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 4

Room: 02.09.023
Wednesday, 20.11.2019, 12:15-14:00
Submission deadline: Monday, 18.11.2019, 16:15, Room 02.09.023

Convex conjugate (14+6 Points)

Exercise 1 (4 points)**.** Let $A \in \mathbb{R}^{n\times n}$ be orthonormal, meaning that $A^\top A = AA^\top = I$. Let the convex set C be given as

$$C := \{u \in \mathbb{R}^n : \|Au\|_\infty \leq 1\}.$$

Compute a formula for the projection onto C given as

$$\Pi_C(v) := \mathrm{argmin}_{u\in\mathbb{R}^n} \frac{1}{2}\|u-v\|_2^2, \quad \text{s.t. } u \in C.$$

Hint: Show that the ℓ_2-norm of a vector is invariant under a multiplication with an orthonormal matrix A, meaning that $\|u\|_2 = \|Au\|_2$.

Solution. We begin proving the hint:

$$\|Ax\|_2^2 = \langle Ax, Ax\rangle = \langle A^\top Ax, x\rangle = \langle x, x\rangle = \|x\|_2^2.$$

The idea is to rewrite the projection onto the set C in terms of the projection $\Pi_{\tilde{C}}$ onto the unit ball of the ℓ_∞-norm $\tilde{C} := \{x \in \mathbb{R}^n : \|x\|_\infty \leq 1\}$. With the substitution

$$w := Au \iff u = A^\top w$$

and using the hint we obtain:

$$\begin{aligned}
\Pi_C(v) &= \mathrm{argmin}_{\|Au\|_\infty\leq 1} \frac{1}{2}\|v-u\|^2 \\
&= A^\top \mathrm{argmin}_{\|w\|_\infty\leq 1} \frac{1}{2}\|v-A^\top w\|^2 \\
&= A^\top \mathrm{argmin}_{\|w\|_\infty\leq 1} \frac{1}{2}\|A(v-A^\top w)\|^2 \\
&= A^\top \mathrm{argmin}_{\|w\|_\infty\leq 1} \frac{1}{2}\|Av-AA^\top w\|^2 \\
&= A^\top \mathrm{argmin}_{\|w\|_\infty\leq 1} \frac{1}{2}\|Av-w\|^2 \\
&= A^\top \Pi_{\tilde{C}}(Av).
\end{aligned}$$

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

- $J(u) = \frac{1}{q}||u||_q^q = \sum_{i=1}^n \frac{1}{q}|u_i|^q$, $q \in [1, +\infty]$.
- $J(u) = \sum_{i=1}^n u_i \text{log} u_i + \delta_{\triangle^{n-1}}(u)$.
- $J(u) = \begin{cases} \frac{1}{2}\|u\|_2^2, & \|u\|_2 \leq \epsilon \\ +\infty, & \text{otherwise} \end{cases}$

Solution.

- Assume $q \in (1, +\infty)$, $J^*(v) = \sup_u \langle u, v\rangle - J(u)$. Since it is separable, we apply first-order optimality condition elementwisely:

$$\sup_{u_i} \langle u_i, v_i\rangle - \frac{1}{q}(|u_i|)^q \Rightarrow 0 = v_i - |u_i|^{q-1}\text{sign}(u_i) \Rightarrow u_i = |v_i|^{1/(q-1)}\text{sign}(v_i)$$

Substitute u_i back to the first equation, we have

$$\begin{aligned} J^*(v)_i &= |v_i|^{q/(q-1)} - \frac{1}{q}|v_i|^{q/(q-1)} \\ &= (1 - \frac{1}{q})|v_i|^{q/(q-1)} \\ &= (1 - \frac{1}{q})|v_i|^{1/(1-\frac{1}{q})} \end{aligned}$$

Substituting $\frac{1}{p} = 1 - \frac{1}{q}$, we get $J^*(v) = \frac{1}{p}||v||_p^p$.
Now, consider $q = 1$, we have $J^*(v) = \sup_u \langle u, v\rangle - \|u\|_1 = \sum_{i=1}^n \sup_{u_i} u_i(v_i - \text{sign}(u))$. The result will be

$$J^*(v) = \begin{cases} 0, & \|v\|_\infty \leq 1 \\ \infty, & \text{otherwise.} \end{cases}$$

For $q = +\infty$, $J(u) = \|u\|_\infty$ and $J^*(v) = \sup_u \langle u, v\rangle - \max_j |u_j|$. Compute the subdifferential we hope to get: $v \in \{x \in \mathbb{R}^n : \|x\| \leq 1, \langle x, u\rangle = \|u\|_\infty\}$.
Therefore, if $\|v\|_1 \leq 1$, we can find a u such that v is in the set. We achieve $J^*(v) = 0$. Otherwise, we can choose $u_i = t\text{sign}(v_i)$ where $t > 0$ is a scalar. We can make $J^*(v) \to +\infty$ as $t \to +\infty$.
The result is then:

$$J^*(v) = \begin{cases} 0, & \|v\|_1 \leq 1 \\ \infty, & \text{otherwise.} \end{cases}$$

- Consider the convex conjugate elementwisely: $J^*(v) = \sup_u \sum_i^n u_i v_i - u_i \text{log} u_i - \delta_{\triangle^{n-1}}(u)$. Let's consider the following minimization problem given v_i:

$$\begin{aligned} &\min_u \sum_i^n u_i \text{log} u_i - u_i v_i \\ &\text{s.t.} \mathbb{1}u = 1 \end{aligned}$$

where $\mathbb{1} = [1, \dots, 1] \in \mathbb{R}^n$. It is obvious that this two problems share the same optimal variable u^* and the domain of log implies $u_i > 0$. Since the feasible set is compact and original energy function is continuous, the KKT condition holds on u^*. Therefore, we have certain $\lambda \in \mathbb{R}$ such that

$$\log u_i^* + 1 - v_i + \lambda = 0, \ \forall i = 1, \dots, n$$

which give $u_i^* = \exp\{-\lambda + v_i - 1\}$. Additionally, $\sum_{i=1}^n u_i^* = 1$. We can get

$$0 = \log(\sum_{i=1}^n \exp\{-\lambda + v_i - 1\}) = \log(\exp\{-\lambda - 1\} \sum_{i=1}^n e^{v_i}) = (-\lambda - 1) + \log(\sum_{i=1}^n e^{v_i})$$

Now, substitute u^* back into the convex conjugate and we can get

$$\begin{aligned} J(v)^* &= \sum_i^n \exp\{-\lambda + v_i - 1\} v_i - \exp\{-\lambda + v_i - 1\}(-\lambda + v_i - 1) \\ &= \sum_i^n -\exp\{-\lambda + v_i - 1\}(-\lambda - 1) \\ &= -(-\lambda - 1) = \log(\sum_{i=1}^n e^{v_i}) \end{aligned}$$

- Rewrite the convex conjugate as $J^*(v) = \sup_{\|u\|_2 \le \epsilon} \langle u, v \rangle - \frac{1}{2} \|u\|_2^2$. We first try to find the corresponding u^*.

$$\begin{aligned} u^* &= \text{argmin}_{\|u\|_2 \le \epsilon} \frac{1}{2} \|u\|_2^2 - \langle u, v \rangle + \frac{1}{2} \|v\|_2^2 \\ &= \text{argmin}_{\|u\|_2 \le \epsilon} \frac{1}{2} \|u - v\|_2^2 \end{aligned}$$

which is a projection problem i.e. project v into a convex set $\{u : \|u\|_2 \le \epsilon\}$. Therefore, if $\|v\|_2 \le \epsilon$, $u^* = v$. Otherwise, $u^* = \epsilon \frac{v}{\|v\|}$.

$$J^*(v) = \begin{cases} \frac{1}{2} \|v\|_2^2, & \|v\|_2 \le \epsilon \\ \epsilon \|v\|_2^2 - \frac{1}{2}\epsilon^2, & \text{otherwise} \end{cases}$$

Exercise 3 (4 points)**.** Show that projection onto a convex set is Lipschitz continuous with constant equals 1, i.e.

$$||\Pi_C(u) - \Pi_C(v)|| \le ||u - v||, \ \forall u, v \in \mathbb{E}$$

where C is a convex set.

Solution. Given a point u, recall the property of projection:

$$\langle u - \Pi_C(u), x - \Pi_C(u) \rangle \le 0, \ \forall x \in C.$$

Since $\Pi_C(v)$ is also an element in C, we get:

$$\langle u - \Pi_C(u), \Pi_C(v) - \Pi_C(u) \rangle \leq 0.$$

As same as above, we can get the ineuqlity for point v:

$$\langle v - \Pi_C(v), \Pi_C(u) - \Pi_C(v) \rangle \leq 0.$$

Sum above inequalities up, we have:

$$\begin{aligned}
&\langle u - \Pi_C(u) + \Pi_C(v) - v, \Pi_C(v) - \Pi_C(u) \rangle \leq 0 \\
\Rightarrow &\langle \Pi_C(v) - \Pi_C(u), \Pi_C(v) - \Pi_C(u) \rangle \leq \langle v - u, \Pi_C(v) - \Pi_C(u) \rangle \\
\Rightarrow &\|\Pi_C(v) - \Pi_C(u)\|^2 \leq \|v - u\| \, \|\Pi_C(v) - \Pi_C(u)\| \\
\Rightarrow &\|\Pi_C(v) - \Pi_C(u)\| \leq \|v - u\|
\end{aligned}$$

Exercise 4 (6 points)**.** Let C_i, $1 \leq i \leq n$ be a family of closed convex sets such that

$$\bigcap_{1 \leq i \leq n} C_i \neq \emptyset.$$

Show that the problem of finding an element u^* in the intersection

$$u^* \in \bigcap_{1 \leq i \leq n} C_i$$

can be formulated as the following optimization problem:

$$u^* \in \arg \min_{u \in \bigcap_{i \in \mathcal{I}} C_i} \sum_{\substack{j \notin \mathcal{I} \\ 1 \leq j \leq n}} d^2(u, C_j),$$

where $\mathcal{I} \subseteq \{1, 2, \ldots, n\}$ can be arbitrary (including the empty set) and $d(z, X)$ is the distance of a point z to the closed convex set X defined as

$$d(z, X) := \min_{x \in X} \|x - z\|_2.$$

Solution. Since all C_i are closed and convex, $d(u, C_i)$ is well defined. Since $d^2(u, C_j) \geq 0$ and $d^2(u, C_j) = 0 \iff u \in C_j$,

$$u \in \bigcap_{\substack{j \notin \mathcal{I} \\ 1 \leq j \leq n}} C_j \iff \sum_{\substack{j \notin \mathcal{I} \\ 1 \leq j \leq n}} d^2(u, C_j) = 0.$$

This yields that

$$0 = \sum_{\substack{j \notin \mathcal{I} \\ 1 \leq j \leq n}} d^2(u^*, C_j) + \delta_{\bigcap_{i \in \mathcal{I}} C_i}(u^*)$$

iff $u^* \in \bigcap_{1 \leq i \leq n} C_i$. Since $\bigcap_{1 \leq i \leq n} C_i$ non-empty

$$\operatorname{argmin}_{u \in \bigcap_{i \in \mathcal{I}} C_i} \sum_{\substack{j \notin \mathcal{I} \\ 1 \leq j \leq n}} d^2(u, C_j) \subset \bigcap_{1 \leq i \leq n} C_i.$$

Programming: SUDOKU(Due: 25.11) (12 Points)

Exercise 5 (12 Points)**.** Solve the SUDOKU given in the file `exampleSudoku1.mat` (`sudoku array` in python) with projected gradient descent. The algorithm is already given, you only need to figure out the correct formulation.

We represent a SUDOKU as a matrix $\boldsymbol{u} \in \{0,1\}^{9\times9\times9}$, where $\boldsymbol{u}_{i,j,k} = 1$ means $\boldsymbol{u}_{i,j} = k$, *i.e.* we fill number k on position (i,j). Therefore, we have following rules, where $f_{i,j}$ are the given entries and B_l is a 9×9 block:

1. Respect given entries: $\boldsymbol{u}_{i,j,k} = 1$, if $f_{i,j} = k$
2. One number for each blank spot: $\sum_k \boldsymbol{u}_{i,j,k} = 1$, $\forall i, j$
3. Numbers occur in a row once: $\sum_j \boldsymbol{u}_{i,j,k} = 1$, $\forall i, k$
4. Numbers occur in a column once: $\sum_i \boldsymbol{u}_{i,j,k} = 1$, $\forall j, k$
5. Numbers occur in a block once: $\sum_{(i,j)\in B_l} \boldsymbol{u}_{i,j,k} = 1$, $\forall B_l, k$

First of all, since the feasible set of u is non-convex, we perform a convex relaxation on it such that $\boldsymbol{u}_{i,j,k} \in [0,1]$.

Since constraints $2-5$ are linear and using the idea from exercise 4, we can vectorize u and try to find a linear operator A, the opimal u^* should satisfy $Au^* = 1$. The problem then can be converted into following convex minimization one:

$$\begin{aligned} u^* &= \text{argmin}_u \frac{1}{2} \left\| Au - 1 \right\|^2, \\ s.t.\ & u_{ijk} \in [0,1], \\ & u_{ijk} = 1 \text{ if } f_{ij} = k. \end{aligned}$$

Your task here is try to figure out the linear operator A, such that if u is a valid solution then $Au = 1$.

Hint: 1. How the optimal u^* look like?

2. We vectorize u, therefore figure out the dimension of A first.

3. Starting from the third constraint, how A should look like only for that constraint?