מודלים לא ליניאריים בחקר ביצועים --- תרגיל בית 2

26 במאי 2024

שאלה 1

ממשו ב־python את ארבע הפונקציות הבאות. הפונקציות מופעלות באופן הבא:

```
x, fv = generic_bisect(f, df, l, u, eps, k)
x, fv = generic_newton(f, df, ddf, x0, k)
x, fv = generic_hybrid(f, df, ddf, l, u, x0, eps, k)
x, fv = generic_gs(f, l, u, eps, k)
```

בהן ממומשות (בהתאמה) ארבע השיטות חציה, ניוטון, ההיברידית וחתך הזהב. המשתנים אותם מקבלות הפונקציות הם:

- `f` --- הפונקציה אותה ממזערים.
- `df` --- הנגזרת של `f`.
- `ddf` --- הנגזרת השנייה של `f`.
- `x0` --- נקודת התחלה עבור שיטת ניוטון.
- `l,u` --- קצוות הקטע ההתחלתי עבור השיטות הדורשות קטע כזה.
- `eps` --- רמות דיוק.
- `k` --- כמות האיטרציות המרבית עבור כל אחת מהשיטות.

כל פונקציה מחזירה את הערכים הבאים:

- `x` --- הקירוב של נקודת המינימום של `f` באיטרציה האחרונה. עבור שיטת החציה -- החזירו את אמצע הקטע האחרון. עבור שיטת חתך הזהב -- החזירו את הקצה התחתון של הקטע.
- `fv` --- וקטור שמכיל את ערכי פונקציית המטרה בכל איטרציה שביצעה השיטה. דאגו שהנקודה הראשונה בווקטור תהיה `f(u)`. עבור שיטת החציה -- החזירו את ערך אמצע הקטע. עבור שיטת חתך הזהב -- החזירו את ערך הקצה העליון של הקטע.

הפעילו את כל ארבע הפונקציות על מנת למצוא את נקודת המינימום היחידה של הפונקציה

$$f(x) = (x-1)^3 + \frac{1}{1-x^2}$$

בקטע $[-0.999, 0.9]$ שגם ישמש כקטע ההתחלתי (אין צורך להוכיח שבקטע זה יש נקודת מינימום יחידה). עבור שיטת ניוטון נקודת ההתחלה תהיה הקצה העליון של הקטע. הגבילו את כמות האיטרציות בכל ניסוי ל־50 והשתמשו ברמת הדיוק $\varepsilon = 10^{-6}$. צרפו תרשים אחד שבו ניתן לראות את הערך של

$$\mathtt{fv}\left(\mathtt{i}\right) + 3.08891254695156$$

כפונקציה של האיטרציות בציר ה־x וטווח לוגריתמי בציר ה־y עבור כל אחת מהשיטות. היעזרו בפונקציה המובנית `semilogy` לשם כך (שימו לב ש־-3.08891254695156 הוא הערך של f בנקודת ההתכנסות). הסבירו את תוצאותיכם בהקשר קצב ההתכנסות: איזו שיטה הכי מהירה במקרה זה? האם יש הבדל מהותי בין שיטת חתך הזהב לשיטת החצייה? האם יש הבדל מהותי בין חתך הזה ושיטת החצייה לבין שיטת ניוטון והשיטה ההיברידית? האם שיטת ניוטון והשיטה ההיברידית מתלכדות? אם הן לא מתלכדות -- מה הסיבה שגרמה להן להיפרד לראשונה?

שאלה 2

נתונה מטריצה $\mathbf{Q} \in \mathbb{R}^{n\times n}$ סימטרית ומוגדרת אי־שלילית ונתון וקטור $\mathbf{0}_n \neq \mathbf{a} \in \mathbb{R}^n$. תהי $f\colon \mathbb{R} \to \mathbb{R}$ המוגדרת ע"י

$$.f\left(\mu\right) = \mathbf{a}^T\left(\mathbf{Q} + \mu\mathbf{I}_n\right)^{-1}\mathbf{a}$$

א. הוכיחו כי f מוגדרת לכל $\mu > 0$.

הכוונה: רישמו את הפירוק הספקטרלי של $\mathbf{Q} + \mu\mathbf{I}_n$ בהינתן הפירוק הספקטרלי של $\mathbf{Q}$.

ב. הוכיחו כי $f\left(\mu\right) > 0$ לכל $\mu > 0$.

ג. הוכיחו כי f יורדת בתחום $\mu > 0$.

ד. יהי $\alpha \in \mathbb{R}$. אנו מעוניינים להפעיל את שיטת החצייה בשביל למצוא $\mu > 0$ המקיים את המשוואה $f\left(\mu\right) = \alpha$. מיצאו תנאי על α עבורו בהכרח קיים פתרון למשוואה זו בתחום $\mu > 0$. הוכיחו את תשובתכם.

בנוסף, כאשר קיים פתרון למשוואה, הגדירו לולאה המוצאת את קצוות הקטע ההתחלתי ממנו ניתן לאתחל את השיטה.

הכוונה: חלקו את תשובתם לפי סימן הערך העצמי הקטן ביותר של $\mathbf{Q}$.

ה. העריכו את כמות האיטרציות הדרושות עד התכנסות השיטה בהסתמך על הקטע שמצאתם בסעיף הקודם.

שאלה 3

נתון פולינום $p\colon \mathbb{R} \to \mathbb{R}$ המוגדר על־ידי

$$.p\left(x\right) = -3.55x^3 + 1.1x^2 + 0.765x - 0.74$$

בנוסף, נתונה הנקודה $x^0 = 0.5554$.

א. הוכיחו של־p יש לפחות שורש אחד ומיצאו קטע סגור $[l, u]$ המכיל שורש של p.

ב. חשבו את שלוש האיטרציות הראשונות של שיטת ניוטון כאשר נקודת ההתחלה היא x^0 הנתונה. חשבו כל נקודה עד כדי ארבע ספרות לאחר הנקודה העשרונית בדיוק. הוכיחו שהסדרה $\left\{x^n\right\}_{n\geq 0}$ הנוצרת ע"י שיטת ניוטון לא מתכנסת לשורש של p.

ג. מיצאו את הנקודות הסטציונריות של p. כעת, נניח והפעלנו את שיטת ניוטון למציאת שורש של p עבור נקודת התחלה כלשהי. הסבירו מדוע לא ניתן לקבוע מראש שכל איטרציות ניוטון אכן מוגדרות היטב.

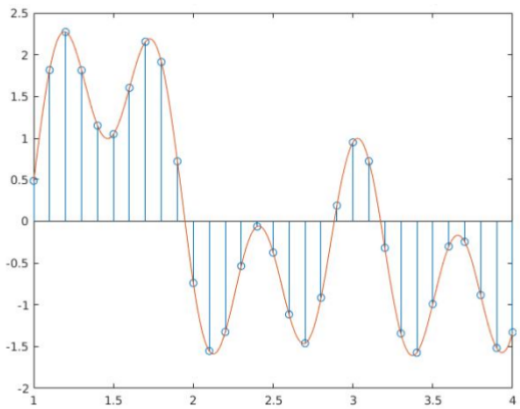
ד. הסבירו מדוע השיטה ההיברידית למציאת שורש של p היא בהכרח בעלת איטרציות מוגדרות היטב ומדוע היא בהכרח מתכנסת לשורש של p. כמו כן הסבירו מדוע היא מונעת את המקרה המתואר בסעיף ג'.

השתמשו בפונקציה אותה כתבתם בשאלה 1 והפעילו אותה עבור הבעיה בשאלה זו, עם הקטע ההתחלתי שמצאתם בסעיף א', ועבור $\varepsilon = 10^{-5}$. הגבילו את כמות האיטרציות ל־50. מהו השורש שמתקבל?

הערה: בשונה משאלה 1 שם חיפשנו נקודת מינימום של פונקציה, כאן אנו מעוניינים למצוא שורש של פונקציה. לכן הזינו $\mathtt{f = p}$, $\mathtt{df = dp}$ ו־$\mathtt{ddf = dp}$.

שאלה 4

שאלה זו עוסקת בסינון רעשים באות (צליל, טמפרטורה ועוד). באופן כללי, במחשבים נהוג לעבוד עם דגימה של האות. הדגימה מיוצגת באמצעות וקטור $\mathbf{s} \in \mathbb{R}^n$ כאשר n היא כמות התצפיות. למשל, בתרשים הבא האות הרציף נדגם 31 פעמים (העיגולים הכחולים), ו־$\mathbf{s} \in \mathbb{R}^{31}$ הוא וקטור ערכי הדגימות.



נשים לב שאם "נחבר את הנקודות" באמצעות קווים נקבל הערכה של האות באמצעות הווקטור $\mathbf{s}$, וככל שנאסוף יותר דגימות נקבל הערכה טובה יותר של האות המקורי. לכן נתייחס ל־$\mathbf{s}$ כאל האות המוערך.
מכיוון שמכשירי מדידה לרוב אינם מדויקים, אותות (מוערכים) במחשב בדרך כלל מכילים רעש. על מנת לסנן רעשים, אנו צריכים מידע כלשהו על האות המקורי. בשאלה זו נניח שהאות אותו מודדים הוא חלק (smooth), במובן שבו דגימה מסוימת קרובה לקודמתה. לכן, על מנת לסנן רעשים מאות מוערך $\mathbf{s} \in \mathbb{R}^n$ כלשהו, נמצא אות $\mathbf{x} \in \mathbb{R}^n$ חלק הדומה ל־$\mathbf{s}$. מתמטית, נפתור את בעיית האופטימיזציה הבאה שנסמנה (P):

$$\min_{\mathbf{x}\in\mathbb{R}^n} \left\{ f\left(\mathbf{x}\right) \equiv \alpha \left\|\mathbf{x}-\mathbf{s}\right\|_2^2 + \sum_{i=2}^{n} \left|\mathbf{x}_i - \mathbf{x}_{i-1}\right| \right\}.$$

מתקיים ש־$f \geq 0$. נשים לב שהמחובר $\|\mathbf{x}-\mathbf{s}\|_2^2$ דואג שהאות $\mathbf{x}$ אותו אנו מנסים למצוא יהיה דומה לאות הנדגם $\mathbf{s}$: הערך המזערי אותו הביטוי $\|\mathbf{x}-\mathbf{s}\|_2^2$ יכול לקבל הוא 0, ומכיוון שזו בעיית מינימיזציה אז נקבל שככל שווקטור הפתרון $\mathbf{x}$ קרוב יותר ל־$\mathbf{s}$, אז f מקבלת ערך נמוך יותר. בנוסף, הערך המזערי של המחובר $\sum_{i=2}^{n} |\mathbf{x}_i - \mathbf{x}_{i-1}|$ הוא גם כן 0, והוא מתקבל כאשר כל הקואורדינטות של וקטור הפתרון $\mathbf{x}$ זהות. מכאן, פתרון של הבעיה הוא וקטור $\mathbf{x}$ הדומה ל־$\mathbf{s}$ וגם שהקואורדינטות שלו לא רחוקות מדיי זו מזו (כלומר, הערכה חלקה ל־$\mathbf{s}$).
הפרמטר $\alpha > 0$ נקבע ע"י המשתמש והוא שולט במשקל אותו אנו נותנים לדמיון ל־$\mathbf{s}$ לעומת החלקות. בסעיפים הבאים ננסה (ולא נצליח) לפתור את הבעיה הנ"ל תוך שימוש בשיטת חתך הזהב.

א. נגדיר את הפונקציה $\phi\colon \mathbb{R}\to\mathbb{R}$ המקיימת

$$\phi_k\left(t\right)=f\left(\mathbf{x}_1,\ldots,\mathbf{x}_{k-1},t,\mathbf{x}_{k+1},\ldots,\mathbf{x}_n\right).$$

כאן אנו מתייחסים ל־$\mathbf{x}_1,\ldots,\mathbf{x}_{k-1},\mathbf{x}_{k+1},\ldots,\mathbf{x}_n$ כאל קבועים. הוכיחו כי לכל k ניתן לכתוב את $\phi_k\left(t\right)$ באופן הבא:

$$\phi_k\left(t\right)=C\left(\mu\left(t-a\right)^2+\left|t-b\right|+\left|t-c\right|\right)+D,$$

עבור $a,b,c,C,D,\mu\in\mathbb{R}$ קבועים.

ב. נתונה הפונקציה $\varphi\colon \mathbb{R}\to\mathbb{R}$ המקיימת

$$\varphi\left(t\right)=\mu\left(t-a\right)^2+\left|t-b\right|+\left|t-c\right|.$$

ונגדיר $l\equiv\min\left\{a,b,c\right\}-1$ ו־$u\equiv\max\left\{a,b,c\right\}+1$. הוכיחו כי אם ל־$\varphi$ קיים מינימום גלובלי $t^*\in\mathbb{R}$, אז $t^*\in\left[l,u\right]$.

ג. ידוע של־φ יש נקודת מינימום יחידה, והיא גלובלית (נדע להוכיח זאת בהמשך הקורס). כיתבו פונקציית python המופעלת באופן הבא:

```
t = gs_denoise_step(mu, a, b, c)
```

כאשר $a,b,c\in\mathbb{R}$ ו־$\mu>0$ ומחזירה את $t=\underset{t\in\mathbb{R}}{\operatorname{argmin}}\varphi\left(t\right)$ באמצעות שיטת חתך הזהב. השיטה תעצור כאשר השגיאה היא לכל היותר $\varepsilon=10^{-10}$. הסבירו מדוע ניתן להפעיל את השיטה על φ ומדוע מובטח שהשיטה אכן תחזיר את נקודת המינימום. במימוש פונקציה זו ניתן גם לקרוא לפונקציה `generic_gs` אותה מימשתם בשאלה 1.

ד. כיתבו פונקציית python המופעלת באופן הבא:

```
x = gs_denoise(s ,alpha ,N)
```

המקבלת את $\mathbf{s}\in\mathbb{R}^n$, $\alpha>0$ ו־$N\in\mathbb{N}$ הפותרת את הבעיה (P) באמצעות האלגוריתם הבא:

$$\begin{array}{l}
\text{set } \mathbf{x}\coloneqq\mathbf{s}\\
\text{for } i=1,2,\ldots,N \text{ do:}\\
\quad\text{for } k=1,2,\ldots,n \text{ do:}\\
\qquad \mathbf{x}_k\coloneqq\underset{t\in\mathbb{R}}{\operatorname{argmin}}\phi_k\left(t\right)\\
\quad\text{end for}\\
\text{end for}
\end{array}$$

השתמשו בפונקציה `gs_denoise_step` מהסעיף הקודם.

הערה: שימו לב שבכל איטרציה של הלולאה הפנימית אנו מורידים את ערכה של f ע"י שינוי של קואורדינטה אחת בלבד של $\mathbf{x}$. הלולאה הפנימית מבצעת "סבב שיפורים" עבור כל משתנה, והלולאה החיצונית מבצעת N סבבים כאלה (שיטות כאלה נקראות Alternating Minimization).

ה. צרפו את הפלט של הקוד script_ex4.ipynb המצורף לתרגיל. שימו לב שהאות המקורי הוא אות בדיד (דיסקרטי) עם ערכים של 1 ו־0 והוא צבוע בשחור.

האם אתם רואים הבדל בין תוצאות האלגוריתם עבור 10, 20 ו־30 איטרציות חיצוניות? איזה אות מקרב לדעתכם טוב יותר את האות האמתי, s או x? האם לדעתכם הפתרון x שמצא האלגוריתם הוא קירוב טוב של האות המקורי? בהמשך הקורס נראה שיטה מוצלחת יותר לקירוב אות בדיד.