Week 11
Spring 2009

Lecture 21. Estimation of Large Covariance Matrices: Lower bound (II)

Observe

$$\mathbf{X}_1, \mathbf{X}_2, \ldots, \mathbf{X}_n \text{ i.i.d. from a } p\text{-variate Gaussian distribution, } N(\mu, \Sigma_{p\times p}).$$

We assume that the covariance matrix $\Sigma_{p\times p} = (\sigma_{ij})_{1\le i,j\le p}$ is contained in the following parameter space,

$$\mathcal{F}(\alpha, \varepsilon, M) = \left\{\Sigma : |\sigma_{ij}| \le M|i-j|^{-(\alpha+1)} \text{ for all } i \ne j \text{ and } \lambda_{\max}(\Sigma) \le 1/\varepsilon\right\} \tag{1}$$

Theorem 1 *Under the assumption (1), we have*

$$\inf_{\hat{\Sigma}} \sup_{\mathcal{F}} \mathbb{E}\left\|\hat{\Sigma} - \Sigma\right\|^2 \ge cn^{-\frac{2\alpha}{2\alpha+1}} + c\frac{\log p}{n}. \tag{2}$$

Last time we have shown

$$\inf_{\hat{\Sigma}} \sup_{\mathcal{F}} \mathbb{E}\left\|\hat{\Sigma} - \Sigma\right\|^2 \ge c\frac{\log p}{n}.$$

In this lecture we will show

$$\inf_{\hat{\Sigma}} \sup_{\mathcal{F}} \mathbb{E}\left\|\hat{\Sigma} - \Sigma\right\|^2 \ge cn^{-\frac{2\alpha}{2\alpha+1}}$$

by the Assouad's lemma.

We shall now define a parameter space that is appropriate for the minimax lower bound argument. For given positive integers k and m with $2k \le p$ and $1 \le m \le k$, define the $p \times p$ matrix $B(m,k) = (b_{ij})_{p\times p}$ with

$$b_{ij} = I\{i = m \text{ and } m+1 \le j \le 2k, \text{ or } j = m \text{ and } m+1 \le i \le 2k\}.$$

Set $k = n^{\frac{1}{2\alpha+1}}$ and $a = k^{-(\alpha+1)}$. We then define the collection of 2^k covariance matrices as

$$\mathcal{H} = \left\{\Sigma(\theta) : \Sigma(\theta) = I_p + \tau a \sum_{m=1}^{k} \theta_m B(m,k), \quad \theta = (\theta_m) \in \{0,1\}^k\right\} \tag{3}$$

where I_p is the $p \times p$ identity matrix and τ is a constant. It is easy to check that as long as $0 < \tau < \min\{M, (1-\varepsilon)/2\}$ the collection $\mathcal{H} \subset \mathcal{F}_\alpha(\varepsilon, M)$. We will show

$$\inf_{\hat{\Sigma}} \sup_{\mathcal{H}} \mathbb{E}\left\|\hat{\Sigma} - \Sigma\right\|^2 \ge cn^{-\frac{2\alpha}{2\alpha+1}} \tag{4}$$

A Lower bound by the Assouad's Lemma

We first prove equation (4). Let $\mathbf{X}_1, \mathbf{X}_2, \ldots, \mathbf{X}_n$ be i.i.d. $N(0, \Sigma(\theta))$ with $\Sigma(\theta) \in \mathcal{H}$. Denote the joint distribution by P_θ. We apply Assouad's Lemma to the parameter space $\mathcal{H}$,

$$\max_{\theta \in \mathcal{H}} 2^2 E_\theta \left\| \hat{\Sigma} - \Sigma(\theta) \right\|^2 \geq \min_{H(\theta, \theta') \geq 1} \frac{\left\| \Sigma(\theta) - \Sigma(\theta') \right\|^2}{H(\theta, \theta')} \frac{k}{2} \min_{H(\theta, \theta') = 1} \| P_\theta \wedge P_{\theta'} \|$$

From Lemma 2 we have

$$\min_{H(\theta, \theta') \geq 1} \frac{\left\| \Sigma(\theta) - \Sigma(\theta') \right\|^2}{H(\theta, \theta')} \geq cka^2$$

and from Lemma 3,

$$\min_{H(\theta, \theta') = 1} \| P_\theta \wedge P_{\theta'} \| \geq c > 0$$

thus

$$\max_{\theta \in \mathcal{F}_{11}} 2^2 E_\theta \left\| \hat{\Sigma} - \Sigma(\theta) \right\|^2 \geq \frac{c^2}{2} k^2 a^2 \geq c_1 n^{-\frac{2\alpha}{2\alpha+1}}.$$

Now we give proofs of auxiliary lemmas.

Lemma 2 *For $\Sigma(\theta)$ defined in (3) we have*

$$\min_{H(\theta, \theta') \geq 1} \frac{\left\| \Sigma(\theta) - \Sigma(\theta') \right\|^2}{H(\theta, \theta')} \geq cka^2.$$

Proof of Lemma 2: We define $v = (1\{k \leq i \leq 2k\})$. Let

$$\left[\Sigma(\theta) - \Sigma(\theta') \right] v = (w_i).$$

There are exactly $H(\theta, \theta')$ number of w_i such that $|w_i| = ka$ (just consider upper half of the matrix), which implies

$$\left\| \left[\Sigma(\theta) - \Sigma(\theta') \right] v \right\|_2^2 \geq H(\theta, \theta') \cdot (ka)^2$$

and so $\left\| \Sigma(\theta) - \Sigma(\theta') \right\|^2 \geq H(\theta, \theta') \cdot (ka)^2 / k \geq cka^2$.

Lemma 3 *Let P_θ be the joint distribution of n i.i.d. $\mathbf{X}_1, \mathbf{X}_2, \ldots, \mathbf{X}_n$ with $\mathbf{X}_1 \sim N(0, \Sigma(\theta))$ and $\Sigma(\theta) \in \mathcal{F}_{11}$. Then for some $c_1 > 0$ we have*

$$\min_{H(\theta, \theta') = 1} \| P_\theta \wedge P_{\theta'} \| \geq c_1.$$

Proof of Lemma 3: When $H(\theta, \theta') = 1$, we will show

$$\begin{aligned} \| P_{\theta'} - P_\theta \|_1^2 &\leq 2K(P_{\theta'} | P_\theta) = 2n \left[\frac{1}{2} tr \left(\Sigma(\theta') \Sigma^{-1}(\theta) \right) - \frac{1}{2} \log \det \left(\Sigma(\theta') \Sigma^{-1}(\theta) \right) - \frac{p}{2} \right] \\ &\leq n \cdot cka^2 \end{aligned}$$

for some small $c > 0$, where $K(\cdot|\cdot)$ is the Kullback–Leibler divergence and the first inequality follows from the well known Pinsker's inequality (see, e.g., Csiszár (1967)). This immediately implies the L_1 distance between two measures is bounded away from 1, and then the lemma follows. Write

$$\Sigma(\theta') = D_1 + \Sigma(\theta).$$

Then

$$\frac{1}{2}tr\left(\Sigma(\theta')\Sigma^{-1}(\theta)\right) - \frac{p}{2} = \frac{1}{2}tr\left(D_1\Sigma^{-1}(\theta)\right).$$

Let λ_i be the eigenvalues of $D_1\Sigma^{-1}(\theta)$. Since $D_1\Sigma^{-1}(\theta)$ is similar to the symmetric matrix $\Sigma^{-1/2}(\theta)D_1\Sigma^{-1/2}(\theta)$, and

$$\left\|\Sigma^{-1/2}(\theta)D_1\Sigma^{-1/2}(\theta)\right\| \le \left\|\Sigma^{-1/2}(\theta)\right\|\|D_1\|\left\|\Sigma^{-1/2}(\theta)\right\| \le c_1\|D_1\| \le c_1\|D_1\|_1 \le c_2 ka,$$

then all eigenvalues λ_i's are real and in the interval $[-c_2 ka, c_2 ka]$, where $ka = k \cdot k^{-(\alpha+1)} = k^{-\alpha} \to 0$. Note that the Taylor expansion yields

$$\log\det\left(\Sigma(\theta')\Sigma^{-1}(\theta)\right) = \log\det\left(I + D_1\Sigma^{-1}(\theta)\right) = tr\left(D_1\Sigma^{-1}(\theta)\right) - R_3$$

where

$$R_3 \le c_3\sum_{i=1}^{p}\lambda_i^2 \text{ for some } c_3 > 0.$$

Write $\Sigma^{-1/2}(\theta) = UV^{1/2}U^T$, where $UU^T = I$ and V is a diagonal matrix. It follows from the fact that the Frobenius norm of a matrix remains the same after an orthogonal transformation that

$$\sum_{i=1}^{p}\lambda_i^2 = \left\|\Sigma^{-1/2}(\theta)D_1\Sigma^{-1/2}(\theta)\right\|_F^2 \le \|V\|^2 \cdot \left\|U^T D_1 U\right\|_F^2 = \left\|\Sigma^{-1}(\theta)\right\|^2 \cdot \|D_1\|_F^2 \le c_4 ka^2. \quad \blacksquare$$