Maximum Likelihood Estimation via Newton-Raphson: Logistic Regression

$y_i \mid \underline{\beta} \sim \text{Binomial}(N_i, p_i)$ where $\text{logit}(p_i) = \underline{x}_i'\underline{\beta}$

i.e. $\log\left(\frac{p_i}{1-p_i}\right) = \underline{x}_i'\underline{\beta}$

$$\Rightarrow \quad \frac{p_i}{1-p_i} = \exp(\underline{x}_i'\beta)$$

$$p_i = \exp(\underline{x}_i'\underline{\beta}) - \exp(\underline{x}_i'\underline{\beta})\, p_i$$

$$(1+\exp(\underline{x}_i'\beta))\, p_i = \exp(\underline{x}_i'\beta)$$

$$p_i = \frac{\exp(\underline{x}_i'\beta)}{1+\exp(\underline{x}_i'\beta)}$$

$$= \frac{1}{1+\exp(-\underline{x}_i'\beta)}$$

$$= (1+\exp(-\underline{x}_i'\beta))^{-1}$$

$$1-p_i = \frac{\exp(-\underline{x}_i'\underline{\beta})}{1+\exp(-\underline{x}_i'\underline{\beta})}$$

$$= \frac{1}{1+\exp(\underline{x}_i'\underline{\beta})}$$

$$= (1+\exp(x_i\beta))^{-1}$$

$$L(\beta \mid y_1, \ldots, y_N) = \prod_{i=1}^{N} \binom{N_i}{y_i} p_i^{y_i} (1-p_i)^{N_i - y_i}$$

$$\ell(\beta \mid y_1, \ldots, y_N) = \log L(\beta \mid y_1, \ldots, y_N) = \overbrace{\sum_{i=1}^{N} \log \binom{N_i}{y_i}}^{c} + \sum_{i=1}^{N} \left[y_i \log p_i + (N_i - y_i)\log(1-p_i) \right]$$

$$= c + \sum_{i=1}^{N} \left[-y_i \log(1+\exp(-\underline{x}_i'\beta)) - (N_i - y_i)\log(1+\exp(\underline{x}_i'\beta) \right.$$

Let $f(\beta) = \ell(\beta \mid y_1, \dots, y_m)$

$$\nabla f(\beta) = \left[\sum_{i=1}^{N} \left[\frac{y_i x_{i0} \exp(-\underline{x}_i'\beta)}{1+\exp(-\underline{x}_i'\beta)} - \frac{(N_i - y_i) x_{i0} \exp(\underline{x}_i'\beta)}{1+\exp(\underline{x}_i'\beta)} \right], \right.$$

gradient

$$\sum_{i=1}^{N} \left[\frac{y_i x_{i1} \exp(-\underline{x}_i'\beta)}{1+\exp(-x_i'\beta)} - \frac{(N_i - y_i) x_{i1} \exp(x_i'\beta)}{1+\exp(x_i'\beta)} \right],$$

$$\dots,$$

$$\left. \sum_{i=1}^{N} \left[\frac{y_i x_{ik} \exp(-\underline{x}_i'\beta)}{1+\exp(-\underline{x}_i'\beta)} - \frac{(N_i - y_i) x_{ik} \exp(x_i'\beta)}{1+\exp(\underline{x}_i'\beta)} \right] \right]$$

$$= \left[\sum_{i=1}^{N} \left[\frac{y_i x_{i0}}{1+\exp(\underline{x}_i'\beta)} - \frac{(N_i - y_i) x_{i0}}{1+\exp(-\underline{x}_i'\beta)} \right], \right.$$

$$\sum_{i=1}^{N} \left[\frac{y_i x_{i1}}{1+\exp(x_i'\beta)} - \frac{(N_i - y_i) x_{i1}}{1+\exp(-x_i'\beta)} \right],$$

$$\dots,$$

$$\left. \sum_{i=1}^{N} \left[\frac{y_i x_{ik}}{1+\exp(x_i'\beta)} - \frac{(N_i - y_i) x_{ik}}{1+\exp(-x_i'\beta)} \right] \right]$$

$D^2 f(\beta) =$ " $(k+1) \times (k+1)$ matrix whose a,b element is

Hessian matrix

$$\sum_{i=1}^{N} \left[\frac{-y_i x_{ia} x_{ib} \exp(\underline{x}_i'\beta)}{(1+\exp(\underline{x}_i'\beta))^2} - \frac{(N_i - y_i) x_{ia} x_{ib} \exp(-x_i'\beta)}{(1+\exp(-\underline{x}_i'\beta))^2} \right]$$