בס"ד

שיטות נומריות, 88-376 ד"ר ג'רמי שיף

מבחן מועד ב', סמסטר א' תשס"ב

זמן המבחן: שעתיים וחצי.
מותר להשתמש בכל חומר עזר ובמחשב כיס מדעי (אבל לא מחשב נייד).
יש לענות על כל 6 השאלות בחלק א' ועל 2 מתוך 4 השאלות בחלק ב'.
ניקוד השאלות: בחלק א' 8 נקודות לכל שאלה, בחלק ב' 25 נקודות לכל שאלה
(סה"כ הניקוד: $100 = 2$ (בונוס אוטומטי) $6 \times 8 + 2 \times 25 + 2$).
יש לנמק היטב כל תשובה!

חלק א'

1. למשוואה $x^3 + x - 1 = 0$ יש שורש קרוב ל-$x = 0.7$. השתמש בשני סיבובים של שיטת ניוטון לשפר קירוב זה.

רקורסיית ניוטון היא

$$x_{n+1} = x_n - \frac{x_n^3 + x_n - 1}{3x_n^2 + 1} = \frac{2x_n^3 + 1}{3x_n^2 + 1}$$

ולכן אם $x_0 = 0.7$, $x_1 \approx 0.6826$, $x_2 \approx 0.6823279$. (היות ורואים 3 ספרות נכונות ב-x_1, סביר לתת כ-6 ספרות דיוק ב-x_2.)

---

2. בשאלה זו r הוא ווקטור שורה ממימד n, c הוא ווקטור עמוד ממימד n ו-A היא מטריצה $n \times n$, כך ש-$\det A \neq 0$. איזה מבין הפקודות הבאות הן פקודות חוקיות ב-Matlab או scilab: $A * (A\backslash c)$, $(A * A)\backslash c$, $A. * (A\backslash c)$, $(A. * A)\backslash c$, $(r/A) * (A\backslash c)$, $(r\backslash A). * (A/c)$?

הפקודות השלישית והששית (מימין לשמאול) לא חוקיות. בשלישית, היות ו-$A\backslash c$ הוא ווקטור עמוד ו-A היא מטריצה ריבועית, לא ניתן לבצע כפל רכיב-רכיב $A. * (A\backslash c)$. בששית, לא ניתן לחלק A על ידי r ($r\backslash A$) ולא A על ידי c (A/c). שאר הפקודות חוקיות, הן מחשבות: $A^{-1}c$, $A \cdot A^{-1}c$, $(A^2)^{-1}c$, $\tilde{A}^{-1}c$ כאשר $\tilde{A}$ הוא הכפל רכיב-רכיב של A עם עצמו, $(rA^{-1}) \cdot (A^{-1}c)$.

---

3. מצא $p(3)$, כאשר $p(x)$ הוא פולינום מדרגה 3 או פחות כך ש- $p(0) = 1$, $p(1) = 0$, $p(2) = 3$, $p(4) = 0$.

טבלת הפרשים מחולקים:

$$\begin{array}{cccccc} 0 & 1 & & & \\ 1 & 0 & -1 & & \\ 2 & 3 & 3 & 2 & \\ 4 & 0 & -\frac{3}{2} & -\frac{3}{2} & -\frac{7}{8} \end{array}$$

ולכן $p(x) = 1 - x + 2x(x-1) - \frac{7}{8}x(x-1)(x-2)$, ו- $p(3) = \frac{19}{4}$.

---

4. מצא את הקו הישר $p(x) = ax + b$ שהוא הקירוב הטוב ביותר, במובן של ריבועים מזעריים, לנקודות

x_i	-1.00	-0.50	0.00	0.50	1.00
y_i	0.67	0.96	1.00	1.04	1.33

$$\begin{pmatrix} \sum 1 & \sum x_i \\ \sum x_i & \sum x_i^2 \end{pmatrix} \begin{pmatrix} b \\ a \end{pmatrix} = \begin{pmatrix} \sum y_i \\ \sum y_i x_i \end{pmatrix}$$
$$\Rightarrow \begin{pmatrix} 5 & 0 \\ 0 & 2.50 \end{pmatrix} \begin{pmatrix} b \\ a \end{pmatrix} = \begin{pmatrix} 5.00 \\ 0.70 \end{pmatrix}$$

ולכן הקו הוא $p(x) = 0.28x + 1.00$.

---

5. מצא את הקו הישר $p(x) = ax + b$ שהוא הקירוב הטוב ביותר על הקטע $[-1, 1]$ לפונקציה $f(x) = 1 + \frac{1}{3}x^3$, במובן ש-

$$\int_{-1}^{1} (f(x) - p(x))^2 dx$$

הוא מינימלי.

$p(x) = a_0 + a_1 x$ כאשר

$$a_0 = \frac{1}{2} \int_{-1}^{1} \left(1 + \frac{1}{3}x^3\right) dx = 1$$

$$a_1 = \frac{3}{2} \int_{-1}^{1} x \left(1 + \frac{1}{3}x^3\right) dx = \frac{1}{5}$$

---

6. חשב את האינטגרל

$$\int_{-1}^{1} e^{-x^2/2} dx$$

על ידי תרבוע גאוס (Gaussian quadrature) עם $n=4$.

אם נכתוב $f(x)=e^{-x^2/2}$, הקירוב הוא $w_1f(x_1)+w_2f(x_2)+w_3f(x_3)+w_4f(x_4)$ כאשר $x_1=-x_3=0.8611363$, $x_2=-x_4=0.3399810$, $w_1=w_3=0.3478548$, $w_2=w_4=0.6521452$. שים לב שבגלל ש-f פונקציה זוגית $w_1f(x_1)=w_3f(x_3)$ ו-$w_2f(x_2)=w_4f(x_4)$. הקירוב ל-7 ספרות יוצא להיות 1.711225, התשובות המדוייקת ל-7 ספרות היא 1.711249.

חלק ב'

7. א. מצא את יחס הנסיגה שמקבלים כאשר מפעילים את שיטת ניוטון למשוואה $x^d=0$. (d מספר שלם נתון גדול מ-1) האם יש התכנסות לשורש $x=0$? האם ההתכנסות היא ליניארית או ריבועית?

ב. מצא את יחס הנסיגה שמקבלים כאשר מפעילים את שיטת ניוטון למשוואה $f(x)=0$, כאשר $f(x)$ היא הפונקציה

$$f(x)=\begin{cases} e^{-1/x^2} & x\neq 0\\ 0 & x=0\end{cases}$$

האם יש התכנסות לשורש $x=0$? יש להתייחס לכל בחירה אפשרית של x_0. כאשר יש התכנסות, האם היא ליניארית, ריבועית או מה?

א. רקורסיית ניוטון לפתרון המשוואה $f(x)=0$ היא

$$x_{n+1}=x_n-\frac{f(x_n)}{f'(x_n)}$$

כאשר $f(x)=x^d$, $f(x)/f'(x)=x/d$, ולכן הרקורסיה היא

$$x_{n+1}=x_n\left(1-\frac{1}{d}\right)$$

אם $d>1$ אזי $1-1/d$ הוא גורם חיובי קטן מ-1. ולכן יש התכנסות ליניארית לשורש $x=0$. ההתכנסות איננה ריבועית (כמו שהיא בדרך כלל עם שיטת ניוטון) בגלל שבשורש $f'(x)=0$.

ב. כאן $f'(x)=(2/x^3)e^{-1/x^2}=(2/x^3)f(x)$ ולכן יחס הנסיגה הוא

$$x_{n+1}=x_n-\frac{x_n^3}{2}$$

יש לשים לב ש-0 היא נקודת השבת היחידה של הרקורסיה. ולכן אם יש התכנסות, ההתכנסות היא ל-0. תחת איזה תנאים יש התכנסות? שים לב ש-

$$|x_{n+1}|=|x_n|\left|1-\frac{x_n^2}{2}\right|$$

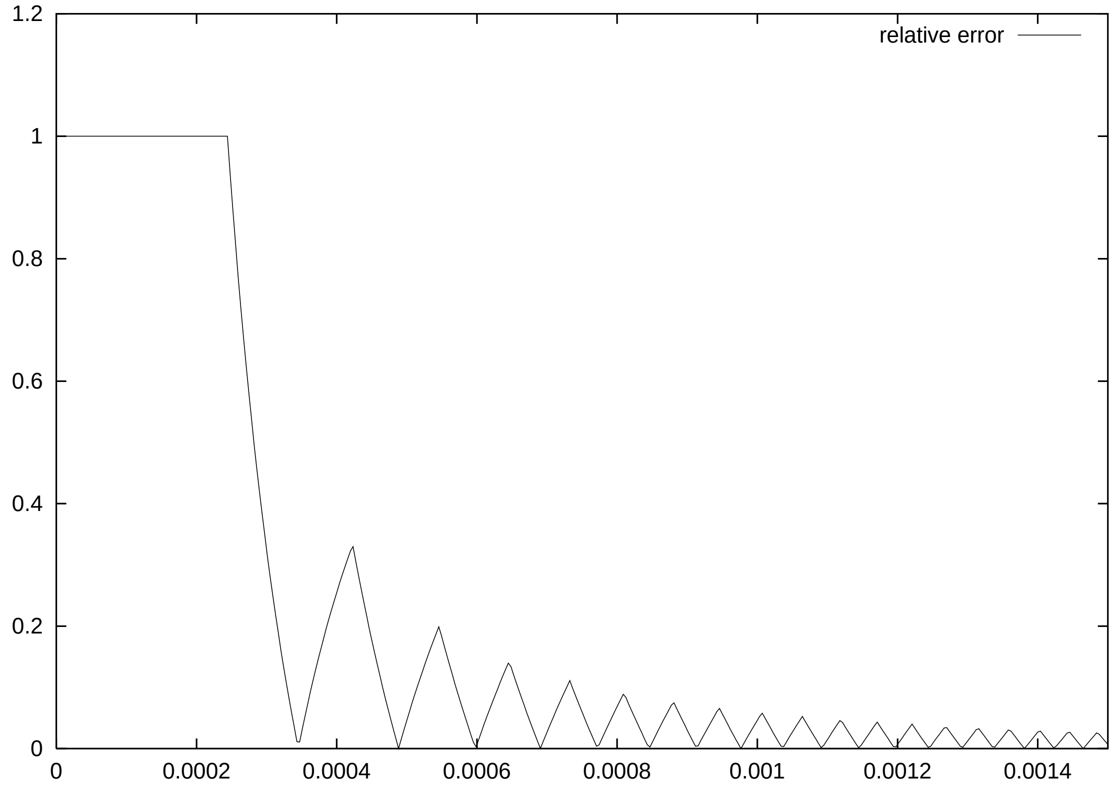
ולכן $|x_{n+1}| < |x_n|$ אם ורק אם $x_n^2 < 4$, כלומר $|x_n| < 2$. מזה נובע שיש התכנסות ל-0 בתנאי ש-$|x_0| < 2$. ההתכנסות היא יותר איטית מליניארית: בהתכנסות ליניארית השינוי בין x_n ל-x_{n+1} (כאשר יש התכנסות ל-0) הוא אחוז קבוע של x_n. כאן השינוי היא $x_n^3/2$, הרבה פחות.

---

8. א. מה הם המספרים הכי גדולים שהם קטנים מ-1 בתקן float ובתקן double?

ב. מצא את ה-ϵ החיובי הכי קטן, כך שכאשר מחשבים את $1 - \epsilon^2/2$ בתקן float מקבלים מספר קטן מ-1. מה היא התשובה אם עובדים בתקן double?

ג. למטה מופיע גרף של הטעות היחסית, כפונקציה של ϵ, בתוצאות שמתקבלות על ידי שימוש בתקן float לחשב $1 - (1 - \epsilon^2/2)$. הסבר את הצורה של הגרף.



א. בתקן float המספר הכי גדול שהוא קטן מ-1 הוא $1.11\ldots11 \times 2^{-1}$ כאשר יש 23 1'ים אחרי הנקודה. כלומר $1 - 2^{-24}$. בתקן double התשובה היא $1 - 2^{-53}$.

ב. כדי לקבל מספר קטן מ-1 כאשר מחשבים את $1 - \epsilon^2/2$ בתקן float, יש לדאוג שהתוצאה המדוייקת הוא יותר מחצי הדרך מ-1 ל-$1 - 2^{24}$. כלומר, צריך $\epsilon^2/2 > 2^{-25}$, כלומר

$$|\epsilon| > 2^{-12} \approx 0.000244$$

. בתקן double יש לדרוש ש-$\epsilon^2/2 > 2^{-54}$, כלומר

$$|\epsilon| > 2^{-27.5} \approx 5.27 \times 10^{-7}$$

ג. הנה כמה נקודות רלוונטיות: כאשר $\epsilon < 0.000244$ המחשב יחזיר 0 ויש 100% טעות, כלומר טעות יחסית 1. כאשר

$$\frac{1}{2}\epsilon^2 = 2^{-24}, 2 \times 2^{-24}, 3 \times 2^{-24}, \ldots$$

לא תהיה טעות בכלל, כלומר כאשר $\epsilon \approx 0.0003453\sqrt{n}$ לכל שלם n. רואים כ-19 נקודות כאלה על הגרף. לדוגמה התשיעית היא בערך 0.00104, מתאים ל-$n = 3$. בין כל זוג נקודות כאלה הטעות עולה ויורדת, כל ש-$\epsilon^2/2$ הוא יותר רחוק מאחד המספרים בתקן float. הטעות המקסימלית יורדת כל ש-ϵ גדל: הטעות המוחלטת היא לא גדולה מ-2^{-25}, ולכן כל שהתוצאה האמיתית יותר גדולה, הטעות היחסית תהיה יותר קטנה. כל הדברים האלה רואים על הגרף.

---

9. א. מצא את ה-natural cubic spline אשר עובר דרך שלושת הנקודות $(-h,\alpha), (0,\beta), (h,\gamma)$, כאשר h,α,β,γ הם מספרים ממשיים כל שהם עם $h > 0$. יש לתת נוסחאות מפורשות ל-cubic spline על הקטעים $(-h,0)$ ו-$(0,h)$.

ב. הוכח שבמקרה ששלושת הנקודות נמצאות על קו ישר אחד, אזי ה-cubic spline שחשבת בסעיף א. שווה ל-linear spline דרך שלושת הנקודות.

ג. הוכח שאם $\alpha \neq \beta = \gamma$ אזי אין ל-cubic spline מסעיף א. נקודות קריטיות בקטע $(-h,0)$. תן הסבר אינטואיטיבי לתופעה זו.

א. מהנוסחאות הכלליות ל-cubic splines

$$4hC_1 = 6\left(\frac{\gamma-\beta}{h} - \frac{\beta-\alpha}{h}\right) \;\Rightarrow\; C_1 = \frac{3(\gamma-2\beta+\alpha)}{2h^2}$$

ולכן

$$f(x) = \begin{cases} \frac{x+h}{h}\beta - \frac{x}{h}\alpha + \frac{x(x+h)(x+2h)}{4h^3}(\gamma-2\beta+\alpha) & -h \le x \le 0 \\ \frac{x}{h}\gamma - \frac{x-h}{h}\beta - \frac{x(x-h)(x-2h)}{4h^3}(\gamma-2\beta+\alpha) & 0 \le x \le h \end{cases}$$

ב. שלושת הנקודות הם על קו ישר אם ורק אם $\gamma-\beta = \beta-\alpha$, ולכן הרכיבים הלא-ליניאריים בנוסחאות למעלה מתאפסים, ומקבלים את ה-spline הליניארי.

ג. אם $\beta = \gamma$, בקטע $-h < x < 0$ יש לנו

$$f(x) = \frac{x+h}{h}\beta - \frac{x}{h}\alpha + \frac{x(x+h)(x+2h)}{4h^3}(\alpha-\beta)$$

ולכן

$$\begin{aligned} f'(x) &= \frac{\beta-\alpha}{h} + \frac{3x^2+6xh+2h^2}{4h^3}(\alpha-\beta) \\ &= (\beta-\alpha)\frac{2h^2-6xh-3x^2}{4h^3} \\ &= (\beta-\alpha)\frac{5h^2-3(x+h)^2}{4h^3} \end{aligned}$$

היות ובקטע $-h < x < 0$, $(x+h)^2 < h^2$, ברור ש-$f'(x) \neq 0$. כלומר, $f(x)$ או עולה מונוטונית על הקטע $(-h,0)$, כאשר $\alpha > \beta$, או יורד מונוטונית, כאשר $\alpha < \beta$. וזה תכונה טבעית לדרוש מפונקציה המחברת את הנקודות $(-h,\alpha),(0,\beta),(h,\beta)$.

---

10. א. הוכח שכאשר מחשבים קירוב לאינטגרל

$$\int_0^1 x\ dx$$

על ידי כלל הטרפז, עם n קטעים, מקבלים את התשובה המדויקת $\frac{1}{2}$. נוסחת עזר:

$$1+2+\ldots+(N-2)+(N-1)=\frac{N(N-1)}{2}$$

ב. הוכח שכאשר מחשבים קירוב לאינטגרל

$$\int_0^1 x^3\ dx$$

על ידי כלל הטרפז, עם n קטעים, מקבלים את התשובה $\frac{1}{4}\left(1+\frac{1}{n^2}\right)$. נוסחת עזר:

$$1^3+2^3+\ldots+(N-2)^3+(N-1)^3=\frac{N^2(N-1)^2}{4}$$

מה מקבלים על ידי כלל סמפסון?

ג. כאשר מחשבים קירוב לאינטגרל

$$\int_0^1 x^5\ dx$$

על ידי כלל הטרפז, עם n קטעים, מקבלים את התשובה $\frac{1}{6}\left(1+\frac{5}{2n^2}-\frac{1}{2n^4}\right)$. מה מקבלים על ידי כלל סמפסון? האם ניתן לשפר על ידי שימוש בשיטת רומברג?

א. כלל טרפז עם n קטעים:

$$\int_0^1 f(x)dx \approx \frac{1}{2n}\left(f(0)+2\sum_{i=1}^{n-1} f\left(\frac{i}{n}\right)+f(1)\right)$$

כאשר $f(x)=x$ מקבלים

$$\int_0^1 xdx \approx \frac{1}{2n}\left(2\sum_{i=1}^{n-1}\frac{i}{n}+1\right)=\frac{1}{2n}\left(\frac{2}{n}\cdot\frac{n(n-1)}{2}+1\right)=\frac{1}{2}$$

ב. כמו כן

$$\int_0^1 x^3 dx \approx \frac{1}{2n}\left(2\sum_{i=1}^{n-1}\frac{i^3}{n^3}+1\right) = \frac{1}{2n}\left(\frac{2}{n^3}\cdot\frac{n^2(n-1)^2}{4}+1\right) = \frac{1}{4}\left(1+\frac{1}{n^2}\right)$$

הדרך הכי קל לקבל את התוצאה של כלל סמפסון הוא להשתמש בזה שמקבלים את התוצאה של סמפסון על ידי

$$\frac{4I_{2n}-I_n}{3}$$

כאשר I_n זה התוצאה של טרפז עם n צעדים, ו-I_{2n} היא התוצאה עם $2n$ צעדים. כאן מקבלים

$$\frac{1}{3}\left(4\cdot\frac{1}{4}\left(1+\frac{1}{4n^2}\right)-\frac{1}{4}\left(1+\frac{1}{n^2}\right)\right)=\frac{1}{4}$$

כלומר סמפסון הוא מדויק!

(אם רוצים להוכיח את זה ישיר: כלל סמפסון ל-$2n$ צעדים הוא

$$\int_0^1 f(x)dx \approx \frac{1}{6n}\left(f(0)+4\sum_{\substack{i=1\\ i\text{ odd}}}^{2n-1} f\left(\frac{i}{2n}\right)+2\sum_{\substack{i=2\\ i\text{ even}}}^{2n-2} f\left(\frac{i}{2n}\right)+f(1)\right)$$

במקרה שלנו

$$\int_0^1 x^3dx \approx \frac{1}{6n}\left(\frac{1}{2n^3}(1^3+3^3+\ldots+(2n-1)^3)+\frac{1}{4n^3}(2^3+4^3+\ldots+(2n-2)^3)+1\right)$$

עכשיו

$$\begin{aligned}2^3+4^3+\ldots+(2n-2)^3 &= 2^3(1^3+2^3+\ldots+(n-1)^3)\\ &= 8\cdot\frac{n^2(n-1)^2}{4}\\ &= 2n^2(n-1)^2\end{aligned}$$

ו-

$$\begin{aligned}1^3+3^3+\ldots+(2n-1)^3 &= (1^3+2^3+\ldots+(2n-1)^3)-(2^3+4^3+\ldots+(2n-2)^3)\\ &= \frac{(2n)^2(2n-1)^2}{4}-2n^2(n-1)^2\\ &= n^2((2n-1)^2-2(n-1)^2)\\ &= n^2(2n^2-1)\end{aligned}$$

ולכן סמפסון נותן

$$\int_0^1 x^3 dx \approx \frac{1}{6n}\left(\frac{1}{2n^3}\cdot n^2(2n^2-1)+\frac{1}{4n^3}\cdot 2n^2(n-1)^2+1\right)$$
$$=\frac{1}{12n^2}\left(2n^2-1+(n-1)^2+2n\right)=\frac{1}{4}$$

גישה זו איננה מומלצת!

ג. כלל טרפז עם n צעדים נותן

$$I_n=\frac{1}{6}\left(1+\frac{5}{2n^2}-\frac{1}{2n^4}\right)$$

ולכן סמפסון נותן

$$\frac{4I_{2n}-I_n}{3}=\frac{1}{3}\left(\frac{2}{3}\left(1+\frac{5}{8n^2}-\frac{1}{32n^4}\right)-\frac{1}{6}\left(1+\frac{5}{2n^2}-\frac{1}{2n^4}\right)\right)=\frac{1}{6}\left(1+\frac{1}{8n^4}\right)$$

בשלב הבא בשיטת רומברג, אם J_{2n} מסמן את התוצאה שמקבלים מסמפסון עם $2n$ צעדים, מחשבים

$$\frac{16J_{4n}-J_{2n}}{15}$$

כאן זה נותן את התוצאה המדוייקת $\frac{1}{6}$.