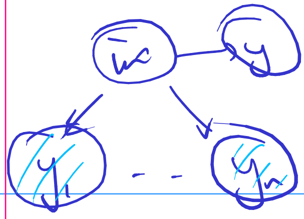
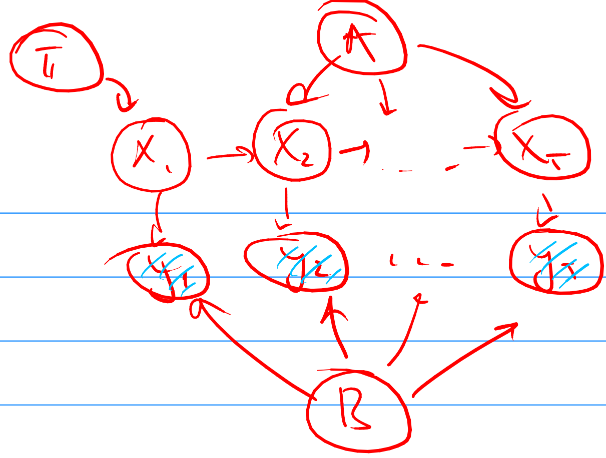
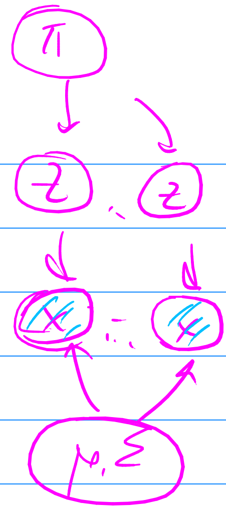
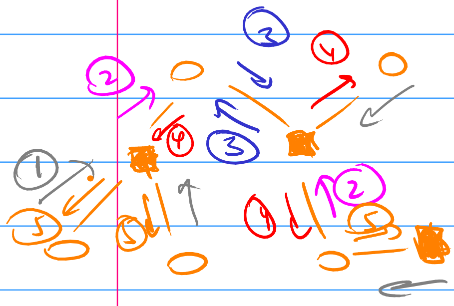




## Approximate inference

$$\mu_{f\to x}(x) = \int f(x,\bar{y})\cdot\prod_{y}\mu_{y\to f}(y)\cdot d\bar{y}$$

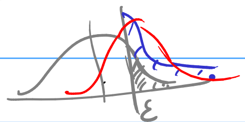
**TrueSkill**

$$F(X) = f_1(X_1)\, f_2(X_2) \,\dots\, f_n(X_n)$$

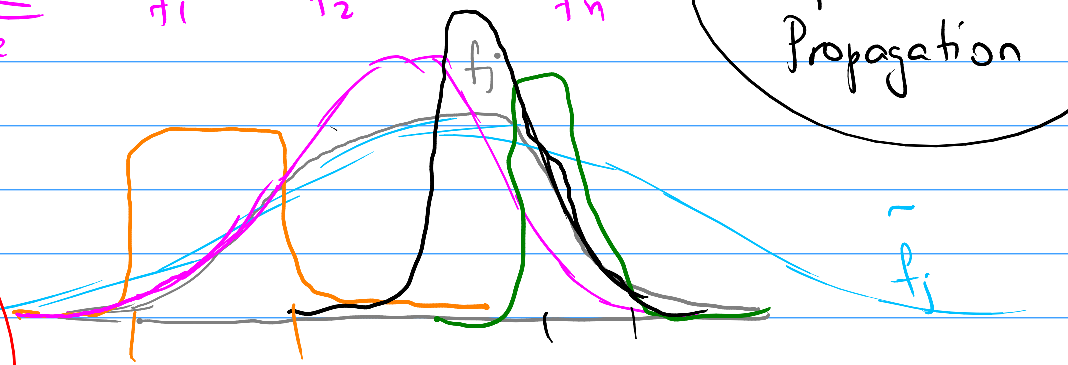
$$\hat{F}_{naive} = \tilde{\hat{f}}_1 \quad \tilde{\hat{f}}_2 \quad f_j \quad \tilde{\hat{f}}_n$$

Expectation Propagation

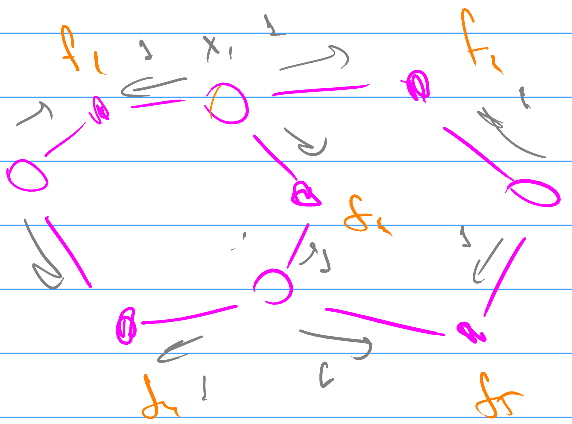
[i>j]



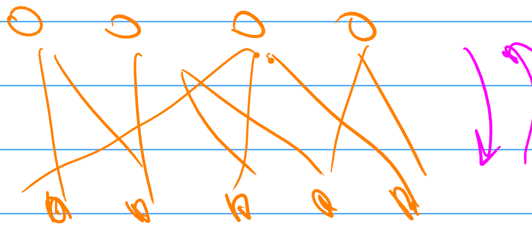
$\tilde{f}_j$




$$\tilde{f}_j = \prod_{i\neq j} f_i \approx f_j \prod_{i\neq j} f_i$$




Loopy belief propagation



## Sampling

$p(\bar{x})$ $\qquad$ $\bar{x} \sim p(\bar{x})$ $\qquad$ $\boxed{\bar{x} \sim Unif([0,1])}$ rand()

$$\bar{x} \; — \blacksquare — \; p(\bar{x})$$

$$\bar{x} \; — \square — \; p^*(\bar{x}) \propto p(\bar{x})$$

$$Unif([0,1]) \rightarrow \boxed{?} \rightarrow \bar{x}_1, \bar{x}_2, \ldots, \bar{x}_r \sim p(\bar{x})$$

$$p^*(\bar{x}) = \boxed{p(\theta)\, p(D|\theta)} \propto p(\theta|D)$$

$$p(\theta|D) = \frac{p(\theta)\, p(D|\theta)}{p(D)}$$

1) $p(\theta|D) \xrightarrow[\theta]{} \max$

2) $p(x|D) = \int p(x|\theta)\, p(\theta|D)\, d\theta = \boxed{\mathbb{E}_{p(\theta|D)}[p(x|\theta)]}$

$$\boxed{\mathbb{E}_{p(\bar{x})}[f(\bar{x})] \approx \frac{1}{R}\sum_{r=1}^{R} f(\bar{x}^{(r)})}, \text{ где } \bar{x}^{(r)} \sim p(\bar{x})$$

3) $Q(\theta, \theta^{(m)}) = \mathbb{E}_{p(z|\theta^{(m)})}[\log p(x, z|\theta)] \approx$

**Monte Carlo EM**

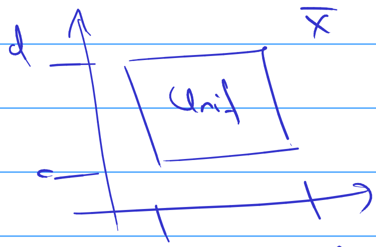
$$\approx \frac{1}{R}\sum_r \log p(x, z^{(r)}|\theta), \quad z^{(r)} \sim p(z|\theta^{(m)})$$

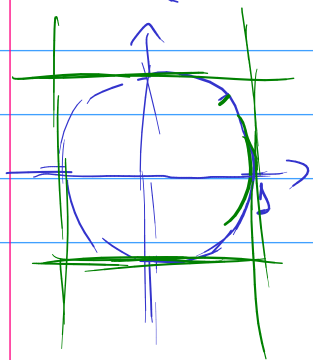
$$\xrightarrow[\theta]{} \max$$

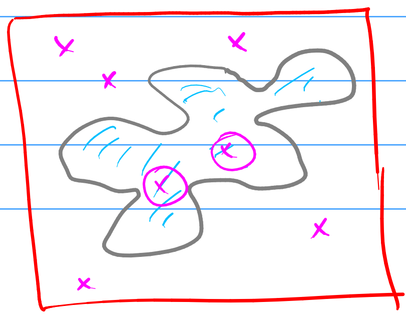
---

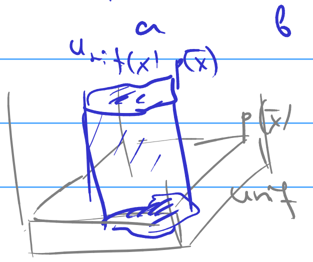
1) Известные распределения

$$x \sim Unif([0,1]) \xrightarrow{a+(b-a)x} x \sim Unif([a,b])$$

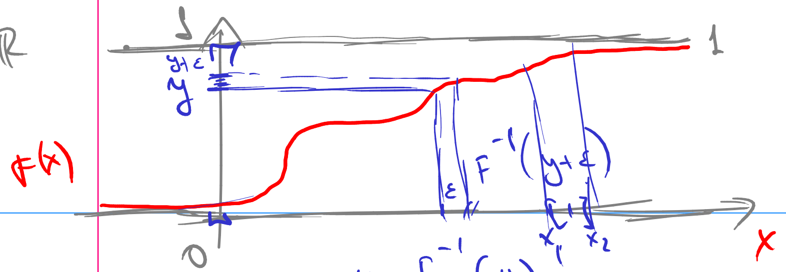






$x \in \mathbb{R}$




$\underline{x = F^{-1}(y)}$

$p(x)$, $x \in \mathbb{R}$

$$F(a) = \int_{-\infty}^{a} p(x)\,dx$$

$$\Pr(x \in [x_1, x_2]) = \Pr(y \in [F(x_1), F(x_2)]) =$$

$$= \underline{F(x_2) - F(x_1)}$$

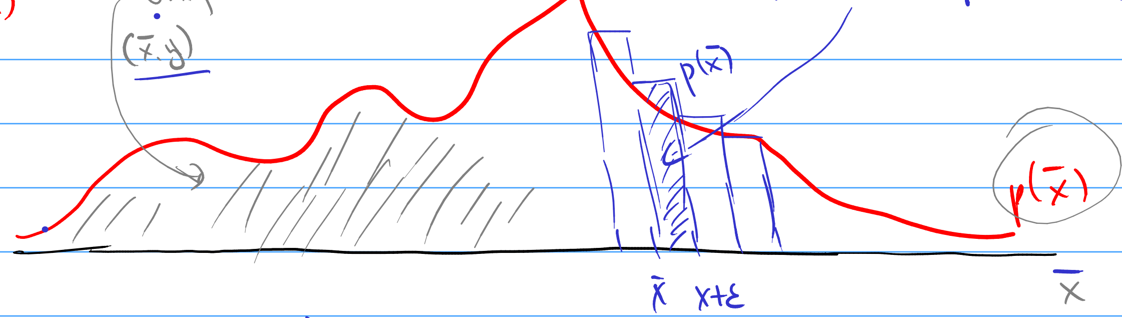
$$\Pr(x \in [x, x+\varepsilon]) = F(x+\varepsilon) - F(x)$$

$$\underline{p(x) = F'(x)}$$

---

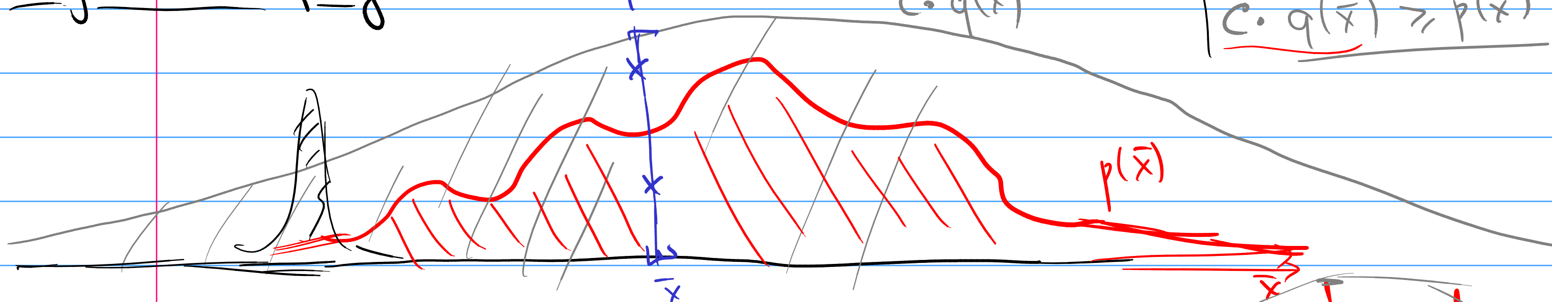
2) $x$ — [?] — $p(x)$

$\Pr[(x,y) \in \boxed{\ }] = \varepsilon \cdot \underline{p(x)}$



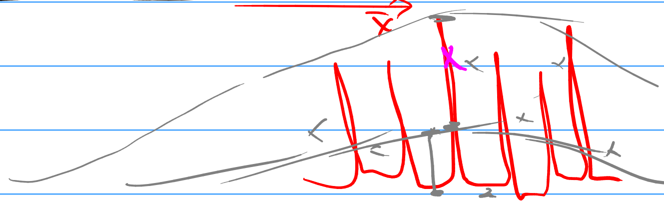

## Rejection sampling

$\boxed{c \cdot q(\bar{x}) \geq p(\bar{x})}$




- $\underline{\bar{x} \sim q(\bar{x})}$
- $y \sim \mathrm{Unif}([0, c \cdot q(\bar{x})])$
- если $y \leq p(\bar{x}) \Rightarrow$ берём $\bar{x}$

  $y > p(\bar{x}) \to$ отклоняем



## Importance sampling

$\bar{x} \sim q(\bar{x})$

Если $p(\bar{x}) > 0 \Rightarrow q(\bar{x}) > 0$

$$\mathbb{E}_{p(\bar{x})}[f(\bar{x})] = \int p(\bar{x}) f(\bar{x})\, d\bar{x} = \int f(\bar{x}) \frac{p(\bar{x})}{q(\bar{x})} \cdot q(\bar{x})\, d\bar{x} =$$

$$= \mathbb{E}_{q(\bar{x})}\left[f \cdot \frac{p}{q}\right] \approx \frac{1}{R} \sum_{r} f(\bar{x}^{(r)}) \frac{p(\bar{x}^{(r)})}{q(\bar{x}^{(r)})}$$

$\bar{x}^{(r)} \sim q(\bar{x})$

importance weights

$p^*(\bar{x}) \propto p(\bar{x})$ , $q^*(\bar{x}) \propto q(\bar{x})$

$$p(\bar{x}) = \frac{1}{Z_p} p^*(\bar{x}) \quad , \quad q(\bar{x}) = \frac{1}{Z_q} q^*(\bar{x})$$

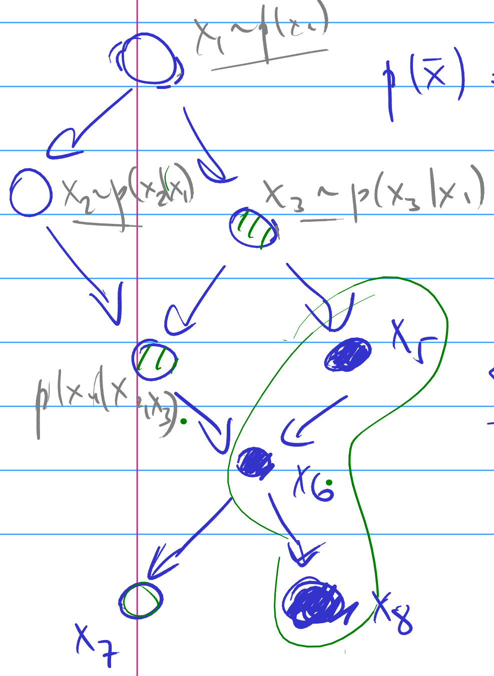
$$E_{p(\bar{x})}[f(\bar{x})] = \int f(\bar{x}) \frac{\frac{1}{Z_p} p^*(\bar{x})}{\frac{1}{Z_q} q^*(\bar{x})} \cdot \left[\frac{1}{Z_q} q^*(\bar{x})\right] d\bar{x} =$$

$$= \underline{E_q\left[f(\bar{x}) \cdot \frac{p^*(\bar{x})}{q^*(\bar{x})} \frac{Z_q}{Z_p}\right]} \approx$$

$$Z_p = \underline{\int p^*(\bar{x}) d\bar{x}} = \int \frac{p^*(\bar{x})}{q(\bar{x})} q(\bar{x}) d\bar{x} \approx \frac{1}{R} \sum \frac{p^*(\bar{x}^{(r)})}{q(\bar{x}^{(r)})}$$

$$Z_q \approx \frac{1}{R} \sum \frac{q^*(\bar{x}^{(r)})}{q(\bar{x}^{(r)})}$$

$$\approx \frac{1}{R} \sum_r f(\bar{x}^{(r)}) \boxed{\frac{p^*(\bar{x}^{(r)})}{q^*(\bar{x}^{(r)})} \cdot \left(\frac{\sum_r q^*(\bar{x}^{(r)})}{\sum_r p^*(\bar{x}^{(r)})}\right)}$$




$$p(\bar{x}) = \prod_i p(x_i | par(x_i))$$

$$x_1 \ldots x_4, x_7 \sim p(x_1 \ldots x_4, x_7 | x_5, x_6, x_8)$$

Rejection $\neq$ Importance sampling

$$\boxed{q(x_1, \ldots x_4, x_7)} = p(x_1) p(x_2|x_1) p(x_3|x_1) p(x_4|x_1,x_3) p(x_7|x_6)$$

$$\frac{p(x_1 \ldots x_4, x_7 \mid x_5, x_6, x_8)}{q(x_1 \ldots x_4, x_7)} = \frac{p(x_1 \ldots x_8)}{p(x_5, x_6, x_8)\, q(x_1 \ldots x_4, x_7)} =$$

$$= \frac{\cancel{p(x_1)}\, \cancel{p(x_2 \mid x_1)} \ldots p(x_5 \mid x_3)\, p(x_6 \mid x_4)\, \cancel{p(x_7 \mid \_)}\, p(x_8 \mid \_)}{p(x_5, x_6, x_8)\, \cancel{p(x_1)}\, \cancel{p(x_2 \mid x_1)}\, \cancel{p(x_3 \mid \_)}\, \cancel{p(x_4 \mid \_)} \ldots \cancel{p(x_7 \mid \_)}} =$$

$$= \frac{\boxed{p(x_5 \mid x_3)\, p(x_6 \mid x_4, x_5)\, p(x_8 \mid x_6)}}{p(x_5, x_6, x_8) = p(D)}$$

---

$\bar{w} - \ldots - p^*(\bar{w})$

$\bar{w}$

$$p(\bar{w} \mid D) \propto \boxed{p(\bar{w}) \prod_n p(y_n \mid \bar{w})} = p^*(\bar{w})$$

$$p(y \mid D) = \int p(y \mid \bar{w})\, p(\bar{w} \mid D)\, d\bar{w} = \mathbb{E}_{p(\bar{w} \mid D)}\left[p(y \mid \bar{w})\right]$$

$$p(y \mid D) \approx \frac{1}{R} \sum_r p(y \mid \bar{w}^{(r)}), \quad \boxed{\bar{w}^{(r)} \sim p(\bar{w} \mid D)}$$

$$\bar{w}^{(r)} \sim q(\bar{w})$$

$$p(y \mid D) \approx \frac{1}{R} \left( \sum_r p(y \mid \bar{w}^{(r)}) \frac{p^*(\bar{w}^{(r)})}{q^*(\bar{w}^{(r)})} \right) \left( \frac{\sum q^*}{\sum p^*} \right)$$

- repeat

$$\bar{w} \sim \boxed{q(\bar{w})}$$

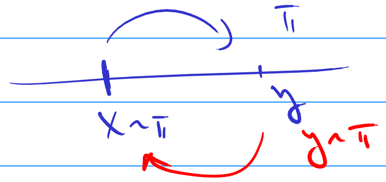
$$p^*(\bar{w}) = p(\bar{w}) \prod_n p(y_n \mid \bar{w})$$

- average with weights

# ③ MCMC Markov Chain Monte Carlo

$$X_1 \to X_2 \to X_3 \to \dots \to X_t \to X_{t+1} \to \dots$$

$$p(x_t|x_{t-1}) = p(x_t|x_{t-1}\dots x_1)$$

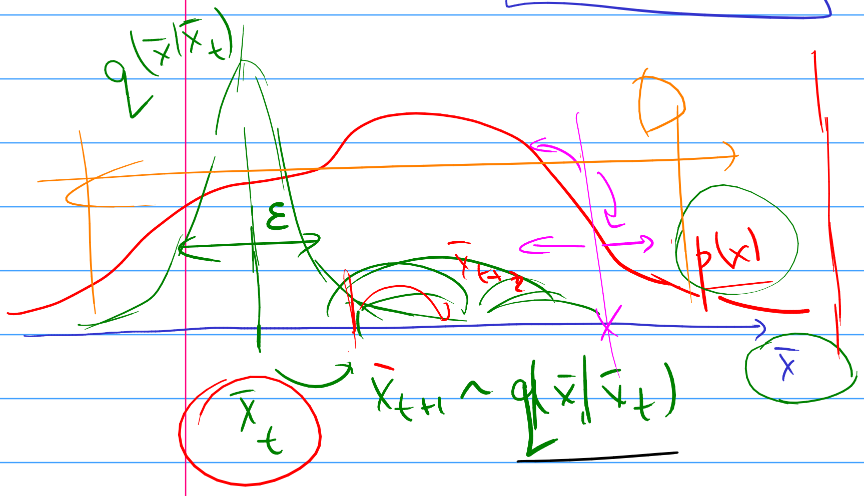



Стацион. распред. $\pi(x) = \int p(x|y)\pi(y)\,dy$

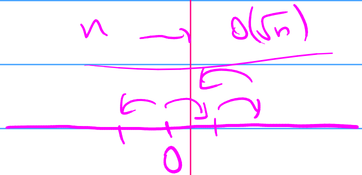
Условие баланса: если $\forall x, y$ $\boxed{\pi(x)p(y|x) = \pi(y)p(x|y)}$,

то $\pi(x)$ — стац. р. для цепи $p(y|x)$

$$\pi(x) \overset{?}{=} \int p(x|y)\pi(y)\,dy = \int \pi(x)\, p(y|x)\,dy = \pi(x)$$




$\bar{x}_{t+1} \sim q(\bar{x}|\bar{x}_t)$




$O(D/\varepsilon^2)$

**Metropolis–Hastings algorithm**

- repeat:
  - sample $\bar{x}' \sim q(\bar{x}'|\bar{x}_t)$
  - $a(\bar{x}_t, \bar{x}') = \dfrac{p^*(\bar{x}')}{p^*(\bar{x}_t)} \cdot \dfrac{q(\bar{x}_t|\bar{x}')}{q(\bar{x}'|\bar{x}_t)}$
  - if $a \geq 1$ then $\bar{x}_{t+1} := \bar{x}'$

    else

    $$\bar{x}_{t+1} := \begin{cases} \bar{x}' & \text{с вер. } a \\ \bar{x}_t & \text{с вер. } 1-a \end{cases}$$

$$p(\bar{x}_t)\, q_{MC}(\bar{x}'|\bar{x}_t) \overset{?}{=} p(\bar{x}')\, q_{MC}(\bar{x}_t|\bar{x}')$$

$a(\bar{x}_t, \bar{x}') \geq 1$ $\quad p(\bar{x}_t)\, \cancel{q(\bar{x}'|\bar{x}_t)} \overset{?}{=} p(\bar{x}') \cdot \cancel{q(\bar{x}_t|\bar{x}')} \cdot \dfrac{p^*(\bar{x}_t)\, q(\bar{x}'|\bar{x}_t)}{p^*(\bar{x}')\, \cancel{q(\bar{x}_t|\bar{x}')}}$

$$a(\bar{x}', \bar{x}_t) = \frac{1}{a(\bar{x}_t, \bar{x}')}$$

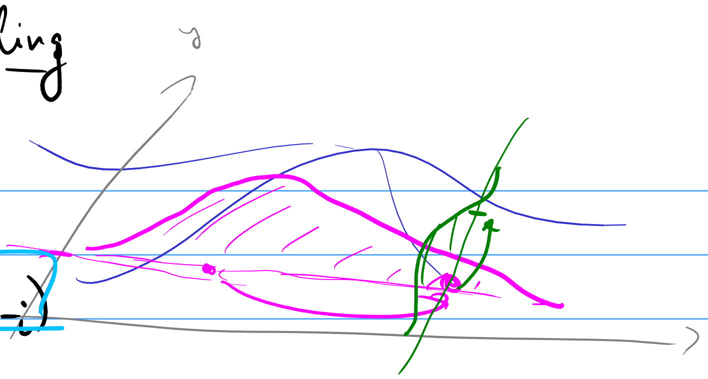
$$\frac{p(x_t)}{p(\bar{x}')} = \frac{p^*(\bar{x}_t)}{p^*(\bar{x}')}$$

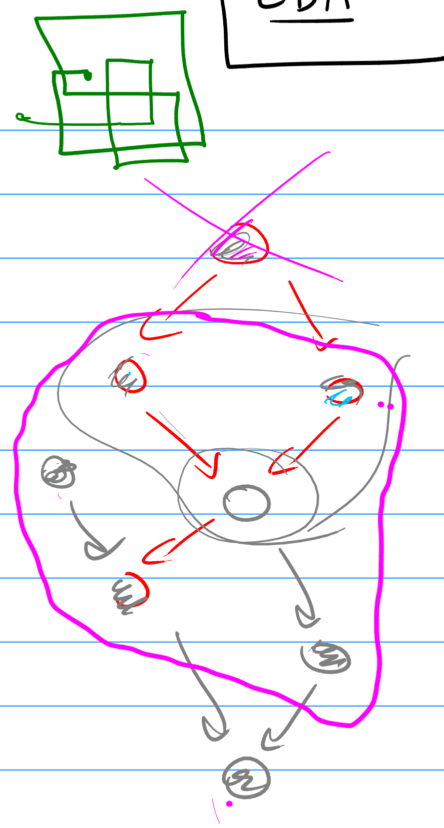
# Gibbs sampling

LDA

$p(x_1 \ldots x_n)$

- repeat:
- for $i = 1..n$:
  - $x_i \sim \boxed{p(x_i \mid \bar{x}_{-i})}$





$\bar{x} = (x_i, \bar{x}_{-i})$

$i$: $q(\bar{x}' \mid \bar{x})$: если $\bar{x}'_{-i} \neq \bar{x}_{-i}$, то $q = 0$

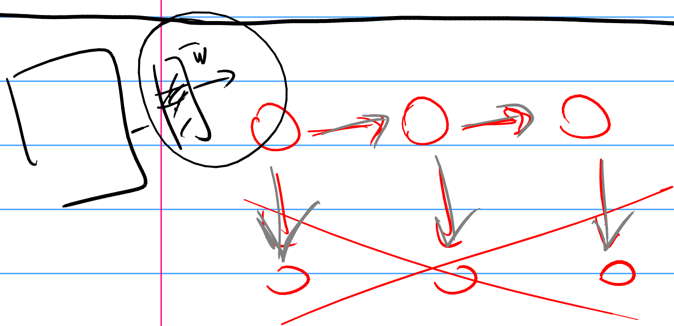
иначе $q(x'_i, \bar{x}_{-i} \mid \bar{x}) = p(x'_i \mid \bar{x}_{-i})$

$\bar{x}' = (x'_i, \bar{x}_{-i})$

$$a = \frac{p(x'_i, \bar{x}_{-i})}{p(x_i, \bar{x}_{-i})} \cdot \frac{q(x_i, \bar{x}_{-i} \mid x'_i, \bar{x}_{-i})}{q(x'_i, \bar{x}_{-i} \mid x_i, \bar{x}_{-i})} =$$

$$= \frac{p(x'_i, \bar{x}_{-i})}{p(x_i, \bar{x}_{-i})} \cdot \frac{p(x_i \mid \bar{x}_{-i})}{p(x'_i \mid \bar{x}_{-i})} = \frac{\cancel{p(\bar{x}_{-i})}\cancel{p(x'_i \mid \bar{x}_{-i})}}{\cancel{p(\bar{x}_{-i})}\cancel{p(x_i \mid \bar{x}_{-i})}} \cdot \frac{\cancel{p(x_i \mid \bar{x}_{-i})}}{\cancel{p(x'_i \mid \bar{x}_{-i})}}$$

$= 1$




$A$ $\quad q_1 \quad q_2 \ldots$

$p(q_i \mid q_j)$