# Failure mode EM detailed process.

$$P(X,Z,T)=\prod_{l=1}^{L} P(X|T,Z)\cdot P(T|Z)\cdot P(Z)$$

① $P(Z)=\prod_{k=1}^{K}(\pi_k)^{\rho_{lk}}$

② $P(T|Z)=\prod_{k=1}^{K}\Big(\underbrace{\prod_{m=1}^{M}N(T_{lm}|\mu_{mk},\Sigma_{mk})\cdot N(T_{ly}|\mu_{yk},\Sigma_{yk})}_{P(T_{lk})}\Big)^{\rho_{lk}}$

③ $P(X|T,Z)=\prod_{k=1}^{K}\Big(\underbrace{\prod_{m=1}^{M}N(X_{lm}|\Phi_l T_{lm},\sigma_{mk}^2 I)\cdot N(Y_l|\Phi_l T_{ly},\sigma_{yk}^2 I)}_{P(X_{lk})}\Big)^{\rho_{lk}}$

后验: $P(Z_l=k|X_l)=\dfrac{P(X_l|Z_l=k)P(Z_l=k)}{\sum_{k=1}^{K}}$

$\hat{\rho}_{lk}\leftarrow E(\rho_{lk})=1\cdot P(Z_l=k|X_l)+0\cdot P(\ )+0\cdot P(\ )\ \cdots=P(Z_l=k|X_l)$ (具体略,和之前一样)

$P(T_{lm}|X_{l1},X_{l2}\cdots X_{lm},Y_l,Z_l=k)$ → 需要 $X_{l1},X_{l2}\cdots Y_l$ 独立,但应该就能这样

$=P(T_{lm}|X_{lm},Z_l=k)\sim N(\hat{T}_{lmk},\hat{\Sigma}_{lmk})$ (和之前一样,换下标就行)

E step

$$P(X,Z,T)=\prod_{l}\Big(\prod_{k}\pi_k\cdot P(T_{lk})\cdot P(X_{lk})\Big)^{\rho_{lk}}$$

$$-2\ln P(X,Z,T)=\sum_{l}\sum_{k}\rho_{lk}\cdot\Big(-2\ln\pi_k+\sum_{m=1}^{M}-2\ln N_T(\ \ )+-2\ln N_T(\ \ )+$$
$$+\sum_{m=1}^{M}-2\ln N_x(\ )+-2\ln N_x(\ \ )\Big)$$

以 $\mu_{mk}, \Sigma_{mk}, \sigma^2_{mk}$ 为例，去除无关项

$$= \sum_{l} \rho_{lk} \Big( const - 2\ln|\Sigma_{mk}|^{-\frac{1}{2}} + (T_{lmk} - \mu_{mk})^T \Sigma_{mk}^{-1} (\quad) - 2\ln|\sigma^2_{mk} I|^{-\frac{1}{2}} + (x_{lm} - \Phi_l T_{lm})^T [\sigma^2_{mk} I]^{-1} (\quad) \Big)$$

$$= \sum_{l} \rho_{lk} \Big( const + \ln|\Sigma_{mk}| + (T_{lmk} - \mu_{mk})^T \Sigma_{mk}^{-1} (\quad) + n_l \cdot \ln \sigma^2_{mk} + \frac{1}{\sigma^2_{mk}} \cdot (x_{lm} - \Phi_l T_{lm})^T (\quad) \Big)$$

$$\frac{\partial E(\quad)}{\partial \mu_{mk}} = \sum_{l} E(\rho_{lk}) \cdot (-2) E(T_{lmk} - \mu_{mk})^T \Sigma_{mk}^{-1} = 0$$

$$\Rightarrow \hat{\mu}_{mk} = \frac{\sum_l \hat{\rho}_{lk} \cdot \hat{T}_{lmk}}{\sum_l \hat{\rho}_{lk}}$$

$$\frac{\partial E(\quad)}{\partial \Sigma_{mk}^{-1}} = \frac{\partial \sum_l E(\rho_{lk}) \cdot \Big( -\ln|\Sigma_{mk}^{-1}| + E\big[ (T_{lmk} - \mu_{mk})^T \Sigma_{mk}^{-1} (\quad) \big] \Big)}{\partial \Sigma_{mk}^{-1}}$$

$$\text{分子} = \sum_l \hat{\rho}_{lk} \cdot \Big( -\ln|\Sigma_{mk}^{-1}| + Tr(\hat{\Sigma}_{mk} \cdot \Sigma_{mk}^{-1}) + (\hat{T}_{lmk} - \hat{\mu}_{mk})^T \Sigma_{mk}^{-1} (\quad) \Big)$$

$$\frac{\partial\, \text{分子}}{\partial \Sigma_{mk}^{-1}} = \sum_l \hat{\rho}_{lk} \Big( -\Sigma_{mk} + \hat{\Sigma}_{mk} + (\hat{T}_{mk} - \hat{\mu}_{mk})(\hat{T}_{mk} - \hat{\mu}_{mk})^T \Big) = 0$$

$$\hat{\Sigma}_{mk} = \frac{\sum_l \hat{\rho}_{lk} \Big[ \hat{\Sigma}_{mk} + (\hat{T}_{mk} - \hat{\mu}_{mk})(\hat{T}_{mk} - \hat{\mu}_{mk})^T \Big]}{\sum_l \hat{\rho}_{lk}}$$

$$\frac{\partial E(\quad)}{\partial \sigma_{mk}^2} = \sum_l E(\rho_{lk})\left(n_l \cdot \frac{1}{\sigma_{mk}^2} - \frac{1}{(\sigma_{mk}^2)^2} E(x_{lm} - \Phi_l T_{lmk})^T(\quad)\right)$$

$$= 0$$

两边乘 $(\sigma_{mk}^2)^2$

$$\sigma_{mk}^2 \sum_l \hat{\rho}_{lk} n_l = \sum_l \hat{\rho}_{lk} (x_{lm} - \Phi_l \hat{T}_{lm})^T(\quad) + tr(\Phi_l \hat{\Sigma}_{lmk} \Phi_l^T)$$

$$\hat{\sigma}_{mk}^2 = \frac{\sum_l \hat{\rho}_{lk}\left((\quad)^T(\quad) + tr(\Phi_l \hat{\Sigma} \Phi_l^T)\right)}{\sum_l \hat{\rho}_{lk} n_l}$$

Update:

$$P(X, T, Z, \theta) = P(X|T, Z, \theta) \cdot P(T|Z, \theta) \cdot P(\theta|Z) \cdot P(Z)$$

$$P(\theta|Z) = \prod_{m=1}^{M+1} (NIW)$$

$$-2 \ln P(\theta|Z) = \sum_{m=1}^{M+1} \Big( (\tau+d+2) \ln |\Sigma_{mk}| + Tr(\Sigma_{mk0} \Sigma_{mk}^{-1}) + \beta (\mu_{mk} - \mu_{mk0})^T \Sigma_{mk}^{-1} (\mu_{mk} - \mu_{mk0}) \Big)$$

$$\frac{\partial E(\ )}{\partial \mu_{mk}} = \sum_i \hat{P}_{ik} (-2) (\hat{T}_{imk} - \mu_{mk}) \Sigma_{mk}^{-1} + 2\beta (\mu_{mk} - \mu_{mk0})^T \Sigma_{mk}^{-1} = 0$$

$$\Rightarrow \hat{\mu}_{mk} = \frac{\sum_i \hat{P}_{ik} \hat{T}_{imk} + \beta \mu_{mk0}}{\sum_i \hat{P}_{ik} + \beta}$$

$$\frac{\partial E(\quad)}{\partial \Sigma_{mk}^{-1}}$$

分子: $\sum_i \hat{\rho}_{ik} \cdot \Big( -\ln|\Sigma_{mk}^{-1}| + Tr(\hat{\Sigma}_{mk} \Sigma_{mk}^{-1}) +$

$(\hat{T}_{imk} - \hat{\mu}_{mk})^T \Sigma_{mk}^{-1} (\quad) \Big)$

$+ Tr(\Sigma_{mk0} \cdot \Sigma_{mk}^{-1}) + \beta(\mu_{mk} - \mu_{mk0})^T \Sigma_{mk}^{-1} (\quad)$

$- (\tau + d + 2) \Sigma_{mk} = 0$

$$\Rightarrow \hat{\Sigma}_{mk} = \frac{\sum_i \hat{\rho}_{ik} \left[ \hat{\Sigma}_{mk} + (\hat{T}_{mk} - \hat{\mu}_{mk})(\hat{T}_{mk} - \hat{\mu}_{mk})^T \right] + \Sigma_{mk0} + \beta(\mu_{mk} - \mu_{mk0})(\quad)^T}{\sum_i \hat{\rho}_{ik} + \tau + d + 2}$$