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	$g(x) = sign(w^Tx + w_0)$	$\ell(w,b) = -r^{(\ell)}(w^Tx^{(\ell)} + b)$	$oldsymbol{w} \leftarrow oldsymbol{w} + \eta r^{(\ell)} x^{(\ell)}$ if $r^{(\ell)} g_{\mathrm{i}} \left(x^{(\ell)} ight) \leq 0$
Sigmoid	$y = \operatorname{sigmoid} \left(w^T x + w_0 \right)$ $= \frac{1}{1 + \exp\left[- \left(w^T x + w_0 \right) \right]}$	$\ell\left(w, w_0 \mid x, r ight) = -r \log y - (1 - r) \log\left(1 - y ight) \ L\left(w, w_0 \mid D ight) = -\sum_{l=1}^N r^{(l)} \log y^{(l)} + \left(1 - r^{(l)} ight) \log\left(1 - y^{(l)} ight)$	$oldsymbol{w} \leftarrow oldsymbol{w} + \sum_{l=1}^{N} ig(r^{(l)} - y^{(l)}ig) x_{j}^{(l)}$
SoftMax	$egin{aligned} y_i &= \operatorname{softmax}\left(oldsymbol{w}_i^{\mathrm{T}}x + w_{i0} ight) \ &= rac{\exp\left[oldsymbol{w}_i^{T}x + w_{i0} ight]}{\sum_{j=1}^{K} \exp\left[oldsymbol{w}_j^{T}x + w_{j0} ight]} (i=1,\ldots,K) \end{aligned}$	$egin{align} L(oldsymbol{w} \mid oldsymbol{x}, oldsymbol{r}) &= -\sum_{i=1}^K r_i \log y_i \ L(oldsymbol{w} \mid D) &= -\sum_{l=1}^N \left[\sum_{i=1}^K r_i^{(l)} \log y_i^{(l)} ight] \end{aligned}$	$oldsymbol{w} \leftarrow oldsymbol{w} + \sum_{l=1}^{N} ig(r^{(l)} - y^{(l)}ig) x_{j}^{(l)}$

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