Practice Exam 2: Continuous Optimisation

1. Let $f_j(\mathbf{x}),\ j = 1, \dots, k\ (1 \le k \in \mathbb{N})$, be convex functions defined on a convex set $\mathcal{C} \subset \mathbb{R}^n$.

   (a) Consider with (given) $\alpha_j \ge 0,\ j = 1, \dots, k$, the function $f(\mathbf{x}) := \sum_{j=1}^k \alpha_j f_j(\mathbf{x})$. Show that f is convex on $\mathcal{C}$. [3 points]

   (b) Show that also $g(\mathbf{x}) := \max_{1 \le j \le k} \{f_j(\mathbf{x})\}$ is a convex function on $\mathcal{C}$. [3 points]

Solution:

(a) For $\mathbf{x}, \mathbf{y} \in \mathcal{C}, \lambda \in [0,1]$ we find using convexity of the f_j's and $\alpha_j \ge 0$:

$$
\begin{aligned}
f(\lambda \mathbf{x} + (1-\lambda)\mathbf{y}) &= \sum_{j=1}^k \alpha_j f_j(\lambda \mathbf{x} + (1-\lambda)\mathbf{y}) \\
\textit{f_j convex, $\alpha_j \ge 0$} \qquad &\le \sum_{j=1}^k \alpha_j [\lambda f_j(\mathbf{x}) + (1-\lambda) f_j(\mathbf{y})] \\
&= \lambda \left[\sum_{j=1}^k \alpha_j f_j(\mathbf{x}) \right] + (1-\lambda) \left[\sum_{j=1}^k \alpha_j f_j(\mathbf{y}) \right] \\
&= \lambda f(\mathbf{x}) + (1-\lambda) f(\mathbf{y})
\end{aligned}
$$

(b) For $\mathbf{x}, \mathbf{y} \in \mathcal{C}, \lambda \in [0,1]$ we find using convexity of the f_j's:

$$
\begin{aligned}
g(\lambda \mathbf{x} + (1-\lambda)\mathbf{y}) &= \max_{1 \le j \le k} \{f_j(\lambda \mathbf{x} + (1-\lambda)\mathbf{y})\} \\
&\le \max_{1 \le j \le k} \{\lambda f_j(\mathbf{x}) + (1-\lambda) f_j(\mathbf{y})\} \\
&\le \lambda \left[\max_{1 \le j \le k} f_j(\mathbf{x}) \right] + (1-\lambda) \left[\max_{1 \le j \le k} \{f_j(\mathbf{y})\} \right] \\
&= \lambda g(\mathbf{x}) + (1-\lambda) g(\mathbf{y})
\end{aligned}
$$

where in the second $\le$ we used that "max of a positive sum of functions $\le$ positive sum of max of the functions".

2. Consider the convex program [3 points]

$$
(CO) \qquad \min_{x \in \mathbb{R}^n} f(\mathbf{x}) \quad \text{s.t.} \quad \mathbf{x} \in \mathcal{F} := \{ \mathbf{x} \in \mathbb{R}^n \mid g_j(\mathbf{x}) \le 0,\ j = 1, \dots, m \},
$$

with convex functions $f, g_j \in C^1(\mathbb{R}^n, \mathbb{R})$.
Show that if $\overline{\mathbf{x}} \in \mathcal{F}$ satisfies the KKT-conditions (Karush-Kuhn-Tucker conditions) for (CO) with a multiplier vector $\overline{\mathbf{y}} \ge 0$ then $(\overline{\mathbf{x}}, \overline{\mathbf{y}})$ is a saddle point for the Lagrangian function $L(\mathbf{x}, \mathbf{y})$ of (CO).

Solution: KKT-conditions means that $\overline{\mathbf{x}} \in \mathcal{F}$ satisfies with $\overline{\mathbf{y}} \geq 0$,

$$(\nabla_x L(\overline{\mathbf{x}}, \overline{\mathbf{y}}) =) \quad \nabla f(\overline{\mathbf{x}}) + \sum_{j \in J} \overline{y}_j \nabla g_j(\overline{\mathbf{x}}) = 0 \quad \text{with } \overline{y}_j g_j(\overline{\mathbf{x}}) = 0 \; \forall j \in J \; .$$

So (by Th. 3.4) $\overline{\mathbf{x}}$ is a global solution of $\min_{x \in \mathbb{R}^n} L(\mathbf{x}, \overline{\mathbf{y}})$ and thus

$$L(\overline{\mathbf{x}}, \overline{\mathbf{y}}) \leq L(\mathbf{x}, \overline{\mathbf{y}}) \quad \forall \mathbf{x} \in \mathbb{R}^n \qquad (\star)$$

Moreover since $\overline{\mathbf{x}}$ is feasible, i.e., $g_j(\overline{\mathbf{x}}) \leq 0 \; \forall j$, and using $\overline{y}_j g_j(\overline{\mathbf{x}}) = 0$ we obviously obtain for all $\mathbf{y} \geq \mathbf{0}$:

$$L(\overline{\mathbf{x}}, \mathbf{y}) = f(\overline{\mathbf{x}}) + \sum_{j \in J} y_j g_j(\overline{\mathbf{x}}) \leq f(\overline{\mathbf{x}}) = f(\overline{\mathbf{x}}) + \sum_{j \in J} \overline{y}_j g_j(\overline{\mathbf{x}}) = L(\overline{\mathbf{x}}, \overline{\mathbf{y}}) \; .$$

Together with $(\star)$ this shows that $(\overline{\mathbf{x}}, \overline{\mathbf{y}})$ is a saddle point of L.

3. Consider the two problems

$$\begin{array}{llll} (P_1) & \min_{\mathbf{x} \in \mathbb{R}^2} f(\mathbf{x}) & \text{s.t.} & g_1(\mathbf{x}) := x_1^2 - x_2 \leq 0 \\ (P_2) & \min_{\mathbf{x} \in \mathbb{R}^2} f(\mathbf{x}) & \text{s.t.} & g_2(\mathbf{x}) := -x_1^2 - x_2 \leq 0 \end{array}$$

both with the same objective $f(\mathbf{x}) = 2x_1^2 + x_2$.

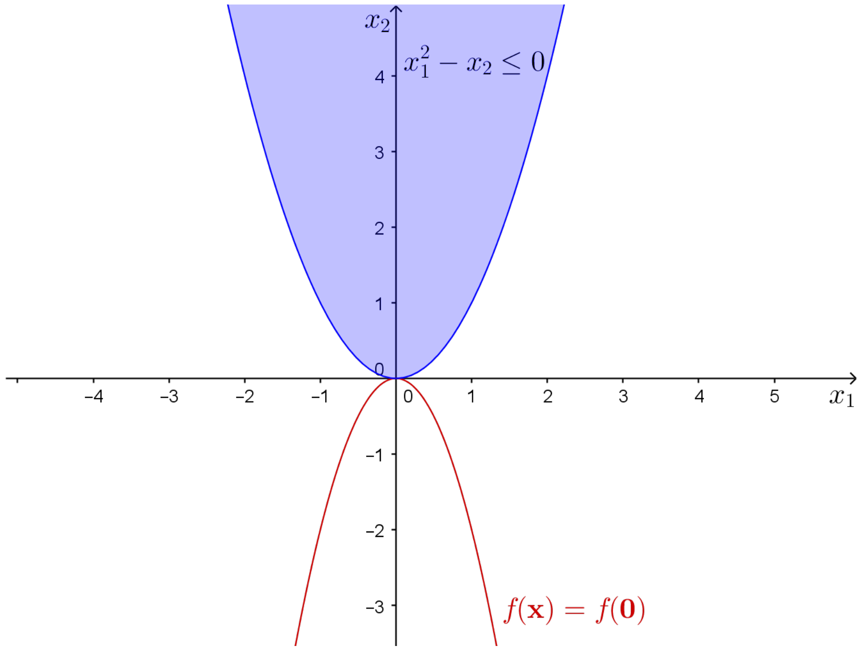
(a) Which of these programs (P_1), (P_2) is a convex problem? Sketch for both problems the feasible set and the level set of f given by $f(\mathbf{x}) = f(0,0)$. [3 points]

(b) Determine for both programs a (the) KKT-point $\overline{\mathbf{x}}$ with corresponding Lagrangean multiplier $\overline{\mu}$. [3 points]

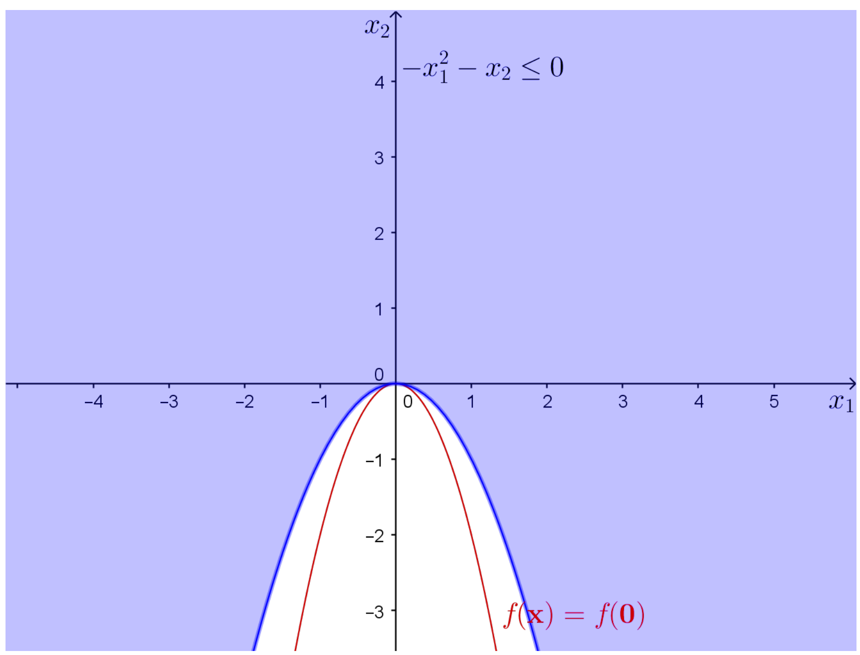
(c) Show for both problems that $\overline{\mathbf{x}}$ is a (local) minimizer. Is it a global minimizer? [4 points]

Solution:

(a) f, g_1 are convex (e.g., show that Hessian is pos. semidef.). But g_2 is not convex, $\nabla^2 g_2(\mathbf{x}) = \begin{pmatrix} -2 & 0 \\ 0 & 0 \end{pmatrix}$ is not positive semidefinite. So (P_1) is convex, (P_2) is not.



Above is a sketch of problem (P_1). The feasible set is coloured blue and the level curve is coloured red.



Above is a sketch of problem (P_2). The feasible set is coloured blue and the level curve is coloured red.

(b) The KKT-conditions read

For (P_1): $\begin{pmatrix}4x_1\\1\end{pmatrix}+\mu_1\begin{pmatrix}2x_1\\-1\end{pmatrix}=0$ with unique solution $\mu_1=1, \overline{x}_1=\overline{x}_2=0$

For (P_2): $\begin{pmatrix}4x_1\\1\end{pmatrix}+\mu_2\begin{pmatrix}-2x_1\\-1\end{pmatrix}=0$ with unique solution $\mu_2=1, \overline{x}_1=\overline{x}_2=0$

Note that for both g_1, g_2 must be active.

(c) Since (P_1) is convex the KKT-point $\overline{\mathbf{x}}=0$ must be a global minimizer (see Th. 3.12).
Since (P_2) is not convex we have to check the second order conditions (in Th. 5.14) (or we can directly argue as below): we compute

$$C_{\overline{\mathbf{x}}}=\{\mathbf{d}\mid \nabla f(\overline{\mathbf{x}})^\mathsf{T}\mathbf{d}\leq 0, \nabla g_2(\overline{\mathbf{x}})^\mathsf{T}\mathbf{d}\leq 0\}=\{\mathbf{d}=(d_1,d_2)\mid d_2=0\}$$

and thus

$$\mathbf{d}^\mathsf{T}\nabla^2 L(\overline{\mathbf{x}},\mu_2)\mathbf{d}=\mathbf{d}^\mathsf{T}\left[\begin{pmatrix}4&0\\0&0\end{pmatrix}+\mu_2\begin{pmatrix}-2&0\\0&0\end{pmatrix}\right]\mathbf{d}=d_1^2(4-2\mu_2)=2d_1^2>0$$

for all $d=(d_1,0)\in C_{\overline{\mathbf{x}}}\setminus\{\mathbf{0}\}$, *i.e.*, $d_1\neq 0$. So $\overline{\mathbf{x}}$ is a local minimizer.
It is a global minimizer since $g_2(\mathbf{x})\leq 0$, or equivalently $-x_1^2\leq x_2$, implies:

$$2x_1^2+x_2\geq 2x_1^2-x_1^2=x_1^2\geq 0=f(\overline{\mathbf{x}})\quad\forall\text{feasible }\mathbf{x}.$$

4. QUESTION OMITTED [3 points]

5. Let $\mathcal{K}=\{\mathbf{x}\in\mathbb{R}^n\mid \|\mathbf{x}\|_2\leq \mathbf{1}^\mathsf{T}\mathbf{x}\}$, where $\mathbf{1}\in\mathbb{R}^n$ is the all-ones vector, and $\|\bullet\|_2$ is the Euclidean norm.

(a) Show that $\mathcal{K}$ is a proper cone. [You may assume closure.] [4 points]

(b) Show that the vectors $\mathbf{1}$ and $(\mathbf{1}-\mathbf{e}_i)$ are in $\mathcal{K}^*$ for all $i=1,\dots,n$, where $\mathbf{e}_i\in\mathbb{R}^n$ is the unit vector with the first entry equal to one and all other entries equal to zero. [2 points]

(c) Show that $\mathcal{K}^*\subseteq\mathbb{R}^n_+$. [1 point]

Solution:

(a) In order to show that $\mathcal{K}$ is a proper cone, we need to show that it is a closed convex pointed full-dimensional cone. We assume closure and will now prove the rest of the properties:

- *Convex cone:*
  Let $\mathbf{x}, \mathbf{y} \in \mathcal{K}$ and $\lambda_1, \lambda_2 > 0$. We have $\|\mathbf{x}\|_2 \leq \mathbf{1}^\mathsf{T}\mathbf{x}$ and $\|\mathbf{y}\|_2 \leq \mathbf{1}^\mathsf{T}\mathbf{y}$. Therefore, letting $\mathbf{z} = \lambda_1\mathbf{x} + \lambda_2\mathbf{y}$, we have

$$\begin{aligned}\|\mathbf{z}\|_2 &= \|\lambda_1\mathbf{x} + \lambda_2\mathbf{y}\|_2 \\ &\leq \|\lambda_1\mathbf{x}\|_2 + \|\lambda_2\mathbf{y}\|_2 \\ &= \lambda_1\|\mathbf{x}\|_2 + \lambda_2\|\mathbf{y}\|_2 \\ &\leq \lambda_1\mathbf{1}^\mathsf{T}\mathbf{x} + \lambda_2\mathbf{1}^\mathsf{T}\mathbf{y} \\ &= \mathbf{1}^\mathsf{T}(\lambda_1\mathbf{x} + \lambda_2\mathbf{y}) \\ &= \mathbf{1}^\mathsf{T}\mathbf{z}.\end{aligned}$$

  This implies then implies that $\mathbf{z} \in \mathcal{K}$.

- *Pointed:*
  Suppose we have $\pm\mathbf{x} \in \mathcal{K}$. Then $\|\mathbf{x}\|_2 \leq \mathbf{1}^\mathsf{T}\mathbf{x}$ and $\|-\mathbf{x}\|_2 \leq \mathbf{1}^\mathsf{T}(-\mathbf{x})$. Therefore $2\|\mathbf{x}\|_2 = \|\mathbf{x}\|_2 + \|-\mathbf{x}\|_2 \leq \mathbf{1}^\mathsf{T}\mathbf{x} - \mathbf{1}^\mathsf{T}\mathbf{x} = 0$, and thus $\mathbf{x} = \mathbf{0}$.

- *Full-dimensional:*
  The vectors $\mathbf{e}_1, \ldots, \mathbf{e}_n \in \mathbb{R}^n$ are n linearly independent vectors and for all i we have $\|\mathbf{e}_i\|_2 = 1 = \mathbf{1}^\mathsf{T}\mathbf{e}_i$.

(b) We have $\mathcal{K}^* = \{\mathbf{y} \in \mathbb{R}^n \mid \mathbf{x}^\mathsf{T}\mathbf{y} \geq 0 \text{ for all } \mathbf{x} \in \mathcal{K}\}$.
For all $\mathbf{x} \in \mathcal{K}$ we have $\mathbf{1}^\mathsf{T}\mathbf{x} \geq \|\mathbf{x}\|_2 \geq 0$, and thus $\mathbf{1} \in \mathcal{K}^*$.
For all $\mathbf{x} \in \mathcal{K}$ we have $(\mathbf{1} - \mathbf{e}_i)^\mathsf{T}\mathbf{x} = \mathbf{1}^\mathsf{T}\mathbf{x} - \mathbf{e}_i^\mathsf{T}\mathbf{x} \geq \|\mathbf{x}\|_2 - \|\mathbf{e}_i\|_2\|\mathbf{x}\|_2 = 0$, and thus $(\mathbf{1} - \mathbf{e}_i) \in \mathcal{K}^*$.

(c) Consider an arbitrary $\mathbf{x} \notin \mathbb{R}^n_+$. Then there exists $i \in \{1, \ldots, n\}$ such that $x_i < 0$. From the proof in part (a) we have that $\mathbf{e}_i \in \mathcal{K}$ and we have $\langle \mathbf{e}_i, \mathbf{x}\rangle = x_i < 0$, which implies that $\mathbf{x} \notin \mathcal{K}^*$.

6. Consider three random variables X_1, X_2, X_3. Suppose that $\text{corr}(X_1, X_2) = 0.5$ and $\text{corr}(X_1, X_3) = -0.6$.
   (a) Formulate as a semidefinite optimisation problem, the problem of finding the minimum possible $\text{corr}(X_2, X_3)$. [1 point]
   (b) What is the dual problem to the problem from part (a)? [2 points]

Solution:

(a)

$$\begin{aligned} \min \quad & y_1 \\ \text{s.t.} \quad & \begin{pmatrix} 1 & 0.5 & -0.6 \\ 0.5 & 1 & y_1 \\ -0.6 & y_1 & 1 \end{pmatrix} \in \mathcal{PSD}^3. \end{aligned}$$

(b) The problem from part (a) is equivalent to

$$\begin{aligned} -\max \quad & -y_1 \\ \text{s.t.} \quad & \begin{pmatrix} 1 & 0.5 & -0.6 \\ 0.5 & 1 & 0 \\ -0.6 & 0 & 1 \end{pmatrix} - y_1 \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & -1 & 0 \end{pmatrix} \in \mathcal{PSD}^3. \end{aligned}$$

The dual to this is then

$$\begin{aligned} -\min \quad & \left\langle \begin{pmatrix} 1 & 0.5 & -0.6 \\ 0.5 & 1 & 0 \\ -0.6 & 0 & 1 \end{pmatrix}, X \right\rangle \\ \text{s.t.} \quad & \left\langle \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & -1 & 0 \end{pmatrix}, X \right\rangle = -1 \\ & X \in \mathcal{PSD}^3. \end{aligned}$$

This is equivalent to

$$\begin{aligned} \max \quad & \left\langle -\begin{pmatrix} 1 & 0.5 & -0.6 \\ 0.5 & 1 & 0 \\ -0.6 & 0 & 1 \end{pmatrix}, X \right\rangle \\ \text{s.t.} \quad & \left\langle \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}, X \right\rangle = 1 \\ & X \in \mathcal{PSD}^3. \end{aligned}$$

7. Consider the following optimisation problem:

$$\begin{aligned} \max \quad & 4x_1x_2 - x_2^2 - 9x_1 + 4x_2 \\ \text{s.t.} \quad & 4x_1^2 + x_2^2 - 8x_1 + 4x_2 + 4 = 0 \qquad \text{(A)} \\ & \mathbf{x} \in \mathbb{R}^2 \end{aligned}$$

(a) Give a finite lower bound to the optimal value of problem (A). [1 point]

(b) Formulate a positive semidefinite optimisation problem to give an upper bound on the optimal value of problem (A). [3 points]

Solution:

(a) To get a finite lower bound we need a feasible point. To narrow down the search for such a feasible point, try setting $x_2 = 0$. Then for $\mathbf{x}$ to be feasible we require $4x_1^2 - 8x_1 + 4 = 0$, or equivalently $x_1 = 1$. Therefore the point $(1, 0)$ is feasible, giving us a lower bound of $4*1*0 - 0^2 - 9*1 + 4*0 = -9$.

(b) Problem (A) is equivalent to

$$\begin{aligned} \max \quad & \left\langle \begin{pmatrix} 0 & 2 \\ 2 & -1 \end{pmatrix}, \mathbf{x}\mathbf{x}^\mathsf{T} \right\rangle - 9x_1 + 4x_2 \\ \text{s.t.} \quad & \left\langle \begin{pmatrix} 4 & 0 \\ 0 & 1 \end{pmatrix}, \mathbf{x}\mathbf{x}^\mathsf{T} \right\rangle - 8x_1 + 4x_2 + 4 = 0 \\ & \begin{pmatrix} 1 & \mathbf{x}^\mathsf{T} \\ \mathbf{x} & \mathbf{x}\mathbf{x}^\mathsf{T} \end{pmatrix} \in \mathcal{PSD}^3 \\ & \mathbf{x} \in \mathbb{R}^2. \end{aligned}$$

This can then be relaxed to

$$\begin{aligned} \max \quad & \left\langle \begin{pmatrix} 0 & 2 \\ 2 & -1 \end{pmatrix}, \mathsf{X} \right\rangle - 9x_1 + 4x_2 \\ \text{s.t.} \quad & \left\langle \begin{pmatrix} 4 & 0 \\ 0 & 1 \end{pmatrix}, \mathsf{X} \right\rangle - 8x_1 + 4x_2 + 4 = 0 \\ & \begin{pmatrix} 1 & \mathbf{x}^\mathsf{T} \\ \mathbf{x} & \mathsf{X} \end{pmatrix} \in \mathcal{PSD}^3 \\ & \mathbf{x} \in \mathbb{R}^2. \end{aligned}$$

8. (Automatic additional points) [4 points]

Question:	1	2	3	4	5	6	7	8	Total
Points:	6	3	10	3	7	3	4	4	40

A copy of the lecture-sheets may be used during the examination.
Good luck!