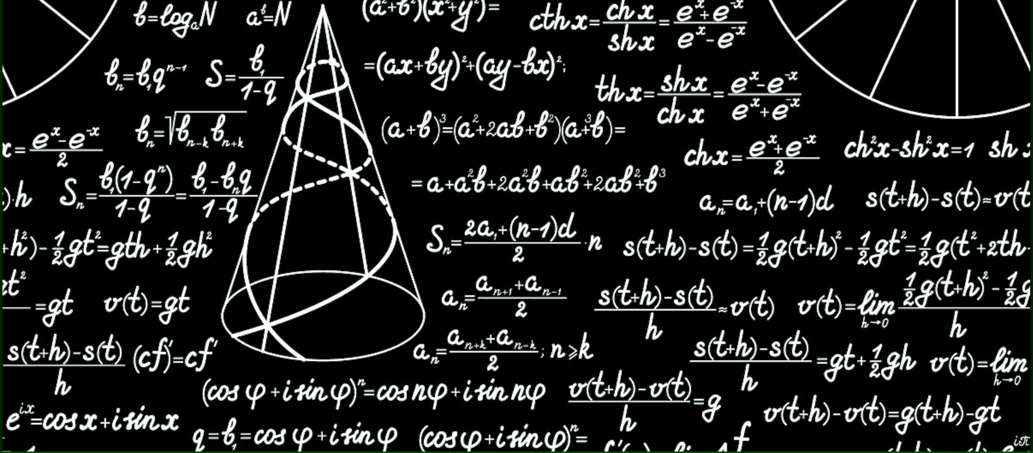


255RC2

Chengyang Shi, Haoran Shen, Jingfan Tang, Ruizhi Deng

1. Transpose

Definition

Given $A = (a_{ij})$, $A^T := (a_{ji})$

Proposition

Let A, B be matrices, $\lambda \in \mathbb{R}$.

1. $(A + B)^T = A^T + B^T$, $(\lambda A)^T = \lambda A^T$
2. $(AB)^T = B^T A^T$
3. $(A^T)^T = A$
4. If A is invertible, so is A^T, and $(A^T)^{-1} = (A^{-1})^T = A^{-T}$

Definition

A matrix A is called **symmetric** if $A = A^T$

A matrix A is called **skew-symmetric** if $A^T = -A \iff A + A^T = 0$

Remark

1. A symmetric or skew-symmetric matrix is necessarily square.
2. For a skew-symmetric matrix, if $i = j$, then $a_{ij} = 0$
3. Any $n \times n$ diagonal matrix is symmetric.

Example

$A = \begin{pmatrix} 1 & -4 & -2 \\ 4 & 1 & 3 \\ 2 & -3 & 2 \end{pmatrix}$ is not skew-symmetric.

2. Matrix as a Function

Matrix as a Function

Definition

An $n \times m$ Matrix can be interpreted as a function $F : M_{m\times p} \to M_{n\times p}$

Property

For $M \in M_{n\times m}, A, B \in M_{m\times p}, \alpha, \beta \in \mathbb{R}$
We have $M(\alpha A + \beta B) = \alpha MA + \beta MB$

Question

1. Do you know other function with such property?
2. Why we want such property?

3. Orthogonal Matrices and Orthonormal Vectors

Orthogonal Matrix

Definition

A matrix $A \in M_n(\mathbb{R})$ is called an orthogonal matrix iff $A^T = A^{-1}$

Properties.

- I_n is an orthogonal matrix.
- If A and B are orthogonal matrices, then AB is also orthogonal.
- If A is an orthogonal matrix, then A^{-1} is also orthogonal matrix.
- If A is an orthogonal matrix, then $det(A) = \pm 1$.
- If A is an orthogonal matrix, then $AA^T = A^T A = I_n$.

Examples.

- Rotation Matrix.

Rotation by θ in $\mathbb{R}^2$ is given by

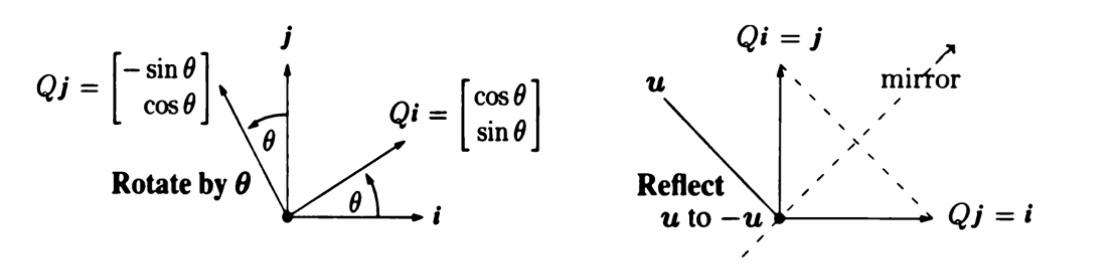
$$\begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}.$$

- Reflection Matrix.

Reflect (x_1, x_2) across $\theta/2$ in $\mathbb{R}^2$ is given by

$$\begin{bmatrix} \cos\theta & \sin\theta \\ \sin\theta & -\cos\theta \end{bmatrix}.$$

You can treat such matrix as a function $Q : \mathbb{R}^2 \to \mathbb{R}^2$



Orthogonal Vectors and Orthonormal Vectors

You may need to distinguish these two concepts clearly ...

Definition

Let $\{v_1, \cdots, v_k\} \in \mathbb{R}$ be a subset of $\mathbb{R}^k$ with k distinct vectors, then $\{v_1, \cdots, v_k\}$ is an orthogonal set of vectors if $\langle v_i, v_j \rangle = 0$ for all $1 \leq i, j \leq k$, $i \neq j$.
Also, $\{q_1, \cdots, q_k\}$ is an orthonormal set of vectors if it is an orthogonal set and all of its vectors are unit vectors (*i.e.*, $\|q_i\| = 1$ for $i \leq i \leq k$).

Remark

Any set containing a single vector is orthogonal; any set containing a single unit vector is orthonormal.

Example.

$\hat{i}, \hat{j}, \hat{k} \in \mathbb{R}^3$.

Orthogonal Vectors and Orthonormal Vectors

Proposition

Given vectors $q_i \in \mathbb{R}^n$, $i = 1, \cdots, m$ such that

$$\langle q_i, q_j \rangle = q_i^T q_j = \delta_{ij} = \begin{cases} \mathbf{1}, & i = j, \\ 0, & i \neq j, \end{cases} \quad (\delta_{ij} : Kronecker)$$

then we call the set of vectors $\{q_i\}$ **orthonormal**, and

$$Q = \begin{bmatrix} | & & | \\ q_1 & \dots & q_m \\ | & & | \end{bmatrix} \Rightarrow Q^T Q = I_m.$$

Orthogonal Vectors and Orthonormal Vectors

Proof

$$Q^TQ = \begin{bmatrix} - & q_1^T & - \\ & \vdots & \\ - & q_m^T & - \end{bmatrix} \begin{bmatrix} | & & | \\ q_1 & \dots & q_m \\ | & & | \end{bmatrix} = \begin{bmatrix} q_1^Tq_1 & \dots & q_1^Tq_m \\ \vdots & \ddots & \vdots \\ q_m^Tq_1 & \dots & q_m^Tq_m \end{bmatrix} = I_m$$

Proposition

Let $Q = \begin{bmatrix} | & | & & | \\ \mathbf{q}_1 & \mathbf{q}_2 & \cdots & \mathbf{q}_n \\ | & | & & | \end{bmatrix} \in M_n(\mathbb{R})$, where $\{\mathbf{q}_1, \mathbf{q}_2, \dots, \mathbf{q}_n\}$ is a set of **orthonormal** vectors in $\mathbb{R}^n$.

Then Q is an orthogonal matrix, $Q^T = Q^{-1}$.

Proof

We verify by showing $QQ^T = I_n$ and $Q^TQ = I_n$ holds.

To show matrix B is the inverse of matrix A, we need to show **both** $AB = I$ and $BA = I$.

Exercise

Determine whether the following matrices are orthogonal matrices

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

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

Proposition

The orthogonal matrices are precisely the matrices that preserve the inner product in $\mathbb{R}^n$ i.e., $\forall x, y \in \mathbb{R}^n$, $\langle x, y\rangle = \langle Ax, Ay\rangle$ or $x^T y = (Ax)^T Ay$.

证明.

If A is orthogonal, i.e., $A^{-1} = A^T$, then $A^T A = I_n$, so
$(Ax)^T(Ay) = x^T A^T Ay = x^T I_n y = x^T y$. $\square$

定义 (Inner Product Preservation)

一个矩阵 Q 保持内积不变，意味着对任意两个向量 $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$，应用 Q 变换后，它们的内积保持不变，即：

$$\langle Q\mathbf{u}, Q\mathbf{v} \rangle = \langle \mathbf{u}, \mathbf{v} \rangle$$

几何意义

- **长度不变**：由于向量范数为

$$\|\mathbf{u}\| = \sqrt{\mathbf{u} \cdot \mathbf{u}},$$

我们有：

$$\|Q\mathbf{u}\| = \sqrt{(Q\mathbf{u}) \cdot (Q\mathbf{u})} = \sqrt{\mathbf{u} \cdot \mathbf{u}} = \|\mathbf{u}\|.$$

因此，正交矩阵保持向量的长度不变。

几何意义

保持内积不变意味着矩阵 Q 代表的线性变换保留了 $\mathbb{R}^n$ 的几何结构，具体包括：

- **角度不变**：内积决定了两向量夹角 θ，因为

$$\mathbf{u} \cdot \mathbf{v} = \|\mathbf{u}\| \|\mathbf{v}\| \cos(\theta).$$

  保持内积不变意味着 $\mathbf{u}$ 和 $\mathbf{v}$ 的夹角与 $Q\mathbf{u}$ 和 $Q\mathbf{v}$ 的夹角相同。

- **正交性不变**：如果 $\mathbf{u}$ 和 $\mathbf{v}$ 垂直（即 $\mathbf{u} \cdot \mathbf{v} = 0$），那么

$$(Q\mathbf{u}) \cdot (Q\mathbf{v}) = 0,$$

  即 $Q\mathbf{u}$ 和 $Q\mathbf{v}$ 也垂直。

Why we need Orthonormal Matrix?

1. **Magnitude of Determinant and Column Norms:**
   - $|\det(A)| = 1$, where A is an orthogonal matrix.
   - Column vectors $\mathbf{q}_i$ have unit norm: $||\mathbf{q}_i|| = 1$.
   - *Implication:* The transformation represented by A is a **rotation** or a **reflection**.
2. **Defining Property and Its Consequences:**
   - $A^T A = I_n$ (which means $A^{-1} = A^T$).
   - The original note mentions this "has very good symmetric properties."
   - *Implications:*
     - Preservation of inner products (hence, lengths and angles are preserved: $\langle Ax, Ay \rangle = \langle x, y \rangle$).
     - The inverse A^{-1} is easy to compute (it's just the transpose A^T).
3. **Applications in Data Processing:**
   - Effective in removing redundant information from data

4. Kernel and Image

Definition

- Given matrix $A \in M_{m\times n}(\mathbb{R})$. The kernel, or nullspace of A, is defined as

$$\ker A = N(A) = \{x \in \mathbb{R}^n \mid Ax = \mathbf{0}\} \subseteq \mathbb{R}^n$$

  where $A : \mathbb{R}^n \to \mathbb{R}^m$.
  $\to$ "Those inputs s.t. the output is 'zero'."

- The image of A, or column space, is defined as

$$\begin{aligned} \operatorname{im} A = C(A) &= \{y \in \mathbb{R}^m \mid \exists x \in \mathbb{R}^n, y = Ax\} \\ &= \{Ax \mid x \in \mathbb{R}^n\} \end{aligned}$$

  where $A : \mathbb{R}^n \to \mathbb{R}^m$.
  $\to$ "The set of all possible outputs"
  (The image is a subspace of $\mathbb{R}^m$).

Injectivity

Definition

A function $f : A \to B$ is called **injective**, if
$\forall x_1, x_2 \in A, f(x_1) = f(x_2) \implies x_1 = x_2$
i.e., $\forall x_1, x_2 \in A, x_1 \neq x_2 \implies f(x_1) \neq f(x_2)$

- 输入与输出一一对应

★ Proposition

Given $A \in M_{m \times n}(\mathbb{R})$, i.e., $A : \mathbb{R}^n \to \mathbb{R}^m$
then, A is injective $\iff \ker A = N(A) = \{0\}$

Proof

($\Rightarrow$) A is injective $\implies \forall x_1, x_2 \in \mathbb{R}^n, Ax_1 = Ax_2 \implies x_1 = x_2$.
Take $v \in \ker A$, i.e., $Av = 0$. We also know that $A0 = 0$, hence $Av = A0$. By injectivity of A, $v = 0$. Therefore $\ker A = \{0\}$.

($\Leftarrow$) Conversely, we know $\ker A = \{0\}$.
Take $v_1, v_2 \in \mathbb{R}^n$, s.t. $Av_1 = Av_2$. We want to show that $v_1 = v_2$. Indeed, we have $A(v_1 - v_2) = 0$, but $\ker A = \{0\}$, hence $v_1 - v_2 = 0$. So $v_1 = v_2$.

Proposition

Given $A \in M_{m\times n}(\mathbb{R})$, $\ker(A) = \ker(A^T A)$

Proof

($\subseteq$) Take $x \in \ker(A)$, i.e., $Ax = 0$. So $(A^T A)x = A^T(Ax) = A^T 0 = 0$. So $x \in \ker(A^T A)$.
($\supseteq$) Take $x \in \ker(A^T A)$, i.e., $A^T Ax = 0$. So $x^T A^T Ax = x^T 0 = 0$.
$(Ax)^T Ax = \|Ax\|^2 = 0. \implies Ax = 0 \implies x \in \ker A.$
$(v_1^2 + v_2^2 + \cdots + v_n^2 = 0 \implies v_i = 0)$

5. Projection Matrix

Definition

A matrix $P \in M_n(\mathbb{R})$ is a **projection matrix**, if $P^2 = P$.
A projection matrix P is called an **orthogonal projection**, if $P = P^T$.

Example

For any orthonormal matrix Q, $QQ^T : \mathbb{R}^n \to \mathbb{R}^n$ is a projection matrix.

- $(QQ^T)^T = (Q^T)^T Q^T = QQ^T$
- $(QQ^T)^2 = (QQ^T)(QQ^T) = Q(Q^T Q)Q^T = QI_k Q^T = QQ^T$

Orthogonal Projection Matrix

Proposition

A projection matrix P is an **orthogonal projection** if

$$\ker P \perp \operatorname{im} P$$

i.e., $\forall x \in \ker P$, $\forall y \in \operatorname{im} P$,

$$\langle x, y \rangle = x^T y = 0 \in \mathbb{R}$$

Exercise

Show that $\ker P \perp \operatorname{im} P$ if $P = P^T$

Example

$$\{q_1, q_2\} = \left\{ \begin{pmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \\ 0 \end{pmatrix}, \begin{pmatrix} 1/\sqrt{2} \\ -1/\sqrt{2} \\ 0 \end{pmatrix} \right\}$$

$$Q = \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ 1/\sqrt{2} & -1/\sqrt{2} \\ 0 & 0 \end{bmatrix}, \quad Q^T = \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} & 0 \\ 1/\sqrt{2} & -1/\sqrt{2} & 0 \end{bmatrix}$$

$$QQ^T = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x \\ y \\ 0 \end{pmatrix}$$

1. VV255, Lecture Notes 24SU, Runze Cai
2. VV255 RC2, Jiayue Huang, et al
3. Introduction to Linear Algebra, Sixth Edition, Gilbert Strang