Exercise Sheet 1

Machine Learning 2, SS16

April 30, 2016

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

Exercise 2

(i) We have to prove $w^\top C w \overset{?}{=} \epsilon \overset{(3)}{=} |x - \sum_j w_j \eta_j|^2$.

$$
\begin{aligned}
&w^\top C w \\
&= w^\top (\mathbb{1}_k x^\top - \eta)(\mathbb{1}_k x^\top - \eta)^\top w \\
&= w^\top (\mathbb{1}_k x^\top - \eta)(x \mathbb{1}_k^\top - \eta^\top) w \\
&= w^\top \mathbb{1}_k x^\top x \mathbb{1}_k^\top w - w^\top \mathbb{1}_k x^\top \eta^\top w - w^\top \eta x \mathbb{1}_k^\top w + w^\top \eta \eta^\top w \\
&= (w^\top \mathbb{1}_k x^\top)(x \mathbb{1}_k^\top w) - 2(w^\top \mathbb{1}_k x^\top)(\eta^\top w) + (w^\top \eta)(\eta^\top w) \\
&= |(w^\top \mathbb{1}_k x^\top) - (w^\top \eta)|^2 \\
&= |w^\top (\mathbb{1}_k x^\top - \eta)|^2 \\
&= |\sum_j w_j (x - \eta_j)|^2
\end{aligned}
$$

Since $\sum_i w_i = 1$ we find $\sum_i w_i x = x$, which leads to the desired result. $\square$

We now perform Lagrange optimization:

$$
\begin{aligned}
\Lambda(w, \lambda) &= w^\top C w - \lambda(w^\top \mathbb{1}_k - 1) \\
\frac{\partial \Lambda}{\partial w} &= 2Cw - \lambda \mathbb{1}_k \overset{!}{=} 0 \quad \Longrightarrow \quad 2Cw = \lambda \mathbb{1}_k \quad \Longrightarrow \quad w = \frac{\lambda}{2} C^{-1} \mathbb{1}_k \\
\frac{\partial \Lambda}{\partial \lambda} &= w^\top \mathbb{1}_k - 1 \overset{!}{=} 0 \quad \Longrightarrow \quad w^\top \mathbb{1}_k = 1
\end{aligned}
$$

We now replace w in the constraint $w^\top \mathbb{1}_k = 1$. Since $C = C^\top$ we find

$$\frac{\lambda}{2} = \frac{1}{\mathbb{1}_k^\top C^{-1} \mathbb{1}_k}$$

(ii) Replacing $\lambda/2$ in the deduced definition of w leads to the desired result. The candidate w is indeed a minimum since $\frac{\partial^2 \Lambda}{\partial w^2} = 2C$ (invertible covariance matrices have positive definite quadratic forms). $\square$

(iii) Multiplying by C from the left in the deduced definition of w leads to $Cw = \frac{\lambda}{2} \mathbb{1}_k$ (resp. $Cw' = \mathbb{1}_k$ for $w' = \frac{2}{\lambda} w$). Both, w' and w, point to the same direction. Therefore, by rescaling w' such that ${w'}^\top \mathbb{1}_k = 1 = w^\top \mathbb{1}_k$ the desired w is identified. $\square$