Mathematics of Machine Learning
Homework 6 - Solutions

Instructors: Afonso S. Bandeira & Nikita Zhivotovskiy
Course Coordinator: Pedro Abdalla Teixeira

April 16, 2021

Try to solve the questions before looking to the answers. Every item must be proved rigorously. Starred problems are harder.

Problem 1

In this exercise we will only focus on $\mathbb{R}^2$. A set $B \subset \mathbb{R}^2$ is said to be convex if for all $x, y \in B$ and $\lambda \in [0, 1]$, the convex combination $\lambda x + (1 - \lambda)y$ belongs to B. Denote $\|x\|_p := (|x_1|^p + |x_2|^p)^{1/p}$ for $p > 0$ and $\|x\|_0$ denotes the number of nonzero entries of the vector.

(a) For which values of $p \geq 0$ is the unit ball of $\|.\|_p$, i.e $B_p := \{x \in \mathbb{R}^2 | \; \|x\|_p \leq 1\}$, convex?

(b) Give a justification for the term "convex relaxation" when we minimize $\|x\|_1$ instead of $\|x\|_0$.

Solution 1

(a) Observe that for $p \geq 1$, the quantity $\|x\|_p$ is the standard ℓ_p norm. In particular, triangular inequality holds, so for every $x, y \in B_p$, the convex combination $\lambda x + (1 - \lambda)y$ lies in B_p, because

$$\|\lambda x + (1 - \lambda)y\|_p \leq \lambda\|x\|_p + (1 - \lambda)\|y\|_p \leq \lambda + (1 - \lambda) = 1.$$

For $p < 1$, take $x = (1, 0)$ and $y = (0, 1)$ and observe that $\frac{x}{2} + \frac{y}{2} = (\frac{1}{2}, \frac{1}{2})$ does not belong to the ball B_p.

(b) When we minimize $\|.\|_1$ instead of $\|.\|_0$, the optimization problem becomes tractable because of the convexity of the ℓ_1 norm. The notion of relaxation is associated with the fact that we simplify the problem and the term "convex" is because now we have a convex optimization.

Problem 2

Let $A \in \mathbb{C}^{d \times N}$ be a matrix. Suppose that every s-sparse vector x can be uniquely recovered by A via $\|.\|_0$ minimization.

(a) Prove that every $2s$ columns of A are linearly independent.

(b) Prove that $d \geq 2s$.

(c) Prove that if a matrix $B \in \mathbb{C}^{d \times N}$ satisfies the condition that every $2s$ columns is linearly independent, then every s-sparse vector x can be uniquely recovered by B via $\|.\|_0$ minimization.

Solution 2

(a) We claim that the only $2s$-sparse vector in the nullspace (kernel) of A is 0. Suppose for instance the claim is true, then for a of indices $i_1, \dots, i_{2s}$ we choose a vector x with support contained in such set of indices, it follows that if $Ax = \sum_{j=1}^{2s} a_{i_l} x_l = 0$, then x is a $2s$-sparse vector in the nullspace of A, the claim implies that $x = 0$ and we conclude that the columns indexed by $\{i_1, \dots, i_{2s}\}$ are linearly independent. To prove the claim, just observe that if v is $2s$-sparse and satisfies $Av = 0$, then we split $v = x - y$ where x, y are s-sparse with disjoint supports. By uniqueness of the $\|.\|_0$ minimization, the fact that $Ax = Ay$ implies $x = y$, but the vectors have disjoint support, so $x = y = 0$.

(b) Observe that $d < 2s$ implies in the existence of a collection of at most $2s$ columns that are linearly dependent, it violates item "a".

(c) Take two s-sparse vectors x, y such that $Bx = By$. The vector $v = x - y$ is $2s$-sparse with $Bv = 0$. Pick a collection of indexes corresponding to the support of the vector v, say $\{i_1, \dots, i_{2s}\}$. Note that $Bv = \sum_{j=1}^{2s} b_{i_l} v_l = 0$ and by linear independence among the columns indexed by $\{i_1, \dots, i_{2s}\}$, we conclude that $v = 0$, so $x = y$.

Problem 3

* For a set S, we denote by $\overline{S}$ its complement and by $|S|$ its cardinality. In the lecture, we have seen conditions that guarantee recovery of sparse vectors via ℓ_1 minimization. The goal of this exercise is to prove one direction of a similar

condition for the recovery of nonnegative signals.

Let $s < d < N$. Given $A \in \mathbb{R}^{d\times N}$, prove that every nonnegative s-sparse vector $x \in \mathbb{R}^N$ is the unique solution of

$$\text{minimize}_{z\in\mathbb{R}^N} \|z\|_1 \text{ subject to } Az = Ax \text{ and } z \geq 0,$$

if the following property holds:

$$v_{\overline{S}} \geq 0 \Longrightarrow \sum_{j=1}^{N} v_j > 0,$$

for all $v \in \ker A \setminus \{0\}$ and all $S \subset [N]$ with $|S| \leq s$.

Solution 3

For every vector $y \neq x$ with nonnegative entries such that $Ay = Ax$ we have $A(y - x) = 0$, i.e, $y - x \in \ker(A)$. If S is the support of the vector x, then $|S| \leq s$ and $(y - x)_{\overline{S}} = (y)_{\overline{S}}$. By assumption $\sum_{i=1}^{N}(y - x)_i > 0$, it follows that

$$\sum_{i\in S}(y - x)_i + \sum_{i\in\overline{S}} y_i > 0.$$

Recall that S is the support of the vector x, it immediately implies that $\sum_{i\in S} x_i = \sum_{i=1}^{N} x_i$. We can conclude that

$$\sum_{i=1}^{N} x_i = \|x\|_1 < \sum_{i=1}^{N} y_i = \|y\|_1,$$

it means that x is the unique minimizer.