# Cost Landscape Plotting of Deep Learning

Data =

| $\mathbf{x}_1$ | $\mathbf{x}_2$ | $\mathbf{y}$ |
|:---:|:---:|:---:|
| -1 | -1 | 2 |
| 0 | -1 | 1 |
| 1 | -1 | 2 |
| -1 | 0 | 1 |
| 0 | 0 | 0 |
| 1 | 0 | 1 |
| -1 | 1 | 2 |
| 0 | 1 | 1 |
| 1 | 1 | 0 |

## 1. กำหนด Architecture
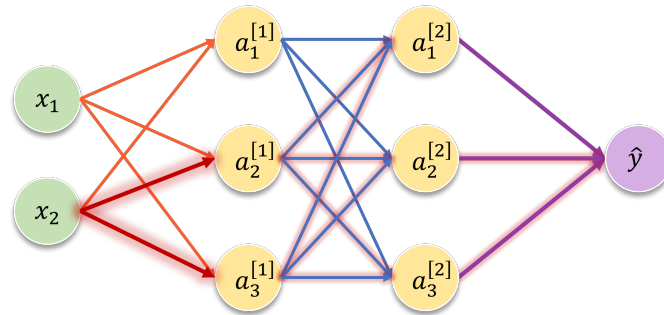


$$\hat{y} = b^{[out]} + w_1^{[out]}a_1^{[2]} + w_2^{[out]}a_2^{[2]} + w_3^{[out]}a_3^{[2]}$$

$$a_1^{[2]} = ReLU(b_1^{[2]} + w_{1,1}^{[2]}a_1^{[1]} + w_{2,1}^{[2]}a_2^{[1]} + w_{3,1}^{[2]}a_3^{[1]})$$
$$a_2^{[2]} = ReLU(b_2^{[2]} + w_{1,2}^{[2]}a_1^{[1]} + w_{2,2}^{[2]}a_2^{[1]} + w_{3,2}^{[2]}a_3^{[1]})$$
$$a_3^{[2]} = ReLU(b_3^{[2]} + w_{1,3}^{[2]}a_1^{[1]} + w_{2,3}^{[2]}a_2^{[1]} + w_{3,3}^{[2]}a_3^{[1]})$$

$$a_1^{[1]} = ReLU(b_1^{[1]} + w_{1,1}^{[1]}x_1 + w_{2,1}^{[1]}x_2)$$
$$a_2^{[1]} = ReLU(b_2^{[1]} + w_{1,2}^{[1]}x_1 + w_{2,2}^{[1]}x_2)$$
$$a_3^{[1]} = ReLU(b_3^{[1]} + w_{1,3}^{[1]}x_1 + w_{2,3}^{[1]}x_2)$$

2. คำนวณ $\hat{y} = b^{[out]} + w_1^{[out]}a_1^{[2]} + w_2^{[out]}a_2^{[2]} + w_3^{[out]}a_3^{[2]}$ จาก $w_{2,2}^{[1]}$ และ $w_{2,3}^{[1]}$ แต่ละชุด โดยเราจะกำหนดให้

$$b^{[1]} = \begin{bmatrix} -2 & -1 & 1 \end{bmatrix}, b^{[2]} = \begin{bmatrix} -2 & -1 & -2 \end{bmatrix}, b^{[out]} = \begin{bmatrix} -1 \end{bmatrix}$$

และ $w^{[1]} = \begin{bmatrix} -1 & 3 & -3 \\ 0 & ? & ? \end{bmatrix}, w^{[2]} = \begin{bmatrix} -1 & 1 & 0 \\ 1 & 1 & 0 \\ 3 & 2 & 0 \end{bmatrix}, w^{[out]} = \begin{bmatrix} -1 \\ -1 \\ 2 \end{bmatrix}$

$$\hat{y} = -1 + (-1)a_1^{[2]} + (-1)a_2^{[2]} + 2a_3^{[2]}$$

$$a_1^{[2]} = ReLU(-2 + (-1)a_1^{[1]} + (1)a_2^{[1]} + (3)a_3^{[1]})$$
$$a_2^{[2]} = ReLU(-1 + (1)a_1^{[1]} + (1)a_2^{[1]} + (2)a_3^{[1]})$$
$$a_3^{[2]} = ReLU(-2 + (0)a_1^{[1]} + (0)a_2^{[1]} + (0)a_3^{[1]})$$

$$a_1^{[1]} = ReLU\left(-2 + (-1)x_1 + (0)x_2\right)$$
$$a_2^{[1]} = ReLU\left(-1 + (3)x_1 + w_{2,2}^{[1]}x_2\right)$$
$$a_3^{[1]} = ReLU\left(1 + (-3)x_1 + w_{2,3}^{[1]}x_2\right)$$

| $x_1$ | $x_2$ | $w_{2,2}^{[1]}=-15$ $w_{2,3}^{[1]}=-5$ | $w_{2,2}^{[1]}=-14$ $w_{2,3}^{[1]}=-5$ | $w_{2,2}^{[1]}=-13$ $w_{2,3}^{[1]}=-5$ | $w_{2,2}^{[1]}=-12$ $w_{2,3}^{[1]}=-5$ | $\cdots$ | $w_{2,2}^{[1]}=12$ $w_{2,3}^{[1]}=5$ | $w_{2,2}^{[1]}=13$ $w_{2,3}^{[1]}=5$ | $w_{2,2}^{[1]}=14$ $w_{2,3}^{[1]}=5$ | $w_{2,2}^{[1]}=15$ $w_{2,3}^{[1]}=5$ |
|---|---|---|---|---|---|---|---|---|---|---|
|  |  | $\hat{y}_1$ | $\hat{y}_2$ | $\hat{y}_3$ | $\hat{y}_4$ | $\cdots$ | $\hat{y}_{338}$ | $\hat{y}_{339}$ | $\hat{y}_{340}$ | $\hat{y}_{341}$ |
| -1 | -1 | -65 | -63 | -61 | -59 | $\cdots$ | -1 | -1 | -1 | -1 |
| 0 | -1 | -56 | -54 | -52 | -50 | $\cdots$ | -1 | -1 | -1 | -1 |
| 1 | -1 | -47 | -45 | -43 | -41 | $\cdots$ | -1 | -1 | -1 | -1 |
| -1 | 0 | -18 | -18 | -18 | -18 | $\cdots$ | -18 | -18 | -18 | -18 |
| 0 | 0 | -3 | -3 | -3 | -3 | $\cdots$ | -3 | -3 | -3 | -3 |
| 1 | 0 | -2 | -2 | -2 | -2 | $\cdots$ | -2 | -2 | -2 | -2 |
| -1 | 1 | -1 | -1 | -1 | -1 | $\cdots$ | -57 | -55 | -53 | -51 |
| 0 | 1 | -1 | -1 | -1 | -1 | $\cdots$ | -48 | -46 | -44 | -42 |
| 1 | 1 | -1 | -1 | -1 | -1 | $\cdots$ | -39 | -41 | -43 | -45 |

3. คำนวณ $Cost = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2$ ของ $\hat{\mathbf{y}}$ แต่ละชุด

ตัวอย่างเช่น

| $w_{2,2}^{[1]} = -13$ $w_{2,3}^{[1]} = -5$ | | | |
|---|---|---|---|
| $\hat{\mathbf{y}}_3$ | $\mathbf{y}$ | $y - \hat{y}_1$ | $(y - \hat{y}_1)^2$ |
| -61 | 2 | -63 | 3969 |
| -54 | 1 | -55 | 2809 |
| -43 | 0 | -43 | 2025 |
| -18 | 1 | -19 | 361 |
| -3 | 0 | -3 | 9 |
| -2 | 1 | -3 | 9 |
| -1 | 2 | -3 | 9 |
| -1 | 1 | -2 | 4 |
| -1 | 0 | -1 | 1 |
| | | $\sum_{i=1}^{n}(y_i - \hat{y}_i)^2 =$ | 9196 |

ดังนั้น $Cost = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2 = \frac{1}{9} \times 9196 = 1021.78$

4. นำ $Cost$ ของ weight แต่ละชุดไป plot ลงกราฟ

| | $w_{2,2}^{[1]} = -15$ $w_{2,3}^{[1]} = -5$ | $w_{2,2}^{[1]} = -14$ $w_{2,3}^{[1]} = -5$ | $w_{2,2}^{[1]} = -13$ $w_{2,3}^{[1]} = -5$ | $w_{2,2}^{[1]} = -12$ $w_{2,3}^{[1]} = -5$ | $\cdots$ | $w_{2,2}^{[1]} = 12$ $w_{2,3}^{[1]} = 5$ | $w_{2,2}^{[1]} = 13$ $w_{2,3}^{[1]} = 5$ | $w_{2,2}^{[1]} = 14$ $w_{2,3}^{[1]} = 5$ | $w_{2,2}^{[1]} = 15$ $w_{2,3}^{[1]} = 5$ |
|---|---|---|---|---|---|---|---|---|---|
| $Cost$ | 1170.22 | 1094.67 | 1021.78 | 951.56 | $\cdots$ | 867.11 | 933.78 | 1003.11 | 1075.11 |