538 Lec. #3

GLM

- How G ?
- What is "under the hood"?

Prelim: Exponential family

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

Moments: $E(Y) = b'(\theta)$

$$\mathrm{Var}(Y) = a(\phi)\, b''(\theta)$$

Defines variance function $v(\,)$ such that

$$\mathrm{Var}(Y) = a(\phi)\, v\{E(Y)\}$$

Prelim: link function

Assume $\underset{\sim}{X}$ influences Y via

linear predictor $\eta = \beta_0 + \beta_1 X_1 + \ldots + \beta_{p-1} X_{p-1}$

How are η and $\mu \equiv E(Y)$ related?

Choose link function $g()$ such that

$$\eta = g(\mu)$$

Canonical link?

Prelim: Fisher Scoring

Applying Newton-Raphson as iterative scheme to solve $\boxed{\ell'(\underset{\sim}{\beta}) = \underset{\sim}{0}}$

$$\hat{\underset{\sim}{\beta}}^{next} = \hat{\underset{\sim}{\beta}}^{prev} + \left[-\ell''(\hat{\underset{\sim}{\beta}}^{prev})\right]^{-1} \ell'(\hat{\underset{\sim}{\beta}}^{prev})$$

FISHER SCORING MODIFICATION?

Prelim / Reminder: Weighted LS

$$\hat{\underset{\sim}{\beta}} = \underset{\underset{\sim}{\beta}}{\operatorname{argmin}} \; (\underset{\sim}{y} - X\underset{\sim}{\beta})^T W (\underset{\sim}{y} - X\underset{\sim}{\beta})$$

$$= (X^T W X)^{-1} X^T W \underset{\sim}{y}$$

e.g. arises if $Y_i \sim N(\underset{\sim}{x}_i^T \underset{\sim}{\beta}, \sigma_i^2)$

with $W = \begin{pmatrix} 1/\sigma_1^2 & & 0 \\ & \ddots & \\ 0 & & 1/\sigma_n^2 \end{pmatrix}$

NOTE: W only matters up to constant of proportionality

Assembling pieces

A $$\frac{\partial}{\partial \underset{\sim}{\beta}} \log f(y_i ; \underset{\sim}{\beta}) = \frac{y_i - \mu_i}{a(\phi)\, v(\mu_i)\, \partial \eta_i / \partial \mu_i} \underset{\sim}{x}_i$$

B $$E\left\{ - \frac{\partial^2}{\partial \underset{\sim}{\beta}^2} \log f(y_i ; \underset{\sim}{\beta}) \right\} = \frac{1}{a(\phi)\, v(\mu_i) (\partial \eta_i / \partial \mu_i)^2} \underset{\sim}{x}_i \underset{\sim}{x}_i^T$$

C $$E\{ -\ell''(\underset{\sim}{\beta}) \} = \frac{1}{a(\phi)} X^T W_\beta X$$

where $$(W_\beta)_{ii} = \frac{1}{v(\mu_i) (\partial \eta_i / \partial \mu_i)^2}$$

D $$\ell'(\underset{\sim}{\beta}) = \frac{1}{a(\phi)} X^T W_\beta \begin{pmatrix} \partial \eta_1 / \partial \mu_1 \, (y_1 - \mu_1) \\ \vdots \\ \partial \eta_n / \partial \mu_n \, (y_n - \mu_n) \end{pmatrix}$$

So Fisher scoring algorithm is

$$\hat{\underset{\sim}{\beta}}^{(t+1)} = \underset{\sim}{\beta}^{(t)} + \left\{X^T W_{\hat{\beta}(t)} X\right\}^{-1} X^T W_{\hat{\beta}(t)} \frac{\partial \eta}{\partial \mu}(Y - \mu)$$

$$= \left\{X^T W_{(t)} X\right\}^{-1} X^T W_{(t)} Y^*_{(t)}$$

where

$$Y^*_{(t)} = X \hat{\underset{\sim}{\beta}}^{(t)} + \frac{\partial \hat{\eta}}{\partial \underset{\sim}{\mu}} \left(\underset{\sim}{Y} - \underset{\sim}{\mu}\right)$$

Algorithm

- for specified $g(\ )$, $v(\ )$
- inputted X, Y