17.02.17 гм II $\qquad p(x)=\frac{1}{Z}\exp(-E(x))$

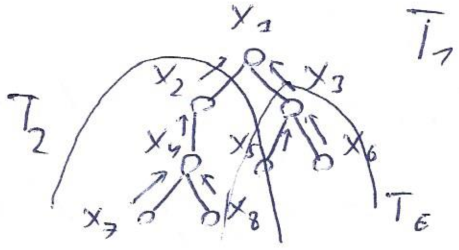
наиболее вероятная конфигурация

1) $X_{MP} = \arg\max_{x} p(x) = \arg\min_{x} (E(x))$

расчёт н.к. и маргинального распр-я

2) Z? $\quad p(x_i)$?

$$p(x_i) = \sum_{x\setminus x_i} p(x), \quad Z = \sum_{x} \hat{p}(x)$$



3) $p(X_{tr}|\Theta) \to \max_{\Theta}$

$$p(x) = \frac{1}{Z}\prod_{i\in V}\psi_i(x_i)\prod_{(i,j)\in\mathcal{E}}\psi_{ij}(x_i,x_j) =$$

$$= \frac{1}{Z}\exp\Big(-\sum_{i\in V}\varphi_i(x_i) - \sum_{(i,j)\in\mathcal{E}}\varphi_{ij}(x_i,x_j)\Big)$$

ф-ия Беллмана $\quad$ задача минимизации энергии на ациклическом графе

$$V^{T_1}(x_1) \triangleq \min_{x\setminus x_1} E^{T_1}(x)\Big|_{x_1}, \quad V^{T_5}(x_5) \triangleq \varphi_5(x_5)$$

$$V^{T_1}(x_1) = \min_{x_2\dots x_8}\Big(\varphi_1(x_1)+\varphi_{12}(x_1,x_2)+\varphi_{13}(x_1,x_3)+\sum+\dots\Big) \underset{\text{ацикл.}}{\overset{\text{дистр.}}{=}}$$

$$= \varphi_1(x_1) + \min_{x_2\dots x_8}\big(\varphi_{12}+\varphi_{13}+\sum+\dots\big) = \varphi_1 + \min_{\substack{x\in T_2\\ x\in T_3}}\big(\varphi_{12}+\varphi_{13}+$$

// $\min(a+b, a+c) = a+\min(b,c)$

$$+E^{T_2}(x)+E^{T_3}(x)\big) \overset{\text{дистр}}{=} \varphi_1 + \min_{x\in T_2}(\varphi_{12}+E^{T_2}) + \min_{x\in T_3}(\varphi_{13}+E^{T_3}) =$$

// $\min_{i,j}(a_i+b_j) = \min_i a_i + \min_j b_j$

$$= \varphi_1(x_1) + \min_{x_2}\Big(\varphi_{12} + \min_{x\in T_2\setminus x_2} E^{T_2}(x)\Big) +$$

$$+ \min_{x_3}\Big(\varphi_{13} + \min_{x\in T_3\setminus x_3} E^{T_3}(x)\Big) = \varphi_1 + \min_{x_2}(\varphi_{12}+V^{T_2}) +$$

$$+ \underbrace{\min_{x_3}(\varphi_{13}+V^{T_3})}_{\text{передача сообщений}} \quad \mu_{3\to1}(x_1) \quad \{\mu_{3\to1}(x_1)\}$$

$V^{T_1}(x_1) = \varphi_1(x_1) + \mu_{2\to1}(x_1) + \mu_{3\to1}(x_1)$

$\mu_{2\to1}(x_1) \stackrel{\Delta}{=} \min\limits_{x_2} (\varphi_{12}(x_1, x_2) + V^T(x_2)) =$

$= \min\limits_{x_2} (\varphi_{12}(x_1, x_2) + \varphi_2(x_2) + \sum\limits_{i \in ch(2)} \mu_{i\to2}(x_2))$

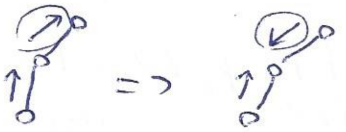
$\mu_{7\to4}(x_4) = \min\limits_{x_7} (\varphi_{47}(x_4, x_7) + \varphi_7(x_7))$

argmin ?

$S_{2\to1} = \operatorname{argmin}\limits_{x_2} (-\,''\,-)$

$x_1^* = \operatorname{argmin}\limits_{x_1} V^{T_1}(x_1)$

$x_2^* = S_{2\to1}(x_1^*)$



$V^{T_2}(x_2)$ переподвесим x_2

$\mu_{1\to2}(x_2) = \min\limits_{x_1} (\varphi_{12}(x_1, x_2) + \varphi_1(x) + \mu_{3\to1}(x_1))$

$\mu_{i\to j}(x_j) = \min\limits_{x_i} (\varphi_{ij}(x_i, x_j) + \varphi_i(x_i) + \sum\limits_{\substack{k=(i,k)\in E \\ k \neq j}} \mu_{k\to i}(x_i))$

Min sum belief propagation

$Z^{T_1}(x_1) \stackrel{\Delta}{=} \sum\limits_{x \setminus x_1} \prod\limits_{i\in V} \psi_i \prod\limits_{(i,j)\in E} \psi_{ij}$

Sum product be BP // $ab + ac = a(b+c)$

// $\sum\limits_{i,j} a_i b_j = \sum\limits_i a_i \sum\limits_j b_j$

$Z^{T_1}(x_1) = \sum\limits_{x\setminus x_1} \psi_1 \psi_{12} \psi_{13} \prod\limits_{i\in T_2} \psi_i \prod\limits_{(i,j)\in T_2} \psi_{ij} \prod\limits_{i\in T_3} \psi_i \prod\limits_{(i,j)\in T_3} \psi_{ij} =$

$= \psi_1 \sum\limits_{\substack{x\in T_2 \\ x \in T_3}} \psi_{12}\psi_{13} \prod\limits_{i\in T_2} \psi_i \prod\limits_{(i,j)\in T_2} \psi_{ij} \prod\limits_{i\in T_3} \psi_i \prod\limits_{(i,j)\in T_3} \psi_{ij} =$

$= \psi_1 \left(\sum\limits_{x\in T_2} \psi_{12} \prod\limits_{\in T_2} \psi_i \prod\limits_{T_2} \psi_{ij} \right) \left(\sum\limits_{x\in T_3} \psi_{13} \prod\limits_{T_3} \psi_i \prod\limits_{T_3} \psi_{ij} \right) =$

$$= \psi_1 \underbrace{\Big(\sum_{x_2} \psi_{12} \overbrace{\psi_2 \sum_{x \in T_2 \setminus x_2} \prod_{T_2} \psi_i \prod_{T_2} \psi_{ij}}^{Z^{T_2}(x_2)} \Big)}_{\mu_{2\to 1}(x_1)}$$

$$\Big(\sum_{x_3} \psi_{13} \psi_3 \sum_{x \in T_3 \setminus x_3} \prod_{T_3} \psi_i \prod_{T_3} \psi_{ij} \Big)$$

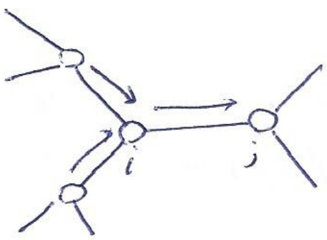
$$Z^{T_1}(x_1) = \psi_1 \, \mu_{2\to 1}(x_1) \mu_{3\to 1}(x_1)$$

$$\mu_{2\to 1}(x_1) = \sum_{x_2} \psi_{12}(x_1, x_2) \cdot Z^{T_2}(x_2) =$$

$$= \sum_{x_2} \Big(\psi_{12}(x_1, x_2) \psi_2(x_2) \prod_{i \in ch(2)} \mu_{i\to 2}(x_2) \Big)$$

$$Z = \sum_{x_1} Z^{T_1}(x_1) \quad , \quad p(x_1) = \frac{Z^{T_1}(x_1)}{Z}$$

$$\mu_{i\to j}(x_j) = \sum_{x_i} \psi_{ij}(x_i, x_j) \psi_i(x_i) \prod_{\substack{k:(i,k)\in\mathcal{E} \\ k \neq j}} \mu_{k\to i}(x_i)$$



$$Z^{T_j}(x_j) = \psi_j(x_j) \prod_{i:(i,j)\in\mathcal{E}} \mu_{i\to j}(x_j)$$

$G = (V, \mathcal{E})$ – ацикличный граф

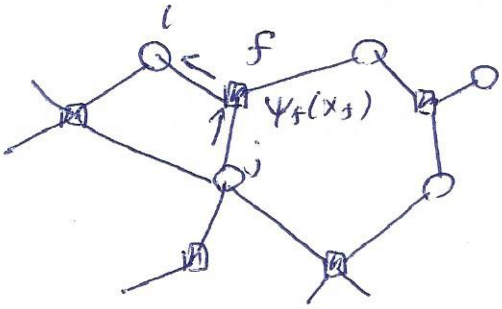
$$\varphi(x_1 \ldots x_n) = \bigotimes_{i \in V} f_i(x_i) \bigotimes_{(i,j)\in\mathcal{E}} f_{ij}(x_i, x_j) \qquad \otimes \quad \oplus \quad \text{полукольцо}$$

$\bigoplus_{x_1 \ldots x_n} \varphi(x_1 \ldots x_n)$ расчёт за линейное время

$\oplus$	min	+	min	max
$\otimes$	+	*	max	min

(mod k)

Loopy Belief Propagation



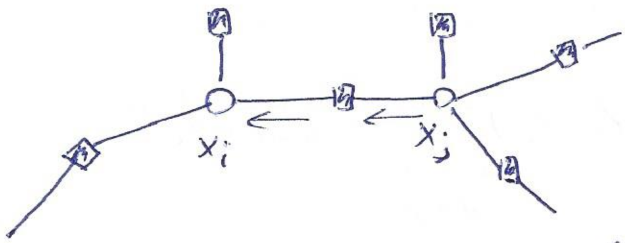
сообщения из вершины в фактор и из фактора в вершину.

Sum product

$$\mu_{j\to f}(x_j) = \prod_{\substack{g\neq f\\(g,j)\in E}} \mu_{g\to j}(x_j)$$

$$\mu_{f\to i}(x_i) = \sum_{x_f\setminus x_i} \psi_f(x_f) \prod_{\substack{j\neq i\\(j,f)\in E}} \mu_{j\to f}(x_j)$$

соответствие между циклическими и ацикличными алгоритмами



$$\mu_{j\to f}(x_j) = \prod_{\substack{g\neq f\\(i,j)\in E}} \mu_{g\to j}(x_j)$$

$$\mu_{f\to i}(x_i) = \sum_{x_f\setminus x_i} \psi_f(x_f) \prod_{\substack{j\neq i\\(j,f)\in E}} \mu_{j\to f}(x_j) =$$

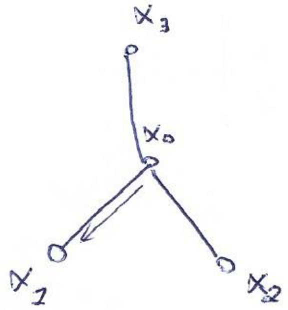
$$= \sum_{x_j} \psi_f(x_i, x_j) \prod_{\substack{j\neq i\\(j,f)\in E}} \mu_{j\to f}(x_j) = \sum_{x_j} \psi_{ij}(x_i, x_j)\, \psi_j(x_j) \prod_{\substack{g\neq f\\(j,g)\in E}} \mu_{g\to j}(x_j)$$

вводится нормировка α

$$b_i(x_i) = \frac{\prod_{f:(i,f)\in E} \mu_{f\to i}(x_i)}{\sum_{x_i} \text{—''—}} \simeq p(x_i)$$

$$b_f(x_f) = \frac{\psi_f(x_f) \prod_{i:(i,f)\in E} \mu_{i\to f}(x_i)}{\sum_{f} \text{—''—}} \simeq p(x_f)$$

17.02.17 2м сем



$$E(x) = \sum (x_i - i)^2 + \sum_{(i,j)\in E} (x_i - x_j)^2$$

$$x_i \in \{1, 2, 3\}$$

$$\mu_{1\to 0}(x_0) = \min_{x_1} \left((x_1 - 1)^2 + (x_1 - x_0)^2 \right) =$$

$$= \begin{cases} 0, & x_0 = 1 \\ 1, & x_0 = 2 \\ 2, & x_0 = 3 \end{cases}$$

$$V_1(x_1) = (x_1 - 1)^2 + \mu_{0\to 1}(x_1)$$

$$\mu_{0\to 1}(x_1) = \min_{x_0} \left(x_0^2 + (x_0 - x_1)^2 + \mu_{2\to 0}(x_0) + \mu_{3\to 0}(x_0) \right) =$$

$$= \cancel{1} + \min_{x_0} \begin{pmatrix} x_0^2 + (x_0 - 1)^2 + \mu_{2\to 0}(x_0) + \mu_{3\to 0}(x_0) \\ x_0^2 + (x_0 - 2)^2 + \mu_{2\to 0}(x_0) + \mu_{3\to 0}(x_0) \\ x_0^2 + (x_0 - 3)^2 + \mu_{2\to 0}(x_0) + \mu_{3\to 0}(x_0) \end{pmatrix}$$

$$\mu_{0\to 1}(1) = \min_{x_0} \left(x_0^2 + (x_0 - 1)^2 + \mu_{2\to 0}(x_0) + \mu_{3\to 0}(x_0) \right) = 4$$

$x_0 = 1$: $1 + 1 + 2 = 4$ $\quad \underline{x_0 = 1}$

$x_0 = 2$: $4 + 1 + 0 + 1 = 6$

$x_0 = 3$: $9 + 4 + 1 + 0 = 14$

$\mu_{0\to 1}(2)$:

$x_0 = 1$: $1 + 1$

$$\mu_{0\to 1}(x_1) = \min \left(\begin{bmatrix} 1 \\ 4 \\ 9 \end{bmatrix} + \begin{bmatrix} (1 - x_1)^2 \\ (2 - x_1)^2 \\ (3 - x_1)^2 \end{bmatrix} + \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} + \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix} \right) =$$

$$= \min \left(\begin{bmatrix} (1 - x_1)^2 + 4 \\ (2 - x_1)^2 + 5 \\ (3 - x_1)^2 + 10 \end{bmatrix} \right) = \{4, 5, 6\}$$

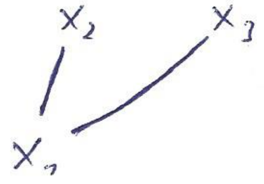
$$V_1(x_1) = (x_1 - 1)^2 + \begin{bmatrix} 4 \\ 5 \\ 6 \end{bmatrix} = \begin{bmatrix} 4 \\ 6 \\ 10 \end{bmatrix}$$

$Ax = b$

$E(x) = \frac{1}{2} x^T A x - b^T x$ $\quad A = \begin{bmatrix} 5 & -2 & 1 \\ -2 & 3 & 0 \\ 1 & 0 & 1 \end{bmatrix}, \quad b = \begin{bmatrix} -6 \\ 0 \\ 2 \end{bmatrix}$

$\varphi_i(x_i) = -b_i x_i + \frac{1}{2} a_{ii} x_i^2$

$\varphi_{ij}(x_i, x_j) = a_{ij} x_i x_j$



$\mu_{2\to1}(x_1) = \min_{x_2} \left(\frac{1}{2} 3x_2^2 - 2x_1x_2 \right) = \cancel{\frac{3}{2} x_2^2 - 3x_2^2 = -\frac{3}{2} x_2^2}$

$\uparrow$

$// \; x_1 = \frac{3}{2} x_2, \quad x_2 = \frac{2}{3} x_1$

$= \frac{3}{2} \left(\frac{2}{3}\right)^2 x_1^2 - \frac{2}{3} \cdot 2x_1^2 = \frac{2}{3} x_1^2 - \frac{4}{3} x_1^2 = -\frac{2}{3} x_1^2$

$\mu_{3\to1}(x_1) = \min_{x_3} \left(-2x_3 + \frac{1}{2} x_3^2 + x_1 x_3 \right) =$

$\cancel{\left(-\frac{3}{2} x_3^2 + x_1 x_3 \right)'_{x_3} = -\frac{3}{2} + x_1 = 0, \quad x_1 = \frac{3}{2}}$

$\frac{\partial}{\partial x_3} : \; -2 + x_3 + x_1 = 0, \quad x_3 = -x_1 + 2 = 2 - x_1$

$= 2(x_1 - 2) + \frac{1}{2}(x_1 - 2)^2 + x_1(2 - x_1) = (x_1 - 2)\left(2 + \frac{1}{2}(x_1 - 2) - x_1\right) =$

$= -\frac{1}{2}(x_1 - 2)^2$

$V(x_1) = \varphi_1(x_1) + \sum_{j \in N(i)} \mu_{j\to i}(x_i) = \varphi(x_1) + \mu_{2\to1}(x_1) + \mu_{3\to1}(x_1) =$

$= 6x_1 + \frac{5}{2} x_1^2 - \frac{2}{3} x_1^2 - \frac{1}{2}(x_1 - 2)^2 = 8x_1 + \frac{4}{3} x_1^2 - 2 = \frac{4}{3} x_1^2 + 8x_1 - 2$

$\frac{d}{dx_1} : \; -\frac{b}{2a} = -\frac{8}{\frac{8}{3}} = -3, \quad \underline{x_1 = -3} = \arg\min_{x_1} V(x_1)$

$x_3 = 2 + 3 = 5$

$x_2 = \frac{2}{3} x_1 = -2$

$\begin{bmatrix} -3 \\ -2 \\ 5 \end{bmatrix}$