Université Paris-Nord
Institut Galilée
Département de Mathématiques

Année 2008-2009
Licence de Mathématiques - L3
C. Basdevant - R. Chong

Corrigé de l'examen partiel d'Analyse Numérique
du mercredi 12 novembre 2008

Durée : 3h
Aucun document n'est autorisé

Exercice 1

Soit B une matrice carrée vérifiant $\|B\| < 1$ pour une certaine norme matricielle.

1. Montrez que $I - B$ est inversible.

   Corrigé : Pour toute norme matricielle (induite par une norme vectorielle ou non) on a $\rho(B) \leq \|B\|$, on en déduit que le rayon spectral de la matrice B est strictement inférieur à 1, et donc 1 ne peut être valeur propre de B, ainsi $I - B$ est inversible.

2. Montrez que la suite $(C_k)_{k\geq 1}$ définie par
$$C_k = I + B + B^2 + \cdots + B^k$$
   est convergente de limite $(I - B)^{-1}$.

   Corrigé : On vérifie facilement que $(I - B)C_k = I - B^{k+1}$, comme $I - B$ est inversible on peut multiplier les deux membres par $(I - B)^{-1}$, soit :
$$C_k = (I - B)^{-1}(I - B^{k+1})$$
   Comme $\|B^p\| \leq \|B\|^p$ et $\|B\| < 1$, on en déduit que B^{k+1} converge vers 0 et ainsi que C_k converge vers $(I - B)^{-1}$.
   Autre approche :
   Pour montrer que la suite C_k est convergente il suffit de montrer qu'elle est de Cauchy, pour cela majorons la norme de $C_{k+p} - C_k$:
$$C_{k+p} - C_k = \sum_{k+1}^{k+p} B^i$$
$$\|C_{k+p} - C_k\| \leq \|B\|^{k+1} \sum_{0}^{p-1} \|B\|^i \leq \frac{\|B\|^{k+1}}{1 - \|B\|}$$
   Comme la norme de B est strictement inférieure à 1, cette dernière quantité peut être rendue arbitrairement petite pour des k assez grands. La suite est donc convergente, soit C sa limite, la suite $(I - B)C_k$ est convergente de limite $(I - B)C$, mais $(I - B)C_k = I - B^{k+1}$ et B^{k+1} a pour limite 0. D'où le résultat demandé : $(I - B)C = I$.

3. Montrez que
$$\|(I-B)^{-1}\| \leq \frac{1}{1-\|B\|}$$

Corrigé : La norme étant une fonction continue, $\|C_k\|$ a pour limite $\|C\|$. Mais on a la majoration :
$$\|C_k\| \leq \sum_0^k \|B\|^k = \frac{1-\|B\|^{k+1}}{1-\|B\|} \leq \frac{1}{1-\|B\|}$$
D'où le résultat demandé.

Exercice 2

Pour résoudre le système linéaire $Ax = b$, où A est une matrice carrée d'ordre n, on considère la méthode suivante, dite méthode de Richardson. Soit $r > 0$ et $x_0 \in \mathbb{R}^n$, on définit la suite $x_k \in \mathbb{R}^n$ par la formule de récurrence :
$$x_{k+1} = x_k - r(Ax_k - b)$$

1. Montrez que cette méthode est du type $Mx_{k+1} = Nx_k + b$ avec $A = M - N$, précisez M et N. Puis donnez la matrice d'itération B de la méthode.

   Corrigé : On trouve $M = r^{-1}I$ et $N = r^{-1}I - A$ et $B = I - rA$.
2. Pour A une matrice symétrique définie positive
   (a) Donnez en la démontrant une condition nécessaire et suffisante sur r pour que la méthode converge.

   Corrigé : Pour que la méthode converge il faut et il suffit que le rayon spectral de B soit inférieur strictement à 1. Les valeurs propres de B se déduisant simplement de celles de A (qui sont réelles), cela s'écrit : $-1 < 1 - r\lambda < 1$ pour toute valeur propre λ de A. Ce qui donne, puisque les valeurs propres de A sont strictement positives, la condition $0 < r < \frac{2}{\lambda}$. La condition nécessaire et suffisante de convergence de la méthode est donc $r < \frac{2}{\rho(A)}$.

   (b) Montrez que la valeur optimale de r est $\frac{2}{\lambda_1+\lambda_n}$ où λ_1 et λ_n sont respectivement la plus petite et la plus grande des valeurs propres de A.

   Corrigé : La méthode est d'autant plus rapide que le rayon spectral de la matrice d'itération est petit. Or $\rho(B) = \max_i |1-r\lambda_i|$ avec λ_i les valeurs propres de A que l'on suppose ici ordonnées de la plus petite à la plus grande (elles sont positives). Cette fonction de r s'écrit $f(r) = \max\{1 - r\lambda_1, r\lambda_n - 1\}$, son graphe est tracé sur la Figure 1 et son minimum est atteint à la valeur indiquée dans l'énoncé.

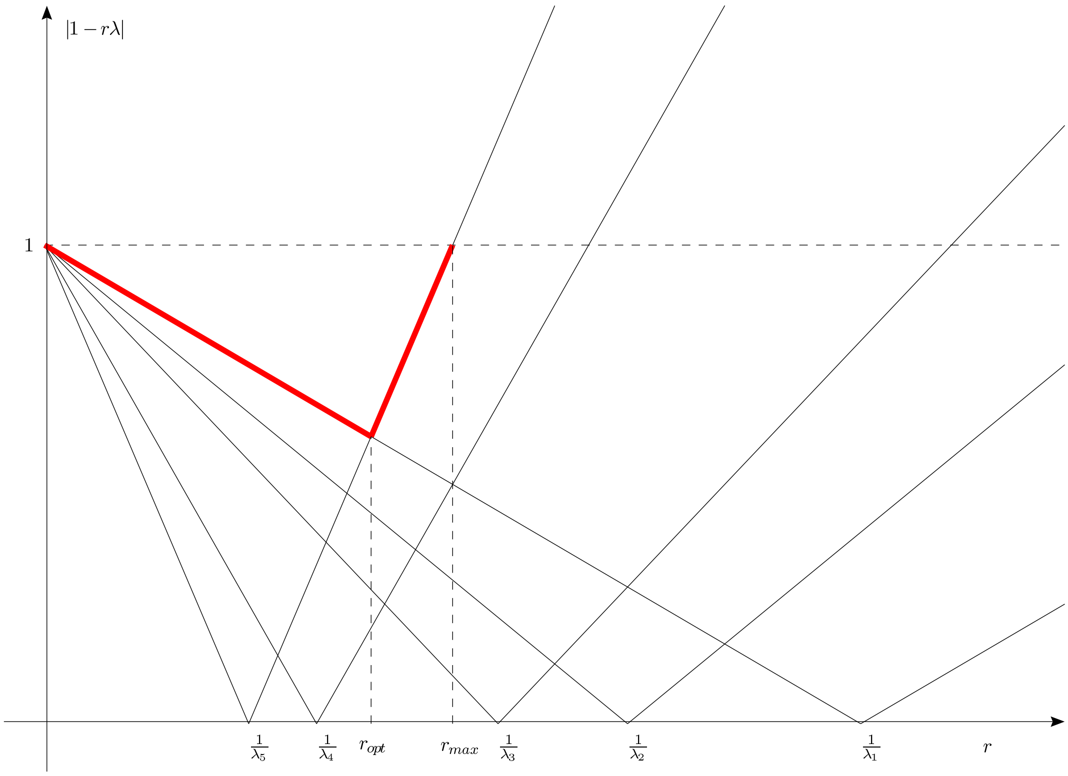

FIG. 1 – $\rho(B)$ en fonction de r

3. On suppose dans cette question que la matrice A est strictement diagonalement dominante avec des éléments diagonaux tous positifs. Montrez que si :
$$0 < r \leq \frac{1}{\max_i a_{i,i}}$$
la méthode de Richardson converge.
Indication : on pourra utiliser la norme $\|B\|_\infty = \max_i(\sum_j |a_{i,j}|)$.

Corrigé : Pour que la méthode soit convergente il suffit de vérifier que $\|B\|_\infty < 1$, soit :
$$|1 - r\, a_{i,i}| + r \sum_{j \neq i} |a_{i,j}| < 1, \quad \forall i$$
mais, compte tenu de la condition sur r on a $1 - r\, a_{i,i} \geq 0$, par ailleurs la stricte diagonale dominance entraîne $r \sum_{j\neq i} |a_{i,j}| < r a_{i,i}$, l'inégalité stricte est donc bien vérifiée pour tout i.

4. Comparez les résultats des questions 2 et 3 pour la matrice tridiagonale d'ordre n, d'éléments diagonaux $a_{i,i} = \alpha > 2$ et d'éléments extra-diagonaux $a_{i,i\pm 1} = -1$, dont les valeurs propres sont :
$$\lambda_k = \alpha - 2\cos(\frac{k\pi}{n+1}) \quad k = 1, \cdots, n$$

Corrigé : La matrice est symétrique, définie positive, en effet toutes ses valeurs propres sont strictement positives. Ses plus petite et plus grande valeurs propres sont respectivement :

$$\lambda_1 = \alpha - 2\cos(\frac{\pi}{n+1}) \quad \text{et} \quad \lambda_n = \alpha + 2\cos(\frac{\pi}{n+1})$$

La question 2 nous dit que la méthode est convergente si et seulement si

$$r < \frac{2}{\alpha + 2\cos(\frac{\pi}{n+1})}$$

et que le paramètre optimal est $r = 1/\alpha$. La matrice est également strictement diagonalement dominante à diagonale positive, la question 3 nous dit que la méthode est convergente pour $r \leq 1/\alpha$, la valeur limite indiquée de r est en fait la valeur optimale.

Exercice 3

On considère le système linéaire de $n > 2$ équations à n inconnues $(x_i \; ; \; i = 1 \ldots n)$:

$$(\alpha + 2\beta)x_i - \beta(x_{i+1} + x_{i-1}) = f_i \quad 1 \leq i \leq n \tag{1}$$

avec $\alpha \geq 0$ et $\beta > 0$, $\{f_i \; ; \; i = 1 \ldots n\}$, x_0 et x_{n+1} des données.

1. Mettre le système sous la forme $Ax = b$ avec $x \in \mathbb{R}^n$, $x = \{x_i \; ; \; i = 1 \ldots n\}$, où A, une matrice $n \times n$ symétrique, et $b \in \mathbb{R}^n$ sont à déterminer.

   Corrigé : En posant $\gamma = \alpha + 2\beta$ le système linéaire s'écrit :

$$\begin{pmatrix} \gamma & -\beta & 0 & 0 & \ldots & \ldots & \ldots & \ldots \\ -\beta & \gamma & -\beta & 0 & \ldots & \ldots & \ldots & \ldots \\ 0 & -\beta & \gamma & -\beta & \ldots & \ldots & \ldots & \ldots \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ \ldots & \ldots & 0 & -\beta & \gamma & -\beta & 0 & \ldots \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ \ldots & \ldots & \ldots & \ldots & 0 & -\beta & \gamma & -\beta \\ \ldots & \ldots & \ldots & \ldots & \ldots & 0 & -\beta & \gamma \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ \ldots \\ x_i \\ \ldots \\ x_{n-1} \\ x_n \end{pmatrix} = \begin{pmatrix} f_1 + \beta x_0 \\ f_2 \\ f_3 \\ \ldots \\ f_i \\ \ldots \\ f_{n-1} \\ f_n + \beta x_{n+1} \end{pmatrix}$$

2. En calculant :

$$(Ax, x) = \sum_{i=1}^{n} (Ax)_i \, x_i$$

   Montrer que la matrice A du système (1) est symétrique, définie positive (indication on pourra poser pour faciliter les calculs $x_0 = x_{n+1} = 0$).

Corrigé : En posant $x_0 = x_{n+1} = 0$, on calcule :

$$\begin{aligned}
(Ax, x) &= \sum_{i=1}^{i=n} (-\beta x_{i-1} + (\alpha + 2\beta)x_i - \beta x_{i+1})\, x_i \\
&= \alpha \sum_{i=1}^{i=n} x_i^2 + \beta \left(\sum_{i=1}^{i=n} (x_i - x_{i-1})\, x_i + \sum_{i=1}^{i=n} (x_i - x_{i+1})\, x_i \right) \\
&= \alpha \sum_{i=1}^{i=n} x_i^2 + \beta \left(\sum_{i=1}^{i=n} (x_i - x_{i-1})\, x_i + \sum_{j=2}^{j=n+1} (x_{j-1} - x_j)\, x_{j-1} \right) \\
&= \alpha \sum_{i=1}^{i=n} x_i^2 + \beta \left((x_1 - x_0)x_1 + \sum_{i=2}^{i=n} (x_i - x_{i-1})^2 + (x_n - x_{n+1})x_n \right) \\
&= \alpha \sum_{i=1}^{i=n} x_i^2 + \beta \left(\sum_{i=1}^{i=n+1} (x_i - x_{i-1})^2 \right)
\end{aligned}$$

On en déduit, puisque $\alpha \geq 0$ et $\beta > 0$, $(Ax, x) \geq 0$, et si $(Ax, x) = 0$ alors $\sum_{i=1}^{i=n+1} (x_i - x_{i-1})^2 = 0$ et compte tenu de $x_0 = 0$, on obtient $(x_i = 0, \forall i)$, ce qui prouve que la matrice symétrique A est définie positive.

3. Indiquez une méthode directe efficace de résolution du système (1) (on citera un théorème assurant que la méthode marche).

   Corrigé : La matrice étant symétrique définie positive, on sait que la décomposition de Gauss peut se faire sans pivotage, car tous les mineurs principaux de la matrice sont non nuls, ce qui conduit à une décomposition LU de la matrice. La forme la plus pratique est la décomposition de Cholesky $A = BB^t$ qui tire efficacement parti du caractère tridiagonal symétrique de la matrice A, la matrice triangulaire inférieure à diagonale positive B n'ayant comme termes non nuls que les termes diagonaux et sous diagonaux. La décomposition de Choleski et la résolution du système se font en $O(n)$ opérations.

4. On considère, pour $k = 1 \ldots n$, les vecteurs x^k de $\mathbb{R}^n$ définis par :

$$x^k = \left\{ x_j^k = \sin(\frac{j\, k\pi}{n+1}) \;;\; j = 1 \ldots n \right\}$$

   Montrez que ces vecteurs sont orthogonaux deux à deux (pour le produit scalaire usuel de $\mathbb{R}^n$) et qu'ils constituent un base de $\mathbb{R}^n$. Indication : on pourra poser par prolongement $x_0 = x_{n+1} = 0$ et utiliser le fait que $\sum_{p=0}^{n} \cos(pa) = \Re(\sum_{p=0}^{n} \exp(ipa))$ où $\Re(z)$ désigne la partie réelle du complexe z.

Corrigé : Calculons le produit scalaire (x^k, x^h) :

$$\begin{aligned}
(x^k, x^h) &= \sum_{j=1}^{j=n} x_j^k\, x_j^h = \sum_{j=0}^{j=n} \sin(\frac{j\,k\pi}{n+1})\, \sin(\frac{j\,h\pi}{n+1}) \\
&= \frac{1}{2}\sum_{j=0}^{j=n}\left(\cos(\frac{j\,(k-h)\pi}{n+1}) - \cos(\frac{j\,(k+h)\pi}{n+1})\right) \\
&= \frac{1}{2}\Re\left(\sum_{j=0}^{j=n}\left(\exp(i\frac{j\,(k-h)\pi}{n+1}) - \exp(i\frac{j\,(k+h)\pi}{n+1})\right)\right)
\end{aligned}$$

Si $k = h$ on obtient :

$$(x^k, x^k) = \frac{1}{2}\Re\left(n+1-\frac{1-\exp(i\frac{(n+1)\,2k\pi}{n+1})}{1-\exp(i\frac{2k\pi}{n+1})}\right) = \frac{n+1}{2}$$

Pour $k \neq h$ on obtient :

$$\begin{aligned}
(x^k, x^h) &= \frac{1}{2}\Re\left(\frac{1-\exp(i\frac{(n+1)\,(k-h)\pi}{n+1})}{1-\exp(i\frac{(k-h)\pi}{n+1})} - \frac{1-\exp(i\frac{(n+1)\,(k+h)\pi}{n+1})}{1-\exp(i\frac{(k+h)\pi}{n+1})}\right) \\
&= \frac{1}{2}\Re\left(\frac{1-\exp(i(k-h)\pi)}{1-\exp(i\frac{(k-h)\pi}{n+1})} - \frac{1-\exp(i(k+h)\pi)}{1-\exp(i\frac{(k+h)\pi}{n+1})}\right)
\end{aligned}$$

Si $k \neq h$ et $k-h$ (et $k+h$) est pair on obtient bien $(x^k, x^h) = 0$ car les numérateurs sont nuls.
Si $k \neq h$ et $k-h$ (et $k+h$) est impair, on obtient :

$$\begin{aligned}
(x^k, x^h) &= \frac{1}{2}\Re\left(\frac{2}{1-\exp(i\frac{(k-h)\pi}{n+1})} - \frac{2}{1-\exp(i\frac{(k+h)\pi}{n+1})}\right) \\
&= \frac{1-\cos(\frac{(k-h)\pi}{n+1})}{(1-\cos(\frac{(k-h)\pi}{n+1}))^2+\sin^2(\frac{(k-h)\pi}{n+1})} - \frac{1-\cos(\frac{(k+h)\pi}{n+1})}{(1-\cos(\frac{(k+h)\pi}{n+1}))^2+\sin^2(\frac{(k+h)\pi}{n+1})} \\
&= \frac{1-\cos(\frac{(k-h)\pi}{n+1})}{2-2\cos(\frac{(k-h)\pi}{n+1})} - \frac{1-\cos(\frac{(k+h)\pi}{n+1})}{2-2\cos(\frac{(k+h)\pi}{n+1})} = 0
\end{aligned}$$

Les vecteurs x^k étant orthogonaux deux à deux forment bien une base de $\mathbb{R}^n$.

5. Montrez que le vecteur x^k est vecteur propre de la matrice A pour une valeur propre λ^k que l'on déterminera. En déduire toutes les valeurs propres de A et retrouvez que A est définie positive.

Corrigé : Calculons Ax^k : pour $1 \leq i \leq n$,

$$\begin{aligned}
(Ax^k)_i &= (\alpha+2\beta)\sin(\frac{i\,k\pi}{n+1}) - \beta\left(\sin(\frac{(i-1)\,k\pi}{n+1}) + \sin(\frac{(i+1)\,k\pi}{n+1})\right) \\
&= (\alpha+2\beta)\sin(\frac{i\,k\pi}{n+1}) - 2\beta\sin(\frac{i\,k\pi}{n+1})\cos(\frac{k\pi}{n+1}) \\
&= \left(\alpha+2\beta\left(1-\cos(\frac{k\pi}{n+1})\right)\right)x_i^k \\
&= \left(\alpha+4\beta\sin^2(\frac{k\pi}{2(n+1)})\right)x_i^k
\end{aligned}$$

On trouve bien que x^k est vecteur propre de la matrice A pour la valeur propre $\lambda^k = \alpha + 4\beta\sin^2(\frac{k\pi}{2(n+1)})$. Comme $\alpha \geq 0$, $\beta > 0$ et $0 < \frac{k\pi}{2(n+1)} < \frac{\pi}{2}$, toutes ces valeurs propres sont distinctes et strictement positives ce qui re-prouve que A est définie positive.

6. Calculez le rayon spectral $\rho(J)$ de la matrice d'itération J de la méthode de Jacobi appliquée à la résolution du système (1) (indication la matrice J a la même forme que la matrice A avec des coefficients α' et β' à déterminer). En déduire que la méthode de Jacobi est convergente.

   Corrigé : La matrice d'itération de la méthode de Jacobi s'écrit formellement, avec les notations du cours, $J = D^{-1}(E+F) = I - D^{-1}A$. Comme ici D est la matrice scalaire $(\alpha+2\beta)I$ on obtient $J = I - \frac{1}{\alpha+2\beta}A$. La matrice J est donc tridiagonale symétrique, similaire à la matrice A, avec sur la diagonale $\alpha' + 2\beta' = 0$ et sur les sous et sur-diagonales $-\beta' = \frac{\beta}{\alpha+2\beta}$. Les vecteurs x^k sont aussi vecteurs propres de J pour les valeurs propres

$$\mu^k = \alpha' + 4\beta'\sin^2(\frac{k\pi}{2(n+1)}) = 2\beta'\left(-1+2\sin^2(\frac{k\pi}{2(n+1)})\right) = -2\beta'\cos(\frac{k\pi}{n+1})$$

   avec $1 \leq k \leq n$. Son rayon spectral est donc

$$\rho(J) = \lambda_1 = -\lambda_n = \frac{2\beta}{\alpha+2\beta}\cos(\frac{\pi}{n+1}) < 1$$

   la méthode de Jacobi est donc convergente, mais de moins en moins efficace quand n est grand car $\rho(J)$ se rapproche alors de 1.

Questions supplémentaires n'ayant pas fait partie du partiel

7. On pose dans la suite $A = D - E - F$ avec les conventions habituelles (D diagonale de A, E triangle inférieur, F triangle supérieur). Montrez que le polynôme caractéristique de la matrice d'itération $\mathcal{L}_1$ de la méthode de Gauss-Seidel s'écrit pour $\mu \neq 0$:

$$P_{\mathcal{L}_1}(\mu) = \frac{\mu^{\frac{n}{2}}}{\det(D-E)}\det(\mu^{-\frac{1}{2}}F + \mu^{\frac{1}{2}}E - \mu^{\frac{1}{2}}D)$$

Corrigé : La matrice d'itération de la méthode de Gauss-Seidel est $\mathcal{L}_1 = (D-E)^{-1}F$, son polynôme caractéristique est donc :

$$\begin{aligned}
P_{\mathcal{L}_1}(\mu) &= \det((D-E)^{-1}F - \mu I) \\
&= \det((D-E)^{-1}) \det(F - \mu(D-E)) \\
&= \frac{\mu^{\frac{n}{2}}}{\det(D-E)} \det(\mu^{-\frac{1}{2}}F + \mu^{\frac{1}{2}}E - \mu^{\frac{1}{2}}D)
\end{aligned}$$

8. Vérifiez que pour toute matrice tridiagonale $A' = D' - E' - F'$ et $\rho \neq 0$ on a l'égalité :

$$\rho^{-1}F' + \rho E' - D' = R(F' + E' - D')R^{-1}$$

avec R la matrice diagonale de terme général $R_{i,i} = \rho^{i-1}$ pour $i = 1 \dots n$. En déduire que le déterminant de $\rho^{-1}F' + \rho E' - D'$ est indépendant de ρ pour $\rho \neq 0$.

Corrigé : Pour C une matrice quelconque, les éléments de la matrice RCR^{-1} sont donnés par :

$$(RCR^{-1})_{i,j} = R_{i,i}C_{i,j}R^{-1}_{j,j} = \rho^{i-j}C_{i,j}$$

Si $A' = D' - E' - F'$ est tridiagonale, dans le produit $RA'R^{-1}$, la diagonale D' est inchangée, la sur-diagonale F' est divisée par ρ et la sous-diagonale E' est multipliée par ρ, ce qui justifie la formule proposée.
Comme

$$\det(\rho^{-1}F' + \rho E' - D') = \det\left(R(F' + E' - D')R^{-1}\right) = \det(F' + E' - D')$$

on en déduit que le déterminant de $\rho^{-1}F' + \rho E' - D'$ est indépendant de ρ.

9. Déduire des questions précédentes que le rayon spectral de la matrice d'itération de Gauss-Seidel est donné par :

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

En déduire que la méthode de Gauss-Seidel est convergente.

Corrigé : Des questions 7 et 8 on déduit pour le polynôme caractéristique de la méthode de Gauss-Seidel (pour $\mu \neq 0$) :

$$\begin{aligned}
P_{\mathcal{L}_1}(\mu) &= \frac{\mu^{\frac{n}{2}}}{\det(D-E)} \det(F + E - \mu^{\frac{1}{2}}D) \\
&= \frac{\mu^{\frac{n}{2}} \det D}{\det(D-E)} \det(D^{-1}(F+E) - \mu^{\frac{1}{2}}I) \\
&= \frac{\mu^{\frac{n}{2}} \det D}{\det(D-E)} P_J(\mu^{\frac{1}{2}})
\end{aligned}$$

Ceci prouve que si μ est valeur propre non nulle de la matrice $\mathcal{L}_1$, $\mu^{\frac{1}{2}}$ est valeur propre de J. On a donc bien $\rho(\mathcal{L}_1) = \rho(J)^2$ et comme on a prouvé que $\rho(J) < 1$ cela entraîne que $\rho(\mathcal{L}_1) < 1$ et la méthode de Gauss-Seidel est convergente.

10. Que peut-on dire de la vitesse de convergence des méthodes de Jacobi et Gauss-Seidel appliquées au sytème (1) quand n augmente ? Quelle méthode vaut il mieux utiliser pour résoudre (1) quand n est grand ?

    Corrigé : Comme $\rho(\mathcal{L}_1) = \rho(J)^2$, la méthode de Gauss-Seidel converge plus vite que la méthode de Jacobi. Asymptotiquement pour n grand :

    $$\rho(\mathcal{L}_1) = \left(\frac{2\beta}{\alpha + 2\beta}\right)^2 \cos^2(\frac{\pi}{n+1}) \xrightarrow[n\to\infty]{} \left(\frac{2\beta}{\alpha + 2\beta}\right)^2$$

    la méthode de Gauss-Seidel ne sera efficace que pour $\alpha > 0$.