Laboratorio de Repaso Álgebra Lineal

Martín Alexis Martínez Andrade - 2049334

4.1 Operaciones con matrices y determinantes

Problema 1. Inversa de la matriz F

Teniendo la matriz

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

Usando eliminación de Gauss–Jordan:

$$\left[\begin{array}{ccc|ccc} 1 & 2 & 3 & 1 & 0 & 0 \\ 0 & 1 & 4 & 0 & 1 & 0 \\ 5 & 6 & 0 & 0 & 0 & 1 \end{array}\right].$$

$$\left[\begin{array}{ccc|ccc} 1 & 2 & 3 & 1 & 0 & 0 \\ 0 & 1 & 4 & 0 & 1 & 0 \\ 0 & -4 & -15 & -5 & 0 & 1 \end{array}\right].$$

$$\left[\begin{array}{ccc|ccc} 1 & 0 & -5 & 1 & -2 & 0 \\ 0 & 1 & 4 & 0 & 1 & 0 \\ 0 & 0 & 1 & -5 & 4 & 1 \end{array}\right].$$

$$\left[\begin{array}{ccc|ccc} 1 & 0 & 0 & -24 & 18 & 5 \\ 0 & 1 & 0 & 20 & -15 & -4 \\ 0 & 0 & 1 & -5 & 4 & 1 \end{array}\right].$$

La parte derecha de la matriz es la inversa de F:

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

Problema 2. Propiedad del determinante: $\det(AB) = \det(A)\det(B)$

Si tenemos dos matrices

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \quad \text{y} \quad B = \begin{pmatrix} e & f \\ g & h \end{pmatrix}.$$

El producto de las matrices es

$$AB = A \cdot B = \begin{pmatrix} ae + bg & af + bh \\ ce + dg & cf + dh \end{pmatrix}.$$

Calculando el determinante de esta matriz AB tenemos

$$\det(AB) = (ae + bg)(cf + dh) - (af + bh)(ce + dg).$$

Que al expandir se obtiene

$$\begin{aligned}
\det(AB) &= \Big[afce + aedh + bfgc + bghd\Big] - \Big[acfe + adfg + bche + bdhg\Big] \\
&= aedh + bgcf - afdg - bhce \\
&= ad(eh) + bc(fg) - ad(fg) - bc(eh) \\
&= ad(eh - fg) - bc(eh - fg) \\
&= (ad - bc)(eh - fg)
\end{aligned}$$

Y dado que

$$\det(A) = ad - bc \quad \text{y} \quad \det(B) = eh - fg$$

se concluye que:

$$\det(AB) = \det(A)\det(B) = (ad - bc)(eh - fg)$$

Y por lo tanto:

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

4.2 Sistemas de ecuaciones lineales

Problema 3. Método de Gauss-Seidel

Tenemos el sistema de ecuaciones:

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

Con Gauss-Seidel se tiene que expresar cada ecuación en forma de actualización para cada variable:

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

Partimos de $x = y = z = 0$, y se itera hasta que la diferencia entre iteraciones consecutivas sea menor que una tolerancia pequeña (1%).

En cada iteración, se obtienen los siguientes resultados:

Iteración 1

Variable	Valor	Error
x	1.75	1
y	1.125	1
z	1.458	1

Iteración 2

Variable	Valor	Error
x	1.666	0.049
y	1.812	0.379
z	1.715	0.149

Iteración 3

Variable	Valor	Error
x	1.774	0.060
y	1.994	0.091
z	1.740	0.014

Iteración 4

Variable	Valor	Error
x	1.813	0.021
y	2.026	0.015
z	1.737	0.0013

Iteración 5

Variable	Valor	Error
x	1.822	0.0047
y	2.030	0.0015
z	1.735	0.001

Dando como solución: $x = 1.822, y = 2.03, z = 1.735$

Problema 4. Solución del sistema homogéneo

Tenemos el sistema:

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

Las ecuaciones son linealmente dependientes:

$2x + 4y + 6z = 2(x + 2y + 3z)$

$3x + 6y + 9z = 3(x + 2y + 3z)$

Así que el sistema tiene una única ecuación independiente:

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

Podemos despejar x en términos de y y z, y queda:

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

Usando como parámetros $y = b$ y $z = c$, la solución general es:

$$(x, y, z) = (-2b - 3c,\ b,\ c)$$

Tal que b y c son números reales.

4.3 Espacios vectoriales y autovalores/auto-vectores

Problema 5. Base y dimensión del subespacio generado por $\{(1,2,3),(2,4,6),(3,6,9)\}$

Analizando los vectores, se puede ver que

$$(2,4,6) = 2\,(1,2,3) \quad \text{y} \quad (3,6,9) = 3\,(1,2,3).$$

Así que todos los vectores son múltiplos de $(1,2,3)$, y entonces el subespacio es

$$\{\, s(1,2,3) : s \in \mathbb{R}\}.$$

Una base para el subespacio es

$$\mathbf{\{(1,2,3)\}}$$

y su **dimensión es 1**.

Problema 6. Autovalores y autovectores de la matriz

Tenemos la matriz

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

Buscamos la solución para la ecuación

$$\det(G - \lambda I) = 0.$$

Calculamos

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

Igualamos su determinante a cero

$$\det(G - \lambda I) = (5-\lambda)^2 - (-2)^2 = (5-\lambda)^2 - 4 = 0.$$

$$(5-\lambda)^2 = 4 \quad \Longrightarrow \quad 5 - \lambda = \pm 2.$$

$$\lambda_1 = 5 - 2 = 3 \quad \text{y} \quad \lambda_2 = 5 + 2 = 7.$$

Cálculo de autovectores:

Para $\lambda = 3$:

$$(G - \lambda_1 I) = (G - 3I) = \begin{bmatrix} 2 & -2 \\ -2 & 2 \end{bmatrix}.$$

Se utiliza la ecuación $(G - 3I)\cdot v = 0$:

$$\begin{cases} 2x - 2y = 0, \\ -2x + 2y = 0. \end{cases}$$

$$2x - 2y = 0 \quad \rightarrow \quad x = y.$$

Un autovector es:

$$v_1 = (1,1)$$

Para $\lambda = 7$:
$$(G - 7I) = \begin{bmatrix} -2 & -2 \\ -2 & -2 \end{bmatrix}.$$
La ecuación $(G - 7I) \cdot v = 0$ da:
$$-2x - 2y = 0 \quad \rightarrow \quad x + y = 0 \quad \rightarrow \quad y = -x.$$
Un autovector es:
$$v_2 = (1, -1).$$
Así que los autovalores y autovectores son:
$$\lambda_1 = 3,\ v_1 = (1, 1) \quad \text{y} \quad \lambda_2 = 7,\ v_2 = (1, -1)$$

4.4 Aplicaciones en IA: reducción de dimensionalidad

7. Uso de PCA en reducción de dimensiones

El Análisis de Componentes Principales es una técnica estadística que usa álgebra lineal para reducir la dimensión de un conjunto de datos, conservando gran parte de su información.

El proceso de PCA es:

1. Estandarizar el rango de variables iniciales continuas: Se resta la media y se divide entre la desviación estándar
2. Calcular la matriz de covarianza para identificar las correlaciones
3. Calcular los autovectores y autovalores de la matriz de covarianza
4. Seleccionar los componentes principales
5. Transformar los datos en el nuevo sistema de coordenadas

Problema 8. Descomposición en valores singulares (SVD)

Tenemos la matriz
$$H = \begin{bmatrix} 3 & 1 \\ 2 & 2 \end{bmatrix}.$$

Valores Singulares

$$H^T = \begin{bmatrix} 3 & 2 \\ 1 & 2 \end{bmatrix} \rightarrow \quad H^T H = \begin{bmatrix} 3 & 2 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} 3 & 1 \\ 2 & 2 \end{bmatrix} = \begin{bmatrix} 9+4 & 3+4 \\ 3+4 & 1+4 \end{bmatrix} = \begin{bmatrix} 13 & 7 \\ 7 & 5 \end{bmatrix}.$$
Obteniendo los autovalores:
$$\det \begin{bmatrix} 13-\lambda & 7 \\ 7 & 5-\lambda \end{bmatrix} = (13-\lambda)(5-\lambda) - 49 = \lambda^2 - 18\lambda + (65 - 49) = \lambda^2 - 18\lambda + 16 = 0.$$
$$\lambda = \frac{18 \pm \sqrt{18^2 - 4 \cdot 16}}{2} = \frac{18 \pm \sqrt{324 - 64}}{2} = \frac{18 \pm \sqrt{260}}{2} = 9 \pm \sqrt{65}.$$
Los valores singulares son:
$$\sigma_1 = \sqrt{9 + \sqrt{65}}, \quad \sigma_2 = \sqrt{9 - \sqrt{65}}.$$

Matriz V
Buscamos los autovectores de $H^\top H$.

Para $\lambda_1 = 9 + \sqrt{65}$:

$$(13 - (9 + \sqrt{65}))x + 7y = (4 - \sqrt{65})x + 7y = 0,$$
$$y = \frac{\sqrt{65} - 4}{7}\, x.$$

Se puede elegir $x = 7$

$$v_1 = \begin{pmatrix} 7 \\ \sqrt{65} - 4 \end{pmatrix}.$$

Su norma es

$$\|v_1\| = \sqrt{7^2 + (\sqrt{65} - 4)^2} = \sqrt{130 - 8\sqrt{65}},$$

Y ahora se puede normalizar para que V sea ortogonal,

$$v_1 = \frac{1}{\sqrt{130 - 8\sqrt{65}}} \begin{pmatrix} 7 \\ \sqrt{65} - 4 \end{pmatrix}.$$

Para $\lambda_2 = 9 - \sqrt{65}$:

$$(13 - (9 - \sqrt{65}))x + 7y = (4 + \sqrt{65})x + 7y = 0,$$
$$y = -\frac{4 + \sqrt{65}}{7}\, x.$$

Eligiendo $x = 7$, se tiene

$$v_2 = \begin{pmatrix} 7 \\ -\left(4 + \sqrt{65}\right) \end{pmatrix},$$

$$\|v_2\| = \sqrt{7^2 + (-4 - \sqrt{65})^2} = \sqrt{130 + 8\sqrt{65}},$$

Y se normaliza igual que con v_1:

$$v_2 = \frac{1}{\sqrt{130 + 8\sqrt{65}}} \begin{pmatrix} 7 \\ -\left(4 + \sqrt{65}\right) \end{pmatrix}.$$

Y resulta en:

$$V = \begin{pmatrix} \dfrac{7}{\sqrt{130 - 8\sqrt{65}}} & \dfrac{7}{\sqrt{130 + 8\sqrt{65}}} \\ \dfrac{\sqrt{65} - 4}{\sqrt{130 - 8\sqrt{65}}} & -\dfrac{4 + \sqrt{65}}{\sqrt{130 + 8\sqrt{65}}} \end{pmatrix}.$$

Matriz U
Se puede encontrar U con los autovectores de

$$HH^\top = \begin{pmatrix} 10 & 8 \\ 8 & 8 \end{pmatrix}.$$

Para $\lambda_1 = 9 + \sqrt{65}$:

$$\Big(10 - (9 + \sqrt{65})\Big)x + 8y = (1 - \sqrt{65})x + 8y = 0,$$

$$y = \frac{\sqrt{65} - 1}{8}x$$

Se elige $x = 8$:

$$u_1 = \begin{pmatrix} 8 \\ \sqrt{65} - 1 \end{pmatrix},$$

$$\|u_1\| = \sqrt{8^2 + (\sqrt{65} - 1)^2} = \sqrt{130 - 2\sqrt{65}}.$$

Se normaliza y se obtiene u_1

$$u_1 = \frac{1}{\sqrt{130 - 2\sqrt{65}}} \begin{pmatrix} 8 \\ \sqrt{65} - 1 \end{pmatrix}.$$

Para $\lambda_2 = 9 - \sqrt{65}$:

$$\Big(10 - (9 - \sqrt{65})\Big)x + 8y = (1 + \sqrt{65})x + 8y = 0,$$

$$y = -\frac{1 + \sqrt{65}}{8}x$$

Eligiendo $x = 8$:

$$u_2 = \begin{pmatrix} 8 \\ -\Big(1 + \sqrt{65}\Big) \end{pmatrix},$$

$$\|u_2\| = \sqrt{8^2 + (1 + \sqrt{65})^2} = \sqrt{130 + 2\sqrt{65}},$$

y se normaliza

$$u_2 = \frac{1}{\sqrt{130 + 2\sqrt{65}}} \begin{pmatrix} 8 \\ -\Big(1 + \sqrt{65}\Big) \end{pmatrix}.$$

La matriz U entonces es

$$U = \begin{pmatrix} \dfrac{8}{\sqrt{130 - 2\sqrt{65}}} & \dfrac{8}{\sqrt{130 + 2\sqrt{65}}} \\ \dfrac{\sqrt{65} - 1}{\sqrt{130 - 2\sqrt{65}}} & -\dfrac{1 + \sqrt{65}}{\sqrt{130 + 2\sqrt{65}}} \end{pmatrix}.$$

Así que entonces la descomposición en valores singulares de H

$$H = U\,\Sigma\,V^\top,$$

es:

$$\Sigma = \begin{pmatrix} \sqrt{9 + \sqrt{65}} & 0 \\ 0 & \sqrt{9 - \sqrt{65}} \end{pmatrix},$$

$$U = \begin{pmatrix} \dfrac{8}{\sqrt{130-2\sqrt{65}}} & \dfrac{8}{\sqrt{130+2\sqrt{65}}} \\ \dfrac{\sqrt{65}-1}{\sqrt{130-2\sqrt{65}}} & -\dfrac{1+\sqrt{65}}{\sqrt{130+2\sqrt{65}}} \end{pmatrix}.$$

$$V = \begin{pmatrix} \dfrac{7}{\sqrt{130-8\sqrt{65}}} & \dfrac{7}{\sqrt{130+8\sqrt{65}}} \\ \dfrac{\sqrt{65}-4}{\sqrt{130-8\sqrt{65}}} & -\dfrac{4+\sqrt{65}}{\sqrt{130+8\sqrt{65}}} \end{pmatrix}.$$

9. Álgebra lineal en el aprendizaje profundo

El Machine Learning es básicamente agentes saltando a diferentes direcciones utilizando fórmulas e inputs en unidades ("neuronas"). Todos los modelos de ML se fundamentan en conceptos de álgebra lineal: optimizaciones utilizando operaciones con matrices (utilizar NumPy en lugar de loops con operaciones escalares en Python, por ejemplo, optimiza mucho el código y el runtime), cálculo vectorial, representación de los datos en vectores y matrices. También, por ejemplo, el forward propagation utiliza multiplicaciones de matrices y sumas de vectores; el backpropagation utiliza operaciones de matrices para obtener derivadas parciales y calcular gradientes; las funciones de costo y el descenso del gradiente se utilizan con matrices para optimizar tiempos y recursos computacionales.

10. Impacto de los espacios vectoriales en la representación de datos en IA

El concepto de espacio vectorial da una forma matemática para representar datos de forma numérica. Los datos (imágenes, audio, texto, video, etc.) se representan como vectores en espacios de muchas dimensiones. Y gracias a que se está manejando un espacio vectorial, se tiene norma y producto interno, lo que se usa para establecer distancia y similitud entre vectores (entre datos), que muchos algoritmos como clustering utilizan. Los espacios vectoriales también permiten aplicar transformaciones lineales como el Análisis de Componentes Principales o embeddings.