Convex Optimization for Machine Learning and Computer Vision

Lecture: T. Wu
Exercises: E. Laude, T. Möllenhoff
Summer Semester 2017

Computer Vision Group
Institut für Informatik
Technische Universität München

# Weekly Exercises 7

Room: 02.09.023
Monday, 26.06.2017, 12:15-14:00
Submission deadline: Wednesday, 21.06.2017, Room 02.09.023

## Majorization minimization and Convex Analysis Revisited (6 + 4 Points)

**Exercise 1** (2 Points)**.** Consider the smooth approximation of the absolute value function $f : \mathbb{R} \to \mathbb{R}$, $x \mapsto \sqrt{x^2 + \varepsilon}$ for some $\varepsilon > 0$. Show that

$$\widehat{f}(x; x_k) = f(x_k) + \frac{1}{2f(x_k)} \left[x^2 - x_k^2\right],$$

is a *majorizing surrogate* at $x_k \in \mathbb{R}$, i.e., prove that

- $\widehat{f}(x_k; x_k) = f(x_k)$,
- $\widehat{f}(x; x_k) \geq f(x)$, $\forall x \in \mathbb{R}$.

**Exercise 2** (2 Points)**.** Let $\|\cdot\|$ be any norm on $\mathbb{R}^n$ and let $\|\cdot\|_*$ denote its dual norm. Let

$$\Delta x_n := \operatorname{argmin}_{v:\|v\|=1} \nabla f(x)^\top v = \operatorname{argmin}_{v:\|v\|\leq 1} \nabla f(x)^\top v,$$

and

$$\Delta x = \|\nabla f(x)\|_* \Delta x_n,$$

be the normalized and unnormalized steepest descent directions at $x$. Prove the following identities.

- $\nabla f(x)^\top \Delta x_n = -\|\nabla f(x)\|_*$
- $\nabla f(x)^\top \Delta x = -\|\nabla f(x)\|_*^2$
- $\Delta x = \operatorname{argmin}_v (\nabla f(x)^\top v + \frac{1}{2}\|v\|^2)$

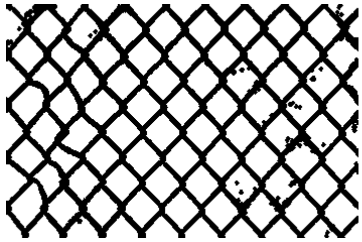
**Exercise 3** (2 Points)**.** Steepest descent method in $\ell_\infty$-norm. Explain how to find a steepest descent direction in the $\ell_\infty$-norm, and give a simple interpretation.

**Exercise 4** (4 Points)**.** Let $f : \mathbb{R}^n \to \mathbb{R} \cup \{\infty\}$. Let conv $f$ be the largest convex function majorized by $f$, meaning that $(\text{conv } f)(x) \leq f(x)$ for all $x \in \mathbb{R}^n$. Show the following identity

$$(\text{conv } f)(x) = \inf \left\{ \sum_{i=1}^{n+1} \lambda_i f(x_i) : \sum_{i=1}^{n+1} \lambda_i x_i = x, \lambda_i \geq 0, \sum_{i=1}^{n+1} \lambda_i = 1 \right\}.$$

Hint: Apply Caratheodory's Theorem to the epigraph of $f$.

Input image



Inpainting mask

Minimizer of (1)

## Programming: Image inpainting (8 Points)

**Exercise 5.** In this exercise the goal is to fill regions specified by a mask in an image by through minimizing the energy

$$E(u) = \frac{\lambda}{2} \|M(u - f)\|^2 + \|Du\|_\varepsilon . \tag{1}$$

Here, $f \in \mathbb{R}^{n_x \cdot n_y \cdot n_c}$ denotes the input image, $M : \mathbb{R}^{n_x \cdot n_y \cdot n_c} \to \mathbb{R}^{n_x \cdot n_y \cdot n_c}$ denotes a diagonal matrix consisting of zero/one values specifying the inpainting mask and $D : \mathbb{R}^{n_x \cdot n_y \cdot n_c} \to \mathbb{R}^{2 \cdot n_x \cdot n_y \cdot n_c}$ is the usual discrete gradient operator from the first exercise sheet. $\lambda > 0$ is a data fidelity parameter determining the smoothness of the solution. As in the second programming exercise, $\|x\|_\varepsilon = \sum_i \sqrt{x^2 + \varepsilon}$ denotes the smoothed $\ell_1$ norm. Your tasks are the following:

1. Find a minimizer of (1) using gradient descent.
2. Minimize (1) using a majorization minimization approach. For this, use the result from exercise 1 to majorize the term $\|\cdot\|_\varepsilon$ at the current solution $u^k$ with a quadratic function. Since the upper bound is quadratic, it can be efficiently minimized by solving a linear system. For that, use the backslash operator in MATLAB.
3. Compare the gradient descent approach to the MM scheme. Which one converges faster?

The deadline for handing in the programming solution is **June 28th, 2017, 23:59pm**.