

Questões sobre projeção - ALN

Beatriz Lúcia

May 2025

1

Considere a decomposição vetorial associada à equação $\mathbf{Ax} = \mathbf{b}$, na qual $\hat{\mathbf{x}}$ é a solução dos mínimos quadrados, \mathbf{p} é a projeção de \mathbf{b} na imagem de A , e $\mathbf{e} = \mathbf{b} - \mathbf{p}$ é o vetor erro:

$$\mathbf{b} = \mathbf{p} + \mathbf{e} \tag{1}$$

Indique a qual dos quatro subespaços fundamentais de uma matriz A pertencem os seguintes vetores:

- \mathbf{e}
- \mathbf{p}
- $\hat{\mathbf{x}}$

Além disso, determine o núcleo de A , isto é, $N(A)$.

Solução.

Sabemos das propriedades da decomposição de mínimos quadrados que:

- O vetor erro \mathbf{e} é ortogonal às colunas de A , ou seja, $\mathbf{e} \in N(A^\top)$;
- A projeção \mathbf{p} está na imagem de A , ou seja, $\mathbf{p} \in C(A)$;
- A solução $\hat{\mathbf{x}}$ está no domínio da aplicação de A , portanto $\hat{\mathbf{x}} \in C(A^\top)$, já que $C(A^\top)$ é a imagem de A^\top e representa o espaço das soluções dos mínimos quadrados;
- Se A tem posto máximo (colunas linearmente independentes), então o núcleo de A é trivial: $N(A) = \{\mathbf{0}\}$.

Portanto:

$$\begin{aligned}\mathbf{e} &\in N(A^\top) \\ \mathbf{p} &\in C(A) \\ \hat{\mathbf{x}} &\in C(A^\top) \\ N(A) &= \{\mathbf{0}\}\end{aligned}$$

Projeção no Problema dos Mínimos Quadrados

Dado um sistema linear superdeterminado (isto é, com mais equações do que incógnitas):

$$A\mathbf{x} = \mathbf{b}, \quad A \in R^{m \times n}, \quad \mathbf{b} \in R^m, \quad (m > n)$$

na maioria dos casos, não existe \mathbf{x} tal que $A\mathbf{x} = \mathbf{b}$, pois $\mathbf{b} \notin C(A)$. Assim, buscamos uma solução que minimize a norma do erro:

$$\hat{\mathbf{x}} = \arg \min_{\mathbf{x} \in R^n} \|A\mathbf{x} - \mathbf{b}\|^2$$

Essa solução $\hat{\mathbf{x}}$ gera um vetor $\mathbf{p} = A\hat{\mathbf{x}}$ que é a **projeção ortogonal** de \mathbf{b} no subespaço coluna de A , isto é:

$$\mathbf{b} = \mathbf{p} + \mathbf{e}, \quad \text{com } \mathbf{p} \in C(A), \quad \mathbf{e} \in N(A^\top)$$

Como \mathbf{e} é ortogonal ao espaço gerado pelas colunas de A , temos:

$$A^\top \mathbf{e} = A^\top (\mathbf{b} - A\hat{\mathbf{x}}) = 0 \quad \Rightarrow \quad A^\top A\hat{\mathbf{x}} = A^\top \mathbf{b}$$

Essa é a chamada **equação normal** do problema de mínimos quadrados. Supondo que $A^\top A$ seja inversível, obtemos a solução explícita:

$$\hat{\mathbf{x}} = (A^\top A)^{-1} A^\top \mathbf{b}$$

Portanto, a projeção de \mathbf{b} sobre o espaço das colunas de A é:

$$\mathbf{p} = A\hat{\mathbf{x}} = A(A^\top A)^{-1} A^\top \mathbf{b}$$

A matriz

$$P = A(A^\top A)^{-1} A^\top$$

é chamada de **matriz de projeção ortogonal** sobre o espaço imagem de A , ou seja, $C(A)$.

Resumo:

- $\mathbf{p} = A\hat{\mathbf{x}} \in C(A)$ é a projeção de \mathbf{b} ;
- $\hat{\mathbf{x}} \in C(A^\top)$ é a solução dos mínimos quadrados;
- $\mathbf{e} = \mathbf{b} - \mathbf{p} \in N(A^\top)$ é o vetor erro, ortogonal às colunas de A .

2

Derive a Alternativa de Fredholm: Se o sistema $A\mathbf{x} = \mathbf{b}$ não possui solução, então existe um vetor \mathbf{y} tal que:

$$A^\top \mathbf{y} = 0 \quad \text{e} \quad \mathbf{y}^\top \mathbf{b} = 1.$$

Dica: Se $\mathbf{b} \notin C(A)$, então \mathbf{b} não é ortogonal a $N(A^\top)$.

Solução:

Suponha que o sistema $A\mathbf{x} = \mathbf{b}$ não tenha solução. Isso significa que $\mathbf{b} \notin C(A)$, ou seja, \mathbf{b} não pertence ao espaço coluna de A .

Pelo Teorema Fundamental da Álgebra Linear, temos:

$$C(A)^\perp = N(A^\top).$$

Logo, como $\mathbf{b} \notin C(A)$, ela não pode ser ortogonal a todos os vetores em $N(A^\top)$. Portanto, existe algum vetor $\mathbf{p} \in N(A^\top)$ tal que:

$$\mathbf{p}^\top \mathbf{b} \neq 0.$$

Seja \mathbf{p} a projeção ortogonal de \mathbf{b} sobre $N(A^\top)$. Como $\mathbf{b} \notin C(A)$, essa projeção é não nula: $\mathbf{p} \neq 0$. Então:

$$\mathbf{p}^\top \mathbf{b} = \mathbf{p}^\top \mathbf{p} = \|\mathbf{p}\|^2 \neq 0.$$

Definimos o vetor:

$$\mathbf{y} = \frac{1}{\mathbf{p}^\top \mathbf{p}} \mathbf{p}.$$

Com isso, temos:

$$A^\top \mathbf{y} = \frac{1}{\mathbf{p}^\top \mathbf{p}} A^\top \mathbf{p} = 0, \quad \text{pois } \mathbf{p} \in N(A^\top),$$

e

$$\mathbf{y}^\top \mathbf{b} = \frac{1}{\mathbf{p}^\top \mathbf{p}} \mathbf{p}^\top \mathbf{b} = \frac{\mathbf{p}^\top \mathbf{b}}{\mathbf{p}^\top \mathbf{p}} = 1.$$

Portanto, existe \mathbf{y} tal que $A^\top \mathbf{y} = 0$ e $\mathbf{y}^\top \mathbf{b} = 1$, como queríamos demonstrar.

3

Justifique as seguintes afirmações verdadeiras:

1. Se $AB = 0$, então o espaço coluna de B está contido no núcleo de A .

Solução: Suponha, por contradição, que exista um vetor $\mathbf{y} = B\mathbf{x}$ que pertence ao espaço coluna de B , mas que não está no núcleo de A . Isso significa que:

$$A(B\mathbf{x}) = (AB)\mathbf{x} \neq 0,$$

o que contradiz a hipótese de que $AB = 0$. Logo, $C(B) \subseteq N(A)$.

2. Se A é uma matriz simétrica, então seu espaço coluna é ortogonal ao seu núcleo.

Solução: Se A é simétrica, então $A = A^\top$. Assim, o espaço coluna de A coincide com seu espaço linha:

$$C(A) = C(A^\top).$$

Pelo Teorema Fundamental da Álgebra Linear, temos:

$$N(A) = C(A^\top)^\perp = C(A)^\perp,$$

logo, o núcleo de A é ortogonal ao seu espaço coluna.

3. Se um subespaço $S \subseteq V$, então $S^\perp \supseteq V^\perp$.

Solução: Seja $\mathbf{v} \in V^\perp$, ou seja, \mathbf{v} é ortogonal a todos os vetores de V . Como $S \subseteq V$, então \mathbf{v} também é ortogonal a todos os vetores de S , o que implica $\mathbf{v} \in S^\perp$. Logo:

$$V^\perp \subseteq S^\perp.$$

4. Para qualquer subespaço V , tem-se $(V^\perp)^\perp = V$.

Solução: Por definição, V^\perp é o conjunto de vetores ortogonais a todos os vetores em V . Assim, qualquer vetor de V é ortogonal a todos os vetores em V^\perp , o que implica:

$$V \subseteq (V^\perp)^\perp.$$

Suponha que $\dim(V) = r$ em um espaço de dimensão n . Então:

$$\dim(V^\perp) = n - r, \quad \text{e} \quad \dim((V^\perp)^\perp) = r.$$

Como $V \subseteq (V^\perp)^\perp$ e ambos têm a mesma dimensão, segue que:

$$(V^\perp)^\perp = V.$$

Alternativamente, como qualquer subespaço V pode ser descrito como o núcleo de alguma matriz A , ou seja, $V = N(A)$, então:

$$V^\perp = C(A^\top), \quad \text{e} \quad (V^\perp)^\perp = N((A^\top)^\top) = N(A) = V.$$

5. Se P é uma matriz de projeção, então $I - P$ também é.

Solução: Suponha que P projete ortogonalmente sobre um subespaço V . Então, $I - P$ projeta ortogonalmente sobre V^\perp . De fato, para qualquer vetor \mathbf{v} :

$$\mathbf{v} - (I - P)\mathbf{v} = \mathbf{v} - \mathbf{v} + P\mathbf{v} = P\mathbf{v},$$

e como $P\mathbf{v} \in V$, temos que $(I - P)\mathbf{v} \in V^\perp$, ou seja, $I - P$ projeta sobre V^\perp .

4

Weighted Least Squares

Neste problema, derivaremos o ajuste por mínimos quadrados ponderados.

Suponha que temos m pontos de dados (t_i, b_i) , que queremos ajustar a uma reta $b = C + Dt$. O método dos mínimos quadrados ordinários escolhe C e D para minimizar o erro quadrático total:

$$\sum_{i=1}^m (C + Dt_i - b_i)^2.$$

No entanto, nem todos os pontos de dados são igualmente confiáveis — muitas vezes, os dados reais vêm acompanhados de uma margem de erro $\sigma_i > 0$ associada a b_i . Para refletir isso, desejamos pesar menos os pontos com maior erro. Assim, buscamos minimizar a seguinte quantidade:

$$\varepsilon = \sum_{i=1}^m \left(\frac{C + Dt_i - b_i}{\sigma_i} \right)^2.$$

(a) **Escreva ε na forma matricial.**

Queremos reescrever ε na forma matricial, como no caso dos mínimos quadrados ordinários. Definimos:

$$A = \begin{pmatrix} 1 & t_1 \\ 1 & t_2 \\ \vdots & \vdots \\ 1 & t_m \end{pmatrix}, \quad \mathbf{x} = \begin{pmatrix} C \\ D \end{pmatrix}, \quad \mathbf{b} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}, \quad W = \begin{pmatrix} \frac{1}{\sigma_1} & 0 & \cdots & 0 \\ 0 & \frac{1}{\sigma_2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \frac{1}{\sigma_m} \end{pmatrix}.$$

Então podemos escrever:

$$\varepsilon = \|W A \mathbf{x} - W \mathbf{b}\|^2.$$

(b) Derive a equação linear cuja solução minimiza ε .

Estamos agora minimizando:

$$\|WA\mathbf{x} - W\mathbf{b}\|^2,$$

o que é um problema de mínimos quadrados ordinários com matriz WA e vetor $W\mathbf{b}$. A solução que minimiza esse erro satisfaz a equação normal:

$$(WA)^\top(WA)\mathbf{x} = (WA)^\top(W\mathbf{b}).$$

Multiplicando os termos:

$$A^\top W^\top WA\mathbf{x} = A^\top W^\top W\mathbf{b}.$$

Como W é diagonal e simétrica, temos $W^\top W = W^2$, portanto:

$$A^\top W^2 A\mathbf{x} = A^\top W^2 \mathbf{b}.$$

Escrevendo isso de forma explícita, obtemos:

$$\begin{pmatrix} \sum \frac{1}{\sigma_i^2} & \sum \frac{t_i}{\sigma_i^2} \\ \sum \frac{t_i}{\sigma_i^2} & \sum \frac{t_i^2}{\sigma_i^2} \end{pmatrix} \begin{pmatrix} C \\ D \end{pmatrix} = \begin{pmatrix} \sum \frac{b_i}{\sigma_i^2} \\ \sum \frac{t_i b_i}{\sigma_i^2} \end{pmatrix}.$$

Conselhos gerais:

- Revisem Jacobi e Seidel. Focar em: coisas básicas, relação entre eles, raio de convergência, número de condicionamento, resultados básicos.
- Revisem resultados básicos de AL
- Revisem condicionamento absoluto, relativo, matricial e normas
- Façam algumas questões sobre estabilidade, só por precaução
- Refaçam a prova!
- Quando virem algum resultado legal sobre os assuntos abordados tente prová-lo.

Atenção: Não vi o teste, são apenas conselhos gerais.