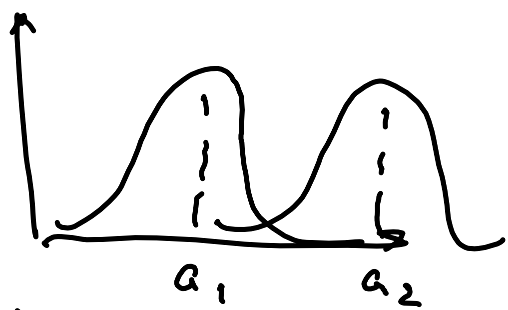
Bayes methods

$$KL(p\|q) = \int_{\mathbb{R}} p(x) \log \frac{p(x)}{q(x)} dx$$

$$\mathcal{N}(x|a, \sigma^2) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-a)^2}{2\sigma^2}}$$

$$KL(\mathcal{N}(x|a_1, \sigma^2) \| \mathcal{N}(x|a_2, \sigma^2)) =$$

$$= \frac{(a_1 - a_2)^2}{2\sigma^2}$$



$$p(x) = \int p(x, y) dy$$

$$p(x, y) = p(x|y) \cdot p(y)$$

EM-алгоритм.

$$\log p(X|\theta) \to \max_{\theta}$$

$$\log p(x|\theta)\cdot 1 = \log p(x|\theta)\cdot \int q(z)\,dz =$$

$$= \int q(z)\cdot \log p(x|\theta)\,dz =$$

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

$$\left\{ p(x,z|\theta) = p(z|x,\theta)\cdot p(x|\theta) \right\}$$

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

$$= \underbrace{\int q(z) \log \frac{p(x,z|\theta)}{q(z)}\,dz}_{\mathcal{L}(q,\theta)} + \underbrace{\int q(z) \log \frac{q(z)}{p(z|x,\theta)}\,dz}_{KL(q(z)\,||\,p(z|x,\theta))}$$

ELBO

$$\log p(x|\theta) = \mathcal{L}(q,\theta) + \underbrace{KL(q\|p)}_{\geq 0}$$

$$\log p(x|\theta) \geq \mathcal{L}(q,\theta)$$

E - шаг: $\mathcal{L} \to \max\limits_{q}$

M - шаг: $\mathcal{L} \to \max\limits_{\theta}$

$$\mathcal{L} = \log p(x|\theta) - KL(q(z)\|p(z|x,\theta))$$

$$\mathcal{L} \to \max_{q} \sim KL(q\|p) \to \min_{q}$$

$$KL(q\|p) = 0 \Leftrightarrow q = p$$

$$\mathcal{L} \to \max_{q} \Rightarrow \boxed{q^*(z) = p(z|x,\theta^{old})}$$

M - шаг: $\mathcal{L} \to \max\limits_{\theta}$

$$\mathcal{L}(q^*,\theta) = \int q^*(z) \cdot \log \frac{p(x,z|\theta)}{q^*(z)}\, dz =$$

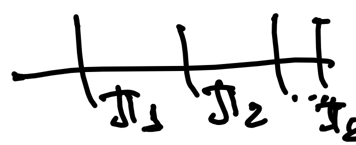
$$= \underbrace{\int q^*(z)\log p(x,z|\theta)\,dz}_{\mathbb{E}_{q^*}\log p(x,z|\theta)} - \underbrace{\int q^*(z)\log q^*(z)\,dz}_{H(q^*)}$$

$$\boxed{\theta^{new} = \arg\max_{\theta} \mathbb{E}_{q^*}\log p(x,z|\theta)}$$

Разделение смеси компонент.

$$p(x) = \sum_{k=1}^{K} \pi_k \cdot p_k(x) \qquad \pi_k \geq 0, \quad \sum_{k=1}^{K}\pi_k = 1$$

$$p_k(x) = \mathcal{N}(x|a_k, \sigma_k^2)$$



$$p(x) = \sum_{k=1}^{K} p(x|H_k=1)\,p(H_k=1)$$

$$p(X|\theta) = \prod_{i=1}^{n} p(x_i|\theta) =$$

$$= \prod_{i=1}^{n} \sum_{k=1}^{K} \pi_k \cdot \mathcal{N}(x_i | a_k, \sigma_k^2)$$

$$\log p(X|\theta) = \sum_{i=1}^{n} \log\left(\sum_{k=1}^{K} \pi_k \cdot \mathcal{N}(x_i | a_k, \sigma_k^2)\right)$$

$z_{ik} = \begin{cases} 1, & \text{i-ый объект из k-ой комп.} \\ 0, & \text{иначе.} \end{cases}$

$$p(X, Z|\theta) = p(X|Z, \theta) \cdot \underbrace{p(Z|\theta)}$$

$$p(z_{ik} = 1) = \pi_k \Rightarrow p(z_i = 1) = \prod_{k=1}^{K} \pi_k^{z_{ik}}$$

$$p(X|Z, \theta) = \prod_{i=1}^{n} p(x_i | z_{ik} = 1, \theta) =$$

$$= \prod_{i=1}^{n} \mathcal{N}(x_i | a_k, \sigma_k^2)$$

$$p(x_i | z_i, \theta) = \prod_{k=1}^{K} \mathcal{N}(x_i | a_k, \sigma_k^2)^{z_{ik}}$$

$$P(X,Z|\Theta) = \underbrace{P(X|Z,\Theta)}_{\prod_{i=1}^{n}\prod_{k=1}^{K} N(\ldots)^{z_{ik}}} \cdot \underbrace{P(Z|\Theta)}_{\prod_{i=1}^{n}\prod_{k=1}^{K} \pi_k^{z_{ik}}} =$$

$$=$$

$$P(X,Z|\Theta) = \prod_{i=1}^{n}\prod_{k=1}^{K} \left(\pi_k \cdot N(x_i|a_k,\sigma_k^2)\right)^{z_{ik}}$$

$$\log P(X,Z|\Theta) = \sum_{i=1}^{n}\sum_{k=1}^{K} z_{ik}\left(\log \pi_k + \log N(\ldots)\right)$$

Е-шаг: $q^*(Z) = P(Z|X,\Theta)$

$$\boxed{P(z_{ik}=1 \mid x_i,\Theta) = \frac{\pi_k \cdot N(x_i|a_k,\sigma_k^2)}{\sum_{j=1}^{K} \pi_j \cdot N(x_i|a_j,\sigma_j^2)}}$$

$$P(Z|X) \propto P(X,Z)$$

$$P(Z|X) = \frac{P(X,Z)}{P(X)} \rightarrow \text{не зависит от } Z.$$

М-шаг:

$$\Theta^* = \arg\max_{\Theta} \mathbb{E}_{q^*} \log P(X,Z|\Theta)$$

$$\mathbb{E}_q \log p(X, Z \mid \theta) = \mathbb{E}_q \sum_i \sum_k z_{ik} \{ \log \pi_k +$$

$$+ \log \mathcal{N}(x_i \mid a_k, b_k^2) \} =$$

$$= \sum_i \sum_k \mathbb{E}(z_{ik}) \{ \log \pi_k + \log \mathcal{N}(\ldots) \} +$$

$$+ \lambda \left(\sum_{k=1}^{K} \pi_k - 1 \right)$$

$$\mathbb{E}_q(z_{ik}) = 1 \cdot p(z_{ik} = 1 \mid x_i, \theta) + 0 \cdot$$

$$\cdot p(z_{ik} = 0 \mid x_i, \theta) = p(z_{ik} = 1 \mid x_i, \theta) =$$

$$= \boxed{\gamma_{ik}}$$

$$Q = \sum_{i=1}^{n} \sum_{k=1}^{K} \gamma_{ik} \{ \log \pi_k - \frac{1}{2} \log 2\pi - \log b_k -$$

$$- \frac{(x_i - a_k)^2}{2 b_k^2} \} + \lambda \left(\sum_{k=1}^{K} \pi_k - 1 \right)$$

$$Q'_{\pi_k} = \ldots = 0 \qquad \pi_k^{new} = \ldots$$

$$Q'_{a_k} = \ldots = 0 \qquad a_k^{new} = \ldots$$

Семинар

$\theta_{old} \; \dots \; \theta_{new}$.

$$\log p(X|\theta_{old}) = \mathcal{L}(q, \theta_{old}) + KL(q \| p)$$

Е-шаг: $q^* = p(Z|X, \theta_{old}) \Rightarrow$

$$\Rightarrow KL(q^* \| p) = 0 \Rightarrow \log p(X|\theta_{old}) = \mathcal{L}(q^*, \theta_{old})$$

М-шаг: $\theta = \arg\max\limits_{\theta} \mathcal{L}(q^*, \theta)$

$$\mathcal{L}(q^*, \theta^{new}) > \mathcal{L}(q^*, \theta_{old})$$

$$\log p(X|\theta_{new}) = \mathcal{L}(q^*, \theta^{new}) + \underbrace{KL(q^* \| p(Z|X, \theta^{new}))}_{\geq 0} > \mathcal{L}(q^*, \theta_{old}) = \log p(X|\theta_{old})$$

① $p_{km} = \frac{1}{|R_m|} \sum_{i: x_i \in R_m} [y_i = k]$

$$G(R_m) = \sum_{k=1}^{K} p_{km}(1 - p_{km})$$

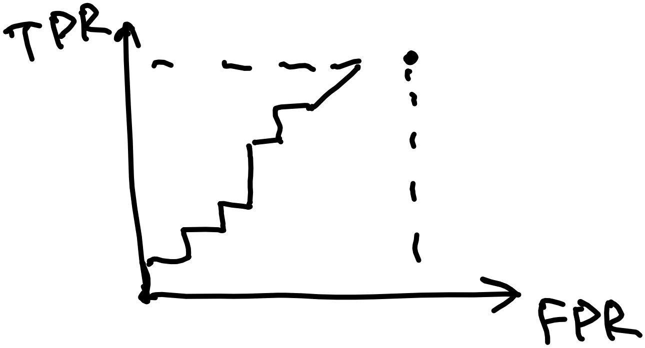
$$Q = G(R_m) - \frac{|R_l|}{|R_m|} \cdot G(R_l) - \frac{|R_r|}{|R_m|} \cdot G(R_r)$$

② $Pr = \frac{TP}{TP + FP}$ $\quad Rec = \frac{TP}{TP + FN}$

$a_i \backslash y_i$	1	0
1	TP	FP
0	FN	TN

③ AUC =

0	1
0.5	0.5

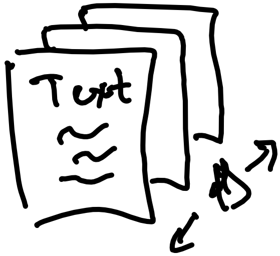


4. $$Q = \frac{1}{n} \sum_{i=1}^{n} \log\Big(1 + \exp\big(\underbrace{-y_i \langle w, x_i \rangle}_{\text{margin}_i}\big)\Big) + \lambda \cdot \|w\|^2$$

5. $$\mathcal{L} = \prod_{t=1}^{T} \prod_{\substack{-m \le j \le m \\ j \ne 0}} \mathbb{P}(w_{t+j} \mid w_j)$$

$$\mathbb{P}(\text{out} \mid \text{center}) = \frac{\exp(u_{\text{out}}^T \cdot v_{\text{center}})}{\sum_{u \in U} \exp(u^T v_{\text{center}})}$$

$$U = \underbrace{\boxed{}}_{|Voc| \times h} \qquad V = \underbrace{\boxed{}}_{|Voc| \times h}$$



$$(d_i, w_i)$$

$$p(D, W) = \prod_{d=1}^{D} \prod_{w \in d} p(d, w)$$

$$p(d, w) = p(w|d) \cdot p(d) =$$

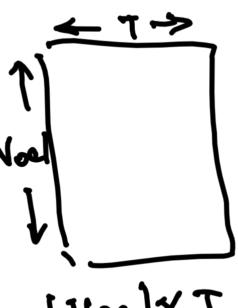
$$p(w|t, d) = p(w|t)$$

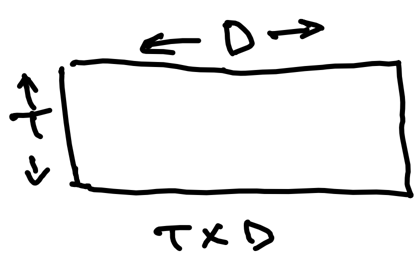
$$= p(d) \cdot \sum_{t=1}^{T} p(w, t|d) =$$

$$P(A, B) = P(A|B)\,P(B)$$

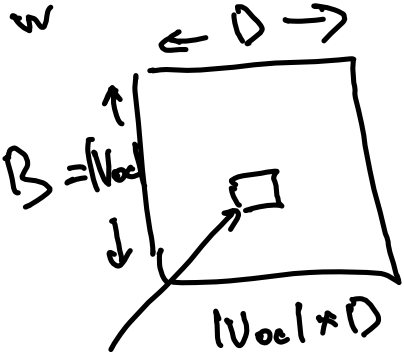
$$\left\{ P(A) = \sum_{B} P(A, B) \right\}$$

$$= p(d) \cdot \sum_{t=1}^{T} p\underbrace{(w|t, \cancel{d})}_{\varphi_{wt}} \cdot p\underbrace{(t|d)}_{\theta_{td}}$$

$\Phi =$ 

$\Theta =$ 

$$\sum_{w} P(w|t) = 1 \qquad \sum_{t} p(w|t) \neq 1$$



$$\approx \Phi \cdot \Theta$$

$p(w|d)$

$$P(D,W) = \prod_{d=1}^{D} \prod_{w \in d} p(d,w) =$$

$$= \prod_{d=1}^{D} \prod_{w \in d} p(d) \cdot p(w|d) =$$

$$= \prod_{d=1}^{D} \prod_{w \in d} p(d) \cdot \sum_{t=1}^{T} p(w|t) \cdot p(t|d)$$

$$\log P(D,W) = \sum_{d=1}^{D} \sum_{w \in d} \log\left(\sum_{t=1}^{T} \varphi_{wt} \cdot \theta_{td}\right) \to \max_{\Phi, \Theta}$$

$$P(D,W|T) = \prod_{d=1}^{D} \prod_{w \in d} \prod_{t=1}^{T} \left(p(w|t) \cdot p(t|d)\right)^{[t=t_d]}$$

$$= \prod_{d=1}^{D} \prod_{w \in d} \prod_{t=1}^{T} \left(p(w|t) \cdot p(t|d)\right)^{[t=t_d]}$$

$$P(D,W,T) \propto P(D,W|T)$$

$$\log P(D,W,T) = \sum_{d=1}^{D} \sum_{w \in d} \sum_{t=1}^{T} [t=t_d] \cdot$$

$$\cdot \log(\varphi_{wt} \cdot \theta_{td})$$

E - шаг:

$$q(T) = p(t_d = t | w,d) \propto p(d,w,t) =$$

$$= \frac{p(w|t) \cdot p(t|d)}{\sum_{t=1}^{T} p(w|t) \cdot p(t|d)} = \boxed{\frac{\varphi_{wt} \cdot \theta_{td}}{\sum_{j=1}^{T} \varphi_{wj} \cdot \theta_{jd}}} =$$

$$= \gamma_{dwt}$$

$\boxed{\text{М - шаг}}$ $\mathbb{E}_{q(T)} \log P(D, W, T) \to \max\limits_{\Phi, \Theta}$

$$Q = \mathbb{E}_q \sum_{d=1}^{D} \sum_{w \in d} \sum_{t=1}^{T} [t_d = t] \cdot \log(\varphi_{wt} \cdot \theta_{td}) +$$

$$+ \sum_{t=1}^{T} \lambda_t \Big(\sum_{w \in Voc} \varphi_{wt} - 1 \Big) +$$

$$+ \sum_{d=1}^{D} \mu_d \Big(\sum_{t=1}^{T} \theta_{td} - 1 \Big) \to \max_{\varphi_{wt}, \theta_{td}}$$

$$\mathbb{E}_q([t_d = t]) = p(t_d = t \mid w, d) = \gamma_{dwt}$$

$$Q = \sum_{d=1}^{D} \sum_{w \in d} \sum_{t=1}^{T} \gamma_{dwt} \cdot \log(\varphi_{wt} \cdot \theta_{td}) +$$

$$+ \sum_{t=1}^{T} \lambda_t \Big(\sum_{w \in Voc} \varphi_{wt} - 1 \Big) + \sum_{d=1}^{D} \mu_d \Big(\sum_{t=1}^{T} \theta_{td} - 1 \Big)$$

$$Q'_{\varphi_{wt}} = \sum_{d=1}^{D} \gamma_{dwt} \cdot n_{dw} \cdot \frac{1}{\varphi_{wt}} + \lambda_t = 0$$

$$\varphi_{wt} = -\frac{1}{\lambda_t} \sum_{d} \gamma_{dwt} \cdot n_{dw} \Rightarrow \boxed{\varphi_{wt} = \frac{\sum_{d=1}^{D} \gamma_{dwt} \cdot n_{dw}}{\sum_{d=1}^{D} \sum_{w} \gamma_{dwt} \cdot n_{dw}}}$$

$$\sum_{w \in Voc} \varphi_{wt} = 1 = -\frac{1}{\lambda_t} \cdot \sum_{d=1}^{D} \sum_{w \in Voc} \gamma_{dwt} \cdot n_{dw} \Rightarrow$$

$$\Rightarrow -\lambda_t = \sum_{d=1}^{D} \sum_{w \in Voc} \cdot \gamma_{dwt} \cdot n_{dw}$$

$$p(t|w) = \frac{p(w|t)\,p(t)}{\sum_t p(w|t)\cdot p(t)} =$$

$$= \frac{p(w|t)\sum_d p(t|d)\cdot p(d)}{\sum_t \left(p(w|t)\cdot \sum_d p(t|d)\cdot p(d)\right)}$$

$\boxed{\text{PLSA}}$