ML Assignment 4

10185101210 陈俊潼

3.7 解：若 $y_i = \pm 1$，令 w 为参数，$\delta(z) = \sigma(w^T x + b)$ 为 sigmoid 函数

∴ 似然函数为 $p(y_i|x_i)w = \sigma(y_i z_i)$，$y_i = -1/1$，$z_i = w^T x_i$

∴ $\underset{w}{\arg\min}\ l(w) y_i x_i$

$= \underset{w}{\arg\min} \sum_{i=1}^{N} \log \sigma(y_i z_i)$

$= \underset{w}{\arg\min} -\sum_{i=1}^{N} \log \frac{1}{\sigma(y_i z_i)} = \underset{w}{\arg\min} -\sum_{i=1}^{N} \log(1+\exp(-$

∴ 对数似然函数为 $-\sum_{i=1}^{N} \log(1+\exp(-y_i(w^T x_i + b)))$

3.8 Sol ∵ $E = \frac{1}{N} \sum_{i=1}^{N} \log(1+\exp(-y_i w^T x_i))$

∴ $\nabla E = -\frac{1}{N} \sum_{i=1}^{N} \frac{y_i x_i \exp(-y_i w^T x_i)}{1+\exp(-y_i w^T x_i)} = -\frac{1}{N} \sum_{i=1}^{N} \frac{y_i x_i}{\exp(y_i w^T x_i)} \cdot \frac{\exp(y_i w^T x_i)}{1+\exp(y_i w^T x_i)}$

$= -\frac{1}{N} \sum_{i=1}^{N} \frac{y_i x_i}{1+\exp(y_i w^T x_i)} = \frac{1}{N} \sum_{n=1}^{N} -y_n x_n \theta(-y_n w^T x_n)$

When y is misclassified, $y_i w^T x_i < 0$, $\theta(-y_n w^T x_n) > 0.5$

While when y is classified, $y_i w^T x_i > 0$, $\theta(-y_n w^T x_n) < 0.5$

So a misclassified example will contribute more gradient.

3.9 解：对感知机而言有误分类次数 $k \le \frac{R^2}{\rho^2}$

其中 $R = \max_n \|x_n\|$，$\rho = \min_n y_n \frac{w_f^T}{\|w_f^T\|} x_n$

对于 T，$R = \max_n \|x_n\| = \sqrt{(-3)^2 + (-9)^2} = 3\sqrt{10}$

$\rho = \frac{w_f^T}{\|w_f^T\|}(-2, 0)$

∴ $k \le \frac{90}{(-2,0)^2} \left(\frac{\|w_f^T\|}{w_f^T}\right)^2 = 22.5 \left(\frac{\|w_f^T\|}{w_f^T}\right)^2$，其中 $\left(\frac{\|w_f^T\|}{w_f^T}\right)^2 < 1$

∴ 最多出现22次误分类。