## Tutorial on Compressed Sensing

## Exercises

### 1. Exercise

Let $\eta \geq 0$. Show that the $\ell_0$ minimization problem

$$(P_{0,\eta}) \qquad x^{\#} = \arg\min \|z\|_0 \text{ s.t. } \|Az - y\|_2 \leq \eta$$

for general $m \times N$-matrices $A$ and $y \in \mathbb{R}^m$ is an NP-hard problem.

*Hint:* You can use the fact that the exact cover problem is NP-hard.
*Exact Cover Problem:* Given as the input a natural number $m$ divisible by 3 and a system $\{T_j : j = 1, ..., N\}$ of subsets of $\{1, ..., m\}$ with $|T_j| = 3$ for all $j \in [N]$. Decide, if there is a subsystem of mutually disjoint sets $\{T_j : j \in J\}$, $J \subset [N]$, such that $\cup_{j \in J} T_j = \{1, ..., m\}$.

### 2. Exercise

Let $q > 1$ and let $A$ be a $m \times N$-matrix with $m < N$. Show that, there is a 1-sparse vector $x$, which is not a solution of the optimization problem

$$(P_q) \qquad x^* = \arg\min \|z\|_q \text{ s.t. } Az = Ax.$$

### 3. Exercise

Let $A$ be a $m \times N$-matrix, $y \in \mathbb{R}^m$, $\eta > 0$ and let $\|\cdot\|$ an arbitrary norm on $\mathbb{R}^m$. Show that the solution of the optimization problem

$$(P_{1,\eta}) \qquad x^* = \arg\min \|z\|_1 \text{ s.t. } \|Az - y\| \leq \eta$$

is $m$-sparse in the case of the uniqueness of the solution.
*Hint:* Show that the system of columns $\{a_j : j \in \text{supp } x^*\}$ is linearly independent.

### 4. Exercise

Let $A$ be an $m \times N$ matrix and $2s \leq m$. Show that the following statements are equivalent:

i) There is a mapping $\Lambda : \mathbb{R}^m \to \mathbb{R}^N$ such that $\Lambda(Ax) = x$ for all $x \in \Sigma_s$. We call such a mapping $\Lambda$ a decoder.

ii) $\Sigma_{2s} \cap \ker(A) = \{0\}$.

iii) For any set $T$ with $\#T = 2s$, the matrix $A_T$ has rank $2s$.

iv) The symmetric non-negative matrix $A_T^t A_T$ is invertible, i.e. positive definite.

## 5. Exercise

[NSP] Given a matrix $A \in \mathbb{R}^{m \times N}$, every vector $x \in \mathbb{R}^N$ supported on a set $T$ is the unique solution of $(P_1)$ with $y = Ax$ if and only if $A$ satisfies the null space property relative to $T$.
*Reminder:* $A$ is said to satisfy the null space property relative to the set $T$ if for all $v \in \ker(A)$ holds

$$\|v_T\|_1 < \|v_{T^C}\|_1,$$

where $(v_T)_i = v_i$ if $i \in T$ and $(v_T)_i = 0$ otherwise.

## 6. Exercise

Given a matrix $A \in \mathbb{R}^{m \times N}$, a vector $x \in \mathbb{R}^N$ with support $T$ is the unique minimizer of $(P_1)$ if and only if $|\sum_{j \in T} \operatorname{sign}(x_j) v_j| < \|v_{T^C}\|_1$ for all $v \in \ker A \setminus \{0\}$.

## 7. Exercise

Show that the RIP implies the NSP.
More explicit: Let $A \in \mathbb{R}^{m,d}$ satisfy the restricted isometry porperty (RIP) of order $2s$ with constant $0 < \delta_{2s} < 1/3$, i.e.

$$(1 - \delta_{2s})\|x\|_2^2 \leq \|Ax\|_2^2 \leq (1 + \delta_{2s})\|x\|_2^2$$

holds for all $2s$-sparse vectors $x$, i.e. for all

$$x \in \Sigma_{2s} = \{v \in \mathbb{R}^d \mid \|v\|_0 = \#\{i \mid v_i \neq 0\} \leq 2s\}.$$

Show that $A$ satisfies the null space property of order $s$ (NSP), i.e. for any $T \subset [d]$ with $\#T \leq s$ and any $v \in \ker A \setminus \{0\}$ it holds

$$2\|v_T\|_1 < \|v\|_1,$$

where $(v_T)_i = v_i$ if $i \in T$ and $(v_T)_i = 0$ otherwise.
*Hint:*

1. First show that

   $$\langle Ax, Ay \rangle \leq \delta_{2s}\|x\|_2\|y\|_2$$

   if $x, y$ are $s$-sparse with disjoint support.

2. For $v \in \ker A \setminus \{0\}$ let $T_0 \subset [d]$ denote the set of indices corresponding to the $s$-largest entries of $v$ (in magnitude). Further let $T^c = T_1 \cup T_2 \cup \ldots$ be a partition of $T^c$ such that $T_1$ contains the indices of $s$-largest entries of $v_{T^c}$, $T_2$ contains the $s$-largest entries of $V_{T^c \setminus T_1}$ etc.

## 8. Exercise

Let $s \in \mathbb{N}$, $0 < \delta < 1$ and let

$$m \geq c\delta^{-2} s \log(ed/s).$$

Further let $A = \tilde{A}/\sqrt{m} \in \mathbb{R}^{m,d}$ with i.i.d. entries $\tilde{a}_{ij} \sim \mathcal{N}(0,1)$. Show that $A$ satisfies the RIP of order $s$ with RIP constant $\delta_s \leq \delta$ with probability at least

$$1 - 2\exp\left(-C\delta^2 m\right).$$

*Hint:* First use a Bernstein inequality to show that

$$\mathbb{P}\left(\left|\|Ax\|_2^2 - \|x\|_2^2\right| \geq t\|x\|_2^2\right) \leq 2\exp\left(-ct^2 m\right)$$

holds for all $t > 0$ and $x \in \mathbb{R}^d$. Then show the desired RIP inequality for a fixed $s$-dimensional subspace using a covering argument.

## Matlab Exercises

### 1. Matlab Exercise

1. Implement the basis pursuit

$$\min_{x\in\mathbb{R}^d} \|x\|_1 \quad \text{subject to} \quad Ax = y$$

   in the form of a linear optimization problem.
   *Hint:* Matlab routine `linprog` can be useful.

2. Test your program for noisy measurements of the form $y = Ax + z$, where $z \in \mathbb{R}^m$ is either deterministic noise (i.e. $\|z\|$ is small) or random Gaussian noise (i.e. $z_i \sim \mathcal{N}(0,\sigma^2)$ and $\sigma > 0$ small).

### 2. Matlab Exercise

Show numerically that the number of measurements $m$ only has to grow logarithmically in the dimension $d$ if we want to recover an $s$-sparse signal $x_0 \in \mathbb{R}^d$ from linear measurements $y = Ax$ with $A \in \mathbb{R}^{m,d}$.
To show this, calculate for increasing values of $d$ and $m$ the error of your approximation and plot the resulting matrix.

### 3. Matlab Exercise

1. Implement the 1-Bit Compressed Sensing Algorithm

$$\max_{x\in\mathbb{R}^d} \sum_{i=1}^{m} y_i\langle a_i, x\rangle \quad \text{subject to} \quad \|x\|_1 \leq R, \quad \|x\|_2 \leq 1,$$

   which recovers the true signal $x_0 \in \mathbb{R}^d$ with $\|x_0\|_1 \leq R$ and $\|x_0\|_2 \leq 1$ from measurements $y_i = \operatorname{sign}\langle a_i, x_0\rangle$, $i = 1, \ldots, m$.
   *Hint:* Matlab package `CVX` can be useful.

2. Test your algorithm with noisy measurements of the form

$$y_i = \begin{cases} \operatorname{sign}\langle a_i, x_0\rangle & \text{with probability } 1-p, \\ -\operatorname{sign}\langle a_i, x_0\rangle & \text{with probability } p \end{cases}$$

for some $0 < p < 1/2$.

## 4. Matlab Exercise

Let $f\colon B_d \to \mathbb{R}$ be a ridge function with $f(x) = g(\langle a, x\rangle)$ for some (unknown) $s$-sparse ridge vector $a \in \mathbb{R}^d$ with $\|a\|_2 = 1$ and some differentiable ridge profile $g\colon \mathbb{R} \to \mathbb{R}$ with $g'(0) \neq 0$. The ridge vector $a$ gets recovered by the following algorithm:

- *Input:* Ridge function $f(x) = g(\langle a, x\rangle)$, $h > 0$ small and $m \in \mathbb{N}$
- Take $\Phi \in \mathbb{R}^{m\times d}$ a normalized Bernoulli matrix (i.e. with entries $\pm 1$, both with probability 1/2.
- Put $\tilde{b}_j := \frac{f(h\varphi_j)-f(0)}{h}$, $j = 1, \dots, m$
- Put $\tilde{a} := \Delta_1(\tilde{b}) = \arg\min_{w\in\mathbb{R}^d} \ \|w\|_1$ s.t. $\Phi w = \tilde{b}$
- Put $\hat{a} := \frac{\tilde{a}}{\|\tilde{a}\|_2}$
- *Output:* $\hat{a}$

Implement this algorithm and show numerically that it indeed recovers the ridge vector $a$.