Instructions: (Please read carefully and follow them!)

Try to solve all problems on your own. If you have difficulties, ask the instructor or TAs.

In this session, we will apply the methods we have developed in the previous labs, to solve a practical problem. The scalability analysis performed in previous labs will be carried out in this lab as well.

The implementation of the optimization algorithms in this lab will involve extensive use of the `numpy` Python package. It would be useful for you to get to know some of the functionalities of `numpy` package. For details on `numpy` Python package, please consult https://numpy.org/doc/stable/index.html

For plotting purposes, please use `matplotlib.pyplot` package. You can find examples in the site https://matplotlib.org/examples/.

Please follow the instructions given below to prepare your solution notebooks:

- Please use different notebooks for solving different Exercise problems.
- The notebook name for Exercise 1 should be `YOURROLLNUMBER_IE684_Lab04_Ex1.ipynb`.
- Similarly, the notebook name for Exercise 2 should be `YOURROLLNUMBER_IE684_Lab04_Ex2.ipynb`, etc and so on.

There are only 2 exercises in this lab. Try to solve all the problems on your own. If you have difficulties, ask the Instructors or TAs.

You can either print the answers using `print` command in your code or you can write the text in a separate text tab. To add text in your notebook, click `+Text`. Some questions require you to provide proper explanations; for such questions, write proper explanations in a text tab. Some questions require the answers to be written in LaTeX notation. (**Write the comments and observations with appropriate equations in LaTeX only.**) Some questions require plotting certain graphs. Please make sure that the plots are present in the submitted notebooks.

After completing this lab's exercises, click File → Download .ipynb and save your files to your local laptop/desktop. Create a folder with name `YOURROLLNUMBER_IE684_Lab04` and copy your `.ipynb` files to the folder. Then zip the folder to create `YOURROLLNUMBER_IE684_Lab04.zip`. Then upload only the `.zip` file to Moodle. **There will be some penalty for students who do not follow the proper naming conventions in their submissions**.

Please check the **submission deadline announced in moodle.**

This lab aims to provide you with applications of the methods learned in previous labs. Here we will try to solve a very common OLS problem used in machine learning for regression tasks using all the previously taught optimization methods. Hope it will be fun! Let's code!

Exercise 1 (15 marks) *Suppose that y is a noisy version of $A\bar{x}$. We will now try to estimate $\bar{x}$ assuming that we are given y and A. One possible approach is to solve the following problem:*

$$\text{minimize}_{\mathbf{x}} \quad f(\mathbf{x}) = \frac{1}{2}\|A\mathbf{x} - \mathbf{y}\|_2^2$$

The loss term $\|A\mathbf{x}-\mathbf{y}\|_2^2$ is called the ordinary least squares (OLS) loss, and the problem is called the OLS Regression problem. This is numbered as Problem (1). Please follow this Google Colab Link for data preparation and some preliminary exercise regarding this problem as a reference to move ahead with this exercise.

1. *Follow the link and solve it.*
2. *Follow the link and solve it.*
3. *With a starting point $\mathbf{x}_0 = \begin{bmatrix}0 & 0 & \dots & 0\end{bmatrix}^\top \in \mathbb{R}^{10}$, solve problem (1) using Newton's method implemented with backtracking line search (use $\alpha_0 = 0.99$, $\rho = 0.5$, $\gamma = 0.5$ for backtracking line search, and $\tau = 10^{-4}$). Comment on difficulties (if any) you face when computing the inverse of Hessian (recall that you need to use an appropriate Python function to compute the inverse of the Hessian). If you face difficulty in computing inverse of Hessian, try to think of some remedy so that you can avoid the issue.*
   - *Let $\mathbf{x}^*$ be the final optimal solution provided by your algorithm. Report the values of $\mathbf{x}^*$ and $\bar{\mathbf{x}}$, and discuss the observations.*
   - *Plot the values $\log(\|\mathbf{x}^k - \mathbf{x}^*\|_2)$ against iterations $k = 0, 1, 2, \dots$.*
   - *Prepare a different plot for plotting $\log(|f(\mathbf{x}^k) - f(\mathbf{x}^*)|)$ obtained from Newton's method against the iterations.*
   - *Comment on the convergence rates of the iterates and the objective function values.*
4. *With a starting point $\mathbf{x}_0 = \begin{bmatrix}0 & 0 & \dots & 0\end{bmatrix}^\top \in \mathbb{R}^{10}$, solve problem (1) using BFGS method implemented with backtracking line search (use $\alpha_0 = 0.99$, $\rho = 0.5$, $\gamma = 0.5$ for backtracking line search, and $\tau = 10^{-4}$).*
   - *Let $\mathbf{x}^*$ be the final optimal solution provided by BFGS algorithm. Report the values of $\mathbf{x}^*$ and $\bar{\mathbf{x}}$, and discuss the observations.*
   - *Plot the values $\log(\|\mathbf{x}^k - \mathbf{x}^*\|_2)$ against iterations $k = 0, 1, 2, \dots$.*
   - *Prepare a different plot for plotting $\log(|f(\mathbf{x}^k) - f(\mathbf{x}^*)|)$ obtained from BFGS method against the iterations.*
   - *Comment on the convergence rates of the iterates and the objective function values.*
5. *Compare and contrast the results obtained by Newton's method and BFGS method and comment on the time taken by both the methods.*

Exercise 2 (60 marks) ***Regularized least squares loss minimization***

1. *Let us now introduce the following regularized problem (with $\lambda > 0$):*

$$\min_x f_\lambda(x) = \frac{\lambda}{2}x^T x + \frac{1}{2}\|A\mathbf{x} - \mathbf{y}\|_2^2 \tag{1}$$

   Comment on the significance of the newly added regularizer term $\frac{\lambda}{2}x^T x$, when compared to problem (1).

2. *Write Python functions to compute the function value, gradient, and Hessian of* f_λ

3. *For* $\lambda \in \{10^{-3}, 10^{-2}, 10^{-1}, 1\}$*, perform the following: with a starting of* $x^0 = [0, 0, \ldots, 0]^T \in \mathbf{R}^{10}$*, solve the equation 1 using Newton and BFGS methods with backtracking line search (use* $\alpha^0 = 0.99$*,* $\rho = 0.5$*,* $\gamma = 0.5$ *for backtracking line search and tau* $= 10^{-5}$*).*

4. *For Newton's method prepare the following plots and discuss relevant observations:*

   - *Prepare a single plot where you depict the values* $log(||x^k - x^*||_2)$ *against iterations* $k = 0, 1, 2, ...$*, for each value of* λ *(use different colors for different* λ *values; if necessary, add zoomed versions of the plots to depict the behavior clearly, and use appropriate legends in your plots). Comment on the convergence rates of the iterates for each value of* λ*.*
   - *Prepare a different plot for plotting* $log(||f(x^k) - f(x^*)||_2)$ *against the iterations, for each value of* λ *(use different colors for different* λ *value; if necessary, add zoomed versions of the plots to depict the behavior clearly and use appropriate legend in your plots). Comment on the convergence rates of the objective function values.*

5. *For BFGS method prepare the following plots and discuss the relevant observations:*

   - *Prepare a single plot where you depict the values* $log(||x^k - x^*||_2)$ *against iterations* $k = 0, 1, 2, ...$*, for each value of* λ *(use different colors for different* λ *values; if necessary, add zoomed versions of the plots to depict the behavior clearly, and use appropriate legends in your plots). Comment on the convergence rates of the iterates for each value of* λ*.*
   - *Prepare a different plot for plotting* $log(||f(x^k) - f(x^*)||_2)$ *against the iterations, for each value of* λ *(use different colors for different* λ *value; if necessary, add zoomed versions of the plots to depict the behavior clearly and use appropriate legend in your plots). Comment on the convergence rates of the objective function values.*

6. *Compare and contrast the results obtained by Newton's method and BFGS method and comment on the time taken by both methods for each value of* λ*.*