233 Analyse numérique matricielle. Résolution approchée de systèmes linéaires, recherche d'éléments propres. Exemples.

Soit $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$ et $n \in \mathbb{N}^*$. Lorsque non défini, $A \in \mathcal{M}_n(\mathbb{K})$ et $b \in \mathbb{K}^n$.

I| Normes matricielles, rayon spectral et conditionnement

1| Normes matricielles

Définition 1 : Soit $\|\cdot\|$ une norme sur $\mathcal{M}_n(\mathbb{K})$. $\|\cdot\|$ est une norme matricielle si :
$$\forall A, B \in \mathcal{M}_n(\mathbb{K}),\quad \|AB\| \leq \|A\| \cdot \|B\|$$

Définition 2 : Soit $\|\cdot\|$ une norme sur $\mathbb{K}^n$. La norme matricielle subordonnée à $\|\cdot\|$ est l'application :
$$\|\cdot\| : \mathcal{M}_n(\mathbb{K}) \to \mathbb{R}_+,\quad A \mapsto \sup_{x \in \mathbb{K}^n \setminus \{0\}} \frac{\|Ax\|}{\|x\|} = \sup_{\|x\|=1} \|Ax\|$$

Exemple 3 : ∘ Normes subordonnées à $\|\cdot\|_1 ; \|\cdot\|_\infty$: $\forall A \in \mathcal{M}_n(\mathbb{K})$,
$$\|A\|_1 = \max_{j \in [\![1;n]\!]} \sum_{i=1}^{n} |a_{i,j}|$$
$$\|A\|_\infty = \max_{i \in [\![1;n]\!]} \sum_{j=1}^{n} |a_{i,j}|$$
∘ Normes non-subordonné à une norme vectorielle :
$$\|A\|_F = \sqrt{\sum_{i=1}^{n} \sum_{j=1}^{n} |a_{i,j}|^2} = \sqrt{\operatorname{tr}(A^*A)}$$

2| Rayon spectral

Définition 4 : Soit $A \in \mathcal{M}_n(\mathbb{K})$. Le rayon spectral de A : $\rho(A) = \max_{\lambda \in \operatorname{Spec}(A)} |\lambda|$

Exemple 5 : Norme subordonnée à $\|\cdot\|_2$: $\forall A \in \mathcal{M}_n(\mathbb{R})$,
$$\|A\|_2 = \sqrt{\rho(A^*A)} = \sqrt{\rho(AA^*)} = \|A^*\|_2$$

Proposition 6 : Soit $A \in \mathcal{M}_n(\mathbb{K})$ et $\varepsilon > 0$.
Alors : (1) Pour toute norme matricielle $\|\cdot\|$, $\rho(A) \leq \|A\|$
(2) Il existe une norme subordonnée $\|\cdot\|$ telle que $\|A\| \leq \rho(A) + \varepsilon$

Proposition 7 : Soit $A \in \mathcal{M}_n(\mathbb{K})$
Alors : (1) $\rho(A) = \lim_{k \to +\infty} \|A^k\|^{\frac{1}{k}}$
(2) $\lim_{i \to +\infty} A^i = 0$ ssi $\forall x \in \mathbb{K}^n$, $\lim_{i \to +\infty} A^i x = 0$
ssi $\rho(A) < 1$
ssi il existe $\|\cdot\|$ tel que : $\|A\| < 1$

3| Conditionnement

But : Quantifier la sensibilité de la solution x du système $Ax = b$ aux perturbations sur A et b.

Définition 8 : Soit $\|\cdot\|$ norme matricielle subordonnée et $A \in GL_n(\mathbb{K})$. Le conditionnement de A par rapport à $\|\cdot\|$ est :
$$\operatorname{cond}(A) = \|A\| \times \|A^{-1}\|$$

Exemple 9 :
$$\begin{pmatrix} 8 & 6 & 4 & 1 \\ 1 & 4 & 5 & 1 \\ 8 & 4 & 1 & 1 \\ 1 & 4 & 3 & 6 \end{pmatrix} x = \begin{pmatrix} 19 \\ 11 \\ 14 \\ 14 \end{pmatrix} \Rightarrow x = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} \text{ or } \begin{pmatrix} 8 & 6 & 4 & 1 \\ 1 & 4 & 5 & 1 \\ 8 & 4 & 1 & 1 \\ 1 & 4 & 3 & 6 \end{pmatrix} x = \begin{pmatrix} 19{,}01 \\ 11{,}05 \\ 14{,}07 \\ 14{,}05 \end{pmatrix} \Rightarrow x = \begin{pmatrix} -2{,}34 \\ 9{,}74 \\ -4{,}85 \\ -1{,}34 \end{pmatrix}$$
Pour la norme $\|\cdot\|_\infty$, $\operatorname{cond}_\infty(A) = 5367{,}5$. Le système est sensible.

Proposition 10 : Soit $A \in GL_n(\mathbb{K})$ et $b \neq 0$
Alors : (1) Si x et $x + \delta x$ sont solutions de $Ax = b$ et $A(x + \delta x) = b + \delta b$
alors :
$$\frac{\|\delta x\|}{\|x\|} \leq \operatorname{cond}(A) \frac{\|\delta b\|}{\|b\|}$$
(2) Si x et $x + \delta x$ sont solutions de $Ax = b$ et $(A + \delta A)(x + \delta x) = b$,
alors :
$$\frac{\|\delta x\|}{\|x + \delta x\|} \leq \operatorname{cond}(A) \frac{\|\delta A\|}{\|A\|}$$
↳ Ces inégalités sont optimales i.e. il existe δA et x ou δA et b pour lesquels on a égalité.

Proposition 11 : Soit $A \in GL_n(\mathbb{K})$, et $U \in \mathcal{M}_n(\mathbb{K})$ unitaire.
Alors : (1) $\forall \alpha \in \mathbb{K} \setminus \{0\}$, $\operatorname{cond}(A) = \operatorname{cond}(A^{-1}) = \operatorname{cond}(\alpha A)$
(2) Si A est normale, alors $\operatorname{cond}_2(A) = \rho(A)\rho(A^{-1}) = \dfrac{|\max_{\lambda \in \operatorname{Spec}(A)} \lambda|}{|\min_{\lambda \in \operatorname{Spec}(A)} \lambda|}$
(3) $\operatorname{cond}_2(U) = 1$.
$\operatorname{cond}_2(UA) = \operatorname{cond}_2(AU) = \operatorname{cond}_2(A)$.

II| Méthodes directes de résolution de $Ax = b$

0| Méthode de Cramer

Proposition 12 : Soit $A = (C_1 | \dots | C_n) \in GL_n(\mathbb{K})$
Alors : $Ax = b$ ssi $\forall i \in [\![1;n]\!]$, $x_i = \dfrac{\det(C_1 | \dots | C_{i-1} | b | C_{i+1} | \dots | C_n)}{\det(A)}$

Coût 13 : On doit calculer $(n+1)$ déterminants de matrices de taille n.
Chaque calcul de déterminant coûte $n!$ opérations au plus.
Par cette méthode, on a au plus $\boxed{(n+1)! \text{ opérations}}$ à réaliser.

1) Méthode de Gauss

Théorème 14 : Soit $A \in \mathcal{M}_n(\mathbb{K})$

Alors : Il existe $M \in GL_n(\mathbb{K})$ tel que MA est triangulaire supérieure.

Coût 15 : Pour l'algorithme donné en annexe, on a :

$$\boxed{\frac{2n^3}{3} + \frac{3n^2}{2} \text{ opérations}}$$ à réaliser.

2) Méthode par décomposition LU

Théorème 16 : Soit $A \in GL_n(\mathbb{K})$ tel que : $\forall k \in [\![1;n]\!]$, $\Delta^k = \begin{pmatrix} a_{1,1} & - & a_{1,k} \\ | & & | \\ a_{k,1} & - & a_{k,k} \end{pmatrix} \in GL_k(\mathbb{K})$.

Alors :

Il existe un UNIQUE couple $(L;U)$ avec U triangulaire supérieure et L triangulaire inférieure avec des 1 sur sa diagonale telles que : $A = LU$

Coût 17 : Pour l'algorithme donné en annexe, on a :

$$\boxed{\frac{2n^3}{3} + \frac{n^2}{2} \text{ opérations}}$$ à réaliser.

3) Méthode de Cholesky

Théorème 18 : Soit $A \in S_n^{++}(\mathbb{R})$

Alors : Il existe une UNIQUE matrice B triangulaire inférieure telle que sa diagonale est positive et $A = BB^*$

Coût 19 : Pour l'algorithme donné en annexe, on a :

$$\boxed{\frac{n^3}{3} + \frac{3n^2}{2} + n \text{ opérations}}$$ à réaliser.

4) Méthode par factorisation QR

Théorème 20 : Soit $A \in GL_n(\mathbb{R})$

Alors : Il existe un UNIQUE couple (Q,R) avec Q orthogonale et R triangulaire supérieure à diagonale positive tel que $A = QR$.

Coût 21 : (Admis) On a besoin de $\boxed{\frac{4n^3}{3} \text{ opérations}}$.

III) Méthodes itératives de résolution de $Ax = b$

1) Cadre de résolution

Définition 22 : Soit $A \in GL_n(\mathbb{K})$. On appelle décomposition régulière de A tout couple $(M;N)$ tel que $A = M - N$ avec $M \in GL_n(\mathbb{K})$.

Une méthode itérative basée sur la décomposition régulière $(M;N)$ est :

$$\begin{cases} x_0 \in \mathbb{K}^n \\ M x_{k+1} = N x_k + b, \ \forall k \geq 1 \end{cases}$$

La méthode itérative converge si : $\forall x_0 \in \mathbb{K}^n$, $x_k \xrightarrow[k \to +\infty]{} x$ avec $Ax = b$.

Théorème 23 : Une méthode itérative converge ssi $\rho(M^{-1}N) < 1$

Théorème 24 : (du point fixe de Picard) Soit $f : \mathbb{K}^n \to \mathbb{K}^n$ une application telle qu'il existe $k \in [0;1[$ tel que : $\forall u, v \in \mathbb{K}^n$, $\|f(u) - f(v)\| \leq k \|u - v\|$

Alors : f possède un UNIQUE point fixe l et pour toute suite $\begin{cases} u_0 \in \mathbb{K}^n \\ u_{n+1} = f(u_n) \end{cases}$

$u_k \xrightarrow[k \to +\infty]{} l$ et $\forall k \in \mathbb{N}$, $\|u_k - l\| \leq k^n \|u_0 - l\|$

Remarque 25 : En pratique, calculer $\rho(M^{-1}N)$ est difficile. On se contente alors de trouver une norme matricielle telle que $\|M^{-1}N\| < 1$.

Théorème 26 : Soit $A \in H_n^{++}(\mathbb{C})$ et $(M;N)$ une décomposition régulière de A.

Alors : $(M^* + N) \in H_n(\mathbb{C})$

↳ Si de plus $(M^* + N) \in H_n^{++}(\mathbb{C})$, alors $\rho(M^{-1}N) < 1$

2) Méthodes de Jacobi, Gauss-Seidel et de relaxation

Méthode	Décomposition $A = M - N$	Matrice d'itération $M^{-1}N$	Itération
Jacobi	$A = D - (E+F)$	$J = D^{-1}(E+F)$	$D u_{k+1} = (E+F) u_k + b$
Gauss-Seidel	$A = (D-E) - F$	$\mathcal{L}_1 = (D-E)^{-1} F$	$(D-E) u_{k+1} = F u_k + b$
Relaxation	$A = \left(\frac{D}{\omega} - E\right) - \left(\frac{1-\omega}{\omega} D + F\right)$	$\mathcal{L}_\omega = \left(\frac{D}{\omega} - E\right)^{-1}\left(\frac{1-\omega}{\omega} D + F\right)$	$\left(\frac{D}{\omega} - E\right) u_{k+1} = \left(\frac{1-\omega}{\omega} D + F\right) u_k + b$

avec $A = \begin{pmatrix} D & & -F \\ & \ddots & \\ -E & & D \end{pmatrix}$ et $\omega \in \mathbb{R}_+^*$

Coût 27 : Pour l'algorithme donné en annexe, pour les méthodes cités, on a besoin d'au plus $\boxed{\frac{3n^2}{2} \text{ opérations}}$ par itération.

3) Cas des matrices tridiagonales

<u>Théorème 28</u> : Soit $A \in \mathcal{M}_n(\mathbb{K})$ tridiagonale.

<u>Alors</u> : $\rho(L_1) = \rho(J)^2$

<u>Théorème 29</u> : Soit $A \in H_n^{++}(\mathbb{C})$ tridiagonale.

<u>Alors</u> : Les trois méthodes convergent et il existe $\omega_{opt} = \dfrac{2}{1+\sqrt{1+\rho(J)^2}}$ tel que $\rho(\omega_{opt}) = \omega_{opt} - 1$ est minimal.

<u>Problème 30</u> : (du Laplacien) Une discrétisation du problème $-u''(x) = f(x)$ conduit à résoudre $A_n x = b$ avec :

$A_n = n^2 \begin{pmatrix} 2 & -1 & & (0) \\ -1 & \ddots & \ddots & \\ & \ddots & \ddots & -1 \\ (0) & & -1 & 2 \end{pmatrix} \in \mathcal{M}_{n-1}(\mathbb{R})$ de valeurs propres : $Spec(A_n) = \left\{ 4n^2 \sin^2\left(\frac{k\pi}{2n}\right) \mid k \in [\![1;n-1]\!] \right\}$

Soit $k \in \mathbb{N}$, $e_k = x_k - x$ l'erreur à l'itération k et $\varepsilon > 0$.

- Pour la méthode de Jacobi, $\forall k \geq \dfrac{\ln(\varepsilon)}{\ln(\rho(J))} \simeq Cn^2$, $\|e_k\|_2 \leq \varepsilon$
- Pour la méthode de G-S, $\forall k \geq \dfrac{\ln(\varepsilon)}{\ln(\rho(L_1))} \simeq Cn^2$, $\|e_k\|_2 \leq \varepsilon$
- Pour la méthode de relaxation, $\forall k \geq \dfrac{\ln(\varepsilon)}{\ln(\rho(L_\omega))} \simeq Cn$, $\|e_k\|_2 \leq \varepsilon$

4) Méthode du gradient conjugué

<u>Définition 31</u> : Soit $A \in \mathcal{M}_n(\mathbb{R})$ et $r_0 \in \mathbb{R}^n$ et $k \in \mathbb{N}$. Le <u>$k^{ième}$ espace de Krylov</u> :

$$K_k = Vect\{r_0 ; Ar_0 ; \ldots ; A^k r_0\}$$

Soit $x_0 \in \mathbb{R}^n$. On défini : $[x_0 + K_k] = \{x \in \mathbb{R}^n \mid (x - x_0) \in K_k\}$

On donne deux définitions de la suite $(x_k)_{k\in\mathbb{N}}$:

$1^{ère}$ définition : On choisit $x_{k+1} \in [x_0 + K_k]$ tel que $r_{k+1} := b - Ax_k \perp K_k$

$2^{ème}$ définition : On choisit x_{k+1} comme le minimum de $f(x) = \frac{1}{2}\langle Ax ; x\rangle - \langle b ; x\rangle$.

<u>Théorème 32</u> : Soit $A \in S_n^{++}(\mathbb{R})$

<u>Alors</u> : À chaque itération k, il existe un UNIQUE et même x_k. De plus ces deux méthodes convergent vers x tel que $Ax = b$ en au plus n itérations.

<u>Coût 33</u> : (Admis) Pour chaque itération, on a besoin de $\boxed{n^3 \text{ opérations}}$

IV) Calcul d'éléments propres

1) Méthode de la puissance

<u>Théorème 34</u> : Soit $A \in \mathcal{M}_n(\mathbb{K})$ diagonalisable de valeurs propres $(\lambda_1 ; \ldots ; \lambda_n)$ telles que $\forall k \in [\![1;n-1]\!]$, $|\lambda_k| < \lambda_n$ et $(e_1 ; \ldots ; e_n)$ des vecteurs propres associés.

Soit $x_0 \in \mathbb{K}^n$ tel que $x_0 = \sum_{i=1}^{n} \beta_i e_i$ avec $\beta_n \neq 0$ et $y_0 = Ax_0$.

<u>Alors</u> : La <u>méthode de la puissance</u> définie par : $\begin{cases} y_k = Ax_{k-1} \\ x_k = \dfrac{y_k}{\|y_k\|} \end{cases}$ converge et $\lim\limits_{k\to+\infty} \|y_k\| = \lambda_n$ et $\lim\limits_{k\to+\infty} x_k = \pm e_n$

<u>Remarque 35</u> : Autrement dit, $\|A^k x_0\| \xrightarrow[k\to+\infty]{} \lambda_n$ et $\dfrac{A^k x_0}{\|A^k x_0\|} \xrightarrow[k\to+\infty]{} \pm e_n$.

La méthode de la puissance donne la plus grande valeur propre en module et un vecteur propre associé.

<u>Remarque 36</u> : En considérant A^{-1} au lieu de A, on peut trouver la plus petite valeur propre en module et un vecteur propre associé. C'est la méthode de la puissance inverse.

2) Méthode de Jacobi

<u>Définition 37</u> : Soit $\theta \in \mathbb{R}$ et $(p;q) \in [\![1;n]\!]$ tels que $p \neq q$. La <u>matrice de Givens</u> :

$$Q(p;q;\theta) = \begin{array}{c} \\ \\ \\ \\ \\ \\ \\ \\ \end{array} \begin{pmatrix} 1 & & & & & & (0) \\ & 1 & & & & & \\ & & \cos(\theta) & - & \sin(\theta) & & \\ & & | & 1 & | & & \\ & & -\sin(\theta) & - & \cos(\theta) & & \\ & & & & & 1 & \\ (0) & & & & & & 1 \end{pmatrix} \begin{array}{l} \\ \\ \leftarrow p \\ \\ \leftarrow q \\ \\ \\ \end{array}$$

$$\begin{array}{ccc} \uparrow & & \uparrow \\ p & & q \end{array}$$

<u>Définition 38</u> : Soit $A \in S_n(\mathbb{R})$. La <u>méthode de Jacobi</u> défini $\begin{cases} A_1 = A \\ A_{k+1} = Q(p_k;q_k;\theta_k)^t A_k Q(p_k;q_k;\theta_k) \end{cases}$ tels que :

(1) $(p_k;q_k)_{k\in\mathbb{N}} \in [\![1;n]\!]^{2\mathbb{N}}$ tels que : $|a^k_{p_k,q_k}| = \max\limits_{i\neq j} |a^k_{i,j}|$, $\forall k \in \mathbb{N}$.

(2) $(\theta_k)_{k\in\mathbb{N}} \in \mathbb{R}^{\mathbb{N}}$ tel que : $a^{k+1}_{p_k,q_k} = 0$

<u>Théorème 39</u> : Soit $A \in S_n(\mathbb{R})$ de valeurs propres $(\lambda_1 ; \ldots ; \lambda_n)$ et $(A_k)_{k\in\mathbb{N}}$ définies par la méthode de Jacobi.

<u>Alors</u> : $\lim\limits_{k\to+\infty} A^k = diag(\lambda_{\sigma(i)})$ avec $\sigma \in S_n$

ANNEXE

Algorithme méthode de Gauss

entrée : A, b
sortie : X tel que $AX=b$

$A = (A|b)$
pour k dans $[\![1;n]\!]$:
  si $a_{k,k} = 0$
    si $\exists i > k \setminus a_{i,i} \neq 0$:
      $L_i \leftrightarrow L_k$
    sinon :
      "La matrice A est singulière"
  $L_k \leftarrow \frac{L_k}{a_{k,k}}$
  pour i dans $[\![k+1;n]\!]$:
    $L_i \leftarrow L_i - a_{i,k} L_k$
pour i dans $[\![n-1;1]\!]$:
  $a_{i,n+1} = a_{i,n+1} - \sum_{j=i+1}^{n} a_{i,j}\, a_{j,n+1}$
rendre $A[:;n+1]$

Algorithme méthode par décomposition LU

entrée : A
sortie : A contenant L et U (sans la diagonale de L)

pour k dans $[\![1;n-1]\!]$:
  pour i dans $[\![k+1;n]\!]$:
    $a_{i,k} = \frac{a_{i,k}}{a_{k,k}}$
    pour j dans $[\![k+1;n]\!]$:
      $a_{i,j} = a_{i,j} - a_{i,k}\, a_{k,j}$
rendre A

Algorithme méthode de Cholesky

entrée : A
sortie : A contenant B

pour j dans $[\![1;n]\!]$:
  pour k dans $[\![1;j-1]\!]$:
    $a_{j,j} = a_{j,j} - a_{j,k}^2$
  $a_{j,j} = \sqrt{a_{j,j}}$
  pour i dans $[\![j+1;n]\!]$:
    pour k dans $[\![1;j-1]\!]$:
      $a_{i,j} = a_{i,j} - a_{j,k}\, a_{i,k}$
    $a_{i,j} = \frac{a_{i,j}}{a_{j,j}}$
rendre A

Algorithme des méthodes itératives

entrée : $A = \Pi - N$, b, ε
sortie : x approximation de la solution de $Ax=b$ à ε près

$x \in \mathbb{R}^n$
$r = b - Ax$
tant que $\|r\|_2^2 > \varepsilon$:
  calculer $y \in \mathbb{R}^n$ solution de $\Pi y = r$
  $x = x + y$
  $r = r - Ay$
rendre x