Stat461 - 2023 Fall

HW3

Tuorui Peng[1]

Exercise 1

1.(a)

We consider the following:

$$\begin{aligned}\frac{1}{d}\left\|X\Delta\right\|_2^2 =& \left\|X(\Delta_S+\Delta_{S^\complement})\right\|_2^2\\ =& \underbrace{\Delta_S'\frac{X'X}{n}\Delta_S}_{\mathrm{I}}+\underbrace{\Delta_{S^\complement}'\frac{X'X}{n}\Delta_{S^\complement}}_{\mathrm{II}}+\underbrace{2\Delta_S'\frac{X'X}{n}\Delta_{S^\complement}}_{\mathrm{III}}\end{aligned}$$

in which each term can be bounded as follows:

I: Using the bound on $\left\|\dfrac{X'X}{n}-I\right\|_\infty$ we have

$$\begin{aligned}\Delta_S'\frac{X'X}{n}\Delta_S =& \Delta_S'\Delta_S+\Delta_S'\Big(\frac{X'X}{n}-I\Big)\Delta_S\\ \geq& \Delta_S'\Delta_S-\left|\Delta_S\right|_1^2\left\|\frac{X'X}{n}-I\right\|_\infty\\ \geq& \left\|\Delta_S\right\|_2^2-\frac{1}{32k}\left|\Delta_S\right|_1^2\end{aligned}$$

II: Similarly, we have

$$\begin{aligned}\Delta_{S^\complement}'\frac{X'X}{n}\Delta_{S^\complement} \geq& \left\|\Delta_{S^\complement}\right\|_2^2-\frac{1}{32k}\left|\Delta_{S^\complement}\right|_1^2\\ \geq& \left\|\Delta_{S^\complement}\right\|_2^2-\frac{9}{32k}\left|\Delta_S\right|_1^2\end{aligned}$$

III: Note that we have $\Delta_S'\Delta_{S^\complement}=0$,

$$\begin{aligned}2\Delta_S'\frac{X'X}{n}\Delta_{S^\complement} =& 2\Delta_S'\Big(\frac{X'X}{n}-I\Big)\Delta_{S^\complement}\\ \geq& -\frac{1}{32k}\left|\Delta_S\right|_1\left\|\Delta_{S^\complement}\right\|_1\\ \geq& -\frac{3}{32k}\left|\Delta_S\right|_1^2\end{aligned}$$

Together, we have

$$\begin{aligned}\frac{1}{n}\left\|X\Delta\frac{1}{n}\right\|_2^2 \geq& \left\|\Delta_S\right\|_2^2-\frac{1}{32k}\left|\Delta_S\right|_1^2+\left\|\Delta_{S^\complement}\right\|_2^2-\frac{9}{32k}\left|\Delta_S\right|_1^2-\frac{3}{32k}\left|\Delta_S\right|_1^2\\ =& \left\|\Delta\right\|_2^2-\frac{1}{2k}\left\|\Delta_S\right\|_1^2\end{aligned}$$

---

[1] TuoruiPeng2028@u.northwestern.edu

Using Cauchy-Schwarz inequality, we have $\|\Delta_S\|_1 \leq |S| \left\|^{1/2}\Delta_S\right\|_2$, and thus

$$\begin{aligned}\frac{1}{n}\left\|X\Delta\frac{1}{n}\right\|_2^2 \geq& \|\Delta\|_2^2 - \frac{1}{2}\|\Delta_S\|_2^2 \\ \geq& \frac{1}{2}\|\Delta\|_2^2\end{aligned}$$

1.(b)

Using Gershgorin circle theorem, we have

$$\gamma(\frac{X_S'X_S}{n}) \in [1-\frac{|S|}{32k}, 1+\frac{|S|}{32k}] \subset [1-\frac{1}{32}, 1+\frac{1}{32}]$$

thus we have

$$\gamma(\frac{X_S'X_S}{n})^{-1} \in [\frac{32}{33}, \frac{32}{31}]$$

We have for any $j \in S^{\complement}$:

$$\begin{aligned}\left\|(X_S'X_S)^{-1}X_SX_j\right\|_1 \leq& \left\|(X_S'X_S)^{-1}\right\|_2 \|X_SX_j\|_\infty \\ \leq& \frac{32}{33}\cdot\frac{1}{32k} = \frac{1}{33k} < 1\end{aligned}$$

1.(c)

Denote $X = \{\varepsilon_{li}\}_{l\in[n],k\in[d]}$. We have for any $i, j \in [d]$:

$$\frac{X_i'X_j}{n} - \delta_{ij} = \frac{1}{n}\sum_{l=1}^n \varepsilon_{li}\varepsilon_{lj} - \delta_{ij}$$

being a zero-mean sub-Gaussian random variable with some parameter σ^2/n, we using the bound on the maximum of sub-Gaussian random variables, we have

$$\mathbb{E}\left[\max_{i,j}\left|\frac{X_i'X_j}{n} - \delta_{ij}\right|\right] \lesssim \sigma\sqrt{\log d/n}$$

now it suffices to bound

$$\max_{i,j}\left|\frac{X_i'X_j}{n} - \delta_{ij}\right| - \mathbb{E}\left[\max_{i,j}\left|\frac{X_i'X_j}{n} - \delta_{ij}\right|\right]$$

Note that function $\{\varepsilon_{ij}\} \mapsto \max_{i,j}\left|\frac{X_i'X_j}{n} - \delta_{ij}\right|$ is $2/n$-Lipschitz, thus we have concentration inequality:

$$\left|\max_{i,j}\left|\frac{X_i'X_j}{n} - \delta_{ij}\right| - \mathbb{E}\left[\max_{i,j}\left|\frac{X_i'X_j}{n} - \delta_{ij}\right|\right]\right| \leq t \quad w.p. \geq 1 - 2\exp(-nt^2/2)$$

set $2\exp(-nt^2/2) = \delta$, we have w.p. at least $1-\delta$:

$$\left|\max_{i,j}\left|\frac{X_i'X_j}{n} - \delta_{ij}\right| \leq \mathbb{E}\left[\max_{i,j}\left|\frac{X_i'X_j}{n} - \delta_{ij}\right|\right]\right| + \sqrt{\frac{2\log 2/\delta}{n}} \leq \frac{1}{32k}$$

i.e.

$$C\sqrt{\frac{\log d}{n}} + \sqrt{\frac{2\log 2/\delta}{n}} \leq \frac{1}{32k}$$

To satisfy the above inequality, we need

$$n \geq 2^{10}k^2(C\sqrt{\log d} + \sqrt{2\log 2/\delta}) \gtrsim Ck^2(\log d + \log\frac{1}{\delta})$$

Exercise 2 MJW 7.11

2.(a)

We have for $\theta \in \mathbb{C}_\alpha(S)$:

$$\begin{aligned}\|\theta\|_1^2 =&(\|\theta_S\|_1 + \left\|\theta_{S^\complement}\right\|_1)^2\\ \leq&(1+\alpha)^2\|\theta_S\|_1^2\\ \leq&(1+\alpha)^2|S|\,\|\theta\|_2^2\end{aligned}$$

Thus using bound (**MJW 7.13**) we have

$$\begin{aligned}\frac{\|X\theta\|_2^2}{n} \geq c_1\left\|\sqrt{\Sigma}\theta\right\|_2^2 - c_2\rho^2(\Sigma)\frac{\log d}{n}\|\theta\|_1^2 \geq& c_1\gamma_{\min}(\Sigma)\|\theta\|_2^2 - c_2\rho^2(\Sigma)\frac{\log d}{n}(1+\alpha)^2|S|\,\|\theta\|_2^2\\ =&\left(c_1\gamma_{\min}(\Sigma) - c_2\rho^2(\Sigma)\frac{\log d}{n}(1+\alpha)^2|S|\right)\|\theta\|_2^2\\ =&\frac{c_1}{2}\gamma_{\min}(\Sigma)\|\theta\|_2^2\end{aligned}$$

which requires

$$|S| \leq \frac{c_1}{2c_2}\frac{\gamma_{\min}(\Sigma)}{\rho^2(\Sigma)}(1+\alpha)^{-2}\frac{n}{\log d}$$

2.(b)

We can just choose $\Sigma^{(d)} = dI$, for which we have

$$\gamma_{\max} \to \infty, \qquad \frac{\gamma_{\min}}{\rho^2} = 1$$

Exercise 3 MJW 7.17

3.(a)

The solution is given by

$$\theta = \arg\min_{\theta\in\mathbb{R}^d}\frac{1}{\sqrt{n}}\|y - X\theta\|_2 + \gamma_n\|\theta\|_1$$

which has equivalent form

$$\begin{aligned}&\Leftrightarrow \arg\min_{\theta\in\mathbb{R}^d}\frac{1}{\sqrt{n}}\|y - X\theta\|_2, \quad \text{s.t. } \|\theta\|_1 \leq R_n\\ &\Leftrightarrow \arg\min_{\theta\in\mathbb{R}^d}\frac{1}{2n}\|y - X\theta\|_2^2, \quad \text{s.t. } \|\theta\|_1 \leq R_n\\ &\Leftrightarrow \arg\min_{\theta\in\mathbb{R}^d}\frac{1}{2n}\|y - X\theta\|_2^2 + \lambda_n\|\theta\|_1\end{aligned}$$

which is the Lagrange form of LASSO problem.

3.(b)

The result is trivial by taking sub-differential of the objective function.

$$\frac{\mathrm{d}}{\mathrm{d}\theta}\frac{1}{\sqrt{n}}\left\|y-X\theta\right\|_2=\frac{\frac{1}{n}X'(X\theta-y)}{\frac{1}{\sqrt{n}}\left\|y-X\theta\right\|_2}$$

3.(c)

Using the previous result, we have

$$\begin{aligned}
&\frac{\frac{1}{n}X'(X\theta-y)}{\frac{1}{\sqrt{n}}\left\|y-X\theta\right\|_2}+\gamma_n\hat{z}=0\\
\Rightarrow&\left\langle\frac{\hat{\Delta}}{\sqrt{n}},\frac{1}{\sqrt{n}}X'(X\theta-X\theta^*-w)\right\rangle+\gamma_n\frac{\left\|X\theta^*+w-X\theta\right\|_2}{\sqrt{n}}\left\langle\hat{\Delta},\hat{z}\right\rangle=0
\end{aligned}$$

in which we use $y=X\theta^*+w$. Then if θ^* is $|S|$-sparse on set S, we have

$$\begin{aligned}
\frac{\left\|X\hat{\Delta}\right\|_2^2}{n}&=\left\langle\frac{\hat{\Delta}}{\sqrt{n}},\frac{1}{\sqrt{n}}X'X\hat{\Delta}\right\rangle\\
&=\left\langle\hat{\Delta},\frac{1}{n}X'w\right\rangle+\gamma_n\frac{\left\|y-X\hat{\theta}\right\|_2}{\sqrt{n}}\left\langle\hat{\Delta},\hat{z}\right\rangle\\
&\leq\left\langle\hat{\Delta},\frac{1}{n}X'w\right\rangle+\gamma_n\frac{\left\|y-X\hat{\theta}\right\|_2}{\sqrt{n}}\left[\left\|\hat{\Delta}_S\right\|_1-\left\|\hat{\Delta}_{S^\complement}\right\|_1\right]
\end{aligned}$$

3.(d)

We have

$$\begin{aligned}
0\leq\frac{\left\|X\hat{\Delta}\right\|_2^2}{n}\leq&\left\langle\hat{\Delta},\frac{1}{n}X'w\right\rangle+\gamma_n\frac{\left\|y-X\hat{\theta}\right\|_2}{\sqrt{n}}\left[\left\|\hat{\Delta}_S\right\|_1-\left\|\hat{\Delta}_{S^\complement}\right\|_1\right]\\
\leq&\frac{\gamma_n}{2\sqrt{n}}\left(\left\|\hat{\Delta}_S\right\|_1+\left\|\hat{\Delta}_{S^\complement}\right\|_1\right)\left\|w\right\|_2+\frac{\gamma_n}{\sqrt{n}}\left\|y-X\hat{\theta}\right\|_2\left(\left\|\hat{\Delta}_S\right\|_1-\left\|\hat{\Delta}_{S^\complement}\right\|_1\right)\\
\leq&\frac{\gamma_n\left\|w\right\|_2}{2\sqrt{n}}\left(3\left\|\hat{\Delta}_S\right\|_1-\left\|\hat{\Delta}_{S^\complement}\right\|_1\right)\\
\Rightarrow&3\left\|\hat{\Delta}_S\right\|_1\geq\left\|\hat{\Delta}_{S^\complement}\right\|_1
\end{aligned}$$

3.(e)

Using the fact that $\|\hat{w}\|_2 \leq \|w\|_2$ we have:

$$\begin{aligned}
s\gamma_n^2 \left\|\hat{\Delta}\right\|_2 \leq& \kappa \left\|\hat{\Delta}\right\|_2^2 \\
\leq& \frac{1}{n}\left\|X\hat{\Delta}\right\|_2^2 \\
\leq& \frac{2w'X\hat{\Delta}}{n} \\
\leq& \frac{2}{n}\left\|X\hat{\Delta}\right\|_\infty \left\|\hat{\Delta}\right\|_1 \\
\leq& \frac{\gamma_n}{\sqrt{n}}\|w\|_2 \sqrt{s}\left\|\hat{\Delta}\right\|_2
\end{aligned}$$

rearrenging the terms, we have

$$\left\|\hat{\Delta}\right\|_2 \leq \frac{\|w\|_2}{\sqrt{n}}\frac{1}{\gamma_n\sqrt{s}}$$

assuming $\kappa = 1$ could yield the result in MJW.

Exercise 4 MJW 13.4

4.(a)

We have

$$\begin{aligned}
\text{star-shaped around } x^* \Leftrightarrow& \alpha(x - x^*) \in C - x^* \quad \forall x \in C, \forall \alpha \in [0,1] \\
\Leftrightarrow& \alpha x + (1-\alpha)x^* \in C \quad \forall x \in C, \forall \alpha \in [0,1]
\end{aligned}$$

4.(b)

$\Rightarrow$ If C is convex, then for any given $x^* \in C$ we have

$$\alpha x + (1-\alpha)x^* \in C \quad \forall x \in C, \forall \alpha \in [0,1] \Rightarrow \text{star-shaped around } x^*$$

$\Leftarrow$ If C is star-shaped around and $x^* \in C$, then for any $x, y \in C$ and $\alpha \in [0,1]$, we have

$$\alpha x + (1-\alpha)y \in C \Rightarrow C \text{ is convex.}$$