DIRO
IFT 2421

EXAMEN INTRA

Max Mignotte

DIRO, Département d'Informatique et de Recherche Opérationnelle, local 2377.
Http : //www.iro.umontreal.ca/∼mignotte/ift2421/
E-mail : mignotte@iro.umontreal.ca

Date : *24/02/2004*

I	Mesure d'Incertitude et Amplification d'Erreur (30 pts).
II	Cancellation de CSE (10 pts).
III	Méthode du Point Fixe et de Newton (39 pts).
IV	Factorisation P^tLU avec Élimination Gaussienne (25 pts).
V	Interpolation de Newton (10 pts).

TOUS DOCUMENTS PERSONNELS, CALCULATRICES ET CALCULATEURS AUTORISÉS

I. Mesure d'Incertitude et Amplification d'Erreur (30 pts)

Soit la fonction suivante,

$$y = \frac{FL^4}{8EI}$$

(qui exprime la déviation y du sommet du mât d'un voilier, où F est le chargement de coté (lb/ft), L est la hauteur (ft), E est le module de l'élaticité et I est le moment d'inertie (ft^4)). Pour chacune de ces variables, on à la mesure approximée et l'incertitude suivante,

$$\begin{aligned}
F &= 50 \text{ lb/ft} \quad \text{à} \ \ 2 \text{ lb/ft près, (i.e.,} \Delta F = 2 \ \ \text{(lb/ft))}, \\
L &= 30 \text{ ft} \quad \text{à} \ \ 0.1 \text{ ft près}, \\
E &= 1.5 \times 10^8 \quad \text{à} \ \ 0.01 \times 10^8 \text{ lb/ft}^2 \text{ près}, \\
I &= 0.06 \text{ ft}^4 \quad \text{à} \ \ 0.0006 \text{ ft}^4 \text{ près}.
\end{aligned}$$

1. Donner le nombre de chiffres significatifs "exact" (cse) existant dans la valeur approximée de F, L, E et I.
2. Donner l'approximation de y (i.e., y^*) et l'incertitude de y (i.e., Δy) par la methode de propagation d'erreur (utilisant l'approximation de taylor de la fonction au premier ordre). Souligner les chiffres significatifs "exacts" (cse) de ce résultat. Arrondir cette approximation au nombre de cse adéquat. Donner finalement l'intervalle de confiance dans lequel se trouverait la vraie valeure de y obtenue par cette méthode (en utilisant les approximations non arrondies).
3. Utiliser la méthode de la fourchette pour obtenir un intervalle de confiance dans lequel se trouverait la vraie valeure de y et en déduire l'incertitude résultante (i.e., erreur absolue) sur y que l'on obtiendrait par cette méthode.
4. Des deux précédentes méthodes, laquelle est la plus précise ? Justifier votre réponse. La méthode de la fourchette est-elle applicable dans le cas de n'importe quelle fonction de plusieurs variables (incertaines) ? Justifier votre réponse.

Réponse

1.

Pour le nombre de cse de chacune de ces variables, on a,

$$\begin{aligned}
F &\approx \underline{5}0 \text{ lb/ft} \ \ \text{car} \ \ \Delta F = 2 < 0.5 \times 10^1, \\
L &\approx \underline{30} \text{ ft} \ \ \text{car} \ \ \Delta L = 0.1 < 0.5 \times 10^0, \\
E &\approx \underline{1.5} \times 10^8 \text{ lb/ft}^8 \ \ \text{car} \ \ \Delta E = 0.01 < 0.5 \times 10^{-1}, \\
I &\approx \underline{0.06} \text{ ft}^4 \ \ \text{car} \ \ \Delta I = 0.0006 < 0.5 \times 10^{-2}. \quad < \textbf{6pts} >
\end{aligned}$$

2.

On a pour valeur approchée,

$$y^* = \frac{50 \times 30^4}{8 \times 1.5 \times 10^8 \times 0.06} = 0.5625 \text{ ft}. \quad < \textbf{2pts} >$$

On obtient la différentielle suivante,

$$\Delta y(F, L, E, I) = \left| \frac{L^4}{8EI} \right| \Delta F + \left| \frac{FL^3}{2EI} \right| \Delta L + \left| -\frac{FL^4}{8E^2 I} \right| \Delta E + \left| -\frac{FL^4}{8EI^2} \right| \Delta I. \quad < \textbf{4pts} >$$

$$
\begin{aligned}
\text{soit} \qquad \Delta y &\approx 0.0225 + 0.0075 + 0.00375 + 0.005625, \\
&\approx 0.039375. \quad < \mathbf{3pts} >
\end{aligned}
$$

On obtient donc $y \approx 0.5625 \pm 0.039375$. Comme $0.039375 < 0.5 \times 10^{-1}$, le rang du dernier chiffre significatif est -1 et on peut écrire $y \approx \underline{0.5}625 \pm 0.039375$ <**1 pt**>. Si on arrondi finalement cette estimation au nombre de cse, on obtient,

$$y \approx 0.6 \ \pm 0.05 \text{ ft.} \quad < \mathbf{1pt} >$$

Si on utilise les approximations non arrondies, y se trouverait dans l'intervalle $[0.523125 \quad 0.601875]$. <**3 pts**>

3.

En utilisant les valeur maximales pour F (i.e., $F + \Delta F$) et L et les valeurs minimales pour E (i.e., $E - \Delta E$) et I, on obtiendrait la borne maximale de cette intervalle de confiance (y_{max}). On peut faire l'inverse pour trouver la borne minimale de cet intervalle de confiance (y_{min}).

$$
\begin{aligned}
y_{min} &= \frac{48 \times (29.9^4)}{8 \times (1.51 \times 10^8) \times 0.0606} = 0.52407, \\
y_{max} &= \frac{52 \times (30.1^4)}{8 \times (1.49 \times 10^8) \times 0.0594} = 0.60285. \quad < \mathbf{5pts} >
\end{aligned}
$$

Qui est raisonablement près de l'intervalle trouvé à la question précédente. La valeur approximée est donc le centre de cette intervalle de confiance et l'incertitude sur cette variable serait la demi largeur de cette intervalle. Ce qui nous permet d'écrire $y \approx 0.56346 \pm 0.03939$. <**2 pts**>

4.

La plus précise des deux méthode est la méthode de la fourchette car celle-ci n'utilise aucune approximation (approximation de Taylor du premier ordre pour la première méthode). Le problème, c'est que cette méthode (dite "de la fourchette") ne peut s'appliquer à toutes les types de fonctions à plusieurs variables ; seulement si la fonction est continue et monotone ; C'est le cas de y, pour laquelle une borne y_{min} et y_{max} peut être aussi facilement définie. <**3 pts**>

II. Cancellation de CSE (10 pts)

Expliquer pourquoi le calcul numérique de ces expressions,

$$
\begin{aligned}
&\ln(x+1) - \ln(x) \quad \text{quand} \ \ x >> 0, \\
&\sqrt{x^2+1} - x \quad \text{quand} \ \ x >> 0, \\
&\cos^2(x) - \sin^2(x) \quad \text{quand} \ \ x \approx \pi/4,
\end{aligned}
$$

peuvent conduire à des problèmes d'erreurs numériques ; Expliquer pourquoi et proposer une formule équivalente qui permettrait d'éviter ce problème pour ces trois relations.

Réponse

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

$$
\begin{aligned}
\ln(x+1) - \ln(x) &= \ln(1 + 1/x), \quad < \textbf{3pts} > \\
\sqrt{x^2+1} - x &= \frac{1}{\sqrt{x^2+1}+x}, \quad < \textbf{3pts} > \\
\cos^2(x) - \sin^2(x) &= \cos(2x), \quad < \textbf{3pts} >
\end{aligned}
$$

(en utilisant respectivement la propriété $\ln(a) - \ln(b) = \ln(a/b)$ dans la premeière relation, en multipliant et divisant la seconde expression par $\sqrt{x^2+1} + x$ et en considérant, pour la troisième expression le fait que $\cos^2(x) = (1 + \cos(2x))/2$ et $\sin^2(x) = (1 - \cos(2x))/2$.

(Nota : D'autres possibilités de formulation sont possibles)

III. Méthode du Point Fixe et de Newton (39 pts)

On se propose de trouver des valeurs approchées de la racine r de l'équation,

$$f(x) = x\exp(x) - 2x^2 - 1 = 0. \tag{1}$$

1. **Méthode du point fixe (15 pts)**.
   (a) Montrer qu'il existe une racine unique r pour cette équation (1) dans l'intervalle $J_1 = [1, 2]$.
   (b) En remarquant que l'équation proposée est équivalente à $g_1(x) = x$ avec $g_1(x) = x\exp(x) - 2x^2 + x - 1$, montrer que l'intervalle $J_1 = [1, 2]$ est un intervalle sur lequel la convergence vers une solution unique par la méthode du pont fixe n'est pas assurée.
   (c) Montrer que la solution de l'équation (1) est aussi solution de l'équation suivante,

   $$h(x) = \ln\Big(\frac{1+2x^2}{x}\Big) - x = 0, \tag{2}$$

   lorsque $x > 0$.
   (d) En remarquant que l'équation proposée est équivalente à $g_2(x) = x$ avec $g_2(x) = \ln(\frac{1+2x^2}{x})$ pour $x > 0$, montrer que l'intervalle $J_1 = [1, 2]$ est un intervalle sur lequel la convergence vers une solution unique par la méthode du pont fixe est cette fois-ci assurée.
   (e) Calculer les 4 premières estimées r_1, r_2, r_3, r_4 en partant de $r_0 = 1.5$.

2. **Méthode de Newton (24 pts)**.
   (a) Soit $g_3(x)$, la fonction intervenant dans la méthode itérative de Newton pour la résolution de l'équation (1) ($f(x) = x\exp(x) - 2x^2 - 1 = 0$). Donner $g_3(x)$ (i.e., donner la relation itérative $r_{n+1} = g_3(r_n)$).
   (b) Nous allons supposer que les conditions sur g_3 et sa dérivée sont remplies pour la convergence de cette méthode vers une solution unique. En déduire une valeur approchée de r après 3 itérations (i.e., donner r_1, r_2, r_3)(avec toujours $r_0 = 1.5$).
   (c) En supposant qu'à la troisième itération (i.e., en r_3), on approxime la solution avec une erreure inférieure à 3×10^{-4}, combien d'itérations de la méthode de la bissection auraient permis d'arriver à la même précision ? (en partant avec le même intervalle de départ).

(d) i. En supposant que le coût calculatoire d'une iteration de la méthode de la bisection ou de la méthode de Newton se compte approximativement en *évaluation de fonction* (et en comptant l'évaluation de la dérivée analytique de l'approche itérative de Newton comme une évaluation de fonction), comparer le nombre d'évaluation de fonctions nécessaire pour une itération de la méthode de Newton et pour une itération de la méthode de la bissection. En déduire celle des deux méthodes qui est effectivement la plus rapide (en terme de coût calculatoire) pour arriver à la précision donnée en 2.(c).

ii. Donner un exemple de formule numérique d'approximation de la dérivée utilisant 2 *évaluations de fonction* et répondre de nouveaux à la question précédente.

(e) Soit $g_4(x)$, la fonction intervenant dans la méthode itérative de Newton pour la résolution de l'équation (2). Donner $g_4(x)$.

(f) Calculer les 4 premières estimées r_1, r_2, r_3 en partant de $r_0 = 1.5$ et en déduire si l'approche iterative semble converger on non. En déduire ce que cela veut dire pour $g_4(x)$.

Nota : $\exp x = e^x$

Réponse

1.

a. Les variations de la fonction $f(x)$ sur $[1, 2]$ montrent que la fonction est continue et croissante sur cet intervalle (donc monotone) ($f'(x) = (x + 1)\exp(x) - 4x$ et $f'(x) > 0 \;\forall x \in [1, 2]$) <**1 pt**>. De plus, on a $f(1) = e - 3 \approx -0.28$ et $f(2) = 2e^2 - 5 \approx 5.77$ ce qui assure l'unicité d'une racine r dans cet interval <**1 pt**>.

b. La dérivée première de $g_1(x)$ s'écrit,

$$g_1'(x) = (x + 1)\exp(x) - 4x + 1.$$

$g_1'(1) = 2e - 3 \approx 2.43$ donc la fonction $g_1(x)$ n'est donc pas contractante sur J_1 (i.e., $|g_1'(x)| \not< 1 \;\forall\, x \in J_1$) et donc la convergence n'est pas assurée dans ce cas. <**2 pts**>.

c. x solution de $x\exp(x) - 2x^2 - 1 = 0$, est solution de l'équation $x\exp(x) = 2x^2 + 1$, aussi solution de $\ln(x) + x = \ln(1 + 2x^2)$ lorsque $x > 0$, et donc solution de $\ln(\frac{1+2x^2}{x}) - x = 0$ lorsque $x > 0$. <**2 pts**>

d. La dérivée première de $g_2(x)$ s'écrit,

$$g_2'(x) \;=\; \frac{4x^2 - (1 + 2x^2)}{x^2}\,\frac{x}{1 + 2x^2} = \frac{2x^2 - 1}{x(1 + 2x^2)} \quad < \textbf{2pts} >.$$

Sur $[1, 2]$, $2x^2 - 1 < 2x^2 + 1$, de même pour $x \in [1, 2]$, $2x^2 - 1 < x(2x^2 + 1)$, donc $g_2'(x) = \frac{2x^2-1}{x(1+2x^2)} < 1$ pour $x \in [1, 2]$. La fonction $g_2(x)$ est donc contractante sur J_1 (i.e., $|g_2'(x)| < 1 \;\forall\, x \in J_1$) et donc la convergence est cette fois-ci assurée <**4 pts**>.

e. En partant de $r_0 = 1.5$, on a, $r_n = g_2(r_{n-1})$ et,

$$\begin{aligned}
r_0 &= 1.5,\\
r_1 &= 1.299282984,\\
r_2 &= 1.214384783,\\
r_3 &= 1.179341467,\\
r_4 &= 1.165215481. \quad < \textbf{3pts} >
\end{aligned}$$

2.
a. On a,

$$g_3(x) = x - \frac{f(x)}{f'(x)} = x - \frac{x\exp(x) - 2x^2 - 1}{(x+1)\exp(x) - 4x}.$$

On a donc la formule itérative suivante,

$$r_{n+1} = r_n - \frac{r_n \exp(r_n) - 2r_n^2 - 1}{(r_n+1)\exp(r_n) - 4r_n}. \quad < \mathbf{4pts} >$$

On suppose que la convergence de l'algorithme est assurée comme nous le suppose l'énoncé et on obtient,

$$\begin{aligned} r_0 &= 1.5, \\ r_1 &= 1.265088144, \\ r_2 &= 1.170062742, \\ r_3 &= 1.156189532. \end{aligned}$$

$< \mathbf{3pts} >$

(NOTA : CELA CONVERGE TRÈS RAPIDEMENT VERS 1.155923589 DÈS LA QUATRIÈME ITÉRATION, DONC LA TROISIÈME ITÉRATION NOUS DONNE UNE ERREURE INFÉRIEURE À 2.66×10^{-4}.)

c. Si on avait utiliser la méthode de la bissection, et en partant de l'intervalle $[1, 2]$, il aurait fallu faire au moins n itérations avec,

$$\frac{b-a}{2^{n+1}} = \frac{1}{2^{n+1}} \leq 3 \times 10^{-4},$$

Soit au moins $n \geqslant 11$ itérations <**3 pts**>.

d. Pour la méthode de la bissection, on a 1 évaluation de fonction à chaque itération et 2 évaluations de fonction pour la méthode de Newton. 11 itérations pour la méthode de la bissection correspont à 11 évaluations de fonctions et 3 itérations de la méthode de Newton correspond à 6 évaluations de fonction. La méthode de Newton converge plus vite (en terme de coût calculatoire) pour cet exemple et pour arriver à la précision donnée en 2.(c) <**3 pts**>.

On pourrait prendre $f'(x) = f[(x+h) - f(x)]/h$ avec h petit pour approximer numériquement la dérivée avec deux évaluations de fonctions. La méthode de Newton resterait cependant encore plus rapide (9 évaluations de fonction contre 11 évaluations de fonction) <**2 pts**>.

e. On a, la formule itérative suivante,

$$r_{n+1} = r_n - \frac{\ln\left(\frac{1+2r_n^2}{r_n}\right) - r_n}{\frac{2r_n^2 - 1}{r_n(1+2r_n^2)} - 1}. \quad < \mathbf{4pts} >$$

f. qui permet d'arriver à convergence, mais très lentement,

$$\begin{aligned} r_0 &= 1.5, \\ r_1 &= 1.384435658, \\ r_2 &= 1.307063019, \\ r_3 &= 1.255075308, \\ &\vdots \quad < \mathbf{3pts} > \end{aligned}$$

Cela veut dire que $|g_4'(x)| < 1$ pour tout $x \in [r_0, r_1, \ldots, r_n]$ <**2 pts**>

IV. Factorisation P^tLU avec Élimination Gaussienne (25 pts)

1. Décomposer la matrice A en produit P^tLU ou P est la matrice de permutation par la méthode d'élimination de Gauss et pivotage partiel.

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 1 \\ 1 & -1 & 1 \end{pmatrix}.$$

2. Calculer le déterminant de A.
3. Donner la complexité algorithmique (i.e., coût calculatoire) lorsque,
   (a) on résoud un système $Ax = b_i$ de dimension n pour K b_i différents (i.e., K membres de droites différents) par la méthode d'élimination de Gauss sans passer par la factorisation P^tLU (i.e., en utilisant K matrices augmentées utilisant chacune son b_i).
   (b) En faisant d'abord une factorisation P^tLU de la matrice A de dimension n et en exploitant efficacement cette factorisation pour minimiser le nombre d'opérations nécessaire à la résolution de ces K systèmes.
4. Indiquez comment, à partir de cette décomposition P^tLU on pourrait calculer l'inverse de la matrice A (Indiquez les différentes étapes, en une dizaine de lignes, sans calculer cette matrice inverse).

Réponse

On intervertit la ligne 1 et la ligne 2 ce qui donne,

$$A = \begin{pmatrix} 2 & 1 & 1 \\ 1 & 2 & 3 \\ 1 & -1 & 1 \end{pmatrix}.$$

L'opération $\text{ligne}_2 = \text{ligne}_2 - (1/2)\text{ligne}_1$ et $\text{ligne}_3 = \text{ligne}_3 - (1/2)\text{ligne}_1$ donne,

$$A = \begin{pmatrix} 2 & 1 & 1 \\ 0 & (3/2) & (5/2) \\ 0 & (-3/2) & (1/2) \end{pmatrix}.$$

L'opération $\text{ligne}_3 = \text{ligne}_3 - (-1)\text{ligne}_2$ donne,

$$A = \begin{pmatrix} 2 & 1 & 1 \\ 0 & (3/2) & (5/2) \\ 0 & 0 & 3 \end{pmatrix}.$$

On a donc la décomposition suivante :

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 1 \\ 1 & -1 & 1 \end{pmatrix} = \underbrace{\begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}}_{P^t<\mathbf{3pts}>} \underbrace{\begin{pmatrix} 1 & 0 & 0 \\ (1/2) & 1 & 0 \\ (1/2) & -1 & 1 \end{pmatrix}}_{L<\mathbf{4pts}>} \underbrace{\begin{pmatrix} 2 & 1 & 1 \\ 0 & (3/2) & (5/2) \\ 0 & 0 & 3 \end{pmatrix}}_{U<\mathbf{4pts}>}.$$

2.

$$\begin{aligned}
\det(\mathrm{A}) &= \det(\mathrm{P}) \times \det(\mathrm{L}) \times \det(\mathrm{U}), \\
&= -1 \times 1 \times (2 \times (3/2) \times 3), \\
&= -9. \quad < \mathbf{4pts} >
\end{aligned}$$

3. Dans le premier cas, on obtient une complexité de $K\big(\frac{o(n^3)}{3} + o(\frac{n^2}{2})\big)$.
Dans le deuxième cas on obtient une complexité de $\big(\frac{o(n^3)}{3} + Ko(\frac{n^2}{2})\big)$.
$<$**4 pts**$>$

4.
On a $A = P^t LU = PLU$ (puisque la matrice P est symétrique), donc,

$$A^{-1} = (PLU)^{-1} = (LU)^{-1}P^{-1} = (LU)^{-1}P,$$

puisque P est orthogonal.

Une possibilité consiste à calculer d'abord $(LU)^{-1}$, c'est à dire faire les étapes de calcul suivants pour une matrice d'ordre n,

1. Poser $y = Ux$.
2. Pour les 3 différents b_i ; $\begin{pmatrix}1\\0\\0\end{pmatrix}, \begin{pmatrix}0\\1\\0\end{pmatrix}, \ldots, \begin{pmatrix}0\\0\\1\end{pmatrix}$,
   (a) Calculer par substitution avant y_i solution de $Ly_i = b_i$.
   (b) Calculer par substitution arrière x_i solution de $Ux_i = y_i$.
   (c) $(LU)^{-1}$ se construit colonne par colonne en considérant que x_i est sa *ime* colonne.
3. Finalement, $A^{-1} = (LU)^{-1}P$.

$<$**6 pts**$>$

V. Interpolation de Newton (10 pts) Soit les points suivants,

x_k	-1	0	5	7
y_k	1	-3	2	2

Appliquez la formule de Newton pour trouver un polynôme de degré trois qui passe par ces points. Évaluez ensuite ce polynôme pour $x = 2$, et $x = 3$.

Réponse

1.
Le tableau des différences divisées s'écrit,

x	y	Δy	$\Delta^2 y$	$\Delta^3 y$
-1	1			
		-4		
0	-3		5/6	
		1		(-41/336)
5	2		-1/7	
		0		
7	2			

<**4 pts**>

On obtient le polynôme suivant,

$$P_3(x) = 1 - 4(x+1) + (5/6)x(x+1) - (41/336)(x+1)x(x-5). < \textbf{4pts} >$$

Pour l'interpolation on trouve,

$$\begin{aligned} P(2) &\approx -3.80357, \\ P(3) &\approx -2.0714. \quad < \textbf{2pts} > \end{aligned}$$