Exercises to Prepare for SVM and Lagrangian Lectures

David S. Rosenberg

1 Equivalent Optimization Problems

Suppose we have two functions $f : \mathbf{R}^d \to \mathbf{R}$ and $g : \mathbf{R}^d \to \mathbf{R}$. Now consider the following optimization problem:

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

This is an unconstrained optimization problem. Let's also consider the following constrained optimization problem:

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

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

We claim that these two problems are "equivalent" in the following sense:

- Suppose the second problem attains a minimum at (x^*, ξ^*), 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 ξ^* for which (x^*, ξ^*) 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 $\xi = g(x)$.

Remark 2. The equivalence shown above is a very strict equivalence. We may also speak more loosely and say that two problems are equivalent if we can easily derive a solution to one of them given a solution to the other one, even if the minimizers and minima are different. For example, if we know that $\arg\min_x \exp\left[f(2x)\right] = x^*$, then we can immediately conclude that $\arg\min_x f(x) = 2x^*$.

Exercise 3. 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

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

is equivalent to

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

which is equivalent to

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

Exercise 4. Convince yourself that the following two optimization problems are equivalent. First problem:

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

for some known c.

Second problem:

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

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

2 Lagrangian Encodes Objective and Constraints (OPTIONAL)

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

We claim that

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

Exercise 5. Convince yourself that this is true. (Hint: Find the sup when $g(x) \leq 0$ and when $g(x) > 0$.)

Exercise 6. Show that the following optimization problems are equivalent:

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

is equivalent to

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

Hint/Solution: Based on the previous exerise, if $g(x) > 0$ (i.e. x is "not feasible" for the first optimization problem), then $\sup_{\lambda \succeq 0} \left(f(x) + \lambda g(x) \right) = \infty$. 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:

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

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

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

and this is exactly the first optimization problem.