatrix} + \begin{vmatrix} 0 & a_2 & 0 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{vmatrix} + \begin{vmatrix} 0 & 0 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{vmatrix}$$

$$\textcircled{2} \cdot \begin{vmatrix} a_1 & 0 & 0 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{vmatrix} = \begin{vmatrix} a_1 & 0 & 0 \\ 0 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{vmatrix} + \begin{vmatrix} 0 & 0 & 0 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{vmatrix}$$

$$\textcircled{3} \cdot \begin{vmatrix} a_1 & 0 & 0 \\ 0 & b_2 & b_3 \\ 0 & c_2 & c_3 \end{vmatrix} = a_1 \begin{vmatrix} b_2 & b_3 \\ c_2 & c_3 \end{vmatrix}$$

Properties continue:

④ Swapping two rows (or two columns) changes the sign of det.

$$\text{e.g. } \det([v_1, v_2, v_3]) = -\det([v_2, v_1, v_3]).$$

Formal Introduction

Def. (Minor)

Let $A \in R^{n \times n}$ be a real square matrix

Denote by $M_{ij} \in R^{(n-1) \times (n-1)}$ a matrix formed by deleting the i^{th} row and j^{th} column of A , called the minor of a_{ij} .

Def. (Determinant)

For a scalar $\alpha \in R$, define $\det(\alpha) = \alpha$.

For any $A \in R^{n \times n}$ with $n \geq 2$ defined $|A| = \det(A)$ as

$$\star \det(A) = \sum_{j=1}^n (-1)^{i+j} \det(M_{ij}) a_{ij} = a_{11} \det(M_{11}) - a_{12} \det(M_{12}) + a_{13} \det(M_{13}) - \dots$$

which is a recursive definition. $\rightarrow (a_{11}, a_{12}, \dots, a_{1n})$: the first row of A .

$$\text{e.g. } \begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc$$

$$\begin{vmatrix} a & b & c \\ d & e & f \\ g & h & i \end{vmatrix} = a \begin{vmatrix} e & f \\ h & i \end{vmatrix} - b \begin{vmatrix} d & f \\ g & i \end{vmatrix} + c \begin{vmatrix} d & e \\ g & h \end{vmatrix}$$

Laplace Expansion

Def. Suppose a matrix $A \in R^{n \times n}$. For any fixed row k ($1 \leq k \leq n$) or any fixed column j ($1 \leq j \leq n$), the determinant $\det(A)$ can be expressed by "expanding" along that row or column as follows:

$$\textcircled{1} \text{ Row } k \text{ expansion: } \det(A) = \sum_{j=1}^n (-1)^{k+j} \det(M_{kj}) a_{kj}.$$

$$\textcircled{2} \text{ Column } j \text{ expansion: } \det(A) = \sum_{i=1}^n (-1)^{i+j} \det(M_{ij}) a_{ij}.$$

Elementary Operations

$$\textcircled{1} \det(E_{R_i \leftrightarrow R_j} A) = -\det(A) \quad \text{Interchange}$$

$$\textcircled{2} \det(E_{R_i + \beta R_j} A) = \det(A) \quad \text{Addition}$$

$$\textcircled{3} \det(E_{\alpha R_i} A) = \alpha \det(A) \quad \text{Scaling}$$

More Properties

① If $A \in R^{n \times n}$ is a triangular matrix with diagonal entries a_{11}, \dots, a_{nn} , $\det(A) = a_{11} \dots a_{nn}$.

② $\forall A, B \in R^{n \times n}$, $\det(AB) = \det(A) \cdot \det(B)$. $\det(A+B) \neq \det(A) + \det(B)$!

③ $\det(A^T) = \det(A)$.

④ $\det(A^{-1}) = \frac{1}{\det(A)}$.

XXII. Linear Transformation

Mapping

Def. Let V and W be two sets. If for any $v \in V$, there is a unique $w = T(v) \in W$, then we say T is a mapping from V to W .

V is called the domain of the mapping.

W is called the codomain of the mapping.

We denote: $T: V \rightarrow W$.

$$v \mapsto T(v).$$

Linear Transformation

Def. If a mapping f from a linear space V to a linear space W satisfies

$$\triangle f(\alpha x + \beta y) = \alpha f(x) + \beta f(y), \forall \alpha, \beta \in \mathbb{R}, \forall x, y \in V.$$

Then we say f is a linear transformation from V to W .

In words, transformation LC of vectors = LC of transformed vectors

Properties:

If f is a linear transformation from V to W , then

- \triangle ① $f(0) = 0$
② $f(\alpha x) = \alpha f(x), \forall \alpha \in \mathbb{R}$.
③ $f(x+y) = f(x) + f(y), \forall x, y \in V$.

e.g. Suppose a matrix $A \in \mathbb{R}^{m \times n}$.

$f: x \mapsto Ax$ is a linear transformation from \mathbb{R}^n to \mathbb{R}^m .

e.g. Suppose two vectors $a \in \mathbb{R}^n, x \in \mathbb{R}^n$.

$f(x) = \langle x, a \rangle$ is a linear transformation from \mathbb{R}^n to \mathbb{R} .

We call it a linear function.

Matrix Representation of Linear Transformation

Def. Consider a linear transformation $f: V \rightarrow W$.

1. Pick a basis $\{v_1, \dots, v_n\}$ of input space V ;

Pick a basis $\{w_1, \dots, w_m\}$ of output space W .

2. Express $f(v_j) = a_{1j}w_1 + \dots + a_{mj}w_m \in W, \forall j$.

Write the coordinates of $T(v_j)$ as $a_j = (a_{1j}, \dots, a_{mj})^T$.

3. Form a matrix $A = [a_1, \dots, a_n] \in \mathbb{R}^{m \times n}$.

Then A is a matrix representation of f .

XXIV. Eigenvalues and Eigenvectors

Eigenvalues and Eigenvectors

Def. Let $A \in C^{n \times n}$ be a square matrix.

If there exists a scalar $\lambda \in R$ or C and a nonzero vector x s.t.

$$Ax = \lambda x$$

Then λ is called a (real or complex) eigenvalue and x is called an eigenvector corresponding (or associated with) λ .

Characteristic Polynomial

Def. let $P_A(\lambda) = \det(\lambda I_n - A)$ is called the characteristic polynomial of A .

and $P_A(\lambda) = 0$ is called the characteristic equation

Q: How many eigenvalues does a matrix A have?

A: The $P_A(\lambda)$, as a polynomial of degree n , has n complex roots, i.e., eigenvalues.

But they may not be distinct!

★ General Procedure to Find Eigenvalues/Eigenvectors

Step 1: Solve $\det(A - \lambda I) = 0$

$\Rightarrow n$ roots $\lambda_1, \dots, \lambda_n$.

Step 2: For each λ_i , find the eigenspace $N(\lambda_i I - A)$.

$N(\lambda_i I - A) = \{ \text{eigenvectors of } A \text{ with } \lambda_i \text{ corresponding to } \lambda_i \} \cup \{0\}$.

Matrix Form of Eigen-equations

Def. Suppose $Av_1 = \lambda_1 v_1, \dots, Av_n = \lambda_n v_n$, where $v_j \in C^n, \lambda_j \in C$.

Then $AV = VD$

where $D = \text{diag}(\lambda_1, \dots, \lambda_n)$, $V = [v_1, \dots, v_n]$.

Eigenbasis

Def. A set of eigenvectors of an $n \times n$ that can form a basis of C^n . Complex!

Spectral Theorem

Theorem:

Any real symmetric matrix A can be written as

$A = VDV^T$ Eigenvalue decomposition

where D is a real diagonal matrix, V is a real orthogonal matrix.

And it implies A has an orthonormal basis.

$$\begin{aligned} A &= VDV^T \\ &= (v_1, \dots, v_n) \cdot \begin{bmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{bmatrix} \cdot \begin{bmatrix} v_1^T \\ \vdots \\ v_n^T \end{bmatrix} \\ &= (v_1, \dots, v_n) \cdot \begin{bmatrix} \lambda_1 v_1^T \\ \vdots \\ \lambda_n v_n^T \end{bmatrix} \end{aligned}$$

$$= \lambda_1 v_1 v_1^T + \dots + \lambda_n v_n v_n^T. \quad \text{Eigendecomposition.}$$

Sum and Product of Eigenvectors

Sum: Let $A \in R^{n \times n}$ (or $C^{n \times n}$) be a square matrix

$$\Delta \sum_{i=1}^n \lambda_i = \sum_{i=1}^n a_{ii}$$

where $\sum_{i=1}^n a_{ii}$ is called the trace of A , denoted by tr(A)

Product:

$$\Delta \prod_{i=1}^n \lambda_i = \det(A)$$

Eigenvalues of Matrix Power and Inverse

Def. Suppose $A \in C^{n \times n}$ is a square matrix, k , a positive integer.

If λ is an eigenvalue of A , then λ^k is an eigenvalue of A^k .

Notation: Denote EIG(A) as the multiset of eigenvalues of matrix A .

Proposition:

Suppose EIG(A) = { $\lambda_1, \lambda_2, \dots, \lambda_n$ }. (multiset, possible $\lambda_1 = \lambda_2$)

Then: (a) EIG(A^k) = { $\lambda_1^k, \lambda_2^k, \dots, \lambda_n^k$ }.

(b) If A^{-1} exists, then EIG(A^{-1}) = { $\lambda_1^{-1}, \lambda_2^{-1}, \dots, \lambda_n^{-1}$ }.

Google PageRank

Additional materials, no need to remember (?)

Similar Matrix

Def. If $A, B \in R^{n \times n}$ satisfy $A = SBS^{-1}$ for some $S \in R^{n \times n}$, invertible, then we say A and B are similar. This relation is denoted $A \sim B$.

Proposition:

Suppose $A, B \in R^{n \times n}$ are similar, then EIG(A) = EIG(B).

Diagonalizable

Def. If a square matrix A is similar to a diagonal matrix, i.e., $A = SDS^{-1}$ for some diagonal matrix D , then we say A is diagonalizable.

Theorem:

An $n \times n$ matrix A is diagonalizable \Leftrightarrow eigenvectors of A can form a basis.

Proposition: Sufficient condition for diagonalizable

Suppose an $n \times n$ matrix A has n distinct eigenvalues, then A is diagonalizable.

If $A = SDS^{-1}$ for some diagonal matrix D , then $A^k = SD^kS^{-1}$.

XXV. Singular Value Decomposition

SVD Theorem

Def. Any matrix $A \in \mathbb{R}^{m \times n}$ can be written as

$$\Delta A = U \Sigma V^T = \sum_{i=1}^{\min(m,n)} u_i \sigma_i v_i^T$$

Where $U \in \mathbb{R}^{m \times m}$, $V \in \mathbb{R}^{n \times n}$ are two orthogonal matrices,

$\Sigma = \text{rec.diag}(\sigma_1, \dots, \sigma_{\min(m,n)})$ is a rectangular diagonal matrix.

Theorem: Compact SVD and vector form

Any matrix $A \in \mathbb{R}^{m \times n}$ with rank r can be written as

$$A = U_r \Sigma_r V_r^T = \sum_{j=1}^r u_j \sigma_j v_j^T$$

where $U_r = [u_1, \dots, u_r] \in \mathbb{R}^{m \times r}$ has orthogonal unit-norm columns,

$V_r = [v_1, \dots, v_r] \in \mathbb{R}^{n \times r}$ has orthogonal unit-norm columns.

$\Sigma_r = \text{diag}(\sigma_1, \dots, \sigma_r) \in \mathbb{R}_+^{r \times r}$ is a diagonal matrix

Singular Values and Singular Vectors

Def. The columns of U are eigenvectors of $AA^T \in \mathbb{R}^{m \times m}$, called left-singular vectors of A .

The columns of V are eigenvectors of $A^TA \in \mathbb{R}^{n \times n}$, called right-singular vectors of A .

The σ_j 's are called the singular values of A , and $\sigma_j = \sqrt{\lambda_j}$, where $\lambda_1, \dots, \lambda_{\min(m,n)}$ are the eigenvalues of both AA^T and A^TA .

Computation of SVD

Q. Find the full SVD of the matrix A

Step 1: Find eigenvalues of A^TA

Suppose they are: $\lambda_1, \dots, \lambda_r > 0$, and $\lambda_{r+1} = \dots = \lambda_n = 0$.

Denote $\sigma_j = \sqrt{\lambda_j}$, where $j = 1, \dots, \min(m, n)$.

Step 2: Find right singular vectors

① Find r unit eigenvectors of A^TA corresponding to $\lambda_1, \dots, \lambda_r$.

② Find an orthonormal basis of $N(A)$, denote as v_{r+1}, \dots, v_n . *

Just useful for full SVD.

Step 3: Find left singular vectors of A .

① Define left singular vectors $u_i = \frac{Av_i}{\sigma_i}$, for $i = 1, \dots, r$.

② Find an orthonormal basis of $N(A^T)$, denote as u_{r+1}, \dots, u_n . *

Write $\Sigma = \text{rec.diag}(\sigma_1, \dots, \sigma_{\min(m,n)})$, $U = (u_1, \dots, u_m)$, $V = (v_1, \dots, v_n)$.

Relationship with Rank

Lemma: $\text{rank}(A) = \text{number of nonzero singular values}$ (counting multiplicity)

Def. Given the compact SVD $A = U \Sigma V^T = \sum_{i=1}^r u_i \sigma_i v_i^T$, where $U \in \mathbb{R}^{m \times r}$, $V \in \mathbb{R}^{n \times r}$ have full rank.

Then $\text{rank}(A) = \dim(\text{C}(A)) = \dim(\text{span}\{u_1, \dots, u_r\}) = r$.

XXVI. Quadratic Form and Positive Definite Matrix

Quadratic Form

Def. For any real symmetric matrix A , we say $f(x) = x^T A x$ is a quadratic form.

e.g. $A = \begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix}$, $f(x) = x^T A x = [x_1 \ x_2] \begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = 2x_1^2 + 3x_2^2$

Definite Matrices

Def. A real symmetric matrix A is

- i. positive semidefinite iff $x^T A x \geq 0$, $\forall x \in \mathbb{R}^n$, denoted as $A \geq 0$. PSD
- ii. negative semidefinite iff $x^T A x \leq 0$, $\forall x \in \mathbb{R}^n$, denoted as $A \leq 0$. NSD
- iii. positive definite iff $x^T A x > 0$, $\forall x \in \mathbb{R}^n \setminus \{0\}$, denoted as $A > 0$. PD
- iv. negative definite iff $x^T A x < 0$, $\forall x \in \mathbb{R}^n \setminus \{0\}$, denoted as $A < 0$. ND
- iii. indefinite if $\exists x, y \in \mathbb{R}^{n \times 1}$ s.t. $x^T A x < 0 < y^T A y$.

Equivalent Conditions

Def. For an arbitrary real symmetric matrix A , the followings are equivalent:

- (a) A is positive definite (semidefinite)
- (b) All eigenvalues of A are positive (non-negative)
- (c) A can be written as $A = R^T R$ for some invertible matrix R .
(square)

Property:

If A is positive definite, then A is invertible.

Curve Classification

Q: What type of curve does a general equation represent?

$$ax^2 + bxy + cy^2 + dx + ey + f = 0$$

A: We express it as a matrix form:

$$u^T A u + w^T u + f = 0$$

where $u = \begin{bmatrix} x \\ y \end{bmatrix}$, $A = \begin{bmatrix} a & \frac{b}{2} \\ \frac{b}{2} & c \end{bmatrix}$, $w = \begin{bmatrix} d \\ e \end{bmatrix}$.

Then we can classify by eigenvalues of A :

Expression	Signs of eigenvalues of A	Conic section (shape)
$x^T A x + w^T x = 0$	$+ +$ or $- -$	Ellipsoid (or empty set)
	$+ -$	Hyperbola

e.g. $5x_1^2 - 6x_1 x_2 + 5x_2^2 = 4$.

1° EVD: Eigenvalues: $\lambda_1 = 8, \lambda_2 = 2$.

$$\text{Eigenvectors: } V = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}$$

2° Change of variables: $Z = V^T x = \frac{1}{\sqrt{2}} \begin{bmatrix} x_1 + x_2 \\ -x_1 + x_2 \end{bmatrix}$

3° Transformed equation: $8Z_1^2 + 2Z_2^2 = 4 \Rightarrow \frac{Z_1^2}{0.5} + \frac{Z_2^2}{2} = 1$

4° Shape: Since $\lambda_1, \lambda_2 > 0$, the conic is an ellipse.