

**TMA947/MMG621
OPTIMIZATION, BASIC COURSE**

- Date:** 15-08-27
- Time:** House V, morning, 8³⁰-13³⁰
- Aids:** Text memory-less calculator, English-Swedish dictionary
- Number of questions:** 7; passed on one question requires 2 points of 3.
Questions are *not* numbered by difficulty.
To pass requires 10 points and three passed questions.
- Examiner:** Michael Patriksson
- Teacher on duty:** Åse Fahlander (0703-088304)
- Result announced:** 15-09-18
Short answers are also given at the end of
the exam on the notice board for optimization
in the MV building.

Exam instructions

When you answer the questions

*Use generally valid theory and methods.
State your methodology carefully.*

*Only write on one page of each sheet. Do not use a red pen.
Do not answer more than one question per page.*

At the end of the exam

*Sort your solutions by the order of the questions.
Mark on the cover the questions you have answered.
Count the number of sheets you hand in and fill in the number on the cover.*

Question 1

(the simplex method)

Consider the following linear program:

$$\begin{aligned}
 &\text{minimize} && z = 2x_1 - x_2, \\
 &\text{subject to} && x_1 + x_2 \geq 1, \\
 & && x_1 - 2x_2 \leq 1. \\
 & && x_1 \geq 0, \\
 & && x_2 \geq 0.
 \end{aligned}$$

- (2p) a) Solve the problem using phase I and phase II of the simplex method.
Aid: Utilize the identity

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}^{-1} = \frac{1}{ad - bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}.$$

- (1p) b) Does its LP dual have an optimal solution?

(3p) Question 2

(linear inequalities)

Consider the system of linear inequalities

$$\mathbf{Ax} \leq \mathbf{b},$$

for which we assume there is at least one solution. Let d be a given scalar. Use linear programming duality to establish the equivalence of the following two statements:

- (a) Every solution \mathbf{x} to the system $\mathbf{Ax} \leq \mathbf{b}$ satisfies $\mathbf{c}^T \mathbf{x} \leq d$.
 (b) There exists some vector $\mathbf{y} \geq \mathbf{0}$ such that $\mathbf{A}^T \mathbf{y} = \mathbf{c}$ and $\mathbf{b}^T \mathbf{y} \leq d$.

(3p) Question 3

(the Frank–Wolfe algorithm)

Consider the problem to

$$\begin{aligned} \underset{x_1, x_2}{\text{minimize}} \quad & (x_1 \ x_2) \begin{pmatrix} 6 & 2 \\ 2 & 9 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} - (52 \ 34) \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \\ \text{subject to} \quad & x_1 + 2x_2 \leq 4. \\ & x_1 + x_2 \leq 3. \\ & 2x_1 \leq 5. \\ & x_1 \geq 0. \\ & x_2 \geq 0. \end{aligned} \tag{1}$$

Solve the problem (1) using the Frank–Wolfe algorithm. Start with the initial guess $\mathbf{x}^{(0)} = (x_1, x_2)^T = (2.5, 0)^T$. The line search should be performed as an exact minimization. If necessary, you are allowed to carry out the calculations approximately with two digits of accuracy.

Hint: You may find it helpful to analyze the problem and the algorithm progress graphically, but this must be augmented with a rigorous analysis.

(3p) Question 4

(modelling)

A small municipality is forced to close one or several schools. Out of ten existing schools, at most three schools can be closed. The annual cost to keep school i open is c_i kr. School i can educate a maximum of k_i students. The municipality is divided into J home areas and there is a requirement that all students in an area belong to the same school. There are b_j students in area j and the average distance from area j to school i is d_{ij} km. The estimated annual cost for student travels is set to m kr per km and student.

Formulate a linear integer program to decide on which schools to keep and which ones to close, such that we minimize the total cost for schools and travels and fulfill the above listed requirements.

Question 5

(true or false)

The below three claims should be assessed. Are they true or false? Provide an answer together with a short but complete motivation.

- (1p) a) *Claim:* A strictly convex function is differentiable.
- (1p) b) *Claim:* For a constrained minimization problem with explicit constraints, any Lagrangian dual formulation provides an upper bound on the optimal value of the original problem.
- (1p) c) *Claim:* In linear programming, at termination of the Simplex method the optimal values of the dual variables are equal to the Lagrange multipliers of the linear constraints in the (primal) linear program.

Question 6

(interior penalty methods)

Consider the problem to

$$\begin{aligned} & \text{minimize } f(\mathbf{x}) := (x_1 - 2)^4 + (x_1 - 2x_2)^2, \\ & \text{subject to } g(\mathbf{x}) := x_1^2 - x_2 \leq 0. \end{aligned}$$

We attack this problem with an interior penalty (barrier) method, using the barrier function $\phi(s) = -s^{-1}$. The penalty problem is to

$$\underset{x \in \mathbb{R}^n}{\text{minimize}} \quad f(\mathbf{x}) + \nu \hat{\chi}_S(\mathbf{x}), \quad (1)$$

where $\hat{\chi}_S(\mathbf{x}) = \phi(g(\mathbf{x}))$, for a sequence of positive, decreasing values of the penalty parameter ν .

We repeat a general convergence result for the interior penalty method below.

THEOREM 1 (convergence of an interior point algorithm) *Let the objective function $f : \mathbb{R}^n \rightarrow \mathbb{R}$ and the functions g_i , $i = 1, \dots, m$, defining the inequality*

constraints be in $C^1(\mathbb{R}^n)$. Further assume that the barrier function $\phi : \mathbb{R}_- \rightarrow \mathbb{R}_+$ is in C^1 and that $\phi'(s) \geq 0$ for all $s < 0$.

Consider a sequence $\{\mathbf{x}_k\}$ of points that are stationary for the sequence of problems (1) with $\nu = \nu_k$, for some positive sequence of penalty parameters $\{\nu_k\}$ converging to 0. Assume that $\lim_{k \rightarrow +\infty} \mathbf{x}_k = \hat{\mathbf{x}}$, and that LICQ holds at $\hat{\mathbf{x}}$. Then, $\hat{\mathbf{x}}$ is a KKT point of the problem at hand.

In other words,

$$\left. \begin{array}{l} \mathbf{x}_k \text{ stationary in (1)} \\ \mathbf{x}_k \rightarrow \hat{\mathbf{x}} \text{ as } k \rightarrow +\infty \\ \text{LICQ holds at } \hat{\mathbf{x}} \end{array} \right\} \implies \hat{\mathbf{x}} \text{ stationary in our problem.}$$

- (1p) a) Does the above theorem apply to the problem at hand and the selection of the penalty function?
- (2p) b) Implementing the above-mentioned procedure, the first value of the penalty parameter was set to $\nu_0 = 10$, which is then divided by ten in each iteration, and the initial problem (1) was solved from the strictly feasible point $(0, 1)^T$. The algorithm terminated after six iterations with the following results: $\mathbf{x}_6 \approx (0.94389, 0.89635)^T$, and the multiplier estimate [given by $\nu_6 \phi'(g(\mathbf{x}_6))$] $\hat{\mu}_6 \approx 3.385$. Confirm that the vector \mathbf{x}_6 is close to being a KKT point. Are the KKT point(s) globally optimal? Why/Why not?

Question 7

(the KKT conditions)

Consider the problem to

$$\begin{array}{ll} \text{minimize} & x_1 x_2 + x_2 x_3 + x_1 x_3 \\ \text{subject to} & x_1 + x_2 + x_3 = 3. \end{array}$$

- (2p) a) Write down the KKT conditions and find *all* KKT points.
- (1p) b) Does the problem have an optimal solution? Motivate!

**TMA947/MMG621
NONLINEAR OPTIMISATION**

Date: 15-08-27

Examiner: Michael Patriksson

Question 1

(the simplex method)

- (2p) a) We first rewrite the problem on standard form. We introduce slack variables s_1 and s_2 . Consider the following linear program:

$$\begin{aligned} \text{minimize} \quad & z = 2x_1 - x_2, \\ \text{subject to} \quad & x_1 + x_2 - s_1 = 1, \\ & x_1 - 2x_2 + s_2 = 1, \\ & x_1, \quad x_2, \quad s_1, \quad s_2 \geq 0. \end{aligned}$$

By introducing an artificial variable a , we get the Phase I problem to

$$\begin{aligned} \text{minimize} \quad & w = a, \\ \text{subject to} \quad & x_1 - 2x_2 + s_2 = 1, \\ & x_1 + x_2 - s_1 + a = 1, \\ & x_1, \quad x_2, \quad s_1, \quad s_2, \quad a \geq 0. \end{aligned}$$

The starting basis is $(s_2, a)^T$. Calculating the reduced costs for the non-basic variables x_1 , x_2 , and s_1 we obtain $\tilde{\mathbf{c}}_N = (-1, -1, 1)^T$, meaning that x_1 enters the basis. From the minimum ratio test, we get that a leaves the basis.

Updating the basis we now have $(s_2, x_1)^T$ in the basis meaning that $w^* = 0$ and the basis found is corresponding to a basic feasible solution of the original problem in the standard form, i.e., the Phase II problem.

Calculating the reduced costs, we obtain $\tilde{\mathbf{c}}_N = (-3, 2)^T$. meaning that x_2 enters the basis. From the minimum ratio test we get that x_1 leaves the basis.

Updating the basis we now have $(s_2, x_2)^T$ in the basis. Calculating the reduced costs, we obtain $\tilde{\mathbf{c}}_N = (3, -1)^T$, meaning that s_1 enters basis. From the minimum ratio test we get that $\mathbf{B}^{-1}\mathbf{N}_{s_1} = (-1, 0)^T \leq \mathbf{0}$, meaning that the problem is unbounded.

- (1p) b) The primal problem is unbounded, implying that $\mathbf{c}^T \mathbf{x}^* = -\infty$. From weak duality we have that $\mathbf{b}^T \mathbf{y} \leq \mathbf{c}^T \mathbf{x}^*$ for all feasible \mathbf{y} , meaning that the dual problem is infeasible.

(3p) Question 2

(linear inequalities)

Consider the linear program to

$$\begin{aligned} & \underset{x}{\text{minimize}} && -c^T x, \\ & \text{subject to} && Ax \leq b, \end{aligned} \tag{1}$$

and its standard form equivalence

$$\begin{aligned} & \underset{x^+, x^-, s}{\text{minimize}} && -c^T x^+ + c^T x^-, \\ & \text{subject to} && Ax^+ - Ax^- + s = b, \\ & && x^+ \geq 0, x^- \geq 0, s \geq 0. \end{aligned} \tag{2}$$

The dual of (2) is to

$$\begin{aligned} & \underset{p}{\text{maximize}} && b^T p, \\ & \text{subject to} && A^T p \leq -c, \\ & && -A^T p \leq c, \\ & && p \leq 0, \end{aligned} \tag{3}$$

and (3) is equivalent to

$$\begin{aligned} & \underset{y}{\text{maximize}} && -b^T y, \\ & \text{subject to} && A^T y = c, \\ & && y \geq 0. \end{aligned} \tag{4}$$

If statement (a) holds, then the objective of (1) and (2) is bounded from below by $-d$. Hence, there exists an optimal solution to the dual of (2), which is (3). Consequently, by strong duality (cf. Theorem 10.6 in the text) the optimal objective values of (3) and (4) are equal to that of (2), which is bounded from below by $-d$. This implies that, for (4), there exists a vector $y \geq 0$ such that $A^T y = c$ and $-b^T y \geq -d$ (i.e., $b^T y \leq d$). This statement is the same as (b).

Conversely, if (b) holds then (3) has at least one feasible solution with an objective value bounded from below by $-d$. Hence, by weak duality (cf. Theorem 10.4 in the text) every x feasible to (1) (i.e., $Ax \leq b$) must satisfy $-c^T x \geq -d$. This implies statement (a).

(3p) **Question 3**

(the Frank–Wolfe algorithm)

Figure 1 shows the feasible set of the problem (i.e., the polyhedron with thick black boundary lines) and some contours of the objective function. The optimal solution is denoted by x^* (i.e., the red dot in the figure). $x^{(k)}$ for $k = 0, 1, 2$ denotes iterates visited by the Frank-Wolfe algorithm.

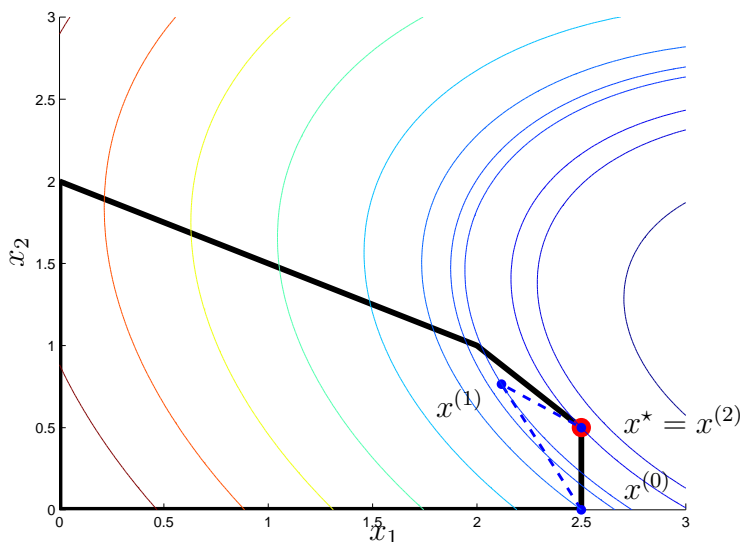


Figure 1: Illustration of the Frank-Wolfe algorithm. The feasible set is a polyhedron with boundary denoted by the thick black lines. Some contours of the objective function are shown. The optimal solution $x^* = (2.5, 0.5)$. The dotted lines show the Frank-Wolfe iterations, with $x^{(k)}$, $k = 0, 1, 2$ denoting the iterates.

The details of the algorithm steps are as follows. Let X denote the feasible set. Let $f(x_1, x_2)$ denote the objective function. For any given iterate $x^{(k)} = (x_1^{(k)}, x_2^{(k)})$. The objective function gradient vector is

$$\nabla f(x_1^{(k)}, x_2^{(k)}) = \begin{bmatrix} 12 & 4 \\ 4 & 18 \end{bmatrix} \begin{bmatrix} x_1^{(k)} \\ x_2^{(k)} \end{bmatrix} - \begin{bmatrix} 52 \\ 34 \end{bmatrix}.$$

The search direction problem is

$$\underset{x \in X}{\text{minimize}} \quad \nabla f(x_1^{(k)}, x_2^{(k)})^\top x. \tag{1}$$

If $\min_{x \in X} \nabla f(x_1^{(k)}, x_2^{(k)})^T x \geq \nabla f(x_1^{(k)}, x_2^{(k)})^T x^{(k)}$, then by optimality conditions (for minimizing a convex function over a convex feasible set) $x^{(k)}$ is optimal. Otherwise, let $y^{(k)}$ denote an optimal solution to the search direction problem. Then the exact minimization line search problem can be expressed into

$$\underset{\alpha \in [0,1]}{\text{minimize}} \quad f(\alpha x^{(k)} + (1 - \alpha)y^{(k)}) \iff \underset{\alpha \in [0,1]}{\text{minimize}} \quad g\alpha^2 + h\alpha,$$

where

$$\begin{aligned} g &= (x^{(k)} - y^{(k)})^T \begin{bmatrix} 6 & 2 \\ 2 & 9 \end{bmatrix} (x^{(k)} - y^{(k)}) \\ h &= (x^{(k)} - y^{(k)})^T \left(\begin{bmatrix} 12 & 4 \\ 4 & 18 \end{bmatrix} y^{(k)} - \begin{bmatrix} 52 \\ 34 \end{bmatrix} \right). \end{aligned} \quad (2)$$

The minimizing value of α , denoted by $\alpha^{(k)}$, can be found using the optimality condition to be

$$\alpha^{(k)} = \begin{cases} 0 & \text{if } -\frac{h}{2g} < 0 \\ -\frac{h}{2g} & \text{if } 0 \leq -\frac{h}{2g} \leq 1. \\ 1 & \text{if } -\frac{h}{2g} > 1 \end{cases} \quad (3)$$

The iterate update formula is

$$x^{(k+1)} = \alpha^{(k)}x^{(k)} + (1 - \alpha^{(k)})y^{(k)}. \quad (4)$$

Now we begin applying the Frank-Wolfe algorithm. At the first iteration with $x^{(0)} = (2.5, 0)$, the objective function gradient is

$$\nabla f(x_1^{(0)}, x_2^{(0)}) = \begin{bmatrix} 12 & 4 \\ 4 & 18 \end{bmatrix} \begin{bmatrix} x_1^{(0)} \\ x_2^{(0)} \end{bmatrix} - \begin{bmatrix} 52 \\ 34 \end{bmatrix} = \begin{bmatrix} 12 & 4 \\ 4 & 18 \end{bmatrix} \begin{bmatrix} 2.5 \\ 0 \end{bmatrix} - \begin{bmatrix} 52 \\ 34 \end{bmatrix} = \begin{bmatrix} -22 \\ -24 \end{bmatrix}.$$

To solve the search direction problem in (1), it is sufficient to restrict the feasible set to the set of all extreme points. That is,

$$\underset{x \in V}{\text{minimize}} \quad \nabla f(x_1^{(0)}, x_2^{(0)})^T x, \quad (5)$$

where V is the set of all extreme points defined as

$$V = \left\{ (0, 0), (0, 2), (2, 1), (2.5, 0.5), (2.5, 0) \right\}.$$

This amounts to finding the minimum among five numbers: 0, -48, -68, -67, -55. The result is that $y^{(0)} = (2, 1)$. Applying the formula in (??) yields

$$\begin{aligned} g &= \left(\begin{bmatrix} 2.5 \\ 0 \end{bmatrix} - \begin{bmatrix} 2 \\ 1 \end{bmatrix} \right)^T \begin{bmatrix} 6 & 2 \\ 2 & 9 \end{bmatrix} \left(\begin{bmatrix} 2.5 \\ 0 \end{bmatrix} - \begin{bmatrix} 2 \\ 1 \end{bmatrix} \right) = 8.5 \\ h &= \left(\begin{bmatrix} 2.5 \\ 0 \end{bmatrix} - \begin{bmatrix} 2 \\ 1 \end{bmatrix} \right)^T \left(\begin{bmatrix} 12 & 4 \\ 4 & 18 \end{bmatrix} \begin{bmatrix} 2 \\ 1 \end{bmatrix} - \begin{bmatrix} 52 \\ 34 \end{bmatrix} \right) = -4 \end{aligned}$$

According to (3), $\alpha^{(0)} = \frac{4}{17}$. Hence, by (4)

$$x^{(1)} = \frac{4}{17}\left(\frac{5}{2}, 0\right) + \left(1 - \frac{4}{17}\right)(2, 1) = \left(\frac{36}{17}, \frac{13}{17}\right) \approx (2.12, 0.76).$$

This is shown in Figure 1.

At the next iteration with $x^{(1)} = \left(\frac{36}{17}, \frac{13}{17}\right)$, we have

$$\nabla f(x_1^{(1)}, x_2^{(1)}) = \begin{bmatrix} 12 & 4 \\ 4 & 18 \end{bmatrix} \begin{bmatrix} x_1^{(1)} \\ x_2^{(1)} \end{bmatrix} - \begin{bmatrix} 52 \\ 34 \end{bmatrix} = \frac{1}{17} \begin{bmatrix} -400 \\ -200 \end{bmatrix} \approx \begin{bmatrix} -23.53 \\ -11.76 \end{bmatrix}.$$

Solving (5) amounts to finding the minimum of 0, -4, -10, -11, -10. This leads to $y^{(1)} = (2.5, 0.5)$. Applying (2) leads to

$$\begin{aligned} g &= \frac{1275}{1156} \approx 1.10 \\ h &= \frac{125}{34} \approx 3.68. \end{aligned}$$

Thus, according to (3) $\alpha^{(1)} = 0$, and from (4) $x^{(2)} = y^{(1)} = (2.5, 0.5)$ as shown in Figure 1.

At the final iteration with $x^{(2)} = (2.5, 0.5)$, we have

$$\nabla f(x_1^{(2)}, x_2^{(2)}) = \begin{bmatrix} -20 \\ -15 \end{bmatrix}.$$

Solving (5) leads to $y^{(2)} = x^{(2)} = (2.5, 0.5)$. Thus, it holds that

$$\min_{x \in X} \nabla f(x_1^{(2)}, x_2^{(2)})^T x \geq \nabla f(x_1^{(2)}, x_2^{(2)})^T x^{(2)}.$$

By optimality conditions, $x^{(2)} = (2.5, 0.5)$ is the optimal solution to our problem.

(3p) Question 4

(modelling)

The decision variables are:

$y_i = 1$, if school i is open, 0 otherwise

$x_{ij} = 1$, if students in area j attend school i , 0 otherwise

Model

$$\begin{aligned}
 & \text{minimize} && \sum_{i=1}^{10} c_i y_i + 2m \sum_{i=1}^{10} \sum_{j=1}^J b_j d_{ij} x_{ij} \\
 & \text{subject to} && \sum_{j=1}^J b_j x_{ij} \leq k_i y_i, \quad i = 1, \dots, 10 \\
 & && \sum_{i=1}^{10} y_i \geq 7, \\
 & && \sum_{i=1}^{10} x_{ij} = 1, \quad j = 1, \dots, J \\
 & && y_i \in \{0, 1\}, \quad i = 1, \dots, 10 \\
 & && x_{ij} \in \{0, 1\}, \quad i = 1, \dots, 10 \\
 & && \quad \quad \quad j = 1, \dots, J
 \end{aligned}$$

The program is linear with integer variables.

Question 5

(true or false)

- (1p) a) False. A simple example has $f(x) = x^2$ for $x \leq 0$, and $x^3 + |x|$ for $x \geq 0$.
- (1p) b) False. It provides a lower bound on the optimal value of the original (primal) problem.
- (1p) c) True. Theorem 10.15 (necessary and sufficient conditions for global optimality) shows that an optimal dual solution is a vector of Lagrange multipliers.

Question 6

(interior penalty methods)

- (1p) a) All functions involved are in C^1 . The conditions on the penalty function are fulfilled, since $\phi'(s) = 1/s^2 \geq 0$ for all $s < 0$. Further, LICQ holds everywhere. The answer is yes.

- (2p) b) With the given data, it is clear that the only constraint is (almost) fulfilled with equality: $(\mathbf{x}_6)_1^2 - (\mathbf{x}_6)_2 \approx -0.005422 \approx 0$. We set up the KKT conditions to see whether it is fulfilled approximately. Indeed, we have the following corresponding to the system $\nabla f(\mathbf{x}_6) + \hat{\mu}_6 \nabla g(\mathbf{x}_6) = \mathbf{0}^2$:

$$\begin{pmatrix} -6.4094265 \\ 3.39524 \end{pmatrix} + 3.385 \begin{pmatrix} 1.88778 \\ -1 \end{pmatrix} \approx \begin{pmatrix} -0.01929 \\ 0.01024 \end{pmatrix},$$

and the right-hand side can be considered near-zero. Since $\hat{\mu}_6 \geq 0$ we approximately fulfill the KKT conditions.

For the last part, we establish that the problem is convex. The feasible set clearly is convex, since g is a convex function and the constraint is on the “ \leq ”-form. The Hessian matrix of f is

$$\begin{pmatrix} 12(x_1 - 2)^2 + 2 & -4 \\ -4 & 8 \end{pmatrix},$$

which is positive semidefinite everywhere (in fact, positive definite outside of the region defined by $x_1 = 2$); hence, f is convex on \mathbb{R}^2 . We conclude that our problem is convex, and hence the KKT conditions imply global optimality. The vector \mathbf{x}_6 therefore is an approximate global optimal solution to our problem.

Question 7

(the KKT conditions)

- (2p) a) The KKT conditions are

$$\nabla f(\mathbf{x}) + \lambda \nabla h(\mathbf{x}) = \begin{pmatrix} x_2 + x_3 \\ x_1 + x_3 \\ x_1 + x_2 \end{pmatrix} + \lambda \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = \mathbf{0}.$$

There is only one feasible point fulfilling the KKT conditions:

$$\bar{\mathbf{x}} = (1, 1, 1)^T \text{ with } \lambda = -2.$$

- (1p) b) Since the eigenvalues of the Hessian of the objective function

$$\nabla^2 f(\mathbf{x}) = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}$$

are $\lambda_1 = \lambda_2 = -1, \lambda_3 = 2$ the objective function is not convex, indicating that the problem is unbounded. The KKT point $\bar{\mathbf{x}}$ is not an optimal solution.

