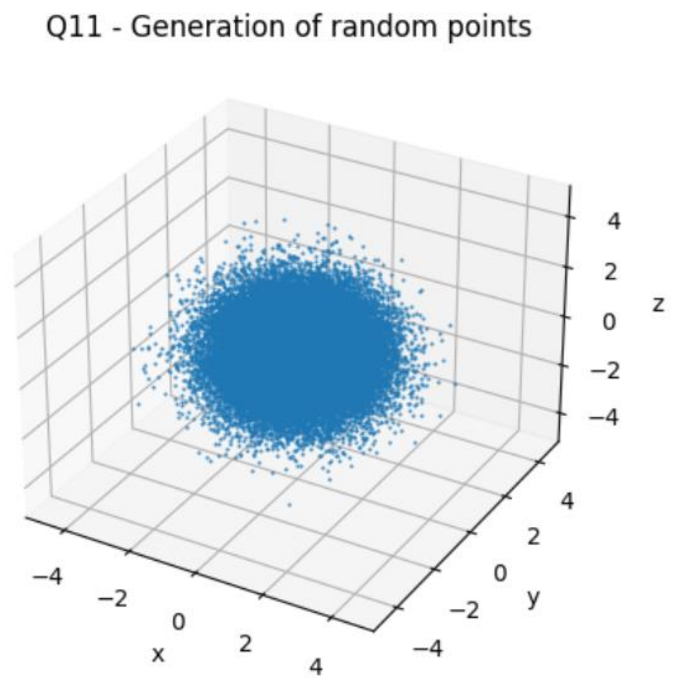
316163260

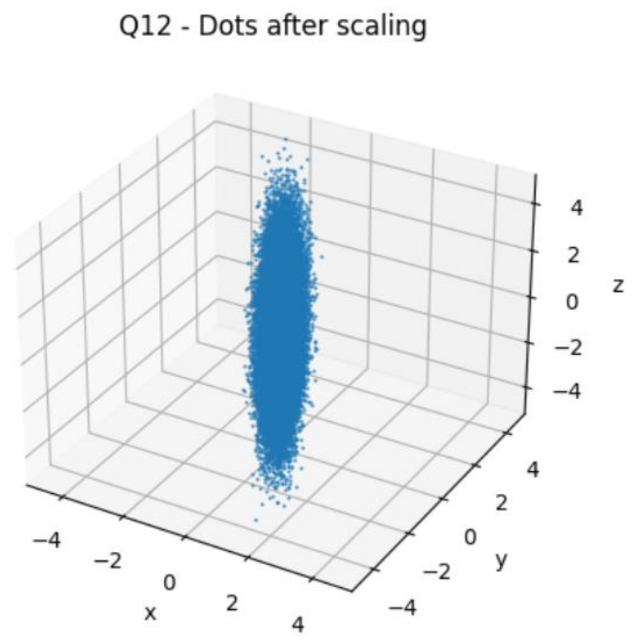
נועה בן דרור

IML – תרגיל 1 - חלק מעשי

(11



(12



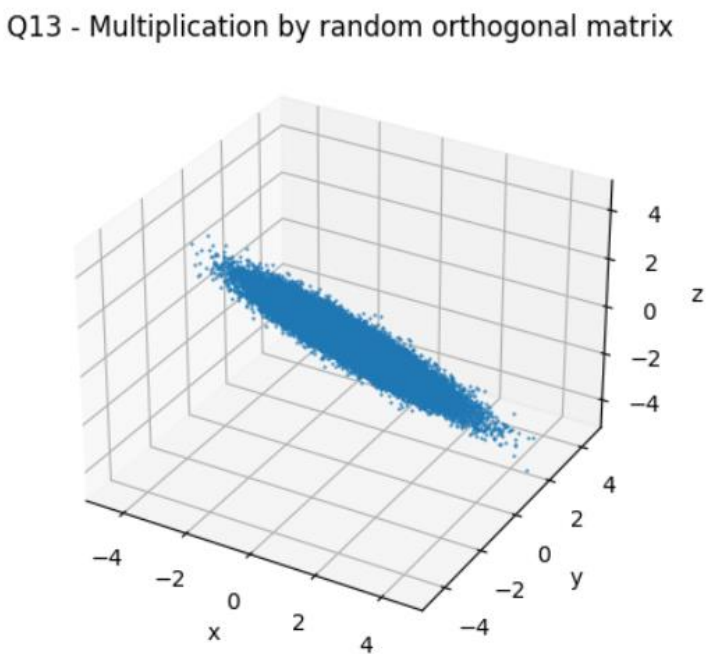
אנליטית, מטריצת השונות המשותפת נראית כך: (SCS^T)

0.010000000000000002	0.0	0.0
0.0	0.25	0.0
0.0	0.0	4.0

נומרית, מטריצת השונות המשותפת נראית כך:

0.01000	-0.00014	-0.00011
-0.00014	0.25106	0.00459
-0.00011	0.00459	3.99784

(13



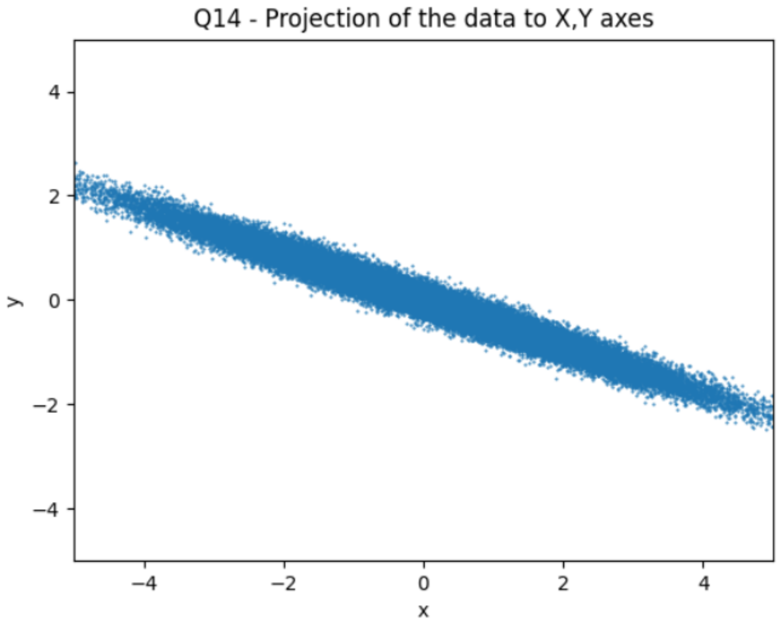
אנליטית, מטריצת השונות המשותפת נראית כך: (SCS^T)

3.29676	-1.45973	0.38232
-1.45973	0.67229	-0.22730
0.38232	-0.22730	0.29095

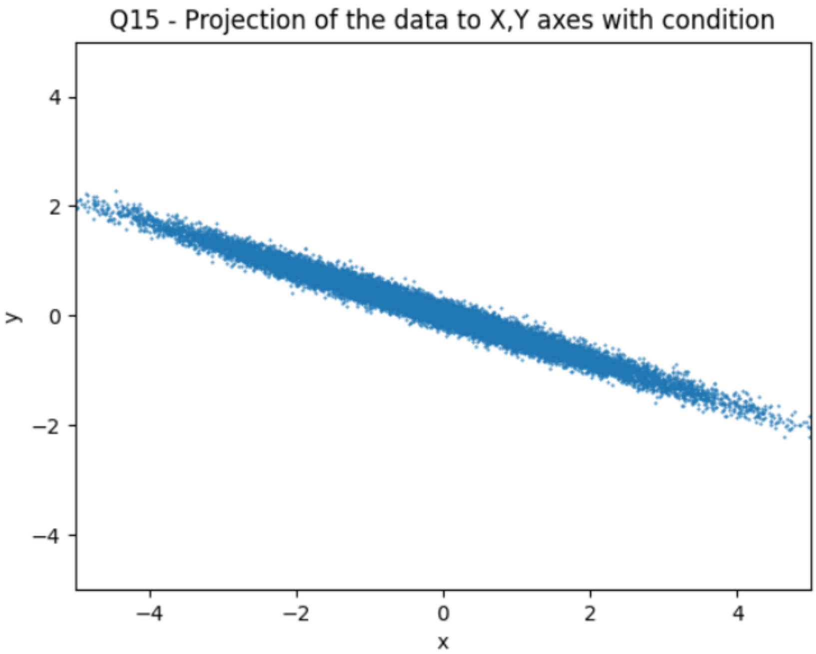
נומרית, מטריצת השונות המשותפת נראית כך:

3.29657	-1.45873	0.37799
-1.45873	0.67142	-0.22535
0.37799	-0.22535	0.29092

(14

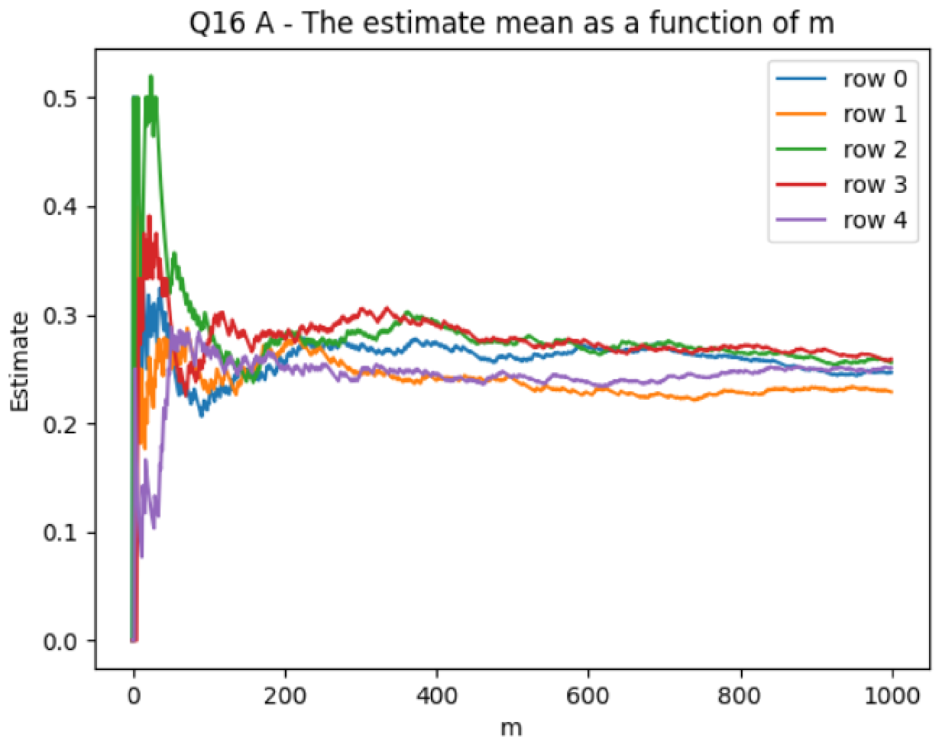


(15



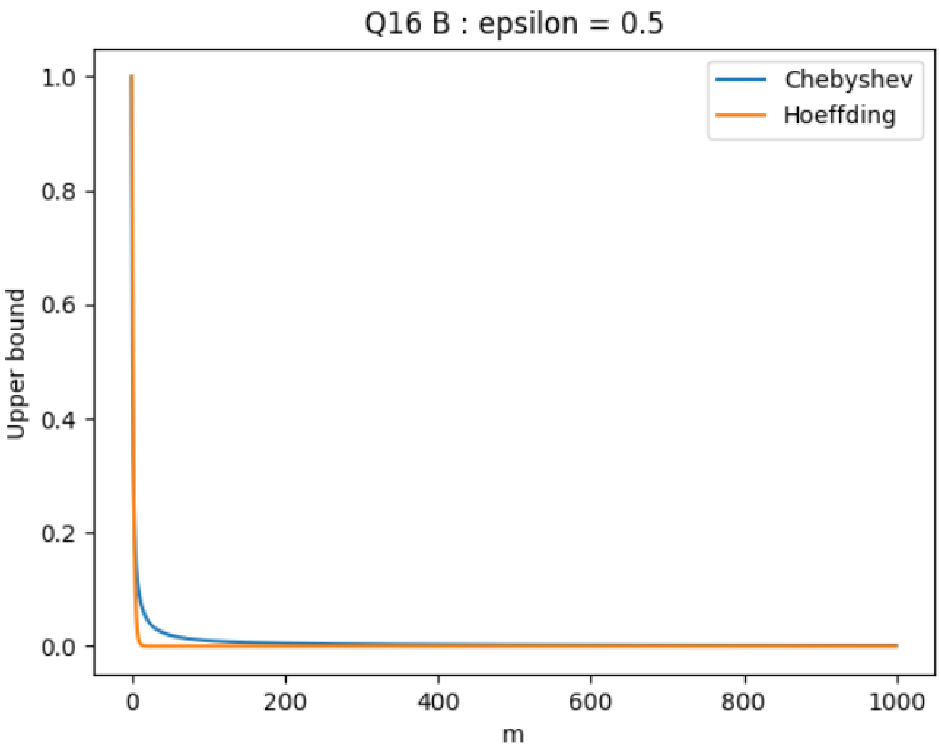
16) א.

לפי חוק המספרים הגדולים, ככל ש- m גדל נצפה לראות שהערכים מתכנסים לתוחלת. במקרה שלנו, התוחלת היא 0.25, ואכן בגרף ניתן לראות כי הערכים מתכנסים פחות או יותר לערך הזה.

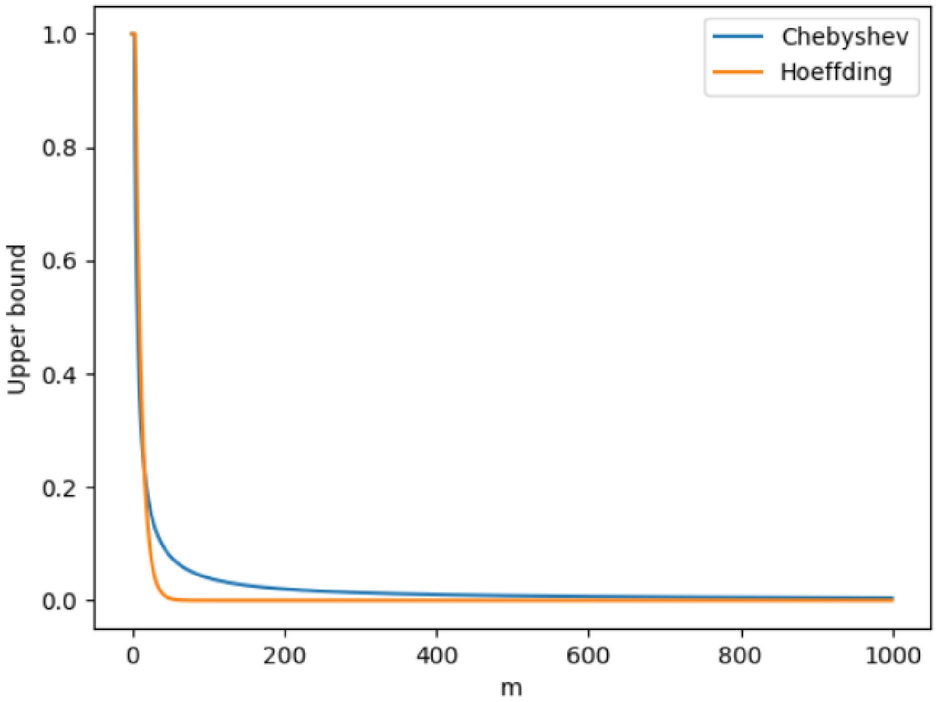


ב.

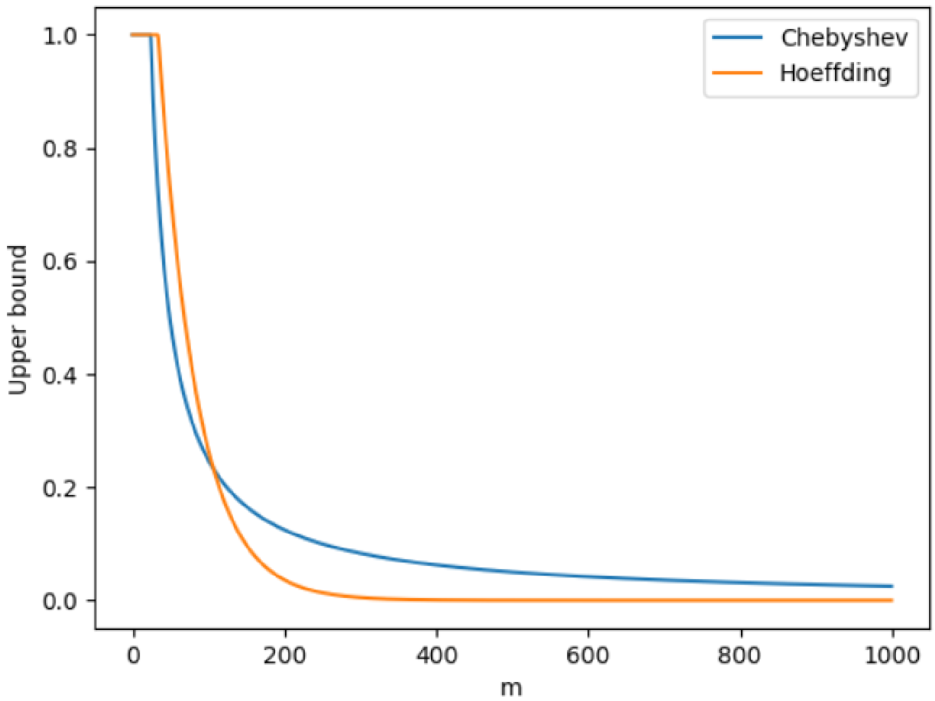
נציג את החסם העליון שצ'בישב והופדינג מספקים כתלות במספר הזריקות:

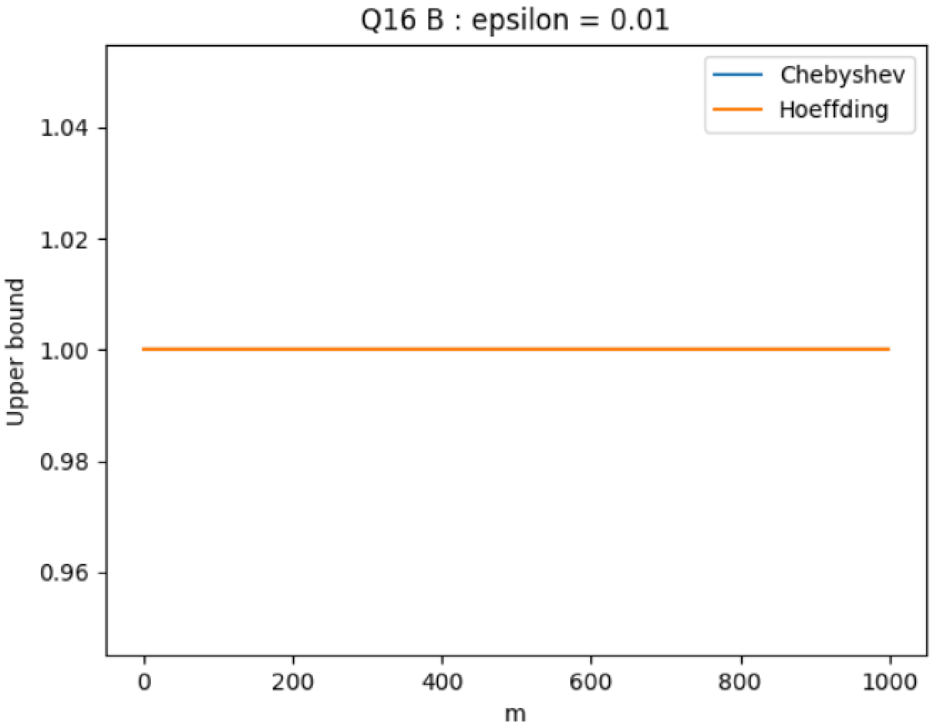


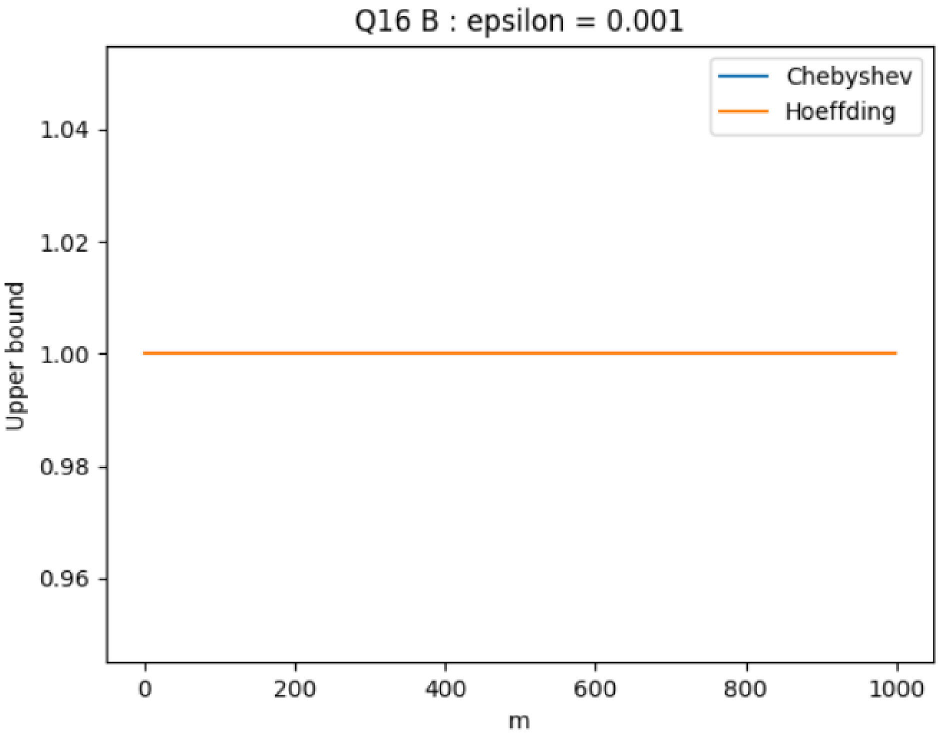
Q16 B : epsilon = 0.25



Q16 B : epsilon = 0.1

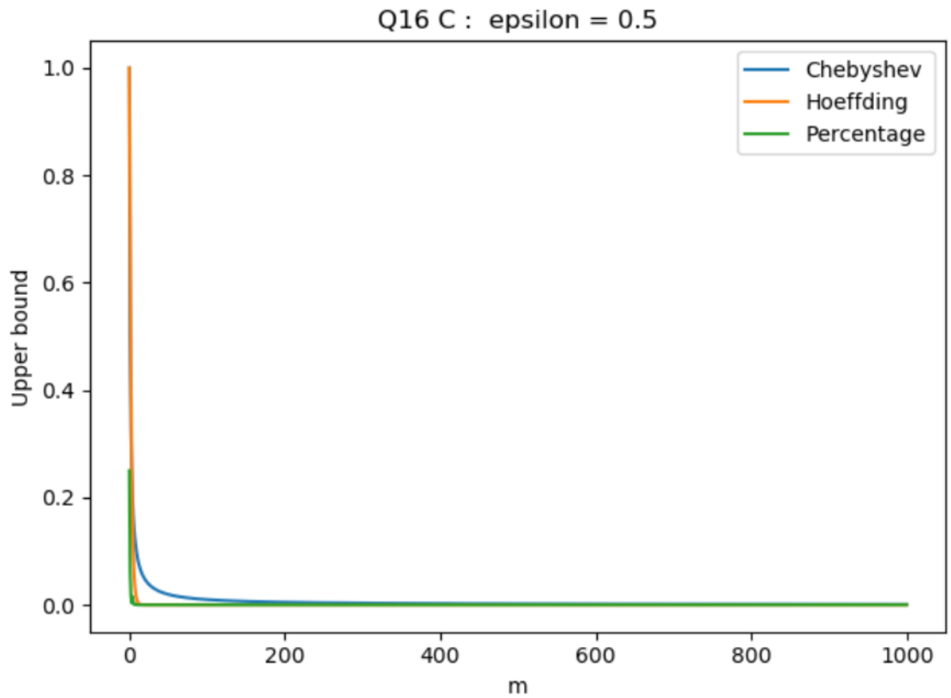


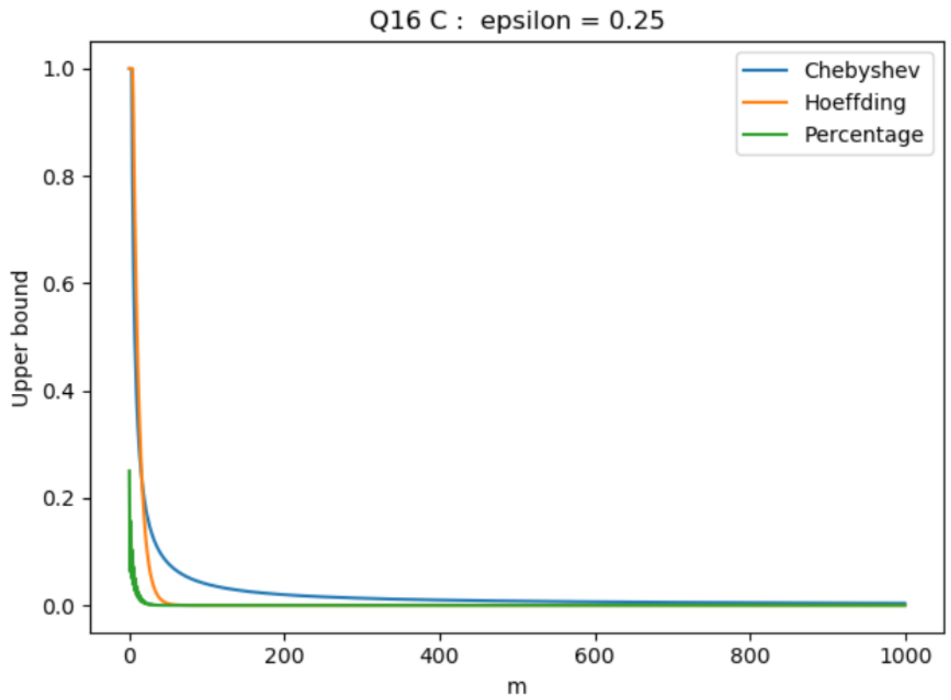

Q16 B : epsilon = 0.01
Chebyshev
Hoeffding
Upper bound
1.04
1.02
1.00
0.98
0.96
0
200
400
600
800
1000
m

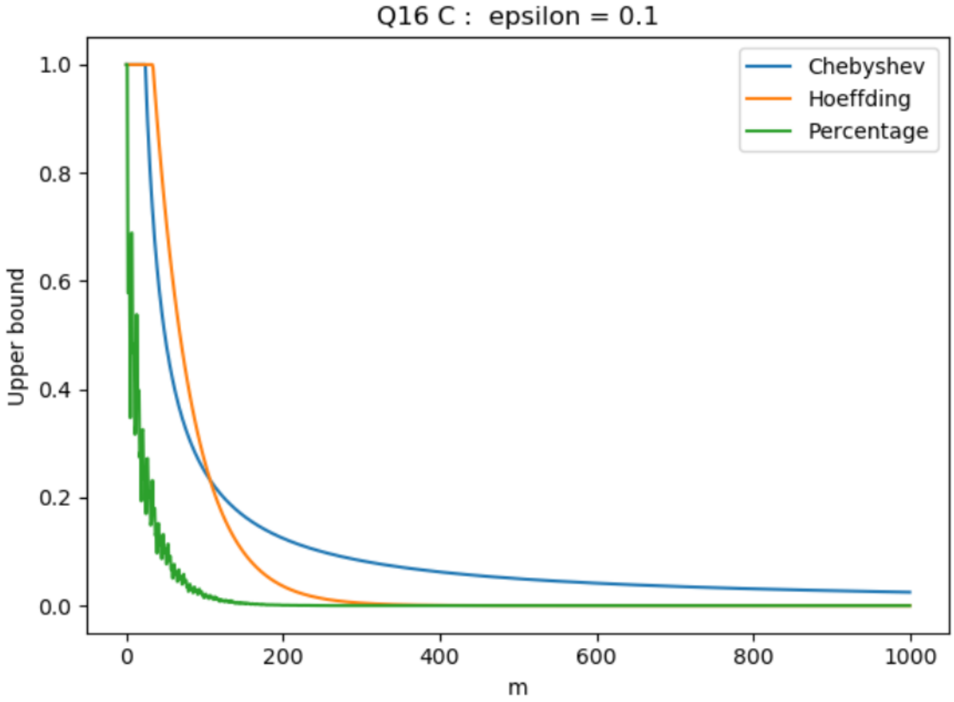

Q16 B : epsilon = 0.001
Chebyshev
Hoeffding
Upper bound
1.04
1.02
1.00
0.98
0.96
0
200
400
600
800
1000
m

ג.

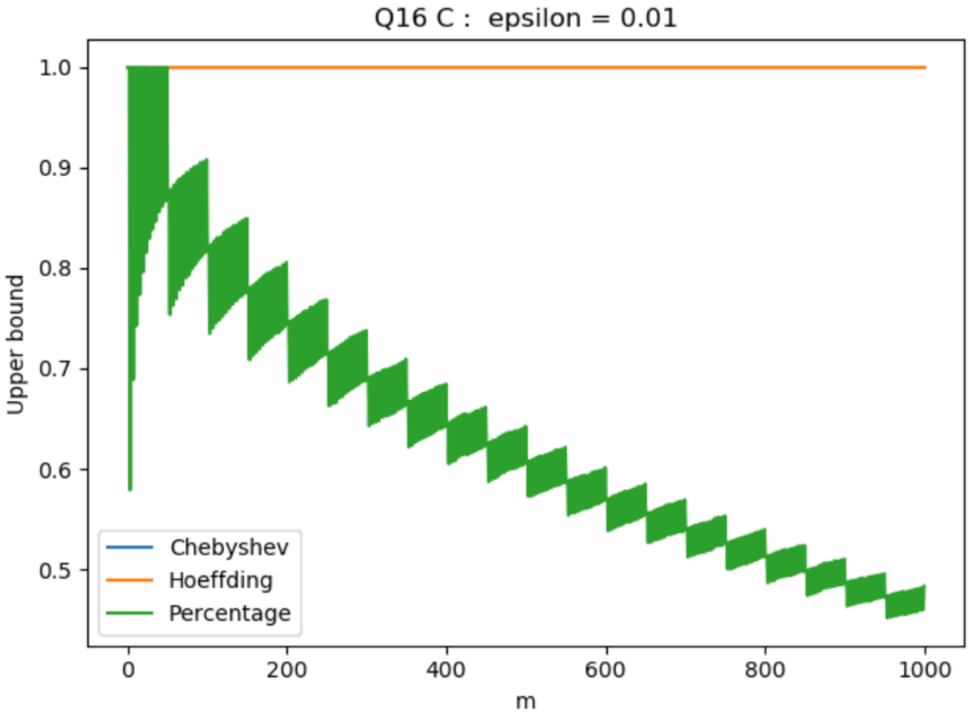
לפי חוק המספרים הגדולים נצפה לראות כי ככל ש- m יגדל, אחוז הרצפים שיספקו את התנאי יקטן. זה תלוי גם באפסילון – ככל שהוא יהיה יותר קטן, כך יהיה יותר קל לספק את התנאי. לכן, אם אפסילון יהיה גדול נצפה לראות אחוז נמוך של רצפים שמספקים את התנאי, ואם אפסילון יהיה קטן נצפה לראות אחוז גבוה יותר של רצפים שמספקים את התנאי.



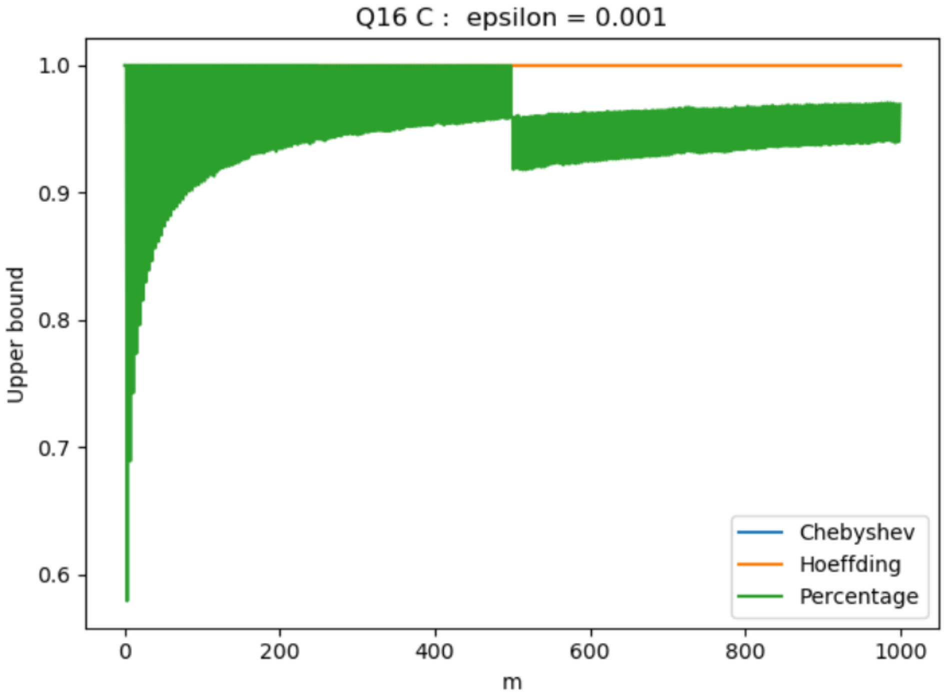




Q16 C : epsilon = 0.1
Chebyshev
Hoeffding
Percentage
Upper bound
m



Q16 C : epsilon = 0.01
Chebyshev
Hoeffding
Percentage
Upper bound
m


Q16 C : epsilon = 0.001
Upper bound
1.0
0.9
0.8
0.7
0.6
0
200
400
600
800
1000
m
Chebyshev
Hoeffding
Percentage

IML - תרגיל 1

נועה בן דרור
316163260

(1) הצלה ש $\underbrace{\begin{bmatrix}1\\2\\3\\4\end{bmatrix}}_{v}$ ו $\underbrace{\begin{bmatrix}0\\-1\\1\\2\end{bmatrix}}_{w}$: לפי התרגול, מתקיים:

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

$$\begin{bmatrix}0\\-\frac{3}{2}\\\frac{3}{2}\\3\end{bmatrix}$$

(2) הצלה ש $\underbrace{\begin{bmatrix}1\\2\\3\\4\end{bmatrix}}_{v}$ ו $\underbrace{\begin{bmatrix}1\\0\\1\\-1\end{bmatrix}}_{w}$. לפי התרגול, מתקיים:

$$p=\frac{\langle v,w\rangle}{\|w\|^2}\cdot w=\frac{1\cdot 1+2\cdot 0+3\cdot 1+4\cdot(-1)}{1^2+0^2+1^2+(-1)^2}w=\frac{0}{3}w=\begin{bmatrix}0\\0\\0\\0\end{bmatrix}$$

(3)

$$V^T w = 0$$

(כי $V^T w = \langle v|w\rangle$) $\Updownarrow$

$$\langle v|w\rangle = 0$$

(כי $\langle v|w\rangle = \|v\|\cdot\|w\|\cdot\cos\theta$) $\Updownarrow$

$$\|v\|\cdot\|w\|\cos\theta = 0$$

(כי $v,w\neq 0$ אז $\|v\|\neq 0$, $\|w\|\neq 0$) $\Updownarrow$

$$\cos\theta = 0$$

$$\Updownarrow$$

$$\theta = \pm 90$$

כנדרש.

(4) $T: V\to W$ ה"ל, A המטריצה המייצגת של ההעתקה.

יהי $x\in V$. אזי:

$$\|Ax\|^2 = \langle Ax|Ax\rangle = (Ax)^T Ax = x^T A^T A x = x^T I x = x^T x = \langle x|x\rangle = \|x\|^2$$

A אורתוגונלית

לכן $A^TA = I$

כלומר קיבלנו $\|Ax\|^2 = \|x\|^2$ ולכן בהכרח $\|Ax\| = \|x\|$

(כי לכל $x\in V$ $\|x\|\geq 0$).

(5) מתקיים:

$$A^{-1} = (U\Sigma V^T)^{-1} = (V^T)^{-1}\Sigma^{-1}U^{-1} = (V^{-1})^{-1}\Sigma^{-1}U^T =$$

(U,V אורתוגונליות ולכן $V^T = V^{-1}$, $U^T = U^{-1}$)

$$V\Sigma^{-1}U^T$$

ה- SVD של מטריצה עוזר למצוא את המטריצה ההפכית בזמן ריצה קצר יותר.
תהליך ידוע לצמצום של מטריצה, ומציאת ההפכית שלה, הוא $O(n^3)$.
לעומת זאת, בעזרת SVD נקבל זמן ריצה $O(n^2)$.

$$A = U\Sigma V^T = \begin{bmatrix} 5 & 5 \\ -1 & 7 \end{bmatrix}$$ 6

U,V אורתוגונליות, Σ אלכסונית.

נחשב $A^T \cdot A$:

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

כעת, נשים לב כי:

$$A^T A = (U\Sigma V^T)^T (U\Sigma V^T) = V\Sigma^T \underbrace{U^T U}_{} \Sigma V^T = V\Sigma^T \Sigma V^T = V\Sigma^2 V^T$$

U אורתוגונלית לכן $U^T U = I$

Σ אלכסונית, לכן $\Sigma^T = \Sigma$

אז קיבלנו

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

נשים לב שזה פירוק ספקטרלי, ולכן נבצע לכסון כדי למצוא את V ואת Σ^2.

לאחר הרצה בפייתון, קיבלנו:

$$\Sigma = \begin{bmatrix}\sqrt{80} & 0\\ 0 & \sqrt{20}\end{bmatrix} \quad \text{ולכן} \quad \Sigma^2 = \begin{bmatrix}80 & 0\\ 0 & 20\end{bmatrix}$$

כמו כן, קיבלנו
$$V = \begin{bmatrix}0.31622777 & 0.9486833\\ 0.9486833 & -0.31622777\end{bmatrix} = \begin{bmatrix}\sqrt{\frac{1}{10}} & \sqrt{\frac{9}{10}}\\ \sqrt{\frac{9}{10}} & -\sqrt{\frac{1}{10}}\end{bmatrix} = V^T$$

כעת, נותר למצוא את U. ידוע כי $AV = U\Sigma$ ולכן $AV\Sigma^{-1} = U$. לאחר חישוב בפייתון קיבלנו:

$$U = \begin{bmatrix}0.707106 & 0.707106\\ 0.707106 & -0.707106\end{bmatrix} = \begin{bmatrix}\sqrt{\frac{1}{2}} & \sqrt{\frac{1}{2}}\\ \sqrt{\frac{1}{2}} & -\sqrt{\frac{1}{2}}\end{bmatrix}$$

(7) יהי b_0 וקטור [redacted] כך ש- $b_0 = \sum_{i=0}^{n} a_i v_i$ $(a_1 \neq 0)$.

ראשית נשים לב כי $C_0 = A^TA$ היא סימטרית משום ש-
$$C_0^T = (A^TA)^T = A^TA = C_0$$

לכן, C_0 צמודה לעצמה, וקיימות

$U, \Sigma \in \mathbb{R}^{n\times n}$ כך ש- U אורתוגונלית ו- Σ אלכסונית וכן מתקיים $C_0 = U\Sigma U^{-1}$.

מהנתונים, $\lambda_1, \dots, \lambda_n$ הם הערכים העצמיים של C_0 ולכן

$$D_{ij} = \begin{cases} \lambda_i & i=j \\ 0 & i \neq j \end{cases}$$

ניתן לחשוב על C_0 כאלכסונית (כי אפשר להשתמש בוקטורים העצמיים כבסיס)

ולכן נניח מעתה כי $C_0 = D$. נסתכל על התמונה של b_{k+1}:

$$C_0^k b_0 = \sum_{i=1}^{n} C_0^k (a_{0i} v_i) \underset{\text{אלכסונית}}{=} \sum_{i=1}^{n} a_{ki} \lambda_i^k v_i =$$

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

עבור $k \to \infty$ נקבל:

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

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

$$\frac{a_1 \lambda_1^k}{|a_1 \lambda_1^k|} = \pm 1$$

$\frac{\lambda_i}{\lambda_1} < 1$ כי $\lambda_1 \geq \lambda_i$

$$\pm \frac{\left(v_1 + \sum_{i=2}^{n} \frac{a_i}{a_1} 0 \cdot v_i\right)}{\left\|\left(v_1 + \sum_{i=2}^{n} \frac{a_i}{a_1} 0 \cdot v_i\right)\right\|} = \pm \frac{v_1}{\|v_1\|} = \pm v_1$$

$\|v_1\| = 1$

(8) נשים לב כי:

$$U = \begin{bmatrix} | & & | \\ u_1 & \cdots & u_n \\ | & & | \end{bmatrix}$$

$$U^T = \begin{bmatrix} - u_1 - \\ \vdots \\ - u_n - \end{bmatrix}$$

$$diag(\sigma) = \Sigma = \begin{bmatrix} \sigma_1 & & \\ & \ddots & \\ & & \sigma_n \end{bmatrix}$$

ולכן:

$$f(\sigma) = \begin{bmatrix} | & & | \\ u_1 & \cdots & u_n \\ | & & | \end{bmatrix} \begin{bmatrix} \sigma_1 & & \\ & \ddots & \\ & & \sigma_n \end{bmatrix} \begin{bmatrix} - u_1 - \\ \vdots \\ - u_n - \end{bmatrix} \cdot X =$$

$$\begin{bmatrix} | & & | \\ \sigma_1 u_1 & \cdots & \sigma_n u_n \\ | & & | \end{bmatrix} \begin{bmatrix} - u_1 - \\ \vdots \\ - u_n - \end{bmatrix} X = \sum_{i=1}^{n} \sigma_i u_i u_i^T \cdot X$$

מכאן מתקיים:

$$\frac{\partial f_j(\sigma_0)}{\partial \sigma_i} = [u_i u_i^T x]_j$$

(9) לפי כלל השרשרת מתקיים:

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

$$\left(f(\sigma)^T - y^T\right)\cdot J_{f(\sigma)}$$

(10) הפונקציה g מקיימת:

$$g\left(\begin{bmatrix} z_1 \\ \vdots \\ z_k \end{bmatrix}\right) = \begin{bmatrix} S_1 \\ \vdots \\ S_k \end{bmatrix}$$

כך ש- $S_i = \dfrac{e^{z_i}}{\sum_{n=1}^{k} e^{z_n}}$.

אנחנו צריכים לחשב את היעקוביאן של g, כלומר $[J_g]_{ij} = \dfrac{\partial S_i}{\partial z_j}$

אם כך:

$$\frac{\partial S_i}{\partial z_j} = \text{derive}\left(\frac{e^{z_i}}{\sum_{n=1}^{k} e^{z_n}}\right) \text{ w.r.t } z_j =$$

$$\sum_{n=1}^{k} e^{z_n} = h$$

$$\text{derive}\left(\frac{e^{z_i}}{h}\right) \text{ w.r.t } z_j$$

נחלק למקרים:

$i=j$:

$$\frac{\partial}{\partial z_j} \cdot \frac{e^{z_i}}{h} = \frac{e^{z_i} \cdot \sum_{n=1}^{k} e^{z_n} - e^{z_i} \cdot e^{z_j}}{\left(\sum_{n=1}^{k} e^{z_n}\right)^2} =$$

$$\frac{e^{z_i}}{\sum_{n=1}^{k} e^{z_n}} \cdot \frac{\sum_{n=1}^{k} e^{z_n} - e^{z_j}}{\sum_{n=1}^{k} e^{z_n}} = S_i - (1 - S_j)$$

$i \neq j$:

$$\frac{\partial}{\partial z_j} \cdot \frac{e^{z_i}}{h} = \frac{0 \cdot \sum_{n=1}^{k} e^{z_n} - e^{z_i} \cdot e^{z_j}}{\left(\sum_{n=1}^{k} e^{z_n}\right)^2} = \frac{-e^{z_i} \cdot e^{z_j}}{\left(\sum_{n=1}^{k} e^{z_n}\right)^2} =$$

$$-\left(\frac{e^{z_i}}{\sum_{n=1}^{k} e^{z_n}} \cdot \frac{e^{z_j}}{\sum_{n=1}^{k} e^{z_n}}\right)$$

לכן היעקוביאן ש g הוא:

$$[J_g]_{ij} = \frac{\partial S_i}{\partial z_j} = \begin{cases} S_i - (1 - S_j) & i = j \\ -(S_i + S_j) & i \neq j \end{cases}$$

Scanned with CamScanner