ĐẠI SỐ TUYẾN TÍNH

PiMA 2019: The Mathematics of Deep Learning

Ngày 29 tháng 07 năm 2019

Mục lục

Vector trong $\mathbb{R}^n$

Vector trong $\mathbb{R}^n$

Trong $\mathbb{R}^n$, một vector là một bộ n số thực có thứ tự:

$$\boldsymbol{x} = (x_1, \ldots, x_n), \; x_i \in \mathbb{R}, i \in \{1, \ldots, n\}.$$

Ví dụ.

- $\boldsymbol{a} = (1, 2)$ là một vector trong $\mathbb{R}^2$.
- $\boldsymbol{e_1} = (1, 0, 0), \boldsymbol{e_2} = (0, 1, 0), \boldsymbol{e_3} = (0, 0, 1)$ là các vector trong $\mathbb{R}^3$.

Vector trong $\mathbb{R}^n$

Cho $\boldsymbol{x} = (x_1, \ldots, x_n), \boldsymbol{y} = (y_1, \ldots, y_n) \in \mathbb{R}^n$ và $\alpha \in \mathbb{R}$, ta định nghĩa phép cộng vector và phép nhân vô hướng với vector như sau:

$$\boldsymbol{x} + \boldsymbol{y} = (x_1 + y_1, \ldots, x_n + y_n),$$

$$\alpha \boldsymbol{x} = (\alpha x_1, \ldots, \alpha x_n).$$

Ngoài ra, ký hiệu $-\boldsymbol{x} = (-1)\boldsymbol{x} = (-x_1, \ldots, -x_n)$.

Ví dụ. Cho $\boldsymbol{u} = (1, 2), \boldsymbol{v} = (3, 1) \in \mathbb{R}^2$, ta có:

$$3\boldsymbol{u} - \boldsymbol{v} = 3(1, 2) + (-1)(3, 1) = (3, 6) + (-3, -1) = (0, 5).$$

Vector trong $\mathbb{R}^n$

Tích vô hướng trong $\mathbb{R}^n$ (Inner product)

Trong $\mathbb{R}^n$, với $\boldsymbol{x} = (x_1, \ldots, x_n), \boldsymbol{y} = (y_1, \ldots, y_n)$, ta định nghĩa:

$$\boldsymbol{x} \cdot \boldsymbol{y} := x_1 y_1 + \ldots + x_n y_n.$$

Khi đó, với mọi $\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z} \in \mathbb{R}^n$ và với mọi $\alpha, \beta \in \mathbb{R}$, các tính chất sau được thỏa mãn:

- $\boldsymbol{x} \cdot \boldsymbol{y} = \boldsymbol{y} \cdot \boldsymbol{x}$;
- $\boldsymbol{x} \cdot \boldsymbol{x} \geq 0$ và $\boldsymbol{x} \cdot \boldsymbol{x} = 0 \Leftrightarrow \boldsymbol{x} = \boldsymbol{0} = (0, \ldots, 0)$;
- $(\alpha \boldsymbol{x} + \beta \boldsymbol{y}) \cdot \boldsymbol{z} = \alpha(\boldsymbol{x} \cdot \boldsymbol{z}) + \beta(\boldsymbol{y} \cdot \boldsymbol{z})$;
- $\boldsymbol{x} \cdot (\alpha \boldsymbol{y} + \beta \boldsymbol{z}) = \alpha(\boldsymbol{x} \cdot \boldsymbol{y}) + \beta(\boldsymbol{x} \cdot \boldsymbol{z})$.

Ta được một **tích vô hướng (inner product)** trong $\mathbb{R}^n$.

Ví du. Trong $\mathbb{R}^2$, cho $\boldsymbol{u} = (1, 2), \boldsymbol{v} = (3, 1)$. Hãy tính $\boldsymbol{u} \cdot \boldsymbol{v}$.

Vector trong $\mathbb{R}^n$

Chuẩn của vector (Norm)

Trong $\mathbb{R}^n$ với tích vô hướng thông thường, với $\boldsymbol{x} = (x_1, \ldots, x_n)$, ta định nghĩa:

$$||\boldsymbol{x}|| := \sqrt{\boldsymbol{x} \cdot \boldsymbol{x}} = \sqrt{x_1^2 + \ldots + x_n^2}.$$

Ta được một **chuẩn (norm)** trên $\mathbb{R}^n$. Nếu $\boldsymbol{v}$ thỏa $||v|| = 1$ thì ta nói $\boldsymbol{v}$ là **vector đơn vị (unit vector)**.

Nhận xét

Với mọi $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$ và với mọi $\alpha \in \mathbb{R}$, ta có:

1. $||\boldsymbol{x}|| \geq 0$ và $||\boldsymbol{x}|| = 0 \Leftrightarrow \boldsymbol{x} = \boldsymbol{0}$;
2. $||\alpha \boldsymbol{x}|| = |\alpha| ||\boldsymbol{x}||$;
3. $||\boldsymbol{x} + \boldsymbol{y}|| \leq ||\boldsymbol{x}|| + ||\boldsymbol{y}||$.

Vector trong $\mathbb{R}^n$

Trong $\mathbb{R}^n$, ngoài chuẩn được định nghĩa như trên, ta còn có một số chuẩn thông dụng như sau:

$$||\boldsymbol{x}||_1 = \sum_{i=1}^{n} |x_i| = |x_1| + \ldots + |x_n|;$$

$$||\boldsymbol{x}||_2^2 = \sum_{i=1}^{n} x_i^2 = x_1^2 + \ldots + x_n^2;$$

$$||\boldsymbol{x}||_\infty = \max_{1 \le i \le n} |x_i| = \max\{|x_1|, \ldots, |x_n|\}.$$

Vector trong $\mathbb{R}^n$

Example

Trong không gian $\mathbb{R}^2$, tìm các vector $\boldsymbol{x}$ sao cho:

a) $||\boldsymbol{x}||_1 = 1$;

b) $||\boldsymbol{x}||_2 = 1$;

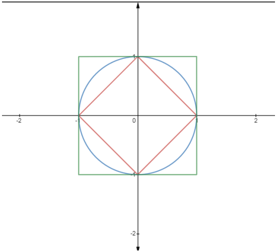
c) $||\boldsymbol{x}||_\infty = 1$.

Vector trong $\mathbb{R}^n$

Example

Trong không gian $\mathbb{R}^2$, tìm các vector $\boldsymbol{x}$ sao cho:

a) $||\boldsymbol{x}||_1 = 1$;
b) $||\boldsymbol{x}||_2 = 1$;
c) $||\boldsymbol{x}||_\infty = 1$.



Vector trong $\mathbb{R}^n$

Ứng dụng của tích vô hướng

Câu hỏi: Tìm hình chiếu của vector $\boldsymbol{x}$ lên vector $\boldsymbol{y}$ trong $\mathbb{R}^n$.

Cho $\boldsymbol{x}$ và $\boldsymbol{y}$ là hai vector khác vector không trong $\mathbb{R}^n$. **Hình chiếu của $\boldsymbol{x}$ lên $\boldsymbol{y}$**, ký hiệu là $\mathrm{proj}_{\boldsymbol{y}}(\boldsymbol{x})$, là một vector $\boldsymbol{z} \in \mathbb{R}^n$ sao cho:

- $\boldsymbol{z}$ cùng phương với $\boldsymbol{y}$, tức là $\exists c \in \mathbb{R} : \boldsymbol{z} = c\boldsymbol{y}$;
- $(\boldsymbol{x} - \boldsymbol{z}) \cdot \boldsymbol{y} = 0$.

Từ định nghĩa trên, hãy lập công thức biểu diễn $\mathrm{proj}_{\boldsymbol{y}}(\boldsymbol{x})$.

Ví dụ. Trong $\mathbb{R}^3$ với tích vô hướng được định nghĩa như từ đầu, hãy tìm hình chiếu của $\boldsymbol{m} = (1, 2, 3)$ lên $\boldsymbol{n} = (3, 2, 1)$.

Siêu phẳng trong $\mathbb{R}^n$

Siêu phẳng (Hyperplane)

Trong $\mathbb{R}^n$ với tích vô hướng thông thường, cho $\boldsymbol{a} = (a_1, \ldots, a_n) \neq \mathbf{0}$ và cho $c \in \mathbb{R}$. Tập nghiệm H của phương trình tuyến tính

$$a_1 x_1 + \ldots + a_n x_n = c$$

là một **siêu phẳng (hyperplane)** của $\mathbb{R}^n$.

$$H = \{\boldsymbol{x} \in \mathbb{R}^n : \boldsymbol{a} \cdot \boldsymbol{x} = c\}.$$

Nhận xét. Siêu phẳng H chia không gian $\mathbb{R}^n$ làm ba miền: $\boldsymbol{a} \cdot \boldsymbol{x} < c; \boldsymbol{a} \cdot \boldsymbol{x} = c; \boldsymbol{a} \cdot \boldsymbol{x} > c$. Như vậy, tích vô hướng còn có thể dùng để xác định vị trí tương đối của một điểm với một siêu phẳng.

Ma trận

Ma trận (Matrix)

Cho K là một trường. Một **ma trận (matrix)** loại m x n trên K là một bảng chữ nhật gồm m dòng, n cột với mn hệ số trong K có dạng

$$A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix}.$$

Ta gọi a_{ij} là phần tử (hệ số) dòng i, cột j của ma trận A, còn được kí hiệu là $[A]_{ij}$. Ký hiệu $M_{mxn}(K)$ là tập hợp tất cả ma trận loại m x n trên trường K. Trong suốt bài trình bày này, chúng ta chỉ làm việc trên trường số thực $\mathbb{R}$.

Ma trận

- Các dòng của ma trận A được gọi là các vector dòng.

$$(a_{i1}, \ldots, a_{in})$$

là dòng thứ i của ma trận A. Các ma trận chỉ có một dòng cũng được gọi là vector dòng.

- Các cột của ma trận A được gọi là các vector cột.

$$\begin{pmatrix} a_{1j} \\ a_{2j} \\ \ldots \\ a_{nj} \end{pmatrix}$$

là cột thứ j của ma trận A. Các ma trận chỉ có một cột cũng được gọi là vector cột.

- Hai ma trận $A = (a_{ij})$ và $B = (b_{ij})$ được gọi là bằng nhau khi và chỉ khi chúng có cùng loại và $a_{xy} = b_{xy}, \forall x, y$.

Các phép toán ma trận

Phép chuyển vị ma trận (Transpose)

Cho A là một ma trận m dòng n cột. **Chuyển vị (transpose)** của ma trận A, ký hiệu A^T, là một ma trận n dòng m cột được xác định như sau:

$$[A^T]_{ij} = [A]_{ji}, \forall i \in \{1, \dots, n\}, j \in \{1, \dots, m\}.$$

Ví dụ. Cho các ma trận A, B, C xác định như sau: $\mathrm{A} = \begin{bmatrix} 0 & 1 & 2 & 3 \\ 4 & 5 & 6 & 7 \\ 8 & 9 & 10 & 11 \end{bmatrix}$, $\mathrm{B} = \begin{bmatrix} 1 & 0 & 2 \\ 0 & 3 & 5 \\ 2 & 5 & 0 \end{bmatrix}$, $\mathrm{C} = \begin{bmatrix} 0 & 1 & 2 \\ -1 & 0 & 3 \\ -2 & -3 & 0 \end{bmatrix}$. Hãy tìm các ma trận A^T, B^T, C^T.

Các phép toán ma trận

Phép cộng ma trận

Cho các ma trận A và B có cùng kích thước m dòng, n cột thuộc $M_{m\times n}(\mathbb{R})$. **Tổng** của các ma trận A và B, ký hiệu là $A+B$, là một ma trận m dòng, n cột thuộc $M_{m\times n}(\mathbb{R})$ được xác định như sau:

$$[A+B]_{ij} = [A]_{ij} + [B]_{ij}, \forall i \in \{1,\dots,m\}, j \in \{1,\dots,n\}.$$

Tích của số thực với ma trận

Cho ma trận A thuộc $M_{m\times n}(\mathbb{R})$ và một số thực α. **Tích** của số thực α với ma trận A, ký hiệu là αA, là một ma trận thuộc $M_{m\times n}(\mathbb{R})$ được xác định như sau:

$$[\alpha A]_{ij} = \alpha [A]_{ij}, \forall i \in \{1,\dots,m\}, j \in \{1,\dots,n\}.$$

Các phép toán ma trận

Ma trận $(-1)A$ được gọi là **ma trận đối** của A, ký hiệu là $-A$.

Tính chất

Với mọi $A, B, C \in M_{m\times n}(\mathbb{R})$ và với mọi $\alpha, \beta \in \mathbb{R}$ thì

1. $A + B = B + A$;
2. $(A + B) + C = A + (B + C)$;
3. $A + \mathbf{0}_{m\times n} = \mathbf{0}_{m\times n} + A = A$ ($\mathbf{0}_{m\times n}$ chỉ ma trận không);
4. $A + (-A) = (-A) + A = \mathbf{0}$;
5. $1.A = A$;
6. $(\alpha\beta)A = \alpha(\beta A)$;
7. $(\alpha + \beta)A = \alpha A + \beta A$;
8. $\alpha(B + C) = \alpha B + \alpha C$.

Các phép toán ma trận

Phép nhân ma trận

Cho các ma trận: $A \in M_{m\,x\,n}(\mathbb{R}), B \in M_{n\,x\,p}(\mathbb{R})$. **Tích của ma trận A và B**, ký hiệu là AB, là một ma trận thuộc $M_{m\,x\,p}(\mathbb{R})$ và được xác định như sau:

$$[AB]_{ij} = \sum_{k=1}^{n} [A]_{ik}[B]_{kj}, \forall k \in \{1, \ldots, n\}$$

hay

$$[AB]_{ij} = [A]_{i1}[B]_{1j} + \ldots + [A]_{in}[B]_{nj}.$$

Điều kiện để thực hiện được phép nhân ma trận A với ma trận B là: **Số cột của ma trận A bằng với số dòng của ma trận B.**

Các phép toán ma trận

Nhận xét: Nếu đặt $\boldsymbol{a}$ là vector dòng i của ma trận A và $\boldsymbol{b}$ là vector cột j của ma trận B:

$$\boldsymbol{a} = ([A]_{i1}, \ldots, [A]_{in}), \boldsymbol{b} = ([B]_{1j}, \ldots, [B]_{nj}),$$

thì tích vô hướng của $\boldsymbol{a}$ và $\boldsymbol{b}$ là:

$$\boldsymbol{a} \cdot \boldsymbol{b} = [A]_{i1}[B]_{1j} + \ldots + [A]_{in}[B]_{nj} = [AB]_{ij}.$$

Như vậy, hệ số dòng i, cột j của ma trận AB bằng với tích vô hướng (thông thường) của vector dòng i của ma trận A và vector cột j của ma trận B.

Các phép toán ma trận

Ví dụ. Cho các ma trận A, B xác định như sau:

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

Hãy tính AB và BA.

Các tính chất của phép nhân ma trận

Cho $M, M_1, M_2 \in M_{m\times n}(\mathbb{R})$, $N, N_1, N_2 \in M_{n\times p}(\mathbb{R})$, $P \in M_{p\times q}(\mathbb{R})$. Ta có:

- $(MN)P = M(NP)$;
- $(M_1 + M_2)N = M_1N + M_2N$; $M(N_1 + N_2) = MN_1 + MN_2$;
- $(MN)^T = N^TM^T$.

Ma trận vuông

Ma trận vuông (Square matrix)

Ma trận vuông là ma trận có số dòng và số cột **bằng nhau**.
Ký hiệu: $M_n(\mathbb{R})$ chỉ tập hợp các ma trận vuông cấp n với các hệ số thuộc trường $\mathbb{R}$ (thay vì $M_{n\times n}(\mathbb{R})$).

Nhận xét: Cho hai ma trận vuông cùng cấp A và B. Khi đó phép nhân ma trận luôn thực hiện được và ma trận kết quả cũng là ma trận vuông cùng cấp với hai ma trận đã cho.

Ma trận vuông

Cho ma trận vuông $A = (a_{ij})$ cấp n. Ta gọi *đường chéo chính* của A là tập hợp các hệ số a_{ii} với $i \in \{1, \ldots, n\}$.

Một số ma trận vuông đặc biệt:

- **Ma trận đường chéo (diagonal matrix)** là ma trận vuông có tính chất: Mọi hệ số không thuộc đường chéo chính đều bằng 0.
- **Ma trận đơn vị (identity matrix)** là ma trận đường chéo có tính chất: Mọi hệ số thuộc đường chéo chính đều bằng 1.
  Ký hiệu I_n chỉ ma trận đơn vị cấp n.

Tính chất

Cho $A \in M_{m\times n}(\mathbb{R})$. Ta luôn có:

$$I_m A = A I_n = A.$$

Ma trận vuông

Một số ma trận vuông đặc biệt (tiếp theo):

- **Ma trận tam giác trên (upper triangular matrix)** là ma trận vuông có tính chất: Mọi hệ số a_{ij} với $n \geq i > j \geq 1$ đều bằng 0.
- **Ma trận tam giác dưới (lower triangular matrix)** là ma trận vuông có tính chất: Mọi hệ số a_{ij} với $1 \leq i < j \leq n$ đều bằng 0.

Tính chất

Tổng và tích của hai ma trận tam giác trên (dưới) cùng cấp là ma trận tam giác trên (dưới).

Ma trận vuông

- **Ma trận đối xứng (symmetric matrix)** là ma trận vuông bằng với ma trận chuyển vị của chính nó. Tức là:

$$A = A^T,$$

hay nói cách khác

$$a_{ij} = a_{ji}, \forall i, j \in \{1, \dots, n\}.$$

- **Ma trận phản xứng (antisymmetric matrix)** là ma trận vuông có tính chất:

$$A = -A^T,$$

hay nói cách khác

$$a_{ij} = -a_{ji}, \forall i, j \in \{1, \dots, n\}.$$

Ma trận nghịch đảo

Ma trận nghịch đảo (Inverse matrix)

Cho ma trận $A \in M_n(\mathbb{R})$. Ta nói ma trận A **khả nghịch** nếu tồn tại một ma trận $B \in M_n(\mathbb{R})$ sao cho:

$$AB = BA = I_n.$$

Khi đó B cũng là ma trận duy nhất thỏa mãn điều kiện trên. Ta nói ma trận B là **nghịch đảo** của ma trận A, ký hiệu là A^{-1}.

Ma trận và hệ phương trình tuyến tính

Hệ phương trình tuyến tính (System of Linear Equations)

Một hệ phương trình tuyến tính trên $\mathbb{R}$ là một hệ thống gồm m phương trình bậc nhất với n ẩn có dạng:

$$\begin{aligned}
a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= b_1 \\
a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= b_2 \\
\vdots \qquad \vdots \qquad\qquad \vdots \quad &\quad \vdots \\
a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n &= b_m
\end{aligned}$$

Ma trận và hệ phương trình tuyến tính

Đặt

$$A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix}, X = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}, B = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}.$$

Ta gọi A là *ma trận các hệ số*, X là *cột các ẩn* và B là *cột các hệ số tự do*. Hệ phương trình đã cho tương đương với phương trình $AX = B$.

TÀI LIỆU THAM KHẢO

Trinh Thanh Deo, Le Van Luyen, Bui Xuan Hai, and Tran Ngoc Hoi.
Đại số Tuyến tính và Ứng dụng (Tập 1).
Nhà Xuất bản Đại học Quốc gia Tp.HCM, 2009.

Nguyen Huu Viet Hung.
Đại số tuyến tính.
Nhà Xuất bản Đại học Quốc gia Hà Nội, 2001.