Ex: $X_1, \dots, X_n \sim \text{Gamma}(\alpha, \beta)$

$$\ell(\alpha, \beta; x) = n\left[\alpha \log \beta - \log \Gamma\alpha + (\alpha-1)\overline{\log x} - \beta \bar{x}\right]$$

when $\Theta = (\alpha, \beta)$ then we want $\hat{\Theta} = (\hat{\alpha}, \hat{\beta})$

(i) First approach via derivatives

$$\frac{\partial \ell}{\partial \alpha} = n\left[\log \beta - \psi(\alpha) + \overline{\log x}\right]$$

$$\frac{\partial \ell}{\partial \beta} = n\left[\frac{\alpha}{\beta} - \bar{x}\right]$$

Solve this system for $\hat{\beta} = \frac{\hat{\alpha}}{\bar{x}}$

but no closed form solution

So we derived the Newton-Ralpson update

$$\alpha^{(t+1)} = \alpha^{(t)} - \frac{g(\alpha^{(t)})}{g'(\alpha^{(t)})}$$

for $g(\alpha) = \log\left(\frac{\hat{\alpha}}{\bar{x}}\right) - \psi(\hat{\alpha}) + \overline{\log x}$

then

$$g'(\alpha) = \frac{1}{\alpha} - \psi'(\alpha)$$

where

$\Gamma(x)$ gamma

$\frac{d}{dx} \log \Gamma(x)$ digamma

$\frac{\partial^2}{\partial x^2} \log \Gamma(x)$ trigamma.

So the update is

$$\alpha^{(t+1)} = \alpha^{(t)} - \frac{\log \alpha - \psi(\alpha^{(t)}) + (\log x - \log \bar{x})}{1/\alpha^{(t)} - \psi'(\alpha^{(t)})}$$

(ii) Second Approch: find them jointly.

$\ell'(\hat{\theta}) = 0$ the newton update

$$\theta^{(t+1)} = \theta^{(t)} - \big(\underbrace{\ell''(\theta^{(t)})}_{\substack{2\times 2 \\ \text{matrix}}}\big)^{-1} \ell'(\theta^{(t)})$$

So we know the first derivative matrix

$$\ell'(\theta) = \begin{bmatrix} \partial \ell / \partial \alpha \\ \partial \ell / \partial \beta \end{bmatrix}$$

{we know this see above}

$$\ell''(\theta) = \begin{bmatrix} \frac{\partial^2 \ell}{\partial \alpha^2} & \partial \ell / \partial \alpha \partial \beta \end{bmatrix}$$

$$\begin{bmatrix} \partial \ell / \partial\alpha\partial\beta & \partial^2 \ell / \partial\beta^2 \end{bmatrix}$$

$$\frac{\partial^2 \ell}{\partial \alpha^2} = -n\,\psi'(\alpha) \qquad \frac{\partial^2 \ell}{\partial\alpha\partial\beta} = n/\beta$$

$$\frac{\partial^2 \ell}{\partial \beta^2} = \frac{-n\alpha}{\beta^2}$$

We want to find x such that

$$\underbrace{-\ell''(\theta)}_{\text{PSD}}\, x = \ell'(\theta)$$

↳ cholesky exists

$$x = C^{-1} C^{-T} \ell'(\theta)$$ Solve using backsolve.

Ex: $Y_i \overset{ind}{\sim} \text{Bern}(\pi_i) \qquad \pi_i = \dfrac{e^{x_i^T\beta}}{1+e^{x_i^T\beta}}$

$$L(\beta ; y, x) = \prod_{i=1} \pi_i^{y_i} (1-\pi_i)^{1-y_i}$$

$$\ell(\beta; y, x) = \sum_{i=1}^{n} y_i \log \pi_i + (1-y_i) \log(1-\pi_i)$$

$$= \sum_{i=1}^{n} y_i \log\left(\frac{\pi_i}{1-\pi_i}\right) + \log(1-\pi_i)$$

$$= \sum_{i=1}^{n} y_i (x_i^T \beta) - \log\left(1+e^{x_i^T \beta}\right)$$

$$\frac{\partial \ell}{\partial \beta_j} = \sum_{i=1}^{n} y_i x_{ij} - \frac{e^{x_i^T\beta}}{1+e^{x_i^T\beta}} x_{ij} = \sum_{i=1}^{n} x_{ij}(y_i - \pi_j)$$

$$\frac{\partial \ell}{\partial \vec{\beta}} = X^T (y - \pi)$$

$$\frac{\partial^2 \ell}{\partial \beta_i \partial \beta_k} = -x_{ij} \frac{\partial \pi_i}{\partial \beta_k} = -\frac{x_{ij} e^{x_i^T\beta}}{(1+e^{x_i^T\beta})^2} x_{ik}$$

$(1 + e \quad)$

$$\frac{d\ell}{d\beta \, d\beta^T} = X^T \, \mathrm{Diag}\left(\pi_i(1-\pi_i)\right) X$$

$$= X^T W X$$

So the corresponding update is

$$\beta^{(t+1)} = \beta^{(t)} + \left[-\ell''(\beta^{(t)})\right]^{-1} \ell'(\beta^{(t)})$$

$$= \beta^{(t)} + \underbrace{\left(X^T W(\beta^{(t)}) X\right)}_{\text{symm. \& P.D.}}^{-1} X^T \left(y - \pi(\beta^{(t)})\right)$$

$$= \beta^{(t)} + C^{-1} C^{-T} X^T \left(y - \pi(\beta^{(t)})\right)$$

How do we calculate $X^T W(\beta^{(k)}) X$?

Crossprod(X, w * X)

Sweep

applies row like multiplication.

Newton-Raphson (tolerance, max # iter.)

1. Initialize $\theta^{(1)}$ arbitrarily compute $\ell(\theta^{(1)})$

2. For $i = 1, \ldots,$ max.iteration

   a. Compute $\ell'(\theta^{(i)})$, $-\ell''(\theta^{(i)})$

   b. Update $\theta^{(i+1)} = \theta^{(i)} + [-\ell''(\theta^{(i)})] \ell'(\theta^{(i)})$

   c. Compute $\ell(\theta^{(t+1)})$

   d. If $\left\| \dfrac{\ell(\theta^{(i+1)}) - \ell(\theta^{(i)})}{\ell(\theta^{(i)})} \right\| <$ tol.

   then break.