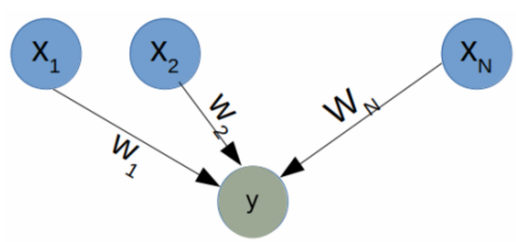
חישוביות וקוגניציה – תרגיל 5

להגשה עד: 12/12/19

<u>שימו לב</u>: שאלה 1 היא שאלה אנליטית ושאלה 2 היא שאלת תכנות

שאלה 1

נתון נוירון לינארי המקבל קלט N מימדי $\mathbf{x}$ (כמתואר בציור) . המשקולות המחברון בין הקלט לנוירון מיוצגים ע"י הוקטור $\mathbf{w}$ כלומר $y = \mathbf{w}^\top \mathbf{x}$. תוחלת הקלט היא אפס, כלומר $\mathbb{E}\left[\mathbf{x}\right] = \mathbf{0}$, ומטריצת הקורלציה של הקלט מסומנת ע"י $\mathbf{C} = \mathbb{E}\left[\mathbf{x}\mathbf{x}^\top\right]$. מטרת הרשת היא למקסם את השונות של הפלט, כפונקציה של $\mathbf{C}, \mathbf{w}$, תחת אילוץ על $\mathbf{w}$.



1. כתבו ביטוי לשונות הפלט (כלומר $\mathrm{Var}\left[y\right]$) כפונקציה של $\mathbf{C}$ ו־ $\mathbf{w}$
2. בהינתן שמוסיפים את האילוץ $\left\|\mathbf{w}\right\|^2 = 1$, מצאו מהו הוקטור $\mathbf{w}$ האופטימלי שממקסם את השונות שמצאתם בסעיף 1
3. נניח שהקלט הוא דו־מימדי ומתפלג באופן הבא: $x_1 \sim \mathcal{N}\left(0,1\right)$ ו $x_2 \sim \mathcal{N}\left(0,4\right)$, ו x_1, x_2 הם בלתי תלויים
   (א) מהי מטריצת הקורלציה של הקלט?
   (ב) מצאו את $\mathbf{w}$ האופטימלי במקרה זה תחת האילוץ של סעיף 2
   (ג) ציירו במישור באופן סכמטי איך ייראה מדגם אופייני של נקודות קלט, וכן ציירו את הכיוון של $\mathbf{w}$
4. נניח כי הקלט הוא חד־מימדי. מהו הערך של $\mathbf{w}$ שממקסם את השונות בפלט, אם אין אילוץ על ערכו של $\mathbf{w}$?
5. השוו את התוצאה שקיבלתם בסעיף 2 לתוצאה שראיתם בכיתה עבור הפתרון האופטימלי להורדת מימד בעזרת נוירון לינארי (PCA). התייחסו לפונקציית המטרה (מטרת האופטימיזציה) ולתוצאה שהתקבלה.

שאלה 2

בשאלה זו תממשו את האלגוריתם של Self Organizing Features Map שנלמד בכיתה.

הדוגמאות שנרצה שהרשת תייצג הינן דו־מימדיות ומיוצרות באופן הבא: ראשית דוגמים את x_1 מהתפלגות אחידה על $[-1,1]$. לאחר מכן בסיכוי $1-p$ דוגמים את x_2 באופן בלתי תלוי מהתפלגות אחידה על $[-1,1]$, ובסיכוי p דוגמים אותו כ $x_2 = \sin\left(f \cdot x_1\right) + \epsilon$ כאשר $\epsilon \sim \mathcal{N}\left(0, \sigma_X^2\right)$.

נגדיר את ה"מפה" להיות רשת של K מייצגים שמסודרים בשרשרת (כלומר, ארגון חד־מימדי). פונקציית הרעש בקידוד, כלומר הסיכוי שדוגמא שהייתה צריכה להיות משוייכת למייצג k תשוייך למייצג $k+v$, מוגדרת כ

$$\pi\left(v\right) = \begin{cases} C\exp\left(-\frac{v^2}{2\sigma^2}\right) & 1 \leq k+v \leq K \\ 0 & \text{otherwise} \end{cases}$$

כאשר C הוא קבוע נרמול שנבחר כך ש $\sum_v \pi\left(v\right) = 1$. שימו לב שהרעש עבור המייצגים בקצוות הוא "קטום". ניתן לממש את הפונקציה π ע"י

```
pifunc = exp(-1*(1/(2*s.^2))*(ks-k).^2)
pifunc = pifunc/sum(pifunc)
```

כאשר ks וקטור שמכיל את האינדקסים של המייצגים ו k האינדקס של המייצג אליו אמורה להיות משוייכת הדוגמא.

<u>שאלות:</u>

1. כתבו פונקציה שמקבלת כפרמטרים את p, f, σ_X ודוגמת נקודה דו מימדית מההתפלגות שהוגדרה.
2. הגרילו $20,000$ דוגמאות מההתפלגות הנ"ל (עם הפרמטרים כפי שמפורט בהמשך), וכן K מייצגים התחלתיים, כאשר הקואורדינטות של כל מייצג נדגמות באקראי ובאופן בלתי תלוי מהתפלגות אחידה על $[-1,1]$.
3. אמנו את הרשת לייצג את הדוגמאות שהוגרלו. עברו על הדוגמאות לפי הסדר ועם כל דוגמא שמוצגת לרשת עדכנו את המייצגים לפי כלל הלמידה של Kohonen.
4. ציירו scatter plot של הנקודות שהוגרלו (לצרכים ויזואליים, הגדירו את השקיפות – alpha – של כל נקודה להיות נמוכה, כדי שניתן יהיה להבדיל בין איזורים במישור עם צפיפות גבוהה של דוגמאות לאיזורים עם צפיפות נמוכה). על אותו גרף, הציגו גם את המייצגים (בצבע שונה) וחברו בקו מקווקוו מייצגים בעלי אינדקסים עוקבים. הציגו גרף אחד כזה עם אתחול הרשת, וגרף אחד כזה לאחר הלמידה.
5. הסבירו את תוצאות הלמידה. בתשובתכם התייחסו גם לנקודות הבאות:
   (א) האם המרחקים בין המייצגים שנלמדו זהים בכל איזור במישור? מדוע?
   (ב) כיצד ישפיע גודל הרעש בקידוד (σ) על התוצאה?

<u>פרמטרים:</u>

לצורך ההתפלגות של הדוגמאות השתמשו בפרמטרים הבאים: $p = 0.95$, $f = 4$, ו $\sigma_X = 0.1$

לצורך אלגוריתם הלמידה, השתמשו בפרמטרים הבאים: $K = 100$, $\eta = 1$, ו $\sigma = 4$ (סטיית התקן של הרעש בקידוד).

<u>מימוש והערות:</u>

- התוצאות עשויות להשתנות מהרצה להרצה עקב המקריות באתחול ובהגרלת הדוגמאות. כדי לקבל "תחושה" לסוג התוצאות, מומלץ להריץ את הקוד מספר פעמים (ניתן לבחור תוצאה "אופיינית" אחת לצורך ההצגה).
- מומלץ (אך לא חובה) לכתוב את כלל העדכון בצורה וקטורית, בלי לעבור בלולאה על המייצגים. שימו לב שההדרכה למימוש של π מגדירה וקטור באורך מספר המייצגים.
- במטלב, הפונקציה min יכולה להחזיר בנוסף לאיבר המינימלי עצמו גם את האינדקס שלו. ב numpy העזרו ב np.argmin.
- בשלב הכתיבה ניתן להתחיל ממספר דוגמאות ומייצגים קטן יותר כדי לחסוך זמן ולמצוא שגיאות בקלות יותר