Prof. Dr. U. Stadtmüller
M. Weyhmüller

Sommer 2012
07.05.2012
Blatt 03

Übungen zur Funktionalen Datenanalyse

(Zu bearbeiten bis Montag, den 14.05.2012)

1. Betrachte erneut die Datei `data02.dat` von der Homepage. Sie enthält (zufällige) Messzeitpunkte x_i in der ersten, und zu diesen Zeitpunkten gemessene Werte y_i in der zweiten Spalte $(1 \leq i \leq 200)$. Man kann davon ausgehen, dass die Messwerte zufällige Messfehler ε_i enthalten, die unabhängig und identisch verteilt sind, mit $\mathrm{E}(\varepsilon_i) = 0$.

   Passe zunächst ein lineares Modell an die Daten an. Verwende dann lokale Polynome vom Grad $p = 1, 2, 3, 15$ mit einem Normalverteilungskern und Bandbreite $h = 0.01$ um die Regressionsfunktion zu schätzen. Plotte die Punkte, sowie die geschätzen Funktionen in ein gemeinsames Schaubild. Plotte 19 weitere Schaubilder, bei denen die Bandbreite jeweils um 20% wächst.

   Hinweis: Evtl. ist die Funktion `locpoly()` aus dem Paket `KernSmooth` hilfreich.

   (6 Punkte)

2. Zeige, dass bei der Regression mit lokalen Polynomen mit $p = 1$ der Schätzer $\hat{\beta}(x; h) := (A_x^\top W_x A_x)^{-1} A_x^\top W_x Y$ durch

$$\hat{\beta}_0(x; h) = \frac{1}{nh} \sum_{i=1}^{n} \frac{\big(s_{2,n}(x; h) - s_{1,n}(x; h)(x - X_i)\big) K\big(\frac{x - X_i}{h}\big) Y_i}{s_{2,n}(x; h) s_{0,n}(x; h) - s_{1,n}^2(x; h)}$$

$$\hat{\beta}_1(x; h) = \frac{1}{nh} \sum_{i=1}^{n} \frac{\big(s_{0,n}(x; h)(x - X_i) - s_{1,n}(x; h)\big) K\big(\frac{x - X_i}{h}\big) Y_i}{s_{2,n}(x; h) s_{0,n}(x; h) - s_{1,n}^2(x; h)}$$

   mit $\hat{\beta}(x; h) = \big(\hat{\beta}_0(x; h), \hat{\beta}_1(x; h)\big)^\top$ gegeben ist.

   Dabei seien die Matrizen A_x und W_x, die Vektoren X und Y, sowie die Funktionen $s_{k,n}(x; h)$ für $k = 0, 1, 2$ wie in der Vorlesung definiert.

   (6 Punkte)