COMP4680/COMP8650: Advanced Topics in Machine Learning

Assignment #3: Convex Optimization and Duality

Due: 11:55pm on Sunday 4 September, 2022.
Submit solutions (showing all working) as a single PDF file or Jupyter notebook via Wattle.
Don't forget to include your name and student ID at the top of the submitted solutions.

1. **Rocket science (20 marks).** You have been tasked with designing the propulsion system for powering a rocket to launch Elon Musk, Jeff Bezos and Richard Branson to Mars. Your propulsion system runs on a combination of solid fuel and liquid oxygen. Each unit of solid fuel weighs 2kg, each unit of liquid oxygen weighs 1kg and the rocket has a maximum weight budget of 500kg for the oxygen and fuel.

   Your engineers tell you that to ignite the system needs at least 30 units of solid fuel, and for efficiency it needs twice as much liquid oxygen to solid fuel (per unit). However, for safe operation you cannot have more than seven times the amount of liquid oxygen to solid fuel (per unit).

   Your job is to choose how much solid fuel and how much liquid oxygen to put in the rocket.

   (a) Write down the set of constraints for your problem in standard form.
   (b) Assume that there may be some measurement error in the amount of oxygen and fuel that gets loaded into the rocket based on your design. Write down a reasonable optimization problem that will maximize the probability that the trip to Mars is successful (i.e., that no constraints are violated assuming that the measurement error is small).
   (c) Is your problem convex?
   (d) Draw the feasible region for your problem and argue for a good solution (you do not need to find the optimal solution).

2. **Optimality (10 marks).** Consider the following optimization problem

$$\begin{array}{ll} \text{minimize} & \frac{1}{2}x^T P x + q^T x + r \\ \text{subject to} & -1 \leq x_i \leq 1, \; i = 1, 2, 3 \end{array}$$

   where

$$P = \begin{bmatrix} 26 & 24 & -4 \\ 24 & 34 & 12 \\ -4 & 12 & 24 \end{bmatrix}, \quad q = \begin{bmatrix} -44 \\ -29 \\ 26 \end{bmatrix}, \quad r = 1.$$

   (a) Prove that $x^\star = (1, 0.5, -1)$ is optimal.
   (b) Are any of the constraints inactive at $x^\star$? If so, which one(s)?

3. **Simple Linear Programs (30 marks).** Give an explicit solution for each of the following linear programs.

   (a) Minimizing a linear function over an affine set.

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & Ax = b. \end{array}$$

(b) Minimizing a linear function over a rectangle.

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & l \preceq x \preceq u. \end{array}$$

(c) Minimizing a linear function over the probability simplex.

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & \mathbf{1}^T x = 1 \\ & x \succeq 0. \end{array}$$

What happens if the equality constraint is replaced by an inequality $\mathbf{1}^T x \leq 1$?

4. **Linear Program with n Constraints (10 marks).** Consider the linear program

$$\begin{array}{ll} \text{minimize} & c^T x + d \\ \text{subject to} & Ax \preceq b \end{array}$$

with A square and nonsingular. Show that the optimal value is given by

$$p^\star = \begin{cases} c^T A^{-1} b + d & \text{if } A^{-T} c \preceq 0 \\ -\infty & \text{otherwise.} \end{cases}$$

5. **Analytic Centering (15 marks).** Consider the primal problem for analytic centering,

$$\text{minimize} \quad -\sum_{i=1}^{m} \log(b_i - a_i^T x)$$

with optimisation variable $x \in \mathbb{R}^n$.

(a) What is the (implicit) domain of the objective function?

(b) Derive an equivalent problem by introducing a variables $y_i = b_i - a_i^T x$ for $i = 1, \ldots, m$.

(c) Write down the Lagrangian for your problem and derive the dual function.

(d) Using the results from (a)–(c) derive a dual problem for the above primal problem.

6. **Supporting hyperplane interpretation of the KKT conditions (15 marks).** Consider a convex problem with no equality constraints,

$$\begin{array}{ll} \text{minimize} & f_0(x) \\ \text{subject to} & f_i(x) \leq 0, \quad i = 1, \ldots, m. \end{array}$$

Assume that $x^\star \in \mathbb{R}^n$ and $\lambda^\star \in \mathbb{R}^m$ satisfy the KKT conditions

$$\begin{aligned} f_i(x^\star) \leq 0, &\quad i = 1, \ldots, m \\ \lambda_i^\star \geq 0, &\quad i = 1, \ldots, m \\ \lambda_i^\star f_i(x^\star) = 0, &\quad i = 1, \ldots, m \\ \nabla f_0(x^\star) + \sum_{i=1}^{m} \lambda_i^\star \nabla f_i(x^\star) = 0. & \end{aligned}$$

Show that

$$\nabla f_0(x^\star)^T (x - x^\star) \geq 0$$

for all feasible x.

Hint: Use the fact that the f_i are convex and $f_i(x) \leq 0$ for any feasible x to show that $f_i(x^\star) + \nabla f_i(x^\star)^T (x - x^\star) \leq 0$. Then combine this expression with complementary slackness and stationarity of the Lagrangian to arrive at the result.