Convex Optimization for Machine Learning and Computer Vision

Lecture: Dr. Tao Wu
Exercises: Yuesong Shen, Zhenzhang Ye
Winter Semester 2018/19

Computer Vision Group
Institut für Informatik
Technische Universität München

Weekly Exercises 11

Room: 01.09.014

Wednesday, 23.01.2019, 12:15-14:00

Submission deadline: Monday, 21.01.2019, 16:15, Room 01.09.014

Convergence Analysis (10+4 Points)

Exercise 1 (4 Points)**.** Show following properties of monotone operator:

- T is monotone, $\lambda \geq 0$. Then λT is monotone.
- R, S are monotone, $\lambda \geq 0$. Then $R + \lambda S$ is monotone.

Solution.
- T is monotone, therefore, $\langle u - v, Tu - Tv\rangle \geq 0$. Since $\lambda \geq 0$, multiplying λ on both sides doesn't change the direction of inequality. We get $\langle u - v, \lambda Tu - \lambda Tv\rangle \geq 0$. This shows that λT is a monotone operator.
- $\langle u - v, (R + \lambda S)u - (R + \lambda S)v\rangle$, we want to show it is larger or equal than 0.

$$\begin{aligned}
&\langle u - v, (R + \lambda S)u - (R + \lambda S)v\rangle \\
=&\langle u - v, Ru - Rv\rangle + \langle u - v, \lambda Su - \lambda Sv\rangle \\
\geq&\, 0
\end{aligned}$$

the inequality holds because R is monotone and λS is monotone from first conclusion.

Exercise 2 (6 Points)**.** Denote $\Pi_C(x)$ as the projection of point x onto a nonempty closed convex set C. Show following properties:

- Π_C is a monotone operator.
- T is firmly nonexpansive, if and only if $||Tx - Ty||^2 \leq \langle x - y, Tx - Ty\rangle$, $\forall x, y$.
- Π_C is firmly nonexpansive.
  Hint: you might use that $\langle y - \Pi_C(x), x - \Pi_C(x)\rangle \leq 0, \forall y \in C$

Solution.
- For any two points u and v, using the definition of projection, we have that $\|u - \Pi_C(u)\|^2 \leq \|u - \Pi_C(v)\|^2$ and $\|v - \Pi_C(v)\|^2 \leq \|v - \Pi_C(u)\|^2$. Summing them up, we have $\|u - \Pi_C(u)\|^2 + \|v - \Pi_C(v)\|^2 \leq \|u - \Pi_C(v)\|^2 + \|v - \Pi_C(u)\|^2$.
  Expanding the squares, we can get the monotonicity of Π_C.

- Recall the proposition of $\frac{1}{2}$-averaged:

$$\|(I-T)x-(I-T)y\|^2+\|Tx-Ty\|^2 \leq \|x-y\|^2$$

Write $\|(I-T)x-(I-T)Y\|^2 = \|x-y\|^2+\|Tx-Ty\|^2-2\langle x-y, Tx-Ty\rangle$, we can achieve the conclusion.

- For two points x and y. Use the hint we get $\langle \Pi_C(y)-\Pi_C(x), x-\Pi_C(x)\rangle \leq 0$ and $\langle \Pi_C(x)-\Pi_C(y), y-\Pi_C y)\rangle \leq 0$. Adding these two yields $\|\Pi_C(x)-\Pi_C(y)\|^2 \leq \langle x-y, \Pi_C(X)-\Pi_C(y)\rangle$. Use the conclusion from second problem, we get the conclusion.

Exercise 3 (4 Points)**.** Prove the theorem from the lecture:

Let C be a nonempty, closed, convex subset of $\mathbb{R}^n$. For each $i \in \{1, ..., m\}$, let $\alpha_i \in (0,1)$, $\omega_i \in (0,1)$ and $\Phi_i : C \to \mathbb{R}^n$ be an α_i-averaged operator. If $\sum_{i=1}^m \omega_i = 1$ and $\alpha = \max_{1\leq i\leq m} \alpha_i$, then

$$\Phi = \sum_{i=1}^m \omega_i \Phi_i$$

is α-averaged.

Solution. Φ_i is α_i-averaged iff

$$\|\Phi_i(u)-\Phi_i(v)\|_2^2 + \frac{1-\alpha_i}{\alpha_i}\|(I-\Phi_i)(u)-(I-\Phi_i)(v)\|_2^2 \leq \|u-v\|_2^2,$$

for all $u, v \in C$. We have the estimate

$$\begin{aligned}
&\|\Phi(u)-\Phi(v)\|_2^2 + \frac{1-\alpha}{\alpha}\|(I-\Phi)(u)-(I-\Phi)(v)\|_2^2 \\
&= \left\|\sum_{i=1}^m \omega_i(\Phi_i(u)-\Phi_i(v))\right\|_2^2 + \frac{1-\alpha}{\alpha}\left\|\left(I-\sum_{i=1}^m \omega_i\Phi_i\right)(u)-\left(I-\sum_{i=1}^m \omega_i\Phi_i\right)(v)\right\|_2^2 \\
&\leq \sum_{i=1}^m \omega_i \|\Phi_i(u)-\Phi_i(v)\|_2^2 + \frac{1-\alpha}{\alpha}\left\|\sum_{i=1}^m \omega_i((I-\Phi_i)(u)-(I-\Phi_i)(v))\right\|_2^2 \\
&\leq \sum_{i=1}^m \omega_i \|\Phi_i(u)-\Phi_i(v)\|_2^2 + \frac{1-\alpha}{\alpha}\sum_{i=1}^m \omega_i \|(I-\Phi_i)(u)-(I-\Phi_i)(v)\|_2^2 \\
&= \sum_{i=1}^m \omega_i \left(\|\Phi_i(u)-\Phi_i(v)\|_2^2 + \frac{1-\alpha}{\alpha}\|(I-\Phi_i)(u)-(I-\Phi_i)(v)\|_2^2\right).
\end{aligned}$$

Since $1 > \alpha \geq \alpha_i > 0$ for all i we have that $\frac{1}{\alpha}-1 \leq \frac{1}{\alpha_i}-1$. Then we can further bound:

$$\begin{aligned}
\ldots &= \sum_{i=1}^{m} \omega_i \left(\|\Phi_i(u) - \Phi_i(v)\|_2^2 + \frac{1-\alpha}{\alpha} \|(I - \Phi_i)(u) - (I - \Phi_i)(v)\|_2^2 \right) \\
&\le \sum_{i=1}^{m} \omega_i \left(\|\Phi_i(u) - \Phi_i(v)\|_2^2 + \frac{1-\alpha_i}{\alpha_i} \|(I - \Phi_i)(u) - (I - \Phi_i)(v)\|_2^2 \right) \\
&\le \sum_{i=1}^{m} \omega_i \|u - v\|_2^2 = \|u - v\|_2^2.
\end{aligned}$$