

Fragenkatalog

3. Test, Multivariate Verfahren 2

SBWL Tourismusanalyse und Freizeitmarketing, VK4

Achim Zeileis

Aus diesem Fragenkatalog werden die Theoriefragen im 3. Test des VK4 gestellt. Die Fragen sind auf Basis der LV-Unterlagen zu beantworten. Zusätzlich zu den Theoriefragen wird der Test Anwendungsbeispiele in R enthalten. In diesen Anwendungsaufgaben sind jeweils R-Output und Grafiken angegeben und müssen anhand von Fragen interpretiert werden. Die R-Aufgaben sind immer an die Probleme angelehnt, die in den Übungen und Tutorien diskutiert wurden.

- Welche Möglichkeiten der Standardisierung einer Datenmatrix mit n Beobachtungen von p Variablen wurden diskutiert? Was ist die Idee der jeweiligen Standardisierung?
- Was ist die Hauptkomponentenzerlegung einer Datenmatrix X ? Was gibt die erste Hauptkomponente an?
- Wie können Abstände zwischen Beobachtungen gemessen werden? Wie sind Manhattan-, Euklidische bzw. Maximumsdistanz definiert?
- Was ist das grundsätzliche Ziel der Clusteranalyse, welche Eigenschaften sollen die Cluster einer Klassifikation haben?
- Was ist eine Partition, was eine Hierarchie von Partitionen?
- Wie wird beim agglomerativen Clustern eine Hierarchie von Partitionen konstruiert? Illustrieren dies anhand der single/complete/average Linkage-Methode
- Durch welche Eigenschaft ist die k -Means-Partition definiert? Beschreiben Sie einen Algorithmus zur Konstruktion.
- Wie wählt man die Anzahl der Hauptkomponenten?
- Wie wählt man die Anzahl Cluster beim hierarchischen Cluster?
- Wie wählt man die Anzahl Cluster bei k -Means?
- Beschreiben Sie wie man mit Hilfe des EM-Algorithmus Mischmodelle schätzen kann.