

Table des matières

I	Résolution des systèmes lineaires	2
1	Rappels sur les matrices	3
1-1	Valeurs, vecteurs propres et valeurs singulières	4
2	Normes matricielles	4
2-1	Normes matricielles induites	5
2-2	Convergence des matrices	6
3	Conditionnement de matrices	7
4	Méthodes directes pour la résolution des systèmes linéaires	11
4-1	Méthode de Gauss	11
4-2	Système quelconque (inversible)	12
4-3	Factorisation LU	15
4-4	Stratégie de pivot dans Gauss	17
4-5	Algorithme de Crout	19
4-6	Méthode de Cholesky	20
4-7	Algorithme de Cholesky	21
5	Méthodes itératives	22
5-1	Convergence des méthodes itératives	23
5-2	Vitesse de convergence, test d'arrêt et nombre d'itérations	23
6	Méthodes de Jacobi, de Gauss Seidel et de Relaxation	24
6-1	Résultats de convergence pour les 3 méthodes	26

Chapitre I

Résolution des systèmes lineaires

Deux problèmes fondamentaux interviennent en analyse matricielle:

- La résolution d'un système linéaire de m équations à n inconnues de la forme $\sum_{j=1}^n a_{ij}x_j = b_i$ pour $1 \leq i \leq m$, qu'on peut écrire sous forme matricielle $AX = b$.
- Le calcul des valeurs et vecteurs propres d'une matrice A c.à.d $AX = \lambda X$ ($X \neq 0$).

On s'intéresse au premier point pour une matrice carrée A à coefficients réels ou complexes. Dans ce cas on est assuré de l'existence et l'unicité de la solution si une des conditions équivalentes suivantes est remplie:

1. A est inversible.
2. $\text{rg}(A) = n$
3. Le système homogène $AX = 0$ admet seulement la solution nulle.

Théoriquement on peut calculer X grâce aux formules de Cramer $X_j = \frac{\Delta_j}{\det A}$ où Δ_j est le déterminant obtenue en remplaçant la $j^{\text{ème}}$ colonne par le second membre b .

C'est une méthode en général à éviter puisqu' un det requiert $(n - 1)n!$ multiplications

$$(\det A = \sum_{\sigma \in \sigma_n} \varepsilon(\sigma) a_{\sigma(1)1} a_{\sigma(2)2} \dots a_{\sigma(n)n}) \quad \sigma_n, \varepsilon(\sigma)$$

représentent resp. l'ensemble des permutations sur $\{1, 2, \dots, n\}$ et la signature de σ . Ainsi pour résoudre un système linéaire de seulement 50 équations, un ordinateur capable d'effectuer 10^9 Flops par secondes mettrait $9.6 \cdot 10^{47}$ années à résoudre ce système. On propose donc d'autres méthodes qui fournissent la solution en un nombre fini d'étapes dites méthodes directes, ou bien de méthodes itératives qui théoriquement nécessitent un nombre infini d'étapes.

1 Rappels sur les matrices

Soit X un vecteur de \mathbb{R}^n ou \mathbb{C}^n , $A = (a_{ij})_{1 \leq i, j \leq n}$ une matrice carée $n \times n$ à coefficients réels ou complexes, le vecteur $b = AX$ est défini par : $b = (b_1, b_2, \dots, b_n)^t$ où $b_i = \sum_{j=1}^n a_{ij} X_j$ pour $1 \leq i \leq n$.

- **Matrice symétrique:**

A est dite symétrique si et seulement si $A = A^t$ où $A^t = (a_{ij}^t)_{1 \leq i, j \leq n} = (a_{ji})_{1 \leq i, j \leq n}$

- **Matrice Hermitienne:**

A est dite hermitienne si et seulement si $A = A^*$ où $A^* = (a_{ij}^*)_{1 \leq i, j \leq n} = (\bar{a}_{ji})_{1 \leq i, j \leq n}$

- **Matrice normale:** A est dite normale si et seulement si $A^*A = AA^*$.

- **Matrice unitaire:**

A est dite unitaire si et seulement si $A^*A = AA^* = I_n$. La matrice est alors inversible et on a $A^{-1} = A^*$.

Dans le cas réel on parle de:

- **Matrice orthogonale:** $AA^t = A^tA = I_n$ et dans ce cas $A^{-1} = A^t$.

- **Matrice symétrique définie positive (resp positive):**

Une matrice symétrique, ou hermitienne est dite positive (resp. définie positive) si et seulement si $\langle AX, X \rangle \geq 0 \quad \forall X \in \mathbb{R}^n \setminus \{0\}$ (resp. $\langle AX, X \rangle > 0 \quad \forall X \in \mathbb{R}^n \setminus \{0\}$).

Pour hermitienne on remplace $\mathbb{R}^n \setminus \{0\}$ par $\mathbb{C}^n \setminus \{0\}$.

- **Matrice diagonales:** Une matrice A est dite diagonale si et seulement si $a_{ij} = 0 \quad \forall i \neq j$.

- **Matrice triangulaire inférieure, supérieure:**

Une matrice A est dite triangulaire inférieure si $a_{ij} = 0$ pour $1 \neq i < j \neq n$.

Une matrice A est dite triangulaire supérieure si $a_{ij} = 0$ pour $1 \neq j < i \neq n$.

Ainsi A est diagonale si et seulement si A est triangulaire inférieure et supérieure.

- **Matrices bandes:**

Une matrice est dite matrice bande si et seulement si $a_{ij} = 0$ pour $j < i - d$ et $j > i + d$. d est appelée demi largeur de bande.

- **Matrices tridiagonales, pentadiagonale:**

Pour $d = 1$ on parle de matrice tridiagonale (si $d = 2$ on dit qu'on a une matrice pentadiagonale).

Propriétés 1.1.

- a) $\forall \alpha \in \mathbb{C}^* \quad (\alpha A)^{-1} = \frac{1}{\alpha} A^{-1}$.
- b) $(AB)^{-1} = B^{-1}A^{-1}$.
- c) $(AB)^t = B^t A^t$.

- d) $(A^t)^{-1} = (A^{-1})^t$

Proposition 1.1.

Soit A une matrice carrée à coefficients complexes ou réels, si A est une matrice triangulaire supérieure (resp triangulaire inférieure) dont les éléments de la diagonale sont tous non nuls alors A est inversible et A^{-1} est aussi triangulaire supérieure (resp. inférieure)

Définition 1.1.

Une matrice carrée A est dite régulière ou non singulière si elle est inversible.

Théorème 1.1.

Soit A une matrice carrée à coefficients réels ou complexes,

- il existe alors une matrice unitaire U telle que U^*AU (ou $U^{-1}AU$) est une matrice triangulaire.
- si A est une matrice normale alors il existe U unitaire telle que $U^{-1}AU$ soit diagonale.
- si A est symétrique alors il existe O orthogonale telle que $O^{-1}AO$ soit diagonale.

1-1 Valeurs, vecteurs propres et valeurs singulières

Proposition 1.2.

Si $AU = \lambda U, U \neq 0$ alors

1. $(A - \mu I)U = (\lambda - \mu)U$.
2. $A^k U = \lambda^k U$
3. Si A est inversible $A^{-1}U = \frac{1}{\lambda}U$
4. $P(A)U = P(\lambda)U$ pour tout polynôme P .

Définition 1.2.

λ est dite valeur singulière de A si $\lambda \geq 0$ et λ^2 est valeur propre de A^*A .

2 Normes matricielles

Définition 2.1.

On appelle norme matricielle toute application notée $\|\cdot\|$ de $M_{n,n}(\mathbb{K})$ (ensemble des matrices à coefficients dans $\mathbb{K} = \mathbb{R}$ ou \mathbb{C}) à valeurs dans \mathbb{R}^+ qui vérifie.

1. $\forall A \in M_{n,n}(\mathbb{K}), \|A\| = 0 \iff A = 0$ (matrice nulle).
2. $\forall A \in M_{n,n}(\mathbb{K})$ et $\forall \alpha \in \mathbb{K} \|\alpha A\| = |\alpha| \|A\|$
3. $\forall A, B \in M_{n,n}(\mathbb{K})$ on a $\|A + B\| \leq \|A\| + \|B\|$
4. $\forall A, B \in M_{n,n}(\mathbb{K})$ on a $\|AB\| \leq \|A\| \|B\|$

Exemple 2.1.

Norme de Shur ou norme de Froebenius noté $\| \cdot \|_S$ ou $\| \cdot \|_F$ définie par:

$$\begin{aligned} \|A\|_S &= \left(\sum_{i,j=1}^n |a_{ij}|^2 \right)^{1/2} \\ \|AB\|_S^2 &= \left(\sum_{i,j=1}^n |(AB)_{ij}|^2 \right) \leq \sum_{i,j=1}^n \left(\sum_{k=1}^n |a_{ik}| |b_{kj}| \right)^2 \\ &\leq \sum_{i,j=1}^n \left(\sum_{k=1}^n |a_{ik}|^2 \right) \left(\sum_{k=1}^n |b_{kj}|^2 \right) \quad (\text{inégalité de Schwartz}) \\ &= \sum_{i,k=1}^n a_{ik}^2 \sum_{j,k=1}^n b_{kj}^2 = \|A\|_S^2 \|B\|_S^2 \end{aligned}$$

2-1 Normes matricielles induites

Sont les normes définies à partir des normes vectorielles par:

$$\|A\| = \sup_{\|x\|=1} \|Ax\| = \sup_{\|x\| \neq 0} \frac{\|Ax\|}{\|x\|}$$

On garde la même notation pour les normes vectorielles. On vérifie facilement que c'est une norme matricielle.

$$\begin{aligned} \|ABx\| &\leq \|A\| \|Bx\| \leq \|A\| \|B\| \text{ pour } \|x\| = 1 \\ \implies \sup_{\|x\|=1} \|ABx\| &\leq \|A\| \|B\| \text{ soit } \|AB\| \leq \|A\| \|B\| \end{aligned}$$

$\| \cdot \|_S$ n'est pas une norme induite ($\|I_n\|_S = \sqrt{n}$).

Rappelons les normes vectorielles suivantes:

$$\|x\|_p = \left(\sum_{i=1}^n |x_i|^p \right)^{1/p} \quad 1 \leq p < \infty \quad \text{et} \quad \|x\|_\infty = \max_{1 \leq i \leq n} |x_i|$$

$\|x\|_1$, $\|x\|_2$ et $\|x\|_\infty$ sont les plus utilisées, on montre que les normes matricielles associées vérifient:

Proposition 2.1.

$$\|A\|_\infty = \max_{1 \leq i \leq n} \left(\sum_{j=1}^n |a_{ij}| \right)$$

$$\|A\|_1 = \max_{1 \leq j \leq n} \left(\sum_{i=1}^n |a_{ij}| \right)$$

$$\|A\|_2 = \sqrt{\rho(A^*A)} = \sqrt{\rho(AA^*)}$$

$$\|A\|_2 = \rho(A) \quad \text{pour une matrice hermetienne}$$

Où $\rho(A)$ est le rayon spectral de A , on montre aussi que

Proposition 2.2.

Si A est une matrice carrée $n \times n$ de norme quelconque $\|\cdot\|$ alors:

- $\rho(A) \leq \|A\|$
- $\rho(A) \leq \inf(\max_{1 \leq i \leq n}(\sum_{j=1}^n |a_{ij}|), \max_{1 \leq j \leq n}(\sum_{i=1}^n |a_{ij}|))$

Preuve:

Soit λ une valeur propre de A , $\exists u \neq 0 / Au = \lambda u$

$$|\lambda| \|u\| \leq \|A\| \|u\| \implies |\lambda| \leq \|A\|$$

$$\implies \rho(A) = \max\{\lambda / \lambda \text{ est valeur propre de } A\} \leq \|A\|$$

Pour le deuxième point il suffit de prendre $\|\cdot\| = \|\cdot\|_1$ ou $\|\cdot\|_\infty$

2-2 Convergence des matrices

On rappelle que $\lim_{k \rightarrow \infty} A^k = 0 \iff \lim_{k \rightarrow \infty} a_{ij}^{(k)} = 0 \iff \lim_{k \rightarrow \infty} \|A^k\| = 0$ où $a_{ij}^{(k)}$ sont les coefficients de la matrice A^k .

Théorème 2.1.

Soit A une matrice carrée d'ordre n , les conditions suivantes sont équivalentes.

- a) $\lim_{k \rightarrow \infty} A^k = 0$.
- b) $\lim_{k \rightarrow \infty} A^k x = 0 \quad \forall x$.
- c) $\rho(A) < 1$.
- d) $\|A\|_p < 1$ pour au moins un $p \in \{1, 2, \infty\}$.

Remarque 2.1.

On sait que si $|\alpha| < 1$ alors

$$\frac{1}{1-\alpha} = 1 + \alpha + \alpha^2 + \alpha^3 + \dots = \sum_{k=0}^{\infty} \alpha^k$$

on a un résultat semblable en remplaçant α par une matrice A avec la condition $\rho(A) < 1$ et on a

Théorème 2.2.

Si $A \in M_{n,n}(\mathbb{K})$ telle que $\rho(A) < 1$ alors $I - A$ et $I + A$ sont inversibles et on a:

$$(I - A)^{-1} = \sum_{k=0}^{\infty} A^k \quad \text{et} \quad \|(I - A)^{-1}\| \leq \frac{1}{1 - \|A\|}$$

c.à.d. la série $\sum_{k \geq 0} A^k$ converge vers $(I - A)^{-1}$ si et seulement si $\rho(A) < 1$.

Preuve:

Les valeurs propres de A sont les λ_i $1 \leq i \leq n$.

Les valeurs propres de $I \pm A$ sont les $1 \pm \lambda_i$ $1 \leq i \leq n$.

$$|\lambda_i| \leq \rho(A) < 1 \implies 1 \pm \lambda_i \neq 0 \text{ donc } I \pm A \text{ inversible}$$

$$\text{et on a } (I - A)(I + A + A^2 + \dots + A^k) = I - A^{k+1} \implies (I - A)^{-1} - (I - A)^{-1}A^{k+1} = \sum_{j=0}^k A^j$$

comme $\rho(A) < 1$ donc $\lim A^{k+1} \rightarrow 0$ c.à.d. $(I - A)^{-1} = \sum_{k=0}^{\infty} A^k$

$$(I - A)(I - A)^{-1} = (I - A)^{-1} - A(I - A)^{-1} \implies (I - A)^{-1} = I + A(I - A)^{-1} \implies$$

$$\|(I - A)^{-1}\| \leq 1 + \|A\| \|(I - A)^{-1}\|$$

$$\implies \|(I - A)^{-1}\| \leq \frac{1}{1 - \|A\|}$$

3 Conditionnement de matrices

Soit la matrice de Hilbert définie par:

$$H = (a_{ij})_{1 \leq i, j \leq n} \text{ où } a_{ij} = \frac{1}{i + j - 1}$$

Pour $n = 4$,

$$H = \begin{pmatrix} 1 & \frac{1}{2} & \frac{1}{3} & \frac{1}{4} \\ \frac{1}{2} & \frac{1}{3} & \frac{1}{4} & \frac{1}{5} \\ \frac{1}{3} & \frac{1}{4} & \frac{1}{5} & \frac{1}{6} \\ \frac{1}{4} & \frac{1}{5} & \frac{1}{6} & \frac{1}{7} \end{pmatrix}$$

Pour

$$b^t = \left(\frac{25}{12}, \frac{77}{60}, \frac{57}{60}, \frac{319}{420} \right) \simeq (2.0833, 1.2833, 0.9500, 0.7599)$$

on calcule $x = H^{-1}b$ et on vérifie que $x^t = (1, 1, 1, 1)$.

Si on change légèrement b en \bar{b} où

$$\bar{b}^t = (2.1, 1.3, 1.0, 0.8)$$

la solution $\bar{x} = H^{-1}\bar{b}$ est donnée par

$$\bar{x}^t = (5.6, -48, 114, -70) \text{ loin de la solution!!!!}$$

Définition 3.1.

le conditionnement mesure la dépendance de la solution d'un problème numérique par rapport aux données du problème, ceci afin de contrôler la validité d'une solution calculée par rapport à ces données. En effet, les données d'un problème numérique dépendent en général de mesures expérimentales et sont donc entachées d'erreurs.

Soit un problème $P : \mathbb{R}^n \rightarrow \mathbb{R}$. Soit aussi une variable perturbée

$\hat{x}_i = x_i(1 + \epsilon_i)$, avec $|\epsilon_i| < \epsilon$. Alors, la condition k du problème est le plus petit nombre tel que :

$$\frac{|P(\hat{x}) - P(x)|}{|P(x)|} \leq k\epsilon + o(\epsilon).$$

Le problème P est bien conditionné si k n'est pas très grand (par rapport à ϵ (précision de la machine)). Sinon, ce problème P est mal conditionné.

Ainsi si on considère une petite variation δb de b , la solution correspondante est $X + \delta X$ solution de $A(X + \delta X) = b + \delta b$ où $AX = b$. On en déduit que $A\delta X = \delta b \implies \delta X = A^{-1}\delta b$

$$\|\delta X\| \leq \|A^{-1}\| \|\delta b\| \quad \text{et} \quad \|b\| \leq \|A\| \|X\|$$

donc $\|\delta X\| \|b\| \leq \|A^{-1}\| \|\delta b\| \|A\| \|X\|$ soit

$$\frac{\|\delta X\|}{\|X\|} \leq \|A^{-1}\| \|A\| \frac{\|\delta b\|}{\|b\|}$$

(optimale puisqu'on peut avoir l'égalité)

Définition 3.2.

Nous appelons conditionnement d'une matrice A relativement à une norme matricielle induite $\|\cdot\|$ le nombre $\|A\| \|A^{-1}\|$ noté $K(A)$ ou $\text{cond}(A)$.

$$K(A) = \|A\| \|A^{-1}\|$$

Pour l'exemple précédent

$$K_1(A) = K_\infty(A) = \|A\|_1 \|A^{-1}\|_1 \simeq 28375 \gg 1$$

$$K_2(A) = \|A\|_2 \|A^{-1}\|_2 \simeq 15514 \gg 1.$$

Nous remarquons ainsi que le conditionnement pour une telle matrice est grand indépendamment de la norme choisie.

On montre aussi que si A est perturbée en $A + \delta A$, la solution X est perturbée par δX et on a :

$$\frac{\|\delta X\|}{\|X + \delta X\|} \leq \|A^{-1}\| \|A\| \frac{\|\delta A\|}{\|A\|}$$

En effet $(A + \delta A)(X + \delta X) = b$

$\implies \delta X = -A^{-1}(\delta A)(X + \delta X)$ d'où le résultat. On montre aussi les propriétés suivantes :

Proposition 3.1.

Soit $A \in M_{n,n}(\mathbb{K})$ inversible et $\|\cdot\|$ une norme matricielle induite, on a alors :

1. $K(\alpha A) = K(A)$ pour tout $\alpha \neq 0$.
2. $K(A^{-1}) = K(A)$
3. $K(I_n) = 1$.
4. $K(A) \geq 1$

5. $K_2(A) = \frac{\max_i(\mu_i)}{\min_i(\mu_i)}$, les μ_i sont les valeurs singulières de A .

6. Si A est normale $K_2(A) = \frac{\max_i(|\lambda_i|)}{\min_i(|\lambda_i|)}$, les λ_i sont les valeurs propres de A .

7. Si A est unitaire $K_2(A) = 1$.

Preuve:

4) $K(I_n) = 1$, $AA^{-1} = I_n$ soit

$$1 = K(AA^{-1}) = \|AA^{-1}\|(AA^{-1})^{-1}\| \\ \leq (\|A\|\|A^{-1}\|)^2 = K(A)^2$$

Donc $K(A) \geq 1$

5)

$$K_2^2(A) = \|A\|_2^2 \|A^{-1}\|_2^2 = \rho(AA^*)\rho((A^{-1})^*A^{-1}) \\ = \max\{|\beta_i| / \beta_i \text{ val propre de } A^*A\} \\ \times \max\{|\beta'_i| / \beta'_i \text{ val propre de } (A^{-1})^*A^{-1} = (AA^*)^{-1}\} \\ = \frac{\max\{\mu_i^2 / \mu_i \text{ val singulière de } A\}}{\min\{\mu_i^2 / \mu_i \text{ val singulière de } A\}}$$

et donc

$$K_2(A) = \frac{\max(\mu_i)}{\min(\mu_i)}$$

$$\rho(A^*A) = \rho(AA^*)$$

en effet soit λ valeur propre de A^*A tel que $|\lambda| = \rho(A^*A)$

$\implies \exists u \neq 0$ vecteur propre de A^*A tel que $A^*Au = \lambda u$

$\implies (AA^*)Au = \lambda Au$ $Au \neq 0$

donc λ est valeur propre de AA^*

$\implies \rho(AA^*) \geq \lambda = \rho(A^*A)$

et $\rho(AA^*) = \rho((A^*)^*A^*) \leq \rho(A^*A)$

6) si A est normale alors il existe U unitaire tel que $U^*AU = D = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_n)$

les λ_i sont les valeurs propres de A ,

$\|A\|_2 = \|U^*AU\|_2$ et on a

$$\|A\|_2 = \rho(A^*A) = \rho(UD^*DU^*) = \rho(D^*D) = \max_i |\lambda_i(A)|^2 = \rho(A)^2$$

c.à.d $K_2(A) = \|A\|_2 \|A^{-1}\|_2 = \rho(A)\rho(A^{-1}) = \frac{\max_i(|\lambda_i|)}{\min_i(|\lambda_i|)}$

7) Si A est unitaire $A^*A = AA^* = I_n$, $K_2(A) = \|A\|_2 \|A^{-1}\|_2$ or $\|A\|_2 = \rho(A^*A) = 1 \implies K_2(A) = 1$

Conséquences:

- La propriété 6) montre que si le spectre de A est large alors A est mal conditionné.
- La propriété 7) justifie l'utilisation de matrices unitaires dans certains procédés de factorisation (ex: Housholder) ainsi le conditionnement du système final est meilleur que celui initial.
- Un système linéaire à matrice unitaire ou orthogonal est bien conditionné.
- Le conditionnement d'une matrice ne change pas si on la multiplie par un scalaire par contre on peut le diminuer en multipliant certaines lignes ou colonnes par des coefficients non nuls c'est ce qu'on appelle équilibrage de matrices. C'est une technique de préconditionnement très utilisée.

Préconditionnement:

Le préconditionnement est une méthode appliquée à une matrice A (généralement mal conditionnée) afin d'améliorer son conditionnement. Pour résoudre $AX = b$ on multiplie le système par une matrice inversible P , $PAX = Pb$. P est choisie de telle sorte que PA soit mieux conditionné (au meilleur cas $P = A^{-1}$), le plus souvent on cherche une matrice P facile à inverser.

Exemple 3.1. Autre exemple de système mal conditionné.

$$\begin{pmatrix} 10 & 7 & 8 & 7 \\ & 5 & 6 & 5 \\ \text{sym} & & 10 & 9 \\ & & & 10 \end{pmatrix} \begin{pmatrix} X_1 \\ X_2 \\ X_3 \\ X_4 \end{pmatrix} = \begin{pmatrix} 32 \\ 23 \\ 33 \\ 31 \end{pmatrix} \text{ de solution } \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}$$

Le système perturbé:

$$A \begin{pmatrix} X_1 + \delta X_1 \\ X_2 + \delta X_2 \\ X_3 + \delta X_3 \\ X_4 + \delta X_4 \end{pmatrix} = \begin{pmatrix} 32.1 \\ 22.9 \\ 33.1 \\ 30.9 \end{pmatrix} \rightarrow \delta b = \begin{pmatrix} 0.1 \\ -0.1 \\ 0.1 \\ -0.1 \end{pmatrix} \text{ donne } \delta X = \begin{pmatrix} 8.2 \\ -13.6 \\ 3.5 \\ -2.1 \end{pmatrix}$$

et donc une solution $(9.2, -10.6, 4.5, -1.1)^t$ une erreur relative de $\frac{\|\delta b\|}{\|b\|} \simeq \frac{1}{200}$ entraîne une erreur relative $\frac{\|\delta X\|}{\|X\|} \simeq 10$ (erreur amplifiée de 2000!!!!) (pourtant matrice sym det = 1

Justification: les valeurs propres de A sont 0.01015, 0.8431, 3.850 et 30.2887. ($K_2(A) \simeq 2984!!$)

De même si on perturbe A soit:

$$A + \delta A = \begin{pmatrix} 10 & 7 & 8.1 & 7.2 \\ 7.08 & 5.04 & 6 & 5 \\ 8 & 5.98 & 9.89 & 9 \\ 6.99 & 4.99 & 9 & 9.98 \end{pmatrix} \begin{pmatrix} X_1 + \delta X_1 \\ X_2 + \delta X_2 \\ X_3 + \delta X_3 \\ X_4 + \delta X_4 \end{pmatrix} = b$$

La solution est: $(-81, 137, -34, 22)^t$!!!!!

On démontre le théorème suivant:

Théorème 3.1.

Soit $A \in M_{n,n}(\mathbb{K})$ inversible, δA une perturbation de A telle que $\|A^{-1}\|\|\delta A\| < 1$, alors $A + \delta A$ est inversible et on a pour δb une perturbation de b et δX une perturbation de X c.à.d $(A + \delta A)(X + \delta X) = b + \delta b$ où $AX = b$. On a les estimations suivantes:

- a) $\frac{\|\delta X\|}{\|X\|} \leq \frac{K(A)}{1 - K(A)\frac{\|\delta A\|}{\|A\|}} \left(\frac{\|\delta A\|}{\|A\|} + \frac{\|\delta b\|}{\|b\|} \right)$ pour toute norme induite.
- b) pour $\delta A = 0$, $\frac{1}{K(A)} \frac{\|\delta b\|}{\|b\|} \leq \frac{\|\delta X\|}{\|X\|} \leq K(A) \frac{\|\delta b\|}{\|b\|}$

Preuve:

$A + \delta A = A(I + A^{-1}\delta A)$ inversible ($\|A^{-1}\|\|\delta A\| < 1$) et on a :

$$\|(I + A^{-1}\delta A)^{-1}\| \leq \frac{1}{1 - \|A^{-1}\delta A\|} \leq \frac{1}{1 - \|A^{-1}\|\|\delta A\|}$$

en plus $(A + \delta A)(X + \delta X) = b + \delta b$ et $AX = b \implies$

$$(A + \delta A)\delta X + \delta AX = \delta b \text{ soit } \delta X = (A + \delta A)^{-1}(\delta b - \delta AX) \\ = (I + A^{-1}\delta A)^{-1}A^{-1}(\delta b - \delta AX)$$

$$\|\delta X\| \leq \frac{\|A^{-1}\|}{1 - \|A^{-1}\|\|\delta A\|} (\|\delta b\| + \|\delta A\|\|X\|)$$

$$\frac{\|\delta X\|}{\|X\|} \leq \frac{\|A^{-1}\|}{1 - K(A)\frac{\|\delta A\|}{\|A\|}} \left(\frac{\|\delta b\|}{\|X\|} + \|\delta A\| \right)$$

$$\leq \frac{K(A)}{1 - K(A)\frac{\|\delta A\|}{\|A\|}} \left(\frac{\|\delta b\|}{\|A\|\|X\|} + \frac{\|\delta A\|}{\|A\|} \right)$$

et comme $\|b\| \leq \|A\|\|X\|$ alors $\frac{1}{\|A\|\|X\|} \leq \frac{1}{\|b\|}$ et donc on a le résultat.

$$\text{b) } \|\delta b\|\|X\| \leq \|\delta X\|\|A\|\|X\| \leq \|\delta X\|\|A\|\|A^{-1}\|\|b\| \text{ donc } \frac{\|\delta b\|}{\|b\|} \leq K(A) \frac{\|\delta X\|}{\|X\|}$$

$$\frac{1}{K(A)} \frac{\|\delta b\|}{\|b\|} \leq \frac{\|\delta X\|}{\|X\|}$$

4 Méthodes directes pour la résolution des systèmes linéaires

4-1 Méthode de Gauss

Systèmes triangulaires

Considérons un système triangulaire inférieure $L = A$, la résolution du système linéaire $LX = b$ conduit au système:

$$\begin{cases} l_{11}x_1 & = b_1 \\ l_{21}x_1 + l_{22}x_2 & = b_2 \\ l_{31}x_1 + l_{32}x_2 + l_{33}x_3 & = b_3 \\ \dots & \dots \\ l_{n1}x_1 + \dots + l_{nn}x_n & = b_n \end{cases}$$

L est inversible une fois que $\det L = \prod_{i=1}^n l_{ii} \neq 0$, on applique une méthode de descente et on obtient :

$$\left\{ \begin{array}{l} x_1 = \frac{b_1}{l_{11}} \\ x_2 = \frac{1}{l_{22}}(b_2 - l_{21}x_1) \\ \cdot \\ x_i = \frac{1}{l_{ii}}(b_i - \sum_{j=1}^{i-1} l_{ij}x_j) \\ \cdot \\ x_n = \frac{1}{l_{nn}}(b_n - \sum_{j=1}^{n-1} l_{nj}x_j) \end{array} \right.$$

Cette résolution demande :

$$\begin{aligned} & n \text{ divisions} + (1 + 2 + \dots + (n - 1)) \text{ soustractions} \\ & + (1 + 2 + \dots + (n - 1)) \text{ multiplications.} \\ & \text{c.à.d. } n + \frac{n(n-1)}{2} + \frac{n(n-1)}{2} = n^2 \text{ opérations.} \end{aligned}$$

Même raisonnement pour un système triangulaire supérieure.

4-2 Système quelconque (inversible)

On a à résoudre $AX = b$ où $A = (a_{ij})_{1 \leq i, j \leq n}$, $X = (x_1, x_2, \dots, x_n)^t$, en posant :

$$A^{(1)} = A, b^{(1)} = b$$

On suppose $a_{11} = a_{11}^{(1)} \neq 0$ (qu'on appelle 1^{er} pivot), après (à partir de la 2^{ème} ligne c.à.d pour $i \geq 2$) on multiplie la 1^{ère} ligne par $m_{i1} = \frac{a_{i1}}{a_{11}}$ et on la retranche de la ligne i ($l_i \rightarrow l_i - m_{i1}l_1$), on obtient :

$$\left\{ \begin{array}{l} a_{11}^{(1)}x_1 + a_{12}^{(1)}x_2 + a_{1n}^{(1)}x_n = b_1^{(1)} \\ 0 + (a_{22}^{(1)} - m_{21}a_{12}^{(1)})x_2 + (a_{2n}^{(1)} - m_{21}a_{1n}^{(1)})x_n = (b_2^{(1)} - m_{21}b_1^{(1)}) \\ \cdot \\ 0 \quad \cdot \\ 0 + (a_{n2}^{(1)} - m_{n1}a_{12}^{(1)})x_2 + (a_{nn}^{(1)} - m_{n1}a_{1n}^{(1)})x_n = (b_n^{(1)} - m_{n1}b_1^{(1)}) \end{array} \right.$$

On obtient ainsi un système de la forme $A^{(2)}X = b^{(2)}$ où

$$A^{(2)} = \begin{pmatrix} a_{11}^{(2)} & a_{12}^{(2)} & \dots & \cdot & a_{1n}^{(2)} \\ 0 & a_{22}^{(2)} & \cdot & \cdot & a_{2n}^{(2)} \\ 0 & & & & \\ 0 & & & & \\ 0 & a_{n2}^{(2)} & \cdot & \cdot & a_{nn}^{(2)} \end{pmatrix} \quad b^{(2)} = \begin{pmatrix} b_1^{(2)} \\ b_2^{(2)} \\ \cdot \\ \cdot \\ b_n^{(2)} \end{pmatrix}$$

$$a_{ij}^{(2)} = \begin{cases} a_{ij}^{(1)} & i = 1 \\ a_{ij}^{(1)} - m_{i1}a_{1j}^{(1)} & i, j = 2, \dots, n \end{cases} \quad \text{on pose } m_{i1} = \frac{a_{i1}^{(1)}}{a_{11}^{(1)}} \text{ pour } i = 2, \dots, n$$

$$b_i^{(2)} = \begin{cases} b_i^{(1)} & i = 1 \\ b_i^{(1)} - m_{i1}b_1^{(1)} & i = 2, \dots, n \end{cases}$$

Nombre d'opérations:

Pour une telle transformation on a besoin de $n - 1$ divisions, $(n - 1)^2 + (n - 1)$ multiplications et $(n - 1)^2 + (n - 1)$ additions.

Par un raisonnement par récurrence, on suppose qu' à l'étape $(k - 1)$ nous avons construit le système $A^{(k)}X = b^{(k)}$

$$\begin{cases} a_{11}^{(1)}x_1 + a_{12}^{(1)}x_2 & + a_{1n}^{(1)}x_n = b_1^{(1)} \\ 0 & + a_{22}^{(2)}x_2 & + a_{2n}^{(2)}x_n = b_2^{(2)} \\ \vdots & \vdots & \vdots \\ 0 & & a_{kk}^{(k)}x_k + \dots + a_{kn}^{(k)}x_n = b_k^{(k)} \\ 0 & & a_{nk}^{(k)}x_k + \dots + a_{nn}^{(k)}x_n = b_n^{(k)} \end{cases}$$

On suppose $a_{kk}^{(k)} \neq 0$ (appelé aussi pivot),
pour $i \geq k + 1$ (les k premières lignes ne changent pas)

$$\begin{cases} m_{ik} & = \frac{a_{ik}^{(k)}}{a_{kk}^{(k)}} & \text{pour } i = k + 1, \dots, n \\ a_{ij}^{(k+1)} & = a_{ij}^{(k)} - m_{ik}a_{kj}^{(k)} & i, j = k + 1, \dots, n \\ b_i^{(k+1)} & = b_i^{(k)} - m_{ik}b_k^{(k)} & i = k + 1, \dots, n \end{cases}$$

Il est clair que pour $k = n$ on obtient le système triangulaire supérieur $A^{(n)}X = b^{(n)}$ suivant:

$$A^{(n)} = \begin{pmatrix} a_{11}^{(1)}x_1 + a_{12}^{(1)}x_2 & & + a_{1n}^{(1)}x_n \\ 0 & + a_{22}^{(2)}x_2 & + a_{2n}^{(2)}x_n \\ \vdots & \vdots & \vdots \\ 0 & & a_{nn}^{(n)} \end{pmatrix} \quad b^{(n)} = \begin{pmatrix} b_1^{(1)} \\ b_2^{(2)} \\ \vdots \\ b_n^{(n)} \end{pmatrix}$$

on note U la matrice $A^{(n)}$.

Remarque 4.1.

b est transformé de la même façon, on assimile dans la pratique le vecteur $b^{(k)}$ à une $(n + 1)$ ième colonne de $A^{(k)}$ on pose alors $a_{i,n+1}^{(k)} = b_i^{(k)}$

On utilise une méthode de remontée pour résoudre le système précédent:

$$\begin{cases} x_n &= \frac{b_n}{a_{nn}^{(n)}} \\ x_{n-1} &= \frac{1}{a_{n-1,n-1}^{(n-1)}} (b_{n-1} - a_{n-1,n}^{(n-1)} x_n) \\ &\vdots \\ x_k &= \frac{1}{a_{kk}^{(k)}} (b_k^{(k)} - \sum_{j=k}^n a_{kj}^{(k)} x_j) \quad k = n-2, \dots, 1 \end{cases}$$

Nombre d'opérations:

Nombre de divisions total: $(n-1) + (n-2) + \dots + 1 = \frac{n(n-1)}{2}$

.....multiplications: $(n-1)^2 + (n-2)^2 + \dots + 1^2 + (n-1) + (n-2) + \dots + 1$
 $= \frac{n(n-1)(2n-1)}{6} + \frac{n(n-1)}{2} = \frac{n^3 - n}{3}$

Nombre d'additions: $\frac{n^3 - n}{3}$.

et pour résoudre $A^{(n)}X = b^{(n)}$ on a besoin de :

n divisions $+ 2 \sum_{k=1}^{n-1}$ multiplications et additions soit n^2 opérations. Au total de l'ordre de $\frac{2n^3}{3} + 2n^2$ opérations.

La méthode de Gauss n'est correctement utilisée que si les pivots $a_{kk}^{(k)} \neq 0$ pour $k = 1, \dots, n$, et le fait que les éléments diagonaux soient non nuls ne suffit pas pour empêcher l'apparition de pivots nuls.

Exemple 4.1. Soit la matrice inversible A et de termes diagonaux non nuls:

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 4 & 5 \\ 7 & 8 & 9 \end{pmatrix}, \quad A^{(2)} = \begin{pmatrix} 1 & 2 & 3 \\ 0 & 0 & -1 \\ 0 & -6 & -12 \end{pmatrix}$$

Des conditions plus restrictives sur A sont donc nécessaires pour assurer que la méthode de Gauss s'applique bien, nous verrons que si les mineurs principaux d_i de A sont non nuls pour $i = 1, \dots, n-1$ alors les pivots correspondants $a_{ii}^{(i)}$ sont également non nuls. (d_i est le déterminant de la i -ième sous matrice principale constituée des i premières lignes et colonnes. La matrice de l'exemple précédent ne satisfait pas cette condition $d_1 = 1$ mais $d_2 = 0$.)

Des exemples de matrices pour lesquelles on peut appliquer Gauss (sans contraintes):

1. Les matrices symétriques définies positives.
2. Les matrices à diagonale strictement dominante par ligne.
3. Les matrices à diagonale strictement dominante par colonne.

où

Définition 4.1. 1. Une matrice $A = (a_{ij})_{i,j}$ est dite à diagonale dominante par lignes (resp par colonnes) si et seulement si

$$|a_{ii}| \geq \sum_{j=1}^n |a_{ij}| \quad \forall i = 1, \dots, n$$

$$(resp.) |a_{ii}| \geq \sum_{j=1}^n |a_{ij}| \quad \forall j = 1, \dots, n$$

2. Une matrice $A = (a_{ij})_{i,j}$ est dite à diagonale fortement dominante par lignes si et seulement si

$$|a_{ii}| \geq \sum_{j=1}^n |a_{ij}| \quad \forall i = 1, \dots, n$$

et il existe $k \in \{1, \dots, n\}$ tel que $|a_{kk}| \geq \sum_{j=1}^n |a_{kj}|$

3. Une matrice $A = (a_{ij})_{i,j}$ est dite à diagonale strictement dominante par lignes si et seulement si

$$|a_{ii}| > \sum_{j=1}^n |a_{ij}| \quad \forall i = 1, \dots, n$$

4-3 Factorisation LU

Elle naît naturellement de la méthode de Gauss si aucune stratégie de pivot n'est utilisée. L'intérêt d'une telle factorisation c'est qu'elle dépend pas du second membre b . Ainsi si on a à résoudre plusieurs systèmes linéaires ayant la même matrice A et différents seconds membre, le nombre d'opérations est considérablement réduit puisque l'effort de calcul le plus important (soit environ $2n^3/3$) est dédiée à la procédure d'élimination.

Théorème 4.1.

Soit $A \in M_{n,n}(\mathbb{K})$, il existe une unique matrice L triangulaire inférieure à diagonale (= 1) et U une matrice triangulaire supérieure telle que $A = LU$ si et seulement si toute les sous matrices principales d'ordre 1 à $n - 1$ de A sont inversibles.

Preuve:

Supposons que les sous matrices principales A_i sont inversibles et montrons par récurrence l'existence et l'unicité de la factorisation LU avec $l_{ii} = 1$ (récurrence sur i).

Pour $i = 1$ la propriété est vraie. Supposons que la propriété est vraie jusqu'à $i - 1$ donc $A_{i-1} = L^{(i-1)}U^{(i-1)}$ avec $L_{kk}^{(i-1)} = 1$ pour $k = 1, \dots, i - 1$. Posons

$$A_i = \begin{pmatrix} A_{i-1} & c \\ d^t & a_{ii} \end{pmatrix}$$

cherchons une factorisation de A_i de la forme:

$$A_i = \begin{pmatrix} A_{i-1} & c \\ d^t & a_{ii} \end{pmatrix} = L^{(i)}U^{(i)} = \begin{pmatrix} L_{i-1} & 0 \\ l^t & 1 \end{pmatrix} \begin{pmatrix} U_{i-1} & u \\ 0^t & u_{ii} \end{pmatrix} \quad (*)$$

Donc on calcule et on trouve:

$$\begin{cases} L_{i-1}u & = c \\ l^t U_{i-1} & = d^t \\ a_{i,i} - l^t u & = u_{ii} \end{cases}$$

et comme A_{i-1} est inversible donc L_{i-1}, U_{i-1} sont inversibles ($\det A_{i-1} = \det L_{i-1} \det U_{i-1}$), donc les vecteurs l et u existents sont uniques, ainsi que u_{ii} d'où l'existence et l'unicité de l'unique factorisation $L^{(i)}U^{(i)}$ de $A^{(i)}$.

Montrons que si la factorisation $A = LU$ avec $L_{ii} = 1, i = 1, \dots, n$; existe et est unique alors les sous matrices principales sont inversibles.

Si A est inversible, donc $u_{ii} \neq 0$ ($\det A = \det L \det U = \det U = u_{11}u_{22} \dots u_{nn}$) et comme $A_i = L_i, U_i; i = 1, \dots, n$ donc $\det A_i = \det L_i \det U_i = \det U_i = u_{11}u_{22} \dots u_{ii} \neq 0$ d'où le résultat.

Si A est singulière et soit k le plus petit indice tel que $u_{kk} = 0$, d'après (*) la factorisation peut être faite jusqu'à l'ordre k , et à partir de cette étape la matrice $U^{(k)}$ étant singulière, on aura pas l'existence et l'unicité du vecteur l^t et donc pour avoir la factorisation complète de A il faut que toutes les sous matrices principales soient inversibles.

Une autre démonstration de l'existence et l'unicité de la factorisation LU est donnée comme suit:

Existence: Posons:

$$m_k = (0, \dots, 0, m_{k+1,k}, \dots, m_{n,k})^T \in \mathbb{R}$$

$$M_k = \begin{pmatrix} 1 & \dots & 0 & 0 & \dots & 0 \\ \vdots & \ddots & \vdots & \vdots & & \vdots \\ 0 & & 1 & 0 & & 0 \\ 0 & & -m_{k+1,k} & 1 & & 0 \\ \vdots & & \vdots & \vdots & \ddots & \vdots \\ 0 & \dots & -m_{n,k} & 0 & \dots & 1 \end{pmatrix} = I_n - m_k e_k^t$$

On a :

$$(M_k)_{il} = \delta_{il} - m_{ik} \delta_{kl} \quad i, l = 1, \dots, n$$

et d'après les transformations vues dans Gauss (passage de $A^{(k)}$ à $A^{(k+1)}$)

$$a_{ij}^{(k+1)} = a_{ij}^{(k)} - m_{ik} \delta_{kk} a_{kj}^{(k)} = \sum_{l=1}^n (\delta_{il} - m_{ik} \delta_{kl}) a_{lj}^{(k)} \quad i, j = k+1, \dots, n$$

c.à.d $A^{(k+1)} = M_k A^{(k)}, k = 1, \dots, n-1$. Donc à la fin du procédé de Gauss, on aura::

$$M_{n-1} M_{n-2} \dots M_1 A = A^{(n)} = U$$

Comme les M_k sont triangulaires inférieures à diagonale égale à l'unité sont donc inversibles et

$$M_k^{(-1)} = 2I_n - M_k = I_n + m_k e_k^t \quad (m_i e_i^t)(m_k e_k^t) = 0 \text{ pour } i \neq k$$

$$\begin{aligned}
A &= M_1^{(-1)} M_2^{(-1)} \dots M_n^{(-1)} U \\
&= (I_n + m_1 e_1^t) (I_n + m_2 e_2^t) \dots (I_n + m_{n-1} e_{n-1}^t) U \\
&= (I_n + \sum_{i=1}^{n-1} m_i e_i^t) U \\
&= \begin{bmatrix} 1 & 0 & \dots & \dots & 0 \\ m_{21} & 1 & & & \vdots \\ \vdots & m_{32} & \ddots & & \vdots \\ \vdots & \vdots & & \ddots & 0 \\ m_{n1} & m_{n2} & \dots & m_{n,n-1} & 1 \end{bmatrix} \quad \text{Donc} \\
A &= LU
\end{aligned}$$

Unicité: supposons qu'il existe L_1, U_1, L_2, U_2 tels que $A = L_1 U_1 = L_2 U_2$
 $L_2^{-1} L_1 = U_2 U_1^{-1} \implies L_2^{-1} L_1$ est diagonale comme elle contient que des 1 sur la diagonale
donc $L_2^{-1} L_1 = U_2 U_1^{-1} = I_n$ donc $U_2 = U_1$ et $L_2 = L_1$

4-4 Stratégie de pivot dans Gauss

Exemple 4.2. Si on utilise la méthode de Gauss à la matrice inversible suivante:

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 4 & 5 \\ 7 & 8 & 9 \end{pmatrix} \quad \text{on trouve} \quad A^{(2)} = \begin{pmatrix} 1 & 2 & 3 \\ 0 & 0 & -1 \\ 0 & -6 & -12 \end{pmatrix}$$

On voit apparaître un pivot nul à la seconde étape, en permutant les lignes 2 et 3 on se ramène (directement) à un système triangulaire supérieur:

$$A^{(2)} = \begin{pmatrix} 1 & 2 & 3 \\ 0 & -6 & -12 \\ 0 & 0 & -1 \end{pmatrix} = U$$

et les matrices de transformation sont données par:

$$M_1 = \begin{pmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ -7 & 0 & 1 \end{pmatrix}, \quad M_2 = I_3$$

Nous avons donc effectué une permutation des lignes de A et donc l'égalité $A = M_1^{(-1)} M_2^{(-1)} U$ est remplacée par $A = M_1^{(-1)} P M_2^{(-1)} U$ où P est la matrice de permutation

$$P = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}$$

En général:

- Si $a_{kk}^{(k)} = 0$ ou trop petit.
On adopte une stratégie de pivot, on en a 2:

- **Stratégie de pivot partiel:**

On cherche l'élément $a_{ik}^{(k)}$ $k \leq i \leq n$ tel que

$$|a_{ik}^{(k)}| = \max_{k \leq p \leq n} |a_{pk}^{(k)}|$$

On permute alors les lignes d'indices i et k pour ramener l'élément $a_{ik}^{(k)}$ en position de pivot.

- **Stratégie de pivot total:**

On cherche l'élément $a_{ij}^{(k)}$ $k \leq i, j \leq n$ tel que

$$|a_{ij}^{(k)}| = \max_{k \leq p, m \leq n} |a_{pm}^{(k)}|$$

On effectue alors une permutation de lignes et de colonnes pour ramener l'élément $a_{ij}^{(k)}$ en position de pivot.

la deuxième stratégie est peu utilisée en pratique sauf pour système assez délicats.

Ainsi si on utilise une stratégie de pivot de lignes ou colonnes nous avons en fait effectué une factorisation de type:

$$PAQ = LU$$

où P et Q sont le résultat des permutations faites sur les lignes et colonnes.

Remarque 4.2. • 1) *Il existe toujours un pivot partiel $\neq 0$, A est inversible car*

$$\det A^{(k)} = a_{11}^{(k)} a_{22}^{(k)} \dots a_{k-1, k-1}^{(k)} \det \begin{vmatrix} a_{kk}^{(k)} & \dots & a_{kn}^{(k)} \\ \vdots & & \vdots \\ a_{nk}^{(k)} & \dots & a_{nn}^{(k)} \end{vmatrix}$$

$$\det A^{(k)} \neq 0 \implies \det \begin{vmatrix} a_{kk}^{(k)} & \dots & a_{kn}^{(k)} \\ \vdots & & \vdots \\ a_{nk}^{(k)} & \dots & a_{nn}^{(k)} \end{vmatrix} \neq 0 \implies \exists k \leq i \leq n / a_{ik}^{(k)} \neq 0$$

- *le déterminant de A peut être calculé simplement $\det A = \pm a_{11}^{(1)} a_{22}^{(2)} \dots a_{nn}^{(n)}$, \pm dépend du nombre de permutations.*

- $\det A = U_{11}^{(1)} U_{22}^{(2)} \dots U_{nn}^{(n)}$.

Exemple 4.3. Exemple de choix de pivot:

$$\begin{cases} 10^{-4}x + y = 1 \\ x + y = 2 \end{cases} \quad \text{dont la solution est} \quad \begin{cases} x = 1.00010 \\ y = 0.99990 \end{cases}$$

Si on prend 10^{-4} comme pivot et 3 chiffres significatifs

$$\begin{cases} 10^{-4}x + y = 1 \\ (x + y) - \frac{1}{10^{-4}}(10^{-4}x + y) = 2 - \frac{1}{10^{-4}} \end{cases}$$

$$\begin{aligned} &\iff \begin{cases} 10^{-4}x + y = 1 \\ (10^4 - 1)y = 10^4 - 2 \end{cases} \\ \iff &\begin{cases} 10^{-4}x + y = 1 \\ 9999y = 9998 \end{cases} \iff \begin{cases} y \simeq 1 \\ x \simeq 0 \end{cases} \end{aligned}$$

En choisissant 1 comme pivot:

$$\begin{aligned} &\begin{cases} x + y = 2 \\ (10^{-4}x + y) - (x + y)10^{-4} = 1 - 2 \cdot 10^{-4} \end{cases} \\ \iff &\begin{cases} x + y = 2 \\ 0.9999y = 0.9999 \end{cases} \implies \begin{cases} y = 1 \\ y = 1 \end{cases} \end{aligned}$$

Exemple 4.4.

$$\begin{cases} \varepsilon x_1 + 2\varepsilon x_2 + 2\varepsilon x_3 = 2\varepsilon \\ 3x_1 + 7x_2 + 8x_3 = 8 \\ x_1 + 3x_2 + 3x_3 = 2 \end{cases}$$

où $\varepsilon > 0$ suffisamment petit. Appliquer Gauss pour calculer $X = (x_1, x_2, x_3)$.

4-5 Algorithme de Crout

C'est une variante de l'algorithme de Gauss avec C.N. qu'aucune stratégie de pivot n'est utilisée, on l'utilise lorsqu'on a à résoudre plusieurs fois le même système linéaire avec différents seconds membres.

- 1^{ère} étape:

$$A = LU = \begin{pmatrix} 1 & \dots & 0 & & 0 \\ L_{21} & 1 & & & \\ L_{31} & L_{32} & \ddots & & \\ \dots & & & \ddots & \\ L_{n1} & \dots & & L_{n,n-1} & 1 \end{pmatrix} \begin{pmatrix} U_{11} & U_{12} & \dots & & U_{1n} \\ & U_{22} & \dots & & U_{2n} \\ & & \ddots & & \\ 0 & & & \ddots & \\ & 0 & & & U_{n,n} \end{pmatrix} \implies L, U$$

Soit $a_{11} = U_{11}$, $a_{12} = U_{12}, \dots, U_{1j} = a_{1j}$
 $a_{21} = L_{21}U_{11}, a_{31} = L_{31}U_{11}, \dots, a_{j1} = L_{j1}U_{11} \implies$

$$U_{1j} = a_{1j}, \quad L_{j1} = \frac{a_{j1}}{a_{11}}$$

En général:

$$\text{Si } i \leq j \quad \begin{cases} a_{ij} = \sum_{k=1}^{i-1} L_{ik}U_{kj} + U_{ii} \\ a_{ji} = \sum_{k=1}^{i-1} L_{jk}U_{ki} + L_{ji}U_{ii} \end{cases}$$

$(L_{ii} = 1 \quad L_{ik} = 0, U_{kj} = 0 \text{ si } k > i, k > j)$

– Si

$$i = 1, \begin{cases} U_{1j} = a_{1j} \\ L_{j1} = \frac{a_{j1}}{a_{11}} \end{cases} \quad 1 \leq j \leq n$$

– Si $j > i$ $U_{ji} = L_{ij} = 0$.

–

$$i = 2, \dots, n \begin{cases} U_{ij} = a_{ij} - \sum_{k=1}^{i-1} L_{ik} U_{kj} \\ L_{ji} = a_{ji} - \frac{\sum_{k=1}^{i-1} L_{jk} U_{ki}}{U_{ii}} \end{cases}$$

• 2^{ème} étape:

Résolution de $Lz = b$

$$\begin{cases} l_{11}z_1 = b_1 \\ l_{21}z_1 + l_{22}z_2 = b_2 \\ l_{31}z_1 + l_{32}z_2 + l_{33}z_3 = b_3 \\ \dots \\ l_{n1}z_1 + \dots + l_{nn}z_n = b_n \end{cases} \implies$$

$$\begin{cases} z_1 = b_1 \\ z_2 = (b_2 - l_{21}z_1) \\ \dots \\ z_i = (b_i - \sum_{j=1}^{i-1} l_{ij}z_j) \\ \dots \\ z_n = (b_n - \sum_{j=1}^{n-1} l_{nj}z_j) \end{cases}$$

• 3^{ème} étape:

Résolution de $UX = z$. On utilise une méthode de remontée pour résoudre le système précédent:

$$\begin{cases} x_n = \frac{z_n}{u_{nn}} \\ x_{n-1} = \frac{1}{u_{n-1,n-1}}(z_{n-1} - u_{n-1,n}x_n) \\ \dots \\ x_k = \frac{1}{u_{kk}}(z_k - \sum_{j=k}^n u_{kj}x_j) \quad k = n-1, \dots, 1 \end{cases}$$

4-6 Méthode de Cholesky

Soit A une matrice symétrique définie positive. La méthode de Gauss s'applique sans stratégie de pivot, en effet les sous matrices principales sont inversibles;

$$\text{Soit } y \in \mathbb{R}^k / A_k y = 0 \text{ et soit } \bar{y} = (y, \underbrace{0, \dots, 0}_{k+1 \dots n})^t$$

$$(\bar{y})_i = \begin{cases} y_i & i \leq k \\ 0 & i > k \end{cases}$$

$$\begin{aligned} {}^t\bar{y}A\bar{y} &= (y_1, \dots, y_k, 0, \dots, 0) \begin{pmatrix} A_k \\ \vdots \\ 0 \end{pmatrix} \begin{pmatrix} y_1 \\ \vdots \\ y_k \\ \vdots \\ 0 \end{pmatrix} \\ &= {}^t y A_k y = 0 \implies \text{puisque } A \text{ est déf. positive } \bar{y} = 0 \implies y = 0 \end{aligned}$$

Donc $\det A_k \neq 0 \forall k \implies$ les pivots sont non nuls. Si on note $A^{(p)}$ la matrice à l'itération p (par Gauss).

$$\det A_p^{(p)} = \det A_p = a_{11}^{(p)} a_{22}^{(p)} \dots a_{pp}^{(p)} \implies a_{pp}^{(p)} \neq 0$$

et donc

Théorème 4.2. *Pour A une matrice symétrique définie positive, il existe une unique matrice triangulaire inférieure B telle que $A = B^t B$.
 B est parfois appelée la racine carrée de A .*

Preuve:

Il existe L et U (uniques) tels que $A = LU$

$$\det A_k = u_{11} \dots u_{kk} = \det U_k > 0 \quad \forall k \implies u_{kk} > 0$$

En posant $D = \text{diag}(\sqrt{u_{ii}})_{1 \leq i \leq n}$ on a $D = (LD)(D^{-1}U) = BC$

$$A = {}^t A \implies BC = {}^t C^t B \implies C({}^t B)^{-1} = (B^{-1})^t C$$

C et B sont resp triangulaires supérieure et inférieure $\implies (B^{-1})^t C$ est triangulaire inférieure et $C({}^t B)^{-1}$ est triangulaire supérieure donc $C({}^t B)^{-1} = (B^{-1})^t C$ est diagonale or les éléments diagonaux sont égaux à 1 $\implies C({}^t B)^{-1} = (B^{-1})^t C = I$ soit $C = {}^t B \implies A = B^t B$.

Supposons qu'il existe B_1 et B_2 tels que:

$$A = B_1^t B_1 = B_2^t B_2 \implies \begin{array}{l} B_1(D_1^{-1})D_1^t B_1 = B_2(D_2^{-1})D_2^t B_2 \\ L_1 U_1 = L_2 U_2 \end{array}$$

où $D_i = \text{diag}(B_i)$ et donc

$$\begin{cases} B_1(D_1^{-1}) &= B_2(D_2^{-1}) \\ D_1^t B_1 &= D_2^t B_2 \end{cases}$$

la diagonale des $D_i^t B_i$ sont égaux, $(B_1)_{ii}^2 = (B_2)_{ii}^2$ et puisque $(B_j)_{ii} > 0$ pour $j = 1, 2 \implies B_1 = B_2$

4-7 Algorithme de Cholesky

En posant

$$A = BB^t =$$

$$\begin{pmatrix} b_{11} & \dots & 0 & & 0 \\ b_{21} & b_{22} & & & \\ & & \ddots & & \\ \dots & & & \ddots & \\ b_{n1} & \dots & b_{n,n-1} & b_{nn} & \end{pmatrix} \begin{pmatrix} b_{11} & b_{21} & \dots & b_{n1} \\ & b_{22} & \dots & b_{n2} \\ & & \ddots & \\ 0 & & & \ddots \\ 0 & & & & b_{n,n} \end{pmatrix} \implies$$

$$a_{ij} = \sum_{k=1}^i b_{ik}b_{jk}, \quad 1 \leq i \leq j \leq n$$

Pour $i = 1$, on détermine la première colonne de B :

$$\begin{aligned} a_{11} &= b_{11}b_{11} & \text{d'où } b_{11} &= \sqrt{a_{11}} \\ a_{1j} &= b_{11}b_{j1} & \text{d'où } b_{j1} &= \frac{a_{1j}}{b_{11}}, \quad 2 \leq j \leq n \end{aligned}$$

On détermine la i -ème colonne de B ($2 \leq i \leq n$), après avoir calculé les $(i - 1)$ premières colonnes :

$$\begin{aligned} a_{ii} &= b_{i1}b_{i1} + \dots + b_{ii}b_{ii} & \text{d'où } b_{ii} &= \sqrt{a_{ii} - \sum_{k=1}^{i-1} b_{ik}^2} \\ a_{ij} &= b_{i1}b_{j1} + \dots + b_{ii}b_{ji} & \text{d'où } b_{ji} &= \frac{a_{ij} - \sum_{k=1}^{i-1} b_{ik}l_{jk}}{b_{ii}}, \quad i+1 \leq j \leq n \end{aligned}$$

Il résulte du théorème précédent qu'il est possible de choisir tous les éléments $b_{ii} > 0$ en assurant que toutes les quantités

$$a_{11}, \dots, a_{ii} - \sum_{k=1}^{i-1} b_{ik}^2, \dots$$

sont positives

Nombre d'opérations

On a une évaluation de $\frac{n^3}{6}$ additions + $\frac{n^3}{6}$ multiplications + $\frac{n^2}{2}$ divisions + n extractions de racines. Donc pour une matrice symétrique définie positive vaut mieux utiliser Cholesky que Gauss.

5 Méthodes itératives

En vue de résoudre $AX = b$, on cherche X comme limite d'une suite vectorielle $(X^{(k)})_k$, en posant $A = M - N$ on construit $(X^{(k)})_k$ ainsi:

$$\begin{cases} MX^{(k+1)} &= NX^{(k)} + b \\ X^{(0)} &\text{donné} \end{cases}$$

$X^{(k+1)} = F(X^{(k)})$ c.à.d. qu'on cherche X comme point fixe de $F(X) = M^{-1}NX + M^{-1}b$. Dans le cas de problèmes linéaires plusieurs questions se posent:

- Pour quelles décomposition de A , a-t-on la convergence de la méthode? (Pour $M = A$ on a la convergence en une itération mais il faut calculer A^{-1})
- Peut-on accélérer la convergence?
- Comment définir le critère d'arrêt? ($\|X^{(k)} - X\|$, X étant inconnu!)

5-1 Convergence des méthodes itératives

Soit une méthode itérative de type:

$$X^{(k+1)} = BX^{(k)} + C \quad (1)$$

pour résoudre $AX = b$, pour A inversible.

Définitions 5.1. *Consistance, convergence.*

- Une méthode est dite consistante si lorsque la suite $(X^{(k)})_k$ converge, elle converge vers la solution. (c.à.d. $A^{-1}b$)
- Une méthode est dite convergente si la suite $(X^{(k)})_k$ converge vers la solution $A^{-1}b$ pour tout $X^{(0)}$. ($\lim_{k \rightarrow \infty} \|X^{(k)} - X\| = 0$ pour toute norme vectorielle sur \mathbb{K}^n .)

Théorème 5.1. *Une méthode itérative (1) consistante est convergente si et seulement si $\rho(B) < 1$ si et seulement si $\|B\| < 1$ pour au moins une norme subordonnée si et seulement si $\lim \|B\|^k = 0$ pour toute norme.*

Preuve

5-2 Vitesse de convergence, test d'arrêt et nombre d'itérations

Définition 5.2. *On définit la vitesse moyenne de convergence par:*

$$R(B^m) = -\log(\|B^m\|_2^{\frac{1}{m}}) = -\frac{\log(\|B^m\|_2)}{m}$$

Ainsi pour 2 méthodes itératives associées à B_1 et B_2 convergentes, la première est dite plus rapide que la seconde en m si et seulement si $R(B_1^m) > R(B_2^m)$. $X^{(k)}$ tend vers X aussi rapidement que B^k tend vers 0 ou $\rho(B)$ est si petit. on a aussi le résultat suivant:

Théorème 5.2. *Soit B la matrice (dite matrice d'itération) associée à une méthode itérative convergente. On a alors*

$$\lim_{m \rightarrow \infty} R(B^m) = R_\infty(B) \quad \text{où} \quad R_\infty(B) = -\log(\rho(B))$$

Définition 5.3. On définit le test d'arrêt à partir du vecteur résidu comme suit:
 $r^{(k)} = b - AX^{(k)}$, pour une précision donnée ε on calcule

$$\frac{\|r^{(k)}\|}{\|b\|} = \frac{\|b - AX^{(k)}\|}{\|b\|} \leq \varepsilon$$

(On peut aussi prendre le test d'arrêt: $\frac{\|X^{(k+1)} - X^{(k)}\|}{\|b\|} \leq \varepsilon$)

Pour calculer le nombre d'itérations maximales en fonction de l'erreur souhaitée on procède comme suit:

Soit une norme donnée $\|\cdot\|$ telle que $\|B\| < 1$ on a $\|e^{(k)}\| \leq \|B^{(k)}\| \|e^{(0)}\|$ où $e^{(k)} = X^{(k)} - X$
 $\|e^{(0)}\| \leq \|X^{(0)} - X^{(1)}\| + \|e^{(1)}\| \leq \|X^{(0)} - X^{(1)}\| + \|B\| \|e^{(0)}\|$

$$\implies \|e^{(k)}\| \leq \frac{\|B\|^k}{1 - \|B\|} \|X^{(1)} - X^{(0)}\|$$

et donc pour que $\|e^{(k)}\| \leq 10^{-N}$ il suffit qu'il existe $K \in \mathbb{N}^*$ tel que

$$\frac{\|B\|^k}{1 - \|B\|} \|X^{(1)} - X^{(0)}\| \leq 10^{-N}$$

c.à.d que

$$K \geq \frac{\log\left(\frac{1 - \|B\|}{\|X^{(1)} - X^{(0)}\|}\right) - N \log 10}{\log(\|B\|)}$$

6 Méthodes de Jacobi, de Gauss Seidel et de Relaxation

On choisit la décomposition

$$A = D - E - F \text{ où}$$

- D : la diagonale de A
- E : la partie inférieure de $-A$
- F : la partie supérieure de $-A$

Pour Jacobi

On prend $M = D$ qu'on suppose inversible et $N = E + F$ ($X^{(k)}$)_k est alors donnée par:

$$\begin{cases} X^{(k+1)} & = D^{-1}(E + F)X^{(k)} + D^{-1}b \\ X^{(0)} & \text{donné} \end{cases}$$

et on a l'algorithme:

$X^{(0)}$ donné.

(1) Pour k

pour $i = 1, \dots, n$

$$X_i^{(k+1)} = \frac{\left(-\sum_{p \neq i}^n a_{ip} X_p^{(k)} + b_i\right)}{a_{ii}}$$

Si test d'arrêt est bon \rightarrow Fin.
 Sinon $k + 1 \rightarrow k$ et aller à (1).

Pour Gauss-Seidel

On a $X^{(0)} = (X_1^{(0)}, X_2^{(0)}, \dots, X_n^{(0)})$ $\xrightarrow{\text{Jacobi}}$ on calcule

$$\begin{aligned} X_1^{(1)} &= f(X_1^{(0)}, X_2^{(0)}, \dots, X_n^{(0)}) \\ X_2^{(1)} &= f(X_1^{(0)}, X_2^{(0)}, \dots, X_n^{(0)}) \\ &\vdots \\ X_n^{(1)} &= f(X_1^{(0)}, X_2^{(0)}, \dots, X_n^{(0)}) \end{aligned}$$

Pour Gauss Seidel:

$$\begin{aligned} X_1^{(1)} &= f(X_1^{(0)}, X_2^{(0)}, \dots, X_n^{(0)}) \\ X_2^{(1)} &= f(X_1^{(1)}, X_2^{(0)}, \dots, X_n^{(0)}) \\ X_3^{(1)} &= f(X_1^{(1)}, X_2^{(1)}, X_3^{(0)}, \dots, X_n^{(0)}) \\ &\vdots \\ X_i^{(1)} &= f(X_1^{(1)}, X_2^{(1)}, \dots, X_{i-1}^{(1)}, X_i^{(0)}, \dots, X_n^{(0)}) \\ &\vdots \end{aligned}$$

Soit l'algorithme:

$X^{(0)} \in \mathbb{K}^n$ donné, ε donné,

(1) A l'itération k

Pour $i = 1, \dots, n$

$$X_i^{(k+1)} = \frac{(-\sum_{p < i} a_{ip} X_p^{(k+1)} - \sum_{p > i} a_{ip} X_p^{(k)} + b_i)}{a_{ii}}$$

Si test d'arrêt est bon \rightarrow Fin.

Sinon $k + 1 \rightarrow k$ et aller à (1).

En écriture matricielle l'algorithme s'écrit:

$$DX^{(k+1)} = EX^{(k+1)} + FX^{(k)} + b$$

$$X^{(k+1)} = (D - E)^{-1}(FX^{(k)} + b) = (D - E)^{-1}FX^{(k)} + C$$

$L_1 = (D - E)^{-1}F$ est la matrice d'itération de Gauss Seidel.

Méthode de relaxation:

L'idée de la méthode de relaxation successives (SSOR) est d'augmenter la vitesse de convergence qui est relativement faible pour les 2 méthodes précédentes, on introduit un paramètre w pour cette raison, et on construit $X^{(k+1)}$ comme combinaison linéaire de $X^{(k)}$ et $\bar{X}^{(k+1)}$ calculée par Gauss Seidel

$$X^{(k+1)} = w\bar{X}^{(k+1)} + (1 - w)X^{(k)} \quad \text{où} \quad \bar{X}^{(k+1)} = D^{-1}(EX^{(k+1)} + FX^{(k)} + b)$$

En éliminant $\bar{X}^{(k+1)}$ dans le système précédent:

$$\bar{X}^{(k+1)} = \frac{1}{w}X^{(k+1)} - \frac{1-w}{w}X^{(k)}$$

$$D\left(\frac{1}{w}X^{(k+1)} - \frac{1-w}{w}X^{(k)}\right) - EX^{(k+1)} = FX^{(k)} + b$$

c.à.d.

$$\left(\frac{D}{w} - E\right)X^{(k+1)} = \left(F + \frac{1-w}{w}D\right)X^{(k)} + b$$

Donc pour $M = \frac{D}{w} - E$ et $N = \frac{1-w}{w}D + F$, on vérifie que la méthode est consistante, en effet si

$$X^{(k)} \longrightarrow X \text{ alors } \left(\frac{D}{w} - E\right)X = \left(F + \frac{1-w}{w}D\right)X + b$$

soit $AX = b$

Algorithme de relaxation:

$X^{(0)} \in \mathbb{K}^n$ donné, ε donné,

(1) A l'itération k

Pour $i = 1, \dots, n$

$$X_i^{(k+1)} = (1-w)X_i^{(k)} - \frac{w}{a_{ii}}\left(-\sum_{p<i} a_{ip}X_p^{(k+1)} - \sum_{p>i} a_{ip}X_p^{(k)} + b_i\right)$$

Si test d'arrêt ($\|AX^{(k+1)} - b\| \leq \varepsilon$) est bon \rightarrow Fin.

Sinon $k+1 \rightarrow k$ et aller à (1).

Nombre d'opérations pour les 3 méthodes:

Pour chaque itération:

$n((n-1)$ multiplications $+(n-1)$ additions $+1$ division) $\simeq 2n^2$ opérations pour une matrice pleine.

Pour comparer à Gauss il faut aboutir à une solution approchée avec une précision demandée en moins de $\frac{n}{3}$ itérations. Dans la pratique les méthodes itératives sont plus utilisées pour des matrices creuses, où à chaque opération le nombre d'opérations est proportionnelle au nombre d'éléments $\neq 0$, en plus il suffit de stocker les éléments non nuls de A c'est ce point le plus essentiel quant au choix des méthodes itératives.

Remarque 6.1. *La méthode de Gauss ou Cholesky nécessite plus d'espace mémoire pour le stockage des éléments même pour des matrices bandes puisqu'il faut remplir tous les éléments dans la bande.*

6-1 Résultats de convergence pour les 3 méthodes

Pour Jacobi et Gauss Seidel

Théorème 6.1. *Soit A une matrice à diagonale strictement dominante, alors les méthodes de Jacobi et Gauss Seidel convergent.*

Avant d'établir des résultats de convergence de la méthode de relaxation, nous donnons les résultats suivants:

Définition 6.1. Soient A, M et N trois matrices carrées à coefficients réels ou complexes. On dit que $A = M - N$ est un éclatement P -régulier de A si et seulement si M est inversible et ${}^tM + N$ est définie positive.

Lemme 6.1. Soit A symétrique définie positive telle que $A - H^tAH$ définie positive alors $\rho(H) < 1$

Preuve:

Soit λ une valeur propre de H / $\rho(H) = |\lambda|$ et soit u le vecteur propre associée,

$${}^tu(A - {}^tHAH)u > 0 \iff {}^t uAu - (\lambda^2){}^t uAu > 0 \iff (1 - \lambda^2) > 0$$

Proposition 6.1. Soit $A = M - N$ un éclatement P -régulier et A symétrique définie positive alors $\rho(M^{-1}N) < 1$

Preuve:

Il suffit de montrer que $A - {}^t(M^{-1}N)A(M^{-1}N)$ est définie positive.

$$\begin{aligned} M^{-1}N &= M^{-1}(M - A) = I - M^{-1}A \\ {}^t(M^{-1}N)A(M^{-1}N) &= {}^t(I - M^{-1}A)A(I - M^{-1}A) \\ &= A - {}^t(M^{-1}A)A - AM^{-1}A + {}^t(M^{-1}A)A(M^{-1}A) \end{aligned}$$

Donc

$$\begin{aligned} A - {}^t(M^{-1}N)A(M^{-1}N) &= {}^t(M^{-1}A)A + AM^{-1}A - {}^t(M^{-1}A)A(M^{-1}A) \\ &= {}^t(M^{-1}A)A + AM^{-1}A - {}^t(M^{-1}A)A(M^{-1}A) \\ &= {}^t(M^{-1}A)[A(M^{-1}A)^{-1} + ({}^t(M^{-1}A))^{-1}A - A]M^{-1}A \\ &= {}^t(M^{-1}A)[M + {}^tM - A]M^{-1}A \\ &= {}^t(M^{-1}A)[N + {}^tM]M^{-1}A \end{aligned}$$

Or $N + {}^tM$ est définie positive donc

${}^t(M^{-1}A)[N + {}^tM]M^{-1}A$ est définie positive et donc $\rho(M^{-1}N) < 1$.

On montre alors le théorème suivant:

Théorème 6.2. (Ostrowski-Reich)

Soit A une matrice symétrique définie positive, alors la méthode de relaxation converge pour tout $w \in]0, 2[$ (en particulier Gauss Seidel converge)

Preuve:

Il suffit de montrer que

$$A = \frac{1}{w}(D - wE) - \frac{1}{w}((1 - w)D + wF)$$

est un éclatement P -régulier ($w \neq 0$).

$D - wE$ est inversible en effet $\det(D - wE) = \prod_{i=1}^n a_{ii} > 0$ car $\forall i \quad a_{ii} = {}^t e_i A e_i > 0$

${}^tM + N = \frac{2 - w}{w}D$ définie positive pour $w \in]0, 2[$

Si $w \in]0, 1[$, on dit qu'on a une sous relaxation.

Si $w \in]1, 2[$, on dit qu'on a une sur-relaxation.

Si $w = 1$ on a Gauss Seidel.

Condition nécessaire de convergence,

$$\rho(L_w) \geq |w - 1| \quad |w \neq 0$$

$$\begin{aligned} \det L_w &= \prod_{i=1}^n \lambda_i(L_w) \\ \det L_w &= \frac{\det(\frac{1-w}{w}D + F)}{\det(\frac{D}{w} - E)} = (1-w)^n \\ \rho(L_w)^n &\geq \prod_{i=1}^n |\lambda_i(L_w)| = |1-w|^n \\ \rho(L_w) &\geq |1-w| \end{aligned}$$

On montre aussi qu'il existe un paramètre de relaxation optimal w^* pour lequel la méthode de relaxation est très efficace, notamment pour des systèmes provenant de la discrétisation d'E.D.P. (pour de tels systèmes on a aussi Gauss Seidel qui converge plus rapidement que Jacobi). En posant

$$\begin{aligned} A &= D(I - L - U) \\ L &= D^{-1}E \quad , \quad U = D^{-1}F \quad , \quad J = L + U \\ L_w &= (I - wL)^{-1}[(1-w)I + wU] \end{aligned}$$

On montre le résultat suivant:

Théorème 6.3. *Soit A une matrice telle que $\forall \alpha \neq 0$, les valeurs propres de $J(\alpha) = \alpha L + \frac{1}{\alpha}U$ sont indépendantes de α alors $\rho(L_1) = \rho(J)^2$. Donc les méthodes de G.Seidel et de Jacobi convergent ou divergent toutes les deux et si on a convergence alors G.S. converge 2 fois plus vite.*

De plus si $sp(J)$ ne contient que des réels et $\rho(J) < 1$ alors $\rho(L_w) < 1$ pour $w \in]0, 2[$ et $\rho(L_w)$ est une fonction de w ayant l'allure suivante:

Rayon spectrale de relaxation en fonction de w .

$$w^* = \frac{2}{1 + \sqrt{1 - \rho(J)^2}} = w_{opt}$$

Exemple 6.1.

$$(1) \quad \begin{cases} -u''(x) = f(x) & 0 < x < 1 \\ u(0) = u(1) = 0 \end{cases}$$

On montre que si f est continue sur $[0, 1]$, (1) admet une solution unique u . Le calcul numérique de u par la méthode des différences finis consiste à remplacer la recherche de u par la recherche d'un vecteur $u_h = (u_1, u_2, \dots, u_n)$ aux points x_1, x_2, \dots, x_n d'une subdivision de $[0, 1]$. Si on suppose qu'on a une subdivision régulière $x_i = ih$ $i = 0, \dots, n+1$ et $h = \frac{1}{n+1}$, en $x_0 = 0$ et en $x_{n+1} = 1$ on prend $u = 0$ c.à.d. $u_0 = u_{n+1} = 0$ (conditions aux limites)

On remplace

$$u''(x_i) \simeq \frac{u(x_{i+1}) + u(x_{i-1}) - 2u(x_i)}{h^2}$$

Donc (1) conduit à

$$(1) \quad \begin{cases} -\frac{u(x_{i+1}) + u(x_{i-1}) - 2u(x_i)}{h^2} = f_i = f(x_i) & i = 0, \dots, n \\ u_0 = u_{n+1} = 0 \end{cases}$$

$$\iff \frac{1}{h^2}AU_h = f_h \quad \text{où} \quad \begin{aligned} f_h &= (f_1, f_2, \dots, f_n)^t \\ U_h &= (u_1, u_2, \dots, u_n)^t \end{aligned}$$

et

$$A = \begin{pmatrix} 2 & -1 & \dots & & 0 \\ -1 & 2 & -1 & \dots & 0 \\ \vdots & \ddots & \ddots & \ddots & \ddots \\ 0 & & & \ddots & -1 \\ 0 & & & -1 & 2 \end{pmatrix}$$

On montre que A est symétrique (définie positive, à diagonale fortement dominante et irréductible et les éléments de la diagonale sont > 0)

donc Jacobi, Gauss Seidel et relaxation pour $w \in]0, 2[$ convergent et on a

$$\rho(L_{w_{\text{opt}}}) = \left(\inf_{0 < w < 2} \rho(L_w) \right) = w_{\text{opt}} - 1 < \rho(L_1) = \rho(J)^2 < \rho(J)$$

si $\rho(J) \neq 0$, et si $\rho(J) = 0$ alors $w_{\text{opt}} = 1 \implies \rho(L_1) = \rho(J)^2 = 0$