The beta - mixture shrinkage prior for sparse covariances with near-minimax posterior convergence

신창수, 오정훈, 장태영, 이경원, 정진욱, 김성민, 백승찬

August 10, 2022

To Do List

- 소개 작성
- 스타일 통일
- 기호 통일 ($R, \mathbb{R}$)
- 백승찬 - Lemma 6 수정
- theorem, lemma 등 한글화
- equation numbering

소개

추가 예정

1 신창수 - Theorem 3 and Lemma 2

1.1 Theorem 3

(The upper bound of the packing number).
If $\zeta^4 \le p,\ p \asymp n^\beta\ for\ some\ 0 \le \beta \le 1,\ s_n = c_1 n\epsilon_n^2/lnp,\ L_n = c_2 n\epsilon_n^2\ \ and\ \delta_n = \epsilon_n/\zeta^3\ for\ some\ constants\ c_1 > 1\ andc_2 > 0, we\ have$

$$lnD(\epsilon_n, P_n, d)\ \le\ (12 + 1/\beta)c_1 n\epsilon_n^2$$

Proof. 이 정리는 Lemma1의 세 가지 조건 중 첫 번째 조건이 성립함을 보이는 정리이다. 증명에 앞서 각각의 용어들에 대해 정의하자.

$D(\epsilon_n, P_n, d)$ 은 P_n 안에서, 각각의 쌍이 이루는 거리들의 거리가 ϵ_n 보다 크거나 같은 점들의 최대 개수로 정의한다. 또한,

$P_n = \{f_\Sigma\ :\ |s(\Sigma, \delta_n)| \le s_n,\ \zeta^{-1} \le \lambda_{min}(\Sigma) \le \lambda_{max}(\Sigma) \le \zeta, ||\Sigma||_{max} \le L_n\}$

$\mathcal{U}(\delta_n, s_n, L_n, \zeta) = \{\Sigma \in \rfloor_p : |s(\Sigma, \delta_n)| \le s_n, \zeta^{-1} \le \lambda_{min}(\Sigma) \le \lambda_{max}(\Sigma) \le \zeta, ||\Sigma||_{max} \le L_n\}$

s_0 : 공분산 행렬의 비대각 원소 중 0이 아닌 원소들의 상한

$\epsilon_n := \{\dfrac{(p + s_0)lnp}{n}\}^{\frac{1}{2}}$ 라고 정의하자.

최종적으로는, $lnD(\epsilon_n, \mathcal{P}_n, d) \leq (12 + 1/\beta)c_1 n\epsilon_n^2$ 임을 보일 건데, 그 전에 먼저 lemma3 소개하자.

[Lemma3](Lemma A.1 in [6])
If P_{Ω_k} is the density of $N_p(0, \Omega_k^{-1})$, k=1,2, then for all $\Omega_k \in \mu_0^+$, k=1,2, and d_i, i=1,2,⋯,p, eigenvalues of $A = \Sigma_1^{-1/2}\Sigma_2\Sigma_1^{-1/2}$, we have that for some $\delta > 0$ and constant $c_o > 0$,

(1)$c_0^{-1}||\Sigma_1 - \Sigma_2||_2^2 \leq \sum_{i=1}^p |d_i - 1|^2 \leq c_o||\Sigma_1 - \Sigma_2||_2^2$

(2)$h(p_{\Omega_1}, p_{\Omega_2}) < \delta$ implies $\max_i |d_i - 1| < 1$ and $||\Omega_1 - \Omega_2||_2 \leq c_o h^2(p_{\Omega_1}, p_{\Omega_2})$

(3) $h^2(p_{\Omega_1}, p_{\Omega_2}) \leq c_o||\Omega_1 - \Omega_2||_2^2$
* 여기서, $h()$는 hellinger distance 로, 우리 논문에서 $d()$와 같음

따라서, lemma3-(3)에 의해, $d(f_{\Sigma_1}, f_{\Sigma_2}) \leq c\zeta||\Omega_1 - \Omega_2||_F$ 임을 알 수 있고, 여기서 Ω는 Σ^{-1} 이다.

위 lemma3-(3)과 우리 논문의 lemma5를 통해, $\Omega_1 = \Sigma_1^{-1}\Sigma_1\Sigma_1^{-1}$ 임을 이용

하면, $d(f_{\Sigma_1}, f_{\Sigma_2}) \leq C\zeta^3||\Sigma_1 - \Sigma_2||_F$ (1) 식을 얻을 수 있다. ϵ-packing의

정의에 의해, $d(f_{\Sigma_i}, f_{\Sigma_j}) \geq \epsilon_n$ 으로부터, (1) 식과 결합하여, $||\Sigma_1 - \Sigma_2||_F \geq \dfrac{\epsilon_n}{C\zeta^3}$ 의
결과를 얻을 수 있다. 따라서, 집합을 $\mathcal{P}_n$ 에서, $\mathcal{U}(\delta_n, s_n, L_n, \zeta)$ 으로 바꾸고,

거리를 Frobenius norm 으로 바꾸어서 위의 결과를 적용하자. 여기서 격자

개념을 생각해보면, 격자의 대각선 부분들까지 고려해주어야 한다. 따라서,

$$lnD(\epsilon_n, P_n, d) \leq lnD(\frac{\epsilon_n}{C\zeta^3}, \mathcal{U}(\delta_n, s_n, L_n, \zeta), ||\cdot||_F)$$

$$\leq ln\left\{\left(\frac{L_n\sqrt{p+j}}{\frac{\epsilon_n}{C\zeta^3}}\right)^p \sum_{j=1}^{s_n}\left(\frac{\sqrt{p+j}\frac{1}{\sqrt{2}}L_n}{\frac{\epsilon_n}{C\zeta^3}}\right)^j \binom{\frac{p}{2}}{j}\right\}$$

여기서, $\sqrt{p+j}$는 격자의 대각선을 고려해준 항이고, $\dfrac{1}{\sqrt{2}}$는 Frobenius norm에서, symmetric term들의 중복을 고려해 준 값이다.

한편, $\left(\dfrac{L_n\sqrt{p+j}}{\frac{\epsilon_n}{C\zeta^3}}\right)$에서 $j \leq s_n \leq p^2$ 임을 통해, j를 p에 대한 부등식으로 적절히 바꾸어주면, $\dfrac{L_n\sqrt{p+j}}{\frac{\epsilon_n}{C\zeta^3}} \leq \dfrac{2p\zeta^3 L_n}{\epsilon_n}$ 을 얻을 수 있다.

따라서, $ln\left\{\left(\frac{L_n\sqrt{p+j}}{\frac{\epsilon_n}{C\zeta^3}}\right)^p \sum_{j=1}^{s_n}\left(\frac{\sqrt{p+j}\frac{1}{\sqrt{2}}L_n}{\frac{\epsilon_n}{C\zeta^3}}\right)^j \binom{\frac{p}{2}}{j}\right\}$

$$= ln\left\{\left(\frac{2pC\zeta^3 L_n}{\epsilon_n}\right)^p \sum_{j=1}^{s_n}\left(\frac{\sqrt{2}C\zeta^3 L_n p}{\epsilon_n}\right)^j \binom{\frac{p}{2}}{j}\right\}$$

$$= ln\left[((2p)^p(\sqrt{2}p)^{s_n})\left(\frac{C\zeta^3L_n}{\epsilon_n}\right)^p\sum_{j=1}^{s_n}\left(\frac{C\zeta^3L_n}{\epsilon_n}\right)^j\binom{\frac{p}{2}}{j}\right]$$

$$= pln2 + plnp + s_n(\frac{1}{2}ln2 + lnp) + pln\left(\frac{CL_n\zeta^3}{\epsilon_n}\right) + ln\left(\sum_{j=1}^{s_n}\left(\frac{2CL_n\zeta^3}{\epsilon_n}\right)(\frac{p^2}{2})^j\right)$$

$$\leq pln2 + plnp + s_n(\frac{1}{2}ln2 + lnp) + pln\left(\frac{CL_n\zeta^3}{\epsilon_n}\right) + s_nln\left(\frac{2CL_n\zeta^3p^2}{\epsilon_n}\right)$$

$$\leq pln2 + plnp + \frac{1}{2}s_nln2 + s_nlnp + (p+s_n)ln(2CL_n) + (p+s_n)ln\zeta^3 + (p+s_n)ln\frac{1}{\epsilon_n} + 2s_nlnp$$

에서 적절한 상수를 곱해주면,

$$\leq 2(p+s_n)lnp + (p+s_n)ln(2CL_n) + (p+s_n)ln\zeta^3 + (p+s_n)ln\frac{1}{\epsilon_n} + 2s_nlnp$$

이다. 먼저, $(p+s_n)ln(2CL_n) \leq 6s_nlnp$ 임을 보일건데,

위에서 정의한 s_n, ϵ_n, L_n 을 통해, $s_n = c_1(p+s_0)$ 이므로,

$p + s_n = (1+c_1)p + c_1s_0$ 임을 알 수 있다. 또한, $2CL_n = 2c_2n\epsilon_n^2$ 이다.

이를 좌변에 대입하면,$(p+s_n)ln(2CL_n) = ((c_1+1) + c_1s_0)ln2c_2n\epsilon_n^2$

$= ((c_1+1)p + c_1s_0)ln2c_2(p+s_0)lnp$ 이다. $c_1 > 1$ 가정에 의해,

$((c_1+1) + c_1s_0)ln2c_2n\epsilon_n^2 \leq 2c_1(p+s_0)ln(2c_2(p+s_0)lnp)$ 이다. 한편, s_0 는

비대각 원소 중 0이 아닌 것들의 개수의 상한이므로 $s_o \leq p^2$ 임을 알 수 있다.

따라서, 적절한 차수 p^3 을 통해 $2c_2(p+s_0)lnp < p^3$ 을 얻을 수 있다. 이를 통해,

$(p+s_n)ln(2CL_n) \leq 2c_1(p+s_0)lnp^3 = 6s_nlnp$ 부등식을 얻을 수 있다.

이를 정리하면,

$$2(p+s_n)lnp + (p+s_n)ln(2CL_n) + (p+s_n)ln\zeta^3 + (p+s_n)ln\frac{1}{\epsilon_n} + 2s_nlnp$$

$$\leq 2(p+s_n)lnp + 6s_nlnp + (p+s_0)ln\zeta^3 + (p+s_n)ln(\frac{1}{\epsilon_n}) + 2s_nlnp$$ 이다.

이제 $(p+s_0)ln\zeta^3 \leq \frac{3}{4}(p+s_n)lnp$ 임을 보이자.

$(p+s_n)ln\zeta^3 = \frac{3}{4}(p+s_n)ln\zeta^4$ 인데, 가정에 의해, $\zeta^4 \leq p$ 이므로,

$(p+s_n)ln\zeta^3 \leq \frac{3}{4}(p+s_n)lnp$ 임을 알 수 있다. 이를 대입하여 정리하면,

$$\leq 2(p+s_n)lnp + 6s_nlnp + (p+s_0)ln\zeta^3 + (p+s_n)ln(\frac{1}{\epsilon_n}) + 2s_nlnp$$

$$\leq 2(p+s_n)lnp + 6s_nlnp + \frac{3}{4}(p+s_n)lnp + (p+s_n)ln(\frac{1}{\epsilon_n}) + 2s_nlnp$$ 이다.

세번째 항에, 앞에서 정의한 ϵ_n 을 대입하여 정리하면,

$(p+s_n)ln(\frac{1}{\epsilon_n}) = \frac{1}{2}(p+s_n)ln\frac{n}{(p+s_o)lnp}$ 임을 알 수 있고,

우변$= \frac{1}{2\beta}(p+s_n)ln\left(\frac{n}{(p+s_o)lnp}\right)^\beta$

$$= \frac{1}{2\beta}(p+s_n)lnn^\beta - \frac{1}{2}(p+s_n)ln(p+s_0)lnp$$

$\leq \frac{1}{2\beta}(p+s_n)lnn^{\beta} = \frac{1}{2\beta}(p+s_n)lnp$ 이다.($\because p \asymp n^{\beta}$)

이를 다시 처음 부등식에서 정리하면,

$\leq 2(p+s_n)lnp + 6s_nlnp + \frac{3}{4}(p+s_n)lnp + \frac{1}{2\beta}(p+s_n)lnp + 2s_nlnp$ 이다.

이 부등식에 적절한 상수배를 해주면,

$\leq 6s_nlnp + \frac{11}{4}(p+s_n)lnp + \frac{1}{2\beta}(p+s_n)lnp + 2s_nlnp$

$(6+\frac{1}{2\beta})(\frac{s_n}{c_1}+s_n)lnp < \frac{1}{2}(12+\frac{1}{\beta})(c_1+c_1)n\epsilon_n^2$ ($\because c_1 > 1$)

$= (12+\frac{1}{2\beta})c_1n\epsilon_n^2$ 이다.

따라서, $lnD(\epsilon_n, P_n, d) \leq (12+\frac{1}{\beta})c_1n\epsilon_n^2$ 이 성립해서,

이를 통해 Lemma1의 첫 번째 조건이 성립함을 알 수 있다.

□

1.2 Lemma 2

If $a = b = \frac{1}{2}, \tau = O(\frac{1}{p^2}\sqrt{\frac{s_0lnp}{n}}),\ s_0lnp = O(n)$ *and* $\zeta > 3$, *we have, for some constant* $C > 0$,

$$\pi^u(\Sigma \in \mathcal{U}(\zeta)) > \left\{\frac{\lambda\zeta}{8}exp(-\frac{\lambda\zeta}{4} - C)\right\}^p$$

Proof. 이 Lemma는 논문의 Theorem4 의 증명에서 활용되는 Lemma이다.

Gershgorin circle Thm에 의해, covariance matrix의 eigenvalue 들은 적어도

$[\sigma_{jj} - \sum_{k\neq j}|\sigma_{kj}|, \sigma_{jj} + \sum_{k\neq j}|\sigma_{kj}|], j \in \{1,2,\cdots,p\}$ 안에 있다는 것을 알 수 있다. 따라서,

$\pi^u(\Sigma \in \mathcal{U}(\zeta)) \geq \pi^u(min_j(\sigma_{jj} - \sum_{k\neq j}|\sigma_{kj}|) > 0, \zeta^{-1} \leq \lambda_{min}(\Sigma) \leq \lambda_{max}(\Sigma) \leq \zeta)$ 임을 보이면 된다.

먼저 $\min_j(\sigma_{jj} - \sum_{k\neq j}|\sigma_{kj}|) > 0$을 살펴보면, 1-norm의 정의에 의해,

$||\Sigma||_1 = \max_{1\leq j\leq n}\sum_{i=1}^{m}|a_{ij}|$ 이므로,

$\lambda_{max}(\Sigma) \leq ||\Sigma||_1 = \max_j(\sigma_{jj} + \sum_{k\neq j}|\sigma_{kj}|) \leq \max_j 2\sigma_{jj}$ 로 표현할 수 있다. 또한, G.C Thm에 의해, $\lambda_{min}(\Sigma) \geq \min_j(\sigma_{jj} - \sum_{k\neq j}|\sigma_{kj}|)$로 표현할 수 있다. 따라서,

이를 위의 식 $\zeta^{-1} \leq \lambda_{min}(\Sigma) \leq \lambda_{max}(\Sigma) \leq \zeta$에 적용해서 다시 표현하면,

$\pi^u(\Sigma \in \mathcal{U}(\zeta)) \geq \pi^u(\zeta^{-1} \leq \min_j(\sigma_{jj} - \sum_{k\neq j}|\sigma_{kj}|) \leq 2\max_j\sigma_{jj} \leq \zeta)$ 이고, $P(A) \geq P(A \cap B) = P(A|B)P(B)$ 성질을 이용하면,

$\pi^u(\zeta^{-1} \leq \min_j(\sigma_{jj} - \sum_{k\neq j}|\sigma_{kj}|) \leq 2\max_j\sigma_{jj} \leq \zeta)$

$\geq \pi^u(\zeta^{-1} \leq \min_j(\sigma_{jj} - \sum_{k\neq j}|\sigma_{kj}|) \leq 2\max_j \sigma_{jj} \leq \zeta \mid \max_{k\neq j}|\sigma_{kj}| < (\zeta p)^{-1})\pi^u(\max_{k\neq j}|\sigma_{kj}| < (\zeta p)^{-1})$ 인데, 조건부에 의해 다음 부등식이 되고,

$\geq \pi^u(\zeta^{-1} \leq \min_j(\sigma_{jj} - \zeta^{-1}) \leq 2\max_j \sigma_{jj} \leq \zeta \mid \max_{k\neq j}|\sigma_{kj}| < (\zeta p)^{-1})\pi^u(\max_{k\neq j}|\sigma_{kj}| < (\zeta p)^{-1})$ 에서, σ_{jj} 항과 σ_{ij} 는 독립이므로 조건부 항을 없앨 수 있다. 따라서,

$= \pi^u(\zeta^{-1} \leq \min_j(\sigma_{jj} - \zeta^{-1}) \leq 2\max_j \sigma_{jj} \leq \zeta)\pi^u(\max_{k\neq j}|\sigma_{kj}| < (\zeta p)^{-1})$ 이다.
다시 정리해보면, 다음과 같은 부등식을 얻을 수 있다.

$$\pi^u(\Sigma \in \mathcal{U}(\zeta)) \geq \underline{\pi^u(\zeta^{-1} \leq \min_j(\sigma_{jj} - \zeta^{-1}) \leq 2\max_j \sigma_{jj} \leq \zeta)} * \underline{\pi^u(\max_{k\neq j}|\sigma_{kj}| < (\zeta p)^{-1})}$$

먼저 첫 번째 밑줄 확률을 계산해보면,

$\pi^u(\zeta^{-1} \leq \min_j(\sigma_{jj} - \zeta^{-1}) \leq 2\max_j \sigma_{jj} \leq \zeta) \geq \pi^u(2\zeta^{-1} \leq \sigma_{jj} \leq \frac{\zeta}{2}, \forall j)$
$\geq \prod_{j=1}^{p} \pi^u(2\zeta^{-1} \leq \sigma_{jj} \leq \frac{\zeta}{2})$ 에서,
σ_{jj} 가 $\Gamma(1, \frac{\lambda}{2})$ 를 따른다는 가정에 의해, $f(\sigma_{jj}) = \frac{\lambda}{2}exp(-\frac{\lambda}{2}\sigma_{jj})$ 의 pdf 를 갖고, 가로 길이가 $\left(\frac{2}{\zeta}, \frac{\zeta}{2}\right)$ 이고 세로 길이가 $\left(0, f\left(\frac{2}{\zeta}\right)\right)$ 인 직사각형을 생각하면,
이는 pdf 의 전체 넓이 보다는 작으므로, 이를 통해 $\prod_{j=1}^{p} \pi^u(2\zeta^{-1} \leq \sigma_{jj} \leq \frac{\zeta}{2})$
$= \left\{\left(\frac{\zeta}{2} - \frac{2}{\zeta}\right)\frac{\lambda}{2}exp(-\frac{\lambda\zeta}{4})\right\}^p \geq \left\{\frac{\lambda\zeta}{8}exp(-\frac{\lambda\zeta}{4})\right\}^p$ 임을 알 수 있다.

이제 두 번째 밑줄 확률인 $\pi^u(\max_{k\neq j}|\sigma_{kj}| < (\zeta p)^{-1})$ 를 계산할 건데, 먼저 이를 위한 lemma 하나를 소개하자.

[lemma 1 in [12]]

The univariate horseshoe density $p(\theta)$ satisfies the following:

$(a)\ \lim_{\theta\to 0} p(\theta) = \infty$

$(b)\ For\ \theta \neq 0, \frac{K}{2}log\left(1 + \frac{4}{\theta^2}\right) < p(\theta) < Klog\left(1 + \frac{2}{\theta^2}\right),\ where\ K = \frac{1}{\sqrt{2\pi^3}}$

따라서, 위의 lemma (b) 를 통해, $\pi^u(\max_{k\neq j}|\sigma_{kj}| < (\zeta p)^{-1})$ 의 계산을 위한 부등식인 $\pi^u(\sigma_{kj}) \leq \frac{1}{\tau\sqrt{2\pi^3}}ln\left(1 + \frac{2\tau^2}{\sigma_{kj}^2}\right)$ 를 알 수 있다.
한편, $|\sigma_{kj}|$ 는 이대일 변환이므로, 2가 곱해져서,

$\pi^u(|\sigma_{kj}| \geq (\zeta p)^{-1}) \leq \frac{1}{\zeta}\sqrt{\frac{2}{\pi^3}}\int_{(\zeta p)^{-1}}^{\infty} ln\left(1 + \frac{2\tau^2}{x^2}\right)dx$ 가 되고,
$\leq \sqrt{\frac{2}{\pi^3}}\int_{(\zeta p)^{-1}}^{\infty} \frac{2\tau^2}{x^2}dx\ (\because ln(1+x) \leq x,\ when\ x \geq 0)$
$= \sqrt{\frac{2}{\pi^3}}\int_{(\zeta p)^{-1}}^{\infty} \frac{2\tau}{x^2}dx = \frac{2\sqrt{2}}{\sqrt{\pi^3}}\tau\zeta p$ 임을 알 수 있다.

이를 통해 이제 $\pi^u(\max_{k\neq j}|\sigma_{kj}| < (\zeta p)^{-1})$ 를 계산해보면,

$$\pi^u(\max_{k\neq j}|\sigma_{kj}| < (\zeta p)^{-1}) = \prod_{k\neq j}\left\{1-\pi^u(|\sigma_{kj}| \geq (\zeta p)^{-1})\right\} = \left(1-\frac{2\sqrt{2}}{\sqrt{\pi^3}}\tau\zeta p\right)^{p(p-1)}$$

$\geq \left(1-\frac{2\sqrt{2}}{\sqrt{\pi^3}}\tau\zeta p\right)^{p^2} \cdots$ (1) ($\because$ 괄호안의 값은확률로 1보다 작으므로)

$\geq exp\left(-\frac{4\sqrt{2}}{\sqrt{\pi^3}}\tau\zeta p^3\right)$ ($\because log(1-x) \geq -2x,\ when\, x \leq \frac{1}{2}$) 이다.

한편, 주어진 조건에서 $\tau = O\left(\frac{1}{p^2}\sqrt{\frac{s_0 lnp}{n}}\right)$, $s_0 lnp = O(n)$ 이라 했으므로,

$\frac{\tau}{\frac{1}{p^2}\sqrt{\frac{s_o lnp}{n}}} \leq c_1, \frac{s_o lnp}{n} \leq c_2,\ c_1, c_2 > 0$ 임을 알 수 있다. 이들을 조합하면

$\Rightarrow\ \tau p^2 \leq \sqrt{c_2}c_1 := c_3 \ \Rightarrow \tau p^3 \leq c_3 p \ \Rightarrow\ -\tau p^3 \geq -c_3 p$ 이다. 이를 (1) 식에서 활용하면, $exp\left(-\frac{4\sqrt{2}}{\sqrt{\pi^3}}\tau\zeta p^3\right) \geq exp\left(-c_3\frac{4\sqrt{2}}{\sqrt{\pi^3}}\zeta p\right) = exp(-Cp)$ 이다. 따라서, 두번째 밑줄 확률의 부등식을 다음과 같이 구할 수 있다.

$\pi^u(\max_{k\neq j}|\sigma_{kj}| < (\zeta p)^{-1}) \geq exp(-Cp)$

이제 첫 번째 밑줄 식과 두 번째 밑줄 식의 결과를 종합하면,

$$\pi^u(\Sigma \in \mathcal{U}(\zeta)) \geq \pi^u(\zeta^{-1} \leq \min_j(\sigma_{jj}-\zeta^{-1}) \leq 2\max_j \sigma_{jj} \leq \zeta)\ *\ \pi^u(\max_{k\neq j}|\sigma_{kj}| < (\zeta p)^{-1})$$

$\geq \left\{\frac{\lambda\zeta}{8}exp(-\frac{\lambda\zeta}{4})\right\}^p exp(-Cp) = \left\{\frac{\lambda\zeta}{8}exp(-\frac{\lambda\zeta}{4}-C)\right\}^p$ 임을 알 수 있다.

□

2 오정훈 - Theorem 4 and Lemma 2

본 절에서는 Lee et al. (2022) 의 사후 수렴속도에 관한 증명인 **Lemma 1** 을 보이기 위하여 필요한 두번째 조건 **Theorem 4** 를 증명하고, **Lemma 4** 의 증명에 필요한 **Lemma 3** 를 증명한다.

Theorem 2.5 (The Upper Bound of the Packing Number)**.** *If* $\zeta^4 \leq p,\ p \asymp n^\beta$ *for some* $0<\beta<1,\ s_n = c_1 n\epsilon_n^2/\ln p,\ L_n = c_2 n\epsilon_n^2$ *and* $\delta_n = \epsilon_n/\zeta^3$ *for some constants* $c_1 > 1$ *and* $c_2 > 0,$ *we have*

$$\ln D(\epsilon_n, \mathcal{P}_n, d) \leq (12+1/\beta)c_1 n\epsilon_n^2.$$

Proof. 정의 $\mathcal{P}_n = \{f_\Sigma : |s(\Sigma,\delta_n)| \leq s_n,\ \zeta^{-1} \leq \lambda_{min}(\Sigma) \leq \lambda_{max}(\Sigma) \leq \zeta,\ ||\Sigma||_{max} \leq L_n\}$ 로부터, 가산반가법성에 의해 다음이 성립한다.

$$\pi(\mathcal{P}_n^c) \leq \pi(|s(\Sigma,\delta_n)| > s_n) + \pi(||\Sigma||_{max} > L_n) \tag{1}$$

부등식 (1)의 첫번째 항의 상계를 구해보자. 앞선 **Lemma 2**의 증명의 결과를 따라가는데, 모형 $\sigma_{kj}|\lambda \sim \mathcal{N}(0,\lambda^2),\ \lambda \sim C^+(0,\tau)$ 에 Carvalho et al. (2010) 의 **Theorem 1** 의 부등식을 적용하면

$$\pi^u(\sigma_{kj}) \leq \frac{1}{\tau\sqrt{2\pi^3}}\ln\left(1+\frac{2\tau^2}{\sigma_{kj}^2}\right), \quad 1 \leq k \neq j \leq p \tag{2}$$

가 성립하므로,

$$\pi^u(|\sigma_{kj}| > \delta_n) \leq \frac{2}{\tau\sqrt{2\pi^3}} \int_{\delta_n}^{\infty} \ln\left(1 + \frac{2\tau^2}{x^2}\right) dx \leq \frac{\sqrt{2}}{\tau\sqrt{\pi^3}} \int_{\delta_n}^{\infty} \frac{2\tau^2}{x^2} dx = \frac{2\sqrt{2}}{\sqrt{\pi^3}} \tau \delta_n^{-1}. \tag{3}$$

를 얻는다. 두번째 부등식에서는 $\ln(1+x) \leq x$ 를 이용하였다. 이제 (3)으로부터, 다음의 사실을 관찰할 수 있다.
적당한 상수 $C > 0$에 대해,

$$\nu_n \equiv \pi^u(|\sigma_{kj}| > \delta_n) \leq \frac{2\sqrt{2}}{\sqrt{\pi^3}} \tau \delta_n^{-1} = \frac{2\sqrt{2}}{\sqrt{\pi^3}} \tau \frac{\sqrt{n}\zeta^3}{\sqrt{(p+s_0)\ln p}} \leq \frac{C}{p^2}\sqrt{\frac{s_0 \ln p}{n}} \frac{\sqrt{n}\zeta^3}{\sqrt{(p+s_0)\ln p}} \lesssim \frac{1}{p^2} \tag{4}$$

이 충분히 큰 모든 n에 대해 성립한다.
두번째 등식은 가정 $\delta_n = \epsilon_n/\zeta^3$ 을 대입하였고, 세번째 부등식은 $\tau = O(\frac{1}{p^2}\sqrt{\frac{s_0 \ln p}{n}})$ 을 사용하였다. 이제까지의 결과를 정리하면, 우리는

$$\nu_n = \pi^u(|\sigma_{kj}| > \delta_n) \lesssim \frac{1}{p^2} \tag{5}$$

의 사실을 얻었다. 이제 위에서 정의된 것들을 살펴보자. ν_n 이 0이 아닌 비대각 원소 σ_{kj} 들의 비율이고, $|s(\Sigma, \delta_n)|$ 이 0이 아닌 σ_{kj} 의 개수라는 해석이 자연스럽다. 따라서 $|s(\Sigma, \delta_n)| \sim Bin(\binom{p}{2}, \nu_n)$ 을 따른다. 여기에 Song and Liang (2017) 의 **Lemma A.3** 의 결과를 이용하면,

$$\begin{aligned} \pi^u(|s(\Sigma, \delta_n)| > s_n) = \mathbb{P}\left(Bin(\binom{p}{2}, \nu_n) > s_n\right) &\leq 1 - \Phi\left(sign(s_n - \binom{p}{2}\nu_n)\sqrt{2\binom{p}{2}H(\nu_n, s_n/\binom{p}{2})}\right) \\ &= 1 - \Phi\left(\sqrt{2\binom{p}{2}H(\nu_n, s_n/\binom{p}{2})}\right) \end{aligned} \tag{6}$$

를 얻는다. (6)의 마지막 등식에서는 최대값이 평균보다 크다는 사실, $s_n \geq \binom{p}{2}\nu_n$ 을 사용하였다. Φ는 표준정규분포의 cdf이고, H 는 다음과 같이 정의된 함수이다.

$$H(\nu, \frac{k}{n}) := \frac{k}{n}\ln\left(\frac{k}{n\nu}\right) + \left(1 - \frac{k}{n}\right)\ln\left(\frac{1 - \frac{k}{n}}{1 - \nu}\right)$$

이제 표준정규분포의 cdf에 관한 부등식 $1 - \Phi(t) \leq \frac{1}{\sqrt{2\pi}}\frac{1}{t}e^{-\frac{t^2}{2}}$ 을 사용하면,

$$1 - \Phi\left(\sqrt{2\binom{p}{2}H(\nu_n, s_n/\binom{p}{2})}\right) \leq \frac{e^{-\binom{p}{2}H(\nu_n, s_n/\binom{p}{2})}}{\sqrt{2\pi}\sqrt{2\binom{p}{2}H(\nu_n, s_n/\binom{p}{2})}} \tag{7}$$

의 상계를 얻는다. (7)의 점근적 성질을 확인해보기 위해, 항 $\binom{p}{2}H(\nu_n, s_n/\binom{p}{2})$ 을 전개해보자.

$$\binom{p}{2}H(\nu_n, s_n/\binom{p}{2}) = s_n \ln\left(\frac{s_n}{\binom{p}{2}\nu_n}\right) + \left(\binom{p}{2} - s_n\right)\ln\left(\frac{\binom{p}{2} - s_n}{\binom{p}{2} - \binom{p}{2}\nu_n}\right) \tag{8}$$

(8)의 첫번째 항을 먼저 보면, 적당한 상수 $C > 0$이 존재하여

$$\begin{aligned} s_n \ln\left(\frac{s_n}{\binom{p}{2}\nu_n}\right) = s_n \ln\left(\frac{c_1(p+s_0)}{\binom{p}{2}\nu_n}\right) \geq s_n \ln(c_1 C(p+s_0)) \geq s_n \ln\left(\sqrt{p+s_0}\right) &\geq \frac{s_n}{2}\ln p \qquad (9) \\ &= \frac{c_1 n\epsilon_n^2}{2} \qquad (10) \end{aligned}$$

이 충분히 큰 모든 n에 대하여 성립함을 알 수 있다. (9)의 첫번째 등식은 $s_n = c_1 n\epsilon_n^2/\ln p = c_1(p+s_0)$를 이용하였고, 두번째 부등식에서는 (5)의 결과 $\nu_n \lesssim \frac{1}{p^2}$를 사용하였다. 다음으로 (8)의 두번째 항을 살펴보면

$$
\begin{aligned}
\left(\binom{p}{2} - s_n\right) \ln\left(\frac{\binom{p}{2} - s_n}{\binom{p}{2} - \binom{p}{2}\nu_n}\right) &= \left(\binom{p}{2} - s_n\right) \ln\left(1 - \frac{s_n - \binom{p}{2}\nu_n}{\binom{p}{2}(1-\nu_n)}\right) && (11)\\
&\geq -2\left(\binom{p}{2} - s_n\right) \frac{s_n - \binom{p}{2}\nu_n}{\binom{p}{2}(1-\nu_n)} && (12)\\
&= -2\left(1 - s_n/\binom{p}{2}\right) \frac{s_n - \binom{p}{2}\nu_n}{(1-\nu_n)} && (13)\\
&\geq -2\left(1 - s_n/\binom{p}{2}\right) \frac{s_n}{(1-\nu_n)} && (14)\\
&= -s_n\left(1 - s_n/\binom{p}{2}\right) \frac{2}{(1-\nu_n)} && (15)\\
&\gtrsim -s_n\left(1 - \frac{c_1(p+s_0)}{p^2}\right) && (16)\\
&\gtrsim -s_n = \frac{-c_1 n\epsilon_n^2}{\ln p} && (17)
\end{aligned}
$$

이 충분히 큰 모든 n에 대해 성립한다. 부등식 (12)는 $\ln(1-x) \geq -2x,\ \forall x \in (0, 1/2)$과 $\nu_n \lesssim 1/p^2$로부터 $\frac{s_n - \binom{p}{2}\nu_n}{\binom{p}{2}(1-\nu_n)} = O(\frac{p+s_0}{p^2}) \searrow 0,\ n \to \infty$임을 이용하였다. 부등식 (16)은 $s_n/\binom{p}{2} = c_1(p+s_0)/\binom{p}{2} = O(\frac{c_1(p+s_0)}{p^2})$인 사실과 $\nu_n \lesssim 1/p^2$로부터 충분히 큰 모든 n에 대해 $\frac{2}{1-\nu_n}$이 상수로 무시가능한 것으로 인해 성립하고, (17)번 부등식은 $1 - \frac{c_1(p+s_0)}{p^2} \to 1$ 때문에 성립한다.

(10)과 (17)의 결과를 가지고 (8)로 되돌아가면 다음을 관찰할수 있다.

$$
\binom{p}{2} H(\nu_n, s_n/\binom{p}{2}) \gtrsim \frac{c_1 n\epsilon_n^2}{2} - \frac{c_1 n\epsilon_n^2}{\ln p} = O(\frac{c_1 n\epsilon_n^2}{2}) \tag{18}
$$

즉 $\binom{p}{2} H(\nu_n, s_n/\binom{p}{2})$은 무한으로 발산하므로, 충분히 큰 모든 n에 대하여 (7) 우변의 분모를 생략할 수 있다. 그리고 적당히 큰 n에 대해,

$$
\binom{p}{2} H(\nu_n, s_n/\binom{p}{2}) \gtrsim \left(\frac{1}{2} - \frac{1}{\ln p}\right) c_1 n\epsilon_n^2 \sim \frac{1}{3} c_1 n\epsilon_n^2 \tag{19}
$$

역시 성립하므로 (18)과 (19)의 결과들을 (6), (7)과 결합하면, 충분히 큰 모든 n에 대하여

$$
\pi^u(|s(\Sigma, \delta_n)| > s_n) \leq \exp\left(-\frac{c_1 n\epsilon_n^2}{3}\right) \tag{20}
$$

이 성립함을 알 수 있다. 여태까지의 모든 결과를 종합하면 (1)에서 제시된 첫번째 항의 상계를 다음과 같이 구할 수 있다.

$$
\begin{aligned}
\pi(|s(\Sigma, \delta_n)| > s_n) &\leq \frac{\pi^u(|s(\Sigma, \delta_n)| > s_n)}{\pi^u(\Sigma \in \mathcal{U}(\zeta))} \leq \frac{\exp\left(-c_1 n\epsilon_n^2/3\right)}{\pi^u(\Sigma \in \mathcal{U}(\zeta))} && (21)\\
&\leq \left\{\frac{8}{\lambda\zeta} \exp\left(\frac{\lambda\zeta}{4} + C\right)\right\}^p \exp\left(-c_1 n\epsilon_n^2/3\right) && (22)\\
&= \exp\left(p\ln 8 - p\ln(\lambda\zeta) + p\frac{\lambda\zeta}{4} + Cp - \frac{c_1 n\epsilon_n^2}{3}\right) && (23)\\
&\lesssim \exp\left(p\ln p - \frac{c_1 n\epsilon_n^2}{3}\right) && (24)\\
&\sim \exp\left(-\frac{(c_1-1)n\epsilon_n^2}{3}\right) && (25)
\end{aligned}
$$

이 충분히 큰 모든 n에 대해 성립한다. (21)의 두번째 부등식에서는 (20)의 결과를, 부등식 (22)에서는 Lee et al. (2022)의 **Lemma 2**를, (24)에서는 $\lambda\zeta \le \ln p$의 가정을 사용하였다.
마지막으로 (1)의 두번째 항인 $\pi(||\Sigma||_{max} > L_n)$의 상계를 구하여보자. 먼저 $||\Sigma||_{max} \le \lambda_{max}(\Sigma)$ 가 성립하는데, 이는 다음으로부터 알 수 있다. 실수인 행렬 A가 양의 준정부호 혹은 정부호 행렬일때,

$$\lambda_{max}(A) = \max_{||\mathbf{u}||=1} \mathbf{u}^\top A\mathbf{u} \ge \mathbf{e}_i^\top A\mathbf{e}_i, \quad \forall i = 1, \cdots, p \tag{26}$$

가 성립. $\{\mathbf{e}_1, \cdots, \mathbf{e}_p\}$는 $\mathbb{R}^p$의 표준기저이다. 그런데 A가 대칭이며 양의 준정부호이면 A의 최대원소는 반드시 대각원소중에 있으므로 $||\Sigma||_{max} \le \lambda_{max}(\Sigma)$. 그리고 $\pi(\Sigma : \lambda_{max}(\Sigma) \le \zeta) = 1$으로부터, 충분히 큰 모든 n에 대하여 $L_n > \zeta$이면,

$$0 = \pi(\lambda_{max}(\Sigma) > L_n) \ge \pi(||\Sigma||_{max} > L_n) \tag{27}$$

이 성립한다. 그러므로 (25)와 (27)에 의해 (1)은

$$\begin{aligned}
\pi(\mathcal{P}_n^c) &\le \pi(|s(\Sigma, \delta_n)| > s_n) + \pi(||\Sigma||_{max} > L_n) && (28)\\
&\le \exp\left(-\frac{(c_1 - 1)n\epsilon_n^2}{3}\right) && (29)
\end{aligned}$$

가 충분히 큰 모든 n에 대해 성립한다. 이로써 증명이 끝났다. □

Lemma 2.6. *If* $\Sigma_0 \in \mathcal{U}(s_0, \zeta_0)$ *and* $\Sigma \in \mathcal{U}(\zeta)$, *we have*

(*i*) $K(f_{\Sigma_0}, f_\Sigma) \le c_0\zeta^4\zeta_0^2||\Sigma - \Sigma_0||_F^2$ *for some* $c_0 > 0$

(*ii*) $V(f_{\Sigma_0}, f_\Sigma) \le \frac{3}{2}\zeta^4\zeta_0^2||\Sigma - \Sigma_0||_F^2$ *for sufficiently small* $||\Sigma - \Sigma||_F^2 \le \frac{1}{c_0^2\zeta^4\zeta_0^2}$.

Proof. **Lemma 3**의 증명은 Banerjee and Ghosal (2015)의 계산을 따라간다. 먼저 다음을 확인하자. $A = \Sigma_0^{\frac{1}{2}}\Sigma^{-1}\Sigma_0^{\frac{1}{2}}$ 라 놓고, d_i를 A의 고유값이라 하자. $d_i = \lambda_i(A),\ i = 1, \cdots, p$. 또, $D = diag(d_i,\ i = 1, \cdots, p)$라 놓는다. 그러면

$$||I - A||_F^2 = tr\left[(I - A)^2\right] = tr\left[U(I - D)^2U^\top\right] = tr\left[(I - D)^2\right] = \sum_{i=1}^p (1 - d_i)^2, \quad UU^\top = U^\top U = I_p$$

이다. 첫번째 등식은 $||B||_F^2 = tr\left(B^\top B\right) = tr\left(BB^\top\right)$ 임을, 두번째 등식은 A의 고유치 분해를 사용하였다. 이로써

$$\begin{aligned}
\sum_{i=1}^n (1 - d_i)^2 &= ||I - A||_F^2 && (30)\\
&= ||\Sigma_0^{\frac{1}{2}}(\Sigma_0^{-1} - \Sigma^{-1})\Sigma_0^{\frac{1}{2}}||_F^2 && (31)\\
&\le ||\Sigma_0^{\frac{1}{2}}||^2\ ||\Sigma_0^{\frac{1}{2}}||^2\ ||\Sigma_0^{-1} - \Sigma^{-1}||_F^2 && (32)\\
&= \lambda_{max}(\Sigma_0) \cdot \lambda_{max}(\Sigma_0) \cdot ||\Sigma_0^{-1} - \Sigma^{-1}||_F^2 && (33)\\
&\le \zeta_0^2 \cdot ||\Sigma_0^{-1} - \Sigma^{-1}||_F^2 && (34)\\
&\le \zeta^2 \cdot ||\Sigma_0^{-1} - \Sigma^{-1}||_F^2 && (35)
\end{aligned}$$

가 성립함을 알 수 있다. (31) 부등식은 아래의 **Lemma 5**의 결과를 이용하였고, 부등식 (35)는 Lee et al. (2022) p.5의 가정 **A3.** $\zeta > \max(3, \zeta_0)$로부터 성립한다. 위와 비슷한 방식으로,

$$||\Sigma_0^{-1} - \Sigma^{-1}||_F \le ||\Sigma^{-1}||\ ||\Sigma_0^{-1}||\ ||\Sigma - \Sigma_0||_F \le \zeta\zeta_0||\Sigma - \Sigma||_F \tag{36}$$

가 성립한다. 먼저 (i)의 쿨백 라이블러 발산을 계산한다. $\mathbb{E}_0$를 밀도함수 f_{Σ_0}에 대한 기대값이라 하자.

$$\begin{aligned}
K(f_{\Sigma_0}, f_\Sigma) &= \mathbb{E}_0\left[-\frac{1}{2}\ln\frac{|\Sigma_0|}{|\Sigma|} - \frac{1}{2}x^\top(\Sigma_0^{-1} - \Sigma^{-1})x\right], \quad x \sim \mathcal{N}(0, \Sigma_0) && (37)\\
&= -\frac{1}{2}\ln\frac{|\Sigma_0|}{|\Sigma|} - \frac{1}{2}tr\left[(\Sigma_0^{-1} - \Sigma^{-1})\Sigma_0\right] && (38)\\
&= -\frac{1}{2}\ln\left(|\Sigma_0^{\frac{1}{2}}||\Sigma|^{-1}|\Sigma_0^{\frac{1}{2}}|\right) - \frac{1}{2}tr(I - \Sigma^{-1}\Sigma_0) && (39)\\
&= -\frac{1}{2}\ln|A| - \frac{1}{2}tr\left[\Sigma_0^{\frac{1}{2}}(I - A)\Sigma_0^{-\frac{1}{2}}\right] && (40)\\
&= -\frac{1}{2}\ln|A| - \frac{1}{2}tr(I - A) && (41)\\
&= -\frac{1}{2}\sum_{i=1}^{p}\ln d_i - \frac{1}{2}\sum_{i=1}^{p}(1 - d_i) && (42)
\end{aligned}$$

등식 (38)에서는 다변량 정규분포의 이차형식의 기대값에 관한 공식을 사용하였다. 다음으로 (ii)의 쿨백 라이블러 분산을 계산해보면

$$\begin{aligned}
V(f_{\Sigma_0}, f_\Sigma) &= \mathbb{E}_0\left[\frac{1}{4}\left\{\ln\frac{|\Sigma_0|}{|\Sigma|} + x^\top(\Sigma_0^{-1} - \Sigma^{-1})x\right\}^2\right], \quad z \overset{let}{=} x^\top(\Sigma_0^{-1} - \Sigma^{-1})x && (43)\\
&= \frac{1}{4}\,\mathbb{E}_0\left[\left\{\ln\frac{|\Sigma_0|}{|\Sigma|} + \mathbb{E}_0(z) + z - \mathbb{E}_0(z)\right\}^2\right] && (44)\\
&= \frac{1}{4}\left[4\big\{K(f_{\Sigma_0}, f_\Sigma)\big\}^2 + Var(z)\right] && (45)\\
&= K^2(f_{\Sigma_0}, f_\Sigma) + \frac{1}{2}tr\left[(\Sigma_0^{-1} - \Sigma^{-1})\Sigma_0(\Sigma_0^{-1} - \Sigma^{-1})\Sigma_0\right] && (46)\\
&= K^2(f_{\Sigma_0}, f_\Sigma) + \frac{1}{2}tr\left[\Sigma_0^{\frac{1}{2}}(\Sigma_0^{-1} - \Sigma^{-1})\Sigma_0^{\frac{1}{2}}\Sigma_0^{\frac{1}{2}}(\Sigma_0^{-1} - \Sigma^{-1})\Sigma_0^{\frac{1}{2}}\right] && (47)\\
&= K^2(f_{\Sigma_0}, f_\Sigma) + \frac{1}{2}tr\big[(I - A)^2\big] && (48)\\
&= K^2(f_{\Sigma_0}, f_\Sigma) + \frac{1}{2}\sum_{i=1}^{p}(1 - d_i)^2 && (49)
\end{aligned}$$

등식 (45)에서는 등식 (37)과 다변량 정규분포의 이차형식의 분산에 관한 공식을 사용하였다. 이제 부등식 (i)을 보인다.

$$\begin{aligned}
K(f_{\Sigma_0}, f_\Sigma) &= -\frac{1}{2}\sum_{i=1}^{p}\ln d_i - \frac{1}{2}\sum_{i=1}^{p}(1 - d_i) && (50)\\
&\leq c_0\sum_{i=1}^{p}(1 - d_i)^2 \quad \text{for some } c_0 > 0 && (51)\\
&\leq c_0\zeta^2||\Sigma_0^{-1} - \Sigma^{-1}||_F^2 && (52)\\
&\leq c_0\zeta^4\zeta_0^2||\Sigma_0 - \Sigma||_F^2 && (53)
\end{aligned}$$

부등식 (51)은 아래의 **참고 1**을 보라. 부등식 (52)는 (35), 부등식 (53)은 (36)의 결과를 이용하였다.

Remark 2.1. 먼저, 다음과 같은 부등식이 성립함을 알 수 있는데,

$$-\frac{1}{2}\ln x - \frac{1}{2}(1 - x) \leq (1 - x)^2, \ \forall x \in [\eta_0, 1] \ \ \textit{for small } \eta_0 > 0.$$

여기서 좌변과 우변의 두 함수는 각각 η_0 과 1 에서 만남을 확인가능하다. 또, A의 고유값 d_i 에 대하여

$$d_i = \lambda_i(A) = \lambda_i(\Sigma^{-1}\Sigma_0) \geq \lambda_{min}(\Sigma_0)\lambda_{min}(\Sigma^{-1}) \geq \zeta_0^{-1}\zeta^{-1}$$

이 성립함을 알 수 있는데, 다음 등식을 만족하는 c_0 를 잡아오는 것을 생각해보자. $x \neq 1$ 이면

$$c_0(1-x)^2 = -\frac{1}{2}\ln x - \frac{1}{2}(1-x) \iff c_0 = -\frac{\ln x + 1 - x}{2(1-x)^2} > 0$$

위에서 $d_i \geq \zeta_0^{-1}\zeta^{-1}$ 을 알고 있으므로

$$c_0 = -\frac{\ln \zeta_0^{-1}\zeta^{-1} + 1 - \zeta_0^{-1}\zeta^{-1}}{2(1-\zeta_0^{-1}\zeta^{-1})^2} > 0$$

과 같이 잡으면, (*51*) 이 성립한다.

마지막으로, 부등식 (ii)를 보인다. 위에서 보인 결과들을 종합하면

$$\begin{aligned} V(f_{\Sigma_0}, f_\Sigma) &= \frac{1}{2}\sum_{i=1}^{p}(1-d_i)^2 + K^2(f_{\Sigma_0}, f_\Sigma) && (54) \\ &\leq \frac{1}{2}\sum_{i=1}^{p}(1-d_i)^2 + \zeta^4\zeta_0^2||\Sigma - \Sigma_0||_F^2 && (55) \\ &\leq \frac{3}{2}\zeta^4\zeta_0^2||\Sigma - \Sigma_0||_F^2 && (56) \end{aligned}$$

가 충분히 작은 $||\Sigma - \Sigma_0||_F^2 \leq 1/(c_0^2\zeta^4\zeta_0^2)$ 에 대해 성립한다.
부등식 (55) 는 위에서 보인 결과 (i) 로부터 $||\Sigma - \Sigma_0||_F^2 \leq 1/(c_0^2\zeta^4\zeta_0^2)$ 로 충분히 작으면, $K^2(f_{\Sigma_0}, f_\Sigma) \leq c_0^2\zeta^8\zeta_0^4||\Sigma - \Sigma_0||_F^4 \leq \zeta^4\zeta_0^2||\Sigma - \Sigma_0||_F^2$ 인 것을 사용하였고, 부등식 (56) 은 부등식 (35) 와 (36) 의 결과에 의해 성립한다. 이로써 **Lemma 3** 가 증명되었다.

□

Lemma 2.7 (Lee et al. (2022)). *For any $p \times p$ matrices A and B, we have*

$$||AB||_F^2 \leq ||A||\,||B||_F, \quad ||AB||_F^2 \leq ||A||_F\,||B||$$

Proof. $\mathbf{b}_j$ 가 행렬 B 의 j 번째 열이라고 하자. 그러면

$$||AB||_F^2 = ||(A\mathbf{b}_1, \cdots, A\mathbf{b}_p)||_F^2 = \sum_{j=1}^{p}||A\mathbf{b}_j||_2^2 \leq ||A||^2\sum_{j=1}^{p}||\mathbf{b}_j||_2^2 = ||A||^2\,||B||_F^2$$

마찬가지로 $\mathbf{a}_i^\top$ 가 행렬 A 의 i 번째 행이라고 하면

$$||AB||_F^2 = ||(\mathbf{a}_1^\top B, \cdots, \mathbf{a}_p^\top B)||_F^2 = \sum_{i=1}^{p}||\mathbf{a}_i^\top B||_2^2 \leq ||B||^2\sum_{i=1}^{p}||\mathbf{a}_i||_2^2 = ||B||^2\,||A||_F^2$$

□

3 장태영 - Lemma 4 and Theorem 5

Lemma (Lemma 3 in the paper). *If $\Sigma_0 \in \mathcal{U}(s_0, \zeta_0)$ and $\Sigma \in \mathcal{U}(\zeta)$ then we have*

1. $K(f_{\Sigma_0}, f_\Sigma) \leq \zeta^4\zeta_0^2\|\Sigma - \Sigma_0\|_F^2$
2. $V(f_{\Sigma_0}, f_\Sigma) \leq \frac{3}{2}\zeta^4\zeta_0^2\|\Sigma - \Sigma_0\|_F^2$

Lemma (Lemma 4 in the paper)**.** *If $a = b = 1/2$, $x > 1$, and $\tau/x > 0$ is sufficiently small then*

$$\pi^u_{ij}(x) \geq \sqrt{\frac{1}{2\pi}\frac{\tau}{x^2}}$$

where $\pi^u_{ij}(\sigma_{ij})$ is the unconstrained marginal prior density of σ_{ij}

Theorem (Theorem 5 in the paper ; The lower bound for $\pi(B_{\varepsilon_n})$)**.** *Here are the conditions we need for this theorem.*

1. $\Sigma_0 \in \mathcal{U}(s_0, \zeta_0)$ *with* $\zeta_0 < \zeta$
2. $p \asymp n^\beta$ *for some* $0 < \beta < 1$
3. $\zeta^4 \leq p$
4. $\zeta^2\zeta_0^2 \leq s_0 \log p$
5. $n \geq \max\{1/\zeta_0^4 \, , \, s_0/(1-\zeta_0/\zeta)^2\} \log p/\zeta^4$
6. $p^{-1} < \lambda < \log p/\zeta_0$
7. $a = b = 1/2$
8. $(p^2\sqrt{n})^{-1} \lesssim \tau \lesssim (p^2\sqrt{n})^{-1}\sqrt{s_0 \log p}$
9. (*Additional, From Thm 1 at page 5 of the paper*) $(p + s_0)\log p = o(n)$ *i.e.* $\varepsilon_n^2 \to 0$
10. (*Additional, From page 4 of the paper*) $p = O(s_0)$

If the conditions above hold, then we have

$$\pi(B_{\varepsilon_n}) \geq \exp\Big\{ -(5 + \frac{1}{\beta})n\varepsilon_n^2 \Big\}$$

Proof of Lemma 4

Proof. Because we have $a = b = 1/2$,

$$\sigma_{ij}|\rho_{ij} \sim N(0, \frac{\rho_{ij}}{1-\rho_{ij}}\tau^2) \ , \ \rho_{ij} \sim \text{Beta}(a, b)$$

is equivalent to

$$\sigma_{ij}|\lambda_{ij} \sim N(0, \lambda_{ij}^2\tau^2) \ , \ \lambda_{ij} \sim \text{C}^+(0, 1)$$

where $\text{C}^+(0, s)$ denotes the standard half-Cauchy distribution on positive real with a scale parameter s.

$$p(\sigma, \rho) = p(\sigma|\rho)p(\rho) = \frac{1}{\sqrt{2\pi\frac{\rho}{1-\rho}\tau^2}} \exp\Big(-\frac{1}{2\frac{\rho}{1-\rho}\tau^2}\sigma^2 \Big) \frac{1}{\pi}\rho^{-1/2}(1-\rho)^{-1/2} \quad \because \ \Gamma(1/2) = \sqrt{\pi}$$

$$\lambda = \sqrt{\frac{\rho}{1-\rho}} \quad (\lambda > 0) \quad \text{and} \quad \rho = \frac{\lambda^2}{\lambda^2 + 1}$$

$$\text{Jacobian} = \left|\frac{d\rho}{d\lambda}\right| = \frac{2\lambda}{(\lambda^2 + 1)^2}$$

$$
\begin{aligned}
p(\sigma,\lambda) &= \frac{1}{\sqrt{2\pi\lambda^2\tau^2}}\exp\Big(-\frac{1}{2\lambda^2\tau^2}\sigma^2\Big)\frac{1}{\pi}\sqrt{\frac{\lambda^2}{\lambda^2+1}\frac{1}{\lambda^2+1}}^{-1}\frac{2\lambda}{(\lambda^2+1)^2} \\
&= \frac{1}{\sqrt{2\pi\lambda^2\tau^2}}\exp\Big(-\frac{1}{2\lambda^2\tau^2}\sigma^2\Big)\frac{2}{\pi}\frac{\lambda^2+1}{\lambda}\frac{\lambda}{(\lambda^2+1)^2} \\
&= \frac{1}{\sqrt{2\pi\lambda^2\tau^2}}\exp\Big(-\frac{1}{2\lambda^2\tau^2}\sigma^2\Big)\frac{2}{\pi}\frac{1}{(\lambda^2+1)} \\
&= p(\sigma|\lambda)p(\lambda)
\end{aligned}
$$

Hence we can conclude that

$$\sigma|\rho \sim N(0, \frac{\rho}{1-\rho}\tau^2)\ ,\ \rho \sim \text{Beta}(1/2, 1/2)$$

is equivalent to

$$\sigma|\lambda \sim N(0, \lambda^2\tau^2)\ ,\ \lambda \sim C^+(0,1)$$

Now we shall derive tight bound for marginal prior density of σ

$$
\begin{aligned}
\pi^u(\sigma) &= \int_0^\infty p(\sigma,\lambda)\,d\lambda \\
&= \int_0^\infty \frac{1}{\sqrt{2\pi\lambda^2\tau^2}}\exp\Big(-\frac{1}{2\lambda^2\tau^2}\sigma^2\Big)\frac{2}{\pi}\frac{1}{(\lambda^2+1)}\,d\lambda \\
&\text{change of variable} : u = 1/\lambda^2 \Leftrightarrow \lambda = u^{-1/2} \quad d\lambda = -\frac{1}{2}u^{-3/2}\,du \\
&= \int_0^\infty \frac{1}{\sqrt{2\pi\tau^2}}u^{1/2}\exp\Big(-\frac{1}{2\tau^2}\sigma^2 u\Big)\frac{2}{\pi}\frac{u}{1+u}\frac{1}{2}u^{-3/2}\,du \\
&= \int_0^\infty \frac{1}{\sqrt{2\pi^3\tau^2}}\exp\Big(-\frac{\sigma^2}{2\tau^2}u\Big)\frac{1}{1+u}\,du \\
&\text{change of variable} : z = 1+u \Leftrightarrow u = z-1 \\
&= \frac{1}{\tau\sqrt{2\pi^3}}\exp\Big(\frac{\sigma^2}{2\tau^2}\Big)\int_1^\infty \frac{1}{z}\exp\Big(-\frac{\sigma^2}{2\tau^2}z\Big)\,dz
\end{aligned}
$$

Define exponential integral E_1 as the following :

$$E_1(x) = \int_1^\infty \frac{1}{z}\exp(-zx)\,dz \quad \forall\, x > 0$$

Then it has tight bound given as

$$\frac{1}{2}\exp(-x)\log\Big(1+\frac{2}{x}\Big) < E_1(x) < \exp(-x)\log\Big(1+\frac{1}{x}\Big) \quad x > 0$$

Note that this tight bound is mentioned in Wikipedia : Exponential integral

Thus, we have

$$\pi^u(\sigma) = \frac{1}{\tau\sqrt{2\pi^3}}\exp\Big(\frac{\sigma^2}{2\tau^2}\Big)E_1\Big(\frac{\sigma^2}{2\tau^2}\Big)$$

Using the tight bound of E_1 given above, we get

$$
\begin{aligned}
\pi^u(\sigma) &< \frac{1}{\tau\sqrt{2\pi^3}}\log\Big(1+\frac{2\tau^2}{\sigma^2}\Big) \\
\pi^u(\sigma) &> \frac{1}{2\tau\sqrt{2\pi^3}}\log\Big(1+\frac{4\tau^2}{\sigma^2}\Big)
\end{aligned}
$$

From now on, we will call these two inequalities as upper and lower bound of marginal prior density of σ_{ij} respectively.

Then, using lower bound of marginal prior density of σ_{ij}, we have the following :

$$\begin{aligned}
\pi^u_{ij}(x) &\geq \frac{1}{2\tau}\sqrt{\frac{1}{2\pi^3}}\log\Big(1+\frac{4\tau^2}{x^2}\Big) \\
&\geq \frac{1}{4\tau}\sqrt{\frac{1}{2\pi^3}}\frac{4\tau^2}{x^2} \quad \because \log(1+x)\geq\frac{1}{2}x \quad \text{when } 0\leq x\leq 1 \quad \text{and} \quad \tau/x \text{ is suff. small} \\
&= \sqrt{\frac{1}{2\pi^3}}\frac{\tau}{x^2}
\end{aligned}$$

□

Proof of Theorem 5

Proof. Note that B_ε is defined as

$$B_\varepsilon = \{f_\Sigma : \Sigma\in\mathcal{C}_p,\ K(f_{\Sigma_0}, f_\Sigma) < \varepsilon^2,\ V(f_{\Sigma_0}, f_\Sigma) < \varepsilon^2\}$$

By Lemma 3, it suffices to show that

$$\pi\Big(\|\Sigma-\Sigma_0\|_F^2 \leq \frac{2}{3\zeta^4\zeta_0^2}\varepsilon_n^2\Big) \geq \exp\big(-Cn\varepsilon_n^2\big)$$

This is because

$$\begin{aligned}
&\|\Sigma-\Sigma_0\|_F^2 \leq \frac{2}{3\zeta^4\zeta_0^2}\varepsilon_n^2 \\
&\Rightarrow K(f_{\Sigma_0,f_\Sigma}) \leq \zeta^4\zeta_0^2\|\Sigma-\Sigma_0\|_F^2 \leq \frac{2}{3}\varepsilon_n^2 < \varepsilon_n^2 \quad \text{and} \quad V(f_{\Sigma_0}, f_\Sigma) \leq \frac{3}{2}\zeta^4\zeta_0^2\|\Sigma-\Sigma_0\|_F^2 \leq \varepsilon_n^2 \\
&\Rightarrow f_\Sigma \in B_{\varepsilon_n}
\end{aligned}$$

so that $\pi(B_{\varepsilon_n}) \geq \pi\Big(\|\Sigma-\Sigma_0\|_F^2 \leq \frac{2}{3\zeta^4\zeta_0^2}\varepsilon_n^2\Big)$

Note that

$$\begin{aligned}
&\pi\Big(\|\Sigma-\Sigma_0\|_F^2 \leq \frac{2}{3\zeta^4\zeta_0^2}\varepsilon_n^2\Big) = \pi\Big(\|\Sigma-\Sigma_0\|_F^2 \leq \frac{2}{3\zeta^4\zeta_0^2}\frac{(p+s_0)\log p}{n}\Big) \\
&\geq \pi\Big(\sum_{i\neq j}(\sigma_{ij}-\sigma^*_{ij})^2 \leq \frac{2}{3\zeta^4\zeta_0^2}\frac{s_0\log p}{n}\ ,\ \sum_{j=1}^p(\sigma_{jj}-\sigma^*_{jj})^2 \leq \frac{2}{3\zeta^4\zeta_0^2}\frac{p\log p}{n}\Big) \quad \because\ x\leq\alpha, y\leq\gamma \Rightarrow x+y\leq\alpha+\gamma \\
&\geq \pi\Big(\max_{i\neq j}(\sigma_{ij}-\sigma^*_{ij})^2 \leq \frac{2}{3\zeta^4\zeta_0^2}\frac{s_0\log p}{p(p-1)n}\ ,\ \max_{1\leq j\leq p}(\sigma_{jj}-\sigma^*_{jj})^2 \leq \frac{2}{3\zeta^4\zeta_0^2}\frac{\log p}{n}\Big) := \pi(A_{n,\Sigma_0})
\end{aligned}$$

where $\Sigma_0 = (\sigma^*_{ij})$

We will introduce Weyl's theorem here. (Source : Wikipedia : Weyl's inequality)

If A, B are $n\times n$ symmetric (or Hermitian) matrices then

$$\lambda_k(A)+\lambda_n(B) \leq \lambda_k(A+B) \leq \lambda_k(A)+\lambda_1(B) \quad \forall\ k=1,\cdots,n$$

where $\lambda_1(M)\geq\cdots\geq\lambda_n(M)$ are eigenvalues of symmetric matrix $M\in\mathbb{R}^{n\times n}$

Here, we will plug in $A=\Sigma_0$, $B=\Sigma-\Sigma_0$ so that $A+B=\Sigma$

Also, we will use two more properties about matrix norm.

The first one is that for symmetric A, we have $-\|A\|_2 \leq \lambda(A) \leq \|A\|_2$ where $\lambda(A)$ is any eigenvalue of A. This is because

$$\lambda(A)^2 = \lambda(A^2) = \lambda(A^TA) \Rightarrow |\lambda(A)| = \sqrt{\lambda(A^TA)} \leq \|A\|_2$$

The second one is the special case of the Hölder inequality $\|A\|_2 \leq \sqrt{\|A\|_1\|A\|_\infty}$ (Source : Wikipedia : Matrix Norm) Also, if A is symmetric, then $\|A\|_1 = \|A\|_\infty$ since the former is maximum absolute column sum and the latter is maximum absolute row sum. Thus we get $\|A\|_2 \le \|A\|_1$ given A is symmetric.

We want to show that

$$\Sigma \in A_{n,\Sigma_0} \Rightarrow \Sigma \in \mathcal{U}(\zeta)$$

Suppose $\Sigma \in A_{n,\Sigma_0}$. Then we have

$$\|\Sigma - \Sigma_0\|_1 \le (p-1)\max_{i\neq j}\left|\sigma_{ij} - \sigma_{ij}^*\right| + \max_{1\le j\le p}\left|\sigma_{jj} - \sigma_{jj}^*\right|$$

$$\begin{aligned}
\lambda_{min}(\Sigma) &\ge \lambda_{min}(\Sigma_0) + \lambda_{min}(\Sigma - \Sigma_0) \quad \because \text{ Weyl's thm} \\
&\ge \lambda_{min}(\Sigma_0) - \|\Sigma - \Sigma_0\|_2 \quad \because -\|A\|_2 \le \lambda(A) \le \|A\|_2 \\
&\ge \lambda_{min}(\Sigma_0) - \|\Sigma - \Sigma_0\|_1 \quad \because \|A\|_2 \le \|A\|_1 \text{ given } A \text{ is symmetric} \\
&\ge \zeta_0^{-1} - \left\{(p-1)\sqrt{\frac{2}{3\zeta^4\zeta_0^2}\frac{s_0\log p}{p(p-1)n}} + \sqrt{\frac{2}{3\zeta^4\zeta_0^2}\frac{\log p}{n}}\right\} \quad \because \Sigma_0 \in \mathcal{U}(s_0,\zeta_0)\ ,\ \Sigma \in A_{n,\Sigma_0} \\
&:= \zeta_0^{-1} - \star
\end{aligned}$$

$$\begin{aligned}
\lambda_{max}(\Sigma) &\le \lambda_{max}(\Sigma_0) + \lambda_{max}(\Sigma - \Sigma_0) \\
&\le \lambda_{max}(\Sigma_0) + \|\Sigma - \Sigma_0\|_2 \\
&\le \lambda_{max}(\Sigma_0) + \|\Sigma - \Sigma_0\|_1 \\
&\le \zeta_0 + \left\{(p-1)\sqrt{\frac{2}{3\zeta^4\zeta_0^2}\frac{s_0\log p}{p(p-1)n}} + \sqrt{\frac{2}{3\zeta^4\zeta_0^2}\frac{\log p}{n}}\right\} \\
&= \zeta_0 + \star
\end{aligned}$$

We shall claim that $\star \to 0$ as $n \to \infty$

$$\star \le \sqrt{\frac{2}{3\zeta^4\zeta_0^2}}\sqrt{\frac{(s_0+1)\log p}{n}} \le \sqrt{\frac{2}{3\zeta^4\zeta_0^2}}\sqrt{\frac{(s_0+p)\log p}{n}} \to 0 \quad \because \varepsilon_n \to 0$$

Thus, combining the fact that $\zeta_0 < \zeta$ and $\star \to 0$, we get

$$\lambda_{min}(\Sigma) \ge \zeta_0^{-1} - \star > \zeta^{-1} \quad \text{and} \quad \lambda_{max}(\Sigma) \le \zeta_0 + \star < \zeta \quad \text{for all suff. large } n$$

Hence, we have shown that

$$\Sigma \in A_{n,\Sigma_0} \Rightarrow \Sigma \in \mathcal{U}(\zeta)$$

as desired.

Using above, we get $\pi(A_{n,\Sigma_0}) \ge \pi^u(A_{n,\Sigma_0})$ since

$$\pi(A_{n,\Sigma_0}) = \frac{\pi^u(A_{n,\Sigma_0})\mathrm{I}(\Sigma \in \mathcal{U}(\zeta))}{\pi^u(\Sigma \in \mathcal{U}(\zeta))} = \frac{\pi^u(A_{n,\Sigma_0})}{\pi^u(\Sigma \in \mathcal{U}(\zeta))} \ge \pi^u(A_{n,\Sigma_0}) \quad \because \pi^u(\Sigma \in \mathcal{U}(\zeta)) \le 1$$

Here, we shall briefly check what we have already shown.

$$\pi(B_{\varepsilon_n}) \ge \pi\Big(\|\Sigma - \Sigma_0\|_F^2 \le \frac{2}{3\zeta^4\zeta_0^2}\varepsilon_n^2\Big) \ge \pi(A_{n,\Sigma_0}) \ge \pi^u(A_{n,\Sigma_0})$$

Hence, from now on, our goal is to prove that

$$\pi^u(A_{n,\Sigma_0}) \ge \exp\big(-Cn\varepsilon_n^2\big)$$

Note that

$$\pi^u(A_{n,\Sigma_0}) = \pi^u\Big(\max_{i\neq j}(\sigma_{ij}-\sigma^*_{ij})^2 \le \frac{2}{3\zeta^4\zeta_0^2}\frac{s_0\log p}{p(p-1)n}\ ,\ \max_{1\le j\le p}(\sigma_{jj}-\sigma^*_{jj})^2 \le \frac{2}{3\zeta^4\zeta_0^2}\frac{\log p}{n}\Big)$$
$$= \pi^u\Big(\max_{i\neq j}(\sigma_{ij}-\sigma^*_{ij})^2 \le \frac{2}{3\zeta^4\zeta_0^2}\frac{s_0\log p}{p(p-1)n}\Big)\times\pi^u\Big(\max_{1\le j\le p}(\sigma_{jj}-\sigma^*_{jj})^2 \le \frac{2}{3\zeta^4\zeta_0^2}\frac{\log p}{n}\Big)$$
$$= \prod_{i<j}\pi^u\Big((\sigma_{ij}-\sigma^*_{ij})^2 \le \frac{2}{3\zeta^4\zeta_0^2}\frac{s_0\log p}{p(p-1)n}\Big)\times\prod_{j=1}^{p}\pi^u\Big((\sigma_{jj}-\sigma^*_{jj})^2 \le \frac{2}{3\zeta^4\zeta_0^2}\frac{\log p}{n}\Big)$$

This is because all elements of Σ are independent to each other given unconstrained setting.

Observe $\prod_{j=1}^{p}\pi^u\Big((\sigma_{jj}-\sigma^*_{jj})^2 \le \frac{2}{3\zeta^4\zeta_0^2}\frac{\log p}{n}\Big)$ first. We want to find a lower bound of this term.

$$\prod_{j=1}^{p}\pi^u\Big((\sigma_{jj}-\sigma^*_{jj})^2 \le \frac{2}{3\zeta^4\zeta_0^2}\frac{\log p}{n}\Big) = \prod_{j=1}^{p}\pi^u\Big(\big|\sigma_{jj}-\sigma^*_{jj}\big| \le \sqrt{\psi}\Big)\quad \text{where } \psi := \frac{2}{3\zeta^4\zeta_0^2}\frac{\log p}{n}$$
$$= \prod_{j=1}^{p}\pi^u(\sigma^*_{jj}-\sqrt{\psi}\le\sigma_{jj}\le\sigma^*_{jj}+\sqrt{\psi})\quad \because\ \psi\to 0 \text{ so that } \sigma^*_{jj}-\sqrt{\psi}\ge 0 \text{ for all suff. large } n$$

Note that $\sigma_{jj}\sim\Gamma(1,\lambda/2)$ and $\pi^u(\sigma_{jj}) = \frac{\lambda}{2}\exp\Big(-\frac{\lambda}{2}\sigma_{jj}\Big)$ is decreasing function

$$\ge \prod_{j=1}^{p}2\sqrt{\psi}\,\pi^u(\sigma^*_{jj}+\sqrt{\psi}) = \prod_{j=1}^{p}2\sqrt{\psi}\frac{\lambda}{2}\exp\Big(-\frac{\lambda}{2}(\sigma^*_{jj}+\sqrt{\psi})\Big) = \prod_{j=1}^{p}\sqrt{\psi}\lambda\exp\Big(-\frac{\lambda}{2}(\sigma^*_{jj}+\sqrt{\psi})\Big)$$
$$\ge \prod_{j=1}^{p}\sqrt{\psi}\lambda\exp\Big(-\frac{\lambda}{2}(\zeta_0+\sqrt{\psi})\Big)\qquad\because\ \sigma^*_{jj}\le\lambda_{max}(\Sigma_0)\le\zeta_0\quad\text{due to energy boundedness}$$
$$= \Big\{\sqrt{\psi}\lambda\exp\Big(-\frac{\lambda}{2}(\zeta_0+\sqrt{\psi})\Big)\Big\}^p$$

Using a condition $\log p/\zeta^4\zeta_0^4\le n$, we have $\lambda\sqrt{\psi}\le\lambda\zeta_0$ since

$$\lambda\sqrt{\psi} = \lambda\sqrt{\frac{2}{3\zeta^4\zeta_0^2}\frac{\log p}{n}} = \lambda\zeta_0\sqrt{\frac{2}{3\zeta^4\zeta_0^4}\frac{\log p}{n}} \le \lambda\zeta_0\sqrt{\frac{2}{3}}\le\lambda\zeta_0$$

Hence, we can proceed the above inequality as the following :

$$\prod_{j=1}^{p}\pi^u\Big((\sigma_{jj}-\sigma^*_{jj})^2 \le \frac{2}{3\zeta^4\zeta_0^2}\frac{\log p}{n}\Big) \ge \Big\{\sqrt{\psi}\lambda\exp\Big(-\frac{\lambda}{2}(\zeta_0+\sqrt{\psi})\Big)\Big\}^p$$
$$= \exp\Big(p\log\lambda\sqrt{\psi}\Big)\exp\Big(-\frac{\lambda}{2}p\zeta_0-\frac{\lambda}{2}p\sqrt{\psi}\Big)$$
$$\ge \exp\Big(p\log\lambda\sqrt{\psi}\Big)\exp\Big(-\frac{\lambda}{2}p\zeta_0-\frac{\lambda}{2}p\zeta_0\Big)\qquad\because\ \lambda\sqrt{\psi}\le\lambda\zeta_0$$
$$= \exp\Big(-p\lambda\zeta_0-p\log\frac{1}{\lambda\sqrt{\psi}}\Big)$$
$$\ge \exp\Big(-p\log p-p\log\frac{1}{\lambda\sqrt{\psi}}\Big)\qquad\because\ \lambda<\log p/\zeta_0 \text{ by assumption}$$

Here, we shall claim that $1/\sqrt{\psi}\le\zeta^3p^{1/2\beta}$ for all sufficiently large n

$$
\begin{aligned}
1/\sqrt{\psi} &= \sqrt{\frac{3}{2}}\zeta_0\zeta^2\sqrt{\frac{n}{\log p}} \\
&< \sqrt{\frac{3}{2}}\zeta^3\sqrt{\frac{n}{\log p}} \quad \because \zeta_0 < \zeta \text{ by assumption} \\
&\le \sqrt{\frac{3}{2}}\zeta^3 C p^{1/2\beta}\frac{1}{\sqrt{\log p}} \quad \because\ p \asymp n^\beta\,,\, n^\beta \le Cp \text{ for some } C > 0 \text{ by assumption} \\
&\le \zeta^3 p^{1/2\beta} \quad \because\ p \text{ gets large enough to attain } \sqrt{3/2}C/\sqrt{\log p} < 1
\end{aligned}
$$

We will complete our process of finding lower bound of $\prod_{j=1}^{p}\pi^u\Big((\sigma_{jj}-\sigma_{jj}^*)^2 \le \frac{2}{3\zeta^4\zeta_0^2}\frac{\log p}{n}\Big)$ as the below.

$$
\begin{aligned}
&\prod_{j=1}^{p}\pi^u\Big((\sigma_{jj}-\sigma_{jj}^*)^2 \le \frac{2}{3\zeta^4\zeta_0^2}\frac{\log p}{n}\Big) \ge \exp\Big(-p\log p - p\log\frac{1}{\lambda\sqrt{\psi}}\Big) \\
&\ge \exp\Big(-p\log p - p\log\frac{\zeta^3 p^{1/2\beta}}{\lambda}\Big) \quad \because\ 1/\sqrt{\psi} \le \zeta^3 p^{1/2\beta} \text{ for all sufficiently large } n \\
&\ge \exp\Big(-p\log p - p(1+\frac{3}{4}+\frac{1}{2\beta})\log p\Big) \quad \because\ p^{-1} < \lambda \text{ and } \zeta^4 \le p \quad \text{by assumption} \\
&\ge \exp\Big(-(3+\frac{1}{2\beta})p\log p\Big)
\end{aligned}
$$

Hence we have

$$
\prod_{j=1}^{p}\pi^u\Big((\sigma_{jj}-\sigma_{jj}^*)^2 \le \frac{2}{3\zeta^4\zeta_0^2}\frac{\log p}{n}\Big) \ge \exp\Big(-(3+\frac{1}{2\beta})p\log p\Big) \tag{57}
$$

for sufficiently large n.

Next, we shall find a lower bound of $\prod_{i<j}\pi^u\Big((\sigma_{ij}-\sigma_{ij}^*)^2 \le \frac{2}{3\zeta^4\zeta_0^2}\frac{s_0\log p}{p(p-1)n}\Big)$. Note that it can be decomposed as the following.

$$
\begin{aligned}
&\prod_{i<j}\pi^u\Big((\sigma_{ij}-\sigma_{ij}^*)^2 \le \frac{2}{3\zeta^4\zeta_0^2}\frac{s_0\log p}{p(p-1)n}\Big) = \prod_{i<j}\pi^u(|\sigma_{ij}-\sigma_{ij}^*| \le \sqrt{\phi}) \quad \text{where } \phi = \frac{2}{3\zeta^4\zeta_0^2}\frac{s_0\log p}{p(p-1)n} \\
&= \prod_{(i,j)\in s(\Sigma_0)}\pi^u(|\sigma_{ij}-\sigma_{ij}^*| \le \sqrt{\phi}) \times \prod_{(i,j)\notin s(\Sigma_0)\,,\,i<j}\pi^u(|\sigma_{ij}| \le \sqrt{\phi})
\end{aligned}
$$

Before finding the lower bound of those two terms above, recall the tight bound of marginal prior density of off diagonal σ_{ij} of covariance matrix.

$$
\begin{aligned}
\pi^u(\sigma_{ij}) &< \frac{1}{\tau\sqrt{2\pi^3}}\log\Big(1+\frac{2\tau^2}{\sigma_{ij}^2}\Big) \\
\pi^u(\sigma_{ij}) &> \frac{1}{2\tau\sqrt{2\pi^3}}\log\Big(1+\frac{4\tau^2}{\sigma_{ij}^2}\Big)
\end{aligned}
$$

We will deal with $\prod_{(i,j)\notin s(\Sigma_0)\,,\,i<j}\pi^u(|\sigma_{ij}| \le \sqrt{\phi})$ first.

$$
\begin{aligned}
&\pi^u(|\sigma_{ij}| > \sqrt{\phi}) = 2\pi^u(\sigma_{ij} > \sqrt{\phi}) = \int_{\sqrt{\phi}}^{\infty} \pi^u(\sigma_{ij})\, d\sigma_{ij} \\
&\leq 2\int_{\sqrt{\phi}}^{\infty} \frac{1}{\tau\sqrt{2\pi^3}} \log\Big(1+\frac{2\tau^2}{\sigma_{ij}^2}\Big)\, d\sigma_{ij} \quad \because \text{ upper bound of marginal prior density of } \sigma_{ij} \\
&\leq 2\int_{\sqrt{\phi}}^{\infty} \frac{1}{\tau\sqrt{2\pi^3}} \frac{2\tau^2}{\sigma_{ij}^2}\, d\sigma_{ij} \quad \because \log(1+x) \leq x \quad \forall\, x > -1 \quad \text{by supporting line lemma} \\
&= 2\tau\sqrt{\frac{2}{\pi^3}} \int_{\sqrt{\phi}}^{\infty} \frac{1}{\sigma_{ij}^2}\, d\sigma_{ij} = 2\tau\sqrt{\frac{2}{\pi^3}}\frac{1}{\sqrt{\phi}}
\end{aligned}
$$

$$
\begin{aligned}
&\prod_{(i,j)\notin s(\Sigma_0)\,,\, i<j} \pi^u(|\sigma_{ij}| \leq \sqrt{\phi}) = \prod_{(i,j)\notin s(\Sigma_0)\,,\, i<j} \big(1 - \pi^u(|\sigma_{ij}| > \sqrt{\phi})\big) \\
&\geq \prod_{(i,j)\notin s(\Sigma_0)\,,\, i<j} \Big(1 - 2\tau\sqrt{\frac{2}{\pi^3}}\frac{1}{\sqrt{\phi}}\Big) \quad \because \pi^u(|\sigma_{ij}| > \sqrt{\phi}) \leq 2\tau\sqrt{\frac{2}{\pi^3}}\frac{1}{\sqrt{\phi}} \\
&\geq \Big(1 - 2\tau\sqrt{\frac{2}{\pi^3}}\frac{1}{\sqrt{\phi}}\Big)^{p^2} \\
&\geq \exp\Big(-4\tau\sqrt{\frac{2}{\pi^3}}\frac{1}{\sqrt{\phi}}\Big)^{p^2} \quad \because \log(1-x) \geq -2x \quad \text{when } 0 \leq x \leq 1/2
\end{aligned}
$$

Note that $\tau/\sqrt{\phi}$ is small enough when n is sufficiently large $\quad \because\ (p^2\sqrt{n})^{-1} \lesssim \tau \lesssim (p^2\sqrt{n})^{-1}\sqrt{s_0 \log p}$

$$
\tau/\sqrt{\phi} \leq C\frac{1}{p^2}\sqrt{\frac{s_0 \log p}{n}}\sqrt{\frac{p(p-1)n}{s_0 \log p}}\frac{2}{3\zeta_0^2\zeta^4} \leq \tilde{C}\frac{1}{p} \to 0
$$

We can proceed the inequality as the following.

$$
\begin{aligned}
&\prod_{(i,j)\notin s(\Sigma_0)\,,\, i<j} \pi^u(|\sigma_{ij}| \leq \sqrt{\phi}) \geq \exp\Big(-4\tau\sqrt{\frac{2}{\pi^3}}\frac{1}{\sqrt{\phi}}\Big)^{p^2} \\
&= \exp\Big(-4\tau p^2\sqrt{\frac{2}{\pi^3}}\frac{1}{\sqrt{\phi}}\Big) \\
&\geq \exp\Big(-4\sqrt{\frac{2}{\pi^3}}\tilde{C}p^2\frac{1}{p}\Big) \quad \because \tau/\sqrt{\phi} \leq \tilde{C}\frac{1}{p} \quad \text{by above} \\
&= \exp(-Cp) \quad \text{for some } C > 0
\end{aligned}
$$

Thus we have

$$
\prod_{(i,j)\notin s(\Sigma_0)\,,\, i<j} \pi^u(|\sigma_{ij}| \leq \sqrt{\phi}) \geq \exp\Big(-4\tau p^2\sqrt{\frac{2}{\pi^3}}\frac{1}{\sqrt{\phi}}\Big) \geq \exp(-Cp) \tag{58}
$$

for sufficiently large n.

Finally, we shall find the lower bound of $\prod_{(i,j)\in s(\Sigma_0)} \pi^u(\big|\sigma_{ij} - \sigma_{ij}^*\big| \leq \sqrt{\phi})$

Recall that marginal prior density of off diagonal σ_{ij} given as $\pi^u(\sigma_{ij})$ is decreasing function with respect to $|\sigma_{ij}|$ since

$$
\pi^u(\sigma_{ij}) = \int_0^{\infty} \frac{1}{\sqrt{2\pi^3\tau^2}} \exp\Big(-\frac{\sigma_{ij}^2}{2\tau^2}u\Big)\frac{1}{1+u}\, du
$$

Also, note that since $\phi = \frac{2}{3\zeta^4\zeta_0^2}\frac{s_0 \log p}{p(p-1)n} \to 0$ as n tends to sufficiently large, we can write

$$
(\big|\sigma_{ij} - \sigma_{ij}^*\big| \leq \sqrt{\phi}) = (\sigma_{ij}^* - \sqrt{\phi} \leq \sigma_{ij} \leq \sigma_{ij}^* + \sqrt{\phi}) \begin{cases} \subset (0,\infty) & \text{if} \quad \sigma_{ij}^* > 0 \\ \subset (-\infty, 0) & \text{if} \quad \sigma_{ij}^* < 0 \end{cases}
$$

for sufficiently large n, since $\sigma^*_{ij} \neq 0$ due to $(i,j) \in s(\Sigma_0)$

Therefore, we have the following inequality.

$$\pi^u(\left|\sigma_{ij} - \sigma^*_{ij}\right| \leq \sqrt{\phi}) \begin{cases} \geq 2\sqrt{\phi}\,\pi^u(\sigma^*_{ij} + \sqrt{\phi}) & \text{if} \quad \sigma^*_{ij} > 0 \\ \geq 2\sqrt{\phi}\,\pi^u(\sigma^*_{ij} - \sqrt{\phi}) & \text{if} \quad \sigma^*_{ij} < 0 \end{cases}$$

Combining three facts, we can yield $\left|\sigma^*_{ij}\right| \leq \zeta_0$ for all $i \neq j$. Those facts are given as the following.

1. The largest entry in magnitude of positive definite matrix lies on the diagonal (Source : Gockenbagh Linear Algebra Lemma 386)
2. Energy boundedness : $\lambda_n \|x\|_2^2 \leq x^T A x \leq \lambda_1 \|x\|_2^2$ if symmetric A has spec$(A) = \{\lambda_1 \geq \cdots \geq \lambda_n\}$
3. $\lambda_{max}(\Sigma_0) \leq \zeta_0$ by $\Sigma_0 \in \mathcal{U}(s_0, \zeta_0)$ assumption

Hence we have

$$\left|\sigma^*_{ij}\right| \leq \max_k |\sigma^*_{kk}| = \max_k \sigma^*_{kk} \leq \lambda_{max}(\Sigma_0) \leq \zeta_0$$

and what follow this are

$$\sigma^*_{ij} + \sqrt{\phi} \leq 2\zeta_0 \quad \text{if} \quad \sigma^*_{ij} > 0$$
$$\left|\sigma^*_{ij} - \sqrt{\phi}\right| \leq 2\zeta_0 \quad \text{if} \quad \sigma^*_{ij} < 0$$

Using this, we get

$$\pi^u(\left|\sigma_{ij} - \sigma^*_{ij}\right| \leq \sqrt{\phi}) \geq 2\sqrt{\phi}\,\pi^u(2\zeta_0)$$

$$\prod_{(i,j)\in s(\Sigma_0)} \pi^u(\left|\sigma_{ij} - \sigma^*_{ij}\right| \leq \sqrt{\phi}) \geq \left(2\sqrt{\phi}\,\pi^u(2\zeta_0)\right)^{s_0} \geq \left(\sqrt{\phi}\,\pi^u(2\zeta_0)\right)^{s_0} \geq \left(\pi^u(2\zeta_0)\sqrt{\frac{2s_0 \log p}{3\zeta^4 \zeta_0^2 p^2 n}}\right)^{s_0}$$
$$= \exp\left(s_0 \log \pi^u(2\zeta_0) + \frac{1}{2} s_0 \log \frac{2s_0 \log p}{3\zeta^4 \zeta_0^2 p^2 n}\right)$$
$$\geq \exp\left(s_0 \log \pi^u(2\zeta_0) + \frac{1}{2} s_0 \log \frac{2}{3}\frac{1}{\zeta^2 p^2 n}\right) \quad \because \ \zeta^2 \zeta_0^2 \leq s_0 \log p \quad \text{by assumption}$$

Note that by taking advantage of Lemma 4, we can write

$$\pi^u(2\zeta_0) \geq \frac{1}{\sqrt{2\pi^3}} \frac{\tau}{4\zeta_0^2}$$

Of course, we should show that $\tau/2\zeta_0$ is sufficiently small to justify the use of Lemma 4. Since ζ_0 is fixed, we shall show that $\tau \to 0$ as $n \to \infty$

$$\begin{aligned} \tau &\lesssim \frac{\sqrt{s_0 \log p}}{p^2 \sqrt{n}} \quad \text{by assumption} \\ &\lesssim \frac{\sqrt{s_0 \log s_0}}{p^2 \sqrt{n}} \quad \because p \lesssim s_0 \\ &\leq \frac{s_0}{p^2 \sqrt{n}} \leq \frac{1}{\sqrt{n}} \quad \because \ \log s_0 \leq s_0 \leq p^2 \end{aligned}$$

Thus $\tau \lesssim \frac{1}{\sqrt{n}}$ so that τ/ζ_0 is sufficiently small as n gets sufficiently large.

Combining with the condition $\tau \gtrsim 1/\sqrt{n}p^2$, we get

$$\pi^u(2\zeta_0) \geq \frac{1}{\sqrt{2\pi^3}} \frac{\tau}{4\zeta_0^2} \gtrsim \frac{1}{\sqrt{n}p^2}$$

Now we shall proceed our target inequality.

$$\begin{aligned}
&\prod_{(i,j)\in s(\Sigma_0)} \pi^u(\left|\sigma_{ij} - \sigma_{ij}^*\right| \leq \sqrt{\phi}) \\
&\geq \exp\Big(s_0 \log \pi^u(2\zeta_0) + \frac{1}{2} s_0 \log \frac{2}{3} \frac{1}{\zeta^2 p^2 n}\Big) \\
&\geq \exp\Big(s_0 \log\Big(\frac{\tilde{C}}{\sqrt{n}p^2}\Big) + \frac{1}{2} s_0 \log \frac{2}{3} \frac{1}{\zeta^2 p^2 n}\Big) \\
&= \exp\Big(\frac{1}{2} s_0 \log\Big(\frac{\tilde{C}^2}{np^4}\Big) + \frac{1}{2} s_0 \log \frac{2}{3} \frac{1}{\zeta^2 p^2 n}\Big) \\
&= \exp\Big(\frac{1}{2} s_0 \log\big(\frac{2\tilde{C}^2}{3\zeta^2 p^6 n^2}\big)\Big)
\end{aligned}$$

Here, we gonna use two inequalities

1. $C^* p^{-2/\beta} \leq n^{-2}$ for some $C^* > 0$
2. $1/\zeta^2 \geq 1/p$

The first one comes from $p \asymp n^\beta$ so that $n^\beta \leq \tilde{C}^* p$ and the second one comes from $\zeta^4 \leq p$ and $1 < \zeta_0 < \zeta$. Using those inequalities, we get

$$\begin{aligned}
&\prod_{(i,j)\in s(\Sigma_0)} \pi^u(\left|\sigma_{ij} - \sigma_{ij}^*\right| \leq \sqrt{\phi}) \\
&\geq \exp\Big(\frac{1}{2} s_0 \log\big(\frac{2\tilde{C}^2}{3\zeta^2 p^6 n^2}\big)\Big) \\
&\geq \exp\Big(\frac{1}{2} s_0 \log\big(\frac{2}{3}\tilde{C}^2 p^{-1} p^{-6} C^* p^{-2/\beta}\big)\Big) \\
&= \exp\Big(\frac{1}{2} s_0 \log\Big(C p^{-(7+\frac{2}{\beta})}\Big)\Big) \\
&= \exp\Big(-\frac{1}{2}(7+\frac{2}{\beta}) s_0 \log p + \frac{1}{2} s_0 \log C\Big) \\
&\geq \exp\Big(-\frac{1}{2}(7+\frac{2}{\beta}) s_0 \log p - \frac{1}{2} s_0 \log p\Big) \quad \text{for suff. large } n \\
&= \exp\Big(-(4+\frac{1}{\beta}) s_0 \log p\Big)
\end{aligned}$$

Hence we have

$$\prod_{(i,j)\in s(\Sigma_0)} \pi^u(\left|\sigma_{ij} - \sigma_{ij}^*\right| \leq \sqrt{\phi}) \geq \exp\Big(-(4+\frac{1}{\beta}) s_0 \log p\Big) \tag{59}$$

At last, combining (57), (58), and (59), we have

$$\begin{aligned}
&\pi^u(A_{n,\Sigma_0}) \geq \exp\Big(-(3+\frac{1}{2\beta}) p \log p\Big) \times \exp(-Cp) \times \exp\Big(-(4+\frac{1}{\beta}) s_0 \log p\Big) \\
&\geq \exp\Big(-(3+\frac{1}{2\beta}) p \log p\Big) \times \exp(-Cp) \times \exp\Big(-(4+\frac{1}{\beta}) s_0 \log p\Big) \times \exp(Cp) \times \exp(-p \log p) \\
&= \exp\Big(-(4+\frac{1}{2\beta}) p \log p\Big) \times \exp\Big(-(4+\frac{1}{\beta}) s_0 \log p\Big) \geq \exp\Big(-(4+\frac{1}{\beta})(p+s_0) \log p\Big) \\
&= \exp\Big(-(4+\frac{1}{\beta}) n\varepsilon_n^2\Big)
\end{aligned}$$

Since we have already shown that

$$\pi(B_{\varepsilon_n}) \geq \pi\Big(\|\Sigma - \Sigma_0\|_F^2 \leq \frac{2}{3\zeta^4\zeta_0^2}\varepsilon_n^2\Big) \geq \pi(A_{n,\Sigma_0}) \geq \pi^u(A_{n,\Sigma_0})$$

we can conclude that

$$\pi(B_{\varepsilon_n}) \geq \exp\Big(-(4+\frac{1}{\beta})n\varepsilon_n^2\Big)$$

□

4 이경원, 정진욱, 김성민 - Theorem 2

4.1 모형

다음의 모형을 생각하자.

$$X_1, \cdots, X_n|\Sigma \sim N(0, \Sigma) \tag{60}$$

양의 정수 s_0와 실수 $\zeta_0 > 1$에 대해 다음과 같은 모수공간을 생각한다.

$$U(s_0, \zeta_0) = \{\Sigma \in C_p : s(\Sigma) \leq s_0,\ \zeta_0^{-1} \leq \lambda_{\min}(\Sigma) \leq \lambda_{\max}(\Sigma) \leq \zeta_0\}$$

여기서 $s(\Sigma)$는 행렬 Σ의 0이 아닌 비대각성분의 개수를 의미한다.

4.2 정리

Theorem 4.1 (Theorem 2 in Lee et al. (2022))**.** 모형 *60*과 양의 정수 s_0 와 실수 $\zeta_0 > 1$에 대해 $\Sigma_0 \in \mathcal{U}(s_0, \zeta_0)$이라 하자. $s_0^2(\log p)^3 = O(p^2 n)$이면 작은 상수 $\epsilon > 0$에 대해 다음이 성립한다.

$$\inf_{\hat{\Sigma}} \sup_{\Sigma_0 \in \mathcal{U}(s_0,\zeta_0)} \mathbb{E}_0\|\hat{\Sigma} - \Sigma_0\|_F^2 \gtrsim \frac{(p+s_0)\log p}{n} I\left(3p < s_0 < p^{3/2-\epsilon/2}\right) + \frac{p+s_0}{n}$$

Remark 4.1. 이 정리는 공분산 추정의 *minimax lower bound*를 알려준다.

논문의 *Theorem 1*에서는 사후수렴속도가 $\frac{(p+s_0)\log p}{n}$ 임을 보였는데, 이는 $3p < s_0 < p^{3/2-\epsilon/2}$ 일 때 베이즈 추론이 *minimax* 이고, 그렇지 않은 경우에도 *nearly minimax* $(\log p)$ 임을 의미한다.

Remark 4.2. 이와 관련된 연구로 *Cai and Zhou* (*2012*)가 있는데, 이 논문에서는 빈도론 관점에서 성긴 공분산 행렬을 추론하는 문제를 다루었다. 다만, 논문에서는 공분산 행렬의 각 **열의** *0*이 아닌 성분에 대한 제약조건을 다루었으나, 본 논문에서는 전체 행렬에서 *0*이 아닌 성분에 대한 제약조건에 대해 다룬다.

4.3 증명

Proof. 다음의 두 항목을 증명하면 된다.

- $3p < s_0 < p^{3/2-\epsilon/2}$ 인 경우,

$$\inf_{\hat{\Sigma}} \sup_{\Sigma_0 \in B_1} \mathbb{E}_0\|\hat{\Sigma} - \Sigma_0\|_F^2 \gtrsim \frac{s_0 \log p}{n} \tag{61}$$

  이 성립하는 $B_1 \subset \mathcal{U}(s_0, \zeta_0)$가 존재함을 보인다. (이경원, 정진욱)

- 나머지 경우,

$$\inf_{\hat{\Sigma}} \sup_{\Sigma_0 \in B_2} \mathbb{E}_0 \|\hat{\Sigma} - \Sigma_0\|_F^2 \gtrsim \frac{s_0 + p}{n} \tag{62}$$

이 성립하는 $B_2 \subset \mathcal{U}(s_0, \zeta_0)$가 존재함을 보인다. (김성민)

먼저 첫 항목을 보이자. $\nu = \sqrt{\epsilon/4}$에 대해 $r = \lfloor p/2 \rfloor$, $\epsilon_{np} = \nu\sqrt{\log p/n}$이라 하자. $A_m(u)$를 m번째 행과 열이 u의 값을 갖고, 나머지에서 모두 0인 대칭행렬이라 하자. 즉,

$$(A_m(u))_{ij} = \begin{cases} u & i = m \text{ or} j = m \\ 0 & \text{otherwise} \end{cases}$$

이라 하자. 이제, 다음과 같이 B_1을 정의한다.

$$B_1 := \left\{ \Sigma(\theta) : \Sigma(\theta) = I_p + \epsilon_{np} \sum_{m=1}^{r} \gamma_m A_m(\lambda_m),\ \theta = (\gamma, \lambda) \in \Theta \right\}$$

여기서 $\gamma = (\gamma_1, \cdots, \gamma_r) \in \Gamma = \{0,1\}^r$, $\lambda = (\lambda_1, \cdots, \lambda_r)^T \in \Lambda \subset \mathbb{R}^{r \times p}$,

$$\Lambda = \Bigg\{ \lambda = (\lambda_{ij}) : \lambda_{mi} \in \{0,1\},\ \|\lambda_m\|_0 = k,\ \sum_{i=1}^{p-r} \lambda_{mi} = 0,$$
$$m \in \{1, \cdots, r\},\ \ \text{satisfying } \max_{1 \le i \le p} \sum_{m=1}^{r} \lambda_{mi} \le 2k \Bigg\},$$

$k = \lceil c_{np}/2 \rceil - 1$, $c_{np} = \lceil s_0/p \rceil$, $\Theta = \Gamma \times \Lambda$이다.

이제 $B_1 \subset \mathcal{U}(s_0, \zeta_0)$과 식 (61)이 성립함을 보이면 된다.

먼저, 임의의 $\zeta_0 > 1$과 충분히 큰 n에 대해 $\Sigma(\theta) \in B_1$의 가장 큰 고유치가 ζ_0보다 작다는 것은 다음과 같이 보일 수 있다.

$$\lambda_{\max}(\Sigma(\theta)) \le \|\Sigma(\theta)\|_1 \le 1 + 2k\epsilon_{np} \le 1 + c_{np}\nu\sqrt{\log p/n} \le \zeta_0$$

두 번째 부등호는 모수공간 Λ의 마지막 조건으로부터, 마지막 부등호는 가정 $s_0^2(\log p)^3 = O(p^2 n)$에 의해 성립한다.

마찬가지의 이유로 임의의 $\zeta_0 > 1$과 충분히 큰 n에 대해

$$2k\epsilon_{np} \le c_{np}\nu\sqrt{\log p/n} \le \left(1 + \frac{s_0}{p}\right)\nu\sqrt{\log p/n} \le 1 - \zeta_0^{-1}$$

가 성립하므로 $\Sigma(\theta) - \zeta_0^{-1} I_p$는 대각지배(diagonally dominant) 행렬이고, 대칭이며 모든 성분이 0보다 크거나 같아 양의 준정부호 행렬이다[1]. 따라서, $\Sigma(\theta)$의 가장 작은 고유치는 ζ_0^{-1}보다 크다.

마지막으로 $\Sigma(\theta)$의 비대각성분은 모두 A_m들에 의해서만 나타나므로

$$s(\Sigma(\theta)) \le 2kp \le s_0$$

에서 $B_1 \subset \mathcal{U}(s_0, \zeta_0)$를 얻는다.

이제 식 (61)이 성립함을 보이자. 이를 위해, 다음의 보조정리를 소개한다. 이 보조정리는 모수공간 Θ에서 거리 d를 갖는 거리공간으로의 변환 $\psi(\theta)$의 최대 위험의 하한을 알려준다.

Lemma 4.2 (Lemma 3 of Cai and Zhou (2012)). *For any $s > 0$ and any estimator T of $\psi(\theta)$ based on an observation from the experiment $\{P_\theta,\ \theta \in \Theta\}$,*

$$\max_{\theta \in \Theta} 2^s \mathbb{E}_\theta d^s(T, \psi(\theta)) \ge \alpha \frac{r}{2} \min_{1 \le i \le r} \|\overline{\mathbb{P}}_{i,0} \wedge \overline{\mathbb{P}}_{i,1}\|,$$

[1] A Hermitian diagonally dominant matrix A with real non-negative diagonal entries is positive semidefinite. From `https://en.wikipedia.org/wiki/Diagonally_dominant_matrix#Applications_and_properties` 혹은, 더 간단하게 Gershgorin circle theorem에 symmetric matrix가 real eigenvalue를 가진다는 사실로도 보일 수 있다.

where $\overline{\mathbb{P}}_{i,a}$ is the mixture distribution over all P_θ with $\gamma_i(\theta)$ fixed to be a while all other components of θ vary over all possible values, i.e.,

$$\overline{\mathbb{P}}_{i,a} = \frac{1}{2^{r-1}|\Lambda|} \sum_{\theta \in \Theta_{i,a}} P_\theta,$$

for $\Theta_{i,a} = \{\theta \in \Theta : \gamma_i(\theta) = a\}$,

$$\|\mathbb{P} \wedge \mathbb{Q}\| = \int (p \wedge q) d\mu,$$

for probability measures $\mathbb{P}$ and $\mathbb{Q}$ which have densities p and q respectively, $\mathbb{E}_\theta$ is expectation with respect to $[X_1, \cdots, X_n|\theta]$, $H(x,y)$ is the Hamming distance defined as

$$H(x,y) = \sum_{j=1}^{r} |x_i - y_i|, \quad x, y \in \{0,1\}^r.$$

$$\alpha = \min_{(\theta,\theta'): H(\gamma(\theta),\gamma(\theta')) \geq 1} d^s(\psi(\theta), \psi(\theta'))/H(\gamma(\theta), \gamma(\theta'))$$

Proof. 최댓값은 평균보다 크거나 같으므로

$$\begin{aligned}
\max_{\theta \in \Theta} 2^s \mathbb{E}_\theta d^s(T, \psi(\theta)) &\geq \frac{1}{2^r|\Lambda|} \sum_{\theta \in \Theta} 2^s \mathbb{E}_\theta d^s(T, \psi(\theta)) \\
&= \frac{1}{2^r|\Lambda|} \sum_{\theta \in \Theta} \mathbb{E}_\theta (2d(T, \psi(\theta)))^s
\end{aligned}$$

$\hat{\theta} := \arg\min d^s(T, \psi(\theta))$ 라 하면 (유일하지 않다면 적당히 하나를 잡으면 된다.)

$$\begin{aligned}
\mathbb{E}_\theta (2d(T, \psi(\theta)))^s &\geq \mathbb{E}_\theta (d(T, \psi(\theta)) + d(T, \psi(\hat{\theta})))^s \\
&\geq \mathbb{E}_\theta (d(\psi(\hat{\theta}), \psi(\theta)))^s
\end{aligned}$$

를 얻는다. 마지막 부등식에서 삼각부등식을 사용하였다.
정리하면 다음을 얻는다.

$$\begin{aligned}
\max_{\theta \in \Theta} 2^s \mathbb{E}_\theta d^s(T, \psi(\theta)) &\geq \frac{1}{2^r|\Lambda|} \sum_{\theta \in \Theta} \mathbb{E}_\theta (d(\psi(\hat{\theta}), \psi(\theta)))^s \\
&\geq \frac{1}{2^r|\Lambda|} \sum_{\theta \in \Theta} \mathbb{E}_\theta \left[\frac{(d(\psi(\hat{\theta}), \psi(\theta)))^s}{H(\gamma(\theta), \gamma(\theta')) \vee 1} H(\gamma(\theta), \gamma(\theta')) \right] \\
&\geq \alpha \frac{1}{2^r|\Lambda|} \sum_{\theta \in \Theta} \mathbb{E}_\theta \left[H(\gamma(\theta), \gamma(\theta')) \right]
\end{aligned}$$

이제

$$\frac{1}{2^r|\Lambda|} \sum_{\theta \in \Theta} \mathbb{E}_\theta \left[H(\gamma(\theta), \gamma(\theta')) \right] \geq \frac{r}{2} \min_{1 \leq i \leq r} \|\overline{\mathbb{P}}_{i,0} \wedge \overline{\mathbb{P}}_{i,1}\|$$

을 보이면 원하는 결과를 얻는다.

$$\frac{1}{2^r|\Lambda|}\sum_{\theta\in\Theta}\mathbb{E}_\theta\left[H(\gamma(\theta),\gamma(\theta'))\right]$$
$$=\frac{1}{2^r|\Lambda|}\sum_{\theta\in\Theta}\sum_{i=1}^{r}\mathbb{E}_\theta\left[|\gamma_i(\theta)-\gamma_i(\theta'))|\right]$$
$$=\sum_{i=1}^{r}\frac{1}{2^r|\Lambda|}\sum_{\rho\in\Gamma}\sum_{\theta:\gamma(\theta)=\rho}\mathbb{E}_\theta\left[|\gamma_i(\theta)-\gamma_i(\theta'))|\right]$$
$$=\frac{1}{2}\sum_{i=1}^{r}\left\{\frac{1}{2^{r-1}|\Lambda|}\sum_{\rho_i=0}\sum_{\theta:\gamma(\theta)=\rho}\mathbb{E}_\theta\left[|\gamma_i(\theta)-\gamma_i(\theta'))|\right]+\frac{1}{2^{r-1}|\Lambda|}\sum_{\rho_i=1}\sum_{\theta:\gamma(\theta)=\rho}\mathbb{E}_\theta\left[|\gamma_i(\theta)-\gamma_i(\theta'))|\right]\right\}$$
$$=\frac{1}{2}\sum_{i=1}^{r}\left\{\frac{1}{2^{r-1}|\Lambda|}\sum_{\rho_i=0}\sum_{\theta:\gamma(\theta)=\rho}\mathbb{E}_\theta\left[|\gamma_i(\theta'))|\right]+\frac{1}{2^{r-1}|\Lambda|}\sum_{\rho_i=1}\sum_{\theta:\gamma(\theta)=\rho}\mathbb{E}_\theta\left[|1-\gamma_i(\theta'))|\right]\right\}$$
$$=\frac{1}{2}\sum_{i=1}^{r}\left\{\frac{1}{2^{r-1}|\Lambda|}\sum_{\rho_i=0}\sum_{\theta:\gamma(\theta)=\rho}\int_{\theta'}\gamma_i(\theta'))d\mathbb{P}_{\theta'}+\frac{1}{2^{r-1}|\Lambda|}\sum_{\rho_i=1}\sum_{\theta:\gamma(\theta)=\rho}\int_{\theta'}1-\gamma_i(\theta'))d\mathbb{P}_{\theta'}\right\}$$
$$=\frac{1}{2}\sum_{i=1}^{r}\left\{\int_{\theta'}\gamma_i(\theta'))\frac{1}{2^{r-1}|\Lambda|}\sum_{\rho_i=0}\sum_{\theta:\gamma(\theta)=\rho}d\mathbb{P}_{\theta'}+\int_{\theta'}\left(1-\gamma_i(\theta'))\right)\frac{1}{2^{r-1}|\Lambda|}\sum_{\rho_i=1}\sum_{\theta:\gamma(\theta)=\rho}d\mathbb{P}_{\theta'}\right\}$$
$$=\frac{1}{2}\sum_{i=1}^{r}\left\{\int_{\theta'}\gamma_i(\theta'))d\overline{\mathbb{P}}_{i,0}+\int_{\theta'}\left(1-\gamma_i(\theta'))\right)d\overline{\mathbb{P}}_{i,0}\right\}$$
$$\geq\frac{1}{2}\sum_{i=1}^{r}\int d\left[\overline{\mathbb{P}}_{i,0}\wedge\overline{\mathbb{P}}_{i,1}\right]$$
$$\geq\frac{r}{2}\min_{1\leq i\leq r}\|\overline{\mathbb{P}}_{i,0}\wedge\overline{\mathbb{P}}_{i,1}\|$$

증명이 끝났다. □

위의 보조정리에 $s=2$를 대입하면 다음의 부등식을 얻는다.

$$\inf_{\hat{\Sigma}}\max_{\theta\in\Theta}2^2\mathbb{E}_\theta\|\hat{\Sigma}-\Sigma(\theta)\|_F^2\geq\alpha\frac{r}{2}\min_{1\leq i\leq r}\|\overline{\mathbb{P}}_{i,0}\wedge\overline{\mathbb{P}}_{i,1}\|$$

여기서

$$\alpha=\min_{(\theta,\theta'):H(\gamma(\theta),\gamma(\theta'))\geq1}\|\Sigma(\theta)-\Sigma(\theta')\|_F^2/H(\gamma(\theta),\gamma(\theta'))$$

이다.

이때 임의의 $\theta,\ \theta'\in\Theta$에 대해

$$\begin{aligned}\|\Sigma(\theta)-\Sigma(\theta')\|_F^2&=\epsilon_{np}^2\left\|\sum_{m=1}^{r}\gamma_m(\theta)A_m(\lambda_m(\theta))-\sum_{m=1}^{r}\gamma_m(\theta')A_m(\lambda_m(\theta'))\right\|_F^2\\&\geq2k\epsilon_{np}^2H(\gamma(\theta),\gamma(\theta'))\end{aligned}$$

이므로 k와 r의 정의($r=\lfloor p/2\rfloor,\ k=\lceil c_{np}/2\rceil-1,\ c_{np}=\lceil s_0/p\rceil$)로부터 다음을 얻는다.

$$\alpha r\geq2k\epsilon_{np}^2r\geq\nu^2\left(\frac{1}{2}-\frac{p}{s_0}\right)\frac{s_0\log p}{n}\asymp\frac{s_0\log p}{n}$$

이제, 다음을 만족하는 적당한 상수 $c_1>0$이 존재함을 보이면 증명이 끝난다.

$$\min_{1\leq i\leq r}\|\overline{\mathbb{P}}_{i,0}\wedge\overline{\mathbb{P}}_{i,1}\|\geq c_1$$

□

이제 어떤 상수 $c_1 > 0$ 가 존재하여 모든 $i = 1, \dots, r$ 에 대해

$$\|\overline{\mathbb{P}}_{i,0} \wedge \overline{\mathbb{P}}_{i,1}\| \geq c_1$$

임을 보이면 되는데 여기서는 $\|\overline{\mathbb{P}}_{1,0} \wedge \overline{\mathbb{P}}_{1,1}\| \geq c_1$, 즉 $i = 1$ 일 때 만을 보일 것이다. i 가 바뀌어도 동일한 상수 c_1 을 얻을 수 있기에 이것만으로 충분하다.

먼저 다음과 같은 변수 공간들을 고려하자.

$$\Lambda_1 := \{\lambda_1(\theta) \in \mathbb{R}^p \,:\, \theta \in \Theta\}, \quad \Lambda_{-1} := \{\lambda_{-1}(\theta) \equiv (\lambda_2(\theta), \dots, \lambda_r(\theta))^T \in \mathbb{R}^{(r-1)\times p} \,:\, \theta \in \Theta\}.$$

각 $a \in \{0,1\}, b \in \{0,1\}^{r-1}, c \in \Lambda_{-1}$ 마다 다음과 같은 확률분포를 정의한다:

$$\overline{\mathbb{P}}_{(1,a,b,c)} := \frac{1}{|\Theta_{(1,a,b,c)}|} \sum_{\theta \in \Theta_{(1,a,b,c)}} \mathbb{P}_\theta, \quad \Theta_{(1,a,b,c)} := \{\theta \in \Theta \,:\, \gamma_1(\theta) = a, \quad \gamma_{-1}(\theta) = b, \quad \lambda_{-1}(\theta) = c\}.$$

여기서 $\gamma_{-1}(\theta) = (\gamma_2(\theta), \dots, \gamma_r(\theta))^T$ 이다. 이제 함수 $f = f(\gamma_{-1}, \lambda_{-1})$ 를 $\Theta_{-1} := \{0,1\}^{r-1} \times \Lambda_{-1}$ 에서 평균을 취한 값을 $\mathbb{E}_{(\gamma_{-1},\lambda_{-1})} f(\gamma_{-1}, \lambda_{-1})$ 라고 쓴다. 즉

$$\mathbb{E}_{(\gamma_{-1},\lambda_{-1})} f(\gamma_{-1}, \lambda_{-1}) := \frac{1}{2^{r-1}|\Lambda|} \sum_{(b,c) \in \Theta_{-1}} |\Theta_{(1,a,b,c)}| f(b, c)$$

이고 a 는 0과 1 중 무엇을 선택해도 무관하다.

이제 어떤 상수 $c_2 \in (0,1)$ 가 존재하여

$$\mathbb{E}_{(\gamma_{-1},\lambda_{-1})} \left\{ \int \left(\frac{d\overline{\mathbb{P}}_{(1,1,\gamma_{-1},\lambda_{-1})}}{d\overline{\mathbb{P}}_{(1,0,\gamma_{-1},\lambda_{-1})}} \right)^2 d\overline{\mathbb{P}}_{(1,0,\gamma_{-1},\lambda_{-1})} - 1 \right\} \leq c_2^2 \tag{16}$$

임을 보이기만 하면 충분하다. 왜냐하면 (Cai and Zhou, 2012, Lemma 8, (ii))의 결과를 이용하면 위 식은

$$\|\overline{\mathbb{P}}_{1,0} \wedge \overline{\mathbb{P}}_{1,1}\| \geq 1 - c_2 > 0$$

을 의미하기 때문이다.

여기서 잠시 이에 대한 증명을 짚고 넘어가자면, 우선 공통의 dominating measure μ 에 대해 두 개의 density q_0 와 q_1 가 있다고 할 때 옌센 부등식을 이용하면

$$\left[\int |q_0 - q_1| d\mu \right]^2 = \left(\int \left| \frac{q_0 - q_1}{q_1} \right| q_1 \right)^2 \leq \int \frac{(q_0 - q_1)^2}{q_1} d\mu = \int \left(\frac{q_0^2}{q_1} - 1 \right) d\mu$$

이다. 위 식과 (16)을 이용하면 알 수 있는 사실은

$$\mathbb{E}_{(\gamma_{-1},\lambda_{-1})} \left\{ \int \left| d\overline{\mathbb{P}}_{(1,1,\gamma_{-1},\lambda_{-1})} - d\overline{\mathbb{P}}_{(1,0,\gamma_{-1},\lambda_{-1})} \right|^2 \right\} \leq c_2^2$$

이고 코시-슈바르츠 부등식을 이용하면

$$\mathbb{E}_{(\gamma_{-1},\lambda_{-1})} \left\{ \int \left| d\overline{\mathbb{P}}_{(1,1,\gamma_{-1},\lambda_{-1})} - d\overline{\mathbb{P}}_{(1,0,\gamma_{-1},\lambda_{-1})} \right| \right\} \leq c_2$$

또한 성립함을 알 수 있다. 여기서 total variation affinity

$$\|\mathbb{P} \wedge \mathbb{Q}\| = 1 - TV(\mathbb{P}, \mathbb{Q})$$

를 이용하면 결국

$$\mathbb{E}_{(\gamma_{-1},\lambda_{-1})}\left\{\left\|\overline{\mathbb{P}}_{(1,1,\gamma_{-1},\lambda_{-1})} \wedge \overline{\mathbb{P}}_{(1,0,\gamma_{-1},\lambda_{-1})}\right\|\right\} \geq 1 - c_2$$

를 얻게 된다. 마지막으로 $\overline{\mathbb{P}}_{1,0}$ 와 $\overline{\mathbb{P}}_{1,1}$ 가 각각 $\overline{\mathbb{P}}_{(1,0,\gamma_{-1},\lambda_{-1})}$, $\overline{\mathbb{P}}_{(1,1,\gamma_{-1},\lambda_{-1})}$ 와 같은 형태의 확률측도들에 대한 가중평균이라는 점을 기억한다면, (Cai and Zhou, 2012, Lemma 4)를 이용하여

$$\|\overline{\mathbb{P}}_{1,0} \wedge \overline{\mathbb{P}}_{1,1}\| \geq \mathbb{E}_{(\gamma_{-1},\lambda_{-1})}\left\{\left\|\overline{\mathbb{P}}_{(1,1,\gamma_{-1},\lambda_{-1})} \wedge \overline{\mathbb{P}}_{(1,0,\gamma_{-1},\lambda_{-1})}\right\|\right\}$$

를 얻게 되므로, 증명이 마무리된다.

이제 (16)을 얻기 위해 $(\gamma_{-1}, \lambda_{-1})$ 가 고정되었을 때 (16)에서 등장하는 각각의 측도 $\overline{\mathbb{P}}_{(1,0,\gamma_{-1},\lambda_{-1})}$ 와 $\overline{\mathbb{P}}_{(1,1,\gamma_{-1},\lambda_{-1})}$ 이 어떠한 형태인지 알 필요가 있다. 먼저 $\overline{\mathbb{P}}_{(1,0,\gamma_{-1},\lambda_{-1})}$ 의 경우 $\gamma_1 = 0$ 으로 설정되어 있는데, 정의를 다시 떠올려본다면

$$\overline{\mathbb{P}}_{(1,0,\gamma_{-1},\lambda_{-1})} = \frac{1}{|\Theta_{(1,0,\gamma_{-1},\lambda_{-1})}|} \sum_{\theta \in \Theta_{(1,0,\gamma_{-1},\lambda_{-1})}} \mathbb{P}_\theta$$

이고 $\mathbb{P}_\theta$ 는 p차원 정규분포 $N_p(0, \Sigma(\theta))$ 를 따르는 n 개의 표본 $X_1, \ldots, X_n$ 에 대한 결합분포이다. 이때 $(\gamma_{-1}, \lambda_{-1})$ 는 고정되어 있으니, $\theta \in \Theta_{(1,0,\gamma_{-1},\lambda_{-1})}$ 에 대응하는 모든 $\Sigma(\theta)$ 들은

$$\Sigma(\theta) = I_p + \epsilon_{np} \sum_{m=2}^{r} \gamma_m A_m(\lambda_m),$$

즉 θ 에 대응하는 $\lambda_1(\theta)$ 가 무슨 값을 취하더라도 $\Sigma(\theta)$ 는 동일한 형태라는 것이다. 따라서 $\overline{\mathbb{P}}_{(1,0,\gamma_{-1},\lambda_{-1})}$ 는 p 차원 정규분포 $N_p(0, \Sigma_0)$ 를 따르는 n 개의 표본에 대한 결합분포임을 알 수 있으며 Σ_0 는 다음과 같이 주어진다:

$$\Sigma_0 := \left(\begin{array}{c|c} 1 & 0_{1\times(p-1)} \\ \hline 0_{(p-1)\times 1} & S_{(p-1)\times(p-1)} \end{array}\right).$$

여기서 $S_{(p-1)\times(p-1)} = \{s_{ij}\}$ 은 $(p-1)\times(p-1)$ 대칭행렬로

$$s_{ij} = \begin{cases} 1 & i = j \\ \epsilon_{np} & \gamma_{i+1} = \lambda_{i+1,j+1} = 1 \\ 0 & \text{otherwise} \end{cases}$$

와 같이 주어진다. 여기서 $i = 1, \ldots, p - r = p - \lfloor p/2 \rfloor$ 에 대해 $\lambda_{mi} = 0$ 이므로 $A_m(\lambda_m)$ 은 대각성분을 가지지 않기에, ϵ_{np} 는 $S_{(p-1)\times(p-1)}$ 의 대각 성분에 나타나지 않음을 유념하자.

이제 $\overline{\mathbb{P}}_{(1,1,\gamma_{-1},\lambda_{-1})}$ 의 경우를 살펴보도록 하자. 우선 임의의 $c \in \Lambda_{-1}$ 에 대해

$$\Lambda_1(c) := \{a \in \mathbb{R}^p \ : \ \lambda_1(\theta) = a, \quad \lambda_{-1}(\theta) = c \ \text{for some} \ \ \theta \in \Theta\}$$

를 정의하자. 또한 $\lambda_{-1} = (\lambda_2(\theta), \ldots, \lambda_r(\theta))^T = \begin{pmatrix} \lambda_2(\theta)^T \\ \hline \vdots \\ \hline \lambda_r(\theta)^T \end{pmatrix} \in \mathbb{R}^{(r-1)\times p}$ 가 하나 주어졌을 때, $n_{\lambda_{-1}}$ 를 λ_{-1} 의 각 열의 성분을 모두 더했을 때 $2k$ 가 되는 열의 갯수라고 하자. 즉

$$n_{\lambda_{-1}} := \left|\left\{i \ : \ \sum_{m=2}^{r} \lambda_{mi} = 2k\right\}\right|$$

이고 $p_{\lambda_{-1}} = r - n_{\lambda_{-1}}$ 이라 두자. 이와 같이 정의하면 $p_{\lambda_{-1}}$ 은 λ_1 의 성분 λ_{1i} 중 0이 될 수도, 1이 될 수도 있는 성분의 갯수라는 것을 알 수 있다. 여기서 임의의 $\lambda_{-1} \in \Lambda_{-1}$ 에 대해

$$|\Lambda_1(\lambda_{-1})| = \binom{p_{\lambda_{-1}}}{k}, \quad p_{\lambda_{-1}} \geq \frac{p}{4} - 1$$

가 성립함을 기억하자. 위 식의 오른쪽은 $n_{\lambda_{-1}} \cdot 2k \leq rk$, 즉 $n_{\lambda_{-1}} \leq r/2$ 임을 이용하면 유도할 수 있다:

$$p_{\lambda_{-1}} = r - n_{\lambda_{-1}} \geq \frac{r}{2} \geq \frac{p}{4} - 1.$$

부등식 $n_{\lambda_{-1}} \cdot 2k \leq rk$ 이 성립한다는 것은 좌변은 행렬 $\lambda_{-1} = \begin{pmatrix} \lambda_2(\theta)^T \\ \hline \vdots \\ \hline \lambda_r(\theta)^T \end{pmatrix}$ 에서 $n_{\lambda_{-1}}$ 에 해당하는 열들만 성분을 더한 것이고 우변은 행렬 $\lambda = \begin{pmatrix} \lambda_1(\theta)^T \\ \hline \lambda_{-1}(\theta)^T \end{pmatrix}$ 의 모든 성분을 더한 것이라는 것을 생각해보면 자명하다. 그러므로 p 가 충분히 크면 $\Lambda_1(\lambda_{-1})$ 은 공집합이 아니다. 이제 $\overline{\mathbb{P}}_{(1,1,\gamma_{-1},\lambda_{-1})}$ 의 정의를 다시 떠올려보면

$$\overline{\mathbb{P}}_{(1,1,\gamma_{-1},\lambda_{-1})} = \frac{1}{|\Theta_{(1,1,\gamma_{-1},\lambda_{-1})}|} \sum_{\theta \in \Theta_{(1,1,\gamma_{-1},\lambda_{-1})}} \mathbb{P}_\theta$$

이고 $\mathbb{P}_\theta$ 는 p차원 정규분포 $N_p(0, \Sigma(\theta))$ 를 따르는 n 개의 표본 $X_1, \ldots, X_n$ 에 대한 결합분포이다. 이때 $\theta \in \Theta_{(1,1,\gamma_{-1},\lambda_{-1})}$ 에 대응하는 $\Sigma(\theta)$ 들은

$$\Sigma(\theta) = I_p + \epsilon_{np} A_m(\lambda_1(\theta)) + \epsilon_{np} \sum_{m=2}^{r} \gamma_m A_m(\lambda_m)$$

의 형태로, 앞선 경우와는 다르게 θ 에 대응하는 $\lambda_1(\theta)$ 가 바뀔 때 마다 $\Sigma(\theta)$ 가 바뀐다는 것을 알 수 있다. 고를 수 있는 $\lambda_1(\theta)$ 의 갯수는 총 $\binom{p_{\lambda_{-1}}}{k}$ 개가 있으므로, 따라서 $\overline{\mathbb{P}}_{(1,1,\gamma_{-1},\lambda_{-1})}$ 는 평균이 0이고 공분산 행렬이 다음과 같은 형태인 p 차원 정규분포를 따르는 n 개의 표본에 대한 결합분포들 $\binom{p_{\lambda_{-1}}}{k}$ 개의 평균임을 알 수 있다:

$$\left(\begin{array}{c|c} 1 & r^T \\ \hline r & S_{(p-1)\times(p-1)} \end{array}\right).$$

이때 $r \in \mathbb{R}^{p-1}$ 은 0이 아닌 성분의 갯수가 k 개이며 0이 아닌 성분들은 모두 ϵ_{np} 이고, $S_{(p-1)\times(p-1)}$ 은 앞서 정의한 것과 같다.

이제 (Cai and Zhou, 2012, p.2411)에서 사용한 논증을 이용한다. 우선 (Cai and Zhou, 2012, Lemma 9)로부터 다음을 알 수 있다: 각 $i = 0, 1, 2$ 에 대해 g_i 를 정규분포 $N(0, \Sigma_i)$ 의 확률밀도함수라 하자. 그러면

$$\int \frac{g_1 g_2}{g_0} = \left[\det(I - \Sigma_0^{-2}(\Sigma_1 - \Sigma_0)(\Sigma_2 - \Sigma_0)\right]^{-1/2}$$

이다.

이때 주어진 $(\gamma_{-1}, \lambda_{-1})$ 에 대해 식 (16)의 좌변에서 적분 $\int \left(\frac{d\overline{\mathbb{P}}_{(1,1,\gamma_{-1},\lambda_{-1})}}{d\overline{\mathbb{P}}_{(1,0,\gamma_{-1},\lambda_{-1})}}\right)^2 d\overline{\mathbb{P}}_{(1,0,\gamma_{-1},\lambda_{-1})}$ 을 생각해본다. $\overline{\mathbb{P}}_{(1,0,\gamma_{-1},\lambda_{-1})}$ 는 p 차원 정규분포 $N_p(0, \Sigma_0)$ 를 따르는 n 개의 i.i.d. 다변량 정규분포들에 대한 결합분포이고 $\overline{\mathbb{P}}_{(1,1,\gamma_{-1},\lambda_{-1})}$ 는 p 차원 정규분포를 따르는 n 개의 표본에 대한 결합분포들 $\binom{p_{\lambda_{-1}}}{k}$ 개의 평균이다. 이제 λ_1 과 λ_1' 을 $\Lambda_1(\lambda_{-1})$ 에서 임의로 뽑고, 이들과 대응하는 공분산 행렬들을 각각 Σ_1 과 Σ_2 라 하자. g_i $(i = 0, 1, 2)$ 를 정규분포 $N_p(0, \Sigma_i)$ 의 확률밀도함수라 하면 적분 $\int \left(\frac{d\overline{\mathbb{P}}_{(1,1,\gamma_{-1},\lambda_{-1})}}{d\overline{\mathbb{P}}_{(1,0,\gamma_{-1},\lambda_{-1})}}\right)^2 d\overline{\mathbb{P}}_{(1,0,\gamma_{-1},\lambda_{-1})}$ 은 $\left(\int \frac{g_1 g_2}{g_0}\right)^n$ 형태의 적분들의 합으로 이루어진다는 것을 알 수 있다. 그래서 $R^{\gamma_{-1},\lambda_{-1}}_{\lambda_1,\lambda_1'}$ 을

$$R^{\gamma_{-1},\lambda_{-1}}_{\lambda_1,\lambda_1'} := -\log\det\left\{I_p - \Sigma_0^{-2}(\Sigma_0-\Sigma_1)(\Sigma_0-\Sigma_2)\right\}$$

와 같이 쓴다면, 식 (16)은 다음과 같이 쓸 수 있다.

$$\begin{aligned}
&\mathbb{E}_{(\gamma_{-1},\lambda_{-1})}\left\{\int\left(\frac{d\overline{\mathbb{P}}_{(1,1,\gamma_{-1},\lambda_{-1})}}{d\overline{\mathbb{P}}_{(1,0,\gamma_{-1},\lambda_{-1})}}\right)^2 d\overline{\mathbb{P}}_{(1,0,\gamma_{-1},\lambda_{-1})}-1\right\}\\
&=\mathbb{E}_{(\gamma_{-1},\lambda_{-1})}\left[\mathbb{E}_{(\lambda_1,\lambda_1')|\lambda_{-1}}\left\{\exp\left(\frac{n}{2}R^{\gamma_{-1},\lambda_{-1}}_{\lambda_1,\lambda_1'}\right)\right\}\right]\\
&=\mathbb{E}_{(\lambda_1,\lambda_1')}\left[\mathbb{E}_{(\gamma_{-1},\lambda_{-1})|(\lambda_1,\lambda_1')}\left\{\exp\left(\frac{n}{2}R^{\gamma_{-1},\lambda_{-1}}_{\lambda_1,\lambda_1'}\right)\right\}\right].
\end{aligned}\tag{18}$$

이때 $\lambda_1,\lambda_1'|\lambda_{-1}\overset{i.i.d}{\sim}\text{Unif}\{\Lambda_1(\lambda_{-1})\}$, $(\gamma_{-1},\lambda_{-1})|(\lambda_1,\lambda_1')\sim\text{Unif}\{\Theta_{-1}(\lambda_1,\lambda_1')\}$ 와 같은 분포가 주어져 있다고 생각하며, $\Theta_{-1}(\cdot,\cdot)$ 은

$$\Theta_{-1}(a_1,a_2) := \{0,1\}^{r-1}\times\{c\in\Lambda_{-1}\ :\ \exists\theta_i\in\Theta,\ i=1,2,\ \text{s.t.}\ \ \lambda_1(\theta_i)=a_i,\ \lambda_{-1}(\theta_i)=c\}$$

와 같이 주어져있다. Lemma 6의 결과를 이용하면 식 (18)은 아래의 식에 의해 유계이다:

$$\begin{aligned}
&\mathbb{E}_J\left[\exp\left\{-n\log\left(1-J\epsilon_{np}^2\right)\right\}\mathbb{E}_{(\lambda_1,\lambda_1')|J}\left\{\mathbb{E}_{(\gamma_{-1},\lambda_{-1})|(\lambda_1,\lambda_1')}\exp\left(\frac{n}{2}R^{\gamma_{-1},\lambda_{-1}}_{1,\lambda_1,\lambda_1'}\right)\right\}\right]\\
&\quad\le\mathbb{E}_J\left[\exp\left\{-n\log\left(1-J\epsilon_{np}^2\right)\right\}\frac{3}{2}-1\right].
\end{aligned}\tag{20}$$

여기서 J는 λ_1과 λ_1' 에서 0이 아닌 성분이 같은 위치에 몇개나 있는지를 센 숫자를 뜻한다. 즉, $J=\lambda_1^T\lambda_1'$ 이다. $\lambda_1,\lambda_1'|\lambda_{-1}\overset{i.i.d}{\sim}\text{Unif}\{\Lambda_1(\lambda_{-1})\}$ 임을 이용하면 각 $j=0,\dots,k$ 마다

$$\mathbb{E}_J\{I(J=j)|\lambda_{-1}\}=\frac{\binom{k}{j}\binom{p_{\lambda_{-1}}-k}{k-j}}{\binom{p_{\lambda_{-1}}}{k}}=\left(\frac{k!}{(k-j)!}\right)^2\frac{\{(p_{\lambda_{-1}}-k)!\}^2}{p_{\lambda_{-1}}!(p_{\lambda_{-1}}-2k+j)!}\frac{1}{j!}\le\left(\frac{k^2}{p_{\lambda_{-1}}-k}\right)^j$$

임을 얻게 된다. 그러므로 모든 λ_{-1}에 대해 $p_{\lambda_{-1}}\ge p/4-1$가 성립한다는 것을 이용하면

$$\mathbb{E}_J I(J=j)=\mathbb{E}_{\lambda_{-1}}\left[\mathbb{E}_J\left\{I(J=j)|\lambda_{-1}\right\}\right]\le\mathbb{E}_{\lambda_{-1}}\left\{\left(\frac{k^2}{p_{\lambda_{-1}}-k}\right)^j\right\}\le\left(\frac{k^2}{p/4-1-k}\right)^j$$

를 얻는다. 따라서, 식 (20)은 p가 충분히 크다면 다음 식에 의해 유계이다:

$$\begin{aligned}
&\sum_{j=0}^{k}\left(\frac{k^2}{p/4-1-k}\right)^j\left[\exp\left\{-n\log\left(1-j\epsilon_{np}^2\right)\right\}\frac{3}{2}-1\right]\\
&=\frac{1}{2}+\sum_{j=1}^{k}\left(\frac{k^2}{p/4-1-k}\right)^j\left[\exp\left\{-n\log\left(1-j\epsilon_{np}^2\right)\right\}\frac{3}{2}-1\right]\\
&\le\frac{1}{2}+\frac{3}{2}\left(\frac{k^2}{p/4-1-k}\right)^j p^{2\nu^2 j}\le\frac{1}{2}+\frac{3C}{2}p^{-\epsilon j}p^{(\epsilon/2)j}\le c_2^2,
\end{aligned}$$

이때 $C>0$는 상수이고 $c_2^2=3/4<1$이며, 두번째 부등식은 $s_0^2=O(p^{3-\epsilon})$과 $\nu=\sqrt{\epsilon/4}$ 임을 이용하면 얻을 수 있다. 이로써 (13)식에 대한 증명이 마무리되었다.

5 백승찬 - Lemma 6

참고문헌

References

Banerjee, S. and Ghosal, S. (2015). Bayesian structure learning in graphical models. *Journal of Multivariate Analysis*, 136:147–162.

Cai, T. T. and Zhou, H. H. (2012). Optimal rates of convergence for sparse covariance matrix estimation. *The Annals of Statistics*, pages 2389–2420.

Carvalho, C. M., Polson, N. G., and Scott, J. G. (2010). The horseshoe estimator for sparse signals. *Biometrika*, 97(2):465–480.

Lee, K., Jo, S., and Lee, J. (2022). The beta-mixture shrinkage prior for sparse covariances with near-minimax posterior convergence rate. *Journal of Multivariate Analysis*, 192:105067.

Song, Q. and Liang, F. (2017). Nearly optimal bayesian shrinkage for high dimensional regression. *arXiv preprint arXiv:1712.08964*.