03

def: ~~Subexponential~~ X mean 0.

sub gaussian: $\exists K>0$ s.t. $Ee^{tX} \le e^{Kt^2}$ for all $t\in\mathbb{R}$.

$\Updownarrow$

X mean 0. $P(|X|>t) \le \cancel{\exp}\, e^{1-\frac{t^2}{4K}}$

$\Updownarrow$

X mean 0 $\left.\begin{matrix} P(X>t) \\ P(X<-t)\end{matrix}\right\} \le e^{-\frac{t^2}{4K}}$

contrast.

Sub exponential: X mean 0 $P(|X|>t) \le e^{1-Kt}$

or $\le 2e^{-Kt}$

$t\in[0,\infty)$

$\Updownarrow$

X mean 0, $C\in(0,\infty)$, $\delta\in(0,\infty)$ s.t. $t\in(-\delta,\delta)$ $Ee^{tY} \le e^{Ct^2}$

values for C, δ: set $\delta = \frac{K}{2}$. $C = \frac{2e}{K^2}$

Thms: ① X, Y subgaussian with mean 0. $\Rightarrow$ X+Y, X−Y are subgaussian.

$$E\{e^{t(X+Y)}\},\ E\{e^{t(X-Y)}\} \le e^{t^2(2K_X+2K_Y)}$$

② X subgaussian $\Rightarrow$ X^2 subexponential. $P(|X^2|>u) \le e^{1-\frac{u}{4K}}$

③ $\{Y_i\}_{i=1}^n$ independent, mean 0 ⊕ subexponential with constant $C_0\in(0,\infty)$. $\max_i P(|Y_i|>t) \le e^{1-C_0t}$ $\forall t\in[0,\infty)$

$\Rightarrow \sum_{i=1}^n Y_i$ subgaussian: $P(|\sum_{i=1}^n Y_i| > nt) \le 2\exp(-Knt^2) < e^{1-Knt^2}$.

(t positive & sufficiently small)

def: random sequence bdd in probability.

X_n of order less than ~~or~~ or equal to the order of a_n: $\Longleftrightarrow$ $X_n = O_p(a_n)$

constant seq.

$\forall \varepsilon>0$. $\exists$ constant M_ε & integer n_ε s.t. $P(|X_n| \le M_\varepsilon |a_n|) \ge 1-\varepsilon$ for $\forall n>n_\varepsilon$.

Thms for Covariance Matrix.

① $X_1 \cdots X_n$ iid copies of $X = \begin{pmatrix} X_{[1]} \\ \vdots \\ X_{[p]}\end{pmatrix}$, mean 0. ⊕ $X_{[1]} \cdots X_{[p]}$ are subgaussian. $\max_j E e^{tX_{[j]}} \le e^{C_0t^2}$ $\forall t\in\mathbb{R}$.

⊕ $EXX^T = \Sigma_* \in S_p^+$, Σ_* bdd diagonally, ($\max_j \Sigma_{*jj} = \max_j E X_{[j]}^2 \le M$)

$\Longrightarrow$ $\exists$ Constant K. $P(|S_{jk} - \Sigma_{*jk}|>t) \le 4\exp(-Knt^2)$

$\rightarrow P(\|S-\Sigma_*\|_{max} \le \nu\sqrt{\frac{\log p_n}{n}}) \ge 1 - 4p_n^{2-K_2\nu^2}$

$\exists K_2$.

$\rightarrow$ if $\log p_n = o(n)$, $\|S-\Sigma_*\|_{max} = O_p(\sqrt{\frac{\log p_n}{n}})$

Q&A for this thm:

1. this is saying: as long as we choose $\nu > \sqrt{\frac{2}{k}}$, P_n can even grow faster than n (as long as $P_n = o(n)$).

2. is it ok that for $M_\varepsilon = \nu$, $\varepsilon = 4P_n^{2-K_2\nu^2}$ depends on P_n?

   I think it's ok, because when $\nu > \sqrt{\frac{2}{K_2}}$, $2 - K_2\nu^2 < 0$, ε decrease with P_n increases, and P_n increases with n.

   So we can still select the smallest n_ε s.t. ε achieves the "minimum".

② $X_1 \cdots X_n$ iid copies of $X = \begin{pmatrix} X_{[1]} \\ \vdots \\ X_{[p]} \end{pmatrix}$, mean 0. ⊕ for some $\alpha > 2$, $\exists\, C_\alpha \in (0, \infty)$ s.t. $\max_{j \in \{1,\cdots,p\}} E(|X_{[j]}|^{2\alpha}) \leq C_\alpha$ ⊕ $EXX^T = \Sigma_* \in S_p^+$, Σ_* bdd diagonaly.

$\Rightarrow \exists$ constant K $P(|S_{jk} - \Sigma_{*jk}| > t) \leq \frac{K}{n^{\frac{\alpha}{2}} t^\alpha}$

$\rightarrow \exists K_2$, $P(\|S - \Sigma_*\|_{max} \leq \nu\sqrt{\frac{p^{\frac{4}{\alpha}}}{n}}) \geq 1 - \frac{K_2}{\nu^\alpha}$ $(n^{-1}p^{\frac{4}{\alpha}}$ sufficiently small.)

$\rightarrow$ if $p^{\frac{4}{\alpha}} = o(n)$, $\|S - \Sigma_*\|_{max} = O_p\left(\sqrt{\frac{p^{\frac{4}{\alpha}}}{n}}\right)$

Q&A for this thm:

1. it gets a weaker result. $\alpha < \frac{4}{\alpha} < 2$, ~~$\alpha = 2$ allows P_n grows the fastest.~~

   $\frac{\alpha}{4} > \frac{1}{2}$. $P_n = O(n^{\frac{\alpha}{4}})$ certainly slower than $O(e^n)$.

   $\frac{\log P_n}{P_n^{\frac{4}{\alpha}}} \rightarrow \infty$. $\forall \alpha$. $(n \to \infty)$

   So what's the point of having this thm?

2. is "$\max_{j=1\cdots p} E(|X_{[j]}|^{2\alpha}) \leq C_\alpha$" a weaker condition than: "$\max_j E e^{tX_{[j]}} \leq e^{C_0 t^2}$ $\forall t \in \mathbb{R}$."

   yes: Gaussian: all moments exist & finite.

   $\rightarrow$ sub gaussian: same,

   So the second condition implies the first.

04

preparation: lasso inverse covariance $\longrightarrow$ lasso inverse correlation.

Covariance: Likelihood: $(2\pi)^{-\frac{p}{2}} |\Sigma|^{-\frac{1}{2}} \exp(-\frac{1}{2}(x-\mu)^T\Sigma^{-1}(x-\mu))$

~~take log: $-\frac{p}{2}\log(2\pi) - \frac{1}{2}\log|\Sigma| - \frac{1}{2}tr(\Sigma^{-1}(x-\mu)(x-\mu)^T)$~~

n copies, & plug in $\hat{\mu} = \frac{1}{n}\sum_{i=1}^{n} x_i \rightarrow |\Sigma|^{-\frac{n}{2}} \exp(-\frac{1}{2}\sum_{i=1}^{n}(x_i-\bar{x})^T\Sigma^{-1}(x_i-\bar{x}))$

take log. $-\frac{n}{2}\log|\Sigma| - \frac{1}{2}tr(\Sigma^{-1}\sum_{i=1}^{n}(x_i-\bar{x})(x_i-\bar{x})^T)$

$S = \frac{1}{n}\sum_{i=1}^{n}(x_i-\bar{x})(x_i-\bar{x})^T \Rightarrow \cancel{\frac{n}{2}} \frac{n}{2}(\log|\Sigma^{-1}| - tr(\Sigma^{-1}S))$

Correlation: R: sample correlation. need to refer to the paper.

now work with correlation matrix. R: sample correlation.

Θ_*^{-1}: population correlation matrix.

Θ estimate the inverse of correlation matrix.

$$\hat{\theta}_\lambda = \underset{\theta \in S_+^p}{\text{argmin}} \{ tr(\theta R) - \log|\theta| + \lambda|\theta^-|_1 \}.$$

interested: $\|\hat{\theta}_\lambda - \theta_*\|_F$.

Thm 1: def: s is # non-zero off-diagonal entrees of θ_*. $\leftarrow$ true precision matrix.

$\exists\, \tau \in (0,1]$ $\tau \le \varphi_{min}(\theta_*^{-1}) \le \varphi_{max}(\theta_*^{-1}) \le \tau^{-1}$ ⊕ given ε positive & sufficiently small (can take $\varepsilon \in (0,\tau)$).

⊕ $\underbrace{\max_{j \neq k} |R_{jk} - (\theta_*^{-1})_{jk}| \le \varepsilon \frac{1}{\sqrt{s}} \frac{\tau^2}{32}}_{\text{this happen.}} \Longrightarrow \|\hat{\theta}_\lambda - \theta_*\|_F \le \varepsilon.$ occurs.

? what is the reason for introducing s?

Cor 1: $X_1 \cdots X_n$ iid copies of $X = \begin{pmatrix} X_{[1]} \\ \vdots \\ X_{[p]} \end{pmatrix}$, mean 0 ⊕ $X_{[1]} \cdots X_{[p]}$ subgaussian.

⊕ $E XX^T = \Sigma_* \in S_p^+$, $\exists$ constant τ $\tau \le \varphi_{min}(\Sigma^*) \le \varphi_{max}(\Sigma^*) \le \frac{1}{\tau}$ all p_n.

Q: does this guarentee Σ^* bdd diagonally?

$\Longrightarrow$ if $\frac{s_n \log p_n}{n}$ sufficiently small. $P(\|\hat{\theta}_{\lambda_n} - \theta_*\|_F \le \nu\sqrt{\frac{s_n \log p_n}{n}}) \ge 1 - c_1 p_n^{2-c_2\nu^2}$

$\downarrow$

therefore $\lambda_n = \nu(n^{-1}\log p_n)^{\frac{1}{2}}\frac{\tau^2}{32}$ small.