变分期望最大化Variational Expectation Maximization

问题描述

对于对称无权重图$G=(V,X)$，其中$V=v_1,...v_n$描述图中节点，$X=\{x_{ij}\}$描述节点i与节点j
之间关联；根据图的SBM模型，我们认为图中共含有C个簇，每个节点i从属于簇m的概率用 $z_{im}\in\mathbf{Z}^{N\times C}, i\in\{1,...,n\}, m\in\{1,...,C\}$
来描述，且z_i相互独立并满足多项分布 $z_i\ ^{iid}M(1,\alpha)$ 基于此，我们使用z_i 来描述节点之间关联。我们认为任意两节点之间的关联相互独立并满足二项分布
$x_{ij}|z_{im},z_jn=1\ B(1,\pi_{z_iz_j})$ 可以看出，只要解出参数π与α, 我们就可以构建出一个SBM模型，继而对图结构加以推断；因此我们给出以下目标：

> 基于观察到的节点关联信息X以及隐变量Z，估计参数$\theta=\{\alpha,\pi\}$
>
> $$\hat{\theta}=argmaxl_{\theta}(X)$$

推导与整理

在节点关联与节点簇信息均已知的情况下，我们首先计算似然函数l_θ，根据边缘概率公式，有 $l_\theta(X)=\sum_Z\log p_\theta(X,Z)=\sum_Z(\log p_\pi(X|Z)+\log p_\alpha(Z))$ 其中 $\sum_Z(\log p_\pi(X|Z)+\log p_\alpha(Z))=\sum_Z(\sum_{i,j}^n\sum_{m,n}^C\log p_\pi(x_{ij}|z_{im}z_{jn})+\sum_i^N\sum_m^C\log p_\alpha($
给定参数时，有 $p_\pi(x_{ij}|z_{im}z_{jn};\pi)=\binom{1}{X_{ij}}[\pi_{mn}(1-\pi_{mn})^{1-X_{ij}}]^{z_{im}z_{jn}}$

$p_\alpha(z_{im})=\alpha_m^{z_{im}}$

$$l_\theta(X)=\sum_Z(\sum_{i,j}^n\sum_{m,n}^C\log p_\pi(x_{ij}|z_{im}z_{jn})+\sum_i^N\sum_m^C\log p_\alpha(z_{im}))$$

可得

$$=\sum_Z(\sum_{i,j}^n\sum_{m,n}^C\log\binom{1}{X_{ij}}[\pi_{mn}(1-\pi_{mn})^{1-X_{ij}}]^{z_{im}z_{jn}}+\sum_i^N\sum_m^C\log\alpha_m^{z_{im}})$$

当Z的数量足够大时，我们无法通过通过最大似然法直接求解参数；如果使用期望最大化算法，构造的辅助函数将会含有$P(Z|X)$，同样无法求解，因此我们需要通过变分法模拟该后验分布$P(Z|X)$；