Assignment 3

E0-230 Computational Methods of Optimization

Instructions

- This is an individual assignment, all the work must be your own.
- No copying or sharing of code is allowed. Any form of academic dishonesty will result in ZERO marks for this assignment.
- The deadline for submission is 20$^{\text{th}}$ November, 2024.
- Submissions will be accepted till 23$^{\text{rd}}$ November, 2024, with a penalty of 5 points per day.
- All the code must be written in Python, along with appropriate comments. The code needs to be submitted for any problem that requires any implementation or programming. Follow the naming convention suggested in the question and feel free to pass additional arguments if necessary with mention. These codes can be reused in subsequent assignments.
- The outputs of the coding problems and other analytical solutions should be reported in a single LaTeX-generated PDF file.
- Submit two separate files, a SRNumber.pdf and a SRNumber.py, directly without creating a zip or rar. Not adhering to this will lead to an additional penalty of 5 points.

1 Systems of Linear Equations (15 points)

Let us revisit the case where solutions to a linear system of equations, $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$, need to be found. In the last assignment, this was cast as the convex optimisation problem given by:

$$\min_{\boldsymbol{x} \in \mathbb{R}^n} ||\boldsymbol{A}\boldsymbol{x} - \boldsymbol{b}||^2.$$

While this method has its own merits (such as finding a point which is "closest" to being a solution if the system is indeterminate, it does not find solutions with a particular property. In this assignment, we will assume that there exists at least one solution, and try to find the particular solution $\boldsymbol{x}^*$ that has the least norm. For the purpose of this question, let:

$$\boldsymbol{A} = \begin{bmatrix} 2 & -4 & 2 & -14 \\ -1 & 2 & -2 & 11 \\ -1 & 2 & -1 & 7 \end{bmatrix} \quad \text{and} \quad \boldsymbol{b} = \begin{bmatrix} 10 \\ -6 \\ -5 \end{bmatrix}.$$

1. **(2 points)** Show that the given system of equations has an infinite number of solutions.
2. **(2 points)** Express the problem of finding the least-norm solution as an optimisation problem (`ConvProb`) with convex constraints and a strongly convex objective function. Show that the constraints and the objective are convex and strongly convex, respectively.
3. **(3 points)** Use the KKT conditions to solve `ConvProb`, and arrive at an expression for $\boldsymbol{x}^*$, and show the intermediate steps. Write code to evaluate the expression and report $\boldsymbol{x}^*$.
4. **(3 points)** Derive a projection operator for the constraint set of `ConvProb`.
5. **(5 points)** Use the derived projection operator and implement projected gradient descent to solve `ConvProb`. Test with different step-sizes and plot $||\boldsymbol{x}^{(t)} - \boldsymbol{x}^*||$ at each iteration t.

2 Support Vector Machines (15 points)

Support vector machines (SVMs) are among the most widely used techniques for classifying data, and are very well studied. The SVM is a linear classifier; that is, we aim to find a function $y = f(\boldsymbol{x}) = \text{sign}\left(\boldsymbol{w}^\top \boldsymbol{x} + b\right)$. We are given data of the form $\mathcal{D} = \{\boldsymbol{x}_i, y_i\}_N$, where the pair $(\boldsymbol{x}_i, y_i) \in \mathbb{R}^n \times \{-1, 1\}$. To learn a support vector machine, we need to solve the following convex optimization problem.

$$\begin{aligned} \boldsymbol{w}^*, \boldsymbol{b}^* = \text{argmin} \ & \frac{1}{2}\|\boldsymbol{w}\|_2^2 \\ \text{s.t.} \ & y_i\left(\boldsymbol{w}^\top \boldsymbol{x}_i + b\right) \geq 1 \quad \text{for } i = 1, \cdots, N. \end{aligned}$$

1. **(2 points)** Use either CVXPY (python) to solve the primal for the data given in "Data.csv" and "Labels.csv", which can be found under the Files tab. What is the value of the primal objective function?
2. **(3 points)** Show that the dual function is of the form

   $$g(\boldsymbol{\Lambda}) = \boldsymbol{\Lambda}^\top \boldsymbol{b} + \frac{1}{2}\boldsymbol{\Lambda}^\top \boldsymbol{A}\boldsymbol{\Lambda},$$

   $\boldsymbol{\Lambda} = (\lambda_1, \cdots, \lambda_k)$. What are $\boldsymbol{A}_{ij}$ and $\boldsymbol{b}_i$? What is k?

3. **(3 points)** Show that

$$\sum_{i:y_i=1} \lambda_i = \sum_{i:y_i=-1} \lambda_i = \gamma.$$

What is the value of γ for the given problem?

4. **(3 points)** Write a program to solve the dual problem. What is the value of the dual objective at optimality?

5. **(2 points)** Which of the primal constraints are active? Describe your answer.

6. **(2 points)** Plot the following:

   (a) The line $\boldsymbol{w}^\top \boldsymbol{x} + b = 0$

   (b) The data points provided. For $y = 1$, mark the points with circles, and for $y = -1$, mark the points with squares.

   (c) The points corresponding to active constraints. Mark these points in red.