MTH603/AM8211
Fall 2023

Assignment Cover Page

Assignment Number: 3

Student Last Name: Walaa

Student First Name: Mariam

Student ID: 500-744-640

If you used any resources other than what is listed on the course outline, or if you collaborated with any of your colleagues, list below these resources or the names of your colleagues along with their id's:

You are expected to know what constitutes academic integrity, to avoid committing academic offences, and to take responsibility for your actions. Provide your signature below to indicate that you are aware of all policies of the course (described in the course outline), including copyright.

Signature: Mariam Walaa

Assignment 3

Mariam Walaa

October 25, 2023

Question 1

[Question 1] Let $f : \mathbb{R}^n \to \mathbb{R}$ have continuous second partial derivatives. Consider $x_0 \in \mathbb{R}^n$ for which $\nabla f(x_0)$ is nonzero. Prove or disprove that for any $d \in \mathbb{R}^n$ for which $\nabla f(x_0) \cdot d < 0$, we have that d is a descent direction starting from x_0.

Solution

To prove this statement, we need to show that if $\nabla f(x_0) \cdot d < 0$, then d is a descent direction, meaning that the directional derivative of f in the direction of d is negative.

The directional derivative of f at a point x_0 in the direction of a vector d is given by:

$$\nabla f(x_0) \cdot d$$

Given that $\nabla f(x_0) \cdot d < 0$, we have a negative directional derivative.

Now, we can use the properties of continuous second partial derivatives. Since f has continuous second partial derivatives, we can apply the Corollary for Taylor's Expansion:

We are given $x_0, d \in \mathbb{R}^{\ltimes}$. Take small enough α approaching 0.

The Taylor expansion of f at a point x_0 is given by:

$$f(x_0 + \alpha d) = f(x_0) + \alpha \nabla f(x_0) \cdot d + O(\alpha^2) < f(x_0)$$

Since $\nabla f(x_0) \cdot d < 0$ and $\alpha > 0$, we can conclude that the second term $\alpha \nabla f(x_0) \cdot d$ is negative. This means that arbitrarily defined d is indeed a descent direction starting from x_0.

Therefore, we have proven that for any $d \in \mathbb{R}^n$ for which $\nabla f(x_0) \cdot d < 0$, d is a descent direction starting from x_0.

Question 2

[Question 2] Let $f : \mathbb{R}^n \to \mathbb{R}$ have continuous second partial derivatives, and assume that the Hessian of f, H_f, is positive for all $x \in \mathbb{R}^n$. Prove that if $\nabla f(x_0) = 0$, then x_0 is the unique global minimizer of f.

Solution:

We want to show that if $\nabla f(x_0) = 0$ and the Hessian of f, $\nabla^2 f(x)$, is positive, then x_0 is a global minimizer, and its uniqueness will follow from the positive definiteness of the Hessian. To prove that x_0 is the unique global minimizer of f, we will use the fact that $\nabla^2 f(x)$ is positive for all x in $\mathbb{R}^n$.

First, since f has continuous second partial derivatives, we can apply Taylor's Expansion of f around x_0:

$$f(x) = f(x_0) + \nabla f(x_0)^\top (x - x_0) + \frac{1}{2}(x - x_0)^\top H(x_0)(x - x_0) + \dots$$

Where $H(x_0) = \nabla^2 f(x)$ is the Hessian matrix of f evaluated at x_0. Since $\nabla f(x_0) = 0$, the first-order term $\nabla f(x_0)^\top (x - x_0)$ disappears, and we are left with:

$$f(x) = f(x_0) + 0 + \frac{1}{2}(x - x_0)^\top H(x_0)(x - x_0) + \dots$$

Now, because the Hessian matrix is positive definite, the quadratic form $(x - x_0)^\top H(x_0)(x - x_0)$ is strictly positive for all $x \neq x_0$. This means that $f(x) > f(x_0)$ for all $x \neq x_0$, which implies that x_0 is a local minimum.

Since we have that x_0 is a local minimum, and the Hessian is positive definite, this implies that f is convex function and x_0 is therefore a global minimum. Furthermore, the uniqueness of the global minimum follows from the positive definiteness of the Hessian, as it guarantees that there are no other local minima near x_0 that could compete with it.

Hence, x_0 is the unique global minimizer of f.

Question 3

[Question 3] Consider the non-linear function

$$f(x, y, z) = x^2y - xy + 3y^2 - y + 2z^2 + 372$$

(a) Find the gradient and Hessian of $f(x, y, z)$.
(b) Using the first order necessary optimality condition, find all candidate minimizers of $f(x, y, z)$.
(c) Find the eigenvalues of the Hessian of $f(x, y, z)$.
(d) Find all strict local minimizers of $f(x, y, z)$.

Solution:

Solution to 3(a):

We have the following partial derivatives

$$\frac{\partial f}{\partial x} = 2xy - y$$

$$\frac{\partial f}{\partial y} = x^2 - x + 6y - 1$$

$$\frac{\partial f}{\partial z} = 4z$$

So, the gradient $\nabla f(x, y, z)$ is a vector given by:

$$\nabla f(x, y, z) = \begin{bmatrix} 2xy - y \\ x^2 - x + 6y - 1 \\ 4z \end{bmatrix}$$

The second partial derivatives of f are as follows:

$$\frac{\partial^2 f}{\partial x^2} = 2y$$

$$\frac{\partial^2 f}{\partial y^2} = 6$$

$$\frac{\partial^2 f}{\partial z^2} = 4$$

$$\frac{\partial^2 f}{\partial x \partial y} = 2x - 1$$

$$\frac{\partial^2 f}{\partial x \partial z} = 0$$

$$\frac{\partial^2 f}{\partial y \partial z} = 0$$

So, the Hessian matrix $\mathbf{H}$ is given by:

$$\mathbf{H} = \begin{bmatrix} 2y & 2x - 1 & 0 \\ 2x - 1 & 6 & 0 \\ 0 & 0 & 4 \end{bmatrix}$$

Solution to 3(b):

To find the candidate minimizers of $f(x, y, z)$, we need to use the first-order necessary optimality condition, which involves setting the gradient of f equal to the zero vector.

Recall that the gradient of $f(x, y, z)$ is given by:

$$\nabla f(x,y,z) = \begin{bmatrix} 2xy - y \\ x^2 - x - 1 + 6y \\ 4z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

To find the candidate minimizers, we set $\nabla f(x, y, z) = \mathbf{0}$, where $\mathbf{0}$ is the zero vector:

$$\begin{bmatrix} 2xy - y \\ x^2 - x - 1 + 6y \\ 4z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

Solving for x, y, and z, we get the following system of equations:

$$\begin{aligned} 2xy - y &= 0 \\ x^2 - x - 1 + 6y &= 0 \\ 4z &= 0 \end{aligned}$$

The third equation $4z = 0$ implies that $z = 0$.

The first two equations are as follows:

1. From the first equation $2xy - y = 0$, we can factor out $2x$ to get:

$$y(2x - 1) = 0$$

This implies that either $2x - 1 = 0$ or $y = 0$.

2. From the second equation $x^2 - x - 1 + 6y = 0$, we can isolate $x^2 - x$ on one side:

$$x^2 - x = 1 - 6y$$

Now, let's consider the cases:

Case 1: $y = 0$

If $y = 0$, then $x^2 - x = 1 - 6(0) = 1$. Substituting this into the second equation:

$$x^2 - x - 1 = 0$$

This gives $x = \frac{1 \pm \sqrt{5}}{2}$ so $(\frac{1+\sqrt{5}}{2}, 0, 0)$ and $(\frac{1-\sqrt{5}}{2}, 0, 0)$ are candidates.

Case 2: $2x - 1 = 0$

If $x = \frac{1}{2}$, then $(\frac{1}{2})^2 - (\frac{1}{2}) = 1 - 6y = -\frac{1}{4}$.

Solving for y gives:

$$y = \frac{1 + \frac{1}{4}}{6} = \frac{5}{24}$$

So, we have three candidates:

Candidate 1: $\vec{x_1} = (x, y, z) = (\frac{1}{2}, \frac{5}{24}, 0)$

Candidate 2: $\vec{x_2} = (x, y, z) = (\frac{1+\sqrt{5}}{2}, 0, 0)$

Candidate 3: $\vec{x_2} = (x, y, z) = (\frac{1-\sqrt{5}}{2}, 0, 0)$

It can be verified that $\nabla f(x_i) = 0$ as required by the first order necessary optimality condition. Therefore, these are the candidate minimizers of the function $f(x, y, z)$ based on the first-order necessary optimality condition.

Solution to 3(c):

To find the eigenvalues of the Hessian matrix of $f(x, y, z)$, we first need to compute the Hessian matrix, and then solve for its eigenvalues.

The Hessian matrix, denoted as $\mathbf{H}$, is given by:

$$\mathbf{H} = \begin{bmatrix} 2y & 2x-1 & 0 \\ 2x-1 & 6 & 0 \\ 0 & 0 & 4 \end{bmatrix}$$

Now, let's find the eigenvalues of $\mathbf{H}$. We can do this by solving the characteristic equation:

$$\det(\mathbf{H} - \lambda \mathbf{I}) = 0$$

Where λ is the eigenvalue, and $\mathbf{I}$ is the identity matrix.

Substituting the values from $\mathbf{H}$, we get:

$$\det\left(\begin{bmatrix} 2y-\lambda & 2x-1 & 0 \\ 2x-1 & 6-\lambda & 0 \\ 0 & 0 & 4-\lambda \end{bmatrix}\right) = 0$$

Expanding the determinant, we have:

$$(2y-\lambda)((6-\lambda)(4-\lambda)-0) - (2x-1)(2x-1)(4-\lambda) = 0$$

Simplifying further gives:

$$(2y-\lambda)(24-10\lambda+\lambda^2) - (2x-1)(2x-1)(4-\lambda) = 0$$

Expanding and simplifying the expression gives:

$$-\lambda^3 + (2y+10)\lambda^2 + (4x^2 - 4x - 20y - 23)\lambda - 16x^2 + 16x + 48y - 4 = 0$$

The eigenvalues λ will depend on the values of x and y.

Thus, the eigenvalues of the Hessian matrix are given by the solutions to the cubic equation above, and the specific values of x and y would be needed to determine the eigenvalues.

Solution to 3(d):

To find the strict local minimizers of $f(x, y, z)$, we need to use the second-order condition, which requires that all eigenvalues of the Hessian matrix be positive. We have already calculated the Hessian matrix in a previous solution, and we will use that information.

For a point (x, y, z) to be a strict local minimizer, both eigenvalues of the Hessian matrix $\mathbf{H}$ at that point must be positive. Using the previously calculated Hessian matrix:

$$\mathbf{H} = \begin{bmatrix} 2y & 2x-1 & 0 \\ 2x-1 & 6 & 0 \\ 0 & 0 & 4 \end{bmatrix}$$

We can now calculate the eigenvalues for the Hessian matrix. For each point (x, y, z) where both eigenvalues are positive, that point is a strict local minimizer.

The eigenvalues are dependent on the values of x and y. We can calculate eigenvalues numerically for specific values of x and y. For example:

```
import numpy as np

# Define the Hessian matrix
x = (1+np.sqrt(5)/2)
y = 0
H = np.array([[2*y, 2*x - 1, 0],[2*x - 1, 6, 0],[0, 0, 4]])

# Calculate eigenvalues
eigenvalues = np.linalg.eigvals(H)

print("Eigenvalues of the Hessian matrix:")
for eigenvalue in eigenvalues:
    print(eigenvalue)

is_strict_minimizer = all(eigenvalues > 0)

print("Is it a strict local minimizer?", is_strict_minimizer)

# Define the Hessian matrix
x = (1-np.sqrt(5)/2)
y = 0
H = np.array([[2*y, 2*x - 1, 0],[2*x - 1, 6, 0],[0, 0, 4]])

# Calculate eigenvalues
eigenvalues = np.linalg.eigvals(H)

print("Eigenvalues of the Hessian matrix:")
for eigenvalue in eigenvalues:
    print(eigenvalue)

is_strict_minimizer = all(eigenvalues > 0)

print("Is it a strict local minimizer?", is_strict_minimizer)

# Define the Hessian matrix
x = 1/2
y = 5/24
H = np.array([[2*y, 2*x - 1, 0],[2*x - 1, 6, 0],[0, 0, 4]])
```

```
# Calculate eigenvalues
eigenvalues = np.linalg.eigvals(H)

print("Eigenvalues of the Hessian matrix:")
for eigenvalue in eigenvalues:
    print(eigenvalue)

is_strict_minimizer = all(eigenvalues > 0)

print("Is it a strict local minimizer?", is_strict_minimizer)
-----------------------------------------------------------------

Eigenvalues of the Hessian matrix:
-1.412724323476323
7.412724323476324
4.0
Is it a strict local minimizer? False

Eigenvalues of the Hessian matrix:
-0.24466701604346763
6.244667016043468
4.0
Is it a strict local minimizer? False

Eigenvalues of the Hessian matrix:
0.4166666666666667
6.0
4.0
Is it a strict local minimizer? True
```

We see that $(\frac{1}{2}, \frac{5}{24}, 0)$ is the only such candidate with all positive eigenvalues, so it is the only strict local minimizer in this problem.