Lecture 05

Dihui Lai

dlai@wustl.edu

October 6, 2020

Generalized Linear Model

Exponential family of probability density function

$$f(y) = \text{exp}\left(\frac{y\theta - b(\theta)}{a(\phi)} + c(y, \phi)\right)$$

The distribution have the following properties

- $E(Y) = b'(\theta)$
- $Var(Y) = b''(\theta)a(\phi)$

Link Function

Assume a linear model that models the mean of the distribution

$$b'(\theta) = g(\eta) = g(\vec{x} \cdot \vec{\beta})$$

here $g^{-1}(\cdot)$ is called the link function.

Log Likelihood Function of GLM

The log likelihood function of GLM

$$\ell = \sum_i \frac{y^i\theta^i - b(\theta^i)}{a(\phi)} + c^i(y^i, \phi)$$

In the model, only θ is depedent on $\vec{x} \cdot \vec{\beta}$. Therefore, "maximize" the likelihood function is equivalent to minimize

$$\ell = -2\sum_i \left[y^i\theta^i - b(\theta^i)\right]$$

Examples of GLM

- Normal: $\log f(y_i, \theta_i, \phi) = -\frac{(y_i - \mu_i)^2}{2\phi} + C$
  Canonical link function
  $$\mu_i = \vec{x^i} \cdot \vec{\beta}$$
- Poisson: $\log f(y_i, \theta_i, \phi) = y_i \log(\lambda_i) - \lambda_i + C$
  Canonical link function
  $$\log(\lambda_i) = \vec{x^i} \cdot \vec{\beta}$$
- Binomial: $\log f(y_i, \theta_i, \phi) = y_i \log(\frac{p_i}{1-p_i}) - \log(1 - p_i) + C$
  Canonical link function
  $$\log(\frac{p_i}{1 - p_i}) = \vec{x^i} \cdot \vec{\beta}$$

The Gradient and Hessian; Optimization quantity

The gradient can be derived as

$$\frac{\partial \ell}{\partial \beta_j} = -2\sum_i (y^i - b'(\theta^i))\frac{\partial \theta^i}{\partial \beta_j}$$

$$\frac{\partial \ell}{\partial \beta_j} = -2\sum_i (y^i - \mu^i)\frac{g'(\eta^i)}{V(\mu^i)} x_j^i$$

The hessian can be derived as

$$\frac{\partial^2 \ell}{\partial \beta_k \partial \beta_j} = 2\sum_i \left[\frac{g'(\eta^i)^2}{V(\mu^i)} - (y^i - \mu^i)\frac{g''(\eta^i)V(\mu^i) - g'(\eta^i)^2 V'(\mu^i)}{V(\mu^i)^2}\right] x_j^i x_k^i$$