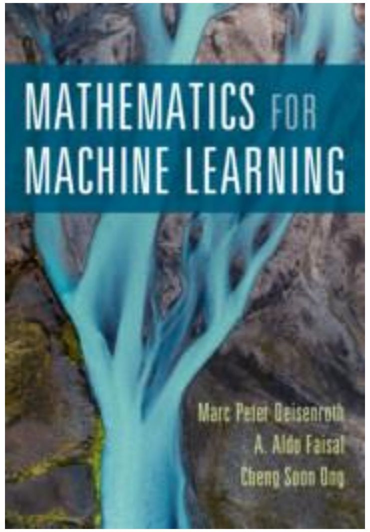
PROGRAMA DE PÓS-GRADUÇÃO EM CIÊNCIA DA COMPUTAÇÃO
TÓPICOS ESPECIAIS EM FUNDAMENTOS DE COMPUTAÇÃO – MATEMÁTICA E ESTATÍSTICA PARA CIÊNCIA DE DADOS
Prof. Dr. Rommel Melgaço Barbosa

Seminários

Decomposição de Matrizes



Hudson Romualdo

Fábio Pereira

André Riccioppo

Gabriel Almeida

Abril/2024

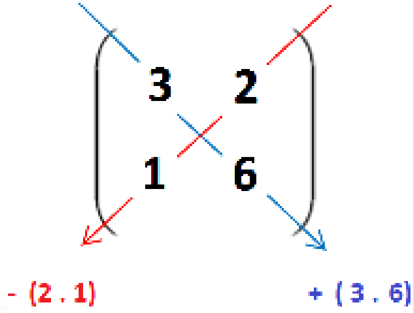
UFG
UNIVERSIDADE
FEDERAL DE GOIÁS

Características numéricas de Matrizes

Hudson de Paula Romualdo

Determinantes

- Os determinantes são conceitos importantes na álgebra linear.
- Um determinante é um objeto matemático na análise e solução de sistemas de equações lineares.
- Os determinantes são definidos <u>apenas para matrizes quadradas</u> $\boldsymbol{A} \in \mathbb{R}^{n\times n}$, ou seja, matrizes com o mesmo número de linhas e colunas.



- O determinante de uma matriz quadrada $\boldsymbol{A} \in \mathbb{R}^{n \times n}$ é uma função que mapeia $\boldsymbol{A}$ em um número real.

$$\det(\boldsymbol{A}) = \begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{vmatrix}$$

- Em uma expressão, a determinante de uma matriz $\boldsymbol{A}$ pode ser representado por $det(\boldsymbol{A})$ou por $|\boldsymbol{A}|$.

Aplicação: Testando a invertibilidade da Matriz

- Verificar a invertibilidade de uma matriz é crucial pois ela determina a existência de soluções para sistemas de equações, entre outros usos.
- Se $\boldsymbol{A}$ é uma matriz 1×1, ou seja, é um número escalar, temos:
  - $\boldsymbol{A} = a \implies \boldsymbol{A}^{-1} = \frac{1}{a}$. Então $a\frac{1}{a} = 1$ se mantem, se e somente se $a \neq 0$.
- Para matrizes 2×2 sabemos que $\boldsymbol{A}\boldsymbol{A}^{-1} = I$. Então, inverso de $\boldsymbol{A}$ é:

$$\boldsymbol{A}^{-1} = \frac{1}{a_{11}a_{22} - a_{12}a_{21}} \begin{bmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{bmatrix}$$

- Portanto, $\boldsymbol{A}$ é inversível se e somente se:

$$a_{11}a_{22} - a_{12}a_{21} \neq 0$$

- O resultado dessa expressão é o determinante de uma matriz $\boldsymbol{A} \in \mathbb{R}^{2\times 2}$:

$$\det(\boldsymbol{A}) = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{12}a_{21}$$

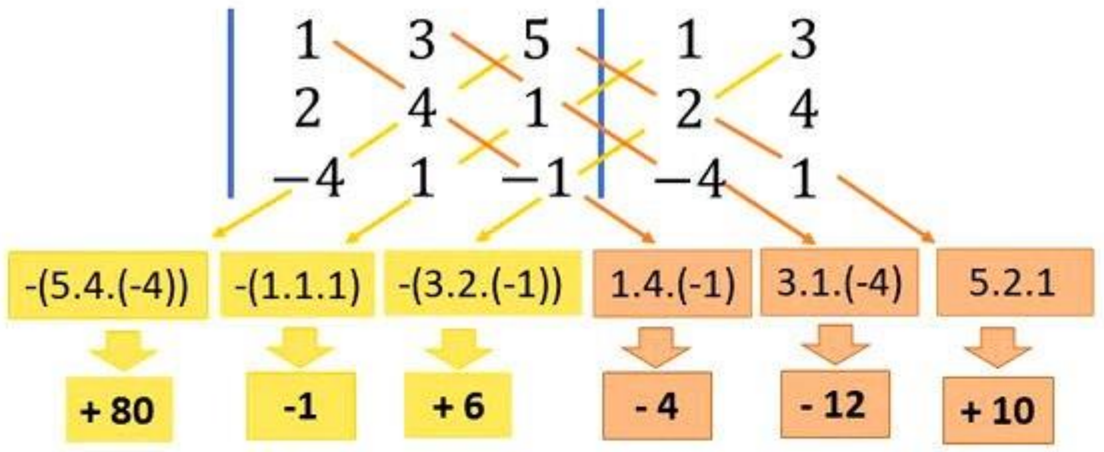
- Só é possível inverter qualquer matriz $\boldsymbol{A} \in \mathbb{R}^{n\times n}$ se e somente se $det(\boldsymbol{A}) \neq 0$.

Regra de Sarrus: determinante de matrizes de ordem 3

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = a_{11}a_{22}a_{33} + a_{21}a_{32}a_{13} + a_{31}a_{12}a_{23}$$

$$- a_{31}a_{22}a_{13} - a_{11}a_{32}a_{23} - a_{21}a_{12}a_{33}$$

- Exemplo:



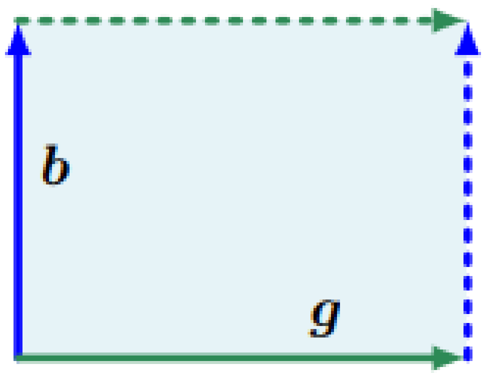
Determinante de matrizes Triangulares

- Chamamos uma matriz quadrada T de uma matriz triangular superior se $T_{ij} = 0$ para todo $i > j$, ou seja, os valores da matriz são zero abaixo da sua diagonal principal. Analogamente, definimos uma matriz triangular inferior como uma matriz com zeros acima da sua diagonal principal.
- Para uma matriz triangular $\boldsymbol{T} \in \mathbb{R}^{n \times n}$, o determinante é o produto dos elementos da diagonal principal, ou seja:

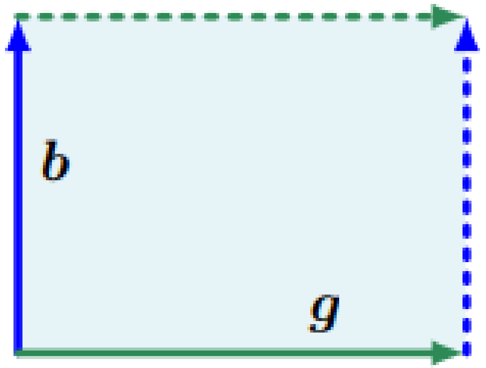
$$\det(\boldsymbol{T}) = \prod_{i=1}^{n} T_{ii}$$

Aplicação: Determinantes como Medida de Área e Volume

- A noção de determinante é natural quando a consideramos como uma função de um conjunto de n vetores que abrangem um objeto em $\mathbb{R}^n$.
- Para $n = 2$, as colunas da matriz formam um paralelogramo.

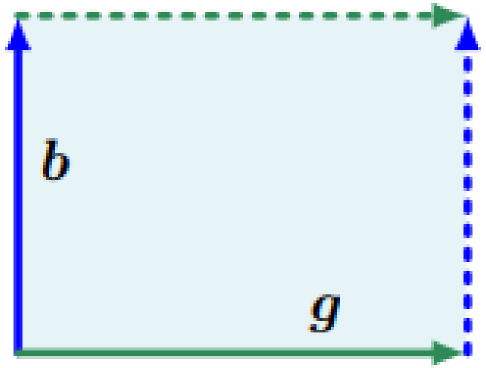


- Considere dois vetores b, g que formam as colunas de uma matriz $\boldsymbol{A} = [b, g]$. Então, o valor absoluto do determinante de A é a área do paralelogramo com vértices 0, b, g, $b + g$.

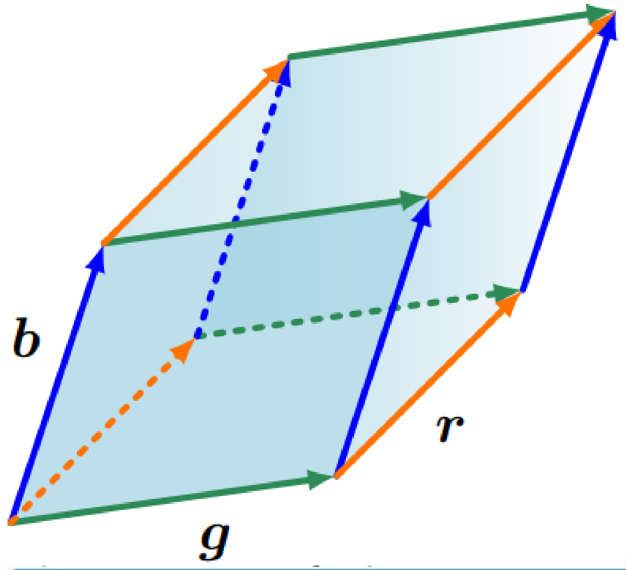


- Em particular, se b, g são linearmente dependentes de forma que $b = \lambda g$ para algum $\lambda \in \mathbb{R}$, eles não formam mais um paralelogramo bidimensional. Portanto, a área correspondente é 0.

- Se b, g são linearmente independentes e são múltiplos dos vetores da base canônica $\boldsymbol{e}_1, \boldsymbol{e}_2$, então eles podem ser escritos como $b\begin{bmatrix} b \\ 0 \end{bmatrix}$e $g\begin{bmatrix} 0 \\ g \end{bmatrix}$, e o determinante é$\begin{vmatrix} b & 0 \\ 0 & b \end{vmatrix} = bg - 0 = bg$.
- A fórmula parece familiar: *área = altura × comprimento*.
- No paralelogramo formado por b, g, a base do paralelogramo tem um comprimento $||g||$ e a altura tem um comprimento $||b|| \operatorname{sen} \theta$. Isso significa que a área (parte sombreada) do paralelogramo no total é exatamente a magnitude do produto vetorial.

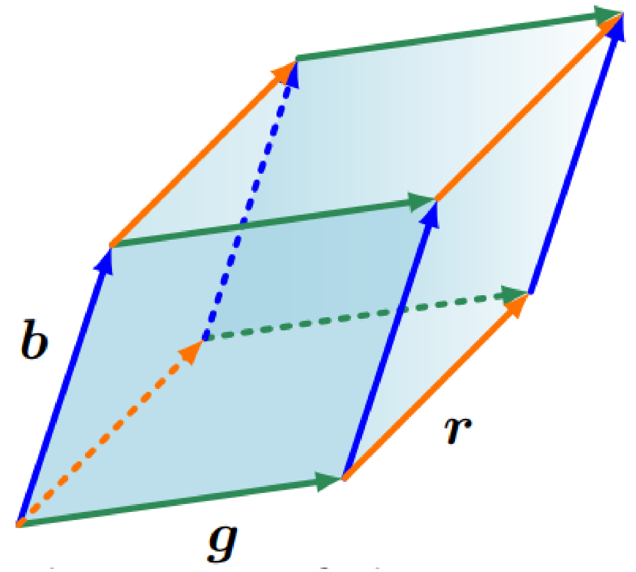


- Essa intuição se estende a dimensões superiores.
- Em $\mathbb{R}^3$, consideramos três vetores $r, b, g \in \mathbb{R}^3$ que abrangem as arestas de um paralelepípedo, ou seja, um sólido com faces que são paralelogramos.
- O valor absoluto do determinante da matriz 3×3 $[r, b, g]$ é o volume do sólido. Assim, o determinante age como uma função que mede o volume assinado formado por vetores coluna compostos em uma matriz.



$$\boldsymbol{r} = \begin{bmatrix} 2 \\ 0 \\ -8 \end{bmatrix}, \quad \boldsymbol{g} = \begin{bmatrix} 6 \\ 1 \\ 0 \end{bmatrix}, \quad \boldsymbol{b} = \begin{bmatrix} 1 \\ 4 \\ -1 \end{bmatrix}$$

$$\boldsymbol{A} = [\boldsymbol{r},\ \boldsymbol{g},\ \boldsymbol{b}] = \begin{bmatrix} 2 & 6 & 1 \\ 0 & 1 & 4 \\ -8 & 0 & -1 \end{bmatrix}$$



Regra de Sarrus:

$$\begin{array}{ccc|cc} 2 & 6 & 1 & 2 & 6 \\ 0 & 1 & 4 & 0 & 1 \\ -8 & 0 & -1 & -8 & 0 \end{array}$$

$$V = |det(\boldsymbol{A})| = |8 - 2 - 192| = 186$$

Expansão de Laplace

- Calcular o determinante de uma matriz $n \times n$ requer um algoritmo geral para resolver os casos para $n > 3$, a expansão de Laplace.
- O uso do teorema reduz o problema de calcular o determinante de uma matriz de ordem superior a 3 aplicando a expansão recursivamente até o calculo de determinantes de matrizes 2×2 no final.

- Considere uma matriz $\boldsymbol{A} \in \mathbb{R}^{n \times n}$. Então, para todo $j = 1, \ldots, n$:
  - Expansão ao longo da coluna *j*

$$\det(\boldsymbol{A}) = \sum_{k=1}^{n} (-1)^{k+j} a_{kj} \det(\boldsymbol{A}_{k,j})$$

  - Expansão ao longo da linha *j*

$$\det(\boldsymbol{A}) = \sum_{k=1}^{n} (-1)^{k+j} a_{jk} \det(\boldsymbol{A}_{j,k})$$

  - Aqui $\boldsymbol{A}_{k,j} \in \mathbb{R}^{(n-1)\times(n-1)}$ é a submatriz de $\boldsymbol{A}$ que obtemos ao excluir a linha k e a coluna j.

$$|\boldsymbol{A}_{3x3}| = \begin{vmatrix} \cancel{a_{11}} & \cancel{a_{12}} & \cancel{a_{13}} \\ \cancel{a_{21}} & a_{22} & a_{23} \\ \cancel{a_{31}} & a_{32} & a_{33} \end{vmatrix} = a_{11} \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix}$$

$$\boldsymbol{A} = [\boldsymbol{r},\ \boldsymbol{g},\ \boldsymbol{b}] = \begin{bmatrix} \cancel{2} & \cancel{6} & \cancel{1} \\ 0 & 1 & 4 \\ -8 & 0 & -1 \end{bmatrix}$$

$$\left.\begin{array}{l} (-1)^{1+1} \cdot 2 \cdot det\left(\begin{vmatrix} 1 & 4 \\ 0 & -1 \end{vmatrix}\right) = 1 \cdot 2 \cdot -1 = -2 \\[2em] (-1)^{1+2} \cdot 6 \cdot det\left(\begin{vmatrix} 0 & 4 \\ -8 & -1 \end{vmatrix}\right) = -1 \cdot 6 \cdot 32 = -192 \\[2em] (-1)^{1+3} \cdot 1 \cdot det\left(\begin{vmatrix} 0 & 1 \\ -8 & 0 \end{vmatrix}\right) = 1 \cdot 1 \cdot 8 = 8 \end{array}\right\} |-2 - 192 + 8| = 186$$

$$\begin{array}{ccc|cc} 2 & 6 & 1 & 2 & 6 \\ 0 & 1 & 4 & 0 & 1 \\ -8 & 0 & -1 & -8 & 0 \end{array}$$

$$V = |det(\boldsymbol{A})| = |8 - 2 - 192| = 186$$

Propriedades

Para $\boldsymbol{A} \in \mathbb{R}^{n \times n}$, o determinante exibe as seguintes propriedades:

- O determinante de um produto de matrizes é o produto dos determinantes correspondentes, $det(\boldsymbol{AB}) = det(\boldsymbol{A}) \cdot det(\boldsymbol{B})$
- Os determinantes são invariantes à transposição, ou seja, $det(\boldsymbol{A}) = det(\boldsymbol{A^T})$.
- Se $\boldsymbol{A}$ é regular (invertível), então $det(\boldsymbol{A}^{-1}) = \frac{1}{det(\boldsymbol{A})}$
- Matrizes semelhantes possuem o mesmo determinante. Portanto, para um mapeamento linear $\Phi : V \rightarrow V$ todas as matrizes de transformação $\boldsymbol{A}_\Phi$ de Φ têm o mesmo determinante. Assim, o determinante é invariante à escolha da base de um mapeamento linear.

- Adicionar um múltiplo de uma coluna/linha a uma outra não altera $det(\boldsymbol{A})$.
- A multiplicação de uma coluna/linha com $\lambda \in \mathbb{R}$ dimensiona $det(\boldsymbol{A})$ por λ. Em particular, $det(\lambda \boldsymbol{A}) = \lambda^n det(\boldsymbol{A})$.
- Trocar duas linhas/colunas altera o sinal de $det(\boldsymbol{A})$.

Devido às últimas três propriedades, podemos usar a Eliminação de Gauss para calcular $det(\boldsymbol{A})$ trazendo $\boldsymbol{A}$ para a forma escalonada por linhas. Podemos parar a eliminação quando tivermos $\boldsymbol{A}$ na forma de uma matriz triangular superior uma vez que o determinante de uma matriz triangular é o produto dos elementos diagonais.

- Uma matriz quadrada $\boldsymbol{A} \in \mathbb{R}^{n \times n}$ tem $det(\boldsymbol{A}) \neq \mathbf{0}$ se e somente se $rk(\boldsymbol{A}) = n$.
- Em outras palavras, $\boldsymbol{A}$ é invertível se e somente se tiver posto completo.
- O posto é o número máximo de colunas/linhas linearmente independentes de uma Matriz. Se a matriz for quadrada, e o posto for igual ao número de linhas, então a matriz tem <u>posto completo</u>.

Traço

O traço de uma matriz quadrada $\boldsymbol{A} \in \mathbb{R}^{n\times n}$ é definido como:

$$\mathrm{tr}(\boldsymbol{A}) := \sum_{i=1}^{n} a_{ii} \,,$$

, isto é, o traço é a soma dos elementos da diagonal principal de $\boldsymbol{A}$.

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ a_{41} & a_{42} & a_{43} & a_{44} \end{pmatrix}$$

$$\mathrm{tr}\, A = a_{11} + a_{22} + a_{33} + a_{44}$$

O Traço satisfaz as seguintes propriedades:

- $tr(\boldsymbol{A} + \boldsymbol{B}) = tr(\boldsymbol{A}) + tr(\boldsymbol{B})$ para $\boldsymbol{A}, \boldsymbol{B} \in \mathbb{R}^{n \times n}$
- $tr(\alpha \boldsymbol{A}) = \alpha \cdot tr(\boldsymbol{A})$, para $\alpha \in \mathbb{R}$ e $\boldsymbol{A} \in \mathbb{R}^{n \times n}$
- $tr(\boldsymbol{I}_n) = n$
- $tr(\boldsymbol{AB}) = tr(\boldsymbol{BA})$ para $\boldsymbol{A} \in \mathbb{R}^{n \times k}$, $\boldsymbol{B} \in \mathbb{R}^{k \times n}$

$$A = \begin{bmatrix} 1 & 3 & 5 \\ 7 & 9 & 2 \\ 4 & 6 & 8 \end{bmatrix}$$

tr (A) = 18

tr (AB) = tr (BA)

tr (kA) = k tr (A)

tr (A) = tr (A^T)

As propriedades do traço dos produtos de matrizes são mais gerais. Especificamente, o traço é invariante sob permutações cíclicas, ou seja,

$$tr(\boldsymbol{AKL}) = tr(\boldsymbol{KLA})$$

para matrizes $\boldsymbol{A} \in \mathbb{R}^{a \times n}, \boldsymbol{K} \in \mathbb{R}^{k \times l}, \boldsymbol{L} \in \mathbb{R}^{l \times a}$.

Esta propriedade generaliza para produtos de um número arbitrário de matrizes.

Dado um mapeamento linear $\Phi : V \rightarrow V$, onde V é um espaço vetorial, definimos o traço desse mapa usando o traço da representação matricial de Φ. Para uma base dada de V, podemos descrever Φ por meio da matriz de transformação $\boldsymbol{A}$. Então, o traço de Φ é o traço de $\boldsymbol{A}$.

Para uma base diferente de V, vale a pena mencionar que a matriz de transformação correspondente $\boldsymbol{B}$ de Φ pode ser obtida por uma mudança de base da forma $\boldsymbol{S}^{-1}\boldsymbol{A}\boldsymbol{S}$ para um $\boldsymbol{S}$ adequado.

Para o traço correspondente de Φ, isso significa que: $tr(\boldsymbol{B}) = tr(\boldsymbol{S}^{-1}\boldsymbol{A}\boldsymbol{S}) = tr(\boldsymbol{A}\boldsymbol{S}\boldsymbol{S}^{-1}) = tr(\boldsymbol{A})$.

Portanto, embora as representações matriciais de mapeamentos lineares sejam dependentes da base, o traço de um mapeamento linear Φ é independente da base.

Polinômio Característico

Para $\lambda \in \mathbb{R}$ e uma matriz quadrada $\boldsymbol{A} \in \mathbb{R}^{n\times n}$

$$pA(\lambda) := det(\boldsymbol{A} - \lambda \boldsymbol{I})$$
$$= c_0 + c_1\lambda + c_2\lambda^2 + \cdots + c_{n-1}\lambda^{n-1} + (-1)^n\lambda^n ,$$

$c_0, \ldots, c_{n-1} \in \mathbb{R}$.

Adicionalmente:

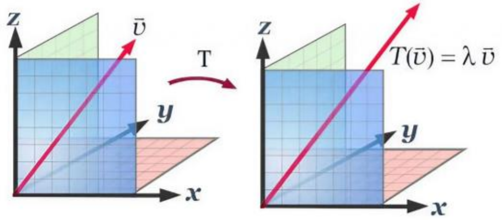
$$c_0 = det(\boldsymbol{A})$$
$$c_{n-1} = (-1)^{n-1}tr(\boldsymbol{A})$$

Este polinômio característico nos permitirá calcular autovalores e autovetores.

Autovalores e Autovetores

Assim como o Determinante e o Traço, Autovalores e Autovetores são características de Matrizes Quadradas.

- Autovalores são valores numéricos que descrevem a **magnitude** da transformação efetuada por uma matriz em uma determinada direção.
- Autovetores são vetores que fornecem a **direção** da transformação na qual a matriz age apenas esticando ou comprimindo o vetor, sem alterar sua orientação.



Motivação: Autovalores e Autovetores em PCA

Análise de Componentes Principais (PCA): é uma técnica que transforma dados de dimensões altas em dimensões inferiores, mantendo o máximo de informações possível.

cod_ibge	distritos	renda	quota	escolaridade	idade	mortalidade	txcresc	causasext	favel	denspop
1	Água Rasa	1961	3.461.999.893	7.599.999.905	32	1.385.999.966	-1.840.000.033	5.297.999.954	0.0	1.256.100.006
12	Alto de Pinheiros	4180	7.595.999.908	8.399.999.619	33	8.680.000.305	-2.519.999.981	3.856.999.969	0.689999998	5.756.000.137
23	Anhanguera	1093	4.5	5.800.000.191	23	1.535.999.966	1.812.000.084	2.268.000.031	0.0	8.569.999.695
34	Aricanduva	1311	2.102.000.046	6.800.000.191	27	1.843.000.031	-1.070.000.052	7.622.000.122	5.380.000.114	1.385.399.933
45	Artur Alvim	1248	1.590.999.985	7.0	27	1.972.999.954	-1.399.999.976	67.25	4.110.000.134	1.673.999.939
56	Barra Funda	2359	3.431.000.137	8.0	31	8.619.999.886	-2.140.000.105	3.790.000.153	4.849.999.905	2.560.000.038
67	Bela Vista	2400	5.365.000.153	8.399.999.619	32	1.240.999.985	-200.999.999	5.986.000.061	0.0	2.496.000.061
78	Belém	1813	3.198.999.977	7.699.999.809	32	1.127.000.046	-3.039.999.962	7.894.000.244	2.680.000.067	7.097.000.122
					...					
90	Vila Mariana	3312	6.354.000.092	8.600.000.381	34	943.999.958	-1.330.000.043	3.318.000.031	0.870000005	1.444.100.037
91	Vila Matilde	1530	26.0	7.400.000.095	29	1.912.999.916	-1.830.000.043	5.872.999.954	0.0	1.116.999.969
92	Vila Medeiros	1405	1.976.000.023	6.800.000.191	27	1.543.000.031	-1.409.999.967	7.798.000.336	249.000.001	1.889.299.927
93	Vila Prudente	1755	3.208.000.183	7.199.999.809	30	1.435.999.966	-2.549.999.952	6.651.000.214	7.429.999.828	1.014.400.024
94	Vila Sônia	2970	4.140.999.985	7.400.000.095	27	1.676.000.023	-0.899999976	7.468.000.031	1.493.000.031	8.012.000.275

1. Padronização dos Dados: Se necessário, padronize os dados para que todas as variáveis tenham média zero e desvio padrão unitário.
2. Cálculo da Matriz de Covariância: Calcule a matriz de covariância dos dados padronizados para capturar as relações entre as variáveis.
3. Cálculo dos <u>Autovalores</u> e <u>Autovetores</u>: Calcule os autovalores e autovetores da matriz de covariância.
4. Ordenação dos Autovalores: Ordene os <u>autovalores</u> em ordem decrescente para identificar as componentes principais mais importantes.
5. Seleção das Componentes Principais: Selecione as primeiras k colunas dos <u>autovetores</u> correspondentes aos k maiores <u>autovalores</u> para obter as k componentes principais.
6. Projeção dos Dados: Projete os dados originais sobre as k componentes principais selecionadas para obter uma representação de menor dimensionalidade.

Seja $\boldsymbol{A} \in \mathbb{R}^{n\times n}$ uma matriz quadrada. Então $\lambda \in \mathbb{R}$ é um autovalor de $\boldsymbol{A}$ e $\boldsymbol{x} \in \mathbb{R}^n\backslash\{0\}$ é o autovetor correspondente de $\boldsymbol{A}$ se:

$$\boldsymbol{Ax} = \lambda\boldsymbol{x}$$

As seguintes afirmações são equivalentes:

- Se λ é um autovalor de $\boldsymbol{A} \in \mathbb{R}^{n\times n}$.
- Existe um $\boldsymbol{x} \in \mathbb{R}^n\backslash\{0\}$ tal que $\boldsymbol{Ax} = \lambda\boldsymbol{x}$, ou de forma equivalente, $(\boldsymbol{A} - \lambda\boldsymbol{I}_n)\boldsymbol{x} = 0$ pode ser resolvido de forma não trivial, ou seja, $\boldsymbol{x} \neq 0$.
- $rk(\boldsymbol{A} - \lambda\boldsymbol{I}_n) < n$.
- $det(\boldsymbol{A} - \lambda\boldsymbol{I}_n) = 0$.

Colinearidade e Coodireção

Dois vetores que apontam na mesma direção são chamados de co-direcionados.

Dois vetores que apontam na mesma direção ou em direções opostas são chamados colineares.

Se $\boldsymbol{x}$ é um autovetor de $\boldsymbol{A}$ associado ao autovalor λ, então para qualquer $\mathrm{c} \in \mathbb{R}\backslash\{0\}$ temos que $c\boldsymbol{x}$ é um autovetor de $\boldsymbol{A}$ com o mesmo autovalor, pois:

$$\boldsymbol{A}(c\boldsymbol{x}) = c\boldsymbol{A}\boldsymbol{x} = c\lambda\boldsymbol{x} = \lambda(c\boldsymbol{x})$$

Assim, todos os vetores que são colineares a $\boldsymbol{x}$ também são autovetores de $\boldsymbol{A}$.

$\lambda \in \mathbb{R}$ é um autovalor de $\boldsymbol{A} \in \mathbb{R}^{n\times n}$ se e somente se λ é uma raiz do polinômio característico $p_{\boldsymbol{A}}(\lambda)$ de $\boldsymbol{A}$.

Considerando que uma matriz quadrada $\boldsymbol{A}$ tenha um autovalor λ_i. A *multiplicidade algébrica* de λ_i é o número de vezes que a raiz aparece no polinômio característico.

O conjunto de todos os autovetores de $\boldsymbol{A}$ associados a um autovalor λ abrange um subespaço de $\mathbb{R}^n$, que é chamado de <u>autoespaço</u> de $\boldsymbol{A}$ com relação λ e é representado por E_λ.

O conjunto de todos os autovalores de $\boldsymbol{A}$ é chamado de <u>autoespectro</u>, ou apenas espectro, de $\boldsymbol{A}$.

Se λ é um autovalor de $\boldsymbol{A} \in \mathbb{R}^{n\times n}$ então o autoespaço correspondente E_λ é o espaço de soluções do sistema homogêneo de equações lineares $(\boldsymbol{A}-\lambda\boldsymbol{I})\boldsymbol{x} = 0$.

Geometricamente, o autovetor correspondente a um autovalor não nulo aponta em uma direção que é esticada pela transformação linear. O autovalor é o fator pelo qual ela é esticada. Se o autovalor for negativo, a direção é invertida.

O Caso da Matriz de Identidade

A matriz identidade $\boldsymbol{I} \in \mathbb{R}^{n\times n}$ possui polinomial característico $p_{\boldsymbol{I}}(\lambda) = det(\boldsymbol{I} - \lambda\boldsymbol{I}) = (1 - \lambda)^n = 0$, o qual possui apenas um autovalor $\lambda = 1$ que ocorre n vezes. Além disso, $\boldsymbol{Ix} = \lambda\boldsymbol{x} = 1\boldsymbol{x}$ é válido para todos os vetores $\boldsymbol{x} \in \mathbb{R}^n \backslash \{0\}$. Devido a isso, o único autoespaço E_1 da matriz identidade abrange n dimensões, e todos os n vetores da base padrão de R^n são autovetores de I.

Os autovalores de uma **matriz triangular** são os elementos da **diagonal**

$$A = \begin{bmatrix} 1/2 & 0 & 0 \\ -1 & 2/3 & 0 \\ 5 & -8 & -1/4 \end{bmatrix}$$

Propriedades úteis

- Uma matriz $\boldsymbol{A}$ e sua transposta $\boldsymbol{A}^T$ possuem os mesmos autovalores, mas não necessariamente os mesmos autovetores.
- O autoespaço E_λ é o espaço nulo de $\boldsymbol{A} - \lambda\boldsymbol{I}$, uma vez que:

$$\boldsymbol{A}\boldsymbol{x} = \lambda\boldsymbol{x} \iff \boldsymbol{A}\boldsymbol{x} - \lambda\boldsymbol{x} = 0$$

$$\iff (\boldsymbol{A} - \lambda\boldsymbol{I})\boldsymbol{x} = 0 \iff \boldsymbol{x} \in \ker(\boldsymbol{A} - \lambda\boldsymbol{I}).$$

- Matrizes semelhantes possuem os mesmos autovalores. Portanto, um mapeamento linear Φ possui autovalores independentes da escolha da base de sua matriz de transformação. Isso torna os autovalores, juntamente com o determinante e o traço, parâmetros característicos principais de uma aplicação linear, pois são todos invariantes em relação à mudança de base.
- Matrizes simétricas e definidas positivas sempre têm autovalores reais e positivos.

Computando autovalores, autovetores e autoespaços

Vamos encontrar os autovalores e autovetores da matriz 2 × 2:

$$\boldsymbol{A} = \begin{bmatrix} 4 & 2 \\ 1 & 3 \end{bmatrix}$$

Passo 1: Polinômio característico. A partir de nossa definição do autovetor $x \neq 0$ e autovalor λ de $\boldsymbol{A}$, haverá um vetor tal que

$\boldsymbol{A}x = \lambda \boldsymbol{x}$, ou seja, $(\boldsymbol{A} - \lambda \boldsymbol{I})\boldsymbol{x} = 0$. Desde que $x \neq 0$, isso requer que o núcleo (espaço nulo) de $\boldsymbol{A} - \lambda \boldsymbol{I}$ contenha mais elementos do que apenas 0. Isso significa que $\boldsymbol{A} - \lambda \boldsymbol{I}$ não é invertível e, portanto, $det(\boldsymbol{A} - \lambda \boldsymbol{I}) = 0$.

Portanto, precisamos calcular as raízes do polinômio característico para encontrar os autovalores.

Passo 2: Autovalores. O polinômio característico é

$$\begin{aligned} p_{\boldsymbol{A}}(\lambda) &= \det(\boldsymbol{A} - \lambda \boldsymbol{I}) \\ &= \det\left(\begin{bmatrix} 4 & 2 \\ 1 & 3 \end{bmatrix} - \begin{bmatrix} \lambda & 0 \\ 0 & \lambda \end{bmatrix}\right) = \begin{vmatrix} 4-\lambda & 2 \\ 1 & 3-\lambda \end{vmatrix} \\ &= (4-\lambda)(3-\lambda) - 2 \cdot 1\,. \end{aligned}$$

Fatorando o polinômio característico e obtemos:

$$p(\lambda) = (4-\lambda)(3-\lambda) - 2 \cdot 1 = 10 - 7\lambda + \lambda^2 = (2-\lambda)(5-\lambda)$$

obtendo as raízes $\lambda_1 = 2$ e $\lambda_2 = 5$.

Passo 3: Autovetores e autoespaços. Encontramos os autovetores que correspondem a esses autovalores examinando os vetores $\boldsymbol{x}$ tais que

$$\begin{bmatrix} 4-\lambda & 2 \\ 1 & 3-\lambda \end{bmatrix} \boldsymbol{x} = \mathbf{0}\,.$$

Para $\lambda = 5$, obtemos

$$\begin{bmatrix} 4-5 & 2 \\ 1 & 3-5 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} -1 & 2 \\ 1 & -2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \mathbf{0}\,.$$

Resolvemos este sistema homogêneo e obtemos um espaço de soluções.

$$E_5 = \mathrm{span}[\begin{bmatrix} 2 \\ 1 \end{bmatrix}]\,.$$

Este autoespaço é unidimensional, pois possui um único vetor de base.

Analogamente, encontramos o autovetor para $\lambda = 2$ resolvendo o sistema homogêneo de equações.

$$\begin{bmatrix} 4-2 & 2 \\ 1 & 3-2 \end{bmatrix} \boldsymbol{x} = \begin{bmatrix} 2 & 2 \\ 1 & 1 \end{bmatrix} \boldsymbol{x} = \mathbf{0}\,.$$

Isso significa que qualquer vetor $\boldsymbol{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$, onde $x_2 = -x_1$, como $\begin{bmatrix} 1 \\ -1 \end{bmatrix}$, é um autovetor com autovalor 2. O autoespaço correspondente é dado por

$$E_2 = \operatorname{span}[\begin{bmatrix} 1 \\ -1 \end{bmatrix}]\,.$$

Os dois autoespaços E_5 e E_2 apresentados são unidimensionais, pois cada um é gerado por um único vetor.

$$E_5 = \text{span}[\begin{bmatrix} 2 \\ 1 \end{bmatrix}] \qquad E_2 = \text{span}[\begin{bmatrix} 1 \\ -1 \end{bmatrix}]$$

No entanto, em outros casos, podemos ter múltiplos autovalores idênticos e o autoespaço pode ter mais de uma dimensão.

Sendo λ_i um autovalor de uma matriz quadrada A. Então, a multiplicidade geométrica de λ_i é o número de autovetores linearmente independentes associados a λ_i. Em outras palavras, é a dimensionalidade do autoespaço gerado pelos autovetores associados a λ_i.

A multiplicidade geométrica de um autovalor específico deve ser pelo menos um, pois cada autovalor tem pelo menos um autovetor associado e não pode exceder sua multiplicidade algébrica, mas pode ser menor.

Exemplo:

A matriz $\boldsymbol{A} = \begin{bmatrix} 2 & 1 \\ 0 & 2 \end{bmatrix}$ tem dois autovalores repetidos $\lambda_1 = \lambda_2 = 2$ e uma multiplicidade algébrica de 2. O autovalor tem, no entanto, apenas um autovetor $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$ e, assim, multiplicidade geométrica 1.

Intuição Gráfica em Duas Dimensões

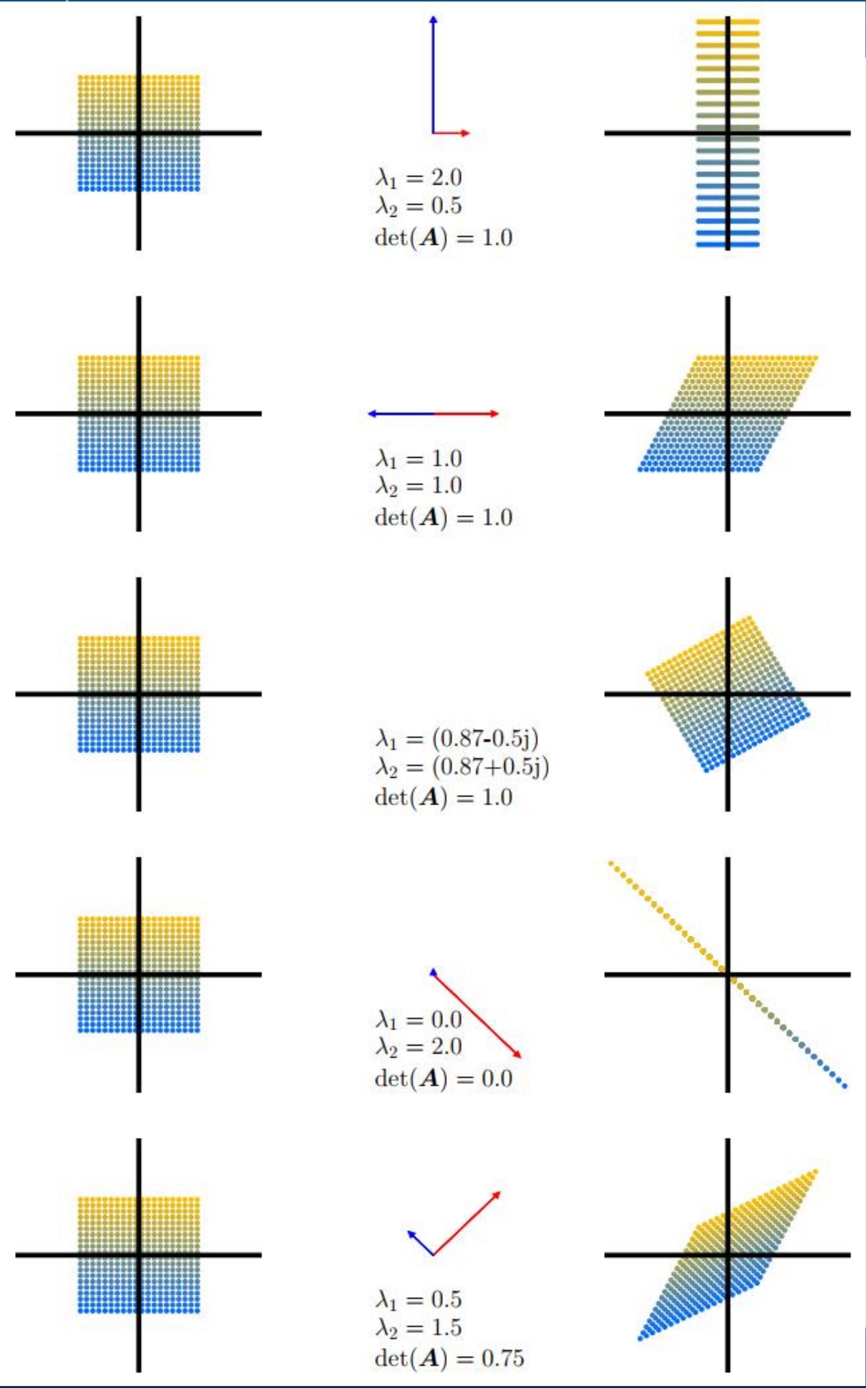
$$\boldsymbol{A}_1 = \begin{bmatrix} \frac{1}{2} & 0 \\ 0 & 2 \end{bmatrix} \quad \det(\boldsymbol{A}_1) = 1 = 2 \cdot \frac{1}{2}$$

$$\boldsymbol{A}_2 = \begin{bmatrix} 1 & \frac{1}{2} \\ 0 & 1 \end{bmatrix} \quad \det(\boldsymbol{A}_2) = 1$$

$$\boldsymbol{A}_3 = \begin{bmatrix} \cos(\frac{\pi}{6}) & -\sin(\frac{\pi}{6}) \\ \sin(\frac{\pi}{6}) & \cos(\frac{\pi}{6}) \end{bmatrix} = \frac{1}{2}\begin{bmatrix} \sqrt{3} & -1 \\ 1 & \sqrt{3} \end{bmatrix}$$

$$\boldsymbol{A}_4 = \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix} \quad \det(\boldsymbol{A}_4) = 2 \cdot 0$$

$$\boldsymbol{A}_5 = \begin{bmatrix} 1 & \frac{1}{2} \\ \frac{1}{2} & 1 \end{bmatrix} \quad \det(\boldsymbol{A}_5) = 1.5 \cdot 0.5$$



Autoespectro de uma Rede Neural Biológica

- Análise e aprendizado através de dados em rede
- Conectividade entre nós de rede
- Se está conectado ou não
- Aplicações em Ciência de Dados

Os autovetores $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n$ de uma matriz $\boldsymbol{A} \in \mathbb{R}^{n \times n}$ com n autovalores distintos $\lambda_1, \ldots, \lambda_n$ são linearmente independentes.

Um matriz $\boldsymbol{A} \in \mathbb{R}^{n \times n}$ é defeituosa se ela possui menos de n autovetores linearmente independentes, não necessariamente com autovalores distintos, mas que ainda assim formam bases de $\mathbb{R}^n$.

Dada uma matriz $\boldsymbol{A} \in \mathbb{R}^{m \times n}$, sempre podemos obter uma matriz simétrica, semidefinida positiva $S \in \mathbb{R}^{n \times n}$ definindo $S := \boldsymbol{A}^T \boldsymbol{A}$

Se $rk(\boldsymbol{A}) = n$, então $S := \boldsymbol{A}^T \boldsymbol{A}$ é simétrica, definita positiva.

$$\begin{bmatrix} 10 & 2 & 1 \\ 2 & 1 & -2 \\ 1 & -2 & 10 \end{bmatrix}$$

Teorema Espectral

Se $\boldsymbol{A} \in \mathbb{R}^{n \times n}$ for simétrica, existe uma base ortonormal do espaço vetorial correspondente V que consiste de autovetores de $\boldsymbol{A}$, e cada autovalor é real.

Uma implicação direta do teorema espectral é que a decomposição de autovalores de uma matriz simétrica $\boldsymbol{A}$ existe (com autovalores reais) e que podemos encontrar uma base ortogonal de autovetores de modo que $\boldsymbol{A} = \boldsymbol{PDP}^T$.

- $\boldsymbol{P}$ é a matriz cujas colunas são os autovetores de $\boldsymbol{A}$.
- $\boldsymbol{D}$ é a matriz diagonal formada pelos autovalores de $\boldsymbol{A}$.

Considere a matriz: $\boldsymbol{A} = \begin{bmatrix} 3 & 2 & 2 \\ 2 & 3 & 2 \\ 2 & 2 & 3 \end{bmatrix}$

O polinômio característico de A é $p_{\boldsymbol{A}}(\lambda) = -(\lambda - 1)^2(\lambda - 7)$, então os autovalores são $\lambda_1 = 1$ e $\lambda_2 = 7$.

Seguindo o procedimento padrão para cálculo de autovetores, obtemos:

$$E_1 = \operatorname{span}[\underbrace{\begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}}_{=:\boldsymbol{x}_1}, \underbrace{\begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}}_{=:\boldsymbol{x}_2}], \quad E_7 = \operatorname{span}[\underbrace{\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}}_{=:\boldsymbol{x}_3}].$$

O teorema espectral afirma que existe uma base ortogonal na decomposição de uma matriz simétrica, mas a base que temos não é ortogonal.

- $\boldsymbol{x}_3$ é ortogonal tanto a $\boldsymbol{x}_1$ quanto a $\boldsymbol{x}_2$.
- $\boldsymbol{x}_1$ e $\boldsymbol{x}_2$ não são ortogonais entre si uma vez que: $\boldsymbol{x}_1^T\boldsymbol{x}_2 = 1 \neq 0$.

Aproveitamos o fato de que $\boldsymbol{x}_1, \boldsymbol{x}_2$ são autovetores associados ao mesmo autovalor λ para construir uma base ortogonal. Considerando que qualquer $\alpha, \beta \in \mathbb{R}$ temos que:

$$\boldsymbol{A}(\alpha\boldsymbol{x}_1 + \beta\boldsymbol{x}_2) = \boldsymbol{A}\boldsymbol{x}_1\alpha + \boldsymbol{A}\boldsymbol{x}_2\beta = \lambda(\alpha\boldsymbol{x}_1 + \beta\boldsymbol{x}_2)$$

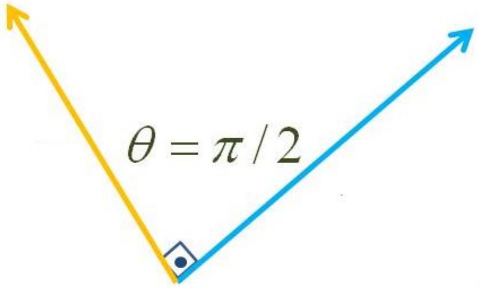
Utilizando o algoritmo de Gram-Schmidt conseguimos obter novos autovetores associados

$$\boldsymbol{x}_1' = \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}, \quad \boldsymbol{x}_2' = \frac{1}{2}\begin{bmatrix} -1 \\ -1 \\ 2 \end{bmatrix}$$

Utilizando o algoritmo de Gram-Schmidt conseguimos obter novos autovetores associados a λ_1:

$$\boldsymbol{x}_1' = \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}, \quad \boldsymbol{x}_2' = \frac{1}{2}\begin{bmatrix} -1 \\ -1 \\ 2 \end{bmatrix},$$

que são ortogonais entre si, ortogonais a $\boldsymbol{x}_3$ e autovetores de $\boldsymbol{A}$ associados a $\lambda_1 = 1$.



PageRank - Ranqueamento de Páginas do Google

O PageRank foi desenvolvido na Universidade de Stanford por Larry Page e Sergey Brin em 1996.

O PageRank utiliza o autovetor correspondente ao autovalor máximo de uma matriz $\boldsymbol{A}$ para determinar a classificação de uma página nos resultados de pesquisa.

O algoritmo calcula o peso (importância) $x_i \geqslant 0$ de um site da web a_i contando o número de páginas que fazem *link* para a_i considerando no calculo desse peso também o peso das páginas quem fazem *link*.

O que aprendemos?

Determinante: é uma medida da "escala" de uma transformação linear representada pela matriz.

Traço: é a soma dos elementos da diagonal principal. Ele fornece informações sobre a soma dos autovalores da matriz.

Autovalores e Autovetores: são os características especiais que utilizamos para entender como uma matriz age sobre o espaço vetorial, identificando as direções principais ao longo das quais ocorrem as transformações e os fatores pelos quais essas transformações ocorrem.

Essas características são fundamentais na teoria das matrizes e têm aplicações em uma variedade de áreas, incluindo física, engenharia, ciência da computação, estatística e muito mais.

CCO0446 - TÓPICOS ESPECIAIS EM FUNDAMENTOS DE COMPUTAÇÃO 1

4.3 - Decomposição de Cholesky

4.4 - Decomposição em autovalores e autovetores e diagonalização

Alunos:
André Morais Riccioppo
Fábio Alves Martins Pereira
Gabriel Matheus Faria de Almeida
Hudson de Paula Romualdo

4.3 - Decomposição de Cholesky

4.3 - Decomposição de Cholesky

- Existem muitas maneiras de fatorar tipos especiais de matrizes que encontramos frequentemente no aprendizado de máquina.
- Nos números reais positivos, temos a operação de raiz quadrada que nos dá uma decomposição do número em componentes idênticos, por exemplo, $9 = 3 \cdot 3$.
- Para matrizes, precisamos ter cuidado para calcular uma operação semelhante a raiz quadrada.
- Para matrizes definidas positivas e simétricas (ver Seção 3.2.3), podemos escolher entre uma série de operações equivalentes de raiz quadrada.
- A decomposição de Cholesky, ou fatoração de Cholesky é uma delas e ela nos fornece uma equivalente à raiz quadrada quando temos matrizes definidas positivas e simétricas

4.3 - Decomposição de Cholesky

Theorem 4.18 (Cholesky Decomposition). *A symmetric, positive definite matrix* $\boldsymbol{A}$ *can be factorized into a product* $\boldsymbol{A} = \boldsymbol{L}\boldsymbol{L}^\top$, *where* $\boldsymbol{L}$ *is a lower-triangular matrix with positive diagonal elements:*

$$\begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{bmatrix} = \begin{bmatrix} l_{11} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ l_{n1} & \cdots & l_{nn} \end{bmatrix} \begin{bmatrix} l_{11} & \cdots & l_{n1} \\ \vdots & \ddots & \vdots \\ 0 & \cdots & l_{nn} \end{bmatrix} . \tag{4.44}$$

$\boldsymbol{L}$ *is called the Cholesky factor of* $\boldsymbol{A}$, *and* $\boldsymbol{L}$ *is unique.*

4.3 - Decomposição de Cholesky

$$\boldsymbol{A} = \begin{bmatrix} a_{11} & a_{21} & a_{31} \\ a_{21} & a_{22} & a_{32} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} = \boldsymbol{L}\boldsymbol{L}^{\top} = \begin{bmatrix} l_{11} & 0 & 0 \\ l_{21} & l_{22} & 0 \\ l_{31} & l_{32} & l_{33} \end{bmatrix} \begin{bmatrix} l_{11} & l_{21} & l_{31} \\ 0 & l_{22} & l_{32} \\ 0 & 0 & l_{33} \end{bmatrix}$$

$$\boldsymbol{A} = \begin{bmatrix} l_{11}^2 & l_{21}l_{11} & l_{31}l_{11} \\ l_{21}l_{11} & l_{21}^2 + l_{22}^2 & l_{31}l_{21} + l_{32}l_{22} \\ l_{31}l_{11} & l_{31}l_{21} + l_{32}l_{22} & l_{31}^2 + l_{32}^2 + l_{33}^2 \end{bmatrix}.$$

4.3 - Decomposição de Cholesky

$$\boldsymbol{A} = \begin{bmatrix} a_{11} & a_{21} & a_{31} \\ a_{21} & a_{22} & a_{32} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$$

$$\boldsymbol{A} = \begin{bmatrix} l_{11}^2 & l_{21}l_{11} & l_{31}l_{11} \\ l_{21}l_{11} & l_{21}^2 + l_{22}^2 & l_{31}l_{21} + l_{32}l_{22} \\ l_{31}l_{11} & l_{31}l_{21} + l_{32}l_{22} & l_{31}^2 + l_{32}^2 + l_{33}^2 \end{bmatrix} .$$

Se compararmos o lado esquerdo de e o lado direito, podemos ver que existe um padrão para os elementos da diagonal

$$l_{11} = \sqrt{a_{11}}\,, \quad l_{22} = \sqrt{a_{22} - l_{21}^2}\,, \quad l_{33} = \sqrt{a_{33} - (l_{31}^2 + l_{32}^2)}$$

4.3 - Decomposição de Cholesky

$$\boldsymbol{A} = \begin{bmatrix} a_{11} & a_{21} & a_{31} \\ a_{21} & a_{22} & a_{32} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$$

$$\boldsymbol{A} = \begin{bmatrix} l_{11}^2 & l_{21}l_{11} & l_{31}l_{11} \\ l_{21}l_{11} & l_{21}^2 + l_{22}^2 & l_{31}l_{21} + l_{32}l_{22} \\ l_{31}l_{11} & l_{31}l_{21} + l_{32}l_{22} & l_{31}^2 + l_{32}^2 + l_{33}^2 \end{bmatrix} .$$

E a mesma coisa vale para os elementos abaixo da diagonal, que também seguem um padrão

$$l_{21} = \frac{1}{l_{11}} a_{21} \, , \quad l_{31} = \frac{1}{l_{11}} a_{31} \, , \quad l_{32} = \frac{1}{l_{22}} (a_{32} - l_{31}l_{21}) \, .$$

4.3 - Decomposição de Cholesky

Calculamos apeans os elementos da diagonal inferior, pois a matriz é simétrica

$$l_{11} = \sqrt{a_{11}}\,, \quad l_{22} = \sqrt{a_{22} - l_{21}^2}\,, \quad l_{33} = \sqrt{a_{33} - (l_{31}^2 + l_{32}^2)}$$

$$\boldsymbol{A} = \begin{bmatrix} a_{11} & a_{21} & a_{31} \\ a_{21} & a_{22} & a_{32} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} = \boldsymbol{L}\boldsymbol{L}^\top = \begin{bmatrix} l_{11} & 0 & 0 \\ l_{21} & l_{22} & 0 \\ l_{31} & l_{32} & l_{33} \end{bmatrix} \begin{bmatrix} l_{11} & l_{21} & l_{31} \\ 0 & l_{22} & l_{32} \\ 0 & 0 & l_{33} \end{bmatrix}$$

4.3 - Decomposição de Cholesky

A principal conclusão é que podemos tirar é que:

Podemos calcular retroativamente quais deveriam ser os componentes l_{ij} para L, se tivermos os valores a_{ij} para A, e os valores previamente calculados de l_{ij}.

$$l_{11} = \sqrt{a_{11}}\,, \quad l_{22} = \sqrt{a_{22} - l_{21}^2}\,, \quad l_{33} = \sqrt{a_{33} - (l_{31}^2 + l_{32}^2)}$$

$$l_{21} = \frac{1}{l_{11}} a_{21}\,, \quad l_{31} = \frac{1}{l_{11}} a_{31}\,, \quad l_{32} = \frac{1}{l_{22}} (a_{32} - l_{31} l_{21})\,.$$

4.3 - Decomposição de Cholesky

2014). The Cholesky decomposition also allows us to compute determinants very efficiently. Given the Cholesky decomposition $\boldsymbol{A} = \boldsymbol{L}\boldsymbol{L}^\top$, we know that $\det(\boldsymbol{A}) = \det(\boldsymbol{L})\det(\boldsymbol{L}^\top) = \det(\boldsymbol{L})^2$. Since $\boldsymbol{L}$ is a triangular matrix, the determinant is simply the product of its diagonal entries so that $\det(\boldsymbol{A}) = \prod_i l_{ii}^2$. Thus, many numerical software packages use the

4.4 - Decomposição em autovalores e autovetores e diagonalização

4.4 - Decomposição em autovalores e autovetores e diagonalização

Uma matriz diagonal é uma matriz que tem valor zero em todos os elementos fora da diagonal, ou seja, eles têm a forma

$$\boldsymbol{D} = \begin{bmatrix} c_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & c_n \end{bmatrix}.$$

4.4 - Decomposição em autovalores e autovetores e diagonalização

$$D = \begin{bmatrix} c_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & c_n \end{bmatrix}.$$

As matrizes diagonais são importantes pois elas permitem o cálculo rápido de:

Determinantes

Potências

inversas.

O determinante é o produto dos elementos da diagonal
A potência D^k é dada por cada elemento diagonal elevado à potência k,
O inverso D^{-1} é o inverso de seus elementos diagonais se todos eles forem diferentes de zero.

4.4 - Decomposição em autovalores e autovetores e diagonalização

$$\boldsymbol{D} = \begin{bmatrix} c_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & c_n \end{bmatrix}.$$

Agora vamos ver como transformar matrizes em sua forma diagonal.

Essa transformação é uma aplicação importante da mudança de base e de autovalores, que foram discutidos nas seções anteriores.

4.4 - Decomposição em autovalores e autovetores e diagonalização

$$\boldsymbol{D} = \begin{bmatrix} c_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & c_n \end{bmatrix}.$$

Lembremos que que duas matrizes A, D são “semelhantes” (Definição 2.22) se existir uma matriz invertível P , tal que D = $P^{-1}AP$.

Mais especificamente, nós queremos examinar matrizes A que são semelhantes às matrizes diagonais D que contêm os autovalores de A na diagonal.

4.4 - Decomposição em autovalores e autovetores e diagonalização

Definition 4.19 (Diagonalizable). A matrix $\boldsymbol{A} \in \mathbb{R}^{n \times n}$ is *diagonalizable* if it is similar to a diagonal matrix, i.e., if there exists an invertible matrix $\boldsymbol{P} \in \mathbb{R}^{n \times n}$ such that $\boldsymbol{D} = \boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}$.

4.4 - Decomposição em autovalores e autovetores e diagonalização

Let $\boldsymbol{A} \in \mathbb{R}^{n \times n}$, let $\lambda_1, \ldots, \lambda_n$ be a set of scalars, and let $\boldsymbol{p}_1, \ldots, \boldsymbol{p}_n$ be a set of vectors in $\mathbb{R}^n$. We define $\boldsymbol{P} := [\boldsymbol{p}_1, \ldots, \boldsymbol{p}_n]$ and let $\boldsymbol{D} \in \mathbb{R}^{n \times n}$ be a diagonal matrix with diagonal entries $\lambda_1, \ldots, \lambda_n$. Then we can show that

$$\boldsymbol{AP} = \boldsymbol{PD} \tag{4.50}$$

if and only if $\lambda_1, \ldots, \lambda_n$ are the eigenvalues of $\boldsymbol{A}$ and $\boldsymbol{p}_1, \ldots, \boldsymbol{p}_n$ are corresponding eigenvectors of $\boldsymbol{A}$.

4.4 - Decomposição em autovalores e autovetores e diagonalização

We can see that this statement holds because

$$\boldsymbol{AP} = \boldsymbol{A}[\boldsymbol{p}_1, \ldots, \boldsymbol{p}_n] = [\boldsymbol{Ap}_1, \ldots, \boldsymbol{Ap}_n] , \tag{4.51}$$

$$\boldsymbol{PD} = [\boldsymbol{p}_1, \ldots, \boldsymbol{p}_n] \begin{bmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{bmatrix} = [\lambda_1 \boldsymbol{p}_1, \ldots, \lambda_n \boldsymbol{p}_n] . \tag{4.52}$$

Thus, (4.50) implies that

$$\boldsymbol{Ap}_1 = \lambda_1 \boldsymbol{p}_1 \tag{4.53}$$

$$\vdots$$

$$\boldsymbol{Ap}_n = \lambda_n \boldsymbol{p}_n . \tag{4.54}$$

Therefore, the columns of $\boldsymbol{P}$ must be eigenvectors of $\boldsymbol{A}$.

4.4 - Decomposição em autovalores e autovetores e diagonalização

Our definition of diagonalization requires that $\boldsymbol{P} \in \mathbb{R}^{n \times n}$ is invertible, i.e., $\boldsymbol{P}$ has full rank (Theorem 4.3). This requires us to have n linearly independent eigenvectors $\boldsymbol{p}_1, \ldots, \boldsymbol{p}_n$, i.e., the $\boldsymbol{p}_i$ form a basis of $\mathbb{R}^n$.

4.4 - Decomposição em autovalores e autovetores e diagonalização

Theorem 4.20 (Eigendecomposition). *A square matrix* $\boldsymbol{A} \in \mathbb{R}^{n \times n}$ *can be factored into*

$$\boldsymbol{A} = \boldsymbol{P}\boldsymbol{D}\boldsymbol{P}^{-1}, \qquad (4.55)$$

where $\boldsymbol{P} \in \mathbb{R}^{n \times n}$ *and* $\boldsymbol{D}$ *is a diagonal matrix whose diagonal entries are the eigenvalues of* $\boldsymbol{A}$*, if and only if the eigenvectors of* $\boldsymbol{A}$ *form a basis of* $\mathbb{R}^n$*.*

4.4 - Decomposição em autovalores e autovetores e diagonalização

O Teorema 4.20 implica que apenas matrizes não defeituosas podem ser diagonalizadas e que as colunas de P são os n autovetores de A.

Uma matriz é dita "não defeituosa" se ela possui um conjunto completo de autovalores e auvetores linearmente independentes.

Em outras palavras, uma matriz não defeituosa é diagonalizável e pode ser decomposta em uma forma diagonal.

Para matrizes simétricas podemos obter resultados ainda mais fortes para a decomposição de autovalores.

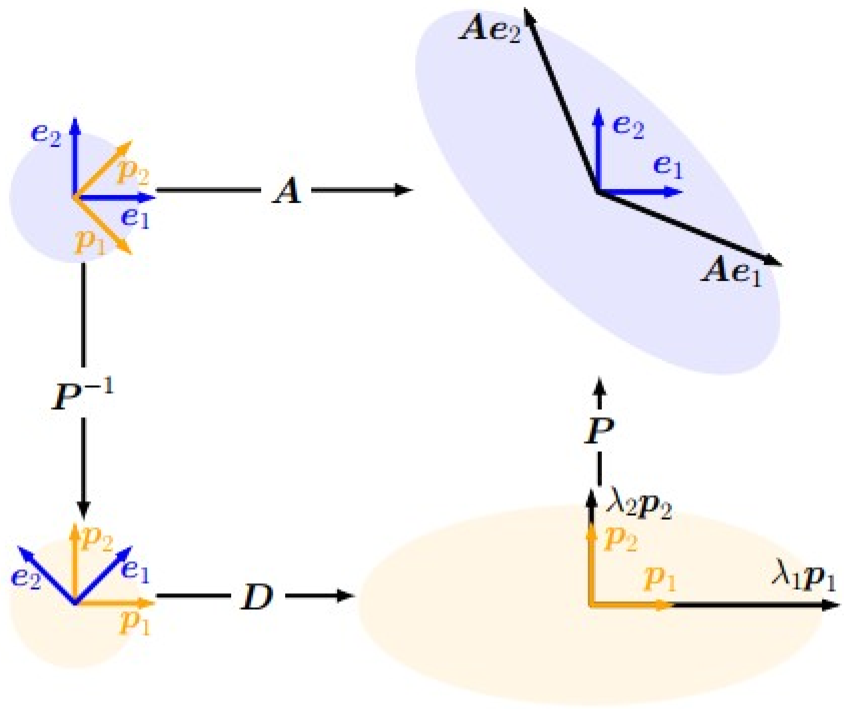
4.4 - Decomposição em autovalores e autovetores e diagonalização

Theorem 4.21. *A symmetric matrix* $\boldsymbol{S} \in \mathbb{R}^{n \times n}$ *can always be diagonalized.*

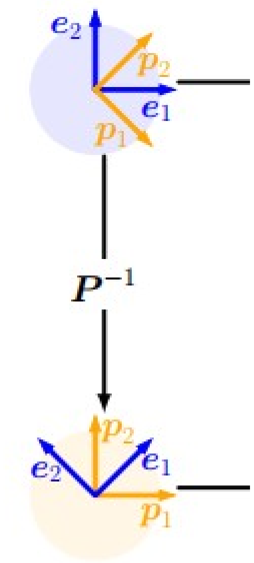
4.4 - Decomposição em autovalores e autovetores e diagonalização

Theorem 4.21. *A symmetric matrix* $\boldsymbol{S} \in \mathbb{R}^{n \times n}$ *can always be diagonalized.*

4.4 - Decomposição em autovalores e autovetores e diagonalização



4.4 - Decomposição em autovalores e autovetores e diagonalização

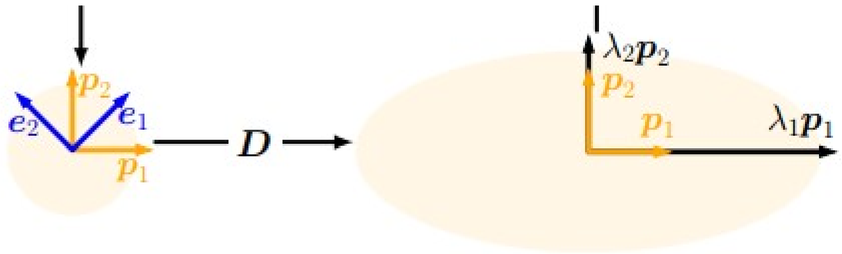


Intuição por trás da decomposição em autovalores e autovetores como uma sequência de transformações.

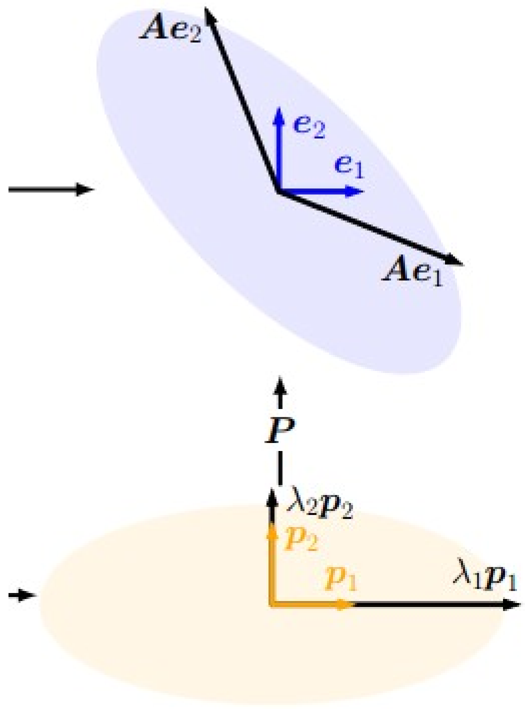
Do canto superior esquerdo para o canto inferior esquerdo: P^{-1} executa uma mudança de base (aqui desenhada em R^2 e representada como uma operação de rotação) da base padrão para a base de auvoteroes.

4.4 - Decomposição em autovalores e autovetores e diagonalização

Da parte inferior esquerda para a parte inferior direita: D realiza uma operação de escala ao longo dos autovetores ortogonais, representados aqui por um círculo sendo esticado em uma elipse.



4.4 - Decomposição em autovalores e autovetores e diagonalização



Da parte inferior direita para a parte superior direita: P desfaz a mudança de base (representada como uma rotação reversa) voltando ao sistema de coordenadas originais.

Decomposição em Valores Singulares

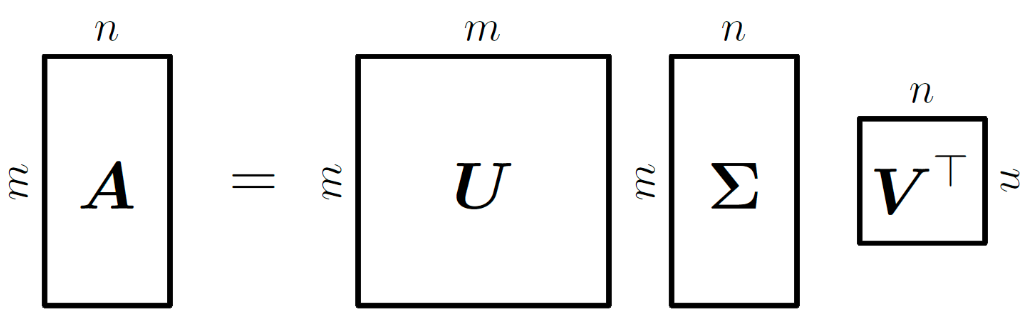
André M. Riccioppo

Decomposição em Valores Singulares

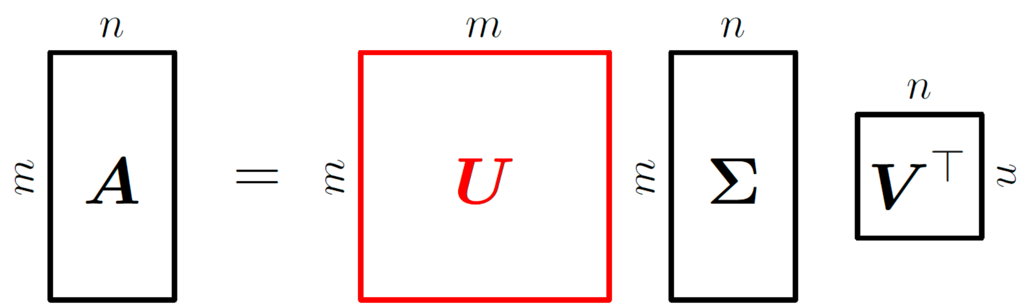
- SVD (Singular Value Decomposition)
- Permite uma análise de uma matriz usando partes mais simples e significativas.
- Como pode ser aplicada a todos tipos de matrizes, não só às matrizes quadradas, já foi chamada de teorema fundamental da álgebra linear.

Teorema

Dada uma matriz retangular $A \in \mathbb{R}^{m\,x\,n}$ de posto $r \in [0, \min(m, n)]$, a SVD de A é uma decomposição na forma

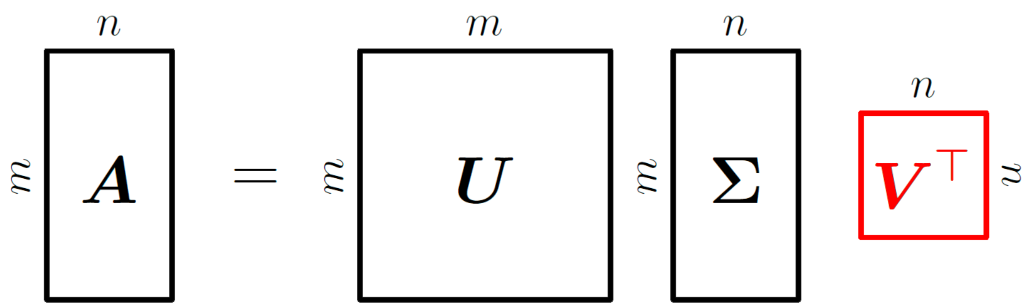


Matriz U



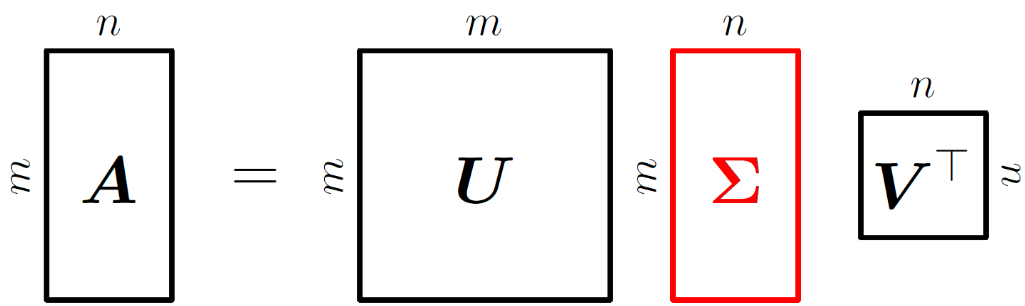
- $U \in \mathbb{R}^{m \times m}$ é uma matriz ortogonal com colunas $u_i, i = 1, \dots, m$
- As colunas u_i são chamadas <u>vetores singulares à esquerda</u>
- São os autovetores da matriz AA^T
- Esses autovetores formam uma base ortogonal para o espaço das <u>linhas</u> de A

Matriz V



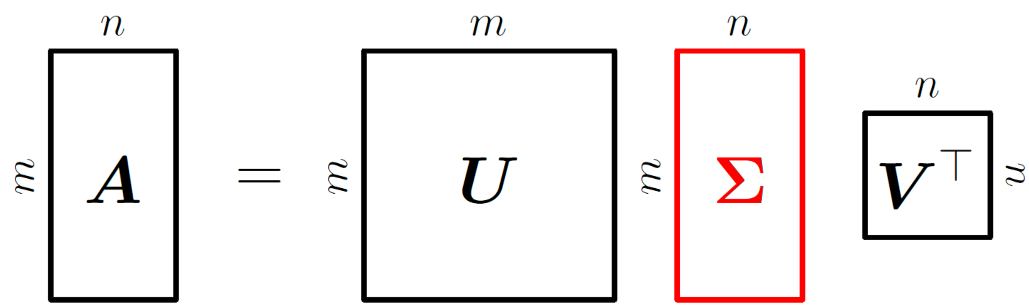
- $V \in \mathbb{R}^{n\,x\,n}$ é uma matriz ortogonal com colunas $v_j, j = 1, \ldots, n$
- As colunas v_i são chamadas <u>vetores singulares à direita</u>
- São os autovetores da matriz $A^T A$
- Esses autovetores formam uma base ortogonal para o espaço das <u>colunas</u> de A

Matriz Σ



- Σ é uma matriz diagonal m x n com $\Sigma_{ii} = \sigma_i \geq 0$ e $\Sigma_{ij} = 0, i \neq j$.
- As entradas diagonais $\sigma_i, i = 1, \ldots, r$ de Σ são chamadas de <u>valores singulares</u>
- Por convenção, os valores singulares são ordenados: $\sigma_1 \geq \sigma_2 \geq \sigma_r \geq 0$.

Matriz Σ



- $\Sigma \in \mathbb{R}^{m\,x\,n}$ é retangular e do mesmo tamanho que A.
- Isso significa que Σ possui uma submatriz diagonal que contém os valores singulares e precisa ser preenchida com 0 se m ≠ n.

se m > n

$$\Sigma = \begin{bmatrix} \sigma_1 & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & \sigma_n \\ 0 & \dots & 0 \\ \vdots & & \vdots \\ 0 & \dots & 0 \end{bmatrix}$$

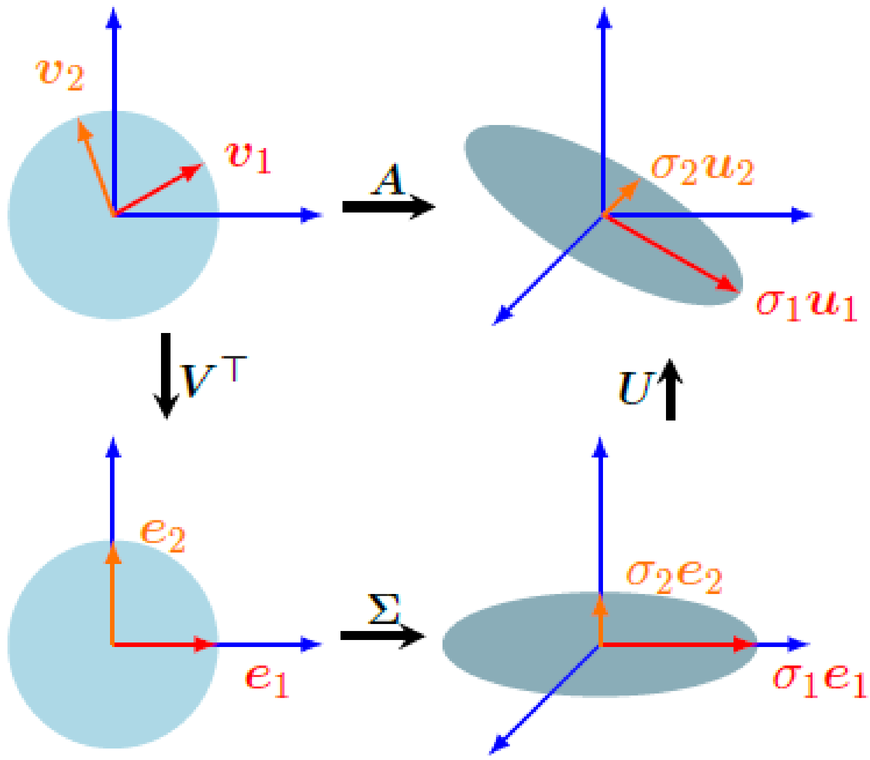
se m < n

$$\Sigma = \begin{bmatrix} \sigma_1 & 0 & 0 & 0 & \dots & 0 \\ 0 & \ddots & 0 & \vdots & & \vdots \\ 0 & 0 & \sigma_m & 0 & \dots & 0 \end{bmatrix}$$

Intepretação geométrica da SVD

- A SVD de uma matriz pode ser interpretada como a decomposição de uma transformação linear $\Phi : \mathbb{R}^n \rightarrow \mathbb{R}^m$ em três operações:
  - Uma mudança de base via V^T (rotação)
  - Uma escala e dilatação (ou contração) de sua dimensionalidade através da matriz de valor singular Σ
  - Ao final, realiza uma segunda mudança de base via U (nova rotação).

Intepretação geométrica da SVD



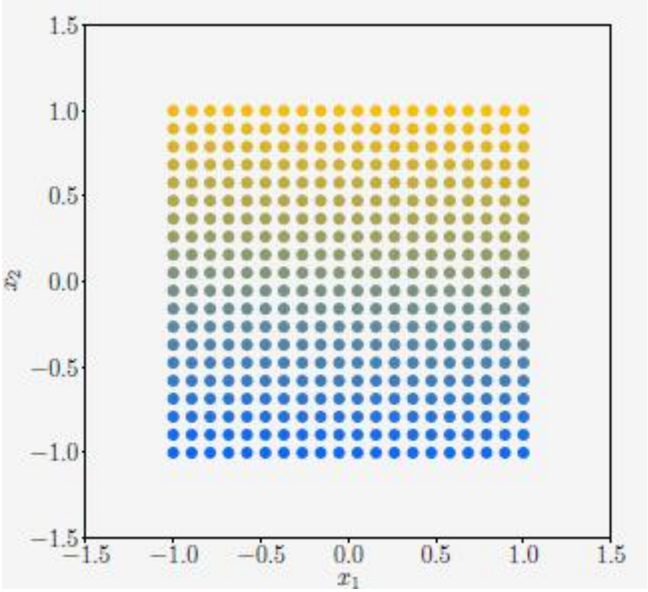
Exemplo

Considere um mapeamento de um grid quadrado de vetores $X \in \mathbb{R}^2$ que se encaixa em um quadrado tamanho 2 x 2 centralizado na origem. Usando a base padrão, os vetores são mapeados usando:

$$A = \begin{bmatrix} 1 & -0{,}8 \\ 0 & 1 \\ 1 & 0 \end{bmatrix} = U\Sigma V^T = \begin{bmatrix} -0{,}79 & 0 & -0{,}62 \\ 0{,}38 & -0{,}78 & -0{,}49 \\ -0{,}48 & -0{,}62 & 0{,}62 \end{bmatrix} \begin{bmatrix} 1{,}62 & 0 \\ 0 & 1{,}0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} -0{,}78 & 0{,}62 \\ -0{,}62 & -0{,}78 \end{bmatrix}$$

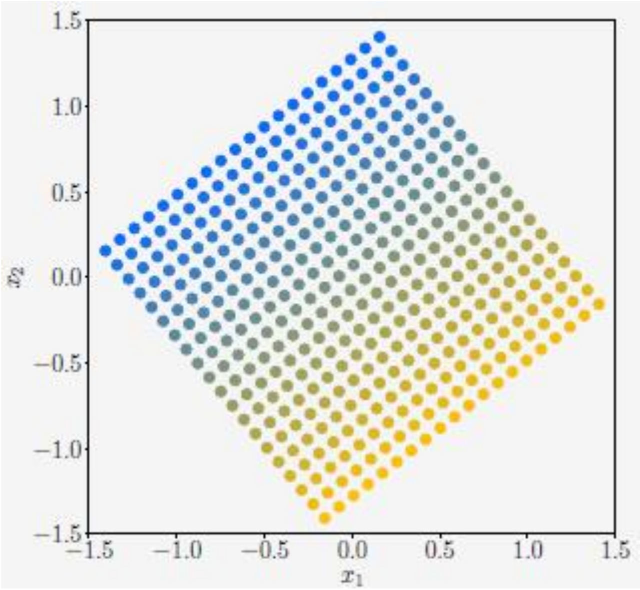
Exemplo

Início: conjunto de vetores X



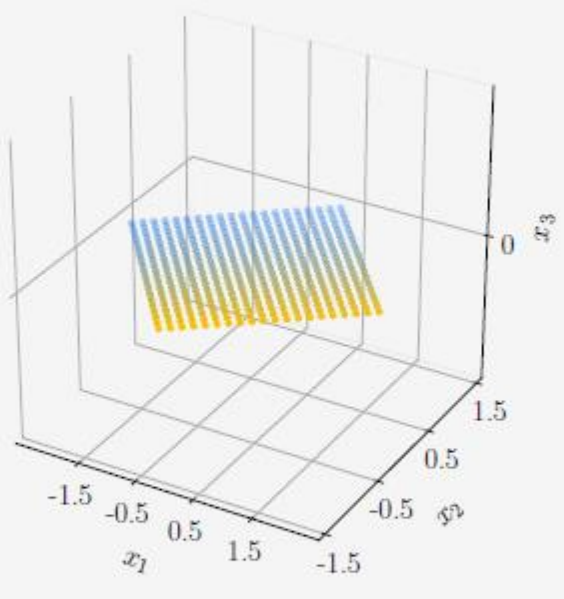
Exemplo

Então aplica $V^T \in \mathbb{R}^{2x2}$, que rotaciona X



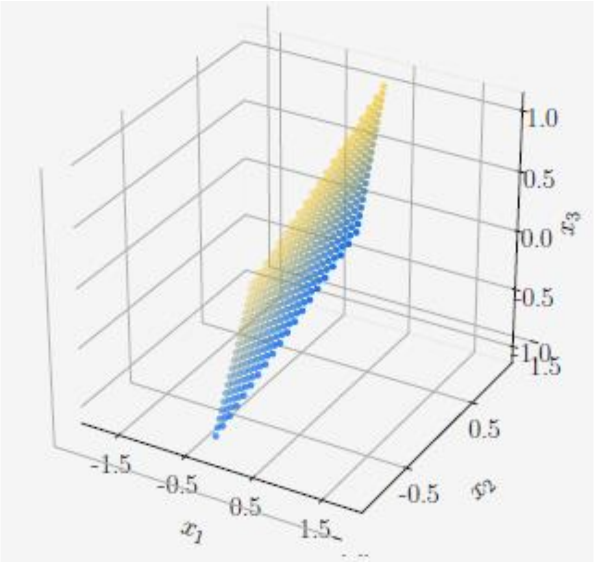
Exemplo

Então esses vetores são mapeados usando a matriz de valor singular Σ para o codomínio $\mathbb{R}^3$. Observe que todos vetores continuam no plano $x_1 - x_2$ e a terceira coordenada sempre é 0. Os vetores no plano $x_1 - x_2$ foram esticados pelos valores singulares.



Exemplo

U faz a rotação no codomínio $\mathbb{R}^3$, de forma que os vetores mapeados não fiquem mais restritos ao plano $x_1 - x_2$, mas se mantém em um plano.



Cálculo da SVD

- O cálculo da SVD de $A \in \mathbb{R}^{m \, x \, n}$ é equivalente a encontrar
  - o conjunto ortonormal de vetores singulares à direita $v_1, \ldots, v_n \in \mathbb{R}^n$, que vão gerar a matriz V
  - Depois, constrói-se o conjunto ortonormal de vetores singulares à esquerda $u_1, \ldots, u_m \in \mathbb{R}^m$, que vão gerar a matriz U
  - Então, liga-se os dois e exige-se que a ortogonalidade do v_i seja preservada sob a transformação A. Isso é importante pois se sabe que as imagens Av_i formam um conjunto de vetores ortogonais. Essas imagens são normalizadas por fatores escalares, que correspondem aos valores singulares.

Cálculo da SVD

1 – Construção dos vetores singulares à direita (V)

- O teorema espectral indica que os autovetores de uma matriz simétrica formam uma base ortonormal, o que significa que podem ser diagonalizados.
- a partir de qualquer matriz retangular $A \in \mathbb{R}^{m\,x\,n}$ podemos diagonalizar $A^T A$ e obter:

$$A^T A = PDP^T = P \begin{bmatrix} \lambda_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_n \end{bmatrix} P^T$$

onde P é uma matriz ortogonal, composta por uma autobase ortonormal. $\lambda_i \geq 0$ são os autovalores de $A^T A$.

Cálculo da SVD

1 – Construção dos vetores singulares à direita (V)

Assumindo-se que a SVD de $A = U\Sigma V^T$ existe:

$$A^T A = (U\Sigma V^T)^T (U\Sigma V^T) = V\Sigma^T U^T U\Sigma V^T$$

onde U, V são matrizes ortogonais.

Cálculo da SVD

1 – Construção dos vetores singulares à direita (V)

Assumindo-se que a SVD de $A = U\Sigma V^T$ existe:

$$A^T A = (U\Sigma V^T)^T (U\Sigma V^T) = V\Sigma^T U^T U\Sigma V^T$$

onde U, V são matrizes ortogonais. Então sendo $U^T U = I$, obtém-se:

$$A^T A = V\Sigma^T \Sigma V^T = V \begin{bmatrix} \sigma_1^2 & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & \sigma_n^2 \end{bmatrix} V^T$$

- Comparando, podemos identificar:

$$V^T = P^T$$

$$\sigma_i^2 = \lambda_i$$

- Então, os autovetores de $A^T A$ que compõe P são os vetores singulares à direita de A (V). Os autovalores de $A^T A$ são os valores singulares de Σ.

Cálculo da SVD

2 – Construção dos vetores singulares à esquerda (U)

- Para obter os vetores singulares à esquerda usa-se um procedimento similar. Começa-se calculando a SVD da matriz simétrica $AA^T \in \mathbb{R}^{m\,x\,m}$

$$AA^T = (U\Sigma V^T)(U\Sigma V^T)^T = U\Sigma V^T V \Sigma^T U^T = U\Sigma\Sigma^T U^T = U\begin{bmatrix}\sigma_1^2 & 0 & 0\\ 0 & \ddots & 0\\ 0 & 0 & \sigma_m^2\end{bmatrix}U^T$$

- O teorema espectral diz que $AA^T = SDS^T$ pode ser diagonalizado e podemos encontrar uma base ortonormal de autovetores de AA^T, que são coletados em S. Os autovetores ortonormais de AA^T são os vetores singulares à esquerda (U) e formam uma base ortonormal no codomínio da SVD.
- Então, como AA^T e A^TA possuem os mesmos autovalores diferentes de zero, as entradas não zero das matrizes Σ na SVD devem ser iguais para ambos os casos.

Cálculo da SVD

2 – Construção dos vetores singulares à esquerda (U)

- Para completar a construção da SVD, são necessários vetores singulares à esquerda que sejam ortonormais. Normaliza-se as imagens dos vetores singulares à direita Av_i, obtendo

$$u_i := \frac{Av_i}{\|Av_i\|} = \frac{1}{\sqrt{\lambda_i}} Av_i = \frac{1}{\sigma_i} Av_i$$

- Onde a última parte é obtida da equação que demonstra que os autovalores de AA^T são tais que $\sigma_i^2 = \lambda_i$.

Cálculo da SVD

3 – Obtenção dos valores singulares

- Então, os autovetores de $A^T A$, que sabemos ser os vetores singulares à direita v_i, e suas imagens normalizadas sob A, os vetores singulares à esquerda u_i formam duas bases ortonormais auto consistentes que são conectados através da matriz de valor singular Σ.

- Concatenando v_i como as colunas de V e u_i como as colunas de U, temos
$$AV = U\,\Sigma$$

- onde Σ tem as mesmas dimensões que A e uma estrutura diagonal de linhas 1, ..., r. Desta forma, temos que $A = U\Sigma V^T$, que é a SVD de A.

Exemplo

Cálculo da SVD de uma matriz

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

Passo 1 – calcular os vetores singulares à direita à partir da base de autovetores de A^TA.

- Primeiro serão calculados os vetores singulares à direita de A^TA. Inicia-se com o cálculo de A^TA.

$$A^TA = \begin{bmatrix} 1 & -2 \\ 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 & 1 \\ -2 & 1 & 0 \end{bmatrix} = \begin{bmatrix} 5 & -2 & 1 \\ -2 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix}$$

Exemplo

- Através da decomposição em autovalores de $A^T A$, obtemos a matriz P formada pelos autovetores de A e matriz diagonal D com os autovalores de A

$$A^T A = \begin{bmatrix} \frac{5}{\sqrt{30}} & 0 & \frac{-1}{\sqrt{6}} \\ \frac{-2}{\sqrt{30}} & \frac{1}{\sqrt{5}} & \frac{-2}{\sqrt{6}} \\ \frac{1}{\sqrt{30}} & \frac{2}{\sqrt{5}} & \frac{1}{\sqrt{6}} \end{bmatrix} \begin{bmatrix} 6 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} \frac{5}{\sqrt{30}} & \frac{-2}{\sqrt{30}} & \frac{1}{\sqrt{30}} \\ 0 & \frac{1}{\sqrt{5}} & \frac{2}{\sqrt{5}} \\ \frac{-1}{\sqrt{6}} & \frac{-2}{\sqrt{6}} & \frac{1}{\sqrt{6}} \end{bmatrix} = PDP^T$$

Exemplo

Os vetores singulares à direita são obtidos pelas colunas de P, portanto

$$V = \begin{bmatrix} \frac{5}{\sqrt{30}} & 0 & \frac{-1}{\sqrt{6}} \\ \frac{-2}{\sqrt{30}} & \frac{1}{\sqrt{5}} & \frac{-2}{\sqrt{6}} \\ \frac{1}{\sqrt{30}} & \frac{2}{\sqrt{5}} & \frac{1}{\sqrt{6}} \end{bmatrix}$$

Exemplo

Passo 2 – calcular a matriz de valores singulares

- Os valores singulares σ_i são as raízes quadradas dos autovalores de $A^T A$, obtidos da matriz D calculada no passo anterior.
- Como rk(A) = 2, existem apenas dois valores singulares diferentes de zero: $\sigma_1 = \sqrt{6}$ e $\sigma_2 = 1$. A matriz de valor singular deve ser do mesmo tamanho que A, então obtemos

$$\Sigma = \begin{bmatrix} \sqrt{6} & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$$

Exemplo

Passo 3 – calcular a os vetores singulares à esquerda como imagem normalizada dos vetores singulares à direita.

- Os vetores singulares à esquerda são obtidos pelo cálculo da imagem dos vetores singulares à direita em A, que são normalizados pela divisão por seu valor singular correspondente.

$$u_1 = \frac{1}{\sigma_1} A v_1 = \frac{1}{\sqrt{6}} \begin{bmatrix} 1 & 0 & 1 \\ -2 & 1 & 0 \end{bmatrix} \begin{bmatrix} \frac{5}{\sqrt{30}} \\ \frac{-2}{\sqrt{30}} \\ \frac{1}{\sqrt{30}} \end{bmatrix} = \begin{bmatrix} \frac{1}{\sqrt{5}} \\ -\frac{2}{\sqrt{5}} \end{bmatrix}$$

$$u_2 = \frac{1}{\sigma_2} A v_2 = \frac{1}{1} \begin{bmatrix} 1 & 0 & 1 \\ -2 & 1 & 0 \end{bmatrix} \begin{bmatrix} 0 \\ \frac{1}{\sqrt{5}} \\ \frac{2}{\sqrt{5}} \end{bmatrix} = \begin{bmatrix} \frac{2}{\sqrt{5}} \\ \frac{1}{\sqrt{5}} \end{bmatrix}$$

$$U = [u_1, u_2] = \frac{1}{\sqrt{5}} \begin{bmatrix} 1 & 2 \\ -2 & 1 \end{bmatrix}$$

Conclusões

- A SVD é comparável à autodecomposição
  - Autodecomposição: só matrizes quadradas com base de autovetores
  - SVD: qualquer matriz retangular
- Ambas são composições de três transformações lineares:
  - uma mudança de base,
  - uma escala e transformação de domínio e
  - uma nova mudança de base.
- A SVD é usada em vários tipos de aplicações:
  - problemas de mínimos quadrados em ajuste de curvas,
  - resolução de sistemas de equações lineares,
  - redução de dimensionalidade,
  - modelagem de tópicos e
  - compressão e clusterização de dados.

Matrizes de aproximação

Gabriel M. Almeida

Decomposição de Valores Singulares (SVD)

- Decomposição espectral ($Ax = \lambda x$)
  - Aplica-se apenas a matrizes quadradas
- Na prática matrizes quadradas não são comuns
  - Matrizes esparsas
  - Estrutura em “blocos”

$$\begin{array}{ccc}
 & \begin{array}{ccccc}\text{Matrix} & \text{Alien} & \text{Serenity} & \text{Casablanca} & \text{Amelie}\end{array} & \\
\begin{array}{c} \\ \\ \text{SciFi} \\ \\ \\ \text{Romance} \\ \\ \end{array} & \begin{bmatrix}1&1&1&0&0\\3&3&3&0&0\\4&4&4&0&0\\5&5&5&0&0\\0&0&0&4&4\\0&0&0&5&5\\0&0&0&2&2\end{bmatrix} & \begin{array}{l}\text{Ana}\\ \text{Bruno}\\ \dots\\ \\ \\ \\ \\ \end{array}
\end{array}$$

Decomposição de Valores Singulares (SVD)

- SVD completo
  - $AV = U\Sigma \Leftrightarrow AVV^{-1} = U\Sigma V^{-1} \Leftrightarrow A = U\Sigma V^T$

$$A \underbrace{\begin{bmatrix} | & \cdots & | & | & \cdots & | \\ v_1 & \cdots & v_r & v_{r+1} & \cdots & v_n \\ | & \cdots & | & | & \cdots & | \end{bmatrix}}_{n \times n} = \underbrace{\begin{bmatrix} | & \cdots & | & | & \cdots & | \\ u_1 & \cdots & u_r & u_{r+1} & \cdots & u_m \\ | & \cdots & | & | & \cdots & | \end{bmatrix}}_{m \times m} \underbrace{\begin{bmatrix} \sigma_1 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & \sigma_r & 0 & \cdots & 0 \\ 0 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & 0 & \cdots & 0 \end{bmatrix}}_{m \times n}$$

Decomposição de Valores Singulares (SVD)

- O SVD de A (mxn) e posto r origirina r valores singulares (σ)
  - As matrizes V, U e ∑, não são matrizes quadradas
  - SVD reduzido

$$A \begin{bmatrix} v_1 & v_2 & \ldots & v_r \end{bmatrix} = \begin{bmatrix} u_1 & u_2 & \ldots & u_r \end{bmatrix} \begin{bmatrix} \sigma_1 & & \\ & \ddots & \\ & & \sigma_r \end{bmatrix}$$

Decomposição de Valores Singulares (SVD)

- Exemplo de aplicação de SVD
  - Matriz A tem posto dois, resultando em 2 valores singulares e matrizes U e V com 2 vetores colunas
- Número de conceitos é relacionado ao posto da matriz

$$
\begin{array}{cc}
 & \begin{matrix} \text{Matrix} & \text{Alien} & \text{Serenity} & \text{Casablanca} & \text{Amelie} \end{matrix} \\
\begin{matrix} \uparrow \\ \text{SciFi} \\ \downarrow \\ \uparrow \\ \text{Romance} \\ \downarrow \end{matrix} &
\begin{bmatrix} 1 & 1 & 1 & 0 & 0 \\ 3 & 3 & 3 & 0 & 0 \\ 4 & 4 & 4 & 0 & 0 \\ 5 & 5 & 5 & 0 & 0 \\ 0 & 0 & 0 & 4 & 4 \\ 0 & 0 & 0 & 5 & 5 \\ 0 & 0 & 0 & 2 & 2 \end{bmatrix}
\end{array}
=
\begin{bmatrix} \mathbf{0.14} & 0.00 \\ \mathbf{0.42} & 0.00 \\ \mathbf{0.56} & 0.00 \\ \mathbf{0.70} & 0.00 \\ 0.00 & \mathbf{0.60} \\ 0.00 & \mathbf{0.75} \\ 0.00 & \mathbf{0.30} \end{bmatrix}
\times
\begin{bmatrix} 12.4 & 0 \\ 0 & 9.5 \end{bmatrix}
\times
\begin{bmatrix} \mathbf{0.58} & \mathbf{0.58} & \mathbf{0.58} & 0.00 & 0.00 \\ 0.00 & 0.00 & 0.00 & \mathbf{0.71} & \mathbf{0.71} \end{bmatrix}
$$

Decomposição de Valores Singulares (SVD)

- Exemplo de aplicação de SVD
  - Matriz A de posto 3 e valor singular σ1, σ2 e σ3

$$
\begin{array}{cc}
 & \begin{matrix}\text{Matrix} & \text{Alien} & \text{Serenity} & \text{Casablanca} & \text{Amelie}\end{matrix} \\
\begin{matrix} \\ \text{SciFi} \\ \\ \\ \\ \text{Romance} \\ \\ \end{matrix} &
\begin{bmatrix}
1 & 1 & 1 & 0 & 0 \\
3 & 3 & 3 & 0 & 0 \\
4 & 4 & 4 & 0 & 0 \\
5 & 5 & 5 & 0 & 0 \\
0 & \boxed{2} & 0 & 4 & 4 \\
0 & 0 & 0 & 5 & 5 \\
0 & \boxed{1} & 0 & 2 & 2
\end{bmatrix}
\end{array}
=
\begin{bmatrix}
\mathbf{0.13} & -0.02 & -0.01 \\
\mathbf{0.41} & -0.07 & -0.03 \\
\mathbf{0.55} & -0.09 & -0.04 \\
\mathbf{0.68} & -0.11 & -0.05 \\
0.15 & \mathbf{0.59} & \mathbf{0.65} \\
0.07 & \mathbf{0.73} & \mathbf{-0.67} \\
0.07 & \mathbf{0.29} & \mathbf{0.32}
\end{bmatrix}
\times
\begin{bmatrix}
\mathbf{12.4} & 0 & 0 \\
0 & \mathbf{9.5} & 0 \\
0 & 0 & \mathbf{1.3}
\end{bmatrix}
\times
\begin{bmatrix}
\mathbf{0.56} & \mathbf{0.59} & \mathbf{0.56} & 0.09 & 0.09 \\
-0.12 & 0.02 & -0.12 & \mathbf{0.69} & \mathbf{0.69} \\
0.40 & \mathbf{-0.80} & 0.40 & 0.09 & 0.09
\end{bmatrix}
$$

Matriz de aproximação

- Removendo um valor singular (σ3)
    - Matriz de aproximação calculada com σ1 e σ2

$$
\begin{array}{c c}
 & \begin{array}{ccccc} \text{Matrix} & \text{Alien} & \text{Serenity} & \text{Casablanca} & \text{Amelie} \end{array} \\
\begin{array}{c} \text{SciFi} \\ \\ \\ \\ \text{Romance} \\ \\ \\ \end{array} &
\begin{bmatrix}
1 & 1 & 1 & 0 & 0 \\
3 & 3 & 3 & 0 & 0 \\
4 & 4 & 4 & 0 & 0 \\
5 & 5 & 5 & 0 & 0 \\
0 & \boxed{2} & 0 & 4 & 4 \\
0 & 0 & 0 & 5 & 5 \\
0 & \boxed{1} & 0 & 2 & 2
\end{bmatrix}
\end{array}
=
\begin{bmatrix}
0.13 & -0.02 & \cancel{-0.01} \\
0.41 & -0.07 & \cancel{-0.03} \\
0.55 & -0.09 & \cancel{-0.04} \\
0.68 & -0.11 & \cancel{-0.05} \\
0.15 & 0.59 & \cancel{0.65} \\
0.07 & 0.73 & \cancel{-0.67} \\
0.07 & 0.29 & \cancel{0.32}
\end{bmatrix}
\times
\begin{bmatrix}
12.4 & 0 & 0 \\
0 & 9.5 & 0 \\
0 & 0 & \cancel{1.3}
\end{bmatrix}
\times
\begin{bmatrix}
0.56 & 0.59 & 0.56 & 0.09 & 0.09 \\
-0.12 & 0.02 & -0.12 & 0.69 & 0.69 \\
\cancel{0.40} & \cancel{-0.80} & \cancel{0.40} & \cancel{0.09} & \cancel{0.09}
\end{bmatrix}
$$

Matriz de aproximação

- Aproximação de matrizes - **Exemplo de aplicação**
  - Algoritmos de recomendação
    - Lojas virtuais, streaming, propagandas
- Considerando k usuários
  - Cada usuário é o resultado da combinação linear dos conceitos e a matriz original A terá posto k
  - Quais filmes recomendar para um usuário baseado em seu perfil?

Matriz de aproximação

- Exemplo de recomendação
  - Excluindo linhas e colunas de U, V e ∑ obtemos uma recomendação dos filmes para os usuários destacados

$$\begin{bmatrix} 1 & 1 & 1 & 0 & 0 \\ 0 & 3 & 3 & 0 & 0 \\ 4 & 4 & 0 & 0 & 0 \\ 5 & 5 & 5 & 0 & 0 \\ 0 & 2 & 0 & 4 & 4 \\ 0 & 0 & 0 & 5 & 5 \\ 0 & 1 & 0 & 2 & 2 \end{bmatrix} = \begin{bmatrix} \mathbf{0.14} & -0.06 & -0.04 \\ \mathbf{0.30} & -0.11 & \mathbf{-0.61} \\ \mathbf{0.43} & -0.16 & \mathbf{0.76} \\ \mathbf{0.74} & -0.31 & -0.18 \\ 0.15 & \mathbf{0.53} & 0.02 \\ 0.07 & \mathbf{0.70} & -0.03 \\ 0.07 & \mathbf{0.27} & 0.01 \end{bmatrix} \times \begin{bmatrix} 12.4 & 0 & 0 \\ 0 & 9.5 & 0 \\ 0 & 0 & 1.3 \end{bmatrix} \times \begin{bmatrix} \mathbf{0.51} & \mathbf{0.66} & \mathbf{0.44} & 0.23 & 0.23 \\ \mathbf{-0.24} & -0.13 & -0.21 & \mathbf{0.66} & \mathbf{0.66} \\ 0.59 & 0.08 & \mathbf{-0.80} & 0.01 & 0.01 \end{bmatrix}$$

Matriz de aproximação

- Exemplo de recomendação
  - Excluindo linhas e colunas de U, V e ∑ obtemos uma recomendação dos filmes para os usuários destacados

$$
\begin{bmatrix}
1 & 1 & 1 & 0 & 0 \\
0 & 3 & 3 & 0 & 0 \\
4 & 4 & 0 & 0 & 0 \\
5 & 5 & 5 & 0 & 0 \\
0 & 2 & 0 & 4 & 4 \\
0 & 0 & 0 & 5 & 5 \\
0 & 1 & 0 & 2 & 2
\end{bmatrix}
=
\begin{bmatrix}
\mathbf{0.14} & -0.06 & \cancel{-0.04} \\
\mathbf{0.30} & -0.11 & \cancel{-0.61} \\
\mathbf{0.43} & -0.16 & \cancel{0.76} \\
\mathbf{0.74} & -0.31 & \cancel{-0.18} \\
0.15 & \mathbf{0.53} & \cancel{0.02} \\
0.07 & \mathbf{0.70} & \cancel{-0.03} \\
0.07 & \mathbf{0.27} & \cancel{0.01}
\end{bmatrix}
\times
\begin{bmatrix}
12.4 & 0 & 0 \\
0 & 9.5 & 0 \\
0 & 0 & \cancel{1.3}
\end{bmatrix}
\times
\begin{bmatrix}
\mathbf{0.51} & \mathbf{0.66} & \mathbf{0.44} & 0.23 & 0.23 \\
-0.24 & -0.13 & -0.21 & \mathbf{0.66} & \mathbf{0.66} \\
\cancel{0.59} & \cancel{0.08} & \cancel{-0.80} & \cancel{0.01} & \cancel{0.01}
\end{bmatrix}
$$

Matriz de aproximação

- Exemplo de recomendação
  - Excluindo linhas e colunas de U, V e ∑ obtemos uma recomendação dos filmes para os usuários destacados

Valores originais

$$\begin{bmatrix} 1 & 1 & 1 & 0 & 0 \\ 0 & 3 & 3 & 0 & 0 \\ 4 & 4 & 0 & 0 & 0 \\ 5 & 5 & 5 & 0 & 0 \\ 0 & 2 & 0 & 4 & 4 \\ 0 & 0 & 0 & 5 & 5 \\ 0 & 1 & 0 & 2 & 2 \end{bmatrix}$$

Valores recomendação

$$\begin{bmatrix} 0.96 & 1.14 & 0.82 & -0.01 & -0.01 \\ 1.94 & 2.32 & 1.66 & 0.07 & 0.07 \\ 2.77 & 3.32 & 2.37 & 0.08 & 0.08 \\ 4.84 & 5.74 & 4.14 & -0.08 & 0.08 \\ 0.40 & 1.42 & 0.33 & 4.06 & 4.06 \\ -0.42 & 0.63 & -0.38 & 4.92 & 4.92 \\ 0.20 & 0.71 & 0.16 & 2.03 & 2.03 \end{bmatrix}$$

Matriz de aproximação

- Aproximação de matrizes - **Exemplo de aplicação**
  - Algoritmos de compressão de imagens
    - Reduzir custo de transmissão de imagens em redes
    - Reduzir tempo de transmissão de imagens
    - Reduzir custo de processamento de imagens por modelos de segmentação
    - Transmissão 4K (30 FPS) sem compressão demandaria uma transmissão de ~ 7 Gbps

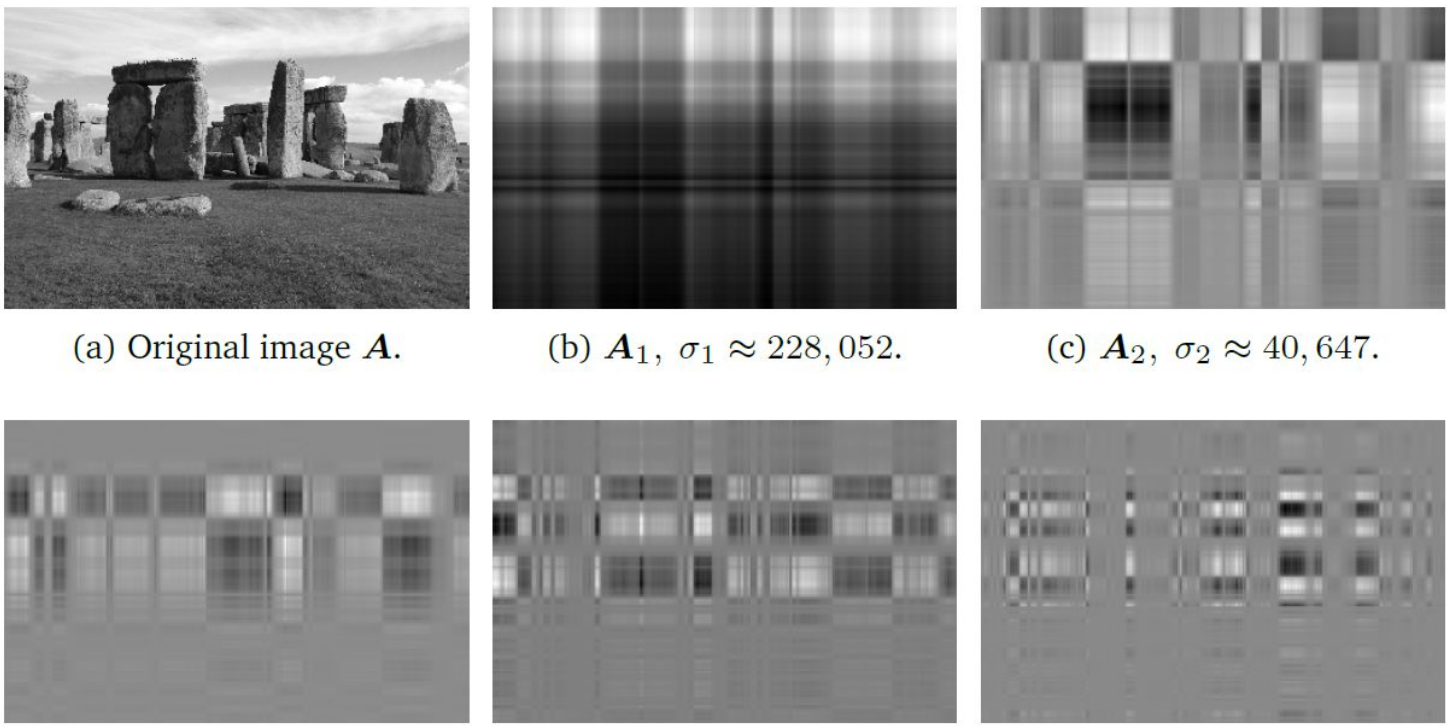
Matriz de aproximação

- Seja a imagem representada por uma matriz A e a matriz de aproximação A_k (de posto k)
  - Podemos usar os k maiores valores singulares

$$A_k = \sigma_1 u_1 v_1^T + \cdots + \sigma_k u_k v_k^T$$

Matriz de aproximação

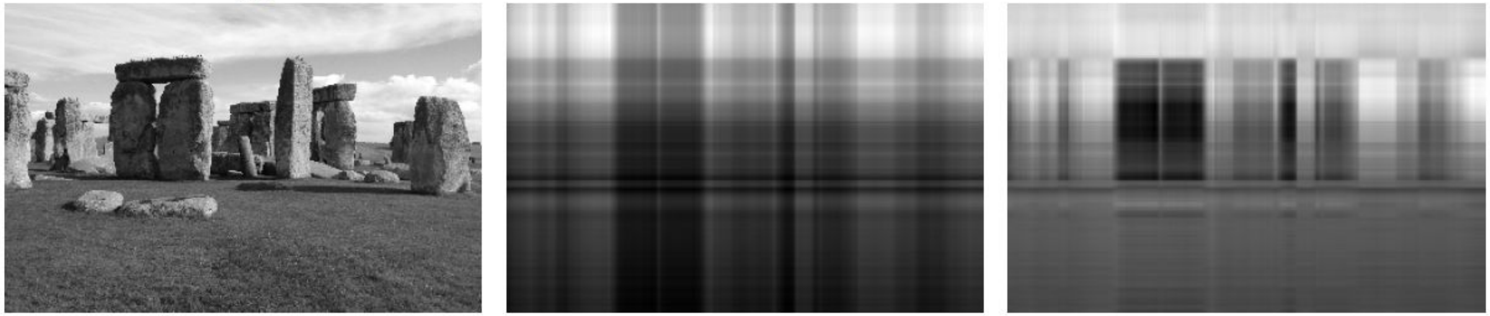
- Compressão de imagens
  - Imagem resultante para cada valor singular calculado



(a) Original image $\boldsymbol{A}$.
(b) $\boldsymbol{A}_1,\ \sigma_1 \approx 228,052.$
(c) $\boldsymbol{A}_2,\ \sigma_2 \approx 40,647.$
(d) $\boldsymbol{A}_3,\ \sigma_3 \approx 26,125.$
(e) $\boldsymbol{A}_4,\ \sigma_4 \approx 20,232.$
(f) $\boldsymbol{A}_5,\ \sigma_5 \approx 15,436.$

Matriz de aproximação

- Compressão de imagens
    - Imagem resultante para cada valor singular calculado



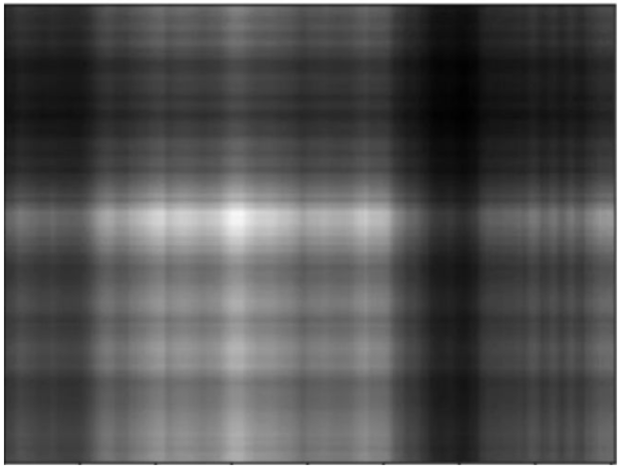
(a) Original image $\boldsymbol{A}$. (b) Rank-1 approximation $\widehat{\boldsymbol{A}}(1)$. (c) Rank-2 approximation $\widehat{\boldsymbol{A}}(2)$.

(d) Rank-3 approximation $\widehat{\boldsymbol{A}}(3)$. (e) Rank-4 approximation $\widehat{\boldsymbol{A}}(4)$. (f) Rank-5 approximation $\widehat{\boldsymbol{A}}(5)$.

Matriz de aproximação

- Compressão de imagens - exemplo cachorro (posto 3024)
    - Imagem resultante para cada valor singular calculado

Posto 1



Posto 10

Posto 100

Matriz de aproximação

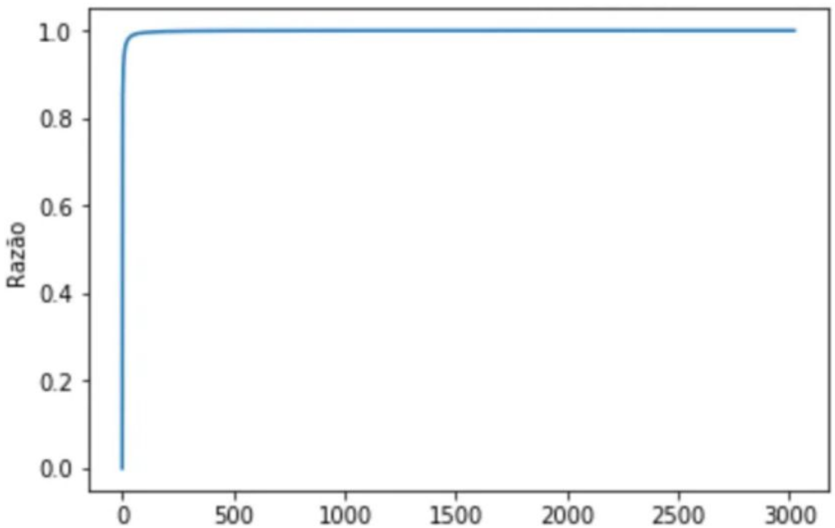
- Compressão de imagens
  - Como verificar se a compressão é suficiente?
- Root Mean Square Error (RMSE)
  - Diferença média entre o pixel da matriz original e o pixel da matriz aproximada

$$RMSE = \sqrt{\frac{1}{MN}\sum_{m=0}^{M-1}\sum_{n=0}^{N-1}\left(I1(m,n) - I2(m,n)\right)^2}$$

Matriz de aproximação

- Compressão de imagens
  - Quando parar de incluir valores singulares na compressão?
  - Regra do polegar: cortar de maneira a manter entre 80-90% da energia total

$$t = \frac{\sum_{i=1}^{k} \sigma^2}{\sum_{i=1}^{n} \sigma^2}$$



Matriz de aproximação

- Compressão de imagens
  - É possível encontrar uma matriz B de posto k que seja mais próxima de A e que não seja A_k?
  - **Teorema de Eckart-Young:** Suponha uma matriz B de posto k. Então B será uma aproximação no máximo tão boa quanto Ak.

$$A_k = \sigma_1 u_1 v_1^T + \cdots + \sigma_k u_k v_k^T$$

Conclusão

- Parte 1 - Hudson Romualdo (hudson_romualdo@discente.ufg.br)
  - Determinante e Traço de matrizes
  - Autovalores e autovetores
- Parte 2 - Fábio Pereira (fabio.pereira@discente.ufg.br)
  - Decomposição de Cholesky
  - Decomposição de autovalores, autovetores e diagonalização
- Parte 3 - André Riccioppo (andre.riccioppo@discente.ufg.br)
  - Decomposição em valores singulares
- Parte 4 - Gabriel Almeida (gabrielmatheus05@discente.ufg.br)
  - Aproximação de matrizes