MÉTODOS NUMÉRICOS I

Tema II: Resolución numérica de sistemas de ecuaciones lineales

Manuel Ruiz Galán

Curso 2019/2020
Doble Grado en Ingeniería Informática y Matemáticas

UNIVERSIDAD DE GRANADA

Departamento de Matemática Aplicada

Índice Tema II

II.1 Métodos directos: Gauss y versiones, factorización de matrices

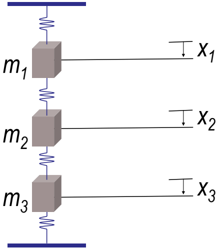
Tratamiento numérico de sistemas de ecuaciones lineales

- Métodos *directos*: Gauss, versiones y factorizaciones
- Métodos *iterativos*: Jacobi y Gauss–Seidel

Interés | herramienta matemática
| problemas vida real

Motivación

4 muelles alineados
3 cuerpos entre los mismos de masas m_1, m_2 y m_3
pesos p_1, p_2 y p_3
sistema en *equilibrio*



Objetivo: determinar los desplazamientos x_1, x_2 y x_3 en función de p_1, p_2 y p_3

d_j deformación del muelle j, $j = 1, 2, 3, 4$
vectores de desplazamientos, deformaciones, fuerzas de reacción de los muelles y pesos:

$$\mathbf{x} := [x_1, x_2, x_3]^T,$$

$$\mathbf{d} := [d_1, d_2, d_3, d_4]^T$$

$$\mathbf{y} := [y_1, y_2, y_3, y_4]^T,$$

$$\mathbf{p} := [m_1 g, m_2 g, m_3 g]^T$$

$\downarrow$ multiplicar por matriz

$$\mathbf{Ax} = \mathbf{d}, \quad \mathbf{Cd} = \mathbf{y}, \quad \mathbf{A}^T\mathbf{y} = \mathbf{p}$$

relación entre desplazamientos y deformaciones

$$d_1 = x_1, \quad d_2 = x_2 - x_1, \quad d_3 = x_3 - x_2, \quad d_4 = -x_3$$

$$\Updownarrow$$

$$\mathbf{d} = \mathbf{A}\mathbf{x}$$

$$\mathbf{A} := \begin{bmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 0 & -1 & 1 \\ 0 & 0 & -1 \end{bmatrix}$$

Ley de Hooke: $c_1, c_2, c_3, c_4 > 0$ (*coeficientes de elasticidad*)

$$j = 1,2,3,4 \ \Rightarrow \ y_j = c_j d_j \ \Leftrightarrow \ \mathbf{y} = \mathbf{C}\mathbf{d}$$

$$\mathbf{C} := \begin{bmatrix} c_1 & 0 & 0 & 0 \\ 0 & c_2 & 0 & 0 \\ 0 & 0 & c_3 & 0 \\ 0 & 0 & 0 & c_4 \end{bmatrix}$$

sistema en equilibrio $\rightsquigarrow$ fuerza neta en cada cuerpo nula

$$j = 1, 2, 3 \;\Rightarrow\; p_j = y_j - y_{j+1} \;\Leftrightarrow\; \mathbf{p} = \mathbf{A}^T\mathbf{y}$$

matriz que relaciona **x** y **d** = transpuesta que conecta **y** con **p**
trabajo interno de deformación de los muelles = trabajo externo de los cuerpos

$$\mathbf{y}^T\mathbf{d} = \mathbf{y}^T\mathbf{A}\mathbf{x} = \mathbf{p}^T\mathbf{x}$$

Resumen

$$\mathbf{A}\mathbf{x} = \mathbf{d}, \quad \mathbf{C}\mathbf{d} = \mathbf{y}, \quad \mathbf{A}^T\mathbf{y} = \mathbf{p} \;\Rightarrow\; \mathbf{A}^T\mathbf{C}\mathbf{A}\mathbf{x} = \mathbf{p}$$

K *matriz de rigidez*

$$\mathbf{K} := \mathbf{A}^T\mathbf{C}\mathbf{A}$$

Solución del problema = solución del sistema de ecuaciones lineales

$$\mathbf{K}\mathbf{x} = \mathbf{p}$$

x minimiza la energía potencial del sistema (Tema V) sistema $N \times N$ unisolvente, N grande $\rightsquigarrow$ *regla de Cramer* ineficiente

complejidad $\rightsquigarrow$ enorme número de operaciones elementales

determinante de $\mathbf{A} \in \mathbb{R}^{N\times N}$

$$\det(\mathbf{A}) = \sum_{\sigma \in \Delta_N} \operatorname{sign}(\sigma) a_{\sigma(1)1} \cdots a_{\sigma(N)N}$$

permutación

$$\underbrace{\sum_{\sigma \in \Delta_N} \operatorname{sign}(\sigma) \underbrace{a_{\sigma(1)1} \cdots a_{\sigma(N)N}}_{N-1 \text{ productos}}}_{N!-1 \text{ sumas}} \rightsquigarrow N!(N-1) + N! - 1 = N!N - 1 \text{ operaciones}$$

(obviando consideraciones sobre la memoria)

regla de Cramer $N+1$ determinantes $+$ N divisiones

$$(N+1)(N!N-1) + N = (N+1)!N - 1 \text{ operaciones}$$

$N = 25$

$$1.008228652816514 \cdot 10^{28} \text{ operaciones}$$

10^9 operaciones/segundo

$$1.008228652816514 \cdot 10^{19} \text{ segundos}$$

(1 año $=$ 365.25 días)

$$3.194883808706981 \cdot 10^{11} \text{ años}$$

II.1.1. Sistemas triangulares

Para SCD, determinante no nulo = ningún cero en la diagonal principal

$$\mathbf{U}\mathbf{x} = \mathbf{b}$$

$\mathbf{U} \in \mathbb{R}^{N\times N}$ triangular superior con elementos diagonales no nulos
$\mathbf{x} \in \mathbb{R}^N$ vector de incógnitas
$\mathbf{b}$ vector de términos independientes

Resolución por sustitución hacia atrás

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

Lo de toda la vida

Si va de N+1 hasta N por convenio es 0.

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

no nulo

este ya lo conoces es el que has calculado antes

despejar de la segunda ecuación

Ídem

$$\mathbf{Lx} = \mathbf{b}$$

$\mathbf{L} \in \mathbb{R}^{N\times N}$ matriz triangular inferior con elementos diagonales no nulos

Resolución por sustitución hacia adelante

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

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

Ejercicio

Dado $N \geq 1$ se tiene que

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

(Indicación: puede procederse por inducción o alternativamente, y de forma constructiva, llamar

$$\alpha := \sum_{j=1}^{N} j$$

y observar que

$$\begin{array}{cccccccc} \alpha & = & 1 & + & 2 & + & \cdots & +N \\ & = & N & + & N-1 & + & \cdots & +1 \end{array}$$

y sumar las "columnas"de esta expresión).

Resolución por sustitución

$$\underbrace{\frac{N(N+1)}{2}}_{\text{multiplicaciones/divisiones}} + \underbrace{\frac{N(N-1)}{2}}_{\text{sumas}} = N^2 \text{ operaciones}$$

¡Regla de Cramer $(N+1)!N-1$ operaciones!

II.1.2. Métodos de Gauss y Gauss–Jordan. Pivotaje

Método de Gauss

$\mathbf{Ax} = \mathbf{b}$ unisolvente $\rightsquigarrow$ $\mathbf{Ux} = \mathbf{c}$ equivalente, $\mathbf{U}$ triangular superior

$\mathbf{A} \in \mathbb{R}^{N\times N}$ regular y $a_{11} \neq 0$
encontrar un sistema equivalente en el que x_1 no aparezca en la ecuación i-ésima $(i = 2, \ldots, N)$

restar

$$\frac{a_{i1}}{a_{11}}\left(a_{11}x_1 + \cdots + a_{1N}x_N = b_1\right)$$

de la ecuación i-ésima $(i = 2, \ldots, N)$

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1N} \\ 0 & \tilde{a}_{22} & \cdots & \tilde{a}_{1N} \\ \vdots & \vdots & & \vdots \\ 0 & \tilde{a}_{N2} & \cdots & \tilde{a}_{NN} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_N \end{bmatrix} = \begin{bmatrix} b_1 \\ \tilde{b}_2 \\ \vdots \\ \tilde{b}_N \end{bmatrix}$$

$\tilde{a}_{22} \neq 0 \rightsquigarrow$ eliminar x_2 de las ecuaciones inferiores
condición de no nulidad en las sucesivas entradas $(i, i) \rightsquigarrow$ sistema triangular superior equivalente al de partida

Método de Gauss

- Datos: $N \geq 1$, $\mathbf{A} \in \mathbb{R}^{N \times N}$, $\mathbf{b} \in \mathbf{R}^N$
- $\mathbf{A}^{(1)} := \mathbf{A}$
- Suponemos $k = 1, \ldots, N \;\Rightarrow\; a_{kk}^{(k)} \neq 0$ (en caso contrario hemos terminado y no es posible llegar a un sistema triangular equivalente), definimos recursivamente los *multiplicadores*

No se puede intercambiar

$$i = k+1, \ldots, N \;\Rightarrow\; m_{ik} := \frac{a_{ik}^{(k)}}{a_{kk}^{(k)}}$$

filas y multiplicar por escalar

e

$$i = k+1, \ldots, N,\; j = k, \ldots, N \;\Rightarrow\; a_{ij}^{(k+1)} := a_{ij}^{(k)} - m_{ik} a_{kj}^{(k)}$$

$$i = k+1, \ldots, N \;\Rightarrow\; b_i^{(k+1)} := b_i^{(k)} - m_{ik} b_k^{(k)}$$

- Sistema triangular superior equivalente

$$\mathbf{Ux} = \mathbf{c}, \quad \mathbf{U} := \mathbf{A}^{(N)}, \quad \mathbf{c} := \mathbf{b}^{(N)},$$

se resuelve por sustitución hacia atrás

$k = 1 \ldots, N$

$$\mathbf{A}^{(k)} = \begin{bmatrix} a_{11}^{(1)} & a_{12}^{(1)} & \cdots & \cdots & \cdots & \cdots & a_{1N}^{(1)} \\ 0 & a_{22}^{(2)} & \cdots & \cdots & \cdots & \cdots & a_{2N}^{(2)} \\ \vdots & & \ddots & & & & \vdots \\ 0 & \cdots & \cdots & 0 & a_{kk}^{(k)} & \cdots & a_{kN}^{(k)} \\ \vdots & & & \vdots & \vdots & & \vdots \\ 0 & \cdots & \cdots & 0 & a_{Nk}^{(k)} & \cdots & a_{NN}^{(k)} \end{bmatrix}$$

Ejemplo

$$\mathbf{A} = \begin{bmatrix} 1 & 1 & 3 \\ 0.1 & 1 & 1 \\ 1 & 2 & 0 \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} 5 \\ 2.1 \\ 3 \end{bmatrix}$$

$$\mathbf{A}^{(1)} = \mathbf{A}, \quad \mathbf{b}^{(1)} = \mathbf{b}$$

$$\mathbf{A}^{(2)} = \begin{bmatrix} 1 & 1 & 3 \\ 0 & 0.9 & 0.7 \\ 0 & 1 & -3 \end{bmatrix}, \quad \mathbf{b}^{(2)} = \begin{bmatrix} 5 \\ 1.6 \\ -2 \end{bmatrix}$$

$$\mathbf{A}^{(3)} = \begin{bmatrix} 1 & 1 & 3 \\ 0 & 0.9 & 0.7 \\ 0 & 0 & -3.\overline{7} \end{bmatrix}, \quad \mathbf{b}^{(3)} = \begin{bmatrix} 5 \\ 1.6 \\ -3.\overline{7} \end{bmatrix}$$

Sustitución hacia atrás

$$x_1 = x_2 = x_3 = 1$$

Solo 1 Transformación que no cambia el determinante

Método de Gauss hasta paso $N \rightsquigarrow \mathbf{A}$ regular

Proposición

Sea $\mathbf{A} \in \mathbb{R}^{N\times N}$ una matriz cuadrada y sea $\mathbf{b} \in \mathbf{R}^N$. Entonces son equivalentes:

(i) El correspondiente método de Gauss puede completarse hasta el paso N-ésimo.

(ii) Para cada $k = 1, \ldots, N$ la k-ésima submatriz principal de $\mathbf{A}$

$$\mathbf{A}_k = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1k} \\ a_{21} & a_{22} & \cdots & a_{1k} \\ \vdots & \vdots & & \vdots \\ a_{k1} & a_{k2} & \cdots & a_{kk} \end{bmatrix}$$

es regular.

Los menores de orden k.

DEMOSTRACIÓN.

$\boxed{\text{(i)} \Rightarrow \text{(ii)}}$ $k = 1, \ldots, N$

$$\det(\mathbf{A}_k) = \det \begin{bmatrix} a_{11}^{(1)} & a_{12}^{(1)} & \cdots & a_{1k}^{(1)} \\ 0 & a_{22}^{(2)} & \cdots & a_{2k}^{(2)} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_{kk}^{(k)} \end{bmatrix}$$

$$\Downarrow$$

$$\det(\mathbf{A}_k) = a_{11}^{(1)} \cdots a_{kk}^{(k)}$$

$$\Downarrow$$

$$\mathbf{A}_k \text{ regular}$$

$\boxed{\text{(ii)} \Rightarrow \text{(i)}}$ Supongamos que no podemos completar el método de Gauss

$$k := \text{mín}\{l \in \{1, \ldots, N\} : \ a_{ll}^{(l)} = 0\}$$

$\mathbf{A}_1$ regular $\Rightarrow k \geq 2$

Método de Gauss hasta $\mathbf{A}^{(k)}$

$$\mathbf{A}^{(k)} = \begin{bmatrix} a_{11}^{(1)} & a_{12}^{(1)} & \cdots & a_{1k}^{(1)} & a_{1k+1}^{(1)} & \cdots & a_{1N}^{(1)} \\ 0 & a_{22}^{(2)} & \cdots & a_{2k}^{(2)} & a_{2k+1}^{(2)} & \cdots & a_{2N}^{(2)} \\ \vdots & \vdots & & \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & 0 & a_{kk+1}^{(k)} & \cdots & a_{kN}^{(k)} \\ 0 & 0 & \cdots & a_{k+1k}^{(k)} & a_{k+1k+1}^{(k)} & \cdots & a_{k+1N}^{(k)} \\ \vdots & \vdots & & \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & a_{Nk}^{(k)} & a_{Nk+1}^{(k)} & \cdots & a_{NN}^{(k)} \end{bmatrix}.$$

Ambas son equivalentes

$$\det(\mathbf{A}_k) = \det(\mathbf{A}^{(k)})_k = 0$$

Me quedo con la submatriz

□

Número de operaciones artiméticas para resolver un sistema mediante el método de Gauss

Ejercicio

Si N es un número natural, entonces

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

(Indicación: o bien se procede por inducción, o bien de forma constructiva, observando la tabla

$$\begin{array}{ccccccccccc}
1^3 & = & (1+0)^3 & = & 1 & + & 0 & + & 0 & + & 0^3 \\
2^3 & = & (1+1)^3 & = & 1 & + & 3\cdot 1 & + & 3\cdot 1^2 & + & 1^3 \\
3^3 & = & (1+2)^3 & = & 1 & + & 3\cdot 2 & + & 3\cdot 2^2 & + & 2^3 \\
\vdots & & \vdots & & \vdots & & & & & & \vdots \\
(N+1)^3 & = & (1+N)^3 & = & 1 & + & 3\cdot N & + & 3\cdot N^2 & + & N^3
\end{array}$$

y llamando

$$\beta = \sum_{j=1}^{N} j^2,$$

deducimos, sumando las "columnas" de la tabla (recuérdese la expresión de la suma de los primeros N números naturales) que

$$(N+1)^3 = (N+1) + \frac{3N(N+1)}{2} + 3\beta,$$

luego

$$\beta = \frac{N(N+1)(2N+1)}{6},$$

como se ha afirmado).

Comprobar

- Paso k:
  - multiplicadores m_{ik}

$$N - k \text{ divisiones}$$

  - coeficientes $\mathbf{A}^{(k+1)}$

$$(N-k)^2 \text{ productos y } (N-k)^2 \text{ sumas}$$

  - coeficientes $\mathbf{b}^{(k+1)}$

$$N - k \text{ productos y } N - k \text{ sumas}$$

- Resolución por sustitución hacia atrás del sistema triangular superior

$$N^2 \text{ operaciones}$$

$$2\sum_{k=1}^{N-1}(N-k)^2 + 3\sum_{k=1}^{N-1}(N-k) + N^2 = \frac{4N^3 + 9N^2 - 7N}{6} \text{ operaciones}$$

¡Regla de Cramer $(N+1)!N - 1$ operaciones!

(i) ¿Cómo podemos reducir los errores de redondeo que afectan al método de Gauss?

(ii) ¿Puede modificarse el método de Gauss de forma que se evite el problema generado cuando $a_{kk}^{(k)} = 0$?

Respuesta simultánea

Efecto errores de redondeo en el algoritmo de Gauss

19/03/20

Ejemplo

→ es casi 0

$$\begin{bmatrix} 10^{-5} & -1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} -1 \\ 3 \end{bmatrix}$$

Solución

$$x_1 = \frac{200000}{100001}, \quad x_2 = \frac{100003}{100001}$$

sustitución hacia atrás sistema triangular método de Gauss (único paso)

$$\begin{bmatrix} 10^{-5} & -1 \\ 0 & 1+10^5 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} -1 \\ 3+10^5 \end{bmatrix}$$

Ordenador $\mathbb{F}(10, 5, -4, 6)$ y redondeo

coeficientes sistema triangular sin overflow (o underflow)

$$[b^{L-1}, b^U(1-b^{-t})] = [10^{-5}, 99999 \cdot 10]$$

redondeos coeficientes del sistema triangular $\rightsquigarrow$ errores considerables

$$\mathrm{rd}(10^{-5}) = (0.1)\cdot 10^{-4}$$

$$\mathrm{rd}(-1) = -(0.1)\cdot 10$$

$$\mathrm{rd}(1+10^{5}) = \mathrm{rd}(3+10^{5}) = (0.1)\cdot 10^{6}$$

sistema triangular en ordenador

$$\begin{bmatrix} (0.1)\cdot 10^{-4} & -(0.1)\cdot 10 \\ 0 & (0.1)\cdot 10^{6} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} -(0.1)\cdot 10 \\ (0.1)\cdot 10^{6} \end{bmatrix}$$

solución por sustitución hacia atrás

$$¡¡¡\ x_1 = 0, \quad x_2 = 1\ !!!$$

Evitar
- dividir por coeficientes relativamente pequeños ⇝ se generan errores de redondeo grandes
- algún $a_{kk}^{(k)}$ del método de Gauss sea nulo

hay un condicional dentro

Método de Gauss con pivotaje (o *pivotaje parcial*), variante *adaptativa* de Gauss

↳ el total tiene más coste de operac

Paso k método de Gauss modificado: antes de definir los multiplicadores m_{ik}, la matriz $\mathbf{A}^{(k+1)}$ y el vector $\mathbf{b}^{(k+1)}$, intercambiar de posición si es necesario dos de las filas $k, \ldots, N$ de la matriz $\mathbf{A}^{(k)}$ de forma que el elemento del vector

Se aplica a A y a b.

$$
\begin{bmatrix} a_{kk}^{(k)} \\ \vdots \\ a_{Nk}^{(k)} \end{bmatrix} \rightsquigarrow \begin{bmatrix} a_{11}^{(1)} & a_{12}^{(1)} & \cdots & \cdots & \cdots & \cdots & a_{1N}^{(1)} \\ 0 & a_{22}^{(2)} & \cdots & \cdots & \cdots & \cdots & a_{2N}^{(2)} \\ \vdots & & \ddots & & & & \vdots \\ 0 & \cdots & \cdots & 0 & a_{kk}^{(k)} & \cdots & a_{kN}^{(k)} \\ \vdots & & & \vdots & \vdots & & \vdots \\ 0 & \cdots & \cdots & 0 & a_{Nk}^{(k)} & \cdots & a_{NN}^{(k)} \end{bmatrix}
$$

que tiene mayor valor absoluto sea $a_{kk}^{(k)}$

↳ evitamos errores

Sistema equivalente al de partida que se resuelve por sustitución hacia atrás

↳ Metemos en este método el intercambio de filas

Si la matriz es regular se puede usar pivotaje, lo hay que tener en cuenta submatrices.

Observación

A diferencia de su versión básica, el método de Gauss con pivotaje siempre tiene sentido hasta el último paso N-ésimo si, y sólo si, la matriz de coeficientes $\mathbf{A}$ es regular, es decir, el sistema en cuestión es compatible determinado.

Ejemplo

$$\begin{bmatrix} 10^{-5} & -1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} -1 \\ 3 \end{bmatrix}$$, sistema ejemplo anterior, $x_1 = \dfrac{200000}{100001}$, $x_2 = \dfrac{100003}{100001}$

Método de Gauss con pivotaje

→ Hemos hecho el cambio.

$$\begin{bmatrix} 1 & 1 \\ 10^{-5} & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 3 \\ -1 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 1 \\ 0 & -(1+10^{-5}) \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 3 \\ -(1+3\cdot 10^{-5}) \end{bmatrix}$$

Números máquina ($\mathbb{F}(10,5,-4,6)$)

$$\begin{bmatrix} (0.1)\cdot 10 & (0.1)\cdot 10 \\ 0 & -(0.1)\cdot 10 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} (0.3)\cdot 10 \\ -(0.1)\cdot 10 \end{bmatrix}$$

solución sustitución hacia atrás (buena)

→ Mejor solución

$$x_1 = 2, \quad x_2 = 1$$

Ejercicio

Diseña e implementa el algoritmo del método de Gauss con pivotaje.

Observación

Otra variante adaptativa del método de Gauss, *pivotaje total o completo*, no solo reordena las filas $k, \ldots, N$ de $\mathbf{A}^{(k)}$, sino también las columnas $k, \ldots, N$ de forma que el elemento de la submatriz de dicha matriz correspondiente a las mencionadas filas y columnas tenga a $a_{kk}^{(k)}$ como el elemento de mayor valor absoluto. Sin embargo, requiere un mayor número de operaciones.

Observación

Para terminar con las variantes del método de Gauss, mencionemos el llamado *método de Gauss–Jordan*, que consiste en hacer ceros no solo debajo de $a_{kk}^{(k)}$ sino también por encima, con el mismo tipo de fórmula. Sin embargo, su coste en operaciones aritméticas es superior.

II.1.3. Métodos de factorización

Sistemas *con la misma matriz de coeficientes* $\rightsquigarrow$ análisis de estructuras

Método de Gauss $\hookrightarrow$ factorización LU → más eficientes

Idea: $\mathbf{Ax} = \mathbf{b}$ compatible determinado

$\mathbf{L}$ triangular inferior, $\mathbf{U}$ triangular superior

→ en Gauss esta es la identidad

$$\mathbf{A} = \mathbf{LU}$$

→ este sistema es más fácil

$$\mathbf{LUx} = \mathbf{b}$$

resolución de dos sistemas triangulares auxiliares

(i) $\mathbf{y} := \mathbf{Ux}$

$$\mathbf{Ly} = \mathbf{b}$$

→ vector de incógnitas

sustitución hacia adelante

(ii)

$$\mathbf{Ux} = \mathbf{y}$$

sustitución hacia atrás

de aquí saco x

Ejemplo

$$\mathbf{b} = \begin{bmatrix} -1 \\ 2 \\ 0 \end{bmatrix}$$

$\mathbf{A} = \mathbf{LU}$

$$\mathbf{A} = \begin{bmatrix} 1 & 3 & -1 \\ 2 & 8 & 4 \\ -1 & 3 & 4 \end{bmatrix}, \quad \mathbf{L} = \begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ -1 & 3 & 1 \end{bmatrix}, \quad \mathbf{U} = \begin{bmatrix} 1 & 3 & -1 \\ 0 & 2 & 6 \\ 0 & 0 & -15 \end{bmatrix}$$

(i) $\mathbf{Ly} = \mathbf{b}$

$$\begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ -1 & 3 & 1 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = \begin{bmatrix} -1 \\ 2 \\ 0 \end{bmatrix}$$

sustitución hacia adelante

$$y_1 = -1 \rightsquigarrow y_2 = 2 - 2y_1 = 4 \rightsquigarrow y_3 = 0 + y_1 - 3y_2 = -13$$

(ii) $\mathbf{Ux} = \mathbf{y}$

$$\begin{bmatrix} 1 & 3 & -1 \\ 0 & 2 & 6 \\ 0 & 0 & -15 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} -1 \\ 4 \\ -13 \end{bmatrix}$$

sustitución hacia atrás

$$x_3 = \frac{13}{15} \rightsquigarrow x_2 = \frac{1}{2}(4 - 6x_3) = \frac{1}{2}(4 - \frac{26}{5}) = -\frac{3}{5}$$

$$\rightsquigarrow x_1 = -1 - 3x_2 + x_3 = \frac{5}{3}$$

solución del sistema inicial $\mathbf{Ax} = \mathbf{b}$

20/03/20

- ¿Cuándo es posible obtener una factorización tipo LU para una matriz regular? → En general no
- ¿Algoritmo?
- ¿Unicidad?

No siempre

$$\left.\begin{array}{r}\mathbf{L}, \mathbf{U} \text{ regulares} \\ \mathbf{A} = \mathbf{LU}\end{array}\right| \Rightarrow a_{11} = l_{11} u_{11} \neq 0$$

por que son Triangulares

Una primera respuesta (parcial): método de Gauss

Proposición

Sean $N \geq 1$, $\mathbf{A} \in \mathbb{R}^{N\times N}$, $\mathbf{b} \in \mathbf{R}^N$ y supongamos que aplicando el método de Gauss al sistema $\mathbf{Ax} = \mathbf{b}$ se obtiene una matriz triangular superior $\mathbf{A}^{(N)}$ y un vector $\mathbf{b}^{(N)}$ de forma que el sistema $\mathbf{A}^{(N)}\mathbf{x} = \mathbf{b}^{(N)}$ es equivalente al de partida. Entonces

$$\mathbf{A} = \mathbf{LU},$$

A nivel práctico:

siendo

$$\mathbf{U} = \mathbf{A}^{(N)}$$

y

Práctico → no sirve
Teórico → cuándo puedo factorizar

$$\mathbf{L} = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ m_{21} & 1 & 0 & \cdots & 0 \\ m_{31} & m_{32} & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ m_{N1} & m_{N2} & m_{N3} & \cdots & 1 \end{bmatrix},$$

donde los coeficientes de la parte inferior de $\mathbf{L}$ son los multiplicadores del método de Gauss definidos recursivamente.

DEMOSTRACIÓN. $k = 1, \ldots, N-1$, notación método de Gauss

no entra

$$\mathbf{A}^{(k+1)} = \mathbf{E}_k \mathbf{A}^{(k)}$$

modifico la identidad

$$\mathbf{E}_k := \begin{bmatrix} 1 & 0 & \cdots & \cdots & \cdots & \cdots & 0 \\ 0 & 1 & \cdots & \cdots & \cdots & \cdots & 0 \\ \vdots & & \ddots & & & & \vdots \\ 0 & \cdots & \cdots & 1 & 0 & \cdots & 0 \\ 0 & \cdots & \cdots & -m_{k+1\ k} & 1 & \cdots & 0 \\ \vdots & & & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & \cdots & -m_{Nk} & 0 & \cdots & 1 \end{bmatrix}.$$

$$\left.\begin{array}{l} \mathbf{e}_k^T := [0, 0, \ldots, 0, \underbrace{1}_{k}, 0, \ldots, 0] \\ \mathbf{m}_k^T := [0, \ldots, 0, m_{k+1\ k}, \ldots, m_{Nk}] \\ \mathbf{I}_N \text{ matriz identidad de orden } N \end{array}\right| \Rightarrow \mathbf{E}_k = \mathbf{I}_N - \mathbf{m}_k \mathbf{e}_k^T$$

$\mathbf{E}_k$ regular (vid. ejercicio)

$$\mathbf{E}_k^{-1} = \mathbf{I}_N + \mathbf{m}_k \mathbf{e}_k^T$$

$\mathbf{A}^{(k+1)} = \mathbf{E}_k \mathbf{A}^{(k)} \wedge$ método de Gauss hasta paso N

$$\mathbf{E}_{N-1}\mathbf{E}_{N-2}\cdots\mathbf{E}_2\mathbf{E}_1\mathbf{A} = \mathbf{U}$$

$$\Downarrow$$

$$\begin{aligned}
\mathbf{A} &= \mathbf{E}_1^{-1}\mathbf{E}_2^{-1}\cdots\mathbf{E}_{N-2}^{-1}\mathbf{E}_{N-1}^{-1}\mathbf{U} \\
&= \left(\mathbf{I}_N + \mathbf{m}_1\mathbf{e}_1^T\right)\left(\mathbf{I}_N + \mathbf{m}_2\mathbf{e}_2^T\right)\cdots\left(\mathbf{I}_N + \mathbf{m}_{N-1}\mathbf{e}_{N-2}^T\right)\left(\mathbf{I}_N + \mathbf{m}_{N-1}\mathbf{e}_{N-1}^T\right)\mathbf{U}
\end{aligned}$$

$$
\begin{aligned}
\mathbf{A} &= \mathbf{E}_1^{-1}\mathbf{E}_2^{-1}\cdots\mathbf{E}_{N-2}^{-1}\mathbf{E}_{N-1}^{-1}\mathbf{U} \\
&= \left(\mathbf{I}_N + \mathbf{m}_1\mathbf{e}_1^T\right)\left(\mathbf{I}_N + \mathbf{m}_2\mathbf{e}_2^T\right)\cdots\left(\mathbf{I}_N + \mathbf{m}_{N-2}\mathbf{e}_{N-2}^T\right)\left(\mathbf{I}_N + \mathbf{m}_{N-1}\mathbf{e}_{N-1}^T\right)\mathbf{U} \\
&= \left(\mathbf{I}_N + \sum_{k=1}^{N-1}\mathbf{m}_k\mathbf{e}_k^T\right)\mathbf{U} \qquad \text{(vid. ejercicio)} \\
&= \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ m_{21} & 1 & 0 & \cdots & 0 \\ m_{31} & m_{32} & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ m_{N1} & m_{N2} & m_{N3} & \cdots & 1 \end{bmatrix}\mathbf{U}
\end{aligned}
$$

□

Ejercicio

Comprueba que cada matriz $\mathbf{E}_k$ de la demostración de la proposición anterior es regular, y de hecho

$$\mathbf{E}_k^{-1} = \mathbf{I}_N + \mathbf{m}_k \mathbf{e}_k^T,$$

y, de forma más general, que si $k \leq l$, entonces

$$\mathbf{m}_k \mathbf{e}_k^T \mathbf{m}_l \mathbf{e}_l^T = \mathbf{0}.$$

Demostración $\rightsquigarrow$ cálculo $\mathbf{L}$ (¡bajo hipótesis proposición!)

(i) Hallar los multiplicadores y construir la matriz directamente

$$\mathbf{L} = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ m_{21} & 1 & 0 & \cdots & 0 \\ m_{31} & m_{32} & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ m_{N1} & m_{N2} & m_{N3} & \cdots & 1 \end{bmatrix}$$

(ii) $[A|I] \rightsquigarrow^{Gauss} \rightsquigarrow [U|L^{-1}]$

Caracterización de la existencia de una factorización LU

Proposición

Sea $\mathbf{A} \in \mathbb{R}^{N\times N}$ una matriz regular. Entonces equivalen

(i) $\mathbf{A}$ admite una factorización LU. = Método Gauss

(ii) Las N submatrices principales de $\mathbf{A}$ son regulares.

DEMOSTRACIÓN.

$\boxed{\text{(i)} \Rightarrow \text{(ii)}}$ $\mathbf{L}, \mathbf{U} \in \mathbb{R}^{N\times N}$ triangulares inferior y superior

$$\mathbf{A} = \mathbf{LU}$$

$\mathbf{A}$ regular $\Rightarrow$ $\mathbf{L}, \mathbf{U}$ regulares

$k = 1, \ldots, N$, k-ésima submatriz principal de $\mathbf{A}$

Todos van a ser regulares

$$\mathbf{A}_k = \begin{bmatrix} l_{11} & 0 & \ldots & 0 \\ l_{21} & l_{22} & \ldots & 0 \\ \vdots & \vdots & & \vdots \\ l_{k1} & l_{k2} & \ldots & l_{kk} \end{bmatrix} \begin{bmatrix} u_{11} & u_{12} & \ldots & u_{1k} \\ 0 & u_{22} & \ldots & u_{2k} \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \ldots & u_{kk} \end{bmatrix}$$

$$\Downarrow$$

$$\det(\mathbf{A}_k) = l_{11} \cdots l_{kk} u_{11} \cdots u_{kk} \neq 0$$

$\boxed{\text{(ii)} \Rightarrow \text{(i)}}$ Dos últimas proposiciones $\square$

Se puede aplicar método de Gauss entonces admite una factorización LU.

Resumen

Teorema

Sea $\mathbf{A} \in \mathbb{R}^{N \times N}$ una matriz regular. Entonces, las siguientes afirmaciones son equivalentes:

(i) Para cualquier $\mathbf{b} \in \mathbb{R}^{N}$, el método de Gauss para el correspondiente sistema de ecuaciones lineales puede completarse hasta el paso N-ésimo.

(ii) $\mathbf{A}$ admite una factorización LU.

(iii) Todas las submatrices principales de $\mathbf{A}$ son regulares.

Proposición → no entra

Si $\mathbf{A} \in \mathbb{R}^{N \times N}$ es una matriz regular entonces hay una matriz $\tilde{\mathbf{A}}$ que se obtiene permutando eventualmente algunas de las filas de $\mathbf{A}$ y que admite una factorización LU.

DEMOSTRACIÓN.

$\mathbf{A}$ regular $\rightsquigarrow$ Gauss con pivotaje sistema con $\mathbf{A}$ matriz coeficientes

$\mathbf{A}^{(1)} = \mathbf{A}$

intercambiar de posición 2 filas $\rightsquigarrow$ posición 1 de columna 1 coeficiente con mayor valor absoluto

matricialmente
$\mathbf{P}_1$ identidad de orden N permutando las mismas filas que en $\mathbf{A}^{(1)}$

$$\mathbf{P}_1\mathbf{A}^{(1)}$$

$\mathbf{A}^{(2)}$ transformaciones correspondientes $\leftrightsquigarrow$ multiplicar por una matriz $\mathbf{E}_1$ (demostración proposición anterior)

$$\mathbf{A}^{(2)} = \mathbf{E}_1\mathbf{P}_1\mathbf{A}^{(1)}$$

Ídem

$$\mathbf{A}^{(3)} = \mathbf{E}_2\mathbf{P}_2\mathbf{A}^{(2)} = \mathbf{E}_2\mathbf{P}_2\mathbf{E}_1\mathbf{P}_1\mathbf{A}^{(1)}$$

sucesivamente $\rightsquigarrow$ matriz triangular superior $\mathbf{U}$

$$\mathbf{U} = \mathbf{A}^{(N)} = \mathbf{E}_{N-1}\mathbf{P}_{N-1}\cdots\mathbf{E}_1\mathbf{P}_1\mathbf{A}$$

$$\mathbf{E} := \mathbf{E}_{N-1}\mathbf{P}_{N-1}\cdots\mathbf{E}_1\mathbf{P}_1$$

$$\mathbf{P} := \mathbf{P}_{N-1}\cdots\mathbf{P}_1$$

$$\mathbf{U} = \mathbf{E}\mathbf{A}$$

$$\Updownarrow$$

$$\mathbf{U} = (\mathbf{E}\mathbf{P}^{-1})\mathbf{P}\mathbf{A}.$$

$$\mathbf{L} := \mathbf{P}\mathbf{E}^{-1} \text{ triangular inferior}$$

(basta usar la expresión de $\mathbf{E}_i^{-1}$, comprobar que $\mathbf{P}_i^{-1} = \mathbf{P}_i$ y que dicha matriz actúa únicamente sobre filas comprendidas entre la i-ésima y la N-ésima)

$\square$

Demostración constructiva

$\mathbf{P}$ matriz que se obtiene al aplicar a la identidad de orden N las mismas permutaciones de filas que a $\mathbf{A}$

$$\mathbf{Ax} = \mathbf{b}$$

$$\Updownarrow$$

$$\mathbf{PAx} = \mathbf{Pb}$$

$\mathbf{PA} = \mathbf{LU}$

$$\mathbf{Ly} = \mathbf{Pb}$$

$$\mathbf{Ux} = \mathbf{y}$$

Práctica 2

Resumen

¿No equivalen?

Teorema

Sea **A** una matriz regular. Entonces:

(i) El método de Gauss con pivotaje es factible, para cualquier sistema de ecuaciones lineales que tenga a **A** por matriz de coeficientes.

(ii) Salvo la eventual permutación de algunas de sus filas, **A** admite una factorización LU.

25/03/20
Covid-19

Algortimos, unicidad factorización LU

$\mathbf{A} \in \mathbb{R}^{N\times N}$ regular

$$\mathbf{A} = \mathbf{LU}$$ Las busco

$\mathbf{L}, \mathbf{U} \in \mathbb{R}^{N\times N}$ triangulares inferior y superior

$$i,j = 1,\ldots,N \;\Rightarrow\; a_{ij} = \sum_{k=1}^{i\wedge j} l_{ik}u_{kj} \tag{1}$$

mínimo de los dos
no cojo los que dan ceros

determinar los coeficientes incógnita de las matrices triangulares a partir de los conocidos de $\mathbf{A}$

$$\frac{N(N+1)}{2} + \frac{N(N+1)}{2} = N(N+1)$$

incógnitas, N^2 datos $\rightsquigarrow$ fijar N incógnitas

Tengo N incógnitas que fijar ~~datos~~ en la primera fila hay uno, en la 2 hay 2. lo conoces, al revés

Tiene que ser regular

- Factorización de *Doolittle*

$$l_{11} = \cdots = l_{NN} = 1$$

- Factorización de *Crout*

$$u_{11} = \cdots = u_{NN} = 1$$

Gauss $\rightsquigarrow$ Doolittle

$$\mathbf{U} = \mathbf{A}^{(N)}$$

$$\mathbf{L} = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ m_{21} & 1 & 0 & \cdots & 0 \\ m_{31} & m_{32} & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ m_{N1} & m_{N2} & m_{N3} & \cdots & 1 \end{bmatrix},$$

o

$$\left[A|I\right] \rightsquigarrow^{Gauss} \rightsquigarrow \left[U|L^{-1}\right]$$

Identificación (1) $\rightsquigarrow$ procedimiento *heurístico* para Doolittle o Crout

Antes de formalizar

Ejemplo

$$\mathbf{A} = \begin{bmatrix} 1 & -2 & 0 & 3 \\ -2 & 3 & 1 & -6 \\ -1 & 4 & -4 & 3 \\ 5 & -8 & 4 & 0 \end{bmatrix}$$

$$\begin{bmatrix} l_{11} & 0 & 0 & 0 \\ l_{21} & l_{22} & 0 & 0 \\ l_{31} & l_{32} & l_{33} & 0 \\ l_{41} & l_{42} & l_{43} & l_{44} \end{bmatrix} \begin{bmatrix} 1 & u_{12} & u_{13} & u_{14} \\ 0 & 1 & u_{23} & u_{24} \\ 0 & 0 & 1 & u_{34} \\ 0 & 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & -2 & 0 & 3 \\ -2 & 3 & 1 & -6 \\ -1 & 4 & -4 & 3 \\ 5 & -8 & 4 & 0 \end{bmatrix}$$

- fila 1 de A

$$l_{11} = 1$$

$$u_{12} = -2$$

$$u_{13} = 0$$

$$u_{14} = 3$$

- fila 2 de A

$$l_{21} = -2$$

$$l_{21}u_{12} + l_{22} = 3 \Leftrightarrow l_{22} = -1$$

$$l_{21}u_{13} + l_{22}u_{23} = 1 \Leftrightarrow u_{23} = -1$$

$$l_{21}u_{14} + l_{22}u_{24} = -6 \Leftrightarrow u_{24} = 0$$

- fila 3 de A

$$l_{31} = -1$$

$$l_{31}u_{12} + l_{32} = 4 \Leftrightarrow l_{32} = 2$$

$$l_{31}u_{13} + l_{32}u_{23} + l_{33} = -4 \Leftrightarrow l_{33} = -2$$

$$l_{31}u_{14} + l_{32}u_{24} + l_{33}u_{34} = 3 \Leftrightarrow u_{34} = -3$$

- fila 4 de A

$$l_{41} = 5$$

$$l_{41}u_{12} + l_{42} = -8 \Leftrightarrow l_{42} = 2$$

$$l_{41}u_{13} + l_{42}u_{23} + l_{43} = 4 \Leftrightarrow l_{43} = 6$$

$$l_{41}u_{14} + l_{42}u_{24} + l_{43}u_{34} + l_{44} = 0 \Leftrightarrow l_{44} = 3$$

$$\mathbf{L} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ -2 & -1 & 0 & 0 \\ -1 & 2 & -2 & 0 \\ 5 & 2 & 6 & 3 \end{bmatrix}, \quad \mathbf{U} = \begin{bmatrix} 1 & -2 & 0 & 3 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 1 & -3 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

Algoritmo factorización de Doolittle

Descripción elementos

Ejemplo anterior: fila de **A** $\leftrightsquigarrow$ correspondiente fila de **U** y **L**

Descripción de elementos más cómoda para programar: fila de **A** $\leftrightsquigarrow$ misma fila de **U** y la correspondiente columna de **L**

Identificación (1), $l_{11} = \cdots = l_{NN} = 1$

$i \leq j$, (1)

$$u_{ij} = a_{ij} - \sum_{k=1}^{i-1} l_{ik} u_{kj}$$

$$\sum_{k=1}^{0} \cdots = 0$$

$j \leq i$, (1)

$$l_{ij} = \frac{1}{u_{jj}} \left(a_{ij} - \sum_{k=1}^{j-1} l_{ik} u_{kj} \right)$$

Alternando ambas expresiones $\rightsquigarrow$ algoritmo

Directamente: en la última expresión intercambiamos los índices

Factorización tipo Doolittle

- Datos: $N \geq 1$, $\mathbf{A} \in \mathbb{R}^{N \times N}$ regular
- $l_{11} = \cdots = l_{NN} = 1$
- $i = 1, \ldots, N$

$$j = i, \ldots, N \;\Rightarrow\; u_{ij} = a_{ij} - \sum_{k=1}^{i-1} l_{ik} u_{kj}$$

y supuesto que $u_{ii} \neq 0$ → esto no dice en lo admite

$$j = i+1, \ldots, N \;\Rightarrow\; l_{ji} = \frac{1}{u_{ii}} \left(a_{ji} - \sum_{k=1}^{i-1} l_{jk} u_{ki} \right)$$

26/03/20

Ejercicio

Comprueba aplicando el algoritmo anterior que la matriz regular

$$\begin{bmatrix} 0.1 & 0.2 & 0.3 & 0.4 \\ 0.2 & 0.9 & 1.2 & 1.5 \\ 0.3 & 1.6 & 2.9 & 3.5 \\ 0.4 & 2.3 & 4.6 & 6.5 \end{bmatrix}$$

admite una factorización tipo Doolittle.

Basta comprobar que el algoritmo da la factorización de la matriz anterior

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 2 & 1 & 0 & 0 \\ 3 & 2 & 1 & 0 \\ 4 & 3 & 2 & 1 \end{bmatrix} \begin{bmatrix} 0.1 & 0.2 & 0.3 & 0.4 \\ 0 & 0.5 & 0.6 & 0.7 \\ 0 & 0 & 0.8 & 0.9 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

Ejercicio

Diseña un algoritmo para determinar la factorización tipo Crout de una matriz regular que la admita.
(Indicación: se puede aprovechar el algoritmo de Doolittle: si $\mathbf{A}^T = \mathbf{LU}$, entonces $\mathbf{A} = \mathbf{U}^T\mathbf{L}^T$).

Unicidad fijando unos en una de las diagonales

Para aprovechar el algoritmo de Doolittle basta aplicarlo a A^T. De esa forma, al trasponer obtenemos la factorización de Crout de A.

Ejercicio

Considera los sistemas de ecuaciones lineales con matriz de coeficientes

$$\begin{bmatrix} 1 & 0 & 1 & 0 \\ 2 & 1 & 3 & 1 \\ 0 & 1 & 3 & 3 \\ 1 & 1 & 4 & 2 \end{bmatrix}$$

y términos independientes

$$\begin{bmatrix} 1 \\ 3 \\ 3 \\ 3 \end{bmatrix}, \quad \begin{bmatrix} 1 \\ 4 \\ 6 \\ 5 \end{bmatrix} \quad \text{y} \quad \begin{bmatrix} 1 \\ 4 \\ 4 \\ 5 \end{bmatrix}.$$

Resuélvelos por el método más eficiente.

Resolución: el método más eficiente es obtener la factorización LU de A.

$$A = \begin{pmatrix} 1 & 0 & 1 & 0 \\ 2 & 1 & 3 & 1 \\ 0 & 1 & 3 & 3 \\ 1 & 1 & 4 & 2 \end{pmatrix} \quad e_1 = \begin{pmatrix} 1 \\ 3 \\ 3 \\ 3 \end{pmatrix} \quad e_2 = \begin{pmatrix} 1 \\ 4 \\ 6 \\ 5 \end{pmatrix} \quad e_3 = \begin{pmatrix} 1 \\ 4 \\ 4 \\ 5 \end{pmatrix}$$

Consideramos $A = LU$:

$$\begin{pmatrix} 1 & 0 & 1 & 0 \\ 2 & 1 & 3 & 1 \\ 0 & 1 & 3 & 3 \\ 1 & 1 & 4 & 2 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ l_{21} & 1 & 0 & 0 \\ l_{31} & l_{32} & 1 & 0 \\ l_{41} & l_{42} & l_{43} & 1 \end{pmatrix} \begin{pmatrix} u_{11} & u_{12} & u_{13} & u_{14} \\ 0 & u_{22} & u_{23} & u_{24} \\ 0 & 0 & u_{33} & u_{34} \\ 0 & 0 & 0 & u_{44} \end{pmatrix}$$

$1 = u_{11}$ $\quad 0 = u_{12}$ $\quad 1 = u_{13}$ $\quad 0 = u_{14}$

$2 = l_{21}\cancel{u_{11}}$ $\quad 1 = \cancel{l_{21}u_{12}} + u_{22}$ $\quad 3 = \underset{2}{l_{21}}\underset{1}{u_{13}} + \underset{1}{u_{23}}$ $\quad 1 = \cancel{l_{21}u_{14}} + u_{24}$ $\quad 3 = l_{31}\cancel{u_{14}} + \underset{1}{l_{32}}\underset{1}{u_{24}} + \underset{2}{u_{34}}$

$0 = l_{31}\cancel{u_{11}}$ $\quad 1 = \cancel{l_{31}u_{12}} + \underset{1}{l_{32}}\cancel{u_{22}}$ $\quad 3 = \underset{0}{l_{31}}\underset{1}{u_{13}} + \underset{1}{l_{32}}\underset{1}{u_{23}} + \underset{2}{u_{33}}$

$1 = l_{41}\cancel{u_{11}}$ $\quad 1 = l_{41}\cancel{u_{12}} + \underset{1}{l_{42}}\cancel{u_{22}}$ $\quad 4 = \underset{1}{l_{41}}\underset{1}{u_{13}} + \underset{1}{l_{42}}\underset{1}{u_{23}} + \underset{1}{l_{43}}\underset{2}{u_{33}}$ $\quad 2 = \cancel{l_{41}u_{14}} + l_{42}u_{24} + l_{43}u_{34} + \underset{2}{u_{44}}$

$u_{44} = -1$

$$L = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 2 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 1 & 1 & 1 & 1 \end{pmatrix} \quad U = \begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 2 & 2 \\ 0 & 0 & 0 & -1 \end{pmatrix}$$

1° $Ly = e_1$ $\quad y = (1, 1, 2, -1)^T$

$Ux = y$ $\quad x = (1, 0, 0, 1)^T$

Los dos sistemas

$x_2 = (1, 0, 0, 2)^T$

$x_3 = (0, 1, 1, 0)^T$

Factorización LU matrices simétricas definidas positivas
válida siempre
perfeccionamiento adicional

$$\mathbf{L} = \mathbf{U}^T$$

$\mathbf{A} \in \mathbb{R}^{N\times N}$ *definida positiva*

$$\mathbf{x} \in \mathbb{R}^N \backslash \{\mathbf{0}\} \ \Rightarrow \ \mathbf{x}^T\mathbf{A}\mathbf{x} > 0$$

Factorización tipo Cholesky

Sea $\mathbf{A} \in \mathbb{R}^{N\times N}$ una matriz simétrica y definida positiva. Entonces existe una matriz triangular superior $\mathbf{U} \in \mathbb{R}^{N\times N}$ con coeficientes positivos en su diagonal principal y de forma que

> 0

$$\mathbf{A} = \mathbf{U}^T\mathbf{U}.$$

Tal matriz triangular es única y, de hecho, para todo $j = 1, \ldots, N$

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

y

$$u_{jj} = \sqrt{a_{jj} - \sum_{k=1}^{j-1} u_{kj}^2}.$$

DEMOSTRACIÓN. Basta proceder por inducción sobre N. Los detalles pueden consultarse en [3, Theorem 3.6]. → no entra □

Algoritmo $\rightsquigarrow$ primer elemento de la diagonal, elemento de la columna 2 sobre la diagonal, elemento 2 de la diagonal, elementos de la columna 3 sobre la diagonal, elemento 3 de la diagonal...

Otro algoritmo heurístico, como con Doolittle, pero eliminando los datos bajo la diagonal principal (redundantes por la simetría del problema) resuelto en prado

Ejercicio

1. Demuestra que una matriz cuadrada $\mathbf{A}$ es simétrica y definida positiva si, y solo si, admite una factorización tipo Cholesky.
2. Comprueba que la matriz

$$\mathbf{A} = \begin{bmatrix} 4 & 2 & -2 \\ 2 & 2 & -3 \\ -2 & -3 & 14 \end{bmatrix}$$

miramos si es tipo cholesky

es definida positiva.

3. Resuelve el sistema lineal $\mathbf{Ax} = \begin{bmatrix} 4 \\ 0 \\ 2 \end{bmatrix}$ por el método de Cholesky.
4. Encuentra una matriz cuadrada de orden 3×3 que sea simétrica pero no definida positiva.

no entra

Matrices banda algoritmos de factorización especialmente eficientes $\rightsquigarrow$ Relación de Ejercicios

Final tema análisis del error

II.2. Métodos iterativos: métodos de Jacobi y Gauss–Seidel

01/04/20

Métodos iterativos $\rightsquigarrow$ solución de un sistema de ecuaciones lineales cuadrado y compatible determinado como límite de una sucesión

Cada término de la sucesión se genera de forma recursiva a partir del anterior $\rightsquigarrow$ *iteradores*

Sistemas de grandes dimensiones y matriz de coeficientes dispersa (número de coeficientes no nulos relativamente pequeño) $\rightsquigarrow$ problemas prácticos: análisis matricial de estructuras, método de elementos finitos...

II.2.1. Métodos iterativos: convergencia y consistencia

Sistema de ecuaciones lineales cuadrado y unisolvente → SCD

$$\mathbf{A}\mathbf{x} = \mathbf{b}$$

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

Método iterativo

Datos x_0, B, c → no cambian; se cambia x_n

Función de vectores a vector $N=1$ que con escalares

$$\left|\begin{array}{l} \mathbf{x}_0 \text{ dado} \\ n \geq 1 \ \Rightarrow \ \mathbf{x}_n = \mathbf{B}\mathbf{x}_{n-1} + \mathbf{c} \end{array}\right.$$

matriz asociada fija

con $\mathbf{B} \in \mathbb{R}^{N\times N}$, $\mathbf{x}_0, \mathbf{c} \in \mathbb{R}^N$

$\mathbf{x}$ solución del sistema → razonamiento heurístico

$$\lim_{n\to\infty} \mathbf{x}_n = \mathbf{x}$$

→ x_0 da igual como lo elija

$$x_n = Bx_{n-1} + c$$
$$\downarrow \quad \downarrow \quad \downarrow$$
$$x \quad Bx \quad c$$

$$\Downarrow \ (\mathbf{u} \in \mathbb{R}^N \mapsto \mathbf{B}\mathbf{u} + \mathbf{c} \in \mathbb{R}^N \text{ continua})$$

$$\mathbf{x} = \mathbf{B}\mathbf{x} + \mathbf{c}$$

consistencia del método con el sistema → Comprobarla a nivel práctico es resolver el sistema, no tiene sentido

$$Ax = b$$

$$x = A^{-1}b$$

Ejercicio

Demuestra que la consistencia del método iterativo con el sistema equivale a

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

Esto sigue siendo resolver el sistema

convergencia a la solución del sistema $\Rightarrow$ consistencia del método con el sistema

Recíproco falso:

Ejercicio

Sea $\mathbf{I} \in \mathbb{R}^{N\times N}$ la matriz identidad de orden N y sea $\mathbf{b} \in \mathbb{R}^{N}$. Dados el sistema de ecuaciones lineales y el método iterativo

$B = -I$

$$A = 2\mathbf{I}\mathbf{x} = \mathbf{b} \quad \text{y} \quad \left| \begin{array}{l} \mathbf{x}_0 \text{ dado} \\ n \geq 1 \;\Rightarrow\; \mathbf{x}_n = -\mathbf{x}_{n-1} + \mathbf{c} \end{array} \right. ,$$

comprueba que el método es consistente con el sistema si, y solo si,

$$\mathbf{c} = \mathbf{b}$$

y que converge a la solución del sistema cuando, y solo cuando,

$$\mathbf{c} = 2\mathbf{x}_0.$$

Pseudorrecíproco:

Proposición

Supongamos que $N \geq 1$, $\mathbf{A}, \mathbf{B} \in \mathbb{R}^{N \times N}$ con $\mathbf{A}$ regular, $\mathbf{x}_0, \mathbf{b}, \mathbf{c} \in \mathbb{R}^N$ y que el método iterativo

$$\left| \begin{array}{l} \mathbf{x}_0 \text{ dado} \\ n \geq 1 \;\Rightarrow\; \mathbf{x}_n = \mathbf{B}\mathbf{x}_{n-1} + \mathbf{c} \end{array} \right.$$

es consistente con el sistema unisolvente $\mathbf{A}\mathbf{x} = \mathbf{b}$. Entonces

el método iterativo converge a la solución del sistema cualquiera sea $\mathbf{x}_0 \in \mathbb{R}^N$

$$\Updownarrow$$

$$\rho(\mathbf{B}) < 1.$$

DEMOSTRACIÓN. Consistencia, $n \geq 1$

Término cualquiera

Por la consistencia

$$\begin{aligned} \mathbf{x}_n - \mathbf{x} &= \mathbf{B}\mathbf{x}_{n-1} + \mathbf{c} - \mathbf{x} \\ &= \mathbf{B}\mathbf{x}_{n-1} + (\mathbf{I} - \mathbf{B})\mathbf{x} - \mathbf{x} \\ &= \mathbf{B}(\mathbf{x}_{n-1} - \mathbf{x}) \end{aligned}$$

$B(x_{n-2} - x)$ ... $B(x_0 - x)$

$$\text{recursivamente} \quad \rightsquigarrow \quad \mathbf{x}_n - \mathbf{x} = \mathbf{B}^n(\mathbf{x}_0 - \mathbf{x}) \tag{2}$$

$\boxed{\Downarrow}$ Relación de recurrencia (2) + convergencia para todo $\mathbf{x}_0 \in \mathbb{R}^N$

$$\lim_{n\to\infty} \mathbf{B}^n(\mathbf{x}_0 - \mathbf{x}) = 0$$

Si pusiera vectores Bu, Todas las columnas convergen a 0

$$\Updownarrow$$

$$\lim_{n\to\infty} \mathbf{B}^n = \mathbf{0}$$

$$\Updownarrow$$

$$\rho(\mathbf{B}) < 1$$

$\boxed{\Uparrow}$ $\mathbf{x}_0 \in \mathbb{R}^N$

$$\rho(\mathbf{B}) < 1 \;\Rightarrow\; \lim_{n\to\infty} \mathbf{B}^n = \mathbf{0}$$

Relación de recurrencia (2), $\|\cdot\|$ en $\mathbb{R}^N$ y su matricial inducida en $\mathbb{R}^{N\times N}$, notada de la misma forma, $n \geq 1$

la que quiera

$$\begin{aligned} 0 \leq \|\mathbf{x}_n - \mathbf{x}\| &= \|\mathbf{B}^n(\mathbf{x}_0 - \mathbf{x})\| \\ &\leq \underbrace{\|\mathbf{B}^n\|}_{0}\|\mathbf{x}_0 - \mathbf{x}\| \end{aligned}$$

norma inducida

$$\lim_{n\to\infty} \mathbf{x}_n = \mathbf{x}$$

$\square$

02/04/20

Observación

Convergencia *para todo* $\mathbf{x}_0 \in \mathbb{R}^N$

Falso si método iterativo consistente con el sistema y no converge para todo $\mathbf{x}_0 \in \mathbb{R}^N$: sistema del Ejercicio anterior

$\mathbf{b} := \mathbf{0} =: \mathbf{c} \Rightarrow$ consistencia

convergencia *solo* cuando

$$\mathbf{c} = 2\mathbf{x}_0 \;\Leftrightarrow\; \mathbf{x}_0 = \mathbf{0}$$

y

$$\rho(\mathbf{B}) = \rho(-\mathbf{I}) = 1$$

II.2.2. Generación de métodos iterativos. Jacobi y Gauss–Seidel

Proposición anterior $\rightsquigarrow$ *procedimiento de diseño de métodos iterativos*

Automáticamente consistentes $\rightsquigarrow$ elimina el grave problema que surge al comprobar dicha condición: hay que conocer a priori la solución del sistema... ¡que es justo lo que se pretende aproximar!

$$\mathbf{Ax} = \mathbf{b}$$

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

Como A) siempre se puede hacer

$$\mathbf{A} = \mathbf{M} - \mathbf{N} = 0$$

$\mathbf{M}$ regular (siempre posible por ser $\mathbf{A}$ regular, métodos de interés $\mathbf{N}$ no nula)

$$\mathbf{Ax} = \mathbf{b} \;\Leftrightarrow\; (\mathbf{M} - \mathbf{N})\mathbf{x} = \mathbf{b} \;\Leftrightarrow\; \mathbf{x} = \mathbf{M}^{-1}\mathbf{N}\mathbf{x} + \mathbf{M}^{-1}\mathbf{b}$$

Sugiere

$$\mathbf{B} = \mathbf{M}^{-1}\mathbf{N}, \quad \mathbf{c} = \mathbf{M}^{-1}\mathbf{b}$$

Matriz fija M es regular Vector fija

Método iterativo

$$\left| \begin{array}{l} \mathbf{x}_0 \text{ dado} \\ n \geq 1 \;\Rightarrow\; \mathbf{x}_n = \mathbf{M}^{-1}\mathbf{N}\mathbf{x}_{n-1} + \mathbf{M}^{-1}\mathbf{b} \end{array} \right.$$

Consistente con el sistema

$$\begin{aligned} (\mathbf{I} - \mathbf{B})\mathbf{A}^{-1}\mathbf{b} &= (\mathbf{I} - \mathbf{M}^{-1}\mathbf{N})\mathbf{A}^{-1}\mathbf{b} \\ &= (\mathbf{M}^{-1}\mathbf{M} - \mathbf{M}^{-1}\mathbf{N})\mathbf{A}^{-1}\mathbf{b} \\ &= \mathbf{M}^{-1}(\mathbf{M} - \mathbf{N})\mathbf{A}^{-1}\mathbf{b} \\ &= \mathbf{M}^{-1}\mathbf{b} \\ &= \mathbf{c} \end{aligned}$$

Corolario

Sean $\mathbf{A}, \mathbf{M}, \mathbf{N} \in \mathbb{R}^{N\times N}$ con $\mathbf{A}$ y $\mathbf{M}$ regulares de forma que $\mathbf{A} = \mathbf{M} - \mathbf{N}$ y sean $\mathbf{b}, \mathbf{x}_0 \in \mathbb{R}^N$. Consideremos el sistema

$$\mathbf{A}\mathbf{x} = \mathbf{b}$$

y el método iterativo

$$\left| \begin{array}{l} \mathbf{x}_0 \text{ dado} \\ n \geq 1 \;\Rightarrow\; \mathbf{x}_n = \mathbf{M}^{-1}\mathbf{N}\mathbf{x}_{n-1} + \mathbf{M}^{-1}\mathbf{b} \end{array} \right. .$$

Entonces

el método iterativo converge a la solución del sistema, cualquiera sea $\mathbf{x}_0 \in \mathbb{R}^N$

$$\Updownarrow$$

$$\rho(\mathbf{M}^{-1}\mathbf{N}) < 1.$$

Todo método iterativo convergente es de esta forma:

Proposición

Sean $\mathbf{A}, \mathbf{B} \in \mathbb{R}^{N\times N}$, con $\mathbf{A}$ regular, sean $\mathbf{b}, \mathbf{c} \in \mathbb{R}^N$ y consideremos el sistema de ecuaciones lineales

$$\mathbf{Ax} = \mathbf{b}$$

y el método

$$\left| \begin{array}{l} \mathbf{x}_0 \text{ dado} \\ n \geq 1 \;\Rightarrow\; \mathbf{x}_n = \mathbf{B}\mathbf{x}_{n-1} + \mathbf{c} \end{array} \right. ,$$

que supondremos que converge hacia la solución del sistema para cualquier estimación inicial $\mathbf{x}_0 \in \mathbb{R}^N$. Entonces existe una descomposición de la matriz de coeficientes

$$\mathbf{A} = \mathbf{M} - \mathbf{N},$$

con $\mathbf{M}, \mathbf{N} \in \mathbb{R}^{N\times N}$ y $\mathbf{M}$ regular, tales que

$$\mathbf{B} = \mathbf{M}^{-1}\mathbf{N}$$

y

$$\mathbf{c} = \mathbf{M}^{-1}\mathbf{b}.$$

DEMOSTRACIÓN. → no entra

convergencia del método para todo (¡basta un!) $\mathbf{x}_0 \in \mathbb{R}^N \Rightarrow$ consistencia

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

Pretendemos que $\mathbf{c}$ sea $\mathbf{M}^{-1}\mathbf{b} \rightsquigarrow \mathbf{M}^{-1}\mathbf{b} = (\mathbf{I} - \mathbf{B})\mathbf{A}^{-1}\mathbf{b}$

Puede conseguirse si

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

equivalentemente ($\mathbf{I} - \mathbf{B}$ es regular por ser $\rho(\mathbf{B}) < 1$)

$$\mathbf{M} := \mathbf{A}(\mathbf{I} - \mathbf{B})^{-1}.$$

Esta elección $\wedge$ debe ser $\mathbf{N}$ con $\mathbf{M}^{-1}\mathbf{N} = \mathbf{B}$

$$\mathbf{N} := \mathbf{A}(\mathbf{I} - \mathbf{B})^{-1}\mathbf{B}$$

Claramente

$$\mathbf{A} = \mathbf{M} - \mathbf{N}$$

$\square$

$$\left| \begin{array}{l} \mathbf{x}_0 \text{ dado} \\ n \geq 1 \;\; \Rightarrow \;\; \mathbf{x}_n = \mathbf{M}^{-1}\mathbf{N}\mathbf{x}_{n-1} + \mathbf{M}^{-1}\mathbf{b} \end{array} \right.$$

con $\mathbf{M}$ regular, $\mathbf{A} = \mathbf{M} - \mathbf{N}$

multiplico por M

Equivalentemente

M Triangular

$$\left| \begin{array}{l} \mathbf{x}_0 \text{ dado} \\ n \geq 1 \;\; \Rightarrow \;\; \mathbf{M}\mathbf{x}_n = \mathbf{N}\mathbf{x}_{n-1} + \mathbf{b} \end{array} \right.$$

-> aquí resuelvo un sistema y no tengo que invertir matrices

- $\rho(\mathbf{M}^{-1}\mathbf{N}) < 1$ -> La convergencia equivale a esto
- iteración $\mathbf{x}_n$ solución del sistema

$$\mathbf{M}\mathbf{x}_n = \mathbf{N}\mathbf{x}_{n-1} + \mathbf{b}$$

sistema resoluble sin un alto coste operativo $\rightsquigarrow$ $\mathbf{M}$ triangular

Métodos iterativos más populares: *Jacobi* y *Gauss–Seidel*

$\mathbf{A} \in \mathbb{R}^{N \times N}$ regular, matrices diagonal y triangulares

$$\mathbf{D} := \begin{bmatrix} a_{11} & 0 & \cdots & 0 \\ 0 & a_{22} & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & a_{NN} \end{bmatrix}$$

$$\mathbf{E} := \begin{bmatrix} 0 & 0 & 0 & \cdots & 0 \\ -a_{21} & 0 & 0 & \cdots & 0 \\ -a_{31} & -a_{32} & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ -a_{N1} & -a_{N2} & -a_{N3} & \cdots & 0 \end{bmatrix}$$

$$\mathbf{F} := \begin{bmatrix} 0 & \cdots & -a_{12} & -a_{1\ N-1} & -a_{1N} \\ \vdots & & \vdots & \vdots & \vdots \\ 0 & \cdots & 0 & -a_{N-2\ N-1} & -a_{N-2\ N} \\ 0 & \cdots & 0 & 0 & -a_{N-1\ N} \\ 0 & \cdots & 0 & 0 & 0 \end{bmatrix}$$

$$i, j = 1, \ldots, N \;\Rightarrow\; d_{ij} := a_{ij}\delta_{ij}$$

δ_{ij} *delta de Kronecker* (1 en la diagonal y 0 fuera)

$$i, j = 1, \ldots, N \;\Rightarrow\; e_{ij} := \left| \begin{array}{ll} -a_{ij}, & \text{si } j < i \\ 0, & \text{en caso contrario} \end{array} \right.$$

$$i, j = 1, \ldots, N \;\Rightarrow\; f_{ij} := \left| \begin{array}{ll} -a_{ij}, & \text{si } i < j \\ 0, & \text{en caso contrario} \end{array} \right.$$

A verifica la *hipótesis adicional*

→ es necesaria

$$a_{11}a_{22}\cdots a_{NN} \neq 0$$

Método de Jacobi $\rightsquigarrow$ $\mathbf{A} = \mathbf{M} - \mathbf{N}$ con

→ es regular

$$\mathbf{M} := \mathbf{D} \quad \text{y} \quad \mathbf{N} := \mathbf{E} + \mathbf{F}$$

Método de Gauss–Seidel $\rightsquigarrow$ $\mathbf{A} = \mathbf{M} - \mathbf{N}$ con

→ es regular

$$\mathbf{M} := \mathbf{D} - \mathbf{E} \quad \text{y} \quad \mathbf{N} := \mathbf{F}$$

¿Dónde se ha usado la hipótesis adicional de no nulidad de los elementos de la diagonal principal de **A**?

Método de Jacobi

$$\left| \begin{array}{l} \mathbf{x}_0 \text{ dado} \\ n \geq 1 \;\Rightarrow\; \mathbf{D}\mathbf{x}_n = (\mathbf{E} + \mathbf{F})\mathbf{x}_{n-1} + \mathbf{b} \end{array} \right.$$

Expresado en coordenadas:

$$\mathbf{x}_0 = [x_{01}, \ldots, x_{0N}]^T$$

→ el que quiera lo guardo, por el nube

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

no puede ser 0

Método de Gauss–Seidel

$$\left|\begin{array}{l} \mathbf{x}_0 \text{ dado} \\ n \geq 1 \;\Rightarrow\; (\mathbf{D}-\mathbf{E})\mathbf{x}_n = \mathbf{F}\mathbf{x}_{n-1} + \mathbf{b} \end{array}\right.$$

Expresado en coordenadas:

$$\mathbf{x}_0 = [x_{01}, \ldots, x_{0N}]^T$$

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

- Similitud: descritos a partir del esquema obtenido al despejar en $\mathbf{Ax} = \mathbf{b}$ x_1 de la primera ecuación, x_2 de la segunda y así hasta x_N de la N-ésima

$$
\left|
\begin{array}{l}
x_1 = \dfrac{1}{a_{11}} \left(b_1 - a_{12}x_2 - a_{13}x_3 - \cdots - a_{1N}x_N \right) \\[2ex]
x_2 = \dfrac{1}{a_{22}} \left(b_2 - a_{21}x_1 - a_{23}x_3 - \cdots - a_{2N}x_N \right) \\[2ex]
\vdots \\
x_N = \dfrac{1}{a_{NN}} \left(b_N - a_{N1}x_1 - a_{N2}x_2 - \cdots - a_{N\ N-1}x_{N-1} \right)
\end{array}
\right.
$$

Jacobi $\rightsquigarrow$ a la derecha las coordenadas de una iteración para obtener a la izquierda la siguiente
Gauss–Seidel $\rightsquigarrow$ a la derecha las coordenadas de una iteración y las que se acaban de hallar de la siguiente más arriba para determinar las de la siguiente a la izquierda

- Diferencia: en Jacobi el vector en cada iteración se calcula a partir del anterior y, en cambio, en Gauss–Seidel, el vector en cada iteración usa las coordenadas que ya se han calculado en la iteración actual

¿Método de Gauss–Seidel más eficiente que método de Jacobi?

03/04/20

Quizás sí...

Ejemplo

$$\begin{bmatrix} 2 & 1 & 3 \\ -1 & 3 & 2 \\ 1 & 4 & 6 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 9 \\ -1 \\ 11 \end{bmatrix}, \qquad \text{solución } x_1 = 1,\ x_2 = -2,\ x_3 = 3$$

Métodos de Jacobi y Gauss–Seidel: coordenadas $\rightsquigarrow$ estimación inicial $\mathbf{x}_0 = [0, 0, 0]^T$

Despejamos x_i de la ecuación i-ésima ($i = 1, 2, 3$)

$$\left| \begin{aligned} x_1 &= 4.5 - 0.5x_2 - 1.5x_3 \\ x_2 &= -\frac{1}{3} + \frac{1}{3}x_1 - \frac{2}{3}x_3 \\ x_3 &= \frac{11}{6} - \frac{1}{6}x_1 - \frac{3}{2}x_2 \end{aligned} \right.$$

Por tanto:

- Jacobi: el algoritmo parte de $\mathbf{x}_0$ y para cada $n \geq 1$

$$\left| \begin{array}{l} x_{n1} = 4.5 - 0.5x_{n-1\ 2} - 1.5x_{n-1\ 3} \\ \\ x_{n2} = -\dfrac{1}{3} + \dfrac{1}{3}x_{n-1\ 1} - \dfrac{2}{3}x_{n-1\ 3} \\ \\ x_{n3} = \dfrac{11}{6} - \dfrac{1}{6}x_{n-1\ 1} - \dfrac{3}{2}x_{n-1\ 2} \end{array} \right. ,$$

obteniendo (truncando)

$$\begin{array}{lll} x_{01} = 0 & x_{02} = 0 & x_{03} = 0 \\ x_{11} = 4.5 & x_{12} = -0.333 & x_{13} = 1.833 \\ \ldots & \ldots & \ldots \\ x_{20\ 1} = 1.308 & x_{20\ 2} = -1.670 & x_{20\ 3} = 2.702 \end{array}$$

- Gauss–Seidel: con la estimación inicial $\mathbf{x}_0$, para todo $n \geq 1$ tenemos que

Iteración actual

$$\left|\begin{array}{l} x_{n1} = 4.5 - 0.5x_{n-1\ 2} - 1.5x_{n-1\ 3} \\[2ex] x_{n2} = -\dfrac{1}{3} + \dfrac{1}{3}x_{n1} - \dfrac{2}{3}x_{n-1\ 3} \\[2ex] x_{n3} = \dfrac{11}{6} - \dfrac{1}{6}x_{n1} - \dfrac{3}{2}x_{n2} \end{array}\right. ,$$

con lo que se generan (truncando) los datos numéricos

$$\begin{array}{lll} x_{01} = 0 & x_{02} = 0 & x_{03} = 0 \\ x_{11} = 4.5 & x_{12} = 1.166 & x_{13} = 0.305 \\ \ldots & \ldots & \ldots \\ x_{20\ 1} = 1.035 & x_{20\ 2} = -1.961 & x_{20\ 3} = 2.968 \end{array} .$$

□

Parece que es mejor Gauss-Seidel, pero no es cierto

...¡Pero no en general!

- Velocidad de convergencia de Jacobi independiente de la de Gauss–Seidel
- Convergencia para Gauss–Seidel independiente de la de Jacobi

Son independientes, no hay uno mejor que otro

- Velocidad de convergencia de Jacobi independiente de la de Gauss–Seidel

Medida de la velocidad de convergencia

$$\mathbf{Ax} = \mathbf{b}, \text{ sistema unisolvente}$$

Método iterativo convergente a la solución del sistema para cualquier estimación inicial (no necesariamente Jacobi o Gauss–Seidel)

$$\left| \begin{array}{l} \mathbf{x}_0 \text{ dado} \\ n \geq 1 \ \Rightarrow \ \mathbf{x}_n = \mathbf{B}\mathbf{x}_{n-1} + \mathbf{c} \end{array} \right.$$

$\rho(\mathbf{B}) < 1 \wedge$ consistencia

$$(\mathbf{I} - \mathbf{B})\mathbf{x} = \mathbf{c}$$

compatible determinado y con la misma solución que $\mathbf{Ax} = \mathbf{b}$

Demostración de la última proposición: norma y su matricial inducida, $\|\mathbf{B}\| < 1$, $n \geq 1$

$$\|\mathbf{x}_n - \mathbf{x}\| \leq \|\mathbf{B}\|^n \|\mathbf{x}_0 - \mathbf{x}\|$$

→ Pero no elegimos la norma, valdría que coger la más conocida

→ La norma de B decide la rapidez o de cuyo es un método iterativo

$\|\mathbf{B}\| < 1$, $n \geq 1$

$$\|\mathbf{x}_n - \mathbf{x}\| \leq \|\mathbf{B}\|^n \|\mathbf{x}_0 - \mathbf{x}\|$$

Cuanto menor sea $\|\mathbf{B}\|$ mejor será la convergencia de la sucesión de iteradores hacia la solución del sistema

Puede probarse ([2, p. 12, Theorem 3]) que si $\mathbf{A} \in \mathbb{R}^{N \times N}$, entonces

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

Nos puede dar una idea de cómo de rápida o lenta es la convergencia

¡Aunque $\mathbf{A}$ real, el resultado es complejo ($\mathbb{C}$)!

Esperable que cuanto menor sea $\rho(\mathbf{B})$ mejor será la convergencia de la sucesión de iteradores hacia la solución del sistema

Dos sistemas en los que la relación de orden entre los radios espectrales de la matriz $\mathbf{M}^{-1}\mathbf{N}$ para el método de Jacobi y Gauss–Seidel van en sentidos distintos:

Ejemplo

$$\mathbf{A}_1 = \begin{bmatrix} 4 & 1 & 1 \\ 2 & -9 & 0 \\ 0 & -8 & -6 \end{bmatrix}, \quad \mathbf{A}_2 = \begin{bmatrix} 7 & 6 & 9 \\ 4 & 5 & -4 \\ -7 & -3 & 8 \end{bmatrix}$$

(redondeando)

- $\mathbf{A}_1$:

$$\rho_{\text{Jacobi}} := \rho(\mathbf{D}^{-1}(\mathbf{E}+\mathbf{F})) = 0.444$$

$$\rho_{\text{Gauss–Seidel}} := \rho((\mathbf{D}-\mathbf{E})^{-1}\mathbf{F}) = 0.019$$

- $\mathbf{A}_2$:

$$\rho_{\text{Jacobi}} := \rho(\mathbf{D}^{-1}(\mathbf{E}+\mathbf{F})) = 0.641$$

$$\rho_{\text{Gauss–Seidel}} := \rho((\mathbf{D}-\mathbf{E})^{-1}\mathbf{F}) = 0.775$$

$\square$

Relación de Ejercicios $\rightsquigarrow$ cálculo iteraciones

- Convergencia para Gauss–Seidel independiente de la de Jacobi

Ejemplo

$$\begin{bmatrix} 3 & 0 & 4 \\ 7 & 1 & 2 \\ -1 & 1 & 9 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 7 \\ 10 \\ 9 \end{bmatrix}, \quad \begin{bmatrix} -3 & 3 & -6 \\ -4 & 7 & -8 \\ 5 & 7 & -9 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} -6 \\ -5 \\ 3 \end{bmatrix}$$

- sistema izquierda $\rightsquigarrow$ convergencia para Gauss–Seidel pero no para Jacobi (redondeando):

$$\rho_{\text{Jacobi}} := \rho(\mathbf{D}^{-1}(\mathbf{E}+\mathbf{F})) = 1.037$$
$$\rho_{\text{Gauss–Seidel}} := \rho((\mathbf{D}-\mathbf{E})^{-1}\mathbf{F}) = 0.963$$

- sistema derecha $\rightsquigarrow$ convergencia para Jacobi pero no para Gauss–Seidel (redondeando):

$$\rho_{\text{Jacobi}} := \rho(\mathbf{D}^{-1}(\mathbf{E}+\mathbf{F})) = 0.813$$
$$\rho_{\text{Gauss–Seidel}} := \rho((\mathbf{D}-\mathbf{E})^{-1}\mathbf{F}) = 1.111$$

$\square$

$\mathbf{A} \in \mathbb{R}^{N\times N}$ es *diagonalmente estrictamente dominante* si

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

se entiende

Tanto Jacobi como Gauss Seidel son convergentes

Método de Jacobi converge para cualquier sistema de ecuaciones lineales que tenga a $\mathbf{A}$ por matriz de coeficientes a su solución, cualquiera sea la elección de la estimación inicial: vid. ~~Relación de Ejercicios~~

Ídem ~~Gauss–Seidel (prueba diferente~~)

Proposición

Sea $\mathbf{A} \in \mathbb{R}^{N\times N}$ una matriz diagonalmente estrictamente dominante. Entonces los métodos de Jacobi y de Gauss–Seidel, para todo sistema de ecuaciones lineales que tenga como matriz de coeficientes $\mathbf{A}$, convergen hacia su solución, independientemente de la estimación inicial que se fije.

Recíproco falso: basta considerar cualquier sistema en el que la matriz de coeficientes sea la matriz $\mathbf{A}_1$ (o $\mathbf{A}_2$) del penúltimo ejemplo

Observaciones

- Si se aplica al sistema de partida una transformación elemental tan simple como intercambiar de posición dos ecuaciones y se usa el mismo método iterativo con los dos sistemas, uno puede converger y otro no. Este hecho se prueba en la Relación de Ejercicios. La idea es que este tipo de transformación elemental no solo puede modificar claramente el hecho de que la matriz de coeficientes sea diagonalmente estrictamente dominante –que es una condición suficiente para la convergencia de Jacobi y Gauss–Seidel– sino que además puede cambiar el radio espectral.

- El número de operaciones que hay que realizar para pasar de una iteración a la siguiente en los métodos de Jacobi y Gauss–Seidel es de N^2 para un sistema de N ecuaciones y N incógnitas. Por tanto, si N es grande, requiere en principio menos operaciones que los directos. Además, y a diferencia de estos últimos, aprovecha la estructura de la matriz de coeficientes cuando es dispersa, tal y como ocurre con los sistemas que surgen en problemas de análisis de estructuras o de elementos finitos.

- La elección que hemos hecho para diseñar los métodos de Jacobi y Gauss–Seidel es la más popular, pero no la única. Una clase de métodos iterativos que los incluye y que son también de uso extendido está constituida por los llamados *métodos de relajación para Jacobi y Gauss–Seidel*, una especie de combinación convexa de ellos. Los detalles pueden consultarse en [3, Sections 4.2.1, 4.2.2, 4.2.3].

No

II.3. Análisis del error

Error cometido al resolver mediante un método numérico un sistema de ecuaciones lineales unisolvente y con igual número de ecuaciones e incógnitas

Primer resultado aplicable a cualquier método directo o iterativo $\rightsquigarrow$ estimaciones del error relativo cometido al resolver de forma aproximada un sistema

Errores derivados del método usado, o los debidos al redondeo, propagación, etc.

Error relativo controlado en función únicamente del vector de términos independientes y de la solución aproximada

Proposición

Sean $\mathbf{A} \in \mathbb{R}^{N\times N}$ una matriz regular y $\mathbf{x}, \mathbf{u}, \mathbf{b} \in \mathbb{R}^N$ con $\|\mathbf{x}\|\|\mathbf{b}\| \neq 0$ y de forma que

$$\mathbf{Ax} = \mathbf{b}.$$

Entonces, para cualquier norma en $\mathbb{R}^N$ se cumplen las desigualdades:

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

donde $c(\mathbf{A})$ es el condicionamiento de la matriz $\mathbf{A}$ relativo a la norma matricial inducida por $\|\cdot\|$.

Demostración. Primera desigualdad

$$\left.\begin{array}{c} \|\mathbf{Au}-\mathbf{b}\| \leq \|\mathbf{A}\|\|\mathbf{x}-\mathbf{u}\| \\ \|\mathbf{x}\| \leq \|\mathbf{A}^{-1}\|\|\mathbf{b}\| \end{array}\right| \Rightarrow \|\mathbf{Au}-\mathbf{b}\|\|\mathbf{x}\| \leq c(\mathbf{A})\|\mathbf{x}-\mathbf{u}\|\|\mathbf{b}\|$$

Segunda desigualdad (razonamiento similar, tomando $\mathbf{v} = \mathbf{Au}$)

$$\begin{aligned}
\frac{\|\mathbf{x}-\mathbf{u}\|}{\|\mathbf{x}\|} &= \frac{\|\mathbf{A}^{-1}\mathbf{b}-\mathbf{A}^{-1}\mathbf{v}\|}{\|\mathbf{x}\|} \\
&\leq \frac{\|\mathbf{A}^{-1}\|\|\mathbf{b}-\mathbf{v}\|}{\|\mathbf{x}\|} \\
&\leq \|\mathbf{A}\|\|\mathbf{A}^{-1}\|\frac{\|\mathbf{b}-\mathbf{v}\|}{\|\mathbf{b}\|} \\
&= c(\mathbf{A})\frac{\|\mathbf{Au}-\mathbf{b}\|}{\|\mathbf{b}\|}
\end{aligned}$$

□

Interpretación

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

$\mathbf{x}$ solución del sistema $\mathbf{Ax}=\mathbf{b}$
$\mathbf{u}$ solución aproximada

Estimación del error relativo de la solución en función del condicionamiento de $\mathbf{A}$ y el error relativo generado al tomar $\mathbf{Au}$ por $\mathbf{b}$

$$\mathbf{Au}-\mathbf{b}$$

residuo, en general no nulo

Tema I: un residuo pequeño no garantiza un error relativo pequeño de la solución

Ejemplo

$0 < \alpha \leq 1$

$$\mathbf{A} = \begin{bmatrix} 1 & 1 \\ 1 & 1-\alpha \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} 2 \\ 2-\alpha \end{bmatrix}, \quad \mathbf{u} = \begin{bmatrix} 0 \\ 2 \end{bmatrix}$$

$$\mathbf{x} = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \text{ solución del sistema } \mathbf{Ax} = \mathbf{b}, \quad \mathbf{Au} = \begin{bmatrix} 2 \\ 2-2\alpha \end{bmatrix}$$

$$\|\mathbf{A}\|_\infty = \left\| \begin{bmatrix} 1 & 1 \\ 1 & 1-\alpha \end{bmatrix} \right\|_\infty = 2, \quad \|\mathbf{A}^{-1}\|_\infty = \left\| \begin{bmatrix} 1-\frac{1}{\alpha} & \frac{1}{\alpha} \\ \frac{1}{\alpha} & -\frac{1}{\alpha} \end{bmatrix} \right\|_\infty = \frac{2}{\alpha}$$

$$\frac{\alpha}{4} \frac{\|\mathbf{Au}-\mathbf{b}\|}{\|\mathbf{b}\|} \leq \frac{\|\mathbf{x}-\mathbf{u}\|}{\|\mathbf{x}\|} \leq \frac{4}{\alpha} \frac{\|\mathbf{Au}-\mathbf{b}\|}{\|\mathbf{b}\|}$$

control bueno si α está próximo a 1 y un mal control si α está cerca de 0

Explícitamente

$$\frac{\alpha^2}{8} \leq 1 \leq 2$$

$\square$

Segundo resultado aplicable a cualquier método iterativo $\rightsquigarrow$ estimaciones del error absoluto cometido al resolver de forma aproximada un sistema

Proposición

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

$$\left| \begin{array}{l} \mathbf{x}_0 \text{ dado} \\ n \geq 1 \ \Rightarrow \ \mathbf{x}_n = \mathbf{B}\mathbf{x}_{n-1} + \mathbf{c} \end{array} \right.$$

converge a la solucion $\mathbf{x}$ del sistema $\mathbf{A}\mathbf{x} = \mathbf{b}$, cualquiera sea $\mathbf{x}_0 \in \mathbb{R}^N$. Si además $\|\cdot\|$ es una norma en $\mathbb{R}^N$ con norma matricial inducida en $\mathbb{R}^{N\times N}$ denotada de igual forma, tal que $\|\mathbf{B}\| < 1$, entonces, para todo $n \geq 1$ se tiene:

(i) $\|\mathbf{x}_n - \mathbf{x}\| \leq \dfrac{\|\mathbf{B}\|^n}{1 - \|\mathbf{B}\|}\|\mathbf{x}_1 - \mathbf{x}_0\|$.

(ii) $\|\mathbf{x}_n - \mathbf{x}\| \leq \|\mathbf{B}\|\|\mathbf{x}_{n-1} - \mathbf{x}\|$.

(iii) $\|\mathbf{x}_n - \mathbf{x}\| \leq \dfrac{\|\mathbf{B}\|}{1 - \|\mathbf{B}\|}\|\mathbf{x}_n - \mathbf{x}_{n-1}\|$.

DEMOSTRACIÓN.

$\boxed{\text{(i)}}$ $n \geq 1$

$$\begin{aligned} \|\mathbf{x}_{n+1} - \mathbf{x}_n\| &= \|\mathbf{B}\mathbf{x}_n - \mathbf{B}\mathbf{x}_{n-1}\| \\ &\leq \|\mathbf{B}\|\|\mathbf{x}_n - \mathbf{x}_{n-1}\| \end{aligned}$$

luego

$$\|\mathbf{x}_{n+1} - \mathbf{x}_n\| \leq \|\mathbf{B}\|^n \|\mathbf{x}_1 - \mathbf{x}_0\|$$

$$\Downarrow$$

$m \geq n \geq 1$

$$\begin{aligned} \|\mathbf{x}_n - \mathbf{x}_m\| &\leq \sum_{j=0}^{m-n-1} \|\mathbf{x}_{j+n+1} - \mathbf{x}_{j+n}\| \\ &\leq \sum_{j=0}^{m-n-1} \|\mathbf{B}\|^{j+n} \|\mathbf{x}_1 - \mathbf{x}_0\| \\ &\leq \frac{\|\mathbf{B}\|^n}{1 - \|\mathbf{B}\|} \|\mathbf{x}_1 - \mathbf{x}_0\| \end{aligned}$$

Tomar límite en $m \to \infty$

$\boxed{\text{(ii)}}$ $n \geq 1$

$$\begin{aligned} \|\mathbf{x}_n - \mathbf{x}\| &= \|\mathbf{B}\mathbf{x}_{n-1} + \mathbf{c} - \mathbf{B}\mathbf{x} - \mathbf{c}\| \\ &\leq \|\mathbf{B}\| \|\mathbf{x}_{n-1} - \mathbf{x}\| \end{aligned}$$

$\boxed{\text{(iii)}}$ $n \geq 1$, desigualdad triangular

$$\|\mathbf{x}_{n-1} - \mathbf{x}\| \leq \|\mathbf{x}_{n-1} - \mathbf{x}_n\| + \|\mathbf{x}_n - \mathbf{x}\|$$

(ii)

$$\|\mathbf{x}_n - \mathbf{x}\| \leq \|\mathbf{B}\| \|\mathbf{x}_n - \mathbf{x}_{n-1}\| + \|\mathbf{B}\| \|\mathbf{x}_n - \mathbf{x}\|$$

estimación pedida (¡reorganizar!) $\square$

Observaciones

1. Las acotaciones (i) y (ii) dan una estimación del error absoluto, aunque hay una diferencia importante entre ambas: la primera no necesita conocer la solución exacta $\mathbf{x}$.

2. La estimación del error absoluto (iii) no solo se obtiene sin necesidad de conocer $\mathbf{x}$ sino que además constituye un criterio de parada cuando se programa el método numérico y se alcanza una tolerancia dada.

3. Las estimaciones obtenidas son caso particular de las que proporciona el Teorema del punto fijo de Banach (véase, por ejemplo, [1, Theorem 5.1.3]), lo cual no es de extrañar, pues estamos usando técnicas inspiradas en dicho teorema.

Ejemplo ilustrativo en Relación de Problemas

II.4. Bibliografía

1. K. Atkinson, W. Ham, *Theoretical numerical analysis. A functional analysis framework*, third edition, Texts in Applied Mathematics **39**, Springer, Dordrecht, 2009.
2. E. Isaacson, H. Keller, *Analysis of numerical methods*, Wiley, New York, 1966.
3. A. Quarteroni, R. Sacco, F. Saleri, *Numerical mathematics*, second edition, Texts in Applied Mathematics **37** Springer–Verlag, Berlin, 2007.