

Laboratorio de Álgebra Lineal

Aldo Hernández

Febrero 2025

1 Operaciones con matrices y determinantes

1. Consideremos la siguiente matriz F

$$\begin{aligned} F &= \left(\begin{array}{ccc|ccc} 1 & 2 & 3 & 1 & 0 & 0 \\ 0 & 1 & 4 & 0 & 1 & 0 \\ 5 & 6 & 0 & 0 & 0 & 1 \end{array} \right) \sim \left(\begin{array}{ccc|ccc} 1 & 2 & 3 & 1 & 0 & 0 \\ 5 & 6 & 0 & 0 & 0 & 1 \\ 0 & 1 & 4 & 0 & 1 & 0 \end{array} \right) \\ &\sim \left(\begin{array}{ccc|ccc} 1 & 2 & 3 & 1 & 0 & 0 \\ 0 & -4 & -15 & -5 & 0 & 1 \\ 0 & 1 & 4 & 0 & 1 & 0 \end{array} \right) \sim \left(\begin{array}{ccc|ccc} 1 & 2 & 3 & 1 & 0 & 0 \\ 0 & 1 & \frac{15}{4} & \frac{5}{4} & 0 & \frac{-1}{4} \\ 0 & 1 & 4 & 0 & 1 & 0 \end{array} \right) \\ &\sim \left(\begin{array}{ccc|ccc} 1 & 0 & \frac{-9}{2} & \frac{-3}{2} & 0 & \frac{1}{2} \\ 0 & 1 & \frac{15}{4} & \frac{5}{4} & 0 & \frac{-1}{4} \\ 0 & 0 & \frac{1}{4} & \frac{-5}{4} & 1 & \frac{1}{4} \end{array} \right) \sim \left(\begin{array}{ccc|ccc} 1 & 0 & \frac{-9}{2} & \frac{-3}{2} & 0 & \frac{1}{2} \\ 0 & 1 & \frac{15}{4} & \frac{5}{4} & 0 & \frac{-1}{4} \\ 0 & 0 & 1 & -5 & 4 & 1 \end{array} \right) \\ &\sim \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & -24 & 18 & 5 \\ 0 & 1 & 0 & 20 & -15 & -4 \\ 0 & 0 & 1 & -5 & 4 & 1 \end{array} \right) \\ \Rightarrow F^{-1} &= \begin{pmatrix} -24 & 18 & 5 \\ 20 & -15 & -4 \\ -5 & 4 & 1 \end{pmatrix} \end{aligned}$$

2. Sean A y B dos matrices de orden n no necesariamente iguales, entonces si alguna de ellas no es invertible (no tiene inversa) podemos afirmar que su determinante es cero

$$\Rightarrow \det AB = 0$$

Además se cumpliría la proposición

$$\det A \cdot \det B = 0$$

En caso de que ambas matrices sean invertibles, entonces tienen inversa y pueden escribirse como productos de matrices elementales (matrices obtenidas al realizar operaciones elementales sobre una matriz identidad)

$$A = E_1 \cdot E_2 \cdots E_n$$

$$B = F_1 \cdot F_2 \cdots F_m$$

$$\begin{aligned}\Rightarrow \det AB &= \det(E_1 \cdot E_2 \cdots E_n \cdots F_1 \cdot F_2 \cdots F_m) \\ &= \det(E_1) \cdot \det(E_2) \cdots \det(E_n) \cdot \det(F_1) \cdot \det(F_2) \cdots \det(F_m) \\ &= \det(E_1 \cdot E_2 \cdots E_n) \cdot \det(F_1 \cdot F_2 \cdots F_m) \\ &= \det(A) \cdot \det(B) \blacksquare\end{aligned}$$

2 Sistemas de ecuaciones lineales

1. Despejamos una variable de cada ecuación del sistema, de manera que tenemos

$$x = \frac{7 + y - z}{4}, y = \frac{1 + 2x + 2z}{4}, z = \frac{5 - x + y}{3}$$

Tomamos como solución inicial $x = y = z = 0$

- Iteración 1

$$x = 1.75$$

$$y = 1.13$$

$$z = 1.46$$

- Iteración 2

$$x = \frac{7 + 1.13 - 1.46}{4} = 1.67$$

$$y = \frac{1 + 3.34 + 2.92}{4} = 1.82$$

$$z = \frac{5 - 1.67 + 1.82}{3} = 1.72$$

- Iteración 3

$$x = \frac{7 + 1.82 - 1.72}{4} = 1.78$$

$$y = \frac{1 + 3.56 + 3.44}{4} = 2$$

$$z = \frac{5 - 1.78 + 2}{3} = 1.74$$

- Iteración 4

$$x = \frac{7 + 2 - 1.74}{4} = 1.82$$

$$y = \frac{1 + 3.64 + 3.48}{4} = 2.03$$

$$z = \frac{5 - 1.82 + 2.03}{3} = 1.74$$

- Iteración 5

$$x = \frac{7 + 2.03 - 1.74}{4} = 1.82$$

$$y = \frac{1 + 3.64 + 3.48}{4} = 2.03$$

$$z = \frac{5 - 1.82 + 2.03}{3} = 1.74$$

Observamos que llegamos a un error aproximado cercano al 0%, entonces concluimos

como resultado lo siguiente

$$x = 1.82$$

$$y = 2.03$$

$$z = 1.74$$

2. Para el sistema homogéneo, podemos definir las siguientes matrices

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 4 & 6 \\ 3 & 6 & 9 \end{pmatrix}, X = \begin{pmatrix} x \\ y \\ z \end{pmatrix}, B = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

De forma que $CX = B$, pero observamos que todos los renglones de C son múltiplos, es decir, son linealmente dependientes, por lo que al realizar operaciones elementales con A, eventualmente llegaremos al siguiente resultado

$$\begin{aligned} A &\sim \begin{pmatrix} 1 & 2 & 3 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \\ \Rightarrow x &= -2y - 3z \\ \Rightarrow X &= \begin{pmatrix} -2y - 3z \\ y \\ z \end{pmatrix} \\ &= y \begin{pmatrix} -2 \\ 1 \\ 0 \end{pmatrix} + z \begin{pmatrix} -3 \\ 0 \\ 1 \end{pmatrix} \end{aligned}$$

Entonces decimos que los vectores constantes $\begin{pmatrix} -2 \\ 1 \\ 0 \end{pmatrix}$ y $\begin{pmatrix} -3 \\ 0 \\ 1 \end{pmatrix}$ son soluciones básicas del sistema de ecuaciones y que el sistema es compatible indeterminado con dos variables libres. Es por esto que hay soluciones infinitas dadas por el siguiente conjunto solución

$$\{x = -2y - 3z, y, z | x, y, z \in \mathbb{R}\}$$

3 Espacios vectoriales y auto-valores/auto-vectores

1. Vemos que los vectores del subespacio dado son linealmente dependientes

$$\left\{ \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}, \begin{pmatrix} 2 \\ 4 \\ 6 \end{pmatrix}, \begin{pmatrix} 3 \\ 6 \\ 9 \end{pmatrix} \right\} = \left\{ \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}, 2 \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}, 3 \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} \right\}$$

Entonces podemos concluir que una base para el subespacio vectorial es

$$B = \left\{ \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} \right\}, \dim B = 1$$

2. Para hallar los eigenvalores de G tenemos que igualar el siguiente determinante a cero

$$\begin{aligned} |G - \lambda I| &= \begin{vmatrix} 5 - \lambda & -2 \\ -2 & 5 - \lambda \end{vmatrix} \\ &= (5 - \lambda)^2 - (-2)^2 \\ \Rightarrow 0 &= (5 - \lambda)^2 - (-2)^2 \\ 4 &= (5 - \lambda)^2 \\ \therefore \lambda_1 &= 3 \\ \therefore \lambda_2 &= 7 \end{aligned}$$

Para encontrar los eigenvectores, basta con sustituir los eigenvalores en $G - \lambda I$

$$\begin{aligned} G - 3I &= \begin{pmatrix} 2 & -2 \\ -2 & 2 \end{pmatrix} \sim \begin{pmatrix} 2 & -2 \\ 0 & 0 \end{pmatrix} \\ \Rightarrow x_1 &= x_2 \end{aligned}$$

$$\begin{aligned} G - 7I &= \begin{pmatrix} -2 & -2 \\ -2 & -2 \end{pmatrix} \sim \begin{pmatrix} -2 & -2 \\ 0 & 0 \end{pmatrix} \\ \Rightarrow x_1 &= -x_2 \end{aligned}$$

$$\therefore \begin{pmatrix} 1 \\ 1 \end{pmatrix}_{\lambda=3} \wedge \begin{pmatrix} -1 \\ 1 \end{pmatrix}_{\lambda=7},$$

4 Aplicaciones en IA: reducción de dimensionalidad

1. El Principal Component Analysis busca reducir el número de dimensiones en conjuntos de datos muy largos a componentes principales que mantengan la mayor cantidad de información original posible. Esto lo hace transformando las variables posiblemente correlacionadas en un conjunto más pequeño de variables no correlacionadas llamadas componentes principales. Básicamente, reduce el número de variables aleatorias con las que vamos a trabajar sin deshacernos de la información original más importante.

Estos componentes son combinaciones lineales de las variables originales que tienen la mayor varianza a comparación de otras combinaciones lineales.

Esta técnica estadística transforma el conjunto de datos original en un nuevo sistema de coordenadas estructurado por los componentes principales, los eigenvalores y eigenvectores de la matriz de covarianza permiten el análisis de estas transformaciones lineales.

Por ejemplo, suponga que tiene un conjunto de datos y lo ha relacionado con múltiples características, resultando en una gráfica multidimensional de comparaciones de todas las características entre ellas mismas para reconocer patrones. Entonces, los eigenvectores otorgan la dirección de varianza en dicha gráfica y los eigenvalores denotan la importancia de esta dirección (a mayor valor, más crítico).

Un caso práctico sería suponer que tenemos un conjunto de información muy grande acerca de algunos modelos de carros y nos interesa saber las relaciones que hay entre todas sus características, después de definirlas todas, estandarizamos todas nuestras variables aleatorias para que todas usen la misma escala. Después, creamos la matriz de covarianza entre todas las variables aleatorias para saber qué tan correlacionadas están entre ellas (en pares) y según los signos de los resultados podremos saber si las combinaciones están correlacionadas o no. Posteriormente obtenemos los eigenvectores (los componentes principales) y eigenvalores de la matriz, y ordenamos los vectores según sus valores propios. Seguido de ello seleccionamos los componentes principales con los que vamos a quedarnos, generalmente son aquellos con los eigenvalores más altos. Finalmente transformamos esta información en un nuevo sistema coordinado definido por los componentes principales, esto es proyectar los vectores de características creados por los eigenvectores de la matriz de covarianza en nuevos ejes definidos por los componentes principales, creando nueva información que captura la mayor parte de la original pero con menos dimensiones.

2. Buscamos las matrices $U_{2 \times 2}$, $D_{2 \times 2}$ y $V_{2 \times 2}$ tales que se cumpla que

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

Definimos

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

Entonces

$$\begin{pmatrix} 3 & 1 \\ 2 & 2 \end{pmatrix} \begin{pmatrix} 3 & 2 \\ 1 & 2 \end{pmatrix} = \begin{pmatrix} 10 & 8 \\ 8 & 8 \end{pmatrix}, \begin{pmatrix} 3 & 2 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} 3 & 1 \\ 2 & 2 \end{pmatrix} = \begin{pmatrix} 13 & 7 \\ 7 & 5 \end{pmatrix}$$

Encontramos los eigenvalores y eigenvectores de AA^T y $A^T A$

$$\begin{aligned} |AA^T - \lambda I| = 0 &= \begin{vmatrix} 10 - \lambda & 8 \\ 8 & 8 - \lambda \end{vmatrix} \\ &= (10 - \lambda)(8 - \lambda) - 64 \\ &= 80 - 18\lambda + \lambda^2 - 64 \\ &= \lambda^2 - 18\lambda + 16 \\ &\Rightarrow \lambda_1 = 9 + \sqrt{65}, \lambda_2 = 9 - \sqrt{65} \\ &\Rightarrow v_1 = \begin{pmatrix} 1.133 \\ 1 \end{pmatrix}, v_2 = \begin{pmatrix} -0.883 \\ 1 \end{pmatrix} \end{aligned}$$

$$\begin{aligned} |A^T A - \lambda I| = 0 &= \begin{vmatrix} 13 - \lambda & 7 \\ 7 & 5 - \lambda \end{vmatrix} \\ &= (13 - \lambda)(5 - \lambda) - 49 \\ &= 65 - 18\lambda + \lambda^2 - 49 \\ &= \lambda^2 - 18\lambda + 16 \\ &\Rightarrow \lambda_1 = 9 + \sqrt{65}, \lambda_2 = 9 - \sqrt{65} \\ &\Rightarrow v_1 = \begin{pmatrix} 1.723 \\ 1 \end{pmatrix}, v_2 = \begin{pmatrix} -0.580 \\ 1 \end{pmatrix} \end{aligned}$$

Al normalizar cada vector característico tenemos que $L_1 = 1.511$, $L_2 = 1.334$, $L_3 = 1.992$ y $L_4 = 1.156$, entonces

$$\begin{aligned} u_1 &= \begin{pmatrix} 0.750 \\ 0.662 \end{pmatrix}, u_2 = \begin{pmatrix} -0.662 \\ 0.750 \end{pmatrix} \\ v_1 &= \begin{pmatrix} 0.865 \\ 0.502 \end{pmatrix}, v_2 = \begin{pmatrix} -0.502 \\ 0.865 \end{pmatrix} \end{aligned}$$

Por tanto tenemos que

$$\begin{aligned} D &= \begin{pmatrix} \sqrt{9 + \sqrt{65}} & 0 \\ 0 & \sqrt{9 - \sqrt{65}} \end{pmatrix} = \begin{pmatrix} 4.130 & 0 \\ 0 & 0.968 \end{pmatrix} \\ U &= (u_1 \ u_2) = \begin{pmatrix} 0.750 & -0.662 \\ 0.662 & 0.750 \end{pmatrix} \end{aligned}$$

$$V = (v_1 \ v_2) = \begin{pmatrix} 0.865 & -0.502 \\ 0.502 & 0.865 \end{pmatrix}$$

Finalmente concluimos la descomposición de la forma $H = UDV^T$

$$\begin{pmatrix} 3 & 1 \\ 2 & 2 \end{pmatrix} = \begin{pmatrix} 0.750 & -0.662 \\ 0.662 & 0.750 \end{pmatrix} \begin{pmatrix} 4.130 & 0 \\ 0 & 0.968 \end{pmatrix} \begin{pmatrix} 0.865 & 0.502 \\ -0.502 & 0.865 \end{pmatrix}$$

3. El álgebra lineal es fundamental para el aprendizaje profundo en redes neuronales por diversas razones, por ejemplo:
 - La forma de representación de los datos son vectores y matrices, al igual que las entradas y salidas de las redes neuronales.
 - En cada capa de la red neuronal se realiza una multiplicación de matrices para transformar los datos antes de pasar a la siguiente capa.
 - Las técnicas como el PCA o la SVD son útiles para reducir la dimensionalidad de los datos antes de alimentarlos a alguna red neuronal para mejorar la eficiencia del aprendizaje y la generalización de los modelos.
4. Los espacios vectoriales permiten representar, manipular y aprender a partir de datos en IA. A través de transformaciones lineales y reducción de dimensionalidad, los datos son proyectados en un espacio que facilita la captura de patrones y relaciones.