Class note

Thursday, February 8, 2018 9:09 AM

$D = \{\vec{x}^1, \vec{x}^2, \dots, \vec{x}^N\}$ $\quad \vec{x}_i \overset{iid}{\sim} P(\vec{x};\theta) = \frac{1}{Z(\theta)} \prod_c \phi_c(x_c;\theta) =$

$\dim(\vec{x}_i) > 1$

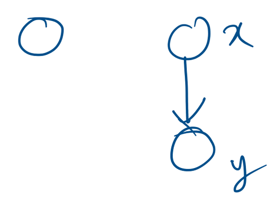
$\max_\theta \log p(D;\theta) = \max_\theta \sum_{n=1}^N \log P(\vec{x}^n;\theta)$

param

$= \max_\theta \sum_{n=1}^N \sum_c \log \phi_c(x_c^n;\theta) - N\log Z(\theta)$

$L(\theta) = \sum_{n=1}^N \sum_c \log \phi_c(x_c^n;\theta) - N\log\left(\sum_y \prod_c \phi_c(y_c;\theta)\right)$

$p(x;\theta) = \exp(\theta^T T(x) - A(\theta))h(x)$



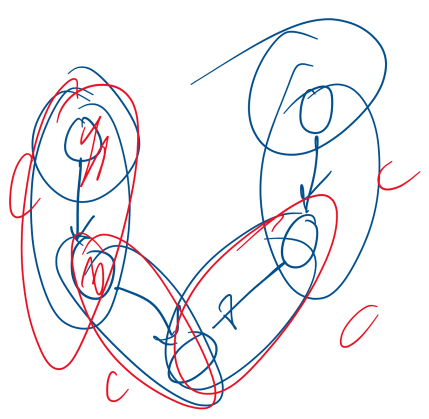
$p(y|x)$

$g^{-1}(\mu) = w^T x$

$x \xrightarrow{w} \eta \rightleftarrows \mu \underset{A^*}{\overset{(A')^{-1}}{\rightleftarrows}} \theta$

S+ $\quad$ natu

$\eta = w^T x$



$\sum_{n=1}^N \log \phi_c(x_c^n;\theta_c) \quad - N\mathbb{E}\left[\frac{\partial}{\partial\theta_c} \log(y_c;\theta_c)\right]$

$\nabla_{\theta_c} L(\theta) = \sum_{n=1}^N \frac{\partial}{\partial\theta_c}\phi_c(x_c^n;\theta_c)\frac{1}{\phi_c(x_c^n;\theta)} - \frac{\nabla_{\theta_c} Z(\theta)}{Z(\theta)}$

$\sum_{n=1}^N \log \phi_c(x_c^n;\theta) - N\mathbb{E}\left[\frac{\partial}{\partial\theta_c}\log\phi_c(y_c;\theta)\right]$

N x empirical $\qquad$ N x real mean

$\nabla L(\theta) = 0$

$\frac{\nabla Z(\theta)}{Z(\theta)} = \nabla \log Z(\theta) = \frac{\prod_{c'\neq c}\phi_{c'}(y_{c'};\theta_{c'}) \frac{\partial\phi(y_c;\theta_c)}{\partial\theta_c}}{\sum_y \prod_{c'} \phi_{c'}(y_{c'};\theta_{c'})}$

$\sum_y \frac{\partial\phi(y_c;\theta_c)}{\partial\theta_c} \cdot \frac{\phi_c(y_c;\theta_c)\prod_{c'\neq c}\phi_{c'}(y_{c'};\theta_{c'})}{\sum_y \prod_{c'}\phi_{c'}(y_{c'};\theta_{c'}) \; \phi_c(y_c;\theta_c)}$

$\sum_y \frac{\frac{\partial\phi(y_c;\theta_c)}{\partial\theta_c}}{\phi_c(y_c;\theta_c)} P(y;\theta)$

$\frac{\partial}{\partial\theta_c}\log\phi(y_c;\theta_c)$

$\sum_y \frac{\partial}{\partial\theta_c}\log\phi_c(y_c;\theta_c)\, P(y;\theta)$

$\mathbb{E}\left[\frac{\partial}{\partial\theta_c}\log\phi_c(y_c;\theta_c)\right]$

$L(\phi) = \sum_n \sum_c \sum_{y_c} \mathbb{1}(y_c = x_c^n)\log\phi_c(y_c) - N\log Z(\phi)$

$\nabla_{\phi_c} L(\phi) = \sum_{n=1}^N \sum_{y_c} \mathbb{1}(y_c = x_c^n)\frac{1}{\phi_c(y_c)} - N\frac{\nabla_{\phi_c} Z(\phi)}{Z(\phi)}$

$\frac{\sum_{y_c}\prod_{c'\neq c}\phi_{c'}(y_{c'})\cdot\phi_c(y_c)}{\sum_{y_{c'}}\prod_{c'}\phi_{c'}(y_{c'}) \; \phi(y_c)}$

empirical $\quad \mathcal{E}(x_c^n) = \frac{1}{N}\sum \mathbb{1}(y_c = x_c^n)$

$P(y_c)$

$= N\frac{\mathcal{E}(x_c^n)}{\phi_c(y_c)} - N\frac{P(y_c)}{\phi_c(y_c)}$

$\sum_y \frac{1}{\phi(y_c)} \cdot P(y_c)$

$\mathbb{E}\left[\frac{1}{\phi(y_c)}\right]$

$\nabla L(\phi_c) = 0$

$\mathcal{E}(x_c^n) = P(x_c^n)$

Condit.

$\log P(x;\theta) = \sum_c \theta_c^T f(x_c) - \log(Z_\theta)$

$\nabla_{\theta_c}(\cdot) = f(x_c) - \frac{\nabla_{\theta_c} Z(\theta)}{Z(\theta)}$ $\qquad Z(\theta) = e^{A(\theta)}$

$\nabla_{\theta_c} Z(\theta) = \frac{\partial A}{\partial\theta_c} e^{A(\theta)}$

$\mathbb{E}[f_c(x_c)]$

$f(x_c) - \mathbb{E}[f_c(x_c)]$

$\mathcal{L}(D;\theta) = \frac{1}{N}\sum_{n=1}^N f_c(x_c) - \mathbb{E}[f_c(x_c)]$

$\mathcal{L}(D;\theta) + Reg(\theta)$

$\|\theta\|_2$

$\|\theta\|_1$