Note: when iterating between (α, β) to find Gamma MLE, we store $\theta = \begin{pmatrix} \alpha \\ \beta \end{pmatrix}$

- Newton update is

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

- Advantage because we can just compute at same time you can get information $I(\theta)$

- Also gain the covariance matrix α, β

Expectation Maximization

- We've seen that we can use a maximum likelihood criterion to estimate θ

Sometimes part of our data is <u>latent</u> – we don't know what z is and we have to estimate θ by using the information in it

<u>Eg</u>: Missing data in linear regression

Ideally we would like

$$L(\theta; x) = \mathbb{P}(x; \theta) = \int \mathbb{P}(x, z; \theta)\, dz$$

and then use MLE estimation as we did before

$$\mathbb{P}(x, z; \theta) = L(\theta; \underbrace{x}_{\text{obs. data}}, \underbrace{z}_{\text{Missing}}) \quad \text{Complete data likel.}$$

– But this is usually hard

– Alternatively we can use the

iterative procedure EM

<u>E-Step:</u> $Q(\theta;\theta^{(t)}) = \mathbb{E}_{Z|X;\theta=\theta^{(t)}}\left[\ell(\theta;X,Z)\right]$

<u>M-Step:</u> $\theta^{(t+1)} = \underset{\theta}{\arg\max}\ Q(\theta;\theta^{(t)})$

<u>Ex:</u> $X, Z \sim \text{iid Exp}(\lambda)$ Z missing

$$\ell(\theta;x,z) = \log f(x,z;\theta)$$

$$= \log f(x;\theta) f(z:\theta)$$

$$= \log f(x;\theta) + \log f(z:\theta)$$

$$= \ell(\theta;x) + \ell(\theta;z)$$

$$= 2\log(\lambda) - \lambda X - \lambda Z$$

$\boxed{E}$ $Q(\lambda,\lambda^{(t)}) = \mathbb{E}_{Z|X,\lambda^{(t)}}\left(\ell(\lambda;X,Z)\right)$

$$= 2\log(\lambda) - \lambda X - \lambda\, \mathbb{E}_{Z|X,\lambda^{(t)}}(Z)$$

$$= 2\log(\lambda) - \lambda x - \frac{\lambda}{\lambda^{(t)}}$$

$\boxed{M}$ $$Q'(\lambda^{(t+1)}, \lambda^{(t)}) = \frac{2}{\lambda^{(t+1)}} - x - \frac{1}{\lambda^{(t)}} = 0$$

$$\boxed{\lambda^{(t+1)} = \frac{2}{x + \frac{1}{\lambda^{(t)}}}}$$

Fixed Point Analysis

$(\lambda^{(t)} \longrightarrow \hat{\lambda})$ $$\frac{2}{\hat{\lambda}} - x - \frac{1}{\hat{\lambda}} = 0$$

$$\hat{\lambda} = \frac{1}{x} = \hat{\lambda}_{MLE}$$

Ex: $X_1, \dots, X_n, Z_1, \dots, Z_m \overset{iid}{\sim} Exp(\lambda)$

but we only observe X_i $\; Y_i = \{Z_i \geq T\}$

$$\ell(\lambda; x, y, z) = \sum_{i=1}^{n} \left(\log(\lambda) - \lambda x_i\right)$$

$$+\sum_{j=1}^{m}\left(\log\lambda \quad -\lambda z_j\right)$$

$$\boxed{E} \quad Q(\lambda,\lambda^{(t)}) = \mathbb{E}_{z|x,y;\lambda^{(t)}}\left(\log(\lambda; X, Y, Z)\right)$$

$$= n\log\lambda - n\lambda\bar{x} + m\log\lambda$$

$$-\lambda\sum_{i=1}^{m}\mathbb{E}_{z|x,y,\lambda^{(t)}}(z_j)$$

$$\mathbb{E}_{z|y,\lambda^{(t)}}(z_j) = \mathbb{E}_{\lambda^{(t)}}\left(z_j \mid z_j \geq T\right)$$

$$= T + \frac{1}{\lambda^{(t)}}$$

$$\Longrightarrow$$

$$= (n+m)\log(\lambda) - \lambda n\bar{x} - \lambda m\left(T + 1/\lambda^{(t)}\right)$$

$$\boxed{M} \quad Q'(\lambda^{(t+1)}, \lambda^{(t)}) = 0$$

$$\frac{m+n}{\quad} - n\bar{x} - m\left(T + \frac{1}{\quad}\right) = 0$$

$\lambda^{(t+1)}$ $\lambda^{(t)}$

$$\boxed{\lambda^{(t+1)} = \frac{n+m}{n\bar{x} + m\left(1 + 1/\lambda^{(t)}\right)}}$$

Again with fixed point analysis

$$\hat{\lambda} = \frac{m+n}{n\bar{x} + m\left(T + \frac{1}{\hat{\lambda}}\right)} \Rightarrow \hat{\lambda} = \frac{n}{n\bar{x} + mT}$$