Université Ibn Khaldoun-Tiaret

Département des Mathématiques

Analyse Numérique Matricielle

Cours et Travaux dirigés destinés aux étudiants de première année Master

par A.Larabi

RAPPELS D'ALGEBRE MATRICIELLE:

Cette section contient des rappels sommaires des propriétés classiques nécessaires à la compréhension des méthodes numériques exposées dans les sections suivantes.

Nous nous intéressons uniquement aux opérateurs linéaires de $\mathbb{C}^n$ dans $\mathbb{C}^n$. $\mathbb{C}^n$ est naturellement un espace vectoriel sur $\mathbb{C}$.

Une application linéaire f de $\mathbb{C}^n$ dans $\mathbb{C}^n$ vérifie

$$\forall x, y \in \mathbb{C}^n, \quad f(x+y) = f(x) + f(y),$$

$$\forall \lambda \in \mathbb{C} \text{ et } \forall x \in \mathbb{C}^n, \quad f(\lambda x) = \lambda f(x).$$

Nous noterons $\mathcal{L}(\mathbb{C}^n,\mathbb{C}^n)$ l'ensemble des applications linéaires de $\mathbb{C}^n$ dans $\mathbb{C}^n$.

Si on choisit des bases $\{e_i\}_{1\leq i\leq n}$ et $\{f_i\}_{1\leq i\leq n}$ de $\mathbb{C}^n$, alors l'application $f \in \mathcal{L}(\mathbb{C}^n,\mathbb{C}^n)$ est caractérisée par la matrice A des a_{ij}.

$$f(e_i) = \sum_{j=1}^{n} a_{ji}\, f_j \text{ pour } i = 1, \ldots, n,$$

$$\begin{array}{c c} & \begin{array}{ccc} f(e_1) & \ldots & f(e_n) \end{array} \\ \begin{array}{c} f_1 \\ f_2 \\ \vdots \\ f_n \end{array} & \left(\begin{array}{ccc} a_{11} & \ldots & a_{1n} \\ a_{21} & \ldots & a_{2n} \\ \ldots & \ldots & \ldots \\ a_{n1} & \ldots & a_{nn} \end{array} \right) = A \end{array}$$

A est dite matrice de l'application linéaire f de $\mathcal{L}(\mathbb{C}^n,\mathbb{C}^n)$ rapportée aux bases $\{e_i\}_{1\leq i\leq n}$ et $\{f_i\}_{1\leq i\leq n}$.

Soit $x \in \mathbb{C}^n$. Les composantes y_i de $y \in \mathbb{C}^n$ tel que $y = f(x)$ sont les composantes de y exprimé dans la base des $\{f_i\}_{1\leq i\leq n}$. Les x_i sont les composantes de x exprimé dans la base des $\{e_i\}_{1\leq i\leq n}$.

$$\begin{aligned} x &= \sum_{i=1}^{n} e_i\, x_i, \\ y &= \sum_{i=1}^{n} x_i (\sum_{j=1}^{n} a_{ji}\, f_j) = \sum_{j=1}^{n} y_j f_j. \end{aligned}$$

D'où

$$y_j = \sum_{i=1}^{n} a_{ji} x_j,$$

et par conséquent $y = Ax$.

Les éléments de la matrice A seront toujours notés a_{ij} pour i et $j = 1, \ldots, n$, avec i indice de ligne et j indice de colonne.

Notation
Nous noterons $\overline{a}$ le complexe conjugué de a. Si $a = x + iy$, alors $\overline{a} = x - iy$.

Nous donnons maintenant un ensemble de définitions de différents types de matrices.

Matrice inversible
Une matrice A est dite inversible s'il existe une matrice A^{-1} telle que

$$AA^{-1} = A^{-1}A = I,$$

où I est la matrice unité ou identité.
Naturellement on a

$$\begin{aligned} (A^{-1})^{-1} &= A, \\ (AB)^{-1} &= B^{-1}A^{-1}. \end{aligned}$$

Matrice conjuguée
B est la matrice conjuguée de A si

$$b_{ij} = \overline{a}_{ij} \;\; \forall i \;\text{ et }\; j = 1, \ldots, n.$$

On a

$$(\overline{A})^{-1} = \overline{(A^{-1})}.$$

Matrice transposée
B est la matrice transposée de A si

$$b_{ij} = a_{ji} \;\; \forall i \;\text{ et }\; j = 1, \ldots, n.$$

On utilisera la notation classique $B = A^T$.
Nous avons

$$\begin{aligned} (A^T)^T &= A, \\ (AB)^T &= B^T A^T, \\ (A^T)^{-1} &= (A^{-1})^T. \end{aligned}$$

Matrice symétrique
Une matrice A est dite symétrique si $A = A^T$, c'est-à-dire si

$$a_{ij} = a_{ji} \; > \forall i \;\text{ et }\; j = 1, \ldots, n.$$

Matrice adjointe
B est l'adjointe de A si

$$b_{ij} = \overline{a}_{ji} \;\; \forall i \;\;\text{ et }\;\; j = 1, \ldots, n.$$

On notera A^* l'adjointe de A.
Nous avons

$$\begin{aligned} A^* &= (\overline{A})^T = \overline{(A^T)}, \\ (A^*)^* &= A, \\ (AB)^* &= B^* A^*. \end{aligned}$$

Matrice hermitienne
Une matrice A est dite hermitienne ou auto-adjointe si

$$A = A^*.$$

Par conséquent $a_{ij} = \overline{a}_{ji}$. On remarquera que la diagonale principale de A est réelle.
Matrice unitaire
Une matrice carrée A est dite unitaire si

$$A^* = A^{-1}.$$

Matrice normale
Une matrice A est dite normale si

$$A^* A = A A^*.$$

Matrice orthogonale
Une matrice unitaire à éléments réels est appelée matrice orthogonale.

Définissons le produit hermitien de deux vecteurs de $\mathbb{C}^n$. Soient x et y deux vecteurs de $\mathbb{C}^n$ de composantes x_i et y_i pour $i = 1, \ldots, n$. Le produit hermitien, noté (x, y), est une application de $\mathbb{C}^n \times \mathbb{C}^n$ dans $\mathbb{C}$, définie par

$$(x, y) = \sum_{j=1}^{n} x_j \, \overline{y}_j.$$

Les propriétés du produit hermitien sont les suivantes :

$$\begin{aligned} (u + v, y) &= (u, y) + (v, y) \;\; \forall u, v \;\text{ et } y \in \mathbb{C}^n, \\ (\lambda x, y) &= \lambda (x, y) \;\; \forall \lambda \in \mathbb{C} \;\text{ et } \forall x \;\text{ et } y \in \mathbb{C}^n, \\ (x, \lambda y) &= \overline{\lambda} (x, y) \;\; \forall \lambda \in \mathbb{C} \;\text{ et } \forall x \;\text{ et } y \in \mathbb{C}^n, \\ (x, y) &= \overline{(y, x)} \;\; \forall x \;\text{ et } y \in \mathbb{C}^n, \\ (x, x) &= \sum_{j=1}^{n} \mid x_j \mid^2 \quad \text{pour } x \neq 0. \end{aligned}$$

Cette dernière relation définit la norme euclidienne $\| \, x \, \|_2^2$.

Matrice hermitienne définie positive
Une matrice hermitienne A est définie positive si

$$\forall x \in C^n - \{0\}, \text{ on a } (Ax, x) > 0.$$

Une matrice hermitienne A est semi-définie positive si

$$\forall x \in C^n - \{0\}, \text{ on a } (Ax, x) \geq 0.$$

Définition 0.0.1 *On appelle sous-matrice principale d'ordre k d'une matrice A la matrice constituée des k premières lignes et des k premières colonnes de A. On la notera A_k.*

Propriété 0.0.2 *Les sous-matrices principales d'ordre k d'une matrice A hermitienne définie positive sont hermitiennes définies positives.*
Les éléments diagonaux de A sont strictement positifs.

Démonstration.
Si x est le vecteur $(x_1, \ldots, x_k, 0, ..., 0)$, le vecteur $y = (x_1, \ldots, x_k) \in \mathbb{C}^k$ est tel que

$$(A_k\, y, y) = (Ax, x) > 0.$$

Donc A_k est définie positive. D'autre part, si e_i est le i$^{\text{ème}}$ vecteur de la base canonique (il n'a qu'une seule composante non nulle, la i$^{\text{ème}}$, qui vaut 1), nous avons

$$(Ae_i, e_i) = a_{ii} > 0.$$

□

Matrice triangulaire supérieure (resp. inférieure)
Une matrice A est triangulaire supérieure (resp. inférieure) si

$$a_{ij} = 0 \text{ si } i > j \text{ (resp. } a_{ij} = 0 \text{ si } i < j).$$

Propriété 0.0.3 *L'ensemble des matrices triangulaires supérieures (resp. inférieures) inversibles forme un sous-groupe multiplicatif du groupe des matrices carrées $n \times n$ inversibles.*

Rappelons qu'un ensemble $v_1, \ldots, v_p$ de vecteurs de $\mathbb{C}^n$ (avec $p \leq n$) sont linéairement indépendants si la relation

$$\sum_{j=1}^{p} \alpha_j\, v_j = 0$$

implique nécessairement que α_j est nul, $\forall j = 1, \ldots, p$.

On sait qu'on peut déduire d'un ensemble de vecteurs linéairement indépendants un autre ensemble de vecteurs linéairement indépendants en réalisant de simples combinaisons linéaires. On peut en plus lui imposer des propriétés d'orthogonalité. La propriété suivante donne une méthode permettant de transformer une suite de vecteurs linéairement indépendants en une autre suite orthogonale qui engendre le même espace vectoriel. C'est le procédé d'orthogonalisation de Gram-Schmidt.

Propriété 0.0.4 *A partir d'une suite de vecteurs de* $\mathbb{C}^n$ *linéairement indépendants* $(q_1, \ldots, q_k)$ *on peut construire une suite de vecteurs non nuls* $(p_1, \ldots, p_k)$ *deux à deux orthogonaux telle que*

$$V(q_1, \ldots, q_i) = V(p_1, \ldots, p_i) \ \forall i \ \textit{ tel que } \ 1 \leq i \leq k.$$

V *est l'espace vectoriel engendré par les vecteurs arguments.*

Démonstration.
On effectuera une démonstration algorithmique.

$$p_1 = q_1$$

et par conséquent $p_1 \neq 0$. Bien évidemment $V(p_1) = V(q_1)$.

$$p_2 = q_2 - \frac{(q_2, p_1)}{\| p_1 \|_2^2} p_1.$$

Pour $i \leq k$,

$$p_i = q_i - \sum_{j=1}^{i-1} \frac{(q_i, p_j)}{\| p_j \|_2^2} p_j.$$

On voit immédiatement que $(p_2, p_1) = (q_2, p_1) - (q_2, p_1) = 0$ et que $V(p_1, p_2) = V(p_1, q_2)$ d'après la définition de p_2. Donc $p_2 \neq 0$.

On démontrera la propriété par récurrence en supposant que $(p_i, p_j) = 0$ si $i \neq j$ et si $i, j < m$, et que $V(q_1, \ldots, q_i) = V(p_1, ..., p_i)$ pour $i < m$. Alors

$$(p_m, p_s) = (q_m, p_s) - \sum_{j=1}^{m-1} \frac{(q_m, p_j)}{\| p_j \|_2^2}(p_j, p_s) = (q_m, p_s) - (q_m, p_s) = 0$$

pour $s < m$, en utilisant l'hypothèse de récurrence.

Enfin

$$\begin{aligned} V(p_1, \ldots, p_{m-1}, p_m) &= V(p_1, \ldots, p_{m-1}, q_m) \text{ d'après la définition de } p_m \\ &= V(q_1, \ldots, q_{m-1}, q_m) \text{ en utilisant l'hypothèse de récurrence.} \end{aligned}$$

Donc $p_m \neq 0$. $\square$

Remarque 0.0.5 *Si on pose* $r_j = \dfrac{p_j}{\| p_j \|_2}$, *alors* $\| r_j \|_2 = 1$ *et la suite* $\{r_j\}$ *est orthonormée.*

Définissons à présent l'image et le noyau d'une application linéaire.

Définition 0.0.6 *L'image de* A, *notée* ImA, *est définie comme l'ensemble des vecteurs* $y \in \mathbb{C}^n$ *tels qu'il existe pour chaque* y *un vecteur* $x \in \mathbb{C}^n$ *tel que* $Ax = y$.
Le noyau de A, *noté* $KerA$, *est l'ensemble des vecteurs* $x \in \mathbb{C}^n$ *tels que* $Ax = 0$.

Définition 0.0.7 *On appelle dimension d'un sous-espace vectoriel le nombre de vecteurs de sa base.*

Le rang d'une matrice A, noté $rg(A)$, est encore égal à la dimension de l'image de A.

Propriété 0.0.8 *Propriété fondamentale.*

$$\boxed{rg(A) + dim(Ker\ A) = n}$$

0.1 NORME D'UN OPERATEUR LINEAIRE

On ne s'intéressera ici encore qu'aux opérateurs linéaires de $\mathbb{C}^n$ dans $\mathbb{C}^n$.

On appelle norme sur $\mathbb{C}^n$ une application ℓ de $\mathbb{C}^n$ dans rrm^+ qui vérifie les axiomes suivants :

i) $\ell(x) \geq 0 \quad \forall x \in \mathbb{C}^n$.
ii) $\ell(x) = 0 \Longleftrightarrow x = 0$.
iii) $\ell(x+y) \leq \ell(x) + \ell(y) \quad \forall x$ et $y \in \mathbb{C}^n$.
iv) $\ell(\lambda x) =\mid \lambda \mid \ell(x) \quad \forall x \in \mathbb{C}^n$ et $\forall \lambda \in \mathbb{C}$.

Définition 0.1.1 *On appelle norme de Holder toute application* ℓ_p *de* $\mathbb{C}^n$ *dans* rrm^+ *définie par*

$$\ell_p(x) = (\sum_{j=1}^{n} \mid x_j \mid^p)^{1/p}.$$

On peut démontrer que ℓ_p satisfait tous les axiomes d'une norme. Les axiomes i), ii) et iv) d'une norme se vérifient trivialement. L'axiome iii), dit inégalité triangulaire, se démontre en utilisant l'inégalité de Holder.

Les trois normes de Holder les plus utilisées sont :

$\ell_\infty(x) =\parallel x \parallel_\infty = \max_j \mid x_j \mid$. C'est la norme de la convergence absolue.

$\ell_2(x) =\parallel x \parallel_2 = (x,x)^{1/2}$. C'est la norme euclidienne.

$\ell_1(x) =\parallel x \parallel_1 = \sum_{j=1}^{n} \mid x \mid_j$. C'est la norme de la convergence en moyenne.

Propriété 0.1.2 *Inégalité de Holder.*

$$\mid \sum_{j=1}^{n} x_j\, y_j \mid \leq \ell_p(x)\, \ell_q(y) \quad \textit{avec } p > 1 \quad \textit{et } \frac{1}{p} + \frac{1}{q} = 1.$$

On définit une norme d'un opérateur linéaire de $\mathcal{L}(\mathbb{C}^n,\mathbb{C}^n)$, dont la matrice est A lorsque l'espace $\mathbb{C}^n$ est rapporté à deux bases par

$$S_{\ell_p\,\ell_q}(A) = \sup_{x \in \mathbb{C}^n - \{0\}} \frac{\ell_p(Ax)}{\ell_q(x)}.$$

On dit encore que $S_{\ell_p\,\ell_q}(A)$ est la norme matricielle induite (par les deux normes vectorielles ℓ_p et ℓ_q).

On peut montrer que ce nombre est toujours défini, que la borne supérieure est atteinte et qu'elle est indépendante du choix des bases. On a donc un nombre qui caractérise l'opérateur linéaire. $S_{\ell_p\,\ell_q}$ est une norme pour l'opérateur linéaire.

En fait on utilise plus fréquemment les normes $S_{\ell_p\,\ell_p}$ dont les plus courantes correspondent à $p = 1, 2$ *ou* ∞.

On a la propriété suivante :

Propriété 0.1.3 *Inégalité de Minkowski.*

$$\ell_p(x+y) \leq \ell_p(x) + \ell_p(y).$$

Propriété 0.1.4

$$\begin{aligned} S_{\ell_1\,\ell_1}(A) &= \max_{1\leq j\leq n} \sum_{i=1}^{n} \mid a_{ij} \mid . \\ S_{\ell_\infty\,\ell_\infty}(A) &= \max_{1\leq i\leq n} \sum_{j=1}^{n} \mid a_{ij} \mid . \\ S_{\ell_2\,\ell_2}(A) &= \sqrt{\rho(A^*A)} \end{aligned}$$

où $\rho(A^*A)$ est le rayon spectral de A^*A, c'est-à-dire si $\lambda_1, \ldots, \lambda_n$ sont les valeurs propres de A^*A, $\rho(A^*A) = \max_j \mid \lambda_j \mid$.

Remarque 0.1.5 *Si A est hermitienne, alors $S_{\ell_2\,\ell_2}(A) = \rho(A)$.*

Lorsqu'il n'y aura pas d'ambiguité, on utilisera la notation $\parallel A \parallel_i$ avec $i = 1, 2, \infty$ ou p à la place de$S_{\ell_p\,\ell_p}(A)$.

Nous pouvons maintenant aborder le problème du conditionnement d'une matrice.

Si on pose $r(x) = \frac{\ell_p(Ax)}{\ell_p(x)}$, alors $S_{\ell_p\,\ell_p}(A)$ est le $\max_{x\in\mathbb{C}^n-\{0\}} r(x)$. Cherchons le $\min_{x\in\mathbb{C}^n-\{0\}} r(x)$. En posant $y = Ax$, on obtient :

$$r(x) = \frac{\ell_p(Ax)}{\ell_p(x)} = \frac{\ell_p(y)}{\ell_p(A^{-1}y)} = \frac{1}{\frac{\ell_p(A^{-1}y)}{\ell_p(y)}}.$$

Si on introduit $S_{\ell_p\,\ell_p}(A^{-1}) = \max_{x\in\mathbb{C}^n-\{0\}} \frac{\ell_p(A^{-1}y)}{\ell_p(y)}$, on voit que l'on a :

$$\min_{x\in C^n-\{0\}} r(x) = \frac{1}{S_{\ell_p\,\ell_p}(A^{-1})}.$$

Définition 0.1.6 *On appelle conditionnement d'un opérateur linéaire de $\mathbb{C}^n$ dans $\mathbb{C}^n$ le rapport, noté $cond_p(A)$,*

$$cond_p(A) = \frac{\max r(x)}{\min r(x)}) = S_{\ell_p\ell_p}(A)\, S_{\ell_p\ell_p}(A^{-1}).$$

Définition 0.1.7 *Une norme est dite multiplicative si*

$$\ell(AB) \leq \ell(A)\,\ell(B).$$

Les normes $S_{\ell_p\ell_p}(A)$ sont des normes multiplicatives.

Propriété 0.1.8 *i)* $\ell_p(Au) \leq S_{\ell_p\ell_p}(A)\,\ell_p(u)\ \ \forall u \in \mathbb{C}^n$.
ii)$\rho(A) \leq S_{\ell_p\ell_p}(A)\ \ \forall p \geq 1$.

Démonstration.
i) La relation est vraie pour $u = 0$. Pour $u \neq 0$ on considère le vecteur $\frac{u}{\ell_p(u)}$ qui est de norme unité.

$$\ell_p(A\frac{u}{\ell_p(u)}) \leq \max_{\ell_p(v)=1} \ell_p(Av) = S_{\ell_p\ell_p}(A)$$

et par conséquent en multipliant les deux membres par $\ell_p(u)$ on obtient le résultat.

ii) Quelle que soit la valeur propre λ de A, on a, si u est son vecteur propre associé

$$\mid \lambda \mid \ell_p(u) = \ell_p(\lambda u) = \ell_p(Au) \leq S_{\ell_p\ell_p}(A)\ell_p(u).$$

Donc $\mid \lambda \mid \leq S_{\ell_p\ell_p}(A)\ \ \forall \lambda$. $\square$

Abordons, pour finir cette section, la convergence d'une suite de matrices.

Définition 0.1.9 *Soit* $\{A^{(k)}\}$ *une suite de matrices. On dira que* $A^{(k)}$ *converge vers* A *si* $\lim_{n\to\infty} \parallel A^{(k)} - A \parallel = 0$ *où la norme choisie* $\parallel \parallel$ *est une norme* $S_{\ell_p\ell_p}$.

Dans certains cas nous générons des suites de matrices telles que

$$A^{(k)} = A^k\ \ \forall k \in \mathbb{N} - \{0\}.$$

Il est intéressant de connaître la condition nécessaire et suffisante de convergence de cette suite vers 0 (matrice nulle).

Théorème 0.1.10 *Une condition nécessaire et suffisante pour que la suite* $\{A^m\}$ *converge vers la matrice nulle lorsque* m *tend vers l'infini, est que* $\rho(A) < 1$.

Démonstration.
Pour toute matrice A il existe une décomposition de Jordan, c'est-à-dire qu'il existe une matrice S inversible telle que $S\,A\,S^{-1} = \tilde{A}$, où $\tilde{A}$ est la forme de

Jordan :

$$\widetilde{A} = \begin{pmatrix} J_1 & & & & & & \\ & J_2 & & & & & \\ & & \ddots & & & 0 & \\ & & & J_i & & & \\ & 0 & & & \ddots & & \\ & & & & & J_{N-1} & \\ & & & & & & J_N \end{pmatrix},$$

où J_i est une matrice $n_i \times n_i$ telle que :

$$J_i = \begin{pmatrix} \lambda_i & 1 & & & \\ & \lambda_i & 1 & & 0 \\ & & \ddots & \ddots & \\ & 0 & & \ddots & 1 \\ & & & & \lambda_i \end{pmatrix}.$$

λ_i est une valeur propre de A et on a :

$$\sum_{i=1}^{k} n_i = n.$$

Les λ_i ne sont pas forcément toutes distinctes. Pour démontrer la convergence on exprime $(\widetilde{A})^m$.

$$(\widetilde{A})^m = S\, A^m\, S^{-1} \Longleftrightarrow A^m = S^{-1}\, (\widetilde{A})^m\, S.$$

On voit donc que montrer que A^m tend vers 0 équivaut à montrer que $(\widetilde{A})^m$ tend vers 0. Or $(\widetilde{A})^m$ est une matrice bloc diagonale. Les blocs de la diagonale sont tout simplement les J_i^m. Par conséquent $(\widetilde{A})^m$ ne tendra vers 0 que si et seulement si tous les J_i^m tendent vers 0.

On peut calculer tous les J_i^mpar récurrence. Si les éléments de J_i^m sont notés $j_{p,q}^{(m,i)}$ pour p et $q = 1, \ldots, n_i$, on voit qu'on obtient :

$$j_{p,q}^{(m,i)} = \begin{cases} 0 \text{ si } q < p, \\ C_m^{q-1}\, \lambda_i^{m-q+p} \text{ si } p \leq q < \max(n_i, m+p), \\ 0 \text{ si } m + p \leq q < n_i. \end{cases}$$

D'où J_i^m tend vers 0 si $\mid \lambda_i \mid < 1\ \forall i.$ $\square$

On peut remarquer que $\parallel J_i^m \parallel \sim C_m^{n_i}\Big(\rho(J_i^{m-(n_i-1)})\Big)$ quand $m \to \infty$.

METHODES ITERATIVES:

Nous cherchons à résoudre le système $Ax = b$. On commence par décomposer la matrice A en

$$A = M - N,$$

de telle façon que M soit inversible. Nous écrirons le système $Ax = b$ sous la forme

$$M\,x = N\,x + b$$

ou encore

$$x = M^{-1}N\,x + M^{-1}b,$$

qui définit une équation de point fixe. Pour la résoudre nous calculerons par récurrence la suite de vecteurs $\{\overset{i}{x}\}$ à partir d'un vecteur $\overset{0}{x}$ choisi arbitrairement à l'aide de la relation suivante :

$$\overset{k+1}{x} = M^{-1}\,N\,\overset{k}{x} + M^{-1}b.$$

Cette relation est une relation de récurrence du premier ordre. Nous pouvons en déduire une relation reliant $\overset{k}{e} = \overset{k}{x} - \overline{x}$ à $\overset{k-1}{e} = \overset{k-1}{x} - \overline{x}$.

$$M(\overset{k}{x} - \overline{x}) = N(\overset{k-1}{x} - \overline{x}),$$

puisque $M\,\overline{x} = N\,\overline{x} + b$ et donc $\overset{k}{e} = M^{-1}\,N\,\overset{k-1}{e}$ pour $k = 1, 2, \ldots$

Posons $B = M^{-1}N$. Nous avons alors

$$\overset{k}{e} = B^k\,\overset{0}{e}\,.$$

Le processus sera convergent si $B^k \longrightarrow 0$ quand $k \longrightarrow \infty$. Le théorème 0.1.10 a montré qu'une condition nécessaire et suffisante pour que B^k tende vers 0 est que $\rho(B) < 1$.

Nous allons maintenant donner deux propriétés qui vont nous guider dans le choix de la décomposition de la matrice A.

Théorème 0.1.11 *Si A est inversible, une condition nécessaire et suffisante pour que la suite $\{\overset{k}{x}\}$ tende vers $\overline{x}$ est que la suite $\overset{k}{r} = b - A\,\overset{k}{x}$ tende vers le vecteur nul de $\mathbb{R}^n$.*

Démonstration.

$$\begin{aligned}\overset{k}{r} &= -A\,\overset{k}{e},\\ \overset{k}{e} &= -A^{-1}\,\overset{k}{r}\,.\end{aligned}$$

En utilisant une norme ℓ nous avons

$$\begin{aligned}\ell(\overset{k}{r}) &\leq S_{\ell\ell}(A)\,\ell(\overset{k}{e}),\\ \ell(\overset{k}{e}) &\leq S_{\ell\ell}(A^{-1})\,\ell(\overset{k}{r}).\end{aligned}$$

Si $x \longrightarrow \overline{x}$, alors $\overset{k}{e}\longrightarrow 0$, ce qui entraîne que $\ell(\overset{k}{e}) \longrightarrow 0$ et donc $\ell(\overset{k}{r}) \longrightarrow 0$ en utilisant la première inégalité, d'où $\overset{k}{r}\longrightarrow 0$. La seconde inégalité permet de démontrer la réciproque. $\square$

La dernière relation $\ell(\overset{k}{e}) \leq S_{\ell\ell}(A^{-1})\,\ell(\overset{k}{r})$ donne une majoration absolue de la norme de l'erreur. Il est préférable de considérer l'erreur relative $\dfrac{\ell(\overset{k}{e})}{\ell(\overset{k}{x})}$.

Théorème 0.1.12 *Le rapport maximum de l'erreur relative au minimum de l'erreur relative, pour un affaiblissement relatif* $\dfrac{\ell(\overset{k}{r})}{\ell(\overset{k}{x})}$ *donné, est égal au conditionnement de* A.

Démonstration.

$$\frac{\ell(\overset{k}{r})}{\ell(\overset{k}{x})} = \frac{\ell(A\ \overset{k}{e})}{\ell(\overset{k}{x})} = \frac{\ell(A\ \overset{k}{e})}{\ell(\overset{k}{e})}\,\frac{\ell(\overset{k}{e})}{\ell(\overset{k}{x})}$$

et donc

$$\frac{\ell(\overset{k}{e})}{\ell(\overset{k}{x})} = \frac{\ell(\overset{k}{r})}{\ell(\overset{k}{x})}\Big(\frac{\ell(A\ \overset{k}{e})}{\ell(\overset{k}{e})}\Big)^{-1}.$$

D'autre part

$$\frac{1}{S_{\ell\ell}(A^{-1})} \leq \frac{\ell(A\,x)}{\ell(x)} \leq S_{\ell\ell}(A).$$

L'erreur relative minimum est donc

$$\frac{\ell(\overset{k}{r})}{\ell(\overset{k}{x})}\frac{1}{S_{\ell\ell}(A)}.$$

L'erreur relative maximum est

$$\frac{\ell(\overset{k}{r})}{\ell(\overset{k}{x})}S_{\ell\ell}(A^{-1}).$$

D'où le rapport est égal à $S_{\ell\ell}(A^{-1})\,S_{\ell\ell}(A) = \ \mathrm{cond}(A)$. $\square$

Les deux résultats suivants sont très intéressants, car ils décrivent l'influence des perturbations des termes du second membre ou de la matrice sur la solution du système linéaire.

Théorème 0.1.13 *Si x et $x+\Delta x$ sont respectivement les solutions des systèmes linéaires suivants :*

$$A\,x = b \quad et\; A\,(x+\Delta x) = b + \Delta b,$$

alors

$$\frac{\parallel \Delta x \parallel}{\parallel x \parallel} \leq \;\; cond(A)\frac{\parallel \Delta b \parallel}{\parallel b \parallel}.$$

Démonstration.
Les deux systèmes donnent par différence $A\,\Delta x = \Delta b$ ou $\Delta x = A^{-1}\,\Delta b$. Nous obtenons donc

$$\parallel \Delta x \parallel \leq \parallel A^{-1} \parallel \parallel \Delta b \parallel .$$

Or $\parallel b \parallel = \parallel A\,x \parallel \leq \parallel A \parallel \parallel x \parallel$, et par conséquent

$$\frac{\parallel \Delta x \parallel}{\parallel x \parallel} \leq \parallel A \parallel \parallel A^{-1} \parallel \frac{\parallel \Delta b \parallel}{\parallel b \parallel}.$$

□

Théorème 0.1.14 *Si x et $x+\Delta x$ sont respectivement les solutions des systèmes linéaires suivants :*

$$A\,x = b \quad et\; (A+\Delta A)\,(x+\Delta x) = b,$$

alors

$$\frac{\parallel \Delta x \parallel}{\parallel x+\Delta x \parallel} \leq \;\; cond(A)\frac{\parallel \Delta A \parallel}{\parallel A \parallel}.$$

Démonstration.

$$A\,x = b = (A+\Delta A)(x+\Delta x)$$

et donc

$$\begin{aligned} 0 &= A\,\Delta x + \Delta A\,(x+\Delta x), \\ \Delta x &= -A^{-1}\,\Delta A\,(x+\Delta x), \\ \parallel \Delta x \parallel &= \parallel A^{-1} \parallel \parallel \Delta A \parallel \parallel x+\Delta x) \parallel . \end{aligned}$$

En multipliant les deux membres par $\parallel A \parallel$ on obtient le résultat. □

Le choix de la décomposition de la matrice A devra respecter les directives suivantes :

1. $\rho(M^{-1}\,N)$ doit être strictement inférieur à 1.
2. La résolution de $M \overset{k}{x} = N \overset{k-1}{x} + b$ doit être simple et nécessiter le moins possible d'opérations.

3. Pour obtenir la meilleure convergence, $\rho(M^{-1}\,N)$ doit être le plus petit possible.

Les décompositions utilisées de A font intervenir :

- la matrice diagonale D obtenue à partir des éléments diagonaux de A.

$$D = \begin{pmatrix} a_{11} & 0 & \cdots \\ & \ddots & \\ \cdots & 0 & a_{nn} \end{pmatrix},$$

- la matrice triangulaire inférieure $-E$:

$$-E = \begin{pmatrix} 0 & & & \\ a_{21} & 0 & 0 & \\ \vdots & & \ddots & \\ a_{n1} & \cdots & a_{n,n-1} & 0 \end{pmatrix},$$

- la matrice triangulaire supérieure $-F$:

$$-F = \begin{pmatrix} 0 & a_{12} & \cdots & a_{1n} \\ & & \ddots & \vdots \\ & 0 & & a_{n-1,n} \\ & & & 0 \end{pmatrix}.$$

Nous avons donc $A = D - E - F$. Nous obtiendrons la décomposition $A = M - N$ à partir de différents types de regroupement de ces matrices D, $-E$ et $-F$.

0.1.1 La méthode de Jacobi

On prend $M = D$ et $N = E + F$.

Le calcul de B est simple, puisque $B = M^{-1}\,N = D^{-1}\,(E+F)$. Nous aurons donc

$$\overset{k+1}{x} = D^{-1}\,(E+F)\ \overset{k}{x} + D^{-1}\,b.$$

Si on exprime cette relation en fonction des éléments de la matrice A, nous avons

$$\overset{k+1}{x_i} = -\sum_{\substack{j=1\\ j\neq i}}^{n} \frac{a_{ij}}{a_{ii}}\ \overset{k}{x_j} + \frac{b_i}{a_{ii}} \text{ pour } i = 1, \ldots, n.$$

Cette expression peut être transformée en utilisant celle du vecteur résidu.

$$\overset{k}{r}_i = -\sum_{j=1}^{n} a_{ij}\ \overset{k}{x}_j\ + b_i = -\sum_{\substack{j=1\\ j\neq i}}^{n} a_{ij}\ \overset{k}{x}_j\ + b_i - a_{ii}\ \overset{k}{x}_i,$$

d'où

$$\begin{aligned}
\overset{k+1}{x}_i &= \frac{\overset{k}{r}_i}{a_{ii}} + \overset{k}{x}_i, \\
\overset{k+1}{r}_i &= -\sum_{\substack{j=1\\ j\neq i}}^{n} \frac{a_{ij}}{a_{jj}}\ \overset{k}{r}_j\ .
\end{aligned}$$

Pour obtenir cette dernière relation on écrit

$$\begin{aligned}
\overset{k+1}{r}_i &= b_i - \sum_{j=1}^{n} a_{ij}\ \overset{k+1}{x}_j \\
&= b_i - \sum_{j=1}^{n} a_{ij}\ \overset{k}{x}_j\ - \sum_{j=1}^{n} a_{ij}\ \frac{\overset{k}{r}_j}{a_{jj}} = \overset{k}{r}_i - \sum_{j=1}^{n} \frac{a_{ij}}{a_{jj}}\ \overset{k}{r}_j\ .
\end{aligned}$$

0.1.2 La méthode de Gauss-Seidel

Philip Ludwig von Seidel : mathématicien allemand : Zweibrücken : 1821 - Munich : 1896.

Cette méthode utilise $M = D - E$ et $N = F$. D'où $B = M^{-1}\, N = (D - E)^{-1}\, F$.
Alors

$$\overset{k+1}{x} = (D - E)^{-1}\, F\ \overset{k}{x}\ + (D - E)^{-1}\, b.$$

Le calcul de l'inverse de $D - E$ peut être évité. Si on écrit $(D - E)\ \overset{k+1}{x} = F\ \overset{k}{x}\ + b$ on obtient

$$\sum_{j=1}^{i} a_{ij}\ \overset{k+1}{x}_j = -\sum_{j=i+1}^{n} a_{ij}\ \overset{k}{x}_j\ + b_i.$$

D'où

$$\overset{k+1}{x}_i = -\frac{1}{a_{ii}} \sum_{j=1}^{i-1} a_{ij}\ \overset{k+1}{x}_j\ - \frac{1}{a_{ii}} \sum_{j=i+1}^{n} a_{ij}\ \overset{k}{x}_j\ + \frac{b_i}{a_{ii}}. \tag{1}$$

On calcule donc itérativement $\overset{k+1}{x}_i$ pour $i = 1, \ldots, n$.

Nous allons introduire des vecteurs intermédiaires $\overset{k,i}{x}$ tels que

$$\overset{k,i}{x}_j = \begin{cases} \overset{k+1}{x}_j & \text{si } j < i, \\ \overset{k}{x}_j & \text{si } j \geq i. \end{cases}$$

Le vecteur residu $\overset{k,i}{r}$ reste égal à $b - A\ \overset{k,i}{x}$. La $i^{ème}$ composante de ce vecteur résidu vaut

$$\begin{aligned}
\overset{k,i}{r_i} &= b_i - \sum_{j=1}^{i-1} a_{ij}\ \overset{k,i}{x_j} - \sum_{j=i+1}^{n} a_{ij}\ \overset{k,i}{x_j} - a_{ii}\ \overset{k,i}{x_i} \\
&= b_i - \sum_{j=1}^{i-1} a_{ij}\ \overset{k+1}{x_j} - \sum_{j=i+1}^{n} a_{ij}\ \overset{k}{x_j} - a_{ii}\ \overset{k}{x_i} \\
&= a_{ii}\ \overset{k+1}{x_i} - a_{ii}\ \overset{k}{x_i} \ \text{ d'après (1).}
\end{aligned}$$

On remarquera que $\overset{k,i+1}{r_i} = 0$. Donc la modification de la $i^{ème}$ composante de $\overset{k,i}{x}$ revient à annuler la même composante du vecteur résidu. En effet

$$\begin{aligned}
\overset{k,i+1}{r_i} &= b_i - \sum_{j=1}^{i} a_{ij}\ \overset{k,i+1}{x_j} - \sum_{j=i+1}^{n} a_{ij}\ \overset{k,i+1}{x_j} \\
&= b_i - \sum_{j=1}^{i} a_{ij}\ \overset{k+1}{x_j} - \sum_{j=i+1}^{n} a_{ij}\ \overset{k}{x_j} = 0 \ \text{ d'après (1).}
\end{aligned}$$

Enfin

$$\begin{aligned}
\overset{k+1}{r_i} &= b_i - \sum_{j=1}^{n} a_{ij}\ \overset{k+1}{x_j} \\
&= b_i - \sum_{j=1}^{i-1} a_{ij}\ \overset{k,i}{x_j} - \sum_{j=i}^{n} a_{ij}\ \overset{k+1}{x_j} \\
&= b_i - \sum_{j=1}^{i-1} a_{ij}\ \overset{k,i}{x_j} - \sum_{j=i}^{n} a_{ij}\ \overset{k,i}{x_j} - \sum_{j=i}^{n} a_{ij}\,(\overset{k+1}{x_j} - \overset{k,i}{x_j}) \\
&= \overset{k,i}{r_i} - \sum_{j=i}^{n} a_{ij}\,(\overset{k+1}{x_j} - \overset{k,i}{x_j}) \\
&= \overset{k,i}{r_i} - \sum_{j=i}^{n} a_{ij}\,(\overset{k+1}{x_j} - \overset{k}{x_j}).
\end{aligned}$$

D'où

$$\begin{aligned}
\overset{k+1}{r_i} &= \overset{k,i}{r_i} - \sum_{j=i}^{n} \frac{a_{ij}}{a_{jj}}\ \overset{k,j}{r_j} \\
&= - \sum_{j=i+1}^{n} \frac{a_{ij}}{a_{jj}}\ \overset{k,j}{r_j}\ . \qquad (2)
\end{aligned}$$

Pour résumer ce que fait la méthode de Gauss-Seidel, on peut dire qu'on passe d'un vecteur $\overset{k,i}{x}$ à un vecteur $\overset{k+1,i}{x}$ en ne changeant qu'une composante, la $i^{ème}$, de telle sorte que la $i^{ème}$ composante du vecteur résidu $\overset{k,i+1}{r}$ soit nulle. Lorsque i = n+1, alors $\overset{k+1,n}{x} = \overset{k+1}{x}$. On a $\overset{k,1}{x} = \overset{k}{x}$. $\overset{k,i+1}{x}$ est calculé par la relation (1.6).

On remarquera qu'il est nécessaire que D soit inversible ($a_{ii} \neq 0 \ \forall i = 1, \ldots, n$).

Une variante de la méthode de Gauss-Seidel, due à Southwell, consiste à modifier la composante i du vecteur $\overset{k}{x}$ qui est telle que $|\overset{k}{r_i}| = \max_j |\overset{k}{r_j}|$. On obtient donc un nouveau vecteur dont le vecteur résidu a une composante i nulle. On recommence en cherchant la composante de module maximum du résidu et ainsi de suite.

Remarque 0.1.15 *Le test d'arrêt utilisé dans les méthodes itératives est*

$$\frac{\|\overset{k}{r}\|}{\| b \|} \leq \varepsilon.$$

Ce test est lié à l'erreur relative de la façon suivante :

$$\frac{\|\overset{k}{r}\|}{\| b \|} = \frac{\| A\ \overset{k}{e}\|}{\| A\,\overline{x} \|} = \frac{\| A^{-1} \|}{\| A^{-1} \|}\frac{\| A\ \overset{k}{e}\|}{\| A\,\overline{x} \|}.$$

Comme $\| \overline{x} \| = \| A^{-1}\, A\,\overline{x} \| \leq \| A^{-1} \| \| A\,\overline{x} \|$ *et* $\| A\ \overset{k}{e}\| \leq \| A \| \|\overset{k}{e}\|$, *alors*

$$\frac{\|\overset{k}{r}\|}{\| b \|} \leq \| A \| \| A^{-1} \| \frac{\|\overset{k}{e}\|}{\| \overline{x} \|} = cond\ A\ \frac{\|\overset{k}{e}\|}{\| \overline{x} \|}.$$

0.1.3 Convergence des méthodes itératives

Nous commencons par donner des proprietés sur certaines matrices et sur la localisation des valeurs propres.

Théorème 0.1.16 *Premier théorème de Gerschgorin-Hadamard. Les valeurs propres de A appartiennent à l'union des n disques D_k pour $k = 1, \ldots, n$ du plan complexe ($\lambda \in \cup_{k=1}^{n} D_k$) où D_k, appelé disque de Gerschgorin, est défini par*

$$| z - a_{kk} | \leq \Lambda_k = \sum_{\substack{j=1 \\ j \neq k}}^{n} | a_{kj} | .$$

Démonstration.
A λ valeur propre on fait correspondre le vecteur propre u. On peut ramener la plus grande de ses composantes à la valeur 1.

Supposons que ce soit la $k^{ème}$ composante.
Donc $1 = \max_i \mid u_i \mid = \mid u_k \mid$.
Alors la $k^{ème}$ composante de $A\,u = \lambda\,u$ s'écrit

$$(\lambda - a_{kk})\,u_k = \sum_{\substack{j=1\\ j\neq k}}^{n} a_{kj}\,u_j.$$

Donc $\mid \lambda - a_{kk} \mid = \mid \sum_{\substack{j=1\\ j\neq k}}^{n} a_{kj}\,u_j \mid \leq \sum_{\substack{j=1\\ j\neq k}}^{n} \mid a_{kj} \mid \mid u_j \mid \leq \sum_{\substack{j=1\\ j\neq k}}^{n} \mid a_{kj} \mid$, ce qui entraîne $\lambda \in D_k$.

En fait k n'est pas connu et par conséquent

$$\lambda \in \cup_{k=1}^{n} D_k.$$

□

Définition 0.1.17 *Une matrice est dite à diagonale dominante si*

$$\forall i,\ 1 \leq i \leq n,\ \mid a_{ii} \mid \geq \sum_{\substack{j=1\\ j\neq i}}^{n} \mid a_{ij} \mid .$$

Une matrice est dite à diagonale strictement dominante si

$$\forall i,\ 1 \leq i \leq n,\ \mid a_{ii} \mid > \sum_{\substack{j=1\\ j\neq i}}^{n} \mid a_{ij} \mid .$$

Une matrice est dite à diagonale fortement dominante si cette matrice est à diagonale dominante et s'il existe au moins un indice i pour lequel

$$\mid a_{ii} \mid > \sum_{\substack{j=1\\ j\neq i}}^{n} \mid a_{ij} \mid .$$

Propriété 0.1.18 *(Théorème d'Hadamard)*

Si une matrice A est à diagonale strictement dominante, alors A est inversible.

Démonstration.
On montre que 0 n'est pas valeur propre. Or $0 \notin \cup_{k=1}^{n} D_k$, puisque

$$\mid 0 - a_{ii} \mid = \mid a_{ii} \mid > \sum_{\substack{j=1\\ j\neq i}}^{n} \mid a_{ij} \mid = \Lambda_i.$$

□

Théorème 0.1.19 *Si A est une matrice à diagonale strictement dominante, alors la méthode de Jacobi est convergente quel que soit le vecteur initial $\overset{0}{x}$.*

Démonstration.
A est inversible. De plus $| a_{ii} |> \sum_{\substack{j=1\\j\neq i}}^{n} | a_{ij} | \Longrightarrow a_{ii} \neq 0$.

Pour démontrer la convergence on doit prouver que

$$\begin{aligned} \rho(B) &< 1 \text{ avec } B = D^{-1}(E+F). \\ B_{ii} &= 0, \\ B_{ij} &= -\frac{a_{ij}}{a_{ii}} \text{ pour } i \neq j. \end{aligned}$$

Par conséquent, puisque A est à diagonale dominante,

$$\forall i, \sum_{j=1}^{n} | B_{ij} |= \sum_{\substack{j=1\\j\neq i}}^{n} \frac{| a_{ij} |}{| a_{ii} |} < 1.$$

D'autre part d'après la propriété 0.1.8 ii,

$$\rho(B) \leq \| B \|_\infty = \max_i \sum_{j=1}^{n} | B_{ij} |< 1.$$

□

De la propriété 0.1.18 on peut déduire trivialement le corollaire suivant :

Corollaire 0.1.20 *Si une matrice Q est singulière, alors il existe au moins un indice j tel que*

$$| q_{jj} |\leq \sum_{\substack{h=1\\h\neq j}}^{n} | q_{jh} | .$$

Théorème 0.1.21 *Théorème de Geiringer. Si la matrice est à diagonale strictement dominante, alors la méthode de Gauss-Seidel est convergente quel que soit le choix du vecteur initial $\overset{0}{x}$.*

Démonstration.
On doit montrer que $\rho(B) < 1$ avec $B = (D-E)^{-1}F$. Or pour toute valeur propre λ de B, $B - \lambda I = (D-E)^{-1}(F - \lambda(D-E))$ est singulière. $(D-E)$ est bien inversible, car c'est une matrice triangulaire inférieure dont la diagonale ne contient que des termes non nuls, les a_{jj}. Donc $Q = F - \lambda(D-E)$ est singulière.

D'après le corollaire 0.1.20,

$$\exists j \text{ tel que } | q_{jj} |\leq \sum_{\substack{h=1\\h\neq j}}^{n} | q_{jh} |= \sum_{h=1}^{j-1} | q_{jh} | + \sum_{h=j+1}^{n} | q_{jh} | .$$

Or

$$q_{jh} = \begin{cases} -a_{jh} \text{ si } j < h, \\ -\lambda\, a_{jh} \text{ si } j \geq h. \end{cases}$$

Par conséquent on a

$$| \lambda | \, | a_{jj} | \leq \sum_{h=1}^{j-1} | a_{jh} | \, | \lambda | + \sum_{h=j+1}^{n} | a_{jh} | \, .$$

S'il existe une valeur propre telle que $| \lambda | \geq 1$, on peut majorer le membre de droite de l'inégalité ci-dessus par

$$| \lambda | \sum_{\substack{h=1 \\ h \neq j}}^{n} | a_{jh} |,$$

ce qui montrerait que A n'est pas à diagonale strictement dominante, d'où une contradiction et par conséquent $| \lambda | < 1$. $\square$

0.1.4 Les méthodes de relaxation.

Lorsque le vecteur $\overset{k}{x}$ est connu, ces méthodes déterminent $\overset{k+1}{x_i}$ par

$$\overset{k+1}{x_i} = (1 - \omega) \overset{k}{x_i} + \omega \overset{k+1^*}{x_i}, \tag{3}$$

où ω est un paramètre réel, appelé le facteur de relaxation, et $\overset{k+1^*}{x_i}$ est déterminé par une méthode de Jacobi ou de Gauss-Seidel de la façon suivante :
Avec la méthode de Jacobi on obtient

$$\overset{k+1}{x_i} = \frac{\omega}{a_{ii}} \Big(b_i - \sum_{\substack{j=1 \\ j \neq i}}^{n} a_{ij} \overset{k}{x_j} \Big) + (1 - \omega) \overset{k}{x_i} \, .$$

Avec la méthode de Gauss-Seidel on obtient

$$\overset{k+1^*}{x_i} = -\frac{1}{a_{ii}} \Big(\sum_{j=1}^{i-1} a_{ij} \overset{k+1}{x_j} + \sum_{j=i+1}^{n} a_{ij} \overset{k}{x_j} - b_i \Big).$$

$\overset{k+1^*}{x_i}$ est donc construit à partir des composantes déjà calculées de $\overset{k+1}{x}$ et de celles de $\overset{k}{x}$ qui doivent être transformées par la suite.
Si $\omega < 1$, on a sous-relaxation.
Si $\omega > 1$, on a sur-relaxation.
Si $\omega = 1$, on retrouve la méthode de Gauss-Seidel ou de Jacobi.
En fait la méthode de Jacobi est peu utilisée dans la méthode de relaxation, car elle n'apporte aucun gain. Par contre la méthode de Gauss-Seidel améliore

souvent la rapidité de la convergence.
Avec la méthode de Gauss-Seidel la relation (3) devient

$$\begin{aligned}
\overset{k+1}{x_i} &= (1-\omega)\,\overset{k}{x_i} - \frac{\omega}{a_{ii}}\Big(\sum_{j=1}^{i-1} a_{ij}\,\overset{k+1}{x_j} + \sum_{j=i+1}^{n} a_{ij}\,\overset{k}{x_j} - b_i\Big)\\
&= -\frac{\omega}{a_{ii}}\Big(\sum_{j=1}^{i-1} a_{ij}\,\overset{k+1}{x_j} + \sum_{j=i}^{n} a_{ij}\,\overset{k}{x_j} - b_i\Big) + \overset{k}{x_i}\ .
\end{aligned}$$

D'où l'écriture matricielle de cette relation

$$\overset{k+1}{x} = \overset{k}{x} - \omega\, D^{-1}(-E\,\overset{k+1}{x} + (D-F)\,\overset{k}{x} - b),$$

ce qui conduit à

$$\frac{1}{\omega}(D-\omega E)\,\overset{k+1}{x} = \frac{1}{\omega}((1-\omega)D+\omega F)\,\overset{k}{x} + b.$$

On voit que cette méthode revient à la décomposition

$$\begin{aligned}
M &= \frac{1}{\omega}(D-\omega E),\\
N &= \frac{1}{\omega}((1-\omega)D+\omega F) \text{ avec } \omega \in \mathbb{R} - \{0\}.
\end{aligned}$$

Par conséquent

$$B = M^{-1}N = (D-\omega E)^{-1}((1-\omega)D+\omega F).$$

On posera $L = D^{-1}E$ et $U = D^{-1}F$. Alors

$$\overset{k+1}{x} = (I-\omega L)^{-1}\Big((1-\omega)I+\omega U\Big)\,\overset{k}{x} + \omega(I-\omega L)^{-1}D^{-1}b.$$

On posera enfin $\mathcal{L}_\omega = (I-\omega L)^{-1}\Big((1-\omega)I+\omega U\Big)$.

La convergence de la méthode est toujours liée au rayon spectral de $\mathcal{L}_\omega$.

Théorème 0.1.22 *Pour toute matrice A, le rayon spectral de $\mathcal{L}_\omega$ est supérieur ou égal à $\mid \omega - 1 \mid$.*

Démonstration.

$$\det(\lambda\, I - \mathcal{L}_\omega) = \det(I-\omega\, L)^{-1}\ \det(\lambda(I-\omega L) - \omega U - (1-\omega)I).$$

Or $\det(I-\omega L) = 1$.

Donc le polynôme caractéristique $p(\lambda) = \det(\lambda\, I - \mathcal{L}_\omega) = \sum_{i=0}^{n} a_i\,\lambda^{n-i}$, avec $a_0 = 1$, est tel que

$$\prod_{i=1}^{n} \lambda_i = (-1)^n \frac{a_n}{a_0} = (-1)^n a_n.$$

Or $a_n = p(0) = \det(-(1-\omega)I - \omega U) = (\omega-1)^n$. Donc $\prod_{i=1}^{n} \mid \lambda_i \mid = \mid \omega - 1 \mid^n$.

Enfin $\rho(\mathcal{L}_\omega) \geq \mid \lambda_i \mid \quad \forall i$. Donc $(\rho(\mathcal{L}_\omega))^n \geq \prod_{i=1}^{n} \mid \lambda_i \mid = \mid \omega - 1 \mid^n$. D'où $\rho(\mathcal{L}_\omega) \geq \mid \omega - 1 \mid$. $\square$

Corollaire 0.1.23 *Pour toute matrice A, une condition nécessaire de convergence de la méthode de relaxation est que*

$$0 < \omega < 2.$$

Démonstration.
Pour avoir $\rho(\mathcal{L}_\omega) < 1$, il est nécessaire que $\mid \omega - 1 \mid < 1$ et donc que $0 < \omega < 2$.
□

Si A est hermitienne définie positive, alors $0 < \omega < 2$ est une condition nécessaire et suffisante pour avoir $\rho(\mathcal{L}_\omega) < 1$.

Théorème 0.1.24 *Si A est hermitienne définie positive, alors une condition nécessaire et suffisante pour que $\rho(\mathcal{L}_\omega) < 1$, est que $0 < \omega < 2$.*

Démonstration.
Soit λ une valeur propre de $\mathcal{L}_\omega$ et x le vecteur propre associé.

$$\mathcal{L}_\omega\, x = \lambda x = (D - \omega E)^{-1}(\omega E^* + (1-\omega)D)x,$$

car $F = E^*$ (A est hermitienne). Par conséquent

$$\begin{aligned} \omega\, E^*\, x + (1-\omega)Dx &= \lambda(D - \omega\, E)\, x, \\ \omega\,(E^*\, x, x) + (1-\omega)(Dx, x) &= \lambda((D - \omega\, E)\, x, x). \end{aligned}$$

Comme $A = D - E - E^*$,

$$\omega\,(E^*\, x, x) = -\omega(Ax, x) + \omega((D - E)x, x).$$

Donc :

$$\begin{aligned} \omega(A\, x, x) &= \omega((D-E)x, x) + (1-\omega)(Dx, x) - \lambda((D - \omega E)x, x) \\ &= (1-\lambda)((D - \omega E)x, x). \end{aligned}$$

$\omega\overline{(Ax, x)} = \omega(x, Ax) = \omega(Ax, x)$ puisque A est hermitienne.

$$\begin{aligned} \omega\overline{(Ax, x)} &= (1-\overline{\lambda})\overline{((D - \omega E)x, x)} = (1-\overline{\lambda})\,(x, (D - \omega E)x) \\ &= (1-\overline{\lambda})((D - \omega E^*)x, x), \end{aligned}$$

puisque la diagonale d'une matrice hermitienne est réelle, et $F = E^*$.

$$= (1-\overline{\lambda})\Big((Dx, x) + (1-\omega)(Dx, x) - \lambda((D - E)x, x)\Big)$$

en utilisant l'expression de $\omega(E^*x, x)$ trouvée précédement,

$$= (2-\omega)(1-\overline{\lambda})(Dx, x) - \lambda(1-\overline{\lambda})((D - \omega E)x, x).$$

On multiplie les deux membres par $1-\lambda$.

$$\begin{aligned}(1-\lambda)\,\omega(Ax,x) &= (2-\omega)\ \mid 1-\lambda\mid^2 (Dx,x)- \\ &\quad \lambda(1-\lambda)(1-\overline{\lambda})((D-\omega E)x,x) \\ &= (2-\omega)\ \mid 1-\lambda\mid^2 (Dx,x)-\lambda(1-\overline{\lambda})\omega(Ax,x)\end{aligned}$$

en utilisant la valeur trouvée pour $\omega(Ax,x)=(1-\overline{\lambda})((D-\omega E)x,x)$.
D'où

$$(1-\mid\lambda\mid^2)(Ax,x)=\frac{2-\omega}{\omega}\mid 1-\lambda\mid^2 (Dx,x).$$

Si le procédé est convergent, alors $\mid\lambda\mid<1$. Puisque A et D sont définies positives, on a donc

$$\frac{2-\omega}{\omega}>0\Longrightarrow 0<\omega<2.$$

Réciproquement, si $\frac{2-\omega}{\omega}>0$ et si $\lambda\neq 1$ alors $1-\mid\lambda\mid^2>0$ et donc $\mid\lambda\mid<1$. Le cas $\lambda=1$ ne peut se produire, car alors $\omega(Ax,x)=(1-\lambda)((D-\omega E)x,x)=0$. Par conséquent on aurait $(Ax,x)=0$, ce qui est contradictoire avec le fait que A est définie positive. $\square$

0.2 RAPPELS

Soit A un opérateur de $\mathbb{C}^n$ dans $\mathbb{C}^n$.

Définition 0.2.1 *On appelle valeur propre de A le nombre $\lambda \in \mathbb{C}$ pour lequel il existe un vecteur u non nul appartenant à $\mathbb{C}^n$, appelé vecteur propre, tel que*

$$A\,u = \lambda\,u. \tag{4}$$

Le couple (valeur propre, vecteur propre associé à λ) est encore appelé élément propre (λ, u).

Définition 0.2.2 *On appelle spectre de A, noté $S_p\,A$, l'ensemble des valeurs propres de A. Le rayon spectral, noté $\rho(A)$, est la plus grande valeur propre en module.*

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

Remarquons que

$$(7.1) \Longleftrightarrow (A - \lambda\,I)u = 0 \Longleftrightarrow u \in \text{Ker } (A - \lambda\,I).$$

Donc l'ensemble des vecteurs propres associés à une même valeur propre λ forme un sous-espace vectoriel, puisque le noyau en est un.

$A - \lambda\,I$ est singulière et donc $\det(A - \lambda\,I) = 0$.

Définition 0.2.3 *On appelle polynôme caractéristique de la matrice A, le polynôme de degré n*

$$P_A(\lambda) = \det(A - \lambda\,I).$$

Nous avons vu que $A - \lambda\,I$ singulière équivaut à $\det(A - \lambda\,I) = 0$. Donc les racines de $P_A(\lambda)$ sont les valeurs propres de la matrice A. Si λ est racine de P_A avec une multiplicité k, on dit que λ est valeur propre de A avec une multiplicité algébrique k.

Théorème 0.2.4 *La dimension du sous-espace propre relatif à une valeur propre est au plus égal à l'ordre de multiplicité de cette valeur propre.*

Démonstration.
Supposons que $\dim \text{Ker } (A - \lambda_1\,I) = k$. On prend la base $u_1, \ldots, u_k$ de ce noyau que l'on compléte par $n - k$ vecteurs $\{u_i\}$, pour $i = k + 1, \ldots, n$, afin de former une base de $\mathbb{R}^n$.

Soit U la matrice de vecteurs colonnes u_i. Alors $A = U\,B\,U^{-1}$ et donc $\det(A - \lambda\,I) = \det(B - \lambda\,I)$. Or les k premières colonnes de B sont égales à $\lambda_1\,e_i$ pour $i = 1, \ldots, k$, où e_i est le $i^{\text{ème}}$ vecteur de la base canonique de $\mathbb{R}^n$.

Par conséquent $\det(B - \lambda\,I) = (\lambda - \lambda_1)^k\,p(\lambda)$, ce qui nous montre que l'ordre de multiplicité de λ_1 est supérieur ou égal à k. $\square$

Définition 0.2.5 *Deux matrices A et B sont dites semblables s'il existe une matrice inversible S telle que*

$$B = S^{-1}\,A\,S.$$

Théorème 0.2.6 *Deux matrices semblables A et B ont les mêmes valeurs propres. Si u et v sont respectivement les vecteurs propres associés à λ pour A et B, on a $v = S^{-1}u$.*

Démonstration.
Pour montrer qu'elles ont les mêmes valeurs propres, il suffit de prouver qu'elles ont le même polynôme caractéristique. Or

$$B - \lambda\,I = S^{-1}\,(A - \lambda\,I)\,S.$$

D'autre part $B\,v = \lambda v \Longrightarrow S^{-1}\,A\,Sv = \lambda\,v \Longrightarrow A\,S\,v = \lambda\,S\,v$. Donc $u = Sv$. $\square$

Il y a donc indépendance des valeurs propres de la matrice A vis-à-vis de la base dans laquelle est représentée la matrice, c'est-à-dire que les valeurs propres sont intrinséquement liées à l'opérateur linéaire A.

Remarque 0.2.7

$$\det A = \prod_{i=1}^{n} \lambda_i.$$

Définition 0.2.8 *On appelle trace de A, notée tr(A), la somme des éléments diagonaux.*

$$\mathrm{tr}(A) = \sum_{i=1}^{n} a_{ii} = \sum_{i=1}^{n} \lambda_i,$$

car le développement de $\det(A - \lambda\,I) = (-1)^n\,\lambda^n + (-1)^{n-1}\lambda^{n-1}\sum_{i=1}^{n} a_{ii} + \ldots$ Par conséquent $\mathrm{tr}(A)$ est indépendante de la base.

Théorème 0.2.9 *Si $Au = \lambda u$, alors*

a) $(A - \mu I)u = (\lambda - \mu)u$.

b) $A^k u = \lambda^k u, \ \forall k \in N$.

c) $A^{-1}u = \frac{1}{\lambda}u$, si A^{-1} existe.

d) $p(A)\,u = p(\lambda)\,u, \ \forall p \in \mathcal{P}$.

Démonstration.
Le a) est évident.

b) $A^2 u = A(Au) = \lambda Au = \lambda(\lambda u)$, et ainsi de suite.

d) Par combinaison linéaire des relations du b), on voit que

$$p(A)\, u = p(\lambda) u, \;\; \forall p \in \mathcal{P}.$$

c) $A^{-1} Au = \lambda A^{-1} u = I\, u$. $\square$

Le b) montre en plus que $\mathrm{tr}(A^k) = \sum_{i=1}^n \lambda_i^k$.

Théorème 0.2.10 *(de Cayley-Hamilton)*

$$P_A(A) = 0.$$

Nous avons introduit les vecteurs propres à droite, mais il est possible de définir des vecteurs propres à gauche.

Les valeurs propres de A^* sont les conjuguées des valeurs propres de A. En effet

$$0 = \overline{\det(A - \lambda I)} = \det(A - \lambda I)^* = \det(A^* - \overline{\lambda} I).$$

En particulier les valeurs propres de A^t sont des valeurs propres de A.

Puisque nous avons $A^* v = \overline{\lambda} v$ avec $v \neq 0$, alors, en prenant la transposée conjuguée, on obtient

$$v^* A = \lambda v^*.$$

v est dit vecteur propre à gauche de A relatif à la valeur propre λ.

Par conséquent, si A est hermitienne, les valeurs propres sont réelles et les vecteurs propres à droite et à gauche sont les mêmes. En effet $A u = \lambda u$ et $u^* A^* = u^* A = \overline{\lambda} u^*$. Donc $u^* A^* u = u^* A u = \overline{\lambda} u^* u = \lambda u^* u$. Par conséquent $\lambda = \overline{\lambda}$. Enfin $u^* A = \overline{\lambda} u^* = \lambda u^*$. Donc u est vecteur propre à gauche.

Remarque 0.2.11 *Lorsque nous parlerons par la suite de vecteur propre, il s'agira de vecteur propre à droite.*

Théorème 0.2.12 *Soit u_i un vecteur propre de A associé à λ_i. Soit v_j un vecteur propre à gauche de A associé à λ_j.*

Si $\lambda_i \neq \lambda_j$, alors $(u_i, v_j) = v_j^* u_i = 0$.

Démonstration.
A partir des relations $A u_i = \lambda_i u_i$ et $v_j^* A = \lambda_j v_j^*$, on obtient $v_j^* A u_i = \lambda_i v_j^* u_i = \lambda_j v_j^* u_i$ en multipliant la première relation à gauche par v_j^* et la seconde à droite par u_i. D'où, si $\lambda_i \neq \lambda_j$, on a $v_j^* u_i = 0$. $\square$

Faisons maintenant quelques rappels sur la diagonalisation des matrices.

Définition 0.2.13 *On dit que A est diagonalisable, s'il existe une matrice inversible P telle que $B = P^{-1}\,A\,P$ soit diagonale.*

La diagonale de B contient les valeurs propres de A et les colonnes de P sont les vecteurs propres de A. Puisque $PB = AP$, alors si $u_j = j^{\text{ème}}$ colonne de P, on trouve

$$Au_j = b_{jj}\,u_j,$$

ce qui prouve que $b_{jj} = \lambda_j$ et u_j est vecteur propre associé à λ_j.

Cette propriété annonce le résultat qui sera démontré ci-après : A est diagonalisable si et seulement s'il existe une base de vecteurs propres.

Théorème 0.2.14 *Les vecteurs propres associés à des valeurs propres distinctes sont linéairement indépendants.*

Démonstration.
Soient $\lambda_1, \ldots, \lambda_s$, s valeurs propres distinctes et $u_1, \ldots, u_s$ les s vecteurs propres associés.

On effectuera une démonstration par récurrence. u_1 est non nul. Il est donc linéairement indépendant. Pour $k < s$, on supposera que les vecteurs propres $u_1, \ldots, u_k$, associés à $\lambda_1, \ldots, \lambda_k$, sont linéairement indépendants. Prenons l'élément propre (λ_{k+1}, u_{k+1}) et supposons que $u_{k+1} = \sum_{i=1}^{k} \alpha_i\,u_i$. Alors

$$A\,u_{k+1} = \lambda_{k+1}\,u_{k+1} = \sum_{i=1}^{k} \alpha_i\,A\,u_i = \lambda_{k+1}\,\sum_{i=1}^{k} \alpha_i\,u_i = \sum_{i=1}^{k} \alpha_i\,\lambda_i\,u_i.$$

Par conséquent $\sum_{i=1}^{k} (\lambda_{k+1} - \lambda_i)\,\alpha_i\,u_i = 0$.

L'indépendance de $u_1, \ldots, u_k$ entraîne que $(\lambda_{k+1} - \lambda_i)\,\alpha_i = 0$ et donc que $\alpha_i = 0$ pour $i = 1, \ldots, k$. $\square$

Théorème 0.2.15 *A est diagonalisable si et seulement si A possède n vecteurs propres linéairement indépendants (il existe donc une base de vecteurs propres).*

Démonstration.
$Au_i = \lambda_i\,u_i$, pour $i = 1, \ldots, n$. Soit S la matrice dont les colonnes sont les u_i.

$$S = (u_1, \ldots, u_n).$$

Alors $A\,S = S\,D$.

Si S est inversible, ce qui n'est vrai que si et seulement si les vecteurs propres sont indépendants, alors $A = S\,D\,S^{-1}$ et A est diagonalisable. $\square$

Corollaire 0.2.16 *Si toutes les valeurs propres sont distinctes, A est diagonalisable.*

Démonstration.
Le théorème 0.2.14 prouve que tous les vecteurs propres sont linéairement indépendants et le théorème 0.2.15 permet de conclure que A est diagonalisable. $\square$

Théorème 0.2.17 *Soit A une matrice diagonalisable . On note λ_j, $\mathcal{U}_j$ et $\mathcal{V}_j$ respectivement une valeur propre de A, le sous-espace propre à droite associé à λ_j et le sous-espace propre à gauche associé à λ_j. Alors pour toute valeur propre λ_i de A, on peut associer à tout vecteur propre à droite u de $\mathcal{U}_i$ un vecteur propre à gauche v de $\mathcal{V}_i$ tel que $v^*\, u = 1$.*

Démonstration.
Démontrons cette propriété par l'absurde. Supposons qu'il existe $u \in \mathcal{U}_i$, $u \neq 0$ tel que $\forall v \in \mathcal{V}_i$, on ait $v^*\, u = 0$. Nous avons également $v_j^*\, u = 0$ quel que soit le vecteur propre à gauche v_j associé à une valeur propre λ_j distincte de λ_i. Or les vecteurs propres à gauche forment une base. On peut exprimer u dans cette base et donc les relations précédentes conduisent à $u^*\, u = 0$, ce qui est impossible. $\square$

Remarque 0.2.18 *En particulier dans le cas où A est hermitienne, nous avons*

$$v_i^*\, u_i = u_i^*\, u_i = \| u_i \|_2^2 \neq 0.$$

Remarque 0.2.19 *Dans le cas où A n'est pas diagonalisable, on ne peut rien dire sur $v^*\, u$.*
Exemple :

$$A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}.$$

1 est valeur propre double. Le sous-espace propre à droite associé est de dimension 1 et est généré par le vecteur $u = e_1$, premier vecteur de la base canonique de $\mathbb{R}^2$. e_2 est le vecteur propre à gauche v de norme euclidienne 1. D'où $v^*\, u = e_2^*\, e_1 = 0$.

Remarque 0.2.20 *Si A est une matrice hermitienne, alors il existe une base de vecteurs propres $\{u_i\}$ tels que $u_i^*\, u_j = \delta_{ij}$.*

0.3 METHODE DE DETERMINATION DIRECTE DU POLYNOME CARACTERISTIQUE

Cette méthode est donnée uniquement à titre de culture générale, car elle est instable. De plus, pour connaître les valeurs propres il est nécessaire d'utiliser des méthodes d'obtention des racines d'un polynôme.

Nous exposerons la méthode de Souriau-Fadeev qui est une modification d'une méthode due à Leverrier.

On pose

$$\left\{\begin{array}{lll} A_1 = A & a_1 = trA_1 & B_1 = A_1 - a_1 I. \\ A_2 = B_1 A & a_2 = \frac{1}{2} trA_2 & B_2 = A_2 - a_2 I. \\ \dots & & \\ A_n = B_{n-1} A & a_n = \frac{1}{n} trA_n & B_n = A_n - a_n I. \end{array}\right. \tag{5}$$

Théorème 0.3.1 *$B_n = 0$ et*

$$P_A(\lambda) = (-1)^n(\lambda^n - a_1\lambda^{n-1} - \dots - a_{n-1}\lambda - a_n).$$

Démonstration.
Supposons que le polynôme caractéristique s'écrive

$$P_A(\lambda) = (-1)^n(\lambda^n - c_1\lambda^{n-1} - \dots - c_{n-1}\lambda - c_n).$$

On démontre par récurrence que $c_k = a_k$.

$$c_1 = \sum_{i=1}^{n} \lambda_i = trA = trA_1 = a_1.$$

Supposons que $c_i = a_i$ pour $i = 1, \dots, k-1$ et montrons que $c_k = a_k$. Grâce aux relations (5) il est simple de voir que

$$A_k = A^k - a_1 A^{k-1} - a_2 A^{k-2} - \dots - a_{k-1} A.$$

Puisque $c_i = a_i$ pour $i = 1, \dots, k-1$, nous avons

$$A_k = A^k - c_1 A^{k-1} - c_2 A^{k-2} - \dots - c_{k-1} A.$$

Par conséquent

$$trA_k = trA^k - c_1\, trA^{k-1} - c_2\, trA^{k-2} - \dots - c_{k-1}\, trA = kc_k.$$

en utilisant les relations de Newton entre les coefficients d'un polynôme et ses racines

$$kc_k = \sum_{i=1}^{n} \lambda_i^k - c_1 \sum_{i=1}^{n} \lambda_i^{k-1} - \dots - c_{k-1} \sum_{i=1}^{n} \lambda_i \text{ pour } k = 2, \dots, n$$

et $c_1 = \sum_{i=1}^{n} \lambda_i$. Mais (5) nous donne aussi $trA_k = ka_k$. Donc $a_k = c_k$. Enfin $B_n = A_n - a_n I = P_A(A) = 0$ d'après le théorème de Cayley-Hamilton. □

Remarque 0.3.2 *Puisque $B_n = 0$, nous avons*

$$A_n = a_n I = B_{n-1} A.$$

Donc $B_{n-1} = a_n A^{-1}$, ce qui nous permet d'obtenir également l'inverse de A par cet algorithme.

Cette méthode de Souriau-Fadeev rend par contre de grands services pour les programmes de calcul formel de l'inverse d'une matrice.

0.3.1 La méthode de Jacobi

La méthode de Jacobi s'applique à une matrice A réelle symétrique. On peut mettre en oeuvre une méthode de décomposition en n'utilisant que les matrices B_k et leurs inverses. Les matrices B_k seront les matrices orthogonales de rotation plane dans le plan (e_p, e_q), où e_p et e_q sont respectivement les $p$ème et $q$ème vecteurs de la base canonique de $\mathbb{R}^n$:

$$B_k = \begin{pmatrix}
1 & & & & & & & & & & \\
& \ddots & & & 0 & & & & & & \\
& & 1 & & & & & & & & \\
& & & \cos\theta & \cdots & \cdots & \cdots & -\sin\theta & & & \\
& & & & 1 & & & & & & \\
& & & 0 & & \ddots & & 0 & & & \\
& & & & & & 1 & & & & \\
& & & \sin\theta & \cdots & \cdots & \cdots & \cos\theta & & & \\
& & & & & & & & 1 & & \\
& & & & 0 & & & & & \ddots & \\
& & & & & & & & & & 1
\end{pmatrix},$$

tous les autres éléments étant nuls.

On formera $A_{k+1} = B_k^T\, A_k\, B_k \;\forall k \in \mathbb{N}$. Puisque B_k est orthogonale, alors toutes les matrices A_k obtenues à partir de A seront symétriques.

Lorsqu'on passe de A_k à A_{k+1}, seuls sont modifiés les éléments des lignes et des colonnes p et q. Si on note $a_{ij}^{(k)}$ les éléments de A_k, nous obtenons les éléments $a_{ij}^{(k+1)}$ suivants pour A_{k+1} :

$a_{ij}^{(k+1)} = a_{ij}^{(k)}$ pour $i, j \neq p, q$,
$a_{ip}^{(k+1)} = a_{ip}^{(k)} \cos\theta + a_{iq}^{(k)} \sin\theta$ pour $i \neq p, q$,
$a_{iq}^{(k+1)} = -a_{ip}^{(k)} \sin\theta + a_{iq}^{(k)} \cos\theta$ pour $i \neq p, q$,
$a_{pp}^{(k+1)} = a_{pp}^{(k)} \cos^2\theta + 2\, a_{pq}^{(k)} \sin\theta\cos\theta + a_{qq}^{(k)} \sin^2\theta$,
$a_{qq}^{(k+1)} = a_{pp}^{(k)} \sin^2\theta - 2\, a_{pq}^{(k)} \sin\theta\cos\theta + a_{qq}^{(k)} \cos^2\theta$,

$a_{pj}^{(k+1)} = a_{pj}^{(k)} \cos\theta + a_{qj}^{(k)} \sin\theta$ pour $j \neq p, q$,
$a_{qj}^{(k+1)} = -a_{pj}^{(k)} \sin\theta + a_{qj}^{(k)} \cos\theta$ pour $j \neq p, q$,
$a_{pq}^{(k+1)} = (a_{qq}^{(k)} - a_{pp}^{(k)}) \sin\theta \cos\theta + a_{pq}^{(k)}(\cos^2\theta - sin^2\theta) = a_{qp}^{(k+1)}$.

Le but est d'annuler $a_{pq}^{(k+1)}$ et $a_{qp}^{(k+1)}$, d'où le choix de l'angle θ.

La divison de la relation précédente par $-\cos^2\theta$ donne

$$(a_{pp}^{(k)} - a_{qq}^{(k)})\mathrm{tg}\theta = a_{pq}^{(k)}(1 - \mathrm{tg}^2\theta).$$

Ensuite

$$\mathrm{tg}2\theta = \frac{2\,\mathrm{tg}\theta}{1 - \mathrm{tg}^2\theta} = \frac{2\,a_{pq}^{(k)}}{a_{pp}^{(k)} - a_{qq}^{(k)}} = \tau$$

conduit à $\mid \theta \mid \leq \frac{\pi}{4}$. Si $a_{pp}^{(k)} = a_{qq}^{(k)}$, on prendra $\theta = \frac{\pi}{4}$. Enfin les relations trigonométriques suivantes :

$$\cos^2 2\,\theta = \frac{1}{1 + \mathrm{tg}^2 2\theta} \text{ et } \cos^2\theta = \frac{\cos 2\,\theta + 1}{2}$$

permettent d'obtenir les valeurs de $\cos\theta$ et de $\sin\theta$.

$$\begin{aligned} \cos\theta &= \frac{1}{\sqrt{2}}\left(1 + \frac{1}{\sqrt{1+\tau^2}}\right)^{\frac{1}{2}}, \\ \sin\theta &= \text{signe } \tau\,\sqrt{1 - \cos^2\theta}. \end{aligned}$$

On peut remarquer que le calcul de $a_{pp}^{(k+1)}$ et $a_{qq}^{(k+1)}$ peut se simplifier. En effet

$$\begin{aligned} a_{pp}^{(k+1)} &= a_{pp}^{(k)} \cos^2\theta + 2\,a_{pq}^{(k)} \sin\theta\cos\theta + a_{qq}^{(k)} \sin^2\theta \\ &= a_{pp}^{(k)} + (2\,a_{pq}^{(k)} \cos^2\theta + (a_{qq}^{(k)} - a_{pp}^{(k)}) \sin\theta\cos\theta)\,\frac{\sin\theta}{\cos\theta} \\ &= a_{pp}^{(k)} + a_{pq}^{(k)}\,\mathrm{tg}\theta \end{aligned}$$

en utilisant la relation qui permet l'annulation de $a_{pq}^{(k+1)}$.

De la même façon on obtient

$$a_{qq}^{(k+1)} = a_{qq}^{(k)} - a_{pq}^{(k)}\,\mathrm{tg}\theta.$$

p et q varient à chaque itération. Donc $a_{pq}^{(k)}$ et $a_{qp}^{(k)}$ qui avaient été annulés lors de la $(k-1)$ème itération ne le seront plus obligatoirement après la $k$ème itération. Cependant en itérant le procédé pour tous les couples (p, q) avec $p \neq q$ on arrive peu à peu à annuler tous les éléments extradiagonaux. La matrice A_k convergera donc vers une matrice diagonale semblable à A dont les valeurs propres seront sur la diagonale.

Nous allons maintenant démontrer la convergence. Posons

$$N(A_k) = \sum_{\substack{i,j=1 \\ i \neq j}}^{n} (a_{ij}^{(k)})^2.$$

Théorème 0.3.3
$$N(A_{k+1}) = N(A_k) - 2\,(a_{pq}^{(k)})^2.$$

Démonstration.
Le calcul de A_{k+1} ne modifie que les éléments des lignes et des colonnes p et q de A_k. D'autre part les matrices sont symétriques ; par conséquent la variation de $N(A_{k+1})$ par rapport à $N(A_k)$ est égale à deux fois la variation due aux colonnes (ou aux lignes) p et q.

Les relations précédentes donnent

$$\begin{aligned}
\sum_{\substack{i=1\\ i\neq p,q}}^{n} ((a_{ip}^{(k+1)})^2 + (a_{iq}^{(k+1)})^2) &= \sum_{\substack{i=1\\ i\neq p,q}}^{n} \Big((a_{ip}^{(k)})^2 \cos^2\theta + 2\,a_{ip}^{(k)} a_{iq}^{(k)} \sin\theta\cos\theta + \\
&\quad (a_{iq}^{(k)})^2 \sin^2\theta + (a_{ip}^{(k)})^2 \sin^2\theta - \\
&\quad 2\,a_{ip}^{(k)} a_{iq}^{(k)} \sin\theta\cos\theta + (a_{iq}^{(k)})^2 \cos^2\theta\Big) \\
&= \sum_{\substack{i=1\\ i\neq p,q}}^{n} ((a_{ip}^{(k)})^2 + (a_{iq}^{(k)})^2)
\end{aligned}$$

ce qui entraîne que

$$N(A_{k+1}) - 2\,(a_{pq}^{(k+1)})^2 = N(A_k) - 2\,(a_{pq}^{(k)})^2.$$

Or θ est choisi de sorte que $a_{pq}^{(k+1)} = 0$, d'où le résultat. □

Le théorème précédent montre que

$$0 \leq N(A_{k+1}) \leq N(A_k).$$

La suite $\{N(A_k)\}$ est donc une suite décroissante bornée inférieurement et par conséquent elle est convergente.

Il est simple de trouver sa limite quand on adopte le choix (fait par Jacobi) de p et q qui correspond à l'élément extradiagonal de plus grand module. Puisque $\mid a_{pq}^{(k)} \mid = \max_{i\neq j} \mid a_{ij}^{(k)} \mid$, on a $N(A_k) \leq (a_{pq}^{(k)})^2\, N_e$, où N_e est le nombre de termes extradiagonaux. Ce nombre vaut $n^2 - n$, d'où $N(A_k) \leq (n^2 - n)(a_{pq}^{(k)})^2$.
Le théorème 0.3.3 donne donc

$$N(A_{k+1}) \leq N(A_k) - \frac{2}{n^2 - n} N(A_k) = (1 - \frac{2}{n^2 - n}) N(A_k).$$

Si on pose $a = 1 - \frac{2}{n^2-n}$, alors $0 \leq a < 1$ dès que $n \geq 2$. Donc $N(A_{k+1}) \leq a^{k+1} N(A)$, ce qui démontre que $\lim_{k\to\infty} N(A) = 0$. La matrice A_∞ est par conséquent diagonale et semblable à A. Ses vecteurs propres z_i sont les vecteurs de base e_i et donc, d'après le 3) de la remarque 7.6.1, les vecteurs propres de A sont les colonnes de la matrice

$$P_\infty = \lim_{k\to\infty} B_0\, B_1 \ldots B_k.$$

EXERCICES:

Exercice 1:

Pour résoudre le système $Ax = b$, on considère la méthode itérative suivante:
$x_0 \in \mathbb{C}^n$ et $x_{k+1} = x_k + \alpha(b - Ax_k), \quad k \in \mathbb{N}$ où $\alpha > 0$ est donné
On suppose de plus que A est à diagonale strictement dominante sur les lignes et que $a_{ii} > 0$ pour tout $i = 1, ..., n$
Montrer que si $0 < \alpha \leq 1/max\{a_{ii}, i = 1, \ldots, n\}$, la méthode itérative est convergente

Exercice 2:

Soit $A = (a_{ij})_{i,j=1,\ldots,n}$ à coefficients réels, on suppose que:
1) $a_{ij} \leq 0, \qquad \forall i, j = 1, \ldots, n, \quad i \neq j$
2) $a_{ii} > 0, \qquad \forall i = 1, \ldots, n$
3) $\sum_{i=1}^{n} a_{ij} = 0, \qquad \forall j = 1, \ldots, n$

I-Soit $\lambda \in \mathbb{R}_+^*$ et pour $x \in \mathbb{R}^n$, on définit

$$||x||_A = \sum_{i=1}^{n} a_{ii}|x_i|$$

1. Montrer que $||.||_A$ est une norme sur $\mathbb{R}^n$

2. Montrer que $A + \lambda.I$ est inversible

II-On considère le système linéaire $(A + \lambda.I)u = b \ldots (1)$

1. Ecrire l'algorithme itérative de Jacobi pour ce système.
2. Montrer que la suite $(u^{(k)})$ définie par l'algorithme vérifie:

$$||u^{(k+1)} - u^{(k)}||_A \leq (1/1 + \alpha)^k ||u^{(1)} - u^{(0)}||_A$$

où $\alpha = min_i(a_{ii})$
3. Montrer que la suite $(u^{(k)})$ est de Cauchy et qu'elle converge vers la solution de (1)

Exercice 3:

Soit A une matrice carrée à coefficients complexe telle que $\rho(A) < 1$,I la matrice identité.
On pose: $B_m = I + A + \ldots + A^m$
Montrer que:

$(I - A)$ est inversible et que $\lim_{m\to+\infty} B_m = (I - A)^{-1}$

Exercice 4:

Soit $||.||$ une norme matricielle et B une matrice carrée à coefficients réels telle que $||B|| < 1$. Montrer que:
1. $I + B$ est inversible
2. $||(I + B)^{-1}|| \leq 1/1 - ||B||$
3. En déduire que si $I + B$ est singulière, alors $||B|| \geq 1$

Exercice 5:

Soit A une matrice carrée d'ordre n, $A = (a_{ij})_{i,j=1\ldots,n}$ régulière et $b \in rrm^n$
On veut résoudre le système linéaire $Ax = b$, on note D la matrice diagonale d'éléments $(a_{ii})_{i=1\ldots,n}$
Pour $\alpha \neq 0$, soit la méthode itérative définie par:

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

1. Montrer que si la méthode (1) converge, sa limite est solution du système $Ax = b$

2. Exprimer les coefficients de la matrice $D^{-1}A$ en fonction de ceux de A

3. On suppose que $0 < \alpha \leq 1$ et que A est à diagonale strictement dominante

3.1. Montrer que $||I - \alpha.D^{-1}A||_\infty < 1$

3.2. En déduire que la méthode itérative (1) est convergente

4. Que peut-on dire du cas $\alpha = 1$

Exercice 6:

Soit A une matrice carrée d'ordre n à coefficients dans $\mathbb{C}$ et inversible pour A_0 donnée, on définit la suite (A_k) par:

$$A_{k+1} = 2A_k - A_k.A.A_k, pour k \in \mathbb{N}$$

1.Montrer que la suite (A_k) converge vers A^{-1} si et seulement si $\rho(I - A.A_0) < 1$

2.Montrer que $A^{-1} - A_{k+1} = (A^{-1} - A_k).A.(A^{-1} - A_k)$

3.Montrer que dans ce cas:

$$\exists M > 0 \text{ tel que } ||A_{k+1} - A^{-1}|| \leq M||A_k - A^{-1}||^2, k \in \mathbb{N}$$

Exercice 7:

Soit $A = (a_{ij})_{i,j=1,...,n}$ à coefficients réels définie par:

$a_{ii} = i + 1, i = 1, ..., n.$ $a_{i+1,i} = 1, i = 1, ..., n - 1.$ $a_{i,i+1} = -i, i = 1, ..., n - 1$

Les autres termes étant nuls, et soit le système linéaire $Ax = b$

1. Montrer que la Méthode de Jacobi peut s'écrire sous la forme

$$x^{(k+1)} = (I - D^{-1}A)x^{(k)} + D^{-1}b$$

2. Calculer La Matrice d'itération de Jacobi pour ce système

3. Montrer que son rayon spectral est inférieur à 1

4. Calculer la matrice d'itération G de Gauss-Seidel

5. Montrer que le polynôme caractéristique de G s'écrit:

$P_G(\lambda) = \lambda^n.det(I - L - \frac{1}{\lambda}U)$ avec $L = D^{-1}E$ et $U = D^{-1}F$

Exercice 8:

Soit $A \in M_3(\mathbb{R})$ définie par: $A = I - E - F$ avec:
$E = \begin{pmatrix} 0 & -2 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$ $F = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ -1 & -1 & 0 \end{pmatrix}$ $I = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$
1. Montrer que A est inversible
2. Soit $0 < \omega < 2$, montrer que $((1/\omega).I - E)$ est inversible ssi $\omega \neq \sqrt{2}/2$
Pour $0 < \omega < 2$ et $\omega \neq \sqrt{2}/2$, on considère la méthode itérative (pour résoudre $Ax = b$) suivante:

$$((1/\omega).I - E)X^{(n+1)} = (F + ((1-\omega)/\omega).I)X^{(n)} + b$$

Posons $L_\omega = ((1/\omega).I - E)^{-1}(F + ((1-\omega)/\omega).I)$
3. Calculer en fonction de ω les valeurs propres et le rayon spectral de L_ω
4. Pour quelles valeurs de ω la méthode est-elle convergente?