Exercise 1 (from S. Boyd & J. Duchi)

For the following convex functions, explain how to calculate a subgradient at a given **x**.

- 1. $\forall \mathbf{x} \in \mathbb{R}^n : f(x) = \max_{1 \le i \le m} (\mathbf{a}_i^T \mathbf{x} + b_i)$, where $\forall i \in \{1, \ldots, m\} : (\mathbf{a}_i, b_i) \in \mathbb{R}^n \times \mathbb{R}$.
- 2. $\forall \mathbf{x} \in \mathbb{R}^n : f(x) = \max_{1 \le i \le m} |\mathbf{a}_i^T \mathbf{x} + b_i|.$
- 3. $\forall \mathbf{x} \in \mathbb{R}^n : f(\mathbf{x}) = \sup_{t \in [0,1]} p(t, \mathbf{x}), \text{ where } p(t, \mathbf{x}) = x_1 + x_2 t + \dots + x_n t^{n-1}.$

Exercise 2 (from S. Boyd & J. Duchi)

Convex functions that are not subdifferentiable. Verify that the following functions, defined on the interval $[0, +\infty)$, are convex, but not subdifferentiable at x = 0.

1.
$$f(0) = 1$$
, and $f(x) = 0$ for $x > 0$

2.
$$f(x) = -\sqrt{x}$$

Exercise 3

Theorem 14.11 in the textbook is a refined bound for Stochastic Gradient Descent (SGD) when the function f is strongly convex. The proof relies on Claim 14.10 which generalizes Lemma 13.5, i.e.,

Claim 1 If f is λ -strongly convex then for every \mathbf{w} , \mathbf{u} and $\mathbf{v} \in \partial f(\mathbf{w})$ we have

$$\langle \mathbf{w} - \mathbf{u}, \mathbf{v} \rangle \ge f(\mathbf{w}) - f(\mathbf{u}) + \frac{\lambda}{2} \|\mathbf{w} - \mathbf{u}\|^2$$

Prove this claim.

Exercise 4

Let $\pi_C(\mathbf{x}) = \arg \min_{\mathbf{y} \in C} \|\mathbf{x} - \mathbf{y}\|$ denote the Euclidean projection of x onto a closed convex set C of a Hilbert space H. Show that the projection is a 1-Lipschitz mapping, that is,

$$\left\|\pi_{\mathcal{C}}(\mathbf{x}_{0}) - \pi_{C}(\mathbf{x}_{1})\right\| \leq \left\|\mathbf{x}_{0} - \mathbf{x}_{1}\right\|,\$$

for all vectors $\mathbf{x}_0, \mathbf{x}_1 \in H$. Show that the Lipschitz constant cannot be improved. *Hint:* First prove the following important property of the projection onto a closed convex.

Lemma 1 If C is a non-empty closed convex subset of a Hilbert space H then

$$\forall (\mathbf{x}, \mathbf{y}) \in H \times C : \langle \mathbf{x} - \pi_C(\mathbf{x}), \mathbf{y} - \pi_C(\mathbf{x}) \rangle \leq 0.$$