Homework 3: Support Vector Machines

Writeup due 23:59 on Friday 27 March 2015

You will do this assignment individually and submit your answers as a PDF via the Canvas course website. There is a mathematical component and a programming component to this homework. Do not submit code.

1. Composing Kernel Functions [4pts]

Prove that

$$K(\boldsymbol{x}, \boldsymbol{x}') = \exp\{-||\boldsymbol{x} - \boldsymbol{x}'||_2^2\},$$

where $\boldsymbol{x}, \boldsymbol{x}' \in \mathbb{R}^D$ is a valid kernel, using only the following properties. If $K_1(\cdot, \cdot)$ and $K_2(\cdot, \cdot)$ are valid kernels, then the following are also valid kernels:

$$\begin{aligned}
K(\boldsymbol{x}, \boldsymbol{x}') &= c\, K_1(\boldsymbol{x}, \boldsymbol{x}') \quad \text{for } c > 0 \\
K(\boldsymbol{x}, \boldsymbol{x}') &= K_1(\boldsymbol{x}, \boldsymbol{x}') + K_2(\boldsymbol{x}, \boldsymbol{x}') \\
K(\boldsymbol{x}, \boldsymbol{x}') &= K_1(\boldsymbol{x}, \boldsymbol{x}')\, K_2(\boldsymbol{x}, \boldsymbol{x}') \\
K(\boldsymbol{x}, \boldsymbol{x}') &= \exp\{K_1(\boldsymbol{x}, \boldsymbol{x}')\} \\
K(\boldsymbol{x}, \boldsymbol{x}') &= f(\boldsymbol{x})\, K_1(\boldsymbol{x}, \boldsymbol{x}')\, f(\boldsymbol{x}') \quad \text{where } f \text{ is any function from } \mathbb{R}^D \text{ to } \mathbb{R}
\end{aligned}$$

2. Slack Variables and Importances [5pts]

Derive the dual problem in the "soft margin" case of the support vector machine, where the data are not linearly separable and there are "slack variables" ξ_n for each datum. Now, imagine that each data point has an "importance" to it in forming our classifier, $r_n > 0$, much like in the regression homework. Derive a dual problem for this case and explain how you arrived at it.

3. SVM as Quadratic Program [5pts]

Quadratic programs are optimization problems that have the following form:

$$\begin{aligned}
\boldsymbol{z}^\star = \arg\min_{\boldsymbol{z}} \frac{1}{2} \boldsymbol{z}^\top \boldsymbol{P} \boldsymbol{z} + \boldsymbol{q}^\top \boldsymbol{z} \\
\text{such that } \boldsymbol{G}\boldsymbol{z} \geq \boldsymbol{h} \text{ and } \boldsymbol{A}\boldsymbol{z} = \boldsymbol{b}
\end{aligned}$$

To solve the primal "hard margin" case of the SVM with a quadratic program, what would we use for $\boldsymbol{P}$, $\boldsymbol{q}$, $\boldsymbol{G}$, $\boldsymbol{h}$, $\boldsymbol{A}$ and $\boldsymbol{b}$? What are these values for the dual version of the problem? [Hint: for the primal problem you'll need to make $\boldsymbol{z}$ include both the weights and the bias term.]

4. Implementing an SVM [10pts]

In the previous homework, you studied a simple data set of fruit measurements. Code up a simple linear SVM to classify lemons from apples. Use a quadratic programming solver, such as `qp` from CVXOPT[1]. First, implement the primal problem and plot your decision boundary and margins. Then, implement the dual version of the problem. Which data are the support vectors?

5. Calibration [1pt]

Approximately how long did this homework take you to complete?

Changelog

- **v1.0** – 14 March 2015 at 12:00
- **v1.1** – 14 March 2015 at 13:00 – Fixed small typo.

[1] http://cvxopt.org/