逻辑回归

其实我们并不需要确切地算出概率密度，算出决策边界足矣。

名称	模型	损失函数	参数更新方程
感知机	$g(x)=sign(w^Tx+w_0)$	$\ell(w,b)=-r^{(\ell)}(w^Tx^{(\ell)}+b)$	$\boldsymbol{w}\leftarrow\boldsymbol{w}+\eta r^{(\ell)}x^{(\ell)}\text{if } r^{(\ell)}g_{\mathrm{i}}\left(x^{(\ell)}\right)\leq 0$
Sigmoid	$y=\mathrm{sigmoid}\left(w^Tx+w_0\right)=\frac{1}{1+\exp\left[-\left(w^Tx+w_0\right)\right]}$	$\ell\left(w,w_0\mid x,r\right)=-r\log y-(1-r)\log(1-y)$ $L\left(w,w_0\mid D\right)=-\sum_{l=1}^{N}r^{(l)}\log y^{(l)}+\left(1-r^{(l)}\right)\log\left(1-y^{(l)}\right)$	$\boldsymbol{w}\leftarrow\boldsymbol{w}+\sum_{l=1}^{N}\left(r^{(l)}-y^{(l)}\right)x_j^{(l)}$
SoftMax	$y_i=\mathrm{softmax}\left(\boldsymbol{w}_{\mathrm{i}}^{\mathrm{T}}x+w_{\mathrm{i}0}\right)=\frac{\exp\left[\boldsymbol{w}_i{}^Tx+w_{i0}\right]}{\sum_{j=1}^{K}\exp\left[\boldsymbol{w}_j{}^Tx+w_{j0}\right]}\quad(i=1,\ldots,K)$	$L(\boldsymbol{w}\mid\boldsymbol{x},\boldsymbol{r})=-\sum_{i=1}^{K}r_i\log y_i$ $L(\boldsymbol{w}\mid D)=-\sum_{l=1}^{N}\left[\sum_{i=1}^{K}r_i^{(l)}\log y_i^{(l)}\right]$	$\boldsymbol{w}\leftarrow\boldsymbol{w}+\sum_{l=1}^{N}\left(r^{(l)}-y^{(l)}\right)x_j^{(l)}$

分析

- 感知机：
    - 对训练数据集要求较高，只能处理线性可分的数据集，并且解不唯一，存在多个解。
    - 0-1这样的离散值不能精确表示一个实例有多接近于某个类C_i