# IML תרגיל 2

2.1 (1) הכלות, נראה כי $\ker(X) \subseteq \ker(X^T X)$:

יהי $v \in V$ כך ש $v \in \ker(X)$, כלומר $Xv = 0$.

אז:

$$X^T X v = X^T (Xv) = X^T 0 = 0$$

כלומר, $v \in \ker(X^T X)$.

כעת נראה כי $\ker(X) \supseteq \ker(X^T X)$:

יהי $v \in V$ כך ש $v \in \ker(X^T X)$, כלומר $X^T X v = 0$.

לכן, $v^T X^T X v = v^T 0 = 0$

וגם $v^T X^T X v = \langle Xv, Xv \rangle = 0$.

מתכונות מכפלה פנימית, $\langle Xv, Xv \rangle = 0 \iff Xv = 0$,

לכן $v \in \ker(X)$.

כיוון שהראינו הכלה בשני הכיוונים, נובע כי $\ker(X) = \ker(X^T X)$

כנדרש.

(2) ראשית, נראה כי $\ker(A)^\perp \supseteq Im(A^T)$:

יהי $v \in V$ כך ש- $A^Tv = y \neq 0$. יהי

$w \in V$ כך ש- $w \in \ker(A)$. נראה כי $w^Ty = 0$:

$$w^Ty = w^TA^Tv \underset{\text{תכונות שחלוף}}{=} (Aw)^T v \overset{w\in\ker(A)}{=} 0v = 0$$

כלומר, $y \in \ker(A)^\perp$.

כעת, נראה כי $\ker(A)^\perp \subseteq Im(A^T)$:

יהי $v \in \ker(A)$ אזי מתקיים:

$$v \in \ker(A) \iff Av = 0 \iff \forall u \in V \ \langle Av, u\rangle = 0 \iff \langle v, A^Tu\rangle = 0 \iff$$

$$\iff v \perp A^Tu \iff v \in Im(A^T)^\perp$$

לכן, $\ker(A) = Im(A^T)^\perp$. ולכן אם ניקח את המשלים

ממשפט 11 בתרגול 2 נקבל כי $\ker(A)^\perp = Im(A^T)^{\perp\perp} = Im(A^T)$

בהשאלה כי לכל $B$ תת מרחב:

ולכן $\ker(A)^\perp \subseteq Im(A^T)$ כנדרש. כלומר, הראינו כי $\ker(A)^\perp = Im(A^T)$

∎

3 $X$ ריבועית ואינה הפיכה $\Leftarrow$ ל-$X$ קיימים $\infty$ פתרונות

או $0$ פתרונות.

כעת, נבחין כי $Xw=y$ ולכן למערכת קיים פתרון

נבחין אם ($y \in Im(X)$ כי). כלומר, נשים לב כי מתקיים

$y \in Ker(X^T)^{\perp} \overset{\text{משלים אורתוגונלי}}{\iff} y \in Im(X) \iff$ למערכת $\infty$ פתרונות

$\iff y \perp Ker(X^T)$

כנדרש

■

(4)

ראשית, נניח כי $X^TX$ הפיכה.

אז $(X^TX)^{-1}$ קיים ומתקיים $(X^TX)^{-1}X^TX=I_n$

לכן, עבור $X^TXw=X^Ty$, נכפיל את שני האגפים

ב- $(X^TX)^{-1}$ ונקבל כי הפתרון יחיד ולמעשה הוא:

$$(X^TX)^{-1}X^TXw=(X^TX)^{-1}Xy \iff w=(X^TX)^{-1}Xy$$

כעת, אם $X^TX$ אינה הפיכה:

ממשפט הדרגה, למערכת יש אינסוף פתרונות

אמ"מ $X^Ty \perp \ker(X^TX)$, ולכן מסעיף 1 $\ker(X^TX)=\ker(X)$

כלומר למערכת יש אינסוף פתרונות אמ"מ $X^Ty \perp \ker(X)$.

נראה כי $X^Ty \perp \ker(X)$:

יהי $v\in\ker(X)$. אז:

$$\langle v, X^Ty\rangle = \langle Xv, y\rangle = \langle 0, y\rangle = 0$$

כלומר, $X^Ty \perp \ker(X)$ ולמערכת יש $\infty$ פתרונות כנדרש.

2.2 (5) (a) צ"ל: $V_i = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_k \end{pmatrix}$ וקטור אורתונורמלי

נכפול הוקטור בעצמו ונקבל P:

$$V_iV_i^T = \begin{bmatrix} x_1x_1 & x_1x_2 & x_1x_3 & \cdots & x_1x_k \\ x_2x_1 & \ddots & & & \vdots \\ x_3x_1 & & \ddots & & \vdots \\ \vdots & & & & \vdots \\ x_kx_1 & \cdots & & & x_kx_k \end{bmatrix}$$

כלומר, $V_iV_i^T$ היא מט' סימטרית, ולכן גם

P סימטרית כסכימת מטריצות סימטריות כנדרש.

(b) יהי $v \in V$. אזי $V$ הוא נפרש ע"י $v_1, \ldots, v_k$

כלומר $\sum_{i=1}^{k} a_i v_i = v = a_1v_1 + a_2v_2 + \ldots + a_kv_k$ כש $a_i \in \mathbb{R}$ $\forall a_i$.

לכן,

$$Pv = \left(\sum_{i=1}^{k} v_iv_i^T\right)v = \sum_{i=1}^{k} v_iv_i^Tv = \sum_{i=1}^{k} v_i(v_i^Tv) =$$

$$= \sum_{i=1}^{k} v_i\left(v_i^T\left(a_1v_1 + a_2v_2 + \ldots + a_kv_k\right)\right) = \sum_{i=1}^{k} v_i\left(a_1v_i^Tv_1 + a_2v_i^Tv_2 + \ldots + a_kv_i^Tv_k\right)$$

נשים לב ש $V_i^T V_j$ הוא המכפלה הפנימית הסטנדרטית,

ונתון ש- $V_i, V_j$ הם אורתונורמליים ולכן

$$V_i^T V_j = \begin{cases} 0 & i \neq j \\ 1 & i = j \end{cases}$$

לכן,

$$\sum_{i=1}^{k} V_i \left(a_1 V_i^T V_1 + a_2 V_i^T V_2 + \dots + a_k V_i^T V_k\right) = \sum_{i=1}^{k} V_i \left(a_i V_i^T V_i\right) = \sum_{i=1}^{k} a_i V_i = V$$

כלומר, הבסיס ש-1 הוא ע"ע לכל $V \in V$. בנוסף

גם 0 הוא ע"ע לכל וקטור בו $V$ שייך למשלים האורתוגונלי

של $V$ ולכן

$$\sum_{i=1}^{k} V_i \left(V_i^T V\right) = \sum_{i=1}^{k} 0 V_i = 0$$

כיוון ש $V_i^T V = 0$ לכל $V$ שבמשלים האורתוגונלי של הבסיס $V$

c) הוכחה בסעיף הקודם.

d)

$$P^2 = \sum_{i=1}^{k} V_i V_i^T \cdot \sum_{i=1}^{k} V_i V_i^T = \sum_{i=1}^{k} \sum_{j=1}^{k} V_i V_i^T V_j V_j^T$$

כיוון שלכל $i,j$ $V_i, V_j$ אורתונורמליים, מתקיים

$$V_i^T V_j = \begin{cases} 0 & i \neq j \\ 1 & i = j \end{cases}$$

לכן

$$\sum_{i=1}^{k} \sum_{j=1}^{k} V_i \overbrace{\left(V_i^T V_j\right)}^{\substack{i \neq j \Rightarrow 0 \\ i=j \Rightarrow 1}} V_j^T = \sum_{i=1}^{k} V_i V_i^T = P$$

כנדרש.

e

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

סעיף 4 קודם

$\blacksquare$

$$(X^TX)^{-1}X^Ty = X^+y$$ צ"ל: (6) 2.3

נשתמש בפירוק ל- SVD:

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

$$= \Big(V\Sigma^T \underbrace{U^TU}_{I_n}\Sigma V^T\Big)^{-1}\left(V\Sigma^TU^T\right)y = \left(V\Sigma^T\Sigma V^T\right)^{-1}\left(V\Sigma^TU^T\right)y =$$

$$= V(\Sigma\Sigma)^{-1}\underbrace{V^TV}_{I_n}\Sigma U^Ty = V\begin{pmatrix} \frac{1}{\sigma_1^2} & & & 0 \\ & \frac{1}{\sigma_2^2} & & \\ & & \ddots & \\ 0 & & & \frac{1}{\sigma_n^2} \end{pmatrix}\begin{pmatrix} \sigma_1 & & & 0 \\ & \sigma_2 & & \\ & & \ddots & \\ 0 & & & \sigma_n \end{pmatrix}U^Ty =$$

$A^T = A$ כי

עבור מט' אלכסונית

(*)

$$= V\begin{pmatrix} \frac{1}{\sigma_1} & & & 0 \\ & \frac{1}{\sigma_2} & & \\ & & \ddots & \\ 0 & & & \frac{1}{\sigma_n} \end{pmatrix}U^Ty = V\Sigma^+U^Ty = X^+y$$

כנדרש.

(*) הרכיבים מהאלכסון גם $\sigma_i = 0$ מוגדרים מהווים בעיה

בפתרון. אם $\sigma_i = 0$ אז בפתרון שם $(\Sigma\Sigma)^{-1}_{ii} = 0$ וכו'

לא משפיע על הפתרון.

7) צ"ל: $sp(\{x_1,\dots,x_m\})=\mathbb{R}^d \iff X^TX$ הפיכה.

הוכחה, נשים לב כי $X$ ריבועית כי מימדי $X^TX$ הם $d\times d$

כמו כן מהתרגול, בנוסף, הראינו כי $rk(X)=rk(X^TX)$

לכן:

$\Longrightarrow$: $X^TX$ הפיכה, לכן $rk(X^TX)$ שווה למימד

מימדה $(d=)$, לכן $rk(X)=d \Leftarrow sp(x_1,\dots,x_m)=\mathbb{R}^d$

$\Longleftarrow$: $sp(x_1,\dots,x_m)=\mathbb{R}^d \Leftarrow rk(X)=d \Leftarrow rk(X^TX)=d$

ועל ריבועית שדרגתה שווה למספר שורות היא הפיכה,

לכן $X^TX$ הפיכה.

כלומר, הראינו את הנדרש.

∎

8) יהי $U\Sigma V^T$ פירוק ה-SVD של $X$.

הוכיחו שמטריצה כ $X^\dagger$ מקיימת את התכונות הבאות:

1) $XX^\dagger X = X$

2) $X^\dagger X X^\dagger = X^\dagger$

3) $(XX^\dagger)^T = XX^\dagger$

4) $(X^\dagger X)^T = X^\dagger X$

הוכחה:

1) $XX^\dagger X = (U\Sigma V^T)(V\Sigma^\dagger U^T)(U\Sigma V^T) = (U\Sigma V^T V \Sigma^\dagger U^T U \Sigma V^T) = U\Sigma\Sigma^\dagger\Sigma V^T =$

$= U\Sigma V^T = X \quad \blacksquare$

2) $\underset{\text{הראינו בתרגול}}{X =} (X^\dagger)^\dagger$ נובע ישירות מ-① ובעובדה ש

$$\Sigma\Sigma^\dagger = \Sigma^\dagger\Sigma$$

3) $(XX^\dagger)^T = (U\Sigma V^T V\Sigma^\dagger U^T)^T = (U\Sigma\Sigma^\dagger U^T)^T = (U\Sigma^\dagger\Sigma U^T)^T =$

$= (U\Sigma^\dagger I_n \Sigma U^T)^T = (U\Sigma^\dagger V^T V\Sigma U^T)^T = ((U\Sigma^\dagger V^T)(V\Sigma U^T))^T =$

$= (V\Sigma U^T)^T (U\Sigma^\dagger V^T)^T = (U\Sigma V^T)(V\Sigma^\dagger U^T) = XX^\dagger \quad \blacksquare$

4) $X = (X^\dagger)^\dagger$ נובע ישירות מ-③ ובעובדה ש

כעת, נשים לב כי מתקיים: $\hat{w} = X^{+}y$

$$X^T\left(X\hat{w} - y\right) = X^T\left(XX^{+}y - y\right) \overset{3}{=} X^T\left(XX^{+}\right)^T b - X^T b =$$

$$= \left(XX^{+}X\right)^T b - X^T b \overset{1}{=} X^T b - X^T b = 0 \quad (*)$$

משמע לכל וקטור $\bar{w} \neq \hat{w}$:

$$\|X\bar{w} - y\|^2 = \|X\bar{w} - X\hat{w} + X\hat{w} - y\|^2 = \|X(\bar{w} - \hat{w}) + X\hat{w} - y\|^2 =$$

$$= \|X(\bar{w} - \hat{w})\|^2 + \|X\hat{w} - y\|^2 + 2\left(X(\bar{w} - \hat{w})\right)^T (X\hat{w} - y) =$$

$$= \|X(\bar{w} - \hat{w})\|^2 + \|X\hat{w} - y\|^2 + 2(\bar{w} - \hat{w})^T X^T(X\hat{w} - y) \overset{*}{=} \|X(\bar{w} - \hat{w})\|^2 + \|X\hat{w} - y\|^2 \geq \|X\hat{w} - y\|^2 \quad (\#)$$

בנוסף, לכל וקטור $\bar{w} \neq \hat{w}$:

$$\langle \hat{w}, \bar{w} - \hat{w} \rangle = \hat{w}^T(\bar{w} - \hat{w}) = (X^{+}y)^T(\bar{w} - \hat{w}) \overset{2}{=} \left(X^{+}XX^{+}y\right)^T(\bar{w} - \hat{w})$$

$$= \left((X^{+}X)(X^{+}y)\right)^T(\bar{w} - \hat{w}) = (X^{+}y)^T(X^{+}X)^T(\bar{w} - \hat{w}) \overset{4}{=} (X^{+}y)^T X^{+}X(\bar{w} - \hat{w}) =$$

$$= (X^{+}y)^T\left(X^{+}\underbrace{X\bar{w}}_{y} - X^{+}X\underbrace{\hat{w}}_{X^{+}y}\right) = (X^{+}y)^T\left(X^{+}y - X^{+}XX^{+}y\right) \overset{2}{=} (X^{+}y)^T\left(X^{+}y - X^{+}y\right) = 0 \quad (\star)$$

ולכן מתקיים לכל פתרון $\bar{w} \neq \hat{w}$:

$$\|\bar{w}\|^2 = \|\hat{w} + \bar{w} - \hat{w}\|^2 = \|\hat{w}\|^2 + \|\bar{w}\|^2 + 2\hat{w}^T(\bar{w} - \hat{w}) = \|\hat{w}\|^2 + \|\bar{w}\|^2 + 0 =$$

$$\geq \|\hat{w}\|^2$$

$\blacksquare$

כנדרש

חלק מעשי:

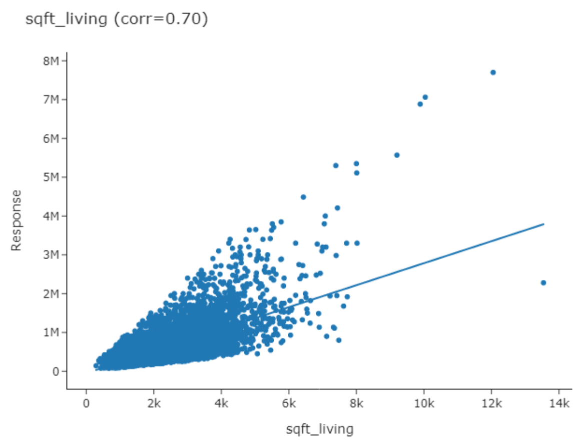
1) - החלטתי להוריד את הפיצ'רים הבאים:

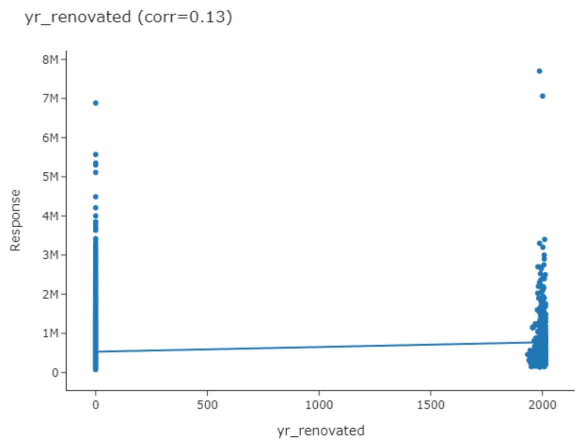
['id', 'sqft_lot15', 'sqft_living15', 'long', 'lat', 'date']

זאת כיוון שלהערכתי הם לא היו רלוונטים לניבוי המחיר הצפוי ל הבית

- לדעתי הפיצ'ר הקטגורי היחיד היה זיפקוד, זאת כיוון שהוא מעיד על מיקום בית שזוהי קטגוריה ולכן היה צורך להשמש ב get_dummies בכדי שניתן יהיה להשתמש בפיצר זה.
- עבור דאטא לא זמין או לא הגיוני ( מס' חדרים קטן מ0 וכו') בסט האימון הסרתי לגמרי. בסט הבחינה אני חישבתי את הממוצע לכל עמודה בסט האימון ובמקומות הלא זמינים בסט הבחינה הכנסתי את הממוצע של אותה עמודה.
- בנוסף, ווידאתי על כל אחד מהשדות שהוא מספרי ובר חישוב

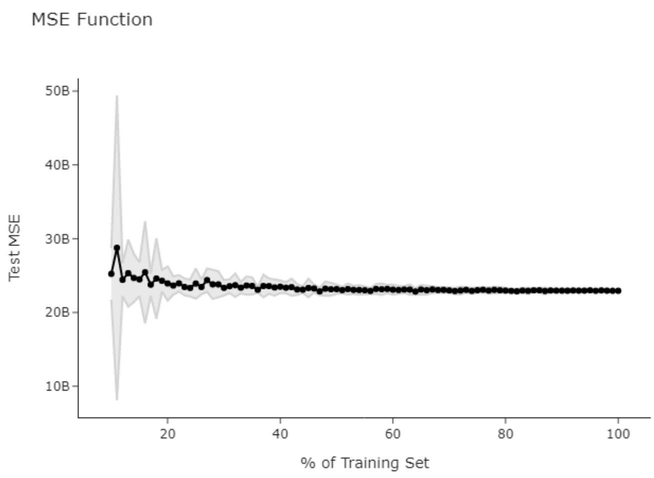
2)



הסקתי האם הם מועילים או לא בכך שניסיתי לראות אם יש צפיפות נקודות גבוהה יותר סביב גרף הפונקציה ככל שמתקרבים אליה ואיש כמות יחסית קטנה של נקודות שרחוקות מהגרף.

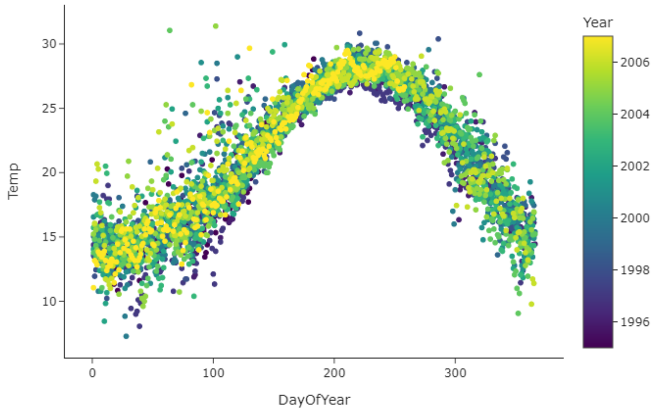
4)




ניתן לראות כי ככל שכמות הדגימות גדלה גם השגיאה של הדגימות קטנה וגם השונות קטנה

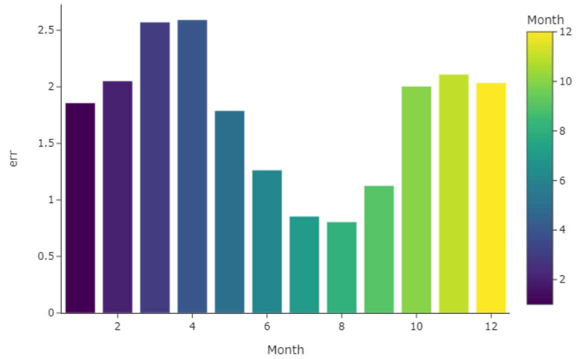
חלק שני:

2)




לפי הגרף, נראה כי דרגת הפונקציה תהיה גדולה מ2, נראה כאילו 3 מתאר את דרגת הפונקציה של הגרף.

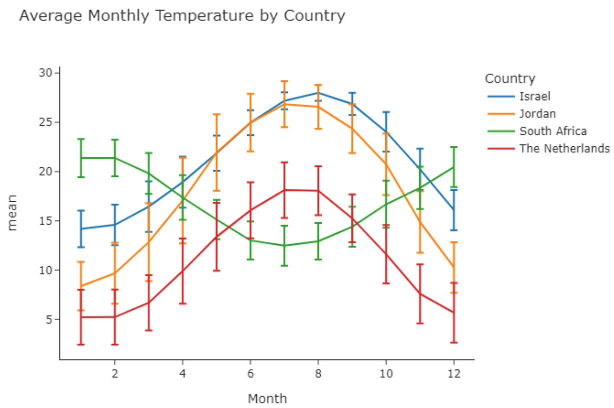
-




-

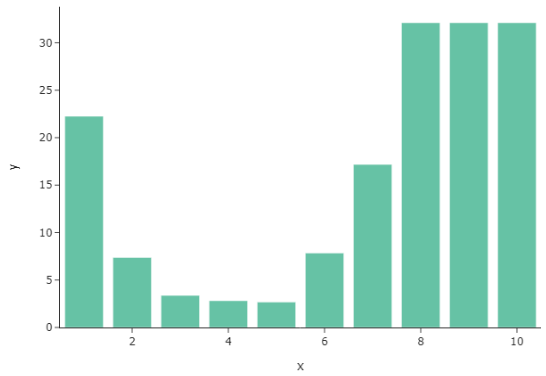
נראה שעבור החודשים 7-8 תהיה יכולת הניבוי הטובה ביותר, ואילו בחודשים 3-4 השונות הגבוהה ביותר ולכן גם יכולת הניבוי טובה פחות.

(3




ניתן לראות כי המדינות שונות ביניהן בגרף המתאר את ממוצע המעלות. ניתן לראות לדוגמא שבדרום אפריקה חודשי החורף והקיץ הפוכים מהולנד.
המדינה שהמודל של ישראל יתאר בצורה הטובה ביותר היא ככל הנראה ירדן, עד כדי כמעט חפיפה בחודשים 5-6. אך בחודשים אחרים (1,12) המודל הישראלי לא מתאר טוב את הטמפרטורה בירדן (בחודש 12 המודל הישראלי ינבא טוב יותר את דרום אפריקה מאשר את ירדן) אם כי צורת המודל עדיין יחסית דומה.

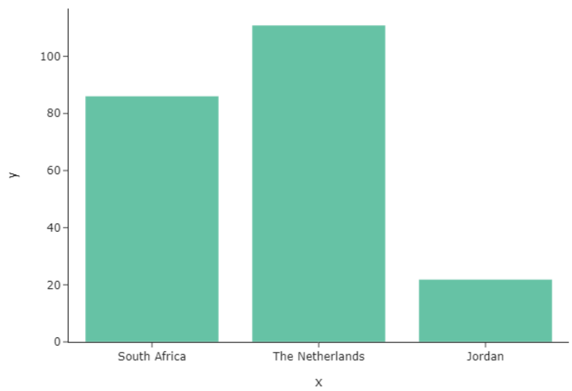
(4




הערכים (בהתאמה): [32.11  32.11  32.11  17.17  7.84  2.67  2.82  3.37  7.38  22.26]

ניתן לראות כי ה-k המתאים ביותר הוא 5 עם שגיאה של 2.67. אך עם זאת גם עבור k של 3 נקבל שגיאה יחסית נמוכה ובהבדל לא גדול מ5

5)




המודל שמתאים בצורה הטובה ביותר למודל הישראלי הוא כפי שנאמר בשאלה 3, ירדן. הגיוני להניח כי חלק מזה הוא בגלל הקרבה בין המדינות