Exercices

Fondamentaux d'algèbre linéaire (suite)

Exercice 1 (norme 2) : Soit A une matrice de $M(n,n)$ à coefficients complexes. On rappelle le produit Hermitien sur $\mathbb{C}^n$:

$$(\mathbf{x},\mathbf{y})_H = \sum_{i=1}^{n} x_i \bar{y}_i$$

1. (a) Montrer que la matrice A^*A est Hermitienne positive (i.e. $(A\mathbf{x},\mathbf{x}) \geq 0$ pour tout $\mathbf{x} \in \mathbb{C}^n$).
   (b) Déduire que A^*A est diagonalisable dans une base de vecteurs propres orthonormaux et que les valeurs propres sont positives.
2. (a) Montrer que $\|A\|_2 = \sqrt{\rho(A^*A)}$. En déduire que si A est Hermitienne, alors $\|A\| = \rho(A)$.
   (b) Si on suppose qu'en plus d'être Hermitienne A est inversible, montrer que $\mathrm{cond}_2(A) = \frac{\lambda_n}{\lambda_1}$ où les λ_i sont les valeurs propres de A ordonnées de sorte que $\lambda_1 \leq \cdots \leq \lambda_n$.
3. (Bonus) Que vaut la norme 2 d'une matrice unitaire ? Et son conditionnement ?

Exercice 2 (Norme matricielle et rayon spectrale) : Soit A une matrice et $\rho(A)$ son rayon spectrale.

1. Montrer que pour toute norme on a $\|A\| \geq \rho(A)$.
2. Soit $\epsilon > 0$ et A une matrice fixée. Posons D^δ la matrice diagonale t.q. $D^\delta_{ii} = \delta^{i-1}$ pour $i \in [\![1,n]\!]$. Le but est de montrer qu'on peut déterminer une norme matricielle (dépendant de A et ϵ) t.q. $\|A\| \leq \rho(A) + \epsilon$.
   (a) Avec la *décomposition de Schur* de $A = UTU^*$, on pose $T^\delta = \left[D^\delta\right]^{-1} T D^\delta$. Montrer qu'on peut choisir δ t.q.
   $$\|T^\delta\|_\infty \leq \rho(A) + \epsilon$$
   (b) Considérons la norme vectorielle $\|\mathbf{x}\| = \|(UD^\delta)^{-1}\mathbf{x}\|_\infty$. Justifier que la norme matricielle subordonnée vérifie $\|A\| \leq \rho(A) + \epsilon$.
   (c) Déduire enfin l'équivalence $\rho(A) < 1$ et il existe une norme $\|\cdot\|$ t.q. $\|A\| < 1$.

Les méthodes itératives

Exercice 3 (D.L. de matrice) : Soit A une matrice et f une fonction analytique, c'est à dire qu'elle peut s'écrire comme une série :

$$f(z) = \sum_{i=0}^{+\infty} a_i z^i \quad \forall z \in \mathbb{C} \ \text{ t.q.} |z| \leq R$$

où $R > 0$ est le rayon de convergence de la série. On définit alors $f(A)$ comme suit :

$$f(A) = \sum_{i=1}^{+\infty} a_i A^i$$

1. (a) Montrer que la série de matrice converge bien si $\rho(A) < R$.
   (b) Justifier que e^A est toujours bien définie. Donner son expression dans le cas d'une matrice diagonalisable.
2. Considérons A et M deux matrices inversibles et posons N t.q. $A = M - N$:
   (a) À l'aide d'un D.L., exprimer A^{-1} en fonction de M et N. À quelle condition la série obtenue converge ?
   (b) Quel est le lien entre le D.L. obtenu et les *méthodes de splitting* ?
   (c) (Bonus) Quel comparaison peut on faire avec les méthodes directes ?

Exercice 4 (Méthode de Richardson) : Soit A une matrice inversible. On cherche à résoudre le système linéaire $A\mathbf{x} = \mathbf{b}$ à l'aide de la méthode itérative suivante :

$$\mathbf{x}^{(k+1)} = \mathbf{x}^{(k)} + \alpha(A\mathbf{x}^k - \mathbf{b})$$

où α est un paramètre réel.

1. (a) Montrer qu'il s'agit d'une *méthode de splitting* en précisant les matrices M et N.

   (b) Montrer que la méthode est convergence ssi

   $$0 \leq 1 + 2\alpha \mathrm{Re}(\lambda) + \alpha^2|\lambda|^2 < 1$$

   pour tout λ valeur propre de A. Indication : Commencer par montrer que λ v.p. de A ssi $\mu = 1 + \alpha\lambda$ v.p. de $M^{-1}N$.

   (c) Déduire que si toutes les v.p. λ de A vérifient $\mathrm{Re}(\lambda) > 0$ ou $\mathrm{Re}(\lambda) < 0$, alors on peut toujours trouver α garantissant la convergence. Montrer également que c'est une condition nécessaire pour garantir la convergence.

2. Supposons maintenant A définie positive et considérons l'algorithme à pas variable :

   $$\mathbf{x}^{(k+1)} = \mathbf{x}^{(k)} + \alpha_k(A\mathbf{x}^k - \mathbf{b})$$

   On notera $\|\mathbf{x}\|_A^2 = (A\mathbf{x}, \mathbf{x})$.

   (a) Montrer que l'algorithme ci-dessus correspond à la minimisation de la fonction $J(\mathbf{x}) = \|\mathbf{x} - \mathbf{x}^*\|_A$ où $\mathbf{x}^*$ est la solution de $A\mathbf{x} = \mathbf{b}$.

   (b) Déterminer α_k minimisant $\|\mathbf{x}^{k+1} - \mathbf{x}^*\|_A$.

   (c) Montrer que pour ce coefficient α_k, on a :

   $$\|\mathbf{e}^{k+1}\|_A^2 = \|\mathbf{e}^k\|_A^2 \left(1 - \frac{\|\mathbf{r}^k\|^4}{\|\mathbf{e}^k\|_A^2 \|\mathbf{r}^k\|_A^2}\right) = \|\mathbf{e}^k\|_A^2 \left(1 - \frac{1}{\|\mathbf{y}^k\|_A^2 \|\mathbf{y}^k\|_{A^{-1}}^2}\right)$$

   où $\mathbf{y}^k$ vérifie $\|\mathbf{y}^k\| = 1$, $\mathbf{e}^k = \mathbf{x}^k - \mathbf{x}^*$ et $\mathbf{r}^k = A\mathbf{x}^k - \mathbf{b}$.

   (d) À l'aide de l'*inégalité de Kantorovitch* (admise) :

   $$\|\mathbf{y}^k\|_A^2 \|\mathbf{y}^k\|_{A^{-1}}^2 \leq \frac{1}{4}\left(\sqrt{\frac{\lambda_{min}}{\lambda_{max}}} + \sqrt{\frac{\lambda_{max}}{\lambda_{min}}}\right)^2$$

   où $0 < \lambda_{min} < \lambda_{max}$ sont la plus petite et la plus grande valeur propre de A, déduire :

   $$\|\mathbf{e}^{k+1}\|_A \leq \|\mathbf{e}^k\|_A \left(\frac{\kappa_2 - 1}{\kappa_2 + 1}\right)$$

   avec κ_2 le conditionnement 2 de A.

Exercice 5 (CG vs descente) : Considérons le système linéaire $A\mathbf{x} = b$ à résoudre où :

$$A = \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix} \quad \text{et} \quad \mathbf{b} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

Partant de $\mathbf{x_0} = [1,\ 1]^t$, calculer les deux premières itérations données :

1. la *méthode de descente à pas optimale*
2. la *méthode du gradient conjugué*

Que constatez-vous ? Interpréter géométriquement les deux méthodes avec les courbes de niveau de $(A\mathbf{x}, \mathbf{x})$.

Exercice 6 (CG préconditionné) : Soit A une *matrice symétrique définie positive*. On souhaite résoudre à l'aide du gradient conjugué le système linéaire $A\mathbf{x} = \mathbf{b}$. Afin d'accélérer la vitesse de l'algorithme, on se propose de préconditionner le système par la matrice P symétrique définie positive.

1. À l'aide de la *décomposition de Cholesky* de la matrice P, déduire une façon de préconditionner le système linéaire initiale pour pouvoir toujours appliquer le CG au système préconditionné (Indication : Penser à poser $\widetilde{A} = G^t A G$).
2. Écrire l'algorithme CG pour le système préconditionné. Comment pouvez-vous l'optimiser ?

Exercices de révision

Exercice 1 (Série de Neumann) : Soit A une matrice de rayon spectral $\rho(A) < 1$.

1. (a) Montrer l'équivalence : B inversible $\Leftrightarrow$ toutes les valeurs propres de B sont non nulles.
   (b) Déduire que la matrice $I - A$ est inversible.
2. Montrer le résultat suivant :
$$(I - A)^{-1} = \sum_{i=0}^{+\infty} A^i$$

Exercice 2 (critère d'arrêt) : Considérons une méthode de splitting $A = M - N$ où M est choisie de sorte que $\rho(M^{-1}N) < 1$ et de sorte que $M^{-1}N$ soit *symétrique*. Afin de stopper l'algorithme, on considère le test d'arrêt suivant :
$$\|\mathbf{x}^{k+1} - \mathbf{x}^k\|_2 \leq \epsilon$$
où $\epsilon > 0$ est une tolérance fixée.

1. Montrer que :
$$\|\mathbf{x}^k - \mathbf{x}\|_2 \leq \|\mathbf{x}^{k+1} - \mathbf{x}\|_2 + \|\mathbf{x}^{k+1} - \mathbf{x}^k\|_2$$
2. Déduire alors que :
$$\|\mathbf{x}^k - \mathbf{x}\|_2 \leq \frac{1}{1 - \rho(M^{-1}N)} \|\mathbf{x}^{k+1} - \mathbf{x}^k\|_2$$
   Dans quelle situation le choix de ce critère d'arrêt vous semble-t-il pertinent ?

Exercice 3 (méthode de relaxation) : Soit A et M deux matrices inversibles. On pose N t.q. $A = M - N$ et on considère la *méthode de splitting* :
$$\mathbf{x}^{k+1} = M^{-1}N\mathbf{x}^k + M^{-1}\mathbf{b}$$
pour résoudre le système linéaire $A\mathbf{x} = \mathbf{b}$. L'idée des *méthodes de relaxation* consiste, à partir d'une méthode itérative, à calculer les itéres comme suit :
$$\widetilde{\mathbf{x}}^{k+1} = (1 - \omega)\widetilde{\mathbf{x}}^k + \omega\mathbf{x}^{k+1}$$
où $\mathbf{x}^{k+1}$ est donné par méthode ci-dessus.

1. (a) Montrer que la *méthode de relaxation* revient à une *méthode de splitting* où $A = \widetilde{M} - \widetilde{N}$ en précisant $\widetilde{M}$ et $\widetilde{N}$.
   (b) Montrer que si pour toute valeur propre λ de $M^{-1}N$ on a $\mathrm{Re}(\lambda) < 1$, alors il existe toujours une valeur de ω t.q. la méthode relaxée converge.
2. Considérons A une *matrice Hermitienne définie positive* et le cas de la *méthode de Jacobi relaxé* :
   (a) Montrer que pour tout élément propre $(\lambda, \mathbf{v}) \in \mathbb{C} \times \mathbb{C}^n$ de $M^{-1}N$ on a :
$$(\lambda - 1)D\mathbf{v} = -A\mathbf{v}$$
   (b) Déduire la convergence de la méthode relaxée.
   (c) (Bonus) Retrouver ce résultat à l'aide du résultat établi à l'exercice 3 de la section précédente.

Exercice 4 (Analyse erreur CG) : On suppose qu'on souhaite résoudre à l'aide du Gradient Conjugué un système linéaire $A\mathbf{x} = \mathbf{b}$ où A est supposée symétrique définie positive. On rappelle que $\mathbf{r}_p = A\mathbf{x}_p - \mathbf{b}$ est le résidu à l'étape p et on note $\mathbf{e}_p = \mathbf{x}_p - \mathbf{x}$ l'erreur à l'étape p.

1. (a) Montrer qu'à l'étape p de l'algorithme, on a :
$$\mathbf{x}_p - \mathbf{x} = \mathbf{e}_0 \sum_{i=1}^{p} a_i A^i \mathbf{e}_0 = Q(A)\mathbf{e}_0$$
   où $Q(\cdot)$ est une polynôme de degré p vérifiant $Q(0) = 1$.
   (b) Déduire ainsi :
$$\|\mathbf{e}_p\|_A = \min_{\mathbf{x}_p \in \mathbf{x}_0 + K(p)} \|\mathbf{x}_p - \mathbf{x}\|_A = \min_{Q \in \mathbb{P}_p^0} \|Q(A)\mathbf{e}_0\|_A$$
   où $\mathbb{P}_p^0$ est l'espace des polynôme de degré p vérifiant $P(0) = 1$.
2. (a) Montrer que :
$$\|\mathbf{e}_p\|_A^2 = \min_{Q \in \mathbb{P}_p^0} \sum_{i=1}^{n} \lambda_i |Q(\lambda_i)(\mathbf{e}_0)_i|^2$$
   où les $(\lambda_i)_{i=1,n}$ sont les n valeurs propres.
   (b) Retrouver ainsi le nombre maximal d'itérations que peut effectuer la méthode avant d'avoir $\mathbf{e}_p = \mathbf{0}$.

Applications et programmation

Pour cette section, je vous invite à faire les programmes en *Python* à l'aide des modules *NumPy* et *MatPlotLib*. L'avantage est que ces outils sont gratuits et multiplate-forme (et sont ceux utilisés en TP !). Il est également possible de faire ces programmes en *Matlab* (payant), *Scilab* (gratuit et aussi multiplate-forme), ou même en C / C++ (les outils de visualisation étant moins facile à mettre en place).

Exercice 1 (Jacobi, Gauss-Seidel et Relaxation) : On considère la résolution du système linéaire obtenu en discrétisant :

$$\begin{cases} -\Delta u + u = f \\ u(0) = u(1) = 0 \end{cases} \quad \leftrightarrow \quad \begin{cases} \dfrac{-u_{i-1} + 2u_i - u_{i+1}}{\Delta x^2} + u_i = f(x_i) \quad i \in \{1, \cdots, n\} \\ u_0 = u_{n+1} = 0 \end{cases}$$

où $x_i = i\Delta x$ et $\Delta x = \frac{1}{n+1}$. L'objectif est de comparer l'efficacité de différentes méthodes de splitting.

1. Quelles sont les propriétés de la matrice A issues du système linéaire à résoudre ?
2. Implémenter la fonction *Jacobi* en exploitant les spécificités de A pour obtenir un code efficace.
3. De même, implémenter la fonction *GaussSeidel* en ajoutant un paramètre permettant d'effectuer la relaxation (cf ex. 3).
4. (Application) À l'aide du programme principale, tester les deux méthodes et étudier l'influence du paramètre de relaxation. Que se passe-t-il lorsque n devient grand ?

Exercice 2 (Fitting de trajectoire) : On considère que l'on mesure la trajectoire (plane) d'un objet soumis uniquement à la gravité en n position $(x_i, y_i)_{i=1,n}$. On supposera ces mesures bruités, c'est à dire qu'elles ne sont pas exactes. L'objet n'étant soumis qu'à la gravité, on sait par les lois de Newton que sa trajectoire $(x(t), y(t))$ au cours du temps vérifie les équations suivantes :

$$x(t) = v_x^0 t + x_0 \quad \text{et} \quad y(t) = v_y^0 t + y_0$$

où (x_0, y_0) est la position initiale et (v_x^0, v_y^0) la vitesse initiale. À l'aide des mesures, on souhaite retrouver les paramètres $\mathbf{x} = (x_0, v_x^0, y_0, v_y^0)$.

1. (a) À l'aide d'une matrice A de dimension $2n \times 4$, formuler le problème ci-dessus sous la forme suivante :

   $$\min_{\mathbf{x} \in \mathbb{R}^4} \|A\mathbf{x} - \mathbf{b}\|_2^2$$

   où $\mathbf{b}$ est un vecteur à préciser.

   (b) Proposer une méthode de type *descente de gradient* pour résoudre ce problème.

2. (a) Écrire une fonction *generatePts* qui génère les observations en prenant en paramètre x_0, y_0, v_0^x et v_0^y et un terme *bruit* pour générer un bruit de mesure (à l'aide de la fonction random).

   (b) Implémenter la méthode de descente de gradient et tester la méthode dans un programme principale (tracer la trajectoire obtenue et les points de mesures).

3. (Bonus) Généraliser au cas d'une trajectoire en 3D (cela ne pose pas de difficulté particulière).

Exercice 3 (PCG) : L'objectif ici est de tester la méthode du Gradient Conjugué préconditionné. En particulier, on étudiera l'impact du préconditionneur.

1. Commencer par implémenter la méthode CG dans une fonction *cg*. Vous pouvez la tester à l'aide du système de l'exercice 1 des applications, et comparer ses performances par rapport aux méthodes de splitting.

2. (a) Adapter la méthode pour le cas où on applique un préconditionnement symétrique, comme dans l'exercice 5.

   (b) Tester avec un préconditionnement de type Jacobi $D^{-1/2}AD^{-1/2}$ où D est la diagonale de A.

   (c) Tester avec le préconditionnement SSOR proposé dans l'exercice 5.