E0 230
Computational Methods in Optimization
Assignment 4

November 24, 2021

Instructions:

- This is an **individual assignment**, and **all work submitted must be your own**!
- Attempt all questions
- You have until **Midnight, December 2nd, 2021** to submit your answers.
- All code must be submitted.
- All long answers **must be submitted in a single PDF.** All numerical problems should be answered to 3 decimal places.
- Your code, your figures, and your PDF must be submitted in a single zip file, which should be called **student_name_cmo21assn1.zip**. Figures should be in .jpg or .png format, or should be embedded and neatly labeled in your report. **Label each separate code and image file with Question_SubQuestion number.**
- Choose the files required for your setup from the concerned directory in the zip file

1. (5 points) Consider the problem of learning a linear classifier with a zero bias term below,

$$\begin{aligned} \arg\min_{w} \ & \frac{1}{2}\|w\|^2 \\ \text{such that } \ & y_i(w^T x_i) \geq 1 \text{ for all} i, \end{aligned}$$

where $y_i \in \{\pm 1\}$, $x_i \in \mathbb{R}^2$ are given data points. We have provided data $\{(x_i, y_i)\}_{i=1}^{10}$ in `a4.csv`.

   (a) (1 point) Consider the problem stated above. We will use active set methods to solve this problem. Initialize the problem with $w0 = [1, -1]$. What is the initial working set? Furthermore, plot the data with red 'x's for those points with $y_i = 1$ and blue 'circles' for those points with $y_i = -1$. Plot w_0 as well.

   (b) (2 points) What are the feasible direction, working set, and next iterate after the first step of the active set method? Add w_1 to the plot you made in the previous part.

   (c) (2 points) Run the active set algorithm for one more step. Detail your results here, and add w_2 to the plot you obtained in the previous part.

2. (15 points) In this problem, we use linear programming to solve a path planning problem for a wheeled robot. The position of the robot updates as follows:

$$\begin{aligned} x_{t+1} &= x_t + v_t \cos(\theta_t) \\ y_{t+1} &= y_t + v_t \sin(\theta_t). \end{aligned}$$

The heading angles θ_t are given to you for all $t \in \{0, ..., T-1\}$ where T is constant. Furthermore, $\theta_{T-1} = \theta_f$, and the sequence $\{\theta_t\}_{t=0}^{T}$ is monotonically decreasing; that is, $\theta_{t+1} \leq \theta_t$ for all t. The goal is to find a sequence of positive speeds $\{v_t\}_{t=1}^{T}$ in such a way that, given (x_f, y_f), we have $(x_T, y_T) = (x_f, y_f)$, and such that some linear cost function is minimized.

   (a) (1 point) Suppose we take the starting point of the robot to be $(0, 0)$ with a heading angle of $\pi/2$. Our goal is to reach x_f, y_f with a heading angle of $\theta_f \in [0, \pi/2)$ in T steps, and we want to take the *longest* path to get there. Last, suppose we require $v_t \geq v_{\min} > 0$ for all t - $v_{\min}$ is a constant that is given to you. State this problem as a linear program in standard form:

$$\begin{aligned} \arg\min \ & c^\top x \\ \text{s.t. } \ & Ax = b \\ & x \geq 0 \end{aligned}$$

   with only two equality constraints. (Hint: recall that *distance* = *speed* $\times$ *time*, and you can assume each unit of time to be 1)

   (b) (3 points) Suppose we take the starting point of the robot to be $(0, 0)$ with a heading angle of $\pi/2$, and choose the final position to be $(3, 4)$ with a heading angle of 0. Furthermore, $T = 100$ and $v_{\min} = 0.01$. The sequence $\{\theta_t\}_{t=1}^{T}$ is given to you in `headingangles.csv`. Find the sequence of speeds that yields the longest path. Plot the trajectory - the sequence of pairs $\{(x_t, y_t)\}$. Repeat this experiment for $\theta_f = \pi/2$. What happens if we set $v_{\min} = 0$? What happens if we set $v_{\min} = 1$? Solve the problem for these two cases, and plot these trajectories as well. You may use `linprog`, `SeDuMi`, `cvx` or `cvxpy` to do so. We strongly encourage you to solve and plot your results for different values of x_f, y_f, and θ_f.

   (c) (5 points) Suppose are given x_f, y_f, and $\theta_f = 0$. Show that if

$$v_{\min} \leq \frac{x_f + y_f}{T\sqrt{2}},$$

then the problem you derived in part (a) is feasible. (Hint: How would you use Farkas' Lemma to tell you about the feasibility of a linear program in standard form? Also, note that for $z_1, z_2 > 0$ and positive constants a, b, we have $az_1 + bz_2 \geq (a+b)\min\{z_1, z_2\}$.)

(d) (6 points) We now wish to apply the simplex method toward solving the problem you derived in part (a). Assume that we have the problem data in part (c) - that is, $\theta_f = 0$. Furthermore, suppose $v_{\min} \leq \frac{x_f + y_f}{T\sqrt{2}}$. Verify that $\{1, T\}$ is a BFS. Is this BFS optimal? Prove your answer. Furthermore, if this BFS is optimal, derive a closed form expression for v_t in terms of $\{\theta_t\}_{t=1}^{T}$, x_f, y_f and $v_{\min}$.

3. (5 points) In this problem, we are given a linear function $f(x) = \langle w^0, x\rangle$, which is a learned linear regression model. We obtain new data, contained in the matrices A, b, and we wish to update our model by finding weights that are at most a distance of r from w^0. To do so, we attempt to solve the following optimization problem.

$$\begin{aligned} \arg\min \ & \frac{1}{2}\|Aw - b\|^2 \\ \text{s.t.} \ & \|w - w^0\|^2 \leq r^2 \end{aligned}$$

(a) (2 points) State the KKT conditions for this problem.

(b) (3 points) Can you find a closed form expression for w in terms of $A,\ b,$ and w^0, when the constraint is inactive, and state that condition as well?

4. (5 points) Consider the following minimization problem:

$$\begin{aligned} \arg\min \ & x^T A_0 x + 2b_0^T x + c_0 \\ \text{s.t.} \ & x^T A_i x + 2b_i^T x + c_i \leq 0, \quad i = 1, 2, \ldots, m, \end{aligned}$$

where A_i is $n \times n$ positive semi-definite matrix, $b_i \epsilon \mathbb{R}^n,\ c_i \epsilon \mathbb{R},\ i = 0, 1, \ldots, m$. Construct the dual of the problem assuming A_0 is positive definite.

5. (20 points) In this problem we will compute the projection of a point on a polytope using gradient projection algorithm.

(a) (2 points) State the dual of the problem of projection of a point y on the polytope $S = \{x \,\epsilon\, \mathbb{R}^n : Ax \leq b\}$ where $A \,\epsilon\, \mathbb{R}^{m\times n},\ b \,\epsilon\, \mathbb{R}^m$.

(b) (3 points) Let us assume strong duality holds, then solving the dual will give us the projection. We can use gradient projection to solve the above dual problem. Derive the general iteration step of the algorithm by taking stepsize $\frac{1}{L}$, where L is the Lipschitz constant of the objective function obtained.

(c) (5 points) Implement the above algorithm, for the following polytope $S = \{(x, y, z) : x + y + z \leq$ [*Last digit of your SR Number*]$, x \geq 0, y \geq z\}$. State the optimal projection point obtained for $(3, -1, 2)$ after 100 iterations.

(d) (5 points) Plot sequence of first 10 points obtained when we try to find the projection of point $(2, 1)$ on the polytope defined by $S = \{(x, y) : x + y \leq 1, x \geq 0, y \geq 0\}$. Repeat the exercise with stepsize $\frac{2}{L}$. Report the plots for the same. By looking at the plot, comment on which stepsize is better and why.

(e) (5 points) Suppose $S = \{x, y, z \,\epsilon\, \mathbb{R} : x + 3y = 1, 2y + z = -1\}$.Write the algorithm if we want to find the projection in an iterative fashion using the gradient projection method, and hence compute the projection of the origin on this set. Verify your solution by finding the analytical solution for the same.