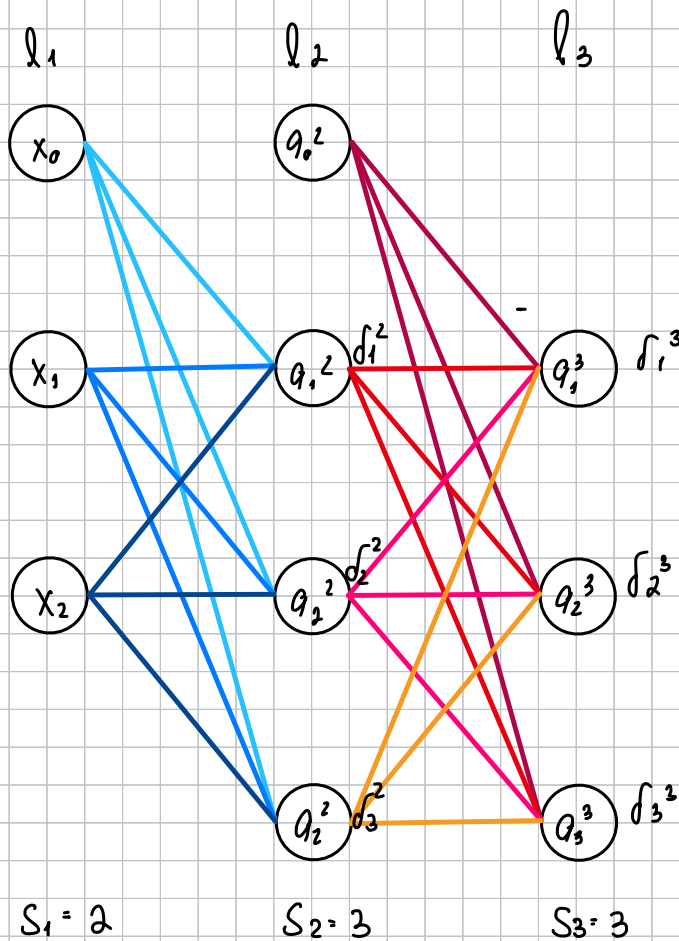


תרגיל 4: ערשלה

2. (30 נקודות) נתונה רשת עם שכבת קלט של שתי יחידות (כל דוגמת אימון עם שתי תכונות), שכבת Hidden עם שלוש יחידות ושכבת פלט עם שלוש יחידות לסווג מזג האוויר לשלושה מצבים אפשריים - "שמש", "מעונן" או "גשום". הניחו כי פונקציית האקטיבציה עבור כל היחידות בשכבת ה- hidden היא שכבת ReLU ובשכבת הפלט היא פונקציית sigmoid.

כתבו באופן מפורש את חישובי כל ה- $\delta_i^{(l)}$ בכל אחת מהשכבות, תחילה עבור כל i בכל אחת מהשכבות, ולאחר מכן הציגו את החישובים באופן וקטורי-מטריצי.

כתבו באופן מפורש את חישובי כל הגרדיאנטים, כלומר את חישובי $\frac{\partial J(\theta)}{\partial \theta_{ij}^{(l)}}$ לכל i ו- j עבור כל אחת מהשכבות הרלבנטיות, ולאחר מכן רשמו את החישובים באופן מטריצי-וקטורי.



$$\delta_i^{(l)} = \begin{cases} g'(z_i^{(l)}) \cdot \sum_{k=0}^{S_{l+1}-1} \theta_{ki}^{(l+1)} \delta_k^{(l+1)}, & l < L \\ a_i^{(l)} - y_i, & l = L \end{cases}$$

$$z_1^1 = x_0 \cdot \theta_{10}^1 + x_1 \cdot \theta_{11}^1 + x_2 \cdot \theta_{12}^1$$

$$z_2^1 = x_0 \cdot \theta_{20}^1 + x_1 \cdot \theta_{21}^1 + x_2 \cdot \theta_{22}^1$$

$$z_3^1 = x_0 \cdot \theta_{30}^1 + x_1 \cdot \theta_{31}^1 + x_2 \cdot \theta_{32}^1$$

$$z_1^2 = \theta_{10}^2 a_0^1 + \theta_{11}^2 a_1^1 + \theta_{12}^2 a_2^1$$

$$z_2^2 = \theta_{20}^2 a_0^1 + \theta_{21}^2 a_1^1 + \theta_{22}^2 a_2^1$$

$$z_3^2 = \theta_{30}^2 a_0^1 + \theta_{31}^2 a_1^1 + \theta_{32}^2 a_2^1$$

$$a_1^1 = \text{relu}(z_1^1)$$

$$a_1^3 = \text{sig}(z_1^3)$$

$$a_2^1 = \text{relu}(z_2^1)$$

$$a_2^3 = \text{sig}(z_2^3)$$

$$a_3^1 = \text{relu}(z_3^1)$$

$$a_3^3 = \text{sig}(z_3^3)$$

$$\delta_1^3 = a_1^3 - y_1$$

$$\delta_2^3 = a_2^3 - y_2$$

$$\delta_3^3 = a_3^3 - y_3$$

$$\delta_1^2 = \text{relu}'(z_1^2) \cdot (\theta_{11}^2 \delta_1^3 + \theta_{21}^2 \delta_2^3 + \theta_{31}^2 \delta_3^3)$$

$$\delta_2^2 = \text{relu}'(z_2^2) \cdot (\theta_{12}^2 \delta_1^3 + \theta_{22}^2 \delta_2^3 + \theta_{32}^2 \delta_3^3)$$

$$\delta_3^2 = \text{relu}'(z_3^2) \cdot (\theta_{13}^2 \delta_1^3 + \theta_{23}^2 \delta_2^3 + \theta_{33}^2 \delta_3^3)$$

חישוב חסר וקטרי :

$$\delta^3 = \begin{pmatrix} a_1^3 \\ a_2^3 \\ a_3^3 \end{pmatrix} - \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} = a^3 - y$$

3×1

$$\Theta^2 = \left(\begin{pmatrix} \theta_{10}^2 \\ \theta_{20}^2 \\ \theta_{30}^2 \end{pmatrix} \begin{pmatrix} \theta_{11}^2 \\ \theta_{21}^2 \\ \theta_{31}^2 \end{pmatrix} \begin{pmatrix} \theta_{12}^2 \\ \theta_{22}^2 \\ \theta_{32}^2 \end{pmatrix} \begin{pmatrix} \theta_{13}^2 \\ \theta_{23}^2 \\ \theta_{33}^2 \end{pmatrix} \right)$$

$\Theta_0^2 \quad \Theta_1^2 \quad \Theta_2^2 \quad \Theta_3^2$

3×4

$$\delta_1^2 = \underset{3 \times 1}{\Theta_1^{2T}} \cdot \underset{1 \times 3}{\delta^3} \cdot \text{relu}'(z_1^2)$$

$$\delta_2^2 = \Theta_2^{2T} \cdot \delta^3 \cdot \text{relu}'(z_2^2)$$

$$\delta_3^2 = \Theta_3^{2T} \cdot \delta^3 \cdot \text{relu}'(z_3^2)$$

חישוב הנגזרת:

$$\frac{\partial J(\theta)}{\partial \theta_{ij}^1} = a_j^1 \cdot \delta_i^{1+1}$$

$$\frac{\partial J(\theta)}{\partial \theta_{ij}^1} = a_j^2 \cdot \delta_i^3 \quad \forall i = \{1, 2, 3\}, j = \{0, 1, 2, 3\}$$

$$\frac{\partial J(\theta)}{\partial \theta_{ij}^1} = a_j^1 \cdot \delta_i^2 \quad \forall i = \{1, 2, 3\}, j = \{0, 1, 2\}$$

חישוב מרחיב וקטורי:

$$a^2 = \begin{pmatrix} a_0^2 \\ a_1^2 \\ a_2^2 \\ a_3^2 \end{pmatrix} \quad \delta^2 = \begin{pmatrix} \delta_1^3 \\ \delta_2^3 \\ \delta_3^3 \end{pmatrix}$$

4×1 3×1

$$D^2 = \delta^3 \cdot a^{2T}$$

3×4 3×1 1×4

$$D^1 = \delta^2 \cdot a^{1T}$$

