חלק תיאורטי

1.1.1 - Linear Algebra

1) $A = \begin{bmatrix} 1 & 1 & 0 \\ 1 & -1 & 2 \end{bmatrix}$

$$A^T A = \begin{bmatrix} 1 & 1 \\ 1 & -1 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} 1 & 1 & 0 \\ 1 & -1 & 2 \end{bmatrix} = \begin{bmatrix} 2 & 0 & 2 \\ 0 & 2 & -2 \\ 2 & -2 & 4 \end{bmatrix}$$

נמצא את הערכים העצמיים של A^TA: ע"י דטרמיננטה והשוואה ל-0:

$$|A^TA - I\lambda| = \left| \begin{bmatrix} 2-\lambda & 0 & 2 \\ 0 & 2-\lambda & -2 \\ 2 & -2 & 4-\lambda \end{bmatrix} \right| = 0$$

פיתוח לפי עמודה ראשונה

$$2-\lambda\left((2-\lambda)\cdot(4-\lambda) - 2\cdot 2\right) - 0\cdot(\dots) + 2\left((0\cdot -2) - (2\cdot(2-\lambda)\right) = 0$$

$$= 2-\lambda\left(8 - 2\lambda - 4\lambda + \lambda^2 - 4\right) + 2\left(-4 + \lambda\right) = 0$$

$$= \underbrace{(2-\lambda)}\left(\lambda^2 - 6\lambda + 4\right) - 4\underbrace{(2-\lambda)} = 0$$

$$= \left(\lambda^2 - 6\lambda\right)\left(2-\lambda\right) = 0$$

$\boxed{\lambda_1 = 2}$

$\lambda^2 - 6\lambda = 0$

$\lambda(\lambda - 6) = 0$

$\boxed{\lambda_3 = 0}$ $\boxed{\lambda_2 = 6}$

הערכים הסינגולריים הם שורשי הע"ע:

$\Rightarrow$

$\sigma_1 = \sqrt{2}$

$\sigma_2 = \sqrt{6}$

$\sigma_3 = 0$

כעת נמצא את הוקטורים העצמיים:

$\lambda_1=2$

$$\begin{bmatrix}2-2 & 0 & 2\\ 0 & 2-2 & -2\\ 2 & -2 & 4-2\end{bmatrix}=\begin{bmatrix}0 & 0 & 2\\ 0 & 0 & -2\\ 2 & -2 & 2\end{bmatrix}\begin{pmatrix}x_1\\x_2\\x_3\end{pmatrix}\Rightarrow\boxed{\begin{pmatrix}1\\1\\0\end{pmatrix}}\cdot\sqrt{2}$$

מתקיים

נורמה

שהוקטור הנ"ל בכפל מימין עם המטריצה שווה לוקטור ה-0.

$\lambda_2=6$

$$\begin{bmatrix}2-6 & 0 & 2\\ 0 & 2-6 & -2\\ 2 & -2 & 4-6\end{bmatrix}=\begin{bmatrix}-4 & 0 & 2\\ 0 & -4 & -2\\ 2 & -2 & -2\end{bmatrix}\begin{pmatrix}x_1\\x_2\\x_3\end{pmatrix}\Rightarrow\boxed{\begin{pmatrix}1\\-1\\2\end{pmatrix}}\cdot\sqrt{6}$$

נורמה

$$-4x_1+2x_3=0\Rightarrow x_3=2x_1$$
$$-4x_2-2x_3=0\Rightarrow x_3=-2x_2$$
$$2x_1-2x_2-2x_3=0\Rightarrow$$

$\lambda_3=0$

$$\begin{bmatrix}2 & 0 & 2\\ 0 & 2 & -2\\ 2 & -2 & 4\end{bmatrix}\begin{pmatrix}x_1\\x_2\\x_3\end{pmatrix}\Rightarrow\boxed{\begin{pmatrix}1\\-1\\-1\end{pmatrix}}\cdot\sqrt{3}\leftarrow$$

נורמה

$$2x_1+2x_3=0\Rightarrow x_3=-x_1$$
$$2x_2-2x_3=0\Rightarrow x_3=x_2$$

V - מטריצת וקטורים עצמיים מנורמלת ולכן:

$$V=\begin{bmatrix}\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{3}}\\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{6}} & -\frac{1}{\sqrt{3}}\\ 0 & \frac{2}{\sqrt{6}} & -\frac{1}{\sqrt{3}}\end{bmatrix}$$

$$\Sigma=\begin{bmatrix}\sqrt{2} & 0 & 0\\ 0 & \sqrt{6} & 0\\ 0 & 0 & 0\end{bmatrix}$$

מטריצה אלכסונית של הערכים הסינגולריים

נשתמש בנוסחה $AV=U\Sigma$ כדי למצוא את U:

$A^{2\times 3} \Rightarrow V^{3\times 3},\ \Sigma^{2\times 3},\ U^{2\times 2}$

Σ^{-1} נכפול מימין $\Rightarrow \quad AV\Sigma^{-1} = U$

$\Downarrow$

$$\begin{bmatrix} 1 & 1 & 0 \\ 1 & -1 & 2 \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{6}} & -\frac{1}{\sqrt{3}} \\ 0 & \frac{2}{\sqrt{6}} & -\frac{1}{\sqrt{3}} \end{bmatrix} \begin{bmatrix} \sqrt{2} & 0 & 0 \\ 0 & \sqrt{6} & 0 \\ 0 & 0 & 0 \end{bmatrix} = U$$

$$\begin{bmatrix} \frac{1}{\sqrt{2}} + \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} - \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{3}} - \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{2}} - \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} + \frac{1}{\sqrt{6}} + \frac{4}{\sqrt{6}} & \frac{1}{\sqrt{3}} + \frac{1}{\sqrt{3}} - \frac{2}{\sqrt{3}} \\ 0 & 0 & 0 \end{bmatrix} \cdot \begin{bmatrix} \sqrt{2} & 0 & 0 \\ 0 & \sqrt{6} & 0 \\ 0 & 0 & 0 \end{bmatrix} = U$$

$\Downarrow$

$$= \begin{bmatrix} \sqrt{2} & 0 & 0 \\ 0 & \sqrt{6} & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{2}} & 0 & 0 \\ 0 & \frac{1}{\sqrt{6}} & 0 \\ 0 & 0 & 0 \end{bmatrix} = \boxed{\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} = U}$$

2) $v \in \mathbb{R}^n,\ u \in \mathbb{R}^m$

$\Downarrow$

מהגדרת מכפלה חיצונית:

$$A = \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} \cdot (u_1, \dots, u_m) = \begin{bmatrix} v_1u_1 & \cdots & v_1u_m \\ \vdots & & \vdots \\ v_nu_1 & \cdots & v_nu_m \end{bmatrix} \Rightarrow [v \otimes u]_{i,j} = v_i \cdot u_j$$

נסתכל על עמודות המטריצה:

$$\begin{pmatrix} v_1u_1 \\ \vdots \\ v_nu_1 \end{pmatrix} \begin{pmatrix} v_1u_2 \\ \vdots \\ v_nu_2 \end{pmatrix} \dots$$

נכתוב את הוקטורים של העמודות כמכפלה סקלרית:

$$u_1\begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} \quad u_2\begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} \quad u_3\begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}$$

כלומר עמודות המטריצה הן כפל בסקלר שונה כל פעם של אותו וקטור:

$\begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}$, כלומר וקטורי העמודות תלויים לינארית (לפי ההגדרה של תלות לינארית)

ולכן $rank(A)=1$ כנדרש.

3) $x=\sum_{i=1}^{n} a_i u_i \in \mathbb{R}_n$ $\quad a_i\in\mathbb{R}, \ (u_1,\dots,u_n)\in \mathbb{R}_n$

$$\langle x,u_j\rangle = \langle \sum_{i=1}^{n} a_i u_i , u_j\rangle \overset{\text{מלינאריות מכפלה פנימית}}{=} \sum_{i=1}^{n} a_i\langle u_i,u_j\rangle$$

כעת, נשים לב שהבסיס היא אורתונורמלי ולכן מתקיים: $i\neq j$ $\langle u_i,u_j\rangle=0$

לכל $i,j\in n$ $\quad i=j$ $\langle u_i,u_j\rangle=1$

ולכן הסכום $\sum_{i=1}^{n} a_i\langle u_i,u_j\rangle$ היא לא 0 רק כאשר $i=j$ ולכן:

$\boxed{a_j=\langle x,u_j\rangle}$ כלומר $\langle x,u_j\rangle = a_j\underbrace{\langle u_j,u_j\rangle}_{=1} = a_j$

כנדרש ☒

4) a) $x=\begin{pmatrix} 3 \\ -4 \\ 1 \\ -2 \end{pmatrix}$,

- $||x||_1 = 3+4+1+2=10$
- $||x||_2=\sqrt{3^2+4^2+1^2+2^2}=\sqrt{9+16+1+4}=\sqrt{30}$

$\cdot \|x\|_\infty = y$

נורמת המקסימום החזירה את הערך הגבוה ביותר, ומאחר וכל ערכי x גבולים שווים מ-1 ולכן שתי הנורמות האחרות בוודאות יחזירו ערך שגדול מ-y בעוד נורמת המקסימום מסיימה בערכים בין 1 ל-y.

b) ℓ_1:

$\|x\|_1 = 1$

$\Downarrow$

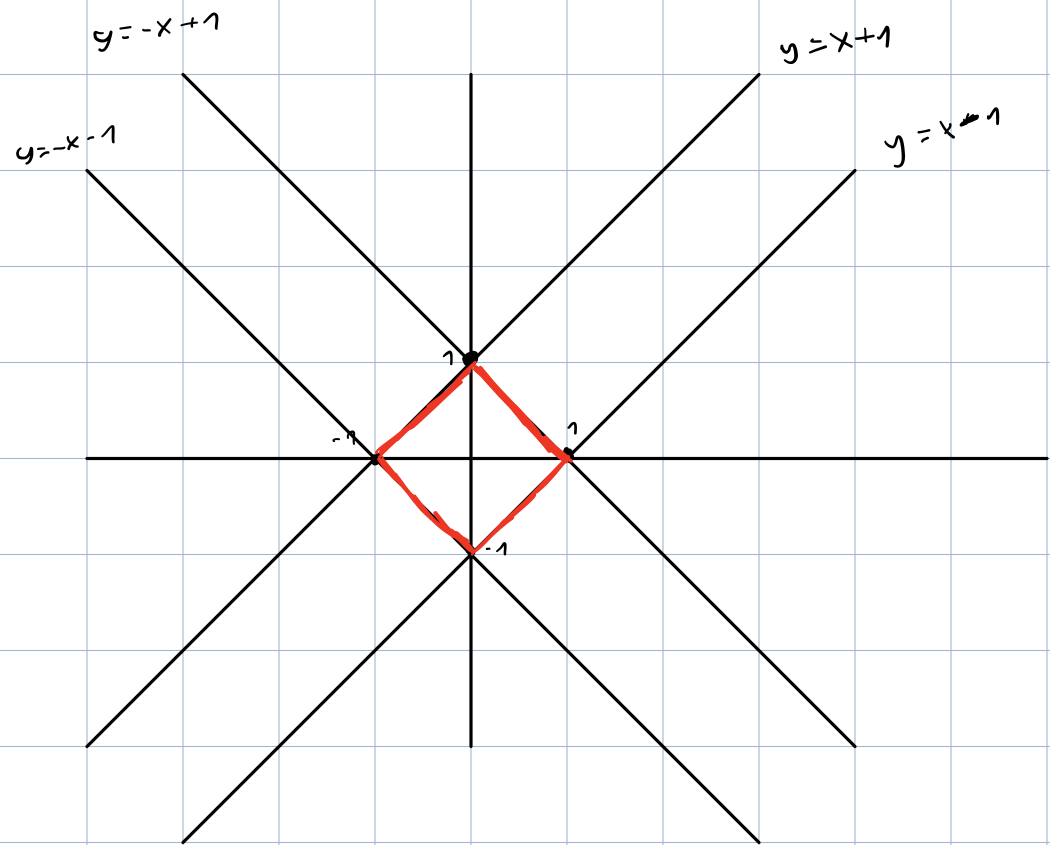
$x=(x,y) \Rightarrow |x|+|y|=1$

$x+y=1 \Rightarrow y=1-x$

$-x+y=1 \Rightarrow y=1+x$

$x-y=1 \Rightarrow y=x-1$

$-x-y=1 \Rightarrow y=-x-1$



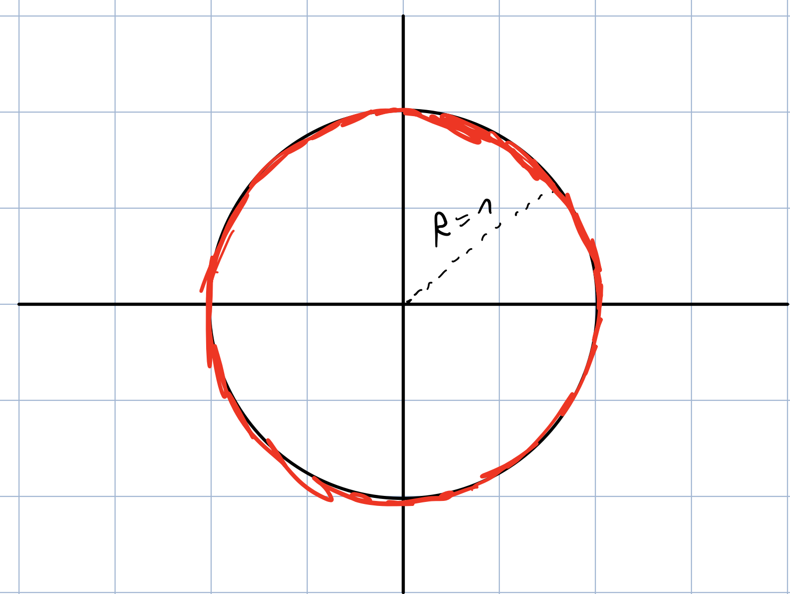
ℓ_2:

$\|x\|_2 = 1$

$\Downarrow$

$\sqrt{x^2+y^2} = 1$

נוסחת מרחק מריבועית הקירים ולכן נקבל מעגל



ℓ_∞:

$\|x\|_\infty = 1$

$\Updownarrow$

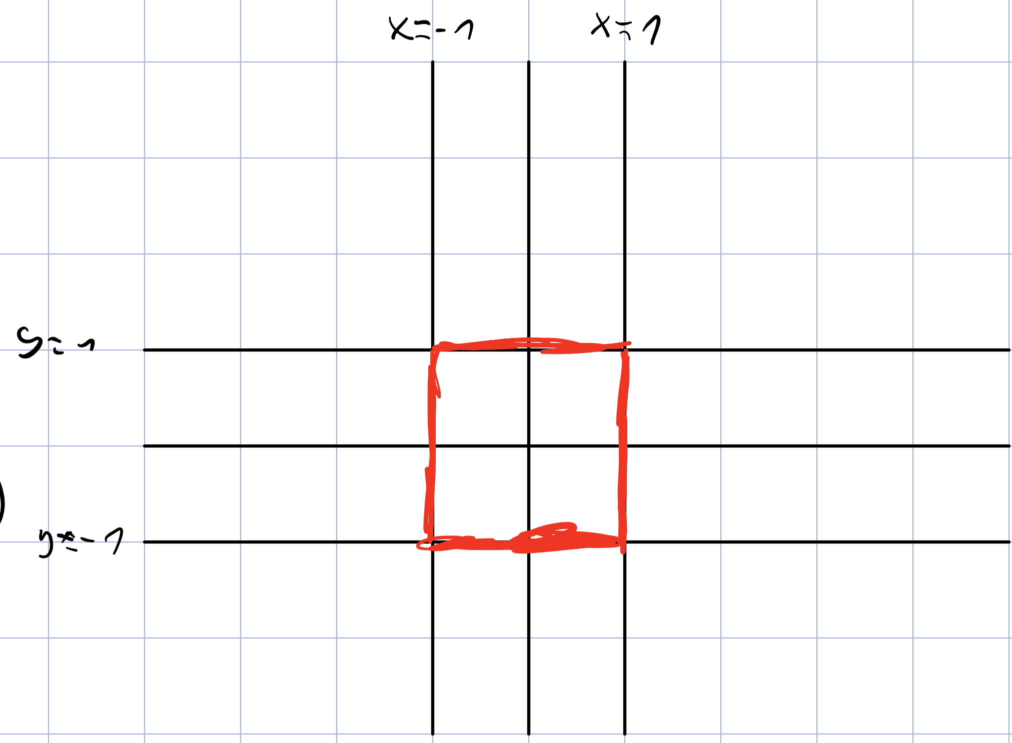
$\max\{x, y\} = 1$

כלומר:

$x = \pm 1,\ -1 < y < 1$

או

$y = \pm 1,\ -1 < x < 1$



השרטוטים בכלל משקפים איך כל נורמה מודדת "אורך" בצורה שונה:

ℓ_1: מעוין: ישנה משמעות גדולה יותר לגדילה בשני הצירים במקביל:
כלומר בצורת אלכסון → כלומר האורך מושפע הכי הרבה משינוי בכל הצירים במקביל

ℓ_2: עיגול: בכל גידול האורך משתנה באותו אופן ולכן גם מודד מרחק מוחלט.

ℓ_∞: ריבוע: האורך מושפע הכי הרבה משינוי בציר (הקואורדינטה) הגדולה ביותר.

1.1.2 - Multivariate Calculus

b) $h(\sigma) = \frac{1}{2}\|f(\sigma) - y\|^2$, $\sigma \in \mathbb{R}^d$, $f : \mathbb{R}^d \to \mathbb{R}^d$

חוק השרשרת:

$$\nabla h(\sigma) = J_f(\sigma)^T \cdot \nabla h(f(\sigma))$$

$$\Downarrow$$

ראינו בתרגול ש: $\nabla_x \|g(x)\|^2 = 2g(x)^T \cdot \nabla_x g(x)$

ולכן:

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

$$= \boxed{J_f(\sigma)^T \cdot (f(\sigma) - y)}$$

c) הגדרה: $S(x)_j = \frac{e^{x_j}}{\sum_{\ell=1}^{k} e^{x_\ell}}$ $\quad S : \mathbb{R}^k \to [0,1]^k$

$$J_S \overset{\text{הגדרה}}{=} \begin{bmatrix} \frac{\partial s_1(x)}{\partial x_1} & \cdots & \frac{\partial s_1(x)}{\partial x_k} \\ \vdots & & \vdots \\ \frac{\partial s_k(x)}{\partial x_1} & \cdots & \frac{\partial s_k(x)}{\partial x_k} \end{bmatrix}$$

בתרגול ראינו שהנגזרת של כל קואורדינטה בה $j=i$, כלומר על האלכסון היא:

$$\frac{\partial s_i}{\partial x_j} = \frac{d}{dx_j}\left(\frac{e^{x_j}}{\sum_k e^{x_k}}\right) \overset{\text{כלל החילוק}}{=} \frac{e^{x_i}\sum e^{x_k} - e^{x_i} \cdot e^{x_j}}{\left(\sum_k e^{x_k}\right)^2} = \frac{e^{x_i}}{\sum e^{x_k}} \cdot \frac{\sum e^{x_k} - e^{x_j}}{\sum e^{x_k}} = S_i(1 - S_i)$$

נסתכל על המקרים בהם $i \neq j$

כלל החילוק

גוזרים לפי משתנה אחר

$$\frac{\partial s_i}{\partial x_j} = \frac{d}{dx_j}\left(\frac{e^{x_i}}{\sum_k e^{x_k}}\right) \overset{\downarrow}{=} \frac{0 \cdot \sum_k e^{x_k} - e^{x_j} \cdot e^{x_i}}{\left(\sum_k e^{x_k}\right)^2} = \frac{-e^{x_j} \cdot e^{x_i}}{\left(\sum_k e^{x_k}\right)^2}$$

$$= \frac{-e^{x_j}}{\sum_k e^{x_k}} \cdot \frac{e^{x_i}}{\sum_k e^{x_k}} = -s_j \cdot s_i$$

ולכן היעקוביאן נראה כך:

$$J_s(x) = \begin{bmatrix} s_1(1-s_1) & -s_1 \cdot s_2 & \cdots & -s_1 s_k \\ -s_2 s_1 & s_2(1-s_2) & & -s_2 s_k \\ \vdots & & \ddots & \vdots \\ -s_k s_1 & \cdots & \cdots & s_k(1-s_k) \end{bmatrix}$$

1.2: Linear Regression:

$X \in R^{m \times d}$, $y \in R^m$

מטריצת הדגימות עם הפיצ'רים — וקטור התוצאות

5) a) $\Rightarrow$

נכפיל משמאל ב X^T

$$v \in \ker(X) \Rightarrow X \cdot v = 0 \Rightarrow X^T \cdot X \cdot v = X^T \cdot 0 = 0 \Rightarrow v \in \ker(X^T X)$$

$\Leftarrow$

נכפול משמאל ב v^T

$$v \in \ker(X^T X) \Rightarrow X^T X \cdot v = 0 \Rightarrow v^T X^T X v = v^T \cdot 0 \Rightarrow \|Xv\|^2 = 0$$

תכונות נורמה

$$\Rightarrow Xv = 0 \Rightarrow v \in \ker(X)$$

$\blacksquare$

b) $A \in \mathbb{R}^{m \times m}$

קיים
$w \in \mathbb{R}^m$

$v \in Im(A^T) \Rightarrow A^T w = v$

$ker(A)^\perp = \{ v \mid v^T w = 0, \forall w \in ker(A) \}$, $Aw = 0$

$\Downarrow$

$w^T A^T = 0$, כלומר כל וקטור ב $ker(A)$ אורתוגונלי לכל עמודה ב A^T

כלומר מימדי A^T ו $ker(A)^\perp$ שווים או במילים אחרות $Im(A^T) = ker(A)^\perp$

c) X עמודות אך בלתי הפיכה כלומר קיים $v \neq 0$ כך ש:

$$X v = 0$$

$$X w = y \iff y \in Im(X^T) \underset{\text{סעיף ב'}}{\iff} y \in ker(X)^\perp \iff y \perp ker(X^T)$$

תנאי זה מקיים פתרון יחיד רק אם X הפיכה ומאחר והיא לא ישנם אינסוף פתרונות.

d) $X^T X w = X^T y$

לפי סעיף a: $ker(X^T X) = ker(X)$, נניח ש $X^T X$ הפיכה אז מתקיים $ker(X) = \{0\}$ כלומר ישנו פתרון יחיד ל X.

נניח ש $X^T x$ לא הפיכה אז מימד הגרעין של X גדול מ ר (ואם מסעיף C) כלומר והיו אינסוף פתרונות שיפתרו את המשוואה הנורמלית.

1.2.2:

1) $x^{+}y \overset{\text{הגדרה}}{=} V\Sigma^{+}U^{T}y \quad (*)$

2) $x^{T}x = (U\Sigma V^{T})^{T}(U\Sigma V^{T}) = V\Sigma^{T}U^{T}U\Sigma V^{T}$

$= V\Sigma^{2}V^{T} \Rightarrow [x^{T}x]^{-1} = V(\Sigma^{2})^{-1}V^{T}$

U אורתוגונלית $\iff U^{T}U = I$

Σ אלכסונית $\iff \Sigma^{T} = \Sigma$

V אורתוגונלית

$\Downarrow$

$[x^{T}x]^{-1}x^{T}y = V(\Sigma^{2})^{-1}V^{T}x^{T}y = V(\Sigma^{2})^{-1}V^{T}V\Sigma U^{T}y$

$= V(\Sigma^{2})^{-1}\Sigma U^{T}y \overset{\text{הגדרת }+}{=} V\Sigma^{+}U^{T}y \overset{(*)}{=} x^{+}y$

V אורתוגונלית $\iff V^{+}V = I$

$\square$

Practical part answers:

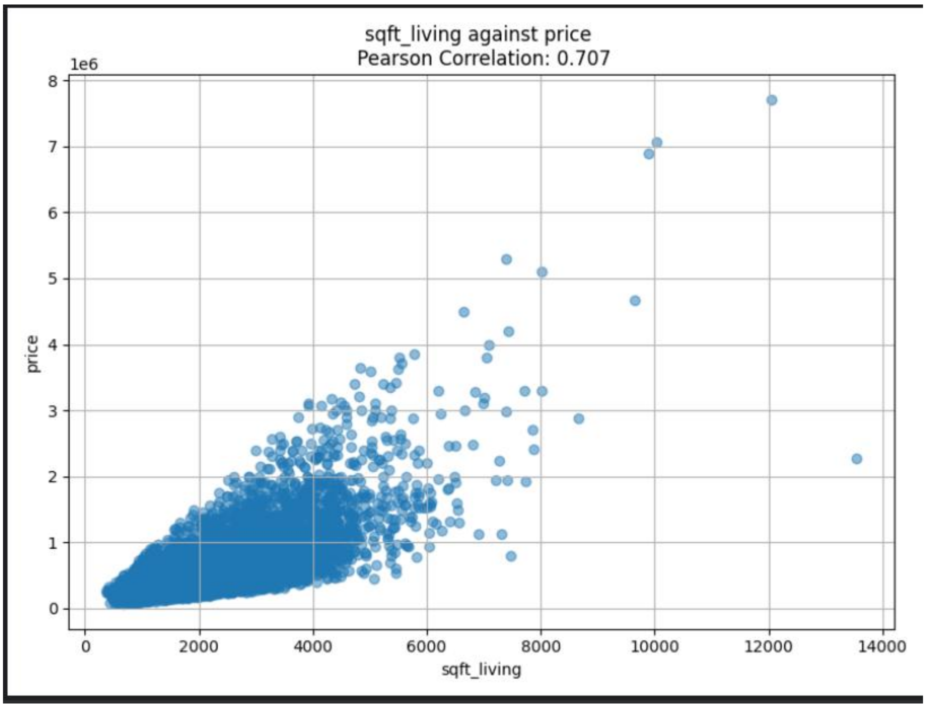
Linear Regression:

3)א) הפיצ'רים שבחרתי להוריד הם ID, DATE, YR_BUILT, YR_RENOVATED מאחר ID הוא שדה שרירותי ללא כל משמעות מתמטית על מחיר הבית וכנ"ל לגבי .DATE לגבי YR_RENOVATED, YR_BUILT הורדתי אותם כי הוספתי פיצ'רים במקומם שמחשבים את ההפרש מהשנה הנוכחית לשנה הנתונה מאחר ולדעתי אלו נתונים שיותר נוח להסתכל עליהם ולהסיק מסקנות(למשל בתים באזור גיל 30 שווים פחות שזה נתון שלא קופץ לעין בכזאת קלות לפי שנת בנייה) מאשר השנה הספציפית שבה הבית נבנה \ שופץ.

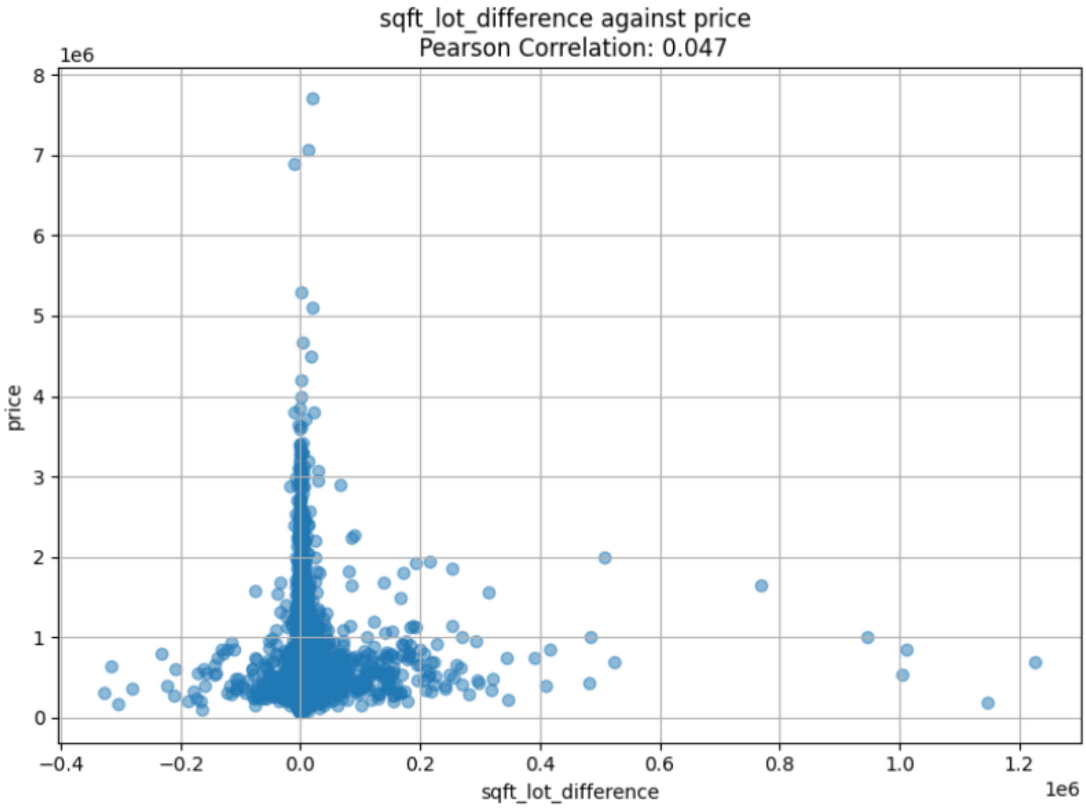
ב) הוספתי חמישה פיצ'רים חדשים: **גיל בית**, **שנים מאז שיפוץ** – נתונים יותר נוחים לניתוח לדעתי מאשר הפיצ'רים המקוריים המקבילים להם. בנוסף הוספתי **הפרש בין גודל מגורים לממוצע 15 קרובים, הפרש בין גודל מגרש לממוצע 15 קרובים.** אלה נתונים שלדעתי יתנו זווית יותר עניינית על הנתונים- איך ההפרש מהגודל הממוצע של 15 השכנים הקרובים משפיע על המחיר.

ג) לגבי נתונים INVALID/MISSING פשוט מחקתי את כל השורה מהנתונים כי לא ניתן לסמוך עליה לדעתי.

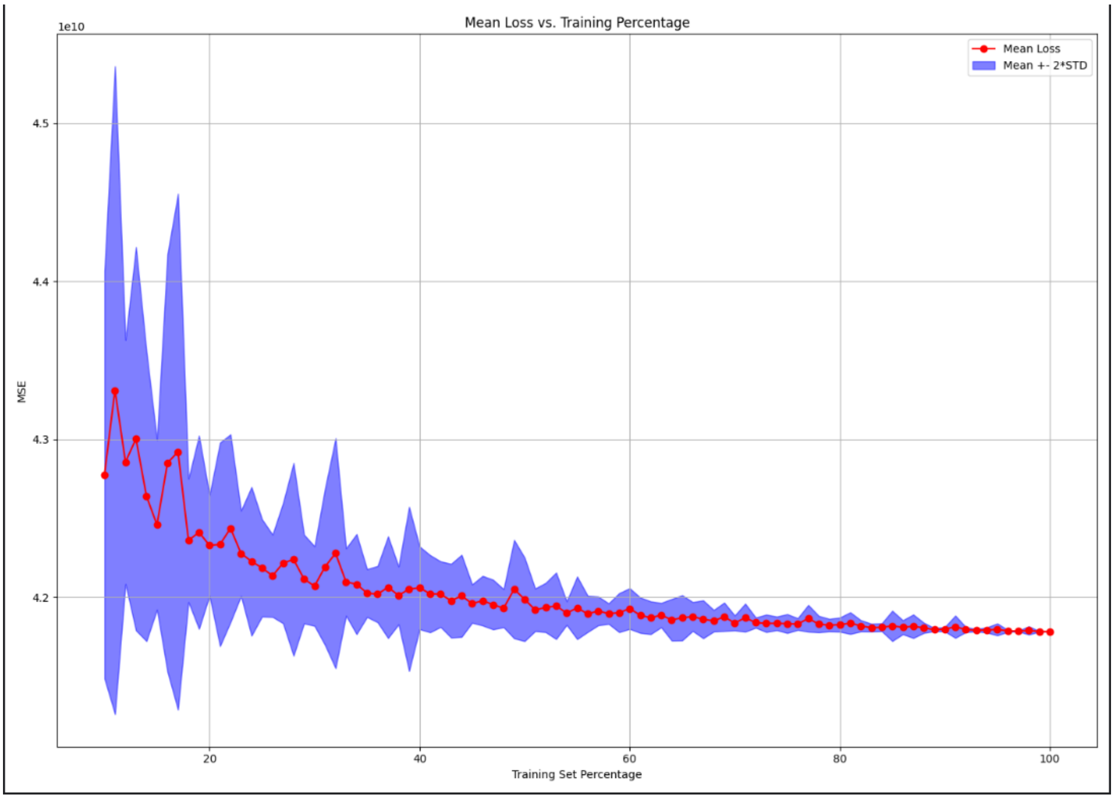
4) גרף 1- שטח אזור בנוי למגורים, ניתן לראות שפיצ'ר זה מאוד רלוונטי למחיר, ציון קורלציית הפירסון שלו הוא 0.7 מה שמעיד על קשר לינארי גבוה וניתן לראות גם בגרף את מגמת העלייה במחיר בהתאם לעליית בשטח הבנוי למגורים. כלומר פיצ'ר זה מאוד רלוונטי למודל.



גרף 2- הפרש בין שטח הבית הכולל לבין שטח הבית הכולל של 15 השכנים הקרובים ביותר אליו. זהו אחד הפיצ'רים שאני הוספתי ולהפתעתי ניתן לראות שהקשר הלינארי מאוד נמוך כלומר אין השפעה של אמיתית על מחיר הבית של ההפרש בין הנתונים הללו. ניתן לראות זאת כי ציון הפירסון קרוב מאוד ל0 וגם בגרף עצמו ניתן לראות שאין קשר לינארי אמיתי ולכן כנראה שאנשים לא מאוד מושפעים מהשוואת שטח ביתם לשטח של בית שכניהם עד כדי עליית מחירים.



6)

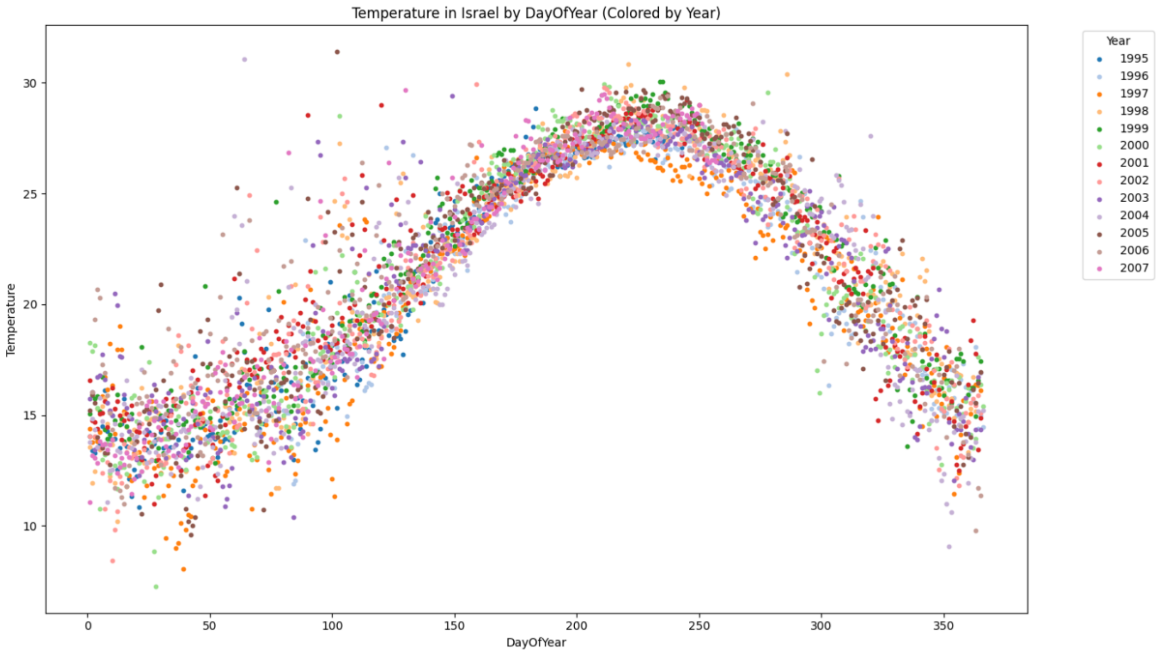


ניתן לראות שMSE נע בין 4.5 ל4.2 le10 ומאחר והוא מחושב בצורה ריבועית הloss הוא בעצם ב אזור 200,000 שזה נשמע טווח סביר בהתאם למחירי בתים(מליונים).

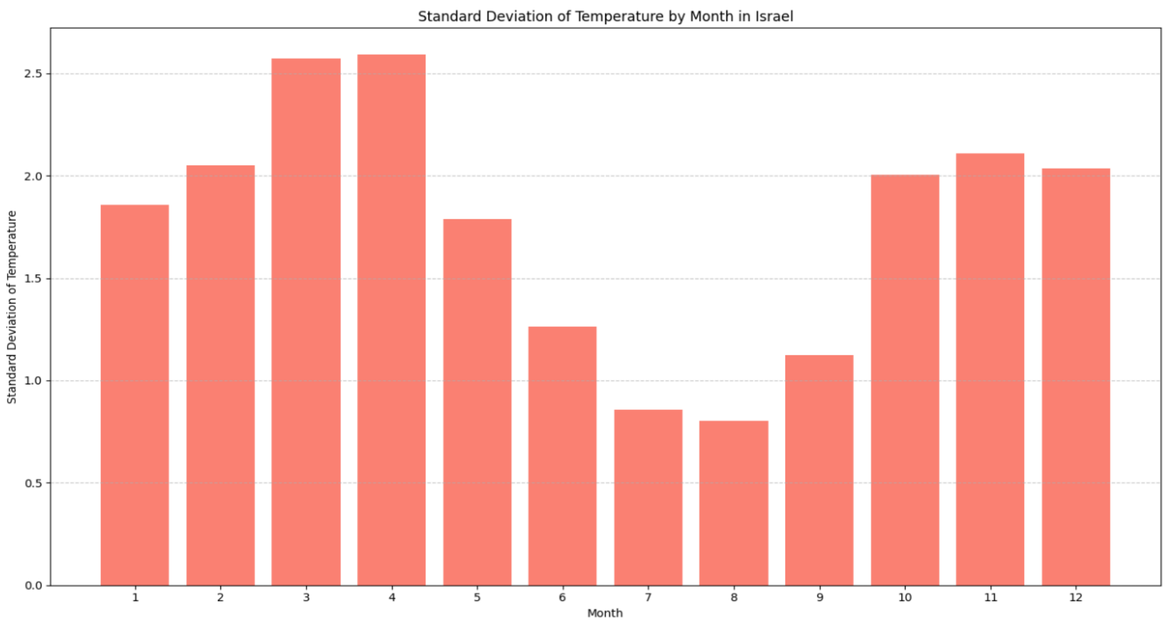
בנוסף ניתן לראות שיש קורלציה בין אחוזי הדגימה לירידה בMSE הכללי, כלומר המודל משתפר ככל שהוא מקבל דגימה יותר גדולה מהנתונים, בהתחלה השינויים חדים אך מאזור ה50 אחוז דגימה הMSE כמעט ולא משתנה כלומר אין לכמות הנתונים שגדלה משמעות בהקטנת הLOSS. לגבי ה2- +Confidence interval of mean(loss) ניתן לראות שבהתחלה הוא בטווח רחב מאוד מהMSE עצמו אך ככל שאחוז הדגימה עולה הטווח שלו מצטמצם והוא נהיה כמעט זהה לMSE האמיתי כלומר טווח הLOSS של דגימה מתקרב לMSE הממוצע שקיבלנו ככל שאחוז הדגימות עולה, זה מעיד על מודל שמשתפר ככל שהוא לומד יותר.

Polynomial Fitting:

(3

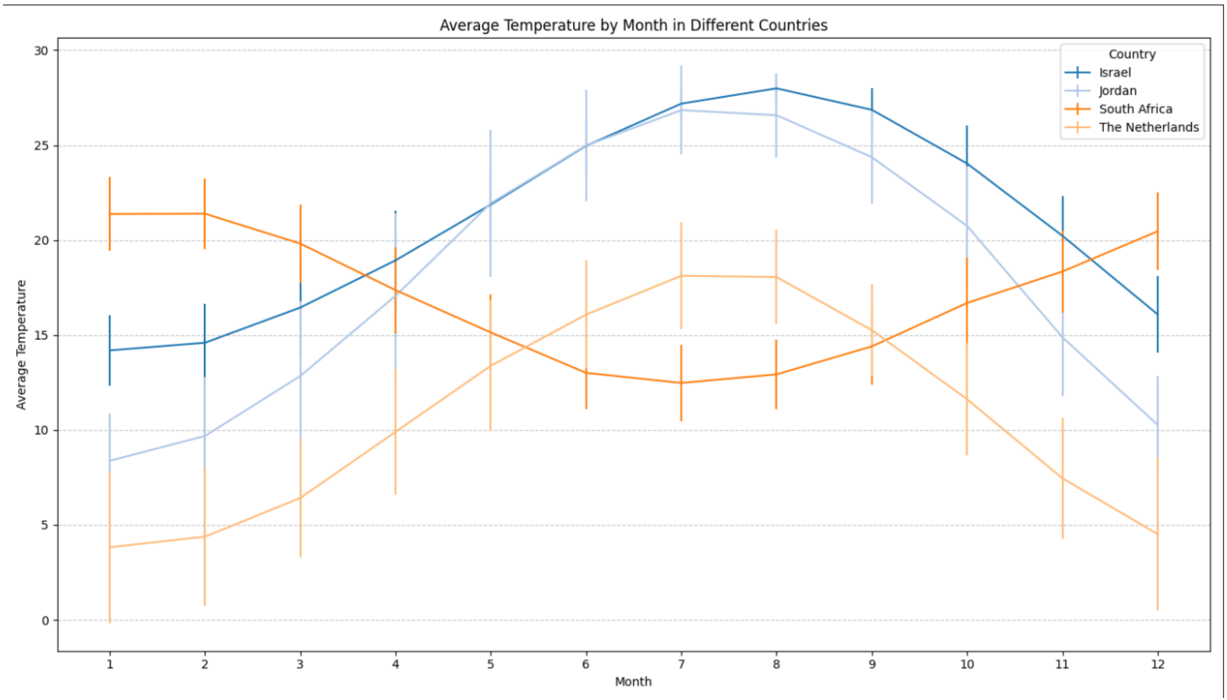


לפי צורת הגרף ניתן לראות שהגרף די מתאים לכל פולינום מדרגה אי זוגית בשל צורתו ולכן אשער שהדרגה המתאימה ביותר היא 3 או 5.



בהתאם לגרף נשים לב שהשונות של טמפרטורה יומית באותו חודש משתנה בין החודשים ולכן נשער שמודל החיזוי יעבוד בצורה הטובה ביותר על חודשים בהם השונות הזאת היא הקטנה ביותר(יולי, אוגוסט) לעומת חודשים בהם השונות היא גדולה (מרץ אפריל) ולכן המודל יצליח במידה שונה לכל חודש.

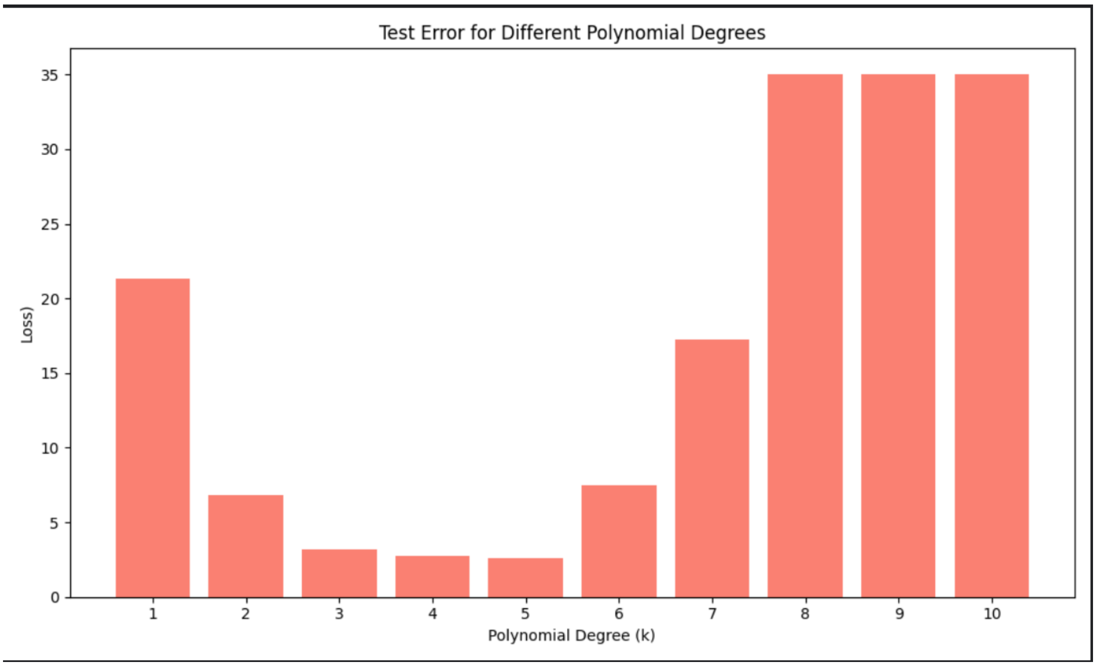
(4



בהתאם לגרף ניתן לראות שישראל ירדן והולנד נצמדות לאותה תבנית אך הולנד עם טמפרטורות כלליות יותר נמוכות לאורך כל השנה. לעומתן דרום אפריקה נצמדת לתבנית הפוכה לגמרי דבר שמסתדר עם העונות ההפוכות בשל הימצאותן בחצי כדור הארץ הדרומי. מפאת תוצאות אלה נסיק שהמודל שיתאים לישראל יתאים כנראה בצורה המקסימלית גם לירדן בשל הדמיון המקסימלי בין הגרפים של שתי המדינות.

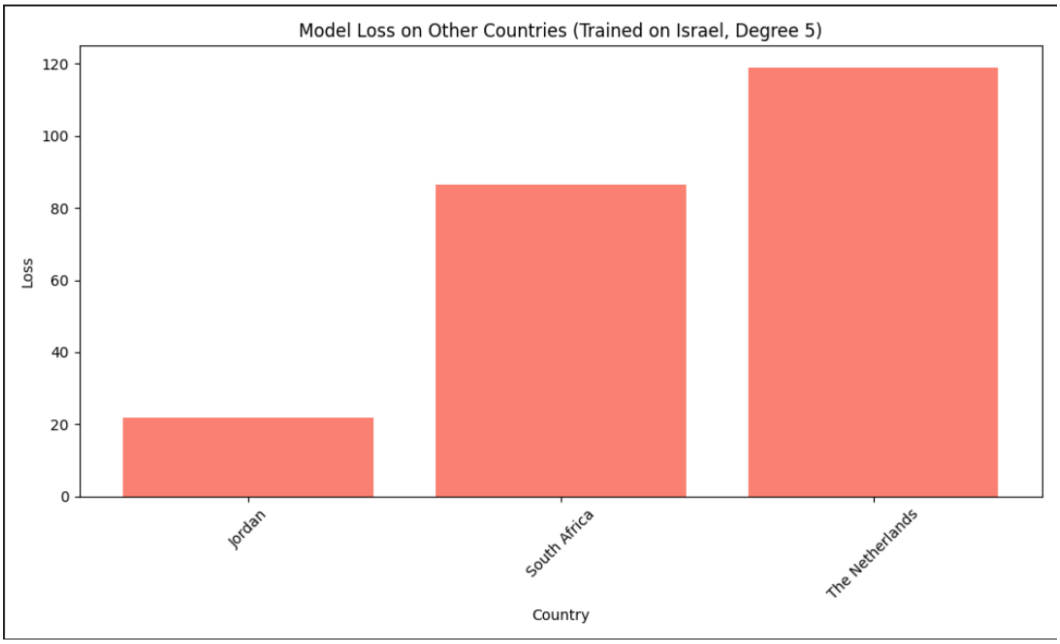
(5

```
k=1, loss=21.320242021509923
k=2, loss=6.825831621422252
k=3, loss=3.171923340889096
k=4, loss=2.7135381423401213
k=5, loss=2.619784398159097
k=6, loss=7.454277989377751
k=7, loss=17.273324151668007
k=8, loss=35.00129019730737
k=9, loss=35.000859274784695
k=10, loss=35.00079890448881
```



ניתן לראות לפי התוצאות שהערך המינימלי של הLOSS מגיע כאשר K=5 כלומר פולינום מדרגה 5 הוא אופטימלי למזער את הLOSS ניתן בנוסף לראות שגם 3 ו4 נותנים ערכים טובים שאולי שווה ליצור מהם מודל פולינומיאלי אך לא כמו 5.

(6



ניתן לראות שבהתאם להשערתנו בשאלה 4 הLOSS הנמוך ביותר לK האופטימלי(5) ירדן היא המדינה שהמודל שפיתחנו הכי מתאים לה כי בה ערך הLOSS הנמוך ביותר שזו השאיפה במודל. מה שכן מפתיע הוא שדרום אפריקה במקום השני על אף שבשאלה 4 צפינו שהולנד תהיה במקום השני כנראה בשל השוני הגדול בטמפרטורות הממוצעות.