QuantCo Hessian Computations

Michael Yin

June 2020

1 Link Functions

Note: f denotes the link function, while g denotes its inverse

Identity:

$$\eta_i = f(\mu_i) = \mu_i$$

$$\mu_i = g(\eta_i) = \eta_i$$

$$g'(\eta_i) = 1$$

$$g''(\eta_i) = 0$$

Log:

$$\eta_i = f(\mu_i) = \ln(\mu_i)$$

$$\mu_i = g(\eta_i) = e_i^\eta$$

$$g'(\eta_i) = e_i^\eta = \mu_i$$

$$g''(\eta_i) = e_i^\eta = \mu_i$$

Logit:

$$\eta_i = f(\mu_i) = \ln\left(\frac{\mu_i}{1-\mu_i}\right)$$

$$\mu_i = g(\eta_i) = \frac{1}{1+e^{-\eta_i}}$$

$$g'(\eta_i) = \frac{e^{-\eta_i}}{(1+e^{-\eta_i})^2} = \mu(1-\mu)$$

$$g''(\eta_i) = \frac{-e^{-\eta_i}(1+e^{-\eta_i})^2 + 2e^{-2\eta_i}(1+e^{-\eta_i})}{(1+e^{-\eta_i})^4} = -\frac{e^{-\eta_i}}{(1+e^{-\eta_i})^2} + \frac{2e^{-2\eta_i}}{(1+e^{-\eta_i})^3}$$

$$= -\mu_i(1-\mu_i) + 2\mu_i(1-\mu_i)^2 = \mu_i(1-\mu_i)(2(1-\mu_i)-1) = \mu_i(1-\mu_i)(1-2\mu_i)$$

2 Distributions

Normal ($\sigma = 1$):

$$L(\mu_i) = \frac{1}{\sqrt{2\pi}} \cdot e^{-\frac{1}{2}(y_i-\mu_i)^2} \to e^{-\frac{1}{2}(y_i-\mu_i)^2}$$

$$LL(\mu_i) = -\frac{1}{2}\left(y_i - \mu_i\right)^2$$

$$LL'(\mu_i) = y_i - \mu_i$$

$$LL''(\mu_i) = -1$$

$$v_i^{-1} = \frac{1}{\mu_i^0} = 1$$

Poisson ($\mu = \lambda$):

$$L(\mu_i) = \frac{\mu_i^{y_i} \cdot e^{-\mu_i}}{y_i!} \to \mu_i^{y_i} \cdot e^{-\mu_i}$$

$$LL(\mu_i) = y_i \ln \mu_i - \mu_i$$

$$LL'(\mu_i) = \frac{y_i}{\mu_i} - 1$$

$$LL''(\mu_i) = -\frac{y_i}{\mu_i^2}$$

$$v_i^{-1} = \frac{1}{\mu_i}$$

Gamma ($\mu = \theta, k = 1$):

$$L(\mu_i) = \frac{1}{\mu_i} \cdot e^{-\frac{y_i}{\mu_i}}$$

$$LL(\mu_i) = -\ln \mu_i - \frac{y_i}{\mu_i}$$

$$LL'(\mu_i) = -\frac{1}{\mu_i} + \frac{y_i}{\mu_i^2}$$

$$LL''(\mu_i) = \frac{1}{\mu_i^2} - \frac{2y_i}{\mu_i^3}$$

$$v_i^{-1} = \frac{1}{\mu_i^2}$$

Binomial ($\mu = p, n = 1$):

$$L(\mu_i) = \mu_i^{y_i} \cdot (1-\mu_i)^{1-y_i}$$

$$LL(\mu_i) = y_i \ln \mu_i + (1-y_i)\ln(1-\mu_i)$$

$$LL'(\mu_i) = \frac{y_i}{\mu_i} - \frac{1-y_i}{1-\mu_i}$$

$$LL''(\mu_i) = -\frac{y_i}{\mu_i^2} - \frac{1-y_i}{(1-\mu_i)^2}$$

$$v_i^{-1} = \frac{1}{\mu_i(1-\mu_i)}$$

3 Formulas

3.1 Gradient

True Gradient:

$$G_k = \frac{\partial LL}{\partial \beta_k} = \sum_i LL'_i \cdot g'_i \cdot x_{ik}$$

QuantCo Gradient:

$$QG_k = \sum_i g'_i \cdot v_i^{-1} \cdot (y_i - \mu_i) \cdot x_{ik}$$

3.2 Hessian

True Hessian:

$$H_{kj} = \frac{\partial^2 LL}{\partial \beta_k \partial \beta_j} = \sum_i [LL''_i \cdot (g'_i)^2 + LL'_i \cdot g''_i] \cdot x_{ik} x_{ij}$$

Gauss-Newton Hessian:

$$GN_{kj} = \sum_i [LL''_i \cdot (g'_i)^2] \cdot x_{ik} x_{ij}$$

QuantCo Hessian:

$$QH_{kj} = \sum_i -(g'_i)^2 \cdot v_i^{-1} \cdot x_{ik} x_{ij}$$

4 Calculations

4.1 Normal, Identity

Gradients:

$$G_k = QG_k = \sum_i (y_i - \mu_i) \cdot x_{ik}$$

Hessians:

$$H_{kj} = GN_{kj} = QC_{kj} = -\sum_i x_{ik}x_{ij}$$

4.2 Poisson, Log

Gradients:

$$G_k = QG_k = \sum_i (y_i - \mu_i) \cdot x_{ik}$$

True Hessian:

$$H_{kj} = \sum_i \left[-\frac{y_i}{\mu_i^2} \cdot \mu_i^2 + \left(\frac{y_i}{\mu_i} - 1 \right) \mu_i \right] \cdot x_{ik}x_{ij} = -\sum_i \mu_i \cdot x_{ik}x_{ij}$$

Gauss-Newton Hessian:

$$GN_{kj} = \sum_i -\frac{y_i}{\mu_i^2} \cdot \mu_i^2 \cdot x_{ik}x_{ij} = -\sum_i y_i \cdot x_{ik}x_{ij}$$

QuantCo Hessian:

$$QH_{kj} = \sum_i -\mu_i^2 \cdot \frac{1}{\mu_i} \cdot x_{ik}x_{ij} = -\sum_i \mu_i \cdot x_{ik}x_{ij}$$

4.3 Gamma, Log

Gradients:

$$G_k = QG_k = \sum_i \left(\frac{y_i}{\mu_i} - 1 \right) \cdot x_{ik}$$

True Hessian:

$$\begin{aligned} H_{kj} &= \sum_i \left[\left(\frac{1}{\mu_i^2} - \frac{2y_i}{\mu_i^3} \right) \mu_i^2 + \left(-\frac{1}{\mu_i} + \frac{y_i}{\mu_i^2} \right) \mu_i \right] \cdot x_{ik}x_{ij} \\ &= \sum_i \left[1 - \frac{2y_i}{\mu_i} - 1 + \frac{y_i}{\mu_i} \right] \cdot x_{ik}x_{ij} = -\sum_i \frac{y_i}{\mu_i} \cdot x_{ik}x_{ij} \end{aligned}$$

Gauss-Newton Hessian:

$$GN_{kj} = \sum_i \left(\frac{1}{\mu_i^2} - \frac{2y_i}{\mu_i^3} \right) \mu_i^2 \cdot x_{ik}x_{ij} = -\sum_i \left(\frac{2y_i}{\mu_i} - 1 \right) \cdot x_{ik}x_{ij}$$

QuantCo Hessian:

$$QH_{kj} = \sum_i -\mu_i^2 \cdot \frac{1}{\mu_i^2} \cdot x_{ik}x_{ij} = -\sum_i x_{ik}x_{ij}$$

4.4 Binomial, Logit

Gradients:

$$G_k = QG_k = \sum_i (y - \mu) \cdot x_{ik}$$

True Hessian:

$$H_{kj} = \sum_i \left[\left(-\frac{y_i}{\mu_i^2} - \frac{1 - y_i}{(1 - \mu_i)^2} \right) \mu_i^2 (1 - \mu_i)^2 + \left(\frac{y_i}{\mu_i} - \frac{1 - y_i}{1 - \mu_i} \right) \mu_i (1 - \mu_i)(1 - 2\mu_i) \right] \cdot x_{ik} x_{ij}$$

$$= \sum_i [-y_i(1-\mu_i)^2 - (1-y_i)\mu_i^2 + y_i(1-\mu_i)(1-2\mu_i) - (1-y_i) \cdot \mu_i(1-2\mu_i)] \cdot x_{ik} x_{ij}$$

$$= \sum_i [-y_i \mu_i (1 - \mu_i) - (1 - y_i)\mu_i(1 - \mu_i)] \cdot x_{ik} x_{ij} = -\sum_i \mu_i (1 - \mu_i) \cdot x_{ik} x_{ij}$$

Gauss-Newton Hessian:

$$GN_{kj} = \sum_i \left(-\frac{y_i}{\mu_i^2} - \frac{1 - y_i}{(1 - \mu_i)^2} \right) \mu_i^2 (1-\mu_i)^2 \cdot x_{ik} x_{ij} = \sum_i [-y_i(1-\mu_i)^2 - (1-y_i)\mu_i^2] \cdot x_{ik} x_{ij}$$

$$= \sum_i [-y_i + 2y_i\mu_i - y_i\mu_i^2 - \mu_i^2 + y_i\mu_i^2] \cdot x_{ik} x_{ij} = -\sum_i (y_i - 2y_i\mu_i + \mu_i^2) \cdot x_{ik} x_{ij}$$

QuantCo Hessian:

$$QH_{kj} = \sum_i -\mu_i^2 (1 - \mu_i)^2 \cdot \frac{1}{\mu_i (1 - \mu_i)} \cdot x_{ik} x_{ij} = -\sum_i \mu_i (1 - \mu_i) \cdot x_{ik} x_{ij}$$