צבי מרמור 206949539

21 רקע מתמטי

2.11 אלגברה לינארית

1 נחשב את ה SVD של המטריצה הנ"ל

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

ראשית, נחשב את:

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

נמצא את הע"ע של A^TA:

$$\det(A^TA-\lambda I)=0\rightarrow\left|\begin{bmatrix}2-\lambda&0&2\\0&2-\lambda&-2\\2&-2&4-\lambda\end{bmatrix}\right|=0\rightarrow\lambda(\lambda-4)(\lambda-6)=0\rightarrow\underline{\lambda=0,2,6}$$

נמצא את הו"ע של $\lambda=4,6$:

$$(A^TA-0I)v=0\Rightarrow\begin{bmatrix}2&0&2\\0&2&-2\\2&-2&4\end{bmatrix}\begin{bmatrix}x_1\\x_2\\x_3\end{bmatrix}=0\Rightarrow\begin{bmatrix}2x_1&0&2x_3&0\\0&2x_2&-2x_3&0\\2x_1&-2x_2&4x_3&0\end{bmatrix}\Rightarrow\begin{bmatrix}x_1&0&x_3&0\\0&x_2&-x_3&0\\x_1&-x_2&2x_3&0\end{bmatrix}\Rightarrow\begin{matrix}x_1=-x_3\\x_2=x_3\\x_3=x_3\end{matrix}\Rightarrow v_0=\begin{pmatrix}-1\\1\\1\end{pmatrix}$$

$$\|v\|=\sqrt{\langle v,v\rangle}=\sqrt{3}$$

נרצה לנרמל את הוקטור כי בסיס זה אורתונורמלי. לכן:

$$\underline{v_0^T=\left(-\frac{1}{\sqrt3},\frac{1}{\sqrt3},\frac{1}{\sqrt3}\right)}$$

לכן

$$(A^TA-2I)v=0\Rightarrow\begin{bmatrix}0&0&2\\0&0&-2\\2&-2&2\end{bmatrix}\begin{bmatrix}x_1\\x_2\\x_3\end{bmatrix}=0\Rightarrow\begin{bmatrix}0&0&2x_3&0\\0&0&-2x_3&0\\2x_1&-2x_2&2x_3&0\end{bmatrix}\Rightarrow\begin{bmatrix}0&0&x_3&0\\0&0&-x_3&0\\x_1&-x_2&x_3&0\end{bmatrix}\quad\begin{matrix}x_1=x_2\\x_3=0\end{matrix}\Rightarrow v=\begin{pmatrix}1\\1\\0\end{pmatrix}$$

$$\|v_4\|=\sqrt{\langle v_4,v_4\rangle}=\sqrt2$$

ננרמל כנ"ל.

לכן

$$\underline{v_2^T=\left(\frac{1}{\sqrt2},\frac{1}{\sqrt2},0\right)}$$

$$(A^TA-6I)v=0\Rightarrow\begin{bmatrix}-4&0&2\\0&-4&-2\\2&-2&-2\end{bmatrix}\begin{bmatrix}x_1\\x_2\\x_3\end{bmatrix}=0\Rightarrow\begin{bmatrix}-4x_1&0&2x_3&0\\0&-4x_2&-2x_3&0\\2x_1&-2x_2&-2x_3&0\end{bmatrix}\Rightarrow\begin{bmatrix}-2x_1&0&x_3&0\\0&2x_2&x_3&0\\x_1&-x_2&-x_3&0\end{bmatrix}\Rightarrow\begin{matrix}x_1=\frac12x_3\\x_2=-\frac12x_3\end{matrix}\Rightarrow v=\begin{pmatrix}1\\-1\\2\end{pmatrix}$$

$$\|v_6\|=\sqrt{\langle v_6,v_6\rangle}=\sqrt{1+1+4}=\sqrt6$$

ננרמל כנ"ל.

לכן:

$$\underline{v_6^T=\left(\frac{1}{\sqrt6},-\frac{1}{\sqrt6},\frac{2}{\sqrt6}\right)}$$

כלומר, מה שנקבל:

$$\underline{V=\begin{bmatrix}\frac{1}{\sqrt6}&\frac{1}{\sqrt2}&-\frac{1}{\sqrt3}\\-\frac{1}{\sqrt6}&\frac{1}{\sqrt2}&\frac{1}{\sqrt3}\\\frac{2}{\sqrt6}&0&\frac{1}{\sqrt3}\end{bmatrix}}\qquad\underline{\Sigma=\begin{bmatrix}\sqrt6&0&0\\0&\sqrt2&0\end{bmatrix}}$$

(לפי מה שלמדנו בתרגול נוריד את $\lambda=0$).

כעת, נמשיך ל-U. נחשב את AA^T ונמצא ערכים עצמיים

תהליך דומה פה למציאת את U:

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

$$\det(AA^T-\lambda I)=0\Rightarrow\begin{vmatrix}2-\lambda&0\\0&6-\lambda\end{vmatrix}=0\Rightarrow(2-\lambda)(6-\lambda)=0\Rightarrow\lambda=2,6$$

כעת, נמשיך עם:

$$(AA^T-2I)U=0\rightarrow\begin{bmatrix}0&0\\0&4\end{bmatrix}\begin{bmatrix}x_1\\x_2\end{bmatrix}=0\Rightarrow\begin{bmatrix}0&0&0\\0&4x_2&0\end{bmatrix}\rightarrow\begin{matrix}x_1=x_1\\x_2=0\end{matrix}\rightarrow U_2=\begin{bmatrix}1\\0\end{bmatrix}$$

$$U_2^T=(1\ \ 0)$$

הוקטור כבר מנורמל, לכן

$$(AA^T-6I)U=0\Rightarrow\begin{bmatrix}-4&0&0\\0&0&0\end{bmatrix}\rightarrow\begin{matrix}x_1=0\\x_2=x_2\end{matrix}\rightarrow U_6=\begin{bmatrix}0\\1\end{bmatrix}$$

כבר מנורמל

לכן

$$U=\begin{bmatrix}0&1\\1&0\end{bmatrix}$$

כלומר

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

② יהיו $U\in R^m$, $V\in R^n$ אזי המכפלה הטנזורית שלהם היא

$$V\otimes U=V\times U^T=\begin{pmatrix}V_1\\\vdots\\V_n\end{pmatrix}(U_1\ \cdots\ U_m)=\begin{pmatrix}V_1U_1&\cdots&V_1U_m\\\vdots&&\\V_nU_1&\cdots&V_nU_m\end{pmatrix}\in R^{n\times m}$$

מהגדרת כפל מטריצות שוויון ההגדרה נכון.

לפי כלל כפל ($M_{m\times1}\times M_{1\times n}$) שימוש

נשים לב כי כל שורות המטריצה ת"ל. עבור כל i שורה i במטריצה היא מהצורה:

$$R_i=(V_i\cdot U_1,V_i\cdot U_2,\dots,V_iU_m)=V_i\cdot(U_1,\dots,U_m)=V_i\cdot\underline{\underline{U}}\rightarrow$$ אותו וקטור בו

ומכאן תלויות לינארית. באופן דומה נשים לב כי כל עמודות המטריצה ת"ל. עמודה j היא מהצורה:

$$C_j^T=(V_1U_j\ \dots\ V_nU_j)=U_j(V_1\ \dots\ V_n)=U_j\cdot\underline{\underline{V}}\rightarrow$$ אותו וקטור בו

ולכן כל שורות ועמודות המטריצה ת"ל.

∎

③ נבו.י מה צריך להוכיח. יהי $V=\mathbb{R}^n$ מ"ו ו-$U=(u_1 \dots u_n)$ בסיס אורתונורמלי של V. צ"ל עבור וקטור כלשהו $X\in\mathbb{R}^n$ שאם מתקיים עבור

כ $X=\sum_{i=1}^{n} a_i u_i$ עבור n סקלרים a_i כי:

$\langle X,u_i\rangle=a_i$. כלומר $X=\sum_{i=1}^{n}\langle X,u_i\rangle u_i$

נוכיח:

$$\sum_{i=1}^{n}\langle X,u_i\rangle u_i=\sum_{i=1}^{n}\Big\langle\sum_{j=1}^{n}a_j u_j,u_i\Big\rangle u_i=\sum_{i=1}^{n}\sum_{j=1}^{n}a_j\langle u_j,u_i\rangle u_i=\sum_{i=1}^{n}a_i u_i=X$$

הגדרת X

מכיוון ש-U בסיס אורתונורמלי. המכפלה הפנימית תתן 0 בכל המקרים בהם $i\neq j$ ו-1 במקרה בו $i=j$. ולכן רק כאשר $i=j$ מתקיים $a_j\langle u_j,u_i\rangle u_i=a_j\cdot 1\cdot u_i$
$=a_i u_i$

הגדרת X.

■

4) **k.** צ"ל כי P הינה סימטרית, כלומר $P=P^T$

$$P=\sum_{i=1}^{k} V_iV_i^T \ , \quad P^T=\left(\sum_{i=1}^{k} V_iV_i^T\right)^T=\sum_{i=1}^{k}\left(V_iV_i^T\right)^T=\sum_{i=1}^{k} V_i^{T^T}V_i^T=\sum_{i=1}^{k} V_iV_i^T=P$$

הגדרת transpose על מטריצות

$V^{T^T}=V$ לכל וקטור זה קורה.

ב. צ"ל: הע"ע של P הם $0,1$ ו $V_1..V_k$ הו"ע של הע"ע 1.

יהי $V_i\in V$ אזי: מתקיים

$$PV_i=\sum_{j=1}^{k} V_jV_j^TV_i=\sum_{i=1}^{k} \mathbb{1}_{i=j}V_i=V_i$$

מכיוון שהוקטורים ב V הם אורתונורמלים אם $i=j$ המכפלה היא 1 ואם לא 0 (הוכחה בשאלה הבאה).

כלומר, $PV_i=V_i$. הערכים העצמיים המתקיימים מכל וקטור 0 הם 1

ו 1 הוא הע"ע של P.

∎

ג. יהי $V\in V$. מתקיים

$$PV=\sum_{i=1}^{k} V_iV_i^TV=\sum_{i=1}^{k} V_i\left(V_i^TV\right)=\sum_{i=1}^{k} V_i\langle V_i,V\rangle=V$$

כלומר כל P_V מחזיר את הוקטור בהתאם לס' 3.

∎

3. צ"ל $P^2=P$.

$$P^2=\sum_{i=1}^{k} V_iV_i^T\cdot\sum_{j=1}^{k} V_jV_j^T=\sum_{i=1}^{k}\sum_{j=1}^{k} V_iV_i^TV_jV_j^T=\sum_{i=1}^{k}\sum_{j=1}^{k} V_i\left(V_i^TV_j\right)V_j^T=\sum_{i=1}^{k}\sum_{j=1}^{k} V_i\langle V_i,V_j\rangle V_j^T=$$

$$\begin{cases} i=j & \langle V_i,V_j\rangle=1 \\ i\neq j & \langle V_i,V_j\rangle=0 \end{cases}$$

כמו שראינו בסעיף 3, מתקיים: סט אורתונורמלי, המכפלה הפנימית מתאפסת

$$=\sum_{i=1}^{k} V_iV_i^T=P$$

∎

ה. צ"ל: $(I-P)P=0$:

$$(I-P)P = IP - P^2 = IP - P = P - P = 0$$

$P^2 = P$ הנתון.

$\blacksquare$

2.1.2 חשבון אינפיניטסימלי רב משתני

5 כאשר נשים:

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

שימוש בכלל השרשרת:

$$\frac{\partial h}{\partial a_i} = \frac{\partial h}{\partial f}\cdot\frac{\partial f}{\partial a_i} = \frac{\partial}{\partial f}\cdot\left(\frac{1}{2}\|f(\sigma)\|^2 - f(\sigma)^T y + \frac{1}{2}\|y\|^2\right)\frac{\partial f}{\partial a_i}$$

$$= \left(\frac{1}{2}\cdot 2f(\sigma) - y\right)\frac{\partial f}{\partial a_i} = \left(f(\sigma)-y\right)\frac{\partial f}{\partial a_i} = (*)$$

כאשר מגדירים כ

$$\frac{\partial f}{\partial a_i} = J_\sigma(f)$$

לכן

$$(*) = \left(f(\sigma)-y\right)\cdot J_\sigma(f)$$

6 מתקיים: $\frac{\partial f_i(x)}{\partial x_j} = [\nabla f(x)_i]_j = [J_x(f)]_{ij}$. לכן: אם כך, ונגזור את S:

$$[\nabla S(x)_i]_j = \frac{\partial}{\partial x_j}\cdot\frac{e^{x_i}}{\sum_{l=1}^{k} e^{x_l}} = \frac{e^{x_i}\left(\sum_{l=1}^{k} e^{x_l}\right) - e^{x_i}e^{x_j}}{\left(\sum_{l=1}^{k} e^{x_l}\right)^2}$$

כמו כן, אם $i=j$:

$$[\nabla S(x)_i]_j = \frac{e^{x_i}}{\sum_{l=1}^{k} e^{x_l}}\cdot\frac{\left(\sum_{l=1}^{k} e^{x_l}\right) - e^{x_j}}{\sum_{l=1}^{k} e^{x_l}} = S(x)_i(1-S(x)_i)$$

אם $i\neq j$:

$$[\nabla S(x)_i]_j = \frac{0\cdot\sum_{l=1}^{k} e^{x_l} - e^{x_i}e^{x_j}}{\left(\sum_{l=1}^{k} e^{x_l}\right)^2} = \frac{-e^{x_i}}{\sum_{l=1}^{k} e^{x_l}}\cdot\frac{e^{x_j}}{\sum_{l=1}^{k} e^{x_l}} = -S(x)_i S(x)_j$$

כלומר:

$$[\nabla S(x)_i]_j = S(x)_i\left(\mathbb{1}_{i=j} - S(x)_j\right)$$

ובצורה מטריצית:

$$J_x(S) = \mathrm{diag}(S) - SS^T,$$

1 א. נר׳ כי $\ker(T)=\ker(X^TX)$. יהי $v\in\ker(T)$ כך ש $v\neq 0$

אז מתקיים מהגדרה $Xv=0$. אז $X^TXv=X^T0=0$.

לכן $v\in\ker(X^TX)$.

יהי $v\in\ker(X^TX)$. אז $X^TXv=0$. לכן $v^TX^TXv=0$

לכן. מתקיים $0=v^TX^TXv=(Xv)^TXv=\|Xv\|$ מהגדרת הנורמה

$\|Xv\|=0 \Leftrightarrow Xv=0$ לכן $v\in\ker(X)$. ■

ב. נניח $A\in M_{k\times k}$ נר׳ כי $\operatorname{Im}(A^T)=\ker(A)^{\perp}$. נראה כי $\operatorname{Im}(A)\subset\ker(A$

יהי $b\in\operatorname{Im}(A^T)$ אז קיים $x\in\mathbb{R}^k$ כך ש $A^Tx=b$.

יהי $w\in\ker(A^T)$ אז מהגדרה $A^Tw=0 \Leftarrow$

$0=\langle A^Tw,x\rangle=\langle w,Ax\rangle=\langle w,b\rangle$

לכן $\operatorname{Im}(A)\subset\ker(A^T)^{\perp}$.

נראה כי $\ker(A^T)^{\perp}\subset\operatorname{Im}(A)$. יהי $b\in\operatorname{Im}(A)$ נניח בשלילה כי $b\notin\ker(A^T)^{\perp}$

אזי קיים $a\in\ker(A^T)$ כך ש $\langle b,a\rangle\neq 0$. משמע כי $b\in\operatorname{Im}(A)^{\perp}$

אז יהי $a\in\operatorname{Im}(A)^{\perp}$ כך ש $\langle b,a\rangle\neq 0$. מכאן ש $a\in\operatorname{Im}(A)^{\perp}$

מתקיים כי [illegible] לכל $v\in\operatorname{Im}(A)$ ובפרט $\langle a,AA^Ta\rangle=0$

לכן $\|A^Ta\|^2=\langle A^Ta,A^Ta\rangle=\langle a,AA^Ta\rangle=0$

לכן מחיוביות הנורמה $A^Ta=0 \Leftarrow a\in\ker(A^T)$ ולכן הגענו לסתירה

אל מניחים. ■

ג. נראה כי X אינה הפיכה. לכן למערכת המשוואות יש 0 פתרונות או ∞ פתרונות (הוכחנו בלינארית 1).

נראה $Xw=y$ כאשר $y\neq 0$. כלומר למערכת יש לפחות פתרון אחד. לכן למערכת ∞ פתרונות.

מכיון ש $y\in Im(X)$ מתקיים $y\in Im(X) \iff y\in ker(X)^{\perp}$ מסעיף ב. ומכאן נובע $\iff$ $y\perp ker(X)$ כנדרש.

■

ד. נניח כי X^TX אינה הפיכה. מסעיף ג מתקיים כי למערכת המשוואות יש ∞ פתרונות אמ"מ $X^Ty\perp ker(X^TX)$. אך מסעיף א נובע כי הגרעינים שווים ($X^Ty\perp ker(X)$

כלומר, זה מה שצריך להוכיח.

יהי $v\in ker(X)$ אז $\langle Xv, y\rangle = 0$ אך זה שווה ל

$\langle v, X^Ty\rangle = 0$ כלומר $X^Ty\perp ker(X^TX)$ כנדרש.

לכן, למערכת המשוואות יש ∞ פתרונות אם המטריצה אינה הפיכה, ו פתרון יחיד אם המטריצה X^TX הפיכה.

2.2.2

2 מתקיים:

$$[X^TX]^{-1}X^T = [(U\Sigma V^T)^T(U\Sigma V^T)]^{-1}(U\Sigma V^T)^T = [V\Sigma^T\cancel{U^TU}\Sigma V^T]^{-1}V\Sigma^TU^T$$

$X=U\Sigma V^T$ U אורתוגונלית

$$= [V\Sigma^T\Sigma V^T]^{-1}V\Sigma^TU^T = V(\Sigma^T\Sigma)^{-1}\cancel{V^TV}\Sigma^TU^T = V(\Sigma^T\Sigma)^{-1}\Sigma^TU^T$$

V אורתוגונלית

נגדיר $A=\Sigma^T\Sigma$ אז מתקיים

$$= VA^{-1}\Sigma^TU^T$$

אז מתקיים ש המטריצה A היא אלכסונית כמו Σ ומתקיים כי $A_{ii}=(\sigma_i)^2$

כמו כן, מכיון שהעמודות של X בת"ל מתקיים כי $\sigma_1\geq\sigma_2\geq\dots\geq\sigma_{d+1}>0$ (ולכן המטריצה A) מתקיים כי המטריצה A הפיכה. כמו כן:

$$[A^{-1}\Sigma^T]_{ii} = \frac{1}{(\sigma_i)^2}\cdot\sigma_i = \frac{1}{\sigma_i} = \Sigma^{+}_{ii}$$

לכן מתקיים לכל $y\in\mathbb{R}^{d+1}$ כי $[X^TX]^{-1}X^Ty = V\Sigma^{+}U^Ty = X^{+T}y$

■

2 מתקיים:

$$[X^TX]^{-1}X^T = \left[(U\Sigma V^T)^T(U\Sigma V^T)\right]^{-1}(U\Sigma V^T)^T = [V\Sigma^T \cancel{U^T U}\Sigma V^T]^{-1}V\Sigma^T U^T$$

$X = U\Sigma V^T$

U אורתוגונלית

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

V אורתוגונלית

נגדיר $A=\Sigma^T\Sigma$ אשר מתקיים

$$= VA^{-1}\Sigma^T U^T$$

אשר מתקיים ש- מטריצה A היא אלכסונית וריבועית, Σ מתקיים ש- $A_{ii}=(\sigma_i)^2$

כמו כן, מכיוון שמטריצת X הינה מדרגה מלאה מתקיים ש- $\sigma_1 \ge \sigma_2 \ge \dots \ge \sigma_{d+1} > 0$ (כלל)

(כלומר A מתקיים ש- איברי A הפיכה. ולכן:

$$[A^{-1}\Sigma^T]_{ii} = \frac{1}{(\sigma_i)^2}\cdot\sigma_i = \frac{1}{\sigma_i} = \Sigma^{+}_{ii}$$

לכן מתקיים עבור כל $y\in\mathbb{R}^{d+1}$ כי

$$[X^TX]^{-1}X^Ty = V\Sigma^{+}U^Ty = X^{+^T}y$$

∎

3. חלק פרקי:

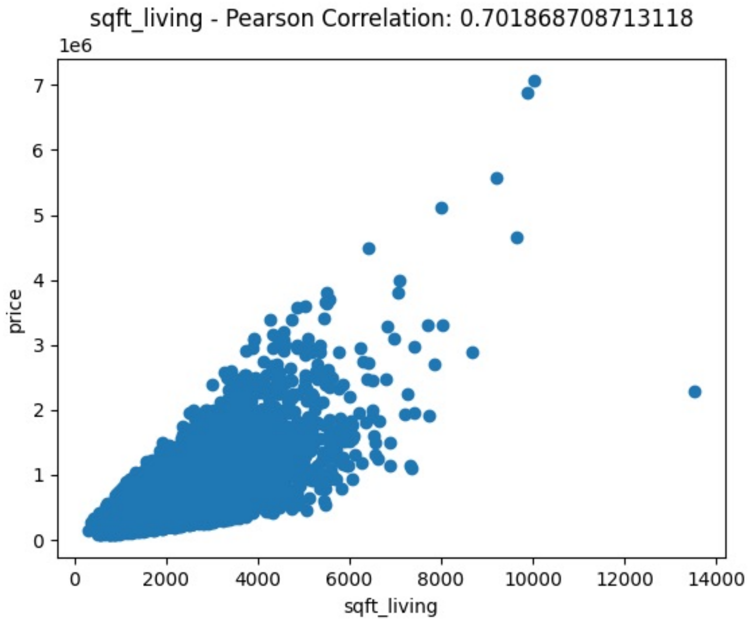
3.1.3 הסבר נרחב לתהליך העיבוד המקדים של המידע שהתקבל לגבי מחירי בתים:
ראשית, בחרתי להוריד כל פיצ׳ר שלא נותן לי מידע כלשהו שתורם מידע לגבי סכום המכירה, כלומר, כל פיצ׳ר שבמבט פשוט אינו קשור ישירות אל סכום המכירה. פיצ׳רים כמו ID, lat' long אשר בהבנה פשוטה לא תרמו לי כל מידע נוסף לגבי המכירה.
בנוסף, הפכתי את המידע על שנת השיפוץ לקטגוריאלי- האם הבית שופץ מאז בנייתו או לא. כמו כן, המרתי את date, year built לפיצ׳ר יחיד שקבע בן כמה היה הבית ברגע מכירתו.

כמו כן, בחרתי להפוך למשתנה קטגוריאלי את הפיצ׳ר המייצג את הבית שופץ או לא, ובחרתי להשאיר כמשתנים קטגוריאליים אשר להם סולם מדידה איכותי את הפיצ׳רים אשר המידע לגיבהם הוצג בצורה קטגוריאלית- פיצ׳רים כמו waterfront, view, grade וכו.

החשיבה שעמדה מאחוריי הפיכתי את תאריך המכירה ותהאריך בניית הבית למשתנה יחיד שמייצג את גיל הבית בעת מכירתו נבע מכך שחשבתי שמדד זה ישקף ביתר בהירות מידע על מחיר הבית או יתן למודל מידע שיעזור לו לחזות את מחיר הבית בצורה טובה יותר. ההנחה שלי הייתה שמשתנה יחיד כמו זה משקף נתונים אלו בצורה בהירה יותר מאשר שנת המכירה או שנת הבנייה.

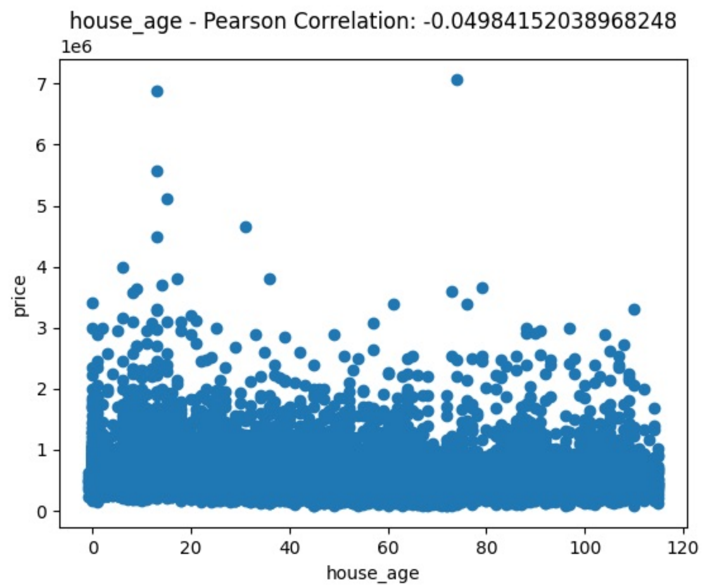
בנוסף, עברתי על הdata כולו ומחקתי ערכים שגויים וערכי nan הן ממטריצת הפיצ׳רים והן מוקטור מחירי הבתים. (X,y בהתאמה).

3.1.4 בחרתי להביא את גרפי הקורלציה של הפיצ׳רים הבאים:

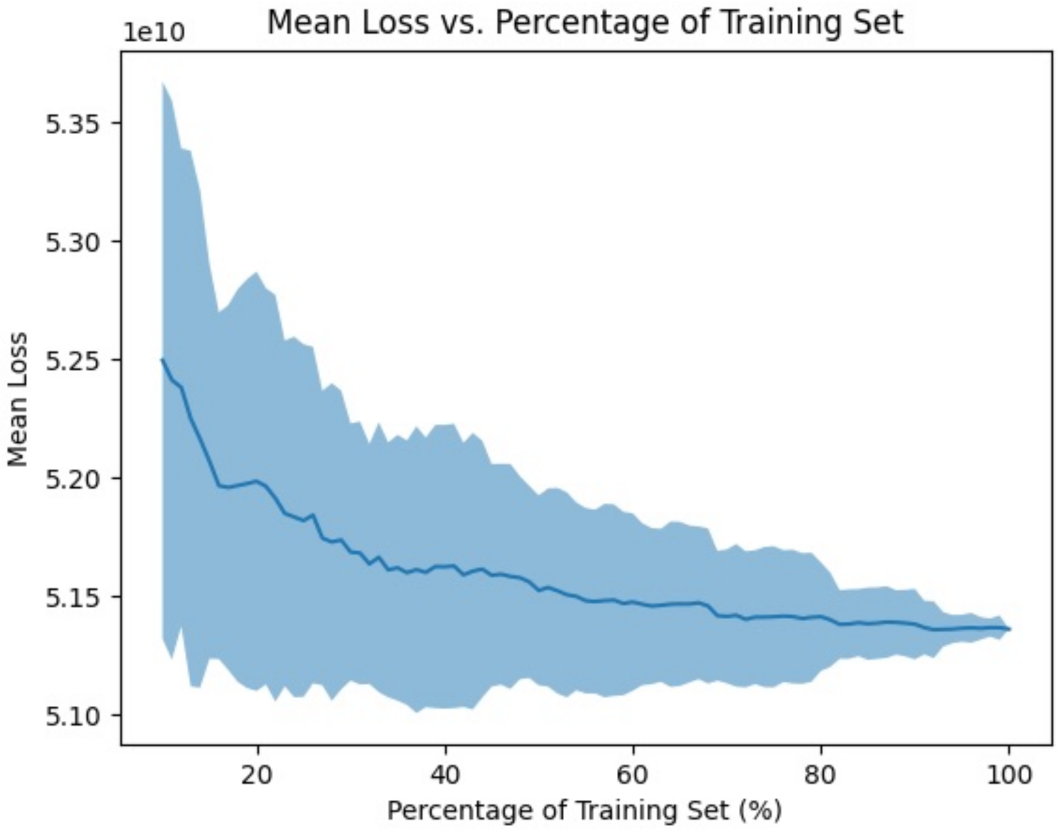


מדד הקורלציה של sq ft living.
ניתן לראות בקלות כי קיימת מגמה בגרף, וככל שגודל שטח המגורים גדל גם מחיר הבית. דבר זה מתבטא גם בציון הקורלציה הגבוה של פיצ׳ר זה- מעל 0.7. דבר זה גורם לכך שפיצ׳ר זה מיטבי עבור המודל ומספק לו מידע ׳חיוני׳ המסייע לו לדעת לחזות את מחיר הבית.

מדד הקורלציה של house age.
בשונה מהגרף הקודם, גרף הקורלציה עבור מחיר הבית אינו מיטבי עבור המודל. ניתן לראות כי אי אפשר לזהות מגמה כלשהי בגרף. כלומר, גיל הבית אינו מעיד בהכרח על מחיר גבוה או נמוך של הבית. דבר שמתבטא גם בצן קורלציה נמוך- מתחת ל0.05. לכן, פיצ׳ר זה אינו ׳תורם׳ הרבה מידע חיוני למודל ומסייע לו לחזות מחירי בתים.

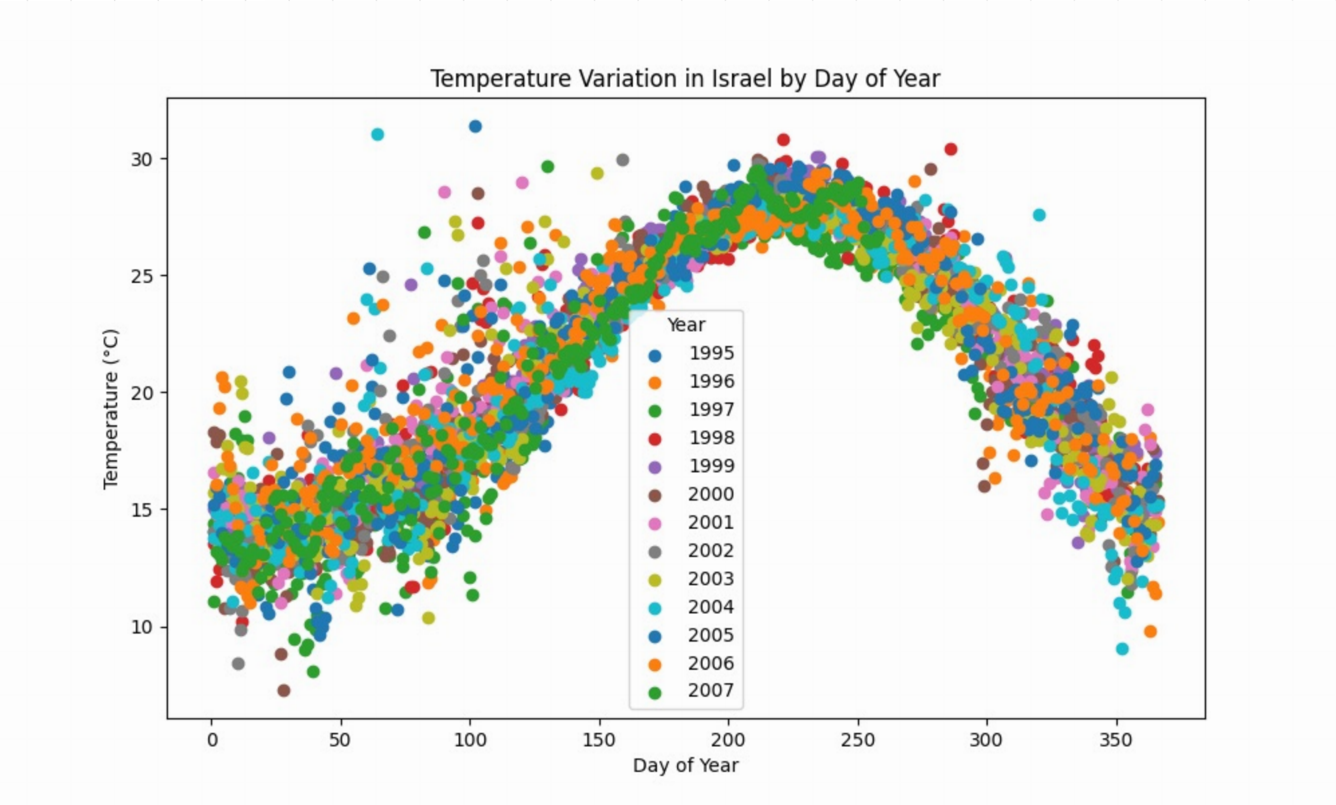


3.1.6 הגרף המוצג מטה מציג את מדד הLoss כתלות בגודל הtraining set.
ניתן בנקל לראות כי ככל שגודל הtraining set גדול יותר כך גודל הloss (המייצג את מספר ה ׳שגיאות׳ בהן שגה המודל) קטן. כלומר, ככל שגודל הנתונים עליהם המודל ׳מתאמן׳ כך מספר השגיאות שלו פוחת בהתאמה.
כמו כן, ניתן לראות כי ככל שגודל הtraining set עולה, כך ה׳איזור התכול׳ מסביב לקן המרכי מתמרכז ומתקרב לקו המרכזי. איזור זה מייצג את הconfidence integral של המודל, כלומר, את אומדן אי הוודאות של המודל. ניתן לראות כי ככל שגודל הtraining set עולה כך איזור זה נהיה קרוב יותר אל הקו המרכזי, דבר המעיד על כך שהמודל מציד תוצאות יותר קונסטנטיביות ויותר קרובות לפתרון נכון.

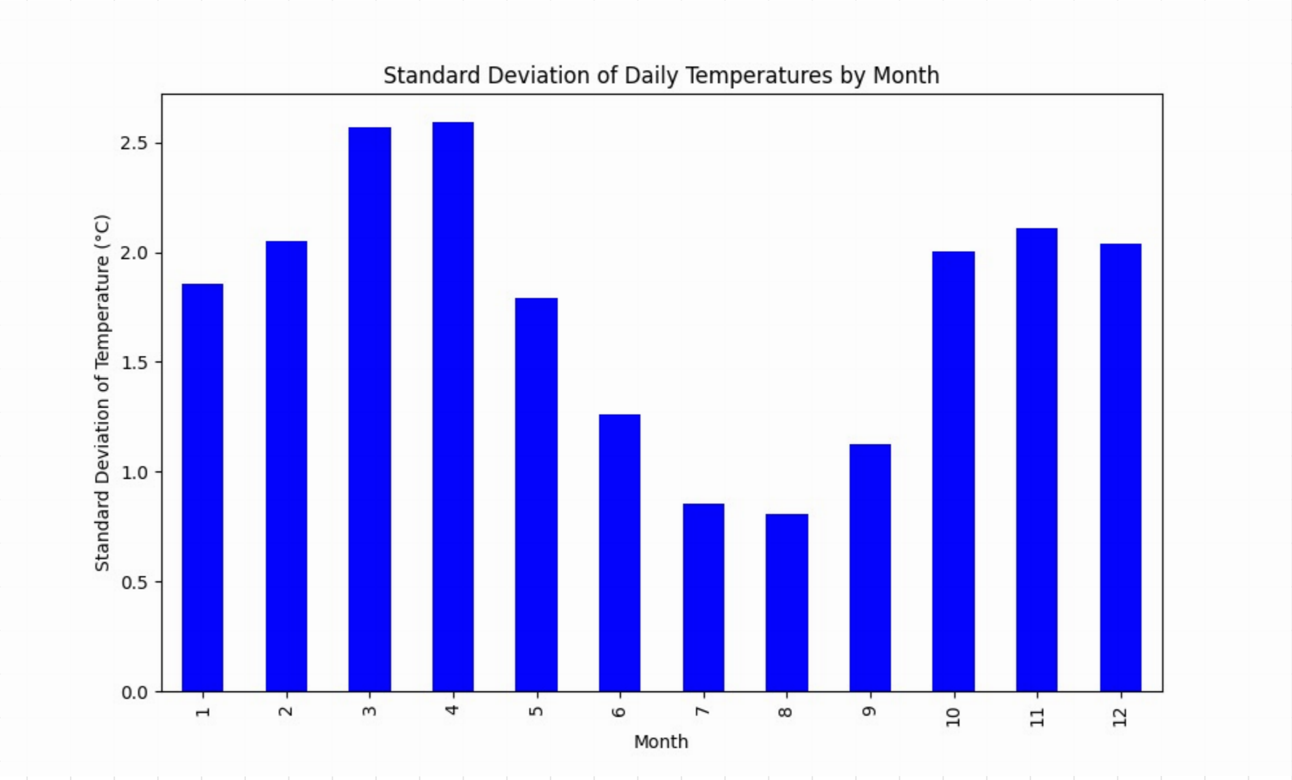


3.2.3 הגרף המוצג מטה מציג פיזור של טמפ׳ יומיות ממוצעות בישראל לפי יום בשנה. הגרף מציג מגמה ברורה של עלייה בטמפ׳ במהלך האביב, הגעה לשיא בקיץ ואז ירידה בסתיו עד לחורף.

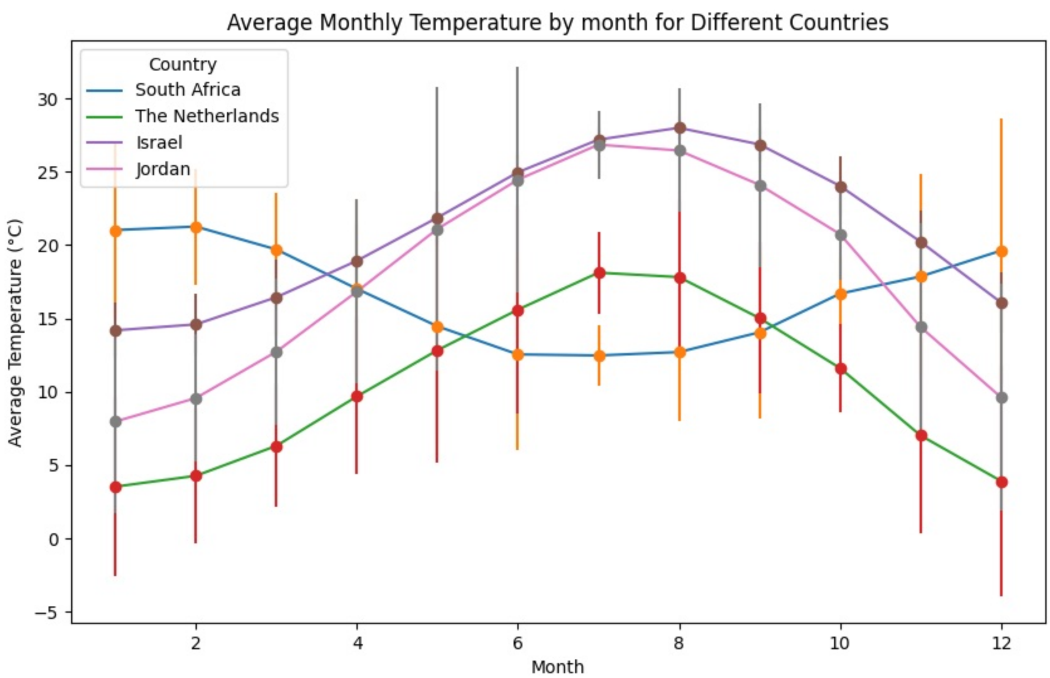
מתוצרת הנתונים נראה שהפולינום הוא מדרגה הדומה לפולינום מדרגה שנייה או שלישית, ולכן יכול להיות שקיים פולינום ממעלה כזו היכול לתאר מגמה כללית של הגרף בצורה קרובה (יחסית) לגרף המקורי.



הגרף המוצג מטה מציג את סטיית התקן של הטמפ׳ היומיות לפי חודשי השנה בישראל. ניתן לראות כי קיימת שונות בסטיית התקן בין חודשי השנה בישראל.
על בסיס הגרף, ניתן לצפות שמודל פולינומי לחיזוי טמפ׳ יצליח טוב יותר בחודשין 4-9 מאשר בשאר חודשי השנה. בחודשים אלו קיימת שונות נמוכה יותר בין טמפ׳ היומיות באותו החודש ולכן המודל יוכל לזהות טוב יותר את המגמה הכללית ולפסק תחזיות מדויקות יותר.
לעומת זאת, בשאר החודשים קיימת שונות גבוהה יותר בין הטמפ׳ היומיות הנמדדות, מה שיכול להוביל לתחזיות פחות מדויקות.

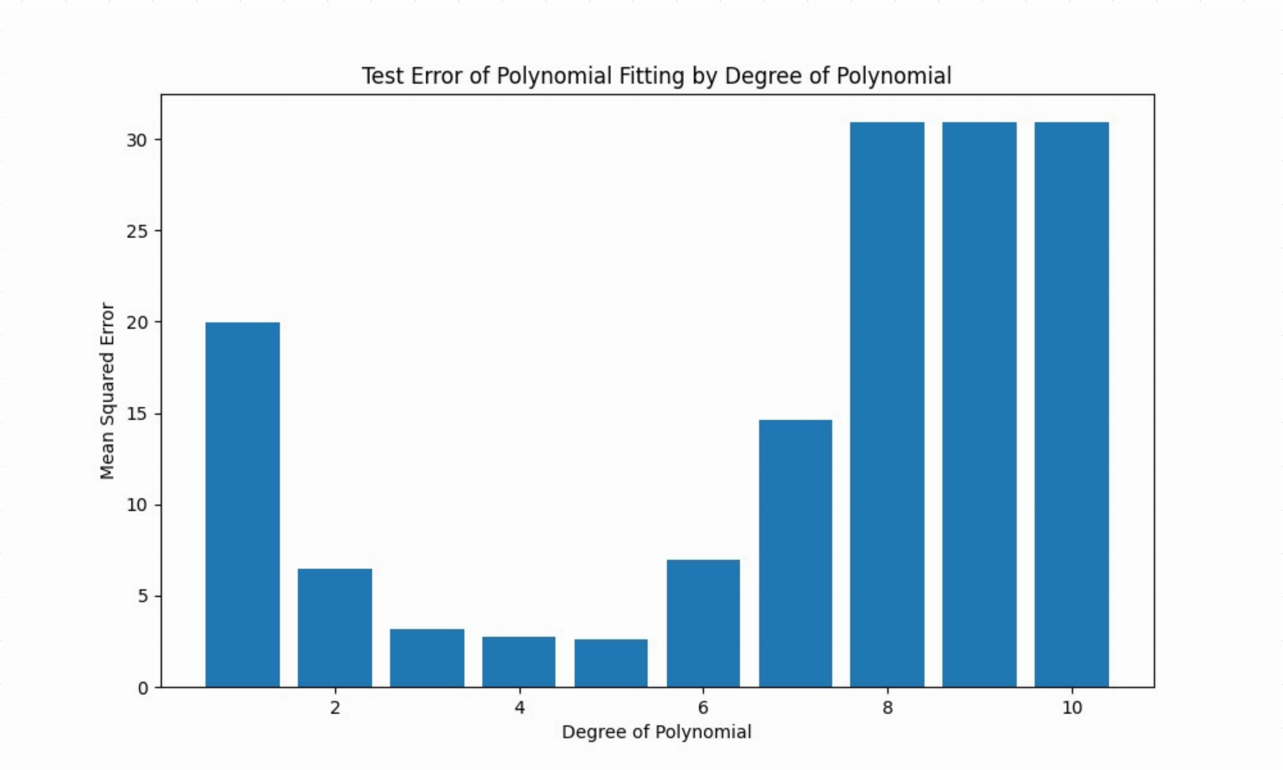


3.2.4 הגרף המוצג מטה מציג את הטמפ׳ החודשיות לאורך השנה במדינות ישראל, ירדן, דרום אפריקה והולנד.
מהגרף ניתן לראות שלא כל המדינות חולקות את אותו דפוס טמפרטורות ממוצע לפי חודשי השנה.
ישנן דמיון בין דפוסי הטמפרטורות של ישראל וירדן, אך דפוסים שונים לגמרי עבור דרום אפריקה או סדר גודל של טמפ׳ שונה בהפרש ניכר עבור הולנד.
עבור ישראל וירדן, ניכרת מגמה דומה של עלייה בטמפרטורות ממוצעות באביב (מרץ-מאי) עד לכ30 מעלות, שמגיעה לשיא בקיץ (יוני-אוגוסט), ואז ירידה בטמפרטורות בסתיו (ספטמבר-נובמבר) והחורף (דצמבר-פברואר) אל בין 12 ל15 מעלות. דפוס זה מצביע על אקלים ממוזג ים-תיכוני דומה במדינות אלו.
לעומת זאת, בדרום אפריקה מופיעה מגמה הפוכה, עם טמפרטורות גבוהות בחורף (יוני-אוגוסט) וטמפרטורות נמוכות יותר בקיץ (דצמבר-פברואר), המאפיינת אקלים טרופי, וכן בהולנד, שם ניכר שקיימת מגמה דומה לאורך השנה, אך קיים הפרש גבוה בין הטמפרטורות בהולנד לטמפרטורות בישראל אשר מגיעות בהולנד לכ17 בשיא ו3 בשיא החורף.
לפיכך, ניתן לצפות שמודל פולינומי שמותאם לנתוני ישראל יתאים גם היטב לחיזוי טמפרטורות בירדן, מכיוון שהם חולקים מגמה עונתית דומה. אך מודל זה לא יתאים כל כך טוב לחיזוי בדרום אפריקה בשל המגמה השונה לחלוטין שם וכן בהולנד עקב הפרשי הטמפ׳.

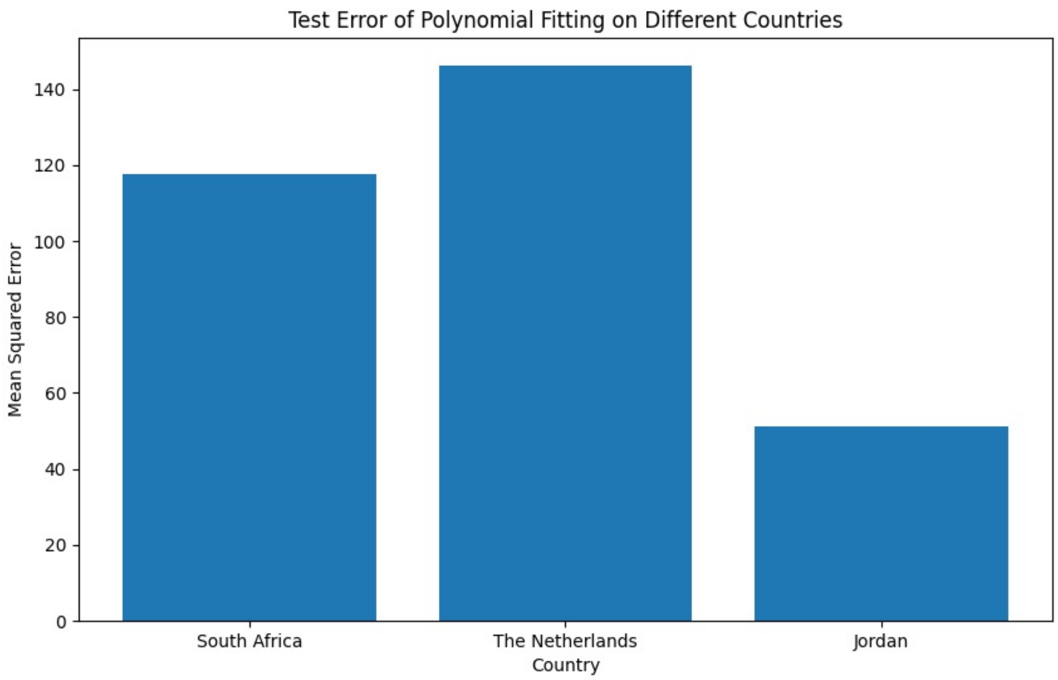


3.2.5 הגרף המוצג מטה מציג את מידת השגיאות בחיזוי המודל עבור דרגות שונות של הפולינום עליו המודל מבוסס.
מהגרף וערכי שגיאת הבחינה, ניתן לראות שהשגיאה יורדת ככל שדרגת הפולינום עולה מ-1 ל-4, ומעבר לדרגה 5, שגיאת הבחינה מתחילה לעלות שוב. זה מרמז שפולינום מדרגה 4 או 5 משיג את הtest error הנמוך ביותר ומספק ככל הנראה את ההתאמה הטובה ביותר לנתונים.
על פי העיקרון של בחירת המודל הפשוט ביותר כאשר מספר מודלים משיגים ביצועים דומים, נבחר בפולינום מדרגה 4 (k=4) כהתאמה הטובה ביותר לנתונים, מכיוון שהוא בעל שגיאת הבחינה הנמוכה ביותר מבין כל הדרגות שנבחנו.
עם זאת, חשוב לציין שההבדל בשגיאת הבחינה בין דרגות 3 (k=3) 4 ו5 הוא יחסית קטן. לכן, פולינום מדרגה 3 יכול גם להיחשב כמועמד פוטנציאלי.

לסיכום, על בסיס המידע שסופק, הפולינום מדרגה 4 (k=4) נראה כהתאמה הטובה ביותר לנתונים, עם שגיאת הבחינה הנמוכה ביותר. עם זאת, פולינום מדרגה 3 (k=3) יכול גם להיות בחירה סבירה, בהתחשב בשגיאת הבחינה הגבוהה במעט אך עדיין נמוכה יחסית שלו ובמורכבות הפוטנציאלית הנמוכה יותר לעומת דרגה 4.



3.2.6 הגרף מציג את שגיאת המודל עבור חיזוי טמפרטורות ממוצעות במדינות שונות בעזרת מודל אשר נבנה על סמך נתוני ישראל.
התוצאות מראות כי המודל מצליח לנבא טוב יותר בירדן בשל הדמיון בדפוסי הטמפרטורות בין ישראל לירדן.
לעומת זאת, השגיאה בהולנד גבוהה מכיוון שההבדל בטמפרטורות המקסימליות משמעותי – בישראל הטמפרטורה בשיא הקיץ מגיעה ל-30 מעלות, בעוד שבהולנד היא מגיעה רק ל-17 מעלות. בדרום אפריקה השגיאה גבוהה אף היא, בשל האקלים הטרופי השונה.
התוצאות מתאימות בצורה יפה אל הדפוסים שזיהינו בתשובה לשאלה 3.
מהגרף המתקבל נובע כי עבור מדינות עם אקלים ודפוסי טמפרטורות שונים, עדיף להתאים מודלים נפרדים המבוססים על נתוניהן המקומיים.



בזמן העבודה התכנותית הפרקטית השתמשתי בchatgpt 3.5 בעזרה בסינטקס ושימוש בספריות matplotlib וsklearn כאשר השתמשתי בהם. כמו כן, נעזרתי בו על מנת למצוא באגים או תקלות בקוד כאשר היו כאלה ולא הצלחתי למצוא את השורה הבעייתית בקוד.