עמוד

ליני דראורין 313308140

IML - תרגיל 1

Warm-up - Algebra Recap -

3. נחשב את הטלת הוקטור $v=(1,2,3,4)$ על $w=(0,-1,1,2)$: $\frac{\langle v\cdot w\rangle}{|w|^2}\cdot w$:

$$|w|^2=0+1+1+4=6$$

$$\langle v\cdot w\rangle=[1\;2\;3\;4]\cdot\begin{bmatrix}0\\-1\\1\\2\end{bmatrix}=0-2+3+8=\boxed{9}$$

לכן וקטור ההטלה הוא: $\frac{6}{9}w=\frac{2}{3}w$ $\leftarrow$ $\begin{bmatrix}0\\-\frac{2}{3}\\\frac{2}{3}\\\frac{4}{3}\end{bmatrix}$

4. נחשב כעת את ההטלה של v על $w=(1,0,1,-1)$:

$$|w|^2=1+0+1+1=3$$

$$\langle v\cdot w\rangle=1\,2\,3\,4\cdot\begin{bmatrix}1\\0\\1\\-1\end{bmatrix}=1+0+3-4=0$$

לכן $\leftarrow$ $0\cdot w=\boxed{0}$ הוא וקטור ההטלה $\leftarrow$ $\begin{bmatrix}0\\0\\0\\0\end{bmatrix}$.

SVD -

7. הנוסחה ל-A היא: $VD^{-1}U^t$ - שימוש כי יותר מהיר בהינתן הפירוק ויותר חסכוני והסיבוכיות יותר טובה מהכפלה מטריצות 3 בצורה נאיבית.

8. עלינו למצוא את הפירוק SVD של המטריצה $UDV^T=C=\begin{pmatrix}5&5\\-1&7\end{pmatrix}$ עם הצעדים:

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

כעת, נחשב את הפירוק EVD של C^tC:

ראה/מצאנו את הערכים העצמיים של C^tC, עם המחשבון בעזרת פייתון וקיבלנו כי: $\lambda_1=20$, $\lambda_2=80$, על התרגול ניתן להדגים את $C^tC=V\underbrace{D^TD}V^t=VD^2V^t$ כפירוק EVD של C^tC

$\leftarrow D^t\cdot D=D^2$ ורק D את כסותינו

$$D^2=\begin{bmatrix}80&0\\0&20\end{bmatrix}\rightarrow D=\begin{bmatrix}\sqrt{80}&0\\0&\sqrt{20}\end{bmatrix}$$

V היא מטריצה אורתוגונלית של וקטורים עצמיים אשר ערכיה הם: $\begin{bmatrix} \frac{1}{3} & -3 \\ 1 & 1 \end{bmatrix}$.

מצאנו את D, V.

נדרש למצוא את U.

עפ"י הנוסחא $CV = UD$ לכן:

$$\overset{C}{\begin{bmatrix} 5 & 5 \\ -1 & 7 \end{bmatrix}} \cdot \overset{V}{\begin{bmatrix} \frac{1}{3} & -3 \\ 1 & 1 \end{bmatrix}} = \begin{bmatrix} U \end{bmatrix} \overset{D}{\begin{bmatrix} \sqrt{80} & 0 \\ 0 & \sqrt{20} \end{bmatrix}}$$

$$D^{-1} \cdot \backslash \quad \begin{bmatrix} \frac{20}{3} & -10 \\ \frac{20}{3} & 10 \end{bmatrix} = \begin{bmatrix} V \end{bmatrix} \cdot \begin{bmatrix} \sqrt{80} & 0 \\ 0 & \sqrt{20} \end{bmatrix}$$

$$= D \begin{bmatrix} 0.7453559925 & -2.2360679775 \\ 0.7453559925 & 2.2360679775 \end{bmatrix} = V$$

מכאן ש- V^t:

$$\begin{bmatrix} 0.745355\ldots & 0.745355 \\ -2.23606 & 2.23606\ldots \end{bmatrix}$$

מכיון שהגעתי לשקלו . את המטריצות שהתקבלו ∎

9. ראשית, שים לב כי המטריצה C_0 הינה סימטרית:

$$C_0^T=(A^TA)^T = A^T A^{TT} = A^T\cdot A = C_0$$

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

נוכיח באינדוקציה על $k\in\mathbb{N}$:

בסיס: $k=1$, מהגדרת b_k נקבל: $b_1=\frac{C_0^1 b_0}{\|C_0^1 b_0\|}$.

נניח נכונות הטענה על $L<k$ ונוכיח עבור $k+1$.

$$\underbrace{b_{k+1}=\frac{C_0 b_k}{\|C_0 b_k\|}}_{\text{נתון}} \underbrace{=}_{\text{הנחת האינדוקציה}} \frac{C_0}{\|C_0 b_k\|}\cdot\frac{C_0^k b_0}{\|C_0^k b_0\|} = \frac{C_0\cdot C_0^k b_0}{\frac{\|C_0\cdot C_0^k b_0\|}{\|C_0^k b_0\|}\|C_0^k b_0\|}$$

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

א. נתון: $\boxed{b_0=\sum_{i=0}^n a_i v_i}$

הערה: נניח כי הוקטורים עצמיים מנורמלים, דבר זה לא משנה (נרמול אותם $\frac{v_i}{\|v_i\|}\to v_i$ החדש) מכיוון שאם av עם $a\neq 0$ עדיין יהיה וקטור עצמי של המטריצה עם λ.

נסמן ב-λ מה שצריך להוכיח:

מאחר ו-C_0 הינה מטריצה סימטרית קיים פירוק evd: $\underline{V^T\begin{pmatrix}\lambda_1 & & \\ & \ddots & \\ & & \lambda_n\end{pmatrix}V}$

ולכן ⇐

$$C_0^k b_0 = V^T\begin{pmatrix}\lambda_1 & & \\ & \ddots & \\ & & \lambda_n\end{pmatrix}^k V b_0 =$$

$$V^T\begin{pmatrix}\lambda_1 & & \\ & \ddots & \\ & & \lambda_n\end{pmatrix}^k\left(\sum_{i=1}^n a_i V v_i\right) = V^T\begin{pmatrix}\lambda_1 & & \\ & \ddots & \\ & & \lambda_n\end{pmatrix}^k\left(\sum_{i=1}^n a_i e_i\right)$$

$$= V^T\left(\sum_{i=1}^n a_i \lambda_i^k e_i\right) = \lambda_1^k\left(a_1 v_1 + \sum_{i=2}^n a_i \frac{(\lambda_i)^k}{\lambda_1^k} v_i\right)$$

כי נתון מתקיים כאשר $i\neq 1$: $\lambda_i<\lambda_1$ ⇐ ולכן: $\lim_{k\to\infty}\left(\frac{\lambda_i}{\lambda_1}\right)^k=0$

$$\lim_{n\to\infty} b_n = \frac{a_1\lambda_1^k v_1}{\|a_1\lambda_1^k v_1\|} = \frac{a_1\lambda_1^k v_1}{a_1\lambda_1^k\|v_1\|} = v_1$$

כי הינו מנורמל.

Projection matrices

14. כדי להראות ש-P סימטרית עליי להראות כי $P_{ij}=P_{ji}$ לכל $1\le i,j\le d$ מאחר

ו-$P\in\mathbb{R}^{d\times d}$, $P=\sum_{i=1}^{k} v_i v_i^T$, עבור ס אשר $1\le i\le k$ מתקיים:

נסמן סימון עזר v_i:

הקואורדינטות: $(v_{im1}, v_{im2}, v_{im3} \dots v_{imd})$

$$v_i\cdot v_i^T=\begin{bmatrix} v_{im1}\cdot v_{im1} & \cdots & v_{im1}\cdot v_{imd} \\ \vdots & & \vdots \\ v_{imd}\cdot v_{im1} & \cdots & v_{imd}\cdot v_{imd}\end{bmatrix}$$

מטריצה סימטרית.

$v_{imi}\cdot v_{imj}=(v_i\cdot v_i^T)_{ij}$ לכל i,j כאשר נסכום k מטריצות כאשר כולן וקטור

מטריצה סימטרית מכיוון שאלו סכום של מטריצות סימטריות.

15. עבור $1\le i\le k$ מתקיים ש-

$$Av_i=\left(\sum_{i=1}^{k} v_i v_j^T\right)v_i=(v_1v_1^T)v_i+\dots(v_iv_i^T)v_i+\dots(v_kv_k^T)v_i=$$

$$v_1(v_i^T\cdot v_i)+\dots v_i(v_i^T\cdot v_i)+\dots v_k(v_k^T\cdot v_i)=$$

$$=0+\dots v_i+\dots+0=\boxed{v_i}$$

מכיוון שלכל $i\ne j$ מתקיים $v_j^T\cdot v_i=0$ ומתקיים $v_i^T\cdot v_i=1$ ומכיוון שמדובר בבסיס אורתונורמלי.

16. ראשית, מה שמצאנו בסעיף הקודם עוזר לנו להרחיב את הבסיס הנ"ל לבסיס אורתונורמלי מלא של $\mathbb{R}^d$ ונסמן את הקואורדינטות $v_{k+1},\dots v_d$ ומתקיים: $Pv_i=\left(\sum_{i=1}^{k} v_i v_j^T\right)v_i=0$

כי עבור $i>k$ מתקיים $v_j^T\cdot v_i=0$

$$\underbrace{\begin{bmatrix} v_1 \dots v_k & v_{k+1} \dots v_d \end{bmatrix}}_{k \quad d-k}\begin{bmatrix}1 & & & & & & \\ & 1 & & & & & \\ & & 1 & & & & \\ & & & 1 & & & \\ & & & & 0 & & \\ & & & & & 0 & \\ & & & & & & 0\end{bmatrix}\begin{bmatrix} — v_1 — \\ — v_2 — \\ \vdots \\ — v_k — \\ \vdots \\ — v_d — \end{bmatrix}=UDU^T$$

17. עלינו להוכיח כי: $P^2 = P^T P = P\cdot P^T = P$

כבר הוכחנו כי מכיוון שהמטריצה סימטרית אז מתקיים $P^T = P$

לכן נקבל כי: $P^2 = P\cdot P = P^T\cdot P = P\cdot P^T$

כעת נוכיח כי $P^2 = P$.

$$P^2 = \left(\sum_{i=1}^{k} v_i v_i^T\right)\left(\sum_{i=1}^{k} v_i v_i^T\right) = \left(v_1 v_1^T + \dots + v_k v_k^T\right)\left(v_1 v_1^T + \dots + v_k v_k^T\right)$$

$$= \sum_{i=1}^{k}\left(v_1 v_1^T \cdot v_1 v_1^T + v_i v_i^T \cdot v_i v_i^T + \dots + v_i v_i^T v_k v_k^T\right) =$$

$$\sum_{i=1}^{k} v_i \underline{v_i^T \cdot v_i} v_i^T = \sum_{i=1}^{k} v_i v_i^T = \boxed{P}$$

מכיוון שמדובר בבסיס אורתונורמלי: $\boxed{\text{לכל } i\neq j \quad v_i^T \cdot v_j = 0}$

$\boxed{\text{וגם } v_i^T \cdot v_i = 1}$

ב. $x \in V \Leftarrow P(x) = x$. נוכיח!

יהי $x \in V$, מכיוון ש-$v_1 \dots v_k$ בסיס אורתונורמלי ל-V קיימים $a_1 \dots a_k \in R$ כך ש-$x = a_1 v_1 + \dots + a_k v_k$ לכן מתקיים ש:

$$P_x = \left(\sum_{i=1}^{k} v_i v_i^T\right)\left(a_1 v_1 + \dots + a_k v_k\right) =$$

$$= \sum_{i=1}^{k}\left(v_i v_i^T a_1 v_1 + \dots v_i v_i^T a_i v_i + \dots + v_i v_i^T a_k v_k\right)$$

$$= \sum_{i=1}^{k}\left(a_1 v_i \overbrace{v_i^T v_1}^{0} + \dots + a_i v_i \overbrace{v_i^T v_i}^{1} + \dots a_k v_i \overbrace{v_i^T v_k}^{0}\right)$$

(בסיס אורתונורמלי)

$$= \sum_{i=1}^{k} a_i v_i = x$$

Multivariat calculus

20) יהי $v_n \dots v_1$ בסיס אורתונורמלי, אזי:

$$f(a) = \begin{bmatrix} | & | & & | \\ v_1 & v_2 & \cdots & v_n \\ | & | & & | \end{bmatrix} \begin{bmatrix} a_1 & & 0 \\ & \ddots & \\ 0 & & a_n \end{bmatrix} \begin{bmatrix} - & v_1 & - \\ - & v_2 & - \\ & \vdots & \\ - & v_n & - \end{bmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

$$= \begin{pmatrix} | & & | \\ a_1 v_1 & \cdots & a_n v_n \\ | & & | \end{pmatrix} \begin{pmatrix} \langle x, v_1 \rangle \\ \vdots \\ \langle x, v_n \rangle \end{pmatrix} \Rightarrow f(a)_j = \sum_{i=1}^{n} \langle x, v_i \rangle a_i \cdot v_{ij}$$

מהגדרת מכפלת
וקטור במטריצה

$$\begin{bmatrix} a_1 v_{11} & \cdots & a_2 v_{21} & \cdots & a_n v_{n1} \end{bmatrix} \begin{bmatrix} \langle x, v_1 \rangle \\ \vdots \\ \langle x, v_n \rangle \end{bmatrix}$$

לכן:

$n \times n$

$$J_\sigma(f) = \begin{pmatrix} \langle x, v_1 \rangle v_{11} & \cdots & \langle x, v_n \rangle v_{n1} \\ \vdots & & \vdots \\ \langle x, v_1 \rangle v_{1n} & \cdots & \langle x, v_n \rangle v_{nn} \end{pmatrix}$$

$$= \begin{pmatrix} | & & | \\ \langle x, v_1 \rangle v_1 & \cdots & \langle x, v_n \rangle v_n \\ | & & | \end{pmatrix}$$

$\blacksquare$

$$h(a)=\frac{1}{2}\|f(a)-y\|^2$$ (2 1

נציג קודם את הגרדיאנט של h:

כשזהו, נחשב את גרדיאנט $L(x):=\frac{1}{2}\|x-y\|^2$

$$L(x)=\frac{1}{2}\|x-y\|^2=\frac{1}{2}\sum_{i=1}^{n}(x_i-y_i)^2$$

הוכחנו בתרגיל

$$\nabla L(x)=\frac{1}{2}\cdot 2(x-y)=\boxed{x-y}$$

נשים לב: $h(a)=L(f(a))$

כי לפי שרשרת מינקו"ק:

$$\underbrace{(\nabla h(a))^T}_{\text{לפי הגדרה}}=J_a(h)=J_{f(a)}(L)\cdot J_a(f)=(\nabla L(f(a)))^T\cdot J_a(f)$$

כשזהו גרדיאנט

$$=(f(a)-y)^T\begin{bmatrix} | & & | \\ \langle x,v_1\rangle v_1 & \cdots & \langle x,v_n\rangle v_n \\ | & & | \end{bmatrix} \quad \text{:: } J_a(f)$$

$$=\left(\sum_{i=1}^{n}\langle x,v_i\rangle a_i v_i-y\right)^T\cdot\begin{pmatrix} | & & | \\ \langle x,v_1\rangle v_1 & \cdots & \langle x,v_n\rangle v_n \\ | & & | \end{pmatrix}$$

(כי V בסיס א.נ. אורתונורמלי)

$$=\left(\langle x,v_1\rangle^2 a_1-y\langle x,v_1\rangle v_1,\ \ldots,\ \langle x,v_n\rangle^2 a_n-y\langle x,v_n\rangle v_n\right)$$

ולכן

$$\nabla h(a)=\begin{pmatrix}\langle x,v_1\rangle^2\cdot a_1-y\langle x,v_1\rangle v_1\\ \vdots\\ \langle x,v_n\rangle^2 a_n-y\langle x,v_n\rangle v_n\end{pmatrix}$$

Scanned by CamScanner

22.

$$g(z)_j = \frac{e^{z_j}}{\sum_{k=1}^{K} e^{z_k}}$$

כאשר $i=j$ זה המקרה שחשבנו בתרגול $S_i(1-S_j)X$

כאשר $i \neq j$;

$$\frac{\partial g}{\partial a_i} = \frac{\partial}{\partial a_i}\frac{e^{a_j}}{\sum_{k=1}^{K} e^{a_k}} = \frac{0 - e^{a_j} \cdot \frac{\partial}{\partial a_i} \cdot \sum_{k=1}^{K} e^{a_k}}{\left(\sum_{k=1}^{K} e^{a_k}\right)^2} =$$

$$\underbrace{\frac{e^{a_j}}{\sum_{k=1}^{K} e^{a_k}}}_{S_j} \cdot \underbrace{\frac{e^{a_i}}{\sum_{k=1}^{K} e^{a_k}}}_{S_i} = S_j S_i$$

מכאן:

$$\boxed{J_{a_i}(x) = \begin{cases} S_i(1-S_j)X & i=j \\ S_i S_j(x) & i \neq j \end{cases}}$$

Practical question

23. כן זה קשר המחזור בהתאם.

24. בהרצה ראינו כי בערך המטריצה $S=\begin{bmatrix} S_x & 0 \\ 0 & S_y \end{bmatrix}$ מקיים שמטריצת השונות המשותפת היא:
$$\begin{bmatrix} (S_x\sigma(x))^2 & 0 \\ 0 & (S_y\sigma(y))^2 \end{bmatrix}$$
ונתון בתרגיל כי $\sigma(x)=\sigma(y)=\sigma(z)=1$, ולכן נקבל:

$$C=\begin{bmatrix} 0.01 & & \\ & 0.25 & 0 \\ 0 & & 4 \end{bmatrix}$$

כאשר לנושא הנומרי: ראינו בהרצאה ש.ו.פ $mean=0$ אז $C=\frac{XX^T}{n-1}$ ומקיים $x=0.1$, $y=0.5$, $z=2$.

25. ראינו מטריצה אורתוגונלית רנדומלית (מסומן A) מוגדרת מכפלה במטריצה שלנו ומסובבת את הנקודות. (הכל מוכפל של המטריצה המוכפלת).

מטריצת השונות היא מטריצה סימטרית לכן קיים פירוק EVD באופן הבא:

$$C=A\cdot\begin{pmatrix} (S_x\sigma(x))^2 & 0 \\ 0 & (S_y\sigma(y))^2 \end{pmatrix}A^T=A\begin{bmatrix} 0.01 & 0 & 0 \\ 0 & 0.25 & 0 \\ 0 & 0 & 4 \end{bmatrix}A^T$$

26. בהצלה 4 הצירים (x,y) ← מקור בהתפלגות נורמלית.

27. בהלפה הצירים x,y ונרמול data לפי הפקודה, אז z מקבל את הסקאלה ונקבל כי מדובר בהתפלגות נורמלית (כמו בכל הצירים).

28.
$$D^m(|\hat{p}-p|\geq\varepsilon)=D^m(|\hat{p}-E(\hat{p})|\geq\varepsilon)\leq\frac{var(x)}{m\varepsilon^2}\leq\frac{1}{4m\varepsilon^2}$$

$var(x)=p(1-p)\leq\frac{1}{4}$ עבור הכל מתקיים.

$$D^m(|\hat{p}-p|>\varepsilon)\overset{\text{הופדינג!}}{\leq}2\exp(-2m\varepsilon^2)\leq\delta$$

$$D^m(|\hat{p}-p|\geq\varepsilon)\leq 2e^{-2m\varepsilon^2}$$

$$\rightarrow m(\varepsilon,\delta)\leq\frac{1}{2\varepsilon^2}\log\left(\frac{2}{\delta}\right)$$

לסכם

29. a) נראה לראות שכאשר ה-m יגדל הגרף יתקרב ל-p המקורי. (עם חוק המספרים הגדולים), כלומר ממוצע הדגימות יתכנס ל-$p=0.25$

b) קורה (נכון.

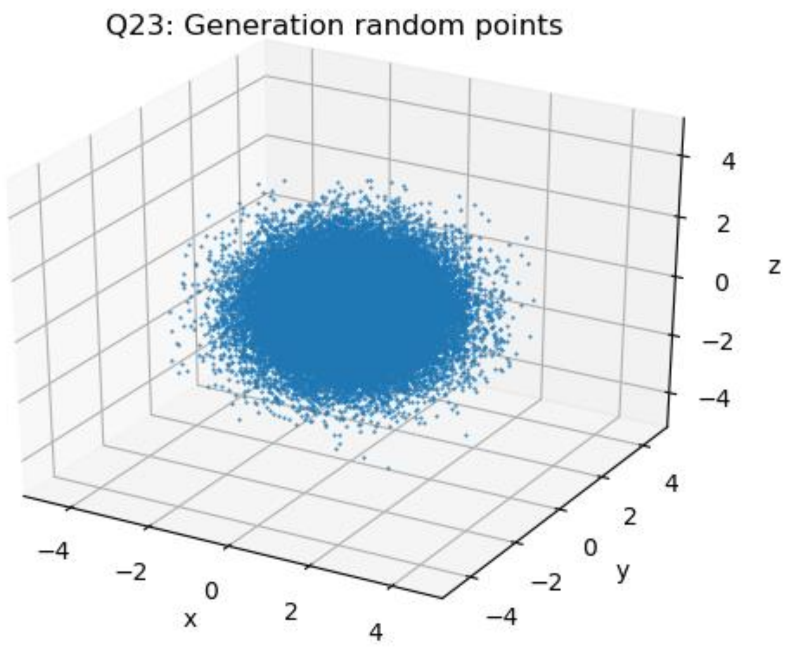
c) כל ה-ε ... ε מקרב ל-0, ואחוז השורות שיקיימו את האי שוויון יגדל. $(|\bar{X}_m - E(x)| \geq \varepsilon)$, כלומר $P(|\bar{X}_m - E(X)| \geq \varepsilon) = 1$

כי למעשה אני "מגדלת" את הטווח ומאפשרת למספר רב להיות בטווח גדול יותר ולכן זה יגרום לשינוי בזמנים הראשונים. $\varepsilon \to 0$

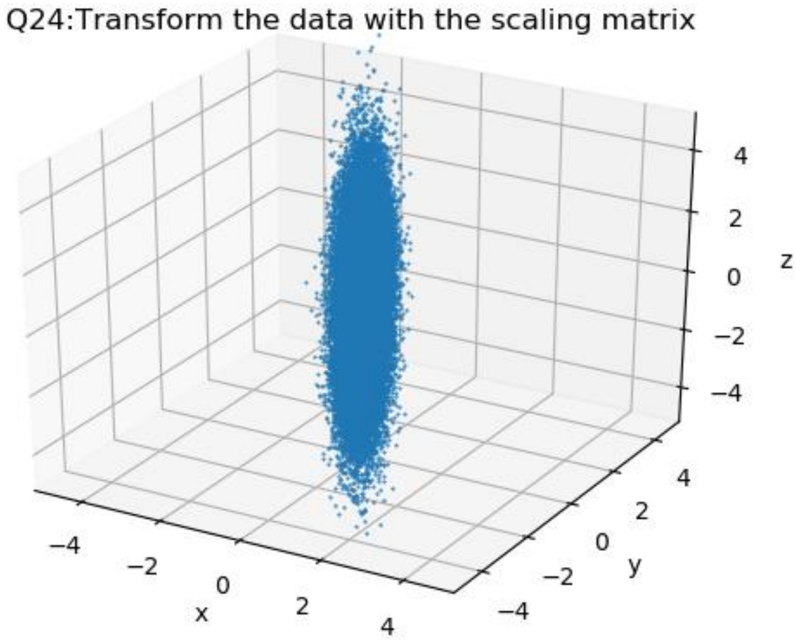
כל הגרפים בדף הבא:

שאלות תכנות –

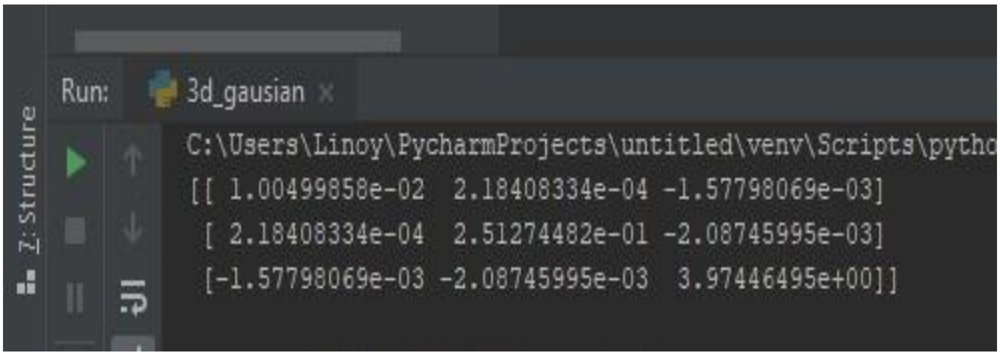
23)



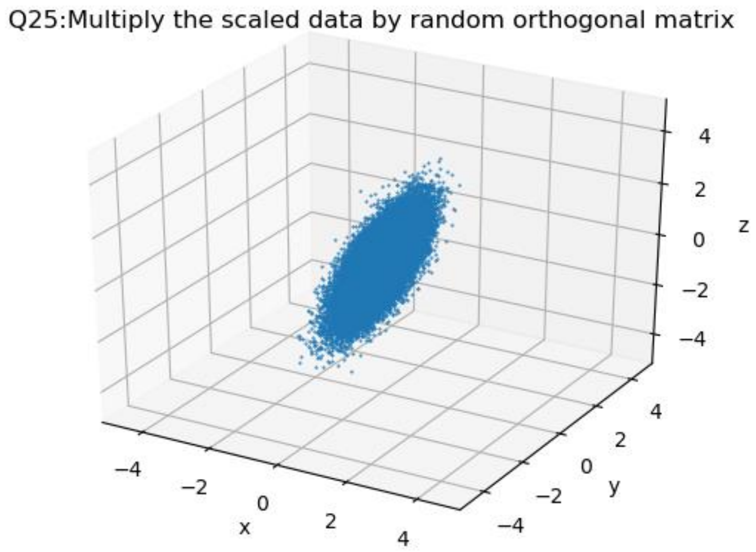
(24



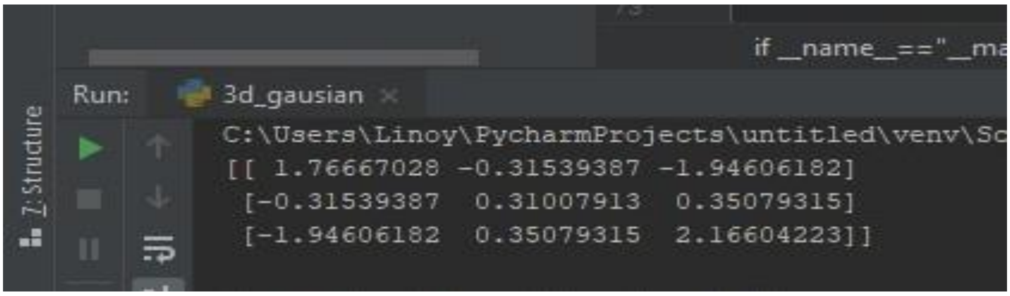
מטריצת השונות -



(25

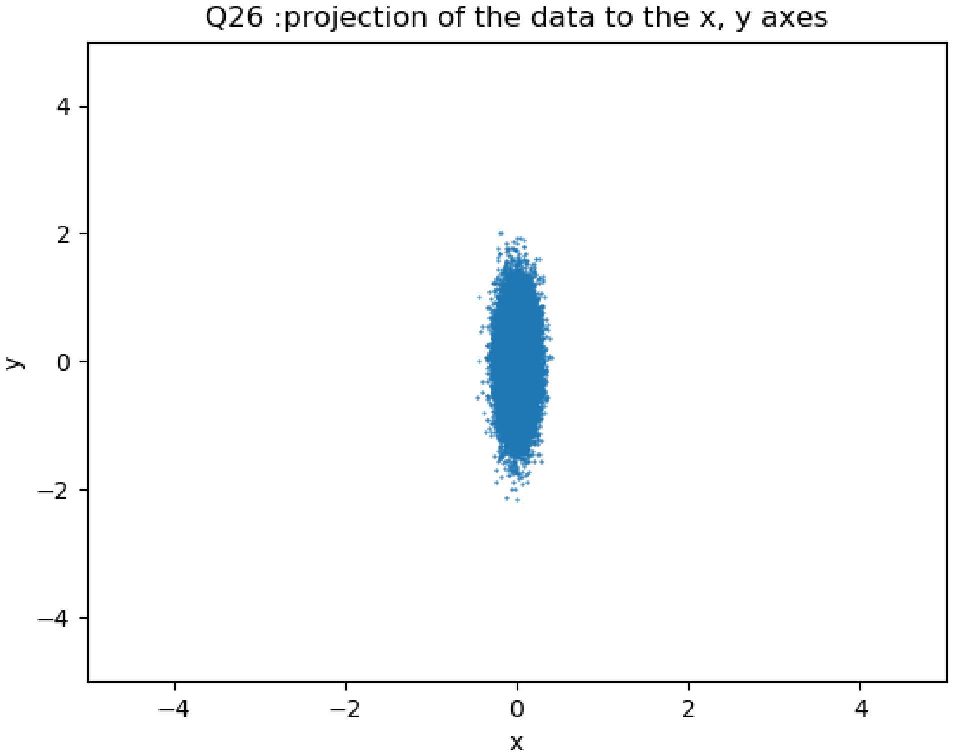


מטריצת שונות –

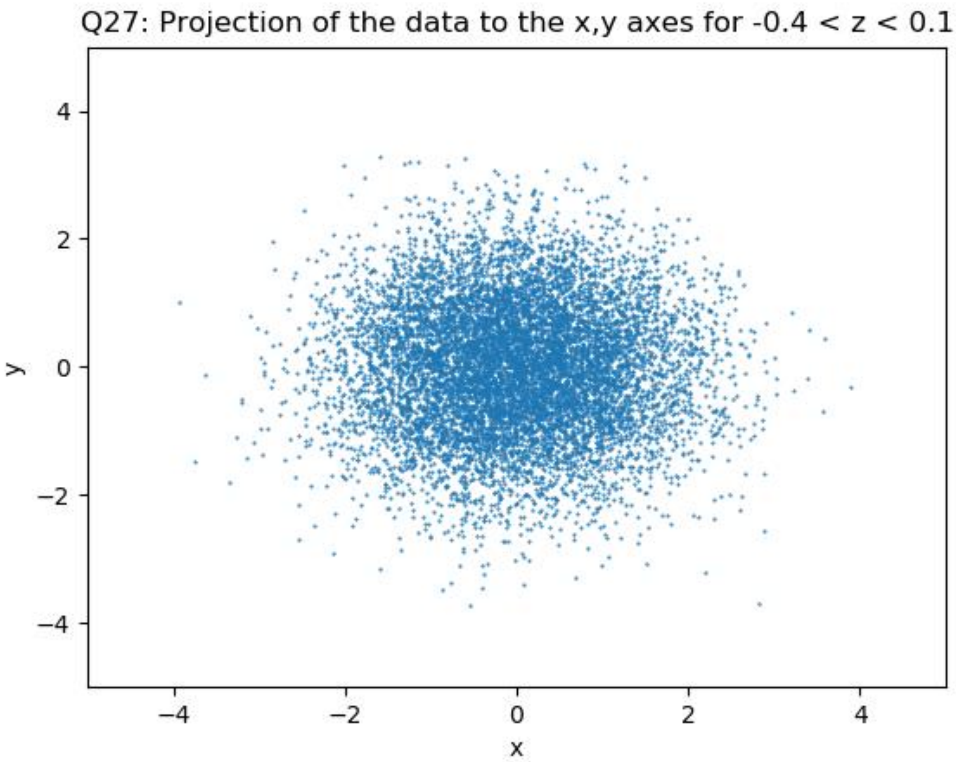



```
C:\Users\Linoy\PycharmProjects\untitled\venv\Sc
[[ 1.76667028 -0.31539387 -1.94606182]
 [-0.31539387  0.31007913  0.35079315]
 [-1.94606182  0.35079315  2.16604223]]
```

(26


Q26 :projection of the data to the x, y axes
y
x
4
2
0
−2
−4

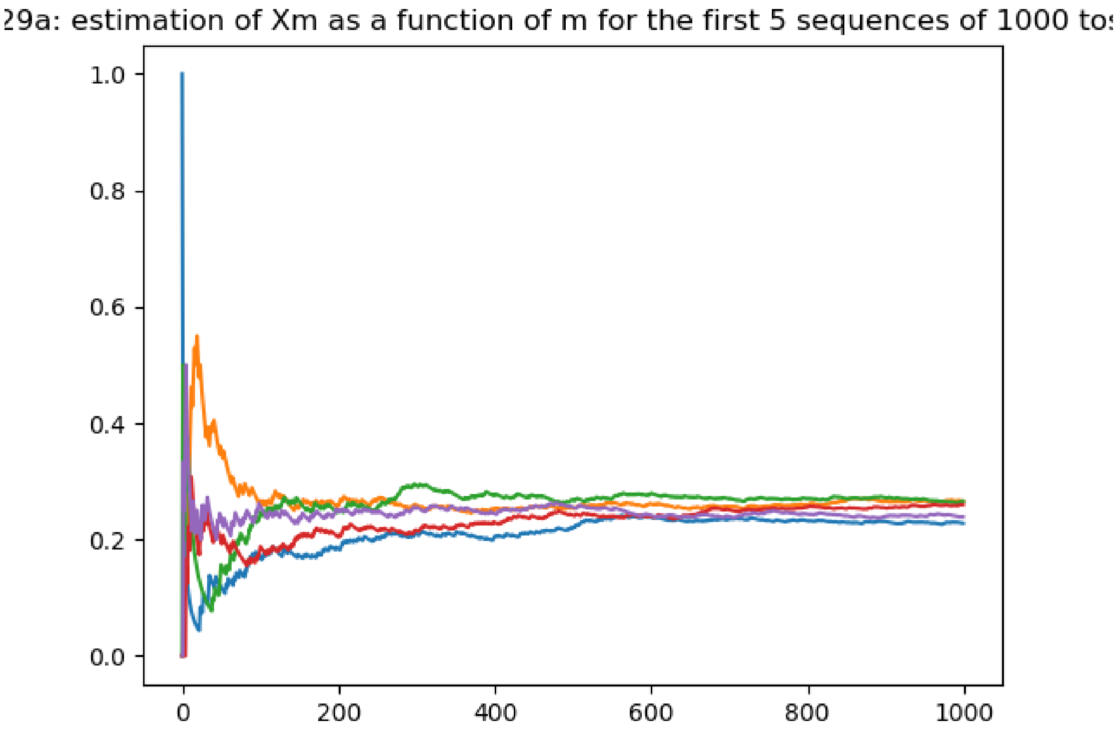
(27


Q27: Projection of the data to the x,y axes for -0.4 < z < 0.1
y
x
4
2
0
−2
−4

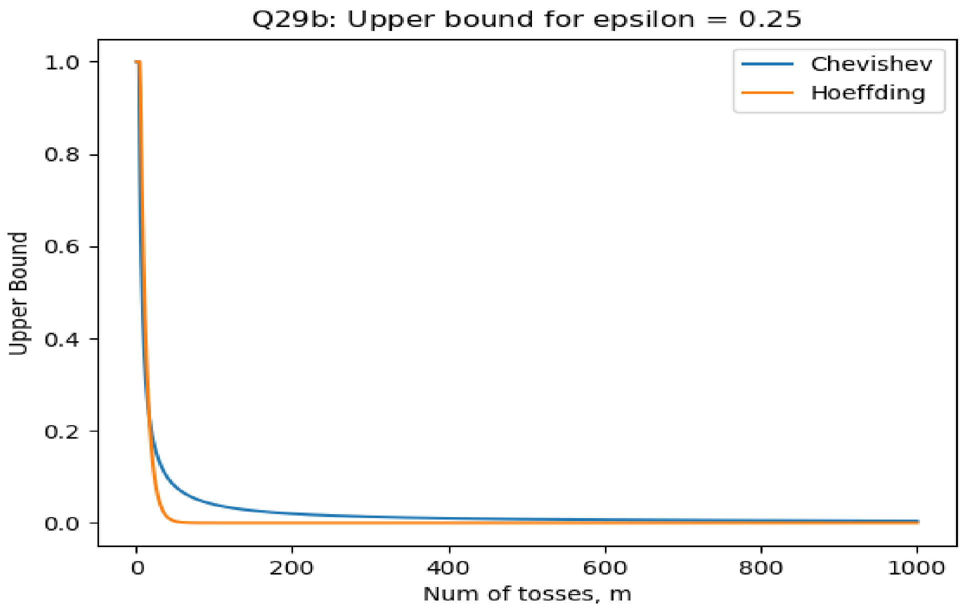
(28בדף נפרד .

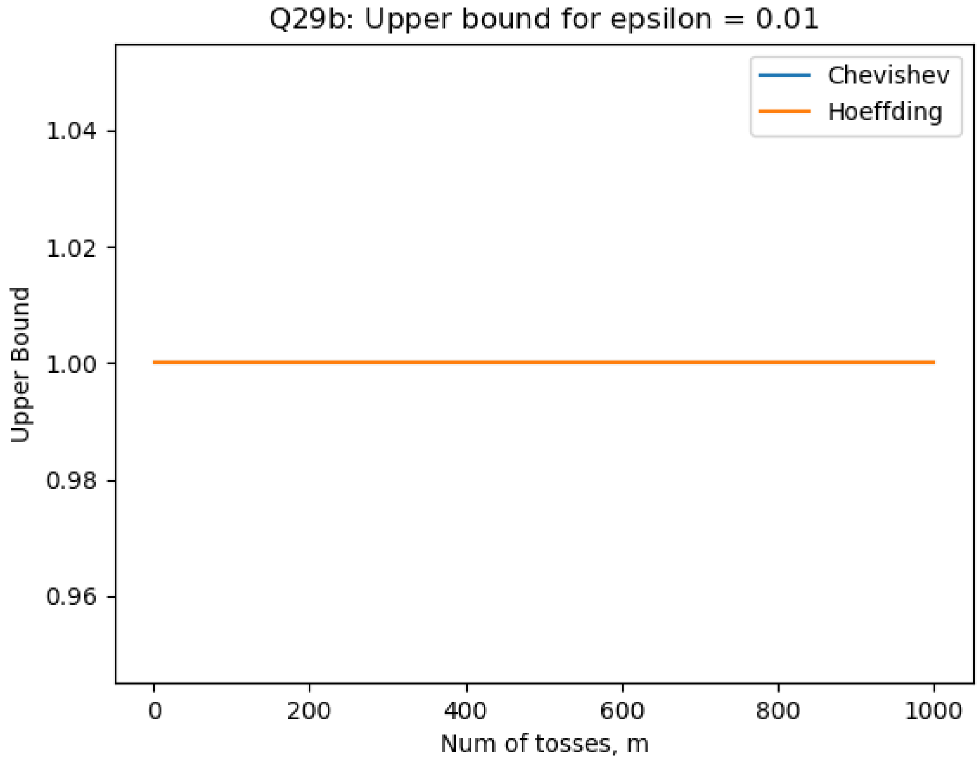
(29

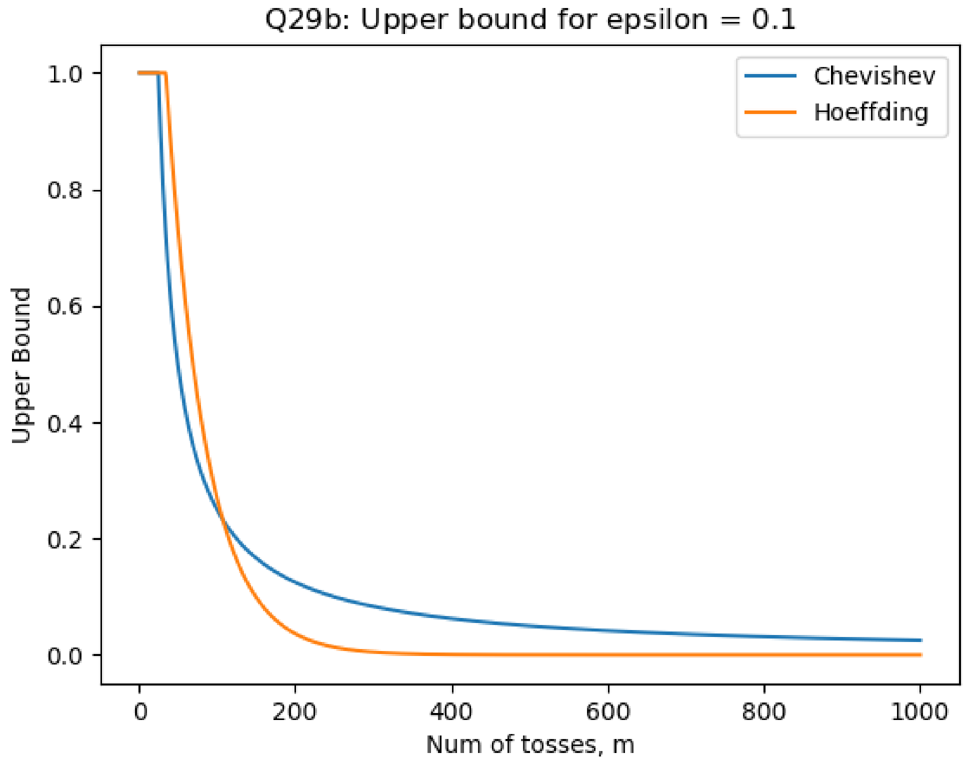
(A

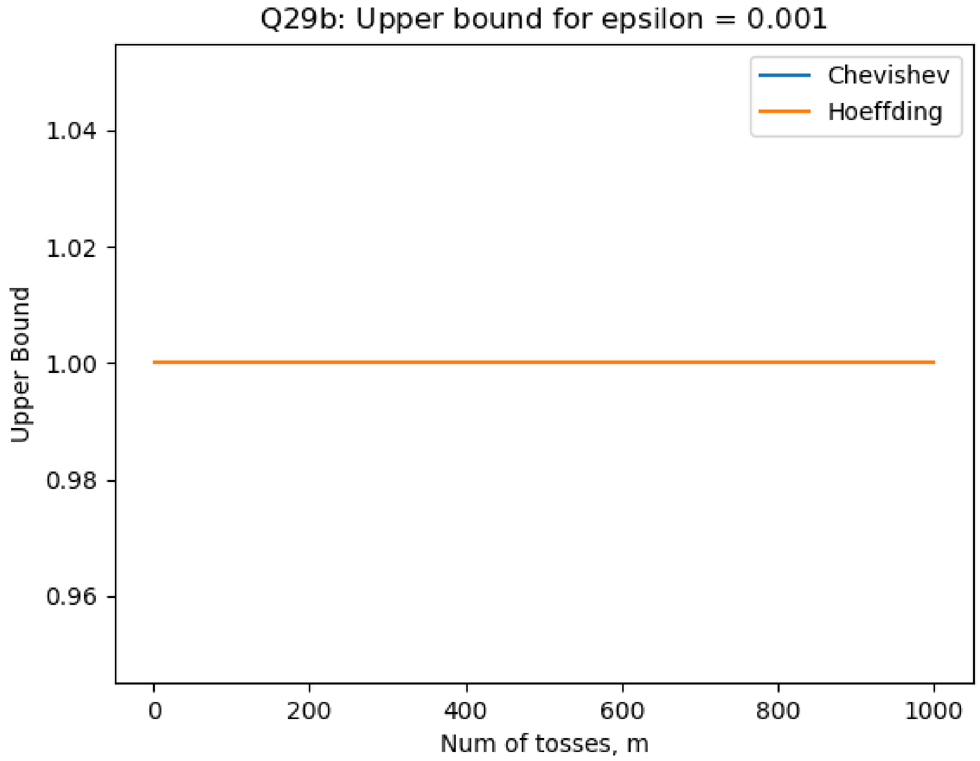


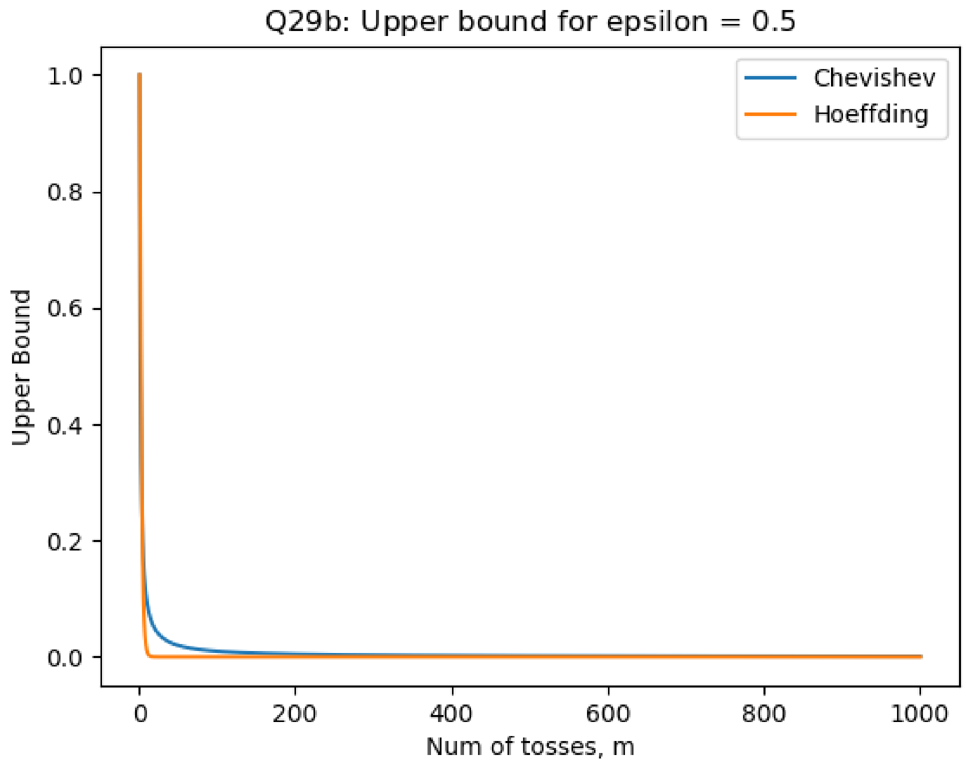
B)



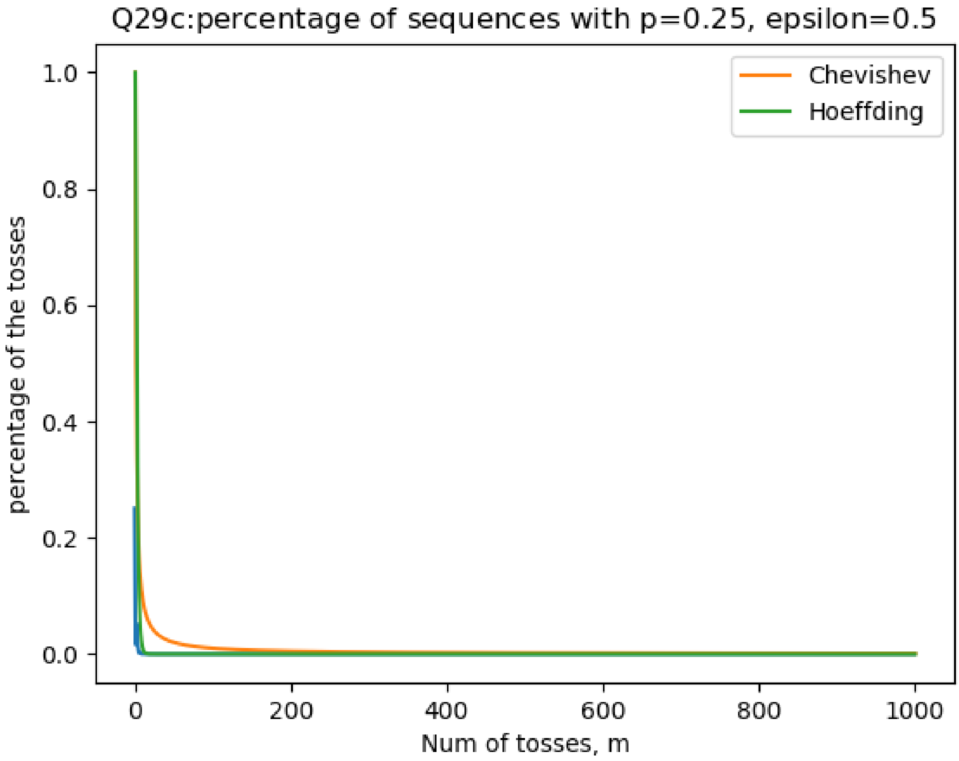

Q29b: Upper bound for epsilon = 0.01
Chevishev
Hoeffding
Upper Bound
1.04
1.02
1.00
0.98
0.96
0
200
400
600
800
1000
Num of tosses, m


Q29b: Upper bound for epsilon = 0.1
Chevishev
Hoeffding
Upper Bound
1.0
0.8
0.6
0.4
0.2
0.0
0
200
400
600
800
1000
Num of tosses, m

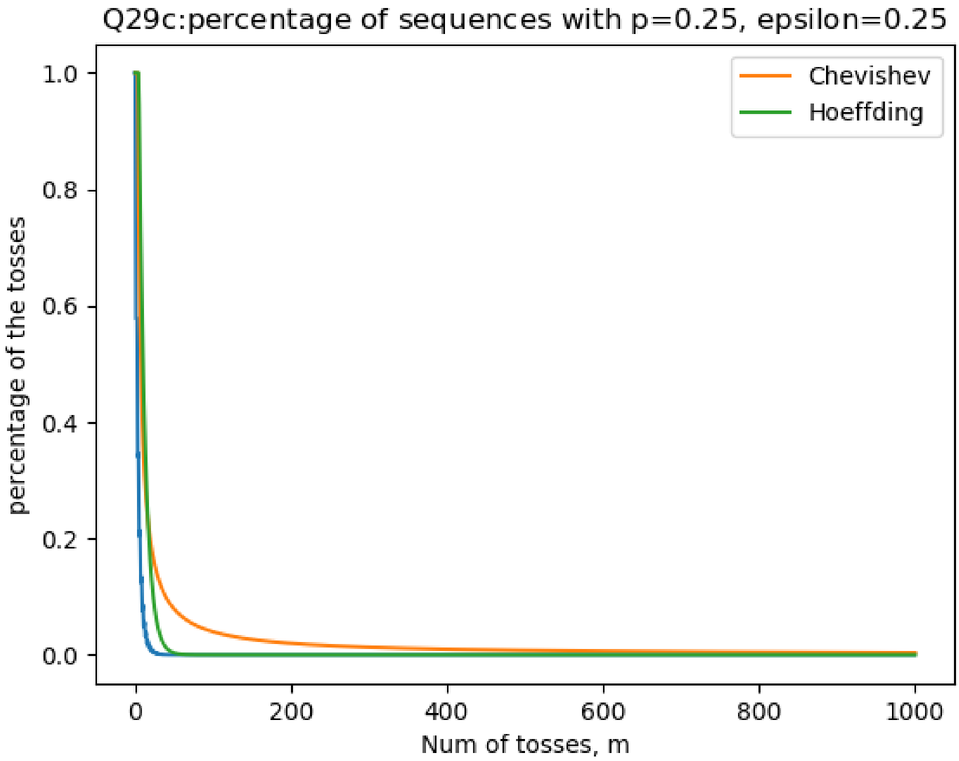

Q29b: Upper bound for epsilon = 0.001
Chevishev
Hoeffding
Upper Bound
1.04
1.02
1.00
0.98
0.96
0
200
400
600
800
1000
Num of tosses, m


Q29b: Upper bound for epsilon = 0.5
Chevishev
Hoeffding
Upper Bound
1.0
0.8
0.6
0.4
0.2
0.0
0
200
400
600
800
1000
Num of tosses, m

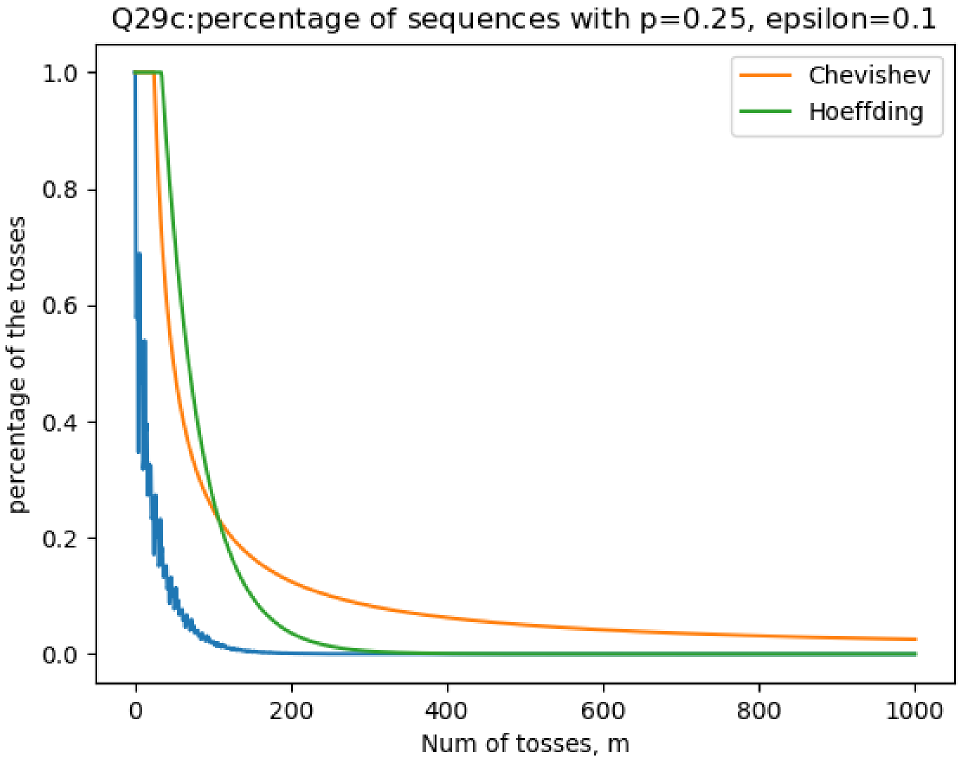
c)

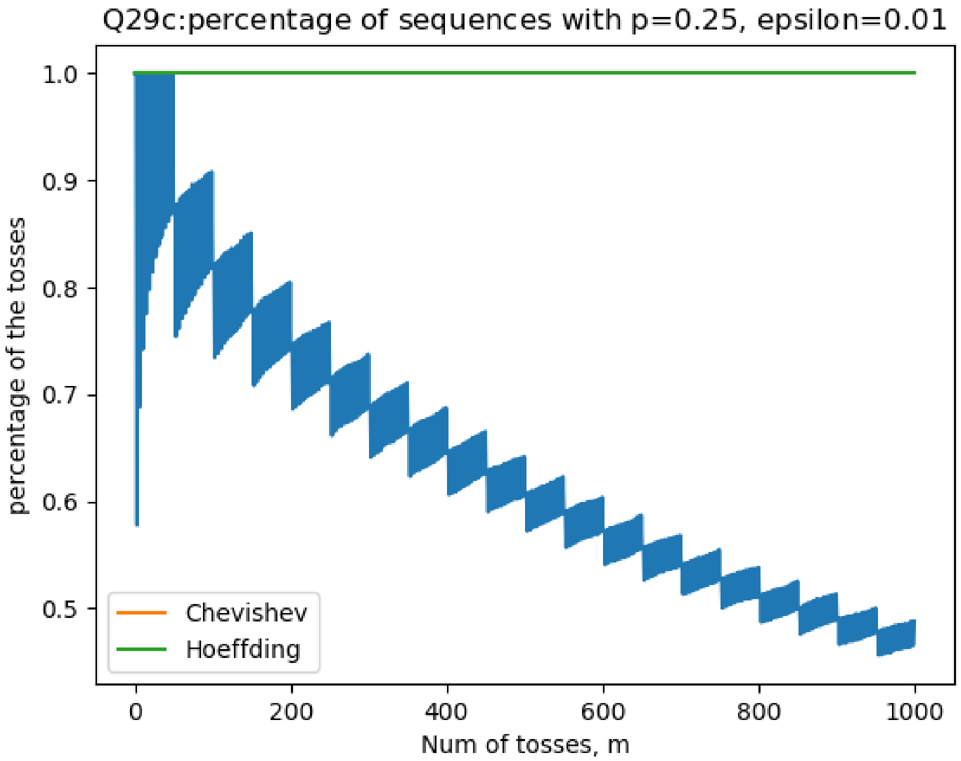


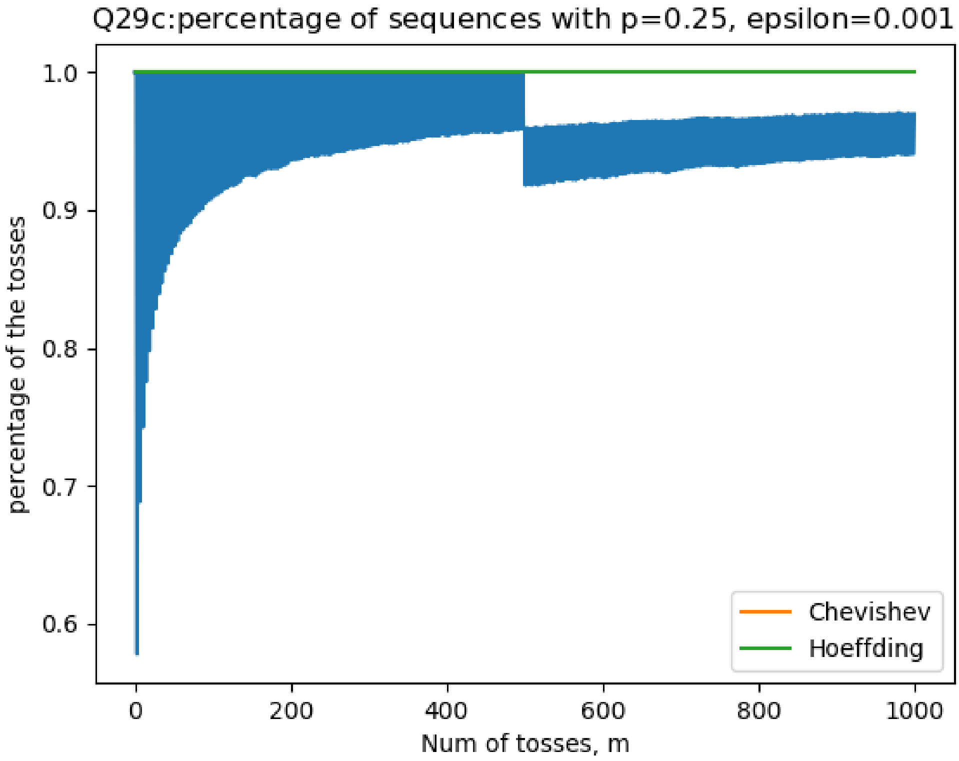
Q29c:percentage of sequences with p=0.25, epsilon=0.5
Chevishev
Hoeffding
1.0
0.8
0.6
0.4
0.2
0.0
percentage of the tosses
0
200
400
600
800
1000
Num of tosses, m



Q29c:percentage of sequences with p=0.25, epsilon=0.25
Chevishev
Hoeffding
1.0
0.8
0.6
0.4
0.2
0.0
percentage of the tosses
0
200
400
600
800
1000
Num of tosses, m


Q29c:percentage of sequences with p=0.25, epsilon=0.1
Chevishev
Hoeffding
percentage of the tosses
1.0
0.8
0.6
0.4
0.2
0.0
0
200
400
600
800
1000
Num of tosses, m


Q29c:percentage of sequences with p=0.25, epsilon=0.01
Chevishev
Hoeffding
percentage of the tosses
1.0
0.9
0.8
0.7
0.6
0.5
0
200
400
600
800
1000
Num of tosses, m


Q29c:percentage of sequences with p=0.25, epsilon=0.001
percentage of the tosses
1.0
0.9
0.8
0.7
0.6
0
200
400
600
800
1000
Num of tosses, m
Chevishev
Hoeffding