UNIVERSIDAD AUTÓNOMA DE NUEVO LEÓN

FACULTAD DE CIENCIAS FÍSICO-MATEMÁTICAS

INTELIGENCIA ARTIFICIAL

ACT 8: LABORATORIO DE ALGEBRA LINEAL

Elaborado por:
Ángel Francisco Hernández Gámez

Introducción a Álgebra lineal

4.1 Operaciones con matrices y determinantes

---

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

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

Usando la fórmula para 3×3, se tiene:

$\det(F) = 1 \cdot (1 \cdot 0 - 4 \cdot 6) - 2 \cdot (0 \cdot 0 - 4 \cdot 5) + 3 \cdot (0 \cdot 6 - 1 \cdot 5)$

$-24 + 40 - 15 = 1$

Como el $\det(F) = 1$, la matriz es invertible y su inversa es la adjunta.

Primero se escribe la matriz F junto con la matriz identidad I de 3×3:

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

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

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

$$\left[\begin{array}{ccc|ccc} 1 & 2 & 0 & 16 & -12 & -3 \\ 0 & 1 & 0 & 20 & -15 & -4 \\ 0 & 0 & 1 & -5 & 4 & 1 \end{array}\right]$$

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

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

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

Para verificar que el resultado obtenido es la inversa de F, debemos comprobar que al multiplicar F por F^{-1} se obtiene la matriz identidad, es decir:

$F \cdot F^{-1} = I$

Dado

$$F = \begin{pmatrix} 1 & 2 & 3 \\ 0 & 1 & 4 \\ 5 & 6 & 0 \end{pmatrix} \quad \text{y} \quad F^{-1} = \begin{pmatrix} -24 & 18 & 5 \\ 20 & -15 & -4 \\ -5 & 4 & 1 \end{pmatrix}$$

Hacemos la multiplicación:

$$F \cdot F^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = I_3$$

Como $F \cdot F^{-1}$ nos da la matriz identidad, podemos verificar que

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

es, efectivamente, la inversa de F.

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

Sabemos que toda matriz invertible A se puede descomponer en un producto de matrices elementales. Es decir, hay una factorización de la forma:

$$A = E_1 E_2 \cdots E_r$$

donde cada E_i es una matriz elemental. Entonces a cada tipo de matriz elemental le corresponde una propiedad bien definida para su determinante:

- Si E es una matriz elemental que intercambia dos filas, $\det(E) = -1$.
- Si E es una matriz elemental que multiplica una fila por un escalar k, $\det(E) = k$
- Si E es una matriz elemental que suma a una fila un múltiplo de otra,$\det(E) = 1$

Para una matriz elemental E y cualquier matriz A se verifica que

$$\det(EA) = \det(E)\det(A)$$

Aplicando esto a la factorización de A, tenemos:

$$\det(AB) = \det(E_1 E_2 \cdots E_r B) = \det(E_1)\det(E_2)\cdots \det(E_r)\det(B)$$

Pero, por definición,

$$\det(A) = \det(E_1 E_2 \cdots E_r) = \det(E_1)\det(E_2)\cdots \det(E_r)$$

Por lo tanto, se sigue que

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

En el caso de matrices que no sean invertibles, se utiliza que las matrices invertibles son densas en el conjunto de todas las matrices y que la función determinante es continua, de modo que la propiedad se extiende a todo $\mathbb{R}^{n\times n}$.

Con eso se demuestra que, para cualquier par de matrices cuadradas A y B,

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

4.2 Sistemas de ecuaciones lineales

---

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

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

Paso 1. Despejamos cada variable

Reordenamos cada ecuación para expresar la variable correspondiente en función de las otras:

- De la primera ecuación:

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

- De la segunda:

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

- De la tercera:

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

Paso 2. Método iterativo

El método de Gauss-Seidel consiste en elegir una aproximación inicial y luego actualizar cada variable usando las fórmulas anteriores. Supongamos las condiciones iniciales:

$$x^{(0)} = 0, \quad y^{(0)} = 0, \quad z^{(0)} = 0.$$

Iteración 1

- $x^{(1)} = \dfrac{7 + y^{(0)} - z^{(0)}}{4} = \dfrac{7 + 0 - 0}{4} = 1.75$
- $y^{(1)} = \dfrac{1 + 2x^{(1)} + 2z^{(0)}}{4} = \dfrac{1 + 2(1.75) + 0}{4} = \dfrac{1 + 3.5}{4} = 1.125$
- $z^{(1)} = \dfrac{5 - x^{(1)} + y^{(1)}}{3} = \dfrac{5 - 1.75 + 1.125}{3} \approx \dfrac{4.375}{3} \approx 1.4583$

Iteración 2

Utilizando los valores de la iteración 1:

- $$x^{(2)} = \frac{7 + y^{(1)} - z^{(1)}}{4} = \frac{7 + 1.125 - 1.4583}{4} \approx \frac{6.6667}{4} \approx 1.6667$$
- $$y^{(2)} = \frac{1 + 2x^{(2)} + 2z^{(1)}}{4} = \frac{1 + 2(1.6667) + 2(1.4583)}{4}$$

Calculamos:
$2(1.6667) = 3.3334 y 2(1.4583) = 2.9166.$
Entonces,

- $$y^{(2)} = \frac{1 + 3.3334 + 2.9166}{4} = \frac{7.25}{4} \approx 1.8125$$
- $$z^{(2)} = \frac{5 - x^{(2)} + y^{(2)}}{3} = \frac{5 - 1.6667 + 1.8125}{3} \approx \frac{5.1458}{3} \approx 1.7153$$

Iteración 3

Para una mejor aproximación:

- $$x^{(3)} = \frac{7 + y^{(2)} - z^{(2)}}{4} \approx \frac{7 + 1.8125 - 1.7153}{4} \approx \frac{7.0972}{4} \approx 1.7743$$
- $$y^{(3)} = \frac{1 + 2x^{(3)} + 2z^{(2)}}{4} \approx \frac{1 + 2(1.7743) + 2(1.7153)}{4} \approx \frac{1 + 3.5486 + 3.4306}{4} \approx 1.9948$$
- $$z^{(3)} = \frac{5 - x^{(3)} + y^{(3)}}{3} \approx \frac{5 - 1.7743 + 1.9948}{3} \approx \frac{5.2205}{3} \approx 1.7402$$

Después de varias iteraciones, el método de Gauss-Seidel converge a una solución aproximada:

$$x \approx 1.77 \text{ a } 1.82, \quad y \approx 1.99 \text{ a } 2.03, \quad z \approx 1.74$$

La precisión va a depender del número de iteraciones y del criterio de tolerancia que se establezca.

4. Encuentre todas las soluciones del sistema homogéneo:

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

Si observamos bien, vemos que la segunda y la tercera ecuación son múltiplos de la primera:

- La segunda es $2(x + 2y + 3z) = 0$
- La tercera es $3(x + 2y + 3z) = 0$

Por lo tanto, el sistema se puede reducir a una única ecuación:

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

Entonces, podemos despejar, por ejemplo, x en función de y y z:

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

Así, la solución general del sistema es

$$(x, y, z) = (-2y - 3z,\ y,\ z) = y\,(-2,1,0) + z\,(-3,0,1)$$

donde y y z son parámetros libres que pueden tomar cualquier valor real.

El conjunto de soluciones del sistema homogéneo está dado por

$$\{\ (-2y - 3z,\ y,\ z) \mid y, z \in \mathbb{R}\ \} = \text{span}\{\ (-2,1,0),\ (-3,0,1)\}$$

Este conjunto describe un plano en $\mathbb{R}^3$.

4.3 Espacios vectoriales y auto-valores/auto-vectores

---

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

Vemos que:

$(2,4,6) = 2\,(1,2,3)$ y $(3,6,9) = 3\,(1,2,3)$

Esto implica que los tres vectores son colineales, por lo que solo aportan una dirección en $\mathbb{R}^3$

- Base: Una base para el subespacio generado es $\{(1,2,3)\}$
- Dimensión: La dimensión del subespacio es 1.

Se puede expresar formalmente como:

span$\{(1,2,3)\}$ con base $\{(1,2,3)\}$ y $\dim = 1$

---

6. Determine los autovalores y autovectores de la matriz:

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

Paso 1: Calcular los autovalores

Buscamos una λ tal que:

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

Calculamos:

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

Asi que el determinante es:

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

Igualamos a cero:

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

De donde se obtienen dos casos:

- Si $5 - \lambda = 2$, entonces$\lambda = 3$.
- Si $5 - \lambda = -2$, entonces $\lambda = 7$.

Por lo tanto, los autovalores son:

$$\lambda_1 = 3 \quad \text{y} \quad \lambda_2 = 7$$

Paso 2: Encontrar los autovectores

1. Para $\lambda = 3$:

Se resuelve:

$$(G - 3I)v = 0$$

Calculamos:

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

El sistema es:

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

De la primera ecuación se tiene $x - y = 0$ o $x = y$.
Un autovector asociado es, por ejemplo,

$$v_1 = (1,1)$$

2. Para $\lambda = 7$:

Se resuelve:

$$(G - 7I)v = 0$$

Calculamos:

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

El sistema es:

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

lo que equivale a $x + y = 0$ o $y = -x$.
Un autovector asociado es, por ejemplo,

$$v_2 = (1, -1)$$

Autovalores: $\lambda = 3$ y $\lambda = 7$.

- Autovectores:

Para $\lambda = 3$, un autovector es $v_1 = (1,1)$
Para $\lambda = 7$, un autovector es $v_2 = (1, -1)$

4.4 Aplicaciones en IA: reducción de dimensionalidad

---

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

El PCA es una técnica que usa el álgebra lineal para transformar un conjunto de datos con variables relacionadas en un nuevo conjunto de variables llamadas componentes principales. Estas son independientes entre sí y están ordenadas según la cantidad de varianza que capturan. Aquí te explico cómo funciona:

1. **Centralización**: Se quita la media de cada variable para centrar los datos en el origen.
2. **Matriz de covarianza**: Se calcula esta matriz para ver cómo varían juntas las variables.
3. **Autovalores y autovectores**: Se encuentran los autovalores y autovectores de la matriz de covarianza. Los autovectores muestran las direcciones con mayor varianza, y los autovalores miden esa varianza en cada dirección.
4. **Selección y proyección**: Se eligen los k autovectores (componentes principales) con los k mayores autovalores y se proyectan los datos originales sobre el espacio que ellos crean. Así, se reduce la cantidad de variables mientras se mantiene la mayor parte de la información (varianza).

---

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

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

El objetivo es expresarla de la forma

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

donde:

- U y V son matrices ortogonales,
- Σ es una matriz diagonal con los valores singulares de H.

Paso 1: Hay que calcular $H^T H$ y sus autovalores

Primero, se calcula

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

Para obtener los autovalores λ se resuelve

$$\det \left(H^T H - \lambda I \right) = \det \begin{pmatrix} 13 - \lambda & 7 \\ 7 & 5 - \lambda \end{pmatrix} = 0$$

Calculando el determinante:

$$(13 - \lambda)(5 - \lambda) - 7^2 = \lambda^2 - 18\lambda + 65 - 49 = \lambda^2 - 18\lambda + 16 = 0$$

La ecuación cuadrática

$$\lambda^2 - 18\lambda + 16 = 0$$

tiene soluciones

$$\lambda_{1,2} = 9 \pm \sqrt{65}$$

Paso 2: Valores singulares

Los valores singulares de H son las raíces cuadradas de los autovalores:

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

Paso 3: Determinar V y U

- Matriz V:

Los vectores propios normalizados de $H^T H$ forman las columnas de V. Para cada autovalor λ_i se resuelve

$$\left(H^T H - \lambda_i I\right) v_i = 0$$

Si bien los cálculos exactos pueden ser laboriosos, se obtiene $V = \begin{pmatrix} | & | \\ v_1 & v_2 \\ | & | \end{pmatrix}$

- Matriz U:

Se calcula usando la relación

$$[u_i = \frac{H, v_i}{\sigma_i}, \quad i = 1{,}2]$$

Así, U es la matriz cuyas columnas son los vectores u_1 y u_2 normalizados.

Entonces

La descomposición SVD de H se expresa como

$$H = U\,\Sigma\,V^T, \quad \text{con} \quad \Sigma = \begin{pmatrix} \sqrt{9+\sqrt{65}} & 0 \\ 0 & \sqrt{9-\sqrt{65}} \end{pmatrix}$$

Los detalles completos de U y V se obtienen resolviendo los sistemas asociados, de modo que se cumple $u_i = \dfrac{H\,v_i}{\sigma_i}$

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

El álgebra lineal es clave en el diseño y entrenamiento de redes neuronales profundas por varias razones:

- **Operaciones matriciales**: Las entradas, pesos y activaciones se representan con vectores y matrices. Las multiplicaciones matriciales nos ayudan a calcular las salidas de cada capa de manera eficiente, aprovechando la paralelización.

- **Backpropagation**: El algoritmo de retropropagación se basa en derivar funciones compuestas, usando gradientes y operaciones matriciales para ajustar los pesos del modelo.

- **Optimización**: Los métodos de optimización, como el gradiente descendente, necesitan calcular y actualizar parámetros en espacios de alta dimensión. El álgebra lineal facilita la representación y manejo de estos datos.

- **Procesamiento en hardware especializado**: Las GPUs y TPUs se benefician de operaciones vectorizadas y matriciales, mejorando el rendimiento en el entrenamiento de modelos de deep learning.

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

Los espacios vectoriales son una herramienta clave para representar datos de manera numérica, lo cual es esencial en muchas aplicaciones de inteligencia artificial. Aquí te explico cómo impactan:

- **Representación y embeddings**: Datos complejos como textos, imágenes o sonidos se convierten en vectores (por ejemplo, usando embeddings) que capturan sus características clave. Esto nos permite medir similitudes y relaciones entre datos con métricas geométricas, como distancias y ángulos.

- **Reducción de dimensionalidad**: Técnicas como PCA y SVD, basadas en la teoría de espacios vectoriales, nos ayudan a reducir el número de características (dimensiones) mientras mantenemos la información más importante. Esto facilita la visualización y mejora el rendimiento de algoritmos de clasificación o agrupamiento.

- **Transformaciones y proyecciones**: Los espacios vectoriales nos permiten realizar transformaciones lineales (rotaciones, escalados, proyecciones) que son cruciales para preprocesar datos, extraer características y adaptar los datos a la estructura que necesitan los modelos de IA.

- **Interpretabilidad**: La estructura geométrica de los espacios vectoriales nos ayuda a entender y visualizar la relación entre diferentes variables y a identificar patrones en los datos.