Conditional Random Fields

July 16, 2018

1 Introduction

Let $x = \{x_1, x_2, ..., x_n\}$ be a sequence of tokens and $y = \{y_1, y_2, ..., y_n\}$ a sequence of labels attributed to these tokens. Let the conditional probability of y be defined by:

$$P(y|x) = \frac{1}{Z(x)} \prod_{t=1}^{T} exp\left\{\sum_{k=1}^{K} \theta_k f_k(y_t, y_{t-1}, x_t)\right\} \tag{1}$$

The partition function is given by:

$$Z(x) = \sum_{y'} \prod_{t=1}^{T} exp\left\{\sum_{k=1}^{K} \theta_k f_k(y'_t, y'_{t-1}, x_t)\right\} \tag{2}$$

$P(y|x;\theta)$ is the conditional likelihood $\mathcal{L}(\theta)$. And the log conditional likelihood is given by:

$$\begin{aligned}\ell(\theta) &= log \prod_{t=1}^{T} exp\left\{\sum_{k=1}^{K} \theta_k f_k(y_t, y_{t-1}, x_t)\right\} - logZ(x) \\ &= \sum_{t=1}^{T} \sum_{k=1}^{K} \theta_k f_k(y_t, y_{t-1}, x_t) - logZ(x)\end{aligned} \tag{3}$$

Now we want to obtain the feature parameters θ by maximum likelihood. For that, we need to find the derivative of $ell(\theta)$ relative to each parameter.

$$\frac{\partial \ell(\theta)}{\partial \theta_k} = \sum_{t=1}^{T} \sum_{k=1}^{K} f_k(y_t, y_{t-1}, x_t) - \frac{\partial logZ(x)}{\partial \theta_k} \tag{4}$$

$$\frac{\partial logZ(x)}{\partial \theta_k} = \frac{1}{Z(x)} \frac{\partial Z(x)}{\partial \theta_k} \tag{5}$$

$$\frac{\partial Z(x)}{\partial \theta_k} = \sum_{y'} \frac{\partial}{\partial \theta_k} \prod_{t=1}^{T} exp\left\{\sum_{k=1}^{K} \theta_k f_k(y'_t, y'_{t-1}, x_t)\right\} \tag{6}$$

$$\frac{\partial}{\partial\theta_k}\prod_{t=1}^{T} exp\left\{\sum_{k=1}^{K}\theta_k f_k(y'_t, y'_{t-1}, x_t)\right\} = \frac{\partial}{\partial\theta_k} exp\left\{\sum_{t=1}^{T}\sum_{k=1}^{K}\theta_k f_k(y'_t, y'_{t-1}, x_t)\right\} \tag{7}$$

$$\frac{\partial}{\partial\theta_k} exp\left\{\sum_{t=1}^{T}\sum_{k=1}^{K}\theta_k f_k(y'_t, y'_{t-1}, x_t)\right\} = exp\left\{\sum_{t=1}^{T}\sum_{k=1}^{K}\theta_k f_k(y'_t, y'_{t-1}, x_t)\right\}\sum_{t=1}^{T} f_k(y'_t, y'_{t-1}, x_t) \tag{8}$$

$$\frac{\partial log Z(x)}{\partial\theta_k} = \frac{1}{Z(x)}\sum_{y'}\prod_{t=1}^{T} exp\left\{\sum_{k=1}^{K}\theta_k f_k(y'_t, y'_{t-1}, x_t)\right\}\sum_{t=1}^{T} f_k(y'_t, y'_{t-1}, x_t) \tag{9}$$

$$\frac{\partial log Z(x)}{\partial\theta_k} = \sum_{y'}\frac{\prod_{t=1}^{T} exp\left\{\sum_{k=1}^{K}\theta_k f_k(y'_t, y'_{t-1}, x_t)\right\}}{\sum_{y''}\prod_{t=1}^{T} exp\left\{\sum_{k=1}^{K} f_k(y''_t, y''_{t-1}, x_t)\right\}}\sum_{t=1}^{T}\theta_k f_k(y'_t, y'_{t-1}, x_t) \tag{10}$$

$$\frac{\partial log Z(x)}{\partial\theta_k} = \sum_{y'} P(y'|x)\sum_{t=1}^{T} f_k(y'_t, y'_{t-1}, x_t) \tag{11}$$

$$\frac{\partial log Z(x)}{\partial\theta_k} = \sum_{t=1}^{T}\sum_{y'} f_k(y'_t, y'_{t-1}, x_t) P(y'|x) \tag{12}$$

$$\frac{\partial log Z(x)}{\partial\theta_k} = \sum_{t=1}^{T}\sum_{y,y'} f_k(y, y', x_t) P(y, y'|x) \tag{13}$$

Essa parte demanda explicação!!!!