

ISI306 - Aprendizaje Automático

Repaso de álgebra lineal

Ing. Jose Eduardo Laruta Espejo

18 de agosto de 2020

Universidad La Salle - Bolivia

1. Conceptos básicos y Notación
2. Multiplicación de matrices
3. Operaciones y propiedades
4. Cálculo y Matrices

Conceptos básicos y Notación

Conceptos básicos y Notación

El campo del álgebra lineal nos presenta formas muy eficientes de representar y realizar operaciones con grupos de números ordenados o conjuntos de ecuaciones lineales, por ejemplo:

$$\begin{aligned}4x_1 - 5x_2 &= -13 \\ -2x_1 + 3x_2 &= 9\end{aligned}$$

El mismo conjunto de ecuaciones se puede representar en forma matricial $Ax = b$ siendo:

$$A = \begin{bmatrix} 4 & -5 \\ -2 & 3 \end{bmatrix}, b = \begin{bmatrix} -13 \\ 9 \end{bmatrix}$$

Se usará la siguiente notación:

- $A \in \mathbb{R}^{m \times n}$ denota una matriz con m filas y n columnas, donde los elementos de A son reales.
- $x \in \mathbb{R}^n$ denota un vector con n elementos. Por convención un vector n -dimensional también se representa como una matriz con n filas y 1 columna.
- El i ésimo elemento de un vector x se denota como x_i .

$$x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \quad (1)$$

- a_{ij} (o $A_{ij}, A_{i,j}$) es un elemento de A en la fila i y la columna j :

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{12} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \quad (2)$$

- a_j o $A_{:,j}$ es la j -ésima columna de la matriz A :

$$A = \begin{bmatrix} | & | & \cdots & | \\ a_1 & a_2 & \cdots & a_n \\ | & | & & | \end{bmatrix} \quad (3)$$

- a_i^T o $A_{i,:}$ es la i -ésima fila de la matriz A :

$$A = \begin{bmatrix} - & a_1^T & - \\ - & a_2^T & - \\ & \vdots & \\ - & a_m^T & - \end{bmatrix} \quad (4)$$

Multipliación de matrices

Multiplicación de matrices

Dadas las matrices $A \in \mathbb{R}^{m \times n}$ y $B \in \mathbb{R}^{n \times p}$ el producto matricial es:

$$C = AB \in \mathbb{R}^{m \times p} \quad (5)$$

donde:

$$C_{ij} = \sum_{k=1}^n A_{ik} B_{kj} \quad (6)$$

nótese que, para que el producto matricial exista, el número de columnas de A debe ser igual al número de filas de B . En tal caso se denominan matrices conformables.

Dados dos vectores $x, y \in \mathbb{R}^n$, el producto $x^T y$ se suele llamar el **producto interno** o **producto escalar**, el resultado es un escalar:

$$x^T y \in \mathbb{R} = [x_1 \ x_2 \ \cdots \ x_n] \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = \sum_{i=1}^n x_i y_i \quad (7)$$

El producto escalar es solamente un caso especial de un producto matricial.

Producto de vectores

Por otro lado, dados 2 vectores $x \in \mathbb{R}^m, y \in \mathbb{R}^n$ (no del mismo tamaño necesariamente) el resultado de xy^T se conoce como el **producto externo** o **producto vectorial** y es una matriz dada por $(xy^T)_{ij} = x_i y_j$

$$xy^T \in \mathbb{R}^{m \times n} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{bmatrix} [y_1 \ y_2 \ \cdots \ y_n] = \begin{bmatrix} x_1 y_1 & x_1 y_2 & \cdots & x_1 y_n \\ x_2 y_1 & x_2 y_2 & \cdots & x_2 y_n \\ \vdots & \vdots & \ddots & \vdots \\ x_m y_1 & x_m y_2 & \cdots & x_m y_n \end{bmatrix} \quad (8)$$

Esta operación puede utilizarse para distintos casos, por ejemplo: repetir un vector en columnas de una matriz.

Producto matriz-vector

Dada la matrices $A \in \mathbb{R}^{m \times n}$ y un vector $x \in \mathbb{R}^n$, el producto da como resultado otro vector $y = Ax \in \mathbb{R}^m$

$$y = Ax = A = \begin{bmatrix} - & a_1^T & - \\ - & a_2^T & - \\ & \vdots & \\ - & a_m^T & - \end{bmatrix} x = \begin{bmatrix} a_1^T x \\ a_2^T x \\ \vdots \\ a_m^T x \end{bmatrix} \quad (9)$$

En este caso, el i -ésimo elemento de y es el producto escalar de la i -ésima fila de A y el vector x . También se puede hacer una pre multiplicación por un vector.

Producto de matrices

Entendiendo los productos entre vectores, el producto matricial se puede expresar en función de distintos productos escalares, dadas las matrices

$A \in \mathbb{R}^{m \times n}$ y $B \in \mathbb{R}^{n \times p}$:

$$C = AB = \begin{bmatrix} \text{---} & a_1^T & \text{---} \\ \text{---} & a_2^T & \text{---} \\ & \vdots & \\ \text{---} & a_m^T & \text{---} \end{bmatrix} \begin{bmatrix} | & | & \cdots & | \\ b_1 & b_2 & \cdots & b_p \\ | & | & & | \end{bmatrix} = \begin{bmatrix} a_1^T b_1 & a_1^T b_2 & \cdots & a_1^T b_p \\ a_2^T b_1 & a_2^T b_2 & \cdots & a_2^T b_p \\ \vdots & \vdots & \ddots & \vdots \\ a_m^T b_1 & a_m^T b_2 & \cdots & a_m^T b_p \end{bmatrix} \quad (10)$$

Es importante construir un entendimiento intuitivo de las operaciones entre matrices y vectores por su importancia en el desarrollo de los algoritmos de aprendizaje automático que se explorarán en el futuro y en su implementación en python. Adicionalmente, la multiplicación de matrices posee las siguientes propiedades:

- Asociatividad: $(AB)C = A(BC)$
- Distributividad: $A(B + C) = AB + AC$
- No es conmutativa: $AB \neq BA$

Operaciones y propiedades

La **matriz identidad**, denotada por $I \in \mathbb{R}^{n \times n}$, es una matriz cuadrada con unos en la diagonal y ceros en otras posiciones:

$$I_{ij} = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases} \quad (11)$$

y tiene la propiedad que para toda matriz $A \in \mathbb{R}^{m \times n}$,

$$AI = A = IA \quad (12)$$

Generalmente, las dimensiones de la matriz identidad I se infieren del contexto en el que se usa.

En una **matriz diagonal** todos los elementos que no pertenecen a la diagonal son iguales a 0. Se suele denotar como $D = \text{diag}(d_1, d_2, \dots, d_n)$, donde:

$$D_{ij} = \begin{cases} d_i & i = j \\ 0 & i \neq j \end{cases} \quad (13)$$

Claramente, $I = \text{diag}(1, 1, \dots, 1)$

Matriz Traspuesta

La traspuesta de una matriz resulta de “intercambiar” las filas con las columnas. Dada una matriz $A \in \mathbb{R}^{m \times n}$, su traspuesta $A^T \in \mathbb{R}^{n \times m}$ es una matriz con dimensiones $n \times m$ donde sus elementos estan definidos por:

$$(A^T)_{ij} = A_{ji} \quad (14)$$

Tiene las siguientes propiedades:

- $(A^T)^T = A$
- $(AB)^T = B^T A^T$
- $(A + B)^T = A^T + B^T$

Una matriz cuadrada es **simétrica** si $A = A^T$. Y es **antisimétrica** si $A = -A^T$.

La inversa de una matriz $A \in \mathbb{R}^{m \times n}$, denotada por A^{-1} y cumple:

$$A^{-1}A = I = AA^{-1} \quad (15)$$

No todas las matrices tienen inversa, por ejemplo, matrices no cuadradas, e incluso en el caso de algunas matrices cuadradas, la inversa puede no existir. Si A posee inversa se dice que es una matriz **invertible** o **no singular**, en otro caso es **no invertible** o **singular**

La inversa de una matriz $A \in \mathbb{R}^{m \times n}$, denotada por A^{-1} y cumple:

$$A^{-1}A = I = AA^{-1} \quad (16)$$

No todas las matrices tienen inversa, por ejemplo, matrices no cuadradas, e incluso en el caso de algunas matrices cuadradas, la inversa puede no existir. Si A posee inversa se dice que es una matriz **invertible** o **no singular**, en otro caso es **no invertible** o **singular**

Cálculo y Matrices

Preguntas?

Preguntas?