Lecture 16
Solving GLMs via IRWLS

09 November 2015

Taylor B. Arnold
Yale Statistics
STAT 312/612

Notes

- problem set 5 posted; due next class

Notes

- problem set 5 posted; due next class
- problem set 6, November 18th

Goals for today

- fixed PCA example from last time
- how to solve logistic regression via weighted least squares
- classification problem on image corpus

fixed PCA example from last time

GLMs

Recall that we define generalized linear models such that the mean of y is some function of $X\beta$, rather than directly equal to it:

$$\mathbb{E}(y|X) = g^{-1}(X\beta)$$

With g, called the *link function*, equal to some fixed and known function.

GLMs

Recall that we define generalized linear models such that the mean of y is some function of $X\beta$, rather than directly equal to it:

$$\mathbb{E}(y|X) = g^{-1}\left(X\beta\right)$$

With g, called the *link function*, equal to some fixed and known function.

Logistic regression

The logistic regression function uses g equal to the logit function to describe a distribution with $y \in \{0, 1\}$. Specifically, we have the following description of the statistical model:

$$\begin{aligned} \mathbb{E}(y|X) &= \operatorname{logit}^{-1}(X\beta) \\ &= \frac{1}{1 + e^{-X\beta}} \end{aligned}$$

Now that we have discussed how to solve the ordinary least squares equation, you may wonder how we would go about solving the logistic regression problem.

Recall that the likelihood of the data point y_i given its mean p_i is equal to:

$$\begin{aligned} L_i(y_i|p_i) &= \exp\left\{ y_i \cdot \log\left(\frac{p_i}{1-p_i}\right) + \log(1-p_i) \right\} \\ &= \exp\left\{ y_i \cdot \theta_i + \frac{e^{\theta_i}}{e^{\theta_i}+1} \right\} \\ &= \exp\left\{ y_i \cdot \theta_i - b(\theta_i) \right\} \end{aligned}$$

Where:

$$\begin{aligned} \theta_i &= \log\left(\frac{p_i}{1-p_i}\right) \\ b(\theta_i) &= -1 \cdot \frac{e^{\theta_i}}{e^{\theta_i}+1} \end{aligned}$$

The derivative of the log-likelihood is then given by:

$$\frac{\partial}{\partial \theta_i} l_i(y_i|\theta_i) = y_i - b'(\theta_i)$$

And

$$\frac{\partial \theta_i}{\partial p_i} = \frac{1 - p_i}{p_i} \cdot \frac{1}{1}$$

Recall that the log-likelihood of the data point y_i given its mean p_i is equal to:

$$l_i(y_i|p_i) = \log(1 - p_i) + y_i \cdot \log\left(\frac{p_i}{1 - p_i}\right)$$

Recall that the log-likelihood of the data point y_i given its mean p_i is equal to:

$$l_i(y_i|p_i) = \log(1-p_i) + y_i \cdot \log\left(\frac{p_i}{1-p_i}\right)$$

Taking the derivative with respect to p_i yields:

$$\frac{\partial}{\partial p_i} l_i(y_i|p_i) = \frac{1}{p_i - 1} + y_i \cdot \frac{1-p_i}{p_i} \cdot$$