Análisis de Errores

Métodos Numéricos

Curso 2021-22

Sección 1

Introducción

¿Para qué este tema?

En la asignatura vamos a analizar cómo aplicar métodos **automatizados** (algoritmos) de forma práctica, para resolver problemas, principalmente AL y AVR.

DATOS   ALGORITMO   RESULTADO

Ejemplo: $\sqrt{17}$ con algoritmos definidos en EMA y AVR

Tipos de Errores

- Errores de entrada
- Errores de almacenamiento
- Errores de algoritmo

- Error absoluto:

$$\|\hat{z} - z\|$$

- Error relativo

$$\frac{\|\hat{z} - z\|}{\|z\|}$$

Sección 2

Error de almacenamiento: números máquina

Error de almacenamiento: números máquina

Cada número se almacena con una cantidad **finita** de decimales.
Los **números máquina** son los que se almacenan de forma **exacta**

NOTACIÓN DECIMAL EN COMA FLOTANTE

Contiene:

- signo
- fracción
- signo exponente
- exponente

Ejemplo: -532’45

$-532'45 + 0;$	$-53'245 + 1;$	$-5324'5 - 1$
$-532'45x10^0;$	$-53'245x10^1;$	$-5324'5x10^{-1}$

Error de almacenamiento: números máquina

De las infinitas que puede haber, nos quedamos con **coma flotante normalizada**: donde la fracción está entre 1 y 10 .

$$\pm m \pm E, 1 \leq m < 10 \Rightarrow \underline{\pm m x 10^{\pm E}}$$

$$-5'3245x10^2; -5'3245 + 2$$

En ordenador se suele usar el **sistema binario**, que solo usa los dígitos 0 y 1.

* Ayuda C. Virtual

Notación coma flotante en sistema binario

(s)	(S)	E	m
signo 0:+ 1:-	signo exponente 0:+ 1:-	número entero base 2	**mantisa** base 2 $1 \leq m < 2$

$$\Rightarrow (-1)^s.m,2^{(-1)^S x E_d}$$

E_d: conversión a decimal de E

Ejemplos:

(1) (0) 10 1′01	$-1'01x2^{2(10 \textit{ en binario})}$
(0) (1) 1 1′1	$1'1x2^{-1} = 0'11$

En decimal:

$$-1'01x2^2 = -101 = -(2^2 + 2^0) = -5$$

$$1'1x2^{-1} = 0,11 = 2^{-1} + 2^{-2} = 0'75$$

Cada número se almacena en una **palabra** de un número de bits. **32 bits precisión simple**

Notación coma flotante en sistema binario

32 bits precisión simple

- **1**: para signo del número
- **8**: para exponente y su signo
- **23**: para mantisa*

* como la primera cifra siempre es 1, se sobreentiende y en realidad se guardan **24** cifras.

Simplificación: no guardar el bit del signo del exponente.
sumar **127** al exponente (el menor exponente negativo que puede tener un número máquina) para que quede seguro positivo.

$$1111111 = -(2^6 + 2^5 + 2^4 + 2^3 + 2^2 + 2^1 + 2^0) = -127$$

Notación coma flotante en sistema binario

Ejemplo

1	10010001	100011100010000....
(1)	(3)	(2)

- **(1)**: bit del signo $\Rightarrow$ negativo
- **(2)**: mantisa, $\mathbf{1}'100011100010.....$*
- **(3)**: exponente 10010001 $\rightarrow$ 145
  Exponente real: 145 -127 $=$ **18**

* *el 1' se sobreetiende y no se ha almacenado*

$$\boxed{-1'10001110001 \text{x} 2^{18}}$$

Notación coma flotante en sistema binario

Ejemplo inverso

$$\boxed{+1011010010001}$$

- nº positivo $\rightarrow$ bit del signo $\underline{\mathbf{0}}$
- normalización del número $\Rightarrow 1'011010010001x2^{12*}$
  **n*º de posiciones que hemos corrido la coma para ser 1'....*
- mantisa $\Rightarrow$ 011010010001... se completa con 0 hasta 23 cifras
- Valor del exponente $\rightarrow \mathbf{12} \Rightarrow 12 + 127 = 139$

$$139 = 2^7 + 2^3 + 2 + 2^0 \rightarrow 10001011$$

Representación del número

0 |	10001011|	01101001000100000000000
1B	8B	23B

- Hoja de ejercicios: 1-5
- Ejemplos de repaso en C Virtual (6 ejercicios usando otras precisiones)

Números no máquina

Los números que se pueden representar en coma flotante de forma exacta son los números máquina.
Cuando no ocurre así, han de aproximarse por **redondeo**
Tipos:

- A la derecha o por exceso: $r(x) = x_d$
- A la izquierda o por defecto: $r(x) = x_i$
- Al más próximo entre x_i o x_d.
  Más usado. Redondeo.

Cota de redondeo absoluto:

$$|r(x) - x| \leq \frac{x_d - x_i}{2} = \frac{2^{\textbf{-23}} x 2^E}{2} = 2^{-24} x 2^E$$

Cota de redondeo relativo:

$$|\frac{r(x) - x}{x}| \leq \frac{x_d - x_i}{2^E} = 2^{\textbf{-24}}$$

Números no máquina

- El número cota de redondeo relativo, 2^{-24}, es la mitad de la distancia entre 1 (m=0) y el siguiente número máquina (m= 00....01)*
Se le llama **precisión o épsilon de la máquina, eps**.
- Los números máquina se concentran de forman irregular: entre dos potencias de 2 consecutivas hay el mismo número de nº máquina, pero las potencias están cada vez más separadas.

** 22 ceros*

Tipos de Números no máquina

a) Exponente fuera de rango $\rightarrow$ **desbordamiento** $\rightarrow$ se toma $+\infty$ ó $-\infty$
b) Mantisa de más de 23 cifras $\rightarrow$ en este caso se hace redondeo
Para **a)**:

$$|E| \leq (11111111)_2 = (255)_{10}$$

Los valores extremos se usan para casos especiales:

- $E_d \in [1, 254]$: **Números normales** $\rightarrow$ $\boxed{x = (-1)^s 1'm \, x \, 2^{E_d - 127}}$
- $E_d = 0$: **Números subnormales** $\rightarrow$ $\boxed{x = (-1)^s 0'm \, x \, 2^{-126}}$ sobreentiende el 1 como parte entera de mantisa y se desplaza la coma para poder usar números más pequeños

Tipos de Números no máquina

- $E_d = 255$: son casos especiales:
  - $\pm\infty$: **desbordamiento**
  - **NaN** *not a number*

Ejercicio 4.b

Sección 3

Propagación de errores.
Errores debidos al algoritmo

Se trabaja con redondeos de números y el resultado de operar a su vez puede ser redondeado.
Denotamos $\oplus$, $\ominus$, $\otimes$, $\oslash$

$$x \oplus y = r(x + y) \qquad x \otimes y = r(xy)$$
$$x \ominus y = r(x - y) \qquad x \oslash y = r(x/y)$$

Hay situaciones anómalas al trabajar con una precisión determinada:

En general no se verifica la propiedad asociativa

$(1 \oplus 2^{-24}) \oplus 2^{-24} = r(r(1 \oplus 2^{-24}) + 2^{-24}) = r(1 \oplus 2^{-24}) = 1$

$1 \oplus (2^{-24} \oplus 2^{-24}) = r(1 \oplus r(2^{-24} + 2^{-24})) = r(1 \oplus 2^{-23}) = 1 + 2^{-23} = 1'0\ldots,1 \neq 1$

Aritmética en coma flotante

Si $x \oplus y = x \not\Rightarrow y = 0$

$y = \theta x, 0 < \theta < 2^{-24}$

Cancelación

Restar dos números muy próximos, en general produce grandes errores relativos

Por ello hay que evitar restar números muy próximos, por ejemplo, cambiando el orden de operaciones o modificando el operador.

Ejercicios 6 y 7 de la hoja

Propagación de error

Tipos:

- Condicionamiento: influencia en el resultado de errores en los datos. Está ligado al problema y no depende del algoritmo que se use.
- Estabilidad: Influencia en el resultado de la acumulación de errores en las sucesivas operaciones.
  Depende del algoritmo.

Condicionamiento

Un problema está mal condicionado si pequeños cambios en los datos $\Rightarrow$ grandes cambios en los resultados.
Sea el problema $y = f(x)$, el **Número de Condicionamiento** $\kappa = \kappa(x) \geq 0/$

$$\frac{\|f(\hat{x}) - f(x)\|}{\|f(x)\|} \sim \kappa(x) \frac{\|\hat{x} - x\|}{\|x\|}$$

Relaciona errores relativos de resultados y datos.
Indica si está bien (~ 1) o mal condicionado

Se trata de operar hasta dejar una expresión parecida a la anterior

Ejemplos

Funciones diferenciables $f : \Re \to \Re$

Se aplica el th. valor medio con $|\hat{x} - x| <<$:

$$f(\hat{x}) - f(x) = f'(\xi)(\hat{x} - x) \sim f'(x)(\hat{x} - x)$$

En términos relativos:

$$\left|\frac{f(\hat{x}) - f(x)}{f(x)}\right| \sim \left|\frac{f'(x)}{f(x)}\right| |\hat{x} - x| = \boxed{\left|\frac{f'(x)x}{f(x)}\right|} \left|\frac{\hat{x} - x}{x}\right|$$

Condicionamiento

Operaciones

Producto (ejercicio 6 hoja de problemas):

$$\frac{\hat{x}_1\hat{x}_2 - x_1x_2}{x_1x_2} = \frac{(x_1+\varepsilon_1)(x_2+\varepsilon_2) - x_1x_2}{x_1x_2} = \frac{x_2\varepsilon_1 + x_1\varepsilon_2 + \varepsilon_1\varepsilon_2}{x_1x_2} \sim$$

$$\sim \frac{x_2\varepsilon_1 + x_1\varepsilon_2}{x_1x_2}$$

Es decir, aproximadamente la suma de errores relativos $\Rightarrow$ el producto de dos números es un problema bien condicionado ($\kappa = 1$).

Sin embargo la suma no siempre lo es (comprobar cuando no)

Ejercicio 8

Estabilidad

Un algoritmo es **inestable** cuando se van acumulando errores y afectan al resultado. Depende del algoritmo
Solo esbozamos la idea, pero se trabajará en aula de informática.

Planteamiento ejercicio 8 de prácticas

Cálculo de $\left(\frac{1}{7}\right)^{100}$

Algoritmo 1

$x_o = 1$
$x_n = \lambda x_{n-1} \rightarrow$ término general $x_n = \lambda^n$
luego es válido con $\lambda = 1/7$ y n=100

Algoritmo 2

$x_o = 1;\ x_1 = \lambda$
$x_{n+1} = (3+\lambda)x_n - 3\lambda x_{n-1} \rightarrow$ término general $x_n = \lambda^n$

Aún siendo ambos válidos, no son igual de estables

Complementos de Análisis Matricial

Métodos Numéricos

Curso 2021-22

Sección 1

Introducción

Introducción

¿Para qué este tema?

- En la primera parte de la asignatura se estudian métodos automatizados para resolver sistemas de ecuaciones lineales
- **Matrices**
- Deseable autonomía para revisar lo menos claro: apoyo hoja de problemas C Virtual

(ejercicios de 1 a 5 para recordar notación de subíndices, productos,...)

Matrices

A^* : Matriz adjunta de A, $A* = (\bar{a}_{ji})$
A^T : Matriz traspuesta de A, $A^T = (a_{ji})$

$(A^*)^* = A$; $((AB)^T)^T = A$
si A es real, $A^* = A^T$
$(AB)^* = B^*A^*$; $(AB)^T = B^TA^T$

(ejercicio 6)

Tipos de Matrices

Inversibles (regular, no singular) $A \in M_n / \exists ! B \in M_n / AB = BA = I$
$B = A^{-1}$ y

$$(AB)^{-1} = B^{-1}A^{-1}$$
$$(A^*)^{-1} = (A^{-1})^*$$
$$(\mathrm{A}^T)^{-1} = (A^{-1})^T$$

Hermítica $A = A^*$ es decir, $a_{ij} = \bar{a}_{ji}$

Simétrica $A = A^T$ es decir, $a_{ij} = a_{ji}$

Unitaria $A^{-1} = A^*$ es decir, $AA^* = A^*A = I$

Normal $AA^* = A^*A$

Ortogonal A real y $A^{-1} = A^T$ es decir, $AA^T = A^TA = I$

Triangular superior $a_{ij} = 0 \, si \; i > j$

Matrices

Tipos de Matrices

Diagonal $a_{ij} = 0 \, si \, i \neq 0$

Banda $\exists p \in 1, 2, ...n / \, a_{ij} = 0 \, si \, |i - j| > p$
p semiancho de banda y 2p-1 ancho de banda

Diagonal estrictamente dominante $|a_{ii}| > \sum_{j=1, j\neq i}^{n} |a_{ij}|$

En algunos casos, estas propiedades están relacionadas:

Toda matriz hermítica es normal
Si A es hermítica e inversible, A^{-1} también lo es
Si A es normal e inversible, A^{-1} también lo es...

(ejercicios de 7 a 11)

Matrices

Traza y determinantes

$A = (a_{ij})_{ij=1}^{n} \quad tr(A) = \sum_{i=1}^{n} a_{ii}$

<u>Propiedades</u>

$tr(AB) = tr(BA)$

$tr(A+B) = tr(A) + tr(B)$

$det(I) = 1$

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

$det(\lambda A) = \lambda^{n} det(A), \ \lambda \in \Re$

$det(A^{*}) = \overline{det(A)}$

A triangular, $det(A) = \prod_{i=1}^{n} a_{ij}$

$det(A^{-1}) = \frac{1}{det(A)}$

Autovalores y Autovectores

$A \in M_n$

Autovalores

Polinomio característico: $|A - \lambda I| = P_A(\lambda)$

Espectro de A: $sp(A) = \{\lambda_1, \lambda_2, ..., \lambda_n\}$

Radio espectral: $\rho(A) = \max\limits_{1 \leq i \leq n} |\lambda_i(A)|, \lambda_i(A) \in sp(A)$

Autovector: $v \in V - \overrightarrow{0} / Av = \lambda v$

- Matrices semejantes: $\exists P$ matriz no singular $/ B = P^{-1}AP$
- Dos matrices semejantes tienen el mismo espectro
- La matriz semejante más sencilla posible si se puede, es la diagonal

Matrices diagonalizables

Una matriz es diagonalizable si $\exists P$ inversible $/ P^{-1}AP$ diagonal

Esto ocurre si $\exists$ una base de autovectores de A

Proposición

Si $A \in M_n$ es una matriz hermítica ($A^* = A$):

- A es definida positiva $\Leftrightarrow sp(A) \subset \Re^+$
- A es semidefinida positiva $\Leftrightarrow sp(A) \subset \Re^+ U\, 0$

(ejercicios de 12 a 14)

Sección 2

Normas Matriciales

Normas Matriciales

Herramienta básica para ver **condicionamientos** de problemas que involucren matrices, y convergencias de procesos iterativos.

Definición

$||| \ ||| : M_n \to \Re^+ U \{0\}$ si verifica:

- $|||A||| = 0 \Leftrightarrow A = 0$
- $|||\lambda A||| = |\lambda| |||A|||$
- $|||A + B||| \leq |||A||| + |||B|||$
- $\boxed{|||A.B||| \leq |||A|||.|||B|||}$

Normas Matriciales

$$||| \; ||| : M_n \to \Re^+ U \{0\} \, / |||A||| = \sup_{v \neq 0} \frac{\|Av\|}{\|v\|} = \sup_{\|v\|=1} \|Av\|$$

Norma matricial subordinada a la norma vectorial $\| \; \|$

Algunas normas matriciales

- $|||A|||_1 = \max_{1 \leq j < n} \sum_{i=1}^{n} |a_{ij}|$ mayor de las cantidades al sumar los módulos de los elementos de cada columna
- $|||A|||_2 = +\sqrt{\rho(A^*A)} = +\sqrt{\rho(AA^*)} = |||A^*|||_2$
- $|||A|||_\infty = \max_{1 \leq i < n} \sum_{j=1}^{n} |a_{ij}|$ mayor de las cantidades al sumar los módulos de los elementos de cada fila
- $|||A|||_F = +\sqrt{\sum_{i,j=1}^{n} |a_{ij}|^2} = +\sqrt{tr(A^*A)}$ Norma de Frobenius

Sección 3

Condicionamiento de una matriz

Condicionamiento de una matriz

Tema anterior: en general no es sencillo determinar el condicionamiento de un problema.
En sistemas lineales (u otros problemas representados por una matriz), sí se puede y está ligado a la matriz del sistema:

$$Au = b, A \in M_n, \exists A^{-1}, b \neq 0$$

Definición: ver qué pasa sobre la solución ante pequeños cambios en b:

$$A(u + \underline{\delta u}) = b + \underline{\delta b}$$

Condicionamiento de una matriz

Tema anterior: en general no es sencillo determinar el condicionamiento de un problema.
En sistemas lineales (u otros problemas representados por una matriz), sí se puede y está ligado a la matriz del sistema:

$$Au = b, A \in M_n, \exists A^{-1}, b \neq 0$$

Definición: ver qué pasa sobre la solución ante pequeños cambios en b:

$$A(u + \underline{\delta u}) = b + \underline{\delta b} \Rightarrow Au + A\delta u = b + \delta b$$

Condicionamiento de una matriz

Tema anterior: en general no es sencillo determinar el condicionamiento de un problema.
En sistemas lineales (u otros problemas representados por una matriz), sí se puede y está ligado a la matriz del sistema:

$$Au = b, A \in M_n, \exists A^{-1}, b \neq 0$$

Definición: ver qué pasa sobre la solución ante pequeños cambios en b:

$$A(u + \underline{\delta u}) = b + \underline{\delta b} \Rightarrow Au + A\delta u = b + \delta b \Rightarrow \cancel{Au} + A\delta u = \cancel{b} + \delta b \Rightarrow$$

$$\underline{\delta u = A^{-1}\delta b}$$

Condicionamiento de una matriz

Para ver el tamaño de la perturbación δu de la solución, usamos el concepto de **norma matricial subordinada**

$$\|\delta u\| \leq |||A^{-1}|||.\|\delta b\|(*)$$

Se acota $\|\delta u\|$

Para tener el número de condicionamiento, se necesitan valores de errores relativos:

$Au = b$ sistema original $\Rightarrow$

$\|b\| \leq |||A|||.\|u\| \Rightarrow \frac{1}{\|u\|} \leq \frac{|||A|||}{\|b\|}(**) \Rightarrow$

Condicionamiento de una matriz

$\Rightarrow$

$$\underbrace{\frac{\|\delta u\|}{\|u\|}}_{\textit{error relativo resultados}} \underset{(*)(**)}{\leq} \frac{|||A^{-1}|||.\|\delta b\|.|||A|||}{\|b\|} = \underbrace{\frac{\|\delta b\|}{\|b\|}}_{\textit{error relativo datos}} |||A|||.|||A^{-1}|||$$

$$\boxed{\mathrm{cond}_p(A) = |||A|||_p|||A^{-1}|||_p}$$

Condicionamiento de una matriz

Propiedades:

- $cond(A) \geq 1$
  por ser norma matricial subordinada:
  $1 = |||I||| = |||AA^{-1}||| \leq |||A|||.|||A^{-1}||| = cond(A)$

- $cond(A) = cond(A^{-1})$
  $cond(A) = |||A|||.|||A^{-1}||| = |||A^{-1}|||.|||A||| = cond(A^{-1})$

- $cond(\lambda A) = cond(A)\, \forall \lambda \in \Re - \{0\}$
  $cond(\lambda A) = |||\lambda A|||.|||(\lambda A)^{-1}||| = |\lambda|.|\lambda^{-1}|.|||A|||.|||A^{-1}||| = cond(A)$

Condicionamiento de una matriz

Propiedad en el caso particular de $||| \ |||_2$:

$$cond_2(A) = +\sqrt{\frac{\lambda_n(A^*A)}{\lambda_1(A^*A)}}$$

con λ_1 y λ_n menor y mayor autovalor de A^*A

Sistemas Lineales. Métodos Directos

Métodos Numéricos

Curso 2021-22

Sección 1

Idea general y conceptos

Idea general y conceptos

¿Para qué este tema?

$Au = b$

- Problemas tipo con solución
- Si se ejecuta en ordenador, esperable error numérico: condicionamiento (Tema1) importante
- Métodos çlásicos": no automatizables
- $u = A^{-1}b$ o Regla de Cramer: costoso operacionalmente (los compararemos después)
- **buscamos métodos que se puedan automatizar**

Ejemplo

Sistema triangular inferior

$$\begin{pmatrix} a_{11} & 0 & \dots & 0 \\ a_{21} & a_{22} & & 0 \\ \vdots & & \ddots & \\ a_{n1} & a_{n2} & \dots l & a_{nn} \end{pmatrix} \begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix}$$

$u_1 = \frac{b_1}{a_{11}}$

$u_2 = \frac{b_2 - a_{21} u_1}{a_{22}}$

...

En general $\to u_i = \frac{b_i - \sum_{j=1}^{i-1} a_{ij} u_j}{a_{ii}}$

Ejemplo

Sistema triangular superior

$$\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ 0 & a_{22} & \dots & a_{2n} \\ & & \ddots & \vdots \\ 0 & 0 & \dots & a_{nn} \end{pmatrix} \begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix}$$

$u_n = \frac{b_n}{a_{nn}}$

$u_{n-1} = \frac{b_{n-1} - a_{n-1,n} u_n}{a_{n-1 n-1}}$

...

En general $\rightarrow u_i = \frac{b_i - \sum_{j=i+1}^{n} a_{ij} u_j}{a_{ii}}$

Vemos que si el sistema fuera de matriz triangular es fácil automatizar.

Vemos que si el sistema fuera de matriz triangular es fácil automatizar.

- **Descomponer una matriz A en producto de una triangular inferior y una superior**
- Nos basamos en método de Gauss
  Ejercicio 1 repaso

Tipos de descomposición

$$A = LU$$
$$PA = LU$$

Descomposición de Cholesky

Para cada uno:

- ¿Cómo se hace?
- **¿Puede hacerse?**

Sección 2

Descomposición LU

LU

L: matriz triangular inferior, con 1s en la diagonal
U: matriz triangular superior

Comenzamos con un ejemplo para visualizar el proceso

$$A = \begin{pmatrix} 1 & 2 & 5 \\ -1 & 0 & 1 \\ 1 & -1 & 1 \end{pmatrix}$$

Hay dos "bucles":

1. en cada columna, hacer ceros los elementos por debajo de la diagonal
2. repetir el proceso para cada columna

LU

$$A = \begin{pmatrix} 1 & 2 & 5 \\ -1 & 0 & 1 \\ 1 & -1 & 1 \end{pmatrix}$$

LU

$$A = \begin{pmatrix} 1 & 2 & 5 \\ -1 & 0 & 1 \\ 1 & -1 & 1 \end{pmatrix}$$

Hacemos cero los elementos de la primera columna debajo de 1:

LU

$$A = \begin{pmatrix} 1 & 2 & 5 \\ -1 & 0 & 1 \\ 1 & -1 & 1 \end{pmatrix}$$

Hacemos cero los elementos de la primera columna debajo de 1:

$$\xrightarrow{2^{a} f + 1 1^{a} f} \begin{pmatrix} 1 & 2 & 5 \\ 0 & 2 & 6 \\ 1 & -1 & 1 \end{pmatrix}$$

$$A = \begin{pmatrix} 1 & 2 & 5 \\ -1 & 0 & 1 \\ 1 & -1 & 1 \end{pmatrix}$$

Hacemos cero los elementos de la primera columna debajo de 1:

$$\xrightarrow{2^a f + 1 1^a f} \begin{pmatrix} 1 & 2 & 5 \\ 0 & 2 & 6 \\ 1 & -1 & 1 \end{pmatrix}$$

$$\xrightarrow{3^a f - 1 1^a f} \begin{pmatrix} 1 & 2 & 5 \\ 0 & 2 & 6 \\ 0 & -3 & -4 \end{pmatrix}$$

Fin de Bucle 1 sobre la columna 1

Repetimos el bucle 1, en tantas columnas como haya hasta llegar a la última (en este caso solo una más)

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

LU

Repetimos el bucle 1, en tantas columnas como haya hasta llegar a la última (en este caso solo una más)

$$\begin{pmatrix} 1 & 2 & 5 \\ 0 & 2 & 6 \\ 0 & -3 & -4 \end{pmatrix} \xrightarrow{3^{\text{a}} f + \frac{3}{2} 2^{\text{a}} f} \begin{pmatrix} 1 & 2 & 5 \\ 0 & 2 & 6 \\ 0 & 0 & 5 \end{pmatrix}$$

Matriz triangular superior: **U**

LU

Comenzamos a generalizar:

$$A = \begin{pmatrix} 1 & 2 & 5 \\ -1 & 0 & 1 \\ 1 & -1 & 1 \end{pmatrix}$$

LU

Comenzamos a generalizar:

$$A = \begin{pmatrix} 1 & 2 & 5 \\ -1 & 0 & 1 \\ 1 & -1 & 1 \end{pmatrix}$$

Primer bucle 1:

$$\xrightarrow{2^a f + 1\, 1^a f} \xrightarrow{3^a f - 1\, 1^a f} \begin{pmatrix} 1 & 2 & 5 \\ 0 & 2 & 6 \\ 0 & -3 & -4 \end{pmatrix} = A_1 = E_1 A$$

Comenzamos a generalizar:

$$A = \begin{pmatrix} 1 & 2 & 5 \\ -1 & 0 & 1 \\ 1 & -1 & 1 \end{pmatrix}$$

Primer bucle 1:

$$\xrightarrow{2^a f + 1 1^a f} \xrightarrow{3^a f - 1 1^a f} \begin{pmatrix} 1 & 2 & 5 \\ 0 & 2 & 6 \\ 0 & -3 & -4 \end{pmatrix} = A_1 = E_1 A$$

con $E_1 = I + l_1 e_1^T$; $l_1 = \begin{pmatrix} 0 & 1 & -1 \end{pmatrix}^T$
(recordad E_i entrega del Tema 2)

LU

Segundo (y último) bucle 1:

$$\xrightarrow{3^{\text{a}} f + \frac{3}{2} 2^{\text{a}} f} \begin{pmatrix} 1 & 2 & 5 \\ 0 & 2 & 6 \\ 0 & 0 & 5 \end{pmatrix} = A_2 = E_2 A_1 = U$$

LU

Segundo (y último) bucle 1:

$$\xrightarrow{3^a f + \frac{3}{2} 2^a f} \begin{pmatrix} 1 & 2 & 5 \\ 0 & 2 & 6 \\ 0 & 0 & 5 \end{pmatrix} = A_2 = E_2 A_1 = U$$

con $E_2 = I + l_2 e_2^T;\, l_2 = \begin{pmatrix} 0 & 0 & +\frac{3}{2} \end{pmatrix}^T$

$$U = E_2 A_1 = E_2 E_1 A \Rightarrow A = \underbrace{E_1^{-1} E_2^{-1}}_{L?} U$$

LU

$$E_1 = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ -1 & 0 & 1 \end{pmatrix} \to E_1^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix} = I - l_1 e_1^T *$$

$$E_2 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 3/2 & 1 \end{pmatrix} \to E_2^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -3/2 & 1 \end{pmatrix} = I - l_2 e_2^T *$$

LU

$$E_1 = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ -1 & 0 & 1 \end{pmatrix} \to E_1^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix} = I - l_1 e_1^T *$$

$$E_2 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 3/2 & 1 \end{pmatrix} \to E_2^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -3/2 & 1 \end{pmatrix} = I - l_2 e_2^T *$$

$$E_1^{-1} E_2^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 1 & -3/2 & 1 \end{pmatrix} = \mathbf{L}$$

LU

$A \in M_n$

En determinadas circunstancias*:

$$U = E_{n-1} ... E_2 E_1 A \Rightarrow A = E_1^{-1} E_2^{-1} ... E_{n-1}^{-1} U$$

con $E_K = I + l_k e_k^T$;

$$\mathrm{l}_k = (\underbrace{0..,0}_{k)} \, l_{k+1k} ... l_{nk})^T$$

siendo k el paso k-ésimo (bucle 2)

$$E_1^{-1} E_2^{-1} ... E_{n-1}^{-1} = L \text{ ¿?}$$

LU

$$\boxed{\text{Si } E_K = I + l_k e_k^T \ \rightarrow E_K^{-1} = I - l_k e_k^T}$$

(Se comprobó en ejercicio 3 de entrega Tema2)

Si es cierto, $E_1^{-1} E_2^{-1} ... E_{n-1}^{-1} = L$

- será una matriz triangular inferior por ser producto de triangulares inferiores
- tendrá 1s en la diagonal por construcción
- se construye colocando en cada elemento debajo de la diagonal de la columna k-ésima (bucle 2) el elemento por el que se ha multiplicado a la fila k-ésima para eliminar el elemento j-ésimo (bucle 1) **cambiado de signo**

Volvemos al ejemplo para concretar esta generalización:

$$A = \begin{pmatrix} 1 & 2 & 5 \\ -1 & 0 & 1 \\ 1 & -1 & 1 \end{pmatrix}$$

Hacemos cero los elementos de la primera columna debajo de 1 (bucle 2, **k=1**:

$$\xrightarrow[\text{bucle 1, } \mathbf{j=2}]{2^{\text{a}} f + 1\, 1^{\text{a}} f} \begin{pmatrix} 1 & 2 & 5 \\ 0 & 2 & 6 \\ 1 & -1 & 1 \end{pmatrix}$$

LU

Volvemos al ejemplo para concretar esta generalización:

$$A = \begin{pmatrix} 1 & 2 & 5 \\ -1 & 0 & 1 \\ 1 & -1 & 1 \end{pmatrix}$$

Hacemos cero los elementos de la primera columna debajo de 1 (bucle 2, **k=1**:

$$\xrightarrow{2^a f + 1 1^a f} \begin{pmatrix} 1 & 2 & 5 \\ 0 & 2 & 6 \\ 1 & -1 & 1 \end{pmatrix} \quad \xrightarrow{3^a f - 1 1^a f} \begin{pmatrix} 1 & 2 & 5 \\ 0 & 2 & 6 \\ 0 & -3 & -4 \end{pmatrix}$$

bucle 1, **j=2**

bucle 1, **j=3=n**, si no se seguiría hasta el final de la columna

LU

$$A_1 = \begin{pmatrix} 1 & 2 & 5 \\ 0 & 2 & 6 \\ 0 & -3 & -4 \end{pmatrix}$$

Hacemos 2, **k=2**:

$$\xrightarrow{3^a f + 3/2\, 2^a f} \begin{pmatrix} 1 & 2 & 5 \\ 0 & 2 & 6 \\ 0 & 0 & 5 \end{pmatrix} = U$$

bucle 1, **j=3** número de fila

$$A_1 = \begin{pmatrix} 1 & 2 & 5 \\ 0 & 2 & 6 \\ 0 & -3 & -4 \end{pmatrix}$$

Hacemos 2, **k=2**:

$$\xrightarrow{3^a f + 3/2 2^a f} \begin{pmatrix} 1 & 2 & 5 \\ 0 & 2 & 6 \\ 0 & 0 & 5 \end{pmatrix} = U$$

bucle 1, **j=3** número de fila

Bucle 2: se haría desde k=1 hasta n-1 con n número de columnas

Bucle 1: para cada columna, se haría desde j= k+1 hasta n con n número de filas $L = \begin{pmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 1 & -3/2 & 0 \end{pmatrix}$ $l_{jk} = \frac{a_{jk}^{k)}}{a_{kk}^{k)}}$

¿Cuándo se puede hacer una descomposición LU?

Si $A \in M_n$

$$\delta_k = det \begin{pmatrix} a_{11} & \dots & a_{1k} \\ & \ddots & \\ & & a_{kk} \end{pmatrix} \neq 0 \ \forall k \in \{1, \dots n\}$$

Lo demostramos tomando un paso genérico k y operando por bloques para comprobar que $a_{kk}^{k)} \neq 0$ **pivote**

$A_k = M_k A$ con $M_K = E_{k-1} \dots E_1$

$^{k)}$ significa de la matriz en el paso k-ésimo de bucle 2

LU

$$A_k = M_k A \text{ con } M_K = E_{k-1} \ldots E_1$$

$$A_k = \left(\begin{array}{ccc|c}
a_{11}^{1)} & & a_{1k}^{1)} & \\
 & \ddots & & * \\
0 & & a_{kk}^{k)} & \\
\hline
 & & a_{k+1,k}^{k)} & \\
 & & \vdots & * \\
 & & a_{k+1,k}^{k)} &
\end{array}\right) =$$

$$\left(\begin{array}{ccc|ccc}
1 & & & & & \\
 & \ddots & & & & \\
* & & 1 & & & \\
\hline
 & & & 1 & & \\
 & * & & & \ddots & \\
 & & & & & 1
\end{array}\right) \cdot \left(\begin{array}{ccc|c}
a_{11} & & a_{1k} & \\
 & \ddots & & * \\
a_{k1} & & a_{kk} & \\
\hline
 & * & & *
\end{array}\right)$$

LU

Operamos por bloques el primero (se demostró que se podía hacer en ejercicio 1 de entrega Tema2):

$$\begin{pmatrix} a_{11}^{1)} & & a_{1k}^{1)} \\ & \ddots & \\ 0 & & a_{kk}^{k)} \end{pmatrix} = \begin{pmatrix} 1 & & \\ & \ddots & \\ * & & 1 \end{pmatrix} . \begin{pmatrix} a_{11} & & a_{1k} \\ & \ddots & \\ a_{k1} & & a_{kk} \end{pmatrix}$$

Tomando determinantes en tema 2 se recordó que $detAB = detAdetB$:
$a_{11}^{1)} \ldots a_{kk}^{k)} = 1.\delta_k \neq 0$
Luego $a_{kk}^{k)} \neq 0$ y se puede tomar como pivote para la siguiente iteración

¿Cuántas descomposiciones LU existen para A?

La descomposición, si existe, es **única**

Dem: Supongamos $L_1 U_1 = A = L_2 U_2$

$\mathbf{det(A) = \delta_n = 1.detU_i \neq 0} \Rightarrow \exists U_i^{-1}, L_i^{-1}$
$L_2^{-1} L_1 = U_2 U_1^{-1}$

La matriz del primer miembro es triangular inferior, con 1s en la diagonal, y la del segundo es triangular superior, luego deben ser la identidad

$L_2^{-1} L_1 = I \Rightarrow L_2 = L_1$
$\mathrm{U}_2 U_1^{-1} = I \Rightarrow U_2 = U_1$

Ejercicios 2-4

¿Cómo se resuelve?

En este caso, la resolución del sistema se convierte en dos subsistemas, cada uno de ellos fácilmente resoluble según lo visto en la introducción.

$$\mathrm{Au} = \mathrm{b}$$

$$L\underbrace{Uu}_{z} = b$$

1. $Lz = b$

   sistema triangular inferior: $\rightarrow z_i = \frac{b_i - \sum_{j=1}^{i-1} l_{ij} z_j}{l_{ii}=1} \rightarrow \mathbf{z}$

2. $Uu = z$

   Sistema triangular superior $\rightarrow u_i = \frac{z_i - \sum_{j=i+1}^{n} u_{ij} u_j}{u_{ii}} \rightarrow \boxed{\mathbf{u}}$

Hasta ahora, hemos visto el caso de un sistema lineal cuya matriz A admite una factorialización LU.

- Pero esto no era posible siempre
- A admite factorialización LU cuando $\delta_k \neq 0, \forall k \in \{1, ..., n\}$
- A matriz inversible (sistema con solución) que no verifica esto ¿?

Sección 3

Descomposición PA=LU

$PA = LU$

L: matriz triangular inferior, con 1s en la diagonal
U: matriz triangular superior

Si no es posible esta factorialización, se debe a que algún pivote a emplear es cero.
La idea es **permutar** filas para evitar que actúe como pivote
En este caso, también se permutan las correspondientes filas del vector de términos independientes

Comenzamos con un ejemplo para visualizar el proceso

$$Au = b \text{ con } A = \begin{pmatrix} 1 & 2 & 5 \\ -1 & -2 & 0 \\ 1 & -1 & 1 \end{pmatrix} \text{ y } b = \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}$$

$\delta_2 = 0$

Hacemos cero los elementos de la primera columna debajo de 1:

Comenzamos con un ejemplo para visualizar el proceso

$$Au = b \text{ con } A = \begin{pmatrix} 1 & 2 & 5 \\ -1 & -2 & 0 \\ 1 & -1 & 1 \end{pmatrix} \text{ y } b = \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}$$

$\delta_2 = 0$

Hacemos cero los elementos de la primera columna debajo de 1:

$$\xrightarrow[3^a f - 1 1^a f]{2^a f + 1 1^a f} \begin{pmatrix} 1 & 2 & 5 \\ 0 & \mathbf{0} & 5 \\ 0 & -3 & -4 \end{pmatrix}$$ No es posible hacer lo mismo en la siguiente columna.

PA=LU

$$\begin{pmatrix} 1 & 2 & 5 \\ \color{red}{0} & \color{red}{\mathbf{0}} & 5 \\ \color{blue}{0} & -3 & -4 \end{pmatrix}$$

Sin embargo, sí se podrían **permutar** la segunda y tercera fila para continuar el proceso (o en este caso, concluirlo):

PA=LU

$$\begin{pmatrix} 1 & 2 & 5 \\ \color{red}{0} & \color{red}{\mathbf{0}} & 5 \\ \color{blue}{0} & -3 & -4 \end{pmatrix}$$

Sin embargo, sí se podrían **permutar** la segunda y tercera fila para continuar el proceso (o en este caso, concluirlo):

$$\xrightarrow{P^{23}} \begin{pmatrix} 1 & 2 & 5 \\ \color{blue}{0} & -3 & -4 \\ \color{red}{0} & \color{red}{\mathbf{0}} & 5 \end{pmatrix} = \mathbf{U}$$

PA = LU

Usamos la notación de la sección LU:

$$A_1 = E_1 P_1 A \text{ con: } E_1 = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ -1 & 0 & 1 \end{pmatrix} \text{ y } P_1 = P^{11} = I$$

PA = LU

Usamos la notación de la sección LU:

$$A_1 = E_1 P_1 A \text{ con: } E_1 = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ -1 & 0 & 1 \end{pmatrix} \text{ y } P_1 = P^{11} = I$$

$$A_2 = E_2 P_2 A_1 \text{ con: } E_2 = I \text{ y } P_2 = P^{23}$$

PA = LU

Usamos la notación de la sección LU:

$$A_1 = E_1 P_1 A \text{ con: } E_1 = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ -1 & 0 & 1 \end{pmatrix} \text{ y } P_1 = P^{11} = I$$

$$A_2 = E_2 P_2 A_1 \text{ con: } E_2 = I \text{ y } P_2 = P^{23}$$

$\boxed{\text{La inversa de una permutación de filas es ella misma: } (\mathrm{P}^{ij})^{-1} = P^{ij}}$
(entrega Tema 2)

PA = LU

$\boxed{PEP = E' \text{ sigue siendo una matriz de tipo E}}$

$$PE_kP = P(I + l_k e_k^T)P = PIP + (Pl_k)(e_K^T P)$$

$$PIP = I$$

PA = LU

$\boxed{PEP = E' \text{ sigue siendo una matriz de tipo E}}$

$PE_kP = P(I + l_k e_k^T)P = PIP + (Pl_k)(e_K^T P)$

$PIP = I$

$k < i \leq j$ se permutan filas por debajo de la diagonal o fila k

$Pl_k = l'_k$ vector de la forma: k primeras componentes cero, y el resto no necesariamente, con las componentes i y j intercambiadas respecto a l_k (i y j no son menores que k, porque se intercambian filas de la diagonal para abajo, nunca hacia arriba que ya hemos .ªrreglado").

PA = LU

$\boxed{PEP = E' \text{ sigue siendo una matriz de tipo E}}$

$PE_kP = P(I + l_k e_k^T)P = PIP + (Pl_k)(e_K^T P)$

$PIP = I$

$k < i \leq j$ se permutan filas por debajo de la diagonal o fila k

$Pl_k = l_k'$ vector de la forma: k primeras componentes cero, y el resto no necesariamente, con las componentes i y j intercambiadas respecto a l_k (i y j no son menores que k, porque se intercambian filas de la diagonal para abajo, nunca hacia arriba que ya hemos .ªrreglado").

$e_k^T P = e_k^T$ es la k-ésima fila de la matriz P o si se intercambian las componentes i y j (debajo de la k) como ambas son cero, queda igual.

PA = LU

$$U = A_2 = E_2 P_2 E_1 P_1 A$$

$$U = E_2 \underbrace{P_2 E_1 (P_2}_{E_1'} P_2) P_1 A = \underbrace{E_2 E_1'}_{E} \underbrace{P_2 P_1}_{P} A$$

$$P = P_2 P_1$$
$$\underbrace{E^{-1}}_{L} U = PA$$

PA = LU

En general:

$U = MA$ con:
$M = E_{n-1}P_{n-1}....E_1P_1A$

M es inversible:

- $det(M) = 1$ si hay un número par de permutaciones $\neq 0$
- $det(M) = -1$ si hay un número impar de permutaciones $\neq 0$

(determinante de producto: producto de determinantes). $\det(E_i) = 1$ y $\det(P_i)$=-1 $\forall i = 1...n-1$.

$$\boxed{Au = b \Longleftrightarrow MAu = \mathbf{M}b}$$

PA = LU

$U = MA$ con:
$M = E_{n-1}P_{n-1}....E_1P_1A$

$U = MA$ con:
$M = E_{n-1}P_{n-1}....P_2E_1(P_2P_2)P_1A = E_{n-1}P_{n-1}....P_3\underbrace{E_2E_1'}_{E_2^*}P_2P_1A =$

PA = LU

$U = MA$ con:
$M = E_{n-1}P_{n-1}....E_1P_1A$

$U = MA$ con:
$$M = E_{n-1}P_{n-1}....P_2E_1(P_2P_2)P_1A = E_{n-1}P_{n-1}....P_3\underbrace{E_2E_1'}_{E_2^*}P_2P_1A =$$
$$= E_{n-1}P_{n-1}....P_3E_2^*(P_3P_3)P_2P_1A = E_{n-1}P_{n-1}....E_2'P_3P_2P_1A =$$

PA = LU

$U = MA$ con:
$M = E_{n-1}P_{n-1}....E_1P_1A$

$U = MA$ con:
$$M = E_{n-1}P_{n-1}....P_2E_1(P_2P_2)P_1A = E_{n-1}P_{n-1}....P_3\underbrace{E_2E_1'}_{E_2^*}P_2P_1A =$$
$$= E_{n-1}P_{n-1}....P_3E_2^*(P_3P_3)P_2P_1A = E_{n-1}P_{n-1}....E_2'P_3P_2P_1A =$$
...
$$= E_{n-1}P_{n-1}E_{n-2}'(P_{n-1}P_{n-1})...P_2P_1A = E'P_{n-1}....P_2P_1A = E'PA$$

PA = LU

$U = MA$ con:
$M = E_{n-1}P_{n-1}....E_1P_1A$

$U = MA$ con:
$M = E_{n-1}P_{n-1}....P_2E_1(P_2P_2)P_1A = E_{n-1}P_{n-1}....P_3\underbrace{E_2E_1'}_{E_2^*}P_2P_1A =$
$= E_{n-1}P_{n-1}....P_3E_2^*(P_3P_3)P_2P_1A = E_{n-1}P_{n-1}....E_2'P_3P_2P_1A =$
...
$= E_{n-1}P_{n-1}E_{n-2}'(P_{n-1}P_{n-1})...P_2P_1A = E'P_{n-1}....P_2P_1A = E'PA$

$\underbrace{E'^{-1}}_{L}U = PA$

PA = LU

Estrategia de resolución de sistemas

Sistema original

$$Au = b \to PAu = L(Uu) = \mathbf{Pb}$$

Sí se permuta ahora el término independiente

$$\to \begin{array}{c} Lw = Pb \to w \\ Uu = w \to \mathbf{u} \end{array}$$

Como LU, se divide en dos sistemas ambos resolubles por remonte (sustitución)

Algunas consideraciones

- Gauss: $\dfrac{n^3 - n}{3}$ sumas y productos y $\dfrac{n(n-1)}{2}$ divisiones
  Remonte: $\dfrac{n(n-1)}{2}$ sumas y productos y n divisiones
  Cramer: $n! - 1$ sumas $(n-a)n!$ productos

n	op. Gauss	op.Cramer
10	805	399167999
100	581550	$\sim 10^{162}$
1000	668165500	$\sim 10^{2573}$

- En la práctica no lo hemos usado, pero existen estrategias para elegir pivote para evitar efectos como los vistos en T1 (aula Informática)

Ejercicios 4-6

Sección 4

Factorialización de Cholesky

Cholesky

- Particularización de lo anterior
- Muy útil cuando se trabaja con sistemas que no sean compatibles determinados como ahora (matrices no cuadradas)

Descomposición de Cholesky

$Au = b$ con A **simétrica y definida positiva**
$\exists A = BB^T$ con B matriz triangular inferior
Además, si $b_{ii} > 0$ la descomposición es única *

** no es la LU, no indica que haya 1s en la diagonal*

Cholesky

A es simétrica y definida positiva.
Si es definida positiva todos sus menores principales son positivos (luego no nulos)
$\Rightarrow \exists LU$

$A = LU$, $\forall k \in \{1, ..., n\}$ tomamos el producto de los primeros bloques de dimensión kxk y calculamos determinantes (tema2 y demostración de LU):

$$\underbrace{\delta_k}_{>0} = 1. \prod_{i=1}^{k} u_{ii} \Rightarrow u_{ii} > 0$$

(se podría ver por inducción, también)

Intercalamos una matriz diagonal
una vez asegurado que existe y que es inversible:

$$\boxed{\text{D} = \text{diag}(\sqrt{u_{11}}, ... \sqrt{u_{nn}})}$$

Existe porque si $u_{ii} > 0$ se puede calcular su raíz y se pueden invertir.

$$\boxed{A = LU = (LD)(D^{-1}U)} = \underbrace{\begin{pmatrix} \sqrt{u_{11}} & & & \\ \mu_{21} & \sqrt{u_{22}} & & \\ \vdots & \vdots & \ddots & \\ \mu_{n1} & \mu_{n2} & \dots & \sqrt{u_{nn}} \end{pmatrix}}_{B} \underbrace{\begin{pmatrix} \sqrt{u_{11}} & \nu_{12} & \dots & \nu_{1n} \\ & \sqrt{u_{22}} & \dots & \nu_{2n} \\ & & \ddots & \\ & & & \sqrt{u_{nn}} \end{pmatrix}}_{C}$$

Así: $A = LU = BC$ con B triangular inferior y C triangular superior

Cholesky

A es simétrica, $A = A^T \Rightarrow A^T = C^T B^T = BC$
$det(B) = det(B^T)$ (tema2). Por ser triangular, este determinante es igual al producto de los elementos de la diagonal:

$$\prod_{1}^{n} \sqrt{u_{ii}} \neq 0 \Rightarrow$$

B y B^T son inversibles y lo mismo ocurre con C:

$$A^T = C^T B^T = BC$$

Cholesky

A es simétrica, $A = A^T \Rightarrow A^T = C^T B^T = BC$
$det(B) = det(B^T)$ (tema2). Por ser triangular, este determinante es igual al producto de los elementos de la diagonal:

$$\prod_1^n \sqrt{u_{ii}} \neq 0 \Rightarrow$$

B y B^T son inversibles y lo mismo ocurre con C:

$$A^T = C^T B^T = BC$$

$$C^T B^T C^{-1} = B$$

Cholesky

A es simétrica, $A = A^T \Rightarrow A^T = C^T B^T = BC$
$det(B) = det(B^T)$ (tema2). Por ser triangular, este determinante es igual al producto de los elementos de la diagonal:

$$\prod_1^n \sqrt{u_{ii}} \neq 0 \Rightarrow$$

B y B^T son inversibles y lo mismo ocurre con C:

$$A^T = C^T B^T = BC$$

$$C^T B^T C^{-1} = B$$

$$\underbrace{B^T C^{-1}}_{i} = \underbrace{(C^T)^{-1} B}_{ii}$$

Cholesky

i: matriz triangular superior con 1s en la diagonal

ii: matriz triangular inferior con 1s en la diagonal

Para que se de la igualdad, debe ser I

Cholesky

i: matriz triangular superior con 1s en la diagonal

ii: matriz triangular inferior con 1s en la diagonal

Para que se de la igualdad, debe ser I

$$\Rightarrow I = B^T C^{-1} = (C^T)^{-1} B \Rightarrow C = B^T \Rightarrow \boxed{A = BB^T}$$

y los elementos de la diagonal: $b_{ii} = \sqrt{u_{ii}}$

Cholesky

Unicidad de factorialización

Suponemos que hay dos distintas:

$B_1 B_1^T = A = B_2 B_2^T$

$(b_1)_{ii} > 0; \ (b_2)_{ii} > 0 \, \forall i$

Unicidad de factorialización

Suponemos que hay dos distintas:

$B_1 B_1^T = A = B_2 B_2^T$

$(b_1)_{ii} > 0;\ (b_2)_{ii} > 0\, \forall i$

$det(B_k) = det(B_k^T),\ (k = 1, 2)\ \ = \prod_1^n (b_k)_{ii}$

Los elementos de la diagonal no son cero: se pueden invertir:

Cholesky

Unicidad de factorialización

Suponemos que hay dos distintas:

$B_1 B_1^T = A = B_2 B_2^T$

$(b_1)_{ii} > 0; \ (b_2)_{ii} > 0 \, \forall i$

$det(B_k) = det(B_k^T), \, (k = 1, 2) \ = \prod_1^n (b_k)_{ii}$

Los elementos de la diagonal no son cero: se pueden invertir:

$B_1 B_1^T = A = B_2 B_2^T \Rightarrow B_2^{-1} B_1 B_1^T = B_2^T$

Cholesky

Unicidad de factorialización

Suponemos que hay dos distintas:

$B_1 B_1^T = A = B_2 B_2^T$

$(b_1)_{ii} > 0; \ (b_2)_{ii} > 0 \, \forall i$

$det(B_k) = det(B_k^T), \ (k = 1, 2) \ = \prod_1^n (b_k)_{ii}$

Los elementos de la diagonal no son cero: se pueden invertir:

$B_1 B_1^T = A = B_2 B_2^T \Rightarrow B_2^{-1} B_1 B_1^T = B_2^T$

$B_2^{-1} B_1 = B_2^T (B_1^T)^{-1}$

Unicidad de factorialización

Suponemos que hay dos distintas:

$B_1 B_1^T = A = B_2 B_2^T$

$(b_1)_{ii} > 0;\ (b_2)_{ii} > 0\, \forall i$

$det(B_k) = det(B_k^T),\ (k = 1, 2)\ = \prod_1^n (b_k)_{ii}$

Los elementos de la diagonal no son cero: se pueden invertir:

$B_1 B_1^T = A = B_2 B_2^T \Rightarrow B_2^{-1} B_1 B_1^T = B_2^T$

$B_2^{-1} B_1 = B_2^T (B_1^T)^{-1}$

Repitiendo el razonamiento anterior, vemos que ambos miembros son la identidad:

$\boxed{\Rightarrow B_1 = B_2}$ $d_{ii} = \frac{b_{ii}^1}{b_{ii}^2} = \frac{b_{ii}^2}{b_{ii}^1} \Rightarrow d_{ii} = 1$

Estrategia de resolución de sistemas

De nuevo es la misma: se resuelven dos sistemas triangulares:
$Au = b \to B \underbrace{B^T u}_{w} = b \to$

$$Bw = b \to w$$
$$B^T u = w \to u$$

Sistemas Lineales. Métodos Iterativos

Métodos Numéricos

Curso 2021-22

Sección 1

Idea general y conceptos

Idea general y conceptos

¿Para qué este tema?

$Au = b$ se busca como límite de una sucesión de vectores (solución aproximada)

- Si A es muy grande y tarda mucho en resolverse
- Si por el tamaño de A o de sus elementos, se van acumulando errores de redondeo
- Si A tiene muchos elementos casi cero que den lugar a errores de algoritmo

Solución

$\{x_k\}$ con

- x_o arbitraria (convergencia más rápida si es más próxima a la solución)
- $x_{k+1} = Bx_k + c$ problema también lineal

B: matriz del método

c: vector del método

Idea general y conceptos

Solución

$\{x_k\}$ con

- x_o arbitraria (convergencia más rápida si es más próxima a la solución)
- $x_{k+1} = Bx_k + c$ problema también lineal

B: matriz del método

c: vector del método

Se eligen a partir de A y b con los métodos de este tema.

Un método es **convergente** si $\forall x_o \in \mathbb{R}^n, \{x_k\}$ es una sucesión convergente, es decir, $\exists \lim_{k\to\infty} \{x_k\} = \mathbf{u}$

Tomamos límites en $x_{k+1} = Bx_k + c$

$u = Bu + c \Rightarrow (I - B)u = c*$

B se elige de modo que I - B sea inversible y * y $Au = b$ sean sistemas equivalentes.
Método bien definido

Idea general y conceptos

Error

El error en cada iteración es:

$$e_k = x_k - u$$

$= (\mathrm{B}x_{k-1} + c) - (Bu + c) = Be_{k-1}$
$= \mathrm{B}((\mathrm{B}x_{k-2} + c) - (Bu + c)) = B^2 e_{k-2} \ldots \ldots \Rightarrow e_k = B^k e_o \Rightarrow$
El error depende de las potencias de B.
Para que haya convergencia $e_k \to 0 \Rightarrow B^k \to 0$

Son equivalentes:

- Un método iterativo es **convergente**
- $\rho(B) < 1$, radio espectral o mayor de autovalores en valor absoluto
- $\exists ||| \ \ ||| / |||B||| < 1$

Idea general y conceptos

En la práctica:

- B? $\rightarrow$ métodos de resolución
- ¿qué B elegir entre los métodos?

1. comprobar si I-B es inversible y que los sistemas son equivalentes **método bien construido**
2. comprobar si $\rho(B) < 1$ o $|||B||| < 1$ **(convergente)**

Cuanto menor sea $\rho(B)$ la convergencia es más rápida.

Los métodos clásicos se basan en descomponer A = M - N con M fácilmente inversible (diagonal, triangular)

$$Au = b \Longleftrightarrow (M - N)u = b \Longleftrightarrow Mu = Nu + b \Longleftrightarrow$$

$$u = \underbrace{M^{-1}N}_{B} u + \underbrace{M^{-1}b}_{c}$$

$$I - B = M^{-1}M - M^{-1}N = M^{-1}(M - N) = M^{-1}A$$

Método bien definido si M tiene inversa

En adelante conviene separar:

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ & & \ddots & \\ a_{n1} & & \dots & a_{nn} \end{pmatrix}$$

$$D = \begin{pmatrix} a_{11} & & & \\ & a_{22} & & \\ & & \ddots & \\ & & & a_{nn} \end{pmatrix} \quad E = \begin{pmatrix} 0 & & & 0 \\ -a_{21} & 0 & & \\ \vdots & & \ddots & \vdots \\ -a_{n1} & & \dots & 0 \end{pmatrix}$$

$$F = \begin{pmatrix} 0 & -a_{12} & & -a_{1n} \\ 0 & 0 & & -a_{2n} \\ \vdots & & \ddots & \vdots \\ 0 & & \dots & 0 \end{pmatrix} \quad \boxed{\mathrm{A} = \mathrm{D} - \mathrm{E} - \mathrm{F}}$$

Sección 2

Método de Jacobi

Jacobi

Como en Tema 3, primero vemos cómo funciona y luego lo formalizamos

- de cada i-ésima ecuación se despeja la i-ésima incógnita
- se parte de una solución inicial $x_o \to x_1$
- se itera

Ejemplo

$x_o = 1, y_o = 2$ no son solución.

$$\left.\begin{array}{rcl} 5x + 2y & = & 1 \\ x - 4y & = & 0 \end{array}\right\}$$

Ejemplo

$x_o = 1, y_o = 2$ no son solución.

$$\left.\begin{array}{rcl} 5x + 2y & = & 1 \\ x - 4y & = & 0 \end{array}\right\}$$

- $$\left.\begin{array}{rcl} x & = & \frac{-2}{5}y + \frac{1}{5} \\ y & = & \frac{1}{4}x \end{array}\right\}$$

Ejemplo

$x_o = 1, y_o = 2$ no son solución.

$$\left.\begin{array}{rcl} 5x + 2y & = & 1 \\ x - 4y & = & 0 \end{array}\right\}$$

- $\left.\begin{array}{rcl} x & = & \frac{-2}{5}y + \frac{1}{5} \\ y & = & \frac{1}{4}x \end{array}\right\}$
- $\begin{array}{rcl} x_1 & = & \frac{-4}{5} + \frac{1}{5} = \mathbf{-0,6} \\ y_1 & = & \frac{1}{4} = \mathbf{0,25} \end{array}$ $\underline{1^{a}\ \text{iteración}}$

Ejemplo

$x_o = 1, y_o = 2$ no son solución.

$$\left.\begin{array}{rcl} 5x + 2y & = & 1 \\ x - 4y & = & 0 \end{array}\right\}$$

- $$\left.\begin{array}{rcl} x & = & \frac{-2}{5}y + \frac{1}{5} \\ y & = & \frac{1}{4}x \end{array}\right\}$$
- $$\begin{array}{rcl} x_1 & = & \frac{-4}{5} + \frac{1}{5} = \mathbf{-0,6} \\ y_1 & = & \frac{1}{4} = \mathbf{0,25} \end{array}$$ 1ª iteración
- $$\begin{array}{rcl} x_2 & = & \frac{-2}{5}y + \frac{1}{5} = \mathbf{0,1} \\ y_2 & = & \frac{1}{4}(-0,6) = \mathbf{-0,15} \end{array}$$ 2ª iteración

......

Jacobi

Recuperando A = D - E -F: $A = M - N \rightarrow \mathbf{M} = \mathbf{D}$ fácilmente inversible

por ser diagonal, y $\mathbf{N} = \mathbf{E} + \mathbf{F}$

Método de Jacobi

Matriz del método $B = M^{-1}N = D^{-1}(E + F)$: **matriz de Jacobi, J**
Vector del método: $c = M^{-1}b = D^{-1}b$

Jacobi

Formalización del Método de Jacobi

Iteración k-ésima + 1:

$$x^{k+1)} \underbrace{=}_{Bx^{k)}+c=M^{-1}Nx^{k)}+M^{-1}b} = D^{-1}(E+F)x^{k)} + D^{-1}b$$

Componente i-ésima:

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

Sección 3

Método de Gauss-Siedel

Gauss -Siedel

Es similar pero según se determinan los valores, se usan en la misma iteración y no en la siguiente:

$x_o = 1, y_o = 2$ no son solución.

$$\left.\begin{array}{rcl} 5x + 2y & = & 1 \\ x - 4y & = & 0 \end{array}\right\}$$

Es similar pero según se determinan los valores, se usan en la misma iteración y no en la siguiente:

$x_o = 1, y_o = 2$ no son solución.

$$\left.\begin{array}{rcl} 5x + 2y & = & 1 \\ x - 4y & = & 0 \end{array}\right\}$$

- $$\left.\begin{array}{rcl} x & = & \frac{-2}{5}y + \frac{1}{5} \\ y & = & \frac{1}{4}x \end{array}\right\}$$

Gauss -Siedel

Es similar pero según se determinan los valores, se usan en la misma iteración y no en la siguiente:

$x_o = 1, y_o = 2$ no son solución.

$$\left.\begin{array}{rcl} 5x + 2y & = & 1 \\ x - 4y & = & 0 \end{array}\right\}$$

- $$\left.\begin{array}{rcl} x & = & \frac{-2}{5}y + \frac{1}{5} \\ y & = & \frac{1}{4}x \end{array}\right\}$$
- $$\begin{array}{rcl} x_1 & = & -4/5 + 1/5 = -\mathbf{0,6} \\ y_1 & = & 1/4(-0,6) = -\mathbf{0,15} \end{array}$$

que era el resultado de la 2ª iteración (se acelera)

Gauss -Siedel

Es similar pero según se determinan los valores, se usan en la misma iteración y no en la siguiente:

$x_o = 1, y_o = 2$ no son solución.

$$\left.\begin{array}{rcl} 5x + 2y & = & 1 \\ x - 4y & = & 0 \end{array}\right\}$$

- $$\left.\begin{array}{rcl} x & = & \frac{-2}{5}y + \frac{1}{5} \\ y & = & \frac{1}{4}x \end{array}\right\}$$
- $$\begin{array}{rcl} x_1 & = & -4/5 + 1/5 = \mathbf{-0,6} \\ y_1 & = & 1/4(-0,6) = \mathbf{-0,15} \end{array}$$

que era el resultado de la 2ª iteración (se acelera)

- $$\begin{array}{rcl} x_2 & = & -2/5(-0,15) + 1/5 = \mathbf{0,26} \\ y_1 & = & 1/4(0,26) = \mathbf{0,065} \end{array}$$

Recuperando A = D - E -F:

Para identificar cada elemento recordamos:
$Mx = Nx + b \to \mathbf{M} = \mathbf{D} - \mathbf{E}$ fácilmente inversible por ser triangular, y $\mathbf{N} = \mathbf{F}$

Método de Gauss-Siedel

Matriz del método $B = M^{-1}N = (D - E)^{-1}F$:
matriz de Gauss-Siedel, $\mathcal{L}_1$
Vector del método: $c = M^{-1}b = (D - E)^{-1}b$

Formalización del Método de Gauss-Siedel

Iteración k-ésima + 1:

$$Dx^{k+1)} = b + Ex^{k+1)} + Fx^{k)} \Rightarrow$$
$$\Rightarrow (D - E)x^{k+1)} = Fx^{k)} + b$$

Componente i-ésima:

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

Ejercicio 1

Sección 4

Método de Relajación

Relajación

- Se trata en realidad de una modificación de lo anterior
- Busca convergencia más rápida (matriz con menor radio espectral)
- Para ello hace la media ponderada de la iteración actual y la anterior

$$x^{k)} \xrightarrow{\text{método}} x^{k+1)} \rightarrow \omega x^{k+1)} + (1-\omega)x^{k)}$$

ω factor de peso

Relajación

Jacobi: $x^{k+1)} = \omega \underset{*}{\mathbf{x_J^{k+1)}}} + (1-\omega)x_J^{k)}$

Es decir:

$$\underbrace{\frac{\omega}{a_{ii}}\left(b_i - \sum_{j=1}^{i-1} a_{ij}x_j^{k)} - \sum_{j=i+1}^{n} a_{ij}x_j^{k)}\right)}_{*} + (1-\omega)x_i^{k)}$$

Matricialmente:

$$\omega D^{-1}(b + (E+F)x^{k)}) + (1-\omega)x^{k)}$$

Juntando elementos con $x^{k)}$:

$$\omega D^{-1}\left[\frac{1-\omega}{\omega}D + E + F\right]x^{k)} + \omega D^{-1}b$$

Se usa poco

Relajación

Gauss - Siedel: $x_i^{k+1)} = \omega \mathbf{x_{iGS}^{k+1)}} + (1-\omega)x_i^{k)}$
Es decir:

$$\underbrace{\frac{\omega}{a_{ii}}\left(b_i - \sum_{j=1}^{i-1} a_{ij}x_j^{k+1)} - \sum_{j=i+1}^{n} a_{ij}x_j^{k)}\right)}_{*} + (1-\omega)x_i^{k)}$$

Matricialmente:

$$x^{k+1)} = \omega D^{-1}(b + Ex^{k+1)}) + \omega D^{-1}Fx^{k)} + (1-\omega)x^{k)}$$

Juntando elementos con $x^{k+1)}$ y multiplicando por D (para evitar la D^{-1}):

$$(D-\omega E)x^{k+1)} = [D(1-\omega) + \omega F]\, x^{k)} + \omega b$$

Gauss - Siedel:

$$(D-\omega E)x^{k+1)} = [D(1-\omega)+\omega F]\, x^{k)} + \omega b$$

Dividiendo entre ω para evitar que b esté multiplicado por ω y así se parezca al sistema que se busca, $Bx = c$ y poder identificar a B :

$$\left(\frac{D}{\omega} - E\right) x^{k+1)} = \left[D\left(\frac{1-\omega}{\omega}\right) + F\right] x^{k)} + b$$

$A = M - N$ y aquí, $\boxed{M = \frac{D}{\omega} - E}$
y $\boxed{N = \frac{1-\omega}{\omega} D + F}$

Comprobación: $M - N = \frac{D}{\omega} - E - \frac{1-\omega}{\omega} D - F = D - E - F = A$

Entonces, el método es:

- x_o
- $x^{k+1)} = \left(\frac{D}{\omega} - E\right)^{-1} \left[\left(D\left(\frac{1-\omega}{\omega}\right) + F\right) x^{k)} + b\right]$

$\mathcal{L}_\omega$, matriz del método es

$$\left(\frac{D}{\omega} - E\right)^{-1} \left(D\left(\frac{1-\omega}{\omega}\right) + F\right) = \boxed{(\text{D-}\omega E)^{-1}\left[(1-\omega)D + F\omega\right]}$$

Ejercicios 2a, 3 (también puede hacerse después de resultados de convergencia)

Sección 5

Métodos por bloques

Métodos por bloques

Es análogo a lo visto: $A = D - E - F$

$$A = \left(\begin{array}{c|c|c|c} A_{11} & A_{12} & \cdots & A_{1p} \\ \hline A_{21} & A_{22} & \cdots & A_{2p} \\ \hline \dots & \dots & \ddots & \dots \\ \hline A_{p1} & A_{p2} & \cdots & A_{pp} \end{array}\right)$$

Se formula todo igual, ya que D (en este caso las matrices que ocupan la diagonal, A_{ii}) son inversibles $i = 1, ...p$.
Desarrollo en problemas.

Sección 6

Convergencia de los métodos

Convergencia

Hemos visto la formulación de los principales métodos iterativos.
Ahora queremos saber de la forma más ágil posible si convengen.

- $I - B$ inversible (o $\exists M^{-1}$)
- $\rho(B) < 1 \leftrightarrow \exists ||| \; ||| / |||B||| < 1$

¿Se puede hacer más sencillo?
De forma general no, pero hay algunos resultados para determinados tipos de matrices.

Método de Jacobi

Si A (matriz original, no B) es diagonal estrictamente dominante, ($|a_{ii}| > \sum_{j\neq i} |a_{ij}|$) **el método de Jacobi converge**

Dem.

$$J = D^{-1}(E+F) \Rightarrow J_{ij} = \begin{cases} \frac{-a_{ij}}{a_{ii}} & \text{si } i \neq j \\ 0 & \text{si } i = j \end{cases}$$

Método de Jacobi

Si A (matriz original, no B) es diagonal estrictamente dominante, ($|a_{ii}| > \sum_{j\neq i} |a_{ij}|$) **el método de Jacobi converge**

Dem.

$$J = D^{-1}(E+F) \Rightarrow J_{ij} = \begin{cases} \frac{-a_{ij}}{a_{ii}} & \text{si } i \neq j \\ 0 & \text{si } i = j \end{cases}$$

$$|||J|||_\infty \underbrace{=}_{\text{suma filas}} \max_{1\leq i\leq n} \sum_1^n |J_{ij}| = \max_{1\leq i\leq n} \sum_{\substack{j=1 \\ i\neq j}}^n \frac{|a_{ij}|}{|a_{ii}|} = \max_{1\leq i\leq n} \left(\frac{1}{|a_{ii}|} \sum_{\substack{j=1 \\ i\neq j}}^n |a_{ij}| \right)$$

Convergencia para método de Jacobi

Conclusión

Es convergente si este valor es $< 1 \iff |a_{ii}| > \sum_{\substack{j=1 \\ i \neq j}}^{n} |a_{ij}|$

(en realidad, con algún detalle vale como resultado también para el método de Gauss Siedel) Ejercicio 3

*El radio espectral de la matriz del método de relajación $\rho(\mathcal{L}_\omega) \geq |1-\omega|$
$\omega \neq 0 \Rightarrow$ el método solo puede ser convergente para $0 < \omega < 2$
(Gauss Siedel se considera relajación con $\omega = 1$).

Dem.

$$det(\mathcal{L}_\omega) = det\left(\left(\frac{D}{\omega} - E\right)^{-1}\left(\frac{1-\omega}{\omega}D + F\right)\right) \underset{\text{determinante de producto}}{=}$$

$$\frac{det\left(\frac{1-\omega}{\omega}D + F\right)}{det\left(\frac{D}{\omega} - E\right)} \underset{\text{por ser matrices triangulares, Tema 2}}{=} \frac{det\left(\frac{1-\omega}{\omega}D\right)}{det\left(\frac{D}{\omega}\right)} =$$

$$\frac{\frac{(1-\omega)^n}{\omega^n}detD}{\frac{1}{\omega^n}detD} = (1-\omega)^n$$

Convergencia para método de Relajación

Por otra parte, $det(\mathcal{L}_\omega) = \prod_1^n \lambda_i$, con λ_i autovalores de $\mathcal{L}_\omega$

$\Rightarrow \prod_1^n |\lambda_i| = |1-\omega|^n \Rightarrow$

$\rho(\mathcal{L}_\omega) \geq \left(\prod_1^n |\lambda_i|\right)^{1/n} = |1-\omega|$

Es decir, solo es condición necesaria. Ejercicios 4 y 5

Convergencia para sistemas con matrices simétricas y definidas positivas

Si A es **hermítica (simétrica en $\mathbb{R}$)** y **definida positiva** con A = M-N, M inversible, si $\underline{M^T + N \text{ es definida positiva y simétrica}}$: $\boxed{\rho(M^{-1}N) < 1} \Rightarrow$ el método definido con $B = M^{-1}N$ es **convergente**

* *este resultado vale para todos los métodos, aunque lo usaremos como necesario para el siguiente resultado sobre Gauss Siedel*

Convergencia para método Gauss -Siedel

Si A es matriz **hermítica** (simétrica si es real) y **definida positiva**, y $0 < \omega < 2$, el método de relajación y en concreto Gauss Siedel, es convergente.

Dem.

$$A = M - N = \left(\left(\frac{D}{\omega} - E \right) - \left(\frac{1-\omega}{\omega} D + F \right) \right), \omega \neq 0 \text{ (i)}$$

A es hermítica (simétrica en reales) $\Rightarrow$
A = D - E - F = A* $= D^* - E^* - F^*$ (ii)

Convergencia para método Gauss -Siedel

Si A es matriz **hermítica** (simétrica si es real) y **definida positiva**, y $0 < \omega < 2$, el método de relajación y en concreto Gauss Siedel, es convergente.

Dem.

$$A = M - N = \left(\left(\frac{D}{\omega} - E\right) - \left(\frac{1-\omega}{\omega}D + F\right)\right), \omega \neq 0 \text{ (i)}$$

A es hermítica (simétrica en reales) $\Rightarrow$
A = D - E - F = A* $= D^* - E^* - F^*$ (ii)
De (ii): $D = D^*$; $E = F^*$

Si A es matriz **hermítica** (simétrica si es real) y **definida positiva**, y $0 < \omega < 2$, el método de relajación y en concreto Gauss Siedel, es convergente.

Dem.

$$A = M - N = \left(\left(\frac{D}{\omega} - E \right) - \left(\frac{1-\omega}{\omega} D + F \right) \right), \omega \neq 0 \text{ (i)}$$

A es hermítica (simétrica en reales) $\Rightarrow$

A = D - E - F = A* $= D^* - E^* - F^*$ (ii)

De (ii): $D = D^*$; $E = F^*$

De (i):

$$M^* + N = \frac{D}{\omega} - E^* + \frac{1-\omega}{\omega} D + F = \frac{D}{\omega} - \cancel{F} + \frac{1-\omega}{\omega} D + \cancel{F} = \frac{2-\omega}{\omega} D$$

Convergencia para método Gauss-Siedel

$$M^* + N = \frac{2-\omega}{\omega} D$$

D es definida positiva (por serlo A, ver álgebra Lineal)
$\Rightarrow M^* + N$ es definida positiva para valores $0 < \omega < 2$

Como es definida positiva: $\Rightarrow \underline{0 < \omega < 2}$
Teorema de Ostrowski-Reich

Ejercicios 6 y en adelante

Sección 7

Test de parada

Test de parada

Cuando un método iterativo es convergente, la solución es $\lim_{k \to \infty} x_k = u$, pero en la práctica no seguimos hasta el infinito sino que se aplica un número finito de veces. ¿Cómo hacerlo para tener una aproximación razonable?

Idea para definir test de parada

Queremos determinar n / x_n sea una buena aproximación a u, solución del sistema.

Entendemos por "buena" **error relativo** $< \epsilon$:

$$\frac{\|x_k - u\|}{\|u\|} < \epsilon$$

El problema es que es una expresión teórica. No conocemos u.

Test de parada

Definición: r_k **residuo** en la iteración k: $Au - Ax_k = A(u - x_k)$

$$\frac{\|r_k\|}{\|b\|} = \frac{\|A(u - x_k)\|}{\|b\| = \|Au\|} < \epsilon \Rightarrow \boxed{\|r_k\| < \epsilon}$$

Vemos como se aplica en cada método

Test de parada

Jacobi:
Se despejaba en la ecuación i-ésima la variable i-ésima o de forma matricial:

$$Dx_{k+1} = b + (E+F)x_k \underset{*}{=} b - Ax_k + \underset{*}{Dx_k} =$$

* como queremos que aparezca r_k hay que hacer que aparezca en la expresión A (en lo que había, falta D). Al introducirlo, aparece un Dx_k que no estaba originalmente, por lo que se compensa con el último término para mantener la igualdad.

$$= r_k + Dx_k \Rightarrow \boxed{\mathrm{r}_k = D(x_{k+1} - x_k)}\, ii)$$

$(Au = b; r_k = Au - Ax_k)$

Test de parada

Es decir, tenemos el sistema: $D \underbrace{d_k} = r_k$ i)

con d_k diferencia entre términos consecutivos: $x_{k+1} = x_k + d_k$

PASOS:

- se calcula r_k $(r_k)_i = b_i - \sum_{j=1}^{n} a_{ij} x_j^k$
- se resuelve i) $d_i^k = \dfrac{r_i^k}{a_{ii}}$
- iteración siguiente ii) $x_i^{k+1} = x_i^k + d_i^k$

Test de parada

Relajación:
Entra aquí Gauss Siedel tomando $\omega = 1$.

$$\left(\frac{D}{\omega} - E\right) x_{k+1} = \left(\frac{1-\omega}{\omega} D + F\right) x_k + b =$$

De nuevo, tratamos de que aparezca lo que definimos como residuo:

$$\frac{D}{\omega} x_{k+1} = E x_{k+1} - D x_k + F x_k + \frac{D}{\omega} x_k + b = \tilde{r}_k + \frac{D}{\omega} x_k$$

Llamamos $\tilde{r}_k$ no exactamente al residuo, sino al valor:
$b - ((D - F) x_k - E x_{k+1})$

Test de parada

Así:

$$\frac{D}{\omega}x_{k+1} = \tilde{r}_k + \frac{D}{\omega}x_k \Rightarrow D(x_{k+1} - x_k) = \tilde{r}_k\omega$$

Con esta definición "modificada" del residuo, sigue siendo válido el test:

$$\|\tilde{r}_k\| < \epsilon\|b\|$$

Test de parada

PASOS:

- se calcula $\tilde{r}_k$: $(\tilde{r}_k)_i = b_i - \sum_{j=1}^{i-1} \underbrace{a_{ij}}_{E} x_j^{k+1)} - \sum_{j=i}^{n} \underbrace{a_{ij}}_{D-F} x_j^{k)}$
- se resuelve el análogo a i): $d_i^k = \dfrac{\omega \tilde{r}_i^k}{a_i i}$
- iteración siguiente, análogo a ii): $x_i^{k+1} = x_i^k + d_i^k$