Labs
Optimization for Machine Learning
Spring 2021

EPFL
School of Computer and Communication Sciences
Martin Jaggi & Nicolas Flammarion
github.com/epfml/OptML_course

Problem Set 10, due May 14, 2021 (Duality)

Prove the following property from the lecture slides:

If f is closed and convex, then for any $\mathbf{x}, \mathbf{y}$,

$$\begin{aligned}\mathbf{y} \in \partial f(\mathbf{x}) &\Leftrightarrow \mathbf{x} \in \partial f^*(\mathbf{y}) \\ &\Leftrightarrow f(\mathbf{x}) + f^*(\mathbf{y}) = \mathbf{x}^\top \mathbf{y}\end{aligned}$$

Hint: if function $f(\mathbf{x})$ is of the following form: $f(\mathbf{x}) = \max_{\alpha \in \mathcal{A}} f_\alpha(\mathbf{x})$, then its subgradient is given by

$$\partial f(\mathbf{x}) = \textbf{Co}\left[\cup \{\partial f_\alpha(\mathbf{x}) | f_\alpha(\mathbf{x}) = f(\mathbf{x})\}\right],$$

where **Co** is taking a convex hull of the set.

Solution:

- First, we will show that if $\mathbf{y} \in \partial f(\mathbf{x})$, then $\mathbf{x} \in \partial f^*(\mathbf{y})$.

  If $\mathbf{y} \in \partial f(\mathbf{x})$, then by definition of subgradient it means that $f(\mathbf{z}) \geq f(\mathbf{x}) + \mathbf{y}^\top(\mathbf{z} - \mathbf{x})\ \forall \mathbf{z}$. Reordering, we get $\mathbf{y}^\top \mathbf{z} - f(\mathbf{z}) \leq \mathbf{y}^\top \mathbf{x} - f(\mathbf{x})\ \forall \mathbf{z}$, which means that $\mathbf{x} \in \mathrm{argmax}_{\mathbf{z}}\{\mathbf{y}^\top \mathbf{z} - f(\mathbf{z})\}$.

  Taking the subgradient of the dual function $f^*(\mathbf{y}) = \max_{\mathbf{z}}\{\mathbf{y}^\top \mathbf{z} - f(\mathbf{z})\}$ (using the formula given in the exercise):

  $$\partial f^*(\mathbf{y}) = \textbf{Co}\left[\cup \left\{\mathbf{z} \;\middle|\; \mathbf{z} \in \underset{\mathbf{z}}{\mathrm{argmax}}\{\mathbf{y}^\top \mathbf{z} - f(\mathbf{z})\}\right\}\right]$$

  But since $\mathbf{x} \in \mathrm{argmax}_{\mathbf{z}}\{\mathbf{y}^\top \mathbf{z} - f(\mathbf{z})\}$, this means that $\mathbf{x} \in \partial f^*(\mathbf{y})$.

- To show that the reverse is also true (i.e. if $\mathbf{x} \in \partial f^*(\mathbf{y})$ then $\mathbf{y} \in \partial f(\mathbf{x})$), we just apply the previous result to the function f^* and use that $f^{**} = f$.

- Now we prove that $\mathbf{y} \in \partial f(\mathbf{x}) \Leftrightarrow f(\mathbf{x}) + f^*(\mathbf{y}) = \mathbf{x}^\top \mathbf{y}$.

  Proof of $\Rightarrow$: As we proved already, $\mathbf{y} \in \partial f(\mathbf{x}) \Leftrightarrow \mathbf{x} \in \arg\max_{\mathbf{z}} \mathbf{y}^\top \mathbf{z} - f(\mathbf{z})$, and we have $f^*(\mathbf{y}) := \max_{\mathbf{z}}\ \mathbf{z}^\top \mathbf{y} - f(\mathbf{z}) =\ \mathbf{x}^\top \mathbf{y} - f(\mathbf{x})$, which implies $f(\mathbf{x}) + f^*(\mathbf{y}) = \mathbf{x}^\top \mathbf{y}$.

  Proof of $\Leftarrow$: We have $f^*(\mathbf{y}) := \max_{\mathbf{z}}\ \mathbf{z}^\top \mathbf{y} - f(\mathbf{z})$ and $f^*(\mathbf{y}) =\ \mathbf{x}^\top \mathbf{y} - f(\mathbf{x})$ so we have $\max_{\mathbf{z}}\ \mathbf{z}^\top \mathbf{y} - f(\mathbf{z}) = \mathbf{x}^\top \mathbf{y} - f(\mathbf{x}) \Rightarrow \mathbf{x} \in \arg\max_{\mathbf{z}} \mathbf{y}^\top \mathbf{z} - f(\mathbf{z}) \Rightarrow \mathbf{y} \in \partial f(\mathbf{x})$