DM de Statistical Learning

Exercice 1: ℓ_1-Coherence and Restricted Isometries

Denote $[d] := \{1, \dots, d\}$. Let $A \in \mathbb{R}^{m\times d}$ be a matrix with ℓ_2-normalized columns $a_1, \dots, a_d$. Recall that the ℓ_1-coherence of matrix $A \in \mathbb{R}^{m\times d}$ is given by

$$\mu_1(s) := \max_{i\in[d]} \left\{ \max \left\{ \sum_{j\in S} |\langle a_i, a_j \rangle| \,, \quad S \subseteq [d]\,,\ |S| = s\,,\ i \notin S \right\} \right\}.$$

1. Prove that

$$\mu_1(s) = \max_{S\subseteq[d],|S|\le s+1} ||A_S^\top A_S - \mathrm{Id}||_{1\to 1} = \max_{S\subseteq[d],|S|\le s+1} ||A_S^\top A_S - \mathrm{Id}||_{\infty\to\infty}$$

   where $||M||_{p\to q} = \sup_{||u||_p \le 1} ||Mu||_q$.

We define the t-th restricted isometry constant δ_t by

$$\delta_t(A) := \max_{S\subseteq[d],|S|\le t} ||A_S^\top A_S - \mathrm{Id}||_{2\to 2}\,.$$

2. Prove that for all $t \ge 2$,

$$\delta_t \le \mu_1(t-1)\,.$$

3. Suppose that $\delta_s < 1$. Prove that, for any $S \subseteq [d]$ such that $|S| \le s$, $A_S^\top A_S$ is invertible and

$$\frac{1}{1+\delta_s} \le ||(A_S^\top A_S)^{-1}||_{2\to 2} \le \frac{1}{1-\delta_s} \text{ and } \frac{1}{\sqrt{1+\delta_s}} \le ||A_S^\dagger||_{2\to 2} \le \frac{1}{\sqrt{1-\delta_s}}$$

   where $A_S^\dagger := (A_S^\top A_S)^{-1} A_S^\top$ is of size $|S| \times d$.

4. Define α, β such that for all s-sparse vectors x it holds

$$\alpha ||x||_2^2 \le ||Ax||_2^2 \le \beta ||x||_2^2\,.$$

   Show that there exists $\rho > 0$ (depending only on (α, β)), such that

$$\delta_s(\rho A) \le \frac{\beta - \alpha}{\beta + \alpha}\,.$$

Exercice 2: Iterative Hard Thresholding

Given a sparsity level s, define the hard thresholding operator H_s by

$$\forall z \in \mathbb{R}^d\,, \quad H_s(z) := z_{L_s(z)} \in \mathbb{R}^d\,,$$

where

$$\forall z \in \mathbb{R}^d\,, \quad L_s(z) := \text{ index of the } s \text{ largest absolute entries of } z\,,$$

breaking ties arbitrarily. Iterative Hard Thresholding (IHT) is defined as follows.

Input: Measurement matrix $A \in \mathbb{R}^{m\times d}$, measurement vector $y \in \mathbb{R}^m$, sparsity level $s \in [d]$.
Initialization: $x^0 = 0$.
Iteration: Repeat until a stopping criterion is met at $n = \bar{n}$:

$$x^{n+1} = H_s(x^n + A^\top(y - Ax^n))$$

Output: s-sparse vector $\hat{x} = x^{\bar{n}}$.

Assume that the $3s$-th restricted isometry constant of A satisfies

$$\delta_{3s} < \frac{1}{2}\,. \qquad \text{(RIP}(3s))$$

1. Given $u, v \in \mathbb{R}^d$ and $S \subseteq [d]$, prove that

$$|\langle u, (\mathrm{Id} - A^\top A)v\rangle| \le \delta_t \|u\|_2 \|v\|_2 \qquad \text{if } |\mathrm{supp}(u) \cup \mathrm{supp}(v)| \le t\,,$$
$$\|((\mathrm{Id} - A^\top A)v)_S\|_2 \le \delta_t \|v\|_2 \qquad \text{if } |S \cup \mathrm{supp}(v)| \le t\,,$$

   where $\mathrm{supp}(u)$ denotes the support of u.

2. Assume that $y = Ax$ for some s-sparse vector x. Set $u^n := x^n + A^\top(y - Ax^n)$, where x^n is the sequence given by IHT. Prove that

$$\|u^n - x^{n+1}\|_2^2 \le \|u^n - x\|_2^2\,.$$

3. Deduce that

$$\|x^{n+1} - x\|_2^2 \le 2\langle u^n - x, x^{n+1} - x\rangle\,.$$

4. Using 1., prove that

$$\langle u^n - x, x^{n+1} - x\rangle \le \delta_{3s}\|x^n - x\|_2 \|x^{n+1} - x\|_2\,.$$

5. Prove that: Under (RIP($3s$)), if $y = Ax$ with $x \in \mathbb{R}^d$ a s-sparse vetor, then IHT sequence x^n converges to x.

6. **[Hard]** Refine the proof above (1-5) to establish the stability and robustness of s-sparse recovery via IHT when $\delta_{3s} < \frac{1}{3}$.