Homework 1

Abigail Wright

STA-4364

Problem 1

$$f(x) = 9x - 4\ln(x-7) \text{ for } x > 7$$

1. Exact Formula for the Newton Iterate

The Newton iterate formula is:

$$x_{n+1} = x_n - \frac{f'(x_n)}{f''(x_n)}$$

where:

- x_{n+1} is the next iterate,
- x_n is the current iterate,
- $f'(x_n)$ is the first derivative of the function at x_n,
- $f''(x_n)$ is the second derivative of the function at x_n.

2. First and Second Derivatives of f

The first derivative of f is:

$$f'(x) = \frac{d}{dx}\left(9x - 4\ln(x-7)\right) = 9 - \frac{4}{x-7}$$

The second derivative of f is:

$$f''(x) = \frac{d}{dx}\left(9 - \frac{4}{x-7}\right) = \frac{4}{(x-7)^2}$$

3. Compute Five Iterations of Newton's Method Starting from $x_0 = 7.80$

Iteration 1:

$$f'(7.8) = 9 - \frac{4}{7.8 - 7} = 9 - \frac{4}{0.8} = 4$$
$$f''(7.8) = \frac{4}{(7.8 - 7)^2} = 6.25$$
$$x_1 = 7.8 - \frac{4}{6.25} = 7.16$$

Iteration 2:

$$f'(7.16) = 9 - \frac{4}{7.16 - 7} = -16$$
$$f''(7.16) = \frac{4}{(7.16 - 7)^2} = 156.25$$
$$x_2 = 7.16 - \frac{-16}{156.25} = 7.2624$$

Iteration 3:

$$f'(7.2624) = 9 - \frac{4}{7.2624 - 7} = -6.2243902439$$
$$f''(7.2624) = \frac{4}{(7.2624 - 7)^2} = 58.09414039$$
$$x_3 = 7.2624 - \frac{-6.2243902439}{58.09414039} = 7.36987904$$

Iteration 4:

$$f'(7.36987904) = 9 - \frac{4}{7.36987904 - 7} = -1.824175215$$
$$f''(7.36987904) = \frac{4}{(7.36987904 - 7)^2} = 29.29069227$$
$$x_4 = 7.36987904 - \frac{-1.824175215}{29.29069227} = 7.4319344$$

Iteration 5:

$$f'(7.4319344) = 9 - \frac{4}{7.4319344 - 7} = -0.2606655085$$
$$f''(7.4319344) = \frac{4}{(7.4319344 - 7)^2} = 21.43998141$$
$$x_5 = 7.4319344 - \frac{-0.2606655085}{21.43998141} = 7.4444444$$

4. Verify Empirically that Newton's Method Will Converge to the Optimal Solution for All Starting Values of x in the Range $(7, 7.8888)$.

This code implements and tests Newton's method for finding the optimal solution (root or minimum) of a function. Here's a breakdown of its functionality:

Function Definitions:

First derivative:

$$f'(x) = 9 - \frac{4}{x - 7} \quad \text{for } x > 7 \tag{1}$$

Second derivative:

$$f''(x) = \frac{4}{(x - 7)^2} \quad \text{for } x > 7 \tag{2}$$

Newton's Method Implementation:

The iteration formula:

$$x_{n+1} = x_n - \frac{f'(x_n)}{f''(x_n)}$$

Convergence criterion:

$$|x_{n+1} - x_n| < \text{tolerance}$$

Testing Convergence:

Starting points are generated in the range:

$$x_0 \in [7.01, 7.8888]$$

Choose 10 equally spaced starting values between the given range of (7,7.8888)

$$f'(x) = \frac{d}{dx}(9x - 4\ln(x-7)) = 9 - \frac{4}{x-7}$$

$$f''(x) = \frac{d}{dx}\left(9 - \frac{4}{x-7}\right) = \frac{4}{(x-7)^2}$$

Example: $x_0 = 7.0100$

$$f'(7.0100) = 9 - \frac{4}{7.0100 - 7} = -391$$

$$f''(7.0100) = \frac{4}{(7.0100-7)^2} = 40000$$

$$x_1 = 7.0100 - \frac{-391}{40000} = 7.019775$$

Continue the above process until convergence is reached

$x_0 = 7.0100$ converges in 11 iterations
$x_0 = 7.1076$ converges in 7 iterations
$x_0 = 7.2053$ converges in 6 iterations
$x_0 = 7.3029$ converges in 5 iterations
$x_0 = 7.4006$ converges in 4 iterations
$x_0 = 7.4982$ converges in 4 iterations
$x_0 = 7.5959$ converges in 5 iterations
$x_0 = 7.6935$ converges in 6 iterations
$x_0 = 7.7912$ converges in 7 iterations
$x_0 = 7.8888$ converges in 17 iterations

- If the point of convergence is the same for each x_0, then our function has converged to the optimal solution. Here our values of x_0 all converge to the same point of 7.4444.
- If the points of convergence are not the same for each x_0, then not all of our starting values within the range $(7, 7.8888)$ converge to the same solution.

5.What behavior does Newton's method exhibit outside of this range?

For our given function of $f, x > 7$ for all values ranged $x <= 7$ would return an error.

- The domain here is restricted to $x > 7$.
- Example: $f(x) = 9x - 4\ln(x-7) = f(6) = 9(6) - 4\ln(6-7) =$ Undefined.
- The Natural Logarithm of 0 or a negative number is undefined, thus cannot be computed.
- For $x > 7.8888$ The point do not converge to the same solution, meaning it is not an optional solution of $f(x) = 9x - 4\ln(x-7)$.

Problem 2

$$f(x_1, x_2) = -\ln(1 - x_1 - x_2) - \ln(x_1) - \ln(x_2)$$

1. Give an exact formula for the Newton iterate for a given value of (x_1, x_2) is computed as:

$$\text{Newton iterate: } (x_1, x_2)_{\text{new}} = (x_1, x_2) - H_f^{-1}(x_1, x_2)\nabla f(x_1, x_2)$$

2. Write a function for the gradient of f, $\nabla f(x_1, x_2)$

The gradient of $f(x_1, x_2)$ is given by:

$$\nabla f(x_1, x_2) = \begin{bmatrix} f_x(x_1, x_2) \\ \\ f_x(x_1, x_2) \end{bmatrix} = \begin{bmatrix} \frac{1}{1-x_1-x_2} - \frac{1}{x_1} \\ \\ \frac{1}{1-x_1-x_2} - \frac{1}{x_2} \end{bmatrix}$$

3. Write a function for the Hessian of f, $H_f(x_1, x_2)$

The Hessian matrix of $f(x_1, x_2)$ is given by:

$$H_f(x_1,x_2) = \begin{bmatrix} f_{xx}(x_1,x_2) & f_{xy}(x_1,x_2) \\ \\ f_{yx}(x_1,x_2) & f_{yy}(x_1,x_2) \end{bmatrix} = \begin{bmatrix} \frac{1}{(1-x_1-x_2)^2} + \frac{1}{x_1^2} & \frac{1}{(1-x_1-x_2)^2} \\ \\ \frac{1}{(1-x_1-x_2)^2} & \frac{1}{(1-x_1-x_2)^2} + \frac{1}{x_2^2} \end{bmatrix}$$

4. Find the minimizer of f with Newton's method using the following starting value $(0.85, 0.05)$. How many iterations are needed to converge?

Newton Iterate The Newton iterate for a given value of $(x_1, x_2) = (0.85, 0.05)$ is computed as:

$$(x_1, x_2)_{\text{new}} = (x_1, x_2) - H_f^{-1}(x_1, x_2)\nabla f(x_1, x_2)$$

$$(x_1, x_2)_{\text{new}} = (0.85, 0.05) - H_f^{-1}(0.85, 0.05)\nabla f(0.85, 0.05)$$

Gradient $\nabla f(x_1, x_2)$ The gradient of $f(0.85, 0.05)$ is given by:

$$\nabla f(0.85, 0.05) = \begin{bmatrix} \frac{1}{1-0.85-0.05} - \frac{1}{0.85} \\ \frac{1}{1-0.85-0.05} - \frac{1}{0.05} \end{bmatrix} = \begin{bmatrix} 8.82352941 \\ -10 \end{bmatrix}$$

Hessian $H_f(x_1, x_2)$ The Hessian matrix of $f(0.85, 0.05)$ is given by:

$$H_f(x_1, x_2) = \begin{bmatrix} 101.38408304 & 100 \\ 100 & 500 \end{bmatrix}$$

Inverse of the Hessian Matrix $H_f^{-1}(x_1, x_2)$ The inverse of the Hessian matrix is calculated as:

$$H_f^{-1}(x_1, x_2) = \frac{1}{(ad - bc)} \times \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$$

Substituting the values of $H_f(0.85, 0.05)$:

$$H_f^{-1}(0.85, 0.05) = \frac{1}{(101.384 \times 500) - (100 \times 100)} \times \begin{bmatrix} 500 & -100 \\ -100 & 101.384 \end{bmatrix}$$

$$= \frac{1}{40692} \times \begin{bmatrix} 500 & -100 \\ -100 & 101.384 \end{bmatrix} = \begin{bmatrix} 0.012287 & -0.002457 \\ -0.002457 & 0.002491 \end{bmatrix}$$

Applying Newton's Method Now, we apply the Newton step:

$$(x_1, x_2)_{\text{new}} = (0.85, 0.05) - \begin{bmatrix} 0.012287 & -0.002457 \\ -0.002457 & 0.002491 \end{bmatrix} \begin{bmatrix} 8.82352941 \\ -10 \end{bmatrix}$$

Performing the matrix multiplication:

$$= (0.85, 0.05) - \begin{bmatrix} 0.13299333 \\ -0.04659866 \end{bmatrix}$$

Thus, the new point is:

$$(x_1, x_2)_{\text{new}} = (0.71700667, 0.09659866) \approx (0.717, 0.097)$$

Convergence Utilizing Newton's Method the function will converge to the minimizer of $f(0.85, 0.05) = (0.33333334361761163, 0.3333333272412798) \approx \left(\frac{1}{3}, \frac{1}{3}\right)$. The given function will converge after 7 iterations.

5. Find the minimizer of f with the method of Steepest Descent using the following starting value $(0.85, 0.05)$. How many iterations are needed to converge?

$$f(x_1, x_2) = -\ln(1 - x_1 - x_2) - \ln(x_1) - \ln(x_2)$$
$$\text{for } x_1 > 0,\ x_2 > 0,\ \text{and } 1 - x_1 - x_2 > 0$$

Steepest Descent Method The method of Steepest Descent is given by:

$$(x_1, x_2)_{\text{new}} = (x_1, x_2) - \lambda \nabla f(x_1, x_2)$$

Here, the step size is $\lambda = 0.01$, and the tolerance is 1×10^{-6}.

Initial Conditions The starting point is:

$$(x_1, x_2) = (0.85, 0.05)$$

Gradient $\nabla f(x_1, x_2)$ The gradient of f at $(0.85, 0.05)$ is calculated as follows:

$$\nabla f(0.85, 0.05) = \begin{bmatrix} \frac{1}{1-x_1-x_2} - \frac{1}{x_1} \\ \frac{1}{1-x_1-x_2} - \frac{1}{x_2} \end{bmatrix} = \begin{bmatrix} \frac{1}{1-0.85-0.05} - \frac{1}{0.85} \\ \frac{1}{1-0.85-0.05} - \frac{1}{0.05} \end{bmatrix} = \begin{bmatrix} 8.82352941 \\ -10 \end{bmatrix}$$

Iteration Process To find the local minimum, we repeat the following steps:

1. Compute the gradient $\nabla f(x_1, x_2)$. 2. Update the point (x_1, x_2) using the steepest descent formula:

$$(x_1, x_2)_{\text{new}} = (x_1, x_2) - \lambda \nabla f(x_1, x_2)$$

3. Continue the iterations until the difference between successive points is less than the tolerance 1×10^{-3}, indicating that the algorithm has converged to a point where the function does not change significantly.

Conclusion The method of Steepest Descent will converge to the minimizer of $f(0.85, 0.05) = (0.333357520.33330812) \approx (\frac{1}{3}, \frac{1}{3})$. The number of iterations required to converge depends on the size of the gradient at each step and the chosen step size $\lambda = 0.01$. Concerning the given function $f(x_1, x_2) = -\ln(1 - x_1 - x_2) - \ln(x_1) - \ln(x_2)$ convergence will occur around 6 iterations.

Problem 3

1. Select a matrix A with n = 5 and p = 3. Then compute AT A and find its determinant.

Let A be a matrix defined as follows:

$$A = \begin{bmatrix} 3 & 7 & 1 \\ 5 & 2 & 8 \\ 4 & 6 & 9 \\ 1 & 3 & 2 \\ 7 & 4 & 5 \end{bmatrix}$$

The transpose of A is:

$$A^T = \begin{bmatrix} 3 & 5 & 4 & 1 & 7 \\ 7 & 2 & 6 & 3 & 4 \\ 1 & 8 & 9 & 2 & 5 \end{bmatrix}$$

Next, we compute $A^T A$:

$$A^T A = \begin{bmatrix} 3 & 5 & 4 & 1 & 7 \\ 7 & 2 & 6 & 3 & 4 \\ 1 & 8 & 9 & 2 & 5 \end{bmatrix} \begin{bmatrix} 3 & 7 & 1 \\ 5 & 2 & 8 \\ 4 & 6 & 9 \\ 1 & 3 & 2 \\ 7 & 4 & 5 \end{bmatrix}$$

Let $(A^T A)^{-1}$ denote the inverse of $A^T A$. Then, the matrix $(A^T A)^{-1} A^T$ is computed as follows:

$$(A^T A)^{-1} A^T$$

Finally, we find $(A^T A)^{-1} (A^T A)^{-1} A^T$ and compute the determinant of $A^T A$:

$$\det(A^T A) = 160871.99999999997$$

2. Select a matrix A with n = 3 and p = 5. Then compute AT A and find its determinant.

Let A be a matrix defined as follows:

$$A = \begin{bmatrix} 8 & 1 & 15 & 11 & 5 \\ 7 & 13 & 6 & 9 & 3 \\ 10 & 4 & 12 & 1 & 7 \end{bmatrix}$$

The transpose of A is:

$$A^T = \begin{bmatrix} 8 & 7 & 10 \\ 1 & 13 & 4 \\ 15 & 6 & 12 \\ 11 & 9 & 1 \\ 5 & 3 & 7 \end{bmatrix}$$

Next, we compute $A^T A$:

$$A^T A = \begin{bmatrix} 8 & 7 & 10 \\ 1 & 13 & 4 \\ 15 & 6 & 12 \\ 11 & 9 & 1 \\ 5 & 3 & 7 \end{bmatrix} \begin{bmatrix} 8 & 1 & 15 & 11 & 5 \\ 7 & 13 & 6 & 9 & 3 \\ 10 & 4 & 12 & 1 & 7 \end{bmatrix}$$

Let $(A^T A)^{-1}$ denote the inverse of $A^T A$. Then, the matrix $(A^T A)^{-1} A^T$ is computed as follows:

$$(A^T A)^{-1} A^T$$

Finally, we find $(A^T A)^{-1}(A^T A)^{-1} A^T$ and compute the determinant of $A^T A$:

$$\det(A^T A) = 0$$

3. What conclusion can you reach based on n and p?

Let matrix A be of dimensions $n \times p$, where n represents the number of rows and p represents the number of columns.

Invertibility Requirements

A matrix is invertible if it satisfies the following conditions:

- **Square Matrix**: The matrix must be $n \times n$.
- **Determinant**: The determinant of the matrix, $\det(A)$, cannot be zero.
- **Linearly Independent**: The columns of the matrix must be linearly independent.
- **Existence of an Inverse**: There must exist a matrix B such that $AB = BA = I$, where I is the identity matrix.

Analysis of Specific Cases

Concerning 3.1: Let A be a 5×3 matrix. Therefore, A^T is a 3×5 matrix and $A^T A$ is a 3×3 square matrix. When the number of rows n is greater than or equal to the number of columns p, $A^T A$ becomes a $p \times p$ matrix, which is a 3×3 matrix in this case. Matrix $A^T A$ meets the requirements for invertibility, as follows:

- The columns of A are linearly independent.
- The determinant of $A^T A$ is non-zero.

Concerning 3.2: Let A be a 3×5 matrix. Therefore, A^T is a 5×3 matrix and $A^T A$ is a 5×5 square matrix. When the number of rows n is less than or equal to the number of columns p, $A^T A$ becomes a $p \times p$ matrix, which is a 5×5 matrix in this case. Matrix $A^T A$ does not meet the requirements for invertibility, as follows:

- The rank of A would have to be $p = 5$ to meet the invertibility condition.
- However, the maximum rank of A is $n = 3$, which is less than p. This means that $A^T A$ is not of full rank and is thus not invertible.
- Consequently, matrix A is singular, and the determinant of $A^T A$ is zero.