EM algorithm

Model $Y \sim f(y;\theta) = \int f(y|u;\theta) f(u;\theta)\,du$ ←

U unobservable

$f(y,u;\theta) = \underbrace{f(y;\theta)}_{\text{likelihood fn}} f(u|y;\theta)$ re-written

see ~~this~~ data from this

Y, U random variables

$E_{\theta'}\{\log f(y,U;\theta) | Y=y\} = E_{\theta'}[\quad \text{RHS} \quad]$

$= \cancel{\log f(y;\theta) + E_{\theta'}\{\log f(U|Y=y;\theta)\}}$

$= \log f(y;\theta) + E_{\theta'}\{\log f(U|Y;\theta) \,|\, Y=y\}$

$\varphi(\theta,\theta') = \ell(\theta;y) + C(\theta,\theta')$

max φ ⇒ get ☆ max ℓ

at each step we $\uparrow \varphi$; ⇒ at each step we $\uparrow \ell(\theta;y)$

$C(\theta,\theta') = E_{\theta'}\{\log f(U|Y;\theta) \,|\, Y=y\}$

$\leq C(\theta',\theta')$

$\theta \leftarrow \theta^{(j+1)} \qquad \theta' \leftarrow \theta^{(j)}$

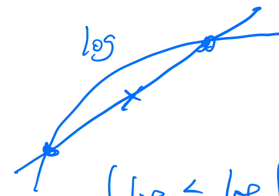
refer back to ch. 4

$$E_{\theta'}\{\log f(u \mid Y;\theta) \mid Y=y\}$$

$$- E_{\theta'}\{\log f(u \mid Y;\theta') \mid Y=y\} \le 0$$

$$\int [\log f(u|y;\theta) \mid Y=y]\, f(u|y;\theta')\, du$$

$$-\int [\log f(u|y;\theta') \mid Y=y]\, f(u|y;\theta')\, du$$



$\int \log \le \log \int$

$$= \int \log\left[\frac{f(u|y;\theta) \mid \cancel{Y=y})}{f(u|y;\theta') \mid \cancel{Y=y})}\right] f(u|y;\theta')\, du$$

$$\le \log \int \frac{f(u|Y=y;\theta)}{f(u\,\cancel{|Y=y};\theta')}\, \cancel{f(u|}y;\theta')\, du \qquad = 0$$

Jensen's inequality.

i) $Q(\theta,\theta') = E_{\theta'}\{\log f(\underline{y, u};\theta) \mid Y=y\}$

E-step "expected value"

2) $\max_{\theta} Q(\theta,\theta')$ $\qquad \theta^{(j+1)}$ $\qquad \theta' = \theta^{(j)}$

M-step

3) $|\ell(\theta^{(j+1)}) - \ell(\theta^{(j)})| < \varepsilon$? or $|\hat{\theta}_c^{(j+1)} - \hat{\theta}^{(j)}| < \varepsilon$

yes → done ; no repeat

Intuition

" E-step : estimate latent variables (missing values)

M-step : max "full likelihood" ↗ $\log f(y,\hat{u};\theta)$

Example $\quad f(y;\beta) = \sum_{r=1}^{p} f_r(y;\beta)\cdot\pi_r$

mixture of p components

$\cancel{f = U = r}$

$$p(U=r \mid \pi) = \pi_r \qquad \text{discrete}$$

$$\log f(y,u;\theta) = \log\left[\prod_{r=1}^{p}\{f_r(y;\beta)\pi_r\}^{1\{u=r\}}\right]$$

$$\theta = (\beta,\pi) \qquad = \sum\left\{1\{u=r\}\log\pi_r + \log f_r(y;\beta)\right\}$$

$$E\{\log f(y,U;\theta) \mid Y=y\} = E\left[\qquad\qquad\right]$$

$Y_1 \dots Y_n$

$$E\log\{f(y;U;\theta \mid Y=y)\} = E\sum_{i=1}^{n}\sum_{r=1}^{p}\left[\qquad \mid y\right]$$

$$P_n\{U=r \mid Y=y;\theta'\} = \frac{\pi_r' f_r(y_i;\beta')}{\sum_{s=1}^{p}\pi_s' f_s(y_i;\beta)} \qquad \text{(Bayes thm.)}$$