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 2

Room: 02.09.023

Friday, 29.04.2016, 09:00-11:00

Submission deadline: Wednesday, 27.04.2016, 14:00, Room 02.09.023

Theory: The Subdifferential, optimality conditions and gradient descent (12+8 Points)

Exercise 1 (4 Points)**.** Let the convex function $f : \mathbb{R}^n \to \mathbb{R} \cup \{\infty\}$ be differentiable at $u \in \operatorname{int}(\operatorname{dom}(f))$. Then

$$\partial f(u) = \{\nabla f(u)\}.$$

Hint: Use the definition of the subdifferential and the directional derivative. For some direction $v \in \mathbb{R}^n$ the directional derivative $\partial_v f$ of f is given as

$$\partial_v f(x) := \lim_{\epsilon \to 0^+} \frac{f(x + \epsilon v) - f(x)}{\epsilon} = \lim_{\epsilon \to 0^-} \frac{f(x) - f(x - \epsilon v)}{\epsilon} = \langle \nabla f(x), v \rangle.$$

Definition (Karush-Kuhn-Tucker KKT conditions)**.** Let $f : \mathbb{R}^n \to \mathbb{R}$, $g : \mathbb{R}^n \to \mathbb{R}^m$ be continuously differentiable. A point $x \in \mathbb{R}^n$ satisfies the KKT-conditions if there exists a Lagrange multiplier $\lambda \in \mathbb{R}^m$ s.t.

- $0 = \nabla f(x) + \sum_{i=1}^m \lambda_i \nabla g_i(x)$
- $\lambda_i \geq 0$, $g_i(x) \leq 0$, $\lambda_i g_i(x) = 0$ for $1 \leq i \leq m$

Definition (Guignard Constraint Qualification GCQ)**.** Let $f : \mathbb{R}^n \to \mathbb{R}$, $g : \mathbb{R}^n \to \mathbb{R}^m$ be continuously differentiable and convex. Let

$$X := \{x \in \mathbb{R}^n : g_i(x) \leq 0,\ 1 \leq i \leq m\}$$

denote the feasible set and $x \in X$. Then the condition

$$\begin{aligned} N_X(x) &:= \{v \in \mathbb{R}^n : \langle v, y - x \rangle \leq 0,\ \forall\, y \in X\} \\ &= \left\{ \sum_{i \in \mathcal{A}(x)} \lambda_i \nabla g_i(x) : \lambda_i \geq 0,\ i \in \mathcal{A}(x) \right\}, \end{aligned}$$

is called GCQ. $N_X(x)$ is called the normal cone of the set X at the point $x \in X$ and $\mathcal{A}(x)$ is the set of active constraints at the point x:

$$\mathcal{A}(x) := \{i : 1 \leq i \leq m,\ g_i(x) = 0\}.$$

Definition (Slater's condition)**.** Let $f : \mathbb{R}^n \to \mathbb{R}$, $g : \mathbb{R}^n \to \mathbb{R}^m$ be continuously differentiable and convex. Let $X := \{x \in \mathbb{R}^n : g_i(x) \leq 0,\, 1 \leq i \leq m\}$ denote the feasible set. The condition

$$\exists\, x \in X \text{ s.t. } g_i(x) < 0,\, \forall\, 1 \leq i \leq m$$

is called Slater's condition.

Proposition. Slater's condition is a constraint qualification CQ, i.e. it implies GCQ.

Exercise 2 (8 Points)**.** Let $f : \mathbb{R}^n \to \mathbb{R}$, $g : \mathbb{R}^n \to \mathbb{R}^m$ be continuously differentiable and convex and let $X := \{x \in \mathbb{R}^n : g_i(x) \leq 0,\, 1 \leq i \leq m\}$ denote the feasible set. Let Slater's condition be satisfied. Show that X is convex and then prove the equivalence of the following statements:

- x solves
$$\min_{x \in \mathbb{R}^n} f(x) + \iota_X(x). \tag{1}$$
- $-\nabla f(x)$ is an element of the normal cone $N_X(x)$ of X at x.
- x satisfies the KKT-conditions.

Hint: Use the proposition stated above. Explain why Slater's condition enables you to apply the sum rule for the subdifferential.

Exercise 3 (4 points)**.** Let the function $E : \mathbb{R}^m \to \mathbb{R}$ be given as

$$E(u) := t(u) + h(u).$$

where the function $h : \mathbb{R}^m \to \mathbb{R}$ is defined as

$$h(u) := g(Du), \qquad g(v) = \sum_{i=1}^{2m} \varphi(v_i), \qquad \varphi(x) = \sqrt{x^2 + \epsilon^2},$$

with $D \in \mathbb{R}^{2m \times m}$ being a finite difference gradient operator and $t : \mathbb{R}^m \to \mathbb{R}$ is defined as

$$t(u) := \frac{\lambda}{2}\|u - f\|^2.$$

1. Show that the function E is L-smooth with $L = \lambda + \frac{\|D\|^2}{\epsilon}$.
2. Show that the function E is m-strongly convex, with $m = \lambda$.

Programming: Image denoising (12 Points)

Exercise 4 (12 Points)**.** Denoise the noisy input image f, given in the file `noisy_input.png` by minimizing the energy from Ex. 3:

$$E(u) = \frac{\lambda}{2} \|u - f\|^2 + \sum_{i=1}^{2m} \sqrt{(Du)_i^2 + \epsilon^2}$$

with gradient descent. To guarantee convergence choose your step size τ so that

$$0 < \tau \leq \frac{2}{m + L}.$$

Use MATLABs `normest` to estimate the norm $\|D\|$ of your finite difference gradient operator D.