DIRO
IFT 2425

EXAMEN INTRA

Max Mignotte

DIRO, Département d'Informatique et de Recherche Opérationnelle, local 2377
Http : //www.iro.umontreal.ca/∼mignotte/ift2425/
E-mail : mignotte@iro.umontreal.ca

Date : *20/02/2014*

I	Mesure d'Incertitude et Amplification d'Erreur (31 pts)
II	Erreur en Arithmétique Flottante (47 pts)
III	Méthode du Point Fixe/Newton (33 pts)
IV	Calcul Matriciel (12 pts)
Total	123 points.

TOUS DOCUMENTS PERSONNELS, CALCULATRICES ET CALCULATEURS AUTORISÉS

I. Mesure d'Incertitude et Amplification d'Erreur (31 pts)

1. Soit un triangle rectangle (cf. Fig. 1) pour lequel on a mesuré la base et la hauteur (en metre) comme étant :

$$L \;=\; 100 \;\text{ avec une erreur relative de }\; 5\%$$
$$H \;=\; \;\text{ dans l'intervalle de confiance :}[140,\; 150]$$

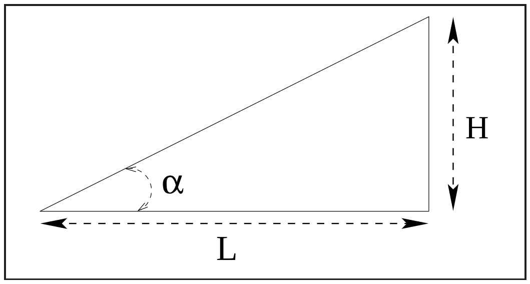

FIG. 1 –

   (a) Donner l'approximation de α (en radians) et l'erreur absolue commise sur α et servez-vous de ce calcul d'erreur pour souligner les chiffres significatifs "exacts" (cse) de la précédente approximation. Arrondir finalement cette approximation au nombre de cse adéquat et indiquer (en le justifiant) quelle est la variable H ou L dont l'amélioration de précision contribuerait le plus possible à l'amélioration de précision de α.
   <**10 pts**>

   (b) Peut-on utiliser la méthode de la fourchette pour obtenir cette incertitude $\Delta\alpha$? Sinon expliquer clairement pourquoi et si oui, utiliser cette méthode (de la fourchette) pour obtenir une estimation de $\Delta\alpha$.
   <**7 pts**>

2. Soit la suite itérative suivante : $x_{[n]} = 1 - n\, x_{[n-1]}$ $\quad$ (avec $x_{[0]} = 1 - \frac{1}{e}$) et n désignant l'itération.

   (a) Si $\Delta x_{[0]}$ désigne l'erreur absolue d'affectation que l'on fait sur le premier terme de la suite, trouver l'erreur que l'on fera au 12-ième terme de la suite (c-a-d si on itère la suite de $x_{[1]}, x_{[2]}, \ldots, x_{[12]}$).
   <**7 pts**>

   (b) Réécrire la suite itérative précédente à l'envers pour avoir une expression du style $x_{[n-1]} = f(x_{[n]})$. Si on part cette fois-ci de $x_{[24]}$ et soit $\Delta x_{[24]}$, l'erreur (absolue) d'affectation que l'on fait sur ce terme. Trouver l'erreur que l'on fera au 12-ième terme de la suite (c-a-d si on itère la suite de $x_{[24]}, x_{[23]}, \ldots, x_{[12]}$). En déduire ce qui se passerait si on partait de $x_{[24]}$ avec une valeur totalement aléatoire.
   <**7 pts**>

Réponse

$\boxed{\textbf{1.(a)}}$

Puisque $\Delta L/L = 5\%$, on a $\Delta L = 5$, d'une part et de plus $H \in [140\;,\;150] \Leftrightarrow H = H^{\star} \pm \Delta H = 145 \pm 5$.
<**2 pts**>

La relation qui lie α et les deux variables incertaines L et H est $\tan(\alpha) = H/L$

<**1 pt**> $$\text{donc :}\quad \alpha(H,L) = \arctan(H/L)$$

On a pour valeur approchée de α : $\alpha^\star(H,L) = \arctan(H^\star/L^\star) = \arctan(145/100) = 0.967047$ en radian.
<1 pt>

Pour calculer l'incertitude de α, on doit calculer la différentielle $\alpha(H,L)$

$$\Delta\alpha(H,L) = \left|\frac{1/L}{1+H^2/L^2}\right| \cdot \Delta H + \left|\frac{-H}{L^2+H^2}\right| \cdot \Delta L$$

<2 pts>

Numériquement, on obtient :

$$\begin{aligned}\Delta\alpha(H,L) &\approx \left|\frac{1/100}{1+145^2/100^2}\right| \cdot 5 + \left|\frac{-145}{100^2+145^2}\right| \cdot 5 \\ &\approx 0.016116 + 0.023368 \approx 0.0395 < 0.5 \times 10^{-1}\end{aligned}$$

<2 pts>

Ce dernier résultat numérique nous montre que la variable L contribue nettement plus à l'imprécision de α que la variable H. Donc si on veut améliorer la précision de α, il est judicieux d'améliorer prioritairement la précision de L.
<1 pt>

De plus, vue l'imprécision que l'on a sur α, on aura qu'un seul chiffre significatif dans le résultat de α donc on écrit $\alpha = \underline{0.9}7 \pm 0.05$.
<1 pt>

1.(b)

On a $\alpha(H,L) = \arctan(H/L)$ avec $H \in [140,\, 150]$ et $L \in [95,\, 105]$. La fonction $\alpha(H,L)$ est continue et monotone (croissante) lorsque H va croitre dans $[140,\, 150]$ et monotone (décroissante) lorsque L va croitre dans $[95,\, 105]$. On peut donc appliquer la méthode de la fourchette et :

$$\begin{aligned}\alpha(H,L) &\in [\arctan(H_{\min}/L_{\max})\,,\, \arctan(H_{\max}/L_{\min})] \\ &\in [0.9273 \;\; 1.0062]\end{aligned}$$

La connaissance de cet intervalle de confiance nous permet de trouver facilement $\Delta\alpha \approx 0.03947 < 0.5 \times 10^{-1}$, soit aiusi une estimation de $\alpha = \underline{0.9}7 \pm 0.05$ avec un cse.
<7 pts>

2.(a)

Lorsqu'on va calculer $x_{[1]} = g(x_{[0]}) = 1 - x_{[0]}$, on aura une amplification d'erreur du à la fonction $g(x) = 1 - x$ de $|\partial g/\partial x| = 1$ puis pour la deuxième itération, lorsqu'on calculera $x_{[2]} = g(x_{[1]}) = 1 - 2x_{[1]}$, une amplification d'erreur du à la fonction $g(x) = 1 - 2x$ de $|\partial g/\partial x| = 2$, puis pour la troisième itération, une amplification d'erreur du à la fonction $g(x) = 1 - 3x$ de $|\partial g/\partial x| = 3$, etc. La propagation des erreurs fera en sorte que Le calcul de $x_{[12]}$ sera entaché d'une erreur de $1 \times 2 \times 3 \times \ldots \times 12 \times \Delta x_{[0]} = 12! \times \Delta x_{[0]} = 479001600 \cdot \Delta x_{[0]}$.

On aurait put le prévoir en calculant le terme général de la suite qui est du style $x_{[n]} = \mathrm{Cst} + (-1)^n\, n! \cdot x_{[0]}$, amplifiant l'imprécision de $x_{[0]}$ de $n!$.
<7 pts>

2.(a)

On trouve facilement $x_{[n-1]} = \frac{1 - x_{[n]}}{n}$ et le même raisonnement que précédemment nous montre que de $x_{[24]}$ à $x_{[23]}$ on aura une atténuation d'erreur de 24 (ou une amplification d'erreur de 1/24), puis lorsqu'on passera du terme $x_{[23]}$ à $x_{[22]}$, une atténuation d'erreur de 23, etc. jusqu'au calcul de $x_{[12]}$ à partir de $x_{[13]}$ avec une atténuation d'erreur de 13, soit une atténuation d'erreur totale de $24 \times 23 \times 22 \times \cdots \times 13 = (24!/12!) = 1.29 \times 10^{15}$ (ou une amplification d'erreur de $12!/24! = 7.72 \times 10^{-16}$).

On pourrait donc partir de $x_{[24]}$ avec une valeur totalement aléatoire ou erronée, de toute façon avec cette stratégie on arriverait à $x_{[12]}$ avec une bonne estimation car l'erreur de la valeur de départ serait atténuée de 1.29×10^{15}.
<7 pts>

Nota : Dans le sens directe, en partant de $x_{[0]} = (1 - \frac{1}{e})$ on trouve, $x_{[11]} \approx 0.442$, $x_{[12]} \approx -4.31$, $x_{[13]} \approx 57.0$ bref, une estimation erronée qui oscille à cause de l'amplification des erreurs numériques. Si on part en sens inverse en prenant une valeur complètement erronée pour $x_{[24]}$ par exemple $x_{[24]} = 0$, on obtient...$x_{[13]} = 0.06694...$, $x_{[12]} = 0.0717732...$, $x_{[11]} = 0.07735...$ qui ne pose aucun problème de convergence. En effet on peut aussi partir de $x_{[24]} = 100$, et on obtient dans ce cas ...$x_{[13]} = 0.06694...$, $x_{[12]} = 0.0717732...$, $x_{[11]} = 0.07735...$ qui converge au même endroit ! malgré deux valeurs totalement différentes de $x_{[24]}$!

II. Erreur en Arithmétique Flottante (47 pts)

1. La façon évidente de calculer le module d'un nombre complexe $z = a + jb$ est d'utiliser la formule mathématique :
$$|z| = |a + jb| = \sqrt{a^2 + b^2}$$
   (a) Pour de très grandes valeurs de a et de b, quel type de problème numérique va t-on rencontrer ? Comment peut on réécrire cette expression pour atténuer ce problème numérique ?
   <5 pts>

   (b) Considérons cette fois-ci que a est très grand et b petit ou inversement. Quel type de problème numérique va t-on encore rencontrer ? Quel serait l'algorithme ou la procédure permettant de calculer cette expression en atténuant ce problème numérique ?
   <5 pts>

2. Expliquer pourquoi le calcul numérique de ces deux expressions

   (a) $\cos^2 x - \sin^2 x$

   (b) $\exp(x - y)$

   peuvent conduire à des problèmes d'erreurs numériques (identifier aussi pour quelle valeur de x). Identifier et citer le problème numérique associé et proposer une expression équivalente qui permettrait d'éviter ces problèmes numériques et/ou augmenter la précision des calculs de ces deux expressions.
   <8 pts>

3. Supposons que l'on désire évaluer dans un programme l'expression : $f(x) = (x - \sin(x))/x^3$ pour des valeurs de x proche de 0. Identifier le(s) problème(s) numérique(s) qui pourrait faire perdre de la précision à ce calcul et proposer une expression permettant d'atténuer ce problème. Trouver analytiquement la limite de cette expression quand x tend vers 0 et indiquer la valeur de x (approximativement en donnant par exemple le premier entier $n > 0$ de $x = 10^{-n}$) à partir de laquelle votre calculatrice vous donne une estimation de ce calcul avec plus de 50% d'erreur relative.
   <10 pts>

4. On aimerait évaluer avec plus de précision que celle donnée par un ordinateur ou une calculatrice l'expression : $E = \ln[\ln(10^9 + 1)] - \ln[\ln(10^9)]$.
   Utiliser pour cela le théorème de la valeur moyenne dont on rappelle l'énoncé :
   si f est continue sur $[a, b]$, $\exists\, c$ tq $[f(b) - f(a)] = f'(c) \cdot (b - a)$ et le fait que si $f'(x)$ est décroissante sur $[a, b]$ (à démontrer), alors $f'(a) > f'(c) > f'(b)$.
   <7 pts>

5. Dans la question I.3 du devoir n^{0}1 (point fixe accéléré & fractales de Mandelbrot) :

(a) on vous demandait d'implémenter deux expressions équivalentes mathématiquement, de l'accélération d'Aitken pour la convergence de la méthode du point fixe, soit :

$$c^{[3]} = c^{[0]} - \frac{(c^{[1]} - c^{[0]})^2}{(c^{[2]} - 2\,c^{[1]} + c^{[0]})} \qquad \text{ou} \qquad c^{[3]} = c^{[1]} - \frac{1}{\frac{1}{(c^{[2]} - c^{[1]})} - \frac{1}{(c^{[1]} - c^{[0]})}}$$

Discuter en le justifiant, laquelle de ces deux expressions (équivalente mathématiquement) est la moins sensible aux erreurs numériques lors de la convergence de la suite du point fixe.
<8 pts>

(b) Dans la question I.1, le point fixe sans accélération d'Aitken converge (pour la précision demandée) en 14 itérations. En moins de combien d'itérations faudrait-il que la méthode du point fixe accéléré par Aitken converge pour que cette technique soit avantageuse en terme de coût calculatoire ?
<4 pts>

Réponse

$\boxed{\textbf{1.(a)}}$

On aura inévitablement un problème d'OVERFLOW pour des nombres a ou b qui sont tous les deux très grand (supérieur à 10^{50} sur ma calculatrice). Pour éviter ce problème d'OVERFLOW, inévitable dans le cas de grands nombres, on peut re-écrire cette expression de la façon suivante :

$$|a + jb| = |a|\sqrt{1 + (b/a)^2}$$

<5 pts>

$\boxed{\textbf{1.(b)}}$

On comprend que l'expression précédente ne marche plus si par exemple $b > 10^{50}$ et si $b < 1$. Dans le cas ou $b = 10^{50}$ et $b = 1$, le carré dans la racine provoquerait un OVERFLOW. Pour résoudre ce problème, on doit donc considérer la procédure générale suivante :

$$|a + jb| = \begin{cases} |a|\sqrt{1 + (b/a)^2} & \text{si} \quad |a| \geq |b| \\ \\ |b|\sqrt{1 + (a/b)^2} & \text{si} \quad |a| < |b| \end{cases}$$

<5 pts>

Nota : Avec ses nouvelles expressions, on peut ainsi calculer le module d'un nombre supérieur à 10^{50} avec ma calculatrice. On est par contre toujours impuissant pour des nombres plus grand que 10^{99}.
Remarquez aussi qu'il existe un problème d'UNDERFLOW lorsque a et b vont être très petit (amplifié par le carré). Ce problème numérique est néanmoins moins grave que le problème d'OVERFLOW.

$\boxed{\textbf{2.}}$

Dans les deux cas, on aura un problème d'annulation de cse (chiffres significatifs exactes) due à la soustraction de deux nombres approximés quasi égaux. Pour éviter ce problème, on doit écrire ces relations de telle façon que plus aucune soustraction de deux nombres incertains (i.e., imprécis) quasi égaux apparaissent.

[-a-] Pour la première expression, cette annulation de cse se passe lorsque $x \to \pi/4$ (dans ce cas $\cos^2(\pi/4)$ et $\sin^2(\pi/4)$ tendent tout les deux vers $1/2$. Pour éviter cela, on peut utiliser les relations trigonométriques pour obtenir

$$\cos^2 x - \sin^2 x = \frac{1}{2}\big(1 + \cos(2x)\big) - \frac{1}{2}\big(1 - \cos(2x)\big) = \cos(2x)$$

<4 pts>

Nota : Si on avait ré-écrit cette expression, $\cos^2 x - \sin^2 x = 2\cos^2 x - 1$, (ou ; $1 - 2\sin^2(x)$) on se serait trouvé face à la soustraction de deux nombres quasi égaux dont un des nombre est approximé et l'autre à une représentation exacte. Bref, une configuration moins intéressant que $\cos(2x)$.

[-b-] Cette soustraction de deux nombres approximés quasi égaux se passera Lorsque $x \to y$. Dans ce cas, il vaut mieux la ré-écrire

$$\exp(x-y) = \exp(x)/\exp(y) \quad (\text{ou } \exp(x).\exp(-y))$$

<4 pts>

3.

Les problèmes sont, d'une part, la soustraction de deux nombres approximés quasi égaux au numérateur puis d'autre part, la forme numérique indéfinie 0/0. Pour atténuer ces problèmes, on pourrait décider que pour des valeurs proches de x, on décide d'évaluer cette expression avec un développement limité de $\sin x$ au voisinage de 0, c-a-d :

$$\frac{(x-\sin(x))}{x^3} = \frac{x - (x - \frac{x^3}{3!} + \frac{x^5}{5!})}{x^3} = \frac{1}{6} - \frac{x^2}{120}$$

Quand $x \to 0$, l'expression tend vers 1/6, ce que l'on aurait pu estimer aussi avec la règle de l'hôpital (appliqué trois fois) car

$$\lim_{x\to 0} \frac{(x-\sin(x))}{x^3} = \lim_{x\to 0} \frac{1-\cos(x)}{3x^2} = \lim_{x\to 0} \frac{\sin(x)}{6x} = \lim_{x\to 0} \frac{\cos(x)}{6} = \frac{1}{6}$$

Si on prend l'expression de départ, ma calculatrice donne pour $f(10^{-6}) = 0.171 \neq 0.1\overline{6}$ et $f(10^{-7}) = 0 \neq 0.1\overline{6}$ donc une erreur relative de plus de 50% d'erreur relative pour $x = 10^{-7}$.

<10 pts>

4.

Avec $f(x) = \ln\ln(x)$ sur $[a,b]$ avec $a = 10^9$ et $b = 10^9 + 1$, et $f'(x) = \frac{1}{x \ln(x)}$, alors le théorème de la valeur moyenne permet d'écrire

$$E = \frac{[f(b) - f(a)]}{b-a} = f'(c) = \frac{1}{c \ln(c)}$$

De plus $f''(x) = -\frac{\ln(x)+1}{x^2(\ln x)^2} < 0 \ \ \forall x \in [a,b]$ donc : $f'(b) < E = f'(c) < f'(a)$, donc numériquement

$$f'(b) = 4.8254942383 \times 10^{-11} < E < f'(a) = 4.8254942434 \times 10^{-11}$$

<7 pts>

Nota : On obtient une précision avec 7 chiffres après la virgules alors qu'avec ma calculatrice j'obtient une précision avec 3 chiffres après la virgule.

5.(a)

Lors de la convergence du point fixe, c-a-d lorsque $c^{[2]} \approx c^{[1]} \approx c^{[0]}$, on peut constater que la première expression se note aussi :

$$c^{[3]} = c^{[0]} - \frac{(c^{[1]} - c^{[0]})^2}{(c^{[2]} - c^{[1]}) + (c^{[0]} - c^{[1]})}$$

Dans cette expression, on a (à la convergence), la différence de deux nombres approximés quasi-égaux amplifié au carré au numérateur et la somme de deux nombres approximés quasi-égaux au dénominateur. A cela on doit ajouté, qu'à la convergence, le numérateur et dénominateur tendent tout les deux vers 0 et on se retrouve aussi avec une forme indéterminée numériquement du style 0/0 à la convergence de la suite de ce point fixe.

Dans la deuxième expression :

$$c^{[3]} = c^{[1]} - \overbrace{\frac{1}{\frac{1}{(c^{[2]}-c^{[1]})} - \frac{1}{(c^{[1]}-c^{[0]})}}}^{\text{2ième terme}}$$

on se retrouve pour le dénominateur du premier et deuxième terme du dénominateur, une différence de deux nombres approximés quasi-égaux. Cependant une fois inversée, ces deux expression donneront, dans le cas d'une perte de précision importante, une valeur très importante pour les deux termes (ou au moins pour l'un des deux), qui une fois inversé encore une fois donneront pour le deuxième terme une valeur presque nulle à $c^{[1]}$. En résumé, à la convergence ($c^{[3]} \approx c^{[2]} \approx c^{[1]}$), on ajoutera une valeur quasi-nulle à $c^{[1]}$, ce qui n'aura dtrictement aucune conséquence nuisible car on sera arrivé à convergence !
<8 pts>

5.(b)

L'accélération utilise deux évaluation de fonction par itération plus une expression d'accélération négligeable par rapport à une évaluation de fonction $g(;y_i)$. Cette méthode serait clairement avantageuse pour moins de 7 itérations. Le TP nous permet de constater que cette méthode accélérée permet une convergence en 4 itérations soit un coût calculatoire presque réduit par deux.
<4 pts>

III. Méthode du Point Fixe/Newton (33 pts)

1. **Méthode du Point Fixe**
   On se propose d'estimer numériquement une solution de l'équation
   $$\sin(x) = 0.2\,x^2 \tag{1}$$
   (a) Montrer qu'il existe une solution unique r pour cette Eq. (1) dans l'intervalle $J = [\pi/2, 2.5]$.
   <5 pts>
   (b) Montrer que la solution r de l'Eq. (1) est solution de l'équation $g_1(x) = x$ avec $g_1(x) = \sin(x) - 0.2\,x^2 + x$.
   <3 pts>
   (c) Montrer, que sur l'intervalle J, la suite $r_n = g_1(r_{n-1})$ converge vers une solution unique r, quel que soit le premier terme de la suite dans J.
   <5 pts>
   (d) Calculer les cinq premiers termes ; $r_1, \dots, r_5$ en partant de $r_0 = 2.5$.
   <5 pts>

2. **Méthode de Newton**
   (a) Donner la relation itérative $r_{n+1} = g_{\text{newt.}}(r_n)$ intervenant dans la méthode itérative de Newton pour la résolution itérative de la racine r de l'Équation (1).
   <5 pts>
   (b) Montrer que dans l'intervalle $J = [\pi/2, 2.5]$, la relation itérative de Newton converge.
   <5 pts>
   (c) Calculer les 5 premières estimées $r_1, \dots, r_4$, en partant de $r_0 = 2.5$.
   <5 pts>

Réponse

$\boxed{\textbf{1.(a)}}$

Trouver une solution de $\sin(x) = 0.2x^2$, c'est trouver une racine de la fonction $f(x) = \sin(x) - 0.2x^2$ qui est continue (comme somme deux fonctions continues).

$f'(x) = \cos(x) - 0.4\,x$, $f''(x) = -\sin(x) - 0.4$. Sur $J = [\pi/2, 2.5]$, $\sin(x) > 0$ et donc $-\sin(x) < 0$ et $f''(x) = -\sin(x) - 0.4x < 0$ sur J. Donc f' est décroissante de $f'(\pi/2) \approx -0.628$ à $f'(2.5) \approx -1.80$ (sur J) ; donc sur J, $f'(x) < 0$ et donc f est décroissante.

f est continue et monotone (décroissante) sur J, de plus $f(\pi/2) \approx 0.50$ et $f(2.5) \approx -0.65$ sont de signes opposés ce qui montre qu'une racine unique de f est dans J.
<5 pts>

$\boxed{\textbf{1.(b)}}$

$\sin(x) = 0.2x^2 \quad\Leftrightarrow\quad \sin(x) - 0.2x^2 = 0 \quad\Leftrightarrow\quad x = \sin(x) - 0.2x^2 + x$
<3 pts>

$\boxed{\textbf{1.(c)}}$

Pour la suite considérée, $g_1'(x) = \cos(x) - 0.4x + 1$. Sur $J = [\pi/2, 2.5]$ (contenant notre racine r, cf. 1.(a)) :

$$\begin{array}{rcl}
\cos(2.5) \approx -0.80114 < & \cos x & < 0 \\
\cos(2.5) \approx -0.80114 < & \cos x - 0.4\,x + 1 & < 1 - 0.2\,\pi \approx 0.37 \\
 & |g_1(x)| & < 0.80114
\end{array}$$

Avec la suite $r_n = g_1(r_{n-1})$, et un premier terme de notre suite dans J, la convergence est donc assurée.
<5 pts>

$\boxed{\textbf{1.(d)}}$

En partant de $r_0 = 2.5$, on a, $r_n = g_1(r_{n-1})$ et :

$$\begin{aligned}
r_1 &= 1.848472144 \\
r_2 &= 2.126797434 \\
r_3 &= 2.071516457 \\
r_4 &= 2.090517456 \\
r_5 &= 2.08442252
\end{aligned}$$

<5 pts>

Nota -1- : Avec une calculatrice CASIO, par exemple, la première suite itérative se programme facilement en affichant 2.5 puis $\boxed{\text{EXE}}$ puis en affichant $\sin(ANS) - 0.2ANS^2 + ANS$, chaque appel à la touche $\boxed{\text{EXE}}$ nous permet d'avoir une itération de la suite.

Nota -2- : Si on continue avec sa calculatrice, on obtient :

$$\begin{aligned}
r_6 &= 2.08642762 \\
\ldots &= \ldots \\
r_{19} &= 2.085934584
\end{aligned}$$

Et sur ma calculatrice, l'estimation se stabilise à partie de r_{19}.

$\boxed{\textbf{2.(a)}}$

La fonction $f(x)$ est dérivable sur $\mathbb{R}$ et la méthode itérative de Newton permet d'écrire :

$$r_{n+1} = r_n - \frac{f(r_n)}{f'(r_n)} = r_n - \frac{\sin(r_n) - 0,2\,(r_n^2)}{\cos(r_n) - 0.4\,r_n} = g_{\text{newt}}(r_n)$$

<5 pts>

2.(b)

Pour montrer que la suite itérative de newton précédente converge, le plus simple est de montrer que la fonction $f(x)$, sur l'intervalle $J = [\pi/2, 2.5]$ (intervalle dans lequel on prendra $r_0 = 2.5$ et dans lequel une racine unique s'y trouve) ne présente pas d'*extrema*, i.e., de valeurs pour laquelle, $f'(x)$ s'annule.

Dans notre cas, il suffit de se rappeler ce que l'on a fait en question 1.(a), c-a-d ;

$f'(x) = \cos(x) - 0.4\,x$, $f''(x) = -\sin(x) - 0.4$. Sur $J = [\pi/2, 2.5]$, $\sin(x) > 0$ et donc $-\sin(x) < 0$ et $f''(x) = -\sin(x) - 0.4x < 0$ sur J. Donc f' est décroissante de $f'(\pi/2) \approx -0.628$ à $f'(2.5) \approx -1.80$ (sur J) .... mais ne s'annule jamais ! donc aucun *extrema* et donc aucun problème !

<5 pts>

Nota : La démonstration utilisant le fait que la méthode de Newton est aussi une méthode du point fixe et nécessitant que $|g'(x) < 1|$ sur J est beaucoup trop difficile ici.

2.(c)

En partant de $r_0 = 2.5$, on a, $r_{n+1} = g_{\text{newt}}(r_n)$ et :

$$\begin{aligned} r_1 &= 2.138269902 \\ r_2 &= 2.08716606 \\ r_3 &= 2.085935309 \\ r_4 &= 2.085934584 \\ r_5 &= 2.085934584 \end{aligned}$$

<5 pts>

Nota : Sur ma calculatrice, l'estimation se stabilise à partir de r_4, ce qui est particulièrement rapide.

IV. Factorisation LU (12 pts)

1. Décomposer la matrice A en produit LU (sans permutation) par la méthode de la factorisation directe.
   <8 pts>

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 7 & 18 \\ 4 & 13 & 38 \end{pmatrix}$$

2. Imaginons qu'un analyste numéricien doit résoudre un système d'équations d'ordre $n = 90000$ sur une machine 64 bits (les coefficients de la matrice sont exprimées en flottant et on rappelle qu'un flottant (FLOAT) correspond à 8 octets sur une machine 64 bits). Il dispose d'un microordinateur possédant 64 Giga-Octet de mémoire vive. Il hésite entre utiliser la méthode de factorisation directe ou la méthode de pivot de Gauss pour décomposer le système en matrice LU puis résoudre ce système par substitution avant et arrière. Il aimerait que tout cela se fasse le plus rapidement possible. Que devrait-il choisir ? Justifier bien votre réponse.
   <4 pts>

Réponse

$\boxed{\textbf{1.}}$

Par factorisation directe, on trouve,

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 7 & 18 \\ 4 & 13 & 38 \end{pmatrix} = \underbrace{\begin{pmatrix} 1 & 0 & 0 \\ 2 & 3 & 0 \\ 4 & 5 & 6 \end{pmatrix}}_{L<\mathbf{4\ pts}>} \underbrace{\begin{pmatrix} 1 & 2 & 3 \\ 0 & 1 & 4 \\ 0 & 0 & 1 \end{pmatrix}}_{U<\mathbf{4\ pts}>}.$$

<8 pts>

$\boxed{\textbf{2.}}$

Un système d'équations d'ordre $n = 90000$ sur une machine 64 bits (avec 8 octets par nombre flottant) nécessite une structure de données de type tableau 2D de 60 Giga-Octets par matrice (sur une machine 64 bits). Ne disposant que de 64 Giga-Octets, l'analyste numéricien aura donc le choix entre la méthode de pivot de Gauss qui va résoudre le système en $O(n^3/3)$ mais qui devra disposer de 60×3 Giga-Octets pour stocker les trois matrices, c-a-d dans notre cas de 120 Giga-Octets de SWAP ou la méthode de Factorisation directe qui résout le système en $O(n^3)$ avec stockage compacte des matrices LU donc ne nécessitant que 60 Giga-Octets de mémoire vive (sans donc avoir besoin de SWAPPER). Le plus intéressant est donc la méthode de factorisation directe malgré le fait qu'elle nécéssite plus de calcul.

<4 pts>