Chalmers/GU
Mathematics

EXAM

TMA947/MAN280
APPLIED OPTIMIZATION

Date:	06–08–31
Time:	House V, morning
Aids:	Text memory-less calculator
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:	Elisabeth Wulcan (0762-721860)
Result announced:	06–09–19 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:

$$\begin{array}{llrrrl} \text{minimize} & z = & x_1 & +\alpha x_2 & +x_3, & \\ \text{subject to} & & x_1 & +2x_2 & -2x_3 \leq & 0, \\ & & -x_1 & & +x_3 \leq & -1, \\ & & x_1, & x_2, & x_3 \geq & 0. \end{array}$$

(2p) a) Solve this problem for $\alpha = -1$ by using phase I and phase II of the simplex method.

[*Aid:* Utilize the identity

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

for producing basis inverses.]

(1p) b) Which values of α leads to an unbounded dual problem? Motivate without additional calculations!

(3p) Question 2

(necessary local and sufficient global optimality conditions)

Consider an optimization problem of the following general form:

$$\text{minimize } f(\boldsymbol{x}), \tag{1a}$$

$$\text{subject to } \boldsymbol{x} \in S, \tag{1b}$$

where $S \subseteq \mathbb{R}^n$ is nonempty, closed and convex, and $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ is in C^1 on S.

Establish the following two results on the local/global optimality of a vector $\boldsymbol{x}^* \in S$ in this problem.

Proposition 1 (necessary optimality conditions, C^1 case) *If $\boldsymbol{x}^* \in S$ is a local minimum of f over S then*

$$\nabla f(\boldsymbol{x}^*)^{\mathrm{T}}(\boldsymbol{x} - \boldsymbol{x}^*) \geq 0, \qquad \boldsymbol{x} \in S \tag{2}$$

holds.

THEOREM 2 (necessary and sufficient global optimality conditions, C^1 case) *Suppose that $f : \mathbb{R}^n \to \mathbb{R}$ is convex on S. Then,*

$$\boldsymbol{x}^* \text{ is a global minimum of } f \text{ over } S \quad \Longleftrightarrow \quad (2) \text{ holds.}$$

Question 3

(Newton's method revisited)

Consider the unconstrained optimization problem to

$$\begin{aligned} &\text{minimize } f(\boldsymbol{x}), \\ &\text{subject to } \boldsymbol{x} \in \mathbb{R}^n, \end{aligned}$$

where $f : \mathbb{R}^n \to \mathbb{R}$ is in C^1 on $\mathbb{R}^n$.

Notice that we may not have access to second derivatives of f at every point of $\mathbb{R}^n$. "Newton's method" referred to below should be understood as follows: in each iteration step, one solves the Newton equation, followed by a line search with respect to f in the direction obtained.

(2p) a) Explain in some detail how Newton's method can be extended to the above problem.

(1p) b) Suppose now that $f \in C^2$ on $\mathbb{R}^n$. Explain why Newton's method *must* be modified when the Hessian matrix is not guaranteed to be positive definite. Also, provide *at least* one such modification.

(3p) Question 4

(modelling)

You are responsible for the planning of a soccer tournament where all 14 teams in the Swedish national league will participate. The teams shall be put into two groups of 7 each, in which all teams will play each other once. The winners of the two groups will then play a final. The decision to make is which teams will play in which group. The objective is to minimize the total expected travelling distance. The distances between the home towns of two teams i and j are given by the constants $d_{ij}(= d_{ji}),\ i,j \in \{1,\dots,14\}$. The constants $p_i,\ i \in \{1,\dots,14\}$, represent the number of points team i took in the national league last year. Assume that the teams are sorted so that the team with the highest point is represented by $i = 1$, the team with second highest point by $i = 2$, and so on. The chance of a team i winning its group is assumed to be the ratio between p_i and the sum of the p_i's in its group. You are not allowed to put the two teams with the highest p_i's (team 1 and team 2) in the same group. Neither are you allowed to arrange the groups so that the difference between the sum of points of the teams in one group compared to the sum of points of the teams in the other group exceeds 15% of the total number of points. All games are played at the home ground of one of the two participating teams; which one is not important since $d_{ij} = d_{ji}$.

Your task is to model this problem as a nonlinear (integer) program. *All functions defined have to be differentiable and explicit!*

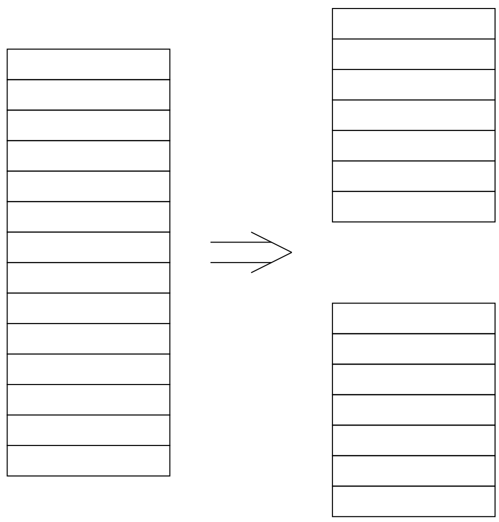


Figure 1: The 14 teams shall be put into two groups of 7 each.

[*Note: Optimization problems of this type have been used, e.g., for the planning of college baseball series in the US.*]

Question 5

(interior penalty methods)

Consider the problem to

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

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

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

where $\hat{\chi}_S(\boldsymbol{x}) = \phi(g(\boldsymbol{x}))$, 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 3 (convergence of an interior point algorithm) *Let the objective function $f : \mathbb{R}^n \to \mathbb{R}$ and the functions g_i, $i = 1, \ldots, m$, defining the inequality constraints be in $C^1(\mathbb{R}^n)$. Further assume that the barrier function $\phi : \mathbb{R} \to \mathbb{R}_+$ is in C^1 and that $\phi'(s) \geq 0$ for all $s < 0$.*

Consider a sequence $\{\boldsymbol{x}_k\}$ 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 $\lim_{k \to +\infty} \boldsymbol{x}_k = \hat{\boldsymbol{x}}$, and that LICQ holds at $\hat{\boldsymbol{x}}$. Then, $\hat{\boldsymbol{x}}$ is a KKT point of the problem at hand.

In other words,

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

(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 $(0,1)^{\mathrm{T}}$. The algorithm terminated after six iterations with the following results: $\boldsymbol{x}_6 \approx (0.94389, 0.89635)^{\mathrm{T}}$, and the multiplier estimate (given by $\nu_6\phi'(g(\boldsymbol{x}_6))$) $\hat{\mu}_6 \approx 3.385$. Confirm that the vector $\boldsymbol{x}_6$ is close to being a KKT point. Is it also near-globally optimal? Why/Why not?

Question 6

(linear programming)

Consider the linear program

$$\begin{aligned} z(b) := \text{maximum} \quad & 2x_1 + 3x_2 + x_3, \\ \text{subject to} \quad & x_1 - x_2 + 2x_3 \leq 1, \\ & 4x_1 + 2x_2 - x_3 \leq b, \\ & x_1, x_2, x_3 \geq 0, \end{aligned}$$

(1p) a) For $b = 2$ its optimal dual solution is claimed to be $\boldsymbol{y} = (5/3, 7/3)^{\mathrm{T}}$. Examine in a suitable way whether this is correct. (Here, it is *not* suitable to first solve the linear program or its corresponding LP dual problem!)

(1p) b) Use linear programming duality to determine the value of $z(b)$ for each $b \geq 0$ and give a principal graphical description of the function $z(b)$. Which are its most important mathematical properties?

(1p) c) Find, for each $b \geq 0$ the marginal value of an *increase* of the right-hand side of the second constraint, that is, find for each $b \geq 0$ the value of the right derivative of the function $z(b)$. Which marginal value is acheived for $b = 2$?

Question 7

(Lagrangian duality)

Consider the following linear programming problem:

$$\begin{aligned} \text{minimize} \quad & z = \quad x_2, & (1)\\ \text{subject to} \quad & x_1 \leq \frac{3}{2}, & (2)\\ & 2x_1 + 3x_2 \geq 6, & (3)\\ & x_1, \quad x_2 \geq 0. & (4) \end{aligned}$$

We will attack this problem by using Lagrangian duality.

(1p) a) Consider Lagrangian relaxing the complicating constraint (3). Write down explicitly the resulting Lagrangian subproblem of minimizing the Lagrange function over the remaining constraints. By varying the multiplier, construct an explicit formula for the Lagrangian dual function. Plot the dual function against the (only) dual variable, and state explicitly the Lagrangian dual problem.

(1p) b) Pick three primal feasible vectors and evaluate their respective objective values. Pick also three dual feasible values and evaluate their respective objective values. Using these six numbers, provide an interval wherein the optimal value of both the primal and dual problem must lie, and thereby also illustrate the Weak Duality Theorem.

(1p) c) Solve the Lagrangian dual problem from a). By using the primal–dual optimality conditions from Chapter 6, generate the (unique) optimal primal solution to the problem given above. Verify the Strong Duality Theorem.

Good luck!

Chalmers/GU
Mathematics

EXAM SOLUTION

TMA947/MAN280
APPLIED OPTIMIZATION

Date: 06–08–31
Examiner: Michael Patriksson

Question 1

(the Simplex method)

(2p) a) After changing sign of the second inequality and adding two slack variables s_1 and s_2, a BFS cannot be found directly. We create the phase I problem through an added artificial variable a_1 in the second linear constraint; the value of a_1 is to be minimized.

We use the BFS based on the variable pair (s_1, a_1) as the starting BFS for the phase I problem. In the first iteration of the Simplex method x_1 is the only variable with a negative reduced cost; hence x_1 is picked as the incoming variable. The minimum ratio test shows that s_1 should leave the basis. In the next iteration the reduced cost for varialbe x_3 is negative, and x_3 is picked as the incoming variable. The minimum ratio test shows that a_1 should leave the basis. We have found an optimal basis, $x_B = (x_1, x_3)^T$, to the phase I problem. We proceed to phase II, since the basis is feasible in the original problem.

Starting phase II with this BFS, we see that all reduced costs are positive, $\tilde{c}_N = (\alpha + 4, 2, 3)^T > 0$, and thus the BFS is optimal. $x_B = B^{-1}b = (2, 1)^T$ so $x^* = (2, 0, 3)^T$ and $z^* = c_B^T x_B = 3$.

(1p) b) For the dual problem to be unbounded, weak duality shows that the primal problem must be infeasible. Since α is in the cost vector of the primal problem, the feasibilty is not affected by α. Hence, no values of α lead to an unbounded feasible problem.

(3p)

Question 2

(necessary local and sufficient global optimality conditions)

See Proposition 4.23 and Theorem 4.24.

Question 3

(Newton's method revisited)

(2p) a) See the text book on quasi-Newton methods.

(1p) b) In order to be certain that the search direction given by the Newton subproblem is (a) defined at all and (b) is a direction of descent, the Hessian matrix must be positive definite. There are several ways in which to modify a matrix that is not positive definite such that the resulting matrix has this property.

The classic one is the Levenberg–Marquardt modification, in which one adds a diagonal matrix to the Hessian matrix such that their sum is positive definite. A second possibility is to replace Newton's method altogether with a quasi-Newton method, as explained in a). Special modifications also include the use of directions of negative curvature, in case the Hessian matrix is indefinite. See the text book for more details.

Question 4

(modelling)

Introduce the binary variables

$$x_i = \begin{cases} 1 & \text{if team } i \text{ is placed in group 1} \\ 0 & \text{otherwise} \end{cases}, \qquad i = 1, \dots, 14$$

The objective function can then be written as

$$\min \sum_{i=1}^{13} \sum_{j=i+1}^{14} d_{ij} \left(x_i x_j + (1 - x_i)(1 - x_j) \right) + \sum_{i=1}^{14} \sum_{j=1}^{14} d_{ij} \frac{p_i x_i}{\sum_{k=1}^{14} p_k x_k} \cdot \frac{p_j (1 - x_j)}{\sum_{k=1}^{14} p_k (1 - x_k)}$$

where the first term represents the travelling within the groups and the second term the expected travelling in the final. The constraints are

$$\begin{aligned}
\sum_{i=1}^{14} x_i &= 7, && (1) \\
x_1 + x_2 &= 1 && (2) \\
\sum_{i=1}^{14} x_i p_i &\le \sum_{i=1}^{14} (1 - x_i) p_i + 0.15 \sum_{i=1}^{14} p_i, && (3) \\
\sum_{i=1}^{14} (1 - x_i) p_i &\le \sum_{i=1}^{14} x_i p_i + 0.15 \sum_{i=1}^{14} p_i, && (4) \\
x_i &\in \{0, 1\} \quad i = 1, \dots, 14. && (5)
\end{aligned}$$

Constraint (1) makes sure that there are 7 teams in each group, constraint (2) that the two best teams are not in the same group and the constraints (3) and

(4) that the groups are arranged so that the difference between the sum of points in the two groups are not bigger than 15% of the total points.

Another possibility (maybe better) is to introduce more binary variables, u_{ij} and v_{ij}, where

$$u_{ij} = \begin{cases} 1 & \text{if team } i \text{ and team } i \text{ are both placed in group 1} \\ 0 & \text{otherwise} \end{cases}, \quad i, j = 1, \dots, 14,$$

$$v_{ij} = \begin{cases} 1 & \text{if team } i \text{ and team } i \text{ are both placed in group 2} \\ 0 & \text{otherwise} \end{cases}, \quad i, j = 1, \dots, 14.$$

We can then add to the previous model the linear forcing constraints

$$\begin{aligned} x_i + x_j &\leq u_{ij} + 1 \quad i, j = 1, \dots, 14, & (6) \\ x_i + x_j &\geq 1 - v_{ij} \quad i, j = 1, \dots, 14, & (7) \end{aligned}$$

the binary constraints

$$\begin{aligned} u_{ij} &\in \{0, 1\} \quad i, j = 1, \dots, 14, & (8) \\ v_{ij} &\in \{0, 1\} \quad i, j = 1, \dots, 14, & (9) \end{aligned}$$

and replace the first term in the previous objective function with the simpler linear term

$$\sum_{i=1}^{13} \sum_{j=i+1}^{14} d_{ij} \left(u_{ij} + v_{ij} \right).$$

Question 5

(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: $(\boldsymbol{x}_6)_1^2 - (\boldsymbol{x}_6)_2 \approx -0.005422 \approx 0$. We set up the KKT conditions to see whether it is fulfilled approximately. Indeed, we have the following corresponding to the system $\nabla f(\boldsymbol{x}_6) + \hat{\mu}_6 \nabla g(\boldsymbol{x}_6) = \mathbf{0}^2$:

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

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 $\boldsymbol{x}_6$ therefore is an approximate global optimal solution to our problem.

Question 6

(linear programming)

(1p) a) By complementarity slackness (Theorem 10.12),

$$\boldsymbol{x}^{\mathrm{T}}(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{y}-\boldsymbol{c}) = 0 \Leftrightarrow \left\{ \begin{array}{l} x_1(y_1+4y_2-2) = x_1 \cdot 0 = 0, \\ x_2(-y_1+2y_2+2) = x_2 \cdot 2 = 0 \Rightarrow x_2 = 0, \\ x_3(2y_1-y_2-1) = x_3 \cdot 0 = 0. \end{array} \right\}$$

Further, it follows that

$$\boldsymbol{y}^{\mathrm{T}}(\boldsymbol{A}\boldsymbol{x}-\boldsymbol{b}) = 0 \Leftrightarrow \left\{ \begin{array}{l} x_1+2x_3 = 1, \\ 4x_1-x_3 = 2, \\ x_2 = 0 \end{array} \right\} \Leftrightarrow \{x_1 = 5/9,\ x_2 = 0,\ x_3 = 2/9\}.$$

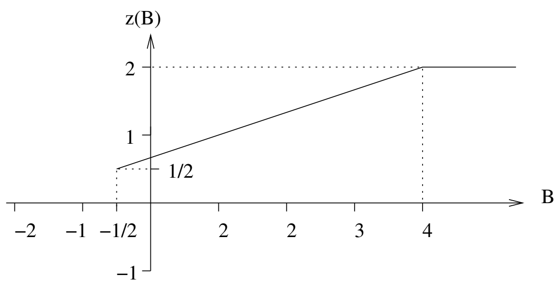
Since $\boldsymbol{x} \geq \boldsymbol{0}$ and $\boldsymbol{y} \geq \boldsymbol{0}$ it follows that $\boldsymbol{y} = (2/3, 1/3)$ is an optimal solution to the LP dual problem.

(2p) b) For $\beta = 2$ the optimal basis is $\boldsymbol{x}_B = (x_1, x_3)^{\mathrm{T}}$. This holds for those values of β such that $\boldsymbol{x}_B$ is feasible and optimal. Here, $\boldsymbol{B} = \begin{pmatrix} 1 & 2 \\ 4 & -1 \end{pmatrix}$, so that $\boldsymbol{x}_B = \boldsymbol{B}^{-1}\boldsymbol{b} = 1/9 \cdot \begin{pmatrix} 1 & 2 \\ 4 & -1 \end{pmatrix} \begin{pmatrix} 1 \\ \beta \end{pmatrix} = 1/9 \cdot \begin{pmatrix} 1+2\beta \\ 4-\beta \end{pmatrix} \geq 0 \Leftrightarrow$ $-1/2 \leq \beta \leq 4$. Optimality follows from $\boldsymbol{c}_N^{\mathrm{T}} - \boldsymbol{c}_B^{\mathrm{T}}\boldsymbol{B}^{-1}\boldsymbol{N} = (-2, 0, 0) -$

$(2,1) \cdot 1/9 \cdot \begin{pmatrix} 1 & 2 \\ 4 & -1 \end{pmatrix} \begin{pmatrix} -1 & 1 & 0 \\ 2 & 0 & 1 \end{pmatrix} = (-2, -2/3, -1/3) \leq 0$. Within the interval $-1/2 \leq \beta \leq 4$, $z(\beta) = (2+\beta)/3$.

For $\beta > 4$, x_3 becomes negative $\Rightarrow$ x_3 is not in the optimal basis for $\beta > 4$. Entering variable (according to the criterion in the dual simplex method) is x_2. The next basis is $\boldsymbol{x}_B = (x_1, x_2) = \boldsymbol{B}^{-1}\boldsymbol{b} = 1/6 \cdot \begin{pmatrix} 2 & 1 \\ -4 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ \beta \end{pmatrix} = 1/6 \cdot \begin{pmatrix} 2+\beta \\ \beta - 4 \end{pmatrix}$, which is optimal, since $\boldsymbol{c}_N^{\mathrm{T}} - \boldsymbol{c}_B^{\mathrm{T}} \boldsymbol{B}^{-1} \boldsymbol{N} = (1,0,0) - (2,-2) \cdot 1/6 \cdot \begin{pmatrix} 2 & 1 \\ -4 & 1 \end{pmatrix} \begin{pmatrix} 2 & 1 & 0 \\ -1 & 0 & 1 \end{pmatrix} = (-3, -2, 0) \leq 0$. Feasibility holds for $\boldsymbol{x}_B \geq 0 \Leftrightarrow \beta \geq 4$. Hence, for $\beta \geq 4$, $z(\beta) = 2$.

The function $z(\beta)$ is piecewise linear and concave on the halfline $\beta \geq -1/2$.



Question 7

(Lagrangian duality)

(1p) a) The Lagrangian subproblem is to, for any $\mu \geq 0$,

$$\begin{aligned} \text{minimize } & x_2 - \mu(2x_1 + 3x_2 - 6), \\ \text{subject to } & x_1 \in [0, 3/2], \\ & x_2 \geq 0. \end{aligned}$$

This problem has the following solution sets for varying values of μ: for $\mu \in [0, 1/3)$, $\boldsymbol{x}(\mu) = (3/2, 0)^{\mathrm{T}}$ uniquely; for $\mu = 1/3$, $x_1(\mu) = 3/2$ while $x_2(\mu) \geq 0$ arbitrarily; finally, for $\mu > 1/3$, there exists no optimal solution to the Lagrangian subproblem.

Inserting these solutions into the Lagrangian subproblem we obtain that the Lagrangian dual function has the following appearance: for $\mu \in [0, 1/3]$, $q(\mu) = 6\mu - 3\mu = 3\mu$, while for $\mu > 1/3$, $q(\mu) = -\infty$.

We can therefore state an explicit linear dual problem as follows:

$$\begin{aligned} &\text{maximize } 3\mu, \\ &\text{subject to } 0 \le \mu \le 1/3. \end{aligned}$$

(1p) b) $\boldsymbol{x} = (1, 2)^{\mathrm{T}} \implies z = 2$; $\boldsymbol{x} = (1, 4/3)^{\mathrm{T}} \implies z = 4/3$; $\boldsymbol{x} = (3/2, 1)^{\mathrm{T}} \implies z = 1$.

$\mu = 0 \implies q(\mu) = 0$; $\mu = 1/6 \implies q(\mu) = 1/2$; $\mu = 1/3 \implies q(\mu) = 1$.

(1p) c) $\mu^* = 1/3$. From a) the optimality conditions for the Lagrangian subproblem yields that $x_1(\mu^*) = x_1^* = 3/2$, while $x_2(\mu^*) \ge 0$. Since $\mu^* \neq 0$, we must satisfy the Lagrangian relaxed constraint with equality; this yields the condition that $3 + 3x_2 = 6$, hence $x_2 = 1$. We verify that $z^* = q^* = 1$.