Econometrics **June 30, 2023**

Topic 18: Eigenvalue and Spike Models

by Sai Zhang

Key points: .

Disclaimer: *The note is built on Prof. Jinchi Lv's lectures of the course at USC, DSO 607, High-Dimensional Statistics and Big Data Problems.*

18.1 Motivation

Consider n independent observations $\mathbf{X}_i \in \mathbb{R}^p$ drawn from a $\mathcal{N}(\mathbf{0}, \mathbf{\Sigma})$, then the covariance can be decomposed into 2 parts, white noise and low rank

$$\mathbf{\Sigma} = \mathrm{Cov}(\mathbf{X}_i) = \mathbf{I} + \sum_{k=1}^{M} \theta_k \nu_k \nu_k' = \mathbf{\Sigma}_0 + \mathbf{\Phi}$$

where M denotes the **number of spikes** in the distribution of eigenvalues. The idea is: spikes deviate from a reference model along a **small fixed number** of unknown directions. If $\mathbf{\Phi} = \mathbf{0}$, then none of the sample eigenvalues is separated from the bulk.

Why a spike model is interesting? A spike model can help determine the latent dimension of the data, some examples being

- Principal component analysis (PCA): spikes are related to the directions of the most variations of the data, i.e., the principal components
- Clustering model: M spikes is equivalent to $M + 1$ clusters
- Economic significance: M is related to the number of factor loadings

Then the question is threefold:

- How to determine M
- How to estimate ν_k
- How to test θ_k

Under rank one alternative, we would like to test the hypothesis

$$H_1 : \mathbf{\Sigma} = \mathbf{I}_p + \theta \boldsymbol{\nu}\boldsymbol{\nu}', \theta > 0$$

against the null

$$H_0 : \mathbf{\Sigma} = \mathbf{I}_p$$

with the key assumptions:

A1 Gaussian error
A2 large p: $p \leq n$ but allows $p/n \to \gamma \in (0, 1)$

Under these assumptions, for the $n \times p$ data matrix $\mathbf{X} = \begin{pmatrix} \mathbf{X}_1' & \cdots & \mathbf{X}_n' \end{pmatrix}'$, $\mathbf{X}'\mathbf{X}$ has a $p-$dimensional **Wishart** distribution $W_p(n, \boldsymbol{\Sigma})$ with the degree of freedom n and covariance matrix $\boldsymbol{\Sigma}$, which is a *random matrix*.

If $\mathbf{Y} = \mathbf{M} + \mathbf{X}$, that is, the sum of the *random matrix* $\mathbf{X}$ and a *deterministic matrix* $\mathbf{M}$ (also $n \times p$), then $\mathbf{Y}'\mathbf{Y}$ has a $p-$dimensional Wishart distribution $W_p(n, \boldsymbol{\Sigma}, \boldsymbol{\Psi})$ with n degrees of freedom, covariance matrix $\boldsymbol{\Sigma}$ and non-centrality matrix $\boldsymbol{\Psi} = \boldsymbol{\Sigma}^{-1}\mathbf{M}'\mathbf{M}$.

Definition 18.1.1: Density of Wishart Distribution

The PDF of Wishart distribution is defined as

$$f(\mathbf{X}) = \frac{1}{2^{np/2}\Gamma_p\left(\frac{n}{2}\right)|\boldsymbol{\Sigma}|^{n/2}} |\mathbf{X}|^{(n-p-1)/2} \exp\left(-\frac{1}{2}\mathrm{tr}\left(\boldsymbol{\Sigma}^{-1}\mathbf{X}\right)\right)$$

where $\mathbf{X}$ is a symmetric positive semidefinite and $\Gamma_p\left(\frac{n}{2}\right)$ is a multivariate gamma function such that

$$\Gamma_p\left(\frac{n}{2}\right) = \pi^{\frac{p(p-1)}{4}} \prod_{j=1}^{p} \Gamma\left(\frac{n}{2} - \frac{j-1}{2}\right)$$

Notice that the sample covariance matrix $\mathbf{S} = \frac{1}{n}\mathbf{X}'\mathbf{X}$ is just a scaled version of Wishart distribution

$$n\mathbf{S} = \mathbf{X}'\mathbf{X} \sim W_p(n, \boldsymbol{\Sigma})$$

For $\boldsymbol{\Sigma} = \mathbf{I}_p$, the empirical distribution fo eigenvalues converges to Marcenko-Pastur distribution

$$f^{\mathrm{MP}}(x) = \frac{1}{2\pi\gamma x}\sqrt{(b_+ - x)(x - b_-)}$$

where $b_{\pm} = (1 \pm \sqrt{\gamma})^2$. Then:

- under $H_0 : \boldsymbol{\Sigma} = \mathbf{I}_p$, we have
$$n^{2/3}\left(\frac{\lambda_1 - \mu(\gamma)}{\sigma(\gamma)}\right) \xrightarrow{d} \mathrm{TW}_1$$
  where TW_1 is the Tracy-Widom distribution
- under $H_1 : \boldsymbol{\Sigma} = \mathbf{I}_p + \theta\boldsymbol{\nu}\boldsymbol{\nu}'$, $\theta > 0$, if θ is strong ($\theta \gg \sqrt{\gamma}$), then
$$n^{1/2}\left(\frac{\lambda_1 - \rho(\theta, \gamma)}{\tau(\theta, \gamma)}\right) \xrightarrow{d} \mathcal{N}(0, 1)$$

Here, the largest eigenvalue test is the best test. But when the signal is weak ($\theta < \sqrt{\gamma}$)

References