

HW2

ששון שמואל למעי 325172351 ויצחק גרינבוים 318837317

12 במאי 2022

Question .1

א

נשים לב שווקטור השארית צריך להיות מאונך לטרנספוז של A :

$$A^T Ax = A^T b \Rightarrow A^T (Ax - b) = 0 \Rightarrow A^T r = 0$$

נבדוק אילו מאונכים :

נקבל שווקטור (1) מאונך :

$$(1 \quad 1 \quad 1 \quad 1) \begin{pmatrix} 2 \\ -1 \\ -2 \\ 1 \end{pmatrix} = 2 - 1 - 2 + 1 = 0$$

$$(1 \quad 2 \quad 5 \quad 10) \begin{pmatrix} 2 \\ -1 \\ -2 \\ 1 \end{pmatrix} = 2 - 2 - 10 + 10 = 0$$

נקבל גם שווקטור (3) מאונך :

$$(1 \quad 1 \quad 1 \quad 1) \begin{pmatrix} -5 \\ 5 \\ 1 \\ -1 \end{pmatrix} = -5 + 5 + 1 - 1 = 0$$

$$(1 \quad 2 \quad 5 \quad 10) \begin{pmatrix} -5 \\ 5 \\ 1 \\ -1 \end{pmatrix} = -5 + 10 + 5 - 10 = 0$$

עם זאת ווטור 2 אינו מאונך:

$$(1 \quad 1 \quad 1 \quad 1) \begin{pmatrix} 0 \\ 1 \\ 4 \\ 9 \end{pmatrix} > 0$$

ב

נדרוש:

$$\begin{aligned} A^T \begin{pmatrix} c \\ d \\ e \\ f \end{pmatrix} = 0 &\Rightarrow \begin{cases} c + d + e + f = 0 \\ c + 2d + 5e + 10f = 0 \end{cases} \\ &\Rightarrow \begin{cases} c = -d - e - f \\ c = -2d - 5e - 10f \end{cases} \\ &\Rightarrow \boxed{\begin{cases} d = -4e = 9f \\ c = 3e + 8f \end{cases}} \end{aligned}$$

ג

עמודות A בת"ל ולכן נוכל להשתמש בנוסחה:

$$x^* = (A^T A)^{-1} A^T b$$

$$x^* = \begin{pmatrix} 4 & 18 \\ 18 & 130 \end{pmatrix}^{-1} \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 2 & 5 & 10 \end{pmatrix} \begin{pmatrix} 3 \\ 8 \\ 15 \\ 22 \end{pmatrix} = \boxed{\begin{pmatrix} 3 \\ 2 \end{pmatrix}}$$

ד

$$r^* = Ax^* - b = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 2 & 5 & 10 \end{pmatrix}^T \begin{pmatrix} 3 \\ 2 \end{pmatrix} - \begin{pmatrix} 3 \\ 8 \\ 15 \\ 22 \end{pmatrix} = \begin{pmatrix} 5 \\ 7 \\ 13 \\ 23 \end{pmatrix} - \begin{pmatrix} 3 \\ 8 \\ 15 \\ 22 \end{pmatrix} = \begin{pmatrix} 2 \\ -1 \\ -2 \\ 1 \end{pmatrix}$$

$$2 = 3 \cdot -2 + 8 \cdot 1 = -6 + 8 = 2 \Rightarrow \checkmark$$

$$\begin{aligned} -1 &= -4 \cdot -2 - 9 \cdot 1 \\ &= 8 - 9 = -1 \Rightarrow \checkmark \end{aligned}$$

Question 2

✕

$$f(x) = (x_1 - 1)^2 + (2x_1 + x_2 - 3)^2$$

$$\begin{aligned}\frac{\partial f(x_1, x_2)}{\partial x_1} &= 2(x_1 - 1) + 2(2x_1 + x_2 - 3) \cdot 2 = 0 \\ &= 10x_1 + 4x_2 - 14 = 0\end{aligned}$$

$$\begin{aligned}\frac{\partial f(x_1, x_2)}{\partial x_2} &= 2(2x_1 + x_2 - 3) = 0 \\ &= 4x_1 + 2x_2 - 6 = 0\end{aligned}$$

$$\begin{aligned}\Rightarrow x_2 &= 3 - 2x_1 \\ \Rightarrow 2x_1 - 2 &= 0 \\ \Rightarrow x_1 = 1 &\Rightarrow x_2 = 1\end{aligned}$$

$$x^* = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

ב

$$A = \begin{pmatrix} 1 & 0 \\ 2 & 1 \end{pmatrix}, b = \begin{pmatrix} 1 \\ 3 \end{pmatrix}$$

עמודות A בת"ל ולכן נשתמש בנוסחה: $x^* = (A^T A)^{-1} A^T b = \underbrace{A^{-1} b}_{\text{הפיכה } A}$

$$x^* = \frac{1}{1} \begin{pmatrix} 1 & 0 \\ -2 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 3 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

ג

למעשה הראינו זאת בסעיף קודם בכך שהשתמשנו בכך ש A בת"ל וריבועית ולכן לכל וקטור b קיים רק פתרון יחיד. ובכך פתרנו את מערכת המשוואות $Ax = b$ וראינו שיש פתרון יחיד.

ד

לא, כתבנו את המשפט המוכיח זאת בסעיף ג. (A בת"ל וריבועית ולכן לכל וקטור b קיים רק פתרון יחיד).

Question 3

א

b. אם $m < n$, אז יש אינסוף פתרונות לבעיית LS

ב

e. לא קיים וקטור \underline{p} שעונה על שתי הדרישות

ג

a. ההטלה הקרובה ביותר של \underline{p} על מרחב העמודות של כל אחת משלוש המטריצות זהה, והינה:

$$\underline{y} = \mathbf{A}_1(\mathbf{A}_1^T \mathbf{A}_1)^{-1} \mathbf{A}_1^T \underline{b}$$

Question 4

א

$$P_A A = A(A^T A)^{-1} A^T A = A(A^T A)^{-1} (A^T A) = AI = A$$

ב

$$P_A P_A = A(A^T A)^{-1} \underbrace{A^T A(A^T A)^{-1} A^T}_I = A(A^T A)^{-1} A^T = P_A$$

ג

$$P_A^T = (A(A^T A)^{-1} A^T)^T \underbrace{=}_{(A^T A)^T = A^T A} A^{TT} (A^T A)^{-1} A^T \underbrace{=}_{A^{TT} = A} A(A^T A)^{-1} A^T = P_A$$

ד

$$H_A = I - P_A$$

$$H_A^T = (I - P_A)^T = I^T - P_A^T \underbrace{=}_{=I^T, \text{ טענה ג}} I - P_A \blacksquare$$

$$H_A H_A = (I - P_A)(I - P_A) = I - 2P_A + P_A P_A \stackrel{\substack{= \\ P_A P_A = P_A}}{=} I - P_A \blacksquare$$

ה

$$A = [M_1 \ M_2]$$

נשים לב כי:

$$[P_A M_1 \ P_A M_2] = P_A [M_1 \ M_2] = P_A \underbrace{A}_{\text{סעיף ג}} \stackrel{=}{=} A = [M_1 \ M_2] \Rightarrow \underbrace{P_A M_2}_{(1)} = M_2$$

$$\begin{aligned} P_{M_2} P_A &= M_2 (M_2^T M_2)^{-1} M_2^T P_A = M_2 (M_2^T M_2)^{-1} (P_A^T M_2)^T \\ &\stackrel{\substack{= \\ P_A^T = P_A}}{=} M_2 (M_2^T M_2)^{-1} (P_A M_2)^T \stackrel{\substack{= \\ (1)}}{=} M_2 (M_2^T M_2)^{-1} (M_2)^T = P_{M_2} \end{aligned}$$

ו

$$H_{M_2} H_A = (I - P_{M_2})(I - P_A) = I - P_A - P_{M_2} + \underbrace{P_{M_2} P_A}_{P_{M_2}} = I - P_A = H_A$$

Question 5

1.

א.

בקוד המקורי אנו בחרנו 4000 תמונות של ספרה 0 ו-4000 של הספרה 7 ועליהן אימנו את המסווג. כעת בחרנו 4000 תמונות של הספרה 0 ועוד 4000 תמונות שמכילות ספרות שאינן 0 ועליהן אימנו את המסווג. שינויים בקוד:

• מחיקת הפרמטר המיוטר digit2.

```
3 %% ===== Parameters =====
4 N = 4000;
5 digit1 = 0;
6 digit2 = 7;
7
```

• בבחירת שאר התמונות שינינו את התנאי כך שהן שונות מהספרה הראשונה (0).

```
20 imagesPerDigit2 = training.images(:,:,training.labels ~= digit1);
21
```

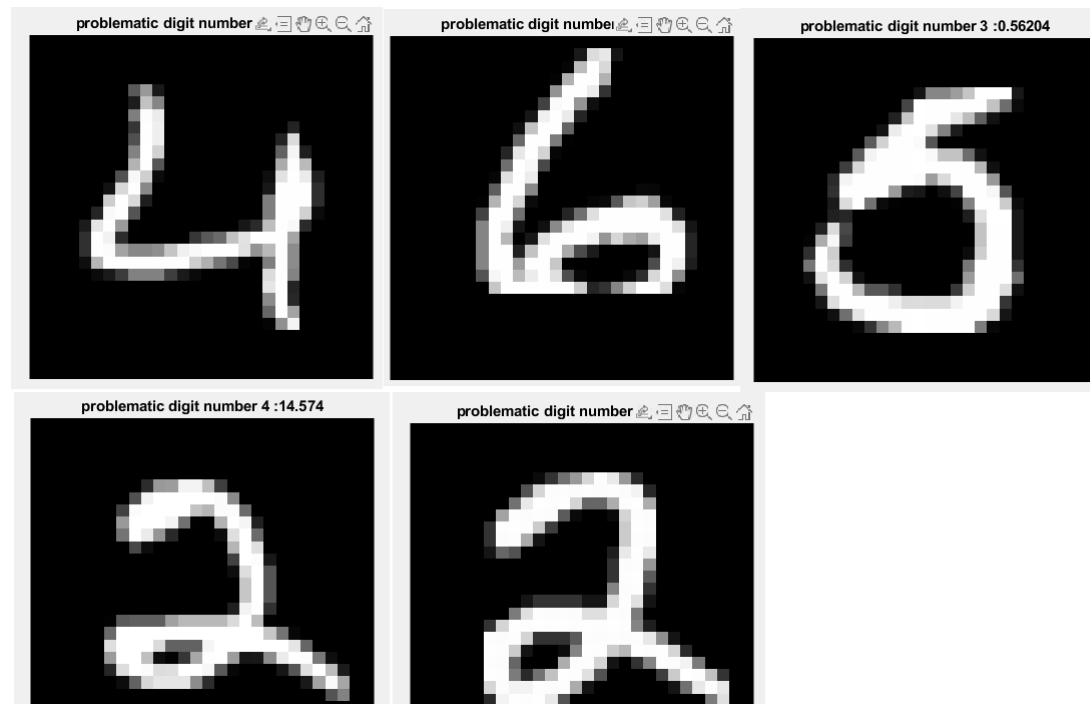
• הכנסנו להערה את הקוד שאמור להדפיס 100 תמונות מהסטים שאנו מאמנים לפיהן. ושינינו גם את הקוד כך שבסוף ידפיס רק 5 תמונות של שגיאה.

ב.

בקבוצת האימון 37 ובקבוצת המבדק 142:

Train Error:
Accuracy=99.075% (37 wrong examples)
Test Error:
Accuracy=96.45% (142 wrong examples)

↗

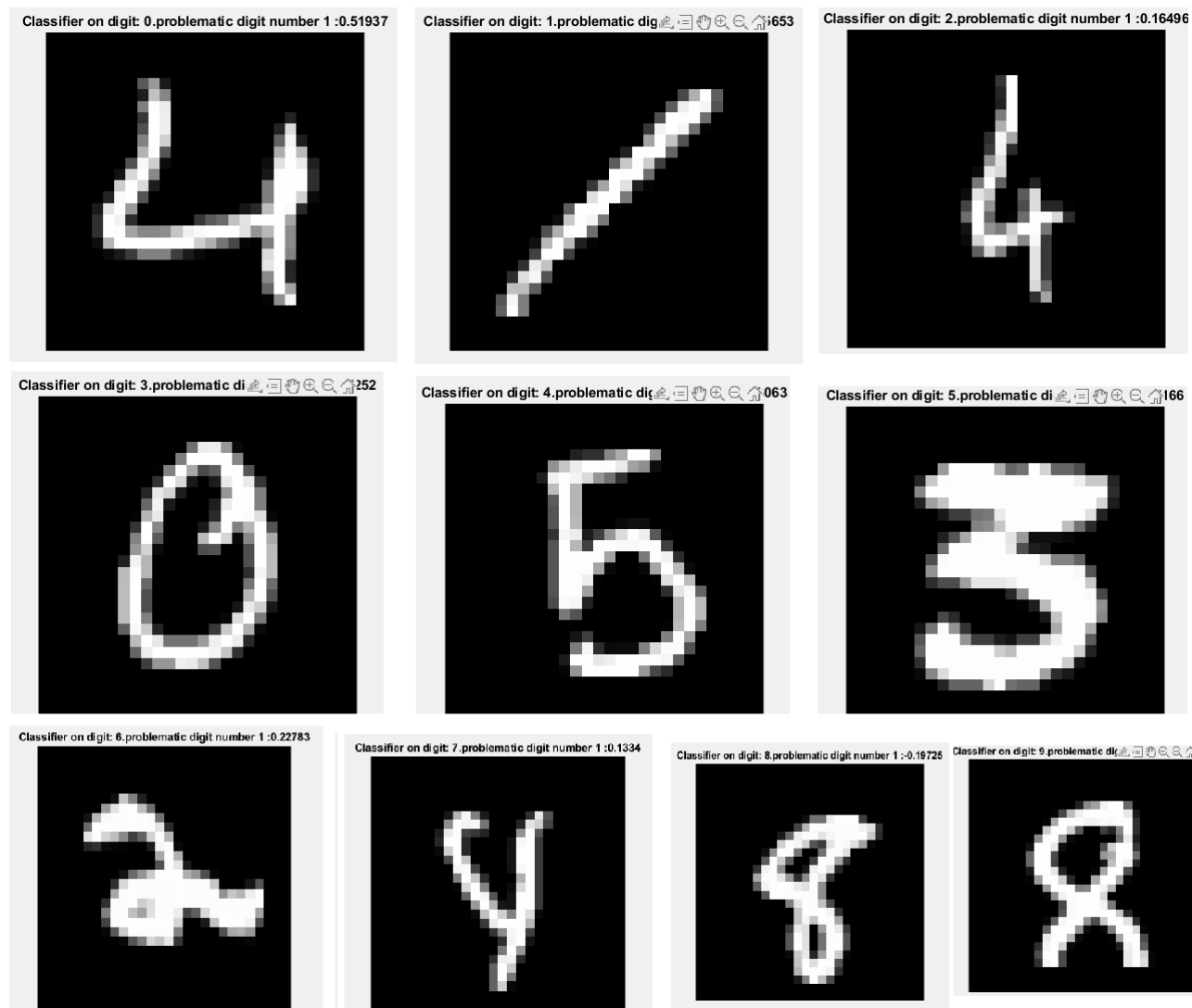


.2

.N

Train Error: Classifier on digit: 0. Accuracy=99.075% (37 wrong examples) Test Error: Classifier on digit: 0. Accuracy=96.45% (142 wrong examples)
Train Error: Classifier on digit: 1. Accuracy=98.325% (67 wrong examples) Test Error: Classifier on digit: 1. Accuracy=95.675% (173 wrong examples)
Train Error: Classifier on digit: 2. Accuracy=95.325% (187 wrong examples) Test Error: Classifier on digit: 2. Accuracy=91.75% (330 wrong examples)
Train Error: Classifier on digit: 3. Accuracy=95.2% (192 wrong examples) Test Error: Classifier on digit: 3. Accuracy=91.5% (340 wrong examples)
Train Error: Classifier on digit: 4. Accuracy=96.875% (125 wrong examples) Test Error: Classifier on digit: 4. Accuracy=94.1% (236 wrong examples)
Train Error: Classifier on digit: 5. Accuracy=95.825% (167 wrong examples) Test Error: Classifier on digit: 5. Accuracy=90.25% (390 wrong examples)
Train Error: Classifier on digit: 6. Accuracy=98.05% (78 wrong examples) Test Error: Classifier on digit: 6. Accuracy=95.475% (181 wrong examples)
Train Error: Classifier on digit: 7. Accuracy=97.45% (102 wrong examples) Test Error: Classifier on digit: 7. Accuracy=94.1% (236 wrong examples)
Train Error: Classifier on digit: 8. Accuracy=92.675% (293 wrong examples) Test Error: Classifier on digit: 8. Accuracy=89.15% (434 wrong examples)
Train Error: Classifier on digit: 9. Accuracy=94.325% (227 wrong examples) Test Error: Classifier on digit: 9. Accuracy=89.925% (403 wrong examples)

ب.



3.

הדרך שאנו מציעים היא לבנות היסטוגרמה (מטריצה של 10×10000) לכל תמונה (מתוך ה-10000) 10 אופציות של ספרות אפשריות. כזכור נשתמש ב 10 המסווגים מסעיף ב. נכפיל את המטריצה $A_{newtest}$ באחד אחד מהמסווגים, נקבל ווקטור שנפעיל עליו את הפעולות $sign$. כעת לכל אינדקס שבו יש 1 עבור מספר מסווג כלשהו, נסמן בהיסטוגרמה עבור האינדקס (שמייצג תמונה) ובעמודה של הספרה שהמסווג מציין. לאחר מכן נעבור על ההיסטוגרמה ולכל תמונה אם יש רק עמודה אחת בה יש 1 אז אנו יודעים לאיזו ספרה צריך לסווג את התמונה הזו.

א.

יש לנו 5060 תמונות (50 אחוזים) שלא קיבלנו לגביהן תשובה חד משמעית. הדרך שנציע לפתרון. ניצור מטריצה של $785 \times 10 \times 10$ אשר מייצגת לכל i, j וקטור עמודה שהוא המסווג בין שני המספרים. כעת עבור כל תמונה שבהיסטוגרמה שלה נראה שיש לה בעמודות שמייצגות את הספרות כמה אופציות לייצוג, נשווה לפי מסווג בין ספרה לספרה (כפי שהוצג בתרגול וכזכור עם דיוק גבוה יותר) נעבור עבור כל שתי אופציות אפשריות עד שנישאר עם אחת בודדת שהיא הנכונה.

ב.

בדרך החדשה זיהינו לא נכון רק 10 אחוזים (כאלף תמונות מתוך עשרת האלפים) [אשר מתוכן 85 בכלל לא סווגו במסווג הראשוני של קדם סעיף 3 לאף ספרה].

.)

