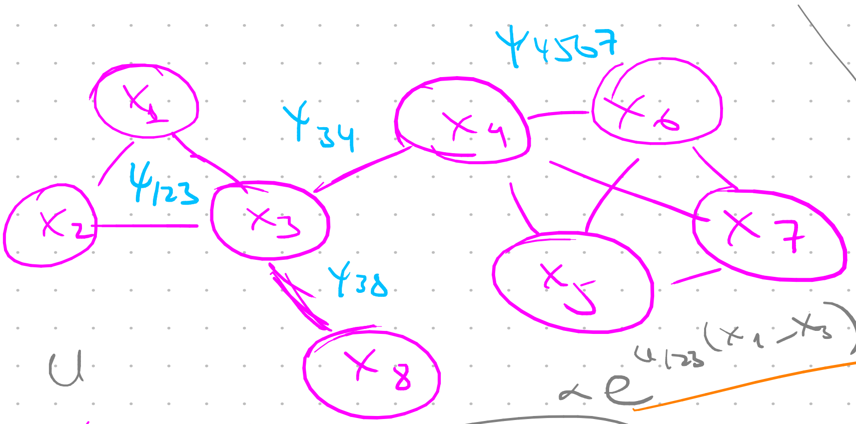
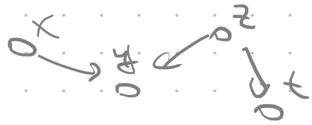
① Undirected graphical models



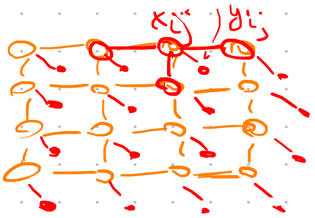


$p(x,y,z,t) = p(x)\,p(z)\,p(y|x,t)\,p(t|z)$

U

$\propto e^{\psi_{123}(x_1 - x_3)}$

$$F(x_1 \dots x_8) = \psi_{123}(x_1,x_2,x_3)\,\psi_{34}(x_3,x_4)\,\psi_{4567}(x_4 \dots x_7)\,\psi_{38}(x_3,x_8)$$

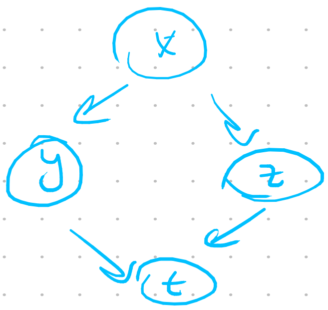


$$p(x,y) = \prod_i \underbrace{p(y_i|x_i)}_{\propto e^{-\alpha \cdot x_i y_i}} \cdot \prod_i \prod_{j \in Nei(i)} \underbrace{p(x_i|x_j)}_{\propto e^{\beta x_i x_j}}$$

$\Longrightarrow \max_{x_i}$

MRF

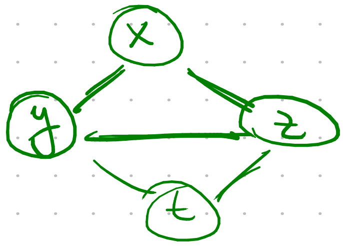
$$\min_{x} \left(\sum_i \alpha\, x_i y_i + \sum_i \sum_{j \in Nei(i)} \beta \cdot x_i x_j \right)$$



$$p(x,y,z,t) = \underbrace{p(x)}_{\psi_x} \underbrace{p(y|x)}_{\psi_{xy}} \underbrace{p(z|x)}_{\psi_{xz}} \underbrace{p(t|y,z)}_{\psi_{yzt}}$$

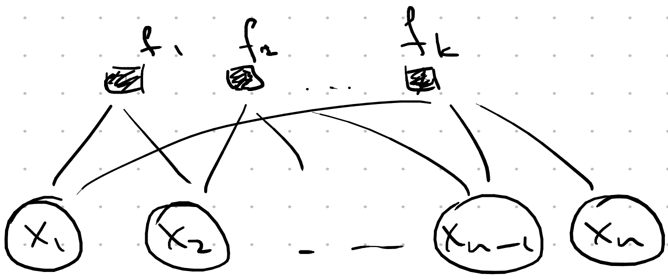
$(\psi_x \to \psi_{xy})$

$\psi_{xyz} \cdot \psi_{yzt}$



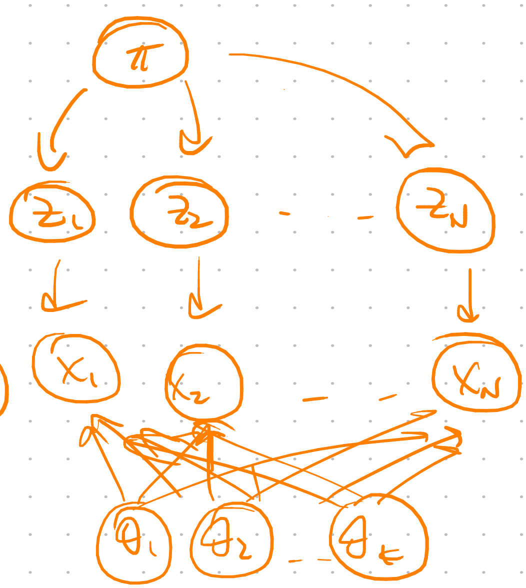
② Factor graphs

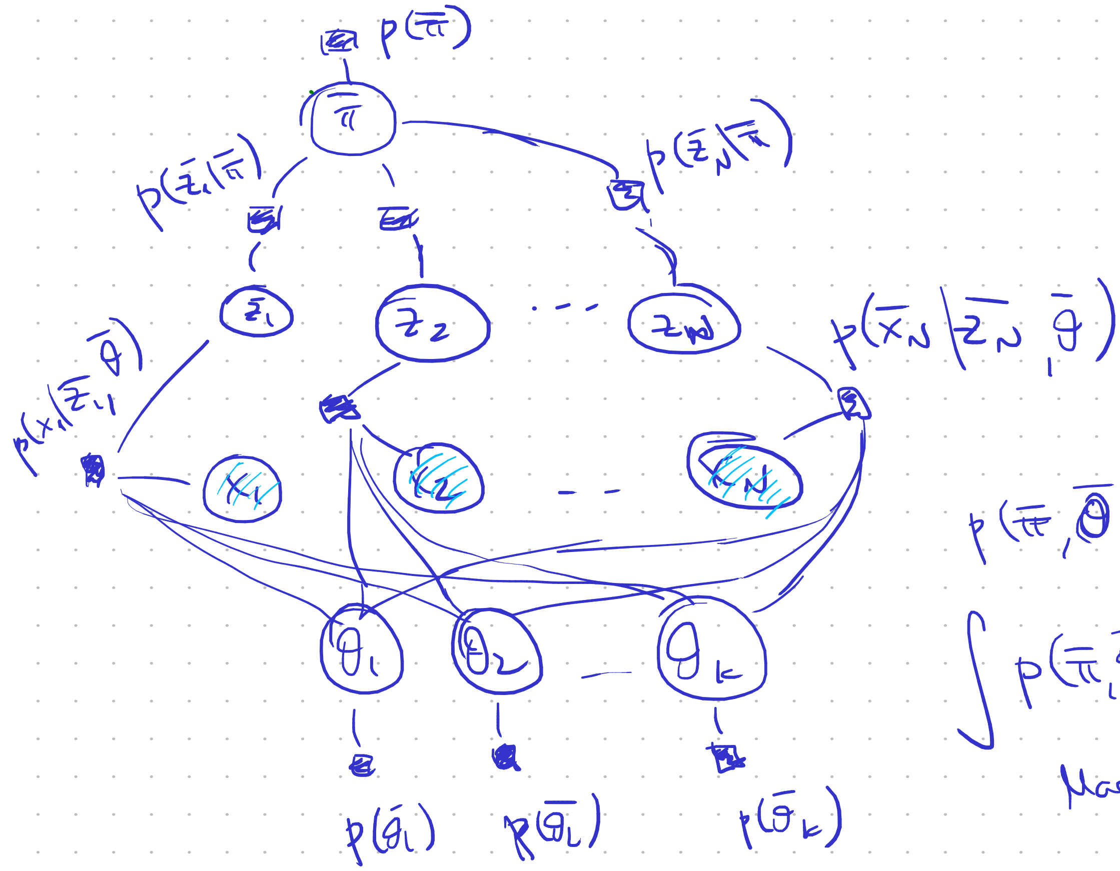
$$F(x_1, \dots, x_n) = f_1(X_1)\,f_2(X_2) \dots f_k(X_k)$$

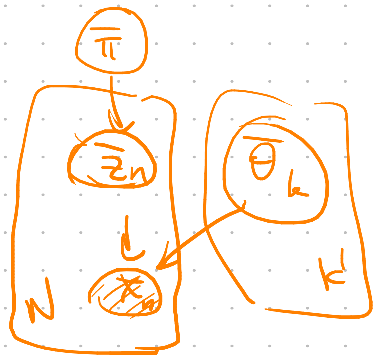


$$p(\bar{\pi}, \bar{\theta}, z, X) = p(\bar{\pi})\,p(\bar{\theta}) \prod_k p(\bar{\theta}_k)$$

$$\prod_n p(z_n|\bar{\pi}) \cdot \prod_n p(x_n|z_n, \bar{\theta})$$







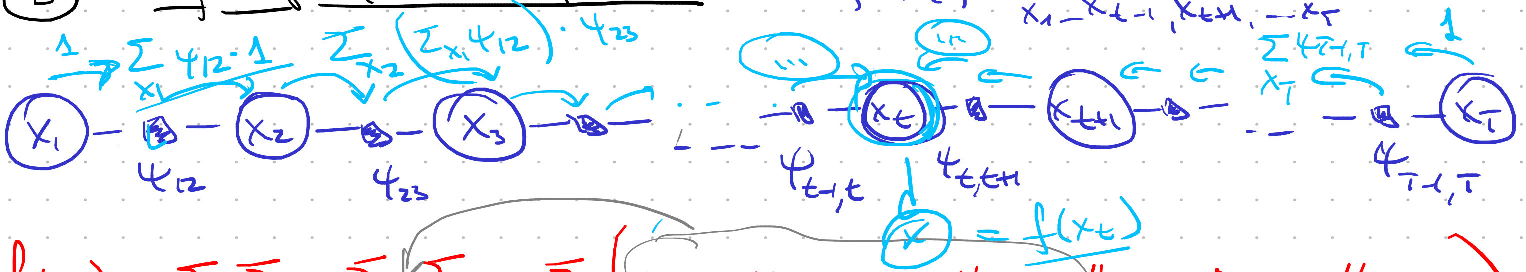
$$p(\bar\pi,\bar\theta|X) = \int p(\bar\pi,\bar\theta,z|x)\,dz$$

Marginalization

$$\sum_{X_{-i}} F(x_1,\dots,x_n) = f(x_i) \quad \text{— Маргинал}$$

③ Пример: цепочка переменных

$$f(x_t) = \sum_{x_1..x_{t-1},x_{t+1},..x_T} F(x_1..x_T)$$



$$f(x_t) = \sum_{x_1}\sum_{x_2}\dots\sum_{x_{t-1}}\sum_{x_{t+1}}\dots\sum_{x_T}\Big(\underbrace{\psi_{12}\cdot\psi_{23}\cdot\ldots\cdot\psi_{t-1,t}}_{F(x_1,\dots,x_{t-1},x_t)}\cdot\underbrace{\psi_{t,t+1}\cdot\ldots\cdot\psi_{T-1,T}}_{F(x_t,x_{t+1},\dots,x_T)}\Big)$$

$$= \sum_{x_1 .. x_{t-1}}\Big[\psi_{12}..\psi_{t-1,t}\cdot\sum_{x_{t+1}..x_T}\psi_{t,t+1}..\psi_{T-1,T}\Big] =$$

$$= \Big[\sum_{x_1}..\sum_{x_{t-1}}\psi_{12}..\psi_{t-1,t}\Big]\times\Big[\sum_{x_{t+1}}..\sum_{x_T}\psi_{t,t+1}..\psi_{T-1,T}\Big]$$

$$\underbrace{\sum_{x_{t-1}}\sum_{x_{t-2}}\dots\sum_{x_2}\sum_{x_1}}_{O(n^{t-1})}\big(\psi_{12}\psi_{23}\dots\psi_{t-1,t}\big) = O(n^t)$$

$$= \sum \cdots \sum_{x_2} \psi_{23} \cdots \psi_{t-1,t} \left(\sum_{x_1} \psi_{12}\right) =$$

2

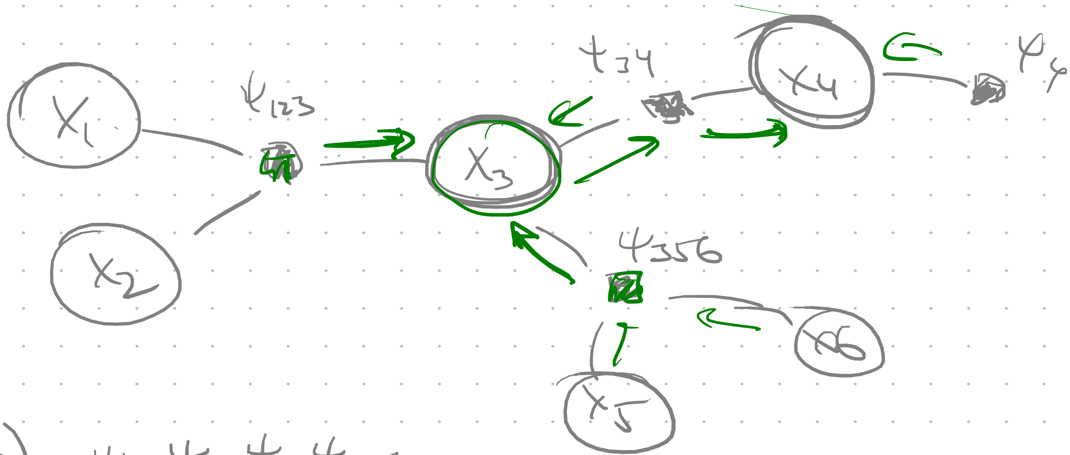
$$= \sum_{x_{t-1}} \cdots \sum_{x_3} \psi_{34} \cdots \psi_{t-1,t} \left(\sum_{x_2} \psi_{23} \left(\sum_{x_1} \psi_{12}\right)\right) = \cdots$$

n^2

$$= \sum_{x_{t-1}} \psi_{t-1,t} \left[\sum_{x_{t-2}} \psi_{t-2,t-1} \left[\sum \cdots \overset{2\cdot 2}{\sum_{x_3}} \psi_{34} \underbrace{\left[\overset{2\cdot 2}{\sum_{x_2}} \psi_{23} \underbrace{\left[\overset{2\cdot 2}{\sum_{x_1}} \psi_{12}\right]}\right]} \cdots \right]\right]$$

$O(n^2 \cdot t)$

$$\sum_{x_{t+1}} \cdots \sum_{x_T} \left(\psi_{t,t+1} \cdots \psi_{T-1,T}\right) = \sum_{x_{t+1}} \psi_{t,t+1} \left[\sum_{x_{t+2}} \psi_{t+1,t+2} \cdots \left[\sum_{x_{T-1}} \psi_{T-2,T-1} \left[\sum_{x_T} \psi_{T-1,T}\right] \cdots \right]\right]$$

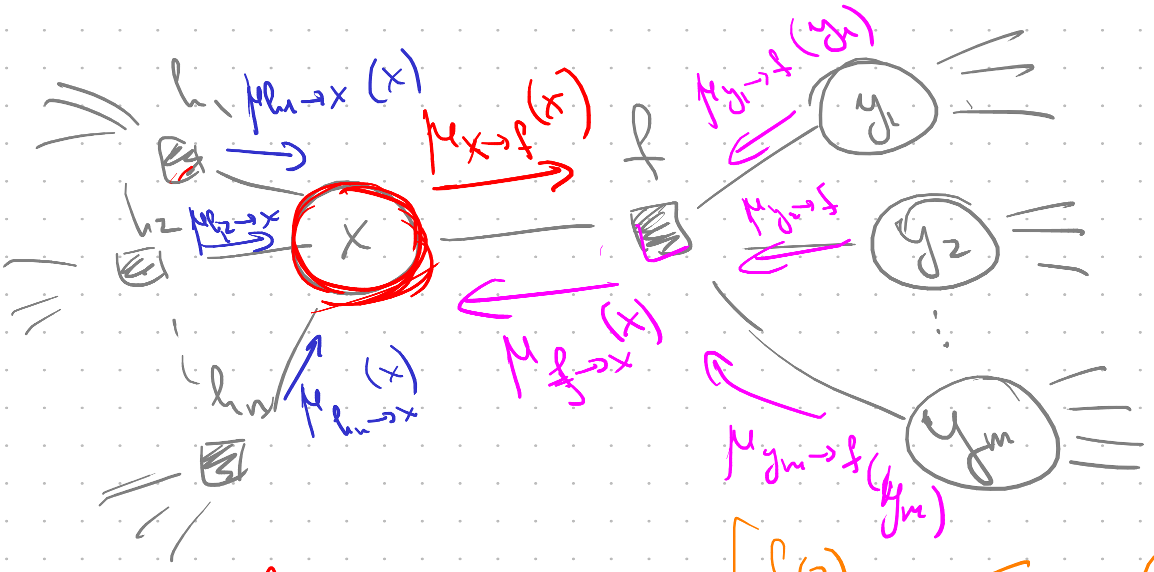


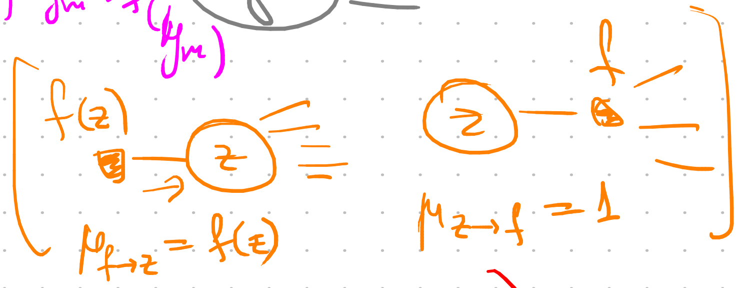
$$F(x_1 \ldots x_6) = \psi_{123}\,\psi_{34}\,\psi_4\,\psi_{356}$$

$$f(x_3) = \sum_{\cdots \cancel{x_3} \cdots} \psi_{123} \cdots \psi_{356} = \left(\sum_{x_1,x_2} \psi_{123}\right) \cdot \left(\sum_{x_5,x_6} \psi_{356}\right) \cdot \left(\sum_{x_4} \psi_4\,\psi_{34}\right)$$

$$f(x_4) = \psi_4 \cdot \sum_{x_3} \left[\psi_{34} \cdot \underbrace{\left(\sum_{x_1,x_2} \psi_{123}\right)} \cdot \underbrace{\left(\sum_{x_5,x_6} \psi_{356}\right)}\right]$$

④ <u>Message passing algorithm</u>





$\mu_{f\to z} = f(z)$

$\mu_{z\to f} = 1$

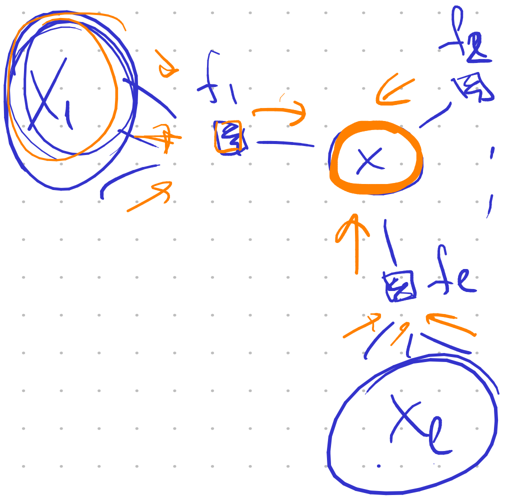
$$\mu_{x\to f}(x) = \prod_{i=1}^{n} \mu_{h_i\to x}(x)$$

$$\mu_{f\to x}(x) = \sum_{y_1\dots y_m} f(x, y_1, \dots, y_m)\cdot \mu_{y_1\to f}(y_1)\dots\mu_{y_m\to f}(y_m)$$

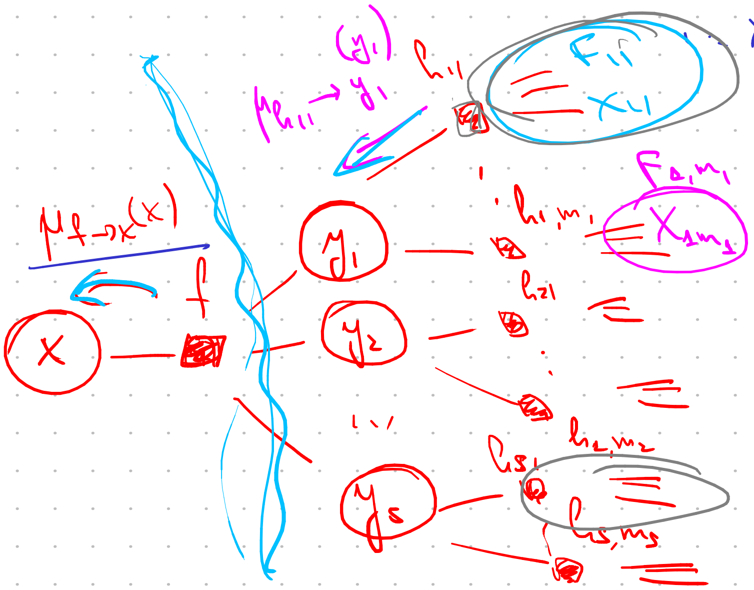
грубой силой

$$f(x) = \sum_{x_1 - x_n} F(x, x_1 - x_n) = \sum_{X_1 - X_N} f_1(x, X_1)\cdot F_1(X_1)\cdot f_2(x, X_2)\cdot F_2(X_2)\cdot \dots \cdot f_l(x, X_l)\cdot F_l(X_l) =$$

$$\sum_{X_1, X_2, \dots, X_l}$$



$$= \left[\sum_{X_1} f_1(x, X_1)\cdot F_1(X_1)\right] \times \dots \times \left[\sum_{X_l} f_l(x, X_l)\, F_l(X_l)\right]$$



$$\mu_{f\to x}(x) \overset{?}{=} \sum_{X} f(x, X)\cdot F(X)$$

Дано:

$$\mu_{h_{ij}\to y_i}(y_i) = \sum_{X_{ij}} h_{ij}(y_i, X_{ij})\cdot F_{ij}(X_{ij})$$

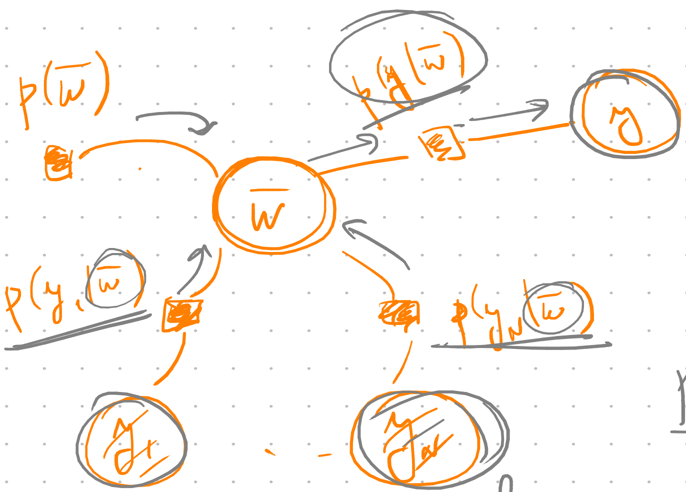
$$\mu_{f\to x}(x) = \sum_{y_1 - y_s} f(x, y_1 - y_s) \cdot \prod_{s=1}^{S} \mu_{y_s \to f}(y_s) =$$

$$= \sum_{y_1 - y_s} f(x, y_1 - y_s) \cdot \prod_{s=1}^{S} \left(\prod_{j=1}^{m_j} \mu_{h_{sj} \to y_s}(y_s) \right) =$$

$$= \sum_{y_1 - y_s} f(x, y_1 - y_s) \prod_{s=1}^{S} \prod_{j=1}^{m_j} \sum_{X_{sj}} h_{sj}(y_s, X_{sj}) \cdot F_{sj}(X_{sj}) =$$

$$= \sum_{y_1 - y_s} \sum_{X_{11}, \dots, X_{s,m_s}} f(x, y_1 - y_s) \cdot \prod_{s=1}^{S} \prod_{j=1}^{m_j} h_{sj}(y_s, X_{sj}) \cdot F_{sj}(X_{sj})$$

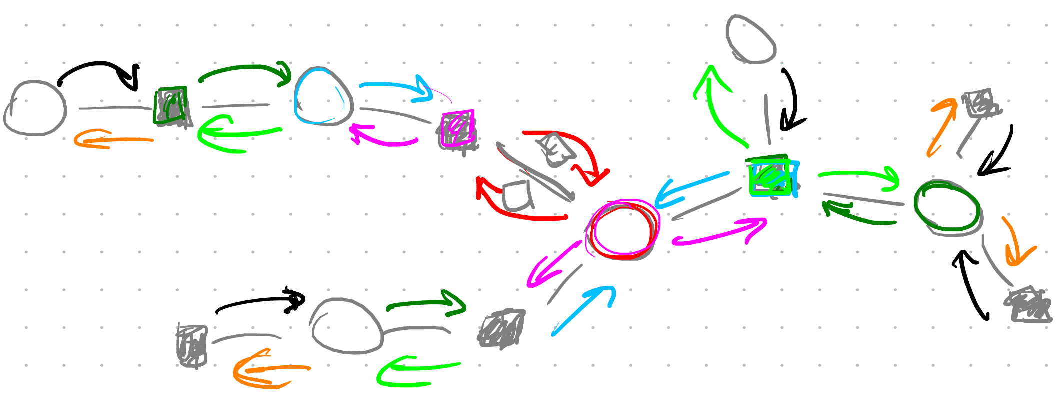
$$= F(X)$$

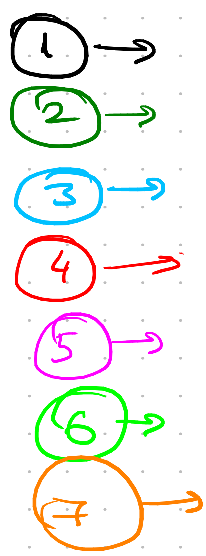


$$p(y | y_1 \dots y_N) = \int p(y, \bar{w} | y_1 \dots y_N) \, d\bar{w}$$

$$\mu_{\bar{w} \to p(y|\bar{w})} = p(\bar{w}) \cdot p(y_1 | \bar{w}) \dots p(y_N | \bar{w})$$

$$\mu_{p(y|\bar{w}) \to y} = \int p(y|\bar{w}) \cdot p(\bar{w}) \cdot p(y_1|\bar{w}) \dots p(y_N|\bar{w}) \, d\bar{w}$$



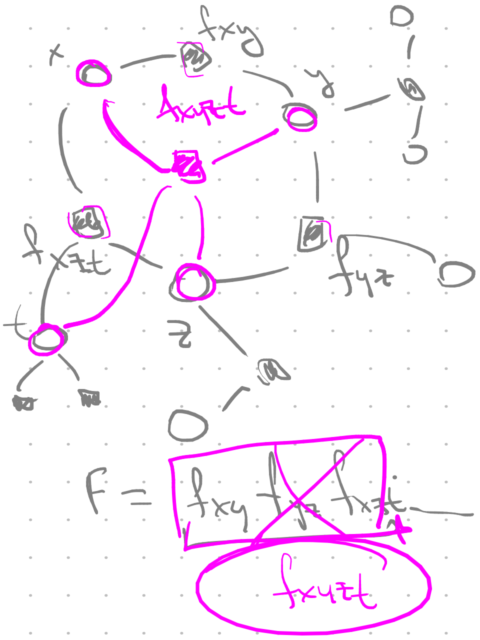


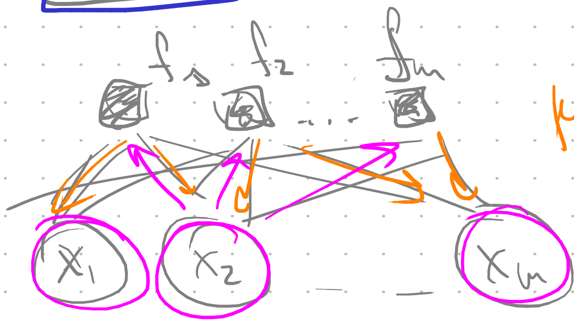
– фактор-граф – дерево
– факторы „простые" → Message Passing

не дерево → можно сделать деревом

дерево, но факторы сложные → Expectation Propagation

Approximate Inference → Loopy Belief Propagation; Sampling; Variational approximations





$\mu_{f\to x}(x) = \sum f(x, \_)$

$\mu_{x\to f}(x) = \prod \_\_$

5 Sampling and MC

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

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

Дано: 0.9

$q(\bar{x}) \propto p(\bar{x})$

900

$\bar{x}$ — ▢ — $q(\bar{x})$

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

$$p(\bar{x}|D) = \int p(\bar{x}|\bar{\theta})\,p(\bar{\theta}|D)\,d\bar{\theta} = E_{\boxed{p(\bar{\theta}|D)}}[p(\bar{x}|\bar{\theta})] \approx \frac{1}{R}\sum_{r=1}^{R} p(\bar{x}|\bar{\theta}_r)$$

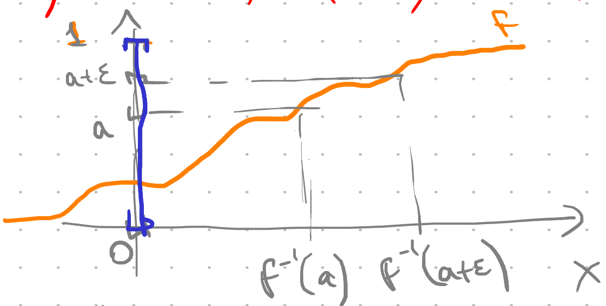
$\sim \bar{\theta}_r$

$$Q(\bar{\theta}, \bar{\theta}^{(m)}) = E_{p(z|\bar{\theta}^{(m)}, x)}[\log p(x, z|\bar{\theta})] \approx \frac{1}{R}\sum_{r} \log p(x, z_r|\bar{\theta})$$

Monte Carlo EM

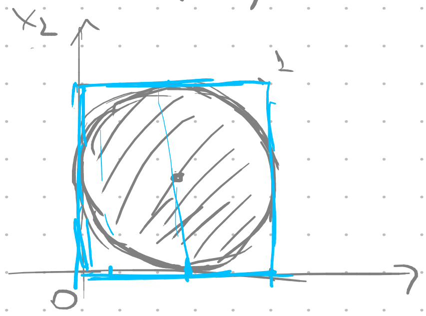
6 „Простые" методы семплирования

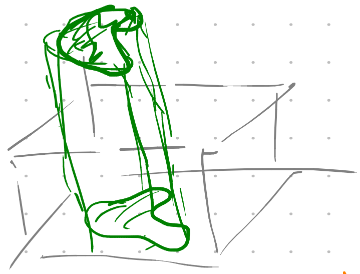
1) $x\in\mathbb{R}$, $p(x)$, $F(a)=\int_{-\infty}^{a} p(x)\,dx$, $z\sim Unif([0,1])$



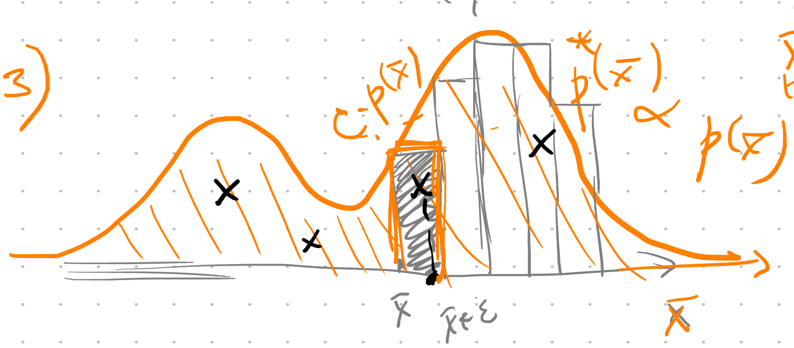
$$p(x\in[F^{-1}(a), F^{-1}(a+\varepsilon)]) = \int_{F^{-1}(a)}^{F^{-1}(a+\varepsilon)} p(x)\,dx = \varepsilon$$

2)





3)



$\in\mathbb{R}^d$ $\qquad$ $\in\mathbb{R}^{d+1}$ $\quad p^*(\bar{x})$

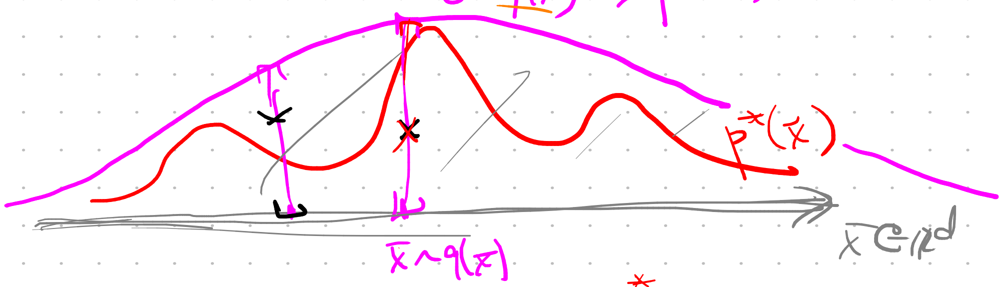
$\bar{x}\sim p(\bar{x}) \iff \bar{x}\sim Unif_{\bar{x}}($  $)$

$p^*(\bar{x}) = c\cdot p(\bar{x})$

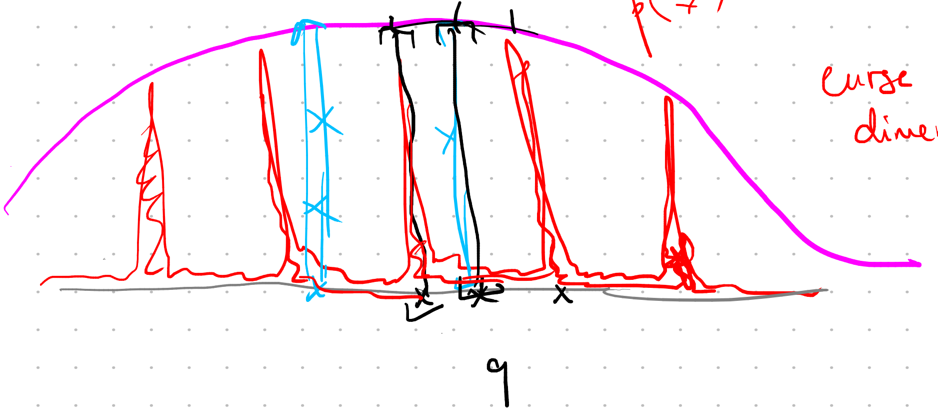
7 Rejection sampling

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

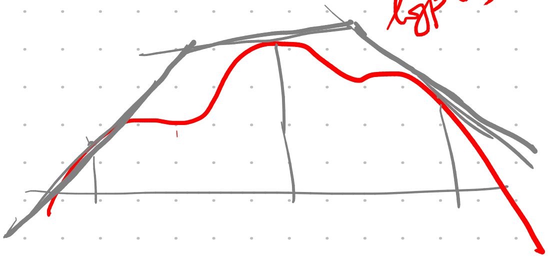
$(\bar{x}, y)$



- $\bar{x}\sim q(\bar{x})$
- $y\sim Unif([0, c\,q(\bar{x})])$
- if $y\leq p^*(\bar{x})$
  accept $\bar{x}$
  else reject $\bar{x}$



curse of dimensionality



8 Importance sampling

$\cancel{\bar{x} \sim p(\bar{x})}$

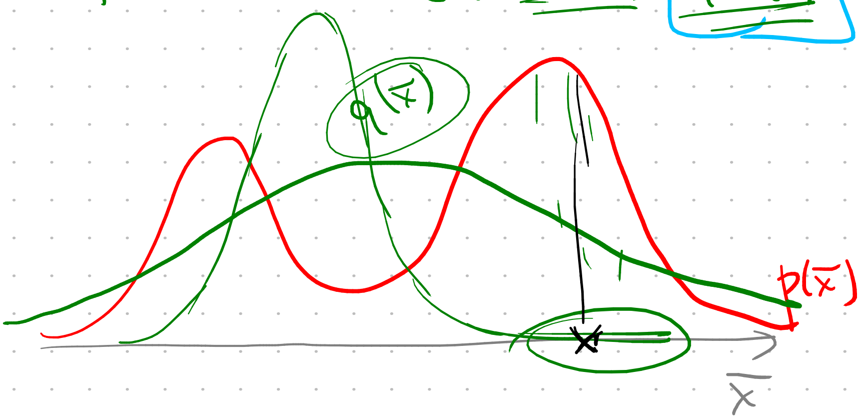
$$\mathbb{E}_{p(\bar x)}[f(\bar x)] = ?$$

$q(\bar x)$, $\bar x \sim q(\bar x)$

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

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

— importance weights



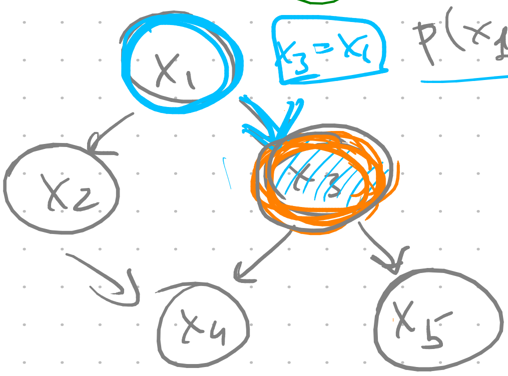
$$p(\bar x) = \frac{1}{Z_p}\,p^*(\bar x) \qquad \mathbb{E}_p[f] = \mathbb{E}_q\left[f\cdot\frac{p^*(\bar x)}{q^*(\bar x)}\cdot\frac{Z_q}{Z_p}\right]$$

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

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

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

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



$x_3 = x_1$

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

$$x_1, x_2, x_4, x_5 \sim p(\bar x_{-3}|x_3=a) = \frac{p(x_1)\,p(x_2|x_1)\,p(a|x_1)\,p(x_4|x_2,a)\,p(x_5|a)}{p(x_3=a)}$$

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

$$\mathbb{E}_p[f(\bar x)] = \mathbb{E}_q\left[f(\bar x)\cdot\frac{\cancel{p(x_1)}\,\cancel{p(x_2|x_1)}\,p(a|x_1)\,\cancel{p(x_4|x_2,a)}\,\cancel{p(x_5|a)}}{p(x_3=a)\cdot \cancel{p(x_1)}\,\cancel{p(x_2|x_1)}\,\cancel{p(x_4|x_2,a)}\,\cancel{p(x_5|a)}}\right]$$

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

Likelihood weighted sampling