

Funktionale Datenanalyse

(Abgabe: Mo., 13.05.2013)

Betrachte die Datei `datenb3.txt` von der Homepage. Sie enthält Messzeiten t_i (in der ersten Spalte) und zu diesen Zeitpunkten gemessene Werte y_i (in der zweiten Spalte), für $i = 1, \dots, 101$. Die Messwerte enthalten zufällige Messfehler, die unabhängig und identisch verteilt mit Mittelwert 0 sind.

1. Verwende den Nadaraya-Watson-Schätzer (mit Normalverteilungskern) und versuche manuell, eine passende Bandbreite zu finden. Plote die Punkte und die geschätzten (mindestens 3) Regressionsfunktionen. Um NW-Schätzer zu berechnen, kann man die Funktion `ksmooth` benutzen.
2. Wiederhole das Ganze mit lokalen Polynomen vom Grad 1 (mit Normalverteilungskern). Hierfür kann man die Funktion `locpoly` benutzen.
3. Schreibe den R-Code, der die Cross-Validation-Prozedur durchführt (für lokale Polynome) und bestimme hiermit eine optimale Bandbreite \hat{h}_{CV} (auf 3 Nachkommastellen genau) für den Schätzer aus Teilaufgabe 2.
Berechne auch zwei Schätzer mit lokalen Polynomen 2. und 3. Grades, die jeweils eine optimale Bandbreite haben. Hängt \hat{h}_{CV} vom Grad der lokalen Polynome ab?
Plote in einem gemeinsamen Bild den Datensatz und die 3 Schätzer mit optimaler Bandbreite.

(2+2+6 Punkte)