

# Übungsvorschlag

[alle Lösungen sind in \multivariate methoden\2multivariate methoden\folien uebung]

## Aufgabe 1

Der Datensatz BOSTON.DTA (Quelle: *Härdle, W./L. Simar: Applied Multivariate Analysis*) enthält die durchschnittlichen Preise von Wohneigentum (= *VALUEOFHOME*) in 506 Distrikten der Stadt Boston zusammen mit einer Reihe von möglicherweise mit diesen Preisen zusammenhängenden Größen (*CRIMRATE* = Kriminalitätsrate, *RESIDENTIALZONE* = Gebietsanteil von Wohnsiedlungen, *NONRETAIL* = Gebietsanteil von Gewerbebetrieben ohne Einzelhandel, *RIVER* = Indikator für Flußnähe, *NOCONCENTRATION* = Stickoxid-Konzentration, *NOOFFROOMS* = durchschnittliche Zimmeranzahl, *AGE* = Anteil Wohnungen mit Baujahr vor 1940, *DISTANCE* = gewichtete Entfernungen zu fünf Bostoner Wirtschaftszentren, *ACCESSTOHIGHWAY* = Verkehrsanbindung zum Highway, *TAXRATE* = Steuersatz Grundsteuer, *PUPILTEACHER* = Verhältnis Schüler/Lehrer, *BLACKPEOPLE* = Maß für Bevölkerungsanteil von Farbigen in nicht überwiegend von Farbigen bewohnten Distrikten, *LOWERSTATUS* = Bevölkerungsanteil unterprivilegierter Schichten).

*Ausnahme:* *RIVER* gelöscht wegen der geringen Variation.

- a) Führen Sie erst eine Schätzung des Faktormodells mit der Methode der Hauptkomponenten um die approximative Anzahl an Faktoren zu bestimmen. Wie werden die Faktorladungen bei jedem Faktor interpretiert? Berechnen Sie die vorhergesagte Faktoren.
- b) Führen Sie nun eine Schätzung des Faktormodells mit der iterativen Methode der Hauptkomponenten gegeben die Anzahl an zugelassenen Faktoren aus der Teilaufgabe 1,a). Interpretieren Sie die geschätzte spezifische Komponenten und vergleichen Sie sie mit der spezifischen Komponenten aus der Teilaufgabe 1,a). Berechnen sie die Vorhergesagte Faktoren und vergleichen sie mit der Vorhersagen der Teilaufgabe 1,a). Welche Methode ist besser?.
- c) Führen Sie schließlich eine ML Schätzung des Faktormodells und mit Hilfe eines LR Tests veranstalten sie eine Modellselektion für die maximale Anzahl an Faktoren. Welches Problem tritt auf? Wie viel Faktoren werden von dem Selektionsverfahren empfohlen? Berechnen Sie die vorhergesagte Faktoren und vergleichen Sie sie mit denen aus der Teilaufgabe 1,b).
- d) Wie viel Faktoren würden Sie endgültig behalten und welche Methode würden sie für die Endgültige Schätzung verwenden?