Convex Optimization for Machine Learning and Computer Vision

Lecture: Dr. Tao Wu
Exercises: Zhenzhang Ye
Winter Semester 2019/20

Computer Vision Group
Institut für Informatik
Technische Universität München

Weekly Exercises 5

Room: 02.09.023

Wednesday, 27.11.2019, 12:15-14:00

Submission deadline: Monday, 25.11.2019, 16:15, Room 02.09.023

Convex conjugate (8+6 Points)

Exercise 1 (4 points). Given a $X \in \mathbb{R}^{m \times n}$, compute the subdifferential of the 1, 2-norm, i.e.

$$\partial ||X||_{1,2} = \partial (\sum_{i=1}^{m} (\sum_{j=1}^{n} X_{i,j}^2)^{1/2}).$$

Solution. We rewrite the original problem into $J(X) = \sum_{i=1}^{m} (\|X_i\|_2)$, where X_i is the i-th row of X. We can apply sum rule on it and get

$$\partial J(X) = \sum_{i=1}^{m} \partial \|X_i\|_2 .$$

Now the problem is given a vector x, compute the subdifferential of its ℓ_2 norm. We consider this problem elementwisely and if $x \neq 0$, we can easily get $\partial \|x\|_2 = \frac{x}{\|x\|}$. Now consider when $x = 0$. Recall the definition of subdifferential:

$$\partial \|0\|_2 = \{p \in \mathbb{E} : \|y\|_2 \geq \langle p, y \rangle, \forall y \in \mathbb{E}\}.$$

If $\|p\| \leq 1$, we have $\|y\|_2 \geq \|p\|_2 \|y\|_2 \geq \langle p, y \rangle$. Therefore, such p is in $\partial \|0\|_2$. Otherwise, for $\|p\| > 1$, choose $y = p$, the inequality doesn't hold any more. Denote $B_1(0) := \{p \in \mathbb{R}^n : \|p\|_2 \leq 1\}$. To summary, we have following equation:

$$\partial \|x\|_2 = \begin{cases} \frac{x}{\|x\|_2}, & x \neq 0 \\ B_1(0), & x = 0 \end{cases}$$

Substitute this back into original problem and write into matrix format, we can get

$$\partial J(X) = \{P \in \mathbb{R}^{m \times n},\ P_i \in \partial \|X_i\|_2\}.$$

Exercise 2 (4 Points). Consider following problems of convex conjugate:

- Let $f : \mathbb{R}^n \to \mathbb{R}$ be convex. Show that the convex conjugate of the perspective function $g : \mathbb{R}^n \times \mathbb{R} \to \mathbb{R} \cup \{\infty\}$

$$g(x,t) = \begin{cases} tf(x/t), & \text{if } t > 0 \\ \infty, & \text{otherwise} \end{cases}$$

is given by

$$g^*(y,s) = \begin{cases} 0, & \text{if } f^*(y) \le -s \\ \infty, & \text{otherwise} \end{cases}$$

- Show that the biconjugate of the persepective function g is given by

$$g^{**}(x,t) = \begin{cases} tf(x/t), & \text{if } t > 0 \\ \sigma_{\mathrm{dom}(f^*)}(x), & \text{if } t = 0 \\ \infty, & \text{if } t < 0 \end{cases}$$

where $\sigma_{\mathrm{dom}(f^*)}(x) = \sup_{y \in \mathrm{dom}(f^*)} \langle x, y \rangle$ is the *support function* of $\mathrm{dom}(f^*)$.

Solution.

-

$$\begin{aligned} g^*(y,s) &= \sup_{x \in \mathbb{R}^n, t \in \mathbb{R}, t > 0} \langle x, y \rangle + st - tf(x/t) \\ &\overset{\xi = x/t}{=} \sup_{\xi \in \mathbb{R}^n, t \in \mathbb{R}, t > 0} t\langle \xi, y \rangle + st - tf(\xi) \\ &= \sup_{t>0} \; t \cdot \left(\left[\sup_{\xi} \langle y, \xi \rangle - f(\xi) \right] + s \right) \\ &= \sup_{t>0} \; t \cdot (f^*(y) + s) = \begin{cases} 0 & \text{if } f^*(y) + s \le 0 \\ \infty & \text{otherwise.} \end{cases} \end{aligned}$$

-

$$g^{**}(x,t) = \sup_{f^*(y) \le -s} \langle y, x \rangle + st$$

For $t < 0$ the supremum is unbounded, since for $s \to -\infty$ we have $st \to \infty$.

For $t = 0$ we have

$$g^{**}(x,t) = \sup_{f^*(y) \le -s} \langle y, x \rangle = \sup_{y \in \mathrm{dom}(f^*)} \langle y, x \rangle = \sigma_{\mathrm{dom}(f^*)}(x)$$

Finally let $t > 0$. The supremum in s is achieved at $\hat{s} = -f^*(y)$ and we have

$$\begin{aligned} g^{**}(x,t) &= \sup_{f^*(y) \le -s} \langle y, x \rangle + st = \sup_{y} \langle y, x \rangle - tf^*(y) \\ &= t \sup_{y} \langle y, x/t \rangle - f^*(y) = tf^{**}(x/t) = tf(x/t). \end{aligned}$$

The last equality holds since f is convex (and is everywhere finite hence continuous and proper) we have $f^{**} = f$.

Exercise 3 (6 Points)**.** Let $A \in \mathbb{R}^{m\times n}$ be a linear operator and $f : \mathbb{R}^n \to \mathbb{R} \cup \{\infty\}$ a convex function. Then $Af : \mathbb{R}^m \to \mathbb{R} \cup \{\infty\}$ defined as

$$(Af)(u) := \begin{cases} \inf\limits_{v\in\mathbb{R}^n,\, Av=u} f(v) & \text{if } \exists\, v \in \mathbb{R}^n \text{ s.t. } Av = u \\ \infty & \text{otherwise.} \end{cases}$$

is called the image of f under A. Here, we assume Af is always proper.

1. Show that the convex conjugate $(Af)^*$ of Af is given as $f^* \circ A^\top$ where $(f^* \circ A^\top)(v) := f^*(A^\top v)$.
2. Name the properties that we require for $A^\top f^* = (f \circ A)^*$ to hold. What theorem from the lecture applies here?
3. Give an example of a closed, convex and non-empty set C and a linear operator A s.t. $AC := \{Ax : x \in C\}$ is not closed.
4. Let f be closed, (convex) and proper. Argue that Af does not need to be closed.

Solution.

1. We find

$$\begin{aligned}(Af)^*(u) &= \sup_{v\in\mathbb{R}^n} \langle u, v\rangle - \inf_{w\in\mathbb{R}^n,\, Aw=v} f(w) \\ &= \sup_{\substack{v\in\mathbb{R}^n \\ w\in\mathbb{R}^n,\, Aw=v}} \langle u, v\rangle - f(w) \\ &= \sup_{w\in\mathbb{R}^n} \langle u, Aw\rangle - f(w) \\ &= \sup_{w\in\mathbb{R}^n} \langle A^\top u, w\rangle - f(w) \\ &= f^*(A^\top u)\end{aligned}$$

2. If $A^\top f^*$ is proper, convex and lsc and $m = n$, it is equal to its biconjugate and using the result from the previous part we find:

$$A^\top f^* = (A^\top f^*)^{**} = (f \circ A)^*.$$

3. Choose $C := \operatorname{epi}(\exp) \subseteq \mathbb{R}^2$. C is closed, convex and non-empty, since it is the epigraph of the continuous, convex and proper function f. Let $A := (0, 1)$ then $AC = (0, \infty)$ which is not closed.

4. Let A, C be defined as in the previous part. Define $f := \delta_C$. Then f is closed, proper and convex. We have

$$\begin{aligned}Af(u) &= \inf_{v\in\mathbb{R}^2,\, Av=u} f(v) \\ &= \inf_{v\in\mathbb{R}^2,\, v_2=u} \delta_C(v) \\ &= \begin{cases} 0 & \text{if } u > 0 \\ \infty & \text{otherwise} \end{cases} \\ &= \delta_{(0,\infty)}(u)\end{aligned}$$

We obtain

$$\operatorname{epi}(Af) = (0, \infty) \times [0, \infty),$$

which is not closed. Therefore Af is not closed.