LUND UNIVERSITY

Department of
AUTOMATIC CONTROL

Exam in Optimization for Learning

2019-10-28

Points and grading

All answers must include a clear motivation. Answers should be given in English. The total number of points is 20. The maximum number of points is specified for each subproblem. Preliminary grading scales:

Grade 3: 12 points on the exam
4: 17 points on exam plus extra-credit handin
5: 22 points on exam plus extra-credit handin

Accepted aid

Authorized *Cheat Sheet.*

Results

Solutions will be posted on the course webpage, and results will be registered in LADOK. Date and location for display of corrected exams will be posted on the course webpage.

1. Determine if the following sets are convex or not:

a. $S_1 = \{(x, y) \in \mathbb{R}^2 : x^2 = y\}$. (1 p)

b. $S_2 = \{x \in \mathbb{R}^n : \max_{i=1,\ldots,n} x_i \le r\}$, where $r \in \mathbb{R}$. (1 p)

c. $S_3 = \{(x, t) \in \mathbb{R}^2 : |x|^2 \le t^2\}$. (1 p)

d. $S_4 = \mathrm{epi}(\exp)$ where exp is the exponential function. (1 p)

e. $S_5 = \{x \in \mathbb{R}^n : Ax \ge b\}$ where $A \in \mathbb{R}^{m \times n}$ and $b \in \mathbb{R}^m$. (1 p)

In each subproblem, you are allowed to assume that norms are convex and that the function $h_p : \mathbb{R} \to \mathbb{R}$ such that

$$h_p(z) = \max(0, z)^p \tag{1}$$

for each $z \in \mathbb{R}$ is convex and nondecreasing, for any $p \ge 1$.

2. Determine whether or not the functions below are convex.

a. $f_1 : \mathbb{R} \to \mathbb{R}$ such that

$$f(x) = \begin{cases} x & \text{if } x > 0, \\ -1 & \text{if } x \le 0. \end{cases}$$

(1 p)

b. $f_2 : \mathbb{R} \to \mathbb{R} \cup \{\infty\}$ such that $f_2 = g^\star$, where $g : \mathbb{R} \to \mathbb{R}$ is such that

$$g(x) = a_4 x^4 + a_3 x^3 + a_2 x^2 + a_1 x^1 + a_0$$

for each $x \in \mathbb{R}$, where $a_0, a_1, a_2, a_3 \in \mathbb{R}$. (1 p)

c. $f_3 : \mathbb{R} \to \mathbb{R} \cup \{\infty\}$ such that

$$f_3(x) = \begin{cases} -\min(\log(x), -e^{-x}) & \text{if } x > 0, \\ \infty & \text{if } x \le 0. \end{cases}$$

(1 p)

d. $f_4 : \mathbb{R} \to \mathbb{R}$ such that

$$f_4(x) = |x|^3$$

for each $x \in \mathbb{R}$. (1 p)

e. $f_5 : \mathbb{R}^n \to \mathbb{R} \cup \{\infty\}$ such that

$$f(x) = \begin{cases} \|x\|_2^2 & \text{if } Ax = b, \\ \infty & \text{if } Ax \ne b, \end{cases}$$

for each $x \in \mathbb{R}^n$, where $A \in \mathbb{R}^{m \times n}$ and $b \in \mathbb{R}^m$. (1 p)

3. Consider the proper, closed and convex function $f : \mathbb{R} \to \mathbb{R}$ such that

$$f(x) = \begin{cases} \frac{1}{2}x^2 & \text{if } |x| \leq \frac{1}{2}, \\ \frac{1}{2}(|x| - \frac{1}{4}) & \text{if } |x| > \frac{1}{2}, \end{cases}$$

for each $x \in \mathbb{R}$, known as the *Huber loss*. Let $\gamma > 0$. Compute the proximal operator $\operatorname{prox}_{\gamma f}$. (1 p)

4. Consider the proper, closed and convex function $f : \mathbb{R}^n \to \mathbb{R}$ such that

$$f(x) = \begin{cases} \|x\|_2 & \text{if } \|x\|_2 \leq 1, \\ \frac{1}{2}\left(\|x\|_2^2 + 1\right) & \text{if } \|x\|_2 > 1 \end{cases}$$

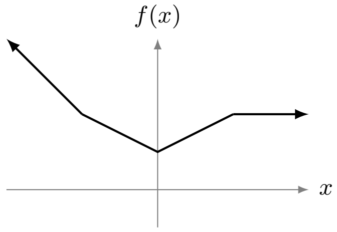
for each $x \in \mathbb{R}$, known as the *reversed Huber* function. Compute ∂f. (1 p)

5. A proper and convex function $f : \mathbb{R} \to \mathbb{R} \cup \{\infty\}$ has the following properties: $f(-2) = 3$, $\partial f(-1) = \{-1\}$, $\partial f(0) = [-1, 0]$. What can you conclude about the following properties?

a. Smoothness. (1 p)

b. Strong convexity. (1 p)

6. Sketch the conjugate of the piecewise linear function showed below. Outside the plotted domain, assume the graph continues in the same direction as on the boundary. (1 p)



7. Let $\gamma > 0$, $b \in \mathbb{R}^n$ and $f : \mathbb{R}^n \to \mathbb{R} \cup \{\infty\}$.

a. Find $(\gamma f)^*$ expressed in terms of f^* and γ. (1 p)

b. Find $(f(\cdot - b))^*$ expressed in terms f^* and b. (1 p)

8. Suppose that $f : \mathbb{R}^n \to \mathbb{R}$ is convex and differentiable, and that $g_i : \mathbb{R} \to \mathbb{R} \cup \{\infty\}$ is proper, closed and convex for each $i = 1, \ldots, n$. Consider the coordinate proximal-gradient method:

1. Pick an initial point $x^0 \in \mathbb{R}^n$ and step-lengths $\gamma_i > 0$ for each $i = 1, \ldots, n$
2. For $k = 0, 1, 2, \ldots$
   (a) Choose an index $i \in \{1, \ldots, n\}$ (according to some schedule)
   (b) Update $x_i^{k+1} = \mathrm{prox}_{\gamma_i g_i}\left(x_i^k - \gamma_i(\nabla f(x^k))_i\right)$
   (c) Update $x_j^{k+1} = x_j^k$ for each $j = 1, \ldots, n$ such that $j \neq i$

Now, suppose that $x^\star \in \mathbb{R}^n$ is a fixed point of the coordinate proximal-gradient method, i.e., if $x^k = x^\star$, then $x^{k+1} = x^\star$, regardless of which coordinate i was chosen. Show that $x^\star$ solves

$$\underset{x=(x_1,\ldots,x_n)\in\mathbb{R}^n}{\text{minimize}} \; f(x) + \sum_{i=1}^n g_i(x_i).$$

(1 p)

9. Consider the primal problem

$$\underset{x\in\mathbb{R}^n}{\text{minimize}} \; \|x\|_1 + \left|\mathbf{1}^T x\right|$$

and the dual problem

$$\underset{\mu\in\mathbb{R}}{\text{minimize}} \; \iota_{[-\mathbf{1},\mathbf{1}]}(-\mathbf{1}\mu) + \iota_{[-1,1]}(\mu)$$

where $\mathbf{1} \in \mathbb{R}^n$ is a vector of all ones. Here we used that

$$(\|\cdot\|_1)^\star = \iota_{[-\mathbf{1},\mathbf{1}]}$$

and

$$(|\cdot|)^\star = \iota_{[-1,1]}.$$

Suppose that $\mu^\star \in \mathbb{R}$ is a solution to the dual problem. Recover the primal solution $x^\star \in \mathbb{R}^n$ using $\mu^\star$. (2 p)