Repaso de Álgebra Lineal y Aprendizaje Supervisado

Instituto Tecnológico de Costa Rica
Escuela de Ingeniería en Computación
Inteligencia Artificial

Mariana Quesada Sánchez
19 de agosto de 2024

***Abstract*— This paper reviews concepts of linear algebra relevant to artificial intelligence, including vectors, norms, distances, dot product, orthogonality, and orthonormality. It also introduces the principles of supervised learning, describing datasets as feature–label pairs and distinguishing between regression and classification tasks through illustrative examples.**

I. INTRODUCTION

El álgebra lineal es la base para representar datos en espacios multidimensionales y para definir operaciones que permiten medir magnitudes, direcciones y similitudes. Estos fundamentos son indispensables en algoritmos de machine learning, en particular dentro del aprendizaje supervisado, donde los datos se representan como vectores de características asociados a etiquetas.

II. ÁLGEBRA LINEAL

A. Vectores

Un vector se define como una entidad matemática caracterizada por magnitud y dirección. En espacios de dos o tres dimensiones, puede visualizarse como un segmento orientado que parte del origen y termina en un punto *(x,y,z)*. En espacios de dimensión n, se representa como una tupla ordenada *(*x_1*,* x_2*, …,* x_n*)*. Los vectores constituyen la base de la representación de datos en espacios multidimensionales y permiten operaciones como la suma, la resta y la multiplicación por escalares.

El desplazamiento de un vector se define como la diferencia entre un punto final $B = (b_1, b_2, \ldots, b_n)$ y un punto inicial $A = (a_1, a_2, \ldots, a_n)$. Formalmente, el vector desplazamiento se expresa como

$$\vec{AB} = B - A = (b_1 - a_1,\ b_2 - a_2,\ \ldots,\ b_n - a_n),$$

lo cual indica cuánto debe recorrerse en cada componente para pasar de A a B. Por ejemplo, si $A = (1, 2)$ y $B = (4, 6)$, entonces $\vec{AB} = (3, 4)$, lo que representa un movimiento de tres unidades en el eje x y cuatro en el eje y.

Es importante distinguir entre un vector de posición y un vector de desplazamiento. Un vector de posición ubica un punto específico en el espacio con respecto al origen, mientras que un vector de desplazamiento describe el cambio necesario para trasladarse de un punto a otro. Por ejemplo, el vector $(4, 3)$ puede interpretarse como la posición de un punto en el plano cartesiano, pero también puede representar el desplazamiento requerido para pasar del origen $(0, 0)$ hasta dicho punto.

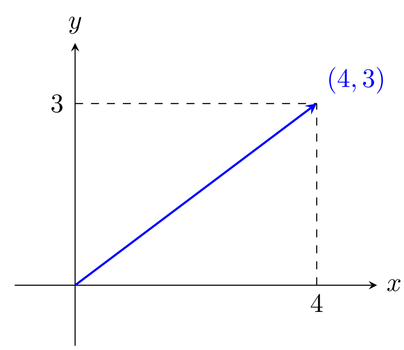

Fig. 1. Representación gráfica del vector $(4, 3)$ en el plano cartesiano.

B. Norma o magnitud

La norma mide longitud de un vector y se denota como $\|x\|$. Geométricamente, puede interpretarse como la distancia desde el punto de origen hasta el punto final definido por x. De esta manera, la norma proporciona una medida cuantitativa de la magnitud del vector, independientemente de su dirección.

Para un vector $x = (x_1, x_2, \ldots, x_n)$, las normas más comunes son:

- **Norma L1 o Manhattan:**
  La distancia Manhattan entre dos puntos $A = (x_1, y_1, z_1, \ldots, n_1)$ y $B = (x_2, y_2, z_2, \ldots, n_2)$ en un espacio n-dimensional se calcula mediante la fórmula:

  $$\|x\|_1 = \sum_{i=1}^{n} |x_i - y_i|$$

  Se interpreta como la distancia recorrida siguiendo los ejes de la cuadrícula.

- **Norma L2 o Euclidiana:**

$$\|x\|_2 = \sqrt{\sum_{i=1}^{n} x_i^2}$$

Corresponde a la distancia en línea recta entre el origen y el punto final, de acuerdo con el teorema de Pitágoras.

Fig. 2. Comparación entre distancia Manhattan y Euclidiana [1].

Una función es considerada una norma si cumple las siguientes propiedades:

1) *Positividad:* $\|x\| \geq 0$ y $\|x\| = 0$ si y sólo si x es el vector nulo.
2) *Homogeneidad:* $\|\alpha x\| = |\alpha|\|x\|$ para cualquier escalar $\alpha \in R$.
3) *Desigualdad triangular:* $\|x + y\| \leq \|x\| + \|y\|$ para todos los vectores x e y.

C. Vectores unitarios

Un vector unitario es aquel cuya norma es igual a uno. Se obtiene normalizando un vector v mediante su magnitud:

$$u = \frac{v}{\|v\|}$$

De esta manera, u conserva la dirección de v, pero con longitud unitaria.

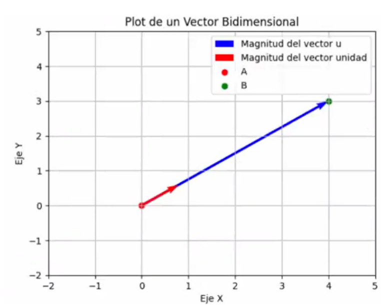

Fig. 3. Vector unitario. [2]

D. Producto punto

El producto punto entre dos vectores es la suma de las multiplicaciones de sus componentes, lo que produce un valor real. Esta operación es fundamental en inteligencia artificial, ya que un vector puede representar características de los datos y otro vector puede representar los pesos asociados a dichas características. Si un peso es cero, la característica correspondiente no contribuye al resultado.

El producto punto entre dos vectores u y v se define de dos formas equivalentes:

- **Definición algebraica:**

$$u \cdot v = \sum_{i=1}^{n} u_i v_i$$

Ej 1. Sea

$$x = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}, \quad y = \begin{bmatrix} 4 \\ 5 \\ 6 \end{bmatrix}.$$

Calculamos el producto punto:

$$x^T y = \begin{bmatrix} 1 & 2 & 3 \end{bmatrix} \begin{bmatrix} 4 \\ 5 \\ 6 \end{bmatrix}$$

$$= 1 \cdot 4 + 2 \cdot 5 + 3 \cdot 6 = 4 + 10 + 18 = \boxed{32}$$

- **Definición geométrica:**

$$u \cdot v = \|u\|\|v\| \cos(\theta)$$

donde θ es el ángulo entre ambos vectores.

Ej 2. Sea

$$u = \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \quad v = \begin{bmatrix} 3 \\ 4 \end{bmatrix}.$$

Paso 1: Calcular el producto punto

$$u \cdot v = 1 \cdot 3 + 2 \cdot 4 = 3 + 8 = 11$$

Paso 2: Calcular las normas

$$\|u\| = \sqrt{1^2 + 2^2} = \sqrt{5}, \qquad \|v\| = \sqrt{3^2 + 4^2} = \sqrt{25} = 5$$

Finalmente, usando la definición geométrica del producto punto:

$$11 = \sqrt{5} \cdot 5 \cdot \cos(\theta) \quad \Rightarrow \quad \cos(\theta) = \frac{11}{5\sqrt{5}} \approx 0.9839$$

$$\theta = \cos^{-1}(0.9839) \approx 10.3^\circ$$

E. Vectores codireccionales

Dos vectores se consideran codireccionales cuando mantienen la misma dirección, aunque difieran en magnitud. Esta relación se cumple si existe un escalar k tal que $v = k \cdot u$. En este caso, el ángulo entre ambos vectores es nulo y el coseno del ángulo es igual a uno. El vector unitario es un caso particular, ya que al ser multiplicado por un escalar recupera la magnitud del vector original sin alterar su dirección.

Se sabe que si dos vectores son codireccionales, el ángulo entre ellos es de 0°. En consecuencia, el producto punto se expresa como

$$u \cdot u = \|u\| \cdot \|u\| \cdot \cos(0) = \|u\|^2.$$

De esta forma, la norma de un vector puede escribirse como

$$\|u\| = \sqrt{u \cdot u}.$$

Asimismo, la distancia euclidiana puede expresarse en términos de producto punto:

$$\sqrt{u \cdot u} = \sqrt{\|u\|^2} = \|u\|.$$

F. Ortogonalidad y ortonormalidad

Dos vectores son ortogonales si su producto punto es cero:

$$u \cdot v = 0$$

Además, un conjunto de vectores es ortonormal si además de ser ortogonales, cada vector es unitario.

III. APRENDIZAJE SUPERVISADO

El aprendizaje supervisado consiste en entrenar un modelo a partir de un conjunto de datos donde cada ejemplo se encuentra representado por un vector de características x_i y una etiqueta asociada y_i. Las características describen propiedades cuantificables del fenómeno observado, mientras que la etiqueta corresponde al valor que se desea predecir.

Existen dos tareas principales dentro del aprendizaje supervisado

La regresión busca predecir valores continuos, como el precio de una vivienda en función de atributos como área, número de habitaciones o ubicación. La *Fig. 4* corresponde a un problema de regresión porque busca ajustar una función que modele la relación entre una variable independiente (carat) y una variable dependiente continua (precio).

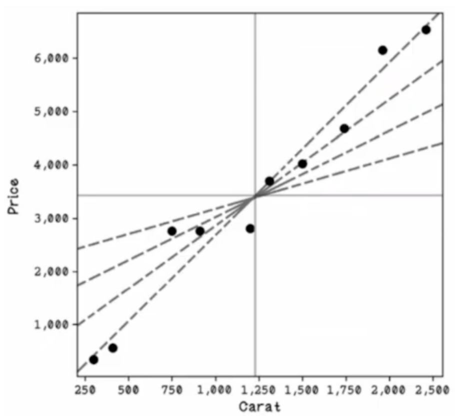

Fig. 4. Ejemplo de regresión. [2].

La clasificación, en cambio, asigna cada instancia a una categoría discreta a partir de sus características. Por ejemplo, predecir el tipo de vehículo dependiendo de cuántas llantas tiene y cuánto pesa.

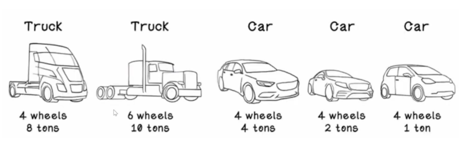

Fig. 5. Ejemplo de clasificación. [2].

En ambos casos, el objetivo es construir un modelo que generalice más allá de los datos de entrenamiento y que logre realizar predicciones confiables sobre ejemplos no observados.

A. Ejercicios de aprendizaje supervisado resueltos

1) Ratas: esperanza de vida vs. obesidad
Dado un conjunto de datos con dos características (esperanza de vida y obesidad) se busca modelar la relación entre ambas.
Tipo: Regresión (la salida es un valor continuo).
2) Animales: identificar aves
Tomando en cuenta datos sobre animales, el peso y si tiene alas, se desea determinar cuáles son pájaros.
Tipo: Clasificación (se clasifica cada ejemplar como ave o no empleando el peso y la presencia de alas).
3) Dispositivos: tablet, laptop o teléfono
Con tamaño de pantalla, peso y sistema operativo, se debe asignar cada dispositivo a una de varias categorías.
Tipo: Clasificación.
4) Meteorología: precipitación $\rightarrow$ humedad.
Con cantidad de precipitación y un valor de humedad, se desea predecir la humedad en distintas épocas del año.
Tipo: Regresión (humedad como variable continua).

REFERENCES

[1] S. Rani and G. Sikka, “Recent Techniques of Clustering of Time Series Data: A Survey,” *Artificial Intelligence Review*, vol. 46, no. 1, pp. 27–44, 2016. Available: https://www.researchgate.net/figure/Comparative-between-Euclidean-and-Manhattan-distance$_{fig}1_3$32432569

[2] S. Pacheco, “Repaso de Matemática: Álgebra Lineal,” Presentación, Instituto Tecnológico de Costa Rica, 2025.