Norwegian University of Science and Technology
Department of Engineering Cybernetics

**Final Exam**
Solution

**Problem 1 (42 %)** LP and Simplex

**a)** The function $f$ is a convex function if its domain $S$ is a convex set and if for any two points $x$ and $y$ the following property is satisfied:

$$f(\alpha x + (1-\alpha)y) \leq \alpha f(x) + (1-\alpha)f(y), \qquad \text{for all } \alpha \in [0,\ 1] \tag{1}$$

$f$ is said to be strictly convex if the above inequality is strict whenever $x \neq y$ and $\alpha$ is in the open interval $(0, 1)$. A function $f$ is said to be concave if $-f$ is convex.

With respect to the optimization problem (4) in the Appendix, we say the problem is convex if

- the objective function is convex;
- the equality constraint functions $c_i(\cdot),\ i \in \mathcal{E}$, are linear, and
- the inequality constraint functions $c_i(\cdot),\ i \in \mathcal{I}$, are concave.

Since LP problems have a linear objective function and linear constraints, LP problems are convex optimization problems. However, since linear functions are not strictly convex, LP problems are not strictly convex problems.

**b)** No, neither the gradient of the Lagrangean nor the gradient of the objective function is used to move to the next point in the Simplex algorithm. The Simplex method jumps from corner point to corner point with no explicit use of gradients.

**c)** We consider the LP problem

$$\min_{x\in\mathbb{R}^3} \quad 3x_1 + 2x_2 + x_3 \tag{2a}$$
$$\text{s.t.} \quad 2x_1 + x_2 + x_3 = 8 \tag{2b}$$
$$x_1 - x_2 - x_3 = 1 \tag{2c}$$
$$x \geq 0 \tag{2d}$$

where

$$c = \begin{bmatrix} 3 \\ 2 \\ 1 \end{bmatrix}, \qquad A = \begin{bmatrix} 2 & 1 & 1 \\ 1 & -1 & -1 \end{bmatrix}, \qquad b = \begin{bmatrix} 8 \\ 1 \end{bmatrix} \tag{3}$$

and define the Lagrangean

$$\mathcal{L}(x, \lambda, s) = c^\top x - \lambda^\top (Ax - b) - s^\top x \tag{4}$$

The derivative of $\mathcal{L}$ with respect to $x$ is

$$\nabla_x \mathcal{L}(x, \lambda, s) = c - A^\top \lambda - s \tag{5}$$

Hence, the KKT conditions (in the same order as (5) in the Appendix) are

$$\begin{aligned}
A^\top \lambda^* + s^* &= c && \text{(6a)}\\
A x^* &= b && \text{(6b)}\\
x^* &\geq 0 && \text{(6c)}\\
s^* &\geq 0 && \text{(6d)}\\
s_i^* x_i^* &= 0, \quad i = 1, 2, \ldots, n && \text{(6e)}
\end{aligned}$$

where $n = 3$, and $A$, $b$, and $c$ given in (3).

**d)** By adding the second equality constraint (2c) to the first (2b), we get $3x_1 = 9$ which means that $x_1 = 3$ for a feasible point. Since there are 3 variables and 2 equality constraints, the basis will contain 2 variables. Hence, for any basic feasible point, one variable will be zero. We know that $x_1$ must be nonzero at a feasible point, so to find the two basic feasible points we require $x_3 = 0$ or $x_2 = 0$. We start with $x_3 = 0$ and get from (2c) that

$$3 - x_2 - 0 = 1 \Rightarrow x_2 = 2 \tag{7}$$

That is, the first basic feasible point $(i)$ is $(3, 2, 0)$. To find the other point we set $x_2 = 0$ and again use (2c) to get

$$3 - 0 - x_3 = 1 \Rightarrow x_3 = 2 \tag{8}$$

Hence, the second basic feasible point $(ii)$ is $(3, 0, 2)$.

We now check the KKT conditions at these two basic feasible points. For the first point $(3, 2, 0)$, the objective function value is $3 \times 3 + 2 \times 2 + 1 \times 0 = 13$. Since both $x_1$ and $x_2$ are nonzero, we know from (6e) that $s_1 = s_2 = 0$. The first KKT condition (6a) can then be written

$$\begin{bmatrix} 2 & 1 \\ 1 & -1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} \lambda_1 \\ \lambda_2 \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ s_3 \end{bmatrix} = \begin{bmatrix} 3 \\ 2 \\ 1 \end{bmatrix} \tag{9}$$

Solving the first two equations of this set gives $\lambda_1 = 5/3$ and $\lambda_2 = -1/3$. The third equation then gives

$$\lambda_1 - \lambda_2 + s_3 = 1 \Rightarrow \frac{5}{3} + \frac{1}{3} + s_3 = 1 \Rightarrow s_3 = -1 \tag{10}$$

Since $s_3 < 0$ violates (6d), the KKT conditions are not satisfied and the basic feasible point $(3, 2, 0)$ can not be optimal.

The second basic feasible point $(3,0,2)$ has objective function value $3\times3+2\times0+1\times2=11$. Here, $x_1$ and $x_3$ are nonzero, so $s_1=s_3=0$ by (6e). The first KKT condition (6a) can then be written

$$\begin{bmatrix}2 & 1\\ 1 & -1\\ 1 & -1\end{bmatrix}\begin{bmatrix}\lambda_1\\ \lambda_2\end{bmatrix}+\begin{bmatrix}0\\ s_2\\ 0\end{bmatrix}=\begin{bmatrix}3\\ 2\\ 1\end{bmatrix} \tag{11}$$

Solving the first and third equations of this set gives $\lambda_1=4/3$ and $\lambda_2=1/3$. The second equation then gives

$$\lambda_1-\lambda_2+s_2=2\Rightarrow\frac{4}{3}-\frac{1}{3}+s_2=2\Rightarrow s_2=1 \tag{12}$$

This means that (6a) is satisfied. Since $s=[0,\ 1,\ 0]^\top$, (6d) is satisfied; the point $(3,0,2)$ satisfies (6c) and (6e) with $s$; finally, the way we found the basic feasible point ensures that (6b) is satisfied. That is, all KKT conditions are satisfied and we can conclude that the solution is $x^*=[3,\ 0,\ 2]^\top$. The fact that this basic feasible point has the lowest objective function value ($11<13$) confirms that this is the solution.

**e)** The LP problem

$$\begin{aligned}\min_{x\in\mathbb{R}^3}\quad & 3x_1+x_2+x_3 && \text{(13a)}\\ \text{s.t.}\quad & 2x_1+x_2+x_3\le 2 && \text{(13b)}\\ & x_1-x_2-x_3\le -1 && \text{(13c)}\\ & x\ge 0 && \text{(13d)}\end{aligned}$$

is not on standard form. It can be transformed into standard form by adding the (positive) slack variables $x_4$ and $x_5$ to (13b) and (13c), respectively:

$$\begin{aligned}2x_1+x_2+x_3+x_4\qquad &=2 && \text{(14a)}\\ x_1-x_2-x_3\qquad +x_5&=-1 && \text{(14b)}\end{aligned}$$

We then have the standard form LP

$$\begin{aligned}\min_{x\in\mathbb{R}^5}\quad & 3x_1+x_2+x_3 && \text{(15a)}\\ \text{s.t.}\quad & 2x_1+x_2+x_3+x_4=2 && \text{(15b)}\\ & x_1-x_2-x_3+x_5=-1 && \text{(15c)}\\ & x\ge 0 && \text{(15d)}\end{aligned}$$

or

$$\begin{aligned}\min_{x\in\mathbb{R}^5}\quad & c^\top x && \text{(16a)}\\ \text{s.t.}\quad & Ax=b && \text{(16b)}\\ & x\ge 0 && \text{(16c)}\end{aligned}$$

with

$$c = \begin{bmatrix} 3 & 1 & 1 & 0 & 0 \end{bmatrix}^\top, \qquad x = \begin{bmatrix} x_1 & x_2 & x_3 & x_4 & x_5 \end{bmatrix}^\top$$
$$A = \begin{bmatrix} 2 & 1 & 1 & 1 & 0 \\ 1 & -1 & -1 & 0 & 1 \end{bmatrix}, \qquad b = \begin{bmatrix} 2 \\ -1 \end{bmatrix} \tag{17}$$

**f)** The dual problem

$$\max_{\lambda \in \mathbb{R}^m} \quad b^\top \lambda \tag{18a}$$
$$\text{s.t.} \quad A^\top \lambda \leq c \tag{18b}$$

can be written as the minimization problem

$$\min_{\lambda \in \mathbb{R}^m} \quad -b^\top \lambda \tag{19a}$$
$$\text{s.t.} \quad c - A^\top \lambda \geq 0 \tag{19b}$$

If we let $x \in \mathbb{R}^n$ be the Lagrange multipliers for $c - A^\top \lambda \geq 0$, the Lagrangean for the dual problem is

$$\bar{\mathcal{L}}(\lambda, x) = -b^\top \lambda - x^\top (c - A^\top \lambda) \tag{20}$$

Differentiating with respect to $\lambda$ and requiring the derivative to be zero gives

$$\nabla_\lambda \bar{\mathcal{L}} = -b + (x^\top A^\top)^\top = Ax - b = 0 \tag{21}$$

From (5) in the Appendix, we then have that the KKT conditions for the dual problem are

$$Ax^* = b \tag{22a}$$
$$A^\top \lambda^* \leq c \tag{22b}$$
$$x^* \geq 0 \tag{22c}$$
$$x_i^*(c - A^\top \lambda^*)_i = 0, \quad i = 1, \dots, n \tag{22d}$$

To show the equivalence to the KKT conditions for the primal problem, consider the dual problem (18) on the form

$$\max_{\lambda \in \mathbb{R}^m} \quad b^\top \lambda \tag{23a}$$
$$\text{s.t.} \quad A^\top \lambda + s = c \tag{23b}$$
$$s \geq 0 \tag{23c}$$

where $s$ has been introduced as a vector of slack variables. We then see that it is natural to define $s = c - A^\top \lambda \geq 0$. Hence, we add the condition $A^\top \lambda^* + s^* = c$, and (22b) can be written $s^* \geq 0$. The KKT conditions for the dual problem are then

$$A^\top \lambda^* + s^* = c \tag{24a}$$
$$Ax^* = b \tag{24b}$$
$$x^* \geq 0 \tag{24c}$$
$$s^* \geq 0 \tag{24d}$$
$$s_i^* x_i^* = 0, \quad i = 1, 2, \dots, n \tag{24e}$$

which are equivalent to the KKT conditions (6) for the primal problem.

**g)** If either the primal or the dual problem has a finite solution, then so does the other, and objective function values are equal. Furthermore, if either the primal or the dual problem, then the other problem is infeasible. (See Theorem 13.1 in Nocedal & Wright).

To show that the primal and dual LP problems have the same optimal objective function values, we will use (24a), (24b), and (24e). Note that (24e) implies that $(s^*)^\top x^* = 0$. We start the derivation from the optimal objective function value of the primal LP:

$$
\begin{aligned}
c^\top x^* &= (A^\top \lambda^* + s^*)^\top x^* && \text{(from equation (24a))} \\
&= (\lambda^*)^\top A x^* + (s^*)^\top x^* && \text{(expanding the parenthesis)} \\
&= (\lambda^*)^\top b + (s^*)^\top x^* && \text{(from equation (24b))} \\
&= (\lambda^*)^\top b + 0 && \text{(from equation (24e))} \\
&= b^\top \lambda^* && \text{(by equivalent forms of the scalar product)}
\end{aligned}
$$

**Problem 2 (30 %)** MPC and optimal control

**a)** When choosing $L$ in the dual-mode MPC, the rationale is that after some time (ideally a shorter time than $L$), the constraints are "resolved" and the LQR controller $u_i = -Kx_i$ is optimal for the rest of the horizon. This assumption must hold.

Furthermore, the control horizon $L$ is chosen to "approximate a desirable predicted curve shape of the controls $u_i$". Moreover, one can argue that $L$ should be chosen to be as large as computational limits permits. The fundamental assumption which must be satisfied when choosing $L$ is then that the time it takes to solve the optimization problem at each sampling instant must be shorter than the sampling interval.

**b)** The optimization problem we must solve to find $c_i$ is a convex quadratic-programming problem with linear inequality constraints. This problem may for instance be solved by an active-set method. Specifically, the optimization problem is

$$
\begin{aligned}
\min_{c} \quad & c^\top W_c c && \text{(25a)} \\
\text{s.t.} \quad & M x_k + N c \leq b && \text{(25b)}
\end{aligned}
$$

where $W_c \geq 0$. Hence, the QP is convex. See the note on MPC by Lars Imslad and the solution to Assignment 9 for more on the matrices in this QP.

**c)** If the linear model is replaced by a nonlinear one, the constraints in the optimization problem become nonlinear. The optimization problem in then a nonlinear program (NLP). Problems of this class can be solved with sequential-quadratic-programming (SQP) methods or interior-point (IP) methods.

**d)** The prediction horizon should be longer than the slowest dynamics in the system (largest time constant) and should also be longer than the control horizon $L$. All predictions of controlled variables should be able to reach a steady state on setpoints or on constraints from movements in the final input block. Figure 1 shows a system with two states, one settles much quicker (dashed) than the other (solid). The slower state determines the prediction horizon; in this case, the prediction horizon (a typical choice is suggested in the figure) is chosen to be long enough so that the slowest state is almost settled at the end of the horizon.

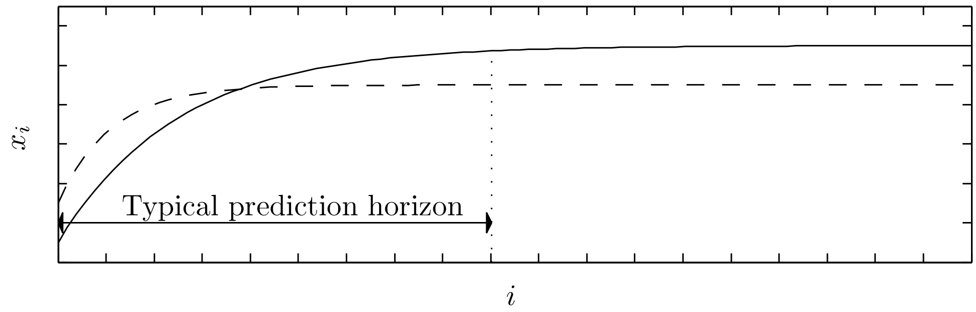


Figure 1: Illustration to Problem 2 d - two states with different settling times. The slower state determines the prediction horizon.

**e)** Control input blocking is to divide the control horizon into blocks or intervals where the controls take on constant values (usually). This reduces the number of optimization variables and hence simplifies the optimization problem and the computational load in the controller. In fact, the MPC computation time is mainly related to the number of control blocks. The intervals are usually short in the beginning of the control horizon and then longer towards the end. Figure 2 shows a possible scheme for input blocking.

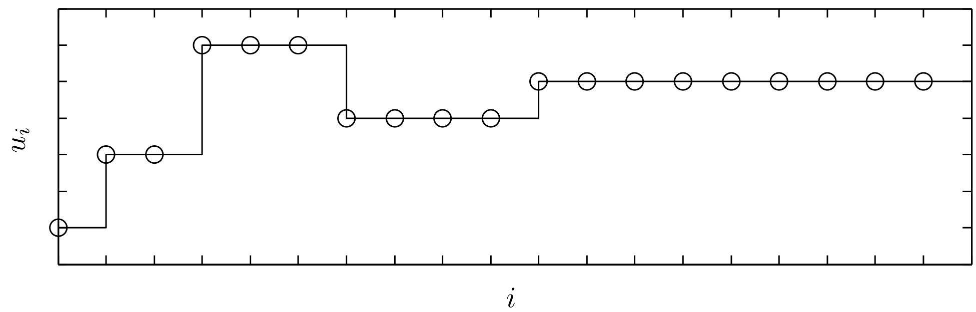


Figure 2: Illustration to Problem 2 e - a possible input-block scheme.

**f)** Denote the acceptable variation of $y_i$ as $\Delta y_i$; that is, $\Delta y_1 = 10$ and $\Delta y_2 = 0.1$. Since we use a quadratic weighting, a reasonable ratio would be

$$\frac{q_1}{q_2} = \left(\frac{\Delta y_2}{\Delta y_1}\right)^2 = \left(\frac{0.1}{10}\right)^2 = 10^{-4} \tag{26}$$

which means that

$$q_1(\Delta y_1)^2 = q_2(\Delta y_2)^2 \tag{27}$$

**Problem 3 (28 %)** Various topics

**a)** The points $x^1$, $x^2$, $x^3$ are ordered so that $f(x^1) \leq f(x^2) \leq f(x^3)$. That is, the points are ordered from best to worst.

Inside contraction is tried if the reflection point is worse than $x^{n+1}$ ($x^3$ for this problem). The inside-contraction point is defined by $g(1/2)$ and is for this problem given by $g(1/2) = (0.65, 0.65)^\top$ (it is hence inside the simplex). If the inside-contraction point is better than the current worst point $x^{n+1}$, the worst point $x^{n+1}$ is replaced by $g(1/2)$; this operation is called inside contraction. The algorithm then proceeds to the next iteration.

If the inside-contraction point is worse than the current worst point, inside contraction is not performed; shrinkage is then performed instead. Since $f(x^{\text{refl}}) = 35 > f(x^3) = 30$, inside contraction will be attempted.

The simplex for this problem is illustrated in Figure 3; the reflection point $x^{\text{refl}}$, the inside-contraction point $x_{1/2} = g(1/2)$, and the simplex after inside contraction are all included.

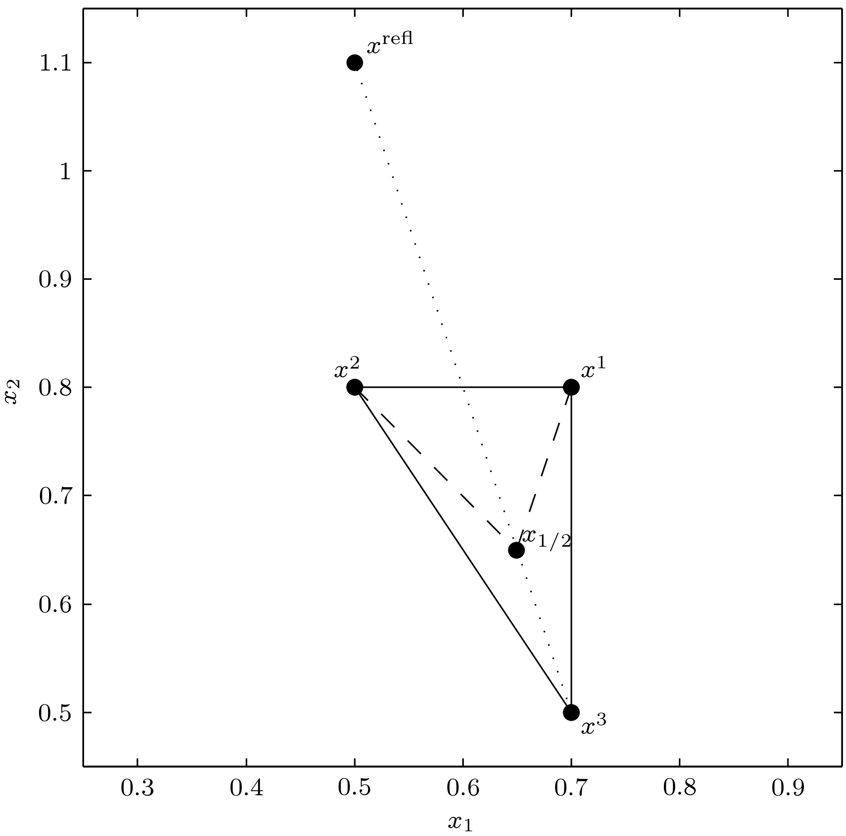


Figure 3: The simplex from Problem 3 a. The dotted line connects the worst point $x^3$ to the reflection point $x^{\text{refl}}$; the dashed line is the simplex after inside contraction where $x^3$ is replaced by the inside-contraction point $x_{1/2} = g(1/2)$.

**b)** We wish to find the step length $\alpha$ that minimizes $f(x_k + \alpha p_k)$. Note that the objective function is convex (since $Q = Q^\top \geq 0$) and that the search direction $p_k$

is a descent direction (since ). By differentiating

$$
\begin{aligned}
f(x_k + \alpha p_k) &= \frac{1}{2}(x_k + \alpha p_k)^\top Q(x_k + \alpha p_k) - b^\top (x_k + \alpha p_k) \\
&= \frac{1}{2}x_k^\top Q x_k + \frac{1}{2}x_k^\top Q \alpha p_k + \frac{1}{2}\alpha p_k^\top Q x_k + \frac{1}{2}\alpha^2 p_k^\top Q p_k - b^\top x_k - \alpha b^\top p_k \\
&= \frac{1}{2}x_k^\top Q x_k + \alpha x_k^\top Q p_k + \frac{1}{2}\alpha^2 p_k^\top Q p_k - b^\top x_k - \alpha b^\top p_k
\end{aligned} \tag{28}
$$

with respect to $\alpha$ and setting the derivative to zero, we get

$$
\frac{\partial f}{\partial \alpha}(x_k + \alpha p_k) = x_k^\top Q p_k + \alpha p_k^\top Q p_k - b^\top p_k = 0 \tag{29}
$$

Rearranging this equation gives

$$
\alpha = \frac{b^\top p_k - x_k^\top Q p_k}{p_k^\top Q p_k} \tag{30}
$$

which minimizes $f(x_k + \alpha p_k)$. See page 42 in Nocedal & Wright for a similar problem involving the steepest descent direction.

**c)** We have the optimization problem

$$
\begin{aligned}
\min_{x \in \mathbb{R}^2} \quad & f(x) = -2x_1 + x_2 && \text{(31a)} \\
\text{s.t.} \quad & c_1(x) = (1 - x_1)^3 - x_2 \geq 0 && \text{(31b)} \\
& c_2(x) = 0.25x_1^2 + x_2 - 1 \geq 0 && \text{(31c)}
\end{aligned}
$$

with solution $x^* = [0,\ 1]^\top$.

To find if the LICQ holds at the solution, we must find the gradients of the constraint that are active at the solution. Since both constraints are active at $x^*$ we find

$$
\nabla c_1(x) = \begin{bmatrix} -3(1 - x_1)^2 \\ -1 \end{bmatrix}, \qquad \nabla c_2(x) = \begin{bmatrix} 0.5x_1 \\ 1 \end{bmatrix} \tag{32}
$$

At $x^*$, these gradients have values

$$
\nabla c_1(x^*) = \begin{bmatrix} -3 \\ -1 \end{bmatrix}, \quad \nabla c_2(x^*) = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \tag{33}
$$

It is clear that these gradients are linearly independent, which means that the LICQ holds.

The Lagrangian for the problem is

$$
\mathcal{L}(x, \lambda) = -2x_1 + x_2 - \lambda_1 c_1(x) - \lambda_2 c_2(x) \tag{34}
$$

To check the first KKT condition, we differentiate with respect to $x$ to obtain and evaluate at $(x^*, \lambda^*)$ to get

$$
\nabla_x \mathcal{L}(x^*, \lambda^*) = \begin{bmatrix} -2 + 3\lambda_1^*(1 - x_1^*)^2 - 0.5\lambda_2^* x_1^* \\ 1 + \lambda_1 - \lambda_2 \end{bmatrix} = 0 \Rightarrow \lambda^* = \begin{bmatrix} 2/3 \\ 5/3 \end{bmatrix} \tag{35}
$$

There are no equality constraints, and both inequality constraints are satisfied. With the positive Lagrange multipliers all KKT conditions are satisfied.

The second-order conditions involve the Hessian of the Lagrangian at the solution. For this problem, we have

$$\nabla^2_{xx}\mathcal{L}(x^*, \lambda^*) = \begin{bmatrix} -6\lambda_1^*(1 - x_1^*) - 0.5\lambda_2^* & 0 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} -29/6 & 0 \\ 0 & 0 \end{bmatrix} \tag{36}$$

We also need to find the critical cone in order to check the second-order conditions. Since there are no equality constraints and both Lagrange multipliers are positive, the critical cone consists of the vectors $w$ which satisfy

$$\nabla c_i(x^*)^\top w = 0, \quad \text{for } i = 1, 2 \tag{37}$$

Starting with the constraint gradient $\nabla c_2(x^*)$, we have that

$$\nabla c_2(x^*)^\top w = \begin{bmatrix} 0 & 1 \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \end{bmatrix} = w_2 = 0 \tag{38}$$

With this information about the vector $w$, we find from

$$\nabla c_1(x^*)^\top w = \begin{bmatrix} -3 & -1 \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \end{bmatrix} = \begin{bmatrix} -3 & -1 \end{bmatrix} \begin{bmatrix} w_1 \\ 0 \end{bmatrix} = -3w_1 = 0 \tag{39}$$

that the critical cone contains only the null vector $w = [0,\, 0]^\top$. The second-order necessary conditions are that the LICQ and the KKT conditions are both satisfied and that

$$w^\top \nabla_{xx}\mathcal{L}(x^*, \lambda^*) w \geq 0, \quad \text{for all } w \text{ in the critical cone} \tag{40}$$

Since $w = [0,\, 0]^\top$ is the only vector in the critical cone, this condition is trivially satisfied. That is, the second-order necessary conditions are satisfied.

The second-order sufficient conditions are that the KKT conditions are satisfied and that

$$w^\top \nabla_{xx}\mathcal{L}(x^*, \lambda^*) w \geq 0, \quad \text{for all } w \text{ in the critical cone, } w \neq 0 \tag{41}$$

Since the Hessian of the Lagrangean is negative semidefinite at the solution and the critical cone only contains $w = 0$, the second-order sufficient conditions are not satisfied.