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 2

Room: 02.09.023

Wednesday, 02.05.2018, 12:15-14:00

Submission deadline: Monday, 30.04.2018, 16:15, Room 02.09.023

Convex sets and functions (12 Points + 4 Bonus)

Exercise 1 (4 Points)**.** Let $J : \mathbb{E} \to \bar{\mathbb{R}}$ be proper. Prove the equivalence of the following statements:

- J is convex.
- $\operatorname{epi}(J) := \left\{ \begin{pmatrix} u \\ \alpha \end{pmatrix} \in \mathbb{E} \times \mathbb{R} : J(u) \leq \alpha \right\}$ is convex.

Solution. Let J be convex, $\lambda \in [0,1]$ and $(u_1, \alpha_1), (u_2, \alpha_2) \in \operatorname{epi}(J)$. This means $J(u_1) \leq \alpha_1 < \infty$ and $J(u_2) \leq \alpha_2 < \infty$ and therefore $u_1, u_2 \in \operatorname{dom}(J)$. Due to the convexity of J we have that:

1. $\operatorname{dom}(J)$ convex and therefore $\lambda u_1 + (1-\lambda)u_2 \in \operatorname{dom}(J)$, and
2. $J(\lambda u_1 + (1-\lambda)u_2) \leq \lambda J(u_1) + (1-\lambda)J(u_2) \leq \lambda\alpha_1 + (1-\lambda)\alpha_2$.

This means that

$$\begin{pmatrix} \lambda u_1 + (1-\lambda)u_2 \\ \lambda\alpha_1 + (1-\lambda)\alpha_2 \end{pmatrix} = \lambda \begin{pmatrix} u_1 \\ \alpha_1 \end{pmatrix} + (1-\lambda) \begin{pmatrix} u_2 \\ \alpha_2 \end{pmatrix} \in \operatorname{epi}(J)$$

and therefore $\operatorname{epi}(J)$ convex. Let conversely $\operatorname{epi}(J)$ be convex and $u_1, u_2 \in \operatorname{dom}(J) := \{u \in \mathbb{E} : J(u) < \infty\}$. By definition of the epigraph set $(u_1, J(u_1)), (u_2, J(u_2)) \in \operatorname{epi}(J)$ and due to the convexity of $\operatorname{epi}(J)$

$$\lambda \begin{pmatrix} u_1 \\ J(u_1) \end{pmatrix} + (1-\lambda) \begin{pmatrix} u_2 \\ J(u_2) \end{pmatrix} \in \operatorname{epi}(J).$$

This means

$$J(\lambda u_1 + (1-\lambda)u_2) \leq \lambda J(u_1) + (1-\lambda)J(u_2).$$

It remains to show that $\operatorname{dom}(J)$ is convex. We have:

$$\begin{aligned} \operatorname{dom}(J) &= \{u \in \mathbb{E} : J(u) < \infty\} \\ &= \{u \in \mathbb{E} : \exists\, \alpha \in \mathbb{R} : J(u) \leq \alpha\} \\ &= \{u \in \mathbb{E} : \exists\, \alpha \in \mathbb{R} \text{ s.t. } (u, \alpha) \in \operatorname{epi}(J)\} \end{aligned}$$

Since $\operatorname{epi}(J)$ is convex it immediatly follows, that $\operatorname{dom}(J)$ is convex. Overall this proves that J convex.

Exercise 2 (4 Points)**.** Show that the following functions $J : \mathbb{R}^n \to \overline{\mathbb{R}}$ are convex:

- $J(u) = \|u\|$, for any norm $\|\cdot\|$ over a normed vector space.
- $J(u) = F(Ku)$, for convex $F : \mathbb{R}^n \to \overline{\mathbb{R}}$ and linear $K : \mathbb{R}^m \to \mathbb{R}^n$.
- (Jensen's inequality) J is convex iff

$$J\left(\sum_{i=1}^{n} \alpha_i u^i\right) \leq \sum_{i=1}^{n} \alpha_i J(u^i),$$

  whenever $\{u^i\}_{i=1}^n \subset \mathbb{R}^n$, $\{\alpha_i\}_{i=1}^n \subset [0,1]$, $\sum_{i=1}^n \alpha_i = 1$.

Solution.

- Take $u, v \in \mathbb{R}^n$, $\lambda \in [0,1]$:

$$\begin{aligned} J(\lambda u + (1-\alpha)v) = \|\lambda u + (1-\lambda)v\| \leq \\ \|\lambda u\| + \|(1-\lambda)v\| = \lambda \|u\| + (1-\lambda)\|v\|. \end{aligned} \tag{1}$$

- Take $u, v \in \mathrm{dom}(J)$, $\lambda \in [0,1]$.

$$\begin{aligned} J(\lambda u + (1-\lambda)v) := F(K(\lambda u + (1-\lambda)v) = \\ F(\lambda Ku + (1-\lambda)Kv)) \leq \\ \lambda F(Ku) + (1-\lambda)F(Kv) = \underbrace{\lambda J(u) + (1-\lambda)J(v)}_{<\infty,\ \text{since } u,v \in \mathrm{dom}(J)} \end{aligned} \tag{2}$$

  This shows that J is convex on its domain and $\mathrm{dom}(J)$ is a convex set.

- "$\Leftarrow$": For $n = 2$ it is precisely the definition of convexity.
  "$\Rightarrow$": We prove this statement using induction. The cases $n = 1$ and $n = 2$ are trivial. Now assume the inequality holds for some $n \geq 1$. Without loss of generality we can assume $\alpha_{n+1} \neq 0$, since the case $\alpha_{n+1} = 0$ follows directly from the assumption.

$$\begin{aligned} J\left(\sum_{i=1}^{n+1} \alpha_i u_i\right) &= J\left(\sum_{i=1}^{n} \alpha_i u_i + \alpha_{n+1} u_{n+1}\right) \\ &= J\left((1-\alpha_{n+1}) \sum_{i=1}^{n} \frac{\alpha_i}{1-\alpha_{n+1}} u_i + \alpha_{n+1} u_{n+1}\right) \\ &\leq (1-\alpha_{n+1}) J\left(\sum_{i=1}^{n} \frac{\alpha_i}{1-\alpha_{n+1}} u_i\right) + \alpha_{n+1} J(u_{n+1}) \\ &\leq (1-\alpha_{n+1}) \sum_{i=1}^{n} \frac{\alpha_i}{1-\alpha_{n+1}} J(u_i) + \alpha_{n+1} J(u_{n+1}) \\ &= \sum_{i=1}^{n} \alpha_i J(u_i) + \alpha_{n+1} J(u_{n+1}) = \sum_{i=1}^{n+1} \alpha_i J(u_i). \end{aligned} \tag{3}$$

  Hence it also holds for $n+1$ and by the principle of induction we are finished.

Exercise 3 (4 Points)**.** Let $U \subset \mathbb{E}$ open and convex and let $J : U \to \mathbb{R}$ be twice continuously differentiable. Prove the equivalence of the following statements:

- J is convex.
- For all $u \in U$ the Hessian $\nabla^2 J(u)$ is positive semidefinite ($\forall\, v \in \mathbb{E} : v^\top \nabla^2 J(u) v \geq 0$).

Hints: You can use that for $u, v \in U$ it holds that J is convex iff

$$(v-u)^\top \nabla J(u) \leq J(v) - J(u).$$

Further recall that there are two variants of the Taylor expansion:

$$J(u+td) = J(u) + td^\top \nabla J(u) + \frac{t^2}{2} d^\top \nabla^2 J(u) d + o(t^2)$$

with $\lim_{t\to 0} \frac{o(t^2)}{t^2} = 0$ and

$$J(u+d) = J(u) + d^\top \nabla J(u) + \frac{1}{2} d^\top \nabla^2 J(u+td) d$$

for appropriate $t \in (0,1)$.

Solution. Let J be convex, $u \in U$ and $d \in \mathbb{R}^n$. Since U is open there exists $\tau > 0$ s.t. for all $t \in (0, \tau]$ we have that $u + td \in U$. Using the Taylor expansion given in the hint we obtain

$$0 \overset{\text{Hint}}{\leq} J(u+td) - J(u) - td^\top \nabla J(u) = \frac{t^2}{2} d^\top \nabla^2 J(u) d + o(t^2)$$

Multiplying both sides with $\frac{2}{t^2}$ yields

$$0 \leq d^\top \nabla^2 J(u) d + 2 \underbrace{\frac{o(t^2)}{t^2}}_{\to 0}.$$

Let conversely $\nabla^2 J(z)$ be positive semidefinite for all $z \in U$ and let $u, v \in X$. Using the Taylor expansion we have

$$J(v) = J(u+(v-u)) = J(u) + (v-u)^\top \nabla J(u) + \frac{1}{2} \underbrace{(v-u)^\top \nabla^2 J(u + t(v-u))(v-u)}_{\geq 0 \text{ by assumption.}}$$

and therefore

$$J(v) - J(u) \geq (v-u)^\top \nabla J(u),$$

which means that J is convex.

Exercise 4 (4 points)**.** Prove the following statement using induction over m: Let $K_1, \ldots, K_m \subset \mathbb{R}^n$, $m \geq n+1$, be convex, such that for all $\mathcal{I} \subset \{1, \ldots, m\}$ with $|\mathcal{I}| = n+1$ it holds that $\bigcap_{i \in \mathcal{I}} K_i \neq \emptyset$. Then $\bigcap_{i=1}^m K_i \neq \emptyset$.

Hint: Use exercise 4 from the first exercise sheet.

Solution. Base case: for $m = n + 1$ the statement clearly holds.

Inductive step: $m \to m + 1$. For any $\mathcal{I} \subset \{1, \dots, m + 1\}$ with $|\mathcal{I}| = n + 1$ assume that $\bigcap_{i \in \mathcal{I}} K_i \neq \emptyset$. Fix $j \in \{1, 2, \dots, m + 1\}$. The assumption implies that for all $\mathcal{I}' \subset \{1, \dots, m + 1\} \setminus \{j\}$ with $|\mathcal{I}'| = n + 1$ it holds that $\bigcap_{i \in \mathcal{I}'} K_i \neq \emptyset$. We may now apply the induction hypothesis to the sets $K_1, \dots, K_{m+1}$ excluding K_j and the sets $\mathcal{I}'$ and conclude that for any $\mathcal{J} \subset \{1, \dots, m + 1\}$ with $\mathcal{J} \neq \emptyset$:

$$x_j \in \bigcap_{i=1, i \neq j}^{m+1} K_i \subset \begin{cases} \bigcap_{i \in \mathcal{J}} K_i & \text{if } j \notin \mathcal{J} \\ \bigcap_{i \notin \mathcal{J}} K_i & \text{if } j \in \mathcal{J}. \end{cases}$$

Now, consider the partitions $A_1 := \{x_j : j \notin \mathcal{J}\}$, $A_2 := \{x_j : j \in \mathcal{J}\}$ of the set $A := \{x_1, x_2, \cdots x_{m+1}\}$ determined via $\mathcal{J}$. Since $m+1 \geq n+2$ we know from exercise 4 of the last sheet that there exists an $\mathcal{J}' \subset \{1, \dots, m + 1\}$ (the proof can easily be adapted to the more general case $m + 1 \geq n + 2$) so that $\operatorname{conv}(A_1) \cap \operatorname{conv}(A_2) \neq \emptyset$. Since the K_i are convex and the intersection of convex sets is convex we have that $\operatorname{conv}(A_1) \subset \bigcap_{i \in \mathcal{J}'} K_i$ and $\operatorname{conv}(A_2) \subset \bigcap_{i \notin \mathcal{J}'} K_i$. Overall we have that

$$\emptyset \neq \operatorname{conv}(A_1) \cap \operatorname{conv}(A_2) \subset \bigcap_{i \in \mathcal{J}'} K_i \cap \bigcap_{i \notin \mathcal{J}'} K_i = \bigcap_{i=1}^{m+1} K_i.$$

The theorem is called Helly's Theorem.

Programming: Inpainting(Due date: 07.05) (12 Points)

Exercise 5 (12 Points)**.** Write a MATLAB program that solves the inpainting problem for the vegetable image:

$$\min_{u \in \mathbb{R}^{n \times m}} \sum_{i,j} (u_{i,j} - u_{i-1,j})^2 + (u_{i,j} - u_{i,j-1})^2 \quad \text{s.t. } u_{i,j} = f_{i,j} \ \forall (i,j) \in I,$$

with index set I of pixels to keep. Those can be identified as the white pixels of the mask image.

Hint: The constrained optimization problem can be reformulated so that it becomes unconstrained: Rewrite the objective as a least squares problem in terms of the unknown intensities $u_{i,j}$, $(i,j) \notin I$ using sparse linear operators: Find linear operators X, Y s.t. u can be decomposed as

$$u = X\tilde{u} + Yf$$

where $\tilde{u}$ contains only the unknown intensities. Optimize for $\tilde{u}$ instead of u. You may use MATALBs `mldivide`.