6 METHODES INDIRECTES

6.0.1 Introduction

Les méthodes directes de résolution de systèmes linéaires fournissent une solution x au problème $Ax = b$ en un nombre fini d'op érations. Si l'ordre n de la matrice A est élevé, le nombre d'opérations est aussi élevé et de plus, le résultat obtenu n'est pas rigoureusement exact. Par ailleurs, il existe des cas où les structures du système liné aire ne sont pas tirés à profit par les méthodes directes. C'est par exemple le cas des systèmes où la matrice A est très creuse. C'est la raison pour laquelle, dans ce cas , on préfère utiliser des méthodes itératives. L'objectif est de construire une suite de vecteurs $\left\{x^{(k)}\right\}_{k=1,2...,n}$ qui tend vers un vecteur $\bar{x}$, solution exacte du problème $Ax = b$. Souvent, on part d'une approximation $\left\{x^{(0)}\right\}$ de $\bar{x}$ obtenue en géné ral par une méthode directe.

6.1 Les méthodes itératives

L'objectif est de résoudre un système du type $Ax = b$. Pour cela, nous allons décomposer la matrice A en

$$A = M - N$$

de sorte que M soit inversible. Ainsi, le système devient :

$$Mx = Nx + b$$

et nous chercherons par récurrence une suite de vecteurs $x^{(i)}$ obtenu à partir d'un vecteur $x^{(0)}$ et de la relation

$$Mx^{(k+1)} = Nx^{(k)} + b$$

c'est-à-dire

$$x^{(k+1)} = M^{-1}Nx^{(k)} + M^{-1}b$$

Cette relation est une relation de récurrence du premier ordre. Nous pouvons en déduire une relation reliant l'erreur $e^{(k)} = x^{(k)} - \bar{x}$ à $e^{(k-1)} = x^{(k-1)} - \bar{x}$:

$$M(x^{(k)} - \bar{x}) = N(x^{(k-1)} - \bar{x})$$

puisque $M\bar{x} = N\bar{x} + b$ et donc $e^{(k)} = M^{-1}Ne^{(k-1)}$ pour $k = 1, 2, ...$ Si on pose $B = M^{-1}N$, nous avons alors

$$e^{(k)} = Be^{(0)}$$

La convergence de la suite $x^{(k)}$ vers la solution $\bar{x}$ est donné par le proposition suivant :

Proposition 202. Le choix de la décomposition de A devra obeir aux règles suivantes :

Remarque 203.

Proposition 204.

1. Le rayon spectral $\rho(M^{-1}N)$ doit être strictement infé rieur à 1.
2. La résiolution de $Mx^{(k)} = Nx^{(k-1)} + b$ doit être simple et n écessiter le moins d'opérations possibles
3. Pour obtenir la meilleure convergence, $\rho(M^{-1}N)$ doit être le plus petit possible.

On voit que la convergence dépend de la décomposition.

6.2 Différentes décomposition de A

On écrit la matrice A sous la forme

$$A = D + E + F$$

avec D la matrice diagonale suivante :

$$D = \begin{pmatrix} a_{11} & 0 & \cdots & 0 \\ 0 & a_{22} & 0 & \vdots \\ \vdots & & \ddots & \vdots \\ 0 & \cdots & 0 & a_{nn} \end{pmatrix}$$

E la matrice triangulaire inférieure suivante

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

et F la matrice triangulaire supérieure

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

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

6.3 Méthode de Jacobi

On pose

$$M = D \quad \text{et} \quad N = -(E + F)$$

ainsi, $B = M^{-1}N = D^{-1}(-E - F)$, ce qui implique :

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

si on exprime cette relation en fonction des éléments de la matrice A nous avons :

$$x_i^{(k+1)} = -\sum_{\substack{j=1 \\ j \neq i}}^{n} \frac{a_{ij}}{a_{ii}} x_j^{(k)} + \frac{b_i}{a_{ii}}, \qquad i = 1, 2, \ldots, n$$

6.4 Méthode de Gauss-Seidel

Cette méthode utilise

$$M = D + E \quad \text{et} \quad N = -F$$

D'où

$$B = -(D + E)^{-1} F,$$

et alors on a :

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

le calcul de l'inverse de $(D+E)$ peut être évité. Si on écrit $(D+E)x^{(k+1)} = -Fx^k + b$, on obtient

$$\sum_{j=1}^{n} a_{ij}x_j^{(k+1)} = -\sum_{j=i+1}^{n} a_{ij}x_j^{(k)}b_i,$$

d'où

$$x_i^{(k+1)} = -\frac{1}{a_{ii}}\sum_{j=1}^{i-1} a_{ij}x_j^{(k+1)} - \frac{1}{a_{ii}}\sum_{j=i+1}^{i-1} a_{ij}x_j^{(k)} + \frac{b_i}{a_{ii}}, \qquad i = 1, 2, ..., n.$$

6.5 Méthode de relaxation

On donne un paramètre $\omega \in]0,2[$, appelé facteur de relaxation, et on pose

$$M = \frac{D}{\omega} + E \quad \text{et} \quad N = \left(\frac{1-\omega}{\omega}\right)D - F$$

et par conséquent

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

d'où

$$x_i^{(k+1)} = (1-\omega)x_i^{(k)} + \frac{\omega}{a_{ii}}\left(-\sum_{j=1}^{i-1} a_{ij}x_j^{(k+1)} - \sum_{j=i+1}^{n} a_{ij}x_j^{(k)} + b_i\right) \qquad i = 1, 2, ..., n.$$

Comme on peut le constater, la méthode de Gauss-Seidel correspond à la méthode de relaxation pour $\omega = 1$.

7 Convergence des méthodes itératives

La convergence des méthodes itératives dépend fortement du rayon spectral de A, Nous étudions d'abord les propriétés de certaines matrices et la localisation de leurs valeurs propres.

Définition 205. Soit $A \in \mathcal{M}_{m,n}(\mathbb{R})$ une matrice. On définit la norme matricielle induite à partir de la norme vectorielle sur $\mathbb{R}^n$ par

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

Proposition 206. Soit A et B deux matrices telles que leur multiplication soit compatible alors on a :

$$\|AB\| \leq \|A\|\,\|B\|$$

pour toute norme induite.

Théorème 207 (Gerschgorin-Hadamard). Les valeurs propres de la matrice A appartiennent à la reunion des n disques D_k pour $k = 1, 2, ..., n$ du plan complexe ($\lambda \in \cup_{k=1}^{n} D_k$ où D_k, appelé disque de Gerschgorin, est dé fini par :

$$|z - a_{kk}| \leq \sum_{\substack{j=1 \\ j\neq i}} \left|a_{kj}\right|$$

7.1 Cas général

On considère une méthode itérative définie comme :

$$\begin{cases} x^{(0)} & \text{donné} \\ x^{(k+1)} & = \quad Cx^{(k)} + D \end{cases}$$

Théorème 208. Soit A une matrice carré d'ordre n, pour que $\lim_{k\to\infty} A^k = 0$, il faut et il suffit que $\rho(A) < 1$.

Théorème 209. Si il existe une norme induite telle que $\|C\| < 1$ alors la méthode itérative décrite ci-dessus est convergente quelque soit $x^{(0)}$ et elle converge vers la solution de :

$$(I_d - C)x = D$$

Théorème 210. Une condition nécessaire et suffisante de convergence de la méthode ci-dessus est que :

$$\rho(C) < 1$$

Remarque 211. la condition de convergence donnée par le rayon spectral n'est pas dé pendante de la norme induite, cependant elle peut être utile car le calcul du rayon spectral peut être difficile.

7.1.1 Cas des matrices à diagonale dominante

Définition 212. Une matrice est dite à diagonale dominante si :

$$\forall i, 1 \leq i \leq n, \quad |a_{ii}| > \sum_{\substack{j=1 \\ j\neq i}}^{n} \left|a_{ij}\right|$$

Théorème 213. Si A est une matrice à diagonale strictement dominante, alors A est inversible et en outre, les méthodes de Jacobi et de Gauss-Seidel convergent.

Démonstration. si A est une matrice à diagonale strictement dominante, on montre que A est inversible en démontrant que 0 n'est pas une valeur propre (c'est-à-dire $KerA = 0$). Posons $B = M^{-1}N$ est soit λ et v tels que $Bv = \lambda v$ avec $v \neq 0$. Puisque l'on s'interesse à $\rho(B) < 1$, on s'interesse en fait à la plus grande valeur propre de plus grand module de B. Ainsi, on peut supposer que $\lambda \neq 0$. L'équation $Bv = \lambda v$ devient :

$$\left(M - \frac{1}{\lambda}N\right)v = 0$$

- Pour Jacobi ; l'équation devient :

$$\left(D + \frac{1}{\lambda}E + \frac{1}{\lambda}F\right)v = 0$$

soit $C = D + \frac{1}{\lambda}E + \frac{1}{\lambda}F$. si $|\lambda| \geq 1$, on aurait :

$$|c_{ii}| = |a_{ii}| > \sum_{\substack{j=1 \\ j\neq i}}^{n} \left|a_{ij}\right| \geq \sum_{\substack{j=1 \\ j\neq i}}^{n} \left|\frac{a_{ij}}{\lambda}\right| = \sum_{\substack{j=1 \\ j\neq i}}^{n} \left|c_{ij}\right|$$

donc C serait à diagonale strictement dominante et par conséquent inversible. C inversible implique que $Cv = 0$ donc $v = 0$. Or $v \neq 0$, d'o ù la contradiction et donc on a bien $|\lambda| < 1$. - Pour Gauss-Seidel ; l'équation devient :

$$\left(D + E + \frac{1}{\lambda}F\right)v = 0$$

en posant encore $C = D + E + \frac{1}{\lambda}F$. et en supposant $|\lambda| \geq 1$, on aurait :

$$|c_{ii}| = |a_{ii}| > \sum_{\substack{j=1 \\ j\neq i}}^{n} \left|a_{ij}\right| \geq \sum_{j<1}^{n} \left|\frac{a_{ij}}{\lambda}\right| + \sum_{j>1}^{n} \left|\frac{a_{ij}}{\lambda}\right| = \sum_{\substack{j=1 \\ j\neq i}}^{n} \left|c_{ij}\right|$$

et on obtient le même type de contradiction. cqfd

7.1.2 Cas des matrices symétriques définies positives

Théorème 214. Si A est une matrice symétrique définie positive, alors les mé thodes de Gauss-Seidel et de relaxation pour ($\omega \in]0,2[$) convergent.

La convergence de la méthode est d'autant plus rapide que $\rho(M^{-1}N)$ est petit. Or cette matrice $B = M^{-1}N$ dépend de ω. Une é tude théorique des valeurs propres de B montre que l'allure de la courbe $\rho(B)$ en fonction de B est décroissante entre 0 et ω_{opt} et croissante entre ω_{opt} et 2. Par ailleurs, on a toujours $1 < \omega_{opt} < 2$. On a donc intérêt à choisir ω le plus proche possible de ω_{opt}.

7.1.3 La méthode de correction

Soit le vecteur reste en x défini comme :

$$r(x) = b - Ax$$

et $\left\{r^{(k)}\right\}$ le reste en $\left\{x^{(k)}\right\}$. On appelle également l'erreur en k le vecteur

$$e^{(k)} = x^{(k)} - \bar{x}$$

où $\bar{x}$ est la solution. si on a une approximation $\left\{x^{(0)}\right\}$ de x, la relation suivante est vérifiée :

$$Ae^{(0)} = A(x^{(0)} - \bar{x}) = A(x^{(0)}) - b = -r^{(0)}$$

ce qui signifie que $e^{(0)}$ est la solution du système $Ax = -r^{(0)}$ et théoriquement, on a $\bar{x} = x^{(0)} - e^{(0)}$. Pratiquement, en appliquant au système $Ax = -r^{(0)}$ la méthode directe qui nous a fourni $x^{(0)}$, on n'obtient pas directement $e^{(0)}$, mais une approximation $y^{(0)}$ de $e^{(0)}$. Si on pose $x^{(1)} = x^{(0)} - y^{(0)}$, $x^{(1)}$ est une nouvelle approximation de $\bar{x}$, en itérant les calculs précédents, on obtient :

$$Ae^{(1)} = A(x^{(1)} - \bar{x}) = A(x^{(1)}) - b = -r^{(1)}$$

la résolution du système $Ax = -r^{(1)}$ donnera une approximation $y^{(1)}$ de $e^{(1)}$, et une nouvelle approximation $x^{(2)}$ de $\bar{x}$:

$$x^{(2)} = x^{(1)} - y^{(1)} = x^{(0)} - y^{(0)} - y^{(1)}$$

Ces calculs peuvent être itérer autant de fois que nécessaire, pour s'arrêter lorsque le reste est suffisament petit. A la $k^{ième}$ itération, les relations suivantes sont vérifiées pour $y^{(k-1)}$ approximation de $e^{(k-1)}$:

$$x^{(k)} = x^{(k-1)} - y^{(k-1)} = x^{(0)} - \sum_{i=0}^{k-1} y^{(i)}$$

avec $y^{(i)}$ une approximation de $e^{(i)}$, solution de $Ax = -r^{(i)}$ et $i = 0,1,2,...,k-1$. Si nous nous arrêtons lorsque $k = N$, il est nécessaire de ré soudre $N+1$ systèmes linéaires : d'abord $Ax = b$, pour obtenir $x^{(0)}$ puis $Ax = -r^{(i)}$ et $i = 0,1,2,...,N-1$ afin d'obtenir $y^{(i)}$. Une fois la matrice A décomposée (en LU ou Cholesky), il s'agit donc de résoudre les systèmes $LUx = -r^{(i)}$ où $-r^{(i)}$ a été calculé par la relation $r^{(i)} = b - Ax^{(i)}$.

8 SERIE D'EXERCICES

Exercice 215. Résoudre le système d'équations linéaires suivant :

$$\begin{cases} 10x_1 - 2x_2 - 2x_3 & = & 6 \\ -x_1 + 10x_2 - 2x_3 & = & 7 \\ -x_1 - x_2 + 10x_3 & = & 8 \end{cases}$$

Par la méthode des approximations successives. Arrêter les calculs dès que :

$$\left|x_i^{(k+1)} - x_i^{(k)}\right| < 10^{-2}$$

Exercice 216. Résoudre le système d'équations linéaires suivant :

$$\begin{cases} 10x_1 - 2x_2 - 2x_3 & = & 6 \\ -x_1 + 10x_2 - 2x_3 & = & 7 \\ -x_1 - x_2 + 10x_3 & = & 8 \end{cases}$$

Par la méthode de Seidel. Arrêter les calculs dès que :

$$\left| x_i^{(k+1)} - x_i^{(k)} \right| < 10^{-2}$$

Exercice 217. Résoudre le système d'équations linéaires suivant :

$$\begin{cases} 10x_1 - 2x_2 - 2x_3 & = & 6 \\ -x_1 + 10x_2 - 2x_3 & = & 7 \\ -x_1 - x_2 + 10x_3 & = & 8 \end{cases}$$

Par la méthode de relaxation. Faire les calculs avec deux décimales.

Exercice 218. Résoudre le système d'équations linéaires suivant :

$$\begin{cases} 10x_1 + x_2 + x_3 & = & 12 \\ 2x_1 + 10x_2 + x_3 & = & 13 \\ 2x_1 + 2x_2 + 10x_3 & = & 14 \end{cases}$$

Par la méthode de relaxation. Faire les calculs avec quatre décimales.