Gradient of Multi-class SVM Loss Function

Given:

- W is weight with a shape of D × C.

- β_{nj} is a binary matrix to indicate the sign of margins.

$$f(x_n, W)_c = x_{nd}W_{dc}$$

$$L_n = \sum_{j \neq y_n} \max(0, x_{nd}W_{dj} - x_{nd}W_{d,y_n} + \Delta)$$

$$L = \frac{1}{N}\sum_n^N \sum_{j \neq y_n} \max(0, x_{nd}W_{dj} - x_{nd}W_{d,y_n} + \Delta) + \lambda \sum_{c=1}^{C}\sum_d W_{c,d}^2$$

The gradient of the loss function with respect to the weight matrix W is:

$$\beta_{nj} = \begin{cases} 1 & \text{if } (x_{nd}W_{dj} - x_{nd}W_{d,y_n} + \Delta) > 0 \\ 0 & \text{otherwise} \end{cases}$$

Thus,

$$L_n = \sum_{j \neq y_n} \beta_{nj}(x_{nd}W_{dj} - x_{nd}W_{d,y_n} + \Delta)$$

$$\frac{\partial L_n}{\partial W_{d,y_n}} = -\sum_{j \neq y_n} \beta_{nj}x_{nd}$$

$$\frac{\partial L_n}{\partial W_{dc}} = \beta_{nc}x_{nd}, \text{ when } c \neq y_n$$

For all samples:

$$\frac{\partial L}{\partial W_{d,y_n}} = -\sum_{n=1}^{N}\sum_{j \neq y_n} \beta_{nj}x_{nd}$$

$$\frac{\partial L}{\partial W_{dc}} = \sum_{n=1}^{N} \beta_{nc}x_{nd}, \text{ when } c \neq y_n$$

The final gradient of the loss function with respect to W is:

$$\frac{\partial L}{\partial W_{cd}} = \frac{1}{N}\sum_{n=1}^{N}\left(\beta_{nc} - \sum_{k \neq y_n}\beta_{nk}\right)x_{nd} + 2\lambda W_{c,d}$$