Convex Optimization for Machine Learning and Computer Vision

Lecture: Dr. Tao Wu
Exercises: Emanuel Laude, Zhenzhang Ye
Summer Semester 2018

Computer Vision Group
Institut für Informatik
Technische Universität München

Weekly Exercises 3

Room: 02.09.023

Wednesday, 09.05.2018, 12:15-14:00

Submission deadline: Monday, 07.05.2018, 16:15, Room 02.09.023

Subdifferential (12+6 Points)

Exercise 1 (4 Points)**.** Let the convex function $J : \mathbb{R}^n \to \mathbb{R} \cup \{\infty\}$ be differentiable at $u \in \operatorname{int}(\operatorname{dom}(J))$. Show that

$$\partial J(u) = \{\nabla J(u)\}.$$

Hint: Use the definition of the subdifferential and the directional derivative. For J being differentiable at the interior of its domain, some direction $v \in \mathbb{R}^n$ and some point $u \in \operatorname{int}(\operatorname{dom}(J))$ the directional derivative $\partial_v J$ of J is given as

$$\partial_v J(u) := \lim_{\epsilon \to 0} \frac{J(u + \epsilon v) - J(u)}{\epsilon} = \lim_{\epsilon \to 0} \frac{J(u) - J(u - \epsilon v)}{\epsilon} = \langle \nabla J(u), v \rangle.$$

Exercise 2 (8 Points)**.** Compute the subdifferential of the following functions:

- $J : \mathbb{R}^n \to \mathbb{R}, J(u) = \|u\|_1$.
- $J : \mathbb{R}^n \to \mathbb{R}, J(u) = \|u\|_\infty$.
- $J : \mathbb{E} \to \overline{\mathbb{R}}, J(u) = \delta_C(u)$ for a closed convex set $C \subset \mathbb{E}$.
- $J : \mathbb{R}^{n \times n} \to \mathbb{R}, J(X) = \sum_{i=1}^n \left(\sum_{j=1}^n (X_{i,j})^2 \right)^{1/2}$.

Exercise 3 (6 points)**.** Compute the subdifferential of nuclear norm:

$$X \in \mathbb{R}^{n \times n} \mapsto \|X\|_{nuclear} = \sum_i \sigma_i(X),$$

i.e., sum of singular values.

Hint: Show that the subdifferential at point $X \in \mathbb{R}^{n \times n}$ with $s \geq 0$ zero singular values is given as

$$\partial \|\cdot\|_{\text{nuc}}(X) = \left\{ U_1 V_1^\top + U_2 M V_2^\top : M \in \mathbb{R}^{s \times s}, \|M\|_{\text{spec}} \leq 1 \right\}, \quad (1)$$

where $U = \begin{bmatrix} U_1 & U_2 \end{bmatrix}$ and $V = \begin{bmatrix} V_1 & V_2 \end{bmatrix}$ are given by the singular value decomposition of $X = U \Sigma V^\top$, with U_1 and V_1 having $n - s$ columns. Furthermore $\|\cdot\|_{\text{spec}}$ denotes the spectral norm, i.e., the largest singular value.