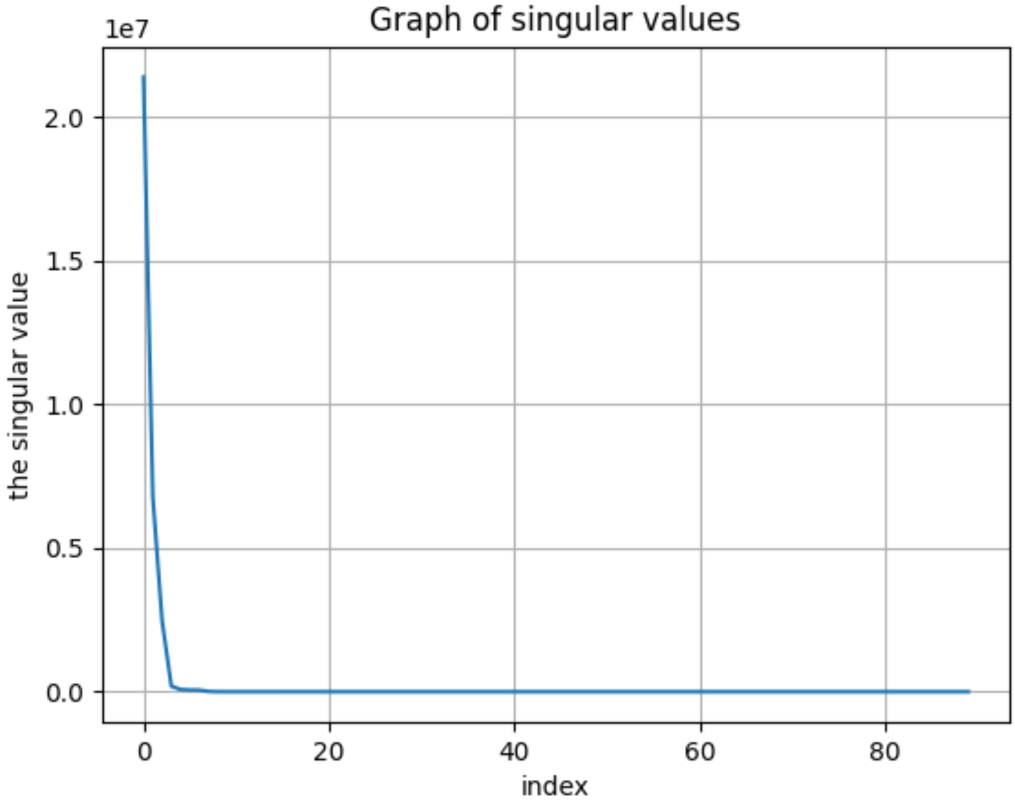
תרגיל בית IML 2

דורון ברודר
204917629
doronbruder

חלק מעשי:

13. תכונת הZIPCODE היא תכונה קטגורית מכיוון שזה כתובת מתוך מרחב כתובות מוגדר של הקודים באותה מדינה . בנוסף זו תכונה בעלת ערך מספרי אך שאינו בכל יחס סדר טבעי שכן מדובר בכתובות שרירותיות , אבל יש להניח כי הוא עלול כן להשפיע על המחיר שכן הוא מקודד את מיקום הנכס ולכן בחרתי בגישת ה-" 1 HOT " וכך באמת ניתן יהיה לגלות קורולציות בין כתובות שונות למחיר

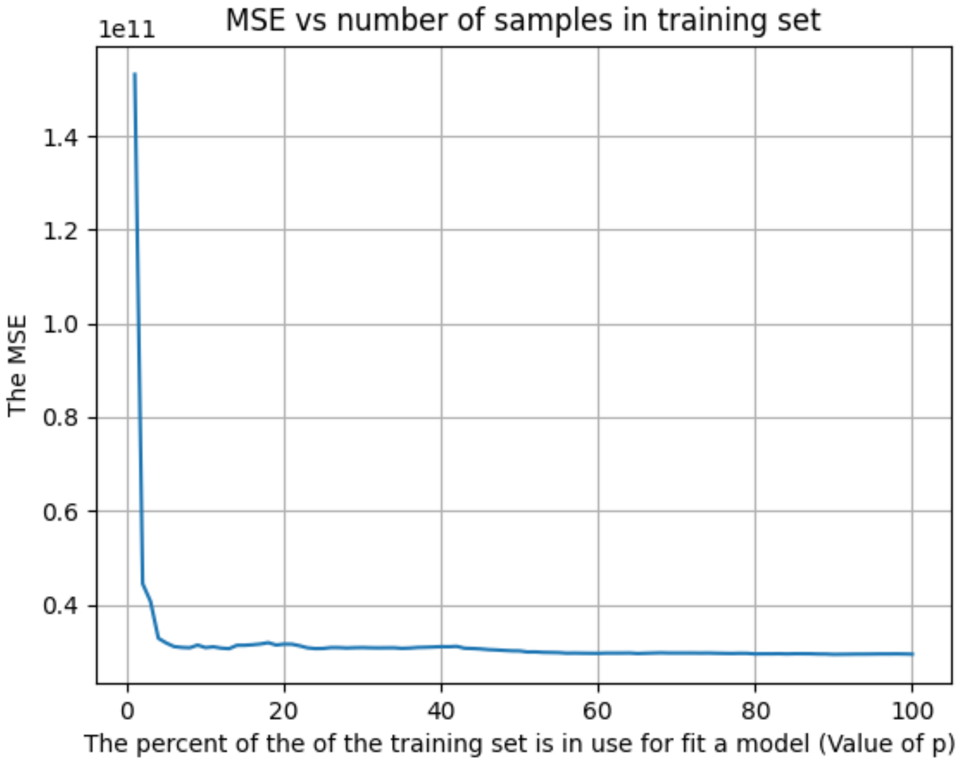
15. ניתן לראות מהגרף שהערכים הסינגולריים שואפים לאפס ככל שהאינדקס גדל אך אף אחד מהם אינו 0 , מכאן ניתן להסיק שהמטריצה XX^T היא מטריצה "כמעט" סינגולרית



תרגיל בית IML 2

דורון ברודר
204917629
doronbruder

.16
ניתן ללמוד מהגרף שהשגיאה הממוצעת הריבועית מתקרבת לאפס ככל שמספר הדגימות (אחוז) גדל
זאת משום שככל שיש יותר דגימות האומדנים משתפרים (חוקי סטטסטטיקה)



תרגיל בית IML 2

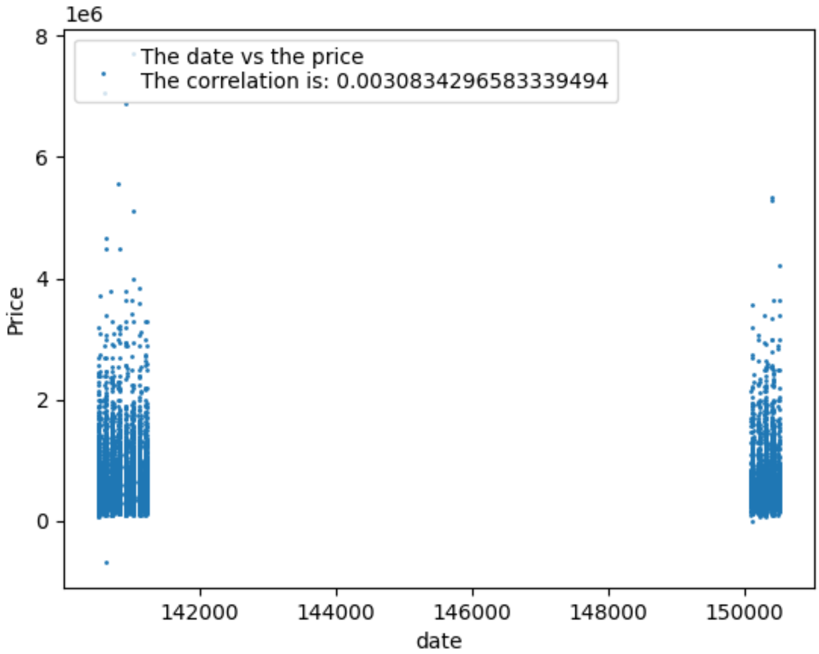
דורון ברודר
204917629
doronbruder

.17
התכונות הלא קטגוריות הן:

```
['date','sqft_living', 'sqft_lot', 'floors','bedrooms','bathrooms','sqft_above','sqft_basement','yr_built','yr_renovated', 'sqft_living15', 'sqft_lot15']
```

שכן התכונות הנ"ל (באופן תאורטי) לא חסומות וכל ערך שלם יכול להיות ערך חוקי עבורן , בניגוד למשל לVIEW שהוא ערך בינארי , או ערכים אחרים שהושמטו.
נציג כעת את הגרפים השונים:

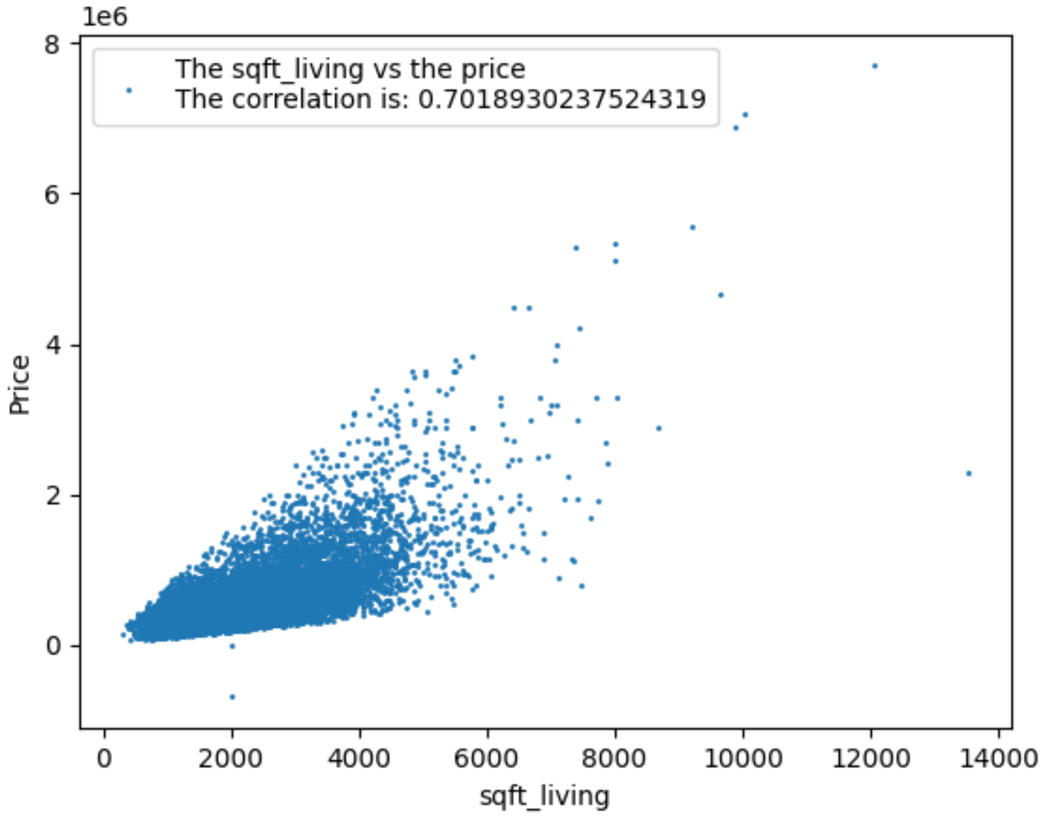
נתחיל בגרף של פיצ'ר שלא משפיע על המודל : התאריך פרסום של המודעה (DATE) ניתן להבין שזו תכונה לא מועילה מכיוון שיש ריכוז זה של מחירים שונים גם בתאריכים ישנים וגם בתאריכים חדשים (בקצוות)



תרגיל בית IML 2

דורון ברודר
204917629
doronbruder

כעת נציג גרף של פיצ׳ר שנראה כמשפיע על המודל : גודל הסלון (sqrft living) ניתן להבין כי זה משפיע על המודל , כלומר על המחיר שכן החל ממחיר מסויים ניתן לראות שבהכרח יש רק חדרי סלון שגודלים מגודל מסויים , למשל מעל 4 מליון הסלון בהכרח מעל 6000



תרגיל בית 2: דורון קריזר 204917629

[נסמן בשאלות הבאות את A ("X") המטריצה של העתקה $T: V \to V$]

: 1.1

(1) ת.ל: $\ker(X) = \ker(X^TX)$

(נראה הכלה דו כיוונית:)

"⇐": יהי $v \in \ker(X)$ אזי $Xv = \underline{0}$ ונקבל:

$$X^TX\underline{v} = X^T(X\underline{v}) = X^T\underline{0} = \underline{0}$$

ולכן $\underline{v} \in \ker(X^TX)$

"⇒": יהי $v \in \ker(X^TX)$ אזי $(X^TX)v = \underline{0}$

נכפיל משמאל ב- v^T ונקבל:

$$v^TX^TXv = \underline{0}$$

$$\Rightarrow (Xv)^T(Xv) = \underline{0}$$

$$\Rightarrow \langle (Xv)^T, (Xv) \rangle = \underline{0}$$

$$\Rightarrow X\underline{v} = 0$$

$$\Rightarrow \underline{v} \in \ker(X)$$

(2) ת.ל עבור מטריצה A (ריבועית) מתקיים: $\mathrm{Im}A^T = (\ker(A))^\perp$

(הערה: מכיון ש-A ריבועית מימד התחום והטווח שווים.)

"⇐": יהי $b \in \mathrm{Im}(A^T)$ אזי קיים $x \in V$ כך ש- $A^Tx = b$

יהי $\underline{v} \in \ker(A)$ אזי $A\underline{v} = \underline{0}$

$$\langle \underline{b}, \underline{v} \rangle = \langle A^T\underline{b}, \underline{v} \rangle = \langle b, A\underline{v} \rangle = \langle b, \underline{0} \rangle = \underline{0}$$

$$\Rightarrow \underline{b} \in (\ker(A))^\perp \Rightarrow \mathrm{Im}(A^T) \subset (\ker(A))^\perp$$

נזכור מימד מרחב השורות שווה למימד מרחב העמודות ולכן:

$$\dim \mathrm{Im}A^T = \dim \mathrm{Im}A \underset{\text{משפט המימד}}{=} \dim V - \dim\ker A$$

כמו כן ידוע: $\dim V = \dim\ker(A) + \dim(\ker A)^\perp \Leftarrow \dim(\ker)^\perp = \dim V - \dim\ker(A)$

מהכלה חד כיוונית ושוויון המימדים נקבל ש- $\mathrm{Im}A^T = \ker(A)^\perp$

Scanned with CamScanner

1.1 המשך:

(3) השתמש עכשיו כי אם A לא הפיכה אז למערכת יש אינסוף פתרונות או אין פתרון.

כעת: מסעיף קודם נקבל ש:

$$y \perp \ker(X) \iff y \in (\ker X)^{\perp} \iff y \in Im(X^T)$$

כלומר y מהווה פתרון של המשוואה הנורמלית עבור המטריצה X.

(אני משוכנע שקיים לפחות פתרון של מערכת המשוואות הנורמליות ... שזה שקול להגיד שקיים פתרון.)

(4) אם X^TX הפיכה (הפיכה) נכפול ב- $(X^TX)^{-1}$

ונקבל את הפתרון היחיד.

אחרת לפי סעיף קודם מספיק להראות כי $Xy \perp \ker(X^TX)$

$\iff$ וזה אמיתי $Xy \perp \ker(X)$

נפלג למקרים:

(א) $y \notin \ker X$ אזי $Xy \in Im X$ (שונה מוקטור האפס ...)

כלומר: $Xy \in (\ker(X^*))^{\perp}$

$$\Rightarrow Xy \perp \ker(X^*)$$

(ב) אחרת אם כל פתרון $w \in \ker(X)$

ובנוסף ידוע ש- $\ker(X^T) \neq \phi$

$\Leftarrow$ X^T לא הפיכה $\Leftarrow$ הגרעין של X הוא לא ריק.

ולכן לפתרון w קיים.

(תרגיל 2 אימון היסק)

ב.1:

5 יהי $V \subseteq R^d$, $\dim(V)=k$

יהי $(v_1 \dots v_k)$ בסיס אורתונורמלי של V

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

(א) ל.3: P סימטרית:

$$P^t = \left(\sum v_i v_i^t\right)^t = \sum (v_i v_i^t)^t = \sum v_i v_i^t = P$$

t לינארית

מכפלה חיצונית של וקטור עם עצמו נותן מטריצה סימטרית

Scanned with CamScanner

1.2 המשך: (סימון $\delta_{i,j} = \begin{cases} 1 & i=j \\ 0 & i \neq j \end{cases}$)

5 (א):

נשים לב כי

$$P^2 = \sum_{i,j} v_i v_i^T v_j v_j^T = \sum_{i,j} v_i \langle v_i, v_j \rangle v_j^T$$

$$= \sum_{i,j} v_i \delta_{ij} v_j^T = \sum_i v_i v_i^T = P$$

אבחנה: אם $P = P^2$ אז כל ערך עצמי של P הוא 0 או 1

הוכחה:

יהי $\underline{v}$ וקטור עצמי של P כלשהו:

$$Pv = \lambda v$$

אבל $Pv = P^2 v = \lambda^2 v$

ולכן $\lambda^2 v = \lambda v$

אבל $\underline{v} \neq 0$ (וקטור עצמי)

ולכן $\lambda = \lambda^2$

$$\lambda \in \mathbb{R}$$

$$\Rightarrow \lambda \in \{0, 1\}$$

כעת נראה שעבור $\forall j$ v_j (וקטור עצמי)

$$Pv_j = v_j = 1 \cdot v_j$$

([illegible])

$$Pv_j = \left(\sum_i v_i v_i^T\right) \cdot v_j = \sum_i v_i \langle v_i^T, v_j \rangle = \sum_i v_i \cdot \delta_{ij}$$

$$= v_j$$

1.2 המשך:

5 (ב):

ג.3: $Pv=v \quad \forall v\in V$

הוכחה: אם $v\in V$ ניתן לרשום כצירוף של $v_1,\dots,v_k$ (בסיס /

(בסיס = ו"ע)

$$\underline{v} = \sum_{i=1}^{k}\lambda_i v_i$$

$$\Rightarrow P\underline{v} = P\sum_{i=1}^{k}\lambda_i v_i$$

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

$$= \underline{v}$$

(הבה שימוש רק בהגדרה של פ v_j לבסיס הנתון הוא 1

והאסוציאטיביות של כפל מטריצות)

(ב) (הוכחת באופן ד'.)

$(P^2=P)$

(ג) ג.3 $(I-P)P=0$

נובע מסעיף (ב)

~~שאלה~~

1.3:

6. בכיתה ראינו את הפתרון: $w = (X^TX)^{-1}X^Ty$

בתרגול ראינו את הפתרון:

$$w = V\Sigma^+U^Ty \overset{def}{=} X^+y$$

$$s.t.\ \Sigma^+_{i,i} = \begin{cases} \frac{1}{\Sigma_{i,i}} & \Sigma_{i,i} \neq 0 \\ 0 & else \end{cases}$$

נרצה להראות ש: אם X הפיכה אזי: $X^+ = (X^TX)^{-1}X^T$

$$XX^T = ((U\Sigma V^T)(U\Sigma V^T)^T)^{-1} = (U\Sigma V^TV\Sigma^TU^T)^{-1}$$

$$= (U\Sigma\Sigma^TU^T)^{-1} = (U^T)^{-1}(\Sigma\Sigma^T)^{-1}U^{-1} = U(\Sigma\Sigma^T)^{-1}U^T$$

השתמשנו בכך ש $VV^T = I$, ובנוסף X הפיכה אז ה"ע $\neq 0$ ולכן $\Sigma^{-1} = \Sigma^+$

אלכסונית היא (המטריצה האלכסונית הן ערכים סינגולרים חיוביים ע"ע)

$$\Rightarrow (XX^T)^{-1}X = U(\Sigma\Sigma^T)^{-1}U^TU\Sigma V^T = U\Sigma^{-1}\Sigma^{-1}\Sigma V^T = U\Sigma^{-1}V^T$$

$$\underset{def}{=} X^+$$

$\cancel{\dim Im A^T = \dim Im A}$

תרגול 1.3:

7. צ"ל: X^TX הפיכה אמ"מ $R^d = Sp(x_1, .. x_m)$

לפי סעיף א' ← 1.1: $\ker(X) = \ker(X^TX)$

נקבל: X^TX הפיכה $\Leftrightarrow \ker(X^TX) = 0 \Leftrightarrow \ker(X) = 0$

$\Leftrightarrow$ X הפיכה $\Leftrightarrow$ עמודות של X בת"ל

$\Leftrightarrow$ d וקטורים בת"ל הם בסיס ל R^d

ג.1.

8. נשים לב שעבור XX^T לא הפיכה אזי מס' הערכים הסינגולריים ב-$\Sigma\Sigma^T$ קטן מ-d, נסמנו ב-r.

לפי פירוק ה-SVD נקבל:

$$(*)\ XX^Tw = Xy$$

$$\Longleftrightarrow U\Sigma\Sigma^TU^Tw = UWV^Ty$$

מכיון שמס' הערכים הסינגולריים ב-$\Sigma\Sigma^T < d$ זה אומר שיש $d-r$ שורות אפסים ונקבל שלכל פתרון של $(*)$ שנסמנו ב-w מתקיים ש- $w_{d-r}, \dots, w_d$ הם משתנים עם דרגת חופש, (ואילו $w_r \dots w_1$ נקבעות באופן יחיד..)

לכן כל פתרון ל-$(*)$ יהיה מהצורה:

$$\begin{bmatrix} w_1 \\ \vdots \\ w_r \\ \alpha_{r+1} \\ \vdots \\ \alpha_d \end{bmatrix}$$

כאשר $\{w_i\}_{i=1}^{r}$ נקבעים וזהים עבור כל הפתרונות ולעומת זאת $\{\alpha\}_{i=r+1}^{d}$ יכולים להיות כלשהם.

כעת נשים לב שבפתרון $\hat{w}$ שמצאנו בתרגיל (כפי שחושב קודם)

$$\hat{w} = X^+y = V\Sigma^+U^+y$$

מתקיים כי $w_{\alpha_{r+1}} \dots w_{\alpha_d} = 0$ (מכיון ש-Σ^+ מכילה $d-r$ שורות אפסים מתאימות מהגדרתה)

ולכן ברור כי:

$$\left\| \begin{bmatrix} w_1 \\ \vdots \\ w_r \\ 0 \\ \vdots \\ 0 \end{bmatrix} \right\|_2 = \|\hat{w}\|_2 \leq \|\tilde{w}\|_2 = \begin{bmatrix} w_1 \\ \vdots \\ w_r \\ \alpha_{r+1} \\ \vdots \\ \alpha_d \end{bmatrix}$$

(עבור כל פתרון שלא חושב בדרך הנ"ל.)

(כיון שהמחוברים α_i לא בהכרח יהיו 0, ואילו האחרים ישארו כפי שהם)