Convex Optimization: Homework 3

Martín Prado

September 10, 2024
Universidad de los Andes – Bogotá Colombia

Exercise 1

Exercise 3, Chapter 4: Güler. Let K_1 and K_2 be convex cones in a vector space E. Show that $K_1 + K_2 \subseteq \text{co}(K_1 \cup K_2)$, and if both cones contain the origin, then $K_1 + K_2 = \text{co}(K_1 \cup K_2)$.

Solution: We're going to transcribe some definitions,

Definition 4.17. A set K in a vector space E is called a cone if $tx \in K$ whenever $t > 0$ and $x \in K$. If K is also a convex set, then it is called a convex cone.

Definition ??? The sum of two sets is defined as follows,

$$K_1 + K_2 = \{x_1 + x_2 \ : \ x_1 \in K_1, \ x_2 \in K_2\}.$$

Definition 4.10. Let $A \subseteq E$ be a nonempty set. The convex hull of A is the set of all convex combinations of points from A, that is,

$$\text{co}(A) = \left\{ \sum_{i=1}^{k} \lambda_i x_i \ : \ x_i \in A, \ \sum_{i=1}^{k} \lambda_i = 1, \ \lambda_i \geq 0, \ k \geq 1 \right\}.$$

From the definitions is clear that since K_1 and K_2 are convex cones, $2x_1 \in K_1$ and $2x_2 \in K_2$. Also, by taking $\lambda_i = \frac{1}{2}$ and $x_i = 2k_i$,

$$k_1 + k_2 = \sum_{i=1}^{2} \frac{1}{2} \cdot 2k_i = \sum_{i=1}^{2} \lambda_i x_i, \qquad x_i \in K_1 \cup K_2, \ \sum_{i=1}^{2} \lambda_i = 1, \ \lambda_i \geq 0.$$

Therefore, it's easy to see that $K_1 + K_2 \subseteq \text{co}(K_1 \cup K_2)$.

On the other hand, for the other inclusion, we use the following theorem:

Theorem 4.11. Let $A \neq \emptyset$ be a subset of an affine space E. Then $\text{co}(A)$ is a convex set; in fact, $\text{co}(A)$ is the smallest convex set containing A.

. $\square$

it's clear that $K_1 \cup K_2 \subset K_1 + K_2$. Now let $\lambda \in [0,1]$ and note that for $(1-\lambda)x + \lambda y$ we have two possible scenarios:

- If both x, y are in the same set K_i, $i = 1, 2$, then $(1-\lambda)x + \lambda y \in K_i \subseteq K_1 + K_2$ because K_1 and K_2 are convex sets.

- If $x \in K_1$ and $y \in K_2$, then $(1-\lambda)x \in K_1$ and $\lambda y \in K_2$ if $\lambda \in (0,1)$. For the case when $\lambda = 0$ or $\lambda = 1$, $(1-\lambda)x = 0$ or $\lambda y = 0$. In this case we use the hypothesis $0 \in K_1 \cap K_2$ to conclude that $(1-\lambda)x + \lambda y \in K_1 + K_2$ for $\lambda \in [0,1]$.

This proves that $K_1 + K_2$ is a convex set that contains $K_1 \cup K_2$. Therefore, using Theorem 4.11 we conclude that $\mathrm{co}(K_1 \cup K_2) \subseteq K_1 + K_2$.

Exercise 2*

Exercise 10, Chapter 2: Boyd. *Solution set of a quadratic inequality.* Let $C \subseteq \mathbb{R}^n$ be the solution set of a quadratic inequality,

$$C = \{x \in \mathbb{R}^n \mid x^T Ax + b^T x + c \leq 0\},$$

with $A \in \mathbb{S}^n, b \in \mathbb{R}^n$, and $c \in \mathbb{R}$.

(a) Show that C is convex if A is semidefinite positive

Exercise 3

Exercise 12, Chapter 2: Boyd. Which of the following sets are convex?

(a) A *slab*, i.e., a set of the form $\{x \in \mathbb{R}^n \mid \alpha \leq a^T x \leq \beta\}$

(b) A *rectangle*, i.e., a set of the form $\{x \in \mathbb{R}^n \mid \alpha_i \leq x_i \leq b_i,\ i = 1, \ldots, n\}$. A rectangle is sometimes called a *hyperrectangle* when $n > 2$.

(c) A *wedge*, i.e., $\{x \in \mathbb{R}^n \mid a_1^T x \leq b_1,\ a_2^T x \leq b_2\}$.

(d) The set of points closer to a given point than a given set, i.e.,

$$\{x \mid \|x - x_0\|_2 \leq \|x - y\|_2 \text{ for all } y \in S\}$$

where $S \subset \mathbb{R}^n$.

(e) The set of points closer to one set than another, i.e.,

$$\{x \mid \operatorname{dist}(x, S) \leq \operatorname{dist}(x, T)\},$$

where $S, T \subseteq \mathbb{R}^n$, and

$$\operatorname{dist}(x, S) = \inf\{\|x - z\|_2 \mid z \in S\}$$

(f) The set $\{x \mid x + S_2 \subseteq S_1\}$, where $S_1, S_2 \subseteq \mathbb{R}^n$ with S_1 convex.

(g) The set of points whose distance to a does not exceed a fixed fraction θ of the distance to b, i.e., the set $\{x \mid \|x - a\|_2 \leq \theta\|x - b\|_2\}$. You can assume $a \neq b$ and $0 \leq \theta \leq 1$.

Solution Item (a)

It is convex

For $x, y \in \{x \in \mathbb{R}^n \mid \alpha \leq a^T x \leq \beta\}$, we have that

$$\alpha \leq \begin{matrix} a^T x \\ a^T y \end{matrix} \leq \beta.$$

Then, for every $t \in [0, 1]$

$$\alpha = (1-t)\alpha + t\alpha \leq (1-t)x + ty \leq (1-t)\beta + t\beta \leq \beta.$$

So for every x, y in the slab, the line $(1-t)x + ty$ is also in the set.

Solution Item (b)

It is convex

Let $x, y \in \{x \in \mathbb{R}^n \mid \alpha_i \leq x_i \leq b_i,\ i = 1, \ldots, n\}$, and define for $t \in [0, 1]$, $z(t) = (1-t)x + ty$ and $z_i(t) = (z(t))_i = (1-t)x_i + ty_i$. Similar to the previous item, for every $i = 1, \ldots, n$:

$$a_i = (1-t)a_i + ta_i \leq \underbrace{(1-t)x_i + ty_i}_{z_i(t)} \leq (1-t)b_i + tb_i = b_i$$

Therefore, $z(t)$ is in the rectangle for every $t \in [0, 1]$

Solution Item (c)

It is convex

Let $x, y \in \{x \in \mathbb{R}^n \mid a_1^T x \leq b_1,\ a_2^T x \leq b_2\}$. We apply the same argument as the previous two items

$$a_i^T((1-t)x + ty) = (1-t)a_i^T x + ta_i^T y \leq (1-t)b_i + tb_i = b_i.$$

Therefore, $(1-t)x + ty$ is also in the wedge for every $t \in [0, 1]$.

Solution Item (d)

It is convex

Since convexity is preserved by the intersection between convex sets, it follows that

$$\{x \mid \|x - x_0\|_2 \leq \|x - y\|_2 \ \forall y \in S\} = \bigcap_{y \in S} \underbrace{\{x \mid \|x - x_0\|_2 \leq \|x - y\|_2\}}_{=C_y}$$

Now, we want to prove that C_y is convex for every $y \in S$. Then, let $R_y = \|y - x_0\|$. Using **Example 2.12** we know that every closed ball is convex and $C_y = \overline{B_{R_y}(x_0)}$. Therefore, the intersection of every ball it's convex.

Solution Item (e)

It is NOT convex

Take the sets $S = \{-2, 2\}$ and $T = \{0\}$ in $\mathbb{R}$. Then, $\{x \mid \operatorname{dist}(x, S) \leq \operatorname{dist}(x, T)\} = (-\infty, -1] \cup [1, \infty)$ which is not connected and thus, it is not convex.

Solution Item (f)

It is convex

Let $x, y \in \{x \mid x + S_2 \subseteq S_1\}$. Let $z \in S_2$ and note that $x + z \in S_1$ and $y + z \in S_1$ by definition. Since S_1 is convex it follows that $(1-t)(x+z) + t(y+z) \in S_1$ for $t \in [0, 1]$. Therefore,

$$\begin{aligned}(1-t)x + ty + z = (1-t)(x+z) + t(y+z) \in S_1, \ \forall z \in S_2 \\ \implies (1-t)x + ty + S_2 \subseteq S_1.\end{aligned}$$

Solution Item (g)

It is convex

Let $x, y \in \{x \mid \|x - a\| \leq \theta \|x - b\|\}$. Then, for $t \in [0, 1]$. Then, using parallelogram rule

$$\begin{aligned}\|x - a\|^2 - \theta^2 \|x - b\|^2 &= \langle x, x \rangle + 2 \langle x, a \rangle + \langle a, a \rangle - \theta^2 \langle x, x \rangle + 2\theta^2 \langle x, b \rangle + \theta^2 \langle b, b \rangle \\ &= (1 - \theta^2) \langle x, x \rangle + 2 \left\langle x, a - \theta^2 b \right\rangle + \langle a, a \rangle - \langle \theta b, \theta b \rangle \leq 0.\end{aligned}$$

This quadratic inequation describes a convex set using exercise 2, because $(1-\theta^2)\langle x, x\rangle = x^T(1-\theta^2)Ix$, and since, $0 \leq \theta^2 \leq 1$, it follows that the matrix $(1-\theta^2)I$ is semidefinite positive.

Exercise 4

Let $f : \mathbb{R} \to \mathbb{R}$. Show that f is not decreasing if and only if there exists φ differentiable and convex such that $f = \varphi'$.

Solution:

$\Longleftarrow$:

We use the following theorem for this implication.

Theorem 4.27. Let C be a convex set in $\mathbb{R}^n$, and let f be a Gâteaux differentiable function on an open set containing C.

Then f is convex on C if and only if the tangent plane at any point $x \in C$ lies below the graph of f, that is,

$$f(y) \geq f(x) + \langle \nabla f(x), y - x\rangle \text{ for all } x, y \in C.$$

. $\square$

If φ is a convex differentiable function, then we have that for every $x < y \in \mathbb{R}$

$$\varphi(x) + \varphi'(x)(y - x) \leq \varphi(y)$$

$$\varphi(y) + \varphi'(y)(x - y) \leq \varphi(x)$$

By adding both inequalities we obtain

$$\varphi(x) + \varphi(y) + (\varphi'(x) - \varphi'(y))(y - x) \leq \varphi(x) + \varphi(y)$$

$$\iff \qquad (\varphi'(x) - \varphi'(y))(y - x) \leq 0.$$

Since $y - x > 0$, we must have that

$$f(x) = \varphi'(x) \leq \varphi'(y) = f(y), \text{ for every } x, y \in \mathbb{R}.$$

$\Longrightarrow$:

I'm going to assume that f is continuous so we can actually define a primitive. Let φ be the primitive of f and define for fixed $x < y \in \mathbb{R}$, the function $\gamma : [0,1] \to \mathbb{R}$ as follows

$$\gamma(t) = \varphi((1-t)x + ty).$$

Note that $\gamma(0) = x$ and $\gamma(1) = y$. Then, we have that

$$\frac{d\gamma}{dt}(t) = \gamma'(t) = (y-x) \cdot f((1-t)x + ty),$$

and by monotonicity of f we have that γ' is non decreasing too. Finally, we have that

$$\begin{aligned}
\varphi(y) &= \gamma(1) - \gamma(0) + \gamma(0) \\
&= \int_0^1 \gamma'(r)dr + \gamma(0) \\
&\geq \int_0^1 \gamma'(0)dr + \gamma(0) \\
&= \gamma'(0) + \gamma(0) \\
&= (y-x)\varphi'(x) + \varphi(x).
\end{aligned}$$

Since x, y were arbitrary, we can use theorem 4.27 to conclude that φ is a convex function.

Exercise 5

Exercise 11, Chapter 4: Güler. Let $f : C \to \mathbb{R}$ be a twice Fréchet differentiable function on a convex open set $C \subseteq \mathbb{R}^n$. The following statements are known to be equivalent:

(a) f is convex.

(b) $f(y) \geq f(x) + \langle \nabla f(x), y - x \rangle$ for all $x, y \in C$.

(c) $\langle \nabla f(y) - \nabla f(x), y - x \rangle \geq 0$ for all $x, y \in C$.

(d) $Hf(x)$ is positive semidefinite at every $x \in C$.

In fact, we have proved that (a), (b), and (d) are equivalent conditions. Give direct proof of

Solution:

(b) $\implies$ (c):

By hypothesis we have for any $x, y \in C$

$$f(y) \geq f(x) + \langle \nabla f(x), y - x \rangle$$
$$f(x) \geq f(y) + \langle \nabla f(y), x - y \rangle$$

By summing both inequalities we obtain

$$\begin{aligned} & f(x) + f(y) \geq f(x) + f(y) + \langle \nabla f(x) - \nabla f(y), y - x \rangle \\ \iff \qquad & 0 \geq \langle \nabla f(x) - \nabla f(y), y - x \rangle . \end{aligned}$$

(c) $\implies$ (b): **Hint:** First, show that the function $g(t) := f(x + t(y - x)) - t \langle \nabla f(x), y - x \rangle$ is differentiable and nondecreasing.

So let $g(t) = f(x + t(y - x)) - t \langle \nabla f(x), y - x \rangle$, and note that g is a composition and sum of differentiable functions. So using chain rule,

$$\begin{aligned} g'(t) &= \langle \nabla f(x + t(y - x)), y - x \rangle - \langle \nabla f(x), y - x \rangle \\ &= \langle \nabla f(x + t(y - x)) - \nabla f(x), y - x \rangle . \end{aligned}$$

From the hypothesis, we know that

$$\begin{aligned} 0 &\leq \langle \nabla f(x + t(y - x)) - \nabla f(x), \;\; x + t(y - x) - x \rangle \\ &= \langle \nabla f(x + t(y - x)) - \nabla f(x), \;\; t(y - x) \rangle \\ &= t g'(t) \end{aligned}$$

Since $t > 0$, it follows that $g'(t) \geq 0$, and thus, g is non decreasing. Therefore,

$$\begin{aligned} f(y) - \langle \nabla f(x), y - x \rangle = g(1) \geq g(0) = f(x) \\ \implies f(y) \geq f(x) + \langle \nabla f(x), y - x \rangle . \end{aligned}$$

(c) $\implies$ (d):

Let $y = x + td$ for a fixed d. Then, by hypothesis

$$\begin{aligned} f(x) + t \langle \nabla f(x), d \rangle + \frac{t^2}{2} \langle Hf(x)(d), d \rangle + O(t^3) &= f(x + td) \\ &= f(y) \\ &\geq f(x) + \langle \nabla f(x), y - x \rangle \\ &= f(x) + t \langle \nabla f(x), d \rangle . \end{aligned}$$

Therefore,

$$\frac{t^2}{2} \langle Hf(x)(d), d \rangle + O(t^3) \geq 0$$

For a sufficiently small t, the absolute value of $\frac{t^2}{2}\langle Hf(x)(d), d\rangle$ is greater that the error $O(t^3)$. Therefore, we can conclude that for every $d \in \mathbb{R}^n$

$$\langle Hf(x)(d), d\rangle \geq 0.$$

(d) $\implies$ (c):

Using mean value theorem, there exists $z_1, z_2 \in C$ such that

$$f(y) = f(x) + \langle \nabla f(x), y - x\rangle + \frac{1}{2}\langle Hf(z_1)(y-x), y-x\rangle \geq f(x) + \langle \nabla f(x), y - x\rangle$$

$$f(x) = f(y) + \langle \nabla f(y), x - y\rangle + \frac{1}{2}\langle Hf(z_2)(x-y), x-y\rangle \geq f(x) - \langle \nabla f(y), y - x\rangle$$

Therefore, by summing both inequalities we obtain

$$f(x) + f(y) \geq f(x) + f(y) + \langle \nabla f(x) - \nabla f(y), y - x\rangle$$
$$\implies \langle \nabla f(y) - \nabla f(x), y - x\rangle \geq 0.$$

Exercise 6*

Exercise 14, Chapter 4: Güler. Let $C \subseteq \mathbb{R}^n$ be a convex set. Show that the following functions are convex:

(a) The indicator function of C defined by

$$\delta_C(x) := \begin{cases} 0, & \text{if } x \in C, \\ +\infty, & \text{otherwise.} \end{cases}$$

(b) The distance function to C defined by

$$d_C(x) := \inf\{\|z - x\| \ : \ z \in C\}.$$

(c) The support function of C defined by

$$\sigma_C(x) := \sup\{\langle z, w\rangle \ : \ z \in C\}.$$

Hint: Consider the epigraph of the function. Furthermore,

(d) Show that $d_C = \|\cdot\| \Box \delta_C$, where $\|\cdot\|(x) = \|x\|$ is the norm function on $\mathbb{R}^n$.

Exercise 7

Consider the function $f(x) = \langle c, x\rangle - \sum_{j=1}^{m} \log(1 - \langle a_j, x\rangle) - \sum_{i=1}^{n} \log(1 - x_i^2)$. Show that f is convex.

Solution:

The sum of convex functions is convex. We must prove that for $f = f_1 + f_2 + f_3$, f_i is convex for $i = 1, 2, 3$. We start with $f_1(x) = \langle c, x\rangle$, which is linear, and thus, convex.

Then,

$$f_2(x) = \sum_{j=1}^{m} -\log(1 - \langle a_j, x\rangle) = \sum_{j=1}^{m} f_{2,j}(x),$$

where $f_{2,j}(x) = -\log(1 - \langle a_j, x\rangle)$. Note that

$$\frac{d^2[-\log(1-y)]}{dy^2} = \frac{1}{(1-y)^2} \geq 0, \; y \in (0, 1).$$

Therefore, $\gamma(y) = -\log(1-y)$ is a convex function and since $f_{2,j} = \gamma \circ \langle a_j, \cdot\rangle$ is a composition of two convex functions, it follows that $f_{2,j}$ is convex function for every $j = 1, \ldots, m$.

Finally, with a similar argument,

$$f_3(x) = \sum_{i=1}^{n} -\log(1 - x_i^2) = \sum_{i=1}^{n} f_{3,i}(x).$$

Then, $f_{3,i} = \gamma \circ (\cdot)^2 \circ \pi_i$. Note that π_i is a projection, and since projections are linear functions, it follows that π_i is a convex function. The composition of these three functions is a convex function too.

In conclusion $f = f_1 + \sum_{j=1}^{m} f_{2,j}(x) + \sum_{i=1}^{n} f_{3,i}(x)$ is a sum of convex functions, so f is convex too.