

# Numérica

Abel Doñate  
abel.donate@estudiantat.upc.edu

## Contents

|          |   |          |
|----------|---|----------|
| <b>1</b> | <b>Sistemas de ecuaciones</b>                   | <b>2</b> |
| 1.1      | Gauss . . . . .                                 | 2        |
| 1.1.1    | Gauss sin pivotamiento (GSP) . . . . .          | 2        |
| 1.1.2    | Gauss con pivotamiento (GCP) . . . . .          | 2        |
| 1.1.3    | Matrices en banda . . . . .                     | 2        |
| 1.2      | Métodos de Factorización . . . . .              | 2        |
| 1.2.1    | Descomposición LU . . . . .                     | 2        |
| 1.2.2    | Descomposición de Crout . . . . .               | 3        |
| 1.2.3    | Descomposición de Cholesky . . . . .            | 3        |
| 1.3      | Ortogonalidad . . . . .                         | 3        |
| 1.3.1    | Descomposición QR . . . . .                     | 3        |
| 1.3.2    | Descomposición SVD . . . . .                    | 3        |
| 1.4      | Aplicaciones Gauss . . . . .                    | 3        |
| 1.5      | Minimización de funciones cuadráticas . . . . . | 3        |
| 1.6      | Multiplicadores de Lagrange . . . . .           | 4        |
| 1.7      | Regresión lineal . . . . .                      | 4        |
| <b>2</b> | <b>Errores</b>                                  | <b>4</b> |
| 2.1      | Aritmética finita y precisión . . . . .         | 4        |
| 2.2      | Normas de matrices . . . . .                    | 4        |
| <b>3</b> | <b>Valores propios</b>                          | <b>5</b> |
| 3.1      | Teorema Espectral . . . . .                     | 5        |
| 3.2      | Propiedades . . . . .                           | 5        |
| 3.2.1    | Deflación . . . . .                             | 5        |
| 3.2.2    | Traslación . . . . .                            | 5        |
| 3.2.3    | Cociente de Rayleigh . . . . .                  | 5        |
| 3.3      | IVD e IVI . . . . .                             | 5        |
| <b>4</b> | <b>Métodos iterativos</b>                       | <b>5</b> |
| 4.1      | Métodos Iterativos Estacionarios . . . . .      | 5        |
| 4.1.1    | Método de Jacobi . . . . .                      | 6        |
| 4.1.2    | Método de Gauss-Seidel . . . . .                | 6        |
| 4.2      | Métodos no estacionarios . . . . .              | 6        |
| 4.2.1    | Descenso del Gradiente . . . . .                | 6        |
| 4.2.2    | Gradientes conjugados . . . . .                 | 6        |
| 4.3      | Precondicionamiento . . . . .                   | 6        |

# 1 Sistemas de ecuaciones

## 1.1 Gauss

### 1.1.1 Gauss sin pivotamiento (GSP)

No hace falta pivotar ningún elemento con una permutación. Podemos encontrarnos con dos condiciones suficientes:

- **Definida positiva**

Usar criterio de Sylvester (menores principales  $> 0$ )

- **Diagonalmente Dominantes**

Los elementos de la diagonal son mayores que la suma de los de la fila

En cada paso se realiza  $A^{(k)} = G^{(k)} A^{(k-1)}$ , análogo a hacer el cambio  $f_i \leftarrow f_i - l_{k+i,k} f_k$ , donde

$$G^{(k)} = \begin{pmatrix} I_{k-1} & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -\bar{l} & I_{n-k} \end{pmatrix}, \quad \bar{l} = \begin{pmatrix} l_{k+1,k} \\ l_{k+2,k} \\ \vdots \end{pmatrix}, \quad l_{k+i,k} = \frac{a_{k+i,k}}{a_{k,k}}$$

Observe que en cada paso el determinante se conserva, así como los menores principales.

La complejidad del algoritmo es  $\sim \frac{2}{3}n^3$

### 1.1.2 Gauss con pivotamiento (GCP)

Cuando  $a_{k,k}^{(k)} = 0$ , entonces tenemos que pivotar con una matriz de permutación. Elegiremos el mayor número en valor absoluto de la columna  $k$ . Si ese elemento es  $a_{2,4}$  Su matriz de permutación es:

$$P = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}, \quad \det(P) = -1 \implies PA = LU$$

### 1.1.3 Matrices en banda

Para matrices que tengan un semiancho de banda  $s$  podemos hacer los mismos pasos en Gauss normal, pero haciendo los dos bucles internos de longitud  $s$ , por lo que la complejidad es de  $\sim 2ns^2$ .

#### Propiedades:

- La inversa de una matriz con semiancho de banda inferior  $s$  es una matriz con semiancho de banda inferior  $s$ . Análogamente con el superior

## 1.2 Métodos de Factorización

### 1.2.1 Descomposición LU

Aprovechamos estas propiedades:

- La inversa de una matriz triangular inferior es triangular inferior
- El producto de dos triangulares inferiores también es triangular inferior
- $(G^{(k)})^{-1} = 2I_n - G^{(k)}$  (unos en la diagonal pero a  $\bar{l}$  se le cambia el signo)

Ahora sabiendo invertimos las  $G^{(k)}$  y las multiplicamos:

$$L = (G^{(1)})^{-1} \dots (G^{(n)})^{-1} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ l_{21} & 1 & 0 & 0 & 0 \\ l_{31} & l_{32} & 1 & 0 & 0 \\ l_{41} & l_{42} & l_{43} & 1 & 0 \\ l_{51} & l_{52} & l_{53} & l_{54} & 1 \end{pmatrix} \implies A = LU$$

Complejidad  $\sim \frac{2}{3}n^3$

### 1.2.2 Descomposición de Crout

Igual que LU, pero la matriz U ahora con unos en la diagonal y podemos añadir  $D$  diagonal tal que  $A = LDU$ .

Complejidad  $\sim \frac{2}{3}n^3$

### 1.2.3 Descomposición de Cholesky

Se explota el hecho de que no solo  $A = LU$ , sino que  $A_{[k]} = L_{[k]}U_{[k]}$ . Por ello podemos realizar un sistema por inducción.

Se realiza para matrices simétricas  $A = LL^T \implies A$  SDP (Simétrica definida positiva)

$$L = (l_{ij}) : \implies \begin{cases} \text{si } i = j \rightarrow l_{ii} = \sqrt{a_{ii} - \sum_{k < i} l_{ik}^2} \\ \text{si } i \neq j \rightarrow l_{ij} = \frac{a_{ij} - \sum_{k < j} l_{ik}l_{jk}}{l_{jj}} \end{cases}$$

El método de Cholesky generalizado descompone la matriz simétrica  $A = LDL^T$  con  $l_{ii} = 1$ . La complejidad es  $\sim \frac{1}{3}n^3$

## 1.3 Ortogonalidad

### 1.3.1 Descomposición QR

Se puede realizar para matrices no cuadradas.  $A_{m \times n} = Q_{m \times m}R_{m \times n}$ , con  $Q$  ortogonal ( $Q^T Q = I$ ) y  $R$  triangular superior.

El algoritmo se hace a partir de matrices de Householder:

$$P = I - 2 \frac{vv^T}{v^T v}$$

que cumple las siguientes propiedades:

- $P$  ortogonal y simétrica
- Sea  $v = x \pm \|x\|e^1 \implies Px = \mp \|x\|e^1$

Cogiendo  $x = A_{[k:n, k]}$ :

$$H_k = \begin{pmatrix} \mathbf{I}_{k-1} & \mathbf{0} \\ \mathbf{0} & \mathbf{P}_k \end{pmatrix}$$

$$H_m H_{m-1} \cdots H_1 A = R \implies Q^T = H_m H_{m-1} \cdots H_1 \implies A = QR$$

Si la descomposición viene dada tenemos una complejidad de  $\sim 3n^2$ .

Si la tenemos que calcular es  $\sim \frac{2}{3}n^3$ .

### 1.3.2 Descomposición SVD

$A_{n \times m} = U_{n \times n} D_{n \times m} V_{m \times m}^T$  con  $V, U$  ortogonales y  $D$  diagonal (valores singulares ordenados).

Si la descomposición viene dada tenemos una complejidad de  $\sim 4n^2$

## 1.4 Aplicaciones Gauss

### 1.5 Minimización de funciones cuadráticas

Queremos buscar el mínimo de la función  $f(x) = \frac{1}{2}x^T A x - x^T b$ .

$$\nabla f = \frac{1}{2}(A + A^T)x - \bar{b} = 0 \implies \frac{1}{2}(A + A^T)x = \bar{b}$$

## 1.6 Multiplicadores de Lagrange

Queremos buscar el mínimo de la función  $f(x)$  con la restricción  $g(x) = 0$ . Para ello definimos el Lagrangiano como:

$$\mathcal{L}(x, \lambda) = f(x) - \lambda g(x) \implies \nabla \mathcal{L} = 0$$

## 1.7 Regresión lineal

Quereos sacar la regresión lineal de los puntos  $Ax = b$ . Para ello minimizamos la función  $f(x) = \|Ax - b\|^2 = (x^T A^T - b^T)(Ax - b)$ :

$$\nabla f(x) = \nabla(x^T A^T - b^T)(Ax - b) = 2A^T Ax - 2A^T b = 0 \implies A^T Ax = A^T b$$

También se puede hacer la SVD o QR, ya que las matrices son ortogonales y no cambian la norma. Hacemos los cambios de variable  $y = V^T x, c = U^T b$ , y nos queda  $\|Dy - c\|$ , que será mínimo cuando  $Dy = c$ .

Para hacerlo con la QR tenemos:

$$\|Ax - b\| = \|QRx - b\| = \|Rx - Q^T b\|$$

Resolviendo el sistema cuadrado  $m \times m$  sin tener en cuenta las últimas componentes, tenemos que  $R_{[m \times m]} x = (Q^T b)_{[m]}$

## 2 Errores

### 2.1 Aritmética finita y precisión

Almacenamiento de números

En los doubles hay 64 bits  $\begin{cases} 52 \text{ para los dígitos} \\ 10 \text{ para el exponente} \\ 2 \text{ para los signos} \end{cases} \implies 2, 1 = (0, 10000011)_2 2^{(10)_2}$

Por tanto los números máximo y mínimo que podemos almacenar son  $M = 0.1 \times 10^{308}, m = 0.56 \times 10^{-308}$

**Errores**

$$\text{Error Absoluto } E_x = x - \bar{x}, \quad \text{Error Relativo } r_x = \frac{E_x}{x} \approx \frac{E_x}{\bar{x}}$$

**Propagación de errores**

Error absoluto para la suma  $E_s = E_x + E_y$

Error absoluto para la resta  $E_r = E_x - E_y$

### 2.2 Normas de matrices

$$\text{Norma de una matriz} \implies \|A\| := \max_{v \neq 0} \frac{\|Av\|}{\|v\|} = \max_{\|v\|=1} \|Av\|$$

Propiedades

- $\|Av\| \leq \|A\| \|v\|$
- $\|AB\| \leq \|A\| \|B\|$
- $\|A + B\| \leq \|A\| + \|B\|$

Estudiamos sistema  $A(x + \delta x) = b + \delta b \implies A\delta x = \delta b$  partiendo del sistema  $Ax = b$

Definiendo el condicionamiento de una matriz como  $\kappa(A) := \|A\| \|A^{-1}\| = \frac{|\lambda_{max}|}{|\lambda_{min}|}$  ( $A$  SDP) y sabiendo que  $r_x = \frac{\|\delta x\|}{\|x\|}$  y  $r_b = \frac{\|\delta b\|}{\|b\|}$

$$r_x \leq \kappa(A) r_b$$

Si  $\kappa(A)$  es muy grande, el error de  $x$  puede ser considerablemente grande.

## 3 Valores propios

### 3.1 Teorema Espectral

**Th:** Si  $A$  es simétrica, diagonaliza en base ortonormal con vaps reales.

**Th:** Si  $Kv = \lambda Mv$  con  $K$  simétrica y  $M$  SDP  $\implies \begin{cases} \lambda_1 \leq \dots \leq \lambda_n \text{ VAPs reales} \\ \text{VEPs M-ortonormales } u_i^T M u_i = \delta_{ij} \\ \text{VEPs K-ortogonales } u_i^T K u_i = \lambda_i \delta_{ij} \end{cases}$

### 3.2 Propiedades

#### 3.2.1 Deflación

Si  $Au_k = \lambda_k u_k$ :

$$B := A - \lambda_k u_k u_k^T \text{ verifica } \begin{cases} Bu_i = \lambda_i u_i, i \neq k \\ Bu_k = 0 \end{cases}$$

Si  $Ku_k = \lambda_k M u_k$

$$K^* := K - \lambda_k u_k u_k^T M \text{ verifica } \begin{cases} K^* u_i = \lambda_i M u_i, i \neq k \\ K^* u_k = 0 \end{cases}$$

#### 3.2.2 Traslación

$A$  tiene VEPs de VAP  $\lambda_i \implies A - pI$  tiene los mismos VEPs de VAP  $\lambda_i - p$

$Kv = \lambda Mv$  tiene VEPs de VAP  $\lambda_i \implies (K - pM)v = \lambda Mv$  tiene los mismos VEPs de VAP  $\lambda_i - p$

#### 3.2.3 Cociente de Rayleigh

$$\rho(v) := \frac{v^T A v}{v^T v} = \frac{\sum \alpha_i^2 \lambda_i}{\sum \alpha_i^2} \implies \begin{cases} \lambda_1 \leq \rho(v) \leq \lambda_n \\ \rho(\alpha u_i) = \lambda_i \end{cases}$$

Si  $Kv = \lambda Mv$ :

$$\text{Descomponemos como } \begin{cases} A^* = L^{-1} K L^{-T} \\ v^* = L^T v \end{cases} \implies \rho(v) := \frac{v^T K v}{v^T M v}$$

## 3.3 IVD e IVI

Si queremos encontrar el VEP de VAP mayor en valor absoluto hacemos IVD con  $A$ , si queremos el mínimo lo hacemos con  $A^{-1}$ :

$$v^{k+1} = \frac{A v^k}{\|A v^k\|} \implies v^k = \frac{A^k v^0}{\|A^k v^0\|}, \quad \textbf{Generalizado: } v^{k+1} = \frac{M^{-1} K v^k}{\|M^{-1} K v^k\|_M}$$

## 4 Métodos iterativos

Los métodos iterativos se usan cuando la dimensión de la matriz es muy grande y tiene muchos coeficientes nulos.

Definimos el residuo  $r(x) := Ax - b$  y el error  $e^k := x^* - x^k$

Definimos  $A = L_A + D_A + U_A$  la descomposición de la matriz

### 4.1 Métodos Iterativos Estacionarios

Genera la sucesión de aproximaciones  $x^{k+1} = Gx^k + h$

- Si  $I - G$  es regular  $\implies \exists! x^*$  punto fijo
- Sistemas de la forma  $x^{k+1} = x^k - C^{-1}(Ax^k - b)$

- Convergente  $\iff \lim e^k = 0$ .  $e^{k+1} = Ge^k \iff \lim G^k = 0 \iff$  Radio espectral  $\rho(G) := \max |\lambda_i| < 1$
- Radio espectral  $\rho(G) := \max |\lambda_i| < 1$

#### 4.1.1 Método de Jacobi

Corresponde a  $C = D_A \implies G = I - D_A^{-1}A$   
 Convergente si  $A$  diagonalmente dominante

#### 4.1.2 Método de Gauss-Seidel

Corresponde a  $C = D_A + L_A \implies G = I - (D_A + L_A)^{-1}A$   
 Convergente si  $\begin{cases} A \text{ diagonalmente dominante} \\ \text{ó} \\ A \text{ SDP} \end{cases}$

### 4.2 Métodos no estacionarios

#### 4.2.1 Descenso del Gradiente

Consideramos la ecuación cuadrática  $\frac{1}{2}x^T Ax - x^T b$ , cuyo mínimo se encuentra en  $Ax^* = b$   
 En cada paso cogemos la dirección de máximo descenso  $r = \nabla\phi$  y realizamos la optimización en esa dirección:  
 Inicializamos en  $x^0$  y seguimos el algoritmo:

$$\begin{aligned} (1) \quad & r^k = \nabla\phi(x^k) = Ax^k - b \\ (2) \quad & \alpha_k = -\frac{\langle r^k, r^k \rangle}{\langle r^k, Ar^k \rangle} \\ (3) \quad & x^{k+1} = x^k + \alpha_k r^k \end{aligned}$$

#### 4.2.2 Gradientes conjugados

Queremos que cada dirección sea  $A$ -ortogonal a las demás:  
 Inicializamos  $x^0, r^0 = Ax^0 - b, p^0 = r^0$  y seguimos el algoritmo:

$$\begin{aligned} (1) \quad & q^k = Ap^k \\ (2) \quad & \alpha_k = -\frac{\langle p^k, r^k \rangle}{\langle p^k, q^k \rangle} \\ (3) \quad & x^{k+1} = x^k + \alpha_k p^k \\ (4) \quad & r^{k+1} = Ax^{k+1} - b (= r^k + \alpha_k q^k) \\ (5) \quad & \beta_{k+1} = -\frac{\langle r^{k+1}, q^k \rangle}{\langle p^k, q^k \rangle} \\ (6) \quad & p^{k+1} = r^{k+1} + \beta_{k+1} p^k \end{aligned}$$

Complejidad  $\sim 2Nn^2$

Propiedades:

- $\langle r^k, p^i \rangle = 0, i < k$
- $\langle r^i, r^i \rangle = \langle r^i, p^i \rangle, i < k$
- $\langle r^k, r^i \rangle = 0, i < k$  (residuos ortogonales)
- $\langle p^k, Ap^i \rangle = 0, i < k$  (direcciones  $A$ -conjugadas)

### 4.3 Precondicionamiento

Sea  $\tilde{A} = P^{-\frac{1}{2}}AP^{-\frac{1}{2}}, \tilde{x} = P^{-\frac{1}{2}}x, \tilde{b} = P^{-\frac{1}{2}}b$ :

$$A \text{ SDP}, P \text{ SDP} \implies \tilde{A} \text{ SDP}, \quad \tilde{A}\tilde{x} = \tilde{b} \iff Ax = b$$