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 3

Room: 02.09.023

Monday, 22.05.2017, 12:15-14:00

Submission deadline: Wednesday, 17.05.2017, Room 02.09.023

Gradient and Subdifferential (12 Points + 4 Bonus)

Exercise 1 (4 Points)**.** Let $X \subset \mathbb{R}^n$ open and convex and let $f : X \to \mathbb{R}$ be twice continuously differentiable. Prove the equivalence of the following statements:

- f is convex.
- For all $x \in X$ the Hessian $\nabla^2 f(x)$ is positive semidefinite ($\forall\, v \in \mathbb{R}^n : v^\top \nabla^2 f(x) v \geq 0$).

Hints: You can use that for $x, y \in X$ it holds that f is convex iff

$$(y - x)^\top \nabla f(x) \leq f(y) - f(x).$$

Further recall that there are two variants of the Taylor expansion:

$$f(x + tv) = f(x) + tv^\top \nabla f(x) + \frac{t^2}{2} v^\top \nabla^2 f(x) v + o(t^2)$$

with $\lim_{t \to 0} \frac{o(t^2)}{t^2} = 0$ and

$$f(x + v) = f(x) + v^\top \nabla f(x) + \frac{1}{2} v^\top \nabla^2 f(x + tv) v$$

for appropriate $t \in (0, 1)$.

Exercise 2 (2 Points)**.** Let $X \subset \mathbb{R}^n$ open and convex, $A \in \mathbb{R}^{n \times n}$ positive semidefinite, $b \in \mathbb{R}^n$, $c \in \mathbb{R}$. Show that that the quadratic form $f : X \to \mathbb{R}$ defined as

$$f(x) := \frac{1}{2} x^\top A x + b^\top x + c,$$

is convex.

Exercise 3 (2 Points)**.** Let $\mathbb{E}$ be an Euclidean space, with norm $\|\cdot\|$. Show that the subdifferential at zero is given by

$$\partial \|\cdot\| (0) = \{y \in \mathbb{E} : \|y\|_* \leq 1\},$$

where $\|\cdot\|_*$ denotes the dual norm given by

$$\|y\|_* = \sup_{\|x\| \leq 1} \langle y, x \rangle.$$

Exercise 4 (4 Points)**.** Compute the subdifferential of the following functions:

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

Exercise 5 (4 Points)**.** Consider the nuclear norm $\|\cdot\|_{\text{nuc}} : \mathbb{R}^{n\times n} \to \mathbb{R}$ given by

$$\|X\|_{\text{nuc}} = \sum_{i=1}^n |\sigma_i(X)| = \|\sigma(X)\|_1 ,$$

where $\sigma_i(X) \in \mathbb{R}$ is the i-th singular value of $X \in \mathbb{R}^{n\times n}$. Show that the subdifferential at point $X \in \mathbb{R}^{n\times n}$ with $s \geq 0$ zero singular values is given as

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

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.