Machine learning

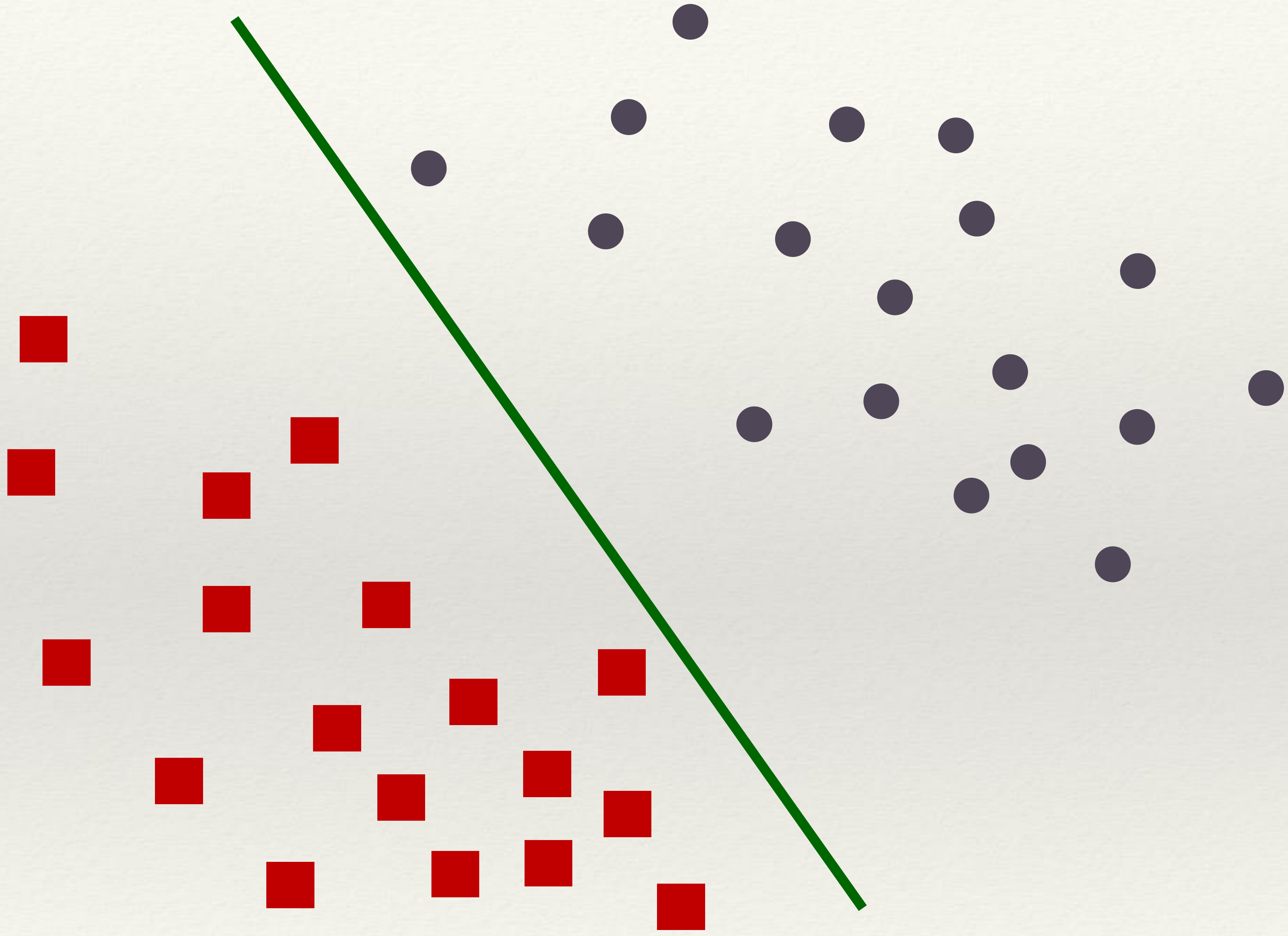
SVM

Exercise X

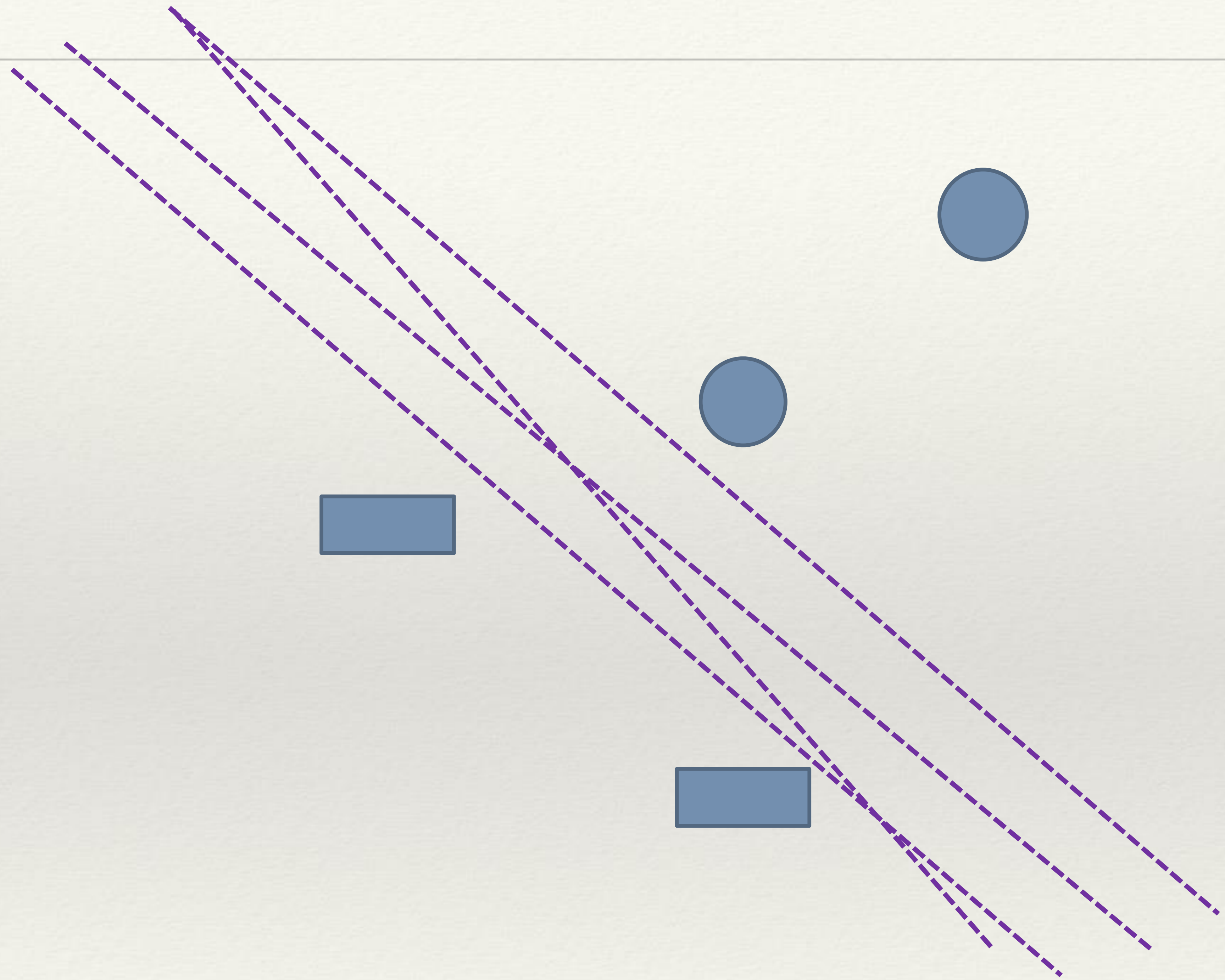
פיתוח:
ד"ר יהונתן שלר
משה פרידמן

מפרידים לינארים
ומושג ה-margin

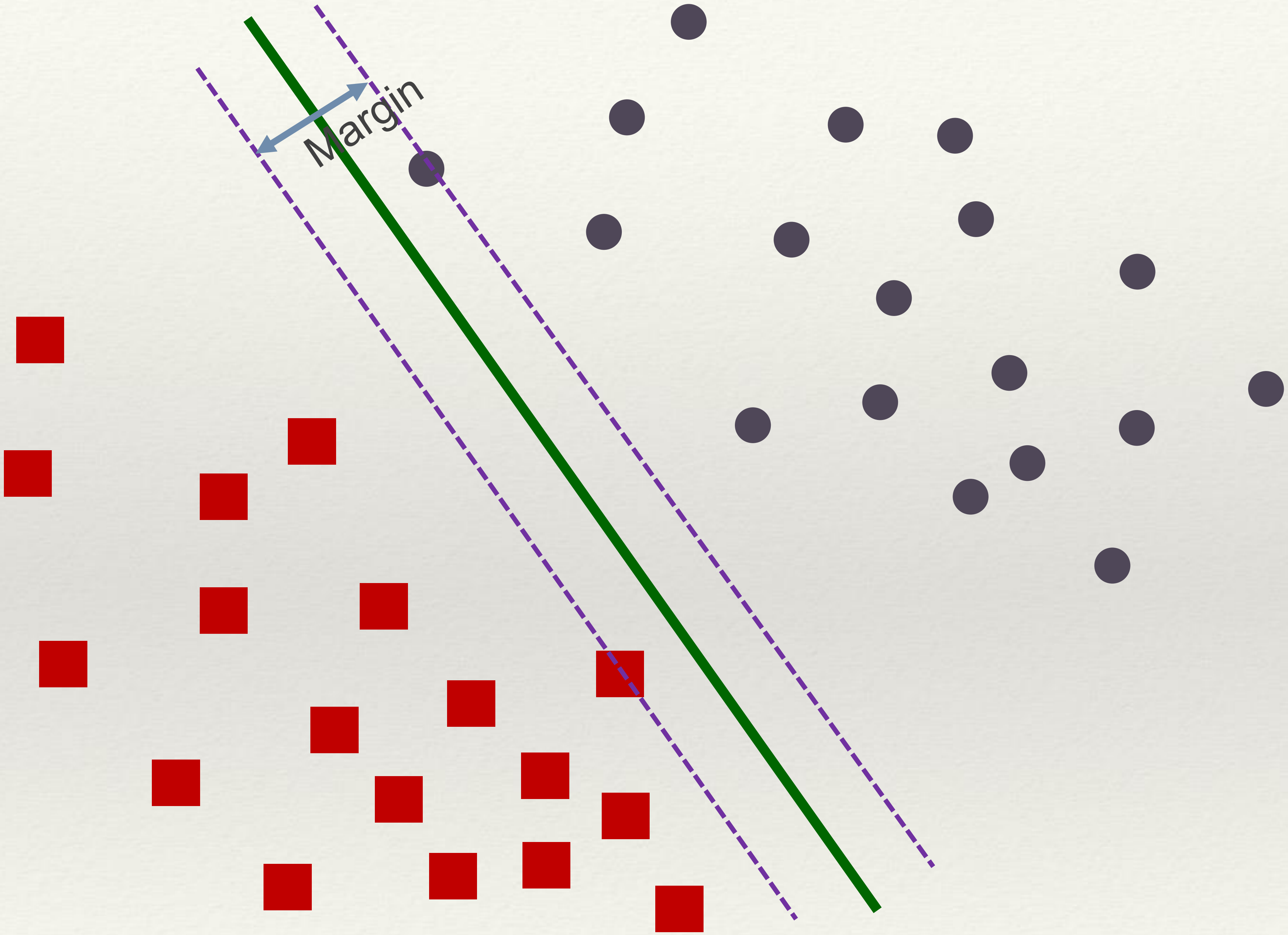
האם יש רק קו מפריד בין המחלקות?



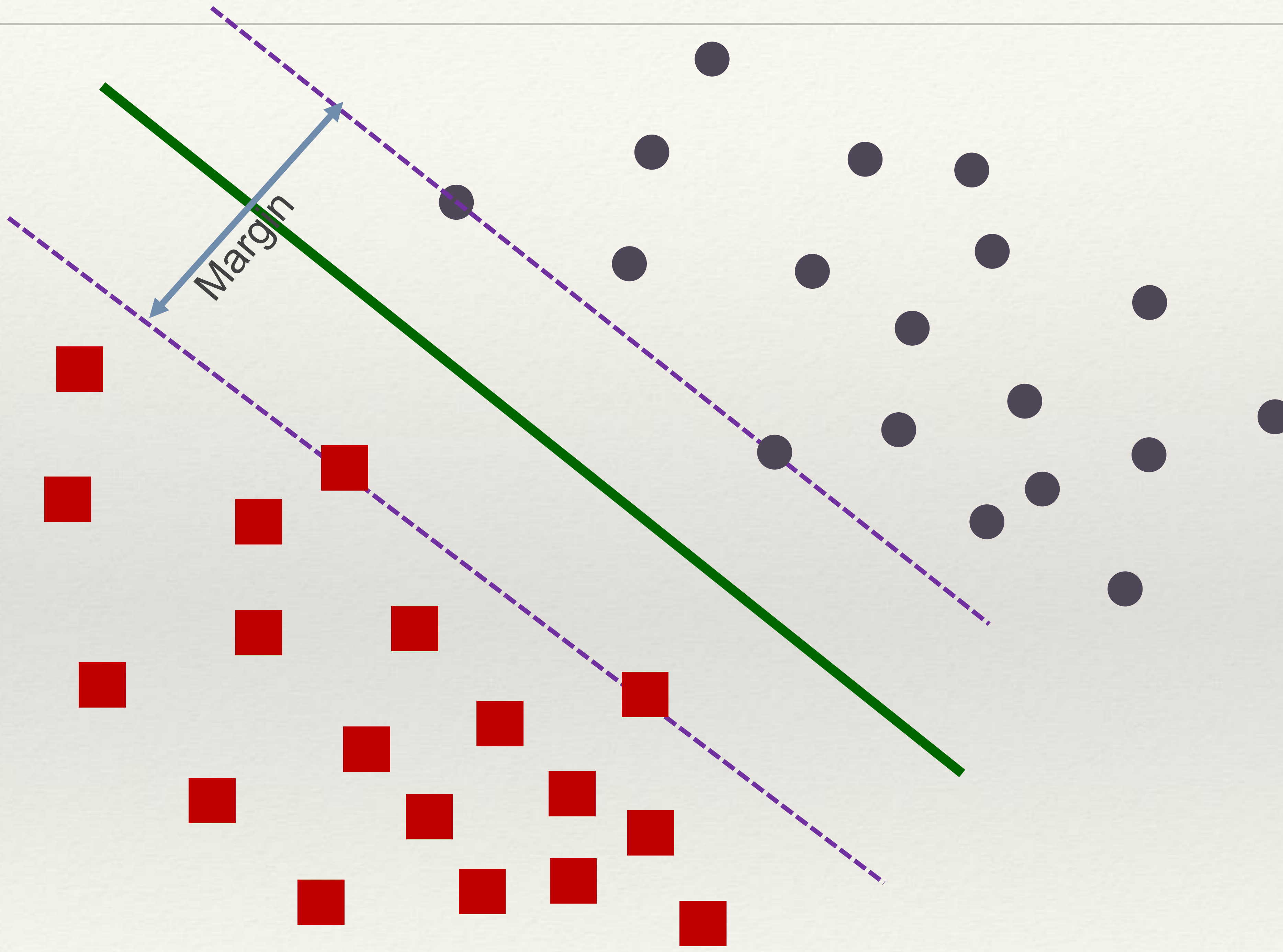
אפשרויות שונות של מפרידים לינארים



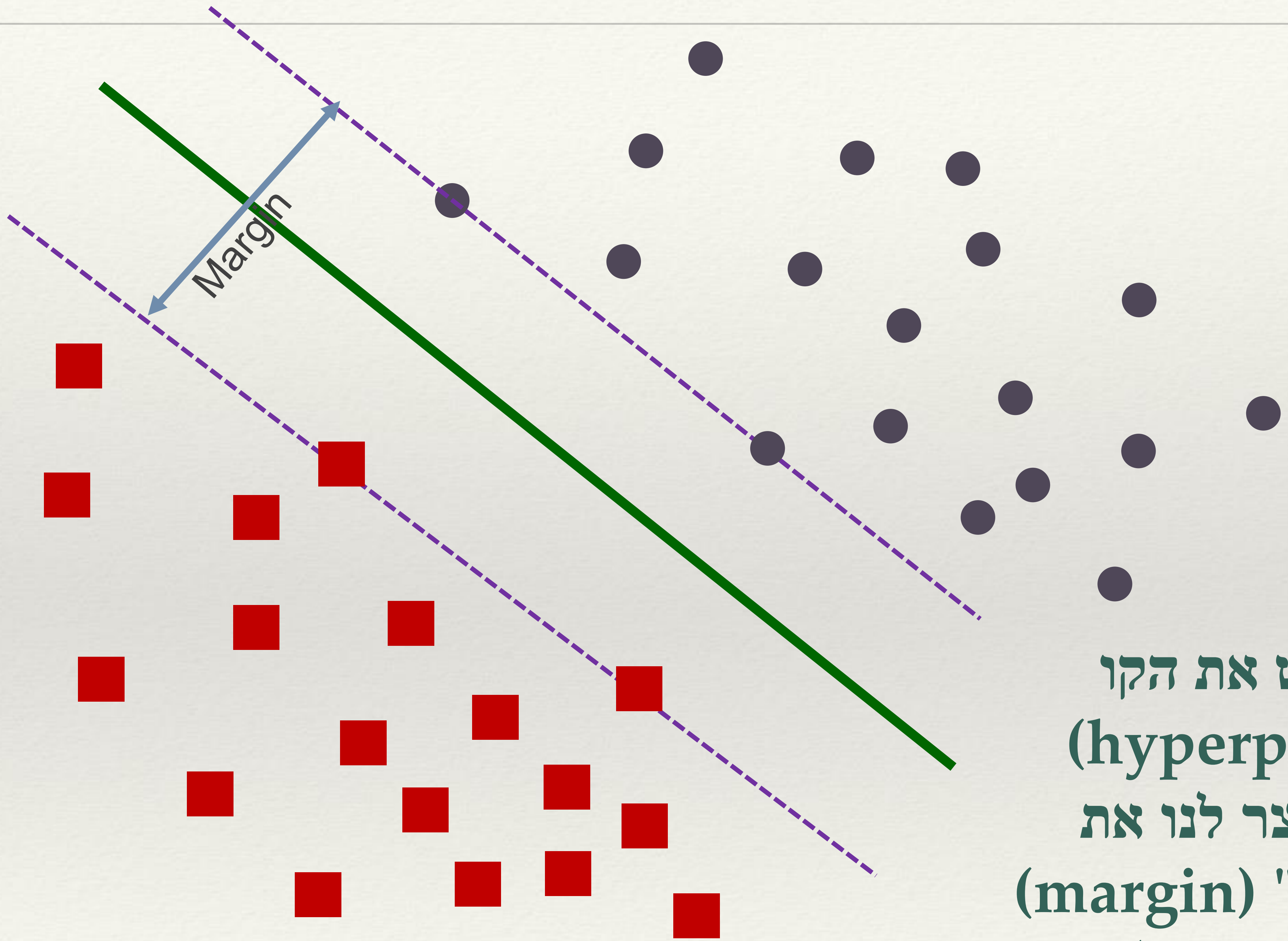
אפשרויות שונות של margin



אפשרויות שונות של margin

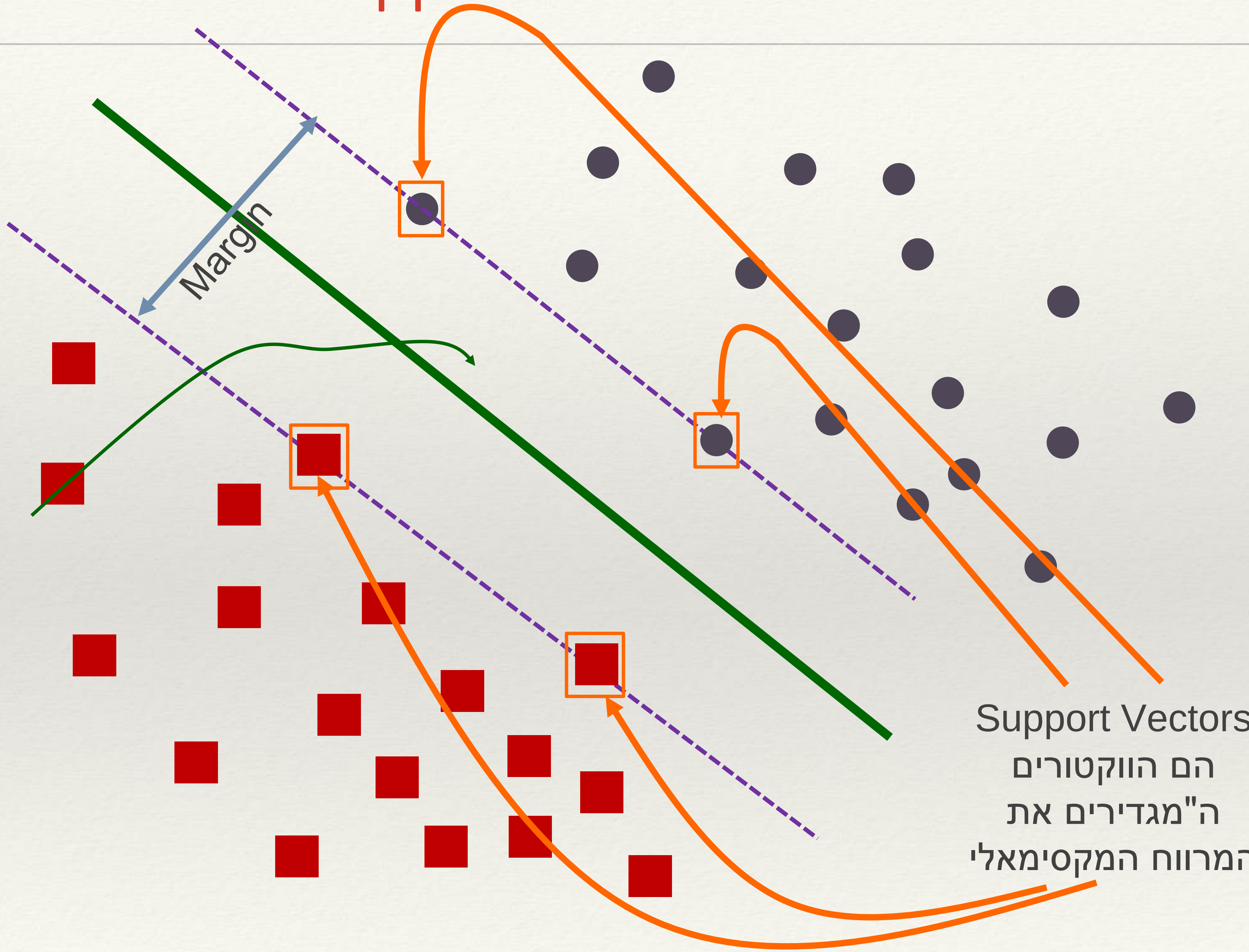


אפשרויות שונות של margin



נחפש את הקו (hyperplane) שמייצר לנו את ה"מרווח" (margin) המקסימאלי

SVM – Support Vector Machines –
מהם ה-Support Vectors?

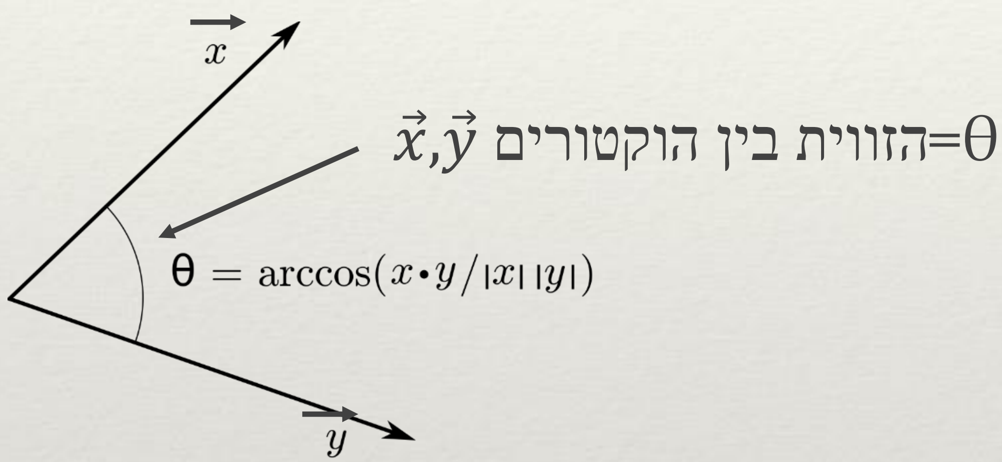


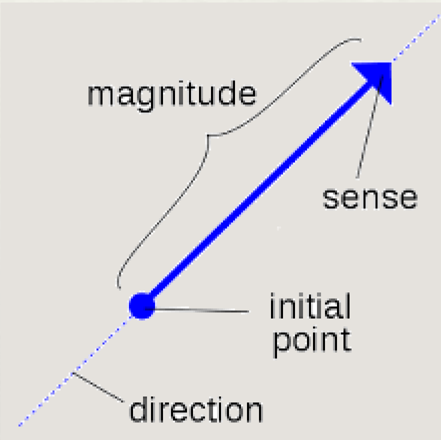
מושגים מתמטיים -
וקטורים וסקלרים

מושגים – וקטור (פרספקטיבה גאומטרית) - תזכורת

<u>וקטור (vector)</u> - נסמן $\vec{x},\vec{y} \in \mathbb{R}^n$

- <u>הסבר פיזיקלי/גאומטרי</u>: ישות מתמטית בעלת גודל וכיוון





<u>נורמה של וקטור</u> - הכללה של מושג ה"אורך" (magnitude) – נסמן $||\vec{x}||$

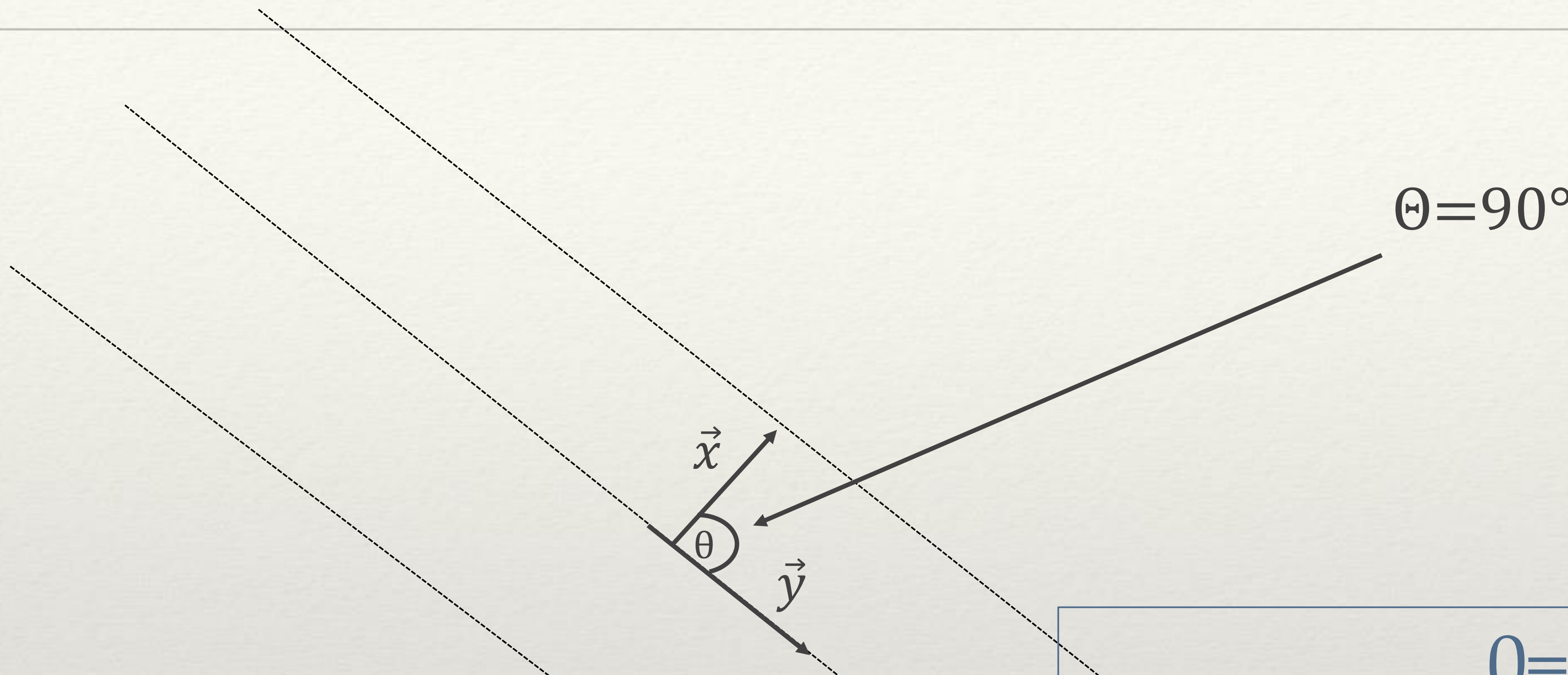
<u>מכפלה סקלרית (dot product) של וקטורים</u> – נסמן $\vec{x} \cdot \vec{y}$

מבחינה גאומטרית $\vec{x} \cdot \vec{y} = ||\vec{x}|| \cdot ||\vec{y}|| \cdot \cos\theta$

- לכן, מכפלה סקלרית של שני וקטורים היא 0 או"א הם ניצבים, כיוון ש-$\cos 90 = 0$

כפל הוקטורים $\vec{x} \cdot \vec{y}$ - למה שווה המכפלה

<u>סקר</u> – א. 0; ב. אי אפשר לדעת; ג. 1



תשובה א. =0
מכיוון, שמכפלה סקלרית של שני וקטורים היא 0 או"א הם ניצבים, כיוון ש-cos 90
$= 0$ <--- $\vec{x},\vec{y}$ ניצבים (אורטוגונלים)

$$\vec{x} \cdot \vec{y} = \|\vec{x}\| \cdot \|\vec{y}\| \cdot \cos\theta$$

מושגים – וקטור (פרספקטיבה אלגברית) - תזכורת

וקטור (vector) - נסמן $\vec{x},\vec{y} \in \mathbb{R}^n$

- מבחינה אלגברית: $\vec{x} = (x_1,x_2, \dots ,x_n),\vec{y} = (y_1,y_2, \dots ,y_n)$

כפל (של וקטור) בסקלר - $\vec{x} \cdot \lambda = \lambda \cdot \vec{x} = (\lambda \cdot x_1,\lambda \cdot x_2, \dots ,\lambda \cdot x_n)$

וקטור כמטריצה:

- וקטור שורה הוא מטריצה בגודל 1xn $\vec{x} = [x_1,x_2, \dots ,x_n]$
- המטריצה המשוחלפת (transpose) של וקטור שורה היא וקטור עמודה nx1

$$\vec{x}^T = [x_1, x_2, \dots , x_n]^T = \begin{bmatrix} x_1 \\ x_2 \\ \dots \\ x_n \end{bmatrix}$$

תרגיל – כפל וקטור בסקלר

חשבו את וקטור w (לפי הנוסחה שמופיעה למטה)

$\alpha_{1\ldots4}$ - הינם סקלרים, $y_{1\ldots4}$ - הינם סקלרים

$x_{1\ldots4}$ - הינם וקטורים

$$X1=(1,1,\sqrt{2},1,-\sqrt{2},-\sqrt{2}),\quad y_1=-1,\ \alpha_1=\frac{1}{8}$$

$$X2=(1,1,-\sqrt{2},1,-\sqrt{2},\sqrt{2}),\quad y_2=+1,\ \alpha_2=\frac{1}{8}$$

$$X3=(1,1,-\sqrt{2},1,\sqrt{2},-\sqrt{2}),\quad y_3=+1,\ \alpha_3=\frac{1}{8}$$

$$X4=(1,1,\sqrt{2},1,\sqrt{2},\sqrt{2}),\quad y_4=-1,\ \alpha_4=\frac{1}{8}$$

$$\mathrm{w}=\sum_{\mathrm{i}}\alpha_i y_i x_i$$

$$\mathrm{w}=\sum_{\mathrm{i}}\alpha_i y_i x_i$$

$$\frac{1}{8}\times-1\times(1,\ 1,\ \sqrt{2},\ 1,\ -\sqrt{2},\ -\sqrt{2})+$$

$$\frac{1}{8}\times+1\times(1,\ 1,-\sqrt{2},1,\ -\sqrt{2},\quad \sqrt{2})+$$

$$\frac{1}{8}\times+1\times(1,\ 1,-\sqrt{2},1,\quad \sqrt{2},\ -\sqrt{2})+$$

$$\frac{1}{8}\times-1\times(1,\ 1,\ \sqrt{2},\ 1,\quad \sqrt{2},\quad \sqrt{2})=$$

$$\boxed{w=(0,\ 0,\ -\frac{\sqrt{2}}{2},\ 0,\quad 0,\quad 0)}$$

מושגים – וקטור (פרספקטיבה אלגברית)

<u>וקטור (vector)</u> - נסמן $\vec{x},\vec{y} \in \mathbb{R}^n$

<u>מכפלה סקלרית (dot product) של וקטורים</u> – נסמן $\vec{x} \cdot \vec{y}$

<u>מבחינה אלגברית</u> - $\langle \vec{x},\vec{y} \rangle = \sum_{i=1}^{n} x_i \cdot y_i = x_1 \cdot y_1 + x_2 \cdot y_2 + \cdots + x_n \cdot y_n$

- בצורה מטריציונית - $\langle \vec{x},\vec{y} \rangle = \vec{x} \cdot \vec{y}^T = \sum_{i=1}^{n} x_i \cdot y_i$

<u>תכונות</u> –

- חוק הפילוג - $\langle \vec{x},\vec{y} + \vec{z} \rangle = \langle \vec{x},\vec{y} \rangle + \langle \vec{x},\vec{z} \rangle = \vec{x} \cdot (\vec{y} + \vec{z})^T = \vec{x} \cdot \vec{y}^T + \vec{x} \cdot \vec{z}^T$
- חוק החילוף - $\langle \vec{x},\vec{y} \rangle = \langle \vec{y},\vec{x} \rangle = \vec{x} \cdot \vec{y}^T = \vec{y} \cdot \vec{x}^T$

תרגיל כפל וקטורי

נתונים הוקטורים $\vec{x} = (1,1,1,1,1,1), \vec{w} = (0,0,\frac{-\sqrt{2}}{2},0,0,0)$

חשבו את הכפל הוקטורי: $\vec{w}^T \cdot \vec{x}$

<u>תשובה</u>:

$$0 \cdot 1 + 0 \cdot 1 + \frac{-\sqrt{2}}{2} \cdot 1 + 0 \cdot 1 + 0 \cdot 1 + 0 \cdot 1 = \frac{-\sqrt{2}}{2}$$

מושגים – וקטור (פרספקטיבה אלגברית)

נורמה של וקטור - הכללה של מושג ה"אורך" (magnitude) – נסמן $\|\vec{x}\|$

- נורמת L_p – דומה למרחק מיניקובסקי $L_p = \|\vec{x}\|_p = \left(\sum_{i=1}^{n}|x_i|^p\right)^{\frac{1}{p}}$
- נורמה סטנדרטית (אוקלדית) - נציב p=2

$$L_2 = \|\vec{x}\|_2 = \left(\sum_{i=1}^{n}|x_i|^2\right)^{\frac{1}{2}} = \sqrt{{x_1}^2 + {x_2}^2 + \cdots + {x_n}^2}$$
$$= \sqrt{x_1 \cdot x_1 + x_2 \cdot x_2 + \cdots + x_n \cdot x_n} = \sqrt{\vec{x} \cdot \vec{x}^T}$$
$$= \sqrt{\langle \vec{x}, \vec{x} \rangle}$$

תכונות הנורמה (הסטדנדרטית):

- אי שלילי - $\|\vec{x}\| \geq 0$
- הכפלה בסקלר - מכפילה גם התוצאה בסקלר - $\|\lambda \cdot \vec{x}\| = \lambda \cdot \|\vec{x}\|$
- מתקיים אי שוויון המשולש - $\|\vec{x} + \vec{y}\| \leq \|\vec{x}\| + \|\vec{y}\|$

לדו'

x=(1,1,1,1,1,1)
x2=2*x=(2,2,2,2,2,2)

נראה ש:
||x2||=2*||x||

||x||=
sqrt(1*1+1*1+1*1+1*1+1*1+
1*1)=sqrt(6)

||x2||=
sqrt(2*2+2*2+2*2+2*2+2*2+
2*2)=
sqrt(4*(1*1+1*1+1*1+1*1+1
*1+1*1))=
sqrt(4)*sqrt(6)=
2*sqrt(6)=2||x||

תרגיל חישוב נורמה

<u>תרגיל נורמה 1</u>:

נתון הווקטור: $\vec{x}=(2,1)$, חשבו את $||\vec{x}||$ כנורמה הסטנדרטית:

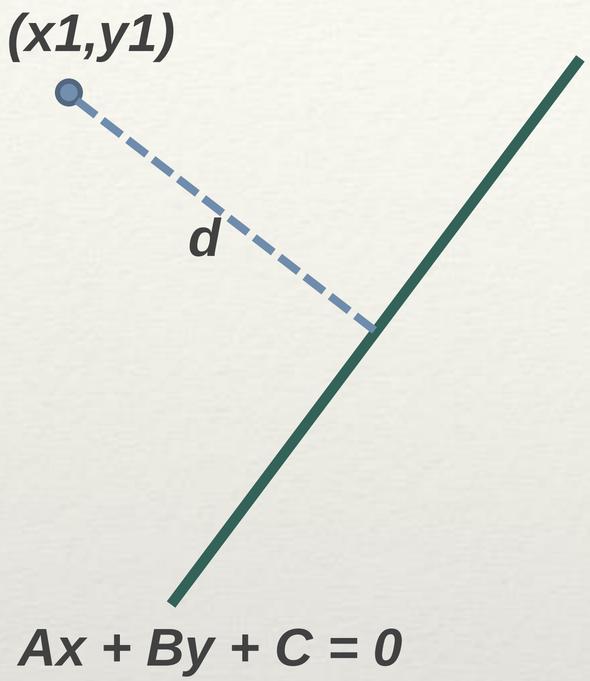
<u>תשובה</u>: $\sqrt{2^2+1^2}=\sqrt{5}$

<u>תרגיל נורמה 2</u>:

נתון הווקטור: $\vec{w}=\left(0,0,\frac{-\sqrt{2}}{2},0,0,0\right)$, חשבו את $||\vec{w}||$ כנורמה הסטנדרטית:

<u>תשובה</u>: $\sqrt{0^2+0^2+(\frac{-\sqrt{2}}{2})^2+0^2+0^2+0^2}=\frac{1}{\sqrt{2}}$

מרחק נקודה מישר



- הישר $Ax + By + C = 0$

$$d = \frac{|Ax_1 + By_1 + C|}{\sqrt{A^2 + B^2}}$$

- במקרה הכללי – משוואת ה-hyperplane

$$\sum w_i x_i + b = 0$$

מרחק הנקודה x_s מה-hyperplane

$$d = \frac{|wx_s + b|}{||w||}$$

תרגיל חישוב מרחק נקודה מhyperplane

נתונה משוואת הישר:

$\vec{w}^T \cdot \vec{x}$=0

נתון הוקטור $\vec{w}$=(2,1)

נתונה הנקודה p=(4,2),

מהו מרחק הנקודה p

מהישר הנ"ל.

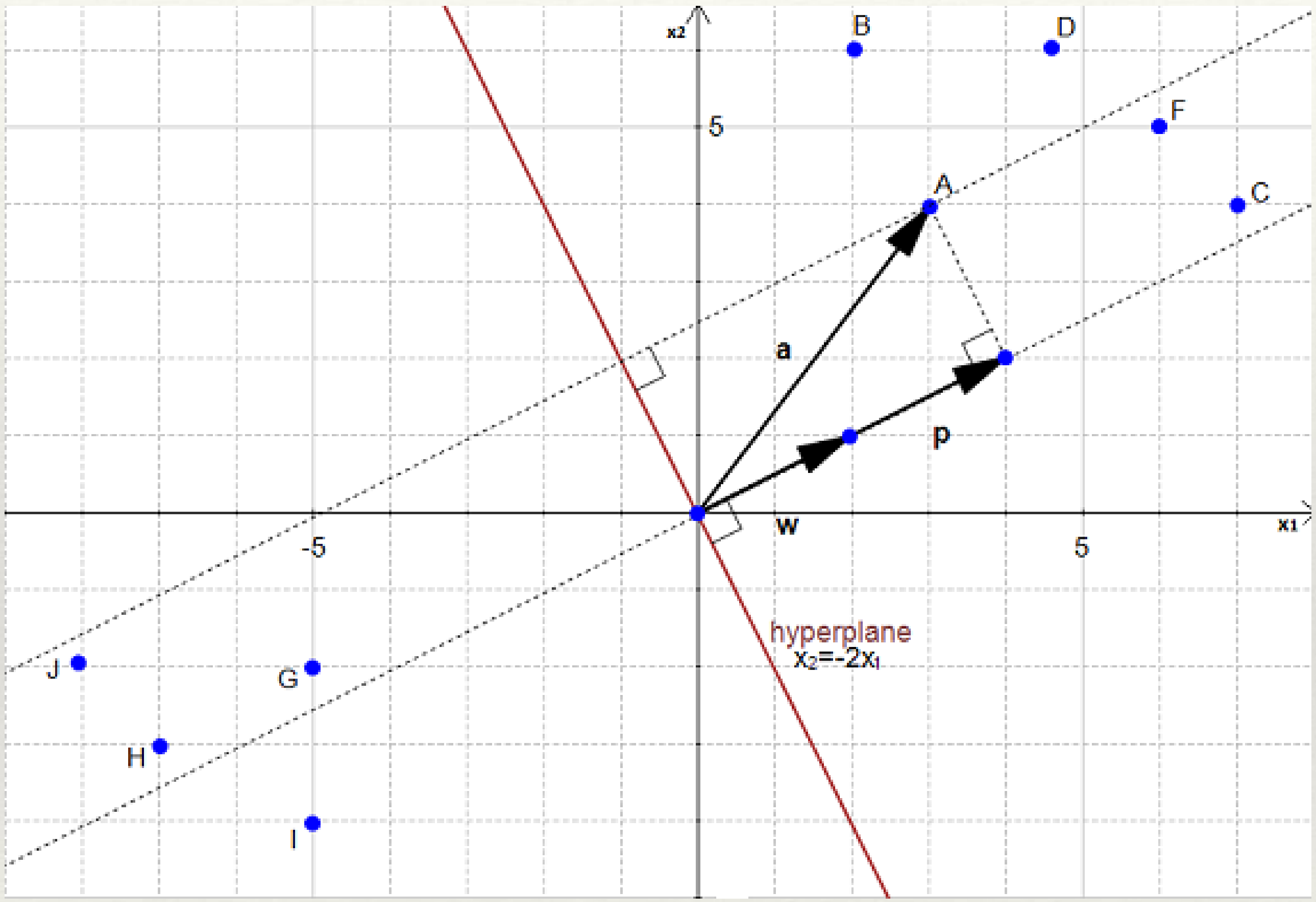
<u>תשובה</u>: נשתמש בנוסחה:

$$\frac{|\vec{w}^T \cdot \vec{p} + b|}{||\vec{w}||}$$

כבר חישבנו $||\vec{w}||=\sqrt{5}$

$$\frac{|2 \cdot 4 + 1 \cdot 2 + 0|}{\sqrt{5}}$$

$$= \frac{10}{\sqrt{5}} = 2\sqrt{5}$$



ניתן לראות גם בעיין ש-p רחוק פי 2 מw

איך מחשבים את ה-margin

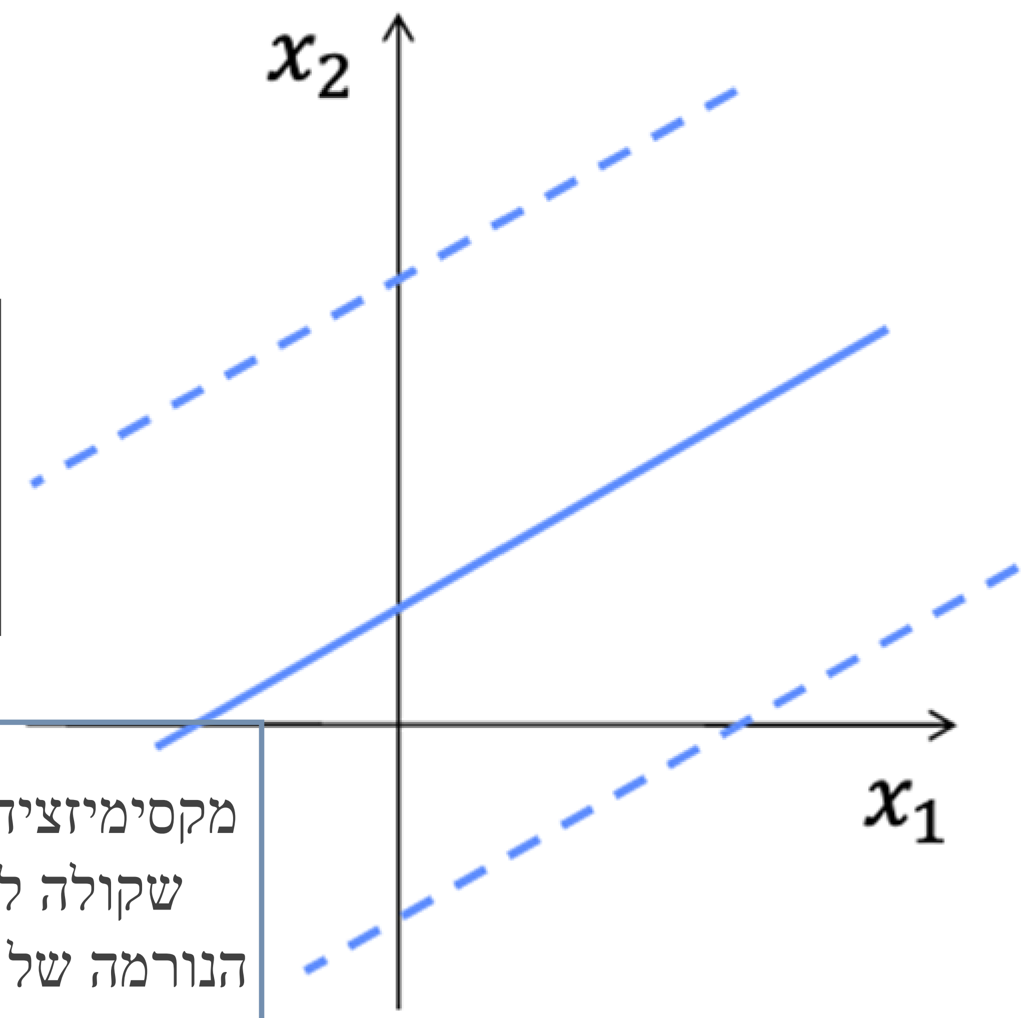
Point-plain distances from the two margins to the origin:

$$d_+ = \frac{|(w\cdot 0)+b+1|}{\|w\|}, d_- = \frac{|(w\cdot 0)+b-1|}{\|w\|}$$

מודדים את המרחק מהקו העליו לראשית הצירים ובאופן דומה לגבי הקו התחתון

$$\Rightarrow \mathrm{M} = \frac{2}{\|w\|}$$

חישוב ה-margin



Objective:

מקסימיזציה של ה-margin שקולה למינימיזציה של הנורמה של וקטור המשקולות

$$max(\mathrm{M}) \rightarrow min(\|w\|^2)$$

$s.t.$

קבלת החלטה – Linear SVM

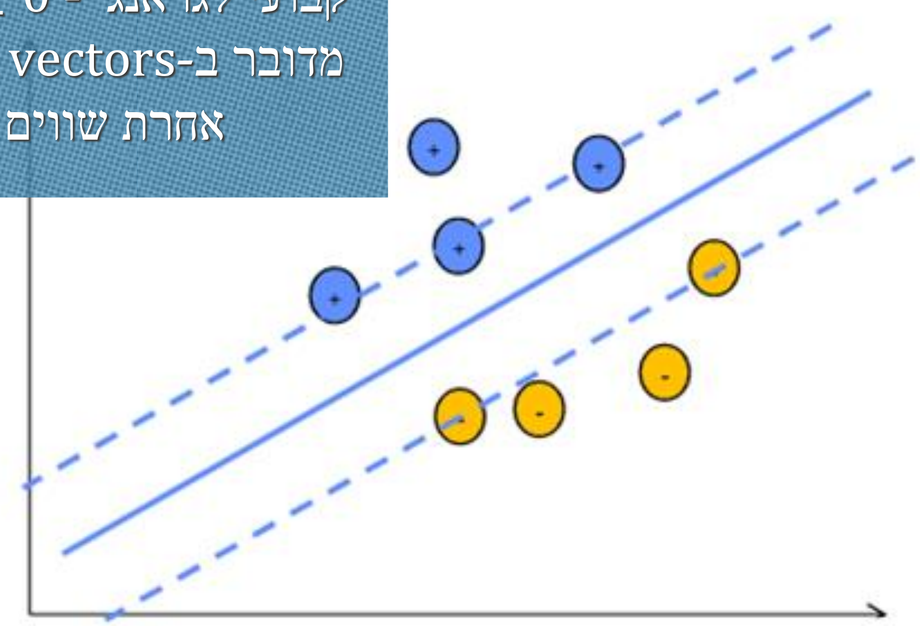
חישוב w

Calculating w, b:

$$w = \sum_i \alpha_i y_i x_i$$

קבועי לגראנג' - $\alpha_i \geq 0$ – אם מדובר ב-support vectors, אחרת שווים ל-0

$$y_i (w \cdot x_i + b) - 1 \geq 0 \ \forall i$$



The average of the nearest positive support vector and the nearest negative

$$b = -\frac{max_{y_i=-1}(\mathbf{w} \cdot \mathbf{x}_i) + min_{y_i=1}(\mathbf{w} \cdot \mathbf{x}_i)}{2}$$

חישוב b, של הקו האמצעי

Final decision function:

$$f(x) = sign(w \cdot x + b) = sign\left(\sum_{i=1}^{l} \alpha_i y_i x_i \cdot x + b \right)$$

החלטת סיווג

תרגיל סיווג SVM 1

נתונים:

- הטבלה הבאה עם נתוני הנקודות.
- קבועי לגראנז' (שכבר עשו עבורים אופטימיזציה)

Table 1: Training Data

x_1	x_2	y_i	α_i
0.38	0.47	+	65.52
0.49	0.61	-	65.52
0.92	0.41	-	0
0.74	0.89	-	0
0.18	0.58	+	0
0.41	0.35	+	0
0.93	0.81	-	0
0.21	0.10	+	0

חשבו עבורן SVM ליניארי, וסווג את הנקודה (0.5 ,0.5)

פתרון (תרגיל סיווג SVM 1)

- אנחנו יכולים לראות שיש 2 SV
- נשתמש במקדמי הלגנרז כדי לחשב את w
- איך?

Table 1: Training Data

x_1	x_2	y_i	α_i
0.38	0.47	+	65.52
0.49	0.61	-	65.52
0.92	0.41	-	0
0.74	0.89	-	0
0.18	0.58	+	0
0.41	0.35	+	0
0.93	0.81	-	0
0.21	0.10	+	0

פתרון (תרגיל סיווג SVM 1)

w=65.52(1*(0.38,0.47)+-1*(0.49,0.61))=
65.52*(-0.11,-0.14)≈(-7.2,-9.17)

$$\mathrm{w} = \sum_{\mathrm{i}} \alpha_i y_i x_i$$

Table 1: Training Data

x_1	x_2	y_i	α_i
0.38	0.47	+	65.52
0.49	0.61	-	65.52
0.92	0.41	-	0
0.74	0.89	-	0
0.18	0.58	+	0
0.41	0.35	+	0
0.93	0.81	-	0
0.21	0.10	+	0

$$w_1 = \sum_i \alpha_i \, .y_i.x_{i1} = 65.52 \times 1 \times 0.38 + 65.52 \times -1 \times 0.49 \approx -7.2$$

$$w_2 = \sum_i \alpha_i \, .y_i.x_{i2} = 65.52 \times 1 \times 0.47 + 65.52 \times -1 \times 0.61 \approx -9.17$$

- ועכשיו נחשב את b
- The average of the nearest positive support vector and the nearest negative support vector

$$b = -\frac{max_{y_i=-1}(\mathbf{w} \cdot \mathbf{x}_i) + min_{y_i=1}(\mathbf{w} \cdot \mathbf{x}_i)}{2}$$

פתרון (תרגיל סיווג SVM 1)

- נתון:

x_1	x_2	y_i	α_i
0.38	0.47	+	65.52
0.49	0.61	-	65.52

- וקיבלנו שערכי המשקולות:

$$w_1 \approx -7.2$$
$$w_2 \approx -9.17$$

- ועכשיו נמצא את הערכים של b עבור הדוגמאות החיוביות והשליליות

$$b = -\frac{max_{y_i=-1}(\mathbf{w}\cdot\mathbf{x}_i) + min_{y_i=1}(\mathbf{w}\cdot\mathbf{x}_i)}{2}$$

-(((-7.2)*0.49+(-9.17)*0.61)+((-7.2)*0.38+(-9.17)*0.47))/2 ≈8.084

לכן ה hyperplane המפריד המקסימלי הוא:

$$-7.2\cdot x_1 - 9.17\cdot x_2 + 8.084 = 0$$

פתרון (תרגיל סיווג SVM 1)

נניח דוגמא חדשה (0.5,0.5). נציב בנוסחה שמצאנו

$$-7.2 \cdot x_1 - 9.17 \cdot x_2 + 8.084 = 0$$

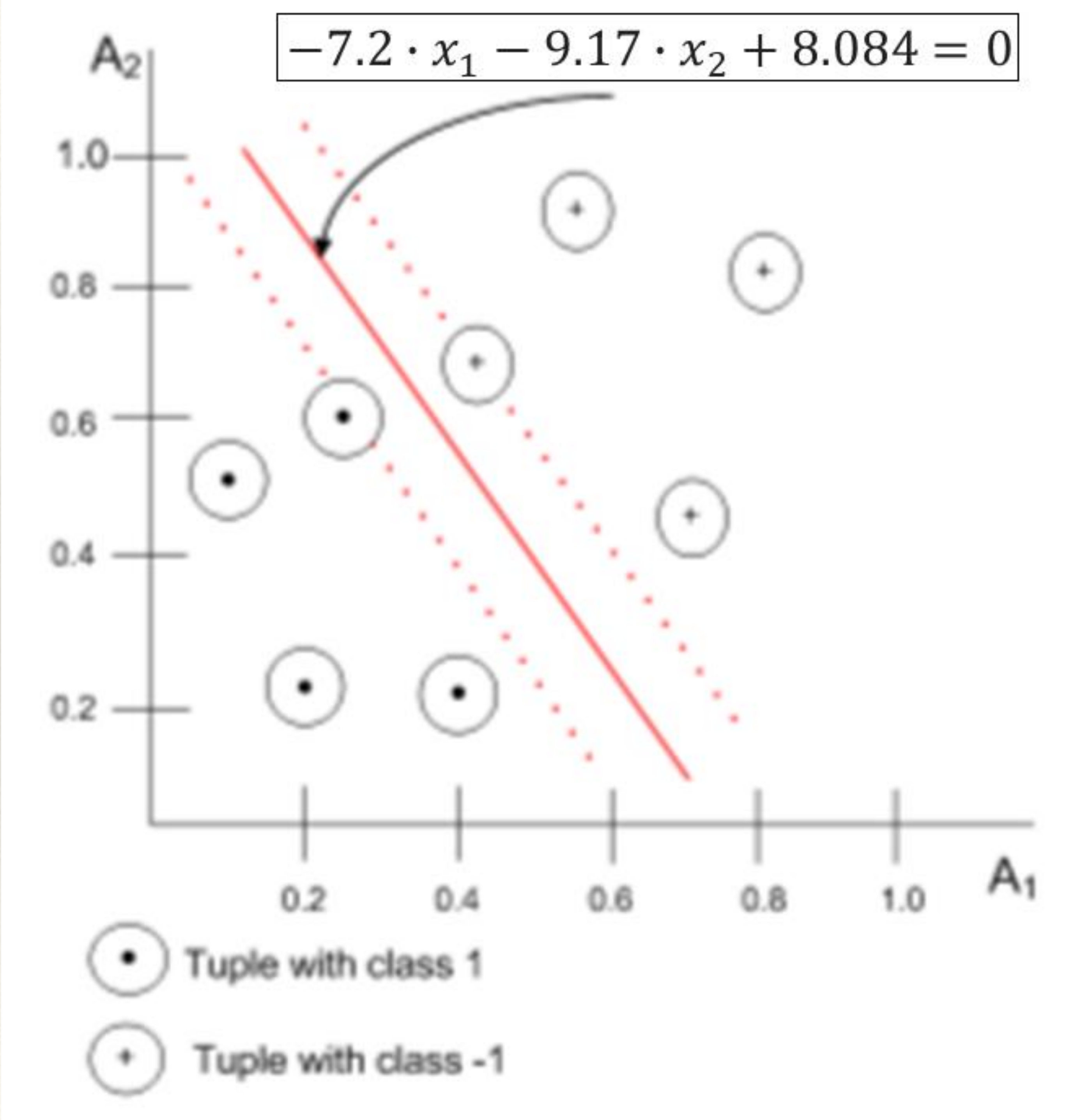
- נקבל

$$-7.2 \cdot 0.5 - 9.17 \cdot 0.5 + 8.084 = \text{-}0.101$$

ולכן הדוגמא מסווגת כשייכת לקטגוריה השלילית..

פתרון (תרגיל סיווג SVM 1)

- תצוגה של התוצר



תרגיל סיווג SVM 2

- נתונים ארבעת הווקטורים הבאים והסיווג שלהם:

$$X1 = (1,1,\sqrt{2},1,-\sqrt{2},-\sqrt{2}) \qquad y1 = -1$$

$$X2 = (1,1,-\sqrt{2},1,-\sqrt{2},\sqrt{2}) \qquad y2 = +1$$

$$X3 = (1,1,-\sqrt{2},1,\sqrt{2},-\sqrt{2}) \qquad y3 = +1$$

$$X4 = (1,1,\sqrt{2},1,\sqrt{2},\sqrt{2}) \qquad y4 = -1$$

- נניח שפתרו עבורינו את בעיית האופטימיזציה של SVM ו"גילו" לנו ש:

$$\alpha_1 = \frac{1}{8}, \alpha_2 = \frac{1}{8}, \alpha_3 = \frac{1}{8}, \alpha_4 = \frac{1}{8}$$

$$b = 0$$

תרגיל סיווג SVM 2 (המשך)

- כיצד נסווג את הווקטור $x_1^{test} = (1,1,1,1,1,1)$

- כיצד נסווג את הווקטור $x_2^{test} = (1,1,\sqrt{2},1,-\sqrt{2},-\sqrt{2})$

פתרון (תרגיל סיווג SVM 2)

- זה שכל האלפות שונות מ-0, אומר לנו שכל הווקטורים הם SV
- סיווג דוגמא חדשה x, מקרה ליניארי פשוט, ללא קרנל נחשב ראשית את ה-W
- $W = \sum_{i=1}^{n} y_i \alpha_i x_i$
- נחשב את b במידת הצורך
- ונסווג: $\text{class}(\mathrm{x}) = \mathrm{SGN}(Wx - b)$

פתרון (תרגיל סיווג SVM 2) – נתחיל בחישוב w – נראה מוכר?

- כלומר, נחשב ראשית את w:

$$\mathrm{w} = \sum_{\mathrm{i}} \alpha_i y_i x_i$$

$$\frac{1}{8} \times -1 \times (1,\ 1,\ \sqrt{2},\ 1,\ -\sqrt{2},\ -\sqrt{2}) +$$
$$\frac{1}{8} \times +1 \times (1,\ 1, -\sqrt{2}, 1,\ -\sqrt{2},\ \ \sqrt{2}) +$$
$$\frac{1}{8} \times +1 \times (1,\ 1, -\sqrt{2}, 1,\ \ \sqrt{2},\ -\sqrt{2}) +$$
$$\frac{1}{8} \times -1 \times (1,\ 1,\ \sqrt{2},\ 1,\ \ \sqrt{2},\ \ \sqrt{2}) =$$

$$\boxed{w = (0,\ 0,\ -\frac{\sqrt{2}}{2},\ 0,\ \ 0,\ \ 0)}$$

$$X1 = (1,1,\sqrt{2},1,-\sqrt{2},-\sqrt{2}),\quad \mathrm{y}_1 = -1, \alpha_1 = \frac{1}{8}$$
$$X2 = (1,1,-\sqrt{2},1,-\sqrt{2},\sqrt{2}),\quad \mathrm{y}_2 = +1, \alpha_2 = \frac{1}{8}$$
$$X3 = (1,1,-\sqrt{2},1,\sqrt{2},-\sqrt{2}),\quad \mathrm{y}_3 = +1, \alpha_3 = \frac{1}{8}$$
$$X4 = (1,1,\sqrt{2},1,\sqrt{2},\sqrt{2}),\quad \mathrm{y}_4 = -1, \alpha_4 = \frac{1}{8}$$
$$\mathrm{w} = \sum_{\mathrm{i}} \alpha_i y_i x_i$$

פתרון (תרגיל סיווג 2 SVM) –
סיווג דוגמא חדשה - ליניארי

$$f(x) = \mathrm{sgn}(\sum_i \alpha_i y_i < x_i^T, x > + b) = \mathrm{sgn}(wx + b)$$

x_1^{test}=(1,1,1,1,1,1)

$$f(x) = \mathrm{sgn}(w^T x + b) = \mathrm{sgn}((0,\ 0,\ \frac{-\sqrt{2}}{2},\ 0,\ \ 0,\ \ 0)^T (1,1,1,1,1,1) + b)$$

$$= \mathrm{sgn}(0\times 1 + 0\times 1 + (-\frac{\sqrt{2}}{2}\times 1) + 0\times 1 + 0\times 1 + 0\times 1 + b) =$$

$$= \mathrm{sgn}(-\frac{\sqrt{2}}{2}) = -1$$

נתון: b=0 ולכן -->

x_2^{test} =(1,1,$\sqrt{2}$, 1, $-\sqrt{2}$, $-\sqrt{2}$)

$$f(x) = \mathrm{sgn}(w^T x + b) = \mathrm{sgn}((0,\ 0,\ -\frac{\sqrt{2}}{2},\ 0,\ \ 0,\ \ 0)^T (1,1,\sqrt{2},1,-\sqrt{2},-\sqrt{2}) + b)$$

$$= \mathrm{sgn}(0\times 1 + 0\times 1 + (-\frac{\sqrt{2}}{2}\times\sqrt{2}) + 0\times 1 + 0\times -\sqrt{2} + +0\times -\sqrt{2} + b) =$$

$$= \mathrm{sgn}(-1) = -1$$

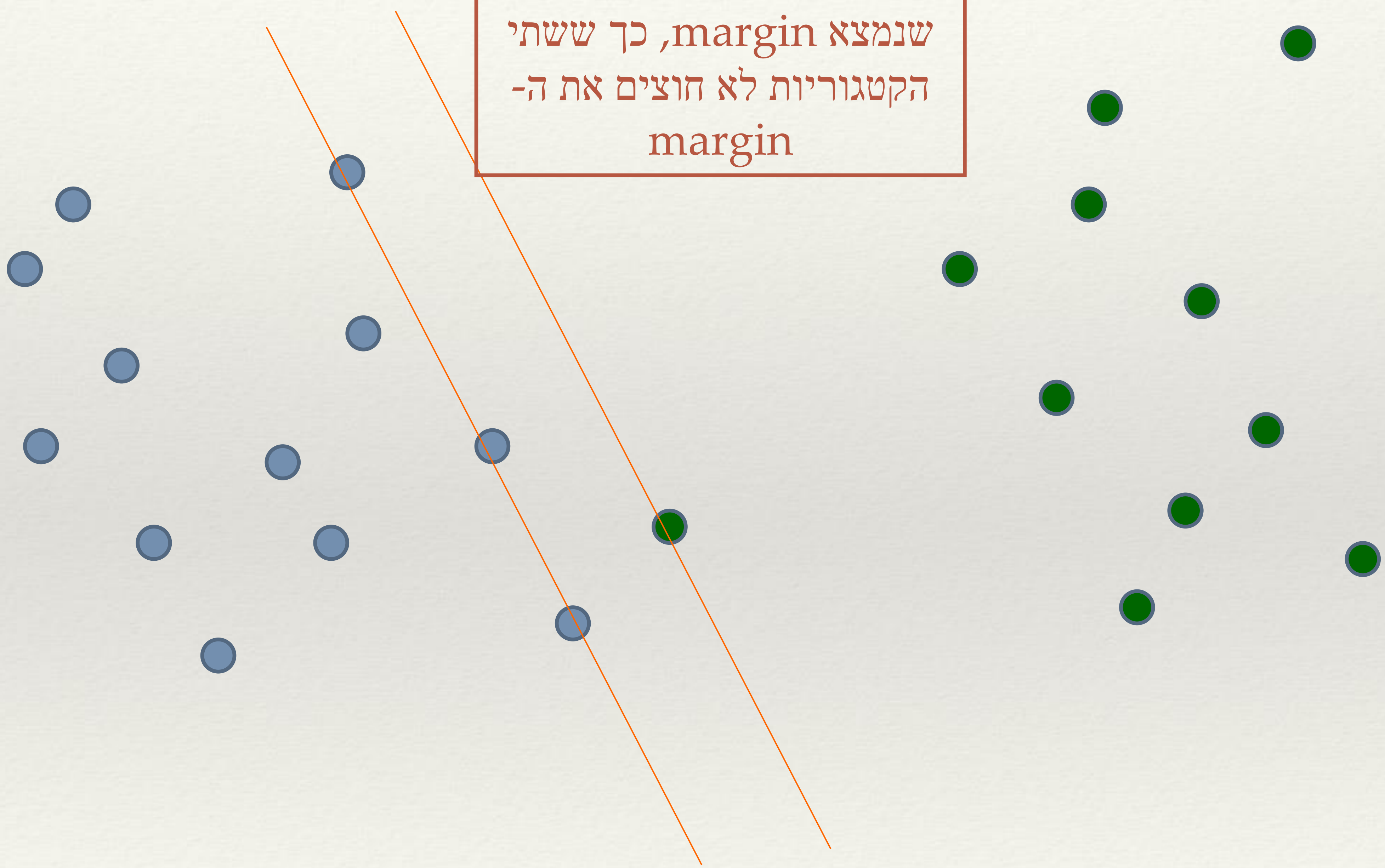
נתון: b=0 ולכן -->

כאן אין צורך לחשב. אנו יודעים שזהו SV , עם סיווג שלילי ולכן ברור גם שהצבה שלו במשוואה תיתן -1

hard margin לעומת
soft margin- הסבר

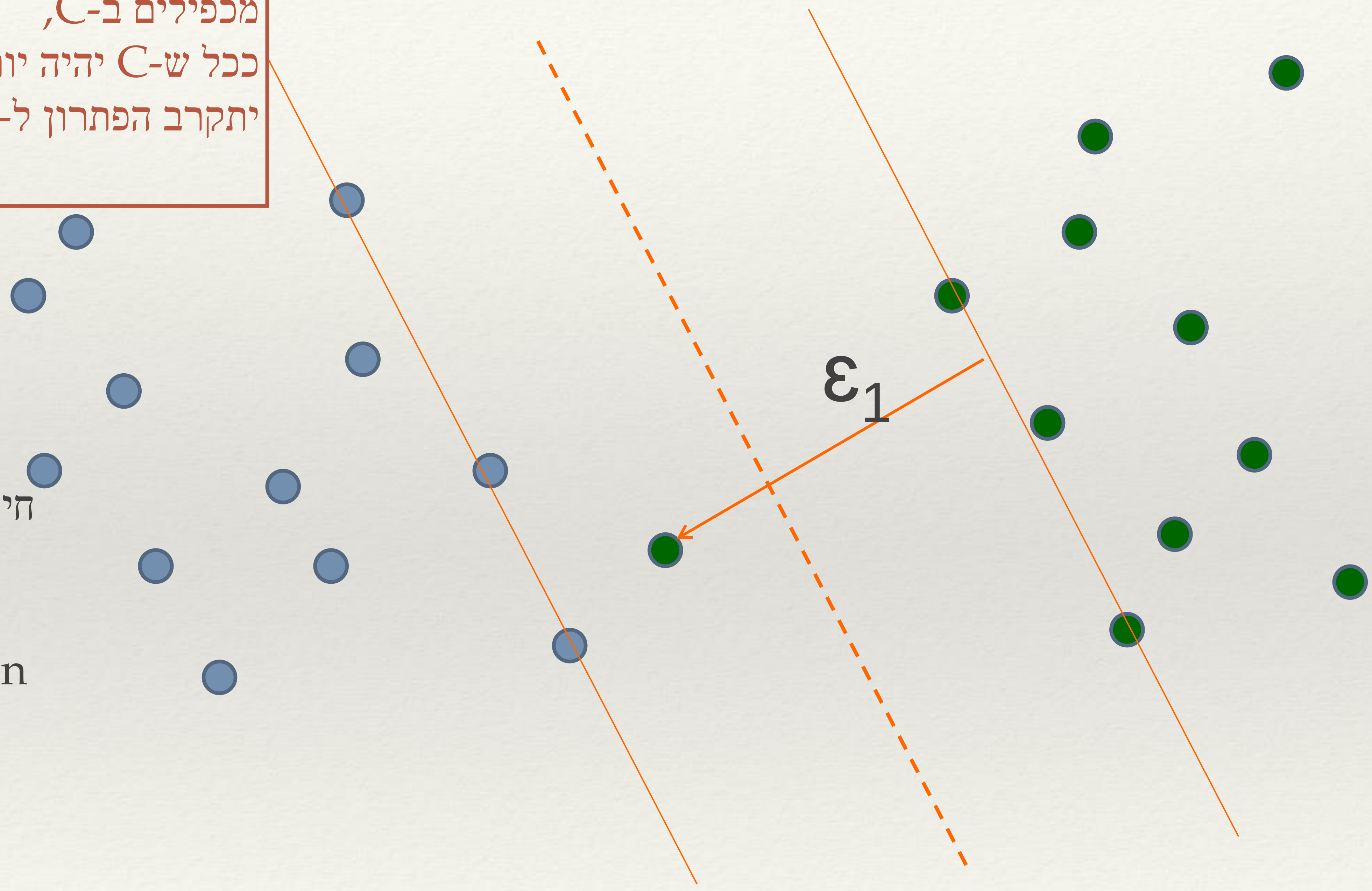
Hard Margin

הפתרון שלמדנו עד כה מחייב שנמצא margin, כך ששתי הקטגוריות לא חוצים את ה-margin



Soft Margin

<u>Soft margin</u> – מאפשר להעניש את הדוגמאות שחוצות את ה-margin בכיוון הלא נכון.
<u>בכמה נעניש</u>?
מכפילים ב-C,
ככל ש-C יהיה יותר גבוהה,
יתקרב הפתרון ל-hard margin



חישוב הטעות ε_j,
נמדדת ביחס
למרחק מה-
margin לכיוון
ה"שגוי"

SVM במקרים שאין הפרדה לינארית

תרגיל

- פתור את הבעיה הבאה באמצעות SVM עם קרנל לינארי.
- נתונה קבוצת האימון הבאה:

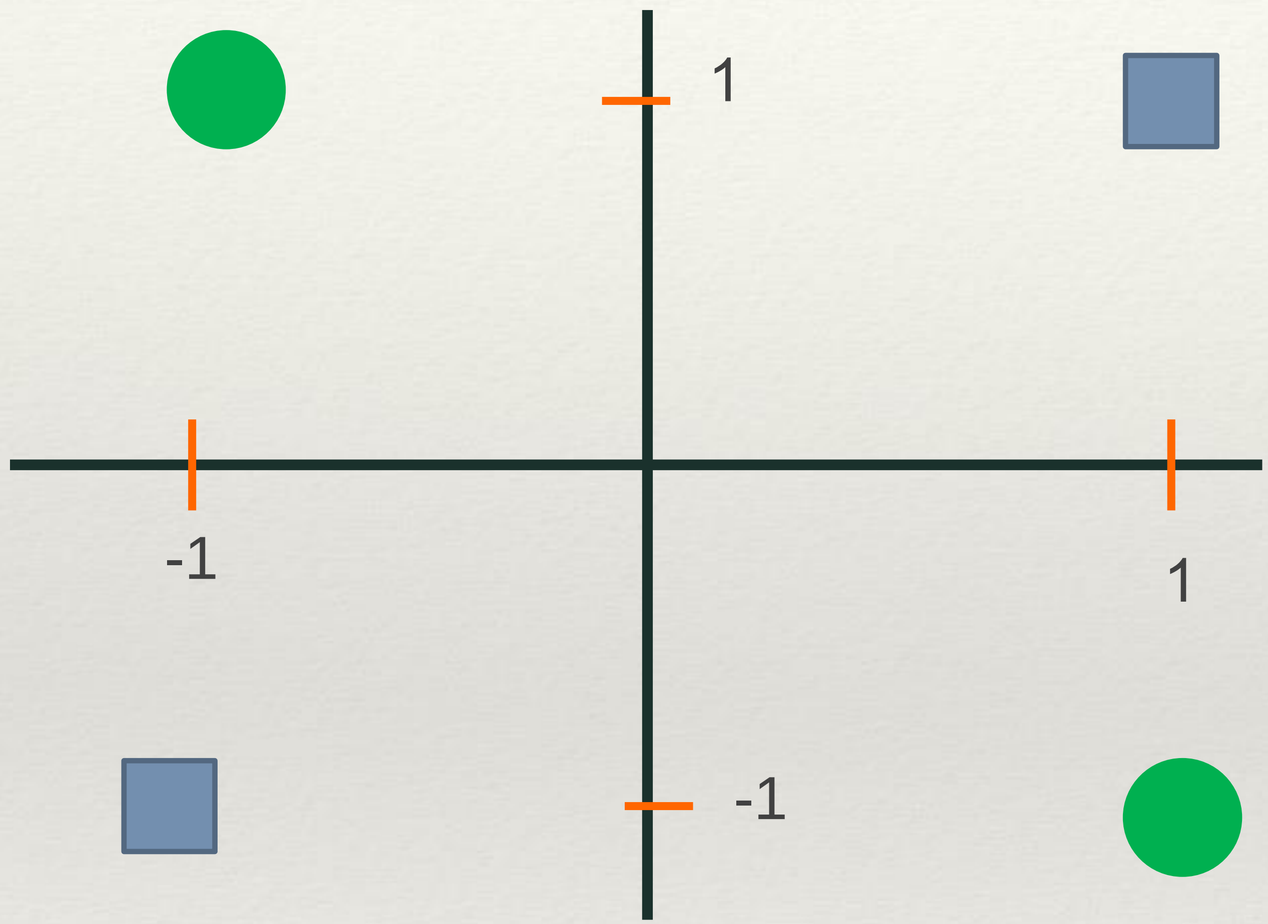
(+1,+1|-1)

(-1,+1|+1)

(+1,-1|+1)

(-1,-1|-1)

תרגיל



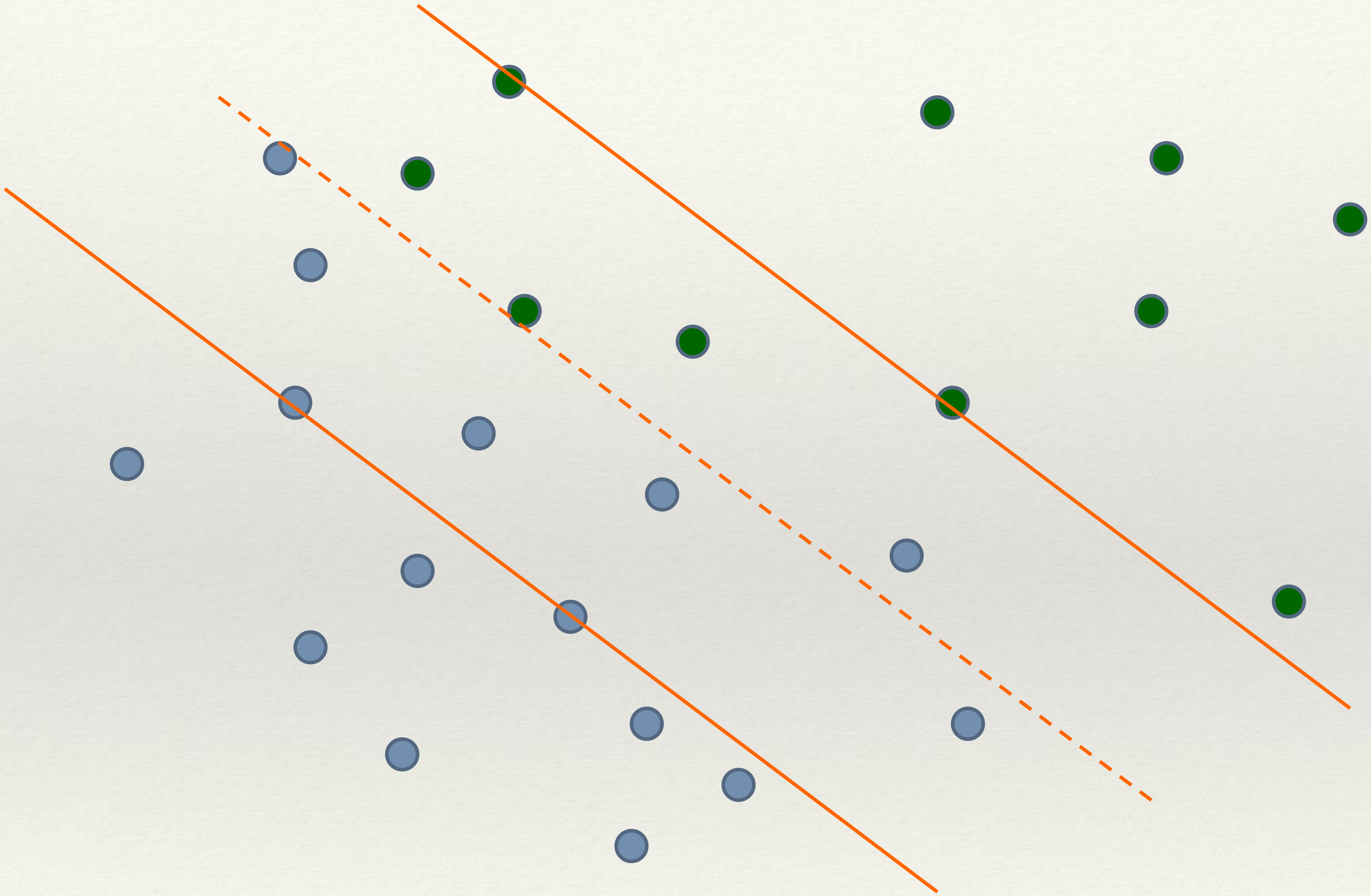
$X_1 = (+1, +1)$
$X_2 = (-1, +1)$
$X_3 = (+1, -1)$
$X_4 = (-1, -1)$

$y_1 = -1$
$y_2 = +1$
$y_3 = +1$
$y_4 = -1$

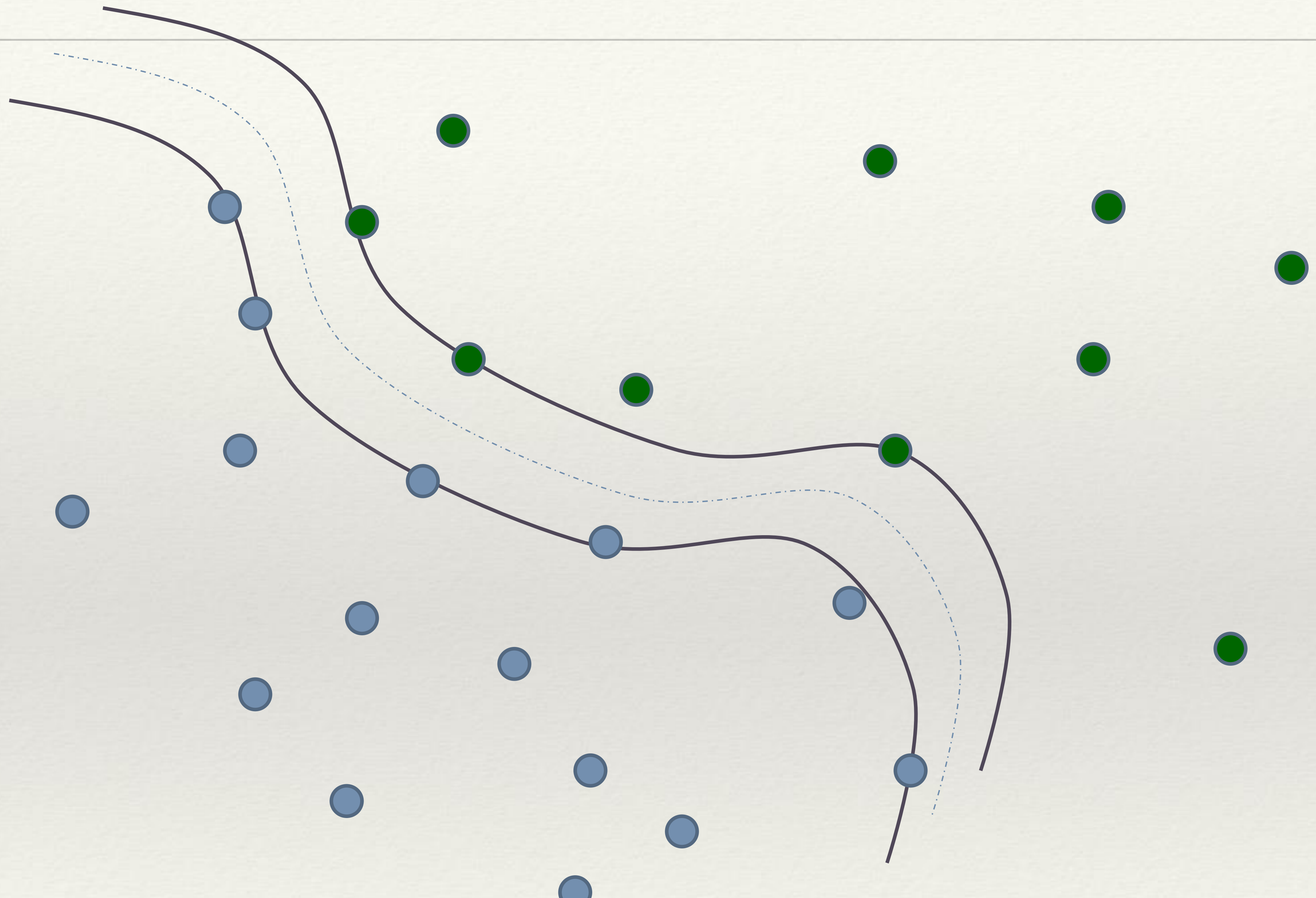
מזכיר את בעיית ה-XOR - אין פתרון ליניארי לבעיה במימד בו ניתנה

--> צריך למצוא פתרון לא ליניארי

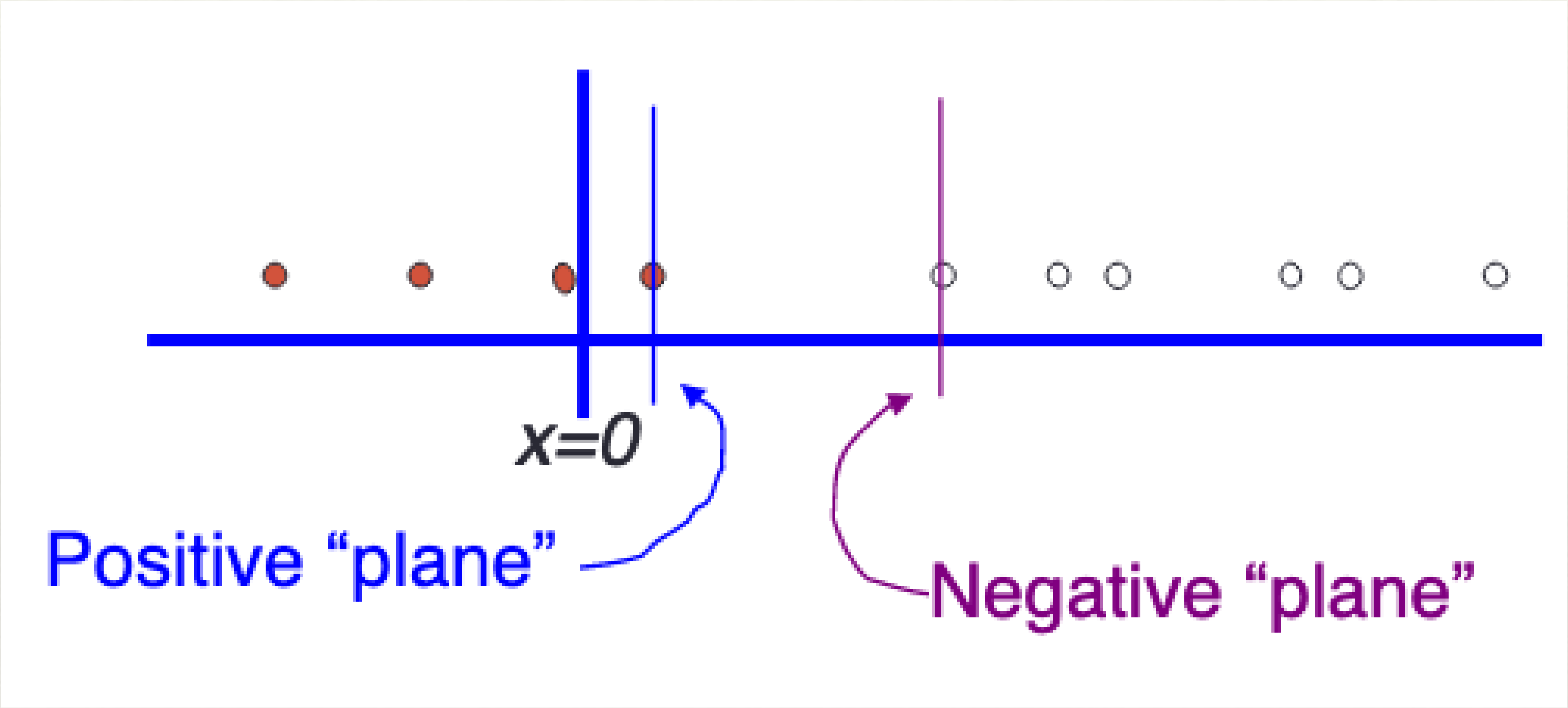
Soft Margin – לעיתים לא מתאים, מדוע?



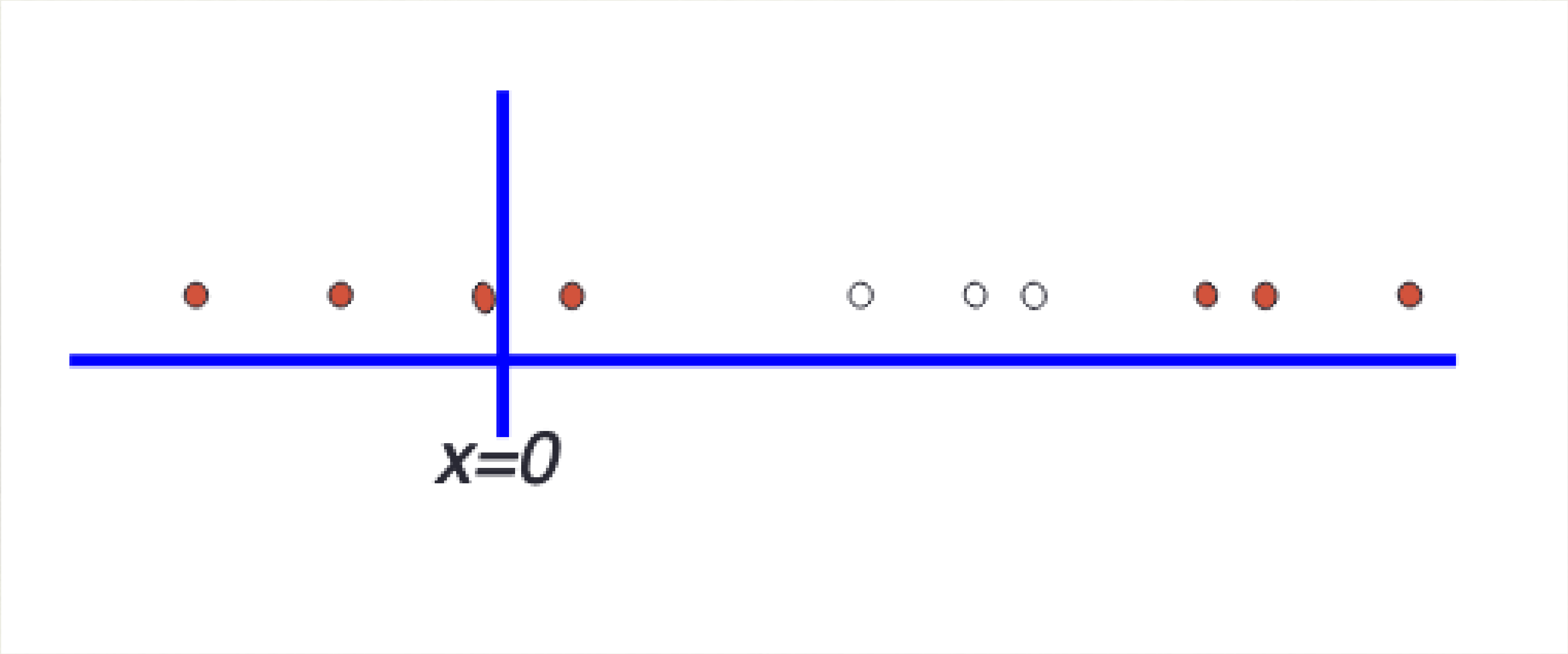
פתרון ע"י "על-מישור" לא לינארי



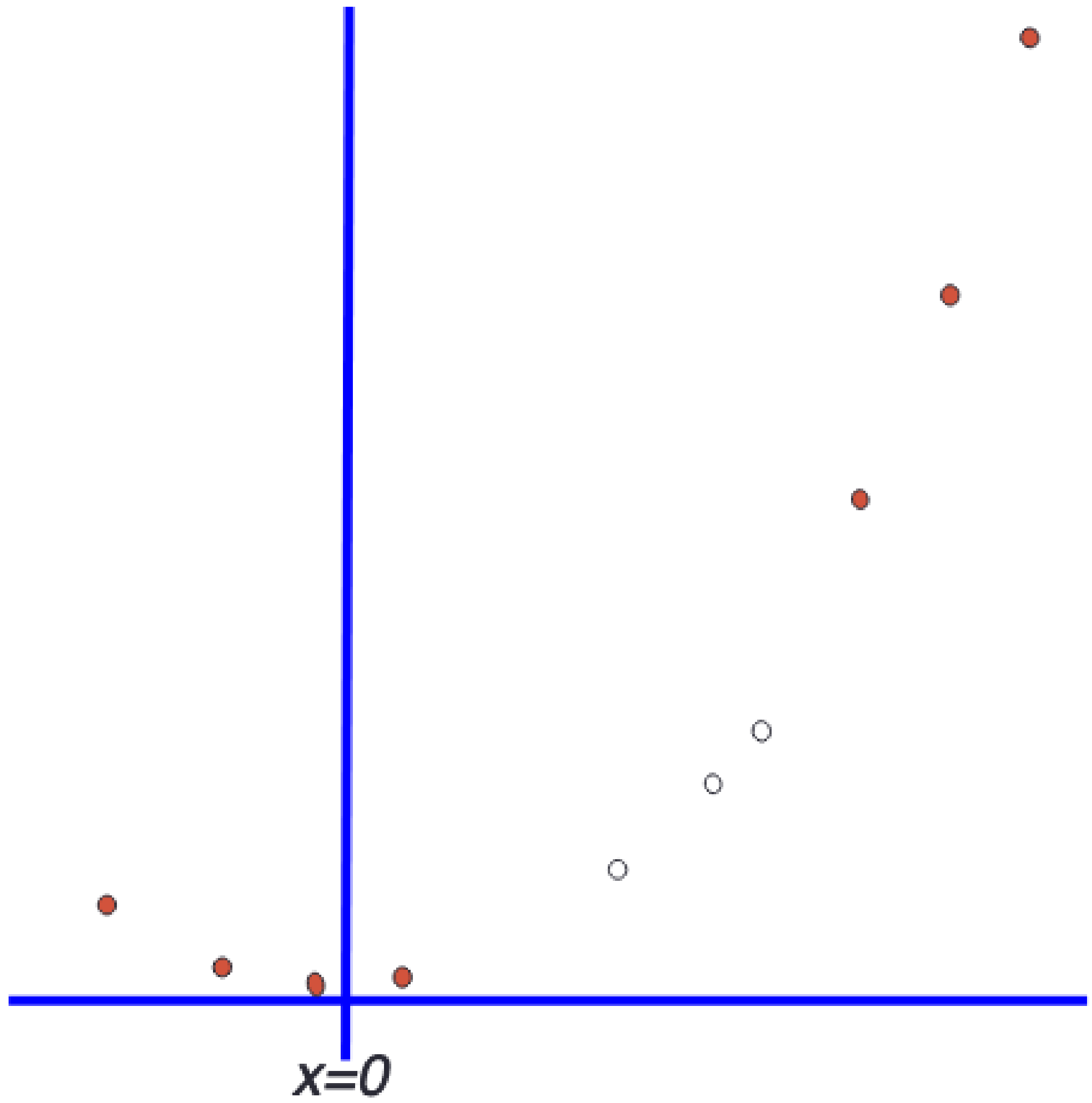
נניח שאנו במרחב חד ממדי – בעיה פשוטה



נניח שאנו במרחב חד ממדי – בעיה מסובבת

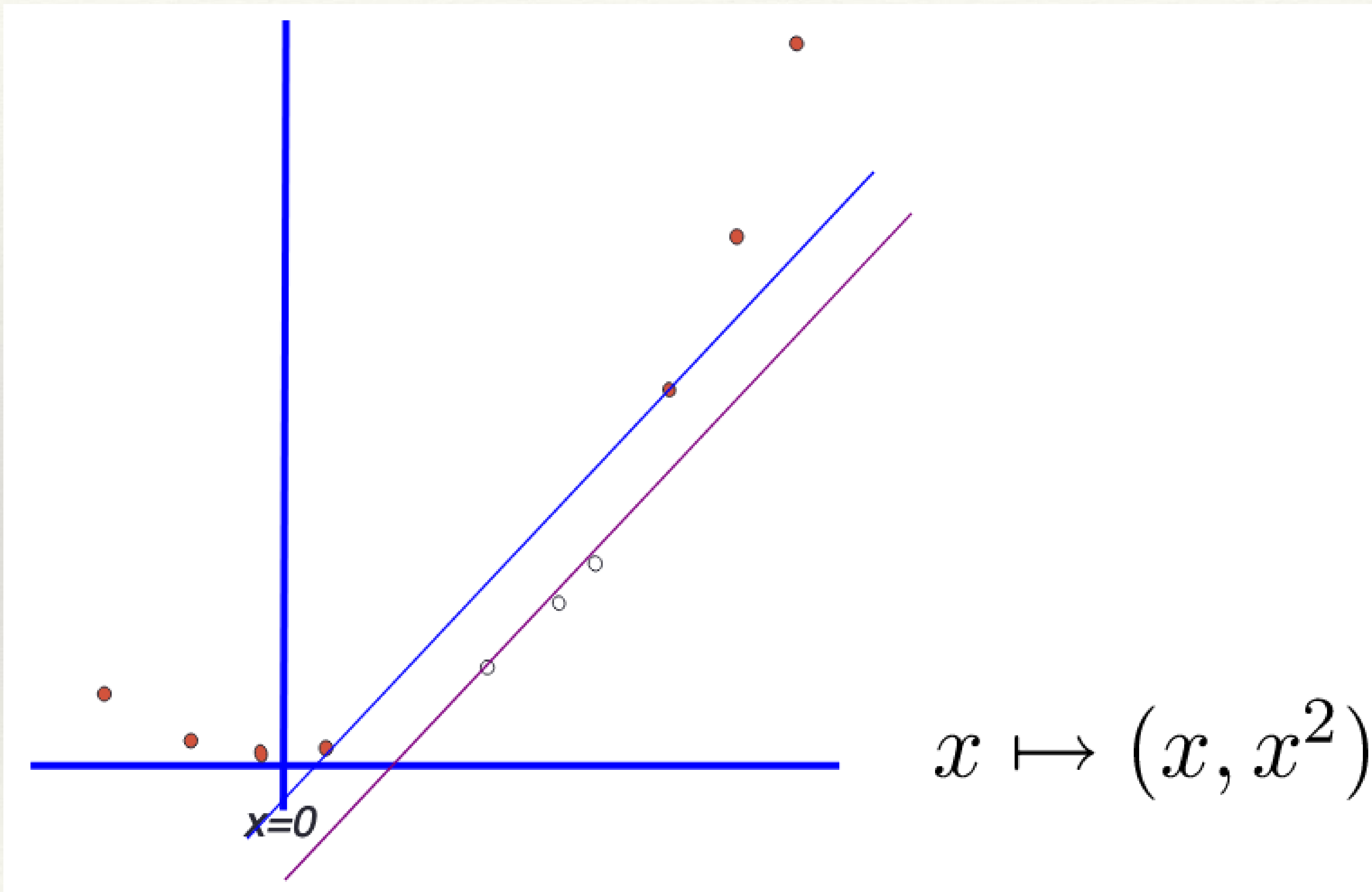


נמיר אותה לבעיה במרחב דו ממדי

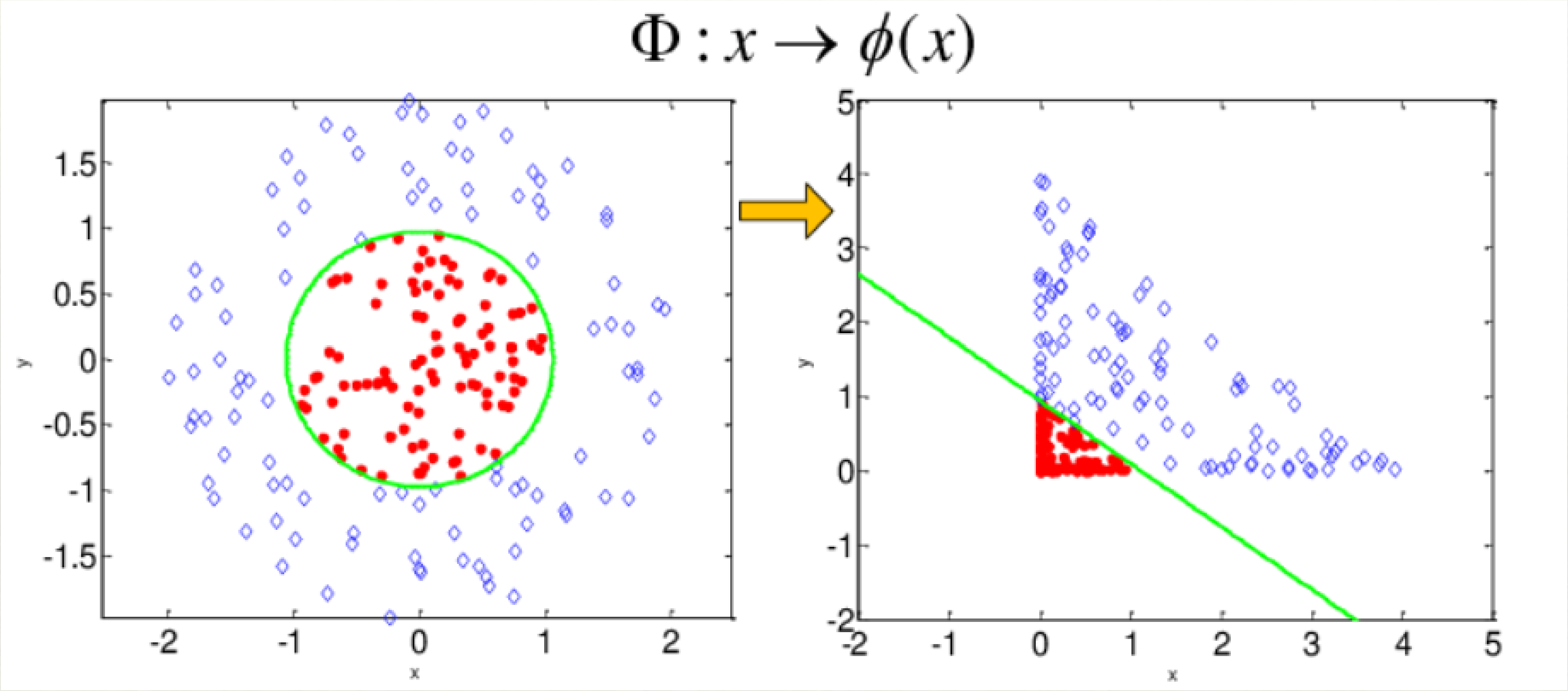


$$x \mapsto (x, x^2)$$

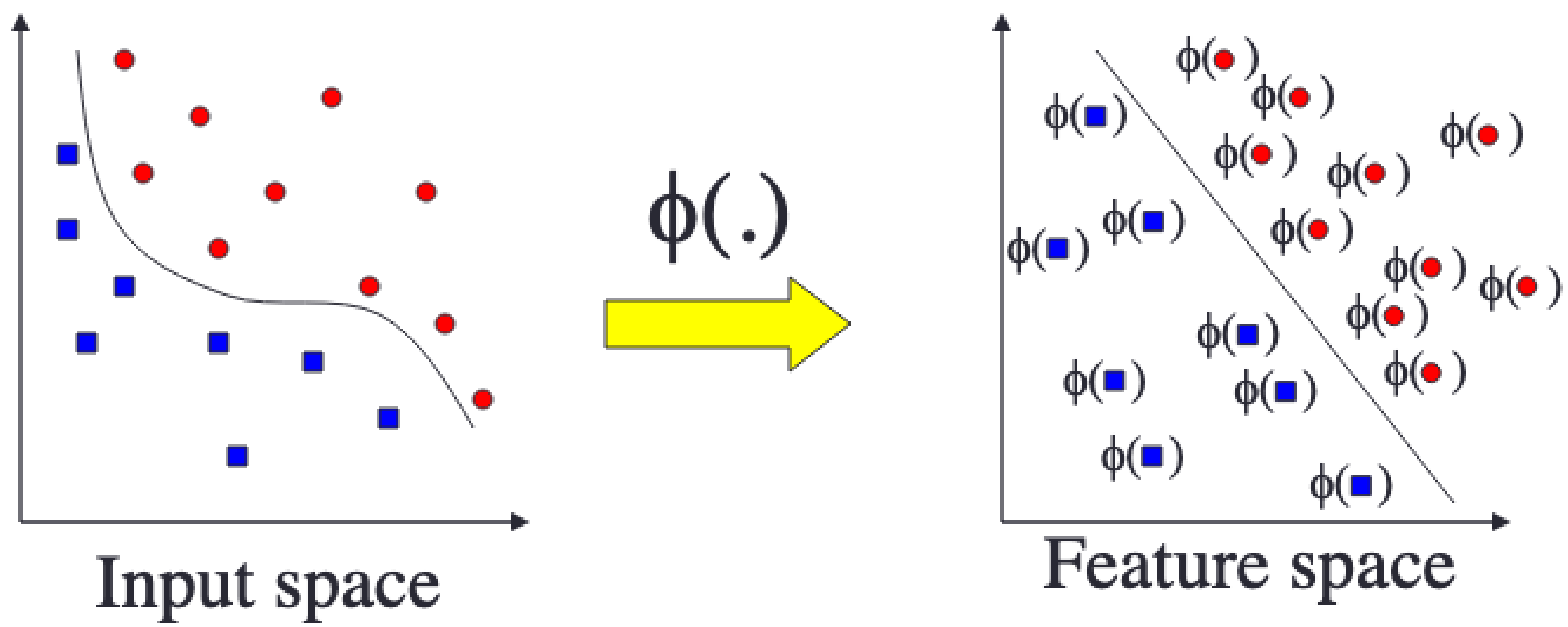
פתרון SVM במרחב דו ממדי



SVM with kernels



טרנספורמצית נתונים בפתרון



Note: The feature space is of a higher dimension than the input space

- Computation in the feature space can be costly because it is high dimensional
  - The feature space can be infinite-dimensional!
- The kernel trick comes to rescue

נמפה כל ווקטור למימד הגבוה

$$\theta(x_1,x_2)=[1,x_1^2,\sqrt{2}x_1x_2,x_2^2,\sqrt{2}x_1,\sqrt{2}x_2]$$

$$example_1=(-1,-1)\longrightarrow\theta(example_1)=(1,1,\sqrt{2},1,-\sqrt{2},-\sqrt{2})$$

$$example_2=(-1,+1)\longrightarrow\theta(example_2)=(1,1,-\sqrt{2},1,-\sqrt{2},\sqrt{2})$$

$$example_3=(+1,-1)\longrightarrow\theta(example_3)=(1,1,-\sqrt{2},1,\sqrt{2},-\sqrt{2})$$

$$example_4=(+1,+1)\longrightarrow\theta(example_4)=(1,1,\sqrt{2},1,\sqrt{2},\sqrt{2})$$

$$w=(0,\ 0,\ -\frac{\sqrt{2}}{2},\ 0,\ \ 0,\ \ \ 0)$$

$$y(example_1)=SGN(w\bullet\theta(example_1))=$$

$$\mathrm{sgn}(0\times1+0\times1+-\frac{\sqrt{2}}{2}\times\sqrt{2}+0\times1+0\times-\sqrt{2}+0\times-\sqrt{2})=SGN(-1)=-1$$

$$y(example_2)=SGN(w\bullet\theta(example_2))=$$

$$\mathrm{sgn}(0\times1+0\times1+-\frac{\sqrt{2}}{2}\times-\sqrt{2}+0\times1+0\times-\sqrt{2}+0\times\sqrt{2})=SGN(1)=1$$

$$y(example_3)=SGN(w\bullet\theta(example_3))=$$

$$\mathrm{sgn}(0\times1+0\times1+-\frac{\sqrt{2}}{2}\times-\sqrt{2}+0\times1+0\times\sqrt{2}+0\times-\sqrt{2})=SGN(1)=1$$

$$y(example_4)=SGN(w\bullet\theta(example_4))=$$

$$\mathrm{sgn}(0\times1+0\times1+-\frac{\sqrt{2}}{2}\times\sqrt{2}+0\times1+0\times\sqrt{2}+0\times\sqrt{2})=SGN(-1)=-1$$

אלו בדיוק הווקטורים מהתרגיל הראשון, רק שבזכות טריק הקרנל חסכנו את המיפוי המפורש למימד הגבוה