Test exam: Continuous Optimisation 2015

3TU- and LNMB-course, Utrecht.

Monday 4^{th} December 2015

1. Let $f : \mathbb{R}^m \to \mathbb{R}$ be a convex function $f(y)$ on $\mathbb{R}^m$ and let $A \in \mathbb{R}^{m \times n}$, $b \in \mathbb{R}^m$ be given.

   (a) Show that the function $g(x) := f(Ax + b)$ is a convex function of x on $\mathbb{R}^n$. [3 points]

   (b) Suppose that f is strictly convex. Show that then $g(x) := f(Ax+b)$ is strictly convex if and only if A has (full) rank n. [4 points]

   Hint: Recall that f is strictly convex if for any $y_1 \neq y_2$, $0 < \lambda < 1$ it holds: $f(\lambda y_1 + (1-\lambda)y_2) < \lambda f(y_1) + (1-\lambda) f(y_2)$.

Solution:

(a) For $x_1, x_2 \in \mathbb{R}^n$, $\lambda \in [0,1]$ we find:

$$\begin{aligned} g(\lambda x_1 + (1-\lambda)x_2) &= f(A(\lambda x_1 + (1-\lambda)x_2) + b) \\ &= f(\lambda A x_1 + (1-\lambda)Ax_2 + \lambda b + (1-\lambda)b) \\ &= f(\lambda(Ax_1 + b) + (1-\lambda)(Ax_2 + b)) \\ f \text{ is convex} \quad &\leq \lambda f(Ax_1 + b) + (1-\lambda) f(Ax_2 + b) \\ &= \lambda g(x_1) + (1-\lambda) g(x_2) \end{aligned}$$

(b) "$\Leftarrow$" $\operatorname{rank}(A) = n$ implies: $x_1 \neq x_2 \quad \Rightarrow \quad Ax_1 \neq Ax_2$.
As in (a) for $x_1 \neq x_2$, $\lambda \in (0,1)$ we obtain:

$$\begin{aligned} g(\lambda x_1 + (1-\lambda)x_2) &= f(\lambda(Ax_1 + b) + (1-\lambda)(Ax_2 + b)) \\ &\qquad \text{"}f \text{ is strict convex, } Ax_1 + b \neq Ax_2 + b\text{"} \\ &< \lambda f(Ax_1 + b) + (1-\lambda) f(Ax_2 + b) \\ &= \lambda g(x_1) + (1-\lambda) g(x_2) \end{aligned}$$

"$\Rightarrow$" Assume $\operatorname{rank}(A) < n$. Then there exist $x_1 \neq x_2$ with $Ax_1 = Ax_2$ and for any $\lambda \in (0,1)$ we obtain:

$$\begin{aligned} g(\lambda x_1 + (1-\lambda)x_2) &= f(\lambda(Ax_1 + b) + (1-\lambda)(Ax_2 + b)) = f(Ax_1 + b) \\ \text{"}g(x_1) = g(x_2)\text{"} \quad &= g(x_1) = \lambda g(x_1) + (1-\lambda) g(x_2). \end{aligned}$$

So g is not strictly convex.

2. For given $S \subset \mathbb{R}^n$ we define the convex hull conv(S) by

$$\text{conv}(S) = \left\{ x = \sum_{i=1}^{m} \lambda_i x_i \;\middle|\; \sum_{i=1}^{m} \lambda_i = 1;\; x_i \in S, \lambda_i \geq 0 \;\forall i;\; m \in \mathbb{N} \right\}$$

Show that conv(S) is the smallest convex set containing S:

(a) Show that the set conv(S) is convex with $S \subset \text{conv}(S)$. [3 points]

(b) Show that for any convex set C containing S we must have $\text{conv}(S) \subset C$. [3 points]
(*Hint: You may use without proof any Lemma, Theorem etc. from the course*)

Solution:

(a) Take $x^1, x^2 \in \text{conv}(S)$, $\lambda \in [0,1]$ (with $x^j = \sum_{i=1}^{m_j} \lambda_i^j x_i^j$, $x_i^j \in S$, $\sum_{i=1}^{m_j} \lambda_i^j = 1$, $\lambda_i^j \geq 0$ for $j = 1, 2$). Then we find:

$$\lambda x^1 + (1-\lambda)x^2 = \sum_{i=1}^{m_1} \lambda\lambda_i^1 x_i^1 + \sum_{i=1}^{m_2} (1-\lambda)\lambda_i^2 x_i^2 \in \text{conv}(S)$$

since $\sum_{i=1}^{m_1} \lambda\lambda_i^1 + \sum_{i=1}^{m_2} (1-\lambda)\lambda_i^2 = 1$ and "coefficients are ≥ 0". Note that (trivially) $S \subset \text{conv}(S)$ holds.

(b) Let $S \subset C$ with convex C: Take any $x \in \text{conv}(S)$, i.e., $x = \sum_{i=1}^{m} \lambda_i x_i$ with $\lambda_i \geq 0, \sum_{i=1}^{m} \lambda_i = 1$ and $x_i \in S$ and thus $x_i \in C$. Since C is convex by Lem.2.5 (Jensen inequality) the convex combination x of points $x_i \in C$ is in C. So $\text{conv}(S) \subset C$.

3. Consider with $0 \neq c \in \mathbb{R}^n$ the program:

$$(P) \qquad \min_{x \in \mathbb{R}^n} c^T x \quad \text{s.t.} \quad x^T x \leq 1 \,.$$

(a) Show that $\overline{x} = -\frac{c}{\|c\|}$ is the minimizer of (P) with minimum value $v(P) = -\|c\|$. [2 points]
($\|x\|$ means here the Euclidian norm.)

(b) Compute the solution $\overline{y}$ of the Lagrangean dual (D) of (P). Show in this way that for the optimal values strong duality holds, i.e., $v(D) = v(P)$. [4 points]

Solution:

(a) Either show this "by a sketch". Or as follows (using Schwarz inequality):
$\|x\| \leq 1$ implies: $c^T x \geq -\|c\|\|x\| \geq -\|c\|$, and " $c^T x = -\|c\|$" holds iff $x = -\frac{c}{\|c\|}$

So $\overline{x} = -\frac{c}{\|c\|}$ is the minimizer with $v(P) = c^T(-\frac{c}{\|c\|}) = -\|c\|$.
(Alternatively find $\overline{x}$ by solving the KKT-conditions.)

(b) The dual (D) is given by

$$(D) \quad \max_{y \geq 0} \ \psi(y) \quad \text{where } \psi(y) := \min_{x \in \mathbb{R}^n} L(x,y)$$

with Lagrangean function $L(x,y) = c^T x + y(x^T x - 1)$.
We find for $y = 0$: $\psi(0) = -\infty$.
for $y > 0$: The minimizer x of $\psi(y)$ satisfies $\nabla_x L(x,y) = c + 2yx = 0$ or $x = -\frac{1}{2y}c$. So (fill in)

$$\psi(y) = -\frac{1}{2y}c^T c + \frac{1}{4y}c^T c - y = -\frac{1}{4y}c^T c - y \ .$$

To find an (unconstrained) maximizer of $\psi(y)$ for $y > 0$ we solve

$$\psi'(y) = \frac{1}{4y^2}c^T c - 1 = 0 \quad \text{with solution} \quad \overline{y} = \frac{1}{2}\|c\| \ .$$

So $v(D) = \psi(\overline{y}) = -\|c\| = v(P)$.

4. Consider the problem (in connection with the design of a cylindrical can with height h, radius r and volume at least 2π such that the total surface area is minimal):

$$(P): \quad \min \ f(h,r) := 2\pi(r^2 + rh) \quad \text{s.t.} \ -\pi r^2 h \leq -2\pi, \quad (\text{and } h > 0, r > 0)$$

(a) Compute a (the) solution $(\overline{h}, \overline{r})$ of the KKT conditions of (P). Show that (P) is not a convex optimization problem. [4 points]

(b) Show that the solution $(\overline{h}, \overline{r})$ in (a) is a local minimizer. Why is it the unique global solution? [3 points]
Hint: Use the sufficient optimality conditions

Solution:

(a) We first note that the functions $f(h,r) = 2\pi(r^2 + rh)$ and $g(h,r) := -\pi r^2 h + 2\pi$ are not convex (for $h > 0$). For the objective function f, e.g., this follows by:

$$\nabla f = 2\pi \begin{pmatrix} r \\ 2r + h \end{pmatrix}, \quad \nabla^2 f = 2\pi \begin{pmatrix} 0 & 1 \\ 1 & 2 \end{pmatrix} \quad \text{and thus:} \quad \det \nabla^2 f < 0$$

We now consider the KKT condition: $(\nabla f = -\mu \nabla g, g \leq 0, \mu \cdot g = 0)$
So consider: $2\pi\binom{r}{2r+h} = \mu\pi\binom{r^2}{2rh}$ $(\star)$:
Case $\mu = 0$: leads to $2\pi\binom{r}{2r+h} = 0$ with solution $(h,r) = (0,0)$ which is not feasible.

Case $\mu > 0$ and thus $\pi r^2 h = 2\pi$:
The 2 equations in $(\star)$ lead to $\mu = 2/r$ and then $2(2r+h) = \frac{2}{r}2rh$ or $h = 2r$. By using the (active) constraint we find $\pi r^2 h = 2\pi r^3 = 2\pi$ with solution $r = 1$. So the unique KKT solution is given by $(\overline{h}, \overline{r}) = (2, 1), \overline{\mu} = 2$.

(b) (We apply the second order sufficient conditions of Th. 5.9 to the nonconvex program (P)). So we will show (for the cone of critical directions $C(\overline{h}, \overline{r})$):

$$d^T \nabla^2_{h,r} L(\overline{h}, \overline{r}, \overline{\mu}) d > 0 \quad \forall d \in C(\overline{h}, \overline{r}) \setminus \{0\} \quad (\star\star)$$

We compute

$$\nabla f(\overline{h}, \overline{r}) = 2\pi \begin{pmatrix} 1 \\ 4 \end{pmatrix}, \qquad \nabla g(\overline{h}, \overline{r}) = -\pi \begin{pmatrix} 1 \\ 4 \end{pmatrix},$$
$$\nabla^2 L(\overline{h}, \overline{r}, \overline{\mu}) = 2\pi \begin{pmatrix} 0 & 1 \\ 1 & 2 \end{pmatrix} + 2(-\pi) \begin{pmatrix} 0 & 2 \\ 2 & 4 \end{pmatrix} = -2\pi \begin{pmatrix} 0 & 1 \\ 1 & 2 \end{pmatrix}$$

and

$$\begin{aligned} C(\overline{h}, \overline{r}) &= \{d \in \mathbb{R}^2 \mid \nabla f(\overline{h}, \overline{r})^T d \leq 0, \nabla g(\overline{h}, \overline{r})^T d \leq 0\} \\ &= \{d \in \mathbb{R}^2 \mid \begin{pmatrix} 1 \\ 4 \end{pmatrix}^T d \leq 0, \ -\begin{pmatrix} 1 \\ 4 \end{pmatrix}^T d \leq 0\} \\ &= \{\lambda \begin{pmatrix} -4 \\ 1 \end{pmatrix} \mid \lambda \in \mathbb{R}\} \end{aligned}$$

For $d = \lambda(-4, 1)^T \neq 0$, (i.e., $\lambda \neq 0$) we obtain (see $(\star\star)$):

$$\lambda(-4, 1)(-2\pi) \begin{pmatrix} 0 & 1 \\ 1 & 2 \end{pmatrix} \lambda \begin{pmatrix} -4 \\ 1 \end{pmatrix} = \ldots = 2\lambda^2 \pi 6 > 0 \quad \forall \lambda \neq 0 \ .$$

So $(\overline{h}, \overline{r}) = (2, 1)$ is a local minimizer.
It is the unique (global) minimizer since the point is the only KKT point. Note that since the linear independency constraint qualification holds ($\nabla g = -\pi \binom{r^2}{2rh} \neq 0$, for $r, h > 0$) any local minimizer must be a KKT point. Also note that for feasible $\|(h, r)\| \to \infty$ also $f \to \infty$ holds. (*To show the latter fact is technically "involved" and was not expected to be done.*)

5. Consider the closed set

$$\mathcal{K} = \{\mathbf{x} \in \mathbb{R}^2 \mid x_1 + 2x_2 \geq 0 \text{ and } 3x_1 + x_2 \geq 0\}$$

(a) Prove that $\mathcal{K}$ is a proper cone. [You may assume closure.] [5 points]

(b) Find the dual cone to $\mathcal{K}$. [1 point]

Solution:

(a) In order for a set to be a proper cone it most be a closed, convex, pointed full-dimensional cone. We will assume closure and prove the rest:

- Convex cone: Consider an arbitrary $\mathbf{x}, \mathbf{y} \in \mathcal{K}$ and $\lambda_1, \lambda_2 > 0$. From Theorem 1.3 of the conic optimisation part of the course, if we can show that $\lambda_1\mathbf{x} + \lambda_2\mathbf{y} \in \mathcal{K}$ then we are done.
  We have

  $$\begin{aligned} x_1 + 2x_2 &\geq 0, & 3x_1 + x_2 &\geq 0, & \lambda_1 &> 0, \\ y_1 + 2y_2 &\geq 0, & 3y_1 + y_2 &\geq 0, & \lambda_2 &> 0. \end{aligned}$$

  This implies that

  $$\begin{aligned} (\lambda_1\mathbf{x} + \lambda_2\mathbf{y})_1 + 2(\lambda_1\mathbf{x} + \lambda_2\mathbf{y})_2 &= \lambda_1(x_1 + 2x_2) + \lambda_2(y_1 + 2y_2) \geq 0, \\ 3(\lambda_1\mathbf{x} + \lambda_2\mathbf{y})_1 + (\lambda_1\mathbf{x} + \lambda_2\mathbf{y})_2 &= \lambda_1(3x_1 + x_2) + \lambda_2(3y_1 + y_2) \geq 0. \end{aligned}$$

  Therefore $\lambda_1\mathbf{x} + \lambda_2\mathbf{y} \in \mathcal{K}$.

- Full-dimensional: Using Definition 1.8, part 2 of the conic optimisation part of the course, this follows from the space being two dimensional and having two linearly independent vectors $\begin{pmatrix}1\\0\end{pmatrix}, \begin{pmatrix}0\\1\end{pmatrix} \in \mathcal{K}$.

- Pointed: We will consider an arbitrary $\mathbf{x} \in \mathbb{R}^2$ such that $\pm\mathbf{x} \in \mathcal{K}$. Using Definition 1.7 of the conic optimisation part of the course, if we can then show that $\mathbf{x} = \mathbf{0}$ then we are done. We have

  $$\left.\begin{aligned} (\mathbf{x})_1 + 2(\mathbf{x})_2 \geq 0 \\ (-\mathbf{x})_1 + 2(-\mathbf{x})_2 \geq 0 \end{aligned}\right\} \quad\Rightarrow\quad x_1 + 2x_2 = 0,$$

  $$\left.\begin{aligned} 3(\mathbf{x})_1 + (\mathbf{x})_2 \geq 0 \\ 3(-\mathbf{x})_1 + (-\mathbf{x})_2 \geq 0 \end{aligned}\right\} \quad\Rightarrow\quad 3x_1 + x_2 = 0.$$

  Therefore

  $$x_1 = \tfrac{2}{5}\underbrace{(3x_1 + x_2)}_{=0} - \tfrac{1}{5}\underbrace{(x_1 + 2x_2)}_{=0} = 0, \qquad x_2 = \underbrace{(3x_1 + x_2)}_{=0} - 3\underbrace{x_1}_{=0} = 0.$$

(b) From Corollary 2.8 of the conic optimisation part of the course and the note on slide 10/31 of the first lecture in the conic optimisation part of the course we have that

$$\mathcal{K}^* = \mathrm{cl}\ \mathrm{conic}\left\{\begin{pmatrix}1\\2\end{pmatrix},\ \begin{pmatrix}3\\1\end{pmatrix}\right\} = \mathrm{conic}\left\{\begin{pmatrix}1\\2\end{pmatrix},\ \begin{pmatrix}3\\1\end{pmatrix}\right\}.$$

6. We will consider bounds to the optimal value of the following problem:

$$\begin{aligned}\min_{\mathbf{x}} \quad & 5x_1^2 - 4x_1x_2 - 2x_1 + x_2^2 + 2 \\ \text{s.t.} \quad & x_1^2 + 5x_2^2 - 4x_1x_2 - 8x_2 = 4 \\ & \mathbf{x} \in \mathbb{R}^2.\end{aligned} \tag{A}$$

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

(b) Formulate a positive semidefinite optimisation problem to give a lower bound on the optimal value of problem (A). [2 points]

(c) Give the dual problem to the positive semidefinite optimisation problem you formulated in part (b) of this question. [1 point]

Solution:

(a) To find an upper bound we can use any feasible point, $\widehat{\mathbf{x}}$. If we limit our search for a feasible point such that $\widehat{x}_2 = 0$ then we would have a feasible point if and only if $4 = \widehat{x}_1^2 + 5 * 0^2 - 4\widehat{x}_1 * 0 - 8 * 0 = \widehat{x}_1^2$. Therefore both $\widehat{\mathbf{x}} = (2, 0)$ and $\widehat{\mathbf{x}} = (-2, 0)$ are feasible points. We only need one point to give us an upper bound, and if we consider the feasible point $\widehat{\mathbf{x}} = (2, 0)$ then this gives us the upper bound of

$$\begin{aligned}5\widehat{x}_1^2 - 4\widehat{x}_1\widehat{x}_2 - 2\widehat{x}_1 + \widehat{x}_2^2 + 2 &= 5 * 2^2 - 4 * 2 * 0 - 2 * 2 + 0^2 + 2 \\ &= 20 - 0 - 4 + 0 + 2 \\ &= 18\end{aligned}$$

(b) Problem (A) is equivalent to

$$\begin{aligned}\min_{\mathbf{x}} \quad & 5x_1^2 - 4x_1x_2 - 2x_1x_3 + x_2^2 + 2x_3^2 \\ \text{s.t.} \quad & x_1^2 + 5x_2^2 - 4x_1x_2 - 8x_2x_3 = 4 \\ & x_3^2 = 1, \qquad \mathbf{x} \in \mathbb{R}^3,\end{aligned}$$

which is in turn equivalent to

$$\begin{aligned}
\min_{\mathbf{x},X} \quad & \left\langle \begin{pmatrix} 5 & -2 & -1 \\ -2 & 1 & 0 \\ -1 & 0 & 2 \end{pmatrix}, X \right\rangle \\
\text{s.t.} \quad & \left\langle \begin{pmatrix} 1 & -2 & 0 \\ -2 & 5 & -4 \\ 0 & -4 & 0 \end{pmatrix}, X \right\rangle = 4 \\
& \left\langle \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}, X \right\rangle = 1, \\
& X = \mathbf{x}\mathbf{x}^T, \qquad \mathbf{x} \in \mathbb{R}^3,
\end{aligned}$$

A lower bound on this is then provided by solving the optimisation problem

$$\begin{aligned}
\min_{\mathbf{x},X} \quad & \left\langle \begin{pmatrix} 5 & -2 & -1 \\ -2 & 1 & 0 \\ -1 & 0 & 2 \end{pmatrix}, X \right\rangle \\
\text{s.t.} \quad & \left\langle \begin{pmatrix} 1 & -2 & 0 \\ -2 & 5 & -4 \\ 0 & -4 & 0 \end{pmatrix}, X \right\rangle = 4 \\
& \left\langle \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}, X \right\rangle = 1, \\
& X \in \mathcal{PSD}^3.
\end{aligned}$$

(c) Considering slide 9/20 of lecture 3 of the conic optimisation part of this course we have that the dual problem is

$$\begin{aligned}
\max_{\mathbf{y}} \quad & 4y_1 + y_2 \\
\text{s.t.} \quad & \begin{pmatrix} 5 & -2 & -1 \\ -2 & 1 & 0 \\ -1 & 0 & 2 \end{pmatrix} - y_1 \begin{pmatrix} 1 & -2 & 0 \\ -2 & 5 & -4 \\ 0 & -4 & 0 \end{pmatrix} - y_2 \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \in \mathcal{PSD}^3.
\end{aligned}$$

7. (Automatic additional points) [4 points]

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

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