Multivariate Methoden: Übungsblatt 4

Aufgabe 1: Asymptotische Eigenschaften der Hauptkomponenten

Theorem Sei $\boldsymbol{\Sigma} > 0$ mit voneinander verschiedenen Eigenwerten mit der Eigenwertzerlegung $\boldsymbol{\Sigma} = \boldsymbol{\Gamma\Lambda\Gamma}'$. Dann gilt

(a) $\sqrt{n-1}(\hat{\boldsymbol{\lambda}} - \boldsymbol{\lambda}) \xrightarrow{\mathcal{L}} N_p(0, 2\boldsymbol{\Lambda}^2)$,

(b) $\sqrt{n-1}(\hat{\boldsymbol{\gamma}}_j - \boldsymbol{\gamma}_j) \xrightarrow{\mathcal{L}} N_p(0, 2\mathbf{V}_j)$,
mit $\mathbf{V}_j = \lambda_j \sum\limits_{k \neq j} \frac{\lambda_k}{(\lambda_k - \lambda_j)^2} \boldsymbol{\gamma}_k \boldsymbol{\gamma}'_k$,

(c) die Elemente in $\hat{\lambda}$ sind von den Elementen von $\hat{\boldsymbol{\Gamma}}$ asymptotisch unabhängig.

Leiten Sie die unter Verwendung des obigen Theorems die asymptotische Verteilung von $\log \hat{\lambda}_j$ her und konstruieren Sie daraus ein Konfidenzintervall für λ_j.

Aufgabe 2

2.1. Führen Sie eine Hauptkomponentenanalyse der Variablen des Datensatzes BANK.DTA durch (d.h. für alle Variablen außer derjenigen, die angibt, ob die betreffende Banknote echt ist). Gehen sie von der Hauptkomponentenanalyse mit der Kovarianzmatrix aus.
2.1,a) Welche Variable geht wie stark in welche Hauptkomponente ein?
2.2,b) Betrachten Sie die Anteile an erklärten Varianz und erstellen Sie ein scree-plot der Ergebnissen der Hauptkomponententransformation. Wieviele Hauptkomponenten würden Sie zur adequaten Beschreibung der Daten für notwendig erachten?