Max Likelihood Estimation for CRFs

Parameter Estimation

$$P_\theta(Y|x) = \frac{1}{Z_x(\theta)} \tilde{P}_\theta(x, Y)$$

$$Z_x(\theta) = \sum_Y \tilde{P}_\theta(x, Y)$$

$D = \{(x[m], y[m])\}_{m=1}^{M}$ — M instances

Data is now a set of pairs

log conditional likelihood

$$\ell_{Y|X}(\theta : D) = \sum_{m=1}^{M} \ln P_\theta(y[m] | x[m], \theta)$$

$$\ell_{Y|X}(\theta : (x[m], y[m])) = \left(\sum_i \theta_i f_i(x[m], y[m]) \right) - \ln Z_{x[m]}(\theta)$$

f_i: feature of that instance

θ_i: parameter weight w_i