Continuous Optimization **Spring 2025**

Exercise 5: Proximal Gradient Descent

Lecturer: Aurelien Lucchi

Problem 1 (Proximal operator for quadratics):

Consider the quadratic function $f(\mathbf{x}) = \frac{1}{2}\mathbf{x}^\top \mathbf{A}\mathbf{x} + \mathbf{b}^\top \mathbf{x} + \mathbf{c}$ where $\mathbf{A} \succcurlyeq 0$.

Prove that

$$\text{prox}_{\eta f}(\mathbf{x}) = (\mathbf{I} + \eta\mathbf{A})^{-1}(\mathbf{x} - \eta\mathbf{b}) \tag{1}$$

Problem 2 (Projected Gradient Descent):

We want to minimize a function f over a convex set $\mathcal{X}$. To do so, we use projected gradient descent that, starting from $\mathbf{x}_0 \in \mathcal{X}$, performs the following updates:

$$\begin{aligned} \mathbf{y}_{k+1} &= \mathbf{x}_k - \eta \nabla f(\mathbf{x}_k) \\ \mathbf{x}_{k+1} &= \Pi_{\mathcal{X}}(\mathbf{y}_{k+1}). \end{aligned}$$

1. Prove that for all $\mathbf{x} \in \mathcal{X}, \mathbf{z} \in \mathbb{R}^d$:
$$(\mathbf{x} - \Pi_{\mathcal{X}}(\mathbf{z}))^\top (\mathbf{z} - \Pi_{\mathcal{X}}(\mathbf{z})) \leq 0.$$
2. If $\mathbf{x}_{k+1} = \mathbf{x}_k$ after the projected gradient descent update, then $\mathbf{x}_k$ is a minimizer of f over $\mathcal{X}$.

Problem 3 (Proximal Gradient Descent: Convergence analysis):

Prove the following result by adapting the proof for the convex case discussed in class.

Theorem 1 (Smooth, strongly-convex)**.** *Assume that f is β-smooth and μ strongly-convex, then*

$$\|\mathbf{x}_{k+1} - \mathbf{x}^*\|^2 \leq (1 - \eta\mu)^k \|\mathbf{x}_1 - \mathbf{x}^*\|^2.$$