

# Chapitre III. Méthodes itératives de résolution des systèmes linéaires

M. Granger

## Table des matières

1	Les méthodes de Jacobi et de Gauss-Seidel	2
2	Etude de conditions de convergence des méthodes itératives	3
3	Méthode de relaxation (hors programme)	5
4	Un exemple de comparaison des vitesses de convergence	7
5	Un exemple (TD)	9

On considère comme dans le chapitre 2 un système

$$Au = b$$

avec  $A$  inversible. Il est parfois plus intéressant de chercher à atteindre l'unique solution  $\underline{u}$  de ce système comme une limite de suite et on se propose dans ce chapitre de décrire deux méthodes parmi les plus classiques pour atteindre ce but.

On écrit  $A$  sous la forme  $A = M - N$ , avec  $M$  inversible, d'inverse réputé facile à calculer, comme par exemple  $A$  diagonale ou triangulaire. On pose  $B = M^{-1}N$ , ce qui donne la formule

$$A = M - N = M(I - M^{-1}N) = M(I - B),$$

et l'inversibilité de  $A$  revient à la condition :

$$I - B \text{ est inversible,}$$

De plus on a, en posant  $c = M^{-1}b$  :

$$Au = b \iff Mu = Nu + b \iff u = Bu + c$$

On note  $\underline{u}$  l'unique solution du système et la méthode itérative associée à ces données, de matrice d'itération  $B = M^{-1}N$ , consiste à définir une suite récurrente :

$$(1) \quad u_0 \in \mathbb{C}, \quad u_{k+1} = Bu_k + c.$$

qu'on peut aussi caractériser si on veut éviter d'introduire l'inverse de  $M$  par la formule :

$$(2) \quad u_0 \in \mathbb{C}, \quad Mu_{k+1} = Nu_k + b.$$

**Définition 0.1** *On dit que la méthode itérative associée à une matrice  $A$  et à une décomposition  $A = M - N$  converge si et seulement si, pour tout second membre  $b$  et toute condition initiale  $u_0$  la suite de terme général  $u_k$  converge vers la solution  $u$ .*

L'objet de ce chapitre est d'étudier quelques exemples de méthodes itératives et quelques conditions suffisantes ou nécessaires et suffisantes de convergence.

*Note.*- La section 3 sur la méthode de relaxation est hors programme, ce qui signifie qu'il n'y aura pas d'exercice portant sur cette section ni de sujet d'examen portant de près ou de loin sur ce sujet. Il en est de même des quelques lignes à la fin de la section 4 et de la section 5 où il est question la méthode de relaxation. Je l'ai toutefois laissée dans ces notes à des fins heuristiques, pour donner une idée sur les raffinements possibles des méthodes basiques qui sont développées dans ce chapitre.

## 1 Les méthodes de Jacobi et de Gauss-Seidel

On suppose que l'écriture  $A = M - N$ , avec  $M$  inversible, est choisie de façon que  $M$  soit facile à inverser. On entend par là que le temps de calcul pour inverser  $A$  est d'un ordre de grandeur inférieur au cas général, qui est  $\frac{2}{3}n^3$  selon le chapitre I.

**Hypothèse et notations.**- On suppose désormais que la condition suivante est satisfaite :

$$\forall i \in \{1, \dots, n\}, A_{i,i} \neq 0.$$

On note respectivement  $D = \text{diag}(A_{1,1}, \dots, A_{n,n})$ ,  $-E$ , et  $-F$  la partie diagonale, triangulaire supérieure stricte et triangulaire inférieure stricte de la matrice  $A$ . On symbolise cela par :

$$A = \begin{pmatrix} a_{1,1} & & & \\ & \ddots & & -F \\ & & -E & \\ & & & \ddots & \\ & & & & a_{n,n} \end{pmatrix}.$$

On entend par là que chacune des notations  $E$  et  $F$  désigne une matrice carrée ayant un triangle de zéros opposé au triangle strict où on a figuré  $-E$  et  $-F$ . Elles sont donc caractérisées par :

$$A = D - E - F, \quad \text{et} \quad \forall (i, j) \in \{1, \dots, n\}^2, \quad \begin{cases} D_{i,j} = 0, & \text{si } i \neq j \\ E_{i,j} = 0 & \text{si } i \leq j \\ F_{i,j} = 0 & \text{si } i \geq j \end{cases}$$

Les deux méthodes itératives les plus simples sont alors :

- *La méthode de Jacobi* : On choisit  $M = D$ , diagonale. La matrice  $B$  correspondante devient ici :

$$\boxed{J = D^{-1}(E + F)}$$

appelée matrice de Jacobi. Notons  $u_k = {}^t(u_{1,k}, \dots, u_{n,k})$  le  $k$ ième vecteur de l'itération. Le passage du vecteur  $u_k$  au suivant  $u_{k+1} = D^{-1}(E + F)u_k + D^{-1}b$  s'obtient en résolvant le système diagonal aux inconnues  $u_{j,k+1}$ ,  $k = 1, \dots, n$  :

$$\begin{cases} a_{1,1}u_{1,k+1} & = & -a_{1,2}u_{2,k} - \dots - a_{1,n}u_{n,k} + b_1 \\ a_{2,2}u_{2,k+1} & = & -a_{2,1}u_{1,k} - \dots & - a_{2,n}u_{n,k} + b_2 \\ \dots & & \dots & \dots \\ \dots & & \dots & \dots \\ a_{n,n}u_{n,k+1} & = & -a_{n,1}u_{1,k} - \dots - a_{n,n-1}u_{n-1,k} & + b_n \end{cases}$$

- *La méthode de Gauss-Seidel* : Dans cette méthode on prend  $M = D - E$ , qui est donc une matrice triangulaire inférieure inversible. La matrice  $B$  devient ici la matrice de Gauss-Seidel notée :

$$\boxed{\mathcal{L}_1 = (D - E)^{-1}F}$$

La complication supplémentaire n'est qu'apparente car l'itération  $u_{k+1} = (D - E)^{-1}Fu_k + (D - E)^{-1}b$ , peut se mettre sous la forme  $(D - E)u_{k+1} + Fu_k + b$  soit finalement  $Du_{k+1} = Eu_{k+1} + Fu_k + b$ , ce qui donne le système :

$$\begin{cases} a_{1,1}u_{1,k+1} & = & -a_{1,2}u_{2,k} & - \cdots & & -a_{1,n}u_{n,k} & + b_1 \\ a_{2,2}u_{2,k+1} & = & -a_{2,1}u_{1,k+1} & - \cdots & & -a_{2,n}u_{n,k} & + b_2 \\ \cdots & & \cdots & & \cdots & & \\ \cdots & & \cdots & & \cdots & & \\ a_{n,n}u_{n,k+1} & = & -a_{n,1}u_{1,k+1} & - \cdots & -a_{n,n-1}u_{n-1,k+1} & & + b_n \end{cases}$$

On constate que la programmation de ce système est en fait plus économique en termes de place mémoire :

on programme le passage du vecteurs  $u_k$  au vecteur  $u_{k+1}$  par une boucle en commençant par la première équation. A la ligne  $j$ , l'équation s'écrit :

$$a_{j,j}u_{j,k+1} = - \sum_{i=1}^{j-1} a_{i,j}u_{i,k+1} - \sum_{i=j+1}^n a_{i,j}u_{i,k}$$

et on voit que le stockage des dernières coordonnées  $u_{j+1,k}, \dots, u_{n,k}$  de  $u_k$  est devenu inutile lorsqu'on passe à la  $j$ ème ligne.

Il suffit donc de déclarer pour tout  $j$  une variable  $j$ ème coordonnée : Au passage de l'équation d'indice  $j$  on calcule et on affecte à la  $j$ ème variable la valeur  $u_{j,k+1}$ . Elle remplace  $u_{j,k}$  qui n'est plus utilisée ensuite.

Remarque : Pour expliquer la notation on verra dans la section suivante que  $\mathcal{L}_1$  est un cas particulier de matrice  $\mathcal{L}_\omega$  liée à la méthode de relaxation.

## 2 Etude de conditions de convergence des méthodes itératives

**Théorème 2.1** .- Il y a équivalence entre les trois conditions suivantes :

1.  $\rho(B) < 1$ .
2. Il existe une norme  $\|\bullet\|$  sur  $\mathbb{C}^n$ , telle que pour la norme subordonnée  $\|B\| < 1$ .
3.  $B^k e_0$  tend vers zéro pour tout vecteur  $e_0$  de  $\mathbb{C}^n$ .
4. La suite  $u_k$  tend vers la solution  $u$  pour toute condition initiale  $u_0$ .

L'équivalence entre les trois premières conditions a été vue au chapitre I.

Notons  $e_k = u_k - u$ , où  $u$  est la solution du système. Des équations  $u_{k+1} = Bu_k + c$  et  $u = Bu + c$ , on tire :

$$e_{k+1} = Be_k, \text{ puis } e_k = B^k e_0$$

La condition 4), qui se lit aussi  $\forall u_0 \in \mathbb{C}^n, \lim_{k \rightarrow \infty} e_k = 0$ , équivaut donc à :

$$\forall e_0 \in \mathbb{C}^n, \lim_{k \rightarrow \infty} B^k e_0 = 0,$$

c'est à dire à la condition 3).

Remarque.- Si la matrice  $B$  d'une méthode itérative possède une valeur propre  $\lambda > 1$ , l'itération diverge presque toujours, i.e. en dehors d'un ensemble de mesure nulle. Il suffit de mettre la matrice sous forme triangulaire supérieure,  $P^{-1}AP = T$ , avec  $T_{n,n} = \lambda$ . Alors si  $e_0$  a une dernière coordonnée non nulle  $\alpha$ , la dernière coordonnée de  $e_k = u_k - u$  est égale à  $\alpha\lambda^k$  et donc tend vers l'infini. Comme l'espace caractéristique est inconnu, on ne peut pas maîtriser la non nullité de cette coordonnée, et les erreurs d'arrondi suffiraient de toutes façons à provoquer le comportement divergent. De même si

$|\lambda| = 1$  les vecteurs  $e_k$  ont une coordonnée de module constante donc ne peuvent encore pas tendre vers zéro.

En conclusion et au vu de la définition donnée en introduction on obtient :

**Proposition 2.2** .- La méthode itérative, associée à  $A = M - N$  et  $B = M^{-1}N$  est convergente si et seulement si :

$$\boxed{\rho(B) < 1}$$

Voici deux exemples de situations où la convergence fonctionne. Un premier exemple pour toute méthode, un deuxième pour les méthodes de Jacobi ou Gauss-Seidel :

**Théorème 2.3** .- On suppose que  $A = M - N$  hermitienne est définie positive. Alors :

1)  $M^* + N$  est hermitienne.

2) Si  $M^* + N$  est en plus définie positive, on a l'inégalité  $\rho(M^{-1}N) < 1$

Dans la situation du théorème, la méthode itérative  $u_{k+1} = M^{-1}Nu_k + M^{-1}b$  converge.

**Démonstration du théorème 2.3** .-

1)  $(M^* + N)^* = M + N^* = A + (N + N^*)$ , est bien hermitienne puisque  $A$  l'est :  $[A + (N + N^*)]^* = A^* + N^* + N = A + N + N^*$ .

2) On va utiliser la norme adaptée au problème, en l'occurrence la norme issue du produit scalaire hermitien associé à la matrice hermitienne  $A$  :

$$v \mapsto \|v\|_A = (v^*Av)^{\frac{1}{2}}$$

On va montrer que pour la norme matricielle associée :

$$\| \|M^{-1}N\| \|_A < 1$$

ce qui suffit à établir la convergence, puisque, comme pour toute norme matricielle associée,  $\rho(M^{-1}N) \leq \| \|M^{-1}N\| \|_A$ .

Remarquons d'abord les égalités  $M^{-1}N = M^{-1}(M - A) = I - M^{-1}A$ , d'où :

$$\| \|M^{-1}N\| \|_A = \sup_{\|v\|_A=1} \|v - M^{-1}Av\|_A$$

On choisit  $v \in \mathbb{C}^n$  tel que  $v^*Av = \|v\|_A^2 = 1$  et on pose  $w = M^{-1}Av$ .

Alors, en appliquant  $Av = Mw$ , et aussi  $v^*A = w^*M^*$  :

$$\begin{aligned} \|v - M^{-1}Av\|_A^2 &= \|v - w\|_A^2 = v^*Av - v^*Aw - w^*Av + w^*Aw \\ &= 1 - w^*M^*w - w^*Mw + w^*Aw \\ &= 1 - w^*(M^* + M - A)w = 1 - w^*(M^* + N)w \end{aligned}$$

Puisque  $M^* + N$  est définie positive, on a  $w^*(M^* + N)w > 0$  ce qui donne bien au final la relation cherchée :

$$\|v - M^{-1}Av\|_A^2 = 1 - w^*(M^* + N)w < 1$$

□

L'exercice 1 propose une application de ce résultat. Le théorème s'applique à la matrice tridiagonale suivante :

$$S_n = \begin{pmatrix} 2 & -1 & 0 & \cdots & \cdots & 0 \\ -1 & 2 & -1 & & & 0 \\ 0 & \ddots & \ddots & \ddots & & \\ & & & \cdots & & \\ & & & & \ddots & \ddots & \ddots \\ 0 & & & & -1 & 2 & -1 \\ 0 & \cdots & & \cdots & 0 & -1 & 2 \end{pmatrix}$$

**Proposition 2.4** .- Si la matrice  $A$  est à diagonale dominante, alors la méthode de Jacobi est convergente.

**Démonstration** Rappelons que pour toute matrice carrée  $M = (m_{i,j})$  le spectre de  $M$ , est contenu dans la réunion des disques, appelés disques de Gershgorin, de centres respectifs  $m_{i,i}$ , et de rayons  $\sum_{1 \leq j \leq n, j \neq i} |m_{i,j}|$ , si on travaille sur les lignes, et aussi dans la réunion des disques de rayons  $\sum_{1 \leq k \leq n, k \neq i} |m_{k,i}|$ . Si  $|a_{i,i}| > \sum_{1 \leq j \leq n, j \neq i} |a_{i,j}|$ , selon l'hypothèse de diagonale dominante, la matrice de Jacobi  $J =$

$$-diag\left(\frac{1}{a_{1,1}}, \dots, \frac{1}{a_{n,n}}\right) \cdot \begin{pmatrix} 0 & a_{1,2} & \cdots & a_{1,n} \\ a_{2,1} & \ddots & & \\ & & \cdots & \\ & & & \ddots \\ a_{n,1} & \cdots & a_{n,n-1} & 0 \end{pmatrix} \text{ a pour coefficients } J_{i,i} = 0, \text{ et pour } i \neq j,$$

$J_{i,j} = -\frac{a_{i,j}}{a_{i,i}}$ . Les cercles de Gershgorin associés à  $J$  sont donc tous centrés en zéro et de rayons respectifs :

$$\sum_{1 \leq j \leq n, j \neq i} |J_{i,j}| = \frac{1}{|a_{i,i}|} \left( \sum_{1 \leq j \leq n, j \neq i} |a_{i,j}| \right) < 1$$

selon l'hypothèse sur  $A$ . La majoration cherchée  $\rho(J) < 1$  s'en déduit aussitôt.  $\square$

Remarque.- Si on suppose que  $A$  est à diagonale dominante sur les colonnes, le résultat est encore vrai : il suffit de considérer la transposée de  $A$ , de remarquer que sa matrice de Jacobi est la transposée de celle de  $A$ , donc a le même spectre et le même rayon spectral.

### 3 Méthode de relaxation (hors programme)

CETTE SECTION INTITULEE METHODE DE RELAXATION EST FACULTATIVE.

Dans cette méthode on modifie la méthode de Gauss-Seidel  $A = M - N$ ,  $N = F$ , dans laquelle  $M = D - E$ , où  $D$  est la diagonale de  $A$  et  $E$  sa partie triangulaire supérieure (on pourrait aussi bien regarder son analogue triangulaire inférieur). On renonce à exiger que pour tout  $(i, j)$ ,  $m_{i,j} = a_{i,j}$  ou  $m_{i,j} = 0$ .

Précisément on pose :

$$M = \frac{D}{\omega} - E.$$

Cette matrice est bien triangulaire supérieure, à diagonale inversible, donc est inversible. Ici  $\omega \neq 0$  est appelé le paramètre de relaxation. Explicitons l'expression de la matrice d'itération qui sera notée  $\mathcal{L}_\omega$  :

$$A = D - E - F = \left(\frac{D}{\omega} - E\right) - N, \text{ avec } N = \left(\frac{1-\omega}{\omega}D + F\right)$$

La méthode itérative consiste donc à résoudre :

$$\left(\frac{D}{\omega} - E\right)u_{k+1} = \left(\frac{1-\omega}{\omega}D + F\right)u_k + b \tag{1}$$

**Notation 3.1** .- On note :  $\mathcal{L}_\omega = (\frac{D}{\omega} - E)^{-1}(\frac{1-\omega}{\omega}D + F)$ , la matrice associée à la méthode de relaxation de paramètre  $\omega$ .

Si  $\omega > 1$ , on parle de sur-relaxation, et si  $\omega < 1$ , on parle de sous-relaxation.

On sait que d'après la proposition 2.2 la convergence de la méthode équivaut à  $\rho(\mathcal{L}_\omega) < 1$ , et le cas  $\omega = 1$  est la méthode de Gauss Seidel.

L'objectif général est de trouver la valeur optimale de  $\omega$ , notée  $\omega_0$  telle que :

$$\boxed{\rho(\mathcal{L}_{\omega_0}) = \inf_{\omega \in I} \rho(\mathcal{L}_\omega)}$$

Dans ce cours nous n'entreprendrons pas d'étude systématique de ce problème. Un exemple significatif sera toutefois donné à la fin.

La complexité des formules théoriques n'est qu'apparente et le système qu'on est amené à résoudre pour mener à bien l'itération est à peu près aussi simple que dans le cas de Gauss Seidel puisque cela donne, à partir de l'équation (1) :

$$\begin{cases} \frac{1}{\omega}a_{1,1}u_{1,k+1} &= \frac{1-\omega}{\omega}u_{1,k} - a_{1,2}u_{2,k} - \cdots - a_{1,n}u_{n,k} + b_1 \\ \frac{1}{\omega}a_{2,2}u_{2,k+1} &= -a_{2,1}u_{1,k+1} - \frac{1-\omega}{\omega}u_{2,k} - \cdots - a_{2,n}u_{n,k} + b_2 \\ \cdots & \cdots \cdots \\ \cdots & \cdots \cdots \\ \frac{1}{\omega}a_{n,n}u_{n,k+1} &= -a_{n,1}u_{1,k+1} - \cdots - a_{n,n-1}u_{n-1,k+1} + \frac{1-\omega}{\omega}u_{n,k} + b_n \end{cases}$$

La remarque sur l'économie de mémoire du programme consistant à déclarer pour TOUS les vecteurs  $u_k$ , une seule variable  $j$ ème coordonnée est entièrement valable. La méthode de relaxation sera donc à préférer dans toute situation où on peut déterminer une diminution de l'ordre de grandeur du rayon spectral, donc du nombre d'itérations à effectuer.

Nous terminons cette section sur un premier résultat concernant le domaine  $I$  dans lequel on peut valablement espérer trouver un paramètre optimal de relaxation  $\omega_0$ .

**Théorème 3.2** .- La méthode de relaxation ne peut converger que sous la condition nécessaire :

$$\boxed{\omega \in ]0, 2[}$$

**Démonstration.**- Soient  $\lambda_1, \dots, \lambda_n$ , les valeurs propres de  $\mathcal{L}_\omega$  dans  $\mathbb{C}$  comptées avec multiplicités.

$$\lambda_1 \cdots \lambda_n = \det(\mathcal{L}_\omega) = \frac{\det(\frac{1-\omega}{\omega}D + F)}{\det(\frac{D}{\omega} - E)}$$

Les matrices en jeu sont triangulaires donc leur déterminants ne dépendent que de leurs diagonales :

$$\lambda_1 \cdots \lambda_n = \frac{(\frac{1-\omega}{\omega})^n \det(D)}{(\frac{1}{\omega})^n \det(D)} = (1 - \omega)^n$$

(On a utilisé la formule  $\det(\lambda A) = \lambda^n \det A$ , pour  $\lambda \in \mathbb{K}$ .)

Or on a la minoration  $|\rho(\mathcal{L}_\omega)| = \max(|\lambda_1|, \dots, |\lambda_n|) = [\max(|\lambda_1|, \dots, |\lambda_n|)^n]^{\frac{1}{n}} \geq (|\lambda_1| \cdots |\lambda_n|)^{\frac{1}{n}} = |\omega - 1|$ .

La condition nécessaire de convergence qu'on en déduit  $|\omega - 1| < 1$  est celle de l'énoncé.  $\square$

Si la matrice  $A$  est définie positive on peut même démontrer en utilisant la proposition 2.3 que la condition  $\omega \in ]0, 2[$  est une condition nécessaire et suffisante de convergence.

## 4 Un exemple de comparaison des vitesses de convergence

On se place exclusivement dans ce qui suit dans la situation où  $A$  est tridiagonale de taille  $q \times q$ , c'est à dire s'écrit

$$A = D - E - F = \begin{pmatrix} D_{1,1} & -F_{1,2} & 0 & \cdots & \cdots & 0 \\ -E_{2,1} & D_{2,2} & \ddots & 0 & & 0 \\ 0 & \ddots & \ddots & & & \\ & & & \cdots & & \\ & & & & \ddots & \ddots & 0 \\ & & & & \ddots & D_{q-1,q-1} & -F_{q-1,q} \\ 0 & \cdots & 0 & -E_{q,q-1} & D_{q,q} & & \end{pmatrix}$$

avec  $D_{i,i}, E_{i,j}, F_{i,j} \in \mathbb{C}$

Tous les résultats de cette section admettent un généralisation à des matrices tridiagonales *par blocs*. Nous n'aborderons pas cela par simple souci de simplicité, les démonstrations étant identiques.

**Théorème 4.1** : *Comparaison des méthodes de Jacobi et Gauss-Seidel.-*

*Soit  $A$  une matrice tridiagonale. Alors,*

$$\boxed{\rho(\mathcal{L}_1) = \rho(J)^2}$$

Par conséquent, les deux méthodes sont simultanément convergentes ou non convergentes et dans le cas de la convergence, la méthode de Gauss-Seidel est plus rapide.

**Démonstration.-**

**Lemme 4.2** .- *Pour  $\mu \neq 0$  on note  $A(\mu)$  la matrice, telle que  $A(1) = A$  et en général :*

$$A(\mu) = \begin{pmatrix} D_{1,1} & -\frac{1}{\mu}F_{1,2} & 0 & \cdots & \cdots & 0 \\ -\mu E_{2,1} & D_{2,2} & \ddots & 0 & & 0 \\ 0 & \ddots & \ddots & & & \\ & & & \cdots & & \\ & & & & \ddots & \ddots & 0 \\ & & & & \ddots & D_{q-1,q-1} & -\frac{1}{\mu}F_{q-1,q} \\ 0 & \cdots & 0 & -\mu E_{q,q-1} & D_{q,q} & & \end{pmatrix}$$

Alors  $\det(A(\mu)) = \det A(1) = \det(A)$ .

**Démonstration du lemme.-** Le lemme se déduit du fait que la matrice  $A(\mu)$  est même semblable à la matrice  $A = A(1)$ . La matrice diagonale rang  $r$ ,  $P = \text{diag}(\mu, \mu^2, \dots, \mu^q)$  fournit en effet la relation de conjugaison :

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

En effet l'action de conjugaison par  $P$  consiste à multiplier la  $i$ ème ligne par  $\mu^{-i}$  et la  $j$ ème colonne par  $\mu^j$ .

□

Considérons alors le polynôme caractéristique de la matrice de Jacobi  $J = D^{-1}(E + F)$ , c'est à dire :

$$p_J(\lambda) = \det(D^{-1}(E + F) - \lambda I).$$

Puisque  $D$  est inversible on peut considérer un autre polynôme  $q_J$  qui a les même racines :

$$q_J(\lambda) := \det(-D)p_J(\lambda) = \det((-D)(D^{-1}(E + F) - \lambda I)) = \det(\lambda D - E - F)$$

Par un calcul analogue le polynôme caractéristique de la matrice de Gauss-Seidel  $\mathcal{L}_1 = (D - E)^{-1}F$  est :

$$p_{\mathcal{L}_1}(\lambda) = \det((D - E)^{-1}F - \lambda I).$$

et il a les mêmes racines que polynôme  $q_{\mathcal{L}_1}$  suivant :

$$q_{\mathcal{L}_1}(\lambda) = \det(-D + E)p_{\mathcal{L}_1}(\lambda) = \det((-D + E)((D - E)^{-1}F - \lambda I) = \det(\lambda(D - E) - F)$$

Une application du lemme donne alors :

$$q_{\mathcal{L}_1}(\lambda^2) = \det(\lambda^2 D - \lambda^2 E - F) = \det(\lambda^2 D - \lambda E - \lambda F) = \lambda^n q_J(\lambda)$$

avec  $\mu = \lambda$ , et  $\lambda^2 D - \lambda^2 E - F = \lambda^2 D - \mu(\lambda E) - \frac{1}{\mu}(\lambda F)$

Cette égalité montre que  $\lambda$  est une racine non nulle de  $p_J$ , c'est à dire un élément non nul du spectre de  $J$ , si et seulement si  $\lambda^2$  est un élément non nul du spectre de  $\mathcal{L}_1$ . Ceci établit la relation cherchée  $\rho(J)^2 = \rho(\mathcal{L}_1)$  et termine la démonstration du théorème 4.1.

□

On remarque que  $\mathcal{L}_1$  admet toujours la valeur propre zéro, ce qui n'est pas obligatoire pour  $J$ . On déduit aussi de la démonstration que  $q_J(\lambda) = 0 \Leftrightarrow q_J(-\lambda) = 0$

LA FIN DE CETTE SECTION EST FACULTATIVE.

Le théorème 4.1 se généralise au cas de la méthode de relaxation, c'est à dire que dans la situation de ce théorème on sait calculer  $\rho(\mathcal{L}_\omega)$  en fonction de  $\rho(J)$ . La démonstration du théorème suivant, assez délicate bien que de nature très élémentaire, sera admise, mais le résultat est assez significatif pour qu'il vaille la peine de l'énoncer :

**Théorème 4.3** *Pour  $A$  tridiagonale par blocs, et  $J$  à valeurs propres réelles, les méthodes  $J$ ,  $\mathcal{L}_\omega$ , convergent ou non simultanément pour tout  $\omega \in ]0, 2[$ . De plus le graphe de la fonction  $\omega \mapsto \rho(\mathcal{L}_\omega)$  a l'allure indiquée dans la figure suivante. Le paramètre optimal  $\omega_0$ , pour lequel  $\rho(\mathcal{L}_\omega)$  est minimal vaut :*

$$\omega_0 = \frac{2}{1 + \sqrt{1 - \rho(J)^2}}$$

et le rayon spectral optimal correspondant est

$$\rho(\mathcal{L}_{\omega_0}) = \omega_0 - 1 = \frac{1 - \sqrt{1 - \rho(J)^2}}{1 + \sqrt{1 - \rho(J)^2}}$$

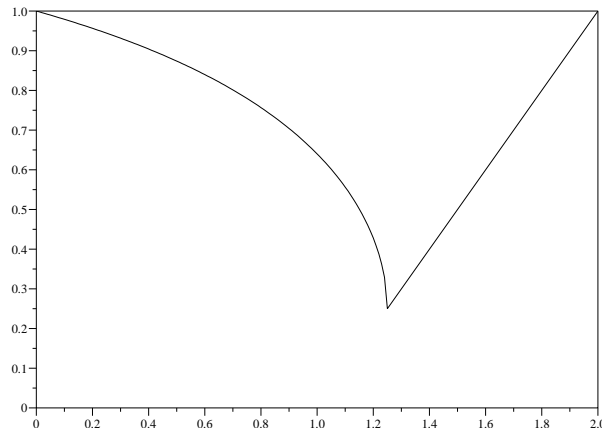


fig.1



Remarque : La fonction de  $\rho = \rho(J)$  ainsi obtenue

$$\rho \rightarrow \frac{1 - \sqrt{1 - \rho^2}}{1 + \sqrt{1 - \rho^2}}$$

est nettement meilleure que la valeur  $\rho(\mathcal{L}_1) = \rho^2$  de la méthode de Jacobi donnée par le théorème 4.1.

## 5 Un exemple (TD)

Le contenu de cette section est le corrigé de l'exercice 4 de la feuille d'exercices associée à ce chapitre. Il est donc vivement conseillé de chercher l'exercice en question à partir de son énoncé avant de lire ce qui suit.

Dans cette section on reprend la matrice

$$S_n = \begin{pmatrix} 2 & -1 & 0 & \cdots & \cdots & 0 \\ -1 & 2 & -1 & & & 0 \\ 0 & \ddots & \ddots & \ddots & & \\ & & & \ddots & \ddots & \ddots \\ 0 & & & & -1 & 2 & -1 \\ 0 & \cdots & \cdots & 0 & -1 & 2 \end{pmatrix}$$

de la section 2. On a déjà vu que les méthodes de Jacobi et de Gauss-Seidel peuvent s'appliquer, avec  $\rho(\mathcal{L}_1) = \rho(J)^2$  ainsi que la méthode de relaxation.

1) Détermination des valeurs propres et des sous espaces propres associés. Le fait que  $\lambda$  soit une valeur propre avec  $x = {}^t(x_1, \dots, x_n)$  comme vecteur propre associé se traduit par le système :

$$\begin{cases} 2x_1 - x_2 & = & \lambda x_1 \\ -x_1 + 2x_2 - x_3 & = & \lambda x_2 \\ & \ddots & \vdots \\ & & -x_{n-2} + 2x_{n-1} - x_n = \lambda x_{n-1} \\ & & -x_{n-1} + 2x_n = \lambda x_n \end{cases}$$

On pose  $x_0 = 0$ , et  $x_{n+1} = 0$ . On constate alors que le  $(n+2)$ -uplet  $(x_0, \dots, x_{n+1})$  peut s'étendre, par des valeurs uniques  $x_{n+i}$  pour  $i \geq 2$  en une suite qui est solution de l'équation de récurrence :

$$x_{i+1} = -x_{i-1} + (2 - \lambda)x_i.$$

On sait que cette équation admet une solution unique pour toute condition initiale,  $(x_0, x_1)$ . Le principe de la démonstration qui suit est d'exprimer les deux conditions "aux bords"  $x_0 = x_{n+1} = 0$ , et de chercher à quelle condition sur  $\lambda$  on perd l'unicité pour trouver une solution non nulle, donc un vecteur propre.

L'ensemble des solutions de cette récurrence linéaire est obtenue à partir de l'équation caractéristique :

$$r^2 + (\lambda - 2)r + 1 = 0$$

de discriminant  $\Delta = (\lambda - 2)^2 - 4 = \lambda^2 - 4\lambda$ , ce qui donne la discussion suivante :

cas 1.-  $\lambda \notin [0, 4]$ , c'est à dire  $\Delta > 0$ . On peut écarter ce résultat d'emblée à cause de ce qu'on a montré dans l'exercice 1). Faisons le quand même à nouveau directement.

Les racines de l'équation caractéristique sont deux réels  $r_+$  et  $r_- = \frac{1}{r_+}$ , avec pour fixer les idées,  $|r_+| > 1$ , puisque  $r_+ \neq r_-$  interdit  $r_+ = \pm 1$ .

Il existe alors  $a, b \in \mathbb{C}$  tels que :

$$\forall k \in \mathbb{N}, x_k = a(r_+)^k + b(r_-)^k.$$

Les conditions aux bords  $x_0 = x_{n+1} = 0$  se traduisent par  $\begin{cases} a + b = 0 \\ a(r_+)^{n+1} + b(r_-)^{n+1} = 0 \end{cases}$  système qui a pour unique solution  $a = b = 0$ . Ainsi le sous espace propre  $E_\lambda$  est réduit à zéro et aucun  $\lambda$  relevant de ce cas 1 n'est une valeur propre.

cas 2.- Si  $\lambda = 0$  ou  $\lambda = 4$ , il y a une racine double  $r$  valant 1 ou  $-1$ , et les solutions de l'équation récurrente sont :

$$x_k = (a + kb)r^k$$

et la condition  $x_0 = 0$ , donne  $a = 0$ . Il reste alors  $0 = x_{n+1} = br^{n+1}$  donc  $b = 0$  et on a la même conclusion que dans le cas 1.

cas 3.-  $\lambda \in ]0, 4[$ . L'équation caractéristique admet deux valeurs propres conjuguées de module 1,  $r = e^{i\theta}$  et  $\bar{r} = e^{-i\theta}$ ,  $\theta$  étant obtenu à partir de  $\lambda$  par

$$r + \bar{r} = 2 \cos \theta = -\lambda + 2$$

Les solutions de l'équation de récurrence sont

$$x_k = ae^{ik\theta} + be^{-ik\theta}$$

et la condition  $x_0 = 0$  se traduit par :  $a = -b$ ,  $x_k = a(e^{ik\theta} - e^{-ik\theta}) = 2ia \sin k\theta$ . La condition  $x_{n+1} = 0$ , donne alors l'équation en  $\theta$  qui fixe les valeurs propres :

$$\sin(n+1)\theta = 0$$

Il suffit de considérer  $\theta \in ]0, \pi[$  et on trouve les solutions :

$$\theta := \theta_m = \frac{m\pi}{n+1}, \text{ pour } m = 1, \dots, n$$

Les valeurs propres sont alors  $\lambda_m = 2 - r - \bar{r} = 2 - 2 \cos \theta_m = 4 \sin^2 \frac{\theta_m}{2}$ .

Le spectre de  $S_n$  est donc constitué par les  $n$  valeurs suivantes :

$$\lambda_m = 4 \sin^2 \frac{m\pi}{2(n+1)}, \text{ pour } m = 1, \dots, n.$$

2) Application aux rayons spectraux respectifs pour  $J, \mathcal{L}_1, \mathcal{L}_\omega$ . L'écriture  $S_n = D - N$ , avec  $D = \text{Diag}(S_n) = 2I_n$ , donne

$$N = E + F = \begin{pmatrix} 0 & 1 & 0 & \cdots & \cdots & 0 \\ 1 & 0 & 1 & & & 0 \\ 0 & \ddots & \ddots & \ddots & & \\ & & & \cdots & & \\ & & & \ddots & \ddots & \ddots \\ 0 & & & & 1 & 0 & 1 \\ 0 & \cdots & \cdots & & 0 & 1 & 0 \end{pmatrix}$$

La matrice de Jacobi est  $J = M^{-1}N = \frac{1}{2}N = I - \frac{A}{2}$ , donc les valeurs propres de  $J$ , notées  $\alpha_n$  se déduisent directement de celles de  $A$  :

$$\alpha_n = 1 - \frac{\lambda_n}{2} = 1 - 2 \sin^2 \frac{m\pi}{2(n+1)} = \cos \frac{m\pi}{n+1}, \text{ et donc } \rho(J) = \cos \frac{\pi}{n+1}$$

Le paramètre de relaxation optimale est  $\omega_0 = \frac{2}{1+\sqrt{1-\rho(J)^2}} = \frac{2}{1+\sqrt{1-\rho(J)^2}} = \frac{2}{1+\sin \frac{\pi}{n+1}}$

On obtient donc :

$$\rho(J) = \cos \frac{\pi}{n+1}, \quad \rho(\mathcal{L}_1) = \rho(J)^2 = (\cos \frac{\pi}{n+1})^2, \quad \rho(\mathcal{L}_{\omega_0}) = \omega_0 - 1 = \frac{1 - \sin \frac{\pi}{n+1}}{1 + \sin \frac{\pi}{n+1}}$$

Les développements limités par rapport à  $\frac{1}{n+1}$  :

$$\begin{aligned} \rho(J) &= 1 - \frac{\pi^2}{2(n+1)^2} + o\left(\frac{1}{(n+1)^2}\right), \\ \rho(\mathcal{L}_1) &= 1 - \frac{\pi^2}{(n+1)^2} + o\left(\frac{1}{(n+1)^2}\right), \\ \rho(\mathcal{L}_{\omega_0}) &= 1 - \frac{\pi}{n+1} + o\left(\frac{1}{n+1}\right) \end{aligned}$$

montrent que le passage à la méthode de relaxation fait ici gagner un ordre de grandeur sur  $1 - \rho$ , donc sur la vitesse de convergence.