G. Montavon
Institute of Computer Science
Department of Mathematics and Computer Science
Freie Universität Berlin

Exercise Sheet 4 (theory part)

Exercise 1: Convex Optimization (10 + 10 + 20 + 20 + 10 P)

The SVDD method for anomaly detection is formulated as a convex optimization problem. Here, we consider a simplified variant of SVDD, the hard-margin SVDD, consisting of finding a maximum enclosing hypersphere of the data:

$$\min_{S,\boldsymbol{c}} \quad S$$

$$\text{s.t.} \quad \forall_{i=1}^N : \|\boldsymbol{c} - \boldsymbol{x}_i\|^2 \leq S$$

where S denotes the squared radius of the hypersphere.

(a) *Verify* that there is a point $(S, \boldsymbol{c})$ such that all constraints above are satisfied.

Always satisfied. Choose S large.

(b) *Write* the Lagrange function associated to this optimization problem.

$$\mathcal{L}(S, \boldsymbol{c}, \boldsymbol{\alpha}) = S + \sum_{i=1}^N \alpha_i \cdot (\|\boldsymbol{c} - \boldsymbol{x}_i\|^2 - S)$$

(c) *State* the KKT conditions for this problem and simplify them.

- Stationarity

$$\frac{\partial \mathcal{L}}{\partial S} \overset{\text{def}}{=} \mathbf{0} \quad \Rightarrow \quad \sum_{i=1}^N \alpha_i = 1$$

$$\frac{\partial \mathcal{L}}{\partial \boldsymbol{c}} \overset{\text{def}}{=} \mathbf{0} \quad \Rightarrow \quad \boldsymbol{c} = \sum_{i=1}^N \alpha_i \boldsymbol{x}_i$$

- Primal constraints

$$\forall_{i=1}^N : \|\boldsymbol{c} - \boldsymbol{x}_i\|^2 \leq S$$

- Dual constraints

$$\forall_{i=1}^N : 0 \leq \alpha_i$$

- Complementary slackness

$$\forall_{i=1}^N : \alpha_i \cdot (\|\boldsymbol{c} - \boldsymbol{x}_i\|^2 - S) = 0$$

(d) *Show* that the dual of SVDD is given by:

$$\max_{\boldsymbol{\alpha}} \quad \sum_{i=1}^{N} \alpha_i \|\boldsymbol{x}_i\|^2 - \sum_{i=1}^{N} \sum_{j=1}^{N} \alpha_i \alpha_j \boldsymbol{x}_i^\top \boldsymbol{x}_j$$

subject to:

$$\sum_{i=1}^{N} \alpha_i = 1,$$

$$\forall_{i=1}^{N} : 0 \leq \alpha_i$$

We express the Lagrange dual:

$$\begin{aligned}
&\arg\max_{\boldsymbol{\alpha}} \inf_{S,\boldsymbol{c}} \Big\{ S + \sum_{i=1}^{N} \alpha_i \cdot (\|\boldsymbol{c} - \boldsymbol{x}_i\|^2 - S) \Big\} \\
&= \arg\max_{\boldsymbol{\alpha}} \inf_{S,\boldsymbol{c}} \Big\{ S \cdot \underbrace{\Big(1 - \sum_{i=1}^{N} \alpha_i\Big)}_{=0} \\
&\qquad + \underbrace{\sum_{i=1}^{N} \alpha_i}_{=1} \boldsymbol{c}^\top \boldsymbol{c} - 2 \underbrace{\sum_{i=1}^{N} \alpha_i \boldsymbol{x}_i^\top}_{\boldsymbol{c}^\top} \boldsymbol{c} + \sum_{i=1}^{N} \alpha_i \|\boldsymbol{x}_i\|^2 \Big\} \\
&= \arg\max_{\boldsymbol{\alpha}} \inf_{S,\boldsymbol{c}} \Big\{ -\boldsymbol{c}^\top \boldsymbol{c} + \sum_{i=1}^{N} \alpha_i \|\boldsymbol{x}_i\|^2 \Big\} \\
&= \arg\max_{\boldsymbol{\alpha}} \Big\{ \sum_{i=1}^{N} \alpha_i \|\boldsymbol{x}_i\|^2 - \sum_{i=1}^{N} \sum_{j=1}^{N} \alpha_i \alpha_j \boldsymbol{x}_i^\top \boldsymbol{x}_j \Big\}
\end{aligned}$$

(e) *Express* the primal variables $S, \boldsymbol{c}$ as a function of dual variables $\alpha_1, \ldots, \alpha_N$.

To get $\boldsymbol{c}$, we simply get from the KKT stationary condition that

$$\boldsymbol{c} = \sum_{i=1}^{N} \alpha_i \boldsymbol{x}_i$$

To get S, we start from the complementary slackness condition

$$\forall_{i=1}^{N} : \alpha_i \cdot (\|\boldsymbol{c} - \boldsymbol{x}_i\|^2 - S) = 0$$

and note that for any index k satisfying $\alpha_k > 0$, the second term of the multiplication must be non-zero for the product to equate zero. In other words,

$$\|\boldsymbol{c} - \boldsymbol{x}_k\|^2 - S = 0.$$

Hence, $S = \|\boldsymbol{c} - \boldsymbol{x}_k\|^2$ for any k satisfying $\alpha_k > 0$.