Question 1. (20 points)

For each of the following functions determine whether it is convex or not.

1- $F(x) = \frac{1}{x}\int_0^x f(t)\, dt$ where $f: \mathbb{R} \to \mathbb{R}$ is convex.
2- $f_\theta(x) = \theta^{-1}x^\theta - \theta^{-1}$ with $\mathbf{dom}\, f_\theta = \mathbb{R}_+$ and $0 < \theta \leq 1$.
3- $f(x_1, x_2) = \frac{x_1^2}{x_2}$ and $x_2 > 0$
4- $f(x_1, x_2, x_3) = -e^{-x_1} + x_2 x_3^2$

Question 2. (50 points)

Consider the likelihood function

$$L(k, c; y_1, y_2, \dots, y_n) = \prod_{i=1}^{n} \frac{k\, c\, y^{c-1}}{(1 + {y_i}^c)^{k+1}}$$

Where c and k are positive numbers, and y_i can be found in "Question2.csv".

a) Write down the log-likelihood function.
b) Write down the corresponding maximum likelihood formulation.
c) Derive the gradient and Hessian of the log-likelihood function.
d) The data file ('Question2.csv') contains 2000 observations. Report your estimated $\hat{k}_{MLE}$, and $\hat{c}_{MLE}$. Plot the value of the parameters $\hat{k}_{MLE}$, and $\hat{c}_{MLE}$ versus the number of iterations. Plot the value of the log-likelihood function versus the number of iterations.

- Gradient descent (Exact line search)
- Accelerated gradient descent (Exact line search)
- Stochastic gradient descent (Exact line search)
- Newton's method (Exact line search)

Question 3. (30 points)

Minimize the following loss functions using Gauss-Newton method.

$$L(\beta) = \min_{\beta_1, \beta_2} \sum_{i=1}^{n} \left(y_i - \frac{y_1\, \beta_1}{y_1 + (\beta_1 - y_1) exp(-\beta_2 d_i)} \right)^2$$

"Question3.csv" contains d_i(the first column) and y_i (the second column). In this question, our goal is to estimate parameters $\hat{\beta} = \begin{bmatrix} \hat{\beta}_1 \\ \hat{\beta}_2 \end{bmatrix}$.

- Decompose $L(\beta)$ such that $L(\beta) = g(\beta)^T g(\beta)$.
- Derive the Jacobian $J_{g(\beta)}$.
- Estimate parameters $\hat{\beta}_1$and $\hat{\beta}_2$ using Gauss-Newton method (adaptive step size).
- Report $\hat{\beta}_1$and $\hat{\beta}_2$.

- Plot y and $\frac{y_1 \beta_1}{y_1+(\beta_1-y_1)exp(-\beta_2 d_i)}$.