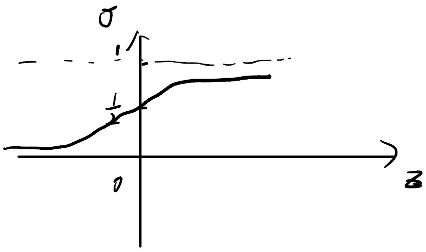
逻辑回归 (Logistic Regression)

Data:

sigmoid : $\sigma(z) = \dfrac{1}{1+e^{-z}}$, $z \to +\infty$, $\lim \sigma(z) = 1$

$z \to 0$, $\sigma(z) = \frac{1}{2}$

$z \to -\infty$, $\lim \sigma(z) = 0$



$\sigma : \mathbb{R} \mapsto (0,1)$

$w^T x \longrightarrow p$

$$\begin{cases} p_1 = P(y=1|x) = \sigma(w^T x) = \dfrac{1}{1+e^{-w^T x}} = \psi(x;w), & y=1 \\ p_2 = P(y=0|x) = 1 - P(y=1|x) = \dfrac{e^{-w^T x}}{1+e^{-w^T x}} \to 1-\psi(x;w), & y=0 \end{cases}$$

$$\frac{\partial \sigma(w^T x)}{\partial w} = \frac{-e^{w^T x}(-x)}{1+e^{-w^T x}}$$

e^{w^T}

$\hookrightarrow P(y|x) = p_1^{y} \cdot p_0^{1-y}$

MLE:
$$\hat{w} = \arg\max_w \log P(Y|X)$$
$$= \arg\max_w \log \prod_{i=1}^{N} P(y_i|x_i)$$
$$= \arg\max_w \sum_{i=1}^{N} \log P(y_i|x_i)$$
$$= \arg\max_w \sum_{i=1}^{N} [y_i \log p_1 + (1-y_i) \cdot \log p_0]$$
$$= \arg\max_w \sum [y_i \cdot \log \psi(w;x) + (1-y_i) \cdot \log[1-\psi(w;x)]]$$

我们通常会选择最小化负对数似然函数作为损失函数

$$J(w) = -L(w) = -\sum_{i=1}^{N} [y_i \cdot \log \psi(x;w) + (1-y_i) \cdot \log(1-\psi(w;x))]$$

梯度下降法更新公式(Gardiant descent)

$$\frac{\partial J(w)}{\partial (w)} = \partial\left[-\sum\left[y_i\cdot\log\left(\frac{1}{1+e^{-w^Tx}}\right)+(1-y_i)\cdot\log\left(1-\frac{1}{1+e^{-w^Tx}}\right)\right]\right]/\partial w$$

$$= -\left[y_i(1-\varphi)\cdot x_i - x_i\cdot\varphi + y_i\cdot x_i\cdot\varphi\right] = -(y_i\cdot x_i - x_i\varphi) = \boxed{\sum x_i(\varphi-y_i)}$$

$$\varphi'(w,x) = \left(\frac{1}{1+e^{-w^Tx}}\right)' = \frac{-e^{-w^Tx}\cdot}{(1+e^{-w^Tx})^2} = \frac{1}{1+e^{-w^Tx}}\cdot\frac{-e^{-w^Tx}}{1+e^{-w^Tx}}\cdot -x = \varphi\cdot(1-\varphi)\cdot x$$

$$[y_i\cdot\log(\varphi)]' = y_i\cdot\frac{1}{\varphi}\cdot\varphi\cdot(1-\varphi)\cdot x = y_i\cdot(1-\varphi)\cdot x_i$$

$$[(1-y_i)\cdot\log(1-\varphi)]' = (1-y_i)\cdot\frac{1}{1-\varphi}\cdot(-\varphi)\cdot(1-\varphi)\cdot x_i = -(1-y_i)\cdot x_i\cdot\varphi$$

梯度下降法参数更新:

$$\theta_{t+1} = \theta_t - \alpha\cdot\frac{\partial J(w)}{\partial w}$$

$$= \theta_t - \alpha\sum_{i=1}^{N} x_i(\varphi - y_i)$$

牛顿法梯度下降

$$H(w) = \frac{\partial^2 J(w)}{\partial w\cdot\partial w^T} = \frac{\partial\sum x_i(\varphi-y_i)}{\partial w^T} = x_i\cdot\varphi' = \sum_{i=1}^{N} x_i\cdot x_i^T\cdot\varphi\cdot(1-\varphi)$$

$$\Rightarrow \theta_{t+1} = \theta_t - \eta\cdot H(w)^{-1}\cdot\nabla J(w)$$

Key: 找到对数似然损失函数，正确求导