Maths for Intelligent Systems
Exercise 8

Marc Toussaint
Machine Learning & Robotics lab, U Stuttgart
Universitätsstraße 38, 70569 Stuttgart, Germany

January 14, 2020

1 Lagrangian and dual function

(Taken roughly from 'Convex Optimization', Ex. 5.1)
Consider the optimization problem

$$\min_x \; x^2 + 1 \qquad \text{s.t.} \quad (x-2)(x-4) \le 0$$

with variable $x \in \mathbb{R}$.

a) Derive the optimal solution x^* and the optimal value $p^* = f(x^*)$ by hand.

b) Write down the Lagrangian $L(x, \lambda)$. Plot (using gnuplot or so) $L(x, \lambda)$ over x for various values of $\lambda \ge 0$. Verify the lower bound property $\min_x L(x, \lambda) \le p^*$, where p^* is the optimum value of the primal problem.

c) Derive the dual function $l(\lambda) = \min_x L(x, \lambda)$ and plot it (for $\lambda \ge 0$). Derive the dual optimal solution $\lambda^* = \text{argmax}_\lambda \, l(\lambda)$. Is $\max_\lambda l(\lambda) = p^*$ (strong duality)?

2 Optimize a constrained problem

Consider the following constrained problem

$$\min_x \sum_{i=1}^n x_i \quad \text{s.t.} \quad g(x) \le 0 \tag{1}$$

$$g(x) = \begin{pmatrix} x^\top x - 1 \\ -x_1 \end{pmatrix} \tag{2}$$

a) First, assume $x \in \mathbb{R}^2$ is 2-dimensional, and draw on paper what the problem looks like and where you expect the optimum.

b) Find the optimum analytically using the Lagrangian. Here, assume that you know apriori that all constraints are active! What are the dual parameters $\lambda = (\lambda_1, \lambda_2)$?

Note: Assuming that you know a priori which constraints are active is a huge assumption! In real problems, this is the actual hard (and combinatorial) problem. More on this later in the lecture.

c) Implement a simple the Log Barrier Method. Tips:

- Initialize $x = (\frac{1}{2}, \frac{1}{2})$ and $\mu = 1$
- First code an inner loop:
  - In each iteration, first compute the gradient of the log-barrier function. Recall that

$$F(x; \mu) = f(x) - \mu \sum_i \log(-g_i(x)) \tag{3}$$

$$\nabla F(x; \mu) = \nabla f - \mu \sum_i (1/g_i(x)) \nabla g_i(x) \tag{4}$$

  - Then perform a backtracking line search along $-\nabla F(x, \mu)$. In particular, backtrack if a step goes beyond the barrier (where $g(x) \not\le 0$ and $F(x, \mu) = \infty$).

  - Iterate until convergence; let's call the result $x^*(\mu)$. Further, compute $\lambda^*(m) = -(\mu/g_1(x), \mu/g_2(x))$ at convergence.
- Decrease $\mu \leftarrow \mu/2$, recompute $x^*(\mu)$ (with the previous x^* as initialization) and iterate this.

Does x^* and λ^* converge to the expected solution?

Note: The path $x^*(\mu) = \text{argmin}_x F(x; \mu)$ (the optimum in dependence of μ) is called *central path*.