Problem Solutions

e-Chapter 8

Pierre Paquay

Problem 8.1

The two separation constraints are

$$(w^T x_+ + b) \geq 1 \text{ and } -(w^T x_- + b) \geq 1;$$

by adding these two constraints, we get that

$$w^T(x_+ - x_-) \geq 2.$$

Then, the Cauchy-Schwarz inequality gives us the following inequalities

$$2 \leq w^T(x_+ - x_-) \leq |w^T(x_+ - x_-)| \leq ||w|| ||x_+ - x_-||;$$

consequently, we get that

$$||w|| \geq \frac{2}{||x_+ - x_-||}.$$

Since we seek to minimize $||w||$, we choose w^* such that

$$||w^*|| = \frac{2}{||x_+ - x_-||}.$$

In this case, as we want w^* to satisfy both constraints, we may note that

$$2 \leq w^{*T}(x_+ - x_-) \leq |w^{*T}(x_+ - x_-)| \leq ||w^*|| ||x_+ - x_-|| = 2.$$

This means that

$$|w^{*T}(x_+ - x_-)| = ||w^*|| ||x_+ - x_-||,$$

which can only happen when $w^* = k(x_+ - x_-)$. Since, we have already established that

$$||w^*|| = \frac{2}{||x_+ - x_-||},$$

we choose k to be

$$k = \frac{2}{||x_+ - x_-||^2}.$$

Now, we may write that

$$w^* = \frac{2(x_+ - x_-)}{||x_+ - x_-||^2}.$$

It remains to determine the value of b^*. To do that we fix the following equality

$$2\left(\frac{(x_+ - x_-)}{||x_+ - x_-||^2}\right)^T x_+ + b^* = 1;$$

which gives us that

$$\begin{aligned} b^* &= 1 - 2\frac{x_+^T x_+ - x_-^T x_+}{||x_+ - x_-||^2} \\ &= \frac{x_-^T x_- - x_+^T x_+}{||x_+ - x_-||^2} \\ &= \frac{||x_-||^2 - ||x_+||^2}{||x_+ - x_-||^2}. \end{aligned}$$

It is now easy to verify that (w^*, b^*) satisfies both constraints and minimizes $||w||$, and therefore gives us the optimal hyperplane.

Problem 8.2

In this case, the constraints are

$$-b \geq 1, \ -(-w_2 + b) \geq 1, \ (-2w_1 + b) \geq 1.$$

If we combine the first and the third ones, we get $w_1 \leq -1$. The quantity we seek to minimize is

$$\frac{1}{2} w^T w = \frac{1}{2}(w_1^2 + w_2^2) \geq \frac{1}{2}(1 + 0) \geq \frac{1}{2},$$

where we have equality when $w_1 = -1$ and $w_2 = 0$; consequently, we choose $w^* = (-1, 0)$. With this in mind, the third constraint becomes

$$1 \leq -2w_1^* + b = 2 + b \Leftrightarrow b \geq -1;$$

so we choose $b^* = -1$. It is now easy to verify that (w^*, b^*) satisfies both constraints and minimizes $||w||$, and therefore gives us the optimal hyperplane. The margin in this case is given by $1/||w^*|| = 1$.

Problem 8.3

(a) We begin by computing the Lagrangian, we get

$$\begin{aligned} \mathcal{L}(\alpha) &= \frac{1}{2}\sum_n \sum_m y_n y_m \alpha_n \alpha_m x_n^T x_m - \sum_n \alpha_n \\ &= \frac{1}{2}(8\alpha_2^2 - 4\alpha_2\alpha_3 - 6\alpha_2\alpha_4 - 4\alpha_2\alpha_3 + 4\alpha_3^2 + 6\alpha_3\alpha_4 - 6\alpha_4\alpha_2 + 6\alpha_3\alpha_4 + 9\alpha_4^2) - \alpha_1 - \alpha_2 - \alpha_3 - \alpha_4 \\ &= 4\alpha_2^2 + 2\alpha_3^2 + \frac{9}{2}\alpha_4^2 - 4\alpha_2\alpha_3 - 6\alpha_2\alpha_4 + 6\alpha_3\alpha_4 - \alpha_1 - \alpha_2 - \alpha_3 - \alpha_4. \end{aligned}$$

Concerning the constraints, we have that

$$0 = \sum_n y_n \alpha_n = -\alpha_1 - \alpha_2 + \alpha_3 + \alpha_4,$$

or equivalently

$$\alpha_1 + \alpha_2 = \alpha_3 + \alpha_4$$

with $\alpha_1, \alpha_2, \alpha_3, \alpha_4 \geq 0$.

(b) If we replace α_1 with $\alpha_3 + \alpha_4 - \alpha_2$, we obtain

$$\mathcal{L}(\alpha) = 4\alpha_2^2 + 2\alpha_3^2 + \frac{9}{2}\alpha_4^2 - 4\alpha_2\alpha_3 - 6\alpha_2\alpha_4 + 6\alpha_3\alpha_4 - 2\alpha_3 - 2\alpha_4.$$

(c) Now, we fix α_3 and α_4 and we take the derivative of $\mathcal{L}(\alpha)$ with respect to α_2, this gives us that

$$\frac{\partial \mathcal{L}}{\partial \alpha_2} = 8\alpha_2 - 4\alpha_3 - 6\alpha_4.$$

By setting the previous expression to 0, we get that

$$\alpha_2 = \frac{\alpha_3}{2} + \frac{3\alpha_4}{4},$$

and also that

$$\alpha_1 = -\alpha_2 + \alpha_3 + \alpha_4 = \frac{\alpha_3}{2} + \frac{\alpha_4}{4}.$$

These expressions are valid since they are both greater or equal to 0, and obviously

$$\alpha_1 + \alpha_2 = \alpha_3 + \alpha_4.$$

(d) It remains to replace α_2 by its new expression (in (c)), we get that

$$\begin{aligned}
\mathcal{L}(\alpha) &= 4\left(\frac{\alpha_3}{2} + \frac{3\alpha_4}{4}\right)^2 + 2\alpha_3^2 + \frac{9}{2}\alpha_4^2 - 4\left(\frac{\alpha_3}{2} + \frac{3\alpha_4}{4}\right)\alpha_3 - 6\left(\frac{\alpha_3}{2} + \frac{3\alpha_4}{4}\right)\alpha_4 + 6\alpha_3\alpha_4 - 2\alpha_3 - 2\alpha_4 \\
&= \alpha_3^2 + (3\alpha_4 - 2)\alpha_3 + \frac{9}{4}\alpha_4^2 - 2\alpha_4 \\
&= \left(\alpha_3 + \frac{3\alpha_4 - 2}{2}\right)^2 + \frac{9}{4}\alpha_4^2 - 2\alpha_4 - \frac{(3\alpha_4 - 2)^2}{4} \\
&= \left(\alpha_3 + \frac{3\alpha_4 - 2}{2}\right)^2 + \alpha_4 - 1 \geq -1.
\end{aligned}$$

The minimum of the Lagrangian is attained when $\alpha_3 = 1$ and $\alpha_4 = 0$, in this case we also have

$$\alpha_1 = \frac{\alpha_3}{2} + \frac{\alpha_4}{4} = \frac{1}{2}$$

and

$$\alpha_2 = \frac{\alpha_3}{2} + \frac{3\alpha_4}{4} = \frac{1}{2}.$$

Problem 8.4

We have

$$X = \begin{pmatrix} 0 & 0 \\ 2 & 2 \\ 2 & 0 \end{pmatrix} \text{ and } y = \begin{pmatrix} -1 \\ -1 \\ 1 \end{pmatrix}.$$

The Lagrangian is equal to

$$\begin{aligned}
\mathcal{L}(\alpha) &= \frac{1}{2}\sum_n \sum_m y_n y_m \alpha_n \alpha_m x_n^T x_m - \sum_n \alpha_n \\
&= 4\alpha_2^2 - 4\alpha_2\alpha_3 + 2\alpha_3^2 - \alpha_1 - \alpha_2 - \alpha_3 \\
&= 2(\alpha_1^2 - \alpha_1) + 2(\alpha_2^2 - \alpha_2) \geq -\frac{1}{2} - \frac{1}{2} \geq -1;
\end{aligned}$$

and the constraints are $\alpha_3 = \alpha_1 + \alpha_2$ with $\alpha_1, \alpha_2, \alpha_3 \geq 0$. The minimum of the Lagrangian is attained when $\alpha_1 = \alpha_2 = 1/2$ which gives us $\alpha_3 = \alpha_1 + \alpha_2 = 1$. Then, the optimal Lagrange multipliers are

$$\alpha_1^* = \frac{1}{2},\ \alpha_2^* = \frac{1}{2},\ \text{and } \alpha_3^* = 1.$$