

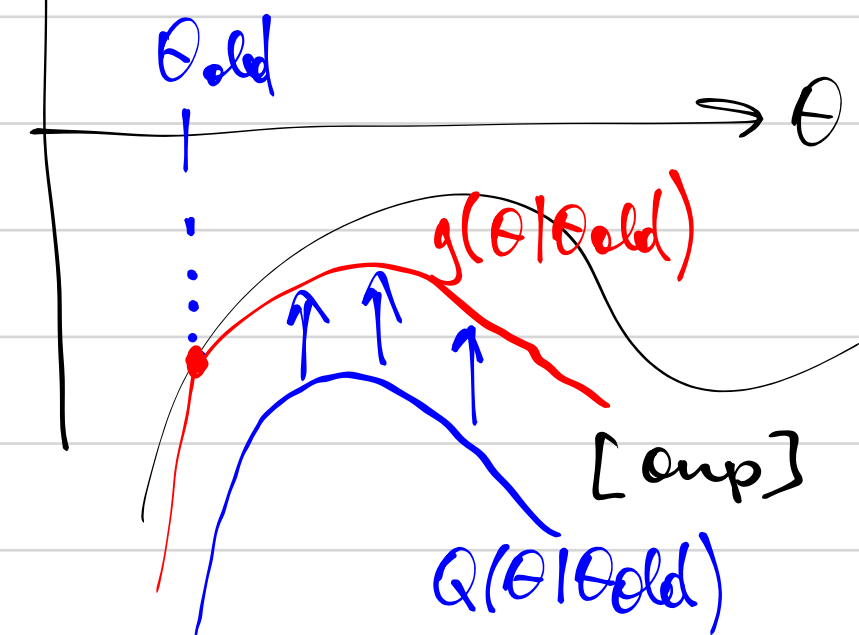
Проблем!

Брегао? Уммуно?

L7. EM2

$\ell(y|\theta)$

$z_i \in \{1, \dots, n_{class}\}$



E. $p(z|y, \theta_{old})$ Тајмса туча $n_{obs} \times n_{class}$

$$Q(\theta|\theta_{old}) = E[\ln p(y, z|\theta) | y, \theta_{old}]$$

$$M. \max_{\theta} Q(\theta|\theta_{old})$$

програмирање!

Урб. б роке $\theta = \theta_{old}$ и права $\ln p(y|\theta) = g(\theta|\theta_{old})$

вспомните нуволек гон-ба. $\ln E \geq E \ln$

$$\ell(y|\theta_{old}) = \dots = \ln E \left[\frac{p(y, z|\theta)}{p(z|y, \theta_{old})} \middle| y, \theta_{old} \right] \geq$$

$$\geq E \ln \left[\frac{p(y, z|\theta)}{p(z|y, \theta_{old})} \right] \quad \text{при } \theta = \theta_{old}$$

$$\frac{p(y, z|\theta_{old})}{p(z|y, \theta_{old})} = p(y|\theta_{old})$$

б роке
 $\theta = \theta_{old}$

$$\ln E \left[p(y|\theta_{old}) \middle| y, \theta_{old} \right] = \ln p(y|\theta_{old})$$

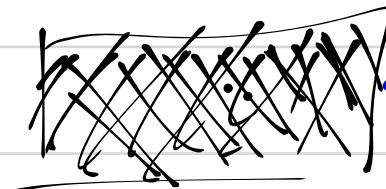
Дайте определение EM-алгоритма
 для заданных q -м $p(A|B)$ ↓

мысли
 $E \quad p(z|y, \theta_{old})$ $q(z)?$
 n_{obs} n_{clust}

E-step заменили на оптимизацию!

E' :

$\max_{q(z)}$



но в типичном случае

$q^*(z) = p(z|y, \theta_{old})$

M-шаг: $\max_{\theta} Q(\theta|\theta_{old})$

[?] ^{прямая} ценность с прямой-ой г. зрения в шаг [?]

мы?

Наиболее решение:

$\max_{q(z)} \left[-KL(q(z) || p(z|y, \theta_{old})) \right]$

$q^* = p(z|y, \theta_{old})$

$KL \geq 0$

$KL(q || q) = 0$

Q. почему не пройдет наиболее решение?

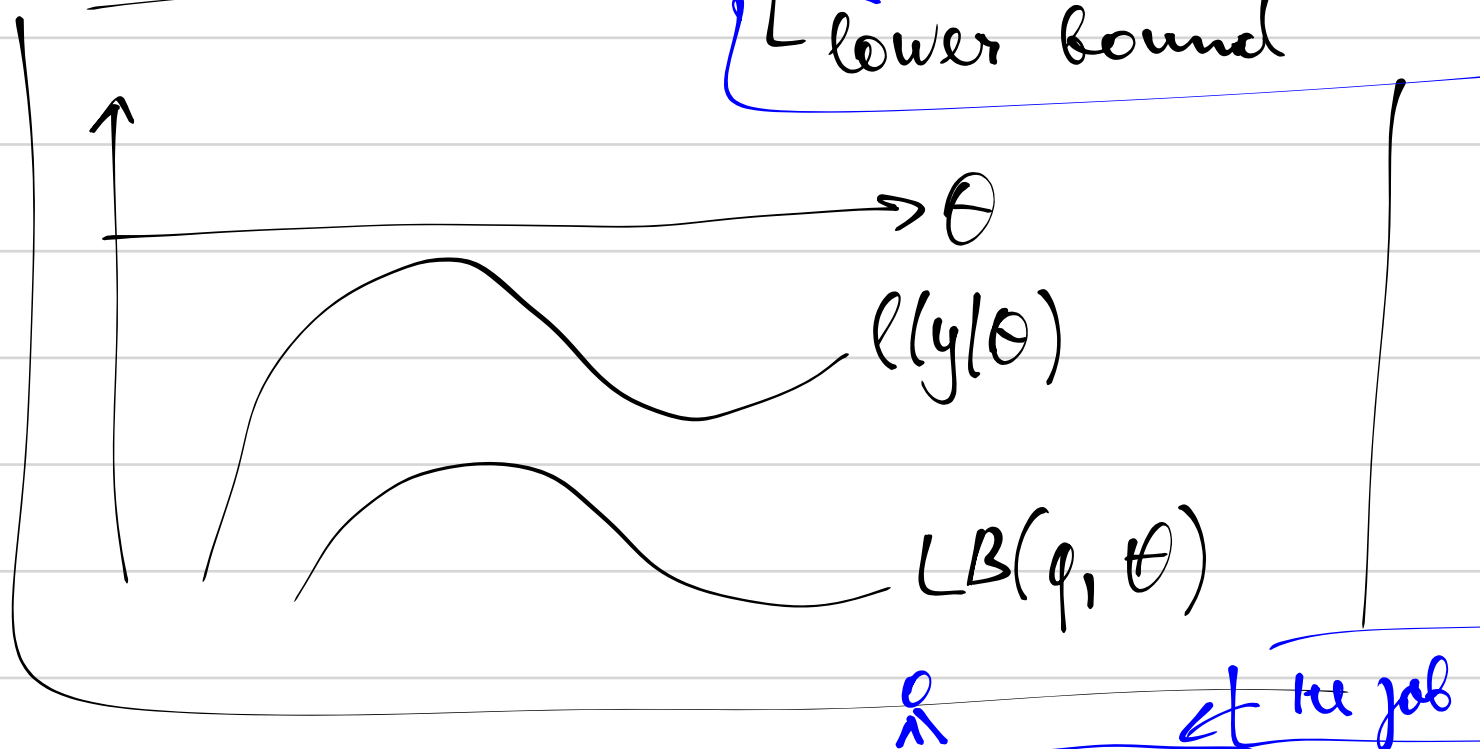
мы?

A. чтобы вычислить эту KL

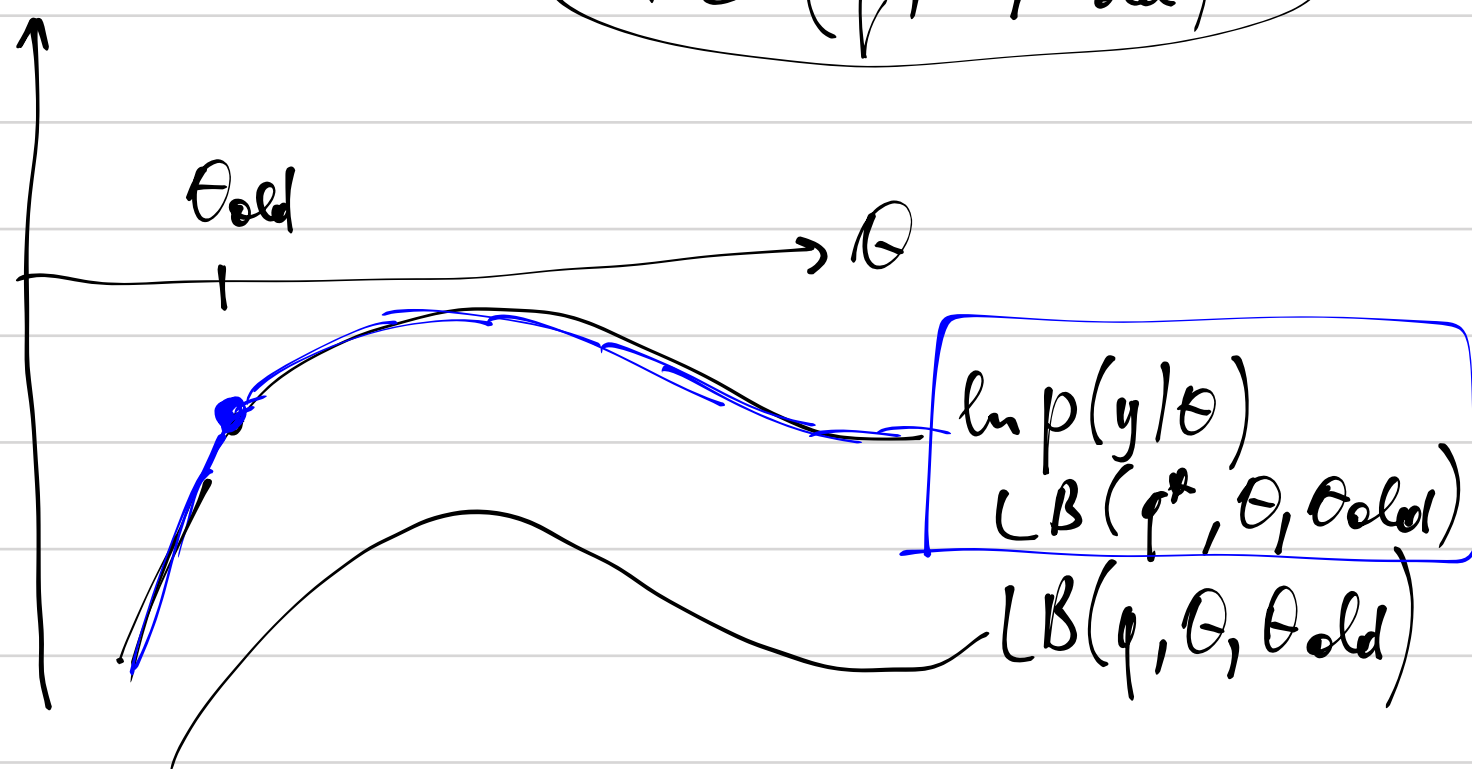
нужно уже $p(z|y, \theta_{old})$ надо
 знать, а мы её не знаем!

(тут же!) (ratio num)

$$\ell(y|\theta) = \ln p(y|\theta) = \underbrace{KL(q(z) \| p(z|y, \theta))}_{\text{lower bound}} + \underbrace{LB(q, \theta)}_{\text{or } y \text{ some function.}}$$



$$\ln p(y|\theta) = \underbrace{KL(q \| p(z|y, \theta_{old}))}_{\text{the job of } \theta} + \underbrace{LB(q, \theta, \theta_{old})}_{\text{the job of } \theta}$$



$$q^* = \underset{q}{\operatorname{argmin}} KL(q \| p(z|y, \theta_{old}))$$

$$\begin{aligned}
 LB(q, \theta) &= \ln p(y|\theta) - \underbrace{KL(q \| p(z|y, \theta))}_{\substack{\forall \\ 0}} = \\
 &= \ln p(y|\theta) + \sum_z q(z) \cdot \ln \frac{p(z|y, \theta)}{q(z)} = \\
 &= \sum_z q(z) \cdot \ln p(y|\theta) + \sum_z q(z) \ln \frac{p(z|y, \theta)}{q(z)} = \\
 &= \sum_z q(z) \cdot \ln \left[\frac{p(z|y, \theta) \cdot p(y|\theta)}{q(z)} \right] = \\
 &= \sum_z q(z) \cdot \ln \frac{p(y, z|\theta)}{q(z)} \quad \text{выгодо} \quad // \\
 &\quad \text{max}_{\theta} LB(q, \theta)
 \end{aligned}$$

$q(z)$ — выбранная q -функция.
 известная!

worse \rightarrow bad π_b
 better \rightarrow good π_g
 $(y_i | z_i = g) \sim \mathcal{N}(\mu_g; 1)$
 $(y_i | z_i = b) \sim \mathcal{N}(\mu_b; 1)$

$$\begin{aligned}
 p(y_i = 2, z_i = b | \theta) &= \\
 &= \pi_b \cdot \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{(2-\mu_b)^2}{2}\right)
 \end{aligned}$$

Новая q -функция EM-алгоритма.
 Init. $\theta_{old} = \{\text{параметры}\}$
 $E \rightarrow E_{max}$ (по сути тоже M-max)

$$\begin{aligned}
 LB(q, \theta) &= \sum_z q(z) \cdot \ln \frac{p(y, z|\theta)}{q(z)} \\
 &= -KL(q(z) \| p(y, z|\theta))
 \end{aligned}$$

известная!
 $\left[-KL(q(z) \| p(z|y, \theta)) \right] \quad \max_{\theta} LB(q, \theta_{old})$

M-uar

$$\max_{\theta} Q(\theta | \theta_{old})$$

$$Q(\theta | \theta_{old}) = \sum_z p(z | y, \theta_{old}) \cdot \ln p(y, z | \theta)$$

M'-uar

$$LB(q, \theta) = \sum_z q(z) \cdot \ln \frac{p(y, z | \theta)}{q(z)}$$

c nreby gnyeno uona

$$q^* = \underset{q}{\operatorname{argmax}} LB(q, \theta_{old})$$

$$q^* = p(z | y, \theta_{old})$$

$$LB(q^*, \theta) = \sum_z p(z | y, \theta_{old}) \cdot \ln \frac{p(y, z | \theta)}{p(z | y, \theta_{old})}$$

$$Q(\theta | \theta_{old}) = \sum_z p(z | y, \theta_{old}) \cdot \ln p(y, z | \theta)$$

$$LB - Q = \sum p(z | y, \theta_{old}) \cdot \ln \frac{1}{p(z | y, \theta_{old})}$$

const [ne job or θ]

$$LB - Q = H(p(z | y, \theta_{old}))$$

$$M': \max_{\theta} LB(q \uparrow \text{c nreby gnyeno uona}, \theta)$$

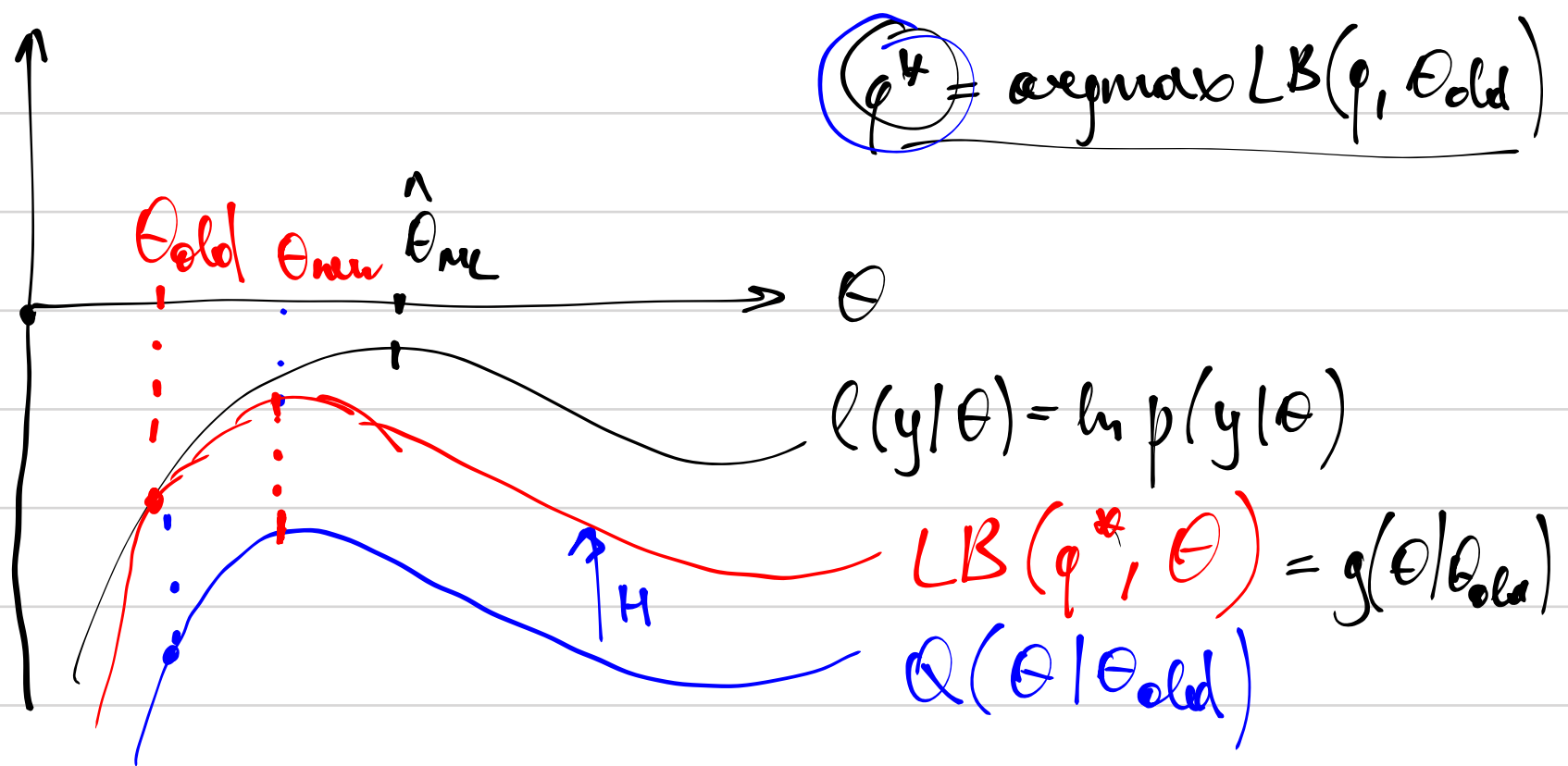
Upro:

Init θ_{old}

$$\rightarrow E' \quad q = \underset{q}{\operatorname{argmax}} LB(q, \theta_{old})$$

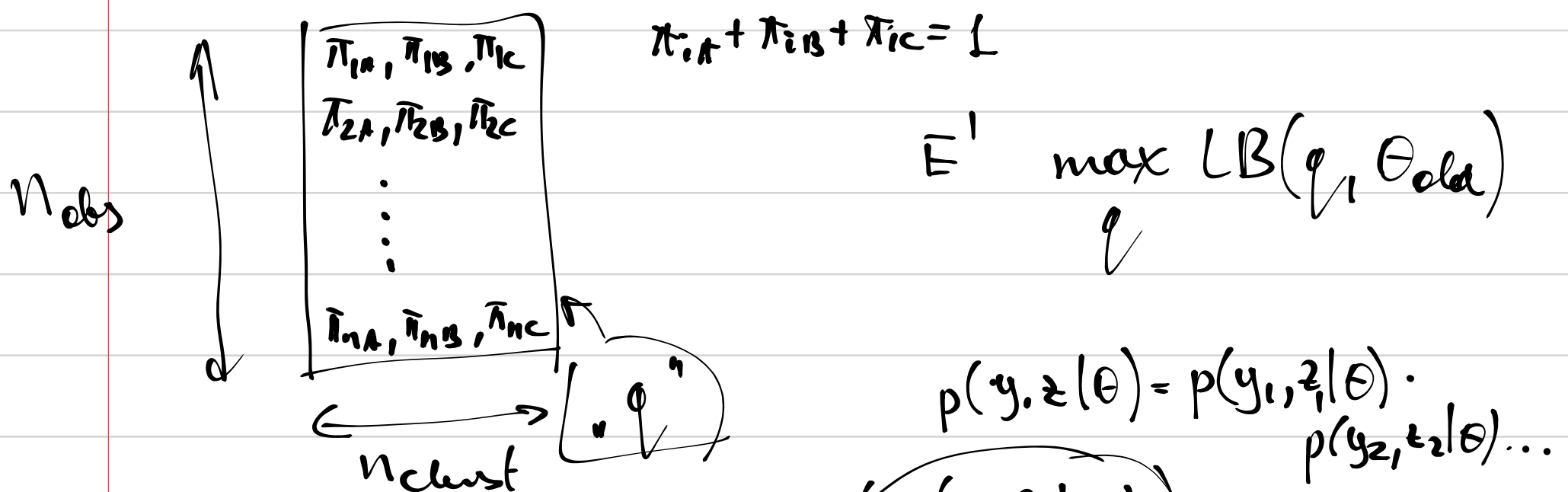
$$M' \quad \theta_{new} = \underset{\theta}{\operatorname{argmax}} LB(q, \theta)$$

noha ke kagolen $\theta_{old} = \theta_{new}$ u nreby gnyeno u E'



$$LB(\varphi^*, \theta) = l(y|\theta) - \underbrace{KL(\varphi^* || p(z|y, \theta))}_{\substack{= 0 \text{ при } \theta = \theta_{old} \\ > 0 \text{ при } \theta \neq \theta_{old}}}$$

3 класт: A, B, C



$$LB(\varphi, \theta) = \sum_z \underbrace{q(z)} \cdot \ln \left(\frac{p(y, z|\theta)}{q(z)} \right)$$

$$z = \begin{pmatrix} z_1 \\ z_2 \\ \vdots \\ z_n \end{pmatrix}$$

$$q(z_1, \dots, z_n) = q(z_1) \cdot q(z_2) \cdot \dots \cdot q(z_n)$$

при независимых наблюдениях задача оптимизации от упрощается.

ещё картинка (от. схем - чная)

