Optimization and Machine Learning
Assignment 1

Due Date: 22 March 2023, 5 pm

Instructions

- Please read each question carefully
- Note that the last question is only for groups with four members
- After completing your assignment compress all your files into a **single** zip file and name the file with your group's name (*e.g.*, `OverAchievers.zip`)

Points for The Questions

	3-member groups	**4-member groups**
Question 1	10 pts.	7.5 pts.
Question 2	10 pts.	10 pts.
Question 3	-	2.5 pts.

Question 1

Recall that the coordinate descent algorithm can also be applied with weighted samples (see slide 31 of the first lecture notes). Show the steps to obtain the following analytical solution of the one-dimensional subproblem for $j \neq 0$:

$$\bar{\beta}_j = \begin{cases} \frac{\hat{\beta}_j - \lambda\alpha}{\sum_{i=1}^n w_i x_{ij}^2 + \lambda(1-\alpha)}, & \text{if } \hat{\beta}_j > 0 \text{ and } \lambda\alpha < \hat{\beta}_j; \\ \frac{\hat{\beta}_j + \lambda\alpha}{\sum_{i=1}^n w_i x_{ij}^2 + \lambda(1-\alpha)}, & \text{if } \hat{\beta}_j < 0 \text{ and } \lambda\alpha < -\hat{\beta}_j; \\ 0, & \text{otherwise.} \end{cases}$$

Here,

$$\hat{\beta}_j = \sum_{i=1}^n w_i x_{ij}(y_i - \bar{\beta}_0 - \sum_{k \neq j} x_{ik}\bar{\beta}_k).$$

What will you turn in?

A PDF file showing the steps that you followed to obtain the solution. The name of the file should be your group's name (*e.g.*, `OverAchievers.pdf`). Add the renamed file into the final compressed file that you will submit.

Evaluation. This question will be evaluated in terms of correctness and clarity of your steps.

Question 2

Implement the coordinate descent algorithm (page 30 of the first lecture notes) in Python to solve the Elastic Net problem. The input and output of your Python function are as follows:

- Input
  - Training data, $\{(x_i, y_i) : i = 1, \ldots, n\}$, $x_i \in \mathbb{R}^p$ (standardized), $y_i \in \mathbb{R}$ (centered)
  - Hyperparameters, $\lambda > 0$ and $\alpha \in [0, 1]$
  - Maximum number of iterations (default value of 1000), *i.e.*, `max_iter=1000`
- Output
  - Weight vector, $\bar{\beta}$
  - Duality gap, $\eta(\bar{\beta}, \bar{v})$ where $\bar{v}$ is the dual feasible vector

Here are some remarks and requirements:

- Recall that the duality gap calculation requires the trick on page 35 of the first lecture.
- Please keep in mind that the lecture notes may skip or overlook some implementation details. If that turns out to be the case, your creativity, persistence, and skills are required.
- Do not forget to fully comment your Python function.

What will you turn in?

Add your function to the designated place in the `test_EN_.py` template script and rename the file after appending your group's name as a **single word** (*e.g.*, `test_EN_OverAchievers.py`). Add the renamed file into the final compressed file that you will submit.

Evaluation. The results of your implementation will be tested against the `sklearn` Elastic Net package. The template script `test_EN_.py` also shows our testing setup as well as the datasets. The clarity and correctness of your implementation along with the comments will determine your score.

Question 3 - Only for groups with four members

Write down the dual of the linear programming model on page 23 of the **second** lecture notes.

What will you turn in?

A PDF file showing the steps that you followed to obtain the dual. The name of the file should be your group's name (**single word**) appended with "Extra" (*e.g.* `OverAchieversExtra.pdf`). Add the renamed file into the final compressed file that you will submit.

Evaluation. This question will be evaluated in terms of correctness and clarity of your steps.