Support Vector Machines cheatsheet

By Jérémy Fix, Hervé Frezza-Buet

Kernels

The separator reads

$$h_{w,b}(x) = \sum_i w_i k(x_i, x) + b$$

Linear	$k(x, x') = < x, x' >$
RBF	$k_\sigma(x, x') = \exp(-\frac{\|x-x'\|^2}{2\sigma^2})$
Polynomial	$k_{\gamma,c,d}(x, x') = (\gamma < x, x' > +c)^d$

Classification SVM : C-SVC

Primal optimization problem:

$$\underset{w,b\xi}{\operatorname{argmin}} \frac{1}{2}\|w\|^2 + C\mathbf{1}^T\xi$$

$$\text{subject to} \begin{cases} y_i h_{w,b}(x_i) \geq 1 - \xi_i & , \forall (x_i, y_i) \in S \\ \xi_i \geq 0 & , \forall i \end{cases}$$

Dual optimization problem:

$$\underset{\alpha}{\operatorname{argmin}} \frac{1}{2}\alpha^T Q \alpha - \mathbf{1}^T \alpha$$

$$\text{subject to} \begin{cases} y^T \alpha = 0 \\ 0 \leq \alpha_i \leq C, \forall i \end{cases}$$

The optimal separator reads :

$$h(x) = \sum_i \alpha_i y_i k(x_i, x) + b$$

For the optimal separator, we have either :

- $\alpha_i = 0$ for **non support vector** x_i, the associated sample does not contribute to the definition of the separating hyperplane,
- $0 < \alpha_i < C$ for **support vectors** x_i just on the margin. These samples are associated with $\xi_i = 0$ and $y_i h_{w,b}(x_i) = 1$
- $\alpha_i = C$ for **support vectors** x_i which involve non null slack variables; they are "on the wrong side" of the margin. These samples are associated with $\xi_i > 0$ and $y_i h_{w,b}(x_i) = 1 - \xi_i$

C-SVC examples