IML- EX2

Amitay Sicherman

תוכן

[Solutions of the Normal Equations: 1](#_Toc38305660)

[Q1: 1](#_Toc38305661)

[Q2: 2](#_Toc38305662)

[Q3: 3](#_Toc38305663)

[Q4: 3](#_Toc38305664)

[Projection Matrices 3](#_Toc38305665)

[Q5: 3](#_Toc38305666)

[Least Squares 6](#_Toc38305667)

[Q6 6](#_Toc38305668)

[Q7 6](#_Toc38305669)

[Q8 7](#_Toc38305670)

[Practical Questions 8](#_Toc38305671)

[Q13 8](#_Toc38305672)

[Q15 9](#_Toc38305673)

[Q16 10](#_Toc38305674)

[Q17 11](#_Toc38305675)

[Q21: 13](#_Toc38305676)

[Q22: 14](#_Toc38305677)

# Solutions of the Normal Equations:

## Q1:

We use the Kerr definition :

and the fact that :

And the proof:

First side:   
let

so we know that from 1) ,

multiple both sides by in the left side and get

From the second side:   
Let

so from 1)

multiple both from the left in and get use the transpose   
from 2) we get that   
and because it is Norm we get   
and from 1)

## Q2:



Basic definitions:

Note   
  
Fist side -   
Let we get that

So by the definition.  
  
Second side-  
there is   
so   
but we see that   
so

Q3:  
X is square and not invertible so there is 0 or solutions.  
so, if there is   
in we see that

## Q4:

if so multiplate both sides in

if so from Q2 there is solutions iff in Q1 we see that , so we need to show that ,  
lets look on , so   
so we get that from Q3 we get that there is

# Projection Matrices

## Q5:

א.  
על פי ההגדרה של הטלה אורתוגונלית נקל כי עבור כל שני ווקטורים מתקיים ש כלומר ש   
 כעת נפתח את הביטוי אלגברית :   
מכיוון שביטוי נכון לכל נקבל כי

נשתמש בזהות הזו ונקבל:

ב.  
הגדרת ערך עצמי ווקטור עצמי היא ש :  
נוכיח כי הם וקטורים עצמיים בעלי ערך עצמי 1:

מכיוון שווקטורי הבסיס אורתוגונליים זה לזה נקבל שהמכפלה הפנימית מתאפסת כאשר ושווה ל1 כאשר ובפרט נקבל כי :

כלומר וקטור עצמי בעל ערך עצמי אחד.  
עבור הווקטורים ששייכים ל נקבל כי אלו ערכים עצמיים עם ערך עצמי 0 לפי הגדרת ה  
נראה כי לא קיימים וקטורים עצמיים נוספים בעלי ערך עצמי שונה מ0 :  
נניח בשלילה קיים וקטור כזה המקיים ש , אז מתקיים :  
מכיוון ש ניתן לכתוב אותו כסכום של וקטור מתוך V ווקטור ניצב לV: נסמן

ונקבל:

עבור חלק 1:   
*נפתח איבר אחד במחוברים :*

ולכן עבור הביטוי הכולל נקבל:   
עבור חלק 2:

ולכן נקבל כי

כלומר בכל מקרה שבו נקבל כי איננו ווקטור עצמי.  
כאשר נקבל כי ולכן הערכים העצמיים המתאימים הם ערכי הבסיס האורתוגונלי.

ג.

*נניח כי ונראה כי :  
מכיוון ש יש לו הצגה יחידה על ידי וקטורי הבסיס: ולכן נקבל :   
נפתח איבר אחד במחוברים :*

ולכן עבור הביטוי הכולל נקבל:   
*ד.  
 ראשית באופן אינטואיטיבי: הטענה נכונה מכיוון שהכפלה במטריצה P מטילה כל ווקטור לתוך תת המרחב V, ולכן אין משמעות להכפלה מטריצה פעם אחת או פעמיים- מכיוון שלאחר שהטלנו את הווקטור לתוך V הכפלה במטריצה P לא תשנה אותו.  
ובאופן פורמלי:*בהתבסס על סעיף ג , נקבל   
שאם נקבל כי   
אם לפי ההגדרה של P נקבל כי ולכן   
ובכל מקרה קיבלנו כי

ה.

על פי סעיף ד:

# Least Squares

## Q6

א.  
נשתמש בהגדרות :  
*כאשר השתמשתי בכך שעבור מטריצה אורתוגונלית X מתקיים*

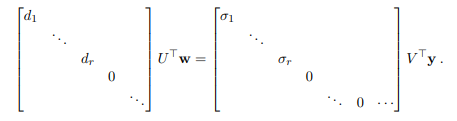
*ב.*

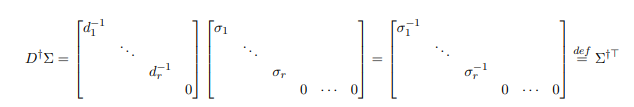
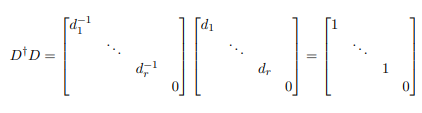
## Q7

*מכיוון שהדרגה של X היא מספר הערכים הסינגולריים שלו, נשתמש בSVD ונראה שהדרגות של X ושל שוות :*

*על פי ההגדרה של D מספר הערכים שאינם מתאפסים בD ו הוא זהה, ולכן הדרגה שלהם שווה.  
כעת, נשים לב כי היא מטריצה ריבועית בעלת מימדים , ולכן אם"ם הגדרה של היא d נקבל שהיא הפיכה. ולכן גם הגדרה של X היא d כלומר*

Q8*:*

על מנת להבין לעומק את הפתרון, אשתמש בפיתוח מהתרגול (מספר 3) משנה שעברה:  
מדוע הוא פתרון?   
הגענו למשוואה: מכאן נוכל לפתח ע"י SVD :  
אם D לא הפיכה- כלומר יש לה ערכים 0 על האלכסון- אז המשוואות בעצם נראות כך:  
  
  
  
*זוהי מערכת משוואות שיש לה יותר מפתרון אחד- כדי להגיע לפתרון יחיד, נגדיר את לכל .  
במצב זה צמצמנו את מערכת המשוואות וקיבלנו פתרון יחיד.  
הפתרון הזה הוא הפתרון שאנו מתבקשים להראות שהוא מינימלי .  
פיתוח המשוואה:*

*כאשר מתקיים :  
*

*ולכן כאשר ביצענו את האיפוס על איברי W נקבל כי אפשר להשמיט את ולקבל את הפתרון:*

*כעת נראה כי הפתרון התקבל ע"י האיפוס הנ"ל הוא המינימלי מכל הפתרונות:*

*כאשר השתמשתי בעובדה ש וש .*

# Practical Questions

## Q13

*חילקתי את הפיצ'רים השונים ל3 סוגים:*

1. *פיצ'רים שלדעתי צריך למחוק- איך תלות (לפחות לא כזו שתוכל להתבטא באופן לינארי) בינם לבין מחיר הדירה. בחלק זה מחקתי את הפי'צרים : 'id','date','price','zipcode','grade','lat','long'*
2. *פיצ'רים של קריטריונים- שאותם פיצלתי לעמודות בינריות: בחרתי לחלק זה את הפיצ'רים שמקיימים שני תנאים- לדעתי הם משפיעים על המחיר אך התלות הזו לא לינארית בגודל הערך של הפיצ'ר- ולכן מעבר ל מתאים כדי להתמודד איתם.  
   בחרתי לעשות* categorical features ל zipcode – מכיוון שהאזור משפיע על המחיר, אבל לא לפי ה'גודל' שלו. ובדומה ה grade- שפה לדעתי זו נקודה לא חד משמעית אבל החלטתי גם פה שיש תלות בקומה- אבל לאו דווקא ככל שהקומה יותר גבוהה המחיר יעלה (או להפך) ובנוסף בגלל שמספר האפשרויות של הקומות קטן יחסית בחרתי לפצל גם את העמודה הזו.
3. את שאר הפיצ'רים השארתי- כאשר חסמתי כל אחד מהם בחסם עליון ותחתון אשר היה נראה לי הגיוני.

## Q15

*הגרפים:*

*תמונה שמכילה צילום מסך

התיאור נוצר באופן אוטומטי*

*ניתן לראות כי ישנם ערכים סינגולריים גדולים מאוד יותר מ6 מליון. ולכן המטריצה רחוקה מלהיות סינגולרית.  
בנוסף , ניתן לשים לב כי הערכים הסינגולריים הולכים ונהיים קטנים יותר בקצב מהיר- וכמו שאפשר לראות בגרף הבא כבר מהערך ה20 הערכים הסינגולריים כבר בגודל של .   
אני לא יכול לומר בוודאות, אך לדעתי זה יכול להגיד שישנים גורמים מסויימים במטריצה אשר מרחיקים אותה מאוד מלהיות סינגולרגית. בעוד הגורמים האחרים לא משפיעים ככ.*

*תמונה שמכילה מפה

התיאור נוצר באופן אוטומטי*

## Q16

תמונה שמכילה צילום מסך

התיאור נוצר באופן אוטומטי

כמו שניתן היה לצפות – מכיוון שהמטריצה רחוקה מלהיות סינגולרית גם ערכי ה MSE גדולים מאוד (וניתן לראות שהם בסדר גודל של הריבוע של הערכים הסינגולריים הגדולים)  
כמו כן ניתן לראות כי ככל שגודל סט האימון גדול יותר כך התוצאה מדוייקת יותר. ובנוסף התוצאה גם יציבה יותר ככל שעולה גודל סט האימון- כלומר המודל יציב יותר.

## Q17

*את הקורלציה הגבוהה ביותר ניתן לראות בין המחיר לבין ה כמו שמתואר בגרף הבא:*

*תמונה שמכילה טקסט, מפה

התיאור נוצר באופן אוטומטי*

*ערך הקורלציה גבוהה (כאשר לקחתי מתוך המטריצה של הקורלציה את הערך שמעיד על הקורלציה בין המשתנים – הערך שלא נמצא על האלכסון הראשי- מכיוון שהערכים על האלכסון מבטאים את השונות של המשתנים עצמם)  
ניתן גם לראות באופן ויזואלי- כי יש התאמה דיי טובה לכך שהמחיר עולה ככל שה עולה.*

*הקורלציה הנמוכה ביותר התקבלה בשנת הבניה של הבית:   
תמונה שמכילה צילום מסך

התיאור נוצר באופן אוטומטי*

*ערך הקורלציה נמוך מאוד, כמו שאפשר לראות בגרך שכמעט אין תלות בין שנת הבניה של הבית לבין המחיר שלו. המחירים דיי אחידים על כל שנות הבניה.*

## Q21:

הגרפים:

תמונה שמכילה צילום מסך

התיאור נוצר באופן אוטומטיתמונה שמכילה מפה, צילום מסך

התיאור נוצר באופן אוטומטי

## Q22:

התהליך שעשינו בשאלה הזו הוא בדיוק הצורה המבוקשת- במקום למצוא ERM חדש- הסתכלנו על הDATA האקספוננציאלית- וביעצנו לה פעולה של לוג.  
לאחר שאנחנו מסתכלים על לוג של הדאטה – אנחנו יכולים לעשות התאמה לינארית רגילה. ולהשתמש בנוסחאות המוכרות לנו.  
אם נרצה להחזיר לאחר מכן את הDATA לצורה האקספוננציאלית – נעלה באקספוננט את הערכים המתקבלים מההתאמה .   
בדיוק כמו שעשיתי בשאלות 19-21