

# Exercise Sheet 11

Machine Learning 2, SS16

July 15, 2016

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

## Exercise 1

The Lagrangian of the given primal problem is:

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

We now differentiate w.r.t. primal variables  $R, c, \xi$ :

$$\begin{aligned}\frac{\partial}{\partial R} \mathcal{L}(R, c, \xi, \alpha, \lambda) &= 2R - 2R \sum_i \alpha_i \stackrel{!}{=} 0 \quad \implies \quad \sum_i \alpha_i = 1 \\ \frac{\partial}{\partial c} \mathcal{L}(R, c, \xi, \alpha, \lambda) &= 2c \sum_i \alpha_i - 2 \sum_i \alpha_i \phi(x_i) \stackrel{!}{=} 0 \quad \implies \quad c = \sum_i \alpha_i \phi(x_i) \\ \frac{\partial}{\partial \xi_i} \mathcal{L}(R, c, \xi, \alpha, \lambda) &= \frac{1}{n\nu} - \alpha_i - \lambda_i \stackrel{!}{=} 0 \quad \implies \quad \frac{1}{n\nu} = \alpha_i + \lambda_i\end{aligned}$$

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

$$\begin{aligned}\mathcal{L}(\alpha, \lambda) &= R^2 + \frac{1}{n\nu} \sum_i \xi_i - \sum_i \alpha_i (R^2 + \xi_i - \phi(x_i)^\top \phi(x_i) - c^\top c + 2c^\top \phi(x_i)) - \sum_i \lambda_i \xi_i \\ &= \cancel{R^2} + \frac{1}{n\nu} \sum_i \xi_i - \cancel{R^2} \underbrace{\sum_i \alpha_i}_1 - \sum_i \alpha_i \xi_i + \sum_i \alpha_i k(x_i, x_i) + c^\top c \underbrace{\sum_i \alpha_i}_1 - 2c^\top \sum_i \alpha_i \phi(x_i) - \sum_i \lambda_i \xi_i \\ &= \cancel{\frac{1}{n\nu} \sum_i \xi_i} - \sum_i \underbrace{(\alpha_i + \lambda_i)}_{\frac{1}{n\nu}} \xi_i + \sum_i \alpha_i k(x_i, x_i) + c^\top c - 2c^\top \underbrace{\sum_i \alpha_i \phi(x_i)}_c \\ &= \sum_i \alpha_i k(x_i, x_i) - c^\top c\end{aligned}$$

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

$$\begin{aligned}\max_{\alpha} \quad & \sum_i \alpha_i k(x_i, x_i) - \sum_i \sum_j \alpha_i \alpha_j k(x_i, x_j) \\ \text{s.t.} \quad & \sum_i \alpha_i = 1 \quad \text{and} \quad \alpha_i \geq 0 \quad \text{and} \quad \lambda_i \geq 0 \quad \text{and} \quad \alpha_i + \lambda_i = \frac{1}{n\nu} \\ \implies \max_{\alpha} \quad & \sum_i \alpha_i k(x_i, x_i) - \sum_i \sum_j \alpha_i \alpha_j k(x_i, x_j) \quad \text{s.t.} \quad \sum_i \alpha_i = 1 \quad \text{and} \quad \frac{1}{n\nu} \geq \alpha_i \geq 0\end{aligned}$$

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

## Exercise 2

Form of the QP problem:

$$\min_{\alpha} \alpha^{\top} P \alpha + q^{\top} \alpha \quad \text{s.t.} \quad G \alpha \preceq h \quad \text{and} \quad A \alpha = b$$

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

$$\begin{aligned} \max_{\alpha} \mathcal{L}(\alpha) &= \max_{\alpha} \sum_i \alpha_i k(x_i, x_i) - \sum_i \sum_j \alpha_i \alpha_j k(x_i, x_j) \quad \text{s.t.} \quad \sum_i \alpha_i = 1 \quad \text{and} \quad \frac{1}{n\nu} \geq \alpha_i \geq 0 \\ \implies \min_{\alpha} -\mathcal{L}(\alpha) &= \min_{\alpha} \sum_i \sum_j \alpha_i \alpha_j k(x_i, x_j) - \sum_i \alpha_i k(x_i, x_i) \quad \text{s.t.} \quad \sum_i \alpha_i = 1 \quad \text{and} \quad \frac{1}{n\nu} \geq \alpha_i \geq 0 \end{aligned}$$

The variables can now be identified as follows:

$$\begin{aligned} P &\hat{=} (K(x_i, x_j))_{i,j} & q &\hat{=} -(K(x_i, x_i))_i & G &\hat{=} \begin{bmatrix} \mathbf{I}_n \\ -\mathbf{I}_n \end{bmatrix} \\ h &\hat{=} \frac{1}{n\nu} \begin{bmatrix} \vec{1}_n \\ \vec{0}_n \end{bmatrix} & A &\hat{=} \begin{bmatrix} 1 & \dots & 1 \\ \phi(x_1) & \dots & \phi(x_n) \end{bmatrix} & b &\hat{=} \begin{bmatrix} 1 \\ \vec{c} \end{bmatrix} \end{aligned}$$