Análisis Matemático para Inteligencia Artificial

Verónica Pastor (vpastor@fi.uba.ar), Martín Errázquin (merrazquin@fi.uba.ar)

Especialización en Inteligencia Artificial

Clase 4

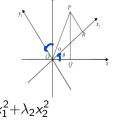
Autovalores y Autovectores: Motivación

Se acuerdan cuando vimos formas cuadráticas:

$$\begin{bmatrix} x & y \end{bmatrix} \begin{bmatrix} 2 & 3 \\ 3 & 9 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \underbrace{2x^2 + 6xy + 9y^2}_{x}$$

Si rotamos un ángulo θ la figura dada por: $2x^2 + 6xy + 9y^2 = 1$ a la forma canónica $\lambda_1 x_1^2 + \lambda_2 x_2^2 = 1$ Esto nos lleva a

una matriz de transición $S = \begin{bmatrix} cos(\theta) & -sen(\theta) \\ sen(\theta) & cos(\theta) \end{bmatrix}$ $x^{T} Ax = (S\tilde{x})^{T} A(S\tilde{x}) = \tilde{x}^{T} (S^{T} AS) \tilde{x} \Rightarrow x^{T} Ax = \lambda_{1} x_{1}^{2} + \lambda_{2} x_{2}^{2}$



XTAZ ER MZERT

$$(x,y) \stackrel{5}{\rightleftharpoons} (x_4, x_5)$$

Obs.: S resulta una matriz ortogonal, $S^T A S = S^{-1} A S = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} = D$

Autovalores y Autovectores: definición

Sea $A \in \mathbb{K}^{n \times n}$, \mathbb{R} o \mathbb{C} diremos que $\lambda \in \mathbb{K}$ es un autovalor (ava) de A y $x \in \mathbb{K}^n, x \neq 0$ es un autovector (ave) asociado a λ si:

$$Ax = \lambda x$$

Interpretación geométrica: A cada vector que se encuentre en la dirección de x, la transformación T(x) = Ax lo contrae (o expende) por un factor λ .

Importante: un autovalor puede ser nulo, un autovector no.

is a rave arrivado a
$$\lambda$$
 de A , rations and principal. x escalar tembria la $x = 3$ de x me importante $y = k \cdot x$ $\Rightarrow A \cdot y = A \cdot (k \cdot x) = k \cdot A \cdot x = k \cdot \lambda \cdot x = k \cdot \lambda$

Autovalores y Autovectores: algunas aplicaciones

eigen volues

eigenvectors

- Geometría: curvas planas o superficies.
- Sistemas dinámicos.
- Análisis del comportamiento de sistemas de ecuaciones diferenciales.
- Análisis de estabilidad.
- Cadenas de Markov.
- Grafos.
- Reducción de dimensiones.
- O Cálculo de resonancias del sistema.
- PageRank.

Hallando los autovalores

Dada
$$A \in \mathbb{K}^{n \times n}$$
, \mathbb{R} o \mathbb{C} , $\lambda \in \mathbb{K}$ es un autovalor de A sii λ es un cero del polinomio característico, $p(\lambda)$, de A :
$$p(\lambda) = \det(A - \lambda I)$$

$$(A - \lambda I)^{n \times n}$$

Observación: Puede ocurrir que algún λ sea raíz múltiple de $p(\lambda)$.

Se llama multiplicidad algebraica, m_a , a la cantidad de veces que λ aparece como raíz.

Se llama multiplicidad geométrica,
$$m_g$$
, del autovalor λ , a la cantidad de autovectores l.i. asociados a λ .

A: $\begin{pmatrix} 4 & 4 \\ 4 & 0 \end{pmatrix} \rightarrow p(\lambda)$: $L(A-\lambda I) = \begin{pmatrix} 4 & 1 \\ 4 & 2 \end{pmatrix} = \begin{pmatrix} 4\lambda + \lambda^2 + 4 \\ 4\lambda + 2\lambda^2 + 4 \end{pmatrix} = \begin{pmatrix} 4\lambda^2 + 4\lambda^2 + 4 \\ 4\lambda^2 + 2\lambda^2 +$

$$A^{-2}z^{2}x^{2} = \begin{pmatrix} 2 & -4 \\ 4 & -2 \end{pmatrix}\begin{pmatrix} 4 \\ 5 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \Rightarrow \begin{cases} 2x - 4y = 0 \\ 4x - 2y = 0 \end{cases} \Rightarrow \begin{cases} 2x - 4y = 0 \\ 4x - 2y = 0 \end{cases} \Rightarrow \begin{cases} 2x - 4y = 0 \\ 4x - 2y = 0 \end{cases} \Rightarrow \begin{cases} 2x - 4y = 0 \\ 4x - 2y = 0 \end{cases} \Rightarrow \begin{cases} 2x - 4y = 0 \\ 4x - 2y = 0 \end{cases} \Rightarrow \begin{cases} 2x - 4y = 0 \\ 4x - 2y = 0 \end{cases} \Rightarrow \begin{cases} 2x - 4y = 0 \\ 4x - 2y = 0 \end{cases} \Rightarrow \begin{cases} 2x - 4y = 0 \\ 4x - 2y = 0 \end{cases} \Rightarrow \begin{cases} 2x - 4y = 0 \\ 4x - 2y = 0 \end{cases} \Rightarrow \begin{cases} 2x - 4y = 0 \\ 4x - 2y = 0 \end{cases} \Rightarrow \begin{cases} 2x - 4y = 0 \\ 4x - 2y = 0 \end{cases} \Rightarrow \begin{cases} 2x - 4y = 0 \\ 4x - 2y = 0 \end{cases} \Rightarrow \begin{cases} 2x - 4y = 0 \\ 4x - 2y = 0 \end{cases} \Rightarrow \begin{cases} 2x - 4y = 0 \\ 4x - 2y = 0 \end{cases} \Rightarrow \begin{cases} 2x - 4y = 0 \\ 4x - 2y = 0 \end{cases} \Rightarrow \begin{cases} 2x - 4y = 0 \\ 4x - 2y = 0 \end{cases} \Rightarrow \begin{cases} 2x - 4y = 0 \\ 4x - 2y = 0 \end{cases} \Rightarrow \begin{cases} 2x - 4y = 0 \\ 4x - 2y = 0 \end{cases} \Rightarrow \begin{cases} 2x - 4y = 0 \\ 4x - 2y = 0 \end{cases} \Rightarrow \begin{cases} 2x - 4y = 0 \\ 4x - 2y = 0 \end{cases} \Rightarrow \begin{cases} 2x - 4y = 0 \\ 4x - 2y = 0 \end{cases} \Rightarrow \begin{cases} 2x - 4y = 0 \\ 4x - 2y = 0 \end{cases} \Rightarrow \begin{cases} 2x - 4y = 0 \\ 4x - 2y = 0 \end{cases} \Rightarrow \begin{cases} 2x - 4y = 0 \\ 4x - 2y = 0 \end{cases} \Rightarrow \begin{cases} 2x - 4y = 0 \\ 4x - 2y = 0 \end{cases} \Rightarrow \begin{cases} 2x - 4y = 0 \\ 4x - 2y = 0 \end{cases} \Rightarrow \begin{cases} 2x - 4y = 0 \\ 4x - 2y = 0 \end{cases} \Rightarrow \begin{cases} 2x - 4y = 0 \\ 4x - 2y = 0 \end{cases} \Rightarrow \begin{cases} 2x - 4y = 0 \\ 4x - 2y = 0 \end{cases} \Rightarrow \begin{cases} 2x - 4y = 0 \\ 4x - 2y = 0 \end{cases} \Rightarrow \begin{cases} 2x - 4y = 0 \\ 4x - 2y = 0 \end{cases} \Rightarrow \begin{cases} 2x - 4y = 0 \\ 4x - 2y = 0 \end{cases} \Rightarrow \begin{cases} 2x - 4y = 0 \\ 4x - 2y = 0 \end{cases} \Rightarrow \begin{cases} 2x - 4y = 0 \\ 4x - 2y = 0 \end{cases} \Rightarrow \begin{cases} 2x - 4y = 0 \\ 4x - 2y = 0 \end{cases} \Rightarrow \begin{cases} 2x - 4y = 0 \\ 4x - 2y = 0 \end{cases} \Rightarrow \begin{cases} 2x - 4y = 0 \\ 4x - 2y = 0 \end{cases} \Rightarrow \begin{cases} 2x - 4y = 0 \\ 4x - 2y = 0 \end{cases} \Rightarrow \begin{cases} 2x - 4y = 0 \\ 4x - 2y = 0 \end{cases} \Rightarrow \begin{cases} 2x - 4y = 0 \\ 4x - 2y = 0 \end{cases} \Rightarrow \begin{cases} 2x - 4y = 0 \\ 4x - 2y = 0 \end{cases} \Rightarrow \begin{cases} 2x - 4y = 0 \\ 4x - 2y = 0 \end{cases} \Rightarrow \begin{cases} 2x - 4y = 0 \\ 4x - 2y = 0 \end{cases} \Rightarrow \begin{cases} 2x - 4y = 0 \\ 4x - 2y = 0 \end{cases} \Rightarrow \begin{cases} 2x - 4y = 0 \\ 4x - 2y = 0 \end{cases} \Rightarrow \begin{cases} 2x - 4y = 0 \\ 4x - 2y = 0 \end{cases} \Rightarrow \begin{cases} 2x - 4y = 0 \\ 4x - 2y = 0 \end{cases} \Rightarrow \begin{cases} 2x - 4y = 0 \\ 4x - 2y = 0 \end{cases} \Rightarrow \begin{cases} 2x - 4y = 0 \\ 4x - 2y = 0 \end{cases} \Rightarrow \begin{cases} 2x - 4y = 0 \\ 4x - 2y = 0 \end{cases} \Rightarrow \begin{cases} 2x - 4y = 0 \\ 4x - 2y = 0 \end{cases} \Rightarrow \begin{cases} 2x - 4y = 0 \\ 4x - 2y = 0 \end{cases} \Rightarrow \begin{cases} 2x - 4y = 0 \\ 4x - 2y = 0 \end{cases} \Rightarrow \begin{cases} 2x - 4y = 0 \\ 4x - 2y = 0 \end{cases} \Rightarrow \begin{cases} 2x - 4y = 0 \\ 4x - 2y = 0 \end{cases} \Rightarrow \begin{cases} 2x - 4y = 0 \\ 4x - 2y = 0 \end{cases} \Rightarrow \begin{cases} 2x - 4y = 0 \\ 4x - 2y = 0 \end{cases} \Rightarrow \begin{cases} 2x - 4y = 0 \\ 4x - 2y = 0 \end{cases} \Rightarrow \begin{cases} 2x - 4y = 0 \\ 4x - 2y = 0 \end{cases} \Rightarrow \begin{cases} 2x - 4y = 0 \\ 4x - 2y = 0 \end{cases} \Rightarrow \begin{cases} 2x - 4y = 0 \end{cases} \Rightarrow \begin{cases} 2x - 4y = 0 \\ 2x - 2y = 0 \end{cases} \Rightarrow \begin{cases} 2x - 4y =$$

Propiedades de los autovectores

Dada $A \in \mathbb{K}^{n \times n}$, el conjunto de todos los autovectores asociados al autovalor λ generan un subespacio de \mathbb{K}^n (autoespacio de A respecto de λ). El conjunto de todos los autovectores de A se llama autoespectro.

A. (a.x+b·y)= a.A.x+b.A.y= a. λ ·×+b· λ ·y= λ ·(a.x+b·y)= λ ·2

Si λ es un autovalor de $A \in \mathbb{K}^{n \times n}$, el autoespacio E_{λ} asociado es la solución al sistema homogéneo $(A - \lambda I)x = 0$.

Teorema: Los autovectores $x_1,...,x_n$ de una matriz $A \in \mathbb{R}^{n \times n}$ con n autovalores distintos de $\lambda_1,...,\lambda_n$ son linealmente independientes. Es decir, forman una base de \mathbb{R}^n .

Teorema: Si $A \in \mathbb{R}^{n \times n}$ es simétrica, entonces existe una base ortonormal correspondiente a un espacio vectorial formado por autovectores de A y además $\lambda_i \in \mathbb{R}, \ \forall i=1,\ldots,n$.

Ejemplo

$$A = \begin{pmatrix} 2 & -5 \\ 2 & -5 \end{pmatrix} \in \mathbb{R}^{2 \times 2} \quad p(\lambda) = \begin{vmatrix} 2 - \lambda & -3 \\ 2 & -5 - \lambda \end{vmatrix} = (2 - \lambda) \begin{pmatrix} -5 - \lambda \end{pmatrix} + 6 = -40 + 3\lambda + \lambda^{2} + 6$$

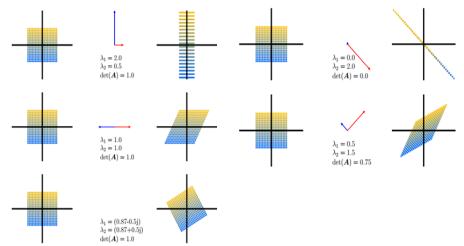
$$= \lambda^{2} + 3\lambda - 4 = 0$$

$$(\lambda - 1)(\lambda + 4) = 0$$

$$(\lambda -$$

Interpretación gráfica (Mathematics for Machine Learning)

Sea
$$A \in \mathbb{K}^{n \times n}$$
: Área: $det(A) = \prod_{i=1}^n \lambda_i$ y Perímetro: $Tr(A) = \sum_{i=1}^n \lambda_i$



Diagonalización

Diagonalización

Definición:

Una matriz $A \in \mathbb{K}^{n \times n}$ es diagonalizable si $\exists S \in \mathbb{K}^{n \times n}$ no singular (es decir, tiene inversa) tal que:

$$S^{-1}AS = D$$
 is inversible, criz diagonal.

donde D es una matriz diagonal. D diegoal/

Teorema: Sea $A \in \mathbb{K}^{n \times n}$ es diagonalizable sii A tiene n autovectores

linealmente independientes.

A B=
$$\frac{1}{2}$$
 $\frac{1}{2}$ $\frac{1}{2}$

Conclusiones para $A \in \mathbb{K}^{n \times n}$

- A es diagonalizable ⇒ los vectores columna de la matriz de diagonalización S son los autovectores de A y los elementos de D son los autovalores de A.
- ② S no es única, se pueden reordenar columnas, etc $\Rightarrow D$ será distinta.
- **3** A tiene n autovalores distintos \Rightarrow los autovectores correspondientes son l.i. \Leftrightarrow A es diagonalizable.
- **1** A tiene menos de n autovectores l.i. $\Leftrightarrow A$ no es diagonalizable.
- Si hay autovalores repetidos y si:
 - $m_a = m_g$ (en los ava repetidos) \Rightarrow los ave son l.i.
 - $m_a > m_g$ (en un ava repetido) $\Rightarrow A$ no es diagonalizable.
- **6** Si $A \in \mathbb{C}^{n \times n}$ es una matriz hermítica ($A = \overline{A^T} = A^H$) $\Rightarrow A$ es diagonalizable y las columnas forman una BON.

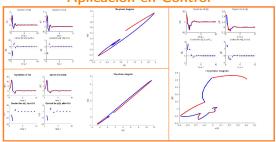


Matrices en bloques de Jordan

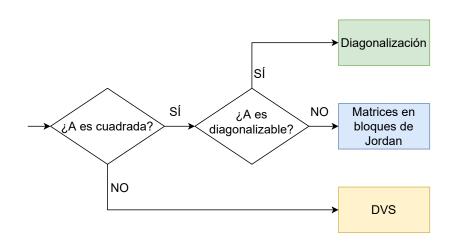
Sea $L: \mathbb{V} \to \mathbb{V}$ endomorfismo sobre $\mathbb{K} - e.v.$ con $dim(\mathbb{V}) = n$. Si $p(\lambda)$ se factoriza en \mathbb{K} entonces existe una base que viene dada por la matriz en bloques (m < n).

Obs.: Cuando es diagonalizable m = n.

Aplicación en Control



¿Qué puede pasar al intentar diagonalizar?



Descomposición en Valores Singulares (DVS)

¿Qué pasa cuando A no es cuadrada?

Teorema: Dada $A \in \mathbb{C}^{m \times n}$, siempre se puede obtener una matriz hermítica, semidefinida positiva $S \in \mathbb{C}^{n \times n}$ como:

$$S \doteq A^H A$$

Obs.: $\tilde{S} \in \mathbb{C}^{m \times m}$ definida como $\tilde{S} \doteq AA^H$ también resulta hermítica y semidefinida positiva.

Obs. (II): Como S, \tilde{S} son semidef. positivas, todos sus autovalores son reales **no negativos**.

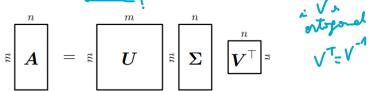
DVS: Definición

Teorema: Sea $A \in \mathbb{R}^{m \times n}$ una matriz de rango $r \in \{0, \dots, \min\{m, n\}\}$. La descomposición en valores singulares (DVS, SVD en inglés) de A se define como:

$$A = U \cdot \Sigma \cdot V^{T}$$
And the man of the

donde:

- $U \in \mathbb{R}^{m \times m}$ es la matriz ortogonal formada por los aves de AA^T .
- $\Sigma \in \mathbb{R}^{m \times n}$ es la matriz "diagonal" de valores singulares $\sigma_i = \sqrt{\lambda_i}$, con λ_i los primeros $p = \min\{m, n\}$ avas de AA^T y A^TA .
- $V \in \mathbb{R}^{n \times n}$ es la matriz ortogonal formada por los aves de $A^T A$.



Ejemplo: Hallar una DVS

Sea
$$A = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \in \mathbb{R}^{2+3}$$

$$A^{T} = \begin{pmatrix} 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \in \mathbb{R}^{2+3}$$

$$A^{T} = \begin{pmatrix} 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 0 \end{pmatrix}$$

Descomposición en Valores Singulares (DVS)

DVS Compacta, Reducida y Truncada

5,=2, 5,=1, 5,=0

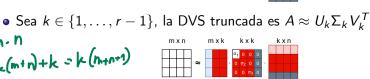
Sea $A \in \mathbb{R}^{m \times n}$ con DVS $A = U \Sigma V^T$:

• Si
$$p = min\{m, n\}$$
, la DVS compacta es $A = U_p \Sigma_p V_p^T$

• Si
$$\sigma_1, \ldots, \sigma_r > 0, \sigma_{r+1}, \ldots, \sigma_p = 0$$
, la DVS reducida es $A = U_r \Sigma_r V_r^T$

Sea
$$k \in \{1, ..., r - h, n \}$$

 $k(m+n)+k = k(n+n+1)$



19 / 21

Pseudoinversa de Moore-Penrose

Sea $A \in \mathbb{R}^{m \times n}$ y el sistema lineal Ax = y. La inversa de A no está definida, pero si m > n una posible solución (cuadrados mínimos) era

$$\hat{x} = (A^T A)^{-1} A^T y$$

¿Pero qué ocurre si A^TA no es invertible? Se define la pseudoinversa de Moore-Penrose como

$$A^{\dagger} = V_r \Sigma_r^{-1} U_r^T$$

y $\hat{x} = A^{\dagger}y$ es (si m < n) la solución de mínima norma euclídea o (si m > n) la aproximación de mínimo error cuadrático/distancia euclídea.

Obs.: Todo lo anteriormente visto también vale para $A \in \mathbb{C}^{m \times n}$, utilizando \cdot^H en vez de \cdot^T . Se utilizaron matrices reales para facilitar la lectura.

Colab

Una aplicación obvia para la DVS es, como vimos recién, la "solución" de sistemas de ecuaciones para los cuales el método de cuadrados mínimos no es factible.

¿Pero es ese el único uso que le podemos dar? Las descomposiciones siempre nos brindan un criterio bajo el cual obtener la mejor aproximación de $orden\ k$ para el objeto a descomponer, y eso podría llegar a ser muy útil.