UNIVERSIDAD AUTÓNOMA DE NUEVO LEÓN

Facultad de Ciencias Físico Matemáticas

Inteligencia artificial

Laboratorio de Repaso de Algebra Lineal

Docente:

Luis Angel Gutierrez Rodriguez

Estudiante:

Jesús Roberto Dávila González

2063584

Grupo: 031

San Nicolas de los Garza, N.L, a 22 de febrero de 2025.

Introducción a Álgebra Lineal

4.1 Operaciones con matrices y determinantes

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

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

2. Demuestre la propiedad de que el determinante de un producto de matrices es igual al producto de los determinantes.

4.2 Sistemas de ecuaciones lineales

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

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

4. Encuentre todas las soluciones del sistema homogéneo:

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

4.3 Espacios vectoriales y auto-valores/auto-vectores

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

6. Determine los autovalores y autovectores de la matriz:

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

4.4 Aplicaciones en IA: reducción de dimensionalidad

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

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

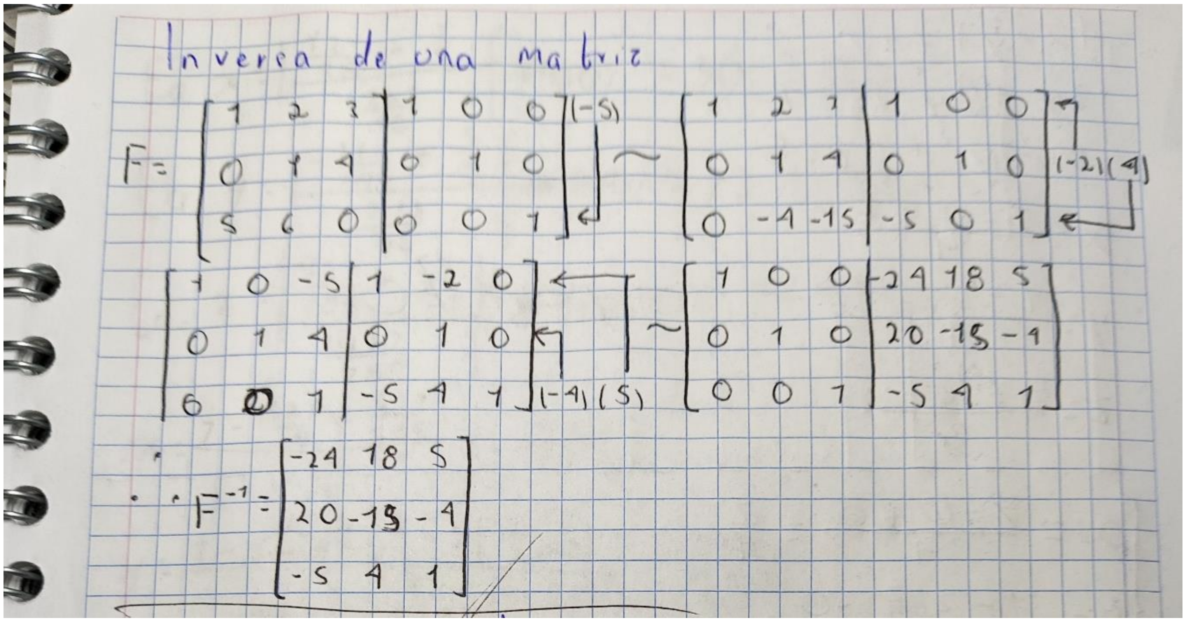
$$H = \begin{bmatrix} 3 & 1 \\ 2 & 2 \end{bmatrix}$$

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

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

1. Operaciones con matrices

1.1 Inversa de una matriz



1.2 El Determinante de un Producto de Matrices

Lema 1. (i) Si A es una matriz elemental que provino de I_n intercambiando en ésta dos de sus líneas, entonces

$$\det A = -1$$

(ii) Si A es una matriz elemental que provino de $I_n$multiplicando en ésta una de sus líneas por por c, entonces

$$\det A = c$$

(iii) Si A es una matriz elemental que provino de In sustituyendo en ésta una de sus líneas por ella misma más un múltiplo de otra línea, entonces

$$\det A = 1$$

Lema 2. Sea B una matriz de orden n y sea A una matriz elemental. Entonces

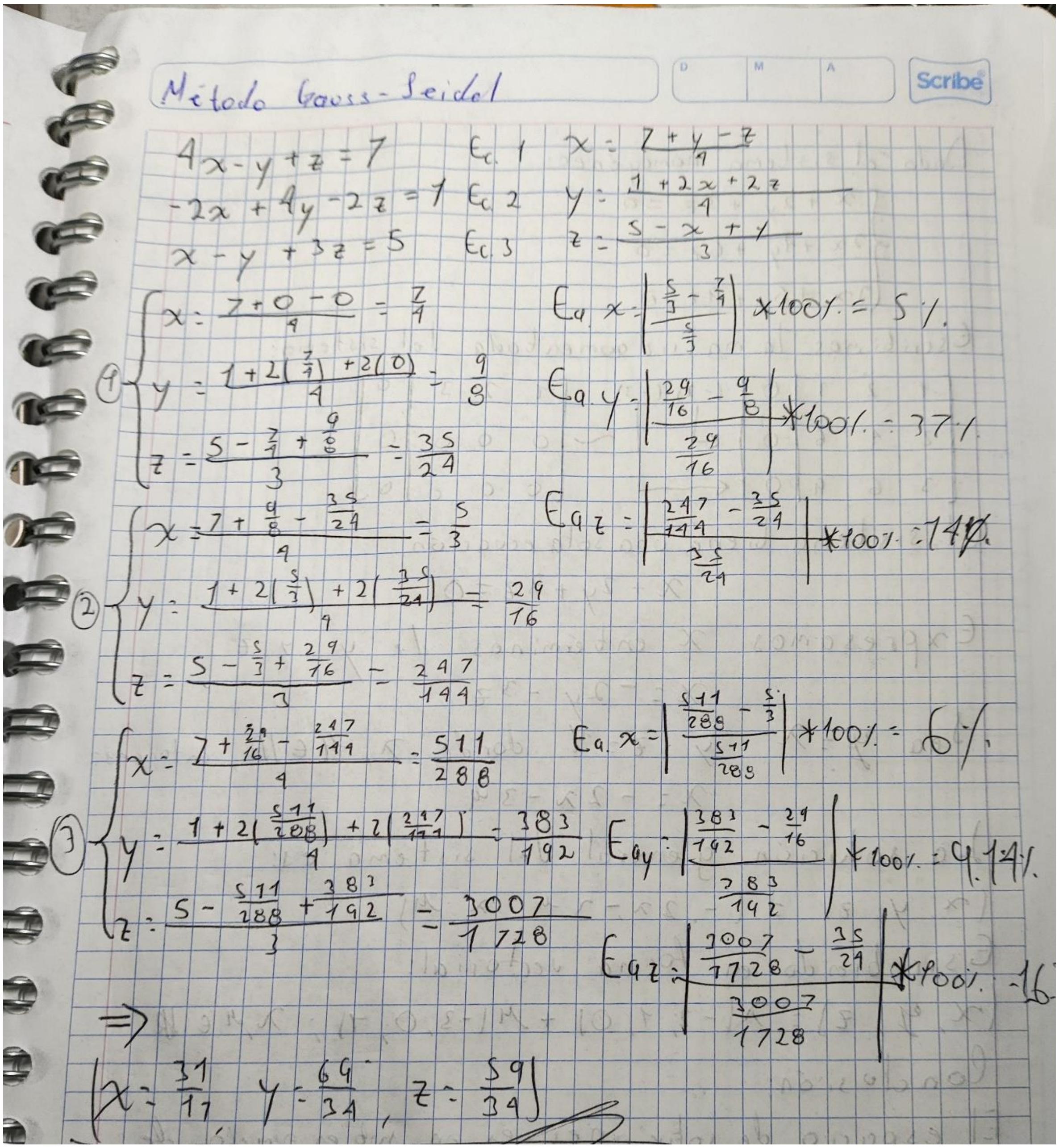
$$\det AB = \det A \det B$$

Demostración. Si B' es la matriz que se obtiene de B realizando en esta una operación elemental, entonces según lo anterior B' = AB, donde A es la matriz elemental que provino de I_n realizando en ésta la misma operación elemental. Entonces

$$\det AB = \det B' = \begin{pmatrix} -1 \\ c \\ 1 \end{pmatrix} \det B = \det A \det B$$

2. Sistemas de ecuaciones lineales

2.1 Método de Gauss-Seidel



Método Gauss-Seidel

$4x - y + z = 7$ Ec 1 $\quad x = \dfrac{7 + y - z}{4}$

$-2x + 4y - 2z = 1$ Ec 2 $\quad y = \dfrac{1 + 2x + 2z}{4}$

$x - y + 3z = 5$ Ec 3 $\quad z = \dfrac{5 - x + y}{3}$

(1) $x = \dfrac{7 + 0 - 0}{4} = \dfrac{7}{4}$, $\quad y = \dfrac{1 + 2\left(\frac{7}{4}\right) + 2(0)}{4} = \dfrac{9}{8}$, $\quad z = \dfrac{5 - \frac{7}{4} + \frac{9}{8}}{3} = \dfrac{35}{24}$

(2) $x = \dfrac{7 + \frac{9}{8} - \frac{35}{24}}{4} = \dfrac{5}{3}$, $\quad y = \dfrac{1 + 2\left(\frac{5}{3}\right) + 2\left(\frac{35}{24}\right)}{4} = \dfrac{29}{16}$, $\quad z = \dfrac{5 - \frac{5}{3} + \frac{29}{16}}{3} = \dfrac{247}{144}$

$E_a\,x = \left|\dfrac{\frac{5}{3} - \frac{7}{4}}{\frac{5}{3}}\right| * 100\% = 5\%$

$E_a\,y = \left|\dfrac{\frac{29}{16} - \frac{9}{8}}{\frac{29}{16}}\right| * 100\% = 37\%$

$E_a\,z = \left|\dfrac{\frac{247}{144} - \frac{35}{24}}{\frac{35}{24}}\right| * 100\% = 14\%$

(3) $x = \dfrac{7 + \frac{29}{16} - \frac{247}{144}}{4} = \dfrac{511}{288}$, $\quad y = \dfrac{1 + 2\left(\frac{511}{288}\right) + 2\left(\frac{247}{144}\right)}{4} = \dfrac{383}{192}$, $\quad z = \dfrac{5 - \frac{511}{288} + \frac{383}{192}}{3} = \dfrac{3007}{1728}$

$E_a\,x = \left|\dfrac{\frac{511}{288} - \frac{5}{3}}{\frac{511}{288}}\right| * 100\% = 6\%$

$E_a\,y = \left|\dfrac{\frac{383}{192} - \frac{29}{16}}{\frac{383}{192}}\right| * 100\% = 9.14\%$

$E_a\,z = \left|\dfrac{\frac{3007}{1728} - \frac{35}{24}}{\frac{3007}{1728}}\right| * 100\% = 16$

$\Rightarrow$

$x = \dfrac{31}{11}, \quad y = \dfrac{69}{34}, \quad z = \dfrac{59}{34}$

2.2 Sistema homogéneo

Dado el sistema homogéneo:

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

Escribimos la matriz aumentada del sistema:

$$\left[\begin{array}{ccc|c} 1 & 2 & 3 & 0 \\ 2 & 4 & 6 & 0 \\ 3 & 6 & 9 & 0 \end{array}\right] \xrightarrow{(-2)\ (-3)} \sim \left[\begin{array}{ccc|c} 1 & 2 & 3 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right]$$

El sistema tiene una sola ecuación

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

Expresamos x en términos de y y z

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

Sea $y = \lambda$ y $z = \mu$ donde $\lambda, \mu \in \mathbb{R}$, entonces:

$$x = -2\lambda - 3\mu$$

La solución general del sistema es:

$$(x, y, z) = (-2\lambda - 3\mu, \lambda, \mu)$$

Escribiendo en forma vectorial:

$$(x, y, z) = \lambda(-2, 1, 0) + \mu(-3, 0, 1) \; ; \; \lambda, \mu \in \mathbb{R}$$

Conclusión:

El espacio de soluciones es un subespacio de dimensión 2, generado por los vectores:

$(-2, 1, 0)$ y $(-3, 0, 1)$

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

3.1 Dado el conjunto de vectores:

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

queremos encontrar una base y la dimensión del subespacio generado por ellos.

Paso 1: Escribir la matriz de los vectores como filas

$$A = \begin{matrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{matrix}$$

Paso 2: Aplicar reducción por filas

Restamos el doble de la primera fila de la segunda fila y el triple de la primera fila de la tercera fila:

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

Paso 3: Determinar la base y la dimensión

Observamos que solo hay **una fila no nula** en la matriz escalonada, lo que indica que el **rango de la matriz es 1**.

Esto significa que todos los vectores originales son combinaciones lineales del primer vector (1, 2, 3), por lo que el **subespacio generado es de dimensión 1** y está generado por el vector $\{(1,2,3)\}$

Conclusión:

- **Base:** $\{(1, 2, 3)\}$
- **Dimensión:** 1

3.2 Dada la matriz G:

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

queremos encontrar sus autovalores y autovectores.

Paso 1: Calcular los autovalores

Los autovalores λ se encuentran resolviendo la ecuación característica:

$$\det G - \lambda I = 0$$

Restamos λI de G:

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

Calculamos su determinante:

$$\begin{matrix} 5-\lambda & -2 \\ -2 & 5-\lambda \end{matrix} = (5-\lambda)(5-\lambda) - (-2)(-2)$$

$$\det G - \lambda I = (5-\lambda)^2 - 4 = 25 - 10\lambda + \lambda^2 - 4 = \lambda^2 - 10\lambda + 21$$

Resolviendo la ecuación característica:

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

Factorizamos:

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

Por lo tanto, los **autovalores** son:

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

Paso 2: Encontrar los autovectores

Para cada autovalor λ, resolvemos $G - \lambda I = 0$

Autovalor $\lambda = 7$

$$G - 7I = \begin{matrix} 5-7 & -2 \\ -2 & 5-7 \end{matrix} = \begin{matrix} -2 & -2 \\ -2 & -2 \end{matrix}$$

Resolvemos:

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

Esto da la ecuación:

$$-2x - 2y = 0 \Longrightarrow x = -y$$

El autovector asociado es cualquier múltiplo de:

$$v_1 = \begin{matrix} 1 \\ -1 \end{matrix}$$

Autovalor $\lambda = 3$

$$G - 3I = \begin{matrix} 5-3 & -2 \\ -2 & 5-3 \end{matrix} = \begin{matrix} 2 & -2 \\ -2 & 2 \end{matrix}$$

Resolvemos:

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

Esto da la ecuación:

$$2x - 2y = 0 \Longrightarrow x = y$$

El autovector asociado es cualquier múltiplo de:

$$v_2 = \begin{matrix} 1 \\ 1 \end{matrix}$$

Conclusión

- **Autovalores:** $\lambda_1 = 7,\ \lambda_2 = 3$
- **Autovectores asociados:** $v_1 = \begin{matrix} 1 \\ -1 \end{matrix}, v_2 = \begin{matrix} 1 \\ 1 \end{matrix}$

4. Aplicaciones en IA: reducción de dimensionalidad

4.1 El **Análisis de Componentes Principales (PCA)** es un método de reducción de dimensionalidad que utiliza el álgebra lineal para transformar un conjunto de datos en un nuevo sistema de coordenadas, donde las variables están ordenadas según su varianza. A continuación, explicamos cómo se aplica PCA paso a paso utilizando conceptos del álgebra lineal:

1. Estandarización de los Datos

Dado un conjunto de datos con nnn observaciones y mmm características, se comienza por centrar los datos restando la media de cada variable. Esto asegura que las nuevas direcciones encontradas en PCA sean independientes de los valores absolutos de los datos.

$$X = X - \mu$$

donde μ es el vector de medias de las variables.

2. Cálculo de la Matriz de Covarianza

PCA busca direcciones de máxima varianza en los datos. Para encontrar estas direcciones, se calcula la matriz de covarianza:

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

donde cada elemento C_{ij} mide la relación lineal entre las variables X_i y X_j.

3. Cálculo de los Autovalores y Autovectores

El siguiente paso es encontrar los **autovalores** y **autovectores** de la matriz de covarianza C. Esto se hace resolviendo:

$$C_v = \lambda_v$$

donde:

- v son los **autovectores** (también llamados **componentes principales**), que indican las nuevas direcciones en las que se proyectarán los datos.
- λ son los **autovalores**, que indican la cantidad de varianza explicada por cada componente principal.

4. Selección de las Principales Componentes

Los autovalores se ordenan en orden **descendente**, y se eligen los primeros k autovectores correspondientes a los mayores autovalores. Esto determina cuántas dimensiones se conservan en la nueva representación de los datos.

5. Transformación de los Datos

Los datos originales X se proyectan en el nuevo espacio utilizando los kkk autovectores seleccionados:

$$X' = XV_k$$

donde $V_k$es la matriz formada por los k autovectores más importantes. Esto da como resultado un conjunto de datos con menor dimensión, pero que conserva la mayor cantidad de información posible.

4.2 La descomposición en valores singulares (SVD) de una matriz HHH se expresa como:

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

donde:

- U es una matriz ortogonal cuyas columnas son los autovectores de HH^T
- V es una matriz ortogonal cuyas columnas son los autovectores de $H^T H$.
- Σ es una matriz diagonal con los valores singulares de H, que son las raíces cuadradas de los autovalores de $H^T H$

Dada la matriz:

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

La descomposición en valores singulares (SVD) de la matriz H es:

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

donde:

$$U = \begin{matrix} -0.7497 & -0.6618 \\ -0.6618 & 0.7497 \end{matrix}$$

$$\sum = \begin{matrix} 4.1306 & 0 \\ 0 & 0.9684 \end{matrix}$$

$$V^T = \begin{matrix} -0.8649 & -0.5019 \\ -0.5019 & 0.8649 \end{matrix}$$

Estos valores corresponden a la descomposición ortogonal de H, donde los valores singulares son 4.1306 y 0.9684.

4.3 El álgebra lineal es fundamental en el aprendizaje profundo (Deep Learning) y en las redes neuronales artificiales. Se utiliza para representar datos, operaciones y transformaciones dentro de la arquitectura de una red neuronal. A continuación, se explican los conceptos clave y su relación con el álgebra lineal:

1. Representación de Datos con Vectores y Matrices

En aprendizaje profundo, los datos (imágenes, texto, sonido, etc.) se representan como **vectores** o **matrices**:

- **Imágenes:** Se representan como **matrices** de píxeles (en escala de grises) o **tensores** si tienen múltiples canales de color (RGB).
- **Texto:** Se transforma en vectores mediante técnicas como **Word Embeddings**.
- **Datos estructurados:** Se representan como matrices donde cada fila es una muestra y cada columna es una característica.

2. Operaciones Matriciales en Redes Neuronales

Las redes neuronales se basan en operaciones matriciales para transformar datos de una capa a otra. Cada neurona realiza cálculos de la forma:

$$Z = WX + b$$

donde:

- X es el **vector de entrada** (características de los datos).
- W es la **matriz de pesos**, que representa la fuerza de conexión entre neuronas.
- b es el **vector de sesgo**.
- Z es la **salida** antes de aplicar la función de activación.

La multiplicación de matrices permite que cada capa de la red transforme las entradas a una nueva representación.

3. Funciones de Activación y Transformaciones Lineales

Para introducir **no linealidad**, se aplican funciones de activación como:

- **ReLU** (Rectified Linear Unit): $f(x) = max\,(0, x)$
- **Sigmoide**: $f(x) = \frac{1}{1+e^{-x}}$
- **Tangente hiperbólica (tanh)**: $f(x) = \frac{e^{x}-e^{-x}}{e^{x}+e^{x}}$.

Estas funciones transforman los valores calculados en cada neurona y permiten que la red neuronal aprenda patrones complejos.

4. Propagación hacia adelante (Forward Propagation)

El álgebra lineal se usa en el **cálculo de la salida** de una red neuronal. La salida de cada capa es el resultado de una **multiplicación de matrices**, seguido de la aplicación de una función de activación:

$$A = \sigma(WX + b)$$

donde σ es la función de activación.

5. Retropropagación y Gradiente Descendente

El **aprendizaje** en redes neuronales ocurre ajustando los pesos W y sesgos b para minimizar el error. Esto se hace mediante:

1. **Cálculo del error** (pérdida) comparando la salida de la red con la salida real.
2. **Cálculo del gradiente** usando **derivadas parciales** y la **regla de la cadena** para determinar cómo cada peso afecta el error.
3. **Ajuste de los pesos** usando el método de **descenso del gradiente**:

$$W = W - \alpha \frac{\partial L}{\partial W}$$

donde α es la tasa de aprendizaje y LL es la función de pérdida.

Este proceso se basa en el **cálculo de derivadas de funciones matriciales** y en la **transposición e inversión de matrices**.

6. Descomposición y Factorización de Matrices

Se utilizan métodos del álgebra lineal como:

- **Descomposición en valores singulares (SVD)** y **autovalores/autovectores** para analizar redes neuronales y reducir dimensionalidad en capas ocultas.
- **Factorización matricial** en técnicas como **aprendizaje de embeddings** y **modelos de recomendación**.

7. Aplicaciones del Álgebra Lineal en Aprendizaje Profundo

- **Redes convolucionales (CNNs):** Usan **productos de matrices** para el filtrado de imágenes.
- **Redes recurrentes (RNNs):** Dependen de **multiplicaciones de matrices** en secuencias de datos.
- **Transformers (GPT, BERT):** Utilizan **productos escalares y atención** para modelar relaciones entre palabras.
- **Autoencoders:** Usan **descomposición matricial** para reducción de dimensión y compresión de datos.

4.4 El impacto de los espacios vectoriales en la representación de datos en Inteligencia Artificial (IA)

Los espacios vectoriales son la base matemática que permite representar datos en Inteligencia Artificial (IA). En este contexto, los datos son transformados en vectores en un espacio de alta dimensión, lo que facilita su procesamiento y análisis mediante modelos matemáticos y algoritmos de aprendizaje automático. A continuación, analizamos su impacto en distintas áreas de la IA:

1. Representación Numérica de Datos

Cualquier tipo de dato (texto, imágenes, audio, etc.) debe convertirse en un vector dentro de un espacio vectorial para ser procesado por algoritmos de IA:

- Texto → Vectores:
  - Modelos como Word2Vec, GloVe y Transformers (BERT, GPT) representan palabras como vectores en un espacio de alta dimensión, donde palabras con significados similares están más cerca entre sí.
  - Ejemplo: "rey" y "reina" estarán cerca en el espacio vectorial, ya que comparten significado semántico.

- Imágenes → Tensores (Vectores Multidimensionales):
  - Cada imagen se representa como una matriz de píxeles en espacios de color (RGB, escala de grises, etc.).
  - Redes neuronales convolucionales (CNN) aplican transformaciones lineales para extraer características relevantes.
- Sonido → Vectores de Características:
  - En procesamiento de voz, una señal de audio se convierte en un vector mediante Transformada de Fourier o MFCC (Mel-Frequency Cepstral Coefficients).

2. Búsqueda y Comparación de Datos

Los espacios vectoriales permiten definir una noción de distancia para comparar datos:

- Distancia Euclidiana (L2L_2):
  - Se usa para medir la similitud entre vectores.
  - Aplicada en reconocimiento facial: si un nuevo rostro está cerca en el espacio vectorial de una persona conocida, se considera una coincidencia.
- Producto Escalar y Cosenos:
  - En modelos de lenguaje, la similitud del coseno mide qué tan similares son dos documentos.
  - En motores de recomendación, se comparan vectores de usuarios y productos para encontrar preferencias similares.

3. Reducción de Dimensionalidad y Representación Eficiente

En IA, los datos suelen tener miles o millones de dimensiones, lo que puede hacer que los cálculos sean costosos. Para solucionar esto, se usan técnicas basadas en espacios vectoriales:

- PCA (Análisis de Componentes Principales): Reduce la dimensionalidad manteniendo la mayor varianza posible en los datos.
- Autoencoders: Redes neuronales que aprenden representaciones comprimidas en espacios vectoriales de menor dimensión.
- t-SNE y UMAP: Métodos para visualizar datos de alta dimensión en 2D o 3D.

4. Aplicaciones en IA Basadas en Espacios Vectoriales

Procesamiento del Lenguaje Natural (PLN)

Los modelos modernos como Word2Vec, BERT y GPT convierten palabras en vectores dentro de un espacio semántico, donde palabras con significados similares están más cerca.

Visión por Computadora

Las imágenes son representadas en espacios vectoriales mediante técnicas como:

- CNN (Redes Neuronales Convolucionales): Transforman imágenes en mapas de características vectoriales.
- Modelos de detección de objetos (YOLO, Faster R-CNN): Representan objetos en espacios de características para clasificarlos.

Motores de Recomendación

- Cada usuario y producto es un vector en un espacio de características.
- Se recomienda contenido basado en la proximidad entre vectores (usuarios con intereses similares).

Aprendizaje No Supervisado y Clustering

Los algoritmos de clustering como K-Means agrupan datos en función de la distancia entre sus vectores en el espacio.