Norwegian University of Science and Technology
Department of Mathematical Sciences

Exercise set 6 solution

1 **a)** We begin by stating the problem in standard form, writing $\mathbf{x} = [x, y]^T$:

$$\min_{\mathbf{x}\in\mathbb{R}^2} f(\mathbf{x}) \quad \text{s.t.} \quad c_i(\mathbf{x}) \geq 0, \quad i = 1, 2, 3,$$

where

$$\begin{aligned} f(\mathbf{x}) &= x^2 + y^2, \\ c_1(\mathbf{x}) &= x + y - 1, \\ c_2(\mathbf{x}) &= 2 - y, \\ c_3(\mathbf{x}) &= y^2 - x. \end{aligned}$$

We then find the Lagrangian function and its gradient:

$$\begin{aligned} \mathcal{L}(\mathbf{x}, \lambda) &= x^2 + y^2 - \lambda_1(x + y - 1) - \lambda_2(2 - y) - \lambda_3(y^2 - x) \\ \nabla_x \mathcal{L}(\mathbf{x}, \lambda) &= \begin{bmatrix} 2x - \lambda_1 + \lambda_3 \\ 2y - \lambda_1 + \lambda_2 - 2y\lambda_3 \end{bmatrix}. \end{aligned}$$

The KKT conditions can now be stated in full as:

$$2x^* - \lambda_1^* + \lambda_3^* = 0 \tag{1a}$$
$$2y^* - \lambda_1^* + \lambda_2^* - 2y^*\lambda_3^* = 0 \tag{1b}$$
$$x^* + y^* - 1 \geq 0 \tag{1c}$$
$$2 - y^* \geq 0 \tag{1d}$$
$$y^{*2} - x^* \geq 0 \tag{1e}$$
$$\lambda_i^* \geq 0, \quad i = 1, 2, 3 \tag{1f}$$
$$\lambda_1^*(x^* + y^* - 1) = 0 \tag{1g}$$
$$\lambda_2^*(2 - y^*) = 0 \tag{1h}$$
$$\lambda_3^*(y^{*2} - x^*) = 0. \tag{1i}$$

b) The feasible set is sketched in figure 1. We will find all KKT points by systematically considering all possible active sets of constraints. Remember that a constraint c_i is active at a point $\mathbf{x}$ if $c_i(\mathbf{x}) = 0$, and that if all λ_i^* are negative at a point, then it is a candidate for a maximizer. Also, the LICQ conditions are satisfied at every point we consider here; with one active constraint, the LICQ conditions hold trivially, and in the cases with two constraints it is not hard to check that the LICQ conditions do hold.

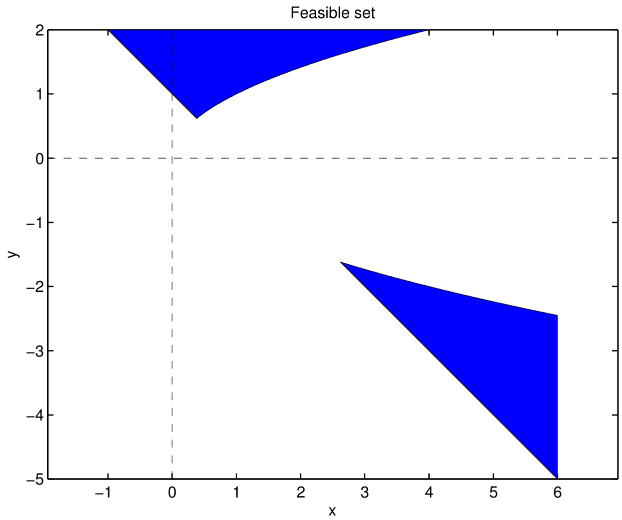

Figure 1: Feasible set. Note: The lower "triangle" extends further toward infinity.

Observe that if $\mathbf{x}^* = [x^*, y^*]^T$ is a KKT point, then from (1a) and (1b) we have:

$$x^* = \frac{\lambda_1^* - \lambda_3^*}{2}, \quad y^* = \frac{\lambda_1^* - \lambda_2^*}{2(1 - \lambda_3^*)}.$$

From here on, we will drop the asterisk in the notation and write x for x^*, etc.

First, if the active set is empty, i.e. neither of (1c)-(1e) are equalities. This corresponds to the interior of the domain. Then, by (1g)-(1i), we have $\lambda_1 = \lambda_2 = \lambda_3 = 0$, and so $x = y = 0$. But this point is not feasible, since it violates condition (1c). Thus, with the active set empty, there are no KKT points.

Next, we consider the case when the active set contains one index, i.e. exactly one of (1c)-(1e) is an equality. This corresponds to the boundaries of the domain, excepting the corner points. If (1c) is active, then $\lambda_2 = \lambda_3 = 0$ while $\lambda_1 \geq 0$. We get

$$x = \frac{\lambda_1}{2}, \quad y = \frac{\lambda_1}{2},$$

and inserting this into (1c) (which is now an equality), we get the condition

$$\frac{\lambda_1}{2} + \frac{\lambda_1}{2} - 1 = 0 \Rightarrow \lambda_1 = 1,$$

giving us the point $(x, y) = (\frac{1}{2}, \frac{1}{2})$. But this point violates condition (1e), so $(\frac{1}{2}, \frac{1}{2})$ is not a KKT point.

If (1d) is active, then $\lambda_1 = \lambda_3 = 0$ while $\lambda_2 \geq 0$, so

$$x = 0, \quad y = -\frac{\lambda_2}{2}.$$

Inserting this into the equality (1d), we get

$$2 + \frac{\lambda_2}{2} = 0 \Rightarrow \lambda_2 = -4.$$

Thus, (0,2) is a candidate for a maximizer. One can then check to verify that all KKT conditions are satisfied, and we find (0,2) to be a KKT point corresponding to a maximizer.

If (1e) is active, then $\lambda_1 = \lambda_2 = 0$ while $\lambda_3 \geq 0$, so

$$x = -\frac{\lambda_3}{2}, \quad y = 0.$$

Inserting this into the equality (1e), we get

$$\frac{\lambda_3}{2} = 0 \Rightarrow \lambda_3 = 0.$$

This gives the candidate point $(0, 0)$, which is not feasible since it violates (1c), and thereby is not a KKT point.

Having considered all possible active sets of one index, we now turn to the cases with two indices, i.e. exactly two of (1c)-(1e) are equalities. This corresponds to the corner points of the domain. First, if (1c) and (1d) are both active, then $\lambda_3 = 0$ while $\lambda_1, \lambda_2 \geq 0$. This gives us

$$x = \frac{\lambda_1}{2}, \quad y = \frac{\lambda_1 - \lambda_2}{2}.$$

Plugging this into equalities (1c) and (1d) yields:

$$\begin{aligned} \frac{\lambda_1}{2} + \frac{\lambda_1 - \lambda_2}{2} - 1 &= 0 \\ 2 - \frac{\lambda_1 - \lambda_2}{2} &= 0, \end{aligned}$$

with solutions $\lambda_1 = -2$ and $\lambda_2 = -6$, yielding the KKT point (-1,2). Note that this is a candidate for a local maximizer, since all multipliers are negative.

Next, if (1c) and (1e) are both active, then $\lambda_2 = 0$ while $\lambda_1, \lambda_3 \geq 0$, which means

$$x = \frac{\lambda_1 - \lambda_3}{2}, \quad y = \frac{\lambda_1}{2(1 - \lambda_3)}.$$

Plugging this into equalities (1c) and (1e) yields:

$$\begin{aligned} \frac{\lambda_1 - \lambda_3}{2} + \frac{\lambda_1}{2(1 - \lambda_3)} - 1 &= 0 \\ \frac{\lambda_1^2}{4(1 - \lambda_3)^2} - \frac{\lambda_1 - \lambda_3}{2} &= 0. \end{aligned}$$

Solving this set of equations yields $\lambda_1 = 5 \pm \frac{9}{\sqrt{5}}$ and $\lambda_3 = 2 \pm \frac{4}{\sqrt{5}}$, thereby giving the candidate points $(x, y) = (\frac{1}{2}(3 \pm \sqrt{5}), \frac{1}{2}(-1 \mp \sqrt{5}))$ which both satisfy the KKT conditions. Since $\lambda_1, \lambda_3 \geq 0$, these points are minimizer candidates. Note: This result can be arrived upon by the easier approach of first finding the points (x, y) where c_1 and c_3 are both active, then working out what λ_1 and λ_3 are.

Finally, we check the case where (1d) and (1e) are both active, i.e. $\lambda_1 = 0$ while $\lambda_2, \lambda_3 \geq 0$. This gives us

$$x = -\frac{\lambda_3}{2}, \quad y = -\frac{\lambda_2}{2(1-\lambda_3)}.$$

Plugging this into equalities (1d) and (1e) yields:

$$2 + \frac{\lambda_2}{2(1-\lambda_3)} = 0$$
$$\frac{\lambda_2^2}{4(1-\lambda_3)^2} + \frac{\lambda_3}{2} = 0,$$

which can be solved to find $\lambda_2 = -28$ and $\lambda_3 = -8$, giving the candidate point $(x, y) = (4, 2)$, which is a candidate for a maximizer, since the multipliers are negative.

Concerning the case with all constraints active, we may conclude that no KKT point exists; all three constraint functions cannot be active at the same point. The KKT points and their corresponding multipliers are summarized in the table below.

Point	λ_1	λ_2	λ_3	Minimizer/maximizer candidate
(0,2)	0	-4	0	Maximizer
$(\frac{1}{2}(3+\sqrt{5}), \frac{1}{2}(-1-\sqrt{5}))$	$5+\frac{9}{\sqrt{5}}$	0	$2+\frac{4}{\sqrt{5}}$	Minimizer
$(\frac{1}{2}(3-\sqrt{5}), \frac{1}{2}(-1+\sqrt{5}))$	$5-\frac{9}{\sqrt{5}}$	0	$2-\frac{4}{\sqrt{5}}$	Minimizer
(-1,2)	-2	-6	0	Maximizer
(4,2)	0	-28	-8	Maximizer

c) To determine whether the KKT points that are minimizer candidates are in fact local minimizers, we check the second order sufficient conditions from Theorem 12.6 in N&W, i.e. whether

$$w^T \nabla^2_{xx} \mathcal{L}(x, \lambda) w > 0 \; \forall w \in \mathcal{C}(x, \lambda), w \neq 0, \tag{2}$$

where, $\mathcal{C}(x, \lambda)$ is the critical cone at x, given by (12.53) in N&W.

For both candidates, i.e. $(\frac{1}{2}(3 \pm \sqrt{5}), \frac{1}{2}(-1 \mp \sqrt{5}))$, we have that the critical cone is simply given as $\mathcal{C}(x, \lambda) = \{0\}$. This is because any $w \in \mathcal{C}(x, \lambda)$ must be orthogonal to the $\nabla c_i(x)$ for which $\lambda_i > 0$, of which there are two for each point. Since the LICQ conditions hold at both points, these two vectors are linearily independent and thus span $\mathbb{R}^2$. The only vector orthogonal to $\mathbb{R}^2$ is the zero vector. Thereby, the only vector in $\mathcal{C}(x, \lambda)$ is the zero vector for these points, and thus condition (2) holds by default. We can conclude that $(\frac{1}{2}(3 \pm \sqrt{5}), \frac{1}{2}(-1 \mp \sqrt{5}))$ are strict local minimizers.

We note that $f(\frac{1}{2}(3-\sqrt{5}), \frac{1}{2}(-1+\sqrt{5})) < f(\frac{1}{2}(3+\sqrt{5}), \frac{1}{2}(-1-\sqrt{5}))$ and $f(\mathbf{x}) \to \infty$ in the unbounded region of the feasible domain. This means that $(\frac{1}{2}(3-\sqrt{5}), \frac{1}{2}(-1+\sqrt{5}))$ is a global minimizer and $(\frac{1}{2}(3+\sqrt{5}), \frac{1}{2}(-1-\sqrt{5}))$ is a local minimizer.

2 We begin by stating the problem in standard form, writing $\mathbf{x} = [x, y]^T$:

$$\min_{\mathbf{x}\in\mathbb{R}^2} f(\mathbf{x}) \quad \text{s.t.} \quad c_i(\mathbf{x}) \geq 0, \quad i = 1, 2$$

where

$$\begin{aligned} f(\mathbf{x}) &= x, \\ c_1(\mathbf{x}) &= y - x^4 \\ c_2(\mathbf{x}) &= x^3 - y \end{aligned}$$

The feasible set is sketched in figure 2.

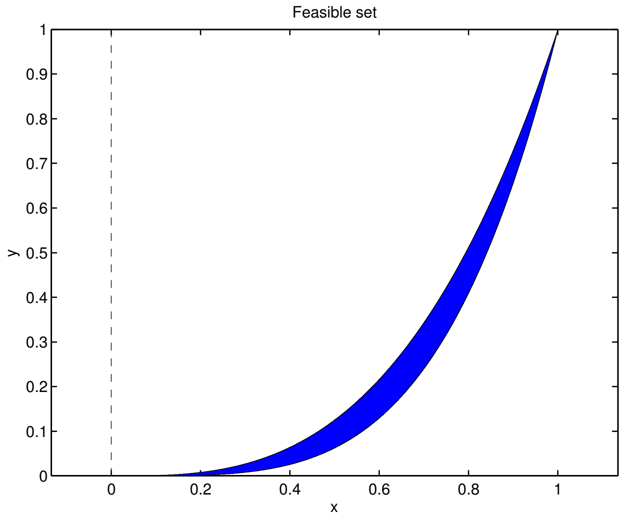

Figure 2: Feasible set.

We then find the Lagrange function and its gradient:

$$\begin{aligned} \mathcal{L}(\mathbf{x}, \lambda) &= x - \lambda_1(y - x^4) - \lambda_2(x^3 - y) \\ \nabla_x \mathcal{L}(\mathbf{x}, \lambda) &= \begin{bmatrix} 1 + 4x^3\lambda_1 - 3x^2\lambda_2 \\ -\lambda_1 + \lambda_2 \end{bmatrix}, \end{aligned}$$

and state the KKT conditions as:

$$1 + 4x^3\lambda_1 - 3x^2\lambda_2 = 0 \tag{3a}$$
$$-\lambda_1 + \lambda_2 = 0 \tag{3b}$$
$$y - x^4 \geq 0 \tag{3c}$$
$$x^3 - y \geq 0 \tag{3d}$$
$$\lambda_i \geq 0, \quad i = 1, 2 \tag{3e}$$
$$\lambda_1(y - x^4) = 0 \tag{3f}$$
$$\lambda_2(x^3 - y) = 0. \tag{3g}$$

Now, we can take a shortcut; from (3b), we see that $\lambda_1 = \lambda_2$, and from (3a) we see that there cannot exist any KKT point for which $\lambda_1 = \lambda_2 = 0$. Therefore, the cases with no active constraints ($\lambda_1 = \lambda_2 = 0$) and one active constraint ($\lambda_1 = 0$ or

$\lambda_2 = 0$) cannot produce KKT points. We are left with considering the case where both constraints are active, i.e. the corner points (0,0) and (1,1).

In the point (1,1), we find (by (3a) and (3b)) that $\lambda_1 = \lambda_2 = -1$, corresponding to a candidate for a maximizer. In fact, the LICQ holds here, with two linearly independent ∇c_i, and know from the discussion in the previous exercise (two linearly independent vectors span $\mathbb{R}^2$ so the critical cone contains only the zero vector) that this is a local maximizer.

The last point is (0,0), for which we cannot write the gradient of f at (0,0) (which is $[1, 0]^T$) as a non-negative linear combination of the gradients of the constraints, and which therefore is not a KKT point. This does not, however, mean that it is not a minimizer. Applying common sense, it is clearly a local minimum, as no other points with $x = 0$ are feasible, and $x = 0$ is the lowest possible value of the objective function.

3 **a)** We begin, as usual, by stating the problem in standard form, writing $\mathbf{x} = [x, y]^T$:

$$\min_{\mathbf{x}\in\mathbb{R}^2} f(\mathbf{x}) \quad \text{s.t.} \quad c_i(\mathbf{x}) \geq 0, \quad i = 1, 2$$

where

$$\begin{aligned} f(\mathbf{x}) &= xy, \\ c_1(\mathbf{x}) &= y - x \\ c_2(\mathbf{x}) &= x^3 - y^4 \end{aligned}$$

The feasible set is sketched in figure 3.

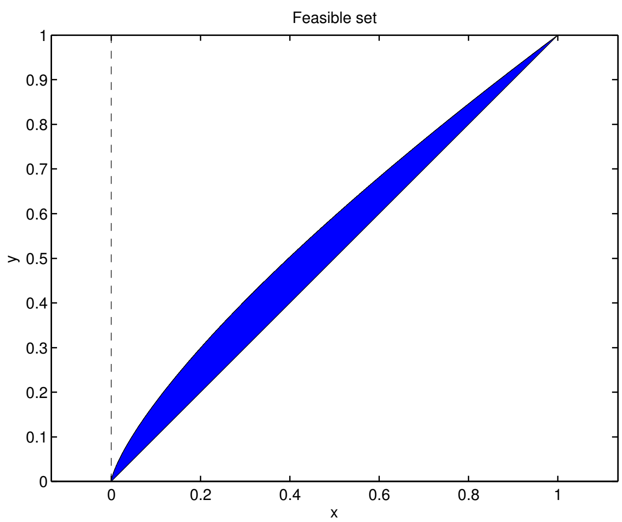

Figure 3: Feasible set, exercise 3.

We find the Lagrange function and its gradient:

$$\begin{aligned} \mathcal{L}(\mathbf{x}, \lambda) &= xy - \lambda_1(y - x) - \lambda_2(x^3 - y^4) \\ \nabla_x \mathcal{L}(\mathbf{x}, \lambda) &= \begin{bmatrix} y + \lambda_1 - 3x^2\lambda_2 \\ x - \lambda_1 + 4y^3\lambda_2 \end{bmatrix}, \end{aligned}$$

and state the KKT conditions as:

$$y + \lambda_1 - 3x^2\lambda_2 = 0 \tag{4a}$$
$$x - \lambda_1 + 4y^3\lambda_2 = 0 \tag{4b}$$
$$y - x \geq 0 \tag{4c}$$
$$x^3 - y^4 \geq 0 \tag{4d}$$
$$\lambda_i \geq 0, \quad i = 1, 2 \tag{4e}$$
$$\lambda_1(y - x) = 0 \tag{4f}$$
$$\lambda_2(x^3 - y^4) = 0. \tag{4g}$$

Now, we can check the different cases of active constraints to find KKT points. First, with no active constraints, i.e. $\lambda_1 = \lambda_2 = 0$, we get the point (0,0). In fact, this is a point with both constraints active; it just so happens that $\lambda_1 = \lambda_2 = 0$ here. One can check that the LICQ does not hold here, but it is still a KKT point because the gradient of f at (0,0) (which is 0) can be written as a non-negative linear combination of the gradients of the constraints. Thereby, we conclude that there exist no KKT points with no active constraints, but that (0,0) is a KKT point. What the failure of the LICQ at (0,0) implies is the following: There exists a function f which has a local minimum at (0,0), but for which (0,0) is not a KKT point. However, we might still be lucky for any given function f, as is the case here.

Next, we check with one active constraint. First, with $\lambda_1 \geq 0, \lambda_2 = 0$, we have from (4a) and (4b) that $x = \lambda_1$ and $y = -\lambda_1$. Inserting into the equality (4c) yields $\lambda_1 = 0$, and therefore $(x, y) = (0, 0)$ again, which has been discussed already.

With $\lambda_2 \geq 0, \lambda_1 = 0$, equations (4a), (4b) and (4d) become

$$\begin{aligned} y - 3x^2\lambda_2 &= 0, \\ x + 4y^3\lambda_2 &= 0, \\ x^3 &= y^4. \end{aligned}$$

Multiplying the first of these by x, the second by y, applying the third and adding the two first gives

$$y^4\lambda_2 = 0.$$

Any solution of this leads to the point (0,0), which we have already found to be a KKT point.

Finally, we check the case with two active constraints, for which there are two points; (0,0), which is already considered, and (1,1). In the point (1,1), we find (by (4a) and (4b)) that $\lambda_1 = -7$ and $\lambda_2 = -2$. All other KKT conditions are satisfied, so (1,1) is a KKT point candidate for a maximizer.

The only minimizer candidate we have is (0,0), for which the LICQ did not hold, and which therefore can be neither confirmed or discarded as a minimizer/maximizer using the second order necessary and sufficient conditions.

However, it is clearly a local (and even global) minimizer, as we are only considering nonnegative values for x and y, and since $f(x, y) = xy$, its global minimum (in the feasible set) is located at (0,0).

b) To find the critical cone C at $(0, 0)$, we use the definition given by equation (12.53) in N&W page 330. First, we find the gradients of the constraints at this point:

$$\nabla c_1(0,0) = \begin{bmatrix} -1 \\ 1 \end{bmatrix}, \qquad \nabla c_2(0,0) = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

Since $\lambda_1 = \lambda_2 = 0$ at this point, we have that $d \in C(0, 0)$ if and only if $\nabla c_1(0, 0)^T d \geq 0$ and $\nabla c_2(0, 0)^T d \geq 0$. The latter condition clearly holds for all d, and so we find that

$$\begin{aligned} C(0,0) &= \{d = (d_1, d_2) \in \mathbb{R}^2 : \nabla c_1(0,0)^T d \geq 0\} \\ &= \{d = (d_1, d_2) \in \mathbb{R}^2 : d_2 \geq d_1\}. \end{aligned}$$

Next, we find that the Hessian of the Lagrangian at $(0, 0)$ with Lagrange multipliers $\lambda^* = (\lambda_1, \lambda_2) = (0, 0)$ is given by

$$\nabla^2 \mathcal{L}((0,0); (\lambda_1, \lambda_2)) = \nabla^2 \mathcal{L}((0,0); (0,0)) = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}.$$

Thus, $d^T \nabla^2 \mathcal{L}((0, 0); (\lambda_1, \lambda_2))d = 2d_1 d_2$. There are clearly directions in the critical cone for which this is negative; one can choose $d_2 > 0$ and $d_1 < 0$.

c) We can find the tangent cone to the feasible set at $(0, 0)$ by looking at the limiting vectors along the lines $c_1(\mathbf{x}) = 0$ and $c_2(\mathbf{x}) = 0$ as $\mathbf{x} \to 0$. Traveling toward (0,0) along $c_1(\mathbf{x}) = 0$ we take, for example,

$$z_k = \begin{bmatrix} 1/k \\ 1/k \end{bmatrix}, \quad t_k = ||z_k|| = \sqrt{2}/k,$$

and find the limiting direction

$$d = \lim_{k \to \infty} \frac{z_k}{t_k} = \begin{bmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix}.$$

Along $c_2(\mathbf{x}) = 0$, we take

$$z_k = \begin{bmatrix} 1/k \\ 1/k^{3/4} \end{bmatrix}, \quad t_k = ||z_k|| = \frac{\sqrt{\sqrt{k} + 1}}{k},$$

and find the limiting direction

$$d = \lim_{k \to \infty} \frac{z_k}{t_k} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}.$$

The tangent cone at (0,0) contains all vectors between these limiting cases, which can be shown to be:

$$T(0,0) = \{d \in \mathbb{R}^2 : d_1 \geq 0 \text{ and } d_2 \geq d_1\}.$$

It is then easy to see that $d^T \nabla^2 \mathcal{L}((0, 0); (\lambda_1, \lambda_2))d = 2d_1 d_2 \geq 0$ for all d in the tangent cone at $(0, 0)$.

4 **a)** We start by writing the problem in the form

$$\min_{(x,y)\in\mathbb{R}^2} f(x,y), \text{ s.t. } c(x,y)=0,$$

where

$$f(x,y)=\frac{1}{2}(x^2+y^2) \text{ and } c(x,y)=xy-1.$$

We then form the Lagrangian

$$\mathcal{L}(x,y,\lambda)=f(x,y)-\lambda c(x,y),$$

and find its gradient

$$\nabla\mathcal{L}(x,y,\lambda)=\begin{bmatrix}x-\lambda y\\ y-\lambda x\end{bmatrix}.$$

We may note that $\nabla c(x,y)\neq 0$ for all $(x,y)\neq(0,0)$, meaning the LICQ holds for all feasible (x,y). We can therefore use the first order KKT conditions to identify candidates for extrema. Solving $\nabla\mathcal{L}(x,y,\lambda)=0$, we find

$$x=\lambda y \text{ and } y(1-\lambda^2)=0.$$

Thus, we may have either $\lambda=\pm1$ or $y=0$. But, if $y=0$, then $c(x,y)=1\neq0$, so this possibility is excluded. Also, taking $\lambda=-1$ yields imaginary values for x and y, which are disregarded. Therefore, we take $\lambda=1$ and continue by observing that

$$c(x,y)=c(y,y)=y^2-1=0\Rightarrow y=\pm1.$$

Thereby, we have the two solutions $(x,y)=(\pm1,\pm1)$, both of which have Lagrange multiplier $\lambda=1$. To verify that they are indeed minima, we check the second order sufficient conditions. We have that

$$\nabla c(x,y)=\begin{bmatrix}y\\ x\end{bmatrix},$$

so the LICQ holds at both points and in addition, for any w in the critical cone at $(\pm1,\pm1)$, we have $w=[\gamma,-\gamma]^T$, $\gamma\in\mathbb{R}$. At the two points we then get

$$w^T\nabla^2\mathcal{L}(x,y,\lambda)w=\begin{bmatrix}\gamma & -\gamma\end{bmatrix}\begin{bmatrix}1 & -1\\ -1 & 1\end{bmatrix}\begin{bmatrix}\gamma\\ -\gamma\end{bmatrix}=4\gamma^2\geq0,$$

with equality if and only if $\gamma=0$, i.e. $w=0$. Thus, $(x,y)=(\pm1,\pm1)$ are indeed minima.

Alternatively, we could argue that since f is coercive and Ω is closed, then there exist minimizers to the problem, and since the points are KKT points where the LICQ holds, these are the only two candidates for minimizers. Then, since f takes the same value at both points, both points must be global minimizers.

b) The quadratic penalty method considers the unconstrained minimization of the objective function

$$g(x,y)=f(x,y)+\frac{\mu}{2}c(x,y)^2=\frac{1}{2}(x^2+y^2)+\frac{\mu}{2}(xy-1)^2.$$

We find that

$$\nabla g(x,y) = \begin{bmatrix} x + \mu(xy^2 - y) \\ y + \mu(x^2 y - x) \end{bmatrix},$$

and solving $\nabla g(x,y) = 0$ gives (from the first component)

$$x = \frac{\mu y}{1 + \mu y^2}.$$

Inserting this into the second equation, we get

$$y + \mu\left(\frac{\mu^2 y^2}{(1+\mu y^2)^2} y - \frac{\mu y}{1+\mu y^2}\right) = 0.$$

One solution of this is $y = 0$, giving $(x,y) = (0,0)$. If $y \neq 0$, we can simplify the equation to

$$1 + \mu y^2 = \mu,$$

with solutions

$$y = \pm\sqrt{1 - \frac{1}{\mu}},$$

which exist (i.e. are real) as long as $\mu \geq 1$. From here, one can check that this also gives

$$x = \pm\sqrt{1 - \frac{1}{\mu}}.$$

It can also be checked that these points are minimizers as long as they exist. The solution $(x,y) = (0,0)$ is a minimizer while $\mu \leq 1$, but becomes a maximizer when $\mu > 1$. Also, we can see that as $\mu \to \infty$, $(x,y) \to (\pm 1, \pm 1)$.

c) The augmented Lagrangian for this problem is

$$L_A(x,y,\lambda,\mu) = \frac{1}{2}(x^2 + y^2) - \lambda(xy - 1) + \frac{\mu}{2}(xy-1)^2,$$

which is coercive and lower semi-continuous such that a minimizer exists, and it has gradient

$$\nabla L_A(x,y,\lambda,\mu) = \begin{bmatrix} x - \lambda y + \mu(xy^2 - y) \\ y - \lambda x + \mu(x^2 y - x) \end{bmatrix}.$$

After a similar computation to that in part b), we find

$$x = \frac{(\mu + \lambda) y}{1 + \mu y^2}$$

and the equation for y:

$$(1 + \mu y^2)^2 = (\lambda + \mu)^2.$$

In addition, we have the solution $(x,y) = (0,0)$. We must be somewhat careful in finding y. First, we have

$$1 + \mu y^2 = \pm(\lambda + \mu),$$

but since the left hand side is positive, we must choose the right hand side positive as well. Therefore, we have

$$1 + \mu y^2 = |\lambda + \mu|$$

and thus

$$y^* = \pm\sqrt{\left|\frac{\lambda}{\mu} + 1\right| - \frac{1}{\mu}},$$

which exist if $|\lambda + \mu| \geq 1$. It can be checked that here, too, we have $x^* = y^*$. The points (x^*, y^*) are the global minimizers if $\lambda + \mu \geq 1$. Otherwise, $(0, 0)$ is the global minimizer. We see that the original solution is obtained when either $\lambda = 1$ or $\mu \to \infty$. The fact that (x^*, y^*) are the global minimizers if $\lambda + \mu \geq 1$ can seen by checking when $\mathcal{L}_A(x^*, y^*, \lambda, \mu) \leq \mathcal{L}_A(0, 0, \lambda, \mu)$. This leads (after some computation) to the condition

$$(\lambda + \mu - 1)(|\lambda + \mu| - 1) \geq \frac{1}{2}(|\lambda + \mu| - 1)^2.$$

Since (x^*, y^*) exist only if $|\lambda + \mu| \geq 1$, and if $|\lambda + \mu| = 1$ then $(x^*, y^*) = (0, 0)$, we can divide by $|\lambda + \mu| - 1$ to obtain the condition

$$(\lambda + \mu - 1) \geq \frac{1}{2}(|\lambda + \mu| - 1),$$

which holds if $\lambda + \mu \geq 1$ but not if $\lambda + \mu \leq -1$.

d) We find the minimizers of

$$\Phi_1(x, y; \mu) = \frac{1}{2}(x^2 + y^2) + \mu|xy - 1|$$

by splitting the domain in three: $xy > 1$, $xy < 1$ and $xy = 1$. First, when $xy = 1$, we see that $x^2 = 1/y^2$, so the objective function takes the form

$$\Phi_1(x, y; \mu) = g(y) = \frac{1}{2}\left(\frac{1}{y^2} + y^2\right).$$

We can see that $g'(y) = 0$ when $y = 1$ or $y = -1$, and $g''(y) = 4$ in both these points, so they are minimizers along the curves $x = 1/y$, and we have the candidates (-1,-1) and (1,1). Furthermore, $\Phi_1(1, 1; \mu) = \Phi_1(-1, -1; \mu) = 1$ for all values of μ.
Next, if $xy > 1$ then

$$\Phi_1(x, y; \mu) = \frac{1}{2}(x^2 + y^2) + \mu(xy - 1),$$

so

$$\nabla\Phi_1(x, y; \mu) = \begin{bmatrix} x + \mu y \\ y + \mu x \end{bmatrix} = 0 \Rightarrow x = -\mu y \Rightarrow (1 - \mu^2)y = 0.$$

If $y = 0$ then $x = 0$, but this is not in the domain considered so we need to take $\mu = \pm 1$. Since $\mu > 0$, the only possibility is $\mu = 1$. This gives us the critical points along the line $x = -y$, but this is still not in the domain considered.

Thus, there are no critical points in the domain $xy > 1$.

Finally, in the domain $xy < 1$, we have

$$\Phi_1(x, y; \mu) = \frac{1}{2}(x^2 + y^2) - \mu(xy - 1),$$

so

$$\nabla\Phi_1(x, y; \mu) = \begin{bmatrix} x - \mu y \\ y - \mu x \end{bmatrix} = 0 \Rightarrow x = \mu y \Rightarrow (1 - \mu^2)y = 0.$$

If $y = 0$ then $x = 0$. This is in the domain and thus a critical point. Also, we may take $\mu = \pm 1$. Since $\mu > 0$, the only possibility is $\mu = 1$. This gives us the critical points along the line $x = y$, which are in the domain considered when $|x| < 1$. We now check whether any of these points are minimizers. Observe that

$$\nabla^2\Phi_1(x, y; \mu) = \begin{bmatrix} 1 & -\mu \\ -\mu & 1 \end{bmatrix}$$

with eigenvalues $\lambda = 1 \pm \mu$. The eigenvalues are positive when $\mu < 1$ and so the point (0,0) is a local minimizer when $\mu < 1$, with value $\Phi_1(0, 0; \mu) = \mu$, which actually makes it a global minimizer.

When $\mu = 1$, we have $\Phi_1(x, y; \mu) = 1$ along the line $x = y$. Also, when $\mu = 1$, we have

$$\Phi_1(x, y; \mu) = \frac{1}{2}(x - y)^2 + 1 \geq 1,$$

so these points are minimizers.

When $\mu > 1$, the global minimizers are found in $(x, y) = (\pm 1, \pm 1)$. This is because $\Phi_1(\pm 1, \pm 1, \mu) = 1$ and $\Phi_1(x, y; \mu) > 1$ elsewhere. This can be seen as following: When $xy > 1$,

$$\begin{aligned} \Phi_1(x, y; \mu) &= \frac{1}{2}(x^2 + y^2) + \mu(xy - 1) \\ &= \frac{1}{2}(x - y)^2 + \mu(xy - 1) + xy \\ &\geq \mu(xy - 1) + xy \\ &> 1, \end{aligned}$$

and when $xy < 1$:

$$\begin{aligned} \Phi_1(x, y; \mu) &= \frac{1}{2}(x^2 + y^2) - \mu(xy - 1) \\ &= \frac{1}{2}(x - y)^2 + xy - \mu(xy - 1) \\ &\geq \mu(1 - xy) + xy \\ &> 1. \end{aligned}$$

To summarize: When $\mu < 1$, we have a global minimizer in $(0, 0)$ with value μ. When $\mu = 1$, the global minimizers can be found on the line $x = y, x \in [-1, 1]$, and with $\mu > 1$, the global minimizers are found in $(x, y) = (\pm 1, \pm 1)$.

5 Forming the log-barrier function is pretty straightforward; following chapter 19.6 in N&W, we define

$$P(x;\mu) = \begin{cases} x + y - \mu \ln(1 - x^2 - y^2), & 1 - x^2 - y^2 \geq 0 \\ \infty, & 1 - x^2 - y^2 < 0 \end{cases}$$

To minimize this, we find that inside the unit circle,

$$\nabla P(x;\mu) = \begin{bmatrix} 1 + \mu \dfrac{2x}{1 - x^2 - y^2} \\ 1 + \mu \dfrac{2y}{1 - x^2 - y^2} \end{bmatrix}.$$

Setting $\nabla P(x;\mu) = 0$ one finds, after some computation, $x = y$ and

$$x = \frac{\mu}{2} \pm \sqrt{\frac{\mu^2}{4} + \frac{1}{2}}.$$

The positive solution is not a minimizer, as $1 - x^2 - y^2 < 0$, meaning the log-barrier function is infinite at this point, so we consider only the negative root,

$$x = \frac{\mu}{2} - \sqrt{\frac{\mu^2}{4} + \frac{1}{2}}.$$

This is a minimizer, as can be checked by computing the Hessian of P and checking that it is positive definite. Thus, the solution for each parameter μ is

$$(x, y) = \left(\frac{\mu}{2} - \sqrt{\frac{\mu^2}{4} + \frac{1}{2}}, \frac{\mu}{2} - \sqrt{\frac{\mu^2}{4} + \frac{1}{2}} \right).$$

It is worth noting that as $\mu \downarrow 0$, $(x, y) \to (-\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}})$, which is the solution of the original problem.

6 Take, for example, the minimization of the function $f(x, y) = -x^6 - y^6$ on the unit circle, i.e. with $c(x, y) = 1 - x^2 - y^2$. We then have

$$\mathcal{L}_A(x, y) = -x^6 - y^6 - \lambda(1 - x^2 - y^2) + \mu(1 - x^2 - y^2)^2,$$

which clearly is not bounded from below.

7 **a)** We are now considering the problem

$$\min_{x \in \mathbb{R}^n} f(x), \text{ s.t. } c(x) = 0,$$

where

$$f(x) = \frac{1}{2} x^T x \text{ and } c(x) = Ax - b,$$

with $b \neq 0$. The Lagrangian is

$$\mathcal{L}(x, \lambda) = \frac{1}{2} x^T x - \lambda^T (Ax - b),$$

where $\lambda \in \mathbb{R}^m$. The KKT conditions become

$$\nabla \mathcal{L}(x, \lambda) = x - A^T\lambda = 0,$$
$$Ax - b = 0.$$

Also, since A has full rank, then the LICQ hold everywhere, meaning the KKT conditions are necessary for minimizers. We therefore look for solutions that satisfy the KKT conditions. If $\lambda = 0$, then $x = 0$ and $Ax = 0$, meaning $Ax - b \neq 0$, so we must have $\lambda \neq 0$. The first condition then gives $x = A^T\lambda$, and inserting this into the second gives $AA^T\lambda = b$. Since A has full rank, AA^T is invertible and we have $\lambda = (AA^T)^{-1}b$, meaning $x = A^T(AA^T)^{-1}b$. Also, since $\nabla^2\mathcal{L}(x, \lambda) = \nabla^2 f(x) = I$, which is positive definite, this is a minimum.

b) The quadratic penalty method considers the unconstrained optimization of

$$g(x) = f(x) + \frac{\mu}{2}c(x)^T c(x),$$

which in our case becomes

$$g(x) = \frac{1}{2}x^T x + \frac{\mu}{2}(Ax - b)^T(Ax - b).$$

Taking the gradient of this, we get

$$\nabla g(x) = x + \mu(A^T Ax - A^T b) = 0$$
$$\Rightarrow \left(\frac{1}{\mu}I + A^T A\right)x = A^T b$$
$$\Rightarrow x = \left(\frac{1}{\mu}I + A^T A\right)^{-1} A^T b.$$

This is, however, not the expression we were looking for. We can easily see that

$$A^T\left(\frac{1}{\mu}I + AA^T\right) = \left(\frac{1}{\mu}I + A^T A\right)A^T.$$

Multiplying both sides from the left by $\left(\frac{1}{\mu}I + A^T A\right)^{-1}$ and from the right by $\left(\frac{1}{\mu}I + AA^T\right)^{-1}$, we see that

$$\left(\frac{1}{\mu}I + A^T A\right)^{-1} A^T = A^T\left(\frac{1}{\mu}I + AA^T\right)^{-1},$$

meaning that we get

$$x = A^T\left(\frac{1}{\mu}I + AA^T\right)^{-1} b.$$

Another way of arriving at the desired expression is by use of the singular value decomposition of A, writing $A = U\Sigma V^T$, where $U \in \mathbb{R}^{m\times m}$ and $V \in \mathbb{R}^{n\times n}$ are orthogonal matrices ($U^{-1} = U^T$ and $V^{-1} = V^T$) and $\Sigma \in \mathbb{R}^{m\times n}$ is the matrix containing the singular values of A along its diagonal. The singular values are all positive. We will write $I_{r\times r}$ for an $r \times r$ identity matrix. Now, we observe that

$$\begin{aligned}
\left(\frac{1}{\mu}I_{n\times n} + A^TA\right)^{-1} A^T &= \left(\frac{1}{\mu}I_{n\times n} + (U\Sigma V^T)^T U\Sigma V^T\right)^{-1} (U\Sigma V^T)^T \\
&= \left(\frac{1}{\mu}I_{n\times n} + V\Sigma^T U^T U\Sigma V^T\right)^{-1} V\Sigma^T U^T \\
&= \left(\frac{1}{\mu}I_{n\times n} + V\Sigma^T \Sigma V^T\right)^{-1} V\Sigma^T U^T \\
&= \left(V\left(\frac{1}{\mu}I_{n\times n} + \Sigma^T \Sigma\right) V^T\right)^{-1} V\Sigma^T U^T \\
&= V\left(\frac{1}{\mu}I_{n\times n} + \Sigma^T \Sigma\right)^{-1} V^T V\Sigma^T U^T \\
&= V\left(\frac{1}{\mu}I_{n\times n} + \Sigma^T \Sigma\right)^{-1} \Sigma^T U^T \\
&= V\Sigma\left(\frac{1}{\mu}I_{m\times m} + \Sigma \Sigma^T\right)^{-1} U^T \\
&= V\Sigma U^T U\left(\frac{1}{\mu}I_{m\times m} + \Sigma \Sigma^T\right)^{-1} U^T \\
&= A^T\left(\frac{1}{\mu}I_{m\times m} + U\Sigma \Sigma^T U^T\right)^{-1} \\
&= A^T\left(\frac{1}{\mu}I_{m\times m} + U\Sigma V^T V \Sigma^T U^T\right)^{-1} \\
&= A^T\left(\frac{1}{\mu}I_{m\times m} + AA^T\right)^{-1}.
\end{aligned}$$

Thereby, we have $x_\mu = A^T\left(\frac{1}{\mu}I_{m\times m} + AA^T\right)^{-1} b$. The fact that

$$\left(\frac{1}{\mu}I_{n\times n} + \Sigma^T\Sigma\right)^{-1}\Sigma^T = \Sigma\left(\frac{1}{\mu}I_{m\times m} + \Sigma\Sigma^T\right)^{-1}$$

can be checked by writing the product componentwise.

c) We now consider the problem

$$\min_{x\in\mathbb{R}^n} f(x), \text{ s.t. } c(x) \geq 0,$$

where

$$f(x) = \frac{1}{2}x^Tx \text{ and } c(x) = \epsilon - \frac{1}{2}||Ax-b||^2,$$

The KKT conditions for this problem are

$$\begin{aligned}
\nabla\mathcal{L}(x,\lambda) = x + \lambda(A^TAx - A^Tb) &= 0 \\
\lambda(\epsilon - \frac{1}{2}||Ax-b||^2) &= 0 \\
\epsilon - \frac{1}{2}||Ax-b||^2 &\geq 0 \\
\lambda &\geq 0.
\end{aligned}$$

With $\lambda = 0$, we get $x = 0$. For the third condition to hold, we must have $\epsilon \geq ||b||^2/2$. This is then a valid KKT point. Also, we have $\nabla^2 \mathcal{L}(x, 0) = I$, which is positive definite, so it is a minimum.
If $\lambda \neq 0$, we get, as in the previous exercise, that

$$\hat{x}_e = A^T \left(\frac{1}{\lambda} I_{m\times m} + AA^T \right)^{-1} b.$$

Here, λ must satisfy the condition that $\lambda > 0$ and λ must solve

$$\epsilon - \frac{1}{2} \left|\left| \left(AA^T \left(\frac{1}{\lambda} I_{m\times m} + AA^T \right)^{-1} - I_{m\times m} \right) b \right|\right|^2 = 0.$$

We can show that such a λ exists; since f is coercive and Ω is bounded and closed, there must exist a global minimizer. Since the LICQ holds, the KKT conditions are necessary for a minimum, and since, if $\epsilon < \frac{1}{2}||b||^2$, our candidate $\hat{x}_e$ is the only KKT point, it must be the global minimum, and thereby have a λ satisfying the above conditions. Thus, by taking $\mu = \lambda$, we get $\hat{x}_\epsilon = x_\mu$.

8

a) We start by finding the minimizer with respect to x of

$$\mathcal{L}(x, \lambda) = \frac{1}{2} x^T x - \lambda^T (Ax - b).$$

One can see that

$$\nabla_x \mathcal{L}(x, \lambda) = x - A^T \lambda = 0 \Rightarrow x = A^T \lambda,$$
$$\nabla_x^2 \mathcal{L}(x, \lambda) = I_{n\times n}.$$

Thus, the (unique) minimizer is $x = A^T \lambda$ and so

$$g(\lambda) = \min_{x\in\mathbb{R}^n} \mathcal{L}(x, \lambda) = \mathcal{L}(A^T\lambda, \lambda) = -\frac{1}{2}\lambda^T AA^T \lambda + \lambda^T b.$$

b) Next, we search for a maximizer of $g(\lambda)$, observing that

$$\nabla g(\lambda) = -AA^T \lambda + b,$$
$$\nabla^2 g(\lambda) = -AA^T.$$

Since $-AA^T$ is negative definite, there exists a unique solution to the maximization problem given by $\nabla g(\lambda) = 0 \Rightarrow AA^T \lambda = b$.

c) First, we observe that

$$\max_{\lambda\in\mathbb{R}^m} \min_{x\in\mathbb{R}^n} \mathcal{L}(x, \lambda) = \max_{\lambda\in\mathbb{R}^m} g(\lambda) = g((AA^T)^{-1} b) = \frac{1}{2} b^T (AA^T)^{-1} b.$$

Concerning the value

$$\min_{x\in\mathbb{R}^n} \max_{\lambda\in\mathbb{R}^m} \mathcal{L}(x, \lambda),$$

we see that if $Ax \neq b$, we can take $\lambda = \gamma(Ax - b)$ and let $\gamma \to -\infty$ to get infinite values of $\mathcal{L}(x, \lambda)$, so that we have

$$\max_{\lambda \in \mathbb{R}^m} \mathcal{L}(x, \lambda) = \begin{cases} \infty, & Ax \neq b \\ \frac{1}{2}||x||^2, & Ax = b. \end{cases}$$

Therefore, we consider only x such that $Ax = b$, and then the problem is equivalent to the one in exercise 7a, i.e.

$$\min_{x \in \mathbb{R}^n} \frac{1}{2}||x||^2 \quad \text{s.t.} \quad Ax = b.$$

This has the solution $x = A^T(AA^T)^{-1}b$, and the value is

$$\frac{1}{2}||A^T(AA^T)^{-1}b||^2 = \frac{1}{2}b^T(AA^T)^{-T}AA^T(AA^T)^{-1}b = \frac{1}{2}b^T(AA^T)^{-1}b,$$

since $(AA^T)^{-T} = ((AA^T)^T)^{-1} = (AA^T)^{-1}$.

9 We have the problem

$$\max_{x,y} c^T x + d^T y, \tag{5}$$
$$A_1 x = b_1 \tag{6}$$
$$A_2 x + B_2 y \leq b_2 \tag{7}$$
$$y \geq l \tag{8}$$
$$y \leq u, \tag{9}$$

with no constraints on x, and we want to write it on standard form

$$\min C^T X \text{ s.t. } AX = B, \quad X \geq 0. \tag{10}$$

There are several modifications needed to achieve this. First, we split x into $x = x^+ - x^-$, where $x^+, x^- \geq 0$. Inequalities (8) and (9) can be taken care of by introducing slack variables $v \geq 0$ and $w \geq 0$ such that they become

$$y - v = l$$
$$y + w = u.$$

In inequality (7) we also introduce a slack variable $z \geq 0$ such that it becomes

$$A_2(x^+ - x^-) + B_2 y + z = b_2,$$

and finally, the equality condition (6) becomes

$$A_1(x^+ - x^-) = b_1.$$

The maximization problem (5) becomes a minimization problem by negating the objective function:

$$\min -c^T(x^+ - x^-) - d^T y.$$

In total, the problem can be written on the standard form (10) if we let

$$X = \begin{bmatrix} x^+ \\ x^- \\ y \\ v \\ w \\ z \end{bmatrix}, C = \begin{bmatrix} -c \\ c \\ d \\ 0 \\ 0 \\ 0 \end{bmatrix}, A = \begin{bmatrix} A_1 & -A_1 & 0 & 0 & 0 & 0 \\ A_2 & -A_2 & B_2 & 0 & 0 & I \\ 0 & 0 & I & -I & 0 & 0 \\ 0 & 0 & I & 0 & I & 0 \end{bmatrix}, B = \begin{bmatrix} b_1 \\ b_2 \\ l \\ u \end{bmatrix}.$$

10 Given the optimization problem

$$\min_x c^T x \quad \text{s.t.} \quad Ax \geq b, \quad x \geq 0, \tag{11}$$

we want to show that its dual is the problem

$$\max_\lambda b^T \lambda \quad \text{s.t.} \quad A^T \lambda \leq c, \quad \lambda \geq 0. \tag{12}$$

We start by defining the Lagrangian function

$$\begin{aligned} \mathcal{L}(x, \lambda, s) &= c^T x - \lambda^T (Ax - b) - s^T x \\ &= \lambda^T b + x^T (c - A^T \lambda - s), \end{aligned}$$

and observe that (11) is equivalent to the problem

$$\min_x \max_{\substack{\lambda \geq 0 \\ s \geq 0}} \mathcal{L}(x, \lambda, s),$$

since

$$\max_{\substack{\lambda \geq 0 \\ s \geq 0}} \mathcal{L}(x, \lambda, s) = \begin{cases} \infty, & Ax - b < 0 \quad \text{or} \quad x < 0 \\ c^T x, & Ax - b \geq 0 \quad \text{and} \quad x \geq 0 \end{cases}.$$

The dual problem is then defined as

$$\max_{\substack{\lambda \geq 0 \\ s \geq 0}} \min_x \mathcal{L}(x, \lambda, s),$$

where we can observe that

$$\min_x \mathcal{L}(x, \lambda, s) = \begin{cases} -\infty, & A^T \lambda + s \neq c \\ \lambda^T b, & A^T \lambda + s = c \end{cases},$$

so that the dual problem is equivalent to

$$\max_{\substack{\lambda \geq 0 \\ s \geq 0}} \lambda^T b, \quad \text{s.t.} \quad A^T \lambda + s = c.$$

Now, we can interpret s as a slack variable in an inequality constraint, so that the dual problem is indeed

$$\max_\lambda b^T \lambda \quad \text{s.t.} \quad A^T \lambda \leq c, \quad \lambda \geq 0.$$

11 We can see that this linear program is already in standard form with

$$c = \begin{bmatrix} 5 \\ 3 \\ 4 \end{bmatrix}, \quad x = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \quad A = \begin{bmatrix} 1 & 1 & 1 \end{bmatrix} \quad b = 1$$

From chapter 13.1 in N&W, we know that this problem has the dual

$$\max b^T \lambda \quad \text{s.t.} \quad A^T \lambda \leq c,$$

in other words,

$$\max \lambda \quad \text{s.t.} \quad \lambda \leq 5, \quad \lambda \leq 3 \quad \text{and} \quad \lambda \leq 4.$$

The solution of this problem is, of course, very easy; it is simply $\lambda = 3$. Note that this gives us a convenient way of solving the primal problem; since only the second constraint is active, we have $x_1 = x_3 = 0$, and so the constraint of the primal problem yields that $x_2 = 1$.

12 **a)** We can see that the problem is a quadratic minimization problem

$$\min \frac{1}{2} x^T G x + c^T x \quad \text{s.t.} \quad a_i^T x - b_i \geq 0,$$

where

$$G = \begin{bmatrix} 8 & 2 \\ 2 & 2 \end{bmatrix}, \quad c = \begin{bmatrix} 2 \\ 3 \end{bmatrix}, \quad a_1 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}, \quad a_2 = \begin{bmatrix} -1 \\ -1 \end{bmatrix}, \quad a_3 = \begin{bmatrix} -1 \\ 0 \end{bmatrix},$$

and $b_1 = 0$, $b_2 = -4$ and $b_3 = -3$. We can check that G is positive definite, so by Theorem 16.4 in N&W, the KKT conditions are necessary and sufficient for minimizers. We therefore set up the KKT conditions:

$$8x + 2y + 2 - \lambda_1 + \lambda_2 + \lambda_3 = 0 \tag{13a}$$
$$2x + 2y + 3 + \lambda_1 + \lambda_3 = 0 \tag{13b}$$
$$\lambda_1(x - y) = 0 \tag{13c}$$
$$\lambda_2(4 - x - y) = 0 \tag{13d}$$
$$\lambda_3(3 - x) = 0 \tag{13e}$$
$$x - y \geq 0 \tag{13f}$$
$$4 - x - y \geq 0 \tag{13g}$$
$$3 - x \geq 0. \tag{13h}$$

We see from (14a) and (14b) that

$$x = \frac{1}{6} + \frac{1}{3}\lambda_1 - \frac{1}{6}\lambda_3$$
$$y = -\frac{5}{3} - \frac{5}{6}\lambda_1 - \frac{1}{2}\lambda_2 - \frac{1}{6}\lambda_3.$$

Now, we can go through the usual procedure of considering all options for active constraints. With no active constraints, i.e. $\lambda_1 = \lambda_2 = \lambda_3 = 0$, we get $(x, y) = (\frac{1}{6}, -\frac{5}{3})$ which is, in fact, a KKT point and as such a global solution of the problem. We should end the search for a minimum here, since the problem is strictly convex and the minimizer is unique, as confirmed by figure 4. However, since the following discussion will prove useful in part b), we carry on looking for KKT points.

Next, we consider cases where only one constraint is active.

First, if $\lambda_1 = \lambda_2 = 0$, i.e. $3 - x = 0$, we get $\lambda_3 = -17$ and $(x, y) = (-\frac{16}{6}, -\frac{27}{16})$, which breaks constraint (14g).
Next, if $\lambda_1 = \lambda_3 = 0$, i.e. $4 - x - y = 0$, we get $\lambda_2 = -11$ and $(x, y) = (\frac{1}{6}, \frac{23}{6})$, which breaks constraint (14f).
Lastly, if $\lambda_2 = \lambda_3 = 0$, i.e. $x - y = 0$, we get $\lambda_1 = -\frac{11}{7}$ and $(x, y) = (-\frac{15}{42}, -\frac{15}{42})$. It is a feasible point, but has a negative Lagrange multiplier, meaning it is a candidate for a maximizer. This will prove useful in part b).

Next, we consider cases where two constraints are active.

First, if $\lambda_1 = 0$, i.e. $3 - x = 0$ and $4 - x - y = 0$, we get $(x, y) = (3, 1)$ with corresponding Lagrange multipliers $\lambda_2 = \frac{1}{3}$ and $\lambda_3 = -17$, meaning it is not a KKT point.
Next, if $\lambda_2 = 0$, i.e. $3 - x = 0$ and $x - y = 0$, we get $(x, y) =$ and $(x, y) = (3, 3)$, which breaks constraint (14g).
Lastly, if $\lambda_3 = 0$, i.e. $x - y = 0$ and $4 - x - y = 0$, we get $(x, y) = (2, 2)$ with corresponding Lagrange multipliers $\lambda_1 = \frac{11}{2}$ and $\lambda_2 = -\frac{33}{2}$, meaning it is not a KKT point.

There are no points in which all three constraints are active. Thus, we have one candidate for a minimizer, $(x, y) = (\frac{1}{6}, -\frac{5}{3})$, which is the global minimizer. Figure 4 shows the feasible domain and the contour lines of the objective function which confirm our observations.

b) Replacing f by $-f$ will turn minima into maxima and vice versa. Especially of note is that since $f \to \infty$ as $x^2 + y^2 \to \infty$, then $-f \to -\infty$, meaning there is no global solution to the minimization problem. However, the maximizer we found in the last problem, $(x, y) = (-\frac{15}{42}, -\frac{15}{42})$, now becomes local minimizer.

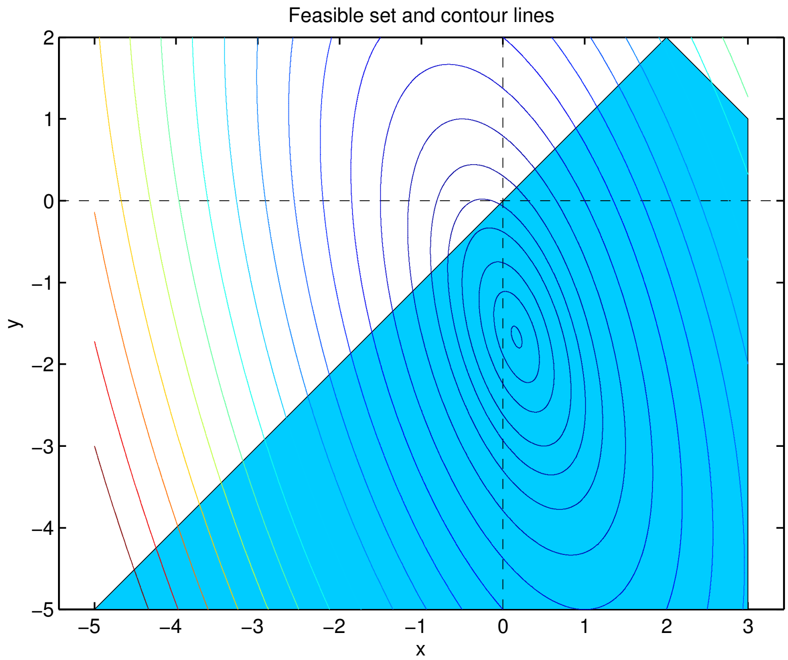

Figure 4: Feasible set (light blue) and contour lines of the function. Note: The feasible set extends further toward infinity.