Home Assignement 1

For May 11, 2017

Instructions: Students are asked to work individually (or in groups of 2 people maximum) and present their solutions clearly. They should make precise references to the results (seen in class or in the literature) invoked to solve the problem. The homework should be handed over on time, hand-written or typed on computer, with name and surname of all members of the group. Reports are expected to be well presented.

Problem 1 – Fast rates under relaxed assumptions on the design matrix (5 points)

In this problem, we use the notations of Lecture 3 on the Lasso estimator. For any $c > 0$ and $J \subset \{1, \dots, p\}$, we define the compatibility constant

$$\kappa(c, J) := \inf \left\{ \frac{c^2 \, |J| \, \|\mathbf{X}\boldsymbol{\beta}\|_2^2}{n(c\|\boldsymbol{\beta}_J\|_1 - \|\boldsymbol{\beta}_{J^c}\|_1)^2} : \boldsymbol{\beta} \in \mathbf{R}^p, \; \|\boldsymbol{\beta}_{J^c}\|_1 < c\|\boldsymbol{\beta}_J\|_1 \right\},$$

not yet studied in class, and restricted eigenvalue constant

$$\mathrm{RE}(c, J) := \inf \left\{ \frac{\|\mathbf{X}\boldsymbol{\beta}\|_2^2}{n\|\boldsymbol{\beta}_J\|_2^2} : \beta \in \mathbf{R}^p, \; \|\boldsymbol{\beta}_{J^c}\|_1 < c\|\boldsymbol{\beta}_J\|_1 \right\},$$

involved in the proof of fast rates for the Lasso.

1. Show that $\kappa(c, J) \geq \mathrm{RE}(c, J)$.

2. Reproducing only the end of the proof, providing fast rates for the Lasso, show that on the event $\{\|\mathbf{X}^\top \boldsymbol{\xi}\|_\infty \leq n\lambda/2\}$, the Lasso estimator $\hat{\boldsymbol{\beta}}$ with smoothing parameter $\lambda > 0$ satisfies, for all $n \geq 1$,

$$\frac{1}{n}\|\mathbf{X}(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}^\star)\|_2^2 \leq \inf_{\boldsymbol{\beta} \in \mathbf{R}^p, J \subset [p]} \left\{ \frac{1}{n}\|\mathbf{X}(\boldsymbol{\beta} - \boldsymbol{\beta}^\star)\|_2^2 + 4\lambda\|\boldsymbol{\beta}_{J^c}\|_1 + \frac{C\lambda^2|J|}{\kappa(3, J)} \right\},$$

for some constant $C > 0$ to be specified.

3. Show that the so-called slow rates for the Lasso estimator may be deduced directly from this result.

4. Suppose that, for some constant $\varepsilon \in (0, 1)$ and some integer $1 \leq s \leq p$, the design matrix $\mathbf{X}$ satisfies the incoherence condition

$$\|\frac{1}{n}\mathbf{X}^\top\mathbf{X} - \mathbf{I}_p\|_\infty \leq \frac{\varepsilon}{s},$$

where, for any matrix $A = (a_{i,j})$, $\|A\|_\infty$ denotes $\max_{i,j} |a_{i,j}|$. It has been shown in class that, under this condition, we get that

$$\mathrm{RE}(c, J) \geq \left(1 - (2c + 1)\frac{\varepsilon|J|}{s}\right)_+,$$

where $x_+ = \max\{0, x\}$. In the same spirit, provide a larger lower bound on $\kappa(3, J)$ depending on ε, $|J|$ and s but independent of n.

5. Suppose that the noise vector $\boldsymbol{\xi}$ is sub-gaussian with variance proxi σ^2. Suppose that $\max\{\|\mathbf{x}^j\|_2 : 1 \leq j \leq p\} \leq \sqrt{n}$. Fix $\delta \in (0, 1)$ and let

$$\lambda = 2\sigma\sqrt{\frac{2}{n}\log\left(\frac{2p}{\delta}\right)}.$$

Supposing that the incoherence condition above holds for $s \geq \|\boldsymbol{\beta}^\star\|_0$, deduce from the previous questions that the Lasso estimator $\hat{\boldsymbol{\beta}}$ with smoothing parameter λ satisfies

$$\frac{1}{n}\|\mathbf{X}(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}^\star)\|_2^2 \leq B\frac{\sigma^2\|\boldsymbol{\beta}^\star\|_0}{n}\log\left(\frac{2p}{\delta}\right),$$

with probability at least $1-\delta$, for a constant $B > 0$ independent of n given explicitly and improving on the constant provided in class.

Problem 2 – Operator norm of bounded random matrices (2 points)

Let A be an $m \times n$ random matrix with i.i.d. entries $a_{i,j}$ satisfying $\mathbf{P}(-b \leq a_{i,j} \leq b) = 1$. For any $\delta \in (0, 1)$, prove that

$$\|A\|_{\mathrm{op}} \leq b\sqrt{2\log\left(\frac{12^{m+n}}{\delta}\right)},$$

with probability larger than $1 - \delta$. The proof details should be provided and the student is free to improve the result if possible.