

AULA 34

- Note bem: a leitura destes apontamentos não dispensa de modo algum a leitura atenta da bibliografia principal da cadeira
- Chama-se a atenção para a importância do trabalho pessoal a realizar pelo aluno resolvendo os problemas apresentados na bibliografia, sem consulta prévia das soluções propostas, análise comparativa entre as suas resposta e a respostas propostas, e posterior exposição junto do docente de todas as dúvidas associadas.

TÓPICOS

Quadrados mínimos e projecção num subespaço.

Matriz pseudo-inversa.

34. Quadrados mínimos e projecção num subespaço.

34.1. Quadrados mínimos e projecção num subespaço.

Se \mathbb{W} é um subespaço de \mathbb{R}^m de dimensão n e \mathbf{A} é uma matriz $m \times n$ cujas colunas formam uma base de \mathbb{W} , a projecção ortogonal de um qualquer vector $\mathbf{u} \in \mathbb{R}^m$ no subespaço \mathbb{W} é dada por

$$\text{proj}_{\mathbb{W}} \mathbf{u} = \mathbf{A}(\mathbf{A}^T \mathbf{A})^{-1} \mathbf{A}^T \mathbf{u}$$

, ou seja, a matriz da transformação $T : \mathbb{R}^m \rightarrow \mathbb{R}^m$ que projecta ortogonalmente \mathbb{R}^m em \mathbb{W} é

$$\mathbf{T} = \mathbf{A}(\mathbf{A}^T \mathbf{A})^{-1} \mathbf{A}^T$$

Exemplos

1. Retomemos o exemplo 1 da aula 34. Como vimos, a melhor aproximação do vector $\mathbf{v} = [0 \ 2 \ 2]^T$ no subespaço \mathbb{W} , de \mathbb{R}^3 , gerado pelos vectores $\mathbf{u}_1 = [2 \ 2 \ 0]^T$ e $\mathbf{u}_2 = [-2 \ 2 \ 0]^T$ é o vector

$$\hat{\mathbf{v}} = \text{proj}_{\mathbb{W}} \mathbf{v} = [0 \ 2 \ 0]^T$$

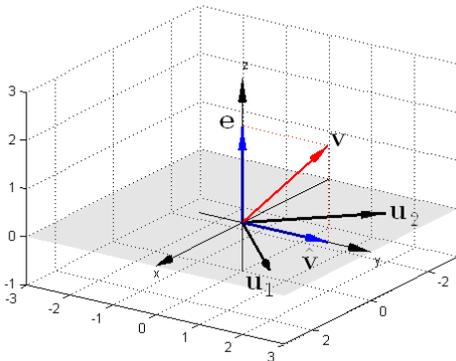


Figura 34.1

Atendendo a que a matriz da transformação $T : \mathbb{R}^m \rightarrow \mathbb{R}^m$ que projecta ortogonalmente \mathbb{R}^m em \mathbb{W} é

$$\mathbf{T} = \mathbf{A}(\mathbf{A}^T \mathbf{A})^{-1} \mathbf{A}^T$$

, temos, neste caso,

$$\mathbf{A} = [\mathbf{u}_1 \ \mathbf{u}_2] = \begin{bmatrix} 2 & -2 \\ 2 & 2 \\ 0 & 0 \end{bmatrix}$$

, e ainda,

$$\mathbf{A}^T \mathbf{A} = \begin{bmatrix} 2 & 2 & 0 \\ -2 & 2 & 0 \end{bmatrix} \begin{bmatrix} 2 & -2 \\ 2 & 2 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 8 & 0 \\ 0 & 8 \end{bmatrix}$$

$$(\mathbf{A}^T \mathbf{A})^{-1} = \begin{bmatrix} 8 & 0 \\ 0 & 8 \end{bmatrix}^{-1} = \frac{1}{8} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

$$\mathbf{A}(\mathbf{A}^T \mathbf{A})^{-1} = \begin{bmatrix} 2 & -2 \\ 2 & 2 \\ 0 & 0 \end{bmatrix} \frac{1}{8} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \frac{1}{8} \begin{bmatrix} 2 & -2 \\ 2 & 2 \\ 0 & 0 \end{bmatrix}$$

, e finalmente

$$\mathbf{T} = \mathbf{A}(\mathbf{A}^T \mathbf{A})^{-1} \mathbf{A}^T = \frac{1}{8} \begin{bmatrix} 2 & -2 \\ 2 & 2 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 2 & 2 & 0 \\ -2 & 2 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

Como tínhamos concluído nos exemplos 1 e 2 da aula 34, os vectores \mathbf{u}_1 e \mathbf{u}_2 geram o subespaço \mathbb{W} de \mathbb{R}^3 correspondente ao plano xy .

Recorde da aula 25, que a matriz de projecção no plano xy corresponde à matriz que acima foi determinada

$$\mathbf{T} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

A matriz é única, não dependendo dos vectores considerados para gerar o subespaço.

Consideremos, por exemplo, os vectores $\mathbf{u}_3 = [2 \ 2 \ 0]^T$ e $\mathbf{u}_4 = [1 \ 2 \ 0]^T$, facilmente se reconhece que os vectores geram o subespaço \mathbb{W} de \mathbb{R}^3 correspondente ao plano xy (basta notar que os vectores existem sobre o plano e têm diferentes direcções). No entanto, ao contrário de \mathbf{u}_1 e \mathbf{u}_2 , \mathbf{u}_3 e \mathbf{u}_4 não são ortogonais,

$$\mathbf{u}_3 \cdot \mathbf{u}_4 = 6$$

Independentemente deste facto, temos

$$\mathbf{A} = [\mathbf{u}_3 \ \mathbf{u}_4] = \begin{bmatrix} 2 & 1 \\ 2 & 2 \\ 0 & 0 \end{bmatrix}$$

, e a matriz de projecção no espaço por eles gerado

$$\mathbf{T} = \mathbf{A}(\mathbf{A}^T \mathbf{A})^{-1} \mathbf{A}^T = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$



```
>> u1=[2 2 0]';
>> u2=[1 2 0]';
>> A=[u1 u2];
```

$$\mathbf{T} = \mathbf{A} \cdot \text{inv}(\mathbf{A}' * \mathbf{A}) * \mathbf{A}'$$

$$\mathbf{T} =$$

$$\begin{array}{ccc} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{array}$$

Suponhamos agora que o conjunto de vectores não é linearmente independente. Por exemplo, os vectores $\mathbf{u}_5 = [2 \ 2 \ 0]^T$, $\mathbf{u}_6 = [1 \ 2 \ 0]^T$, e $\mathbf{u}_7 = [3 \ 4 \ 0]^T$. Facilmente se reconhece que $\mathbf{u}_7 = \mathbf{u}_5 + \mathbf{u}_6$. A matriz \mathbf{A} tem colunas linearmente dependentes, de onde resulta que $\mathbf{A}^T \mathbf{A}$ é singular, pelo que não podemos calcular $(\mathbf{A}^T \mathbf{A})^{-1}$.

Podemos contornar o problema, procurando uma base para o espaço gerado por \mathbf{u}_5 , \mathbf{u}_6 e \mathbf{u}_7 (o que, no presente exemplo, é obviamente simples: basta ignorar \mathbf{u}_7). Veremos na próxima secção como abordar a questão sem necessidade de procurar uma base do subespaço.

34.2. Matriz pseudo-inversa.

Seja \mathbf{A} uma matriz $m \times n$ com colunas linearmente independentes. Designa-se por matriz pseudo-inversa de \mathbf{A} a matriz $n \times m$

$$\mathbf{A}^+ = (\mathbf{A}^T \mathbf{A})^{-1} \mathbf{A}^T$$

A matriz pseudo-inversa de qualquer matriz \mathbf{A} (mesmo não sendo $\mathbf{A}^T \mathbf{A}$ invertível) pode ser calculada a partir da decomposição em valores singulares. Se $\mathbf{A} = \mathbf{U} \Sigma \mathbf{V}^T$ então

$$\mathbf{A}^+ = \mathbf{V} \Sigma^+ \mathbf{U}^T$$

em que Σ^+ é a matriz

$$\Sigma^+ = \begin{bmatrix} \mathbf{D}^{-1} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix}$$

, e \mathbf{D} é a matriz dos valores singulares de \mathbf{A} .

Propriedades da matriz pseudo inversa

$$\mathbf{A} \mathbf{A}^+ \mathbf{A} = \mathbf{A}$$

$$\mathbf{A}^+ \mathbf{A} \mathbf{A}^+ = \mathbf{A}^+$$

$\mathbf{A} \mathbf{A}^+$ e $\mathbf{A}^+ \mathbf{A}$ são matrizes simétricas.

Exemplos

2. Com base na definição da matriz pseudo-inversa podemos reformular o exemplo 1 e os exemplos da aula 34.

Podemos dizer, por exemplo, que, sendo \mathbf{A} a matriz cujas colunas formam uma base de \mathbb{W} , a matriz da transformação $T: \mathbb{R}^m \rightarrow \mathbb{R}^m$ que projecta ortogonalmente \mathbb{R}^m em \mathbb{W} é

$$\mathbf{T} = \mathbf{A} \mathbf{A}^+$$

, e ainda que, a solução de mínimo erro quadrático do sistema $\mathbf{A} \mathbf{x} = \mathbf{b}$ é

$$\hat{\mathbf{x}} = \mathbf{A}^+ \mathbf{b}$$

Note que, se \mathbf{A} for regular obtemos

$$\mathbf{x} = \mathbf{A}^{-1} \mathbf{b}$$

Independentemente da formulação mais sintética, não deixa de estar em causa o cálculo de $\mathbf{A}^+ = (\mathbf{A}^T \mathbf{A})^{-1} \mathbf{A}^T$ e, portanto, o problema da inversão da matriz $\mathbf{A}^T \mathbf{A}$. No entanto, recorrendo ao conceito de decomposição em valores singulares, podemos resolver a questão.

3. Retomemos o conjunto de vectores $\{\mathbf{u}_5, \mathbf{u}_6, \mathbf{u}_7\}$, com $\mathbf{u}_5 = [2 \ 2 \ 0]^T$, $\mathbf{u}_6 = [1 \ 2 \ 0]^T$, $\mathbf{u}_7 = [3 \ 4 \ 0]^T$ e o vector $\mathbf{v} = [0 \ 2 \ 2]^T$. Procuremos a melhor aproximação do vector \mathbf{v} no subespaço \mathbb{W} de \mathbb{R}^3 gerado pelos vectores \mathbf{u}_5 , \mathbf{u}_6 , \mathbf{u}_7 que, como sabemos, corresponde à projecção ortogonal de \mathbf{v} em \mathbb{W} .

Temos a matriz

$$\mathbf{A} = [\mathbf{u}_5 \quad \mathbf{u}_6 \quad \mathbf{u}_7] = \begin{bmatrix} 2 & 1 & 3 \\ 2 & 2 & 4 \\ 0 & 0 & 0 \end{bmatrix}$$

, e podemos verificar que a matriz

$$\mathbf{A}^T \mathbf{A} = \begin{bmatrix} 2 & 2 & 0 \\ 1 & 2 & 0 \\ 3 & 4 & 0 \end{bmatrix} \begin{bmatrix} 2 & 1 & 3 \\ 2 & 2 & 4 \\ 0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 8 & 6 & 14 \\ 6 & 5 & 11 \\ 14 & 11 & 25 \end{bmatrix}$$

não é invertível

$$\begin{bmatrix} 8 & 6 & 14 & 1 & 0 & 0 \\ 6 & 5 & 11 & 0 & 1 & 0 \\ 14 & 11 & 25 & 0 & 0 & 1 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & 1 & 0 & -11/4 & 5/4 \\ 0 & 1 & 1 & 0 & 7/2 & -3/2 \\ 0 & 0 & 0 & 1 & 1 & -1 \end{bmatrix}$$



```
>> u1=[2 2 0]';
>> u2=[1 2 0]';
>> u3=[3 4 0]';
>> A=[u1 u2 u3];
>> inv(A'*A)
```

Warning: Matrix is close to singular or badly scaled.

Results may be inaccurate. RCOND = 1.184238e-017.

ans =

```
      *           *           *
      *           *           *
      *           *           *
```

No entanto, a matriz pseudo-inversa de qualquer matriz \mathbf{A} , mesmo não sendo $\mathbf{A}^T \mathbf{A}$ invertível, pode ser calculada a partir da decomposição em valores singulares da matriz \mathbf{A} , sendo

$$\mathbf{A}^+ = \mathbf{V} \Sigma^+ \mathbf{U}^T$$

em que Σ^+ é

$$\Sigma^+ = \begin{bmatrix} \mathbf{D}^{-1} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix}$$

, sendo \mathbf{D} a matriz dos valores singulares de \mathbf{A} .

Dada a morosidade dos cálculos, vamos recorrer ao MatLab para decompor \mathbf{A} em valores singulares.



```
>> [U S V]=svd(A);
>> S(1:2,1:2)=inv(S(1:2,1:2))
>> Am=V*S*U'
```

$$\mathbf{A}_m = \begin{bmatrix} 1.0000 & -0.6667 & 0 \\ -1.0000 & 0.8333 & 0 \\ -0.0000 & 0.1667 & 0 \end{bmatrix}$$

O cálculo é dado meramente a título de exemplo, dado que existe em MatLab a função `pinv(A)` que calcula a pseudo-inversa de uma matriz \mathbf{A} .



```
>> pinv(A)
ans =
    1.0000   -0.6667    0
   -1.0000    0.8333    0
   -0.0000    0.1667    0
```

$$\mathbf{A}^+ = \begin{bmatrix} 1 & -2/3 & 0 \\ -1 & 5/6 & 0 \\ 0 & 1/6 & 0 \end{bmatrix}$$

Dada a pseudo-inversa, podemos agora calcular a melhor aproximação do vector $\mathbf{v} = [0 \ 2 \ 2]^T$ no subespaço \mathbb{W} , de \mathbb{R}^3 , gerado pelos vectores \mathbf{u}_5 , \mathbf{u}_6 , \mathbf{u}_7 , que, como sabemos, corresponde à projecção ortogonal de \mathbf{v} em \mathbb{W} .

Temos

$$\begin{aligned} \hat{\mathbf{v}} &= \mathbf{A}\mathbf{A}^+\mathbf{v} = \begin{bmatrix} 2 & 1 & 3 \\ 2 & 2 & 4 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & -2/3 & 0 \\ -1 & 5/6 & 0 \\ 0 & 1/6 & 0 \end{bmatrix} \begin{bmatrix} 0 \\ 2 \\ 2 \end{bmatrix} \\ &= \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 0 \\ 2 \\ 2 \end{bmatrix} \\ &= \begin{bmatrix} 0 \\ 2 \\ 0 \end{bmatrix} \end{aligned}$$

O que confirma os cálculos efectuados nos exemplos anteriores $\hat{\mathbf{v}} = [0 \ 2 \ 0]^T$.