LICENCE 3 MATHEMATIQUES

Expédition dans la semaine n°	Etape	Code UE	N° d'envoi de l'UE
50	**2L3MAT**	SMI6U01T	**2**

Nom de l'UE : Analyse numérique et optimisation

Le cours contient 3 chapitres (systèmes linéaires, systèmes non linéaires, optimisation). Pour chaque semaine, il est proposé d'étudier une partie du cours, de faire des exercices (corrigés) et, éventuellement, de réaliser un TP en python. Les TP sont conseillés mais non obligatoires. Deux devoirs sont à rendre afin de bénéficier d'une note de contrôle continu.
note finale=max(note-examen, 1/3(2 note-examen + note-contrôle-continu)).

- Contenu de l'envoi : Polycopié, chapitre 1, paragraphe 1 à 4. TP 1 et 2

- Guide du travail à effectuer

Semaine 1 :
Etudier les paragraphes 1.5.1 (méthodes itératives, définition et propriétés)
Exercice proposé (avec corrigé) : 55 (convergence de suites)

Semaine 2 :
Etudier le paragraphe 1.5.2, méthodes de Richardson et Jacobi.
Exercices proposés (avec corrigés) : 56 (sur Richardson) et 57 (sur Jacobi)

Semaine 3 :
Etudier le paragraphe 1.5.2, méthode de Gauss-Seidel
Exercices proposés (avec corrigés) : 58 (comparaison Jacobi/Gauss-Seidel)
et 59 (Jacobi avec relaxation). commencer le deuxième TP 3 (Jacobi, Gauss-Seidel)

Semaine 4 :
Etudier les paragraphes 1.5.2, méthodes SOR et SSOR, et 1.5.3 (méthodes par blocs)
Exercice proposés(avec corrigé) : 65 (Jacobi et Gauss-Seidel). Finir le TP 3 (SOR)
L'exercice 72 fait partie du premier devoir (à rendre avant fin janvier)

-Coordonnées de l'enseignant responsable de l'envoi
R. Herbin, I2M, 39 rue Joliot Curie, 13453 marseille cedex 13
email : raphaele.herbin@univ-amu.fr
Vous pouvez aussi consulter la page web: http://www.i2m.univ-amu.fr/~herbin
et me poser des questions par email

1.5 Méthodes itératives

Les méthodes directes sont très efficaces : elles donnent la solution exacte (aux erreurs d'arrondi près) du système linéaire considéré. Elles ont l'inconvénient de nécessiter une assez grande place mémoire car elles nécessitent le stockage de toute la matrice en mémoire vive. Si la matrice est pleine, c.à.d. si la plupart des coefficients de la matrice sont non nuls et qu'elle est trop grosse pour la mémoire vive de l'ordinateur dont on dispose, il ne reste plus qu'à gérer habilement le "swapping" c'est–à–dire l'échange de données entre mémoire disque et mémoire vive pour pouvoir résoudre le système.
Cependant, si le système a été obtenu à partir de la discrétisation d'équations aux dérivés partielles, il est en général "creux", c.à. d. qu'un grand nombre des coefficients de la matrice du système sont nuls ; de plus la matrice a souvent une structure "bande", *i.e.* les éléments non nuls de la matrice sont localisés sur certaines diagonales. On a vu au chapitre précédent que dans ce cas, la méthode de Choleski "conserve le profil" (voir à ce propos page 44). Si on utilise une méthode directe genre Choleski, on aura donc besoin de la place mémoire pour stocker la structure bande.
Lorsqu'on a affaire à de très gros systèmes issus par exemple de l'ingénierie (calcul des structures, mécanique des fluides, ...), où n peut être de l'ordre de plusieurs milliers, on cherche à utiliser des méthodes nécessitant le moins de mémoire possible. On a intérêt dans ce cas à utiliser des méthodes itératives. Ces méthodes ne font appel qu'à des produits matrice vecteur, et ne nécessitent donc pas le stockage du profil de la matrice mais uniquement des termes non nuls. Par exemple, si on a seulement 5 diagonales non nulles dans la matrice du système à résoudre, système de n équations et n inconnues, la place mémoire nécessaire pour un produit matrice vecteur est $6n$. Ainsi pour les gros systèmes, il est souvent avantageux d'utiliser des méthodes itératives qui ne donnent pas toujours la solution exacte du système en un nombre fini d'itérations, mais qui donnent une solution approchée à coût moindre qu'une méthode directe, car elles ne font appel qu'à des produits matrice vecteur.

Remarque 1.46 (Sur la méthode du gradient conjugué)**.**
Il existe une méthode itérative "miraculeuse" de résolution des systèmes linéaires lorsque la matrice A est symétrique définie positive : c'est la méthode du gradient conjugué, découverte dans les années 50[8]. Elle est miraculeuse en ce sens qu'elle donne la solution exacte du système $A\boldsymbol{x} = \boldsymbol{b}$ en un nombre fini d'opérations (en ce sens c'est une méthode directe) : moins de n itérations où n est l'ordre de la matrice A, bien qu'elle ne nécessite que des produits matrice vecteur ou des produits scalaires. La méthode du gradient conjugué est en fait une méthode d'optimisation pour la recherche du minimum dans $\mathbb{R}^n$ de la fonction de $\mathbb{R}^n$ dans $\mathbb{R}$ définie par : $f(\boldsymbol{x}) = \frac{1}{2}A\boldsymbol{x} \cdot \boldsymbol{x} - \boldsymbol{b} \cdot \boldsymbol{x}$. Or on peut montrer que lorsque A est symétrique définie positive, la recherche de $\boldsymbol{x}$ minimisant f dans $\mathbb{R}^n$ est équivalent à la résolution du système $A\boldsymbol{x} = \boldsymbol{b}$ (Voir paragraphe 3.2.2 page 183). Malheureusement, la méthode du gradient conjugué n'est pas si miraculeuse que cela en pratique : en effet, le nombre n est en général très grand et on ne peut en géneral pas envisager d'effectuer un tel nombre d'itérations pour résoudre le système. De plus, si on utilise la méthode du gradient conjugué brutalement, non seulement elle ne donne pas la solution en n itérations en raison de l'accumulation des erreurs d'arrondi, mais plus la taille du système croît et plus le nombre d'itérations nécessaires devient élevé. Ces problèmes ont été résolus grâce On a alors recours aux techniques dites de "préconditionnement". Nous reviendrons sur ce point au chapitre 3.
La méthode itérative du gradient à pas fixe, qui est elle aussi obtenue comme méthode de minimisation de la fonction f ci-dessus, fait l'objet de l'exercice 59 page 97 et du théorème 3.19 page 191.

1.5.1 Définition et propriétés

Soit $A \in \mathcal{M}_n(\mathbb{R})$ une matrice inversible et $\boldsymbol{b} \in \mathbb{R}^n$, on cherche toujours ici à résoudre le système linéaire (1.1) c'est–à–dire à trouver $\boldsymbol{x} \in \mathbb{R}^n$ tel que $A\boldsymbol{x} = \boldsymbol{b}$, mais de façon itérative, c.à.d. par la construction d'une suite.

8. Hestenes, Magnus R. ; Stiefel, Eduard (December 1952). "Methods of Conjugate Gradients for Solving Linear Systems". Journal of Research of the National Bureau of Standards. 49 (6).

Définition 1.47 (Méthode itérative)**.** *On appelle méthode itérative de résolution du système linéaire* (1.1) *une méthode qui construit une suite* $(\boldsymbol{x}^{(k)})_{k\in\mathbb{N}}$*, où l'itéré* $\boldsymbol{x}^{(k)}$ *est calculé à partir des itérés* $\boldsymbol{x}^{(0)}\ldots\boldsymbol{x}^{(k-1)}$*, censée converger vers* x *solution de* (1.1)*).*

Bien sûr, on souhaite que cette suite converge vers la solution $\boldsymbol{x}$ du système.

Définition 1.48 (Méthode itérative convergente)**.** *On dit qu'une méthode itérative est convergente si pour tout choix initial* $\boldsymbol{x}^{(0)}\in\mathbb{R}^n$*, on a :*

$$\boldsymbol{x}^{(k)} \longrightarrow \boldsymbol{x} \text{ quand } k\to+\infty$$

Enfin, on veut que cette suite soit simple à calculer. Une idée naturelle est de travailler avec une matrice P inversible qui soit "proche" de A, mais plus facile que A à inverser. On écrit alors $A = P-(P-A) = P-N$ (avec $N = P-A$), et on réécrit le système linéaire $A\boldsymbol{x} = \boldsymbol{b}$ sous la forme

$$P\boldsymbol{x} = (P-A)\boldsymbol{x}+\boldsymbol{b} = N\boldsymbol{x}+\boldsymbol{b}. \tag{1.83}$$

Cette forme suggère la construction de la suite $(\boldsymbol{x}^{(k)})_{k\in\mathbb{N}}$ à partir d'un choix initial $\boldsymbol{x}^{(0)}$ donné, par la formule suivante :

$$\begin{aligned} P\boldsymbol{x}^{(k+1)} &= (P-A)\boldsymbol{x}^{(k)}+\boldsymbol{b} \\ &= N\boldsymbol{x}^{(k)}+\boldsymbol{b}, \end{aligned} \tag{1.84}$$

ce qui peut également s'écrire :.

$$\boldsymbol{x}^{(k+1)} = B\boldsymbol{x}^{(k)}+\boldsymbol{c},\ \text{avec } B = P^{-1}(P-A) = \mathrm{Id}-P^{-1}A = P^{-1}N \text{ et } \boldsymbol{c} = P^{-1}\boldsymbol{b}. \tag{1.85}$$

Remarque 1.49 (Convergence vers $A^{-1}\boldsymbol{b}$)**.** *Si* $P\boldsymbol{x}^{(k+1)} = (P-A)\boldsymbol{x}^{(k)}+b$ *pour tout* $k\in\mathbb{N}$ *et* $\boldsymbol{x}^{(k)}\longrightarrow\bar{\boldsymbol{x}}$ *quand* $k\longrightarrow+\infty$ *alors* $P\bar{\boldsymbol{x}} = (P-A)\bar{\boldsymbol{x}}+b$*, et donc* $A\bar{\boldsymbol{x}} = \boldsymbol{b}$*, c.à.d.* $\bar{\boldsymbol{x}} = \boldsymbol{x}$*. En conclusion, si la suite converge, alors elle converge bien vers la solution du système linéaire.*

On introduit l'erreur d'approximation $\boldsymbol{e}^{(k)}$ à l'itération k, définie par

$$\boldsymbol{e}^{(k)} = \boldsymbol{x}^{(k)}-\boldsymbol{x},\ k\in\mathbb{N} \tag{1.86}$$

où $\boldsymbol{x}^{(k)}$ est construit par (1.85) et $\boldsymbol{x} = A^{-1}\boldsymbol{b}$. Il est facile de vérifier que $\boldsymbol{x}^{(k)}\to\boldsymbol{x} = A^{-1}\boldsymbol{b}$ lorsque $k\to+\infty$ si et seulement si $\boldsymbol{e}^{(k)}\to\boldsymbol{0}$ lorsque $k\to+\infty$

Lemme 1.50. *La suite* $(\boldsymbol{e}^{(k)})_{k\in\mathbb{N}}$ *définie par* (1.86) *est également définie par*

$$\begin{aligned} \boldsymbol{e}^{(0)} &= \boldsymbol{x}^{(0)}-\boldsymbol{x} \\ \boldsymbol{e}^{(k)} &= B^k\boldsymbol{e}^{(0)} \end{aligned} \tag{1.87}$$

DÉMONSTRATION – Comme $\boldsymbol{c} = P^{-1}\boldsymbol{b} = P^{-1}A\boldsymbol{x}$, on a

$$\boldsymbol{e}^{(k+1)} = \boldsymbol{x}^{(k+1)}-\boldsymbol{x} = B\boldsymbol{x}^{(k)}-\boldsymbol{x}+P^{-1}A\boldsymbol{x} \tag{1.88}$$

$$= B(\boldsymbol{x}^{(k)}-\boldsymbol{x}). \tag{1.89}$$

Par récurrence sur k,

$$\boldsymbol{e}^{(k)} = B^k(\boldsymbol{x}^{(0)}-\boldsymbol{x}),\quad \forall k\in\mathbb{N}. \tag{1.90}$$

■

Théorème 1.51 (Convergence de la suite)**.** *Soit A et $P \in \mathcal{M}_n(\mathbb{R})$ des matrices inversibles. Soit $\boldsymbol{x}^{(0)}$ donné et soit $(\boldsymbol{x}^{(k)})_{k\in\mathbb{N}}$ la suite définie par* (1.85)*.*

1. *La suite $(\boldsymbol{x}^{(k)})_{k\in\mathbb{N}}$ converge, quel que soit $\boldsymbol{x}^{(0)}$, vers $\boldsymbol{x} = A^{-1}\boldsymbol{b}$ si et seulement si $\rho(B) < 1$.*
2. *La suite $(\boldsymbol{x}^{(k)})_{k\in\mathbb{N}}$ converge, quel que soit $\boldsymbol{x}^{(0)}$, si et seulement si il existe une norme induite notée $\|\cdot\|$ telle que $\|B\| < 1$.*

DÉMONSTRATION –

1. On a vu que la suite $(\boldsymbol{x}^{(k)})_{k\in\mathbb{N}}$ définie par (1.85) converge vers $\boldsymbol{x} = A^{-1}\boldsymbol{b}$ si et seulement si la suite $\boldsymbol{e}^{(k)}$ définie par (1.87) tend vers $\mathbf{0}$. On en déduit par le lemme 1.34 que la suite $(\boldsymbol{x}^{(k)})_{k\in\mathbb{N}}$ converge (vers $\boldsymbol{x}$), pour tout $x^{(0)}$, si et seulement si $\rho(B) < 1$.
2. Si il existe une norme induite notée $\|\cdot\|$ telle que $\|B\| < 1$, alors en vertu du corollaire 1.34, $\rho(B) < 1$ et donc la méthode converge pour tout $x^{(0)}$.
   Réciproquement, si la méthode converge alors $\rho(B) < 1$, et donc il existe $\eta > 0$ tel que $\rho(B) = 1 - \eta$. Prenons maintenant $\varepsilon = \frac{\eta}{2}$ et appliquons la proposition 1.33 : il existe une norme induite $\|\cdot\|$ telle que $\|B\| \le \rho(B) + \varepsilon < 1$, ce qui démontre le résultat.

■

Pour trouver des méthodes itératives de résolution du système (1.1), on cherche donc une décomposition de la matrice A de la forme : $A = P - (P - A) = P - N$, où P est inversible et telle que le système $Py = d$ soit un système facile à résoudre (par exemple P diagonale ou triangulaire).

Estimation de la vitesse de convergence Soit $\boldsymbol{x}^{(0)} \in \mathbb{R}^n$ donné et soit $(\boldsymbol{x}^{(k)})_{k\in\mathbb{N}}$ la suite définie par (1.85). On a vu que, si $\rho(B) < 1$, $\boldsymbol{x}^{(k)} \to x$ quand $k \to \infty$, où $\boldsymbol{x}$ est la solution du système $A\boldsymbol{x} = b$. On montre à l'exercice 78 page 118 que (sauf cas particuliers)

$$\frac{\|\boldsymbol{x}^{(k+1)} - \boldsymbol{x}\|}{\|\boldsymbol{x}^{(k)} - \boldsymbol{x}\|} \longrightarrow \rho(B) \quad \text{lorsque } k \to +\infty,$$

indépendamment de la norme choisie sur $\mathbb{R}^n$. Le rayon spectral $\rho(B)$ de la matrice B est donc une bonne estimation de la vitesse de convergence. Pour estimer cette vitesse de convergence lorsqu'on ne connaît pas $\boldsymbol{x}$, on peut utiliser le fait (voir encore l'exercice 78 page 118) qu'on a aussi

$$\frac{\|\boldsymbol{x}^{(k+1)} - \boldsymbol{x}^{(k)}\|}{\|\boldsymbol{x}^{(k)} - \boldsymbol{x}^{(k-1)}\|} \longrightarrow \rho(B) \text{ lorsque } k \to +\infty,$$

ce qui permet d'évaluer la vitesse de convergence de la méthode par le calcul des itérés courants.

1.5.2 Quelques exemples de méthodes itératives

Une méthode simpliste

Le choix le plus simple pour le système $Px = (P - A)x + b$ soit facile à résoudre (on rappelle que c'est un objectif dans la construction d'une méthode itérative) est de prendre pour P la matrice identité (qui est très facile à inverser !). Voyons ce que cela donne sur la matrice

$$A = \begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix}. \tag{1.91}$$

On a alors $B = P - A = \begin{bmatrix} -1 & 1 \\ 1 & -1 \end{bmatrix}$. Les valeurs propres de B sont 0 et -2 et on a donc $\rho(B) = 2 > 1$. La suite $(\boldsymbol{e}^{(k)})_{k\in\mathbb{N}}$ définie par $\boldsymbol{e}^{(k)} = B^k\boldsymbol{e}^{(0)}$ n'est donc en général pas convergente. En effet, si $\boldsymbol{e}^{(0)} = a\boldsymbol{u}_1 + b\boldsymbol{u}_2$, où $\boldsymbol{u}_1 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$ est vecteur propre de B associé à la valeur propre $\lambda = -2$, on a $\boldsymbol{e}^{(k)} = (-2)^k a$ et donc $|\boldsymbol{e}^{(k)}| \to +\infty$ lorsque $k \to \infty$ dès que $a \neq 0$. Cette première idée n'est donc pas si bonne…

La méthode de Richardson

Affinons un peu et prenons maintenant $P = \beta\mathrm{Id}$, avec $\beta \in \mathbb{R}$. On a dans ce cas $P - A = \beta\mathrm{Id} - A$ et $B = \mathrm{Id} - \frac{1}{\beta}A = \mathrm{Id} - \alpha A$ avec $\alpha = \frac{1}{\beta}$. Les valeurs propres de B sont de la forme $1 - \alpha\lambda$, où λ est valeur propre de A. Pour la matrice A définie par (1.91), les valeurs propres de A sont 1 et 3, et les valeurs propres de

$$B = \begin{bmatrix} 1-2\alpha & \alpha \\ \alpha & 1-2\alpha, \end{bmatrix}$$

sont $1 - \alpha$ et $1 - 3\alpha$. Le rayon spectral de la matrice B, qui dépend de α est donc $\rho(B) = \max(|1-\alpha|, |1-3\alpha|)$, qu'on représente sur la figure ci-dessous. La méthode itérative s'écrit

$$\begin{aligned} &\boldsymbol{x}^{(0)} \in \mathbb{R}^n \text{ donné },\\ &\boldsymbol{x}^{(k+1)} = B\boldsymbol{x}^{(k)} + \boldsymbol{c}, \text{ avec } \boldsymbol{c} = \alpha\boldsymbol{b}. \end{aligned} \tag{1.92}$$

Pour que la méthode converge, il faut et il suffit que $\rho(B) < 1$, c.à.d. $3\alpha - 1 < 1$, donc $\alpha < \frac{2}{3}$. On voit que le choix $\alpha = 1$ qu'on avait fait au départ n'était pas bon. Mais on peut aussi calculer le meilleur coefficient α pour avoir la meilleure convergence possible : c'est la valeur de α qui minimise le rayon spectral ρ ; il est atteint pour $1 - \alpha = 3\alpha - 1$, ce qui donne $\alpha = \frac{1}{2}$. Cette méthode est connue sous le nom de *méthode de Richardson*[9]. Elle est souvent écrite sous la forme :

$$\begin{aligned} &\boldsymbol{x}^{(0)} \in \mathbb{R}^n \text{ donné },\\ &\boldsymbol{x}^{(k+1)} = \boldsymbol{x}^{(k)} + \alpha\boldsymbol{r}^{(k)}, \end{aligned}$$

où $\boldsymbol{r}^{(k)} = \boldsymbol{b} - A\boldsymbol{x}^{(k)}$ est le résidu. On vérifie facilement que cette forme est équivalente à la forme (1.92) qu'on vient d'étudier.

La méthode de Jacobi

Dans le cas de l'exemple de la matrice A donné par (1.91), la méthode de Richardson avec le coefficient optimal $\alpha = \frac{1}{2}$ revient à prendre comme décomposition de $A = P + A - P$ avec comme matrice $P = D$, où D est la matrice diagonale dont les coefficients sont les coefficients situés sur la diagonale de A. La *méthode de Jacobi*[10] consiste justement à prendre $P = D$, et ce même si la diagonale de A n'est pas constante.

9. Lewis Fry Richardson, (1881-1953) est un mathématicien, physicien, météorologue et psychologue qui a introduit les méthodes mathématiques pour les prévisions métérologiques. Il est également connu pour ses travaux sur les fractals. C'était un pacifiste qui a abandonné ses travaux de météorologie en raison de leur utillisation par l'armée de l'air, pour se tourner vers l'étude des raisons des guerres et de leur prévention.

10. Carl G. J. Jacobi, (1804 - 1851), mathématicien allemand. Issu d'une famille juive, il étudie à l'Université de Berlin, où il obtient son doctorat à 21 ans. Sa thèse est une discussion analytique de la théorie des fractions. En 1829, il devient professeur de mathématique à l'Université de Königsberg, et ce jusqu'en 1842. Il fait une dépression, et voyage en Italie en 1843. À son retour, il déménage à Berlin où il sera pensionnaire royal jusqu'à sa mort. Sa lettre du 2 juillet 1830 adressée à Legendre est restée célèbre pour la phrase suivante, qui a fait couler beaucoup d'encre : "M. Fourier avait l'opinion que le but principal des mathématiques était l'utilité publique et l'explication des phénomènes naturels ; mais un philosophe comme lui aurait dû savoir que le but unique de la science, c'est l'honneur de l'esprit humain, et que sous ce titre, une question de nombres vaut autant qu'une question du système du monde." C'est une question toujours en discussion. . . .

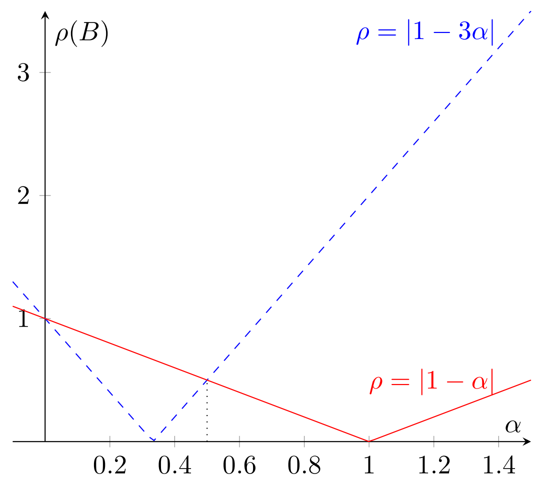

FIGURE 1.4: Rayon spectral de la matrice B de Richardson en fonction du coefficient α.

Elle n'est équivalente à la méthode de Richardson avec coefficient optimal que dans le cas où la diagonale est constante ; c'est le cas de l'exemple (1.91), et donc dans ce cas la méthode de Jacobi s'écrit

$$\boldsymbol{x}^{(0)} = \begin{bmatrix} x_1^{(0)} \\ x_2^{(0)} \end{bmatrix} \in \mathbb{R}^2 \text{ donné },$$
$$\boldsymbol{x}^{(k+1)} = \begin{bmatrix} x_1^{(k+1)} \\ x_2^{(k+1)} \end{bmatrix} = B_J \boldsymbol{x}^{(k)} + \boldsymbol{c}, \text{ avec } B_J = \begin{bmatrix} 0 & \frac{1}{2} \\ \frac{1}{2} & 0, \end{bmatrix} \text{ et } \boldsymbol{c} = \frac{1}{2}\boldsymbol{b}. \tag{1.93}$$

Dans le cas d'une matrice A générale, on décompose A sous la forme $A = D - E - F$, où D représente la diagonale de la matrice A, $(-E)$ la partie triangulaire inférieure et $(-F)$ la partie triangulaire supérieure :

$$D = \begin{bmatrix} a_{1,1} & 0 & \dots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & & 0 & a_{n,n} \end{bmatrix}, \; -E = \begin{bmatrix} 0 & 0 & \dots & 0 \\ a_{2,1} & \ddots & & \vdots \\ \vdots & \ddots & \ddots & 0 \\ a_{n,1} & \dots & a_{n-1,n} & 0 \end{bmatrix} \text{ et } -F = \begin{bmatrix} 0 & a_{1,2} & \dots & a_{1,n} \\ \vdots & \ddots & & \vdots \\ \vdots & \ddots & \ddots & a_{n,n-1} \\ 0 & \dots & 0 & -0 \end{bmatrix}. \tag{1.94}$$

La méthode de Jacobi s'écrit donc :

$$\begin{cases} \boldsymbol{x}^{(0)} \in \mathbb{R}^n \\ D\boldsymbol{x}^{(k+1)} = (E+F)\boldsymbol{x}^{(k)} + \boldsymbol{b}. \end{cases} \tag{1.95}$$

Lorsqu'on écrit la méthode de Jacobi comme sous la forme (1.85) on a $B = D^{-1}(E+F)$; on notera B_J cette matrice :

$$B_J = \begin{bmatrix} 0 & -\frac{a_{1,2}}{a_{1,1}} & \dots & -\frac{-a_{1,n}}{a_{1,1}} \\ -\frac{a_{2,1}}{a_{2,2}} & \ddots & & -\frac{-a_{2,n}}{a_{2,2}} \\ \vdots & \ddots & \ddots & \vdots \\ -\frac{a_{n,1}}{a_{n,n}} & \dots & -\frac{-a_{n-1,n}}{a_{n,n}} & 0 \end{bmatrix}.$$

La méthode de Jacobi s'écrit aussi :

$$\begin{cases} \boldsymbol{x}^{(0)} \in \mathbb{R}^n \\ a_{i,i}x_i^{(k+1)} = -\sum_{j<i} a_{i,j}x_j^{(k)} - \sum_{j>i} a_{i,j}x_j^{(k)} + b_i \;\; i = 1, \dots, n. \end{cases} \tag{1.96}$$

La méthode de Gauss-Seidel

Dans l'écriture (1.96) de la méthode de Jacobi, on pourrait remplacer les composantes $x_j^{(k)}$ dans la somme pour $j < i$ par les composantes $x_j^{(k+1)}$, puisqu'elles sont déjà calculées au moment où l'on calcule $x_i^{(k+1)}$. C'est l''idée de la méthode de Gauss-Seidel [11] qui consiste à utiliser le calcul des composantes de l'itéré $(k+1)$ dès qu'il est effectué. Par exemple, pour calculer la deuxième composante $x_2^{(k+1)}$ du vecteur $x^{(k+1)}$, on pourrait employer la "nouvelle" valeur $x_1^{(k+1)}$ qu'on vient de calculer plutôt que la valeur $x_1^{(k)}$ comme dans (1.96) ; de même, dans le calcul de $x_3^{(k+1)}$, on pourrait employer les "nouvelles" valeurs $x_1^{(k+1)}$ et $x_2^{(k+1)}$ plutôt que les valeurs $x_1^{(k)}$ et $x_2^{(k)}$. Cette idée nous suggère de remplacer dans (1.96) $x_j^{(k)}$ par $x_j^{(k+1)}$ si $j < i$. On obtient donc l'algorithme suivant :

$$\begin{cases} \boldsymbol{x}^{(0)} \in \mathbb{R}^n \\ a_{i,i}x_i^{(k+1)} = -\sum_{j<i} a_{i,j}x_j^{(k+1)} - \sum_{i<j} a_{i,j}x_j^{(k)} + b_i, \quad i = 1, \ldots, n. \end{cases} \tag{1.97}$$

La méthode de Gauss–Seidel s'écrit donc sous la forme $P\boldsymbol{x}^{(k+1)} = (P-A)\boldsymbol{x}^{(k)} + \boldsymbol{b}$, avec $P = D - E$ et $P - A = F$:

$$\begin{cases} \boldsymbol{x}_0 \in \mathbb{R}^n \\ (D-E)\boldsymbol{x}^{(k+1)} = F\boldsymbol{x}^{(k)} + b. \end{cases} \tag{1.98}$$

Si l'on écrit la méthode de Gauss–Seidel sous la forme $\boldsymbol{x}^{(k+1)} = B\boldsymbol{x}^{(k)} + \boldsymbol{c}$, on voit assez vite que $B = (D-E)^{-1}F$; on notera B_{GS} cette matrice, dite matrice de Gauss-Seidel.
Ecrivons la méthode de Gauss-Seidel dans le cas de la matrice A donnée par (1.91) : on a dans ce cas $P = D - E = \begin{bmatrix} 2 & 0 \\ -1 & 2 \end{bmatrix}$, $F = \begin{bmatrix} 0 & 1 \\ 0 & 0, \end{bmatrix}$. L'algorithme de Gauss-Seidel s'écrit donc :

$$\boldsymbol{x}^{(0)} = \begin{bmatrix} x_1^{(0)} \\ x_2^{(0)} \end{bmatrix} \in \mathbb{R}^2 \text{ donné },$$

$$\boldsymbol{x}^{(k+1)} = \begin{bmatrix} x_1^{(k+1)} \\ x_2^{(k+1)} \end{bmatrix} = B_{GS}\boldsymbol{x}^{(k)} + \boldsymbol{c}, \text{ avec } B_{GS} = \begin{bmatrix} 0 & 0 \\ 0 & \frac{1}{4} \end{bmatrix} \text{ et } \boldsymbol{c} = \begin{bmatrix} \frac{1}{2} & 0 \\ \frac{1}{4} & \frac{1}{2} \end{bmatrix} \boldsymbol{b}. \tag{1.99}$$

On a donc $\rho(B_{GS}) = \frac{1}{4}$. Sur cet exemple la méthode de Gauss-Seidel converge donc beaucoup plus vite que la méthode de Jacobi : Asymptotiquement, l'erreur est divisée par 4 à chaque itération au lieu de 2 pour la méthode de Jacobi. On peut montrer que c'est le cas pour toutes les matrices tridiagonales, comme c'est énoncé dans le théorème suivant :

Théorème 1.52 (Comparaison de Jacobi et Gauss-Seidel pour les matrices tridiagonales)**.** *On considère une matrice $A \in \mathcal{M}_n(\mathbb{R})$ tridiagonale, c.à.d. telle que $a_{i,j} = 0$ si $|i-j| > 1$; soient B_{GS} et B_{J} les matrices d'itération respectives des méthodes de Gauss-Seidel et Jacobi, alors :*

$$\rho(B_{\mathrm{GS}}) = (\rho(B_{\mathrm{J}}))^2.$$

Pour les matrices tridiagonales, la méthode de Gauss–Seidel converge (ou diverge) donc plus vite que celle de Jacobi.

La démonstration de ce résultat se fait en montrant que dans le cas tridiagonal, λ est valeur propre de la matrice d'itération de Jacobi si et seulement si λ^2 est valeur propre de la matrice d'itération de Gauss-Seidel. Elle est laissée à titre d'exercice.

11. Philipp Ludwig von Seidel (Zweibrücken, Allemagne 1821 – Munich, 13 August 1896) mathématicien allemand dont il est dit qu'il a découvert en 1847 le concept crucial de la convergence uniforme en étudiant une démonstration incorrecte de Cauchy.

Méthodes SOR et SSOR

L'idée de la méthode de sur-relaxation (SOR = Successive Over Relaxation) est d'utiliser la méthode de Gauss-Seidel pour calculer un itéré intermédiaire $\tilde{x}^{(k+1)}$ qu'on "relaxe" ensuite pour améliorer la vitesse de convergence de la méthode. On se donne $0 < \omega < 2$, et on modifie l'algorithme de Gauss–Seidel de la manière suivante :

$$\begin{cases} x_0 \in \mathbb{R}^n \\ a_{i,i}\tilde{x}_i^{(k+1)} = -\sum_{j<i} a_{i,j}x_j^{(k+1)} - \sum_{i<j} a_{i,j}x_j^{(k)} + b_i \\ x_i^{(k+1)} = \omega\tilde{x}_i^{(k+1)} + (1-\omega)x_i^{(k)}, \quad i = 1,\ldots,n. \end{cases} \tag{1.100}$$

(Pour $\omega = 1$ on retrouve la méthode de Gauss–Seidel.)
L'algorithme ci-dessus peut aussi s'écrire (en multipliant par $a_{i,i}$ la ligne 3 de l'algorithme (1.100)) :

$$\begin{cases} x^{(0)} \in \mathbb{R}^n \\ a_{i,i}x_i^{(k+1)} \quad = \quad \omega\left[-\sum_{j<i} a_{i,j}x_j^{(k+1)} - \sum_{j>i} a_{i,j}x_j^{(k)} + b_i\right] \\ \qquad\qquad\qquad +(1-\omega)a_{i,i}x_i^{(k)}. \end{cases} \tag{1.101}$$

On obtient donc

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

La matrice d'itération de l'algorithme SOR est donc

$$B_\omega = \left(\frac{D}{\omega} - E\right)^{-1} (F + \left(\frac{1-\omega}{\omega}\right) D) = P^{-1}N, \text{ avec } P = \frac{D}{\omega} - E \text{ et } N = F + \left(\frac{1-\omega}{\omega}\right) D.$$

Il est facile de vérifier que $A = P - N$.

Proposition 1.53 (Condition nécessaire de convergence de la méthode SOR)**.**
Soit $A \in \mathcal{M}_n(\mathbb{R})$ et soiet D, E et F les matrices définies par (1.94) *; on a donc $A = D - E - F$. Soit B_ω la matrice d'itération de la méthode SOR (et de la méthode de Gauss–Seidel pour $\omega = 1$) définie par :*

$$B_\omega = \left(\frac{D}{\omega} - E\right)^{-1} \left(F + \frac{1-\omega}{\omega}D\right), \quad \omega \neq 0.$$

Si $\rho(B_\omega) < 1$ alors $0 < \omega < 2$.

DÉMONSTRATION – Calculons $\det(B_\omega)$. Par définition,

$$B_\omega = P^{-1}N, \text{ avec } P = \frac{1}{\omega}D - E \text{ et } N = F + \frac{1-\omega}{\omega}D.$$

Donc $\det(B_\omega) = (\det(P))^{-1}\det(N)$. Comme P et N sont des matrices triangulaires, leurs déterminants sont les produits ceofficients diagonaux (voir la remarque 1.60 page 95). On a donc :

$$\det(B_\omega) = \frac{(\frac{1-\omega}{\omega})^n \det(D)}{(\frac{1}{\omega})^n \det(D)} = (1-\omega)^n.$$

Or le déterminant d'une matrice est aussi le produit des valeurs propres de cette matrice (comptées avec leur multiplicités algébriques), dont les valeurs absolues sont toutes inférieures au rayon spectral. On a donc : $|\det(B_\omega)| = |(1-\omega)^n| \leq (\rho(B_\omega))^n$, d'où le résultat. ■

On a un résultat de convergence de la méthode SOR (et donc également de Gauss-Seidel) dans le cas où A est symétrique définie positive, grâce au lemme suivant :

Lemme 1.54 (Condition suffisante de convergence pour la suite définie par (1.85))**.** *Soit $A \in \mathcal{M}_n(\mathbb{R})$ une matrice symétrique définie positive, et soient P et $N \in \mathcal{M}_n(\mathbb{R})$ telles que $A = P - N$ et P est inversible. Si la matrice $P^t + N$ est symétrique définie positive alors $\rho(P^{-1}N) = \rho(B) < 1$, et donc la suite définie par* (1.85) *converge.*

DÉMONSTRATION – On rappelle (voir le corollaire (1.37) page 64) que si $B \in \mathcal{M}_n(\mathbb{R})$, et si $\|\cdot\|$ est une norme induite sur $\mathcal{M}_n(\mathbb{R})$ par une norme sur $\mathbb{R}^n$, on a toujours $\rho(B) \le \|B\|$. On va donc chercher une norme sur $\mathbb{R}^n$, notée $\|\cdot\|_*$ telle que

$$\|P^{-1}N\|_* = \max\{\|P^{-1}N\boldsymbol{x}\|_*,\ \boldsymbol{x} \in \mathbb{R}^n,\ \|\boldsymbol{x}\|_* = 1\} < 1,$$

(où on désigne encore par $\|.\|_*$ la norme induite sur $\mathcal{M}_n(\mathbb{R})$) ou encore :

$$\|P^{-1}N\boldsymbol{x}\|_* < \|x\|_*,\ \ \forall \boldsymbol{x} \in \mathbb{R}^n,\ \ \boldsymbol{x} \neq 0. \tag{1.102}$$

On définit la norme $\|\cdot\|_*$ par $\|\boldsymbol{x}\|_* = \sqrt{A\boldsymbol{x}\cdot\boldsymbol{x}}$, pour tout $\boldsymbol{x} \in \mathbb{R}^n$. Comme A est symétrique définie positive, $\|\cdot\|_*$ est bien une norme sur $\mathbb{R}^n$, induite par le produit scalaire $(\boldsymbol{x}|\boldsymbol{y})_A = A\boldsymbol{x}\cdot\boldsymbol{y}$. On va montrer que la propriété (1.102) est vérifiée par cette norme. Soit $\boldsymbol{x} \in \mathbb{R}^n$, $\boldsymbol{x} \neq 0$, on a : $\|P^{-1}N\boldsymbol{x}\|_*^2 = AP^{-1}N\boldsymbol{x}\cdot P^{-1}N\boldsymbol{x}$. Or $N = P - A$, et donc : $\|P^{-1}N\boldsymbol{x}\|_*^2 = A(\mathrm{Id} - P^{-1}A)\boldsymbol{x}\cdot(Id - P^{-1}A)\boldsymbol{x}$. Soit $\boldsymbol{y} = P^{-1}A\boldsymbol{x}$; remarquons que $\boldsymbol{y} \neq 0$ car $\boldsymbol{x} \neq 0$ et $P^{-1}A$ est inversible. Exprimons $\|P^{-1}N\boldsymbol{x}\|_*^2$ à l'aide de $\boldsymbol{y}$.

$$\|P^{-1}N\boldsymbol{x}\|_*^2 = A(\boldsymbol{x}-\boldsymbol{y})\cdot(\boldsymbol{x}-\boldsymbol{y}) = A\boldsymbol{x}\cdot\boldsymbol{x} - 2A\boldsymbol{x}\cdot\boldsymbol{y} + A\boldsymbol{y}\cdot\boldsymbol{y} = \|\boldsymbol{x}\|_*^2 - 2A\boldsymbol{x}\cdot\boldsymbol{y} + A\boldsymbol{y}\cdot\boldsymbol{y}.$$

Pour que $\|P^{-1}N\boldsymbol{x}\|_*^2 < \|\boldsymbol{x}\|_*^2$ (et par suite $\rho(P^{-1}N) < 1$), il suffit donc de montrer que $-2A\boldsymbol{x}\cdot\boldsymbol{y} + A\boldsymbol{y}\cdot\boldsymbol{y} < 0$. Or, comme $P\boldsymbol{y} = A\boldsymbol{x}$, on a : $-2A\boldsymbol{x}\cdot\boldsymbol{y} + A\boldsymbol{y}\cdot\boldsymbol{y} = -2P\boldsymbol{y}\cdot\boldsymbol{y} + A\boldsymbol{y}\cdot\boldsymbol{y}$. En écrivant : $P\boldsymbol{y}\cdot\boldsymbol{y} = \boldsymbol{y}\cdot P^t\boldsymbol{y} = P^t\boldsymbol{y}\cdot\boldsymbol{y}$, on obtient donc que : $-2A\boldsymbol{x}\cdot\boldsymbol{y} + A\boldsymbol{y}\cdot\boldsymbol{y} = (-P - P^t + A)\boldsymbol{y}\cdot\boldsymbol{y}$, et comme $A = P - N$ on obtient $-2A\boldsymbol{x}\cdot\boldsymbol{y} + A\boldsymbol{y}\cdot\boldsymbol{y} = -(P^t + N)\boldsymbol{y}\cdot\boldsymbol{y}$. Comme $P^t + N$ est symétrique définie positive par hypothèse et que $\boldsymbol{y} \neq \boldsymbol{0}$, on en déduit que $-2A\boldsymbol{x}\cdot\boldsymbol{y} + A\boldsymbol{y}\cdot\boldsymbol{y} < 0$, ce qui termine la démonstration. ∎

Théorème 1.55 (CNS de convergence de la méthode SOR pour les matrices s.d.p.)**.**
Soit $A \in \mathcal{M}_n(\mathbb{R})$ une matrice symétrique définie positive, et soient D, E et F les matrices définies par (1.94) *; on a donc $A = D - E - F$. Soit B_ω la matrice d'itération de la méthode SOR (et de la méthode de Gauss–Seidel pour $\omega = 1$) définie par :*

$$B_\omega = \left(\frac{D}{\omega} - E\right)^{-1}\left(F + \frac{1-\omega}{\omega}D\right),\quad \omega \neq 0.$$

Alors :

$$\rho(B_\omega) < 1 \textit{ si et seulement si } 0 < \omega < 2.$$

En particulier, si A est une matrice symétrique définie positive, la méthode de Gauss–Seidel converge.

DÉMONSTRATION – On sait par la proposition 1.53 que si $\rho(B_\omega) < 1$ alors $0 < \omega < 2$. Supposons maintenant que A est une matrice symétrique définie positive, que $0 < \omega < 2$ et montrons que $\rho(B_\omega) < 1$. Par le lemme 1.54 page 93, il suffit pour cela de montrer que $P^t + N$ est une matrice symétrique définie positive. Or,

$$P^t = \left(\frac{D}{\omega} - E\right)^t = \frac{D}{\omega} - F,$$

$$P^t + N = \frac{D}{\omega} - F + F + \frac{1-\omega}{\omega}D = \frac{2-\omega}{\omega}D.$$

La matrice $P^t + N$ est donc bien symétrique définie positive. ∎

Remarque 1.56 (Comparaison Gauss–Seidel/Jacobi)**.** *On a vu (théorème 1.55) que si A est une matrice symétrique définie positive, la méthode de Gauss–Seidel converge. Par contre, même dans le cas où A est symétrique définie positive, il existe des cas où la méthode de Jacobi ne converge pas, voir à ce sujet l'exercice 60 page 98.*

Remarquons que le résultat de convergence des méthodes itératives donné par le théorème précédent n'est que partiel, puisqu'il ne concerne que les matrices symétriques définies positives et que les méthodes Gauss-Seidel et SOR. On a aussi un résultat de convergence de la méthode de Jacobi pour les matrices à diagonale dominante stricte, voir exercice 65 page 100, et un résultat de comparaison des méthodes pour les matrices tridiagonales par blocs, voir le théorème 1.57 donné ci-après. Dans la pratique, il faudra souvent compter sur sa bonne étoile. . .

Estimation du coefficient de relaxation optimal de SOR La question est ici d'estimer le coefficient de relaxation ω optimal dans la méthode SOR, c.à.d. le coefficient $\omega_0 \in]0,2[$ (condition nécessaire pour que la méthode SOR converge, voir théorème 1.55) tel que

$$\rho(B_{\omega_0}) \leq \rho(B_\omega), \forall \omega \in]0,2[.$$

Ce coefficient ω_0 donnera la meilleure convergence possible pour SOR. On sait le faire dans le cas assez restrictif des matrices tridiagonales (ou tridiagonales par blocs, voir paragraphe suivant). On ne fait ici qu'énoncer le résultat dont la démonstration est donnée dans le livre de Ph.Ciarlet conseillé en début de cours.

Théorème 1.57 (Coefficient optimal, matrice tridiagonale)**.** *On considère une matrice $A \in \mathcal{M}_n(\mathbb{R})$ qui admet une décomposition par blocs définie dans la définition 1.103 page 95 ; on suppose que la matrice A est tridiagonale par blocs, c.à.d. $A_{i,j} = 0$ si $|i-j| > 1$; soient $B_{\rm GS}$ et $B_{\rm J}$ les matrices d'itération respectives des méthodes de Gauss-Seidel et Jacobi. On suppose de plus que toutes les valeurs propres de la matrice d'itération J de la méthode de Jacobi sont réelles et que $\rho(B_{\rm J}) < 1$. Alors le paramètre de relaxation optimal, c.à.d. le paramètre ω_0 tel que $\rho(B_{\omega_0}) = \min\{\rho(B_\omega), \omega \in]0,2[\}$, s'exprime en fonction du rayon spectral $\rho(B_{\rm J})$ de la matrice J par la formule :*

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

et on a : $\rho(B_{\omega_0}) = \omega_0 - 1$.

La démonstration de ce résultat repose sur la comparaison des valeurs propres des matrices d'itération. On montre que λ est valeur propre de B_ω si et seulement si

$$(\lambda + \omega - 1)^2 = \lambda\omega\mu^2,$$

où μ est valeur propre de $B_{\rm J}$ (voir [Ciarlet] pour plus de détails).

Remarque 1.58 (Méthode de Jacobi relaxée)**.** *On peut aussi appliquer une procédure de relaxation avec comme méthode iérative "de base" la méthode de Jacobi, voir à ce sujet l'exercice 62 page 98). Cette méthode est toutefois beaucoup moins employée en pratique (car moins efficace) que la méthode SOR.*

Méthode SSOR En "symétrisant" le procédé de la méthode SOR, c.à.d. en effectuant les calculs SOR sur les blocs dans l'ordre 1 à n puis dans l'ordre n à 1, on obtient la méthode de sur-relaxation symétrisée (SSOR = Symmetric Successive Over Relaxation) qui s'écrit dans le formalisme de la méthode I avec

$$B_{SSOR} = \underbrace{\left(\frac{D}{\omega} - F\right)^{-1}\left(E + \frac{1-\omega}{\omega}D\right)}_{\text{calcul dans l'ordre } n\ldots 1}\underbrace{\left(\frac{D}{\omega} - E\right)^{-1}\left(F + \frac{1-\omega}{\omega}D\right)}_{\text{calcul dans l'ordre } 1\ldots n}.$$

1.5.3 Les méthodes par blocs

Décomposition par blocs d'une matrice

Dans de nombreux cas pratiques, les matrices des systèmes linéaires à résoudre ont une structure "par blocs", et on se sert alors de cette structure lors de la résolution par une méthode itérative.

Définition 1.59. *Soit $A \in \mathcal{M}_n(\mathbb{R})$ une matrice inversible ; une décomposition par blocs de A est définie par un entier $S \leq n$, des entiers $(n_i)_{i=1,\ldots,S}$ tels que $\sum_{i=1}^{S} n_i = n$, et S^2 matrices $A_{i,j} \in \mathcal{M}_{n_i,n_j}(\mathbb{R})$ (ensemble des matrices rectangulaires à n_i lignes et n_j colonnes, telles que les matrices $A_{i,i}$ soient inversibles pour $i = 1, \ldots, S$ et*

$$A = \begin{bmatrix} A_{1,1} & A_{1,2} & \ldots & & \ldots & A_{1,S} \\ A_{2,1} & \ddots & \ddots & & & \vdots \\ \vdots & \ddots & \ddots & \ddots & & \\ & & \ddots & \ddots & \ddots & \vdots \\ \vdots & & & \ddots & \ddots & A_{S-1,S} \\ A_{S,1} & \ldots & & \ldots & A_{S,S-1} & A_{S,S} \end{bmatrix} \tag{1.103}$$

Remarque 1.60.

1. *Si $S = n$ et $n_i = 1$ $\forall i \in \{1, \ldots, S\}$, chaque bloc est constitué d'un seul coefficient, et on retrouve la structure habituelle d'une matrice. Les méthodes que nous allons décrire maintenant sont alors celles que nous avons vu dans le cas de matrices sans structure particulière.*
2. *Si A est symétrique définie positive, la condition $A_{i,i}$ inversible dans la définition 1.59 est inutile car $A_{i,i}$ est nécessairement symétrique définie positive donc inversible. Pour s'en convaincre, prenons par exemple $i = 1$; soit $y \in \mathbb{R}^{n_1}$, $y \neq 0$ et $x = (y, 0 \ldots, 0)^t \in \mathbb{R}^n$. Alors $A_{1,1}y \cdot y = Ax \cdot x > 0$ donc $A_{1,1}$ est symétrique définie positive.*
3. *Si A est une matrice triangulaire par blocs, c.à.d. de la forme (1.103) avec $A_{i,j} = 0$ si $j > i$, alors*
$$\det(A) = \prod_{i=1}^{S} \det(A_{i,i}).$$
Par contre si A est décomposée en 2×2 blocs carrés (i.e. tels que $n_i = m_j$, $\forall (i,j) \in \{1, 2\}$), on a en général :
$$\det(A) \neq \det(A_{1,1})\det(A_{2,2}) - \det(A_{1,2})\det(A_{2,1}).$$

Méthode de Jacobi

On cherche une matrice P tel que le système $Px = (P - A)x + b$ soit facile à résoudre (on rappelle que c'est un objectif dans la construction d'une méthode itérative). On avait pris pour P une matrice diagonale dans la méthode de Jacobi. La méthode de Jacobi par blocs consiste à prendre pour P la matrice diagonale D formée par les blocs diagonaux de A :

$$D = \begin{bmatrix} A_{1,1} & 0 & \ldots & & \ldots & 0 \\ 0 & \ddots & \ddots & & & \vdots \\ \vdots & \ddots & \ddots & \ddots & & \\ & & \ddots & \ddots & \ddots & \vdots \\ \vdots & & & \ddots & \ddots & 0 \\ 0 & \ldots & & \ldots & 0 & A_{S,S} \end{bmatrix}.$$

Dans la matrice ci-dessus, 0 désigne un bloc nul.

On a alors $N = P - A = E + F$, où E et F sont constitués des blocs triangulaires inférieurs et supérieurs de la matrice A :

$$E = \begin{bmatrix} 0 & 0 & \dots & & \dots & 0 \\ -A_{2,1} & \ddots & \ddots & & & \vdots \\ \vdots & \ddots & \ddots & \ddots & & \\ & & \ddots & \ddots & \ddots & \vdots \\ \vdots & & & \ddots & \ddots & 0 \\ -A_{S,1} & \dots & & \dots & -A_{S,S-1} & 0 \end{bmatrix}, \; F = \begin{bmatrix} 0 & -A_{1,2} & \dots & & \dots & -A_{1,S} \\ 0 & \ddots & \ddots & & & \vdots \\ \vdots & \ddots & \ddots & \ddots & & \\ & & \ddots & \ddots & \ddots & \vdots \\ \vdots & & & \ddots & \ddots & -A_{S-1,S} \\ 0 & \dots & & \dots & 0 & 0 \end{bmatrix}.$$

On a bien $A = P - N$ et avec D, E et F définies comme ci-dessus, la méthode de Jacobi s'écrit :

$$\begin{cases} x^{(0)} \in \mathbb{R}^n \\ Dx^{(k+1)} = (E+F)x^{(k)} + b. \end{cases} \tag{1.104}$$

Lorsqu'on écrit la méthode de Jacobi comme sous la forme (1.85) on a $B = D^{-1}(E+F)$; on notera J cette matrice. En introduisant la décomposition par blocs de x, solution recherchée de (1.1), c.à.d. : $x = [x_1, \dots, x_S]^t$, où $x_i \in \mathbb{R}^{n_i}$, on peut aussi écrire la méthode de Jacobi sous la forme :

$$\begin{cases} x_0 \in \mathbb{R}^n \\ A_{i,i}x_i^{(k+1)} = -\sum_{j<i} A_{i,j}x_j^{(k)} - \sum_{j>i} A_{i,j}x_j^{(k)} + b_i \;\; i = 1, \dots, S. \end{cases} \tag{1.105}$$

Si $S = n$ et $n_i = 1 \; \forall i \in \{1, \dots, S\}$, chaque bloc est constitué d'un seul coefficient, et on obtient la méthode de Jacobi par points (aussi appelée méthode de Jacobi), qui s'écrit donc :

$$\begin{cases} x_0 \in \mathbb{R}^n \\ a_{i,i}x_i^{(k+1)} = -\sum_{j<i} a_{i,j}x_j^{(k)} - \sum_{j>i} a_{i,j}x_j^{(k)} + b_i \;\; i = 1, \dots, n. \end{cases} \tag{1.106}$$

Méthode de Gauss-Seidel

La même procédure que dans le cas $S = n$ et $n_i = 1$ donne :

$$\begin{cases} x^{(0)} \in \mathbb{R}^n \\ A_{i,i}x_i^{(k+1)} = -\sum_{j<i} A_{i,j}x_j^{(k+1)} - \sum_{i<j} A_{i,j}x_j^{(k)} + b_i, \quad i = 1, \dots, S. \end{cases} \tag{1.107}$$

La méthode de Gauss–Seidel s'écrit donc sous forme la forme $Px^{(k+1)} = (P - A)x^{(k)} + b$, $P = D - E$ et $P - A = F$:

$$\begin{cases} x_0 \in \mathbb{R}^n \\ (D-E)x^{(k+1)} = Fx^{(k)} + b. \end{cases} \tag{1.108}$$

Si l'on écrit la méthode de Gauss–Seidel sous la forme $x^{(k+1)} = Bx^{(k)} + c$, on voit assez vite que $B = (D - E)^{-1}F$; on notera B_{GS} cette matrice, dite matrice de Gauss-Seidel.

Méthodes SOR et SSOR

La méthode SOR s'écrit aussi par blocs : on se donne $0 < \omega < 2$, et on modifie l'algorithme de Gauss–Seidel de la manière suivante :

$$\begin{cases} x_0 \in \mathbb{R}^n \\ A_{i,i}\tilde{x}_i^{(k+1)} = -\sum_{j<i} A_{i,j}x_j^{(k+1)} - \sum_{i<j} A_{i,j}x_j^{(k)} + b_i \\ x_i^{(k+1)} = \omega\tilde{x}_i^{(k+1)} + (1-\omega)x_i^{(k)}, \quad i = 1, \dots, S. \end{cases} \tag{1.109}$$

(Pour $\omega = 1$ on retrouve la méthode de Gauss–Seidel.)
L'algorithme ci-dessus peut aussi s'écrire (en multipliant par $A_{i,i}$ la ligne 3 de l'algorithme (1.100)) :

$$\left\{\begin{array}{lcl} x^{(0)} \in \mathbb{R}^n & & \\ A_{i,i}x_i^{(k+1)} & = & \omega\left[-\sum_{j<i} A_{i,j}x_j^{(k+1)} - \sum_{j>i} A_{i,j}x_j^{(k)} + b_i\right] \\ & & +(1-\omega)A_{i,i}x_i^{(k)}. \end{array}\right. \tag{1.110}$$

On obtient donc

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

L'algorithme SOR s'écrit donc comme une méthode II avec

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

Il est facile de vérifier que $A = P - N$.
L'algorithme SOR s'écrit aussi comme une méthode I avec

$$B = \left(\frac{D}{\omega} - E\right)^{-1} (F + \left(\frac{1-\omega}{\omega}\right) D).$$

Remarque 1.61 (Méthode de Jacobi relaxée)**.** *On peut aussi appliquer une procédure de relaxation avec comme méthode iérative "de base" la méthode de Jacobi, voir à ce sujet l'exercice 62 page 98). Cette méthode est toutefois beaucoup moins employée en pratique (car moins efficace) que la méthode SOR.*

En "symétrisant" le procédé de la méthode SOR, c.à.d. en effectuant les calculs SOR sur les blocs dans l'ordre 1 à n puis dans l'ordre n à 1, on obtient la méthode de sur-relaxation symétrisée (SSOR = Symmetric Successive Over Relaxation) qui s'écrit dans le formalisme de la méthode I avec

$$B = \underbrace{\left(\frac{D}{\omega} - F\right)^{-1}\left(E + \frac{1-\omega}{\omega}D\right)}_{\text{calcul dans l'ordre } S\ldots 1} \underbrace{\left(\frac{D}{\omega} - E\right)^{-1}\left(F + \frac{1-\omega}{\omega}D\right)}_{\text{calcul dans l'ordre } 1\ldots S}.$$

1.5.4 Exercices (méthodes itératives)

Exercice 58 (Convergence de suites)**.** *Corrigé en page 108*

Etudier la convergence de la suite $(\boldsymbol{x}^{(k)})_{k\in\mathbb{N}} \subset \mathbb{R}^n$ définie par $\boldsymbol{x}^{(0)}$ donné, $\boldsymbol{x}^{(k)} = B\boldsymbol{x}^{(k)} + \boldsymbol{c}$ dans les cas suivants :

$$\text{(a)} \quad B = \begin{bmatrix} \frac{2}{3} & 1 \\ 0 & \frac{2}{3} \end{bmatrix},\ \boldsymbol{c} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \qquad \text{(b)} \quad B = \begin{bmatrix} \frac{2}{3} & 1 \\ 0 & 2 \end{bmatrix},\ \boldsymbol{c} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

Exercice 59 (Méthode de Richardson)**.** *Suggestions en page 107, corrigé en page 108*

Soit $A \in \mathcal{M}_n(\mathbb{R})$ une matrice symétrique définie positive, $b \in \mathbb{R}^n$ et $\alpha \in \mathbb{R}$. Pour trouver la solution de $Ax = b$, on considère la méthode itérative suivante :

— Initialisation : $x^{(0)} \in \mathbb{R}^n$,
— Iterations : $x^{(k+1)} = x^{(k)} + \alpha(b - Ax^{(k)})$.

1. Pour quelles valeurs de α (en fonction des valeurs propres de A) la méthode est-elle convergente ?
2. Calculer α_0 (en fonction des valeurs propres de A) t.q. $\rho(Id - \alpha_0 A) = \min\{\rho(Id - \alpha A), \alpha \in \mathbb{R}\}$.

Commentaire sur la méthode de Richardson : On peut la voir comme une méthode de gradient à pas fixe pour la minimisation de la fonction f définie de $\mathbb{R}^N$ dans $\mathbb{R}$ par : $\boldsymbol{x} \mapsto f(\boldsymbol{x}) = \frac{1}{2}A\boldsymbol{x} \cdot \boldsymbol{x} - \boldsymbol{b} \cdot \boldsymbol{x}$, qui sera étudiée au chapitre Optimisation. On verra en effet que grâce qu caractère symétrique définie positif de A, la fonction $\boldsymbol{f}$ admet un unique minimum, caractérisé par l'annulation du gradient de f en ce point. Or $\nabla f(\boldsymbol{x}) = A\boldsymbol{x} - \boldsymbol{b}$, et annuler le gradient consiste à résoudre le système linéaire $A\boldsymbol{x} = \boldsymbol{b}$.

Exercice 60 (Non convergence de la méthode de Jacobi)**.** *Suggestions en page 107. Corrigé en page 109.*
Soit $a \in \mathbb{R}$ et

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

Montrer que A est symétrique définie positive si et seulement si $-1/2 < a < 1$ et que la méthode de Jacobi converge si et seulement si $-1/2 < a < 1/2$.

Exercice 61 (Jacobi et Gauss–Seidel : cas des matrices tridiagonales)**.**
Soit $A \in \mathcal{M}_n(\mathbb{R})$ une matrice carrée d'ordre n inversible et tridiagonale ; on note $a_{i,j}$ le coefficient de la ligne i et la ligne j de la matrice A. On décompose en $A = D - E - F$, où D représente la diagonale de la matrice A, $(-E)$ la partie triangulaire inférieure stricte et $(-F)$ la partie triangulaire supérieure stricte.
On note B_J et B_{GS} les matrices d'itération des méthodes de Jacobi et Gauss-Seidel pour la résolution d'un système linéaire de matrice A.

1. Calculer les matrices B_J et B_{GS} pour la matrice particulère $A = \begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix}$ et calculer leurs rayons spectraux.Montrer que les méthodes convergent. Citer les résultats du cours qui s'appliquent pour cette matrice.
2. Montrer que λ est valeur propre de B_J si et seulement s'il existe un vecteur complexe $\boldsymbol{x} = (x_1, \ldots x_n) \in \mathbb{C}^n$, $\boldsymbol{x} \neq \boldsymbol{0}$, tel que
$$-a_{p,p-1}x_{p-1} - a_{p,p+1}x_{p+1} = \lambda a_{p,p}x_p, \quad p = 1, \ldots, n.$$
avec $x_0 = x_{n+1} = 0$.
3. Soit $\boldsymbol{y} = (y_1, \ldots y_n) \in \mathbb{C}^n$ défini par $y_p = \lambda^p x_p$, où λ est une valeur propre non nulle de B_J et $\boldsymbol{x} = (x_1, \ldots, x_n)$ un vecteur propre associé. On pose $y_0 = y_{n+1} = 0$. Montrer que
$$-\lambda^2 a_{p,p-1}y_{p-1} - a_{p,p+1}y_{p+1} = \lambda^2 a_{p,p}y_p, \quad p = 1, \ldots, n.$$
4. Montrer que μ est valeur propre de B_{GS} associée à un vecteur propre $\boldsymbol{z} \neq \boldsymbol{0}$ si et seulement si
$$(F - \mu(D - E))\,\boldsymbol{z} = \boldsymbol{0}.$$
5. Montrer que λ est valeur propre non nulle de B_J si et seulement si λ^2 est valeur propre de B_{GS}, et en déduire que $\rho(B_{GS}) = \rho(B_J)^2$.
6. On considère la matrice :
$$A = \begin{bmatrix} 1 & \frac{3}{4} & \frac{3}{4} \\ \frac{3}{4} & 1 & \frac{3}{4} \\ \frac{3}{4} & \frac{3}{4} & 1 \end{bmatrix}$$
Montrer que cette matrice est symétrique définie positive. Montrer que $\rho(B_{GS}) \neq \rho(B_J)^2$. Quelle est l'hypothèse mise en défaut ici ?

Exercice 62 (Méthode de Jacobi et relaxation)**.** *Suggestions en page 107, corrigé en page 110*
Soit $n \geq 1$. Soit $A = (a_{i,j})_{i,j=1,\ldots,n} \in \mathcal{M}_n(\mathbb{R})$ une matrice symétrique. On note D la partie diagonale de A, $-E$ la partie triangulaire inférieure de A et $-F$ la partie triangulaire supérieure de A, c'est-à-dire :

$$\begin{array}{l} D = (d_{i,j})_{i,j=1,\ldots,n},\ d_{i,j} = 0 \text{ si } i \neq j,\ d_{i,i} = a_{i,i}, \\ E = (e_{i,j})_{i,j=1,\ldots,n},\ e_{i,j} = 0 \text{ si } i \leq j,\ e_{i,j} = -a_{i,j} \text{ si } i > j, \\ F = (f_{i,j})_{i,j=1,\ldots,n},\ f_{i,j} = 0 \text{ si } i \geq j,\ f_{i,j} = -a_{i,j} \text{ si } i < j. \end{array}$$

Noter que $A = D - E - F$. Soit $b \in \mathbb{R}^n$. On cherche à calculer $x \in \mathbb{R}^n$ t.q. $Ax = b$. On suppose que D est définie positive (noter que A n'est pas forcément inversible). On s'intéresse ici à la méthode de Jacobi (par points), c'est–à–dire à la méthode itérative suivante :

Initialisation. $x^{(0)} \in \mathbb{R}^n$
Itérations. Pour $n \in \mathbb{N}$, $Dx^{(k+1)} = (E + F)x^{(k)} + b$.

On pose $J = D^{-1}(E + F)$.

1. Montrer, en donnant un exemple avec $n = 2$, que J peut ne pas être symétrique.
2. Montrer que J est diagonalisable dans $\mathbb{R}$ et, plus précisement, qu'il existe une base de $\mathbb{R}^n$, notée $\{f_1, \ldots, f_n\}$, et il existe $\{\mu_1, \ldots, \mu_n\} \subset \mathbb{R}$ t.q. $J\boldsymbol{f}_i = \mu_i \boldsymbol{f}_i$ pour tout $i \in \{1, \ldots, n\}$ et t.q. $D\boldsymbol{f}_i \cdot \boldsymbol{f}_j = \delta_{i,j}$ pour tout $i, j \in \{1, \ldots, n\}$.

   En ordonnant les valeurs propres de J, on a donc $\mu_1 \leq \ldots \leq \mu_n$, on conserve cette notation dans la suite.
3. Montrer que la trace de J est nulle et en déduire que $\mu_1 \leq 0$ et $\mu_n \geq 0$.

   On suppose maintenant que A et $2D - A$ sont symétriques définies positives et on pose $x = A^{-1}b$.
4. Montrer que la méthode de Jacobi (par points) converge (c'est-à-dire $x^{(k)} \to x$ quand $n \to \infty$). [Utiliser un théorème du cours.]

   On se propose maintenant d'améliorer la convergence de la méthode par une technique de relaxation. Soit $\omega > 0$, on considère la méthode suivante :
   Initialisation. $x^{(0)} \in \mathbb{R}^n$
   Itérations. Pour $n \in \mathbb{N}$, $D\tilde{x}^{(k+1)} = (E + F)x^{(k)} + b$, $x^{(k+1)} = \omega\tilde{x}^{(k+1)} + (1 - \omega)x^{(k)}$.
5. Calculer les matrices M_ω (inversible) et N_ω telles que $M_\omega x^{(k+1)} = N_\omega x^{(k)} + b$ pour tout $n \in \mathbb{N}$, en fonction de ω, D et A. On note, dans la suite $J_\omega = (M_\omega)^{-1}N_\omega$.
6. On suppose dans cette question que $(2/\omega)D - A$ est symétrique définie positive. Montrer que la méthode converge (c'est-à-dire que $x^{(k)} \to x$ quand $n \to \infty$.)
7. Montrer que $(2/\omega)D - A$ est symétrique définie positive si et seulement si $\omega < 2/(1 - \mu_1)$.
8. Calculer les valeurs propres de J_ω en fonction de celles de J. En déduire, en fonction des μ_i, la valeur "optimale" de ω, c'est-à-dire la valeur de ω minimisant le rayon spectral de J_ω.

Exercice 63 (Jacobi pour une matrice 3×3 particulière)**.**
Soit $A = \begin{bmatrix} a & 0 & \alpha \\ 0 & b & 0 \\ \alpha & 0 & c \end{bmatrix}$. On suppose que A est symétrique définie positive. Montrer que la méthode de Jacobi converge pour n'importe quel second membre et n'importe quel choix initial.

Exercice 64 (Une matrice cyclique)**.** *Suggestions en page 107*

Soit $\alpha \in \mathbb{R}$ et soit $A \in \mathcal{M}_4(\mathbb{R})$ la matrice définie par

$$A = \begin{pmatrix} \alpha & -1 & 0 & -1 \\ -1 & \alpha & -1 & 0 \\ 0 & -1 & \alpha & -1 \\ -1 & 0 & -1 & \alpha \end{pmatrix}$$

Cette matrice est dite cyclique : chaque ligne de la matrice peut être déduite de la précédente en décalant chaque coefficient d'une position.

1. Déterminer les valeurs propres de A.
2. Pour quelles valeurs de α la matrice A est-elle symétrique définie positive ? singulière ?

3. On suppose ici que $\alpha \neq 0$. Soit $b = (b_1, b_2, b_3, b_4)^t \in \mathbb{R}^4$ donné. On considère la méthode de Jacobi pour la résolution du système $Ax = b$. Soit $(x^{(k)})_{n\in\mathbb{N}}$ la suite de vecteurs donnés par l'algorithme. On note $x_i^{(k)}$ pour $i = 1, \ldots, 4$ les composantes de $x^{(k)}$. Donner l'expression de $x_i^{(k+1)}$, $i = 1, \ldots, 4$, en fonction de $x_i^{(k)}$ et $b_i^{(k)}$, $i = 1, \ldots, 4$. Pour quelles valeurs de α la méthode de Jacobi converge-t-elle ?
4. On suppose maintenant que A est symétrique définie positive. Reprendre la question précédente pour la méthode de Gauss-Seidel.

Exercice 65 (Jacobi pour les matrices à diagonale dominante stricte)**.** *Suggestions en page 107, corrigé en page 109*

Soit $A = (a_{i,j})_{i,j=1,\ldots,n} \in \mathcal{M}_n(\mathbb{R})$ une matrice à diagonale dominante stricte (c'est-à-dire $|a_{i,i}| > \sum_{j\neq i} |a_{i,j}|$ pour tout $i = 1, \ldots, n$). Montrer que A est inversible et que la méthode de Jacobi (pour calculer la solution de $Ax = b$) converge.

Exercice 66 (Jacobi pour pour un problème de diffusion)**.**

Soit $f \in C([0,1])$; on considère le système linéaire $Ax = b$ issu de la discrétisation par différences finies de pas uniforme égal à $h = \frac{1}{n+1}$ du problème suivant :

$$\begin{cases} -u''(x) + \alpha u(x) = f(x), \ x \in [0,1], \\ u(0) = 0, u(1) = 1, \end{cases} \tag{1.111}$$

où $\alpha \geq 0$.

1. Donner l'expression de A et $\boldsymbol{b}$.
2. Montrer que la méthode de Jacobi appliquée à la résolution de ce système converge (distinguer les cas $\alpha > 0$ et $\alpha = 0$).

Exercice 67 (Jacobi et diagonale dominance forte)**.**

1. Soit $A \in \mathcal{M}_n(\mathbb{R})$ une matrice symétrique définie positive.
   (a) Montrer que tous les coefficients diagonaux de A sont strictement positifs.
   (b) En déduire que la méthode de Jacobi pour la résolution du système linéaire $Ax = b$, avec $b \in \mathbb{R}^n$, est bien définie.

Soit $M \in \mathcal{M}_n(\mathbb{R})$ une matrice carrée d'ordre n, avec $n > 1$. On dit que la matrice M est irréductible si :

$$\text{pour tous ensembles d'indices } I \subset \{1, \ldots, n\}, I \neq \emptyset, \text{ et } J = \{1 \ldots, n\} \setminus I, J \neq \emptyset, \ \exists i \in I, \exists j \in J; a_{i,j} \neq 0. \tag{1.112}$$

2 (a) Montrer qu'une matrice diagonale n'est pas irréductible. En déduire qu'une matrice inversible n'est pas forcément irréductible.

2 (b) Soit $M \in \mathcal{M}_n(\mathbb{R})$ une matrice carrée d'ordre n, qui s'écrit sous la forme :

$$M = \begin{bmatrix} A & 0 \\ B & C \end{bmatrix}$$

où A et C sont des matrices carrées d'ordre p et q, avec $p + q = n$, et $B \in \mathcal{M}_{q,p}(\mathbb{R})$. La matrice M peut-elle être irréductible ?

3. Soit $A \in \mathcal{M}_n(\mathbb{R})$, $n > 1$ une matrice irréductible qui vérifie de plus la propriété suivante :

$$\forall i = 1, \ldots, n, a_{i,i} \geq \sum_{j\neq i} |a_{i,j}| \tag{1.113}$$

(On dit que la matrice est à diagonale dominante). Montrer que la méthode de Jacobi pour la résolution du système linéaire $Ax = b$, avec $b \in \mathbb{R}^n$, est bien définie.

4. Soit $A \in \mathcal{M}_n(\mathbb{R})$, $n > 1$ une matrice irréductible qui vérifie la propriété (1.113). On note B_J la matrice d'itération de la méthode de Jacobi pour la résolution du système linéaire $Ax = b$, avec $b \in \mathbb{R}^n$, et $\rho(B_\mathrm{J})$ son rayon spectral. On suppose que A vérifie la propriété supplémentaire suivante :

$$\exists i_0 ; a_{i_0,i_0} > \sum_{j \neq i_0} |a_{i,j}|. \tag{1.114}$$

(a) Montrer que $\rho(B_\mathrm{J}) \leq 1$.

(b) Montrer que si $Jx = \lambda x$ avec $|\lambda| = 1$, alors $|x_i| = \|x\|_\infty$, $\forall i = 1, \ldots, n$, où $\|x\|_\infty = \max_{k=1,\ldots,N} |x_k|$. En déduire que $x = 0$ et que la méthode de Jacobi converge.

(c) Retrouver ainsi le résultat de la question 2 de l'exercice 66.

5. En déduire que si A est une matrice qui vérifie les propriétés (1.112), (1.113) et (1.114), alors A est inversible.

6. Montrer que la matrice A suivante est symétrique définie positive et vérifie les propriétés (1.113) et (1.114).

$$A = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 2 & 1 & 1 \\ 0 & 1 & 2 & 1 \\ 0 & 1 & 1 & 2 \end{bmatrix}$$

La méthode de Jacobi converge-t-elle pour la résolution d'un système linéaire dont la matrice est A ?

Exercice 68 (Méthodes de Jacobi et Gauss Seidel pour une matrice 3×3)**.** *Corrigé détaillé en page 112*

On considère la matrice $A = \begin{bmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{bmatrix}$ et le vecteur $b = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$. Soit $x^{(0)}$ un vecteur de $\mathbb{R}^3$ donné.

1. *Méthode de Jacobi*

1.a Ecrire la méthode de Jacobi pour la résolution du système $A\boldsymbol{x} = \boldsymbol{b}$, sous la forme $\boldsymbol{x}^{(k+1)} = B_J \boldsymbol{x}^{(k)} + \boldsymbol{c}_J$.

1.b Déterminer le noyau de B_J et en donner une base.

1.c Calculer le rayon spectral de B_J et en déduire que la méthode de Jacobi converge.

1.d Calculer $x^{(1)}$ et $x^{(2)}$ pour les choix suivants de $x^{(0)}$:

$$(i)\, x^{(0)} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \qquad (ii) x^{(0)} = \begin{bmatrix} 0 \\ 1 \\ 2 \end{bmatrix}.$$

2. *Méthode de Gauss-Seidel.*

2.a Ecrire la méthode de Gauss-Seidel pour la résolution du système $A\boldsymbol{x} = b$, sous la forme $\boldsymbol{x}^{(k+1)} = B_{GS} \boldsymbol{x}^{(k)} + \boldsymbol{c}_{GS}$.

2.b Déterminer le noyau de B_{GS}.

2.c Calculer le rayon spectral de B_{GS} et en déduire que la méthode de Gauss-Seidel converge.

2.d Comparer les rayons spectraux de B_{GS} et B_J et vérifier ainsi un résultat du cours.

2.d Calculer $\boldsymbol{x}^{(1)}$ et $\boldsymbol{x}^{(2)}$ pour les choix suivants de $\boldsymbol{x}^{(0)}$:

$$(i) \;\; \boldsymbol{x}^{(0)} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \qquad (ii) \;\; \boldsymbol{x}^{(0)} = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}.$$

Exercice 69 (Convergence en un nombre fini d'itérations)**.**

1 Soit α et β des réels. Soit $u^{(0)} \in \mathbb{R}$ et $(u^{(k)})_{k\in\mathbb{N}}$ la suite réelle définie par $u^{(k+1)} = \alpha u^{(k)} + \beta$.

1.a Donner les valeurs de α et β pour lesquelles la suite $(u^{(k)})_{k\in\mathbb{N}}$ converge.

1.b On suppose que $\alpha \neq 0$, et que la suite $(u^{(k)})_{k\in\mathbb{N}}$ converge vers une limite qu'on note $\overline{u}$. Montrer que s'il existe $K \in \mathbb{N}$ tel que $u_K = \overline{u}$, alors $u^{(k)} = \overline{u}$ pour tout $k \in \mathbb{N}$.

2 Soit $n > 1$, B une matrice réelle carrée d'ordre n et $b \in \mathbb{R}^n$. Soit $u^{(0)} \in \mathbb{R}^N$ et $(u^{(k)})_{k\in\mathbb{N}}$ la suite définie par $u^{(k+1)} = Bu^{(k)} + c$.

2.a Donner les conditions sur B et c pour que la suite $(u^{(k)})_{k\in\mathbb{N}}$ converge vers une limite indépendante du choix initial $u_0 \in \mathbb{R}^N$.

2.b On suppose que la suite $(u^{(k)})_{k\in\mathbb{N}}$ converge vers une limite qu'on note $\overline{u}$. Montrer qu'on peut avoir $u^{(1)} = \overline{u}$ avec $u^{(0)} \neq \overline{u}$.

Exercice 70 (SOR et Jacobi pour une matrice tridiagonale)**.**

Soit $A = (a_{i,j})_{i,j=1,\ldots,n} \in \mathcal{M}_n(\mathbb{R})$ une matrice carrée d'ordre n tridiagonale, c'est-à-dire telle que $a_{i,j} = 0$ si $|i-j| > 1$, et dont la matrice diagonale extraite $D = \mathrm{diag}(a_{i,i})_{i=1,\ldots,n}$ est inversible.
Soit B_ω la matrice d'itération de la méthode SOR associée à A. Montrer que λ est valeur propre de J si et seulement si ν_ω est valeur propre de B_ω, où $\nu_\omega = \mu_\omega^2$ et μ_ω vérifie $\mu_\omega^2 - \lambda\omega\mu_\omega + \omega - 1 = 0$.
En déduire que

$$\rho(B_\omega) = \max_{\lambda \text{ valeur propre de } J} \{|\mu_\omega|; \mu_\omega^2 - \lambda\omega\mu_\omega + \omega - 1 = 0\}.$$

Exercice 71 (Méthode de Jacobi pour des matrices particulières)**.** *Suggestions en page 107, corrigé en page 114*

On note $\mathcal{M}_n(\mathbb{R})$ l'ensemble des matrices carrées d'ordre n à coefficients réels, et Id la matrice identité dans $\mathcal{M}_n(\mathbb{R})$. Soit $A = [a_{i,j}]_{i,j=1,\ldots,n} \in \mathcal{M}_n(\mathbb{R})$. On suppose que :

$$a_{i,j} \leq 0, \forall i,j = 1,\ldots,n,\ i \neq j, \tag{1.115}$$

$$a_{i,i} > 0, \forall i = 1,\ldots,n. \tag{1.116}$$

$$\sum_{i=1}^{n} a_{i,j} = 0, \forall j = 1,\ldots,n. \tag{1.117}$$

Soit $\lambda \in \mathbb{R}_+^*$.

1. Pour $x \in \mathbb{R}^n$, on définit
$$\|x\|_A = \sum_{i=1}^{n} a_{i,i}|x_i|.$$
Montrer que $\|\cdot\|_A$ est une norme sur $\mathbb{R}^n$.
2. Montrer que la matrice $\lambda\mathrm{Id} + A$ est inversible.
3. On considère le système linéaire suivant :
$$(\lambda\mathrm{Id} + A)u = b \tag{1.118}$$
Montrer que la méthode de Jacobi pour la recherche de la solution de ce système définit une suite $(u^{(k)})_{k\in N}$ de $\mathbb{R}^n$.
4. Montrer que la suite $(u^{(k)})_{k\in\mathbb{N}}$ vérifie :
$$\|u^{(k+1)} - u^{(k)}\|_A \leq (\frac{1}{1+\alpha})^k \|u^{(1)} - u^{(0)}\|_A,$$
où $\alpha = \frac{\lambda}{\max_{i=1,\ldots,n} a_{i,i}}$.

5. Montrer que la suite $(u^{(k)})_{k\in\mathbb{N}}$ est de Cauchy, et en déduire qu'elle converge vers la solution du système (1.118).

Exercice 72 (Une méthode itérative particulière)**.**

Soient $\alpha_1, \ldots, \alpha_n$ des réels strictement positifs, et A la matrice $n \times n$ de coefficients $a_{i,j}$ définis par :

$$\begin{cases} a_{i,i} = 2 + \alpha_i \\ a_{i,i+1} = a_{i,i-1} = -1 \\ a_{i,j} = 0 \text{ pour tous les autres cas.} \end{cases}$$

Pour $\beta > 0$ on considère la méthode itérative $Px^{(k+1)} = Nx^{(k)} + b$ avec $A = P - N$ et $N = \mathrm{diag}(\beta - \alpha_i)$ (c.à.d $\beta - \alpha_i$ pour les coefficients diagonaux, et 0 pour tous les autres).

1. Soit $\lambda \in \mathbb{C}$ une valeur propre de la matrice $P^{-1}N$; montrer qu'il existe un vecteur $x \in \mathbb{C}^n$ non nul tel que $Nx \cdot \overline{x} = \lambda Px \cdot \overline{x}$ (où $\overline{x}$ désigne le conjugué de x). En déduire que toutes les valeurs propres de la matrice $P^{-1}N$ sont réelles.

2. Montrer que le rayon spectral $\rho(P^{-1}N)$ de la matrice vérifie : $\rho(P^{-1}N) \leq \max_{i=1,n} \dfrac{|\beta - \alpha_i|}{\beta}$

3. Déduire de la question 1. que si $\beta > \dfrac{\overline{\alpha}}{2}$, où $\overline{\alpha} = \max_{i=1,n} \alpha_i$, alors $\rho(P^{-1}N) < 1$, et donc que la méthode itérative converge.

4. Trouver le paramètre β minimisant $\max_{i=1,n} \dfrac{|\beta - \alpha_i|}{\beta}$.
(On pourra d'abord montrer que pour tout $\beta > 0$, $|\beta - \alpha_i| \leq \max(\beta - \underline{\alpha}, \bar{\alpha} - \beta)$ pour tout $i = 1, \ldots, n$, avec $\underline{\alpha} = \min_{i=1,\ldots,n} \alpha_i$ et $\bar{\alpha} = \max_{i=1,\ldots,n} \alpha_i$ et en déduire que $\max_{i=1,n} |\beta - \alpha_i| = \max(\beta - \underline{\alpha}, \bar{\alpha} - \beta))$.

Exercice 73 (Méthode des directions alternées)**.**

Soit $n \in \mathbb{N}$, $n \geq 1$ et soit $A \in \mathcal{M}_n(\mathbb{R})$ une matrice carrée d'ordre n symétrique inversible et $b \in \mathbb{R}^n$. On cherche à calculer $u \in \mathbb{R}^n$, solution du système linéaire suivant :

$$Au = b, \tag{1.119}$$

On suppose connues des matrices X et $Y \in \mathcal{M}_n(\mathbb{R})$, symétriques. Soit $\alpha \in \mathbb{R}_+^*$, choisi tel que $X + \alpha\mathrm{Id}$ et $Y + \alpha\mathrm{Id}$ soient définies positives (où Id désigne la matrice identité d'ordre n) et $X + Y + \alpha\mathrm{Id} = A$.
Soit $u^{(0)} \in \mathbb{R}^n$, on propose la méthode itérative suivante pour résoudre (1.119) :

$$(X + \alpha Id)u^{(k+1/2)} = -Yu^{(k)} + b, \tag{1.120a}$$
$$(Y + \alpha Id)u^{(k+1)} = -Xu^{(k+1/2)} + b. \tag{1.120b}$$

1. Montrer que la méthode itérative (1.120) définit bien une suite $(u^{(k)})_{k\in\mathbb{N}}$ et que cette suite converge vers la solution u de (1.1) si et seulement si
$$\rho\big((Y + \alpha Id)^{-1}X(X + \alpha Id)^{-1}Y\big) < 1.$$
(On rappelle que pour toute matrice carrée d'ordre n, $\rho(M)$ désigne le rayon spectral de la matrice M.)
2. Montrer que si les matrices $(X + \frac{\alpha}{2}Id)$ et $(Y + \frac{\alpha}{2}Id)$ sont définies positives alors la méthode (1.120) converge. On pourra pour cela (mais ce n'est pas obligatoire) suivre la démarche suivante :
   (a) Montrer que
   $$\rho\big((Y + \alpha Id)^{-1}X(X + \alpha Id)^{-1}Y\big) = \rho\big(X(X + \alpha Id)^{-1}Y(Y + \alpha Id)^{-1}\big).$$
   (On pourra utiliser l'exercice 41 page 70).

(b) Montrer que

$$\rho\big(X(X+\alpha Id)^{-1}Y(Y+\alpha Id)^{-1}\big) \leq \rho\big(X(X+\alpha Id)^{-1}\big)\rho\big(Y(Y+\alpha Id)^{-1}\big).$$

(c) Montrer que $\rho\big(X(X+\alpha Id)^{-1}\big) < 1$ si et seulement si la matrice $(X + \frac{\alpha}{2}Id)$ est définie positive.

(d) Conclure.

3. Soit $f \in C([0,1]\times[0,1])$ et soit A la matrice carrée d'ordre $n = M \times M$ obtenue par discrétisation de l'équation $-\Delta u = f$ sur le carré $[0,1]\times[0,1]$ avec conditions aux limites de Dirichlet homogènes $u = 0$ sur $\partial\Omega$, par différences finies avec un pas uniforme $h = \frac{1}{M}$, et $\boldsymbol{b}$ le second membre associé.
   (a) Donner l'expression de A et $\boldsymbol{b}$.
   (b) Proposer des choix de X, Y et α pour lesquelles la méthode itérative (1.120) converge dans ce cas et qui justifient l'appellation "méthode des directions alternées" qui lui est donnée.

Exercice 74 (Systèmes linéaires, "Mauvaise relaxation")**.** Soit $A = (a_{i,j})_{i,j=1,\ldots,n} \in \mathcal{M}_n(\mathbb{R})$ une matrice s.d.p.. On note D la partie diagonale de A, $-E$ la partie triangulaire inférieure stricte de A et $-F$ la partie triangulaire supérieure stricte de A, c'est-à-dire :

$$\begin{aligned}
&D = (d_{i,j})_{i,j=1,\ldots,n},\ d_{i,j} = 0 \text{ si } i \neq j,\ d_{i,i} = a_{i,i},\\
&E = (e_{i,j})_{i,j=1,\ldots,n},\ e_{i,j} = 0 \text{ si } i \leq j,\ e_{i,j} = -a_{i,j} \text{ si } i > j,\\
&F = (f_{i,j})_{i,j=1,\ldots,n},\ f_{i,j} = 0 \text{ si } i \geq j,\ f_{i,j} = -a_{i,j} \text{ si } i < j.
\end{aligned}$$

Soit $b \in \mathbb{R}^n$. On cherche à calculer $x \in \mathbb{R}^n$ t.q. $Ax = b$. Pour $0 < \omega < 2$, on considère la méthode itérative suivante :

(a) Initialisation : On choisit $x^{(0)} \in \mathbb{R}^n$.

(b) Itérations : Pour $k \in \mathbb{N}$,
On calcule $\tilde{x}^{(k+1)}$ dans $\mathbb{R}^n$ solution de $(D-E)\tilde{x}^{(k+1)} = Fx^{(k)} + b$,
On pose $x^{(k+1)} = \omega\tilde{x}^{(k+1)} + (1-\omega)x^{(k)}$.

Enfin, on pose $M = \frac{D-E}{\omega}$ et $N = M - A$.

1. Montrer que la méthode s'écrit aussi $Mx^{(k+1)} = Nx^{(k)} + b$ et que la suite $(x^{(k)})_{k\in\mathbb{N}}$ est bien définie (c'est-à-dire que M est inversible).
2. On suppose dans cette question que $0 < \omega \leq 1$. Montrer que $M^t + N$ est s.d.p..
   N.B. : Un lemme du polycopié (lemme 1.54) donne alors que la méthode est convergente.
3. On suppose, dans cette question, que $n = 2$. On pose $A = \begin{bmatrix} \alpha & \gamma \\ \gamma & \beta \end{bmatrix}$.
   (a) Montrer que $\alpha > 0$, $\beta > 0$ et $\gamma^2 < \alpha\beta$.
   (b) Montrer que la méthode est convergente pour $0 < \omega < 2$. [Indication : Soit μ une valeur propre de $M^{-1}N$, montrer que $\mu \in]-1,1[$.]
   (c) Montrer, en donnant un exemple, que $M^t + N$ n'est pas toujours s.d.p.. [Prendre $\gamma \neq 0$.]
4. On suppose, dans cette question, que $n = 3$ et on prend $A = \begin{bmatrix} 1 & a & a \\ a & 1 & a \\ a & a & 1 \end{bmatrix}$, avec $-\frac{1}{2} < a < 1$.
   (a) Vérifier que A est bien s.d.p..
   (b) Montrer que si $a = -1/2$ et $\omega = 2$, la matrice M est toujours inversible (mais A n'est plus s.d.p.) et que $\rho(M^{-1}N) > 1$. En déduire qu'il existe $a \in]-1/2,1[$ et $\omega \in]0,2[$ tels que $\rho(M^{-1}N) > 1$.
5. On suppose $n \geq 3$. Donner un exemple de matrice A ($A \in M_n(\mathbb{R})$, A s.d.p) pour laquelle la méthode est non convergente pour certains $\omega \in]0,2[$. [Utiliser la question précédente.]

Exercice 75 (Vitesse de convergence pour la méthode de Jacobi)**.**
Soient A une matrice carrée d'ordre n, inversible, et $b \in \mathbb{R}^n$, $n > 1$. On pose $\bar{x} = A^{-1}b$. On note D la partie diagonale de A, $-E$ la partie triangulaire inférieure stricte de A et $-F$ la partie triangulaire supérieure stricte de A.
On suppose que D est inversible et on note B_J la matrice des itérations de la méthode de Jacobi, c'est-à-dire $B_J = D^{-1}(E+F)$. On rappelle que la méthode de Jacobi s'écrit
Initialisation : $x^{(0)} \in \mathbb{R}^n$,
Itérations : pour tout $k \geq 0$, $Dx^{(k+1)} = (E+F)x^{(k)} + b$.
On munit $\mathbb{R}^n$ d'une norme notée $\|\cdot\|$. On note ρ le rayon spectral de B_J.

1. On suppose que B_J est diagonalisable dans $\mathbb{R}$ (c'est-à-dire qu'il existe une base de $\mathbb{R}^n$ formée de vecteurs propres de B_J). Montrer qu'il existe $C > 0$, dépendant de A, b, $x^{(0)}$ et de la norme choisie sur $\mathbb{R}^n$, mais indépendant de k, telle que
$$\|x^{(k)} - \overline{x}\| \leq C\rho^k \text{ pour tout } k \geq 0.$$
2. On ne suppose plus que B_J est diagonalisable. Montrer que pour tout $\varepsilon > 0$, il existe $C_\varepsilon > 0$, dépendant de A, b, $x^{(0)}$, ε et de la norme choisie sur $\mathbb{R}^n$, mais indépendant de k, telle que
$$\|x^{(k)} - \overline{x}\| \leq C_\varepsilon(\rho+\varepsilon)^k \text{ pour tout } k \geq 0.$$

Dans la suite de cet exercice on prend $n = 2$ et $A = \begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix}$.

3. Dans cette question, on choisit, pour norme dans $\mathbb{R}^2$, la norme euclidienne, c'est-à-dire $\|x\|^2 = x_1^2 + x_2^2$ si $x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$. Montrer qu'il existe C (dépendant de b et $x^{(0)}$, mais non de k) telle que
$$\|x^{(k)} - \overline{x}\| = C\rho^k \text{ pour tout } k \geq 0. \tag{1.121}$$
4. Montrer qu'il existe des normes dans $\mathbb{R}^2$ pour lesquelles la conclusion de la question 3 est fausse (c'est-à-dire pour lesquelles la suite $(\|x^{(k)} - \overline{x}\|/\rho^k)_{k\in\mathbb{N}}$ n'est pas une suite constante sauf éventuellement pour des valeurs particulières de $x^{(0)}$).

Exercice 76 (Convergence d'une méthode itérative)**.**
Soit $A \in M_3(\mathbb{R})$ définie par $A = I - E - F$ avec

$$I = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} E = -\begin{bmatrix} 0 & 2 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \text{ et } F = -\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 1 & 1 & 0 \end{bmatrix}.$$

1. Montrer que A est inversible.
2. Soit $0 < \omega < 2$. Montrer que $(\frac{I}{\omega} - E)$ est inversible si et seulement si $\omega \neq \sqrt{2}/2$.

Pour $0 < \omega < 2$, $\omega \neq \sqrt{2}/2$, on considère (pour trouver la solution de $Ax = b$) la méthode itérative suivante :

$$(\frac{I}{\omega} - E)x^{(n+1)} = (F + \frac{1-\omega}{\omega}I)x^{(n)} + b.$$

On note $B_\omega = (\frac{I}{\omega} - E)^{-1}(F + \frac{1-\omega}{\omega}I)$. (De sorte que $x^{(n+1)} = B_\omega x^{(n)} + (\frac{I}{\omega} - E)^{-1}b$.)

3. Calculer, en fonction de ω, les valeurs propres de B_ω.
4. Donner l'ensemble des valeurs de ω pour lesquelles la méthode est convergente (quelquesoit $x^{(0)}$).
5. Déterminer $\omega_0 \in]0,2[$ t.q. $\rho(B_{\omega_0}) = \min\{\rho(B_\omega), \omega \in]0,2[, \omega \neq \sqrt{2}/2\}$.
   Pour cette valeur ω_0, montrer que la méthode donne la solution exacte de $Ax = b$ après un nombre fini d'itérations (quelquesoit $x^{(0)}$).

Exercice 77 (Une méthode itérative à pas optimal)**.**
Soit $n > 1$ et $A \in \mathcal{M}_n(\mathbb{R})$. On suppose que A est inversible. Pour calculer la solution du système linéaire $Ax = b$, avec $b \in \mathbb{R}^n$ donné, on se donne une matrice M appartenant à $\mathcal{M}_n(\mathbb{R})$, inversible (et plus simple à "inverser" que A) et on considère la méthode itérative suivante :
Initialisation : $x^{(0)}$ vecteur donné de $\mathbb{R}^n$, on pose $r^{(0)} = b - Ax^{(0)}$.
Itérations : pour $k \geq 0$, on choisit un réel α_k, on résout $M(x^{(k+1)} - x^{(k)}) = \alpha_k r^{(k)}$ et on pose $r^{(k+1)} = b - Ax^{(k+1)}$.
Pour conclure la description de la méthode, il reste à donner le choix de α_k, ceci est fait à la question 3.

1. Montrer que la méthode considérée peut aussi s'écrire de la manière suivante :
   Initialisation : $x^{(0)}$ vecteur donné de $\mathbb{R}^n$, $r^{(0)} = b - Ax^{(0)}$, $My^{(0)} = r^{(0)}$.
   Itérations : pour $k \geq 0$, On choisit un réel α_k,
   $x^{(k+1)} = x^{(k)} + \alpha_k y^{(k)}$, $r^{(k+1)} = r^{(k)} - \alpha_k Ay^{(k)}$, $My^{(k+1)} = r^{(k+1)}$.
2. On suppose, dans cette question, que α_k ne dépend par de k (c'est-à-dire $\alpha_k = \alpha$ pour tout $k \geq 0$). Déterminer $B \in \mathcal{M}_n(\mathbb{R})$ et $c \in \mathbb{R}^n$ tels que
   $$x^{(k+1)} = Bx^{(k)} + c \quad \text{pour tout } k \geq 0.$$
   Montrer que si la suite $(x^{(k)})_{k\in\mathbb{N}}$ converge, sa limite est solution du système linéaire $Ax = b$.

Pour la suite de l'exercice, on suppose que M est une matrice symétrique définie positive et on note $\|\cdot\|_M$ la norme sur $\mathbb{R}^n$ induite par M, c'est-à-dire $\|z\|_M^2 = Mz \cdot z$ (où $y \cdot z$ désigne le produit scalaire usuel de y et z)

3. (Choix de α_k) Soit $k \geq 0$.
   Pour $x^{(k)}$, $r^{(k)}$ et $y^{(k)}$ connus, on pose, pour $\alpha \in \mathbb{R}$, $f(\alpha) = \|M^{-1}(r^{(k)} - \alpha Ay^{(k)})\|_M^2$.
   Si $y^{(k)} \neq 0$, montrer qu'il existe un unique $\bar{\alpha} \in \mathbb{R}$ tel que $f(\bar{\alpha}) \leq f(\alpha)$ pour tout $\alpha \in \mathbb{R}$. Donner cette valeur de $\bar{\alpha}$.
   Dans la suite de l'exercice, si $y^{(k)} \neq 0$, on choisit $\alpha_k = \bar{\alpha}$ lors de l'itération k (si $y^{(k)} = 0$, on prend, par exemple, $\alpha_k = 0$).
4. Soit $k \geq 0$. Si $y^{(k)} \neq 0$, montrer que
   $$\|y^{(k)}\|_M^2 - \|y^{(k+1)}\|_M^2 = \frac{(Ay^{(k)} \cdot y^{(k)})^2}{M^{-1}Ay^{(k)} \cdot Ay^{(k)}}.$$
   En déduire que la suite $(\|y^{(k)}\|_M)_{k\in\mathbb{N}}$ converge dans $\mathbb{R}$.
5. (question indépendante des précédentes) Montrer, en donnant un exemple avec $n = 2$, que la matrice $A + A^t$ est symétrique mais pas nécessairement inversible

On suppose dans la suite que $A + A^t$ est (symétrique) définie positive.

6. Montrer qu'il existe $\beta > 0$ tel que
   $$Az \cdot z \geq \beta z \cdot z \quad \text{pour tout } z \in \mathbb{R}^n.$$
   En déduire qu'il existe $\gamma > 0$ tel que
   $$\frac{(Az \cdot z)^2}{M^{-1}Az \cdot z} \geq \gamma \|z\|_M^2.$$
7. Montrer que $y^{(k)} \to 0$ quand $k \to +\infty$ et en déduire que la suite $(x^{(k)})_{k\in\mathbb{N}}$ tend vers x, solution de $Ax = b$, quand $k \to +\infty$.

1.5.5 Exercices, suggestions

Exercice 59 page 97 (Méthode itérative du "gradient à pas fixe".)

1. Calculer le rayon spectral $\rho(B)$ de la matrice d'itération $B = \mathrm{Id} - \alpha A$. Calculer les valeurs de α pour lesquelles $\rho(B) < 1$ et en déduire que la méthode itérative du gradient à pas fixe converge si $0 < \alpha < \frac{2}{\rho(A)}$.
2. Remarquer que $\rho(Id - \alpha A) = \max(|1 - \alpha\lambda_1|, |1 - \alpha\lambda_n - 1|$, où $\lambda_1, \dots, \lambda_n$ sont les valeurs propres de A ordonnées dans le sens croissant. En traçant les graphes des valeurs prises par $|1-\alpha\lambda_1|$ et $|1-\alpha\lambda_n-1|$ en fonction de α, en déduire que le min est atteint pour $\alpha = \frac{2}{\lambda_1+\lambda_n}$.

Exercice 60 page 98 (Non convergence de la méthode de Jacobi)

Considérer d'abord le cas $a = 0$.
Si $a \neq 0$, pour chercher les valeurs de a pour lesquelles A est symétrique définie positive, calculer les valeurs propres de A en cherchant les racines du polynôme caractéristique. Introduire la variable μ telle que $a\mu = 1 - \lambda$.
Pour chercher les valeurs de a pour lesquelles la méthode de Jacobi converge, calculer les valeurs propres de la matrice d'itération J définie en cours.

Exercice 64 page 99 (Une matrice cyclique)

1. On peut trouver les trois valeurs propres (dont une double) sans calcul en remarquant que pour $\alpha = 0$ il y a 2 fois 2 lignes identiques, que la somme des colonnes est un vecteur constant et par le calcul de la trace.
2. Une matrice A est symétrique définie positive si et seulement si elle est diagonalisable et toutes ses valeurs propres sont strictement positives.
3. Appliquer le cours.

Exercice 65 page 100 (Jacobi et diagonale dominante stricte.)

Pour montrer que A est inversible, montrer que $Ax = 0$ si et seulement si $x = 0$. Pour montrer que la méthode de Jacobi converge, montrer que toutes les valeurs propres de la matrice A sont strictement inférieures à 1 en valeur absolue.

Exercice 62 page 98 (Méthode de Jacobi et relaxation.)

1. Prendre pour A une matrice (2,2) symétrique dont les éléments diagonaux sont différents l'un de l'autre.
2. Appliquer l'exercice 12 page 19 en prenant pour T l'application linéaire dont la matrice est D et pour S l'application linéaire dont la matrice est $E + F$.
4. Remarquer que $\rho(B_{\mathrm{J}}) = \max(-\mu_1, \mu_n)$, et montrer que :
si $\mu_1 \leq -1$, alors $2D - A$ n'est pas définie positive,
si $\mu_n \geq 1$, alors A n'est pas définie positive.
6. Reprendre le même raisonnement qu'à la question 2 à 4 avec les matrices M_ω et N_ω au lieu de D et $E + F$.
7. Chercher une condition qui donne que toutes les valeurs propres sont strictement positives en utilisant la base de vecteurs propres ad hoc. (Utiliser la base de $\mathbb{R}^n$, notée $\{f_1, \dots, f_n\}$, trouvée à la question 2.)
8. Remarquer que les f_i de la question 2 sont aussi vecteurs propres de J_ω et en déduire que les valeurs propres $\mu_i^{(\omega)}$ de J_ω sont de la forme $\mu_i^{(\omega)} = \omega(\mu_i - 1 - 1/\omega)$. Pour trouver le paramètre optimal ω_0, tracer les graphes des fonctions de $\mathbb{R}_+$ dans $\mathbb{R}$ définies par $\omega \mapsto |\mu_1^{(\omega)}|$ et $\omega \mapsto |\mu_n^{(\omega)}|$, et en conclure que le minimum de $\max(|\mu_1^{(\omega)}|, |\mu_n^{(\omega)}|)$ est atteint pour $\omega = \frac{2}{2-\mu_1-\mu_n}$.

Exercice 71 page 102 (Méthode de Jacobi et relaxation.)

2. Utiliser l'exercice 65 page 100

1.5.6 Exercices, corrigés

Exercice 58 page 97

(a) La valeur propre double est $\frac{2}{3}$ et donc le rayon spectral est $\frac{2}{3}$ qui est strictement inférieur à 1, donc la suite converge vers $\overline{\boldsymbol{x}} = (Id - B)^{-1}\boldsymbol{c} = \begin{bmatrix} 9 \\ 3 \end{bmatrix}$.

(b) Les valeurs propres sont $\frac{2}{3}$ et 2 et donc le rayon spectral est 2 qui est strictement supérieur à 1, donc la suite diverge (sauf si $x^{(0)} = Bx^{(0)} + c$).

Exercice 59 page 97 (Méthode itérative de Richardson)

1. On peut réécrire l'itération sous la forme : $x_{k+1} = (Id - \alpha A)x_k + \alpha b$. La matrice d'itération est donc $B = Id - \alpha A$. La méthode converge si et seulement si $\rho(B) < 1$; or les valeurs propres de B sont de la forme $1 - \alpha\lambda_i$ où λ_i est v.p. de A. On veut donc :
$$-1 < 1 - \alpha\lambda_i < 1, \ \ \forall i = 1, \dots, n.$$
c'est–à–dire $-2 < -\alpha\lambda_i$ et $-\alpha\lambda_i < 0, \forall i = 1, \dots, n$.
Comme A est symétrique définie positive, $\lambda_i > 0, \forall i = 1, \dots, n$, donc il faut $\alpha > 0$.
De plus, on a :
$$(-2 < -\alpha\lambda_i \ \ \forall i = 1, \dots, n) \iff (\alpha < \frac{2}{\lambda_i} \ \ \forall i, 1, \dots, n) \iff (\alpha < \frac{2}{\lambda_n}).$$
La méthode converge donc si et seulement si $0 < \alpha < \dfrac{2}{\rho(A)}$.
2. On a : $\rho(Id - \alpha A) = \sup_i |1 - \alpha\lambda_i| = \max(|1 - \alpha\lambda_1|, |1 - \alpha\lambda_n|)$. Le minimum de $\rho(Id - \alpha A)$ est donc obtenu pour α_0 tel que $1 - \alpha_0\lambda_1 = \alpha_0\lambda_n - 1$, c'est–à–dire (voir Figure (1.5) $\alpha_0 = \dfrac{2}{\lambda_1 + \lambda_n}$.

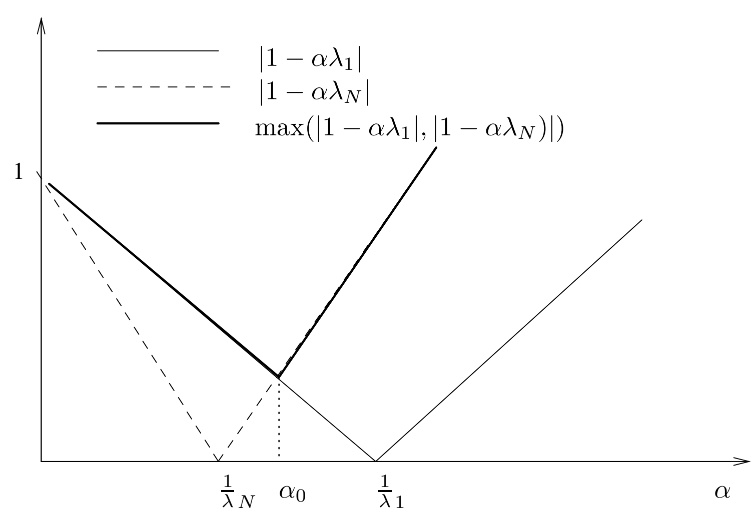

FIGURE 1.5: Graphes de $|1 - \alpha\lambda_1|$ et $|1 - \alpha\lambda_n|$ en fonction de α.

Exercice 60 page 98 (Non convergence de la méthode de Jacobi)

— Si $a = 0$, alors $A = Id$, donc A est s.d.p. et la méthode de Jacobi converge.
— Si $a \neq 0$, posons $a\mu = (1 - \lambda)$, et calculons le polynôme caractéristique de la matrice A en fonction de la variable μ.

$$P(\mu) = det \begin{vmatrix} a\mu & a & a \\ a & a\mu & a \\ a & a & a\mu \end{vmatrix} = a^3 \det \begin{vmatrix} \mu & 1 & 1 \\ 1 & \mu & 1 \\ 1 & 1 & \mu \end{vmatrix} = a^3(\mu^3 - 3\mu + 2).$$

On a donc $P(\mu) = a^3(\mu - 1)^2(\mu + 2)$. Les valeurs propres de la matrice A sont donc obtenues pour $\mu = 1$ et $\mu = 2$, c'est–à–dire : $\lambda_1 = 1 - a$ et $\lambda_2 = 1 + 2a$.
La matrice A est définie positive si $\lambda_1 > 0$ et $\lambda_2 > 0$, c'est–à–dire si $-\frac{1}{2} < a < 1$.
La méthode de Jacobi s'écrit :

$$X^{(k+1)} = D^{-1}(D - A)X^{(k)},$$

avec $D = Id$ dans le cas présent ; donc la méthode converge si et seulement si $\rho(D - A) < 1$.
Les valeurs propres de $D - A$ sont de la forme $\nu = 1 - \lambda$ où λ est valeur propre de A. Les valeurs propres de $D - A$ sont donc $\nu_1 = -a$ (valeur propre double) et $\nu_2 = 2a$.. On en conclut que la méthode de Jacobi converge si et seulement si $-1 < -a < 1$ et $-1 < 2a < 1$, *i.e.* $-\frac{1}{2} < a < \frac{1}{2}$.
La méthode de Jacobi ne converge donc que sur l'intervalle $]-\frac{1}{2}, \frac{1}{2}[$ qui est strictement inclus dans l'intervalle $]-\frac{1}{2}, 1[$ des valeurs de a pour lesquelles la matrice A est s.d.p..

Exercice 65 page 100 (Jacobi pour les matrices à diagonale dominante stricte)

Pour montrer que A est inversible, supposons qu'il existe $x \in \mathbb{R}^n$ tel que $Ax = 0$; on a donc

$$\sum_{j=1}^{n} a_{ij} x_j = 0.$$

Pour $i \in \{1, \ldots, n\}$, on a donc

$$|a_{i,i}||x_i| = |a_{i,i}x_i| = |\sum_{j;i\neq j} a_{i,j}x_j| \leq \sum_{j;i\neq j} |a_{i,j}|\|x\|_\infty, \quad \forall i = 1, \ldots, n.$$

Si $x \neq 0$, on a donc

$$|x_i| \leq \frac{\sum_{j;i\neq j} a_{i,j}x_j|}{|a_{i,i}|}\|x\|_\infty < \|x\|_\infty, \quad \forall i = 1, \ldots, n,$$

ce qui est impossible pour i tel que

$$|x_i| = \|x\|_\infty.$$

Montrons maintenant que la méthode de Jacobi converge : Si on écrit la méthode ous la forme $Px^{(k+1)} = (P - A)x^{(k)} + b$ avec , on a

$$P = D = \begin{bmatrix} a_{1,1} & & 0 \\ & \ddots & \\ 0 & & a_{n,n} \end{bmatrix}.$$

La matrice d'itération est

$$\begin{aligned} B_J = P^{-1}(P - A) = D^{-1}(E + F) &= \begin{bmatrix} a_{1,1}^{-1} & & 0 \\ & \ddots & \\ 0 & & a_{n,n}^{-1} \end{bmatrix} \begin{bmatrix} 0 & & -a_{i,j} \\ & \ddots & \\ -a_{i,j} & & 0 \end{bmatrix} \\ &= \begin{bmatrix} 0 & -\frac{a_{1,2}}{a_{1,1}} & \ldots \\ & \ddots & \\ -\frac{a_{1,1}}{a_{n,n}} & \ldots & 0 \end{bmatrix}. \end{aligned}$$

Cherchons le rayon spectral de $B_{\rm J}$: soient $x \in \mathbb{R}^n$ et $\lambda \in \mathbb{R}$ tels que $B_{\rm J}x = \lambda x$, alors

$$\sum_{j;i\neq j} -\frac{a_{i,j}}{a_{i,i}} x_j = \lambda x_i, \text{ et donc } |\lambda||x_i| \le \sum_{j;i\neq j} |a_{i,j}| \frac{\|x\|_\infty}{|a_{i,i}|}.$$

Soit i tel que $|x_i| = \|x\|_\infty$ et $x \neq 0$, on déduit de l'inégalité précédente que

$$|\lambda| \le \frac{\sum_{j;i\neq j} |a_{i,j}|}{|a_{ii}|} < 1 \text{ pour toute valeur propre } \lambda.$$

On a donc $\rho(B_{\rm J}) < 1$ ce qui prouve que la méthode de Jacobi converge.

Exercice 62 page 98 (Méthode de Jacobi et relaxation)

1. $J = D^{-1}(E+F)$ peut ne pas être symétrique, même si A est symétrique :
   En effet, prenons $A = \begin{pmatrix} 2 & -1 \\ -1 & 1 \end{pmatrix}$.
   Alors
   $$J = D^{-1}(E+F) = \begin{pmatrix} \frac{1}{2} & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 0 & \frac{1}{2} \\ 1 & 0 \end{pmatrix} \neq \begin{pmatrix} 0 & 1 \\ \frac{1}{2} & 0 \end{pmatrix}.$$
   donc J n'est pas symétrique.
2. On applique l'exercice 12 pour l'application linéaire T de matrice D, qui est, par hypothèse, définie positive (et évidemment symétrique puisque diagonale) et l'application S de matrice $E+F$, symétrique car A est symétrique.
   Il existe donc $(f_1 \ldots f_n)$ base de E et $(\mu_1 \ldots \mu_n) \in \mathbb{R}^n$ tels que
   $$Jf_i = D^{-1}(E+F)f_i = \mu_i f_i, \quad \forall i = 1, \ldots, n, \text{ et } Df_i \cdot f_j = \delta_{ij}.$$
3. La définition de J donne que tous les éléments diagonaux de J sont nuls et donc sa trace également. Or $Tr(J) = \sum_{i=1}^{n} \mu_i$. Si $\mu_i > 0 \ \forall i = 1, \ldots, n$, alors $Tr(J) > 0$, donc $\exists i_0; \ \ \mu_{i_0} \le 0$ et comme $\mu_1 \le \mu_{i_0}$, on a $\mu_1 \le 0$. Un raisonnement similaire montre que $\mu_n \ge 0$.
4. La méthode de Jacobi converge si et seulement si $\rho(J) < 1$ (théorème 1.51 page 88). Or, par la question précédente, $\rho(J) = \max(-\mu_1, \mu_n)$. Supposons que $\mu_1 \le -1$, alors $\mu_1 = -\alpha$, avec $\alpha \ge 1$. On a alors $D^{-1}(E+F)f_1 = -\alpha f_1$ ou encore $(E+F)f_1 = -\alpha D f_1$, ce qui s'écrit aussi $(D+E+F)f_1 = D(1-\alpha)f_1$ c'est–à–dire $(2D-A)f_1 = \beta D f_1$ avec $\beta \le 0$. On en déduit que $(2D-A)f_1 \cdot f_1 = \beta \le 0$, ce qui contredit le fait que $2D-A$ est définie positive. En conséquence, on a bien $\mu_1 > -1$.
   Supposons maintenant que $\mu_n = \alpha \ge 1$. On a alors $D^{-1}(E+F)f_1 = -\alpha f_n$, soit encore $(E+F)f_n = -\alpha D f_n$. On en déduit que $Af_n = (D-E-F)f_n = D(1-\alpha)f_n = D\beta f_n$ avec $\beta \le 0$. On a alors $Af_n \cdot f_n \le 0$, ce qui contredit le fait que A est définie positive.
5. Par définition, on a :$D\tilde{x}^{(k+1)} = (E+F)x^{(k)} + b$ et $x^{(k+1)} = \omega\tilde{x}^{(k+1)} + (1-\omega)x^{(k)}$. On a donc $x^{(k+1)} = \omega[D^{-1}(E+F)x^{(k)} + D^{-1}b] + (1-\omega)x^{(k)}$ c'est–à–dire $x^{(k+1)} = [Id - \omega(Id - D^{-1}(E+F))]x^{(k)} + \omega D^{-1}b$,, soit encore $\frac{1}{\omega}Dx^{(k+1)} = [\frac{1}{\omega}D - (D-(E+F))]x^{(k)} + b$. On en déduit que $M_\omega x^{(k+1)} = N_\omega x^{(k)} + b$ avec $M_\omega = \frac{1}{\omega}D$ et $N_\omega = \frac{1}{\omega}D - A$.
6. La matrice d'itération est donc maintenant $J_\omega = M_\omega^{-1}N_\omega$. En reprenant le raisonnement de la question 2 avec l'application linéaire T de matrice M_ω, qui est symétrique définie positive, et l'application S de matrice N_ω, qui est symétrique, il existe une base $(\tilde{f}_1, \ldots, \tilde{f}_n)$ de $\mathbb{R}^n$ et $(\tilde{\mu}_1, \ldots \tilde{\mu}_n) \subset \mathbb{R}$ tels que
   $$J_\omega \tilde{f}_i = {M_\omega}^{-1} N_\omega \tilde{f}_i = \omega D^{-1}\left(\frac{1}{\omega}D - A\right)\tilde{f}_i = \tilde{\mu}_i \tilde{f}_i, \quad \forall i = 1, \ldots, n,$$

$$\text{et } \frac{1}{\omega} D\tilde{f}_i \cdot \tilde{f}_j = \delta_{ij}, \quad \forall i, = 1, \dots, n, \forall j, = 1, \dots, n.$$

Supposons $\tilde{\mu}_1 \le -1$, alors $\tilde{\mu}_1 = -\alpha$, avec $\alpha \ge 1$ et $\omega D^{-1}(\frac{1}{\omega}D - A)\tilde{f}_1 = -\alpha \tilde{f}_1$, ou encore $(\frac{1}{\omega}D - A)\tilde{f}_1 = -\alpha \dfrac{1}{\omega} D\tilde{f}_1$. On a donc $(\frac{2}{\omega}D - A)\tilde{f}_1 = (1-\alpha)\dfrac{1}{\omega}D\tilde{f}_1$, ce qui entraîne $(\dfrac{2}{\omega}D - A)\tilde{f}_1 \cdot \tilde{f}_1 \le 0$. Ceci contredit l'hypothèse $\dfrac{2}{\omega}D - A$ définie positive.
De même, si $\tilde{\mu}_n \ge 1$, alors $\tilde{\mu}_n = \alpha$ avec $\alpha \ge 1$. On a alors

$$(\frac{1}{\omega}D - A)\tilde{f}_n = \alpha \frac{1}{\omega} D\tilde{f}_n,$$

et donc $A\tilde{f}_n = (1-\alpha)\frac{1}{\omega}D\tilde{f}_n$ ce qui entraîne en particulier que $A\tilde{f}_n \cdot \tilde{f}_n \le 0$; or ceci contredit l'hypothèse A définie positive.

7. On cherche une condition nécessaire et suffisante pour que

$$\left(\frac{2}{\omega}D - A\right) x \cdot x > 0, \quad \forall x \ne 0, \tag{1.122}$$

On va montrer que (1.122) est équivalent à

$$\left(\frac{2}{\omega}D - A\right) f_i \cdot f_i > 0, \quad \forall i = 1, \dots, n, \tag{1.123}$$

où les $(f_i)_{i=1,n}$ sont les vecteurs propres de $D^{-1}(E+F)$ donnés à la question 2.
La famille $(f_i)_{i=1,\dots,n}$ est une base de $\mathbb{R}^n$, et

$$\begin{aligned} \left(\frac{2}{\omega}D - A\right) f_i &= \left(\frac{2}{\omega}D - D + (E+F)\right) f_i \\ &= \left(\frac{2}{\omega} - 1\right) Df_i + \mu_i Df_i \\ &= \left(\frac{2}{\omega} - 1 + \mu_i\right) Df_i. \end{aligned} \tag{1.124}$$

On a donc en particulier $\left(\frac{2}{\omega}D - A\right) f_i \cdot f_j = 0$ si $i \ne j$, ce qui prouve que (1.122) est équivalent à (1.123).
De (1.123), on déduit aussi, grâce au fait que $Df_i \cdot f_i = 1$,

$$\left(\frac{2}{\omega}D - A\right) f_i \cdot f_i = \frac{2}{\omega} - 1 + \mu_i.$$

Une condition nécessaire et suffisante pour avoir (1.122) est donc $\frac{2}{\omega} - 1 + \mu_1 > 0$ car $\mu_1 = \inf \mu_i$, c'est–à–dire : $\frac{2}{\omega} > 1 - \mu_1$, ce qui est équivalent à : $\omega < \dfrac{2}{1-\mu_1}$.

8. La matrice d'itération J_ω s'écrit :

$$J_\omega = \left(\frac{1}{\omega}D\right)^{-1} \left(\frac{1}{\omega}D - A\right) = \omega I_\omega, \text{ avec } I_\omega = D^{-1}(\frac{1}{\omega}D - A).$$

Soit λ une valeur propre de I_ω associée à un vecteur propre u ; alors :

$$D^{-1}\left(\frac{1}{\omega}D - A\right) u = \lambda u, \ \textit{i.e.}\ \left(\frac{1}{\omega}D - A\right) u = \lambda Du.$$

On en déduit que

$$(D-A)u + \left(\frac{1}{\omega} - 1\right) Du = \lambda Du, \text{ soit encore}$$

$$D^{-1}(E+F)u = \left(1 - \frac{1}{\omega} + \lambda\right) u.$$

Ceci montre que u est un vecteur propre de J associé à la valeur propre $(1 - \frac{1}{\omega} + \lambda)$. Il existe donc $i \in \{1, \ldots, n\}$ tel que $(1 - \frac{1}{\omega} + \lambda) = \mu_i$. Les valeurs propres de I_ω sont donc les nombres $(\mu_i - 1 + \frac{1}{\omega})$ pour $i \in \{1, \ldots, n\}$. Finalement, les valeurs propres de J_ω sont donc les nombres $(\omega(\mu_i - 1) + 1)$ pour $i \in \{1, \ldots, n\}$.
On cherche maintenant à minimiser le rayon spectral

$$\rho(J_\omega) = \sup_i |\omega(\mu_i - 1) + 1)|$$

On a, pour tout i,

$$\omega(\mu_1 - 1) + 1) \leq \omega(\mu_i - 1) + 1) \leq \omega(\mu_n - 1) + 1),$$

et

$$-(\omega(\mu_n - 1) + 1) \leq -(\omega(\mu_i - 1) + 1) \leq -(\omega(\mu_1 - 1) + 1),$$

donc

$$\rho(J_\omega) = \max\left(|\omega(\mu_1 - 1) + 1)|, (|\omega(\mu_n - 1) + 1)|\right)$$

dont le minimum est atteint (voir Figure 1.6) pour

$$\omega(\mu_n - 1) + 1 = -\omega(\mu_1 - 1) - 1 \text{ c'est–à–dire } \omega = \frac{2}{2 - \mu_1 - \mu_n}.$$

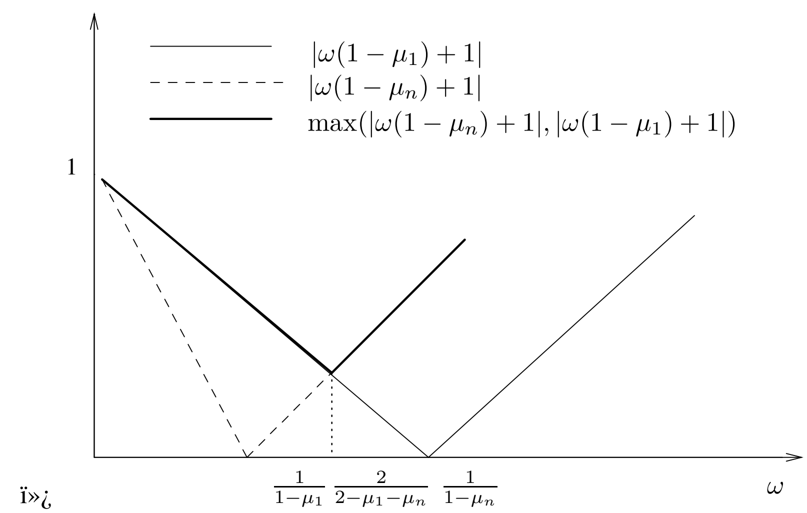

FIGURE 1.6: Détermination de la valeur de ω réalisant le minimum du rayon spectral.

Exercice 68 page 101 (Jacobi et Gauss-Seidel pour une matrice tridiagonale)

1.a La méthode de Jacobi s'écrit $D\boldsymbol{x}^{(k+1)} = (E+F)\boldsymbol{x}^{(k)} + \boldsymbol{b}$ avec

$$D = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix}, E = \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \text{ et } F = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}.$$

La méthode de Jacobi s'écrit donc $x^{(k+1)} = B_{\mathrm{J}}x^{(k)} + c_J$ avec $B_{\mathrm{J}} = D^{-1}(E+F) = \begin{bmatrix} 0 & \frac{1}{2} & 0 \\ \frac{1}{2} & 0 & \frac{1}{2} \\ 0 & \frac{1}{2} & 0 \end{bmatrix}$ et $c_J = \begin{bmatrix} \frac{1}{2} \\ 0 \\ \frac{1}{2} \end{bmatrix}$.

1.b On remarque que $x \in \mathrm{Ker}(B_{\mathrm{J}})$ si $x_2 = 0$ et $x_1 + x_3 = 0$. Donc $\mathrm{Ker} B_{\mathrm{J}} = \{t \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}, t \in \mathbb{R}\}$.

1.c Le polynôme caractéristique de B_{J} est $P_J(\lambda) = \det(B_{\mathrm{J}} - \lambda Id) = (-\lambda(-\lambda^2 + \frac{1}{2})$ et donc $\rho(B_{\mathrm{J}}) = \frac{\sqrt{2}}{2} < 1$. On en déduit que la méthode de Jacobi converge.

1.d Choix (i) : $x^{(1)} = \begin{bmatrix} \frac{1}{2} \\ 0 \\ \frac{1}{2} \end{bmatrix}, x^{(2)} = \begin{bmatrix} \frac{1}{2} \\ \frac{1}{2} \\ \frac{1}{2} \end{bmatrix}$.

Choix (ii) : $x^{(1)} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, x^{(2)} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$.

2.a La méthode de Gauss-Seidel s'écrit $(D-E)x^{(k+1)} = Fx^{(k)} + b$, où D, E et F ont été définies à la question 1.a. La méthode s'écrit donc $x^{(k+1)} = B_{GS}x^{(k)} + c_{GS}$ avec $B_{GS} = (D-E)^{-1}F$ et $c_{GS} = (D-E)^{-1}b$. Calculons $(D-E)^{-1}F$ et $(D-E)^{-1}b$ par échelonnement.

$$\left[\begin{array}{ccccccc} 2 & 0 & 0 & 0 & 1 & 0 & 1 \\ -1 & 2 & 0 & 0 & 0 & 1 & 0 \\ 0 & -1 & 2 & 0 & 0 & 0 & 1 \end{array}\right] \rightsquigarrow \left[\begin{array}{ccccccc} 2 & 0 & 0 & 0 & 1 & 0 & 1 \\ 0 & 2 & 0 & 0 & \frac{1}{2} & 1 & \frac{1}{2} \\ 0 & -1 & 2 & 0 & 0 & 0 & 1 \end{array}\right] \rightsquigarrow \left[\begin{array}{ccccccc} 2 & 0 & 0 & 0 & 1 & 0 & 1 \\ 0 & 2 & 0 & 0 & \frac{1}{2} & 1 & \frac{1}{2} \\ 0 & 0 & 2 & 0 & \frac{1}{4} & \frac{1}{2} & \frac{5}{4} \end{array}\right]$$

On a donc

$$B_{GS} = \begin{bmatrix} 0 & \frac{1}{2} & 0 \\ 0 & \frac{1}{4} & \frac{1}{2} \\ 0 & \frac{1}{8} & \frac{1}{4} \end{bmatrix} \text{ et } c_{GS} = \begin{bmatrix} \frac{1}{2} \\ \frac{1}{4} \\ \frac{5}{8} \end{bmatrix}.$$

2.b Il est facile de voir que $x \in \mathrm{Ker}(B_{GS})$ si et seulement si $x_2 = x_3 = 0$. Donc $\mathrm{Ker} B_{GS} = \{t \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, t \in \mathbb{R}\}$.

2.c Le polynôme caractéristique de B_{GS} est $P_{GS}(\lambda) = \det(B_{GS} - \lambda Id)$. On a donc

$$P_{GS}(\lambda) = \begin{vmatrix} -\lambda & \frac{1}{2} & 0 \\ 0 & \frac{1}{4} - \lambda & \frac{1}{2} \\ 0 & \frac{1}{8} & \frac{1}{4} - \lambda \end{vmatrix} = -\lambda \left((\frac{1}{4} - \lambda)^2 - \frac{1}{16} \right) = \lambda^2 (\frac{1}{2} - \lambda)$$

et donc $\rho(B_{GS}) = \frac{1}{2} < 1$. On en déduit que la méthode de Gauss-Seidel converge.

2.d On a bien $\rho(B_{\mathrm{GS}}) = \frac{1}{2} = \rho(B_{\mathrm{J}})^2$, ce qui est conforme au théorème 1.36 du cours.

2.e Choix (i) : $x^{(1)} = \begin{bmatrix} \frac{1}{2} \\ \frac{1}{4} \\ \frac{5}{8} \end{bmatrix}, x^{(2)} = \begin{bmatrix} \frac{5}{8} \\ \frac{5}{8} \\ \frac{13}{16}s \end{bmatrix}$.

Choix (ii) : $x^{(1)} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, x^{(2)} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$.

Exercice 71 page 102 (Méthode de Jacobi pour des matrices particulières)

1. Soit $x \in \mathbb{R}^n$, supposons que
$$\|x\|_A = \sum_{i=1}^{n} a_{i,i}|x_i| = 0.$$
Comme $a_{i,i} > 0$, $\forall i = 1, \ldots, n$, on en déduit que $x_i = 0$, $\forall i = 1, \ldots, n$. D'autre part, il est immédiat de voir que $\|x + y\|_A \le \|x\| + \|y\|_A$ pour tout $(x, y) \in \mathbb{R}^n \times \mathbb{R}^n$ et que $\|\lambda x\|_A = |\lambda|\|x\|_A$ pour tout $(x, \lambda) \in \mathbb{R}^n \times \mathbb{R}$. On en déduit que $\|\cdot\|_A$ est une norme sur $\mathbb{R}^n$.
2. Posons $\tilde{A} = \lambda \mathrm{Id} + A$ et notons $\tilde{a}_{i,j}$ ses coefficients. Comme $\lambda \in \mathbb{R}_+^*$,et grâce aux hypothèses (1.115)–(1.117), ceux-ci vérifient :
$$\begin{aligned} \tilde{a}_{i,j} &\le 0, \forall i, j = 1, \ldots, n,\ i \ne j, && (1.125)\\ \tilde{a}_{i,i} &> \sum_{\substack{i=1\\ j\ne i}}^{n} a_{i,j}, \forall i = 1, \ldots, n. && (1.126)\end{aligned}$$
La matrice $\tilde{A}$ est donc à diagonale dominante stricte, et par l'exercice 65 page 100, elle est donc inversible.
3. La méthode de Jacobi pour la résolution du système (1.118) s'écrit :
$$\tilde{D}u^{(k+1)} = (E + F)u^{(k)} + b, \tag{1.127}$$
avec $\tilde{D} = \lambda \mathrm{Id} + D$, et $A = D - E - F$ est la décomposition habituelle de A en partie diagonale, triangulaire inférieure et triangulaire supérieure. Comme $a_{i,i} \ge 0$ et $\lambda \in \mathbb{R}_+^*$, on en déduit que $\tilde{D}$ est inversible, et que donc la suite $(u^{(k)})_{k\in\mathbb{N}}$ est bien définie dans $\mathbb{R}$.
4. Par définition de la méthode de Jacobi, on a :
$$u_i^{(k+1)} = \frac{1}{a_{i,i} + \lambda}(-\sum_{\substack{j=1,n\\ j\ne i}} a_{i,j}u_j^{(k)} + b_i).$$
On en déduit que
$$u_i^{(k+1)} - u_i^{(k)} = \frac{1}{a_{i,i} + \lambda}\sum_{\substack{j=1,n\\ j\ne i}} -a_{i,j}(u_j^{(k)} - u_j^{(k-1)}).$$
et donc
$$\|u^{(k+1)} - u^{(k)}\|_A \le \sum_{i=1}^{n} \frac{a_{i,i}}{a_{i,i} + \lambda}|\sum_{\substack{j=1,n\\ j\ne i}} a_{i,j}(u_j^{(k)} - u_j^{(k-1)})|.$$
Or $\dfrac{a_{i,i}}{a_{i,i} + \lambda} \le \dfrac{1}{1 + \frac{\lambda}{a_{i,i}}} \le \dfrac{1}{1 + \alpha}$. On a donc
$$\|u^{(k+1)} - u^{(k)}\|_A \le \frac{1}{1 + \alpha}\sum_{j=1}^{n} |u_j^{(k)} - u_j^{(k-1)}] \sum_{\substack{j=1,n\\ j\ne i}} -a_{i,j}.$$
Et par hypothèse, $-\sum_{\substack{j=1,n\\ j\ne i}} a_{i,j} = a_{j,j}$. On en déduit que
$$\|u^{(k+1)} - u^{(k)}\|_A \le \frac{1}{1 + \alpha}\|u^{(k)} - u^{(k-1)}\|_A.$$
On en déduit le résultat par une récurrence immédiate.

Université de Marseille, année 2018-2019
Licence de Mathématiques, analyse numérique
Travaux Pratiques 3, en python

Soient $A \in \mathcal{M}_n(\mathbb{R})$ $(n > 1)$ une matrice inversible et $\boldsymbol{b} \in \mathbb{R}^n$ donnés ; on s'intéresse ici aux méthodes itératives permettant de calculer la solution du système linéaire $A\boldsymbol{x} = \boldsymbol{b}$.

On considère des méthodes qui s'écrivent sous le formalisme suivant :

$$P\boldsymbol{x}^{(k+1)} = N\boldsymbol{x}^{(k)} + \boldsymbol{b}, \text{ pour tout } k \geq 0.$$

On rappelle que pour que la suite $(\boldsymbol{x}^{(k)})_{k\in N}$ converge (quelque soit le choix initial $\boldsymbol{x}^{(0)}$ et quelque soit le second membre $\boldsymbol{b}$) il faut et il suffit que le rayon spectral $\rho(P^{-1}N)$ de la matrice $P^{-1}N < 1$ satisfasse $\rho(P^{-1}N)N < 1$.

On met A sous la forme $A = D - E - F$, où D est la diagonale de A, E est la partie triangulaire inférieure de A (diagonale exclue) et F est la partie triangulaire supérieure de A (diagonale exclue).
Avec python, on peut définir les matrices D ,E et F à partir de A en utilisant les commandes **numpy.diag**, **numpy.tril** et **numpy.triu**.

Pour P, N et ε donnés avec $\varepsilon > 0$, la méthode s'écrit :

- **Initialisation** Choisir $x^{(0}$ dans $\mathbb{R}^n$,
- **Itération** k Pour $k \geq 0$, si $|\boldsymbol{r}^{(k}| > \varepsilon$, avec $\boldsymbol{r}^{(k)} = \boldsymbol{b} - A\boldsymbol{x}^{(k)}$,
  calculer $\boldsymbol{x}^{(k+1)}$ solution de $P\boldsymbol{x}^{(k+1)} = N\boldsymbol{x}^{(k)} + b$.

Pour tester et comparer les méthodes de Jacobi, Gauss-Seidel et SOR, on va s'intéresser à des problèmes dont on connait la solution exacte, notée $\boldsymbol{x}$. Les deux problèmes considérés sont :

- **Premier problème** $A = \begin{bmatrix} 10 & 7 & 8 & 7 \\ 7 & 5 & 6 & 5 \\ 8 & 6 & 10 & 9 \\ 7 & 5 & 9 & 10 \end{bmatrix}$ et $\boldsymbol{b} = \begin{bmatrix} 32 \\ 23 \\ 33 \\ 31 \end{bmatrix}$.
- **Second problème** On prend pour A la matrice de l'exercice 1 du TP1. Pour $n \in \mathbb{N}^\star$, on note A_n cette matrice (qui était notée A_h au tp1, avec $h = 1/(n+1)$).

  Le second membre $\boldsymbol{b}$ correspond à $f(x) = \pi^2 \sin(2\pi x)$ pour $x \in [0,1]$ (il est différent de celui des TP 1 et 2). Pour $n \in \mathbb{N}^\star$, on note $\boldsymbol{b}_n$ ce second membre. Le problème discrétisé s'écrit donc $A_n x_n = \boldsymbol{b}_n$.

  On pourra se limiter à considérer le cas $n = 10$.

1. On utilise ici la méthode de Jacobi ($P = D$, $N = E + F$).

(a) Construire une fonction dont les paramètres sont A, $\boldsymbol{b}$, ε et $\boldsymbol{x}^{(0)}$ et qui calcule, avec la méthode de Jacobi, la solution approchée du système linéaire $A\boldsymbol{x} = \boldsymbol{b}$ et retourne le résidu ($|\boldsymbol{r}^{(k}|$) et l'erreur ($|\boldsymbol{x} - \boldsymbol{x}^{(k)}|$) à chaque itération de la méthode.

(b) Pour les deux problèmes considérés, calculer $\rho(P^{-1}N)$.

Si $\rho(P^{-1}N) < 1$, construire pour $\varepsilon = 10^{-6}$ et deux choix différents de $x^{(0)}$ (par exemple, $x^{(0)} = 0$ et $x^{(0)}$ dont toutes les composantes sont égales à 1) un graphique, avec une échelle semilog, donnant l'évolution de l'erreur en fonction de k et l'évolution de la quantité $\rho(P^{-1}N)^k$ [Utiliser **matplotlib.pyplot.semilogy**].

L'erreur se comporte-t-elle comme $\rho(P^{-1}N)^k$? Donner aussi le nombre d'itérations utilisées.

2. Reprendre les questions 1a et 1b avec la méthode de Gauss-Seidel ($P = D - E$, $N = F$).

3. On s'intéresse maintenant à la méthode SOR, c'est-à-dire $P = \frac{1}{\omega}D - E$, $N = F + (\frac{1}{\omega} - 1)D$ avec $0 < \omega < 2$.

(a) Reprendre la question 1a en ajoutant le paramètre ω.

(b) Calculer une approximation de la valeur de ω minimisant $\rho(P^{-1}N)$.

(c) Reprendre la question 1b avec la valeur de ω trouvée dans la question précédente.

4. Comparer l'efficacité des trois méthodes (Jacobi, Gauss-Seidel et SOR) sur ces deux exemples.

Pourquoi dans certains cas la méthode de Jacobi donne-t-elle la solution (approchée) avec un nombre d'itérations inférieur au nombre d'itérations nécessaire pour la méthode de Gauss-Seidel ?