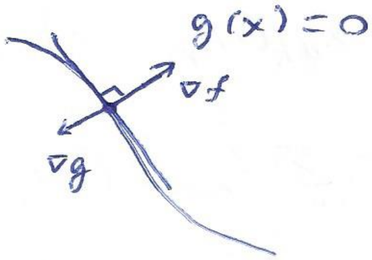
10.03.17 лк IV

$$\begin{cases} f(x) \to extr \\ g(x) = 0 \end{cases}$$

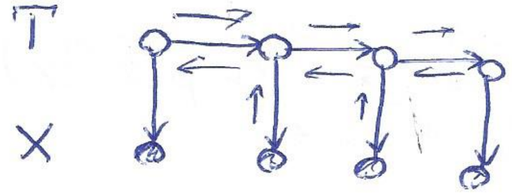


$$\mathcal{L}(x, \lambda) = f(x) + \lambda g(x)$$

$$\nabla f(x) = -\lambda \nabla g(x)$$

$$\frac{\partial \mathcal{L}}{\partial x} = 0, \quad \frac{\partial \mathcal{L}}{\partial \lambda} = 0$$

HMM {Hidden Marcov Models}



байесовская сеть

$$x_n \in \mathbb{R}^d, \quad t_n \in \{0, 1\}^K$$

$$\sum_{k=1}^{K} t_{nk} = 1$$

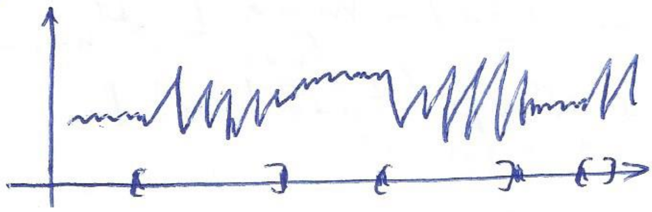
$$p(X, T) = p(t_1) \prod_{n=2}^{N} p(t_n | t_{n-1}) \prod_{n=1}^{N} p(x_n | t_n)$$

$$p(t_{1k} = 1) = \pi_k \qquad \sum_k \pi_k = 1, \ \pi_k \geq 0$$

$$p(t_{nl} = 1 | t_{n-1,k} = 1) = A_{kl}, \quad \sum_{l=1}^{K} A_{kl} = 1, \ A_{kl} \geq 0$$

$$p(x_n | t_{nk} = 1) = p(x_n | \varphi_k)$$

$$\Theta = \{\pi, A, \{\varphi_k\}\}$$



$A_{kk} \gg 0$

диагональное преобладание

1) $p(X_{tr}, T_{tr} | \Theta) \to \max\limits_{\Theta}$ обучение с учителем

2) $p(T | X, \Theta) \to \max\limits_{T}$ прогнозирование

3) $p(X_{tr} | \Theta) \to \max\limits_{\Theta}$ обучение без учителя

$$p(X, T) = \underbrace{\prod_{k=1}^{K} \pi_k^{t_{1k}}}_{p(t_1)} \underbrace{\prod_{n=2}^{N} \prod_{k=1}^{K} \prod_{l=1}^{K} A_{kl}^{t_{n-1,k} t_{nl}}}_{p(t_n | t_{n-1})} \prod_{n=1}^{N} \prod_{k=1}^{K} p(x_n | \varphi_k)^{t_{nk}}$$

1) $\ln p(X_{tr}, T_{tr} | \Theta) = \sum_k t_{1k} \ln \pi_k + \sum_{n,k,l} t_{n-1,k} \cdot t_{nl} \ln A_{kl} +$

$$+ \sum_{n,k} t_{nk} \ln p(x_n | \varphi_k) \to \max_{\pi, A, \varphi}$$

$$\sum_k \pi_k = 1, \quad \sum_{l} A_{kl} = 1$$

$$\mathcal{L}(\theta, \lambda) = \ln p(X_{t2}, T_{t2} \mid \theta) + \lambda_0 \left(\sum_k \pi_k - 1 \right) +$$

$$+ \sum_{k,l} \lambda_{k}(A_{kl} - 1) \to \text{extr}$$

$$\frac{\partial \mathcal{L}}{\partial A_{kl}} = \sum_n t_{n-1,l} \cdot t_{nl} \frac{1}{A_{kl}} + \lambda_k = 0$$

$$\frac{\partial \mathcal{L}}{\partial \lambda_k} = \sum_l A_{kl} - 1 = 0$$

BP - max sum

$$A_{kl} = - \frac{\sum_{n=2}^{N} t_{n-1,k} t_{nl}}{\lambda_k} \quad \Big| \sum_l \cdot = 1$$

$$A_{kl} = \frac{\sum_{n=2}^{N} t_{n-1,k} t_{nl}}{\sum_{n=1}^{N} t_{n-1,k}}, \quad \varphi_k = \arg\max_{\varphi} \sum_{n: t_{nk}=1} \ln p(x_n \mid \varphi)$$

2) $\arg\max_T p(T \mid X, \theta) = \arg\max_T p(X, T \mid \theta) = \arg\max_T \ln p(X, T \mid \theta)$

$$\mu_{n \to n+1}(t_{n+1}) = \max_{t_n} [\mu_{n-1 \to n}(t_n) + \ln p(x_n \mid t_n) + \ln p(t_{n+1} \mid t_n)]$$

$$\mu_{n+1 \to n}(t_n) = \max_{t_{n+1}} [\mu_{n+2 \to n+1}(t_{n+1}) + \ln p(x_{n+1} \mid t_{n+1}) + \ln p(t_{n+1} \mid t_n)]$$

$$t_n^* = \arg\max_{t_n} (\mu_{n-1 \to n}(t_n) + \mu_{n+1 \to n}(t_n) + \ln p(x_n \mid t_n))$$

3) $$\ln p(X \mid \theta) = \underbrace{\int q(T) \ln \frac{p(X, T \mid \theta)}{q(T)} dT}_{\mathcal{L}(q, \theta)} + \int q(T) \ln \frac{q(T)}{p(T \mid X, \theta)} dT$$

$$\ln p(X \mid \theta) \to \max_{\theta} \Leftrightarrow \mathcal{L}(q, \theta) \to \max_{q, \theta}$$

E-step: $\mathcal{L}(q, \theta) \to \max_q \Leftrightarrow q(T) = p(T \mid X, \theta_0)$

M-step: $\mathcal{L}(q_0, \theta) \to \max_{\theta} \Leftrightarrow \mathbb{E}_{q(T)} \ln p(X, T \mid \theta) \to \max_{\theta}$

$\cancel{\mathbb{E}_{q(T)}}$ $\cancel{\mathbb{E}_T}$ | M max

$$\mathbb{E}_T \ln p(X,T|\Theta) = \mathbb{E}_T \Big(\sum_{k=1}^{K} t_{1k} \ln \pi_{k} + \sum_{n,k,l} t_{n-1,k} t_{nl} \ln A_{kl} + \sum_{n,k} t_{nk} \ln p(x_n|\varphi_k) \Big) = \sum_{k=1}^{K} \mathbb{E} t_{1k} \ln \pi_k +$$

$$+ \sum_{n,k,l} \mathbb{E}\, t_{n-1,k} t_{nl} \cdot \ln A_{kl} + \sum_{n,k} \ln p(x_n|\varphi_k) \cdot \mathbb{E}\, t_{nk}$$

$$\xi \sim Be(p),\ \mathbb{E}\xi = p$$

$$\mathbb{E}\, t_{1k} = \mathbb{P}\{t_{1k} = 1\},\quad \cancel{\mathbb{E}_{t_{nk}}}\ \mathbb{E}\, t_{nk} = \mathbb{P}\{t_{nk} = 1\}$$

$$\mathbb{E}\, t_{n-1,k} t_{nl} = \mathbb{P}\{t_{n-1,k} = 1, t_{nl} = 1\} \neq \mathbb{P}(t_{n-1,k} = 1)\mathbb{P}(t_{nl} = 1)$$

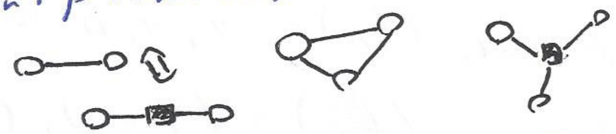
$$p(t_n),\ p(t_{n-1}, t_n)$$

sum-prod BP

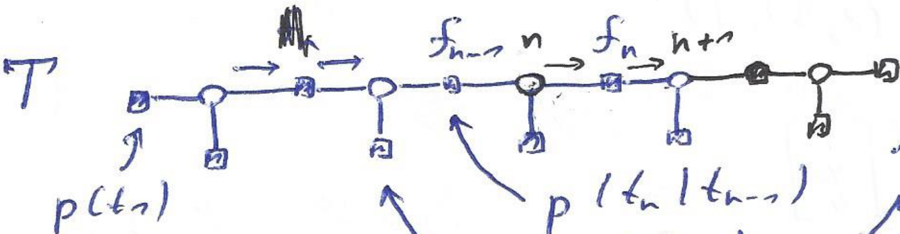
$$\mu_{n\to n+1}(t_{n+1}) = \sum_{t_n} \mu_{n-1\to n}(t_n)\, p(x_n|t_n)\, p(t_{n+1}|t_n)$$

$$\mu_{n+1\to n}(t_n) = \sum_{t_{n+1}} \mu_{n+2\to n+1}(t_{n+1})\, p(x_{n+1}|t_{n+1})\, p(t_{n+1}|t_n)$$

$$p(t_n) \propto \mu_{n-1\to n}(t_n)\, \mu_{n+1\to n}(t_n)\, p(x_n|t_n)$$



T



$$\mu_{\uparrow}(t_n) = p(x_n|t_n)$$

$$\mu_{n\to f_n}(t_n) = \mu_{f_{n-1}\to n}(t_n)\, \mu_{\uparrow}(t_n)$$

$$\mu_{f_n\to n+1}(t_{n+1}) = \sum_{t_n} p(t_{n+1}|t_n)\, \mu_{n\to f_n}(t_n)$$

$$p(t_n, t_{n+1}) \propto p(t_{n+1}|t_n)\, \cancel{\varphi}\, \mu_{n\to f}(t_n)\, \mu_{n+1\to f}(t_{n+1})$$

$$\mathbb{E} \ln p(X,T|\Theta) = \sum_k \mathbb{E}\, t_{1k} \ln \pi_k + \sum_{k,n,l} \mathbb{E}\, t_{n-1,k} t_{nl} \ln A_{kl} +$$

$$+ \sum_{n,k} \mathbb{E}\, t_{nk} \ln p(x_n|\varphi_k) \to \max_{\pi, A, \varphi}$$

$$\sum \pi_k = 1,\ \sum_l A_{kl} = 1$$

$$A_{kl} = \frac{\sum_n \mathbb{E}\, t_{n-1,k} t_{nl}}{\sum_{n=2} \mathbb{E}\, t_{n-1,k}} \qquad \varphi_k = \arg\max_{\varphi} \sum_{n=1}^{N} \mathbb{E}\, t_{nk} \ln p(x_n|\varphi)$$

10.03.17 ГМ сем

$\varphi_{ij}(t_i, t_j) = [t_i \neq t_j]$

φ_i	1	2	3	4	5
1	1	3	1	1	1
2	1	1	0	1	1
3	0	2	2	1	0

$$t_1 \to t_2 \to t_3 \to t_4 \to t_5$$

$$\mu_{i\to j}(x_j) = \min_{x_i}\Big(\varphi_i(x_i) + \varphi_{ij}(x_i, x_j) + \sum_{k\in N(i)\setminus j} \mu_{k\to i}(x_i)\Big)$$

$$V_i(x_i) = \varphi_i(x_i) + \sum_{j\in N(i)} \mu_{j\to i}(x_i)$$

$$\mu_{t_1\to t_2} = \min_{t_1}\left(\underbrace{\begin{bmatrix}1\\1\\0\end{bmatrix}}_{\varphi_1(x_1)} + \underbrace{\begin{bmatrix}0\\0\\0\end{bmatrix}}_{\Sigma\mu} + \begin{bmatrix}0&1&1\\1&0&1\\1&1&0\end{bmatrix}\Big\downarrow t_1\right) =$$

$$\xrightarrow[t_2]{} \varphi_{ij}(x_i, x_j)$$

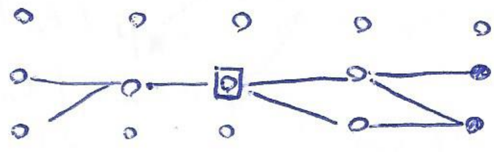
$$= \min_{t_1}\begin{pmatrix}1&2&2\\2&1&2\\1&1&0\end{pmatrix} = (1\,1\,0)$$

$\mu_{t_1\to t_2}(t_2) = 1, 1, 0$

$S_{t_1\to t_2}(t_1) = ((1,3), (2,3), 3)$

$$V_5(t_5) = \begin{bmatrix}1\\1\\0\end{bmatrix} + \begin{bmatrix}4\\3\\4\end{bmatrix} = \begin{bmatrix}5\\4\\4\end{bmatrix}$$

S

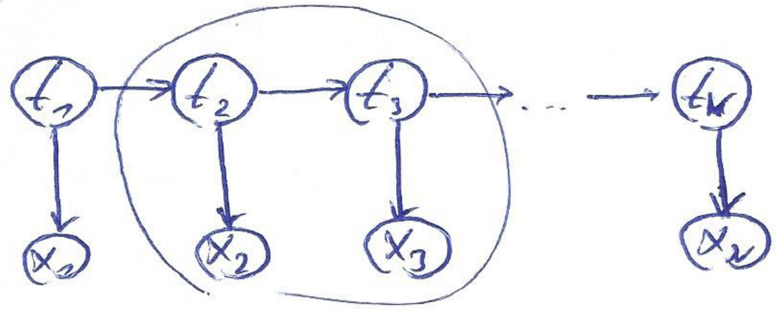


решётка

~~$V_4(t_5)$~~

$$\min_T p(t_1 \dots t_k \dots t_N) = \min_T p(t_1 \dots t_k)\, p(t_{k+1} \dots t_N \mid t_k) =$$

$$= \min_{t_1 \dots t_k} p(t_1 \dots t_k) \min_{t_{k+1} \dots t_N} p(t_{k+1} \dots t_N \mid t_k^*)$$



$p(t_i \mid x)$

$p(t_{i+1}, t_i \mid x)$

1) $\sum_y p(x,y) = p(x)$

2) $p(x \mid y) = \dfrac{p(x,y)}{p(y)}$

3) $f(x) = \prod_{i \in V} \varphi_i(x_i) \prod_{(i,j) \in E} \varphi_{ij}(x_i, x_j)$

4) $V_i(x_i) = \sum_{x \setminus i} f(x) = \varphi_i(x_i) \prod_{j \in N(i)} \mu_{j \to i}(x_i)$

$$p(x,T) = p(t_1) \cdot p(x_1 \mid t_1) \prod_{i=2}^{N} p(t_i \mid t_{i-1})\, p(x_i \mid t_i)$$

$$p(t_i \mid x) = \sum_{t \setminus i} p(T \mid x) = \sum_{t \setminus i} \frac{p(x,T)}{p(x)} = \sum_{t \setminus i} \frac{p(x,T)}{\sum_t p(x,T)} \propto$$

$$\propto \sum_{t \setminus i} p(x,T)$$

$$p(t_i \mid x) = \mu_{t_{i-1} \to t_i}(t_i) \cdot \mu_{t_{i+1} \to t_i}(t_i) \cdot p(x_i \mid t_i) p(t_i)$$

$$p(t_i, t_{i+1} \mid x)$$

$$A_{ii} = 1 - \alpha \quad , \quad \alpha \in [0,1] \quad , \quad \alpha = \partial(p)$$

$$A_{ij} = \frac{\alpha}{k-1} \quad , \quad i \neq j$$

$$p(x,T) = \prod_k \pi_k^{z_{1k}} \prod_{n,k,l} A_{kl}^{z_{nk} \cdot z_{n-1,l}} \prod_{n,k} p(x_n \mid \varphi_k)^{z_{nk}}$$

$$\ln p(x,T) = \sum_k z_{1k} \ln \pi_k + \sum_{n,k,l} z_{nk} \cdot z_{n-1,l} \ln A_{kl} + C$$