Exam: Continuous Optimisation 2015

3TU- and LNMB-course, Utrecht.

Monday 4th January 2016

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

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

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

Solution:

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

$$\begin{array}{rcl}
f(\lambda x + (1-\lambda)y) & = & \sum_{j=1}^{k} \alpha_j f_j(\lambda x + (1-\lambda)y) \\
f_j \textit{ convex, } \alpha_j \geq 0 \qquad & \leq & \sum_{j=1}^{k} \alpha_j [\lambda f_j(x) + (1-\lambda) f_j(y)] \\
& = & \lambda [\sum_{j=1}^{k} \alpha_j f_j(x)] + (1-\lambda)[\sum_{j=1}^{k} \alpha_j f_j(y)] \\
& = & \lambda f(x) + (1-\lambda) f(y)
\end{array}$$

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

$$\begin{array}{rcl}
g(\lambda x + (1-\lambda)y) & = & \max_{1 \leq j \leq k} \{f_j(\lambda x + (1-\lambda)y)\} \\
& \leq & \max_{1 \leq j \leq k} \{\lambda f_j(x) + (1-\lambda) f_j(y)\} \\
& \leq & \lambda [\max_{1 \leq j \leq k} f_j(x)] + (1-\lambda) \max_{1 \leq j \leq k} \{f_j(y)\} \\
& = & \lambda g(x) + (1-\lambda) g(y)
\end{array}$$

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

2. Consider the convex program [3 points]

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

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

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

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

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

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

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

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

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

3. Consider the two problems

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

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

(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 $f(x) = f(0,0)$. [3 points]

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

(c) Show for both problems that $\overline{x}$ 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, $\nabla^2 g_2(x) = \begin{pmatrix} -2 & 0 \\ 0 & 0 \end{pmatrix}$ is not positive semidefinite. So (P_1) is convex, (P_2) is not. (Give two complete sketches).

(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 $\overline{x} = 0$ must be a global minimizer (see Th. 3.7).
Since (P_2) is not convex we have to check the second order conditions (in Th. 5.9) (or we can directly argue as below): we compute

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

and thus

$$d^T \nabla^2 L(\overline{x}, \mu_2) = d^T \Big(\begin{pmatrix} 4 & 0 \\ 0 & 0 \end{pmatrix} + \mu_2 \begin{pmatrix} -2 & 0 \\ 0 & 0 \end{pmatrix} \Big) d = 2d_1^2 > 0$$

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

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

4. Consider the auxiliary program of the SQP-method (for solving a nonlinear program (P)) with some $x_k \in \mathbb{R}^n$: [3 points]

$$(Q_k) \quad \min_{d\in\mathbb{R}^n} \quad \nabla f(x_k)^T d + \frac{1}{2} d^T L_k d \qquad \text{s.t.} \quad \nabla g_j(x_k)^T d + g_j(x_k) \le 0 \quad \forall j \in J$$

Assume x_k is feasible for (P), i.e., $g_j(x_k) \le 0 \forall j \in J$, and L_k is positive definite. Show that if $d_k \neq 0$ is a solution of (Q_k) then d_k is a descent direction for f, i.e., $\nabla f(x_k)^T d < 0$.

Solution: Since x_k is feasible for (P), i.e., $g_j(x_k) \le 0, \forall j$, obviously $\overline{d} = 0$ is feasible for (Q_k). Since d_k is a global minimizer of (Q_k) (why global?; (Q_k) is convex!) we must have:

$$\nabla f(x_k)^T d_k + \frac{1}{2} d_k^T L_k d_k \le \nabla f(x_k)^T \overline{d} + \frac{1}{2} \overline{d}^T L_k \overline{d} = 0$$

Positive definiteness of L_k implies for $d_k \neq 0$: $\nabla f(x_k)^T d_k \le -\frac{1}{2} d_k^T L_k d_k < 0$.

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 $i = 1, \ldots, n$, 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 $\operatorname{corr}(X_1, X_2) = 0.5$ and $\operatorname{corr}(X_1, X_3) = -0.6$.
   (a) Formulate as a semidefinite optimisation problem, the problem of finding the minimum possible $\operatorname{corr}(X_2, X_3)$. [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:

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

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

(b) Give the standard completely positive approximation for this problem, the solution of which would provide an upper bound to the optimal value of problem (A). [3 points]

Solution:

(a) To get a finite upper 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 an upper bound of $4*1*0 - 0^2 - 9*1 + 4*0 = -9$.

(b) Problem (A) is equivalent to

$$\begin{aligned} \max \quad & 4x_1x_2 - x_2^2 - 9x_1x_3 + 4x_2x_3 \\ \text{s.t.} \quad & 4x_1^2 + x_2^2 - 8x_1x_3 + 4x_2x_3 = -4 \\ & x_3^2 = 1 \qquad \mathbf{x} \in \mathbb{R}^3_+. \end{aligned} \tag{1}$$

This is in turn equivalent to

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

This can then be relaxed to

$$
\begin{aligned}
\max \quad & \left\langle \begin{pmatrix} 0 & 2 & -9/2 \\ 2 & -1 & 2 \\ -9/2 & 2 & 0 \end{pmatrix}, X \right\rangle \\
\text{s.t.} \quad & \left\langle \begin{pmatrix} 4 & 0 & -4 \\ 0 & 1 & 2 \\ -4 & 2 & 0 \end{pmatrix}, X \right\rangle = -4 \\
& \left\langle \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}, X \right\rangle = 1 \\
& X \in \mathcal{CP}^3.
\end{aligned} \tag{3}
$$

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!