Homework 3 模型

陈久宁

2018 年 12 月 8 日

* **预测模型 (forward)**

$$\hat{y} = \phi(W^T x) \tag{1}$$

其中 $W \in \mathbb{R}^{3\times 1}$, $\phi(x) = 1/(1 + e^{-x})$.

* **误差计算 (forwardLoss)**

$$
\begin{aligned}
L &= \ell_2(\hat{Y}, Y) \\
&= \ell_2(\phi(W^T X), Y) \\
&= \frac{1}{2} \left\| \phi(W^T X) - Y \right\|_2^2 \\
&= \frac{1}{2n} \sum_{i=1}^n (\phi(W^T X_i) - Y_i)^2
\end{aligned}
\tag{2}
$$

* **梯度计算 (backwardLoss)**

$$
\begin{aligned}
\nabla_W L &= \frac{1}{n} \sum_{i=1}^n \frac{\partial(\phi(W^T X_i) - Y_i)}{\partial W}(\phi(W^T X_i - Y_i) \\
&= \frac{1}{n} \sum_{i=1}^n \frac{\partial(\phi(W^T X_i))}{\partial W^T X_i} \frac{\partial W^T X_i}{\partial W}(\phi(W^T X_i - Y_i)
\end{aligned}
\tag{3}
$$

其中

$$\frac{\partial \phi(x)}{\partial x} = \frac{e^x}{(1 + e^x)^2} \in \mathbb{R}$$

$$\frac{\partial W^T x}{\partial W} = \frac{\partial x^T W}{\partial W} = x \in \mathbb{R}^{3\times 1}$$

* 更新权重 (update)

$$W = W - t\nabla_W L$$

其中 t 为学习率 (LearnRate)

* 梯度下降 (trainNetwork)

Data: $X \in \mathbb{R}^{3 \times n}, Y \in \mathbb{R}^{1 \times n}$

Result: $W \in \mathbb{R}^{3 \times 1}$

init: $W = $ initializeNetwork()

while *epoch < MaxEpoch* **do**

    $\hat{Y} = $ forward(X, W)

    $\nabla_W L = $ backwardLoss$(X, Y, \hat{Y}, W)$

    $W = $ update$(W, \nabla_W L)$

end