STAT 620: Asymptotic Statistics **Spring 2022**

Lecture: Apr 26

Lecturer: Xianyang Zhang

theorem

1 Theorem

Let $Y \sim N_p\big(\mu\,, \sigma^2\,\mathrm{I}_p\big)$ and recall from the previous lecture that the ideal risk is given by

$$R^I(\mu) = \sum_{i=1}^{p} \min\big(\mu_i^2\,, \sigma^2\big).$$

Let $\widehat{\mu}$ be hard-thresholding or soft-thresholding estimator, that is $\widehat{\mu} = \eta_S(Y, \lambda)$ or $\widehat{\mu} = \eta_H(Y, \lambda)$ with $\lambda = \sigma\sqrt{2\,\log\,p}$. Then

$$\mathbb{E}\|\widehat{\mu} - \mu\|^2 \leq \big(2\,\log\,p + \delta\big)\,\big(\sigma^2 + R^I(\mu)\big),$$

where

$$\delta = \begin{cases} 1.0 & \text{for soft-thresholding,} \\ 1.2 & \text{for hard-thresholding.} \end{cases}$$

2 Proof

We shall give the proof of the theorem for soft-thresholding estimator and the case of hard-thresholding estimator can be derived in a similar fashion.

WLOG, we let $\sigma = 1$. The risk for each coordinate is

$$r_s(\lambda, \mu) = \mathbb{E}\big(\eta_S(Y, \lambda) - \mu\big)^2\ ,\ Y \sim N(\mu, 1).$$

Claim 1

$$r_s(\lambda, \mu)\ \leq\ \min\big(r_s(\lambda, 0) + \mu^2\,,\,1 + \lambda^2\big) \leq r_s(\lambda, 0) + \min\big(\mu^2\,,\,1 + \lambda^2\big).$$

Proof of Claim 1: The second inequality follows easily. Here we shall show the steps to derive the first inequality. Note

$$\begin{aligned} r_s(\lambda, \mu) &= \int_{\lambda}^{\infty} (y - \lambda - \mu)^2\,\phi(y - \mu)\,dy\ +\ \int_{-\infty}^{-\lambda} (y + \lambda - \mu)^2\,\phi(y - \mu)\,dy\ +\ \mu^2\,\mathbb{P}(-\lambda - \mu \leq Z \leq \lambda - \mu) \\ &= \int_{\lambda-\mu}^{\infty} (u - \lambda)^2\,\phi(u)\,du\ +\ \int_{-\infty}^{-\lambda-\mu} (u + \lambda)^2\,\phi(u)\,du\ +\ \mu^2\,\mathbb{P}(-\lambda - \mu \leq Z \leq \lambda - \mu) \end{aligned}$$

where $u = y - \mu$ and $Z \sim N(0, 1)$. Thus we get

$$0 \leq \frac{\partial r_s(\lambda, \mu)}{\partial \mu}\ =\ 2\,\mu\,\mathbb{P}(-\lambda - \mu \leq Z \leq \lambda - \mu)\ \leq\ 2\,\mu$$

which implies that

$$r_s(\lambda, \mu)\ \leq\ \lim_{\mu\to\infty} r_s(\lambda, \mu)\ =\ 1 + \lambda^2.$$

Also, note that

$$r_s(\lambda,\mu) - r_s(\lambda,0) = \int_0^{\mu} \frac{\partial r_s(\lambda,u)}{du}\, du \;\le\; \int_0^{\mu} 2\,u\,du \;=\; \mu^2$$

which indicates that $r_s(\lambda,\mu) \;\le\; r_s(\lambda,0) + \mu^2$.

Claim 2:

$$r_s(\lambda,0) \;\le\; \frac{2\,\phi(\lambda)}{\lambda}.$$

Proof of Claim 2:

$$\begin{aligned} r_s(\lambda,0) &= 2\int_{\lambda}^{\infty} (y-\lambda)^2\,\phi(y)\,dy \\ &= 2\lambda^2\,\mathbb{P}(Z>\lambda) + 2\int_{\lambda}^{\infty} y^2\,\phi(y)\,dy - 4\,\lambda\int_{\lambda}^{\infty} y\,\phi(y)\,dy. \end{aligned}$$

As $\phi'(y) = -y\,\phi(y)$, we have

$$\lambda\int_{\lambda}^{\infty} y\,\phi(y)\,dy = -\lambda\int_{\lambda}^{\infty} d\phi(y) \;= \lambda\,\phi(\lambda).$$

Moreover using integration by parts, we have

$$\begin{aligned} \int_{\lambda}^{\infty} y^2\,\phi(y)\,dy &= -\int_{\lambda}^{\infty} y\,d\phi(y) \\ &= \lambda\,\phi(\lambda) + \int_{\lambda}^{\infty} \phi(y)\,dy \\ &= \lambda\,\phi(\lambda) + \mathbb{P}(Z>\lambda). \end{aligned}$$

Also,

$$\mathbb{P}(Z>\lambda) = \mathbb{E}\Big[\mathbf{1}\{Z>\lambda\}\Big] \le \mathbb{E}\Big[\frac{Z}{\lambda}\mathbf{1}\{Z>\lambda\}\Big] = \frac{1}{\lambda}\int_{\lambda}^{\infty} y\,\phi(y)\,dy = \frac{\phi(\lambda)}{\lambda}.$$

Then,

$$r_s(\lambda,0) = 2\,(1+\lambda^2)\,\mathbb{P}(Z>\lambda) - 2\,\lambda\,\phi(\lambda) \;\le\; 2\,(1+\lambda^2)\,\frac{\phi(\lambda)}{\lambda} - 2\,\lambda\,\phi(\lambda) \;=\; \frac{2\,\phi(\lambda)}{\lambda}.$$

2.1 Back to main proof

Putting $\lambda = \sqrt{2\,\log\,p}$, we have:

$$r_s(\lambda,0) \le 2\,\frac{1}{\sqrt{2\,\log\,p}}\,\frac{1}{\sqrt{2\,\pi}}\,\frac{1}{p} = \frac{1}{p\,\sqrt{\pi\,\log\,p}} \le \frac{2\,\log\,p+1}{p}.$$

Now,

$$\begin{aligned} \mathbb{E}\|\widehat{\mu}-\mu\|^2 &= \sum_{i=1}^{p} \mathbb{E}\big(\widehat{\mu}_i - \mu_i\big)^2 \\ &\le p\,r_s(\lambda,0) + \sum_{i=1}^{p} \min\big(\mu_i^2, \lambda^2+1\big) \\ &\le \big(2\,\log\,p+1\big) + \sum_{i=1}^{p} \min\big(\mu_i^2, 2\,\log\,p+1\big) \\ &\le \big(2\,\log\,p+1\big)\,\big(1+R^I(\mu)\big). \end{aligned}$$

3 Risk inflation

Foster and George (1994) proved that

$$\inf_{\widehat{\mu}} \sup_{\mu} \frac{R(\mu, \widehat{\mu})}{\sigma^2 + R^I(\mu)} \geq 2 \log p \left(1 + o(1)\right)$$

where the infimum is over all hard-thresholding estimators:

$$\widehat{\mu}_i^{HT} = \begin{cases} Y_i & \text{if } |Y_i| \geq \lambda, \\ 0 & \text{if } |Y_i| < \lambda. \end{cases}$$

Assuming $\sigma = 1$, then for k-sparse μ, we have

$$R^I(\mu) = \sum_{i=1}^{p} \min\left(\mu_i^2\,,\, 1\right) \leq k \implies \frac{1}{1+k} \leq \frac{1}{1+R^I(\mu)}.$$

It is enough to show that $\forall \lambda \geq 0$,

$$RI(\lambda) = \max_k \sup_{\|\mu\|_0 = k} \frac{R(\mu, \widehat{\mu})}{1+k} \geq 2 \log p \left(1 + o(1)\right).$$

To this end, we note that

$$\begin{aligned} r_H(\lambda, \mu) &= \mathbb{E}\Big[(Y-\mu)^2\, \mathbf{1}\{|Y| \geq \lambda\} + \mu^2\, \mathbf{1}\{|Y| < \lambda\}\Big] \\ &= \mathbb{E}\Big[Z^2\, \mathbf{1}\{|Z+\mu| \geq \lambda\} + \mu^2\, \mathbf{1}\{|Z+\mu| < \lambda\}\Big] \\ &= \mathbb{E}\big[Z^2\, \mathbf{1}\{|Z+\mu| \geq \lambda\}\big] + \mu^2\, \mathbb{P}\big(|Z+\mu| < \lambda\big) \\ &\geq \mu^2 \mathbb{P}\big(\mu + Z \leq \lambda\big) \end{aligned}$$

where $Z \sim N(0,1)$ and $Y \sim N(\mu, 1)$. For k-sparse μ, we let

$$f(k) = \sup_{\|\mu\|_0 = k} R\big(\mu, \widehat{\mu}\big) = (p-k)\, r_H(\lambda, 0) + k \sup_{\mu} r_H(\lambda, \mu).$$

Then we have

$$RI(\lambda) = \max_{0 \leq k \leq p} \frac{f(k)}{1+k} \geq \max\left(f(0)\,,\, \frac{f(p)}{1+p}\right) = \max\left(p\, r_H(\lambda, 0)\,,\, \frac{p}{1+p} \sup_{\mu} r_H(\lambda, \mu)\right).$$

Now,

$$r_H(\lambda, 0) = \mathbb{E}\Big[Z^2\, \mathbf{1}\{|Z| > \lambda\}\Big] = 2\, \mathbb{E}\Big[Z^2\, \mathbf{1}\{Z > \lambda\}\Big] = 2 \int_{\lambda}^{\infty} y^2\, \phi(y)\, dy \approx 2\, \lambda\, \phi(\lambda),$$

where $\lambda = \sqrt{2 \log p}$. Recall from the above that

$$\sup_{\mu} r_H(\lambda, \mu) \geq \sup_{\mu} \mu^2\, \mathbb{P}\big(\mu + Z \leq \lambda\big).$$

Then

$$\begin{aligned} \sup_{\mu} r_H(\lambda, \mu) &\geq \sup_{\mu} \mu^2\, \mathbb{P}\big(\mu + Z \leq \lambda\big) \\ &= \sup_{\mu} \mu^2\, \Phi\big(\lambda - \mu\big) \\ &\geq \sup_{0 \leq u \leq \lambda} (\lambda - u)^2\, \Phi(u)\,, \quad \text{where } u = \lambda - \mu \\ &= \sup_{0 \leq u \leq \lambda} \Big[\lambda^2\, \Phi(u) + u^2\, \Phi(u) - 2\, \lambda\, u\, \Phi(u)\Big] \\ &\geq \lambda^2 - 4\, \lambda \sqrt{\log \lambda} + o\big(\lambda \sqrt{\log \lambda}\big)\,, \end{aligned}$$

where we set $u = \sqrt{2\,\log\,\lambda^2}$ to get the last inequality. Finally, we get

$$RI(\lambda) \gtrsim \max\left(2\,p\,\lambda\,\phi(\lambda)\,,\,\lambda^2\right) \sim 2\,\log\,p\,(1 + o(1)).$$