16/11/09

PMR: DERIVATIVE OF THE LOG LIKELIHOOD OF A LOG-LINEAR MODEL

$$\log p(x) = \log \left\{ \sum_y \exp \left[\sum_k \theta_k \phi_k(x,y) \right] \right\} - \log Z$$

where $$Z = \sum_{x,y} \exp \left[\sum_k \theta_k \phi_k(x,y) \right]$$

$$\frac{\partial \log p(x)}{\partial \theta_\ell} = \frac{\sum_y \phi_\ell(x,y) \exp \left[\sum_k \theta_k \phi_k(x,y) \right]}{\sum_{y'} \exp \left[\sum_k \theta_k \phi_k(x,y') \right]}$$

$$- \frac{\sum_{x,y} \phi_\ell(x,y) \exp \left[\sum_k \theta_k \phi_k(x,y) \right]}{\sum_{x',y'} \exp \left[\sum_k \theta_k \phi_k(x',y') \right]}$$

$$= \sum_y \phi_\ell(x,y) \, p(y|x,\theta) - \sum_{x,y} \phi_\ell(x,y) \, p(x,y|\theta)$$

$$= \langle \phi_\ell(x,y) \rangle^+ - \langle \phi_\ell(x,y) \rangle^-$$