date: 20230508

goal: 半监督的高斯混合模型的EM算法推导过程

$L=\prod_{D_l}P(X_j,Y_j|\theta)\prod_{D_u}P(X_j|\theta)=\prod_{D_l}P(X_j|\theta)P(Y_j|\theta,X_j)\prod_{D_u}P(X_j|\theta)$
$=\prod_{D_l}[\sum_{i=1}^{N}\alpha_iP(X_j|\mu_i,\Sigma_i)P(Y_j|\mu_i,\Sigma_i,X_j)]\prod_{D_u}[\sum_{i=1}^{N}\alpha_iP(X_j|\mu_i,\Sigma_i)]$

$P(X_j|\mu_i,\Sigma_i)=\frac{1}{(2\pi)^{\frac{p}{2}}|\Sigma_i|^{\frac{1}{2}}}\exp\left(-\frac{1}{2}(X_j-\mu_i)^T\Sigma_i^{-1}(X_j-\mu_i)\right)$

$$P(Y_j|\mu_i,\Sigma_i,X_j)=\begin{cases}0 & i=Y_j\\ 1 & i\neq Y_j\end{cases}$$

$$l=\log L=\sum_{D_l}\log(\sum_{i=1}^{N}\alpha_iP(X_j|\mu_i,\Sigma_i)I(Y_j=i))+\sum_{D_u}\log(\sum_{i=1}^{N}\alpha_iP(X_j|\mu_i,\Sigma_i))$$

记$A=\sum_{D_l}\log(\sum_{i=1}^{N}\alpha_iP(X_j|\mu_i,\Sigma_i)I(Y_j=i))$
$B=\sum_{D_u}\log(\sum_{i=1}^{N}\alpha_iP(X_j|\mu_i,\Sigma_i))$

$\frac{\partial A}{\partial\mu_i}=\sum_{D_l}\frac{\alpha_i\frac{\partial P(X_j|\mu_i,\Sigma_i)}{\partial\mu_i}I(Y_j=i)}{\sum_{i=1}^{N}\alpha_iP(X_j|\mu_i,\Sigma_i)I(Y_j=i)}=\sum_{D_l}\frac{\alpha_iP(X_j|\mu_i,\Sigma_i)(X_j-\mu_i)I(Y_j=i)}{\sum_{i=1}^{N}\alpha_iP(X_j|\mu_i,\Sigma_i)I(Y_j=i)}$

$$\begin{cases}\sum_{D_l,Y_j=i}(X_j-\mu_i) & i=Y_j\\ 0 & i\neq Y_j\end{cases}$$

$\frac{\partial A}{\partial\mu_i}=\sum_{D_l,Y_j=i}(X_j-\mu_i)$

$\frac{\partial B}{\partial\mu_i}=\sum_{D_u}\frac{\alpha_i\frac{\partial P(X_j|\mu_i,\Sigma_i)}{\partial\mu_i}}{\sum_{i=1}^{N}\alpha_iP(X_j|\mu_i,\Sigma_i)}=\sum_{D_u}\frac{\alpha_iP(X_j|\mu_i,\Sigma_i)(X_j-\mu_i)}{\sum_{i=1}^{N}\alpha_iP(X_j|\mu_i,\Sigma_i)}$

$\frac{\partial l}{\partial\mu_i}=\frac{\partial A}{\partial\mu_i}+\frac{\partial B}{\partial\mu_i}=\sum_{D_l,Y_j=i}(X_j-\mu_i)+\sum_{D_u}\frac{\alpha_iP(X_j|\mu_i,\Sigma_i)(X_j-\mu_i)}{\sum_{i=1}^{N}\alpha_iP(X_j|\mu_i,\Sigma_i)}=0$

（1） $\mu_i=\frac{\sum_{D_l,Y_j=i}X_j+\sum_{D_u}\frac{\alpha_iP(X_j|\mu_i,\Sigma_i)X_j}{\sum_{i=1}^{N}\alpha_iP(X_j|\mu_i,\Sigma_i)}}{\sum_{D_l,Y_j=i}1+\sum_{D_u}\frac{\alpha_iP(X_j|\mu_i,\Sigma_i)}{\sum_{i=1}^{N}\alpha_iP(X_j|\mu_i,\Sigma_i)}}$

$\partial\Sigma_i(m,n)$为Σ_i的m行n列的元素

$\frac{\partial B}{\partial\Sigma_i(m,n)}=\sum_{D_u}\frac{\alpha_i\frac{\partial P(X_j|\mu_i,\Sigma_i)}{\partial\Sigma_i(m,n)}}{\sum_{i=1}^{N}\alpha_iP(X_j|\mu_i,\Sigma_i)}$

$\frac{\partial P(X_j|\mu_i,\Sigma_i)}{\partial\Sigma_i(m,n)}=\frac{1}{(2\pi)^{\frac{p}{2}}|\Sigma_i|^{\frac{3}{2}}}(-\frac{1}{2}\frac{\partial|\Sigma_i|}{\partial\Sigma_i(m,n)})\exp\left(-\frac{1}{2}(X_j-\mu_i)^T\Sigma_i^{-1}(X_j-\mu_i)\right)+$
$\frac{1}{(2\pi)^{\frac{p}{2}}|\Sigma_i|^{\frac{1}{2}}}\exp\left(-\frac{1}{2}(X_j-\mu_i)^T\Sigma_i^{-1}(X_j-\mu_i)\right)(-\frac{1}{2}\frac{\partial D}{\partial\Sigma_i(m,n)}),whereD=(X_j-\mu_i)^T\Sigma_i^{-1}(X_j-\mu_i)$

先看前半部分，由于$|\Sigma_i|=\sum_{(m或n)}\Sigma_{i(m,n)}*A_{i(m,n)},A_{i(m,n)}$为对应元素的代数余子式
所以$\frac{\partial|\Sigma_i|}{\partial\Sigma_i(m,n)}=A_{i(m,n)}$, 所以前半部分为： $-\frac{1}{2}P(X_j|\mu_i,\Sigma_i)\frac{A_{i(m,n)}}{|\Sigma_i|}$

再看后半部分$\frac{\partial D}{\partial\Sigma_i(m,n)}=(X_j-\mu_i)^T\frac{\partial\Sigma_i^{-1}}{\partial\Sigma_i(m,n)}(X_j-\mu_i)$

由于$\frac{\partial I}{\partial\Sigma_i(m,n)}=0$,得到$\frac{\partial\Sigma_i\Sigma_i^{-1}}{\partial\Sigma_i(m,n)}=\frac{\partial\Sigma_i}{\partial\Sigma_i(m,n)}\Sigma_i^{-1}+\Sigma_i\frac{\partial\Sigma_i^{-1}}{\partial\Sigma_i(m,n)}=0$
移项得$\frac{\partial\Sigma_i^{-1}}{\partial\Sigma_i(m,n)}=-\Sigma_i^{-1}\frac{\partial\Sigma_i}{\partial\Sigma_i(m,n)}\Sigma_i^{-1}$，代入上式得

$\frac{\partial D}{\partial \Sigma_{i(m,n)}}=-(X_j-\mu_i)^T\Sigma_i^{-1}\frac{\partial \Sigma_i}{\partial \Sigma_{i(m,n)}}\Sigma_i^{-1}(X_j-\mu_i)=-[\Sigma_i^{-1}(X_j-\mu_i)]^T\frac{\partial \Sigma_i}{\partial \Sigma_{i(m,n)}}[\Sigma_i^{-1}(X_j-\mu_i)]$，where，$\frac{\partial \Sigma_i}{\partial \Sigma_{i(m,n)}}$为在m,n处为1，其余为0的矩阵

所以后半部分为$\frac{1}{2}P(X_j|\mu_i,\Sigma_i)[\Sigma_i^{-1}(X_j-\mu_i)]^T\frac{\partial \Sigma_i}{\partial \Sigma_{i(m,n)}}[\Sigma_i^{-1}(X_j-\mu_i)]$

$\frac{\partial P(X_j|\mu_i,\Sigma_i)}{\partial \Sigma_{i(m,n)}}=-\frac{1}{2}P(X_j|\mu_i,\Sigma_i)\frac{A_{i(m,n)}}{|\Sigma_i|}+\frac{1}{2}P(X_j|\mu_i,\Sigma_i)[\Sigma_i^{-1}(X_j-\mu_i)]^T\frac{\partial \Sigma_i}{\partial \Sigma_{i(m,n)}}[\Sigma_i^{-1}(X_j-\mu_i)]$

所以，$\frac{\partial B}{\partial \Sigma_{i(m,n)}}=\sum_{D_u}\frac{\alpha_i(-\frac{1}{2}P(X_j|\mu_i,\Sigma_i)\frac{A_{i(m,n)}}{|\Sigma_i|}+\frac{1}{2}P(X_j|\mu_i,\Sigma_i)[\Sigma_i^{-1}(X_j-\mu_i)]^T\frac{\partial \Sigma_i}{\partial \Sigma_{i(m,n)}}[\Sigma_i^{-1}(X_j-\mu_i)])}{\sum_{i=1}^N\alpha_iP(X_j|\mu_i,\Sigma_i)}$

可以发现前半部分$\frac{A_{i(m,n)}}{|\Sigma_i|}$为$\Sigma_i^{-1}$的$m,n$位置的元素；且由于$\frac{\partial \Sigma_i}{\partial \Sigma_{i(m,n)}}$为在m,n处为1，其余为0的矩阵，所以后半部分为$\Sigma_i^{-1}(X_j-\mu_i)(X_j-\mu_i)^T\Sigma_i^{-1}$的$m,n$元素

所以$\frac{\partial B}{\partial \Sigma_i}=\sum_{D_u}\frac{\alpha_iP(X_j|\mu_i,\Sigma_i)[-\frac{1}{2}\Sigma_i^{-1}+\frac{1}{2}(\Sigma_i^{-1}(X_j-\mu_i)(X_j-\mu_i)^T\Sigma_i^{-1})]}{\sum_{i=1}^N\alpha_iP(X_j|\mu_i,\Sigma_i)}$

同理，可以证明得到如下:

$\frac{\partial A}{\partial \Sigma_{i(m,n)}}=\sum_{D_l,Y_j=i}\frac{\alpha_i\frac{\partial P(X_j|\mu_i,\Sigma_i)}{\partial \Sigma_{i(m,n)}}}{\alpha_iP(X_j|\mu_i,\Sigma_i)}=\sum_{D_l,Y_j=i}-\frac{1}{2}\frac{A_{i(m,n)}}{|\Sigma_i|}+\frac{1}{2}[\Sigma_i^{-1}(X_j-\mu_i)]^T\frac{\partial \Sigma_i}{\partial \Sigma_{i(m,n)}}[\Sigma_i^{-1}(X_j-\mu_i)]$

所以$\frac{\partial A}{\partial \Sigma_i}=\sum_{D_l,Y_j=i}[-\frac{1}{2}\Sigma_i^{-1}+\frac{1}{2}(\Sigma_i^{-1}(X_j-\mu_i)(X_j-\mu_i)^T\Sigma_i^{-1})]$

$\frac{\partial l}{\partial \Sigma_i}=\frac{\partial A}{\partial \Sigma_i}+\frac{\partial B}{\partial \Sigma_i}=\sum_{D_l,Y_j=i}[-\frac{1}{2}\Sigma_i^{-1}+\frac{1}{2}(\Sigma_i^{-1}(X_j-\mu_i)(X_j-\mu_i)^T\Sigma_i^{-1})]$
$+\sum_{D_u}\frac{\alpha_iP(X_j|\mu_i,\Sigma_i)[-\frac{1}{2}\Sigma_i^{-1}+\frac{1}{2}(\Sigma_i^{-1}(X_j-\mu_i)(X_j-\mu_i)^T\Sigma_i^{-1})]}{\sum_{i=1}^N\alpha_iP(X_j|\mu_i,\Sigma_i)}$

$\sum_{D_l,Y_j=i}(-1)+\sum_{D_l,Y_j=i}(X_j-\mu_i)(X_j-\mu_i)^T\Sigma_i^{-1}-\sum_{D_u}\frac{\alpha_iP(X_j|\mu_i,\Sigma_i)}{\sum_{i=1}^N\alpha_iP(X_j|\mu_i,\Sigma_i)}+$
$\sum_{D_u}\frac{\alpha_iP(X_j|\mu_i,\Sigma_i)(X_j-\mu_i)(X_j-\mu_i)^T\Sigma_i^{-1}}{\sum_{i=1}^N\alpha_iP(X_j|\mu_i,\Sigma_i)}=0$

(2) $\Sigma_i=\frac{\sum_{D_l,Y_j=i}(X_j-\mu_i)(X_j-\mu_i)^T+\sum_{D_u}\frac{\alpha_iP(X_j|\mu_i,\Sigma_i)(X_j-\mu_i)(X_j-\mu_i)^T}{\sum_{i=1}^N\alpha_iP(X_j|\mu_i,\Sigma_i)}}{\sum_{D_l,Y_j=i}1+\sum_{D_u}\frac{\alpha_iP(X_j|\mu_i,\Sigma_i)}{\sum_{i=1}^N\alpha_iP(X_j|\mu_i,\Sigma_i)}}$

由于$\sum\alpha_i=1$，$\alpha_i\geq 0$, 所以使用拉格朗日乘子法求解:
$U=l+\lambda(\sum_{i=1}^N\alpha_i-1)$

$\frac{\partial U}{\partial \alpha_i}=\sum_{D_l,Y_j=i}\frac{1}{\alpha_i}+\sum_{D_u}\frac{P(X_j|\mu_i,\Sigma_i)}{\sum_{i=1}^N\alpha_iP(X_j|\mu_i,\Sigma_i)}+\lambda=0$
将上式乘以α_i,并从i=1,累加到i=N，可以计算得到$\lambda=l_o+u_o, l_o$为D_l的个数，u_o为D_u的个数

(3) 所以$\alpha_i=\frac{\sum_{D_l,Y_j=i}1+\sum_{D_u}\frac{P(X_j|\mu_i,\Sigma_i)}{\sum_{i=1}^N\alpha_iP(X_j|\mu_i,\Sigma_i)}}{l_o+u_o}$