Error in GLSNet proof

Jiaxin Hu

June 12, 2021

1 Sub-Gaussian Case

By the definition, $\boldsymbol{Y}_{kl}^{(ij)} = \mathbb{E}[A_{jli}] - A_{jli}$ for $k = j$ and is equal to 0 otherwise. Since $\boldsymbol{Y}^{(ij)}$ is rank 1, we have $\left\|\boldsymbol{Y}^{(ij)}\right\|_2 = \left\|\boldsymbol{Y}^{(ij)}\right\|_F$. Then,

$$\left\|(\boldsymbol{Y}^{(ij)})^2\right\|_2^2 = \sum_{l=1}^{n} [\mathbb{E}[A_{jli}] - A_{jli}]^4 \leq n \left[\max_l A_{jli} - A_{jli}\right]^4.$$

Note that A_{jli} are sub-Gaussian variables with $\sigma = 1$, we have

$$\max_l |A_{jli} - A_{jli}| \leq t_0,$$

with probability $1 - 2ne^{-t_0^2}$. Hence, we have

$$\left\|(\boldsymbol{Y}^{(ij)})^2\right\|_2 \leq \sqrt{n}t_0^2, \tag{1}$$

with probability $1 - 2ne^{-t_0^2}$. To let the inequality (1) holds for all (i, j), we have

$$\max_{i\in[N], j\in[n]} \left\|(\boldsymbol{Y}^{(ij)})^2\right\|_2 \leq \sqrt{n}t_0^2,$$

with probability $(1 - 2ne^{-t_0^2})^{nN}$.
Note that $\sigma_y^2 = \left\|\sum_{ij} \mathbb{E}[\left(\boldsymbol{Y}^{(ij)}\right)^2]\right\|_2 \leq v_0^2 nN$. Then, by Lemma S1, we have

$$\begin{aligned} P\left(\left\|\frac{1}{N}\sum_{i=1}^{N}\sum_{j=1}^{n}\left(\boldsymbol{Y}^{(ij)}\right)^2\right\|_2 \geq \frac{t}{N}\right) &\leq n\exp\left\{-\frac{t^2/2}{v_0^2 nN + \sqrt{n}t_0^2 t/3}\right\}(1 - 2ne^{-t_0^2})^{nN} + 1 - (1 - 2ne^{-t_0^2})^{nN} \\ &\leq n\exp\left\{-\frac{u^2 N/2}{v_0^2 n + \sqrt{n}t_0^2 u/3}\right\}(1 - 2ne^{-t_0^2})^{nN} + 1 - (1 - 2ne^{-t_0^2})^{nN}, \end{aligned} \tag{2}$$

where $u = t/N$.

Let $t_0^2 = 3\log n + 2\log N$. Then, the probability $(1 - 2ne^{-t_0^2})^{nN} = \mathcal{O}(e^{-1/nN})$ and $1 - (1 - 2ne^{-t_0^2})^{nN} = \mathcal{O}(1 - e^{-1/nN}) = \mathcal{O}(1/nN)$.

Let $u = \mathcal{O}\left(\sqrt{nt_0^4 \log^2 n/N}\right)$. The right hand side is propotion to

$$n \exp\left\{-\frac{2nt_0^4 \log^2 n}{nt_0^4 \log n}\right\}\left(1 - \frac{1}{nN}\right) + \frac{1}{nN} = \mathcal{O}(1/n).$$

Therefore, we have

$$\left\|\frac{1}{N}\sum_{i=1}^{N}\sum_{j=1}^{n}\left(\boldsymbol{Y}^{(ij)}\right)^2\right\|_2 \leq \sqrt{\frac{n(3\log n + 2\log N)^2 \log^2 n}{N}},$$

with probability $\mathcal{O}(1/n)$.

2 Poisson Case

For Poisson data, the inequality (1) holds with probability $1 - ne^{-t_0}$. Then, we should plug $t_0 = 3\log n + 2\log N$ in the inequality (2) and have the result

$$P\left(\left\|\frac{1}{N}\sum_{i=1}^{N}\sum_{j=1}^{n}\left(\boldsymbol{Y}^{(ij)}\right)^2\right\|_2 \geq \frac{t}{N}\right) \lesssim \sqrt{n(3\log n + 2\log N)^4 \log^2 n/N},$$

with probability $1 - \mathcal{O}(n^{-1})$.

References