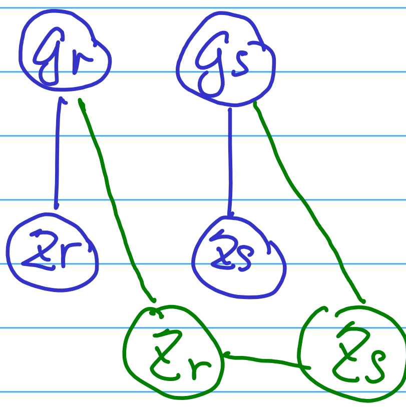
Parameter estimation of Hierarchical Model

$$\theta = \arg\max E[-\log P(G, Z)]$$



$$P(G,Z) = P(G)\cdot P(Z|G)$$

$$= P(G)\cdot \prod_j P(Z^j|G)$$

$$= \frac{1}{Z_G}\exp\Big\{-\beta_g \sum_{(r,s)} [g_r \neq g_s]\Big\}$$

$$\cdot \prod_j \frac{1}{Z_z}\exp\Big\{-\alpha \sum_s [g_s \neq z_s^j] - \beta_z \sum_{(r,s)} [z_r^j \neq z_s^j]\Big\}$$

(Using Pseudo-likelihood approximation)

$$= \prod_{s\in G} P(g_s|\cdot)\cdot \prod_{s\in Z^j} P(g_s|\cdot)$$

$$= \prod_{s\in G} \frac{1}{Z_g}\exp\Big\{-\beta_g \sum_{r\in N_s} [g_r \neq g_s] - \alpha \sum_{j\in J} [g_s \neq z_s^j]\Big\}$$ group nodes

$$\cdot \prod_{s\in Z^j} \frac{1}{Z_z}\exp\Big\{-\alpha [g_s \neq z_s^j] - \beta_z \sum_{r\in N_s} [z_r^j \neq z_s^j]\Big\}$$ subject nodes

$$Z_g = \sum_k \exp\Big\{-\beta_g \sum_{r\in N_s} [g_r \neq k] - \alpha \sum_j [z_s^j \neq k]\Big\}$$

$$Z_z = \sum_k \exp\Big\{-\alpha [g_s \neq k] - \beta_z \sum_{r\in N_s} [z_r^j \neq k]\Big\}$$

Let $a = -\sum_j [g_s \neq z_s^j]$, $b = -\sum_{r\in N_s} [g_r \neq g_s]$

$a_1 = -[g_s \neq z_s^j]$, $b_1 = -\sum_{r\in N_s} [z_r^j \neq z_s^j]$

$$-\log p(G, \xi) = \sum_{s \in G} \left(-a\alpha - b\beta_g + \log \sum_k e^{a\alpha + b\beta_g}\right)$$

$$+ \sum_{s \in \xi j} \left(-a_1\alpha - b_1\beta_\xi + \log \sum_k e^{a_1\alpha + b_1\beta_\xi}\right) = \sum_{s \in G} q_1 + \sum_{s \in \xi j} q_2$$

$$q_1 = -a\alpha - b\beta_g + \log \sum_k e^{a\alpha + b\beta_g}$$

$$q_2 = -a_1\alpha - b_1\beta_\xi + \log \sum_k e^{a_1\alpha + b_1\beta_\xi}$$

$$E[-\log P(G, \xi)] \approx \frac{1}{M} \sum_m \left(-\log P(G^m, \xi^m)\right)$$

$$= \frac{1}{M} \sum_{m, s \in G} \underset{\Downarrow}{q_1} + \frac{1}{M} \sum_{m, s \in \xi j} \underset{\Downarrow}{q_2}$$

$$q_1(\alpha, \beta_g) \qquad q_2(\alpha, \beta_\xi)$$

$$\frac{\partial q_1}{\partial \beta_g} = -b + \frac{\sum_k e^{a\alpha + b\beta_g} \cdot b}{\sum_k e^{a\alpha + b\beta_g}}$$

$$\frac{\partial^2 q_1}{\partial \beta_g^2} = \frac{\left(\sum_k e^{a\alpha + b\beta_g} \cdot b^2\right)\left(\sum_k e^{a\alpha + b\beta_g}\right) - \left(\sum_k e^{a\alpha + b\beta_g} \cdot b\right)^2}{\left(\sum_k e^{a\alpha + b\beta}\right)^2}$$

$$M_0 = \sum_k e^{a\alpha + b\beta}, \quad M_1 = \sum_k e^{a\alpha + b\beta_g} \cdot b, \quad M_2 = \sum_k e^{a\alpha + b\beta_g} \cdot b^2$$

$$\frac{\partial q_1}{\partial \beta_g} = -b + \frac{M_1}{M_0}, \quad \frac{\partial^2 q_1}{\partial \beta_g^2} = \frac{M_2 \cdot M_0 - M_1^2}{M_0^2}$$

$$\frac{\partial q_1}{\partial \alpha} = -a + \frac{\sum_k e^{a\alpha + b\beta_g} \cdot a}{\sum_k e^{a\alpha + b\beta_g}} =$$

$$\frac{\partial^2 q_1}{\partial \alpha^2} = \frac{\left(\sum_k e^{a\alpha + b\beta_g} \cdot a^2\right)\left(\sum_k e^{a\alpha + b\beta_g}\right) - \left(\sum_k e^{a\alpha + b\beta_g} \cdot a\right)^2}{\left(\sum_k e^{a\alpha + b\beta_g}\right)^2}$$

$$M_0=\sum_K e^{a\alpha+b_1\beta}, \quad M_1=\sum_k e^{a\alpha+b\beta_g}\cdot a, \quad M_2=\sum_k e^{a\alpha+b\beta_g}\cdot a^2$$

$$\frac{\partial q_1}{\partial \alpha}=-a+\frac{M_1}{M_0}, \quad \frac{\partial^2 q_1}{\partial \alpha^2}=\frac{M_2\cdot M_0-M_1^2}{M_0^2}$$

$$\frac{\partial q_2}{\partial \alpha}=-a_1+\frac{M_1}{M_0}, \quad \frac{\partial^2 q_2}{\partial \alpha^2}=\frac{M_2\cdot M_0-M_1^2}{M_0^2}$$

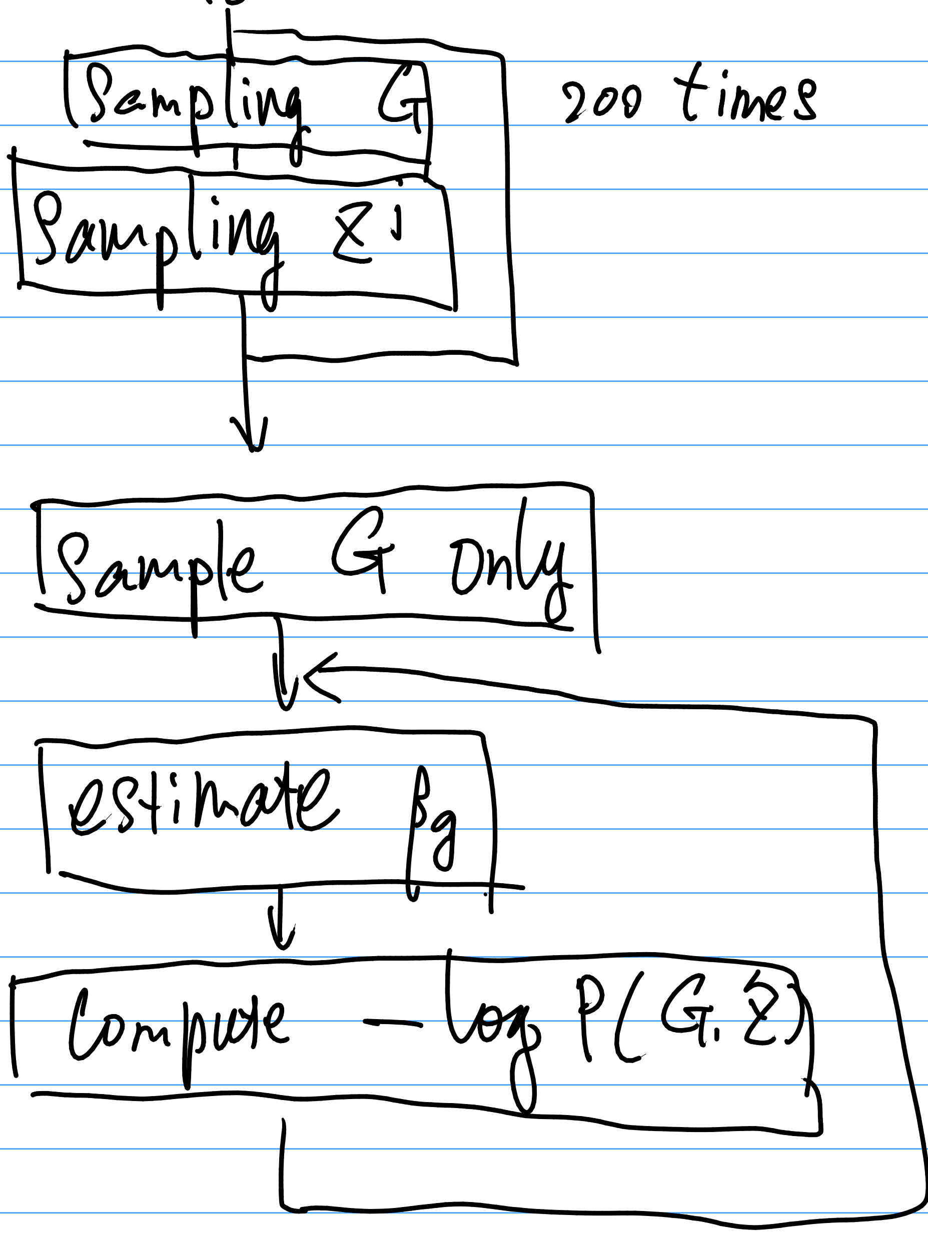
Replacing $\begin{cases} a \text{ with } a_1 \\ b \text{ with } b_1 \end{cases}$ in M_0, M_1, M_2

$$\frac{\partial q_2}{\partial \beta_z}=-b+\frac{M_1}{M_0}, \quad \frac{\partial^2 q_2}{\partial \beta_z^2}=\frac{M_2M_0-M_1^2}{M_0^2}$$

$$M_0=\sum_K e^{a_1\alpha+b_1\beta_z}, \quad M_1=\sum_K e^{a_1\alpha+b_1\beta_z}\cdot b_1, \quad M_2=\sum_K e^{a_1\alpha+b_1\beta_z}\cdot b_1^2$$

genest:

input: α, β_g, β_z



Define hidden variable $Y = (G, Z)$

$\theta^* = \underset{\theta}{\operatorname{argmax}} \log P(Y; \theta)$

EM: E step, given θ^{old}

$Q = E_{p(Y|X, \theta^{old})} [\log P(X, Y | \theta)]$

M step: $\theta^{new} = \underset{\theta}{\operatorname{argmax}} Q(\theta, \theta^{old})$

Monte-Carlo EM:

E step: $\tilde{Q}(\theta, \theta^{old}) = \frac{1}{M} \sum_m \log P(X, Y^m | \theta)$

Y^m is sample of $P(Y | X, \theta^{old})$

for a single sample Y^m,

$\log P(X, Y) = \log P(Y) + \log(X | Y)$

$E = -\log P(X, Y) = -\log P(Y) - \log P(X|Y)$

$E(X, Y) = E(Y) + E(X|Y)$

$E(Y) = E(G, Z)$

$$E(G,Z) = -\log p(G,Z)$$

$$p(G,Z) \approx \prod_{s\in G} p(g_s|\cdot) \cdot \prod_{s\in Z^j} p(z_s|\cdot)$$

$$= \prod_{s\in G} \frac{1}{Z_g} \exp\left\{ -\beta_g \sum_{r\in N_s} [g_r \neq g_s] - \alpha \sum_{j\in J} (g_s \neq z_s^j] \right\}$$ grp

$$\cdot \prod_{s\in Z^j} \frac{1}{Z_z} \exp\left\{ -\alpha [z_s \neq g_s] - \beta_z \sum_{r\in N_s} [z_s^j \neq z_r^j] \right\}$$ pub

$$Z_g = \sum_k \exp\left\{ -\beta_g \sum_{r\in N_s} [g_r \neq k] - \alpha \sum_{j\in J} [z_s^j \neq k] \right\}$$

$$Z_z = \sum_k \exp\left\{ -\alpha [g_s \neq k] - \beta_z \sum_{r\in N_s} [z_r^j \neq k] \right\}$$

Let $a_1 = -\sum_{j\in J} [g_s \neq z_s]$, $b_1 = -[g_s \neq g_r]$

$a_2 = -[z_s \neq g_s]$, $b_2 = -\sum_{r\in N_s} [z_s^j \neq z_r^j]$

$$E(Y) = E(G,Z) = -\log P(G,Z)$$

$$= \sum_{s\in G} \left(-a_1\alpha - b_1\beta_g + \log \sum_k e^{a_1\alpha + b_1\beta_g} \right)$$

$$+ \sum_{s\in Z^j} \left(-a_2\alpha - b_2\beta_z + \log \sum_k e^{a_2\alpha + b_2\beta_z} \right)$$