Homework 2

Elaheh Torkashvand[a,1]

[a] *Department of Applied and Computational Mathematics and Statistics, University of Notre Dame, Notre Dame, US, 46556*

There are many approaches available to find a root of a function computationally. In this Homework, the Newton-Raphson method is analyzed, see the R code for the function.

The **Newton method** finds the solution of $f(x_0) = 0$ using the Taylor expansion, this scenario might happen in optimization processes. The first-order Taylor expansion of f under the assumptions of existence and continuity of $f'(x) = \frac{d}{dx}f(x)$ with $f''(x_0^{(k)}) \neq 0$ is given by

$$f(x) \approx f\left(x_0^{(k)}\right) + \left(x - x_0^{(k)}\right) f'(x_0^{(k)}). \tag{0.1}$$

For $x = x_0^{(k+1)}$ and also assuming $f(x_0^{(k+1)}) \approx 0$, the Newton iteration method follows

$$x_0^{(k+1)} = x_0^{(k)} - \frac{f(x_0^{(k)})}{f'(x_0^{(k)})}, \quad \text{for} \quad k \geq 0. \tag{0.2}$$

The iteration procedure stops if and only if

$$\max_{1 \leq i \leq n} |x_i^{(k)} - x_i^{(k-1)}| < \epsilon, \text{ or,} \tag{0.3}$$

$$|f(x_i^{(k)})| \leq \epsilon, \tag{0.4}$$

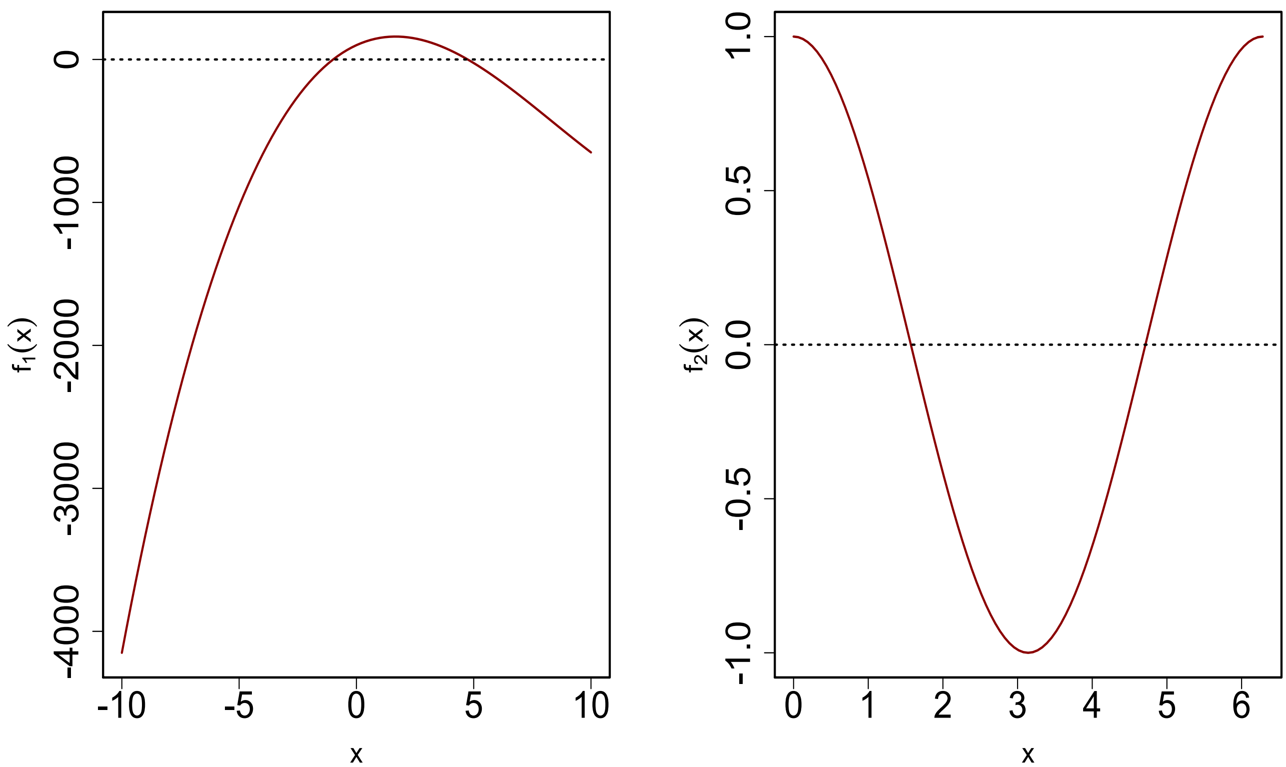
for an arbitrary small value of ϵ, say $\epsilon = 10^{-5}$. Conditions (0.3) or (0.4) imply the distance between two subsequent solution of a problem is negligible. Looking at the diagrams, we could detect whether the $x_0^{(k+1)}$ is the solution or not.

There are cases where the Newton method takes us to nowhere and we need to implement another methodology, bisection method, fixed point iteration, or, etc., to find the solution. For now, this is beyond the scope of our class.

Question 1. Write a program module to implement the Newton method to find the root of given functions, which is input together with values that bracket a root, and an epsilon as the stopping criterion. Your program should check that the two starting values are legitimate, in other words, try the method considering two different starting values.

[1]Etorkash@nd.edu

- Consider $f_1(x) = x^3 - 25x^2 + 75x + 100$, for $x \in [-10, 10]$.
- Also, $f_2(x) = \cos(x)$, for $x \in [0, 2\pi]$.



1. Draw the diagram of each function.
2. Graphically find the roots considering a horizontal line at $X = 0$.
3. Put the solution on each curve, are they valid?

Question 2. Mathematically, show that the k-nearest-neighbour method can be seen as a Kernel estimator with the following data driven metric

$$K_k(x, x_0) = I\left(\|x - x_0\| \le \|x_{(k)} - x_0\|\right), \tag{0.5}$$

where $x_{(k)}$ is the training observations ranked k'th in distance from x_0 and

$$I(t) = \begin{cases} 1, & t \in S, \\ 0, & \text{otherwise}. \end{cases} \tag{0.6}$$

The **Method of Moments** is one of the popular statistical methods to estimate parameters of a parametric family with finite expectation; i.e.

$$|E\left(X^j\right)| < \infty, \tag{0.7}$$

where $j \geq 1$. Equation (0.7) implies finite values for $E(X), E(X^2), \ldots, E(X^{j-1})$.

The procedure works for $i = 1, \ldots, j$ as solving the following system of equations

$$E(X^i) = M_i', \quad \text{for} \quad i = 1, \ldots, j, \tag{0.8}$$

where $M_i' = \frac{1}{n}\sum_{\ell=1}^{N} X_\ell^i$ for $X_1, X_2, \ldots, X_N$ as a sample obtained from the parametric distribution f_Θ.

Question 3. (a). Consider the beta density with parameters $\alpha > 0$ and $\beta > 0$, Beta(α, β). Find the method of moments estimate of α and β, $\widehat{\alpha}$ and $\widehat{\beta}$, where $X_1, X_2, \ldots, X_N$ is a set of observations from this density. Consider $j = 2$ in Equation (0.8). Can you have a closed form expression of the solution? If not, the Newton-Raphson method for a system of equations is a possible solution, this is beyond the scope of our class.

(b). Let $N = 100$, $\alpha = 5$, and $\beta = 5$. Generate $R = 500$ sets of data with length N from the corresponding beta distribution. Estimate the mean squared error (mse), $\frac{1}{R}\sum_{r=1}^{R}(\widehat{\alpha}_r - \alpha)^2$, bias, $\frac{1}{R}\sum_{r=1}^{R}(\widehat{\alpha}_r - \alpha)$, and variance, mse$(\alpha)$-bias$^2(\alpha)$, assuming $\beta = 5$ is known.

Question 4. Show that the sample mean $\bar{X}$ and sample standard deviation, s^2, are the method of moments estimate of μ and σ^2 for an arbitrary distribution with $E(X^2) < \infty$.

Suppose $X_1, \ldots, X_N$ is a random sample from the population with pdf $f(x;\theta)$. Define the likelihood function as

$$L(\theta) = \prod_{i=1}^{N} f(x_i;\theta), \quad \theta \in \boldsymbol{\Theta}. \tag{0.9}$$

The log likelihood function is defined accordingly as

$$l(\theta) = \log(L(\theta)), \quad \theta \in \boldsymbol{\Theta}.$$

The maximum likelihood estimate (MLE) of θ, $\hat{\theta}_N(x_1, \ldots, x_N)$, is the value which maximizes $L(\theta)$. In other words, $\hat{\theta}_N$ finds a value from the parameter space such that $x_1, \ldots, x_N$ has a highest chance of occurrence under it. The ML estimator is defined by $\hat{\theta}_N(X_1, \ldots, X_N)$.

Question 5. Repeat **Question 3** for the MLE. Which method do you prefer at the end based on different measures of accuracy? Explain your answer.

Question 6. Consider the gamma density, $\Gamma(\alpha = 5, \beta)$ where N observations are randomly selected; i.e. $X_1, \ldots, X_N$. Find the MM and ML estimate of β.

For purpose of simulation and comparison between the MME and MLE, assume $\beta = 10$. Find different measures of accuracy, MSE, bias, and variance and then, choose between MME and MLE.

Question 7. Go to Scenario 2. on page 13 of ESLII. Suppose all distributions of the mixture model have the same weight, 0.1. Use the codes of Lecture Notes 1 to generate from this mixture distribution while their mean is generated from $N(0,2)$ and their standard deviation is obtained from $\Gamma(2,8)$. Apply the K-nearest neighbour method and classification accordingly.