Chaudhary ML, PS # 4

Ashwin Acharya

May 4, 2016

1 Problem 3

Consider a k-class logistic regression model in which we fit vectors θ_j $(j = 1, 2, ..., k-1)$ under the assumption that $ln\frac{\pi_j(x)}{\pi_k(x)} = \theta_j^T x$.

1.1 3a

From the above, we know that $\pi_j(x) = \pi_k(x) * e^{\theta_j^T x}$

We know that the class probabilities for x sum to 1, so that $\pi_k(x) = 1 - \Sigma_{j=1}^{k-1}\pi_j(x) = 1 - \Sigma_{j=1}^{k-1}\pi_k(x) * e^{\theta_j^T x} = 1 - \pi_k(x)\Sigma_{j=1}^{k-1}e^{\theta_j^T x} = 1 - \pi_k(x)S$.

Therefore, $\pi_k(x)(1+S) = 1$, so $\mathbf{\pi_k(x) = \frac{1}{1+S} = \frac{1}{1+\Sigma_{j=1}^{k-1}e^{\theta_j^T x}}}$.

Then, $\mathbf{\pi_j(x) = \pi_k(x) * e^{\theta_j^T x} = \frac{e^{\theta_j^T x}}{1+\Sigma_{j=1}^{k-1}e^{\theta_j^T x}}}$.

1.2 3b

For a logistic regression model as detailed above, label the zero-vector $\vec{0}$ as θ_k, so that $e^{theta_k^T x}$ always equals 1. Then, for all classes $j \in \{1, .., k\}$, $P(j|\text{model,x}) = \frac{e^{\theta_j^T x}}{\Sigma_{j=1}^{k} e^{\theta_j^T x}}$. Thus, the likelihood of a given input-output pair (x, y), $P((x,y)| < \theta_1, \theta_2, ...\theta_k >)$, equals $P(y|\text{model}, x)$, which equals $\frac{e^{\theta_y^T x}}{\Sigma_{j=1}^{k} e^{\theta_j^T x}}$.

The likelihood of the dataset is the product of the likelihoods of each of the results. This equals $\Pi_i \frac{e^{\theta_{y_i}^T x_i}}{\Sigma_{j=1}^{k} e^{\theta_j^T x_i}} = \mathbf{\Pi_i \frac{e^{\theta_{y_i}^T x_i}}{1+\Sigma_{j=1}^{k-1} e^{\theta_j^T x_i}}}$. (Where, again, θ_k is defined as the zero-vector $\vec{0} =< 0, 0, ..., 0 >$.)