“Machine Learning estadístico para interfaces Cerebro-Computadora”

Repaso álgebra lineal

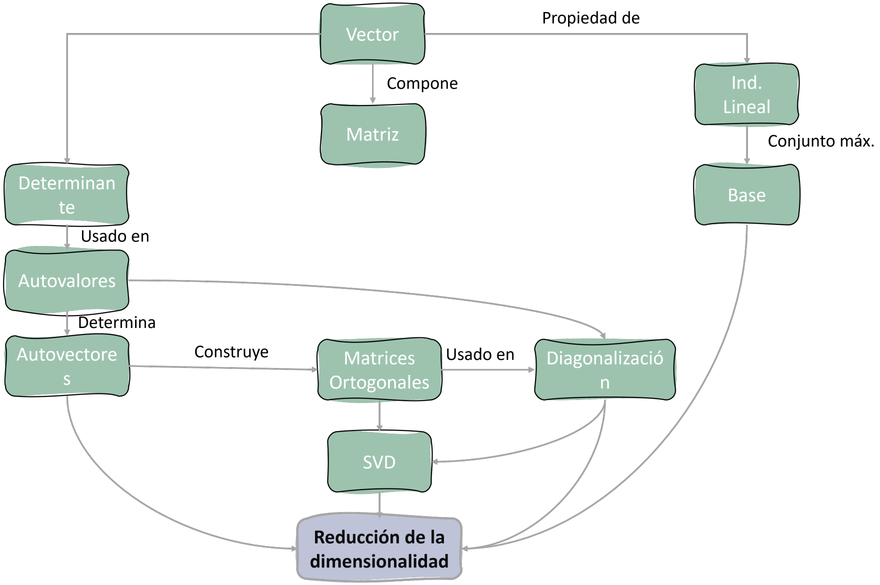
Dra. Victoria Peterson
vpeterson@santafe-conicet.gov.ar

Nov 2023

NiCALab

CONICET
UNL
I M A L

Álgebra lineal all around



Autovalores y autovectores

Valores *propios*

"Caracterizadores de una matriz"

Definición

Def. 4.6. MML Book

Sea $\mathbf{A} \in \mathbb{R}^{n\times n}$ una matriz cuadrada. Luego, $\lambda \in \mathbb{R}$ es un autovalor de $\mathbf{A}$ y $\mathbf{x} \in \mathbb{R}^n \setminus \{\mathbf{0}\}$ es el correspondiente autovector de $\mathbf{A}$ si

$$\mathbf{A}\mathbf{x} = \lambda\mathbf{x}$$

o que es lo mismo

$$(\mathbf{A} - \lambda\mathbf{I}_n)\,\mathbf{x} = \mathbf{0}$$

'

Autovalores y autovectores

Valores *propios*

"Caracterizadores de una matriz"

Definición

Def. 4.6. MML Book

Sea $\mathbf{A} \in \mathbb{R}^{n \times n}$ una matriz cuadrada. Luego, $\lambda \in \mathbb{R}$ es un autovalor de $\mathbf{A}$ y $\mathbf{x} \in \mathbb{R}^n \setminus \{\mathbf{0}\}$ es el correspondiente autovector de $\mathbf{A}$ si

$$\mathbf{A}\mathbf{x} = \lambda \mathbf{x}$$

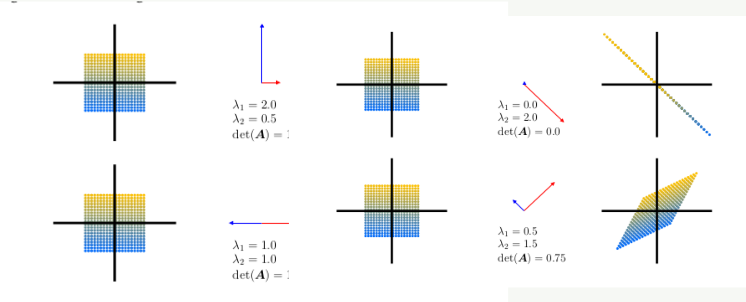
o que es lo mismo

$$(\mathbf{A} - \lambda \mathbf{I}_n)\,\mathbf{x} = \mathbf{0}$$

"Multiplicación matriz-vector igual a una multiplicación escalar"

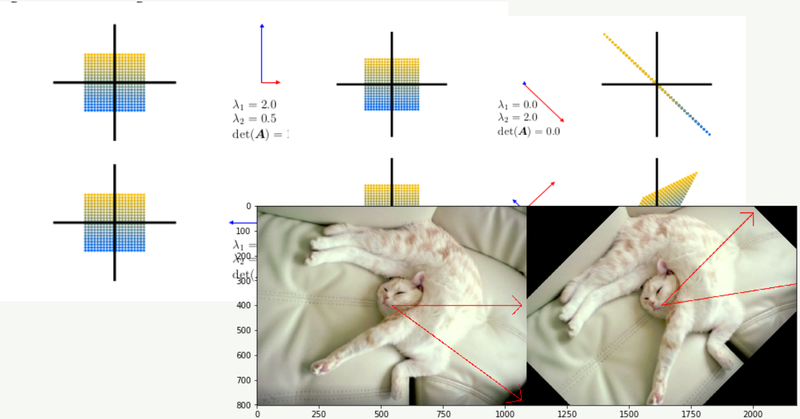
Autovalores y autovectores

Matrices como transformaciones lineales



Autovalores y autovectores

Matrices como transformaciones lineales



Autovalores y autovectores

Computo

- Resolver $\det\left(\mathbf{A} - \lambda\mathbf{I}\right) = 0$
- Hallar los autovectores correspondientes a cada autovalor hallado mediante:

$$\begin{bmatrix} a_{11} - \lambda & a_{12} \\ a_{21} & a_{22} - \lambda \end{bmatrix} \mathbf{x} = \mathbf{0}$$

Descomposición en valores propios (EVD)

Teorema

Def. 4.20. MML Book

Una matriz cuadrada $\mathbf{A} \in \mathbb{R}^{n \times n}$ puede ser factorizada en

$$\mathbf{A} = \mathbf{P}\mathbf{D}\mathbf{P}^{-1},$$

donde $\mathbf{P} \in \mathbb{R}^{n \times n}$ y $\mathbf{D}$ es una matriz diagonal cuyos valores en la diagonal son los autovalores de $\mathbf{A}$, si y sólo si los autovectores de $\mathbf{A}$ forman una base en $\mathbb{R}$.

Descomposición en valores propios (EVD)

Teorema

Def. 4.20. MML Book

Una matriz cuadrada $\mathbf{A} \in \mathbb{R}^{n \times n}$ puede ser factorizada en

$$\mathbf{A} = \mathbf{P}\mathbf{D}\mathbf{P}^{-1},$$

donde $\mathbf{P} \in \mathbb{R}^{n \times n}$ y $\mathbf{D}$ es una matriz diagonal cuyos valores en la diagonal son los autovalores de $\mathbf{A}$, si y sólo si los autovectores de $\mathbf{A}$ forman una base en $\mathbb{R}$.

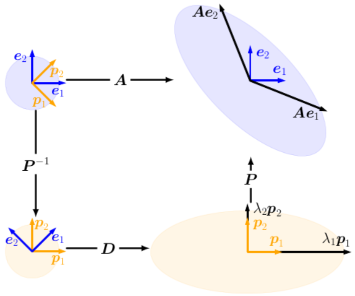


Fig. 4.7 - MML Book

Descomposición en valores propios (EVD)

Teorema

Def. 4.20. MML Book

Una matriz cuadrada $\mathbf{A} \in \mathbb{R}^{n \times n}$ puede ser factorizada en

$$\mathbf{A} = \mathbf{P}\mathbf{D}\mathbf{P}^{-1},$$

donde $\mathbf{P} \in \mathbb{R}^{n \times n}$ y $\mathbf{D}$ es una matriz diagonal cuyos valores en la diagonal son los autovalores de $\mathbf{A}$, si y sólo si los autovectores de $\mathbf{A}$ forman una base en $\mathbb{R}$.

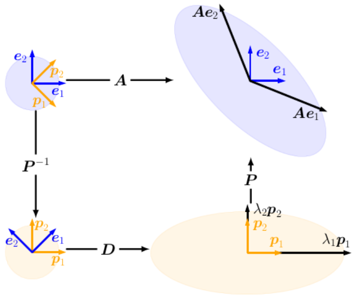


Fig. 4.7 - MML Book

Transformaciones Lineales

Descomposición en valores propios (EVD)

Teorema

Def. 4.20. MML Book

Una matriz cuadrada $\mathbf{A} \in \mathbb{R}^{n \times n}$ puede ser factorizada en

$$\mathbf{A} = \mathbf{P}\mathbf{D}\mathbf{P}^{-1},$$

donde $\mathbf{P} \in \mathbb{R}^{n \times n}$ y $\mathbf{D}$ es una matriz diagonal cuyos valores en la diagonal son los autovalores de $\mathbf{A}$, si y sólo si los autovectores de $\mathbf{A}$ forman una base en $\mathbb{R}$.

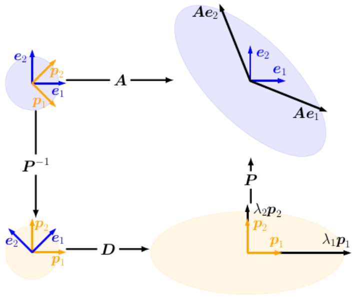


Fig. 4.7 - MML Book

Transformaciones Lineales

SVD

Formulación

Sea $\mathbf{A} \in \mathbb{R}^{m \times n}$, una matriz rectangular de rango $r \in [0, \min(m, n)]$ la descomposición en SVD de $\mathbf{A}$ es una descomposición de la forma:

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

donde $\mathbf{U} \in \mathbb{R}^{m \times m}$ matriz ortogonal con vectores columna $\mathbf{u}_i$, $i = 1, \ldots, m$, $\mathbf{V} \in \mathbb{R}^{n \times n}$ matriz ortogonal con vectores columna $\mathbf{v}_j$, $j = 1, \ldots, n$ y $\Sigma \in \mathbb{R}^{m \times n}$ matriz diagonal tal que $\mathbf{\Sigma}_i = \sigma_i \geq 0$.

SVD

Formulación

Sea $\mathbf{A} \in \mathbb{R}^{m \times n}$, una matriz rectangular de rango $r \in [0, \min(m,n)]$ la descomposición en SVD de $\mathbf{A}$ es una decomposición de la forma:

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

donde $\mathbf{U} \in \mathbb{R}^{m \times m}$ matriz ortogonal con vectores columna $\mathbf{u}_i$, $i = 1, \ldots, m$, $\mathbf{V} \in \mathbb{R}^{n \times n}$ matriz ortogonal con vectores columna $\mathbf{v}_j$, $j = 1, \ldots, n$ y $\Sigma \in \mathbb{R}^{m \times n}$ matriz diagonal tal que $\mathbf{\Sigma}_i = \sigma_i \geq 0$.

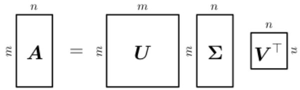


Fig. 10.2 - MML Book

SVD

Formulación

Sea $\mathbf{A} \in \mathbb{R}^{m \times n}$, una matriz rectangular de rango $r \in [0, \min(m, n)]$ la descomposición en SVD de $\mathbf{A}$ es una descomposición de la forma:

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

donde $\mathbf{U} \in \mathbb{R}^{m \times m}$ matriz ortogonal con vectores columna $\mathbf{u}_i$, $i = 1, \ldots, m$, $\mathbf{V} \in \mathbb{R}^{n \times n}$ matriz ortogonal con vectores columna $\mathbf{v}_j$, $j = 1, \ldots, n$ y $\Sigma \in \mathbb{R}^{m \times n}$ matriz diagonal tal que $\mathbf{\Sigma}_i = \sigma_i \geq 0$.

$\mathbf{u}_i$, $i = 1, \ldots, m$: vectores singulares izquierdos
$\mathbf{v}_j$, $j = 1, \ldots, n$: vectores singulares derechos
σ_i, $i = 1, \ldots, r$: valores singulares.

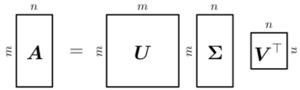


Fig. 10.2 - MML Book

SVD

Computo

1. Calculo de los vectores singulares derechos mediante descomposición en valores propios.
2. Cálculo de la matriz de valores singulares.
3. Cálculo de los vectores singulares izquierdos como la imagen normalizada de los vectores singulares derechos.

SVD

Aproximación de matrices

Una matriz $\mathbf{A} \in \mathbb{R}^{m \times n}$ de rango r puede escribirse como la suma de matrices $\mathbf{A}_i$ de rango-1 tal que:

$$\mathbf{A} = \sum_{i=1}^{r} \sigma_i \mathbf{u}_i \mathbf{v}_i^T = \sum_{i=1}^{r} \sigma_i \mathbf{A}_i.$$

Por lo tanto es fácil de ver que podemos obtener una aproximación de $\mathbf{A}$ mediante la suma de las primeras k matrices de rango-1:

$$\hat{\mathbf{A}}(k) \triangleq \sum_{i=1}^{k} \sigma_i \mathbf{u}_i \mathbf{v}_i^T = \sum_{i=1}^{k} \sigma_i \mathbf{A}_i.$$

SVD

Sección 4.6 MML Book

Aproximación de matrices

Una matriz $\mathbf{A} \in \mathbb{R}^{m \times n}$ de rango r puede escribirse como la suma de matrices $\mathbf{A}_i$ de rango-1 tal que:

$$\mathbf{A} = \sum_{i=1}^{r} \sigma_i \mathbf{u}_i \mathbf{v}_i^T = \sum_{i=1}^{r} \sigma_i \mathbf{A}_i.$$

Por lo tanto es fácil de ver que podemos obtener una aproximación de $\mathbf{A}$ mediante la suma de las primeras k matrices de rango-1:

$$\hat{\mathbf{A}}(k) \triangleq \sum_{i=1}^{k} \sigma_i \mathbf{u}_i \mathbf{v}_i^T = \sum_{i=1}^{k} \sigma_i \mathbf{A}_i.$$

La aproximación de matrices mediante SVD es una aproximación de mínima norma (ver Teo 4.25)

SVD

Aproximación de matrices: error de aproximación

Sea $\mathbf{A} \in \mathbb{R}^{m \times n}$ una matriz de rango r y sea $\mathbf{B} \in \mathbb{R}^{m \times n}$ una matriz de rango k, para cualquier $k \leq r$ con $\hat{\mathbf{A}}(k) = \sum_{i=1}^{k} \sigma_i \mathbf{u}_o \mathbf{v}_i^T$ se muestra que:

$$\hat{\mathbf{A}}(k) = argmin_{rk(\mathbf{B})=k} \|\mathbf{A} - \mathbf{B}\|$$

$$\|\mathbf{A} - \hat{\mathbf{A}}\|_2 = \sigma_{k+1}^*$$

SVD

Aproximación de matrices: error de aproximación

Sea $\mathbf{A} \in \mathbb{R}^{m \times n}$ una matriz de rango r y sea $\mathbf{B} \in \mathbb{R}^{m \times n}$ una matriz de rango k, para cualquier $k \leq r$ con $\hat{\mathbf{A}}(k) = \sum_{i=1}^{k} \sigma_i \mathbf{u}_o \mathbf{v}_i^T$ se muestra que:

$$\hat{\mathbf{A}}(k) = argmin_{rk(\mathbf{B})=k} \|\mathbf{A} - \mathbf{B}\|$$

$$\|\mathbf{A} - \hat{\mathbf{A}}\|_2 = \sigma^*_{k+1}$$

* La norma espectral de $\mathbf{A}$ es igual a su valor singular más grande σ_1

Notación

Notación

$\mathbf{X} \in \mathbb{R}^{N \times D}$: matriz de N observaciones
$\mathbf{x} \in \mathbb{R}^{D}$: vector D-dimensional de atributos/features/covariables
$f(\cdot)$: función $f : \mathbb{R}^D \to \mathbb{R}$
$\mathbf{y} \in \mathbb{R}^{N}$: vector N-dimensional. Variable dependiente, target.

Definiciones

Matriz ortogonal

Def. 3.8. MML Book

Una matriz cuadrada $\mathbf{A} \in \mathbb{R}^{n\times n}$ es una matriz ortogonal si y sólo si sus columnas son ortogonales, es decir:

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

$\Rightarrow$

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

Definiciones

Teorema Espectral

Teo. 4.15. MML Book

Si $\mathbf{A} \in \mathbb{R}^{n \times n}$ es simétrica, existe una base ortonormal del espacio vectorial V correspondiente que consiste en los autovectores de $\mathbf{A}$ y cada autovalor es real.

Colorario

La descomposición en valores propios de una matriz $\mathbf{A}$ existe, y puede encontrarse una base ortogonal de autovectores tal que $\mathbf{A} = \mathbf{PDP}^T$, donde $\mathbf{D}$ es diagonal y las columnas de $\mathbf{P}$ contienen los autovectores.

Definiciones

Norma espectral de una matriz

Teo. 4.23. MML Book

Para $\mathbf{x} \in \mathbb{R}^n \setminus \{\mathbf{0}\}$, la norma espectral de una matriz $\mathbf{A} \in \mathbb{R}^{m \times n}$ está definida cómo:

$$\|\mathbf{A}\|_2 \triangleq \max_{\mathbf{x}} \frac{\|\mathbf{A}\mathbf{x}\|}{\|\mathbf{x}\|}$$

La norma espectral determina la longitud máxima que puede tomar cualquier vector $\mathbf{x}$ al multiplicarse por $\mathbf{A}$.