MS&E 111X & 211X
Introduction to Optimization (Accelerated)
Homework 4
Course Instructor: Yinyu Ye
Due Date: 11:59 pm Nov 30, 2021

Please submit your homework through Gradescope. If you haven't already been added to Gradescope, you can use the entry code **2RJNKV** to join. Please note: late homework will not be accepted. Each problem will be graded out of 10 points.
Some problems allow group work. Groups should be no larger than 4. If you decide to work together, provide the names of those you worked with.

Problem 1

For parts a)-c) below, label them as True or False. If true, provide a short reason; if false, provide reasoning or a counter example.

a) *True or False:* The simplex method (with cycle breaking rules) for a general linear program with n variables always converges to the optimal solution after a finite amount of steps and it takes at most a polynomial number of steps in n to converge.

b) Consider the following optimization problem:

$$\min f(x_1, x_2) = -\log(x_1 x_2) + x_1^2 + x_2 + 3(x_1 - x_2)^4.$$

True or False: The gradient descent procedure with line search on this problem is guaranteed to converge to the global minimum.

c) Consider the following optimization problem:

$$\min f(x_1, x_2) = (x_1 - 1)^2 + (x_2 - 2)^2.$$

True or False: Newton's method on this problem will converge to the global minimum faster than the gradient descent method provided that we use line search.

Problem 2

Solve the transportation problem described in pages 12-14, Lecture 10 using the simplex method. Specifically, write down the shadow prices, reduced costs, the termination test and the new BFS for each iteration. Clearly state your optimal solution at the end. [Hint: You can begin with the BFS provided in page 21, Lecture 10.]

Problem 3

Recall the logistic regression problem in Homework 2 Problem 6 whose objective is to minimize the function

$$f(\boldsymbol{x}, x_0) = \sum_i \log(1 + \exp(-\boldsymbol{a}_i^T \boldsymbol{x} - x_0)) + \sum_i \log(1 + \exp(\boldsymbol{b}_i^T \boldsymbol{x} + x_0))$$

with training data $\boldsymbol{a}_1 = (0; 0)$, $\boldsymbol{a}_2 = (1; 0)$, $\boldsymbol{a}_3 = (0; 1)$, $\boldsymbol{b}_1 = (0; 0)$, $\boldsymbol{b}_2 = (-1; 0)$, $\boldsymbol{b}_3 = (0; -1)$, where $\boldsymbol{a}_i$ are points we negatively label and $\boldsymbol{b}_i$ are points we positively label. We consider numerically solving this problem using gradient descent. You are free to use whatever programming language you prefer, but you cannot use any off-the-shelf implementation such as `scipy.optimize` in python. You may work in teams for this problem.

a) We fix step size as 0.1 and we stop iterating at the earliest round when the Euclidean norm of gradient drops below 10^{-3} or we reach 1000th round. Write code for fixed step size gradient descent for the problem in Homework 2 Problem 6. You are free to choose your starting point of $(\boldsymbol{x}^{(0)}, x_0^{(0)})$. After running the algorithm, please produce and report the following 2 plots:

   1) horizontal axis: the iterations k, vertical axis: Euclidean norm of $[\boldsymbol{x}; x_0]$
   2) horizontal axis: the iterations k, vertical axis: the value of negative log-likelihood function $f(\boldsymbol{x}^{(k)}, x_0^{(k)})$;

b) Modify your code to add the regularization proposed by Homework 2 Problem 6 Part (c), which is

$$\min \quad f(\boldsymbol{x}, x_0) + \mu ||\boldsymbol{x}||^2.$$

   Produce the same plots as Part (a) with $\mu_1 = 0.01$ and $\mu_2 = 1$ (i.e. please report 4 plots for this subsection, 2 plots for each regularization scale).

c) Compare the results in Parts (a) and (b) and summarize what effects of regularization you notice.

d) Attach/print your code for implementations of this problem.

Problem 4

Consider the following variant of the portfolio management quadratic program in Lecture 12 Slide 16:

$$\begin{aligned} \min \quad & x_1^2 + 2x_2^2 + 3x_3^2 - x_1x_2 - x_2x_3 - x_1x_3 - x_1 - 2x_2 - 3x_3 \\ \text{subject to} \quad & x_1 + x_2 + x_3 = 1 \end{aligned}$$

We consider numerically solving this problem using Newton's method. You are free to use whatever programming language you prefer, but you cannot use any off-the-shelf implementation such as `scipy.optimize` in python. Write code to perform Newton's method and report your final (x_1, x_2, x_3).

Problem 5

Consider the LP problem

$$\begin{aligned} \text{minimize} \quad & x_1 + x_2 & \\ \text{subject to} \quad & x_1 + x_2 + x_3 & = 1, \\ & (x_1, x_2, x_3) & \geq 0. \end{aligned}$$

a) Formulate the barrierred problem using a logarithmic barrier and derive the KKT optimality conditions for the barrierred problem. Resolve the optimal solution in terms of the barrier parameter $\mu : x_1^*(\mu)$, $x_2^*(\mu)$ and $x_3^*(\mu)$.

b) The trajectory defined by $(x_1^*, x_2^*, x_3^*)(\mu)$ is known as the central path to the optimum of the problem. Find the location of 3 points on this path by evaluating the solution to the barrierred problem for $\mu \to \infty$, $\mu = 1$, and $\mu \to 0$. Show that the solution obtained with $\mu \to 0$ is the optimal solution of the original problem using LP duality.