Solución de sistemas de ecuaciones lineales

Condicionamiento. Métodos directos: eliminación gaussiana, factorización LU. Pivoteo. Métodos indirectos. Gauss-Seidel.

Manuel Carlevaro

Departamento de Ingeniería Mecánica

Grupo de Materiales Granulares - UTN FRLP

manuel.carlevaro@gmail.com

$$a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n = b_1$$

$$a_{21}x_2 + a_{22}x_2 + \dots + a_{2n}x_n = b_2$$

$$\vdots$$

$$a_{n1}x_1 + a_{n2}x_2 + \dots + a_{nn}x_n = b_n$$

$$a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n = b_1$$

$$a_{21}x_2 + a_{22}x_2 + \dots + a_{2n}x_n = b_2$$

$$\vdots$$

$$a_{n1}x_1 + a_{n2}x_2 + \dots + a_{nn}x_n = b_n$$

O, en forma matricial:

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

$$a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n = b_1$$

$$a_{21}x_2 + a_{22}x_2 + \dots + a_{2n}x_n = b_2$$

$$\vdots$$

$$a_{n1}x_1 + a_{n2}x_2 + \dots + a_{nn}x_n = b_n$$

O, en forma matricial:

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

Matriz de coeficientes aumentada:

$$\begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} & b_1 \\ a_{21} & a_{22} & \dots & a_{2n} & b_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{N1} & a_{N2} & \dots & a_{nn} & b_n \end{bmatrix}$$

$$a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n = b_1$$

$$a_{21}x_2 + a_{22}x_2 + \dots + a_{2n}x_n = b_2$$

$$\vdots$$

$$a_{n1}x_1 + a_{n2}x_2 + \dots + a_{nn}x_n = b_n$$

O, en forma matricial:

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

Matriz de coeficientes aumentada:

$$\begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} & b_1 \\ a_{21} & a_{22} & \dots & a_{2n} & b_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{N1} & a_{N2} & \dots & a_{nn} & b_n \end{bmatrix}$$

Si $\mathbb{A} \in \mathbb{R}^{n \times n}$ y $\mathbf{b} \in \mathbb{R}$, la existencia y unicidad de la solución está asegurada si una de las siguientes condiciones se cumple:

- ▶ A es invertible (no singular)
- Prg(A) = n
- lacktriangle El sistema homogéneo $\mathbb{A} oldsymbol{x} = oldsymbol{0}$ admite solo la solución nula.

$$a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n = b_1$$

$$a_{21}x_2 + a_{22}x_2 + \dots + a_{2n}x_n = b_2$$

$$\vdots$$

$$a_{n1}x_1 + a_{n2}x_2 + \dots + a_{nn}x_n = b_n$$

O, en forma matricial:

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

Matriz de coeficientes aumentada:

$$\begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} & b_1 \\ a_{21} & a_{22} & \dots & a_{2n} & b_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{N1} & a_{N2} & \dots & a_{nn} & b_n \end{bmatrix}$$

Si $\mathbb{A} \in \mathbb{R}^{n \times n}$ y $\mathbf{b} \in \mathbb{R}$, la existencia y unicidad de la solución está asegurada si una de las siguientes condiciones se cumple:

- ▶ A es invertible (no singular)
- $ightharpoonup \operatorname{rg}(A) = n$
- El sistema homogéneo $\mathbb{A}\boldsymbol{x}=\mathbf{0}$ admite solo la solución nula.

Solución: regla de Cramer

$$x_j = \frac{\Delta_j}{\det \mathbb{A}}$$

Esfuerzo computacional: $\mathcal{O}((n+1)!)$. n=50, Intel i7: 200 Gflops $\approx 5\times 10^{45}$ años.

Estabilidad de la solución:

$$(\mathbb{A} + \delta \mathbb{A})(\boldsymbol{x} + \frac{\delta \boldsymbol{x}}{\delta \boldsymbol{x}}) = \boldsymbol{b} + \delta \boldsymbol{b}$$

Unicidad de la solución: \mathbb{A} es no singular: $|\mathbb{A}| \neq 0$. Número de condición: $\operatorname{cond}(\mathbb{A}) = ||\mathbb{A}|| ||\mathbb{A}^{-1}||$. Se puede demostrar¹, que si

$$\operatorname{cond}(\mathbb{A}) \frac{\|\delta \mathbb{A}\|}{\|\mathbb{A}\|} < 1$$

se cumple:

$$\frac{\|\delta \boldsymbol{x}\|}{\|\boldsymbol{x}\|} \leq \frac{\operatorname{cond}(\mathbb{A})}{1 - \operatorname{cond}(\mathbb{A}) \frac{\|\delta \mathbf{A}\|}{\|\mathbf{A}\|}} \left(\frac{\|\delta \boldsymbol{b}\|}{\|\boldsymbol{b}\|} + \frac{\|\delta \mathbb{A}\|}{\|\mathbf{A}\|} \right)$$

En general es muy costoso evaluar $\|\mathbb{A}\|$. Usualmente se compara $|\mathbb{A}|$ con a_{ij} .

$$\begin{cases} 2x + y = 3 \\ 2x + 1.001y = 0 \end{cases}, \quad |\mathbb{A}| = 0.002$$

Solución: x = 1501.5, y = -3000.

Ejemplo:

¹Ver González, *Introducción al cálculo numérico* (2014), sección 2.3., y Quarteroni, Sacco y Saleri, *Numerical Mathematics* (2000), sección 3.1.

Estabilidad de la solución:

$$(\mathbb{A} + \delta \mathbb{A})(\boldsymbol{x} + \frac{\delta \boldsymbol{x}}{\boldsymbol{x}}) = \boldsymbol{b} + \delta \boldsymbol{b}$$

Unicidad de la solución: \mathbb{A} es no singular: $|\mathbb{A}| \neq 0$. Número de condición: $\operatorname{cond}(\mathbb{A}) = \|\mathbb{A}\| \|\mathbb{A}^{-1}\|$. Se puede demostrar¹, que si

$$\operatorname{cond}(\mathbb{A})\frac{\|\delta\mathbb{A}\|}{\|\mathbb{A}\|} < 1$$

se cumple:

$$\frac{\|\delta \boldsymbol{x}\|}{\|\boldsymbol{x}\|} \leq \frac{\operatorname{cond}(\mathbb{A})}{1 - \operatorname{cond}(\mathbb{A})\frac{\|\delta\mathbb{A}\|}{\|\mathbb{A}\|}} \left(\frac{\|\delta \boldsymbol{b}\|}{\|\boldsymbol{b}\|} + \frac{\|\delta\mathbb{A}\|}{\|\mathbb{A}\|}\right)$$

En general es muy costoso evaluar $\|\mathbb{A}\|$. Usualmente se compara $|\mathbb{A}|$ con a_{ij} .

Ejemplo:

$$\begin{cases} 2x + y = 3\\ 2x + 1.001y = 0 \end{cases}, \quad |\mathbb{A}| = 0.002$$

Solución: x = 1501.5, y = -3000.

$$\begin{cases} 2x + y = 3 \\ 2x + 1.002y = 0 \end{cases}, \quad |\mathbb{A}| = 0.004$$

Solución: x = 751.5, y = -1500.

0.1% de cambio en $a_{ij}\mapsto$ 100% de cambio en ${m x}$.

Mal condicionamiento

Si la solución de un sistema lineal cambia mucho cuando el problema cambia muy poco, la matriz está mal condicionada.

¹Ver González, Introducción al cálculo numérico (2014), sección 2.3., y Quarteroni, Sacco y Saleri, Numerical Mathematics (2000), sección 3.1.

Otro ejemplo:

$$\mathbb{A} = \begin{bmatrix} 1 & 100 \\ 0 & 1 \end{bmatrix}, \quad |\mathbb{A}| = 1$$

$$\mathbb{A}^{-1} = \begin{bmatrix} 1 & -100 \\ 0 & 1 \end{bmatrix}, \quad |\mathbb{A}^{-1}| = 1$$

Otro ejemplo:

$$\mathbb{A} = \begin{bmatrix} 1 & 100 \\ 0 & 1 \end{bmatrix}, \quad |\mathbb{A}| = 1$$

$$\mathbb{A}^{-1} = \begin{bmatrix} 1 & -100 \\ 0 & 1 \end{bmatrix}, \quad |\mathbb{A}^{-1}| = 1$$

Soluciones:

$$m{b} = egin{bmatrix} 100 \\ 1 \end{bmatrix}
ightarrow m{x} = egin{bmatrix} 0 \\ 1 \end{bmatrix}, \ m{b} = egin{bmatrix} 100 \\ 0 \end{bmatrix}
ightarrow m{x} = egin{bmatrix} 100 \\ 0 \end{bmatrix}$$

1% de cambio en $\boldsymbol{b}\mapsto$ 100% de cambio en \boldsymbol{x} .

Otro ejemplo:

$$\mathbb{A} = \begin{bmatrix} 1 & 100 \\ 0 & 1 \end{bmatrix}, \quad |\mathbb{A}| = 1$$

$$\mathbb{A}^{-1} = \begin{bmatrix} 1 & -100 \\ 0 & 1 \end{bmatrix}, \quad |\mathbb{A}^{-1}| = 1$$

Soluciones:

$$m{b} = egin{bmatrix} 100 \ 1 \end{bmatrix}
ightarrow m{x} = egin{bmatrix} 0 \ 1 \end{bmatrix}, \ m{b} = egin{bmatrix} 100 \ 0 \end{bmatrix}
ightarrow m{x} = egin{bmatrix} 100 \ 0 \end{bmatrix}$$

1% de cambio en $b \mapsto 100\%$ de cambio en x.

```
1 #!/usr/bin/env python3
2
3 import numpy as np
4 from numpy import linalg as la
5
6 a = np.array([[1, 100],[0, 1]])
7 ai = np.array([[1, -100],[0, 1]])
8 a_norm = la.norm(a, ord=2)
9 ai_norm = la.norm(ai, ord=2)
10 print(f"||a|| = {a_norm}, ||ai|| = {ai_norm}")
11 print(f"cond(a) = {a_norm * ai_norm}")
12 print(f"cond(a) = {la.cond(a)}")
```

```
$ ./cond.py
||a|| = 100.00999900019995, ||ai|| = 100.00999900019995
cond(a) = 10001.999900019995
cond(a) = 10001.999900019995
```

Métodos:

- ▶ Directos: alcanzan la solución en un número finito de pasos. $\mathcal{O}(2/3N^3)$.
 - > Eliminación gaussiana
 - ightharpoonup Factorización LU
 - \rightarrow LDM T
 - > Factorización de Cholesky
 - ightharpoonup Factorización QR

- Iterativos: más eficientes en casos particulares. $\mathcal{O}(N^2)$.
 - > Jacobi
 - > Gauss-Seidel
 - > Subespacios de Krylov
 - > GMRES

Métodos:

- ▶ Directos: alcanzan la solución en un número finito de pasos. $\mathcal{O}(2/3N^3)$.
 - > Eliminación gaussiana
 - ightharpoonup Factorización LU
 - ightharpoonup LDM T
 - > Factorización de Cholesky
 - > Factorización ${\cal Q}{\cal R}$

- Iterativos: más eficientes en casos particulares. $\mathcal{O}(N^2)$.
 - Jacobi
 - > Gauss-Seidel
 - > Subespacios de Krylov
 - ➤ GMRES

Métodos directos:

- $\bullet \ a_{ik} \rightleftharpoons a_{jk}, |\mathbb{A}| \to -|\mathbb{A}|$
- $k \times a_{ij}, |\mathbb{A}| \to k \times |\mathbb{A}|$
- $ightharpoonup a_{ik} k \, a_{ik}, |\mathbb{A}|$ no cambia

Métodos:

- ▶ Directos: alcanzan la solución en un número finito de pasos. $\mathcal{O}(2/3N^3)$.
 - > Eliminación gaussiana
 - ightharpoonup Factorización LU
 - \rightarrow LDM T
 - > Factorización de Cholesky
 - ightharpoonup Factorización QR

- lterativos: más eficientes en casos particulares. $\mathcal{O}(N^2)$.
 - Jacobi
 - > Gauss-Seidel
 - > Subespacios de Krylov
 - > GMRES

Métodos directos:

- $k \times a_{ij}, |\mathbb{A}| \to k \times |\mathbb{A}|$
- $ightharpoonup a_{ik} k \, a_{ik}, |\mathbb{A}|$ no cambia

Método	Forma inicial	Forma final
Eliminación gaussiana	$\mathbb{A}\mathbf{x} = \mathbf{b}$	$\mathbb{U}\mathbf{x}=\mathbf{c}$
Descomposición LU	$\mathbb{A}\mathbf{x}=\mathbf{b}$	$\mathbb{L}\mathbb{U}\mathbf{x}=\mathbf{b}$
Eliminación de Gauss-Jordan	$\mathbb{A}\mathbf{x} = \mathbf{b}$	$\mathbb{I}\mathbf{x}=\mathbf{c}$

Dos fases: eliminación y solución.

Fase de eliminación:

Utiliza solo una operación elemental:

$$Ec(i) \leftarrow Ec(i) - \lambda \times Ec(j)$$

 $\ \, {\rm donde}\,\, \, {\rm Ec}(j)\,\, {\rm se}\,\, {\rm denomina}\,\, ecuaci\'on\,\, pivote.$

Dos fases: eliminación y solución.

Fase de eliminación:

Utiliza solo una operación elemental:

$$\mathsf{Ec}(i) \leftarrow \mathsf{Ec}(i) - \lambda \times \mathsf{Ec}(j)$$

donde Ec(j) se denomina ecuación pivote.

Ejemplo:

$$\begin{cases} 4x_1 - 2x_2 + x_3 = 11 \\ -2x_1 + 4x_2 - 2x_3 = -16 \\ x_1 - 2x_2 + 4x_3 = 17 \end{cases}$$

$$\begin{bmatrix} 4 & -2 & 1 & 11 \\ -2 & 4 & -2 & -16 \\ 1 & -2 & 4 & 17 \end{bmatrix}$$

Dos fases: eliminación y solución.

Fase de eliminación:

Utiliza solo una operación elemental:

$$\mathsf{Ec}(i) \leftarrow \mathsf{Ec}(i) - \lambda \times \mathsf{Ec}(j)$$

donde Ec(j) se denomina ecuación pivote.

Ejemplo:

$$\begin{cases} 4x_1 - 2x_2 + x_3 = 11 \\ -2x_1 + 4x_2 - 2x_3 = -16 \\ x_1 - 2x_2 + 4x_3 = 17 \end{cases}$$

$$\begin{bmatrix} 4 & -2 & 1 & 11 \\ -2 & 4 & -2 & -16 \\ 1 & -2 & 4 & 17 \end{bmatrix}$$

Ec(1): pivote; Ec(2) \leftarrow Ec(2) $-(-0.5) \times$ Ec(1); Ec(3) \leftarrow Ec(3) $-0.25 \times$ Ec(1)

$$\begin{bmatrix} 4 & -2 & 1 & 11 \\ 0 & 3 & -1.5 & -10.5 \\ 0 & -1.5 & 3.75 & 14.25 \end{bmatrix}$$

Dos fases: eliminación y solución.

Fase de eliminación:

Utiliza solo una operación elemental:

$$\mathsf{Ec}(i) \leftarrow \mathsf{Ec}(i) - \lambda \times \mathsf{Ec}(j)$$

donde Ec(j) se denomina ecuación pivote.

Ejemplo:

$$\begin{cases} 4x_1 - 2x_2 + x_3 = 11 \\ -2x_1 + 4x_2 - 2x_3 = -16 \\ x_1 - 2x_2 + 4x_3 = 17 \end{cases}$$

$$\begin{bmatrix} 4 & -2 & 1 & | & 11 \\ -2 & 4 & -2 & | & -16 \\ 1 & -2 & 4 & | & 17 \end{bmatrix}$$

Ec(1): pivote; Ec(2) \leftarrow Ec(2) -

$$Ec(2) \leftarrow Ec(2) - (-0.5) \times Ec(1);$$

 $Ec(3) \leftarrow Ec(3) - 0.25 \times Ec(1)$

$$\begin{bmatrix} 4 & -2 & 1 & 11 \\ 0 & 3 & -1.5 & -10.5 \\ 0 & -1.5 & 3.75 & 14.25 \end{bmatrix}$$

Ec(2): pivote;

$$Ec(3) \leftarrow Ec(3) - (-0.5) \times Ec(2)$$

$$\begin{bmatrix} 4 & -2 & 1 & 11 \\ 0 & 3 & -1.5 & -10.5 \\ 0 & 0 & 3 & 9 \end{bmatrix}$$

Dos fases: eliminación y solución.

Fase de eliminación:

Utiliza solo una operación elemental:

$$\mathsf{Ec}(i) \leftarrow \mathsf{Ec}(i) - \lambda \times \mathsf{Ec}(j)$$

donde Ec(j) se denomina ecuación pivote.

Ejemplo:

$$\begin{cases} 4x_1 - 2x_2 + x_3 = 11 \\ -2x_1 + 4x_2 - 2x_3 = -16 \\ x_1 - 2x_2 + 4x_3 = 17 \end{cases}$$

$$\begin{bmatrix} 4 & -2 & 1 & 11 \\ -2 & 4 & -2 & -16 \\ 1 & -2 & 4 & 17 \end{bmatrix}$$

Ec(1): pivote;

$$Ec(2) \leftarrow Ec(2) - (-0.5) \times Ec(1);$$

 $Ec(3) \leftarrow Ec(3) - 0.25 \times Ec(1)$

$$\begin{bmatrix} 4 & -2 & 1 & 11 \\ 0 & 3 & -1.5 & -10.5 \\ 0 & -1.5 & 3.75 & 14.25 \end{bmatrix}$$

Ec(2): pivote;

$$Ec(3) \leftarrow Ec(3) - (-0.5) \times Ec(2)$$

$$\begin{bmatrix} 4 & -2 & 1 & 11 \\ 0 & 3 & -1.5 & -10.5 \\ 0 & 0 & 3 & 9 \end{bmatrix}$$

Bonus: no se altera $|\mathbb{A}|$:

$$|\mathbb{A}| = |\mathbb{U}| = U_{11} \times U_{22} \times \cdots U_{NN}$$

Fase de solución: sustitución hacia atrás.

Algoritmo:

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1k} & \cdots & a_{1j} & \cdots & a_{1n} & b_1 \\ 0 & a_{22} & a_{23} & \cdots & a_{2k} & \cdots & a_{2j} & \cdots & a_{2n} & b_2 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & a_{kk} & \cdots & a_{kj} & \cdots & a_{kn} & b_k \\ \vdots & \vdots & \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & a_{ik} & \cdots & a_{ij} & \cdots & a_{in} & b_i \\ \vdots & \vdots & \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & a_{nk} & \cdots & a_{nj} & \cdots & a_{nn} & b_N \end{bmatrix}$$

Eliminación:

1: for
$$i \leftarrow k+1, k+2, \dots n$$
 do

2:
$$\lambda \leftarrow a_{ik}/a_{kk}$$

$$3: \quad \text{for } j \leftarrow k, n \text{ do}$$

$$a_{ij} \leftarrow a_{ij} - \lambda a_{kj}$$

6:
$$b_i \leftarrow b_i - \lambda b_k$$

7: end for

Solución:

$$x_n = b_n/a_{nn}$$

$$x_k = \left(b_k - \sum_{j=k+1}^n a_{kj}x_j\right) \frac{1}{a_{kk}}$$

$$k = n - 1, n - 2, \dots, 1$$

Factorización o descomposición LU:

$$\mathbb{A}=\mathbb{L}\mathbb{U}$$

Nombre	Restricción	
Descomposición de Doolittle	$L_{ii} = 1, i = 1, 2, \cdots, n$	
Descomposición de Crout	$U_{ii}=1,\ i=1,2,\cdots,n$	
Descomposición de Choleski	$\mathbb{T}_{\star}=\mathbb{T}^T$	
($\mathbb A$ debe ser simétrica y definida positiva)	$\mathbb{L} = \mathbb{U}^{-}$	

Luego de la factorización:

$$\mathbb{L}\mathbb{U}\,oldsymbol{x} = oldsymbol{b}$$

La solución consiste en:

$$\mathbb{L}\, \boldsymbol{y} = \boldsymbol{b}$$

resuelta por sustitución hacia adelante, seguida de:

$$\mathbb{U} \, oldsymbol{x} = oldsymbol{y}$$

que da el resultado $oldsymbol{x}$ obtenido por sustitución hacia atrás.

Método de Doolittle:

$$\mathbb{L} = \begin{bmatrix} 1 & 0 & 0 \\ L_{21} & 1 & 0 \\ L_{31} & L_{32} & 1 \end{bmatrix} \qquad \mathbb{U} = \begin{bmatrix} U_{11} & U_{12} & U_{13} \\ 0 & U_{22} & U_{23} \\ 0 & 0 & U_{33} \end{bmatrix}$$

Encontramos A:

$$\mathbb{A} = \left[\begin{array}{ccc} U_{11} & U_{12} & U_{13} \\ U_{11}L_{21} & U_{12}L_{21} + U_{22} & U_{13}L_{21} + U_{23} \\ U_{11}L_{31} & U_{12}L_{31} + U_{22}L_{32} & U_{13}L_{31} + U_{23}L_{32} + U_{33} \end{array} \right]$$

Método de Doolittle:

$$\mathbb{L} = \begin{bmatrix} 1 & 0 & 0 \\ L_{21} & 1 & 0 \\ L_{31} & L_{32} & 1 \end{bmatrix} \qquad \mathbb{U} = \begin{bmatrix} U_{11} & U_{12} & U_{13} \\ 0 & U_{22} & U_{23} \\ 0 & 0 & U_{33} \end{bmatrix}$$

Encontramos A:

$$\mathbb{A} = \left[\begin{array}{ccc} U_{11} & U_{12} & U_{13} \\ U_{11}L_{21} & U_{12}L_{21} + U_{22} & U_{13}L_{21} + U_{23} \\ U_{11}L_{31} & U_{12}L_{31} + U_{22}L_{32} & U_{13}L_{31} + U_{23}L_{32} + U_{33} \end{array} \right]$$

Aplicamos ahora eliminación gaussiana con las siguientes operaciones elementales:

fila 2
$$\leftarrow$$
 fila 2 $-L_{21} \times$ fila 1 (elimina a_{21}) fila 3 \leftarrow fila 3 $-L_{31} \times$ fila 1 (elimina a_{31})

$$\mathbb{A}' = \begin{bmatrix} U_{11} & U_{12} & U_{13} \\ 0 & U_{22} & U_{23} \\ 0 & U_{22}L_{32} & U_{23}L_{32} + U_{33} \end{bmatrix}$$

Método de Doolittle (cont.): tomamos ahora la segunda fila como pivote:

$$\mathbb{A}' = \left[\begin{array}{ccc} U_{11} & U_{12} & U_{13} \\ 0 & U_{22} & U_{23} \\ 0 & U_{22}L_{32} & U_{23}L_{32} + U_{33} \end{array} \right]$$

fila 3 — fila 3 —
$$L_{32} \times$$
 fila 2 (elimina a_{32})
$$\mathbb{A}'' = \mathbb{U} = \begin{bmatrix} U_{11} & U_{12} & U_{13} \\ 0 & U_{22} & U_{23} \\ 0 & 0 & U_{33} \end{bmatrix}$$

Método de Doolittle (cont.): tomamos ahora la segunda fila como pivote:

$$\mathbb{A}' = \begin{bmatrix} U_{11} & U_{12} & U_{13} \\ 0 & U_{22} & U_{23} \\ 0 & U_{22}L_{32} & U_{23}L_{32} + U_{33} \end{bmatrix}$$

fila 3
$$\leftarrow$$
 fila 3 $-L_{32} imes$ fila 2 (elimina a_{32})

$$\mathbb{A}'' = \mathbb{U} = \left[\begin{array}{ccc} U_{11} & U_{12} & U_{13} \\ 0 & U_{22} & U_{23} \\ 0 & 0 & U_{33} \end{array} \right]$$

Características del método de Doolittle:

- La matriz U es idéntica a la matriz triangular superior que resulta de la eliminación gaussiana
- Los elementos no diagonales de $\mathbb L$ son los factores que multiplican a la ecuación pivote durante la eliminación gaussiana: L_{ij} es el multiplicador que elimina a_{ij}

Método de Doolittle (cont.): tomamos ahora la segunda fila como pivote:

$$\mathbb{A}' = \begin{bmatrix} U_{11} & U_{12} & U_{13} \\ 0 & U_{22} & U_{23} \\ 0 & U_{22}L_{32} & U_{23}L_{32} + U_{33} \end{bmatrix}$$

Características del método de Doolittle:

- La matriz U es idéntica a la matriz triangular superior que resulta de la eliminación gaussiana
- lackbox Los elementos no diagonales de $\mathbb L$ son los factores que multiplican a la ecuación pivote durante la eliminación gaussiana: L_{ij} es el multiplicador que elimina a_{ij}

fila 3
$$\leftarrow$$
 fila 3 $-L_{32} \times$ fila 2 (elimina a_{32})
$$\mathbb{A}'' = \mathbb{U} = \begin{bmatrix} U_{11} & U_{12} & U_{13} \\ 0 & U_{22} & U_{23} \\ 0 & 0 & U_{33} \end{bmatrix}$$

Almacenamiento:

$$\mathbb{L}/\mathbb{U} = \left[\begin{array}{ccc} U_{11} & U_{12} & U_{13} \\ L_{21} & U_{22} & U_{23} \\ L_{31} & L_{32} & U_{33} \end{array} \right]$$

Sistema:

$$2x_1 - x_2 = 1$$
$$-x_1 + 2x_2 - x_3 = 0$$
$$-x_2 + x_3 = 0$$

Solución:
$$x_1 = x_2 = x_3 = 1$$

$$\left[\begin{array}{cc|cc}
2 & -1 & 0 & 1 \\
-1 & 2 & -1 & 0 \\
0 & -1 & 1 & 0
\end{array} \right]$$

Sistema:

$$2x_1 - x_2 = 1$$
$$-x_1 + 2x_2 - x_3 = 0$$
$$-x_2 + x_3 = 0$$

Solución:
$$x_1 = x_2 = x_3 = 1$$

$$\begin{bmatrix}
2 & -1 & 0 & | & 1 \\
-1 & 2 & -1 & | & 0 \\
0 & -1 & 1 & | & 0
\end{bmatrix}$$
OK

$$\left[\begin{array}{ccc|c} 0 & -1 & 1 & 0 \\ -1 & 2 & -1 & 0 \\ 2 & -1 & 0 & 1 \end{array}\right]$$

NOK

$$[\mathbb{A}|\mathbf{b}] = \begin{bmatrix} \varepsilon & -1 & 1 & 0 \\ -1 & 2 & -1 & 0 \\ 2 & -1 & 0 & 1 \end{bmatrix} \to [\mathbb{A}'|\mathbf{b}'] = \begin{bmatrix} \varepsilon & -1 & 1 & 0 \\ 0 & 2 - 1/\varepsilon & -1 + 1/\varepsilon & 0 \\ 0 & -1 + 2/\varepsilon & -2/\varepsilon & 1 \end{bmatrix}$$

$$[\mathbb{A}|\mathbf{b}] = \begin{bmatrix} \varepsilon & -1 & 1 & 0 \\ -1 & 2 & -1 & 0 \\ 2 & -1 & 0 & 1 \end{bmatrix} \rightarrow [\mathbb{A}'|\mathbf{b}'] = \begin{bmatrix} \varepsilon & -1 & 1 & 0 \\ 0 & 2 - 1/\varepsilon & -1 + 1/\varepsilon & 0 \\ 0 & -1 + 2/\varepsilon & -2/\varepsilon & 1 \end{bmatrix}$$

Almacenamiento en la memoria ($\varepsilon \ll 1$):

$$[\mathbb{A}'|\mathbf{b'}] = \begin{bmatrix} \varepsilon & -1 & 1 & 0 \\ 0 & -1/\varepsilon & 1/\varepsilon & 0 \\ 0 & 2/\varepsilon & -2/\varepsilon & 1 \end{bmatrix}$$

$$[\mathbb{A}|\mathbf{b}] = \begin{bmatrix} \varepsilon & -1 & 1 & 0 \\ -1 & 2 & -1 & 0 \\ 2 & -1 & 0 & 1 \end{bmatrix} \rightarrow [\mathbb{A}'|\mathbf{b}'] = \begin{bmatrix} \varepsilon & -1 & 1 & 0 \\ 0 & 2 - 1/\varepsilon & -1 + 1/\varepsilon & 0 \\ 0 & -1 + 2/\varepsilon & -2/\varepsilon & 1 \end{bmatrix}$$

Almacenamiento en la memoria ($\varepsilon \ll 1$):

$$[\mathbb{A}'|\mathbf{b'}] = \begin{bmatrix} \varepsilon & -1 & 1 & 0 \\ 0 & -1/\varepsilon & 1/\varepsilon & 0 \\ 0 & 2/\varepsilon & -2/\varepsilon & 1 \end{bmatrix}$$

 \mathbb{A} es diagonalmente dominante si:

$$|a_{ii}| > \sum_{\substack{j=1\\i\neq i}}^{N} |a_{ij}|; \ (i=1,2,\ldots,N)$$

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

$$\left[
\begin{array}{ccc}
4 & -2 & 1 \\
-2 & 4 & -1 \\
1 & -1 & 3
\end{array}
\right]$$

NO diagonalmente dominante

Diagonalmente dominante

En todos los casos: si $a_{ii} \neq 0 \mapsto$ no intercambiar filas.

▶ Pivoteo trivial:

▶ $a_{ii} = 0 \mapsto$ buscar el primer $a_{ki} \neq 0 (k > i)$ e intercambiar filas $i \leftrightarrows k$.

En todos los casos: si $a_{ii} \neq 0 \mapsto$ no intercambiar filas.

- ▶ Pivoteo trivial:
 - ▶ $a_{ii} = 0 \mapsto$ buscar el primer $a_{ki} \neq 0 (k > i)$ e intercambiar filas $i \leftrightarrows k$.
- ▶ Pivoteo parcial:
 - $oldsymbol{\lambda}$ $a_{ii}=0\mapsto$ buscar la fila k tal que $|a_{ki}|=\max_{j>i}|a_{ji}|\wedge a_{ji}
 eq 0$ e intercambiar filas $i\leftrightarrows k$.

En todos los casos: si $a_{ii} \neq 0 \mapsto$ no intercambiar filas.

- ▶ Pivoteo trivial:
 - ▶ $a_{ii} = 0 \mapsto$ buscar el primer $a_{ki} \neq 0 (k > i)$ e intercambiar filas $i \leftrightarrows k$.
- ▶ Pivoteo parcial:
 - $m{\lambda}$ $a_{ii}=0\mapsto$ buscar la fila k tal que $|a_{ki}|=\max_{j>i}|a_{ji}|\wedge a_{ji}\neq 0$ e intercambiar filas $i\leftrightarrows k$.
- ▶ Pivoteo parcial escalado:
 - lacksquare Calcular $s_i = \max_{1 \leq j \leq N} |a_{ij}|, \ i=1,\ldots,N$
 - $\ \ \, a_{ii} = 0 \mapsto \text{buscar la fila k tal que } \frac{|a_{ki}|}{s_k} = \max_{j>i} \frac{|a_{ji}|}{s_j} \wedge a_{ji} \neq 0 \text{ e intercambiar filas } i \leftrightarrows k \text{ y } s_i \leftrightarrows s_k.$

En todos los casos: si $a_{ii} \neq 0 \mapsto$ no intercambiar filas.

- ▶ Pivoteo trivial:
 - ▶ $a_{ii} = 0 \mapsto$ buscar el primer $a_{ki} \neq 0 (k > i)$ e intercambiar filas $i \leftrightarrows k$.
- ▶ Pivoteo parcial:
 - $m{\lambda}$ $a_{ii}=0\mapsto$ buscar la fila k tal que $|a_{ki}|=\max_{j>i}|a_{ji}|\wedge a_{ji}\neq 0$ e intercambiar filas $i\leftrightarrows k$.
- ▶ Pivoteo parcial escalado:
 - lacksquare Calcular $s_i = \max_{1 \leq j \leq N} |a_{ij}|, \ i=1,\ldots,N$
 - $\ \, a_{ii} = 0 \mapsto \text{buscar la fila } k \text{ tal que } \frac{|a_{ki}|}{s_k} = \max_{j>i} \frac{|a_{ji}|}{s_j} \wedge a_{ji} \neq 0 \text{ e intercambiar filas } i \leftrightarrows k \text{ y } s_i \leftrightarrows s_k.$
- ▶ Pivoteo completo o maximal:
 - **>** $a_{ii} = 0$ → buscar la fila j > i y columna k > i tal que $|a_{jk}| = \max_{\substack{l > i \\ m > i}} |a_{lm}| \land a_{jk} \neq 0$ e intercambiar filas $i \leftrightarrows j$ y columnas $i \leftrightarrows k$.

Pivoteo: estrategias

En todos los casos: si $a_{ii} \neq 0 \mapsto$ no intercambiar filas.

▶ Pivoteo trivial:

▶ $a_{ii} = 0 \mapsto$ buscar el primer $a_{ki} \neq 0 (k > i)$ e intercambiar filas $i \leftrightarrows k$.

▶ Pivoteo parcial:

 $m{\lambda}$ $a_{ii}=0\mapsto$ buscar la fila k tal que $|a_{ki}|=\max_{j>i}|a_{ji}|\wedge a_{ji}\neq 0$ e intercambiar filas $i\leftrightarrows k$.

▶ Pivoteo parcial escalado:

- lacksquare Calcular $s_i = \max_{1 \leq j \leq N} |a_{ij}|, \ i = 1, \dots, N$
- $\ \, a_{ii} = 0 \mapsto \text{buscar la fila } k \text{ tal que } \frac{|a_{ki}|}{s_k} = \max_{j>i} \frac{|a_{ji}|}{s_j} \wedge a_{ji} \neq 0 \text{ e intercambiar filas } i \leftrightarrows k \text{ y } s_i \leftrightarrows s_k.$

▶ Pivoteo completo o maximal:

> $a_{ii} = 0$ → buscar la fila j > i y columna k > i tal que $|a_{jk}| = \max_{\substack{l > i \\ m > i}} |a_{lm}| \land a_{jk} \neq 0$ e intercambiar filas $i \leftrightarrows j$ y columnas $i \leftrightarrows k$.

Nota: en matemática " $x=0,\,x\neq0$ ", en mundo real " $|x|<arepsilon,\,|x|>arepsilon$ ".

Ejemplo con Python:

```
1 #!/usr/bin/env pvthon3
 2
 3 import numpy as np
 4 from scipy import linalg
 5
 6 = np.array([[3, 2, 0], [1, -1, 0], [0, 5, 1]])
 7 b = np.array([2, 4, -1])
 8 \times = linalg.solve(a. b)
 9 print('x = ', x)
1.0
11 # Verificación
12 print('a @ x ?', (np.dot(a, x) == b))
13 print('a @ x ?', np.allclose(a @ x, b))
14 print('a @ x =', a @ x)
15 # Descomposición LU
16 p. l. u = linalg.lu(a)
17 print(p)
18 print(l)
19 print(u)
```

```
$ ./eiemplo.pv
x = [2. -2. 9.]
a@x?[True True]
a @ x ? True
a @ x = [2, 4, -1]
[[1. 0. 0.]
[0. 0. 1.]
[0. 1. 0.]]
[[ 1.
           0
                  0 .
ΓΘ.
          1.
                    0.
[ 0.33333333 -0.33333333 1.
         2.
[[3.
                  0.
[0.
         5. 1.
[0.
         0
                  0.3333333311
```

Ventajas:

- ▶ Es posible almacenar solo los elementos no nulos de la matriz
- ▶ Las iteraciones son auto-correctivas

Ventajas:

- ▶ Es posible almacenar solo los elementos no nulos de la matriz
- ▶ Las iteraciones son auto-correctivas

Desventajas:

- ▶ Más lentos que los métodos directos
- No siempre convergen (garantizado cuando la matriz es diagonalmente dominante)

Objetivo: a partir de una conjetura inicial $\boldsymbol{x}^{(0)}$, producir una aproximación mejorada $\boldsymbol{x}^{(k+1)}$ a partir de la anterior $\boldsymbol{x}^{(k)}$.

Ventajas:

- ▶ Es posible almacenar solo los elementos no nulos de la matriz
- ▶ Las iteraciones son auto-correctivas

Desventajas:

- ▶ Más lentos que los métodos directos
- No siempre convergen (garantizado cuando la matriz es diagonalmente dominante)

Objetivo: a partir de una conjetura inicial $\boldsymbol{x}^{(0)}$, producir una aproximación mejorada $\boldsymbol{x}^{(k+1)}$ a partir de la anterior $\boldsymbol{x}^{(k)}$.

Si $\mathbb{A} = \mathbb{S} - \mathbb{T}$, $\mathbb{A}x = b$ es igual a $\mathbb{S}x = \mathbb{T}x + b$, entonces:

Iteración:
$$x^{(k)} o x^{(k+1)}: \quad \mathbb{S} {m x}^{(k+1)} = \mathbb{T} {m x}^{(k)} + {m b}$$

Una descomposición útil $\mathbb{S} - \mathbb{T}$ debe cumplir:

- $x^{(k+1)}$ debe ser **fácil de calcular**: \mathbb{S} debe ser una matriz simple e invertible (diagonal o triangular).
- La secuencia $\boldsymbol{x}^{(k)}$ debe **converger** a la solución \boldsymbol{x} . Definiendo el error en la iteración k como:

$$\boldsymbol{e}^{(k)} = \boldsymbol{x} - \boldsymbol{x}^{(k)}$$

podemos sustraer la ecuación de iteración de la ecuación original:

Ecuación de error:
$$\mathbb{S}oldsymbol{e}^{(k+1)} = \mathbb{T}oldsymbol{e}^{(k)}$$

Iniciando con un error $e^{(0)}$, después de k iteraciones se produce el error $e^{(k)} = (\mathbb{S}^{-1}\mathbb{T})^k e^{(0)}$. La convergencia y estabilidad determinan que

$$oldsymbol{x}^{(k)}
ightarrow oldsymbol{x}$$
 exactamente cuando $oldsymbol{e}^{(k)}
ightarrow oldsymbol{0}$

El método iterativo es **convergente** si y solo si todo autovalor de $\mathbb{S}^{-1}\mathbb{T}$ satisface $|\lambda|<1$, la velocidad de convergencia depende del radio espectral:

$$\rho(\mathbb{S}^{(-1)}\mathbb{T}) = \max_{i} |\lambda_{i}|$$

Dado que una solución típica de $e^{(k+1)} = (\mathbb{S}^{-1}\mathbb{T})e^{(k)}$ es una combinación lineal de autovectores, el error luego de k iteraciones es:

$$\boldsymbol{e}^{(k)} = c_1 \lambda_1^k \boldsymbol{v}_1 + \dots + c_n \lambda_n^k \boldsymbol{v}_n$$

el radio espectral $ho=|\lambda_{\max}|$ gobierna la velocidad a la que $e^{(k)}$ converge a cero.

Tres posibilidades:

- 1. \mathbb{S} : parte diagonal de \mathbb{A} (método de Jacobi).
- 2. S: parte triangular de A (método de Gauss-Seidel).
- 3. S: combinación de 1 y 2 (sobrerrelajación sucesiva SOR).

Método de Jacobi

Escribimos las ecuaciones $\mathbb{A}\mathbf{x} = \mathbf{b}$ en notación escalar:

$$\sum_{j=1}^{N} a_{ij} x_j = b_i, \ i = 1, 2, \dots, n$$

Extraemos el término que contiene x_i :

$$a_{ii}x_i + \sum_{\substack{j=1\\j\neq i}}^{N} a_{ij}x_j = b_i, \ i = 1, 2, \dots, n$$

Resolviendo para x_i tenemos:

$$x_i = \frac{1}{a_{ii}} \left(b_i - \sum_{\substack{j=1\\j \neq i}}^{N} a_{ij} x_j \right), i = 1, 2, \dots, n$$

Esquema iterativo: para cada $k \geq 1$, generamos $x_i^{(k)}$ de ${\pmb x}^{(k)}$ a partir de ${\pmb x}^{(k-1)}$ mediante

$$x_i^{(k)} = \frac{1}{a_{ii}} \left(b_i - \sum_{\substack{j=1\\j\neq i}}^N a_{ij} x_j^{(k-1)} \right), \ i = 1, 2, \dots, n$$

Método de Jacobi

Escribimos las ecuaciones Ax = b en notación escalar:

$$\sum_{i=1}^{N} a_{ij} x_j = b_i, \ i = 1, 2, \dots, n$$

Extraemos el término que contiene x_i :

$$a_{ii}x_i + \sum_{j=1}^{N} a_{ij}x_j = b_i, \ i = 1, 2, \dots, n$$

Resolviendo para x_i tenemos:

$$x_i = \frac{1}{a_{ii}} \left(b_i - \sum_{\substack{j=1\\j \neq i}}^{N} a_{ij} x_j \right), i = 1, 2, \dots, n$$

Esquema iterativo: para cada $k \geq 1$, generamos $x_i^{(k)}$ de ${\pmb x}^{(k)}$ a partir de ${\pmb x}^{(k-1)}$ mediante

$$x_i^{(k)} = \frac{1}{a_{ii}} \left(b_i - \sum_{\substack{j=1\\j \neq i}}^N a_{ij} x_j^{(k-1)} \right), \ i = 1, 2, \dots, n$$

Técnica de relajación:

donde p es un entero positivo.

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

Determinación de ω : Si $\Delta x^{(k)} = |x^{(k-1)} - x^{(k)}|$ calculado con $\omega = 1$:

$$\omega_{
m opt}pprox rac{2}{1+\sqrt{1-(\Delta x^{(k+p)})/\Delta x^{(k)})^{1/p}}}$$

16

MÉTODO DE JACOBI (CONT.)

Esquema general con relajación:

- lacktriangle Realizar k iteraciones con $\omega=1$ (k=10). Luego de la iteración k almacenar $\Delta x^{(k)}$.
- $lackbox{$\blacktriangleright$}$ Realizar p iteraciones adicionales y almacenar $\Delta x^{(k+p)}$ para la última iteración.
- lacktriangle Realizar las iteraciones siguientes con $\omega=\omega_{
 m opt}$.

MÉTODO DE JACOBI (CONT.)

Esquema general con relajación:

- Realizar k iteraciones con $\omega = 1$ (k = 10). Luego de la iteración k almacenar $\Delta x^{(k)}$.
- \triangleright Realizar p iteraciones adicionales y almacenar $\Delta x^{(k+p)}$ para la última iteración.
- Realizar las iteraciones siguientes con $\omega = \omega_{\text{opt}}$.

Algoritmo:

Requiere \mathbb{A} , $\mathbf{b} \cup \omega$ Devuelve x

▶ inputs output

1: $\mathbf{x} \leftarrow \text{valores aleatorios, ceros}$

- 2: repeat
- for $i \leftarrow 1, \ldots, N$ do
 - $\sigma \leftarrow 0$
- for $i \leftarrow 1, \ldots, N$ do
- if $i \neq i$ then
- $\sigma \leftarrow \sigma + a_{ij}x_j$
- 8. end if
 - end for
- g. $x_i \leftarrow x_i + \omega \left(\frac{b_i - \sigma}{a_{i,i}} - x_i \right)$
- end for 11:
 - Verificar convergencia
- 13: until Convergencia alcanzada

⊳ Fin *j*-loop ⊳ Fin *i*-loop

17

MÉTODO DE GAUSS-SEIDEL

Para i>1 tenemos $x_1^{(k)}, x_2^{(k)}, \cdots, x_{i-1}^{(k)}$ de ${\boldsymbol x}^{(k)}$:

$$x_i^{(k)} = \frac{1}{a_{ii}} \left(b_i - \sum_{j=1}^{i-1} a_{ij} x_j^{(k)} - \sum_{j=i+1}^{N} a_{ij} x_j^{(k-1)} \right), \ i = 1, 2, \dots, n$$

Para i>1 tenemos $x_1^{(k)},x_2^{(k)},\cdots,x_{i-1}^{(k)}$ de ${\boldsymbol x}^{(k)}$:

$$x_i^{(k)} = \frac{1}{a_{ii}} \left(b_i - \sum_{j=1}^{i-1} a_{ij} x_j^{(k)} - \sum_{j=i+1}^{N} a_{ij} x_j^{(k-1)} \right), \ i = 1, 2, \dots, n$$

Con el método de **sobrerrelajación sucesiva** (SOR):

$$x_i^{(k)} = \frac{\omega}{a_{ii}} \left(b_i - \sum_{j=1}^{i-1} a_{ij} x_j^{(k)} - \sum_{j=i+1}^{N} a_{ij} x_j^{(k-1)} \right) + (1 - \omega) x_i^{(k-1)}, \ i = 1, 2, \dots, n$$

Para i>1 tenemos $x_1^{(k)},x_2^{(k)},\cdots,x_{i-1}^{(k)}$ de ${\boldsymbol x}^{(k)}$:

$$x_i^{(k)} = \frac{1}{a_{ii}} \left(b_i - \sum_{j=1}^{i-1} a_{ij} x_j^{(k)} - \sum_{j=i+1}^{N} a_{ij} x_j^{(k-1)} \right), \ i = 1, 2, \dots, n$$

Con el método de **sobrerrelajación sucesiva** (SOR):

$$x_i^{(k)} = \frac{\omega}{a_{ii}} \left(b_i - \sum_{j=1}^{i-1} a_{ij} x_j^{(k)} - \sum_{j=i+1}^{N} a_{ij} x_j^{(k-1)} \right) + (1 - \omega) x_i^{(k-1)}, \ i = 1, 2, \dots, n$$

EJEMPLO PYTHON

```
1 #!/usr/bin/env python3
 2 import numpy as np
 3
 4 def gauss seidel(A, b, w=1.0, tol=1e-10, max iter=10000):
       n = len(b)
 5
       x = np.zeros like(b, dtype=np.float64)
       for it in range(max_iter):
 7
           for i in range(n):
 8
               sigma = 0
 9
               for j in range(n):
1.0
                   if i != i:
11
12
                        sigma += A[i, i] * x[i]
13
               x[i] = x[i] + 
14
1.5
                   w * ((b[i] - sigma) / A[i, i] - x[i])
16
17
           if np.linalg.norm(A @ x - b, ord=np.inf) < tol:</pre>
               break
18
19
20
       return x. it
```

```
$ ./gauss-seidel.py
Iteración: 374, x: [ 3. -2. 2. 1.]
Residuo: 9.55289e-11
```

- ▶ R.L. Burden, D.J. Faires y A.M. Burden. *Análisis numérico*. 10.ª ed. Mexico: Cengage Learning, 2017. Capítulo 6.
- ▶ Gilbert Strang. *Linear Algebra and Its Applications, 4th Edition.* 4th. Brooks Cole, 2006. Capítulo 7.
- ▶ A.J. Salgado y S.M. Wise. *Classical Numerical Analysis*. Cambridge, United Kingdom: Cambridge University Press, 2023. poi: 10.1017/9781108942607. Capítulo 3.
- ▶ A. Quarteroni, R. Sacco y F. Saleri. *Numerical Mathematics*. New York, United States: Springer-Verlag, 2000. Capítulo 3.
- ▶ Gilbert Strang. *Linear Algebra and Its Applications, 4th Edition.* 4th. Brooks Cole, 2006. Capítulo 7.