Non-asymptotic and Asymptotic viewpoint of LASSO

DL Theory Talk #3

December 17, 2024

Goals

- Resources of this slide: https://www.stat.berkeley.edu/~songmei/Teaching/STAT260_Spring2021/Lecture_notes/scribe_lecture1.pdf
- Get a flavor of the difference between the non-asymptotic theory and the asymptotic theory using the example of LASSO.
- Introduce mean-field theory.

LASSO: A Non-asymptotic View

Let $\beta_0 \in \mathbb{R}^d$, $X \in \mathbb{R}^{n \times d}$, $\varepsilon \in \mathbb{R}^n$ and $Y = X\beta_0 + w \in \mathbb{R}^n$. We consider the case $d \gg n$ but hope that β_0 is sparse in some sense (e.g., β_0 is k-sparse if β_0 has k non-zero elements). To recover β_0 given A and Y, we solve the following LASSO problem:

$$\hat{\beta} = \arg\min_{\beta} \frac{1}{2n} \|Y - X\beta\|_2^2 + \frac{\lambda}{n} \|\beta\|_1.$$

Restricted Strong Convexity

Definition: Restricted Strong Convexity(RSC)

We say a matrix $A \in \mathbb{R}^{n \times d}$ satisfies the restricted strong convexity property, if there exists universal constants c_1 and c_2, such that for any $v \in \mathbb{R}^d$, we have

$$\frac{\|Av\|_2^2}{n} \geq c_1 \|v\|_2^2 - c_2 \frac{\log d}{n} \|v\|_1^2. \tag{2}$$

Why this property is called restricted strong convexity? If we define $f(x) = (1/2n)\|y - Ax\|_2^2$, strong convexity property says that $\nabla^2 f(x) \succeq c_1 I_d$, so that for any direction v, we have

$$\frac{\|Av\|_2^2}{n} \geq c_1 \|v\|_2^2.$$

Restricted strong convexity simply says that f is strongly convex in the direction v when $\|v\|_1$ is small.

Non-asymptotic Result

Theorem (Negahban et al. 2012)

For any $X \in \mathbb{R}^{n \times d}$ satisfying the RSC property with constant c_1 and c_2, there exists universal constant $c < \infty$ (depending only on c_1, c_2), such that as long as $\lambda \geq 2\|X^\top w\|_\infty$, for any $\beta_0 \in \mathbb{R}^d$ and $S \subseteq [d]$ with $|S| \leq n/(c \log d)$, the LASSO estimator $\hat{\beta}$ satisfies

$$\|\hat{\beta} - \beta_0\|_2^2 \leq c\frac{\lambda^2 |S|}{n^2} + c\frac{\lambda}{n}\|\beta_{0,S^c}\|_1 + c\frac{\log d}{n}\|\beta_{0,S^c}\|_1^2.$$

Asymptotic Result

Theorem (Bayati and Montanari 2015)

We consider the asymptotic limit when $n/d \to \delta \in (0, \infty)$ as $d \to \infty$. Let $X \in \mathbb{R}^{n \times d}$ with $X_{ij} \sim \mathcal{N}(0, 1/n)$. Let $\beta_0 \in \mathbb{R}^d$ with $\beta_{0,i} \sim_{\text{iid}} \mathbb{P}_0$. Let $\sim \mathcal{N}(0, \sigma^2 I_n)$. Let $\hat{\beta}$ be the LASSO estimator. Then we have

$$\lim_{d,n\to\infty} \frac{1}{d} \|\hat{\beta} - \beta_0\|_2^2 = \mathbb{E}_{(\mathbf{X}_0, Z)\sim \mathbb{P}_0 \times \mathcal{N}(0,1)} \left[\left(\eta(\mathbf{X}_0 + \tau_* Z; \theta_*) - \mathbf{X}_0 \right)^2 \right].$$

Asymptotic Result

Remark

The asymptotic error for high-dimensional LASSO estimator is equivalent to

$$\mathbb{E}_{\hat{X},X_0}\left[(\hat{X}-X_0)^2\right],$$

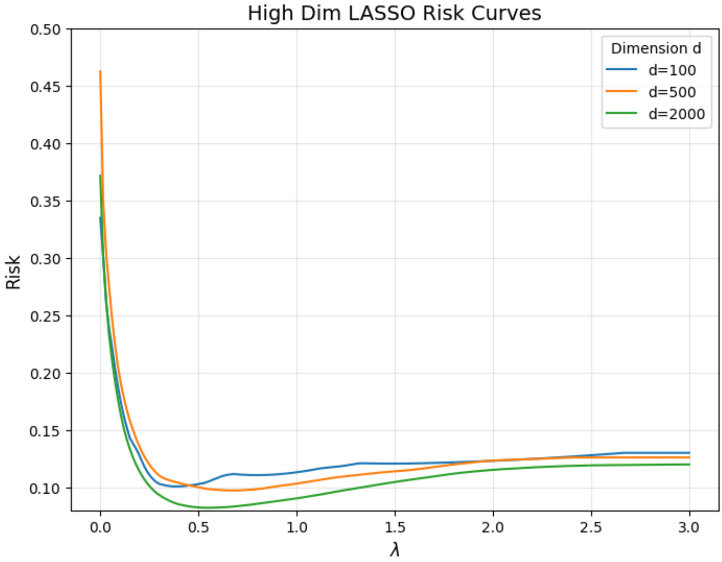
where $(\hat{X}, X_0)$ following the distribution of

$$(X_0, Z) \sim \mathbb{P}_0 \times \mathcal{N}(0,1), \qquad Y = X_0 + \tau_* Z,$$

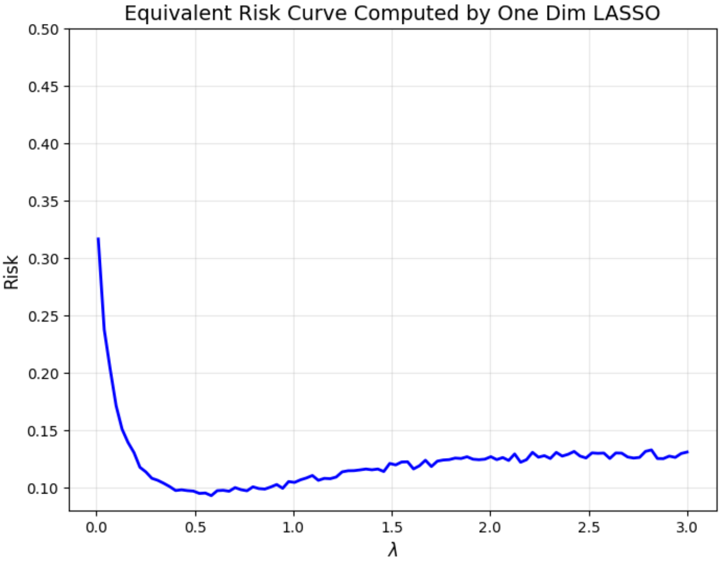
$$\hat{X} = \arg\min_v \left\{(Y-v)^2 + \tau_*\alpha_*|v|\right\} = \eta(Y, \tau_*\alpha_*).$$

This can be interpreted as an one dimensional LASSO problem.

Simulations



Simulations

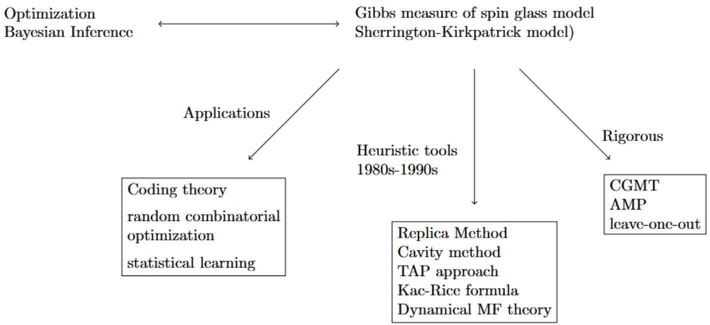


Mean-field Theory Introduction

In physics and probability theory, mean-field theory studies the behavior of high-dimensional random (stochastic) models by studying a simpler model that approximates the original by averaging over degrees of freedom.

In our example, the LASSO problem is a high dimensional random model, while the one dimensional model in the previous remark is the simpler model that approximates the original one.

Tools developed in statistical physics



References I

Bayati, Mohsen and Andrea Montanari (2015). *The LASSO risk for gaussian matrices*. arXiv: 1008.2581 [math.ST]. URL: https://arxiv.org/abs/1008.2581.

Negahban, Sahand N. et al. (Nov. 2012). "A Unified Framework for High-Dimensional Analysis of M-Estimators with Decomposable Regularizers". In: *Statistical Science* 27.4. ISSN: 0883-4237. DOI: 10.1214/12-sts400. URL: http://dx.doi.org/10.1214/12-STS400.