Algorithmique numérique

Contrôle de connaissances - mercredi 27 mars 2013 à 10h.

Durée : deux heures.
Documents autorisés : notes prises en cours.

Dans toute la suite $\mathbb{K}$ désigne un corps et $\mathbb{F}$ désigne l'ensemble des nombres flottants en base 2 et précision p (sans contrainte sur la plage de l'exposant). Les notations $\mathrm{fl}(\cdot)$, $\mathrm{fl}_\downarrow(\cdot)$ et $\mathrm{fl}_\uparrow(\cdot)$, précisent le fait que l'opération entre parenthèses est effectuée respectivement en arrondi au plus proche, vers le bas ou vers le haut.

1. (a) Rappeler la définition de "matrice fortement régulière".
   (b) Soit $A \in \mathbb{K}^{n\times n}$ triangulaire. Montrer que A est fortement régulière si et seulement si A est inversible.
   (c) Soit $A \in \mathbb{K}^{n\times n}$ fortement régulière. En combien d'opérations dans $\mathbb{K}$ l'élimination de Gauss produit-elle le facteur U ? Quel est le surcoût arithmétique pour obtenir le facteur L ? Pourquoi ?

2. Montrer que le produit de deux matrices de dimensions $n \times n$ et à coefficients dans $\mathbb{K}$ n'est, asymptotiquement, pas plus coûteuse que l'inversion d'une matrice de dimensions $n \times n$ et à coefficients dans $\mathbb{K}$.

3. Montrez que le problème de déterminer le zéro d'une fonction f est équivalent au problème de chercher le point fixe d'une autre fonction que vous préciserez.

4. On recherche la racine positive de l'équation $x^2 - 4 = 0$: il s'agit de $x^* = 2$. Pour cela on utilise la méthode de Newton.
   (a) Écrivez l'itération.
   (b) Donnez une représentation graphique de la fonction $f : x \mapsto x^2 - 4$, dessinez quelques itérés en partant de $x_0 = 1$ et de $x_0 = -1$.
   (c) Calculez x_1 et x_2 en partant de $x_0 = 2$, de $x_0 = 1$, puis de $x_0 = 4$.
   (d) Montrez que l'itération peut s'écrire

   $$\begin{cases} x_0 \text{ donné}, \\ x_{n+1} = \frac{x_n}{2} + \frac{2}{x_n}. \end{cases}$$

   En utilisant l'arithmétique par intervalles, montrez que si $x_0 \in [\frac{1}{2}, 8]$ alors tous les itérés successifs appartiennent à $[\frac{1}{2}, 8]$.
   (e) On suppose désormais que $x_0 \in [\frac{1}{2}, 8]$. Calculez $|x_{n+1} - x^*|$ en fonction de $|x_n - x^*|$. Montrez sur cet exemple que la méthode de Newton est d'ordre 2.

5. Soit $x = (x_1, \ldots, x_n)^T \in \mathbb{F}^n$ ($n \geqslant 2$). On considère l'algorithme suivant, qui (vous allez le montrer) calcule un encadrement de $S := \sum_{i=1}^{n} x_i$.

$\hat{s}_1 = x_1$
for $i = 2 : n$
    $[\widehat{s}_i, e_i] = \mathrm{TwoSum}(\widehat{s}_{i-1}, x_i)$ // $\widehat{s}_i = \mathrm{fl}(\widehat{s}_{i-1} + x_i)$ et $\widehat{s}_i + e_i = \widehat{s}_{i-1} + x_i$
end
$\underline{c}_2 = e_2$
for $i = 3 : n$, $\underline{c}_i = \mathrm{fl}_\downarrow(\underline{c}_{i-1} + e_i)$
$\underline{s} = \mathrm{fl}_\downarrow(\widehat{s}_n + \underline{c}_n)$
$\overline{c}_2 = e_2$
for $i = 3 : n$, $\overline{c}_i = \mathrm{fl}_\uparrow(\overline{c}_{i-1} + e_i)$
$\overline{s} = \mathrm{fl}_\uparrow(\widehat{s}_n + \overline{c}_n)$
Renvoyer $(\underline{s}, \overline{s})$

(a) En supposant qu'un appel à TwoSum nécessite 6 opérations flottantes, quel est le nombre d'opérations flottantes effectuées par cet algorithme ?

(b) Montrez que l'on a $S = \widehat{s}_n + \sum_{i=2}^{n} e_i$.

(c) En déduire que $\underline{s} \leqslant S \leqslant \overline{s}$.

6. Rappeler la définition d'erreur relative, celle des deux modèles flottants standards (MS1 et MS2), et celle de γ_n. Quelle hypothèse implicite fait-on à chaque fois qu'on énonce des bornes d'erreur impliquant une quantité comme γ_n?

7. Soient $a, b, c, d \in \mathbb{F}$. On considère le calcul d'une approximation $\hat{x}$ de $x = ad - bc$ à l'aide de l'algorithme **A** ci-dessous.

A :
$\hat{w} \leftarrow \mathrm{fl}(bc)$;
$\hat{f} \leftarrow \mathrm{fl}(ad - \hat{w})$; $e \leftarrow \mathrm{fl}(\hat{w} - bc)$;
$\hat{x} \leftarrow \mathrm{fl}(\hat{f} + e)$;
Renvoyer $\hat{x}$

(a) Justifier brièvement l'égalité $e = \hat{w} - bc$.

(b) En appliquant MS1, montrer que l'erreur relative $|\hat{x} - x|/|x|$ est majorée par $\mathbf{B} = 2u + u^2 + u^2(1 + u)|bc|/|x|$ avec $u = 2^{-p}$.

(c) Donner une condition suffisante sur $|ad|/|bc|$ pour que l'erreur relative soit au plus $3u + O(u^2)$. Cette majoration est-elle satisfaisante lorsqu'on sait que ad et bc sont de signes opposés ?

(d) Soient $N = 2^p - 1$ et $(a, b, c, d) = (N - 1, N, N, N + 1)$. Pourquoi a, b, c, d sont-ils dans $\mathbb{F}$? En admettant que $bc = N^2$ et $\hat{x} = x = -1$, quelles conclusions cet exemple permet-il de tirer concernant la qualité de la borne d'erreur **B** et celle de l'algorithme **A** ?

* * *