Exam: Continuous Optimisation 2016

Monday 12$^{\text{th}}$ December 2016

1. We will consider the first step in iterative methods from $\mathbf{x}_0 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ to attempt to minimise the function $f : \mathbb{R}^2 \to \mathbb{R}, \quad f(\mathbf{x}) = 2x_1^2 + x_2^2 \exp(x_1) - x_1 - x_2 \quad \text{over } \mathbb{R}^2$.
   (a) Starting from $\mathbf{x}_0$, considering the direction of steepest descent, $\mathbf{d}_S$, as the search direction and exact line search (i.e. $\lambda_0 \in \arg\min_{\lambda \in \mathbb{R}}\{f(\mathbf{x}_0 + \lambda \mathbf{d}_S)\}$), evaluate $\mathbf{x}_1 = \mathbf{x}_0 + \lambda_0 \mathbf{d}_S$. [2 points]
   (b) Starting from $\mathbf{x}_0$, considering Newton's direction, $\mathbf{d}_N$, as the search direction (not normalised), and $\lambda_0 = 1$, evaluate $\mathbf{x}_1 = \mathbf{x}_0 + \mathbf{d}_N$. [2 points]

Solution:

(a) We have

$$\nabla f(\mathbf{x}) = \begin{pmatrix} 4x_1 + x_2^2 \exp(x_1) - 1 \\ 2x_2 \exp(x_1) - 1 \end{pmatrix}, \qquad \nabla f(\mathbf{x}_0) = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \qquad \mathbf{d}_S = -\nabla f(\mathbf{x}_0) = -\begin{pmatrix} 0 \\ 1 \end{pmatrix},$$

$$\lambda_0 \in \arg\min_{\lambda}\{f(0, 1-\lambda)\} = \arg\min_{\lambda}\{(1 - \lambda - \tfrac{1}{2})^2 - \tfrac{1}{4}\} = \{\tfrac{1}{2}\}, \qquad \lambda_0 = \tfrac{1}{2},$$

$$\mathbf{x}_1 = \mathbf{x}_0 + \lambda_0 \mathbf{d}_S = \begin{pmatrix} 0 \\ 1/2 \end{pmatrix}. \qquad \left(f(\mathbf{x}_0) = 0, \qquad f(\mathbf{x}_1) = -1/4 \right)$$

(b) We have

$$\nabla^2 f(\mathbf{x}) = \begin{pmatrix} 4 + x_2^2 \exp(x_1) & 2x_2 \exp(x_1) \\ 2x_2 \exp(x_1) & 2\exp(x_1) \end{pmatrix}, \qquad \nabla^2 f(\mathbf{x}_0) = \begin{pmatrix} 5 & 2 \\ 2 & 2 \end{pmatrix},$$

$$\left[\nabla^2 f(\mathbf{x}_0)\right]^{-1} = \frac{1}{6}\begin{pmatrix} 2 & -2 \\ -2 & 5 \end{pmatrix},$$

$$\mathbf{d}_0 = -\left[\nabla^2 f(\mathbf{x}_0)\right]^{-1} \nabla f(\mathbf{x}_0) = -\frac{1}{6}\begin{pmatrix} 2 & -2 \\ -2 & 5 \end{pmatrix}\begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 1/3 \\ -5/6 \end{pmatrix},$$

$$\mathbf{x}_1 = \mathbf{x}_0 + \lambda_0 \mathbf{d}_0 = \begin{pmatrix} 1/3 \\ 1/6 \end{pmatrix}.$$

$$\left(f(\mathbf{x}_1) = -5/18 + \exp(1/3)/36 \approx -0.239 \right),$$

As a point of interest, at global minimiser: $\mathbf{x}^* \approx \begin{pmatrix} 0.199 \\ 0.410 \end{pmatrix}$ and $f(\mathbf{x}^*) \approx -0.325$

2. (a) Consider two convex sets $\mathcal{A} \subseteq \mathbb{R}^n$ and $\mathcal{B} \subseteq \mathbb{R}$, and two convex functions $h : \mathcal{A} \to \mathcal{B}$ and $g : \mathcal{B} \to \mathbb{R}$, with g also being a monotonically increasing function on $\mathcal{B}$. [3 points]
For $f : \mathcal{A} \to \mathbb{R}$ given by $f(\mathbf{x}) = g(h(\mathbf{x}))$, show that f is a convex function.

(b) For a norm $\| \bullet \|$ on $\mathbb{R}^n$ and a convex function $f : \mathbb{R}^n \to \mathbb{R}$, consider using the barrier method to solve the problem $\min_{\mathbf{x}}\{f(\mathbf{x}) : \|\mathbf{x}\| \leq 1\}$.
Let $\widehat{\mathcal{F}} = \{\mathbf{x} \in \mathbb{R}^n : \|\mathbf{x}\| < 1\}$ and $b : \widehat{F} \to \mathbb{R}$ be given by $b(\mathbf{x}) = (1 - \|\mathbf{x}\|)^{-2}$.

i. Justify that b is a valid barrier function for this problem. [1 point]

ii. Show that b is a convex function. [2 points]

Solution:

(a) Consider arbitrary $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ and $0 \leq \lambda \leq 1$. We need to show that

$$f(\lambda\mathbf{x} + (1-\lambda)\mathbf{y}) \leq \lambda f(\mathbf{x}) + (1-\lambda) f(\mathbf{y}).$$

We have

$$\begin{aligned}
h(\lambda\mathbf{x} + (1-\lambda)\mathbf{y}) &\leq \lambda h(\mathbf{x}) + (1-\lambda)h(\mathbf{y}) && \text{as } h \text{ is convex}\\
g(h(\lambda\mathbf{x} + (1-\lambda)\mathbf{y})) &\leq g(\lambda h(\mathbf{x}) + (1-\lambda)h(\mathbf{y})) && \text{as } g \text{ is monotonically increasing}\\
g(\lambda h(\mathbf{x}) + (1-\lambda)h(\mathbf{y})) &\leq \lambda g(h(\mathbf{x})) + (1-\lambda)g(h(\mathbf{y})) && \text{as } g \text{ is convex}\\
f(\lambda\mathbf{x} + (1-\lambda)\mathbf{y}) &\leq \lambda f(\mathbf{x}) + (1-\lambda) f(\mathbf{y}) && \text{combining these inequalities.}
\end{aligned}$$

(b) i. As norms are continuous, so is b. We also thus have that $\mathbf{y} \in \mathrm{bd}(\widehat{\mathcal{F}})$ if and only if $\|\mathbf{y}\| = 1$, and thus $\lim_{\substack{\mathbf{x}\in\widehat{\mathcal{F}}\\ \mathbf{x}\to\mathbf{y}}} b(\mathbf{x}) = \infty$.

ii. Let $\mathcal{A} = \widehat{\mathcal{F}}$ and $\mathcal{B} = [0, 1)$, and consider the functions $h : \mathcal{A} \to \mathcal{B}$ and $g : \mathcal{B} \to \mathbb{R}$ given by $h(\mathbf{x}) = \|\mathbf{x}\|$ and $g(y) = (1 - y)^{-2}$.
We have that $\mathcal{B}$ is trivially a convex set, and $\mathcal{A}$ is a convex set by Corollary 1.16.
We have that h is a convex function by Exercise 1.4
We have $g'(y) = 2(1-y)^{-3} > 0$ for all $y \in \mathcal{B}$, and thus g is monotonically increasing.
We have $g''(y) = 6(1-y)^{-4} > 0$ for all $y \in \mathcal{B}$, and thus g is convex.
Therefore, by part *(a)* of this question, b is a convex function.

3. Consider the problem

$$\begin{aligned}
\min_{\mathbf{x}} \quad & x_2 \\
\text{s. t.} \quad & x_1^2 \leq x_1 + x_2 \\
& 2x_1 \leq x_1^2 + x_2
\end{aligned} \tag{P}$$

(a) Is (P) a convex optimisation problem? Justify your answer. [2 points]

(b) Find a strictly feasible descent direction for the problem (P) at $\widehat{\mathbf{x}} = \begin{pmatrix} 2 \\ 2 \end{pmatrix}$. [2 points]

(c) i. Show that the Linear Independency Constraint Qualification holds at all feasible points of (P). [2 points]

ii. Find the KKT points for (P). [3 points]

iii. Given that the optimal solution to (P) is attained, find the global minimiser and optimal value to this problem. Justify your answer. [1 point]

iv. Provide justification for this global minimiser being a strict local minimiser of order 1. [1 point]

(d) Formulate and solve the Lagrangian dual problem to (P). Is there strong duality? [4 points]

Solution:

(a) We have

$$f(\mathbf{x}) = x_2, \qquad \nabla f(\mathbf{x}) = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \qquad \nabla^2 f(\mathbf{x}) = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix},$$

$$g_1(\mathbf{x}) = x_1^2 - x_1 - x_2, \quad \nabla g_1(\mathbf{x}) = \begin{pmatrix} 2x_1 - 1 \\ -1 \end{pmatrix}, \quad \nabla^2 g_1(\mathbf{x}) = \begin{pmatrix} 2 & 0 \\ 0 & 0 \end{pmatrix},$$

$$g_2(\mathbf{x}) = 2x_1 - x_1^2 - x_2, \quad \nabla g_2(\mathbf{x}) = \begin{pmatrix} 2 - 2x_1 \\ -1 \end{pmatrix}, \quad \nabla^2 g_2(\mathbf{x}) = \begin{pmatrix} -2 & 0 \\ 0 & 0 \end{pmatrix},$$

There exists $\mathbf{x} \in \mathbb{R}^2$ such that $\nabla^2 g_2(\mathbf{x})$ is not positive semidefinite, and thus the problem is not convex. (In fact the matrix is not positive semidefinite at all $\mathbf{x} \in \mathbb{R}^2$.)

(b) At $\widehat{\mathbf{x}} = \begin{pmatrix} 2 \\ 2 \end{pmatrix}$ we have

$$\nabla f(\widehat{\mathbf{x}}) = \begin{pmatrix} 0 \\ 1 \end{pmatrix},$$

$$g_1(\widehat{\mathbf{x}}) = 2^2 - 2 - 2 = 0, \qquad \nabla g_1(\widehat{\mathbf{x}}) = \begin{pmatrix} 2 * 2 - 1 \\ -1 \end{pmatrix} = \begin{pmatrix} 3 \\ -1 \end{pmatrix},$$

$$g_2(\widehat{\mathbf{x}}) = 2 * 2 - 2^2 - 2 = -2 < 0.$$

The active set at $\widehat{\mathbf{x}}$ is thus $\mathcal{J}_{\widehat{\mathbf{x}}} = \{1\}$, and we are looking for $\mathbf{h} \in \mathbb{R}^2$ such that $\nabla f(\widehat{\mathbf{x}})^\mathsf{T}\mathbf{h} < 0$ and $\nabla g_1(\widehat{\mathbf{x}})^\mathsf{T}\mathbf{h} < 0$. Equivalently, we want $3h_1 < h_2 < 0$. For example, $\mathbf{h} = \begin{pmatrix} -1 \\ -2 \end{pmatrix}$ is a strictly feasible descent direction at $\widehat{\mathbf{x}}$.

(c) i. We have

$$g_1(\mathbf{x}) = x_1^2 - x_1 - x_2, \qquad \nabla g_1(\mathbf{x}) = \begin{pmatrix} 2x_1 - 1 \\ -1 \end{pmatrix},$$

$$g_2(\mathbf{x}) = 2x_1 - x_1^2 - x_2, \qquad \nabla g_2(\mathbf{x}) = \begin{pmatrix} 2 - 2x_1 \\ -1 \end{pmatrix}.$$

Suppose for the sake of contradiction that LICQ does not hold at $\mathbf{x}$. Then we must have that $\mathcal{J}_{\mathbf{x}} = \{1,2\}$, and $\nabla g_1(\mathbf{x})$ and $\nabla g_2(\mathbf{x})$ are not linearly independent.

Therefore $\begin{pmatrix} 2x_1 - 1 \\ -1 \end{pmatrix} = \mu \begin{pmatrix} 2 - 2x_1 \\ -1 \end{pmatrix}$ for some $\mu \in \mathbb{R}$, implying that $\mu = 1$ and $2x_1 - 1 = 2 - 2x_1$, or equivalently $x_1 = 3/4$. We have $1 \in \mathcal{J}_{\mathbf{x}}$ and thus $0 = g_1(\mathbf{x}) = x_1^2 - x_1 - x_2 = 9/16 - 3/4 - x_2 = -3/16 - x_2$, or equivalently $x_2 = -3/16$. Finally, as $2 \in \mathcal{J}_{\mathbf{x}}$, we get the contradiction $0 = g_2(\mathbf{x}) = 2x_1 - x_1^2 - x_2 = 3/2 - 9/16 + 3/16 = 24/16 - 9/16 + 3/16 = 18/16$.

(Alternatively: If $\mathcal{J}_{\mathbf{x}} = \{1,2\}$ then $0 = g_1(\mathbf{x}) = x_1^2 - x_1 - x_2$ and $0 = g_2(\mathbf{x}) = 2x_1 - x_1^2 - x_2$. Therefore $x_1^2 - x_1 = 2x_1 - x_1^2$, or equivalently $0 = 2x_1^2 - 3x_1 = 2x_1\left(x_1 - \frac{3}{2}\right)$. We now consider two cases:

1. If $x_1 = 0$ then $x_2 = x_1^2 - x_1 = 0$. We then have $\nabla g_1(\mathbf{x}) = \begin{pmatrix} -1 \\ -1 \end{pmatrix}$ and $\nabla g_2(\mathbf{x}) = \begin{pmatrix} 2 \\ -1 \end{pmatrix}$, which are clearly linearly independent vectors.
2. If $x_1 = 3/2$ then $x_2 = x_1^2 - x_1 = 9/4 - 6/4 = 3/4$. We then have $\nabla g_1(\mathbf{x}) = \begin{pmatrix} 2 \\ -1 \end{pmatrix}$ and $\nabla g_2(\mathbf{x}) = \begin{pmatrix} -1 \\ -1 \end{pmatrix}$, which are clearly linearly independent vectors.)

ii. We have that $\mathbf{x} \in \mathbb{R}^2$ is a KKT point if it is feasible and there exists $\boldsymbol{\lambda} \in \mathbb{R}^2_+$ such that

$$\begin{aligned} 0 &= \lambda_1 g_1(\mathbf{x}), \\ 0 &= \lambda_2 g_2(\mathbf{x}), \\ \nabla f(\mathbf{x}) &= -\lambda_1 \nabla g_1(\mathbf{x}) - \lambda_2 \nabla g_2(\mathbf{x}). \end{aligned}$$

We final equality is equivalent to

$$\begin{pmatrix} 0 \\ 1 \end{pmatrix} = -\lambda_1 \begin{pmatrix} 2x_1 - 1 \\ -1 \end{pmatrix} - \lambda_2 \begin{pmatrix} 2 - 2x_1 \\ -1 \end{pmatrix},$$

which is in turn equivalent to

$$0 = \lambda_1(1 - 2x_1) + \lambda_2(2x_1 - 2),$$
$$1 = \lambda_1 + \lambda_2.$$

We now consider 3 cases:

1. $\underline{\lambda_1 = 0}$: Then we have $\lambda_2 = 1 - \lambda_1 = 1$ and $0 = \lambda_1(1 - 2x_1) + \lambda_2(2x_1 - 2) = 2x_1 - 2$, implying that $x_1 = 1$. As $\lambda_2 > 0$ we also require $0 = g_2(\mathbf{x}) = 2x_1 - x_1^2 - x_2 = 2 - 1 - x_2$, implying that $x_2 = 1$. We then check $g_1(\mathbf{x}) = x_1^2 - x_1 - x_2 = 1 - 1 - 1 = -1 \leq 0$. Therefore $\mathbf{x} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ is a KKT point, with multipliers $\boldsymbol{\lambda} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$.
2. $\underline{\lambda_2 = 0}$: Then we have $\lambda_1 = 1 - \lambda_2 = 1$ and $0 = \lambda_1(1 - 2x_1) + \lambda_2(2x_1 - 2) = 1 - 2x_1$, implying that $x_1 = 1/2$. As $\lambda_1 > 0$ we also require $0 = g_1(\mathbf{x}) = x_1^2 - x_1 - x_2 = 1/4 - 1/2 - x_2 = -1/4 - x_2$, implying that $x_2 = -1/4$. We have $g_2(\mathbf{x}) = 2x_1 - x_1^2 - x_2 = 1 - 1/4 + 1/4 = 1 > 0$. Therefore this point is infeasible, and there is no KKT point in this case.
3. $\underline{\lambda_1, \lambda_2 > 0}$: Then $0 = g_1(\mathbf{x}) = g_2(\mathbf{x})$ and thus $x_1^2 - x_1 = x_2 = 2x_1 - x_1^2$. Therefore $0 = 2x_1^2 - 3x_1 = 2x_1(x_1 - 3/2)$, and thus $x_1 = 0$ or $x_1 = 3/2$. We consider these two cases separately:
   (a) $\underline{x_1 = 0}$: Then $x_2 = x_1^2 - x_1 = 0$. We then have $1 = \lambda_1 + \lambda_2$ and $0 = \lambda_1(1 - 2x_1) + \lambda_2(2x_1 - 2) = \lambda_1 - 2\lambda_2$. This implies that $\boldsymbol{\lambda} = \begin{pmatrix} 2/3 \\ 1/3 \end{pmatrix}$ and $\mathbf{x} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$ is a KKT point.
   (b) $\underline{x_1 = 3/2}$: Then $x_2 = x_1^2 - x_1 = 9/4 - 3/2 = 3/4$. We then have $1 = \lambda_1 + \lambda_2$ and $0 = \lambda_1(1 - 2x_1) + \lambda_2(2x_1 - 2) = -2\lambda_1 + \lambda_2$. This implies that $\boldsymbol{\lambda} = \begin{pmatrix} 1/3 \\ 2/3 \end{pmatrix}$ and $\mathbf{x} = \begin{pmatrix} 3/2 \\ 3/4 \end{pmatrix}$ is a KKT point.

iii. Two alternative answers:

1. Any global minimiser is also a local minimiser. As LICQ holds everywhere in this problem, from Remark 5.11, a local minimiser is also a KKT point. We have three KKT points as possible global minimisers, and by comparison we have that the global minimiser is $\mathbf{x}^* = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$, and the optimal value is zero.
2. For any $\mathbf{x} \in \mathbb{R}^2$ feasible we have $x_2 \geq x_1^2 - x_1 = x_1(x_1 - 1) \geq 0$ for $x_1 \in [0, 1]$ and $x_2 \geq 2x_1 - x_1^2 = x_1(2 - x_1) > 0$ for $x_1 \notin [0, 2]$. Therefore $x_2 \geq 0$ for all $\mathbf{x}$ feasible, with equality if and only if $\mathbf{x} = \mathbf{0}$.

iv. The conditions of Theorem 5.13 hold at $\mathbf{x}^* = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$, i.e. LICQ holds and $\mathcal{J}_{\mathbf{x}} = \{1, 2\}$, and thus this is a strict local minimiser of order 1.

(d) We have

$$\begin{aligned} L(\mathbf{x};\mathbf{y}) &= f(\mathbf{x}) + y_1 g_1(\mathbf{x}) + y_2 g_2(\mathbf{x}) \\ &= x_2 + y_1(x_1^2 - x_1 - x_2) + y_2(2x_1 - x_1^2 - x_2) \\ &= (y_1 - y_2)x_1^2 + (2y_2 - y_1)x_1 + (1 - y_1 - y_2)x_2, \\ \psi(\mathbf{y}) &= \inf_{\mathbf{x}\in\mathbb{R}^2} L(\mathbf{x};\mathbf{y}). \end{aligned}$$

We now consider the following four cases:

1. $\underline{y_1 + y_2 \neq 1}$: Then considering $x_2 \to \pm\infty$ we get $\psi(\mathbf{y}) = -\infty$.
2. $\underline{y_1 + y_2 = 1 \text{ and } y_1 < y_2}$: Then from the negative coefficient of the x_1^2 term of $L(\mathbf{x};\mathbf{y})$, we see that considering $x_1 \to \infty$ we have $\psi(\mathbf{y}) = -\infty$.
3. $\underline{y_1 + y_2 = 1 \text{ and } y_1 = y_2}$: Then $y_1 = y_2 = 1/2$ and $L(\mathbf{x};\mathbf{y}) = x_1/2$. Considering $x_1 \to -\infty$, this implies that $\psi(\mathbf{y}) = -\infty$.
4. $\underline{y_1 + y_2 = 1 \text{ and } y_1 > y_2}$: Then

$$L(\mathbf{x};\mathbf{y}) = (y_1 - y_2)x_1^2 + (2y_2 - y_1)x_1,$$

which, when considering $\mathbf{y} \in \mathbb{R}^2$ fixed, is a quadratic function in x_1 with a strictly positive coefficient on the x_1^2 term. From the example on the minimum of a Quadratic function from the slides, we then have

$$\psi(\mathbf{y}) = \frac{-(2y_2 - y_1)^2}{4(y_1 - y_2)} = \frac{-(2 - 3y_1)^2}{4(2y_1 - 1)}.$$

The dual problem is thus

$$\begin{aligned} \max_{\mathbf{y}\in\mathbb{R}^2} \quad & \frac{-(2y_2 - y_1)^2}{4(y_1 - y_2)} \\ \text{s.t.} \quad & y_1 + y_2 = 1, \qquad y_1 > y_2 \geq 0. \end{aligned}$$

For all feasible points of this problem the objective function is less than or equality to zero, with equality if and only if $2y_2 = y_1 = 1 - y_2$. Therefore the optimal solution to the dual problem is $\mathbf{y} = \begin{pmatrix} 2/3 \\ 1/3 \end{pmatrix}$, and its optimal value is zero. We thus have strong duality.

4. For $n \in \mathbb{N}$, consider a proper cone $\mathcal{L} \subseteq \mathbb{R}^n$ and a nonsingular matrix $\mathsf{A} \in \mathbb{R}^{n\times n}$. We then let $\mathcal{K} = \mathsf{A}\mathcal{L} := \{\mathsf{A}\mathbf{x} : \mathbf{x} \in \mathcal{L}\} \subseteq \mathbb{R}^n$.
   (a) Show that $\mathcal{K}$ is a convex cone. [1 point]
   (b) Show that $\mathcal{K}$ is pointed. [1 point]
   (c) Find $\mathcal{K}^*$, the dual cone to $\mathcal{K}$, in terms of $\mathcal{L}^*$. [2 points]
   (d) Show that $\mathcal{K}^*$ is pointed. [2 points]
   (e) Show that $\mathcal{K}$ is a proper cone. (You may assume that $\mathcal{K}$ is closed.) [1 point]

Solution:

(a) Consider arbitrary $\mathbf{u}, \mathbf{v} \in \mathcal{K}$ and $\boldsymbol{\lambda} \in \mathbb{R}^2_+$. We need to show that $\lambda_1\mathbf{u} + \lambda_2\mathbf{v} \in \mathcal{K}$.

There exists $\mathbf{x}, \mathbf{y} \in \mathcal{L}$ such that $\mathbf{u} = \mathsf{A}\mathbf{x}$ and $\mathbf{v} = \mathsf{A}\mathbf{y}$. As $\mathcal{L}$ is a convex cone we have $\lambda_1\mathbf{x} + \lambda_2\mathbf{y} \in \mathcal{L}$, and thus $\lambda_1\mathbf{u} + \lambda_2\mathbf{v} = \mathsf{A}(\lambda_1\mathbf{x} + \lambda_2\mathbf{y}) \in \mathcal{K}$.

(b) Consider an arbitrary $\mathbf{u} \in \mathbb{R}^n$ such that $\pm\mathbf{u} \in \mathcal{K}$. We need to show that $\mathbf{u} = \mathbf{0}$.

There exists $\mathbf{x}, \mathbf{y} \in \mathcal{L}$ such that $\mathbf{u} = \mathsf{A}\mathbf{x}$ and $-\mathbf{u} = \mathsf{A}\mathbf{y}$. Therefore $\mathbf{0} = \mathbf{u} + (-\mathbf{u}) = \mathsf{A}(\mathbf{x} + \mathbf{y})$ and $\mathbf{x} + \mathbf{y} = \mathsf{A}^{-1}\mathbf{0} = \mathbf{0}$, or equivalently $\mathbf{y} = -\mathbf{x}$. Therefore $\pm\mathbf{x} \in \mathcal{L}$, implying that $\mathbf{x} = \mathbf{0}$ and $\mathbf{u} = \mathsf{A}\mathbf{0} = \mathbf{0}$.

(c)
$$\begin{aligned}\mathcal{K}^* &= \{\mathbf{y} \in \mathbb{R}^n : \mathbf{u}^\mathsf{T}\mathbf{y} \geq 0 \text{ for all } \mathbf{u} \in \mathcal{K}\}\\ &= \{\mathbf{y} \in \mathbb{R}^n : (\mathsf{A}\mathbf{x})^\mathsf{T}\mathbf{y} \geq 0 \text{ for all } \mathbf{x} \in \mathcal{L}\}\\ &= \{\mathbf{y} \in \mathbb{R}^n : \mathbf{x}^\mathsf{T}(\mathsf{A}^\mathsf{T}\mathbf{y}) \geq 0 \text{ for all } \mathbf{x} \in \mathcal{L}\}\\ &= \{\mathbf{y} \in \mathbb{R}^n : \mathsf{A}^\mathsf{T}\mathbf{y} \in \mathcal{L}^*\} = \{\mathsf{A}^{-\mathsf{T}}\mathbf{z} : \mathbf{z} \in \mathcal{L}^*\} = \mathsf{A}^{-\mathsf{T}}\mathcal{L}^*\end{aligned}$$

Any of the answers from the final line are correct.

(d) Three alternative proofs:

1. Consider arbitrary $\mathbf{u} \in \mathbb{R}^n$ such that $\pm\mathbf{u} \in \mathcal{K}^*$. We need to show that $\mathbf{u} = \mathbf{0}$.
   As $\mathcal{L}$ is a proper cone, so is $\mathcal{L}^*$.
   We have $\mathsf{A}^\mathsf{T}\mathbf{u} \in \mathcal{L}^*$ and $-(\mathsf{A}^\mathsf{T}\mathbf{u}) = \mathsf{A}^\mathsf{T}(-\mathbf{u}) \in \mathcal{L}^*$, and thus $\mathsf{A}^\mathsf{T}\mathbf{u} = \mathbf{0}$. Therefore $\mathbf{u} = \mathsf{A}^{-\mathsf{T}}\mathbf{0} = \mathbf{0}$.
2. Consider arbitrary $\mathbf{u} \in \mathbb{R}^n$ such that $\pm\mathbf{u} \in \mathcal{K}^*$. We need to show that $\mathbf{u} = \mathbf{0}$.
   As $\pm\mathbf{u} \in \mathcal{K}^*$ we have $\langle\mathbf{u}, \mathbf{y}\rangle \geq 0$ and $\langle-\mathbf{u}, \mathbf{y}\rangle \geq 0$ for all $\mathbf{y} \in \mathcal{K}$, and thus $\langle\mathbf{u}, \mathbf{y}\rangle = 0$ for all $\mathbf{y} \in \mathcal{K}$.
   Therefore $0 = \langle\mathbf{u}, \mathsf{A}\mathbf{x}\rangle = \mathbf{u}^\mathsf{T}\mathsf{A}\mathbf{x} = (\mathsf{A}^\mathsf{T}\mathbf{u})^\mathsf{T}\mathbf{x}$ for all $\mathbf{x} \in \mathcal{L}$, and as $\mathcal{L}$ is full dimensional, this implies that $\mathbf{0} = \mathsf{A}^\mathsf{T}\mathbf{u}$ and we get the contradiction $\mathbf{0} = \mathsf{A}^{-\mathsf{T}}\mathsf{A}^\mathsf{T}\mathbf{u} = \mathbf{u}$.

3. We will show the equivalent result that $\mathcal{K}$ is full-dimensional.

   As $\mathcal{L}$ is full dimensional there exists linearly independent vectors $\mathbf{x}_1, \dots, \mathbf{x}_n \in \mathcal{L}$. Then letting $\mathbf{u}_i = \mathsf{A}\mathbf{x}_i$ for all i, we have $\mathbf{u}_1, \dots, \mathbf{u}_n \in \mathcal{K}$. The proof is completed if we can show that $\mathbf{u}_1, \dots, \mathbf{u}_n$ are linearly independent.

   Suppose for the sake of contradiction that $\mathbf{u}_1, \dots, \mathbf{u}_n$ are not linearly independent. Then $\exists \boldsymbol{\lambda} \in \mathbb{R}^n \setminus \{\mathbf{0}\}$ such that $\mathbf{0} = \sum_{i=1} \lambda_i \mathbf{u}_i = \mathsf{A}(\sum_{i=1}^n \lambda_i \mathbf{x}_i)$. Therefore $\mathbf{0} = \mathsf{A}^{-1}\mathbf{0} = \sum_{i=1} \lambda_i \mathbf{x}_i$. As $\mathbf{x}_1, \dots, \mathbf{x}_n$ are linearly independent, this implies the contradiction $\boldsymbol{\lambda} = \mathbf{0}$.

(e) From Definition 7.9, $\mathcal{K}$ is a proper cone if it is a closed convex cone which is pointed and full dimensional.

We can assume that $\mathcal{K}$ is a closed set, and from part *(a)* we have that $\mathcal{K}$ is a convex cone.

From part *(b)* we have that $\mathcal{K}$ is a pointed set.

From part *(d)* we have that $\mathcal{K}^*$ is a pointed set, and thus by Theorem 8.11, we have that $\mathcal{K}$ is full-dimensional.

5. For $\mathbf{b} \in \mathbb{R}^m$ and $\mathsf{A}_1, \dots, \mathsf{A}_m \in \mathcal{S}^n$, consider the problem of varying $\mathbf{y} \in \mathbb{R}^m$ in order to minimise $\mathbf{b}^\mathsf{T}\mathbf{y}$, with the constraint that all the eigenvalues of $\sum_{i=1}^m y_i \mathsf{A}_i$ are between minus one and plus two.

   (a) Formulate this problem as a conic optimisation problem in a standard form. [2 points]

   (b) Find the dual problem to this conic optimisation problem. [2 points]

   If you were unable to solve part (a), then as an alternative question to (b): Find the dual problem to $\max_{\mathbf{y}}\{\mathbf{b}^\mathsf{T}\mathbf{y} : (\ \mathbf{c},\ \mathsf{C}\ ) + \sum_{i=1}^m y_i\ (\ \mathbf{a}_i,\ \mathsf{A}_i\ ) \in \mathbb{R}^p_+ \times \mathcal{PSD}^n\}$, with the vectors $\mathbf{c}, \mathbf{a}_1, \dots, \mathbf{a}_m \in \mathbb{R}^p$ and the matrix $\mathsf{C} \in \mathcal{S}^n$.

Solution:

(a)

$$\begin{aligned} \min_{\mathbf{y}} \quad & \mathbf{b}^\mathsf{T}\mathbf{y} \\ \text{s.t.} \quad & -\mathrm{I} \preceq \sum_{i=1}^m y_i \mathsf{A}_i \preceq 2\mathrm{I} \end{aligned}$$

$$\begin{aligned} -\max_{\mathbf{y}} \quad & -\mathbf{b}^\mathsf{T}\mathbf{y} \\ \text{s.t.} \quad & (\ \mathrm{I},\ 2\mathrm{I}\ ) - \sum_{i=1}^m y_i\ (\ -\mathsf{A}_i,\ \mathsf{A}_i\ ) \ \in\ \mathcal{PSD}^n \times \mathcal{PSD}^n \end{aligned}$$

Equivalent answer:

$$\begin{aligned} -\max_{\mathbf{y}} \quad & -\mathbf{b}^\mathsf{T}\mathbf{y} \\ \text{s.t.} \quad & \begin{pmatrix} \mathrm{I} & \mathrm{O} \\ \mathrm{O} & 2\mathrm{I} \end{pmatrix} - \sum_{i=1}^{m} y_i \begin{pmatrix} -\mathsf{A}_i & \mathrm{O} \\ \mathrm{O} & \mathsf{A}_i \end{pmatrix} \in \mathcal{PSD}^{2n} \end{aligned}$$

(b)

$$\begin{aligned} -\min_{\mathsf{V},\mathsf{W}} \quad & \langle (\ \mathrm{I},\ 2\mathrm{I}\ ), \mathsf{X} \rangle \\ \text{s.t.} \quad & \langle (\ -\mathsf{A}_i,\ \mathsf{A}_i\ ), \mathsf{X} \rangle = -b_i \text{ for all } i = 1, \dots, m \\ & \mathsf{X} \in \mathcal{PSD}^n \times \mathcal{PSD}^n, \end{aligned}$$

$$\begin{aligned} \max_{\mathsf{V},\mathsf{W}} \quad & -\langle \mathrm{I}, \mathsf{V} \rangle - 2\langle \mathrm{I}, \mathsf{W} \rangle \\ \text{s.t.} \quad & -\langle \mathsf{A}_i, \mathsf{V} \rangle + \langle \mathsf{A}_i, \mathsf{W} \rangle = -b_i \text{ for all } i = 1, \dots, m \\ & \mathsf{V}, \mathsf{W} \in \mathcal{PSD}^n. \end{aligned}$$

Equivalent answer:

$$\begin{aligned} \max_{\mathsf{X}} \quad & -\left\langle \begin{pmatrix} \mathrm{I} & \mathrm{O} \\ \mathrm{O} & 2\mathrm{I} \end{pmatrix}, \mathsf{X} \right\rangle \\ \text{s.t.} \quad & \left\langle \begin{pmatrix} -\mathsf{A}_i & \mathrm{O} \\ \mathrm{O} & \mathsf{A}_i \end{pmatrix}, \mathsf{X} \right\rangle = -b_i \text{ for all } i = 1, \dots, m \\ & \mathsf{X} \in \mathcal{PSD}^{2n}. \end{aligned}$$

Solution to alternative question:

$$\begin{aligned} \max_{\mathbf{y}} \quad & \mathbf{b}^\mathsf{T}\mathbf{y} \\ \text{s.t.} \quad & (\ \mathbf{c},\ \mathsf{C}\ ) - \sum_{i=1}^{m} y_i\ (\ -\mathbf{a}_i,\ \ -\mathsf{A}_i\ ) \in \mathbb{R}^p_+ \times \mathcal{PSD}^n \end{aligned}$$

$$\begin{aligned} \min_{\mathbf{x},\mathsf{X}} \quad & \langle \mathbf{c}, \mathbf{x} \rangle + \langle \mathsf{C}, \mathsf{X} \rangle \\ \text{s.t.} \quad & -\langle \mathbf{a}_i, \mathbf{x} \rangle - \langle \mathsf{A}_i, \mathsf{X} \rangle = b_i \text{ for all } i = 1, \dots, m \\ & \mathbf{x} \in \mathbb{R}^p_+,\ \mathsf{X} \in \mathcal{PSD}^n. \end{aligned}$$

6. (Automatic additional points) [4 points]

Question:	1	2	3	4	5	6	Total
Points:	4	6	15	7	4	4	40

A copy of the lecture-sheets may be used during the examination. You may use any results from the lecture slides in your answers (Lemmas, Theorems, Corollaries, Exercises, etc.), however you should reference the result.

Good Luck!

Hints:

1. *g is a monotonically increasing function on $\mathcal{B} \subseteq \mathbb{R}$ if for all $a, b \in \mathcal{B}$ with $a \leq b$ we have $g(a) \leq g(b)$.*
2. *If g is differentiable in $\mathcal{B} \subseteq \mathbb{R}$ then g is a monotonically increasing function on $\mathcal{B}$ if and only if $g'(z) \geq 0$ for all $z \in \mathcal{B}$.*
3. *One of the properties of a norm is that it is a continuous function.*
4. $\begin{pmatrix} a & b \\ b & c \end{pmatrix}^{-1} = \frac{1}{ac - b^2} \begin{pmatrix} c & -b \\ -b & a \end{pmatrix}$
5. *The following are equivalent for $\mathsf{A} \in \mathbb{R}^{n \times n}$:*
   - A *is a nonsingular matrix;*
   - A *has an inverse matrix;*
   - A^T *has an inverse matrix.*