Convex Optimization for Machine Learning and Computer Vision

Lecture: Dr. Tao Wu
Exercises: Emanuel Laude, Zhenzhang Ye
Summer Semester 2018

Computer Vision Group
Institut für Informatik
Technische Universität München

Weekly Exercises 4

Room: 02.09.023

Wednesday, 16.05.2018, 12:15-14:00

Submission deadline: Monday, 14.05.2018, 16:15, Room 02.09.023

Convex cone (10+10 Points)

Exercise 1 (4 points)**.** Assume $J : \mathbb{E} \to \mathbb{R}$, prove following facts of convex conjugate:

- $\tilde{J}(\cdot) = \alpha J(\cdot) \Rightarrow \tilde{J}^*(\cdot) = \alpha J^*(\cdot/\alpha)$, $\alpha > 0$.
- $\tilde{J}(\cdot) = J(\cdot - z) \Rightarrow \tilde{J}^*(\cdot) = J^*(\cdot) + \langle \cdot, z\rangle$.

Exercise 2 (6 points)**.** Assume $J : \mathbb{R}^n \to \mathbb{R}$, compute the convex conjugate of following functions:

- $J(u) = \frac{1}{q}||u||_q^q = \sum_{i=1}^n \frac{1}{q} u_i^q$, $q \in [1, +\infty]$.
- $J(u) = \sum_{i=1}^n u_i \mathrm{log} u_i + \delta_{\triangle^{n-1}}(u)$.
- $J(u) = \begin{cases} \frac{1}{2}u^2, & -\epsilon \leq u \leq \epsilon \\ +\infty, & \text{otherwise} \end{cases}$

Exercise 3 (10 Points)**.**

Definition (Slater's condition)**.** Let $J : \mathbb{R}^n \to \mathbb{R}$, $G : \mathbb{R}^n \to \mathbb{R}^m$ be continuously differentiable and convex, and $H : \mathbb{R}^n \to \mathbb{R}^l$ be affine linear i.e. $Au + b = 0$. Let $U := \{u \in \mathbb{R}^n : g_i(u) \leq 0, h_j(u) = 0,\ 1 \leq i \leq m, 1 \leq j \leq l\}$ denote the feasible set. The condition

$$\exists\, u \in U \text{ s.t. } g_i(u) < 0, h_j(u) = 0, \forall\, 1 \leq i \leq m, 1 \leq j \leq l$$

is called Slater's condition

Definition (Polar cone)**.** For a set C, the polar cone of C is defined as

$$C^o = \{y \in \mathbb{E} : \langle y, d\rangle,\ \forall d \in C\}.$$

Definition (Tangent cone)**.** Let $U \subset \mathbb{E}$ be convex and $u \in U$. Then the tangent cone $T_U(u)$ is defined as

$$T_U(u) = \{d \in \mathbb{E} : \exists u_i \in U \text{ with } u_i \to u \text{ and } \exists t_i \to 0^+, \text{ s.t. } \lim_{i \to +\infty} \frac{u_i - u}{t_i} = d\}$$

Now consider following constrainted optimization problem:

$$\begin{aligned}
\min_u \ & J(u) \\
s.t. \ & g_i(u) \leq 0, && i = 1, \dots, m \\
& h_j(u) = Au + b = 0, && j = 1, \dots, l
\end{aligned}$$

where J and g_i are continuously differentiable and convex functions and h_j are affine linear. Let U be the feasible set defined as before and $U_1 := \{u \in \mathbb{R}^n : G(u) \leq 0\}$ and $U_2 := \{u \in \mathbb{R}^n : H(u) = 0\}$. Assume Slater's condition holds in U.

1. Using following theorem:

   Theorem 1. Let $f_1, \dots, f_n$ are proper convex functions on $\mathbb{R}^n$, and let $f = f_1 + \cdots + f_m$. If the convex sets $\mathrm{ri}(\mathrm{dom} f_i)$, $i = 1, \dots, m$ have a point in common, then
   $$\partial f(u) = \partial f_1(u) + \cdots + \partial f_n(u), \ \forall u.$$

   prove that $N_U(u) = N_{U_1}(u) + N_{U_2}(u)$ where $N_U(u)$ is the normal cone of U at u.

2. Prove that $N_{U_2}(u) = \{\sum_{j=1}^{l} \mu_j \nabla h_j(u) : \mu \in \mathbb{R}^l\}$.

3. Deduce that $T_{U_1}(u) = \{d \in \mathbb{E} : \nabla G_{\mathcal{A}}(u) d \leq 0\}$, where $\mathcal{A}(u) = \{i : g_i(u) = 0, i = 1, \dots, m\}$ is called active set.

   Hint: Firstly, show that $\{d \in \mathbb{E} : \nabla G_{\mathcal{A}}(u) d \leq 0\} \subset \mathrm{cl}(\{d \in \mathbb{E} : \nabla G_{\mathcal{A}}(u) d < 0\}) \subset T_{U_1}(u)$. For the first "$\subset$" relation, consider the linear combination of a boundary point and an inner point. Then show $T_{U_1}(u) \subset \{d \in \mathbb{E} : \nabla G_{\mathcal{A}}(u) d \leq 0\}$.

4. Show that $N_{U_1}(u) = \{\sum_{i=1}^{m} \lambda_i \nabla g_i(u) : \lambda_i \geq 0, \lambda_i g_i(u) = 0, i = 1, \dots, m\}$. You can use following two theorems:

   Theorem 2. If a set $C \subset \mathbb{E}$ is closed and convex, then the bipolar cone is itself i.e. $C^{oo} = C$.

   Theorem 3. Let $C \subset \mathbb{E}$ be a nonempty, convex set and let $u \in C$. Then the normal cone of C at u is the polar cone of the tangent cone of C at u. That is
   $$N_c(u) = (T_c(u))^o.$$

5. Show that $u^* \in U$ satisfies that $-\nabla J(u^*) \in N_U(u^*)$ if and only if u^* is a minimizer.