נניח . בכל שלב המעדכן

$$w^{(t+1)} = w^{(t)} + \eta \, y_i x_i \qquad \eta > 0$$

נוכיח שלכל מספר צעדים ה $k-i$, יקיים:

$$\underbrace{w^{(k+1)} \cdot w^*} \geq w^{(k)} \cdot w^* + r$$

$$w^{(k)} \cdot w^* + \eta \, y_i x_i \cdot v \geq w^{(k)} \cdot w^* + r$$

$$y_i x_i \cdot w^* \geq 0 \qquad \text{(כיון של טעות)}$$

$$\eta \geq 0 \qquad \text{(צעד חיובי)}$$

לכל סכום שונה M צעדים (כל שלב)

$$w^{(k+1)} \cdot w^* \geq M r \qquad ②$$

(כלומר k צעדים $w^{(k+1)} \cdot w^* \geq kr$.)

נחסום את הנורמה מלמעלה:

$$\| w^{(k+1)} \|^2 \leq M R^2$$

$$\| w^{(k)} + \eta \, y_i x_i \|^2 = \| w^{(k)} \|^2 + 2 \eta y_i \| w^{(k)} x_i \| + \underbrace{\eta^2 \| x_i \|^2}_{y^2 = 1}$$

כי אם מניח $\| x_i \| \leq R$ לכל

$$\| w^{(k+1)} \|^2 \leq \| w^{(k)} \|^2 + \underbrace{2 \eta y_i \| w^{(k)} x_i \|}_{\eta > 0 \;\; \text{כי}} + \eta^2 \| x_i \|^2 \leq \| w^{(k)} \|^2 + \eta^2 R^2$$

$$\Leftarrow \text{לכל} \; M \; \text{צעדים}$$

$$y_i \| w^{(k)} x_i \| \leq 0 \;\; \text{כי טעות}$$

$$\| w^{(k+1)} \|^2 \leq M \cdot R^2 \cdot \eta^2$$

$$\sqrt{M} R \eta \geq \| w^{(k+1)} \|^2 \geq w^{(k+1)} \cdot w^* \geq M r$$

$$\overline{\text{עד כאן עובד}}$$

$$\sqrt{M} \leq \frac{R \eta}{r}$$

$$M \leq \left(\frac{R \eta}{r} \right)^2 \quad \Leftarrow = \left(\frac{R}{r} \right)^2$$

טענה: לא ניתן ל perceptron יחיד לממש את פונקציית parity

על קלט של 3-features.

הוכחה: נניח שקיים perceptron כזה שמקבל את טבלת האמת הבאה של ה perceptron:

X_1	X_2	X_3	f
0	0	0	1
0	0	1	0
0	1	0	0
0	1	1	1
1	0	0	0
1	0	1	1
1	1	0	1
1	1	1	0

נניח כי W_1, W_2, W_3 הם המשקלות של ה perceptron

θ הוא הסף (threshold) של ה perceptron

כאשר מבצעים חישוב של parity

א. נסתכל על טבלת האמת ונקבל:

$(1)\quad X_1=0\;\; X_2=0\;\; X_3=0 \;\Rightarrow\; W_1X_1+W_2X_2+W_3X_3 = 0 \;\Rightarrow\; 0 > \theta$

$(2)\quad X_1=0\;\; X_2=0\;\; X_3=1 \;\Rightarrow\; W_1X_1+W_2X_2+W_3X_3 = W_3 \;\Rightarrow\; W_3 < \theta$

$(3)\quad X_1=0\;\; X_2=1\;\; X_3=0 \;\Rightarrow\; W_1X_1+W_2X_2+W_3X_3 = W_2 \;\Rightarrow\; W_2 < \theta$

$(4)\quad X_1=0\;\; X_2=1\;\; X_3=1 \;\Rightarrow\; W_1X_1+W_2X_2+W_3X_3 = W_2+W_3 \;\Rightarrow\; W_2+W_3 > \theta$

$(5)\quad X_1=1\;\; X_2=0\;\; X_3=0 \;\Rightarrow\; W_1X_1+W_2X_2+W_3X_3 = W_1 \;\Rightarrow\; W_1 < \theta$

$(6)\quad X_1=1\;\; X_2=0\;\; X_3=1 \;\Rightarrow\; W_1X_1+W_2X_2+W_3X_3 = W_1+W_3 \;\Rightarrow\; W_1+W_3 > \theta$

$(7)\quad X_1=1\;\; X_2=1\;\; X_3=0 \;\Rightarrow\; W_1X_1+W_2X_2+W_3X_3 = W_1+W_2 \;\Rightarrow\; W_1+W_2 > \theta$

$(8)\quad X_1=1\;\; X_2=1\;\; X_3=1 \;\Rightarrow\; W_1X_1+W_2X_2+W_3X_3 = W_1+W_2+W_3 \;\Rightarrow\; W_1+W_2+W_3 < \theta$

כעת נשים לב לסימן של θ לפי (1) לכל W_1, W_2, W_3 כי

לכן ($W_x < \theta < 0 \;\Leftarrow$ לפי $(1),(3),(5)$)

מכאן של W_1, W_2, W_3 לפי (4) לכל הצירופים (וכן $(6),(7)$)

אך ישנה סתירה ביחס ל-1 $W_2+W_3 > \theta$ [ולכן $W_2+W_3 < 2\theta$]

$\overset{\wedge}{\theta}\;\;\overset{\wedge}{\theta}$

לכן הגענו לסתירה של perceptron לא יכול לממש את פונקציית parity על

input של 3-features. $\blacksquare$