Analyse numérique, Matmeca 1ere année
Corrigé de la feuille 3[1]

1. Méthode itérative : un exemple

A est une matrice régulière dont la diagonale D est inversible. On propose de résoudre le système $\mathrm{A}x = b$ par la méthode itérative suivante :

$$\begin{cases} x_0 \in \mathbb{R}^n \\ x_{k+1} = (\mathrm{I} - \alpha \mathrm{D}^{-1}\mathrm{A})x_k + \alpha \mathrm{D}^{-1} b \end{cases}$$

1.a Supposons que la suite $(x_k)_{k\in\mathbb{N}}$ soit convergente, et notons x sa limite. Alors en passant à la limite dans l'égalité, on trouve :

$$x = (\mathrm{I} - \alpha \mathrm{D}^{-1}\mathrm{A})x + \alpha \mathrm{D}^{-1} b$$

d'où

$$\mathrm{D}x = \mathrm{D}x - \alpha \mathrm{A}x + \alpha b$$

soit $\mathrm{A}x = b$.

1.b

$$\left(\mathrm{D}^{-1}\mathrm{A}\right)_{i,j} = \frac{a_{i,j}}{a_{i,i}}$$

1.c A est à diagonale strictement dominante, c'est à dire

$$\forall i \in [\![1\,;\,n]\!] \qquad |a_{i,i}| > \sum_{j\neq i} |a_{i,j}| \tag{1}$$

Comme on l'a vu à la feuille 1, pour une matrice M donnée,

$$\|\mathrm{M}\|_1 = \max_{j=1\ldots\mathrm{N}} \sum_{i=1}^{\mathrm{N}} |m_{i,j}|$$
$$\|\mathrm{M}\|_\infty = \max_{i=1\ldots\mathrm{N}} \sum_{j=1}^{\mathrm{N}} |m_{i,j}|$$

Dans notre cas, on a

$$(\mathrm{I} - \alpha \mathrm{D}^{-1}\mathrm{A})_{i,j} = \delta_{i,j} - \alpha \frac{a_{i,j}}{a_{i,i}}$$

On en déduit immédiatement

$$\sum_{j=1}^{\mathrm{N}} \left|(\mathrm{I} - \alpha \mathrm{D}^{-1}\mathrm{A})_{i,j}\right| = \alpha \sum_{j\neq i} \left|\frac{a_{i,j}}{a_{i,i}}\right|$$

D'après (1), on a

$$\sum_{j\neq i} \left|\frac{a_{i,j}}{a_{i,i}}\right| < 1$$

[1] généré avec $\LaTeX\,2_\varepsilon$. Tous les commentaires, compléments, insultes et remarques désobligeantes sont les bienvenus à `perrier@math.u-bordeaux1.fr`

donc si $0 < \alpha \leqslant 1$

$$\alpha \sum_{j \neq i} \left| \frac{a_{i,j}}{a_{i,i}} \right| < \alpha \leqslant 1$$

d'où l'on déduit

$$||\mathrm{I} - \alpha \mathrm{D}^{-1}\mathrm{A}||_\infty < 1$$

Comme $||\mathrm{I} - \alpha \mathrm{D}^{-1}\mathrm{A}||_\infty < 1$, on en déduit que le rayon spectral de la matrice d'itération est strictement inférieur à 1, donc d'après le cours, la méthode itérative est convergente.

2. Méthodes de Jacobi et de Gauss-Seidel

A est égale à

$$\begin{pmatrix}
2 & -1 & 0 & \dots & \dots & \dots & 0 \\
-1 & 3 & -2 & \ddots & & & \vdots \\
0 & -1 & 4 & -3 & \ddots & & \vdots \\
\vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \\
\vdots & & \ddots & \ddots & \ddots & \ddots & 0 \\
\vdots & & & \ddots & -1 & n & -(n-1) \\
0 & \dots & \dots & \dots & 0 & -1 & n+1
\end{pmatrix}$$

$\boxed{\textbf{2.a}}$ On reprend les notations classiques : $\mathrm{A} = \mathrm{D} - \mathrm{E} - \mathrm{F}$. La méthode de Jacobi s'écrit alors

$$\left\{ \begin{array}{l} x_0 \in \mathbb{R}^n \\ \mathrm{D}x_{k+1} = (\mathrm{E} + \mathrm{F})x_k + b \end{array} \right.$$

ou encore

$$\left\{ \begin{array}{l} x_0 \in \mathbb{R}^n \\ x_{k+1} = \mathrm{D}^{-1}(\mathrm{E} + \mathrm{F})x_k + \mathrm{D}^{-1}b \end{array} \right.$$

soit $\mathrm{J} = \mathrm{D}^{-1}(\mathrm{E} + \mathrm{F})$. Après calcul, J est égale à

$$\begin{pmatrix}
0 & \frac{1}{2} & 0 & \dots & \dots & \dots & 0 \\
\frac{1}{3} & 0 & \frac{2}{3} & \ddots & & & \vdots \\
0 & \frac{1}{4} & 0 & \frac{3}{4} & \ddots & & \vdots \\
\vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \\
\vdots & & \ddots & \ddots & \ddots & \ddots & 0 \\
\vdots & & & \ddots & \frac{1}{n} & 0 & \frac{n-1}{n} \\
0 & \dots & \dots & \dots & 0 & \frac{1}{n+1} & 0
\end{pmatrix}$$

On applique le théorème de Gershgörin afin de localiser les valeurs propres de J. On a

$$\begin{array}{rl}
 & \Lambda_1 = \frac{1}{2} \\
\forall i \in [\![2\,;\, n-1]\!] & \Lambda_i = 1 \\
 & \Lambda_n = \frac{1}{n+1}
\end{array}$$

On en déduit que les valeurs propres de J sont dans la boule de centre 0, et de rayon 1. Supposons qu'une valeur propre soit de module 1. Alors comme J est irréductible, cette valeur propre est dans

$$\partial \mathrm{B}(0, \frac{1}{2}) \bigcap \partial \mathrm{B}(0,1) \bigcap \partial \mathrm{B}(0, \frac{1}{n+1}) = \varnothing$$

Donc il n'existe pas de valeur propre de module 1, donc $\rho(\mathrm{J}) < 1$. On en déduit que la méthode de Jacobi est convergente.

> **Remarque**
> Cette matrice rentre dans le cadre du théorème disant que la méthode de Jacobi pour le système $\mathrm{A}x = b$ converge si la matrice A est irréductible à forte dominante diagonale.

2.b En conservant les mêmes notations, la méthode de Gauss–Seidel s'écrit

$$\begin{cases} x_0 \in \mathbb{R}^n \\ (\mathrm{D}-\mathrm{E})x_{k+1} = \mathrm{F}x_k + b \end{cases}$$

ou encore

$$\begin{cases} x_0 \in \mathbb{R}^n \\ x_{k+1} = (\mathrm{D}-\mathrm{E})^{-1}\mathrm{F}x_k + (\mathrm{D}-\mathrm{E})^{-1}b \end{cases}$$

La matrice d'itération est donc $\mathrm{G} = (\mathrm{D}-\mathrm{E})^{-1}\mathrm{F}$. Afin d'étudier la convergence de la méthode de Gauss–Seidel, on étudie le spectre de G. Pour cela, on commence par manipuler l'expression formelle du polynôme caractéristique afin de ne retrouver que des matrices déjà calculées ou facilement calculables (D'une manière générale, il est hors de question de calculer explicitement $(\mathrm{D}-\mathrm{E})^{-1}$).

$$\begin{aligned}
\det(\lambda \,\mathrm{Id} - \mathrm{G}) &= \det(\lambda \,\mathrm{Id} - (\mathrm{D}-\mathrm{E})^{-1}\mathrm{F}) \\
&= \lambda^n \det(\mathrm{Id} - \frac{(\mathrm{D}-\mathrm{E})^{-1}\mathrm{F}}{\lambda}) \\
&= \lambda^n \det((\mathrm{D}-\mathrm{E})^{-1}) \det(\mathrm{D}-\mathrm{E}-\frac{\mathrm{F}}{\lambda}) \\
&= \lambda^n \det((\mathrm{D}-\mathrm{E})^{-1}) \det(\mathrm{D}) \det(\mathrm{Id} + \mathrm{D}^{-1}\mathrm{E} - \frac{\mathrm{D}^{-1}\mathrm{F}}{\lambda}) \\
&= \lambda^n \det(\mathrm{Id} - \mathrm{L} - \frac{\mathrm{U}}{\lambda})
\end{aligned}$$

où on a noté $\mathrm{L} = \mathrm{D}^{-1}\mathrm{E}$ et $\mathrm{U} = \mathrm{D}^{-1}\mathrm{F}$. On a également utilisé le fait que $\det((\mathrm{D}-\mathrm{E})^{-1}) = \det(\mathrm{D}^{-1})$ (ceci est vrai car $(\mathrm{D}-\mathrm{E})^{-1}$ est triangulaire inférieure).

Il reste à étudier l'inversibilité de la matrice $\mathrm{M}(\lambda) = \mathrm{Id} - \mathrm{L} - \frac{\mathrm{U}}{\lambda}$, que l'on explicite :

$$\begin{pmatrix}
1 & -\frac{1}{2\lambda} & 0 & \cdots & \cdots & \cdots & 0 \\
-\frac{1}{3} & 1 & \frac{2}{3\lambda} & \ddots & & & \vdots \\
0 & -\frac{1}{4} & 1 & \frac{3}{4\lambda} & \ddots & & \vdots \\
\vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \\
\vdots & & \ddots & \ddots & \ddots & \ddots & 0 \\
\vdots & & & \ddots & -\frac{1}{n} & 1 & \frac{n-1}{n\lambda} \\
0 & \cdots & \cdots & \cdots & 0 & -\frac{1}{n+1} & 1
\end{pmatrix}$$

On calcule les Λ_i du théorème de Gershgörin :

$$\begin{aligned}
\Lambda_1 &= \frac{1}{2\,|\lambda|} \\
\Lambda_i &= \frac{1}{i+1} + \frac{i}{|\lambda|\,(i+1)} \qquad \forall i \in [\![\, 2\,;\, n-1 \,]\!] \\
\Lambda_n &= \frac{1}{n+1}
\end{aligned}$$

Si l'on suppose que $|\lambda| \geqslant 1$, alors on voit que $\mathrm{M}(\lambda)$ est à forte dominante diagonale. De plus, elle est irréductible, on en déduit que si $|\lambda| \geqslant 1$, alors $\mathrm{M}(\lambda)$ est inversible, donc le déterminant $\det(\lambda\,\mathrm{Id} - \mathrm{G})$ est non nul. Il vient alors que $\rho(\mathrm{G}) < 1$, donc que la méthode est convergente.

2.c Les algorithmes de résolution itérative de systeme suivent en général

1. Une phase d'initialisation : On choisit une donnée initiale. On calcule de résidu [2] associé à cette valeur initiale
2. Une boucle `Tant que` dans laquelle on effectue les itérations sur x_n. Cette boucle dépend d'un paramètre « petit » (ici `erreur`) qui est une variable « globale » et contrôle le résidu maximal que l'on veut avoir

```
#####################################################
#Initialisation de l'algorithme
#####################################################
#On initialise l'algorithme à b (ou à zéro, ou à n'importe lequel
#de vos vecteurs préférés)
POUR i de 1 à n FAIRE
   x(i)=b(i)
FIN POUR i

#On calcule le résidu, c'est à dire r=A*x-b
r(1)=2*x(1)-x(2)-b(1)
r(n)=-x(n-1)+(n+1)*x(n)-b(n)
POUR i de 1 à n FAIRE
   r(i)=-x(i-1)+(i+1)*x(i)-i*x(i+1)-b(i)
FIN POUR i

#On calcule la norme du résidu
S=0
POUR i de 1 à n FAIRE
   S=S+|r(i)|^2
FIN POUR i

#####################################################
#On rentre à proprement parler dans l'algorithme
#####################################################

TANT QUE (S>erreur)

     #On commence par calculer u=F*x_k + b
```

[2] **Rappel :** on appelle résidu la quantité $b - \mathrm{A}x$

```
    POUR i de 1 à n-1 FAIRE
       u(i)=i*x(i-1)+b(i)
    FIN POUR i
    u(n)=b(n)

    #Il reste à résoudre(D-E)*x=u, ce qui est
    #facile car (D+E) est triangulaire (et meme bidiagonale)
    x(1)=0.5*u(1)
    POUR i de 2 à n FAIRE
       x(i)=(u(i)+x(i-1))/(i+1)
    FIN POUR i

    #FIN de l'itération, on calcule le nouveau
    #résidu, ainsi que sa norme
    r(1)=2*x(1)-x(2)-b(1)
    r(n)=-x(n-1)+(n+1)*x(n)-b(n)
    POUR i de 1 à n FAIRE
       r(i)=-x(i-1)+(i+1)*x(i)-i*x(i+1)-b(i)
    FIN POUR i

    #On calcule la norme du résidu
    S=0
    POUR i de 1 à n FAIRE
       S=S+|r(i)|^2
    FIN POUR i

FIN TANT QUE
```

Remarque : On voit bien ici qu'il est beaucoup plus simple de calculer x_{k+1} en multipliant par F, puis en ajoutant b, et enfin en inversant le système $(\mathrm{D}-\mathrm{E})x = u$, ce qui revient à faire (à la louche) $3n$ opérations plutôt que de multiplier par $(\mathrm{D}-\mathrm{E})^{-1}\mathrm{F}$, qui revient à faire n^2 opérations puisque la matrice $(\mathrm{D}-\mathrm{E})^{-1}\mathrm{F}$ est triangulaire inférieure, mais pleine.

3. Méthode itérative et M–matrices

On s'intéresse au problème aux limites

$$\begin{cases} -u''(x) = f(x) & \forall x \in \,]\,0\,;1\,[\\ u(0) = u(1) = 0 \end{cases}$$

$\boxed{\textbf{3.i}}$ Si f est positive sur $[\,0\,;1\,]$, alors u est concave. On en déduit que le graphe de f est au dessus de toutes ses cordes, donc en particulier au dessus de la corde joignant le point $(0, u(0)) = (0, 0)$ et le point $(1, u(1)) = (1, 0)$, qui est incluse dans l'axe des abscisses. On en déduit que $u \geqslant 0$ sur$[\,0\,;1\,]$.

Cette équation est une équation de diffusion : on la trouve par exemple

- en effectuant un bilan d'énergie sur un matériau rectiligne fini dont les extrémités ont une température imposée. Ce matériau est chauffé par une source volumique f. Alors on a

$$-\operatorname{div}(\overrightarrow{q}) + f = 0$$

où $\overrightarrow{q}$ est le vecteur flux de chaleur. En supposant de plus que le matériau suit la **Loi de Fourier**

$$\overrightarrow{q} = -\lambda \overrightarrow{\nabla} \mathrm{T}$$

et que $\lambda > 0$ est constant, alors T vérifie

$$-\mathrm{T}'' = \frac{f}{\lambda}$$

- en effectuant un bilan de matière sur un milieu rectiligne fini dont les extrêmités ont une densité donnée. Ce milieu est alimenté par une source de matière f. Alors on a

$$-\operatorname{div}(\overrightarrow{\jmath}) + f = 0$$

où $\overrightarrow{\jmath}$ est le vecteur flux de matière. En supposant de plus que le milieu obéit à la **Loi de Fick**

$$\overrightarrow{\jmath} = -\mathrm{D} \overrightarrow{\nabla} \rho$$

et que $\mathrm{D} > 0$ est constant, alors ρ vérifie

$$-\rho'' = \frac{f}{\mathrm{D}}$$

Le résultat de cette question nous dit que dans ce cas, la température ou la densité restent positifs.

3.ii La discrétisation de l'exercice 4 de la feuille 1 avait amené au système $\mathrm{AU} = \mathrm{B}$ avec

$$\mathrm{A} = \begin{pmatrix} 2 & -1 & 0 & 0 & \dots & 0 \\ -1 & 2 & -1 & 0 & \dots & 0 \\ 0 & \ddots & \ddots & \ddots & & \vdots \\ \vdots & & \ddots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & -1 & 2 & -1 \\ 0 & \dots & 0 & 0 & -1 & 2 \end{pmatrix}$$

et

$$\mathrm{B} = \begin{pmatrix} h^2 f(x_1) \\ h^2 f(x_2) \\ \vdots \\ h^2 f(x_n) \end{pmatrix}$$

On suppose que les B_i sont positifs (ce qui est le cas si f est positive). Montrons que dans ce cas les U_i sont positifs. Pour cela, on écrit la méthode de Jacobi associée à ce système : on considère la suite $\left(\mathrm{U}^{(k)}\right)_{k \in \mathbb{N}}$

$$\left\{ \begin{array}{l} \mathrm{U}^{(0)} \text{ est donnée} \\ \mathrm{DU}^{(k+1)} = (\mathrm{E} + \mathrm{F})\mathrm{U}^{(k)} + \mathrm{B} \end{array} \right.$$

où D est la matrice des coefficients diagonaux de A, et E et F sont les opposés des matrices triangulaires supérieure et inférieure de A. Les coefficients de D, E et F sont donc positifs. Supposons que les coefficients de $\mathrm{U}^{(k)}$ soient tous positifs. Alors les coefficients de $(\mathrm{E} + \mathrm{F})\mathrm{U}^{(k)}$ sont tous positifs, et ceux de B sont tous positifs par hypothèse. On en déduit que les coefficients de $(\mathrm{E} + \mathrm{F})\mathrm{U}^{(k)} + \mathrm{B}$ sont tous positifs. Comme les coefficients de D, donc de D^{-1} sont également positifs (parce que D est diagonale), il vient que les coefficients de $\mathrm{U}^{(k+1)}$ sont tous positifs. Si l'on choisit par exemple $\mathrm{U}^{(0)} = \mathrm{B}$, alors tous les coefficients de $\mathrm{U}^{(k)}$ sont positifs, et ce, quel que soit k. De plus, A est irréductible à diagonale fortement dominante, donc la méthode de Jacobi converge : les $\mathrm{U}^{(k)}$ convergent vers U, la solution de $\mathrm{AU} = \mathrm{B}$. Comme tous les coefficients de $\mathrm{U}^{(k)}$ sont positifs, on en déduit que tous les coefficients de U sont positifs.

3.iii On considère le système $\mathrm{A}x = b$, où A est une M–matrice. Montrons que si tous les coefficients de b sont positifs, alots tous les coefficients de x sont positifs. Pour cela, on considère la méthode de Jacobi :

$$\begin{cases} x^{(0)} \\ \mathrm{M}x^{(k+1)} = \mathrm{N}x^{(k)} + b \end{cases}$$

où M est la matrice diagonale dont les éléments diagonaux sont les mêmes que ceux de A, et N est telle que $\mathrm{N} = \mathrm{M} - \mathrm{A}$. On en déduit que les coefficients de M et de N sont tous positifs. Supposons que les coefficients de $x^{(0)}$ sont tous positifs. Montrons la propriété suivante par récurrence

$$\mathscr{P}(k) \ : \ « \forall i \in [\![1 \,;\, n]\!] \quad x_i^{(k)} \geqslant 0 »$$

- $\underline{\mathscr{P}(0)}$ est vraie par hypothèse
- $\underline{\mathscr{P}(k) \Longrightarrow \mathscr{P}(k+1)}$ Supposons que $\mathscr{P}(k)$ soit vraie. Alors

$$\forall i \in [\![1 \,;\, n]\!] \qquad x_i^{(k)} \geqslant 0$$

  Par hypothèse, on sait que les coefficients de b sont positifs. De plus, comme les coefficients de N sont positifs, et que par hypothèse de récurrence, les coefficients de $x^{(k)}$ sont positifs, les coefficients de $\mathrm{N}x^{(k)}$ sont tous positifs. On en déduit que les coefficients de $\mathrm{N}x^{(k)} + b$ sont tous positifs. Enfin, comme M est diagonale et que tous ses coefficients sont positifs, les coefficients de M^{-1} sont également tous positifs. Ainsi, les coefficients de $\mathrm{M}^{-1}(\mathrm{N}x^{(k)} + b)$, c'est à dire les coefficients de $x^{(k+1)}$ sont tous positifs, donc $\mathscr{P}(k+1)$ est vraie.
- <u>Conclusion :</u> Si pour tout i dans $[\![1 \,;\, n]\!]$ $x_i^{(0)} \geqslant 0$, alors pour tout k dans $\mathbb{N}$, et pour tout i dans $[\![1 \,;\, n]\!]$, $x_i^{(k)} \geqslant 0$.

Comme on a supposé que A est soit à diagonale strictement dominante, soit irréductible à diagonale fortement dominante, la méthode de Jacobi converge, c'est à dire que

$$\forall i \in \mathbb{N} \qquad x_i = \lim_{k \to \infty} x_i^{(k)} \geqslant 0$$

On en déduit que les M–matrices, c'est à dire les matrices telles que

- A est à diagonale strictement dominante ou irréductible à diagonale fortement dominante
- les coefficients de la diagonale de A sont positifs, et les coefficients hors diagonale sont tous négatifs.

suivent la propriété suivante

> Lors de la résolution de $\mathrm{A}x = b$, si pour tout i $b_i \geqslant 0$ alors pour tout i, $x_i \geqslant 0$.

4. Problèmes de convection

On s'intéresse au problème de convection–réaction suivant

$$\begin{cases} v\partial_x \rho(x) = -c\rho(x) & \forall x \in]\,0\,;1\,[\\ \rho(1) = \rho^\star \end{cases}$$

où v et c sont des constantes, avec $v < 0$ et $c > 0$. Une approximation de ρ_i est donnée par

$$\begin{cases} v(\rho_{i+1} - \rho_i) + ch\rho_i = 0 \\ -v\rho_n + ch\rho_n = -v\rho^\star \end{cases}$$

4.i Afin de trouver les ρ_i, on est amené à résoudre le système $\mathrm{A}\rho = b$, avec

$$\mathrm{A} = \begin{pmatrix} ch - v & v & 0 & \dots & 0 \\ 0 & ch - v & v & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ \vdots & & \ddots & ch - v & v \\ 0 & \dots & \dots & 0 & ch - v \end{pmatrix}$$

et $b = {}^t(0, \dots, 0, -v\rho^\star)$. La matrice A est telle que

- les coefficients diagonaux sont positifs, et les coefficients extra diagonaux sont négatifs
- $|ch - v| = ch + |v| > |-v| = |v|$, donc la matrice A est à stricte dominante diagonale.

A est donc une M–matrice. Le système que l'on a à résoudre est tel que les b_i sont tous positifs ou nuls (en fait, ils sont tous nuls, sauf le dernier qui est égal à $-v\rho^\star$, avec $\rho^\star > 0$ et $v < 0$, et qui est donc positif), et A est une M–matrice. On en déduit que les ρ_i sont tous positifs ou nuls.

4.ii Le système que l'on a à résoudre est triangulaire supérieur, il peut donc se résoudre par un algorithme de remontée.

4.iii On décompose la matrice sous la forme $\mathrm{A} = \mathrm{D} - \mathrm{E} - \mathrm{F}$, où D est la partie diagonale de A, E est l'opposée de la partie triangulaire inférieure de A (ici, on a $\mathrm{E} = 0$), et F est l'opposée de la partie triangulaire supérieure de A. Alors la méthode de Jacobi s'écrit sous la forme

$$\begin{cases} \rho^{(0)} \\ \mathrm{D}\rho^{(k+1)} = (\mathrm{E} + \mathrm{F})\rho^{(k)} + b \end{cases}$$

La matrice d'itération de Jacobi est $\mathrm{D}^{-1}(\mathrm{E} + \mathrm{F}) = \mathrm{D}^{-1}\mathrm{F}$, elle est donc triangulaire supérieure avec des 0 sur la diagonale. Comme cette matrice est triangulaire supérieure, on peut lire ses valeurs propres sur la diagonale. La seule valeur propre de J est donc 0. On en déduit que $\rho(\mathrm{J}) = 0$. Soit $\bar\rho$ la solution du système $\mathrm{A}\rho = b$. Alors $\bar\rho$ vérifie

$$\mathrm{D}\bar\rho = (\mathrm{E} + \mathrm{F})\bar\rho + b$$

donc on trouve

$$\rho^{(k+1)} - \bar\rho = \mathrm{J}(\rho^{(k)} - \bar\rho)$$

ou encore

$$\rho^{(k)} = \bar\rho + \mathrm{J}^k(\rho^{(0)} - \bar\rho)$$

La matrice J n'ayant que 0 comme valeur propre, elle est nilpotente. On en déduit que $\mathrm{J}^n = 0$, donc que $\rho^{(n)} = \bar{\rho}$, c'est à dire que la méthode de Jacobi converge en au plus n itérations. Il en est de même pour la méthode de Gauss–Seidel, car dans notre cas, la méthode de Jacobi et la méthode de Gauss–Seidel sont les mêmes (c'est toujours le cas lorsque $\mathrm{E} = 0$).

4.iv Si l'on inverse la numérotation du maillage, alors on est amené à résoudre le système $\widetilde{\mathrm{A}}\widetilde{\rho} = \widetilde{b}$ avec

$$\begin{pmatrix} ch - v & 0 & \dots & \dots & 0 \\ v & ch - v & \ddots & & \vdots \\ 0 & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & v & ch - v & 0 \\ 0 & \dots & 0 & v & ch - v \end{pmatrix}$$

et $\widetilde{b} = {}^t(-v\rho^\star, 0, \dots, 0)$. Dans ce cas, la méthode de Gauss-Seidel s'écrit

$$\left\{ \begin{array}{l} \widetilde{\rho}^{(0)} \\ (\widetilde{\mathrm{D}} - \widetilde{\mathrm{E}})\widetilde{\rho}^{(k+1)} = \widetilde{\mathrm{F}}\widetilde{\rho}^{(k)} + \widetilde{b} \end{array} \right.$$

avec $\widetilde{\mathrm{F}} = 0$. On trouve donc à la première itération $\widetilde{\rho}^{(1)} = \bar{\rho}$.

5. Méthode de Jacobi par blocs

5.a Dans la feuille 2, on a vu que la discrétisation du Laplacien en dimension 2 sur une grille cartésienne $((n+2)\times(n+2))$ menait à la matrice suivante

$$\mathrm{A} = \begin{pmatrix} \widetilde{\mathrm{A}} & -\mathrm{I}_n & 0_n & \dots & \dots & 0_n \\ -\mathrm{I}_n & \widetilde{\mathrm{A}} & -\mathrm{I}_n & \ddots & & \vdots \\ 0_n & -\mathrm{I}_n & \widetilde{\mathrm{A}} & -\mathrm{I}_n & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & 0_n \\ \vdots & & \ddots & -\mathrm{I}_n & \widetilde{\mathrm{A}} & -\mathrm{I}_n \\ 0_n & \dots & \dots & 0_n & -\mathrm{I}_n & \widetilde{\mathrm{A}} \end{pmatrix}$$

avec

$$\widetilde{\mathrm{A}} = \begin{pmatrix} 4 & -1 & 0 & \dots & \dots & 0 \\ -1 & 4 & -1 & \ddots & & \vdots \\ 0 & -1 & 4 & -1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & 0 \\ \vdots & & \ddots & -1 & 4 & -1 \\ 0 & \dots & \dots & 0 & -1 & 4 \end{pmatrix}$$

5.b On décompose A sous la forme $\mathrm{A} = \mathrm{D} - \mathrm{E} - \mathrm{F}$, avec

$$\mathrm{D} = \begin{pmatrix} \widetilde{\mathrm{A}} & 0_n & 0_n & \dots & \dots & 0_n \\ 0_n & \widetilde{\mathrm{A}} & 0_n & \ddots & & \vdots \\ 0_n & 0_n & \widetilde{\mathrm{A}} & 0_n & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & 0_n \\ \vdots & & \ddots & 0_n & \widetilde{\mathrm{A}} & 0_n \\ 0_n & \dots & \dots & 0_n & 0_n & \widetilde{\mathrm{A}} \end{pmatrix}$$

et

$$\mathrm{E}+\mathrm{F} = \begin{pmatrix} 0_n & \mathrm{I}_n & 0_n & \dots & \dots & 0_n \\ \mathrm{I}_n & 0_n & \mathrm{I}_n & \ddots & & \vdots \\ 0_n & \mathrm{I}_n & 0_n & \mathrm{I}_n & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & 0_n \\ \vdots & & \ddots & \mathrm{I}_n & 0_n & \mathrm{I}_n \\ 0_n & \dots & \dots & 0_n & \mathrm{I}_n & 0_n \end{pmatrix}$$

Et la méthode de Jacobi par blocs s'écrit alors

$$\begin{cases} x_0 \in \mathbb{R}^{n^2} \\ \mathrm{D}x_{k+1} = (\mathrm{E}+\mathrm{F})x_k + b \end{cases}$$

Après le calcul des composantes de $y = (\mathrm{E}+\mathrm{F})x_k + b$, on voit qu'il reste à résoudre le système $\mathrm{D}x_{k+1} = y$. Si l'on note $\bar{y}^1$ le vecteur des n premières composantes de y, $\bar{y}^2$ le vecteur des n suivantes, etc ..., $\bar{y}^n$ le vecteur des n dernières composantes de y, et en faisant de même pour x_{k+1}, on voit qu'on devra résoudre les systèmes linéaires suivants (pour tout $1 \leqslant i \leqslant n$) :

$$\widetilde{\mathrm{A}}\bar{x}^i_{k+1} = \bar{y}^i$$

5.c On voit que l'on va avoir à résoudre un grand nombre de fois le système $\widetilde{\mathrm{A}}x = b$ (n fois à chaque itération). Il peut donc être utile d'effectuer décomposition LU de $\widetilde{\mathrm{A}}$. En effet, même si cela représente un investissement assez lourd au départ, on aura ensuite à se contenter d'un algorithme de descente et d'un algorithme de remontée pour chaque résolution du système $\widetilde{\mathrm{A}}x = b$, ce qui est très peu. De plus, vu le profil de la matrice A, sa décomposition LU sera très simple, et encore creuse.

Pour l'algorithme, on propose de stocker les données de x dans un tableau à double entrée $n \times n$ de telle sorte que la première colonne représente les n premières composantes de x, la deuxième colonne les n suivantes, etc .... Dans l'algorithme, $\widetilde{\mathrm{A}}$ sera notée C. On ne va pas écrire l'algorithme en entier, ce serait beaucoup trop fastidieux. On suppose donc avoir déjà programmé un algorithme de décomposition LU, et des algorithmes de montée, descente (résolution de systèmes linéaires triangulaires supérieurs et inférieurs) de multiplication de matrices $n \times n$ par vecteur, et de norme d'un vecteur de $\mathbb{R}^n$. On note (comme en `fortran` par exemple) $\mathrm{A}(:, i)$ la i^{eme} colonne de la matrice A.

```
#####################################################
#Une fois pour toutes, on effectue la décomposition LU de C
#####################################################
(L,U)=LUDecomposition(C)




#####################################################
#Initialisation de l'algorithme itératif
#####################################################
#On choisit d'initialiser  l'algorithme à B
POUR i de 1 à n FAIRE
     X(:,i)=B(:,i)
FIN POUR i

#On calcule le résidu, c'est à dire R=A*x-B
R(:,1)=C*X(:,1)-X(:,2)
R(:,n)=C*X(:,n)-X(:,n-1)
POUR i de 1 à n FAIRE
     R(:,i)=-X(:,i-1)+C*X(:,i)-X(:,i+1)
FIN POUR i


#On calcule la norme 2 du résidu
S=0
POUR i de 1 à n FAIRE
```

```
   POUR j de 1 à n FAIRE
        S=S+|R(i,j)|^2
   FIN POUR j
FIN POUR i


#####################################################
#Boucle itérative
#####################################################

TANT QUE (S>erreur)

     #On commence par calculer V=(E+F)*x_k + B
     POUR i de 1 à n-1 FAIRE
        V(:,i)=X(:,i-1)+X(:,i+1)+B(:,i)
     FIN POUR i
     V(:,1)=-X(:,2)
     V(:,n)=-X(:,n-1)

     #Il reste à résoudre  D*X=V, ce qui n'est pas de la tarte
     #(En fait, on le résout par blocs)
     POUR i de 1 à n FAIRE
          #On résout C*X(:,i)=V(:,i)
          #ou encore... L*U*X(:,i)=V(:,i)
          Vprime=algorithme_de_remontée(L,V(:,i))
          X(:,i)=algorithme_de_descente(U,Vprime)
     FIN POUR i

     #FIN de l'itération, on calcule le nouveau
     #résidu, ainsi que sa norme
     R(:,1)=C*X(:,1)-X(:,2)
     R(:,n)=C*X(:,n)-X(:,n-1)
     POUR i de 1 à n FAIRE
          R(:,i)=-X(:,i-1)+C*X(:,i)-X(:,i+1)
     FIN POUR i
     S=0
     POUR i de 1 à n FAIRE
          POUR j de 1 à n FAIRE
             S=S+|R(i,j)|^2
          FIN POUR j
     FIN POUR i

FIN TANT QUE
```

6. Méthode de relaxation

Dans cet exercice, on étudie des méthodes itératives de résolution de $\mathrm{A}x = b$, avec

$$\mathrm{A} = \begin{pmatrix} 4 & -1 & -1 & 0 & \dots & \dots & 0 \\ -2 & 5 & -2 & -1 & \ddots & & \vdots \\ 0 & -2 & 6 & -3 & -1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & 0 \\ \vdots & & \ddots & -2 & n+1 & -(n-2) & -1 \\ \vdots & & & \ddots & -2 & n+2 & -(n-1) \\ 0 & \dots & \dots & \dots & 0 & -2 & n+3 \end{pmatrix}$$

6.a On prend les notations classiques $\mathrm{A} = \mathrm{D} - \mathrm{E} - \mathrm{F}$. La méthode de Jacobi s'écrit alors

$$\begin{cases} x_0 \in \mathbb{R}^n \\ \mathrm{D}x_{k+1} = (\mathrm{E} + \mathrm{F})x_k + b \end{cases}$$

La matrice d'itération de Jacobi est donc $\mathrm{J} = \mathrm{D}^{-1}\mathrm{E} + \mathrm{D}^{-1}\mathrm{F}$, et on a

$$\begin{cases} \mathrm{J}_{i,i-1} = \dfrac{2}{i+3} & \text{pour} \quad i = 2 \dots n \\ \mathrm{J}_{i,i+1} = \dfrac{i}{i+3} & \text{pour} \quad i = 1 \dots n-1 \\ \mathrm{J}_{i,i+2} = \dfrac{1}{i+3} & \text{pour} \quad i = 1 \dots n-2 \end{cases},$$

les autres coefficients étant nuls.

Pour calculer $\|\mathrm{J}\|_\infty$ et $\|\mathrm{J}\|_1$, il suffit de calculer les normes 1 des lignes et des colonnes de J (on note L_i et C_i respectivement les vecteurs lignes et colonnes de J).

- **Norme 1 des colonnes** Pour une colonne i qui n'est pas trop près du bord, on a

$$\begin{aligned} \|\mathrm{C}_i\|_1 &= \frac{2}{i+4} + \frac{i-1}{i+2} + \frac{1}{i+1} \\ &= \frac{2}{i+4} + 1 - \frac{3}{i+2} + \frac{1}{i+1} \\ &= 1 + \frac{2(i+2)(i+1) - 3(i+4)(i+1) + (i+2)(i+4)}{(i+1)(i+2)(i+4)} \\ &= 1 - \frac{5i}{(i+1)(i+2)(i+4)} \end{aligned}$$

D'où

$$\begin{aligned} \|\mathrm{C}_1\|_1 &= \frac{1}{2} \\ \|\mathrm{C}_2\|_1 &= \frac{3}{5} \\ \|\mathrm{C}_n\|_1 &= \frac{n}{n+3} = 1 - \frac{3}{n+3} \\ \|\mathrm{C}_i\|_1 &= 1 - \frac{5i}{(i+1)(i+2)(i+4)} \qquad \text{pour} \quad i = 3 \dots n-1 \end{aligned}$$

- **Norme 1 des lignes**

$$\begin{aligned}
\|\mathrm{L}_1\|_1 &= \frac{1}{2} \\
\|\mathrm{L}_{n-1}\|_1 &= \frac{n+1}{n+2} = 1 - \frac{1}{n+2} \\
\|\mathrm{L}_n\|_1 &= \frac{2}{n+3} \\
\|\mathrm{L}_i\|_1 &= 1 \qquad\qquad \text{pour} \quad i = 2 \dots n-2
\end{aligned}$$

On en déduit

$$\|\mathrm{J}\|_\infty < 1 \qquad \text{et} \qquad \|\mathrm{J}\|_1 = 1$$

Comme $\|\mathrm{J}\|_\infty < 1$, la méthode de Jacobi est convergente.

6.b On note $\mathrm{A} = \mathrm{D} - \mathrm{E} - \mathrm{F}$, et on utilise la méthode de relaxation suivante :

$$\left\{\begin{array}{ll}
x_0 \in \mathbb{R}^n & \\
\mathrm{D}y_{k+\frac{1}{2}} = b + \mathrm{E}x_{k+1} + \mathrm{F}x_k & (\mathbf{1}) \\
\text{avec } \; x_{k+1} = \frac{1}{2}\left(y_{k+\frac{1}{2}} + x_k\right) & (\mathbf{2})
\end{array}\right.$$

6.b.i Si $x_k \to x$, alors d'après ($\mathbf{1}$), la suite $\left(y_{k+\frac{1}{2}}\right)_{k\in\mathbb{N}}$ converge et on note y sa limite. En faisant passer ($\mathbf{2}$) à la limite, on trouve la relation $y = x$. En reportant dans ($\mathbf{1}$) que l'on a également fait passer à la limite, on trouve

$$\mathrm{D}x = b + \mathrm{E}x + \mathrm{F}x$$

ou encore $\mathrm{A}x = b$.

6.b.ii La relation ($\mathbf{2}$) peut se réécrire

$$y_{k+\frac{1}{2}} = 2x_{k+1} - x_k$$

En reportant dans ($\mathbf{1}$), on trouve

$$2\mathrm{D}x_{k+1} - \mathrm{D}x_k = b + \mathrm{E}x_{k+1} + \mathrm{F}x_k$$

d'où

$$(2\mathrm{D} - \mathrm{E})x_{k+1} = (\mathrm{F} + \mathrm{D})x_k + b$$

soit

$$x_{k+1} = \mathrm{G}x_k + \mathrm{H}b$$

avec

$$\left\{\begin{array}{l}
\mathrm{G} = (2\mathrm{D} - \mathrm{E})^{-1}(\mathrm{F} + \mathrm{D}) \\
\mathrm{H} = (2\mathrm{D} - \mathrm{E})^{-1}b
\end{array}\right.$$

6.b.iii Commençons par calculer G en fonction de $\mathrm{L} = \mathrm{D}^{-1}\mathrm{E}$ et $\mathrm{U} = \mathrm{D}^{-1}\mathrm{F}$:

$$\begin{aligned}
\mathrm{G} &= (2\mathrm{D} - \mathrm{E})^{-1}(\mathrm{F} + \mathrm{D}) \\
&= \left(2\mathrm{D}\left(\mathrm{Id} - \frac{\mathrm{D}^{-1}\mathrm{E}}{2}\right)\right)^{-1} \mathrm{D}\left(\mathrm{Id} + \mathrm{D}^{-1}\mathrm{F}\right) \\
&= \left(\mathrm{Id} - \frac{\mathrm{L}}{2}\right)^{-1} \frac{1}{2}\mathrm{D}^{-1}\mathrm{D}\left(\mathrm{Id} + \mathrm{U}\right) \\
\mathrm{G} &= \frac{1}{2}\left(\mathrm{Id} - \frac{\mathrm{L}}{2}\right)^{-1} (\mathrm{Id} + \mathrm{U})
\end{aligned}$$

On a alors

$$\begin{aligned}
\mathrm{P_G}(\lambda) &= \det\left(\lambda\,\mathrm{Id} - \frac{1}{2}\left(\mathrm{Id} - \frac{\mathrm{L}}{2}\right)^{-1}(\mathrm{Id} + \mathrm{U})\right) \\
&= \det\left(\left(\mathrm{Id} - \frac{\mathrm{L}}{2}\right)^{-1}\left(\lambda\left(\mathrm{Id} - \frac{\mathrm{L}}{2}\right) - \frac{1}{2}(\mathrm{Id} + \mathrm{U})\right)\right) \\
&= \det\left(\left(\mathrm{Id} - \frac{\mathrm{L}}{2}\right)^{-1}\right)\det\left(\left(\lambda - \frac{1}{2}\right)\mathrm{Id} - \frac{\lambda}{2}\mathrm{L} - \frac{1}{2}\mathrm{U}\right) \\
&= \left(\lambda - \frac{1}{2}\right)^n \det\left(\mathrm{Id} - \frac{\lambda}{2\lambda - 1}\mathrm{L} - \frac{1}{2\lambda - 1}\mathrm{U}\right)
\end{aligned}$$

(On a utilisé le fait que $\det\left(\left(\mathrm{Id} - \frac{\mathrm{L}}{2}\right)^{-1}\right) = 1$, car il s'agit d'une matrice triangulaire avec uniquement des 1 sur la diagonale).

6.b.iv Les coefficients des différentes matrices s'écrivent

$$\mathrm{L}_{i,i-1} = \frac{2}{i+3} \quad \text{pour} \quad i = 2 \dots n$$

$$\begin{cases} \mathrm{U}_{i,i+1} = \dfrac{i}{i+3} \\ \mathrm{U}_{i,i+2} = \dfrac{1}{i+3} \end{cases}$$

Montrons que la matrice $\mathrm{M}(\lambda) = \mathrm{Id} - \frac{\lambda}{2\lambda-1}\mathrm{L} - \frac{1}{2\lambda-1}\mathrm{U}$ est inversible. Ceci revient à montrer que ${}^t\mathrm{M}(\lambda)$ est inversible, et il suffit donc de montrer que ${}^t\mathrm{M}(\lambda)$ est à stricte dominante diagonale (on n'essaie pas avec $\mathrm{M}(\lambda)$, car pour $\lambda = 1$, la somme des coefficients hors diagonale est égale à 1, et elle n'est donc pas à diagonale strictement dominante). Sur une colonne i qui n'est ni la première, ni la deuxième, ni la dernière, la somme des modules des coefficients hors diagonale de $\mathrm{M}(\lambda)$ est égale à

$$\mathrm{S}_i = \frac{|\lambda|}{|2\lambda - 1|}\frac{2}{i+4} + \frac{1}{|2\lambda - 1|}\left(\frac{i-1}{i+2} + \frac{1}{i+1}\right)$$

Tout d'abord, il convient de trouver des majorations de $\frac{|\lambda|}{|2\lambda-1|}$ et $\frac{1}{|2\lambda-1|}$ lorsque $|\lambda| \geqslant 1$.

D'une part, on a

$$1 \leqslant |\lambda| = \left|\lambda - \frac{1}{2} + \frac{1}{2}\right| \leqslant \left|\lambda - \frac{1}{2}\right| + \left|\frac{1}{2}\right|$$

d'où

$$\frac{1}{2} \leqslant \left|\lambda - \frac{1}{2}\right|$$

soit

$$\frac{1}{|2\lambda - 1|} \leqslant 1.$$

D'autre part, on a

$$|\lambda| = \left|\lambda - \frac{1}{2} + \frac{1}{2}\right| \leqslant \left|\lambda - \frac{1}{2}\right| + \frac{1}{2} \leqslant \left|\lambda - \frac{1}{2}\right| + \frac{|\lambda|}{2}$$

d'où

$$\frac{|\lambda|}{2} \leqslant \left|\lambda - \frac{1}{2}\right|$$

soit

$$\frac{|\lambda|}{|2\lambda - 1|} \leqslant 1.$$

On en déduit

$$\mathrm{S}_i \leqslant \frac{2}{i+4} + \frac{i-1}{i+2} + \frac{1}{i+1}$$

Et cette quantité est strictement inférieure à 1 d'après ce qu'on a fait pour Jacobi. En utilisant de la même manière ce qu'on a vu pour Jacobi (et en utilisant aussi les inégalités sur λ), on montre qu'il en est de même pour la première, la deuxième, et la dernière colonne.

6.b.v Comme pour $|\lambda| \geqslant 1$ la matrice B est inversible, on en déduit que $\mathrm{P_G}(\lambda)$ est non nul pour $|\lambda| \geqslant 1$. Cela veut dire que le rayon spectral de G, matrice d'itération, est strictement inférieur à 1, donc que la méthode de relaxation est convergente.

7. Méthode de splitting

A est une matrice symétrique définie positive ; on reprend les notations $\mathrm{A} = \mathrm{D} - \mathrm{E} - \mathrm{F}$, et on a $\mathrm{F} = {}^t\mathrm{E}$. On étudie la méthode itérative suivante

$$\begin{cases} x_0 \in \mathbb{R}^n & \\ (\mathrm{D} - \mathrm{E})x_{k+\frac{1}{2}} = \mathrm{F}x_k + b & \quad (\mathbf{1}) \\ (\mathrm{D} - \mathrm{F})x_{k+1} = \mathrm{E}x_{k+\frac{1}{2}} + b & \quad (\mathbf{2}) \end{cases}$$

7.a La matrice A étant symétrique définie positive, on en déduit que pour tout vecteur e_i de la base canonique, on a $\langle \mathrm{A}e_i, e_i\rangle > 0$. Or on voit que $\langle \mathrm{A}e_i, e_i\rangle = a_{i,i}$. Par conséquent, les coefficients de la diagonale de A sont tous strictement positifs, donc non nuls. On en déduit que les matrices D, $\mathrm{D} - \mathrm{E}$ et $\mathrm{D} - \mathrm{F}$ sont inversibles, et la méthode est donc bien définie.

Supposons que la suite $(x_k)_{k\in\mathbb{N}}$ converge vers une limite x. Alors d'après l'équation (**1**), la suite $\left(x_{k+\frac{1}{2}}\right)_{k\in\mathbb{N}}$ converge également, vers une limite $\bar{x}$. Alors en faisant passer (**2**) et (**1**) à la limite, on trouve

$$\begin{cases} (\mathrm{D} - \mathrm{E})\bar{x} = \mathrm{F}x + b \\ (\mathrm{D} - \mathrm{F})x = \mathrm{E}\bar{x} + b \end{cases}$$

En retranchant ces deux équations, on trouve

$$(\mathrm{D} - \mathrm{E})\bar{x} - (\mathrm{D} - \mathrm{F})x = \mathrm{F}x - \mathrm{E}\bar{x}$$

d'où $\mathrm{D}\bar{x} = \mathrm{D}x$. La matrice D étant inversible, on en déduit $x = \bar{x}$. En reprenant (**1**) et en le faisant passer à la limite, on obtient

$$(\mathrm{D} - \mathrm{E})x = \mathrm{F}x + b$$

donc $\mathrm{A}x = b$.

$\boxed{\textbf{7.b}}$ L'équation (**1**) donne

$$x_{k+\frac{1}{2}} = (\mathrm{D}-\mathrm{E})^{-1}\mathrm{F}x_k + (\mathrm{D}-\mathrm{E})^{-1}b$$

En reportant dans (**1**), on trouve

$$x_{k+1} = (\mathrm{D}-\mathrm{F})^{-1}\mathrm{E}(\mathrm{D}-\mathrm{E})^{-1}\mathrm{F}x_k + (\mathrm{D}-\mathrm{F})^{-1}(\mathrm{D}-\mathrm{E})^{-1}b + (\mathrm{D}-\mathrm{F})^{-1}b$$

On en déduit ainsi

$$\begin{cases} \mathrm{G} = (\mathrm{D}-\mathrm{F})^{-1}\mathrm{E}(\mathrm{D}-\mathrm{E})^{-1}\mathrm{F} \\ \mathrm{H} = (\mathrm{D}-\mathrm{F})^{-1}(\mathrm{D}-\mathrm{E})^{-1}b + (\mathrm{D}-\mathrm{F})^{-1}b \end{cases}$$

$\boxed{\textbf{7.c}}$ Comme on l'a vu à la question **(a)**, la diagonale de A est constituée d'éléments strictement positifs. La notation $\mathrm{D}^{-\frac{1}{2}}$ a donc bien un sens. À présent, martyrisons un peu G :

$$\begin{aligned}
\mathrm{G} = &\ (\mathrm{D}-\mathrm{F})^{-1}\mathrm{E}(\mathrm{D}-\mathrm{E})^{-1}\mathrm{F} \\
= &\ \left(\mathrm{D}^{1/2}\mathrm{D}^{1/2} - \mathrm{D}^{1/2}\mathrm{D}^{-1/2}\mathrm{F}\mathrm{D}^{-1/2}\mathrm{D}^{1/2}\right)^{-1}\mathrm{E} \\
&\ \times \left(\mathrm{D}^{1/2}\mathrm{D}^{1/2} - \mathrm{D}^{1/2}\mathrm{D}^{-1/2}\mathrm{F}\mathrm{D}^{-1/2}\mathrm{D}^{1/2}\right)^{-1}\mathrm{F} \\
= &\ \left(\mathrm{D}^{1/2}(\mathrm{Id} - \mathrm{D}^{-1/2}\mathrm{F}\mathrm{D}^{-1/2})\mathrm{D}^{1/2}\right)^{-1}\mathrm{E} \\
&\ \times \left(\mathrm{D}^{1/2}(\mathrm{Id} - \mathrm{D}^{-1/2}\mathrm{F}\mathrm{D}^{-1/2})\mathrm{D}^{1/2}\right)\mathrm{F} \\
= &\ \mathrm{D}^{-1/2}\,(\mathrm{Id} - \mathrm{U})^{-1}\,\mathrm{D}^{-1/2}\mathrm{E}\mathrm{D}^{-1/2}\,(\mathrm{Id} - \mathrm{L})^{-1}\,\mathrm{D}^{-1/2}\mathrm{F}\mathrm{D}^{-1/2}\mathrm{D}^{1/2} \\
= &\ \mathrm{D}^{-1/2}\,(\mathrm{Id} - \mathrm{U})^{-1}\,\mathrm{L}\,(\mathrm{Id} - \mathrm{L})^{-1}\,\mathrm{U}\mathrm{D}^{1/2}
\end{aligned}$$

Après avoir bien souffert, G est donc bien semblable à $(\mathrm{Id} - \mathrm{U})^{-1}\,\mathrm{L}\,(\mathrm{Id} - \mathrm{L})^{-1}\,\mathrm{U}$.

$\boxed{\textbf{7.d}}$ Les matrices L et $(\mathrm{Id} - \mathrm{L})^{-1}$ commutent, donc G est en fait semblable à $(\mathrm{Id} - \mathrm{U})^{-1}\,(\mathrm{Id} - \mathrm{L})^{-1}\,\mathrm{L}\mathrm{U} = (\mathrm{Id} - \mathrm{U})^{-1}\left(\mathrm{Id} - {}^t\mathrm{U}\right)^{-1}\,{}^t\mathrm{U}\,\mathrm{U}$. Considérons λ une valeur propre de $(\mathrm{Id} - \mathrm{U})^{-1}\left(\mathrm{Id} - {}^t\mathrm{U}\right)^{-1}\,{}^t\mathrm{U}\,\mathrm{U}$ et X un vecteur propre associé à λ. On a alors

$$(\mathrm{Id} - \mathrm{U})^{-1}\left(\mathrm{Id} - {}^t\mathrm{U}\right)^{-1}\,{}^t\mathrm{U}\,\mathrm{U}\mathrm{X} = \lambda\mathrm{X}$$

d'où

$${}^t\mathrm{U}\,\mathrm{U}\mathrm{X} = \lambda\left(\mathrm{Id} - {}^t\mathrm{U}\right)(\mathrm{Id} - \mathrm{U})\,\mathrm{X}$$

On multiplie scalairement par X, pour trouver, après calcul

$$\|\mathrm{U}\mathrm{X}\|_2^2 = \lambda\|\mathrm{X}\|_2^2 - 2\lambda(\mathrm{U}\mathrm{X} \mid \mathrm{X}) + \lambda\|\mathrm{X}\|_2^2 \qquad (\mathbf{3})$$

ce qui amène à

$$\|\mathrm{U}\mathrm{X}\|_2^2 = \lambda\|\mathrm{U}\mathrm{X} - \mathrm{X}\|_2^2$$

et on en déduit que $\lambda \in \mathbb{R}_+$. En reprenant (**3**), on trouve

$$\begin{aligned}
(1-\lambda)\|\mathrm{U}\mathrm{X}\|_2^2 &= \|\mathrm{X}\|_2^2 - 2(\mathrm{U}\mathrm{X} \mid \mathrm{X}) \\
&= \left(\left(\mathrm{Id} - {}^t\mathrm{U} - \mathrm{U}\right)\mathrm{X} \mid \mathrm{X}\right) \\
&= {}^t\mathrm{X}\left(\mathrm{Id} - {}^t\mathrm{U} - \mathrm{U}\right)\mathrm{X} \\
&= {}^t\mathrm{X}\left(\mathrm{Id} - \mathrm{D}^{-\frac{1}{2}}\,{}^t\mathrm{E}\,\mathrm{D}^{-\frac{1}{2}} - \mathrm{D}^{-\frac{1}{2}}\mathrm{E}\mathrm{D}^{-\frac{1}{2}}\right)\mathrm{X} \\
&= {}^t\left(\mathrm{D}^{-\frac{1}{2}}\mathrm{X}\right)\left(\mathrm{D} - {}^t\mathrm{E} - \mathrm{E}\right)\mathrm{D}^{-\frac{1}{2}}\mathrm{X} \\
&= {}^t\left(\mathrm{D}^{-\frac{1}{2}}\mathrm{X}\right)\mathrm{A}\left(\mathrm{D}^{-\frac{1}{2}}\mathrm{X}\right) > 0
\end{aligned}$$

car A est symétrique définie positive. On en déduit que $1-\lambda > 0$, donc que $0 < \lambda < 1$. On en déduit que $\rho(\lambda) < 1$, donc que la méthode est convergente.

8. Méthode des directions alternées

On désire résoudre le système $(\mathrm{H}+\mathrm{V})x = b$, où H est une matrice symétrique définie positive, et où V est une matrice semi définie positive. Pour cela, on utilise la méthode itérative

$$\begin{cases} u_0 \in \mathbb{R}^n \quad r > 0 & \\ (\mathrm{H} + r\,\mathrm{Id}\,)u_{k+\frac{1}{2}} = (r\,\mathrm{Id}\, - \mathrm{V})u_k + b & (\mathbf{1}) \\ (\mathrm{V} + r\,\mathrm{Id}\,)u_{k+1} = (r\,\mathrm{Id}\, - \mathrm{H})u_{k+\frac{1}{2}} + b & (\mathbf{2}) \end{cases}$$

$\boxed{\textbf{8.a}}$ Les matrices $\mathrm{H}+r\,\mathrm{Id}$ et $\mathrm{V}+r\,\mathrm{Id}$ sont inversibles, si et seulement si $-r$ n'est pas valeur propre de ces matrices. Or H et V sont symétriques positives ; leurs valeurs propres sont donc positives, et pour tout $r > 0$, $\mathrm{H}+r\,\mathrm{Id}$ et $\mathrm{V}+r\,\mathrm{Id}$ sont inversibles. La méthode est donc bien définie.

Supposons que la suite $(u_k)_{k\in\mathbb{N}}$ soit convergente et notons u sa limite. Alors d'après l'équation $(\mathbf{1})$, la suite $\left(u_{k+\frac{1}{2}}\right)_{k\in\mathbb{N}}$ est également convergente, vers une limite que l'on note $\bar{u}$. En passant à la limite dans $(\mathbf{1})$ et $(\mathbf{2})$, on trouve

$$\begin{cases} (\mathrm{H} + r\,\mathrm{Id}\,)\bar{u} = (r\,\mathrm{Id}\, - \mathrm{V})u + b \\ (\mathrm{V} + r\,\mathrm{Id}\,)u = (r\,\mathrm{Id}\, - \mathrm{H})\bar{u} + b \end{cases}$$

En retranchant la seconde équation trouvée à la première, on trouve

$$\mathrm{H}\bar{u} + r\,\mathrm{Id}\;\bar{u} - \mathrm{V}u - ru = ru - \mathrm{V}u - r\bar{u} + \mathrm{H}\bar{u}$$

d'où

$$2r(\bar{u} - u) = 0$$

d'où l'on tire $\bar{u} = u$. En passant à la limite dans $(\mathbf{1})$, on trouve

$$(\mathrm{H} + r\,\mathrm{Id}\,)u = (r\,\mathrm{Id}\, - \mathrm{V})u + b$$

d'où l'on déduit $(\mathrm{H} + \mathrm{V})u = b$, et la méthode est donc consistante.

$\boxed{\textbf{8.b}}$ L'équation $(\mathbf{1})$ donne

$$u_{k+\frac{1}{2}} = (\mathrm{H} + r\,\mathrm{Id}\,)^{-1}(r\,\mathrm{Id}\, - \mathrm{V})u_k + (\mathrm{H} + r\,\mathrm{Id}\,)^{-1}b$$

En utilisant l'équation $(\mathbf{2})$, on trouve

$$u_{k+1} = \mathrm{T}_r u_k + \mathrm{S}_r$$

avec

$$\begin{cases} \mathrm{T}_r = (\mathrm{V} + r\,\mathrm{Id}\,)^{-1}(r\,\mathrm{Id}\, - \mathrm{H})(\mathrm{H} + r\,\mathrm{Id}\,)^{-1}(r\,\mathrm{Id}\, - \mathrm{V}) \\ \mathrm{S}_r = (\mathrm{V} + r\,\mathrm{Id}\,)^{-1}(r\,\mathrm{Id}\, - \mathrm{H})b + b \end{cases}$$

On a

$$\mathrm{T}_r = (\mathrm{V} + r\,\mathrm{Id}\,)^{-1}\widetilde{\mathrm{T}}_r(\mathrm{V} + r\,\mathrm{Id}\,)$$

avec

$$\widetilde{\mathrm{T}}_r = (r\,\mathrm{Id}\, - \mathrm{H})(\mathrm{H} + r\,\mathrm{Id}\,)^{-1}(r\,\mathrm{Id}\, - \mathrm{V})(\mathrm{V} + r\,\mathrm{Id}\,)^{-1}$$

Comme dans cet exercice on joue avec des matrices symétriques, on choisit de trouver un majorant de $\|\widetilde{\mathrm{T}}_r\|_2$. Déjà, on a

$$\|\widetilde{\mathrm{T}}_r\|_2 \leqslant \|(r\,\mathrm{Id} - \mathrm{H})(\mathrm{H} + r\,\mathrm{Id})^{-1}\|_2 \; \|(r\,\mathrm{Id} - \mathrm{V})(\mathrm{V} + r\,\mathrm{Id})^{-1}\|_2$$

Il reste à évaluer $\|(r\,\mathrm{Id} - \mathrm{H})(\mathrm{H}+r\,\mathrm{Id})^{-1}\|_2$ et $\|(r\,\mathrm{Id} - \mathrm{V})(\mathrm{V}+r\,\mathrm{Id})^{-1}\|_2$. Comme ces deux matrices sont symétriques (car ce sont des produits de matrices symétriques qui commutent entre elles), il suffit de connaître leur spectre pour connaître leur $\|\cdot\|_2$.

$$\begin{aligned}\det\big((r\,\mathrm{Id} - \mathrm{V})(\mathrm{V} + r\,\mathrm{Id})^{-1} - \lambda\,\mathrm{Id}\big) &= \det(r\,\mathrm{Id} - \mathrm{V} - \lambda(\mathrm{V} + r\,\mathrm{Id}))\\ &\quad \times \det((\mathrm{V} + r\,\mathrm{Id})^{-1})\\ &= \det(r(1-\lambda)\,\mathrm{Id} - (\lambda+1)\mathrm{V})\\ &\quad \times \det((\mathrm{V} + r\,\mathrm{Id})^{-1})\end{aligned}$$

Ce dernier déterminant est nul, si et seulement si $r\,\dfrac{1-\lambda}{1+\lambda}$ est valeur propre de V. Comme V est symétrique et positive, les valeurs propres de V sont toutes positives. On en déduit que $\det\big((r\,\mathrm{Id} - \mathrm{V})(\mathrm{V} + r\,\mathrm{Id})^{-1} - \lambda\,\mathrm{Id}\big)$ est non nul si $r\,\dfrac{1-\lambda}{1+\lambda} < 0$, c'est à dire si $\lambda \in \,]-\infty\,;-1\,[\cup]\,1\,;+\infty\,[$. Les valeurs propres de $(r\,\mathrm{Id} - \mathrm{V})(\mathrm{V}+r\,\mathrm{Id})^{-1}$ sont donc dans la boule de centre 0 et de rayon 1. On a donc

$$\|(r\,\mathrm{Id} - \mathrm{V})(\mathrm{V} + r\,\mathrm{Id})^{-1}\|_2 \leqslant 1$$

Le même calcul donne

$$\begin{aligned}\det\big((r\,\mathrm{Id} - \mathrm{H})(\mathrm{H} + r\,\mathrm{Id})^{-1} - \lambda\,\mathrm{Id}\big) &= \det(r(1-\lambda)\,\mathrm{Id} - (\lambda+1)\mathrm{H})\\ &\quad \times \det((\mathrm{H} + r\,\mathrm{Id})^{-1})\end{aligned}$$

Comme H est symétrique définie positive, ses valeurs propres sont strictement positives. On en déduit que si λ est valeur propre de $(r\,\mathrm{Id} - \mathrm{H})(\mathrm{H}+r\,\mathrm{Id})^{-1}$, alors $r\,\dfrac{1-\lambda}{1+\lambda}$ est valeur propre de H, donc est strictement positif. Il vient alors $\lambda \in \,]-1\,;1\,[$, et donc

$$\|(r\,\mathrm{Id} - \mathrm{H})(\mathrm{H} + r\,\mathrm{Id})^{-1}\|_2 < 1$$

on en déduit immédiatement

$$\|\widetilde{\mathrm{T}}_r\|_2 < 1$$

et donc $\widetilde{\mathrm{T}}_r$ est de rayon spectral strictement inférieur à 1. La méthode itérative converge donc quelque soit $r > 0$.

8.c On découpe la matrice du laplacien 2D sous la forme

$$\mathrm{V} = \begin{pmatrix} 2\mathrm{I}_n & -\mathrm{I}_n & 0_n & \dots & \dots & 0_n \\ -\mathrm{I}_n & 2\mathrm{I}_n & -\mathrm{I}_n & \ddots & & \vdots \\ 0_n & -\mathrm{I}_n & 2\mathrm{I}_n & -\mathrm{I}_n & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & 0_n \\ \vdots & & \ddots & -\mathrm{I}_n & 2\mathrm{I}_n & -\mathrm{I}_n \\ 0_n & \dots & \dots & 0_n & -\mathrm{I}_n & 2\mathrm{I}_n \end{pmatrix}$$

$$
\mathrm{H} = \begin{pmatrix}
\bar{\mathrm{A}} & 0_n & 0_n & \dots & \dots & 0_n \\
0_n & \bar{\mathrm{A}} & 0_n & \ddots & & \vdots \\
0_n & 0_n & \bar{\mathrm{A}} & 0_n & \ddots & \vdots \\
\vdots & \ddots & \ddots & \ddots & \ddots & 0_n \\
\vdots & & \ddots & 0_n & \bar{\mathrm{A}} & 0_n \\
0_n & \dots & \dots & 0_n & 0_n & \bar{\mathrm{A}}
\end{pmatrix}
$$

avec $\bar{\mathrm{A}}$ la matrice du laplacien 1D. Les matrices H et V sont définies positives (attention, si on devait le montrer, il faudrait revenir à la définition, car ici, les matrices H et V sont à forte dominante diagonale, mais pas irréductibles). Pour calculer $u^{m+1/2}$ à partir de u^m, il suffit de résoudre n fois le système du Laplacien 1D (sauf que le coefficient de la diagonale est $r+2$ au lieu de 2). Pour le calcul de u^m à partir de $u^{m+1/2}$, c'est exactement le même système que l'on doit résoudre (pour en être vraiment convaincu, il suffit de réordonner les variables dans l'ordre suivant :$1, n+1, 2n+1, \dots, (n-1)n+1, 2, n+2, 2n+2\dots$)

> **Remarque**
> Le terme de « méthode des directions alternées » est justifié ici par le fait que résoudre un système issu de la discrétisation par différences finies du Laplacien 2D peut être remplacé par la résolution du Laplacien 1D alternativement dans une et dans l'autre des deux directions.

9. Méthode de Richardson

Soit A une matrice symétrique définie positive. On considère la méthode

$$
\begin{cases} x_0 \in \mathbb{R}^n \\ x_{k+1} = x_k + \alpha r_k \quad \text{avec} \quad r_k = b - \mathrm{A}x_k \end{cases}
$$

9.a On a

$$
\begin{aligned}
x_{k+1} &= x_k + \alpha r_k \\
&= x_k + \alpha(b - \mathrm{A}x_k) \\
&= (\mathrm{Id} - \alpha \mathrm{A})x_k + \alpha b
\end{aligned}
$$

Comme A est symétrique définie positive, A est diagonalisable, donc la matrice $(\mathrm{Id} - \alpha\mathrm{A})$ est également diagonalisable et ses valeurs propres sont les $1-\alpha\lambda_i$, où les λ_i sont les valeurs propres de A. De plus, on a

$$
\lambda_i \leqslant \lambda_{\max} \quad \Longrightarrow \quad \alpha\lambda_i \leqslant \alpha\lambda_{\max} < 2
$$

on en déduit

$$
1 - \alpha\lambda_i > -1
$$

Et comme A est définie positive, on a

$$
1 - \alpha\lambda_i < 1
$$

On en déduit que $\mathrm{S}p(\mathrm{Id} - \alpha\mathrm{A}) \subset \,]-1\,;1\,[$. La matrice $(\mathrm{Id} - \alpha\mathrm{A})$ est donc de rayon spectral strictement inférieur à 1, donc la méthode est convergente.

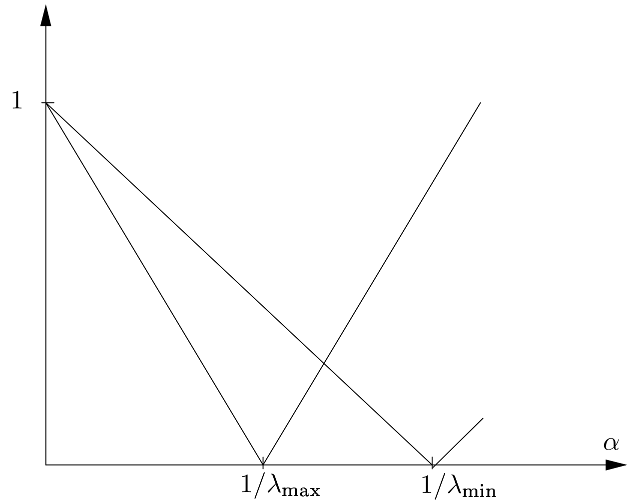
$\boxed{\textbf{9.b}}$ Comme pour tout $\lambda_i \in \mathrm{S}p(\mathrm{Id} - \alpha\mathrm{A})$, on a

$$\lambda_{\min} \leqslant \lambda_i \leqslant \lambda_{\max}$$

On en déduit

$$\rho(\mathrm{Id} - \alpha\mathrm{A}) = \max\left(|1 - \alpha\lambda_{\max}|, |1 - \alpha\lambda_{\min}|\right)$$

En traçant les fonctions $\alpha \mapsto |1 - \alpha\lambda_{\max}|$ et $\alpha \mapsto |1 - \alpha\lambda_{\min}|$ entre $\left[0\,; \dfrac{2}{\lambda_{\max}}\right[$, on trouve le dessin suivant :



Le maximum de ces deux fonctions atteint son minimum à l'intersection de $\alpha \mapsto |1 - \alpha\lambda_{\max}|$ dans sa partie croissante ($\alpha \geqslant \dfrac{1}{\lambda_{\max}}$) et de $\alpha \mapsto |1 - \alpha\lambda_{\min}|$ dans sa partie décroissante ($\alpha \leqslant \dfrac{1}{\lambda_{\min}}$). α_{opt} vérifie donc

$$\alpha_{\mathrm{opt}}\lambda_{\max} - 1 = 1 - \alpha_{\mathrm{opt}}\lambda_{\min}$$

On en déduit

$$\alpha_{\mathrm{opt}} = \frac{1}{\lambda_{\max} + \lambda_{\min}}$$

Et on a alors

$$\rho(\mathrm{Id} - \alpha\mathrm{A}) = \frac{\lambda_{\max} - \lambda_{\min}}{\lambda_{\max} + \lambda_{\min}}$$

ou encore

$$\rho(\mathrm{Id} - \alpha\mathrm{A}) = \frac{\dfrac{\lambda_{\max}}{\lambda_{\min}} - 1}{\dfrac{\lambda_{\max}}{\lambda_{\min}} + 1} = \frac{\mathrm{Cond}_2(\mathrm{A}) - 1}{\mathrm{Cond}_2(\mathrm{A}) + 1}$$

10. Méthode des gradients

A est une matrice symétrique définie positive, b un vecteur de $\mathbb{R}^n$, et $\bar{x}$ est tel que $\mathrm{A}\bar{x} = b$.

10.a La fonctionnelle

$$\mathrm{J} : x \mapsto \mathrm{J}(x) = (\mathrm{A}x \mid x) - 2(b \mid x)$$

est une fonctionnelle strictement convexe (car sa dérivée seconde, A, est définie positive). Elle admet donc un unique minimum sur $\mathbb{R}^n$, caractérisé par $\mathrm{J}'(\bar{x}) = 0$, c'est à dire $\mathrm{A}\bar{x} = 0$. Par ailleurs on a

$$\begin{aligned}
\mathrm{E}(x) &= (\mathrm{A}(x - \bar{x}) \mid x - \bar{x}) \\
&= (\mathrm{A}x \mid x) - (\mathrm{A}x \mid \bar{x}) - (\mathrm{A}\bar{x} \mid x) + (\mathrm{A}\bar{x} \mid \bar{x}) \\
&= (\mathrm{A}x \mid x) - 2(b \mid x) + (\mathrm{A}\bar{x} \mid \bar{x}) \\
&= \mathrm{J}(x) + (b \mid \mathrm{A}^{-1}b)
\end{aligned}$$

La fonctionnelle E n'est donc que la fonctionnelle J translatée, et elle a donc les mêmes propriétés de minimum. $\bar{x}$ est donc également l'unique minimum de E.

10.b

$$\begin{aligned}
\mathrm{E}(x) &= (\mathrm{A}(x - \bar{x}) \mid x - \bar{x}) \\
&= (\mathrm{A}x - \mathrm{A}\bar{x} \mid \mathrm{A}^{-1}(\mathrm{A}x - \mathrm{A}\bar{x})) \\
&= (\mathrm{A}x - b \mid \mathrm{A}^{-1}(\mathrm{A}x - b)) \\
&= (r(x) \mid \mathrm{A}^{-1}r(x))
\end{aligned}$$

10.c On cherche à minimiser $\mathrm{E}(x_{k+1})$, où x_{k+1} est défini par

$$x_{k+1} = x_k + \alpha_k p_k$$

$\mathrm{E}(x_{k+1})$ est une fonction d'une variable réelle α_k, elle est donc minimale si et seulement si sa dérivée par rapport à α_k est nulle.

$$\begin{aligned}
\mathrm{E}(x_{k+1}) &= (\mathrm{A}(x_k + \alpha_k p_k - \bar{x}) \mid x_k + \alpha_k p_k - \bar{x}) \\
&= (\mathrm{A}(x_k - \bar{x}) \mid x_k - \bar{x}) + \alpha_k(\mathrm{A}p_k \mid x_k - \bar{x}) + \alpha_k(x_k - \bar{x} \mid \mathrm{A}p_k) \\
&\quad + {\alpha_k}^2(\mathrm{A}p_k \mid p_k) \\
&= \mathrm{E}_k + 2\alpha_k(\mathrm{A}p_k \mid x_k - \bar{x}) + {\alpha_k}^2(\mathrm{A}p_k \mid p_k)
\end{aligned}$$

La dérivée de E_{k+1} par rapport à α_k s'annule donc en

$$\alpha_k = -\frac{(\mathrm{A}p_k \mid x_k - \bar{x})}{(\mathrm{A}p_k \mid p_k)} = \frac{(\mathrm{A}p_k \mid \bar{x} - x_k)}{(\mathrm{A}p_k \mid p_k)} = \frac{(p_k \mid \mathrm{A}\bar{x} - \mathrm{A}x_k)}{(\mathrm{A}p_k \mid p_k)} = \frac{(p_k \mid r_k)}{(\mathrm{A}p_k \mid p_k)}$$

On a alors

$$r_{k+1} = b - \mathrm{A}x_{k+1} = b - \mathrm{A}(x_k + \alpha_k p_k) = b - \mathrm{A}x_k - \alpha_k \mathrm{A}p_k = r_k - \alpha_k \mathrm{A}p_k$$

et

$$\begin{aligned}
(p_k \mid r_{k+1}) &= (p_k \mid r_k) - \alpha_k(\mathrm{A}p_k \mid p_k) \\
&= (p_k \mid r_k) - \frac{(p_k \mid r_k)}{(\mathrm{A}p_k \mid p_k)}(p_k \mid \mathrm{A}p_k) \\
&= 0
\end{aligned}$$

Enfin on a

$$\begin{aligned}
\mathrm{E}_{k+1} &= (r_{k+1} \mid \mathrm{A}^{-1} r_{k+1}) \\
&= (r_{k+1} \mid \mathrm{A}^{-1}(r_k - \alpha_k \mathrm{A} p_k)) \\
&= (r_{k+1} \mid \mathrm{A}^{-1} r_k) - \alpha_k (r_{k+1} \mid \mathrm{A}^{-1}\mathrm{A} p_k) \\
&= (r_{k+1} \mid \mathrm{A}^{-1} r_k) \\
&= (r_k - \alpha_k \mathrm{A} p_k \mid \mathrm{A}^{-1} r_k) \\
&= (r_k \mid \mathrm{A}^{-1} r_k) - \alpha_k (\mathrm{A} p_k \mid \mathrm{A}^{-1} r_k) \\
&= \mathrm{E}_k - \alpha_k (p_k \mid r_k) \\
&= \mathrm{E}_k - \frac{(p_k \mid r_k)}{(\mathrm{A} p_k \mid p_k)} (p_k \mid r_k) \\
&= \mathrm{E}_k - \frac{(p_k \mid r_k)^2}{(\mathrm{A} p_k \mid p_k)(\mathrm{A}^{-1} r_k \mid r_k)} (\mathrm{A}^{-1} r_k \mid r_k) \\
&= \left(1 - \frac{(p_k \mid r_k)^2}{(\mathrm{A} p_k \mid p_k)(\mathrm{A}^{-1} r_k \mid r_k)}\right) \mathrm{E}_k
\end{aligned}$$

10.d D'après l'inégalité de Cauchy–Schwarz

$$\left|(\mathrm{A} p_k \mid p_k)(\mathrm{A}^{-1} r_k \mid r_k)\right| \leqslant \|\mathrm{A}^{-1} r_k\|_2 \|r_k\|_2 \|\mathrm{A} p_k\|_2 \|p_k\|_2 \leqslant \mathrm{Cond}_2(\mathrm{A}) \|r_k\|_2^2 \|p_k\|_2^2$$

On en déduit

$$\frac{(p_k \mid r_k)^2}{(\mathrm{A} p_k \mid p_k)(\mathrm{A}^{-1} r_k \mid r_k)} \leqslant \frac{(p_k \mid r_k)^2}{\mathrm{Cond}_2(\mathrm{A}) \|r_k\|_2^2 \|p_k\|_2^2} \leqslant \frac{1}{\mathrm{Cond}_2(\mathrm{A})}$$

d'où

$$\mathrm{E}_{k+1} \leqslant \left(1 - \frac{1}{\mathrm{Cond}_2(\mathrm{A})}\right) \mathrm{E}_k$$

10.e D'après la question (**d**), on a

$$\mathrm{E}_k \leqslant \left(\frac{\mathrm{Cond}_2(\mathrm{A}) - 1}{\mathrm{Cond}_2(\mathrm{A})}\right)^k \mathrm{E}_0$$

Le membre de droite est une suite géométrique de raison inférieure à 1, donc elle converge vers 0, et la méthode de gradient à pas optimal converge.

10.f

$$\begin{aligned}
(\mathrm{A} p_{k-1} \mid p_k) &= (\mathrm{A} p_{k-1} \mid r_k + \beta_k p_{k-1}) \\
&= (\mathrm{A} p_{k-1} \mid r_k) + \beta_k (\mathrm{A} p_{k-1} \mid p_{k-1}) \\
&= (\mathrm{A} p_{k-1} \mid r_k) - \frac{(\mathrm{A} p_{k-1} \mid r_k)}{(\mathrm{A} p_{k-1} \mid p_{k-1})} (\mathrm{A} p_{k-1} \mid p_{k-1}) \\
&= 0
\end{aligned}$$

$$\begin{aligned}
(r_{k+1} \mid r_k) &= (r_k - \alpha_k \mathrm{A} p_k \mid r_k) \\
&= \|r_k\|_2^2 - \alpha_k (\mathrm{A} p_k \mid r_k) \\
&= \|r_k\|_2^2 - \alpha_k (\mathrm{A} p_k \mid p_k - \beta_k p_{k-1}) \\
&= \|r_k\|_2^2 - \alpha_k (\mathrm{A} p_k \mid p_k) + \alpha_k \beta_k (\mathrm{A} p_k \mid p_{k-1}) \\
&= \|r_k\|_2^2 - \alpha_k (\mathrm{A} p_k \mid p_k) \\
&= \|r_k\|_2^2 - (r_k \mid p_k)
\end{aligned}$$

Par ailleurs, comme on l'a vu à la question (**d**), on a $(r_k \mid p_{k-1}) = 0$ donc $(p_k \mid r_k) = \|r_k\|_2^2$. On en déduit que $(r_{k+1} \mid r_k) = 0$.

D'après la question (**d**), on a

$$\mathrm{A}p_{k-1} = \frac{1}{\alpha_{k-1}}\,(r_{k-1} - r_k)$$

ce qui permet de calculer $(\mathrm{A}p_{k-1} \mid p_{k-1})$ et $(\mathrm{A}p_{k-1} \mid r_k)$:

$$(\mathrm{A}p_{k-1} \mid p_{k-1}) = \frac{1}{\alpha_{k-1}}(\mathrm{A}p_{k-1} \mid r_{k-1} - r_k) = \frac{1}{\alpha_{k-1}}(\mathrm{A}p_{k-1} \mid r_{k-1}) = \frac{\|r_{k-1}\|_2^2}{\alpha_{k-1}}$$

$$(\mathrm{A}p_{k-1} \mid r_k) = \frac{-1}{\alpha_{k-1}}(r_{k-1} - r_k \mid r_k) = -\frac{\|r_k\|_2^2}{\alpha_{k-1}}$$

ce qui donne immédiatement

$$\beta_k = \frac{\|r_k\|_2^2}{\|r_{k-1}\|_2^2}$$

Pour l'algorithme, on précise en plus que $\beta_0 = 0$, et on se donne un vecteur x_0. La mise en place de l'algorithme ne pose pas de grande difficulté, si ce n'est qu'il faut bien faire attention à l'ordre dans lequel on calcule les quantités intermédiaires p_k, α_k et r_k. On suppose que l'on a déjà programmé des fonctions de multiplication matrice–vecteur et une fonction `produitscalaire` qui calcule le produit scalaire de deux vecteurs.

```
#On choisit d'initialiser  l'algorithme à b (i.e. x0=0)
#Attention! les quantités p,b,r,x sont des vecteurs
#tandis que alpha, beta et N sont des scalaires

p=b
r=b
x=0
beta=0
alpha=produitscalaire(p,p)/produitscalaire(A*p,p)


#On calcule la norme 2 du résidu au carré
N=produitscalaire(r,r)

TANT QUE (N>erreur) FAIRE
     x=x+alpha*p
     r=b-A*x
     N=produit_scalaire(r,r)
     beta=-produitscalaire(A*p,r)/produitscalaire(A*r,r)
     p=r+beta*p
     alpha=produitscalaire(p,r)/produitscalaire(A*p,p)
FIN TANT QUE
```

11. Méthode de la puissance

A est une matrice diagonalisable. le complexe λ_1 est la valeur propre de plus grand module de A, et on suppose que λ_1 est simple. On se donne $q_0 \in \mathbb{C}^n$, et on définit les suites $(x_k)_{k\in\mathbb{N}}$ et $(q_k)_{k\in\mathbb{N}}$ de $\mathbb{C}^n$ de la manière suivante :

$$x_k = \mathrm{A}q_{k-1} \qquad \text{et} \qquad q_k = \frac{x_k}{\|x_k\|}$$

11.a On raisonne par récurrence

- <u>$k = 1$</u>

  On a $x_1 = \mathrm{A}q_0$ donc $q_1 = \dfrac{x_1}{\|x_1\|} = \dfrac{\mathrm{A}q_0}{\|\mathrm{A}q_0\|}$, donc le résultat est vrai pour $k = 1$.
- <u>$k \implies k+1$</u>

  On suppose que $q_k = \dfrac{\mathrm{A}^k q_0}{\|\mathrm{A}^k q_0\|}$. On a alors

$$q_{k+1} = \frac{x_{k+1}}{\|x_{k+1}\|} = \frac{\mathrm{A}q_k}{\|\mathrm{A}q_k\|} = \frac{\mathrm{A}^{k+1}q_0}{\|\mathrm{A}^k q_0\|}\frac{1}{\left\|\dfrac{\mathrm{A}^{k+1}q_0}{\|\mathrm{A}^k q_0\|}\right\|} = \frac{\mathrm{A}^{k+1}q_0}{\|\mathrm{A}^{k+1}q_0\|}$$

  D'où le résultat au rang $k+1$.
- <u>Conclusion</u>

$$\forall k \in \mathbb{N}^\star \qquad q_k = \frac{\mathrm{A}^k q_0}{\|\mathrm{A}^k q_0\|}$$

11.b On cherche la limite de

$$\left(\frac{\overline{\lambda_1}}{|\lambda_1|}\right)^k q_k = \left(\frac{1}{\lambda_1}\right)^k |\lambda_1|^k \frac{\mathrm{A}^k q_0}{\|\mathrm{A}^k q_0\|} = \frac{\dfrac{\mathrm{A}^k q_0}{\lambda_1^k}}{\left\|\dfrac{\mathrm{A}^k q_0}{\lambda_1^k}\right\|}$$

Étudions $\dfrac{\mathrm{A}^k q_0}{\lambda_1^k}$. Pour celà, on décompose q_0 dans la base $(u_1 \dots u_n)$ de vecteurs propres de A (où u_1 est associé à la valeur propre λ_1) sous la forme

$$q_0 = \sum_{i=1}^{n} \alpha^{(i)} u_i$$

$$\begin{aligned}
\frac{\mathrm{A}^k q_0}{\lambda_1^k} &= \frac{1}{\lambda_1^k}\mathrm{A}^k q_0 \\
&= \frac{1}{\lambda_1^k}\sum_{i=1}^{n}\alpha^{(i)}\mathrm{A}^k u_i \\
&= \sum_{i=1}^{n}\alpha^{(i)}\left(\frac{\lambda_i}{\lambda_1}\right)^k u_i \\
&= \alpha^{(1)}u_1 + e_k
\end{aligned}$$

avec

$$e_k = \sum_{i=2}^{n} \alpha^{(i)} \left(\frac{\lambda_i}{\lambda_1}\right)^k u_i$$

La suite de vecteurs $(e_k)_{k\in\mathbb{N}}$ tend vers 0 (car $|\lambda_i| < |\lambda_1|$ pour tout i). On peut donc écrire

$$\left(\frac{\overline{\lambda}_1}{|\lambda_1|}\right) q_k = \frac{\alpha^{(1)} u_1 + e_k}{\|\alpha^{(1)} u_1 + e_k\|}$$

Comme $\alpha^{(1)}$ est non nul, le membre de droite admet une limite qui est u_1 normalisé.

11.c D'après la question précédente

$$\lim_{k\to+\infty} \|\mathrm{A}q_k - \lambda_1 q_k\| = 0$$

d'après l'inégalité triangulaire

$$\|\mathrm{A}q_k - \lambda_1 q_k\| \geqslant \Big| \|\mathrm{A}q_k\| - |\lambda_1|\, \|q_k\| \Big| = \Big| \|\mathrm{A}q_k\| - |\lambda_1| \Big|$$

On en déduit

$$\lim_{k\to+\infty} \|\mathrm{A}q_k\| = |\lambda_1|$$

11.d

$$x_{k+1}(j) - \lambda_1 q_k(j) = (\mathrm{A}q_k - \lambda_1 q_k)(j)$$

or d'après la question (**b**)

$$\lim_{k\to+\infty} \|\mathrm{A}q_k - \lambda_1 q_k\| = 0$$

On en déduit

$$\lim_{k\to+\infty} |(x_{k+1}(j) - \lambda_1 q_k(j))| = 0$$

donc que

$$\lim_{k\to+\infty} \frac{x_{k+1}(j)}{q_k(j)} = \lambda_1$$

11.e Si $(\lambda_1, \dots, \lambda_n)$ sont les valeurs propres de A, alors les valeurs propres de A^{-1} sont $\left(\frac{1}{\lambda_1}, \dots, \frac{1}{\lambda_n}\right)$. D'après les questions précédentes, la méthode décrite par

$$\begin{cases} q_0 \in \mathbb{C}^n \\ \mathrm{A}x_{k+1} = q_k \\ q_k = \dfrac{x_k}{\|x_k\|} \end{cases}$$

est telle que

$$\lim_{k\to+\infty} \|x_k\| = \frac{1}{|\lambda_n|}$$

où λ_n est la valeur propre de module le plus petit de A.

11.f Pour trouver la valeur propre la plus proche de μ, il suffit d'appliquer la méthode à la matrice $\mathrm{A}' = \mathrm{A} - \mu\,\mathrm{Id}$. En effet, si $(\lambda_1, \dots, \lambda_n)$ est l'ensemble des valeurs propres de A, alors les valeurs propres de A' sont les $\lambda_1 - \mu, \dots \lambda_n - \mu$. Ainsi, la valeur propre de plus petit module de A' est telle que $|\lambda_i - \mu|$ soit minimal, c'est donc la valeur propre de A la plus proche de μ.

12. Méthode de Lanczos

A est une matrice symétrique réelle, α_k et β_k sont deux familles de réels, et q_k est une famille de vecteurs, définis par

$$\begin{cases} q_0 = 0 \\ q_1 \quad \text{quelconque, de norme } 1 \\ \mathrm{A}q_k = \beta_{k-1}q_{k-1} + \alpha_k q_k + \beta_k q_{k+1} \end{cases}$$

12.a Supposons que la famille q_k soit orthonormée. Alors en multipliant la relation (**1**) respectivement par q_k et q_{k+1} on trouve

$$\begin{aligned} (\mathrm{A}q_k \mid q_k) &= \beta_{k-1}(q_{k-1} \mid q_k) + \alpha_k(q_k \mid q_k) + \beta_k(q_{k+1} \mid q_k) = \alpha_k \|q_k\|^2 = \alpha_k \\ (\mathrm{A}q_k \mid q_{k+1}) &= \beta_{k-1}(q_{k-1} \mid q_{k+1}) + \alpha_k(q_k \mid q_{k+1}) + \beta_k(q_{k+1} \mid q_{k+1}) = \beta_k \|q_{k+1}\|^2 \\ &= \beta_k \end{aligned}$$

Réciproquement, supposons que

$$\begin{aligned} \alpha_k &= (\mathrm{A}q_k \mid q_k) \\ \beta_k &= (\mathrm{A}q_k \mid q_{k+1}) \end{aligned}$$

Montrons par récurrence sur k que sous les hypothèses précédentes la famille q_k est orthonormée

- $\underline{k = 1}$ Par hypothèse, la famille q_1 est orthonormée.
- $\underline{k \implies k+1}$ On suppose que la famille q_k est orthonormée. Alors sous réserve que $(\mathrm{A}q_k \mid q_{k+1})$ soit non nul, on peut définir

$$q_{k+1} = \frac{\mathrm{A}q_k - \beta_{k-1}q_k - \alpha_k q_k}{(\mathrm{A}q_k \mid q_{k+1})}$$

on a alors

$$\begin{aligned} (q_{k+1} \mid q_k) &= \frac{1}{\beta_k}\left((\mathrm{A}q_k \mid q_k) - \alpha_k(q_k \mid q_k) - \beta_{k-1}(q_{k-1} \mid q_k)\right) \\ &= -\frac{\beta_{k-1}}{\beta_k}(q_{k-1} \mid q_k) \\ &= 0 \end{aligned}$$

car q_{k-1} est orthogonal à q_k. Par ailleurs

$$\begin{aligned} (q_{k+1} \mid q_{k-1}) &= \frac{1}{\beta_k}\left((\mathrm{A}q_k \mid q_{k-1}) - \alpha_k(q_k \mid q_{k-1}) - \beta_{k-1}(q_{k-1} \mid q_{k-1})\right) \\ &= -\frac{\alpha_k}{\beta_k}(q_{k-1} \mid q_k) \\ &= 0 \end{aligned}$$

q_{k+1} est donc orthogonal à q_k et q_{k-1}. Montrons maintenant que q_{k+1} est orthogonal à q_i pour $i \leqslant k-2$.

$$\begin{aligned} (q_{k+1} \mid q_i) &= \frac{1}{\beta_k}\left((\mathrm{A}q_k \mid q_i) - \alpha_k(q_k \mid q_i) - \beta_{k-1}(q_{k-1} \mid q_i)\right) \\ &= \frac{1}{\beta_k}(\mathrm{A}q_k \mid q_i) \\ &= \frac{1}{\beta_k}(q_k \mid \mathrm{A}q_i) \\ &= \frac{1}{\beta_k}(q_k \mid \beta_{i-1}q_{i-1} + \alpha_i q_i + \beta_i q_{i+1}) \\ &= \frac{1}{\beta_k}\left(\beta_{i-1}(q_k \mid q_{i-1}) + \alpha_i(q_k \mid q_i) + \beta_i(q_k \mid q_{i+1})\right) \\ &= 0 \end{aligned}$$

car par hypothèse de récurrence, $(q_k \mid q_i) = 0$ pour $i \leqslant k-1$.

- Conclusion : la famille q_k est orthonormée.

12.b

```
#les q(i) sont des vecteurs tandis que les alpha(i)
#et les beta(i) sont des scalaires
q(0)=0
q(1)=vecteur aléatoire, de norme 1
k=1
alpha(1)=scalaire(A*q(1),q(1))
v=A*q(1) - alpha(1)*q(1)

TANT QUE (norme(v)>0) FAIRE
     beta(k)=norme(v)
     q(k+1)=v/beta(k)
     alpha(k+1)=scalaire(A*q(k+1),q(k+1))
     v= A*q(k+1) - beta(k)*q(k) - alpha(k+1)*q(k+1)
     k=k+1
FIN TANT QUE
```

12.c Dans l'algorithme précédent, il est possible que l'on s'arrête avant d'avoir construit une famille de n vecteurs orthonormés (par exemple, si on parvient à un q_k qui est un vecteur propre de A). Dans ce dernier cas, si on a construit une famille de j vecteurs, alors on poursuit l'algorithme en choisissant un vecteur $j+1$ aléatoire, de norme 1, et appartenant à l'orthogonal de Vect$(q_1, \ldots, q_j)$. Et dans la base q_i, la matrice de A est tridiagonale, avec les α_i sur la diagonale et les β_i sur la sur et la sous–diagonale.