Convex Optimization for Machine Learning and Computer Vision

Lecture: Dr. Tao Wu
Exercises: Zhenzhang Ye
Winter Semester 2019/20

Computer Vision Group
Institut für Informatik
Technische Universität München

Weekly Exercises 10

Room: 02.09.023

Wednesday, 29.01.2020, 12:15-14:00

Submission deadline: Monday, 27.01.2020, 16:15, Room 02.09.023

Primal-Dual Methods (8+6 Points)

Exercise 1 (4 Points)**.** Consider the optimization problem

$$\min_{x\in\mathbb{R}^n} \; g(x) + \sum_{i=1}^{k} f_i(K_i x), \tag{1}$$

with $g : \mathbb{R}^n \to \overline{\mathbb{R}}$, $f_i : \mathbb{R}^{m_i} \to \overline{\mathbb{R}}$ closed, proper, convex and $K_i : \mathbb{R}^n \to \mathbb{R}^{m_i}$ linear. Assume that g and all f_i are *simple* in the sense that their proximal mapping

$$\operatorname{prox}_{\tau f_i}(y) := \operatorname{argmin}_{x\in\mathbb{R}^{m_i}} \; f_i(x) + \frac{1}{2\tau}\left\|x-y\right\|^2,$$

can be efficiently computed. Explain how (1) can be solved with PDHG and write down the explicit update equations.
Hint: Stack the individual K_i into a single matrix K.

Exercise 2 (4 Points)**.** Prove that the algorithm

$$\begin{aligned} u^{k+1} &= \operatorname{prox}_{\tau G}(u^k - \tau K^* \bar{p}^k), \\ p^{k+1} &= \operatorname{prox}_{\sigma F^*}(p^k + \sigma K u^{k+1}), \\ \bar{p}^{k+1} &= 2p^{k+1} - p^k. \end{aligned} \tag{PDHG*}$$

converges, and the limit of the u^k is a minimizer of $G(u) + F(Ku)$ (with the same assumptions on F, G, and K as in the lecture).

Hint: Show that (PDHG*) is equivalent to an algorithm we discussed in the lecture applied to a reformulated problem!

Exercise 3 (6 Points)**.** Consider the *consensus optimization* problem:

$$\begin{aligned} \min_{\{x_i\}_{i=1}^l \subset \mathbb{R}^n,\, x_0\in\mathbb{R}^n} \quad & \sum_{i=1}^{l} f_i(x_i) \\ \text{subject to} \quad & x_i = x_0 \quad \forall i \in \{1,2,...,l\}. \end{aligned} \tag{2}$$

Here each function $f_i : \mathbb{R}^n \to \mathbb{R} \cup \{\infty\}$ is proper, convex, and lower-semicontinuous.

- Write down the augmented Lagrangian functional for (2) (which will involve multipliers $\{y_i\}_{i=1}^l \subset \mathbb{R}^n$).
- Formulate an alternating direction of multipliers (ADMM) method for (2). Update the variables in the order of $\{x_i\}_{i=1}^l$, $\{y_i\}_{i=1}^l$, x_0.