



מבוא למערכות לומדות (236756)

סמסטר חורף תשפ"ב – 07 במרץ 2022

מרצה: ד"ר יונתן בלינקוב

מבחן מסכם מועד ב'

הנחיות הבחינה:

- **משך הבחינה:** 3 שעות.
- **חומר עזר:** המבחן בחומר סגור (ללא ספרים, מחברות, דפי נוסחאות).
- אין צורך במחשבון.
- מותר לכתוב בעט **בלבד**.
- מותר לענות בעברית או באנגלית.
- יש לכתוב את התשובות **על גבי שאלון זה** בכתב יד קריא. תשובה בכתב יד לא קריא – לא תיבדק.
- במבחן 16 עמודים ממוספרים סה"כ, כולל עמוד שער זה שמספרו 1 ושלושה עמודי טיוטה בסוף הגיליון.
- נא לכתוב רק את המבוקש ולצרף הסברים קצרים עפ"י ההנחיות.
- **בתום המבחן יש להגיש את שאלון זה בלבד.**

מבנה הבחינה:

- **חלק א' [76 נק']:** 4 שאלות פתוחות.
- **חלק ב' [24 נק']:** 4 שאלות סגורות (אמריקאיות) [כל אחת 6 נק'].

בהצלחה!

חלק א' – שאלות פתוחות [76 נק']

שאלה 1 – Linear regression & Optimization [14 נק']

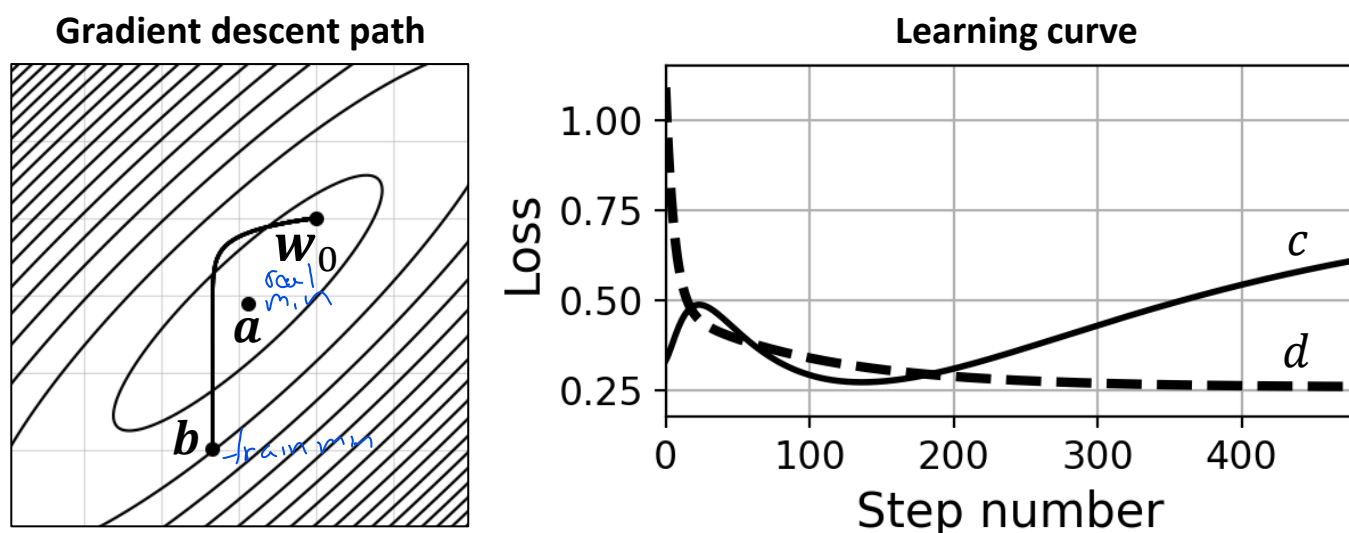
נתונה בעיית רגרסיה ליניארית דו-ממדית (בעיה בשני פרמטרים): $\operatorname{argmin}_{\mathbf{w} \in \mathbb{R}^2} \frac{1}{m} \sum_{i=1}^m (\mathbf{w}^T \mathbf{x}_i - y_i)^2$.

אוספים דאטה S ומחלקים אותו לסט אימון ולסט ואלידציה.

מתחילים מווקטור $\mathbf{w}_0 = \mathbf{0}$ ופותרים את הבעיה (עבור סט האימון) בעזרת gradient descent (לא SGD) עם גודל צעד η .

בתרשים השמאלי: המסלול המלא שנוצר מאימון עם GD החל מ- \mathbf{w}_0 (המסלול מתואר ע"י עקומה במרחב \mathbb{R}^2 , ומראה את כל הפתרונות $\mathbf{w}_0 = \mathbf{0}, \mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_{480} = \mathbf{b}$). קווי המתאר (ה-level sets) מתארים loss landscape שהמינימום שלו הוא בנקודה a . עליכם להבין האם מדובר ב-loss על ה-training או ה-validation.

בתרשים הימני: מופיע גרף ההתכנסות המראה את ה-training loss ואת ה-validation loss.



א. [4 נק'] התאימו בין הפתרונות a, b שבתרשים השמאלי לבין הפיתרון האופטימלי על סט האימון $\mathbf{w}_{\text{train}}^*$ והפיתרון האופטימלי על סט האלידציה $\mathbf{w}_{\text{val}}^*$. התאימו בין העקומות c, d לבין ה-training loss וה-validation loss. מלאו את המקומות הריקים באותיות a, b, c, d כנדרש.

$\mathbf{w}_{\text{train}}^*$ matches b,
השלימו

$\mathbf{w}_{\text{val}}^*$ matches a,
השלימו

training loss matches d,
השלימו

validation loss matches c,
השלימו

שימו לב כיצד שני ה-losses (ובפרט ה-training loss) לא יורדים מתחת 0.25.

ב. [5 נק'] אילו מהפתרונות הבאים עשויים לשפר את ה-training loss בסוף האימון (ביחס לתוצאות המוצגות לעיל)?

סמנו את כל התשובות המתאימות.

a. ☒ הוספת רגולריזציה ℓ^2 .

b. ☒ שימוש במדיניות early stopping (עצירת ה-GD לפני התכנסות, לפי קריטריון כלשהו).

c. ☒ אימון עם SGD (עם $\text{batch_size}=1$) במקום GD, במשך מספר צעדים זהה (480).

d. ☒ מיפוי של שני ה-features המקוריים ל-feature mapping פולינומיאלי.

e. ☒ סיבוב מערכת הצירים של ה-features המקוריים (ב-dataset S כולו) ב- 45° סביב ראשית הצירים ($\mathbf{w}_0 = \mathbf{0}$).

ג. [5 נק'] אילו מהפתרונות הבאים עשויים לשפר את ה-validation loss בסוף האימון (ביחס לתוצאות המוצגות לעיל)?

סמנו את כל התשובות המתאימות.

a. ☒ הוספת רגולריזציה ℓ^2 .

b. ☒ שימוש במדיניות early stopping (עצירת ה-GD לפני התכנסות, לפי קריטריון כלשהו).

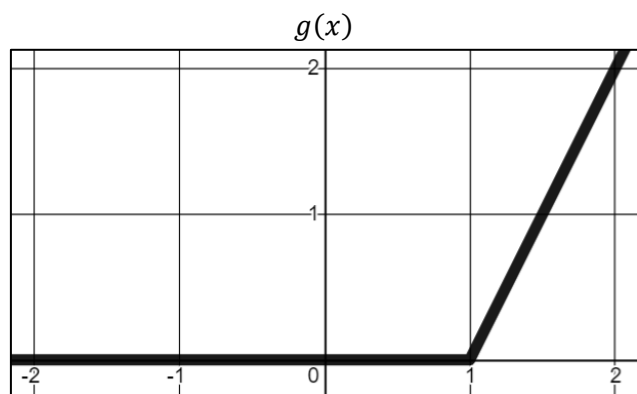
c. ☒ אימון עם SGD (עם $\text{batch_size}=1$) במקום GD, במשך מספר צעדים זהה (480).

d. ☒ מיפוי של שני ה-features המקוריים ל-feature mapping פולינומיאלי.

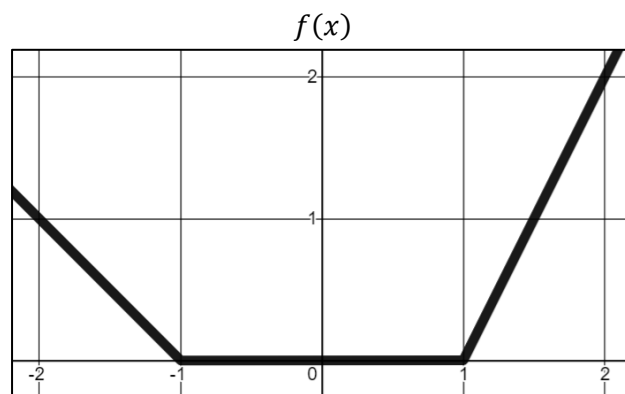
e. ☒ סיבוב מערכת הצירים של ה-features המקוריים (ב-dataset S כולו) ב- 45° סביב ראשית הצירים ($\mathbf{w}_0 = \mathbf{0}$).

שאלה 2 – Deep learning [20 נק']

נתונות שתי הפונקציות הרציפות $f, g: \mathbb{R} \rightarrow \mathbb{R}$ שבתרשימים הבאים:



בתחום $(-\infty, 1]$ הפונקציה היא אפס.
בתחום $(1, \infty)$ השיפוע של g הוא 2.

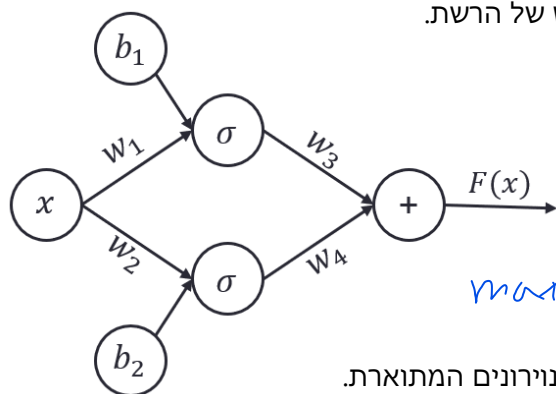


בתחום $[-1, 1]$ הפונקציה היא אפס.
בתחום $(1, \infty)$ השיפוע הוא 2 ובתחום $(-\infty, -1)$ הוא -1.

נרצה ללמוד את הפונקציות f, g בעזרת רשת הנורונים הבאה:

$$F(x) = w_3 \cdot \sigma(w_1 \cdot x + b_1) + w_4 \cdot \sigma(w_2 \cdot x + b_2)$$

כאשר $w_1, w_2, w_3, w_4, b_1, b_2 \in \mathbb{R}$ הם פרמטרים סקלריים ו- $x \in \mathbb{R}$ הוא הקלט של הרשת.



פונקציית האקטיבציה σ יכולה להיות אחת מהשתיים:

1. סיגמואיד, משמע $\sigma(z) = \frac{1}{1+e^{-z}}$

2. ReLU, משמע $\sigma(z) = \max\{0, z\}$

$$\max\{0, 2x-2\}$$

בשני הסעיפים הבאים נראה שניתן לממש את הפונקציות f, g בעזרת רשת הנורונים המתוארת.

א. [4 נק'] נבחר $w_2 = b_2 = w_4 = 0$. כתבו את ערכי w_1, w_3, b_1 ואת הבחירה של σ שמקיימים: $\forall x \in \mathbb{R}: F(x) = g(x)$.

תשובה סופית (לרשותכם דפי טיוטה בסוף הגיליון):

First layer: $w_1 = \frac{1}{2}$; $b_1 = -2$; **Second layer:** $w_3 = \frac{2}{1}$.

Activation: Sigmoid or ReLU (circle your choice).

$$\max\{0, -x-1\}$$

ב. [5 נק'] כתבו את ערכי $w_1, w_2, w_3, w_4, b_1, b_2$ ואת הבחירה של σ שמקיימים: $\forall x \in \mathbb{R}: F(x) = f(x)$.

תשובה סופית:

First layer: $w_1 = \frac{2}{2}$; $w_2 = -1$; $b_1 = -2$; $b_2 = -1$.

Second layer: $w_3 = \frac{1}{1}$; $w_4 = \frac{1}{1}$.

Activation: Sigmoid or ReLU (circle your choice).

הערה: סעיף ג' לא תלוי בסעיפים הקודמים.

ג. [4 נק'] הציגו פונקציית loss שמתאימה לבעיית הרגרסיה שהוגדרה, כך שמזעור של $\sum_{i=1}^m \ell(f(x_i), F(x_i))$ יביא ללמידת פרמטרים מתאימים שיקיימו $F(x) = f(x)$.

Answer: $\ell(a, b) = \underbrace{(a-b)^2}_{\text{השלימו}}.$
 $\ell: \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}_{\geq 0}$

הערה: סעיף ד' תלוי בסעיף ב' רק דרך הבחירה של פונקציית האקטיבציה.

בסעיף הבא נבחן את הקמירות של הבעיה שנוצרה.

תזכורת: הפונקציה $g: \mathbb{R} \rightarrow \mathbb{R}$ נקראת פונקציה קמורה אם מתקיים
 $\forall z_1, z_2 \in \mathbb{R}, \forall t \in [0, 1]: tg(z_1) + (1-t)g(z_2) \geq g(tz_1 + (1-t)z_2)$

ד. [7 נק'] בסעיף זה נניח שוב $w_2 = b_2 = w_4 = 0$.

הוכיחו / הפריכו: הפונקציה $\ell(a, F(x))$ קמורה ביחס לפרמטר w_1 (בהינתן כל בחירה של (b_1, w_3, a, x)).

ניתן לענות לפי הגדרת הקמירות או לפי מאפיינים שלמדנו (אך יש לציין אותם במפורש).

שימו לב: עליכם להשתמש בבחירה של σ מסעיף ב' ובבחירה של ℓ מסעיף ג'. אין להציב ערכים מסעיף א'.

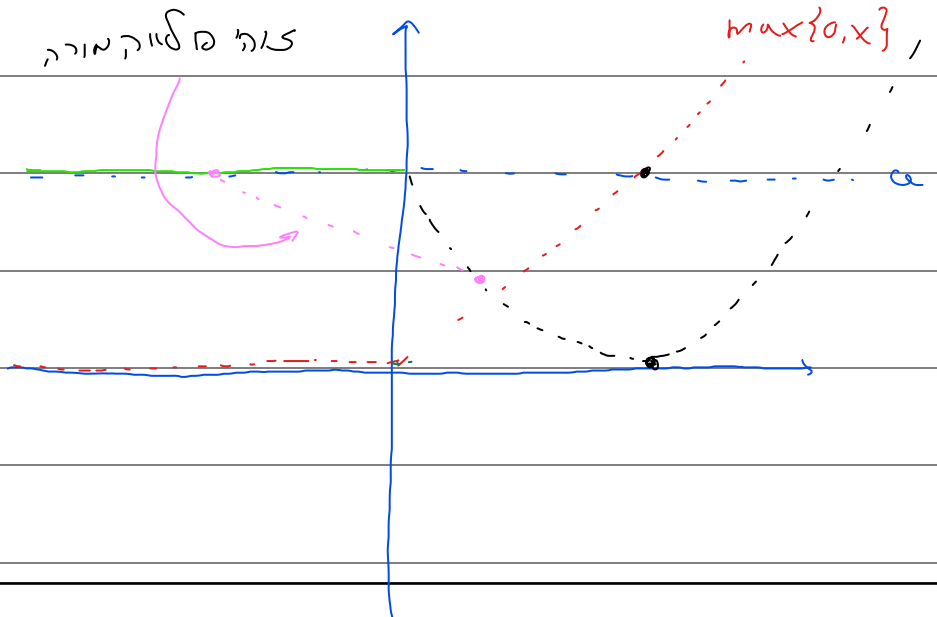
תשובה (לרשותכם דפי טיוטה בסוף הגיליון):

$$F(x) = w_3 \cdot \max\{0, w_1 x + b_1\}$$

$$\ell(a, f(x)) = (a - w_3 \max\{0, w_1 x + b_1\})^2$$

זוג פונקציות

$\max\{0, x\}$



שאלה 3 – Naïve Bayes [18 נק']

בשאלה זו נראה ש-Gaussian Naïve Bayes הינו מסווג לינארי.

נתון דאטה $S = \{(x_i, y_i)\}_{i=1}^m$ עם דוגמאות $x_i \in \mathbb{R}^d$ ותיוגים $y_i \in \{0,1\}$.

לפי מידול Gaussian Naïve Bayes, נניח שההתפלגות המותנה של כל כניסה $k = 1, \dots, d$ הינה:

$$X[k] | Y = y \sim \mathcal{N}(\mu_y[k], \sigma[k]^2)$$

כאשר $X[k]$ הוא המשתנה המקרי המתאים לכניסה ה- k ב- x , המסומנת $x[k]$.

תזכורת: פונקציית הצפיפות של התפלגות גאוסיאנית $\mathcal{N}(\mu, \sigma^2)$ נתונה ע"י:

$$P(z) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(z-\mu)^2}{2\sigma^2}\right)$$

א. [4 נק'] הוכיחו שמתקיים:

$$P(X = x | Y = 1) = \left(\prod_{k=1}^d \frac{1}{\sigma[k]\sqrt{2\pi}}\right) \exp\left(-\sum_{k=1}^d \frac{1}{2\sigma[k]^2} (x[k] - \mu_1[k])^2\right)$$

הוכחה מנומקת:

$$P(Y = 1 | X = x) = \frac{1}{1 + \frac{P(Y = 0)P(X = x | Y = 0)}{P(Y = 1)P(X = x | Y = 1)}}$$
$$\frac{P(Y = 0)P(X = x | Y = 0)}{P(Y = 1)P(X = x | Y = 1)} = \frac{1-p}{p} \cdot \exp\left(\sum_{k=1}^d \left(\frac{\mu_0[k] - \mu_1[k]}{\sigma[k]^2} x[k] + \frac{\mu_1[k]^2 - \mu_0[k]^2}{2\sigma[k]^2}\right)\right)$$

ב. [7 נק'] בעזרת האמור לעיל, הוכיחו שמתקיים:

הוכחה:

[illegible]

ג. [7 נק'] בהסתמך על האמור לעיל, הוכיחו כי מתקבל כלל החלטה ליניארי.

משמע, ההיפותזה $h(x) = \operatorname{argmax}_{y \in \{0,1\}} P(Y = y | X = x)$ מְשָׁרָה גבול החלטה ליניארי (decision boundary).

הוכחה תמציתית:

שאלה 4 – SVM [24 נק']

היזכרו בבעיות ה-SVM במקרה ההומוגני (נניח שמתקיים $\lambda = 1$ בבעיה ה-Soft):

Hard SVM

$$\operatorname{argmin}_{\mathbf{w} \in \mathbb{R}^d} \|\mathbf{w}\|_2^2$$

$$\text{s.t. } y_i \cdot \mathbf{w}^\top \mathbf{x}_i \geq 1, \forall i \in [m]$$

Soft SVM

$$\operatorname{argmin}_{\mathbf{w} \in \mathbb{R}^d} \left(\frac{1}{m} \sum_{i=1}^m \max\{0, 1 - y_i \cdot \mathbf{w}^\top \mathbf{x}_i\} + \|\mathbf{w}\|_2^2 \right)$$

נתון דאטה d -ממדי $\{(x_i, y_i)\}_{i=1}^m$ עם סיווגים בינאריים (± 1). ידוע שהדאטה פריד ליניארית ע"י מפריד הומוגני.

צוות מחקר פתר את שתי בעיות האופטימיזציה שלמעלה, וקיבל את $\mathbf{w}_{\text{hard}}, \mathbf{w}_{\text{soft}} \in \mathbb{R}^d$ כפתרונות.

א. [8 נק'] האם ניתן לומר שאחד מהמקרים $\|\mathbf{w}_{\text{hard}}\|_2 \leq \|\mathbf{w}_{\text{soft}}\|_2$ או $\|\mathbf{w}_{\text{hard}}\|_2 \geq \|\mathbf{w}_{\text{soft}}\|_2$ מתקיים בהכרח?

אם כן, איזה מהם? בכל מקרה – הסבירו בקצרה.

Handwritten note: Bigger Norm \rightarrow smaller Margin

הסבר תמציתי:

Handwritten: $\|\mathbf{w}_{\text{hard}}\|_2 = \|\mathbf{w}_{\text{soft}}\|_2$ מכיוון שדאטה קטנה, לכן ברצון חזק

Handwritten: פשוט הפתרונות של Hard שונים מאלו של Soft ולכן במקרה זה $\|\mathbf{w}_{\text{hard}}\|_2 = \|\mathbf{w}_{\text{soft}}\|_2$

Handwritten: $\lambda = 1$ ולכן הנוסחה היא $\lambda \|\mathbf{w}\|_2^2$ ולכן נרצו למצוא ונוע. ~~אנו~~ וירר

ב. [8 נק'] נוסף עוד feature ממקור לא ידוע. הצוות פתר את בעיית ה-Hard-SVM המעודכנת וקיבל את $\mathbf{w}'_{\text{hard}} \in \mathbb{R}^{d+1}$. האם ניתן לומר שאחד מהמקרים $\|\mathbf{w}_{\text{hard}}\|_2 \leq \|\mathbf{w}'_{\text{hard}}\|_2$ או $\|\mathbf{w}_{\text{hard}}\|_2 \geq \|\mathbf{w}'_{\text{hard}}\|_2$ מתקיים בהכרח? אם כן, איזה מהם? בכל מקרה – הסבירו בקצרה.

הסבר תמציתי:

$\|\mathbf{w}'\|_2 \geq \|\mathbf{w}\|_2$ מכיון שאם הפיצ'ר לא נוסף – נניח, לאט אט

ואז, אט-אט, הפיצ'ר \mathbf{w}_{hard} היה נורמלי (אט אט אט הפיצ'ר ה- $d+1$)

אם הפיצ'ר נוסף $\|\mathbf{w}\|_2 \geq \|\mathbf{w}'\|_2$

ג. [8 נק'] יהי $\{(x_i, y_i)\}_{i=1}^m$ סט אימון כלשהו (לאו דווקא פריד ליניארית) עבורו $x_i \in \mathbb{R}^d, y_i \in \{\pm 1\}$ ויהי $\mathbf{w}_{\text{soft}} \in \mathbb{R}^d \setminus \{\mathbf{0}\}$ הפתרון של Soft-SVM על סט זה.

עליכם להוכיח **אחת** מבין שתי הטענות הבאות (השנייה מזכה בניקוד חלקי בלבד):

(i) [8 מתוך 8 נק'] לכל סט כזה ניתן להוסיף feature חדש (בלי לשנות את ה-features האחרים ואת התיגים הנתונים), כך שמתקיים $\|\mathbf{w}'_{\text{soft}}\|_2 > \|\mathbf{w}_{\text{soft}}\|_2$, כאשר $\mathbf{w}'_{\text{soft}} \in \mathbb{R}^{d+1}$ אופטימלי עבור הבעיה המעודכנת (עם ה-feature החדש).

או

(ii) [4 מתוך 8 נק'] לכל סט כזה ניתן להוסיף feature חדש (בלי לשנות את ה-features האחרים ואת התיגים הנתונים), כך שקיים $\mathbf{w}' \in \mathbb{R}^{d+1}$ עבורו $\|\mathbf{w}'\|_2 > \|\mathbf{w}_{\text{soft}}\|_2$ וגם $\mathcal{L}(\mathbf{w}_{\text{soft}}) \geq \mathcal{L}(\mathbf{w}')$ כאשר מגדירים

$$\mathcal{L}(\mathbf{w}) \triangleq \frac{1}{m} \sum_{i=1}^m \max\{0, 1 - y_i \cdot \mathbf{w}^T \mathbf{x}_i\}, \quad \mathcal{L}'(\mathbf{w}) \triangleq \frac{1}{m} \sum_{i=1}^m \max\{0, 1 - y_i \cdot \mathbf{w}^T \mathbf{x}'_i\}$$

הדוגמאות המעודכנות
עם ה-feature החדש

הוכחה (יש לציין איזו טענה מוכיחים):

(i) להגדיל את הפיצ'ר שלנו כך שיוזגד הטווח הכולל

$$\sqrt{w_1^2 + w_2^2 + \dots + w_d^2}$$

$$K < \frac{1}{\|\mathbf{w}_{\text{soft}}\|_2} \quad \text{כפול} \quad K = \frac{1}{\|\mathbf{w}_{\text{soft}}\|_2 + 1} \quad \text{לכפול}$$

$$\left(\sqrt{w_1^2 + w_2^2 + \dots + w_d^2} \right)$$

$$\mathbf{w}' = [0, 0, \dots, 0, \frac{1}{K}] \quad \text{יהי} \quad \forall i: X_i = K y_i \quad \text{נשא-וגר הפיצ'ר סהוד}$$

$$\mathcal{L}(\mathbf{w}') = \frac{1}{m} \sum_{i=1}^m \max\{0, 1 - K \cdot \frac{1}{K} \cdot y_i \cdot y_i\} + \left(\frac{1}{K}\right)^2 = \left(\frac{1}{K}\right)^2 \quad \text{ומכאן}$$

$$y_i \cdot w'_i \cdot x_i = y_i \cdot K \cdot y_i \cdot \frac{1}{K} = 1$$

$$K y_i y_i$$

$$\frac{1}{K^2} = \|\mathbf{w}'\|_2^2 = \mathcal{L}(\mathbf{w}') \geq \mathcal{L}(\mathbf{w}_{\text{soft}}) \geq \|\mathbf{w}_{\text{soft}}\|_2^2$$

חלק ב' – שאלות רב-ברירה [24 נק']

בשאלות הבאות סמנו את התשובות המתאימות (לפי ההוראות). בחלק זה אין צורך לכתוב הסברים.

א. [6 נק'] היכרו בבעיות רגרסיה ליניארית עם Least squares (LS), וסמנו את כל התשובות הנכונות. שימו לב: בסעיף זה, הגזירות מתייחסות להגדרת ה-gradients המסורתית, ולא ל-subgradients.

(a) כאשר פותרים LS עם הפיצ'רים המקוריים, הבעיה קמורה ביחס ל- w .

(b) כאשר פותרים LS עם feature mapping שנקבע מראש (למשל פולינומיאלי), הבעיה קמורה ביחס ל- w .

(c) Ridge regression (ℓ^2) היא בעיה קמורה אך לא גזירה ביחס ל- w .

(d) Lasso (ℓ^1) היא בעיה קמורה אך לא גזירה ביחס ל- w .

$$\nabla_w^2 \|Xw - y\|^2 \geq 0$$

הא-ה

$$\lambda \|w\|_2^2$$

ב. [6 נק'] הטענות הבאות עוסקות במודלים מסוג Linear Soft SVM, Perceptron, and Logistic Regression. סמנו את כל הטענות הנכונות.

(a) Logistic Regression יכול ללמוד גם מפרידים לא ליניאריים בגלל שה-sigmoid מכניס non-linearity. ~~x~~

(b) ניתן להכליל Logistic Regression לבעיות multiclass בעזרת פונקציית Softmax. (b)

(c) כל עוד הדאטה פריד ליניארית, Soft SVM ופרספטרון מחזירים את אותו המפריד. ~~x~~

(d) בשלושת האלגוריתמים ניתן להשתמש ב-feature mapping כדי ללמוד מפרידים לא ליניאריים. (d)

(e) פרספטרון לומד בעזרת GD (לא stochastic), ואילו Soft SVM יש ללמוד באמצעות SGD. ~~x~~

??

$$\mathcal{L}(z) = (1 - z)^2$$

$$\mathcal{L}'(z) = -2(1 - z) = 2z - 2$$

$$\mathcal{L}''(z) = 2$$

ג. [6 נק'] נגדיר את פונקציית ה-squared loss הבאה:

סמנו את כל הטענות הנכונות ביחס לפונקציה זו.

(a) הפונקציה קמורה (convex) ביחס ל- z . (a)

(b) הנגזרת של הפונקציה היא $\frac{\partial}{\partial z} \mathcal{L} = 2z - 2$. (b)

(c) עבור בעיות סיווג ליניארי (ה-loss מחושב ע"י $z = y_i w^T x_i$ והפרדיקציות ניתנות ע"י $h(x_i) = \text{sgn}(w^T x_i)$). (c)

כאשר ה-training loss הוא 0, גם ה-training error הוא 0.

(d) עבור בעיות סיווג ליניארי (ה-loss מחושב ע"י $z = y_i w^T x_i$ והפרדיקציות ניתנות ע"י $h(x_i) = \text{sgn}(w^T x_i)$). ~~x~~

כאשר ה-training error הוא 0, גם ה-training loss הוא 0.

אם יכול להיות - חיובי אז הט"ל נכון $\leftarrow te = 0$

te

ד. [6 נק'] נגדיר מחלקת היפותזות שמכלילה את המסווגים שמחזיר Adaboost לאחר T צעדים עם מחלקת היפותזות \mathcal{B} בתור מסווגי בסיס. משמע:

$$\mathcal{H}_{\mathcal{B},T} = \{ h_{\text{strong}}(x) = \text{sign}(\sum_{t=1}^T \alpha_t h_t(x)) \mid \alpha \in \mathbb{R}^T, h_t \in \mathcal{B} \}$$

לפניכם מספר טענות על ההשפעה של T ו- $\text{VCdim}(\mathcal{B})$ על $\text{VCdim}(\mathcal{H}_{\mathcal{B},T})$.

בחרו בטענה היחידה הנכונה (השאלה אינה עוסקת במקרי קצה אלא במקרה הסביר).

a) $\text{VCdim}(\mathcal{H}_{\mathcal{B},T}) \Leftarrow T$ ☒ $\text{VCdim}(\mathcal{H}_{\mathcal{B},T}) \Leftarrow \text{VCdim}(\mathcal{B})$ ☐

b) $\text{VCdim}(\mathcal{H}_{\mathcal{B},T}) \Leftarrow T$ ☒ $\text{VCdim}(\mathcal{H}_{\mathcal{B},T}) \Leftarrow \text{VCdim}(\mathcal{B})$ ☐

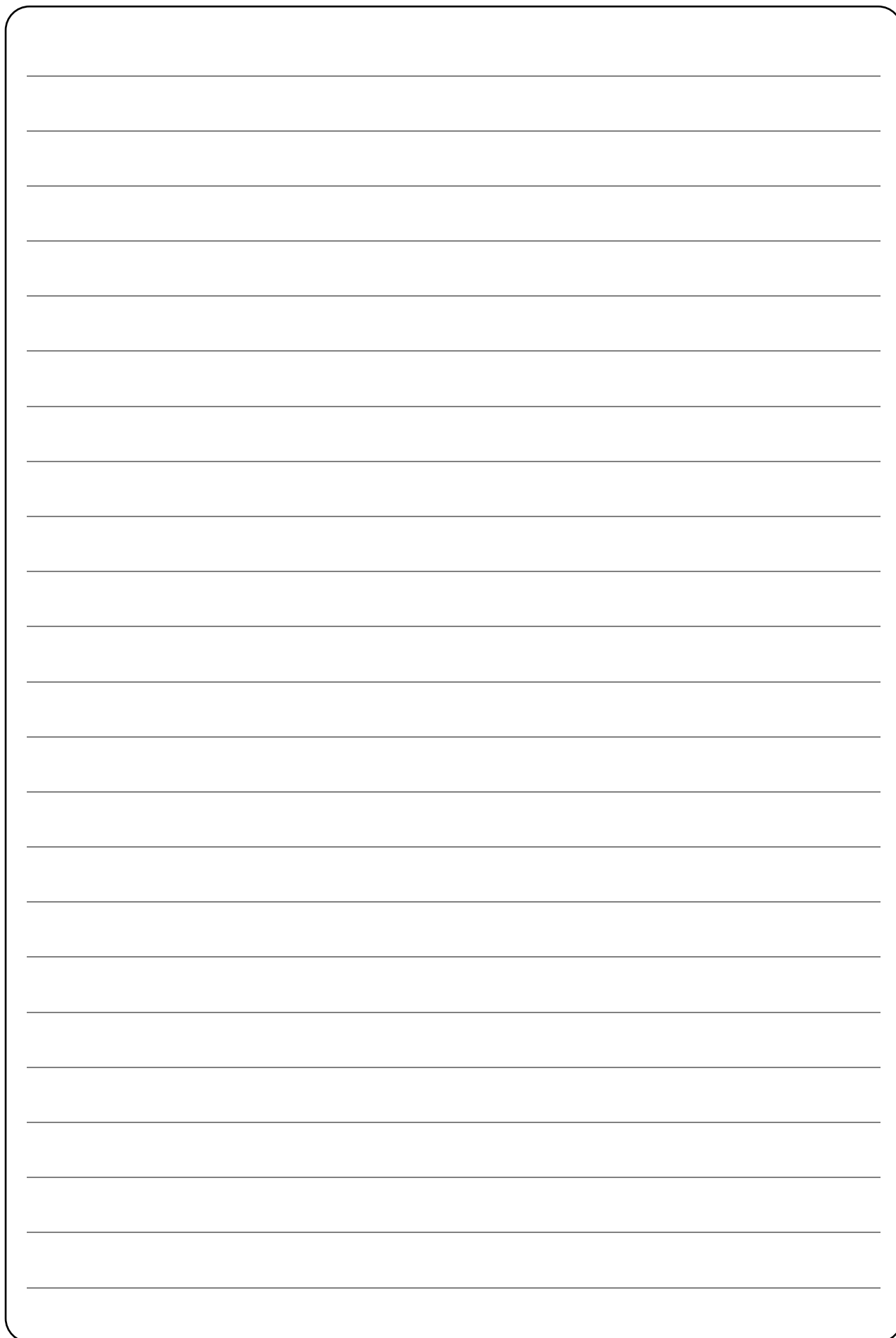
c) $\text{VCdim}(\mathcal{H}_{\mathcal{B},T}) \Leftarrow T$ ☒ $\text{VCdim}(\mathcal{H}_{\mathcal{B},T}) \Leftarrow \text{VCdim}(\mathcal{B})$ ☐

d) $\text{VCdim}(\mathcal{H}_{\mathcal{B},T}) \Leftarrow T$ ☒ $\text{VCdim}(\mathcal{H}_{\mathcal{B},T}) \Leftarrow \text{VCdim}(\mathcal{B})$ ☐

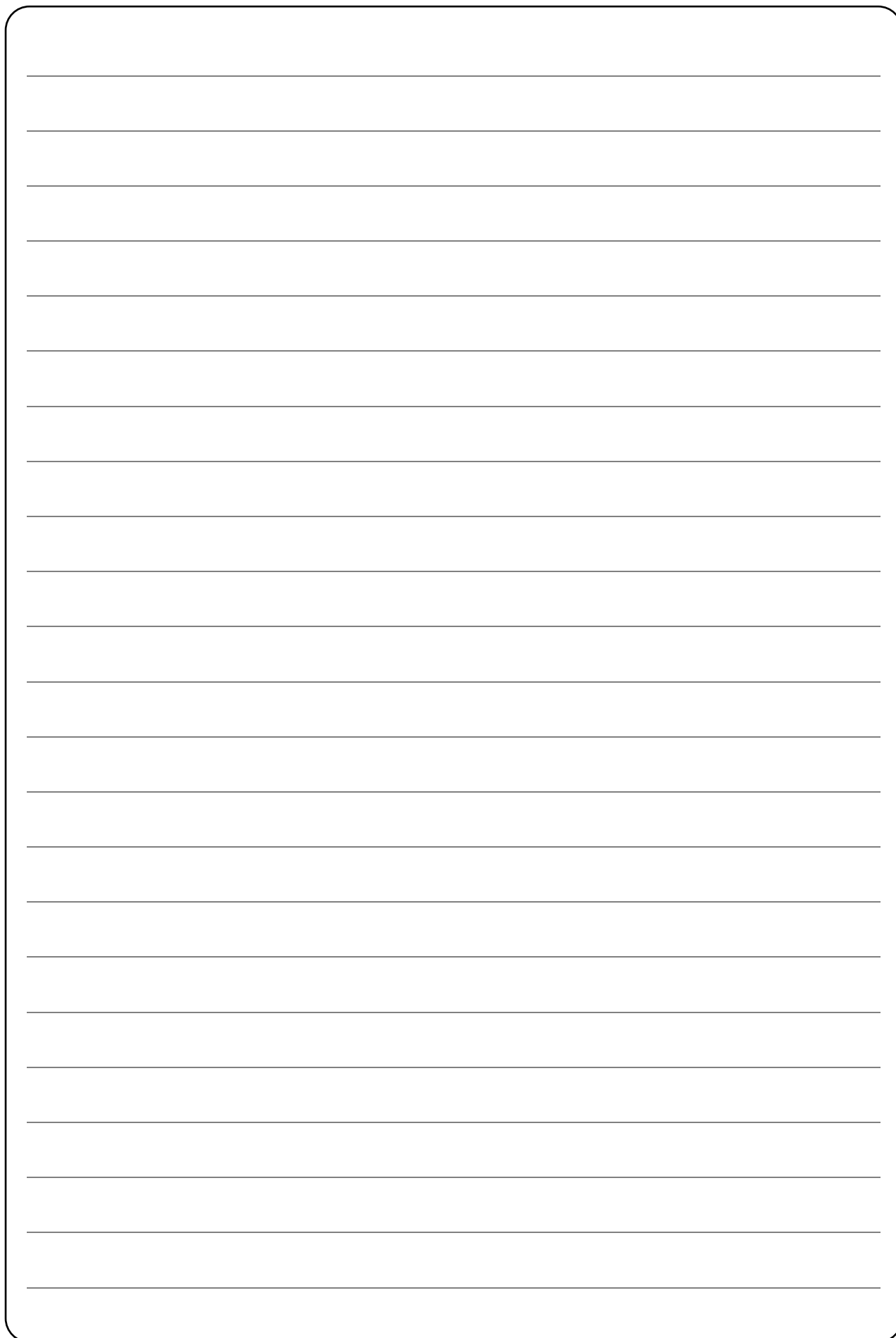
e. רק $\text{VCdim}(\mathcal{B})$ משפיע על $\text{VCdim}(\mathcal{H}_{\mathcal{B},T})$.

f. רק T משפיע על $\text{VCdim}(\mathcal{H}_{\mathcal{B},T})$. ☒

מסגרת נוספת (יש לציין אם מדובר בטייטה או בהמשך לתשובה אחרת):

A large rectangular box with rounded corners, containing 25 horizontal lines for writing. The lines are evenly spaced and extend across the width of the box. The box is intended for providing a second answer or further explanation.

מסגרת נוספת (יש לציין אם מדובר בטייטה או בהמשך לתשובה אחרת):

A large rectangular box with rounded corners, containing 25 horizontal lines for writing. The lines are evenly spaced and extend across the width of the box. The box is intended for a student to provide a second answer or further details to a question.

מסגרת נוספת (יש לציין אם מדובר בטייטה או בהמשך לתשובה אחרת):

A large rectangular box with rounded corners, containing 25 horizontal lines for writing. The lines are evenly spaced and extend across the width of the box.