Week 4

Thursday, February 1, 2024 2:04 PM

Assume that $y = X\theta^* + w$

covariates $X \in \mathbb{R}^{n\times d}$, θ^* true parameter $\in \mathbb{R}^d$, w : noise vector $\in \mathbb{R}^n$

response $y \in \mathbb{R}^n$

MSE　　　　Control the spasity

Lagrangian Lasso : $\min_{\theta\in\mathbb{R}^d} \underbrace{\frac{1}{2n}\|y - X\theta\|_2^2}_{} + \underbrace{\lambda_n\|\theta\|_1}_{}$

Goal: Analyse the consistency of solution $\hat{\theta}$ of LL

i.e. $\|\hat{\theta}_n - \theta\|_2 \overset{?}{\to} 0$　　Asymptotically unbiased.

unbiased estimator : $\hat{\theta}$ for true parameter θ, if

$$E[\hat{\theta}] = \theta$$

Example: $X_1, \dots, X_n \overset{iid}{\sim} N(\theta, 1)$, $\hat{\theta} = \frac{1}{n}\sum_{i=1}^n X_i$

$$E[\hat{\theta}] = E\left[\frac{1}{n}\sum_{i=1}^n X_i\right] = \frac{1}{n}\sum_{i=1}^n E[X_i] = \frac{1}{n}\sum_{i=1}^n \theta = \theta$$

$\Rightarrow \hat{\theta}$ unbiased.

---

① Restricted nullspace. $\mathbb{C}_\alpha(S) = \{\Delta \in \mathbb{R}^d : \|\Delta_{S^c}\|_1 \le \alpha\|\Delta_S\|_1\}$

Example : for $d = 4$, $S = \{1, 2\}$, $\alpha = 3$

$\mathbb{C}_3(S) = \{(x_1, x_2, x_3, x_4) : |x_3| + |x_4| \le 3(|x_1| + |x_2|)$

$1 \xrightarrow{\alpha \ge 10} 0.1$

$1 \xrightarrow{\alpha \ge 100} 0.01$

only have values on the indexes of

Assumption:

(A1) θ^* is supported on $S \in \{1, \dots, d\}$, $|S| = s < d$

(A2) X satisfies RE condition with parameter $(K, 3)$,

$$\frac{1}{n}\|X\Delta\|_2^2 \ge K\|\Delta\|_2^2 \quad \text{for all } \Delta \in \mathbb{C}_3(S).$$

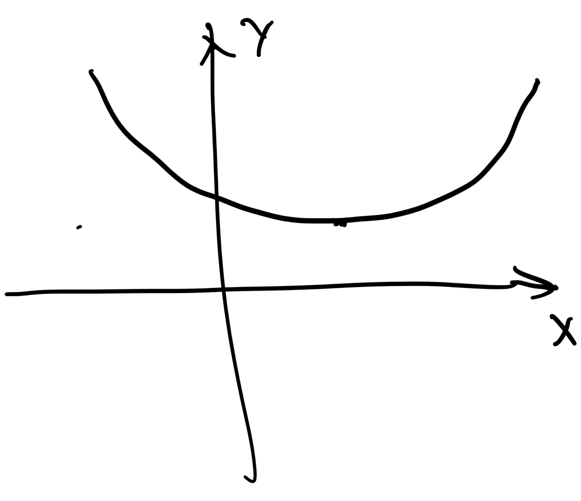
Example: For $X \in \mathbb{R}^{d\times d}$, if X positive definite,

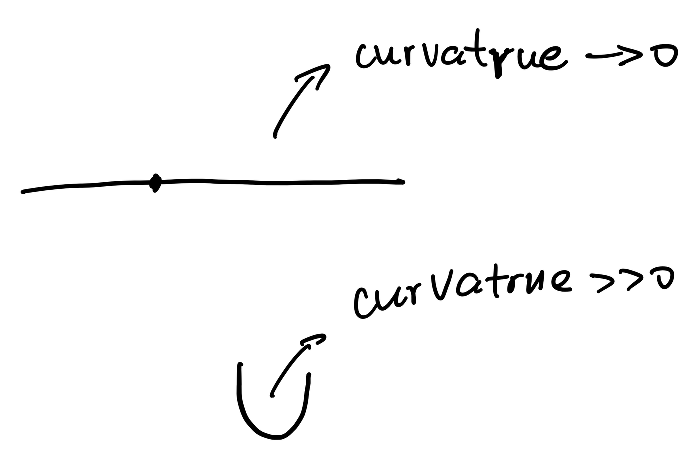
eigenvalues of X are all positive.

for $x, y \in \mathbb{R}^d$, inner product : $x^T y$.

generalized product : $x^T X y$.

$\lambda_{min}\|y\|_2^2 \le \|Xy\|_2^2 \le \lambda_{max}\|y\|_2^2$





Thm: If (A1), (A2) hold, and $\lambda_n \ge 2\left\|\frac{X^T w}{n}\right\|_\infty$ ($\in \mathbb{R}^d$),

$$\|\hat{\theta} - \theta^*\|_2^2 \le \frac{2}{K}\sqrt{s}\,\lambda_n.$$

Pf: $\min_{\theta\in\mathbb{R}^d} \frac{1}{2n}\|y - X\theta\|_2^2 + \lambda_n\|\theta\|_1 = L(\theta; \lambda_n)$

$$L(\hat{\theta}; \lambda_n) \le L(\theta^*; \lambda_n)$$

$$\frac{1}{2n}\|\underset{X\theta^* + w}{y} - X\hat{\theta}\|_2^2 + \lambda_n\|\hat{\theta}\|_1 \le \frac{1}{2n}\|\underbrace{y - X\theta^*}_{= w}\|_2^2 + \lambda_n\|\theta^*\|_1$$

$\hat{\Delta} = \hat{\theta} - \theta^*$

$$\frac{1}{2n}\|X\hat{\Delta} - w\|_2^2 + \lambda_n\|\hat{\theta}\|_1 \le \frac{1}{2n}\|w\|_2^2 + \lambda_n\|\theta^*\|_1$$

$$0 \le \frac{1}{2n}\|X\hat{\Delta}\|_2^2 \le \frac{1}{n}w^T X\hat{\Delta} + \lambda_n\{\|\theta^*\|_1 - \|\hat{\theta}\|_1\}$$

$$0 \le \frac{1}{n}\|X\hat{\Delta}\|_2^2 \le 2\frac{1}{n}\|X^T w\|_\infty\|\hat{\Delta}\|_1 + 2\lambda_n\{\|\theta^*\|_1 - \|\hat{\theta}\|_1\}$$

use condition

$\hat{\Delta} = \hat{\theta} - \theta^*$　$\le \lambda_n\{\|\hat{\Delta}\|_1 + 2(\|\theta^*\|_1 - \|\hat{\theta}\|_1)\}$

$\|\hat{\Delta} + \theta^*\|_1 = \sum_{i=1}^d |\hat{\Delta}_i + \theta_i^*|$　$\le \lambda_n\{\|\hat{\Delta}\|_1 + 2(\|\theta_S^*\|_1 - \|\hat{\Delta} + \theta^*\|_1)\}$

$= \left(\sum_{i\in S} + \sum_{i\in S^c}\right)|\hat{\Delta}_i + \theta_i^*|$　$= \lambda_n\{\|\hat{\Delta}\|_1 + 2\|\theta_S^*\|_1 - 2(\|\hat{\Delta}_S + \theta_S^*\|_1 + \|\hat{\Delta}_{S^c} + \theta_{S^c}^*\|_1)\}$

$\ge -\|\hat{\Delta}_S\|_1 + \|\theta_S^*\|_1$　　$= \|\hat{\Delta}_{S^c}\|_1$

$$\le \lambda_n\{\|\hat{\Delta}_S\|_1 + \|\hat{\Delta}_{S^c}\|_1 + 2\|\hat{\Delta}_S\|_1 - 2\|\hat{\Delta}_{S^c}\|_1\}$$

$$\le \lambda_n\{3\|\hat{\Delta}_S\|_1 - \|\hat{\Delta}_{S^c}\|_1\} \quad \Rightarrow \|\hat{\Delta}_{S^c}\|_1 \le 3\|\hat{\Delta}_S\|_1$$

$$\Rightarrow \hat{\Delta} \in C_3(S)$$

$$K\|\hat{\Delta}\|_2^2 \le \frac{1}{n}\|X^T\hat{\Delta}\|_2^2 \le \lambda_n\{3\|\hat{\Delta}_S\|_1 - \|\hat{\Delta}_{S^c}\|_1\}$$

$$\le 3\lambda_n\|\hat{\Delta}_S\|_1 \quad \left(\sum_{i=1}^s|\Delta_i|\right) \le \left(s\cdot\sum_{i=1}^s|\Delta_i|^2\right)^{\frac{1}{2}}$$

$$\le 3\lambda_n\sqrt{s}\,\|\hat{\Delta}_S\|_2$$

$$\le 3\lambda_n\sqrt{s}\,\|\hat{\Delta}\|_2$$

$\hat{\theta} - \theta^*$ ⇑

$$\|\hat{\Delta}\|_2 \le \frac{3\lambda_n\sqrt{s}}{K} \quad \square$$