

Maschinelles Lernen

11. Übung

Prof. Tobias Scheffer
Dr. Niels Landwehr
Christoph Sawade
Paul Prasse

WS09/10

Ausgabe am: 26.01.10
Besprechung am: 02.02.10

Aufgabe 1 (1/3 Punkt):

Sei $f(t)$ eine polynomielle Funktion

$$f(t) = \sum_{k=0}^p \alpha_k t^k$$

des Grades p . Zeigen sie, dass durch die Anwendung des Differenzenfilters ($g(t) = f(t) - f(t-1)$) der Polynomgrad von $g(t)$ kleiner gleich $p-1$ ist. Vielleicht hilft ihnen die Binomische Formel.

Aufgabe 2 (2/3 Punkt):

Sie sind Besitzer einer Strandbar am Wannsee. Sie befürchten, dass in den nächsten Jahren der Wasserspiegel des Sees so stark ansteigen wird, dass ihre Bar überschwemmt wird. Sie haben in den letzten Jahren an immer der gleichen Stelle die folgenden Wasserstände gemessen:

t	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009
x_t	1,26	1,32	1,61	1,42	1,59	1,63	1,93	1,82	1,97	1,99

Sie nehmen an, dass die stationäre Komponente der Wasserstände einem AR_1 -Prozess folgt und möchten zuerst die Parameter schätzen und dann eine Prognose für die kommenden Jahre abgeben, um zu entscheiden ob Sie den Cocktail-Stand ein Stück die Böschung hinaufschieben.

- Prüfen Sie ob die Wasserstände Realisierungen eines stationären Prozesses sind. Berechnen sie dazu die Mittelwert-, Varianz- und Autokovarianzfunktion.
- Bestimmen Sie die Zeitreihe $\{y_t = x_t - \mu_X(t)\}$, wobei $\mu_X(t)$ eine aus den Daten geschätzte lineare Trendgerade ist.
- Bestimmen Sie die Zeitreihe $\{\Delta x_t\}$, die sich nach einmaliger Anwendung des Differenzenfilters ergibt und bereinigen sie die Zeitreihe bzgl. ihres Mittelwertes, d.h. $\{z_t = \Delta x_t - \mu_{\Delta x}\}$.
- Eine Zeitreihe $\{x_t \in \mathbb{R}, t = 1, \dots, n\}$ mit Mittelwert 0 sei die Realisierung eines AR_1 -Prozesses. Zeigen Sie, dass (bei Minimierung des mittleren quadratischen Fehlers) für den Schätzer des Prozess-Parameters β_1 gilt

$$\beta_1 = \frac{\sum_{t=2}^n x_{t-1} x_t}{\sum_{t=2}^n x_{t-1}^2}.$$

- Schätzen Sie den AR_1 -Parameter der Wasserstandszeitreihe für die trendbereinigte (Teilaufgabe b) und die gefilterte Sequenz (Teilaufgabe c).
- Prognostizieren Sie die Wasserstände für 2010 und 2011 unter Verwendung der beiden Modelle. Vergessen Sie am Ende nicht die Datenbereinigung (Trendbereinigung bzw. Filterung & Mittelwertbereinigung) rückgängig zu machen.