מבוא למערכות לומדות - תרגיל 2 חלק תיאורטי

[illegible] 208648154

1) טענה: $\ker(X) = \ker(X^TX)$

הוכחה: נוכיח הכלה דו-כיוונית.

$\subseteq$: נניח כי $v \in \ker(X)$ אזי $Xv = 0$

נכפיל ונקבל כי

$$(X^TX)v = X^T(Xv) = X^T \cdot 0 = 0$$

$\Leftarrow$ $v \in \ker(X^TX)$ ולכן $\ker(X) \subseteq \ker(X^TX)$

$\supseteq$: נניח כעת כי $v \in \ker(X^TX)$, צ"ל כי

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

הנחנו כי $X \neq 0$ ולכן $X^T \neq 0$ $\Leftarrow$ $Xv = 0$ זה יתקיים

הכלה, קרי $v \in \ker(X) \Leftarrow \ker(X^TX) \subseteq \ker(X)$

$\Downarrow$

$$\ker(X) = \ker(X^TX)$$

מש"ל! $\blacksquare$

2) תהי A מט' ריבועית

טענה: $Im(A^T) = \ker(A)^\perp$

הוכחה: יהי $x \in Im(A^T)$ ו- $y \in \ker(A)$ צריך להראות

$$\langle x, y \rangle = 0$$

ובכך נוכיח הכלה דו-כיוונית:

$\subseteq$: ראשית, כי ניתן לכתוב $x = A^Tu$ נכתוב ונקבל

$$x^Ty = x^TA^Tu = (Ax)^Tu$$

אך כפתרון $x \in ker(A)$ ולכן $Ax=0$

⇐ $(Ax)^T u = 0$

צ"ל כי קיבלנו ש- $Im A^T \subseteq ker(A)^\perp$

ב: כעת נבחר $z \in Im(A^T)^\perp$ וצ"ל כי

$$z^T y = 0$$

אך

$$z^T(A^T u) = (Az)u = 0 \,:$$

מכך קיבלנו כי $z \in ker A$ ⇐

⇐ $Im(A^T)^\perp \subseteq ker(A)$

מעבר דו כיווני
כלומר ערך בו!

⇐ $Im(A^T) \subseteq ker(A)^\perp$

$\Downarrow$

קיבלנו כי $Im(A^T) = ker(A)^\perp$ כנדרש!

3) יהי $y = Xu$ תהי מערכת משוואות עם הנעלמים.
ו- X מט' המקדמים של המערכת.

צ"ל: שלמערכת משוואות יש פתרונות אמ"מ $y \perp ker(X^T)$

הוכחה:

נזכיר כי עמודות ה- A של המערכת צ' תחום של
מטריצה ולכן מקדמים 2 מצבים או יש ∞ פתרונות
או 0.

כי יש ∞ פתרונות
עם מקדם והאר חזיה.

⇐ נקבל את שרשרת הגרירות:

סעיף ב'

⇐ $y \in ker(X^T)^\perp \iff y \in Im(X) \iff$ יש ∞ פתרונות

$\Leftarrow y \perp \ker(X^T)$

$\blacksquare$

4) נבדוק שמתקיים, אילו X^TX הפיכה אז פתרון ריבועים פחותים
אלגברה לינארית כי למערכת קיים פתרון יחיד (הרי
הוא נמצא).
אחרת X^TX אינה הפיכה, נ"ל שהדרגה לא מלאה
ולכן קיים לה פתרון לא יחיד.

נזכר כי $\ker(X^TX) = \ker(X)$

לכן בפרט $v \in \ker(X^TX) \Leftarrow v \in \ker(X)$

ונקבל כי $v \perp \ker(X)^{\perp}$

נ"ל מכיוון שקיים $x \neq 0$ בגרעין אז יש פתרון ה-
$\ker(X)^{\perp} \perp x$ אך הוכחנו כי הפתרון ה"ל אחר כי
בשאלה יש אינסוף פתרונות (3), כנדרש!

$\blacksquare$

ב.) (ז) לפי הגדרה זו ρ מוגדרת ע"י מכפלה חיצונית
ולכן בהינתן בסיס אז $(v_1,\dots,v_k)$ הינו כי ρ מוגדרת
כסכום φ ע"י כאשר σ ע"י היא $v_i\otimes v_i$ $1\le i\le k$.

מכך נשים לב כי

$$\rho_{ij}=\sum_{t=1}^{k} v_{i,t}\cdot v_{j,t}$$

כאשר $v_{i,t}$ הינה
הקואורדינטה ה-i בוקטור v_t $1\le t\le k$, $1\le i,j\le d$.
$\Leftarrow$ נשים לב כי $\rho_{i,j}=\rho_{j,i}$, דבר נובע מהסימטריות
φ בפונקציה הכפל.
קרי, ρ סימטרית כנדרש!

ח) נוכיח כי הע"ע φ ρ הינם $1,0$ ו-4 הו כבסיס תמונה 3
שנראה אותה לעיל. נרחיב את הבסיס $B=(v_1,\dots,v_k)$
לבסיס למרחב ע"י תהליך גרם-שמידט כך שנקבל
$B'=(v_1,\dots,v_k,u_{k+1},\dots,u_d)$ ונראה מהן הע"ע של ρ עם וקטור
הבסיס.

והי $V=\mathrm{Span}(v_1,\dots,v_k)$, $W=\mathrm{Span}(u_{k+1},\dots,u_d)$

נשים לב כי לכל $v\in\mathbb{R}^d$ ניתן לרשום אותו כסכום
φ $u\in V$ ו-$w\in W$ וכך σ

$$\rho(u+w)=\rho u+\rho w$$

$$\rho u=\sum_{i=1}^{k} v_i v_i^T\cdot u \Rightarrow u=\alpha_1 v_1+\dots+\alpha_k v_k \quad \Leftarrow$$

$$\Rightarrow \sum_{i=1}^{k} v_i\sum_{j=1}^{k} v_i^T\cdot\alpha_j v_j=\sum_{i=1}^{k} v_i\cdot\sum_{j=1}^{k}\alpha_j\cdot\delta_{ij}=$$

$= \sum_{i=1}^{k} \alpha_i v_i = u$

⇐ כי לכל $u\in V$ מתקיים כי $Pu=u$, וקרי

$\underline{\lambda=1}$

ולכל $u\in W$

$$Pu = \sum_{i=1}^{k} v_i v_i^T \sum_{j=k+1}^{d} \beta_j u_j = 0 \Rightarrow \lambda=0$$

כל הוקטורים הם
ניצבים זה לזה.

⇐ מכיוון ש- $\mathbb{R}^d = V\oplus W$ אזי אלו המספרים
היחידים ונקבל כי ע"ע של P הם $0,1$ <u>כנדרש</u>!

III. צ"ל להוכיח כי $\lambda=0,1$ הראנו כי לכל $v\in V$ $Pv=1v$
קרי $Pv=v$ (וזהו בהגדרות כפי ע"ע).

IV.

$$P^2 = P\cdot P = \sum_{i=1}^{k} v_i v_i^T \cdot \sum_{j=1}^{k} v_j \cdot v_j^T =$$

$$= \sum_{i=1}^{k}\sum_{j=1}^{k} v_i \underbrace{v_i^T v_j}_{\text{מ"פ}} v_j^T = \sum_{i=1}^{k}\sum_{j=1}^{k} v_i \underbrace{\delta_{ij}}_{\text{בהגדרת מ"פ של בסיס א"נ.}} v_j^T =$$

בגלל δ
הסכומים
נורכים 0.

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

V. <u>נשים</u> לב:

$$(I-P)P = IP - P\cdot P =$$

$$= P - P^2$$

אך ב- IV הוכחנו כי $P=P^2$ ולכן $P-P^2=0$
קרי $(I-P)P=0$ <u>כנדרש</u>! ☑

⇐ נניח כי X^TX היא הפיכה. נ"ל קיימת ההופכית $[X^TX]^{-1}$. בהרצאה ראינו כי

$$w^* = (X^TX)^{-1}X^Ty$$

⇐ כדי הפתרונות שלנו מגיעים לאותה תשובה, נראה כי $(X^TX)^{-1}X^T = X^+$ ⇐ מה שיוכיח כי האלו אותו פתרון.

נזכר כי:

$$X = U\Sigma V^T$$
$$X^+ = V\Sigma^+U^T$$

כמו כן, מכיוון ש- X^TX הפיכה אז אנו יודעים כי ערכי $\sigma_i \neq 0$, קרי $\Sigma^+ = \Sigma^{-1}$ דבר שיועיל לנו בהמשך ההוכחה.

⇐

$$X^TX = (U\Sigma V^T)^T(U\Sigma V^T) =$$
$$= V\Sigma^TU^TU\Sigma V^T = V\Sigma^2V^T$$

⇐ זהו פירוק SVD גם כן ולכן לפי מה שראינו, בהינתן פירוק זה

$$(X^TX)^{-1} = V(\Sigma^2)^{-1}V^T$$

⇐

$$(X^TX)^{-1}X^T = V(\Sigma^2)^{-1}V^TV\Sigma U^T =$$
$$= V(\Sigma^2)^{-1}\Sigma U^T$$

ונשים לב כי

$$(\Sigma^2)^{-1} = \begin{pmatrix} 1/\sigma_1^2 & & \\ & \ddots & \\ & & 1/\sigma_d^2 \end{pmatrix}$$

ומכיוון ש- $(\Sigma^2)^{-1}$ ו- Σ אלכסוניות מתקיים כי

$$(\Sigma^2)^{-1}\cdot\Sigma = \begin{pmatrix} 1/\sigma_1 & & \\ & \ddots & \\ & & 1/\sigma_d \end{pmatrix} = \Sigma^{-1} =$$

$$= \Sigma^+$$

$\Leftarrow$

$$(X^TX)^{-1}X^T = V\Sigma^+U^T = X^+$$

לפי <u>הגדרה</u>.

קרי קיבלנו כי גם במקרה זה הפתרון שהוצג הינו X^+, <u>כנדרש</u>!

■

7) נראה כי X^TX הפיכה אמ"מ $span(x_1,\dots,x_m)=\mathbb{R}^d$.

בתרגיל 1 ראינו כי דרגת X היא דרגת תת-המרחב הנפרש ע"י העמודות, קרי $span\{x_1,\dots,x_m\}$

עכשיו לגבי X^TX מכיוון שראינו שדרגה שווה לדרגת X (הוכח בתרגיל 1, ראינו כי הדרגות שוות ולכן לא נפרט את ההוכחה כאן).

זו גם לא נכון דרגה את הדרגה הפיכה.

?

שוויון לא הוסבר

$$X^TX \text{ הפיכה } \Leftrightarrow Rank(X^TX)=d = Rank(X) \Leftrightarrow span(x_1,\dots,x_m)=\mathbb{R}^d$$

מה קשר בין הגדרנו את דרגת הגן!

ז"א כי קיבלנו ש-

$$X^TX \text{ הפיכה } \Leftrightarrow span(x_1,\dots,x_m)=\mathbb{R}^d$$

<u>כנדרש</u>!

■

ב.) עלינו להראות כעת כי בכל מקרה הפתרון שמצאנו $u^1 = X^+y$ הינו המינימלי, קרי אם פתרון $\tilde{u}$ מקיים $\|u^1\| \le \|\tilde{u}\|$.

<u>הוכחה</u>: נזכר כי כבר מצאנו את u^1 בתרגול, ראינו כי כאשר פתרנו את המשוואות הנורמליות בעזרת compact SVD אז קיבלנו תחת הפירוק כי:

$$\Sigma = \begin{pmatrix} S & 0 \\ 0 & 0 \end{pmatrix}, \quad S = diag(\sigma_1, \dots, \sigma_p)$$

ש"ל כי וקטור הפתרון u^1 בקואורדינטות $1, \dots, p$ וקבע באופן יחודי ע"פ הערכים הסינגולריים של X, בנוסף כאשר ושאר מהפירוק זהה נקבל את u^1 ככלל 0 X^+y. ⇐ נזכר גם כי $\tilde{u}$ פתרון ל- $X^TXu = X^Ty$, נקבל את הנוסח מכיוון ש- $\tilde{u} = \tilde{V}\tilde{\Sigma}^{-1}\tilde{U}^T y$.

מכך גם $\tilde{u}$ פתרון אך יודעים כי $u^1_i = \tilde{u}_i$ כאשר $1 \le i \le p$ מהפירוק שראינו בתרגול.

ויתר ה $u_{p+1}, \dots, u_{d+1}$ ערכים חופשיים.

⇐ $\|\tilde{u}\|^2 = \sum_{i=1}^{d+1} \tilde{u}_i^2 =$

$$= \sum_{i=1}^{r} \tilde{u}_i^2 + \sum_{i=p+1}^{d+1} \tilde{u}_i^2 \ge \sum_{i=1}^{p} \tilde{u}_i^2 + 0 = \sum_{i=1}^{p} {u^1_i}^2 =$$

הורדנו פה ערכים אי-שליליים!

$= \|u^1\|$

⇐ $\|\tilde{u}\| \ge \|u^1\|$ כנדרש!

■

מבוא למערכות לומדות 67577
רועי זהבי 208648154

Fitting A Linear Regression Model

1. ראשית נרצה לבנות את הפונקציה $loadData()$ אשר תבצע תהליך קדם-עיבוד למידע שקיבלנו עבור הבתים. אציין באילו פיצ'רים השתמשי, על איזו החלטתי לוותר ואיזה עיבוד בוצע.

- **שדות שלא נעשה בהם שימוש**: $id, Date$ אלו שדות שנועדו ליצר סדר בתוך בסיס הנתונים שקיבלנו – לא רק שאינם קטגורים הם אינם נותנים לנו כל מידע כלל על הנכס ולכן אינם רלוונטיים כלל למודל הלינארי שלנו.

  כמו כן, החלטתי כמו כן להתעלם מעמודות **שקשורות למיקום** – בעולם האמיתי למקום הנכנס יש חשיבות גבוהה מאוד ולכן אלו שדות שנרצה מאוד שישפיעו על המודל שלנו, אך במקרה זה בו אין אנו יודעים את המיקום או הקשר הספציפי שלו לשווי הנכס אי הכנסתו לשקלול לא עלולה להשפיע רבות על יכולת החיזוי של המודל שלנו (ניתן לדוגמא למצע את המיקום ולהסתכל על המרחק בין הנקודות, אך שוב אין זה נותן אינדיקציה).

  שיקול דומה הפעלתי עבור שדה ה-$zipecode$, מהניתוח עולה כי ישנה רק דגימת אחת עם קוד 0, אך זו דגימה לא תקינה (20671) ולכן אתמודד עם מקרים אלו בעיבוד שנבצע בהמשך.

  שדות sqft_living15 ו- sqft_lot15, ממעבר על המידע שהתקבל, נראה כי שדות אלו רגישים לגודל המרתף או השטח העליון אם זה קיים. לעיתים מתקיים קשר ישיר ולפעמים יש תנודות במידע שלא ניתן להבין מהו הקשר. מכיוון שאנו משתמשים בכל שאר השדות, החלטתי לוותר על חישוב זה ולהתחשב במדידה המקורית, כאשר הגדלים האחרים שמסופקים לנו ישפיעו על המודל באופן זהה.

- **שדות שנלקחו כפי שהם**: מספר החדרים, קומות, מספר חדרי השירותים דירוגים כאלו ואחרים שמופיעים (כמו מצב, ציון ונוף) אלו שדות שכנראה יתארו יחסית בצורה טובה את הקשר הליניארי בניהם לבין ה-$response$ שלנו – המחיר. לכן, לא נעבד אותם ונשתמש בהם כפי שקיבלנו במודל.

- **שדות שדרושים עיבוד:** נשים לב כי ישנם שדות, כגון שנת שיפוץ וגודל מרתף, אשר מתארים מאורע אם קרה ואם לא מופיע 0. נרצה להימנע ממקרה זה מכיוון שהנ"ל ייצר לנו רעש גבוהה בדגימות שלנו מה שישפיע על יכולת חיזוי המודל. לכן, נדרש לחשוב על כללי עיבוד שיעזרו לנו להחליט מה לעשות בכל שדה:

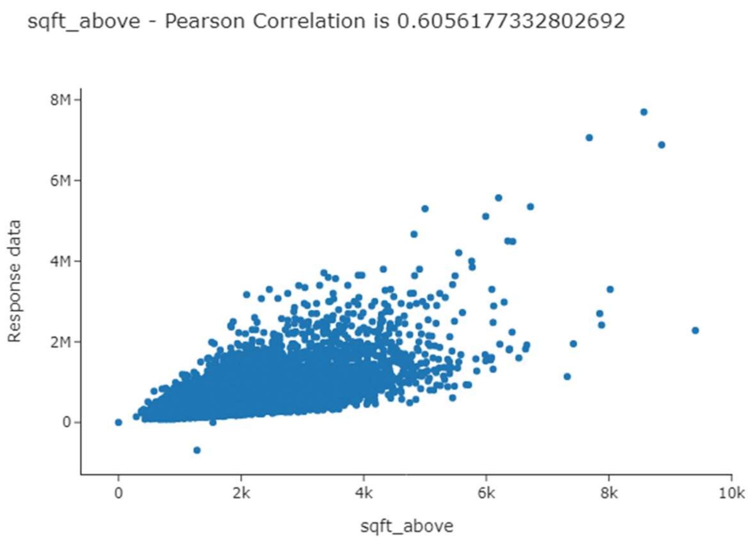
  i. **שנת השיפוץ** – נייצר משתנה דמה $renovated$ אשר יקבל ערכים $\{0,1\}$, כאשר 1 מצביע על כך שהיה שיפוץ ו-0 על כך שלא היה.

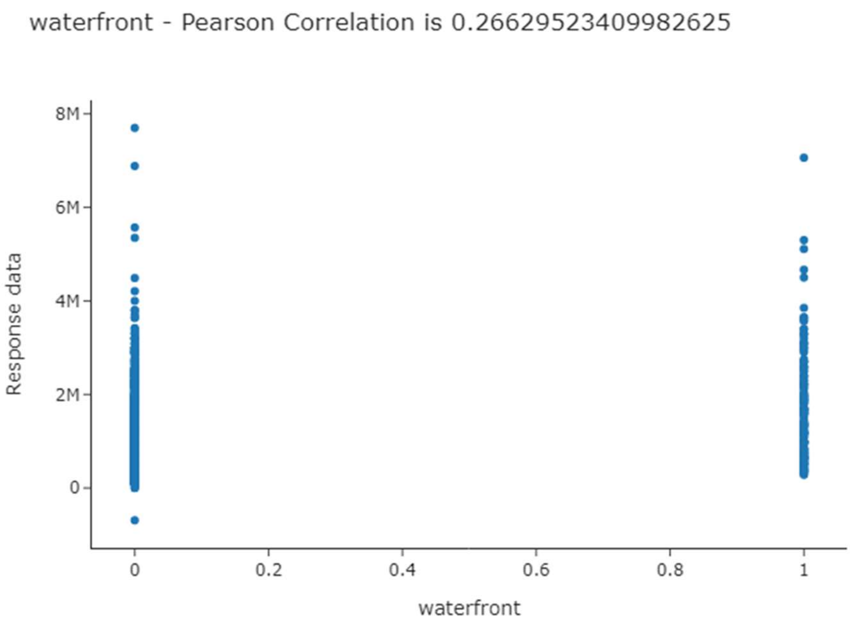
  ii. **גודל המרתף** – כדי להקטין את הפיזור ניתן שוב להגדיר משתנה מסווג עם קיומו של מטרף, אך נוכל לחלק את הגודל ב-1000, ולהקטין משמעותית את הפזור בעת חישוב המודל כך נוכל בכל זאת להתחשב במקרים בהם יש מרתף ממש גדול. חשבתי לבצע מהלך זה מכיוון שלאחר שבצעתי ממוצע לעמודה, מתקבל כי הממוצע הינו 291.482 רגל מרובע, ולכן חלוקה ב-1000 תשמר אותנו יחסית קרובים ל-0 אך בכל זאת תייצר את ההבדל שהקטן שאנו רוצים עבור קיום המרתף. **נשים לב** כי זהו לא המצב בגודל השטח העליון ולכן לא נשנה שדה זה, כאשר ה-0 היחיד שהם הינו טעות במידע.

iii. **שנת בנייה** – נרצה להמיר את שנת הבנייה לגיל הנכס כאשר ככל שהנכס יותר חדש כך נקבל ערך גבוהה יותר, קרי $\frac{1}{2022-age}$ תוך וידאו שאין אנו מחלקים מ-0 (לא בוצעה בדיקתי כי בדקתי במידע כי זה לא ייתכן).

2. נתבונן בשני פיצ׳רים מתוך מטריצת הדגימות שלנו. נבחר את עמודת sqft_above ו – waterfront. מכיוון שהמודל שלנו מחפש קשר ליניארי, ניתן לראות כי שדות בהם יש קשר לינארי כלשהו (כמו גודל מול מחיר) הינם שדות שיועילו יותר למודל שלנו מאשר שדות בינארים, כגון יש/אין מרתף.

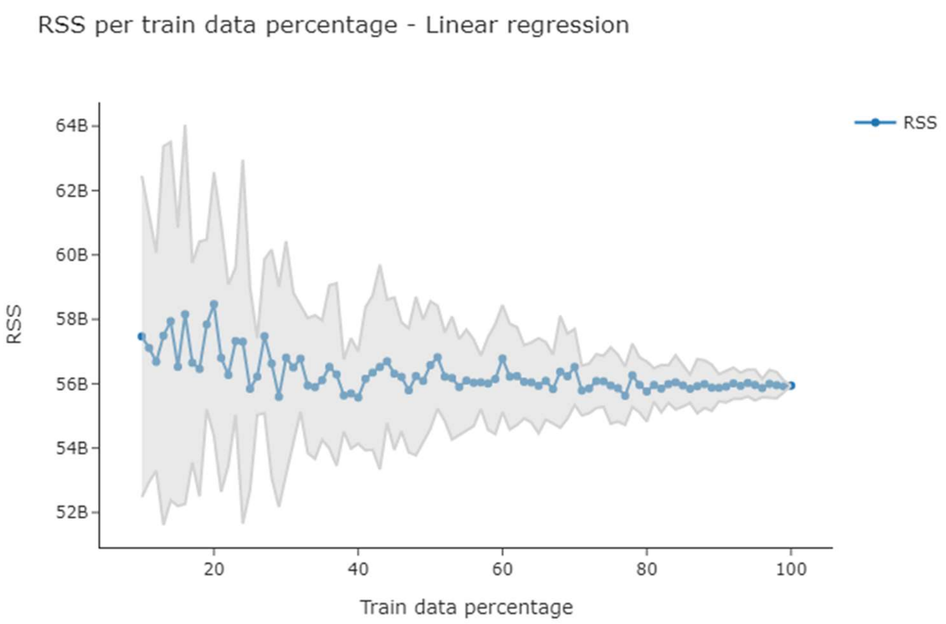
מתאם פירסון שחישבנו נותן לנו את עוצמת הקשר הליניארי בין X, Y (חיובי, שלילי או לא קיים). ניתן לראות זאת בברור שנסתכל על הפלטים שהוצאנו בחישוב של המתאם לכל פי׳צר.
עבור הפי׳צרים שנבחרו מתקבל:





נשים לב כי ההערכה שלנו נכונה ובאמת המתאם עומד מול מה שציפינו, עבור פיצר גודל השטח העליון, נקבל כי המתאם גובהה יותר – זאת גם בהתאם למגמה הכללית שאנו רואים בגרף, ביחס מתאם נמוך עבור פיצ׳ר בינארי שלא מגלם כל קשר לינארי.
לכן, ניתן לראות כי באמת שדה waterfront פחות יועיל למודל שלנו מכיוון שהוא לא תורם למציאת הקשר הלינארי, זאת ביחס לקשר שברור נוכח בשדה השטח.

3. הפונקציה המתאימה מומשה בקבצים
4. מצב גרף הפלט עבור שאלה זו.



נתבונן בגרף הפלט, באופן ישיר אנו רואים ככל שאנו מאמנים את המודל שלנו על יותר מידע, כך השגיאה מתכנסת וקטנה לדיוק המודל, שגיאה ריבועית של כ-56 מיליארד, שאילו נוציא שורש נקבל כי השגיאה MSE הינה 230~ אלף דולר בהערכת מחיר בית עבור מידע שעוד לא נראה לפני.

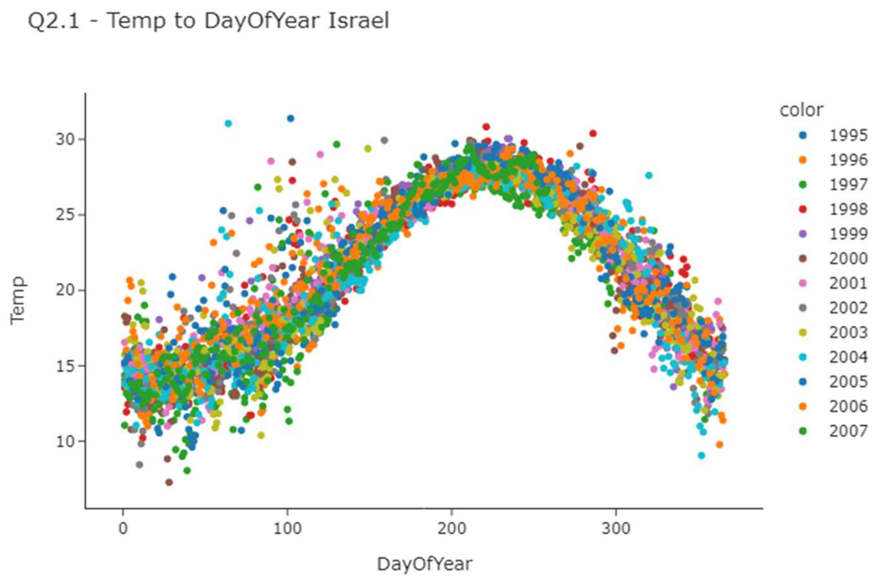
מכיוון שלא הגדרנו מהי הטעות שאנו מאפשרים למודל להשיג, מבחינתו זה נתון נוסף שניתן ללמוד ממנו ובהמשך לשנות אולי את פרמטרי עיבוד המידע שלנו, או לנסות לשנות פיצרים נוספים ע״מ לקבל אומד טוב יותר, אם זה אפשרי – תוך ניסיון להימנע מלבצע overfit, כאשר כפי שציפינו שימוש בכל סט האימון הקטין את השגיאה וצמצם את מרחב הבחירה של המודל.

כמו כן, מבחינת confidence interval – ראינו בהרצאה כי זה נתון שמאפשר לנו להמחיש את השונות, כאשר אנו יכולים ללמוד מן הפלט כי תחום זה הולך ומצטמצם ככל שאחוז עולה, קרי מתקיים כי היכולת של המודל לחזות עבור $\hat{y}$ ערכים גדולים ורחוקים יותר מצטמצם משמעותית, מה שמאפשר לנו לשפר את הדיוק, תחת סט האימון שקיבלנו, אך במחיר של גמישות במודל (אותו תשלום שראינו בין ה-$bias$ ל-$variance$).

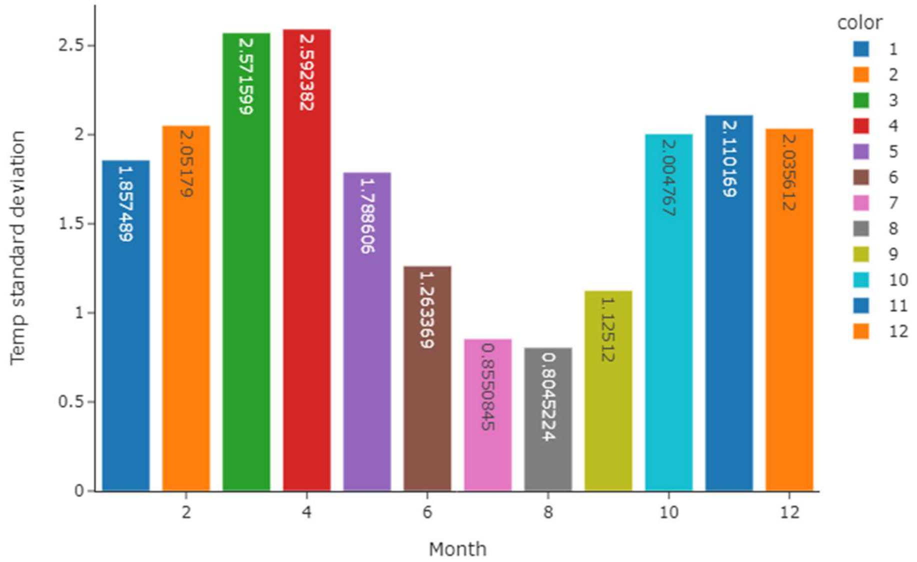
Polynomial Fitting

1. מומש בקבצים.
2.

a. מצ״ב גרף הפלט של הסעיף המבוקש. ניתן לראות כי הגרף מייצג מגמה כללית של פולינום מדרגה 3 או 4, על בסיס העיקול וירידה בגודל השיפוע כאשר הערכים מתרכבים ל-0.



b. להלן גרף המייצג את סטיית התקן ביחס לחודשים כפי שהתבקשנו לעשות

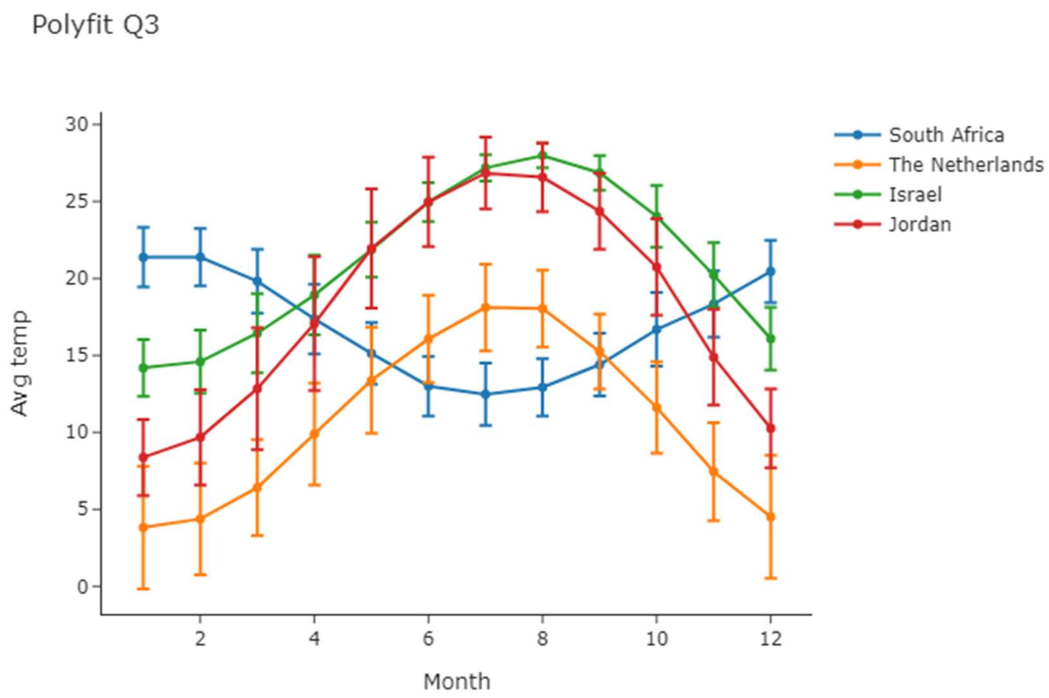


כפי שנתן לראות בגרף, אנו יכולים להעריך כי המודל שלנו לא יוכל לבצע פרדיקציה זהה על כל החודשים, מכיוון שסטיות התקן משתנות. ראינו כי סטיית התקן והשונות הן המדד לנו להבין את פזור המידע ולכן כ״כ שאלו גבוהים יותר אנו למדים כי המידע שיקבלנו יותר

מפוזר ולכן יהיה יותר קשה לקבל פרדיקציה טובה עבור מודלים שמתבססים על רגרסיה ליניארית.

לכן, לפי תובנה זו ניתן לראות כי המודל שלנו יתקשה לבצע חיזוי מדויק בחודשי החורף ביחס לחודשי הקיץ – תובנה הגיונית עבור המעלות שבישראל, כאשר ישנן שנים עם שינויים גדולים בחורף ביחס לקיץ שהמעלות בו יחסית זהות.

3. להלן הפלט הנדרש:

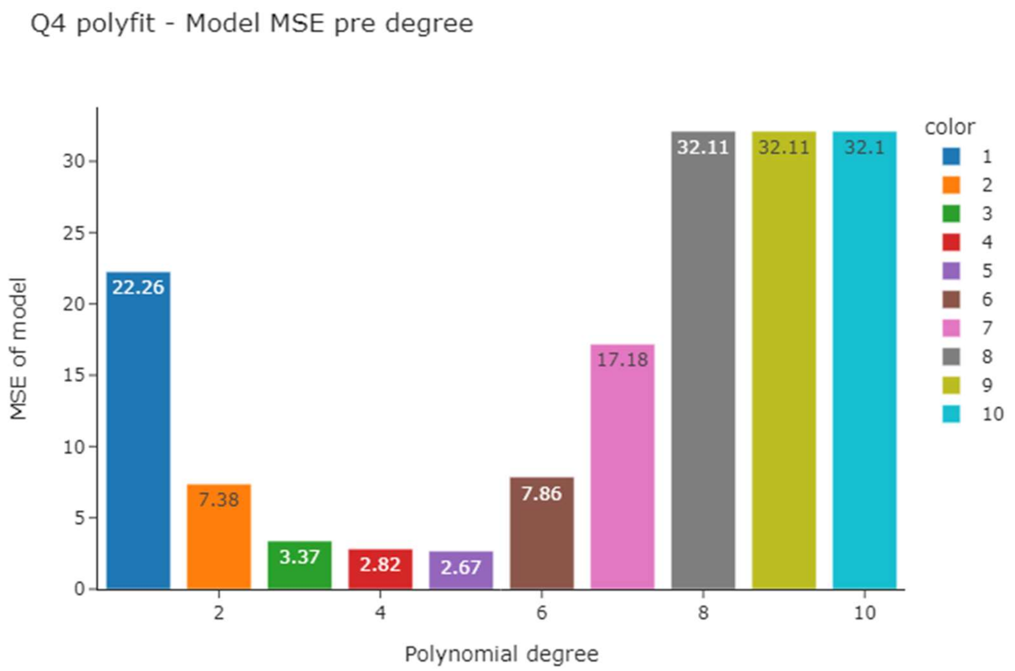


בהתבסס על גרף זה ניתן לראות כי לא לכל המדינות הדפוס זהה, לדוגמא דרום אפריקה הפוכה מישראל, ירדן או הולנד (דבר הגיוני שאנו רוצים לראות מכיוון שזו בצד השני של קו המשווה). לכן, ניתן לראות כי אילו נתאים מודל עבור ישראל בלבד, בהסתברות גבוהה מאוד שהמודל יספק פרדיקציה טובה מאוד עבור ירדן – ניתן לראות כי תוך התחשבות בסטיות התקן, קיימת חפיפה רבה בין ממוצעים ולכן הגיוני כי המודל יאפשר לקבל התאמה טובה גם במקרה זה.

מנגד, המודל לא יעבוד טוב כלל עבור דרום אפריקה מכיוון שסט האימון עבורו לא רלוונטי כלל. הנ״ל מתנהג בצורה שונה ועל כן לא נוכל להתבסס – שזו סה״כ תובנה שציפינו לה.

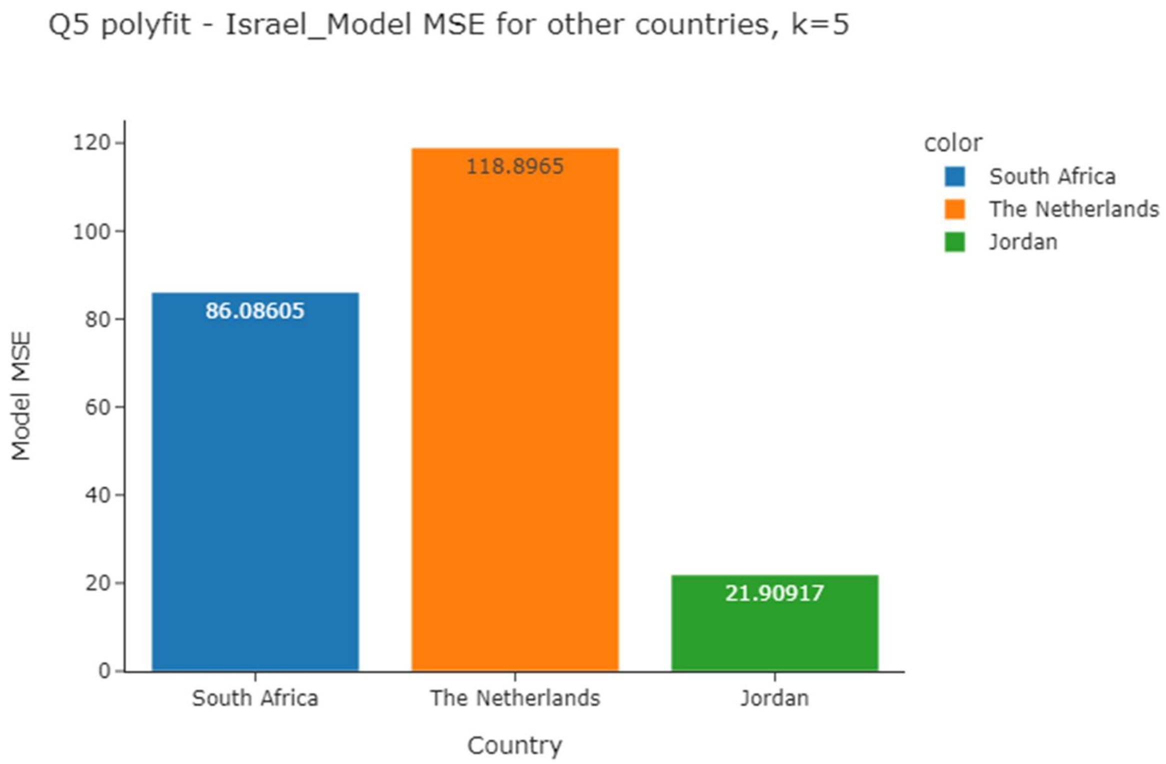
עבור הולנד נשים לב כי הפונקציה בצורתה דומה מאוד לפונקציה שמיוצרת ע״י הממוצעים של ישראל אך קיים פער בין התוצאות. לכן מודל של ישראל כנראה ישגה להשיג פרדיקציה טובה עבור הולנד, אך אילו נצליח לנרמל את הקבוע שמפריד בין הפונקציות נוכל לבצע התאמה ולקבל מודל שחוזה בצורה טובה.

4. להלן גרף הפלט שמסכם את שגיאות המודל תחת כל דרגה



ניתן ראות כי פולינום מדרגה 5 השיג את השגיאה הקטנה ביותר, כאשר דרגה 4 מאוד קרובה בשגיאתה. פולינום מדרגה מאוד גבוהה כבר לא משיגים קירוב מתאים. נשים לב כי לא נרצה להמשיך ולהגדיל את דרגות הפולינום מכיוון שאנו נצבע overfit לסט האימון שלנו – ולכן נשאף להשתמש בפולינום מדרגה 4 או 5 כפונקציית הפרדיקציה שלנו עבור המקרה הנ״ל.

5. להלן פלט שגיאות המודל, כאשר התאמנו פולינום מדרגה 5 על כל המידע שהגיע מישראל



ניתן לראות כי מה שצפינו בשאלה 3 בהחלט מתקיים, המודל הצליח לחזות יחסית בצורה טובה את הטמפרטורה עבור ירדן, אך השגיאה הריבועית עלתה בצורה משמעותית כאשר ניסינו לחזות את עבור דרום אפריקה והולנד.

ניתן גם להסביר מדוע השגיאה מדרום אפריקה אפילו קטנה יותר מזו של הולנד למרות שראינו שהגרפים של ישראל והולנד היו דומים – שגיאה זו עלתה בהפרשי הטמפרטורה, אומנם עונות השנה הפוכות בין ישראל לדרום אפריקה, אך באופן ממוצע ראינו כי הטמפרטורות יחסית קרובות, ביחס להולנד. לכן המודל השיג תוצאות שבאופן כללי קרובות יותר למידע עליו הוא אומן – ולכן השגיאה במקרה זה קטנה יותר.

אילו היינו רוצים, ניתן היה להוסיף התערבות נוספת ולנרמל את תוצאת המודל במרחק הממוצע בין הטמפרטורות ולקבל חיזוי שככה״נ משמעותית טוב יותר עבור הולנד מן המודל הקיים.