引入 1 从缺失数据入手

- 观察data X
  缺失data Z
  完全data $Y = (X, Z)$
  参数 θ.

似然：

- $L(\theta|X) = P(X|\theta) \rightarrow$ 直接求解困难.
  $L(\theta|Y) = P(Y|\theta) \rightarrow$ 简单，容易计算. 利用
  $P(Z|Y, \theta) \rightarrow$ 简单

计算已知完全data下θ的后验期望.

已知 $\theta^{(t)}$ 下，在完全数据下θ的期望. 完全似然 缺失data.

E-step:

- $$Q(\theta|\theta^{(t)}) = \int \log L(\theta|Y)\, p(z|X, \theta^{(t)})$$
  $$= E\left[\int \log L(\theta|Y) \,\middle|\, X, \theta^{(t)}\right].$$

M-step:

- $$\theta^{(t+1)} = \arg\max_{\theta} Q(\theta|\theta^{(t)}),$$

极大似然

↓

Q函数.

- $$\log P(X|\theta) = \int \log P(x, z|\theta)\, dz$$
  $$= \log \int \frac{p(x,z|\theta)}{q(z)} q(z)\, dz \qquad \text{令 } q(z) = P(z|x, \theta^{(t)}).$$
  $$\geq \int \log \frac{p(x,z|\theta)}{p(z|x,\theta^{(t)})} \cdot p(z|x,\theta^{(t)})\, dz$$
  $$= \underbrace{\int \log p(x,z|\theta)\, p(z|x,\theta^{(t)})\, dz}_{Q(\theta,\theta^{(t)})} - \underbrace{\int \log p(z|x,\theta^{(t)})\, dz}_{\text{与}\theta\text{无关 const.}}$$

EM 直观

EM

- E-step：

$$Q(\theta \mid \theta^{(t)}) = \int \log L(\theta \mid Y)\, p(Z \mid X, \theta^{(t)})$$

$$= E[\log L(\theta \mid Y) \mid X, \theta^{(t)}]$$

$$= E[\log p(X, Z \mid \theta) \mid X, \theta^{(t)}]$$

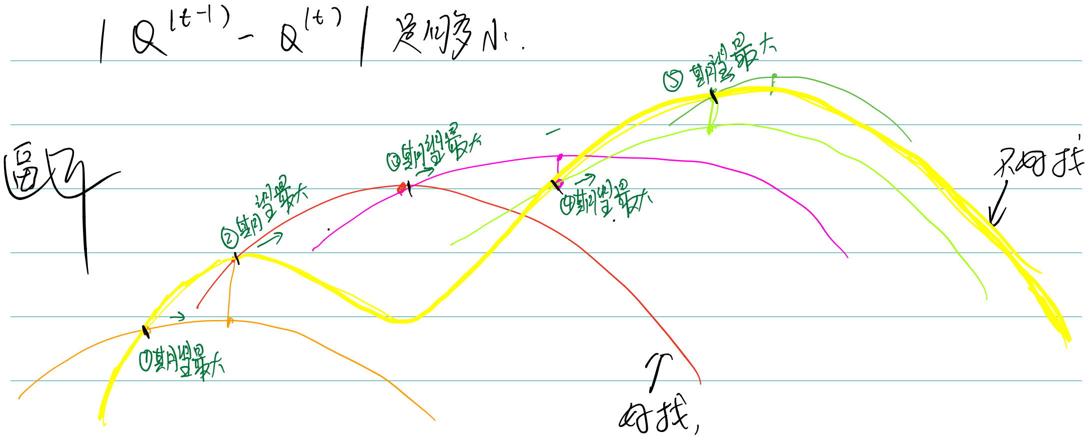
M-step：

$$\theta^{(t+1)} = \arg\max_{\theta} Q(\theta \mid \theta^{(t)}).$$

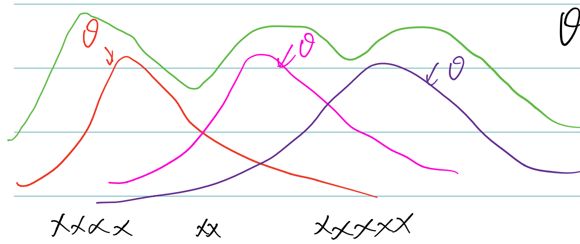
停止条件：

$$|Q^{(t-1)} - Q^{(t)}| \text{ 足够小.}$$

逐渐逼近



引入2 从高斯混合模型入手.



$\theta = \{\mu_1, \cdots, \Sigma_1, \cdots, \alpha_1, \cdots, \alpha_{k-1}\}$

- $\theta^{(g+1)} = \arg\max_\theta \int \log P(x, z|\theta) \cdot P(z|X, \theta^{(g)})$

  X：观测数据

  z：隐变量.

- 对于混合模型 $\theta = \{\alpha_1 \cdots \alpha_{k-1} \; \mu_1 \cdots \mu_k, \Sigma_1 \cdots \Sigma_k\}$

  $\theta_{MLE} = \arg\max_\theta L(\theta|X)$

  $= \arg\max_\theta \left(\Sigma \log \sum \alpha_l N(X|\mu_l, \Sigma_l) \right)$

  直接求解难.

- $P(x_i) = \int \underbrace{P(x_i|z_i)}_{N(x_i|\mu_i, \Sigma_i)} \cdot \underbrace{P(z_i)}_{\alpha_{z_i}} dz_i$

  $= \sum_{z_i=1}^{k} \alpha_{z_i} N(x_i|\mu_i, \Sigma_i)$.