BatchInhib

Rui Lu

Bigeye

Date: Jul. 4 2016

1 Inhibition via Mini-Batch Statistics

在一个 $mini - batch$ 中，考虑某个 $feature\ map$ $\{x_i^{(j)}\}(i = 1, ..., N$ 表示 $feature$ 的维度，$j = 1, ..., m$ 表示 $batch\ size)$，相互抑制的作用可以表示如下：

$$\mu_{x_i} = \frac{1}{m}\sum_{j=1}^{m} x_i^{(j)} \tag{1.1}$$

$$\sigma_{x_i} = \frac{1}{m}\sum_{j=1}^{m} (x_i^{(j)} - \mu_{x_i})^2 \tag{1.2}$$

$$r_{ik} = \frac{\frac{1}{m}[\sum_{j=1}^{m}(x_i^{(j)} - \mu_{x_i})(x_k^{(j)} - \mu_{x_k})]}{\sigma_{x_i}\sigma_{x_k}} \tag{1.3}$$

$$R_{ik} = \frac{g(\alpha_{ik})r_{ik}}{\sum_{l\neq i} g(\alpha_{il})r_{il}} \tag{1.4}$$

$$y_i^{(j)} = x_i^{(j)} - \sum_{l\neq i} R_{il}x_l^{(j)} \tag{1.5}$$

其中，$g(x)=\frac{1}{\rho}log(1+e^{\rho x})$

对上面的式子求导:

$$\frac{\partial L}{\partial R_{ik}}=\sum_{j=1}^{m}\frac{\partial L}{\partial y_i^{(j)}}\frac{\partial y_i^{(j)}}{\partial R_{ik}}=\sum_{j=1}^{m}\frac{\partial L}{\partial y_i^{(j)}}(-x_k^{(j)}) \tag{1.6}$$

$$\frac{\partial L}{\partial \alpha_{ik}}=\frac{\partial L}{\partial R_{ik}}\frac{\partial R_{ik}}{\partial \alpha_{ik}}=\frac{\partial L}{\partial R_{ik}}\frac{r_{ik}[\sum_{l\neq i}g(\alpha_{il})r_{il}-g(\alpha_{ik})r_{ik}]}{[\sum_{l\neq i}g(\alpha_{il})r_{il}]^2}g'(\alpha_{ik}) \tag{1.7}$$

$$\frac{\partial L}{\partial r_{ik}}=\frac{\partial L}{\partial R_{ik}}\frac{\partial R_{ik}}{\partial r_{ik}}=\frac{g(\alpha_{ik})[\sum_{l\neq i}g(\alpha_{il})r_{il}-g(\alpha_{ik})r_{ik}]}{[\sum_{l\neq i}g(\alpha_{il})r_{il}]^2} \tag{1.8}$$

$$\frac{\partial L}{\partial \sigma_{x_i}}=\frac{\partial L}{\partial r_{ik}}\frac{\partial r_{ik}}{\partial \sigma_{x_i}}=\frac{(-1)}{\sigma_{x_i}^2}\frac{\frac{1}{m}[\sum_{j=1}^{m}(x_i^{(j)}-\mu_{x_i})(x_k^{(j)}-\mu_{x_k})]}{\sigma_{x_k}} \tag{1.9}$$

$$\begin{aligned}\frac{\partial L}{\partial \mu_{x_i}}&=\frac{\partial L}{\partial r_{ik}}\frac{\partial r_{ik}}{\partial \mu_{x_i}}+\frac{\partial L}{\partial \sigma_{x_i}}\frac{\partial \sigma_{x_i}}{\partial \mu_{x_i}}=\\ &\frac{\partial L}{\partial r_{ik}}\frac{\frac{1}{m}[\sum_{j=1}^{m}(-1)(x_k^{(j)}-\mu_{x_k})]}{\sigma_{x_i}\sigma_{x_k}}+\frac{\partial L}{\partial \sigma_{x_i}}\frac{(-1)}{2m}\sum_{j=1}^{m}(x_i^{(j)}-\mu_{x_i})\end{aligned} \tag{1.10}$$

$$\begin{aligned}\frac{\partial L}{\partial x_i^{(j)}}&=\frac{\partial L}{\partial \mu_{x_i}}\frac{\partial \mu_{x_i}}{\partial x_i^{(j)}}+\frac{\partial L}{\partial \sigma_{x_i}}\frac{\partial \sigma_{x_i}}{\partial x_i^{(j)}}+\frac{\partial L}{\partial r_{ik}}\frac{\partial r_{ik}}{\partial x_i^{(j)}}+\frac{\partial L}{\partial y_i^{(j)}}\frac{\partial y_i^{(j)}}{\partial x_i^{(j)}}=\\ &\frac{\partial L}{\partial \mu_{x_i}}\frac{1}{m}+\frac{\partial L}{\partial \sigma_{x_i}}\frac{2}{m}(x_i^{(j)}-\mu_{x_i})+\frac{\partial L}{\partial r_{ik}}\frac{1}{m}\frac{(x_k^{(j)}-\mu_{x_k})}{\sigma_{x_i}\sigma_{x_k}}+\frac{\partial L}{\partial y_i^{(j)}}\end{aligned} \tag{1.11}$$