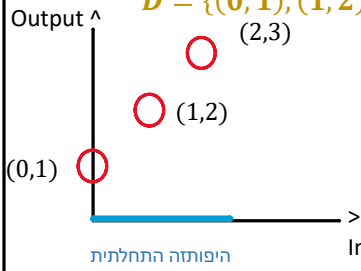


נריץ Epoch על דוגמה אחד:

$D = \{(0, 1), (1, 2), (2, 3)\}$, [Example: in the first example (0 is the input, 1 is the True Output)]
תנאי התחלה:

1. משקולות מאופסות (גם בIAS) $h = y(x) = w_1x + w_0$

2. קצב למידה $\lambda = 0.1$



נחשב את השגיאה בתלות בהיפוטזה ובנקודות

$$MSE_{D,h}(w_0, w_1) = \frac{1}{2m} \{ [t_1 - (w_1x_1 + w_0)]^2 + [t_2 - (w_1x_2 + w_0)]^2 + [t_3 - (w_1x_3 + w_0)]^2 \}$$

נשפר את w_0 : 1. נציב את ערכי w_1 למשוואה שלנו,

$$MSE_{D,h}(w_0, w_1) = \frac{1}{2m} \{ [1 - (w_1 \cdot 0 + w_0)]^2 + [2 - (w_1 \cdot 1 + w_0)]^2 + [3 - (w_1 \cdot 2 + w_0)]^2 \}$$

1. נגזור את המשוואה לפי w_0 נשים לב שגזירה של סכום, הוא סכום הנגזרות, ויש w_1 הוא קבוע פה chainRule

$$\frac{d([1 - (w_1 \cdot 0 + w_0)])^2}{dw_0} \Rightarrow 2 * \frac{d[1 - (w_1 \cdot 0 + w_0)]}{dw_0} * [1 - (w_1 \cdot 0 + w_0)] = 2 * -1 * [1 - (w_1 \cdot 0 + w_0)]$$

$$\frac{d([2 - (w_1 \cdot 1 + w_0)])^2}{dw_0} \Rightarrow 2 * \frac{d[2 - (w_1 \cdot 1 + w_0)]}{dw_0} * [2 - (w_1 \cdot 1 + w_0)] = 2 * -1 * [2 - (w_1 \cdot 1 + w_0)]$$

$$\frac{d([3 - (w_1 \cdot 2 + w_0)])^2}{dw_0} \Rightarrow 2 * \frac{d[3 - (w_1 \cdot 2 + w_0)]}{dw_0} * [3 - (w_1 \cdot 2 + w_0)] = 2 * -1 * [3 - (w_1 \cdot 2 + w_0)]$$

$$\frac{MSE_{D,h}(w_0, w_1)}{dw_0} = \{-2 * [1 - (w_1 \cdot 0 + w_0)]\} + \{-2 * [2 - (w_1 \cdot 1 + w_0)]\} + \{-2 * [3 - (w_1 \cdot 2 + w_0)]\}$$

3. נגזור את המשוואה לפי w_1 , נשים לב שגזירה של סכום, הוא סכום הנגזרות, ויש w_0 הוא קבוע פה

$$\frac{d([1 - (w_1 \cdot 0 + w_0)])^2}{dw_1} \Rightarrow 2 * \frac{d[1 - (w_1 \cdot 0 + w_0)]}{dw_1} * [1 - (w_1 \cdot 0 + w_0)] = 2 * 0 * []$$

$$\frac{d([2 - (w_1 \cdot 1 + w_0)])^2}{dw_1} \Rightarrow 2 * \frac{d[2 - (w_1 \cdot 1 + w_0)]}{dw_1} * [2 - (w_1 \cdot 1 + w_0)] = 2 * -1 * [2 - (w_1 \cdot 1 + w_0)]$$

$$= -4 + 2w_1 + 2w_0$$

$$\frac{d([3 - (w_1 \cdot 2 + w_0)])^2}{dw_1} \Rightarrow 2 * \frac{d[3 - (w_1 \cdot 2 + w_0)]}{dw_1} * [3 - (w_1 \cdot 2 + w_0)] = 2 * -2 * [3 - (w_1 \cdot 2 + w_0)]$$

$$= -12 + 8w_1 + 4w_0$$

$$\frac{MSE_{D,h}(w_0, w_1)}{dw_1} = 0 + (-4 + 2w_1 + 2w_0) + (-12 + 8 + -4w_0) = -16 + 10w_1 + 6w_0$$

$$\frac{MSE_{D,h}(w_0, w_1)}{dw}$$

קיבלנו את הגרדיאנט של הביטוי לפי כל המשקולות

$$\frac{MSE_{D,h}(w_0, w_1)}{dw_0} = \{-2 * [1 - (w_1 \cdot 0 + w_0)]\} + \{-2 * [2 - (w_1 \cdot 1 + w_0)]\} + \{-2 * [3 - (w_1 \cdot 2 + w_0)]\}$$

$$\frac{MSE_{D,h}(w_0, w_1)}{dw_1} = -16 + 10w_1 + 6w_0$$

נציב את ערכי המשקולות ההתחלתיים שלנו $w_0 = 0, w_1 = 0$

$$w_0 \text{step} = (-2) + (-4) + (-6) * \lambda = -12 * 0.1 = -1.2$$

$$w_1 \text{step} = [-16 + 10 * 0 + 6 * 0] * \lambda = -16 * 0.1 = -1.6$$

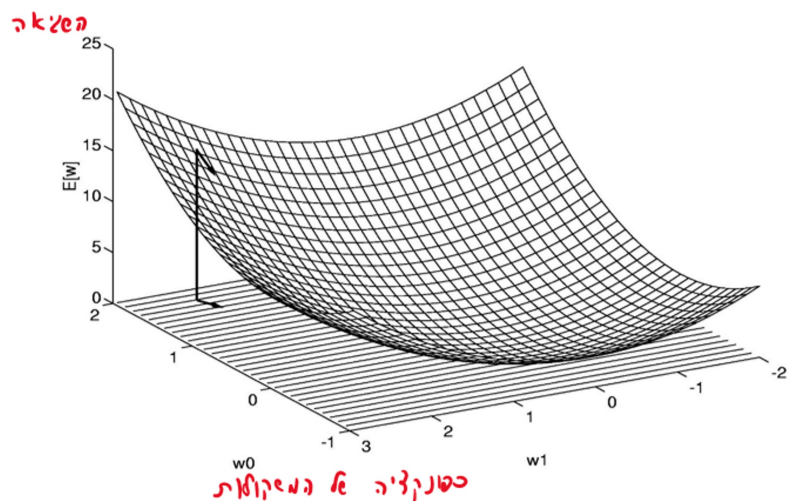
נחשב את ערכי המשקולות החדשים:

$$w_0 \text{New} = w_0 \text{old} - w_0 \text{step} = 0 - (-1.2) = 1.2$$

$$w_1 \text{New} = w_1 \text{old} - w_1 \text{step} = 0 - (-1.6) = 1.6$$

D - Set of given Points
 t, x - scalars
 w_i - variables

$$MSE_{D,h}(w_0, w_1) = \frac{1}{2m} \sum_{i \in D} \underbrace{(t_i - (w_1 x_i + w_0))}_{\text{error}}^2$$



11. [6] נתונה קבוצת אימון $D = \{(0,1), (1,2), (2,3)\}$ ומעונינים לבצע רגרסיה ליניארית למשתנה בודד (Single x (feature) בשיטת GD. מתחילים במשקולות מאופסות (גם הביאס) וקצב הלמידה 0.1. מהן המשקולות שתתקבלנה לאחר EPOCH אחד של GD? לשם הבהרה: הדוגמא (0,1) שב D מורכבת מקלט $x=0$ ותיוג $t=1$

$$W1 = 0.1 \text{mean}(t-y)x = 0.1(1/3)((1-0)0 + (2-0)1 + (3-0)2) = (1/30)(2+6) = 8/30$$

$$W0 = 0.1 \text{mean}(t-y) = 0.1(1/3)((1-0) + (2-0) + (3-0)) = (1/30)(1+2+3) = 6/30$$