EISTI - DÉPARTEMENT MATHÉMATIQUES

EXAMEN D'ANALYSE NUMÉRIQUE

février 2008 – **DURÉE 1h30**
Correction.

Exercice 1

Pour résoudre l'équation $ax^2 + 2bx + c = 0$ on doit calculer $\sqrt{b^2 - ac}$. Supposons que nous utilisons pour le codage des nombres en machine la norme IEEE-754 et le calcul de $m(x)$, étant donné la valeur x, se fait par troncature. Est-il possible que la vraie valeur de $b^2 - ac$ soit non négative et la valeur calculée négative ? Justifier votre réponse.

SOLUTION :

Notons $m(x)$ la valeur calculée de $b^2 - ac$. On a :

$$m(x) = m\left(m\left(b^2\right) - m(ac)\right)$$

D'après le cours on sait que tout nombre y est transformé en machine avec

$$m(y) = y(1+\eta) \text{ avec } |\eta| \leq \text{eps}$$

donc ici

$$m(x) = \left(m\left(b^2\right) - m(ac)\right)(1+\eta) \text{ avec } |\eta| \leq \text{eps}$$

Comme $|\eta| \leq \text{eps}$, on a $1 + \eta > 0$ donc $m(x)$ est du même signe que $m\left(b^2\right) - m(ac)$. Sachant que $b^2 - ac \geq 0 \Leftrightarrow b^2 \geq ac$, et que la troncature est une opération monotone, on a $m(b^2) \geq m(ac)$, soit $m\left(b^2\right) - m(ac) \geq 0$. On en déduit que $m(x) \geq 0$. Il n'est donc pas possible que la valeur calculée soit négative.

Exercice 2

Nous avons un ordinateur qui utilise comme base de numérotation β et dispose de p places pour la mantisse. On sait que si un nombre réel $x > 0$ peut être stocké dans cet ordinateur, alors il peut être représenté sous la forme $x = w \times \beta^n$, avec w nombre entier tel que $\beta^{p-1} \leq w < \beta^p$. Par exemple le réel 123.456 peut être représenté par 123456×10^{-3} et il peut être stocké dans un ordinateur avec $\beta = 2$ et $p = 18$, car $2^{17-1} = 65536 < 13456 < 2^{17} = 131072$. Un tel nombre on l'appellera nombre *digits p.racine* β.

Par la suite on notera $\text{ulp}(x) = \beta^n$ (ulp = Unit in the Last Place de x).

1. Donner la propriété commune de tous les nombres *digits p.racine* β en fonction de ulp.

   SOLUTION.- Si un réel a est un nombre *digits p.racine* β alors $a = w \times \beta^n = w \times \text{ulp}(a)$, avec $w \in \mathbb{N}$: le nombre a est un multiple de $\text{ulp}(a)$.

2. Montrer à l'aide d'un contre-exemple pour $\beta = 2$ que tout réel n'est pas un nombre *digits p.racine* β pour p fini.

   SOL.- Prenons $a = \frac{1}{3} = 0.333...$ Il n'existe pas de dernier chiffre significatif de a, donc a n'est pas un nombre *digits p.racine* β

3. Considérons maintenant deux réels a et b tels que $0 < b < a$ et qui sont des nombres *digits p.racine* β. Montrer que a est un multiple de $\text{ulp}(b)$.

   SOL.- On a

   $$a = w \times \text{ulp}(a) > b = w' \times \text{ulp}(b) \tag{1}$$

   Montrons qu'alors $\text{ulp}(a) \geq \text{ulp}(b)$. (Il ne me semble pas que ce soit évident mais je n'ai pas exigé la démonstration)

   Plusieurs cas peuvent se présenter a priori :

   (a) si $w = w'$ alors $\text{ulp}(a) > \text{ulp}(b)$, cqfd

(b) si $w < w'$, alors $\frac{1}{w} > \frac{1}{w'}$ donc en multipliant (1) par cette minoration on obtient $\mathrm{ulp}\,(a) \geq \mathrm{ulp}\,(b)$, cqfd

(c) si $w > w'$, raisonnons par l'absurde, et supposons que $\mathrm{ulp}\,(a) < \mathrm{ulp}\,(b)$ soit

$$\beta^{n_a} < \beta^{n_b} \text{ si on a } a = w \times \beta^{n_a} \text{ et } b = w \times \beta^{n_b}$$

Alors $n_a < n_b$ soit $n_a + 1 \leq n_b$. Alors

$$a = w \times \beta^{n_a} = w' \times \frac{w}{w'} \times \beta^{n_a} \tag{2}$$

Or on a

$$\beta^{p-1} \leq w' < w < \beta^p$$

soit

$$\frac{1}{w'} \leq \beta^{1-p} \text{ et } w < \beta^p$$

donc

$$\frac{w}{w'} < \beta$$

En reprenant la majoration (2), on obtient

$$a = w \times \beta^{n_a} = w' \times \frac{w}{w'} \times \beta^{n_a} < w' \times \beta \times \beta^{n_a} = w' \times \beta^{n_a+1} \leq w' \times \beta^{n_b} = b$$

or $a \leq b$ contredit l'hypothèse, donc on ne peut avoir que $\mathrm{ulp}\,(a) \geq \mathrm{ulp}\,(b)$.

$\mathrm{ulp}\,(a) \geq \mathrm{ulp}\,(b)$. donc on a $\mathrm{ulp}\,(a) = \beta^k$ et $\mathrm{ulp}\,(b) = \beta^l$ avec $k \geq l$. Donc $\mathrm{ulp}\,(a) = \beta^l \beta^{k-l}$ avec $\beta^{k-l} \in \mathbb{N}$, par conséquent, a est un multiple de $\mathrm{ulp}\,(b)$.

4. On continue avec les données de la question précédente. Démontrer que si $\frac{a}{b} \leq 2$, alors $a - b$ est aussi un nombre *digits p.racine* β.

   SOL.- Il faut montrer que $a-b$ est un multiple de $\mathrm{ulp}\,(a-b)$. On a $a \leq 2b \Rightarrow a - b \leq b$. Par conséquent si $a-b = w'' \times \beta^m$, on a (en reprenant la démonstration ci-dessus) $l \geq m$. De plus, on a toujours $k \geq l$ d'après ce qui précède, soit au final $k \geq l \geq m$. Alors $b - a = w \times \beta^k - w' \times \beta^l = w \times \beta^m \beta^{k-m} - w' \times \beta^m \beta^{l-m} = \left(w \times \beta^{k-m} - w' \times \beta^{l-m}\right) \beta^m$, avec $\beta^{k-m} \in \mathbb{N}$ car $k \geq m$, et $\beta^{l-m} \in \mathbb{N}$ car $l \geq m$ donc $b - a$ est un multiple de $\mathrm{ulp}\,(a - b)$.

Exercice 3
On s'intéresse à la série

$$S = 1 + \sum_{i=1}^{\infty} \frac{1}{i^2 + i}$$

dont la somme partielle d'ordre n est notée S_n :

$$S_n = 1 + \sum_{i=1}^{n} \frac{1}{i^2 + i}$$

1. En décomposant la fraction en éléments simples, montrer que

   $$S_n = 2 - \frac{1}{n+1}$$

   SOL : $\frac{1}{i^2+i} = \frac{1}{i} - \frac{1}{i+1}$ donc $\sum_{i=1}^{n} \frac{1}{i^2+i} = \sum_{i=1}^{n} \left(\frac{1}{i} - \frac{1}{i+1}\right) = 1 - \frac{1}{n+1}$ d'où le résultat

2. En déduire la valeur de S.

   SOL : le calcul de la limite donne

   $$\lim_{n\to\infty} \frac{1}{n+1} = 0$$

   donc

   $$\lim_{n\to\infty} S_n = 2$$

3. On peut calculer la valeur de la somme partielle S_n de deux façons. Soit en sommant dans l'ordre croissant des indices, dans ce cas on évalue :

$$S_{1,n} = 1 + \frac{1}{2} + \frac{1}{6} + \dots \frac{1}{n^2+n}$$

soit dans l'ordre décroissant des indices, dans ce cas on évalue

$$S_{2,n} = \frac{1}{n^2+n} + \dots + \frac{1}{6} + \frac{1}{2} + 1$$

Les résultats calculés avec un programme en langage C en simple précision sont résumés dans le tableau ci-dessous pour différentes valeurs de n.

n	$S_{1,n}$	$S_{2,n}$	$S_n = 2 - \frac{1}{n+1}$
99	$1,989999890$	$1,990000010$	$1,99$
999	$1,999001026$	$1,999000072$	$1,999$
9999	$1,999791980$	$1,999899983$	$1,9999$

(a) Quel est le calcul de somme partielle le plus précis ?
SOL : le calcul de $S_{2,n}$ se rapproche davantage de la valeur exacte $1,9999$ donc il est plus précis que celui de $S_{1,n}$.

(b) Donnez une explication et une justification théorique de ce résultat.
SOL : Lorsqu'on effectue les calculs dans $S_{1,n}$ on obtient des sommes de plus en plus grandes devant le terme qui leur est ajouté $\frac{1}{i^2+i}$. Celui-ci vient à être négligé puisqu'on ajoute un grand nombre avec un petit. Par contre lorsqu'on procède au calcul avec $S_{2,n}$ on ajoute progressivement des petits nombres, qui se cumulent progressivement de façon à ce que le phénomène précédent ne se produise pas.
On peut vérifier sur le cas de l'addition de trois nombres a, b et c tels que $a > b > c$, que faire $(a + b) + c$ n'est pas équivalent à faire $a + (b + c)$. (Cet exercice a été fait en TD).

Exercice 4
Considérons une matrice carrée $\mathbf{A} \in \mathbb{R}^{n\times n}$ et un vecteur $\mathbf{x} \in \mathbb{R}^n$.

1. Montrer que si $\mathbf{A}$ est régulière, il existe un réel $r > 0$ tel que

$$\|\mathbf{Ax}\| \geq r\,\|\mathbf{x}\|$$

et réciproquement.

Vérifier aussi que $r = \frac{1}{\|\mathbf{A}^{-1}\|}$.

Indication : On pourra montrer que pour tout vecteur $\mathbf{x}_0 \in \mathbb{R}^n$ tel que $A\mathbf{x}_0 = 0$, on a forcément $\mathbf{x}_0 = 0$.

SOL.-

Montrons que si $\mathbf{A}$ est regulière, alors il existe un réel $r > 0$ tel que

$$\|\mathbf{Ax}\| \geq r\,\|\mathbf{x}\|$$

On remarque tout d'abord que r ne peut pas dépendre de $\mathbf{x}$
$\mathbf{A}$ est régulière $\Rightarrow$ $\mathbf{A}^{-1}$ existe $\Rightarrow \forall \mathbf{x} \in \mathbb{R}^n$: $\|\mathbf{x}\| = \left\|\mathbf{A}^{-1}\mathbf{Ax}\right\| \leq \left\|\mathbf{A}^{-1}\right\| \|\mathbf{Ax}\| \Rightarrow \|\mathbf{Ax}\| \geq \frac{1}{\|\mathbf{A}^{-1}\|}\|\mathbf{x}\| = r\,\|\mathbf{x}\|$. donc $r = \frac{1}{\|\mathbf{A}^{-1}\|}$
Montrons maintenant le réciproque.

On a $\|\mathbf{Ax}\| \geq r\,\|\mathbf{x}\|$. Soit $\mathbf{x}_0 \in \mathbb{R}^n$ tel que $\mathbf{Ax}_0 = \mathbf{0}$. Nous allons démontrer que $\mathbf{x}_0 = \mathbf{0}$. On a $0 = \|\mathbf{Ax}_0\| \geq r\,\|\mathbf{x}_0\|$. Comme $r > 0$, il s'ensuit que $\mathbf{x}_0 = \mathbf{0}$. La seule solution de $\mathbf{Ax}_0 = \mathbf{0}$ étant $\mathbf{x}_0 = \mathbf{0}$. suffit à démontrer que $\mathbf{A}$ est régulière.

2. Supposons que la matrice $\mathbf{A}$ est régulière. Alors il existe un réel $\delta > 0$, tel que pour toute matrice $\mathbf{E}$ vérifiant $\|\mathbf{E}\| \leq \delta$, la matrice $\mathbf{A}+\mathbf{E}$ est régulière.

   Exprimer δ en fonction de $\left\|\mathbf{A}^{-1}\right\|$.

   Indication : Utiliser le résultat précédent.

   SOL.- $\mathbf{A}$ est régulière $\Rightarrow$ il existe un réel $r > 0$ tel que $\|\mathbf{Ax}\| \geq r\|\mathbf{x}\|$. Par ailleurs, nous avons

$$\|(\mathbf{A}+\mathbf{E})\,\mathbf{x}\| \geq \|\mathbf{Ax}\| - \|\mathbf{Ex}\| \tag{3}$$

   car si l'on prend deux matrices $\mathbf{A}$ et $\mathbf{B}$ on a

$$\begin{aligned} \|\mathbf{Ax}\| &= \|(\mathbf{A}+\mathbf{B}-\mathbf{B})\,\mathbf{x}\| \\ &= \|(\mathbf{A}+\mathbf{B})\,\mathbf{x}+(-\mathbf{B})\,\mathbf{x}\| \leq \|(\mathbf{A}+\mathbf{B})\,\mathbf{x}\| + \|-\mathbf{Bx}\| \\ &= \|(\mathbf{A}+\mathbf{B})\,\mathbf{x}\| + \|\mathbf{Bx}\| \end{aligned}$$

   soit

$$\|(\mathbf{A}+\mathbf{B})\,\mathbf{x}\| \geq \|\mathbf{Ax}\| - \|\mathbf{Bx}\|$$

   De (3) on déduit que $\|(\mathbf{A}+\mathbf{E})\,\mathbf{x}\| \geq \|\mathbf{Ax}\| - \|\mathbf{Ex}\| \geq \|\mathbf{Ax}\| - \|\mathbf{E}\|\,\|\mathbf{x}\| \geq (r - \|\mathbf{E}\|)\,\|\mathbf{x}\|$. Cette quantité est positive (et peut donc jouer le rôle du r dans la question précédente, et par la même entrainer que $\mathbf{A}+\mathbf{E}$ soit régulière) si $r - \|\mathbf{E}\| > 0$ soit $\|\mathbf{E}\| < r$, ce qui est vérifié si on prend un δ majorant $\|\mathbf{E}\|$ et tel que $\delta < r = \frac{1}{\|\mathbf{A}^{-1}\|}$.
   On a alors $\|(\mathbf{A}+\mathbf{E})\,\mathbf{x}\| \geq (r - \|\mathbf{E}\|)\,\|\mathbf{x}\|$

3. Supposons que la matrice $\mathbf{A}$ est regulière. Montrer que si la matrice $\mathbf{B}$ est singulière, alors

$$\frac{1}{cond\,(\mathbf{A})} \leq \frac{\|\mathbf{A}-\mathbf{B}\|}{\|\mathbf{A}\|}$$

   Indication : Utiliser le résultat précédent.

   SOL.- À la question précédente, nous avons montré que si $\|\mathbf{E}\| \leq \frac{1}{\|\mathbf{A}^{-1}\|}$, alors la matrice $(\mathbf{A}+\mathbf{E})$ est régulière. La contraposée de cette affirmation se lit : si $(\mathbf{A}+\mathbf{E})$ est singulière, alors $\|\mathbf{E}\| > \frac{1}{\|\mathbf{A}^{-1}\|}$

   Posons $\mathbf{E} = \mathbf{B}-\mathbf{A}$. Donc si $\mathbf{B} = \mathbf{A}+\mathbf{E}$ est singulière, on a $\frac{\|\mathbf{A}-\mathbf{B}\|}{\|\mathbf{A}\|} = \frac{\|\mathbf{E}\|}{\|\mathbf{A}\|} \geq \frac{1}{\|\mathbf{A}^{-1}\|\|\mathbf{A}\|} = \frac{1}{cond(\mathbf{A})}$.