Support Vector Machines

Aarti Singh, HMW-Alexander

February 13, 2017

Back to Index

Contents

Resources

- Lecture

Pick the one with largest **margin**.
Separate line: ($\vec{w}$ is perpendicular to this line)

$$\vec{w}^T\vec{x} + b = 0$$

Margin:

$$\gamma = \frac{2a}{||\vec{w}||}$$

Confidence:

$$(\vec{w}^T\vec{x}_i + b)y_i$$

Object function of SVM:

$$\max_{\vec{w},b} \gamma = \frac{2a}{||\vec{w}||}, \quad s.t. \ (\vec{w}^T\vec{x}_i + b)y_i \geq a, \ \forall \ i$$

or

$$\min_{\vec{w}} \vec{w}^T\vec{w}, \quad s.t. \ (\vec{w}\vec{x}_i + b)y_i \geq 1, \ \forall \ i$$

1 Support Vectors

Linear yperplane defined by "support vectors". Only need to store the support vectors to predict labels of new points.

2 Data is ont linearly separable

Not quadratic programming:

$$\min_{\vec{w}} \vec{w}^T \vec{w} + C \# mistakes, \quad s.t.\ (\vec{w}^T \vec{x}_i + b) y_i \geq 1,\ \forall\ i$$

C is a tradeoff parameter.
Soft margin approach (QP):

$$\min_{\vec{w}} \vec{w}^T \vec{w} + C \sum_j \xi, \quad s.t.\ (\vec{w}^T \vec{x}_i + b) y_i \geq 1 - \xi_i\ \forall\ i$$

- ξ: slack variables = (¿1 if x_i misclassified)
- C: tradeoff parameter (chosen by cross-validation)

Hinge loss:

$$\xi_i = (1 - (\vec{w}^T \vec{x}_i + b) y_i)_+$$

Regularized hinge loss:

$$\min_{\vec{w}, b} \vec{w}^T \vec{w} + C \sum_i (1 - (\vec{w}^T \vec{x}_i + b) y_i)_+$$