## Warm Up

1. נחשב את ההטלה של v =(1,2,3,4) על w=(0,-1,1,2):

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

2. נחשב את ההטלה של v =(1,2,3,4) על $w=(1,0,1,-1)$:

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

3. יהיו $v,w\in R^m$ המקיימים $v,w\neq 0$. נסמן את הזווית ביניהם θ.

<u>נוכיח</u>: $\langle v,w\rangle=0\ \leftrightarrow\theta\ =\ \mp 90°$:

←: מתקיים $\left(\frac{\langle v,w\rangle}{||v||||w||}\right)$ θ = arccos ולכן $\cos(\theta)=\frac{\langle v,w\rangle}{||v||||w||}$. ומכך נובע ש: $\langle v,w\rangle=0$

→: מההנחה, $v,w\neq 0$, ולכן $||v||,||w||\neq 0$ אזי גם $||v||\cdot||w||\neq 0$. מתקיים:
$\theta=\arccos\left(\frac{\langle v,w\rangle}{||v||||w||}\right)=\arccos\left(\frac{0}{||v||||w||}\right)=\arccos(0)$, וזה מתקיים אמ"מ
$\cos(\theta)=0$, כלומר כאשר: $\theta=\frac{\pi}{2}=90°$, ומאי זוגיות פונקציית הקוסינוס נכליל ונקבל שהנדרש מתקיים אמ"מ $\theta=\ \mp 90°$.

■

4. תהא $T:V\to W$ ה"ל ונסמן את המטריצה המייצגת אותה ב-A, והיא אורתונורמלית (ולכן מקיימת: $A^TA=I$). יהא $v\in V$, מתקיים:
$(||Ax||_2)^2\ =Ax^TAx=\ x^TA^TAx=x^Tx=\ (||x||_2)^2$ ומכך נובע: $||Ax||_2\ =\ ||x||_2$, כנדרש.

■

## SVD

5. תהא A מטריצה הפיכה ויהא $UDV^T$ פירוק ה-SVD שלה. ז"א ש: U-אורתוגונלית, T-אורתוגונלית ו-D- אלכסונית א"ש. מתקיים:

$$A^{-1}=(UDV^T)^{-1}=(V^T)^{-1}D^{-1}U^{-1}=\ (V^{-1})^TD^{-1}U^{-1}$$

נבחין שגדלים אלו קלים לחישוב:

* אם $D=diag(\sigma_1,\dots,\sigma_n)$ אזי $D^{-1}=diag(\sigma_1{}^{-1},\dots,\sigma_n{}^{-1})$.

* מאורתונורמליות V,U: $U^T=U^{-1},V^T=V^{-1}$.

6. נמצא את פירוק SVD של: $C=\begin{bmatrix}5&5\\-1&7\end{bmatrix}$.

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

נמצא ע"ע:

$$f_{C^TC}=\ |I\lambda-C^TC|=\ \left|\begin{bmatrix}\lambda-26&-18\\-18&\lambda-74\end{bmatrix}\right|=(\lambda-80)(\lambda-20)$$

ולכן הע"ע הם: $\lambda_1=20$, $\lambda_2=80$ ומצאנו את: **D** $=\begin{bmatrix}\sqrt{20}&0\\0&\sqrt{30}\end{bmatrix}$.

הו"ע הם:

$$N(C^TC-80I)=N\left(\begin{bmatrix}-54&18\\18&-6\end{bmatrix}\right)=\ N\left(\begin{bmatrix}1&-\frac{1}{3}\\0&0\end{bmatrix}\right)=sp\left\{\begin{pmatrix}1\\3\end{pmatrix}\right\}$$

$$N(C^TC-20I)=N\left(\begin{bmatrix}6&18\\18&54\end{bmatrix}\right)=\ N\left(\begin{bmatrix}1&3\\0&0\end{bmatrix}\right)=sp\left\{\begin{pmatrix}-3\\1\end{pmatrix}\right\}$$

נורמל אותם: $v_1 = \frac{1}{\sqrt{10}}\binom{-3}{1}$, $v_2 = \frac{1}{\sqrt{10}}\binom{1}{3}$. ומצאנו גם את $\boldsymbol{V} = \begin{bmatrix} \frac{-3}{\sqrt{10}} & \frac{1}{\sqrt{10}} \\ \frac{1}{\sqrt{10}} & \frac{3}{\sqrt{10}} \end{bmatrix}$.

מתקיים:

$CV = \begin{bmatrix} 5 & 5 \\ -1 & 7 \end{bmatrix}\begin{bmatrix} \frac{-3}{\sqrt{10}} & \frac{1}{\sqrt{10}} \\ \frac{1}{\sqrt{10}} & \frac{3}{\sqrt{10}} \end{bmatrix} = \begin{bmatrix} \frac{-10}{\sqrt{10}} & \frac{20}{\sqrt{10}} \\ \frac{10}{\sqrt{10}} & \frac{20}{\sqrt{10}} \end{bmatrix} = \begin{bmatrix} -\sqrt{10} & 2\sqrt{10} \\ \sqrt{10} & 2\sqrt{10} \end{bmatrix}$, ומכיוון ש:

$UD = \begin{bmatrix} a & b \\ c & d \end{bmatrix}\begin{bmatrix} \sqrt{20} & 0 \\ 0 & \sqrt{80} \end{bmatrix} = \begin{bmatrix} a\sqrt{20} & b\sqrt{80} \\ c\sqrt{20} & d\sqrt{80} \end{bmatrix}$, אז: $\boldsymbol{U} = \begin{bmatrix} \frac{-1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix}$.

7. תהא $A \in R_{m\times n}$ ונגדיר: $C_0 = A^T A$. יהיו $\lambda_1, .., \lambda_n$ הע"ע של $C_0$ כך שמתקיים:
(*) $\lambda_1 \geq .. \geq \lambda_n \geq 0$ ו- $v_1, .., v_n$ הו"ע העצמיים המתאימים להם. נסמן את פירוק ה-*SVD* של: $A = UDV^T$. אז מתקיים:

$C_0 = A^T A = (UDV^T)^T(UDV^T) = VD^T U^T UDV^T = VD^T DV^T := VSV^T$.

כאשר *S* אלכסונית.

נבחין שקיבלנו את פירוק ה-*EVD* של $C_0$, ולכן אלכסון *S* מכיל את הע"ע של $C_0$.

נחשב:

$b_{k+1} = \frac{C_0 b_k}{\|C_0 b_k\|} = \frac{VSV^T VSV^T b_{k-1}}{\|VSV^T VSV^T b_{k-1}\|} = \frac{VS^2V^T b_{k-1}}{\|VS^2V^T b_{k-1}\|} = \cdots = \frac{VS^kV^T b_0}{\|VS^kV^T b_0\|}$.

נפתח את המונה:

$$\begin{bmatrix} | & \dots & | \\ v_1 & \dots & v_n \\ | & \dots & | \end{bmatrix}\begin{bmatrix} \lambda_1^k & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & \lambda_n^k \end{bmatrix}\begin{bmatrix} - & v_1 & - \\ \vdots & \vdots & \vdots \\ - & v_n & - \end{bmatrix}\begin{bmatrix} | \\ b_0 \\ | \end{bmatrix} = V \cdot \begin{bmatrix} - & \sum_{i=1}^{n} v_{i,1}\lambda_i^k\langle v_1, b_0\rangle & - \\ \vdots & \vdots & \vdots \\ - & \sum_{i=1}^{n} v_{i,1}\lambda_i^k\langle v_n, b_0\rangle & - \end{bmatrix}$$

נבחין שמ-(*): לכל *i*: $\lambda_i^k$ זניח לעומת $\lambda_1^k$, ולכן נוכל להתעלם מכפל ב- $\lambda_i^k$ שאינם $\lambda_1^k$, ונקבל:

$$\begin{bmatrix} - & v_{1,1}\lambda_1^k\langle v_1, b_0\rangle & - \\ \vdots & \vdots & \vdots \\ - & v_{n,n}\lambda_1^k\langle v_n, b_0\rangle & - \end{bmatrix}.$$

אזי $$b_{k+1} = \frac{\begin{bmatrix} - & v_{1,1}\lambda_1^k\langle v_1,b_0\rangle & - \\ \vdots & \vdots & \vdots \\ - & v_{n,n}\lambda_1^k\langle v_n,b_0\rangle & - \end{bmatrix}}{\left\|\begin{bmatrix} - & v_{1,1}\lambda_1^k\langle v_1,b_0\rangle & - \\ \vdots & \vdots & \vdots \\ - & v_{n,n}\lambda_1^k\langle v_n,b_0\rangle & - \end{bmatrix}\right\|} = \frac{v_1\lambda_1^k\langle v_1,b_0\rangle}{\|v_1\lambda_1^k\langle v_1,b_0\rangle\|} = \frac{v_1\lambda_1^k\langle v_1,b_0\rangle}{\|v_1\|\lambda_1^k\langle v_1,b_0\rangle} = \frac{v_1}{\|v_1\|} = v_1$$

ונכונות המעבר האחרון נובעת מכך ש-V מטריצה אורתונורמלית.

■

8. $$= \begin{bmatrix} | & \dots & | \\ u_1 & \dots & u_n \\ | & \dots & | \end{bmatrix}\begin{bmatrix} \sigma_1 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & \sigma_n \end{bmatrix}\begin{bmatrix} - & u_1 & - \\ \vdots & \vdots & \vdots \\ - & u_n & - \end{bmatrix}\begin{bmatrix} x_n \\ \vdots \\ x_n \end{bmatrix} = f(\sigma)$$

$$\begin{bmatrix} | & \dots & | \\ u_1 & \dots & u_n \\ | & \dots & | \end{bmatrix}\begin{bmatrix} \sigma_1\ \langle u_1, x\rangle \\ \vdots \\ \sigma_n\ \langle u_n, x\rangle \end{bmatrix} = \sum_{i=1}^{n} \sigma_i\langle u_i, x\rangle u_i$$

מתקיים: $\langle u_i, x\rangle u_i = \frac{\partial f_i}{\partial \sigma_i}$ ולכן: $J_\sigma(f) = \begin{bmatrix} | & \cdots & | \\ \langle u_1, x\rangle u_1 & \cdots & \langle u_n, x\rangle u_n \\ | & \cdots & | \end{bmatrix}$.

9. נגדיר את $g(\sigma) = \frac{\|\sigma\text{-}y\|^2}{2}$, אזי נבחין ש: $h(\sigma) = g^\circ f(\sigma)$, וכן:

$$g(\sigma) = \frac{1}{2}\sum_{i=1}^{n}(\sigma_i\text{-}y_i)^2 \ = \ \frac{1}{2}\sum_{i=1}^{n}\sigma_i{}^2 \text{ - } 2\sigma_i y_i + y_i{}^2$$

אזי:

$$\frac{\partial g_i}{\partial \sigma_i}(\sigma) \ = \ \frac{(2\sigma_i\text{-}2y_i)}{2} = \ \sigma_i\text{-}y_i$$

ולכן: $J_\sigma(g) = (\sigma\text{-}y)^T$, ומכלל השרשרת:

$$J_\sigma(h) = \ J_\sigma(g \circ f) = \ J_{f(\sigma)}(g)\, J_\sigma(f) = (f(\sigma) - y)^T \begin{bmatrix} | & \cdots & | \\ \langle u_1, x\rangle u_1 & \cdots & \langle u_n, x\rangle u_n \\ | & \cdots & | \end{bmatrix} =$$

$$(\langle f(\sigma) - y, \langle u_1, x\rangle u_1 \ \rangle, \ldots, \langle f(\sigma) - y, \langle u_1, x\rangle u_1 >)$$

$$= (\langle u_1, x\rangle\langle f(\sigma)\text{-}y, u_1\rangle, \ldots, \langle u_n, x\rangle\langle f(\sigma)\text{-}y, u_n\rangle)$$

$$= \begin{bmatrix} \langle u_1, x\rangle\langle f(\sigma)\text{-}y, u_1\rangle \\ \vdots \\ \langle u_n, x\rangle\langle f(\sigma)\text{-}y, u_n\rangle \end{bmatrix}\rangle$$

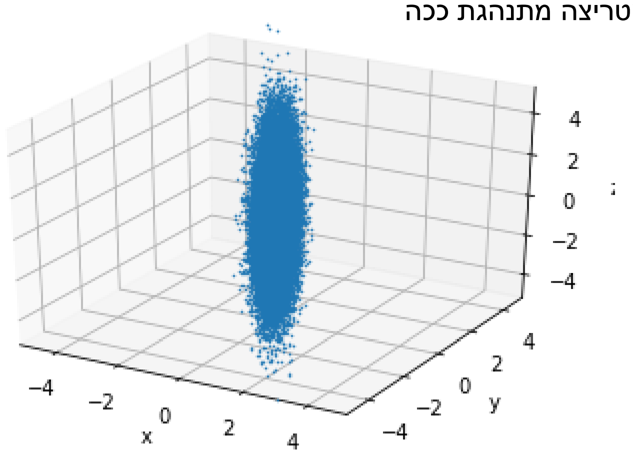
10. נחשב עכשיו את היעקוביאן עבור i שונה מ-j:

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

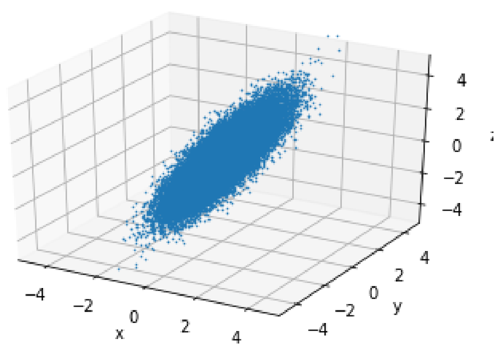
$$= -\frac{e^{a_i}}{\sum_{k=1}^{N} e^{a_k}} \cdot \frac{e^{a_j}}{\sum_{k=1}^{N} e^{a_k}} = -S_i \cdot S_j$$

11. כלום

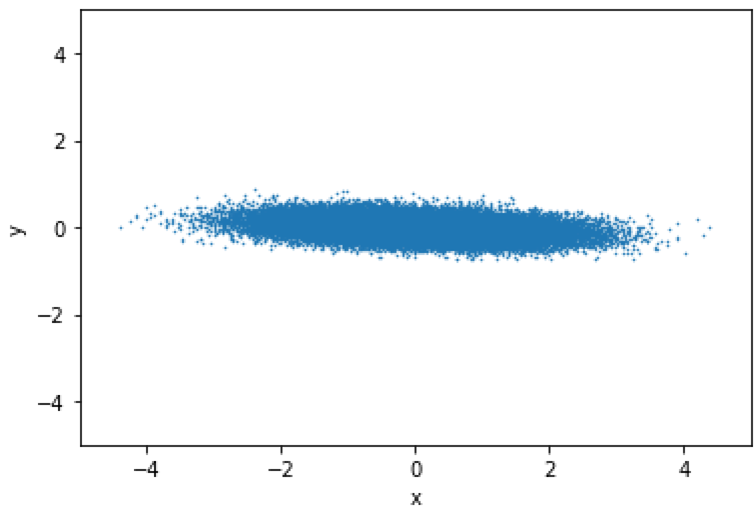
12. המטריצה מתנהגת ככה



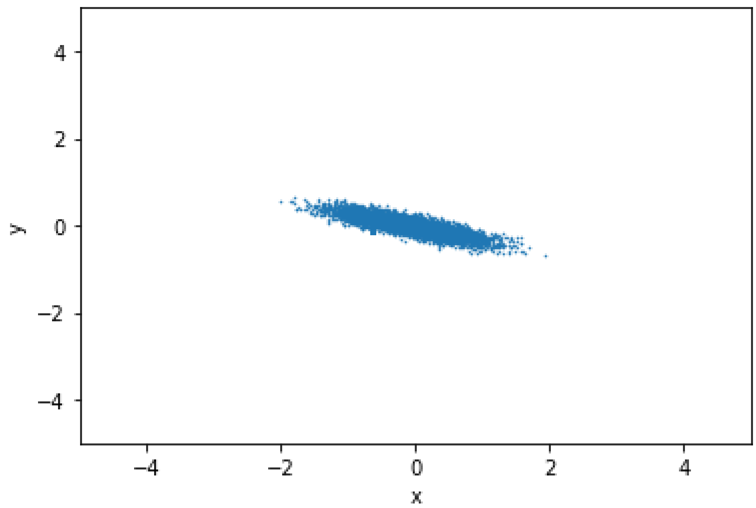

13. המטריצה מנהגת ככה



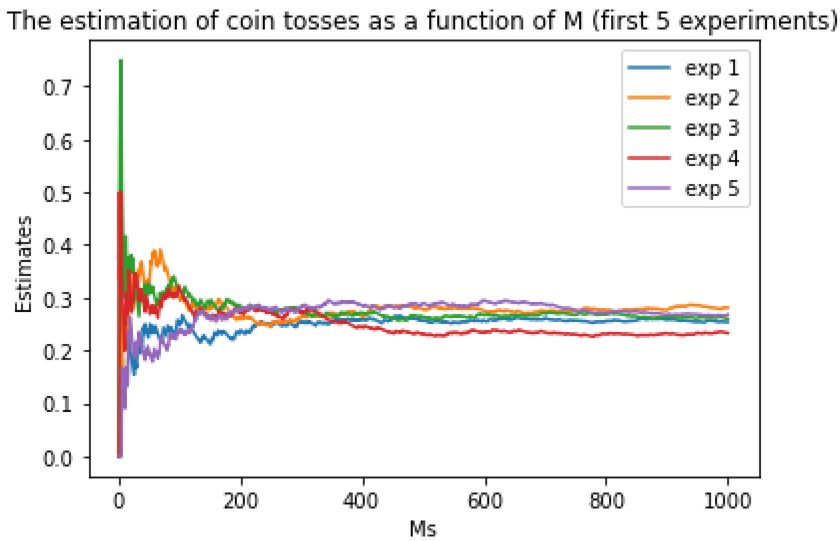
.14

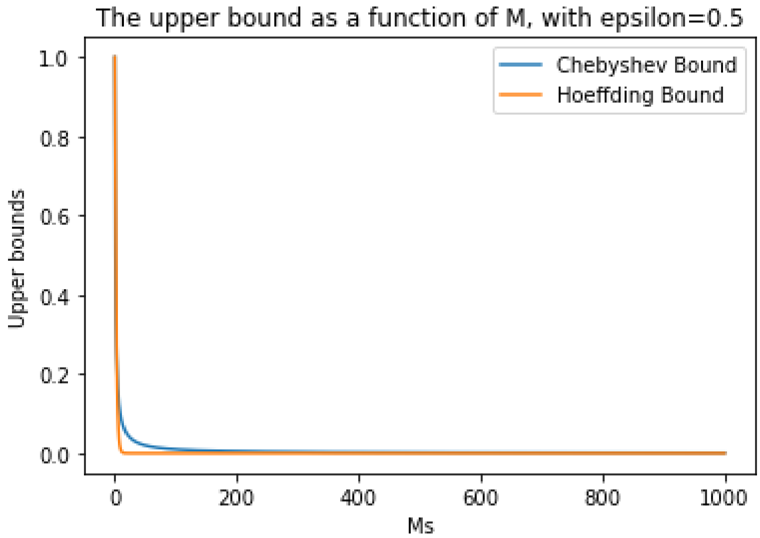

y
x

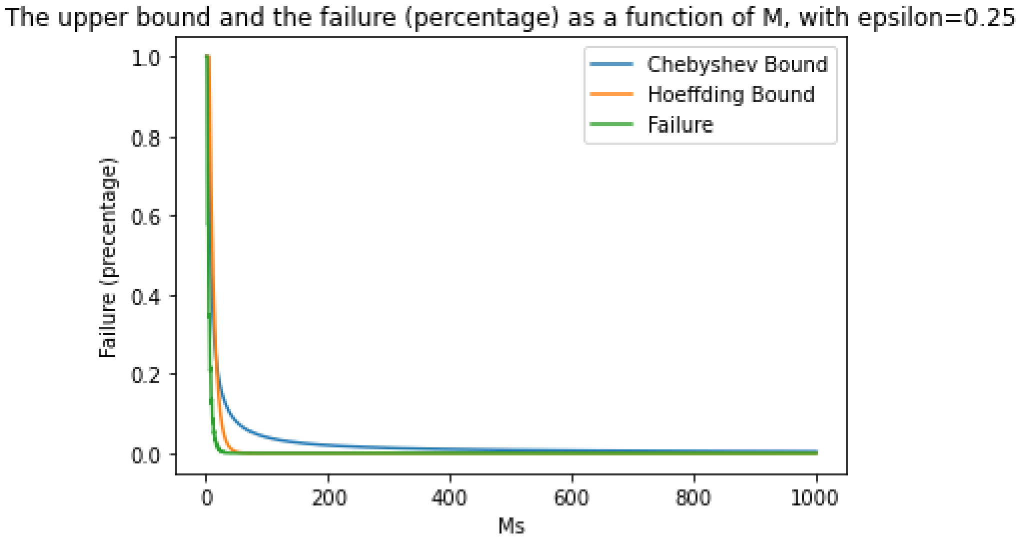
.15

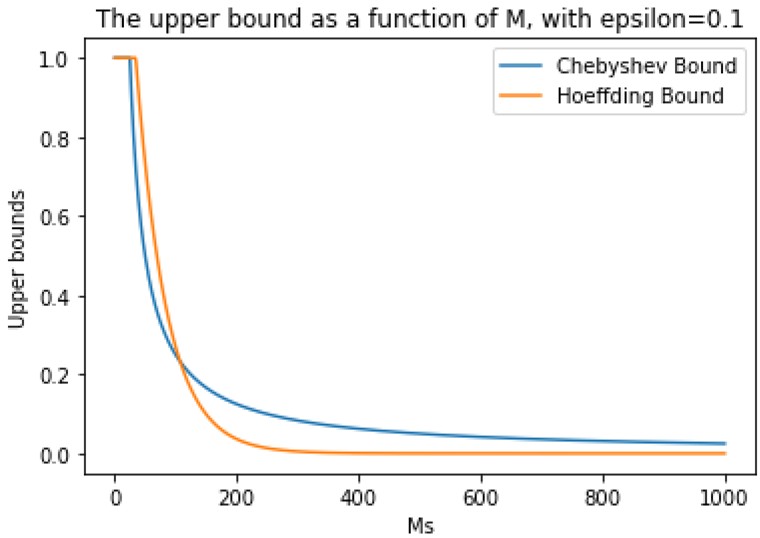

y
x

.16

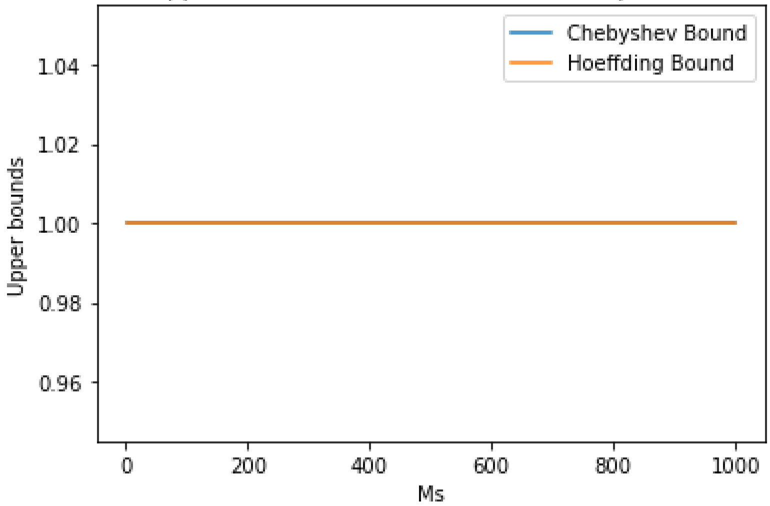

The estimation of coin tosses as a function of M (first 5 experiments)
exp 1
exp 2
exp 3
exp 4
exp 5
Estimates
Ms


The upper bound as a function of M, with epsilon=0.5
Chebyshev Bound
Hoeffding Bound
Upper bounds
Ms
0.0
0.2
0.4
0.6
0.8
1.0
0
200
400
600
800
1000

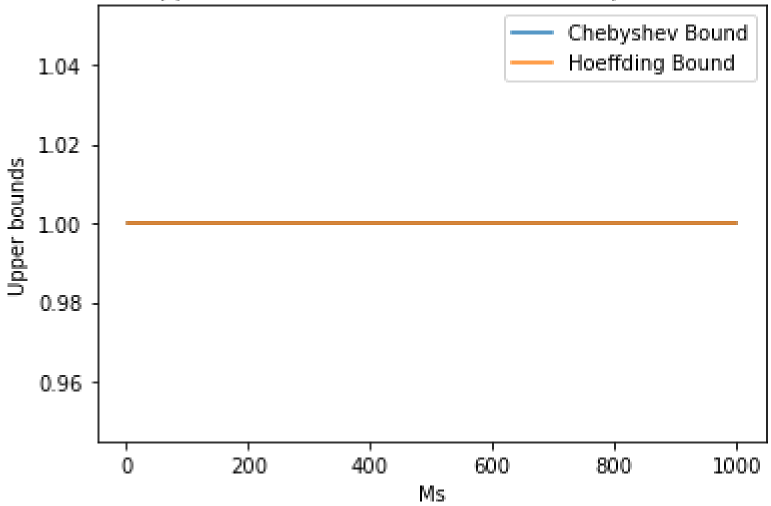

The upper bound and the failure (percentage) as a function of M, with epsilon=0.25
Chebyshev Bound
Hoeffding Bound
Failure
Failure (precentage)
Ms
0.0
0.2
0.4
0.6
0.8
1.0
0
200
400
600
800
1000


The upper bound as a function of M, with epsilon=0.1
Chebyshev Bound
Hoeffding Bound
Upper bounds
Ms
0.0
0.2
0.4
0.6
0.8
1.0
0
200
400
600
800
1000

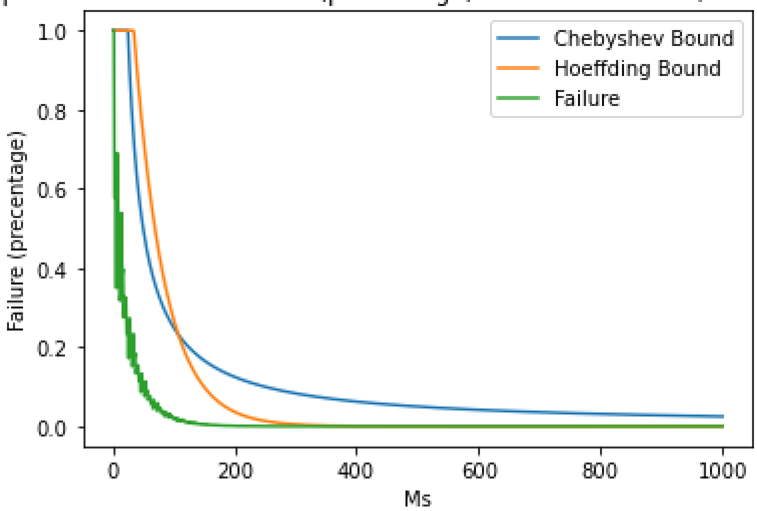
The upper bound as a function of M, with epsilon=0.01

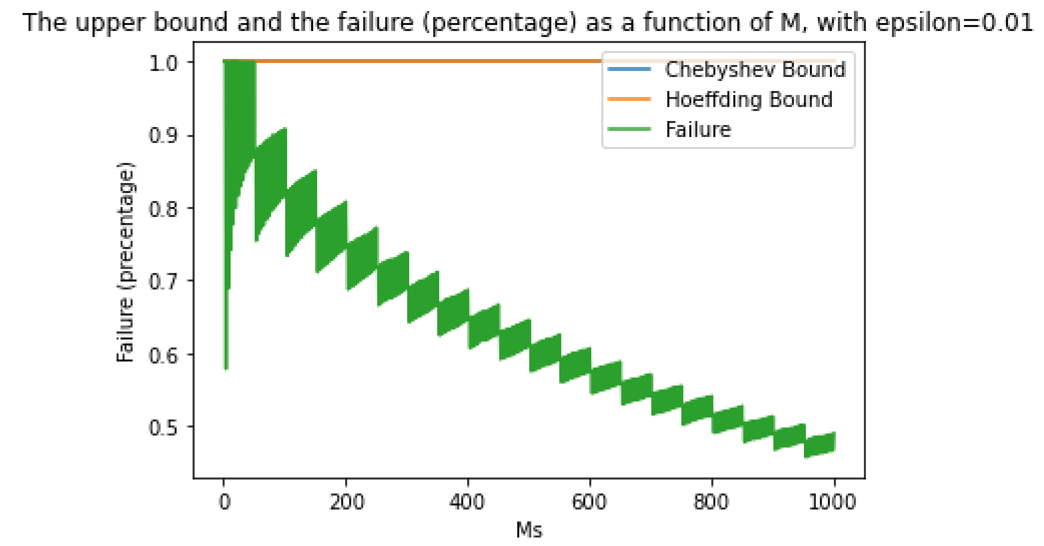


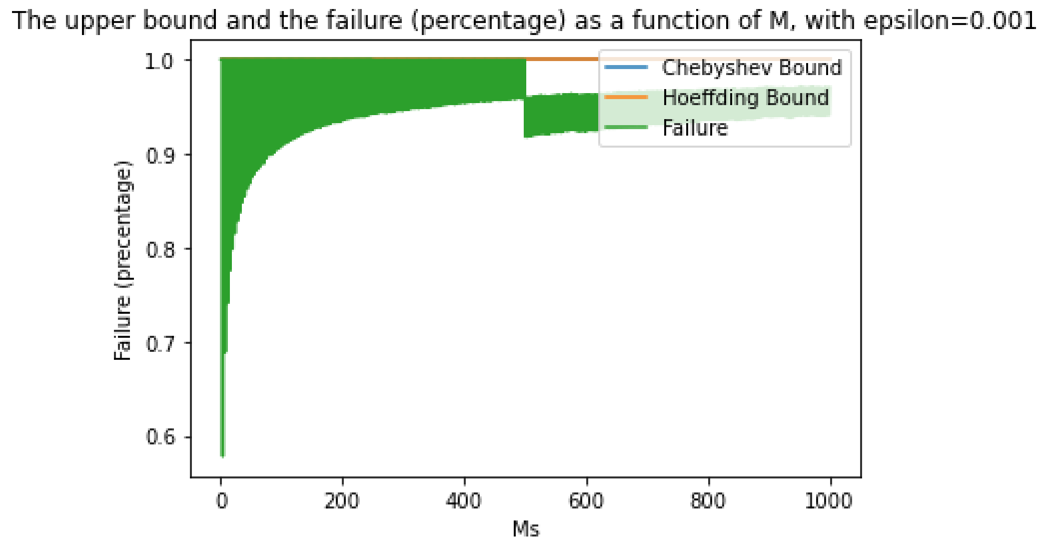

The upper bound as a function of M, with epsilon=0.001

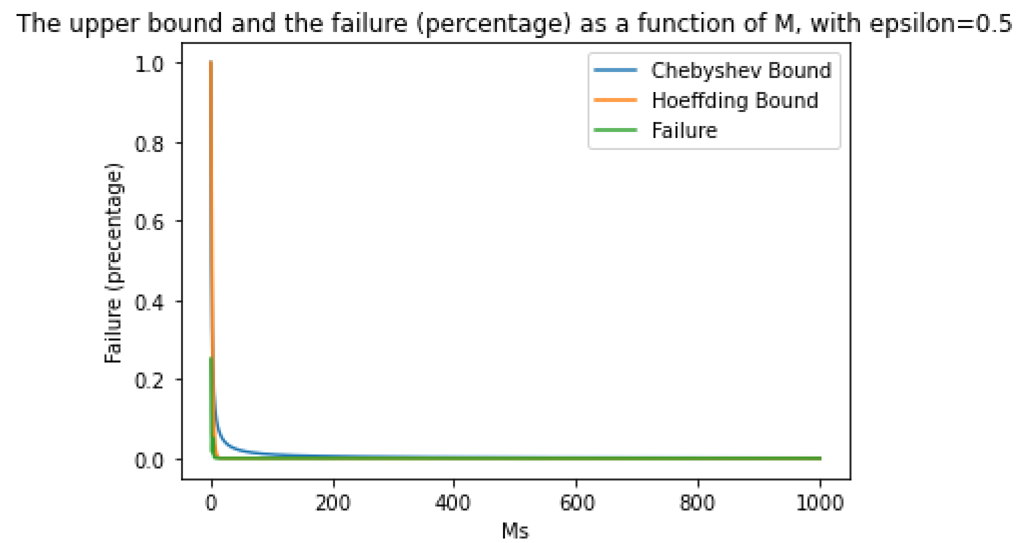


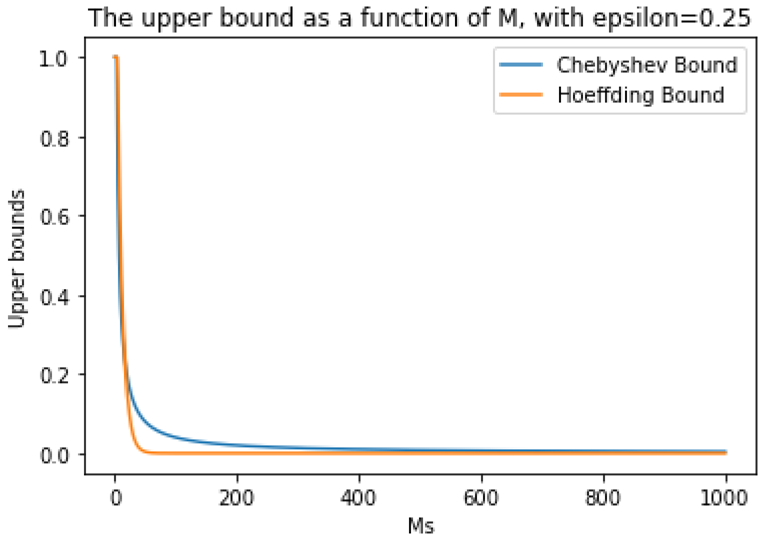

The upper bound and the failure (percentage) as a function of M, with epsilon=0.1




The upper bound and the failure (percentage) as a function of M, with epsilon=0.01
Chebyshev Bound
Hoeffding Bound
Failure
Failure (precentage)
Ms


The upper bound and the failure (percentage) as a function of M, with epsilon=0.001
Chebyshev Bound
Hoeffding Bound
Failure
Failure (precentage)
Ms


The upper bound and the failure (percentage) as a function of M, with epsilon=0.5
Chebyshev Bound
Hoeffding Bound
Failure
Failure (precentage)
Ms


The upper bound as a function of M, with epsilon=0.25
Chebyshev Bound
Hoeffding Bound
Upper bounds
1.0
0.8
0.6
0.4
0.2
0.0
0
200
400
600
800
1000
Ms