Laboratorio de Repaso Álgebra Lineal

13 de Febrero 2025

1 Introducción a Álgebra Lineal

1.1 Operaciones con matrices y determinantes

1. Encuentre la inversa de la siguiente matriz y verifique su resultado:

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

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

$$\begin{pmatrix} 1 & 2 & 3 & | & 1 & 0 & 0 \\ 0 & 1 & 4 & | & 0 & 1 & 0 \\ 5 & -4 & -15 & | & -5 & 0 & 1 \end{pmatrix}$$

$$\begin{pmatrix} 1 & 0 & -5 & | & 1 & -2 & 0 \\ 0 & 1 & 4 & | & 0 & 1 & 0 \\ 0 & 0 & 1 & | & -5 & 4 & 1 \end{pmatrix}$$

$$\begin{pmatrix} 1 & 0 & 0 & | & -25 & 18 & 5 \\ 0 & 1 & 0 & | & 20 & -15 & -4 \\ 0 & 0 & 1 & | & -5 & 4 & 1 \end{pmatrix}$$

Gracias al método montante podemos calcular la adjunta y la determinante teniendo que la determinante es $= 1$ y la adjunta es:

$$\begin{pmatrix} -25 & 18 & 5 \\ 20 & -15 & -4 \\ -5 & 4 & 1 \end{pmatrix}$$

al dividir esta entre la determinante obtenemos la inversa

$$\begin{pmatrix} -25 & 18 & 5 \\ 20 & -15 & -4 \\ -5 & 4 & 1 \end{pmatrix}$$

2. Demuestre la propiedad de que el determinante de un producto de matrices es igual al producto de los determinantes. Teniendo las siguientes matrices realizamos una multiplicación de ellas

$$\begin{pmatrix} a & b & | & e & f \\ c & d & | & g & h \end{pmatrix}$$

la determinante de la primer matriz es
ad - cb y de la matriz 2 es
eh - gf y su multiplicación es
adeh - adgf - cbeh + cbgf

$$\begin{pmatrix} ae + bg & af + bh \\ ce + dg & cf + dh \end{pmatrix}$$

posterior sacamos la determinante de esta:
aecf + aedh + bgcf + bgdh - (ceaf + cebh + dgaf + dgbh)
simplificando
aedh + bgcf - cebh - dgaf
Por ultimo comparamos
aedh + bgcf - cebh - dgaf = adeh - adgf - cbeh + cbgf
y vemos que son iguales

1.2 Sistemas de ecuaciones lineales

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

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

iniciamos x, y, z = 0

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

posterior sustituimos los valores de x, y, z en la primera ecuación y con los nuevos valores en la segunda, y repetimos el proceso en la tercera, repetimos este proceso hasta que lleguemos a una solución
Primera Iteración

$$\begin{aligned} x &= \frac{7 + y - z}{4} = \frac{7}{4} \\ y &= \frac{1 + 2x + 2z}{4} = \frac{9}{8} \\ z &= \frac{5 - x + y}{3} = \frac{35}{24} \end{aligned}$$

Segunda Iteración

$$x = \frac{7+y-z}{4} = \frac{5}{3}$$
$$y = \frac{1+2x+2z}{4} = \frac{29}{16}$$
$$z = \frac{5-x+y}{3} = \frac{247}{48}$$

Calculamos el error de cada uno

$$\epsilon_x = \left| \frac{\frac{5}{3} - \frac{7}{4}}{\frac{5}{3}} \right| * 100 = 5\%$$
$$\epsilon_y = \left| \frac{\frac{29}{16} - \frac{9}{8}}{\frac{29}{16}} \right| * 100 = 37.93\%$$
$$\epsilon_z = \left| \frac{\frac{247}{48} - \frac{35}{24}}{\frac{247}{48}} \right| * 100 = 71.65\%$$

Tercera Iteración

$$x = \frac{7+y-z}{4} = \frac{11}{12}$$
$$y = \frac{1+2x+2z}{4} = \frac{105}{32}$$
$$z = \frac{5-x+y}{3} = \frac{707}{288}$$

Calculamos el error de cada uno

$$\epsilon_x = \left| \frac{\frac{11}{12} - \frac{5}{3}}{\frac{11}{12}} \right| * 100 = 81.81\%$$
$$\epsilon_y = \left| \frac{\frac{105}{32} - \frac{29}{16}}{\frac{105}{32}} \right| * 100 = 44.76\%$$
$$\epsilon_z = \left| \frac{\frac{707}{288} - \frac{247}{48}}{\frac{707}{288}} \right| * 100 = 109.61\%$$

Cuarta Iteración

$$x = \frac{7+y-z}{4} = \frac{1127}{576}$$
$$y = \frac{1+2x+2z}{4} = \frac{943}{384}$$
$$z = \frac{5-x+y}{3} = 1.833043981$$

Calculamos el error de cada uno

$$\epsilon_x = \left| \frac{\frac{1127}{567} - \frac{11}{12}}{\frac{1127}{576}} \right| * 100 = 53.88\%$$

$$\epsilon_y = \left| \frac{\frac{943}{384} - \frac{105}{32}}{\frac{943}{384}} \right| * 100 = 33.61\%$$

$$\epsilon_z = \left| \frac{1.833043981 - \frac{707}{288}}{1.833043981} \right| * 100 = 33.92\%$$

Quinta Iteración

$$x = \frac{7 + y - z}{4} = 1.905671296$$

$$y = \frac{1 + 2x + 2z}{4} = 2.119357639$$

$$z = \frac{5 - x + y}{3} = 1.737895448$$

Calculamos el error de cada uno

$$\epsilon_x = \left| \frac{1.905671296 - \frac{1127}{567}}{1.905671296} \right| * 100 = 4.30\%$$

$$\epsilon_y = \left| \frac{2.119357639 - \frac{943}{384}}{2.119357639} \right| * 100 = 15.87\%$$

$$\epsilon_z = \left| \frac{1.737895448 - 1.833043981}{1.737895448} \right| * 100 = 5.47\%$$

Sexta Iteración

$$x = \frac{7 + y - z}{4} = 1.845365548$$

$$y = \frac{1 + 2x + 2z}{4} = 2.041630498$$

$$z = \frac{5 - x + y}{3} = 1.732088317$$

Calculamos el error de cada uno

$$\epsilon_x = \left| \frac{1.845365548 - 1.905671296}{1.845365548} \right| * 100 = 3.26\%$$

$$\epsilon_y = \left| \frac{2.041630498 - 2.119357639}{2.041630498} \right| * 100 = 3.81\%$$

$$\epsilon_z = \left| \frac{1.732088317 - 1.737895448}{1.732088317} \right| * 100 = 0.34\%$$

Séptima Iteración

$$\begin{aligned} x &= \frac{7+y-z}{4} = 1.82738554525 \\ y &= \frac{1+2x+2z}{4} = 2.029736931125 \\ z &= \frac{5-x+y}{3} = 1.734117128625 \end{aligned}$$

Calculamos el error de cada uno

$$\begin{aligned} \epsilon_x &= \left| \frac{1.82738554525 - 1.845365548}{1.82738554525} \right| * 100 = 0.98\% \\ \epsilon_y &= \left| \frac{2.029736931125 - 2.041630498}{2.029736931125} \right| * 100 = 0.58\% \\ \epsilon_z &= \left| \frac{1.734117128625 - 1.732088317}{1.734117128625} \right| * 100 = 0.12\% \end{aligned}$$

Al obtener un margen de error cercano al 0% podemos decir que los valores están muy cercanos a los que buscamos, por lo tanto

$$\begin{aligned} x &\approx 1.82738554525 \\ y &\approx 2.029736931125 \\ z &\approx 1.734117128625 \end{aligned}$$

4. Encuentre todas las soluciones del sistema homogéneo:

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

Podemos ver que el sistema anterior descrito vemos que la ecuación 1 es múltiplo de la ecuación 2 y 3 por lo tanto al intentar resolverlo obtendremos una solución trivial 0 donde todos los valores x, y, z obtienen valores de 0 y esto da solución a la matriz, sin embargo, este tambien tiene soluciones infinitas y para esto obtenemos

$$\begin{aligned} x &= -2y - 3z \\ y &= t \\ z &= s \end{aligned}$$

esto para t, s $\epsilon\mathbb{R}$

1.3 Espacios vectoriales y autovalores/autovectores

5. Encuentre la base y la dimensión del subespacio generado por los vectores:

$$\{(1,2,3),(2,4,6),(3,6,9)\}$$

Convertimos los valores dados en una matriz 3 x 3 y realizamos una eliminación de Gauss

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

para lo cual multiplicamos la primer fila por 2 y se lo restamos a la segunda y posteriormente multiplicamos la primer fila por 3 y lo restamos a la tercera obteniendo la siguiente matriz

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

Observamos que la única fila restante es la primera, de esta manera podemos decir que estos valores tienen como base el vector $\{1,2,3\}$ y con una dimensión $= 1$

6. Determine los autovalores y autovectores de la matriz:

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

Polinomio característico:

$$\det(G - \lambda I) = \lambda^2 - 10\lambda + 21$$

Solución de la ecuación cuadrática:

$$\lambda = \frac{-(-10) \pm \sqrt{(-10)^2 - 4(1)(21)}}{2(1)}$$

$$\lambda = \frac{10 \pm \sqrt{100 - 84}}{2} = \frac{10 \pm \sqrt{16}}{2} = \frac{10 \pm 4}{2}$$

Autovalores:

$$\lambda_1 = 7, \quad \lambda_2 = 3$$

Autovectores para $\lambda_1 = 7$:

$$\begin{bmatrix} -2 & -2 \\ -2 & -2 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = 0 \quad \Rightarrow \quad v_1 + v_2 = 0 \quad \Rightarrow \quad v_1 = -v_2$$

Autovectores para $\lambda_2 = 3$:

$$\begin{bmatrix} 2 & -2 \\ -2 & 2 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = 0 \quad \Rightarrow \quad v_1 = v_2$$

Por lo tanto obtenemos que:
Los autovalores de G son:
$\lambda_1 = 7$
$\lambda_2 = 3$.

Los autovectores correspondientes son:
Para $\lambda_1 = 7$, un autovector es $\begin{bmatrix} -1 \\ 1 \end{bmatrix}$.
Para $\lambda_2 = 3$, un autovector es $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$.

1.4 Aplicaciones en IA: reducción de dimensionalidad

7. Explique cómo el PCA (Análisis de Componentes Principales) utiliza el álgebra lineal para reducir dimensiones.

El análisis de componentes principales (PCA) utiliza álgebra lineal para reducir dimensiones en un conjunto de datos mediante vectores propios y valores propios.

Paso 1:
Identificar las características más importantes Se calcula la matriz de covarianza, que representa la relación entre los diferentes elementos de los datos
Se analizan los vectores propios y los valores propios de la matriz de covarianza

Paso 2:
Crear un nuevo espacio de características
Se eligen los k vectores propios superiores en función de los k valores propios más grandes
Se proyectan los datos en el nuevo espacio de características

Paso 3
Descomponer las variables en componentes principales
Se encuentran los componentes principales linealmente independientes que explican la mayor parte de la varianza de los datos originales

El PCA es una técnica de reducción de dimensionalidad que se utiliza para transformar grandes conjuntos de datos en otros más pequeños que contienen la mayor parte de la información.

Se utiliza en bioestadística, marketing, sociología y otros campos.

8. Calcule la descomposición en valores singulares (SVD) de la matriz:

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

Empezamos encontrando la transpuesta de la matriz

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

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

Posterior a esto realizamos la multiplicación de matrices entre ellas

$$H^T H = \begin{bmatrix} 3 & 2 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} 3 & 1 \\ 2 & 2 \end{bmatrix} = \begin{bmatrix} 13 & 7 \\ 7 & 5 \end{bmatrix}$$

Multiplicamos la matriz resultante por λI y calculamos su determinante

$$\det(H^T H - \lambda I) = (13 - \lambda)(5 - \lambda) - 49 = \lambda^2 - 18\lambda + 16$$

Usamos la formula general para determinar los valores de λ

$$\lambda = \frac{18 \pm \sqrt{324 - 64}}{2} = \frac{18 \pm \sqrt{260}}{2} = \frac{18 \pm 2\sqrt{65}}{2}$$

$$\lambda_1 = 9 + \sqrt{65}, \quad \lambda_2 = 9 - \sqrt{65}$$

Por ultimo los valores singulares de esta matriz son representados por las raíces de λ

$$\sigma_1 = \sqrt{9 + \sqrt{65}}, \quad \sigma_2 = \sqrt{9 - \sqrt{65}}$$

9. Analice el uso de álgebra lineal en el aprendizaje profundo con redes neuronales.

El **álgebra lineal** juega un papel fundamental en el aprendizaje profundo, particularmente en el contexto de las **redes neuronales artificiales**. A continuación, se presenta un análisis de su uso clave.

1. Representación de datos

Los datos de entrada a una red neuronal suelen representarse en forma de **vectores** o **matrices**. Por ejemplo, una imagen en escala de grises de 28×28 píxeles se puede representar como una matriz de 28×28 de números. Cuando se tiene un conjunto de datos de múltiples imágenes, se usa una **matriz** para representar todo el conjunto de datos. Cada fila puede ser un vector de características de una imagen.

Las **etiquetas** o **salidas deseadas** también se representan de manera similar, y el proceso de aprendizaje ajusta los pesos de la red para minimizar la diferencia entre la salida de la red y estas etiquetas.

2. Operaciones en redes neuronales

Las redes neuronales consisten en múltiples capas de nodos (neuronas). Entre cada capa, se realizan operaciones de **producto de matrices** y **sumas** para transformar las entradas a través de pesos, sesgos y activaciones.

3. Autovalores y autovectores

El **análisis espectral** de las matrices de peso también se encuentra presente en el entrenamiento de redes neuronales. Por ejemplo, algunas técnicas de optimización, como el **descenso de gradiente estocástico**, pueden aprovechar el conocimiento de los autovalores y autovectores de las matrices de peso para acelerar el proceso de convergencia, o bien para estabilizar la red durante el entrenamiento.

4. Descomposición en valores singulares (SVD)

SVD es una técnica que descompone cualquier matriz en tres matrices: U, Σ y V^T. En el contexto de redes neuronales, SVD se usa en ciertas técnicas de compresión de modelos, donde las grandes matrices de peso se descomponen en matrices más pequeñas para reducir la complejidad computacional y el sobreajuste.

5. Reducción de dimensionalidad y regularización

La **reducción de dimensionalidad**, como la técnica de **Análisis de Componentes Principales (PCA)**, se basa en conceptos de álgebra lineal. En redes neuronales, se usa para reducir el número de características en los datos de entrada antes de la fase de entrenamiento.

Técnicas de **regularización**, como **Dropout** y **L2 Regularization**, tienen bases algebraicas que restringen el espacio de soluciones posibles, ayudando a prevenir el sobreajuste.

6. Optimización de redes neuronales

En **aprendizaje profundo**, se optimizan las matrices de pesos utilizando **métodos de optimización** que dependen de la álgebra lineal, como el **descenso de gradiente** y sus variantes (Adam, RMSprop). Estos algoritmos calculan gradientes utilizando derivadas parciales de las matrices, lo cual requiere el uso eficiente de productos de matrices.

7. Redes neuronales convolucionales (CNN)

En las **redes neuronales convolucionales** (CNN), las operaciones de convolución también se pueden representar como productos de matrices. En una operación de convolución, el filtro se desliza sobre la imagen de entrada y, mediante multiplicación y suma, produce una nueva imagen o mapa de características.

8. Redes neuronales recurrentes (RNN)

En las **redes neuronales recurrentes** (RNN), las matrices se utilizan para representar las conexiones entre las neuronas en diferentes momentos de tiempo. Las propagaciones hacia adelante y hacia atrás en el tiempo (backpropagation through time) implican productos de matrices para calcular los gradientes y ajustar los pesos.

10. Explique el impacto de los espacios vectoriales en la representación de datos en IA.

    Los **espacios vectoriales** son fundamentales en la **representación de datos** en **inteligencia artificial (IA)**, pues permiten estructurar y manipular información de manera eficiente, facilitando tareas como el aprendizaje automático, el procesamiento de lenguaje natural, y la visión computacional. A continuación, se detalla su impacto en estos contextos.

1. Representación de Datos Numéricos

En IA, la información (como imágenes, texto, audio, etc.) se representa comúnmente en forma de **vectores** o **matrices**. Los datos originales, que pueden estar en cualquier formato (como píxeles de una imagen, palabras en un texto o muestras de audio), se transforman en **vectores** de características que pueden ser manipulados matemáticamente.

Las **etiquetas** o **salidas deseadas** también se representan de manera similar, y el proceso de aprendizaje ajusta los pesos de la red para minimizar la diferencia entre la salida de la red y estas etiquetas.

2. Espacios Vectoriales y Aprendizaje Automático

Los algoritmos de **aprendizaje automático** (machine learning) utilizan espacios vectoriales para representar los datos de entrada y aprender patrones o relaciones a partir de esos datos. Algunos ejemplos relevantes incluyen:

Regresión Lineal y Clasificación

La **regresión lineal** y muchos algoritmos de clasificación, como **máquinas de soporte vectorial** (SVM), operan en un espacio vectorial donde las **características** de los datos se representan como vectores, y el modelo aprende a predecir una salida (un valor numérico o una clase) a partir de esos vectores.

En **SVM**, la idea central es encontrar un hiperplano en un espacio vectorial de alta dimensión que separe las diferentes clases de datos. Esto se logra mediante un **producto interno** y la **proyección** de los datos en ese espacio.

Reducción de Dimensionalidad

La **reducción de dimensionalidad**, como la **Análisis de Componentes Principales (PCA)**, utiliza la idea de proyección de datos en un espacio de menor dimensión manteniendo la mayor parte de la varianza. PCA encuentra una nueva base ortogonal en el espacio vectorial de los datos, lo que facilita tareas posteriores, como la clasificación o el clustering.

3. Similitud y Distancia en Espacios Vectoriales

En IA, medir la **similitud** o **distancia** entre vectores es clave para tareas como el clustering (agrupamiento), la recomendación, y la búsqueda de información. Algunas de las medidas más comunes incluyen:

- **Distancia Euclidiana**: Es la medida más común para la similitud entre vectores. Se utiliza en muchos algoritmos de aprendizaje automático, como **k-means**.
- **Coseno de similitud**: Utilizado especialmente en **procesamiento de texto** para medir la similitud entre documentos. Se calcula el coseno del ángulo entre dos vectores, lo que da una idea de cuán similares son los documentos representados como vectores.

4. Transformaciones Lineales y Redes Neuronales

Las redes neuronales, en especial las **redes neuronales profundas (DNN)**, utilizan los **espacios vectoriales** en cada capa de la red. Los datos pasan a través de transformaciones lineales y no lineales mientras viajan de una capa a otra:

Transformaciones lineales

Son operaciones donde el input (vectores de características) se multiplica por una matriz de pesos y se le añade un vector de sesgo, que es una forma de **proyección lineal** en un espacio vectorial de mayor dimensión. Este proceso permite que el modelo aprenda representaciones complejas de los datos.

Funciones de activación

Aunque las transformaciones lineales son fundamentales, las funciones de activación (como **ReLU**, **sigmoide** o **tanh**) introducen no linealidades que permiten a las redes neuronales aprender representaciones no lineales de los datos en espacios vectoriales de alta dimensión.

5. Reducción de Dimensionalidad y Regularización

La **reducción de dimensionalidad**, como la técnica de **Análisis de Componentes Principales (PCA)**, se basa en conceptos de álgebra lineal. En redes neuronales, se usa para reducir el número de características en los datos de entrada antes de la fase de entrenamiento.

Técnicas de **regularización**, como **Dropout** y **L2 Regularization**, tienen bases algebraicas que restringen el espacio de soluciones posibles, ayudando a prevenir el sobreajuste.

6. Optimización de Redes Neuronales

En **aprendizaje profundo**, se optimizan las matrices de pesos utilizando **métodos de optimización** que dependen de la álgebra lineal, como el **descenso de gradiente** y sus variantes (Adam, RMSprop). Estos algoritmos calculan gradientes utilizando derivadas parciales de las matrices, lo cual requiere el uso eficiente de productos de matrices.

7. Redes Neuronales Convolucionales (CNN)

En las **redes neuronales convolucionales** (CNN), las imágenes (que son matrices) se transforman en **vectores** a medida que las convoluciones extraen características de la imagen. Estas características luego se utilizan para tomar decisiones (como clasificación).

Cada filtro convolucional extrae un conjunto de características que puede representarse como un **vector de activación**. Estos vectores se procesan a través de capas adicionales para crear representaciones más abstractas de la imagen.

8. Redes Neuronales Recurrentes (RNN)

En las **redes neuronales recurrentes** (RNN) y sus variantes, como **LSTMs** (Long Short-Term Memory) y **GRUs** (Gated Recurrent Units), operan sobre secuencias de datos, como texto o series temporales. Estas redes procesan vectores en cada paso de tiempo y mantienen una representación del estado a lo largo del tiempo.

Cada paso de tiempo procesa un vector de entrada y, usando el **espacio vectorial**, actualiza su representación del estado oculto (que es otro vector). Esto permite a la red aprender patrones secuenciales o temporales.