

## Weekly Exercises 4

Room: 02.09.023

Friday, 13.5.2016, 09:00-11:00

Submission deadline: Wednesday, 11.5.2016, 14:00, Room 01.09.014

### Theory: Projected gradient descent (12+4 Points)

**Exercise 1** (4 Points). Let  $A \in \mathbb{R}^{n \times n}$  be orthonormal, meaning that  $A^T A = A A^T = I$ . Let the convex set  $C$  be given as

$$C := \{u \in \mathbb{R}^n : \|Au\|_\infty \leq 1\}.$$

Compute a formula for the projection onto  $C$  given as

$$\Pi_C(v) := \arg \min_{u \in \mathbb{R}^n} \|u - v\|_2^2, \quad \text{s.t. } u \in C.$$

Hint: Show that the  $\ell_2$ -norm of a vector is invariant under a multiplication with an orthonormal matrix  $A$ , meaning that  $\|u\|_2 = \|Au\|_2$ .

**Exercise 2** (8 Points). Let  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  be coercive, differentiable and let  $\nabla f$  be locally Lipschitz continuous i.e. for each  $x \in \mathbb{R}^n$  there exists  $\epsilon > 0$  and a constant  $L_\epsilon > 0$ , such that  $\nabla f$  is  $L_\epsilon$ -Lipschitz on  $B_\epsilon(x)$ .

1. Give an example of a function  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  that meets the above assumptions, but for which  $\nabla f$  is not Lipschitz continuous. (With a proof).
2. Let  $f$  be an arbitrary function meeting the above assumptions. Show that for any  $\alpha \geq \min_x f(x)$  there exists a constant  $L_\alpha$  such that  $f$  is  $L_\alpha$ -smooth on the sublevel set  $S_\alpha := \{x : f(x) \leq \alpha\}$ .  
Hint: You can use the topological definition of compactness: Any subset  $X \subset \mathbb{R}^n$  is called compact (closed and bounded in the Euclidean case that we consider) if each of its open covers has a finite subcover (see [https://en.wikipedia.org/wiki/Compact\\_space](https://en.wikipedia.org/wiki/Compact_space)). One possible open cover of  $X$  is the union of all  $\epsilon$ -balls with any  $\epsilon$ :  $\bigcup_{x \in X} B_\epsilon(x)$ .
3. Show that  $x^+ := x - \tau \nabla f(x) \in S_\alpha$  if  $\tau < \frac{1}{L_\alpha}$  and  $x \in S_\alpha$ .
4. Conclude that for each initialization  $x_0$ , there exists a  $\tau_{x_0}$  such that gradient descent with the constant step size  $\tau_{x_0}$  converges.

**Exercise 3** (4 Points). Let  $C_i$ ,  $1 \leq i \leq n$  be a family of closed convex sets such that

$$\bigcap_{1 \leq i \leq n} C_i \neq \emptyset.$$

Show that the problem of finding an element  $u^*$  in the intersection

$$u^* \in \bigcap_{1 \leq i \leq n} C_i$$

can be formulated as the following optimization problem:

$$u^* \in \arg \min_{u \in \bigcap_{i \in \mathcal{I}} C_i} \sum_{\substack{j \notin \mathcal{I} \\ 1 \leq j \leq n}} d^2(u, C_j),$$

where  $\mathcal{I} \subseteq \{1, 2, \dots, n\}$  can be arbitrary (including the empty set) and  $d(z, X)$  is the distance of a point  $z$  to the closed convex set  $X$  defined as

$$d(z, X) := \min_{x \in X} \|x - z\|_2.$$

## Programming: SUDOKU (12 Points)

**Exercise 4** (12 Points). Solve the SUDOKUs given in the files `exampleSudoku1.mat` and `exampleSudoku2.mat` with projected gradient descent. For that you need to find a point

$$u^* \in \bigcap_{1 \leq i \leq n+m+1} C_i$$

where the convex sets in the intersection are given as

$$C_i := \{u \in \mathbb{R}^{729} : \langle a_i, u \rangle = 1\}, \quad 1 \leq i \leq n,$$

$$C_i := \{u \in \mathbb{R}^{729} : u_j = 1, j \in \mathcal{B}\}, \quad n+1 \leq i \leq n+m,$$

and  $\mathcal{B}$  is the set of indexes corresponding to the known numbers and

$$C_{n+m+1} := \{u \in \mathbb{R}^{729} : u_j \in [0, 1], \forall 1 \leq j \leq 729\}.$$

For a more precise definition of the constraint sets see lecture.

Solve the programming assignment in the spirit of exercise 3 using the following two partitions of the indexes  $\{1, 2, \dots, n+m+1\}$ :

1.  $\mathcal{I}_1 := \{n+m+1\}$  and
2.  $\mathcal{I}_2 := \{n+1, n+2, \dots, n+m+1\}$ ,

and plot the resulting energy decays.

Hint: Show that for the linear constraint sets  $C_i$ ,  $1 \leq i \leq n$  the distance  $d(z, C_i)$  of a point  $z$  to the set  $C_i$  is equal to  $|\langle a_i, z \rangle - 1|$ .