Compte-rendu en *Méthodes numériques de base*
Résultats sur l'identification de conductivité

Aurélien PEPIN, Antonin KLOPP-TOSSER

2 mai 2017

1 Méthode des différences finies

QUESTION 1. L'écriture sous forme matricielle du θ-schéma *(10)* exprime l'itération $k+1$ du vecteur U en fonction de l'itération k. On isole donc, dans le θ-schéma, les termes en $u_i^{(k+1)}$.

$$\begin{aligned} &u_i^{(k+1)} - \mu\theta \left(C_{i+1/2}\, u_{i+1}^{(k+1)} - (C_{i+1/2} + C_{i-1/2})\, u_i^{(k+1)} + C_{i-1/2}\, u_{i-1}^{(k+1)}\right) \\ &= \mu(1-\theta) \left(C_{i+1/2}\, u_{i+1}^{k} - (C_{i+1/2} + C_{i-1/2})\, u_i^{k} + C_{i-1/2}\, u_{i-1}^{k}\right) + u_i^k \end{aligned}$$

$$\begin{aligned} \iff &\left(1 + \theta\mu(C_{i+1/2} + C_{i-1/2})\right)\, u_i^{(k+1)} - \theta\mu(C_{i+1/2}\, u_{i+1}^{(k+1)}) - \theta\mu(C_{i-1/2}\, u_{i-1}^{(k+1)}) \\ &= \left(1 + (\theta-1)\mu(C_{i+1/2} + C_{i-1/2})\right)\, u_i^{k} - (\theta-1)\mu\, C_{i+1/2}\, u_{i+1}^{k} - (\theta-1)\mu\, C_{i-1/2}\, u_{i-1}^{k} \end{aligned}$$

Sachant que $U^{(k)}$ est la matrice-colonne des $u_i^{(k)}$ $\forall i \in [\![1;n]\!]$, on identifie les termes un à un dans le membre de gauche puis dans le membre de droite :

$$\Bigg(I + \theta\mu \underbrace{\begin{bmatrix} C_{1+1/2} + C_{1-1/2} & -C_{1+1/2} & & 0 \\ -C_{2-1/2} & \ddots & \ddots & \\ & \ddots & \ddots & -C_{(n-1)+1/2} \\ 0 & & -C_{n-1/2} & C_{n+1/2} + C_{n-1/2} \end{bmatrix}}_{A}\Bigg) \underbrace{\begin{bmatrix} u_1^{(k+1)} \\ u_2^{(k+1)} \\ \vdots \\ u_n^{(k+1)} \end{bmatrix}}_{U^{(k+1)}}$$

$$\Bigg(I + (\theta-1)\mu \underbrace{\begin{bmatrix} C_{1+1/2} + C_{1-1/2} & -C_{1+1/2} & & 0 \\ -C_{2-1/2} & \ddots & \ddots & \\ & \ddots & \ddots & -C_{(n-1)+1/2} \\ 0 & & -C_{n-1/2} & C_{n+1/2} + C_{n-1/2} \end{bmatrix}}_{A}\Bigg) \underbrace{\begin{bmatrix} u_1^{(k)} \\ u_2^{(k)} \\ \vdots \\ u_n^{(k)} \end{bmatrix}}_{U^{(k)}} + \mu \underbrace{\begin{bmatrix} B_1^{(k)} \\ 0 \\ \vdots \\ 0 \end{bmatrix}}_{B^{(k)}}$$

$B_1^{(k)}$ exprime les conditions aux limites. Pour $i = 1$, le terme $(A)_{i,i-1} = (A)_{1,0}$ n'est pas défini et vaut 0 par convention. Pour conserver l'égalité, on ajoute dans B un coefficient simulé :

$$B_1^{(k)} = C_{1/2}\ (\theta\ u_0(t_k) + (1-\theta)\ u_0(t_k+1))$$

QUESTION 2. Une matrice M est dite définie positive si, pour toute matrice-colonne non nulle x à coefficients réels, ${}^t xMx > 0$. Soit $x = (x_1, x_2, ..., x_n)$ non nul :

$$ {}^t xAx = \sum_{i=1}^{n} (C_{i+1/2} + C_{i-1/2}) x_i^2 \; - \; 2 \sum_{i=1}^{n-1} C_{i+1/2} \; x_i \; x_{i+1} $$

On pose $j = i - 1$ en développant le premier terme :

$$ \begin{aligned} {}^t xAx &= \sum_{i=1}^{n} C_{i+1/2} \; x_i^2 + \sum_{j=0}^{n-1} C_{j+1/2} \; x_{j+1}^2 - 2 \sum_{i=1}^{n-1} C_{i+1/2} \; x_i \; x_{i+1} \\ &= \sum_{i=1}^{n-1} C_{i+1/2} (x_i - x_{i+1})^2 + C_{1/2} \; x_{1/2}^2 + C_{n+1/2} \; x_n^2 \end{aligned} \tag{1} $$

La fonction $C(x)$ est strictement positive par définition donc ${}^t xAx > 0 \; \forall x$ non nul.
La matrice A est bien définie positive.

2 Factorisation de Cholesky dans le cas tridiagonal

QUESTION 3. Phase de factorisation de Cholesky.

Se référer à : `Q3_Factorisation_Cholesky.sce`

La décomposition de Cholesky permet ensuite de résoudre des systèmes linéaires en une phase de descente puis une phase de remontée.

QUESTION 4. Première phase de descente.

Se référer à : `Q4_Descente_Cholesky.sce`

QUESTION 5. Seconde phase de remontée.

Se référer à : `Q5_Remontee_Cholesky.sce`

On peut appliquer dans l'ordre ces trois fonctions pour résoudre l'exemple ci-dessous, noté $MX = B$. Le TD 5 a montré que la matrice M était aussi symétrique définie positive donc elle admet une unique factorisation de Cholesky.

$$ \begin{pmatrix} 1 & -1 & 0 & 0 & 0 \\ -1 & 2 & -1 & 0 & 0 \\ 0 & -1 & 2 & -1 & 0 \\ 0 & 0 & -1 & 2 & -1 \\ 0 & 0 & 0 & -1 & 2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{pmatrix} = \begin{pmatrix} 10 \\ 20 \\ 30 \\ 40 \\ 50 \end{pmatrix} \quad \text{où la solution est } X = \begin{pmatrix} 350 \\ 340 \\ 310 \\ 250 \\ 210 \end{pmatrix} $$

Pour calculer et vérifier cet exemple, cf. `Q5_Exemple_Cholesky.sce`

3 Problème stationnaire

QUESTION 6. L'équation (9) de l'énoncé fournit une appromixation du système (13) :

$$\frac{\partial}{\partial x}[C(x)\frac{\partial u}{\partial x}](x_i,t_k) \approx \frac{C_{i+1/2}\ u_{i+1}^{(k)} - (C_{i+1/2}+C_{i-1/2})\ u_i^{(k)} + C_{i-1/2}\ u_{i-1}^{(k)}}{\delta_x^2}$$

En écrivant ce système sous forme matricielle pour tout $\forall i \in [\![1;n]\!]$, on obtient :

$$\frac{\partial}{\partial x}[C(x)\frac{\partial u}{\partial x}] = \frac{1}{\delta_x^2}(-Au+B) = 0 \iff Au = B$$

où, par identification :

- A est la matrice tridiagonale, symétrique et définie positive de la **question 1** ;
- u est la matrice-colonne des $u_i^{(k)}$ à l'itération k ;
- B est la matrice-colonne des conditions aux limites déterminée de la façon suivante :

Pour i = 1.

$$\begin{aligned}(Au)_1 &= -C_{3/2}\ u_2^{(k)} + (C_{3/2}+C_{1/2})\ u_1^{(k)} \\ &\iff (B)_1 = C_{1/2}\ u_0\end{aligned}$$

Pour i > 1.

$$\begin{aligned}(Au)_i &= -C_{i+1/2}\ u_{i+1}^{(k)} + (C_{i+1/2}+C_{i-1/2})\ u_i^{(k)} - C_{i-1/2}\ u_{i-1}^{(k)} \\ &= \frac{\delta_x^2}{\delta_x^2}(-C_{i+1/2}\ u_{i+1}^{(k)} + (C_{i+1/2}+C_{i-1/2})\ u_i^{(k)} - C_{i-1/2}\ u_{i-1}^{(k)}) \\ &\simeq \delta_x^2 \cdot 0 \\ &\iff (B)_i = 0\ \forall i \in [\![2;n]\!]\end{aligned}$$

La matrice-colonne B est celle de la **question 1** avec la fonction u_0 constante.

Puisque A est définie positive, elle n'admet pas 0 comme valeur propre.
Elle est donc inversible et le système $Au = B$ possède par conséquent une unique solution.

QUESTION 7. Solution exacte du problème stationnaire et convergence.

Se référer à : `Q7_Probleme_stationnaire.sce`

$$\begin{aligned}\frac{\partial}{\partial x}[C(x)\frac{\partial u}{\partial x}] = 0 &\iff -\frac{1}{l}\exp(-\frac{x}{l})\ \frac{\partial u}{\partial x} + \exp(-\frac{x}{l})\ \frac{\partial^2 u}{\partial x^2} = 0 \\ &\iff \frac{\partial^2 u}{\partial x^2} - \frac{1}{l}\frac{\partial u}{\partial x} = 0\end{aligned}$$

donc $u(x)$ est de la forme :

$$u(x) = \alpha + \beta\exp(\frac{x}{l})$$

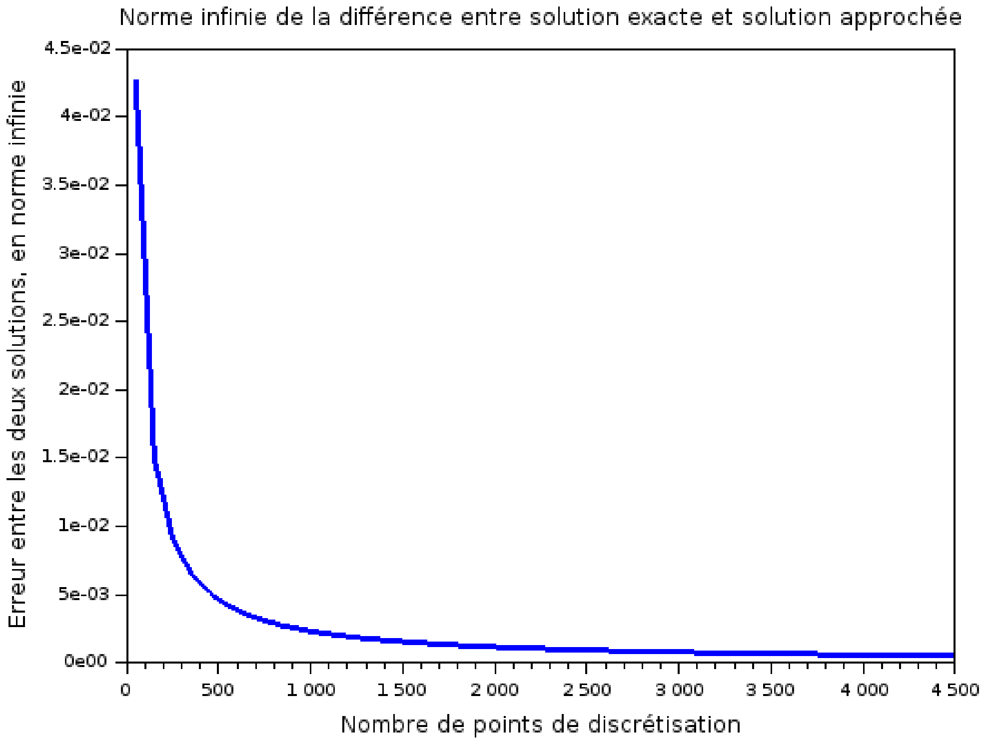
Détermination des constantes α et β. À partir des conditions initiales, on résout :

$$\begin{cases} u(-l) = \alpha + \beta e^{-1} = u_0 = 1 \\ u(l) = \alpha + \beta e = 0 \end{cases} \iff \begin{cases} \beta = \frac{1}{e^{-1}-e} \\ \alpha = -\frac{e}{e^{-1}-e} \end{cases}$$

Finalement, la solution exacte de l'équation différentielle (13) est :

$$u(x) = -\frac{e}{e^{-1} - e} + \frac{1}{e^{-1} - e} \exp(\frac{x}{l}) = \frac{e(e - e^{\frac{x}{l}})}{e^2 - 1}$$

Il est possible de mesurer la précision de la méthode des différences finies en comparant cette solution exacte avec la solution approchée pour différents pas $\delta_x \to 0$.



Plus le nombre de points de discrétisation x_i augmente, plus le pas δ_x tend vers 0. La solution numérique converge vers la solution exacte quand $\delta_x \to 0$.

4 Évolution d'une donnée initiale

QUESTION 8. On cherche à montrer que $\rho(M^{-1}N)$ est strictement inférieur à 1. On applique pour cela à M^{-1} et à N le résultat suivant, démontré plus bas.

> **Proposition**. Soit P un polynôme d'endomorphisme (ou de matrice), soit S une matrice. Si v est vecteur propre de S pour la valeur propre λ, alors il est vecteur propre de $P(S)$ pour la valeur $P(\lambda)$.

Soient P et Q deux polynômes de matrice. Dire en effet que v est vecteur propre de S associé à la valeur propre λ revient à écrire :

$$\begin{aligned}
Sv = \lambda v &\iff S^p\ v = S^{p-1} \cdot Sv = S^{p-1} \cdot \lambda v = \lambda^p v \quad \forall p \in \mathbb{N} \\
&\iff P(S)v = \sum_{k=0}^{p} a_k\ S^k(v) = \sum_{k=0}^{p} a_k\ \lambda^k v = P(\lambda)v \\
&\iff P^{-1}(S)v = \frac{1}{P(\lambda)} v \\
&\iff P^{-1}(S)\ Q(S)v = \frac{Q(\lambda)}{P(\lambda)} v
\end{aligned}$$

Soit x_A un vecteur propre de A et λ_A sa valeur propre associée.
M^{-1} et N peuvent être exprimés comme des polynômes de matrice :

> $$\begin{aligned}
> M^{-1} = \theta\mu A + I &\iff M^{-1}(A) = \theta\mu A^1 + A^0 \\
> N = (\theta - 1)\mu A + I &\iff N(A) = (\theta - 1)\mu A^1 + A^0
> \end{aligned}$$
>
> On en déduit donc la forme des valeurs propres de M et de N en fonction de celles de A :
>
> $$\lambda_M = 1 + \theta\mu\lambda_A \qquad \text{et} \qquad \lambda_N = 1 + (\theta - 1)\mu\lambda_A$$
>
> Les valeurs propres de $M^{-1}N$ sont donc de la forme :
>
> $$\frac{\lambda_N}{\lambda_M} = \frac{1 + (\theta - 1)\mu\lambda_A}{1 + \theta\mu\lambda_A} = \frac{1 + \theta\mu\lambda_A - \mu\lambda_A}{1 + \theta\mu\lambda_A}$$
>
> Or, $\mu > 0$ et A est définie positive donc $\lambda_A > 0$, d'où $\mu\lambda_A > 0$.
> Le quotient est donc strictement inférieur à 1.
>
> De plus, $\theta \in [0; 1]$ donc le quotient est toujours strictement supérieur à 1.

Toutes les valeurs propres de $M^{-1}N$ sont comprises dans l'intervalle $]-1; 1[$ donc :

$$\rho(M^{-1}N) = \max_i |\lambda_i| < 1$$

QUESTION 9. D'après le chapitre 3 du cours, la méthode de Crank-Nicolson converge si et seulement si $\rho(M^{-1}N) < 1$. La **question 8** a montré que c'était le cas.

Soit X la limite des $U^{(k)}$. Pour $k \to +\infty$, on a :

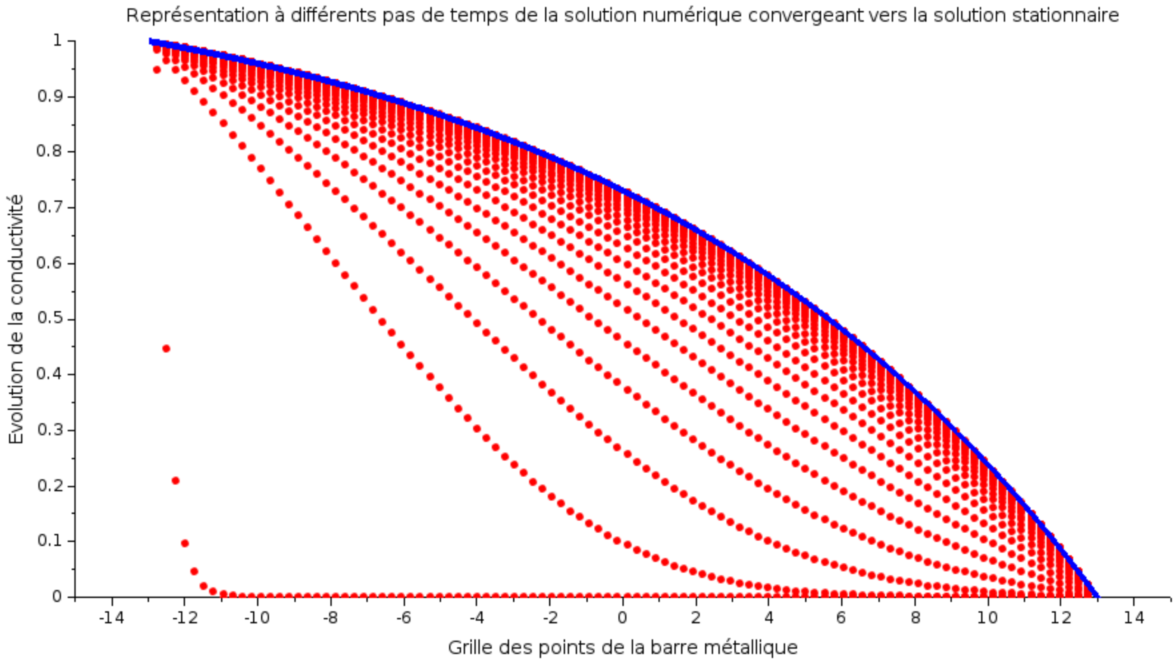
> $$\begin{aligned}
> MX = NX + \mu B &\iff (M - N)X = \mu B \\
> &\iff (I + \theta\mu A - I - (\theta - 1)\mu A)X = \mu B \\
> &\iff \mu AX = \mu B \\
> &\iff AX = B
> \end{aligned}$$

A et B sont les matrices de la **question 6** donc la solution discrétisée converge bien vers la solution stationnaire.

QUESTION 10. Dans le graphique ci-dessous, on trace en bleu la solution stationnaire trouvée dans la **question 6**. Puis on trace en rouge la solution approchée à différents pas t_k.

Se référer à : `Q10_CrankNicolson.sce`

On choisit les paramètres suivants :
- $n = 100$, nombre de points de discrétisation de la grille ;
- $n_t = 6000$, nombre d'itérations assez grand pour approcher de près la solution exacte ;
- $T = 500$, ce qui donne un pas de temps $\delta_t \simeq 0.083$.



Quand $k \to +\infty$, $U^{(k)}$ tend effectivement vers la solution stationnaire. La convergence de la méthode de Crank-Nicolson est rapide : sur le graphique, la deuxième mesure, déjà éloignée de $U^{(0)}$, a été relevée à $k = 200$. Chaque courbe rouge correspond à 200 itérations supplémentaires.

5 Étude du problème inverse

QUESTION 11. Mesure du flux à $t_{inter} = \frac{2}{3}T$ et $t_{fin} = T$.

Se référer à : `Q11_Flux.sce`

La fonction `flux(x_d)` est gourmande en temps. Sa complexité temporelle est en $O(n_t \cdot n)$ mais elle cache trois étapes en $O(n)$: le calcul du second membre, la descente et la remontée.

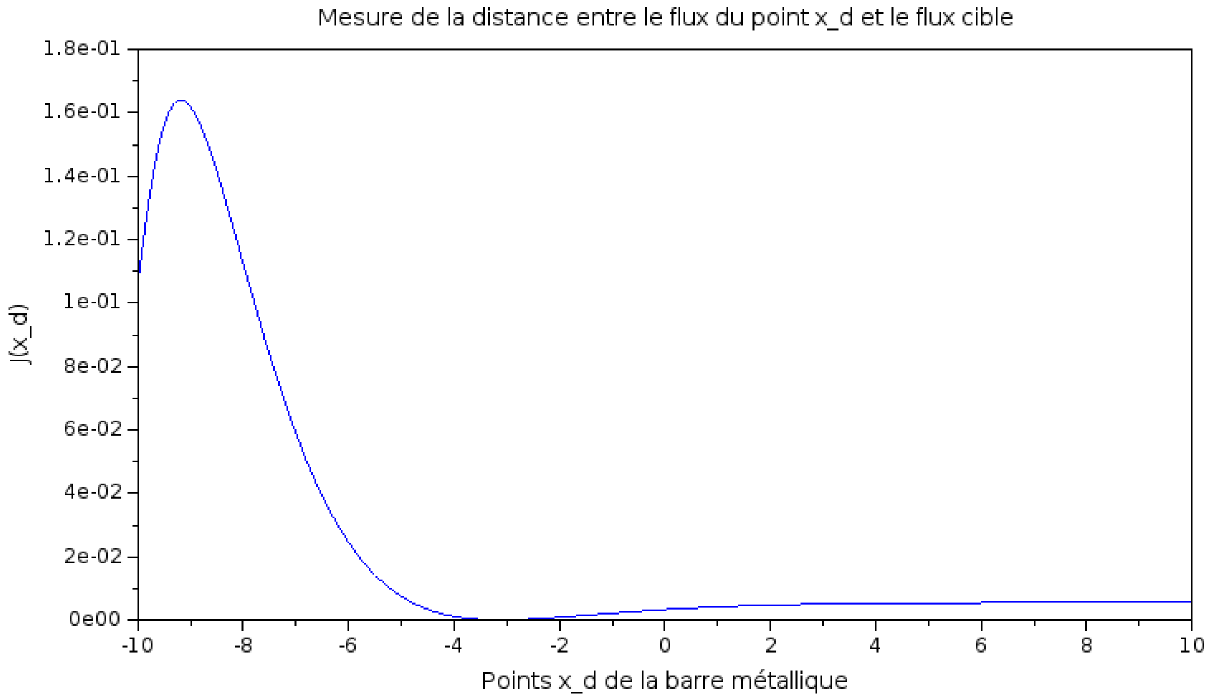
On en optimise pour cela quelques parties comme le calcul du vecteur $B^{(k)}$ ramené au calcul de sa seule composante non nulle $B_1^{(k)}$. On économie ainsi de l'espace mémoire et n itérations pour l'affectation.

QUESTION 12. Recherche de x_d^* avec la fonction J.

> Se référer à : `Q12_Fonction_J.sce`

La fonction $J(x_d)$ mesure l'écart (en norme euclidienne) entre le flux $\mathcal{F}$ mesuré à différents points x_d et le flux cible $\mathcal{F}_{cible} = (-0.1, -0.18)$.

Le graphique ci-dessous représente la fonction J sur la barre métallique $]-10\,;10[$ à raison de 200 points x_d d'échantillonnage.



La fonction J admet un unique minimum. C'est le flux égal au flux cible pour $x_d \simeq -3.5$. On implémente désormais deux méthodes pour trouver par le calcul cette valeur x_d^* : la méthode de la dichotomie et la méthode de Gauss-Newton.

QUESTION 13. Méthode de la dichotomie.

> Se référer à : `Q13_J_Dichotomie.sce`

$J(x_d)$ est une mesure de la distance entre le flux en x_d et le flux cible. On en déduit :

Données : $J(x_1), J(x_2), J(x_3)$
Sorties : I, intervalle décidé
si $J(x_1) = \min(J(x_1), J(x_2), J(x_3))$ **alors**
| $I = [a\,; x_2]$
sinon si $J(x_2) = \min(J(x_1), J(x_2), J(x_3))$ **alors**
| $I = [x_1\,; x_3]$
sinon
| $I = [x_2\,; b]$
fin

QUESTION 14. Méthode de Gauss-Newton.

Annexes

6 RESTE

Soit $x = (x_1, x_2, ..., x_n)$
$x^T Ax = \sum_{i=1}^{n}(C_{i+1/2} + C_{i-1/2})x_i^2 - 2\sum_{i=1}^{n-1} C_{i+1/2}x_i x_{i+1}$
$= \sum_{i=1}^{n} C_{i+1/2}x_i^2 + \sum_{j=0}^{n-1} C_{j+1/2}x_{j+1}^2 + -2\sum_{i=1}^{n-1} C_{i+1/2}x_i x_{i+1}$ On pose j = i-1
$= \sum_{i=1}^{n-1} C_{i+1/2}(x_i - x_{i+1})^2 + C_{1/2}x_{1/2}^2 + C_{n+1/2}x_n^2 > 0$
Comme $x^T Ax > 0 \forall x \in \mathbb{R}^n$ alors la matrice A est symétrique définie positive.

D'après (9) : $\frac{\partial}{\partial x}[C(x)\frac{\partial u}{\partial x}](x_i, t_k) = \frac{C_{i+1/2}u_{i+1}^{(k)} - (C_{i+1/2}+C_{i-1/2})u_i^{(k)} + C_{i-1/2}u_{i-1}^{k}}{\delta^2}$ Pour résoudre (13), on résout 9 pour chaque coefficient u_i^k

$$\begin{array}{c} B \\ \begin{pmatrix} u_0 C_{1/2} \\ 0 \\ 0 \\ 0 \end{pmatrix} \end{array}$$

Question 6.
D'après l'équation (9),
$\frac{\partial}{\partial x}[C(x)\frac{\partial u}{\partial x}](x_i, t_k) \approx \frac{C_{i+1/2}u_{i+1}^{(k)} - (C_{i+1/2}+C_{i-1/2})u_i^{(k)} + C_{i-1/2}u_{i-1}^{(k)}}{\delta_x^2}$
Si on écrit ce système, $\forall i$ sous forme matricielle, on obtient :

$\frac{\partial}{\partial x}[C(x)\frac{\partial u}{\partial x}] = \frac{1}{\delta^2}(-Au + \begin{array}{c} B \\ \begin{pmatrix} u_0 C_{1/2} \\ 0 \\ 0 \\ 0 \end{pmatrix} \end{array})$ car $u(-l) = u_0$ avec $u = (u_1^{(k)}, ..., u_n^{(k)})$ et A la matrice

trouvée à la question 1.
Donc A u = B, avec $B = (b_0, ..., b_n)$ où, $b_0 = C_{1/2} u_0$ et $b_i = 0 \ \forall i \in [2, n]$

Ce système admet une solution unique car A est définie et donc inversible.

Question 7.
On résout l'équation :

$$\begin{cases} \frac{\partial}{\partial x}[C(x)\frac{\partial u}{\partial x}] = 0 \\ u(-l) = u_0 \\ u(l) = 0 \end{cases}$$

$\frac{\partial}{\partial x}[C(x)\frac{\partial u}{\partial x}] = 0 \ -\frac{1}{l}e^{-\frac{x}{l}}\frac{\partial u}{\partial x} + e^{-\frac{x}{l}}\frac{\partial^2 u}{\partial x^2} = 0 \ \frac{\partial^2 u}{\partial x^2} - \frac{1}{l}\frac{\partial u}{\partial x} = 0$
$u(x) = \alpha + \beta e^{\frac{x}{l}}$
Calcul des constantes :

$$\begin{cases} u(-l) = \alpha + \beta e^{-1} = u_0 = 1 \\ u(l) = \alpha + \beta e = 0 \end{cases}$$

$$\begin{cases} \beta = \frac{1}{e^{-1} - e} \\ \alpha = -\frac{e}{e^{-1} - e} \end{cases}$$

Donc : $u(x) = -\frac{e}{e^{-1}-e} + \frac{1}{e^{-1}-e}e^{\frac{x}{l}}$
Question 8.

$M = I + \theta\mu A$ et $N = I + (\theta - 1)\mu A$
M et N sont donc des polynomes de matrices.
Donc, soit x_A un vecteur propre de A et λ_A une valeur propre de A, alors x_A est aussi un vecteur propre de M avec comme valeur propre $M\lambda_A$
Les valeurs propres de M sont $\lambda_M = 1 + \theta\mu\lambda_A$.
Les valeurs propres de N sont $\lambda_N = 1 + (\theta - 1)\mu\lambda_A$
Les valeurs propres de $M^{-1}N$ sont $\frac{\lambda_N}{\lambda_M} = \frac{1+(\theta-1)\mu\lambda_A}{1+\theta\mu\lambda_A} \ \forall\lambda_A$. $1 + (\theta - 1)\mu\lambda_A < 1 + \theta\mu\lambda_A$
Donc $\rho(M^{-1}N) = max(\frac{\lambda_N}{\lambda_M}) < 1$

Question 9.
$MU^{(k+1)} = NU^{(k)} + \mu B$ Cette méthode de décomposition converge uniquement si $\rho(M^{-1}N) < 1$, ce qui est le cas ici(question 8).
Cette décomposition converge vers $Ax = b$, avec A = M - N.
$M - N = I + \theta\mu A - I - (\theta - 1)\mu A = \mu A$
$\mu Ax = \mu B \ Ax = B$ On retrouve ici les matrices A et B de la question (6).