Inteligencia Artificial

Act8: Laboratorio de Álgebra Lineal

Arturo Garza Rodríguez

February 2025

1. Operaciones con matrices y determinantes

1.1. Matriz inversa

1.1.1. Obtener la matriz inversa

Obtener la matriz inversa de:

$$F = \begin{pmatrix} 1 & 2 & 3 \\ 0 & 1 & 4 \\ 5 & 6 & 0 \end{pmatrix}$$

Para obtener la matriz inversa de una matriz, usamos el método adjunto definido como:

$$A^{-1} = \frac{Adj(A)}{|A|} \tag{1}$$

Donde A^{-1} es la matriz inversa, $Adj(A)$ la matriz A adjunta de A y $|A|$ el determinante de A. Por simplicidad, renombramos la matriz F como A.
Determinante de A
Para obtener el determinante de la matriz de 3x3 podemos utilizar el método de la lluvia (Regla de Sarrus).

$$|A| = (1)(1)(0)+(2)(4)(5)+(3)(0)(6)-(3)(1)(5)-(2)(0)(0)-(1)(4)(6) = 0+40+0-15-0-24$$

$$|A| = 40 - 39 = 1$$

Adjunta y matriz de cofactores
Obtener la matriz adjunta solo es posible cuando la matriz a tartar es cuadrada, en este caso es de 3x3, por lo que sí es posible. El proceso siguiere aplicar la siguiente fórmula a cada elemento de A para obtener la matriz de cofactores, y luego transponer esta matriz resultante para obtener la adjunta de A:

$$a_{i,j} = (-1)^{i+j} \cdot |m_{i,j}|$$

Donde $a_{i,j}$ es el elemento de la matriz adjunta y $|m_{i,j}|$ es el determinante de la matriz menor complementaria. Para evitar cálculos extra, podemos definir una matriz de signos que irán alternando entre positivo y negativo por la paridad entre filas y columnas.

$$Signos = \begin{pmatrix} + & - & + \\ - & + & - \\ + & - & + \end{pmatrix}$$

Estos signos serán aplicados a los determinantes que encontremos asociados a cada término para obtener el valor de $a_{i,j}$. Entonces, definimos A^t como:

$$Cof(A) = \begin{pmatrix} +\begin{vmatrix} 1 & 4 \\ 6 & 0 \end{vmatrix} & -\begin{vmatrix} 0 & 4 \\ 5 & 0 \end{vmatrix} & +\begin{vmatrix} 0 & 1 \\ 5 & 6 \end{vmatrix} \\ -\begin{vmatrix} 2 & 3 \\ 6 & 0 \end{vmatrix} & +\begin{vmatrix} 1 & 3 \\ 5 & 0 \end{vmatrix} & -\begin{vmatrix} 1 & 2 \\ 5 & 6 \end{vmatrix} \\ +\begin{vmatrix} 2 & 3 \\ 1 & 4 \end{vmatrix} & -\begin{vmatrix} 1 & 3 \\ 0 & 4 \end{vmatrix} & +\begin{vmatrix} 1 & 2 \\ 0 & 1 \end{vmatrix} \end{pmatrix}$$

$$Cof(A) = \begin{pmatrix} -24 & 20 & -5 \\ 18 & -15 & 4 \\ 5 & -4 & 1 \end{pmatrix}$$

$$Adj(A) = Cof(A)^t = \begin{pmatrix} -24 & 18 & 5 \\ 20 & -15 & -4 \\ -5 & 4 & 1 \end{pmatrix}$$

Inversa de A

Para obtener la inversa de A, sustituimos los resultados en 1:

$$A^{-1} = \frac{\begin{pmatrix} -24 & 18 & 5 \\ 20 & -15 & -4 \\ -5 & 4 & 1 \end{pmatrix}}{1} = \begin{pmatrix} -24 & 18 & 5 \\ 20 & -15 & -4 \\ -5 & 4 & 1 \end{pmatrix}$$

Por lo tanto, la inversa de A antes F es:

$$F^{-1} = \begin{pmatrix} -24 & 18 & 5 \\ 20 & -15 & -4 \\ -5 & 4 & 1 \end{pmatrix}$$

1.1.2. Verificación de resultado

Para comprobar el que la matriz resultante es la inversa de F, podemos hacer uso de la siguiente propiedad:

$$A \cdot A^{-1} = I_n$$

Donde n viene de nxn, el tamaño de la matriz A. Se cumple que para toda matriz, la multiplicación de la matriz original y la inversa deber ser igual a la identidad de 'grado' (tamaño) n.

$$\rightarrow \begin{pmatrix} 1 & 2 & 3 \\ 0 & 1 & 4 \\ 5 & 6 & 0 \end{pmatrix} \cdot \begin{pmatrix} -24 & 18 & 5 \\ 20 & -15 & -4 \\ -5 & 4 & 1 \end{pmatrix}$$

$$= \begin{pmatrix} -24+40-15 & 18-30+12 & 5-8+3 \\ 0+20-20 & 0-15+16 & 0-4+4 \\ -120+120+0 & 90-90+0 & 25-24+0 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Se ha demostrado, que la matriz obtenida es la inversa de la original. $\blacksquare$

1.2. Demostración de propiedad de que el determinante de un producto de matrices es igual al producto de los determinantes

Para demostrar esto, haremos uso de las matrices elementales.

1.2.1. Descomposición de una matriz en matrices elementales

Sabemos que cualquier matriz cuadrada A de orden n no singular puede reescribirse como un producto de matrices elementales:

$$A = E_1 E_2 \dots\ E_k$$

donde cada E_i es una matriz elemental obtenida mediante oepraciones elementales sobre filas.
Análogamente, si B también es invertible, se puede escribir como:

$$B = F_1 F_2 \dots\ F_m$$

donde F_j son matrices elementales.

1.2.2. Determinantes de matrices elementales

Las matrices elementales tienen propiedades sencillas respecto al determinante:

1. Si E_i es una matriz de intercambio de filas, entonces $\det E_i = -1$.
2. Si E_i es una matriz de escalamiento de fila por un escalar λ, entonces $\det E_i = \lambda$.
3. Si E_i es una matriz de una fila multiplicada por un escalar a otra fila, entonces $\det E_i = 1$.

Para cualquier matriz elemental E_i:

$$\det(E_i A) = \det E_i \cdot \det A$$

1.2.3. Producto de matrices A y B

Dado que $A = E_1 E_2 \dots E_k$ y $B = F_1 F_2 \dots F_m$, su producto es:

$$AB = (E_1 E_2 \dots E_k)(F_1 F_2 \dots F_m)$$

Obtenemos el determinante de ambos lados:

$$\det(AB) = \det(E_1 E_2 \dots E_k F_1 F_2 \dots F_m)$$

Usando la propiedad mencionada anteriormente:

$$\det(AB) = \det(E_1)\det(E_2)\dots\det(E_k)\det(F_1)\det(F_2)\dots\det(F_m)$$

Pero:

$$\det(A) = \det(E_1)\det(E_2)\ldots\det(E_k)$$

y

$$\det(B) = \det(F_1)\det(F_2)\ldots\det(F_m)$$

Entonces:

$$\det(AB) = \det(A)\det(B)$$

Se ha demostrado que el determinante de un producto de matrices es igual al producto de detemrinantes ■

2. Sistemas de ecuaciones lineales

2.1. Método de Gauss-Seidel

Resuelva el siguiente sistema por el método de Gauss-Seidel:

$$\begin{cases} 4x - y + z = 7 \\ -2x + 4y - 2z = 1 \\ x - y + 3x = 5 \end{cases}$$

El método de Gauss-Seidel es un algoritmo iterativo para resolver sistemas de ecuaciones lineales $Ax = b$ que actualiza cada variable en función de los valores más recientes, acelerando la convergencia en matrices diagonales dominantes.

Paso 1

Despejamos las 3 variables, una de cada ecuación:

$$x = \frac{7 + y - z}{4}$$

$$y = \frac{1 + 2x + 2x}{4}$$

$$z = \frac{5 - x + y}{3}$$

Definimos valores iniciales para cada variable como: $x = 0, y = 0, z = 0$.

Proceso

El proceso consiste en sustituir los valores más recientes encontrados en las ecuaciones despejadas: Por cuestiones de presentación he redondeado cada resultado a

Iteración	x_1	x_2	x_3
0	0.0000	0.0000	0.0000
1	1.7500	1.1250	1.4583
2	1.6667	1.8125	1.7153
3	1.7743	1.9948	1.7402
4	1.8137	2.0269	1.7378
5	1.8223	2.0300	1.7359
6	1.8235	2.0297	1.7354
7	1.8236	2.0295	1.7353

los primeros 4 decimales. Se ha detenido el proceso en la iteración 7 puesto que utilicé una tolerancia de 0,0005, esto significa que entre las iteraciones 6 y 7 existe una diferencia menor a 0,0005 en todas las variables, por lo que se concluye que el valor obtenido es suficiente. Por lo tanto, los valores finales son:

$x = 1{,}82358$ $\quad$ $y = 2{,}02949$ $\quad$ $z = 1{,}73530$

2.2. Sistema homogéneo

Encuentre las soluciones al sistema homogéneo:

$$\begin{cases} x + 2y + 3z = 0 \\ 2x + 4y + 6z = 0 \\ 3x + 6y + 9z = 0 \end{cases}$$

Si nos ponemos a analizar el sistema, nos encontramos con que en realidad todas las ecuaciones son un múltiplo de la primera, la ecuación2 es 2*ecuación1, mientras que la ecuación3 = 3*ecuación1. Por lo que podemos afirmar que no existe una única solución, sino infinitas.

Como simplificamos todo a un solo vector, entonces, para obtener una solución factible para el problema, es necesario definir dos valores, y en base a ellos se definirá la otra variable, veamos:

$$x + 2y + 3z = 0$$

$$x = -2y - 3z$$

Si $y = 2, z = 1 \rightarrow x = -2(2) - 3(1) = -7$.

Sustituimos los valores en la ecuacion:

$$(-7) + (2)(2) + (3)(1) = -7 + 4 + 3 = 0.$$

Por lo tanto, si queremos obtener una solución factible, solo tenemos que definir dos variables (y, z) y obtener x dado:

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = y \begin{bmatrix} -2 \\ 1 \\ 0 \end{bmatrix} + z \begin{bmatrix} -3 \\ 0 \\ 1 \end{bmatrix}, \quad \text{con } y, z \in \mathbb{R}.$$

3. Espacios vectoriales y auto-valores/auto-vectores

3.1. Bases y dimensión

Encuentre la base y la dimensión del subespacio generado por los vectores (1,2,3),(2,4,6),(3,6,9).

3.1.1. Notación

Para encontrar la base y dimensión de un conjunto de vectores, escribimos estos como una matriz, obteniendo:

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 4 & 6 \\ 3 & 6 & 9 \end{pmatrix}$$

Y hacemos las operaciones necesarias para llevar la matriz a su forma escalonada. Como son vectores dependientes, entonces:

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

3.1.2. Dimensión

La matriz escalonada tiene una única fila **no nula**, lo que indica que el rango de la matriz es 1 ($\rho(A) = 1$), esto significa que la dimensión del subespacio generado (o dimensión de la imagen de A) es 1.

3.1.3. Base

La Ãºnica fila no nula en la matriz escalonada es (1,2,3), lo que indica que este vector es suficiente para generar el subespacio.

Por lo tanto, la base del subespacio es:

$$B(A) = \left\{ \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} \right\}$$

3.2. Autovalores y autovectores

Determine los autovalores y autovectores de la matriz:

$$G = \begin{bmatrix} 5 & -2 \\ -2 & 5 \end{bmatrix}$$

Definición Sea A una matriz de nxn con componentes reales, el número λ (real o complejo) se llama eigenvalor (autovalor) de A si existe un vector diferente de 0 $v \in \mathbb{C}^n$ tal que:

$$A \cdot v = \lambda v$$

Como $A \cdot v = \lambda v$, entonces:

$$\rightarrow A \cdot v - \lambda I v = 0$$

$$\rightarrow (A - \lambda I) \cdot v = 0$$

con v como e núcleo de $A - \lambda I$.

Teorema Sea A una matriz de dimensión nxn, entonces λ es un valor propio de A si y solo si:

$$p(\lambda) = |A - \lambda I| = 0$$

$$\rightarrow \det(A - \lambda I) = 0$$

3.2.1. Obtención de autovalores

Sea $G = \begin{bmatrix} 5 & -2 \\ -2 & 5 \end{bmatrix}$ y $\lambda I = \begin{bmatrix} \lambda & 0 \\ 0 & \lambda \end{bmatrix}$, entonces: $G - \lambda I = \begin{bmatrix} 5-\lambda & -2 \\ -2 & 5-\lambda \end{bmatrix}$

Aplicando el determinante a la matriz resultante:

$$\det(G - \lambda I) = \begin{vmatrix} 5-\lambda & -2 \\ -2 & 5-\lambda \end{vmatrix} = (5-\lambda)^2 - (-2)^2 = 0$$

$$\rightarrow (5-\lambda)^2 - (-2)^2 = 0 = (25 - 10\lambda + \lambda^2) - (4) = 0$$

$$\rightarrow \lambda^2 - 10\lambda + 21 = 0$$

Factorizando:

$$\rightarrow (\lambda - 3)(\lambda - 7) = 0$$

Resolviendo: $\lambda = 3$ $\lambda = 7$

3.2.2. Obtención de autovectores

Para $\lambda = 3$

$$\begin{bmatrix} 5-3 & -2 \\ -2 & 5-3 \end{bmatrix} = \begin{bmatrix} 2 & -2 \\ -2 & 2 \end{bmatrix}$$

Multiplicamos la matriz resultante por (x,y) e igualamos a 0:

$$\begin{bmatrix} 2 & -2 \\ -2 & 2 \end{bmatrix} \cdot \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

Y resolviendo el sistema de ecuaciones, encontramos que el eigenvector asociado a $\lambda = 3$ es:

$$v_1 = k_1 \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

Para $\lambda = 7$

$$\begin{bmatrix} 5-7 & -2 \\ -2 & 5-7 \end{bmatrix} = \begin{bmatrix} -2 & -2 \\ -2 & -2 \end{bmatrix}$$

Multiplicamos la matriz resultante por (x,y) e igualamos a 0:

$$\begin{bmatrix} -2 & -2 \\ -2 & -2 \end{bmatrix} \cdot \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

Y resolviendo el sistema de ecuaciones, encontramos que el eigenvector asociado a $\lambda = 7$ es:

$$v_2 = k_2 \begin{bmatrix} -1 \\ 1 \end{bmatrix}$$

4. Aplicaciones de la IA: reducción de dimensionalidad

4.1. Análisis de Componentes Principales (PCA)

Reduce la dimensionalidad de los datos usando álgebra lineal, maximizando la varianza retenida.

4.1.1. Centralización

Se resta la media de cada variable para centrar los datos en el origen.

4.1.2. Matriz de covarianza

Se calcula como:

$$C = \frac{1}{n-1} X^T X$$

para analizar la relación entre variables.

4.1.3. Autovalores y autovectores

Se resuelve:

$$Cv = \lambda v$$

- **Autovalores**: indican la varianza explicada.
- **Autovectores**: definen las direcciones principales.

4.1.4. Proyección

Se eligen los k autovectores con mayores autovalores y se proyectan los datos:

$$X' = XW$$

reduciendo la dimensión sin perder mucha información.

4.2. Descomposición en valores singulares (SVD)

Calcule la descomposición en valores singulares de la matriz:

$$H = \begin{pmatrix} 3 & 1 \\ 2 & 2 \end{pmatrix}$$

4.2.1. Calcular $H^T H$ y sus valores propios

$$H = \begin{pmatrix} 3 & 1 \\ 2 & 2 \end{pmatrix} \qquad H^T = \begin{pmatrix} 3 & 2 \\ 1 & 2 \end{pmatrix}$$

$$H^T H = \begin{pmatrix} 3 & 2 \\ 1 & 2 \end{pmatrix} \cdot \begin{pmatrix} 3 & 1 \\ 2 & 2 \end{pmatrix} = \begin{pmatrix} 9+4 & 3+4 \\ 3+4 & 1+4 \end{pmatrix} = \begin{pmatrix} 13 & 7 \\ 7 & 5 \end{pmatrix}$$

4.2.2. Valores y vectores propios de H^TH

$$\begin{vmatrix} 13-\lambda & 7 \\ 7 & 5-\lambda \end{vmatrix} = (13-\lambda)(5-\lambda) - 7^2 = 65 - 18\lambda + \lambda^2 - 49$$
$$= \lambda^2 - 18\lambda + 16 = 0$$

Resolviendo con la fórmula general:

$$\lambda = \frac{18 \pm \sqrt{(-18)^2 - 4(1)(16)}}{2(1)} = \frac{18 \pm \sqrt{260}}{2} = 9 \pm \sqrt{65}$$

$\rightarrow$ $\lambda_1 = 17{,}062$ $\lambda_2 = 0{,}9377$

Vectores propios

Para $\lambda_1 = 17{,}062$:

$$\begin{bmatrix} -4{,}062 & 7 \\ 7 & -12{,}062 \end{bmatrix} \cdot \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

Obteniendo:

$$v_1 = k_1 \begin{bmatrix} \frac{1}{7}(4+\sqrt{65}) \\ 1 \end{bmatrix}$$

Para $\lambda_2 = 0{,}9377$:

$$\begin{bmatrix} 12{,}0623 & 7 \\ 7 & 4{,}0623 \end{bmatrix} \cdot \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

Obteniendo:

$$v_2 = k_2 \begin{bmatrix} \frac{1}{7}(4-\sqrt{65}) \\ 1 \end{bmatrix}$$

4.2.3. Obtencióndе U y V

Partimos de:

$$H = U\Sigma V^T$$

donde U y V son matrices ortogonales y Σ es una matriz diagonal con los valores singulares.

Los vectores propios v_1 y v_2 forman las columnas de la matriz V:

$$V = \begin{pmatrix} \frac{1}{7}(12{,}062) & \frac{1}{7}(-4{,}062) \\ 1 & 1 \end{pmatrix}$$

La matriz Σ está definida por los valores singulares (raíz cuadrada de los valores propios) en la diagonal:

$$\Sigma = \begin{pmatrix} \sqrt{17{,}062} & 0 \\ 0 & \sqrt{0{,}9377} \end{pmatrix} \approx \begin{pmatrix} 4{,}13 & 0 \\ 0 & 0{,}97 \end{pmatrix}$$

La matriz U se obtiene usando la relación $HV = U\Sigma \to HV\Sigma^T = U$, se definen las columnas de U como:

$$u_1 = \frac{1}{\sigma_1} H v_1 = \frac{1}{4{,}13} \begin{pmatrix} 3 & 1 \\ 2 & 2 \end{pmatrix} \begin{pmatrix} \frac{1}{7}(12{,}062) \\ 1 \end{pmatrix}$$

$$u_2 = \frac{1}{\sigma_2} H v_2 = \frac{1}{0{,}97} \begin{pmatrix} 3 & 1 \\ 2 & 2 \end{pmatrix} \begin{pmatrix} \frac{1}{7}(-4{,}062) \\ 1 \end{pmatrix}$$

Encontramos que:

$$U = \begin{pmatrix} \frac{\frac{3}{7}(12{,}062)+1}{4{,}13} & \frac{\frac{3}{7}(-4{,}062)+1}{0{,}97} \\ \frac{\frac{2}{7}(12{,}062)+2}{4{,}13} & \frac{\frac{2}{7}(-4{,}062)+2}{0{,}97} \end{pmatrix}$$

Descomposición

Por último, la descomposición en valores singulares queda de la siguiente forma:

$$H = U\Sigma V^T =$$

$$\begin{pmatrix} \frac{\frac{3}{7}(12{,}062)+1}{4{,}13} & \frac{\frac{3}{7}(-4{,}062)+1}{0{,}97} \\ \frac{\frac{2}{7}(12{,}062)+2}{4{,}13} & \frac{\frac{2}{7}(-4{,}062)+2}{0{,}97} \end{pmatrix} \cdot \begin{pmatrix} \sqrt{9+\sqrt{65}} & 0 \\ 0 & \sqrt{9-\sqrt{65}} \end{pmatrix} \cdot \begin{pmatrix} \frac{1}{7}(4+\sqrt{65}) & 1 \\ \frac{1}{7}(4-\sqrt{65}) & 1 \end{pmatrix}$$

4.3. Álgebra Líneal en aprendizaje profundo

4.3.1. Representación de Datos

Los datos de entrada se representan como vectores o matrices. Una imagen, por ejemplo, se almacena como una matriz de píxeles, y un conjunto de datos completo se modela como una matriz donde cada fila es una muestra y cada columna una característica.

4.3.2. Pesos y Operaciones Lineales

Las conexiones entre neuronas tienen pesos que se almacenan en matrices. La propagación de la información en la red se realiza mediante productos de matrices:

$$Z = WX + b$$

donde W es la matriz de pesos, X la entrada y b el sesgo.

4.3.3. Funciones de Activación

Las funciones de activación transforman la salida de cada neurona. Aunque algunas son no lineales (ReLU, sigmoide, tanh), sus derivadas se utilizan en el cálculo del gradiente.

4.3.4. Propagación Hacia Atrás y Gradientes

El entrenamiento de redes neuronales usa el algoritmo de retropropagación, basado en el cálculo del gradiente mediante la regla de la cadena. Se utilizan derivadas parciales y el producto de matrices para actualizar los pesos:

$$W^{(t+1)} = W^{(t)} - \alpha \nabla L$$

donde α es la tasa de aprendizaje y ∇L el gradiente de la función de pérdida.

4.3.5. Descomposición en Valores Singulares (SVD)

SVD permite reducir la dimensionalidad de los datos y comprimir modelos. Se descompone una matriz A en:

$$A = U \Sigma V^T$$

donde U y V son ortogonales, y Σ contiene los valores singulares.

4.4. Impacto de espacios vectoriales en representación de datos en IA

4.4.1. Representación de Datos

En IA, los datos se representan como vectores en espacios multidimensionales. Palabras, imágenes y señales se transforman en vectores numéricos para facilitar su procesamiento.

4.4.2. Distancias y Similitud

El álgebra lineal permite medir similitudes entre datos mediante distancias euclidianas y coseno del ángulo entre vectores, fundamentales en algoritmos de clasificación y clustering.

4.4.3. Reducción de Dimensionalidad

Técnicas como PCA proyectan datos a subespacios de menor dimensión, eliminando redundancias sin perder información clave, optimizando modelos de IA.

4.4.4. Espacios Latentes

Modelos como autoencoders y Word2Vec aprenden representaciones compactas de datos en espacios latentes, permitiendo tareas como generación de texto y compresión de imágenes.