חישוביות וקוגניציה – תרגיל 3

להגשה עד: 25/11/2019

<u>שימו לב</u>: שאלות 1 ו 2 הן שאלות אנליטיות, ושאלה 3 היא שאלת תכנות

שאלה 1

בשאלה זו נשלים את הוכחת השקילות בין ההצגה של רגרסיה לינארית פשוטה לבין פרספטרון לינארי.
נניח שנתונות לנו נקודות $\{(x^\mu, y^\mu)\}_{\mu=1}^{P}$, ואנו רוצים לתאר את הקשר בין x ל y כפונקציה לינארית $\hat{y} = ax + b$.
למדתם בקורס סטטיסטיקה או שיטות מחקר שהערכים שממזערים את השגיאה הריבועית (סכום ריבועי ההפרשים בין $\hat{y}^\mu$ ל y^μ) הם:

$$b^* = \bar{y} - a^* \bar{x}$$

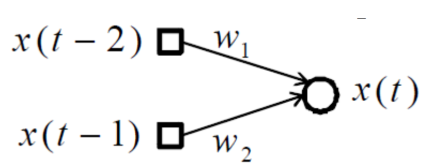
$$a^* = \frac{\mathrm{Cov}\left[x, y\right]}{\mathrm{Var}\left[x\right]} = r_{xy} \frac{s_y}{s_x}$$

כאשר $\bar{x}$ ו $\bar{y}$ הם הממוצעים.

הניחו כעת שכל דוגמא חד־מימדית ממופה לוקטור דו־מימדי באופן הבא: $x^\mu \mapsto \begin{bmatrix} x^\mu \\ 1 \end{bmatrix}$ (וערכי ה y^μ לא משתנים).
הראו כי הפתרון שמתקבל בשיטת היפוך המטריצה שלמדתם בשיעור, $w^* = C^{-1}u$ זהה לפתרון המוצג לעיל במונחי a^*, b^*. כלומר, הוכיחו כי מתקיים $w_1^* = a^*$ ו $w_2^* = b^*$.

שאלה 2

נתון פרספטרון לינארי המנסה לחזות את ערכו של סיגנל מחזורי חד־מימדי בזמן t, באמצעות הערכים בשני צעדי הזמן הקודמים $t-1$ ו $t-2$. כלומר, הפלט של הפרספטרון בזמן t הוא $x\left(t\right) = w_1 x\left(t-2\right) + w_2 x\left(t-1\right)$, כמתואר בציור:



בסעיפים הבאים, נתונים שלושה סוגים של הפלט הרצוי מהפרספטרון כפונקציה של הזמן. שימו לב שהסיגנל מחזורי, ונניח שהוא נמשך זמן ממושך. עליכם למצוא בכל סעיף את w_1, w_2 אשר ממזערים את השגיאה הריבועית בין הפלט הרצוי לפלט המצוי, וכן לחשב את ערכה של שגיאת ההכללה.

1. הסיגנל הרצוי: $1, 1, 1 \ldots$ כלומר $y\left(1\right) = 1$, $y\left(2\right) = 1$, $y\left(3\right) = 1$, $\ldots$
2. הסיגנל הרצוי: $1, 2, 1, 2, 1 \ldots$ כלומר $y\left(1\right) = 1$, $y\left(2\right) = 2$, $y\left(3\right) = 1$, $\ldots$
3. הסיגנל הרצוי: $0, 1, 2, 0, 1, 2 \ldots$ כלומר $y\left(1\right) = 0$, $y\left(2\right) = 1$, $y\left(3\right) = 2$, $\ldots$

שאלה 3

בשאלה זו תשוו בין מספר אלגוריתמי למידה שהוצגו בשיעור לפתרון של בעיית למידה מפוקחת.
הפונקציה שננסה ללמוד היא

$$y = 1 + x + x^2 + x^3$$

כאשר $x \sim \mathcal{U}(-5,5)$ – כלומר x מתפלג בהתפלגות אחידה רציפה בין -5 ל 5.
<u>שימו לב</u>: בכל הסעיפים הבאים, עליכם ללמוד פרספטרון לינארי עם סף – כלומר וקטור המשקולות w יהיה דו־מימדי.

1. צרו $P = 500$ דוגמאות מההתפלגות הנ"ל, וחשבו את הפלט של הפרספטרון הלינארי בשיטת היפוך מטריצת הקורלציה.
   (א) הציגו גרף ובו מופיעה הפונקציה הרצויה (y) לעומת הפלט של הפרספטרון ($\hat{y}$) על כל התחום $x \in [-5,5]$.
   (ב) חשבו את שגיאת האימון ואת שגיאת ההכללה. האם הערכים דומים אחד לשני?
2. חזרו על הלמידה עבור מספר דוגמאות משתנה, P בין 5 ל 100 בקפיצות של 5. לכל ערך של P, ערכו 100 חזרות ומצעו את ערכי שגיאת האימון ושגיאת ההכללה שמתקבלים. הציגו גרף של ממוצעי שגיאת האימון ושגיאת ההכללה כפונקציה של מספר הדוגמאות, והסבירו את התוצאה.
3. השוו בין אלגוריתמי הלמידה שנלמדו בכיתה. ראשית, צרו 100 דוגמאות מההתפלגות, בהן תשתמשו בכל אחד מהסעיפים הבאים:
   (א) <u>למידת גרדיאנט (batch)</u>: הריצו את האלגוריתם 100 צעדי עדכון, וחשבו את שגיאת שגיאת האימון ושגיאת ההכללה לאחר כל צעד.
   (ב) <u>למידת גרדיאנט (online)</u>: הריצו את האלגוריתם על 100 צעדי עדכון (כך שכל דוגמא מוצגת פעם אחת בדיוק), וחשבו את שגיאת שגיאת האימון ושגיאת ההכללה לאחר כל צעד.
   (ג) <u>היפוך מטריצת הקורלציה</u>: מצאו את הפתרון שמתקבל ע"י היפוך מטריצת הקורלציה, וחשבו את שגיאת האימון ושגיאת ההכללה.
4. הציגו את התוצאות של סעיף 3 על אותו גרף, כלומר הציגו את שגיאת האימון ושגיאת ההכללה של כל אחד מהאלגוריתמים כפונקציה של מספר צעדי העדכון (לא לשכוח להוסיף מקרא ברור לגרף), ודונו בהבדלים בין הביצועים של האלגוריתמים השונים.

<u>הערות</u>:

- על מנת לחשב את שגיאת ההכללה, חשבו את ממוצע השגיאה על פני כל התחום הנתון, -5 עד 5, בקפיצות של 0.01.
- עבור אלגוריתמי הגרדיאנט, אתחלו את w באופן שרירותי (למשל $w = \begin{bmatrix}1\\1\end{bmatrix}$, או הגרילו ערכים באקראי).
- עבור אלגוריתמי הגרדיאנט, השתמשו בקצב לימוד של $\eta = 0.01$.