Análisis Matemático para Inteligencia Artificial

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

Especialización en Inteligencia Artificial

Clase 3

Colab

El mundo del Data Science y las competencias de Machine Learning cobraron mucha relevancia con la competencia de Netflix, que prometía **1M USD** a quien pudiera obtener una performance de un 10% por sobre su algoritmo de base para el problema de recomendación de películas.

En este contexto surgió una nueva variante de modelo en base a factorización de matrices.

¿Lo interesante? Con las dos clases de AM ya nos alcanza para entenderlo!

Demostremos que: $x^T A x > 0 \quad \forall x \in \mathbb{R}^n / \{0\}$

Si $A \in \mathbb{R}^{n \times n}$ es una matriz definida positiva, entonces es invertible:
Dem.: (por método del absurdo)

Supongamos que el determinante de A es cero, y llegaremos a una solución absurda. $det(A) = 0 \Leftrightarrow \nexists A^{-1} \Longleftrightarrow \exists \tilde{x} \neq 0_v \,/\, A\tilde{x} = 0$

Si $det(A) = 0$ y queremos resolver un sistema $Ax = 0$ significa que además de la solución trivial hay infinitas soluciones, es decir $rango(A) < n$, entonces

$$\exists \tilde{x} \neq 0 : A\tilde{x} = 0 \Rightarrow \tilde{x}^T A \tilde{x} = \tilde{x}^T 0 = 0$$

y entonces no se cumple que $\tilde{x}^T A \tilde{x} > 0$.

Us A no invertible $\Rightarrow$ A NO def.+

$\Updownarrow$

A def+ $\Rightarrow$ A invertible

¿Por qué funciona la inversa? (2/2)

Demostremos que:

$A = x \cdot x^T + \lambda I_k$ **es definida positiva**, es decir, $y^T A y > 0, \ \forall y \in \mathbb{R}^k - \{0\}$

$= \sum x_i x_i^T + \lambda I_k = x_1 \cdot x_1^T + x_2 \cdot x_2^T + \dots + x_j \cdot x_j^T + \lambda I_k$

Dem.: Sea $y \in \mathbb{R}^k - \{0\}$

$$y^T \overbrace{(x \cdot x^T + \lambda I_k)}^{A} y = \underbrace{y^T x}_{<y,x>} \underbrace{x^T y}_{<x,y>} + y^T \lambda I_k y$$

(recordemos que el p.i. es $< u, v >= u^T v$), prop. asociativa,

$$< y, x >< x, y > + \lambda < y, y >= \underbrace{< x, y >^2}_{\geq 0} + \overset{\geq 0}{\lambda} \underbrace{||y||^2}_{>0} > 0.$$

$y \neq 0$

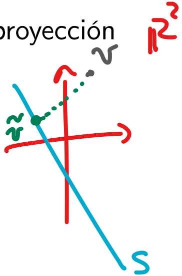
$\lambda \in \mathbb{R}^+$ por construcción

$\therefore x \cdot x^T + \lambda I_k$ es definida positiva.

Proyección Ortogonal

Proyección

Sea $\mathbb{V}$ un EV y $S \subset \mathbb{V}$ un SEV. Una transformación lineal $\Pi : \mathbb{V} \to S$ es una proyección si $\Pi^2 = \Pi \circ \Pi = \Pi$
Esto significa que la matriz de transformación asociada a la proyección cumple con la propiedad de idempotencia: $[\Pi]^2 = [\Pi]$.



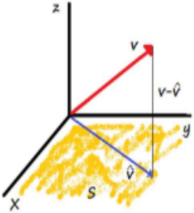
Proyección Ortogonal

Dado $\mathbb{V}$ un EV con p.i. y $S \subset \mathbb{V}$ un SEV, el objetivo es dado $v \in \mathbb{V}$ hallar $\tilde{v} \in S$ que sea "lo más parecido posible" a v.

$v - \tilde{v} \in S^{\perp}$

$$\tilde{v} \in S : \tilde{v} = arg\ min_{s \in S} ||v - s||$$

Además vale que $\langle v - \tilde{v}, s \rangle = 0, \ \forall s \in S$



Obs: La ortogonalidad de la proyección tiene que ver con el p.i. que se use.

Teorema de proyección

Sea $\mathbb{V}$ un EV de dimensión finita con p.i. $\langle .,. \rangle$, S un SEV. Dado $v \in \mathbb{V}$ existe un único $\tilde{v} \in S$ tal que

$\exists$ $!$

$$||v - \tilde{v}|| \leq ||v - u||, \quad \forall u \in \mathbb{S}$$

la proy. ortogonal es única $\forall v$

¿Cómo hallar la proyección?

Sea $\mathbb{V}$ un EV de dimensión n con p.i. $\langle .,. \rangle$, y $S \subset \mathbb{V}$ un SEV, $dim(S) = m \geq 1$, y sea $B = \{s_1, ..., s_m\}$ una BON de S. Buscamos encontrar la proyección de $\tilde{v} \in S$ de $v \in \mathbb{V}$ ($\tilde{v} = \Pi_S(v)$).

Como $\tilde{v} \in S$, $\tilde{v} = \sum_{i=1}^{m} \alpha_i s_i \Rightarrow$ busco los coeficientes que minimizan $||v - \underbrace{\sum_{i=1}^{m} \alpha_i s_i}_{\tilde{v}}||$. El problema puede escribirse como:

$$\tilde{v} = \Pi_S(v) = \sum_{i=1}^{m} \alpha_i s_i = B\alpha, \quad B = [s_1, ..., s_m], \ \alpha = [\alpha_1, ..., \alpha_m]^T$$

x ej. $\tilde{v}, v \in \mathbb{R}^n$

$B = \begin{pmatrix} | & | & & | \\ s_1 & s_2 & \cdots & s_m \\ | & | & & | \end{pmatrix} \in \mathbb{R}^{n \times m}$ $\quad \alpha = \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_m \end{pmatrix} \in \mathbb{R}^m$

¿Cómo hallar la proyección?

Como por definición $\langle v - \Pi_S(v), s \rangle = 0$, $\forall s \in S$, debo resolver el siguiente sistema de ecuaciones:

$$\begin{cases} \langle v - \Pi_S(v), s_1 \rangle = 0 \\ \vdots \\ \langle v - \Pi_S(v), s_m \rangle = 0 \end{cases} \Rightarrow \begin{cases} s_1^T(v - B\alpha) = 0 \\ \vdots \\ s_m^T(v - B\alpha) = 0 \end{cases} \Rightarrow \overbrace{\begin{pmatrix} -s_1^T- \\ \vdots \\ -s_m^T- \end{pmatrix}}^{B^T} (v - B\alpha) = 0$$

$$B^T(v - B\alpha) = 0 \Leftrightarrow B^Tv = B^TB\alpha \Leftrightarrow \alpha = (B^TB)^{-1}B^Tv$$

$$\Pi_S = B(B^TB)^{-1}B^T$$

$$\tilde{v} = B\cdot\alpha = \underbrace{B\cdot(B^TB)^{-1}\cdot B^T}_{\Pi_S} v$$

Observación: Si B es una BON entonces $P_{\Pi_S} = BB^T$.

n

m $\quad$ m>n

$s \in S \quad v \in V$

Supongamos que tenemos un sistema sobredeterminado de la forma: $\min \|\tilde{y} - y\|$

$$Ab = y, \; A \in \mathbb{R}^{m \times n}, \; b \in \mathbb{R}^n, \; y \in \mathbb{R}^m, \; m > n.$$

$S = Col(A)$ $\qquad$ *b* es la solución de cuadrados mínimos $\qquad$ $\underbrace{A}_{B} \cdot \underbrace{\hat{b}}_{\alpha} = \tilde{y}$

Como $m > n$, puede que no exista b que satisfaga todas las m ecuaciones, entonces busco la solución que más se acerque (busco $Proy_{Col(A)}y$)

$$b = \underbrace{(A^TA)^{-1}A^Ty}_{\hat{\beta}} \Rightarrow P = \underbrace{A(A^TA)^{-1}A^T}_{\Pi_S}$$

$\hat{y} = B \cdot \alpha = A \cdot \hat{\beta}$

Colab

Ya vimos que el ajuste por cuadrados mínimos funciona en la teoría, ahora veámoslo funcionar en la práctica!