Exercise Sheet 10

Machine Learning 2, SS16

July 14, 2016

Mario Tambos, 380599; Viktor Jeney, 348969; Sascha Huk, 321249; Jan Tinapp, 0380549

Exercise 1

The Lagrangian of the given primal problem is:

$$\mathcal{L}(R,c,\xi,\alpha,\lambda) = R^2 + \frac{1}{n\nu}\sum_i \xi_i - \sum_i \alpha_i(R^2 + \xi_i - \|\phi(x_i) - c\|) - \sum_i \lambda_i \xi_i$$

$$= R^2 + \frac{1}{n\nu}\sum_i \xi_i - \sum_i \alpha_i(R^2 + \xi_i - (\phi(x_i)^\top \phi(x_i) + c^\top c - 2c^\top \phi(x_i))) - \sum_i \lambda_i \xi_i$$

$$= R^2 + \frac{1}{n\nu}\sum_i \xi_i - \sum_i \alpha_i(R^2 + \xi_i - \phi(x_i)^\top \phi(x_i) - c^\top c + 2c^\top \phi(x_i)) - \sum_i \lambda_i \xi_i$$

We now differentiate w.r.t. primal variables R, c, ξ:

$$\frac{\partial}{\partial R}\mathcal{L}(R,c,\xi,\alpha,\lambda) = 2R - 2R\sum_i \alpha_i \stackrel{!}{=} 0 \qquad \Longrightarrow \qquad \sum_i \alpha_i = 1$$

$$\frac{\partial}{\partial c}\mathcal{L}(R,c,\xi,\alpha,\lambda) = 2c\sum_i \alpha_i - 2\sum_i \alpha_i \phi(x_i) \stackrel{!}{=} 0 \qquad \Longrightarrow \qquad c = \sum_i \alpha_i \phi(x_i)$$

$$\frac{\partial}{\partial \xi_i}\mathcal{L}(R,c,\xi,\alpha,\lambda) = \frac{1}{n\nu} - \alpha_i - \lambda_i \stackrel{!}{=} 0 \qquad \Longrightarrow \qquad \frac{1}{n\nu} = \alpha_i + \lambda_i$$

The Lagrangian of the dual problem then can be obtained by plugging in the derived results:

$$\mathcal{L}(\alpha,\lambda) = R^2 + \frac{1}{n\nu}\sum_i \xi_i - \sum_i \alpha_i(R^2 + \xi_i - \phi(x_i)^\top \phi(x_i) - c^\top c + 2c^\top \phi(x_i)) - \sum_i \lambda_i \xi_i$$

$$= \cancel{R^2} + \frac{1}{n\nu}\sum_i \xi_i - \cancel{R^2}\underbrace{\cancel{\sum_i \alpha_i}}_{1} - \sum_i \alpha_i \xi_i + \sum_i \alpha_i k(x_i,x_i) + c^\top c\underbrace{\sum_i \alpha_i}_{1} - 2c^\top \sum_i \alpha_i \phi(x_i) - \sum_i \lambda_i \xi_i$$

$$= \cancel{\frac{1}{n\nu}\sum_i \xi_i} - \cancel{\sum_i \underbrace{(\alpha_i + \lambda_i)}_{\frac{1}{n\nu}} \xi_i} + \sum_i \alpha_i k(x_i,x_i) + c^\top c - 2c^\top \underbrace{\sum_i \alpha_i \phi(x_i)}_{c}$$

$$= \sum_i \alpha_i k(x_i,x_i) - c^\top c$$

Lastly, by using the definition of c, we obtain the dual program:

$$\max_\alpha \sum_i \alpha_i k(x_i,x_i) - \sum_i \sum_j \alpha_i \alpha_j k(x_i,x_j)$$

$$\text{s.t.} \qquad \sum_i \alpha_i = 1 \quad \text{and} \quad \alpha_i \geq 0 \quad \text{and} \quad \lambda_i \geq 0 \quad \text{and} \quad \alpha_i + \lambda_i = \frac{1}{n\nu}$$

$$\Longrightarrow \max_\alpha \sum_i \alpha_i k(x_i,x_i) - \sum_i \sum_j \alpha_i \alpha_j k(x_i,x_j) \qquad \text{s.t.} \qquad \sum_i \alpha_i = 1 \quad \text{and} \quad \frac{1}{n\nu} \geq \alpha_i \geq 0$$

The primal variable c is determined by $c = \sum_i \alpha_i \phi(x_i)$. R, the support vectors, then can be found by using the constraint of the primal problem (by finding the points on the gutter, i.e. solving $\|\phi(x_i) - c\| = R$).

Exercise 2

Form of the QP problem:

$$\min_x \frac{1}{2}^\top Px + q^\top x \qquad \text{s.t.} \quad Gx \preceq h \quad \text{and} \quad Ax = b$$

The variables have been identified as follows:

$$\begin{array}{llll} x \widehat{=} \alpha & P \widehat{=} (K_{i,j} y_i y_j)_{i,j} & q \widehat{=} -\mathbf{1} & G \widehat{=} -\mathbf{I} \\ h \widehat{=} \mathbf{0} & A \widehat{=} y & b \widehat{=} \mathbf{0} & \end{array}$$

The objective of the dual problem can be further derived like this:

$$\begin{aligned} \min_{w,b} \max_{\alpha} L(w,b,\alpha) &= \min_{w,b} \max_{\alpha} \frac{1}{2}||w||^2 - \sum_{i=1}^{n} \alpha_i (y_i(wx_i + b) - 1) \\ \Longrightarrow \max_{\alpha > \mathbf{0}} L(\alpha) &= \max_{\alpha > \mathbf{0}} \sum_{i=1}^{n} \alpha_i - \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} \alpha_i \alpha_j y_i y_j K_{i,j} \quad \text{s.t.} \quad \sum_i \alpha_i y_i = 0 \\ \Longrightarrow \min_{\alpha > \mathbf{0}} -L(\alpha) &= \min_{\alpha > \mathbf{0}} \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} \alpha_i \alpha_j y_i y_j K_{i,j} - \sum_{i=1}^{n} \alpha_i \quad \text{s.t.} \quad \sum_i \alpha_i y_i = 0 \end{aligned}$$

It is now easy to match the pattern of the QP problem.