Exam: Continuous Optimisation 2015

3TU- and LNMB-course, Utrecht.

Monday 25th January 2016

1. Let $f : \mathcal{C} \to \mathbb{R}$, $\mathcal{C} \subset \mathbb{R}^n$ convex, be a convex function. Show that then the following holds: [3 points]
A local minimizer of f on $\mathcal{C}$ is a global minimizer on $\mathcal{C}$. And a strict local minimizer of f on $\mathcal{C}$ is a strict global minimizer on $\mathcal{C}$.

Solution: for a local minimizer $\overline{x}$: Suppose $\overline{x}$ is not a global one. Then with some $y \in \mathcal{C}$ we have $f(\overline{x}) > f(y)$. Thus for $0 < \lambda \leq 1$ we find with $x_\lambda := \overline{x} + \lambda(y - \overline{x})$ using convexity of f:

$$f(x_\lambda) \leq f(\overline{x}) + \lambda[f(y) - f(\overline{x})] < f(\overline{x})$$

So letting $\lambda \downarrow 0$, $\overline{x}$ cannot be a local minimizer.
for a strict local minimizer $\overline{x}$: Suppose it is not a strict global one. Then with some $y \in \mathcal{C}, \overline{x} \neq y$ we have $f(\overline{x}) \geq f(y)$. Thus for $0 < \lambda \leq 1$ we find with $x_\lambda := \overline{x} + \lambda(y - \overline{x})$ using convexity of f:

$$f(x_\lambda) \leq f(\overline{x}) + \lambda[f(y) - f(\overline{x})] \leq f(\overline{x})$$

So letting $\lambda \downarrow 0$, $\overline{x}$ cannot be a strict local minimizer.

2. (a) Show that for $d \in \mathbb{R}^n$ it holds: [2 points]

$$d^T x \geq 0 \ \forall x \in \mathbb{R}^n \quad \Leftrightarrow \quad d = 0.$$

(b) Let $c, a_i \in \mathbb{R}^n, i = 1, \ldots, m$ $(m \geq 1)$. Show using the Farkas Lemma (lecture sheets, Th. 3.12) that precisely one of the following alternatives (I) or (II) is true: [3 points]
(I): $c^T x < 0,\ a_i^T x \leq 0, i = 1, \ldots, m$ has a solution $x \in \mathbb{R}^n$.
(II): there exist $\mu_1 \geq 0, \ldots, \mu_m \geq 0$ such that: $c + \sum_{i=1}^m \mu_i a_i = 0$

Solution:

(a) "$\Rightarrow$":

$$d^T x \geq 0 \ \forall x \in \mathbb{R}^n \Rightarrow \pm d^T e_j \geq 0 \forall j \Rightarrow d^T e_j = 0 \ \forall j \Rightarrow d = 0$$

"$\Leftarrow$":

$$d = 0 \Rightarrow d^T x = 0 \ \forall x \in \mathbb{R}^n \Rightarrow d^T x \geq 0 \ \forall x \in \mathbb{R}^n$$

(b) For linear functions $f(x) := c^T x, g_i(x) := a_i^T x$ and $\mathcal{C} = \mathbb{R}^n$ under the Slater condition by Farkas Lemma precisely one of I or II is true:
(I): $c^T x < 0,\ \ a_i^T x \le 0, i = 1, \dots, m$ has a solution x.
(II): there exist $\mu_1 \ge 0, \dots, \mu_m \ge 0$ such that:

$$c^T x + \sum_{i=1}^m \mu_i a_i^T x = \left[c + \sum_{i=1}^m \mu_i a_i\right]^T x \ge 0 \quad \forall x \in \mathbb{R}^n.$$

In view of (a), (II) is equivalent with
(II): there exist $\mu_1 \ge 0, \dots, \mu_m \ge 0$ such that: $\quad c + \sum_{i=1}^m \mu_i a_i = 0.$
Note that the Slater condition is satisfied: In the linear case only feasibility is required and obviously $x = 0$ is feasible for $g_i(x) \le 0$.

3. Given is the problem

$$(P) \qquad \min_{x \in \mathbb{R}^2} \ (-2x_1 - x_2) \qquad \text{s.t.} \qquad x_1 \le 0, \ \text{and} \ -(x_1-1)^2 - (x_2-1)^2 + 2 \le 0\ .$$

(a) Is (P) a convex problem? Sketch the feasible set and the level set of f given by $f(x) = f(\overline{x})$ with $\overline{x} = 0$. Is LICQ (constraint qualification) satisfied at $\overline{x}$? [3 points]

(b) Show that the point $\overline{x} = 0$ is a KKT-point of (P). Determine the corresponding Lagrangean multipliers. [3 points]

(c) Show that $\overline{x}$ is a local minimizer. What is the order of this minimizer? Is it a global minimizer? [3 points]

(d) Consider now the program (objective f and constraint function g_2 interchanged): [2 points]

$$(\widetilde{P}) \qquad \min_{x \in \mathbb{R}^2} -(x_1-1)^2 - (x_2-1)^2 + 2 \qquad \text{s.t.} \qquad x_1 \le 0, \ \text{and} \ -2x_1 - x_2 \le 0\ .$$

Explain (without any further calculations) why $\overline{x} = 0$ is also a local minimizer of $(\widetilde{P})$.

Solution:

(a) (P) is not a convex program since g_2 is not convex: $\nabla^2 g_2(x) = \begin{pmatrix} -2 & 0 \\ 0 & -2 \end{pmatrix}$ is negative definite. LICQ holds at $\overline{x} = 0$:

$$\nabla g_1(x) = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad \nabla g_2(\overline{x}) = \begin{pmatrix} 2 \\ 2 \end{pmatrix} \quad \text{are linearly independent}$$

Give a complete sketch.

(b) The KKT condition for $\overline{x} = 0$ (g_1 and g_2 active) read:

$$\begin{pmatrix} -2 \\ -1 \end{pmatrix} + \mu_1 \begin{pmatrix} 1 \\ 0 \end{pmatrix} + \mu_2 \begin{pmatrix} 2 \\ 2 \end{pmatrix} = 0$$

With (unique) solution $\mu_1 = 1, \mu_2 = 1/2$.

(c) Since the assumptions of Th 5.4 are satisfied, $\overline{x} = 0$ is a local minimizer of order $p = 1$.
It is not a global minimizer since $f(\overline{x}) = 0$ and e.g. for feasible $x = (0, x_2),\ x_2 \geq 2$ we have $f(0, x_2) \to -\infty$ for $x_2 \to \infty$.

(d) The KKT condition at $\overline{x} = 0$ for (P) directly yields a corresponding KKT condition for $(\widetilde{P})$ at $\overline{x}$ (feasible for $(\widetilde{P})$!!) which again satisfies the assumption of Theorem 5.4 for $(\widetilde{P})$.

4. Consider the (nonlinear) program: [3 points]

$$(P) \qquad \min_x \ f(x) \qquad \text{s.t.} \quad x \in \ \mathcal{F} := \{x \in \mathbb{R}^n \mid g_j(x) \leq 0,\ j \in J\}$$

with $f, g_j \in C^1(\mathbb{R}^n, \mathbb{R}),\ J = \{1, \ldots, m\}$. Let d_k be a strictly feasible descent direction in $x_k \in \mathcal{F}$. Show that for $t > 0$, small enough, it holds:

$$f(x_k + td_k) < f(x_k) \quad \text{and} \quad x_k + td_k \in \mathcal{F}$$

Solution: By using Taylor around x_k we find for $j \in J_{x_k}$ (use $g_j(x_k)^T d_k < 0;\ g_j(x_k) = 0$):

$$g_j(x_k+td_k) = g_j(x_k)+t\nabla g_j(x_k)^T d_k+o(t) = t\nabla g_j(x_k)^T d_k+o(t) < 0 \quad \text{for } t > 0 \text{ small enough.}$$

By continuity also for $j \notin J_{x_k}$ we have $g_j(x_k + td_k) < 0$ for $t > 0$ small enough. So $x_k + td_k \in \mathcal{F}$. In view of $f(x_k)^T d_k < 0$ we also find

$$f(x_k + td_k) = f(x_k) + t\nabla f(x_k)^T d_k + o(t) < f(x_k) \quad \text{for } t > 0 \text{ small enough.}$$

5. For a given nonempty set $\mathcal{A} \subseteq \mathbb{R}^n$ we define its conic hull, conic($\mathcal{A}$) by

$$\text{conic}(\mathcal{A}) := \left\{ \sum_{i=1}^{m} \mu^i \mathbf{x}^i \;\middle|\; \mathbf{x}^i \in \mathcal{A},\ \mu^i \geq 0 \text{ for all } i,\ m \in \mathbb{N} \right\}.$$

(a) Show that conic($\mathcal{A}$) is a convex cone. [2 points]

(b) Show that if $\mathcal{A} \subseteq \mathcal{B} \subseteq \mathbb{R}^n$, with $\mathcal{B}$ being a convex cone, then conic($\mathcal{A}$) $\subseteq \mathcal{B}$. [3 points]

(c) Show that conic($\mathcal{A}$) is full dimensional if and only if there does not exist $\mathbf{y} \in \mathbb{R}^n \setminus \{\mathbf{0}\}$ such that $\langle \mathbf{y}, \mathbf{x} \rangle = 0$ for all $\mathbf{x} \in \mathcal{A}$. [1 point]

Solution:

(a) By Theorem 1.3, equivalently we want to show that for all $\mathbf{u}, \mathbf{v} \in \text{conic}(\mathcal{A})$ and $\lambda_1, \lambda_2 > 0$ we have $\lambda_1\mathbf{u} + \lambda_2\mathbf{v} \in \text{conic}(\mathcal{A})$.

Considering an arbitrary $\mathbf{u}, \mathbf{v} \in \text{conic}(\mathcal{A})$ and $\lambda_1, \lambda_2 > 0$ we have

$$\mathbf{u} = \sum_{i=1}^{m} \mu^i \mathbf{x}^i, \qquad \mathbf{v} = \sum_{i=1}^{p} \nu^i \mathbf{y}^i,$$

$$\begin{aligned} \text{for some } & \mathbf{x}^1, \dots, \mathbf{x}^m, \mathbf{y}^1, \dots, \mathbf{y}^p \in \mathcal{A}, \\ & \mu^1, \dots, \mu^m, \nu^1, \dots, \nu^p \geq 0, \\ & p, m \in \mathbb{N}. \end{aligned}$$

Therefore

$$\lambda_1\mathbf{u} + \lambda_2\mathbf{v} = \sum_{i=1}^{m} \underbrace{\lambda_1\mu^i}_{\geq 0} \mathbf{x}^i + \sum_{i=1}^{p} \underbrace{\lambda_2\nu^i}_{\geq 0} \mathbf{y}^i \in \text{conic}(\mathcal{A}).$$

(b) For $k \in \mathbb{N}$, let $\mathcal{L}^k := \left\{ \sum_{i=1}^{k} \mu^i \mathbf{x}^i \;\middle|\; \mathbf{x}^i \in \mathcal{A},\ \mu^i \geq 0 \text{ for all } i \right\}$. We will prove by induction that $\mathcal{L}^k \subseteq \mathcal{B}$ for all $k \in \mathbb{N}$, and thus $\mathcal{B} \supseteq \bigcup_{k\in\mathbb{N}} \mathcal{L}^k = \text{conic}(\mathcal{A})$.

We start by proving the case of $k = 1$. If $\mathbf{y} \in \mathcal{L}^1$ then $\mathbf{y} = \mu\mathbf{x}$ for some $\mu \geq 0$ and $\mathbf{x} \in \mathcal{A}$. We thus have $\mathbf{x} \in \mathcal{B}$, and as $\mathcal{B}$ is a cone we have $\mathbf{y} = \mu\mathbf{x} \in \mathcal{B}$.

We now suppose the statement is true for k, and show it is also true for $k+1$. If $\mathbf{y} \in \mathcal{L}^{k+1}$ then $\mathbf{y} = \sum_{i=1}^{k+1} \mu^i \mathbf{x}^i$ where $\mathbf{x}^i \in \mathcal{A}$ and $\mu^i \geq 0$ for all i. Letting $\mathbf{z}^1 = \sum_{i=1}^{k} 2\mu^i \mathbf{x}^i \in \mathcal{L}^k \subseteq \mathcal{B}$ and $\mathbf{z}^2 = 2\mu^{k+1}\mathbf{x}^{k+1} \in \mathcal{L}^1 \subseteq \mathcal{B}$, the set $\mathcal{B}$ being convex implies that $\mathcal{B} \ni \frac{1}{2}\mathbf{z}^1 + \frac{1}{2}\mathbf{z}^2 = \mathbf{y}$.

Alternatively:

$$\begin{aligned} \text{conic}(\mathcal{A}) &= \left\{ \sum_{i=1}^{m} \mu^i \mathbf{x}^i \;\middle|\; \mathbf{x}^i \in \mathcal{A},\ \mu^i \geq 0 \text{ for all } i,\ m \in \mathbb{N} \right\} \\ &= \{\mathbf{0}\} \cup \left\{ \sum_{i=1}^{m} \mu^i \mathbf{x}^i \;\middle|\; \mathbf{x}^i \in \mathcal{A},\ \mu^i \geq 0 \text{ for all } i,\ m \in \mathbb{N},\ \lambda = \sum_{i=1}^{m} \mu^i > 0 \right\} \\ &= \{\mathbf{0}\} \cup \left\{ \lambda \sum_{i=1}^{m} \theta^i \mathbf{x}^i \;\middle|\; \mathbf{x}^i \in \mathcal{A},\ \theta^i \geq 0 \text{ for all } i,\ m \in \mathbb{N},\ 1 = \sum_{i=1}^{m} \theta^i,\ \lambda > 0 \right\} \\ &= \{\mathbf{0}\} \cup \mathbb{R}_{++}\text{conv}(\mathcal{A}) = \mathbb{R}_{+}\text{conv}(\mathcal{A}). \end{aligned}$$

As $\mathcal{B}$ is convex, we have $\text{conv}(\mathcal{A}) \subseteq \mathcal{B}$. As $\mathcal{B}$ is a cone we then get

$$\mathcal{B} \supseteq \mathbb{R}_{+}\text{conv}(\mathcal{A}) = \text{conic}(\mathcal{A}).$$

(c) We will prove the equivalent statement that conic($\mathcal{A}$) is not full dimensional if and only if there exists $\mathbf{y} \in \mathbb{R}^n \setminus \{\mathbf{0}\}$ such that $\langle \mathbf{y}, \mathbf{x} \rangle = 0$ for all $\mathbf{x} \in \mathcal{A}$.

($\Rightarrow$) Suppose conic($\mathcal{A}$) is not full-dimensional. Then by definition 1.8.3 there exists $\mathbf{y} \in \mathbb{R}^n \setminus \{\mathbf{0}\}$ such that $\langle \mathbf{y}, \mathbf{x} \rangle = 0$ for all $\mathbf{x} \in \operatorname{conic} \mathcal{A}$. We trivially have $\mathcal{A} \subseteq \operatorname{conic}(\mathcal{A})$ and thus $\langle \mathbf{y}, \mathbf{x} \rangle = 0$ for all $\mathbf{x} \in \mathcal{A}$.

($\Leftarrow$) Suppose there exists $\mathbf{y} \in \mathbb{R}^n \setminus \{\mathbf{0}\}$ such that $\langle \mathbf{y}, \mathbf{x} \rangle = 0$ for all $\mathbf{x} \in \mathcal{A}$. Then for all $\mathbf{z} \in \operatorname{conic}(\mathcal{A})$ we have $\mathbf{z} = \sum_{i=1}^m \mu^i \mathbf{x}^i$ for some $\mathbf{x}^i \in \mathcal{A}$ and $\mu^i \geq 0$ for all i, $m \in \mathbb{N}$, and thus $\langle \mathbf{y}, \mathbf{z} \rangle = \sum_{i=1}^m \mu^i \langle \mathbf{y}, \mathbf{x}^i \rangle = 0$. Therefore, by definition 1.8.3, we have that conic($\mathcal{A}$) is not full-dimensional.

6. In this question we will consider the proper cone $\mathcal{K} \subseteq \mathbb{R}^{n+2}$ defined as

$$\mathcal{K} = \left\{ \begin{pmatrix} x \\ \mathbf{y} \\ z \end{pmatrix} \,\middle|\, \mathbf{y} \in \mathbb{R}^n,\ x, z \in \mathbb{R},\ \|\mathbf{y}\|_2 \leq x,\ z \geq 0 \right\}.$$

(a) Consider a ray $\mathcal{R} = \{\mathbf{c} - y_1 \mathbf{a} \mid y_1 \in \mathbb{R}_+\}$ with fixed $\mathbf{a}, \mathbf{c} \in \mathbb{R}^n$. We wish to find the distance between the origin and the closest point in this ray. Formulate this problem as a conic optimisation problem over $\mathcal{K}$. [2 points]

(b) Give an explicit characterisation of $\mathcal{K}^*$. [1 point]
[Justification for your answer must be provided]

(c) What is the dual problem to your formulation in part (a)? [2 points]
[If you were not able to answer parts (a) and (b) then instead find the dual to: $\min_y \quad y \quad \text{s.t.} \quad \mathbf{c} + y\mathbf{a} \in \mathbb{R}^n_+$.]

Solution:

(a) This problem is equivalent to the following problems

$$\min_{y_1} \quad \|\mathbf{c} - y_1 \mathbf{a}\|_2 \quad \text{s.t.} \quad y_1 \geq 0,$$

$$\begin{aligned} \min_{\mathbf{y}} \quad & y_2 \\ \text{s.t.} \quad & \|\mathbf{c} - y_1 \mathbf{a}\|_2 \leq y_2, \quad y_1 \geq 0, \end{aligned}$$

$$\begin{aligned} \min_{\mathbf{y}} \quad & y_2 \\ \text{s.t.} \quad & \begin{pmatrix} 0 \\ \mathbf{c} \\ 0 \end{pmatrix} - y_1 \begin{pmatrix} 0 \\ \mathbf{a} \\ -1 \end{pmatrix} - y_2 \begin{pmatrix} -1 \\ 0 \\ 0 \end{pmatrix} \in \mathcal{K} \end{aligned}$$

$$\begin{aligned} -\max_{\mathbf{y}} \quad & 0y_1 - y_2 \\ \text{s.t.} \quad & \begin{pmatrix} 0 \\ \mathbf{c} \\ 0 \end{pmatrix} - y_1 \begin{pmatrix} 0 \\ \mathbf{a} \\ -1 \end{pmatrix} - y_2 \begin{pmatrix} -1 \\ 0 \\ 0 \end{pmatrix} \in \mathcal{K} \end{aligned}$$

The correct answer is either of the last two formulations, or equivalent.

(b) We have that $\mathcal{K} = \mathcal{A}_0 \times \mathbb{R}_+$, and thus $\mathcal{K}^* = \mathcal{A}_0^* \times \mathbb{R}_+^* = \mathcal{A}_0 \times \mathbb{R}_+ = \mathcal{K}$.

(c) Considering

$$\begin{aligned} -\max_{\mathbf{y}} \quad & 0y_1 - y_2 \\ \text{s.t.} \quad & \begin{pmatrix} 0 \\ \mathbf{c} \\ 0 \end{pmatrix} - y_1 \begin{pmatrix} 0 \\ \mathbf{a} \\ -1 \end{pmatrix} - y_2 \begin{pmatrix} -1 \\ 0 \\ 0 \end{pmatrix} \in \mathcal{K} \end{aligned}$$

the dual problem is

$$\begin{aligned} -\min_{x,\mathbf{y},z} \quad & \left\langle \begin{pmatrix} 0 \\ \mathbf{c} \\ 0 \end{pmatrix}, \begin{pmatrix} x \\ \mathbf{y} \\ z \end{pmatrix} \right\rangle \\ \text{s.t.} \quad & \left\langle \begin{pmatrix} 0 \\ \mathbf{a} \\ -1 \end{pmatrix}, \begin{pmatrix} x \\ \mathbf{y} \\ z \end{pmatrix} \right\rangle = 0 \\ & \left\langle \begin{pmatrix} -1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} x \\ \mathbf{y} \\ z \end{pmatrix} \right\rangle = -1, \\ & \begin{pmatrix} x \\ \mathbf{y} \\ z \end{pmatrix} \in \mathcal{K}^* \end{aligned}$$

This can be simplified to

$$\begin{aligned} \max_{x,\mathbf{y},z} \quad & -\langle \mathbf{c}, \mathbf{y} \rangle \\ \text{s.t.} \quad & z = \langle \mathbf{a}, \mathbf{y} \rangle \\ & x = 1, \quad z \geq 0, \quad \|\mathbf{y}\|_2 \leq x \end{aligned}$$

which in turn is equivalent to

$$\max_{\mathbf{y}} \quad \langle -\mathbf{c}, \mathbf{y} \rangle \quad \text{s.t.} \quad \langle \mathbf{a}, \mathbf{y} \rangle \geq 0, \quad \|\mathbf{y}\|_2 \leq 1.$$

Alternative question:

The problem is equivalent to $-\max_y \quad -y \quad \text{s.t.} \quad \mathbf{c} - y(-\mathbf{a}) \in \mathbb{R}^n_+$.

The dual to this is $-\min_{\mathbf{x}} \quad \langle \mathbf{c}, \mathbf{x} \rangle \quad \text{s.t.} \quad \langle -\mathbf{a}, \mathbf{x} \rangle = -1, \quad \mathbf{x} \in \mathbb{R}^n_+$,

which is equivalent to $\max_{\mathbf{x}} \quad \langle -\mathbf{c}, \mathbf{x} \rangle \quad \text{s.t.} \quad \langle \mathbf{a}, \mathbf{x} \rangle = 1, \quad \mathbf{x} \in \mathbb{R}^n_+$

7. Consider the following optimisation problem: [3 points]

$$\begin{aligned} \min_{\mathbf{x}} \quad & 2x_2^2 + 5x_1x_2 - 4x_2 \\ \text{s.t.} \quad & 2x_1^2 + x_1 + 3x_2^2 - 2x_1x_2 = 3 \\ & \mathbf{x} \in \mathbb{R}^2. \end{aligned} \tag{A}$$

Give the standard positive semidefinite approximation for this problem, the solution of which would provide a lower bound to the optimal value of problem (A).

Solution:

$$\begin{aligned} \min_{\mathbf{x}} \quad & 2x_2^2 + 5x_1x_2 - 4x_2x_3 \\ \text{s.t.} \quad & 2x_1^2 + x_1x_3 + 3x_2^2 - 2x_1x_2 = 3 \\ & x_3^2 = 1 \\ & \mathbf{x} \in \mathbb{R}^2, \end{aligned}$$

$$\begin{aligned} \min_{\mathbf{x}} \quad & \left\langle \begin{pmatrix} 0 & 5/2 & 0 \\ 5/2 & 2 & -2 \\ 0 & -2 & 0 \end{pmatrix}, \mathbf{x}\mathbf{x}^\mathsf{T} \right\rangle \\ \text{s.t.} \quad & \left\langle \begin{pmatrix} 2 & -1 & 1/2 \\ -1 & 3 & 0 \\ 1/2 & 0 & 0 \end{pmatrix}, \mathbf{x}\mathbf{x}^\mathsf{T} \right\rangle = 3 \\ & \left\langle \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \mathbf{x}\mathbf{x}^\mathsf{T} \right\rangle = 1 \\ & \mathbf{x} \in \mathbb{R}^3, \end{aligned}$$

$$\begin{aligned} \min_{\mathbf{x}} \quad & \left\langle \begin{pmatrix} 0 & 5/2 & 0 \\ 5/2 & 2 & -2 \\ 0 & -2 & 0 \end{pmatrix}, X \right\rangle \\ \text{s.t.} \quad & \left\langle \begin{pmatrix} 2 & -1 & 1/2 \\ -1 & 3 & 0 \\ 1/2 & 0 & 0 \end{pmatrix}, X \right\rangle = 3 \\ & \left\langle \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}, X \right\rangle = 1 \\ & X \in \mathcal{PSD}^3. \end{aligned}$$

8. (Automatic additional points) [4 points]

Question:	1	2	3	4	5	6	7	8	Total
Points:	3	5	11	3	6	5	3	4	40

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