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

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

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.

Solution. The optimization problem (1) can be rewritten in the standard form as

$$\min_{x\in\mathbb{R}^n} \; g(x) + f(Kx), \tag{2}$$

where $K = \begin{bmatrix} K_1 \\ \vdots \\ K_k \end{bmatrix}$, and $f(z_1,\ldots,z_k) = \sum_{i=1}^k f_i(z_i)$. The PDHG updates are given by:

$$\begin{aligned} x^{t+1} &= \operatorname{prox}_{\tau g}(x^t - \tau \sum_{i=1}^{k} K_i^\top y_i^t), \\ y_i^{t+1} &= \operatorname{prox}_{\sigma f_i^*}(y_i^t + \sigma K_i(2x^{t+1} - x^t)), \text{ for } 1 \le i \le k. \end{aligned} \tag{3}$$

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!

Solution. By Fenchel's Duality Theorem, computing $\min_u G(u) + F(Ku)$ is the same as computing $-\min_p G^*(-K^*p) + F^*(p)$, where the minimizers of the first and the second problem are related via $p \in \partial F(Ku)$. By replacing $G^*(-K^*p) = \sup_u \langle -K^*p, u\rangle - G(u)$, we find

$$\begin{aligned}\min_u G(u) + F(Ku) &= -\min_p G^*(-K^*p) + F^*(p) \\ &= -\min_p \max_u F^*(p) + \langle -Ku, p\rangle - G(u)\end{aligned}$$

Applying the usual (PDHG) algorithm to the $\min_p \max_u$ in the second line yields (PDHG*) for which we established the convergence in the lecture.

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

$$\begin{aligned}\min_{\{x_i\}_{i=1}^l \subset \mathbb{R}^n,\, x_0 \in \mathbb{R}^n} \quad & \sum_{i=1}^l f_i(x_i) \\ \text{subject to} \quad & x_i = x_0 \quad \forall i \in \{1, 2, ..., l\}.\end{aligned} \tag{4}$$

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 (4) (which will involve multipliers $\{y_i\}_{i=1}^l \subset \mathbb{R}^n$).
- Formulate an alternating direction of multipliers (ADMM) method for (4). Update the variables in the order of $\{x_i\}_{i=1}^l$, $\{y_i\}_{i=1}^l$, x_0.

Solution.

- Augmented Lagrangian is defined as:

$$\mathcal{L}_\tau(x_0, \{x_i\}_{i=1}^l, \{y_i\}_{i=1}^l) = \sum_{i=1}^l \left(f_i(x_i) - \langle y_i, x_i - x_0\rangle + \frac{\tau}{2}\|x_i - x_0\|_2^2 \right)$$

with $\tau > 0$.

- ADMM can be formulated as:

$$\begin{aligned} x_i^{k+1} &= \arg\min_{x_i} f_i(x_i) - \langle y_i^k, x_i\rangle + \frac{\tau}{2}\|x_i - x_0^k\|_2^2 \\ &= (\partial f_i + \tau I)^{-1}(\tau x_0^k + y_i^k) \quad \forall i \in \{1, ..., l\}, \qquad (5)\\ y_i^{k+1} &= y_i^k - \tau(x_i^{k+1} - x_0^k) \quad \forall i \in \{1, ..., l\}, \qquad (6)\\ x_0^{k+1} &= \arg\min_{x_0} \sum_{i=1}^l \left(\langle y_i^{k+1}, x_0\rangle + \frac{\tau}{2}\|x_i^{k+1} - x_0\|_2^2 \right) \\ &= \frac{1}{l}\sum_{i=1}^l \left(x_i^{k+1} - \frac{1}{\tau} y_i^{k+1} \right). \qquad (7)\end{aligned}$$