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 3

Room: 02.09.023

Friday, 6.5.2016, 09:00-11:00

Submission deadline: Wednesday, 4.5.2016, 14:00, Room 02.09.023

## Theory: Strong convexity, Lipschitz continuity and subgradient descent (12 Points)

**Exercise 1** (4 Points)**.** Prove the following theorem: If $E \in \mathcal{S}^{1,1}_{m,L}(\mathbb{R}^n)$, then for any $u, v \in \mathbb{R}^n$ we have

$$\langle \nabla E(u) - \nabla E(v), u - v \rangle \geq \frac{mL}{m+L} \|u - v\|^2 + \frac{1}{m+L} \|\nabla E(u) - \nabla E(v)\|^2$$

**Solution.** Let $E$ be $L$-smooth and $m$-strongly convex. In case $L = m$, combining cocoercivity and strong monotonicity of $\nabla E$ directly gives:

$$\frac{1}{2}\langle \nabla E(u) - \nabla E(v), u - v \rangle \geq \frac{1}{2L} \|\nabla E(u) - \nabla E(v)\|^2, \tag{1}$$

$$\frac{1}{2}\langle \nabla E(u) - \nabla E(v), u - v \rangle \geq \frac{m}{2} \|u - v\|^2, \tag{2}$$

and adding the above yields the desired result.

Now let $m < L$ (note that it always holds that $m \leq L$). Then

$$g(u) = E(u) - \frac{m}{2} \|x\|^2$$

is convex and $(L - m)$-smooth.

Cocoercivity of $g$ then gives

$$\langle \nabla g(u) - \nabla g(v), u - v \rangle \geq \frac{1}{L-m} \|\nabla g(u) - \nabla g(v)\|^2 \tag{3}$$

$$\Leftrightarrow \langle \nabla E(u) - \nabla E(v) - m(u - v), u - v \rangle \geq \frac{1}{L-m} \|\nabla E(u) - \nabla E(v) - m(u - v)\|^2 \tag{4}$$

$$\Leftrightarrow \langle \nabla E(u) - \nabla E(v), u - v \rangle \geq m \|u - v\|^2 + \frac{1}{L-m} \|\nabla E(u) - \nabla E(v) - m(u - v)\|^2 \tag{5}$$

The right-hand side can be rewritten as

$$m\|u-v\|^2 + \frac{1}{L-m}\|\nabla E(u) - \nabla E(v) - m(u-v)\|^2 \tag{6}$$

$$= m\|u-v\|^2 + \frac{m^2}{L-m}\|u-v\|^2 + \frac{1}{L-m}\|\nabla E(u) - \nabla E(v)\|^2 - \tag{7}$$

$$\frac{2m}{L-m}\langle \nabla E(u) - \nabla E(v), u-v\rangle \tag{8}$$

Multiplying both sides with $L-m$ and rearranging and combining things yields

$$(L+m)\langle \nabla E(u) - \nabla E(v), u-v\rangle \geq \tag{9}$$

$$m(L-m)\|u-v\|^2 + m^2\|u-v\|^2 + \|\nabla E(u) - \nabla E(v)\|^2 \tag{10}$$

Dividing by $L+m$ finally gives the desired result

$$\langle \nabla E(u) - \nabla E(v), u-v\rangle \geq \frac{mL}{m+L}\|u-v\|^2 + \frac{1}{m+L}\|\nabla E(u) - \nabla E(v)\|^2. \tag{11}$$

**Exercise 2** (4 Points)**.** Compute the subdifferentials of the following convex functions:

1. $f : \mathbb{R}^n \to \mathbb{R}$ with $f(x) = \|x\|_1$.
2. $f : \mathbb{R}^n \to \mathbb{R}$ with $f(x) = \|x\|_2$.
3. $f : \mathbb{R}^{n\times m} \to \mathbb{R}$ with $f(X) = \|X\|_{2,1} := \sum_{i=1}^m \|x^i\|_2$, where $x^i \in \mathbb{R}^n$ is the $i$-th column of $X$.

**Solution.**

1. We have that $f(x) = \|x\|_1 = \sum_{i=1}^n |x_i|$. Since $\mathrm{dom}(|(\cdot)_i|) = \mathbb{R}^n$, we can apply the sum rule for the subdifferential and obtain

$$\partial f(x) = \sum_{i=1}^n |(\cdot)_i|,$$

and therefore

$$\partial f(x) = \{p \in \mathbb{R}^n : p_i \in \partial|x|,\ 1 \leq i \leq n\}.$$

It remains to compute the subdifferential $\partial|\cdot|$ of the absolute value function $|\cdot| : \mathbb{R} \to \mathbb{R}$. Since $|x| = \|x\|_2$ for $x \in \mathbb{R}$ we refer to part 2 of this exercise.

2. For $x \neq 0$ $f$ is differentiable and we have $\partial f(x) = \left\{\frac{x}{\|x\|_2}\right\}$. For $p \in \mathbb{R}^n$ with $\|p\|_2 \leq 1$ we have $f(y) - f(x) = \|y\|_2 \geq \|y\|_2 \cdot \|p\|_2 \geq \langle y, p\rangle$. Therefore $p \in \partial f(0)$. For $\|p\|_2 > 1$ and $y = p$ we have

$$f(p) - f(0) = \|p\|_2 < \|p\|_2^2 = \langle p, p\rangle.$$

Together this yields

$$\partial\|x\|_2 = \begin{cases} \frac{x}{\|x\|_2} & \text{if } x \neq 0 \\ B_1(0) & \text{if } x = 0. \end{cases}$$

3. For $f(X) := \|X\|_{2,1} = \sum_{i=1}^{m} \|x^i\|_2$ we can again apply the sum rule of the subdifferential. Together with part 2 of the exercise we get
$$\partial f(X) := \{P \in \mathbb{R}^{n\times m} : p^i \in \partial \|x^i\|_2\}.$$

**Exercise 3** (4 Points)**.** Let $f \in \mathbb{R}^n$. Show that the $\ell_1$-norm proximity operator of $f$ defined as the solution $u$ of the convex optimization problem
$$\arg\min_{u\in\mathbb{R}^n} \frac{1}{2\lambda}\|u-f\|^2 + \|u\|_1,$$
is given as
$$u \in \mathbb{R}^n, \quad u_i := \begin{cases} f_i + \lambda & \text{if } f_i < -\lambda \\ 0 & \text{if } f_i \in [-\lambda, \lambda] \\ f_i - \lambda & \text{if } f_i > \lambda. \end{cases}$$
Hint: Note that the above optimization problem is decoupled in the sense that one can look for the individual entries $u_i$ of the optimal $u$ separately.

**Solution.** We begin reformulating the optimality condition
$$0 \in \partial\left(\frac{1}{2\lambda}(u_i - f_i)^2 + |u_i|\right)$$
of the optimal $u_i$
$$0 = \frac{1}{\lambda}(u_i - f_i) + p, \quad p \in \partial|u_i| := \begin{cases} -1 & \text{if } u_i < 0 \\ [-1, 1] & \text{if } u_i = 0 \\ 1 & \text{if } u_i > 0 \end{cases}$$
$$f_i \in u_i + \begin{cases} -\lambda & \text{if } u_i < 0 \\ [-\lambda, \lambda] & \text{if } u_i = 0 \\ \lambda & \text{if } u_i > 0. \end{cases}$$
Recall that we are looking for a $u_i$ that satisfies the condition above given a fixed $f_i$. We distinguish the following cases:

1. Assume $f_i \in [-\lambda, \lambda]$. Choosing $u_i := 0$ satisfies the condition above.
2. Assume $f_i > \lambda$. Choosing $u_i := f_i - \lambda$ again satisfies the condition.
3. Assume $f_i < -\lambda$. Choosing $u_i := f_i + \lambda$ is the right choice.

# Programming: TV-$\ell_1$-denoising (12 Points)

**Exercise 4** (12 Points)**.** Denoise the noisy input image $f$, given in the file `fish_saltpepper.png` by minimizing the following robust denoising energy:
$$E(u) = \frac{\lambda}{2}\|u - f\|_1 + \|Du\|_{2,1}$$
with subgradient descent, where $D$ is a finite difference color gradient operator, and the $\ell_{2,1}$-norm is defined as on exercise sheet 0.