KHÔNG GIAN EUCLID

Nguyễn Mạnh Hùng

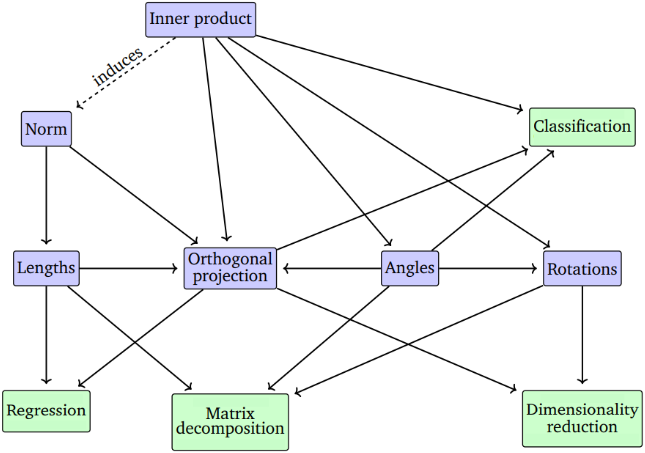
AI Academy Vietnam

July, 2024

Nội dung

1. Không gian Euclid
2. Khái niệm hình học trong không gian Euclid
3. Cơ sở trực chuẩn và phép chiếu trực giao
4. Phân tích QR

Sơ đồ các khái niệm cơ bản



Tích vô hướng (inner product)

Trường hợp $\mathbb{R}^n$

Cho u và v là hai véc tơ trong $\mathbb{R}^n$: $\mathrm{u} = \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix}$ và $\mathrm{v} = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix}$.

Khi đó, tích vô hướng (*dot product*) của u và v là

$$\mathrm{u} \cdot \mathrm{v} = \mathrm{u}^T \mathrm{v} = u_1 v_1 + u_2 v_2 + \cdots + u_n v_n$$

Ví dụ: Xét $\mathrm{u} = \begin{bmatrix} 2 \\ -5 \\ -1 \end{bmatrix}$ và $\mathrm{v} = \begin{bmatrix} 3 \\ 2 \\ -3 \end{bmatrix}$. Suy ra

$$\mathrm{u} \cdot \mathrm{v} = (2)(3) + (-5)(2) + (-1)(-3) = -1$$

Tích vô hướng

Trường hợp tổng quát

Cho một không gian véc tơ V. Một ánh xạ $\varphi : V \times V \to \mathbb{R}$ được gọi là một tích vô hướng trong V nếu thỏa mãn:

1. *Tính đối xứng*: $\varphi(x, y) = \varphi(y, x)$ với mọi x, y $\in V$.
2. *Tính song tuyến tính*: với mọi x, y, z $\in V$ và λ, $\beta \in \mathbb{R}$,

$$\varphi(\lambda x + \beta y, z) = \lambda\varphi(x, z) + \beta\varphi(y, z)$$
$$\varphi(x, \lambda y + \beta z) = \lambda\varphi(x, y) + \beta\varphi(x, z)$$

3. *Tính xác định dương*: $\forall x \in V \setminus \{0\} : \varphi(x, x) > 0, \quad \varphi(0, 0) = 0.$

Chú ý: Tích vô hướng thường được viết $\langle x, y\rangle$ thay cho $\varphi(x, y)$.

Tích vô hướng

Không gian Euclid (Euclidean space)

Không gian véc tơ V có tồn tại một tích vô hướng $\langle \cdot, \cdot \rangle$ được gọi là một *không gian tích vô hướng* hay một *không gian Euclid*.

Độ dài, chuẩn, khoảng cách

Chuẩn trong không gian Euclid (norm)

Chuẩn trên không gian véc tơ V là một hàm $\|\cdot\| : V \to \mathbb{R}$, gán cho mỗi véc tơ x một số $\|x\|$ gọi là *độ dài*, sao cho với mọi $\lambda \in \mathbb{R}$ và $x, y \in V$ các tính chất sau thỏa mãn:

1. $\|\lambda x\| = |\lambda|.\|x\|$
2. $\|x + y\| \leq \|x\| + \|y\|$
3. $\|x\| \geq 0$ và $\|x\| = 0 \Leftrightarrow x = 0$

Độ dài, chuẩn, khoảng cách

Ví dụ về chuẩn của $\mathrm{x} \in \mathbb{R}^n$

- Chuẩn **Manhattan** hay chuẩn ℓ_1 trên $\mathbb{R}^n$ được định nghĩa bởi

$$\|\mathrm{x}\|_1 = \sum_{i=1}^{n} |x_i|$$

- Chuẩn **Euclid** hay chuẩn ℓ_2 trên $\mathbb{R}^n$ được xác định bởi

$$\|\mathrm{x}\|_2 = \sqrt{\mathrm{x} \cdot \mathrm{x}} = \sqrt{\sum_{i=1}^{n} x_i^2}$$

Chẳng hạn, $\mathrm{u} = [1, -2, 2]^T \implies \|\mathrm{u}\|_1 = 1 + 2 + 2 = 5$

$$\text{và} \quad \|\mathrm{u}\|_2 = \sqrt{1 + 4 + 4} = 3$$

Độ dài, chuẩn, khoảng cách

Véc tơ chuẩn hóa

Một véc tơ có độ dài bằng 1 được gọi là *véc tơ đơn vị*. Nếu một véc tơ $\mathrm{u} \neq 0$, ta có thể *chuẩn hóa* nó để thu được véc tơ đơn vị *cùng hướng*:

$$\mathrm{e}_{\mathrm{u}} = \frac{1}{\|\mathrm{u}\|}\mathrm{u}$$

Ví dụ: Véc tơ chuẩn hóa của $\mathrm{u} = [1, -2, 2]^T$ là

$$\mathrm{e}_{\mathrm{u}} = \frac{1}{\|\mathrm{u}\|}\mathrm{u} = \frac{1}{3}[1, -2, 2]^T = \left[\frac{1}{3}, -\frac{2}{3}, \frac{2}{3}\right]$$

Độ dài, chuẩn, khoảng cách

Khoảng cách (distance)

Cho V là một không gian véc tơ với tích vô hướng $\langle \cdot, \cdot \rangle$. Khi đó

$$d(\mathrm{x}, \mathrm{y}) = \|\mathrm{x} - \mathrm{y}\|$$

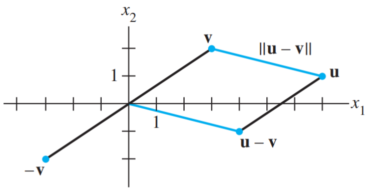
được gọi là *khoảng cách* giữa x và y trong V.

Ví dụ:

Tính khoảng cách giữa $\mathrm{u} = [7, 1]^T$ và $\mathrm{v} = [3, 2]^T$:

$$\mathrm{u} - \mathrm{v} = \begin{bmatrix} 7 \\ 1 \end{bmatrix} - \begin{bmatrix} 3 \\ 2 \end{bmatrix} = \begin{bmatrix} 4 \\ -1 \end{bmatrix}$$

$$\Rightarrow \|\mathrm{u} - \mathrm{v}\| = \sqrt{17}$$



Góc giữa hai véc tơ

Góc (angle)

Cho x và y là hai véc tơ khác 0 trong không gian véc tơ V với tích vô hướng $\langle\cdot,\cdot\rangle$. Khi đó, tồn tại duy nhất một góc $\omega \in [0, \pi]$ sao cho

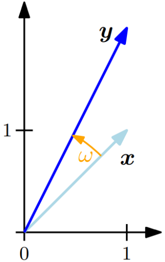
$$\cos\omega = \frac{\langle x, y\rangle}{\|x\|\,\|y\|}$$

Số ω được gọi là *góc* giữa hai véc tơ x và y.

Ví dụ: Góc giữa $u = [1, 1]^T$ và $v = [1, 2]^T$

$$\cos\omega = \frac{x\cdot y}{\|x\|\,\|y\|} = \frac{3}{\sqrt{10}}$$

$\Longrightarrow \omega \approx 0.32$ (rad) hay 18^o.



Trực giao (orthogonality)

Định nghĩa

Hai véc tơ x và y được gọi là *trực giao* nếu: $\langle x, y \rangle = 0$, ta viết $x \perp y$. Hơn nữa, nếu $\|x\| = \|y\| = 1$, tức các véc tơ là véc tơ đơn vị, thì x và y được gọi là *trực chuẩn*.

Định lý Pythagoras

Nếu hai véc tơ x và y trong không gian Euclid V là trực giao thì

$$\|x + y\|^2 = \|x\|^2 + \|y\|^2$$

Trực giao

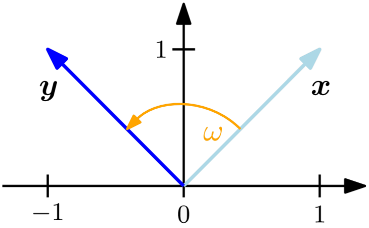
Ví dụ

Hai véc tơ $\mathrm{x} = [1, 1]^T$ và $\mathrm{y} = [-1, 1]^T$ trực giao với nhau:

$$\mathrm{x} \cdot \mathrm{y} = (1)(-1) + (1)(1) = 0$$

Tuy nhiên, nếu chọn tích vô hướng

$$\langle \mathrm{x}, \mathrm{y} \rangle = \mathrm{x}^T \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix} \mathrm{y}$$



thì x và y không trực giao với nhau, do $\langle \mathrm{x}, \mathrm{y} \rangle = 2(1)(-1) + (1)(1) = -1$.

Hệ trực giao (orthogonal set)

Định nghĩa

Một hệ véc tơ $\mathbf{u}_1, \ldots, \mathbf{u}_p$ trong không gian Euclid V được gọi *hệ trực giao* nếu từng đôi một véc tơ trong hệ là trực giao với nhau, tức là

$$\langle \mathbf{u}_i, \mathbf{u}_j \rangle = 0, \quad \text{với mọi } i \neq j$$

Ví dụ: Xét các véc tơ $\mathbf{u}_1 = \begin{bmatrix} 3 \\ 1 \\ 1 \end{bmatrix}$, $\mathbf{u}_2 = \begin{bmatrix} -1 \\ 2 \\ 1 \end{bmatrix}$, $\mathbf{u}_3 = \begin{bmatrix} 1 \\ 4 \\ -7 \end{bmatrix}$. Hệ véc tơ $\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\}$ là trực giao vì

$$\begin{aligned}
\mathbf{u}_1 \cdot \mathbf{u}_2 &= (3)(-1) + (1)(2) + (1)(1) = 0 \\
\mathbf{u}_1 \cdot \mathbf{u}_3 &= (3)(1) + (1)(4) + (1)(-7) = 0 \\
\mathbf{u}_2 \cdot \mathbf{u}_3 &= (-1)(1) + (2)(4) + (1)(-7) = 0
\end{aligned}$$

Cơ sở trực giao, trực chuẩn

Định nghĩa (orthogonal/orthonormal basis)

- Một hệ véc tơ $\mathcal{B} = \{b_1, \ldots, b_n\}$ được gọi là *cơ sở trực giao* của không gian Euclid V nếu $\mathcal{B}$ là một cơ sở của V, đồng thời là một hệ trực giao.
- Nếu các véc tơ trong cơ sở trực giao $\mathcal{B}$ là các véc tơ đơn vị (độ dài bằng 1) thì $\mathcal{B}$ được gọi là một *cơ sở trực chuẩn* của V.

Biểu diễn véc tơ trong cơ sở trực giao

Nếu $\mathcal{B}$ là một cơ sở trực giao của V, mọi véc tơ $x \in V$ đều có biểu diễn duy nhất:

$$x = \frac{\langle x, b_1\rangle}{\langle b_1, b_1\rangle} b_1 + \frac{\langle x, b_2\rangle}{\langle b_2, b_2\rangle} b_2 + \cdots + \frac{\langle x, b_n\rangle}{\langle b_n, b_n\rangle} b_n$$

Biểu diễn véc tơ trong cơ sở trực giao

Ví dụ

Trong $\mathbb{R}^3$, cho các véc tơ

$$b_1 = \frac{1}{3}\begin{bmatrix}1\\2\\2\end{bmatrix},\ b_2 = \frac{1}{3}\begin{bmatrix}2\\-2\\1\end{bmatrix},\ b_3 = \frac{1}{3}\begin{bmatrix}2\\1\\-2\end{bmatrix},\ x = \begin{bmatrix}6\\1\\8\end{bmatrix}$$

Dễ dàng kiểm tra hệ $\mathcal{B} = \{b_1, b_2, b_3\}$ là một cơ sở trực chuẩn của $\mathbb{R}^3$. Biểu diễn của x trong cơ sở $\mathcal{B}$ như sau:

$$\begin{aligned} x &= \frac{x \cdot b_1}{b_1 \cdot b_1} b_1 + \frac{x \cdot b_2}{b_2 \cdot b_2} b_2 + \frac{x \cdot b_3}{b_3 \cdot b_3} b_3 \\ &= \frac{8}{1} b_1 + \frac{6}{1} b_2 + \frac{-1}{1} b_3 = 8b_1 + 6b_2 - b_3 \end{aligned}$$

Trực giao hóa Gram-Schmidt

The Gram–Schmidt Process

Given a basis $\{\mathbf{x}_1, \ldots, \mathbf{x}_p\}$ for a nonzero subspace W of $\mathbb{R}^n$, define

$$\begin{aligned}
\mathbf{v}_1 &= \mathbf{x}_1 \\
\mathbf{v}_2 &= \mathbf{x}_2 - \frac{\mathbf{x}_2 \cdot \mathbf{v}_1}{\mathbf{v}_1 \cdot \mathbf{v}_1}\mathbf{v}_1 \\
\mathbf{v}_3 &= \mathbf{x}_3 - \frac{\mathbf{x}_3 \cdot \mathbf{v}_1}{\mathbf{v}_1 \cdot \mathbf{v}_1}\mathbf{v}_1 - \frac{\mathbf{x}_3 \cdot \mathbf{v}_2}{\mathbf{v}_2 \cdot \mathbf{v}_2}\mathbf{v}_2 \\
&\vdots \\
\mathbf{v}_p &= \mathbf{x}_p - \frac{\mathbf{x}_p \cdot \mathbf{v}_1}{\mathbf{v}_1 \cdot \mathbf{v}_1}\mathbf{v}_1 - \frac{\mathbf{x}_p \cdot \mathbf{v}_2}{\mathbf{v}_2 \cdot \mathbf{v}_2}\mathbf{v}_2 - \cdots - \frac{\mathbf{x}_p \cdot \mathbf{v}_{p-1}}{\mathbf{v}_{p-1} \cdot \mathbf{v}_{p-1}}\mathbf{v}_{p-1}
\end{aligned}$$

Then $\{\mathbf{v}_1, \ldots, \mathbf{v}_p\}$ is an orthogonal basis for W. In addition

$$\text{Span}\,\{\mathbf{v}_1, \ldots, \mathbf{v}_k\} = \text{Span}\,\{\mathbf{x}_1, \ldots, \mathbf{x}_k\} \qquad \text{for } 1 \le k \le p$$

Trực giao hóa Gram-Schmidt

Ví dụ

Cho các véc tơ $x_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$, $x_2 = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}$ và $x_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$. Hãy xây dựng một cơ sở trực giao cho $W = \text{span}(x_1, x_2, x_3)$.

- Bước 1: Đặt $v_1 = x_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$
- Bước 2: Tìm v_2 dưới dạng

$$v_2 = x_2 - \frac{x_2 \cdot v_1}{v_1 \cdot v_1} v_1 = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} - \frac{2}{3} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} -2/3 \\ 1/3 \\ 1/3 \end{bmatrix}$$

Trực giao hóa Gram-Schmidt

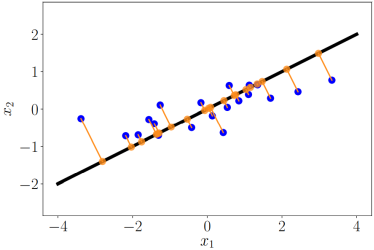
- Bước 3: Tìm v_3 dưới dạng

$$v_3 = x_3 - \frac{x_3 \cdot v_1}{v_1 \cdot v_1} v_1 - \frac{x_3 \cdot v_2}{v_2 \cdot v_2} v_2 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} - \frac{1}{3} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} - \frac{1}{2} \begin{bmatrix} -2/3 \\ 1/3 \\ 1/3 \end{bmatrix} = \begin{bmatrix} 0 \\ -1/2 \\ 1/2 \end{bmatrix}$$

- Kết luận: $\{v_1, v_2, v_3\}$ là cơ sở trực giao của W.

Phép chiếu trực giao

- Phép chiếu trực giao là lớp các phép biến đổi tuyến tính quan trọng trong đồ họa, lý thuyết mã hóa, thống kê và học máy.
- Trong học máy, dữ liệu có số chiều lớn, trong đó chỉ có một số chiều nắm giữ lượng lớn thông tin.
- Phép chiếu dữ liệu gốc với số chiều lớn lên không gian các thuộc tính có số chiều thấp hơn sẽ hỗ trợ cho việc học dữ liệu và trích xuất thông tin phù hợp.



Phép chiếu trực giao

Định nghĩa (orthogonal projection)

Cho V là không gian Euclid và W là một không gian con của V. Mọi véc tơ $y \in V$ đều có biểu diễn duy nhất dưới dạng:

$$y = \widehat{y} + z$$

ở đó $\widehat{y} \in W$ và $z \perp W$. Véc tơ $\widehat{y}$ gọi là *hình chiếu trực giao* của y lên W. Ánh xạ $\mathrm{proj}_W(y) = \widehat{y}$ gọi là *phép chiếu trực giao* của V lên W.

Phép chiếu với cơ sở trực chuẩn

Nếu $\{u_1, \ldots, u_k\}$ là một cơ sở trực chuẩn của W thì

$$\widehat{y} = \langle y, u_1 \rangle u_1 + \cdots + \langle y, u_k \rangle u_k$$

Phép chiếu trực giao với cơ sở trực chuẩn

Ví dụ

Trong $\mathbb{R}^3$ cho các véc tơ:

$$\mathrm{u}_1 = \frac{1}{3}\begin{bmatrix} 2 \\ -1 \\ 2 \end{bmatrix}, \ \mathrm{u}_2 = \frac{1}{3}\begin{bmatrix} -1 \\ 2 \\ 2 \end{bmatrix}, \ \mathrm{y} = \begin{bmatrix} 2 \\ 1 \\ 3 \end{bmatrix}$$

Cho W là không gian sinh bởi $\{\mathrm{u}_1, \mathrm{u}_2\}$. Rõ ràng $\{\mathrm{u}_1, \mathrm{u}_2\}$ là một cơ sở trực chuẩn. Do đó hình chiếu trực giao của y trên W có dạng:

$$\mathrm{proj}_W(\mathrm{y}) = (\mathrm{y}\cdot \mathrm{u}_1)\mathrm{u}_1 + (\mathrm{y}\cdot \mathrm{u}_2)\mathrm{u}_2 = \begin{bmatrix} 2 \\ -1 \\ 2 \end{bmatrix} + \frac{2}{3}\begin{bmatrix} -1 \\ 2 \\ 2 \end{bmatrix} = \begin{bmatrix} 4/3 \\ 1/3 \\ 10/3 \end{bmatrix}$$

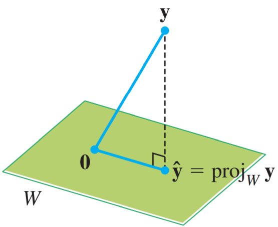
Xấp xỉ tốt nhất

Định lý

Cho V là không gian Euclid, W là không gian con của V và $\mathrm{y} \in V$ tùy ý. Kí hiệu $\widehat{\mathrm{y}}$ là hình chiếu trực giao của y lên W. Khi đó, $\widehat{\mathrm{y}}$ là véc tơ nằm trong W gần với y nhất, theo nghĩa

$$\|\mathrm{y} - \widehat{\mathrm{y}}\| < \|\mathrm{y} - \mathrm{v}\|$$

với mọi véc tơ $\mathrm{v} \in W$ khác $\widehat{\mathrm{y}}$.



Thực hành với Python

NumPy cung cấp các hàm để tính toán trong không gian Euclid:

- **numpy.inner**(a, b): trả về tích vô hướng của hai véc tơ a và b.
- **numpy.linalg.norm**(x): trả về chuẩn Euclid của véc tơ x.
- **numpy.linalg.qr**(A): trả về hai ma trận Q và R trong phân tích ma trận $A = QR$.

Phân tích QR

Định lý

Nếu một ma trận A cỡ $m \times n$ với các cột độc lập tuyến tính, A có thể được phân tích thành

$$A = QR$$

ở đó Q là ma trận $m \times n$ có các cột tạo thành một cơ sở trực chuẩn của không gian cột $\mathrm{Col}\, A$ và R là một ma trận tam giác trên cỡ $n \times n$ khả nghịch có các phần tử đường chéo là các số dương.

Chú ý: Không gian cột $\mathrm{Col}\, A$ là không gian sinh bởi các véc tơ cột của ma trận A. Chẳng hạn:

$$A = \begin{bmatrix} 1 & 3 & 9 \\ 0 & 1 & 4 \\ 2 & 1 & -2 \end{bmatrix} \Rightarrow \mathrm{Col}\, A = \mathrm{span}\left\{ \begin{bmatrix} 1 \\ 0 \\ 2 \end{bmatrix}, \begin{bmatrix} 3 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 9 \\ 4 \\ -2 \end{bmatrix} \right\}$$

Phân tích QR

Tìm phân tích QR của ma trận

1. Các cột của A tạo thành một cơ sở của $\mathrm{Col}\, A$: $\{x_1, \ldots, x_n\}$.
2. Xây dựng một cơ sở trực chuẩn (Gram-Schmidt): $\{u_1, \ldots, u_n\}$.
3. Đặt $Q = [u_1\ u_2\ \cdots\ u_n]$.
4. Xác định R: Với mỗi $k = 1, 2, \ldots, n$; $x_k \in \mathrm{span}\{u_1, \ldots, u_k\}$ nên

$$x_k = r_{1k}u_1 + \cdots + r_{kk}u_k + 0u_{k+1} + \cdots + 0u_n = Qr_k$$

trong đó $r_k = [r_{1k}\ \cdots\ r_{kk}\ 0\ \cdots\ 0]^T$. Có thể giả sử $r_{kk} > 0$, nếu không chỉ cần đổi dấu u_k. Đặt

$$R = [r_1\ r_2\ \cdots\ r_n]$$

Phân tích QR

Ví dụ

Tìm phân tích QR của ma trận $A = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{bmatrix}$.

Ta thực hiện như sau:

1. Cơ sở của không gian $\mathrm{Col}\, A$ là: $x_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$, $x_2 = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}$ và $x_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$.
2. Sử dụng trực giao hóa Gram-Schmidt, thu được cơ sở trực chuẩn:

$$u_1 = \frac{1}{\sqrt{3}} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, u_2 = \frac{1}{\sqrt{6}} \begin{bmatrix} -2 \\ 1 \\ 1 \end{bmatrix}, u_3 = \frac{1}{\sqrt{2}} \begin{bmatrix} 0 \\ -1 \\ 1 \end{bmatrix}$$

Phân tích QR

3. Đặt $Q = \begin{bmatrix} 1/\sqrt{3} & -2/\sqrt{6} & 0 \\ 1/\sqrt{3} & 1/\sqrt{6} & -1/\sqrt{2} \\ 1/\sqrt{3} & 1/\sqrt{6} & 1/\sqrt{2} \end{bmatrix}$

4. Từ quá trình trực giao hóa Gram-Schmidt, ta xác định được R:

$$R = \begin{bmatrix} \sqrt{3} & 2/\sqrt{3} & \sqrt{3}/3 \\ 0 & \sqrt{6}/3 & \sqrt{6}/6 \\ 0 & 0 & 1/\sqrt{2} \end{bmatrix}$$

Cuối cùng, ta thu được phân tích QR:

$$\begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{bmatrix} = \begin{bmatrix} 1/\sqrt{3} & -2/\sqrt{6} & 0 \\ 1/\sqrt{3} & 1/\sqrt{6} & -1/\sqrt{2} \\ 1/\sqrt{3} & 1/\sqrt{6} & 1/\sqrt{2} \end{bmatrix} \begin{bmatrix} \sqrt{3} & 2/\sqrt{3} & \sqrt{3}/3 \\ 0 & \sqrt{6}/3 & \sqrt{6}/6 \\ 0 & 0 & 1/\sqrt{2} \end{bmatrix}$$

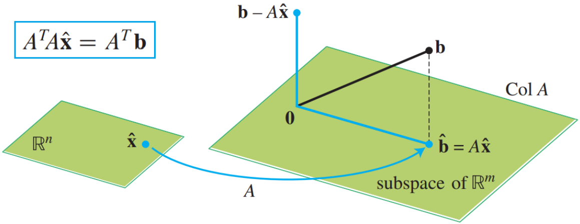
Bài toán bình phương cực tiểu

Định nghĩa

Cho $A \in \mathbb{R}^{m \times n}$ và $b \in \mathbb{R}^m$, *nghiệm bình phương cực tiểu* của $Ax = b$ là véc tơ $\widehat{x} \in \mathbb{R}^n$ sao cho

$$\|b - A\widehat{x}\| \leq \|b - Ax\|$$

với mọi $x \in \mathbb{R}^n$.



Bài toán bình phương cực tiểu

Trong một số trường hợp, phương trình tìm nghiệm bình phương cực tiểu có *điều kiện xấu*; tức là sai lệch nhỏ của phần tử trong A cũng dẫn đến sai số lớn trong nghiệm $\widehat{x}$. Phương pháp phân tích QR có thể được sử dụng hiệu quả hơn để tìm nghiệm bình phương tối thiểu.

Sử dụng phân tích QR

Cho $A \in \mathbb{R}^{m \times n}$ có các cột độc lập tuyến tính. Giả sử phân tích QR của A có dạng $A = QR$. Khi đó với mỗi $b \in \mathbb{R}^m$, nghiệm bình phương cực tiểu của phương trình $Ax = b$ được cho bởi

$$\widehat{x} = R^{-1}Q^T b$$

Thực hành với Python

Bài toán

Cho ma trận: $A = \begin{bmatrix} 1 & 3 & -7 \\ 2 & 1 & 1 \\ -2 & 7 & 11 \\ 9 & 11 & 3 \end{bmatrix}$, $b = \begin{bmatrix} 3 \\ 1 \\ -1 \\ 3 \end{bmatrix}$

1. Thực hiện phân tích QR cho ma trận A.
2. Tìm nghiệm bình phương cực tiểu của $Ax = b$ bằng phân tích QR.
3. Tìm một cơ sở trực chuẩn của không gian cột của A.

Thực hành với Python

```
import numpy as np
a=np.array([[1,3,-7],[2,1,1],[-2,7,11],[9,11,3]])
q,r=np.linalg.qr(a)
print("Ma trận A, Q, R theo thứ tự:")
print(a,q,r,sep="\n")
```

```
Ma trận A, Q, R theo thứ tự:
[[ 1  3 -7]
 [ 2  1  1]
 [-2  7 11]
 [ 9 11  3]]
[[-0.10540926  0.21081851  0.9486833 ]
 [-0.21081851 -0.10540926 -0.21081851]
 [ 0.21081851  0.9486833  -0.21081851]
 [-0.9486833   0.21081851 -0.10540926]]
[[-9.48683298 -9.48683298  0.        ]
 [ 0.          9.48683298  9.48683298]
 [ 0.          0.         -9.48683298]]
```

Thực hành với Python

```
b=np.array([3,1,-1,3])
rinv=np.linalg.inv(r)
qt=q.T
print("Nghiệm:",np.dot(np.dot(rinv,qt),b))
```

```
Nghiệm: [ 0.08888889  0.28888889 -0.26666667]
```

```
print("Cơ sở trực chuẩn:")
for i in range(q.shape[1]):
    print(q[:,i])
```

```
Cơ sở trực chuẩn:
[-0.10540926 -0.21081851  0.21081851 -0.9486833 ]
[ 0.21081851 -0.10540926  0.9486833   0.21081851]
[ 0.9486833  -0.21081851 -0.21081851 -0.10540926]
```

Tài liệu tham khảo

1. Charu C. Aggarwal; Linear Algebra and Optimization for Machine Learning, Springer, 2020.
2. Stephen Boyd, Lieven Vandenberghe; Introduction to Applied Linear Algebra: Vectors, Matrices, and Least Squares, Cambridge University Press, 2018.
3. David C. Lay, Steven R. Lay, Judi J. McDonald; Linear Algebra and Its Applications, Fifth edition, Pearson, 2016
4. Tom Lyche; Numerical Linear Algebra and Matrix Factorizations, Springer, 2020.
5. Gilbert Strang; Linear Algebra and Learning from Data, Wellesley- Cambridge Press, 2019.