

Sparse Linear Regression

$$Y_i = \langle \mathbf{x}_i, \boldsymbol{\beta} \rangle + w_i, \quad \text{for } i \in [n]$$

where

$$\mathbf{x}_i \stackrel{i.i.d.}{\sim} \mathcal{N}(0, 1)$$

$$w_i \stackrel{i.i.d.}{\sim} \mathcal{N}(0, \sigma^2)$$

$\boldsymbol{\beta} \in \{0, 1\}^p$ uniformly k -sparse

Goal: Exact recovery of $\boldsymbol{\beta}$ w.h.p as $p \rightarrow \infty$

Associated Detection Problem

$$\mathbb{Q} : \begin{bmatrix} \mathbf{x}_i \\ y_i \end{bmatrix} = \begin{bmatrix} \mathbf{x}_i \\ \sqrt{k/\sigma^2 + 1} w_i \end{bmatrix}$$

$$\mathbb{P} : \begin{bmatrix} \mathbf{x}_i \\ y_i \end{bmatrix} = \begin{bmatrix} \mathbf{x}_i \\ \langle \mathbf{x}_i, \boldsymbol{\beta} \rangle + w_i \end{bmatrix}$$

Goal: Construct a function f such that
 $f(\mathbf{y}) \xrightarrow{p \rightarrow \infty} 1$ under \mathbb{P} , $f(\mathbf{y}) \xrightarrow{p \rightarrow \infty} 0$ under \mathbb{Q} .

The Low-Degree Method

Neyman-Pearson Lemma (Informal):

$L(\mathbf{y}) = \frac{d\mathbb{P}}{d\mathbb{Q}}(\mathbf{y})$ is the “optimal” test statistic.

Low-Degree Conjecture [1] (Informal):

$\mathcal{P}^{\leq D} L(\mathbf{y})$ is the “optimal” computationally bounded test statistic.

If $\mathbb{E}_{\mathbb{Q}}(\mathcal{P}^{\leq D} L(\mathbf{y}))^2$ is bounded for $D \sim \log n$, then no poly-time algorithms exist.

$$\mathbb{E}_{\mathbb{Q}}(\mathcal{P}^{\leq D} L(\mathbf{y}))^2 \leq \mathbb{E}_{\beta^{(1)}, \beta^{(2)}} \exp \left(\langle \beta^{(1)}, \beta^{(2)} \rangle (1 - 2\theta) \log p \right)$$

Polynomial-time reduction to recovery

Statistical-to-Computational Gap

When $k \in o(p^\theta)$ for $\theta \in [0, 1/2]$,

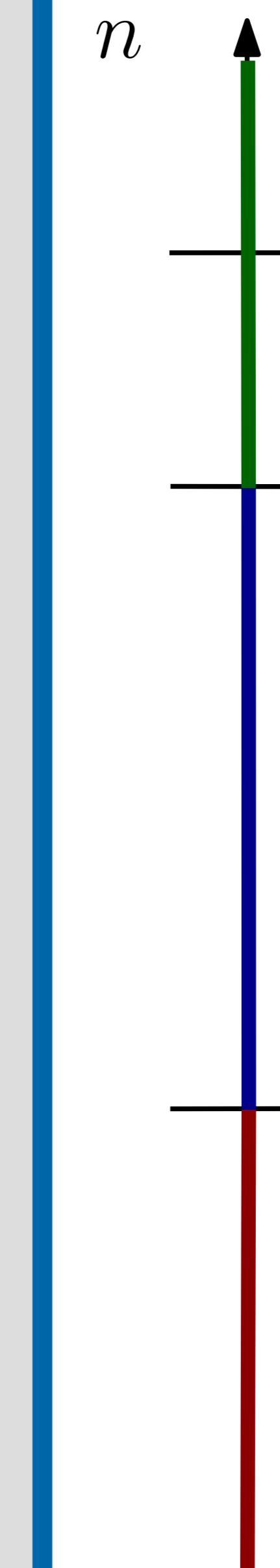
Exact Recovery

$$n_{\text{CORR}} \approx n_{\text{LASSO}} = (2k + \sigma^2) \log p$$

Approximate Recovery [2]

$$n_{\text{LD}} = (1 - 2\theta)(k + \sigma^2) \log p$$

Low-Degree Hard [3]



$$n_{\text{IT}} = \frac{2k \log p}{\log(1+k/\sigma^2)}$$

Information-Theoretically Impossible [4]

Achievability

Column Correlation Algorithm (CORR) [2]

$$\hat{\beta}_j = \mathbb{1} \left\{ \frac{\langle X_j, y \rangle}{\|y\|_2} \geq \tau \right\}$$

$$\begin{bmatrix} \times \\ \times \end{bmatrix} = \begin{bmatrix} \times \\ \times \end{bmatrix} \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} + \begin{bmatrix} & & \\ & & \end{bmatrix}$$

This project is funded by the Cambridge Trust.

[1] D. Kunisky, A. S. Wein, and A. S. Bandeira, “Notes on Computational Hardness of Hypothesis Testing: Predictions Using the Low-Degree Likelihood Ratio,” in Mathematical Analysis, its Applications and Computation, Cham, 2022, pp. 1–50. doi: 10.1007/978-3-030-97127-4_1.

[2] A. S. Bandeira, A. E. Alaoui, S. B. Hopkins, T. Schramm, A. S. Wein, and I. Zadik, “The Franz-Parisi Criterion and Computational Trade-offs in High Dimensional Statistics,” arXiv, arXiv:2205.09727, May 2022. doi: 10.48550/arXiv.2205.09727.

[3] G. Arpino, “Computational Hardness of Sparse High-Dimensional Linear Regression,” Zürich, ETH-Zürich, 2021. Accessed: Jun. 27, 2022. [1 Online-Ressource]. Available: https://gabrielarpino.github.io/files/masters_thesis.pdf

[4] W. Wang, M. J. Wainwright, and K. Ramchandran, “Information-Theoretic Limits on Sparse Signal Recovery: Dense versus Sparse Measurement Matrices,” IEEE Transactions on Information Theory, vol. 56, no. 6, pp. 2967–2979, Jun. 2010, doi: 10.1109/TIT.2010.2046199.