TMA947 Assignment Exercises 3

Edoardo Mangia

September 2024

Contents

1 HW 3.1

Consider the problem

$$\begin{aligned} \text{minimize} \quad & \frac{1}{2}x_1^2 + \frac{1}{2}x_2^2 && (3.1) \\ \text{subject to} \quad & x_1 \leq 4, && (3.2) \\ & x_2 \leq 4, && (3.3) \\ & x_1 + x_2 \geq 4. && (3.4) \end{aligned}$$

Lagrangian relax constraint (3.4), and formulate the dual function $q(\mu)$ explicitly. State the dual problem, and verify that it is a convex optimization problem.

Solution

Lagrangian Relaxation

The constraint $x_1 + x_2 \geq 4$ is relaxed using a Lagrange multiplier $\mu \geq 0$. The Lagrangian function is:

$$L(x_1, x_2, \mu) = \frac{1}{2}x_1^2 + \frac{1}{2}x_2^2 + \mu(4 - x_1 - x_2).$$

Dual Function

The dual function $q(\mu)$ is obtained by minimizing the Lagrangian over the primal variables x_1 and x_2 subject to the remaining constraints:

$$q(\mu) = \inf_{x_1, x_2} \left\{ \frac{1}{2}x_1^2 + \frac{1}{2}x_2^2 + \mu(4 - x_1 - x_2) \text{ subject to } x_1 \leq 4,\ x_2 \leq 4 \right\}.$$

Partial derivatives of L with respect to x_1 and x_2 are taken and set to zero to find the stationary points:

$$\frac{\partial L}{\partial x_1} = x_1 - \mu = 0 \quad \Rightarrow \quad x_1 = \mu,$$

$$\frac{\partial L}{\partial x_2} = x_2 - \mu = 0 \quad \Rightarrow \quad x_2 = \mu.$$

Applying the Constraints

The conditions $x_1 = \mu$ and $x_2 = \mu$ are checked against the constraints $x_1 \leq 4$ and $x_2 \leq 4$.

- If $\mu \leq 4$, then $x_1 = \mu$ and $x_2 = \mu$ satisfy the constraints.
- If $\mu > 4$, then $x_1 = 4$ and $x_2 = 4$ are set.

Calculating $q(\mu)$

- **Case 1:** $\mu \leq 4$

$$q(\mu) = \frac{1}{2}\mu^2 + \frac{1}{2}\mu^2 + \mu(4 - \mu - \mu) = -\mu^2 + 4\mu.$$

- **Case 2:** $\mu > 4$

$$q(\mu) = \frac{1}{2}(4)^2 + \frac{1}{2}(4)^2 + \mu(4 - 4 - 4) = 16 - 4\mu.$$

Thus, the dual function $q(\mu)$ is given by:

$$q(\mu) = \begin{cases} -\mu^2 + 4\mu & \text{if } \mu \leq 4, \\ 16 - 4\mu & \text{if } \mu > 4. \end{cases}$$

Dual Problem

The dual problem is formulated as:

$$\text{Maximize} \quad q(\mu) = \min_{\mu \geq 0} \begin{cases} -\mu^2 + 4\mu & \text{if } \mu \leq 4, \\ 16 - 4\mu & \text{if } \mu > 4. \end{cases}$$

Convexity of the Dual Problem

The dual problem is verified to be a convex optimization problem based on the concavity of the dual function $q(\mu)$:

- For $\mu \leq 4$, $q(\mu) = -\mu^2 + 4\mu$ is a concave quadratic function.
- For $\mu > 4$, $q(\mu) = 16 - 4\mu$ is a linear (and thus concave) function.

Since the dual function is concave and the dual problem involves maximizing $q(\mu)$ subject to $\mu \geq 0$, the dual problem is a convex optimization problem.

2 E 3.4

Consider the problem to

$$\begin{aligned} \text{minimize} \quad & -x_1^3 + x_2^2 - 2x_1x_2x_3, && (1) \\ \text{subject to} \quad & 2x_1 + x_2^2 + x_3 = 5, && (2) \\ & 5x_1^2 - x_2^2 - x_3 \leq 2, && (3) \\ & x_1, x_2, x_3 \geq 0. && (4) \end{aligned}$$

(a) State the KKT conditions for the problem.

(b) Verify that the KKT conditions are satisfied at $(1, 0, 3)^T$.

a)

The given optimization problem is:

$$\text{Minimize } f(x) = -x_1^3 + x_2^2 - 2x_1x_2x_3$$

subject to:

$$g_1(x) = 2x_1 + x_2^2 + x_3 - 5 = 0$$

$$h_1(x) = 5x_1^2 - x_2^2 - x_3 - 2 \geq 0$$

$$x_1 \geq 0, \quad x_2 \geq 0, \quad x_3 \geq 0$$

Lagrangian Function

The Lagrangian function $\mathcal{L}(x, \lambda, \mu)$ is defined as:

$$\mathcal{L}(x, \lambda, \mu) = f(x) + \lambda g_1(x) - \mu h_1(x)$$

where:

- λ is the Lagrange multiplier for the equality constraint $g_1(x) = 0$.
- μ is the Lagrange multiplier for the inequality constraint $h_1(x) \geq 0$.

Substituting the given functions:

$$\mathcal{L}(x_1, x_2, x_3, \lambda, \mu) = -x_1^3 + x_2^2 - 2x_1x_3^2 + \lambda(2x_1 + x_2^2 + x_3 - 5) - \mu(5x_1^2 - x_2^2 - x_3 - 2)$$

KKT Conditions

The Karush-Kuhn-Tucker (KKT) conditions are necessary conditions for a solution $x = (x_1, x_2, x_3)$ to be optimal for a problem with equality and inequality constraints. The conditions include:

1. **Stationarity:**

   The gradient of the Lagrangian with respect to x must be zero at the optimal point:

   $$\nabla_x \mathcal{L}(x, \lambda, \mu) = \nabla f(x) + \lambda \nabla g_1(x) - \mu \nabla h_1(x) = 0$$

   This condition can be broken down into components:

   $$\frac{\partial \mathcal{L}}{\partial x_1} = -3x_1^2 - 2x_3^2 + 2\lambda - 10\mu x_1 = 0$$
   $$\frac{\partial \mathcal{L}}{\partial x_2} = 2x_2 + 2\lambda x_2 + 2\mu x_2 = 0$$
   $$\frac{\partial \mathcal{L}}{\partial x_3} = -4x_1x_3 + \lambda + \mu = 0$$

2. **Primal Feasibility:**

   The constraints must be satisfied at the optimal point:

   $$g_1(x) = 2x_1 + x_2^2 + x_3 - 5 = 0$$

   $$h_1(x) = 5x_1^2 - x_2^2 - x_3 - 2 \geq 0$$

$$x_1 \geq 0, \quad x_2 \geq 0, \quad x_3 \geq 0$$

3. **Dual Feasibility:**

   The Lagrange multiplier μ associated with the inequality constraint must be non-negative:

$$\mu \geq 0$$

4. **Complementary Slackness:**

   The product of the multiplier μ and the constraint $h_1(x)$ must be zero:

$$\mu \cdot h_1(x) = 0$$

   This condition ensures that either the multiplier is zero, or the constraint is active (i.e., $h_1(x) = 0$).

b)

The point $(x_1, x_2, x_3) = (1, 0, 3)$ is verified against the KKT conditions.

1. Calculation of $f(x)$, $g_1(x)$, and $h_1(x)$ at $(1, 0, 3)$:

$$\begin{aligned}
f(1,0,3) &= -1^3 + 0^2 - 2(1)(3)^2 = -1 - 18 = -19 \\
g_1(1,0,3) &= 2(1) + 0^2 + 3 - 5 = 0 \\
h_1(1,0,3) &= 5(1)^2 - 0^2 - 3 - 2 = 0
\end{aligned}$$

Thus, $g_1(x) = 0$ and $h_1(x) = 0$.

2. Calculation of Gradients $\nabla f(x)$, $\nabla g_1(x)$, and $\nabla h_1(x)$:

$$\nabla f(x) = \left[-3x_1^2 - 2x_3^2,\ 2x_2,\ -4x_1x_3\right]$$

At $(1, 0, 3)$:

$$\nabla f(1,0,3) = \left[-3(1)^2 - 2(3)^2,\ 2(0),\ -4(1)(3)\right] = [-3 - 18,\ 0,\ -12] = [-21,\ 0,\ -12]$$

$$\nabla g_1(x) = [2,\ 2x_2,\ 1]$$

At $(1, 0, 3)$:

$$\nabla g_1(1,0,3) = [2,\ 0,\ 1]$$

$$\nabla h_1(x) = [10x_1,\ -2x_2,\ -1]$$

At $(1,0,3)$:

$$\nabla h_1(1,0,3) = [10,\ 0,\ -1]$$

3. Stationarity Condition:

$$\nabla f(x) + \lambda \nabla g_1(x) - \mu \nabla h_1(x) = 0$$

Substituting the values:

$$[-21,\ 0,\ -12] + \lambda\,[2,\ 0,\ 1] - \mu\,[10,\ 0,\ -1] = 0$$

This results in the following equations:

$$\begin{aligned} -21 + 2\lambda - 10\mu &= 0 \quad (1) \\ 0 + 0\lambda + 0\mu &= 0 \quad (2) \\ -12 + \lambda + \mu &= 0 \quad (3) \end{aligned}$$

From equation (2), it is satisfied for any values of λ and μ. Solving equations (1) and (3) together:

From equation (1):

$$2\lambda - 10\mu = 21 \quad \Rightarrow \quad \lambda = \frac{21 + 10\mu}{2}$$

Substituting λ into equation (3):

$$-12 + \left(\frac{21 + 10\mu}{2}\right) + \mu = 0$$

Simplifying:

$$-12 + 10.5 + 5\mu + \mu = 0 \quad \Rightarrow \quad -1.5 + 6\mu = 0 \quad \Rightarrow \quad \mu = 0.25$$

Substituting $\mu = 0.25$ back into the expression for λ:

$$\lambda = \frac{21 + 10(0.25)}{2} = \frac{21 + 2.5}{2} = 11.75$$

Thus, the values of the multipliers are:

$$\lambda = 11.75, \quad \mu = 0.25$$

4. Verification of Complementary Slackness:

The complementary slackness condition states:

$$\mu \cdot h_1(x) = 0$$

Given:

$$h_1(1, 0, 3) = 0 \quad \text{and} \quad \mu = 0.25$$

Thus, the complementary slackness condition is satisfied because:

$$0.25 \cdot 0 = 0$$

3 E 3.10

A company produces loading pallets in two models: standard and extra long. Each model consists of three cross-sectional beams with a length corresponding to the length of the pallet. The details are as follows:

- **Standard Model:**
  - Number of boards on top: 5
  - Number of boards on bottom: 5
  - Total boards: $5 + 5 = 10$
  - Time to assemble: 0.25 hours
  - Total production time for boards: $10 \cdot 0.002 = 0.02$ hours
  - Total production time for beams: $3 \cdot 0.005 = 0.015$ hours
  - Total production time: $0.02 + 0.015 = 0.035$ hours
- **Extra Long Model:**
  - Number of boards on each side: 9 (on the top and bottom)
  - Total boards: $9 + 9 = 18$

- Time to assemble: 0.30 hours
- Total production time for boards: $18 \cdot 0.002 = 0.036$ hours
- Total production time for beams: $3 \cdot 0.007 = 0.021$ hours
- Total production time: $0.036 + 0.021 = 0.057$ hours

LP Formulation

Decision Variables

Let:

$$\begin{aligned} x_1 &: \text{ Number of standard pallets produced} \\ x_2 &: \text{ Number of extra long pallets produced} \end{aligned}$$

Objective Function

Maximize the profit:

$$\text{Maximize } Z = 50x_1 + 70x_2$$

Constraints

Assembly Time Constraint:

$$0.25x_1 + 0.30x_2 \leq 200$$

Production Time Constraint:

$$0.035x_1 + 0.057x_2 \leq 40$$

Non-negativity Constraints:

$$x_1 \geq 0, \quad x_2 \geq 0$$

Complete LP Formulation

The complete LP problem is formulated as:

$$\begin{aligned} \text{Maximize } & Z = 50x_1 + 70x_2 \\ \text{subject to } & \\ & 0.25x_1 + 0.30x_2 \leq 200 \\ & 0.035x_1 + 0.057x_2 \leq 40 \\ & x_1 \geq 0 \\ & x_2 \geq 0 \end{aligned}$$

Lagrangian Dual Problem

The Lagrangian is defined as:

$$L(x_1, x_2, \lambda_1, \lambda_2) = 50x_1+70x_2+\lambda_1(200-0.25x_1-0.30x_2)+\lambda_2(40-0.035x_1-0.057x_2)$$

Dual Problem

1. Differentiate L with respect to x_1 and x_2 and set the derivatives equal to zero:

$$\frac{\partial L}{\partial x_1} = 50 - 0.25\lambda_1 - 0.035\lambda_2 = 0$$
$$\frac{\partial L}{\partial x_2} = 70 - 0.30\lambda_1 - 0.057\lambda_2 = 0$$

2. The dual problem is expressed as:

$$\begin{aligned} \text{Minimize } & W = 200\lambda_1 + 40\lambda_2 \\ \text{subject to} & \\ & 0.25\lambda_1 + 0.035\lambda_2 \geq 50 \\ & 0.30\lambda_1 + 0.057\lambda_2 \geq 70 \\ & \lambda_1 \geq 0 \\ & \lambda_2 \geq 0 \end{aligned}$$

Global Primal-Dual Optimality Conditions

The global primal-dual optimality conditions state that:

- The primal and dual variables satisfy their respective constraints.
- The complementary slackness conditions:
  - If $x_1 > 0$, then λ_1 must satisfy $200 - 0.25x_1 - 0.30x_2 = 0$.
  - If $x_2 > 0$, then λ_2 must satisfy $40 - 0.035x_1 - 0.057x_2 = 0$.
- The duality gap is zero at optimality.

Primal-Dual Optimality Conditions

For a solution (x_1^*, x_2^*) of the primal problem and $(\lambda_1^*, \lambda_2^*)$ of the dual problem to be optimal, the following conditions must be satisfied:

1. **Primal Problem**

$$\begin{aligned} \max \quad & 50x_1 + 70x_2 \\ \text{s.t.} \quad & 0.25x_1 + 0.3x_2 \leq 200 \quad \text{(Constraint 1)} \\ & 0.035x_1 + 0.057x_2 \leq 40 \quad \text{(Constraint 2)} \\ & x_1, x_2 \geq 0. \end{aligned}$$

2. **Dual Problem**

$$\begin{aligned} \min_{\lambda_1,\lambda_2} \quad & 200\lambda_1 + 40\lambda_2 \\ \text{s.t.} \quad & 0.25\lambda_1 + 0.035\lambda_2 \geq 50, \\ & 0.3\lambda_1 + 0.057\lambda_2 \geq 70, \\ & \lambda_1, \lambda_2 \geq 0. \end{aligned}$$

3. **Primal-Dual Optimality Conditions (KKT Conditions)**

- **Primal Feasibility:** The primal variables x_1^* and x_2^* must satisfy the constraints of the primal problem:

$$\begin{aligned} 0.25x_1^* + 0.3x_2^* &\leq 200, \\ 0.035x_1^* + 0.057x_2^* &\leq 40, \\ x_1^* &\geq 0, \\ x_2^* &\geq 0. \end{aligned}$$

- **Dual Feasibility:** The dual variables λ_1^* and λ_2^* must satisfy the constraints of the dual problem:

$$\begin{aligned} 0.25\lambda_1^* + 0.035\lambda_2^* &\geq 50, \\ 0.3\lambda_1^* + 0.057\lambda_2^* &\geq 70, \\ \lambda_1^* &\geq 0, \\ \lambda_2^* &\geq 0. \end{aligned}$$

- **Complementary Slackness:** For each inequality constraint, either the constraint is active (i.e., holds with equality) or the corresponding Lagrange multiplier is zero:

$$\begin{aligned} \lambda_1^* (200 - 0.25x_1^* - 0.3x_2^*) &= 0, \\ \lambda_2^* (40 - 0.035x_1^* - 0.057x_2^*) &= 0. \end{aligned}$$

- **Stationarity:** The gradient of the Lagrangian with respect to x_1 and x_2 must be zero at the optimal points (x_1^*, x_2^*):

$$\begin{aligned} \frac{\partial \mathcal{L}}{\partial x_1} &= 50 - 0.25\lambda_1^* - 0.035\lambda_2^* = 0, \\ \frac{\partial \mathcal{L}}{\partial x_2} &= 70 - 0.3\lambda_1^* - 0.057\lambda_2^* = 0. \end{aligned}$$

4 E 3.11

Consider the problem to minimize $f(x)$ subject to $x \in S$, where $f : \mathbb{R}^n \to \mathbb{R}$ is a C^1 function, and S is some subset of $\mathbb{R}^n$. Assume that $\bar{x}$ satisfies the geometric optimality condition, and let γ be an arbitrary smooth curve through S starting at $\bar{x} \in S$, i.e., $\gamma : [0,1] \to S$ is a differentiable function with $\gamma(0) = \bar{x}$. Show that

$$\frac{d}{dt} f(\gamma(t))\bigg|_{t=0} \leq 0.$$

The definition of the tangent cone is:

$$T_S(x) := \{p \mid \exists \{x_k\}_{k=1}^{\infty} \subset S, \{\lambda_k\}_{k=1}^{\infty} \subset (0,\infty), \text{ such that } \lim_{k\to\infty} x_k = x, \quad \lim_{k\to\infty} \lambda_k(x_k - x) = p\}$$

In simpler terms, it captures the directions in which one can move from the point x while remaining in the set S, scaled by small positive factors.

Then, $\gamma'(0) \in T_S(x)$ since the definition of the (one-sided) derivative is

$$\gamma'(0) = \lim_{t\to 0} \frac{\gamma(t) - \gamma(0)}{t}. \tag{4}$$

Fixing any sequence $t_k \to 0$, and letting $x_k := \gamma(t_k)$, $\lambda_k = \frac{1}{t_k}$, the sequences required in the definition of $T_S(x)$ are defined.

A notion of descent directions to the objective function f can be formulated. The cone of descent directions is a concept used to describe the set of feasible directions in which one can move from a current point in order to decrease the objective function value.

$$\overset{\circ}{F}(x) := \{p \in \mathbb{R}^n \mid \nabla f(x)^\top p < 0\}$$

Theorem 1 (Geometric Optimality Conditions) *Consider the problem (1), where $f \in C^1$. Then*

$$x^* \text{ is a local minimum of } f \text{ over } S \implies \overset{\circ}{F}(x^*) \cap T_S(x^*) = \emptyset.$$

If the example with smooth curves is returned to, it is shown that for any smooth curve γ through S starting at x^*, $\gamma'(0) \in T_S(x^*)$. The geometric optimality condition reduces to the statement that

$$\frac{d}{dt}\bigg|_{t=0} f(\gamma(t)) \geq 0$$

when applied to this tangent vector.