Cơ sở trực giao – Không gian con trực giao

Nguyễn Hoàng Thạch
nhthach@math.ac.vn

Tóm tắt

Tóm tắt

Cơ sở trực giao

Giả sử V là một không gian Euclid.

Định nghĩa

Một tập hợp S gồm các vector của V được gọi là một hệ trực giao *nếu các vector trong S đôi một vuông góc với nhau. Một hệ trực giao S gồm toàn các vector đơn vị được gọi là một* hệ trực chuẩn.
Một cơ sở S là một cơ sở trực giao *(tương ứng* cơ sở trực chuẩn*) nếu nó là một hệ trực giao (tương ứng trực chuẩn).*

Chú ý: Nếu $S = \{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k\}$ thì S là một hệ trực chuẩn nếu và chỉ nếu

$$\langle \mathbf{v}_i, \mathbf{v}_j \rangle = \begin{cases} 0 \text{ if } i \neq j \\ 1 \text{ if } i = j \end{cases}$$

Thí dụ:

- Trong $\mathbb{R}^n$ với tích vô hướng thông thường, cơ sở chính tắc là một cơ sở trực chuẩn.
- Trong $\mathbb{R}^3$, $S = \{(\cos\theta, \sin\theta, 0), (-\sin\theta, \cos\theta, 0)\}$ là một hệ trực chuẩn.

Hệ trực giao và độc lập tuyến tính

Định lý

Nếu $S = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\}$ *là một hệ trực giao gồm các vector* khác không *thì* S *độc lập tuyến tính.*

Chứng minh:
Giả sử $c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_k\mathbf{v}_k = \mathbf{0}$ là một ràng buộc tuyến tính của S. Lấy tích vô hướng của cả hai vế với $\mathbf{v}_i$ ta được:

$$c_1\langle \mathbf{v}_1, \mathbf{v}_i\rangle + c_2\langle \mathbf{v}_2, \mathbf{v}_i\rangle + \cdots + \langle \mathbf{v}_n, \mathbf{v}_i\rangle = 0\,.$$

Vì S là hệ trực giao nên vế trái bằng $c_i\langle \mathbf{v}_i, \mathbf{v}_i\rangle = c_i\|\mathbf{v}_i\|^2$, suy ra $c_i\|\mathbf{v}_i\|^2 = 0$.
Vì $\mathbf{v}_i \neq \mathbf{0}$ nên $c_i = 0$.

Hệ trực giao và độc lập tuyến tính

Hệ quả

Nếu $\dim(V) = n$ *thì mọi hệ trực giao gồm n vector khác không của V là một cơ sở.*

Thí dụ:
Trong $\mathbb{R}^4$, xét
$S = \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \mathbf{v}_4\} = \{(2, 3, 2, -3), (1, 0, 0, 1), (-1, 0, 2, 1), (-1, 2, -1, 1)\}$.
Có thể kiểm tra rằng các vector của S đôi một vuông góc với nhau:

$$\mathbf{v}_1 \cdot \mathbf{v}_2 = \mathbf{v}_1 \cdot \mathbf{v}_3 = \mathbf{v}_1 \cdot \mathbf{v}_4 = \mathbf{v}_2 \cdot \mathbf{v}_3 = \mathbf{v}_2 \cdot \mathbf{v}_4 = \mathbf{v}_3 \cdot \mathbf{v}_4 = 0\,.$$

Suy ra S là một hệ trực giao, do đó là một cơ sở của $\mathbb{R}^4$.

Tọa độ trong cơ sở trực chuẩn

Định lý

Nếu $B = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ *là một cơ sở trực chuẩn của* V *thì tọa độ của một vector* $\mathbf{u}$ *trong cơ sở* B *là*

$$[\mathbf{u}]_B = [\langle \mathbf{u}, \mathbf{v}_1 \rangle \ \langle \mathbf{u}, \mathbf{v}_2 \rangle \ \ldots \langle \mathbf{u}, \mathbf{v}_n \rangle]^T .$$

Chứng minh:
Giả sử $\mathbf{u} = c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_n\mathbf{v}_n$. Khi đó $\langle \mathbf{u}, \mathbf{v}_i \rangle = c_i \langle \mathbf{v}_i, \mathbf{v}_i \rangle = c_i$.

Thí dụ: Trong $\mathbb{R}^3$, cho cơ sở trực chuẩn
$B = \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\} = \{(3/5, 4/5, 0), (-4/5, 3/5, 0), (0, 0, 1)\}$ và $\mathbf{u} = (5, -5, 2)$.
Ta có $\mathbf{u} \cdot \mathbf{v}_1 = -1$, $\mathbf{u} \cdot \mathbf{v}_2 = -7$, $\mathbf{u} \cdot \mathbf{v}_3 = 2$. Từ đó $[\mathbf{u}]_B = [-1 \ 7 \ 2]^T$.

Tóm tắt

Quá trình trực giao hóa Gram-Schmidt

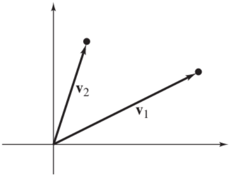
Cho $B = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ là một cơ sở của không gian Euclid V. *Quá trình trực giao hóa Gram-Schmidt* biến B thành $B' = \{\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_n\}$ như sau:

$$
\begin{aligned}
\mathbf{w}_1 &= \mathbf{v}_1 \\
\mathbf{w}_2 &= \mathbf{v}_2 - \frac{\langle \mathbf{v}_2, \mathbf{w}_1 \rangle}{\|\mathbf{w}_1\|^2}\mathbf{w}_1 \\
\mathbf{w}_3 &= \mathbf{v}_3 - \frac{\langle \mathbf{v}_3, \mathbf{w}_1 \rangle}{\|\mathbf{w}_1\|^2}\mathbf{w}_1 - \frac{\langle \mathbf{v}_3, \mathbf{w}_2 \rangle}{\|\mathbf{w}_2\|^2}\mathbf{w}_2 \\
\ldots \\
\mathbf{w}_n &= \mathbf{v}_n - \frac{\langle \mathbf{v}_n, \mathbf{w}_1 \rangle}{\|\mathbf{w}_1\|^2}\mathbf{w}_1 - \frac{\langle \mathbf{v}_n, \mathbf{w}_2 \rangle}{\|\mathbf{w}_2\|^2}\mathbf{w}_2 - \cdots - \frac{\langle \mathbf{v}_n, \mathbf{w}_{n-1} \rangle}{\|\mathbf{w}_{n-1}\|^2}\mathbf{w}_{n-1}
\end{aligned}
$$

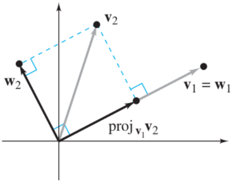
Quá trình trực giao hóa Gram-Schmidt

Định lý

1. *Tập hợp B' là một cơ sở trực giao của V.*
2. *Tập hợp $B'' = \{\mathbf{u}_1, \ldots, \mathbf{u}_n\}$, ở đó $\mathbf{u}_i = \frac{\mathbf{w}_i}{\|\mathbf{w}_i\|}$, là một cơ sở trực chuẩn của V.*
3. *Với mọi $k = 1, \ldots, n$,*
   $span\{\mathbf{v}_1, \ldots, \mathbf{v}_k\} = span\{\mathbf{w}_1, \ldots, \mathbf{w}_k\} = span\{\mathbf{u}_1, \ldots, \mathbf{u}_k\}$.



$\{\mathbf{v}_1, \mathbf{v}_2\}$ is a basis for R^2.



$\mathbf{w}_2 = \mathbf{v}_2 - \text{proj}_{\mathbf{v}_1}\mathbf{v}_2$
is orthogonal to $\mathbf{w}_1 = \mathbf{v}_1$.

Hình: Larson et al., p. 314

Quá trình trực giao hóa Gram-Schmidt

Thí dụ:

- Trong $\mathbb{R}^3$, xét cơ sở $B = \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\} = \{(1,1,0), (1,2,0), (0,1,2)\}$.

$$\begin{aligned}
\mathbf{w}_1 &= \mathbf{v}_1 = (1,1,0) \\
\mathbf{w}_2 &= \mathbf{v}_2 - \frac{\mathbf{v}_2 \cdot \mathbf{w}_1}{\|\mathbf{w}_1\|^2}\mathbf{w}_1 = \left(-\frac{1}{2}, \frac{1}{2}, 0\right) \\
\mathbf{w}_3 &= \mathbf{v}_3 - \frac{\mathbf{v}_3 \cdot \mathbf{w}_1}{\|\mathbf{w}_1\|^2}\mathbf{w}_1 - \frac{\mathbf{v}_3 \cdot \mathbf{w}_2}{\|\mathbf{w}_2\|^2}\mathbf{w}_2 = (0,0,2)
\end{aligned}$$

  Tập hợp $B' = \{\mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_3\}$ là một cơ sở trực giao của $\mathbb{R}^3$. Chuẩn hóa B' ta được một cơ sở trực chuẩn $B'' = \left\{\left(\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2}, 0\right), \left(-\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2}, 0\right), (0,0,1)\right\}$.

- Trong $\mathbb{R}^3$, xét $\mathbf{v}_1 = (0,1,0), \mathbf{v}_2 = (1,1,1)$. Trực giao hóa $\{\mathbf{v}_1, \mathbf{v}_2\}$ ta được $\mathbf{w}_1 = (0,1,0), \mathbf{w}_2 = (1,0,1)$. Chuẩn hóa hai vector này ta thu được một cơ sở trực chuẩn của *span* $\{\mathbf{v}_1, \mathbf{v}_2\}$: $\left\{(0,1,0), \left(\frac{\sqrt{2}}{2}, 0, \frac{\sqrt{2}}{2}\right)\right\}$.

Quá trình trực giao hóa Gram-Schmidt

Chú ý: Việc chuẩn hóa có thể được thực hiện ngay trong từng bước của quá trình trực giao hóa như sau

$$
\begin{aligned}
\mathbf{u}_1 &= \frac{\mathbf{w}_1}{\|\mathbf{w}_1\|} \text{ với } \mathbf{w}_1 = \mathbf{v}_1 \\
\mathbf{u}_2 &= \frac{\mathbf{w}_2}{\|\mathbf{w}_2\|} \text{ với } \mathbf{w}_2 = \mathbf{v}_2 - \langle \mathbf{v}_2, \mathbf{u}_1 \rangle \mathbf{u}_1 \\
\mathbf{u}_3 &= \frac{\mathbf{w}_3}{\|\mathbf{w}_3\|} \text{ với } \mathbf{w}_3 = \mathbf{v}_3 - \langle \mathbf{v}_3, \mathbf{u}_1 \rangle \mathbf{u}_1 - \langle \mathbf{v}_3, \mathbf{u}_2 \rangle \mathbf{u}_2 \\
\ldots & \\
\mathbf{u}_n &= \frac{\mathbf{w}_n}{\|\mathbf{w}_n\|} \text{ với } \mathbf{w}_n = \mathbf{v}_n - \langle \mathbf{v}_n, \mathbf{u}_1 \rangle \mathbf{u}_1 - \cdots - \langle \mathbf{v}_n, \mathbf{u}_{n-1} \rangle \mathbf{u}_{n-1}
\end{aligned}
$$

Quá trình trực giao hóa Gram-Schmidt

Thí dụ:

Tìm một cơ sở trực chuẩn của không gian nghiệm của hệ:

$$\begin{cases} x + y \qquad + 7t = 0 \\ 2x + y + 2z + 6t = 0 \end{cases}$$

- Nghiệm tổng quát của hệ là $(-2s + r, 2s - 8r, s, r)$ với $s, r \in \mathbb{R}$, từ đó một cơ sở của không gian nghiệm của hệ là $B = \{\mathbf{v}_1, \mathbf{v}_2\} = \{(-2, 2, 1, 0), (1, -8, 0, 1)\}$.
- Trực chuẩn hóa cơ sở B:

$$\begin{aligned} \mathbf{u}_1 &= \frac{\mathbf{v}_1}{\|\mathbf{v}_1\|} = \left(-\frac{2}{3}, \frac{2}{3}, \frac{1}{3}, 0\right) \\ \mathbf{w}_2 &= \mathbf{v}_2 - \langle \mathbf{v}_2, \mathbf{u}_1 \rangle \mathbf{u}_1 = (-3, -4, 2, 1) \\ \mathbf{u}_2 &= \frac{\mathbf{v}_2}{\|\mathbf{v}_2\|} = \left(-\frac{3}{\sqrt{30}}, -\frac{4}{\sqrt{30}}, \frac{2}{\sqrt{30}}, \frac{1}{\sqrt{30}}\right) \end{aligned}$$

Vậy $B' = \{\mathbf{u}_1, \mathbf{u}_2\}$ là một cơ sở trực chuẩn của không gian nghiệm của hệ đã cho.

Tóm tắt

Không gian con trực giao

Xét không gian Euclid V với $\dim(V) = n$.

Định nghĩa

Giả sử S là một không gian con của V và $\mathbf{v} \in V$. Ta nói rằng $\mathbf{v}$ vuông góc (hay trực giao) với S, và viết $\mathbf{v} \perp S$, nếu $\mathbf{v}$ vuông góc với mọi vector $\mathbf{w} \in S$.

Nhận xét: $\mathbf{v} \perp S$ nếu và chỉ nếu $\mathbf{v}$ vuông góc với mọi vector trong một cơ sở (hoặc hệ sinh) của S.

Định nghĩa

Hai không gian con S_1, S_2 của V được gọi là vuông góc (hay trực giao) với nhau, ký hiệu là $S_1 \perp S_2$, nếu với mọi $\mathbf{u} \in S_1$ và với mọi $\mathbf{v} \in S_2$, $\mathbf{u}$ và $\mathbf{v}$ vuông góc với nhau.

Nhận xét: Nếu $S_1 \perp S_2$ thì $S_1 \cap S_2 = \{\mathbf{0}\}$.

Thí dụ: Trong $\mathbb{R}^3$, xét $S_1 = span\{(1,0,0),(1,1,0)\}$, $\mathbf{u} = (-1,1,1)$, $S_2 = span\{\mathbf{u}\}$. Ta có $\mathbf{u} \perp S_1$ và $S_1 \perp S_2$.

Phần bù trực giao

Định nghĩa

Phần bù trực giao *của một không gian con* S, *ký hiệu* $S^{\perp}$, *được định nghĩa bởi*

$$S^{\perp} = \{\mathbf{u} \mid \mathbf{u} \perp S\} = \{\mathbf{u} \mid \mathbf{u} \text{ vuông góc với mọi } \mathbf{v} \in S\} .$$

Định lý

Nếu S *là một không gian con của* V *thì* $S^{\perp}$ *là một không gian con của* V.

Nhận xét: $V^{\perp} = \{\mathbf{0}\}$, $\{\mathbf{0}\}^{\perp} = V$.

Tìm phần bù trực giao của một không gian con

Thí dụ: Trong $\mathbb{R}^4$, tìm phần bù trực giao của $S = span\{\mathbf{v}_1, \mathbf{v}_2\} = span\{(1,2,1,0),(0,0,0,1)\}$.

Ta có $\mathbf{u} \in S^\perp$ nếu và chỉ nếu $\mathbf{u} \cdot \mathbf{v}_1 = \mathbf{u} \cdot \mathbf{v}_2 = 0$, tức là $A^T\mathbf{u} = \mathbf{0}$, với

$$A = [\mathbf{v}_1 \ \mathbf{v}_2] = \begin{pmatrix} 1 & 0 \\ 2 & 0 \\ 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Giải hệ phương trình tuyến tính, ta thu được $S^\perp = span\{\mathbf{u}_1, \mathbf{u}_2\}$ với $\mathbf{u}_1 = (-2,1,0,0)$, $\mathbf{u}_2 = (-1,0,1,0)$.

Nhận xét: Các vector $\mathbf{v}_1, \mathbf{v}_2, \mathbf{u}_1, \mathbf{u}_2$ tạo thành một cơ sở của $\mathbb{R}^4$, do đó mọi vector của $\mathbb{R}^4$ biểu diễn được một cách duy nhất dưới dạng tổng của một vector của S và một vector của $S^\perp$!

Tổng trực tiếp

Định nghĩa

Giả sử S_1, S_2 *là hai không gian con của* V*. Nếu mọi vector* $\mathbf{v} \in V$ *có thể viết được một cách* duy nhất *dưới dạng* $\mathbf{v} = \mathbf{v}_1 + \mathbf{v}_2$ *với* $\mathbf{v}_1 \in S_1$, $\mathbf{v}_2 \in S_2$ *thì ta nói rằng* V *là* tổng trực tiếp *của* S_1 *và* S_2 *và viết là* $V = S_1 \oplus S_2$*.*

Nhận xét:

- Nếu $V = S_1 \oplus S_2$ thì $S_1 \cap S_2 = \{\mathbf{0}\}$.
- Nếu $V = S_1 \oplus S_2$ thì hợp của một cơ sở của S_1 và một cơ sở của S_2 là một cơ sở của V.

Thí dụ:

- $\mathbb{R}^3 = span\{(1,0,1),(1,1,0)\} \oplus span\{(-1,1,1)\}$.
- $\mathbb{R}^4 = span\{(1,2,1,0),(0,0,0,1)\} \oplus span\{(-2,1,0,0),(-1,0,1,0)\}$.

Phép chiếu vuông góc lên một không gian con

Định lý

Nếu S là một không gian con của V thì:

1. $\dim(S) + \dim(S^{\perp}) = \dim(V)$.
2. $V = S \oplus S^{\perp}$.
3. $(S^{\perp})^{\perp} = S$.

Định nghĩa

Cho S là một không gian con của V và $\mathbf{v} \in V$. Giả sử $\mathbf{v} = \mathbf{v}_1 + \mathbf{v}_2$, với $\mathbf{v}_1 \in S, \mathbf{v}_2 \in S^{\perp}$, là biểu diễn duy nhất của $\mathbf{v}$ dưới dạng tổng của một vector của S và một vector của $S^{\perp}$. Ta nói rằng $\mathbf{v}_1$ là hình chiếu vuông góc *của $\mathbf{v}$ lên không gian con S và ký hiệu $\pi_S(\mathbf{v}) = \mathbf{v}_1$.*

Tìm hình chiếu vuông góc lên một không gian con

Định lý

Nếu $\{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_k\}$ *là một cơ sở trực chuẩn của không gian con* S *thì với mọi* $\mathbf{v} \in V$*, hình chiếu vuông góc của* $\mathbf{v}$ *lên* S *được cho bởi công thức*

$$\pi_S(\mathbf{v}) = \langle \mathbf{v}, \mathbf{u}_1 \rangle \mathbf{u}_1 + \langle \mathbf{v}, \mathbf{u}_2 \rangle \mathbf{u}_2 + \cdots + \langle \mathbf{v}, \mathbf{u}_k \rangle \mathbf{u}_k \,.$$

Thí dụ: Trong $\mathbb{R}^3$, xét $S = span\,\{\mathbf{w}_1, \mathbf{w}_2\} = span\,\{(0,3,1),(2,0,0)\}$ và $\mathbf{v} = (1,1,3)$. Tìm $\pi_S(\mathbf{v})$.

Áp dụng Gram-Schmidt, ta tìm được một cơ sở trực chuẩn của S từ hệ sinh $\{\mathbf{w}_1, \mathbf{w}_2\}$:

$$\{\mathbf{u}_1, \mathbf{u}_2\} = \left\{ \frac{1}{\sqrt{10}}\mathbf{u}_1, \frac{1}{2}\mathbf{u}_2 \right\} .$$

Từ đó

$$\pi_S(\mathbf{v}) = (\mathbf{v} \cdot \mathbf{u}_1)\mathbf{u}_1 + (\mathbf{v} \cdot \mathbf{u}_2)\mathbf{u}_2 = (1, \frac{9}{5}, \frac{3}{5}) \,.$$

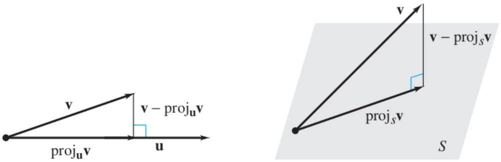
"Khoảng cách" đến một không gian con

Định lý

Cho S là một không gian con của V và $\mathbf{v} \in V$. Khi đó, với mọi $\mathbf{u} \in S$:

$$\|\mathbf{v} - \mathbf{u}\| \geq \|\mathbf{v} - \pi_S(\mathbf{v})\| .$$

Chú ý: Đại lượng $\|\mathbf{v} - \pi_S(\mathbf{v})\|$ có thể được coi là *khoảng cách* từ vector **v** đến không gian con S.



Hình: Larson et al., p. 325

Tóm tắt

Các không gian con cơ bản của một ma trận

Cho A là một ma trận $m \times n$.

Định nghĩa

Các không gian con cơ bản *của ma trận A là các không gian sau:*

1. *Không gian nghiệm $N(A)$ của A;*
2. *Không gian cột $R(A)$ của A;*
3. *Không gian nghiệm $N(A^T)$ của A^T;*
4. *Không gian cột $R(A^T)$ của A^T.*

Nhận xét: Các không gian $N(A), R(A^T)$ là các không gian con của $\mathbb{R}^n$; các không gian $R(A), N(A^T)$ là các không gian con của $\mathbb{R}^m$.

Thí dụ: $A = \begin{pmatrix} 1 & 2 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$

- $R(A) = span\{(1,0,0,0),(0,1,0,0)\}$;
- $R(A^T) = span\{(1,2,0),(0,0,1)\}$;
- $N(A) = span\{(-2,1,0)\}$;
- $N(A^T) = span\{(0,0,1,0),(0,0,0,1)\}$.

Các không gian con cơ bản của một ma trận

Định lý

1. *Các không gian $R(A)$ và $N(A^T)$ là các không gian con trực giao của $\mathbb{R}^m$.*
2. *Các không gian $N(A)$ và $R(A^T)$ là các không gian con trực giao của $\mathbb{R}^n$.*
3. $R(A) \oplus N(A^T) = \mathbb{R}^m$.
4. $N(A) \oplus R(A^T) = \mathbb{R}^n$.

Nhận xét: Một cách phát biểu khác của định lý trên là $R(A)$ và $N(A^T)$ là phần bù trực giao của nhau trong $\mathbb{R}^m$; $N(A)$ và $R(A^T)$ là phần bù trực giao của nhau trong $\mathbb{R}^n$.

Tóm tắt

Bài toán bình phương tối thiểu

Bài toán: Có n quan sát, với dữ liệu đầu vào là x_i và số đo đầu ra tương ứng là y_i. Ta muốn xấp xỉ mối liên hệ giữa đầu vào và đầu ra bởi một đa thức $p(x) = c_1 + c_2 x + c_3 x^2 + \cdots + c_d x^{d-1}$. Với những hệ số c_j như thế nào thì xấp xỉ là "tốt"?

Ta đánh giá chất lượng của xấp xỉ thông qua vector *sai số*:

$$\mathbf{e}_p = [y_1 - p(x_1), \ldots, y_n - p(x_n)]^T$$

Sai số càng "nhỏ" thì xấp xỉ càng "tốt". Như vậy, ta cần giải bài toán

$$\min_{c_1, \ldots, c_d} \|\mathbf{e}_p\| .$$

Bài toán bình phương tối thiểu

Đặt $A = \begin{pmatrix} 1 & x_1 & x_1^2 & \dots & x_1^{d-1} \\ 1 & x_2 & x_2^2 & \dots & x_2^{d-1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_n & x_n^2 & \dots & x_n^{d-1} \end{pmatrix}$, $\mathbf{y} = [y_1 \ y_2 \ \dots \ y_n]^T$ và $\mathbf{u} = [c_1 \ c_2 \ \dots \ c_d]^T$. Ta có:

$$\mathbf{e}_p = \mathbf{y} - A\mathbf{u}\,.$$

Bài toán trở thành:

$$\min_{\mathbf{u}\in\mathbb{R}^d} \|\mathbf{y} - A\mathbf{u}\|\,.$$

Vì $\mathbf{y}$ cố định và $A\mathbf{u} \in R(A)$ nên $\|\mathbf{y} - A\mathbf{u}\|$ nhỏ nhất khi $A\mathbf{u}$ là hình chiếu vuông góc của $\mathbf{y}$ lên $R(A)$. Điều này tương đương với:

$$\begin{aligned} & \mathbf{y} - A\mathbf{u} \in R(A)^\perp = N(A^T) \\ \iff \quad & A^T(\mathbf{y} - A\mathbf{u}) = \mathbf{0} \\ \iff \quad & (A^TA)\mathbf{u} = A^T\mathbf{y} \end{aligned}$$