

# Exercises to Prepare for SVM and Lagrangian Lectures

*David S. Rosenberg*

## 1 Equivalent Optimization Problems

Suppose we have two functions  $f : \mathbf{R}^d \rightarrow \mathbf{R}$  and  $g : \mathbf{R}^d \rightarrow \mathbf{R}$ . And suppose we have the following optimization problem:

$$\min_{x \in \mathbf{R}^d} f(x) + g(x).$$

This is an unconstrained optimization problem. Now consider the following constrained optimization problem:

This problem is “equivalent” to the following problem:

$$\begin{array}{ll} \text{minimize} & f(x) + \xi \\ \text{subject to} & \xi \geq g(x). \end{array}$$

When an optimization problem is presented in this form, it should be understood as a minimization over all variables that are unknown. This case, we are minimizing over  $x \in \mathbf{R}^d$  and  $\xi \in \mathbf{R}$ .

We now claim that these two problems are “equivalent” in the following sense:

- Suppose the second problem attains a minimum at  $(x^*, \xi^*)$ , and that minimum is  $M$ . Then the first problem also has a minimum value of  $M$  and it is attained at  $x^*$ . [It follows that  $\xi^* = g(x^*)$ .]
- Conversely, if the first problem attains a minimum at  $x^*$ , then there is a  $\xi^*$  for which  $(x^*, \xi^*)$  is a minimizer of the second problem, and the minimum values are the same.

**Exercise 1.** Convince yourself that these two problems are equivalent. [Hint/Answer: In the second problem, for any fixed value of  $x$ , the objective is always minimized (subject to the constraint) by  $\xi = g(x)$ .]

**Exercise 2.** Recall the definition of the “positive part” of a number:

$$(x)_+ = x1(x \geq 0) = \begin{cases} x & \text{if } x \geq 0 \\ 0 & \text{otherwise.} \end{cases}$$

Convince yourself that the problem

$$\min_{w \in \mathbf{R}^d} f(w) + \sum_{i=1}^n (1 - y_i [w^T x_i + b])_+$$

is equivalent to

$$\begin{aligned} & \text{minimize} && f(w) + \sum_{i=1}^n \xi_i \\ & \text{subject to} && \xi_i \geq (1 - y_i [w^T x_i + b])_+ \text{ for } i = 1, \dots, n, \end{aligned}$$

which is equivalent to

$$\begin{aligned} & \text{minimize} && f(w) + \sum_{i=1}^n \xi_i \\ & \text{subject to} && \xi_i \geq 0 \text{ for } i = 1, \dots, n \\ & && \xi_i \geq 1 - y_i [w^T x_i + b] \text{ for } i = 1, \dots, n. \end{aligned}$$

**Exercise 3.** Convince yourself that the following two optimization problems are equivalent. First problem:

$$\begin{aligned} & \text{minimize} && f(x) \\ & \text{subject to} && x_i + \alpha_i = c \text{ for } i = 1, \dots, n, \\ & && x_i \geq 0, \alpha_i \geq 0 \text{ for } i = 1, \dots, n, \end{aligned}$$

for some known  $c$ .

Second problem:

$$\begin{aligned} & \text{minimize} && f(x) \\ & \text{subject to} && x_i \in [0, c] \text{ for } i = 1, \dots, n. \end{aligned}$$

(Hint: Figure out what value  $\alpha_i$  is for any given  $x_i$ . And what constraints to we need on  $x_i$  to satisfy the constraints, and so that the corresponding  $\alpha_i$  also satisfies its constraints?)

## 2 Lagrangian Encodes Objective and Constraints

First some shorthand: If  $\lambda \in \mathbf{R}^d$ , we write  $\lambda \succeq 0$  as a shorthand for  $\lambda_i \geq 0 \ i = 1, \dots, d$ . Similarly, if  $c \in \mathbf{R}^d$ , then  $\lambda \succeq c$  is shorthand for  $\lambda - c \succeq 0$ .

We claim that

$$\sup_{\lambda \succeq 0} (f(x) + \lambda g(x)) = \begin{cases} f(x) & \text{for } g(x) \leq 0 \\ \infty & \text{otherwise.} \end{cases}$$

**Exercise 4.** Convince yourself that this is true. (Hint: Find the sup when  $g(x) \leq 0$  and when  $g(x) > 0$ .)

**Exercise 5.** Show that the following optimization problems are equivalent:

$$\begin{array}{ll} \text{minimize} & f(x) \\ \text{subject to} & g(x) \leq 0 \end{array}$$

is equivalent to

$$\inf_x \left( \sup_{\lambda \succeq 0} (f(x) + \lambda g(x)) \right).$$

Hint/Solution: Based on the previous exercise, if  $g(x) > 0$  (i.e.  $x$  is “not feasible” for the first optimization problem), then  $\sup_{\lambda \succeq 0} (f(x) + \lambda g(x)) = \infty$ . So the infimum of the second optimization problem will not occur at any  $x$  where  $g(x) > 0$ . Thus the following problem is equivalent to the second problem:

$$\inf_{\{x | g(x) \leq 0\}} \left( \sup_{\lambda \succeq 0} (f(x) + \lambda g(x)) \right).$$

But when  $g(x) \leq 0$ , we know from the previous exercise that the supremum evaluates to  $f(x)$ . Thus the second optimization problem is also equivalent to

$$\inf_{\{x | g(x) \leq 0\}} f(x),$$

and this is exactly the first optimization problem.