Correction des exercices du TD 3

Exercice 3.1

Soient trois nombres réels positifs a, b, c avec $a > b > c$. On calcule leur somme de deux façons :

1. $S = (a + b) + c$
2. $S = a + (b + c)$

Quel résultat est le meilleur ?

SOLUTION :

Pour deux nombres quelconques a et b, la somme s est entachée d'une erreur d'entrée et d'une erreur de calcul. On a donc :

$$\eta(s) = \eta(a+b) = \eta^I(a+b) + \eta^C(a+b) \tag{1}$$

L'erreur d'entrée, inérente à la représentation du résultat de l'opération, se décompose en fonction des erreurs de même type commises sur les opérandes, à savoir :

$$\eta^I(a+b) = \frac{a}{a+b}\eta^I(a) + \frac{b}{a+b}\eta^I(b) \tag{2}$$

Donc (1) donne :

$$\begin{aligned}\eta(a+b) &= \eta^I(a+b) + \eta^C(a+b)\\ &= \frac{a}{a+b}\eta^I(a) + \frac{b}{a+b}\eta^I(b) + \eta^C(a+b)\end{aligned}$$

Si on utilise ce résultat pour la somme de trois réels, on obtient :

$$\begin{aligned}\eta\left((a+b)+c\right) &= \eta\left(s+c\right) = \eta^I(s+c) + \eta^C(s+c)\\ &= \frac{a+b}{a+b+c}\eta(a+b) + \frac{c}{a+b+c}\eta^I(c) + \eta^C(a+b+c)\end{aligned}$$

car le résultat de l'opération $a + b$ se récpercute au niveau de la valeur de s.

Or d'après (2)

$$\eta(a+b) = \frac{a}{a+b}\eta^I(a) + \frac{b}{a+b}\eta^I(b) + \eta^C(a+b)$$

donc

$$\eta\left((a+b)+c\right) = \frac{a+b}{a+b+c}\left[\frac{a}{a+b}\eta^I(a) + \frac{b}{a+b}\eta^I(b) + \eta^C(a+b)\right] + \frac{c}{a+b+c}\eta^I(c) + \eta^C(a+b+c)$$

d'où

$$\begin{aligned}\eta\left((a+b)+c\right) &= \frac{a}{a+b+c}\eta^I(a) + \frac{b}{a+b+c}\eta^I(b) + \frac{a+b}{a+b+c}\eta^C(a+b) + \frac{c}{a+b+c}\eta^I(c) + \eta^C(a+b+c)\\ &= \frac{1}{a+b+c}\left[a\eta^I(a) + b\eta^I(b) + c\eta^I(c) + (a+b)\eta^C(a+b) + (a+b+c)\eta^C(a+b+c)\right]\end{aligned}$$

Si on introduit

$$R_I = \max(\eta^I(a), \eta^I(b), \eta^I(c))$$

le maximum des erreurs d'entrée des trois constantes, et comme par ailleurs les erreurs de calculs sont majorées par l'epsilon machine :

$$\max\left\{\left|\eta^C(a+b)\right|, \left|\eta^C(a+b+c)\right|\right\} \leq eps$$

on obtient :

$$|\eta\left((a+b)+c\right)| \leq \frac{1}{a+b+c}\left[(a+b+c)R_I + (a+b)eps + (a+b+c)eps\right]$$

donc

$$|\eta\left((a+b)+c\right)| \leq R_I + \frac{2a+2b+c}{a+b+c}eps \tag{3}$$

Pour le calcul de l'erreur commise lorsqu'on calcule $S = a+(b+c)$ il suffit de faire une permutation circulaire :

$\begin{array}{ccc} a & b & c \\ \downarrow & \downarrow & \downarrow \\ b & c & a \end{array}$ ce qui conduit au résultat dédut de (3) suivant :

$$|\eta\left(a+(b+c)\right)| \leq R_I + \frac{2b+2c+a}{a+b+c}eps \tag{4}$$

Pour comparer les bornes supérieures, on doit donc comparer les quantités $2a+2b+c$ et $2b+2c+a$.

On a

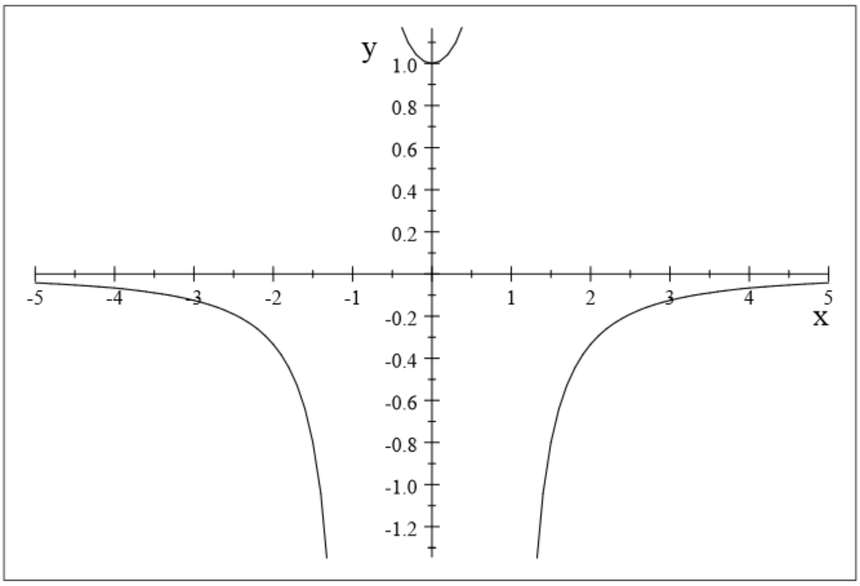
$$\begin{aligned} 2a+2b+c &< 2b+2c+a \\ &\Leftrightarrow 2a+c < 2c+a \\ &\Leftrightarrow a < c \end{aligned}$$

Or on est dans le cas où $a > c$ donc $2a+2b+c > 2b+2c+a$ donc la borne supérieure de $|\eta\left((a+b)+c\right)|$ est plus grande que celle de $|\eta\left(a+(b+c)\right)|$. On en déduit qu'il y a un risque que le calcul de $(a+b)+c$ produise une erreur plus grande que celui de $a+(b+c)$. En conclusion: lorsqu'on somme des nombres il faut commencer par les plus petits.

Exercice Supplémentaire

Calculer le nombre condition des fonctions :

1. $f(x) = \sqrt{x}$
2. $g(x) = \frac{1}{1-x^2}$



SOLUTION :

1. Par définition
$$\kappa = \left| \frac{f'(x)}{f(x)}.x \right|$$
Comme
$$f'(x) = \frac{1}{2\sqrt{x}}$$
on a
$$\kappa = \left| \frac{1}{2\sqrt{x}.\sqrt{x}}.x \right| = \frac{1}{2}$$
On en déduit que quel que soit l'intervalle d'étude, la fonction f est bien conditionnée.

2. Comme
$$g'(x) = \frac{2x}{(1-x^2)^2}$$
on a :
$$\kappa = \left| \frac{2x}{(1-x^2)^2}\left(1-x^2\right).x \right| = \left| \frac{2x}{1-x^2}.x \right| = \left| \frac{2x^2}{1-x^2} \right|$$
On s'aperçoit de suite que si x est proche de 1, le dénominateur peut exploser et entraîner un conditionnement très grand.

Exercice Supplémentaire

Soit la fonction f définie sur $_+$ par
$$f(x) = \sqrt{x+1} - \sqrt{x}$$

1. Décomposer en transformations élémentaires le calcul de f
2. Cacluler pour chaque trnasformation son nombre condition
3. Etablir la consition qui rend l'algorithme stable
4. En utilisant le résultat précédent, établir un autre algorithme qui est toujours stable.

SOLUTION

1. $\phi_1 : t_1 = \phi_1(x) = x+1$
   $\phi_2 : t_2 = \phi_2(x) = \sqrt{x}$
   $\phi_3 : t_3 = \phi_3(t_1) = \sqrt{t_1}$
   $\phi_4 : t_4 = \phi_4(t_2, t_3) = t_3 - t_2$
2.
$$\kappa_1 = \left| \frac{\phi_1'(x)}{\phi_1(x)}.x \right| = \frac{x}{x+1}$$
$$\kappa_2 = \left| \frac{\phi_2'(x)}{\phi_2(x)}.x \right| = 1/2$$
$$\kappa_3 = \left| \frac{\phi_3'(x)}{\phi_3(x)}.x \right| = 1/2$$
$$\kappa_4 = \left| \frac{1}{\phi_4(x)} J\phi_4(t_2,t_3).\begin{pmatrix} t_2 \\ t_3 \end{pmatrix} \right| = \frac{1}{t_3 - t_2}\begin{pmatrix} -1 & 1 \end{pmatrix}\begin{pmatrix} t_2 \\ t_3 \end{pmatrix} = \frac{t_3 - t_2}{t_3 - t_2} = 1$$
car $J\phi_4(t_2,t_3) = \begin{pmatrix} -1 & 1 \end{pmatrix}$
3. Il n'y a aucun cas dans lequel l'algorithme soit instable.

Exercice 3.2

Considérons le calcul $y = a^2 - b^2$ selon deux algorithmes :

Algorithme 1 : $y \leftarrow a^2 - b^2$

Algorithme 2 : $y \leftarrow (a-b)(a+b)$

En utilisant les formules de la propagation des erreurs :

1. Calculer l'erreur Δy pour le premier algorithme
2. Calculer l'erreur Δy pour le second algorithme
3. Trouver s'il y a un algorithme plus crédible que l'autre
4. Etudier la stabilité numérique des algorithmes

SOLUTION :

1. **Algorithme 1**

   (a) Etapes :

      1. $v \leftarrow a * a$
      2. $w \leftarrow b * b$
      3. $y \leftarrow v * w$

      L'algorithme comporte $r = 3$ étapes

   (b) Détail des trois étapes :

      1. $x^{(0)} = \begin{pmatrix} a \\ b \end{pmatrix} \to \phi^{(1)}\left(x^{(0)}\right) = \begin{pmatrix} a^2 \\ b \end{pmatrix} = x^{(1)}$
      2. $x^{(1)} = \begin{pmatrix} v \\ b \end{pmatrix} \to \phi^{(2)}\left(x^{(1)}\right) = \begin{pmatrix} v \\ b^2 \end{pmatrix} = x^{(2)}$
      3. $x^{(2)} = \begin{pmatrix} v \\ w \end{pmatrix} \to \phi^{(3)}\left(x^{(2)}\right) = v - w = x^{(3)}$

      On a bien $y = \phi^{(3)} \bigcirc \phi^{(2)} \bigcirc \phi^{(1)}(x)$

   (c) On calcule les différentes fonctions Ψ de la formule (1.8.4), $\Psi^{(k)}$ représente l'application qui reste à appliquer après avoir traité les k premières étapes.

   Comme $r = 3$, on calcule $\Psi^{(k)}$ pour $k = 1, 2, ...r - 1$ soit pour $k = 1, 2$.

   $\Psi^{(1)} = \phi^{(3)} \bigcirc \phi^{(2)}$ donc $\Psi^{(1)}\begin{pmatrix} v \\ b \end{pmatrix} = \phi^{(3)}\left[\phi^{(2)}\begin{pmatrix} v \\ b \end{pmatrix}\right] = \phi^{(3)}\begin{pmatrix} v \\ b^2 \end{pmatrix} = v - b^2$.

   $\Psi^{(2)} = \phi^{(3)}$ donc $\Psi^{(2)}\begin{pmatrix} v \\ w \end{pmatrix} = \phi^{(3)}\begin{pmatrix} v \\ w \end{pmatrix} = v - w$

   (d) On calcule les jacobiens des fonctions Ψ ainsi que celui de ϕ, et la valeur de Δx.

   $\Psi^{(1)} : IR^2 \to IR$ donc $J\Psi^{(1)}$ est une matrice de une ligne et deux colonnes. De même pour $\Psi^{(2)}$ et ϕ.

   $J\Psi^{(1)}\begin{pmatrix} v \\ b \end{pmatrix} = \begin{pmatrix} \frac{\partial \Psi^{(1)}}{\partial x} & \frac{\partial \Psi^{(1)}}{\partial y} \end{pmatrix} = \begin{pmatrix} 1 & -2b \end{pmatrix}$

   $J\Psi^{(2)}\begin{pmatrix} v \\ w \end{pmatrix} = \begin{pmatrix} \frac{\partial \Psi^{(2)}}{\partial x} & \frac{\partial \Psi^{(2)}}{\partial y} \end{pmatrix} = \begin{pmatrix} 1 & -1 \end{pmatrix}$

   $J\phi\begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} 2a & -2b \end{pmatrix}$ car $\phi : \begin{matrix} IR^2 \to IR \\ (a,b) \mapsto a^2 - b^2 \end{matrix}$

   Δx est l'erreur qui s'applique sur les paramètres d'entrée de l'algorithme : $\Delta x = \begin{pmatrix} \Delta a \\ \Delta b \end{pmatrix}$

(e) On calcule les matrices des erreurs $H_{k+1} = diag\left[\eta_i\left(m\left(x^{(k)}\right)\right)\right]$ = erreur élémentaire sur l'application $\phi^{(k+1)}$. Ces matrices sont toujours diagonales et comportent un seul élément non nul, qui est positionné au niveau de la composante sur laquelle porte la modification. Ainsi on a :

$H_1 = \begin{pmatrix} \eta_1(a*a) & 0 \\ 0 & 0 \end{pmatrix}, H_2 = \begin{pmatrix} 0 & 0 \\ 0 & \eta_2(b*b) \end{pmatrix}, H_3 = \eta_3(v-w) = \eta_3\left(a^2-b^2\right).$

On note que $\left|\eta_i\left(m\left(x^{(k)}\right)\right)\right| \leq \varepsilon$.

(f) On applique la formule (1.8.19), soit :

$$\begin{aligned}
\Delta y &\simeq J\phi(x).\Delta x + J\Psi^{(1)}\left(x^{(1)}\right).H_1.\left(x^{(1)}\right) + J\Psi^{(2)}\left(x^{(2)}\right).H_2.\left(x^{(2)}\right) + H_3.y \qquad (5)\\
&= \begin{pmatrix} 2a & -2b \end{pmatrix}.\begin{pmatrix} \Delta a \\ \Delta b \end{pmatrix} + \begin{pmatrix} 1 & -2b \end{pmatrix}.\begin{pmatrix} \eta_1(a*a) & 0 \\ 0 & 0 \end{pmatrix}.\begin{pmatrix} a^2 \\ b \end{pmatrix} + \\
&\quad \begin{pmatrix} 1 & -1 \end{pmatrix}.\begin{pmatrix} 0 & 0 \\ 0 & \eta_2(b*b) \end{pmatrix}.\begin{pmatrix} a^2 \\ b^2 \end{pmatrix} + \eta_3\left(a^2-b^2\right).\left(a^2-b^2\right) \\
&= 2a\Delta a - 2b\Delta b + \begin{pmatrix} 1 & -2b \end{pmatrix}.\begin{pmatrix} \eta_1(a*a).a^2 \\ 0 \end{pmatrix} + \begin{pmatrix} 1 & -1 \end{pmatrix}.\begin{pmatrix} 0 \\ \eta_2(b*b).b^2 \end{pmatrix} + \eta_3\left(a^2-b^2\right).\left(a^2-b^2\right) \\
&= 2a\Delta a - 2b\Delta b + \eta_1(a*a).a^2 - \eta_2(b*b).b^2 + \eta_3\left(a^2-b^2\right).\left(a^2-b^2\right)
\end{aligned}$$

avec $\left|\eta_i\left(m\left(x^{(k)}\right)\right)\right| \leq \varepsilon, i = 1, ...3$.

2. **Algorithme 2**

(a) Etapes :

1. $v \leftarrow a+b$
2. $w \leftarrow a-b$
3. $y \leftarrow v*w$

L'algorithme comporte $r=3$ étapes

(b) Détail des trois étapes :

1. $x^{(0)} = \begin{pmatrix} a \\ b \end{pmatrix} \rightarrow \phi^{(1)}\left(x^{(0)}\right) = \begin{pmatrix} a \\ b \\ a+b \end{pmatrix} = x^{(1)}$
2. $x^{(1)} = \begin{pmatrix} a \\ b \\ v \end{pmatrix} \rightarrow \phi^{(2)}\left(x^{(1)}\right) = \begin{pmatrix} a-b \\ v \end{pmatrix} = \begin{pmatrix} w \\ v \end{pmatrix} = x^{(2)}$
3. $x^{(2)} = \begin{pmatrix} w \\ v \end{pmatrix} \rightarrow \phi^{(3)}\left(x^{(2)}\right) = v*w = (a+b)(a-b) = x^{(3)}$

On a bien $y = \phi^{(3)} \bigcirc \phi^{(2)} \bigcirc \phi^{(1)}(x)$

(c) On calcule les différentes fonctions Ψ de la formule (1.8.4), $\Psi^{(k)}$ représente l'application qui reste à appliquer après avoir traité les k premières étapes.

Comme $r=3$, on calcule $\Psi^{(k)}$ pour $k = 1, 2, ...r-1$ soit pour $k = 1, 2$.

$\Psi^{(1)} = \phi^{(3)} \bigcirc \phi^{(2)}$ donc $\Psi^{(1)}\begin{pmatrix} a \\ b \\ v \end{pmatrix} = \phi^{(3)}\left[\phi^{(2)}\begin{pmatrix} a \\ b \\ v \end{pmatrix}\right] = \phi^{(3)}\begin{pmatrix} a-b \\ v \end{pmatrix} = v(a-b).$

$\Psi^{(2)} = \phi^{(3)}$ donc $\Psi^{(2)}\begin{pmatrix} w \\ v \end{pmatrix} = \phi^{(3)}\begin{pmatrix} w \\ v \end{pmatrix} = v*w$

(d) On calcule les jacobiens des fonctions Ψ ainsi que celui de ϕ, et la valeur de Δx.

$\Psi^{(1)} : IR^3 \rightarrow IR$ donc $J\Psi^{(1)}$ est une matrice de une ligne et trois colonnes. $\Psi^{(2)} : IR^2 \rightarrow IR$ donc $J\Psi^{(2)}$ est une matrice de une ligne et deux colonnes, de même pour ϕ.

$J\Psi^{(1)}\begin{pmatrix} a \\ b \\ v \end{pmatrix} = \begin{pmatrix} \frac{\partial\Psi^{(1)}}{\partial x} & \frac{\partial\Psi^{(1)}}{\partial y} & \frac{\partial\Psi^{(1)}}{\partial z} \end{pmatrix} = \begin{pmatrix} v & -v & a-b \end{pmatrix}$

$J\Psi^{(2)}\begin{pmatrix} v \\ w \end{pmatrix} = \begin{pmatrix} \frac{\partial\Psi^{(2)}}{\partial x} & \frac{\partial\Psi^{(2)}}{\partial y} \end{pmatrix} = \begin{pmatrix} v & w \end{pmatrix}$