High-Dimensional Estimation with Constraints

Andrew Nobel

April, 2023

Basic Problem

Problem: Given a known, bounded set $K \subseteq \mathbb{R}^n$, wish to estimate an unknown vector $x \in K$ from m random measurements of the form

$$y_i = \langle x, V_i \rangle \text{ where } V_1, \ldots, V_m \text{ iid } \mathcal{N}_n(0, I)$$

Idea: Set K reflects assumed structure of the unknown vector x. Set K, observations y_i, and vector V_i are known to us

Punch line

- In many cases the number of samples m^* needed to reliably estimate x is roughly $w(K)^2 \approx$ effective dimension of K
- In particular, m^* is often independent of the ambient dimension n, or depends only logarithmically on n

Estimation from Random Projections

A. No Noise

- Generate $V_1, \ldots, V_m$ iid $\sim \mathcal{N}_n(0, I)$
- Observe $y_1 = \langle x, V_1 \rangle, \ldots, y_m = \langle x, V_m \rangle$

In matrix form, $y = Wx$ where $y = (y_1, \ldots, y_m)^t$ and $W = (V_1^t, \ldots, V_m^t)^t \in \mathbb{R}^{m \times n}$ is a Gaussian random matrix with iid $\mathcal{N}(0, 1)$ entries.

B. Bounded Noise. Assume that $y = Wx + \nu$, or equivalently, $y_i = \langle x, V_i \rangle + \nu_i$ for $1 \leq i \leq m$ where $\nu \in \mathbb{R}^m$ is a noise vector such that

$$m^{-1}||\nu||_1 = m^{-1} \sum_{i=1}^{m} |\nu_i| \leq \varepsilon$$

where ε is a *known* noise level

Estimation of Unknown $x \in K$

A. Noiseless case

- We know that $x \in K$ and $Wx = y$
- Natural estimate is any $\hat{x} \in \mathbb{R}^n$ such that $\hat{x} \in K$ and $W\hat{x} = y$

B. Bounded noise

- We know that $x \in K$ and $m^{-1}||Wx - y||_1 \leq \varepsilon$
- Natural estimate is any $\hat{x} \in \mathbb{R}^n$ such that $\hat{x} \in K$ and $m^{-1}||W\hat{x} - y||_1 \leq \varepsilon$

Note: Bounded noise case reduces to noiseless case when $\varepsilon = 0$. In each setting there is at least one solution $\hat{x}$

Geometry

Note: Feasibility conditions for estimates in noiseless case equivalent to

$$\hat{x} \in K \cap \{u : Wu = y\}$$

This is the intersection of the constraint set K and a (random) m-dimensional affine subspace of $\mathbb{R}^n$

- Error $||\hat{x} - x|| \le$ diam$(K \cap \{u : Wu = Wx\})$
- Will see: Maximum error (over all $x \in K$) is controlled by mean width

Key Theorem

Theorem: Let $L \subseteq \mathbb{R}^n$ be bounded and let $W \in \mathbb{R}^{m \times n}$ have iid $\mathcal{N}(0,1)$ entries. For $\varepsilon \geq 0$ define the random set

$$L_\varepsilon = \{u \in L : m^{-1}||Wu||_1 \leq \varepsilon\} \subseteq \mathbb{R}^n$$

Then for each fixed $\varepsilon \geq 0$ we have

$$\mathbb{E} \sup_{u \in L_\varepsilon} ||u||_2 \leq \sqrt{\frac{8\pi}{m}}\, \mathbb{E} \sup_{u \in L} |\langle u, V\rangle| + \sqrt{\frac{\pi}{2}}\varepsilon$$

Recovery from Random Projections

Recall: Observe $y = Wx + \nu \in \mathbb{R}^m$ where $W \in \mathbb{R}^{m \times n}$ is a Gaussian random matrix, $x \in K \subseteq \mathbb{R}^n$, and $\nu \in \mathbb{R}^m$ satisfies $m^{-1}||\nu||_1 \leq \varepsilon$.

Theorem: If estimate $\hat{x} \in K$ and $m^{-1}||W\hat{x} - y||_1 \leq \varepsilon$ then

$$\mathbb{E} \sup_{x \in K} ||\hat{x} - x||_2 \leq \sqrt{8\pi} \left(\frac{w(K)}{\sqrt{m}} + \varepsilon \right)$$

Note

- For $\varepsilon = 0$ bound is small when effective dimension $w(K)^2 \leq cm$
- Bound considers worst case behavior over all possible $x \in K$

Recovery: High Probability Version

Fact: Under conditions of the recovery theorem, for $t > 0$, with probability at least

$$1 - 2\exp(-mt^2/2\,\mathsf{diam}(K)^2)$$

we have

$$\sup_{x \in K} ||\hat{x} - x||_2 \le c_1 \left(\frac{w(K)}{\sqrt{m}} + \varepsilon \right) + c_2 t$$