תרגיל 1

חלק מתמטי

1. צריך להוכיח:

$$\forall T: V \to W, M_T = A\ orthogonal \Rightarrow \forall x \in V, \|Ax\| = \|x\|$$

הוכחה:

$$\|Ax\| = \langle Ax|Ax\rangle \overset{(*)}{=} \langle x|A^t \cdot Ax\rangle = \langle x|x\rangle = \|x\|$$

מעבר (*) נובע מתכונת האופרטור הצמוד.

2. נמצא את הפירוק, לפי ההנחיות:

$$\mathrm{V}\Sigma^{\mathrm{T}}\Sigma\mathrm{V}^{\mathrm{T}} = \begin{bmatrix} 1 & 1 \\ 1 & -1 \\ 0 & 2 \end{bmatrix} \cdot \begin{bmatrix} 1 & 1 & 0 \\ 1 & -1 & 2 \end{bmatrix} = \begin{bmatrix} 2 & 0 & 2 \\ 0 & 2 & -2 \\ 2 & -2 & 4 \end{bmatrix}$$

והע"ע העצמיים שלה הם:

$$\begin{aligned}
\det\left(\begin{bmatrix} 2-\lambda & 0 & 2 \\ 0 & 2-\lambda & -2 \\ 2 & -2 & 4-\lambda \end{bmatrix}\right)
&= (2-\lambda)\det\left(\begin{bmatrix} 2-\lambda & -2 \\ -2 & 4-\lambda \end{bmatrix}\right) + 2\det\left(\begin{bmatrix} 0 & 2 \\ 2-\lambda & -2 \end{bmatrix}\right) \\
&= (2-\lambda)\big((2-\lambda)(4-\lambda) - 4 - 4\big) = (2-\lambda)(\lambda^2 - 6\lambda + 8 - 8) \\
&= \lambda(2-\lambda)(\lambda-6)
\end{aligned}$$

עם וקטורים עצמיים:

$$\ker\left(\begin{bmatrix} 2 & 0 & 2 \\ 0 & 2 & -2 \\ 2 & -2 & 4 \end{bmatrix}\right) = \ker\left(\begin{bmatrix} 1 & 0 & 1 \\ 1 & 1 & 0 \\ -1 & -1 & 0 \end{bmatrix}\right) \to \begin{cases} x_1 = -x_3 \\ x_1 = -x_2 \end{cases} \to v_1 = \begin{bmatrix} -\frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \end{bmatrix}$$

$$\ker\left(\begin{bmatrix} 0 & 0 & 2 \\ 0 & 0 & -2 \\ 2 & -2 & 0 \end{bmatrix}\right) \to \begin{cases} x_3 = 0 \\ x_1 = x_2 \end{cases} \to v_1 = \begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \\ 0 \end{bmatrix}$$

$$\ker\left(\begin{bmatrix} -4 & 0 & 2 \\ 0 & -4 & -2 \\ 2 & -2 & -4 \end{bmatrix}\right) \to \begin{cases} x_1 = \frac{1}{2}x_3 \\ x_1 = -x_2 \\ x_1 - x_2 - 2x_3 = 0 \end{cases} \to v_1 = \begin{bmatrix} \frac{1}{2} \\ -\frac{1}{2} \\ 1 \end{bmatrix}$$

נשים לב שנרמלנו את הוקטורים תוך כדי, כדי שעמודות המטריצה תהיינה בסיס נ"ס.

לפיכך, נסיק כי:

$$V = \begin{bmatrix} -\frac{1}{\sqrt{3}} & \frac{1}{\sqrt{2}} & \frac{1}{2} \\ \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{2}} & -\frac{1}{2} \\ \frac{1}{\sqrt{3}} & 0 & 1 \end{bmatrix}$$

וכן:

$$\mathrm{U\Sigma\Sigma^T U^T} = \begin{bmatrix}1 & 1 & 0\\ 1 & -1 & 2\end{bmatrix}\cdot\begin{bmatrix}1 & 1\\ 1 & -1\\ 0 & 2\end{bmatrix} = \begin{bmatrix}2 & 0\\ 0 & 6\end{bmatrix}$$

וניתן לראות בקלות ש:

$$\Sigma = \begin{bmatrix}2 & 0 & 0\\ 0 & 6 & 0\end{bmatrix}, U = \begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix}$$

ולכן:

$$A = \begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix}\begin{bmatrix}2 & 0 & 0\\ 0 & 6 & 0\end{bmatrix}\begin{bmatrix}-\frac{1}{\sqrt{3}} & \frac{1}{\sqrt{2}} & \frac{1}{2}\\ \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{2}} & -\frac{1}{2}\\ \frac{1}{\sqrt{3}} & 0 & 1\end{bmatrix}$$

3. נראה את ההוכחה לפי הגדרה:

$$\lim_{k\to\infty} b_k = \pm v_1 \Leftrightarrow \lim_{k\to\infty}\frac{\mathrm{C}_0 b_k}{\|C_0 b_k\|} = \pm v_1 \Leftrightarrow \lim_{k\to\infty}\frac{\mathrm{V\Sigma^T\Sigma V^T} b_k}{\|\mathrm{V\Sigma^T\Sigma V^T} b_k\|} = \pm v_1$$
$$\Leftrightarrow \lim_{k\to\infty}\frac{\left(\mathrm{V\Sigma^T\Sigma V^T}\right)^{k-1} b_0}{\|(\mathrm{V\Sigma^T\Sigma V^T})^{k-1} b_0\|} = \pm v_1 \Leftrightarrow \lim_{k\to\infty}\frac{\mathrm{V}\left(\Sigma^{\mathrm{T}}\Sigma\right)^{\mathrm{k}}\mathrm{V^T}\sum_{i=0}^{n} a_i v_i}{\|\mathrm{V}(\Sigma^{\mathrm{T}}\Sigma)^{k}\mathrm{V^T}\sum_{i=0}^{n} a_i v_i\,\|}$$
$$= \pm v_1$$

ולפי הגדרה של כפל בבסיס נורמלי סטנדרטי:

$$\Leftrightarrow \lim_{k\to\infty}\frac{\mathrm{V}\left(\Sigma^{\mathrm{T}}\Sigma\right)^{\mathrm{k}}\sum_{i=1}^{n} a_i e_i}{\|\mathrm{V}(\Sigma^{\mathrm{T}}\Sigma)^{\mathrm{k}}\sum_{i=1}^{n} a_i e_i\,\|} = \pm v_1$$

נמשיך לפשט את הביטוי, נוציא את המקדמים המתאימים החוצה ונגיע לכך ש:

$$\Leftrightarrow \lim_{k\to\infty}\left(\frac{\lambda_1}{|\lambda_1|}\right)^k\left(\frac{a_1}{|a_1|}\right)\frac{v_1 + \frac{1}{a_1}V\left(\frac{1}{\lambda_1}\Sigma^{\mathrm{T}}\Sigma\right)^k\sum_{i=2}^{n} a_i e_i}{\left\|v_1 + \frac{1}{a_1}V\left(\frac{1}{\lambda_1}\Sigma^{\mathrm{T}}\Sigma\right)^k\sum_{i=2}^{n} a_i e_i\right\|} = \pm v_1$$

ידוע לנו לגבי הממשיים כי:

$$\left(\frac{\lambda_1}{|\lambda_1|}\right)^k = \pm 1)$$

וכאשר משאיפים לאינסוף, מקבלים:

$$\left(\frac{1}{\lambda_1}\Sigma^{\mathrm{T}}\Sigma\right)^k \xrightarrow{k\to\infty} 0$$

$$\Leftrightarrow \lim_{k\to\infty} \pm\frac{a_1}{|a_1|}\frac{v_1}{\|v_1\|} = \pm v_1$$

והביטוי הנ"ל נכון עד כדי קבוע, ובפרט פורש את אותו תת-מרחב.

4. ראשית, נבין לעומק את הפונקציה:

$$f(\sigma) = U \cdot diag(\sigma) \cdot U^T \cdot x$$

נסמן:

$$x = \begin{bmatrix} a_1 \\ \vdots \\ a_n \end{bmatrix}, U = \begin{bmatrix} u_{11} & \dots & u_{1n} \\ \vdots & \ddots & \vdots \\ u_{n1} & \dots & u_{nn} \end{bmatrix}$$

לפיכך:

$$f(\sigma) = U \cdot diag(\sigma) \cdot U^T \cdot x = \begin{bmatrix} u_{11} & \dots & u_{1n} \\ \vdots & \ddots & \vdots \\ u_{n1} & \dots & u_{nn} \end{bmatrix} \begin{bmatrix} \sigma_1 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & \sigma_n \end{bmatrix} \begin{bmatrix} u_{11} & \dots & u_{n1} \\ \vdots & \ddots & \vdots \\ u_{1n} & \dots & u_{nn} \end{bmatrix} \begin{bmatrix} a_1 \\ \vdots \\ a_n \end{bmatrix}$$

$$= \begin{bmatrix} \sigma_1 \sum_{i=1}^{n} u_{1i}u_{i1} & \dots & \sigma_1 \sum_{i=1}^{n} u_{1i}u_{in} \\ \vdots & \ddots & \vdots \\ \sigma_n \sum_{i=1}^{n} u_{ni}u_{i1} & \dots & \sigma_n \sum_{i=1}^{n} u_{ni}u_{in} \end{bmatrix} \begin{bmatrix} a_1 \\ \vdots \\ a_n \end{bmatrix}$$

$$= \begin{bmatrix} \sum_{k=1}^{n} \left(\sigma_1 \sum_{i=1}^{n} u_{1i}u_{ik} \right) a_k \\ \vdots \\ \sum_{k=1}^{n} \left(\sigma_n \sum_{i=1}^{n} u_{ni}u_{ik} \right) a_k \end{bmatrix}$$

ולפי זה ש-U היא אורתוגונלית, אז בפרט:

$$\forall k, j \in [n], \sum_{i=1}^{n} u_{ji}u_{ik} = 1$$

שכן המכפלה של המטריצה בשחלופה הוא מטריצת היחידה.

$$= \begin{bmatrix} \sum_{k=1}^{n} (\sigma_1) a_k \\ \vdots \\ \sum_{k=1}^{n} (\sigma_n) a_k \end{bmatrix} = \begin{bmatrix} \sigma_1 \sum_{k=1}^{n} a_k \\ \vdots \\ \sigma_n \sum_{k=1}^{n} a_k \end{bmatrix}$$

לפי הגדרת הדיפרנציאל והיעקוביאן:

$$J_f(\sigma_1, \dots, \sigma_n) = \det\left(\begin{bmatrix} \frac{\partial f_1}{\partial \sigma_1} & \dots & \frac{\partial f_1}{\partial \sigma_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial f_1}{\partial \sigma_1} & \dots & \frac{\partial f_n}{\partial \sigma_n} \end{bmatrix} \right) = \det\left(\begin{bmatrix} \sum_{k=1}^{n} a_k & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & \sum_{k=1}^{n} a_k \end{bmatrix} \right)$$

$$= \left(\sum_{k=1}^{n} a_k \right)^n$$

כנדרש.

5. נפתור זאת בעזרת כלל השרשרת:

$$h(\sigma) = \frac{1}{2} \|f(\sigma) - y\| = \frac{1}{2} \sqrt{\langle f(\sigma) - y | f(\sigma) - y \rangle}$$

$$h'(\sigma)=\frac{1}{2}D_{f(\sigma)-y}(\|\cdot\|^2)D_\sigma(f(\sigma)-y)$$

כאשר עבור $z=\begin{bmatrix}b_1\\ \vdots\\ b_n\end{bmatrix}$ כלשהו,

$$\frac{\mathrm{dg}}{dz}(\|z\|^2)=2z\Rightarrow\frac{1}{2}D_{f(\sigma)-y}(\|\cdot\|^2)=\frac{1}{2}\cdot 2(f(\sigma)-y)=\begin{bmatrix}\sigma_1\sum_{k=1}^{n}a_k-y_1\\ \vdots\\ \sigma_n\sum_{k=1}^{n}a_k-y_n\end{bmatrix}$$

לפיכך:

$$\Delta\left(\frac{1}{2}\|\cdot\|^2\right)=\begin{bmatrix}\sigma_1\sum_{k=1}^{n}a_k-y_1 & \dots & 0\\ \vdots & \ddots & \vdots\\ 0 & \dots & \sigma_n\sum_{k=1}^{n}a_k-y_n\end{bmatrix}\rightarrow\det\left(\frac{1}{2}\|\cdot\|^2\right)$$
$$=\left(\prod_{i=1}^{n}\sigma_n\right)\left(\sum_{k=1}^{n}a_k-y\right)^n$$

$$\Delta(f(\sigma)-y)=\begin{bmatrix}\sum_{k=1}^{n}a_k & \dots & 0\\ \vdots & \ddots & \vdots\\ 0 & \dots & \sum_{k=1}^{n}a_k\end{bmatrix}\rightarrow\det\big(\Delta\big(f(\sigma)\big)-y\big)=\left(\sum_{k=1}^{n}a_k\right)^n$$

ולפי דטרמיננטה היא פונקציה כפלית, אזי:

$$J_{h(x)}=\left(\prod_{i=1}^{n}\sigma_n\right)\left(\sum_{k=1}^{n}a_k-y\right)^n\left(\sum_{k=1}^{n}a_k\right)^n$$

6. ראשית, נתבונן בפונקציה:

$$S_j(x)=\frac{e^{x_j}}{\sum_{\ell=1}^{k}e^{x_\ell}}$$

כעת, נגזור אותה קואורדינטה-קואורדינטה, לפי כלל השרשרת:

$$\frac{\partial S_j}{\partial x_i}=D_{y=e^{x_j}}\left(\frac{y}{\sum_{\ell=1}^{k}e^{x_\ell}}\right)D_{x_i}(e^{x_j})=-\frac{1}{2}\frac{e^{x_j}}{e^{2x_i}}\cdot 0=0$$

אך באלכסון:

$$\frac{\partial S_j}{\partial x_j}=D_{y=e^{x_j}}\left(\frac{y}{\sum_{\ell=1}^{k}e^{x_\ell}}\right)D_{x_j}(e^{x_j})=-\frac{1}{2}\frac{e^{x_j}}{e^{2x_j}}\cdot 1=-\frac{1}{2e^{x_j}}$$

כלומר, לכל כניסה:

$$\Delta(S(x))=\begin{bmatrix}-\frac{1}{2e^{x_1}} & \dots & 0\\ \vdots & \ddots & \vdots\\ 0 & \dots & -\frac{1}{2e^{x_k}}\end{bmatrix}\rightarrow J_S=\prod_{i=1}^{k}-\frac{1}{2e^{x_i}}$$

7. נחשב את ההסיאן:

$$Hessian_{f(x,y)} = \begin{bmatrix} \frac{\partial f}{\partial x^2} & \frac{\partial f}{\partial x \partial y} \\ \frac{\partial f}{\partial y \partial x} & \frac{\partial f}{\partial y^2} \end{bmatrix} = \begin{bmatrix} 6x & -5 \\ -5 & 20y^3 \end{bmatrix}$$

8. נסמן X מ"מ עבור $x_1, x_2 \dots x_n$, עבור אוסף כלשהו בגודל n מתוך המדגמים.
לפי הגדרת תוחלת, ולפי שהמשתנים הם $i.i.d$,

$$\mathbb{E}[X] = \frac{1}{n}\sum_{i=1}^{n} x_i$$

לפי שהמשתנים הם $i.i.d$, ובפרט בלתי תלויים, נובע כי:

$$Var(X) = \mathbb{E}\big[X - \mathbb{E}[X]\big] = \frac{1}{n}\sum_{i=1}^{n}\left(x_i - \frac{1}{n}\sum_{i=1}^{n} x_i\right)$$

נזכיר את אי-שוויון צ'בישב:

$$\Pr(|X - \mathbb{E}[X]| \geq C) \leq \frac{Var(X)}{C^2}$$

לפיכך, עבור המשתנים שלפנינו, וכל ϵ, מתקיים:

$$\forall \epsilon > 0, \Pr\left(\left|X - \frac{1}{n}\sum_{i=1}^{n} x_i\right| \geq \epsilon\right) \leq \frac{Var(X)}{\epsilon^2}$$

כלומר:

$$\forall \epsilon > 0, \Pr(|X - \hat{\mu}_n| \geq \epsilon) \leq \frac{\frac{1}{n}\sum_{i=1}^{n}\left(x_i - \frac{1}{n}\sum_{i=1}^{n} x_i\right)}{\epsilon^2}$$

ובהשאפה לאינסוף:

$$\lim_{n\to\infty}\left(\frac{\frac{1}{n}\sum_{i=1}^{n}\left(x_i - \frac{1}{n}\sum_{i=1}^{n} x_i\right)}{\epsilon^2}\right) = 0$$

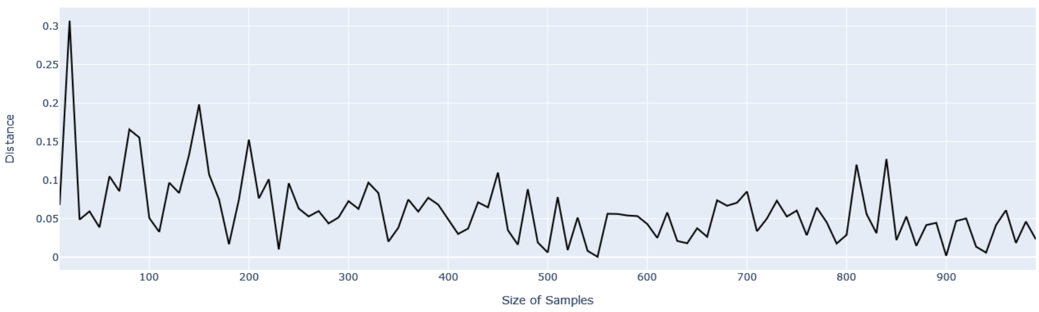
כלומר, $\hat{\mu}_n$ הוא קונסיסטנטי לפי הגדרה

9. בהמשך

חלק יישומי – גרפים

שאלה 2

(1)Distance Between Estimated and True Value of the Expectation as a Function of the Sample Size



(2)the empirical PDF function under the fitted model