The SVM loss is set up so that SVM "wants" the correct class for each image to have a score higher than the incorrect classes by some fixed margin Δ.

Let the score for the j-th class is the j-th element: $s_j = f(x_i, W)_j$. The Multiclass SVM loss for the i-th example is then formalized as follows:

$$L_i = \sum_{j \neq y_i} max(0, s_j - s_{y_i} + \Delta)$$

The threshold at zero $max(0, -)$ function is often called the *hinge loss*.

The total loss would be *data loss* plus *regularization loss*:

$$L = \frac{1}{N} \sum_i L_i + \lambda R(W)$$

The loss function quantifies our unhappiness with predictions on the training set.

The SVM "wants" the score of the correct class to be higher than all other scores by at least a margin of delta. If any class has a score inside the red region (or higher), then there will be accumulated loss. Otherwise the loss will be zero.

Our objective will be to find the weights that will simultaneously satisfy this constraint for all examples in the training data and give a total loss that is as low as possible.