Convex Optimization for Machine Learning and Computer Vision

Lecture: Dr. Virginia Estellers
Exercises: Emanuel Laude
Winter Semester 2017/18

Computer Vision Group
Institut für Informatik
Technische Universität München

Weekly Exercises 5

Room: 02.09.023
Friday, 02.12.2017, 09:15-11:00
Submission deadline theory: Monday, 27.11.2017, 10:15, Room 02.09.023
Submission deadline coding: Monday, 04.12.2017, 10:15, Room 02.09.023

Theory: Fenchel Duality (0+16 Points)

Exercise 1 (4 Points)**.** Let $X, Y \in \mathbb{R}^{m\times n}$ be matrices and let $Y_i, X_i \in \mathbb{R}^m$ denote the i-th columns of X, Y. Then, the Frobenius scalar product is defined as follows:

$$\langle X, Y\rangle_F := \sum_{i=1}^{n} \langle X_i, Y_i\rangle, \tag{1}$$

where $\langle X_i, Y_i\rangle$ is the classical vector scalar product. For notational convenience we often omit the subscript F in $\langle\cdot,\cdot\rangle_F$. Compute the convex conjugates of the following functions:

1. $f_1 : \mathbb{R}^{m\times n} \to \mathbb{R}\cup\{\infty\}$ where $f_1(X) = \|X\|_{2,\infty} := \max_{1\le i\le n} \|X_i\|_2$.
2. $f_2 : \mathbb{R}^{m\times n} \to \mathbb{R}\cup\{\infty\}$ where

$$f_2(X) := \delta_{\|\cdot\|_{2,1}\le 1}(X) = \begin{cases} 0 & \text{if } \|X\|_{2,1} := \sum_{i=1}^{n} \|X_i\|_2 \le 1, \\ \infty & \text{otherwise.} \end{cases} \tag{2}$$

Solution. 1. Let $X \in \mathbb{R}^{m\times n}$, $\|X\|_{2,\infty} \le 1$. We have any for $Y \in \mathbb{R}^{m\times n}$:

$$\begin{aligned} \langle X, Y\rangle_F &= \sum_{i=1}^{n} \langle X_i, Y_i\rangle \\ &\le \sum_{i=1}^{n} |X_i| \cdot |Y_i| \\ &\le \sum_{i=1}^{n} |X_i| \cdot \max_{1\le j\le n} |Y_j| \\ &= \|X\|_{2,1} \cdot \|Y\|_{2,\infty}. \end{aligned}$$

This implies that

$$\langle X, Y\rangle_F - \|Y\|_{2,\infty} \le \|X\|_{2,1} \cdot \|Y\|_{2,\infty} - \|Y\|_{2,\infty} = (\|X\|_{2,1} - 1) \cdot \|Y\|_{2,\infty} \le 0$$

Since $\langle X, 0\rangle_F - \|0\|_{2,\infty} = 0$ we get

$$f_1^*(X) = \sup_{Y \in \mathbb{R}^{m\times n}} \langle X, Y\rangle_F - \|Y\|_{2,\infty} = 0.$$

Now let $\|X\|_{2,1} > 1$. Define $Y \in \mathbb{R}^{m\times n}$ so that the i-th column Y_i of Y, $1 \le i \le n$ is given as $Y_i := \frac{X_i}{\|X_i\|_2}$, which implies $\|Y\|_{2,\infty} = 1$. We get

$$\langle X, Y\rangle_F = \sum_{i=1}^{n} \|X_i\|_2 = \|X\|_{2,1}.$$

For $\alpha > 0$ we get

$$\langle X, \alpha Y\rangle_F - \|\alpha Y\|_{2,\infty} = \alpha \underbrace{(\|X\|_{2,1} - 1)}_{>1}.$$

Therefore,

$$f_1^*(X) = \sup_{Y \in \mathbb{R}^{m\times n}} \langle X, Y\rangle_F - \|Y\|_{2,\infty} = \infty.$$

Altogether we obtain

$$f_1^*(X) = \delta_{\|\cdot\|_{2,1}\le 1}(X).$$

2. We have $f_2 = f_1^*$ and since f_1 is closed, proper and convex we have

$$f_2^* = f_1^{**} = f_1.$$

Exercise 2 (8 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.

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 closed, proper and convex 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 $C := \mathrm{epi}(\exp) \subseteq \mathbb{R}^2$. 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

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

We obtain

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

which is not closed. Therefore Af is not closed.

Exercise 3 (4 Points)**.** Let $H : \mathbb{R}^n \to \mathbb{R} \cup \{\infty\}$ and $R : \mathbb{R}^m \to \mathbb{R} \cup \{\infty\}$ be proper, closed, convex functions and $K \in \mathbb{R}^{m \times n}$ a linear operator. Let there exist a $u \in \text{ri}(\text{dom}(H))$ such that $Ku \in \text{ri}(\text{dom}(R))$. Let $f(u) := H(u) + R(Ku) = \tilde{f}(Au)$, where

$$A := \begin{pmatrix} I \\ K \end{pmatrix} \in \mathbb{R}^{n+m \times n}, \quad \tilde{f}(u,d) := H(u) + R(d).$$

Prove Fenchel's Duality Theorem, i.e. show that

$$\inf_{u \in \mathbb{R}^n} H(u) + R(Ku) = \sup_{q \in \mathbb{R}^m} -H^*(-K^\top q) - R^*(q)$$

Hint: You can assume that the conditions above guarantee that $A^\top \tilde{f}^*$ is closed proper and convex. Argue that $\tilde{f}^*(u,d) = H^*(u) + R^*(d)$. Which result from the lecture applies here? Begin your computation with

$$\inf_{u \in \mathbb{R}^n} f(u) = -\sup_{u \in \mathbb{R}^n} \langle u, 0 \rangle - f(u) = -f^*(0) \dots$$

Solution. Using the result from the lecture for the convex conjugate of a decoupled sum we obtain for $\tilde{f}^*$:

$$\tilde{f}^*(u,d) = H^*(u) + R^*(d).$$

Further we have

$$\begin{aligned}
\inf_{u \in \mathbb{R}^n} H(u) + R(Ku) &= \inf_{u \in \mathbb{R}^n} f(u) \\
&= -\sup_{u \in \mathbb{R}^n} \langle u, 0 \rangle - f(u) \\
&= -f^*(0) \\
&= -(\tilde{f} \circ A)^*(0) \\
&= -(A^\top \tilde{f}^*)(0) \\
&= -\inf_{(q,p),\, q + K^\top p = 0} H^*(q) + R^*(p) \\
&= -\inf_{p \in \mathbb{R}^m} H^*(-K^\top p) + R^*(p) \\
&= \sup_{p \in \mathbb{R}^m} -H^*(-K^\top p) - R^*(p)
\end{aligned}$$

Programming: Denoising with Duality (Due on 04.12.2017) (12 Points)

Exercise 4 (12 Points)**.** Denoise the noisy input image f, given in the file `noisy_input.png` by solving the dual problem of:

$$\min_u \frac{1}{2} \|u - f\|^2 + \alpha \|Du\|_{2,1}$$

with projected gradient descent. For details of the derivation of the dual problem cf. the lecture.