Métodos Numéricos

- Si E e.v. diremos que una app $\|\cdot\|: E \to \mathbb{R}$ es una norma en E si verifica:

n1) $x \in E \Rightarrow \|x\| \geq 0$; $\|x\| = 0 \iff x = 0$

n2) $x, y \in E \Rightarrow \|x+y\| \leq \|x\| + \|y\|$ (desigualdad triangular)

n3) $x \in E$, $\lambda \in \mathbb{R} \Rightarrow \|\lambda x\| = |\lambda| \, \|x\|$

- $E = \mathbb{R}^N$ $p \geq 1$

$$\|x\|_p := \left(\sum_{j=1}^{N} |x_j|^p \right)^{1/p} \quad x \in \mathbb{R}^N$$

cuando $p=1$ se denomina norma 1 y es la suma de los valores absolutos.

cuando $p=2$ se denomina norma euclídea.

- $E = \mathbb{R}^N$ se denomina norma del máximo (Tomar el mayor valor absoluto)

$$\|x\|_\infty := \max_{j=1,\dots,N} |x_j| \quad x \in \mathbb{R}^N$$

- $E = \mathbb{R}^{M \times N}$ se denomina norma de Frobenius

$$\|A\|_F := \sqrt{\sum_{i=1}^{M} \sum_{j=1}^{N} a_{ij}^2} \quad A \in \mathbb{R}^{M \times N}$$

- $E = C([a,b])$ se denomina norma del máximo

$$\|f\|_\infty := \max_{a \leq x \leq b} |f(x)| \quad f \in C([a,b])$$

- $E = C^k([a,b])$ $k \in \mathbb{N}$ se denomina norma

$$\|f\|_k := \max_{j=0,1\dots k} \|f^{(j)}\|_\infty \quad f \in C^k([a,b])$$

- Error absoluto $\|x^* - x\|$
- Error relativo $\dfrac{\|x^* - x\|}{\|x\|}$

- Distancia entre dos vectores $dist(x,y) := \|x - y\|$

- $x \in \mathbb{R}^N \Rightarrow \|x\|_\infty \leq \|x\|_1 \leq N \|x\|_\infty$

- $\|\cdot\|$, $\|\cdot\|_*$ equivalentes si $\exists \, c_1, c_2 > 0$: $c_1 \|x\| \leq \|x\|_* \leq c_2 \|x\|$

- Todas las normas en un esp normado finito dimensional son equivalentes.

- Sea $M, N \in \mathbb{N}$, consideremos sendas normas en $\mathbb{R}^N$ y $\mathbb{R}^M$

$$\|A\| := \sup_{x \in \mathbb{R}^N,\ \|x\|=1} \|Ax\| \qquad \text{(Def. de inducción)}$$

- $A \in \mathbb{R}^{M \times N}$ $\qquad \|A\| = \sup_{x \in \mathbb{R}^N,\ x \neq 0} \dfrac{\|Ax\|}{\|x\|}$, en part. $\|Ax\| \leq \|A\| \|x\|$

- $\|\cdot\|_1$ inducida en $\mathbb{R}^{M \times N}$

$$\|A\|_1 = \max_{j=1 \dots N} \sum_{i=1}^{M} |a_{ij}|$$

Suma de los valores abs de las columna y escoger el mayor

$1 \quad \infty$

- $\|\cdot\|_\infty$ inducida en $\mathbb{R}^{M \times N}$

$$\|A\|_\infty = \max_{i=1 \dots M} \sum_{j=1}^{N} |a_{ij}|$$

Suma de los valores abs de las filas y escoges el mayor.

- $A \in \mathbb{R}^{M \times N} \Rightarrow \|A\|_1 = \|A^T\|_\infty$

- Radio Espectral: $\rho(A) := \max\{ |\lambda| : \lambda \in \mathbb{C} \wedge \det(A - \lambda I) = 0 \}$

- La norma euclídea en $\mathbb{R}^N$ también se puede expresar como

$$\|x\|_2 = \sqrt{x^T x} \qquad x \in \mathbb{R}^N$$

- Si $A \in \mathbb{R}^{M \times N}$ entonces $\|A\|_2 = \sqrt{\rho(A^T A)}$

- Una norma en $\mathbb{R}^{N \times N}$, esto es, cuadrada, se dice <u>matricial</u>

$$A, B \in \mathbb{R}^{N \times N} \Rightarrow \|AB\| \leq \|A\| \|B\|$$

- Norma en $\mathbb{R}^{N \times N}$ inducida por una norma en $\mathbb{R}^N$ es matricial.

- Sea $A \in \mathbb{R}^{N \times N}$ se cumple $\lim_{n \to \infty} A^n = 0 \iff \rho(A) < 1$.

- Un problema está bien planteado cuando es <u>unisolvente</u> y <u>estable</u>

i) Unisolvente $\iff \exists!\, x_0 \in X : f(x_0) = y_0$

ii) estable $\iff$ x_0 depende continuamente de los datos y_0.

- (P) $\forall\, y \in Y$ unisolvente $\rightarrow$ inversa de f, resolvente, g
- Estabilidad $\rightarrow$ a pequeñas variaciones de datos y_0 corresponden pequeñas perturbaciones de la solución x_0.
- ¿Estabilidad (P) $\Leftrightarrow$ continuidad de g?

Pequeños cambios de los datos $y \in \mathbb{R}_+ \rightarrow$ cambios cercanos de las correspondientes soluciones $x \in \mathbb{R}_+$ pero no controlados, se aproximan a velocidad diferente.

- Estabilidad $\rightarrow$ condición que fuerce un control de los valores de las soluciones en función de los datos, de manera que pequeñas perturbaciones de y_0 generen perturbaciones pequeñas y controladas de x_0.
- Estabilidad $\sim$ lipschitzianidad local.
- Sean X, Y subconjuntos $\neq \emptyset$ de esp. normados. $g: Y \rightarrow X$ una app y $y_0 \in Y$. Diremos que g es estable en y_0:

$$\exists\, \delta, M > 0 : \sup_{y \in Y,\, 0 < \|y - y_0\| < \delta} \frac{\|g(y) - g(y_0)\|}{\|y - y_0\|} \leq M$$

g es estable si lo es en todos y cada uno de los elementos de Y.

- El problema (P) es estable en $y_0 \in Y$ si su resolvente $g: Y \rightarrow X$ lo es en dicho punto, y es estable si lo es en cualquier dato de Y.
- g estable en $y_0 \Rightarrow g$ continua en y_0.

$\not\Leftarrow$ contraej: raíz en 0.

- Cociente entre el error relativo cometido cerca de $g(y_0)$ y el error relativo de y_0.

$$\left| \frac{\frac{g(y) - g(y_0)}{g(y_0)}}{\frac{y - y_0}{y_0}} \right| = \left| \frac{g(y) - g(y_0)}{y - y_0} \right| \cdot \left| \frac{y_0}{g(y_0)} \right| \qquad \left| \frac{g(y) - g(y_0)}{y - y_0} \right|$$

si $y_0 g(y_0) \neq 0$ $\Downarrow$ $\qquad$ $\Downarrow$ si $y_0 g(y_0) = 0$

- Condicionamiento relativo $\quad c(g, y_0) := \left| \frac{g'(y_0)\, y_0}{g(y_0)} \right|$
- Condicionam. absoluto $\quad C(g, y_0) = |g'(y_0)|$

• Si en el problema (P) la resolvente es de clase C^1 e $y_0 \in Y$, el condicionamiento relativo o absoluto de (P) son los de su resolvente en dicho punto.

• Si el condicionamiento es pequeño se dice que el problema está bien condicionado.

• Si $\|\cdot\|$ denota la norma matricial en $\mathbb{R}^{N\times N}$ inducida por una norma en $\mathbb{R}^N$ que notaremos igual, definimos el condicionamiento $c(A)$ de una matriz regular $A \in \mathbb{R}^{N\times N}$ como:

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

• Algoritmo procedimiento que describe de manera precisa y mediante un nº finito de op. aritméticas y lógicas, la resolución de un problema.

• Complejidad de un algoritmo: medida del tiempo de ejecución que suele expresarse en términos de un parámetro asociado al problema

• Algoritmo PageRank. de Google.
i) contemplamos el nº de enlaces que recibe la pg. i.
Desv: no se tiene en cuenta que la pg sea más relevante si ofrece enlaces con otras pg
ii) Contemplamos las pgs con enlaces entrantes. (que recibe)
Desv: no influye el nº de enlaces de cada pg (que presenta)
Sol: dividir cada valor de relevancia x_i por el nº de enlaces que salen de la pag.

• Los ordenadores trabajan con números máquina y pecan de un error de redondeo que provoca una propagación del error.

• Pasar de base 2 a base 10 → trivial.
Pasar de base 10 a base 2:
*) enteros: algoritmo de Euclides
*) decimales: multiplicar por adecuadas potencias de $\frac{1}{2}$.

- Recordemos:

$$\alpha \neq 1,\ n > k \geq 0 \implies \sum_{j=k}^{n} \alpha^{j} = \frac{\alpha^{k} - \alpha^{n+1}}{1-\alpha}$$

En particular; si $|\alpha|<1$

$$\sum_{j=k}^{\infty} \alpha^{k} = \frac{\alpha^{k}}{1-\alpha}$$ esto es,

la suma de una serie geométrica es igual al primer término menos el último término por la razón entre 1 menos la razón.

- Representación binaria finita $\Rightarrow$ representación decimal finita.

- Representación posicional:

$$x = (-1)^{s} \sum_{n=-M}^{N} x_n b^{n}$$

$$\boxed{C(A) = \|A\| \, \|A^{-1}\|}$$

- Representación con punto fijo:

$$(-1)^{s} b^{-k} \sum_{n=0}^{n-1} a_n b^{n} \Longleftrightarrow (-1)^{s} (a_{n-2} \cdots a_k, a_{k-1} \cdots a_0)$$

- Representación con punto flotante.

$$(-1)^{s} b^{e} \sum_{n=1}^{t} a_n b^{-n} \Longleftrightarrow (-1)^{s} (0, a_1 \dots a_t) \cdot b^{e} = (-1)^{s} m \cdot b^{e-}$$

t = nº cifras significativas
m = mantisa
e = exponente.

- Sean $t \in \mathbb{N}$, $L, U \in \mathbb{Z}$ con $L \leq U$ y $x \in \mathbb{F}(b,t,L,U)$:

i) $-x \in \mathbb{F}(b,t,L,U)$

ii) $b^{L-1} \leq |x| \leq b^{U}(1-b^{-t})$

iii) $\text{card}(\mathbb{F}(b,t,L,U)) = 2(b-1)b^{t-1}(U-L+1)+1$

$\varepsilon_M = b^{1-t}$ (epsilon maquina)

$u = \frac{1}{2}\varepsilon_M$ (precisión maquina)

• Sea $F(b,t,L,U)$, el redondeo del nº real.

$$x=(-1)^s\, b^e \sum_{n=1}^{\infty} a_n b^{-n}$$

es:
$$rd(x) = tr\left(x + (-1)^s \frac{b}{2} \frac{b^e}{b^{t+1}}\right)$$

• $tr(x) = (-1)^s\,(0, a_1 \dots a_t) \cdot b^e$

$$\boxed{\begin{array}{ccccc} f \in C^1 & \Longrightarrow & f \text{ estable} & \Longrightarrow & f \text{ continua} \\ \cancel{\Leftarrow} & & & \cancel{\Leftarrow} & \\ |\cdot| & & & \text{raiz en } 0. & \end{array}}$$

• Sea el sist. punto flotante: $F(b,t,L,U)$ y sea
$x = (-1)^s\, b^e \sum_{i=1}^{t} a_i b^{-i}$

i) $|x - tr(x)| \leq b^{e-t}$

ii) $\dfrac{|x - tr(x)|}{|x|} \leq \varepsilon_M$

iii) $|x - rd(x)| \leq \frac{1}{2} b^{e-t}$

iv) $\dfrac{|x - rd(x)|}{|x|} \leq \dfrac{\varepsilon_M}{2} = u$

• sea $F(b,t,L,U)$, $x \in \mathbb{R}$ con $b^{L-1} \leq |x| \leq b^U(1-b^{-t})$

$rd(x) = (1+\mu)x$ para un $\mu \in \mathbb{R} : |\mu| \leq u$

MÉTODOS NUMÉRICOS TEMA 2

Métodos para resolver Sst. ecuaciones
- Directos
  - Gauss
  - Versiones
    - Pivotaje
    - Total
    - Gauss-Jordan
  - Factorizaciones
    - LU
    - Doolittle
    - Crout
    - Cholesky
- ITERATIVOS
  - Jacobi
  - Gauss-Seidel.

ALGORITMOS QUE APRENDER

- Resolución por sustitución hacia atrás:

$$x_N = \frac{b_N}{u_{NN}}$$

$$i = N-1, \dots 1 \Rightarrow \quad x_i = \frac{1}{u_{ii}} \left(b_i - \sum_{j=i+1}^{N} u_{ij} x_j \right)$$

- Resolución por sustitución hacia adelante.

$$x_1 = \frac{b_1}{l_{11}}$$

$$i = 2, \dots, N \Rightarrow \quad x_i = \frac{1}{l_{ii}} \left(b_i - \sum_{j=1}^{i-1} l_{ij} x_j \right)$$

Un multiplicador es un término por el que, en el método de Gauss, al multiplicarlo por una n-upla (fila) obtenemos otra que nos sirve para simplificar.
Gauss se resuelve por sustitución hacia atrás.

- Gauss:
  1º) Debemos calcular el multiplicador; para ello buscamos la segunda fila que no tenga ceros, y dividimos ese término entre el que está encima, ese es el multiplicador.
  2) Para calcular los nuevos valores de la segunda fila, debemos restarla con el término que tienen justamente encima multiplicador por el multiplicador.
  3) Así obtenemos el valor de ese término. Repetimos el proceso en todas las filas que no tengan ceros.
  (Ejemplo diap 20)

→ Si $A \in \mathbb{R}^{N\times N}$, si el método de Gauss llega hasta el paso N, entonces A es regular.

→ Sea $A \in \mathbb{R}^{N\times N}$ matriz cuadrada, es equivalente:

i) El mét. Gauss puede completarse hasta el paso n-ésimo.

ii) $\forall K=1 \dots N$, la K-ésima submatriz principal de A es regular

→ Nº operaciones aritméticas para resolver un sistema mediante Gauss

$$\sum_{j=1}^{N} j^2 = \frac{N(N+1)(2N+1)}{6}$$

Con Gauss debemos evitar dividir por coeficientes pequeños o que algún $a_{kk}^{(k)}$ sea nulo.

• Método de Gauss con pivotaje.

Este método tiene sentido hasta el paso n-ésimo <=> A es regular

La diferencia es que en las filas que no tienen ceros, debemos situar en primer lugar la que tenga como primer coeficiente el mayor valor absoluto.

Existe además el método de Gauss con pivotaje total, el cual no solo reordena las filas $k, \dots, N$ sino que también reordena las columnas $k \dots N$ de manera que el primer elemento tiene como valor el mayor valor abs de todas las filas y columnas.

Existe también el método de Gauss-Jordan, que consiste en hacer ceros no solo debajo sino también por encima.

• Factorización LU

sea $Ax = b$, SCD $\Rightarrow A = LU$; $LUx = b$

i) $y := Ux$; $\boxed{Ly = b}$ sustitución hacia delante

ii) $\boxed{Ux = y}$ sustitución hacia atrás

Debe verificarse que L, U regulares y $A = LU$.

→ Sea $A \in \mathbb{R}^{N\times N}$ matriz regular, equivale

i) A admite factorización LU

ii) Las N submatrices principales de A son regulares.

iii) $\forall b \in \mathbb{R}^{N\times N}$ el método de Gauss para el correspondiente SEL puede completarse hasta el paso N-ésimo

→ Si $A \in \mathbb{R}^{N\times N}$ una matriz regular $\Rightarrow \exists$ una matriz $\tilde{A}$ que se obtiene una matriz permutando filas o columnas y admite factorización LU.

→ Sea A una matriz regular; equivale:

i) Método de Gauss con pivotaje es factible, ∀ SEL que tenga A como matriz de coeficientes.

ii) salvo alguna permutación de algunas de sus filas, A admite factorización LU.

Factorizaciones

Doolittle $l_{11} = \dots = l_{nn} = 1$ (1^a matriz) [avanzamos verticalmente

Crout $u_{11} = \dots = u_{nn} = 1$ (2^a matriz) [avanzamos horizontalmente

• Doolittle

$j = i, \dots, N \Rightarrow u_{ij} = a_{ij} - \sum_{k=1}^{i-1} l_{ik} u_{kj}$; $j = i+1, \dots N . \Rightarrow l_{ji} = \frac{1}{u_{ii}} \left(a_{ji} - \sum_{k=1}^{i-1} l_{jk} u \dots \right)$

→ Si $A^T = LU \Rightarrow A = U^T L^T$

→ La factorización LU es el método más eficiente

• Cholesky

$A \in \mathbb{R}^{N \times N}$ def - positiva ⇒ $x \in \mathbb{R}^N \setminus \{0\} \Rightarrow x^T A x > 0$

$L = U^T$

Sea $A \in \mathbb{R}^{N \times N}$ matriz simétrica y def ⊕ ⇒ $\exists ! U \in \mathbb{R}^{N \times N}$ con coef ⊕ diagonal:

$$A = U^T U$$

$$i = 1, \dots, j-1 \Rightarrow u_{ij} = \frac{1}{u_{ii}} \left(a_{ij} - \sum_{k=1}^{i-1} u_{ki} u_{kj} \right) \quad ; \quad u_{jj} = \sqrt{a_{jj} - \sum_{k=1}^{j-1} u_{kj}^2}$$

Métodos iterativos solución de SEL y SCD como límite de una sucesión. Cada término se obtiene de manera recursiva del anterior.

$$\left.\begin{array}{l} x_0 \text{ dado} \\ x_n = B x_{n-1} + c \end{array}\right\} n \geq 1 \qquad x_0, c \in \mathbb{R}^N, \; B \in \mathbb{R}^{N \times N}$$

$$\boxed{c = (I - B) A^{-1} b}$$

Consistencia de un sistema

$\lim_{x \to \infty} x_n = x \Rightarrow x = Bx + c$

→ convergencia a la sol. del sistema $\overset{*}{\Rightarrow}$ consistencia del método con el sistema.

→ Sea $A, B \in \mathbb{R}^{N \times N}$, A regular, $\left.\begin{array}{l} x_0 \text{ dado} \\ x_n = B x_{n-1} + c \end{array}\right\}$ $x_0, b, c \in \mathbb{R}^N$ $N \geq 1$ $n \geq 1$

es consistente con el sistema unisolvente $Ax = b$.

$\forall x_0 \in \mathbb{R}^N$ el método iterativo converge a la solución ⇔ $\rho(B) < 1$

→ convergencia solo cuando $c = 2x_0 \iff x_0 = 0$

y $\rho(B) = \rho(-I) = 1$

- Procedimiento de diseño de métodos iterativos

$Ax = b$, $A \in \mathbb{R}^{N\times N}$ regular, $b \in \mathbb{R}^N$

$A = M - N$, M regular y N no nula

$Ax = b \iff (M-N)x = b \iff x = M^{-1}Nx + M^{-1}b$

Entonces: $B = M^{-1}N$ $\quad c = M^{-1}b$

Sugiere:
$$\left.\begin{array}{l} x_0 \text{ dado} \\ x_n = M^{-1}Nx_{n-1} + M^{-1}b \end{array}\right\} n \geq 1$$

Para entenderlo:

$$D = \begin{pmatrix} a_{11} & & 0 \\ & \ddots & \\ 0 & & a_{nn} \end{pmatrix} \quad E = \begin{pmatrix} 0 & & 0 \\ & \ddots & \\ -\text{Inf} & & 0 \end{pmatrix} \quad F = \begin{pmatrix} 0 & & -\text{Sup} \\ & \ddots & \\ 0 & & 0 \end{pmatrix}$$

$A = D - E - F$; $Ax = b$.

① $A = M - N$

② Decides cual es M y N

Método	$A = D - E - F$
Jacobi	$A = D - (E+F)$
Gauss-Seidel	$A = \underbrace{D - E}_{M} - \underbrace{F}_{N}$

③ Método converge $\iff \rho(M^{-1}N) < 1$

- Jacobi: $$x_{ni} = \frac{1}{a_{ii}}\left(b_i - \sum_{\substack{j=1 \\ j\neq i}}^{N} a_{ij}\, x_{(n-1)j}\right)$$

- Gauss-Seidel $$x_{ni} = \frac{1}{a_{ii}}\left(b_i - \sum_{j=1}^{i-1} a_{ij}x_{nj} - \sum_{j=i+1}^{N} a_{ij}\, x_{(n-1)j}\right)$$

→ sean $A, M, N \in \mathbb{R}^{N\times N}$, A y M regulares; $A = M - N$, $b, x_0 \in \mathbb{R}^N$

consideremos $Ax = b$. y el método $\left.\begin{array}{l} x_0 \text{ dado} \\ x_n = M^{-1}Nx_{n-1} + M^{-1}b \end{array}\right\} n \geq 1$

El método converge a la solución $\forall x_0 \in \mathbb{R}^N \iff \rho(M^{-1}N) < 1$

→ La velocidad de convergencia de Jacobi es independiente a la de Gauss-Seidel.

→ El método iterativo converge a la solución para cualquier estimación inicial.

→ $A \in \mathbb{R}^{N \times N}$ es simétrica y def ⊕ $\Longleftrightarrow$ admite fact. Cholesky.

→ Cholesky nos dice que su diagonal es > 0.

→ $A \in \mathbb{R}^{N \times N}$ es diag. estrictamente dominante si

$$i = 1 \ldots N \Rightarrow \sum_{\substack{j=1 \\ j \neq i}}^{N} |a_{ij}| < |a_{ii}|$$

Esto es, el término situado en la diagonal en valor absoluto es mayor que la suma de los valores absolutos de los términos de esa misma fila.

→ Sea $A \in \mathbb{R}^{N \times N}$ una matriz diag. est. dominante $\Longrightarrow$ Jacobi y Gauss-Seidel convergen hacia su solución $\forall$ SEL que tenga A como matriz de coef. independientemente de la estimación inicial que se fije.

→ Sea la transformación elemental "intercambiar dos ecuaciones de lugar". Esta transf. elemental no solo modifica que la matriz sea DED, sino que puede cambiar el valor del radio espectral

→ Sea $A \in \mathbb{R}^{N \times N}$ regular, $x, u, b \in \mathbb{R}^N$ con $\|x\| \|b\| \neq 0$ de la forma $Ax = b \Longrightarrow \forall$ norma en $\mathbb{R}^N$ se verifica

$$\frac{1}{c(A)} \frac{\|Au - b\|}{\|b\|} \leq \frac{\|x - u\|}{\|x\|} \leq c(A) \frac{\|Au - b\|}{\|b\|}$$

siendo $c(A)$ el condicionamiento de la matriz A
siendo x la solución del sistema $Ax = b$; u solución aprox.

Interpretación: estimación del error relativo en función de $c(A)$ y error relativo generado al tomar Au por b

$Au - b$ es el denominado <u>residuo</u>.

→ Sean $N \geq 1$, $A, B \in \mathbb{R}^{N\times N}$, A regular, $b, c \in \mathbb{R}^N$ y supongamos que el método

$$\left.\begin{array}{l} x_0 \text{ dado} \\ x_n = B x_{n-1} + c \end{array}\right\} n \geq 1$$

converge a la solución x del sistema $Ax = b$ $\forall x_0 \in \mathbb{R}^N$.

Si además $\|\cdot\|$ es una norma en $\mathbb{R}^N$ con norma matricial inducida en $\mathbb{R}^{N\times N}$: $\|B\| < 1$, entonces se verifica:

i) $\|x_n - x\| \leq \dfrac{\|B\|^n}{1 - \|B\|} \|x_1 - x_0\|$

ii) $\|x_n - x\| \leq \|B\| \, \|x_{n-1} - x\|$

iii) $\|x_n - x\| \leq \dfrac{\|B\|}{1 - \|B\|} \|x_n - x_{n-1}\|$

$\rho(A) = \inf\{\|A\| : \|\cdot\|$ es una norma matricial inducida en $\mathbb{R}^{N\times N}\}$

- (rojo) Se comprueba simplemente viendo si las submatrices son o no regulares
- (negro) Submatrices principales son regulares (se pueden permutar filas)
- (verde) El método iterativo converge a la sol $\iff \rho(B) < 1$ o directamente si es DED.

Métodos para resolver sistemas

- Directos
  - Gauss
  - Versiones
    - Pivotaje
    - Total
    - Gauss Jordan
  - Factorizaciones (↗ método más eficiente.)
    - LU
    - Doolittle
    - Crout
    - Cholesky

    $LUx = b$; $Ly = b$; $Ux = y$
- Iterativos
  - Jacobi
  - Gauss-Seidel

TEMA 5 MÉTODOS NUMÉRICOS

- **PRINCIPIO DEL MÍNIMO** $A \in \mathbb{R}^{N\times N}$ simétrica y def$\oplus$, $b \in \mathbb{R}^N$ y $f: \mathbb{R}^N \to \mathbb{R}$ es la función cuadrática definida $\forall x \in \mathbb{R}^N$

$$f(x) = \frac{1}{2} x^T A x - b^T x$$

La solución del sistema $Ax=b$ y el mínimo de la función es único y coincide. $-\frac{1}{2} b^T A^{-1} b$

- A def$\oplus$ $\Rightarrow$ A regular.

$Ax = 0$; $x^T A x = 0 \Leftrightarrow x = 0$; $A^{-1}b$ solución de $Ax=b$.

~Dem~

$x := A^{-1}b$, $y \in \mathbb{R}^N$

$$\begin{aligned} f(y) - f(x) &= \frac{1}{2} y^T A y - b^T y - \left[\frac{1}{2} x^T A x - b^T x\right] \\ &= \frac{1}{2} y^T A y - b^T y - \frac{1}{2} x^T A x + b^T x \quad \downarrow b = Ax \\ &= \frac{1}{2} y^T A y - (Ax)^T y - \frac{1}{2} x^T A x + (Ax)^T x \quad \downarrow A \text{ simétrica.} \\ &= \frac{1}{2} y^T A y - x^T A y - \frac{1}{2} x^T A x + x^T A x \\ &= \frac{1}{2} y^T A y - x^T A y + \frac{1}{2} x^T A x \\ &= \frac{1}{2} (y-x)^T A (y-x) \end{aligned}$$

A def$\oplus$ alcanza mínimo en $y = x = A^{-1}b$; $f(A^{-1}b) = -\frac{1}{2} b^T A^{-1} b$

- $A \in \mathbb{R}^{N\times N}$ simétrica y def$\oplus$ $\Leftrightarrow$ A admite factorización LU Cholesky
- $A \in \mathbb{R}^{N\times N}$ regular $\Rightarrow$ $A^T A$ simétrica y def$\oplus$
- $A \in \mathbb{R}^{M\times N}$ rang$(A) = N$ $\Leftrightarrow$ columnas de A l.i. $\Rightarrow$ $A^T A$ simét. def$\oplus$
- $C \in \mathbb{R}^{M\times N}$ simétrica y def$\oplus$, $A \in \mathbb{R}^{M\times N}$ rang$(A) = N$ $\Leftrightarrow$ columnas A l.i. $\Rightarrow$ $A^T C A$ simétrica y def$\oplus$

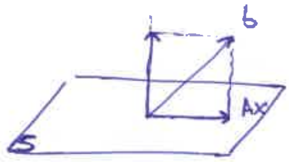
• Aplicación II del Principio del mínimo.

S s.v. de $\mathbb{R}^M$ $\qquad \left\{ \begin{bmatrix} a_{11}\\ \vdots \\ a_{M1}\end{bmatrix}, \dots, \begin{bmatrix} a_{1N}\\ \vdots\\ a_{MN}\end{bmatrix}\right\}$ base de S.

rang(A) = N. $\qquad S = \{ Ax : x\in\mathbb{R}^N \}$.

$b\in\mathbb{R}^M$; $\quad b\in S \iff Ax=b$ compatible.

$\qquad b\notin S \iff Ax=b$ incompatible



• ¿$\exists\, Ax\in S$: Ax más próximo a b?

¿$\exists\, x\in\mathbb{R}^N : y\in\mathbb{R}^N \Rightarrow \|Ax-b\|_2 \le \|Ay-b\|_2$?

¿$\exists\, x\in\mathbb{R}^N : y\in\mathbb{R}^N \Rightarrow \|Ax-b\|_2^2 \le \|Ay-b\|_2^2$?

$Ax=b$

¿$\exists\, x\in\mathbb{R}^N : y\in\mathbb{R}^N \Rightarrow (Ax-b)^T(Ax-b) \le (Ay-b)^T(Ay-b)$? $\qquad x^TA^T=b^T$

¿$\exists\, x\in\mathbb{R}^N : y\in\mathbb{R}^N \Rightarrow x^TA^TAx + b^Tb - b^TAx \le y^TA^TAy + b^Tb - 2b^TAy$?

¿$\exists\, x\in\mathbb{R}^N : y\in\mathbb{R}^N \Rightarrow \frac{1}{2}x^TA^TAx - b^TAx \le \frac{1}{2}y^TA^TAy - b^TAy$? P. Mínim

Por el p. mínimo $\Rightarrow$ $\quad f(x) = \frac{1}{2}x^TA^TAx - b^TAx \quad$ Respuesta afirmativa

Ax es la mejor aproximación de b en S.

Demostración NORMA P

$x \in \mathbb{R}^N$; $x = (x_1, \dots, x_N)$

se define

$$\|x\|_p = \left(\sum_{j=1}^{N} |x_j|^p\right)^{\frac{1}{p}} \quad ; \quad p > 1$$

Demostremos que es una norma.

$\boxed{p=1}$

$$\|x\|_1 = \sum_{j=1}^{N} |x_j| \quad \text{es obvio}$$

- $\|x\|_1 = 0$; $x_1 = x_2 = \dots = x_N = 0$; $x = 0$.

  $\|x\|_1 > 0$ por def. , esto es, ser suma de valores abs.

- $\|\alpha x\|_1 \overset{(?)}{=} |\alpha| \, \|x\|_1$; $\alpha \in \mathbb{R}$

$$\|\alpha x\|_1 = \sum_{j=1}^{N} |\alpha x_j| = \sum_{j=1}^{N} |\alpha| |x_j|$$

$$= |\alpha| \sum_{j=1}^{N} |x_j| = |\alpha| \, \|x\|_1$$

- $\left.\begin{array}{l} x = (x_1, \dots, x_N) \in \mathbb{R}^N \\ y = (y_1, \dots, y_N) \in \mathbb{R}^N \end{array}\right\} \quad x + y = (x_1 + y_1, \dots, x_N + y_N)$

¿ $\|x+y\|_1 \leq \|x\|_1 + \|y\|_1$?

$$\|x+y\|_1 = \sum_{j=1}^{N} |x_j + y_j| \leq \sum_{j=1}^{N} \left(|x_j| + |y_j|\right)$$

$$= \sum_{j=1}^{N} |x_j| + \sum_{j=1}^{N} |y_j| =$$

$$= \|x\|_1 + \|y\|_1$$

$\boxed{p > 1}$

$p \in \mathbb{R}$

Se dice que p' es un exponente conjugado de p si

$$\frac{1}{p} + \frac{1}{p'} = 1$$

siendo $p, p' > 1$

(i) $¿\, x y \leq \frac{x^p}{p} + \frac{y^{p'}}{p'}$? ; $x, y \in \mathbb{R}$, $x, y \geq 0$

Definimos la función:

$$f(x) = \frac{x^p}{p} + \frac{y^{p'}}{p'} - xy$$

$\forall x \in \mathbb{R}$, $x > 0$ || siendo y fijo.

Queremos demostrar $f(x) \geq 0 \quad \forall x \geq 0$

Queremos calcular los ptos críticos. y fijo ; $y \geq 0$

$$f'(x) = x^{p-1} - y = 0 \; ; \quad x = y^{\frac{1}{p-1}}$$

$$f''(x) = (p-1)\, x^{p-2} \; ; \quad f''\left(y^{\frac{1}{p-1}}\right) = \underset{\overset{\vee}{0}}{(p-1)}\; \underset{\overset{\vee}{0}}{y^{\frac{p-2}{p-1}}} > 0$$

entonces $y^{\frac{1}{p-1}}$ es un mínimo relativo.

Veamos $f\left(y^{\frac{1}{p-1}}\right) = \frac{y^{\frac{p}{p-1}}}{p} + \frac{y^{p'}}{p'} - y^{\frac{1}{p-1}}\, y$

$$\frac{1}{p-1} + 1 = \frac{1 + p - 1}{p-1} = \frac{p}{p-1}$$

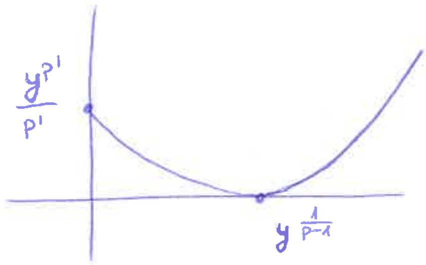
$$\frac{1}{p} + \frac{1}{p'} = 1 \; : \quad p' + p = pp' \; ; \quad p = pp' - p'$$

$$p = p'(p-1) \; ; \quad \frac{p}{p-1} = p'$$

$$f\left(y^{\frac{1}{p-1}}\right) = \frac{y^{p'}}{p} + \frac{y^{p'}}{p'} - y^{p'} =$$

$$= y^{p'}\left(\frac{1}{p} + \frac{1}{p'} - 1\right) = 0$$

$$f(0) = \frac{y^{p'}}{p'}$$



Hemos demostrado que $f(x) \geq 0 \quad \forall\, x \geqslant 0$

es decir,

$$x\,y \leq \frac{x^p}{p} + \frac{y^{p'}}{p'}$$

[es obvio si $x=0$, o $y=0$ se verifica la desigualdad]

ii) $x_1, \dots, x_N \in \mathbb{R}$

$y_1, \dots, y_N \in \mathbb{R}$

$$¿\sum_{j=1}^{N} |x_j y_j| \leq \|x\|_p \|y\|_{p'}?$$

$$\|x\|_p = \left(\sum_{j=1}^{N} |x_j|^p\right)^{\frac{1}{p}}$$

$$\|y\|_{p'} = \left(\sum_{j=1}^{N} |y_j|^{p'}\right)^{\frac{1}{p'}}$$

ii.1 Si $x=0$ o $y=0$ entonces se verifica la desigualdad.

ii.2 Suponemos que $x \neq 0$, $y \neq 0$

$$\|x\|_p = 1 = \|y\|_{p'}$$

$$\sum_{j=1}^{N} |x_j y_j| = \sum_{j=1}^{N} |x_j| |y_j| \leq \sum_{j=1}^{N} \left[\frac{|x_j|^p}{p} + \frac{|y_j|^{p'}}{p'} \right] =$$

$$= \sum_{j=1}^{N} \frac{|x_j|^p}{p} + \sum_{j=1}^{N} \frac{|y_j|^{p'}}{p'}$$

MAGIA $\rightarrow$

$$= \frac{1}{p} + \frac{1}{p'} = 1 \; ; \quad \sum_{j=1}^{N} |x_j y_j| \leq \underset{\substack{\| \\ \|x\|_p \; \|y\|_{p'} \\ \| \quad\;\; \| \\ 1 \quad\;\; 1}}{1}$$

ii. 3

Supongamos que $x \neq 0$; $y \neq 0$, $\|x\|_p \neq 1$

$\|y\|_{p'} \neq 1$

suponemos $\frac{x}{\|x\|_p}$; $\frac{y}{\|y\|_{p'}}$

$$\left\| \frac{x}{\|x\|_p} \right\|_p \underset{\text{Def}}{=} \left[\sum_{j=1}^{N} \left| \frac{x_j}{\|x\|_p} \right|^p \right]^{\frac{1}{p}} = \left[\sum_{j=1}^{N} \frac{1}{\|x\|_p^p} |x_j|^p \right]^{\frac{1}{p}} =$$

*Aclaración

$$\frac{x}{\|x\|_p} = \left(\frac{x_1}{\|x\|_p}, \frac{x_2}{\|x\|_p}, \dots, \frac{x_N}{\|x\|_p} \right)$$

* Fin aclaración.

$$= \left[\frac{1}{\|x\|_p^p} \sum |x_j|^p \right]^{\frac{1}{p}} = \left(\frac{1}{\|x\|_p^p} \right)^{\frac{1}{p}} \left(\sum_{j=1}^{N} |x_j|^p \right)^{\frac{1}{p}} =$$

$$= \frac{1}{\|x\|_p} \cdot \|x\|_p = 1$$

Analogamente para y. suponemos $\frac{y}{\|y\|_{p'}} \in \mathbb{R}^N$

$$\left\| \frac{y}{\|y\|_p} \right\| = 1$$

$$\sum_{j=1}^{N} \frac{|x_j y_j|}{\|x\|_p \|y\|_{p'}} \underset{\substack{\downarrow \\ ii.2}}{\leq} 1 ;$$

$$\sum_{j=1}^{N} |x_j y_j| \leq \|x\|_p \|y\|_{p'}$$

(iii) $\|x+y\|_p \leq \|x\|_p + \|y\|_p$

$$\|x+y\|_p := \left[\sum_{j=1}^{N} |x_j + y_j|^p \right]^{\frac{1}{p}}$$

Partimos de:

$$\|x+y\|_p^p = \sum_{j=1}^{N} |x_j + y_j|^p = \sum_{j=1}^{N} |x_j + y_j|^{p-1} |x_j + y_j|$$

$$\leq \sum_{j=1}^{N} |x_j + y_j|^{p-1} \cdot (|x_j| + |y_j|)$$

$$= \sum_{j=1}^{N} |x_j + y_j|^{p-1} |x_j| + \sum_{j=1}^{N} |x_j + y_j|^{p-1} |y_j|$$

$$= \sum_{j=1}^{N} |x_j| |x_j + y_j|^{p-1} + \sum_{j=1}^{N} |y_j| |x_j + y_j|^{p-1}$$

(ii) $$\leq \|x\|_p \left[\sum_{j=1}^{N} (|x_j + y_j|^{p-1})^{p'} \right]^{\frac{1}{p'}} + \|y\|_p \left[\sum_{j=1}^{N} ((x_j + y_j)^{p-1})^{p'} \right]$$

——— ACLARACIONES ———

$$\frac{1}{p} + \frac{1}{p'} = 1 ; \quad p' + p = pp' ; \quad p = pp' - p' ; p = p'(p-1)$$

——— FIN ACLARACIONES ———

$$= \|x\|_p \left[\sum_{j=1}^{N} |x_j + y_j|^p \right]^{\frac{1}{p'}} + \|y\|_p \left[\sum |x_j + y_j|^p \right]^{\frac{1}{p'}}$$

$$= \left[\sum_{j=1}^{N} |x_j + y_j|^p \right]^{\frac{1}{p'}} (\|x\|_p + \|y\|_p)$$

Hemos llegado a

$$\sum_{j=1}^{N} |x_j + y_j|^p \leq \left[\sum_{j=1}^{N} |x_j + y_j|^p \right]^{\frac{1}{p'}} \left(\|x\|_p + \|y\|_p \right)$$

$$\left[\sum_{j=1}^{N} |x_j + y_j|^p \right]^{\left(1 - \frac{1}{p'}\right)} \leq \|x\|_p + \|y\|_p$$

$$\left[\sum_{j=1}^{N} |x_j + y_j|^p \right]^{\frac{1}{p}} \leq \|x\|_p + \|y\|_p$$

$$\|x + y\|_p \leq \|x\|_p + \|y\|_p$$