Pràctica 2
Resolució de sistemes d'equacions lineals: mètodes iteratius

Índex

1 Introducció

En matemàtiques computacionals, un mètode iteratiu intenta resoldre un problema (per exemple, trobar la solució d'una equació o sistema d'equacions) trobant aproximacions successives a la solució partint d'una estimació inicial. Aquesta aproximació contrasta amb els mètodes directes, que tracten de resoldre el problema mitjançant una successió finita d'operacions i, en absència d'errors d'arrodoniment, donarien una solució exacta (com la resolució d'un sistema d'equacions lineals $\mathsf{A}\vec{x} = \vec{b}$ per eliminació gaussiana). Els mètodes iteratius són freqüentment útils fins i tot per a problemes lineals que involucren un gran nombre de variables (de vegades de l'ordre de milions), on els mètodes directes serien prohibitivament cars (i en alguns casos impossibles), fins i tot amb el poder computacional més gran.

Estudiarem dos mètodes iteratius per resoldre sistemes d'equacions lineals: els mètodes de Jacobi i Gauss-Seidel. En ambdós casos treballarem amb un *sistema quadrat (el mateix nombre d'equacions que d'incògnites) amb solució única*. A més a més, **tots els elements de la diagonal de la matriu de coeficients han de ser no nuls**.

2 El mètode de Jacobi

Considerem un sistema d'equacions lineals que satisfà les hipòtesis abans esmentades. Suposem que la seua expressió matricial és

$$\mathsf{A}\vec{x} = \vec{b}.$$

Considerem la següent descomposició de la matriu de coeficients A:

$$\mathsf{A} = \mathsf{L} + \mathsf{D} + \mathsf{U},$$

on la matriu L és la «part triangular inferior» de A, U és la «part triangular superior» de A i D és la «part diagonal» de A. Vegem un exemple amb l'objectiu d'aclarir aquesta descomposició:

Exemple 1. Prenem la matriu

$$\mathsf{A} = \begin{bmatrix} 10 & 3 & 1 \\ 2 & -10 & 3 \\ 1 & 3 & 10 \end{bmatrix}$$

Aleshores $\mathsf{A} = \mathsf{L} + \mathsf{D} + \mathsf{U}$, on

$$\mathsf{L} = \begin{bmatrix} 0 & 0 & 0 \\ 2 & 0 & 0 \\ 1 & 3 & 0 \end{bmatrix}, \quad \mathsf{D} = \begin{bmatrix} 10 & 0 & 0 \\ 0 & -10 & 0 \\ 0 & 0 & 10 \end{bmatrix}, \quad \mathsf{U} = \begin{bmatrix} 0 & 3 & 1 \\ 0 & 0 & 3 \\ 0 & 0 & 0 \end{bmatrix}.$$

Observem la següent successió d'equivalències:

$$\begin{aligned} \text{El vector } \vec{x} \text{ és una solució del sistema } \mathsf{A}\vec{x} = \vec{b} &\Leftrightarrow (\mathsf{L} + \mathsf{D} + \mathsf{U})\vec{x} = \vec{b} \\ &\Leftrightarrow \mathsf{D}\vec{x} = \vec{b} - (\mathsf{L} + \mathsf{U})\vec{x} \\ &\Leftrightarrow \vec{x} = \mathsf{D}^{-1}[\vec{b} - (\mathsf{L} + \mathsf{U})\vec{x}] \end{aligned} \tag{1}$$

Notem que D és invertible perquè els elements de la seua diagonal són no nuls (per hipòtesi). Així D^{-1} pot calcular-se molt fàcilment perquè és una matriu diagonal: D^{-1} és una matriu diagonal que té com a entrades de la diagonal els inversos dels elements respectius de la diagonal de D.

El mètode de Jacobi és una tècnica iterativa basada en la igualtat (1). Consta dels següents passos: prenem una aproximació inicial de la solució $\vec{x}_0$, la substituïm en el primer membre de (1) i calculem una altra aproximació $\vec{x}_1$, la substituïm en el primer membre de (1) i calculem una altra aproximació $\vec{x}_2$, i així successivament. Fent açò, obtenim una successió de vectors $\vec{x}_0$, $\vec{x}_1$, $\vec{x}_2, \ldots$ tal que

$$\vec{x}_{k+1} = \mathsf{D}^{-1}[\vec{b} - (\mathsf{L} + \mathsf{U})\vec{x}_k], \qquad k = 0, 1, 2, 3, \ldots \tag{2}$$

El fet important ací és la següent propietat:

Proposició 1. Si aquesta successió de vectors és convergent, aleshores el vector límit és una solució del sistema d'equacions lineals.

La demostració d'aquesta propietat és molt senzilla: siga $\vec{v}$ el vector límit de la successió $(\vec{x}_k)$. Prenent límits a ambdós costats de la igualtat (1), obtenim que $\vec{v} = \mathsf{D}^{-1}[\vec{b} - (\mathsf{L}+\mathsf{U})\vec{v}]$, que és equivalent a dir que $\vec{v}$ és una solució del sistema d'equacions lineals (tenint en compte l'anterior successió d'equivalències).

Exemple 2. Considerem el següent sistema d'equacions lineals:

$$\left.\begin{array}{rcl} 10x+3y+z &=& 14 \\ 2x-10y+3z &=& -5 \\ x+3y+10z &=& 14 \end{array}\right\},$$

La matriu de coeficients d'aquest sistema és la matriu A de l'exemple 1. La relació de recurrència (2) és, en aquest cas:

$$\vec{x}_{k+1} = \underbrace{\begin{bmatrix} 1/10 & 0 & 0 \\ 0 & -1/10 & 0 \\ 0 & 0 & 1/10 \end{bmatrix}}_{\mathsf{D}^{-1}} \left(\underbrace{\begin{bmatrix} 14 \\ -5 \\ 14 \end{bmatrix}}_{\vec{b}} - \underbrace{\begin{bmatrix} 0 & 3 & 1 \\ 2 & 0 & 3 \\ 1 & 3 & 0 \end{bmatrix}}_{\mathsf{L+U}} \vec{x}_k \right)$$

Apliquem ara el mètode de Jacobi elegint un vector inicial qualsevol. Prenem, per exemple, $\vec{x}_0 = (0,0,0)$. Aleshores

$$\begin{aligned} \vec{x}_1 &= \begin{bmatrix} 1/10 & 0 & 0 \\ 0 & -1/10 & 0 \\ 0 & 0 & 1/10 \end{bmatrix} \left(\begin{bmatrix} 14 \\ -5 \\ 14 \end{bmatrix} - \begin{bmatrix} 0 & 3 & 1 \\ 2 & 0 & 3 \\ 1 & 3 & 0 \end{bmatrix} \vec{x}_0 \right) \\ &= \begin{bmatrix} 1/10 & 0 & 0 \\ 0 & -1/10 & 0 \\ 0 & 0 & 1/10 \end{bmatrix} \left(\begin{bmatrix} 14 \\ -5 \\ 14 \end{bmatrix} - \begin{bmatrix} 0 & 3 & 1 \\ 2 & 0 & 3 \\ 1 & 3 & 0 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \right) = \begin{bmatrix} 7/5 \\ 1/2 \\ 7/5 \end{bmatrix} \end{aligned}$$

Calculem ara $\vec{x}_2$:

$$\begin{aligned} \vec{x}_2 &= \begin{bmatrix} 1/10 & 0 & 0 \\ 0 & -1/10 & 0 \\ 0 & 0 & 1/10 \end{bmatrix} \left(\begin{bmatrix} 14 \\ -5 \\ 14 \end{bmatrix} - \begin{bmatrix} 0 & 3 & 1 \\ 2 & 0 & 3 \\ 1 & 3 & 0 \end{bmatrix} \vec{x}_1 \right) \\ &= \begin{bmatrix} 1/10 & 0 & 0 \\ 0 & -1/10 & 0 \\ 0 & 0 & 1/10 \end{bmatrix} \left(\begin{bmatrix} 14 \\ -5 \\ 14 \end{bmatrix} - \begin{bmatrix} 0 & 3 & 1 \\ 2 & 0 & 3 \\ 1 & 3 & 0 \end{bmatrix} \begin{bmatrix} 7/5 \\ 1/2 \\ 7/5 \end{bmatrix} \right) = \begin{bmatrix} 111/100 \\ 6/5 \\ 111/100 \end{bmatrix} \end{aligned}$$

Volem continuar fent iteracions per tal de «veure» si el procés és convergent o no. Amb Scilab és més fàcil:

```
-->A=[10 3 1; 2 -10 3; 1 3 10];

-->D=diag([diag(A)])
```

```
 D  =

    10.     0.      0.
    0.    - 10.     0.
    0.      0.      10.

-->L=tril(A)-D
 L  =

    0.     0.     0.
    2.     0.     0.
    1.     3.     0.

-->U=triu(A)-D
 U  =

    0.     3.     1.
    0.     0.     3.
    0.     0.     0.
```

Amb l'ajuda de les funcions **diag**, **tril** i **triu** hem calculat fàcilment les matrius L, D i U.

```
-->F=inv(D)
 F  =

    0.1     0.      0.
    0.    - 0.1     0.
    0.      0.      0.1

-->R=L+U
 R  =

    0.     3.     1.
    2.     0.     3.
    1.     3.     0.

-->x0=[0; 0; 0];

-->x1=F*(b-R*x0)
 x1  =

    1.4
```

```
    0.5
    1.4

-->x2=F*(b-R*x1)
 x2  =

    1.11
    1.2
    1.11

-->x3=F*(b-R*x2)
 x3  =

    0.929
    1.055
    0.929

-->x4=F*(b-R*x3)
 x4  =

    0.9906
    0.9645
    0.9906

-->x5=F*(b-R*x4)
 x5  =

    1.01159
    0.9953
    1.01159

-->x6=F*(b-R*x5)
 x6  =

    1.000251
    1.005795
    1.000251
```

«Veiem» que el procés sembla ser convergent al vector $(1, 1, 1)$. Podeu comprovar que aquest és, de fet, una solució del sistema.

Podem fer el mateix però fent servir un codi de Scilab més eficient (més curt) (amb l'ajuda de la comanda **for**):

```
-->x=[0; 0; 0];

-->for i=1:6
-->x=F*(b-R*x);
-->end;

-->x
 x  =

    1.000251
    1.005795
    1.000251
```

El vector que ens retorna després del bucle és, directament, $\vec{x}_6$. Si volem estar «més segurs» de la convergència, podem aplicar més iteracions:

```
-->x=[0; 0; 0];

-->for i=1:50
-->x=F*(b-R*x);
-->end;

-->x
 x  =

    1.
    1.
    1.
```

3 El mètode de Gauss-Seidel

Aquest mètode és una lleugera modificació del mètode de Jacobi. En la majoria dels casos, el nombre d'iteracions necessari per obtenir una solució aproximada és més menut que per al mètode de Jacobi.

Donat un sistema d'equacions lineals $\mathsf{A}\vec{x} = \vec{b}$ (que satisfaça les hipòtesis establides al començament de la secció), es fa servir la mateixa descomposició de la matriu A explicada en el mètode de Jacobi:

$$\mathsf{A} = \mathsf{L} + \mathsf{D} + \mathsf{U}.$$

Tanmateix, ací reescrivim la igualtat $\mathsf{A}\vec{x} = \vec{b}$ com

$$(\mathsf{L} + \mathsf{D})\vec{x} = \vec{b} - \mathsf{U}\vec{x}. \tag{3}$$

Aquesta és la igualtat crucial del mètode de Gauss-Seidel. Comencem amb un vector inicial qualsevol $\vec{x}_0$ i calculem $\vec{x}_1$ de manera que

$$(\mathsf{L}+\mathsf{D})\vec{x}_1 = \vec{b} - \mathsf{U}\vec{x}_0.$$

Aleshores, calculem $\vec{x}_2$ de manera que

$$(\mathsf{L}+\mathsf{D})\vec{x}_2 = \vec{b} - \mathsf{U}\vec{x}_1,$$

i així successivament: per a cada $k = 0, 1, 2, \ldots$ calculem $\vec{x}_{k+1}$ de manera que

$$(\mathsf{L}+\mathsf{D})\vec{x}_{k+1} = \vec{b} - \mathsf{U}\vec{x}_k \tag{4}$$

Com que la matriu $\mathsf{L}+\mathsf{D}$ és triangular inferior, els components del vector $\vec{x}_{k+1}$ es calculen a partir dels components de $\vec{x}_k$ per **substitució progressiva**. Com en el cas del mètode de Jacobi, si la successió de vectors $\vec{x}_0$, $\vec{x}_1, \ldots$ és convergent, aleshores el vector límit és la solució del sistema. Aclarim aquests fets amb un exemple.

Exemple 3. Aplicarem el mètode de Gauss-Seidel al sistema de l'exemple 2.

$$\mathsf{L}+\mathsf{D} = \begin{bmatrix} 10 & 0 & 0 \\ 2 & -10 & 0 \\ 1 & 3 & 10 \end{bmatrix}, \quad \mathsf{U} = \begin{bmatrix} 0 & 3 & 1 \\ 0 & 0 & 3 \\ 0 & 0 & 0 \end{bmatrix}, \quad \vec{b} = \begin{bmatrix} 14 \\ -5 \\ 14 \end{bmatrix}$$

Comencem també amb el vector inicial $\vec{x}_0 = (0, 0, 0)$. La primera iteració és:

$$\underbrace{\begin{bmatrix} 10 & 0 & 0 \\ 2 & -10 & 0 \\ 1 & 3 & 10 \end{bmatrix}}_{\mathsf{L}+\mathsf{D}} \vec{x}_1 = \underbrace{\begin{bmatrix} 14 \\ -5 \\ 14 \end{bmatrix}}_{\vec{b}} - \underbrace{\begin{bmatrix} 0 & 3 & 1 \\ 0 & 0 & 3 \\ 0 & 0 & 0 \end{bmatrix}}_{\mathsf{U}} \underbrace{\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}}_{\vec{x}_0} \tag{5}$$

Si anomenem

$$\vec{x}_1 = \begin{bmatrix} x \\ y \\ z \end{bmatrix},$$

la igualtat (5) es pot escriure com

$$\underbrace{\begin{bmatrix} 10 & 0 & 0 \\ 2 & -10 & 0 \\ 1 & 3 & 10 \end{bmatrix}}_{\mathsf{L}+\mathsf{D}} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 14 \\ -5 \\ 14 \end{bmatrix}.$$

I aquest és un sistema d'equacions lineals la matriu de coeficients del qual és triangular inferior:

$$\left.\begin{aligned} 10x &= 14 \\ 2x - 10y &= -5 \\ x + 3y + 10z &= 14 \end{aligned}\right\}.$$

Ara, per substitució progressiva obtenim $x = 7/5$, $y = 39/50$ i $z = 513/500$, o siga:

$$\vec{x}_1 = \begin{bmatrix} 7/5 \\ 39/50 \\ 513/500 \end{bmatrix}.$$

La segona iteració és:

$$\underbrace{\begin{bmatrix} 10 & 0 & 0 \\ 2 & -10 & 0 \\ 1 & 3 & 10 \end{bmatrix}}_{\text{L+D}} \vec{x}_2 = \underbrace{\begin{bmatrix} 14 \\ -5 \\ 14 \end{bmatrix}}_{\vec{b}} - \underbrace{\begin{bmatrix} 0 & 3 & 1 \\ 0 & 0 & 3 \\ 0 & 0 & 0 \end{bmatrix}}_{\text{U}} \underbrace{\begin{bmatrix} 7/5 \\ 39/50 \\ 513/500 \end{bmatrix}}_{\vec{x}_1} \tag{6}$$

Si anomenem ara

$$\vec{x}_2 = \begin{bmatrix} x \\ y \\ z \end{bmatrix}$$

l'anterior igualtat pot escriure's com

$$\underbrace{\begin{bmatrix} 10 & 0 & 0 \\ 2 & -10 & 0 \\ 1 & 3 & 10 \end{bmatrix}}_{\text{L+D}} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 5317/500 \\ -4039/500 \\ 14 \end{bmatrix}.$$

I aquest és un sistema d'equacions lineals la matriu de coeficients del qual és triangular inferior:

$$\left.\begin{array}{rcl} 10x & = & 5317/500 \\ 2x - 10y & = & -4039/500 \\ x + 3y + 10z & = & 14 \end{array}\right\}.$$

Ara, per substitució progressiva, obtenim $x = 5317/5000$, $y = 3189/3125$ i $z = 246879/250000$, és a dir:

$$\vec{x}_2 = \begin{bmatrix} 5317/5000 \\ 3189/3125 \\ 246879/250000 \end{bmatrix}.$$

I així successivament...

Fem-ho amb Scilab. Primer, hem d'**executar** el fitxer *SustitucionProgresiva.sci*, on hem definit una funció, **SustitucionProgresiva**, que calcula la solució **per substitució progressiva** d'un sistema T$\vec{x} = \vec{b}$, on T és una matriu triangular inferior invertible. La sintaxi és `SustitucionProgresiva(T,b)`, on T és la matriu de coeficients i `b` és el vector de termes independents.

```
-->M=L+D; x0=[0; 0; 0];
```

```
-->x1=SustitucionProgresiva(M,b-U*x0)
 x1  =

    1.4
    0.78
    1.026

-->x2=SustitucionProgresiva(M,b-U*x1)
 x2  =

    1.0634
    1.02048
    0.987516

-->x3=SustitucionProgresiva(M,b-U*x2)
 x3  =

    0.9951044
    0.9952757
    1.0019069

-->x4=SustitucionProgresiva(M,b-U*x3)
 x4  =

    1.0012266
    1.0008174
    0.9996321

-->x5=SustitucionProgresiva(M,b-U*x4)
 x5  =

    0.9997916
    0.9998480
    1.0000665

-->x6=SustitucionProgresiva(M,b-U*x5)
 x6  =

    1.000039
    1.0000277
    0.9999878
```

Fent servir la funció **for** podem efectuar més iteracions, si volem:

```
-->x=x0;

-->for(i=1:50) x=SustitucionProgresiva(M,b-U*x); end;

-->x
 x  =

    1.
    1.
    1.
```

4 Criteri de convergència

És possible que, quan s'aplique el mètode de Jacobi o Gauss-Seidel, la successió de vectors obtinguda siga divergent. No obstant això, quan la matriu de coeficients és d'un tipus especial, podem garantir la convergència d'ambdós mètodes.

Definició 1. Una matriu quadrada A és **estrictament diagonalment dominant** si per a totes les files el valor absolut de l'element de la diagonal és més gran que la suma dels valors absoluts de la resta dels elements d'aquesta fila.

Algunes vegades la matriu de coeficients d'un sistema d'equacions lineals no és estrictament diagonalment dominant, però si canviem l'ordre de les equacions i/o canviem l'ordre de les incògnites el nou sistema té una matriu de coeficients que és estrictament diagonalment dominant.

Aquesta és la condició que assegura la convergència dels mètodes de Jacobi i de Gauss-Seidel:

Teorema 1. Si una matriu quadrada A és estrictament diagonalment dominant, aleshores els mètodes de Jacobi i Gauss-Seidel són convergents.

Nota 1: Comprovant que la matriu quadrada A és estrictament diagonalment dominant queden garantides totes les condicions que hem assumit al començament del butlletí: el sistema és compatible determinat perque A té rang màxim i les entrades diagonals no poden ser nul·les.

Nota 2: Si la matriu original A no és estrictament diagonalment dominant però intercanviant l'ordre de les files (és a dir, intercanviant l'ordre de les equacions) i/o de les columnes (és a dir, de les incògnites) aleshores sí que és estrictament diagonalment dominant, pots resoldre el sistema resultant (que equival a l'original) amb els mètodes de Jacobi o Gauss-Seidel perque la seua convergència està garantida.

Nota 3: En cas contrari, és a dir, quan la matriu "no pot ser estrictament diagonalment dominant", els mètodes iteratius poden ser caòtics. Com a exemple, aplica el mètode de Jacobi amb la matriu de coeficients

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