Modern Optimization Techniques

Retake-Exam (online)

Prof. Dr. Dr. Lars Schmidt-Thieme
M.Sc. Mohsan Jameel

Jun 2^{nd}, 2020
ISMLL Universitt Hildesheim

Problem	A	B	C	Σ
1. Unconstrained Optimization				/10
2. Constrained Optimization				/10
3. Equality Constrained Optimization				/10
4. Inequality Constrained Optimization				/10

First Name: ______________________ Bonus: /04

Last Name: ______________________ Total: /40

Matrikel: ______________________ Grade:

Note:

- Time: 120 minutes
- The first upload phase will be from 11:00 - 11:10. In this phase, you must upload your first draft
- The second upload phase will be from 12:10 - 12:20. In the second and final phase, you have to upload your final solution on the learnweb.
- You cannot communicate with another person via any means
- You can only use mouse to interact with your laptop/pc, typing/talking are not allowed.

1. Unconstrained Optimization

1A. Second Order Convexity Condition (2 points)

Show that the following functions are convex by using the second order convexity condition:

(i) $f_1 : \mathbb{R} \longrightarrow \mathbb{R}^+ \quad f_1(x) = e^{ax} \quad a \in \mathbb{R}$

(ii) $f_2 : \mathbb{R} \longrightarrow \mathbb{R} \quad f_2(x) = x \cdot \log x$

(iii) $f_3 : \mathbb{R}^2 \longrightarrow \mathbb{R} \quad f_3\left(x_1, x_2\right) = x_1^2 + x_2^2$

1B. Stochastic Gradient Descent for unconstrained problem (5 points)

In this question you have to use Stochastic Gradient Descent optimization method to find solution for a least squared objective function.

$$\min_{\beta} \frac{1}{2}\|Y - X\beta\|_2^2$$

Perform two iterations (one epoch) starting with $\beta^{(0)} = (2, 1)^T$ and $\mu = 0.1$. What is the overall loss (RMSE) after each iteration?

$$X = \begin{pmatrix} 1 & 2 \\ 1 & 3 \end{pmatrix} \quad Y = \begin{pmatrix} 7 \\ 9 \end{pmatrix}$$

1C. Convexity / Concavity of Functions (3 points)

c) Suppose $f : \mathbb{R} \longrightarrow \mathbb{R}$ is increasing and convex on its domain (a, b). Let g denote its inverse, i.e., the function with domain $(f(a), f(b))$ and $g(f(x)) = x$ for $a < x < b$ What can you say about convexity or concavity of g?

2. Constrained Optimization

2A. Problem formulation (Linear Programing) (3 points)

A healthy diet contains m different nutrients in quantities at least equal to $b_1, \cdots, b_m$. This diet can be composed by choosing nonnegative quantities $x_1, \cdots, x_n$ of n different foods. One unit quantity of food j contains an amount a_{ij} of nutrient i, and has a cost of c_j.

Formulate this problem as an optimization problem in order to determine the cheapest diet that satisfies the nutritional requirements. What if we want to determine the most nutritious diet that does not exceed the cost given by the different foods? What is the relation between both formulations?

2B. Simplex Method (5 points)

Use the simplex method to solve the following optimization problem:

$$
\begin{aligned}
\max \quad & Z = 4x_1 - x_2 + 2x_3 \\
\text{s.t.} \quad & 2x_1 + x_2 + 2x_3 \leq 6 \\
& x_1 - 4x_2 + 2x_3 \leq 0 \\
& 5x_1 - 2x_2 - 2x_3 \leq 4 \\
& x_1, x_2, x_3 \geq 0
\end{aligned}
$$

2C. Duality (2 points)

The SVM primal problem can be written as follows:

$$\min \frac{1}{2}\|\beta\|^2$$
$$\text{s.t.} y_i \left(\beta_0 + \beta^T x_i\right) \geq 0, \quad i = 1, .., n$$
$$\beta \in \mathbb{R}^p, \quad \beta_0 \in \mathbb{R}$$

Show that the dual problem can be written as follows:

$$\max -\frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n}\alpha_i\alpha_j y_i y_j x_i^T x_i$$
$$\text{s.t.} \sum_{i=n}^{n}\alpha_i y_i = 0, \quad \alpha_i \geq 0$$

What is the advantage of using the dual problem over the primal in this case?

3. Equality Constrained Optimization

3A. KKT Conditions (3 points)

Consider the following optimization problem:

$$\begin{aligned} \text{minimize} \quad & f(x_1, x_2) = x_1^2 + 2x_2^2 + 2x_1 \\ \text{subject to} \quad & x_1 + 2x_2 \geq 5 \\ & x_1, x_2 \in \mathbb{R} \end{aligned}$$

a. Make a sketch of the level sets of f and the constraint.

b. Are the KKT conditions sufficient for guaranteeing optimality for this problem? Find the points satisfying the KKT conditions and find the optimal solution.

3B. Newton's Method for equality constraints (5 points)

Consider the following optimization problem:

$$\begin{aligned} \text{minimize} \quad & f(x_1, x_2, x_3) = 3x_1^2 + x_2^2 + 4x_3^2 \\ \text{subject to} \quad & x_1 + 2x_2 - 3x_3 = 0 \\ & -2x_1 - x_2 + 12x_3 = 2 \end{aligned}$$

Use the equality constrained Newton Algorithm with a starting point $x^{(0)} = (1, 0, 1/3)^\top$ and learning rate $\mu = 0.05$. Write down the system of equations that you would have to solve in the first step of the algorithm. **Solve one step by computing** $x^{(1)}$.

Hint: The inverse matrix of: $\begin{bmatrix} 6 & 0 & 0 & 1 \\ 0 & 2 & 0 & 2 \\ 0 & 0 & 8 & -3 \\ 1 & 2 & -3 & 0 \end{bmatrix}$ is $\begin{bmatrix} 0.16 & -0.05 & 0.02 & 0.05 \\ -0.05 & 0.2 & 0.11 & 0.3 \\ 0.02 & 0.11 & 0.08 & -0.11 \\ 0.05 & 0.3 & -0.11 & -0.3 \end{bmatrix}$

3C. Subdifferential (2 points)

Given the following function

$$f(x) = \max\left\{0, \frac{1}{2}\left(x^2 - 2x - 3\right)\right\}$$

Draw a sketch of this function. Compute its subdifferential!

4. Inequality Constrained Optimization

4A. Backtracking Line Search (2 points)

Consider the following minimization problem for $a \leq b$ and $a, b \in \mathbb{R}$

$$\begin{array}{ll} \text{minimize} & f_0(x) = \mathrm{e}^{(x-5)^2} \\ \text{subject to} & x \in [a, b] \end{array}$$

What is the derived optimization problem that you would solve using the logarithmic barrier function?

4B. Inequality Constrained: Active Set Method (5 points)

Solve the following optimization problem using the Active Set method with the equality constraint Newton method to find the search direction and the starting point $x^{(0)} = (2,0)^T$. Stop when the working set is empty.

$$\begin{aligned}
\text{minimize} \quad & f(x_1, x_2) = x_1^2 + x_2^2 - 2x_1 - 5x_2 \\
\text{subject to} \quad & -x_1 + 2x_2 - 2 \leq 0 \\
& x_1 + 2x_2 - 6 \leq 0 \\
& x_1 - 2x_2 - 2 \leq 0 \\
& -x1, -x2 \leq 0
\end{aligned}$$

Hint: For simplicity, use the most negative value of the Lagrangian multiplier λ to drop the constraint and set the step size $\mu = 1$. Stop when the working set is empty.

4C. Barrier Function (3 points)

What are the advantages of the penalty methods over the barrier methods? Is it possible to combine them in one method? Propose one mixed Barrier-Penalty function and discuss your choice.

Reserve Page – clearly indicate which problem you are working on!