# Álgebra Linear - Lista de Exercícios 8

Caio Lins e Tiago da Silva

19 de outubro de 2021

1. Escreva as 3 equações para a reta $b = C + Dt$ passar pelos pontos $(-1, 7)$, $(1, 7)$, $(2, 21)$. Ache a solução de mínimos quadrados $\hat{x}$ e a projeção $p = A\hat{x}$.

**Resolução:**
Precisamos, nesse exercício, escrever as condições necessárias para que a reta $\{(t, C + Dt) : t \in \mathbb{R}\}$ contenha os pontos $(-1, 7)$, $(1, 7)$ e $(2, 21)$; desta forma, escolhas subsequentes de $t$ no conjunto $\{-1, 1, 2\}$ culminam em

$$\begin{cases} C - D = 7 \\ C + D = 7 \\ C + 2D = 21, \end{cases} \tag{1}$$

que é, para $(C, D) \in \mathbb{R}^2$, absolutamente inadmissível (por exemplo, as equações iniciais implicam que $C = 7$ e $D = 0$; estas condições, contudo, violam a outra igualdade). Nesse contexto, como a Equação (1) é equivalente a

$$A \begin{bmatrix} C \\ D \end{bmatrix} = \begin{bmatrix} 7 \\ 7 \\ 21 \end{bmatrix},$$

com $A = \begin{bmatrix} \mathbf{a} & \mathbf{w} \end{bmatrix}$, em que $\mathbf{a} = \begin{bmatrix} 1 & 1 & 1 \end{bmatrix}^T$ e $\mathbf{w} = \begin{bmatrix} -1 & 1 & 2 \end{bmatrix}^T$, podemos projetar (ortogonalmente) o vetor $\mathbf{b} = \begin{bmatrix} 7 & 7 & 21 \end{bmatrix}^T$ no espaço coluna de $A$; para isso, precisamos (sistematicamente) enfrentar a equação

$$A^T A \begin{bmatrix} \hat{C} \\ \hat{D} \end{bmatrix} = A^T \mathbf{b}$$

(com $(\hat{C}, \hat{D}) \in \mathbb{R}^2$), que é igual a

$$\begin{bmatrix} 3 & 2 \\ 2 & 6 \end{bmatrix} \begin{bmatrix} \hat{C} \\ \hat{D} \end{bmatrix} = \begin{bmatrix} 35 \\ 42 \end{bmatrix};$$

logo, $\hat{C} = 9$ e $\hat{D} = 4$. Agora, o vetor $\hat{x}$, descrito no enunciado, é precisamente igual ao vetor $\begin{bmatrix} \hat{C} & \hat{D} \end{bmatrix}^T$; portanto, como $p = A\hat{x}$, temos que

$$\hat{x} = \begin{bmatrix} 9 \\ 4 \end{bmatrix} \text{ e } p = \begin{bmatrix} 5 \\ 13 \\ 17 \end{bmatrix}.$$

---

2. Dado o problema acima, quais dos quatro subespaços fundamentais contêm o vetor erro $e = b - p$? E o vetor $p$? E o vetor $\hat{x}$? Qual é o núcleo de $A$?

**Resolução:**
Perceba, inicialmente, que, utilizando as notações do exercício um (isto é, escrevemos $\mathbf{b}$ em oposição a $b$), $e = \mathbf{b} - p$ é ortogonal ao espaço coluna de $A$ (com efeito, temos que $p = A(A^T A)^{-1} A^T \mathbf{b} = P\mathbf{b}$, em que $P$ é simétrica; logo, se $v \in C(A)$,

$$
\begin{aligned}
&\langle \mathbf{b} - p, v\rangle = \\
&= \langle \mathbf{b}, v\rangle - \langle P\mathbf{b}, v\rangle = \\
&= \langle \mathbf{b}, v\rangle - \langle \mathbf{b}, P^T v\rangle = \\
&= \langle \mathbf{b}, v\rangle - \langle \mathbf{b}, Pv\rangle = \\
&= \langle \mathbf{b}, v\rangle - \langle \mathbf{b}, v\rangle = 0;
\end{aligned}
$$

portanto, $e = \mathbf{b} - p$ é ortogonal a qualquer vetor em $C(A)$ e, nesse sentido, $e \in C(A)^{\perp}$); isto é, $e \in C(A)^{\perp} = N(A^T)$ (isso porque, se $x \in C(A)^{\perp}$, então, em particular, $x$ é ortogonal às colunas de $A$ e, dessa maneira, $A^T x = 0$). Por outro lado, $p$ está no espaço coluna de $A$; ele é, por definição, a projeção de $\mathbf{b}$ neste espaço. Além disso, como $A^T A$ é invertível (porquanto, por exemplo, seu determinante é não nulo), temos que seu posto é igual a dois e, portanto, o posto de $A^T$ é igual a dois, o que, desse modo, garante que o espaço coluna de $A^T$ é igual a $\mathbb{R}^2$ e, portanto, $\hat{x}$ pertence a $C(A^T)$. Nessas condições, temos também que, como $N(A)$ é igual ao complemento ortogonal, em $\mathbb{R}^2$, de $C(A^T)$, $N(A) = \{0\} \subset \mathbb{R}^2$.

---

3. Ache a melhor reta que se ajusta aos pontos $t = -2, -1, 0, 1, 2$ e $b = 4, 2, -1, 0, 0$.

**Resolução:**
Contemplamos, neste exercício, a caracterização da otimalidade da reta pela otimização do desvio quadrático; isto é, a reta $f(t) = wt + a$ é ótima no sentido de que, se $\tilde{f}(t) = \tilde{w}t + \tilde{a}$ é outra reta que objetiva se ajustar aos dados $\{(t_i, b_i) : 1 \leq i \leq 5\}$, então

$$
\sum_{1 \leq i \leq 5} |f(t_i) - b_i|^2 \leq \sum_{1 \leq i \leq 5} |\tilde{f}(t_i) - b_i|^2
$$

(estou escrevendo isso porque, em geral, a verificação de que uma reta, um modelo linear, é mais apropriada que outra trascende o cômputo do desvio quadrático). Nessas condições, seja

$$
A = \begin{bmatrix} -2 & 1 \\ -1 & 1 \\ 0 & 1 \\ 1 & 1 \\ 2 & 1 \end{bmatrix};
$$

temos, dessa forma, que computar $\mathbf{w} = \begin{bmatrix} w \\ b \end{bmatrix}$ tal que $\|A\mathbf{w} - b\|^2 \leq \|A\mathbf{x} - b\|^2$ para qualquer $\mathbf{x} \in \mathbb{R}^2$, o que é equivalente a abordar a equação $A^T A\mathbf{w} = b$[1]; isto é,

$$
\begin{bmatrix} 10 & 0 \\ 0 & 5 \end{bmatrix} \begin{bmatrix} w \\ b \end{bmatrix} = \begin{bmatrix} -10 \\ 5 \end{bmatrix}.
$$

Portanto, $w = -1$ e $b = 1$, o que, em particular, implica que a reta descrita no enunciado é igual a $\{(x, -x + 1), x \in \mathbb{R}\}$.

---

4. Dados os vetores

$$
v_1 = [1 \ -1 \ 0 \ 0], \ v_2 = [0 \ 1 \ -1 \ 0] \text{ e } v_3 = [0 \ 0 \ 1 \ -1],
$$

use o método de Gram-Schmidt para achar uma base ortornormal que gera o mesmo espaço de $v_1, v_2, v_3$.

**Resolução:**
O algoritmo de Gram-Schmidt contempla, em sua descrição, duas etapas: projeção e normalização. Explicitamente, seja $\{v_1, \cdots, v_n\}$ um conjunto de vetores linearmente independentes; o método de

[1] *Maybe you don't believe this assertion; in this case:* o gradiente, com respeito a $\mathbf{w}$, de $\|A\mathbf{w} - b\|^2$ é igual a $A^T(A\mathbf{w} - b)$; as condições de otimalidade, nesse sentido, exigem que $A^T A\mathbf{w} - A^T b = 0$ e, como $A^T A$ é, nesse exercício, positiva definida, essas condições são suficientes para a garantia de que $\|A\mathbf{w} - b\| < \|A\mathbf{x} - b\|$ para qualquer $\mathbf{x} \in \mathbb{R}^2$.

Gram-Schmidt, nesse sentido, permite, deterministicamente, computar um conjunto $\{u_1, \cdots, u_n\}$ de vetores ortonormais tais que, para $1 \leq i \leq n$, $\text{span}(\{u_1, \cdots, u_i\}) = \text{span}(\{v_1, \cdots, v_i\})$. Iniciamos, nesse contexto, escrevendo $u_i = \frac{v_1}{\|v_1\|}$ e, iterativamente, computamos

$$\tilde{u}_m = v_m - \sum_{1 \leq i \leq m-1} \langle u_i, v_m \rangle u_i; \tag{2}$$

em seguida, $u_m = \frac{\tilde{u}_m}{\|\tilde{u}_m\|}$ para $m \geq 2$ (e $m \leq n$). Munidos, portanto, do conjunto de vetores $\{v_1, v_2, v_3\}$, descritos no enunciado, podemos computar $u_1 = \frac{1}{\sqrt{2}} v_1$ e

$$\tilde{u}_2 = v_2 - \langle u_1, v_2 \rangle u_1 = v_2 - \frac{1}{2} \langle v_1, v_2 \rangle v_1 = \begin{bmatrix} 1/2 \\ 1/2 \\ -1 \\ 0 \end{bmatrix},$$

o que nos direciona a $u_2 = \left(\frac{2}{3}\right)^{1/2} \tilde{u}_2$ (isso porque $\|\tilde{u}_2\|^2 = \frac{3}{2}$). Utilizando, agora, a Equação (2), temos que

$$\begin{aligned} \tilde{u}_3 &= v_3 - \langle u_1, v_3 \rangle u_1 - \langle u_2, v_3 \rangle u_2 = \\ &= v_3 - \frac{1}{2} \langle v_1, v_3 \rangle v_1 - \frac{2}{3} \langle \tilde{u}_2, v_3 \rangle \tilde{u}_2 = \\ &= v_3 + \frac{2}{3} \tilde{u}_2 = \begin{bmatrix} 1/3 \\ 1/3 \\ 1/3 \\ -1 \end{bmatrix}; \end{aligned}$$

isto é, $u_3 = \frac{\sqrt{3}}{2} \tilde{u}_3$. Portanto,

$$u_1 = \begin{bmatrix} 1/\sqrt{2} \\ -1/\sqrt{2} \\ 0 \\ 0 \end{bmatrix}, u_2 = \begin{bmatrix} 1/\sqrt{6} \\ 1/\sqrt{6} \\ \sqrt{2/3} \\ 0 \end{bmatrix} \text{ e } u_3 = \begin{bmatrix} 1/(2\sqrt{3}) \\ 1/(2\sqrt{3}) \\ 1/(2\sqrt{3}) \\ \sqrt{3}/2 \end{bmatrix}$$

são os vetores que, nas condições do algoritmo de Gram-Schimdt, formam uma base ortonormal para o espaço vetorial $\text{span}(\{v_1, v_2, v_3\})$. Aliás, o `numpy`, em Python, enseja o cômputo destes vetores; as linhas seguintes, por exemplos, executam essa tarefa.

```python
import numpy as np

x = np.array([1, -1, 0, 0]).reshape(-1, 1)
y = np.array([0, 1, -1, 0].reshape(-1, 1)
w = np.array([0, 0, 1, -1]).reshape(-1, 1)
# Concatena os vetores
A = np.hstack([x, y, w])

# QR
q, r = np.linalg.qr(A)
# q = [u_{1} u_{2} u_{3}]
print(q)
```

---

5. Se os elementos de cada linha de uma matriz $A$ somam zero, ache uma solução para $Ax = 0$ e conclua que $\det A = 0$. Se esses elementos somam 1, conclua que $\det(A - I) = 0$.

**Resolução:**
Seja $A \in \mathbb{R}^{n \times n}$ ($A$ é, nas condições do enunciado, quadrada; isso porque a definição de autovalores e, em particular, de determinante é realizada nesta classe de matrizes); seja, além disso, $\{\mathbf{a}_1, \cdots, \mathbf{a}_n\} \subset \mathbb{R}^n$

o conjunto das linhas de $A$, que identificamos, desta vez, como vetores em $\mathbb{R}^n$. Seja, logo, $\mathbf{v} = (v_i)_{1 \leq i \leq n} \in \mathbb{R}^n$, em que $v_i = 1$ para $1 \leq i \leq n$; a verificação de que a soma dos elementos em cada linha de $A$ é nula culmina, portanto, em $\mathbf{a}_i^T \mathbf{v} = \mathbf{0}$; isto é, $\mathbf{v}$ é ortogonal às linhas de $A$, o que garante que $\mathbf{v} \in C(A^T)^\perp = N(A)$. Temos, desta maneira, que o núcleo de $A$ contém algum vetor não nulo (especificamente, $\mathbf{v}$), o que, neste cenário, implica a nulidade do seu determinante, $\det A = 0$ (isso porque, por exemplo, 0 é um autovalor de $A$; logo, como o determinante é igual ao produto dos autovalores, $\det A = 0$; por outro lado, o posto de $A$, quadrada, é igual à sua dimensão se, e somente se, seu determinante é não nulo; contudo, como a dimensão do núcleo de $A$ é positiva, temos que $\text{posto}(A) < n$, o que também implica a nulidade do determinante). Correlativamente, se as somas dos elementos das linhas de $A$ é unitária, temos que $(\mathbf{a}_i - \mathbf{e}_i)^T \mathbf{v} = \mathbf{0}$, em que $\mathbf{e}_i$ é a $i$-ésima linha da matriz identidade, e, logo, $\mathbf{v}$ pertence ao núcleo de $A - I$, garantindo, deste modo, que $\det(A - I) = 0$.

---

6. Use as propriedades do determinante (e não suas fórmulas) para mostrar que

$$\det \begin{bmatrix} 1 & a & a^2 \\ 1 & b & b^2 \\ 1 & c & c^2 \end{bmatrix} = (b-a)(c-a)(c-b).$$

**Resolução:**
Seja $A(a,b,c)$ a matriz do enunciado: ela é chamada de *matriz de Vandermonde*; seu determinante, aliás, é convenientemente descrito por *determinante de Vandermonde* e é nulo se, e somente se, a cardinalidade do conjunto $\{a,b,c\}$ for distinta de três (essa verificação tem aplicações, por exemplo, na teoria de interpolação de polinômios, garantindo, com efeito, a existência e a unicidade de um polinômio de grau (até) $n$ que contempla, em sua curva, $n+1$ pares distintos de coordenadas informados a priori). Como o determinante é, por definição (em $\mathbb{R}^n$), uma $n$-forma antissimétrica que é unitária na matriz identidade[2], temos que ele é um polinônmio de grau dois em $a$, em $b$ e em $c$; com efeito, se escrevermos $f : \mathbb{R} \to \mathbb{R}$ para a função $f(a) = \det A(a,b,c)$,

$$f(a) = \det \begin{bmatrix} 1 & a & a^2 \\ 1 & b & b^2 \\ 1 & c & c^2 \end{bmatrix} =$$
$$= \det \begin{bmatrix} 1 & 0 & a^2 \\ 1 & b & b^2 \\ 1 & c & c^2 \end{bmatrix} + \det \begin{bmatrix} 1 & a & a^2 \\ 1 & 0 & b^2 \\ 1 & 0 & c^2 \end{bmatrix} =$$
$$= \det \begin{bmatrix} 1 & 0 & 0 \\ 1 & b & b^2 \\ 1 & c & c^2 \end{bmatrix} + \det \begin{bmatrix} 1 & 0 & a^2 \\ 1 & b & 0 \\ 1 & c & 0 \end{bmatrix} + \det \begin{bmatrix} 1 & a & 0 \\ 1 & 0 & b^2 \\ 1 & 0 & c^2 \end{bmatrix}$$

(isso porque

$$\det \begin{bmatrix} 1 & a & a^2 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix} = 0,$$

porquanto o determinante é antissimétrico e a coluna três é diretamente proporcional à coluna dois), e os dois componentes, nessa expressão, à direita são polinômios de grau (até) dois em $a$ (na medida em que o determinante é uma $n$-forma). Agora, se $a = b$ e $a = c$, a estabilidade do determinante à transposição implica que $b$ e $c$ são raízes de $f$; logo, o teorema fundamental da álgebra garante que $f(a) = \alpha(a-b)(a-c)$ para algum $\alpha \in \mathbb{R}$. Escolhemos, portanto, $a = 0$ e verificamos, por ser estável, com respeito às operações elementares, o determinante, que

$$\det \begin{bmatrix} 1 & 0 & 0 \\ 1 & b & b^2 \\ 1 & c & c^2 \end{bmatrix} = \det \begin{bmatrix} 1 & 0 & 0 \\ 1 & b & 0 \\ 1 & c & c^2 - cb \end{bmatrix} = cb(c-b)$$

[2]Explicitamente, uma $n$-forma antissimétrica é uma função $f : (\mathbb{R}^n)^n \to \mathbb{R}$, $\mathbf{v} = (v_1, \cdots, v_n) \mapsto \alpha \in \mathbb{R}$, $v_i \in \mathbb{R}^n$, que é linear em cada $v_i$ (isto é, $f(v_1, \cdots, v_i + \beta u_i, \cdots, v_n) = f(v_1, \cdots, v_i, \cdots, v_n) + \beta f(v_1, \cdots, u_i, \cdots, v_n)$) e que satisfaz, para $i \neq j$, $f(v_1, \cdots, v_i, \cdots, v_j, \cdots, v_n) = -f(v_1, \cdots, v_j, \cdots, v_i, \cdots, v_n)$ (esta é a antissimetria; perceba, em particular, que, se existirem $i$ e $j$, $i \neq j$, tais que $v_i = v_j$, então $f(\mathbf{v}) = 0$).

(utilizamos, nesta expressão, a verificação de que o determinante de uma matriz triangular é igual ao produto dos elementos da sua diagonal; isso é verdade porque (1) ele é uma $n$-forma, (2) ele é unitário na matriz identidade e (3) ele é estável à transposição); portanto, $\alpha = (c-b)$. Temos, desse modo, que $\det A = (c-b)(a-b)(a-c)$.

---

7. Calcule

$$\det \begin{bmatrix} 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}.$$

**Resolução:**
Seja $A$ a matriz do enunciado; perceba que, como

$$A = \begin{bmatrix} \mathbf{e}_2 & \mathbf{e}_3 & \mathbf{e}_4 & \mathbf{e}_1 \end{bmatrix},$$

a caracterização do determinante como uma $n$-forma garante que

$$\begin{aligned} \det A = -\det \begin{bmatrix} \mathbf{e}_1 & \mathbf{e}_3 & \mathbf{e}_4 \mathbf{e}_2 \end{bmatrix} &= \\ = \det \begin{bmatrix} \mathbf{e}_1 & \mathbf{e}_2 & \mathbf{e}_4 & \mathbf{e}_3 \end{bmatrix} &= \\ = -\det \begin{bmatrix} \mathbf{e}_1 & \mathbf{e}_2 & \mathbf{e}_3 & \mathbf{e}_4 \end{bmatrix} &= -1 \end{aligned}$$

por definição; isto é, $\det A = -1$.

---

8. Use o fato de que

$$\det \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 2 & 3 & 4 \\ 1 & 3 & 6 & 10 \\ 1 & 4 & 10 & 20 \end{bmatrix} = 1$$

para mostrar que

$$\det \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 2 & 3 & 4 \\ 1 & 3 & 6 & 10 \\ 1 & 4 & 10 & \mathbf{19} \end{bmatrix} = 0.$$

**Resolução:**
Sejam, agora, $A = (a_{ij})_{1 \leq i,j \leq 4}$ e $B = (b_{ij})_{1 \leq i,j \leq 4}$, com

$$A = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 2 & 3 & 4 \\ 1 & 3 & 6 & 10 \\ 1 & 4 & 10 & 20 \end{bmatrix} \text{ e } B = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 2 & 3 & 4 \\ 1 & 3 & 6 & 10 \\ 1 & 4 & 10 & 19 \end{bmatrix};$$

se escrevermos, nesse sentido, $M = (a_{ij})_{1 \leq i,j \leq 3}$, temos que, pelo método dos cofatores,

$$\det A - \det B = (-1)^{4+4}(20 - 19) \det M = \det M$$

e, logo, como $\det M = 1$ ($M$ é uma matriz de Pascal simétrica; seu determinante é unitário) e $\det A = 1$, $\det B = 0$.

---

9. Ache o determinante da seguinte matriz:

$$A = \begin{bmatrix} 1 & 1 & 4 \\ 1 & 2 & 2 \\ 1 & 2 & 5 \end{bmatrix}$$

usando cofatores. O que acontece quando mudamos o valor 4 para 100?

**Resolução:**
O método dos cofatores, aplicado na linha inicial, culmina em

$$\det A = 1 \cdot (10 - 4) - 1 \cdot (5 - 2) + 4 \cdot (2 - 2) = 3;$$

por outro lado, como o cofator correspondente ao elemento $a_{13}$ é nulo, o determinante é estável com respeito a este elemento: isto é, a modificação de 4 para 100 é inócua a $\det A$.