Apuntes Metnum

October 8, 2023

Gian

Índice

Elementos del álgebra lineal

1.1 Vectores

Un **vector es un conjunto ordenado de números reales**, que se pueden representar como una lista de números. Por ejemplo, el vector $v \in \mathbb{R}^n$ se puede representar como $v = (1, 2, 3)$.

1.1.1 Suma

Para **sumar dos vectores**, se suman las componentes correspondientes:

$$w = v + u \quad \text{con } w_i = v_i + u_i \quad \text{para } i = 1, 2, 3, ..., n$$

La suma de vectores es **conmutativa** y **asociativa**.

1.1.2 Multiplicación por escalares

Los vectores se pueden **multiplicar por escalares**: Sea $\alpha \in \mathbb{R}$ y $v \in \mathbb{R}^n$, entonces $\alpha \cdot v = (\alpha \cdot v_1, \alpha \cdot v_2, ..., \alpha \cdot v_n)$ para

1.1.3 Producto interno

El **producto interno** de dos vectores $v, u \in \mathbb{R}^n$ se define como:

$$v \cdot u = \sum_{i=1}^{n} v_i u_i = \|v\| \ \|u\| \cos \theta$$
$$\text{donde } \theta \ \text{ es el angulo entre } v \ \text{ y } u$$

Graficamente, el producto interno se puede interpretar como la **proyección de un vector sobre otro**.

1.1.4 Combinación Lineal

Una **combinación lineal** w de vectores $v_1, v_2, ..., v_n$ es un vector de la forma

$$w = \sum_{i=1}^{n} \alpha_i v_i \quad \text{con } \alpha_i \in \mathbb{R}^n$$

Decimos que $v_1, v_2, ..., v_n$ son **linealmente independientes** si la única combinación lineal que da el vector nulo es la trivial, es decir, si $\alpha_1 = \alpha_2 = ... = \alpha_n = 0$.

En cambio, si existe una combinación lineal no trivial (algún $\alpha_i \neq 0$) que da el vector nulo, entonces los vectores son **linealmente dependientes**.

Nombramos **espacio generado** por un conjunto de vectores $v_1, v_2, ..., v_n$ al conjunto de todas las combinaciones lineales de esos vectores:

$$\boldsymbol{S} = \left\{ \boldsymbol{x} \in \mathbb{R}^{\boldsymbol{n}} \ \ \textbf{tal que } \boldsymbol{x} = \sum_{\boldsymbol{i}=\mathbf{1}}^{\boldsymbol{n}} \boldsymbol{\alpha}_{\boldsymbol{i}} \boldsymbol{v}_{\boldsymbol{i}} \right\}$$

y su **dimensión** es la cantidad máxima de vectores linealmente independientes que lo generan.

1.1.5 Base vectorial

Una **base vectorial** $\boldsymbol{B}$ es un conjunto de vectores linealmente independientes que generan un espacio vectorial. En otras palabras, $\boldsymbol{B}$ es una base de $\boldsymbol{S}$ si todos los vectores $v \in S$ **se pueden escribir como una combinación lineal de los vectores de** B.

1.2 Matrices

Una matriz es un **arreglo rectangular de números reales**. Por ejemplo, la matriz $A \in \mathbb{R}^{m\times n}$ se puede representar como:

$$A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{i1} & a_{i2} & \dots & a_{\in} \\ \vdots & \vdots & \vdots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix}$$

1.2.1 Suma

Sean $A, B \in \mathbb{R}^{m\times n}$, entonces $A + B = C \in \mathbb{R}^{m\times n}$, donde $C_{ij} = A_{ij} + B_{ij}$.

$$A + B = \begin{bmatrix} a_{11} + b_{11} & a_{12} + b_{12} & \dots & a_{1n} + b_{1n} \\ a_{21} + b_{21} & a_{22} + b_{22} & \dots & a_{2n} + b_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{i1} + b_{i1} & a_{i2} + b_{i2} & \dots & a_{\in} + b_{in} \\ \vdots & \vdots & \vdots & \vdots \\ a_{m1} + b_{m1} & a_{m2} + b_{m2} & \dots & a_{mn} + b_{mn} \end{bmatrix}$$

La suma de matrices es **conmutativa** y **asociativa**.

Notar que para poder sumar dos matrices, **deben tener la misma dimensión**.

1.2.2 Producto por escalares

Sean $A \in \mathbb{R}^{m\times n}$ y $\alpha \in \mathbb{R}$, entonces $\alpha A = B \in \mathbb{R}^{m\times n}$, donde $B_{ij} = \alpha A_{ij}$.

$$\alpha A = \begin{bmatrix} \alpha a_{11} & \alpha a_{12} & \dots & \alpha a_{1n} \\ \alpha a_{21} & \alpha a_{22} & \dots & \alpha a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \alpha a_{i1} & \alpha a_{i2} & \dots & \alpha a_{\in} \\ \vdots & \vdots & \vdots & \vdots \\ \alpha a_{m1} & \alpha a_{m2} & \dots & \alpha a_{mn} \end{bmatrix}$$

1.2.3 Producto de matrices

Sean $A \in \mathbb{R}^{m\times n}$ y $B \in \mathbb{R}^{n\times p}$, entonces $AB = C \in \mathbb{R}^{m\times p}$, donde

$$C_{ij} = \sum_{k=1}^{n} a_{ik} b_{kj}$$

$$AB = \begin{bmatrix} \sum_{k=1}^{n} a_{1k} b_{k1} & \sum_{k=1}^{n} a_{1k} b_{k2} & \dots & \sum_{k=1}^{n} a_{1k} b_{kp} \\ \sum_{k=1}^{n} a_{2k} b_{k1} & \sum_{k=1}^{n} a_{2k} b_{k2} & \dots & \sum_{k=1}^{n} a_{2k} b_{kp} \\ \vdots & \vdots & \ddots & \vdots \\ \sum_{k=1}^{n} a_{ik} b_{k1} & \sum_{k=1}^{n} a_{ik} b_{k2} & \dots & \sum_{k=1}^{n} a_{ik} b_{kp} \\ \vdots & \vdots & \vdots & \vdots \\ \sum_{k=1}^{n} a_{mk} b_{k1} & \sum_{k=1}^{n} a_{mk} b_{k2} & \dots & \sum_{k=1}^{n} a_{mk} b_{kp} \end{bmatrix}$$

1.2.4 Rango de una matriz

El **rango de una matriz** $A \in \mathbb{R}^{m \times n}$ es la **cantidad máxima de columnas linealmente independientes** que tiene.

1.2.5 Determinante de una matriz

El **determinante** $\det(A)$ de una matriz $A \in \mathbb{R}^{n \times n}$ es un **número real** que se calcula como:

$$\det(A) = \sum_{i=1}^{n} (-1)^{i+j} a_{ij} \det(A_{ij}) \quad \text{para cualquier } j \in \{1, 2, ..., n\}$$

donde A_{ij} es la matriz que se obtiene de A al eliminar la fila i y la columna j.

Graficamente, **es el área del paralelogramo que forman las filas de** A. En un espacio, el determinante es el volumen del paralelepípedo correspondiente.

1.2.6 Espacio imagen

El **espacio imagen** de una matriz $A \in \mathbb{R}^{m \times n}$ es el conjunto de todos los vectores $b \in \mathbb{R}^m$ que se pueden escribir como $b = Ax$ para algún $x \in \mathbb{R}^n$.

$$Im(A) = \{b \in \mathbb{R}^m \ \text{ tal que } b = Ax \ \text{ para algún } x \in \mathbb{R}^n\}$$

Los vectores $b \in Im(A)$ son **combinaciones lineales de las columnas de** A.

1.2.7 Espacio nulo

El **espacio nulo** de una matriz $A \in \mathbb{R}^{m\times n}$ es el conjunto de todos los vectores $x \in \mathbb{R}^n$ tales que $Ax = 0$

$$Nu(A) = \{x \in \mathbb{R}^n \text{ tal que } Ax = 0\}$$

Propiedad

$$N(A) \neq \{0\} \Leftrightarrow \text{las columnas de A son linealmente dependientes}$$

1.2.8 Matriz inversa

Sea $A \in \mathbb{R}^{n\times n}$, entonces A es **inversible** si existe una matriz $A^{-1} \in \mathbb{R}^{n\times n}$ tal que

$$AA^{-1} = A^{-1}A = I$$

.

Propiedad

$$A \text{ es inversible} \Leftrightarrow \operatorname{rang}(A) = n \Leftrightarrow \det(A) \neq 0$$

Cuando A es inversible, decimos que A es una **matriz no singular**

Propiedad

La **inversa de un matriz diagonal** (si existe), es una **matriz diagonal**.

Propiedad

La inversa de un **matriz triangular superior** (si existe), es una **matriz triangular superior**.

Analogamente, la inversa de un **matriz triangular inferior** (si existe), es una **matriz triangular inferior**.

1.2.9 Matriz traspuesta

La **matriz traspuesta** de una matriz $A \in \mathbb{R}^{m \times n}$ es la matriz $A^T \in \mathbb{R}^{n \times m}$ tal que

$$A^T_{ij} = A_{ji}$$

.

Propiedad

$$(A^T)^T = A$$

$$(A + B)^T = A^T + B^T$$

$$(AB)^T = B^T A^T$$

$$(A^{-1})^T = (A^T)^{-1}$$

1.2.10 Submatriz principal

Una **submatriz principal** de una matriz $A \in \mathbb{R}^{m \times n}$ de orden k es una matriz $A^{(k)}$ que se obtiene de A al eliminar las ultimas $m - k$ filas y las ultimas $n - k$ columnas. Por ejemplo, si

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ a_{41} & a_{42} & a_{43} & a_{44} \end{bmatrix}$$

entonces:

$$A^{(2)} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \quad A^{(3)} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \quad A^{(4)} = \begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ a_{41} & a_{42} & a_{43} & a_{44} \end{bmatrix}$$

1.2.11 Matrices especiales

Matriz Identidad: La **matriz identidad** $I \in \mathbb{R}^{n \times n}$ es la matriz cuadrada que tiene 1 en la diagonal y 0 en el resto de las posiciones:

$$I = \begin{bmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{bmatrix}$$

Matriz diagonal: Una **matriz diagonal** $D \in \mathbb{R}^{n \times n}$ es una matriz cuadrada que tiene 0 en todas las posiciones excepto en la diagonal:

$$D = \begin{bmatrix} d_{11} & 0 & \dots & 0 \\ 0 & d_{22} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & d_{nn} \end{bmatrix}$$

Matriz triangular superior: Una **matriz triangular superior** $U \in \mathbb{R}^{n \times n}$ es una matriz cuadrada que tiene 0 en todas las posiciones por debajo de la diagonal:

$$U = \begin{bmatrix} u_{11} & u_{12} & \dots & u_{1n} \\ 0 & u_{22} & \dots & u_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & u_{nn} \end{bmatrix}$$

Matriz triangular inferior: Una **matriz triangular inferior** $L \in \mathbb{R}^{n \times n}$ es una matriz cuadrada que tiene 0 en todas las posiciones por encima de la diagonal:

$$L = \begin{bmatrix} l_{11} & 0 & \dots & 0 \\ l_{21} & l_{22} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ l_{n1} & l_{n2} & \dots & l_{nn} \end{bmatrix}$$

Propiedad

El **producto de dos matrices triangulares superiores** es una matriz **triangular superior**.

Analogamente, el **producto de dos matrices triangulares inferiores** es una **matriz triangular inferior**.

Matriz estrictamente diagonal dominante: Una matriz $A \in \mathbb{R}^{n \times n}$ es **estrictamente diagonal dominante** si para todo $i \in \{1, 2, ..., n\}$ se cumple que

$$|a_{ii}| > \sum_{\substack{j=1 \\ j \neq i}}^{n} |a_{ij}|$$

Matriz de permutación: Una **matriz de permutación** $P \in \mathbb{R}^{n \times n}$ es una matriz que se obtiene de la matriz identidad $I \in \mathbb{R}^{n \times n}$ al **intercambiar dos o más filas (o columnas)** de I.

Al multiplicar una matriz $A \in \mathbb{R}^{m \times n}$ por una matriz de permutación $P \in \mathbb{R}^{n \times n}$, se obtiene:

$$P = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix} \quad PA = \begin{bmatrix} — & \mathrm{fila}_2(A) & — \\ — & \mathrm{fila}_4(A) & — \\ — & \mathrm{fila}_3(A) & — \\ — & \mathrm{fila}_1(A) & — \end{bmatrix}$$

$$AP = \begin{bmatrix} | & | & | & | \\ \mathrm{col}_4(A) & \mathrm{col}_2(A) & \mathrm{col}_3(A) & \mathrm{col}_1(A) \\ | & | & | & | \end{bmatrix}$$

Matriz elemental (tipo 1): Una **matriz elemental (tipo 1)** es la matriz identidad con una **fila multiplicada por un escalar no nulo**:

$$E = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \alpha & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \qquad EA = \begin{bmatrix} - & \text{fila}_1(A) & - \\ - & \alpha\,\text{fila}_2(A) & - \\ - & \text{fila}_3(A) & - \\ - & \text{fila}_4(A) & - \end{bmatrix}$$

$$AE = \begin{bmatrix} | & | & | & | \\ \alpha\,\text{col}_2(A) & \text{col}_1(A) & \text{col}_3(A) & \text{col}_4(A) \\ | & | & | & | \end{bmatrix}$$

Matriz elemental (tipo 2): Una **matriz elemental (tipo 2)** es la matriz identidad con **un elemento no nulo fuera de la diagonal**:

$$E = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ \alpha & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \qquad EA = \begin{bmatrix} - & \text{fila}_1(A) & - \\ - & \text{fila}_2(A) & - \\ - & \text{fila}_3(A) + \alpha\ \text{fila}_1(A) & - \\ - & \text{fila}_4(A) & - \end{bmatrix}$$

$$AE = \begin{bmatrix} | & | & | & | \\ \text{col}_1(A) + \alpha\,\text{col}_3(A) & \text{col}_2(A) & \text{col}_3(A) & \text{col}_4(A) \\ | & | & | & | \end{bmatrix}$$

Sistema de ecuaciones lineales

2.1 Definición

Un sistema de ecuaciones lineales es un **conjunto de ecuaciones lineales que se deben cumplir simultáneamente**.

$$
\begin{gathered}
a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n = b_1 \\
a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n = b_2 \\
\vdots \\
a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n = b_m
\end{gathered}
$$

Podemos armar una matriz $A \in \mathbb{R}^{m \times n}$ con los coeficientes de las incógnitas, un vector $x \in \mathbb{R}^n$ con las incógnitas y un vector $b \in \mathbb{R}^m$ con los resultados de las ecuaciones:

$$
A = \begin{bmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \ldots & a_{mn} \end{bmatrix} \quad x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \quad b = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}
$$

Entonces, el sistema de ecuaciones lineales **se puede representar como una operación matricial**, en la que se busca encontrar el vector x que cumple

$$
Ax = b
$$

.

Si $b \notin Im(A)$ entonces el sistema **no tiene solución**.

Si $b \in Im(A)$ entonces:
- Si $\operatorname{rang}(A) = n$ entonces el sistema **tiene una única solución**.
- Si $\operatorname{rang}(A) < n$ entonces el sistema **tiene infinitas soluciones**.

Sistemas equivalentes: Sean $A, B \in \mathbb{R}^{n \times n}$, y $b, d \in \mathbb{R}^n$, entonces $Ax = b$ y $Bx = d$ son **sistemas equivalentes** si tienen el mismo conjunto de soluciones.

2.2 Resolución

2.2.1 Sistemas de ecuaciones diagonales

Sea $A \in \mathbb{R}^{n\times n}$ una matriz diagonal y $b \in \mathbb{R}^n$, entonces $Ax = b$ es un **sistema de ecuaciones diagonales** y se puede resolver despejando cada incógnita por separado:

$$\begin{aligned} a_{11}x_1 &= b_1 \\ a_{22}x_2 &= b_2 \\ &\vdots \\ a_{nn}x_n &= b_n \end{aligned}$$

- **Si $a_{ii} \neq 0$** para todo $i \in \{1, 2, ..., n\}$, el sistema **tiene única solución** y

$$x_i = \frac{b_i}{a_{ii}}$$

- **Si $a_{ii} = 0$** para algún $i \in \{1, 2, ..., n\}$:
  - Si $b_i = 0$ entonces el sistema **tiene infinitas soluciones**.
  - Si $b_i \neq 0$ entonces el sistema **no tiene solución**.

2.2.2 Sistemas de ecuaciones triangulares

Backward Substituion: Sea $A \in \mathbb{R}^{n\times n}$ una **matriz triangular superior** y $b \in \mathbb{R}^n$, entonces $Ax = b$ es un sistema de ecuaciones triangulares de la forma:

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + ... + a_{1n}x_n &= b_1 \\ a_{22}x_2 + ... + a_{2n}x_n &= b_2 \\ a_{33}x_3 + ... + a_{3n}x_n &= b_3 \\ &\vdots \\ a_{nn}x_n &= b_n \end{aligned}$$

- **Si $a_{ii} \neq 0$** para todo $i \in \{1, 2, ..., n\}$, el sistema **tiene una única solución** y se puede resolver usando **backward substitution**:

$$
\begin{aligned}
x_n &= \frac{b_n}{a_{nn}} \\
x_{n-1} &= \frac{b_{n-1} - a_{n-1n} x_n}{a_{n-1n-1}} \\
&\vdots \\
x_i &= \frac{b_i - \sum_{j=i+1}^{n} a_{ij} x_j}{a_{ii}} \\
&\vdots \\
x_1 &= \frac{b_1 - \sum_{j=2}^{n} a_{1j} x_j}{a_{11}}
\end{aligned}
$$

Esté método tiene **complejidad $\mathcal{O}(n^2)$**.

- **Si $a_{ii} = 0$** para algún $i \in \{1, 2, ..., n\}$:
  - Ejecutamos el algoritmo de backward substitution hasta llegar a la fila i. Osea que obtenemos los valores para $x_n, x_{n-1}, ..., x_{i+1}$.
  - Como todos estos valores son conocidos, simplemente hacemos la cuenta $\text{fila}_i(A)x$.
    - Si b_i es el valor que obtenemos, entonces el sistema **tiene infinitas soluciones**.
    - Si no lo es, entonces el sistema **no tiene solución**.

Forward Substituion: Sea $A \in \mathbb{R}^{n \times n}$ una **matriz triangular inferior** y $b \in \mathbb{R}^n$, entonces $Ax = b$ es un sistema de ecuaciones triangulares que se resuelve de forma similar al los sistemas triangulares superiores, la única diferencia es que se resuelve de arriba hacía abajo:

$$
\begin{aligned}
a_{11} x_1 &= b_1 \\
a_{21} x_1 + a_{22} x_2 &= b_2 \\
a_{31} x_1 + a_{32} x_2 + a_{33} x_3 &= b_3 \\
&\vdots \\
a_{n1} x_1 + a_{n2} x_2 + ... + a_{nn} x_n &= b_n
\end{aligned}
$$

2.3 Sistemas de ecuaciones generales

2.3.1 Eliminación gaussiana

Sea $A \in \mathbb{R}^{n \times n}$ y $b \in \mathbb{R}^n$, entonces usaremos el **método de eliminación gaussiana** para transformar el sistema original en un **sistema equivalente** que sea más fácil de resolver. En particular, vamos a transformar el sistema $Ax = b$ en uno de la forma $Ux = c$, donde U es una **matriz triangular superior**.

Para esto, se aplican las siguientes **operaciones elementales**:

- **Multiplicar una fila por un escalar no nulo** usando una matriz elemental (tipo 1).
- **Intercambiar dos filas** usando una matriz de permutación.
- **Sumar una fila multiplicada por un escalar no nulo a otra** fila usando una matriz elemental (tipo 2).

Debemos aplicar estas operaciones de forma tal que al final del proceso obtengamos un sistema de ecuaciones triangular superior. Para esto, vamos a aplicar el siguiente esquema:

ELIMINACIÓNGAUSSIANA(A: Matriz):

Para $i \leftarrow 1$ a n **hacer**
    Para $j \leftarrow i + 1$ a n **hacer**
        $m_{ij} = \frac{a_{ji}}{a_{ii}}$
        $\text{fila}_j(A) = \text{fila}_j(A) - m_{ij}\,\text{fila}_i(A)$
    Fin
Fin

La version mostrada **asume que** $a_{ii} \neq 0$ para todo $i \in \{1, 2, ..., n\}$, en todo momento.

Propiedad

El algoritmo propuesto tiene complejidad $\mathcal{O}(n^3)$.

2.3.2 Eliminación Gaussiana con pivoteo

Si en alguna iteración, nos encontramos con que $a_{ii} = 0$ entonces pueden darse dos posibles situaciones:

- $a_{ji} = 0$ para todo $j \in \{i + 1, i + 2, ..., n\}$: En este caso, la fila i es nula, y podemos **pasar a la siguiente iteración**.

- Existe algún $\boldsymbol{j} \in \{\boldsymbol{i}+\mathbf{1}, \boldsymbol{i}+\mathbf{2}, ..., \boldsymbol{n}\}$ tal que $a_{ji} \neq 0$: En este caso, **intercambiamos la fila i con la fila j**, y continuamos con el algoritmo.

En la práctica debemos tener en cuenta que **los números de coma flotante tienen precisión finita** por lo que debemos elegir de manera cuidadosa el pivote. Usaremos dos estrategias para esto:

- **Pivoteo parcial**: En cada iteración, elegimos entre las filas $\boldsymbol{i}, \boldsymbol{i}+\mathbf{1}, ..., \boldsymbol{n}$ aquella que tiene el mayor valor absoluto en la columna $\boldsymbol{i}$. Luego, intercambiamos la fila $\boldsymbol{i}$ con la fila elegida.
- **Pivoteo completo**: En cada iteración, buscamos la celda que tiene el mayor valor absoluto entre todas las filas $\boldsymbol{i}, \boldsymbol{i}+\mathbf{1}, ..., \boldsymbol{n}$ y todas las columnas $\boldsymbol{i}, \boldsymbol{i}+\mathbf{1}, ..., \boldsymbol{n}$. Luego, intercambiamos la fila $\boldsymbol{i}$ con la fila de la celda elegida, y la columna $\boldsymbol{i}$ con la columna de la celda elegida.

Factorización LU

3.1 Objetivo

Resolver varios sistemas de ecuaciones lineales con Eliminación Gaussiana tiene complejidad $\mathcal{O}(n^3)$ por cada uno. Existen técnicas de factorización de matrices que nos permiten mejorar esta complejidad.

La **factorización LU** de una matriz $A \in \mathbb{R}^{n \times n}$ es una factorización de la forma $A = LU$, donde L es una **matriz triangular inferior** y U es una **matriz triangular superior**.

Podemos usar esta factorización para resolver el sistema $Ax = b$ de la siguiente forma:

- Factorizamos $A = LU$: Entonces $Ax = b \Leftrightarrow LUx = b$.
- Definimos $y = Ux$: Entonces podemos resolver $Ly = b$ usando **forward substitution**.
- Luego, resolvemos el sistema $Ux = y$ usando **backward substitution**.

Osea que resolvemos dos sistemas triangulares, lo cual tiene **complejidad $\mathcal{O}(n^2)$ por cada uno**. Hay que ver cuanto nos cuesta factorizar A.

3.2 Método

Sea $A \in \mathbb{R}^{n \times n}$, supongamos que aplicamos eliminación gaussiana y se verifica que $a_{ii} \neq 0$ para todo $i \in \{1, 2, ..., n\}$.

Sea E la matriz elemental (tipo 2) que representa la **primer operación de la eliminación gaussiana**:

$$E = \begin{bmatrix} 1 & 0 & 0 & ... & 0 \\ -m_{21} & 1 & 0 & ... & 0 \\ 0 & 0 & 1 & ... & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & ... & 1 \end{bmatrix}$$

Vemos que $EA = A_1$ representa la operación

$$\text{fila}_2(A) = \text{fila}_2(A) - m_{21}\,\text{fila}_1(A)$$

. Podemos juntar todas las matrices elementales (tipo 2) que representan las operaciones que anulan todos los elementos por debajo de la diagonal, y obtener una matriz M_1:

$$M_1 = \begin{bmatrix} 1 & 0 & 0 & ... & 0 \\ -m_{21} & 1 & 0 & ... & 0 \\ -m_{31} & 0 & 1 & ... & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ -m_{n1} & 0 & 0 & ... & 1 \end{bmatrix}$$

.

Entonces $M^1 A = A^1$ donde A^1 es la matriz que se obtiene de A al aplicar el primer paso de eliminación gaussiana.

De la misma manera, podemos definir M^i **como la matriz que representa las operaciones que anulan todos los elementos por debajo de la diagonal** en la columna i:

$$M^i = \begin{bmatrix} 1 & ... & 0 & 0 & ... & 0 \\ \vdots & \ddots & \vdots & \vdots & & \\ 0 & ... & 1 & 0 & ... & 1 \\ 0 & ... & -m_{i+1i} & 1 & ... & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \\ 0 & ... & -m_{ni} & 0 & ... & 1 \end{bmatrix}$$

Entonces $M^i A^{i-1} = A^i$ donde A^i es la matriz que se obtiene de A al aplicar el paso i de eliminación gaussiana.

Luego, podemos pensar al proceso de eliminación gaussiana como una **secuencia de multiplicaciones de matrices**: Sea $U \in \mathbb{R}^{n \times n}$ la matriz triangular superior que se obtiene de aplicar eliminación gaussiana a A, entonces:

$$U = M^n M^{n-1} ... M^2 M^1 A$$

Propiedad

Sea $I \in \mathbb{R}^{n \times n}$ la matriz identidad, e_i el vector que tiene 1 en la posición i y 0 en el resto, y $m_i = [0, ..., m_{i+1i}, ..., m_{ni}]$, entonces:

$$M^i = I - m_i^t e_i$$

Propiedad

M^i es una matriz **triangular inferior** con 1 en la diagonal.

Propiedad

M^i es **inversible** y su inversa es

$$(M^i)^{-1} = I + m_i^t e_i$$

Demostración

$$\begin{aligned} M_i (M_i)^t &= (I - m_i^t e_i)(I + m_i^t e_i) \\ &= I + m_i^t e_i - m_i^t e_i - m_i^t e_i m_i^t e_i \\ &= I - m_i^t e_i m_i^t e_i \end{aligned}$$

Pero $e_i m_i^t = 0$ porque e_i tiene 0 en todas las posiciones salvo en la i y m_i^t tiene 0 en todas las posiciones hasta la posición i. Entonces $e_i m_i^t = 0$:

$$I - m_i^t e_i m_i^t e_i = I - m_i^t 0 e_i = I$$

Como habiamos dicho que $U = M^n M^{n-1} ... M^2 M^1 A$ y ahora sabemos que M^i es inversible para todo $i = \{1, ..., n\}$, entonces:

$$\boldsymbol{A} = (\boldsymbol{M}^{\mathbf{1}})^{-\mathbf{1}}(\boldsymbol{M}^{\mathbf{2}})^{-\mathbf{1}}...(\boldsymbol{M}^{\boldsymbol{n}})^{-\mathbf{1}}\boldsymbol{U}$$

Entonces, si definimos $\boldsymbol{L} = (\boldsymbol{M}^{\mathbf{1}})^{-\mathbf{1}}(\boldsymbol{M}^{\mathbf{2}})^{-\mathbf{1}}...(\boldsymbol{M}^{\boldsymbol{n}})^{-\mathbf{1}}$, tenemos que $\boldsymbol{A} = \boldsymbol{LU}$ obtenemos la factorización LU de $\boldsymbol{A}$ asociada a la eliminación gaussiana.

¡Cuidado!

La factorización LU no siempre existe. Si en algún paso de la eliminación gaussiana, nos encontramos con que $\boldsymbol{a_{ii}} = \mathbf{0}$ para algún $\boldsymbol{i} \in \{\mathbf{1}, \mathbf{2}, ..., \boldsymbol{n}\}$, entonces la factorización LU no existe.

3.3 Propiedades

Propiedad

Si $\boldsymbol{A} \in \mathbb{R}^{n \times n}$ es **no singular** y tiene **factorización LU**, entonces esa **factorización es única**.

Demostración

Supongamos que existen al menos dos factorizaciones LU de $\boldsymbol{A}$: $\boldsymbol{A} = \boldsymbol{L}_1 \boldsymbol{U}_1$ y $\boldsymbol{A} = \boldsymbol{L}_2 \boldsymbol{U}_2$.

Como $\boldsymbol{A}$ es inversible, entonces $\boldsymbol{U}_1$ y $\boldsymbol{U}_2$ son inversibles. Además, las inversas son tambien triangulares superiores. Tanto $\boldsymbol{L}_1$ como $\boldsymbol{L}_2$ son triangulares inferiores con $\boldsymbol{1}$ en la diagonal. Entonces, partiendo de las dos factorizaciones:

$$
\begin{aligned}
\boldsymbol{L}_1 \boldsymbol{U}_1 &= \boldsymbol{L}_2 \boldsymbol{U}_2 \\
\boldsymbol{L}_1^{-1} \boldsymbol{L}_1 \boldsymbol{U}_1 &= \boldsymbol{L}_1^{-1} \boldsymbol{L}_2 \boldsymbol{U}_2 \\
\boldsymbol{U}_1 &= \boldsymbol{L}_1^{-1} \boldsymbol{L}_2 \boldsymbol{U}_2 \\
\boldsymbol{U}_1 \boldsymbol{U}_2^{-1} &= \boldsymbol{L}_1^{-1} \boldsymbol{L}_2 \boldsymbol{U}_2 \boldsymbol{U}_2^{-1} \\
\boldsymbol{U}_1 \boldsymbol{U}_2^{-1} &= \boldsymbol{L}_1^{-1} \boldsymbol{L}_2
\end{aligned}
$$

$\boldsymbol{U}_1 \boldsymbol{U}_2^{-1}$ es una matriz triangular superior por ser producto de dos matrices triangulares superiores. $\boldsymbol{L}_1^{-1} \boldsymbol{L}_2$ es una matriz triangular inferior por ser producto de dos matrices triangulares inferiores.

Como $\boldsymbol{D} = \boldsymbol{U}_1 \boldsymbol{U}_2^{-1} = \boldsymbol{L}_1^{-1} \boldsymbol{L}_2$ entonces $\boldsymbol{D}$ necesariamente tiene que ser una matriz diagonal.

Tambien sabemos que $\boldsymbol{L}_1^{-1} \boldsymbol{L}_2$ tiene $\boldsymbol{1}$ en la diagonal. Entonces $\boldsymbol{D} = \boldsymbol{I}$.

Luego, $\boldsymbol{U}_1 \boldsymbol{U}_2^{-1} = \boldsymbol{I}$ y $\boldsymbol{U}_1 = \boldsymbol{U}_2$. Entonces $\boldsymbol{L}_1 = \boldsymbol{L}_2$. ■

Propiedad

Sea $\boldsymbol{A} \in \mathbb{R}^{n \times n}$ no singular.

$$
\begin{gathered}
\boldsymbol{A} \ \text{ tiene factorización } \boldsymbol{LU} \\
\Leftrightarrow \\
\text{todas sus matrices principales son no singulares}
\end{gathered}
$$

Demostración

$\Rightarrow$) Si $\boldsymbol{A}$ es no signular y tiene factorización LU, tanto $\boldsymbol{L}$ como $\boldsymbol{U}$ son no singulares. Los elementos de la diagona de $\boldsymbol{L}$ son todos $\mathbf{1}$ y los elementos de la diagonal de $\boldsymbol{U}$ son todos no nulos. Las submatrices principales de $\boldsymbol{L}$ son tambien triangulares con $\mathbf{1}$ en la diagonal, por lo tanto son no singulares. Las submatrices principales de $\boldsymbol{U}$ son triangulares superiores con elementos no nulos en la diagonal, por lo tanto son no singulares.

Como $\boldsymbol{A} = \boldsymbol{LU}$, entonces las submatriz de orden $\boldsymbol{k}$ de $\boldsymbol{A}$ es el resultado del producto de la submatriz de orden $\boldsymbol{k}$ de $\boldsymbol{L}$ y la submatriz de orden $\boldsymbol{k}$ de $\boldsymbol{U}$. Como ambas son no singulares, entonces la submatriz de orden $\boldsymbol{k}$ de $\boldsymbol{A}$ tambien es no singular.

$\Leftarrow$) Demostramos por inducción en la dimensión de la matriz $\boldsymbol{A}$.

- **Caso base**: $\boldsymbol{n} = \mathbf{2}$

  Como $\boldsymbol{a}_{\mathbf{11}}$ no es nulo por ser la submatriz principal de orden $\mathbf{1}$ de $\boldsymbol{A}$, entonces el primer (y único) paso de la eliminación Gaussiana se puede realizar sin inconvenientes y se encuentra la factorización LU.

- **Paso inductivo**: Supongamos que vale para matrices de imensión2 $\boldsymbol{n}$ y veamos que vale para matrices de orden $\boldsymbol{n} + \mathbf{1}$.

  Consideremos $\boldsymbol{A} \in \mathbb{R}^{(\boldsymbol{n}+\mathbf{1})\times(\boldsymbol{n}+\mathbf{1})}$ con todas sus submatrices principales no singulares. Veamos que $\boldsymbol{A}$ tiene factorización LU.

  Si escribimos $\boldsymbol{A}$ como:

  $$\boldsymbol{A} = \begin{bmatrix} \boldsymbol{A}^{(\boldsymbol{n})} & \boldsymbol{c}_{\boldsymbol{n}+\mathbf{1}} \\ \boldsymbol{f}^{\boldsymbol{t}}_{\boldsymbol{n}+\mathbf{11}} & \boldsymbol{a}_{\boldsymbol{n}+\mathbf{1}\boldsymbol{n}+\mathbf{1}} \end{bmatrix}$$

  donde $\boldsymbol{A}^{(\boldsymbol{n})} \in \mathbb{R}^{\boldsymbol{n}\times\boldsymbol{n}}$, $\boldsymbol{c}_{\boldsymbol{n}+\mathbf{1}} \in \mathbb{R}^{\boldsymbol{n}}$, $\boldsymbol{f}_{\boldsymbol{n}+\mathbf{11}} \in \mathbb{R}^{\boldsymbol{n}}$ y $\boldsymbol{a}_{\boldsymbol{n}+\mathbf{1}\boldsymbol{n}+\mathbf{1}} \in \mathbb{R}$.

  Como todas las submatrices principales de $\boldsymbol{A}$ son no singulares, entonces $\boldsymbol{A}^{(\boldsymbol{n})}$ y todas sus submatrices principales son no singulares. Entonces, por hipótesis inductiva, $\boldsymbol{A}^{(\boldsymbol{n})}$ tiene factorización LU. Sea $\boldsymbol{A}^{(\boldsymbol{n})} = \boldsymbol{L}^{(\boldsymbol{n})}\boldsymbol{U}^{(\boldsymbol{n})}$.

  Propongamos una factorización LU para $\boldsymbol{A}$, se resaltan los valores que necesitamos calcular:

$$A = \begin{bmatrix} A^{(n)} & c_{n+1} \\ f_{n+1}^t & a_{n+1,n+1} \end{bmatrix} = \begin{bmatrix} L^{(n)} & 0 \\ l_{n+11}^t & 1 \end{bmatrix} \begin{bmatrix} U^{(n)} & u_{n+1} \\ 0 & u_{n+1,n+1} \end{bmatrix}$$

Realizando el producto en bloques, tenemos que verificar que:

1. $A^{(n)} = L^{(n)}U^{(n)}$

   Se pueden calcular usando eliminación gaussiana ya que es la factorización LU de $A^{(n)}$.

2. $c_{n+1} = L^{(n)}u_{n+1}$

   Como $L^{(n)}$ es no signular, entonces este sistema tiene solución y es única, por lo que es posible determinar u_{n+1}.

3. $f_{n+1}^t = l_{n+1}^t U^{(n)}$

   La matriz $U^{(n)}$ es no singular ya que $A^{(n)}$ es no singular, por lo tanto el tercer sistema tambien tiene solución y es única, por lo que es posible determinar l_{n+1}^t.

4. $a_{n+1n+1} = l_{n+11}^t u_{n+1} + u_{n+1,n+1}$

   Como l_{n+11}^t y u_{n+1} son conocidos, entonces es posible determinar unívocamente $u_{n+1,n+1}$.

Concluimos entonces que A tiene factorización LU, ya que todos los sistemas propuestos tienen solución y es única. ■

Propiedad

Si $A \in \mathbb{R}^{n \times n}$ es **estrictamente diagonal dominante**, entonces A **tiene factorización LU**.

Demostración

Vamos a demostrar que A es no singular y que todas sus submatrices principales son no singulares. De esta manera, estamos en condiciones de aplicar la propiedad anterior y concluir que A tiene factorización LU.

- **A es no singular**: Supongamos que A es singular, entonces existe $x \in \mathbb{R}^n$, $x \neq 0$ tal que $Ax = 0$.

Como $\boldsymbol{x} \neq \mathbf{0}$, entonces tiene una coordenada con máximo valor absoluto, o dicho de otra manera existe $\boldsymbol{k_0} \in \{\mathbf{1}, \mathbf{2}, ..., \boldsymbol{n}\}$ tal que

$$\left|\boldsymbol{x_{k_0}}\right| = \max_{\boldsymbol{j}=\mathbf{1}...\boldsymbol{n}} \left|\boldsymbol{x_j}\right|$$

con $\left|\boldsymbol{x_{k_0}}\right| \neq \mathbf{0}$.

Consideremos la ecuación $\boldsymbol{k_0}$ del sistema $\boldsymbol{Ax} = \mathbf{0}$:

$$\sum_{j=1}^{n} a_{k_0 j} x_j = 0$$

Separamos el término $\boldsymbol{k_0}$:

$$\sum_{\substack{j=1 \\ j \neq k_0}}^{n} a_{k_0 j} x_j + a_{k_0 k_0} x_{k_0} = 0$$

Pasamos restando:

$$\sum_{\substack{j=1 \\ j \neq k_0}}^{n} a_{k_0 j} x_j = -a_{k_0 k_0} x_{k_0}$$

Tomamos valor absoluto, como $\left|-a_{k_0 k_0} x_{k_0}\right| = \left|a_{k_0 k_0}\right|\left|x_{k_0}\right|$:

$$\left|\sum_{\substack{j=1 \\ j \neq k_0}}^{n} a_{k_0 j} x_j\right| = \left|a_{k_0 k_0}\right|\left|x_{k_0}\right|$$

Aplicamos desigualdad triangular al lado izquierdo de la equación:

$$\sum_{\substack{j=1 \\ j \neq k_0}}^{n} \left|a_{k_0 j} x_j\right| \geq \left|\sum_{\substack{j=1 \\ j \neq k_0}}^{n} a_{k_0 j} x_j\right| = \left|a_{k_0 k_0}\right|\left|x_{k_0}\right|$$

Entonces:

$$\sum_{\substack{j=1 \\ j \neq k_0}}^{n} \left|a_{k_0 j} x_j\right| \geq \left|a_{k_0 k_0}\right| \left|x_{k_0}\right|$$

Pasamos $\left|x_{k_0}\right|$ dividiendo:

$$\sum_{\substack{j=1 \\ j \neq k_0}}^{n} \frac{\left|a_{k_0 j}\right| \left|x_j\right|}{\left|x_{k_0}\right|} \geq \left|a_{k_0 k_0}\right|$$

Pero $\left|x_j\right| \leq \left|x_{k_0}\right|$ para todo $j \in \{1, 2, ..., n\}$, entonces $\frac{\left|x_j\right|}{\left|x_{k_0}\right|} \leq 1$:

$$\sum_{\substack{j=1 \\ j \neq k_0}}^{n} \left|a_{k_0 j}\right| \left|x_j\right| \geq \sum_{\substack{j=1 \\ j \neq k_0}}^{n} \frac{\left|a_{k_0 j}\right| \ \left|x_j\right|}{\left|x_{k_0}\right|} \geq \left|a_{k_0 k_0}\right|$$

Entonces:

$$\sum_{\substack{j=1 \\ j \neq k_0}}^{n} \left|a_{k_0 j}\right| \left|x_j\right| \geq \left|a_{k_0 k_0}\right|$$

Pero A es estrictamente diagonal dominante, entonces $\left|a_{k_0 k_0}\right| > \sum_{\substack{j=1 \\ j \neq k_0}}^{n} \left|a_{k_0 j}\right|$ por lo que llegamos a una contradicción. Entonces A es no singular. ■

Demostración alternativa

Vamos a desmostrar que es posible realizar el primer paso de la eliminación Gaussiana y que la matriz conformada por las filas 2 a n y columnas 2 a n es estrictamente diagonal dominante. De esta manera, podremos afirmar que la eliminación Gaussiana se puede aplicar sin inconvenientes y por lo tanto existe la factorización LU.

- **Primer paso de la eliminación Gaussiana**: Como A es estrictamente diagonal dominante, entonces podemos afirmar que $a_{11} \neq 0$. Entonces, el primer paso de la eliminación Gaussiana es:

$$\tilde{F}_i = F_i - \frac{a_{i1}}{a_{11}} F_1 \rightarrow \tilde{a}_{ij} = a_{ij} - \frac{a_{i1}}{a_{11}} a_{1j}$$

- **La parte de la matriz que queda por triangular es estrictamente diagonal dominante**: Tenemos que ver que

$$\tilde{a}_{ii} \geq \sum_{\substack{j=2 \\ j \neq i}}^{n} |\tilde{a}_{ij}| \quad \text{para todo } i \in \{2, 3, ..., n\}$$

.

Analicemos el término de la sumatoria:

$$\sum_{\substack{j=2 \\ j \neq i}}^{n} |\tilde{a}_{ij}| = \sum_{\substack{j=2 \\ j \neq i}}^{n} \left| a_{ij} - \frac{a_{i1}}{a_{11}} a_{1j} \right|$$

Aplicando la desigualdad triangular, tenemos:

$$\sum_{\substack{j=2 \\ j \neq i}}^{n} |\tilde{a}_{ij}| = \sum_{\substack{j=2 \\ j \neq i}}^{n} \left| a_{ij} - \frac{a_{i1}}{a_{11}} a_{1j} \right| \leq \sum_{\substack{j=2 \\ j \neq i}}^{n} \left(|a_{ij}| + \left| \frac{a_{i1}}{a_{11}} a_{1j} \right| \right)$$

$$\sum_{\substack{j=2 \\ j \neq i}}^{n} |\tilde{a}_{ij}| \leq \sum_{\substack{j=2 \\ j \neq i}}^{n} \left(|a_{ij}| + \left| \frac{a_{i1}}{a_{11}} a_{1j} \right| \right)$$

$$\sum_{\substack{j=2 \\ j \neq i}}^{n} |\tilde{a}_{ij}| \leq \sum_{\substack{j=2 \\ j \neq i}}^{n} \left(|a_{ij}| + \left| \frac{a_{i1}}{a_{11}} a_{1j} \right| \right) = \sum_{\substack{j=2 \\ j \neq i}}^{n} |a_{ij}| + \left| \frac{a_{i1}}{a_{11}} \right| \sum_{\substack{j=2 \\ j \neq i}}^{n} |a_{1j}|$$

$$\sum_{\substack{j=2 \\ j \neq i}}^{n} |\tilde{a}_{ij}| \leq \sum_{\substack{j=2 \\ j \neq i}}^{n} |a_{ij}| + \left| \frac{a_{i1}}{a_{11}} \right| \sum_{\substack{j=2 \\ j \neq i}}^{n} a_{1j}$$

Como A es estrictamente diagonal dominante, entonces

$$|a_{ii}| > \sum_{\substack{j=1 \\ j \neq i}}^{n} |a_{ij}| \Rightarrow |a_{ii}| - |a_{i1}| > \sum_{\substack{j=2 \\ j \neq i}}^{n} |a_{1j}|$$

Entonces, remplazamos en la desigualdad anterior:

$$\sum_{\substack{j=2\\j\neq i}}^{n}|a_{ij}| + \left|\frac{a_{i1}}{a_{11}}\right|\sum_{\substack{j=2\\j\neq i}}^{n} a_{1j} < |a_{ii}| - |a_{i1}| + \left|\frac{a_{i1}}{a_{11}}\right|(|a_{11}| - |a_{1i}|)$$

$$\sum_{\substack{j=2\\j\neq i}}^{n}|\tilde{a}_{ij}| < |a_{ii}| - |a_{i1}| + \left|\frac{a_{i1}}{a_{11}}\right|(|a_{11}| - |a_{1i}|)$$

$$\sum_{\substack{j=2\\j\neq i}}^{n}|\tilde{a}_{ij}| < |a_{ii}| - |a_{i1}| + \left|\frac{a_{i1}}{a_{11}}\right||a_{11}| - \left|\frac{a_{i1}}{a_{11}}\right||a_{1i}|$$

$$\sum_{\substack{j=2\\j\neq i}}^{n}|\tilde{a}_{ij}| < |a_{ii}| - |a_{i1}| + |a_{i1}| - \left|\frac{a_{i1}}{a_{11}}\right||a_{1i}|$$

$$\sum_{\substack{j=2\\j\neq i}}^{n}|\tilde{a}_{ij}| < |a_{ii}| - \left|\frac{a_{i1}}{a_{11}}\right||a_{1i}|$$

$$\sum_{\substack{j=2\\j\neq i}}^{n}|\tilde{a}_{ij}| \leq \left|a_{ii} - \frac{a_{i1}}{a_{11}}a_{1i}\right| = |\tilde{a}_{ii}|$$

Concluimos entonces que la matriz conformada por las filas 2 a n y columnas 2 a n que resulta del primer paso de eliminación Gaussiana es estrictamente diagonal dominante por lo que exist efactorización LU. ■

3.4 Factorización PLU

En caso de que la factorización LU no exista, podemos usar **pivoteo parcial** para obtener una factorización PLU que es una factorización LU la **matriz original con sus filas permutadas**:

$$\boldsymbol{PA = LU}$$

Propiedad

Toda matriz $\boldsymbol{A} \in \mathbb{R}^{\boldsymbol{n \times n}}$ **tiene factorización PLU**.

Demostración

Si aplicamos eliminación Gaussiana con pivoteo parcial, aplicando permutaciones cuando sea necesario por la presencia de elementos nulos en la diagonal durante el proceso, se obtiene el siguiente producto de matrices:

$$\boldsymbol{M^{n-1}P^{n-1}M^{n-2}P^{n-2}...M^{i}P^{i}...M^{2}P^{2}M^{1}P^{1}A = U}$$

donde $\boldsymbol{M^i = I - m_i^t e_i}$ y $\boldsymbol{P^i}$ es una matriz de permutación que indican los intercambios realziados entre las filas.

Tenemos que encontrar una forma de llegar desde esta ecuación hasta una ecuación de la forma $\boldsymbol{PA = LU}$

Como cada $\boldsymbol{P^i}$ es una matriz de permutación entre filas, entonces $\boldsymbol{P^i}$ es no singular y su inversa es ella misma. Osea que $\boldsymbol{P^iP^i = I}$. Podemos agregar entonces los siguientes terminos a la ecuación, sin modificar su resultado:

$$\boldsymbol{M^{n-1}P^{n-1}M^{n-2}P^{n-1}P^{n-1}P^{n-2}...P^{i+2}...P^{n-1}P^{n-1}...P^{i+1}P^{i}}$$
$$\boldsymbol{...M^{2}P^{3}...P^{n-1}P^{n-1}...P^{3}P^{2}M^{1}P^{2}...P^{n-1}P^{n-1}...P^{2}P^{1}A = U}$$

Notemos ahora $\boldsymbol{\widetilde{M}^i = (P^{n-1}...P^{i+2}P^{i+1}M^{i(P^{i+1}...P^{n-1})}}$, entonces tenemos que:

$$\boldsymbol{M^{n-1}\widetilde{M}^{n-2}...\widetilde{M}^{i}...\widetilde{M}^{2}\widetilde{M}^{1}(P^{n-1}...P^{2}P^{1})A = U}$$

Veamos que estructura tiene $\boldsymbol{\widetilde{M}^i}$:

$$
\begin{aligned}
\widetilde{M}^i &= \left(P^{n-1}...P^{i+2}P^{i+1}\right)\left(I - m_i^t e_i\right)\left(P^{i+1}...P^{n-1}\right) \\
&= \left(P^{n-1}...P^{i+2}P^{i+1}\right)I\left(P^{i+1}...P^{n-1}\right) \\
&\quad -\left(P^{n-1}...P^{i+2}P^{i+1}\right)\left(m_i^t e_i\right)\left(P^{i+1}...P^{n-1}\right) \\
&= I - \left(P^{n-1}...P^{i+2}P^{i+1}\right)\left(m_i^t e_i\right)\left(P^{i+1}...P^{n-1}\right)
\end{aligned}
$$

Como $P^{i+1}...P^{n-1}$ son matrices de permutación que realizan intercambiamos entre las filas $i+1$ a n, entonces $e_i\left(P^{i+1}...P^{n-1}\right) = e_i$.

Además nombremos $\widetilde{m}_i = \left(P^{n-1}...P^{i+2}P^{i+1}\right)m_i^t$ al entonces tenemos que:

$$
\widetilde{M}^i = I - \widetilde{m}_i^t e_i
$$

Entonces, vemos que $\widetilde{M}^i$ son matrices triangules inferiores con 1 en la diagonal. Además, como $P^{n-1}...P^{i+2}P^{i+1}$ son matrices de permutación, entonces $\widetilde{M}^i$ son no singulares. Por lo que, podemos escribir:

$$
\left(P^{n-1}...P^2P^1\right)A = \left(\widetilde{M}^1\right)^{-1}\left(\widetilde{M}^2\right)^{-1}...\left(\widetilde{M}^i\right)^{-1}U
$$

Definimos $L = \left(\widetilde{M}^1\right)^{-1}\left(\widetilde{M}^2\right)^{-1}...\left(\widetilde{M}^i\right)^{-1}$ y $P = P^{n-1}...P^2P^1$, entonces:

$$
PA = LU\blacksquare
$$

Normas vectoriales y matriciales

4.1 Normas vectoriales

Una **norma vectorial** es una función $f : \mathbb{R}^n \to \mathbb{R}$ que cumple las siguientes propiedades:

- $f(x) > 0$ para todo $x \neq 0 \in \mathbb{R}^n$
- $f(x) = 0 \Leftrightarrow x = 0$
- $f(\alpha x) = |\alpha|\ f(x)$ para todo $\alpha \in \mathbb{R}$ y $x \in \mathbb{R}^n$
- $f(x + y) \leq f(x) + f(y)$ para todo $x, y \in \mathbb{R}^n$

Norma 1 (norma Manhattan):

$$\|x\|_1 = \sum_{i=1}^{n} |x_i|$$

Norma 2 (norma Euclídea):

$$\|x\|_2 = \sqrt{\sum_{i=1}^{n} x_i^2}$$

Norma p:

$$\|x\|_p = \left(\sum_{i=1}^{n} |x_i|^p\right)^{\frac{1}{p}}$$

Norma infinito:

$$\|x\|_{\text{inf}} = \max_{\{1 \leq i \leq n\}} |x_i|$$

4.2 Normas matriciales

Una **norma matricial** es una función $f : \mathbb{R}^{m\times n} \to \mathbb{R}$ que cumple las siguientes propiedades:

- $f(A) > 0$ para todo $A \neq 0 \in \mathbb{R}^{m\times n}$
- $f(A) = 0 \Leftrightarrow A = 0$
- $f(\alpha A) = |\alpha|\ f(A)$ para todo $\alpha \in \mathbb{R}$ y $A \in \mathbb{R}^{m\times n}$
- $f(A+B) \leq f(A) + f(B)$ para todo $A, B \in \mathbb{R}^{m\times n}$

Adicionalmente, si f cumple que $f(AB) \leq f(A)f(B)$ para todo $A \in \mathbb{R}^{m\times n}$ y $B \in \mathbb{R}^{n\times p}$ entonces diremos que f es una **norma submultiplicativa**.

Norma de Frobenius:

$$\|A\|_F = \sqrt{\sum_{i=1}^{m}\sum_{j=1}^{n} a_{ij}^2}$$

Normas matficiales inducidas: Sean f_1 una norma vctiral definida en $\mathbb{R}^m$ y f_2 una norma vectorial definida en $\mathbb{R}^n$, entonces la función $F : \mathbb{R}^{m\times n}$ es una **norma inducida** si:

$$F(A) = \max_{x\neq 0} \frac{f_1(Ax)}{f_2(x)} = \max_{x:f_2(x)=1} f_1(Ax)$$

Número de condición: Sea $A \in \mathbb{R}^{n\times n}$ una matriz no singular y $\|.\|$ una norma matricial. Se define el **número de condición** de A como:

$$\kappa(A) = \|A\|\ \|A^{-1}\|$$

Propiedad

Si $\|.\|$ es una norma matricial inducida, entonces $\kappa(I) = 1$.

Propiedad

Si $\|.\|$ es una norma submultiplicativa, entonces $\kappa(A) \geq 1$

4.2.2 Cota del error

Si $\boldsymbol{A} \in \mathbb{R}^{\boldsymbol{n}\times\boldsymbol{n}}$ matriz no singular y $\|.\|$ una norma matricial inducida. Sea $\tilde{\boldsymbol{x}}$ una solución aproximada de $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$ con $\boldsymbol{b} \neq \boldsymbol{0}$ y sea $\boldsymbol{A}\tilde{\boldsymbol{x}} = \tilde{\boldsymbol{b}}$ entonces:

$$\frac{\| \boldsymbol{x} - \tilde{\boldsymbol{x}} \|}{\| \boldsymbol{x} \|} \leq \| \boldsymbol{A}^{-}\boldsymbol{1} \| \frac{\| \boldsymbol{b} - \tilde{\boldsymbol{b}} \|}{\| \boldsymbol{b} \|}$$

Factorización de Cholesky

5.1 Matrices Simétricas Definidas Positivas

Sea $A \in R^{\{n \times n\}}$ una matriz cuadrada, se dice que es **simétrica definida positiva (SDP)** si:

- $A = A^T$ (es simétrica)
- $x^T A x > 0$ para todo $x \in R^n$ con $x \neq 0$ (es definida positiva)

Propiedad

Si A es SDP, entonces **A es no singular**.

Demostración

Supongamos que A es SDP y singular. Entonces existe $x \neq 0 \in \mathbb{R}^n$ tal que $Ax = 0$. Entonces $x^T A x = 0$ lo cual es absurdo pues contradice la definición de SDP ■

Propiedad

Si A es SDP, entonces $a_{ii} > 0$ para todo $i = 1, 2, ..., n$.

Demostración

Sea e_i el i —ésimo vector canónico. Entonces $e_i^T A e_i = a_{ii} > 0$ pues A es SDP ■

Propiedad

Si A es SDP, **todas sus submatrices principales son SDP**.

Demostración

Sea $A^{(k)}$ la submatriz principal de A de orden k. Tenemos que ver que cumple con las dos condiciones de SDP:

- $A^{(k)} = \left(A^{(k)}\right)^T$

$$a_{ij}^{(k)} = a_{ij} = a_{ji} = \left(a_{ji}^{(k)}\right)$$

- $x^T A^{(k)} x > 0$ para todo $x \in R^k$ con $x \neq 0$

  Supongamos que esto no sucede, entonces existe $\bar{x} \neq 0 \in R^k$ tal que $\bar{x}^T A^{(k)} \bar{x} = 0$. Armemos un vector $x \in R^n$ tal que $x = (\bar{x}, 0, ..., 0)$. Entonces:

$$x^T A x = [\bar{x}^t \;\; 0 \;\; ... \;\; 0] \begin{bmatrix} A^{(k)} & * & ... & * \\ * & * & ... & * \\ \vdots & \vdots & ... & \vdots \\ * & * & ... & * \end{bmatrix} \begin{bmatrix} \bar{x} \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

$$= [\bar{x}^t \;\; 0 \;\; ... \;\; 0] \begin{bmatrix} A^{(k)} \bar{x} \\ * \\ \vdots \\ * \end{bmatrix} = x^t A^{(k)} x \leq 0$$

  Entonces $x^t A x \leq 0$, lo cual es absurdo pued A era una matriz SDP.

Entonces $A^{(k)}$ cumple con ambas condiciones y es SDP ■

Propiedad

A es SDP $\Leftrightarrow$ **$\forall B \in R^{n \times n}$ no singular vale que $B^T A B$ es SDP**.

Demostración

Supongamos que A es SDP. Entonces tenemos que ver que $B^T A B$ cumple las condiciones de SDP:

- $B^T A B = (B^T A B)^T$

$$(B^T A B)^t = ((B^T A) B)^T = B^T (B^T A)^T = B^T A^T B$$

  Como $A^T = A$, queda:

$$(B^T A B)^t = B^T A^T B = B^T A B$$

- $x^T B^T A B x > 0$ para todo $x \in R^n$ con $x \neq 0$

  Sea $x \neq 0 \in \mathbb{R}^n$, entonces:

$$x^T B^T A B x = (Bx)^T A (Bx)$$

Si nombramos $y = Bx$, entonces $y \neq 0$ pues B es no singular y $x \neq 0$. Entonces resulta $x^T B^T ABx = y^T Ay > 0$ pues A es SDP.

Luego $B^T AB$ es SDP ■

Propiedad

Si A es SDP, entonces la submatriz conformada por las filas 2 a n y las columnas 2 a n despues del primer paso de la eliminación gaussiana es SDP.

Demostración

Sea M_1 la matriz asociada al primer paso de la eliminación gaussiana y $\tilde{A} \in \mathbb{R}^{n-1 \times n-1}$ conformada por las filas 2 a n y las columnas 2 a n de $M_1 A$.

Realicemos el producto $M_1 A M_1^T$:

$$M_1 A M_1^T = \begin{bmatrix} 1 & 0 & \dots & 0 \\ -\frac{a_{21}}{a_{11}} & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ -\frac{a_{n1}}{a_{11}} & 0 & \dots & 1 \end{bmatrix} A \begin{bmatrix} 1 & -\frac{a_{21}}{a_{11}} & \dots & -\frac{a_{n1}}{a_{11}} \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{bmatrix}$$

$$= \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ 0 & & \tilde{A} & \end{bmatrix} \begin{bmatrix} 1 & -\frac{a_{21}}{a_{11}} & \dots & -\frac{a_{n1}}{a_{11}} \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{bmatrix}$$

$$= \begin{bmatrix} a_{11} & 0 \\ 0 & \tilde{A} \end{bmatrix}$$

Por la propiedad anterior, podemos afirmar que $M_1 A M_1^T$ es SDP. Entonces $\tilde{A}$ es SDP

Corolario

Si A es SDP, entonces **tiene factorización LU**.

Corolario

Si A es SDP, entonces se puede aplicar el **método de eliminación gaussiana sin pivoteo**.

5.2 Factorización de Cholesky

Sea $A \in R^{\{n\times n\}}$ **una matriz SDP**. Entonces existe una única **matriz triangular inferior L tal que $A = LL^T$**.

5.2.1 Demostración

Sea $A \in R^{\{n\times n\}}$ una matriz SDP. Entonces A tiene factorización LU

$$A = LU \Rightarrow A^t = (LU)^t = U^t L^t$$

Además como es SDP, $A^t = A \Rightarrow U^t L^t = LU$

Como L es triangular inferior con 1s en la diagonal y L^t es triangular superior con 1s en la diagonal, ambas son inversibles. Entonces:

$$\begin{aligned} LU = U^t L^t &\Rightarrow L^{-1} LU = L^{-1} U^t L^t \Rightarrow U = L^{-1} U^t L^t \\ &\Rightarrow U (L^t)^{-1} = L^{-1} U^t L^t (L^t)^{-1} \\ &\Rightarrow U (L^t)^{-1} = L^{-1} U^t \end{aligned}$$

$U(L^t)^{-1}$ es triangular superior pues ambas matrices son triangulares superiores. Además $L^{-1}U^t$ es triangular inferior pues ambas matrices son triangulares inferiores. Por lo tanto, la igualdad a la que llegamos solo se puede dar si ambas matrices son diagonales:

$$U(L^t)^{-1} = L^{-1} U^t = D \;\; \text{matriz diagonal}$$

Además, podemos escribir U como:

$$U = DL^t$$

Entonces $A = LU = LDL^t$

Sea ahora $x \neq 0$ tal que $L^t x = e_i$. Entonces:

$$0 < x^t A x = x^t LDL^t x = (L^t x)^t D (L^t x) = e_i^t D e_i = d_{ii}$$

Esto implica que todos los elementos de la diagonal son distintos de cero, por lo tanto D es no singular. Además:

$$D = \sqrt{D}\sqrt{D}$$

donde $\sqrt{D}$ es la matriz diagonal con la raíz cuadrada de los elementos de D en la diagonal. Entonces:

$$A = LDL^t = L\sqrt{D}\sqrt{D}L^t = \left(L\sqrt{D}\right)\left(L\sqrt{D}\right)^t = \tilde{L}\tilde{L}^t$$

5.2.2 Algoritmo

El algoritmo para calcular la factorización de Cholesky es el siguiente:

Si l_{ij} son los elementos de L y a_{ij} son los elementos de A, entonces:

CHOLESKY(A: Matriz):

$l_{11} = \sqrt{a_{11}}$ **Para** $j \leftarrow 2$ a n **hacer**

$l_{j1} = \frac{a_{j1}}{l_{11}}$ **Para** $i \leftarrow 2$ a $n-1$ **hacer**

$$l_{ii} = \sqrt{a_{ii} - \sum_{k=1}^{i-1} l_{ik}^2}$$

Para $j \leftarrow i+1$ a n **hacer**

$$l_{ji} = \frac{a_{ji} - \sum_{k=1}^{i-1} l_{jk} l_{ik}}{l_{ii}}$$

Fin

$$l_{nn} = \sqrt{a_{nn} - \sum_{k=1}^{n-1} l_{nk}^2}$$

Fin

Bibliografía

6.1 Videos de clases

- Algebra Lineal
- Sistemas Lineales
- Factorización LU
- Normas y error
- Factorización SDP
- Factorización QR
- Autovalores
- Factorización SVD
- Métodos Iterativos
- Cuadrados Mínimos Lineales
- Interpolación

6.2 Enlaces

- Métodos Numéricos, CubaWiki

6.3 Libros

- R. Burden y J.D.Faires, **Análisis numérico**, International Thomson Editors, 2002.
- V. Chvatal, **Linear programming**, Freeman, 1983.
- G. Dahlquist, A. Bjorck, **Numerical methods**, Dover, 2003.
- J. Demmel, **Applied Numerical Linear Algebra**, SIAM,1997.
- J. Dennis y J. More, **Numerical methods for unconstrained optimization and nonlinear equations**, Prentice- Hall, 1983.
- P. Gill, W. Murray and M. Wright, **Numerical Linear Algebra and Optimization** , Addison Wesley, 1991.
- G. H. Golub, **Matrix Computations**, Charles F. Van Loan,JHU Press, 2013.
- G. Jerónimo, J. Sabia, S. Tesauri, **Algebra lineal**, Depto de Matemática, FCEN - UBA, 2008.
- M. Heath, **Scientific computing: an introductory survey, Philosophical Transactions**. Series A, Mathematical, Physical, and Engineering Sciences, 2002
- N. Higham, **Accuracy and Stability of Numerical Algorithms**, SIAM, 2002.
- K. Hoffman y R. Kunze, **Algebra lineal**, Prentice- Hall, 1977.
- R. Horn and C. Johnson, **Matrix Analysis**, Cambridge University Press, 2012.
- E. Isaacson and H. Keller, **Analysis of Numerical Methods**, Dover Publications, 1994.
- D. Kincaid y W. Cheney, **Análisis numérico**, Addison Wesley Iberoamericana, 1994.
- B. Kernighan y R. Pike, **The Practice of Programming**, Addison Wesley, 1999.
- C. Meyer, **Matrix analysis and applied linear algebra**, SIAM, 2010.
- P. J. Olver, C. Shakiban, **Applied Linear Algebra**, Second Edition, Springer International Publishing, 2018.
- T. Sauer, **Numerical Analysis**, Pearson, 3rd Edition, 2017.
- G. Stewart, **Introduction to matrix computations**, Academic Press, 1973.
- G. Strang, **Algebra lineal y sus aplicaciones**, Ediciones Paraninfo, 4ta ed., 2007.
- E. Süli, David F.Mayers, **An Introduction to Numerical Analysis**, Cambridge University Press, 2003.+
- L. N. Trefethen, **Numerical Linear Algebra**, SIAM, 1997.
- R. Varga, **Matrix Iterative Analysis**, Springer, 2000.

- D. Watkins, **Fundamentals of matrix computations**, John Wiley & Sons, 2010