exercise a) $L_N(\pi) = \frac{1}{N}\sum_{n=1}^{N} \log(\exp(-a_n w^T \phi(x_n)) + 1)$

Prove that:

$(=) -\frac{1}{N}\sum_{n=1}^{N} \log \pi(a_n | x_n) = \frac{1}{N}\sum_{n=1}^{N} \log(\exp(-a_n w^T \phi(x_n)) + 1)$

$(=) -\frac{1}{N}\sum_{n=1}^{N} \log\left(\frac{1}{1+\exp(-a w^T \phi(x))}\right) (=) -\frac{1}{N}\sum_{n=1}^{N} \log(1) - \log(1+\exp(-a w^T \phi))$

$(=) \boxed{+\frac{1}{N}\sum_{n=1}^{N} +\log(1+\exp(-a w^T \phi(x))}$

$\square$.

Exercise 1| B)

$$g = \nabla_W \hat{L}_N(\pi) = \nabla_W \left[\frac{1}{N} \sum_{n=1}^{N} \log(\exp(-a_n w^T \phi(x_n)) + 1) \right]$$

a) $e^x \cdot e^{-x} = 1$

b) $\log(x) = \frac{x'}{x}$

c) $\pi(-a|x) = \frac{1}{1+\exp(a_n w^T \phi(x_n))}$

d) $\pi(-a|x) = 1 - \pi(a|x)$

$$\stackrel{\text{rule b}}{=} \frac{1}{N} \sum_{n=1}^{N} \frac{\exp(-a_n w^T \phi(x_n)) \cdot (a_n \phi(x_n))}{\exp(-a_n w^T \phi(x_n) + 1)} \cdot \frac{\exp(a_n w^T \phi(x_n))}{\exp(a_n w^T \phi(x_n))}$$

$$\stackrel{\text{rule a}}{=} \frac{1}{N} \sum_{n=1}^{N} \frac{-a_n \phi(x_n)}{\exp(a_n w^T \phi(x_n)) + 1}$$

$$\stackrel{\text{rule c+d}}{=} \frac{1}{N} \sum_{n=1}^{N} a_n \phi(x_n) (\pi(a_n|x_n) - 1) \quad \square.$$

Exercise 1) c

a) $e^{x} \cdot e^{-x} = 1$

b) $a_n \cdot a_n = 1$

c) $\log(x) = \frac{x'}{x}$

d) $\frac{U}{V} = \frac{U'-V'}{V^2}$

e) $\pi(-a_n|x_n) = \frac{1}{1+\exp(a_n w^T \phi(x_n))}$

f) $\pi(-a_n|x_n) = 1 - \pi(a_n|x_n)$

$$H = \nabla_w^2 \hat{L}_N(\pi) = \nabla_w g$$

$$= \nabla_w \left[\frac{1}{N} \sum_{n=1}^{N} a_n \phi(x_n) \left(\pi(a_n|x_n) - 1 \right) \right]$$

$$= \frac{1}{N} \sum_{n=1}^{N} a_n \phi(x_n) \frac{1}{1+\exp(-a_n w^T \phi(x_n))}$$

rule d
$$= \frac{1}{N} \sum_{n=1}^{N} a_n \phi(x_n) \frac{-(1+\exp(-a_n w^T \phi(x_n)))'}{(1+\exp(-a_n w^T \phi(x_n)))^2}$$

$$= \frac{1}{N} \sum_{n=1}^{N} a_n \phi(x_n) \frac{-\exp(-a_n w^T \phi(x_n)) \cdot a_n \phi(x_n)}{(1+\exp(-a_n w^T \phi(x_n)))^2}$$

rule a
$$= \frac{1}{N} \sum_{n=1}^{N} \frac{a_n \phi(x_n) \cdot \exp(-a_n w^T \phi(x_n))}{(1+\exp(-a_n w^T \phi(x_n)))^2} \cdot \frac{\exp(a_n w^T \phi(x_n))}{\exp(a_n w^T \phi(x_n))}$$

rule b
$$= \frac{1}{N} \sum_{n=1}^{N} \phi(x_n) \phi^T(x_n) \frac{1}{1+\exp(-a_n w^T \phi(x_n))} \cdot \frac{1}{e(a_n w^T \phi(x_n))+1}$$

rule e, f
$$= \frac{1}{N} \sum_{n=1}^{N} \phi(x_n) \phi^T(x_n) \pi(a_n|x_n) (1-\pi(a_n|x_n)) \quad \square.$$

Scanned by CamScanner