Notes of [Paul, 2007]

Rui Wang, Xingzhong Xu

August 8, 2017

Abstract

Key words: High dimension, Huber loss.

1 Weak concentration inequalities for quadratic forms

Lemma 1. *Suppose X and Y are iid $N_n(0, I_n)$, C is a $n \times n$ matrix, $\|C\|$ is the largest singular value of C. Then for $0 < \delta < 1$,*

$$\Pr\left(\frac{1}{n}|X^T CY| > t\right) \le 2\exp\Big(-\frac{(1-\delta^2)nt^2}{2\|C\|^2}\Big), \quad \textit{for } 0 \le t \le \frac{\delta}{1-\delta^2}\|C\|.$$

Proof. For $\lambda \le \frac{1}{\|C\|}$, we have

$$\begin{aligned}\Pr(X^T CY > \lambda nt) &\le \exp(-nt)\operatorname{E}\exp(\lambda X^T CY)\\ &= \exp(-\lambda nt)\operatorname{E}\exp\left(\frac{\lambda^2}{2}\|CY\|^2\right)\\ &= \exp(-\lambda nt)|I_n - \lambda^2 C^T C|^{-1/2}.\end{aligned} \tag{1}$$

Now, use the fact that $\log|I_n - \lambda^2 C^T C| = \sum_{i=1}^n \log(1 - \lambda^2\lambda_i(C^T C))$. We have

$$\begin{aligned}-\log|I_n - \lambda^2 C^T C| &= \sum_{i=1}^n -\log(1 - \lambda^2\lambda_i(C^T C))\\ &\le -n\log(1 - \lambda^2\|C\|^2)\\ &\le n\frac{\lambda^2\|C\|^2}{1 - \lambda^2\|C\|^2}.\end{aligned}$$

Hence with (1), for $0 < \delta < 1$ and $0 \le \lambda \le \frac{\delta}{\|C\|}$, we have,

$$\begin{aligned}
\Pr(X^T CY > \lambda nt) &\le \exp(-n\lambda t + \frac{n}{2}\frac{\lambda^2\|C\|^2}{1-\lambda^2\|C\|^2}) \\
&= \exp\Big\{n(-\lambda\|C\|\frac{t}{\|C\|} + \frac{1}{2}\frac{(\lambda\|C\|)^2}{1-(\lambda\|C\|)^2})\Big\} \qquad (2) \\
&= \exp\Big\{n(-\lambda\|C\|\frac{t}{\|C\|} + \frac{1}{2}\frac{(\lambda\|C\|)^2}{1-\delta^2})\Big\}.
\end{aligned}$$

The last expression is a quadratic form of $\lambda\|C\|$, it takes it's minimum when $\lambda\|C\| = \frac{1-\delta^2}{\|C\|}t$. To make this value fall in the interval $(0, \delta)$, we require $t \le \frac{\delta}{1-\delta^2}\|C\|$. Thus, for $0 \le t \le \frac{\delta}{1-\delta^2}\|C\|$, we have

$$\begin{aligned}
\Pr(X^T CY > \lambda nt) &\le \exp\Big\{n(-\frac{1-\delta^2}{\|C\|}t\frac{t}{\|C\|} + \frac{1}{2}\frac{(\frac{1-\delta^2}{\|C\|}t)^2}{1-\delta^2})\Big\} \\
&= \exp\Big\{n(-\frac{1-\delta^2}{\|C\|}t\frac{t}{\|C\|} + \frac{1}{2}\frac{(\frac{1-\delta^2}{\|C\|}t)^2}{1-\delta^2})\Big\} \\
&= \exp\Big\{-\frac{n}{2}\frac{1-\delta^2}{\|C\|^2}t^2\Big\}.
\end{aligned}$$

□

Lemma 2. *Suppose X is distributed as $N_n(0, I_n)$. Then for $0 < \delta < 1$,*

$$\Pr(\frac{1}{n}(X^T CX - \operatorname{tr} C) > t) \le 2\exp(-\frac{\delta n t^2}{4\|C\|^2}) \quad \textit{for } t \le \frac{\|C\|(1-\delta)}{\delta}.$$

Proof.

$$\begin{aligned}
\Pr(\frac{1}{n}(X^T CX - \operatorname{tr} C) > t) &= \Pr(\lambda X^T CX > \lambda(nt + \operatorname{tr} C)) \\
&\le \exp(-\lambda(nt + \operatorname{tr} C))\,\mathrm{E}(\exp(\lambda X^T CX)) \\
&= \exp(-\lambda(nt + \operatorname{tr} C))|I_n - 2\lambda C|^{-1/2} \\
&= \exp\big(-\lambda(nt + \operatorname{tr} C) - \frac{1}{2}\log|I_n - 2\lambda C|\big) \\
&\le \exp\big(-\lambda nt - \lambda \operatorname{tr} C - \frac{1}{2}\log|I_n - 2\lambda C|\big).
\end{aligned}$$

We have the inequality

$$-\log(1-u) - u \le \frac{u^2}{2(1-|u|)} \quad \text{for } u \in (-1, 1).$$

Hence

$$\Pr(\frac{1}{n}(X^T CX - \operatorname{tr} C) > t) \le \exp\big(-\lambda nt + \frac{1}{2}\sum_{i=1}^{n} \frac{(2\lambda\lambda_i(C))^2}{2(1-2|\lambda\lambda_i(C)|)}\big).$$

For $\lambda \le \frac{1-\delta}{2\|C\|}$

$$\begin{aligned}
-\lambda nt + \frac{1}{2}\sum_{i=1}^{n} \frac{(2\lambda\lambda_i(C))^2}{2(1-2|\lambda\lambda_i(C)|)} &\le -\lambda nt + \frac{1}{2}\sum_{i=1}^{n} \frac{(2\lambda\lambda_i(C))^2}{2\delta} \\
&\le -\lambda nt + \frac{n\lambda^2\|C\|^2}{\delta}
\end{aligned}$$

For $t \le \frac{\|C\|(1-\delta)}{\delta}$, we let $\lambda = \frac{\delta t}{2\|C\|^2}$, then

$$-\lambda nt + \frac{n\lambda^2\|C\|^2}{\delta} \le -\frac{\delta nt^2}{4\|C\|^2}$$

$\square$

References

References

[Paul, 2007] Paul, D. (2007). Asymptotics of sample eigenstruture for a large dimensional spiked covariance model. *Statistica Sinica*, 17:1617.