BP + Learning MRFs and CRFs

Wednesday, September 28, 2022 11:30 AM

MRF w/ feature based parametrization

Parameterization: $\psi_c(y_c;\theta) = \exp\left(\vec{\theta} \cdot \vec{f}_c(y_c)\right) = \exp\left(\sum_{m=1}^{M} \theta_m f_{c,m}(y_c)\right)$

Data (MRF): $D = \{\vec{y}^{(i)}\}_{i=1}^{N}$

Model (MRF): $p(y^{(i)}|\theta) = \frac{1}{Z}\prod_c \psi_c(y_c^{(i)};\theta)$

Log-likelihood (MRF): $\ell_D(\theta) = \frac{1}{N}\sum_{i=1}^{N} \log p(\vec{y}^{(i)}|\vec{\theta}) = \frac{1}{N}\sum_{i=1}^{N}\left(\sum_c \log(\psi_c(\vec{y}_c^{(i)})) - \log(Z)\right)$

Derivatives (MRF):

$$\frac{\partial \ell_D(\theta)}{\partial \theta_k} = \underbrace{\frac{1}{N}\left(\sum_{i=1}^{N}\sum_c f_{c,k}(\vec{y}_c^{(i)})\right)}_{\mathbb{E}_{\vec{y}\sim p(\cdot|D)}[f_{\bullet,k}(\vec{y})]} - \underbrace{\frac{1}{N}\left(\sum_{i=1}^{N}\sum_c\sum_{\vec{y}_c'} \overbrace{p_\theta(\vec{y}_c')}^{\text{factor marginal obtained by running BP}} f_{c,k}(\vec{y}_c')\right)}_{\mathbb{E}_{\vec{y}\sim p(\cdot|\theta)}[f_{\bullet,k}(\vec{y})]}$$

where $f_{\bullet,k}(\vec{y}) = \sum_c f_{c,k}(\vec{y}_c)$