321123788 IML - תרגיל 1

אליסון זלמן

1) $v=\begin{bmatrix}1\\2\\3\\4\end{bmatrix}$, $w=\begin{bmatrix}0\\-1\\1\\2\end{bmatrix}$

נמצא כעת את ההטלה של v על w :

$$p=\frac{\langle v,w\rangle}{\|w\|^2}\cdot w=\frac{\left\langle\begin{bmatrix}1\\2\\3\\4\end{bmatrix},\begin{bmatrix}0\\-1\\1\\2\end{bmatrix}\right\rangle}{\left\|\begin{bmatrix}0\\1\\-1\\2\end{bmatrix}\right\|^2}\cdot\begin{bmatrix}0\\-1\\1\\2\end{bmatrix}=\frac{9}{6}\begin{bmatrix}0\\-1\\1\\2\end{bmatrix}=\begin{bmatrix}0\\-1.5\\1.5\\3\end{bmatrix}$$

2) $v=\begin{bmatrix}1\\2\\3\\4\end{bmatrix}$, $w=\begin{bmatrix}1\\0\\1\\-1\end{bmatrix}$

$$p=\frac{\langle v,w\rangle}{\|w\|^2}\cdot w=\frac{0}{3}\cdot w=\vec{0}$$

3) צ"ל: הזווית בין v ל-w היא $\pm 90^\circ$ $\Longleftrightarrow$ $\langle v,w\rangle=0$ (בהנחה $v,w\neq\vec{0}$)

$\Leftarrow$: נתון $\langle v,w\rangle=0$.

מהתרגול הוכחנו $\langle x,y\rangle=\|x\|\cdot\|y\|\cdot\cos\theta$

כאשר θ היא הזווית בין x ל-y. לכן מתקיים:

$$\begin{cases}\langle v,w\rangle=0\\ \langle v,w\rangle=\|v\|\cdot\|w\|\cdot\cos\theta\end{cases}$$
$$\Downarrow$$
$$\|v\|\cdot\|w\|\cdot\cos\theta=0$$

נתון ש- $v,w\neq\vec{0}$, לכן מתכונות הנורמה מתקיים $\|v\|\neq 0 \wedge \|w\|\neq 0$, ונותר להסיק $\cos\theta=0$.

לכן $\theta=\frac{\pi}{2}+\pi k,\ k\in\mathbb{Z}$ $\Longrightarrow$ $\theta=\pm 90^\circ$.

$\Rightarrow$: נתון $\theta=\pm 90^\circ$. נחשב את $\langle v,w\rangle$. לפי ההוכחה

מהתרגול: $\langle v,w\rangle=\|v\|\cdot\|w\|\cdot\cos\theta$, כמו כן

$\cos 90^\circ=\cos(-90^\circ)=0$, ולכן $\langle v,w\rangle=\|v\|\cdot\|w\|\cdot 0=0$

4) נוכיח את השמירה ~~[illegible]~~ האורתונורמליות במטריצה A:

$$\|Ax\|\overset{(1)}{=}\sqrt{(Ax)^T\cdot Ax}=\sqrt{x^TA^T\cdot A\cdot x}\overset{(2)}{=}\sqrt{x^T\cdot x}=\|x\|$$

(1) ראינו בהגדרת ~~[illegible]~~ $\|x\|_2=\sqrt{x\cdot x}$

(2) מתכונות מטריצה אורתוגונלית $A^T\cdot A=I$

5) נפרק SVD: קיימות U ו-V אורתוגונליות

כך ש:

$$A = UDV^T$$

$$\Updownarrow$$

$$AA^{-1} = UDV^TA^{-1}$$

$$\Updownarrow$$

$$I = UDV^TA^{-1}$$

$$\Updownarrow$$

$$U^T = DV^TA^{-1}$$

$$\Updownarrow$$

$$D^{-1}U^T = V^TA^{-1}$$

$$\Updownarrow$$

$$VD^{-1}U^T = A^{-1}$$

מהפיכות U, V, A נובע שגם D הפיכה (אחרת הייתה לה דטרמיננטה 0 ומכפל דטרמיננטות גם $det(A)=0$ בסתירה), לכן נובע ש-D היא מהצורה:

$$D = \begin{bmatrix} \frac{1}{D_{11}} & & \\ & \ddots & \\ & & \frac{1}{D_{nn}} \end{bmatrix}$$

ג. יעיל: פירוק ה-SVD שימושי עבור אלגוריתמים למיניהם

שצריכים למצוא מטריצה הפיכה מפני שהדברים היחידים שצריך

לחשב מתוך פירוק SVD הוא חישוב V, D^{-1} ו-U^T

($O(n^2)$ לכל אחד) וגם כפל מטריצות ב-$O(n^3)$, כך שהאלגוריתם

יהיה בעל זמן ריצה פולינומיאלי ולכן יעיל.

6)*

$$C=\begin{bmatrix}5 & 5\\ -1 & 7\end{bmatrix}, \quad C^T=\begin{bmatrix}5 & -1\\ 5 & 7\end{bmatrix}$$

$$C^TC=\begin{bmatrix}5 & -1\\ 5 & 7\end{bmatrix}\begin{bmatrix}5 & 5\\ -1 & 7\end{bmatrix}=\begin{bmatrix}26 & 18\\ 18 & 74\end{bmatrix}$$

* נמצא את הע"ע של C^TC:

$$\det\left(\begin{bmatrix}\overset{26-x}{\cancel{x-26}} & -18\\ -18 & \underset{74-x}{\cancel{x-74}}\end{bmatrix}\right)=x^2-100x+1924-324=(x-20)(x-80)=0$$

לכן נקבל: $\lambda=20,80$

ונסיק ש:

$$D=\begin{bmatrix}\sqrt{80} & 0\\ 0 & \sqrt{20}\end{bmatrix}$$

כעת נמצא את הו"ע:

$$\lambda=80: \left[\begin{array}{cc|c}26-80 & 18 & 0\\ 18 & 74-80 & 0\end{array}\right]\Rightarrow\left[\begin{array}{cc|c}-54 & 18 & 0\\ 18 & -6 & 0\end{array}\right]\Rightarrow$$

$$\Rightarrow\left[\begin{array}{cc|c}1 & -1/3 & 0\\ 1 & -1/3 & 0\end{array}\right]\Rightarrow\left[\begin{array}{cc|c}1 & -1/3 & 0\\ 0 & 0 & 0\end{array}\right]$$

לכן $V_{80}=\mathrm{span}\left(\begin{bmatrix}1/3\\ 1\end{bmatrix}\right)$

$$\lambda=20: \left[\begin{array}{cc|c}26-20 & 18 & 0\\ 18 & 74-20 & 0\end{array}\right]\Rightarrow\left[\begin{array}{cc|c}6 & 18 & 0\\ 18 & 54 & 0\end{array}\right]\Rightarrow\left[\begin{array}{cc|c}1 & 3 & 0\\ 0 & 0 & 0\end{array}\right]$$

לכן $V_{20}=\mathrm{span}\left(\begin{bmatrix}-3\\ 1\end{bmatrix}\right)$

נבצע תהליך גרם-שמידט למערכת הו"ע $\left\{\begin{bmatrix}1/3\\ 1\end{bmatrix},\begin{bmatrix}-3\\ 1\end{bmatrix}\right\}$:

$$u_1 = \frac{V_1}{\|V_1\|} = \frac{\begin{bmatrix} 1/3 \\ 1 \end{bmatrix}}{\sqrt{\frac{10}{9}}} = \frac{\begin{bmatrix} 1 \\ 3 \end{bmatrix}}{\sqrt{10}} = \begin{bmatrix} 1/\sqrt{10} \\ 3/\sqrt{10} \end{bmatrix}$$

$$w_2 = v_2 - \langle v_2 | u_1 \rangle u_1 = \begin{bmatrix} -3 \\ 1 \end{bmatrix} - \left\langle \begin{bmatrix} -3 \\ 1 \end{bmatrix} \middle| \begin{bmatrix} 1/\sqrt{10} \\ 3/\sqrt{10} \end{bmatrix} \right\rangle \begin{bmatrix} 1/\sqrt{10} \\ 3/\sqrt{10} \end{bmatrix} =$$

$$= \begin{bmatrix} -3 \\ 1 \end{bmatrix} - 0 \cdot \begin{bmatrix} 1/\sqrt{10} \\ 3/\sqrt{10} \end{bmatrix} = \begin{bmatrix} -3 \\ 1 \end{bmatrix} \Rightarrow u_2 = \frac{1}{\sqrt{10}} \begin{bmatrix} -3 \\ 1 \end{bmatrix} = \begin{bmatrix} -3/\sqrt{10} \\ 1/\sqrt{10} \end{bmatrix}$$

לכן:

$$V = \begin{bmatrix} 1/\sqrt{10} & -3/\sqrt{10} \\ 3/\sqrt{10} & 1/\sqrt{10} \end{bmatrix}$$

$$CV = \begin{bmatrix} 5 & 5 \\ -1 & 7 \end{bmatrix} \begin{bmatrix} 1/\sqrt{10} & -3/\sqrt{10} \\ 3/\sqrt{10} & 1/\sqrt{10} \end{bmatrix} = \begin{bmatrix} 20/\sqrt{10} & -\frac{10}{\sqrt{10}} \\ 20/\sqrt{10} & 10/\sqrt{10} \end{bmatrix}$$

$$UD = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} \sqrt{80} & 0 \\ 0 & \sqrt{20} \end{bmatrix} = \begin{bmatrix} a\sqrt{80} & b\sqrt{20} \\ c\sqrt{80} & d\sqrt{20} \end{bmatrix}$$

$$\begin{cases} \frac{20}{\sqrt{10}} = a\sqrt{80} \\ \frac{-10}{\sqrt{10}} = b\sqrt{20} \\ \frac{20}{\sqrt{10}} = c\sqrt{80} \\ \frac{10}{\sqrt{10}} = d\sqrt{20} \end{cases} \Rightarrow \begin{cases} a = \sqrt{\frac{1}{2}} \\ b = -\sqrt{\frac{1}{2}} \\ c = \sqrt{\frac{1}{2}} \\ d = \sqrt{\frac{1}{2}} \end{cases}$$

לסיכום:

$$U = \begin{bmatrix} 1/\sqrt{2} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix}, \quad D = \begin{bmatrix} \sqrt{80} & 0 \\ 0 & \sqrt{20} \end{bmatrix}, \quad V = \begin{bmatrix} \frac{1}{\sqrt{10}} & -\frac{3}{\sqrt{10}} \\ \frac{3}{\sqrt{10}} & \frac{1}{\sqrt{10}} \end{bmatrix}$$

7) תחילה: נראה ש-

$$b_{k+1} = \frac{C_0^{k+1} b_0}{\|C_0^{k+1} b_0\|}$$

נוכיח זאת באינדוקציה על k:

$k=1$:

לפי הגדרת b_{k+1}:

$$b_1 = \frac{C_0 b_0}{\|C_0 b_0\|}$$

נניח ל-k ונוכיח עבור $k+1$:

$$b_{k+1} = \frac{C_0 b_k}{\|C_0 b_k\|} \overset{\text{ה.א.}}{=} \frac{\frac{C_0 \cdot C_0^k \cdot b_0}{\|C_0^k b_0\|}}{\left\| \frac{C_0 \cdot C_0^k b_0}{\|C_0^k \cdot b_0\|} \right\|} = \frac{\frac{C_0^{k+1} b_0}{\|C_0^k b_0\|}}{\frac{1}{\|C_0^k b_0\|} \cdot \|C_0^{k+1} b_0\|} =$$

$$= \frac{C_0^{k+1} b_0}{\|C_0^{k+1} b_0\|}$$

כעת נתבונן ב-b_k:

$$b_k = \frac{C_0^k b_0}{\|C_0^k b_0\|} \overset{(*)}{=} \frac{U D^k U^T b_0}{\|U D^k U^T b_0\|}$$

(*) לפי פירוק EVD של מטריצה סימטרית C_0 ומכך שעבור U אורתונורמלית מתקיים: $U^T U = I$

כעת נציב ב $b_0 = (a_1 v_1 + \dots + a_n v_n)$ כאשר $(v_1, \dots, v_n)$

גם ז"א ניתן לכתוב את אלכסון D ומרכיבים את עמודות המטריצה U.

$$= \frac{UD^k U^T (a_1 v_1 + \dots + a_n v_n)}{\| UD^k U^T (a_1 v_1 + \dots + a_n v_n) \|} = \frac{UD^k (a_1 e_1 + \dots + a_n e_n)}{\| UD^k (a_1 e_1 + \dots + a_n e_n) \|}$$

המעבר הזה נובע מכך ש:

$$\forall i \in [n] \quad U^T a_i v_i = \begin{bmatrix} - & v_1 & - \\ & \vdots & \\ - & v_n & - \end{bmatrix} a_i v_i = \sum_{j=1}^{n} v_j \cdot a_i v_i = a_i e_i$$

מכיוון ש $(v_1, \dots, v_n)$ אורתונורמלים.

$$= \frac{UD^k a_1 e_1 + UD^k (a_2 e_2 + \dots + a_n e_n)}{\| UD^k a_1 e_1 + UD^k (a_2 e_2 + \dots + a_n e_n) \|}$$

כעת נשים לב ש:

$$UD^k a_1 e_1 = \begin{bmatrix} | & & | \\ v_1 & \cdots & v_n \\ | & & | \end{bmatrix} \begin{bmatrix} \lambda_1^k & & \\ & \ddots & \\ & & \lambda_n^k \end{bmatrix} \begin{bmatrix} a_1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} = \begin{bmatrix} | & & | \\ v_1 & \cdots & v_n \\ | & & | \end{bmatrix} \begin{bmatrix} a_1 \lambda_1^k \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

$$= a_1 \lambda_1^k v_1$$

לכן:

$$= \frac{a_1 \lambda_1^k v_1 + UD^k (a_2 e_2 + \dots + a_n e_n)}{\| a_1 \lambda_1^k v_1 + UD^k (a_2 e_2 + \dots + a_n e_n) \|} = \left(\frac{\lambda_1}{|\lambda_1|} \right)^k \cdot \frac{a_1}{|a_1|} \cdot \frac{v_1 + \frac{1}{a_1} U \left(\frac{1}{\lambda_1} D \right)^k \cdot \sum_{i=2}^{n} a_i e_i}{\| v_1 + \frac{1}{a_1} U \left(\frac{1}{\lambda_1} D \right)^k \cdot \sum_{i=2}^{n} a_i e_i \|}$$

מכיוון שגם מתקיים ש:

$$\lim_{k \to \infty} \left(\frac{1}{\lambda_1} D \right)^k = \lim_{k \to \infty} \begin{bmatrix} 1 & & & \\ & \left(\frac{\lambda_2}{\lambda_1}\right)^k & & \\ & & \ddots & \\ & & & \left(\frac{\lambda_n}{\lambda_1}\right)^k \end{bmatrix} = \begin{bmatrix} 1 & & \\ & 0 & \\ & & \ddots & \\ & & & 0 \end{bmatrix}$$

מכיוון שאם $|\lambda_1| > |\lambda_2| \geq \dots \geq |\lambda_n|$

לכן נקבל:

$$\frac{1}{a_1} U \cdot \left(\frac{1}{\lambda_1} D\right)^k (a_2 e_2 + \dots + a_n e_n) \xrightarrow[k\to\infty]{} \frac{1}{a_1} U \cdot 0_{n\times 1} = 0$$

ומכיוון ש:

$$\left(\frac{\lambda_1}{|\lambda_1|}\right)^k = \pm 1 \ , \quad \frac{a_1}{|a_1|} = \pm 1$$

נקבל אם כן:

$$\lim_{k\to\infty} b_k = \lim_{k\to\infty} \left(\frac{\lambda_1}{|\lambda_1|}\right)^k \frac{a_1}{|a_1|} \frac{v_1 + \frac{1}{a_1} U \left(\frac{1}{\lambda_1} D\right)^k (a_2 e_n + \dots + a_n e_n)}{\| v_1 + \frac{1}{a_1} U \left(\frac{1}{\lambda_1} D\right)^k (a_2 e_2 + \dots + a_n e_n) \|} =$$

$$= (\pm 1) \cdot (\pm 1) \cdot \frac{v_1 + 0}{\| v_1 + 0 \|} = \pm v_1$$

8) נכתוב הפונק' את $f(\sigma)$:

$$f(\sigma) = U\,\mathrm{diag}(\sigma)\,U^T x = \begin{bmatrix} | & & | \\ u_1 & \cdots & u_n \\ | & & | \end{bmatrix} \begin{bmatrix} \sigma_1 & & 0 \\ & \ddots & \\ 0 & & \sigma_n \end{bmatrix} \begin{bmatrix} - & u_1 & - \\ & \vdots & \\ - & u_n & - \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} =$$

$$= \begin{bmatrix} | & & | \\ u_1 & \cdots & u_n \\ | & & | \end{bmatrix} \begin{bmatrix} \sigma_1 & & \\ & \ddots & \\ & & \sigma_n \end{bmatrix} \begin{bmatrix} \langle u_1 | x \rangle \\ \vdots \\ \langle u_n | x \rangle \end{bmatrix} = \begin{bmatrix} | & & | \\ u_1 & \cdots & u_n \\ | & & | \end{bmatrix} \begin{bmatrix} \sigma_1 \langle u_1 | x \rangle \\ \vdots \\ \sigma_n \langle u_n | x \rangle \end{bmatrix} =$$

$$= \begin{bmatrix} \sum_{j=1}^{n} (U_{1j} \cdot \sigma_j \langle u_1 | x \rangle) \\ \vdots \\ \sum_{j=1}^{n} (U_{nj} \cdot \sigma_j \langle u_n | x \rangle) \end{bmatrix}$$

כעת נוכל לחשב קואורדינטה כללי i,j במטריצת היעקוביאן:

$$[J_\sigma(f)]_{ij} = \frac{\partial f_i(\sigma)}{\partial \sigma_j} = \left(\sum_{j=1}^{n} U_{ij} \cdot \sigma_j \langle u_i | x \rangle \right) \frac{\partial}{\partial \sigma_j} = \boxed{U_{ij} \langle u_j | x \rangle}$$

ט) נחשב תחילה את $h(\sigma)$:

$$h(\sigma)=\frac{1}{2}\|f(\sigma)-y\|^2=\frac{1}{2}\left((f(\sigma)-y)(f(\sigma)-y)^T\right)$$

$$=\frac{1}{2}\left(\|f(\sigma)\|^2-2f(\sigma)^Ty+\|y\|^2\right)=\frac{1}{2}\|f(\sigma)\|^2-f(x)^Ty+\frac{1}{2}\|y\|^2$$

לפי כלל השרשרת:

$$\nabla h=\nabla(g\circ f)^T\cdot\nabla f$$

ולכן:

$$\nabla h=\left(\frac{1}{2}\cdot 2\cdot f(\sigma)-y\right)^T\cdot\nabla f(\sigma)=$$

$$=(f(\sigma)-y)^T\cdot\nabla f(\sigma)$$

י) בתרגיל פתרנו המקרה $i=j$. כעת עבור $j\neq i$ מתקיים:

$$\frac{\partial g(z)_j}{\partial z_j}=\frac{\partial}{\partial z_j}\frac{e^{z_i}}{\sum_{k=1}^{K}e^{z_k}}=\frac{0-e^{z_i}e^{z_j}}{\left(\sum_{k=1}^{K}e^{z_k}\right)^2}=$$

$$=-\left(\frac{e^{z_i}}{\sum_{k=1}^{K}e^{z_i}}\cdot\frac{e^{z_j}}{\sum_{k=1}^{K}e^{z_j}}\right)=-g(z)_j\cdot g(z)_i$$

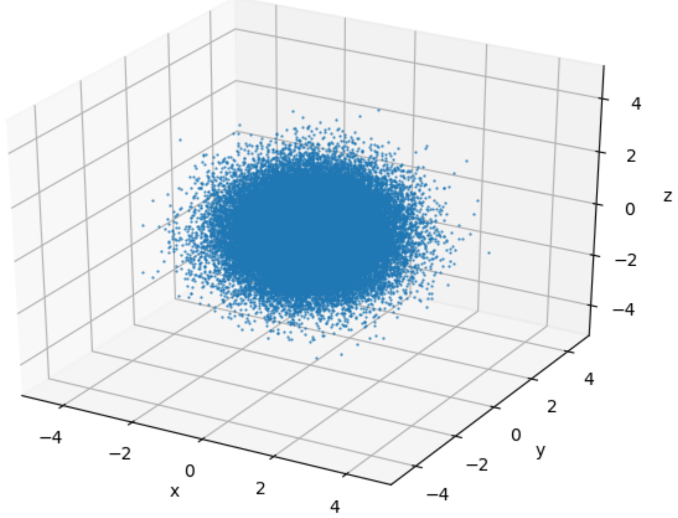
12) הגרף הצטמצם וקטן יותר על צירי ה-x וה-y, והתרחב על פני ציר ה-z, כלומר מה שהגדרנו במטריצה הם הכפלות מכפלה: את x ב-0.1, y ב-0.5, ו-z ב-2.

13) הגרף נשאר זהה מבחינת הצורה שלו, אבל כעת הוא מסובב בזווית.

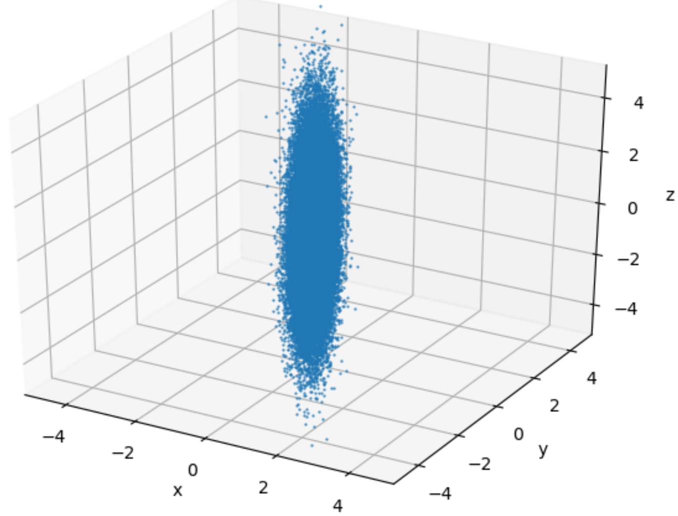
14) הגרף נראה כמו הגרף משאלה 13 רק משוקף על המישור הנוצר ע"י הצירים x ו-y.

15) הגרף כעת נראה מצומצם יותר.

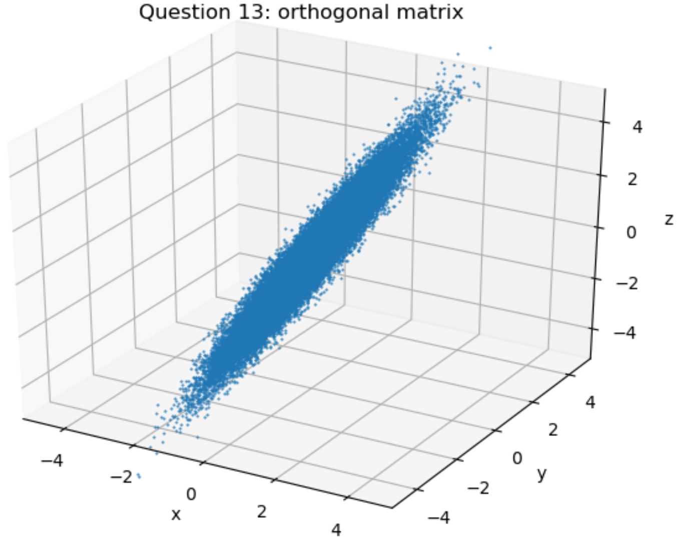
Question 11: identity matrix



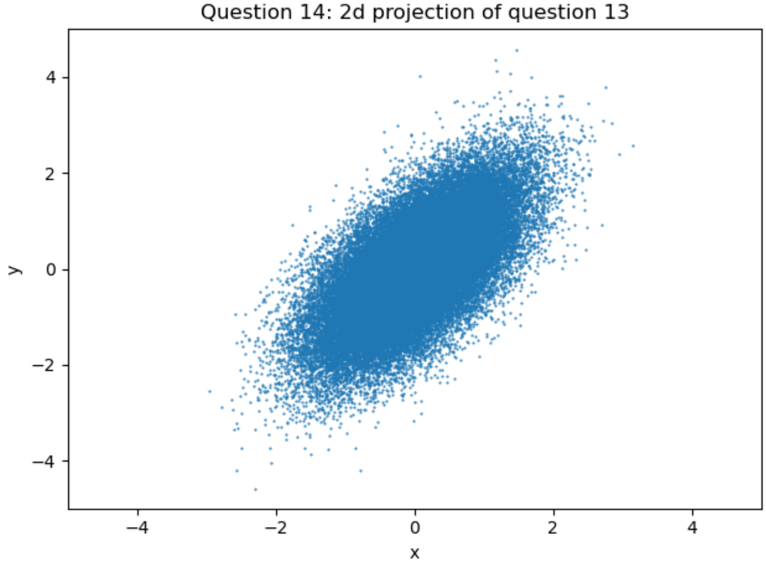
Question 12: scaling matrix



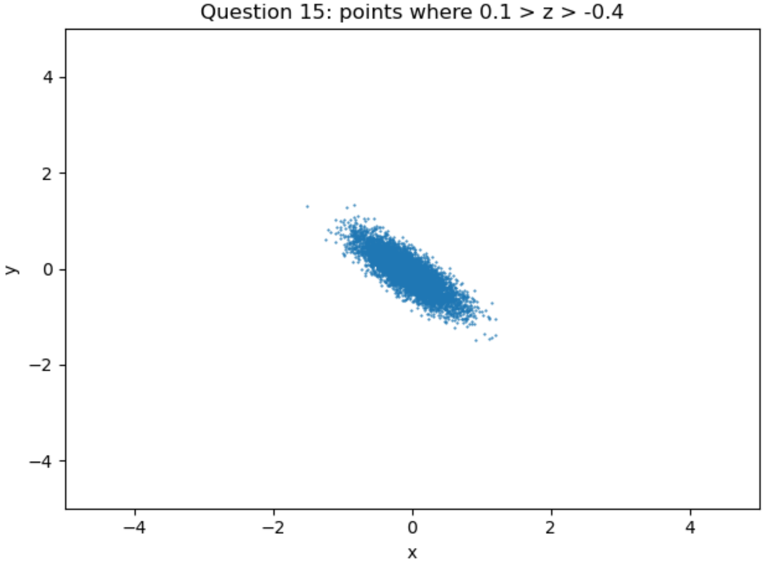
Question 13: orthogonal matrix



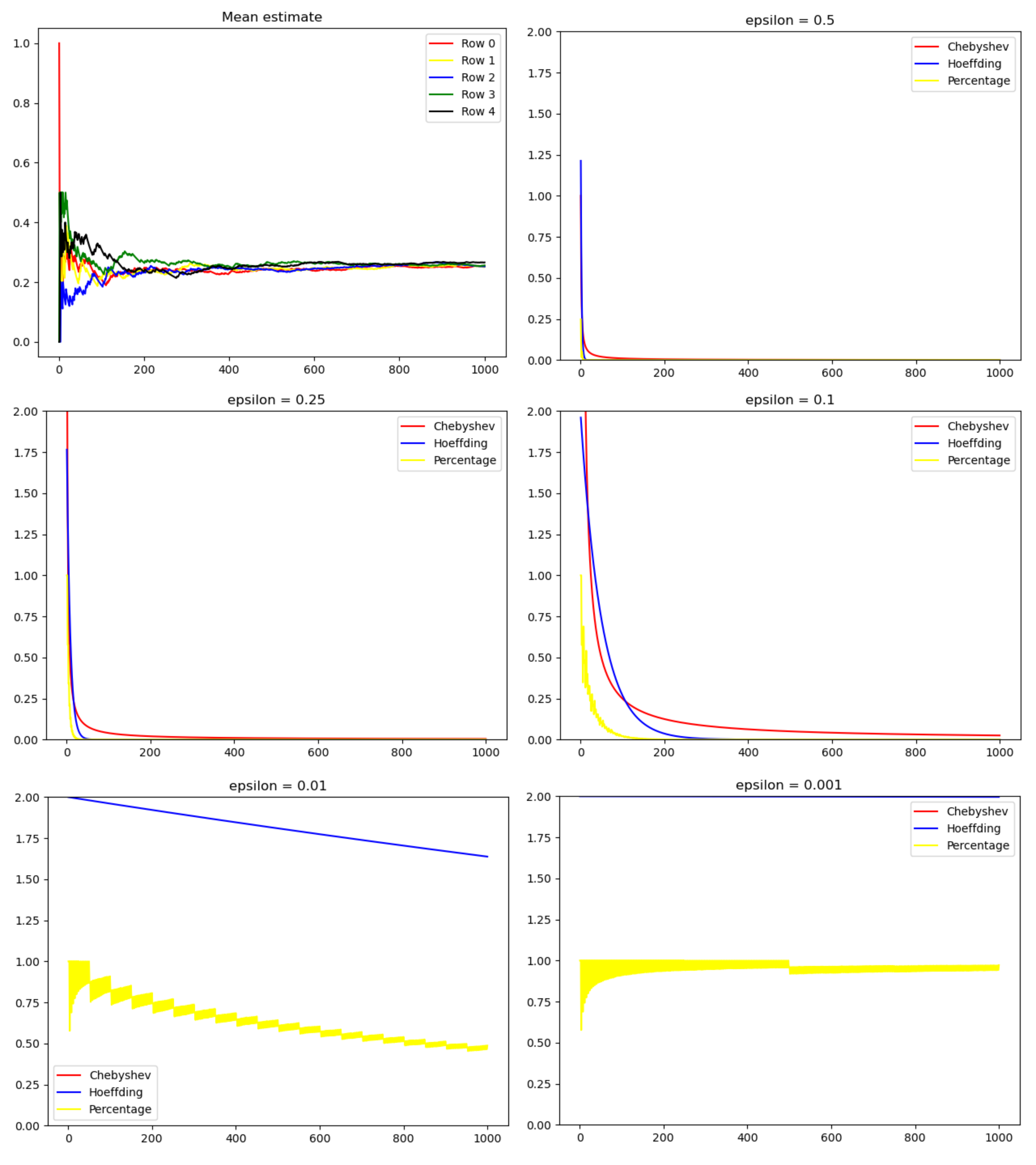
Question 14: 2d projection of question 13



Question 15: points where 0.1 > z > -0.4



... שיש בו יותר חישוב הערכים יתכנסו
לסביבת הערך 0.25 שזאת התוצאה עבור אפסילון.
(6) אני מצפה שככל שאפסילון קטן, כך הגרף יהיה קרוב יותר
לערך $y=1$ (ככל ש-m גדול יותר). זאת מכיוון שאפסילון
בעצם מייצג את הדיוק ~~[illegible]~~ שאנחנו דורשים מהמצב, ולכן
ככל שנדרוש תוצאה מדויקת יותר, כך יהיו דומים יותר
הצפיים אשר יפרו אותה.



Mean estimate
Row 0
Row 1
Row 2
Row 3
Row 4
epsilon = 0.5
epsilon = 0.25
epsilon = 0.1
epsilon = 0.01
epsilon = 0.001
Chebyshev
Hoeffding
Percentage