DDA 6010 Assignment 1

2024-9-17

Problem 1 Coerciveness and Optimality

(a) (1) $f(x, y, z) = x^2 - 2xy + y^2 + z^2$

Coercivity: Since f is a quadratic function, rewrite it as:

$$f(\mathbf{x}) = \mathbf{x}^{\mathbf{T}}\mathbf{A}\mathbf{x}$$

Where

$$\mathbf{A} =$$

is positive definite, Thus f is coercive

Global Minimizers:

$$f(x, y, z) = 0 \implies x = y \text{ and } z = 0$$

The global minimizers are all points of the form $(x, x, 0)$ for $x \in \mathbb{R}$.

(2) $f(x, y) = x^4 + y^4 - 4xy$

Coercivity: As $|x|$ or $|y|$ becomes large, x^4 and y^4 dominate, ensuring $f(x, y) \to +\infty$. Thus, f is coercive.

Global Minimizers: Find critical points by solving:

$$\frac{\partial f}{\partial x} = 4x^3 - 4y = 0 \quad \Rightarrow \quad y = x^3$$

$$\frac{\partial f}{\partial y} = 4y^3 - 4x = 0 \quad \Rightarrow \quad x = y^3$$

Substituting $y = x^3$ into $x = y^3$:

$$x = (x^3)^3 \implies x^9 = x \implies x(x^8 - 1) = 0 \implies x = 0, \pm 1$$

Corresponding y values:

$$y = 0, \pm 1$$

Critical points: $(0, 0)$, $(1, 1)$, $(-1, -1)$. Evaluating f at these points:

$$f(0, 0) = 0, \quad f(1, 1) = -2, \quad f(-1, -1) = -2$$

The global minimizers are $(1, 1)$ and $(-1, -1)$ with minimum value -2.

(b) **Finding Stationary Points:** Compute the partial derivatives:

$$f_x = 4x^3 - y^2 - x = 0$$

$$f_y = -2xy + 2y = 0 \implies y(-2x + 2) = 0$$

Case 1: $y = 0$

$$4x^3 - x = 0 \implies x(4x^2 - 1) = 0 \implies x = 0, \pm\frac{1}{2}$$

Stationary points: $(0,0)$, $\left(\frac{1}{2}, 0\right)$, $\left(-\frac{1}{2}, 0\right)$

Case 2: $-2x + 2 = 0 \implies x = 1$

$$4(1)^3 - y^2 - 1 = 0 \implies y^2 = 3 \implies y = \pm\sqrt{3}$$

Stationary points: $(1, \sqrt{3})$, $(1, -\sqrt{3})$

Classification of Stationary Points: Compute the second partial derivatives:

$$f_{xx} = 12x^2 - 1, \quad f_{xy} = -2y, \quad f_{yy} = -2x + 2$$

The Hessian determinant:

$$D = f_{xx}f_{yy} - (f_{xy})^2$$

(i) **Point** $(0,0)$**:**

$$f_{xx} = -1, \quad f_{xy} = 0, \quad f_{yy} = 2$$

$$D = (-1)(2) - 0 = -2 < 0 \implies \text{Saddle Point}$$

(ii) **Point** $\left(\frac{1}{2}, 0\right)$**:**

$$f_{xx} = 12\left(\frac{1}{2}\right)^2 - 1 = 2, \quad f_{xy} = 0, \quad f_{yy} = 1$$

$$D = (2)(1) - 0 = 2 > 0 \quad \text{and} \quad f_{xx} > 0 \implies \text{Local Minimum}$$

(iii) **Point** $\left(-\frac{1}{2}, 0\right)$**:**

$$f_{xx} = 12\left(-\frac{1}{2}\right)^2 - 1 = 2, \quad f_{xy} = 0, \quad f_{yy} = 3$$

$$D = (2)(3) - 0 = 6 > 0 \quad \text{and} \quad f_{xx} > 0 \implies \text{Local Minimum}$$

(iv) **Point** $(1, \sqrt{3})$**:**

$$f_{xx} = 11, \quad f_{xy} = -2\sqrt{3}, \quad f_{yy} = 0$$

$$D = (11)(0) - (-2\sqrt{3})^2 = -12 < 0 \implies \text{Saddle Point}$$

(v) **Point** $(1, -\sqrt{3})$**:**

$$f_{xx} = 11, \quad f_{xy} = 2\sqrt{3}, \quad f_{yy} = 0$$

$$D = (11)(0) - (2\sqrt{3})^2 = -12 < 0 \implies \text{Saddle Point}$$

Summary:

- **Coercivity:**
  - $f(x,y,z) = x^2 - 2xy + y^2 + z^2$ is coercive.
  - $f(x,y) = x^4 + y^4 - 4xy$ is coercive.
- **Global Minimizers:**
  - For $f(x,y,z)$: All points $(x,x,0)$.
  - For $f(x,y)$: Points $(1,1)$ and $(-1,-1)$ with $f = -2$.
- **Stationary Points of** $f(x,y) = x^4 - xy^2 - \frac{1}{2}x^2 + y^2$**:**

- $(0,0)$: Saddle Point
- $\left(\frac{1}{2},0\right)$: Local Minimum
- $\left(-\frac{1}{2},0\right)$: Local Minimum
- $(1,\sqrt{3})$: Saddle Point
- $(1,-\sqrt{3})$: Saddle Point

Problem 2 Quadratic Optimization Problems

(a) **Verification of Global Optimal Solution for Problem (1):**

Consider the optimization problem:

$$\min_{\mathbf{x}\in\mathbb{R}^n, y\in\mathbb{R}} g(\mathbf{x},y) := \frac{\lambda}{2}\|\mathbf{x}\|_2^2 + \sum_{i=1}^{m} \max\{0, 1-b_i(\mathbf{a}_i^\top \mathbf{x}+y)\}, \tag{1}$$

where $\lambda > 0$, $\mathbf{A} = (\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_m) \in \mathbb{R}^{n\times m}$, and $b_i \in \{-1, 1\}$ for all i.

Coercivity:

- The term $\frac{\lambda}{2}\|\mathbf{x}\|_2^2$ is coercive since $\lambda > 0$ and $\|\mathbf{x}\|_2^2 \to \infty$ as $\|\mathbf{x}\|_2 \to \infty$.
- The hinge loss $\max\{0, 1-b_i(\mathbf{a}_i^\top \mathbf{x}+y)\}$ is non-negative for all i.
- Thus, as $\|(\mathbf{x},y)\| \to \infty$, $g(\mathbf{x},y) \to \infty$.

Lower Boundedness and Convexity:

- Each $\max\{0, 1-b_i(\mathbf{a}_i^\top \mathbf{x}+y)\}$ is convex.
- The sum of convex functions is convex.
- The quadratic term is convex due to $\lambda > 0$.
- Therefore, $g(\mathbf{x},y)$ is convex and coercive.

Existence of Global Minimizer:

- A convex, coercive function on $\mathbb{R}^{n+1}$ attains its global minimum.
- Hence, Problem (1) possesses at least one global optimal solution.

(b) **Derivation of Optimal α for the Unconstrained Quadratic Optimization Problem:**

Consider the problem:

$$\min_{\alpha\in\mathbb{R}} f(\mathbf{x}-\alpha\nabla f(\mathbf{x})), \quad \text{where} \quad f(\mathbf{x}) = \frac{1}{2}\mathbf{x}^\top\mathbf{A}\mathbf{x} + \mathbf{b}^\top\mathbf{x} + c,$$

with $\mathbf{A} \in \mathbb{R}^{n\times n}$ symmetric and positive definite, $\mathbf{b} \in \mathbb{R}^n$, and $c \in \mathbb{R}$.

Gradient Computation:

$$\nabla f(\mathbf{x}) = \mathbf{A}\mathbf{x} + \mathbf{b}$$

Define the Update Step:

$$\mathbf{x}_{\text{new}} = \mathbf{x} - \alpha\nabla f(\mathbf{x}) = \mathbf{x} - \alpha(\mathbf{A}\mathbf{x}+\mathbf{b})$$

Express $f(\mathbf{x}_{\text{new}})$:

$$f(\mathbf{x}_{\text{new}}) = \frac{1}{2}(\mathbf{x}-\alpha(\mathbf{A}\mathbf{x}+\mathbf{b}))^\top\mathbf{A}(\mathbf{x}-\alpha(\mathbf{A}\mathbf{x}+\mathbf{b})) + \mathbf{b}^\top(\mathbf{x}-\alpha(\mathbf{A}\mathbf{x}+\mathbf{b})) + c$$

Simplify $f(\mathbf{x_{new}})$:

$$f(\mathbf{x}_{\text{new}}) = \frac{1}{2}\mathbf{x}^\top \mathbf{A}\mathbf{x} - \alpha \mathbf{x}^\top \mathbf{A}(\mathbf{A}\mathbf{x} + \mathbf{b}) + \frac{1}{2}\alpha^2 (\mathbf{A}\mathbf{x} + \mathbf{b})^\top \mathbf{A}(\mathbf{A}\mathbf{x} + \mathbf{b}) + \mathbf{b}^\top \mathbf{x} - \alpha \mathbf{b}^\top (\mathbf{A}\mathbf{x} + \mathbf{b}) + c$$

Identify Quadratic Form in α:

$$f(\mathbf{x}_{\text{new}}) = \underbrace{\frac{1}{2}\mathbf{x}^\top \mathbf{A}\mathbf{x} + \mathbf{b}^\top \mathbf{x} + c}_{\text{Constant}} + \underbrace{\left(-\alpha \mathbf{x}^\top \mathbf{A}(\mathbf{A}\mathbf{x} + \mathbf{b}) - \alpha \mathbf{b}^\top (\mathbf{A}\mathbf{x} + \mathbf{b})\right)}_{\text{Linear in } \alpha} + \underbrace{\frac{1}{2}\alpha^2 (\mathbf{A}\mathbf{x} + \mathbf{b})^\top \mathbf{A}(\mathbf{A}\mathbf{x} + \mathbf{b})}_{\text{Quadratic in } \alpha}$$

Express $f(\mathbf{x_{new}})$ as a Quadratic Function:

$$f(\mathbf{x}_{\text{new}}) = C - \alpha \|\nabla f(\mathbf{x})\|^2 + \frac{1}{2}\alpha^2 \nabla f(\mathbf{x})^\top \mathbf{A} \nabla f(\mathbf{x})$$

where C is a constant independent of α.

Find Optimal α:

To minimize $f(\mathbf{x}_{\text{new}})$ with respect to α, take the derivative and set it to zero:

$$\frac{df}{d\alpha} = -\|\nabla f(\mathbf{x})\|^2 + \alpha \nabla f(\mathbf{x})^\top \mathbf{A} \nabla f(\mathbf{x}) = 0$$

$$\Rightarrow \alpha^* = \frac{\|\nabla f(\mathbf{x})\|^2}{\nabla f(\mathbf{x})^\top \mathbf{A} \nabla f(\mathbf{x})}$$

Final Expression for Optimal α:

$$\alpha^* = \frac{(\mathbf{A}\mathbf{x} + \mathbf{b})^\top (\mathbf{A}\mathbf{x} + \mathbf{b})}{(\mathbf{A}\mathbf{x} + \mathbf{b})^\top \mathbf{A}(\mathbf{A}\mathbf{x} + \mathbf{b})}$$

Problem 3 Optimality Condition

(a) To find the stationary points of $f(x_1, x_2)$, we compute the partial derivatives and set them to zero.

Partial Derivatives:

$$\frac{\partial f}{\partial x_1} = f(x_1, x_2) \cdot \left(2x_1 - 3x_2^2(1 - x_1)^2\right)$$

$$\frac{\partial f}{\partial x_2} = f(x_1, x_2) \cdot \left(2x_2(1 - x_1)^3\right)$$

Setting Partial Derivatives to Zero:

$$\frac{\partial f}{\partial x_1} = 0 \quad \text{and} \quad \frac{\partial f}{\partial x_2} = 0$$

Since $f(x_1, x_2) > 0$ for all (x_1, x_2), the equations reduce to:

$$2x_1 - 3x_2^2(1 - x_1)^2 = 0 \quad \text{and} \quad 2x_2(1 - x_1)^3 = 0$$

Solving the System:

From the second equation:

$$2x_2(1 - x_1)^3 = 0$$

This implies either:

$$x_2 = 0 \quad \text{or} \quad (1 - x_1)^3 = 0 \implies x_1 = 1$$

Case 1: $x_2 = 0$

Substitute $x_2 = 0$ into the first equation:

$$2x_1 - 3(0)^2(1 - x_1)^2 = 2x_1 = 0 \implies x_1 = 0$$

Thus, one stationary point is $z = (0, 0)$.

Case 2: $x_1 = 1$

Substitute $x_1 = 1$ into the first equation:

$$2(1) - 3x_2^2(1 - 1)^2 = 2 = 0$$

This is a contradiction. Hence, there are no stationary points with $x_1 = 1$ except possibly when $x_2 = 0$, which leads back to $z = (0, 0)$.

Conclusion:

The only solution to the system is $z = (0, 0)$. Therefore, $z = (0, 0)$ is the unique stationary point of f.

(b) To determine whether $z = (0, 0)$ is a local minimum, we examine the second-order partial derivatives and the Hessian matrix at z.

First-Order Partial Derivatives:

$$f_{x_1} = \frac{\partial f}{\partial x_1} = f(x_1, x_2) \cdot \left(2x_1 - 3x_2^2(1 - x_1)^2\right)$$

$$f_{x_2} = \frac{\partial f}{\partial x_2} = f(x_1, x_2) \cdot \left(2x_2(1 - x_1)^3\right)$$

Second-Order Partial Derivatives:

$$f_{x_1x_1} = \frac{\partial^2 f}{\partial x_1^2} = f \cdot \left(2 - 3x_2^2 \cdot (-2)(1 - x_1)\right) + f \cdot \left(2x_1 - 3x_2^2(1 - x_1)^2\right)^2$$

$$f_{x_1x_2} = \frac{\partial^2 f}{\partial x_1 \partial x_2} = \frac{\partial}{\partial x_2}\left(f \cdot \left(2x_1 - 3x_2^2(1 - x_1)^2\right)\right)$$

$$f_{x_2x_2} = \frac{\partial^2 f}{\partial x_2^2} = \frac{\partial}{\partial x_2}\left(f \cdot \left(2x_2(1 - x_1)^3\right)\right)$$

However, evaluating the Hessian at $z = (0, 0)$ simplifies the calculations:

$$f(0, 0) = \exp(0 + 0) = 1$$

$$f_{x_1x_1}(0, 0) = 2 \cdot f(0, 0) = 2$$

$$f_{x_1x_2}(0, 0) = 0 \quad \text{(since terms involving } x_2 \text{ vanish)}$$

$$f_{x_2x_2}(0, 0) = 2 \cdot f(0, 0) = 2$$

Hessian Matrix at $z = (0, 0)$:

$$H = \begin{bmatrix} f_{x_1x_1} & f_{x_1x_2} \\ f_{x_1x_2} & f_{x_2x_2} \end{bmatrix} = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}$$

Determining Definiteness:

- The Hessian matrix H is diagonal with positive entries.
- The eigenvalues of H are both 2, which are positive.
- Therefore, H is positive definite.

Conclusion:

Since the Hessian is positive definite at $z = (0, 0)$, z is a local minimum of f.

(c) **Boundedness from Below:**

The exponential function $\exp(t)$ is always positive for all real numbers t. Thus:

$$f(x_1, x_2) = \exp\left(x_1^2 + x_2^2(1 - x_1)^3\right) > 0 \quad \forall (x_1, x_2) \in \mathbb{R}^2$$

Hence, f is bounded from below by 0.

Analyzing Global Minimality:

Local Minimum at $z = (0, 0)$:

From part (b), $z = (0, 0)$ is a local minimum with:

$$f(0, 0) = 1$$

Existence of Other Points with Lower Function Values:

To determine whether z is a global minimum, we seek points (x_1, x_2) where $f(x_1, x_2) < 1$.

Consider $x_1 \to \infty$,

$$x_1^2 + x_2^2(1 - x_1)^3 \to -\infty$$

This shows that $f(x_1, x_2)$ can take values **arbitrarily close to 0**, which is **less than 1**.

Thus proved

Problem 4: Nonlinear Least Squares

(a): Formulation

To fit the quadratic function $y = ax^2 + bx + c$ to the data points $\{(x_i, y_i)\}_{i=1}^{50}$, we can express the problem in matrix form.

$$X = \begin{pmatrix} x_1^2 & x_1 & 1 \\ x_2^2 & x_2 & 1 \\ \vdots & \vdots & \vdots \\ x_{50}^2 & x_{50} & 1 \end{pmatrix}, \quad z = \begin{pmatrix} a \\ b \\ c \end{pmatrix}, \quad y = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_{50} \end{pmatrix}$$

The least squares problem is:

$$\min_{z \in \mathbb{R}^3} \|Xz - y\|^2$$

(b): Analytical Solution

The analytical solution to the least squares problem is given by the normal equations:

$$X^T X z = X^T y$$

Solving for z:

$$z = (X^T X)^{-1} X^T y$$

This yields the coefficients a, b, and c that minimize the sum of squared residuals.

(c): Python Code and Plot

Below is the Python code to generate the data, perform the least squares fitting, and plot the results.

```python
import numpy as np
import matplotlib.pyplot as plt

# Seed for reproducibility
np.random.seed(314)

# Generate 50 points between 0 and 1
x = np.linspace(0, 1, 50)

# True quadratic parameters
a_true = 2
b_true = -3
c_true = 1

# Generate noisy y-values
noise = 0.05 * np.random.randn(len(x))
y = a_true * x**2 + b_true * x + c_true + noise

# Construct design matrix X
X = np.vstack((x**2, x, np.ones(len(x)))).T

# Compute least squares solution
z = np.linalg.inv(X.T @ X) @ X.T @ y
a_est, b_est, c_est = z

print("Estimated Parameters:")
print(f"a = {a_est:.4f}")
print(f"b = {b_est:.4f}")
print(f"c = {c_est:.4f}")

# Generate fitted y-values
x_fit = np.linspace(0, 1, 200)
y_fit = a_est * x_fit**2 + b_est * x_fit + c_est

# Plotting
plt.figure(figsize=(10, 6))
plt.scatter(x, y, color='blue', label='Data Points')
plt.plot(x_fit, y_fit, color='red', linewidth=2, label='Fitted Quadratic')
plt.plot(x_fit, a_true * x_fit**2 + b_true * x_fit + c_true,
         color='green', linestyle='--', linewidth=2, label='True Quadratic')
plt.title('Nonlinear Least Squares Fit of Quadratic Function')
plt.xlabel('$x$')
plt.ylabel('$y$')
plt.legend()
plt.grid(True)
plt.show()
```

Listing 1: Nonlinear Least Squares Fitting

Explanation:

- **Data Generation:** Generates 50 data points based on the true quadratic function with added Gaussian noise.
- **Least Squares Fitting:** Constructs the design matrix X and computes the estimated parameters using the normal equations.
- **Plotting:** Displays the original data points, the fitted quadratic curve, and the true quadratic curve for comparison.

Sample Output:

```
Estimated Parameters:
a = 2.0012
b = -2.9985
c = 1.0013
```

Sample Plot: