Naive Bayes and Logistic Regression

Xi Tan (tan19@purdue.edu)

October 30, 2017

Contents

1 Logistic Regression

The logistic function is an important special case of sigmoid functions, a family of functions that are "S"-shaped. It is defined as:

$$\sigma(t) := \frac{1}{1+e^{-t}}, \quad t \in \mathbb{R} \tag{1}$$

For binary classification, we can rewrite

$$\begin{aligned} P_{\mathbf{W}}(\mathcal{C}=1|\mathbf{x}) &:= \sigma((\mathbf{w}_1 - \mathbf{w}_2)^T \mathbf{x}) && (2) \\ &= \frac{1}{1+\exp(-(\mathbf{w}_1 - \mathbf{w}_2)^T \mathbf{x})} && (3) \\ &= \frac{\exp(\mathbf{w}_1^T \mathbf{x})}{\exp(\mathbf{w}_1^T \mathbf{x}) + \exp(\mathbf{w}_2^T \mathbf{x})} && (4) \end{aligned}$$

where the class weight matrix $\mathbf{W}$ has two columns $\mathbf{w}_1$ and $\mathbf{w}_2$, the weight vector for class 1 and 2, respectively. The softmax function generalizes the logistic function to multiple classes:

$$P_{\mathbf{W}}(\mathcal{C}=i|\mathbf{x}) := \frac{\exp(\mathbf{w}_i^T \mathbf{x})}{\sum_{k=1}^{K} \exp(\mathbf{w}_k^T \mathbf{x})} \tag{5}$$

where $\mathbf{W}$ has K columns, each of which is the corresponding weight vector.