פתרון תרגיל מספר 1 ־ מערכות לומדות

שם: מיכאל גרינבאום, **ת.ז:** 211747639

9 באפריל 2020

1. **צ"ל:** הטלה של $v=(1,2,3,4)$ על הוקטור $w=(0,-1,1,2)$
הוכחה:

$$\langle v,w\rangle = 1\cdot 0+2\cdot -1+3\cdot 1+4\cdot 2=-2+3+8=9$$
$$\|w\|^2=\langle w,w\rangle = 0^2+(-1)^2+1^2+2^2=2+4=6$$

אמרנו בתרגול שההטלה של v על w היא

$$\langle v,w\rangle\cdot\frac{w}{\|w\|^2}=\frac{9}{6}\cdot(0,-1,1,2)=\frac{3}{2}\cdot(0,-1,1,2)=\left(0,\frac{-3}{2},\frac{3}{2},3\right)$$

כלומר ההטלה היא $\boxed{\left(0,\frac{-3}{2},\frac{3}{2},3\right)}$

מ.ש.ל.☺

2. **צ"ל:** הטלה של $v=(1,2,3,4)$ על הוקטור $w=(1,0,1,-1)$
הוכחה:

$$\langle v,w\rangle = 1\cdot 1+2\cdot 0+3\cdot 1+4\cdot -1=1+3-4=0$$
$$\|w\|^2=\langle w,w\rangle = 1^2+0^2+1^2+(-1)^2=3$$

אמרנו בתרגול שההטלה של v על w היא

$$\langle v,w\rangle\cdot\frac{w}{\|w\|^2}=\frac{0}{3}\cdot(1,0,1,-1)=(0,0,0,0)$$

כלומר ההטלה היא $\boxed{(0,0,0,0)}$

מ.ש.ל.☺

3. **צ"ל:** $\langle v,w\rangle=0$ אם"ם הזווית בין v ל־ w זה $\theta=\pm 90$
הוכחה:
יהיו $0\neq v,w\in\mathbb{R}^m$
$\Leftarrow$: נניח כי הזווית בין v ל־ w זה $\theta=\pm 90$, אזי

$$\langle v,w\rangle=\|v\|\cdot\|w\|\cdot\cos(\theta)=\|v\|\cdot\|w\|\cdot 0=0$$

כנדרש
$\Rightarrow$: נניח כי $\langle v,w\rangle=0$, נסמן את הזווית בינהם ב־ θ, נשים לב כי $-180\leq\theta\leq 180$, וגם

$$0=\langle v,w\rangle=\|v\|\cdot\|w\|\cdot\cos(\theta)$$

נשים לב כי $\|v\|,\|w\|\neq 0$ כי הנחנו כי $v,w\neq 0$ ולכן $\cos(\theta)=\frac{0}{\|v\|\cdot\|w\|}=0$,
עתה נשים לב שהזוויות היחידות המקיימות את השוויון בטווח המתאים הוא $\theta=\pm 90$, כנדרש

מ.ש.ל.☺

4. **צ"ל:** $\|Ax\|_2 = \|x\|_2$
הוכחה:
נשים לב כי A אורתוגונלית ולכן $A \cdot A^T = A^T \cdot A = I_n$,
במקרה הפרטי של נורמה אוקלידית נקבל

$$\|Ax\|_2 \stackrel{\text{def}}{=} \sqrt{\langle Ax, Ax\rangle} \stackrel{\text{def}}{=} \sqrt{(Ax)^T \cdot Ax} = \sqrt{x^T \cdot A^T \cdot A \cdot x} = \sqrt{x^T \cdot I_n \cdot x} = \sqrt{x^T \cdot x} \stackrel{\text{def}}{=} \sqrt{\langle x, x\rangle} \stackrel{\text{def}}{=} \|x\|_2$$

עתה במקרה הכללי למכפלה פנימית שמושרת מנורמה מתקיים $A^* = A^T$ כי A אורתוגונלית

$$\|Ax\|_2 \stackrel{\text{def}}{=} \sqrt{\langle Ax, Ax\rangle} = \sqrt{\langle A^* \cdot Ax, x\rangle} = \sqrt{\langle I_n \cdot x, x\rangle} = \sqrt{\langle x, x\rangle} \stackrel{\text{def}}{=} \|x\|_2$$

מ.ש.ל.☺

5. **צ"ל:** נוסחא ל- A^{-1} ונימוק מדוע זה חשוב
הוכחה:
נשים לב כי $A = U \cdot D \cdot V^T$,
נתון ש- U, V הן אורתוגונליות ולכן $U^{-1} = U^T, V^T = V^{-1}$ וגם $|\det(U)| = |\det(V)| = 1$,
עתה נשים לב כי $0 \neq \det(A)$ כי A הפיכה ולכן

$$\begin{aligned} 0 \neq |\det(A)| &= \left|\det\left(U \cdot D \cdot V^T\right)\right| = \left|\det(U) \cdot \det\left(V^T\right) \cdot \det(D)\right| \\ &= |\det(U)| \cdot \left|\det\left(V^T\right)\right| \cdot |\det(D)| = 1 \cdot |\det(D)| = |\det(D)| \end{aligned}$$

לכן קיבלנו כי $|\det(D)| \neq 0$, כלומר $\det(D) \neq 0$, כלומר D הפיכה, כלומר U, V, D הפיכות, עתה נשים לב כי

$$A^{-1} = \left(U \cdot D \cdot V^T\right)^{-1} = \left(V^T\right)^{-1} \cdot D^{-1} \cdot U^{-1} = V \cdot D^{-1} \cdot U^T$$

כלומר הנוסחא שקיבלנו היא $\boxed{A^{-1} = V \cdot D^{-1} \cdot U^T}$,
נשים לב שאת V, U^T אפשר לחשב בזמן n^2 ואת D^{-1} בזמן n כי רק צריך לעשות $\frac{1}{\sigma_i}$ לכל איבר באלכסון כדי לחשב הופכית של מטריצה אלכסונית,
ולכן אפשר לחשב את A^{-1} בזמן n^2 + זמן הכפלת המטריצות = $O\left(n^2 + n^\omega\right) = O\left(n^\omega\right) \sim O\left(n^{2.31}\right)$,
נשים לב שלפי דירוג גאוס, היינו יכולים לחשב הופכית בזמן $O\left(n^3\right)$, כלומר מצאנו שיטה יותר יעילה למצוא מטריצה הפוכה.

מ.ש.ל.☺

6. **צ"ל:** מציאת פירוק SVD של C
הוכחה:

$$V \cdot D^2 \cdot V^T = C^T \cdot C = \begin{bmatrix} 26 & 18 \\ 18 & 74 \end{bmatrix}$$

נשים לב כי עבור $v_1 = (-3, 1)$ מתקיים

$$C^T \cdot C \cdot v_1 = \begin{bmatrix} 26 & 18 \\ 18 & 74 \end{bmatrix} \cdot \begin{bmatrix} -3 \\ 1 \end{bmatrix} = \begin{bmatrix} 18 - 78 \\ 74 - 54 \end{bmatrix} = \begin{bmatrix} -60 \\ 20 \end{bmatrix} = 20 \cdot \begin{bmatrix} -3 \\ 1 \end{bmatrix} = 20 \cdot v_1$$

לכן v_1 הוא וקטור עצמי לערך 20 ולכן $\hat{v}_1 = \frac{v_1}{|v_1|} = \left(\frac{-3}{\sqrt{10}}, \frac{1}{\sqrt{10}}\right)$ ערך עצמי עם נורמה 1
נשים לב כי עבור $v_1 = (1, 3)$ מתקיים

$$C^T \cdot C \cdot v_1 = \begin{bmatrix} 26 & 18 \\ 18 & 74 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 3 \end{bmatrix} = \begin{bmatrix} 26 + 54 \\ 18 + 222 \end{bmatrix} = \begin{bmatrix} 80 \\ 240 \end{bmatrix} = 80 \cdot \begin{bmatrix} 1 \\ 3 \end{bmatrix} = 80 \cdot v_2$$

לכן v_2 הוא וקטור עצמי לערך 80 ולכן $\hat{v_2}=\frac{v_2}{|v_2|}=\left(\frac{1}{\sqrt{10}},\frac{3}{\sqrt{10}}\right)$ ערך עצמי עם נורמה 1
ולכן $\{\hat{v_1},\hat{v_2}\}$ זה בסיס אורתונורמלי של וקטורים עצמיים, כלומר נוכל למצוא כי

$$V=V^T=\begin{bmatrix}\frac{-3}{\sqrt{10}} & \frac{1}{\sqrt{10}}\\ \frac{1}{\sqrt{10}} & \frac{3}{\sqrt{10}}\end{bmatrix}, D^2=\begin{bmatrix}20 & 0\\ 0 & 80\end{bmatrix}\Rightarrow D=\begin{bmatrix}\sqrt{20} & 0\\ 0 & \sqrt{80}\end{bmatrix}$$

עכשיו אנו יודעים כי

$$\begin{aligned}
C\cdot V=U\cdot D &\Rightarrow \boxed{U=C\cdot V\cdot D^{-1}}\\
U &= \begin{bmatrix}5 & 5\\ -1 & 7\end{bmatrix}\cdot\begin{bmatrix}\frac{-3}{\sqrt{10}} & \frac{1}{\sqrt{10}}\\ \frac{1}{\sqrt{10}} & \frac{3}{\sqrt{10}}\end{bmatrix}\cdot\begin{bmatrix}\frac{1}{\sqrt{20}} & 0\\ 0 & \frac{1}{\sqrt{80}}\end{bmatrix}\\
&= \begin{bmatrix}-\sqrt{10} & 2\cdot\sqrt{10}\\ \sqrt{10} & 2\cdot\sqrt{10}\end{bmatrix}\cdot\frac{1}{\sqrt{20}}\cdot\begin{bmatrix}1 & 0\\ 0 & \frac{1}{2}\end{bmatrix}\\
&= \begin{bmatrix}-\sqrt{10} & \sqrt{10}\\ \sqrt{10} & \sqrt{10}\end{bmatrix}\cdot\frac{1}{\sqrt{20}}=\begin{bmatrix}\frac{-\sqrt{2}}{2} & \frac{\sqrt{2}}{2}\\ \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2}\end{bmatrix}
\end{aligned}$$

לכן קיבלנו כי

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

מ.ש.ל.☺

7. **צ"ל:** הוכחה שהאלגוריתם שואף ל - $\pm v_1$
הוכחה:
תחילה נוכיח באינדוקציה כי $b_{k+1}=\frac{C_0^k\cdot b_0}{\|C_0^k\cdot b_0\|}$
<u>בסיס:</u> $k=0$: נשים לב כי

$$b_1\overset{\text{def}}{=}\frac{C_0\cdot b_0}{\|C_0\cdot b_0\|}$$

כנדרש
<u>צעד:</u> $k-1\to k$: נשים לב כי

$$\begin{aligned}
b_{k+1}=\frac{C_0\cdot b_k}{\|C_0\cdot b_k\|} &\overset{\text{induction}}{=}\frac{C_0\cdot\frac{C_0^{k-1}\cdot b_0}{\|C_0^{k-1}\cdot b_0\|}}{\left\|C_0\cdot\frac{C_0^{k-1}\cdot b_0}{\|C_0^{k-1}\cdot b_0\|}\right\|}=\frac{C_0\cdot C_0^{k-1}\cdot b_0}{\left\|C_0\cdot\left\|C_0^{k-1}\cdot b_0\right\|\cdot\frac{C_0^{k-1}\cdot b_0}{\|C_0^{k-1}\cdot b_0\|}\right\|}\\
&=\frac{C_0\cdot C_0^{k-1}\cdot b_0}{\left\|C_0\cdot C_0^{k-1}\cdot b_0\right\|}=\frac{C_0^k\cdot b_0}{\left\|C_0^k\cdot b_0\right\|}
\end{aligned}$$

כנדרש,
עתה נסמן $b_0=\sum_{i=1}^n a_i\cdot v_i$, לכן מתקיים

$$C_0^k\cdot b_0=C_0^k\cdot\left[\sum_{i=1}^n a_i\cdot v_i\right]=\sum_{i=1}^n a_i\cdot C_0^k\cdot v_i=\sum_{i=1}^n a_i\cdot\lambda_i^k\cdot v_i=a_1\cdot\lambda_1^k\cdot\left[v_1+\sum_{i=2}^n\frac{a_i}{a_1}\cdot\left(\frac{\lambda_i}{\lambda_1}\right)^k\cdot v_i\right]$$

נציב ונקבל

$$b_{k+1}=\frac{C_0^k\cdot b_0}{\left\|C_0^k\cdot b_0\right\|}=\frac{a_1\cdot\lambda_1^k\cdot\left[v_1+\sum_{i=2}^n\frac{a_i}{a_1}\cdot\left(\frac{\lambda_i}{\lambda_1}\right)^k\cdot v_i\right]}{\left\|a_1\cdot\lambda_1^k\cdot\left[v_1+\sum_{i=2}^n\frac{a_i}{a_1}\cdot\left(\frac{\lambda_i}{\lambda_1}\right)^k\cdot v_i\right]\right\|}$$
$$=\frac{a_1\cdot\lambda_1^k\cdot\left[v_1+\sum_{i=2}^n\frac{a_i}{a_1}\cdot\left(\frac{\lambda_i}{\lambda_1}\right)^k\cdot v_i\right]}{\left\|a_1\cdot\lambda_1^k\right\|\cdot\left\|\left[v_1+\sum_{i=2}^n\frac{a_i}{a_1}\cdot\left(\frac{\lambda_i}{\lambda_1}\right)^k\cdot v_i\right]\right\|}$$
$$=\operatorname{sign}\left(a_1\cdot\lambda_1^k\right)\cdot\frac{\left[v_1+\sum_{i=2}^n\frac{a_i}{a_1}\cdot\left(\frac{\lambda_i}{\lambda_1}\right)^k\cdot v_i\right]}{\left\|\left[v_1+\sum_{i=2}^n\frac{a_i}{a_1}\cdot\left(\frac{\lambda_i}{\lambda_1}\right)^k\cdot v_i\right]\right\|}$$

עתה נשים לב כי $\frac{\lambda_i}{\lambda_1}<1$ לכל $2\le i\le n$ מהנתון $\lambda_1>\lambda_2\ge\cdots\ge\lambda_n$, ולכן $\left(\frac{\lambda_i}{\lambda_1}\right)^k\overset{k\to\infty}{\to}0$, נציב ונקבל מאש"ג כי

$$\lim_{k\to\infty}\left[\sum_{i=2}^n\frac{a_i}{a_1}\cdot\left(\frac{\lambda_i}{\lambda_1}\right)^k\cdot v_i\right]=\sum_{i=2}^n\frac{a_i}{a_1}\cdot 0\cdot v_i=0_{\mathbb{R}^m}$$
$$\lim_{k\to\infty}\left[v_1+\sum_{i=2}^n\frac{a_i}{a_1}\cdot\left(\frac{\lambda_i}{\lambda_1}\right)^k\cdot v_i\right]=v_1+0_{\mathbb{R}^m}=v_1$$

ולכן נקבל כי

$$\lim_{k\to\infty}b_{k+1}=\lim_{k\to\infty}\left[\operatorname{sign}\left(a_1\cdot\lambda_1^k\right)\cdot\frac{\left[v_1+\sum_{i=2}^n\frac{a_i}{a_1}\cdot\left(\frac{\lambda_i}{\lambda_1}\right)^k\cdot v_i\right]}{\left\|\left[v_1+\sum_{i=2}^n\frac{a_i}{a_1}\cdot\left(\frac{\lambda_i}{\lambda_1}\right)^k\cdot v_i\right]\right\|}\right]$$
$$=\operatorname{sign}\left(a_1\cdot\lambda_1^k\right)\cdot\lim_{k\to\infty}\left[\frac{\left[v_1+\sum_{i=2}^n\frac{a_i}{a_1}\cdot\left(\frac{\lambda_i}{\lambda_1}\right)^k\cdot v_i\right]}{\left\|\left[v_1+\sum_{i=2}^n\frac{a_i}{a_1}\cdot\left(\frac{\lambda_i}{\lambda_1}\right)^k\cdot v_i\right]\right\|}\right]$$
$$=\operatorname{sign}\left(a_1\cdot\lambda_1^k\right)\frac{\lim_{k\to\infty}\left[v_1+\sum_{i=2}^n\frac{a_i}{a_1}\cdot\left(\frac{\lambda_i}{\lambda_1}\right)^k\cdot v_i\right]}{\left\|\lim_{k\to\infty}\left[v_1+\sum_{i=2}^n\frac{a_i}{a_1}\cdot\left(\frac{\lambda_i}{\lambda_1}\right)^k\cdot v_i\right]\right\|}$$
$$=\operatorname{sign}\left(a_1\cdot\lambda_1^k\right)\cdot\frac{v_1}{\|v_1\|}=\operatorname{sign}\left(a_1\cdot\lambda_1^k\right)\cdot\frac{v_1}{\|1\|}=\pm v_1$$

כלומר הראנו שהאלגוריתם מתכנס ל- $\pm v_1$, כנדרש

מ.ש.ל.☺

8. **צ"ל:** היאקוביאן של $f(\sigma)$
הוכחה:

תחילה נסמן את העמודות של U ב־ u_i, ונסמן $\sigma = (\sigma_1, \dots, \sigma_n)$ וננסה לפשט את הביטוי:

$$f(\sigma) = U \cdot \mathrm{diag}(\sigma) \cdot U^T \cdot x = U \cdot \mathrm{diag}(\sigma) \cdot \begin{bmatrix} \langle u_1, x\rangle \\ \vdots \\ \langle u_n, x\rangle \end{bmatrix}$$
$$= U \cdot \begin{bmatrix} \sigma_1 \cdot \langle u_1, x\rangle \\ \vdots \\ \sigma_n \cdot \langle u_n, x\rangle \end{bmatrix} = \sum_{i=1}^{n} \sigma_i \cdot u_i \cdot \langle u_i, x\rangle$$

עתה נרצה לחשב את

$$\frac{\partial f_i}{\partial \sigma_j} = \frac{\partial}{\partial \sigma_j}\left[\sum_{k=1}^{n} \sigma_k \cdot u_k \cdot \langle u_k, x\rangle\right]_i = \sum_{k=1}^{n} [u_k \cdot \langle u_k, x\rangle]_i \cdot \frac{\partial \sigma_k}{\partial \sigma_j}$$
$$= \sum_{k=1}^{n} [u_k \cdot \langle u_k, x\rangle]_i \cdot \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases} = [u_j \cdot \langle u_j, x\rangle]_i$$

כלומר קיבלנו שבשורה ה־ i יש לנו את הוקטור $u_j \cdot \langle u_j, x\rangle$, שבצורה מטריציות אפשר לרשום:

$$J_\sigma(f) = U \cdot \begin{bmatrix} \langle u_1, x\rangle & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & \langle u_n, x\rangle \end{bmatrix} = U \cdot \mathrm{diag}\left(U^T \cdot x\right)$$

כלומר קיבלנו כי $\boxed{J_\sigma(f) = U \cdot \mathrm{diag}\left(U^T \cdot x\right)}$

מ.ש.ל.☺

9. **צ"ל:** הגרדיאנט של $h(\sigma) = \frac{1}{2}\|f(\sigma) - y\|^2$

הוכחה:

נשתמש בחוק השרשרת ובכך שבתרגול ראינו כי הגרדיאנט של $\|x\|^2 = 2x$ ונקבל כי

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

לכן נקבל כי

$$\boxed{\nabla h(\sigma) = J_\sigma(h)^T = J_\sigma(f)^T \cdot (f(\sigma) - y)}$$

מ.ש.ל.☺

10. **צ"ל:** היאקוביאן של $softmax$

הוכחה:

נסמן ב־ $g_i = e^{a_i}, h = \sum_{i=1}^{n} e^{a_i}$, לכן $f_i = \frac{g_i}{h}$,
נחלק ל־ 2 מקרים: אם $i \neq j$ אז

$$\frac{\partial f_i}{\partial a_j} = \frac{\partial}{\partial a_j} \cdot \frac{g_i}{h} = \frac{\partial}{\partial a_j} \cdot \left[\frac{e^{a_i}}{\sum_{i=1}^{n} e^{a_i}}\right] = \frac{0 \cdot \left[\sum_{i=1}^{n} e^{a_i}\right] - e^{a_i} \cdot e^{a_j}}{\left[\sum_{i=1}^{n} e^{a_i}\right]^2} = \frac{-e^{a_i} \cdot e^{a_j}}{\left[\sum_{i=1}^{n} e^{a_i}\right]^2}$$
$$= -\frac{e^{a_i}}{\sum_{i=1}^{n} e^{a_i}} \cdot \frac{e^{a_j}}{\sum_{j=1}^{n} e^{a_j}} = -f_i \cdot f_j$$

עכשיו אם $i=j$:

$$\frac{\partial f_i}{\partial a_j}=\frac{\partial}{\partial a_j}\cdot\frac{g_i}{h}=\frac{\partial}{\partial a_j}\cdot\left[\frac{e^{a_i}}{\sum_{i=1}^{n}e^{a_i}}\right]=\frac{e^{a_i}\cdot\left[\sum_{i=1}^{n}e^{a_i}\right]-e^{a_i}\cdot e^{a_i}}{\left[\sum_{i=1}^{n}e^{a_i}\right]^2}=\frac{e^{a_i}\cdot\left[\sum_{i=1}^{n}e^{a_i}-e^{a_i}\right]}{\left[\sum_{i=1}^{n}e^{a_i}\right]^2}$$
$$=\frac{e^{a_i}}{\sum_{i=1}^{n}e^{a_i}}\cdot\frac{\sum_{i=1}^{n}e^{a_i}-e^{a_i}}{\sum_{i=1}^{n}e^{a_i}}=f_i\cdot\left[1-\frac{e^{a_i}}{\sum_{i=1}^{n}e^{a_i}}\right]=f_i\cdot(1-f_i)$$

כלומר קיבלנו כי

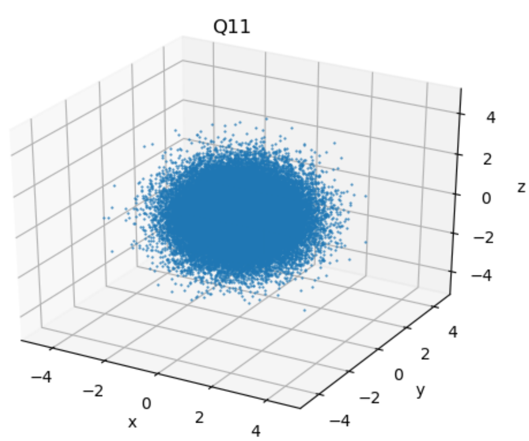
$$\boxed{\frac{\partial f_i}{\partial a_j}=\begin{cases}f_i\cdot(1-f_i) & i=j\\ -f_i\cdot f_j & i\neq j\end{cases}}$$

אם נגדיר $\delta_{i,j}=\begin{cases}1 & i=j\\ 0 & i\neq j\end{cases}$, נוכל לכתוב את הנוסחא באופן הבא:

$$\boxed{\frac{\partial f_i}{\partial a_j}=f_i\cdot(\delta_{i,j}-f_j)}$$

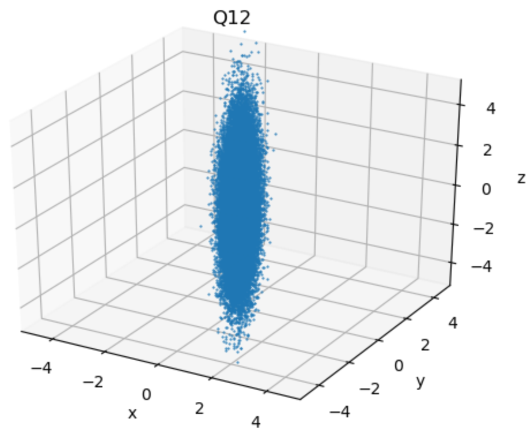
מ.ש.ל.☺

11. **פתרון:**



מ.ש.ל.☺

12. **פתרון:**

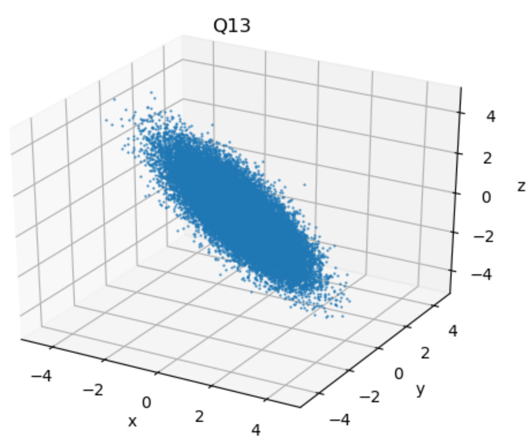


היינו מצפים שתצא המטריצה $\begin{bmatrix} 0.01 & 0 & 0 \\ 0 & 0.25 & 0 \\ 0 & 0 & 4 \end{bmatrix}$ בגלל שהגדלנו ב־ $0.1, 0.5, 2$ ו$variance$ מגדיל את הסקלר בריבוע והמטריצה ההתחלתית הייתה I_3.

המטריצה שיצאה היא $\begin{bmatrix} 0.011001 & 0.001 & 0.005 \\ 0.001 & 0.249 & 0.07 \\ 0.005 & 0.07 & 4.019 \end{bmatrix} \sim \begin{bmatrix} 0.01 & 0 & 0 \\ 0 & 0.25 & 0 \\ 0 & 0 & 4 \end{bmatrix}$, שזה ממש דומה למטריצה המקורית עד כדי שגיאה קטנה במחשב.

מ.ש.ל.☺

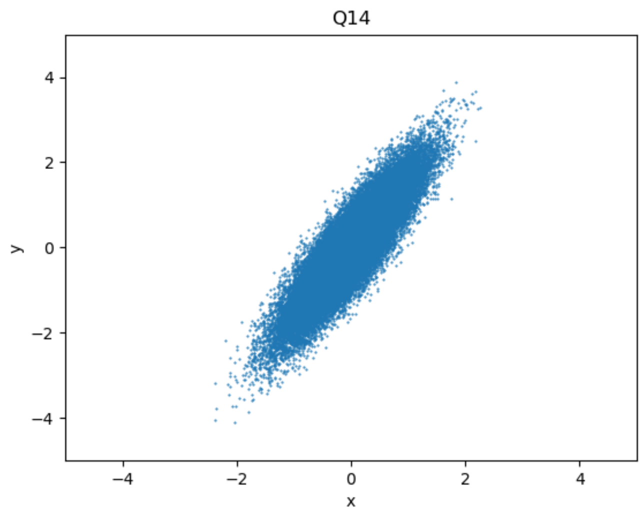
13. **פתרון:**



המטריצה שיצאה היא $\begin{bmatrix} 0.86 & -1.42 & 0.02 \\ -1.42 & 3.3 & 016 \\ 0.02 & 0.16 & 0.05 \end{bmatrix}$, שלא מכילה הרבה משמעות לדעתי, אבל זה פשוט סיבוב של המידע מהמטריצה הקודמת, אז אני מנחש ש־ $\det$ לדוגמא נשמר.

מ.ש.ל.☺

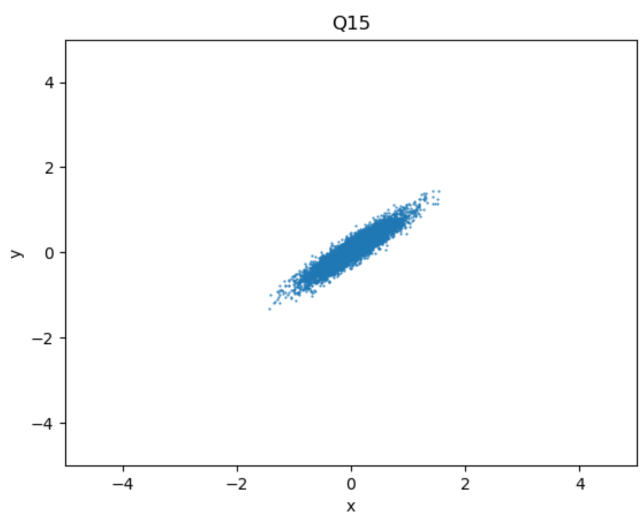
14. **פתרון:**



ההטלה נראת כמו גאוסיאן לאחר סקייל וסיבוב, זה הגיוני כי ראינו שהטלה של גאוסיאן היא גאוסיאן. ואם היינו עושים הטלה למטריצה המקורית היינו מקבלים בדיוק גאוסיאן.

מ.ש.ל.☺

15. **פתרון:**

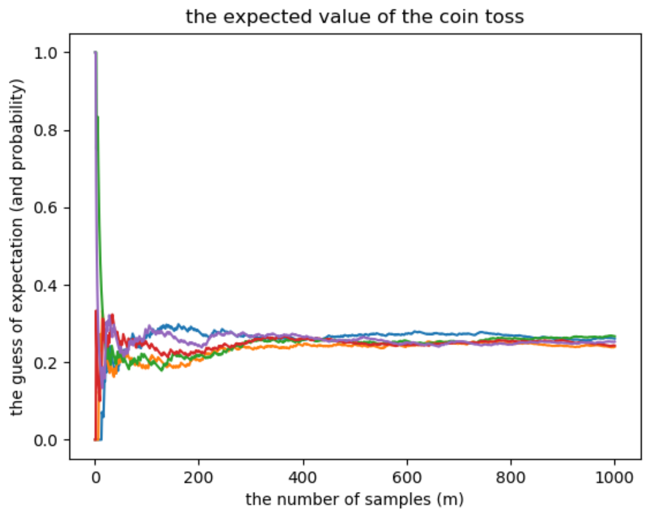


ההטלה נראת כמו גאוסיאן לאחר סקייל וסיבוב, זה הגיוני כי ראינו שהטלה של גאוסיאן בהתנייה היא גאוסיאן. ואם היינו עושים הטלה למטריצה המקורית היינו מקבלים בדיוק גאוסיאן.

מ.ש.ל.☺

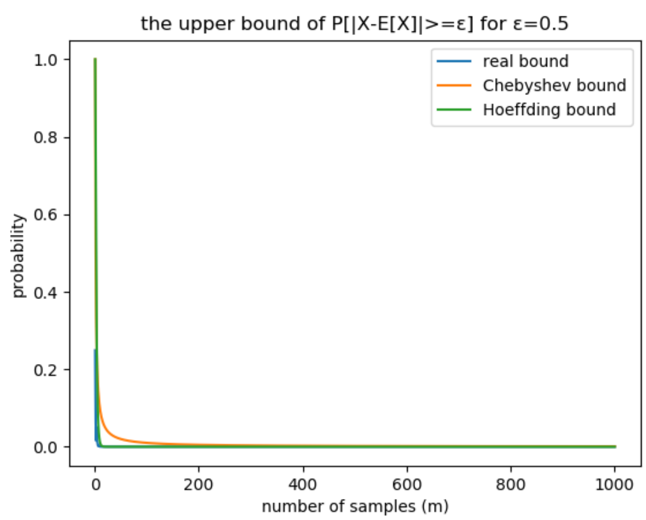
16. **פתרון:**

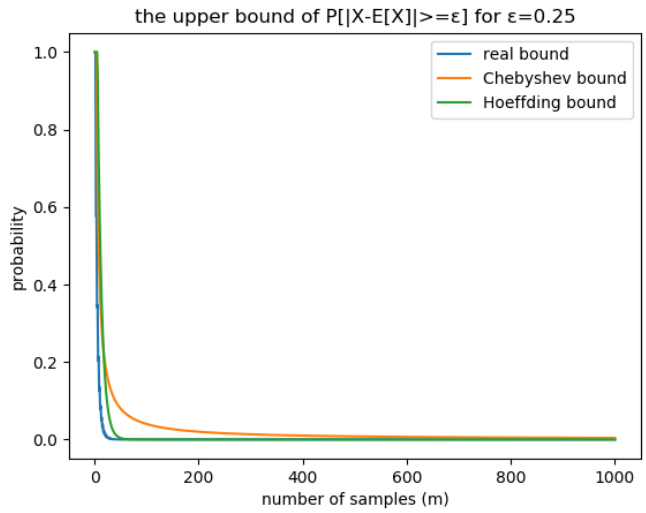
(א) **פתרון:**
אני מצפה שבאינסוף הערך יהיה 0.25 בגלל שעבור משתנה ברנולי מתקיים $\mathbb{E}\left[X\right] = p$ ובקוד שלנו מתקיים $p = 0.25$, ובאינסוף אנחנו יודעים שמחוק המספרים הגדולים, אנחנו נתקרב יותר ויותר לתוחלת.

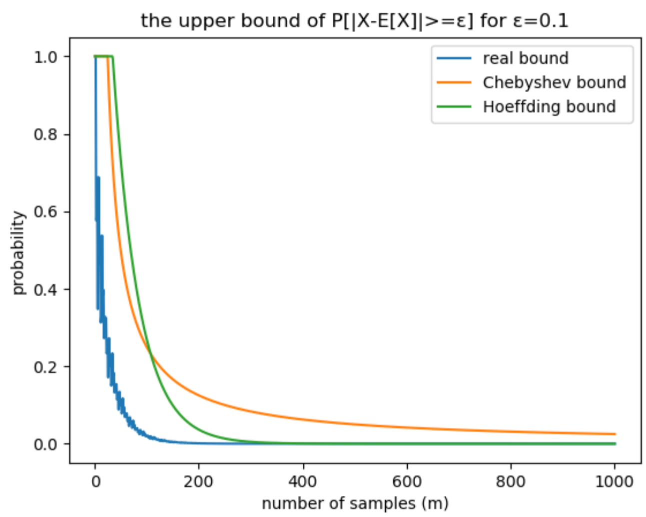


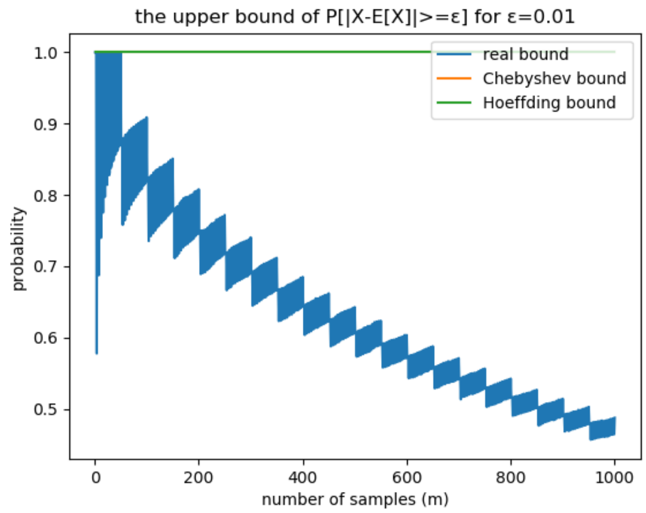
מ.ש.ל.א.☺

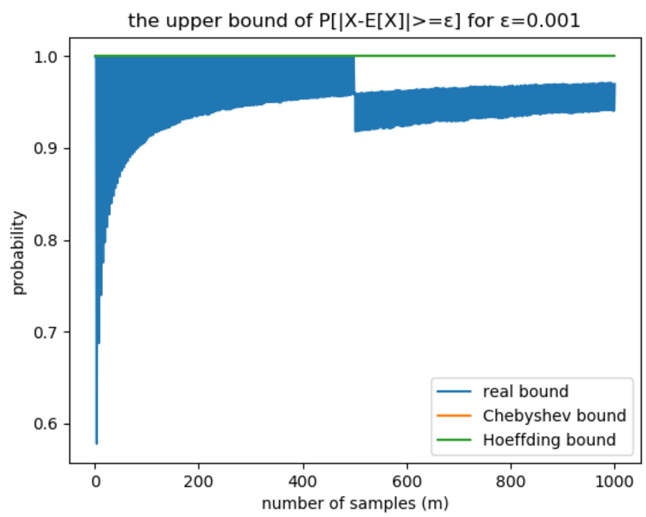
(ב) **פתרון:**
אני מצפה לראות שהחסם האמיתי תמיד יהיה מתחת לחסמים של צ'בישב והופדינג ונראה יותר שינויים ככל שהאפסילון קטן יותר (נראה שהופדינג יותר חזק)




the upper bound of P[|X-E[X]|>=ε] for ε=0.25
real bound
Chebyshev bound
Hoeffding bound
probability
number of samples (m)
0.0
0.2
0.4
0.6
0.8
1.0
0
200
400
600
800
1000


the upper bound of P[|X-E[X]|>=ε] for ε=0.1
real bound
Chebyshev bound
Hoeffding bound
probability
number of samples (m)
0.0
0.2
0.4
0.6
0.8
1.0
0
200
400
600
800
1000


the upper bound of P[|X-E[X]|>=ε] for ε=0.01
real bound
Chebyshev bound
Hoeffding bound
probability
number of samples (m)
1.0
0.9
0.8
0.7
0.6
0.5
0
200
400
600
800
1000


the upper bound of P[|X-E[X]|>=ε] for ε=0.001
real bound
Chebyshev bound
Hoeffding bound
probability
number of samples (m)
1.0
0.9
0.8
0.7
0.6
0
200
400
600
800
1000

מ.ש.ל.ב.☺