Apuntes Metnum

October 8, 2023

Gian

Contents

1 Elementos del álgebra lineal 5
1.1 Vectores 6
1.1.1 Suma 6
1.1.2 Multiplicación por escalares 6
1.1.3 Producto interno 6
1.1.4 Combinación Lineal 6
1.1.5 Base vectorial 6
1.2 Matrices 8
1.2.1 Suma 8
1.2.2 Producto por escalares 8
1.2.3 Producto de matrices 8
1.2.4 Rango de una matriz 9
1.2.5 Determinante de una matriz 9
1.2.6 Espacio imagen 9
1.2.7 Espacio nulo 9
1.2.8 Matriz inversa 10
1.2.9 Matriz traspuesta 11
1.2.10 Submatriz principal 11
1.2.11 Matrices especiales 11
2 Sistema de ecuaciones lineales 14
2.1 Definición 15
2.2 Resolución 16
2.2.1 Sistemas de ecuaciones diagonales 16
2.2.2 Sistemas de ecuaciones triangulares 16
2.3 Sistemas de ecuaciones generales 18
2.3.1 Eliminación gaussiana 18
2.3.2 Eliminación Gaussiana con pivoteo 18
3 Factorización LU 20
3.1 Objetivo 21
3.2 Método 22
3.3 Propiedades 24
3.4 Factorización PLU 30
4 Normas vectoriales y matriciales 32
4.1 Normas vectoriales 33
4.2 Normas matriciales 34
4.2.2 Cota del error 35
5 Factorización de Cholesky 36
5.1 Matrices Simétricas Definidas Positivas 37
5.2 Método 41
5.2.1 Algoritmo 41
6 Factorización QR 43
6.1 Marices Ortogonales 44
6.1.1 Métodos 44
6.2 Método de Givens 46
6.2.1 Rotaciones de Givens 46
6.2.2 Factorización QR en el plano $(\mathbf{2} \times \mathbf{2})$ 46
6.2.3 Factorización en $\mathbb{R}^{n \times n}$ 46
6.3 Método de Householder 49
6.3.1 Reflexiones de Householder 49

6.3.2 Factorización en el plano $(\mathbf{2} \times \mathbf{2})$ 50
6.3.3 Factorización en $\mathbb{R}^{n \times n}$ 51
6.4 Propiedades 52
7 Factorización SVD 53
7.1 Autovalores 54
7.1.1 Disco de Gershgorin 58
7.1.2 Matriz semejantes 60
7.1.3 Propiedades de autovalores 62
7.2 Métodos para calcular autovalores 67
7.2.1 Método de la potencia 67
7.2.2 Método de la deflación 67
7.3 Descomposición en valores singulares (SVD) 68
7.3.1 Método 68
7.3.2 Demostración 68
7.3.3 Propiedades 70
8 Métodos iterativos para sistemas de ecuaciones lineales 72
8.1 Introducción 73
8.1.1 Método de Jacobi 73
8.1.2 Método de Gauss-Seidel 74
8.2 Análisis de convergencia 75
8.2.1 Matriz convergente 75
8.2.2 Teorema de convergencia 75
8.2.3 Cota del error 77
9 Cuadrados Mínimos Lineales 80
9.1 El problema 81
9.1.1 Ecuaciones normales 83
9.2 Resolución usando la factorización QR 85
9.3 Resolucion usando la factorización SVD 87
9.4 Estimación del error 89
10 Interpolación 91
10.1 Definiciones 92
10.1.1 Polinomio de Lagrange 92
10.2 Diferencias divididas 96
10.2.1 Demostración 96
10.3 Interpolación de Neville 99
10.4 Interpolación segmentaria 101
10.4.1 Interpolación lineal segmentaria 101
10.4.2 Interpolación cuadrática segmentaria 101
10.4.3 Interpolación cúbica segmentaria 101
11 Bibliografía 105
11.1 Videos de clases 106
11.2 Enlaces 107
11.3 Libros 108

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:

$$S = \left\{ x \in \mathbb{R}^n \quad \text{tal que } x = \sum_{i=1}^{n} \alpha_i v_i \right\}$$

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 $Ax = 0$

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

Propiedad

$$\mathbf{Nu}(\boldsymbol{A}) \neq \{0\} \Leftrightarrow \textbf{las columnas de A son linealmente dependientes}$$

Propiedad

$$\boldsymbol{Im}(\boldsymbol{A}) \bigoplus \boldsymbol{Nu}(\boldsymbol{A}) = \mathbb{R}^{\boldsymbol{m}}$$

1.2.8 Matriz inversa

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

$$\boldsymbol{A}\boldsymbol{A}^{-1} = \boldsymbol{A}^{-1}\boldsymbol{A} = \boldsymbol{I}$$

Propiedad

$$\boldsymbol{A} \ \ \textbf{es inversible} \Leftrightarrow \textbf{rang}(\boldsymbol{A}) = \boldsymbol{n} \Leftrightarrow \textbf{det}(\boldsymbol{A}) \neq \mathbf{0}$$

Cuando $\boldsymbol{A}$ es inversible, decimos que $\boldsymbol{A}$ 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

$$A_{ij}^T = A_{ji}$$

.

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 $i \in \{1, 2, \dots, n\}$ se cumple que

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

Matriz de permutación

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

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

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

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

Matriz elemental (tipo 1)

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

$$\boldsymbol{E} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \alpha & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad \boldsymbol{EA} = \begin{bmatrix} — & \mathrm{fila}_1(\boldsymbol{A}) & — \\ — & \alpha\,\mathrm{fila}_2(\boldsymbol{A}) & — \\ — & \mathrm{fila}_3(\boldsymbol{A}) & — \\ — & \mathrm{fila}_4(\boldsymbol{A}) & — \end{bmatrix}$$

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

Matriz elemental (tipo 2)

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

$$\boldsymbol{E} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ \alpha & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad \boldsymbol{EA} = \begin{bmatrix} — & \mathrm{fila}_1(\boldsymbol{A}) & — \\ — & \mathrm{fila}_2(\boldsymbol{A}) & — \\ — & \mathrm{fila}_3(\boldsymbol{A}) + \alpha\ \mathrm{fila}_1(\boldsymbol{A}) & — \\ — & \mathrm{fila}_4(\boldsymbol{A}) & — \end{bmatrix}$$

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

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 $\text{rang}(A) = n$ entonces el sistema **tiene una única solución**.
- Si $\text{rang}(A) < n$ 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 $x_n, x_{n-1}, ..., x_{i+1}$.
  - Comos 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 $\boldsymbol{A} \in \mathbb{R}^{n \times n}$ y $\boldsymbol{b} \in \mathbb{R}^{n}$, 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 $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$ en uno de la forma $\boldsymbol{U}\boldsymbol{x} = \boldsymbol{c}$, donde $\boldsymbol{U}$ 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($\boldsymbol{A}$: Matriz):

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

La version mostrada **asume que $a_{ii} \neq 0$** para todo $i \in \{1, 2, \ldots, n\}$, 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 $j \in \{i + 1, i + 2, \ldots, n\}$: En este caso, la fila i es nula, y podemos **pasar a la siguiente iteración**.
- Existe algún $j \in \{i + 1, i + 2, \ldots, n\}$ 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 $i, i + 1, ..., n$ aquella que tiene el mayor valor absoluto en la columna i. Luego, intercambiamos la fila i con la fila elegida.
- **Pivoteo completo**: En cada iteración, buscamos la celda que tiene el mayor valor absoluto entre todas las filas $i, i + 1, ..., n$ y todas las columnas $i, i + 1, ..., n$. Luego, intercambiamos la fila i con la fila de la celda elegida, y la columna 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 $\boldsymbol{I} \in \mathbb{R}^{n \times n}$ la matriz identidad, $\boldsymbol{e}_i$ el vector que tiene $\mathbf{1}$ en la posición $\boldsymbol{i}$ y $\mathbf{0}$ en el resto, y $\boldsymbol{m}_i = [\mathbf{0}, ..., \boldsymbol{m}_{i+1i}, ..., \boldsymbol{m}_{ni}]$, entonces:

$$\boldsymbol{M}^i = \boldsymbol{I} - \boldsymbol{m}_i^t \boldsymbol{e}_i$$

Propiedad

$\boldsymbol{M}^i$ es una matriz **triangular inferior** con $\mathbf{1}$ en la diagonal.

Propiedad

$\boldsymbol{M}^i$ es **inversible** y su inversa es

$$(\boldsymbol{M}^i)^{-1} = \boldsymbol{I} + \boldsymbol{m}_i^t \boldsymbol{e}_i$$

Demostración

$$\begin{aligned} \boldsymbol{M}_i (\boldsymbol{M}_i)^t &= (\boldsymbol{I} - \boldsymbol{m}_i^t \boldsymbol{e}_i)(\boldsymbol{I} + \boldsymbol{m}_i^t \boldsymbol{e}_i) \\ &= \boldsymbol{I} + \boldsymbol{m}_i^t \boldsymbol{e}_i - \boldsymbol{m}_i^t \boldsymbol{e}_i - \boldsymbol{m}_i^t \boldsymbol{e}_i \boldsymbol{m}_i^t \boldsymbol{e}_i \\ &= \boldsymbol{I} - \boldsymbol{m}_i^t \boldsymbol{e}_i \boldsymbol{m}_i^t \boldsymbol{e}_i \end{aligned}$$

Pero $\boldsymbol{e}_i \boldsymbol{m}_i^t = \mathbf{0}$ porque $\boldsymbol{e}_i$ tiene $\mathbf{0}$ en todas las posiciones salvo en la $\boldsymbol{i}$ y $\boldsymbol{m}_i^t$ tiene $\mathbf{0}$ en todas las posiciones hasta la posición $\boldsymbol{i}$. Entonces $\boldsymbol{e}_i \boldsymbol{m}_i^t = \mathbf{0}$:

$$\boldsymbol{I} - \boldsymbol{m}_i^t \boldsymbol{e}_i \boldsymbol{m}_i^t \boldsymbol{e}_i = \boldsymbol{I} - \boldsymbol{m}_i^t \mathbf{0} \boldsymbol{e}_i = \boldsymbol{I}$$

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

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

Entonces, si definimos $\boldsymbol{L} = (\boldsymbol{M}^1)^{-1} (\boldsymbol{M}^2)^{-1} ... (\boldsymbol{M}^n)^{-1}$, tenemos que $\boldsymbol{A} = \boldsymbol{L}\boldsymbol{U}$ 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 $\mathbf{1}$ 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 $\mathbf{1}$ 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.

$\boldsymbol{A}$ tiene factorización $\boldsymbol{LU}$

$\Leftrightarrow$

todas sus matrices principales son no singulares

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 U son triangulares superiores con elementos no nulos en la diagonal, por lo tanto son no singulares.

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

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

- **Caso base**: $n = 2$

  Como a_{11} no es nulo por ser la submatriz principal de orden 1 de 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 n y veamos que vale para matrices de orden $n + 1$.

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

  Si escribimos A como:

  $$A = \begin{bmatrix} A^{(n)} & c_{n+1} \\ f^t_{n+11} & a_{n+1n+1} \end{bmatrix}$$

  donde $A^{(n)} \in \mathbb{R}^{n\times n}$, $c_{n+1} \in \mathbb{R}^n$, $f_{n+11} \in \mathbb{R}^n$ y $a_{n+1n+1} \in \mathbb{R}$.

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

  Propongamos una factorización LU para A, se resaltan los valores que necesitamos calcular:

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

  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. $f^t_{n+1} = l^t_{n+1}U^{(n)}$

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

4. $a_{n+1n+1} = \boldsymbol{l}_{n+11}^t \boldsymbol{u}_{n+1} + u_{n+1,n+1}$

   Como $\boldsymbol{l}_{n+11}^t$ y $\boldsymbol{u}_{n+1}$ son conocidos, entonces es posible determinar unívocamente $u_{n+1,n+1}$.

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

Propiedad

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

Demostración

Vamos a demostrar que $\boldsymbol{A}$ 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 $\boldsymbol{A}$ tiene factorización LU.

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

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

$$\left|x_{k_0}\right| = \max_{j=1\ldots n} \left|x_j\right|$$

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

  Consideremos la ecuación k_0 del sistema $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{0}$:

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

  Separamos el término 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. ■

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

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

.

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**:

$$\boldsymbol{PA} = \boldsymbol{LU}$$

Propiedad

Toda matriz $\boldsymbol{A} \in \mathbb{R}^{n \times n}$ **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:

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

donde $\boldsymbol{M}^i = \boldsymbol{I} - \boldsymbol{m}_i^t \boldsymbol{e}_i$ 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 $\boldsymbol{PA} = \boldsymbol{LU}$

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 $\boldsymbol{P}^i\boldsymbol{P}^i = \boldsymbol{I}$. Podemos agregar entonces los siguientes terminos a la ecuación, sin modificar su resultado:

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

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

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

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

$$\begin{aligned}
\widetilde{\boldsymbol{M}}^i &= (\boldsymbol{P}^{n-1}...\boldsymbol{P}^{i+2}\boldsymbol{P}^{i+1})(\boldsymbol{I} - \boldsymbol{m}_i^t\boldsymbol{e}_i)(\boldsymbol{P}^{i+1}...\boldsymbol{P}^{n-1}) \\
&= (\boldsymbol{P}^{n-1}...\boldsymbol{P}^{i+2}\boldsymbol{P}^{i+1})\boldsymbol{I}(\boldsymbol{P}^{i+1}...\boldsymbol{P}^{n-1}) \\
&\quad -(\boldsymbol{P}^{n-1}...\boldsymbol{P}^{i+2}\boldsymbol{P}^{i+1})(\boldsymbol{m}_i^t\boldsymbol{e}_i)(\boldsymbol{P}^{i+1}...\boldsymbol{P}^{n-1}) \\
&= \boldsymbol{I} - (\boldsymbol{P}^{n-1}...\boldsymbol{P}^{i+2}\boldsymbol{P}^{i+1})(\boldsymbol{m}_i^t\boldsymbol{e}_i)(\boldsymbol{P}^{i+1}...\boldsymbol{P}^{n-1})
\end{aligned}$$

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

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

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

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

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

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

$$\boldsymbol{P}\boldsymbol{A} = \boldsymbol{L}\boldsymbol{U}\blacksquare$$

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

Cuando representamos una matriz $A \in \mathbb{R}^{n \times n}$ en una computadora, usamos aritmética de punto flotante finita. Esto quiere decir que los coeficientes que usamos y los cálculos que hacemos no son exactos. El **número de condición** de una matriz nos da una idea de cuánto puede variar la solución de un sistema de ecuaciones lineales si se dan errores de redondeo:

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

Mientras más chico sea este número, mejor condicionada estará la matriz para que la podamos usar en un algoritmo que resuelve sistemas de ecuaciones.

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 $\boldsymbol{A} \in \mathbb{R}^{n \times n}$ 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:

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

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:

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

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

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 $B^T AB$ es SDP**.

Demostración

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

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

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

  Como $A^T = A$, queda:

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

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

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

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

  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 $M_1 A M_1^T$:

$$
\begin{aligned}
M_1 A M_1^T &= \begin{bmatrix} 1 & 0 & \dots & 0 \\ -\frac{a_{21}}{a_{11}} & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ -\frac{a_{n1}}{a_{11}} & 0 & \dots & 1 \end{bmatrix} A \begin{bmatrix} 1 & -\frac{a_{21}}{a_{11}} & \dots & -\frac{a_{n1}}{a_{11}} \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{bmatrix} \\
&= \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ 0 & & \tilde{A} & \end{bmatrix} \begin{bmatrix} 1 & -\frac{a_{21}}{a_{11}} & \dots & -\frac{a_{n1}}{a_{11}} \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{bmatrix} \\
&= \begin{bmatrix} a_{11} & 0 \\ 0 & \tilde{A} \end{bmatrix}
\end{aligned}
$$

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**.

Propiedad

Sea $\boldsymbol{A} \in \mathbb{R}^{n \times m}$, entonces $\boldsymbol{A}^t \boldsymbol{A}$ es una matriz semidefinida positiva.

Demostración

Sea $\boldsymbol{x} \in \mathbb{R}^m$, entonces:

$$\boldsymbol{x}^t \boldsymbol{A}^t \boldsymbol{A} \boldsymbol{x} = (\boldsymbol{A} \boldsymbol{x})^t \boldsymbol{A} \boldsymbol{x} = \|\boldsymbol{A} \boldsymbol{x}\|^2 \geq 0$$

5.2 Método

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

Como $\boldsymbol{A}$ tiene factorización LU

$$\boldsymbol{A} = \boldsymbol{L}\boldsymbol{U} \Rightarrow \boldsymbol{A}^{t} = (\boldsymbol{L}\boldsymbol{U})^{t} = \boldsymbol{U}^{t}\boldsymbol{L}^{t}$$

Además como es SDP, $\boldsymbol{A}^{t} = \boldsymbol{A} \Rightarrow \boldsymbol{U}^{t}\boldsymbol{L}^{t} = \boldsymbol{L}\boldsymbol{U}$

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

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

$\boldsymbol{U}(\boldsymbol{L}^{t})^{-1}$ es triangular superior pues ambas matrices son triangulares superiores. Además $\boldsymbol{L}^{-1}\boldsymbol{U}^{t}$ 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:

$$\boldsymbol{U}(\boldsymbol{L}^{t})^{-1} = \boldsymbol{L}^{-1}\boldsymbol{U}^{t} = \boldsymbol{D} \quad \textbf{matriz diagonal}$$

Además, podemos escribir $\boldsymbol{U}$ como:

$$\boldsymbol{U} = \boldsymbol{D}\boldsymbol{L}^{t}$$

Entonces $\boldsymbol{A} = \boldsymbol{L}\boldsymbol{U} = \boldsymbol{L}\boldsymbol{D}\boldsymbol{L}^{t}$

Sea ahora $\boldsymbol{x} \neq \boldsymbol{0}$ tal que $\boldsymbol{L}^{t}\boldsymbol{x} = \boldsymbol{e}_{i}$. Entonces:

$$\boldsymbol{0} < \boldsymbol{x}^{t}\boldsymbol{A}\boldsymbol{x} = \boldsymbol{x}^{t}\boldsymbol{L}\boldsymbol{D}\boldsymbol{L}^{t}\boldsymbol{x} = (\boldsymbol{L}^{t}\boldsymbol{x})^{t}\boldsymbol{D}(\boldsymbol{L}^{t}\boldsymbol{x}) = \boldsymbol{e}_{i}^{t}\boldsymbol{D}\boldsymbol{e}_{i} = \boldsymbol{d}_{ii}$$

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

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

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

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

5.2.1 Algoritmo

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

Si $\boldsymbol{l}_{ij}$ son los elementos de $\boldsymbol{L}$ y $\boldsymbol{a}_{ij}$ son los elementos de $\boldsymbol{A}$, entonces:

CHOLESKY($\boldsymbol{A}$: 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

Factorización QR

6.1 Marices Ortogonales

Sea $\boldsymbol{Q} \in \mathbb{R}^{n \times n}$, $\boldsymbol{Q}$ es **ortogonales** si y solo si $\boldsymbol{Q}\boldsymbol{Q}^T = \boldsymbol{Q}^T\boldsymbol{Q} = \boldsymbol{I}$

Propiedad

Sea $\boldsymbol{Q} \in \mathbb{R}^{n \times n}$ ortogonal, entonces $\|\,\mathbf{col}_i(\boldsymbol{Q})\|_2 = 1$ para todo $i = 1 \ldots n$

Propiedad

Sea $\boldsymbol{Q} \in \mathbb{R}^{n \times n}$ ortogonal, entonces $\|\,\mathbf{fila}_i(\boldsymbol{Q})\|_2 = 1$ para todo $i = 1 \ldots n$

Propiedad

Sea $\boldsymbol{Q} \in \mathbb{R}^{n \times n}$ ortogonal, entonces $\|\boldsymbol{Q}\|_2 = 1$

Propiedad

Sea $\boldsymbol{Q} \in \mathbb{R}^{n \times n}$ ortogonal, entonces $\kappa(\boldsymbol{Q}) = 1$

Propiedad

Sea $\boldsymbol{Q} \in \mathbb{R}^{n \times n}$ ortogonal y $\boldsymbol{x} \in \mathbb{R}^n$ entonces $\|\boldsymbol{Q}\boldsymbol{x}\|_2 = \|\boldsymbol{x}\|_2$

Propiedad

Sea $\boldsymbol{Q}, \boldsymbol{R} \in \mathbb{R}^{n \times n}$ ortogonales entonces $\boldsymbol{Q}\boldsymbol{R}$ es ortogonal

6.1.1 Métodos

Sea $\boldsymbol{A} \in \mathbb{R}^{n \times n}$, $\boldsymbol{Q} \in \mathbb{R}^{n \times n}$ **ortogonal y** $\boldsymbol{R} \in \mathbb{R}^{n \times n}$ **triangular superior tal que** $\boldsymbol{A} = \boldsymbol{Q}\boldsymbol{R}$

Entonces podemos resolver el sistema de ecuaciones $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$ de la siguiente manera:

$$\boldsymbol{Ax = b \Rightarrow QRx = b \Rightarrow Q^t QRx = Q^t b \Rightarrow Rx = Q^t b}$$

Entonces **podemos resolver el sistema $Rx = Q^t b$ en $\mathcal{O}(n^2)$ operaciones**.

Veamos como obtener $\boldsymbol{Q}$ y $\boldsymbol{R}$ usando dos métodos distintos: **Método de Givens** y **Método de Householder**.

6.2 Método de Givens

6.2.1 Rotaciones de Givens

Dado un angulo θ definimos la transformación lineal **$W : \mathbb{R}^2 \to \mathbb{R}^2$ que rota al vector θ grados en el sentido horario**:

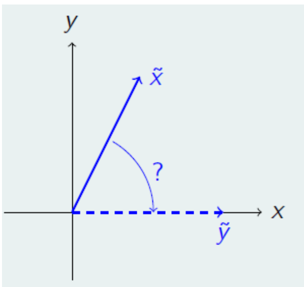
$$W = \begin{bmatrix} \cos(\theta) & \sin(\theta) \\ -\sin(\theta) & \cos(\theta) \end{bmatrix}$$

Entonces **W es ortogonal** y $\|Wx\|_2 = \|x\|_2$ para todo $x \in \mathbb{R}^2$.

Sean $\tilde{x}, \tilde{y} \in \mathbb{R}^2$ tal que

$$\tilde{y} = \begin{bmatrix} \|\tilde{x}\|_2 \\ 0 \end{bmatrix}$$

Queremos encontrar W tal que $W\tilde{x} = \tilde{y}$. Proponemos:



$$W = \begin{bmatrix} \frac{\tilde{x}_1}{\|\tilde{x}\|_2} & \frac{\tilde{x}_2}{\|\tilde{x}\|_2} \\ -\frac{\tilde{x}_2}{\|\tilde{x}\|_2} & \frac{\tilde{x}_1}{\|\tilde{x}\|_2} \end{bmatrix}$$

6.2.2 Factorización QR en el plano (2×2)

Si $A \in \mathbb{R}^{2\times 2}$ y tomamos

$$\tilde{x} = \text{col}_1(A) \begin{bmatrix} a_{11} \\ a_{21} \end{bmatrix} \qquad \tilde{y} = \begin{bmatrix} \|\tilde{x}\|_2 \\ 0 \end{bmatrix}$$

entonces podemos armar $W \in \mathbb{R}^{2\times 2}$ tal que $W\tilde{x} = \tilde{y}$. Además:

$$WA = \begin{bmatrix} \|\tilde{x}\|_2 & * \\ 0 & * \end{bmatrix} = R \Rightarrow WA = R \Rightarrow W^tWA = W^tR \Rightarrow A = W^tR$$

Si renombramos $W^t = Q$ entonces $A = QR$.

6.2.3 Factorización en $\mathbb{R}^{n\times n}$

Tenemos que ver como adapatar este proceso a más dimensiones. Sean $A \in \mathbb{R}^{n\times n}$, nuestro primer objetivo es anular el elemento a_{21}. Sabemos que si tomamos un vector

$$\tilde{x} = \begin{bmatrix} a_{11} \\ a_{21} \end{bmatrix}$$

entonces existe $W \in \mathbb{R}^{2\times 2}$ tal que

$$W\tilde{x} = \begin{bmatrix} \|\tilde{x}\|_2 \\ 0 \end{bmatrix}$$

.

Vamos a inventar una matriz $W_{12} \in \mathbb{R}^{n\times n}$ que realize esta misma operación sobre A. Definimos entonces $W_{12} \in \mathbb{R}^{n\times n}$ como:

$$W_{12} = \begin{bmatrix} w_{11} & w_{12} & 0 & \dots & 0 \\ w_{21} & w_{22} & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 1 \end{bmatrix} \Rightarrow W_{12}A = \begin{bmatrix} * & * & \dots & * \\ 0 & * & \dots & * \\ a_{31} & a_{32} & \dots & a_{3n} \\ \vdots & \vdots & \dots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{bmatrix} = A^1$$

Ahora queremos anular el elemento a_{31}, para eso vamos a usar una matriz $W_{23} \in \mathbb{R}^{n\times n}$ que realize esta misma operación sobre A^1. Nuestro nuevo es

$$\tilde{x} = \begin{bmatrix} a_{11} \\ a_{31} \end{bmatrix}$$

y nuestro nuevo $W \in \mathbb{R}^{2\times 2}$ tal que

$$W\tilde{x} = \begin{bmatrix} \|\tilde{x}\|_2 \\ 0 \end{bmatrix}$$

.

Definimos entonces $W_{13} \in \mathbb{R}^{n\times n}$ como:

$$W_{23} = \begin{bmatrix} w_{11} & 0 & w_{12} & \dots & 0 \\ 0 & 1 & & \dots & 0 \\ w_{21} & 0 & w_{22} & \dots & 0 \\ 0 & 0 & 0 & \dots & 1 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 0 \end{bmatrix} \Rightarrow W_{23}A^1 = \begin{bmatrix} * & * & \dots & * \\ 0 & * & \dots & * \\ 0 & * & \dots & * \\ a_{41} & a_{42} & \dots & a_{4n} \\ \vdots & \vdots & \dots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{bmatrix} = A^2$$

Podemos repetir este proceso, por cada elemento que queremos anular: Supongamos que queremos anular el elemento a_{ij}, entonces elegimos

$$\tilde{x} = \begin{bmatrix} a_{ii} \\ a_{ij} \end{bmatrix}$$

y $W \in \mathbb{R}^{2\times 2}$ tal que

$$W\tilde{x} = \begin{bmatrix} \|\tilde{x}\|_2 \\ 0 \end{bmatrix}$$

y definimos $W_{ij} \in \mathbb{R}^{n\times n}$ como:

$$
W_{ij} = \begin{bmatrix} 1 & \dots & 0 & \dots & 0 & \dots & 0 \\ \vdots & \ddots & \vdots & \dots & \vdots & \dots & \vdots \\ 0 & \dots & w_{ii} & \dots & w_{ij} & \dots & 0 \\ \vdots & \ddots & \vdots & \ddots & \vdots & \dots & \vdots \\ 0 & \dots & w_{ji} & \dots & w_{jj} & \dots & 0 \\ \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & \dots & 0 & \dots & 0 & \dots & 1 \end{bmatrix}
$$

En las coordenadas resaltadas, ubicamos los elementos de W.

Tras realizar este proceso $n-1$ veces, obtenemos una matriz $R \in \mathbb{R}^{n \times n}$ triangular superior:

$$
W_{n-1n} W_{n-2n} W_{n-2n-1} \dots W_{1n} \dots W_{12} A = R
$$

Ademas como cada W_{ij} es ortogonal, entonces es inversible y podemos armar una matriz $Q \in \mathbb{R}^{n \times n}$ tal que:

$$
Q = W_{12}^t \dots W_{1n}^t \dots W_{n-2n-1}^t W_{n-2n}^t W_{n-1n}^t
$$

Luego:

$$
A = QR
$$

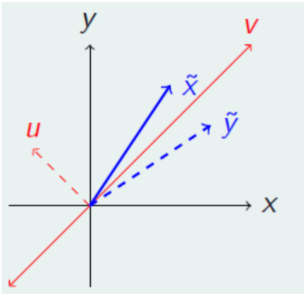
Propiedad

Ejecutar el método de Givens sobre una matriz $A \in \mathbb{R}^{n \times n}$ cuesta $\mathcal{O}\left(\frac{4}{3} n^3\right)$ **operaciones**

6.3 Método de Householder

6.3.1 Reflexiones de Householder

Dado tres vectores $\boldsymbol{u}$, $\boldsymbol{v} \in \mathbb{R}^2$ ortogonales entre si y $\tilde{\boldsymbol{x}} \in \mathbb{R}^2$.Buscamos la transformación lineal $\boldsymbol{H} : \mathbb{R}^n \to \mathbb{R}^n$ **que refleja al vector x respecto del plano definido por u y v**:



Esta matriz debe cumplir las siguientes condiciones:

1. $\boldsymbol{Hu} = -\boldsymbol{u}$
2. $\boldsymbol{Hv} = \boldsymbol{v}$
3. $\boldsymbol{H}\tilde{\boldsymbol{x}} = \tilde{\boldsymbol{y}}$

Como $\boldsymbol{u}$ y $\boldsymbol{v}$ son ortogonales, forman una base de $\mathbb{R}^2$, entonces podemos escribir $\tilde{\boldsymbol{x}}$ como combinación lineal de $\boldsymbol{u}$ y $\boldsymbol{v}$:

$$\tilde{\boldsymbol{x}} = \alpha \boldsymbol{v} + \beta \boldsymbol{u}$$

Además, la reflexion de $\tilde{\boldsymbol{x}}$ respecto del plano definido por $\boldsymbol{u}$ y $\boldsymbol{v}$ se puede escribir como $\tilde{\boldsymbol{y}} = \alpha \boldsymbol{v} - \beta \boldsymbol{u}$.

Entonces, si remplazamos $\tilde{\boldsymbol{y}}$ en la ecuación $\boldsymbol{H}\tilde{\boldsymbol{x}} = \tilde{\boldsymbol{y}}$ obtenemos que:

$$\begin{aligned}
\boldsymbol{H}\tilde{\boldsymbol{x}} &= \tilde{\boldsymbol{y}} \\
&= \alpha \boldsymbol{v} - \beta \boldsymbol{u} \\
&= \alpha \boldsymbol{v} + \beta \boldsymbol{u} - \beta \boldsymbol{u} - \beta \boldsymbol{u} \quad (\textbf{agregamos un termino que es cero}) \\
&= \alpha \boldsymbol{v} + \beta \boldsymbol{u} - 2\beta \boldsymbol{u} \\
&= \tilde{\boldsymbol{x}} - 2\beta \boldsymbol{u} \\
&= \boldsymbol{I}\tilde{\boldsymbol{x}} - \boldsymbol{W}\tilde{\boldsymbol{x}} \quad \textbf{con } \boldsymbol{W}\tilde{\boldsymbol{x}} = 2\beta \boldsymbol{u} \\
&= (\boldsymbol{I} - \boldsymbol{W})\tilde{\boldsymbol{x}}
\end{aligned}$$

Entonces $\boldsymbol{H} = \boldsymbol{I} - \boldsymbol{W}$. Ahora tenemos que encontrar $\boldsymbol{W}$. Como $\tilde{\boldsymbol{x}} = \alpha \boldsymbol{v} + \beta \boldsymbol{u}$ entonces:

$$\begin{aligned}
\boldsymbol{W}\tilde{\boldsymbol{x}} = \boldsymbol{W}(\alpha \boldsymbol{v} + \beta \boldsymbol{u}) = \alpha \boldsymbol{W}\boldsymbol{v} + \beta \boldsymbol{W}\boldsymbol{u} = 2\beta \boldsymbol{u} \\
\Rightarrow \alpha = 0 \quad \textbf{y} \quad \boldsymbol{W}\boldsymbol{u} = 2\boldsymbol{u}
\end{aligned}$$

Por ahora, asumamos por simplicidad que $\|\boldsymbol{u}\|_2 = 1$. Sea

$$\boldsymbol{P} = \boldsymbol{u}\boldsymbol{u}^t = \begin{bmatrix} u_1^2 & u_1 u_2 \\ u_1 u_2 & u_2^2 \end{bmatrix}$$

Entonce **$\boldsymbol{P}$ es simétrica** y cumple las siguientes propiedades:

- $\mathbb{P}^t = \boldsymbol{P}$
- $\boldsymbol{P}\boldsymbol{P}^t = \boldsymbol{P}$
- $\boldsymbol{P}\boldsymbol{u} = \boldsymbol{u}$
- $\boldsymbol{P}\boldsymbol{v} = 0$

Entonces, si tomamos $\boldsymbol{W} = 2\boldsymbol{P}$ entonces $\boldsymbol{W}\boldsymbol{u} = 2\boldsymbol{P}\boldsymbol{u} = 2\boldsymbol{u}$.

Finalmente, $\boldsymbol{H} = \boldsymbol{I} - 2\boldsymbol{P}$.

Propiedad

$\boldsymbol{H}$ es simétrica y ortogonal

Propiedad

Sean $\tilde{\boldsymbol{x}}, \tilde{\boldsymbol{y}} \in \mathbb{R}^{\boldsymbol{n}}$ tal que $\|\boldsymbol{x}\|_{\mathbf{2}} = \|\boldsymbol{y}\|_{\mathbf{2}}$ entonces existe una reflexión $\boldsymbol{H}$ tal que $\boldsymbol{H}\tilde{\boldsymbol{x}} = \tilde{\boldsymbol{y}}$:

$$\boldsymbol{H} = \boldsymbol{I} - \mathbf{2}\frac{(\tilde{\boldsymbol{x}} - \tilde{\boldsymbol{y}})(\tilde{\boldsymbol{x}} - \tilde{\boldsymbol{y}})^{\boldsymbol{t}}}{\|\tilde{\boldsymbol{x}} - \tilde{\boldsymbol{y}}\|_{\mathbf{2}}^{\mathbf{2}}}$$

6.3.2 Factorización en el plano (2×2)

Sea $\boldsymbol{A} \in \mathbb{R}^{\mathbf{1} \times \mathbf{2}}$ buscamos una reflexión $\boldsymbol{H}$ tal que $\boldsymbol{H}\boldsymbol{A}$ resulte en una matriz triangular. Tomemos

$$\tilde{\boldsymbol{x}} = \mathbf{col}_{\mathbf{1}}(\boldsymbol{A}) = \begin{bmatrix} \boldsymbol{a}_{\mathbf{11}} \\ \boldsymbol{a}_{\mathbf{21}} \end{bmatrix} \qquad \tilde{\boldsymbol{y}} = \begin{bmatrix} \|\tilde{\boldsymbol{x}}\|_{\mathbf{2}} \\ \mathbf{0} \end{bmatrix}$$

Entonces, por la propiedad anterior, existe una reflexión $\boldsymbol{H}$ tal que $\boldsymbol{H}\tilde{\boldsymbol{x}} = \tilde{\boldsymbol{y}}$. Entonces:

$$\boldsymbol{H}\boldsymbol{A} = \begin{bmatrix} \|\tilde{\boldsymbol{x}}\|_{\mathbf{2}} & * \\ \mathbf{0} & * \end{bmatrix} = \boldsymbol{R} \Rightarrow \boldsymbol{H}\boldsymbol{A} = \boldsymbol{R} \Rightarrow \boldsymbol{H}^{\boldsymbol{t}}\boldsymbol{H}\boldsymbol{A} = \boldsymbol{H}^{\boldsymbol{t}}\boldsymbol{R} \Rightarrow \boldsymbol{A} = \boldsymbol{H}^{\boldsymbol{t}}\boldsymbol{R}$$

Como $\boldsymbol{H}^{\boldsymbol{t}}$ es ortogonal, podemos renombrar $\boldsymbol{H}^{\boldsymbol{t}} = \boldsymbol{Q}$ y entonces $\boldsymbol{A} = \boldsymbol{Q}\boldsymbol{R}$.

6.3.3 Factorización en $\mathbb{R}^{n \times n}$

Se puede extender el mismo concepto a múltiples dimensiones. En este caso, nuestro primer objetivo es anular todos los elementos de la primer fila por debajo de la diagonal. Para eso, tomamos

$$\tilde{x} = \mathrm{col}_1(A) = \begin{bmatrix} a_{11} \\ a_{21} \\ \dots \\ a_{n1} \end{bmatrix} \qquad \tilde{y} = \begin{bmatrix} \|\tilde{x}\|_2 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

y buscamos una reflexión H_1 tal que $H_1 \tilde{x} = \tilde{y}$.

Por la misma propiedad anterior, existe una reflexión H_1 tal que $H_1 \tilde{x} = \tilde{y}$. Entonces:

$$H_1 A = \begin{bmatrix} \|\tilde{x}\|_2 & * & \dots & * \\ 0 & * & \dots & * \\ \vdots & \vdots & \dots & \vdots \\ 0 & * & \dots & * \end{bmatrix} = A^1$$

En los siguientes paso, buscamos H_2 tal que $H_2 A^1$ anule todos los elementos de la segunda fila por debajo de la diagonal.

En cada paso, buscamos la reflexión H_i que permite anular las entradas de la fila i por debajo de la diagonal.

Al final del proceso, obtenemos una secuencia de productos matriciales que nos permiten obtener una matriz $R \in \mathbb{R}^{n \times n}$ triangular superior:

$$H_n H_{n-1} \dots H_1 A = R$$

Además, como cada H_i es ortogonal, entonces es inversible y podemos armar una matriz $Q \in \mathbb{R}^{n \times n}$ tal que:

$$Q = H_1^t H_2^t \dots H_{n-1}^t$$

6.4 Propiedades

Propiedad

Ejecutar el método de Householder sobre una matriz $\boldsymbol{A} \in \mathbb{R}^{n \times n}$ cuesta $\mathcal{O}\left(\frac{2}{3}n^3\right)$ **operaciones**

Propiedad

Sea $\boldsymbol{A} \in \mathbb{R}^{n \times n}$, $\boldsymbol{A}$ no singular. Existen únicas $\boldsymbol{Q} \in \mathbb{R}^{n \times n}$ ortogonal y $\boldsymbol{R} \in \mathbb{R}^{n \times n}$ triangular superior tal que $\boldsymbol{A} = \boldsymbol{Q}\boldsymbol{R}$

Factorización SVD

7.1 Autovalores

Sea $A \in \mathbb{C}^{n \times n}$, $\lambda \in \mathbb{C}$ es **autovalor de A** si y solo si existe $v \in \mathbb{C}^n$ tal que $v \neq 0$ y

$$Av = \lambda v$$

Si λ es autovalor de A, entonces v es **autovector asociado a λ**.

Vamos a llamar **radio espectral** de A al número

$$\rho(A) = \max\{|\lambda| : \lambda \text{ es autovalor de } A\}$$

Propiedad

$A - \lambda I$ es una matriz singular (no inversible)

Vamos a definir el **polinomio característico** de A como:

$$P(\lambda) = \det(A - \lambda I)$$

Propiedad

λ es autovalor de A si y solo si es **raiz de $P(\lambda)$** (si $P(\lambda) = 0$)

Propiedad

A tiene n autovalores, algunos pueden tener multiplicidad mayor a 1 (pueden estar repetidos)

Propiedad

Si v es autovector entonces **αv también es autovector**.

Propiedad

Si $\boldsymbol{\lambda}$ es autovalor de $\boldsymbol{A} \Rightarrow$ **$\lambda - \alpha$ es autovalor de $A - \alpha I$**.

Demostración

$$(\boldsymbol{A} - \boldsymbol{\alpha I})\boldsymbol{v} = \boldsymbol{Av} - \boldsymbol{\alpha v} = \boldsymbol{\lambda v} - \boldsymbol{\alpha v} = (\boldsymbol{\lambda} - \boldsymbol{\alpha})\boldsymbol{v}$$

Propiedad

Si $\boldsymbol{\lambda}$ es autovalor de $\boldsymbol{A}$ y $\boldsymbol{v}$ autovector de asociado, entonces **$(\lambda)^{(k)}$ es autovalor de $A^{(k)}$ y v es autovector asociado**.

Demostración

Vamos a demostrar por inducción en $\boldsymbol{k}$:

- **Caso base ($k = 1$):** $\boldsymbol{Av} = \boldsymbol{\lambda v} \Rightarrow \boldsymbol{A}^{1}\boldsymbol{v} = \boldsymbol{\lambda}_{1}\boldsymbol{v}$ por definición de autovector.
- **Paso inductivo:** Supongamos que vale para $\boldsymbol{k}$, entonces $\boldsymbol{A}^{k}\boldsymbol{v} = \boldsymbol{\lambda}^{k}\boldsymbol{v}$. Queremos ver que $\boldsymbol{A}^{k+1}\boldsymbol{v} = \boldsymbol{\lambda}^{k+1}\boldsymbol{v}$.

$$\boldsymbol{A}^{k+1}\boldsymbol{v} = \boldsymbol{A}^{k}(\boldsymbol{Av}) = \boldsymbol{A}^{k}(\boldsymbol{\lambda v}) = \boldsymbol{\lambda A}^{k}\boldsymbol{v}$$

Por hipótesis inductiva $\boldsymbol{A}^{k}\boldsymbol{v} = \boldsymbol{\lambda}^{k}\boldsymbol{v}$ entonces

$$\boldsymbol{\lambda A}^{k}\boldsymbol{v} = \boldsymbol{\lambda\lambda}^{k}\boldsymbol{v} = \boldsymbol{\lambda}^{k+1}\boldsymbol{v} \quad \blacksquare$$

Propiedad

Sea **$Q \in \mathbb{R}^{n \times n}$ una matriz ortogonal**, entonces **sus autovalores reales son 1 o -1**

Demostración

Por ser $\boldsymbol{Q}$ ortogonal, sabeemos que $\|\boldsymbol{Av}\|_{2} = \|\boldsymbol{v}\|_{2}$. Supongamos que $\boldsymbol{v}$ es autovector asociado a $\boldsymbol{\lambda}$ entonces $\boldsymbol{Av} = \boldsymbol{\lambda v}$ y $\|\boldsymbol{Av}\|_{2} = \|\boldsymbol{\lambda v}\|_{2} = |\boldsymbol{\lambda}|\ \|\boldsymbol{v}\|_{2}$. Entonces $|\boldsymbol{\lambda}| = \boldsymbol{1} \quad \blacksquare$

Propiedad

Si $\lambda_1, \lambda_2, ..., \lambda^n$ **son los autovalores distintos** con autovalores asociados $v^1, v^2, ..., v^n$ **son linealmente independientes**.

Demostración

Vamos a demostrar por inducción en k:

- **Caso base ($k = 1$):** es válido ya que v_1 no es el vector nulo.
- **Paso inductivo:** Supongamos que vale para k, entonces $v^1, v^2, ..., v^k$ son linealmente independientes. Queremos ver que $v^1, v^2, ..., v^{k+1}$ tambien lo son.

Supongamos que no, entonces v^{k+1} se puede escribir como combinación lineal de $v^1, v^2, ..., v^k$:

$$v^{k+1} = \sum_{i=1}^{k} c_i v^i$$

Los coeficientes c_i no son todos nulos ya que $v^{k+1} \neq 0$ por definición de autovector.

Multiplicando por A a ambos lados de la igualdad queda:

$$A v^{k+1} = \lambda^{k+1} v^{k+1}$$

y

\` $A \sum_{i=1}^{k} c_i v^i = \sum_{i=1}^{k} c_i A v^i = \sum_{i=1}^{k} c_i \lambda^i v^i$

Osea que:

$$\lambda^{k+1} v^{k+1} = \sum_{i=1}^{k} c_i \lambda^i v^i$$

Por otro lado, si multiplicamos la igualdad original por λ^{k+1} queda:

$$\lambda^{k+1} v^{k+1} = \lambda^{k+1} \sum_{i=1}^{k} c_i v^i = \sum_{i=1}^{k} c_i \lambda^{k+1} v^i$$

Entonces podemos combinar las igualdades y nos queda:

$$\begin{aligned} \sum_{i=1}^{k} c_i \lambda^i v^i &= \sum_{i=1}^{k} c_i \lambda^{k+1} v^i \\ \Rightarrow \sum_{i=1}^{k} c_i \lambda^i v^i - \sum_{i=1}^{k} c_i \lambda^{k+1} v^i &= 0 \\ \Rightarrow \sum_{i=1}^{k} c_i (\lambda^i - \lambda^{k+1}) v^i &= 0 \end{aligned}$$

Ahora, como $v^1, v^2, ..., v^k$ son linealmente independientes (por hipotesis induvtiva), la única manera de que esta combinación lineal resulte en el vector nulo es que $c_{j(\lambda^j - \lambda^{k+1})} = 0$ para todo $j = 1...k$.

Sabemos que existe algún $\boldsymbol{c_j \neq 0}$, lo que implica que para ese $\boldsymbol{j}$ tiene que valer que $\boldsymbol{\lambda^j - \lambda^{k+1} = 0 \Rightarrow \lambda^j = \lambda^{k+1}}$, lo cual es una contradicción ya que por hipótesis $\boldsymbol{\lambda_1, \lambda_2, \ldots, \lambda^n}$ son distintos. ■

Propiedad

$\boldsymbol{A}$ y $\boldsymbol{A^t}$ tienen los **mismos autovalores**

Demostración

Sabemos que los autovalores de $\boldsymbol{A}$ son las raíces del polinomio característico $\boldsymbol{P(\lambda) = \det(A - \lambda I)}$. Además sabemos que *la determinante de una matriz y su traspuesta son iguales*, entonces

$$\boldsymbol{\det(A - \lambda I) = \det\left((A - \lambda I)^t\right) = \det(A^t - \lambda I)}$$

Entonces el polinomio característico de $\boldsymbol{A}$ y $\boldsymbol{A^t}$ es el mismo y sus raíces también lo son.

7.1.1 Disco de Gershgorin

Sea $A \in \mathbb{C}^{n\times n}$ y

$$r_i = \sum_{\substack{k=1 \\ k\neq i}}^{n} |a_{\{ik\}}|$$

definimos el disco de Gershgorin

$$D_i = \{x \in \mathbb{C} : |x - a_{ii}| \leq r_i\}$$

Propiedad

Sea λ autovalor de A entonces $\lambda \in D_i$ **para algún** $i = 1, 2, ..., n$

Demostración

Sea v el autovector asociado a λ. Como $v \neq 0$ entonces

$$\|v\|_\infty = \max_{1\leq i\leq n} |v_i| \neq 0$$

Sea k_0 el índice de una coordenada de v tal que $\|v\|_\infty = \left|v_{k_0}\right|$. Sabemos que $Av = \lambda v$. En particular, considerando la fila k_0 de A, tenemos que:

$$\sum_{j=1}^{n} a_{k_0 j} v_j = \lambda v_{k_0}$$

Separando de la sumatoria el término k_0:

$$a_{k_0 k_0} v_{k_0} + \sum_{\substack{j=1 \\ j\neq k_0}}^{n} a_{k_0 j} v_j = \lambda v_{k_0}$$

$$\Rightarrow \sum_{\substack{j=1 \\ j\neq k_0}}^{n} a_{k_0 j} v_j = \lambda v_{k_0} - a_{k_0 k_0} v_{k_0} = \left(\lambda - a_{k_0 k_0}\right) v_{k_0}$$

Si tomamos módulo a ambos lados de la igualdad nos queda:

$$\left|\sum_{\substack{j=1 \\ j\neq k_0}}^{n} a_{k_0 j} v_j\right| = |\lambda - a_{k_0 k_0}|\ |v_{k_0}| = |\lambda - a_{k_0 k_0}|\ \|v\|_\infty$$

$$\Rightarrow \sum_{\substack{j=1 \\ j\neq k_0}}^{n} |a_{k_0 j} v_j| \geq |\lambda - a_{k_0 k_0}|\ |v_{k_0}|$$

Como $|v_{k_0}| \neq 0$ podemos pasar dividiendo y nos queda:

$$\sum_{\substack{j=1 \\ j\neq k_0}}^{n} |a_{k_0 j}| \frac{|v_j|}{|v_{k_0}|} \geq |\lambda - a_{k_0 k_0}|$$

Como $\|v\|_\infty = |v_{k_0}|$ entonces $\frac{|v_j|}{|v_{k_0}|} \leq 1$ y por lo tanto:

$$\sum_{\substack{j=1 \\ j \neq k_0}}^{n} |a_{k_0 j}| \geq \sum_{\substack{j=1 \\ j \neq k_0}}^{n} |a_{k_0 j}| \frac{|v_j|}{|v_{k_0}|} \geq |\lambda - a_{k_0 k_0}|$$

Luego λ cumple la condición para pertenecer a D ∎

Propiedad

Si $M = D_{i_1} \cup D_{i_2} \cup \ldots \cup D_{i_k}$ **es disjunto con la unión de los restantes discos** D_i **entonces hay exactamente** m **autovalores de** A (contados con su multiplicidad) en M.

7.1.2 Matriz semejantes

Sean $\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{C} \in \mathbb{C}^{n \times n}$ son **matrices semejantes** si existe $\boldsymbol{P} \in \mathbb{C}^{n \times n}$ matriz inversible, tal que:

$$\boldsymbol{A} = \boldsymbol{P}^{-1}\boldsymbol{B}\boldsymbol{P}$$

Propiedad

Si $\boldsymbol{A}$ y $\boldsymbol{B}$ son semejantes entonces **tienen los mismos autovalores**

Demostración

Sea $\boldsymbol{\lambda}$ autovalor de $\boldsymbol{A}$ y $\boldsymbol{v}$ autovector asociado. Queremos ver que $\boldsymbol{\lambda}$ es autovalor de $\boldsymbol{B}$. Sabemos que

$$\boldsymbol{A}\boldsymbol{v} = \boldsymbol{\lambda}\boldsymbol{v}$$

Multiplicando a ambos lados por $\boldsymbol{P}^{-1}$ nos queda:

$$\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{v} = \boldsymbol{\lambda}\boldsymbol{P}^{-1}\boldsymbol{v}$$

Como $\boldsymbol{A}$ y $\boldsymbol{B}$ son semejantes:

$$\begin{gathered}\boldsymbol{P}^{-1}\boldsymbol{P}\boldsymbol{B}\boldsymbol{P}^{-1}\boldsymbol{v} = \boldsymbol{\lambda}\boldsymbol{P}^{-1}\boldsymbol{v} \\ \Rightarrow \boldsymbol{B}\boldsymbol{P}^{-1}\boldsymbol{v} = \boldsymbol{\lambda}\boldsymbol{P}^{-1}\boldsymbol{v}\end{gathered}$$

Como $\boldsymbol{P}$ es inversible y $\boldsymbol{v} \neq \boldsymbol{0}$, entonces $\boldsymbol{P}^{-1}\boldsymbol{v} \neq \boldsymbol{0}$ y por lo tanto $\boldsymbol{\lambda}$ Si nombramos $\boldsymbol{u} = \boldsymbol{P}^{-1}\boldsymbol{v}$ entonces $\boldsymbol{B}\boldsymbol{u} = \boldsymbol{\lambda}\boldsymbol{u}$ podemos concluir que $\boldsymbol{\lambda}$ es autovalor de $\boldsymbol{B}$ y $\boldsymbol{u}$ su autovector asociado ■

Diagonalización por semenjanza

Dada $\boldsymbol{A} \in \mathbb{C}^{n \times n}$, $\boldsymbol{A}$ es **diagonizable por semejanza** si es semejante a una matriz diagonal.

Propiedad

$\boldsymbol{A}$ es diagonizable por semejanza si y solo si sus **autovectores forman una base**.

Demostración

$\Rightarrow$) Si $\boldsymbol{A}$ es diagonizable por semejanza, entonces existe $\boldsymbol{P}$ matriz inversible tal que $\boldsymbol{A} = \boldsymbol{P}\boldsymbol{D}\boldsymbol{P}^{-1}$. Por la propiedad anterior, $\boldsymbol{D}$ tiene los mismos autovalores que $\boldsymbol{A}$. Además si $\boldsymbol{v}$ es autovector de $\boldsymbol{D}$, entonces $\boldsymbol{P}\boldsymbol{v}$ es autovector de $\boldsymbol{A}$.

Los autovectores de una matriz diagonal son los vectores canónicos $\boldsymbol{e}_1, \boldsymbol{e}_2, \dots, \boldsymbol{e}_n$. Entonces los autovectores de $\boldsymbol{A}$ son $\boldsymbol{P}\boldsymbol{e}_1, \boldsymbol{P}\boldsymbol{e}_2, \dots, \boldsymbol{P}\boldsymbol{e}_n$.

Ahora, $\boldsymbol{P}\boldsymbol{e}_i = \mathbf{col}_{i(\boldsymbol{P})}$. Como $\boldsymbol{P}$ es inversible, sus columnas son linealmente independiente, por lo tanto, $\boldsymbol{P}\boldsymbol{e}_1, \boldsymbol{P}\boldsymbol{e}_2, \dots, \boldsymbol{P}\boldsymbol{e}_n$ son linealmente independientes y forman una base.

Luego, los auvotes de $\boldsymbol{A}$ forman una base.

$\Leftarrow$) Si $\boldsymbol{A}$ tiene base de autovectores, entonces existen $\boldsymbol{v_1}, \boldsymbol{v_2}, ..., \boldsymbol{v_n}$ autovectores linealmente independientes tal que $\boldsymbol{A}\boldsymbol{v_i} = \boldsymbol{\lambda_i}\boldsymbol{v_i}$ siendo $\boldsymbol{\lambda_i}$ el autovalor al que están asociados.

Definamos $\boldsymbol{P} \in \mathbb{R}^{n \times n}$ a la matriz cuyas columnas son los autovectores $\boldsymbol{v_1}, \boldsymbol{v_2}, ..., \boldsymbol{v_n}$. Como $\boldsymbol{v_1}, \boldsymbol{v_2}, ..., \boldsymbol{v_n}$ son linealmente independientes $\Rightarrow \boldsymbol{P}$ es inversible.

$$\begin{aligned}
\boldsymbol{AP} &= \boldsymbol{A}[\boldsymbol{v_1} \;\; \boldsymbol{v_2} \;\; \dots \;\; \boldsymbol{v_n}] = [\boldsymbol{Av_1} \;\; \boldsymbol{Av_2} \;\; \dots \;\; \boldsymbol{Av_n}] \\
&= [\boldsymbol{\lambda_1 v_1} \;\; \boldsymbol{\lambda_2 v_2} \;\; \dots \;\; \boldsymbol{\lambda_n v_n}] \\
&= [\boldsymbol{v_1} \;\; \boldsymbol{v_2} \;\; \dots \;\; \boldsymbol{v_n}] \begin{bmatrix} \boldsymbol{\lambda_1} & \boldsymbol{0} & \dots & \boldsymbol{0} \\ \vdots & \vdots & \ddots & \vdots \\ \boldsymbol{0} & \boldsymbol{0} & \dots & \boldsymbol{\lambda_n} \end{bmatrix} \\
&= \boldsymbol{PD}
\end{aligned}$$

Entonces logramos conseguir $\boldsymbol{D} \in \mathbb{R}^{n \times n}$ diagonal tal que $\boldsymbol{AP} = \boldsymbol{PD}$ y como $\boldsymbol{P}$ es inversible, si multiplicamos a derecha por su inversa en ambos lados nos queda:

$$\boldsymbol{APP^{-1}} = \boldsymbol{PDP^{-1}} \Rightarrow \boldsymbol{A} = \boldsymbol{PDP^{-1}}$$

Luego $\boldsymbol{A}$ es diagonizable por semejanza ■

7.1.3 Propiedades de autovalores

Sea $\boldsymbol{A} \in \mathbb{R}^{n \times n}$:

Propiedad

Si $\boldsymbol{A}$ es **simétrica**, sus **autovolores son reales**.

Demostración

Sea $\boldsymbol{\lambda}$ autovalor de $\boldsymbol{A}$ y $\boldsymbol{v} \neq \mathbf{0}$ su autovector asociado. Entonces $\boldsymbol{Av} = \boldsymbol{\lambda v}$. *Conjugemos* (Si $\boldsymbol{a} = \boldsymbol{x} + \boldsymbol{yi} \Rightarrow \bar{\boldsymbol{a}} = \boldsymbol{x} - \boldsymbol{yi}$) ambos lados de la igualdad:

$$\boldsymbol{Av} = \boldsymbol{\lambda v} \Rightarrow \bar{\boldsymbol{A}}\boldsymbol{v} = \bar{\boldsymbol{\lambda}}\boldsymbol{v} \Rightarrow \bar{\boldsymbol{A}}\bar{\boldsymbol{v}} = \bar{\boldsymbol{\lambda}}\bar{\boldsymbol{v}}$$

Como $\boldsymbol{A} \in \mathbb{R}^{n \times n}$, $\bar{\boldsymbol{A}} = \boldsymbol{A}$:

$$\bar{\boldsymbol{A}}\bar{\boldsymbol{v}} = \bar{\boldsymbol{\lambda}}\bar{\boldsymbol{v}} \Rightarrow \boldsymbol{A}\bar{\boldsymbol{v}} = \bar{\boldsymbol{\lambda}}\bar{\boldsymbol{v}}$$

Multipliquemos por $\boldsymbol{v}^t$ ambos lados de la igualdad:

$$\boldsymbol{v}^t\boldsymbol{A}\bar{\boldsymbol{v}} = \boldsymbol{v}^t\bar{\boldsymbol{\lambda}}\bar{\boldsymbol{v}} \Rightarrow \boldsymbol{v}^t\boldsymbol{A}\bar{\boldsymbol{v}} = \bar{\boldsymbol{\lambda}}\boldsymbol{v}^t\bar{\boldsymbol{v}} \Rightarrow \left(\boldsymbol{A}^t\boldsymbol{v}\right)^t\bar{\boldsymbol{v}} = \bar{\boldsymbol{\lambda}}\boldsymbol{v}^t\bar{\boldsymbol{v}}$$

Como $\boldsymbol{A}$ es simétrica:

$$\left(\boldsymbol{A}^t\boldsymbol{v}\right)^t\bar{\boldsymbol{v}} = \bar{\boldsymbol{\lambda}}\boldsymbol{v}^t\bar{\boldsymbol{v}} \Rightarrow (\boldsymbol{Av})^t\bar{\boldsymbol{v}} = \bar{\boldsymbol{\lambda}}\boldsymbol{v}^t\bar{\boldsymbol{v}}$$

Como $\boldsymbol{v}$ es autovector asociado a $\boldsymbol{\lambda}$:

$$(\boldsymbol{Av})^t\bar{\boldsymbol{v}} = \bar{\boldsymbol{\lambda}}\boldsymbol{v}^t\bar{\boldsymbol{v}} \Rightarrow \boldsymbol{\lambda}\boldsymbol{v}^t\bar{\boldsymbol{v}} = \bar{\boldsymbol{\lambda}}\boldsymbol{v}^t\bar{\boldsymbol{v}}$$

Ahora como $\boldsymbol{v} \neq \mathbf{0}$:

$$\boldsymbol{v}^t\bar{\boldsymbol{v}} = \boldsymbol{v}_1 * \bar{\boldsymbol{v}}_1 + \boldsymbol{v}_2 * \bar{\boldsymbol{v}}_2 + \ldots + \boldsymbol{v}_n * \bar{\boldsymbol{v}}_n = |\boldsymbol{v}_1|^2 + |\boldsymbol{v}_2|^2 + \ldots + |\boldsymbol{v}_n|^2 > \mathbf{0}$$

Entonces:

$$\boldsymbol{\lambda}\boldsymbol{v}^t\bar{\boldsymbol{v}} = \bar{\boldsymbol{\lambda}}\boldsymbol{v}^t\bar{\boldsymbol{v}} \Rightarrow \boldsymbol{\lambda} = \bar{\boldsymbol{\lambda}}\frac{\boldsymbol{v}^t\bar{\boldsymbol{v}}}{\boldsymbol{v}^t\bar{\boldsymbol{v}}} \Rightarrow \boldsymbol{\lambda} = \bar{\boldsymbol{\lambda}}$$

Luego $\boldsymbol{\lambda}$ es real ■

Propiedad

Si $\boldsymbol{A}$ tiene un **autovalor real** entonces existe un **autovector asociado con coeficientes reales**.

Demostración

Sea $\boldsymbol{\lambda}$ autovalor real de $\boldsymbol{A}$ y $\boldsymbol{v}$ autovector asociado. Entonces $\boldsymbol{v}$ es solución al sistema $(\boldsymbol{A} - \boldsymbol{\lambda I})\boldsymbol{v} = \boldsymbol{0}$. Este sistema puede resolverse utilizando elminiación gaussiana. Como $\boldsymbol{A} \in \mathbb{R}^{n \times n}$, todos los coeficientes que se usan en el proceso son reales, por lo tanto, la solución que se obtiene tendrá coeficientes reales ■

Propiedad

Si $\boldsymbol{A}$ es **simétrica** y **λ_1 y λ_2 autovalores distintos** con $\boldsymbol{v_1}$ y $\boldsymbol{v_2}$ autovectores asociados. Entonces **v_1 y v_2 son ortogonales**.

Demostración

Sea $\boldsymbol{\lambda_1}$ y $\boldsymbol{\lambda_2}$ autovalores distintos de $\boldsymbol{A}$ y $\boldsymbol{v_1}$ y $\boldsymbol{v_2}$ autovectores asociados. Entonces $\boldsymbol{A v_1} = \boldsymbol{\lambda_1 v_1}$ y $\boldsymbol{A v_2} = \boldsymbol{\lambda_2 v_2}$.

Multiplicando a ambos lados de la primera igualdad por $\boldsymbol{v_2}$ y a ambos lados de la segunda igualdad por $\boldsymbol{v_1}$ nos queda:

$$\begin{aligned} v_2^t A v_1 &= v_2^t \lambda_1 v_1 \\ v_1^t A v_2 &= v_1^t \lambda_2 v_2 \end{aligned}$$

Como $\boldsymbol{A}$ es simétrica, $\boldsymbol{A^t} = \boldsymbol{A}$ y entonces:

$$v_2^t A v_1 = (A^t v_2)^t v_1 = (A v_2)^t v_1 = (v_1^t A v_2)^t = v_1^t A v_2$$

Entonces:

$$\lambda_1 v_2^t v_1 = \lambda_2 v_1^t v_2$$

Como por hipotesis $\boldsymbol{\lambda_1} \neq \boldsymbol{\lambda_2}$ y además $\boldsymbol{v_2^t v_1} = \boldsymbol{v_1^t v_2}$, entonces $\boldsymbol{v_1^t v_2} = \boldsymbol{0}$ y por lo tanto $\boldsymbol{v_1}$ y $\boldsymbol{v_2}$ son ortogonales ■

Propiedad

Sea $Q \in \mathbb{R}^{n\times n}$ **ortogonal, λ es autovalor de A** si y solo si **λ es autovalor de QAQ^t**.

Teorema de Schur (simplificado)

Si **A tiene todos tus autovalores reales**, **existe $Q \in \mathbb{R}^{n\times n}$ ortogonal** tal que

$$Q^t A Q = T$$

con **$T \in \mathbb{R}^{n\times n}$ triangular superior**.

Demostración

Sea λ_1 autovalor de A y v_1 su autovector asociado, asumamos que $\|v_1\|_2 = 1$.

Sabemos que existe una transformación de Householder H_1 tal que $H_1 v_1 = e_1$ con e_1 el primer vector canónico. Como H_1 es una transformación de householder es simétirca y ortogonal, por lo tanto $H_1^t = H_1^{-1} = H_1$. Entonces:

$$H_1 A H_1^t e_1 = H_1 A v_1 = H_1 \lambda_1 v_1 = \lambda_1 H_1 v_1 = \lambda_1 e_1$$

Osea que

$$H_1 A H_1^t = \begin{bmatrix} \lambda_1 & * & \cdots & * \\ 0 & & \tilde{A} & \end{bmatrix}$$

Además A es semejante a $H_1 A H_1^t$ pues:

$$\begin{aligned} A &= H_1^{-1} H_1 A H_1^{-1} H_1 \\ &= H_1^t H_1 A H_1^t H_1 \quad (\text{porque } H_1 \text{ es ortogonal}) \\ &= H_1^t (H_1 A H_1^t) H_1 \end{aligned}$$

Luego, como A es semejante a $H_1 A H_1^t$, ambas matrices tienen los mismos autovalores.

Como $\mathrm{col}_1(H_1 A H_1^t) = \lambda_1 e_1$, sabemos tambien que el resto de los autovalores están determinados por los autovalores de la matriz $\tilde{A} \in \mathbb{R}^{(n-1)\times(n-1)}$ que se obtiene al eliminar la primera fila y la primera columna de A:

Repitiendo el proceso para $\tilde{A}$ obtenemos una reflexión de Householder $\widetilde{H}_2 \in \mathbb{R}^{(n-1)\times(n-1)}$ que tiene exactametne las mismas propiedades que H_1:

- $\widetilde{H}_2 \tilde{A} \widetilde{H}_2^t = \lambda_2 e_2$ con $e_2 \in \mathbb{R}^{n-1}$
- $\tilde{A}$ y $\widetilde{H}_2 \tilde{A} \widetilde{H}_2^t$ son semejantes y tienen los mismos autovalores.

Armemos entonces $H_2 \in \mathbb{R}^{n\times n}$ de la siguiente manera:

$$H_2 = \begin{bmatrix} 1 & 0^t \\ 0 & \widetilde{H}_2 \end{bmatrix}$$

Queda entonces que:

$$H_2 H_1 A H_1^t H_2^t = \begin{bmatrix} \lambda_1 & * & \cdots & * \\ 0 & \lambda_2 & \cdots & * \\ 0 & 0 & \cdots & * \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & * \end{bmatrix}$$

Repitiendo este proceso $n-1$ llegaremos a una matriz T triangular superior tal que

$$H_{n-1} ... H_2 H_1 A H_1^t H_2^t ... H_{n-1}^t = T$$

Solo queda renombrar $Q = H_{n-1} ... H_2 H_1$ y listo ■

Propiedad

Si **A es simétrica**, entonces **T es diagonal**. **Los elementos** de la diagonal de T **son los autovalores** y **las columnas de Q son los autovectores**.

Demostración

Como A es simétrica, todos sus autovalores son reales. Entonces, por el teorema de Schur, existe $Q \in \mathbb{R}^{n \times n}$ ortogonal y T traiangular superior tal que

$$A = QTQ^t$$

Además

$$A^t = (QTQ^t)^t = (TQ^t)^t Q^t = (Q^t)^t T^t Q^t = QT^t Q^t$$

Como A es simétrica, $A^t = A$ y entonces:

$$QTQ^t = QT^t Q^t$$

Como Q es ortogonal, $Q^t = Q^{-1}$:

$$QTQ^t = QT^t Q^t \Rightarrow Q^t QTQ^t Q = Q^t QT^t Q^t Q \Rightarrow T = T^t$$

Luego T^t es una matriz triangular superior y simétrica, por lo tanto es diagonal, es decir A es diagonizable por semejanza ■

Corolario

Si **A es simétrica**, entonces **tiene base (ortonormal) de autovectores**.

Demostración

Sabemos que $\boldsymbol{A}$ es diagonizable por semejanza a $\boldsymbol{Q T Q^t}$, esto implica que tiene base de autovecotores. Por la propiedad anterior, tambien sabemos que las columnas de $\boldsymbol{Q}$ son los autovectores de $\boldsymbol{A}$. Como $\boldsymbol{Q}$ es ortogonal, sus columnas son linealmente independientes y además $\|\mathbf{col}_{\boldsymbol{i}}(\boldsymbol{Q})\|_{\mathbf{2}} = \mathbf{1}$ para todo $\boldsymbol{i} = \mathbf{1}, \mathbf{2}, \ldots, \boldsymbol{n}$ concluimos que esa base es ortonormal ■

7.2 Métodos para calcular autovalores

7.2.1 Método de la potencia

Sea $A \in \mathbb{R}^{n \times n}$, $\lambda_1, \lambda_2, \ldots, \lambda^n$ sus n autovalores con **autovectores asociados $v^1, v^2, \ldots, v^n$ que forman una base**.

Supongamos que $|\lambda_1| > |\lambda_2| > \ldots \geq |\lambda^n|$. Sea $q_0 \in \mathbb{R}^n$ tal que $\|q_0\| = 1$, definimos la sucesión $\{q_k\}$ donde cada uno de sus elementos se define como:

$$q^k = \frac{Aq^{k-1}}{\|Aq^{k-1}\|_2}$$

Esta sucesión **converge al autovector v^1**. Además $\lambda_k = \left(q^k\right)^t Aq^k$ **converge a λ_1**.

7.2.2 Método de la deflación

Sea $A \in \mathbb{R}^{n \times n}$, λ_1 autovalor de A con autovector asociado v^1 y $\|v^1\| = 1$.

Sea $H \in \mathbb{R}^{n \times n}$ **ortogonal** tal que $H^t v^1 = e_1$, entonces:

$$H^t AH = \begin{bmatrix} \lambda_1 & a^t \\ 0 & B \end{bmatrix}$$

Como A y $H^t AH$ tienen los mismos autovalores, **los otros autovalores de A son los autovalores de B**.

7.3 Descomposición en valores singulares (SVD)

7.3.1 Método

Sea $A \in \mathbb{R}^{m \times n}$, $r = \text{rang}(A)$, existen matrices $U \in \mathbb{R}^{m \times m}$, $V \in \mathbb{R}^{n \times n}$ ortogonales y $\Sigma \in \mathbb{R}^{m \times n}$ **diagonal** tal que:

$$A = U \Sigma V^t$$

Además, Σ tiene la forma:

$$\Sigma = \begin{bmatrix} \sigma_1 & 0 & \dots & 0 & 0 & \cdots & 0 \\ 0 & \sigma_2 & \dots & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \sigma_r & 0 & \cdots & 0 \\ 0 & 0 & \dots & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 0 & 0 & \cdots & 0 \end{bmatrix}$$

con $\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_r > 0$. Llamamos a los σ_i **valores singulares** de A.

Proponemos la siguientes condiciones para las matrices de esta descomposición:

- **U tiene como columnas los autovectores de AA^t**
- **V tiene como columnas los autovectores de $A^t A$**
- $\sigma_i = \sqrt{\lambda_i}$ donde λ_i es el i —ésimo autovalor de $A^t A$ si los tomamos ordenados de mayor a menor: $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_r$.

7.3.2 Demostración

Sea $u_1, u_2, \dots, u_m$ **a las columnas de U** y $v_1, v_2, \dots, v_n$ **a las columnas de V**. Si la descomposición SVD existe, entonces $A = U \Sigma V^t$ y $A^t = V \Sigma^t U^t$. Entonces:

$$AV = U\Sigma \underbrace{V^t V}_{\text{ortogonales}} = U\Sigma \Rightarrow Av_i = \begin{cases} \sigma_i u_i \text{ si } i \leq r \\ 0 \text{ sino} \end{cases}$$

$$A^t U = V\Sigma^t \underbrace{U^t U}_{\text{ortogonales}} = V\Sigma^t \Rightarrow A^t u_i = \begin{cases} \sigma_i v_i \text{ si } i \leq r \\ 0 \text{ sino} \end{cases}$$

Columnas de V

Supongamos $i \leq r$ multiplicamos la primera relación por A^t obtenemos:

$$A^t A v_i = \sigma_i A^t u_i = \sigma_i^2 v_i$$

Entonces **v_i es autovector de $A^t A$**, además **σ_i^2 deben ser sus autovalores**.

$A^t A$ es una matriz simétrica definida positiva (ver Section 5.1) y su rango es r. Por la condición de símetria sabemos que **existe una base ortonormal de autovectores** (ver Section 7.1.3). Además por ser semidefinida positiva, **todos sus autovalores son reales y positivos**.

Sabemos que r de esos autovalores son no nulos y el 0 es autovalor con multiplicidad $n - r$. Sean $\lambda_1, ..., \lambda_r$ los autovalores no nulos de $A^t A$ y $v_1, ..., v_r$ los autovectores asociados y $v_{r+1}, ..., v_n$ los autovectores asociados al 0, entonces **$v_1, ..., v_r, v_{r+1}, ..., v_n$ es una base ortonormal de autovectores de $A^t A$**.

Estos vectores son los candidatos a conformar las columnas de V y definimos $\sigma_i = \sqrt{\lambda_i} > 0$.

Columnas de U

De la relación $Av_i = \sigma_i u_i$ **podemos despejar u_i** para $i \leq r$:

$$u_i = \frac{1}{\sigma_i} A v_i$$

Para que esta relación sea correcta, **debemos ver que $u_1, ..., u_r$ son otornomales**:

$$\begin{aligned}
u_i^t u_j &= \left(\frac{1}{\sigma_i} A v_i\right)^t \left(\frac{1}{\sigma_j} A v_j\right) = \frac{1}{\sigma_i}(Av_i)^t \frac{1}{\sigma_j} A v_j = \frac{1}{\sigma_i}\frac{1}{\sigma_j}(Av_i)^t A v_j \\
&= \frac{1}{\sigma_i}\frac{1}{\sigma_j} v_i^t A^t A v_j = \frac{1}{\sigma_i}\frac{1}{\sigma_j} v_i^t \lambda_j v_j \qquad (v_j \text{ autovector de } A^t A) \\
&= \frac{1}{\sigma_i}\frac{1}{\sigma_j} \lambda_j v_i^t v_j = 0 \qquad (v_i \;\text{ y } v_j \text{ son ortonormales})
\end{aligned}$$

Luego **$u_1, ..., u_r$ son ortogonales**. Falta ver que $\| u_i \|_2 = 1$:

$$\begin{aligned}
u_i^t u_i &= \left(\frac{1}{\sigma_i} A v_i\right)^t \left(\frac{1}{\sigma_i} A v_i\right) = \frac{1}{\sigma_i}(Av_i)^t \frac{1}{\sigma_i} A v_i = \frac{1}{\sigma_i^2}(Av_i)^t A v_i \\
&= \frac{1}{\sigma_i^2} v_i^t A^t A v_i = \frac{1}{\sigma_i^2} \lambda_i v_i^t \qquad (v_i \text{ autovector de } A^t A) \\
&= \frac{1}{\sigma_i^2} \lambda_i v_i^t v_i = \frac{1}{\sigma_i^2} \lambda_i \qquad (v_i^t v_i = 1) \\
&= \frac{1}{\left(\sqrt{\lambda_i}\right)^2} \lambda_1 \qquad (\text{por definición de } \sigma_i) \\
&= \frac{1}{\lambda_i} \lambda_i = 1
\end{aligned}$$

Entonces **$u_1, ..., u_r$ son ortonormales**. Falta definir los u_i con $i > r$:

Como $\dim(Im(A)) = r$ y los $u_i \in Im(A)$, entonces estos vectores conforman una base otornomal de $Im(A)$. Además $Nu(A^t) = Im(A)^\perp$. Entonces toda base ortonormal de $Im(A)$ se puede extender a una base ortonormal de $\mathbb{R}^m$ con vectores que pertenecen a $Nu(A^t)$. Sean $u_{r+1}, ..., u_m$ dicha extensión, entonces **$u_1, ..., u_m$ es una base ortonormal de $\mathbb{R}^m$**.

Veamos que cumplen la relaciones definidas para u_i con $i > r$:

- $Av_i = \sigma_i u_i$ para $i = 1, ..., r$ se cumple por definición de u_i.
- $Av_i = 0$ para $i = r + 1, ..., n$. Como v_i es autovector de $A^t A$ del autovalor 0, entonces $A^t A v_i = 0 \Rightarrow v_i^t A^t A v_i = 0 \Rightarrow \|Av_i\|_2 = 0 \Rightarrow A^t v_i = 0$.
- $A^t u = \sigma_i v$ para $i = 1, ..., r$, se cumple por definición de u_i.
- $A^t u = 0$ para $i = r + 1, ..., n$. Como $u_i \in Nu(A^t)$, entonces $A^t u_i = 0$.

7.3.3 Propiedades

Propiedad

$$\|A\|_2 = \sigma_1$$

Demostración

$$\begin{aligned}
\|A\|_2 &= \max_{\|x\|_2=1} \|Ax\|_2 \\
&= \max_{\|x\|_2=1} \|U\Sigma V^t x\|_2 \\
&= \max_{\|x\|_2=1} \|\Sigma V^t x\|_2 \text{ (porque } U \text{ es ortogonal)} \\
&= \max_{\|x\|_2=1} \|\Sigma x\|_2 \text{ (porque } V \text{ es ortogonal)} \\
&= \max_{\|x\|_2=1} \|\Sigma x\|_2 \\
&= \max_{\|x\|_2=1} \sqrt{\sum_{i=1}^{r} \sigma_i^2 x_i^2} \\
&\leq \max_{\|x\|_2=1} \sigma_1 \sqrt{\sum_{i=1}^{r} x_i^2} \\
&\leq \sigma_1
\end{aligned}$$

Además,

$$\|A\|_2 = \max_{\|x\|_2=1} \|Ax\|_2 \geq \|Av_1\| = \sqrt{v_1^t A v_1} = \sqrt{v_1^t \sigma_1 v_1} = \sigma_1$$

Luego, concluimos que $\|A\|_2 = \sigma_1$ ■

Propiedad

Si $\boldsymbol{A}$ es inversible entonces:

$$\kappa(A) = \frac{\sigma_1}{\sigma_r}$$

Demostración

Como A es inversible, $r = n$ y tiene n valores singulares no nulos. Además, los valores singulares de A^{-1} son $\frac{1}{\sigma_n}, \ldots, \frac{1}{\sigma_1}$ de los valores singulares de A. Por definición de κ:

$$\kappa(A) = \|A\|_2 \, \|A^{-1}\|_2 = \sigma_1 \frac{1}{\sigma_n} = \frac{\sigma_1}{\sigma_n}$$

Propiedad

$$\|A\|_F = \sqrt{\sum_{i=1}^{r} \sigma_i^2}$$

Métodos iterativos para sistemas de ecuaciones lineales

8.1 Introducción

Sean $\boldsymbol{A} \in \mathbb{R}^{n \times n}$ y $\boldsymbol{b} \in \mathbb{R}^{n}$. Un **método iterativo** para resolver el sistema $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$ **es un algoritmo que** produce una secuencia de vectores $\{\boldsymbol{x}^k\}$ que **se espera converja a la solución del sistema**.

Estos métodos, en la práctica, son **muy útiles para sistemas de gran tamaño**, en los que los métodos exactos son muy costosos en términos de tiempo y memoria.

Vamos a estudiar dos métodos iterativos muy conocidos: el **método de Jacobi** y el **método de Gauss-Seidel** que

8.1.1 Método de Jacobi

Sean $\boldsymbol{x}^0 \in \mathbb{R}^n$. Supongamos que $\boldsymbol{a}_{ii} \neq \boldsymbol{0}\ \forall \boldsymbol{i} = \boldsymbol{1} \ldots \boldsymbol{n}$. Definimos $\boldsymbol{x}^1$ de la siguiente manera:

$$
\begin{aligned}
\boldsymbol{x}_1^1 &= \frac{\boldsymbol{b}_1 - \sum_{\substack{j=1 \\ j \neq 1}}^{n} \boldsymbol{a}_{1j} \boldsymbol{x}_j^0}{\boldsymbol{a}_{11}} \\
\boldsymbol{x}_2^1 &= \frac{\boldsymbol{b}_2 - \sum_{\substack{j=1 \\ j \neq 2}}^{n} \boldsymbol{a}_{2j} \boldsymbol{x}_j^0}{\boldsymbol{a}_{22}} \\
&\vdots \\
\boldsymbol{x}_i^1 &= \frac{\boldsymbol{b}_i - \sum_{\substack{j=1 \\ j \neq i}}^{n} \boldsymbol{a}_{ij} \boldsymbol{x}_j^0}{\boldsymbol{a}_{ii}} \\
&\vdots \\
\boldsymbol{x}_n^1 &= \frac{\boldsymbol{b}_n - \sum_{\substack{j=1 \\ j \neq n}}^{n} \boldsymbol{a}_{nj} \boldsymbol{x}_j^0}{\boldsymbol{a}_{nn}}
\end{aligned}
$$

A partir de $\boldsymbol{x}^1$, definimos $\boldsymbol{x}^2, \boldsymbol{x}^3, \ldots$ de la misma manera:

$$
\boldsymbol{x}_i^{k+1} = \frac{\boldsymbol{b}_1 - \sum_{\substack{j=1 \\ j \neq 1}}^{n} \boldsymbol{a}_{ij} \boldsymbol{x}_j^k}{\boldsymbol{a}_{ii}}
$$

Si quisieramos escribir usando operaciones matriciales, debemos encontrar una forma de escribir $\boldsymbol{A}$ un poco más cómoda. Para ello, descompogamosla de la siguiente manera:

$$
\begin{aligned}
\boldsymbol{A} &= \boldsymbol{D} - \boldsymbol{L} - \boldsymbol{U} \\
&= \begin{bmatrix} \boldsymbol{a}_{11} & \ldots & \boldsymbol{0} & \ldots & \boldsymbol{0} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ \boldsymbol{0} & \ldots & \boldsymbol{a}_{ii} & \ldots & \boldsymbol{0} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ \boldsymbol{0} & \ldots & \boldsymbol{0} & \ldots & \boldsymbol{a}_{nn} \end{bmatrix} - \begin{bmatrix} \boldsymbol{0} & \ldots & \boldsymbol{0} & \ldots & \boldsymbol{0} \\ \vdots & \ddots & \ldots & \ddots & \vdots \\ -\boldsymbol{a}_{i1} & \ldots & \boldsymbol{0} & \ldots & \boldsymbol{0} \\ \vdots & \ddots & \ldots & \ddots & \vdots \\ -\boldsymbol{a}_{n1} & \ldots & -\boldsymbol{a}_{nn-1} & \ldots & \boldsymbol{0} \end{bmatrix} - \begin{bmatrix} \boldsymbol{0} & \ldots & -\boldsymbol{a}_{12} & \ldots & -\boldsymbol{a}_{1n} \\ \vdots & \ddots & \ldots & \ddots & \vdots \\ \boldsymbol{0} & \ldots & \boldsymbol{0} & \ldots & -\boldsymbol{a}_{in} \\ \vdots & \ddots & \ldots & \ddots & \vdots \\ \boldsymbol{0} & \ldots & \boldsymbol{0} & \ldots & \boldsymbol{0} \end{bmatrix}
\end{aligned}
$$

Entonces:

$$
\begin{gathered}
\boldsymbol{Ax} = \boldsymbol{b} \\
(\boldsymbol{D} - \boldsymbol{L} - \boldsymbol{U})\boldsymbol{x} = \boldsymbol{b} \\
\boldsymbol{Dx} - (\boldsymbol{L} + \boldsymbol{U})\boldsymbol{x} = \boldsymbol{b} \\
\boldsymbol{Dx} = (\boldsymbol{L} + \boldsymbol{U})\boldsymbol{x} + \boldsymbol{b} \\
\boldsymbol{x} = \boldsymbol{D}^{-1}((\boldsymbol{L} + \boldsymbol{U})\boldsymbol{x} + \boldsymbol{b})\boldsymbol{x} = \boldsymbol{D}^{-1}(\boldsymbol{L} + \boldsymbol{U})\boldsymbol{x} + \boldsymbol{D}^{-1}\boldsymbol{b}
\end{gathered}
$$

Si ahora vemos como es la forma que tiene cada uno de los elementos de $\boldsymbol{x}$, vemos que es igual a la forma al principio de cada sección, por lo que cada iteración de Jacobi se puede escribir como:

$$
\boldsymbol{x}^{k+1} = \boldsymbol{D}^{-1}(\boldsymbol{L} + \boldsymbol{U})\boldsymbol{x}^k + \boldsymbol{D}^{-1}\boldsymbol{b}
$$

8.1.2 Método de Gauss-Seidel

Sea $\boldsymbol{x}^0 \in \mathbb{R}^n$. Supongamos que $a_{ii} \neq 0 \ \forall i = 1 ... n$. Definimos $\boldsymbol{x}^1$ de la siguiente manera:

$$
\begin{gathered}
x_1^1 = \frac{b_1 - \sum_{\substack{j=1 \\ j \neq 1}}^n a_{1j} x_j^0}{a_{11}} \\
x_2^1 = \frac{b_2 - a_{21} x_1^1 - \sum_{j=3}^n a_{2j} x_j^1}{a_{22}} \\
\vdots \\
x_i^1 = \frac{b_i - \sum_{j=1}^{i-1} a_{ij} x_j^1 - \sum_{j=i+1}^n a_{ij} x_j^1}{a_{ii}} \\
\vdots \\
x_n^1 = \frac{b_n - \sum_{j=1}^{n-1} a_{nj} x_j^1}{a_{nn}}
\end{gathered}
$$

De manera similar, los siguientes valores de $\boldsymbol{x}$ se definen como:

$$
x_i^{k+1} = \frac{b_i - \sum_{j=1}^{i-1} a_{ij} x_j^k - \sum_{j=i+1}^n a_{ij} x_j^k}{a_{ii}}
$$

Y usando la descomposición de $\boldsymbol{A}$ en $\boldsymbol{D} - \boldsymbol{L} - \boldsymbol{U}$, podemos despejar $\boldsymbol{x}$ de tal manera que nos permitirá escribir el método de Gauss-Seidel de manera matricial:

$$
\begin{gathered}
\boldsymbol{Ax} = \boldsymbol{b} \\
(\boldsymbol{D} - \boldsymbol{L} - \boldsymbol{U})\boldsymbol{x} = \boldsymbol{b} \\
(\boldsymbol{D} - \boldsymbol{L})\boldsymbol{x} - \boldsymbol{Ux} = \boldsymbol{b} \\
(\boldsymbol{D} - \boldsymbol{L})\boldsymbol{x} = \boldsymbol{Ux} + \boldsymbol{b} \\
\boldsymbol{x} = (\boldsymbol{D} - \boldsymbol{L})^{-1}\boldsymbol{Ux} + (\boldsymbol{D} - \boldsymbol{L})^{-1}\boldsymbol{b}
\end{gathered}
$$

Luego, la iteración de Gauss-Seidel se puede escribir como:

$$
\boldsymbol{x}^{k+1} = (\boldsymbol{D} - \boldsymbol{L})^{-1}\boldsymbol{U}\boldsymbol{x}^k + (\boldsymbol{D} - \boldsymbol{L})^{-1}\boldsymbol{b}
$$

8.2 Análisis de convergencia

Dado un vector inicial $x^0 \in \mathbb{R}^{n \times n}$ y $T \in \mathbb{R}^{n \times n}$ y $c \in \mathbb{R}^n$,

$$x^{k+1} = Tx^k + c$$

En Jacobi, $T = D^{-1}(L + U)$ y $c = D^{-1}b$. En Gauss-Seidel, $T = (D - L)^{-1}U$ y $c = (D - L)^{-1}b$.

Si x^* **es la solución del sistema**, vamos a ver condiciones necesarias para que la secuencia $\{x^k\}$ converja a x^*.

8.2.1 Matriz convergente

Sea $A \in \mathbb{R}^{n \times n}$, decimos que A es convergente si

$$\lim_{k \to \infty} A^k = 0$$

Propiedad

$$\begin{aligned} A \text{ es convergente} &\Leftrightarrow \rho(A) < 1 \\ &\Leftrightarrow \lim_{k \to \infty} \|A_k\| = 0 \quad \text{para toda norma inducida} \\ &\Leftrightarrow \lim_{k \to \infty} A^k x = 0 \quad \text{para todo } x \in \mathbb{R}^n \end{aligned}$$

Propiedad

Si $\rho(A) < 1 \Rightarrow I - A$ es no singular y

$$\sum_{k=0}^{\infty} A^k = (I - A)^{-1}$$

8.2.2 Teorema de convergencia

La sucesión $\{x^k\}$ definida por $x^{k+1} = Tx^k + c$ converge a x^* para cualquier x^0 inicial a la solución del sistema $x = Tx + c$ si y solo si $\rho(T) < 1$.

Demostración

$\Leftarrow$) Como $\rho(T) < 1$, entonces $(I - T)$ es inversible. Entonces

$$x = Tx + c \Leftrightarrow (I - T)x = c \Leftrightarrow x^* = (I - T)^{-1}c$$

Sea la sucesión $x^k = Tx^{k-1} + c$. Entonces podemos expresar a x^{k-1} en función de x^{k-2} quedandonos:

$$x^k = Tx^{k-1} + c = T(Tx^{k-2} + c) + c = T^2 x^{k-2} + Tc + c$$

De manera similar, podemos expresar a x^{k-2} en función de x^{k-3} y así sucesivamente. Entonces, podemos expresar a x^k en función de x^0 de la siguiente manera:

$$x^k = T^k x^0 + T^{k-1}c + T^{k-2}c + \ldots + c = T^k x^0 + (T^{k-1} + T^{k-2} + \ldots + I)c$$

Como $\rho(T) < 1$, entonces $\lim_{k\to\infty} T^k = 0$. Además $\sum_{k=0}^{\infty} T^k = (I - T)^{-1}$. Entonces

$$\lim_{k\to\infty} x^k = \lim_{k\to\infty} T^k x^0 + \lim_{k\to\infty} (T^{k-1} + T^{k-2} + \ldots + I)c = 0 + (I - T)^{-1}c.$$

Luego el límite de la sucesión existe, no depende de x^0 y es igual a $x^* = (I - T)^{-1}c$.

$\Rightarrow)$ Debemos ver que $\rho(T) < 1$ sabiendo que la sucesión converge independientemente del x^0 inicial.

Sabemos que

$$\rho(T) < 1 \Leftrightarrow A \;\text{ es convergente} \Leftrightarrow \lim_{k\to\infty} T^k z = 0 \;\text{ para todo } z \in \mathbb{R}^n$$

Demostremos entonces que $\lim_{k\to\infty} T^k z = 0$ **para todo** $z \in \mathbb{R}^n$:

Sea $z \in \mathbb{R}^n$. Consideremos $x^0 = x^* - z$. con x^* el límite de la suseción $\{x^k\}$:

$$\begin{aligned}
\lim_{k\to\infty} T^k z &= \lim_{k\to\infty} T^k (x^* - x^0) \\
&= \lim_{k\to\infty} T^{k-1}(Tx^* - Tx^0) \\
&= \lim_{k\to\infty} T^{k-1}(x^* - c - Tx^0) \\
&= \lim_{k\to\infty} T^{k-1}(x^* - c - x^1 + c) \\
&= \lim_{k\to\infty} T^{k-1}(x^* - x^1) \\
&= \lim_{k\to\infty} T^{k-2}(Tx^* - Tx^1) \\
&= \lim_{k\to\infty} T^{k-2}(x^* - c - Tx^1) \\
&= \lim_{k\to\infty} T^{k-2}(x^* - c - x^2 + c) \\
&= \lim_{k\to\infty} T^{k-2}(x^* - x^2)
\end{aligned}$$

Si seguimos haciendo los remplazos correspondientes, llegaremos a que

$$\lim_{k\to\infty} T^k z = \lim_{k\to\infty} (x^* - x^k)$$

Este último límite es igual a 0, ya que la sucesión converge a x^*. Por lo tanto, $\rho(T) < 1$.

Propiedad

Si $A \in \mathbb{R}^{n\times n}$ es estrictamente diagonal dominante, entonces el método de Jacobi converge.

Propiedad

Si $\boldsymbol{A} \in \mathbb{R}^{n\times n}$ es estrictamente diagonal dominante, entonces el método de Gauss-Seidel converge.

Propiedad

Si $\boldsymbol{A} \in \mathbb{R}^{n\times n}$ es simétrica definida positiva, entonces el método de Gauss-Seidel converge.

Propiedad

Sea $\boldsymbol{A} \in \mathbb{R}^{n\times n}$ tal que $a_{ij} \leq 0 \; \forall i \neq j$ y $a_{ii} > 0 \; \forall i$. Se satisface una sola de las siguientes propiedades:

- $\rho(T_{GS}) < \rho(T_J) < 1$
- $1 < \rho(T_J) < \rho(T_{GS})$
- $\rho(T_{GS}) = \rho(T_J) = 0$
- $\rho(T_{GS}) = \rho(T_J) = 1$

Propiedad

Sea $\boldsymbol{A} \in \mathbb{R}^{n\times n}$ tal que $a_{ij} \leq 0 \; \forall i \neq j$ y $a_{ii} > 0 \; \forall i$. Entonces, ambos métodos divergen o ambos convergen. En el segundo caso, el método de Gauss-Seidel converge más rápido.

8.2.3 Cota del error

Sea $\boldsymbol{T} \in \mathbb{R}^{n\times n}$ tal que $\|\boldsymbol{T}\| < 1$ para una norma inducida. Entonces:

Propiedad

$x^{k+1} = Tx^k + c$ converge independientemente del x^0 inicial.

Demostración

Hay una propiedad que estables que $|\rho(A)| \leq \|A\|$ para toda norma inducida. Por lo tanto $\rho(T) < \|T\| < 1$ y la sucesión converge independientemente del x^0 inicial (por el teorema de convergencia).

Propiedad

$\|x^* - x^k\| \leq \|T\|^k \; \|x^* - x^0\|$

Demostración

Sabemos que $x^* = Tx^* + c^1$ y que la suceción $x^k = T^k x^{k-1} + c$:

$$\begin{aligned} \| \; x^* - x^k \; \| &= \| \; Tx^* + c - T^k x^{k-1} - c \; \| \\ &= \| \; T\left(x^* - x^{k-1}\right) \; \| \leq \| \; T \; \| \; \| \; x^* - x^{k-1} \; \| \end{aligned}$$

Volviendo a aplicar el mismo razonamiento, llegamos a que:

$$\begin{aligned} \| \; x^* - x^k \; \| &\leq \| \; T \; \| \; \| \; Tx^* + c - T^k x^{k-1} - c \; \| \\ &= \| \; T \; \| \; \| \; T\left(x^* - x^{k-1}\right) \; \| \leq \| \; T\|^2 \; \| \; x^* - x^{k-1} \; \| \end{aligned}$$

Repetimos el proceso k veces y llegamos a que:

$$\| \; x^* - x^k \; \| \leq \| \; T\|^k \; \| \; x^* - x^0 \; \| \; \blacksquare$$

Propiedad

$$\|x^* - x^k\| \leq \frac{\|T\|^k}{1 - \|T\|} \; \|x^1 - x^0\|$$

Demostración

Veamos primero la diferentes dos iteraciones sucesivas de la serie:

$$\begin{aligned} \| \; x^{k+1} - x^k \; \| &= \| \; Tx^{k+1} + c - T^k x^{k-1} - c \; \| \\ &= \| \; T\left(x^k - x^{k-1}\right) \; \| \leq \| \; T \; \| \; \| \; x^k - x^{k-1} \; \| \end{aligned}$$

Si seguimos remplazando, nos queda que:

$$\| \; x^{k+1} - x^k \; \| \leq \| \; T\|^k \; \| \; x^1 - x^0 \; \| \leq$$

Ahora tomamos x^i y x^j con $j > k$:

$$\| \boldsymbol{x}^j - \boldsymbol{x}^k \| \leq \left(\|\boldsymbol{T}\|^{j-1} + \|\boldsymbol{T}\|^{j-2} + \ldots + \|\boldsymbol{T}\|^k\right) \|\boldsymbol{x}^1 - \boldsymbol{x}^0\|$$
$$\leq \|\boldsymbol{T}\|^k \left(\sum_{i=0}^{j-1-k} \|\boldsymbol{T}\|^i \right) \|\boldsymbol{x}^1 - \boldsymbol{x}^0\|$$

Si ahora tomamos el límite cuando $\boldsymbol{j} \to \infty$, como $\{\boldsymbol{x}^j\}_{j=0}^{\infty}$ converge a $\boldsymbol{x}^*$ y

$$\sum_{i=0}^{j-1-k} \|\boldsymbol{T}\|^i = \sum_{i=0}^{\infty} \|\boldsymbol{T}\|^i \quad \textbf{que es la serie geométrica}$$

Entonces, como $\|\boldsymbol{T}\| < \mathbf{1}$, la serie converge a $\frac{1}{1 - \|\boldsymbol{T}\|}$. Por lo tanto,

$$\| \boldsymbol{x}^* - \boldsymbol{x}^k \| \leq \frac{\|\boldsymbol{T}\|^k}{1 - \|\boldsymbol{T}\|} \|\boldsymbol{x}^1 - \boldsymbol{x}^0\| \blacksquare$$

Cuadrados Mínimos Lineales

9.1 El problema

Dado un conjunto de m pares ordenados (x_i, y_i) con $i = 1, 2, l..., m$, **buscamos una función $f(x)$** perteneciente a una familia $\mathcal{F}$ tal que "mejor aproxime" a los datos.

Hay diferentes formas de realizar esta aproximación:

- Se puede intentar minimizar el error máximo, es decir, buscar la función $f(x)$ que minimiza la siguiente función:

  $$\min_{f \in \mathcal{F}} \max_{i=1,...,m} |f(x_i) - y_i|$$

  Sin embargo, al considerar el error máximo, la función obtenida es sensible a valores atípicos que pueden aparecer en la muestra.

- Se puede intentar minimizar la suma de los errores absolutos, es decir, buscar la función $f(x)$ que minimiza la siguiente función:

  $$\min_{f \in \mathcal{F}} \sum_{i=1}^{m} |f(x_i) - y_i|$$

  En este caso, los valores atípicos deja de dominar a la muestra y reflejan mejor la tendencia de los datos. Sin embargo, la función obtenida puede no ser continua o diferenciable en todos los puntos.

- En general, **la forma más común de realizar la aproximación es minimizar la suma de los cuadrados de los errores**. Es decir, buscar la función $f(x)$ que minimiza la siguiente función:

  $$\min_{f \in \mathcal{F}} \sum_{i=1}^{m} (f(x_i) - y_i)^2$$

Dado un conjunto de funciones $\{\phi_1, ..., \phi_n\}$ **linealmente independientes**, la familia $\mathcal{F}$ se define como:

$$\mathcal{F} = \left\{ f(x) = \sum_{j=1}^{n} c_j \phi_{j(x)} \right\}$$

Donde **c_j son los coeficientes** que se deben determinar **para minimizar la función de error**. Entonces:

$$\min_{f \in F} \sum_{i=1}^{m} (f(x_i) - y_i)^2 = \min_{c_1,...,c_n} \sum_{i=1}^{m} \left(\sum_{j=1}^{n} c_j \phi_{j(x_i)} - y_i \right)^2$$

Sean $A \in \mathbb{R}^{m \times n}$, $b \in R^m$ y $x \in \mathbb{R}^n$ definidos como:

$$A = \begin{bmatrix} \phi_1(x_1) & \dots & \phi_{n(x_1)} \\ \vdots & \ddots & \vdots \\ \phi_1(x_m) & \dots & \phi_{n(x_m)} \end{bmatrix} \qquad b = [y_1 \ \dots \ y_m] \qquad x = [c_1 \ \dots \ c_n]$$

Entonces podemos reescribir **el problema de minimización de manera matricial** como:

$$\bar{x} = \min_{x \in \mathbb{R}^n} \|Ax - b\|_2^2$$

Propiedad

Existe $\bar{x}$ solución del problema de cuadrados mínimos lineales (CML).

Demostración

Hay un resultado de álgebra lineal que establece que dada una matriz $A \in \mathbb{R}^{m\times n}$, $\mathbf{Im}(A) \oplus \mathbf{Nu}(A) = \mathbb{R}^m$. Es decir, todo vector de $\mathbb{R}^m$ se puede escribir como la suma de dos vectores, uno en la imagen de A y otro en el núcleo de A^t. Además, el $\mathbf{Nu}(A^t) = \mathbf{Im}(A)^\perp$

Aplicando este resultado al vector b, nos queda que $b = b^1 + b^2$ con $b^1 \in \mathbf{Im}(A)$ y $b^2 \in \mathbf{Nu}(A^t)$. Entonces, el problema de minimización se puede reescribir como:

$$\begin{aligned}
\min_{x\in\mathbb{R}^n} \|Ax - b\|_2^2 &= \min_{x\in\mathbb{R}^n} \|Ax - b^1 - b^2\|_2^2 \\
&= \min_{x\in\mathbb{R}^n} ((Ax - b^1) + b^2)^t((Ax - b^1) - b^2) \quad (\text{por def de norma 2}) \\
&= \min_{x\in\mathbb{R}^n} (Ax - b^1)^t(Ax - b^1) - 2(Ax - b^1)^t b^2 + {b^2}^t b^2
\end{aligned}$$

Como $Ax - b^1 \in \mathbf{Im}(A)$ y $b^2 \in \mathbf{Nu}(A^t) = \mathbf{Im}(A)^\perp$, entonces $(Ax - b^1)^t b^2 = 0$ y:

$$\begin{aligned}
\min_{x\in\mathbb{R}^n} \|Ax - b\|_2^2 &= (Ax - b^1)^t(Ax - b^1) + {b^2}^t b^2 \\
&= \|Ax - b^1\|_2^2 + \|b^2\|_2^2
\end{aligned}$$

Como el segundo termino es constante (no depende de x), el minimo será aquel que minimice $\|Ax - b^1\|_2^2$.

Ahora como $b^1 \in \mathbf{Im}(A)$, entonces existe $\bar{x} \in \mathbb{R}^n$ tal que $A\bar{x} = b^1$. Luego, $\|A\bar{x} - b^1\|_2^2 = 0$ que es lo mínimo que puede valer una norma.

Por lo tanto, siempre existe $\bar{x}$ solución del problema de cuadrados mínimos lineales (CML). ∎

Propiedad

La solución es única si y solo si las columnas de A son linealmente independientes.

Demostración

$\Rightarrow$) Sea $\bar{x}$ una solución única del problema. Supongamos las columnas de A son linealmente dependientes, entonces existe $z \neq 0$ tal que $Az = 0$. Entonces, $A(\bar{x} + z) = A\bar{x} + Az = A\bar{x} = b^1$. Luego, $\bar{x} + z$ también es solución del problema, lo que contradice la unicidad de la solución.

$\Leftarrow$) Supongamos que las columnas de A son linealmente independientes y existen al menos dos soluciones distintas $\bar{x}_1$ y $\bar{x}_2$. Entonces, $A(\bar{x}_1 - \bar{x}_2) = A\bar{x}_1 - A\bar{x}_2 = b^1 - b^1 = 0$. Luego, $\bar{x}_1 - \bar{x}_2$ es un vector no nulo que es solución del sistema homógeneo ($Az = 0$). Esto contradice que las columnas de A sean linealmente independientes. ∎

9.1.1 Ecuaciones normales

La caracterización de la solución de CML depende del valor de $\boldsymbol{b^1}$ (proyección de $\boldsymbol{b}$ en "Im"(A)), sin embargo **nos gustaría conseguir un sistema que dependa de los valores conocidos originales**: $\boldsymbol{A}$ y $\boldsymbol{b}$.

Si $\bar{x} \in \mathbb{R}^n$ es la solución de CML, entonces sabemos que

$$\begin{aligned} A\bar{x} = b^1 - b^2 &\Rightarrow b^1 - A\bar{x} = b^2 \\ &\Rightarrow A^t(b^1 - A\bar{x}) = A^t b^2 \end{aligned}$$

Como $b^2 \in \mathrm{N}(A^t)$ entonces $A^t b^2 = 0$:

$$\begin{aligned} A^t(b - A\bar{x}) = A^t b = 0 &\Rightarrow A^t(b - A\bar{x}) = 0 \\ &\Rightarrow A^t b - A^t A\bar{x} = 0 \\ &\Rightarrow A^t A\bar{x} = A^t b \end{aligned}$$

Este sistema de ecuaciones es conocido como las **ecuaciones normales** del problema de CML.

Esta formulación nos da **$A^t A$ que es una matríz simétrica semidifinida positiva** (y si todas las columnas de $\boldsymbol{A}$ son linealmente independientes es simétrica definida positiva). Gracias a esto podemos utilizar el método de Cholesky para resolver el sistema.

Sin embargo, **debido a la aritmética finita se pueden presentar algunos problemas**. Por ejemplo:

$$A = \begin{bmatrix} 1 & 1 \\ \varepsilon & 0 \\ 0 & \varepsilon \end{bmatrix} \Rightarrow A^t A = \begin{bmatrix} 1+\varepsilon^2 & 1 \\ 1 & 1+\varepsilon^2 \end{bmatrix}$$

En este caso, la matriz original $\boldsymbol{A}$ tiene dos columnas iguales, pero problemas de representación numérica, las columnas guardadas en la computadora no lo son y la representación de $\boldsymbol{A}$ termina teniendo dos columnas linealmente independientes. Osea que **este sistema aproximado tiene una única solución pero el sistema original tiene infinitas soluciones**.

Por otro lado, es posible que $A^t A$ esté mal condicionada, veamos su número de condición relativo a la norma 1:

$$(A^t A)^{-1} = \begin{bmatrix} \frac{1+\varepsilon^2}{\varepsilon^2(2+\varepsilon^2)} & \frac{-1}{\varepsilon^2(2+\varepsilon^2)} \\ -\frac{1}{\varepsilon^2(2+\varepsilon^2)} & \frac{1}{\varepsilon^2(2+\varepsilon^2)} \end{bmatrix}$$

$$\|A^t A\|_1 = \max_{j=1,\dots,n} \sum_{i=1}^{n} |(A^t A)_{i,j}| = 2 + \varepsilon^2$$

$$\|(A^t A)^{-1}\|_1 = \max_{j=1,\dots,n} \sum_{i=1}^{n} |(A^t A)^{-1}_{i,j}| = \frac{2+\varepsilon^2}{\varepsilon^2}$$

$$\kappa(A^t A) = \|A^t A\|_1 \ \|(A^t A)^{-1}\|_1 = \frac{(2+\varepsilon^2)^2}{\varepsilon^2}$$

Entonces, para valores de ε muy pequeños, el número de condición de $A^t A$ es muy grande.

Propiedad

Sea $\boldsymbol{A} \in \mathbb{R}^{\boldsymbol{m} \times \boldsymbol{n}}$ con $\mathbf{rang}(\boldsymbol{A}) = \boldsymbol{n}$. Entonces $\boldsymbol{A}^{\boldsymbol{t}}\boldsymbol{A}$ es inversible y la solución de CML es

$$\bar{\boldsymbol{x}} = \left(\boldsymbol{A}^{\boldsymbol{t}}\boldsymbol{A}\right)^{-\mathbf{1}}\boldsymbol{A}^{\boldsymbol{t}}\boldsymbol{b}$$

9.2 Resolución usando la factorización QR

Sea $\boldsymbol{A} \in \mathbb{R}^{m\times n}$, $\mathbf{rang}(\boldsymbol{A}) = \boldsymbol{n}$, $\boldsymbol{Q}' \in \mathbb{R}^{m\times m}$ matriz ortogonal y $\boldsymbol{R}' \in \mathbb{R}^{m\times n}$ triangular superior tal que $\boldsymbol{AP} = \boldsymbol{QR}$. Además, $\mathbf{rang}(\boldsymbol{A}) = \mathbf{rang}(\boldsymbol{R})$

Sea $\boldsymbol{P} \in \mathbb{R}^{n\times n}$ matriz de permutación tal que $\boldsymbol{AP} = \boldsymbol{QR}$ tal que $\boldsymbol{R}$ es $\boldsymbol{R}'$ reorganizada de la siguiente manera:

$$\boldsymbol{R} = \begin{bmatrix} r_{11} & r_{12} & \dots & r_{1p} & r_{1p+1} & \dots & r_{1n} \\ 0 & r_{22} & \dots & r_{2p} & r_{2p+1} & \dots & r_{2n} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & r_{pp} & r_{pp+1} & \dots & r_{pn} \\ 0 & 0 & \dots & 0 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 0 & 0 & \dots & 0 \end{bmatrix} = \begin{bmatrix} \boldsymbol{R}_1 \\ 0^{n-p} \end{bmatrix}$$

donde $\boldsymbol{R}_1 \in (\boldsymbol{p} \times \boldsymbol{n})$ es la submatriz que toma desde la fila $\mathbf{1}$ hasta la fila $\boldsymbol{p}$ y desde la columna 1 hasta $\boldsymbol{n}$ de $\boldsymbol{R}'$

$$\boldsymbol{AP} = \boldsymbol{Q}\begin{bmatrix} \boldsymbol{R}_1 \\ 0^{n-p} \end{bmatrix}$$

Entonces el problema de Minimos Cuadrados Lineales (MCL) nos queda:

$$\begin{aligned} \min_{x\in\mathbb{R}^n} \|\boldsymbol{Ax} - \boldsymbol{b}\|_2^2 &= \min_{x\in\mathbb{R}^n} \left\|\boldsymbol{Q}^t(\boldsymbol{Ax} - \boldsymbol{b})\right\|_2^2 \\ &= \min_{x\in\mathbb{R}^n} \left\|\boldsymbol{Q}^t\boldsymbol{Ax} - \boldsymbol{Q}^t\boldsymbol{b}\right\|_2^2 \\ &= \min_{x\in\mathbb{R}^n} \left\|\boldsymbol{Q}^t\boldsymbol{APP}^{-1}\boldsymbol{x} - \boldsymbol{Q}^t\boldsymbol{b}\right\|_2^2 \\ &= \min_{x\in\mathbb{R}^n} \left\|\boldsymbol{RP}^{-1}\boldsymbol{x} - \boldsymbol{Q}^t\boldsymbol{b}\right\|_2^2 \end{aligned}$$

Simplifiquemos un poco más el problema, definamos $\boldsymbol{y} = \boldsymbol{P}^{-1}\boldsymbol{x}$ y $\boldsymbol{c} = \left[\begin{smallmatrix} c^p \\ c^{m-p} \end{smallmatrix}\right] = \boldsymbol{Q}^t\boldsymbol{b}$ entonces:

- Si $\boldsymbol{A}$ tiene columnas linealmente independientes, entonces $\mathbf{rang}(\boldsymbol{A}) = \boldsymbol{n} = \boldsymbol{p}$ y $\boldsymbol{P} = \boldsymbol{I}$:

  $$\begin{aligned} \min_{x\in\mathbb{R}^n} \|\boldsymbol{Ax} - \boldsymbol{b}\|_2^2 &= \min_{x\in\mathbb{R}^n} \left\|\boldsymbol{Rx} - \boldsymbol{Q}^t\boldsymbol{b}\right\|_2^2 \\ &= \min_{x\in\mathbb{R}^n} \left\|\boldsymbol{R}_1\boldsymbol{x} - \begin{bmatrix} c_{1..n} \\ c_{n+1...m} \end{bmatrix}\right\|_2^2 \\ &= \min_{x\in\mathbb{R}^n} \|\boldsymbol{R}_1\boldsymbol{x} - \boldsymbol{c}\|_2^2 + \|\boldsymbol{c}^{m-n}\|_2^2 \end{aligned}$$

  El segundo término es una constante, por lo que podemos eliminarlo:

  $$\min_{x\in\mathbb{R}^n} \|\boldsymbol{Ax} - \boldsymbol{b}\|_2^2 = \min_{x\in\mathbb{R}^n} \|\boldsymbol{R}_1\boldsymbol{x} - \boldsymbol{c}\|_2^2 \geq 0$$

  Como $\boldsymbol{R}_1$ es triangular superior con elementos no nulos en ladiagonal, entonces podemos afirmar que existe $\boldsymbol{x}$ tal que $\boldsymbol{R}_1\boldsymbol{x} - \boldsymbol{c} = \mathbf{0}$ y este $\boldsymbol{x}$ es solución del problema de MCL.

- Ahora si $\boldsymbol{A}$ tiene columnas linealmente dependientes, entonces $\mathbf{rang}(\boldsymbol{A}) = \boldsymbol{p} < \boldsymbol{n}$ y

$$
\begin{aligned}
\min_{\boldsymbol{x}\in\mathbb{R}^n} \|\boldsymbol{A}\boldsymbol{x}-\boldsymbol{b}\|_2^2 &= \min_{\boldsymbol{x}\in\mathbb{R}^n} \left\|\boldsymbol{R}\boldsymbol{P}^{-1}\boldsymbol{x}-\boldsymbol{Q}^t\boldsymbol{b}\right\|_2^2 \\
&= \min_{\boldsymbol{y}\in\mathbb{R}^n} \left\|\boldsymbol{R}\boldsymbol{y}-\boldsymbol{Q}^t\boldsymbol{b}\right\|_2^2 \\
&= \min_{\boldsymbol{y}\in\mathbb{R}^n} \left\|\begin{bmatrix}\boldsymbol{R}_1\\ \boldsymbol{0}^{n-p}\end{bmatrix}\begin{bmatrix}\boldsymbol{y}^p\\ \boldsymbol{y}^{n-p}\end{bmatrix}-\begin{bmatrix}\boldsymbol{c}^p\\ \boldsymbol{c}^{m-p}\end{bmatrix}\right\|_2^2 \\
&= \min_{\boldsymbol{y}\in\mathbb{R}^n} \|\boldsymbol{R}_1\boldsymbol{y}^p-\boldsymbol{c}^p\|_2^2+\|\boldsymbol{c}^{m-p}\|_2^2 \\
&= \min_{\boldsymbol{y}\in\mathbb{R}^n} \|\boldsymbol{R}_1\boldsymbol{y}^p-\boldsymbol{c}^p\|_2^2
\end{aligned}
$$

Luego el problema se reduce a encontrar $\boldsymbol{y}^p$ tal que $\boldsymbol{R}_1\boldsymbol{y}^p-\boldsymbol{c}^p=\boldsymbol{0}$ y asignar valores arbitrarios al resto de las coordenadas de $\boldsymbol{y}$. De esta manera, una vez obtenido $\boldsymbol{y}$, se puede resolver el sistema $\boldsymbol{y}=\boldsymbol{P}^{-1}\boldsymbol{x}$

9.3 Resolucion usando la factorización SVD

Sea $A \in \mathbb{R}^{m\times n}$, $\text{rang}(A) = r$, por la descomposición SVD sabemos que existen $U \in \mathbb{R}^{m\times m}$ y $V \in \mathbb{R}^{n\times n}$ matrices ortogonales y $\Sigma \in \mathbb{R}^{m\times n}$ matriz diagonal:

$$\Sigma = \begin{bmatrix} \sigma_1 & 0 & \cdots & 0 & \cdots & 0 \\ 0 & \sigma_2 & \cdots & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sigma_r & \cdots & 0 \\ 0 & 0 & \cdots & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 0 & 0 & 0 \end{bmatrix}$$

y $\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_r > 0$ tal que $A = U\Sigma V^t$. Analicemos el problema de Cuadrados Mínimos Lineales:

$$\begin{aligned} \min_{x\in\mathbb{R}^n} \|Ax - b\|_2^2 &= \min_{x\in\mathbb{R}^n} \left\|U^t(Ax - b)\right\|_2^2 \\ &= \min_{x\in\mathbb{R}^n} \left\|U^t Ax - U^t b\right\|_2^2 \\ &= \min_{x\in\mathbb{R}^n} \left\|\Sigma V^t x - U^t b\right\|_2^2 \end{aligned}$$

Definiendo $y = V^t x$ y $c = \begin{bmatrix} c^r \\ c^{m-r} \end{bmatrix} = U^t b$ entonces:

$$\min_{x\in\mathbb{R}^n} \left\|\Sigma V^t x - U^t b\right\|_2^2 = \min_{y\in\mathbb{R}^n} \|\Sigma y - c\|_2^2$$

- Si A tiene todas sus columnas linealmente independientes entonces $r = n$ y:

  $$\min_{y\in\mathbb{R}^n} \|\Sigma y - c\|_2^2 = \min_{y\in\mathbb{R}^n} \|\Sigma y - c^n\|_2^2 + \|c^{m-n}\|_2^2$$

  El segundo término es una constante y lo mínimo que puede valer el primer término es cero. Como Σ tiene n elementos no nulos, entonces podemos afirmar que existe único y tal que $\Sigma y - c^n = 0$. Basta tomar

  $$y = \frac{c_i^n}{\sigma_i} \quad \text{para } i = 1, \dots, n$$

  y luego resolver el sistema $y = V^t x$ para obtener x.

- Si A tiene columnas linealmente dependientes entonces $r < n$: sea Σ^p la submatriz de Σ con las primeras r filas y r columnas e y^p el vector formado por las primeras r coordenadas de y e y^{n-p} el vector que contiene al resto.

  $$\min_{y\in\mathbb{R}^n} \|\Sigma y - c\|_2^2 = \min_{y\in\mathbb{R}^n} \|\Sigma^p y^p - c^p\|_2^2 + \|c^{m-r}\|_2^2$$

  El segundo término es una constante y lo mínimo que puede valer el primer término es cero. Como Σ^p tiene r elementos no nulos, entonces podemos afirmar que existe único y^p tal que $\Sigma^p y^p - c^p = 0$. Basta tomar

  $$y^p = \frac{c_i^p}{\sigma_i} \quad \text{para } i = 1, \dots, r$$

El resto de las coordenadas de $\boldsymbol{y}$ no influyen en el mínimo y pueden tomar valores arbitarios. De esta manera se pueden obtener las infinitas soluciones $\boldsymbol{x}$ del problema resolviendo el sistema $\boldsymbol{y} = \boldsymbol{V}^t \boldsymbol{x}$.

9.4 Estimación del error

Vamos a proponer una **generalización del número de condición** que funcion para matrices no cuadradas:

$$\chi(A) = \|A\|_2 \; \|(A^tA)^{-1}A^t\|_2$$

Este número lo podremos utilizar para medir la relación entre pequeños cambios del vector b con los cambios en la solución del problema de Cuadrados Mínimos Lineales.

Propiedad

Sea $A \in \mathbb{R}^{m\times n}$ con $\operatorname{rang}(A) = n$. Sean $b, \bar{b} \in \mathbb{R}^m$ tales que $b = b_1 + b_2$, $\bar{b} = \bar{b}_1 + \bar{b}_2$ con $b_1, \bar{b}_1 \in \operatorname{Im}(A)$ y $b_2, \bar{b}_2 \in \operatorname{Nu}(A^t)$. Si $b^1 \neq 0$ entonces:

$$\frac{\|x - \bar{x}\|}{\|\bar{x}\|} = \frac{\|(A^tA)^{-1}A^tb - (A^tA)^{-1}A^t\bar{b}\|_2}{\|(A^tA)^{-1}A^tb\|_2} \leq \chi(A)\frac{\|b^1 - \bar{b}^1\|_2}{\|b^1\|_2}$$

Demostración

Como $\operatorname{rang}(A) = n$, la solución del problema de Cuadrados Mínimos Lineales es única y sabemos que verifica:

$$x = (A^tA)^{-1}A^tb$$
$$\bar{x} = (A^tA)^{-1}A^t\bar{b}$$

Entonces:

$$\begin{aligned}
\|x - \bar{x}\|_2 &= \left\|(A^tA)^{-1}A^tb - (A^tA)^{-1}A^t\bar{b}\right\|_2 \\
&= \left\|(A^tA)^{-1}A^t(b^1 + b^2) - (A^tA)^{-1}A^t\left(\bar{b}^1 + \bar{b}^2\right)\right\|_2 \\
&= \left\|(A^tA)^{-1}A^tb^1 - (A^tA)^{-1}A^t\bar{b}^1\right\|_2 \quad \text{porque } b^2, \bar{b}^2 \in \operatorname{Nu}(A^t) \\
&= \left\|(A^tA)^{-1}A^t\left(b^1 - \bar{b}^1\right)\right\|_2 \\
&\leq \left\|(A^tA)^{-1}A^t\right\|_2\left\|b^1 - \bar{b}^1\right\|_2 \quad \text{por def de norma}
\end{aligned}$$

Por otro lado, $Ax = b^1$ entonces $\|b_1\|_2 = \|Ax\|_2 \leq \|A\|_2 \; \|x\|_2$:

$$\frac{1}{\|x\|_2} \leq \frac{\|A\|_2}{\|b^1\|_2}$$

Multiplicando los dos términos del mismo lado de las desigualdades tenemos:

$$\frac{\left\|(A^tA)^{-1}A^tb - (A^tA)^{-1}A^t\bar{b}\right\|_2}{\|x\|_2} \leq \frac{\|A\|_2\left\|(A^tA)^{-1}A^t\right\|_2\left\|b^1 - \bar{b}^1\right\|_2}{\|b^1\|_2} \blacksquare$$

Propiedad

$$\chi(\boldsymbol{A})^{\mathbf{2}} = \chi(\boldsymbol{A}^{\boldsymbol{t}}\boldsymbol{A})$$

Interpolación

10.1 Definiciones

Dado un conjunto de pares ordenadas de valores (x_i, y_i) para $i = 0, 1, 2, ..., n$, **buscamos un polinomio $P(x)$ de grado menor o igual a n** que interpole a los puntos dados, es decir, **tal que**:

$$P(x_i) = y_i \quad \forall i = 0, 1, 2, ..., n$$

10.1.1 Polinomio de Lagrange

Definamos el polinomio

$$L_{nk} = \prod_{i=0, i \neq k}^{n} \frac{x - x_i}{x_k - x_i}$$

Entonces L_{nk} es un polinomio de grado n tal que

- $L_{nk}(x_i) = 0 \quad \forall i = 0, ..., n \land i \neq k$
- $L_{nk}(x_k) = 1$

Nos falta hacer que L_{nk} sea igual a y_k en x_k. Entonces:

$$P(x) = \sum_{k=0}^{n} y_k L_{nk}(x)$$

Este último polinomio cumple con la condición de interpolación, es decir, $P(x_i) = y_i$ para $i = 0, 1, 2, ..., n$ y es conocido como el **polinomio interpolador**.

Propiedad

Dados (x_i, y_i) para $i = 0, 1, 2, ..., n$, el polinomio interpolador de grado menor o igual a n es único.

Error de interpolación

Sea $f : \mathbb{R} \to \mathbb{R} \in C^{n+1}[a, b]$ (una función $n + 1$ veces derivable definida en el intervalo cerrado $[a, b]$) y sean $x_0, ..., x_n \in [a, b]$.

Consideremos $P(x)$, el polinomio interpolador de grado menor o igual a n que la interpola y $\bar{x} \in [a, b]$ entonces **existe $\xi(\bar{x})$ tal que**:

$$f(\bar{x}) = P(\bar{x}) + \frac{f^{n+1}(\xi(\bar{x}))}{(n + 1)!} \prod_{i=0}^{n} (\bar{x} - x_i)$$

Demostración

Supongamos que P es el polinomio interpolador de f que pasa por los puntos $(x_i, f(x_i))$ para $i = 0, 1, 2, ..., n$.

- Si $\bar{x} = x_k$ para algun $k \in \{0, ..., n\}$ entonces $P(\bar{x}) = f(\bar{x})$ porque para armar P usamos a $\bar{x}$. Por otro lado, $\prod_{i=0}^{n}(\bar{x} - x_i)$ se anula. Entonces no importa que valor tome $\xi(\bar{x})$, la identidad siempre es verdadera.
- Supongamos ahora que $\bar{x}$ no es ninguno de los x_i originales, entonces definimos la función

$$g(t) = f(t) - P(t) - (f(\bar{x}) - P(\bar{x})) \prod_{i=0}^{n} \frac{t - x_i}{\bar{x} - x_i} \quad \text{para } t \in [a, b]$$

Veamos sus propiedades:

Propiedad

$$g \in C^{n+1}[a, b]$$

Demostración

Tenemos que analizar la derivabilidad de cada uno de los términos de $g(t)$:

- $f(t) \in C^{n+1}[a, b]$ por hipótesis.
- $P(t) \in C^{n+1}[a, b]$ porque es un polinomio de grado menor o igual a n.
- $f(\bar{x}) - P(\bar{x})$ por que es resta de funciones en $C^{n+1}[a, b]$.
- $\prod_{i=0}^{n} \frac{t-x_i}{\bar{x}-x_i} \in C^{n+1}[a, b]$ porque es un polinomio.

Propiedad

Si x_k es uno de los puntos usados para armar el polinomio, entonces $g(x_k) = 0$

Demostración

$$\begin{aligned} g(x_k) &= f(x_k) - P(x_k) - (f(\bar{x}) - P(\bar{x})) \prod_{i=0}^{n} \frac{x_k - x_j}{\bar{x} - x_j} \\ &= f(x_i) - f(x_i) - (f(\bar{x}) - P(\bar{x}))(x_k - x_k) \prod_{\substack{j=0 \\ j \neq k}}^{n} \frac{x_i - x_j}{\bar{x} - x_j} \\ &= 0 \end{aligned}$$

Propiedad

$$g(\bar{x}) = 0$$

Demostración

$$
\begin{aligned}
g(\bar{x}) &= f(\bar{x}) - P(\bar{x}) - (f(\bar{x}) - P(\bar{x})) \prod_{i=0}^{n} \frac{\bar{x} - x_j}{\bar{x} - x_j} \\
&= f(\bar{x}) - P(\bar{x}) - (f(\bar{x}) - P(\bar{x})) \\
&= 0
\end{aligned}
$$

Teorema de Rolle

Sea $h : \mathbb{R} \to \mathbb{R}$ una función continua en $[c, d]$ y diferenciable en (c, d) tal que $h(c) = h(d)$, entonces existe $\xi \in (c, d)$ tal que $h'(\xi) = 0$

Sabemos que g tiene al menos $n + 2$ puntos donde se anula. Si ordenamos $x_0, \ldots, x_n, \bar{x}$ de menor a mayor, podemos aplicar el teorema de Roll en cada intervalo definido por dos puntos sucesivos. Entonces, podemos afirmar que g' se anula en al menos un punto en cada intervalo. Por lo tanto podemos afirmar que g' se anula en al menos $n + 1$ puntos.

Siguiendo este mismo razonamiento sobre g' podemos concluir que g'' se anula en al menos n puntos, g''' en al menos $n - 1$ puntos y así sucesivamente hasta llegar a g^{n+1} que se anula en al menos un punto. Este punto depende de los valores de $x_0, \ldots, x_n, \bar{x}$ y se lo denotará como $\xi(\bar{x})$.

Calculemos la derivada de orden $n + 1$ de g:

$$
g^{n+1}(t) = f^{n+1}(t) - P^{n+1}(t) - (f(\bar{x}) - P(\bar{x})) \left(\prod_{i=0}^{n} \frac{t - x_i}{\bar{x} - x_i} \right)^{n+1}
$$

Sabemos que P es un polinomio de grado menor o igual a n, por lo que $P^{n+1}(t) = 0$. Además $\left(\prod_{i=0}^{n} \frac{t - x_i}{\bar{x} - x_i} \right)$ es un polinomio de grado $n + 1$, por lo tanto la derivada de orden $n + 1$ es igual al coeficiente que acompaña a la potencia de orden $n + 1$ que es igual $(n - 1)! \prod_{i=0}^{n} \frac{1}{\bar{x} - x_i}$. Luego:

$$
g^{n+1}(t) = f^{n+1}(t) - (f(\bar{x}) - P(\bar{x}))(n - 1)! \left(\prod_{i=0}^{n} \frac{1}{\bar{x} - x_i} \right)
$$

Si ahora evaluamos esta expresión en $\xi(\bar{x})$ obtenemos que:

$$
\begin{aligned}
g^{n+1}(t) = 0 = f^{n+1}(\xi(\bar{x})) - (f(\bar{x}) - P(\bar{x}))(n-1)! \left(\prod_{i=0}^{n} \frac{1}{\bar{x} - x_i} \right) \\
\Rightarrow f^{n+1}(\xi(\bar{x})) = (f(\bar{x}) - P(\bar{x}))(n-1)! \left(\prod_{i=0}^{n} \frac{1}{\bar{x} - x_i} \right) \\
\Rightarrow \frac{f^{n+1}(\xi(\bar{x}))}{(n+1)!} = (f(\bar{x}) - P(\bar{x})) \left(\prod_{i=0}^{n} \frac{1}{\bar{x} - x_i} \right) \\
\Rightarrow \frac{f^{n+1}(\xi(\bar{x}))}{(n+1)!} \prod_{i=0}^{n} (\bar{x} - x_i) = f(\bar{x}) - P(\bar{x}) \\
\Rightarrow \frac{f^{n+1}(\xi(\bar{x}))}{(n+1)!} \prod_{i=0}^{n} (\bar{x} - x_i) + P(\bar{x}) = f(\bar{x})
\end{aligned}
$$

Luegos llegamos a la expresión propuest al comienzo de esta sección ■

10.2 Diferencias divididas

Si bien el polinomio que encontramos con el método de Lagrange es el único que interpola a los puntos dados, la forma que tenemos de armarlo no es la más eficiente. Además, si queremos agregar un punto más a la interpolación, debemos rehacer todo el cálculo.

Para evitar esto, se utiliza el método de las **diferencias divididas que nos permite conseguir el polinomio interpolador de manera recursiva**.

Definimos la secuencia de **diferencias divididas** como:

- **Orden 0**:

$$f[x_i] = f(x_i)$$

- **Orden 1**:

$$f[x_i, x_{i+1}] = \frac{f[x_{i+1}] - f[x_i]}{x_{i+1} - x_i}$$

- **Orden k**:

$$f[x_i, x_{i+1}, ..., x_{i+k}] = \frac{f[x_{i+1}, x_{i+2}, ..., x_{i+k}] - f[x_i, x_{i+1}, ..., x_{i+k-1}]}{x_{i+k} - x_i}$$

Entonces podemos escribir al polinomio interpolador $P(x)$ como:

$$\begin{aligned} P(x) = f[x_0] + f[x_0, x_1](x - x_0) + f[x_0, x_1, x_2](x - x_0)(x - x_1) \\ +... + f[x_0, x_1, ..., x_n](x - x_0)(x - x_1)...(x - x_{n-1}) \end{aligned}$$

Supongamos ahora que queremos agregar un punto más a la interpolación, es decir, queremos encontrar el polinomio interpolador que tambien tenga en cuenta al punto (x_{n+1}, y_{n+1}). En este caso, **simplemente debemos calcular la diferencia dividida de orden 0 a n para x_{n+1} y agregar el término correspondiente a $P(x)$**.

10.2.1 Demostración

Vamos a demostrar por inducción que el polinomio obtenido con diferencias divididas es el mismo polinomio que se obtiene con el método de Lagrange:

Caso base

Para $n = 1$, el polinomio interpolador de Lagrange es:

$$\begin{aligned} P(x) &= f(x_0)\frac{x - x_1}{x_0 - x_1} + f(x_1)\frac{x - x_0}{x_1 - x_0} \\ &= f(x_0)\frac{x - x_0 + x_0 - x_1}{x_0 - x_1} + f(x_1)(x - x_0) \\ &= f(x_0)\frac{(x_0 - x_1) + (x - x_0)}{x_0 - x_1} + f(x_1)\frac{x - x_0}{x_1 - x_0} \\ &= f(x_0)\underbrace{\frac{x_0 - x_1}{x_0 - x_1}}_{=1} + f(x_0)\frac{x - x_0}{x_0 - x_1} + f(x_1)\frac{x - x_0}{x_1 - x_0} \end{aligned}$$

$$
\begin{aligned}
P(x) &= f(x_0) + f(x_0)\frac{x - x_0}{x_0 - x_1} + f(x_1)\frac{x - x_0}{x_1 - x_0} \\
&= f(x_0) + (x - x_0)\left(\frac{f(x_0)}{x_0 - x_1} + \frac{f(x_1)}{x_1 - x_0}\right) \\
&= f(x_0) + (x - x_0)\left(\frac{-f(x_0)}{x_1 - x_0} + \frac{f(x_1)}{x_1 - x_0}\right) \\
&= f(x_0) + (x - x_0)\frac{f(x_1) - f(x_0)}{x_1 - x_0} \\
&= f(x_0) + (x - x_0)f[x_0, x_1] \quad \blacksquare
\end{aligned}
$$

Caso inductivo

Sea $P_n(x)$ el polinomio interpolador en los puntos $x_0, ..., x_n$, es decir, $P_n(x_i) = f(x_i)$ para $i = 0, 1, 2, ..., n$.

Hipotesis inductiva:

$$
\begin{aligned}
P_n(x) = f[x_0] + f[x_0, x_1](x - x_0) + f[x_0, x_1, x_2](x - x_0)(x - x_1) \\
+ ... + f[x_0, x_1, ..., x_n](x - x_0)(x - x_1)...(x - x_{n-1})
\end{aligned}
$$

Queremos ver que:

Si $P_{n+1}(x)$ el polinomio interpolador en los puntos $x_0, ..., x_n, x_{n+1}$ entonces:

$$P_{n+1} = f[x_0] + ... + f[x_0, ..., x_n](x - x_0)...(x - x_{n-1}) + f[x_0, ..., x_{n+1}](x - x_0)...(x - x_n)$$

Definamos

$$P(x) = P_n(x) + a(x - x_0)...(x - x_n)$$

Entonces P es un polinomio de grado de grado menor o igual a $n + 1$ tal que $P(x_i) = f(x_i)$ para $i = 0, 1, 2, ..., n$. Debemos encontrar el valor de a tal que $P(x_{n+1}) = f(x_{n+1})$ para poder considerarlo como el polinomio interpolador en los puntos $x_0, ..., x_n, x_{n+1}$:

$$
\begin{aligned}
P(x_{n+1}) = f(x_{n+1}) &\Leftrightarrow P_n(x_{n+1}) + a(x_{n+1} - x_0)...(x_{n+1} - x_n) = f(x_{n+1}) \\
&\Leftrightarrow a(x_{n+1} - x_0)...(x_{n+1} - x_n) = f(x_{n+1}) - P_n(x_{n+1}) \\
&\Leftrightarrow a = \frac{f(x_{n+1}) - P_n(x_{n+1})}{(x_{n+1} - x_0)...(x_{n+1} - x_n)}
\end{aligned}
$$

Entonces $P_{n+1} = P$, si demostramos que $a = f[x_0, ..., x_{n+1}]$ habremos demostrado la propiedad.

Consideremos $Q_n(x)$ el polinomio interpolador de los puntos $x_1, ..., x_{n+1}$, por hipoteis inductiva:

$$Q_n(x) = f[x_1] + f[x_1, x_2](x - x_1) + ... + f[x_1, ..., x_{n+1}](x - x_1)...(x - x_n)$$

Definamos

$$Q(x) = Q_n(x) + \frac{x - x_{n+1}}{x_{n+1} - x_0}(Q_n(x) - P_n(x))$$

Entonces Q es un polinomio de grado menor o igual a $n + 1$. Además:

- Q interpola a los puntos $x_1, ..., x_n$.

$$Q(x_i) = Q_n(x_i) + \frac{x_i - x_{n+1}}{x_{n+1} - x_0} \underbrace{(Q_n(x_i) - P_n(x_i))}_{=0} = f(x_i) \quad \forall i = 1, 2, ..., n$$

- $Q(x_{n+1}) = f(x_{n+1})$

$$Q(x_{n+1}) = Q_n(x_{n+1}) + \underbrace{\frac{x_{n+1} - x_{n+1}}{x_{n+1} - x_0}}_{=0}(Q_n(x_{n+1}) - P_n(x_{n+1})) = f(x_{n+1})$$

- $Q(x_0) = P_n(x_0)$

$$\begin{aligned} Q(x_0) &= Q_n(x_0) + \underbrace{\frac{x_0 - x_{n+1}}{x_{n+1} - x_0}}_{=-1}(Q_n(x_0) - P_n(x_0)) \\ &= Q_n(x_0) - (Q_n(x_0) - P_n(x_0)) = P_n(x_0) = f(x_0) \end{aligned}$$

Tanto P como Q son polinomios interpoladores de $x_0, ..., x_{n+1}$, osea que son el mismo polinomio (ya que el polinomio interpolador es único).

Queriamos demostrar que a (el coeficiente que acompaña a la potencia de orden $n + 1$ de P es $f[x_1, ..., x_n + 1]$

El coeficiente $c_{(n,Q)}$ que acompaña a la potencia $n + 1$ de Q es el coeficiente que acompaña a la potencia n de Q_n es:

$$a = c_{(n,Q)} = \frac{c_{(n,Q_n)} - c_{(n,P_n)}}{x_{n+1} - x_0}$$

Pero, por hipotesis inductiva: $c_{(n,Q_n)} = f[x_1, ..., x_{n+1}]$ y $c_{(n,P_n)} = f[x_0, ..., x_n]$. Entonces:

$$a = \frac{f[x_1, ..., x_{n+1}] - f[x_0, ..., x_n]}{x_{n+1} - x_0} = f[x_0, ..., x_{n+1}] \quad \blacksquare$$

10.3 Interpolación de Neville

Sean $f : [a, b] \to \mathbb{R}$ y $(x_i, f(x_i))$ un conjunto de pares ordenados tal que $x_i \in [a, b]$ para $i = 0, \ldots, n$. Sabemos que existe un único polinomio P de grado menor o igual a n tal que $P(x_i) = f(x_i)$ para $i = 0, \ldots, n$.

Buscamos una forma de calcular $P(x)$ en función de dos polinomios que interpolan en un punto menos que nos permita calcular P de manera recursiva.

Notación

Vamos a notar $P_{m_1 m_2 \ldots m_k}$ al polinomio interpolador en los puntos x_{m_1}, x_{m_2}, ..., $x_{m_k} \in [a, b]$.

Propiedad

Dados $x_0, \ldots, x_k \in [a, b]$ e $i, j \in \{0, \ldots, k\}$. Si $P_{01\ldots(j-1)(j+1)\ldots k}$ es el polinomio que interpola todos los puntos salvo x_i y $P_{0\ldots(i-1)(i+1)\ldots k}$ es el polinomio que interpola todos los puntos salvo x_j, entonces el polinomio interpolador $P_{01\ldots k}$ puede expresarse como:

$$P_{01\ldots k}(x) = \frac{(x - x_j)P_{01\ldots(j-1)(j+1)\ldots k}(x) - (x - x_i)P_{0\ldots(i-1)(i+1)\ldots k}(x)}{x_i - x_j}$$

Demostración

Tenemos que ver que $P_{01\ldots k}(x_i) = f(x_i)$ para $i = 0, 1, 2, \ldots, n$.

Sea x_r con $r \in \{0, \ldots, k\}$, entonces:

- **Si $r \neq i, j$:**

$$\begin{aligned}
P(x_r) &= \frac{(x_r - x_j)P_{01\ldots(j-1)(j+1)\ldots k}(x_r) - (x_r - x_i)P_{0\ldots(i-1)(i+1)\ldots k}(x_r)}{x_i - x_j} \\
&= \frac{(x_r - x_j)f(x_r) - (x_r - x_i)f(x_r)}{x_i - x_j} \\
&= \frac{f(x_r)(x_r - x_j - x_r + x_i)}{x_i - x_j} \\
&= f(x_r)\frac{x_i - x_j}{x_i - x_j} \\
&= f(x_r)
\end{aligned}$$

- **Si $r = i$:**

$$
\begin{aligned}
P(x_i) &= \frac{(x_i - x_j)P_{01\ldots(j-1)(j+1)\ldots k}(x_i) - (x_i - x_i)P_{0\ldots(i-1)(i+1)\ldots k}(x_i)}{x_i - x_j} \\
&= \frac{(x_i - x_j)P_{01\ldots(j-1)(j+1)\ldots k}(x_i)}{x_i - x_j} \\
&= (x_j - x_i)P_{01\ldots(i-1)(i+1)\ldots k}(x_i)\frac{)}{x_i - x_j} \\
&= f(x_i)
\end{aligned}
$$

- **Si $r = j$:**

$$
\begin{aligned}
P(x_j) &= \frac{(x_j - x_j)P_{01\ldots(j-1)(j+1)\ldots k}(x_j) - (x_j - x_i)P_{0\ldots(i-1)(i+1)\ldots k}(x_j)}{x_i - x_j} \\
&= \frac{(x_j - x_i)P_{0\ldots(i-1)(i+1)\ldots k}(x_j)}{x_i - x_j} \\
&= P_{0\ldots(i-1)(i+1)\ldots k}(x_j) \\
&= f(x_j)
\end{aligned}
$$

Luego $P = P_{0,\ldots,k}$ ■

Definimos ahora el polinomio Q_{ij} ($i \geq j$) como el polinomio que interpola en los puntos x_{i-j} hasta el punto x_i como:

$$
Q_{ij}(x) = P_{i-j\ldots i}(x) = \frac{(x - x_{i-j})Q_{ij-1} - (x - x_i)Q_{i-1j-1}}{x_i - x_{i-j}}
$$

Entonces:

$$
P_{0,\ldots,n} = Q_{nn} = \frac{(x - x_0)Q_{nn-1} - (x - x_n)Q_{n-1n-1}}{x_n - x_0}
$$

Y para calcularlo podemos armar el siguiente arbol de recursión:

$$
\begin{array}{llllll}
x_0 & f(x_0) = Q_{00}(x) & & & & \\
 & & Q_{11}(x) & & & \\
x_1 & f(x_1) = Q_{10}(x) & & Q_{22}(x) & & \\
 & & Q_{21}(x) & & Q_{33}(x) & \\
x_2 & f(x_2) = Q_{20}(x) & & Q_{32}(x) & & Q_{44}(x) \\
 & & Q_{31}(x) & & Q_{33}(x) & \\
x_3 & f(x_3) = Q_{30}(x) & & Q_{42}(x) & & \\
 & & Q_{41}(x) & & & \\
x_4 & f(x_4) = Q_{40}(x) & & & &
\end{array}
$$

10.4 Interpolación segmentaria

10.4.1 Interpolación lineal segmentaria

Sean (x_i, y_i) con $x_i < x_{i+1}$ para $i = 0, 1, 2, ..., n$. Por cada par de puntos (x_i, y_i) y (x_{i+1}, y_{i+1}) para $i = 0, 1, 2, ..., n - 1$, realizamos una interpolación lineal:

$$L_{i(x)} = a_i x + b_i$$

Entonces por cada x_i, tenemos dos incognitas a_i y b_i y dos ecuaciones:

$$L_i(x_i) = y_i$$
$$L_i(x_{i+1}) = y_{i+1}$$

Entonces tenemos un sistema de $2n$ ecuaciones con $2n$ incognitas en la que cada L_i queda univocamente determinado.

10.4.2 Interpolación cuadrática segmentaria

Sean (x_i, y_i) con $x_i < x_{i+1}$ para $i = 0, 1, 2, ..., n$. Por cada par de puntos (x_i, y_i) y (x_{i+1}, y_{i+1}) para $i = 0, 1, 2, ..., n - 1$, realizamos una interpolación cuadrática:

$$Q_i = a_i + b_i(x - x_i) + c_i(x - x_i)^2$$

Entonces por cada x_i, tenemos tres incognitas a_i, b_i y c_i y dos ecuaciones:

$$Q_i(x_i) = y_i$$
$$Q_i(x_{i+1}) = y_{i+1}$$

Además podemos pedir que la función derivada de Q_i en x_{i+1} sea igual a la derivada de Q_{i+1} en x_{i+1}, es decir:

$$Q_i'(x_{i+1}) = Q_{i+1}'(x_{i+1}) \quad \text{para } i = 0, 1, 2, ..., n - 2$$

Tenemos un sistema de $2n + n - 1$ ecuaciones con $3n$ incognitas, nos falta una ecuación para poder terminar de definir el sistema. La condición que agreamos, en general, se utiliza para controlar el comportamiento de la derivada en alguno de los extremos (x_0 o $x?n$). Sin embargo, esto resulta en una solución asímetrica ya que es imposible pedir condiciones simulteáneas en ambos extremo y mantener la certeza de que el sistema resultante sea compatible.

10.4.3 Interpolación cúbica segmentaria

Sean (x_i, y_i) con $x_i < x_{i+1}$ para $i = 0, 1, 2, ..., n$. Por cada par de puntos (x_i, y_i) y (x_{i+1}, y_{i+1}) para $i = 0, 1, 2, ..., n - 1$, realizamos una interpolación cúbica:

$$S_i(x) = a_i + b_i(x - x_i) + c_i(x - x_i)^2 + d_i(x - x_i)^3$$

Luego la interpolacion $S(x) = S_{i(x)}$ para $x \in [x_i, x_{i+1}]$

Por cada x_i, tenemos cuatro incognitas a_i, b_i, c_i y d_i y las siguientes ecuaciones:

- $S_i(x_i) = y_i$

- $S_i(x_{i+1}) = y_{i+1}$
- $S_i'(x_{i+1}) = S_{i+1}'(x_{i+1})$ para $i = 0, 1, 2, ..., n-1$
- $S_i''(x_{i+1}) = S_{i+1}''(x_{i+1})$ para $i = 0, 1, 2, ..., n-1$

Tenemos un sistema de $4n$ incognitas con $2n + n - 1 + n - 1 = 4n - 2$ ecuaciones, nos falta dos condiciones:

- **Frontera sujeta**: $S_0''(x_0) = f''(x_0)$ y $S_n''(x_n) = f''(x_n)$
- **Frontera natural**: $S_0''(x_0) = 0$ y $S_n''(x_n) = 0$

En ambos casos, el sistema resultante es tridiagonal y estrictamente diagonal dominante, por lo que es posible resolverlo de manera eficiente y tiene solución única.

Demostración

Vamos a demostrar que siempre existe una función que cumple las condiciones impuestas por el sistema de ecuaciones, vamos a escribir en rojo, las incognitas:

Tomemos la primer condición $S_i(x_i) = y_i$ entonces:

$$S_i(x_i) = a_i + b_i(x_i - x_i) + c_i(x_i - x_i)^2 + d_i(x_i - x_i)^3 = a_i = y_i \quad \text{para } i = 0, 1, 2, ..., n-1$$

$$\begin{aligned} S(x_n) = S_{n-1}(x) &= \overbrace{a_{n-1}}^{= f(x_{n-1})} + b_{n-1}(x_n - x_{n-1}) + c_{n-1}(x_n - x_{n-1})^2 + d_{n-1}(x_n - x_{n-1})^3 \\ &= y_{n-1} + b_{n-1}(x_n - x_{n-1}) + c_{n-1}(x_n - x_{n-1})^2 + d_{n-1}(x_n - x_{n-1})^3 = f(x_n) \\ &\Rightarrow b_{n-1}(x_n - x_{n-1}) + c_{n-1}(x_n - x_{n-1})^2 + d_{n-1}(x_n - x_{n-1})^3 = y_n - y_{n-1} \end{aligned}$$

De la segunda condición $S_i(x_{i+1}) = S_{i+1}(x_{i+1})$ para $i = 0, 1, 2, ..., n-2$:

$$\begin{aligned} &S_i(x_{i+1}) = S_{i+1}(x_{i+1}) \\ &\Rightarrow a_i + b_i(x_{i+1} - x_i) + c_i(x_{i+1} - x_i)^2 + d_i(x_{i+1} - x_i)^3 \\ &\qquad = a_{i+1} + b_{i+1}(x_{i+1} - x_{i+1}) + c_{i+1}(x_{i+1} - x_{i+1})^2 + d_{i+1}(x_{i+1} - x_{i+1})^3 \\ &\Rightarrow a_i + b_i(x_{i+1} - x_i) + c_i(x_{i+1} - x_i)^2 + d_i(x_{i+1} - x_i)^3 = a_{i+1} \\ &\Rightarrow y_i + b_i(x_{i+1} - x_i) + c_i(x_{i+1} - x_i)^2 + d_i(x_{i+1} - x_i)^3 = y_{i+1} \end{aligned}$$

De la tercera condición $S_i'(x_{i+1}) = S_{i+1}'(x_{i+1})$ para $i = 0, 1, 2, ..., n-2$:

$$S_i'(x_{i+1}) = S_{i+1}'(x_{i+1})$$

$$\Rightarrow b_i + 2c_i(x_{i+1} - x_i) + 3d_i(x_{i+1} - x_i)^2 = b_{i+1} + 2c_{i+1}(x_{i+1} - x_{i+1}) + 3d_{i+1}(x_{i+1} - x_{i+1})^2$$

$$\Rightarrow b_i + 2c_i(x_{i+1} - x_i) + 3d_i(x_{i+1} - x_i)^2 = b_{i+1}$$

De la cuarta condición $S_i''(x_{i+1}) = S_{i+1}''(x_{i+1})$ para $i = 0, 1, 2, ..., n-2$:

$$\begin{aligned} &S_i''(x_{i+1}) = S_{i+1}''(x_{i+1}) \\ &\Rightarrow 2c_i + 6d_i(x_{i+1} - x_i) = 2c_{i+1} + 6d_{i+1}(x_{i+1} - x_{i+1}) \\ &\Rightarrow 2c_i + 6d_i(x_{i+1} - x_i) = 2c_{i+1} \end{aligned}$$

De la última condición $S_0''(x_0) = 0$ y $S_n''(x_n) = 0$:

$$S_0''(x_0) = 2c_0 = 0$$
$$S_{n-1}''(x_n) = 2c_{n-1} + 6d_{n-1}(x_n - x_{n-1}) = 0$$

Resumiendo, nos quedaron las siguientes ecuaciones:

1. $a_i = y_i$ para $i = 0, 1, 2, ..., n-1$
2. $b_{n-1}(x_n - x_{n-1}) + c_{n-1}(x_n - x_{n-1})^2 + d_{n-1}(x_n - x_{n-1})^3 = y_n - y_{n-1}$
3. $y_i + b_i(x_{i+1} - x_i) + c_i(x_{i+1} - x_i)^2 + d_i(x_{i+1} - x_i)^3 = y_{i+1}$ para $i = 0, 1, 2, ..., n-1$
4. $b_i + 2c_i(x_{i+1} - x_i) + 3d_i(x_{i+1} - x_i)^2 = b_{i+1}$ para $i = 0, 1, 2, ..., n-2$
5. $2c_i + 6d_i(x_{i+1} - x_i) = 2c_{i+1}$ para $i = 0, 1, 2, ..., n-2$
6. $c_0 = 0$
7. $2c_{n-1} + 6d_{n-1}(x_n - x_{n-1}) = 0$

Nombremos $h_i = x_i - x_{i-1}$, entonces:

De (7) podemos despejar

$$d_{n-1} = -\frac{c_{n-1}}{3h_n}$$

De (2) podemos despejar:

$$\begin{aligned} b_{n-1} &= \frac{y_n - y_{n-1} - c_{n-1}h_n^2 - d_{n-1}h_n^3}{h_n} \\ &= \frac{y_n - y_{n-1} - c_{n-1}h_n^2 + c_{n-1}h_n^2 + c_{n-1}\frac{h_n^3}{3h_n}}{h_n} \\ &= \frac{y_n - y_{n-1}}{h_n} - \frac{2}{3}c_{n-1}h_n \end{aligned}$$

De (5) podemos despejar:

$$d_i = \frac{2c_{i+1} - 2c_i}{6h_i} \quad \text{para } i = 0, ..., n-2$$

De (3) podemos despejar:

$$\begin{aligned} b_i &= \frac{y_{i+1} - y_i - c_ih_{i+1}^2 - d_ih_{i+1}^3}{h_{i+1}} \quad \text{para } i = 0, ..., n-2 \\ &= \frac{y_{i+1} - y_i - c_ih_{i+1}^2 + \frac{2c_{i+1} - 2c_i}{6h_i}h_{i+1}^3}{h_{i+1}} \\ &= \frac{y_{i+1} - y_i}{h_{i+1}} - \frac{2}{3}c_ih_{i+1} - \frac{c_{i+1}}{3}h_{i+1} \end{aligned}$$

Finalmente, usamos (4) remplazando los valores despejados:

$$b_i + 2c_ih_{i+1} + 3d_ih_{i+1}^2 = b_{i+1} \quad \text{para } i = 0, ..., n-2$$
$$\frac{y_{i+1} - y_i}{h_{i+1}} - \frac{2}{3}c_ih_{i+1} - \frac{c_{i+1}}{3}h_{i+1} + 2c_ih_{i+1} + 3h_{i+1}\frac{2c_{i+1} - 2c_i}{6h_i} = \frac{y_{i+2} - y_{i+1}}{h_{i+2}} - \frac{y_{i+1} - y_i}{h_{i+1}}$$

$$c_i\left(-\frac{2}{3}h_i + 2h_{i+1} - h_{i+1}\right) + c_{i+1}\left(-\frac{1}{3}h_{i+1} + h_{i+1} + \frac{2}{3}h_{i+2}\right) + c_{i+2}\left(\frac{1}{3}h_{i+2}\right) = \frac{y_{i+2} - y_{i+1}}{h_{i+2}} - \frac{y_{i+1} - y_i}{h_{i+1}}$$

$$\frac{c_i}{3}h_{i+1} + 2\frac{c_{i+1}}{3}(h_{i+1} + h_{i+2}) + \frac{c_{i+2}}{3}h_{i+2} = \frac{y_{i+2} - y_{i+1}}{h_{i+2}} - \frac{y_{i+1} - y_i}{h_{i+1}}$$

Nos queda el caso $i = n - 2$:

$$b_{n-2} + 2c_{n-2}h_{n-1} + 3d_{n-2}h_{n-1}^2 = b_{n-1}$$

$$\frac{y_{n-1} - y_{n-2}}{h_{n-1}} - \frac{2}{3}c_{n-2}h_{n-1} - \frac{c_{n-1}}{3}h_{n-1} + 2c_{n-2}h_{n-1} + 3h_{n-1}^2\frac{2c_{n-1} - 2c_{n-2}}{6h_{n-2}} = \frac{y_n - y_{n-1}}{h_n} - \frac{2}{3}c_{n-1}h_n$$

$$c_{n-2}\left(-\frac{2}{3}h_{n-1} + 2h_{n-1} - h_{n-1}\right) + c_{n-1}\left(-\frac{1}{3}h_{n-1} + h_{n-1} + \frac{2}{3}h_n\right) = \frac{y_n - y_{n-1}}{h_n} - \frac{y_{n-1} - y_{n-2}}{h_{n-1}}$$

$$\frac{c_{n-2}}{3}h_{n-1} + 2\frac{c_{n-1}}{3}(h_{n-1} + h_n) = \frac{y_n - y_{n-1}}{h_n} - \frac{y_{n-1} - y_{n-2}}{h_{n-1}}$$

Luego la matriz asociada al sistema es:

$$\begin{array}{c}
\begin{matrix} c_0 & c_1 & c_2 & c_3 & \dots & c_i & c_{i+1} & c_{i+2} & \dots & c_{n-2} & c_{n-1} \end{matrix} \\
\begin{bmatrix}
1 & 0 & 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 & 0 \\
\frac{1}{3}h_1 & \frac{2}{3}(h_1 + h_2) & \frac{1}{3}h_2 & 0 & \dots & 0 & 0 & 0 & \dots & 0 & 0 \\
0 & \frac{1}{3}h_2 & \frac{2}{3}(h_2 + h_3) & \frac{1}{3}h_3 & \dots & 0 & 0 & 0 & \dots & 0 & 0 \\
\vdots & \vdots & \vdots & \vdots & \dots & \vdots & \vdots & \vdots & \dots & \vdots & \vdots \\
0 & 0 & 0 & 0 & \dots & \frac{1}{3}h_{i+1} & \frac{2}{3}(h_{i+1} + h_{i+2}) & \frac{1}{3}h_{i+2} & \dots & 0 & 0 \\
\vdots & \vdots & \vdots & \vdots & \dots & \vdots & \vdots & \vdots & \dots & \vdots & \vdots \\
0 & 0 & 0 & 0 & \dots & 0 & 0 & 0 & \dots & \frac{1}{3}h_{n-1} & \frac{2}{3}(h_{n-1} + h_n)
\end{bmatrix}
\end{array}$$

Que resulta ser estrictamente diagonal dominante, por lo que existe una solución y es única. Al obtener todos los c_i podemos calcular el resto de las incognitas que quedaron expresadas en función de estos coeficientes.

Bibliografía

11.1 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

11.2 Enlaces

- Métodos Numéricos, CubaWiki

11.3 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