Advanced Topics in Control 2020: Large-Scale Convex Optimization

Solution to Exercise 2

Goran Banjac, Mathias Hudoba de Badyn, Andrea Iannelli,
Angeliki Kamoutsi, Ilnura Usmanova

March 20, 2020

---

1 Convex Functions

(a) We will show that the Hessian of f_1 is positive semidefinite at all $x \in \mathbb{R}^n$.
Let $\beta(x) = e^{x_1} + \ldots + e^{x_n}$. Then a straightforward calculation yields

$$z^\top \nabla^2 f_1(x) z = \frac{1}{\beta(x)^2} \sum_{i=1}^{n} \sum_{j=1}^{n} e^{x_i + x_j} (z_i - z_j)^2 \geq 0, \ \ \forall z \in \mathbb{R}^n.$$

Hence by Proposition on Slide 12 of Lecture 3, f_1 is convex.

(b) The function $f_2(x) = \|x\|^p$ can be viewed as a composition $h(g(x))$ of the scalar function

$$h(t) = \begin{cases} t^p, & \text{if } t \geq 0 \\ 0, & \text{otherwise,} \end{cases}$$

with $p \geq 1$, and the function $g(x) = \|x\|$. In this case h is convex and monotically increasing, while g is convex over $\mathbb{R}^n$ (since any vector norm is convex, by the triangle inequality and positive homogeneity). Using the composition rules on Slide 17, it follows that f_2 is convex over $\mathbb{R}^n$.

(c) The fuction $f_3(x)$ can be viewed as a composition $h(g(x))$ of the function

$$h(t) = \begin{cases} \frac{1}{t}, & \text{if } t > 0 \\ +\infty, & \text{otherwise,} \end{cases}$$

and the function $g(x) = f(x)$, for $x \in \mathbb{R}^n$. In this case, h is convex and monotonically decreasing over $\mathbb{R}$, while g is concave over $\mathbb{R}^n$. Therefore, by the composition rules of Slide 17, it follows that f_3 is convex over $\mathbb{R}^n$.

(d) As we have already shown in 1b), the map $x \to \|x\|_2^2$ is convex. Threfore, $x \to \|Ax - b\|_2^2$ is convex as it is a composition with the affine mapping $x \to Ax - b$. Moreover, $x \to \|x\|_1$ is convex since it is a norm, Finally, sum of convex functions is convex.

(e) For each $i = 1, \ldots, n$, the map $x \to 1 + x_i$ is convex (since it is affine), thus $x \to \max(0, 1 + x_i)$ is convex as max of convex functions. The sum over i is convex, as (finite) sum of convex functions. The second term $x \to \|x\|_2^2$ is convex by 1b). Finally, the sum f_5 is convex.

(f) For a fixed $y \in \mathbb{R}^n$, we consider the affine map $f_y(x) := x^\top y - g(y)$. Then we have $f_6(x) = \sup_{y\in\mathbb{R}^n} f_y(x)$ which is convex as the supremum of (uncountably many) convex functions.

(g) Note that
$$d_C^2(x) = \min_{y\in C} \|x-y\|_2^2 = \|x\|_2^2 - \max_{y\in C}\{2x^\top y - \|y\|_2^2\},$$
where we used that $\|x-y\|_2^2 = (x-y)^\top(x-y)$ and $\min_{y\in C}\{-2y^\top x+\|y\|_2^2\} = -\max_{y\in C}\{2x^\top y- \|y\|_2^2\}$. Hence,
$$\phi_C(x) = \max_{y\in C}\left[y^\top x - \frac{1}{2}\|y\|_2^2\right].$$
Therefore, since ϕ_C is a maximization of affine - and hence convex - functions, it is necessarily convex.

(h) Note that since h_1 and h_2 have a common affine minorant, we ensure that $g(x) > -\infty$, for all $x \in \mathbb{R}^n$. Indeed, for arbitrary $x \in \mathbb{R}^n$ and x_1, x_2 such that $x_1 + x_2 = x$, we have by assumption that $h_1(x_1) + h_2(x_2) \geq s^\top x - 2r > -\infty$. So, $g(x) = \inf\{h_1(x_1) + h_2(x_2) : x_1 + x_2 = x\} \geq s^\top x - 2r > -\infty$.

Moreover, g is proper. Indeed, choose specific $x_1 \in \operatorname{dom} h_1$ and $x_2 \in \operatorname{dom} h_2$. Then $h_1(x_1) + h_2(x_2) < \infty$. Thus $x := x_1 + x_2 \in \operatorname{dom} g$.

Next, define $f : \mathbb{R}^n\times\mathbb{R}^n \to (-\infty,\infty]$ by $f(x, y) := h_1(x)+h_2(x-y)$. The convexity of h_1 and h_2 implies that f is convex. Indeed, for all $\theta \in [0, 1]$ and for all $(x_1, y_1), (x_2, y_2) \in \mathbb{R}^n \times \mathbb{R}^n$, we have

$$\begin{aligned}
f\left(\theta(x_1, y_1) + (1-\theta)(x_2, y_2)\right) &= h_1\left(\theta x_1 + (1-\theta)x_2\right) + h_2\left(\theta(x_1 - y_1) + (1-\theta)(x_2 - y_2)\right)\\
&\leq \theta h_1(x_1) + (1-\theta h_1(x_2)) + \theta h_2(x_1 - y_1) + (1-\theta)h_2(x_2 - y_2)\\
&= \theta f(x_1, y_1) + (1-\theta)f(x_2, y_2),
\end{aligned}$$

where in the first and third line we used the definition of f, while in the second line we used the convexity of h_1 and h_2.

All in all, we have that $f(x, y)$ is jointly convex in (x, y) and the partial minimization over the second variable $g(x) = \min_{y\in\mathbb{R}^n} f(x, y) > -\infty$, for all $x \in \mathbb{R}^n$. Hence, by the partial minimization rule on Slide 16, we conclude that g is convex.

2 Strictly Convex Functions and Unique Minimizers

(a) Assume for the sake of contradiction that there exists $x \neq x^\star$ that satisfies $f(x) = f(x^\star)$. Then, since f is strictly convex we get
$$f(\frac{1}{2}x + \frac{1}{2}x^\star) < \frac{1}{2}(f(x) + f(x^\star)) = f(x^\star),$$
which is a contradiction. Hence, at most one minimizer exists.

(b) The function $f(x) = \frac{1}{x}$ with domain $\{x : x > 0\}$ is strictly convex with infimum 0. But no x exists that attains the infimum.

3 Subdifferentials

(a) Assume first that $s = (s_1, \ldots, s_n) \in \partial g(x)$. Then
$$\sum_{i=1}^n g_i(y_i) = g(y) \geq g(x) + s^\top(y-x) = \sum_{i=1}^n (g_i(x_i) + s_i(x_i - y_i)), \tag{1}$$

for all $y = (y_1, \ldots, y_n) \in \mathbb{R}^n$.

For fixed $j = 1, \ldots, n$, set $y_i = x_i$, for all $i \neq j$ and let the j-th component y_j be a free scalar. Then, (1) gives

$$g_j(y_j) \geq g_j(x_j) + s_j(y_j - x_j),$$

for all $y_j \in \mathbb{R}$. Equivalently, $s_j \in \partial g_j(x_j)$. Since this holds for arbitrary $j = 1, \ldots, n$ the proof of one direction is complete.

For the other direction let $s = (s_1, \ldots, s_n)$ and suppose that $s_i \in \partial g_i(x_i)$, for each $i = 1, \ldots, n$. Then,

$$g_i(y_i) \geq g_i(x_i) + s_i(y_i - x_i),$$

for all $y_i \in \mathbb{R}$. Summing over i

$$g(y) \geq g(x) + \sum_{i=1}^{n} s_i(y_i - x_i) = g(x) + s^\top (y - x),$$

for all $y = (y_1, \ldots, y_n) \in \mathbb{R}^n$ and thus $s = (s_1, \ldots, s_n) \in \partial g(x)$.

(b) A vector $s \in \partial \mathcal{I}_C(x)$ if and only if

$$\mathcal{I}_C(y) \geq \mathcal{I}_C(x) + s^\top (y - x), \tag{2}$$

for all $y \in \mathbb{R}^n$. Assume first that $x \in C$, then (2) is equivalent to $\mathcal{I}_C(y) \geq s^\top (y - x)$, for all y. This is equivalent to $s^\top (y - x) \leq 0$, for all $y \in C$. Equivalently $s \in N_C(x)$. Thus, $\partial \mathcal{I}_C(x) = N_C(x)$, for $x \in C$.

Assume next that $x \notin C$. For $y \in C$, relation (2) gives $0 \geq \infty + s^\top (y - x)$. No such s exists and $\partial \mathcal{I}_C(x) = \emptyset$.

4 Conjugate Functions

(a) Let $f(x) = \|x\|_2$. We then have that $f^*(s) = \sup_{x \in \mathbb{R}^n} (s^\top x - \|x\|_2)$. We will compute the conjugate f^* via the following steps.

(1) We have that $f^*(s) \geq 0$, for all s. Indeed by selecting $x = 0$, we get $f^*(s) \geq s^\top 0 - \|0\|_2 = 0$.

(2) Next, we show that $f^*(s) \leq 0$, for all s with $\|s\|_2 \leq 1$. Indeed, by the Cauchy-Schwarz inequality, we have

$$f^*(s) = \sup_x (s^\top x - \|x\|_2) \leq \sup_x (\|s\|_2 \|x\|_2 - \|x\|_2) = \sup((\|s\|_2 - 1)\|x\|_2).$$

Hence, if $\|s\|_2 \leq 1$, then $f^*(s) \leq 0$.

(3) Moreover, we show that $f^*(s) = \infty$, for all s with $\|s\|_2 > 1$. Set $x = ts$ with $t \geq 0$ to get $f^*(s) = \sup(s^\top x - \|x\|) \geq t\|s\|_2(\|s\|_2 - 1)$. Whenever, $\|s\|_2 > 1$, we let $t \to \infty$ to conclude that $f^*(s) = \infty$.

(4) Finally, we combine the above results to derive

$$f^*(s) = \begin{cases} 0, & \text{if } \|s\|_2 \leq 1 \\ +\infty, & \text{otherwise,} \end{cases}$$

(b) We use the formula $\partial f(x) = \arg\max_s (s^\top x - f^*(s)) = \arg\max_{\|s\|_2 \le 1}(s^\top x)$. For $x = 0$ the objective is 0 and all feasible points are maximizers. Thus, $\partial f(0) = \{s : \|s\|_2 \le 1\}$.

For $x \neq 0$, the maximum is attained when s and x are aligned, i.e., when $s = \frac{x}{\|x\|_2}$.

Therefore,

$$\partial f(x) = \begin{cases} B_{\|\cdot\|_2}(0,1), & \text{if } x = 0 \\ \{\frac{x}{\|x\|_2}\}, & \text{otherwise.} \end{cases}$$