Convex Optimization for Computer Vision
Lecture: M. Möller and T. Möllenhoff
Exercises: E. Laude
Summer Semester 2016

Computer Vision Group
Institut für Informatik
Technische Universität München

Weekly Exercises 1

Room: 02.09.023

Friday, 22.04.2016, 09:00-11:00

Submission deadline: Wednesday, 20.04.2016, 14:00, Room 02.09.023

Theory: Convex Sets and Functions (12+8 Points)

Exercise 1 (4 Points)**.** Let $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ be proper. Prove the equivalence of the following statements:

- f is convex.
- $\text{epi}(f) := \left\{ \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^{n+1} : f(x) \leq y \right\}$ is convex.

Solution. Let f be convex, $\lambda \in [0,1]$ and $(x_1, y_1), (x_2, y_2) \in \text{epi}(f)$. This means $f(x_1) \leq y_1 < \infty$ and $f(x_2) \leq y_2 < \infty$ and therefore $x_1, x_2 \in \text{dom}(f)$. Due to the convexity of f we have that:

1. $\text{dom}(f)$ convex and therefore $\lambda x_1 + (1-\lambda)x_2 \in \text{dom}(f)$, and
2. $f(\lambda x_1 + (1-\lambda)x_2) \leq \lambda f(x_1) + (1-\lambda)f(x_2) \leq \lambda y_1 + (1-\lambda)y_2$.

This means that

$$\begin{pmatrix} \lambda x_1 + (1-\lambda)x_2 \\ \lambda y_1 + (1-\lambda)y_2 \end{pmatrix} = \lambda \begin{pmatrix} x_1 \\ y_1 \end{pmatrix} + (1-\lambda) \begin{pmatrix} x_2 \\ y_2 \end{pmatrix} \in \text{epi}(f)$$

and therefore $\text{epi}(f)$ convex. Let conversely $\text{epi}(f)$ be convex and $x_1, x_2 \in \text{dom}(f) := \{x \in \mathbb{R}^n : f(x) < \infty\}$. By definition of the epigraph set $(x_1, f(x_1)), (x_2, f(x_2)) \in \text{epi}(f)$ and due to the convexity of $\text{epi}(f)$

$$\lambda \begin{pmatrix} x_1 \\ f(x_1) \end{pmatrix} + (1-\lambda) \begin{pmatrix} x_2 \\ f(x_2) \end{pmatrix} \in \text{epi}(f).$$

This means

$$f(\lambda x_1 + (1-\lambda)x_2) \leq \lambda f(x_1) + (1-\lambda)f(x_2).$$

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

$$\begin{aligned} \text{dom}(f) &= \{x \in \mathbb{R}^n : f(x) < \infty\} \\ &= \{x \in \mathbb{R}^n : \exists\, y \in \mathbb{R} : f(x) \leq y\} \\ &= \{x \in \mathbb{R}^n : \exists\, y \in \mathbb{R} \text{ s.t. } (x,y) \in \text{epi}(f)\} \end{aligned}$$

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

Exercise 2 (4 Points)**.** Let $g : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ be convex. Show that the perspective function $f : \mathbb{R}^n \times \mathbb{R} \to \mathbb{R} \cup \{+\infty\}$ of g given as

$$f(x,t) := \begin{cases} t\, g\left(\frac{x}{t}\right) & \text{if } t > 0 \text{ and } \frac{x}{t} \in \operatorname{dom}(g) \\ +\infty & \text{otherwise,} \end{cases}$$

is convex.

Solution. Let $\lambda \in [0,1]$, and $(x_1,t_1),(x_2,t_2) \in \operatorname{dom}(f)$. That means $x_1, x_2 \in \mathbb{R}^n$ and $t_1, t_2 > 0$ s.t. $\frac{x_1}{t_1}, \frac{x_2}{t_2} \in \operatorname{dom}(g)$. We have:

$$\begin{aligned} f(\lambda x_1 + (1-\lambda)x_2, \lambda t_1 + (1-\lambda)t_2) &= (\lambda t_1 + (1-\lambda)t_2)\, g\left(\frac{\lambda x_1 + (1-\lambda)x_2}{\lambda t_1 + (1-\lambda)t_2}\right) \\ &= (\lambda t_1 + (1-\lambda)t_2)\, g\left(\frac{\lambda t_1 \frac{x_1}{t_1} + (1-\lambda)t_2 \frac{x_2}{t_2}}{\lambda t_1 + (1-\lambda)t_2}\right) \\ &= (\lambda t_1 + (1-\lambda)t_2) \\ &\quad g\left(\underbrace{\frac{\lambda t_1}{\lambda t_1 + (1-\lambda)t_2}\frac{x_1}{t_1} + \frac{(1-\lambda)t_2}{\lambda t_2 + (1-\lambda)t_2}\frac{x_2}{t_2}}_{:=\alpha\frac{x_1}{t_1} + (1-\alpha)\frac{x_2}{t_2},\, 0\le\alpha\le 1}\right) \\ &\le (\lambda t_1 + (1-\lambda)t_2) \\ &\quad \left(\frac{\lambda t_1}{\lambda t_1 + (1-\lambda)t_2}\, g\left(\frac{x_1}{t_1}\right) + \frac{(1-\lambda)t_2}{\lambda t_2 + (1-\lambda)t_2}\, g\left(\frac{x_2}{t_2}\right)\right) \\ &= \lambda t_1\, g\left(\frac{x_1}{t_1}\right) + (1-\lambda)t_2\, g\left(\frac{x_2}{t_2}\right) < +\infty \end{aligned}$$

The above computation shows that both f is convex on its domain

$$\operatorname{dom}(f) = \left\{(x,t) \in \mathbb{R}^{n+1} : t > 0, \, \frac{x}{t} \in \operatorname{dom}(g)\right\}$$

and $\operatorname{dom}(f)$ is a convex set. This implies that f is convex.

Exercise 3 (4 Points)**.** Let $\emptyset \neq X \subset \mathbb{R}^n$. Prove the equivalence of the following statements:

- X is closed.
- Every convergent sequence $\{x_n\}_{n\in\mathbb{N}} \subset X$ attains its limit in X.

Solution. Let X be closed. By definition this means that the complement of X given as $X_C := \mathbb{R}^n \setminus X$ is open meaning that for all $x \in X_C$ there exists $\epsilon > 0$ s.t. the ball $B_\epsilon(x)$ is entirely contained in X_C:

$$B_\epsilon(x) \cap X = \emptyset.$$

Suppose that there exists a convergent sequence $X \supset \{x_n\}_{n\in\mathbb{N}} \to x$ with $x \notin X$. However, by definition of convergence for all $\epsilon > 0$ there exists $N \in \mathbb{N}$ s.t.

$$X \ni x_n \in B_\epsilon(x)$$

for all $n \geq N$, which contradicts the assumption. Let conversely X not be closed (not the same as open). That means there exists $x \notin X$ s.t. for all $\epsilon > 0$ it holds that $B_\epsilon(x) \cap X \neq \emptyset$. This means that for all $\epsilon_n := \frac{1}{n} > 0$ there exists $x_n \in B_\epsilon(x) \cap X$. By construction we have a sequence $\{x_n\}_{n\in\mathbb{N}}$ converging to $x \notin X$ but with elements in X.

Exercise 4 (4 Points)**.** Let $X \subset \mathbb{R}^n$ open and convex and let $f : X \to \mathbb{R}$ be twice continuously differentiable. Prove the equivalence of the following statements:

- f is convex.
- For all $x \in X$ the Hessian $\nabla^2 f(x)$ is positive semidefinite ($\forall\, v \in \mathbb{R}^n : v^\top \nabla^2 f(x) v \geq 0$).

Hints: You can use that for $x, y \in X$ it holds that f is convex iff

$$(y - x)^\top \nabla f(x) \leq f(y) - f(x).$$

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

$$f(x + tv) = f(x) + tv^\top \nabla f(x) + \frac{t^2}{2} v^\top \nabla^2 f(x) v + o(t^2)$$

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

$$f(x + v) = f(x) + v^\top \nabla f(x) + \frac{1}{2} v^\top \nabla^2 f(x + tv) v$$

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

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

$$0 \overset{\text{Hint}}{\leq} f(x + tv) - f(x) - tv^\top \nabla f(x) = \frac{t^2}{2} v^\top \nabla^2 f(x) v + o(t^2)$$

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

$$0 \leq v^\top \nabla^2 f(x) v + 2 \underbrace{\frac{o(t^2)}{t^2}}_{\to 0}.$$

Let conversely $\nabla^2 f(z)$ be positive semidefinite for all $z \in X$ and let $x, y \in X$. Using the Taylor expansion we have

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

and therefore

$$f(y) - f(x) \geq (y-x)^\top \nabla f(x),$$

which means that f is convex.

Exercise 5 (4 Points)**.** Let $X \subset \mathbb{R}^n$ open and convex, $A \in \mathbb{R}^{n \times n}$ positive semidefinite, $b \in \mathbb{R}^n$, $c \in \mathbb{R}$. Show that that the quadratic form $f : X \to \mathbb{R}$ defined as

$$f(x) := \frac{1}{2} x^\top A x + b^\top x + c,$$

is convex.

Solution. To show that f is convex it suffices to show that the Hessian $\nabla^2 f(x)$ is positive semidefinite, since f is twice continuously differentiable. We start rewriting $f(x)$ in terms of finite sums:

$$\begin{aligned} f(x) &= \frac{1}{2} \sum_{i=1}^n x_i \sum_{j=1}^n a_{ij} x_j + \sum_{i=1}^n x_i b_i + c \\ &= \frac{1}{2} \sum_{i=1}^n x_i \sum_{\substack{j=1, \\ j \neq i}}^n a_{ij} x_j + \frac{1}{2} \sum_{i=1}^n a_{ii} x_i^2 + \sum_{i=1}^n x_i b_i + c \end{aligned}$$

We now proceed computing the first and second order partial derivatives:

$$\begin{aligned} \frac{\partial f(x)}{\partial x_k} &= \frac{1}{2} \sum_{\substack{j=1, \\ j \neq k}} a_{kj} x_j + \frac{1}{2} \sum_{\substack{i=1, \\ i \neq k}} a_{ik} x_i + a_{kk} x_k + b_k \\ &= \frac{1}{2} \sum_{j=1} a_{kj} x_j + \frac{1}{2} \sum_{i=1} a_{ik} x_i + b_k \end{aligned}$$

Then we have for the gradient of f:

$$\nabla f(x) = \frac{1}{2}(A + A^\top) x + b.$$

The second order derivatives are given as:

$$\frac{\partial^2 f(x)}{\partial x_k^2} = \frac{1}{2} a_{kk} + \frac{1}{2} a_{kk} = a_{kk},$$

and

$$\frac{\partial^2 f(x)}{\partial x_k \partial x_l} = \frac{1}{2} a_{kl} + \frac{1}{2} a_{lk}.$$

The Hessian is then given as

$$\nabla^2 f(x) = \frac{1}{2}(A + A^\top).$$

Since A is positive semidefinite also the Hessian $\nabla^2 f(x)$ is positive semidefinite:

$$v^\top \frac{1}{2}(A + A^\top) v = v^\top A v \geq 0.$$

Programming: Inpainting (12 Points)

Exercise 6 (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`.