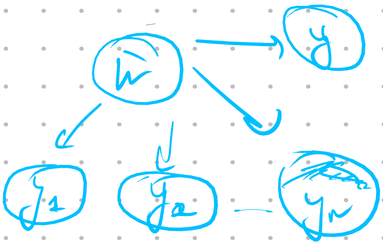
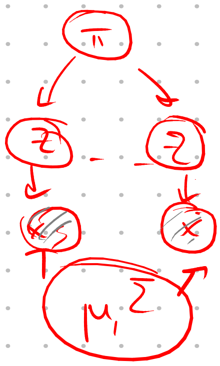
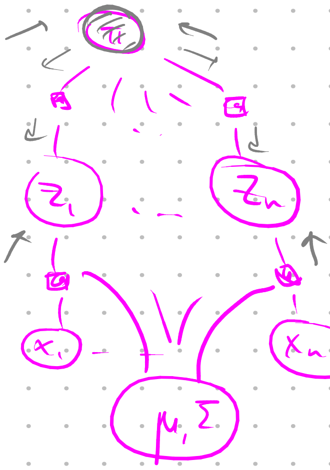






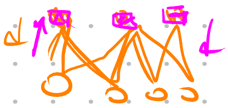
Approximate inference

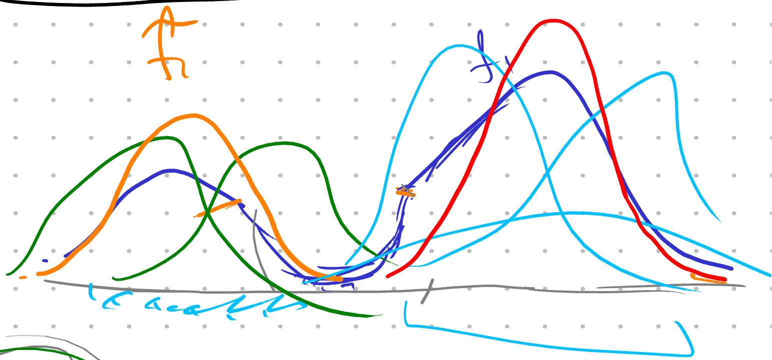
Complex factors

Factor graphs w/ cycles

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

Loopy belief propagation





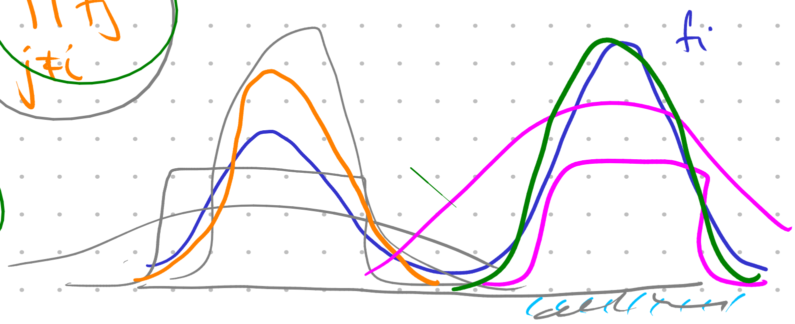
Expectation Propagation

$$F(X) = \underline{f_1(X_1)\, f_2(X_2) \ldots f_n(X_n)}$$

$$\tilde{F}_{naive} = \tilde{f}_1 \times \tilde{f}_2 \times \ldots \times \tilde{f}_n$$

$$\tilde{f}_i \prod_{j\neq i} f_j \approx f_i \cdot \prod_{j\neq i} f_j$$

$$\tilde{f}_i(x_1, x_2, x_3) \cdot \prod f_j(\_\ \_)$$



Sampling $\quad p(\bar{x}) \rightsquigarrow \bar{x} \sim p(\bar{x})$ $\qquad z \sim Unif([0,1])$ rand()

$$\bar{x} \text{ --- } \boxed{?} \text{ --- } p_{\gamma}(\bar{x}) = \underline{p(\theta \mid D)}$$

$$\bar{x} \text{ --- } \boxed{?} \text{ --- } p^*(\bar{x}) = p(\theta)\, p(D \mid \theta)$$

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

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

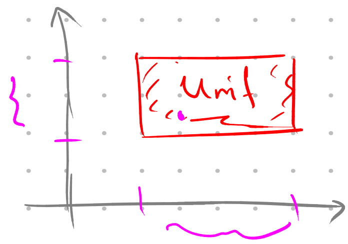
2) $Q(\theta, \theta^{(m)}) = \mathbb{E}_{p(z|\theta^{(m)})}\left[\log p(x,z|\theta)\right] \approx$

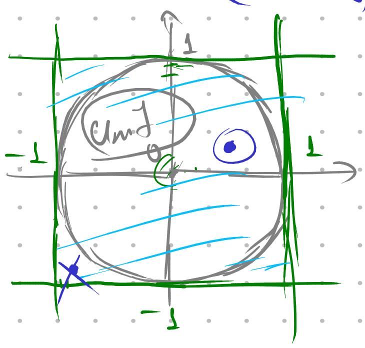
$$\approx \boxed{\frac{1}{R}\sum_{r} \log p(x, z^{(r)}|\theta)} \xrightarrow[\theta]{} \max, \quad \text{где } z^{(r)} \sim p(z|\theta^{(m)}).$$

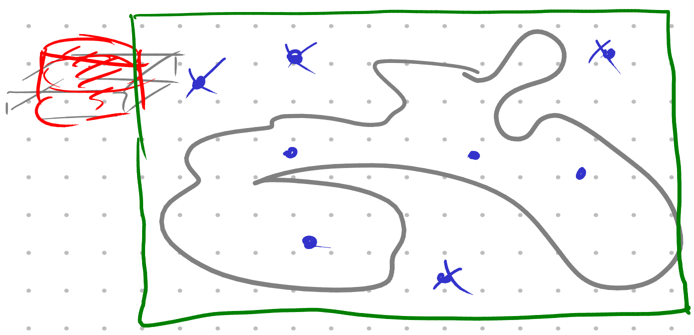
Monte Carlo EM

① Известные распределения

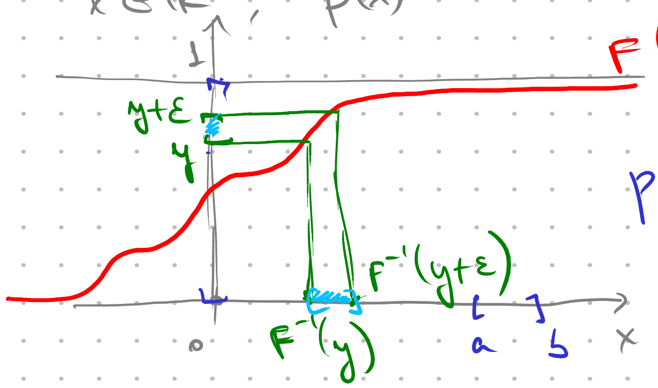
$x \sim Unif([0,1]) \xrightarrow{a+(b-a)x} Unif([a,b])$







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



$$F(x) = \int_{-\infty}^{x} p(y)\,dy$$

$$Pr(x \in [a,b]) = Pr(y \in [F(a), F(b)]) = F(b) - F(a)$$

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

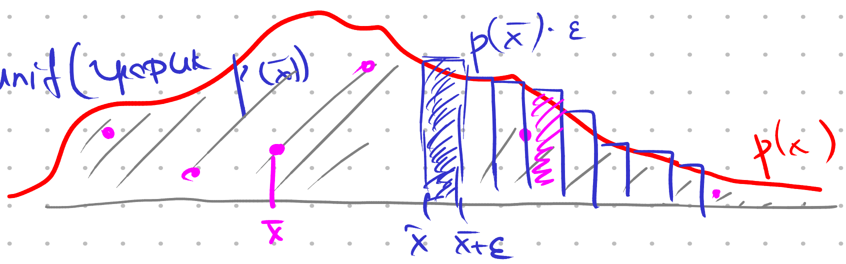
$\varepsilon \to 0$

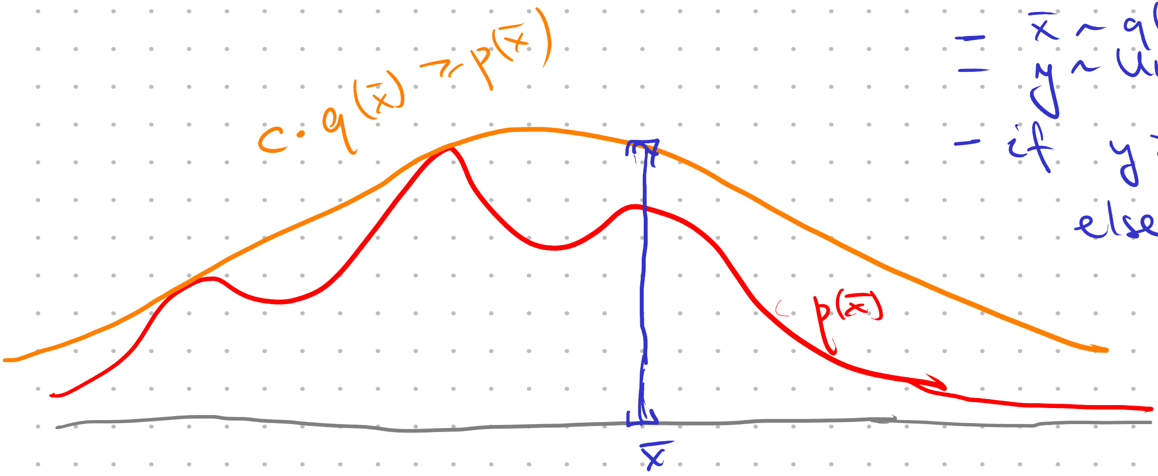
$$q(x) = F'(x) = p(x)$$

$y \sim Unif([0,1]) \Rightarrow x = F^{-1}(y)$

② Rejection sampling

$x \sim p(\bar{x}) \Leftrightarrow (\bar{x}, y) \sim Unif($подграфик





- $\bar{x} \sim q(\bar{x})$
  $y \sim \mathrm{Unif}([0, c \cdot q(\bar{x})])$
- if $y > p(\bar{x})$ drop
  else accept $\bar{x}$

3 Importance sampling

$x \to \boxed{?} \to p(\bar{x})$! $\bar{x} \sim q(\bar{x})$

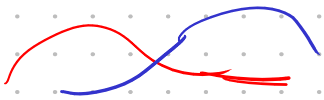
$\boxed{\text{если } p \neq 0 \text{, то и } q \neq 0}$

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

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

importance weights

где $x^{(r)} \sim q(\bar x)$



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

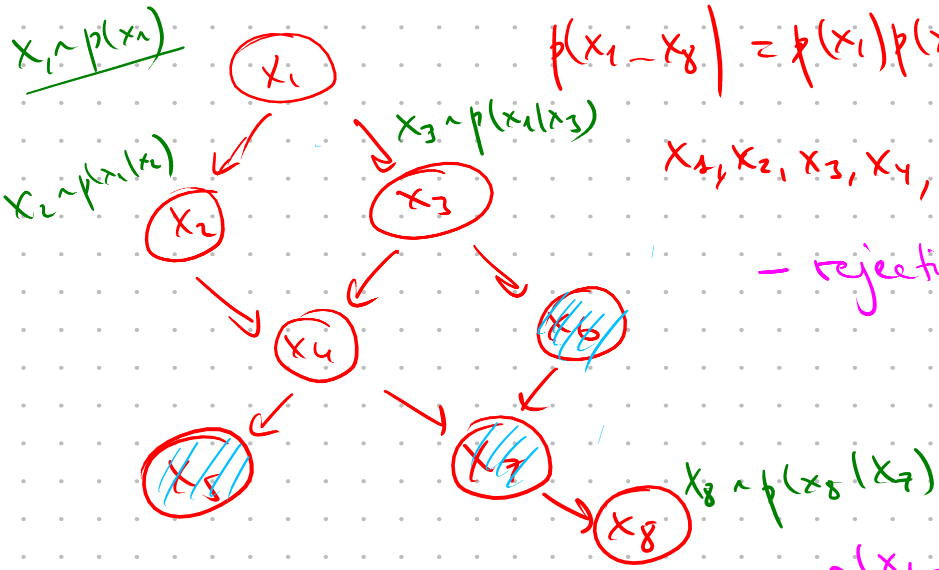
$$= \frac{1}{R}\left(\sum_r f(\bar x^{(r)}) \cdot \frac{p^*(\bar x^{(r)})}{q^*(\bar x^{(r)})}\right) \cdot \frac{Z_q}{Z_p}$$

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

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

$$= \mathbb{E}_q\left[\frac{p^*}{q^*}\right] \approx \frac{1}{R}\sum \frac{p^*(\bar x^{(r)})}{q^*(\bar x^{(r)})}$$

$$\approx \frac{\sum_{r=1}^{R} f(\bar x^{(r)}) \overbrace{\boxed{\frac{p^*(\bar x^{(r)})}{q^*(\bar x^{(r)})}}}^{= w^{(r)}}}{\sum_{r=1}^{R} \frac{p^*(\bar x^{(r)})}{q^*(\bar x^{(r)})}}$$



$p(x_1 \ldots x_8) = p(x_1)p(x_2|x_1) \ldots p(x_8|x_7)$

$x_1, x_2, x_3, x_4, x_8 \sim p(x_1 \ldots x_4, x_8 | x_5, x_6, x_7)$

– rejection sampling

$q(x_1 \ldots x_8)$, $x_5, x_6, x_7 \sim \delta(x_5)$

– importance sampling

$q(x_1 \ldots x_4, x_8) = p(x_1)p(x_2|x_1)p(x_3|x_1)p(x_4|x_2,x_3)p(x_8|x_7)$

likelihood weighted sampling

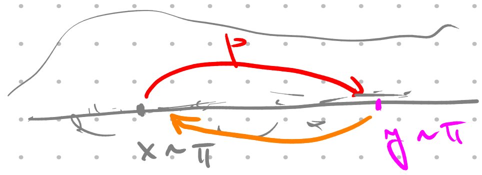
$$W = \frac{p(\bar{x})}{q(\bar{x})} = \frac{p(x_1 \ldots x_4, x_5, \ldots, x_8)}{q(x_1 \ldots x_4, x_8)} = \frac{p(x_1 \ldots x_8)}{q(\ldots)} =$$

$$= \frac{\cancel{p(x_1)}\,\cancel{p(x_2|x_1)} \ldots \cancel{p(x_8|x_7)}}{\cancel{p(x_1)} \ldots \cancel{p(x_8|x_7)}} = \frac{p(x_5|x_4)p(x_6|x_3)p(x_7|x_4,x_6)}{1}$$

4 Markov Chain Monte Carlo (MCMC)

$x_1 \to x_2 \to \ldots \to x_t \to x_{t+1} \to \ldots$

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



Стационарное распр. $\pi(x) = \int p(x|y)\pi(y)\,dy$

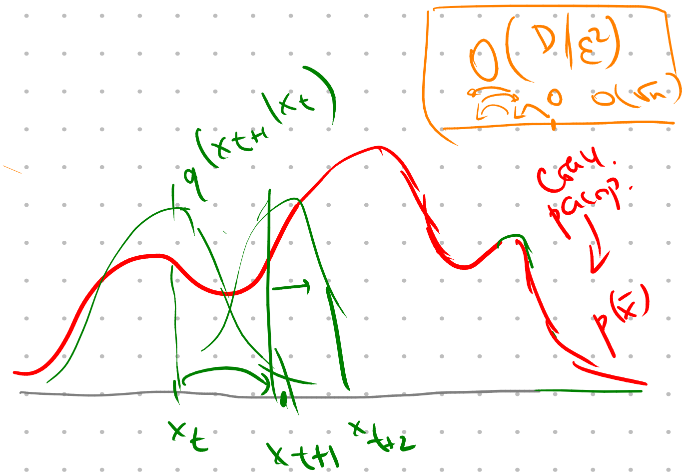
$x_1 \sim p(x)$, $x_2 \sim p(x_2|x_1)$, ..., ; ... $p(x_t) \xrightarrow[t \to \infty]{} \pi$

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

то π – стац. распр. для $p(y|x)$

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

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



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

Metropolis – Hastings algorithm:

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

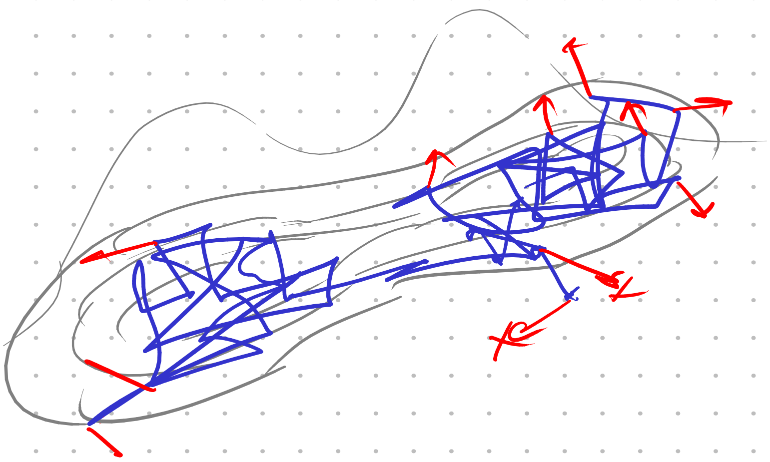
    else

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

$$p(\bar{x}) \cdot p_{MC}(\bar{y} \mid \bar{x}) \overset{?}{=} p(\bar{y})\, p_{MC}(\bar{x} \mid \bar{y})$$

$$a(\bar{x}, \bar{y}) \geq 1 \Rightarrow \cancel{p(\bar{x})} \cdot q(\cancel{\bar{y} \mid \bar{x}}) \overset{?}{=} \cancel{p(\bar{y})}\, q(\cancel{\bar{x} \mid \bar{y}}) \cdot \frac{\cancel{p(\bar{x})}}{\cancel{p(\bar{y})}} \cdot \frac{q(\cancel{\bar{y} \mid \bar{x}})}{q(\cancel{\bar{x} \mid \bar{y}})}$$

$=$

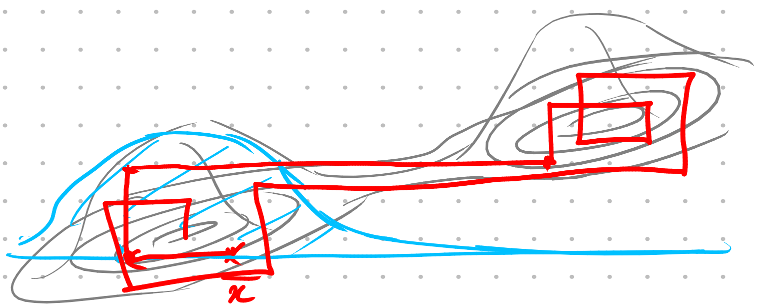


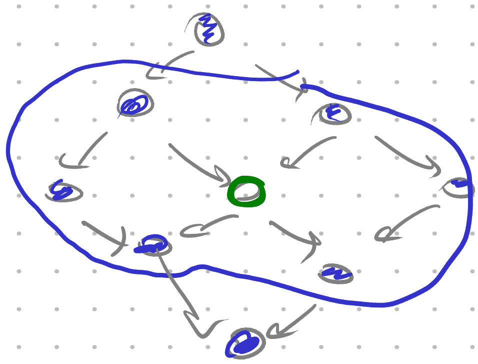
5 Gibbs sampling

$p(x_1, x_2, \ldots, x_n)$

нужно мы умеем $x_i \sim p(x_i \mid \bar{x}_{-i})$

- repeat, $i = 1, \ldots, n, 1, \ldots, n$ —
  - $x_i \sim p(x_i \mid \bar{x}_{-i})$





$$q(\bar{x}' \mid \bar{x}) : \begin{cases} 0 & \text{если } \bar{x}'_{-i} \neq \bar{x}_{-i} \\ p(x_i \mid \bar{x}_{-i}) & \text{если } \bar{x}'_{-i} = \bar{x}_{-i} \end{cases}$$

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

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

$$= 1$$