Repaso matemáticas general

Sandra Mingo Ramírez

2024/25

Índice general

1. Álgebra lineal

1.1. Vectores

Un vector es una cantidad que tiene tanto magnitud como dirección. En el contexto de álgebra lineal, un vector es una lista ordenada de números. Un vector se puede representar como:

$$\mathbf{v} = \begin{pmatrix} v_1 \\ v_2 \\ \dots \\ v_n \end{pmatrix}$$

Suma de vectores Dos vectores del mismo tamaño se pueden sumar componente a componente.

$$\mathbf{u} + \mathbf{v} = \begin{pmatrix} u_1 + v_1 \\ u_2 + v_2 \\ \dots \\ u_n + v_n \end{pmatrix}$$

Multiplicación por un escalar Un vector se puede multiplicar por un número (escalar) multiplicando cada componente por ese número.

$$c\mathbf{v} = \begin{pmatrix} cv_1 \\ cv_2 \\ \dots \\ cv_n \end{pmatrix}$$

Producto punto (o producto escalar) Dados dos vectores $\mathbf{u}$ y $\mathbf{v}$, su producto punto es:

$$\mathbf{u} \cdot \mathbf{v} = u_1 v_1 + u_2 v_2 + \ldots + u_n v_n$$

El producto punto es útil para calcular ángulos entre vectores y para proyectar un vector sobre otro.

1.2. Matrices

Una matriz es una disposición rectangular de números en filas y columnas. Las matrices son fundamentales en álgebra lineal y se utilizan para representar transformaciones lineales, sistemas de ecuaciones lineales, y más. Una matriz $\mathbf{A}$ de tamaño $m \times n$ se representa como:

$$\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix}$$

donde a_{ij} es el elemento en la fila i y la columna j.

Suma de matrices Dos matrices del mismo tamaño se pueden sumar elemento a elemento.

$$\mathbf{A} + \mathbf{B} = \begin{pmatrix} a_{11} + b_{11} & a_{12} + b_{12} & \dots & a_{1n} + b_{1n} \\ a_{21} + b_{21} & a_{22} + b_{22} & \dots & a_{2n} + b_{2n} \\ \dots & \dots & \dots & \dots \\ a_{m1} + b_{m1} & a_{m2} + b_{m2} & \dots & a_{mn} + b_{mn} \end{pmatrix}$$

Multiplicación por un escalar Una matriz se puede multiplicar por un escalar multiplicando cada elemento de la matriz por ese escalar.

$$c\mathbf{A} = \begin{pmatrix} ca_{11} & ca_{12} & \dots & ca_{1n} \\ ca_{21} & ca_{22} & \dots & ca_{2n} \\ \dots & \dots & \dots & \dots \\ ca_{m1} & ca_{m2} & \dots & ca_{mn} \end{pmatrix}$$

Multiplicación de matrices Dadas dos matrices $\mathbf{A}$ de tamaño $m \times n$ y $\mathbf{B}$ de tamaño $n \times p$, su producto $\mathbf{C} = \mathbf{AB}$ es una matriz de tamaño $m \times p$ donde cada elemento c_{ij} se calcula como:

$$c_{ij} = \sum_{k=1}^{n} a_{ik} b_{kj}$$

La multiplicación de matrices no es conmutativa, es decir, $\mathbf{AB} \neq \mathbf{BA}$ en general.

Transpuesta de una matriz La transpuesta de una matriz $\mathbf{A}$, denotada como $\mathbf{A}^T$, se obtiene intercambiando filas por columnas.

$$\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix} \rightarrow \mathbf{A}^T = \begin{pmatrix} a_{11} & a_{21} & \dots & a_{m1} \\ a_{12} & a_{22} & \dots & a_{m2} \\ \dots & \dots & \dots & \dots \\ a_{1n} & a_{2n} & \dots & a_{mn} \end{pmatrix}$$

Determinante de una matriz El determinante es un valor escalar que se puede calcular a partir de los elementos de una matriz cuadrada. Es útil para determinar si una matriz tiene inversa y en la resolución de sistemas de ecuaciones lineales.

Para una matriz 2×2:

$$\det \mathbf{A} = \begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc$$

Para matrices más grandes, el cálculo del determinante es más complejo y se puede hacer mediante expansión por cofactores.

Inversa de una matriz La inversa de una matriz cuadrada $\mathbf{A}$, denotada como $\mathbf{A}^{-1}$, es una matriz tal que:

$$\mathbf{AA}^{-1} = \mathbf{A}^{-1}\mathbf{A} = \mathbf{I}$$

donde $\mathbf{I}$ es la matriz de identidad. No todas las matrices tienen inversa; una matriz tiene inversa si y solo si su determinante es distinto de cero.

Norma de una matriz Las normas de una matriz son herramientas matemáticas que nos permiten cuantificar el "tamaño" o la "magnitud" de una matriz. En el contexto del aprendizaje automático (Machine Learning, ML), las normas de matrices son útiles para regularizar modelos, medir errores, y entender la estabilidad de los algoritmos.

Una norma de matriz es una función que asigna un número real no negativo a una matriz, cumpliendo ciertas propiedades:

1. **No negatividad:** $||\mathbf{A}|| \geq 0$ y $||\mathbf{A}|| = 0$ si y solo si $\mathbf{A} = 0$
2. **Homogeneidad:** $||c\mathbf{A}|| = |c| \cdot ||\mathbf{A}||$ para cualquier escalar c

3. **Desigualdad triangular:** $||\mathbf{A} + \mathbf{B}|| \leq ||\mathbf{A}|| + ||\mathbf{B}||$
4. **Submultiplicatividad** (para normas de matrices): $||\mathbf{AB}|| \leq ||\mathbf{A}|| \cdot ||\mathbf{B}||$

En ML, las normas más utilizadas son las **normas vectoriales** (aplicadas a matrices) y las **normas matriciales específicas**.

- **Norma L_1 (Norma de Manhattan)** En ML se utiliza en la regularización L_1 (Lasso) para obtener modelos dispersos (sparse model).

  Para un vector $\mathbf{v} = (v_1, v_2, \ldots, v_n)$, la norma L_1 se define como:

  $$||\mathbf{v}||_1 = \sum_{i=1}^{n} |v_i|$$

  Para una matriz $\mathbf{A}$, la norma L_1 es la máxima suma absoluta de las columnas:

  $$||\mathbf{A}||_1 = max_{1 \leq j \leq n} \sum_{i=1}^{m} |a_{ij}|$$

2. Ecuaciones diferenciales

Una ecuación diferencial es una ecuación que relaciona una función con sus derivadas. En otras palabras, describe cómo cambia una cantidad en función de otra (por ejemplo, cómo cambia la concentración de una proteína en función del tiempo). La ecuación diferencial más básica es:

$$y'(t) = \frac{dy}{dt} = f(t, y)$$

Aquí, y es la función que queremos encontrar, t es la variable independiente (como el tiempo), y $f(t, y)$ es una función que describe cómo cambia y con respecto a t.

2.1. Ecuaciones diferenciales ordinarias (EDO)

Estas ecuaciones involucran una sola variable independiente, y la función desconocida depende solo de esa variable. **Ejemplo:** modelar el crecimiento de una población:

$$\frac{dP}{dt} = kP$$

donde $P(t)$ es la población en el tiempo t, y k es una constante de crecimiento.

2.2. Ecuaciones diferenciales parciales (EDP)

Las EDP involucran múltiples variables independientes, de forma que la función desconocida depende de varias variables. **Ejemplo:** modelar la difusión de una sustancia en un medio:

$$\frac{\delta u}{\delta t} = D\frac{\delta^2 u}{\delta x^2}$$

donde $u(x, t)$ es la concentración de la sustancia en la posición x y el tiempo t, y D es el coeficiente de difusión.

2.3. Orden de una ecuación diferencial

El orden de una ecuación diferencial es el orden de la derivada más alta que aparece en la ecuación.

- **Primer orden**: solo involucra la primera derivada

$$\frac{dy}{dt} = f(t, y)$$

- **Segundo orden**: involucra la segunda derivada

$$\frac{d^2y}{dt^2} + a\frac{dy}{dt} + by = 0$$

2.4. Solución de una ecuación diferencial

Resolver una ecuación diferencial significa encontrar la función $y(t)$ que satisface la ecuación. La **solución general** es una familia de soluciones que incluye constantes arbitrarias. La **solución particular** se obtiene al asignar valores específicos a las constantes, generalmente usando condiciones iniciales o de contorno.

Dependiendo del tipo de ecuación, existen diferentes métodos de resolver las ecuaciones diferenciales. Las ecuaciones diferenciales parciales (EDP) son más complejas y suelen requerir métodos avanzados como separación de variables, transformadas de Fourier o métodos numéricos. Para las ecuaciones diferenciales ordinarias (EDO), se pueden seguir los siguientes pasos:

1. **Separacion de variables**: se usa cuando la ecuación se puede separar en términos que dependen solo de y y solo de t.

$$\frac{dy}{dt} = ky \rightarrow \frac{dy}{y} = kdt$$

   Integrando ambos lados:

$$ln|y| = kt + C \rightarrow y(t) = Ce^{kt}$$

2. **Ecuaciones lineales de primer orden**: se resuelven usando un factor integrante.

$$\frac{dy}{dt} + P(t)y = Q(t)$$

3. **Ecuaciones de segundo orden con coeficientes constantes**: se resuelven asumiendo una solución de la forma $y(t) = e^{rt}$, donde r es una constante a determinar.

$$\frac{d^2y}{dt^2} + a\frac{dy}{dt} + by = 0$$