多分类问题的交叉熵

杨航锋

在多分类问题中，损失函数为交叉熵损失函数。假设数据集 $T=\{(x^{(1)},y^{(1)}),(x^{(2)},y^{(2)}),\cdots,(x^{(m)},y^{(m)})\}$，输入空间 $\mathcal{X}\in\mathbb{R}^n$ 为 n 维向量的集合，特征向量 $x^{(i)}\in\mathcal{X}$，标签 $y^{(i)}\in\mathbb{R}^k$ 为 k 维 $One\ Hot$ 向量，模型的预测输出为 $p^{(i)}$ 其中 $p^{(i)}\in\mathbb{R}^k$，综上该数据集的损失函数为

$$Loss(Y,P)=-\sum_{i=1}^{m}\sum_{j=1}^{k}y_j^{(i)}\log p_j^{(i)}$$

举例说明

$y^{(1)}=[1,0,0]$
$y^{(2)}=[0,1,0]$
$p^{(1)}=[0.8,0.1,0.1]$
$p^{(2)}=[0.3,0.6,0.1]$
$Loss(y,p)=-(1\times\log 0.8+0\times\log 0.1+0\times\log 0.1)-(0\times\log 0.3+1\times\log 0.6+0\times\log 0.1)$