EISTI - DÉPARTEMENT MATHÉMATIQUES

EXAMEN D'ANALYSE NUMÉRIQUE

29 mai 2006 – **DURÉE 3h00**

Exercice 1 – Une idée d'Archimède

En 250 av. J.C., Archimède a estimé le nombre π comme suit : Dans un cercle de diamètre 1 et donc de circonférence π, il a inscrit un carré. Le périmètre du carré est inférieur à la circonférence du cercle, et par conséquent Archimède a obtenu ainsi une limite inférieure pour π. Ensuite il a considéré un octogone, un 16-gone et ainsi de suite, en doublant chaque fois le nombre de côtés du polygone . De cette façon il obtenait des estimations de plus en plus meilleures pour π. Finalement il a montré que $\frac{223}{71} < \pi < \frac{223}{70}$ ce qui correspond à une estimation avec une erreur comprise entre 10^{-4} et 10^{-2}.

Traduisons l'idée d'Archimède à notre langage d'analyse numérique. Le périmètre p_n d'un polygone de 2^n côtés inscrit dans un cercle de diamètre 1 est donné par la formule

$$\begin{aligned} p_{n+1} &= 2^n \sqrt{2\left(1-\sqrt{1-\left(\frac{p_n}{2^n}\right)^2}\right)} \qquad (1.1) \\ \text{avec } p_2 &= 2\sqrt{2} \end{aligned}$$

Si on exécute avec Scilab la relation (1.1) pour $n = 3, 4, \ldots, 30$ on obtient les résultats donnés par le tableau suivant :

n	Valeurs de p_n						
3 - 9	2.8752328	2.8870061	2.889954	2.8906912	2.8908755	2.8909216	2.8909332
10 - 16	2.890936	2.8909368	2.8909369	2.890937	2.890937	2.890937	2.8909371
17 - 23	2.8909371	2.8909364	2.8909417	2.8909628	2.8909206	2.8912584	2.8912584
24 - 30	2.8939592	2.9154759	3.	2.8284271	4.	0.	0.

Manifestement nos ordinateurs modernes ne peuvent pas faire aussi bien qu'Archimède faisait avec la règle et le compas !

Heureusement que l'analyse numérique peut venir en aide aux archimèdes en herbe. Les deux questions suivantes vous permettront de faire mieux qu'Archimède.

1. D'abord il faut trouver la cause de l'annulation de p_{n+1} pour $n > n_0$. À votre avis quelles sont les raisons de cette annulation ?

   SOL. La quantité $\frac{p_n}{2^n}$ devient négligeable pour $n \geq 29$. (En effet $\frac{p_{29}}{2^{29}} = \frac{4}{2^{29}} = 7.45 \times 10^{-9}$). Donc $1 - \sqrt{1 - \left(\frac{p_n}{2^n}\right)^2} \simeq 1 - 1 = 0$.

2. Après la cause, le remède. La formule (1.1) n'est pas bonne pour un calcul numérique à l'aide d'un ordinateur. Une idée serait de poser

$$q_{n+1} = 2\left(1 - \sqrt{1 - \left(\frac{p_n}{2^n}\right)^2}\right)$$

   et d'examiner s'il est possible d'exprimer p_{n+1} en fonction de q_{n+1}.

   (a) Exprimer p_{n+1} en fonction de q_{n+1}.
   SOL. L'idée est d'éviter de diviser par 2^n. Nous allons donc remplacer le calcul de $2\left(1 - \sqrt{1 - \left(\frac{p_n}{2^n}\right)^2}\right)$ par une formule recursive.
   On pose

$$q_{n+1} = 2\left(1 - \sqrt{1 - \left(\frac{p_n}{2^n}\right)^2}\right)$$

   On a donc

$$p_{n+1} = 2^n \sqrt{q_{n+1}}$$

   (b) Exprimer q_{n+1} en fonction de q_n.
   SOL. La relation précédente permet d'écrire

$$\begin{aligned} q_{n+1} &= 2\left(1 - \sqrt{1 - \left(\frac{2^{n-1} q_n}{2^n}\right)^2}\right) \\ &= 2\left(1 - \sqrt{1 - \frac{q_n}{4}}\right) \\ &= 2\frac{\left(1 - \sqrt{1 - \frac{q_n}{4}}\right)\left(1 + \sqrt{1 - \frac{q_n}{4}}\right)}{\left(1 + \sqrt{1 - \frac{q_n}{4}}\right)} \\ &= \frac{q_n}{2 + \sqrt{4 - q_n}} \end{aligned}$$

   (c) En utilisant les deux résultats précédents montrer que la quantité p_n ne s'annule pas.
   SOL. Pour montrer que la quantité p_n ne s'annule pas il faut montrer que même si la quantité q_n devient négligeable, le terme $p_{n+1} = 2^n \sqrt{q_n}$ ne tend pas vers 0.

Nous pouvons d'abord remplacer le calcul initial de p_n par le suivant

$$p_{n+1} = 2^n \sqrt{q_{n+1}} \tag{1.2}$$

où

$$\begin{aligned} q_{n+1} &= \frac{q_n}{2+\sqrt{4-q_n}} && (1.3) \\ \text{avec } q_3 &= \frac{2}{2+\sqrt{2}} \end{aligned}$$

De ces formules, on conclut que $q_n > 0$.
Nous avons maintenant d'après (1.2) et (1.3) $p_{n+1} = 2^n \sqrt{q_{n+1}} = 2^n \sqrt{\dfrac{q_n}{2+\sqrt{4-q_n}}} \geq 2^n \sqrt{\dfrac{q_n}{2+\sqrt{4}}} = 2^{n-1}\sqrt{q_n} = p_n$. Donc $(p_n)_n$ est une suite croissante et $p_2 > 0$.

Exercice 2 – Décomposition $\mathbf{LDL}^\top$

1. Soit

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

Calculer la décomposition $\mathbf{LDL}^\top$ de $\mathbf{A}$.

SOL. On pose

$$\mathbf{D} = \begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix} \quad \mathbf{L} = \begin{pmatrix} 1 & 0 \\ c & 1 \end{pmatrix}$$

en multipliant on obtient par identification

$$\mathbf{A} = \begin{pmatrix} 2 & 1 \\ 1 & 0 \end{pmatrix} = \mathbf{LDL}^\top = \begin{pmatrix} a & ac \\ ac & ac^2+b \end{pmatrix}$$

donc

$$a = 2, \quad b = -\frac{1}{2}, \quad c = \frac{1}{2}$$

2. Existe-t-il une décomposition $\mathbf{LL}^\top$ de $\mathbf{A}$? Quelle est la condition que A ne vérifie pas ?

SOL. On reprend sans la matrice $\mathbf{D}$, avec $\mathbf{L} = \begin{pmatrix} a & 0 \\ b & c \end{pmatrix}$ on obtient

$$\mathbf{A} = \begin{pmatrix} 2 & 1 \\ 1 & 0 \end{pmatrix} = \mathbf{LL}^\top = \begin{pmatrix} a & b \\ 0 & c \end{pmatrix} \begin{pmatrix} a & 0 \\ b & c \end{pmatrix} = \begin{pmatrix} a^2+b^2 & bc \\ bc & c^2 \end{pmatrix}$$

Il y a une contradiction entre les égalités $bc = 1$ et $c^2 = 0$. La matrice $\mathbf{A}$ n'admet donc pas de décomposition $\mathbf{LL}^\top$.
Soit $x = (1, -2)^\top$. On a

$$\mathbf{Ax} \cdot \mathbf{x} = \begin{pmatrix} 2 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 1 \\ -2 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ -2 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ -2 \end{pmatrix} = -2 < 0$$

On remarque que $\mathbf{Ax} \cdot \mathbf{x} < 0$ donc la matrice $\mathbf{A}$ n'est pas définie positive. La condition nécessaire pour que $\mathbf{A}$ admette une décomposition $\mathbf{LU}$ n'est donc pas vérifiée.

3. Écrire l'algorithme de décomposition $\mathbf{LDL}^\top$.
SOL. Sachant que $\mathbf{A} = \mathbf{LDL}^\top$ on a

$$\forall \quad 1 \le i, j \le n, \quad A_{ij} = \sum_{k=1}^{n} L_{ik} D_{kk} \left(L^T_{kj} \right) = \sum_{k=1}^{n} L_{ik} D_{kk} L_{jk}$$

mais $\mathbf{L}$ est triangulaire inférieure donc $L_{ik} = 0, \forall k > i$ et $L_{jk} = 0, \forall k > j$ soit

$$A_{ij} = \sum_{k=1}^{\min(i,j)} L_{ik} D_{kk} L_{jk}$$

1ère colonne:
Pour $i = j = 1$ on a :

$$A_{11} = L_{11} D_{11} L_{11} = D_{11} \tag{1}$$

Pour $j = 1$ et $i \ge 1$:

$$A_{i1} = \sum_{k=1}^{1} L_{ik} D_{kk} L_{jk} = L_{i1} D_{11} L_{11} = L_{i1} D_{11} \Rightarrow L_{i1} = \frac{A_{i1}}{D_{11}} \tag{2}$$

Colonnes suivantes :
On suppose que l'on a calculé les p premières colonnes de $\mathbf{L}$ et de $\mathbf{D}$.
Pour $i = j = p + 1$ on a :

$$A_{p+1,p+1} = \sum_{k=1}^{p+1} L_{p+1,k} D_{kk} L_{p+1,k} = \sum_{k=1}^{p} \left(L_{p+1,k} \right)^2 D_{kk} + D_{p+1,p+1}$$

Les quantités $L_{p+1,k}$ ont déjà été calculées dans les colonnes précédentes, donc on obtient :

$$D_{p+1,p+1} = A_{p+1,p+1} - \sum_{k=1}^{p} \left(L_{p+1,k} \right)^2 D_{kk} \tag{3}$$

Pour $j = p+1$, $i > j$, on a :

$$A_{i,p+1} = \sum_{k=1}^{p+1} L_{ik} D_{kk} L_{p+1,k} \tag{4}$$

$$= \sum_{k=1}^{p} L_{ik} D_{kk} L_{p+1,k} + L_{i,p+1} D_{p+1,p+1} \underbrace{L_{p+1,p+1}}_{=1} \tag{5}$$

d'où

$$L_{i,p+1} = \frac{1}{D_{p+1,p+1}} \left(A_{i,p+1} - \sum_{k=1}^{p} L_{ik} D_{kk} L_{p+1,k} \right) \tag{6}$$

L'algorithme s'établit donc avec les formules précédentes :

```
Calcul de D_11 par 1
Pour i=2 à n faire
Calcul de L_i1 par 2
FPour
Pour j=2 à n faire
Calcul de D_jj par 3
Pour i=j+1 à n faire
Calcul de L_ij par 6
Fpour
Fpour
```

Exercice 3 – Approximation linéaire

Nous avons un ensemble $(\mathbf{x}_i)_{i=1,\ldots,n}$, des points du plan $\mathbb{R}^2$. On cherche à déterminer une droite représentée par $\mathbf{a}^\top \mathbf{x} = b$. L'approximation au sens des moindres carrés donne $\widehat{\mathbf{a}}, \widehat{b}$ comme valeurs des coefficients de la droite.

On pose $\mathbf{a} = \dfrac{\widehat{\mathbf{a}}}{\|\widehat{\mathbf{a}}\|_2}$ tel que $\|\mathbf{a}\|_2 = 1$ et $b = \dfrac{\widehat{b}}{\|\widehat{\mathbf{a}}\|_2}$

Admettons que l'approximation au sens de moindres carrées revient à chercher à minimiser l'erreur ε de l'approximation

$$\varepsilon = \min_{\substack{\mathbf{a} \in \mathbb{R}^2 \\ b \in \mathbb{R}}} \sum_{i=1}^{n} \frac{\left(\mathbf{a}^\top \mathbf{x}_i - b\right)^2}{\|\mathbf{a}\|_2^2}$$

1. Notons par $\mathbf{P} = [\mathbf{x}_1, \ldots, \mathbf{x}_n]^\top \in \mathbb{R}^{n\times 2}$ la matrice formée par les points du l'ensemble des points. Montrer que

$$\varepsilon = \min_{\substack{\mathbf{a}, b \\ \|\mathbf{a}\|_2^2 = 1}} \|\mathbf{P}\mathbf{a} - b\mathbf{1}_n\|_2^2$$

où $\mathbf{1}_n = [1, \ldots, 1]^\top \in \mathbb{R}^n$.

SOL. Soient $\widehat{\mathbf{a}}, \widehat{b}$, les valeurs optimales des paramètres, c'est-à-dire que l'on a

$$\varepsilon = \sum_{i=1}^{n} \frac{\left(\widehat{\mathbf{a}}^\top \mathbf{x}_i - \widehat{b}\right)^2}{\|\widehat{\mathbf{a}}\|_2^2}$$

On pose $\mathbf{a} = \dfrac{\widehat{\mathbf{a}}}{\|\widehat{\mathbf{a}}\|}$ tel que $\|\mathbf{a}\|_2 = 1$ et $b = \dfrac{\widehat{b}}{\|\widehat{\mathbf{a}}\|}$ et on obtient

$$\varepsilon = \sum_{i=1}^{n} \left(\frac{\widehat{\mathbf{a}}^\top}{\|\widehat{\mathbf{a}}\|_2} \mathbf{x}_i - \frac{\widehat{b}}{\|\widehat{\mathbf{a}}\|_2} \right)^2 = \sum_{i=1}^{n} \left(\mathbf{a}^\top \mathbf{x}_i - b\right)^2 = \|\mathbf{P}\mathbf{a} - b\mathbf{1}_n\|_2^2$$

ce qui prouve que

$$\min_{\substack{\mathbf{a}, b \\ \|\mathbf{a}\|_2^2 = 1}} \|\mathbf{P}\mathbf{a} - b\mathbf{1}_n\|_2^2 \leq \varepsilon$$

car on a obtenu le minimum avec des paramètres $\mathbf{a}$ de norme 1.

Nous avons aussi l'inégalité inverse puisque les paramètres $\mathbf{a}$ de norme 1 sont inclus dans l'ensemble des vecteurs $\mathbf{a}$ de norme quelconque. D'où finalement la relation cherchée.

2. Calculer cette même erreur minimale en utilisant la décomposition en valeurs singulières de la matrice $\mathbf{P}\mathbf{a} - b\mathbf{1}_n$.

   SOL. D'après le résultat précédent, on a

   $$\varepsilon = \min_{\substack{\mathbf{a}, b \\ \|\mathbf{a}\|_2^2 = 1}} \|\mathbf{P}\mathbf{a} - b\mathbf{1}_n\|_2^2$$

   En faisant la DVS de $\mathbf{P}\mathbf{a} - b\mathbf{1}_n$, on a

   $$\min_{\substack{\mathbf{a}, b \\ \|\mathbf{a}\|_2^2 = 1}} \|\mathbf{P}\mathbf{a} - b\mathbf{1}_n\|_2^2 = \mathbf{U}\boldsymbol{\Delta}\mathbf{V}^\top = \mathbf{U}\left[\sigma_1\right]\mathbf{V}^\top$$

   avec $\mathbf{U}, \mathbf{V}$ matrices orthogonales de taille $1xn$, et σ_1 valeur singulière de $\mathbf{P}\mathbf{a} - b\mathbf{1}_n$.

Exercice 4 – Du côté de chez Newton

On cherche à calculer l'inverse d'un nombre réel $\alpha > 0$

1. Écrire le terme général de la suite $(x_n)_{n \in \mathbb{N}}$ définie par la méthode de Newton appliquée à la fonction $f(x) = \frac{1}{x} - \alpha$
   $\left(\text{Rappel : La méthode de Newton est: } x_{n+1} = x_n - \dfrac{f(x_n)}{f'(x_n)}\right)$.

SOL. Le terme général de la suite $(x_n)_{n\in\mathbb{N}}$ définie par la méthode de Newton appliquée à la fonction

$$f(x) = \frac{1}{x} - \alpha$$

est donné par

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)} = 2x_n - \alpha x_n^2$$

2. On introduit la suite intermédiaire définie par $y_n = 1 - \alpha x_n$

(a) Donner l'expression de y_{n+1} en fonction de y_n.
SOL. L'expression de y_{n+1} en fonction de x_n est :

$$y_{n+1} = 1 - \alpha x_{n+1} = 1 - \alpha\left(2x_n - \alpha x_n^2\right)$$

Remarquant que $y_n = 1 - \alpha x_n \Leftrightarrow x_n = \frac{1-y_n}{\alpha}$ on obtient après substitution et simplification

$$y_{n+1} = y_n^2$$

(b) En déduire une condition sur x_0 pour que la suite x_n converge.
SOL. La condition sur y_0 pour que la suite y_n converge est que $|y_0| < 1$, sinon la suite est stationnaire (si $y_0 = 1$) ou divergente. Elle équivaut au fait que $0 < x_0 < \frac{\alpha}{2}$ d'après $x_n = \frac{1-y_n}{\alpha}$ et sachant que $\alpha > 0$.

(c) Établir l'expression de x_n en fonction de n, α, x_0.
SOL. L'expression de x_n en fonction de n, α, x_0 se déduit de $y_{n+1} = y_n^2$. En effet on a $y_n = y_0^{2^n}$ (démonstration facile par récurrence) dont on déduit que $1-\alpha x_n = (1-\alpha x_0)^{2^n}$ puis que $x_n = \frac{1}{\alpha}\left[1 - (1-\alpha x_0)^{2^n}\right]$

Exercice 5 – Runge-Kutta, est-ce vraiment nécessaire ?

Soit l'équation différentielle

$$\begin{aligned} y' &= f(x,y) \qquad (5.1)\\ \text{avec conditions initiales } y_0 &= y(x_0) \text{ et } y_1 = y(x_0+h) \end{aligned}$$

où $f : [a,b] \to \mathbb{R}$. On partage l'intervalle $[a,b]$ en n sous-intervalles égaux de longueur $h = \frac{b-a}{n}$. On utilise pour la résoudre, non pas la méthode de Runge-Kutta, mais la suivante :

$$y_{k+1} = y_k + \frac{h}{2}\left(3f(x_k,y_k) - f(x_{k-1},y_{k-1})\right)\ ;\ k \geq 0 \qquad (5.2)$$

qui est beaucoup plus simple.

Bien sûr nous savons que la simplicité seule ne suffit pas. Il faut aussi que l'erreur entre la valeur y_n obtenue en appliquant la méthode (5.2) et la valeur $y(x_n)$ obtenue en utilisant la solution analytique, soit bornée et petite.

On commence par doter la fonction f de bonnes propriétés. Par exemple f est lipschitzienne en y :

$$|f(x,y_1) - f(x,y_2)| \leq L|y_1 - y_2| \; ; \; y_1, y_2 \in \mathbb{R} \; ; \; x \in [a,b] \text{, et avec } L > 0 \text{ constante}$$

L'erreur dont nous voulons calculer sa borne supérieure est

$$\varepsilon(x_k; h) = y(x_k) - y_k \tag{5.3}$$

D'autre part on suppose que la dérivée numérique $\dfrac{y(x_{n+1}) - y(x_n)}{h}$ au point x_{n+1} est une bonne approximation de $y'(x_n)$. L'erreur de cette approximation s'écrit :

$$\begin{aligned} e(x_k; h) &= \frac{y(x_{k+1}) - y(x_k)}{h} - y'(x_{k+1}) \\ &= \frac{y(x_{k+1}) - y(x_k)}{h} - f(x_{k+1}, y(x_{k+1})) \end{aligned} \tag{5.4}$$

Posons $e_{\max}(h) = \max\limits_{k=1,\ldots,n} |e(x_k, h)|$.

Enfin, posons $y''_{\max} = \max\limits_{k=1,\ldots,n} |y''_k|$

1. Montrer que

$$y(x_{k+1}) - y_{k+1} = (y(x_k) - y_k) + h(y'(x_k) - y'_k) + \frac{h^2}{2} y''_k + he(x_k; h)$$

SOL. D'après (5.4) on a

$$y(x_{k+1}) = y(x_k) + hy'(x_k) + he(x_k; h)$$

et (5.2) s'écrit

$$y_{k+1} = y_k + \frac{h}{2}\left(3y'_k - y'_{k-1}\right)$$

On a donc

$$y(x_{k+1}) - y_{k+1} = (y(x_k) - y_k) + h\left[y'(x_k) - \frac{1}{2}\left(3y'_k - y'_{k-1}\right)\right] + he(x_k; h)$$

Il reste d'exprimer y'_{k-1} en fonction y'_k . On utilise le développement de Taylor. On a

$$y'_{k-1} = y'_k - hy''_k + O\left(h^2\right)$$

et donc

$$\begin{aligned} y(x_{k+1}) - y_{k+1} &= (y(x_k) - y_k) + h\left[y'(x_k) - \frac{1}{2}(3y'_k - y'_k + hy''_k)\right] + he(x_k; h) \\ &= (y(x_k) - y_k) + h(y'(x_k) - y'_k) + \frac{h^2}{2} y''_k + he(x_k; h) \end{aligned}$$

2. Montrer que

$$|\varepsilon(x_n;h)| \leq h\left(e_{\max}+\frac{h}{2}y''_{\max}\right)\sum_{l=0}^{n-1}(1+L)^l \tag{5.5}$$

   SOL. D'après la dernière relation, on a

$$\varepsilon(x_{k+1};h)=\varepsilon(x_k;h)+h\left[f(x_k,y(x_k))-f(x_k,y_k)\right]+\frac{h^2}{2}y''_k+he(x_k;h)$$

   et par conséquent

$$|\varepsilon(x_{k+1};h)| \leq (1+L)\,|\varepsilon(x_k;h)|+\frac{h^2}{2}y''_{\max}+he_{\max}(h)$$

   et finalement

$$|\varepsilon(x_n;h)| \leq h\left(e_{\max}+\frac{h}{2}y''_{\max}\right)\sum_{l=0}^{n-1}(1+L)^l$$

   car $\varepsilon(x_0;h)=0$.

3. Remarques concernant la qualité de l'approximation.

   SOL. On peut améliorer la qualité de l'approximation en diminuant la valeur de h. La seule limitation pour cette diminution est la valeur de eps.