Análisis Matemático para Inteligencia Artificial

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

Especialización en Inteligencia Artificial

4/11/2022

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!

¿Por qué funciona la inversa? (1/2)

simétrica y $x^T A x > 0$ $\forall \tilde{x} \neq 0$

Demostremos que:

Si $A \in \mathbb{R}^{n\times n}$ es una matriz definida positiva, entonces es invertible:

$AA^T = A^T A = I$

Dem.: (por método del absurdo)

Supongamos que el determinante de A es cero, y llegaremos a una solución absurda.

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 $\tilde{x}^T A\tilde{x} > 0$.

$$\underbrace{u\left(\sum_{j\in\Omega_i} \underbrace{v_j v_j^T}_{k\times k} + \lambda I_k\right)}_{} = \sum_{j\in\Omega_i} r_{ij} v_j^T{}_{1\times k}$$

$$u_{1\times k} \cdot \underbrace{A_{k\times k}}_{\text{¿es def. positiva?}} = b_{1\times k}$$

A^{-1} A^{-1}

$xA = b$

$(Ax)^T = b^T$

Demostremos que:

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

A es simétrica ?

simétrica

Suma de MATRICES SIMÉTRICAS ES SIMÉTRICA

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

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

distrib. (esp. vectorial!)

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

$\lambda > 0$ parámetro del modelo

$$< y, x >< x, y > + \lambda < y, y > = < x, y >^2 + \lambda ||y||^2 > 0.$$

p.i conmutativo (en $\mathbb{R}^k$)

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

Proyección

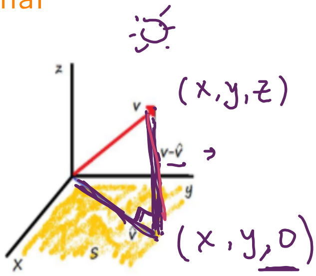
Sea $\mathbb{V}$ un EV y $S \subset \mathbb{V}$ un SEV. Una transformación lineal $\Pi : \mathbb{V} \to S$ es una proyección si $\Pi^2 = \Pi \circ \Pi = \Pi$ idempotente
Esto significa que la matriz de transformación asociada a la proyección cumple con la propiedad de idempotencia: $[\Pi]^2 = [\Pi]$.

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.

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

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

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

$s_i \in \mathbb{V}$ tiene n comp.

¿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 $v \in \mathbb{V}$ ($\tilde{v} = \Pi_S(v)$).

α_i

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

matricialmente

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

$$\underset{n \times m}{\boxed{\begin{bmatrix} s_1, \dots & s_m \end{bmatrix}}} \underset{m \times 1}{\begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_m \end{bmatrix}} =$$

$\tilde{v} = B\alpha \rightarrow$?? matriz ¿$\exists B^{-1}$? (*)

(*) $\exists B^{-1}$ porque los vectores de B son l.i y $\det(B) \neq 0$

¿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:

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

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

$B^T v - B^T B\alpha = 0$ $\quad$ I $\quad$ $\exists B^{-1} \Rightarrow (B^{-1})^T = (B^T)^{-1}$

$\tilde{v} = \Pi_S(v)$ $\quad$ $\cancel{P_{\Pi_S}} = \underbrace{B(B^T B)^{-1}B^T v}_{\alpha}$

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

$\|v\| = v^T v = 1$

Supongamos que tenemos un sistema sobredeterminado de la forma:

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

b es la solución de cuadrados mínimos

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

proyección sobre S donde S es sub. Col(A)

mxn nxm mxn nxn

$$b = (A^T A)^{-1} A^T y \;\Rightarrow\; P = \underbrace{A(\underbrace{A^T A}_{n \times n})^{-1} A^T}$$

MATRIZ de la T.L. Proyección ortogonal

P es MATRIZ de mxm

Colab

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