SVM Derivative

Saif

February 21, 2018

Let us start with the loss function of Support Vector Machine. SVM just wants the score of correct class to be higher than incorrect class by the margin (delta).

SVM Loss Function

$$Li = \sum \max(0, s_j - s_{yi} + \Delta]) \tag{1}$$

$$\begin{aligned} s_{yi} &= scoreofcorrectclass \\ s_j &= scoreofincorrectclass. \end{aligned} \tag{2}$$

We sum across all incorrect classes. Because this loss function is the convex function, we can use derivative and optimization technique to reach at global minimum. This is the reason we calculate derivative which can be either done using calculus (analytically) or numerically. We can use both the techniques to check if we have calculated our gradient correctly.

Here we will discuss about calculating gradient using calculus.

Let us begin for single data point.

$$\begin{bmatrix} x_{11} & x_{12} & x_{13} \end{bmatrix} . \begin{bmatrix} w_{11} & w_{21} & w_{31} \\ w_{12} & w_{22} & w_{32} \\ w_{13} & w_{23} & w_{33} \end{bmatrix}$$

K=3 (Number of Classes) so at the end of calculation we expect 3 scores as output for 1 example

D=3(bias term added)

Dimension for single data point is 1x3 and the dimensions of Weights are 3x3 You can use different ways for calculation and for taking dot product.

$$Score1 = w_{11}x_{11} + w_{12}x_{12} + w_{13}x_{13} \tag{3}$$

$$Score2 = w_{21}x_{11} + w_{22}x_{12} + w_{23}x_{13} \tag{4}$$

These 2 represents score for 2 classes and in the similar manner you can calculate score for class 3 for particular example.

Loss Function (Let us say Class 2 is the correct class)

$$\begin{aligned} f = \max(0, ((w_{11}x_{11} + w_{12}x_{12} + w_{13}x_{13}) - (w_{21}x_{11} + w_{22}x_{12} + w_{23}x_{13}) + 1) \\ + \max(0, (w_{31}x_{11} + w_{32}x_{12} + w_{33}x_{13}) - (w_{21}x_{11} + w_{22}x_{12} + .w_{23}x_{13}) + 1)) \end{aligned} \tag{5}$$

If we arrange this equation in proper manner and gather like terms together this becomes

$$\begin{aligned} f = \max(0, (w_{11}x_{11} - w_{21}x_{11}) + (w_{12}x_{12} - w_{22}x_{12}) + (w_{13}x_{13} - w_{23}x_{13}) + 1) + \\ \max(0, (w_{31}x_{11} - w_{21}x_{11}) + (w_{32}x_{12} - w_{22}x_{12}) + (w_{33}x_{13} - w_{23}x_{13}) + 1) \end{aligned} \tag{6}$$

Now we have equation of loss function in simpler form and we can calculate derivative with respect to each weights. (Check if margin is violated and then calculate derivative).

$$\begin{gathered} \frac{\partial f}{\partial w_{11}} = x_{11} \\ \frac{\partial f}{\partial w_{12}} = x_{12} \\ \frac{\partial f}{\partial w_{13}} = x_{13} \end{gathered} \tag{7}$$

This particular example has class 2 as correct class so let us calculate derivative with respect to w2

$$\frac{\partial f}{\partial w_{21}} = -x_{11} - x_{11}.... = -x_{11} * NumberoftimesmarginViolated \tag{8}$$

$$\begin{aligned} \frac{\partial f}{\partial w_{22}} = -x_{12} - x_{12}.... = -x_{12} * NumberoftimesmarginViolated \\ . \\ . \end{aligned} \tag{9}$$

So if we see here, we have two cases for calculating derivative. We can generalise this easily so as you see right over here that if my class 2 is the correct class we will calculate derivative with respect to w2 differently than w1 and w3. You will calculate derivative with respect to w3 in the same manner as we calculated derivative with respect to w1 above.

You can derive rest of it with respect to several Ws. I hope that you have clear idea now on how to calculate derivative of SVM. You can extend this concept for single example for multiple examples.