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 : *16/02/2011*

I	Mesure d'Incertitude et Amplification d'Erreur (32 pts)
II	Erreur en Arithmétique Flottante (35 pts)
III	Méthode du Point Fixe et de Newton (33 pts)
IV	Factorisation LU (17 pts)
Total	117 points.

Tous documents personnels, calculatrices et calculateurs autorisés

I. Mesure d'Incertitude et Amplification d'Erreur (32 pts)

Soit un tube dont la hauteur est H et le rayon R et dont la surface est donnée par

$$S = 2\pi RH$$

On suppose que pour les variables de cette formule, on a les mesures approximées et les incertitudes suivantes :

π	$= 3.14$	avec tous les chiffres donnés considérés comme significatif "exact" (cse)
R	$= 10.2$ mètres	avec une erreur absolue de 0.1 mètre
H	$\in [49.9\,,\,50.1]$	i.e., H est une valeur élément de cette intervalle de confiance

On vous demande de :

1. Exprimer, pour chacune de ces trois variables, l'incertitude que l'on a sur elle en la mettant sous la forme $x = x^\star \pm \Delta x$ où $x^\star$ et Δx seront déterminés et en soulignant le nombre de chiffres significatifs "exact" (cse) pour $x^\star$.
   <5 pts>
2. Donner l'approximation de cette surface S, et l'incertitude sur S (i.e., ΔS) par la méthode de propagation d'erreur (utilisant l'approximation de Taylor de la fonction au premier ordre).
   <5 pts>
   Souligner les chiffres significatifs "exacts" (cse) de ce résultat. Arrondir cette approximation aux nombres de cse adéquat.
   <2 pts>
3. Quelle est la variable qui engendre ici le plus d'incertitude (ou erreur absolue) sur le résultat final ? Justifier votre réponse.
   <3 pts>
   Si on avait eu une précision infinie pour π, aurait-on eu plus de chiffres significatifs ? Justifier votre réponse.
   <3 pts>
4. Peut on utiliser la méthode de la fourchette pour obtenir cette incertitude ? Si oui expliquer pourquoi et utiliser cette méthode (de la fourchette) pour obtenir l'incertitude résultante (i.e., erreur absolue) sur S que l'on obtiendrait par cette méthode et le nombre de chiffres significatifs "exact" (cse) pour la valeur approximée de S.
   <5 pts>
   Devrait-on retrouver le même résultat que celui de la question 2. (de cet exercice I.) ou non ? Justifier votre réponse.
   <2 pts>
5. Dans le cours, on a vu que l'erreur relative sur la multiplication (ou la division) de deux valeurs (approximées) est ... et je vous laisse retrouver la suite. Appliquer ce corollaire pour retrouver l'incertitude sur S.
   <5 pts>
   Devrait-on retrouver le même résultat que celui de la question 2 ou non ? Justifier votre réponse.
   <2 pts>

Réponse

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

Le chiffre des centièmes “4” (de rang -2) nous permet de trouver une borne supérieure de l’erreur absolue ou de l’incertitude que l’on a sur la constante π, i.e., $\Delta\pi = 0.5 \times 10^{-2}$ et donc d’écrire $\pi = \underline{3.14} \pm 0.005$. <**2 pts**>

Puisque $\Delta R = 0.1 < 0.5 \times 10^0$, le rang du dernier chiffre significatif est 0, correspondant au chiffre des unités, ce qui nous permet d’écrire $R = \underline{10}.2 \pm 0.1$. <**1 pt**>

Finalement, la connaissance d’un intervalle de confiance pour la mesure de H nous permet d’écrire ; d’une part que la valeur approximée pour cette variable sera le centre de cette intervalle et d’autre part qu’une borne supérieure de son erreur absolue est la demi largeur de cet intervalle, i.e., $\Delta H = 0.1$. Comme $\Delta H < 0.5 \times 10^0$, cela nous permet aussi d’écrire aussi que $H = \underline{50} \pm 0.1$. <**2 pts**>

$\boxed{\mathbf{2.}}$

On a pour valeur approchée de S

$$S^* = 2 \times 3.14 \times 10.2 \times 50 = 3202.8 \quad \mathrm{m}^2 \qquad < \textbf{2 pts} >$$

Pour calculer l’incertitude de S, on doit calculer la différentielle $\Delta S(\pi, R, H)$, i.e.,

$$\Delta S(\pi, R, H) = |2R\,H|\ \Delta\pi\ + |2\pi\, H|\ \Delta R\ + |2\pi\, R|\ \Delta H \qquad < \textbf{2 pts} >$$

Soit

$$\begin{aligned}
&\Delta S(\pi, R, H) = |2 \times 10.2 \times 50| \times 0.005\ + |2 \times 3.14 \times 50| \times 0.1 + |2 \times 3.14 \times 10.2| \times 0.1\\
&= 5.1 + 31.4 + 6.4056\\
&\approx 42.9056 \qquad < \textbf{2 pt} >
\end{aligned}$$

On obtient donc $S = 3202.8 \pm 42.9$. Comme $42.9 < 0.5 \times 10^2$, le rang du dernier chiffre significatif est 2 (i.e, les centaines) et on peut écrire $S = \underline{32}02 \pm 42.9$. Si on arrondi finalement cette estimation aux nombres de cse, adéquat, on obtient

$$S = \underline{32}\ \times 10^2 \pm 43 \quad (\text{en}\ \ \mathrm{m}^2). \quad < \textbf{1 pt} >$$

$\boxed{\mathbf{3.}}$

Le calcul précédent (i.e., $42.9 \approx 5.1{+}31.4{+}6.4056$) nous montre que c’est l’imprécision sur R qui engendre le plus d’erreur d’imprécision sur le résultat. <**3 pts**>

Si on avait eu une précision infinie pour π, i.e., $\Delta\pi = 0$, on aurait eu $\Delta S = 31.4 + 6.4056 \approx 37.4$, ce qui n’aurait rien changé au nombre de chiffres significatifs du résultat puisque $37.4 < 0.5 \times 10^2$; le rang du dernier chiffre significatif aurait toujours été les centaines. <**3 pts**>

$\boxed{\mathbf{4.}}$

La méthode de la fourchette peut être utilisée ici car la fonction $S(\pi, R, H) = 2\pi \times R \times H$ est strictement monotone sur chacun des intervalles dans lequel se trouve les variables incertaines π, R, et H. Il est donc possible, en utilisant les valeurs minimales pour π, R et H de déterminer la borne minimale pour S et inversement. <**1.5 pts**>

$$\begin{aligned}
S_{\min} &= 2 \times (3.14 - 0.005) \times (10.2 - 0.1) \times (50 - 0.1) = 3160.0173\\
S_{\max} &= 2 \times (3.14 + 0.005) \times (10.2 + 0.1) \times (50 + 0.1) = 3245.8287
\end{aligned}$$

<**1.5 pts**>

La valeur approximée est donc le centre de cette intervalle de confiance et l’incertitude sur S serait la demi largeur de cette intervalle. Ce qui nous permet d’écrire $S \approx 3202.923 \pm 42.9057$. et puisque $42.9 < 0.5 \times 10^2$, $S = \underline{32}\ \times 10^2 \pm 43$. <**2 pts**>

Il est normal de trouver un résultat différent de la question 2. car la méthode de la fourchette (quand celle-ci est applicable) est exacte contrairement à la méthode de l'approximation de Taylor. On peut cependant constater que la méthode de Taylor est une excellente approximation. $<$**2 pts**$>$

$\boxed{\textbf{5.}}$

Dans le cours, on a vu que l'erreur relative sur la multiplication (ou la division) de deux valeurs (approximées) est la somme des erreures relatives. $<$**2 pts**$>$

Dans le cas de $S(\pi, R, H)$, l'erreur relative sur S est donc la somme des trois erreures relatives asociées à chacune de ces 3 valeurs incertaines. On a donc

$$\frac{\Delta S(\pi, R, H)}{S^\star} = \frac{\Delta S(\pi, R, H)}{3202.8} = \frac{0.005}{3.14} + \frac{0.1}{10.2} + \frac{0.1}{50} = 0.013396278$$

donc $\Delta S(\pi, R, H) = 0.013396278 \times 3202.8 \approx 42.9056$ $<$**3 pts**$>$

Devrait-on retrouver le même résultat que celui de la question 2 ? Oui ! car ce corollaire vient directement de l'application de la formule de propagation d'erreur utilisant l'approximation de Taylor au premier ordre ! $<$**2 pts**$>$

II. Erreur en Arithmétique Flottante (35 pts)

1. On voudrait implémenter la fonction SOFTMAX, aussi connue sous le nom de fonction MULTINOMIAL LOGIT ou FONCTION LOGISTIQUE MULTIPLE :

   ALGORITHM ; SOFTMAX

   | $[x_1, x_2, \ldots, x_n]$ Tableau de flottants positifs

   • A partir du tableau $[x_1, x_2, \ldots, x_n]$
   on veut obtenir :

   • $\left[\dfrac{\exp(x_1)}{Z}, \dfrac{\exp(x_2)}{Z}, \ldots, \dfrac{\exp(x_n)}{Z}\right]$

   Avec $Z = \exp(x_1) + \exp(x_2) + \ldots \exp(x_n)$

   Cette fonction est souvent utilisée pour transformer un vecteur de flottants (stockant des mesures d'énergie ou les sorties d'un réseau de neurones (fonction d'activation) ou d'un modèle de Markov caché ou de regression logistique) en vraies probabilitées ou plus précisement en valeurs discrétisées de mesures de densité de probabilité (dont la somme intègre donc à un). Cela est précieux pour interpréter les résultats obtenus en différentes classes probabilistes ou pour d'autres raisons.

   Si on n'y prend pas garde, et même si on considère un tableau de DOUBLE au lieu de FLOAT, l'implémentation de cette fonction va générer des problèmes numériques qui pourraient s'avérer catastrophiques. Identifiez le (ou les) problème(s) numérique(s) qui serait l'origine de ce bug et proposez une solution pour y remedier en proposant un pseudo-code sans erreur numérique qui implémenterait correctement cette fonction en justifiant bien votre réponse.
   $<$**10 pts**$>$

2. On cherche à évaluer numériquement l'intégrale suivante :

$$I_n = \int_0^1 \frac{x^n}{10 + x}\, dx$$

pour des valeurs de $n \geq 12$. Pour cela le calcul de I_0 et une intégration par partie nous permet de trouver la relation de récurence suivante :

$$\begin{cases} I_0 & = \ln\left(\frac{11}{10}\right) \\ \\ I_n & = \frac{1}{n} - 10\, I_{n-1} \end{cases}$$

(a) Retrouver I_0 et la relation de récurence $I_n = \frac{1}{n} - 10\, I_{n-1}$ par calcul analytique
<**6 pts**>

(b) Montrer que, pour de forte valeur de n (par exemple $n \geq 12$), cette méthode d'estimation conduit à une catastrophe numérique.
<**6 pts**>

(c) En supposant que l'on fasse une erreur absolue de ΔI_0 sur I_0 et que l'on s'apercoive expérimentalement que cette suite itérative sur ordinateur diverge completement à partir de I_8, i.e., en utilisant des FLOATS, on obtient le double de ce que l'on obtient en DOUBLE pour cette itération. Peux t-on donner une grossiere estimation de ΔI_0 lorsque I_0 est représenté en FLOAT ?
<**4 pts**>

3. Pour chacune de ces fonctions, il existe un point x_0 pour lequel une évaluation de $f(x)$ avec x proche (mais différent de x_0) pourrait causer des problèmes numériques.

   (a) $\frac{1}{1-x} - \frac{1}{1+x}$ pour $x \neq \pm 1$

   (b) $\sqrt{1+x} - \cos x$

   (c) $\sqrt{1+\cos x}$

   Trouver ce point x_0 et expliquer pourquoi (i.e., identifier et citer le problème numérique associé) et proposer une formule équivalente plus fiable qui permettrait d'éviter ces problèmes numériques et qui permettrait d'augmenter la précision des calculs de ces trois expressions dans un voisinage de x_0.
   <**9 pts**>

Réponse

$\boxed{\textbf{1}}$

Le problème numérique qui pourrait poser problème ici serait l'OVERFLOW. En effet, au dela d'une certaine valeur pour x, la fonction exponentielle n'est pas définie et va générer une erreur numérique d'OVERFLOW contaminant ainsi le calcul de Z (NaN ou INF ajouté à une suite de nombre) puis le calcul de chacune des valeurs du vecteur que l'on veut obtenir qui ne sommera plus à un. <**4 pts**>

Nota : Le problèm énoncé plus haut est beaucoup plus important dans ce cas particulier au problème de l'addition de nombres qui pourraient ici aussi être de grandeur très différentes dans le calcul de Z et qui provoquerait une imprécision de Z.

Une façon de corriger ce problème est de s'apercevoir que :

$$\text{SOFTMAX}(< x_1, x_2, \ldots, x_n >) = \text{SOFTMAX}(< x_1 + a, x_2 + a, \ldots, x_n + a >)$$

En effet on a bien $\forall\, i \in [0, \ldots n]$:

$$\begin{aligned} \frac{\exp(x_i + a)}{\exp(x_1 + a) + \exp(x_2 + a) + \ldots + \exp(x_n + a)} &= \frac{\exp(a)\exp(x_i)}{\exp(a)\big[\exp(x_1) + \exp(x_2) + \ldots + \exp(x_n)\big]} \\ &= \frac{\exp(x_i)}{\exp(x_1) + \exp(x_2) + \ldots + \exp(x_n)} \qquad < \textbf{2 pts} > \end{aligned}$$

Ainsi en considérant $a = -\max\{x_1, x_2, \ldots, x_n\}$, on n'évitera le problème d'OVERFLOW sans, bien sur, modifier le résultat par rapport au calcul analytique. On convertit ainsi le problème d'OVERFLOW catastrophique en quelques UNDERFLOWS qui ne constituent pas de problème ici ; les petites valeurs d'énergies seront convertit, après l'exponentielle (pour seulement quelques x_i) en valeurs nulles qui n'affecteront pas le calcul de Z (toujours assurée d'être non nulle).
Le pseudo-code de la fonction SOFTMAX correctement implémenté sera donc :

ALGORITHM ; SOFTMAX

| $[x_1, x_2, \ldots, x_n]$ Tableau de flottants positifs

- A partir du tableau: $[x_1, x_2, \ldots, x_n]$
- On calcul: $x_{\max} = \max\{x_1, x_2, \ldots, x_n\}$
- On considère le nouveau tableau:
$[x_1^\circ = x_1 - x_{\max}, x_2^\circ = x_2 - x_{\max}, \ldots, x_n^\circ = x_n - x_{\max}]$
- On calcul: $\left[\dfrac{\exp(x_1^\circ)}{Z}, \dfrac{\exp(x_2^\circ)}{Z}, \ldots, \dfrac{\exp(x_n^\circ)}{Z}\right]$

Avec $Z = \exp(x_1^\circ) + \exp(x_2^\circ) + \ldots \exp(x_n^\circ)$

<4 pts>

2.(a)

$$I_0 = \int_0^1 \frac{1}{x+10}\, dx = \big[\ln(x+10)\big]_0^1 = \ln(11) - \ln(10) = \ln\Big(\frac{11}{10}\Big)$$

<2 pts>

$$\begin{aligned} I_n &= \int_0^1 \frac{x^n}{10+x} = \int_0^1 \frac{x^{n-1}(x+10) - 10\, x^{n-1}}{10+x} \\ &= \int_0^1 x^{n-1}\, dx \; - 10 \int_0^1 \frac{x^{n-1}}{10+x}\, dx = \Big[\frac{x^n}{n}\Big]_0^1 - 10\, I_{n-1} \\ &= \frac{1}{n} - 10\, I_{n-1} \end{aligned}$$

<4 pts>

2.(b)

En considérant chaque étape de cette suite itérative comme le résultat de la fonction $f_n(x) = \frac{1}{n} - 10\, x$, et en estimant la propagation d'erreur donnée par l'approximation de Taylor de la fonction f_n au premier ordre, on a :

$$\Delta f_n(x) = |-10|\, \Delta_x$$

Soit une amplification d'erreur par un facteur 10 à chaque itération de cette suite ! et donc l'erreur initiale sur I_0 (i.e., ΔI_0) est propagée à toute les itérations de cette suite et par cumul d'erreurs :

$$\Delta I_n = 10 \Delta I_{n-1} = 10^2 \Delta I_{n-2} = \ldots = 10^n \Delta I_0$$

jusqu'à une instabilité complète !
<6 pts>

Nota : Sur mon ordinateur, en utilisant des FLOAT on trouve $I_0 = 0.095310$, $I_1 = 0.046898$, $I_2 = 0.031018$, $I_3 = 0.023152$, $I_4 = 0.018480$, $I_5 = 0.015195$, $I_6 = 0.014712$, $I_7 = -0.004263$, $I_8 = 0.167631$, $I_9 = -1.565204$, $I_{10} = 15.752038$, $I_{11} = -157.429474$, ...
et en utilisant des DOUBLE, on a ce même problème vers $I_{18}(= 74.839443)$ alors que, normalement, cette integrale devrait converger vers 0 pour de fortes valeurs de n.

2.(c)

On a donc $\Delta I_8 \approx 10^8 \Delta I_0$. En considérant que la suite commence à diverger en FLOAT à partir de ce $n = 8$ (et que pour ce $n = 8$, la valeur donnée en DOUBLE est encore raisonablement juste), l'erreur ΔI_0 (sur I_0) est de l'ordre de 10^{-8}.
<4 pts>

3

Dans les trois cas, on aura un problème d'annulation de cse due à la soustraction de deux nombres approximés quasi égaux **<1 pt>**. Pour éviter ce problème, on peut écrire ces relations de telle façon que plus aucune soustraction de deux nombres incertains (i.e., imprécis) quasi égaux apparaissent.

On écrit la première expression en considérant le même dénominateur $(1 - x)(1 + x)$, on utilise une approximation de Taylor au voisinage de x_0 pour la seconde et on multiplie et divise par $\sqrt{1 - \cos(x)}$ pour la troisième. On obtient les trois expressions équivalentes suivantes

$$\frac{1}{1-x} - \frac{1}{1+x} = \frac{2x}{1-x^2} \qquad \text{plus de problème de précision au voisinage de } x_0 = 0$$

$$\sqrt{1+x} - \cos x \approx 1 + \frac{x}{2} - (1 - \frac{x^2}{4}) = \frac{x}{2} + \frac{x^2}{4} \qquad \text{plus de problème de précision au voisinage de } x_0 = 0$$

$$\text{ou} \quad \frac{1 + x - \cos^2(x)}{\sqrt{1+x} + \cos(x)} \quad \text{ou} \quad \frac{x + \sin^2(x)}{\sqrt{1+x} + \cos(x)}$$

$$\sqrt{1 + \cos x} = \frac{|\sin(x)|}{\sqrt{1 - \cos(x)}}$$

plus de problème de précision au voisinage de $x_0 = \pi$

<9 pts>
Nota : D'autres possibilités de formulation sont possibles

III. Méthode du Point Fixe et de Newton (33 pts)

1. Une application très célèbre de la méthode de Newton (utilisé de nos jours dans certains calculateurs) est le calcul de la racine carré d'un réel positif α. Soit donc la racine (positive) de la fonction

$$f(x) = x^2 - \alpha$$

   (a) Montrer que la méthode de Newton permet d'aboutir à la suite itérative suivante
   <4 pts>

$$x_{n+1} = \frac{1}{2}\Big(x_n + \frac{\alpha}{x_n}\Big)$$

   (b) Montrer, par les deux conditions de convergence vues en cours, que cette suite itérative de Newton converge sur l'interval $[1 \;\; \alpha]$ pour $\alpha > 1$.
   <8 pts>

(c) On pose $e_n = \frac{x_n - \sqrt{\alpha}}{\sqrt{\alpha}}$, pour $n = 0, 1, \ldots$

Montrer que l'on a nécessairement la relation $e_{n+1} = \frac{e_n^2}{2\,(1+e_n)}$

<5 pts>

(d) En supposant que l'on parte avec une erreur relative initiale $e_0 < 1$ (i.e., une erreur inférieure de $100\,\%$), montrer numériquement (i.e., sans démonstration) que l'on obtiendra une erreur relative de l'ordre de 10^{-15} en 5 itérations.

<3 pts>

(e) En déduire donc le nombres d'opérations élémentaires (i.e., additions, soustraction, multiplication et division de flottants) pour atteindre cette précision.

<3 pts>

(f) En utilisant la méthode de Newton, donner les cinq premières itérations $(x_0, x_1, \ldots, x_5)$ de l'estimation de la racine positive de $\alpha = 7$, en partant de $x_0 = \alpha/2$.

<3 pts>

2. Soit la fonction $f(x) = x^2 - \ln(1+x)$, qui admet deux racines, la première évidente $r_1 = 0$ et la deuxième r_2 qui semble comprise entre $[0.6 \;\; 1,0]$ que l'on désire trouver par une méthode de point fixe. Trouver une suite itérative $x_{n+1} = g(x_n)$ de type "point fixe" convergente (et démontrer bien que celle-ci converge) vers la deuxième racine de cette équation pour un choix adéquat de x_0 que vous préciserez. Calculer les trois premières itérations (x_0, x_1, x_2, x_3)

<7 pts>

Réponse

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

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

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

<4 pts>

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

La méthode de Newton est une méthode du point fixe, associée à la suite itérative $x_{n+1} = g(x_n)$ avec $g(x) = \frac{1}{2}(x + \frac{\alpha}{x})$. Calculons $|g'(x)|$

$$|g'(x)| = \left|\frac{1}{2} - \frac{\alpha}{2\,x^2}\right|$$

La racine positive de α (avec $\sqrt{\alpha} > 1$ précisée par l'énoncée) est nécessairement dans l'interval $J = [1 \;\; \alpha]$, et dans cet interval dans lequel on devra prendre x_0, $|g'(\xi)| < 1 \;\forall \xi$ ce qui assure la convergence de cette suite en tant que point fixe.

<4 pts>

Une autre façon de s'assurer de la convergence de cette suite est de revenir à la forme $x_{n+1} = x_n - \frac{(x_n^2 - \alpha)}{2x_n}$ et de s'apercevoir que l'interval $J = [1 \;\; \alpha]$ contient la racine (avec $\sqrt{\alpha} > 1$ précisée par l'énoncée), devra aussi contenir x_0 (le premier élément de cette suite itérative) et que $f'(x) = 2\,x$ ne s'annule jamais dans J. La racine et le premier élément de cette suite itérative ne sont donc pas séparé par un extrema de la fonction f ; la convergence est donc assurée en tant que méthode de Newton.

<4 pts>

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

$$e_{n+1} = \frac{x_{n+1} - \sqrt{\alpha}}{\sqrt{\alpha}} = \frac{\frac{1}{2}\Big(x_n + \frac{\alpha}{x_n}\Big) - \sqrt{\alpha}}{\sqrt{\alpha}} = \frac{(x_n - \sqrt{\alpha})^2}{2\sqrt{\alpha}\,x_n} = \frac{(x_n - \sqrt{\alpha})^2}{\alpha \cdot 2\left(\frac{x_n - \sqrt{\alpha} + \sqrt{\alpha}}{\sqrt{\alpha}}\right)} = \frac{e_n^2}{2\,(1 + e_n)}$$

<**5 pts**>

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

En utilisant $e_{n+1} = \frac{e_n^2}{2\,(1+e_n)}$, on obtient

$$\begin{aligned} e_0 &= 1 \\ e_1 &= 0.25 \\ e_2 &= 0.025 \\ e_3 &= 3 \times 10^{-4} \\ e_4 &= 4.64 \times 10^{-8} \\ e_5 &= 1.08 \times 10^{-15} \end{aligned}$$

<**3 pts**>

Nota : On aurait put aussi simplement dire que la méthode de Newton est convergente quadratiquement dans la cas d'une racine simple (i.e., notre cas) et calculer e_0 et e_1, voir que l'on gagnait 1 cse, dans cette itération et, avec la convergence quadratique de Newton, forcément 2 cse entre e_1 et e_2 et 4 cse entre e_2 et e_3 et 8 cse entre e_3 et e_4 et enfin 16 cse entre e_4 et e_5 ; soit une approximation de l'ordre de 10^{-15} pour la cinquième itération.

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

Il y a trois opération élémentaires (i.e., une division, une addtion et une multiplication) pour chaque itération de la suite itérative de newton ; donc au maximum $3 \times 5 = 15$ opérations élémentaires pour arriver à la racine d'un réel avec une erreur relative inférieure à 10^{-15} !

<**3 pts**>

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

$$\begin{aligned} x_0 &= 3.5 \\ x_1 &= 2.75 \\ x_2 &= 2.647727273 \\ x_3 &= 2.645752048 \\ x_4 &= 2.645751311 \\ x_5 &= 2.645751311 \end{aligned}$$

<**3 pts**>

$\boxed{\textbf{2}}$

L'énoncée nous dit que la deuxième racine est comprise dans l'interval $[0.6 \;\; 1.0]$. On peut mettre $f(x) = 0$, sous la forme $x = \sqrt{\ln(1+x)}$ ce qui nous conduirait à la suite itérative

$$x_{n+1} = \sqrt{\ln(1+x_n)}$$

reste à démontrer que sous cette forme, la convergence est assurée. sur $J = [0.6 \;\; 1.0]$. Pour cela calculons $|g'(x)|$

$$|g'(x)| = \Big|\frac{1}{x+1} \times \frac{1}{\sqrt{\ln(1+x)}}\Big|$$

La fonction $1/x+1$ est décroissante et positive sur J, la fonction $1+x$ est croissante et positive sur J, même chose pour la fonction $\sqrt{\ln(1+x)}$ et donc $\sqrt{\ln(1+x)}$ est aussi décroissante et positive sur J. En conclusion $|g'(x)|$ est décroissante et positive sur J et $|g'(\xi)| < g'(0.6) \approx 0.912 < 1$ sur J, la convergence est donc assurée si on prend $x_0 \in J$. On trouve pour $x_0 = 0.8$

$$\begin{aligned} x_1 &= 0.766672462 \\ x_2 &= 0.754385718 \\ x_3 &= 0.749745808 \end{aligned}$$

<**7 pts**>

Nota : Au bout de la 19ième itération, cela converge vers .746881742.

Nota : On aurait put prendre la suite itérative $x_{n+1} = \frac{\ln(1+x_n)}{x_n}$ qui converge ($x_0 = 0.8 \quad x_1 = 0.7347 \quad x_2 = 0.7497 \quad x_3 = 0.7462...$) mais la démonstration de convergence aurait été plus compliqué que celle donnée dan notre cas.

IV. Factorisation LU (17 pts)

1. Décomposer la matrice A en produit LU (sans permutation) par la méthode de la factorisation directe (celle assurant des "1" sur la diagonale de la matrice U).
   <**8 pts**>

$$A = \begin{pmatrix} 1 & 2 & 2 \\ 2 & 5 & 6 \\ 2 & 6 & 9 \end{pmatrix}$$

2. Décomposer la matrice A en produit LU (sans permutation) par la méthode d'élimination de Gauss.
   <**6 pts**>
3. Calculer le déterminant de A.
   <**3 pts**>

Réponse

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

Par factorisation directe (la méthode assurant des "1" sur la diagonale de U), on trouve

$$A = \begin{pmatrix} 1 & 2 & 2 \\ 2 & 5 & 6 \\ 2 & 6 & 9 \end{pmatrix} = \underbrace{\begin{pmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 2 & 2 & 1 \end{pmatrix}}_{L<\textbf{4 pts}>} \underbrace{\begin{pmatrix} 1 & 2 & 2 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{pmatrix}}_{U<\textbf{4 pts}>}.$$

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

Puisque la décomposition LU précédente a aussi des "1" sur la diagonale de la matrice L, le résultat précédent sera donc forcément le résultat donné par la méthode d'élimination de Gauss (à cause de l'unicité de la représentation).

<**6 pts**>

Nota : Les 6 points sont aussi attribués si on a refait les calcul avec la méthode du pivot de Gauss et que le résultat est correct !

$\boxed{\textbf{3.}}$

$$\begin{aligned} \det(\mathrm{A}) &= \det(\mathrm{L}) \times \det(\mathrm{U}) \\ &= 1 \times (1 \times 1) \\ &= 1 \end{aligned}$$

<**3 pts**>