Exercícios de Redução de Dimensionalidade, Análise de Componentes Principais e Clusterização K-means

2025

Exercício 1:

Use propriedades algébricas e a definição de autovalores e autovetores para provar que os autovalores de $A^{\mathsf{T}}A$ são todos positivos.

Exercício 2:

Determine a melhor fatoração de posto 1 de $A = \begin{bmatrix} 1 & 0 \\ 1 & 3 \\ 1 & 1 \end{bmatrix}$ e o erro associado usando o PCA.

Exercício 3:

Dado os pontos $a_1 = \begin{bmatrix} 3 \\ 0 \\ 1 \end{bmatrix}$, $a_2 = \begin{bmatrix} 1 \\ 0 \\ 3 \end{bmatrix}$, $a_3 = \begin{bmatrix} -3 \\ 0 \\ -1 \end{bmatrix}$ e $a_4 = \begin{bmatrix} -1 \\ 0 \\ -3 \end{bmatrix}$, determine a redução dos pontos em dimensão 1 e 2 com o PCA.

Exercício 4:

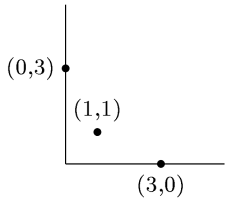
Seja B uma matriz simétrica e com autovalores distintos, prove que os seus autovetores são perpendiculares entre si (ortogonais entre si).

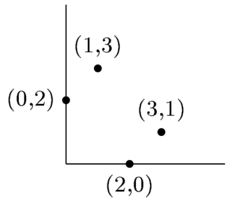
Exercício 5:

Determine uma matriz M de posto 1 (tal que $M = bc^{\mathsf{T}}$) que melhor representa a matriz A na norma de Frobenius:

1. $A = \begin{bmatrix} 0 & 1 & 3 \\ 3 & 1 & 0 \end{bmatrix}$

2. $A = \begin{bmatrix} 0 & 1 & 3 & 2 \\ 2 & 3 & 1 & 0 \end{bmatrix}$





Dica: utilize as simetrias do desenho.

Exercício 6:

Determine uma matriz de posto 1 que aproxima $A = \begin{bmatrix} 1 & 2 & 1 \\ 2 & 4 & 1 \\ 3 & 0 & 0 \end{bmatrix}$ na norma de Frobenius melhor que $\begin{bmatrix} 1 & 2 & 1 \\ 2 & 4 & 2 \\ 3 & 6 & 3 \end{bmatrix}$. Justifique como você achou essa matriz.

Exercício 7:

Dadas duas tabelas (matrizes):

Matriz Usuário-gênero	**Drama**	**Ação**
User A	0.9	0.1
User B	0.8	0.2
User C	0.3	0.7
User D	0.6	0.4
User E	0.0	1.0

Matriz Filme-gênero	**Drama**	**Ação**
Titanic	0.9	0.1
Rocky	0.1	0.9
The Hobbit	0.5	0.5
Fight Club	0.0	1.0
Jurassic Park	0.2	0.8

1. Calcule a tabela (matriz) de usuários por filmes.
2. Usando a matriz do item (a), determine a aproximação de posto 1 em Julia usando a função autovalores e autovetores.
3. Desenhe os filmes em dimensão 1.
4. Desenhe os usuários em dimensão 1.
5. Qual filme você recomendaria para quem gostou de Titanic usando os itens anteriores?

Exercício 8:

Temos que $A = BC$, onde A é uma matrix $m \times n$, B é uma matrix $m \times p$ e C é uma matrix $p \times n$ tal que p é o menor valor possível (p é o posto). Em cada item, ache "no olho"(sem usar um algoritmo) matrizes B e C que não foram dadas.

1. Seja $A = \begin{bmatrix} 2 & 10 & 20 & -2 & 0 \\ 3 & 15 & 30 & -30 & 0 \\ 5 & 25 & 50 & -5 & 0 \\ 7 & 35 & 70 & -7 & 0 \end{bmatrix}$

2. Seja $A = \begin{bmatrix} 1 & 3 & 5 & 5000 & 31 \\ 4 & 5 & 9 & 9000 & 54 \\ 3 & 5 & 8 & 8000 & 53 \end{bmatrix}$

Exercício 9:

Filmes Seja U uma matriz com a preferência de 4 usuários por 5 filmes levando em consideração somente o nível de comédia:

$$U = \begin{bmatrix} 2 & 10 & 20 & 200 & -10 \\ -3 & -15 & -30 & -300 & 15 \\ 5 & 25 & 50 & 500 & -25 \\ 7 & 35 & 70 & 700 & -35 \end{bmatrix}$$

Determine uma possível solução para quanto cada usuário gosta (ou não gosta) de comédia e quanto cada filme é (ou não é) de comédia.

Exercício 10:

Considere a bandeira da Grécia como uma imagem azul (pixel azul vale 1) e branca (pixel branco vale 0) e modele com uma matriz A.

Figura F1: Bandeira da Grécia.

1. Qual é o posto da bandeira da Grécia (tal que $A = BC^\top$)? O que a matriz B e matriz $C^\top$ representam nesse caso? Explique com as suas próprias palavras.

2. Determine dois países tais quais suas bandeiras tem posto 1.

3. Determine dois países tais quais suas bandeiras tem posto 2.

4. Determine dois países tais quais suas bandeiras tem posto 3.

Exercício 11:

Qual é a melhor matriz de posto 1 que representa a matriz $A = \begin{bmatrix} 2 & 3 \\ 1 & 0 \end{bmatrix}$?

Exercício 12:

Escreva uma matriz $A_{(4\times 6)}$ de posto 3 tal que a fatoração aproximada dela de posto 2 vai ter erro menor que 0.1.

Exercício 13:

Determine o primeiro componente principal v (em $\mathbb{R}^2$) de $A = \begin{bmatrix} 1 & 0 & 3 \\ 1 & 3 & 0 \end{bmatrix}$

Dica: Usa a simetria do gráfico para chutar o primeiro componente principal e depois faça com autovalores e autovetores.

Exercício 14:

Seja A uma matriz simétrica e com autovalores distintos, prove que os seus autovetores são perpendiculares.

Exercício 15:

Determine uma matriz M de posto 1 (tal que $M = bc^\mathsf{T}$) que melhor representa a matriz A e o erro:

1. $A = \begin{bmatrix} 0 & 1 & 3 \\ 3 & 1 & 0 \end{bmatrix}$

2. $A = \begin{bmatrix} 0 & 1 & 3 & 2 \\ 2 & 3 & 1 & 0 \end{bmatrix}$

Dica: usa a simetria dos pontos

Exercício 16:

Considere a seguinte imagem:

1. Insira esta imagem em uma matriz (5×5). Suponha que as sombras na imagem preta e branca que você vê são apenas valores 0, 0.5 e 1.

2. Qual é o posto da imagem?

3. Determine a imagem da matriz B tal que $A = BC$.

4. Faça a análise em componentes principais com posto 1, 2, 3, 4 e 5 e verifique os seus erros de aproximação.

Exercício 17:

A melhor aproximação de posto 1 da matriz A é xy^T. Determine a melhor matriz de posto 1 de A^T. Prove que sua resposta está correta.

Exercício 18:

Quanto se economiza de memória em porcentagem ao fatorar uma matriz 1000×1000 que tem posto 5 e armazenar a fatoração?

Exercício 19:

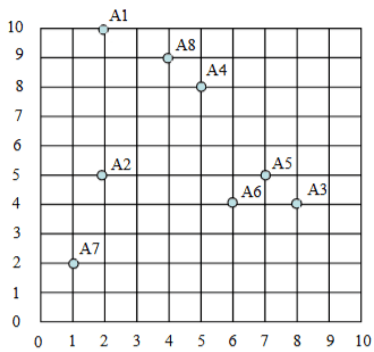
Seja $\begin{bmatrix} 2 & 2 & 10 \\ 5 & 6 & 25 \end{bmatrix} \approx \begin{bmatrix} 2 \\ 6 \end{bmatrix} \begin{bmatrix} 1 & 1 & 5 \end{bmatrix}$. Essa é a melhor fatoração de posto 1? Justifique sua resposta.

Exercício 20:

Encontre (ou desenhe) uma imagem A na internet, com mais ou menos 200×200 pixels (pode ser um pouco maior), que

1. exija mais que 3 componentes e menos que 6 no PCA para recuperar pelo menos 99% da norma de Frobenius total.
2. exija mais que 40 componentes para recuperar pelo menos 99% da norma de Frobenius total.

Exercício 21:



	A1	A2	A3	A4	A5	A6	A7	A8
A1		$\sqrt{25}$	$\sqrt{72}$	$\sqrt{13}$	$\sqrt{50}$	$\sqrt{52}$	$\sqrt{65}$	$\sqrt{5}$
A2			$\sqrt{37}$	$\sqrt{18}$	$\sqrt{25}$	$\sqrt{17}$	$\sqrt{10}$	$\sqrt{20}$
A3				$\sqrt{25}$	$\sqrt{2}$	$\sqrt{4}$	$\sqrt{53}$	$\sqrt{41}$
A4					$\sqrt{13}$	$\sqrt{17}$	$\sqrt{52}$	$\sqrt{2}$
A5						$\sqrt{2}$	$\sqrt{45}$	$\sqrt{25}$
A6							$\sqrt{29}$	$\sqrt{29}$
A7								$\sqrt{58}$

Use os pontos a_1, a_2, ..., a_8 e a tabela de distâncias acima para o exercício:

1. Faça só uma iteração (um passo) do algoritmo K-means com as sementes sendo a_1, a_2 e a_7.
2. Determine os clusters, os centróides, a fatoração matricial depois de uma iteração.
3. Represente graficamente o erro.

4. Quantas iterações você ia precisar fazer a mais até o algoritmo ter convergência?

Exercício 22:

Represente matricialmente o que seria pegar um dado com 100 pontos em dimensão 5, fazer uma redução de dimensionalidade com o PCA para duas dimensões e depois usar o K-means para clusterizar os pontos em dimensão 2.

Exercício 23:

Seja $A = \begin{bmatrix} 0 & 10 & 1 & 1 & 9 \\ 1 & 9 & 0 & 1 & 10 \end{bmatrix}$. Determine duas matrizes B e C tal que BC aproxima A na norma de Frobenius e C é uma matriz indicadora ("matriz de clusters").