NMFM 402, Practical 3, GLM - param. ~~est~~ estim.

Ex. 1 (a) recall the equations for MLE:

$$\sum_{i=1}^{n}\left[y_i x_{ij} - \exp\left(\sum_{l=0}^{q} x_{il}\beta_l\right)\cdot x_{ij}\right] = 0, \quad j=0,\dots,q \quad (*)$$

consider the mapping $F: \mathbb{R}^{q+1} \to \mathbb{R}^{q+1}$ defined by

$$F(\beta_0,\dots,\beta_q) = \begin{pmatrix} f_0(\beta_0,\dots,\beta_q) \\ \dots \\ f_q(\beta_0,\dots,\beta_q) \end{pmatrix}, \text{ where}$$

$$\cancel{f_j(\beta_0,\dots,\beta_q) = \sum_{i=1}^{n}\left[y_i x_{ij} - \exp\left(\sum_{l=0}^{q} x_{il}\beta_l\right)\right.}$$

$$f_j(\beta_0,\dots,\beta_q) = \sum_{i=1}^{n}\left[y_i x_{ij} - \exp\left(\sum_{l=0}^{q} x_{il}\beta_l\right)\cdot x_{ij}\right], \quad j=0,\dots,q,$$

so that equation (*) can be written as

$$F(\beta_0,\dots,\beta_q) = \begin{pmatrix} 0 \\ \dots \\ 0 \end{pmatrix}.$$

Calculate Jacobian matrix of F:

$$J_F(\beta_0,\dots,\beta_q) = \begin{pmatrix} \frac{\partial f_0}{\partial \beta_0}, & \dots, & \frac{\partial f_0}{\partial \beta_q} \\ \dots & & \\ \frac{\partial f_q}{\partial \beta_0}, & \dots, & \frac{\partial f_q}{\partial \beta_q} \end{pmatrix}, \quad \text{where}$$

$$\frac{\partial f_j}{\partial \beta_r} = \sum_{i=1}^{n} -\exp\left(\sum_{l=0}^{q} x_{il}\beta_l\right)\cdot x_{ir}\cdot x_{ij}$$

in result, the update rule is

$$\underline{\beta}^{[n+1]} = \underline{\beta}^{[n]} - J_F^{-1}(\underline{\beta}^{[n]}) \cdot F(\underline{\beta}^{[n]}), \quad \text{where}$$

$\underline{\beta}^{[n]} = (\beta_0^{[n]}, \dots, \beta_2^{[n]})$ is the approximation of the solution in n-th step and J_F and F are defined above.

(b) Recall the update rule from lecture notes

$$\underline{\beta}^{[n+1]} = \underline{\beta}^{[n]} - (H^{-1})^{[n]} \cdot (X' \cdot W^{[n]} \cdot y - X' W^{[n]} \mu^{[n]})$$

and calculate

$$W^{[n]} = \operatorname{diag}\{w_{ii}^{[n]}\}, \quad \text{where}$$

$$w_{ii}^{[n]} = \frac{w_i}{v(\mu_i^{[n]}) \cdot g'(\mu_i^{[n]})} = \frac{1}{\mu_i^{[n]} \cdot \frac{1}{\mu_i^{[n]}}} = 1$$

see Practical 2

$$\mu^{[n]} = \begin{pmatrix} \exp(\sum_{l=0}^{2} x_{1l}\beta_l) \\ \dots \\ \exp(\sum_{l=0}^{2} x_{nl}\beta_l) \end{pmatrix}$$

and so $X' \cdot W^{[n]} \cdot y = X' \cdot y = \begin{pmatrix} \sum_{i=1}^{n} x_{i0} y_i \\ \dots \\ \sum_{i=1}^{n} x_{i2} y_i \end{pmatrix}$ and

$$X' \cdot W^{[n]} \cdot \mu^{[n]} = X' \cdot \mu^{[n]} = \begin{pmatrix} \sum_{i=1}^{n} x_{i0} \cdot \exp(\sum_{l=0}^{2} x_{il}\beta_l^{[n]}) \\ \dots \\ \sum_{i=1}^{n} x_{i2} \cdot \exp(\sum_{l=0}^{2} x_{il}\beta_l^{[n]}) \end{pmatrix}$$

altogether, we have

$$X' \cdot W^{[n]} \cdot y - X' \cdot W^{[n]} \cdot \mu^{[n]} = \begin{pmatrix} \sum_{i=1}^{n} y_i x_{i0} - \exp\left(\sum_{l=0}^{q} x_{il} \beta_l^{[n]}\right) \cdot x_{i0} \\ \cdots \\ \sum_{i=1}^{n} y_i x_{iq} - \exp\left(\sum_{l=0}^{q} x_{il} \beta_l^{[n]}\right) \cdot x_{iq} \end{pmatrix}$$

Now calculate element of matrix $H = \begin{pmatrix} \frac{\partial^2 l}{\partial \beta_0 \partial \beta_0} & \cdots & \frac{\partial^2 l}{\partial \beta_0 \partial \beta_q} \\ \cdots & & \\ \frac{\partial^2 l}{\partial \beta_q \partial \beta_0} & \cdots & \frac{\partial^2 l}{\partial \beta_q \partial \beta_q} \end{pmatrix}$,

being the Hessian matrix of ~~log likelihood~~ log-likelihood function ~~&~~ $l(\beta_0 \ldots, \beta_q)$.

By (2.16) and above, we have

$$\frac{\partial^2 l}{\partial \beta_r \partial \beta_j} = -\sum_{i=1}^{n} \frac{w_i}{\varphi} \cdot a_i \, x_{ij} \, x_{ir} \,, \quad \text{where}$$

$w_i = 1$, $\varphi = 1$, and

$$a_i = \frac{1}{v(\mu_i) \cdot [g'(\mu_i)]^2} \cdot \left[1 + (y_i - \mu_i) \cdot \frac{v(\mu_i) \cdot g''(\mu_i) + v'(\mu_i) \cdot g'(\mu_i)}{v(\mu_i) \cdot g'(\mu_i)} \right]$$

recall $v(\mu_i) = \mu_i$, $g(\mu_i) = \log(\mu_i)$, $g'(\mu_i) = \frac{1}{\mu_i}$, $g''(\mu_i) = -\frac{1}{\mu_i^2}$

$v'(\mu_i) = 1$

and so $v(\mu_i) \cdot g''(\mu_i) + v'(\mu_i) \cdot g'(\mu_i) = \mu_i \cdot \left(-\frac{1}{\mu_i^2}\right) + 1 \cdot \frac{1}{\mu_i} = 0$

and $a_i = \dfrac{1}{v(\mu_i) \cdot [g'(\mu_i)]^2} = \dfrac{1}{\mu_i \cdot \left(\frac{1}{\mu_i}\right)^2} = \mu_i = \exp\left(\sum_{l=0}^{q} x_{il} \beta_l\right)$

In result, we have

$$\frac{\partial^2 l}{\partial \beta_r \partial \beta_j} = -\sum_{i=1}^{n} \exp\left(\sum_{l=0}^{q} x_{il}\beta_l\right) x_{ij} x_{ir}$$

This concludes the update rule for Newton-Raphson.

Now compare (a) and (b):

$$X' W^{[m]} y - X' W^{[m]} \mu^{[m]} = F\left(\beta^{[m]}\right)$$

$$H^{[m]} = J_F\left(\beta^{[m]}\right)$$

so that we have same update rules in (a) and (b).

(c) We substitute H by its expectation $-J = EH$ (J is Fisher information matrix). The elements of $-J$ are

$$-J = \left(\sum_{i=1}^{n} \frac{w_i}{\varphi} \cdot \left(- \frac{1}{v(\mu_i) \cdot [g'(\mu_i)]^2} \right) \cdot x_{ir} x_{ij} \right)_{r, j = 0, \dots, q}$$

In our case, we have

$$\sum_{i=1}^{n} \frac{w_i}{\varphi} \left(- \frac{1}{v(\mu_i) \cdot [g'(\mu_i)]^2} \right) \cdot x_{ir} \cdot x_{ij} = -\sum_{i=1}^{n} \exp\left(\sum_{l=0}^{q} x_{il}\beta_l\right) x_{ij} x_{ir} =$$

$$= \frac{\partial^2 l}{\partial \beta_r \partial \beta_j} \Longrightarrow -J = H$$ and the Fisher method of scores coincides with the Newton-Raphson algorithm (NOTE: this is always true if we chose canonical link function)