Chapitre II. Normes matricielles. Conditionnement

M. Granger

Table des matières

1 Réduction des matrices

1.1 Rappel sur les valeurs propres et les vecteurs propres

Définition 1.1 .- *Soit $f : E \longrightarrow E$ un endomorphisme de E. On dit que $\lambda \in \mathbb{K}$ est une valeur propre de f et v un vecteur propre associé à λ si :*

$$v \neq 0 \ \textit{et} \ f(v) = \lambda.v$$

On représente f sur une base $\mathcal{B}$ (**la même à la source et au but !**) par une matrice carrée $A = Mat_{\mathcal{B},\mathcal{B}}(f)$. Si $V = [v]_{\mathcal{B}}$ on écrit $A.V = \lambda V$ et on dit que λ est une valeur propre de A

Les valeurs propres de f (ou de A) sont exactement les racines dans $\mathbb{K}$ du polynôme caractéristique de f, ou de sa matrice A qui est le polynôme de degré n :

$$P(X) = det(XI - A) = \begin{vmatrix} X - a_{1,1} & & \dots & & -a_{1,n} \\ & \dots & \dots & -a_{k,l} & \\ & \dots & X - a_{i,i} & \dots & \\ & \dots & & \dots & \\ a_{n,1} & & \dots & & X - a_{n,n} \end{vmatrix}$$

P est un polynôme unitaire qui s'écrit : $X^n - \alpha_1 X^{n-1} + \ldots + (-1)^i \alpha_i X^{n-i} + \ldots + (-1)^n \alpha_n$
avec $\alpha_1 = tr(A) = \sum_{i=1}^n a_{i,i}$, et $\alpha_n = detA$.

Si $\mathbb{K} = \mathbb{C}$, il résulte du théorème de D'Alembert-Gauss, que P a n racines distinctes ou confondues : $\lambda_1, \ldots, \lambda_n$. On note $spec(A) = \lambda_1, \ldots, \lambda_n)$.
Rappelons aussi qu'on sait que deux matrices semblables ont le même déterminant, donc le même

polynôme caractéristique $P_{U^{-1}AU} = det(XI - U^{-1}AU) = det(U^{-1}(XI - A)U) = P_A$ et par conséquent le même spectre.

1.2 Produit scalaire ou hermitien, adjoint d'un endomorphisme.

On considère un espace vectoriel E sur $\mathbb{K}$, et une application :

$$< .,. >: E \times E \longrightarrow \mathbb{K}.$$

Définition 1.2 .- *Cas où $\mathbb{K} = \mathbb{R}$. On dit que $(x, y) \longrightarrow < x, y >$ est un* produit scalaire *sur E si et seulement si :*

1. *C'est une application bilinéaire.*
2. $\forall x \in E, < x, x > \geq 0$
3. $\forall x \in E, < x, x >= 0 \Leftrightarrow x = 0$

Définition 1.3 .- *Cas où $\mathbb{K} = \mathbb{C}$. On dit que $(x, y) \longrightarrow < x, y >$ est un* produit scalaire hermitien *sur E si et seulement si :*

1. *C'est une application sesquilinéaire, c'est à dire :*
   $< \lambda_1 x_1 + \lambda_2 x_2, y >= \lambda_1 < x_1, y > + \lambda_2 < x_2, y >$
   et $< x, y >= \overline{< y, x >}$ *où $z \mapsto \overline{z}$ désigne la conjugaison des nombres complexes.*
2. $\forall x \in E, < x, x > \geq 0$
3. $\forall x \in E, < x, x >= 0 \Leftrightarrow x = 0$

N.B. En utilisant les deux premières propriétés on voit que la sesquilinéarité entraine aussi :

$$< x, \lambda_1 y_1 + \lambda_2 y_2 >= \overline{\lambda_1} < x, y_1 > + \overline{\lambda_2} < x, y_2 >$$

Proposition-définition 1.4 .- *On dit qu'une base $v_1, \dots, v_n$ d'un espace euclidien ou hermitien, c'est à dire muni d'un produit scalaire ou scalaire hermitien est orthonormée, si et seulement si :* $\forall i, j, < v_i, v_j >= \begin{cases} 0 & si\ i \neq j \\ 1 & si\ i = j \end{cases}$.

Tout espace euclidien ou hermitien possède des bases orthonormées.

La dernière assertion s'obtient par le procédé d'orthogonalisation de Gram-Schmid.

L'exemple standard : Le produit hermitien canonique sur $\mathbb{C}^n$. On désigne par $v = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ resp. $w = \begin{pmatrix} y_1 \\ \vdots \\ y_m \end{pmatrix}$ deux vecteurs de $\mathbb{C}^n$, et par w^* la matrice ligne $(\overline{y_1}, \dots, \overline{y_n})$ et on définit :

$$< v, w >= \sum x_i \overline{y_i} = w^*.v$$

Noter que le dernier terme présente $< v, w >$ comme le produit matriciel d'une ligne par une colonne.

Notation 1.5 .- *La transconjuguée de la matrice $A \in \mathcal{M}_{m,n}(\mathbb{K})$ est la matrice : $A^* = {}^t\overline{A} \in \mathcal{M}_{n,m}(\mathbb{K})$ obtenue en transposant A et en conjuguant ses coefficients. En d'autres termes : $(A^*)_{i,j} = \overline{A_{j,i}}$ Par exemple :*

$$\begin{pmatrix} 1 & 2 \\ i & 1+i \end{pmatrix}^\star = \overline{\begin{pmatrix} 1 & i \\ 2 & 1+i \end{pmatrix}} = \begin{pmatrix} 1 & -i \\ 2 & 1-i \end{pmatrix}$$

Dans le cas réel le produit scalaire canonique s'exprime de façon similaire :

$$< v, w >= \sum x_i y_i = \ {}^t w.v = \ {}^t v.w$$

Noter que si A est une matrice réelle $A^\star$ est la trnsposée de A.

Proposition-définition.- *Soit $f : E \longrightarrow F$ une application linéaire entre deux espaces vectoriels munis chacun d'un produit scalaire (cas $\mathbb{K} = \mathbb{R}$) ou d'un produit scalaire hermitien (cas $\mathbb{K} = \mathbb{C}$). Il existe une unique application linéaire $f^* : F \longrightarrow E$, appelée l'adjointe de f telle que :*

$$\forall x \in E,\ \forall y \in F, \quad < f(x), y >=< x, f^*(y) >$$

Nous traiterons ne que le cas des produits hermitiens canoniques, en montrant qu'en termes matriciels $f \rightarrow f^\star$ correspond à l'opération de transconjugaison :

Lemme 1.6 .- *Soit $A \in \mathcal{M}_{m,n}(\mathbb{K})$ qu'on identifie à l'application linéaire $\varphi_A :\ \mathbb{K}^n \longrightarrow \mathbb{K}^m$. On munit les espaces de leurs produits scalaires (resp. hermitiens) canoniques. Alors l'application définie par la matrice A^*, est adjointe à φ_A. $(\varphi_A)^\star = (\varphi_{A^\star})$*

Démonstration.- Exercice. □

On a utilisé librement dans cette démonstration deux règles de calcul concernant les adjointes de matrices. Leur démonstration est élémentaire, mais nous allons quand même en donner un énoncé formel en raison de leur importance pratique :

Lemme 1.7 .- *Soit $A \in \mathcal{M}_{m,n}(\mathbb{K})$ et $A \in \mathcal{M}_{p,m}(\mathbb{K})$, alors on a les formules :*

$$(A^*)^* = A \quad {}^t({}^tA)) = A$$

$$(B.A)^* = A^*.B^* \quad {}^tAB = {}^tB^tA$$

Démonstration.- Exercice □

1.3 Réduction des matrices

Rappelons le principe de la réduction des matrices :

On considère une matrice A et l'application linéaire $\varphi_A : \mathbb{K}^n \longrightarrow \mathbb{K}^m$ qu'elle définit par rapport aux bases canoniques respectives $\mathcal{C}a_n$ et $\mathcal{C}a_m$. Par rapport à deux bases quelconques $\mathcal{B}$ et $\mathcal{C}$ avec P et Q pour matrices de passage respectives [1] la matrice de φ_A est

$$B = Q^{-1}AP$$

On dit alors que A et B sont *équivalentes* .

Dans le cas des matrices carrées et si on prend $\mathcal{B} = \mathcal{C}$, et $P = Q$, on dit alors que $B = P^{-1}AP$ est *semblable* à A. Cette notion est, pour deux matrice carrées, distinctes de celle d'équivalence et les deux notions sont envisageables, mais c'est la similitude qui permet de traiter les changements de bases pour les endomorphismes.

[1] Rappelons que $P = P_{\mathcal{C}a_n \rightarrow \mathcal{B}}$, est la matrice dont les colonnes sont les coordonnées des vecteurs de $\mathcal{B}$ dans la base $\mathcal{C}a_n$

Définition 1.8 .- *Soit $A \in \mathcal{M}_{n,n}(\mathbb{K})$ une matrice carrée. On dit que A est* triangulable *si elle est semblable à une matrice triangulaire*

$$T = \begin{pmatrix} t_{1,1} & \cdots & & t_{1,n} \\ & \ddots & t_{i,j} & \vdots \\ & 0 & \ddots & \\ & & & t_{n,n} \end{pmatrix}$$

c'est à dire telle que $t_{i,j} = 0$ pour $i < j$.

On dit que A est diagonalisable si elle est semblable à une matrice diagonale, c'est à dire telle que $t_{i,j} = 0$ pour $i \neq j$.

Remarques Le fait que la matrice de l'endomorphisme φ_A, soit triangulaire (resp. diagonale), dans la base $\mathcal{B} = (e_1, \ldots, e_n)$ équivaut au fait que chaque sous-espace $(e_1, \ldots, e_j)$ est stable par A (resp. que chaque e_i est propre pour A).

Z.- Ne pas confondre diagonale et diagonalisable, triangulaire et triangulable.

Définition 1.9 *Soit $A \in \mathcal{M}_{n,n}(\mathbb{C})$. On dit que A est :*

- *hermitienne ou à symétrie hermitienne si $A^* = A$.*
- *unitaire si $A^*A = AA^* = I$. Ou encore : $A \in GL_n(\mathbb{C})$ et $A^* = A^{-1}$.*
- *normale si $A^*A = AA^*$.*

Soit $A \in \mathcal{M}_{n,n}(\mathbb{R})$. On dit que A est :

- *symétrique si ${}^tA = A$*
- *orthogonale si ${}^tA.A = A\,{}^tA = I$*

Par exemple $\begin{pmatrix} 1 & a \\ a & 2 \end{pmatrix}$ est symétrique, $\begin{pmatrix} 1 & 1+i \\ 1-i & 1 \end{pmatrix}$ est hermitienne, mais $\begin{pmatrix} 1 & i \\ i & 1 \end{pmatrix}$, $\begin{pmatrix} 1 & -i \\ i & i \end{pmatrix}$ ne le sont pas.

Remarques i) Toute matrice unitaire et toute matrice hermitienne est normale.

ii) L'ensemble des matrice unitaires de taille $n \times n$ est un groupe noté $U(n)$. De même l'ensemble des matrices orthogonales, est un groupe noté $O(n)$.

Le groupe $U(n)$ (resp. $O(n)$) est l'ensemble des matrices de changements de bases de $\mathbb{C}^n$ (resp $\mathbb{R}^n$), entre deux bases orthonormées pour le produit scalaire hermitien canonique (resp. le produit scalaire canonique).

iii) Les matrices orthogonales sont les matrices unitaires à coefficients réels. Les matrices symétriques sont les matrices hermitiennes à coefficients réels.

Théorème 1.10 .- *1) Toute matrice carrée réelle ou complexe est triangulable dans une base orthonormée de $\mathbb{C}^n$. Autrement dit :*

$$\exists U \in U(n), \ \ U^{-1}AU = T \ \text{ est triangulaire (supérieure)}$$

(et on a aussi $T = U^\star AU$, puisque $U^\star = U^{-1}$)

2) Toute matrice normale est diagonalisable dans une base orthonormée de $\mathbb{C}^n$. En particulier, toute matrice hermitienne, symétrique, unitaire ou orthogonale, est unitairement diagonalisable.

3) Si A est hermitienne, ses valeurs propres sont réelles, et elle est diagonalisable dans une base orthonormée. Ainsi, il existe $U \in U(n)$, telle que $U^{-1}AU = D$ soit diagonale réelle.

4) Si de plus A est symétrique réelle, elle est diagonalisable dans une base orthonormée de $\mathbb{R}^n$.

Démonstration *du théorème 1.10* .- **1)** Comme tout polynôme à coefficients dans $\mathbb{C}$, le polynôme caractéristique de A possède au moins un racine λ_1, valeur propre de A. Soit $f_1 \in \mathbb{C}^n$ un vecteur propre associé :

$$Af_1 = \lambda_1 f_1$$

qu'on supposera unitaire, c'est à dire $\|f_1\| = 1$. Soit $H = f_1^\perp$ l'hyperplan orthogonal à f_1, et $(\epsilon_2, \ldots, \epsilon_n)$ une base orthonormée de H. On obtient ainsi une base orthonormé $\mathcal{B}_1 = (f_1, \epsilon_2, \ldots, \epsilon_n)$, de $\mathbb{C}^n$. La matrice U_1 de changement de base entre la base canonique (qui est orthonormée !) et la base $\mathcal{B}_1$ est unitaire, et on obtient :

$$U_1^{-1} A U_1 = Mat(\varphi_A)_{\mathcal{B}_1} = \begin{pmatrix} \lambda_1 & a_{1,2} & \ldots & a_{1,n} \\ 0 & & & \\ \vdots & & A' & \\ 0 & & & \end{pmatrix}$$

En fait, si on note $pr : \mathbb{C}^n = \mathbb{C}.f_1 \oplus H \mapsto H$ la projection orthogonale, A' est la matrice de l'application linéaire $\tilde{f} = pr \circ f : H \to H$, dans la base $(\epsilon_2, \ldots, \epsilon_n)$ de H, d'après le calcul suivant :

$$f(\epsilon_j) = (\sum_{i=2}^{n} a'_{i,j} \epsilon_i) + a'_{1,j} f_1 = \tilde{f}(\epsilon_j) + a'_{1,j} f_1$$

Par récurrence sur la dimension, il existe une base orthonormée $(f_2, \ldots, f_n)$ de H, dans laquelle la matrice de $\tilde{f}$ est une matrice triangulaire supérieure T'. Le résultat en découle car :

1. $\mathcal{C} = (f_1, \ldots, f_n)$ est une base orthonormée.
2. $Mat(\varphi_A)_{\mathcal{C}} = \begin{pmatrix} \lambda_1 & a_{1,2} & \ldots & & a_{1,n} \\ 0 & t'_{2,2} & & & t'_{2,n} \\ & & \ddots & t'_{i,j} & \\ & 0 & & \ddots & \\ 0 & & & 0 & t'_{n,n} \end{pmatrix}$ est triangulaire.

2) Si A est normale la matrice triangulaire $T = U^\star AU$ obtenue en 1) est aussi normale. En effet, en utilisant la relation $U^\star U = UU^\star = I$ (U est unitaire), on obtient :

$$T^\star T = U^\star A^\star UU^\star AU = U^\star A^\star IAU = U^\star A^\star AU$$

et de même

$$TT^\star = U^\star AA^\star U$$

ce qui montre l'implication $A^\star A = AA^\star \Rightarrow T^\star T = TT^\star$.

On conclut à l'aide du résultat suivant qui fait l'objet de la question c) de l'exercice 5) déjà évoqué plus haut :

Soit T une matrice carrée. Alors T est triangulaire et normale si et seulement si elle est diagonale.

3) Le résultat du 1) permet d' écrire, sans hypothèse sur A :

$$T = U^\star AU, \text{ et } T^\star = U^\star A^\star U$$

Donc si A est hermitienne ($A = A^\star$), il en est de même de T. Comme $T = T^\star$, elle est la fois triangulaire supérieure comme T, et inférieure comme $T^\star$, donc diagonale et la diagonale est réelle car $T = T^\star$ signifie alors que pour tout i, $T_{i,i} = \overline{T_{i,i}}$.

4) Dans le cas où A est une matrice réelle, ET où A n'a que des valeurs propres réelles (c'est à dire le polynôme caractéristique de A est scindé sur $\mathbb{R}$), on peut dans la démonstration du point 1) supposer que le vecteur propre f_1 est dans $\mathbb{R}^n$. On peut donc refaire la démonstration avec une base orthonormée de $\mathbb{R}^n$ et la matrice de passage obtenue est orthogonale.

Ceci s'applique à une matrice A symétrique réelle : en effet une telle matrice est en particulier hermitienne donc a toute ses valeurs propres réelles selon le point 3). $\square$

Complément : Quotients de Rayleigh des matrices hermitiennes

Soit A une matrice hermitienne. Ses valeurs propres sont réelles, ce qui permet de les ordonner par ordre croissant : $\lambda_1 \leq \ldots \leq \lambda_n$.
Considérons l'application :

$$R_A : \mathbb{C} \setminus \{0\} \to \mathbb{R}$$

définie par $R_A(v) = \frac{v^\star A v}{v^\star v}$. Le fait que R_A soit à valeurs dans $\mathbb{R}$ est dû au fait que A est hermitienne. En effet $v^* v = \|v\|^2$ est réel et $v^\star A v$ est est réel car égal à son conjugué, d'après le calcul suivant : $\overline{v^\star A v} = (v^\star A v)^\star = v^\star A^\star (v^\star)^\star = v^\star A v$.

Puisque A est hermitienne on peut trouver une base orthonormée dans laquelle elle est représentée par la matrice diagonale : $\Delta = diag(\lambda_1, \cdots, \lambda_n)$, et si $(\alpha_1, \cdots, \alpha_n)$ sont les coordonnées d'un vecteur v de $\mathbb{C}^n$ dans cette base on a $v^\star A v = (\alpha_1, \cdots, \alpha_n)\Delta \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix} = \sum_{i=1}^n \lambda_i |\alpha_i|^2$. On en déduit le résultat suivant d'encadrement des quotients de Rayleigh :

Proposition 1.11 *.- Soient A et les λ_i comme ci-dessus et $(v_1, \ldots, v_n)$ une base* orthonormée *de vecteurs propres avec $Av_i = \lambda_i v_i$. Alors :*

1. *$\lambda_k = R_A(v_k)$ et l'image de R_A est l'intervalle $[\lambda_1, \lambda_n]$*
2. *Soit V_k le sous-espace de $\mathbb{C}^n$ engendré par $\{v_1, \ldots, v_k\}$ et $W_k = Vect < v_{k+1}, \ldots, v_n >= V_k^\perp$. On a :*

$$\lambda_k = \sup_{x \in V_k} (R_A(v)) \quad et \quad \lambda_k = \inf_{x \in W_{k-1}} (R_A(v))$$

Démonstration.- Soient $v = (\alpha_1, \ldots, \alpha_n)$ les coordonnées de $v \in \mathbb{C}^n$ dans la base $(v_1, \ldots, v_n)$. Tout découle de la formule due aux calculs pécédant l'énoncé :

$$R_A(v) = \frac{\sum_{i=1}^n \lambda_i |\alpha_i|^2}{\sum_{i=1}^n |\alpha_i|^2}$$

En effet elle implique que $R_A(vect < \{v_i, \ldots, v_j\} >) = [\lambda_i, \lambda_j]$, ce qui donne les deux résultats de l'énoncé.

$\square$

Terminons cette section par une notion qui interviendra en force dans le paragraphe suivant pour les questions de normes :

Définition 1.12 *.- Soit $A \in \mathcal{M}_{n,n}(\mathbb{C})$, et $\lambda_1, \ldots, \lambda_n$, la famille des racines du polynôme caractéristique. Le* rayon spectral *de A est le nombre réel ≥ 0 :*

$$\rho(A) = \max_{1 \leq i \leq n} |\lambda_i|$$

2 Normes, normes matricielles

Les problèmes concrets de l'analyse numérique font intervenir des matrices de grande taille, pour lesquelles on ne peut espérer obtenir que des solutions approchées.

Ceci soulève trois sortes de problèmes : précision du résultat obtenu, temps de calcul nécessaire, et stabilité de la solution obtenue par rapport à des perturbations des données. Il existe des systèmes dits mal conditionnées pour lesquels d'infimes perturbations (dues à des incertitudes de mesures par exemple) provoquent des grands changements dans la solution. On examinera ce problème au chapitre III.

2.1 Rappels de définitions

On suppose connue la notion de norme sur un espace vectoriel E sur $\mathbb{K}$.

Définition 2.1 *On dit que l'application* $|||\bullet||| : \mathcal{M}_{n,n}(\mathbb{K}) \to \mathbb{R}$ *est une Rappelons norme matricielle si :*
- *C'est une norme de l'espace vectoriel* $\mathcal{M}_{n,n}(\mathbb{K})$
- *Quel que soient* $A, B \in \mathcal{M}_{n,n}(\mathbb{K}), |||AB||| \leq |||A|||.|||B|||,$

La deuxième propriété est parfois appelée propriété de sous-multiplicativité. Les normes matricielles les plus importante (les seules en fait qui seront utilisées dans ce cours) sont les normes subordonnées à une norme sur $E = \mathbb{K}^n$. Rappelons cette notion :

Proposition-définition 2.2 *.- Soit* $\| \bullet \|$ *une norme sur* $E = \mathbb{K}^n$*. La norme subordonnée à cette norme est l'application* $\mathcal{L}(E) = \mathcal{M}_{n,n}(\mathbb{K}) \xrightarrow{|||\bullet|||} \mathbb{R}$ *définie par* $|||A||| = \sup_{v \neq 0} \frac{\|Av\|}{\|v\|}$
1) Une norme subordonnéee à une norme sur E *est bien une norme et c'est une norme matricielle.*
2) On a

$$\forall v \in E, \|Av\| \leq |||A|||.\|v\|$$

De plus $|||A|||$ *est le plus petit des réels* $\lambda \in \mathbb{R}$ *tel que* $\forall v \in E, \|Av\| \leq \lambda.\|v\|$

Les démonstrations de ces énoncés seront proposées dans l'exercice 6).

2.2 Les normes usuelles et leurs normes subordonnées

Théorème 2.3 *.- On note* $\|\bullet\|_1$*,* $\|\bullet\|_\infty$ *et* $\|\bullet\|_2$ *les trois normes usuelles sur* $\mathbb{K}^n$*, rappelées ci-dessous, et* $|||\bullet|||_1$*,* $|||\bullet|||_\infty$ *et* $|||\bullet|||_2$ *les normes matricielles subordonnées associées. Soit* $v = (\alpha_1, \ldots, \alpha_n) \in \mathbb{K}^n$*, et* $A = (a_{i,j})_{1 \leq i,j \leq n}$*. Alors ces normes sont données par les formules suivantes :*
1)$\|v\|_1 = \sum_{i=1}^n |\alpha_i| \qquad |||A|||_1 = \max_j(\sum_{i=1}^n |a_{i,j}|)$
2)$\|v\|_\infty = Max_i|\alpha_i| \qquad |||A|||_\infty = \max_i(\sum_{j=1}^n |a_{i,j}|)$
3)$\|v\|_2 = \sum_{i=1}^n |\alpha_i|^2 \qquad |||A|||_2 = \sqrt{\rho(A^\star.A)} = \sqrt{\rho(A.A^\star)} = |||A^\star|||_2$

Les deux premiers points seront proposés en exercice voir l'exercice 7) et on va démontrer ici le troisième.

Remarquons d'abord pour justifier la formule que la matrice $A^\star.A$ est hermitienne (puisque $(A^\star.A)^\star = A^\star(.A^\star)^\star = A^\star.A$) et que toutes ses valeurs propres sont ≥ 0. Soit en effet X un vecteur propre de valeur propre λ pour $A^\star A$.
Alors $\lambda\|X\|^2 = X^\star(\lambda X) = X^\star A^\star.AX = \|AX\|^2 \geq 0$. Par définition on a :

$$|||A|||_2 = \sup_{v \neq 0} \frac{< Av, Av >}{< v, v >} = \sup_{v \neq 0} \frac{< v^\star(A^\star.A)v >}{v^\star.v}$$

On reconnait ici les quotients de Rayleigh $R_{A^\star.A}(v) = \frac{<v^\star(A^\star.A)v>}{v^\star.v}$, qui sont tous ≥ 0, et dont le sup est bien $\rho(A^\star.A)$, selon le paragraphe précédent.

2.3 Normes matricielles et rayon spectral

Nous allons montrer que le rayon spectral d'une matrice A est égal à la borne inférieure des $N(A)$, lorsque N parcourt l'ensemble de toutes les normes subordonnées. Prendre garde aussi au fait que $A \mapsto \rho(A)$ n'est pas une norme (question : caractériser les matrices carrées telles que $\rho(A) = 0$). La caractérisation annoncée de $\rho(A)$, comme une borne inférieure des $N(A)$, est un corollaire immédiat du théorème suivant :

Théorème 2.4 .- *Soit A une matrice carrée complexe de rayon spectral $\rho(A)$. Alors :*
1) Pour toute norme matricielle $|||\bullet|||$ (subordonnée à une norme $\|\bullet\|$), sur E, $\rho(A) \leq |||A|||$.
2) Pour tout $\epsilon > 0$, il existe un norme matricielle subordonnée $|||\bullet|||$, (dépendant de ϵ, ET de A !!) telle que :

$$|||A||| \leq \rho(A) + \epsilon$$

Démonstration.- 1) Soit λ une valeur propre de A de module $|\lambda| = \rho(A)$, et v une vecteur propre associé. De l'égalité $Av = \lambda v$, on tire

$$\rho(A)\|v\| = \|Av\| \leq |||A|||.v\|$$

Donc en simplifiant par $\|v\| \neq 0$: $\rho(A) \leq |||A|||$.
2) Par application du théorème 1.10, on choisit U unitaire telle que $T = U^{-1}AU = (t_{i,j})$ soit une matrice triangulaire. Ainsi $t_{i,j} = 0$ si $i > j$. On note $t_{i,i} = \lambda_i$, ième valeur propre de A. Considérons alors la matrice diagonale $D_\delta = D = diag(1, \delta, \ldots, \delta^{n-1})$.

Petit lemme : La matrice $D^{-1}U^{-1}AUD = D^{-1}TD$ est la matrice triangulaire dont le coefficient à la place (i, j) est $\delta^{j-i}t_{i,j}$.

Ceci est le résutat d'une calcul élémentaire [2], et grâce à la condition "triangulaire", les coefficients non diagonaux de cette dernière matrice sont nuls ou tendent vers zéro quand $\delta \mapsto 0$.

Donc, $\lim\limits_{\delta \mapsto 0} D^{-1}U^{-1}AUD = diag(\lambda_1, \cdots, \lambda_n)$, avec $\lambda_i = t_{i,i}$.
Pour l'une quelconque des trois normes usuelles, on remarque que si Δ est diagonale $|||\Delta||| = \rho(\Delta)$. Considérons alors la norme sur $\mathcal{M}_{n,n}(\mathbb{K})$:

$$A \mapsto N_\delta(A) = |||(UD_\delta)^{-1}AUD_\delta|||_\infty$$

On obtient : $\lim\limits_{\delta \mapsto 0} N_\delta(A) = |||diag(\lambda_1, \cdots, \lambda_n)|||_\infty = \rho(A)$, et par conséquent :

$$\exists \delta > 0, \ N_\delta(A) < \rho(A) + \epsilon$$

Pour compléter, il reste à vérifier que N_δ est subordonnée à une norme sur $\mathbb{K}^n$. Vérifier que $v \mapsto \|(UD_\delta)^{-1}.v\|_\infty$ est une norme et que N_δ est la norme subordonnée. □

Pour conclure nous donnons un corollaire dont on retiendra surtout, pour s'en servir dans les chapitres suivants que la condition numérique $\rho(A) < 1$ sur le rayon spectral de B, gouverne la convergence des suites géométriques matricielles, c'est à dire du type B^k où B est une matrice carrée fixée.

Corollaire 2.5 .- *Soit $A \in \mathcal{M}_{n,n}(\mathbb{K})$. Les quatre conditions suivantes sont équivalentes*

1. $\lim\limits_{k \mapsto \infty} A^k = 0$
2. $\forall v \in \mathbb{K}^n, \ \lim\limits_{k \mapsto \infty} A^k v = 0$
3. $\rho(A) < 1$
4. *Il existe au moins une norme matricielle $\|\bullet\|$, telle que $\|A\| < 1$*

Le démonstration de ce corollaire est quasi immédiate et permet de récapituler tout ce qui précède. Elle est proposée sous forme de l'exercice 10).

[2] Il suffit d'utiliser le fait que l'effet sur A de la multiplication à gauche (resp à droite) par la matrice *diagonale* $diag(d_1, \cdots, d_n)$ est de multiplier la $i^{ème}$ ligne (resp colonne) de A par d_i

3 Questions de conditionnement

3.1 Exemple introductif

Dans ce court chapitre, on se propose d'examiner la sensibilité d'une méthode de résolution :
-aux erreurs d'arrondis
-aux "petites perturbations" des données du problème.

On mesurera cette sensibilité par un nombre appelé conditionnement, qui dépend d'une norme sur $\mathbb{K}^n$. Lorsqu'il est très élevé, cela peut mettre cause la pertinence du résultat numérique obtenu.

On part d'un système $Au = b$, de taille $n \times n$, et on se donne par ailleurs ΔA et Δb, présumés petits par rapport à A et b et tel que $A + \Delta A$ soit aussi inversible. On se propose de comparer les solutions x et y des systèmes :

$$Ax = b \quad \text{ et } \ (A + \Delta A)y = (b + \Delta b)$$

On écrit aussi $y - x = \Delta x$ et on se demande si le taux d'erreur sur x mesuré par $\frac{\|\Delta x\|}{\|x\|}$, s'obtient à partir de $\frac{|||\Delta A|||}{|||A|||} + \frac{\|\Delta b\|}{\|b\|}$ par un "facteur d'amplification" multiplicatif raisonable.

Voici un exemple numérique qui montre les difficultés auxquelles on peut se heurter :

$$\text{Soit} \qquad A = \begin{pmatrix} 10 & 7 & 8 & 7 \\ 7 & 5 & 6 & 5 \\ 8 & 6 & 10 & 9 \\ 7 & 5 & 9 & 10 \end{pmatrix}$$

Alors le système $Ax = b$ admet lorsque $b = \begin{pmatrix} 32 \\ 23 \\ 33 \\ 31 \end{pmatrix}$, la solution $x = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}$, et le système $Ay = b + \Delta b$ admet lorsque $b + \Delta b = \begin{pmatrix} 32,1 \\ 22,9 \\ 33,1 \\ 30,9 \end{pmatrix}$ la solution $y = \begin{pmatrix} 9,2 \\ -12,6 \\ 4,5 \\ -1,1 \end{pmatrix}$.

La première solution s'obtient par vérification immédiate (par addition des coefficients des lignes de A) et la deuxième par un calcul à la main (avec un peu de patience), ou avec une calculette.

On constate sur cet exemple, en choisissant pour simplifier la norme du sup $\|\| = \|\|_\infty$, ce qui ne change pas l'ordre de grandeur des évaluations, par rapport à la norme euclidienne standard :

$$\|\Delta b\| = 0,1 \ \|b\| = 33, \ \|x\| = 1, \ \text{ mais } \|\Delta x\| = \|y - x\| = 13,6$$

$$\text{donc } \frac{\|\Delta b\|}{\|b\|} = 0,1/33 \sim 0,00303 \quad \frac{\|\Delta x\|}{\|x\|} = 13,6.$$

Le taux d'amplification de l'erreur est donc de l'ordre de $13.6/0,00303$, supérieur à 4000.

–> inv(A)
ans =

25.	−41.	10.	−6.
−41.	68.	−17.	10.
10.	−17.	5.	−3.
−6.	10.	−3.	2.

3.2 Définition du conditionnement

On fixe une norme $\|\|$, sur $\mathbb{C}^n$, et on note $|||\bullet|||$, la norme subordonnée sur les matrices carrées.

Définition 3.1 *Le conditionnement de la matrice carrée inversible A, est le nombre $cond(A) = |||A||| \times |||A^{-1}|||$*

Bien sûr le conditionnement dépend de la norme choisie. Le choix par défaut des logiciels est celui de la norme euclidienne, et on notera parfois $cond_2$ le conditionnement associé.

Exemple Soit $A = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}$. On a : $A^{-1} = \begin{pmatrix} 1 & -1 & 1 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{pmatrix}$

Pour la norme du sup, on trouve : $|||A||| = 2$ et $|||A^{-1}||| = 3$, d'où $cond(A) = 6$. Pour la norme euclidienne on trouve $cond(A) \sim 4,04$, en utilisant le point 3 de la proposition 3.2 suivante.

Proposition 3.2 .-
1) Pour toute matrice carrée inversible, $c(A) \geq 1,\ c(A) = c(A^{-1}),\ c(\alpha A) = c(A),$ pour tout $\alpha \in \mathbb{C}$.
2) Si $Q \in O(n)$, on a $c_2(Q) = 1$, et de plus pour toute matrice carrée inversible A, $c_2(A) = c_2(AQ) = c_2(QA)$
3) Pour toute matrice carrée inversible A, $A^\star A$ est diagonalisable de valeurs propres μ_i^2, avec $0 < \mu_1^2 \leq \cdots \leq \mu_n^2$, $\mu_i > 0$, et alors le conditionnement pour la norme euclidienne canonique est

$$c_2(A) = \frac{\mu_n}{\mu_1}.$$

Les propriétés 1) 2) sont des exercices élémentaires sur les normes subordonnées, : voir l'exercice 1).

Voyons 3). La matrice $A^\star A$ est hermitienne donc diagonalisable avec des valeurs propres λ_i réelles. Le fait que pour tout i, $\lambda_i > 0$, a été vu dans la démonstration du théorème 18 du chapitre I (il faut préciser que $\lambda_i \neq 0$ puisque A est inversible). Ceci permet de noter ces valeurs propres $\lambda_i = \mu_i^2$, avec $\mu_i > 0$.

Selon le résultat du point 3) de ce même théorème la norme euclidienne de A, est liée au rayon spectral de $A^\star A$ qui vaut λ_n, par la relation $|||A||| = \sqrt{\rho(A^\star A)} = \mu_n$. De même $|||A^{-1}||| = \sqrt{\rho((A^{-1})^\star A^{-1})} = \sqrt{\rho((AA^\star)^{-1})}$. A priori la matrice $AA^\star$ a pour la même raison que $A^\star A$ un spectre de la forme $(\nu_1^2, \ldots, \nu_n^2)$, avec $0 < \nu_1^2 \leq \ldots \leq \nu_n^2$ et par passage à l'inverse les valeurs propres de $(AA^\star)^{-1} \lesssim$ sont les $\frac{1}{\nu_i^2}$ et donc $\rho((A^{-1})^\star A^{-1}) = \frac{1}{\nu_1}$. La formule sur $c_2(A)$ en résulte parce que pour tout i $\mu_i = \nu_i$, à cause du lemme suivant :

Lemme 3.3 .- *Si A est une matrice carrée inversible, les matrice $A^\star A$ et $AA^\star$ ont le même spectre.*

Démonstration.- On constate la validité de l'égalité $(XI - A^\star A)A^{-1} = A^{-1}(XI - AA^\star)$, en observant que chacun des deux membres de cette égalité est égal à la même matrice polynomiale en X : $XA^{-1}A$. En égalant les déterminants on trouve :

$$det(XI - A^\star A)det(A^{-1}) = det(A^{-1})det(XI - AA^\star)$$

ce qui en simplifiant par le scalaire $det(A^{-1}) = \frac{1}{det(A)}$, fournit l'égalité requise entre les polynômes caractéristiques :

$$P_{A^\star A}(X) = P_{AA^\star}(X)$$

$\square$

A titre d'illustration de la proposition 3.2 vérifier que $c_2 \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} = \sqrt{\frac{3+\sqrt{5}}{3-\sqrt{5}}} \sim 2,61$

C'est l'objet de l'exercice 4).

Dans l'exercice 5) on établit que $c_2(A) = 1 \Leftrightarrow A \in \mathbb{C}.U(n)$, ce qui montre que les matrices unitaires forment exactement l'ensemble des matrices de conditionnement optimal.

4 Estimations d'erreurs

Théorème 4.1 *: Estimation de l'erreur dans un système linéaire. Supposons que $Au = b$ et $Au' = b'$. On note $u' - u = \Delta u$ et $b' - b = \Delta b$. Alors :*

$$\frac{\|\Delta u\|}{\|u\|} \leq cond(A)\frac{\|\Delta b\|}{\|b\|}$$

Si au contraire on modifie A en $A' = A + \Delta A$ avec $b = b'$, donc $(A + \Delta A)u' = b$:

$$\frac{\|\Delta u\|}{\|u'\|} \leq cond(A)\frac{|||\Delta A|||}{|||A|||}, \quad \textit{et aussi}$$

$$\frac{\|\Delta u\|}{\|u\|} \leq cond(A)\frac{|||\Delta A|||}{|||A|||}.(1 + O(|||\Delta A|||))$$

Démonstration 1) Les formules $Au = b$, $Au' = b'$, donnent par soustraction $\Delta b = A\Delta u$. Les relations $\Delta u = A^{-1}\Delta b$ et $b = Au$ conduisent aux majorations :

$$\|\Delta u\| \leq |||A^{-1}|||.\|\Delta b\| \ ; \text{ et } \|b\| \leq |||A|||.\|u\| \ \text{ ou encore } \frac{1}{\|u\|} \leq \frac{|||A|||}{\|b\|}$$

qui entraînent le résultat cherché :

$$\frac{\|\Delta u\|}{\|u\|} \leq |||A^{-1}|||.\|\Delta b\| \times \frac{|||A|||}{\|b\|} = cond(A) \lhd \frac{\|\Delta b\|}{\|b\|}$$

2) Des égalités : $0 = Au - A'u' = A(u - u') + (A - A')u'$, on tire aussitôt $u - u' = -A^{-1}(A - A')u'$ qui fournit la première relation :

$$\|u - u'\| = \|A^{-1}(A - A')u'\| \leq |||A^{-1}|||.|||A - A'|||.\|u'\| = c(A)\frac{|||\Delta A|||}{|||A|||}\|u'\|$$

Pour la deuxième relation on échange les rôles de A et A', ce qui n'altère pas $|||\Delta A|||$:

$$\frac{\|\Delta u\|}{\|u\|} \leq |||A'^{-1}|||.|||\Delta A||| = c(A)\frac{|||\Delta A|||}{|||A|||}.\frac{|||(A + \Delta A)^{-1}|||}{|||A^{-1}|||}$$

et le dernier terme $\frac{|||(A+\Delta A)^{-1}|||}{|||A^{-1}|||}$ tend bien vers 1 quand $|||\Delta A|||$ tend vers zéro. □

Dans l'exercice 6) on montre que si A est normale de spectre $\{\lambda_1, \cdots, \lambda_n\}$, alors $c_2(A) = \frac{\max|\lambda_i|}{\min|\lambda_i|}$.
On vérifie directement ce résultat dans le cas où A est une matrice diagonale. Se ramener à ce cas particulier par une transformation unitaire qui ne modifie pas le conditionnement.

N.B. On se gardera d'étendre cette formule simplfiée à des matrices non normales pour lesquelles il convient de revenir à la formule générale du point 3) de la proposition 3.2.

Dans le cas d'une matrice normale, on peut aussi fournir de façon particulièrement accessible une interprétation géométrique du fait qu'un rapport $\frac{\max|\lambda_i|}{\min|\lambda_i|}$ élevé conduit à un mauvais conditionnement :

Soient b_1 et b_2, des vecteurs propres unitaires, pour les valeurs propres de modules extrêmes $|\lambda_1| > |\lambda_2|$. Ces vecteurs sont orthogonaux puisque A est unitairement diagonalisable. Considérons x et $x' = x + \Delta x$, solutions des deux systèmes :

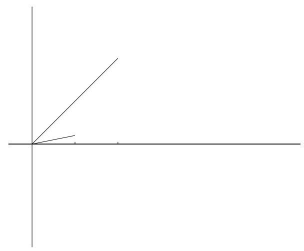
$$Ax = b_1, \quad \text{et } Ax' = b_1 + \epsilon b_2 = b_1 + \Delta b, \quad \epsilon > 0$$

Un calcul immédiat donne $x = \frac{b_1}{\lambda_1}$, et $\Delta x = \epsilon\frac{b_2}{\lambda_2}$. Comme $\epsilon = \|\Delta b\| = \frac{\|\Delta b\|}{\|b\|}$, cela montre l'égalité

$$\frac{\|\Delta x\|}{\|x\|} = \frac{|\epsilon|}{|\lambda_2|}.|\lambda_1| = \frac{|\lambda_1|}{|\lambda_2|}\frac{\|\Delta b\|}{\|b\|}.$$

Ce calcul fait apparaître directement le conditionnement $c(A) = \frac{|\lambda_1|}{|\lambda_2|}$ comme le facteur d'amplification de l'erreur pour deux systèmes ayant des seconds membres arbitrairement proches, $b = b_1,$ et $b + \Delta b = b_1 + \epsilon b_2$, convenablement choisis par rapport aux deux sous espaces propres extêmes.

Voir la figure jointe avec $b_1 = (1, 0),\ \lambda_1 = 0.5,\ \lambda_2 = (0.1)$, et $b_1 + \Delta b = (1, 0.2)$. On trouve $u_1 = (2, 0)$ et $u_2 = (2, 2)$.



$fig.1$