תומר אליהו
208027780

אישוגות וקוונטיזציה תרגיל 6

(1) נראה ש $\bar{w}=0$ היא נקודת שבת אך לא יציבה של האלגוריתם של Oja (בממוצע).

באמר נראה ש $\langle \Delta \bar{w} \rangle = 0$ כאשר $\langle \Delta \bar{w} \rangle = \langle \eta (y\bar{x} - y^2 \bar{w}) \rangle$

ראשית, נראה ש $\bar{w}=0$ היא נקודת שבת של האלגוריתם:

$$\Delta w = \eta(y\bar{x} - y^2 \bar{w}) \underset{\text{עבור } \bar{w}=0}{=} \eta y \bar{x} = \eta \bar{w}^T \bar{x} \bar{x} \underset{\text{שכן } \bar{w}=0}{=} 0$$

$\rightarrow$ בעבור $\bar{w}=0$ אין שינוי ב Δw ($\Delta w = 0$) ולכן $\bar{w}=0$ היא נקודת שבת של האלגוריתם

כעת נעזר באנליזת יציבות כדי להראות ש $\bar{w}=0$ אינה יציבה;

ראינו שתנאי ההתכנסות בממוצע שקול ל: $\omega\omega^T C \omega = C\bar{\omega}$

וכי קורה רק בשני מקרים, האחד טיפלנו בתרגול, בשני $\bar{w}=0$ נטפל עכשיו;

נגדיר $\bar{w} = 0 + \delta\bar{w}$ כש $\delta\bar{w} = \sum_{k=1}^{n} \varepsilon^k \bar{u}^k$;

וקטורים עצמיים

$$\langle \Delta \bar{w} \rangle = \langle \Delta(0 + \delta\bar{w}) \rangle = \langle \Delta(\delta w) \rangle = \eta\left(C\delta\bar{w} - (\delta\bar{w}^T C \delta\bar{w})\delta\bar{w}\right)$$

$$\approx \eta(C\delta\bar{w}) + O(\|\delta w\|^3)$$

זניח

מסדר ראשון

$$\delta\bar{w} = \sum_{k=1}^{n} \varepsilon^k \bar{u}^k = \sum$$

לצמצם ... של וקטורים עצמיים

$$\Rightarrow \langle \Delta w \rangle \approx \eta \langle C \tilde{w} \rangle = \eta \left\langle C \sum_{k=1}^{n} \varepsilon^k \tilde{u}^k \right\rangle =$$

$$= \eta \left\langle \sum_{k=1}^{n} \varepsilon^k C u^k \right\rangle = \eta \left\langle \sum_{k=1}^{n} \varepsilon^k \lambda^k u^k \right\rangle$$

$\eta > 0$, $\lambda^k > 0$ (כי λ^k ערך עצמי) וקיימים $\varepsilon^k \neq 0$

אז $\langle \Delta w \rangle$ לא יהיה 0 אף פעם ובגלל איטרציה:

מתרחק או מתקרב (Δw ישתנה תמיד).

(2) (א)

$\langle \bar{x} \rangle = 0$

$C = \langle \bar{x} \bar{x}^T \rangle$

$$Var(y) = \mathbb{E}[(y - \mathbb{E}[y])^2] = \mathbb{E}[y^2] - \mathbb{E}[y]^2$$

$$\left(\mathbb{E}[y] = \mathbb{E}[\sum x_i w_i] = \sum \mathbb{E}[x_i] w_i = 0 \right)$$

כי לכל: $\mathbb{E}[x_i] = 0$

$$= \mathbb{E}[y^2] = \mathbb{E}[(\bar{w}^T x)^2] = \mathbb{E}[(\bar{w}^T \bar{x})(\bar{w}^T \bar{x})]$$

$$= \mathbb{E}[\bar{w}^T x x^T w] = w^T \mathbb{E}[x x^T] w$$

$$= w^T C w$$

(ב) נחפש w עבורו $w^T C w$ מקסימלי על הגבלה $\|w\| = 1$ שקול לאילוץ $w^T w = 1 \iff w^T w - 1 = 0$

אז נגדיר $g(w) = w^T w - 1$

ונעבוד עם כופלי לגרנז'

נגדיר:

$$L(w, \lambda) = w^T C w + \lambda (w^T w - 1)$$

נגזור לפי w:

$$\frac{\partial L}{\partial w} = 2C\bar{w} + 2\lambda \bar{w}$$

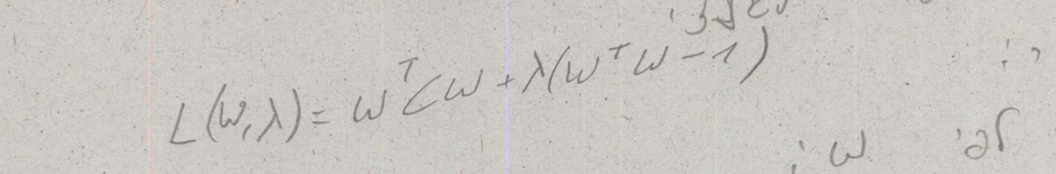
נשווה ל 0

$$2C\bar{w} + 2\lambda \bar{w} = 0$$

$$\Rightarrow C\bar{w} = -\lambda \bar{w}$$

כלומר המקסימום מתקבל עבור $\bar{w}$ שהוא וקטור עצמי (וניקח את הוקטור המתאים לערך העצמי הגבוה ביותר כדי לקבל באמת את המקסימום)

(ג).



(כל הכיוונים עבור w_1, w_2 אופטימליים)

שרירותי

$X_2 \sim N(0,5)$ $\quad$ $X_1 \sim N(0,2)$

נחשב את מטריצת הקורלציה:

$$C = \langle X \cdot X^T \rangle = \left\langle \begin{pmatrix} X_1^2 & X_1X_2 \\ X_1X_2 & X_2^2 \end{pmatrix} \right\rangle$$

$$= \begin{pmatrix} \langle X_1^2 \rangle & \langle X_1X_2 \rangle \\ \langle X_1X_2 \rangle & \langle X_2^2 \rangle \end{pmatrix}$$

$X_1 X_2 = 0$ $\quad$ כי X_1, X_2 בלתי תלויים ולכן בלתי מתואמים.

$$\langle X_1^2 \rangle = Var(X_1) - \langle X_1 \rangle^2 = 1$$

$$Var(X) = \langle X^2 \rangle - \langle X \rangle^2$$

$$\Longrightarrow \langle X_2^2 \rangle = 5$$

$$\Longrightarrow \quad C = \begin{pmatrix} 1 & 0 \\ 0 & 5 \end{pmatrix}$$

אז הוקטורים העצמיים הם $\begin{pmatrix}5\\0\end{pmatrix}, \begin{pmatrix}0\\1\end{pmatrix}$ (כי המטריצה אלכסונית)

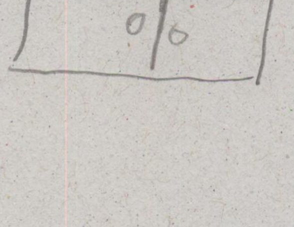
ואם ננרמל כדי לקבל נורמה 1 אז $\begin{pmatrix}0\\1\end{pmatrix}, \begin{pmatrix}1\\0\end{pmatrix}$

הוקטור העצמי $\begin{pmatrix}0\\1\end{pmatrix}$ מתאים לערך העצמי 5 כי $\begin{pmatrix} 1 & 0 \\ 0 & 5 \end{pmatrix}\begin{pmatrix}0\\1\end{pmatrix} = 5 \cdot \begin{pmatrix}0\\1\end{pmatrix}$

ולכן הוא הערך המקסימלי אף השונות של הכיוון הלינארי

(1,5 ערכים עצמיים ו-5 המקסימלי).

הוקטור שמצאנו מתאים ←



ד

נחשב את שונות הפלט למקרה זה

$$Var(y) = w^T C w = w^T \langle x^2 \rangle w = w \langle x^2 \rangle w = w^2 \langle x^2 \rangle$$

משסעיף א

$w^T = w = (w_1)$

נשים לב ש $\langle x^2 \rangle > 0$ (ממוצע של מספרים חיוביים) $\Longrightarrow$

$w^2 \langle x^2 \rangle$ יהיה גדול ככל ש w גדול $\Longrightarrow$

w המקסימלי להיטל זה יהיה $w \to \pm\infty$

תשובות לשאלות מהחלק התנועתי

ט) כן, וקטור המשקלות שמצאנו הגיוני, יש לו נורמה קרובה מאוד ל-1 כפי שהיינו מצפים.

י) ראינו שהרשת מתכנסת לוקטור העצמי שמתאים לערך העצמי הגדול ביותר, ולכן נצפה שאם הפרש הערכים בין שני הוקטורים (העצמיים) יתכנסו ל-0 לעומת הפרש בו מתכנס ל-2 כי אז הוקטור אנו (w) יכול לקיים $w = -u_{max}$).