$\ell(x|\theta) \to \max\limits_{\Theta}$; $\hat{\Theta}_{ML}$

z – латентная (ненабл.) перем.

$\in 1 \dots K$

$$\underbrace{\ell(x|\theta)}_{\text{неполное}} \longrightarrow \underbrace{\ell(x,z|\Theta)}_{\text{полное}} \qquad p(x) \nearrow \mathbb{P}\{X=x\}, \quad p(x) \searrow f_X(x)$$

$$\ell(x|\theta) = \sum_i \ln p(x_i|\Theta) = \sum_i \sum_j \mathbb{P}\{z=j\} \ln \underbrace{p(x_i|\Theta)}{} =$$

$$= \sum_i \sum_j \mathbb{P}\{z=j\} \ln \frac{p(z=j, x_i|\Theta) \cdot \mathbb{P}\{z=j\}}{p(z=j|x_i,\Theta) \cdot \mathbb{P}\{z=j\}} =$$

$$= \sum_i \sum_j \mathbb{P}\{z=j\} \ln \frac{p(z=j, x_i|\Theta)}{\mathbb{P}\{z=j\}} +$$

$M(p(z), p(z,x|\Theta))$

$$+ \underbrace{\sum_i \sum_j \mathbb{P}\{z=j\} \ln \frac{\mathbb{P}\{z=j\}}{p(z=j|x_i,\Theta)}}_{KL(\underbrace{p(z)}{} \,\|\, \underbrace{p(z|X,\Theta)}{})}$$

$$\ell(x|\Theta) = \underbrace{M(p(z), p(z,x|\Theta))}_{\text{нижняя оценка на } \ell(x|\Theta)} + \underbrace{KL(p(z) \| p(z|X,\Theta))}_{\geq 0}$$

$$\begin{array}{cccc} z_1 & z_2 & \dots & z_k \\ p_1 & p_2 & \dots & p_k \end{array}$$

E-шаг: $M(p(z), p(z,x|\Theta)) \to \max\limits_{p(z)}$

$$\Longleftrightarrow$$

$$KL(p(z) \| p(z|X,\Theta)) \to \min\limits_{p(z)}$$

$$\Rightarrow p(z)^* = p(z|X, \Theta_{old}) = g$$

M-шаг: $M(\dots, \dots) \to \max\limits_{\Theta}$

$$M = \sum_i \sum_j \underbrace{\mathbb{P}\{z=j\}}_{p(z)^* = g_{ij}} \ln \underline{p(z=j, x_i|\Theta)} \sim$$

$$\underbrace{\mathbb{P}\{z=j\}}_{\quad} \quad p(z)^* = g_{ij}$$

$$\sim \sum_i \sum_j \underbrace{\mathbb{P}\{z=j\}}_{g_{ij}} \ln p(z=j, x_i|\Theta) \to \max_{\Theta}$$

$$= Q(\Theta, \Theta_{old})$$

EM-алгоритм

Init: $\Theta_{old} = [\underbrace{-\ \Theta\ -}_{\text{парам. модели}}, \underbrace{p_1 \dots p_k}_{p(z)}]$

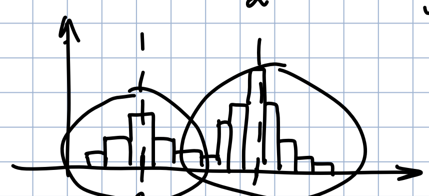
E-шаг: $\boxed{p(z)^* = p(z \mid X, \Theta_{old})}$ $\to$ $p(z=1)^*$ $\vdots$ $p(z=k)^*$

M-шаг: $\boxed{Q(\Theta, \Theta_{old})} \to \Theta$

Разделение смеси норм. распред.

$$X_1 \dots X_n \xrightarrow{z=1} N(\mu_1, \sigma_1^2) \qquad z \in \{1, 2\}$$

$$X_1 \dots X_n \xrightarrow{z=2} N(\mu_2, \sigma_2^2)$$



$\boxed{p_1 := \mathbb{P}\{z=1\}}$

$p_2 := \mathbb{P}\{z=2\} = 1 - p_1$

Init: $\Theta_{old} := [\underbrace{\mu_1}_{3};\ \sigma_1^2;\ \underbrace{\mu_2}_{10};\ \sigma_2^2;\ \underbrace{p_1}_{0.5}]$

E-шаг: $p(z=1)^* := \underbrace{p(z=1 \mid X, \Theta)}$

$$\underbrace{p(z_i = 1 \mid x_i, \Theta)} = \frac{p(z_i = 1, x_i \mid \Theta)}{p(x_i \mid \Theta)} =$$

$$= \frac{p(x_i \mid z_i = 1, \Theta) \cdot p(z_i = 1)}{\underbrace{p(x_i \mid z_i = 1, \Theta)} \cdot \underbrace{p(z_i = 1)} + p(x_i \mid z_i = 2, \Theta)\, p(z_i = 2)} =$$

$$= \frac{\frac{1}{\sqrt{2\pi\sigma_1^2}} e^{-\frac{1}{2}\left(\frac{x_i - \mu_1}{\sigma_1}\right)^2} \cdot p_1}{\left[\frac{1}{\sqrt{2\pi\sigma_1^2}} e^{-\frac{1}{2}\left(\frac{x_i - \mu_1}{\sigma_1}\right)^2} \cdot p_1 + \frac{1}{\sqrt{2\pi\sigma_2^2}} e^{-\frac{1}{2}\left(\frac{x_i - \mu_2}{\sigma_2}\right)^2} \cdot (1 - p_1)\right]} =$$

$$= g_i$$

<u>М-шаг</u>: $Q(\Theta, \Theta_{old})$

$$Q(\Theta, \Theta_{old}) = \sum_i \sum_j g_{ij} \times \ln p(z = j, x_i \mid \Theta) \to \max_{\Theta}$$

$$Q = \sum_i \overset{P\{z=1\}}{g_i} \ln\left(\underbrace{p(x_i \mid z = 1, \Theta) \cdot p_1}\right) +$$

$$+ (1 - g_i) \ln\left(p(x_i \mid z = 2, \Theta) \cdot p_2\right) =$$

$$\parallel \quad P\{z=2\} = 1 - P\{z=1\}$$

$$= \sum_i g_i \left(\ln \frac{1}{\sqrt{2\pi\sigma_1^2}} e^{-\frac{1}{2}\left(\frac{x_i-\mu_1}{\sigma_1}\right)^2} + \ln p_1 \right) +$$

$$+ (1-g_i) \left(\ln \frac{1}{\sqrt{2\pi\sigma_2^2}} e^{-\frac{1}{2}\left(\frac{x_i-\mu_2}{\sigma_2}\right)^2} + \ln (1-p_1) \right).$$

$$\to \max_{\Theta} \equiv \max_{\mu_1, \sigma_1^2, \mu_2, \sigma_2^2, p_1}$$

$$Q'_{\mu_1} = \left[\sum_i g_i \left(-\frac{1}{2} \ln (2\pi\sigma_1^2) - \frac{1}{2} \left(\frac{x_i - \mu_1}{\sigma_1} \right)^2 \right]'_{\mu_1} =$$

$$= \sum_i g_i \frac{(x_i - \mu_1)}{\sigma_1^2}$$

$$\sum_i g_i \frac{(x_i - \hat{\mu}_1)}{\cancel{\hat{\sigma}_1^2}} = 0 \Rightarrow \sum_i g_i x_i - \sum_i g_i \hat{\mu}_1 = 0$$

$$\hat{\mu}_1 = \frac{\sum_i g_i x_i}{\sum_i g_i}$$

$$\Theta_{new} := (\hat{\mu}_1, \hat{\sigma}_1^2, \hat{\mu}_2, \hat{\sigma}_2^2, \hat{p}_1)$$

$\hookrightarrow$ E-war

$$\hat{\mu}_2 = \frac{\sum_i (1-g_i) x_i}{\sum_i (1-g_i)}$$

$$\hat{\sigma}_1^2 = \dots \quad , \hat{\sigma}_2^2 = \dots$$

$$\hat{p}_1 = \dots$$