318442241 ת.ז. יאיר שטרן ex2 – IML

.1

הפיצ'רים שחשבתי שאין להם השפעה על מחיר הדירה והורדתי:

id, lat, long, date

כמו כן, סיננתי ערכים לא הגיוניים (שקטנים או שווים לאפס) עבור:

price, sqrt_lot, sqrt_living, bedrooms, floors

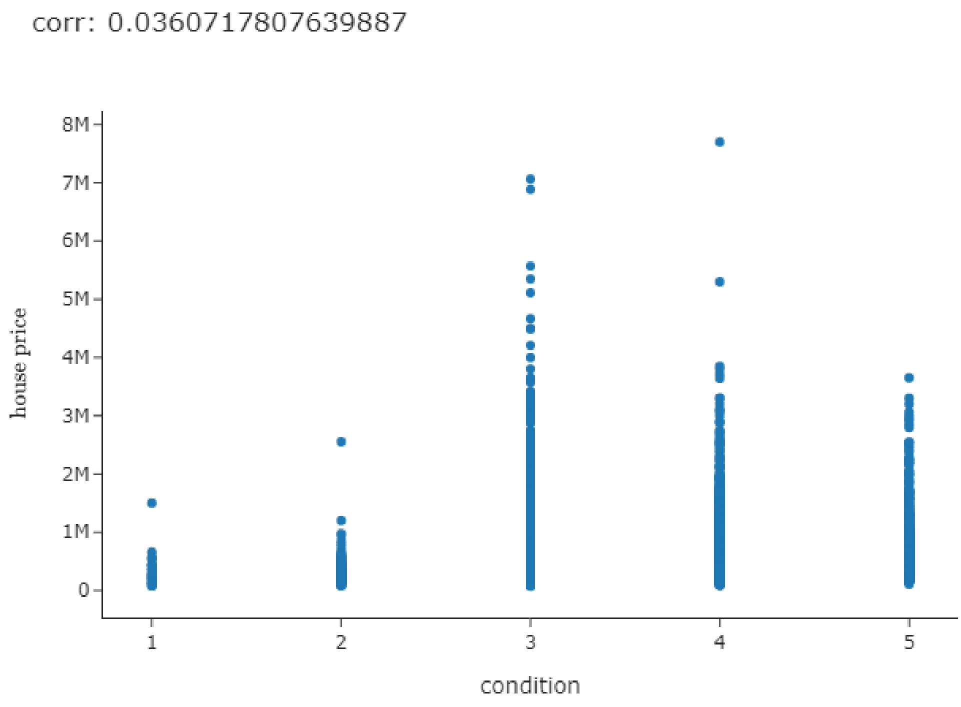
כמו כן הסרתי ערכי Nan, וחזרות.

הוספתי עמודה חדשה "bedrooms_for_floor" המכילה את מספר החדרים בכל קומה, שחשבתי שהוא יכול להוסיף מידע על שווי הבית (שהרי בית עם הרבה קומות עם מעט חדרים בכל קומה לא בהכרח שווה יותר מבית עם פחות קומות והרבה חדרים בכל קומה).

הפיצ'ר "zipcode" הוא ערך קטגוריאלי ולכן יצרתי dummies שמחלק את הערכים לעמודות המכילות 1/0 אם הבית נמצא בזיפ קוד הספציפי.

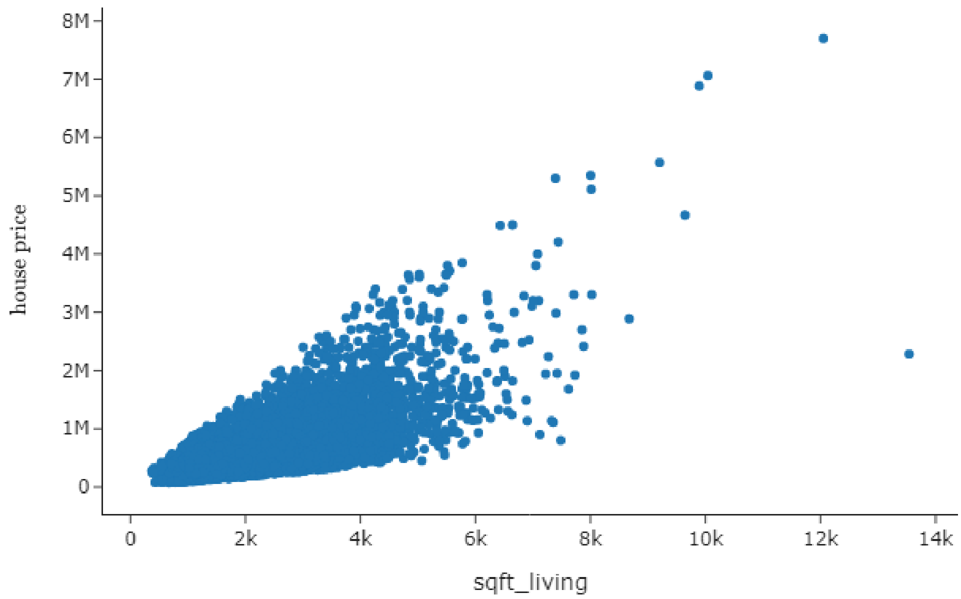
.2

הקורולציה של הפיצ'ר conditions נמוכה וניתן לראות בגרף שאין קשר בין מצב הדירה למחיר.



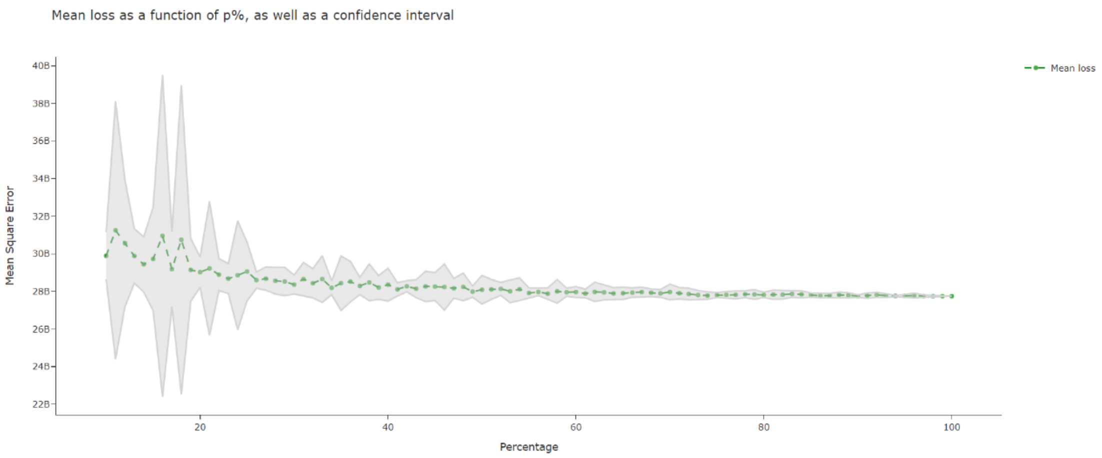
לעומת זאת בפיצ'ר sqrt_living שהקורולציה שלו גבוהה ניתן לראות בגרף שיש קשר בין הפיצ'ר למחיר הבית.

corr: 0.7019749231432295



.4

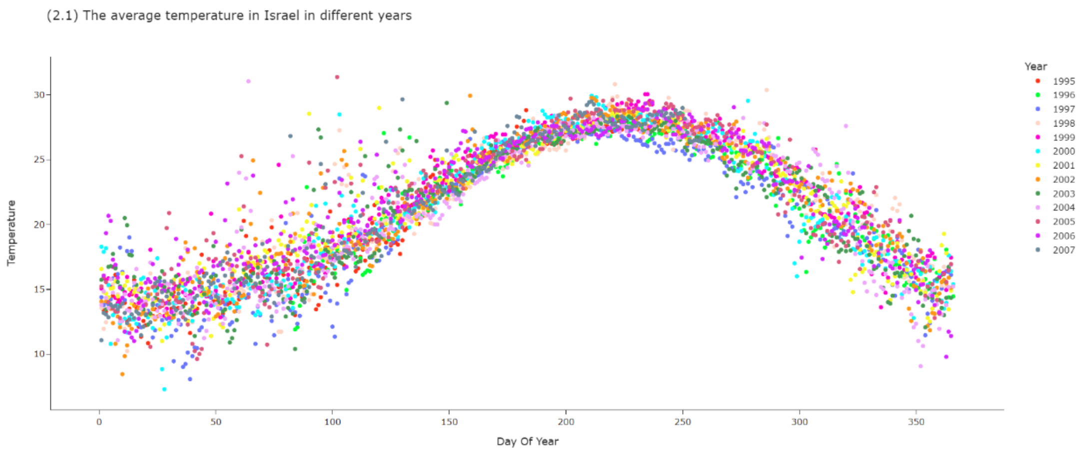
הקו הירוק מתאר את ה loss הממוצע, השטח האפור מתאר את ה confidence interval. ניתן לראות שככל שאנחנו לוקחים יותר אחוזים מהדגימות ה loss הממוצע מתכנס לערך כל שהוא כאשר גם ה confidence interval מתכנס אל אותו הערך.



חלק 2:

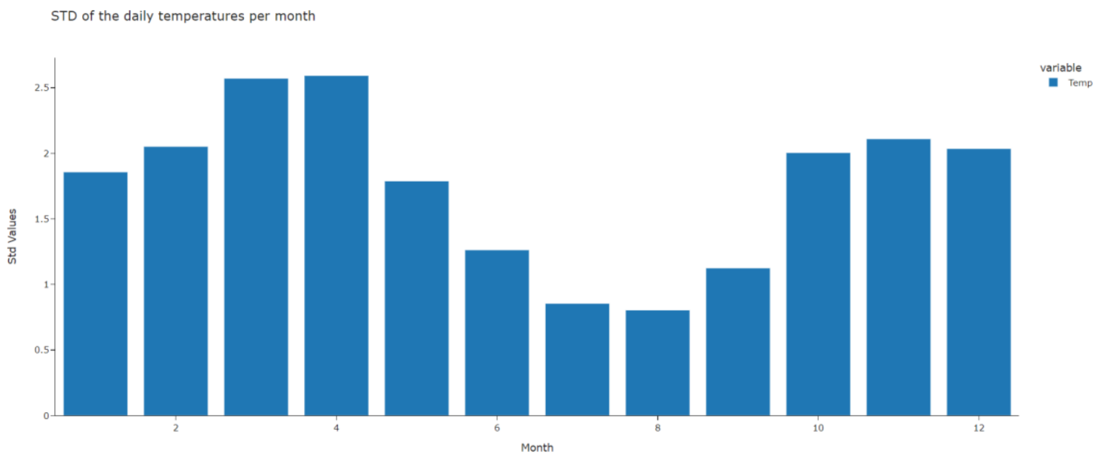
1.2

כפי שנראה הגרף מתאים לפולינום מדרגה הגבוהה מ-2



2.2

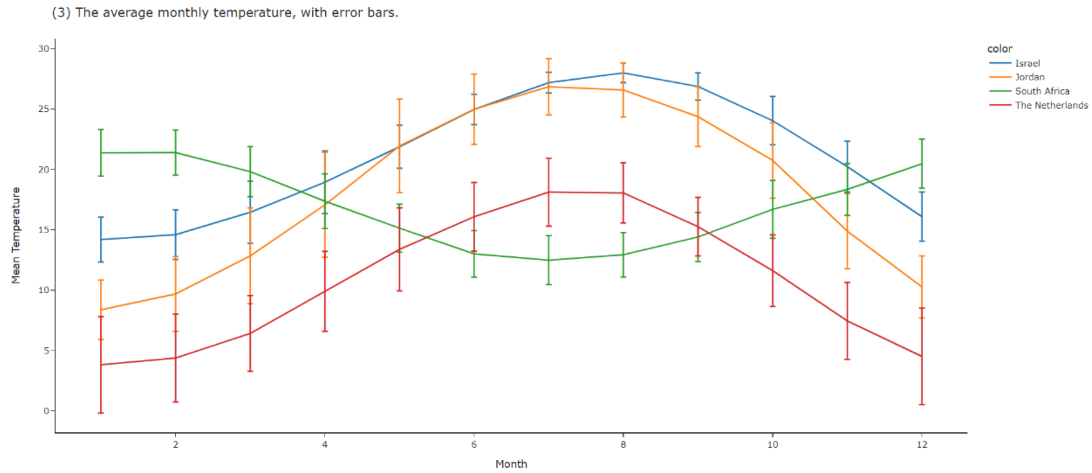
ניתן לראות שסטיית התקן בחודשים 7,8 היא נמוכה יותר (ניתן לשער שזה כיוון שבחודשים האלו התחזית יחסית קבועה בניגוד לשאר החודשים ולתקופות מעבר.)
לכן ניתן לשער שהמודל יחזה טוב יותר בחודשים אלו.



.3

ראשית, ניתן לראות מהגרף שהדפוס של ירדן דומה לשל ישראל, כמו כן הדפוס של הולנד דומה לשל ישראל אך נמוך יותר (קר שם יותר...). לעומת זאת הדפוס של דרום אפריקה ממש שונה משניהם כאשר בחודשים הקרים בישראל חם באפריקה ולהפך.

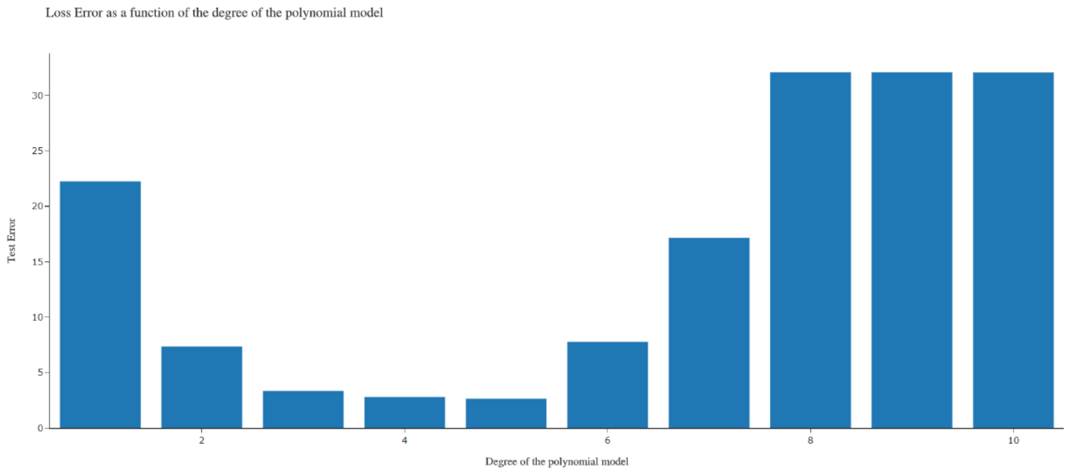
ניתן להסיק שהמודל של ישראל יכול לעבוד טוב (יחסית) עבור ירדן אך לא עבור שאר המדינות.



.4

```
for polynomial model of degree  1  the loss is :  22.26
for polynomial model of degree  2  the loss is :  7.38
for polynomial model of degree  3  the loss is :  3.37
for polynomial model of degree  4  the loss is :  2.82
for polynomial model of degree  5  the loss is :  2.67
for polynomial model of degree  6  the loss is :  7.8
for polynomial model of degree  7  the loss is :  17.18
for polynomial model of degree  8  the loss is :  32.11
for polynomial model of degree  9  the loss is :  32.11
for polynomial model of degree  10  the loss is :  32.1
```

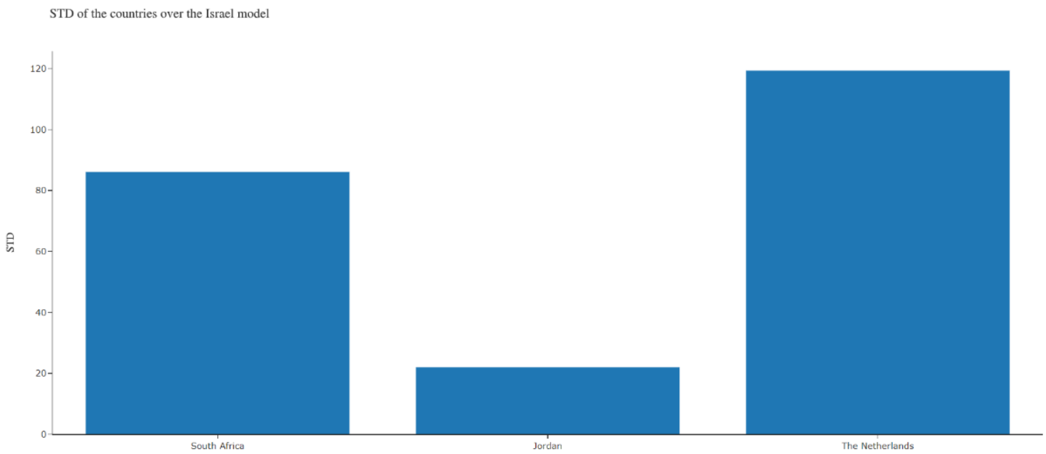
ניתן לראות שהמעלה המביאה את ה loss הנמוך ביותר היא 5 לכן מודל ממעלה 5 הכי טוב. (3,4 נותנים גם כן loss נמוך, לכן הם גם אפשריים).



.5

כמו שראינו בגרף בשאלה 3, ניתן לראות שחיזוי לפי ישראל עבור ירדן מביא לשגיאה נמוכה ביחס לשאר המדינות.

נשים לב, שהשגיאה עבור הולנד גדולה יותר מהשגיאה עבור דרום אפריקה (למרות שראינו שהדפוס דומה) כי אמנם הדפוס דומה לשל ישראל אך ברוב השנה הערכים של דרום אפריקה קרובים יותר לשל ישראל מאשר ערכי הולנד.



318442241 יאיר שטכן

IML חלק תיאורטי:

שאלה 1:

כיוון ראשון: יהי $v\in\ker(X) \Rightarrow Xv=0 \Rightarrow X^TXv=0$

$\Rightarrow v\in\ker(X^TX)$

כיוון שני: יהי $v\in\ker(X^TX) \Rightarrow X^TXv=0 \Rightarrow v^TX^TXv=0$

$\Rightarrow (v^TX^T)Xv=0 = (Xv)^TX=0 \Rightarrow \langle Xv,Xv\rangle=0$

$\Rightarrow Xv=0 \Rightarrow v\in\ker(X)$.

שאלה 2:

$v\in\ker(A) \iff Av=0 \iff \langle Av,w\rangle=0$ לכל $w\in V$

$\iff \langle v,A^Tw\rangle=0 \quad w\in V \iff v\perp A^Tw$

$\iff v\in \mathrm{Im}(A^T)^{\perp}$

ולפי המשלימים האורת׳ נקבל את הנדרש.

שאלה 3:

למערכת קיים פתרון יחיד אמ"מ היא הפיכה

אחרת יש אינסוף פתרונות או שלא קיים פתרון

לכן, עבור מערכת $Xw=y$ יש אינסוף פתרונות אמ"מ קיים

$u\in\mathbb{R}^n$ כך ש $Xu=y \iff y\in Im(X)$

משאלה קודמת, $y\in Im(X^T)^\perp$ אמ"מ $y\in Im(X)$

לכן $y\in Im(X)$ אמ"מ $y\perp ker(X^T)$

שאלה 4:

נניח ש- X^TX הפיכה ← קיימת $(X^TX)^{-1}$ ההופכית. נקבל:

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

צ"ל קיים w כך שלמערכת יש פתרון יחיד.

כעת נניח ש X^TX לא הפיכה.

משאלה קודמת למערכת $X^TXw=X^Ty$ יש אינסוף פתרונות

אמ"מ $X^Ty\perp ker(X^TX) \underset{\text{שאלה 1}}{\iff} X^Ty\perp ker(X)$

לכן, עבור $w\in ker(X)$ נקבל: $\langle X^Ty|w\rangle = y^TX\underbrace{w}_{\in ker(X)}=0$

← קיבלנו ש $X^Ty\perp ker(X)$.

שאלה 5:

א) לכל $i\in[n]$ $v\in\mathbb{R}^d$ בסיס אורתוגונלי.

$$P=\sum_{i=1}^{n} v_i \otimes v_i^T = \begin{pmatrix} v_1^1 v_1^1 & \cdots & v_1^1 v_d^1 \\ \vdots & & \vdots \\ v_d^1 v_1^1 & \cdots & v_d^1 v_d^1 \end{pmatrix} + \cdots +$$

$$\cdots + \begin{pmatrix} v_1^n v_1^n & \cdots & v_1^n v_d^n \\ \vdots & & \vdots \\ v_d^n v_1^n & \cdots & v_d^n v_d^n \end{pmatrix} = \begin{pmatrix} \sum_{i=1}^n v_1^i v_1^i & \cdots & \sum_{i=1}^n v_1^i v_d^i \\ \vdots & & \vdots \\ \sum_{i=1}^n v_d^i v_1^i & \cdots & \sum_{i=1}^n v_d^i v_d^i \end{pmatrix}$$

כל אחד מהמחוברים זו מטריצה סימטרית ולכן גם החיבור שלהם.

ב)

עבור $i\in[k]$ נראה ש- $Pv_i=v_i$

$$Pv_i=\sum_{j=1}^{n}(v_j v_j^T)v_i=\sum_{j=1}^{n} v_j \langle v_j | v_i\rangle = v_i\langle v_i|v_i\rangle$$

$$= v_i \|v_i\|^2 = v_i$$

$\|v_i\|=1$

שונה מאפס רק כאשר $i=j$ כי הם אורתונורמלים.

עבור $i\in\{k+1,\dots,d\}$ נקבל מכך שהם אורתו' ש: $Pv_i=0$

סה"כ נקבל שקיים בסיס $v_1,\dots,v_d$ בסיס מו"ע עם ע"ע $0,1$.

ג)

יהי $v\in V$, v הוא צ"ל של הבסיס האורתונורמלי $v_1,\dots,v_k$

כך ש $v=\sum_{i=1}^{k}\alpha_i v_i$

מסעיף קודם הם וע"ע ונקבל $Pv=P(\sum_{i=1}^{k}\alpha_i v_i)=\sum_{i=1}^{k}\alpha_i \underbrace{Pv_i}_{v_i}=v$

ג)

נשתמש ב- EVD, לפי קיימת מטריצה U אורתוגונלית

כשהעמודות שלה וקטורים עצמיים של P, D אלכסונית כאשר על האלכסון הע"ע של P

מתקיים $P = UDU^T$

$$P^2 = (UDU^T)^2 = UD\underbrace{U^TU}_{I}DU^T = UDU^T = P$$

ד)

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

סעיף קודם $P = P^2$

שאלה 6:

נרצה להראות ש- $(X^TX)^{-1}X^T = X^\perp$

נשתמש ב- SVD על X ונקבל:

$$X^TX = (U\Sigma V^T)^T(U\Sigma V^T) = V\Sigma^T\underbrace{U^TU}_{I}\Sigma V^T = V\Sigma^T\Sigma V^T$$

כעת:

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

וסה"כ:

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

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

שאלה 7:

X^TX הפיכה אמ"מ $\ker(X^TX)=0$ אמ"מ (שאלה 1) $\ker(X)=0$

וזה אמ"מ $\dim(Im(X))=d$ אמ"מ $\dim(col(X))=d$

וזה אמ"מ $\dim(row(X))=d \iff Span\{x_1,\dots,x_m\} \ni \mathbb{R}^d$

שאלה 8:

נסתכל על פתרון w של המשוואה $\underset{w\in\mathbb{R}^d}{\arg\min}\|Xw-y\|^2$

ויהי $X=U\Sigma V^T$ פירוק SVD של X.

$$\|Xw-y\|^2=\|U\Sigma V^Tw-y\|^2=\left\|U\Sigma\begin{bmatrix}V_1^Tw\\V_2w\end{bmatrix}-UU^Ty\right\|^2=$$

$$=\left\|U\left(\Sigma\begin{bmatrix}V_1^Tw\\V_2w\end{bmatrix}-U^Ty\right)\right\|^2=\left\|\begin{pmatrix}\Sigma&0\\0&0\end{pmatrix}\begin{bmatrix}V_1^Tw\\V_2w\end{bmatrix}-\underbrace{[V_1V_2]^T}_{U}y\right\|^2$$

Σ החלק שאינם אפסים

$$=\left\|\begin{bmatrix}\Sigma_1V_1^Tw\\0\end{bmatrix}-\begin{bmatrix}V_1^Ty\\0\end{bmatrix}-\begin{bmatrix}0\\V_2^Ty\end{bmatrix}\right\|^2=\left\|\begin{bmatrix}\Sigma_1V_1^Tw-V_1^Ty\\0\end{bmatrix}\right\|^2+\left\|\begin{pmatrix}0\\V_2^Ty\end{pmatrix}\right\|^2$$

כעת זה שווה אפס אמ"מ $\Sigma_1V_1^Tw-V_1^Ty=0$

$$\iff \quad V_1^Tw=(\Sigma_1)^{-1}V_1^Ty$$

ועבור $w=X^+y$ נקבל $V_1^TX^+y=(\Sigma_1)^{-1}V_1^Ty$

$$V_2^TX^+y=V_2^T(V\Sigma^+V^T)y=0$$

לכן נקבל שלפי פתרון $\hat{w}$ מתקיים ש- "$\tilde{w}\|\varepsilon\|w$"