Cours no 8
Le 16 mars 2012

Mathématiques appliquées et numériques

Licence 3, Dpt Géosciences
Année 2011-2012, 2e semestre

Présentation synthétique du cours

Janvier – Juin 2012

Cours donné en 3^e année de
Licence de Sciences de la planète Terre
par Michael Ghil et Jean Roux
TD par Mohamadou Diallo
École normale supérieure, Paris

Huitième cours

Algèbre linéaire III. Résolution des systèmes linéaires $Ax = b$ par des méthodes numériques itératives.

Nous avons déjà signalé que le domaine d'application privilégié de ces méthodes est, essentiellement, celui de la résolution des systèmes linéaires à matrice creuse ou, au moins, possédant une structure (tridiagonale, pentadiagonale, structure bande, tridiagonale par blocs, etc.). Généralement ce type de matrice est obtenu par discrétisation des équations aux dérivées partielles qui fournit usuellement les trois types de matrices suivantes :

- symétriques définies positives,
- strictement diagonalement dominantes,
- irréductibles (voir livre pour cette notion) et diagonalement dominantes.

8.1 Généralités

Construction et convergence des méthodes itératives Il s'agit toujours de résoudre

$$Ax = b \tag{8.1.1}$$

où $A \in \mathbb{C}^{n,n}$ est inversible et $b \in \mathbb{C}^n$. Résoudre (8.1.1) par une méthode itérative, c'est construire une suite de vecteurs $x^{(m)}$ telle que $x^{(m)} \underset{m\to\infty}{\longrightarrow} x$ où x est solution de (8.1.1).

Dans tout ce cours 8, la suite $x^{(m)}$ est définie, de façon générale, par le schéma itératif

$$x^{(m+1)} = Bx^{(m)} + c, \tag{8.1.2}$$

où la matrice B est construite à partir de la matrice A, la matrice B et le vecteur c sont *indépendants* de l'indice m de l'itération. La matrice B est appelée *matrice d'itération*. Pour que la limite de la suite soit égale à x il faut que

$$c = (I - B)x = (I - B)A^{-1}b \tag{8.1.3}$$

cette condition doit être vérifiée pour toute construction de la matrice B.

Définition 8.1.1. *La méthode itérative (8.1.2) est dite convergente si et seulement si, pour tout $x^{(0)} \in \mathbb{C}^n$, alors $x^{(m)} \to x$ lorsque $m \to \infty$.*

Donnons immédiatement des critères de convergence pour une méthode itérative. Ces critères s'appuient sur le théorème ci-après (voir livre) :

Théorème 8.1.1. *Soit A une matrice carrée donnée, les quatre conditions suivantes sont équivalentes :*

- *(a)* $A^k \xrightarrow[k\to\infty]{} 0$,
- *(b)* $A^k x \xrightarrow[k\to\infty]{} 0$, *ceci pour tout* x,
- *(c)* $\rho(A) < 1$,
- *(d) Il existe (au moins) une norme matricielle* $\|.\|$ *telle que* $\|A\| < 1$.

Alors, pour une méthode itérative, on en déduit les critères de convergence suivants :

Théorème 8.1.2. *Les trois propositions suivantes sont équivalentes*

- *(1) La méthode est convergente.*
- *(2)* $\rho(B) < 1$ *(on rappelle que* $\rho(B)$ *est le rayon spectral de* B*).*
- *(3) Il existe une norme matricielle* $\|.\|$ *telle que* $\|B\| < 1$.

Preuve : Soit $x = A^{-1}b$, on pose $\epsilon^{(m)} = x^{(m)} - x$, $m = 0, 1, \cdots$. On remarque, en passant à la limite dans (8.1.2), que x est solution de

$$x = Bx + c.$$

Par soustraction, d'après (8.1.2), il vient donc, pour $m = 0, 1, \cdots$

$$\epsilon^{(m+1)} = B\epsilon^{(m)},$$

soit

$$\epsilon^{(m)} = B^m \epsilon^{(0)} \tag{8.1.4}$$

Pour que $\epsilon^{(m)} \to 0$ lorsque $m \to \infty$, quel que soit $\epsilon^{(0)} \in \mathbb{C}^n$, il faut et il suffit que $\lim_{m\to\infty} B^m = 0$. Les équivalences du Théorème 8.1.2 se déduisent alors du Théorème 8.1.1. $\square$

La question est maintenant de construire simplement la matrice d'itération B et le vecteur c associé donné par (8.1.3).

Nous verrons que toutes les méthodes qui seront exhibées entrent dans le cadre de la définition suivante

Définition 8.1.2. *Soit* $A \in \mathbb{C}^{n,n}$, *on appelle décomposition de* A *toute manière d'écrire* A *sous la forme* $A = M - N$, $M \in \mathbb{C}^{n,n}$ *et* $N \in \mathbb{C}^{n,n}$, *où la matrice* M *est* régulière.

À une décomposition choisie $M-N$ de la matrice A on associe la méthode itérative

$$Mx^{(m+1)} = Nx^{(m)} + b \tag{8.1.5}$$

qui est bien de la forme (8.1.2), (8.1.3) avec $B = M^{-1}N$ et

$$c = M^{-1}b = M^{-1}AA^{-1}b = M^{-1}(M-N)A^{-1}b = (I-B)A^{-1}b.$$

La formule (8.1.5) fournit donc une méthode qui convient.

L'étude des méthodes itératives se ramène à l'étude des deux problèmes suivants

- 1. Étant donné une méthode itérative de matrice d'itération B, déterminer si la méthode est convergente, c'est-à-dire examiner si $\rho(B) < 1$ ou, de façon équivalente, exhiber une norme matricielle $\|.\|$ telle que $\|B\| < 1$.

- 2. Étant donné deux méthodes itératives convergentes, de matrices d'itération B_1 et B_2, les comparer. La méthode la plus rapide sera celle dont la matrice d'itération aura le plus petit rayon spectral.

8.2 Les méthodes itératives de Jacobi, Gauss-Seidel et de relaxation successive

Il y a deux types de méthodes itératives, les méthodes ponctuelles et les méthodes par blocs inspirées des premières. Nous n'étudierons pas ces dernières (voir livre).

Après avoir introduit les deux méthodes de Jacobi et de Gauss-Seidel, on généralisera Gauss-Seidel pour obtenir les méthodes de relaxation.

Ces méthodes itératives ont en commun que chaque itération nécessite un nombre d'op.él. du même ordre de grandeur que celui nécessaire au produit d'un vecteur par une matrice (i.e. environ n^2 op.él. par itération). Il faut ici insister sur le fait que n peut être très grand, quelquefois de l'ordre du million (ou plus), et on a grand intérêt à avoir des méthodes qui convergent vite, ou au moins à utiliser la méthode la plus rapide possible.

Soit $A = (a_{ij})$ une matrice complexe d'ordre n. On pose

$$D = \mathrm{diag}(a_{ii}), E = (e_{ij}) \text{ où } e_{ij} = \begin{cases} 0 & \text{si} \quad j \geq i \\ -a_{ij} & \text{si} \quad j < i \end{cases}, \tag{8.2.1}$$

$$F = (f_{ij}) \text{ où } f_{ij} = \begin{cases} -a_{ij} & \text{si} \quad j > i \\ 0 & \text{si} \quad j \leq i \end{cases}. \tag{8.2.2}$$

ce que l'on peut schématiser ainsi (Figure 8.1).

On a donc

$$A = D - E - F.$$

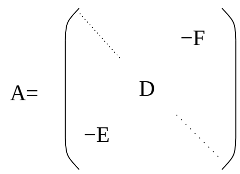

Figure 8.1:

Hypothèse fondamentale On a supposé A inversible. On suppose de plus que $a_{ii} \neq 0$, $1 \leq i \leq n$; de sorte que D est inversible et $D^{-1} = \text{diag}(1/a_{ii})$. $\square$

Dans le cadre de la discrétisation des équations aux dérivées partielles (EDP) par différences finies ou éléments finis, l'hypothèse ci-dessus est vérifiée dans les bons cas où on peut exhiber des théorèmes d'existence et d'unicité de la solution de l'EDP. De ce point de vue ce n'est pas une hypothèse restrictive.

Nous allons considérer trois exemples de décomposition de A faisant intervenir les matrices D, E et F dans la définition des matrices M et N intervenant dans la méthode itérative générale (8.1.5).

La méthode de Jacobi

On utilise la décomposition $M = D$ et $N = E + F = D - A$. On a bien $A = M - N$ avec M inversible. La matrice d'itération associée est alors

$$B = M^{-1}N = D^{-1}(E+F) = D^{-1}(D-A) = I - D^{-1}A,$$

et

$$c = M^{-1}b = D^{-1}b = (D^{-1}A)(A^{-1}b) = (I-B)A^{-1}b;$$

on retrouve la condition nécessaire (8.1.3) sur le vecteur c.

La méthode (8.1.5) est alors appelée la méthode itérative de *Jacobi*, elle s'écrit

$$Dx^{(m+1)} = (E+F)x^{(m)} + b, \quad x^{(0)} \in \mathbb{C}^n \quad \text{donné ;} \tag{8.2.3}$$

ou encore

$$a_{ii}x_i^{(m+1)} = -\sum_{\substack{j=1 \\ j \neq i}}^{n} a_{ij}x_j^{(m)} + b_i, \quad 1 \leq i \leq n, \tag{8.2.4}$$

soit

$$x_i^{(m+1)} = \frac{1}{a_{ii}}\left(-\sum_{j=1, j\neq i}^{n} a_{ij}x_j^{(m)} + b_i\right), \quad 1 \leq i \leq n. \tag{8.2.5}$$

La méthode utilise $2n$ mémoires pour stocker toutes les composantes des deux vecteurs itérés successifs $x^{(m)}$ et $x^{(m+1)}$.

La méthode de Gauss-Seidel

Intuitivement, il semble que la convergence de la méthode précédente sera améliorée si pour calculer $x_i^{(m+1)}$ on utilise les $(i-1)$ premières composantes de $x^{(m+1)}$ déjà calculées (ce que l'on appelle les *composantes actualisées*) au lieu d'utiliser les $(i-1)$ premières composantes de $x^{(m)}$. On montre (voir livre pour plus d'informations) que, pour une large classe de matrices, cette amélioration de la convergence est théoriquement justifiée. Par exemple, lorsque la matrice A est tridiagonale et définie positive, les méthodes ponctuelles de Jacobi et de Gauss-Seidel convergent simultanèment et $\rho(\mathcal{L}_1) \leq (\rho(B))^2$. La méthode de Gauss-Seidel converge plus vite que Jacobi (en un certain sens deux fois plus vite (voir livre)). Les relations (8.2.4) sont donc modifiées comme suit

$$a_{ii}x_i^{(m+1)} = -\sum_{j=1}^{i-1} a_{ij}x_j^{(m+1)} - \sum_{j=i+1}^{n} a_{ij}x_j^{(m)} + b_i, \quad 1 \leq i \leq n, \tag{8.2.6}$$

en faisant la convention $\sum_p^q = 0$ si $p > q$.

Sous forme matricielle cela s'écrit

$$(D-E)x^{(m+1)} = Fx^{(m)} + b, \tag{8.2.7}$$

ce qui correspond à la décomposition $M = D - E$ et $N = F$ (on a bien $A = D - E - F = M - N$). La matrice M est bien inversible car $(D-E)$ est inversible puisqu'elle est triangulaire supérieure avec des éléments diagonaux $a_{ii} \neq 0$ (D est supposée inversible).

C'est la méthode ponctuelle dite de *Gauss-Seidel* dite aussi des *déplacements successifs.*

La matrice d'itération $\mathcal{L}_1 = (D-E)^{-1}F$ est appelée la matrice de Gauss-Seidel associée à A. En posant $L = D^{-1}E$ et $U = D^{-1}F$ on a

$$\mathcal{L}_1 = (I-L)^{-1}U. \tag{8.2.8}$$

N.B. : On verra ci-après la raison de la notation $\mathcal{L}_1$. $\square$

Remarque 8.2.1. *Cette méthode n'exige que n mémoires pour conserver les composantes des deux vecteurs itérés successifs, la i^{ieme} composante de $x^{(m+1)}$ venant "écraser", dès qu'elle est calculée, la i^{ieme} composante de $x^{(m)}$ devenue inutile.*

Remarque 8.2.2. *L'examen des formules (8.2.4) et (8.2.6), définissant les méthodes itératives de Jacobi et de Gauss-Seidel, montre que ces méthodes sont très faciles à programmer.*

La méthode de relaxation successive

Supposons qu'à partir du vecteur $x^{(m)}$ on ait calculé, par une méthode qui reste à définir, les $(i-1)$ premières composantes de $x^{(m+1)}$. On calcule alors le nombre $\tilde{x}_i^{(m+1)}$ comme par la méthode de Gauss-Seidel, c'est-à-dire que

$$a_{ii}\tilde{x}_i^{(m+1)} = -\sum_{j=1}^{i-1} a_{ij}x_j^{(m+1)} - \sum_{j=i+1}^{n} a_{ij}x_j^{(m)} + b_i, \quad 1 \le i \le n, \tag{8.2.9}$$

la i^{ieme} composante de $x^{(m+1)}$ est alors donnée par

$$x_i^{(m+1)} = x_i^{(m)} + \omega(\tilde{x}_i^{(m+1)} - x_i^{(m)}) \tag{8.2.10}$$

où $\omega \in \mathbb{R}$, $\omega \neq 0$. Le nombre $x_i^{(m+1)}$ est donc une moyenne pondérée de l'ancienne composante $x_i^{(m)}$ et du nombre auxiliaire $\tilde{x}_i^{(m+1)}$. Si $\omega = 1$, $x_i^{(m+1)} = \tilde{x}_i^{(m+1)}$ avec $\tilde{x}_i^{(m+1)}$ calculé, à partir des composantes de $x^{(m)}$, par Gauss-Seidel. L'idée est donc de faire mieux que Gauss-Seidel, sachant que l'on retrouve cette méthode pour $\omega = 1$. On espère trouver un ω appartenant à un intervalle de confiance, encadrant la valeur 1, où la méthode converge.

En combinant les formules (8.2.9) et (8.2.10) on obtient, en multipliant (8.2.10) par a_{ii} (par hypothèse $a_{ii} \neq 0$), pour $i = 1, 2, \cdots$

$$a_{ii}x_i^{(m+1)} = a_{ii}x_i^{(m)} + \omega \left\{ -\sum_{j=1}^{i-1} a_{ij}x_j^{(m+1)} - \sum_{j=i+1}^{n} a_{ij}x_j^{(m)} + b_i - a_{ii}x_i^{(m)} \right\}, \tag{8.2.11}$$

relation qui donne directement $x_i^{(m+1)}$ en fonction des $(i-1)$ premières composantes de $x^{(m+1)}$ déjà calculées, et des $(n-i+1)$ dernières composantes de $x^{(m)}$.

Remarque 8.2.3. *La phrase introductive à cette méthode est maintenant éclaircie !*

Remarque 8.2.4. *On fait la même observation que pour Gauss-Seidel pour l'encombrement mémoire.*

En notation matricielle on exprime (8.2.11) par

$$(D - \omega E)x^{(m+1)} = ((1-\omega)D + \omega F)\, x^{(m)} + \omega b, \tag{8.2.12}$$

la matrice $(D - \omega E)$ est inversible pour tout ω car $a_{ii} \neq 0$ et (8.2.12) correspond à la décomposition

$$M = \frac{1}{\omega}(D - \omega E) \text{ et } N = \frac{1}{\omega}\left((1-\omega)D + \omega F\right). \tag{8.2.13}$$

On vérifie que

$$M - N = \frac{D}{\omega} - E - \frac{1-\omega}{\omega}D - F = D - E - F = A.$$

On a ainsi défini la méthode de relaxation successive associée à A. On a

$$\mathcal{L}_\omega = M^{-1}N = (D - \omega E)^{-1}\left((1-\omega)D + \omega F\right), \tag{8.2.14}$$

en posant $L = D^{-1}E$ et $U = D^{-1}F$, il vient

$$\mathcal{L}_\omega = (I - \omega L)^{-1}\left((1-\omega)I + \omega U\right). \tag{8.2.15}$$

Le paramètre ω est le paramètre de relaxation. Si $\omega > 1$ (resp. $\omega < 1$) on dit qu'il y a sur-relaxation (resp. sous-relaxation). Si $\omega = 1$ on a déjà fait remarquer que l'on retrouve la méthode de Gauss-Seidel, ce qui justifie la notation $\mathcal{L}_1$ utilisée à son sujet.

Pour cette méthode il s'agit de trouver

- Un intervalle de confiance $[\omega_m, \omega_M]$ (contenant le nombre 1) tel que $\rho(\mathcal{L}_\omega) < 1$ pour $\omega_m < \omega < \omega_M$.
- Dans cet intervalle, s'il est non vide, un paramètre optimal ω_{opt}, s'il existe, tel que

$$\rho(\mathcal{L}_{\omega_{opt}}) = \inf\left\{\rho(\mathcal{L}_\omega); \quad \omega_m < \omega < \omega_M\right\},$$

  en espérant que $\rho(\mathcal{L}_{\omega_{opt}}) < \rho(\mathcal{L}_1)$.

On prouve (voir livre)

Corollaire 8.2.1. *Pour toute matrice A, une condition nécessaire de convergence de la méthode de relaxation successive est que $0 < \omega < 2$.*

8.3 Théorèmes de convergence dans le cas des matrices strictement diagonalement dominantes

Généralement on ne sait rien dire de la convergence et de la comparaison des méthodes itératives lorsque la matrice A n'a pas de propriété(s) particulière(s), telle(s) que symétrique définie positive ou diagonalement dominante. On peut exhiber une matrice A_1 telle que la méthode de Jacobi converge alors que Gauss-Seidel diverge ; on peut trouver aussi une autre matrice A_2 telle que, inversement, Gauss-Seidel converge alors que Jacobi diverge.

Cependant il existe grossièrement deux grandes classes de résultats de convergence. L'une relative aux matrices symétriques (hermitiennes) définies positives, l'autre aux matrices strictement diagonalement dominantes (voir livre pour les définitions).

Naturellement ces résultats sont distincts entre eux. Par exemple, une matrice (symétrique) peut-être diagonalement dominante sans être définie positive et inversement. Comme, en pratique, les deux classes de matrices précédentes sont fréquentes (lors de la discrétisation des équations aux dérivées partielles), il faut examiner les deux catégories de résultats. Dans le cadre très restreint de ce cours nous examinerons, pour donner l'idée, que le cas d'une matrice strictement digonalement dominante.

Soit la

Définition 8.3.1. *Une matrice $A \in \mathcal{C}^{n,n}$ est dite à diagonale dominante si et seulement si*

$$|a_{i,i}| \geq \sum_{j=1,j\neq i}^{n} |a_{i,j}|, \qquad i = 1, 2, \cdots, n \tag{8.3.1}$$

avec inégalité stricte pour au moins un indice i.

Elle est dite à diagonale strictement dominante (SDD) s'il y a inégalité stricte dans (8.3.1) pour tout i.

Remarque 8.3.1. *Certains auteurs préfèrent dire fortement dominante au lieu de dominante. Réservant le nom de "dominante" au cas où il n'y a jamais d'inégalité stricte dans (8.3.1).*

On prouve (voir livre) qu'une matrice SDD est régulière. Notre hypothèse fondamentale est satisfaite.

Pour ces matrices on a le

Théorème 8.3.1. *Soit A une matrice strictement diagonalement dominante, alors la méthode ponctuelle de Jacobi converge.*

Preuve : Notons d'abord, qu'avec cette hypothèse, l'écriture de Jacobi, $B = M^{-1}N$ avec $M = D$ et $N = E + F$, a un sens car la matrice diagonale est régulière car nécessairement $|a_{i,i}| > 0$ lorsque A est SDD.

Une condition suffisante de convergence est $\rho(B) < 1$. Prouvons-le en raisonnant par l'absurde.

Soit λ une valeur propre de B, par définition nous avons

$$\det(M^{-1}N - \lambda I) = 0.$$

Comme la matrice M est régulière, de façon équivalente, il vient

$$\det(N - \lambda M) = \det(M) \det(M^{-1}N - \lambda I) = 0. \tag{8.3.2}$$

Posons $C = N - \lambda M$, $C = (c_{i,j})$. Supposons que $|\lambda| \geq 1$. On a, d'une part, puisque M est diagonale

$$\sum_{j=1,j\neq i}^{n} |c_{i,j}| = \sum_{j=1,j\neq i}^{n} |a_{i,j}|, \quad \text{ et } |c_{i,i}| = |\lambda||a_{i,i}| \;; \tag{8.3.3}$$

d'autre part, puisque A est SDD et $|\lambda| \geq 1$,

$$\sum_{j=1,j\neq i}^{n} |a_{i,j}| < |a_{i,i}| \leq |\lambda||a_{i,i}| = |c_{i,i}|. \tag{8.3.4}$$

Les relations (8.3.3) et (8.3.4) impliquent que la matrice C est aussi SDD si $|\lambda| \geq 1$. La matrice C est alors régulière et donc $\det(N - \lambda M) \neq 0$, ce qui est absurde d'après (8.3.2). Donc λ telle que $|\lambda| \geq 1$ ne peut pas être valeur propre de B et nécessairement $\rho(B) < 1$. La méthode ponctuelle de Jacobi converge si A est SDD. $\square$

Ce théorème n'énonce qu'une condition suffisante de convergence. Il se peut que Jacobi converge sans qu'elle soit SDD !

On montre (voir livre), par la même technique, que la méthode ponctuelle de Gauss-Seidel et la méthode de relaxation successive pour $0 < \omega \leq 1$ convergent si A est SDD.

Théorème 8.3.2. *Lorsque A est SDD, la méthode ponctuelle de Gauss-Seidel converge.*

Théorème 8.3.3. *Si A est SDD, la méthode de relaxation successive $\mathcal{L}_\omega$ est convergente pour $0 < \omega \leq 1$.*

Remarque 8.3.2. *Ce théorème n'est qu'une condition suffisante de convergence. Il ne dit pas que si $\omega > 1$ la méthode de relaxation diverge. Soit en effet la matrice*

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

Pour $0 < a < 1$ cette matrice est SDD. Il est facile de voir que la matrice $\mathcal{L}_\omega$ s'écrit

$$\mathcal{L}_\omega = \begin{pmatrix} 1-\omega & -\omega a \\ 0 & 1-\omega \end{pmatrix},$$

sa valeur propre (double) est $1-\omega$. La méthode de relaxation converge si $|1-\omega| < 1$, c'est-à-dire pour $0 < \omega < 2$.