Applied Multivariate Statistical Analysis

June 21, 2013

8-9. *(A test that all variables are independent.)*

(a) Consider that the normal theory likelihood ratio test of H_0: $\mathbf{\Sigma}$ is the diagonal matrix

$$\begin{bmatrix} \sigma_{11} & 0 & \cdots & 0 \\ 0 & \sigma_{22} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sigma_{pp} \end{bmatrix}, \quad \sigma_{ii} > 0$$

Show that the test is as follows: Reject H_0 if

$$\Lambda = \frac{|\mathbf{S}|^{n/2}}{\prod_{i=1}^{p} s_{ii}^{n/2}} = |\mathbf{R}|^{n/2} < c$$

For a large sample size, $-2\ln\Lambda$ is approximately $\chi^2_{p(p-1)/2}$. Bartlett[3] suggests that the test statistic $-2[1-(2p+11)/6n]\ln\Lambda$ be used in place of $-2\ln\Lambda$. This results in an improved chi-square approximation. The large sample α critical point is $\chi^2_{p(p-1)/2}(\alpha)$. Note that testing $\mathbf{\Sigma} = \mathbf{\Sigma_0}$ is the same as testing $\boldsymbol{\rho} = \boldsymbol{I}$.

(b) Show that the likelihood ratio test of H_0: $\mathbf{\Sigma} = \boldsymbol{\sigma}^2\mathbf{I}$ rejects H_0 if

$$\Lambda = \frac{|\mathbf{S}|^{n/2}}{(\mathrm{tr}(\mathbf{S})/p)^{np/2}} = \left[\frac{\prod_{i=1}^{p}\hat{\lambda}_i}{\left(\frac{1}{p}\sum_{i=1}^{p}\hat{\lambda}_i\right)^p}\right]^{n/2} = \left[\frac{\text{geometric mean } \hat{\lambda}_i}{\text{arithmetic mean } \hat{\lambda}_i}\right]^{np/2} < c$$

For a large sample size, Bartlett[3] suggests that

$$-2[1-(2p^2+p+2)/6pn]\ln\Lambda$$

is approximately $\chi^2_{(p+2)(p-1)/2}$. Thus, the large sample α critical point is $\chi^2_{(p+2)(p-1)/2}(\alpha)$. This test is called a *sphericity test*, because the constant density contours are spheres where $\mathbf{\Sigma} = \boldsymbol{\sigma}^2\mathbf{I}$.

Proof 8-9(a): The key is to find $\max_{\boldsymbol{\mu}} L(\boldsymbol{\mu}, \mathbf{\Sigma_0})$

By 5-12:

$$\text{likelihood ratio} = \Lambda = \frac{\max_{\boldsymbol{\mu}} L(\boldsymbol{\mu}, \mathbf{\Sigma_0})}{\max_{\boldsymbol{\mu},\mathbf{\Sigma}} L(\boldsymbol{\mu}, \mathbf{\Sigma})}$$

Since

$$L(\boldsymbol{\mu}, \mathbf{\Sigma_0}) = \frac{1}{(2\pi)^{np/2} \cdot |\mathbf{\Sigma_0}|^{n/2}} \cdot \exp\left(-\sum_{j=1}^{n}(x_j - \boldsymbol{\mu})' \cdot \mathbf{\Sigma_0^{-1}} \cdot (x_j - \boldsymbol{\mu}) \Big/ 2\right) \quad \text{(by 4-11)}$$

$$= \frac{1}{(2\pi)^{np/2}\prod_{i=1}^{p}\sigma_{ii}^{n/2}} \cdot \exp\left(-\sum_{j=1}^{n}\sum_{i=1}^{p}(x_{ji} - \mu_i)^2 \Big/ \sigma_{ii} \cdot 2\right) \quad \text{(plug in } \mathbf{\Sigma_0}\text{)}$$

$$= \prod_{i=1}^{p}\left(\frac{1}{(2\pi)^{n/2}\sigma_{ii}^{n/2}} \cdot \exp\left(-\sum_{j=1}^{n}(x_{ji} - \mu_i)^2 \Big/ 2\sigma_{ii}\right)\right)$$

Hence

$$\max_{\boldsymbol{\mu}} L(\boldsymbol{\mu}, \mathbf{\Sigma_0}) = \prod_{i=1}^{p}\left(\frac{1}{(2\pi)^{n/2}\sigma_{ii}^{n/2}} \cdot \exp\left(-\sum_{j=1}^{n}(x_{ji} - \mu_i)^2 \Big/ 2\sigma_{ii}\right)\right) \quad (1)$$

Hence maximum estimators are

$$\hat{\mu_i} = \frac{1}{n}\sum_{j=1}^{n} x_{ji}, \quad \hat{\sigma_{ii}} = \frac{1}{n}\sum_{j=1}^{n}(x_{ji} - \bar{x_i})^2 \quad \text{(result 4.11)}$$

Hence

$$\max_{\boldsymbol{\mu}} L(\boldsymbol{\mu}, \mathbf{\Sigma_0}) = \prod_{i=1}^{p}(2\pi)^{-n/2} \cdot \hat{\sigma_{ii}}^{-n/2} \cdot \exp(-n/2)$$

$$= \prod_{i=1}^{p}(2\pi)^{-n/2} \cdot s_{ii}^{-n/2} \cdot \left(\frac{n-1}{n}\right)^{-n/2} \cdot \exp(-n/2)$$

From 5-10:

$$\max_{\boldsymbol{\mu}, \mathbf{\Sigma}} = \frac{1}{(2\pi)^{np/2} \cdot |\hat{\mathbf{\Sigma}}|^{n/2}} \cdot e^{-np/2}$$

$$= (2\pi)^{-np/2} \cdot \left|\frac{n-1}{n} \cdot S\right|^{-n/2} \cdot e^{-np/2}$$

We know

$$\Lambda = \frac{\max_{\boldsymbol{\mu}} L(\boldsymbol{\mu}, \mathbf{\Sigma_0})}{\max_{\boldsymbol{\mu}, \mathbf{\Sigma}} L(\boldsymbol{\mu}, \mathbf{\Sigma})}$$

$$= \frac{\prod_{i=1}^{p}(2\pi)^{-n/2} \cdot \left(\frac{n-1}{n}s_{ii}\right)^{-n/2} \cdot \exp(-n/2)}{(2\pi)^{-np/2}\left|\frac{n-1}{n}S\right|^{-n/2} \cdot \exp(-np/2)}$$

$$= \frac{|\mathbf{S}|^{n/2}}{\prod_{i=1}^{p} s_{ii}^{n/2}}$$

$$= |\mathbf{R}|^{n/2} \quad \text{(8-28)}$$

8-9(b): From 8-9(a) formula (1)

$$\max_{\boldsymbol{\mu}} L(\boldsymbol{\mu}, \boldsymbol{\Sigma_0}) = \prod_{i=1}^{p} (2\pi)^{-n/2} \cdot \sigma_{ii}^{-n/2} \cdot \exp\left(-\sum_{j=1}^{n} (x_{ji} - \mu_i)^2 \Big/ 2\sigma_{ii} \right)$$

Now under H_0, all the $\sigma_{ii} = \sigma^2$

Hence

$$\max_{\boldsymbol{\mu}} L(\boldsymbol{\mu}, \boldsymbol{\Sigma_0}) = (2\pi)^{-np/2} \cdot (\sigma^2)^{-np/2} \cdot \exp\left(-\sum_{j=1}^{n} \sum_{i=1}^{p} (x_{ji} - \mu_i)^2 \Big/ 2\sigma^2 \right)$$

To find the maximum estimators $\hat{\mu_i}$ & $\hat{\sigma}^2$ take derivatives with respect to μ_i & σ^2 respectively and set them to 0 to find:

$$\begin{aligned} \hat{\mu_i} &= \frac{1}{n} \sum_{j=1}^{n} x_{ji} \\ \hat{\sigma}^2 &= \sum_{j=1}^{n} \sum_{i=1}^{p} (x_{ji} - \bar{x_i})^2 / np \\ &= \frac{n-1}{n} tr(\mathbf{S})/p \end{aligned}$$

Hence:

$$\begin{aligned} \Lambda &= \frac{\max_{\boldsymbol{\mu}} L(\boldsymbol{\mu}, \boldsymbol{\Sigma_0})}{\max_{\boldsymbol{\mu}, \boldsymbol{\Sigma}} L(\boldsymbol{\mu}, \boldsymbol{\Sigma})} \\ &= \frac{(2\pi)^{-np/2} (\hat{\sigma}^2)^{-np/2} \cdot \exp(-np/2)}{(2\pi)^{-np/2} \cdot |\hat{\boldsymbol{\Sigma}}|^{-n/2} \cdot \exp(-np/2)} \\ &= \frac{|\hat{\boldsymbol{\Sigma}}^{n/2}|}{(\hat{\boldsymbol{\sigma}}^2)^{np/2}} = \frac{\left| \frac{n-1}{n} \cdot \mathbf{S} \right|^{n/2}}{\left(\frac{n-1}{n} tr(\mathbf{S})/p \right)^{np/2}} \\ &= \frac{|\mathbf{S}|^{n/2}}{(tr(\mathbf{S})/p)^{np/2}} = \left[\frac{\prod \hat{\lambda}_i}{\left(\frac{1}{p} \sum \hat{\lambda}_i \right)^p} \right]^{n/2} = \left(\frac{\text{geo mean } \ \hat{\lambda}_i}{\text{ari mean } \ \hat{\lambda}_i} \right)^{np/2} \end{aligned}$$

9-5. Let $\mathbf{S}$ be a matrix with eigenvalue-eigenvector pairs $(\lambda_1, \mathbf{e}_1)$, $(\lambda_2, \mathbf{e}_2)$, ..., $(\lambda_p, \mathbf{e}_p)$, where $\lambda_1 \geq \lambda_2 \geq ... \geq \lambda_p$. Let $m < p$ and define

$$\mathbf{L} = \{l_{ij}\} = \left[\sqrt{\lambda_1 \mathbf{e}_1} \mid \sqrt{\lambda_2 \mathbf{e}_2} \mid ... \mid \sqrt{\lambda_m \mathbf{e}_m}\right]$$

and

$$\mathbf{\Psi} = \begin{pmatrix} \psi_1 & 0 & ... & 0 \\ 0 & \psi_2 & ... & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & ... & \psi_p \end{pmatrix} \text{ with } \psi_i = s_{ii} - \sum_{j=1}^{m} l_{ij}^2$$

Prove the inequality

$$\text{Sum of squared entries of } (\mathbf{S} - (\mathbf{LL}' + \mathbf{\Psi})) \leq \lambda_{m+1}^2 + \cdots + \lambda_p^2$$

Proof: By definition of ψ_i, we know that the diagonal of $(\mathbf{S} - (\mathbf{LL}' + \mathbf{\Psi}))$ are all zeroes. Since $(\mathbf{S} - (\mathbf{LL}' + \mathbf{\Psi})))$ and $(\mathbf{S} - \mathbf{LL}')$ have the same elements except on the diagonal, we know that

$$(\text{Sum of squared entries of } (\mathbf{S} - (\mathbf{LL}' + \mathbf{\Psi}))) \leq \text{ Sum of squared entries of } (\mathbf{S} - \mathbf{LL}')$$

Since $\mathbf{S} = \lambda_1 \mathbf{e}_1 \mathbf{e}_1' + \cdots + \lambda_p \mathbf{e}_p \mathbf{e}_p'$ and $\mathbf{LL}' = \lambda_1 \mathbf{e}_1 \mathbf{e}_1' + \cdots + \lambda_m \mathbf{e}_m \mathbf{e}_m'$, $\mathbf{S} - \mathbf{LL}' = \lambda_{m+1} \mathbf{e}_{m+1} \mathbf{e}_{m+1}' + \cdots + \lambda_p \mathbf{e}_p \mathbf{e}_p'$. Writing it in matrix form, this is saying $\mathbf{S} - \mathbf{LL}' = \mathbf{P}_2 \mathbf{\Lambda}_2 \mathbf{P}_2'$ where $\mathbf{P}_2 = [\mathbf{e}_{m+1} | \cdots | \mathbf{e}_p]$ and $\mathbf{\Lambda}_2 = Diag(\lambda_{m+1}, \cdots, \lambda_p)$. Now

$$\begin{aligned}
&(\text{Sum of squared entries of } \mathbf{S} - \mathbf{LL}') \\
=&\text{tr}((\mathbf{S} - \mathbf{LL}')(\mathbf{S} - \mathbf{LL}')') && \text{(matrix calculation)} \\
=&\text{tr}((\mathbf{P}_2 \mathbf{\Lambda}_2 \mathbf{P}_2')(\mathbf{P}_2 \mathbf{\Lambda}_2 \mathbf{P}_2')') \\
=&\text{tr}(\mathbf{P}_2 \mathbf{\Lambda}_2 \mathbf{\Lambda}_2 \mathbf{P}_2') && (\mathbf{P}_2' \mathbf{P}_2' = \mathbf{I}) \\
=&\text{tr}(\mathbf{\Lambda}_2 \mathbf{\Lambda}_2) \\
=&\lambda_{m+1}^2 + \cdots + \lambda_p^2.
\end{aligned}$$

and this completes the proof.