

CHƯƠNG 4: TÍCH VÔ HƯỚNG, ĐỘ LỚN VÀ ĐỘ GIỐNG NHAU/ KHOẢNG CÁCH GIỮA CÁC VECTOR

Karlie Le

March 2019

1 Hệ cơ sở \Rightarrow Không gian tọa độ:

- Ví dụ 1: $R^3 \ni v = (x, y, z)$: Basis $u = (u_x, u_y, u_z) = (i, j, k)$

$$u_x = (1, 0, 0)$$

$$u_y = (0, 1, 0)$$

$$u_z = (0, 0, 1)$$

$$\Rightarrow [v]_u = (x, y, z)$$

Tổng quát: $R^n \ni v = (x_1, \dots, x_n)$:

$$u_1 = (1, 0, \dots, 0)$$

...

$$u_n = (0, \dots, 0, 1)$$

$$\Rightarrow [v]_u = (x_1, \dots, x_n)$$

- Ví dụ 2: $R^{2 \times 2} \ni v = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$

$$u_1 = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}; u_2 = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$$

$$u_3 = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}; u_4 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

$$\Rightarrow [v]_u = (a, b, c, d) \in R^4$$

- Ví dụ 3: (Không gian đa thức)

$$P_n(R) \ni v(z) = \sum_{i=0}^n a_i z^i, a_i \in R \forall i = 0, \dots, n:$$

$$\varepsilon = \{e_i = z^i\}_{i=0}^n \Rightarrow [v]_\varepsilon = (a_0, \dots, a_n) \in R^{n+1}$$

LƯU Ý:

- Mọi vector space đều có thể quy về R^n
- Dimension n là số basis vector nhỏ nhất để miêu tả $\forall v \in V$

2 Hàm tuyến tính \Rightarrow Ma trận:

- Hàm $f: V \rightarrow W$ (biến đổi từ không gian vector V qua W)
 Chọn Basis B cho $V \xrightarrow[\text{space}]{\text{coordinate}} [v]_B \in R^n, \forall v \in V$
 Chọn Basis D cho $W \xrightarrow[\text{space}]{\text{coordinate}} [w]_D \in R^m, \forall w \in W$
- $f: V \rightarrow W$ là hàm tuyến tính (linear) nếu $f(au + bv) = af(u) + bf(v)$
"Cộng và scale đầu vào \rightarrow đầu ra cộng và scale tương ứng"
- $f(au + bv) \neq af(u) + bf(v)$: hàm phi tuyến (non-linear)
Ví dụ: Hàm sigmoid $\sigma(z) = \frac{1}{1+e^{-z}} = \frac{e^z}{e^z+1}$ với $z \in R^n$

3 Một số phép toán ma trận:

- Ma trận $A = [a_{ij}] \in R^{m \times n}$ gồm m hàng và n cột
 $A^T = [a_{ji}] \in R^{n \times m}$. $(A+B)^T = A^T + B^T$
- Vector b theo quy ước mặc định là vector **cột**
- $A = [c_1] \dots [c_n]$ gồm n cột, cột thứ j là c_j
 $A = [r_1^T] \dots [r_n^T]$ gồm n cột, cột thứ j là r_j^T
- Các cột còn lại là độc lập tuyến tính (linear independent) khi không có cột nào là linear combination của các cột còn lại:
 $\sum_{j=1}^n a_j c_j = 0_n \iff a_j = 0 \forall j=1, \dots, n$
- Matrix product: $AB = Q: [m,n][n,p] = [m,p]; q_{ik} = \sum_{j=1}^n a_{ij} b_{jk}$
 Các vector cột: $q_k = A_{b_k}$. $(AB)^T = A^T B^T$
- Tích vô hướng: $u \cdot v := u^T v = \sum_i u_i v_i \in R$ (**inner/dot/scalar product**)
 $uxv := uv^T \in R^{m \times m}$ (**Outer product**)
- Nghịch đảo M^{-1} của ma trận vuông: $MM^{-1} = I_m = \text{diag}(1, \dots, 1)$. $(AB)^{-1} = B^{-1}A^{-1}$

- Ma trận vuông đối xứng (symmetric): $H^T = H$
- Ma trận vuông đối xứng $M \in R^{m \times n}$ là xác định dương (**positive definite**) nếu $x^T A x > 0, \forall x \neq 0_m$.
- M có trace $\text{tr}(M) := \sum_{i=1}^n a_{ii}$. Frobenius norm $A_F := \sqrt{\sum_{i=1}^m \sum_{j=1}^n a_{ij}^2} := \sqrt{\text{tr}(A^T A)}$
- Quadratic function:

$$f(x) = a + b^T x + \frac{1}{2} x^T C x = \frac{1}{2} \sum_{i,j=1}^n c_{ij} x_i x_j + \sum_{k=1}^n b_k x_k + a$$
với $a \in R, x, b \in R^n, C \in R^{n \times n}$ symmetric và ít nhất một hệ số $c_y \neq 0$
- Quadratic form:
 $q(x) = x^T C x = \sum_{i,k=1}^n c_{ik} x_i x_k$ hàm của các đa thức cùng bậc 2 $x_i x_j$

4 Tích vô hướng và khoảng cách của 2 vector trong R^n

- a_i là mức độ giống/khác giữa v và basis vector e_i trong V .
- $a \cdot b = \sum_{i=1}^n a_i b_i \in R$ là **mức độ liên quan** giữa $a, b \in R^n$
 \Rightarrow Là cơ sở để tính độ lớn, khoảng cách, góc,..., giữa các vectors.
- Orthogonal (perpendicular) vectors $a \perp b \Rightarrow \cos \theta = 0 \leftrightarrow a \cdot b = 0$
- Normalized (unit/direction) vector: $\|b\| = 1$
- Orthonormal vectors: $a \cdot b = 0$ và $\|a\| = \|b\| = 1$

5 Khái quát hóa lên abstract vector spaces

- Inner product tổng quát là 1 hàm $\langle \cdot, \cdot \rangle: V \times V \rightarrow R$ thỏa 3 tính chất trong dot product
- Trong $R^n \ni a, b$ và symmetric, positive-definite $\Sigma \in R^{n \times n}$:
 $\langle a, b \rangle_\Sigma = a^T \Sigma b$