

Wong Wen Tai (1002323)

$$J = (-1) \sum_{i=1}^n y_i \log(h(x_i)) + (1-y_i) \log(1-h(x_i))$$

$$\nabla_w J = \nabla_w \left((-1) \sum_{i=1}^n y_i \log(s(w \cdot x_i)) + (1-y_i) \log(1-s(w \cdot x_i)) \right)$$

$$= (-1) \sum_{i=1}^n x_i y_i (1-s(w \cdot x_i)) + (1-y_i) (-s(w \cdot x_i)) (x_i)$$

$$= (-1) \sum_{i=1}^n x_i y_i - x_i y_i s(w \cdot x_i) + (-x_i s(w \cdot x_i)) + x_i y_i s(w \cdot x_i)$$

$$= \sum_{i=1}^n -x_i (y_i - s(w \cdot x_i))$$

$$= \sum_{i=1}^n x_i (s(w \cdot x_i) - y_i)$$

$$= \sum_{i=1}^n x_i (h(x_i) - y_i) \quad \# QED$$

2. 1) $C_{jk} = \sum_i A_{ijk} b_i = A_{jk} b_i$
 $ijk, i \rightarrow jk, [A, b]$

2) $C_j = \sum_{i,k} A_{ijk} b_{ik} = A_{ijk} b_{ik}$
 $ijk, ik \rightarrow j, [A, b]$

3) $A_{ik} = \sum_j A_{ijk} = A_{ijk}$
 $ijkl \rightarrow ik, [A]$

4) $A_{ki} = \sum_j A_{ijk} = A_{ijk}$
 $ijkl \rightarrow ki, [A]$

5) $C_i = \sum_{j,k} A_{ijk} A_{ijk} = A_{ijk} A_{ijk}$
 $ijk, ijk \rightarrow i, [A]$

6) $C = x^T A x, x \in \mathbb{R}^d, 1\text{-tensor}, A \in \mathbb{R}^{d \times d}, 2\text{-tensor}$

$$C = \sum_{j,k} x_j A_{jk} x_k = x_j A_{jk} x_k$$

$ijk, k \rightarrow, [x, A, x]$

7) $C = A G^T B, A \in \mathbb{R}^{d \times e}, 2\text{-tensor}, G \in \mathbb{R}^{f \times e}, 1\text{-tensor}, B \in \mathbb{R}^{f \times d}, 2\text{-tensor}$

$$C_{de} = \sum_{f,l} A_{de} G_{fl} B_{fl} = A_{de} G_{fl} B_{fl}$$

$$de, fl, fl \rightarrow d, [A, G, B]$$

8) $C_{abef} = \sum_{c,d} A_{abcd} B_{bcde} E_{cdef}$

$$abcd, bcde, cdef \rightarrow abef, [A, B, E]$$

3. 1) No

1) Yes. (3, 2, 5, 3, 4)

3) Yes. (3, 2, 5, 3, 4)

4) Yes. (3, 2, 5, 3, 2)

5) Yes. (3, 2, 3, 3, 2)

6) Yes. (7, 7)