$\{\bar{x}_1, \dots \bar{x}_N\} = X$ – данные $\quad p(X|\theta) \to \max$ трудно

$\{\bar{z}_1, \dots, \bar{z}_N\} = Z$ – скрытые перем. $\quad p(X, Z|\theta) \to \max$ легко

$$\log p(X|\theta) - \underbrace{\log p(X|\theta^{(m)})}_{= \ell^{(m)}} =$$

max

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

$$= \underbrace{\log}_{f} \int \left(\frac{p(X|Z,\theta)p(Z|\theta)}{p(Z|X,\theta^{(m)})} \right) p(Z|X,\theta^{(m)})\, dz - \ell^{(m)} \geq$$

$\mathbb{E}_{p(Z|X,\theta^{(m)})}[\dots]$

$f(\alpha x + (1-\alpha) y) \geq$
$\alpha f(x) + (1-\alpha) f(y)$

$f(\mathbb{E}\, g(x)) \geq \mathbb{E}[f(g(x))]$

$f\left(\int g(x) p(x) dx\right) \geq$
$\geq \int p(x) f(g(x)) dx$

$$\geq \int p(Z|X,\theta^{(m)}) \log \frac{p(X|Z,\theta)p(Z|\theta)}{p(Z|X,\theta^{(m)})}\, dz - \log p(X|\theta^{(m)})$$

$$= \int \log \frac{\overbrace{p(X|Z,\theta)p(Z|\theta)}^{= p(X,Z|\theta)}}{\underbrace{p(Z|X,\theta^{(m)})\, p(X|\theta^{(m)})}_{= p(Z,X|\theta^{(m)})}}\, p(Z|X,\theta^{(m)})\, dz$$

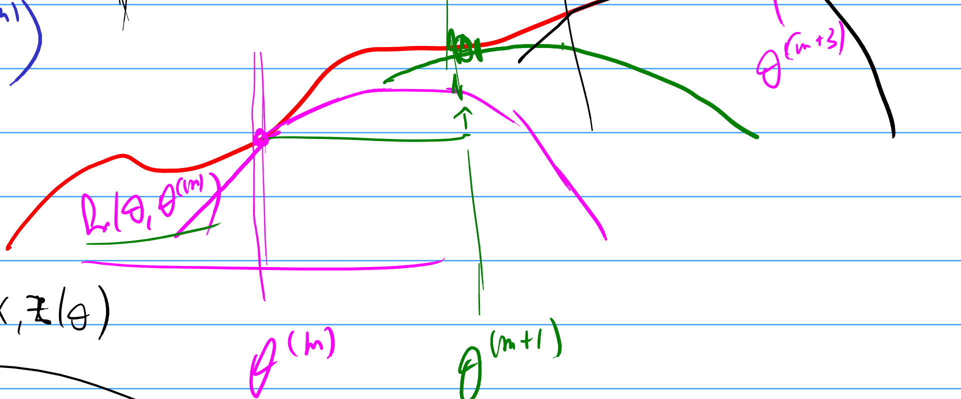
$$\log p(X|\theta) \geq \underbrace{\log p(X|\theta^{(m)}) + \int \text{—}\, dz}_{= h(\theta,\theta^{(m)})}$$

$\forall \theta \quad \ell(\theta) \geq h(\theta, \theta^{(m)})$

$\ell(\theta^{(m)}) = h(\theta^{(m)}, \theta^{(m)})$

$$\theta^{(m+1)} = \arg\max_\theta h(\theta, \theta^{(m)})$$

MM-algorithms
minorization-
maximization

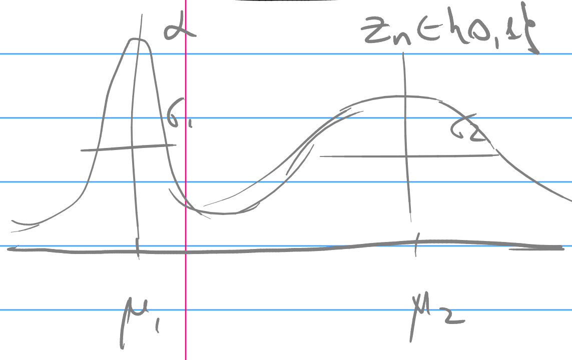


$\max_\theta$

$$h(\theta,\theta^{(m)}) = \int \log \frac{\overbrace{p(X|Z,\theta)p(Z|\theta)}^{= p(X,Z|\theta)}}{p(Z|X,\theta^{(m)})\, p(X|\theta^{(m)})}\, p(Z|X,\theta^{(m)})\, dz$$

EM-algorithm: M-шаг: $\theta^{(m+1)} := \arg\max_{\theta} Q(\theta, \theta^{(m)})$, где

E-шаг: $Q(\theta, \theta^{(m)}) = \int \log p(X, Z|\theta)\, p(Z|X, \theta^{(m)})\, dZ$



$z_n \in \{0, 1\}$

$p(x_n|\theta) = \alpha N(x_n|\mu_1, \sigma_1) + (1-\alpha) N(x_n|\mu_2, \sigma_2)$

$p(x_n, z_n|\theta) = \left[\alpha N(x_n|\mu_1, \sigma_1)\right]^{z_n} \left[(1-\alpha) N(x_n|\mu_2, \sigma_2)\right]^{1-z_n}$

$\log p(X, Z|\theta) = \log \prod_n p(x_n, z_n|\theta) =$

$= \sum_n \left[z_n \log \alpha N(x_n|\mu_1, \sigma_1) + (1-z_n) \log\left[(1-\alpha) N(x_n|\mu_2, \sigma_2)\right]\right]$

$\dots, z_n^2$

$\mathbb{E}_{\dots|\theta^{(m)}}[z_n^2]$

$Q(\theta, \theta^{(m)}) = \int \underbrace{\sum_n [\dots]}_{\text{лин. комб. } z_n} p(Z|X, \theta^{(m)})\, dZ =$

$= \sum_n \left(\log \alpha N(x_n|\mu_1, \sigma_1) \cdot \mathbb{E}_{Z|X, \theta^{(m)}}[z_n] + \log\left((1-\alpha) N(x_n|\mu_2, \sigma_2)\right)\left(1 - \mathbb{E}_{Z|X, \theta^{(m)}}[z_n]\right)\right)$

E-шаг: $\mathbb{E}_{Z|X, \theta^{(m)}}[z_n] = p(z_n = 1|X, \theta^{(m)}) = \dfrac{\alpha^{(m)} N_1(x_n)}{\alpha^{(m)} N_1(x_n) + (1-\alpha)^{(m)} N_2(x_n)}$

---

1) $Q(\theta, \theta^{(m)}) \to \max$

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

$Y = f(X, Z)$

$y_n = f(x_n, z_n)$ $\quad \int \log p(f(X, Z)|\theta)\, p(Z|X, \theta^{(m)})\, dZ$

$p(X|\underset{(X,Z)}{Y}, \theta) = p(X|Y)$

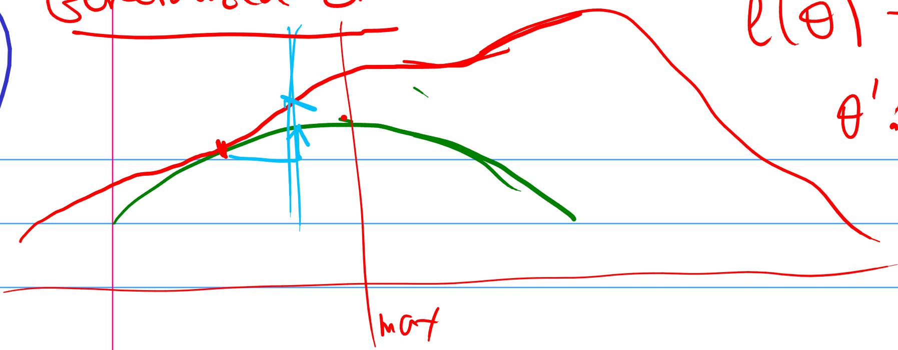
$\prod_n p(x_n, z_n|\theta)$

$Q(\theta, \theta^{(m)}) = \mathbb{E}_{Z|X, \theta^{(m)}}\left[\log p(X, Z|\theta)\right]$

2) Если $Z = \{\bar z_1, \dots, \bar z_n\}$, $X = \{\bar x_1, \dots, \bar x_N\}$, ∞

$= \sum_{n=1}^{N} \mathbb{E}_{\bar z_n|\bar x_n, \theta^{(m)}}\left[\log p(\bar x_n, \bar z_n|\theta)\right]$

3) Generalized EM



$\ell(\theta) \geq h(\theta, \theta^{(m)})$

θ': $h(\theta', \theta^{(m)}) > h(\theta^{(m)}, \theta^{(m)})$

$\underline{\ell(\theta') > \ell(\theta)}$

4) Stochastic EM $\quad Q(\theta, \theta^{(m)}) = E_{Z|X,\theta^{(m)}}\left[\log p(X, Z|\theta)\right]$

Monte Carlo EM $\quad Z_1^{(m)}, \dots, Z_R^{(m)} \sim p(Z|X, \theta^{(m)})$

$$\hat{Q}(\theta, \theta^{(m)}) = \frac{1}{R} \sum_{r=1}^{R} \log p(X, Z_r^{(m)}|\theta) \xrightarrow[\theta]{} \max$$

$$p(X, Z|\theta) = \prod_n (\alpha N_1(\bar{x}_n))^{z_n} ((1-\alpha) N_2(\bar{x}_n))^{1-z_n}$$

$\bar{x}_n$: $z_{n1}, z_{n2}, \_ \; z_{nR}$

$$\sum_n \sum_r \left[z_{nr} \log(\_) + (1 - z_{nr}) \log(\_) \right]$$

5) $\theta_{MAP} = \arg\max_\theta \log p(\theta|X) = \arg\max_\theta \left[\log p(X|\theta) + \log p(\theta)\right]$

$Q(\theta, \theta^{(m)}) = E_{Z|X,\theta^{(m)}}\left[\log p(X, Z|\theta)\right] \underset{\uparrow \text{const}}{\leq} \log p(X|\theta)$

$\theta^{(m+1)} = \arg\max_\theta \left[Q(\theta, \theta^{(m)}) + \log p(\theta)\right]$

6) Semi-supervised learning

- fix z_n