引入

$$\theta = \arg\max_{\theta} P(X|\theta) = \arg\max_{\theta} E[P(X,Z|\theta)] = \arg\max_{\theta} E[\ln P(X,Z|\theta)]$$

←引入隐变量,

$\ln P(X,Z|\theta)$ 的期望计算中，对于序列X是已知的，而 θ' 的概率是由旧参数值 所估计的，因此上式可以表示为：

旧参数, 要估计的新参数

$$\theta = \arg\max_{\theta} \sum_{Z} P(Z|X,\theta') \ln P(X,Z|\theta)$$

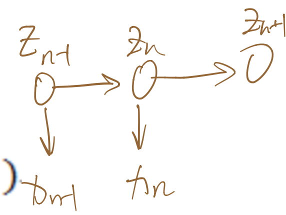
为了方便表示，以下定义：

$$Q(\theta,\theta') = \sum_{Z} P(Z|X,\theta') \ln P(X,Z|\theta)$$

而 $P(X,Z|\theta)$ 可以表示为：

初始值, 隐状态的转移.

$$P(X,Z|\theta) = P(Z|\theta)P(X|Z,\theta) = P(z_1|\theta)\prod_{n=2}^{N} P(z_n|z_{n-1},\theta)\prod_{n=1}^{N} P(x_n|z_n,\theta)$$



根据HMM的结构定义，其参数θ主要分为三部分：隐藏状态的先验分布π（同 $P(z_1)$ 相关），各隐藏状态之间的转移概率Λ（同 $P(z_n|z_{n-1})$ 相关），即已知隐藏状态确定观测值的发射概率参数∅（同 $P(x_n|z_n)$ 相关）。由此可以得出

初始值 先验分布 转移矩阵 发射矩阵

$$\ln P(X,Z|\theta) = \underbrace{\ln P(z_1|\pi)}_{\text{item 1}} + \underbrace{\sum_{n=2}^{N} \ln P(z_n|z_{n-1},\Lambda)}_{\text{item 2}} + \sum_{n=1}^{N} \underbrace{\ln P(x_n|z_n,\emptyset)}_{\text{item 3.}}$$

如果发射概率是高斯概率的情况（以下都按此设置，其它概率假设也是类似的），这里的K表示隐藏状态的可能个数，N表示序列的长度。π_k表示初始隐藏状态为k的先验概率，Λ_{jk}表示由状态j转移到状态k的概率，μ_k, Σ_k表示隐藏状态为k的发射概率。发射概率为高斯.

第n个序列为k状态.

$$\ln P(X,Z|\theta) = \sum_{k=1}^{K} \ln \pi_k^{z_{1k}} + \sum_{n=2}^{N}\sum_{j=1}^{K}\sum_{k=1}^{K} \ln \Lambda_{jk}^{z_{n-1,j}\cdot z_{n,k}} + \sum_{n=1}^{N}\sum_{k=1}^{K} \ln N(x_n|\mu_k,\Sigma_k)^{z_{n,k}}$$

状态转移

第n个序列为k状态时发射概率.

此时综合得出：

$$Q(\theta,\theta')=\sum_{k=1}^{K}\sum_{Z}P(Z|X,\theta')\ln\pi_k^{z_{1k}}+\sum_{n=2}^{N}\sum_{j=1}^{K}\sum_{k=1}^{K}\sum_{Z}P(Z|X,\theta')\ln\Lambda_{jk}^{z_{n-1,j}\cdot z_{n,k}}+\sum_{n=1}^{N}\sum_{k=1}^{K}\sum_{Z}P(Z|X,\theta')z_{n,k}\ln N\left(x_n|\mu_k,\Sigma_k\right)$$

$$Q(\theta,\theta')=\underbrace{\sum_{k=1}^{K}\gamma(z_{1k})\ln\pi_k}_{\text{初始}}+\underbrace{\sum_{n=2}^{N}\sum_{j=1}^{K}\sum_{k=1}^{K}\vartheta(z_{n-1,j},z_{n,k})\ln\Lambda_{jk}}_{\text{转移}}+\underbrace{\sum_{n=1}^{N}\sum_{k=1}^{K}\gamma(z_{nk})\boxed{\ln N(x_n|\mu_k,\Sigma_k)}}_{\text{发射}}$$

即 $P(x_n|\phi_k)$

这里定义：

已知θ'与X下，第n个状态为k的概率

$$\gamma(z_{n,k})=\sum_{Z}P(Z|X,\theta')z_{n,k}=P(z_{n,k}|X,\theta')$$

已知θ'下，n-1状态为j n状态为k 的概率，

$$\vartheta(z_{n-1,j},z_{n,k})=\sum_{Z}P(Z|X,\theta')z_{n-1,j}\cdot z_{n,k}=P(z_{n-1,j},z_{n,k}|X,\theta')$$

① $Q(\theta,\theta')$的定义式中可以看出，$\boxed{\gamma(z_{n,k})\text{和}\vartheta(z_{n-1,j},z_{n,k})}$是同旧参数值$\theta'$相关的

② 而先验概率π、转移概率Λ和发射概率参数$\emptyset$是需要估计的新参数

因此可以建立 EM 算法，E 步骤估计$\gamma(z_{n,k})$和$\vartheta(z_{n-1,j},z_{n,k})$，而 M 步骤计算新参数$\pi$、$\Lambda$和$\emptyset$。

与旧参数有关的估计　　新参数有关，