Homework 6 for Math 173B - Winter 2025

Questions:

(0) Remember to write the names of all people you discussed this assignment with.

(0') You must submit all your computer codes as part of this assignment. In particular, for each question, your code must be presented as part of your answer.

(1) Consider the optimization function

$$f(x) = x_1^2 + 4x_2^2 \quad \text{s.t.} \quad \begin{cases} x_1 + 2x_2 & \geq 4 \\ x_1 & \geq 0 \\ x_2 & \geq 0 \end{cases}$$

a) Write the associated Lagrangian function.

b) Find the Lagrangian dual function.

c) Verify that the Lagrangian dual is concave and differentiable.

d) Solve the dual optimization problem (analytically).

e) Can you use the dual optimal solution to help solve the primal optimization solution? Why or why not?

2) Let $A \in \mathbb{R}^{m \times n}$, and suppose we we wish to solve the optimization problem

$$\min_{x \in \mathbb{R}^n} x^T x \quad \text{s.t.} \quad Ax = b$$

(a) Write (analytically) a dual ascent algorithm for solving this problem. Show your work.

(b) Write (analytically) a method of multipliers algorithm for solving this problem. Show your work.

(c) Write a computer program implementing your algorithm from part (a). In particular, let A be a random matrix with standard gaussian entries, with $m = 500, n = 1000$, and let $b = \mathbb{1}$, the all ones vector. Choose the step-size to be $\alpha = \frac{1}{\|A\|^2}$.

You should produce **two plots:**

- one plot showing the value of the cost function $(x^{(t)})^T x^{(t)}$ against the iteration number t.
- another plot showing $\|Ax^{(t)} - b\|$ (as a proxy for infeasibility), against the iteration t.
- it would improve visibility if your y-axis is in log-scale.

(d) Repeat part (c), but now implementing your algorithm from (b). *Keep the same matrix A and vector b as in part (c).* Set $\rho = 0.01$.

(e) Try different choices of α and ρ in your codes above. Compare the performance of dual ascent and the method of multipliers (by plotting $(x^{(t)})^T x^{(t)}$ from each method and $\|Ax^{(t)} - b\|$ from each method) on the same two plots, with the best choices of α and ρ that you could find. Comment on the performance. For example, which method seems to be more computationally efficient for this problem? Is the same method more efficient across all choices of ρ and α that you tried?