Modern Optimization Techniques

Final Exam

Prof. Dr. Dr. Lars Schmidt-Thieme
M.Sc. Nourhan Ahmed

Feb 18^{th}, 2022
ISMLL Universität Hildesheim

Note:

- Time: 120 minutes
- Add your name and matriculation number on top of every page.
- Please provide clear and detailed answers to get full points.
- Please make sure the provided solution is clearly written and scanned.
- Plagiarism, cheating and group submissions are not allowed. Any suspicious solution will be further investigated and the student will fail the course for this semester.

1. Unconstrained Optimization and Convexity

1A. Convexity of functions (3 points)

For what values of the parameter a is the quadratic form $Q(x, y, z) = ax^2 + 4ay^2 + 4az^2 + 4xy + 2axz + 4yz$ is:

- Positive definite.
- Negative definite.

1B. Coordinate Descent (4 points)

For the following data A, learn a linear regression model

$$\hat{y}(a) = \sum_{i=1}^{2} \beta_i a_i$$

using Coordinate Descent. Initialize your parameters as $\beta = (1, 1)^\top$, and do **two iterations**. Write down β after each iteration.

$$A = \begin{pmatrix} 1 & 4 \\ 2 & 2 \end{pmatrix} \qquad y = \begin{pmatrix} 5 \\ 2 \end{pmatrix}$$

1C. Differentials (3 points)

Prove whether or not the following function $g(x)$ is subgradient of the corresponding function $f(x)$: $f(x) = \max\{f_1(x), f_2(x)\}, x \in \mathbb{R}^n, f_1(x)$ *and* $f_2(x)$ are convex and continuously differentiable:

$$g(x) = \begin{cases} \nabla f_1(x) & if \;\; f_1(x) > f_2(x) \\ \nabla f_2(x) & if \;\; f_1(x) \leq f_2(x) \end{cases}$$

2. Constrained Optimization and Duality

2A. Linear Programming (3 points)

A company makes two products (X and Y) using two machines (A and B). Each unit of X that is produced requires 50 minutes processing time on machine A and 30 minutes processing time on machine B. Each unit of Y that is produced requires 24 minutes processing time on machine A and 33 minutes processing time on machine B.

At the start of the current week there are 30 units of X and 90 units of Y in stock. Available processing time on machine A is forecast to be 40 hours and on machine B is forecast to be 35 hours.

The demand for X in the current week is forecast to be 75 units and for Y is forecast to be 95 units. Company policy is to maximise the combined sum of the units of X and the units of Y in stock at the end of the week.

- Formulate the problem of deciding how much of each product to make in the current week as a linear program.

2B. Simplex Method (5 points)

Use the simplex method to solve the following optimization problem:

$$\begin{array}{ll} \text{maximize:} & z = 3x_1 + 2x_2 + x_3 \\ \text{subject to the constraints:} & 4x_1 + x_2 + x_3 \leq 30 \\ & 2x_1 + 3x_2 + x_3 \leq 60 \\ , & x_1 + 2x_2 + 3x_3 \leq 40 \\ & x_1 \geq 0,\, x_2 \geq 0,\, and\, x_3 \geq 0 \end{array}$$

2C. Linear Programming and Duality (2 points)

Consider the linear programming problem:

$$\begin{array}{ll} \text{minimize} & z = c^T x + d^T v \\ \text{subject to:} & A_1 x + Bv \geq b_1, \\ & A_2 x = b_2, \\ & \sum_{k=1}^{l} v_k = a, \\ & x \geq 0^n, \\ & v \geq 0^l, \end{array}$$

$$where \quad x \in R^n, \quad v \in R^l, \quad c \in R^n, \quad d \in R^l, \quad A_1 \in R^{m_1 \times n}, \quad A_2 \in R^{m_2 \times n}, \\ \mathrm{B} \in R^{m_1 \times l}, \quad b_1 \in R^{m_1}, \quad b_2 \in R^{m_2}, \quad a \in R.$$

- Formulate its dual problem.

3. Constrained Minimization and Duality

3A. Active Set Method (2 points)

Explain in your words how Active Set method works and **two** main limitations of this method.

3B. Duality (5 points)

Consider the strictly convex quadratic optimization problem:

$$\begin{aligned} \text{minimize} \quad & f(x_1, x_2) = 2x_1^2 + x_2^2 - 4x_1 - 6x_2 \\ \text{subject to} \quad & -x_1 + 2x_2 \leq 4 \end{aligned}$$

1. Explicitly state its dual Lagrangian function g as a function of λ.
2. Solve this dual Lagrangian problem and provide the optimal Lagrange multiplier λ
3. Prove that strong duality holds.

3C. Penalty Method (3 points)

For quadratic penalty functions, the penalty is equal to the sum of square of all constraints, such that the penalty function is defined as below, where ρ is the penalty weight:

$$P(x, \rho) = \rho g(x)^\top g(x)$$

1. Can this formulation of the penalty function $P(x, \rho)$ handle inequality constraints?
   If yes, explain how. If not, reformulate the function $P(x, \rho)$ by defining a function $\phi(x, \rho)$ that can handle inequality constraints.
2. Show how penalty functions can suffer from ill conditioning, when considering inequality constraints.
   Hint: Consider that the penalty weight increases over time during the optimization procedure.

4. Constrained Optimization

4A. Penalty Method (2 points)

Consider the following optimization problem:

$$\begin{aligned} \text{maximize} \quad & f(x_1, x_2) = x_1^2 + 2x_2^2 \\ \text{subject to} \quad & 1 + x_1 - x_2 \geq 0 \\ & x_1 \geq 0 \end{aligned}$$

1. Write down the derived optimization problem for the penalty method using the quadratic penalty function starting at $x = (0, 3)$

4B. Active Set Method (5 points)

Solve the following optimization problem using Active Set method using quadratic programming to find the search direction and the starting point is $x^{(0)} = (0, 0)^T$. Write down the system of equations and all steps needed to solve.

$$\begin{aligned} \text{minimize} \quad & f(x_1, x_2) = (x_1 - 1)^2 + (x_2 - 2.5)^2 \\ \text{subject to} \quad & x_1 - 2x_2 \leq 0 \\ & -x_1 - 2x_2 + 6 \leq 0 \\ & -x_1 \leq 0 \end{aligned}$$

Hints:

- The result of the first step: $\begin{bmatrix} \triangle x \\ \lambda \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ -2.5 \\ -4.5 \end{bmatrix}$
- The result of the second step: $\begin{bmatrix} \triangle x \\ \lambda \end{bmatrix} = \begin{bmatrix} 1.8 \\ 0.9 \\ -1.6 \end{bmatrix}$
- Stop when the working set is empty

4C. Cutting Plans and Barrier Methods (3 points)

Explain in your words 3 main differences between Cutting plans methods and Barrier Methods