Correction TD3 : Systèmes linéaires et équations différentielles

1 Résolution de systèmes linéaires

Exercice 1 (Décompositions). Voici 2 matrices décomposées de la manière suivante (on ne demande pas de vérifier l'égalité)

$$\mathbf{A} = \begin{bmatrix} 2 & -2 & 2 \\ -1 & 2 & -9 \\ -2 & 0 & -4 \end{bmatrix} = \begin{bmatrix} -\frac{2}{3} & \frac{1}{3} & \frac{2}{3} \\ \frac{1}{3} & -\frac{2}{3} & \frac{2}{3} \\ \frac{2}{3} & \frac{2}{3} & \frac{1}{3} \end{bmatrix} \begin{bmatrix} -3 & 2 & -7 \\ 0 & -2 & 4 \\ 0 & 0 & -6 \end{bmatrix}$$

$$\mathbf{B} = \begin{bmatrix} -8 & 21 & 3 \\ 4 & -3 & 3 \\ -2 & 3 & -1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ -\frac{1}{2} & 1 & 0 \\ \frac{1}{4} & -\frac{3}{10} & 1 \end{bmatrix} \begin{bmatrix} -8 & 21 & 3 \\ 0 & \frac{15}{2} & \frac{9}{2} \\ 0 & 0 & -\frac{4}{10} \end{bmatrix}$$

1. Sous quelles formes sont décomposées les matrices ? Justifier.

La première décomposition est une décomposition QR avec

$$\mathbf{Q} = \begin{bmatrix} -\frac{2}{3} & \frac{1}{3} & \frac{2}{3} \\ \frac{1}{3} & -\frac{2}{3} & \frac{2}{3} \\ \frac{2}{3} & \frac{2}{3} & \frac{1}{3} \end{bmatrix}, \qquad \mathbf{R} = \begin{bmatrix} -3 & 2 & -7 \\ 0 & -2 & 4 \\ 0 & 0 & -6 \end{bmatrix}$$

On voit facilement que **R** est une matrice triangulaire supérieure. Il nous faut aussi vérifier que **Q** est une matrice orthogonale. Pour cela, on effectue le calcul suivant

$$\mathbf{Q}\mathbf{Q}^T = \begin{bmatrix} -\frac{2}{3} & \frac{1}{3} & \frac{2}{3} \\ \frac{1}{3} & -\frac{2}{3} & \frac{2}{3} \\ \frac{2}{3} & \frac{2}{3} & \frac{1}{3} \end{bmatrix} \begin{bmatrix} -\frac{2}{3} & \frac{1}{3} & \frac{2}{3} \\ \frac{1}{3} & -\frac{2}{3} & \frac{2}{3} \\ \frac{2}{3} & \frac{2}{3} & \frac{1}{3} \end{bmatrix} = \begin{bmatrix} \frac{4}{9}+\frac{1}{9}+\frac{4}{9} & -\frac{2}{9}-\frac{2}{9}+\frac{4}{9} & -\frac{4}{9}+\frac{2}{9}+\frac{2}{9} \\ -\frac{2}{9}-\frac{2}{9}+\frac{4}{9} & \frac{1}{9}+\frac{4}{9}+\frac{4}{9} & \frac{2}{9}-\frac{4}{9}+\frac{2}{9} \\ -\frac{4}{9}+\frac{2}{9}+\frac{2}{9} & \frac{2}{9}-\frac{4}{9}+\frac{2}{9} & \frac{4}{9}+\frac{4}{9}+\frac{1}{9} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 0 \end{bmatrix}.$$

On a donc bien $\mathbf{Q}\mathbf{Q}^T = \mathbf{I}$ (matrice identité : des 1 sur la diagonale et des 0 partout ailleurs). Donc **Q** est une matrice orthogonale.
La seconde décomposition est une décomposition LU avec

$$\mathbf{L} = \begin{bmatrix} 1 & 0 & 0 \\ -\frac{1}{2} & 1 & 0 \\ \frac{1}{4} & -\frac{3}{10} & 1 \end{bmatrix}, \qquad \mathbf{U} = \begin{bmatrix} -8 & 21 & 3 \\ 0 & \frac{15}{2} & \frac{9}{2} \\ 0 & 0 & -\frac{4}{10} \end{bmatrix}.$$

En effet, **L** est une matrice triangulaire inférieure et **U** une matrice triangulaire supérieure.

2. Quel est le rang de chaque matrice ? Justifier.

On peut montrer que le rang des deux matrices étudiées est 3. Étudions d'abord la première matrice

$$\mathbf{A} = \begin{bmatrix} 2 & -2 & 2 \\ -1 & 2 & -9 \\ -2 & 0 & -4 \end{bmatrix}$$

Clairement, les deux premières lignes sont indépendantes (on ne peut pas exprimer la seconde ligne en fonction de la première de manière linéaire) et le rang de la matrice est forcément supérieur ou égal à 2 (il est forcément inférieur ou égal à 3 car on travaille sur une matrice 3×3). Cherchons maintenant si la troisième ligne l_3 peut être écrite comme une combinaison linéaire des deux premières lignes l_1 et l_2. Cela à chercher deux réels λ et μ tel que

$$l_3 = \lambda l_1 + \mu l_2$$

et donc de voir si le système suivant a une solution

$$\left\{\begin{array}{rcr} 2\lambda - \mu & = & -2 \\ -2\lambda + 2\mu & = & 0 \\ 2\lambda - 9\mu & = & -4 \end{array}\right.$$

Ce système n'admet aucune solution. On note l'_1, l'_2 et l'_3 les lignes de ce système. On a du côté gauche de l'égalité

$$l'_3 + 7l'_1 - 8l'_2 = 2\lambda - 9\mu + 7(2\lambda - \mu) + 8(-2\lambda + 2\mu) = (2 + 14 - 16)\lambda + (-9 - 7 + 16)\mu = 0$$

et du côté droit de l'égalité

$$l'_3 + 7l'_1 - 8l'_2 = -4 - 14 + 0 = -18$$

D'où, $0 = -18$ si ce système admet des solutions, ce qui n'est pas possible. Ce système n'admet pas de solutions donc les lignes de la matrices sont linéairement indépendantes et le **rang de la matrice est** 3.

Étudions maintenant la deuxième matrice

$$\mathbf{B} = \begin{bmatrix} -8 & 21 & 3 \\ 4 & -3 & 3 \\ -2 & 3 & -1 \end{bmatrix}$$

Clairement, les deux premières lignes sont indépendantes et le rang de la matrice est forcément supérieur ou égal à 2. Cherchons maintenant si la troisième ligne l_3 peut être écrite comme une combinaison linéaire des deux premières lignes l_1 et l_2. Cela à chercher deux réels λ et μ tel que

$$l_3 = \lambda l_1 + \mu l_2$$

et donc de voir si le système suivant a une solution

$$\left\{\begin{array}{rcr} -8\lambda + 4\mu & = & -2 \\ 21\lambda - 3\mu & = & 3 \\ 3\lambda + 3\mu & = & -1 \end{array}\right.$$

Ce système n'admet aucune solution. On note l'_1, l'_2 et l'_3 les lignes de ce système. On a du côté gauche de l'égalité

$$l_2 + 2l_1 - \frac{5}{3}l_3 = \lambda(21 - 16 - 5) + \mu(-3 + 8 - 5) = 0$$

et du côté droit

$$l_2 + 2l_1 - \frac{5}{3}l_3 = 3 - 4 + \frac{5}{3} = \frac{2}{3}$$

D'où, $0 = \frac{2}{3}$ si ce système admet des solutions, ce qui n'est pas possible. Ce système n'admet pas de solutions donc les lignes de la matrices sont linéairement indépendantes et le **rang de la matrice est 3**.

3. Résoudre les systèmes suivants en utilisant les questions précédentes :

$$\mathbf{Ax} = \begin{bmatrix} 2 & 4 & -3 \end{bmatrix}^\top$$
$$\mathbf{Bx} = \begin{bmatrix} 8 & 20 & -8 \end{bmatrix}^\top$$

On veut résoudre le système suivant

$$\begin{bmatrix} 2 & -2 & 2 \\ -1 & 2 & -9 \\ -2 & 0 & -4 \end{bmatrix} \mathbf{x} = \begin{bmatrix} 2 \\ 4 \\ -3 \end{bmatrix}.$$

Or dans la question 1, on a étudié la décomposition QR de la matrice du système et pour résoudre le système, il faut tout d'abord calculer

$$\mathbf{y} = \mathbf{Q}^T \begin{bmatrix} 2 \\ 4 \\ -3 \end{bmatrix} = \begin{bmatrix} -\frac{2}{3} & \frac{1}{3} & \frac{2}{3} \\ \frac{1}{3} & -\frac{2}{3} & \frac{2}{3} \\ \frac{2}{3} & \frac{2}{3} & \frac{1}{3} \end{bmatrix} \begin{bmatrix} 2 \\ 4 \\ -3 \end{bmatrix} = \begin{bmatrix} -\frac{4}{3} + \frac{4}{3} - 2 \\ \frac{2}{3} - \frac{8}{3} - 2 \\ \frac{4}{3} + \frac{8}{3} - 1 \end{bmatrix} = \begin{bmatrix} -2 \\ -4 \\ 3 \end{bmatrix}$$

Puis, il faut résoudre le système $\mathbf{R}\mathbf{x} = \mathbf{y}$. Notons $\mathbf{x} = \begin{bmatrix} x_1 & x_2 & x_3 \end{bmatrix}^\top$, on veut donc résoudre le système suivant

$$\begin{cases} -3x_1 + 2x_2 - 7x_3 & = & -2 \\ \quad\quad -2x_2 + 4x_3 & = & -4 \\ \quad\quad\quad\quad -6x_3 & = & 3 \end{cases}$$

On a immédiatement que $x_3 = -\frac{1}{2}$, d'où avec la seconde ligne $x_2 = 1$ et finalement avec la première ligne $x_1 = \frac{5}{2}$.
On veut maintenant résoudre le système suivant

$$\begin{bmatrix} -8 & 21 & 3 \\ 4 & -3 & 3 \\ -2 & 3 & -1 \end{bmatrix} \mathbf{x} = \begin{bmatrix} 8 \\ 20 \\ -8 \end{bmatrix}.$$

Or dans la question 1, on a étudié la décomposition LU de la matrice du système. Pour le résoudre, il faut tout d'abord résoudre le système $\mathbf{L}\mathbf{y} = \mathbf{b}$. Si on note $\mathbf{y} = \begin{bmatrix} y_1 & y_2 & y_3 \end{bmatrix}^\top$, cela revient à résoudre le système suivant

$$\begin{cases} y_1 & = & 8 \\ -\frac{1}{2}y_1 + y_2 & = & 20 \\ \frac{1}{4}y_1 - \frac{3}{10}y_2 + y_3 & = & -8 \end{cases}$$

On a immédiatement $y_1 = 8$, d'où avec la deuxième ligne $y_2 = 24$ et avec la troisième ligne $y_3 = -10 + \frac{36}{5} = -\frac{14}{5}$. On doit ensuite résoudre le système $\mathbf{U}\mathbf{x} = \mathbf{y}$. Notons $\mathbf{x} = \begin{bmatrix} x_1 & x_2 & x_3 \end{bmatrix}^\top$, on veut donc résoudre le système suivant

$$\begin{cases} -8x_1 + 21x_2 + 3x_3 & = & 8 \\ \quad\quad \frac{15}{2}x_2 + \frac{9}{2}x_3 & = & 24 \\ \quad\quad\quad\quad -\frac{4}{10}x_3 & = & -\frac{14}{5} \end{cases}$$

On a directement $x_3 = \frac{14}{5} \times \frac{10}{4} = 7$. Puis

$$x_2 = \frac{2}{15}(24 - \frac{9 \times 7}{2}) = \frac{1}{15}(48 - 63) = -1$$

et

$$x_1 = -\frac{1}{8}(8 - 21 + 21) = -1.$$

Exercice 2 (Factorisation LU d'une matrice tridiagonale)**.** Une matrice tridiagonale est une matrice de la forme :

$$\begin{bmatrix} a_1 & c_1 & 0 & \dots & 0 \\ b_2 & a_2 & c_2 & \ddots & \vdots \\ 0 & \ddots & \ddots & \ddots & 0 \\ \vdots & \ddots & \ddots & \ddots & c_{n-1} \\ 0 & \dots & 0 & b_n & a_n \end{bmatrix}$$

1. Soit $\mathbf{A}$ la matrice tridiagonale $n \times n$ suivante ($a_i = 2, b_j = c_k = -1$) :

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

On cherche une factorisation $\mathbf{A} = \mathbf{LU}$ de $\mathbf{A}$, où $\mathbf{L}$ et $\mathbf{U}$ sont des matrices bidiagonales de la forme :

$$\mathbf{L} = \begin{bmatrix} 1 & 0 & 0 & \dots & 0 \\ \beta_2 & 1 & 0 & \ddots & \vdots \\ 0 & \ddots & \ddots & \ddots & 0 \\ \vdots & \ddots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & \beta_n & 1 \end{bmatrix}, \quad \mathbf{U} = \begin{bmatrix} \alpha_1 & -1 & 0 & \dots & 0 \\ 0 & \alpha_2 & -1 & \ddots & \vdots \\ 0 & \ddots & \ddots & \ddots & 0 \\ \vdots & \ddots & \ddots & \ddots & -1 \\ 0 & \dots & 0 & 0 & \alpha_n \end{bmatrix}$$

Montrer que cette factorisation existe et que les coefficients α_i et β_i s'obtiennent avec les relations suivantes :

$$\alpha_1 = 2, \quad \beta_i = \frac{-1}{\alpha_{i-1}}, \quad \alpha_i = 2 + \beta_i, \quad \forall i = 2, \cdots, n$$

Pour montrer les relations, on calcule le produit matriciel $\mathbf{LU}$. Pour simplicité, on va prendre $n = 5$ (le calcul est similaire pour un n général) :

$$\mathbf{LU} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ \beta_2 & 1 & 0 & 0 & 0 \\ 0 & \beta_3 & 1 & 0 & 0 \\ 0 & 0 & \beta_4 & 1 & 0 \\ 0 & 0 & 0 & \beta_5 & 1 \end{bmatrix} \begin{bmatrix} \alpha_1 & -1 & 0 & 0 & 0 \\ 0 & \alpha_2 & -1 & 0 & 0 \\ 0 & 0 & \alpha_3 & -1 & 0 \\ 0 & 0 & 0 & \alpha_4 & -1 \\ 0 & 0 & 0 & 0 & \alpha_5 \end{bmatrix} = \begin{bmatrix} \alpha_1 & -1 & 0 & 0 & 0 \\ \alpha_1\beta_2 & -\beta_2+\alpha_2 & -1 & 0 & 0 \\ 0 & \alpha_2\beta_3 & -\beta_3+\alpha_3 & -1 & 0 \\ 0 & 0 & \alpha_3\beta_4 & -\beta_4+\alpha_4 & -1 \\ 0 & 0 & 0 & \alpha_4\beta_5 & -\beta_5+\alpha_5 \end{bmatrix}.$$

On a donc

$$\mathbf{A} = \begin{bmatrix} 2 & -1 & 0 & 0 & 0 \\ -1 & 2 & -1 & 0 & 0 \\ 0 & -1 & 2 & -1 & 0 \\ 0 & 0 & -1 & 2 & -1 \\ 0 & 0 & 0 & -1 & 2 \end{bmatrix} = \begin{bmatrix} \alpha_1 & -1 & 0 & 0 & 0 \\ \alpha_1\beta_2 & -\beta_2+\alpha_2 & -1 & 0 & 0 \\ 0 & \alpha_2\beta_3 & -\beta_3+\alpha_3 & -1 & 0 \\ 0 & 0 & \alpha_3\beta_4 & -\beta_4+\alpha_4 & -1 \\ 0 & 0 & 0 & \alpha_4\beta_5 & -\beta_5+\alpha_5 \end{bmatrix} = \mathbf{LU}.$$

En comparant les deux matrices élément par élément, on retrouve les relations cherchées.

2. Soit $\mathbf{f} = (f_1, \dots, f_n)^\top$. Montrer que la solution $\mathbf{x} = (x_1, \dots, x_n)^\top$ du système linéaire $\mathbf{Ax} = \mathbf{f}$ se calcule par des formules de récurrence, en deux temps :

(i) Calcul d'un vecteur $\mathbf{y} = (y_1, \dots, y_n)^\top$ vérifiant :

$$y_1 = f_1, \quad y_i = f_i - \beta_i y_{i-1}, \quad \forall i = 2, \dots, n$$

(ii) Calcul de la solution $\mathbf{x}$ par :

$$x_n = \frac{y_n}{\alpha_n}, \quad x_i = \frac{y_i + x_{i+1}}{\alpha_i}, \quad \forall i = n-1, \dots, 1$$

On veut calculer la solution du système $\mathbf{Ax} = \mathbf{f}$ ce qui est équivalent à résoudre

$$\mathbf{LUx} = \mathbf{f}.$$

Notons $\mathbf{y} = \mathbf{Ux}$ ($\mathbf{y}$ est un vecteur colonne). On peut maintenant réécrire notre système comme deux systèmes suivants :

$$\mathbf{Ly} = \mathbf{f} \tag{1}$$
$$\mathbf{Ux} = \mathbf{y}. \tag{2}$$

Il faut d'abord résoudre le premier système car $\mathbf{f}$ est connu. En reprenant l'exemple pour $n = 5$:

$$\mathbf{Ly} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ \beta_2 & 1 & 0 & 0 & 0 \\ 0 & \beta_3 & 1 & 0 & 0 \\ 0 & 0 & \beta_4 & 1 & 0 \\ 0 & 0 & 0 & \beta_5 & 1 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \\ y_5 \end{bmatrix} = \begin{bmatrix} f_1 \\ f_2 \\ f_3 \\ f_4 \\ f_5 \end{bmatrix} = \mathbf{f}.$$

La première ligne donne tout de suite $y_1 = f_1$. En progressant aux lignes suivantes, on retrouve les relations cherchées :

$$y_i = f_i - \beta_i y_{i-1}.$$

On peut maintenant résoudre le deuxième système

$$\mathbf{Ux} = \begin{bmatrix} \alpha_1 & -1 & 0 & 0 & 0 \\ 0 & \alpha_2 & -1 & 0 & 0 \\ 0 & 0 & \alpha_3 & -1 & 0 \\ 0 & 0 & 0 & \alpha_4 & -1 \\ 0 & 0 & 0 & 0 & \alpha_5 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \\ y_5 \end{bmatrix} = \mathbf{y}.$$

La dernière ligne donne tout de suite $x_5 = \frac{y_5}{\alpha_5}$. En progressant aux lignes précédentes, on retrouve les relations cherchées :

$$x_i = \frac{y_i + x_{i+1}}{\alpha_i}.$$

3. Préciser les matrices **L** et **U** dans le cas $n = 3$.

$$\mathbf{L} = \begin{bmatrix} 1 & 0 & 0 \\ -\frac{1}{2} & 1 & 0 \\ 0 & -\frac{2}{3} & 1 \end{bmatrix} \qquad \mathbf{U} = \begin{bmatrix} 2 & -1 & 0 \\ 0 & \frac{3}{2} & -1 \\ 0 & 0 & \frac{4}{3} \end{bmatrix}$$

2 Résolution d'équations différentielles

Exercice 3. Soit f une fonction continue sur $[0,1]$. On considère le problème suivant :
Trouver $u \in \mathcal{C}^2([0,1])$ tel que :

$$\begin{cases} -u''(x) &= f(x) \qquad x \in]0,1[\\ u(0) &= 0 \\ u(1) &= 0 \end{cases} \tag{3}$$

1. Dans cette question on pose $f(x) = 1,\ \forall x \in [0,1]$. Calculer la solution $u(x)$ du problème (3).

On veut résoudre l'équation $-u''(x) = 1$. Pour cela, il suffit d'intégrer deux fois :

$$\begin{aligned} u''(x) &= -1 \\ u'(x) &= -x + c \\ u(x) &= -\frac{x^2}{2} + cx + d \qquad c, d \in \mathbb{R} \end{aligned}$$

On calcule les constantes c et d en utilisant les conditions de bord :

$$\begin{aligned} 0 = u(0) &= -\frac{0^2}{2} + c \times 0 + d = d \\ 0 = u(1) &= -\frac{1^2}{2} + c \times 1 + d = -\frac{1}{2} + c + d = -\frac{1}{2} + c \end{aligned}$$

On voit que $c = \frac{1}{2}$ et $d = 0$, on a donc la fonction

$$u(x) = -\frac{x^2}{2} + \frac{x}{2} = \tfrac{1}{2}x\,(1-x)\,.$$

2. On suppose que la fonction u est de classe $\mathcal{C}^4$ sur $[0,1]$. Montrer, en utilisant la formule de Taylor, que $u''(x)$ peut s'écrire en tout point $x \in]0,1[$:

$$u''(x) = \frac{u(x+h) - 2u(x) + u(x-h)}{h^2} + h^2\varepsilon(x), \qquad \forall h \neq 0 \text{ tel que } x+h,\ x-h \in]0,1[$$

où la fonction ε est bornée sur $[0,1]$.

La formule de Taylor pour une fonction $u(x)$ exprimé à l'abscisse $x \pm h$ donne

$$u(x+h) = u(x) + \frac{h^1}{1!}u'(x) + \frac{h^2}{2!}u''(x) + \frac{h^3}{3!}u'''(x) + h^4\varepsilon_+(x) \tag{4}$$

$$u(x-h) = u(x) - \frac{h^1}{1!}u'(x) + \frac{h^2}{2!}u''(x) - \frac{h^3}{3!}u'''(x) + h^4\varepsilon_-(x) \tag{5}$$

où les deux fonctions $\varepsilon_\pm(x)$ sont bornées sur $[0,1]$. En prenant la somme de (4) et (5), on obtient

$$u(x+h) + u(x-h) = 2u(x) + h^2u''(x) - h^4\varepsilon(x)$$

avec $-\varepsilon(x) = \varepsilon_+(x) + \varepsilon_-(x)$. Finalement, si on divise cette dernière équation par h^2 et exprime $u''(x)$, on a

$$u''(x) = \frac{u(x+h) - 2u(x) + u(x-h)}{h^2} + h^2\varepsilon(x).$$

3. On revient au cas général. On subdivise de façon régulière l'intervalle $[0,1]$. Soit $n \in \mathbb{N}^*$ et $h = \frac{1}{n}$. On pose $x_i = ih$, et $f_i = f(x_i)$ pour $i = 0, \dots, n$. Les inconnues sont désormais $u_i = u(x_i)$ pour $i = 0, \dots, n$.

(i) Montrer que l'on a l'approximation :

$$u''(x_i) = \frac{u_{i+1} - 2u_i + u_{i-1}}{h^2}$$

En utilisant la question 2 avec $h = \frac{1}{n}$, on peut écrire

$$u''(x_i) = \frac{u(x_i+h) - 2u(x_i) + u(x_i-h)}{h^2} = \frac{u(x_{i+1}) - 2u(x_i) + u(x_{i-1})}{h^2} = \frac{u_{i+1} - 2u_i + u_{i-1}}{h^2}.$$

(ii) Montrer que le problème (3) écrit aux noeuds x_i :

$$\left\{\begin{array}{rcll} -u''(x_i) & = & f(x_i) & \quad x_i = 1, \dots, n-1 \\ u(0) & = & 0 & \\ u(1) & = & 0 & \end{array}\right. \tag{6}$$

peut être approché par le système linéaire :

$$A\mathbf{u} = \mathbf{f}, \quad u_0 = 0,\ u_n = 0$$

avec $\mathbf{u} = (u_1, u_2, \dots, u_{n-1})^\top$, $\mathbf{f} = (f_1, f_2, \dots, f_{n-1})^T$, et A une matrice que l'on précisera.

On voit tout de suite que $0 = u(0) = u(x_0) = u_0$ et $0 = u(1) = u(x_n) = u_n$. En utilisant l'approximation de question précédente, on a le système suivant :

$$-\frac{u_0 - 2u_1 + u_2}{h^2} = f_1; \qquad -\frac{u_{i-1} - 2u_i + u_{i+1}}{h^2} = f_i \quad \text{pour} \quad i = 2, \dots, n-2; \qquad -\frac{u_{n-2} - 2u_{n-1} + u_n}{h^2} = f_{n-1}.$$

Avec les conditions de bord $u_0 = u_n = 0$, ceci est équivalent au système

$$\begin{aligned} 2u_1 - u_2 &= h^2 f_1 \\ -u_{i-1} + 2u_i - u_{i+1} &= h^2 f_i \qquad i = 2, \dots, n-2 \\ -u_{n-2} + 2u_{n-1} &= h^2 f_{n-1} \end{aligned}$$

Le système s'écrit sous la forme matricielle suivante :

$$\begin{bmatrix} 2 & -1 & 0 & \dots & 0 \\ -1 & 2 & -1 & \ddots & \vdots \\ 0 & \ddots & \ddots & \ddots & 0 \\ \vdots & \ddots & \ddots & \ddots & -1 \\ 0 & \dots & 0 & -1 & 2 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_{n-2} \\ u_{n-1} \end{bmatrix} = h^2 \begin{bmatrix} f_1 \\ f_2 \\ \vdots \\ f_{n-2} \\ f_{n-1} \end{bmatrix}$$

On remarque que la matrice de système est la matrice tridiagonale **A** de l'Exercice 2.

4. On considère de nouveau $f = 1$, et on prend $n = 4$. En utilisant Exercice 2, résoudre le système linéaire :

$$\left\{\begin{array}{rcrcrcl} 2\,u_1 & - & u_2 & & & = & \frac{1}{16} \\ -u_1 & + & 2\,u_2 & - & u_3 & = & \frac{1}{16} \\ & - & u_2 & + & 2\,u_3 & = & \frac{1}{16} \end{array}\right.$$

Comparer les valeurs u_i obtenues avec les valeurs de la solution exacte trouvée à la question 1, $u(x_i)$ pour $i = 1, \ldots, 3$. Commentaires ?

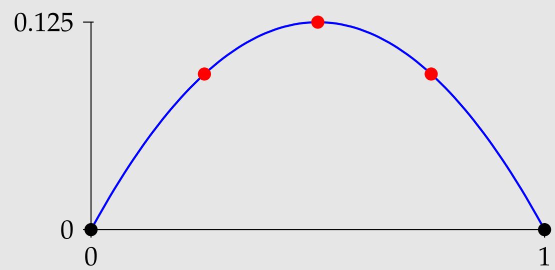
On résoud le système en deux temps en utilisant la décomposition LU calculé dans l'Exercice 2, question 3.

$$\begin{bmatrix} 1 & 0 & 0 \\ -\frac{1}{2} & 1 & 0 \\ 0 & -\frac{2}{3} & 1 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = \begin{bmatrix} \frac{1}{16} \\ \frac{1}{16} \\ \frac{1}{16} \end{bmatrix} \quad \Longrightarrow \quad \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = \frac{1}{32} \begin{bmatrix} 2 \\ 3 \\ 4 \end{bmatrix} = \begin{bmatrix} 0.0625 \\ 0.09375 \\ 0.125 \end{bmatrix}$$

$$\begin{bmatrix} 2 & -1 & 0 \\ 0 & \frac{3}{2} & -1 \\ 0 & 0 & \frac{4}{3} \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} \quad \Longrightarrow \quad \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix} = \frac{1}{32} \begin{bmatrix} 3 \\ 4 \\ 3 \end{bmatrix} = \begin{bmatrix} 0.09375 \\ 0.125 \\ 0.09375 \end{bmatrix}$$

Le graphe à droite montre la fonction $u(x) = 0.5\,x\,(1 - x)$ calculée analytiquement dans question 1 pour $x \in [0, 1]$, ainsi que les valeurs de u_0, u_4 (les conditions de bord) et u_1, u_2, u_3 (la solution numérique). On voit que même si la subdivision utilisé est assez brute, la solution numérique approxime bien la solution analytique. En effet, l'erreur commise est

$$\begin{aligned} u(x_1) - u_1 = u(0.25) - u_1 &= (0.5 \times 0.25 \times (1 - 0.25)) - 0.09375 &= 0 \\ u(x_2) - u_2 = u(0.50) - u_2 &= (0.5 \times 0.50 \times (1 - 0.50)) - 0.12500 &= 0 \\ u(x_3) - u_3 = u(0.75) - u_3 &= (0.5 \times 0.75 \times (1 - 0.75)) - 0.09375 &= 0 \end{aligned}$$



3 Extrait de l'examen de 2012

3.1 Représentation des nombres (4 points)

Lors d'un calcul, un ordinateur fournit des réponses approximatives pour des raisons de cardinalité. En effet, il utilise un nombre fini de bits (chiffre binaire c'est à dire 0 ou 1) pour représenter les entiers ou les réels. Pour comprendre comment l'ordinateur effectue ses calculs et éviter les situations pathogènes, il est intéressant de comprendre le calcul en système binaire et la représentation des réels en nombres à virgule flottante.

Question 1. *"There are only 10 types of people in the world : those who understand binary and those who don't." Expliquer la blague.*

La réponse est simple. En binaire, 10 veut dire 2. Si on traduit en français la phrase, on a donc : "Il y a seulement 10 (= 2 en binaire) types de personnes dans le monde : ceux qui comprennent le binaire et ceux qui ne le comprennent pas". A priori, si vous avez compris la blague, c'est que vous comprenez le binaire.

Question 2. *Convertir les nombres 12 et 13 en binaire. Effectuer l'addition 12+13 en binaire.*

Convertissons tout d'abord 12 en binaire. Cela donne

$$\begin{array}{rcccccl} 12 & = & 2 & \times & 6 & + & \mathbf{0} \\ 6 & = & 2 & \times & 3 & + & \mathbf{0} \\ 3 & = & 2 & \times & 1 & + & \mathbf{1} \\ 1 & = & 2 & \times & 0 & + & \mathbf{1} \end{array}$$

On a donc $(12)_{10} = (1100)_2$. Convertissons maintenant 13 en binaire. On a

$$\begin{array}{ccccccc} 13 & = & 2 & \times & 6 & + & \mathbf{1} \\ 6 & = & 2 & \times & 3 & + & \mathbf{0} \\ 3 & = & 2 & \times & 1 & + & \mathbf{1} \\ 1 & = & 2 & \times & 0 & + & \mathbf{1} \end{array}$$

On a ainsi $(13)_{10} = (1101)_2$. On effectue maintenant l'addition de $(1100)_2$ et $(1101)_2$. Pour rappel, l'addition en binaire fonctionne de la manière suivante

+	0	1
0	0	1
1	1	10

D'où l'opération suivante

$$\begin{array}{cccccc} & & 1 & 1 & 0 & 0 \\ + & & 1 & 1 & 0 & 1 \\ \hline = & 1 & {}^{1}1 & {}^{1}0 & 0 & 1 \end{array}$$

On a $(1100)_2 + (1101)_2 = (11001)_2$. Or $(12)_{10} + (13)_{10} = (25)_{10}$, vérifions si $(25)_{10} = (11001)_2$.

$$\begin{array}{ccccccc} 2 & \times & 0 & + & \mathbf{1} & = & 1 \\ 2 & \times & 1 & + & \mathbf{1} & = & 3 \\ 2 & \times & 3 & + & \mathbf{0} & = & 6 \\ 2 & \times & 6 & + & \mathbf{0} & = & 12 \\ 2 & \times & 12 & + & \mathbf{1} & = & 25 \end{array}$$

On a bien $(25)_{10} = (11001)_2$, le résultat obtenu en binaire est bien conforme au résultat obtenu en base 10.

Question 3. *Comment sont représentés les réels sur un ordinateur ? Quels sont les avantages et les inconvénients d'une telle représentation ? (4 lignes maximum)*

Les réels sont stockés sur un ordinateur sous forme de flottants en base 2, soit par 2 nombres écrits en binaire, la mantisse contenant la valeur normalisée du nombre que l'on veut représenter et l'exposant qui définit le décalage par rapport à la normalisation. L'avantage principal est la rapidité des calculs sur ces nombres (opérations en binaires faciles), l'inconvéniant principal est la perte de précision sur certains réels.

Question 4. *Expliquer les résultats de la session Scilab présentée ci-dessous :*

commande :	résultat :
`» a=1`	`a =  1.`
`» b=10^20`	`b =  1.000D+20`
`» c=-b`	`c = -1.000D+20`
`» (a+b)+c`	`ans =  0.`
`» a+(b+c)`	`ans =  1.`

Les 3 premières lignes sont évidentes à comprendre, on affecte correctement 1 à a, 10^{20} à b et -10^{20} à c. La quatrième ligne ne donne pas le résultat attendu car il effectue tout d'abord l'opération $a+b$ mais en pratique du fait de l'écart trop important entre a et b, l'ordinateur va évaluer $a + b$ à 10^{20} et non $10^{20} + 1$. Puis l'ordinateur soustrait 10^{20} au résultat précédent et on obtient 0 au lieu du 1 attendu. La cinquième ligne permet un ordre des calcul évitant ce problème, mais il suffit de remplacer a par b et vice versa dans cette ligne pour retrouver le même genre de problème qu'à la ligne précédente.

3.2 Intégration (3 points)

Soit g une fonction de [0,1] dans $\mathbb{R}$, en général les méthodes analytiques d'évaluation de $I = \int_0^1 g(x)dx$ se basent sur le calcul d'une primitive de g. Cependant, pour la plupart des fonctions, on ne connaît pas d'expression analytique des primitives. Il est alors intéressant d'utiliser une méthode numérique. On considère la méthode de quadrature de la forme :

$$\int_0^1 g(x)dx \approx g(0) \tag{7}$$

Question 5. *Donner l'interprétation géométrique d'une telle méthode.*

Il faut tout d'abord remarquer que l'approximation vaut

$$\int_0^1 g(x)dx \approx g(0) = g(0) \times (1-0)$$

$g(0) \times (1-0)$ peut être vu comme l'aire du rectangle de longueur $g(0)$ en ordonnée et de largeur $1-0$ entre les abscisses 0 et 1. On retrouve ici la formule des rectangles (formule de Newton-Cotes fermé $n=0$)

Question 6. *Quelle est le degré de précision de ce schéma ?*

On a vu dans le cours que la formule des rectangles (dans le cas Newton-Cotes fermé) était de degré de précision 0. Néanmoins, si on n'avait pas reconnu la formule des rectangles ici, on pouvait directement le vérifier. En effet, si $g(x) = 1$ sur $[0,1]$,

$$\int_0^1 g(x)\,dx = 1 \qquad g(0) = 1$$

La formule de quadrature est donc exacte pour $g(x) = 1$, elle est de degré de précision au moins 0. Si $g(x) = x$ sur $[0,1]$,

$$\int_0^1 g(x)\,dx = \frac{1}{2} \qquad g(0) = 0$$

La formule de quadrature n'est pas exacte pour $g(x) = x$, son degré de précision est donc bien 0.

Question 7. *Cette méthode parait-elle intéressante ? Justifier.*

Cette méthode n'est pas intéressante car l'erreur commise est rapidement importante (voir à la question précédente pour $g(x) = x$) même pour des fonctions g vraiment simple. Elle serait plus intéressante sous forme composite mais cela implique forcément plus d'évaluation de la fonction g que la méthode simple (soit 1 seule fois).

3.3 Méthode itérative du point fixe (3 points)

En analyse numérique, une méthode itérative est un procédé algorithmique. Pour résoudre un problème donné, après le choix d'une valeur initiale considérée comme une première ébauche de solution, la méthode procède par itérations au cours desquelles elle détermine une succession de solutions approximatives raffinées qui se rapprochent graduellement de la solution cherchée.

Par exemple, on cherche à résoudre numériquement l'équation $\cos(x^*) = x^*$ sur [0,1].

Question 8. *Montrer en utilisant le théorème du point fixe que l'itération $x_n = \cos(x_{n-1})$ avec $x_0 \in [0,1]$ converge vers la solution unique d'une telle équation. On rappelle que $\forall x \in [0,1], \sin(x) \leq \sin(1) < 1$.*

On suppose que l'équation $\cos(x^*) = x^*$ admet une unique solution sur $[0,1]$ (implicitement supposé dans l'énoncé, une démonstration simple existe, pour cela regarder dans la correction du TP 1). Pour tout n, $x_n \in [0,1]$, donc pour montrer que la suite converge bien vers x^*, il suffit de montrer que la suite (x_n) converge.

On pose $g(x) = \cos(x)$ est une fonction continue et dérivable sur $[0,1]$ et $g'(x) = \sin(x)$ est une fonction continue sur $[0,1]$. g' admet donc un maximum sur $[0,1]$ et comme elle est strictement croissante sur $[0,1]$

$$\text{Pour tout } x \text{ de } [0,1] \qquad 0 \leq g'(x) \leq g'(1) < 1$$

donc par passage au maximum

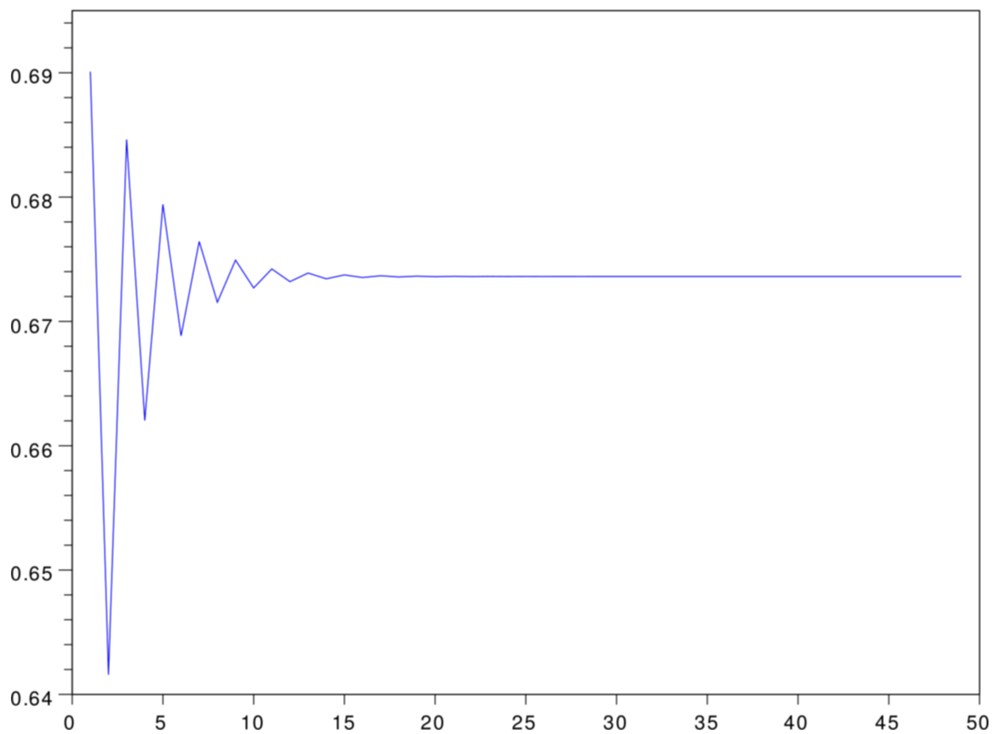
$$\max_{x\in[0,1]} |g'(x)| \leq g'(1) < 1$$

Le maximum de $|g'|$ sur $[0,1]$ est strictement inférieur à 1, donc d'après le cours (théorème du point fixe), la suite (x_n) converge vers x^* unique solution de l'équation $x^* = g(x^*)$ sur $[0,1]$.

Question 9. *Compléter l'algorithme qui permet de calculer x_n :*

```
function [ res ] = pointfixe(x0,n)
  res = x0;
    for i=1:n,
      res = cos(y);
    end
endfunction
```

Question 10. *On note $e_n = |x_n - x^*|$ l'erreur de la méthode après n itérations. La figure présente l'évolution de e_{n+1}/e_n en fonction de n. Pourquoi est-ce un graphique intéressant ? Pouvait-t-on prédire ce comportement ?*



Vitesse de convergence de la méthode du point fixe appliquée à la fonction $cos(x)$.

Ce graphique est intéressant, il permet de montrer que e_{n+1}/e_n admet une limite réelle strictement positive lorsque n tend vers l'infini donc que la méthode utilisée est d'ordre 1. Ce comportement était prévisible. En effet puisque $g'(x^*) \neq 0$, on démontre de manière théorique que la méthode est d'ordre 1.