① Why sampling?

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

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

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

$$\approx \frac{1}{R}\sum_{r=1}^{R} p(\bar{x}|\bar{\theta}_r), \quad \boxed{\bar{\theta}_r \sim p(\bar{\theta}|D)}$$

$$p(\bar{\theta}|D) \propto p(\bar{\theta}) \cdot \prod_n p(\bar{x}_n|\bar{\theta})$$

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

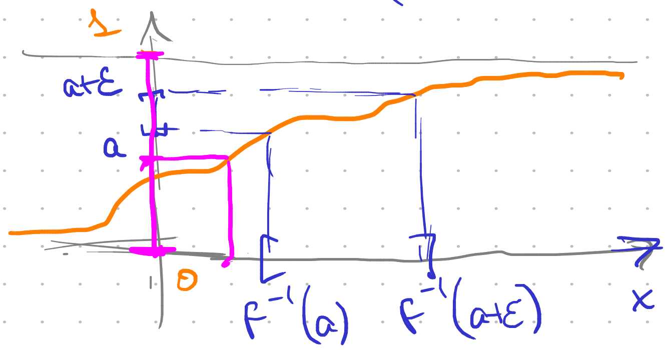
$$\approx \frac{1}{R}\sum_{r=1}^{R} \log p(x, z_r|\bar{\theta}) \rightarrow \max_{\bar{\theta}}$$

Monte Carlo EM

② $x \in \mathbb{R}$, $p(x)$, $F(a) = \int_{-\infty}^{a} p(x)\,dx$

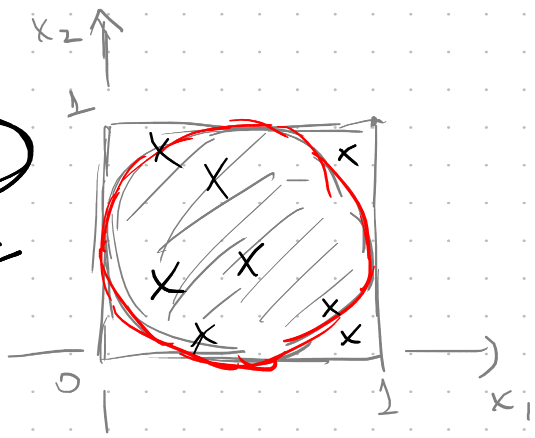
$z \sim \mathrm{Unif}([0,1])$

$z = \mathrm{rand}()$



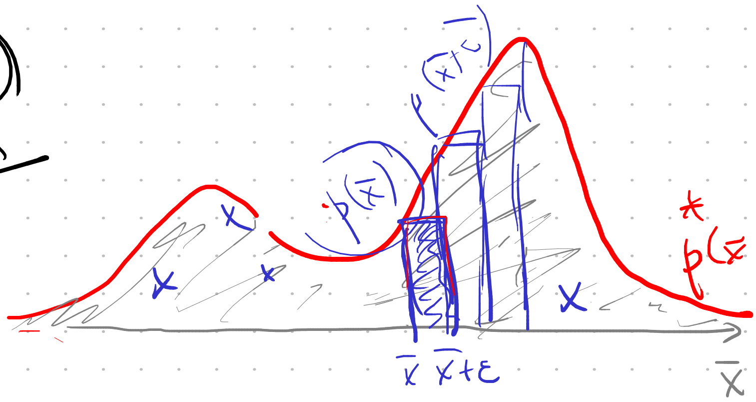
$$\mathbb{P}\left(x \in [F^{-1}(a), F^{-1}(a+\varepsilon)]\right) = \int_{F^{-1}(a)}^{F^{-1}(a+\varepsilon)} p(x)\,dx = \varepsilon$$

③ Idea



Unif

4 Idea



$\bar{x} \sim p(\bar{x}) \iff \bar{x} \sim \mathrm{Unif}(\;)$

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

$p^*(\bar{x}) = \boxed{C\cdot}\, p(\bar{x})$

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

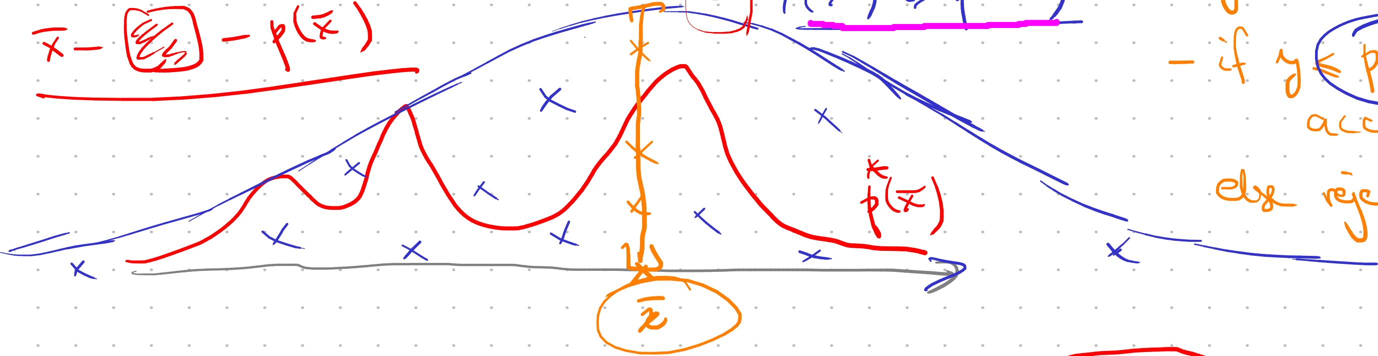
5 Rejection sampling

$\bar{x} - \boxed{} - p(\bar{x})$

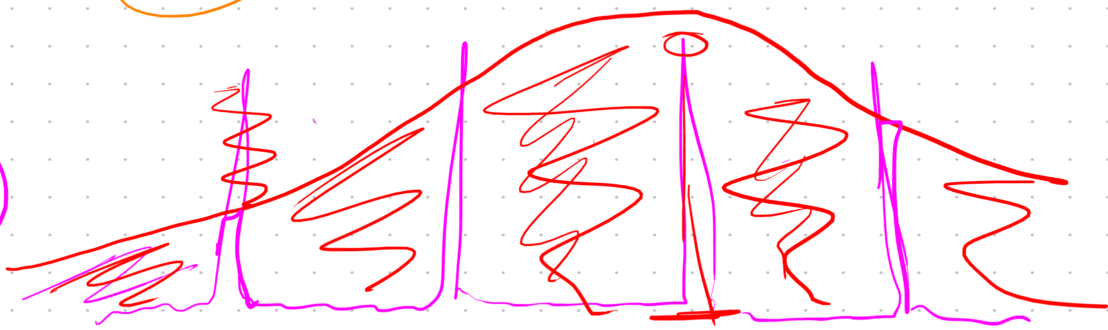
$C\cdot q(\bar{x}) \geq p^*(\bar{x})$

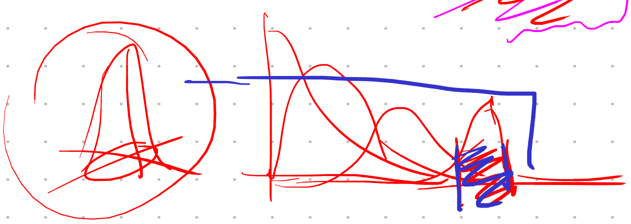
- $\bar{x} \sim q(\bar{x})$
- $y \sim \mathrm{Unif}[0, C\cdot q(\bar{x})]$
- if $y \leq p(\bar{x})$ accept $\bar{x}$

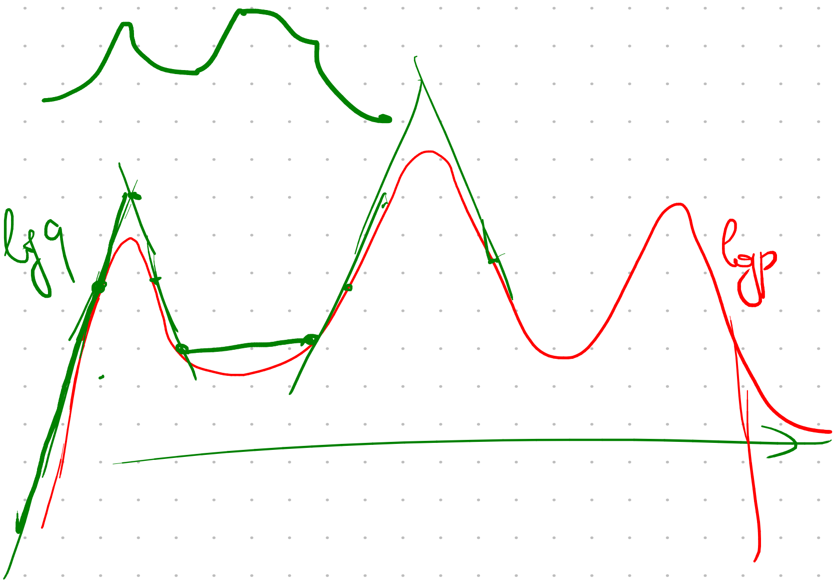
else reject $\bar{x}$



$p^*(\bar{\theta}) = p(\bar{\theta}|D)$







6 Importance sampling

~~$\tilde{x} \sim p(\tilde{x})$~~

$$E_{p(\bar{x})}[f(\bar{x})] = ?$$

~~$\tilde{x} \sim p(\tilde{x})$~~

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

Idea

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

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

importance weights

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

$\int p^*(\bar{x})\,d\bar{x}$

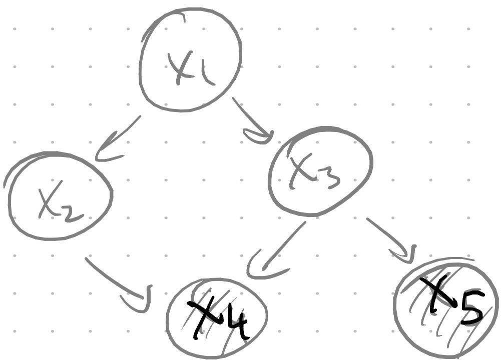
$$p(\bar{x}) = \frac{1}{Z_p}\,p^*(\bar{x}) \qquad E_p[f] = E_q\left[f(\bar{x})\cdot\frac{p^*(\bar{x})}{q^*(\bar{x})}\cdot\boxed{\frac{Z_q}{Z_p}}\right]$$

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

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

$\bar{x}_r \sim q(\bar{x})$: $\frac{Z_p}{Z_q} \approx \frac{1}{R}\sum \frac{p^*(\bar{x}_r)}{q^*(\bar{x}_r)}$

7



$$p(x_1 \ldots x_5) = p(x_1)\,p(x_2|x_1)\,p(x_3|x_1)\,p(x_4|x_2,x_3)\,p(x_5|x_3)$$

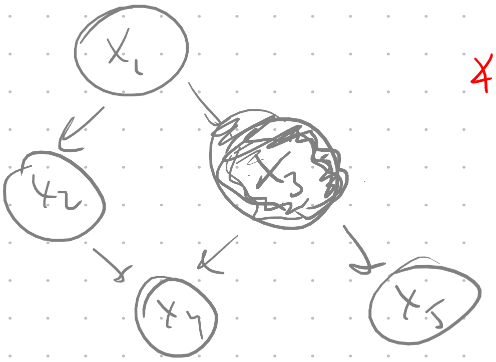
$x_1 \ldots x_5 \sim p(x_1 \ldots x_5)$?

$x_1 \sim p(x_1)$

$x_2 \sim p(x_2|x_1)$ $x_3 \sim p(x_3|x_1)$ $x_4 \sim$ ____

$$x_1, x_2, x_3 \sim p(x_1, x_2, x_3 \mid x_4, x_5)$$ $\Leftarrow$ rejection sampling

Importance sampling



$$p(x_1, x_2, x_4, x_5 \mid x_3 = a)$$

$$\not\propto \left[q(x_1, x_2, x_4, x_5) = p(x_1)\, p(x_2 \mid x_1)\, p(x_4 \mid x_2, a)\, p(x_5 \mid a) \right]$$

$$p(x_1, \dots, x_5 \mid x_3 = a) = \frac{p(x_1)\, p(x_2 \mid x_1)\, p(a \mid x_1)\, p(x_4 \mid x_2, a)\, p(x_5 \mid a)}{p(x_3 = a)}$$

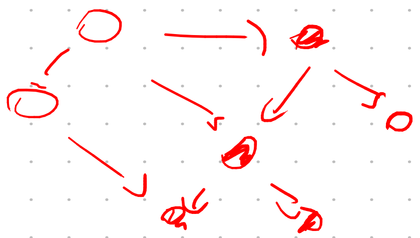
$$p(x_3 = a) \Rightarrow Z_p$$

$$E_p[f(\bar{x})] = E_q\left[f(\bar{x}) \cdot \frac{p^*(\bar{x})}{q(\bar{x})} \cdot \frac{1}{Z_p} \right] =$$

$$= E_q\left[\frac{1}{Z_p} f(\bar{x}) \cdot \frac{p(x_1)\, p(a \mid x_1) \dots}{p(x_1) \dots} \right] =$$

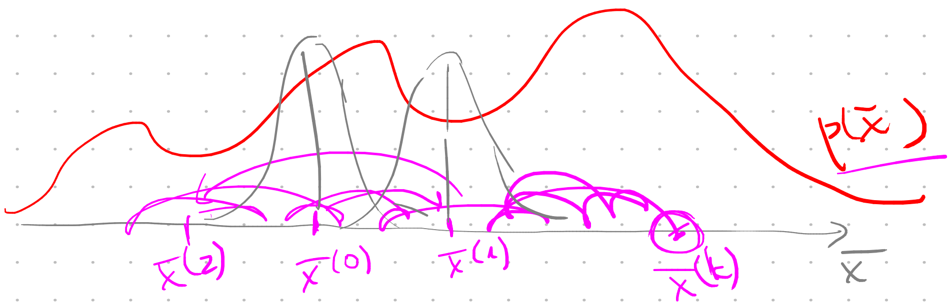
$$= E_q\left[\frac{1}{Z_p} f(\bar{x})\, p(x_3 = a \mid x_1) \right]$$

Likelihood weighted sampling



$$\frac{p^*(\bar{x})}{q(\bar{x})} = \prod_{x_i \in \text{Evidence}} p(x_i \mid par(x_i))$$

(8) MCMC – Markov chain Monte Carlo



$$\bar{x}^{(0)}, \bar{x}^{(1)}, \bar{x}^{(2)}, \dots$$

$$q(\bar{x}^{(t)} \mid \bar{x}^{(t-1)})$$

$q^{(0)}(\bar x)$

$\bar x^{(0)} \sim q^{(0)}(\bar x)$

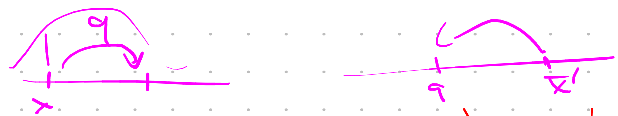
$$q^{(1)}(\bar x) = \int q(\bar x^{(1)}, \bar x^{(0)})\, d\bar x^{(0)} = \int q^{(0)}(\bar x^{(0)})\, \underbrace{q(\bar x | \bar x^{(0)})}_{T(\bar x, \bar x')}\, d\bar x^{(0)}$$

$$\boxed{q^{(t+1)}(\bar x) = \int q(\bar x|\bar x')\, q^{(t)}(\bar x')\, d\bar x'}$$

$q^{(0)}, q^{(1)}, q^{(2)}, \dots, q^{(t)}, \dots \to \pi(\bar x)$ – stationary distribution

$\pi(\bar x)$: $$\boxed{\pi(\bar x) = \int \boxed{q(\bar x|\bar x')}\, \pi(\bar x')\, d\bar x'}$$

$\pi(\bar x) = p(\bar x)$



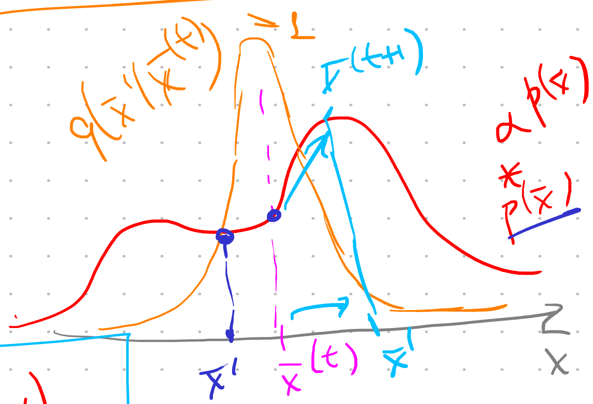
Условие баланса: если $\forall \bar x, \bar x'$ $p(\bar x)\, q(\bar x'|\bar x) = p(\bar x')\, q(\bar x|\bar x')$,

то $p(\bar x)$ — стац. распр. для $q(\bar x'|\bar x)$

$$p(\bar x) \overset{?}{=} \int p(\bar x')\, q(\bar x|\bar x')\, d\bar x' = \int p(\bar x)\, \underbrace{q(\bar x'|\bar x)\, d\bar x'}_{=1} = p(\bar x)$$

9 Metropolis-Hastings algorithm

Markov chain: $\bar x^{(0)}, \bar x^{(1)}, \dots, x^{(t)}, \dots$



- sample $\bar x' \sim q(\bar x'|\bar x^{(t)})$
- $$a(\bar x', \bar x^{(t)}) = \frac{p^*(\bar x')}{p^*(\bar x^{(t)})} \cdot \boxed{\frac{q(\bar x^{(t)}|\bar x')}{q(\bar x'|\bar x^{(t)})}}$$

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

- if $a \geq 1$ then accept $\bar x^{(t+1)} = \bar x'$

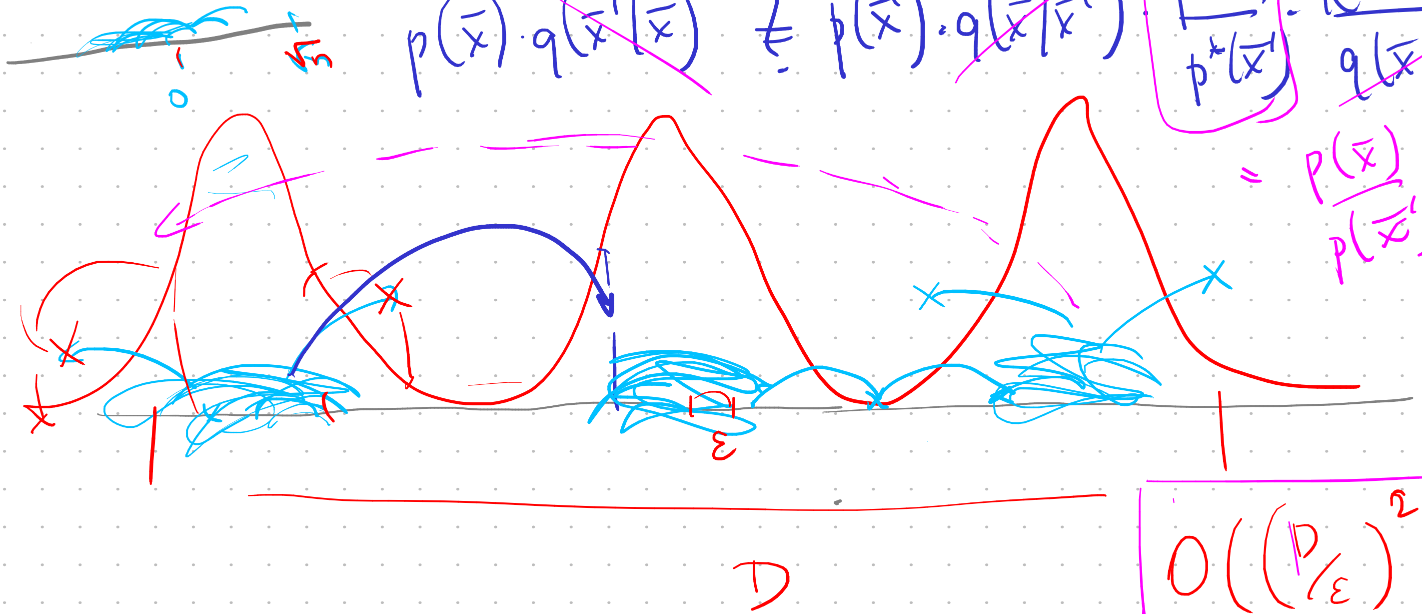
else $$\bar x^{(t+1)} = \begin{cases} \bar x^{(t)} & \text{с вероятн. } 1-a \\ \bar x' & \text{с вероятн. } a \end{cases}$$

$\forall \bar{x}, \bar{x}' \quad p(\bar{x}) \cdot p_{MC}(\bar{x}'|\bar{x}) \stackrel{?}{=} p(\bar{x}')\, p_{MC}(\bar{x}|\bar{x}')$

$a(\bar{x}', \bar{x}) \geq 1$: $\quad p_{MC}(\bar{x}'|\bar{x}) = q(\bar{x}'|\bar{x})$

$$p(\bar{x}) \cdot \cancel{q(\bar{x}'|\bar{x})} \stackrel{?}{=} p(\bar{x}') \cdot \cancel{q(\bar{x}|\bar{x}')} \cdot \overbrace{\boxed{\frac{p^*(\bar{x})}{p^*(\bar{x}')}} \cdot \frac{\cancel{q(\bar{x}'|\bar{x})}}{\cancel{q(\bar{x}|\bar{x}')}}}^{a(\bar{x}, \bar{x}')}$$

$$= \frac{p(\bar{x})}{p(\bar{x}')}$$



$$\boxed{O\left(\left(D/\varepsilon\right)^2\right)}$$

$$\boxed{\bar{\theta} \sim p(\bar{\theta}|D)}$$

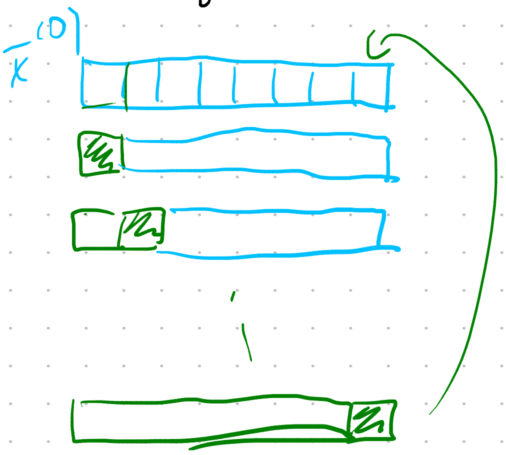
10 Gibbs sampling

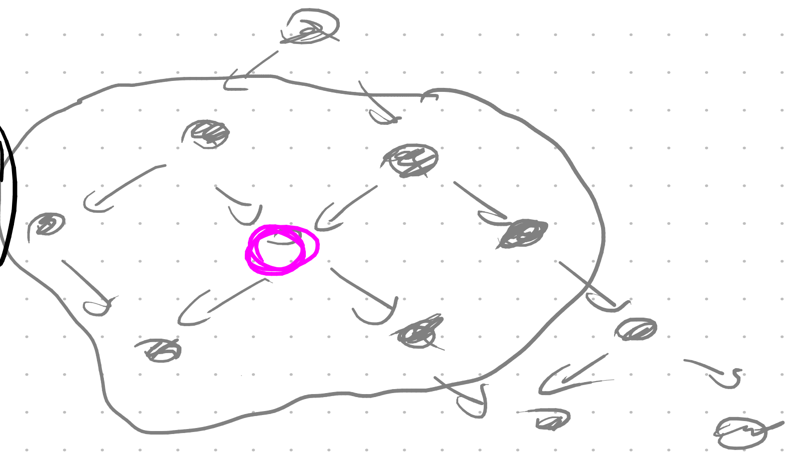
$$p(\bar{x}) = p(x_1, x_2, \ldots, x_n)$$

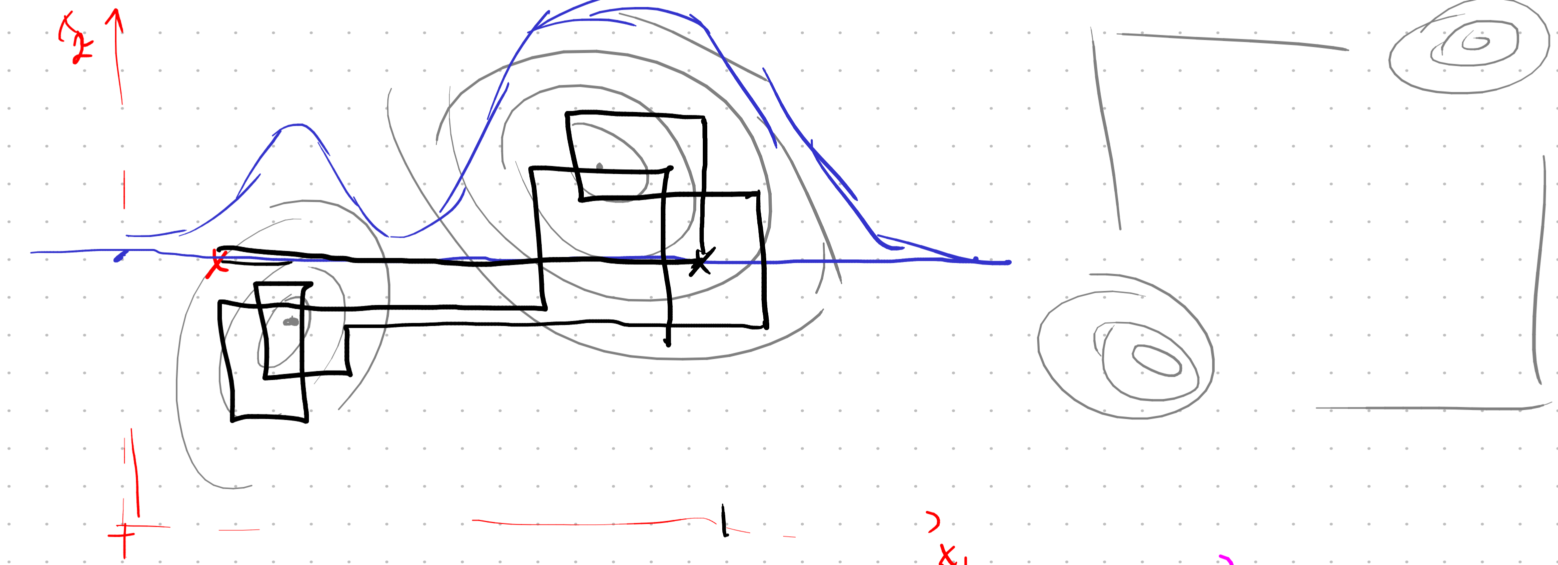
$$x_i \sim p(x_i | x_1 \ldots x_{i-1}, x_{i+1}, \ldots, x_n)$$

- init $\bar{x}^{(0)}$
- loop over k:
  - for $i = 1 \ldots n$:

$$x_i^{(k+1)} \sim p\left(x_i \,\middle|\, x_1^{(k+1)}, \ldots, x_{i-1}^{(k+1)}, x_{i+1}^{(k)}, \ldots, x_n^{(k)}\right)$$







i-й шаг:

$$q(\bar{x}' | \bar{x}) = \begin{cases} 0, & \bar{x}'_{-i} \neq \bar{x}_{-i} \\ p(x'_i | \bar{x}_{-i}), & \bar{x}'_{-i} = \bar{x}_{-i} \end{cases}$$

$= (x_1 \dots x_{i-1}, x_{i+1} \dots x_n)$

← всегда принимает

$\boxed{\bar{x}'_{-i} = \bar{x}_{-i}}$

$$a(\bar{x}', \bar{x}) = \frac{p^*(\bar{x}')}{p^*(\bar{x})} \cdot \frac{q(\bar{x} | \bar{x}')}{q(\bar{x}' | \bar{x})} = \frac{p(\bar{x}')}{p(\bar{x})} \cdot \frac{p(x_i | \bar{x}'_{-i})}{p(x'_i | \bar{x}_{-i})} = \frac{p(x'_i | \bar{x}'_{-i})\, \cancel{p(\bar{x}'_{-i})}}{p(x_i | \bar{x}_{-i})\, \cancel{p(\bar{x}_{-i})}} \times \frac{p(x_i | \bar{x}'_{-i})}{p(x'_i | \bar{x}_{-i})}$$

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

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

11 Slice sampling

$p^*(\bar{x}^{(k)})$

$\bar{x} \sim p(\bar{x}) \Leftrightarrow (\bar{x}, y) \sim \mathrm{Unif}(\text{график } p^*)$