

Activity 19



Loss functions

Classifying new data:

features \rightarrow weights
 $\hat{y} = \text{sign}(\mathbf{x}^T \mathbf{w})$

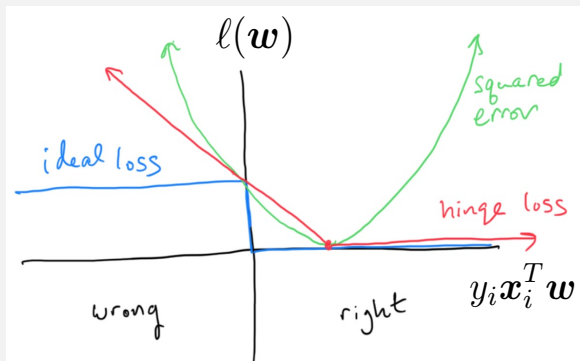
Training a classifier:

loss function
 $\min_{\mathbf{w}} \ell(\mathbf{w}) + \lambda r(\mathbf{w})$

$\ell(\mathbf{w})$

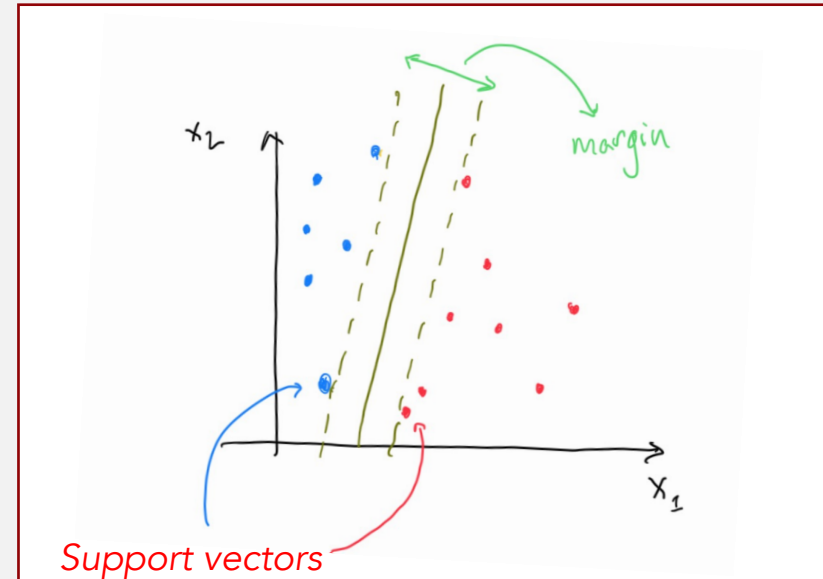
- squared error $\|\mathbf{X}\mathbf{w} - \mathbf{y}\|_2^2$
- ideal (0-1) loss $\sum_i \frac{1}{2} |y_i - \text{sign}(\mathbf{x}_i^T \mathbf{w})|$
- hinge loss $\sum_i (1 - y_i \mathbf{x}_i^T \mathbf{w})_+$
- logistic loss $\log(1 + e^{-y_i \mathbf{x}_i^T \mathbf{w}})$

Hinge Loss



$y_i \mathbf{x}_i^T \mathbf{w}$
 positive when correct
 negative when wrong

Support Vector Machines



maximize **margin**
 s.t. correct classification

minimize $\|\tilde{\mathbf{w}}\|^2$
 s.t. $y_i \mathbf{x}_i^T \mathbf{w} \geq 1$ for $i = 1, \dots$

For non-separable data: $\min_{\mathbf{w}} \sum_i (1 - y_i \mathbf{x}_i^T \mathbf{w})_+ + \lambda \|\tilde{\mathbf{w}}\|_2^2$