207807082 יאיר [illegible]

EX1

linear algebra - 1 שאלה

1 נחשב את ה-SVD של A^TA

$$A^TA = \begin{pmatrix}1&1\\1&-1\\0&2\end{pmatrix}\begin{pmatrix}1&1&0\\1&-1&2\end{pmatrix} = \begin{pmatrix}2&0&2\\0&2&-2\\2&-2&4\end{pmatrix}$$

נמצא את הע"ע

$$\det\begin{pmatrix}x-2&0&-2\\0&x-2&2\\-2&2&x-4\end{pmatrix} = (x-2)\det\begin{pmatrix}x-2&2\\2&x-4\end{pmatrix} - 2\det\begin{pmatrix}0&-2\\x-2&2\end{pmatrix} =$$

$$= (x-2)\big((x-2)(x-4)-4\big) - 4(x-2) = x^3 - 8x^2 + 12x = x(x-6)(x-2)$$

לכן $\lambda_1=0$ $\lambda_2=2$ $\lambda_3=6$

נמצא את הו"ע לכל ע"ע

עבור λ_1: $\ker(A^TA) = \ker\begin{pmatrix}2&0&2\\0&2&-2\\2&-2&4\end{pmatrix}$

$$\begin{pmatrix}2&0&2\\0&2&-2\\2&-2&4\end{pmatrix} \to \begin{pmatrix}1&0&1\\0&2&-2\\0&-2&2\end{pmatrix} \to \begin{pmatrix}1&0&1\\0&1&-1\\0&0&0\end{pmatrix}$$

לכן $V_0 = \ker\begin{pmatrix}2&0&2\\0&2&-2\\2&-2&4\end{pmatrix} = span\left\{\begin{pmatrix}-1\\1\\1\end{pmatrix}\right\}$

עבור λ_2: $\ker\begin{pmatrix}0&0&2\\0&0&-2\\2&-2&2\end{pmatrix}$

$$\begin{pmatrix}0&0&2\\0&0&-2\\2&-2&2\end{pmatrix} \to \begin{pmatrix}1&-1&1\\0&0&1\\0&0&0\end{pmatrix}$$

לכן $V_2 = span\left\{\begin{pmatrix}1\\1\\0\end{pmatrix}\right\}$

עבור λ_6: $\ker\begin{pmatrix}-4&0&2\\0&-4&-2\\2&-2&-2\end{pmatrix}$

$$\begin{pmatrix}-4&0&2\\0&-4&-2\\2&-2&-2\end{pmatrix} \to \begin{pmatrix}1&0&-0.5\\0&1&0.5\\0&-1&-0.5\end{pmatrix} \to \begin{pmatrix}1&0&-0.5\\0&1&0.5\\0&0&0\end{pmatrix}$$

$$V_6 = \ker\begin{pmatrix}-4&0&2\\0&-4&-2\\2&-2&-2\end{pmatrix} = span\left\{\begin{pmatrix}1\\-1\\2\end{pmatrix}\right\}$$

כעת ננרמל את הוקטורים שקיבלנו ונכניס ל-V ונקבל

$$V = \begin{pmatrix}\frac{1}{\sqrt6}&\frac{1}{\sqrt2}&-\frac{1}{\sqrt3}\\-\frac{1}{\sqrt6}&\frac{1}{\sqrt2}&\frac{1}{\sqrt3}\\\frac{2}{\sqrt6}&0&\frac{1}{\sqrt3}\end{pmatrix}$$

וכן $\Sigma^T\Sigma = \begin{pmatrix}6&0&0\\0&2&0\\0&0&0\end{pmatrix}$ ומכאן $\Sigma = \begin{pmatrix}\sqrt6&0&0\\0&\sqrt2&0\end{pmatrix}$

כעת נעבור ל-evd של AA^T

$$AA^T = \begin{pmatrix}1&1&0\\1&-1&2\end{pmatrix}\begin{pmatrix}1&1\\1&-1\\0&2\end{pmatrix} = \begin{pmatrix}2&0\\0&6\end{pmatrix}$$

כבר מטריצה אלכסונית ולכן הוקטורים העצמיים שלה.

נסדר את הערכים כך ש $U = \begin{pmatrix}6&0\\0&2\end{pmatrix}$ לאחר נרמול נקבל $U = \begin{pmatrix}0&1\\1&0\end{pmatrix}$

ולכן

$$A = \begin{pmatrix}0&1\\1&0\end{pmatrix}\begin{pmatrix}\sqrt6&0&0\\0&\sqrt2&0\end{pmatrix}\begin{pmatrix}\frac{1}{\sqrt6}&\frac{1}{\sqrt2}&-\frac{1}{\sqrt3}\\-\frac{1}{\sqrt6}&\frac{1}{\sqrt2}&\frac{1}{\sqrt3}\\\frac{2}{\sqrt6}&0&\frac{1}{\sqrt3}\end{pmatrix}$$

ב) א' $v\in\mathbb{R}^n$ $u\in\mathbb{R}^m$ אזי המטריצה המוגדרת באופן הבא מדרגה 1

$$A = v\otimes u = \begin{pmatrix}v_1u_1&v_1u_2&\cdots&v_1u_m\\\vdots&\vdots&&\vdots\\v_nu_1&v_nu_2&\cdots&v_nu_m\end{pmatrix}$$

יהיו שני עמודות z_i, z_j לא תלויות לינארית הנ"ל

אך ניתן לראות כי $z_i = v\cdot u_i$ וכן $z_j = v\cdot u_j$ העמודות תלויות אחת בשניה.

לכן כל העמודות תלויות אחת בשניה ולכן הדרגה של A היא 1 $rank(A)=1$

ⓑ כעין התרגיל נחשב את המכפלה הפנימית בין x ל-u_j

$$\langle x,u_j\rangle=\left\langle\sum_{i=1}^{n}a_iu_i,u_j\right\rangle\overset{\text{לינאריות}}{=}\sum_{i=1}^{n}a_i\langle u_i,u_j\rangle=a_j\langle u_j,u_j\rangle\overset{\text{נורמה של כל וקטור שווה 1}}{=}a_j$$

עבור כל $i\neq j$ המכפלה הפנימית היא 0 כי יוצרים בסיס אורתו'

ⓐ נשים לב כי $P=\sum_{i=1}^{k}v_iv_i^T=\sum_{i=1}^{k}v_i\otimes v_i$

כמו כן $v_i\otimes v_i$ היא סימטרית כי עבור כל תא ij ו-ji נקבל בדיוק אותו ערך. לכן נקבל סכום של מטריצות סימטריות שהיא סימטרית (הוכחה: עבור כל תא נקבל סכום של ערכים זהים)

ⓑ עבור כל $v_i\in\{v_1,\dots,v_k\}$

$$Pv_i=\sum_{j=1}^{k}v_jv_j^Tv_i\overset{\substack{\text{כי עבור כל } j\neq i\\ \langle v_j,v_i\rangle=0}}{=}v_iv_i^Tv_i=v_i$$

לכן עבור כל v_i כזה הוא וקטור עצמי עם ע״ע 1

כעת יהי x אשר אורתוגונלי לבסיס. לכן אפשר לבטא אותו להיות $x=\sum_{i=k+1}^{n}v_i$

אזי

$$Px=\sum_{j=1}^{k}v_jv_j^T\cdot\sum_{i=k+1}^{n}v_i=$$

כיוון שהוקטורים של בסיס אורתונורמלי אז המכפלה הפנימית שלהם תהיה 0 ולכן

$$=\sum_{j=1}^{k}v_j\sum_{i=k+1}^{n}v_j^Tv_i=\sum_{j=1}^{k}v_j\cdot 0=0$$

לכן 0 הוא ע״ע

ⓒ יהי $v\in V$ אזי $v=\sum_{i=1}^{k}a_iv_i$

$$PV=\sum_{i=1}^{k}v_iv_i^T\sum_{j=1}^{k}a_jv_j\overset{\text{לינאריות}}{=}\sum_{i=1}^{k}v_i\sum_{j=1}^{k}a_jv_i^Tv_j\overset{\text{מכפלה פנימית}}{=}\sum_{i=1}^{k}v_i\sum_{j=1}^{k}a_j\langle v_i,v_j\rangle\overset{\text{עבור כל } i\neq j \text{ המכפלה הפנימית היא 0}}{=}\sum_{i=1}^{k}v_i\cdot a_i\cdot v_i^T\cdot v_i=\sum_{i=1}^{k}v_ia_i=V$$

ⓓ

$$P^2=\sum_{i=1}^{k}v_iv_i^T\cdot\sum_{j=1}^{k}v_jv_j^T=\sum_{i=1}^{k}\sum_{j=1}^{k}v_iv_i^Tv_jv_j^T\overset{*}{=}\sum_{i=1}^{k}v_iv_i^Tv_iv_i^T\underset{v_iv_i^T=1}{=}\sum_{i=1}^{k}v_iv_i^T=P$$

* כי $i=j$ אזי $v_i^Tv_j=1$ אחרת שווה ל-0 ולכן

ⓔ הראינו ש-P [illegible], כי $P^2=P \Leftarrow P-P^2=0 \Leftarrow P(I-P)=0$ ■

multivariate calculus שאלה 2

(5)

$$h(\sigma) = g(f(\sigma)) \Rightarrow h'(\sigma) = g'(f(\sigma)) \cdot f'(\sigma) \qquad g(x) = \frac{1}{2}\|x-y\|^2$$

ונחשב את הנגזרת של g לכניסה x_i

$$\frac{\partial g(x)}{\partial x_j} = \frac{\partial}{\partial x_j}\sum_{i\in[d]}(x_i-y)^2 = \sum_{i\in[d]}\frac{\partial (x_i-y)^2}{\partial x_j} = \sum_{i\in[d]}\frac{\partial(x_i^2-2x_iy+y^2)}{\partial x_j} = 2x_j-2y$$

עבור x_i^2 נקבל $2x_i$
ועבור $-2x_iy$ נקבל $-2y$

לכן

$$Dh(\sigma) = \frac{1}{2}Dg(f(\sigma)-y)\cdot Df(\sigma) = \frac{1}{2}(2f(\sigma)-y)\cdot Df(\sigma) = (f(\sigma)-y)\cdot Df(\sigma)$$

לכן

$$\nabla h(\sigma) = Df(\sigma)^T(f(\sigma)-y)$$

(6)

$$S_j(x) = \frac{e^{x_j}}{\sum_{i=1}^{k} e^{x_i}}$$

נחלק את הנגזרת של $S_j(x)$ עבור i למקרים?

$$\frac{\partial S_j}{\partial x_i} = \frac{e^{x_j}\sum_{l=1}^{k}e^{x_l} - e^{2x_j}}{\left(\sum_{l=1}^{k}e^{x_l}\right)^2}$$ במקרה $i=j$ נקבל

$$\frac{\partial S_j}{\partial x_i} = \frac{-e^{x_i}e^{x_j}}{\left(\sum_{l=1}^{k}e^{x_l}\right)^2}$$ במקרה $i\neq j$ נקבל

לכן נוכל לבנות את המטריצה $[J(s)]_{ij}$ כאשר עבור כל i,j נבחר בין הנגזרות.

7) (א) נוכיח הכלה דו כיוונית:

יהי $v\in\ker(X^TX)$:

$$0=X^TX(v)\Rightarrow v^TX^TXv=0\Rightarrow (vX)^TXv=0\Rightarrow \|Xv\|=0\Rightarrow Xv=0$$

יהי $v\in\ker(X)$:

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

(ב) יהי $v\in Im(A^T)$

נראה $\forall u\in\ker(A)$ מתקיים $\langle v,u\rangle=0$

קיים w כך ש $A^Tw=v$ ומתקיים $Au=0$ ולכן

$$\langle v,u\rangle=\langle A^Tw,u\rangle=\langle w,Au\rangle=\langle w,0\rangle=0$$

לכן $Im(A^T)\subseteq\ker(A)^\perp$

לכן נשאר להראות שהממדים שלהם שווים

מתקיים ממשפט הממדים $n=\dim(\ker(A)^\perp)+\dim(\ker(A))$ וגם $n-\dim\ker(A)=rk(A)$

לכן $\dim(\ker(A)^\perp)=\dim(Im(A))=rk(A)$

(ג) מתקיים מסעיף ב $\ker(X^T)^\perp=Im(X^{T^T})=Im(X)$

לכן $y\in Im X$ אמ"מ $y\perp\ker(X^T)$ לכן יש פתרון למשוואה של $Xw=y$

X אינו הפיך או X אינו ריבועי לכן $\ker(X)\neq\{0\}$. לכן אמ"מ יש אינסוף פתרונות Ω_2

כלומר $u\in\ker(X)$ מתקיים $X(u+w)=y$

(ד) נראה ש X^TX הפיך מתקיים:

$$X^TXw=X^Ty\Rightarrow(X^TX)^{-1}X^TXw=(X^TX)^{-1}X^Ty\Rightarrow w=(X^TX)^{-1}X^Ty$$

נראה כי $X^Ty\perp\ker(X^TX)$

יהי $u\in\ker(X^TX)$

$$\langle u,X^Ty\rangle=\langle Xu,y\rangle\underset{u\in\ker(X)\Leftarrow u\in\ker(X^TX)}{=}\langle 0,y\rangle=0$$

לכן יש פתרונות ל $X^TXw=X^Ty$ כי קיים פתרון

(8)

$X = U\Sigma V^T$

$$X^Tx = (U\Sigma V^T)^T U\Sigma V^T = V\Sigma^T\Sigma V^T$$

אם X^Tx הפיכה אז $V\Sigma^T\Sigma V^T$ הפיכה

ולכן בהכרח $\Sigma^T\Sigma$ הפיכה

מכאן נובע $(X^Tx)^{-1} = V(\Sigma^T\Sigma)^{-1}V^T$ ולכן כדי להוכיח $X^+y = (X^Tx)^{-1}X^Ty$

צריך שיתקיים ⊛ $V\Sigma^+U^Ty = V(\Sigma^T\Sigma)^{-1}\Sigma^TU^Ty$

בנוסף: ① $(X^Tx)^{-1}x^Ty = V(\Sigma^T\Sigma)^{-1}\Sigma^TU^Ty$ ② $X^+y = V\Sigma^+U^Ty$

⊛ נראה ש $\Sigma^+ = (\Sigma^T\Sigma)^{-1}(\Sigma)^T$

נרצה להראות כי $\Sigma^T\Sigma$ אלכסונית כיוון שהיא ריבועית

$$[\Sigma^T\Sigma]_{i,j} = \sum_{l=1}^{n}(\Sigma)^T_{i,l}\cdot(\Sigma)_{l,j} = (\Sigma)^T_{i,i}\cdot(\Sigma)_{i,j}$$

ומכאן נסיק כי כאשר $i=j$ נקבל σ_i^2 ואחרת 0 ולכן האלכסונית.
$\sigma_i\neq0$ כי היא מטריצה אלכסונית הפיכה ולכן ערכי $(\Sigma^T\Sigma)^{-1}$ יהיו כאשר $i=j$ $\frac{1}{\sigma_i^2}$
ואחרת 0.

בנוסף מתקיים

$$((\Sigma^T\Sigma)^{-1}\Sigma^T)_{i,j} = \sum_{l=1}^{m}(\Sigma^T\Sigma)_{i,l}\cdot\Sigma^T_{l,j} = (\Sigma^T\Sigma)_{i,i}\cdot(\Sigma^T)_{i,j}$$

ולכן כאשר $i=j$ נקבל $\frac{1}{\sigma_i}$ ואחרת 0

זו מטריצה $d\times m$ ולפי ההגדרה מתקיים כי $\Sigma^+ = (\Sigma^T\Sigma)^{-1}(\Sigma)^T$

ולכן $X^+y = (X^Tx)^{-1}X^Ty$ כנדרש.

practical part

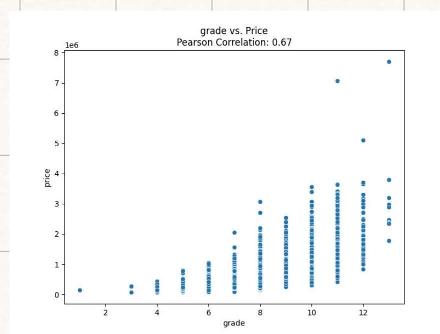
3.1.3 * לכן הוצאתי את ה-id כיון שלו לא היה שום ערך בהכרעה
בערך שלו. ובנוסף את ה date
כמו כן החלפתי את השנה שבוצעה שיפוץ ל 0 אם שופץ כדי לא
להיות ספציפיים מדי
בדקתי כי אין שנה שבה שופץ יכול להיות לפני השנה שנבנה
בדקתי שגודל הבית כללי הוא חיובי
בדקתי שאין ערכי null או ערכים שליליים

* השנה שנבנה הבית יש לו סדר שאפשר לכן לנסות הוא קטגוריאלי.

* הזיפקוד מאחר והערכים שלו הם שווה מספר לזה נקודה כבורה יותר
יותר משנה ספציפית

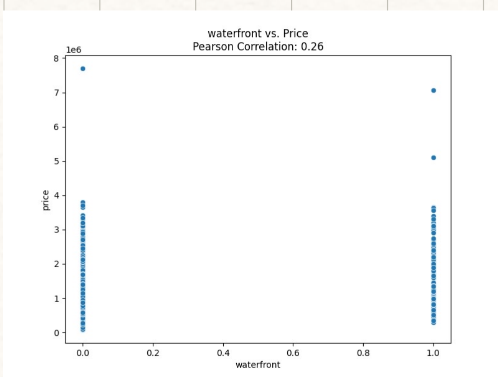
* הורדתי את השורה עם 33 חדרים כחול

3.1.4 טוב:



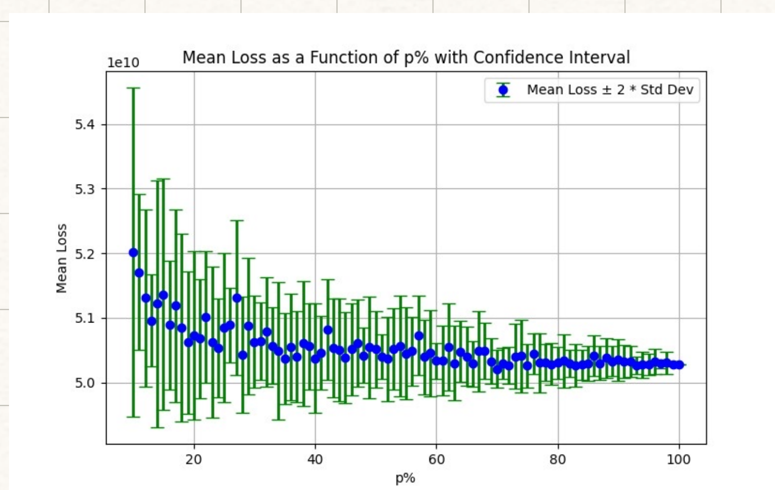
זה טוב שיש להן לציון מחיר ממוצע מוגדל
ולכן כן ביון המחירים יעזור לנבא טובה
יותר טובה

לא טוב:



זה לא כי אין יותר כל כך בין
המחיר ולכן לא יעזור לנבא טובה
יותר טובה

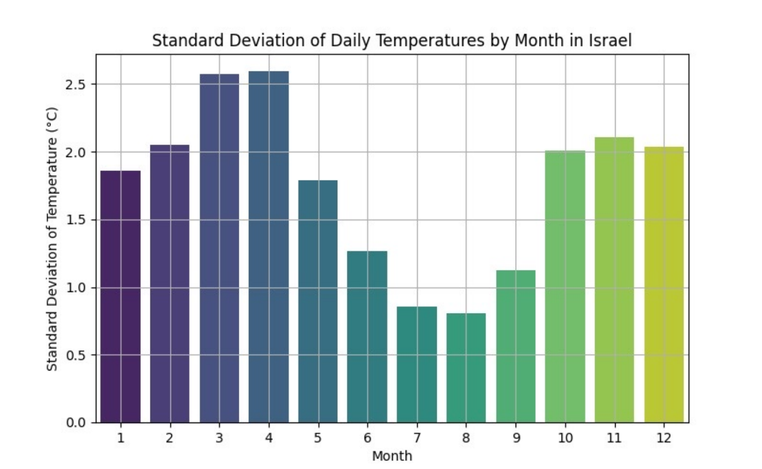
3.1.6

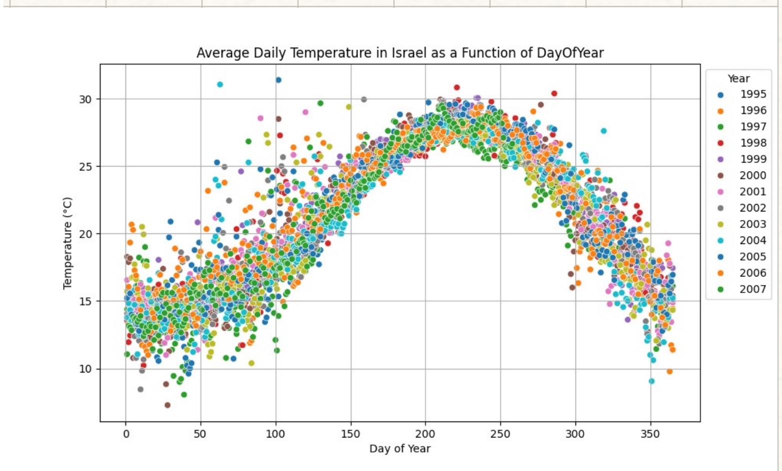


מה שאנחנו רואים הוא לראות הוא כי ככל שהמחוז
גדל ככה ה-loss יותר כבורה שמאלה
K's שנים ב-training ככל שנוני
לומדים יותר בהרצאות הלמידה לכן יותר טוב
מסבירים את ה-loss בצורה גבוהה.

Polinomial fitting

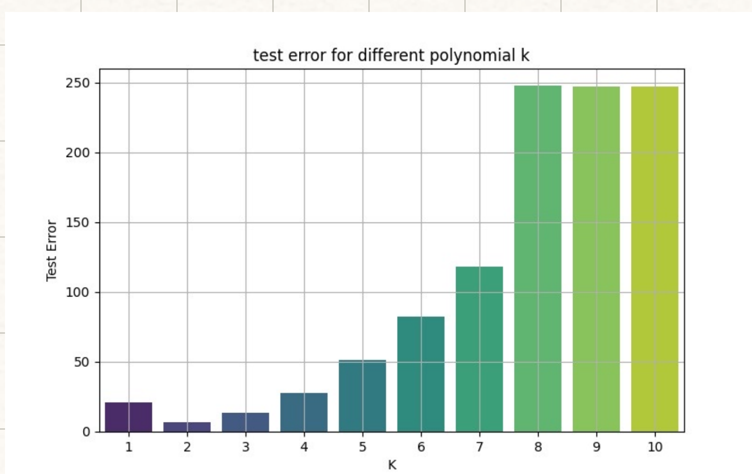
(3) ניתן לראות מהגרף שיש חודשים (לדוגמא חודשי הקיץ) אשר בהם הסטיית תקן היא
יותר נמוכה ולכן הניבוי של ה-מודל עבור חודשים אלו יהיה מדויק יותר מאשר בחודשים אחרים
יותר.

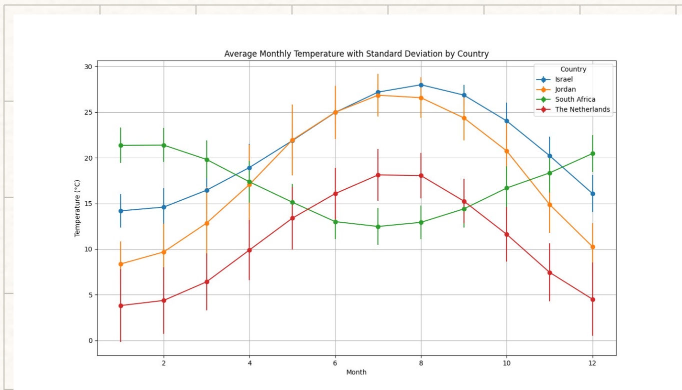


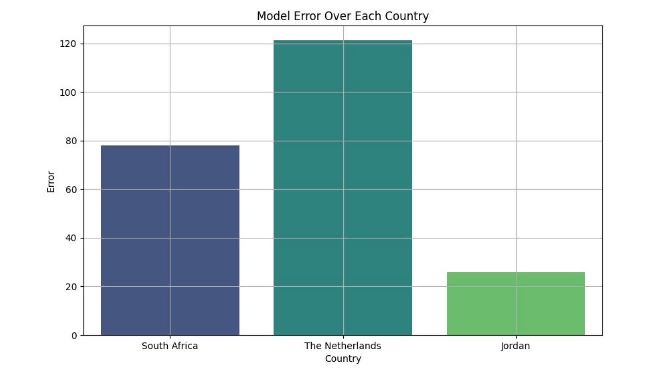


(4) לכן לסיכום יש את אותה פונק' לאורך שנה וכך כל שנה
כלומר שהעונות יוצרות את הסינוס הכי מתאים על המודל שמתאים לישראל הוא לא בהכרח מתאים
הכי פחות מכיוון שלהולנד ויורדן יש פונק' שונה מאוד מההתפלגות לזו של ישראל
ולכן המודל של ישראל יכול להתאים רק במידה מסוימת

(5) ראינו בסעיף 3 כי הדוח המייצג את הטמפרטורות כתלות ביום בשנה הוא מתנהג כמו
$x^2+\dots$ יותר מאשר כל דרגה אחרת לכן לא מתאים לפולינום עבור $k=2$ נקבל את ה-error הנמוך ביותר
וכן גם עבור כל k שגדול ומרוחק מ-2 נקבל error יותר גדול, כמו כל מה שלמדנו בתרגול כמו $k=8$
מעניין שגם עד ה-נקודה הקודמת לו
לאחר מכן קיבלנו כי כל יחידה מעל לכך לא גדל.
עבור $k=8/9/10$ נקבל את המודל הכי גרוע
גדול רק $k=8$







6) ניתן לראות בגרף שהמודל מסוגל 4
כי הפער אצל בישראל יחסית דומה לירדן
ורק אחריה ולעומת זאת דאמסטרדם זה
יותר. לכן מה שניתן לראות שבגלל מה ש
אצלו מתנהגת יותר ויותר מישראל המודל מפספס
ולמעשה, כפי שניתן לראות בגרף התחתון