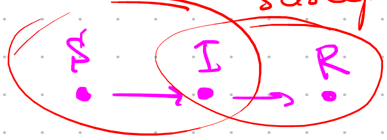
① SIR models

$\uparrow$ susceptible – infected – recovered / removed

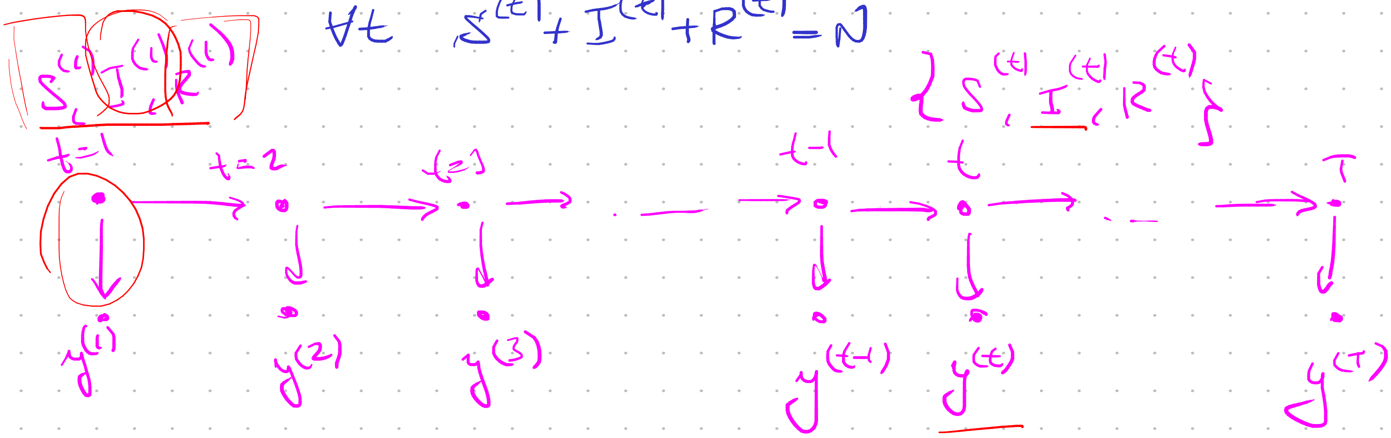


$$X = \{x_1, x_2, \dots, x_N\} \qquad y = \{y^{(1)}, y^{(2)}, \dots, y^{(T)}\}$$

$$x_i^{(t)} \in \{S, I, R\} \qquad y^{(t)} = \#\ \text{выявл. заболевш. во вр. } t$$

$$S^{(t)} = \#\{i \mid x_i^{(t)} = S\} \quad I^{(t)} = \text{---}\,''\text{---}\ I, \quad R^{(t)} = \text{---}\,''\text{---}\ R.$$

$$\forall t \quad S^{(t)} + I^{(t)} + R^{(t)} = N$$



Параметры модели:

1) $R^{(1)} = 0, \quad p(x_i^{(1)} = I) = \boxed{\pi}, \quad p(x_i^{(1)} = S) = 1 - \pi$

2) $p(x_i^{(t)} \in y^{(t)} \mid x_i^{(t)} = I) = \boxed{\rho}$

$$p(y^{(t)} \mid I^{(t)}, \rho) = \mathrm{Binom}(y^{(t)} \mid I^{(t)}, \rho)$$

3) $p(x_i^{(t+1)} = R \mid x_i^{(t)} = I) = \boxed{\mu}$

$$p(x_i^{(t+1)} = I \mid x_i^{(t)} = I) = 1 - \mu$$

4) $p(\text{заразиться от одного контакта}) = \boxed{\beta}$

$$p(x_i^{(t+1)} = S \mid x_i^{(t)} = S) = (1-\beta)^{I^{(t)}} \qquad \Theta = \{\pi, \rho, \mu, \beta\}$$

$$p(x_i^{(t+1)} = I \mid x_i^{(t)} = S) = 1 - (1-\beta)^{I^{(t)}}$$

$$\begin{array}{c|c}
 & \begin{array}{ccc} \quad S \quad & \quad\quad I \quad\quad & \quad R \end{array} \\
\begin{array}{c} S \\ I \\ R \end{array} & \begin{pmatrix} (1-\beta)^{I^{(t)}} & 1-(1-\beta)^{I^{(t)}} & 0 \\ 0 & 1-\mu & \mu \\ 0 & 0 & \boxed{1} \end{pmatrix}
\end{array} = p(x_i^{(t+1)} \mid x_i^{(t)})$$

② Inference

$$p(y, x \mid \bar\theta) = p(x^{(1)} \mid \pi)\, p(y^{(1)} \mid x^{(1)}, p)\, p(x^{(2)} \mid x^{(1)}, \beta, \mu) \ldots$$

$$\ldots\, p(x^{(T)} \mid x^{(T-1)}, \beta, \mu)\, p(y^{(T)} \mid x^{(T)}, p)$$

$$= \left[\pi^{I^{(1)}} (1-\pi)^{S^{(1)}}\right] \left[\binom{I^{(1)}}{y^{(1)}} p^{y^{(1)}} (1-p)^{I^{(1)} - y^{(1)}}\right] \left[\ldots\right] \times \ldots =$$

$$= \pi^{I^{(1)}} (1-\pi)^{S^{(1)}} \cdot \left[\prod_{t=1}^{T} \overbrace{\binom{I^{(t)}}{y^{(t)}}}^{\text{const}} p^{y^{(t)}} (1-p)^{I^{(t)} - y^{(t)}}\right] \times$$

$$\times \left[\prod_{t=1}^{T-1} \prod_{j=1}^{N} p(\overbrace{x_j^{(t+1)}}^{I, R} \mid \overbrace{x_j^{(t)}}^{I, J}, \mu, \beta, I_{-j}^{(t)})\right]$$

$$p(y \mid \bar\theta) = \int p(x, y \mid \bar\theta)\, dx = \int p(y \mid x, \theta)\, p(x \mid \theta)\, dx =$$

$$= E_{p(x \mid \bar\theta)}\left[p(y \mid x, \theta)\right] \approx \frac{1}{R} \sum_{r=1}^{R} p(y \mid x_r, \theta)$$

sampling

$\theta \to \max$

$$\theta_0 \xrightarrow{\text{Gibbs}} X_{0,r}$$

MAP

$$\theta_1 \xrightarrow{\text{Gibbs}} X_{1,r}$$

MAP

$$\theta_2 \longrightarrow X_{2,r}$$

$$\ldots$$

$$\theta^*$$

Gibbs sampling:

$$X = (\bar x_1, \bar x_2, \ldots, \bar x_j, \ldots, \bar x_N)$$

$$\bar x_j = (x_j^{(1)}, x_j^{(2)}, \ldots, x_j^{(T)})$$

$$\begin{cases} 1 \\ \ldots \\ j: \boxed{\bar x_j \sim p(\bar x_j \mid X_{-j}, y, \theta)} \\ \ldots \\ N \end{cases}$$

③ Stochastic Viterbi algorithm $\quad \bar{x}_j \sim p(\bar{x}_j \mid X_{-j}, y, \theta)$

$$
\begin{tikzcd}
x_j^{(1)} \arrow[r] \arrow[d] & x_j^{(2)} \arrow[r] \arrow[d] & \cdots \arrow[r] & x_j^{(t-1)} \arrow[r] \arrow[d] & x_j^{(t)} \arrow[r] \arrow[d] & \boxed{x_j^{(t+1)}} \arrow[r] & x_j^{(T)} \arrow[d] \\
y^{(1)} & y^{(2)} & & y^{(t-1)} & y^{(t)} & & y^{(T)}
\end{tikzcd}
$$

$a, a' \in \{S, I, R\}$

$$q^{(t)}_{j,a,a'} = p\left(x_j^{(t)} = a,\ x_j^{(t+1)} = a' \mid y^{(1)}, \ldots, y^{(t)}, X_{-j}, \bar{\theta}\right)$$

$$\bar{q}_j^{(t)} = \begin{array}{c|c} & \begin{array}{ccc} S & I & R \end{array} \\ \hline \begin{array}{c} S \\ I \\ R \end{array} & \begin{pmatrix} & & \\ & & \\ & & \end{pmatrix} \end{array}$$

$$p\left(x_j^{(t)} \mid x_j^{(t+1)}, \ldots, x_j^{(T)}, y, X_{-j}, \bar{\theta}\right) = p\left(x_j^{(t)} \mid x_j^{(t+1)}, y^{(1)} \ldots y^{(t)}, X_{-j}, \bar{\theta}\right)$$

$$= \frac{p\left(x_j^{(t)}, x_j^{(t+1)} \mid y^{(1)} \ldots y^{(t)}, X_{-j}, \bar{\theta}\right)}{\text{Cst}\ \boxed{p\left(x_j^{(t+1)} \mid \ldots\right)}} \propto q^{(t)}_{j, x_j^{(t)}, x_j^{(t+1)}}$$

$$q^{(t)}_{j,a,a'} = p\left(x_j^{(t)} = a, x_j^{(t+1)} = a' \mid y^{(1)} \ldots y^{(t)}, X_{-j}, \bar{\theta}\right) \propto$$

$$\begin{array}{c|c} & \begin{array}{ccc} S & I & R \end{array} \\ \hline \begin{array}{c} S \\ I \\ R \end{array} & \begin{pmatrix} \ldots & \ldots & 0 \\ 0 & 1-\mu & \mu \\ 0 & 0 & 1 \end{pmatrix} \end{array}$$

$$\propto p\left(x_j^{(t)} = a, x_j^{(t+1)} = a', y^{(t)} \mid y^{(1)} \ldots y^{(t-1)}, X_{-j}, \bar{\theta}\right) =$$

$$= p\left(x_j^{(t)} = a \mid y^{(1)} \ldots y^{(t-1)}, X_{-j}, \bar{\theta}\right) \cdot p\left(y^{(t)} \mid x_j^{(t)} = a, X_{-j}, \theta\right) \cdot p\left(x_j^{(t+1)} = a' \mid x_j^{(t)} = a, X_{-j}, \bar{\theta}\right)$$

$$\sum_{a''} p\left(x_j^{(t-1)} = a'', x_j^{(t)} = a \mid \ldots\right) = \qquad p\left(y^{(t)} \mid x_j^{(t)} = a, X_{-j}, \theta\right) = \mathrm{Binom}\left(y^{(t)} \mid I_{-j}^{(t)} + [a = I], f\right)$$

$$= \sum_{a''} q^{(t-1)}_{j, a'', a}$$

④ **Gibbs sampling**

① $Q_j^{(t)} = \left(- \; q^{(t)}_{j,a,a'} \; - \right)$ – stoch. Viterbi

② Sample $x_j^{(T)}, x_j^{(T-1)}, \dots, x_j^{(1)}$

$$x_j^{(T)} \sim p(x_j^{(T)} = a \mid y, X_{-j}, \theta) = \sum_{a'} q^{(T)}_{j,a,a'}$$

$$\vdots$$

$$x_j^{(t)} \sim p(x_j^{(t)} = a \mid x_j^{(t+1)} = a', y, X_{-j}, \theta) \propto q^{(t)}_{j,a,a'}$$

③ Output $\bar{x}_j = \left(x_j^{(1)} \dots x_j^{(T)} \right)$

⑤ **Learning**

1) Init θ, X

2) for $j = 1 \dots N, 1, \dots, N, \dots$

(i) $\bar{x}_j \sim$ Gibbs

(ii) replace $\bar{x}_j$ in X

(iii) every M steps:

update $\bar{\theta} = (\pi, \mu, \beta, \gamma)$ from X (M-max)

$\propto \pi^{a_\pi - 1} (1-\pi)^{b_\pi - 1}$

$$p(\bar{\theta} \mid X, y) \propto p(\bar{\theta})\, p(X, y \mid \bar{\theta})$$

$$p(\pi)\, p(\mu)\, p(\beta)\, p(\gamma)$$

$p(\pi) = \mathrm{Beta}(a_\pi, b_\pi)$

$$\log p(\bar{\theta} \mid X, y) = \text{Const} + (a_\pi - 1)\log \pi + (b_\pi - 1)\log(1-\pi) + I^{(1)} \log \pi + S^{(1)} \log(1-\pi)$$

π

$$+ (a_\mu - 1)\log \mu + (b_\mu - 1)\log(1-\mu) + \sum_{j=1}^{N} \sum_{t=1}^{T-1} \left([x_j^{(t)} = I, x_j^{(t+1)} = R] \log \mu + [x_j^{(t)} = I, x_j^{(t+1)} = I] \log(1-\mu) \right)$$

μ

$$+ (a_\gamma - 1)\log\gamma + (b_\gamma - 1)\log(1-\gamma)$$

$$+ \sum_{t=1}^{T} \left(y^{(t)} \cdot \log\gamma + (I^{(t)} - y^{(t)}) \log(1-\gamma) \right) \quad \Big] \; \gamma$$

$$+ (a_\beta - 1)\log\beta + (b_\beta - 1)\log(1-\beta)$$

$$+ \sum_{j=1}^{N} \sum_{t=1}^{T-1} \Big([x_j^{(t)} = S, x_j^{(t+1)} = S] \cdot I^{(t)} \cdot \log(1-\beta) + [x_j^{(t)} = S, x_j^{(t+1)} = I] \cdot ? \Big) \quad \Big] \; \beta$$

presence-only data

$I^{(t)}$ контактов с ≥ 1 заражённым

$$? = \Big(\underbrace{E\left[\text{конт., привели к заражен.}\right]}_{=\, I^{(t)} \cdot p(\text{заразился при 1 контакте} \,|\, \text{заразился}) \,=\, I^{(t)} \cdot \frac{\beta}{1-(1-\beta)^{I^{(t)}}}} \cdot \log\beta + \underbrace{E\left[\text{—"— не —"—}\right]}_{=\, I^{(t)}\left(1 - \frac{\beta}{1-(1-\beta)^{I^{(t)}}}\right)} \cdot \log(1-\beta) \Big)$$

$$p(\pi | \ldots) = \text{Beta}(a'_\pi, b'_\pi)$$

$$a'_\pi = a_\pi + I^{(1)}, \quad b'_\pi = b_\pi + S^{(1)}$$

$$a'_\mu = a_\mu + \sum_{j=1}^{N} \sum_{t=1}^{T-1} [x_j^{(t)} = I, x_j^{(t+1)} = R] \qquad b'_\mu = b_\mu + \sum_{j=1}^{N} \sum_{t=1}^{T-1} [x_j^{(t)} = I, x_j^{(t+1)} = I]$$

$$a'_\gamma = a_\gamma + \sum_{t=1}^{T} y^{(t)} \qquad b'_\gamma = b_\gamma + \sum_{t=1}^{T} (I^{(t)} - y^{(t)})$$

$$a'_\beta = a_\beta + \sum_{j=1}^{N} \sum_{t=1}^{T-1} [x_j^{(t)} = S, x_j^{(t+1)} = I] \cdot I^{(t)} \cdot \frac{\beta}{1-(1-\beta)^{I^{(t)}}}$$

$$b'_\beta = b_\beta + \sum_{j=1}^{N} \sum_{t=1}^{T-1} [x_j^{(t)} = S] \left([x_j^{(t+1)} = S] \cdot I^{(t)} + [x_j^{(t+1)} = I] \cdot I^{(t)} \left(1 - \frac{\beta}{1-(1-\beta)^{I^{(t)}}}\right) \right)$$

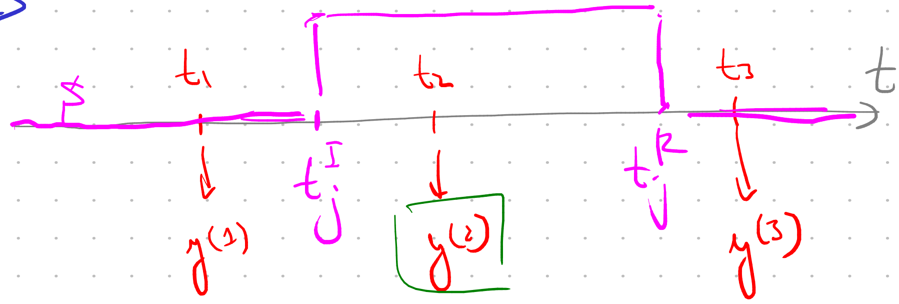
6 Extensions

1) SIR → SEIR → SEIRS

exposed

SIR → SIRS → SEIRS

2) Continuous time



3) Модель заражения

β

$$p(x_j^{(t+1)} = S \mid x_j^{(t)} = S) = (1-\beta)^{|I^{(t)} \cap Nei(j)|}$$

---

1 EM-algorithm

$$p(X|\bar\theta) \to \max \qquad p(X|\bar\theta) = \int \overbrace{p(X,Z|\bar\theta)}^{\text{проще}} dZ$$

$$p(X,Z|\bar\theta) = p(X|\bar\theta)\cdot p(Z|X,\bar\theta)$$

$$\log p(X|\bar\theta) = \log p(X,Z|\bar\theta) - \log p(Z|X,\bar\theta) \quad \Big\} \; E_Z$$

$$\log p(X|\bar\theta) = E_{p(Z|X,\bar\theta^{(m)})}\left[\log p(X,Z|\bar\theta) - \log p(Z|X,\bar\theta)\right]$$

$$= \int \log p(X,Z|\bar\theta)\, p(Z|X,\bar\theta^{(m)})\, dZ - \int \log p(Z|X,\bar\theta)\, p(Z|X,\bar\theta^{(m)})\, dZ$$

$$+ \int \log p(Z|X,\bar\theta^{(m)})\, p(Z|X,\bar\theta^{(m)})\, dZ - \int \ldots dZ$$

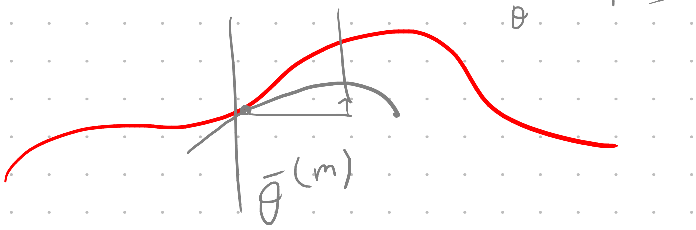
$\bar\theta \to \max$

$Q(\bar\theta, \bar\theta^{(m)})$

$$= \underbrace{\int \log \frac{p(X,Z|\bar\theta)}{p(Z|X,\bar\theta^{(m)})}\, p(Z|X,\bar\theta^{(m)})\, dZ}_{h(\bar\theta,\bar\theta^{(m)})} + \underbrace{\int \log \frac{p(Z|X,\bar\theta^{(m)})}{p(Z|X,\bar\theta)}\, p(Z|X,\bar\theta^{(m)})\, dZ}_{KL(p(Z|X,\bar\theta^{(m)}) \,\|\, p(Z|X,\bar\theta))}$$

$$\log p(X|\bar\theta) = \mathcal{L}(\bar\theta, \bar\theta^{(m)}) + \underbrace{KL\left(p(Z|X,\bar\theta^{(m)}) \,\|\, p(Z|X,\bar\theta)\right)}_{\geq 0,\ =0 \text{ if } \bar\theta = \bar\theta^{(m)}}$$

$$\underline{\log p(X|\bar\theta)} \geq \mathcal{L}(\bar\theta, \bar\theta^{(m)}) \xrightarrow[\bar\theta]{} \max$$



② Variational approximation

$\theta \to p(x, z) \leftarrow \bar\theta, z$

$p(z|x) = ?$

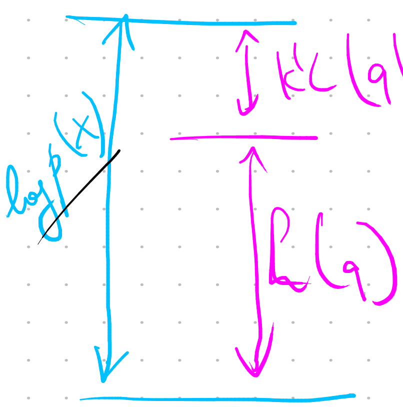
$$p(x,z) = p(x)\, p(z|x)$$

$$\log p(x) = \log p(x,z) - \log p(z|x) \quad \Big|\ \mathbb{E}_{q(z)}$$

$$\log p(x) = \mathbb{E}_{q(z)}[\log p(x,z)] - \mathbb{E}_{q(z)}[\log p(z|x)] \pm \mathbb{E}_{q(z)}[\log q(z)]$$

$$\log p(x) = \int \log \frac{p(x,z)}{q(z)}\, q(z)\, dz + \int \log \frac{q(z)}{p(z|x)}\, q(z)\, dz$$

$$\underbrace{\log p(x)}_{\text{evidence}} = \underbrace{\mathcal{L}(q)}_{\substack{\text{evidence lower bound (ELBO)} \ /\ \text{variational lower bound}}} + KL(q(z) \,\|\, p(z|x))$$



$$KL(q\|p(z|x)) \xrightarrow[q]{} \min$$

$$\Updownarrow$$

$$\xrightarrow[q]{} \max$$

Variational approximation

$$q(z) \approx p(z|x)$$

$$\Updownarrow$$

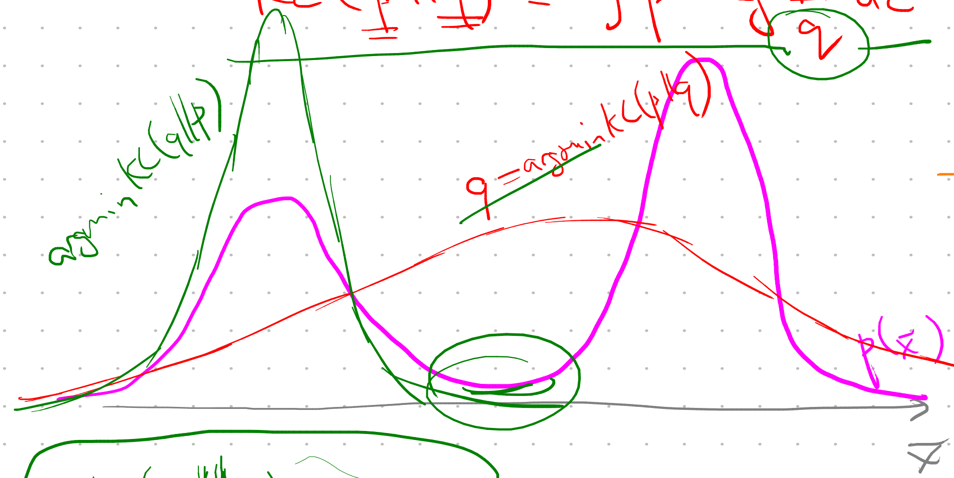
$$\mathcal{L}(q) \xrightarrow[q]{} \max$$

③ KL divergence

$$KL(p\|q) = \int p \log \frac{p}{q} \, dz$$

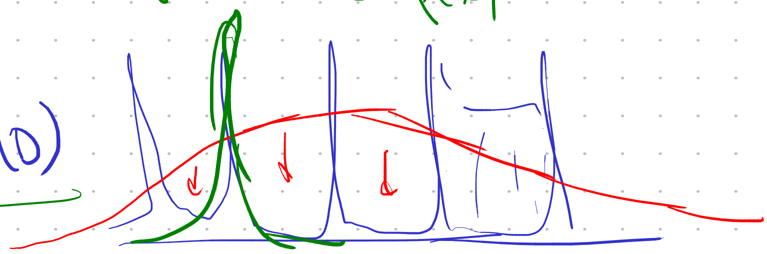
$KL(p\|q) \to \min$

$KL(q\|p) \to \min$



$$KL(q\|p) \to \min = \int q(\bar{x}) \log \frac{q(\bar{x})}{p(\bar{x})} d\bar{x}$$

$p(\bar{\theta}|D)$



$$KL(p\|q) \to \min = \int p \log \frac{p}{q} d\bar{x} = \int p \log p \, d\bar{x} - \int p \log q \, d\bar{x} = \text{const} - \int p \log q \, d\bar{x}$$

$p = p_{data}$

$$\int p \log q \, d\bar{x} = \underbrace{\sum_n \log q(\bar{x}_n)}_{\to \max}$$

④

$$q(z) = \prod_{i=1}^{M} q_i(z_i), \quad \text{где } z_i \cap z_j = \emptyset, \quad z_i \subseteq z$$

$$L(q) = \int \log \frac{p(x,z)}{\prod_{i=1}^M q_i(z_i)} \cdot \prod_{i=1}^{M} q_i(z_i) \, dz_1 \ldots dz_M =$$

$$= \int \left(\prod q_i\right) \cdot \log p(x,z) \, dz - \sum_{i=1}^{M} \int \log q_i(z_i) \cdot \prod_{j=1}^{M} q_j(z_j) \, dz_1 \ldots dz_M$$

$$\int q_i \log q_i \, dz_i$$

$$\int \log q_i(z_i) \prod_j q_j(z_j) \, dz = \left(\int q_i \log q_i \, dz_i\right) \cdot \underbrace{\prod_{j \neq i} \int q_j(z_j) \, dz_j}_{=1}$$

$$\mathcal{L}(q) = \underbrace{\int \log p(x,z) \cdot \prod_{i=1}^{M} q_i(z_i)\, dz}\; - \underbrace{\sum_{j=1}^{M} \int q_j(z_j) \log q_j(z_j)\, dz_j}$$

$$\mathcal{L}(q) \xrightarrow[q_j(z_j)]{} \max \quad !$$

$$\mathcal{L}(q_j) = \text{Const} - \int q_j \log q_j \, dz_j + \int q_j(z_j) \cdot \underbrace{\left[\int \log p(x,z) \cdot \prod_{i \neq j} q_i \, dz_{-j} \right]} dz_j$$

$$= \text{Const} - \int q_j \log \frac{q_j}{e^{\int \dots}} \, dz_j \xrightarrow{q_j} \max$$

$q_j^*(z_j) = \tilde{p}(z_j)$, где

$$\log \tilde{p}(z_j) = \mathbb{E}_{q_{-j}(z_{-j})} \left[\log p(x,z) \right]$$

$$\boxed{\log q_j^*(z_j) = \mathbb{E}_{q_i^*(z_i),\, i \neq j} \left[\log p(x,z) \right] + \text{Const}}$$