

Diplomarbeit

Asymptotik eines nicht-parametrischen Kernschätzers für  
zeitvariable autoregressive Prozesse

von

Carsten Jentsch



Carl-Friedrich-Gauß-Fakultät für Mathematik und Informatik  
der Technischen Universität Carolo-Wilhelmina  
zu Braunschweig, im Oktober 2006

Betreut durch:  
Prof. Dr. J.-P. Kreiß  
Institut für Mathematische Stochastik

## Eidesstattliche Erklärung

Hiermit bestätige ich an Eides statt, dass ich die vorliegende Arbeit selbständig und nur unter Verwendung der angegebenen Hilfsmittel verfasst habe.

Braunschweig, 30. Oktober 2006

.....  
Unterschrift

# Inhaltsverzeichnis

<b>1</b>	<b>Einleitung</b>	<b>4</b>
1.1	Anmerkung . . . . .	4
1.2	Zusammenfassung . . . . .	4
1.3	Notation . . . . .	6
<b>2</b>	<b>Mathematische Grundlagen</b>	<b>7</b>
2.1	Zeitreihen . . . . .	7
2.2	Konvergenzarten . . . . .	11
<b>3</b>	<b>Das zeitabhängige <math>AR</math>-Modell (<math>tvAR</math>-Modell)</b>	<b>17</b>
3.1	Das stationäre $AR$ -Modell . . . . .	18
3.2	Motivation des $tvAR$ -Modells . . . . .	19
3.3	Simulation von $tvAR$ -Prozessen . . . . .	24
3.4	Eigenschaften von $tvAR$ -Prozessen . . . . .	28
<b>4</b>	<b>Schätztheorie</b>	<b>34</b>
4.1	Parametrische Kurvenschätzung . . . . .	34
4.2	Nicht-parametrische Kurvenschätzung . . . . .	37
4.3	Kleinste-Quadrate-Kernschätzer im $tvAR$ -Modell . . . . .	42
4.4	Darstellung des Kernschätzers als Funktion empirischer Momente . . . . .	46
4.5	Zerlegung des Schätzfehlers . . . . .	48
<b>5</b>	<b>Die Beveridge-Nelson-Zerlegung</b>	<b>52</b>
5.1	$BN/tvBN$ -Zerlegung . . . . .	52
5.2	$BN/tvBN$ -Zerlegung zweiter Ordnung . . . . .	56
<b>6</b>	<b>Asymptotische Betrachtungen</b>	<b>60</b>
6.1	Asymptotik der empirischen Momente . . . . .	61
6.2	Asymptotik des Bias-Terms und des Restterms . . . . .	73
6.3	Asymptotische Normalität des Schätzfehlers . . . . .	75

---

<b>7</b>	<b>Anhang</b>	<b>87</b>
7.1	Ungleichungen und Doppelsummen . . . . .	87
7.2	Gleichmäßige Beschränktheit . . . . .	89

# Abbildungsverzeichnis

3.1	$AR(1)$ -Reihe mit $n = 500$ und $\alpha = 0,9$ . . . . .	18
3.2	$AR(1)$ -Reihe mit $n = 500$ und $\alpha = -0,3$ . . . . .	19
3.3	Relative Ein-Tages-Veränderungen des DAX für den Zeitraum 03.01.2003 bis 30.12.2004 . . . . .	20
3.4	$tvAR(4)$ -Reihe mit Anfangswerten gleich null . . . . .	25
3.5	$tvAR(4)$ -Reihe mit erzeugten Anfangswerten . . . . .	26
3.6	$tvAR(1)$ -Reihe mit $n = 500$ und $\alpha(u) = 0,9 \sin(2\pi u)$ . . . . .	27
3.7	$tvAR(1)$ -Reihe mit $n = 500$ und $\alpha(u) = 1,8u - 0,9$ . . . . .	28
4.1	Sinusfunktion $\alpha(u) = 0,9 \sin(2\pi u)$ . . . . .	36
4.2	Epanechnikov-Kern $K(x) = \frac{3}{4}(1 - x^2)I_{[-1,1]}(x)$ . . . . .	39
4.3	Biweight-Kern $K(x) = \frac{15}{16}(1 - x^2)^2I_{[-1,1]}(x)$ . . . . .	40
4.4	Biweight-Kern $K_h(x)$ mit Bandweite $h = 0,5$ . . . . .	41
4.5	Tatsächliche Koeffizientenfunktion $\alpha(u) = 0,9 \sin(2\pi u)$ und deren Kur- venschätzungen für $n = 100, 500, 1000$ . . . . .	45
4.6	Tatsächliche Koeffizientenfunktion $\alpha(u) = 1,8u - 0,9$ und deren Kur- venschätzungen für $n = 100, 500, 1000$ . . . . .	46

# Kapitel 1

## Einleitung

### 1.1 Anmerkung

Der Kern dieser Diplomarbeit stammt aus dem „discussion paper“ *Nonparametric kernel estimation of evolutionary autoregressive processes* des Sonderforschungsbereichs 373 der Humboldt-Universität zu Berlin von Woocheol Kim aus dem Jahre 2001 (siehe [14]), welches sich leider als unvollständig und fehlerhaft erwies.

### 1.2 Zusammenfassung

In der klassischen Zeitreihenanalyse beschäftigt man sich hauptsächlich mit sogenannten stationären Zeitreihen. Stationarität eines Systems oder einer Zeitreihe bedeutet hier, dass sich gewisse stochastische Eigenschaften nicht mit der Zeit ändern (siehe Definition 2.4). Diese Annahme einer unveränderlichen Struktur hat zur Folge, dass mehr verfügbare Daten eines Systems auch gleichzeitig mehr Information über dessen stochastische Eigenschaften bedeutet.

Zum einen stellt die Annahme der Stationarität eine wesentliche Einschränkung der betrachteten Klasse von Zeitreihen dar, und zum anderen kann sie bestenfalls als eine mathematische Idealisierung angesehen werden, welche oftmals zu einfach ist, um etwa die komplizierte dynamische Struktur eines ökonomischen Prozesses zu erfassen (siehe Abschnitt 3.2). Besonders bei der Untersuchung von Datenreihen aus der Wirtschaft mit langer Vergangenheit werden die Zweifel groß, ob ein stationäres Modell die Wirklichkeit angemessen widerspiegeln kann. Dort lassen sich nämlich häufig strukturelle Schwankungen über die Zeit beobachten, welche nicht vernünftig mit einem stationären Ansatz untersucht werden können.

In dieser Arbeit werden wir uns mit einer allgemeineren Klasse bestehend aus zeitabhä-

gen Prozessen beschäftigen. Wir erlauben dabei den Zeitreihen, dass ihre statistischen Merkmale von der Zeit abhängen und sie somit nicht mehr die Annahme der Stationarität erfüllen. Um sich jedoch nicht zu sehr von der weit entwickelten Theorie stationärer Prozesse zu entfernen, verallgemeinert man das klassische autoregressive Modell

$$X_t = \sum_{k=1}^p \alpha_k X_{t-k} + \epsilon_t, \quad \forall t \in \mathbb{Z}, \quad (1.1)$$

indem man den  $AR$ -Koeffizienten  $\alpha_k$  eine Abhängigkeit von der Zeit  $t$  gestattet. Setzt man in einem solchen zeitvariablen Modell den Prozess in die Zukunft fort, indem man die Anzahl der Beobachtungen  $n$  größer werden lässt, so führt dieses nicht notwendigerweise zu einem Gewinn an Information über die statistische Struktur des Prozesses in der Gegenwart. Deshalb skaliert man ihn auf das Intervall  $[0, 1]$  um, damit eine sinnvolle asymptotische Betrachtung überhaupt erst möglich wird (siehe Abschnitt 3.2). Eine zeitabhängige  $AR$ -Reihe der Ordnung  $p$ , auch als  $tvAR(p)$ -Zeitreihe bezeichnet, erfüllt dann per Definition die Modellgleichung

$$X_{t,n} = \sum_{k=1}^p \alpha_k\left(\frac{t}{n}\right) X_{t-k,n} + \epsilon_t, \quad t = 1, \dots, n \quad (1.2)$$

mit Koeffizientenfunktionen  $\alpha_k : [0, 1] \rightarrow \mathbb{R}$  und einem unabhängig und identisch verteilten weißen Rauschen  $\epsilon_t \sim (0, \sigma_\epsilon^2)$  (siehe Dahlhaus [6]).

Die Klasse der möglichen Koeffizientenfunktionen  $\alpha_k(\cdot)$  schränkt man noch ein, indem man von ihnen einen *hinreichend glatten* Verlauf fordert, welcher lokal ein stationäres Verhalten zur Folge hat. Macht man nämlich keine weiteren Einschränkungen bzgl. dem Grad der Nicht-Stationarität und lässt damit beliebige reellwertige Funktionen auf  $[0, 1]$  als Koeffizienten zu, so ist es in diesem Fall nicht möglich eine bedeutsame Theorie zu entwickeln (siehe Abschnitt 4.2). Dahlhaus konnte in [3] den Begriff der lokalen Stationarität einer Zeitreihe definieren, welche es möglich macht, diese lokal durch stationäre Prozesse zu approximieren (siehe Satz 3.2).

Das Hauptanliegen dieser Arbeit ist es nun, in dem  $tvAR(p)$ -Modell (1.2) einen nicht-parametrischen Kernschätzer  $\widehat{\underline{\alpha}}(u) = (\widehat{\alpha}_1(u), \dots, \widehat{\alpha}_p(u))^T$  für die Koeffizientenfunktionen  $\alpha_k(u)$  anzugeben und die asymptotische Verteilung des Schätzfehlers  $\widehat{\underline{\alpha}}(u) - \underline{\alpha}(u)$  herzuleiten.

Nachdem wir in Kapitel 2 einige wichtige Begriffe und Resultate angegeben haben, wenden wir uns in Kapitel 3 der Modellbildung zu. Dabei erläutern wir kurz die Grenzen eines stationären autoregressiven Ansatzes und zeigen die Vorzüge eines  $tvAR(p)$ -Modells gemäß (1.2) auf. In den Abschnitten 3.3 und 3.4 widmen wir uns der Simulation solcher Prozesse bzw. können einige erste Eigenschaften beweisen. Anschließend befassen

wir uns in Kapitel 4 mit der relevanten Schätztheorie, indem wir zunächst knapp auf die allgemeine Idee der nicht-parametrischen Regression eingehen und dann einen nicht-parametrischen Kernschätzer für die Koeffizientenfunktionen  $\underline{\alpha}(u)$  herleiten. An die Kernfunktion  $K$  gestellte Annahmen notieren wir in Abschnitt 4.2 und alternative Darstellungen des Schätzers bzw. des Schätzfehlers in Abhängigkeit empirischer Momente  $\mathbf{S}_{n,l}(u)$  (siehe Lemma 4.2) beweisen wir in den Abschnitten 4.4 und 4.5. In Kapitel 5 führen wir die sogenannte Beveridge-Nelson-Zerlegung (siehe [1]) für lineare Prozesse der Form

$$X_t = \sum_{j=0}^{\infty} c_j \epsilon_{t-j} \quad (1.3)$$

(erster Ordnung) bzw. für Ausdrücke der Gestalt  $X_t X_{t+h}$ ,  $h \in \mathbb{Z}$  (zweiter Ordnung) ein und verallgemeinern diese für unsere Zwecke auf den Fall zeitabhängiger Koeffizienten, welche wir in Kapitel 6 benutzen, um die Asymptotik der empirischen Momente  $\mathbf{S}_{n,l}(u)$  herzuleiten. Abschließend wenden wir einen Zentralen Grenzwertsatz für Martingaldifferenzschemata von Liptser und Shirjaev aus [17] an, um die asymptotische Normalität des Schätzfehlers  $\widehat{\underline{\alpha}}(u) - \underline{\alpha}(u)$  nachzuweisen und fassen in Theorem 6.11 die in Kapitel 6 erzielten Resultate zusammen.

### 1.3 Notation

Ab Kapitel 3 kennzeichnen wir Vektoren grundsätzlich mit einem Unterstrich und Matrizen, indem wir sie fett schreiben. Beispielsweise bezeichnet  $\widehat{\underline{\alpha}}(u) = (\widehat{\alpha}_1(u), \dots, \widehat{\alpha}_p(u))^T$  den  $p$ -dimensionalen Schätzvektor für die Koeffizientenfunktionen  $\alpha_k(u)$  aus Abschnitt 4.3 und  $\mathbf{S}_{n,l}(u)$  ist die  $p \times p$ -dimensionale Matrix empirischer Momente aus Lemma 4.2.

Alle Vektoren in dieser Arbeit sind als Spaltenvektoren zu verstehen. Immer wenn Spaltenvektoren untereinander in einen Vektor schreiben, meinen wir wieder einen Vektor (siehe etwa die Definition von  $\underline{\tau}_n(u)$  in Lemma 4.3) und wenn wir Zeilenvektoren untereinander schreiben, ist eine Matrix gemeint (siehe etwa die Definition in (4.25)).

Sind  $A$  und  $B$  zwei  $n \times n$ -Matrizen, so steht die Schreibweise  $(A, B)$  für die  $n \times 2n$ -Matrix, deren linke Hälfte gleich  $A$  und deren rechte gleich  $B$  ist. Multiplikation mit einer dritten  $n \times n$ -Matrix  $C$  bedeutet dann  $(A, B)C = (AC, BC)$  (siehe etwa (4.28)).

# Kapitel 2

## Mathematische Grundlagen

In diesem Kapitel sind alle für diese Arbeit grundlegenden Begriffe erklärt sowie einige benötigte Sätze notiert. Diese Zusammenfassung mathematischer Grundlagen stützt sich dabei im Wesentlichen auf die Bücher *Time Series: Theory and Methods* von Brockwell und Davis [2] und *Einführung in die Zeitreihenanalyse* von Kreiß und Neuhaus [15].

In Abschnitt 2.1 wird zunächst der wichtige Begriff des stochastischen Prozesses eingeführt und definiert, was unter einer Zeitreihe verstanden werden soll, bevor dann Begriffe und wichtige Resultate der Zeitreihenanalyse vorgestellt werden. Auf die verschiedenen Konvergenzarten der Stochastik und wichtige Sätze über diese Konvergenzen wird in Abschnitt 2.2 eingegangen.

### 2.1 Zeitreihen

#### **Definition 2.1 (Stochastischer Prozess).**

*Ein stochastischer Prozess ist eine Familie von Zufallsvariablen  $\{X_t\}_{t \in T}$  auf einem Wahrscheinlichkeitsraum  $(\Omega, \mathcal{F}, P)$ .*

#### **Definition 2.2 (Realisierung eines stochastischen Prozesses).**

*Die Funktionen  $\{X_\cdot(\omega)\}_{\omega \in \Omega}$  auf  $T$  werden als Realisierungen oder Pfade des Prozesses  $\{X_t\}_{t \in T}$  bezeichnet.*

#### **Definition 2.3 (Zeitreihe).**

*Der Begriff Zeitreihe wird gleichermaßen für die zugrundeliegenden Beobachtungen sowie für den stochastischen Prozess  $\{X_t\}_{t \in T}$ , wovon die Beobachtungen eine Realisierung darstellen, benutzt.*

In der Definition der Zeitreihe wird zwar die Indexmenge nicht weiter eingeschränkt, jedoch werden wir nur Zeitreihen mit diskreten, also höchstens abzählbaren Indexmengen

$T$  untersuchen. Insbesondere interessiert man sich häufig für Lösungen einer Modellgleichung (siehe unten) auf der Menge der ganzen Zahlen, also gerade für den Fall, dass  $T = \mathbb{Z}$  gilt.

Nun kommen wir zu dem in der Zeitreihenanalyse überaus wichtigen Begriff der Stationarität. Man bezeichnet ein System (Zeitreihe) stationär, falls sich in ihm die stochastischen Eigenschaften wie Erwartungswert und Varianz sowie die Abhängigkeitsstruktur nicht mit der Zeit ändern.

**Definition 2.4 (Stationarität).**

Die Zeitreihe  $\{X_t\}_{t \in \mathbb{Z}}$  heißt (schwach) stationär, falls sie die Eigenschaften

$$(i) \quad E[|X_t|^2] < \infty, \quad \forall t \in \mathbb{Z}, \quad (2.1)$$

$$(ii) \quad E[X_t] = \mu, \quad \forall t \in \mathbb{Z}, \quad (2.2)$$

$$(iii) \quad Cov[X_r, X_s] = Cov[X_{r+t}, X_{s+t}], \quad \forall r, s, t \in \mathbb{Z} \quad (2.3)$$

erfüllt.

Das wohl einfachste Beispiel einer stationären Zeitreihe ist das weiße Rauschen. Es ist zugleich auch eines der wichtigsten, denn in vielen Zeitreihenmodellen wird ein weißes Rauschen benutzt, um dort einen zufälligen um null streuenden Fehler auszudrücken. Mit seiner Hilfe führen wir danach die wichtige Klasse der *ARMA*-Modelle ein.

**Definition 2.5 (Weißes Rauschen).**

Ein Prozess  $\{\epsilon_t\}_{t \in \mathbb{Z}}$  heißt weißes Rauschen oder *White-Noise-Prozess* mit Varianz  $\sigma_\epsilon^2$ , in Zeichen  $\epsilon_t \sim (0, \sigma_\epsilon^2)$ , falls die  $\{\epsilon_t\}_{t \in \mathbb{Z}}$  stochastisch unabhängig sind und

$$(i) \quad E[\epsilon_t] = 0, \quad \forall t \in \mathbb{Z}, \quad (2.4)$$

$$(ii) \quad Var[\epsilon_t] = \sigma_\epsilon^2, \quad \forall t \in \mathbb{Z} \quad (2.5)$$

gelten.

**Bemerkung 2.6 (Unabhängigkeit des weißen Rauschens).**

In der Literatur wird häufig nur die Unkorreliertheit der Zufallsvariablen  $\{\epsilon_t\}_{t \in \mathbb{Z}}$  gefordert. Die weiteren Ausführungen in dieser Arbeit erfordern jedoch die stärkere Annahme der Unabhängigkeit, so dass wir dieses bereits in der Definition des weißen Rauschens haben einfließen lassen.  $\square$

**Definition 2.7 (ARMA( $p, q$ )-Prozess).**

Der Prozess  $\{X_t\}_{t \in \mathbb{Z}}$  heißt *ARMA( $p, q$ )-Prozess (autoregressive moving average)*, falls  $\{X_t\}_{t \in \mathbb{Z}}$  stationär ist und

$$X_t - \sum_{k=1}^p \alpha_k X_{t-k} = \epsilon_t + \sum_{k=1}^q \beta_k \epsilon_{t-k}, \quad \forall t \in \mathbb{Z} \quad (2.6)$$

gilt, wobei  $\epsilon_t \sim (0, \sigma_\epsilon^2)$  ein weißes Rauschen bezeichnet.

Die Definitionen 2.5 und 2.7 sind zwar nicht auf den reellen Fall beschränkt, jedoch werden wir uns im Folgenden nur mit reellem weißem Rauschen und mit  $ARMA(p, q)$ -Reihen mit reellen Koeffizienten  $\alpha_j$  und  $\beta_j$  befassen.

**Bemerkung 2.8 (Eine alternative Darstellungen der  $ARMA$ -Gleichung).**

Definiert man für einen Prozess  $\{X_t\}_{t \in \mathbb{Z}}$  durch

$$L^j X_{t-j} := X_{t-j}, \quad j \in \mathbb{Z}, \quad (2.7)$$

den Lag-Operator  $L$ , so lässt sich mit seiner Hilfe und den sogenannten  $z$ -Transformationen

$$A(z) := 1 - \sum_{k=1}^p \alpha_k z^k, \quad z \in \mathbb{C} \quad (2.8)$$

und

$$B(z) := 1 + \sum_{k=1}^q \beta_k z^k, \quad z \in \mathbb{C} \quad (2.9)$$

die Gleichung (2.6) kompakt schreiben als

$$A(L)X_t = B(L)\epsilon_t, \quad \forall t \in \mathbb{Z}. \quad (2.10)$$

□

Da wir uns in dieser Arbeit nur mit einem Spezialfall aus der Klasse der  $ARMA$ -Modelle, dem  $AR$ -Modell, befassen werden, wollen wir dieses hier in einer eigenen Definition hervorheben. Der rechte Teil der  $ARMA$ -Gleichung (2.6) vereinfacht sich dabei zu  $\epsilon_t$ , d.h. der  $MA$ -Anteil verschwindet sonst.

**Definition 2.9 ( $AR(p)$ -Prozess).**

Der Prozess  $\{X_t\}_{t \in \mathbb{Z}}$  heißt  $AR(p)$ -Prozess (*autoregressive*), falls  $\{X_t\}_{t \in \mathbb{Z}}$  stationär ist und

$$X_t = \sum_{k=1}^p \alpha_k X_{t-k} + \epsilon_t, \quad \forall t \in \mathbb{Z} \quad (2.11)$$

gilt, also falls in Definition 2.7 gerade  $q = 0$  ist.

**Satz 2.10 (Existenz von Lösungen der  $ARMA$ -Gleichung).**

Wenn  $A(z) \neq 0$  für alle  $z \in \mathbb{C}$  mit  $|z| = 1$  gilt, dann gibt es eine  $ARMA(p, q)$ -Reihe  $\{X_t\}_{t \in \mathbb{Z}}$ , welche (2.6) bzw. (2.10) erfüllt. Diese Lösung ist eindeutig bestimmt.

Der obige Satz stellt unter der Annahme, dass das Polynom  $A$  keine Nullstellen auf dem Einheitskreis hat, die Existenz einer stationären Lösung im  $ARMA$ -Modell sicher. Besonders interessiert man sich für stationäre Lösungen  $X_t$ , welche nur noch vom weißen Rauschen abhängen und dabei speziell für solche Darstellungen von  $X_t$ , die nur von der Vergangenheit, also von  $\epsilon_s$ ,  $s \leq t$  abhängen. Solche Lösungen bzw. Prozesse werden kausal bezeichnet. Bevor wir die Frage nach der Existenz einer solchen Lösung im  $ARMA$ - bzw.  $AR$ -Modell beantworten werden, wollen wir noch den Begriff der Kausalität präzisieren.

**Definition 2.11 (Kausalität).**

Ein  $ARMA(p, q)$ -Prozess, definiert durch  $A(L)X_t = B(L)\epsilon_t$ , heißt kausal, falls eine Folge von Konstanten  $\{\psi_k\}_{k=0}^{\infty}$  existiert mit

$$X_t = \sum_{k=0}^{\infty} \psi_k \epsilon_{t-k}, \quad t \in \mathbb{Z}, \quad \sum_{k=0}^{\infty} |\psi_k| < \infty, \quad (2.12)$$

wobei man  $\{X_t\}_{t \in \mathbb{Z}}$  dann auch als  $MA(\infty)$ -Reihe bezeichnet.

**Satz 2.12 (Kausalität des  $ARMA(p, q)$ -Prozesses).**

Sei  $\{X_t\}_{t \in \mathbb{Z}}$  ein  $ARMA(p, q)$ -Prozess, dessen  $z$ -Transformationen  $A$  und  $B$  keine gemeinsamen Nullstellen besitzen. Dann ist  $\{X_t\}_{t \in \mathbb{Z}}$  genau dann kausal, wenn  $A(z) \neq 0$  für alle  $z \in \mathbb{C}$  mit  $|z| \leq 1$  gilt.

**Korollar 2.13.**

Ein  $AR(p)$ -Prozess ist genau dann kausal, wenn  $A(z) \neq 0$  für alle  $z \in \mathbb{C}$  mit  $|z| \leq 1$  gilt.

Abschließen werden wir diesen Abschnitt mit einigen Resultaten aus der Spektraltheorie. Angewandt auf die oben eingeführten Modelle ermöglicht sie es, die  $ARMA$ -Gleichung (2.6) bzw. (2.10) mit Hilfe von stochastischen Integralen auszudrücken oder auch den Prozess selbst als Integral bzgl. eines stochastischen Prozesses darzustellen. Dazu werden wir zunächst definieren, was wir unter einem Prozess mit orthogonalen Zuwächsen verstehen, bevor wir zu einer weiteren Darstellung von (2.6) über stochastische Integrale bzgl. solcher Prozesse (mit orthogonalen Zuwächsen) gelangen.

**Definition 2.14 (Prozess mit orthogonalen Zuwächsen).**

Sei  $\langle X, Y \rangle := E[X\bar{Y}]$  definiert. Ein Prozess mit orthogonalen Zuwächsen ist ein komplexwertiger stochastischer Prozess  $\{Z(\lambda) : \lambda \in (-\pi, \pi]\}$ , so dass für alle  $\lambda \in (-\pi, \pi]$

$$\langle Z(\lambda), Z(\lambda) \rangle < \infty, \quad (2.13)$$

$$\langle Z(\lambda), 1 \rangle = 0 \quad (2.14)$$

gilt sowie für alle  $\lambda_1 < \lambda_2, \lambda_3 < \lambda_4 \in (-\pi, \pi]$  mit  $(\lambda_1, \lambda_2] \cap (\lambda_3, \lambda_4] = \emptyset$

$$\langle Z(\lambda_4) - Z(\lambda_3), Z(\lambda_2) - Z(\lambda_1) \rangle = 0 \quad (2.15)$$

erfüllt ist.

**Bemerkung 2.15 (Eine weitere Darstellung der ARMA-Gleichung).**

Definiert man mit Hilfe der  $z$ -Transformationen  $A$  und  $B$  aus Bemerkung 2.8 die Funktionen  $\varphi(\omega) := A(e^{-i\omega})$  und  $\psi(\omega) := B(e^{-i\omega})$ ,  $\omega \in (-\pi, \pi]$ , so lässt sich (2.6) schreiben als

$$\int_{(-\pi, \pi]} e^{it\lambda} \varphi(\lambda) dZ_X(\lambda) = \int_{(-\pi, \pi]} e^{it\lambda} \psi(\lambda) dZ_\epsilon(\lambda), \quad \forall t \in \mathbb{Z}, \quad (2.16)$$

wobei  $\{Z_X(\lambda) : \lambda \in (-\pi, \pi]\}$  der zu  $\{X_t\}_{t \in \mathbb{Z}}$  bzw.  $\{Z_\epsilon(\lambda) : \lambda \in (-\pi, \pi]\}$  der zu  $\{\epsilon_t\}_{t \in \mathbb{Z}}$  gehörige Prozess mit orthogonalen Zuwächsen ist.  $\square$

**Satz 2.16 (Spektraldarstellung des ARMA( $p, q$ )-Prozesses).**

Sei  $\{X_t\}_{t \in \mathbb{Z}}$  ein ARMA( $p, q$ )-Prozess, definiert durch

$$A(L)X_t = B(L)\epsilon_t, \quad t \in \mathbb{Z}, \quad \epsilon_t \sim (0, \sigma_\epsilon^2) \quad (2.17)$$

mit  $A$  und  $B$  wie in Bemerkung 2.8. Falls  $A$  und  $B$  keine gemeinsamen Nullstellen haben und  $A(z) \neq 0$  für alle  $z \in \mathbb{C}$  mit  $|z| = 1$  gilt, dann besitzt  $\{X_t\}_{t \in \mathbb{Z}}$  die Spektraldichte

$$f(\lambda) = \frac{\sigma_\epsilon^2 |B(e^{-i\lambda})|^2}{2\pi |A(e^{-i\lambda})|^2} = \frac{\sigma_\epsilon^2 |\psi(\lambda)|^2}{2\pi |\varphi(\lambda)|^2}, \quad \lambda \in (-\pi, \pi] \quad (2.18)$$

bzw. die Transferfunktion (oder den Filter)

$$A(\lambda) = \frac{\sigma_\epsilon B(e^{-i\lambda})}{\sqrt{2\pi} A(e^{-i\lambda})} = \frac{\sigma_\epsilon \psi(\lambda)}{\sqrt{2\pi} \varphi(\lambda)}, \quad \lambda \in (-\pi, \pi]. \quad (2.19)$$

Der Prozess  $\{X_t\}_{t \in \mathbb{Z}}$  besitzt dann die Spektraldarstellung

$$X_t = \int_{(-\pi, \pi]} e^{it\lambda} \frac{\sigma_\epsilon \psi(\lambda)}{\sqrt{2\pi} \varphi(\lambda)} dZ_\epsilon(\lambda), \quad t \in \mathbb{Z}. \quad (2.20)$$

## 2.2 Konvergenzarten

In der Stochastik muss man sorgsam zwischen den verschiedenen Konvergenzarten unterscheiden. Wir werden die fast sichere Konvergenz, die Konvergenz nach Wahrscheinlichkeit, die schwache Konvergenz und die Konvergenz in Verteilung definieren sowie deren in dieser Arbeit benötigten Eigenschaften und die Zusammenhänge der Konvergenzen untereinander ansprechen. Besonders häufig werden wir in den folgenden Kapiteln die Aussagen der Lemmata 2.26 und 2.27 über hinreichende Bedingungen für Konvergenz und Beschränktheit nach Wahrscheinlichkeit benutzen, weshalb wir diese auch kurz beweisen werden.

Abschließend zitieren wir einen Zentralen Grenzwertsatz für Martingaldifferenzschemata aus einer Arbeit von Liptser und Shiryaev (siehe [17]), welcher später in Kapitel 6 das Fundament der asymptotischen Betrachtungen sein wird.

**Definition 2.17 (*P*-fast sichere Konvergenz).**

Es seien  $\{X_n\}_{n \in \mathbb{N}_0}$  Zufallsvariablen auf einem Wahrscheinlichkeitsraum  $(\Omega, \mathcal{A}, P)$ . Dann konvergiert  $\{X_n\}_{n \in \mathbb{N}}$  (*P*-)fast sicher gegen  $X_0$ , falls

$$P(\{\omega \in \Omega \mid X_n(\omega) \xrightarrow[n \rightarrow \infty]{} X_0(\omega)\}) = 1 \quad (2.21)$$

gilt. Man schreibt dann

$$X_n \xrightarrow{f.s.} X_0. \quad (2.22)$$

**Definition 2.18 (Konvergenz nach Wahrscheinlichkeit).**

Es seien  $\{X_n\}_{n \in \mathbb{N}_0}$  Zufallsvariablen auf einem Wahrscheinlichkeitsraum  $(\Omega, \mathcal{A}, P)$ . Dann konvergiert  $\{X_n\}_{n \in \mathbb{N}}$  stochastisch (nach Wahrscheinlichkeit) gegen  $X_0$ , falls

$$\forall \epsilon > 0 : P(|X_n - X_0| \geq \epsilon) \xrightarrow[n \rightarrow \infty]{} 0 \quad (2.23)$$

gilt. Man schreibt dann

$$X_n \xrightarrow{P} X_0 \quad \text{oder} \quad X_n - X_0 = o_P(1). \quad (2.24)$$

**Definition 2.19 (schwache Konvergenz).**

Es sei  $\{P_n\}_{n \in \mathbb{N}_0}$  eine Folge von Verteilungen auf  $(\mathbb{R}^k, \mathbb{B}^k)$ , wobei  $\mathbb{B}^k$  die  $k$ -dimensionale Borel- $\sigma$ -Algebra bezeichne. Dann konvergiert  $\{P_n\}_{n \in \mathbb{N}}$  schwach gegen  $P_0$ , falls

$$\int f dP_n \xrightarrow[n \rightarrow \infty]{} \int f dP_0, \quad \forall f \in C^b(\mathbb{R}^k), \quad (2.25)$$

wobei

$$C^b(\mathbb{R}^k) := \{f : \mathbb{R}^k \rightarrow \mathbb{R}, f \text{ stetig und beschränkt}\}. \quad (2.26)$$

Man schreibt dann

$$P_n \Rightarrow P_0. \quad (2.27)$$

**Definition 2.20 (Konvergenz in Verteilung).**

$\{X_n\}_{n \in \mathbb{N}_0}$  seien Zufallsvariablen auf den Wahrscheinlichkeitsräumen  $(\Omega_n, \mathcal{A}_n, P_n)_{n \in \mathbb{N}_0}$  und sei  $\{P_n\}_{n \in \mathbb{N}_0}$  eine Folge von Verteilungen auf  $(\mathbb{R}^k, \mathbb{B}^k)$ . Dann konvergiert  $\{X_n\}_{n \in \mathbb{N}}$  in Verteilung gegen  $X_0$ , falls  $\{P_n^{X_n}\}_{n \in \mathbb{N}}$  schwach gegen  $P_0^{X_0}$  konvergiert, also wenn

$$\int f dP_n^{X_n} \xrightarrow{n \rightarrow \infty} \int f dP_0^{X_0}, \quad \forall f \in C^b(\mathbb{R}^k) \quad (2.28)$$

gilt. Man schreibt dann

$$X_n \xrightarrow{D} X_0. \quad (2.29)$$

Um auszudrücken, dass  $X_n$  gegen eine  $\mathcal{N}(\mu, \sigma^2)$ -verteilte Zufallsvariable konvergiert, so schreibt man auch

$$X_n \xrightarrow{D} \mathcal{N}(\mu, \sigma^2). \quad (2.30)$$

**Bemerkung 2.21.**

Aus der P-fast sicheren Konvergenz folgt die Konvergenz nach Wahrscheinlichkeit und aus der Konvergenz nach Wahrscheinlichkeit folgt die Konvergenz in Verteilung. Die Umkehrungen gelten im Allgemeinen nicht.  $\square$

**Satz 2.22 (Slutsky).**

Seien  $\{X_n\}_{n \in \mathbb{N}_0}$  und  $\{Y_n\}_{n \in \mathbb{N}}$   $\mathbb{R}^k$ -wertige Folgen von Zufallsvariablen auf  $(\Omega, \mathcal{A}, P)$  mit  $X_n \xrightarrow{D} X_0$  und  $Y_n - a = o_P(1)$ ,  $a \in \mathbb{R}^k$ . Dann folgt

$$(i) \quad X_n + Y_n \xrightarrow{D} X_0 + a, \quad (2.31)$$

$$(ii) \quad X_n \cdot Y_n \xrightarrow{D} X_0 \cdot a. \quad (2.32)$$

**Definition 2.23 (Beschränktheit nach Wahrscheinlichkeit).**

Eine Folge von  $\mathbb{R}^k$ -wertigen Zufallsvariablen  $\{X_n\}_{n \in \mathbb{N}}$  auf einem Wahrscheinlichkeitsraum  $(\Omega, \mathcal{A}, P)$  heißt beschränkt nach Wahrscheinlichkeit (oder straff), falls

$$\forall \epsilon > 0 \exists M < \infty \forall n \in \mathbb{N} : P(\|X_n\| > M) < \epsilon \quad (2.33)$$

gilt. Man schreibt dann

$$X_n = O_P(1). \quad (2.34)$$

**Bemerkung 2.24 (Stochastische Landau-Symbole).**

Die Symbole  $o_P(\cdot)$  und  $O_P(\cdot)$  heißen stochastische Landau-Symbole. In den Definitionen 2.18 und 2.23 sind  $o_P(1)$  und  $O_P(1)$  erklärt, so dass man für eine Folge  $\{a_n\}_{n \in \mathbb{N}}$  mit Werten in  $(0, \infty)$  die Ausdrücke  $o_P(a_n)$  und  $O_P(a_n)$  durch

$$X_n = O_P(a_n) \quad :\Leftrightarrow \quad X_n = Y_n \cdot a_n \wedge Y_n = O_P(1), \quad (2.35)$$

$$X_n = o_P(a_n) \quad :\Leftrightarrow \quad X_n = Y_n \cdot a_n \wedge Y_n = o_P(1) \quad (2.36)$$

definieren kann. Handelt es sich bei den  $X_n$  nicht um Zufallsvektoren, sondern um deterministische Vektoren, so reduzieren sich die stochastischen Landau-Symbole auf die üblichen Landau-Symbole  $o(a_n)$  und  $O(a_n)$ .  $\square$

**Satz 2.25 (Wahrscheinlichkeitstheoretische Ungleichungen).**

Seien  $X$  und  $Y$  reellwertige Zufallsvariablen auf einem Wahrscheinlichkeitsraum  $(\Omega, \mathcal{A}, P)$ . Dann gilt für alle  $\epsilon > 0$

$$(i) \quad P(|X| \geq \epsilon) \leq \frac{E[|X|^r]}{\epsilon^r}, \quad r > 0, \quad (\text{allg. Markov-Ungleichung}) \quad (2.37)$$

$$(ii) \quad P(|X - E[X]| \geq \epsilon) \leq \frac{\text{Var}[X]}{\epsilon^2}, \quad (\text{Tschebyscheff-Ungleichung}) \quad (2.38)$$

$$(iii) \quad E[|X \cdot Y|] \leq \sqrt{E[X^2]} \sqrt{E[Y^2]}. \quad (\text{Cauchy-Schwarz-Ungleichung}) \quad (2.39)$$

**Lemma 2.26 (Hinreichende Bedingungen für  $o_P(1)$ ).**

Es seien  $\{X_n\}_{n \in \mathbb{N}}$  reellwertige Zufallsvariablen auf  $(\Omega, \mathcal{A}, P)$ . Gilt eine der beiden Aussagen

$$(i) \quad E[|X_n|] \xrightarrow{n \rightarrow \infty} 0, \quad (2.40)$$

$$(ii) \quad E[X_n^2] \xrightarrow{n \rightarrow \infty} 0, \quad (2.41)$$

so folgt die Konvergenz gegen null nach Wahrscheinlichkeit von  $X_n$ , d.h.  $X_n = o_P(1)$ .

*Beweis.*

Wendet man die allgemeine Markov-Ungleichung (2.37) für  $r = 1$  und  $r = 2$  auf die Definition von  $o_P(1)$  an, so erhält man

$$P(|X_n| \geq \epsilon) \leq \frac{E[|X_n|]}{\epsilon} \quad (2.42)$$

bzw.

$$P(|X_n| \geq \epsilon) \leq \frac{E[|X_n|^2]}{\epsilon^2}. \quad (2.43)$$

Nach Voraussetzung streben die jeweiligen Erwartungswerte für alle  $\epsilon > 0$  mit wachsendem  $n$  gegen null, so dass  $X_n = o_P(1)$  folgt.  $\square$

**Lemma 2.27 (Hinreichende Bedingungen für  $O_P(1)$ ).**

Es seien  $\{X_n\}_{n \in \mathbb{N}}$  reellwertige Zufallsvariablen auf  $(\Omega, \mathcal{A}, P)$ . Ist für eine endliche Konstante  $C$  eine der beiden Beschränktheitsbedingungen

$$(i) \quad \exists C < \infty \forall n \in \mathbb{N} : E[|X_n|] < C, \quad (2.44)$$

$$(ii) \quad \exists C < \infty \forall n \in \mathbb{N} : E[X_n^2] < C \quad (2.45)$$

erfüllt, so folgt die Beschränktheit nach Wahrscheinlichkeit von  $X_n$ , d.h.  $X_n = O_P(1)$ .

*Beweis.*

Analog zum Beweis von Lemma 2.26 ergibt sich die Aussage aus

$$P(|X_n| \geq M) \leq \frac{E[|X_n|]}{M} < \frac{C}{M} \quad (2.46)$$

bzw. aus

$$P(|X_n| \geq M) \leq \frac{E[X_n^2]}{M^2} < \frac{C}{M^2} \quad (2.47)$$

für hinreichend großes  $M$ . □

**Lemma 2.28 (Rechenregeln für Konvergenz nach Wahrscheinlichkeit).**

Seien  $\{X_n\}_{n \in \mathbb{N}_0}$  und  $\{Y_n\}_{n \in \mathbb{N}_0}$   $\mathbb{R}^k$ -wertige Folgen von Zufallsvariablen auf  $(\Omega, \mathcal{A}, P)$  mit  $X_n = X_0 + o_P(1)$  und  $Y_n = Y_0 + o_P(1)$ . Dann folgt

$$(i) \quad X_n + Y_n = X_0 + Y_0 + o_P(1), \quad (2.48)$$

$$(ii) \quad X_n^T Y_n = X_0^T Y_0 + o_P(1). \quad (2.49)$$

Die beiden folgenden Sätze zeigen, wie sich Konvergenz und Beschränktheit nach Wahrscheinlichkeit vereinen lassen (Slutsky) sowie die Verträglichkeit der Konvergenz in Verteilung mit stetigen Abbildungen bzw. mit Abbildungen, deren Menge der Unstetigkeitsstellen die Wahrscheinlichkeit null besitzen.

**Satz 2.29 (Slutsky).**

Seien  $\{X_n\}_{n \in \mathbb{N}}$  und  $\{Y_n\}_{n \in \mathbb{N}}$   $\mathbb{R}^k$ -wertige Folgen von Zufallsvariablen auf  $(\Omega, \mathcal{A}, P)$  mit  $X_n = O_P(1)$  und  $Y_n = o_P(1)$ . Dann folgt

$$X_n \cdot Y_n = o_P(1). \quad (2.50)$$

**Satz 2.30 (Continuous Mapping Theorem).**

Es sei  $\{X_n\}_{n \in \mathbb{N}_0}$  eine  $\mathbb{R}^k$ -wertige Folge von Zufallsvariablen. Ist  $g : \mathbb{R}^k \rightarrow \mathbb{R}^m$  eine messbare Abbildung und für die stets Borel-messbare Menge  $D_g := \{x \in \mathbb{R}^k : g \text{ ist unstetig in } x\}$  gelte  $P(X_0 \in D_g) = 0$ , dann gilt die Implikation

$$X_n \xrightarrow{D} X_0 \quad \Rightarrow \quad g(X_n) \xrightarrow{D} g(X_0). \quad (2.51)$$

Will man für eine Folge  $k$ -dimensionaler Zufallsvektoren Konvergenz in Verteilung nachweisen, so erleichtert die Cramér-Wold-Technik die Situation erheblich, indem sie das Problem auf eine Konvergenz von eindimensionalen Zufallsvariablen zurückführt.

**Satz 2.31 (Cramér-Wold-Technik).**

$\{X_n\}_{n \in \mathbb{N}_0}$  seien  $\mathbb{R}^k$ -wertige Zufallsvariablen. Dann gilt

$$X_n \xrightarrow{D} X_0 \quad \Leftrightarrow \quad c^T X_n \xrightarrow{D} c^T X_0, \quad \forall c \in \mathbb{R}^k. \quad (2.52)$$

**Satz 2.32 (Zentraler Grenzwertsatz für Martingaldifferenzschemata).**

Für jedes  $n \in \mathbb{N}_0$  sei die Folge  $\xi^n := \{\xi_{n,t}, \mathcal{F}_t^n\}_{t=0}^n$  eine quadratisch integrierbare Martingaldifferenz, d.h. für alle  $t \in \{1, \dots, n\}$  seien  $E[\xi_{n,t}^2] < \infty$  und  $E[\xi_{n,t} | \mathcal{F}_{t-1}^n] = 0$  erfüllt. Gilt für alle  $n \in \mathbb{N}$  zusätzlich

$$\sum_{t=1}^n E[\xi_{t,n}^2] = 1 \quad (2.53)$$

sowie  $\xi_{n,0} = 0$ , dann sind die beiden Bedingungen

$$\sum_{t=1}^n E[\xi_{t,n}^2 I(|\xi_{t,n}| > \delta) | \mathcal{F}_{t-1}^n] \xrightarrow{P} 0, \quad (2.54)$$

$$\sum_{t=1}^n E[\xi_{t,n}^2 | \mathcal{F}_{t-1}^n] \xrightarrow{P} 1 \quad (2.55)$$

hinreichend (und notwendig) für die Verteilungskonvergenz

$$\sum_{t=1}^n \xi_{t,n} \xrightarrow{D} \mathcal{N}(0, 1). \quad (2.56)$$

# Kapitel 3

## Das zeitabhängige $AR$ -Modell ( $tvAR$ -Modell)

Die klassische Zeitreihenanalyse beschäftigt sich mit der Untersuchung von Abhängigkeitsstrukturen einzelner Beobachtungen. Weit verbreitet sind Modelle wie z.B.  $AR$ -Modelle oder die allgemeinere Klasse der  $ARMA$ -Modelle (siehe Abschnitt 2.1), bei denen man eine über die Zeit *unveränderliche* funktionale Abhängigkeit der Beobachtungen von ihren Vorgängern unterstellt. Zusätzlich geht man davon aus, dass diese Abhängigkeit mit einem Fehler behaftet ist, welcher etwa durch Messfehler oder Fehlspezifikation des Modells erklärt werden kann. Diese Fehlereinflüsse werden dabei durch geeignete Zufallsvariablen, etwa einem weißen Rauschen (siehe Definition 2.5), modelliert. Daher spricht man auch von stochastischer, statt von funktionaler Abhängigkeit. Man kommt jedoch schnell zu der Erkenntnis, dass die Annahme einer über die Zeit unveränderlichen Abhängigkeit der Beobachtungen reale Sachverhalte oft nur unzureichend beschreiben kann. Daher macht es Sinn in einem zeitvariablen autoregressiven Modell, dem sogenannten  $tvAR$ -Modell (time varying), diese Forderung zu lockern und eine über die Zeit *veränderliche* Abhängigkeitsstruktur zuzulassen.

In Abschnitt 3.1 untersuchen wir einführend das klassische stationäre  $AR$ -Modell und stellen zur Veranschaulichung die Realisierungen zweier  $AR(1)$ -Modelle graphisch dar und diskutieren deren Eigenschaften kurz. Im zweiten Abschnitt motivieren wir zunächst die Notwendigkeit eines zeitvariablen  $AR$ -Modells ( $tvAR$ -Modell) durch das nicht-stationäre Verhalten einer Finanzzeitreihe, bevor wir den Ansatz zeitabhängiger Spektraldarstellungen von Priestley [21] und dessen Weiterentwicklung von Dahlhaus [6] präsentieren. Bevor wir im letzten Teil dieses Kapitels die Frage nach der Existenz kausaler Lösungen beantworten und einen Zusammenhang im  $tvAR$ -Modell zwischen Stationarität und Nicht-Stationarität herstellen, beschäftigen wir uns in Abschnitt 3.3 mit dem Problem der Simulation von  $AR$ - und  $tvAR$ -Prozessen.

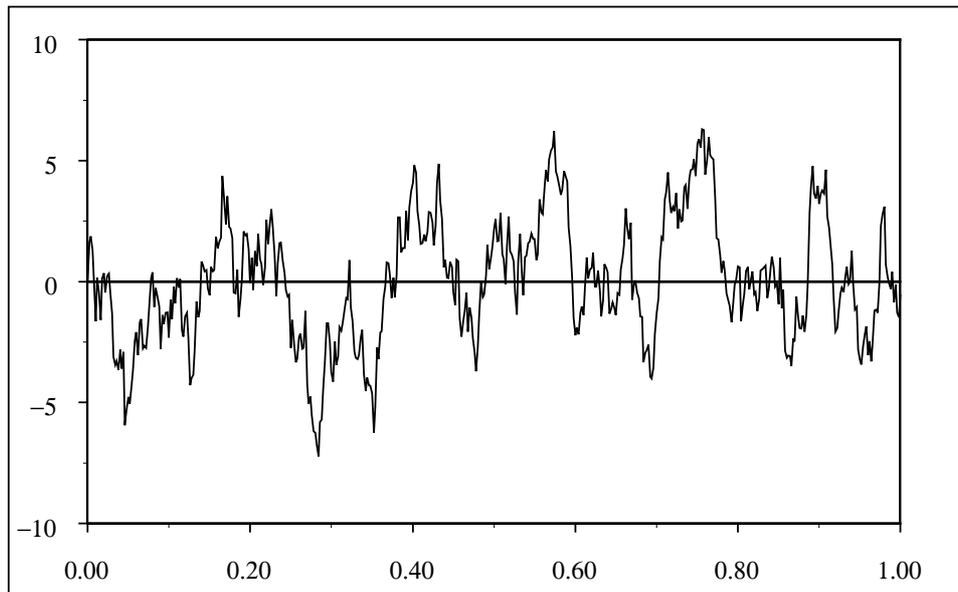


Abbildung 3.1:  $AR(1)$ -Reihe mit  $n=500$ ,  $\alpha = 0,9$  und  $\epsilon_t \sim \mathcal{N}(0, 1)$ .

### 3.1 Das stationäre $AR$ -Modell

Im stationären autoregressiven Modell  $p$ -ter Ordnung ( $AR(p)$ -Modell, siehe Definition 2.9) errechnen sich die Werte einer Zeitreihe als gewichtete Linearkombinationen der  $p$  Vorgänger plus Fehler. Dabei unterstellt man den Koeffizienten bzw. Gewichten des Modells, dass sie über die Zeit konstant bleiben. Es liegt also eine unveränderliche Abhängigkeitsstruktur vor. Genauer beschäftigt man sich mit Lösungen der Differenzgleichungen

$$X_t = \sum_{k=1}^p \alpha_k X_{t-k} + \epsilon_t, \quad \forall t \in \mathbb{Z}, \quad (3.1)$$

wobei  $\epsilon_t \sim (0, \sigma_\epsilon^2)$  ein weißes Rauschen beschreibt. Nach Korollar 2.13 existiert eine stationäre kausale Lösung genau dann, wenn  $1 - \sum_{k=1}^p \alpha_k z^k \neq 0$  für alle  $z$  mit  $|z| \leq 1$  gilt.

Um die Bedeutung der Koeffizienten  $\alpha_k$ ,  $k = 1, \dots, p$  in Gleichung (3.1) besser verstehen zu können, wollen wir den Spezialfall  $p = 1$  mit  $\alpha := \alpha_1$ , also

$$X_t = \alpha X_{t-1} + \epsilon_t, \quad \forall t \in \mathbb{Z} \quad (3.2)$$

näher untersuchen. Für das  $AR(1)$ -Modell (3.2) existiert eine stationäre kausale Lösung genau dann, wenn  $|\alpha| < 1$  gilt.

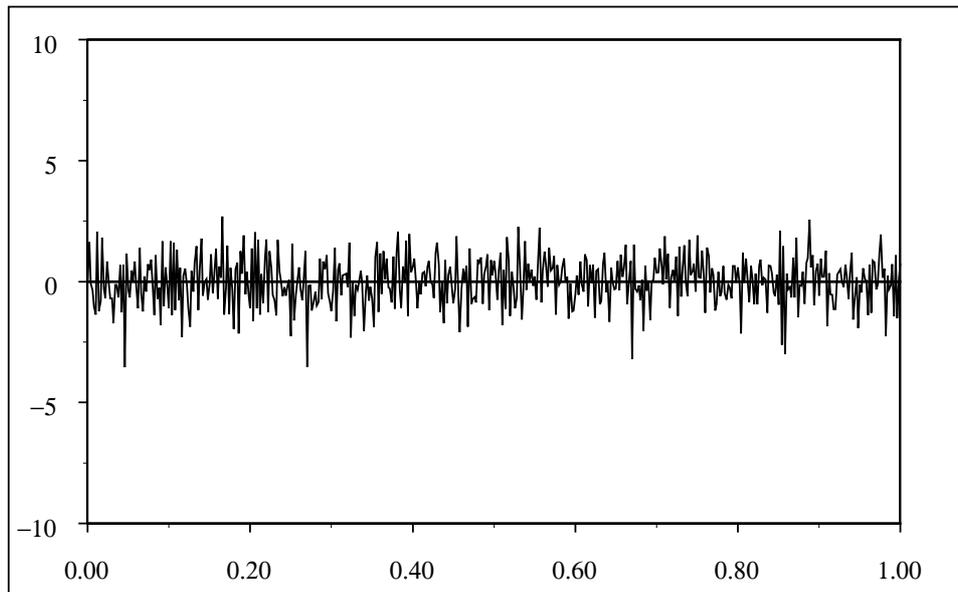


Abbildung 3.2:  $AR(1)$ -Reihe mit  $n=500$ ,  $\alpha = -0,3$  und  $\epsilon_t \sim \mathcal{N}(0,1)$ .

Zur Veranschaulichung haben wir zwei  $AR(1)$ -Reihen mit  $n = 500$  simuliert und graphisch dargestellt (siehe auch Abschnitt 3.3). Abbildung 3.1 zeigt eine Realisierung einer  $AR(1)$ -Zeitreihe mit positivem Koeffizienten  $\alpha = 0,9$  nahe bei eins und Abbildung 3.2 eine mit negativem Koeffizienten  $\alpha = -0,3$  nahe bei null. Bei beiden Simulationen haben wir ein standard-normalverteiltes weißes Rauschen  $\epsilon_t \sim \mathcal{N}(0,1)$  verwendet. Die Kurvenverläufe beider Realisierungen haben als charakteristische Merkmale der Stationarität zum einen eine in gewisser Weise gleichbleibende Struktur und zum anderen schwanken sie um null.

Am Anfang des folgenden Abschnitts gehen wir näher auf die Kurvenverläufe der Pfade aus den Abbildungen 3.1 und 3.2 ein und interpretieren deren Unterschiede bzw. Gemeinsamkeiten, bevor wir die Probleme bei der Modellierung realer Daten diskutieren und die Grenzen stationärer autoregressiver Modelle aufzeigen.

## 3.2 Motivation des $tvAR$ -Modells

Vergleicht man die Kurvenverläufe der beiden Realisierungen aus Abbildung 3.1 und 3.2, so fallen einem im Wesentlichen zwei Unterschiede auf. Zum einen verläuft die erste Kurve mit positivem Koeffizienten deutlich ruhiger als die zweite mit negativem Koeffizienten. Das liegt gerade daran, dass beim Übergang von einem Zeitpunkt zum nächsten der

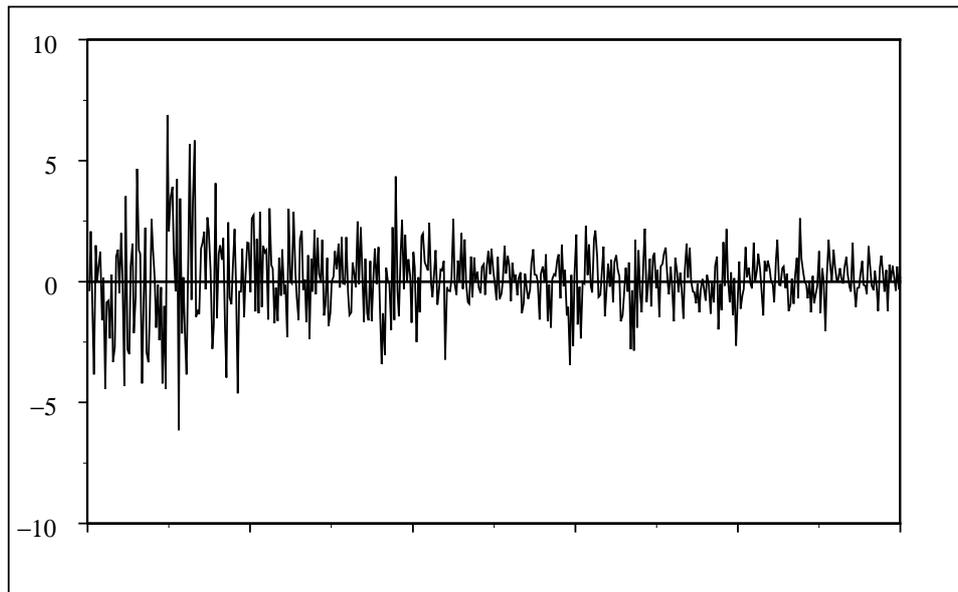


Abbildung 3.3: Relative Ein-Tages-Veränderungen des DAX für den Zeitraum 03.01.2003 bis 30.12.2004 in Prozent %.

Vorgänger prozentual mit gleichem Vorzeichen übernommen wird und der dazu addierte Fehler um null streut. Dies hat zur Folge, dass die Kurve in Abbildung 3.1 weniger oszilliert als die in Abbildung 3.2, wo jeder Wert prozentual mit umgekehrtem Vorzeichen in seinen Nachfolger eingeht und sich daher die Vorzeichen der Zeitreihe oft ändern und die Kurve stark um null schwankt. Zum anderen bewirkt der vom Betrag her größere Koeffizient  $\alpha = 0,9$ , dass die Amplitude in Abbildung 3.1 größer ist als in der zweiten Abbildung.

Beide Kurven haben aber die Gemeinsamkeit, dass sie ihre Struktur, also im ersten Beispiel den ruhigen und im zweiten den stärker oszillierenden Verlauf, über die Zeit beibehalten, was gerade charakteristisch für stationäre Prozesse ist. Diese Eigenschaft macht es möglich, die *AR*-Koeffizienten in gewisser Weise als Durchschnitt über die Zeit konsistent zu schätzen und erlaubt somit einen relativ einfachen asymptotischen Zugang (siehe Kreiß [15], Kapitel 11).

Bei praktischen Anwendungen lassen sich allerdings häufig strukturelle Schwankungen über die Zeit beobachten, welche nicht vernünftig mit einem stationären Ansatz zu erklären sind. Als prominentes Beispiel haben wir für den Zeitraum 03.01.2003 bis 30.12.2004 aus den Tagesschlusskursen des Deutschen Aktien Indizes (DAX) die relativen Ein-Tages-Veränderungen errechnet und diese in Abbildung 3.3 aufgetragen. Man erkennt hier das

nicht-stationäre Verhalten dieser Zeitreihe deutlich an den hohen Ausschlägen in der linken Hälfte der Abbildung und daran, dass die Ausschlagshöhe im Laufe der Zeit immer mehr abnimmt.

Es drängt sich daher die Frage auf wie sich solche Phänomene geeignet modellieren lassen, ohne sich dabei komplett von der weit entwickelten Theorie autoregressiver Prozesse zu lösen. Nahe liegend scheint hier sich von der einschränkenden Annahme der Stationarität zu befreien und die bestehende Theorie auf den nicht-stationären Fall auszudehnen, indem man die Forderung nach konstanten Koeffizienten  $\alpha_k$  lockert und ihnen eine Abhängigkeit von der Zeit gestattet, man also Koeffizientenfunktionen  $\alpha_k(t)$  betrachtet. Dabei fordert man von diesen Funktionen, dass sie sich *langsam* mit der Zeit ändern. Diese Glattheit ermöglicht es nämlich lokal die Theorie stationärer  $AR$ -Reihen anzuwenden (siehe Satz 3.6). Man will also *global* nicht-stationäre Prozesse betrachten, welche *lokal* ein annähernd stationäres Verhalten aufweisen.

Die Idee eines lokal approximativ stationären Prozesses fand erstmalig 1965 in der Arbeit von Priestley [21] Erwähnung, wo der Autor eine Theorie basierend auf zeitabhängigen Spektren (für Spektraldarstellung im zeitunabhängigen Fall siehe Abschnitt 2.1) entwickelte. Priestley betrachtete Prozesse mit einer zeitabhängigen Spektraldarstellung

$$X_t = \int_{(-\pi, \pi]} e^{i\lambda t} A(t, \lambda) d\xi(\lambda), \quad t \in \mathbb{Z}, \quad (3.3)$$

wobei  $\xi(\lambda)$  einen Prozess mit orthogonalen Zuwächsen bezeichnet und  $A(t, \lambda)$  eine zeitvariable glatte Transferfunktion ist.

Innerhalb der von Priestley entwickelten Theorie für nicht-stationäre Prozesse ist es aber unmöglich eine sinnvolle asymptotische Theorie aufzusetzen. Beispielsweise ist es nicht möglich wie im stationären Fall einen konsistenten Schätzer über ein gewisses arithmetisches Mittel zu definieren. Asymptotische Betrachtungen werden aber gebraucht, um die statistischen Untersuchungen zu vereinfachen, da es hoffnungslos ist, Berechnungen für eine endliche Anzahl von Daten anzustellen. Seien etwa die Beobachtungen  $X_1, \dots, X_n$  eines zeitvariablen  $AR(p)$ -Modells nach Priestley

$$X_t = \sum_{k=1}^p \alpha_k(t) X_{t-k} + \epsilon_t, \quad t = 1, \dots, n \quad (3.4)$$

mit weißem Rauschen  $\epsilon_t \sim (0, \sigma_\epsilon^2)$  gegeben, wobei  $\alpha_k(\cdot) : \mathbb{N} \rightarrow \mathbb{R}$  zeitabhängige Koeffizientenfunktionen sind. Lässt man nun  $n$  gegen  $\infty$  streben, d.h. man setzt den Prozess in die Zukunft fort, so ergibt sich aufgrund der Nicht-Stationarität *keine* Zunahme an Information über das Verhalten des Prozesses in der Gegenwart. Dies liegt daran, dass sich die Struktur der Zeitreihe mit  $t$  ändert und daher zukünftige Daten keine oder nur

wenig Aussagekraft über die Gestalt der Reihe in der Gegenwart haben.

Hier ist also ein anderer Zugang von Nöten. Dahlhaus [6] wählte den aus der nicht-parametrischen Kurvenschätzung bekannten Ansatz, dass für größer werdendes  $n$  der Prozess auf einem immer feineren Gitter des selben Intervalls beobachtet wird. Er skaliert den Prozess auf das Intervall  $[0, 1]$  um und betrachtet für das zeitabhängige  $AR$ -Modell das Dreiecksschema

$$X_{t,n} = \sum_{k=1}^p \alpha_k\left(\frac{t}{n}\right) X_{t-k,n} + \epsilon_t, \quad t = 1, \dots, n \quad (3.5)$$

mit  $\epsilon_t \sim (0, \sigma_\epsilon^2)$  und Koeffizientenfunktionen  $\alpha_k(\cdot) : [0, 1] \rightarrow \mathbb{R}$ . Lässt man nun in diesem neuen Modell  $n$  immer größer werden, so stehen einem für jeden Zeitpunkt  $t \in [0, 1]$  immer mehr Beobachtungen in der Nähe von  $t$  zur Verfügung, man erhält also immer *mehr* Information über die *lokale* Struktur der Zeitreihe.

Weiter oben haben wir bereits die Klasse von Modellen, die wir hier untersuchen wollen, eingeschränkt, indem wir von nicht-stationären Zeitreihen gesprochen haben, die lokal ein stationäres Verhalten zeigen. Denn machen wir keinerlei Einschränkungen hinsichtlich des Grades der Nicht-Stationarität, so ist es einfach unmöglich, hierfür eine bedeutsame Theorie zu entwickeln. Um pathologische Fälle, hervorgerufen etwa durch stark oszillierende Koeffizientenfunktionen  $\alpha_k(\cdot)$ , auszuschließen, schränken wir das Verhalten des Prozesses ein, indem wir fordern, dass er sich lokal nicht wesentlich von einem stationären Prozess unterscheiden soll. Man erreicht dies durch Einbetten einer stationären Struktur zu jedem Zeitpunkt  $t$ . Dabei ist der Ansatz vergleichbar mit der nicht-parametrischen Technik eine Gerade lokal an eine nicht-lineare Kurve anzulegen (siehe Kapitel 4). Genauso wie man dort eine Glattheitsbedingung der Kurve benötigt, um den Ansatz zu rechtfertigen, müssen wir in unserem Fall auch eine gewisse Glattheit der  $\alpha_k(\cdot)$  fordern.

Dahlhaus gelang es in [3] den Begriff der lokalen Stationarität gründlich zu definieren und damit die vorangegangenen Überlegungen zu präzisieren. Dabei beschreibt er die geforderte Glattheit mit Hilfe der Spektraldarstellung bzw. mit der Stetigkeit einer approximierenden Transferfunktion (siehe Satz 2.16).

**Definition 3.1 (Lokale Stationarität).**

Das Dreiecksschema  $X_{t,n}$ ,  $t = 1, \dots, n$  heißt lokal stationär mit Transferfunktion  $A^0(\cdot)$  und Trend  $\mu(\cdot)$ , falls eine Spektraldarstellung

$$X_{t,n} = \mu\left(\frac{t}{n}\right) + \int_{(-\pi, \pi]} \exp(i\lambda t) A_{t,n}^0(\lambda) d\xi(\lambda) \quad (3.6)$$

existiert, mit

(i)  $\xi(\lambda)$  ist ein stochastischer Prozess auf  $[-\pi, \pi]$  mit  $\overline{\xi(\lambda)} = \xi(-\lambda)$  und

$$\text{cum}\{d\xi(\lambda_1), \dots, d\xi(\lambda_k)\} = \eta \left( \sum_{j=1}^k \lambda_j \right) h_k(\lambda_1, \dots, \lambda_{k-1}) d\lambda_1 \cdots d\lambda_k, \quad (3.7)$$

wobei  $\text{cum}\{\dots\}$  den Kumulanten  $k$ -ter Ordnung bezeichne,  $h_1 = 0$ ,  $h_2(\lambda) = 1$ ,  $|h_k(\lambda_1, \dots, \lambda_{k-1})| \leq \text{const}_k$  für alle  $k$  und  $\eta(\lambda) = \sum_{j=-\infty}^{\infty} \delta(\lambda + 2\pi j)$  die  $2\pi$ -periodische Fortsetzung der Dirac Delta-Funktion ist.

(ii) Es existiert eine Konstante  $K$  und eine  $2\pi$ -periodische Funktion  $A : [0, 1] \times \mathbb{R} \rightarrow \mathbb{C}$  mit  $A(u, -\lambda) = \overline{A(u, \lambda)}$  und

$$\sup_{t, \lambda} |A_{t,n}^0(\lambda) - A(\frac{t}{n}, \lambda)| \leq Kn^{-1} \quad (3.8)$$

für alle  $n$ .  $A(u, \lambda)$  und  $\mu(u)$  seien stetig in  $u$ .

In seiner Arbeit betrachtet Dahlhaus unter gewissen zusätzlichen Annahmen die Differenzgleichungen (3.5), wo er jedoch wie in obiger Definition eine veränderliche Trendfunktion  $\mu(\cdot)$  sowie eine veränderliche Funktion der Standardabweichung des weißen Rauschens  $\sigma(\cdot)$  zulässt. In dieser Arbeit wollen wir aber nur den Fall ohne Trend  $\mu(\cdot) \equiv 0$  behandeln sowie eine über die Zeit konstante Standardabweichung  $\sigma(\cdot) \equiv \sigma_\epsilon$ , welche bei uns in das weiße Rauschen  $\epsilon_t \sim (0, \sigma_\epsilon^2)$  einfließt, zulassen.

Es ergibt sich als Spezialfall des Theorems 2.3 aus [3] die lokale Stationarität von (3.5) aus folgendem Satz. Zu beachten ist dabei, dass die Modellgleichung (3.9) nicht mehr nur auf  $[0, 1]$ , also für  $t \in \{1, \dots, n\}$ , sondern auf  $\mathbb{R}$  betrachtet wird, indem man  $t \in \mathbb{Z}$  zulässt.

### Satz 3.2 (Lokale Stationarität des tvAR-Modells).

Betrachte das System von Differenzgleichungen

$$X_{t,n} = \sum_{k=1}^p \alpha_k(\frac{t}{n}) X_{t-k,n} + \epsilon_t, \quad t \in \mathbb{Z} \quad (3.9)$$

mit  $\epsilon_t \sim (0, \sigma_\epsilon^2)$ . Für  $k = 1, \dots, p$  seien die Koeffizientenfunktionen  $\alpha_k(\cdot)$  stetig auf  $\mathbb{R}$ , wobei  $\alpha_k(u) = \alpha_k(0)$  für  $u < 0$  und  $\alpha_k(u) = \alpha_k(1)$  für  $u > 1$  gelte, und differenzierbar für  $u \in (0, 1)$  mit beschränkten Ableitungen. Falls  $1 - \sum_{k=1}^p \alpha_k(u) z^k \neq 0$  für alle  $z \in \mathbb{C}$  mit  $|z| \leq 1 + c$  mit  $c > 0$  gleichmäßig in  $u$ , dann hat das obige Differenzgleichungssystem eine Lösung der Form (3.6) mit Transferfunktion

$$A(u, \lambda) = \frac{\sigma_\epsilon}{\sqrt{2\pi}} \left( 1 - \sum_{k=1}^p \alpha_k(u) e^{-ij\lambda} \right)^{-1} \quad (3.10)$$

und zeitvariabler Spektraldichte

$$f(u, \lambda) = \frac{\sigma_\epsilon^2}{2\pi} \left| 1 - \sum_{k=1}^p \alpha_k(u) e^{-ij\lambda} \right|^{-2}. \quad (3.11)$$

Unter den Voraussetzungen von Satz 3.2 ist das zeitabhängige  $tvAR(p)$ -Modell also lokal stationär im Sinne der Definition 3.1 von Dahlhaus.

**Bemerkung 3.3.**

Die Eigenschaft der lokalen Stationarität des  $tvAR$ -Modells ist wesentlich für die folgenden Kapitel. Daher werden wir uns im Weiteren mit den Begriffen  $tvAR$ -Modell bzw.  $tvAR$ -Prozess auf die Modellgleichung (3.9) beziehen, wobei die Koeffizientenfunktionen  $\alpha_k(\cdot)$  wie in Satz 3.2 stetig durch  $\alpha_k(u) = \alpha_k(0)$  für  $u < 0$  und  $\alpha_k(u) = \alpha_k(1)$  für  $u > 1$  auf  $\mathbb{R}$  fortgesetzt sind.  $\square$

Um ein Gespür zu bekommen, was lokal stationär gemäß Definition 3.1 eigentlich bedeutet bzw. welche Auswirkungen das Erlauben von zeitabhängigen Koeffizienten auf die Pfade einer  $tvAR$ -Zeitreihe hat, werden wir uns im nächsten Abschnitt mit der Simulation solcher Prozesse beschäftigen.

### 3.3 Simulation von $tvAR$ -Prozessen

Sowohl im stationären  $AR$ -Modell, als auch im zeitabhängigen  $tvAR$ -Modell tritt bei der Simulation einer Realisierung das Problem auf, dass man keinen Anfang hat. Im  $AR(1)$ -Modell etwa steht einem keine anfängliche Beobachtung  $X_t$  zur Verfügung aus der sich der zeitliche Nachfolger  $X_{t+1}$  berechnen lässt, da man ja die Zeitreihe als Lösung auf der Menge der ganzen Zahlen  $\mathbb{Z}$  betrachtet.

Im  $AR(p)$ -Modell löst man dieses Problem, indem man  $p$  Anfangswerte gleich null setzt und die Reihe einschwingen lässt. Möchte man beispielsweise die Realisierung eines  $AR(2)$ -Prozesses mit  $n = 500$  simulieren, so setzt man zwei Anfangswerte gleich null, berechnet die Werte der Zeitreihe für  $n = 600$  und benutzt dann nur die letzteren 500 Beobachtungen. Dieses Vorgehen kann man damit rechtfertigen, dass die Koeffizienten konstant sind und sich der Prozess exponentiell schnell einschwingt (siehe etwa Schlittgen [22], S.106f).

Die Situation im  $tvAR$ -Modell unterscheidet sich von der oben beschriebenen dadurch, dass die Koeffizienten nicht konstant sind, sondern veränderliche Funktionen auf  $[0, 1]$  bzw. konstant fortgesetzt auf  $\mathbb{R}$  darstellen. Ein analoges Vorgehen wie bei  $AR$ -Reihen ist hier also zunächst nicht möglich, da nicht klar ist wie man den Prozess sinnvoll einschwingen

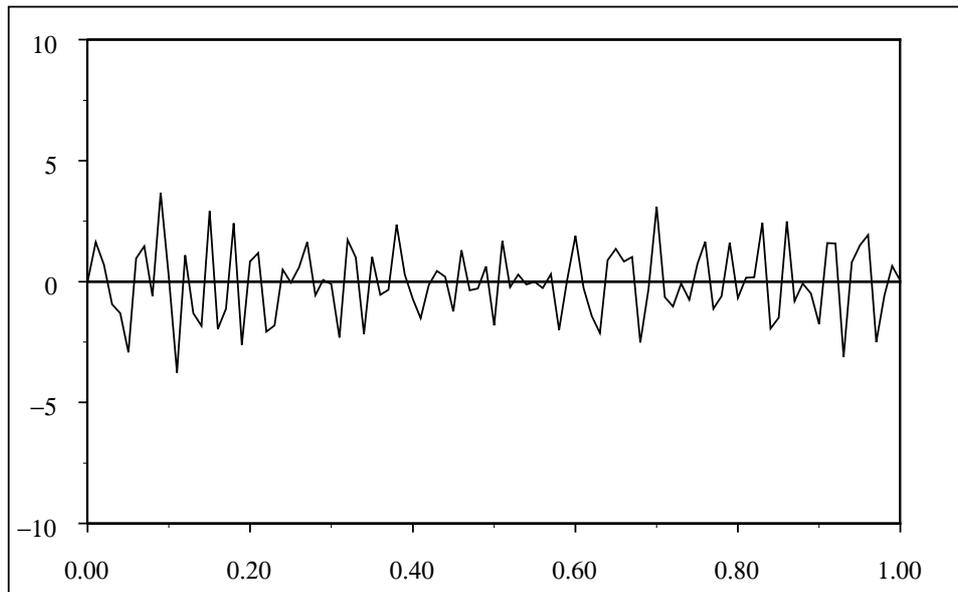


Abbildung 3.4:  $tvAR(4)$ -Reihe *ohne* Einschwingen mit  $n = 100$ ,  $\alpha_1(u) = -0,9u + 0,2$ ,  $\alpha_2(u) = -0,4u - 0,4$ ,  $\alpha_3(u) = -0,8u + 0,2$ ,  $\alpha_4(u) = 0,9u - 1$  und  $\epsilon_t \sim \mathcal{N}(0, 1)$ .

lassen soll. Ein auf den ersten Blick adäquater Ausweg wäre es, die ersten Beobachtungen, also  $X_{1,n}, \dots, X_{p,n}$ , gleich null zu setzen und damit die Reihe zu simulieren. Dieser Ansatz ist aber nicht sonderlich zufriedenstellend, denn selbst für großes  $n$ , also eine große Anzahl von Beobachtungen, würde man den Prozess aufgrund der Einschwingphase an der linken Intervallgrenze nicht richtig wiedergeben und den Kurvenverlauf verfälschen. Ein ebenfalls recht nahe liegendes Setzen der Beobachtungen  $X_{1-p,n} = \dots = X_{0,n} = 0$ , welche nicht mehr in  $[0, 1]$  liegen, würde zwar das Problem des Einschwingens etwas weiter an den Rand des Intervalls schieben, jedoch immer noch die Kurve nicht richtig darstellen. Insbesondere wenn man es mit Koeffizientenfunktionen zu tun hat, für welche nicht gerade  $\alpha_k(0) = 0$  gilt, also wo  $\alpha_k(\cdot)$  nicht glatt an null anschließt.

Man löst dieses Problem, indem man die stetige Fortsetzung der  $\alpha_k(\cdot)$  aus Satz 3.2 benutzt und damit den Prozess außerhalb von  $[0, 1]$  als stationär betrachtet. Dies macht es möglich, das oben beschriebene Einschwingverfahren für  $AR$ -Zeitreihen zu benutzen, um sich sinnvolle Anfangswerte  $X_{1-p,n}, \dots, X_{0,n}$  zu erzeugen, womit man dann die Simulation für das Intervall  $[0, 1]$  vornimmt. Durch das glatte Fortsetzen bei null erreicht man ein schnelleres Einschwingen und erhält einen besseren Kurvenverlauf in der Nähe der linken Intervallgrenze.

Um die Unterschiede der eben angesprochenen Simulationsverfahren von  $tvAR$ -Prozessen

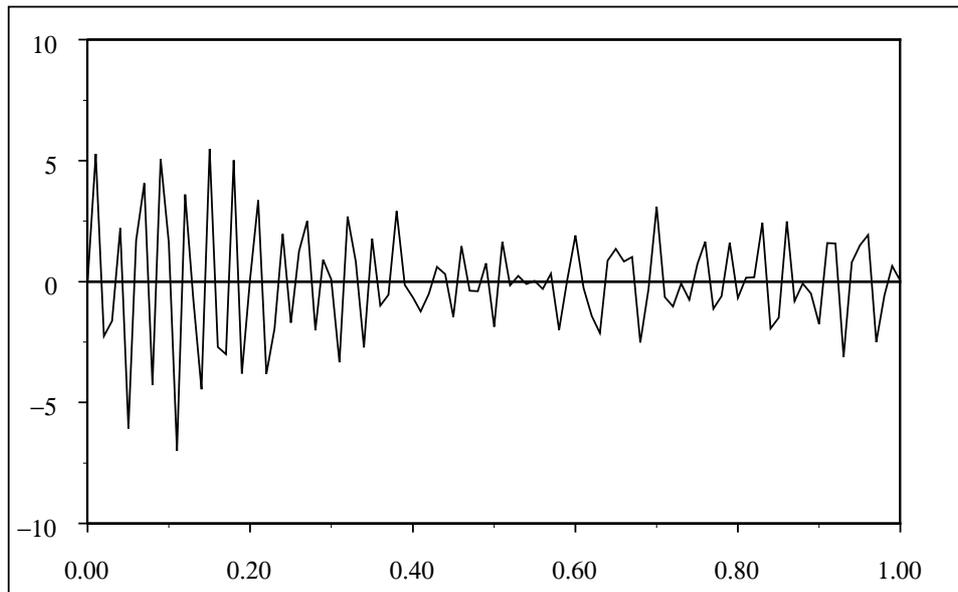


Abbildung 3.5:  $tvAR(4)$ -Reihe mit Einschwingen mit  $n = 100$ ,  $\alpha_1(u) = -0,9u + 0,2$ ,  $\alpha_2(u) = -0,4u - 0,4$ ,  $\alpha_3(u) = -0,8u + 0,2$ ,  $\alpha_4(u) = 0,9u - 1$  und  $\epsilon_t \sim \mathcal{N}(0, 1)$ .

graphisch zu veranschaulichen, haben wir dieselbe  $tvAR(4)$ -Reihe einmal mit Startwerten  $X_{-3,100} = \dots = X_{0,100} = 0$  (Abbildung 3.4) und einmal mit Anfangswerten, welche wir zuvor durch die Simulation der stationären  $AR(4)$ -Reihe  $X_t = \alpha_1(0)X_{t-1} + \alpha_2(0)X_{t-2} + \alpha_3(0)X_{t-3} + \alpha_4(0)X_{t-4} + \epsilon_t$  erzeugt haben (Abbildung 3.5), simuliert. Man sieht deutlich, dass das Verfahren, welches in Abbildung 3.4 angewandt wurde, den Kurvenverlauf im ersten Drittel des Intervalls  $[0, 1]$  verfälscht, weil es dort die Höhe der Ausschläge zu klein darstellt.

Zur genaueren Veranschaulichung der Auswirkungen zeitabhängiger Koeffizienten haben wir in den Abbildungen 3.6 und 3.7 zwei weitere  $tvAR$ -Reihen graphisch dargestellt. In Abbildung 3.6 sieht man die Realisierung einer  $tvAR(1)$ -Reihe mit  $\alpha(u) = 0,9 \sin(2\pi u)$  (siehe auch Abbildung 4.1), wobei die Simulation mit Anfangswert  $X_{0,500} = 0$  vorgenommen wurde. Dieses hat hier keine signifikanten Auswirkungen auf den Kurvenverlauf, da für die Koeffizientenfunktion  $\alpha(0) = 0$  gilt. Abbildung 3.7 zeigt den Pfad eines  $tvAR(1)$ -Prozesses mit linearer Koeffizientenfunktion  $\alpha(u) = 1,8u - 0,9$ . Da in diesem Fall  $\alpha(0) = -0,9$  gilt, haben wir hier durch Simulation der stationären  $AR(1)$ -Reihe  $X_t = -0,9X_{t-1} + \epsilon_t$  einen Anfangswert erzeugt, womit wir dann die  $tvAR(1)$ -Reihe generiert haben. Betrachtet man die beiden Abbildungen, so ist als Merkmal der Nicht-Stationarität eine Veränderung der Struktur über die Zeit deutlich erkennbar.

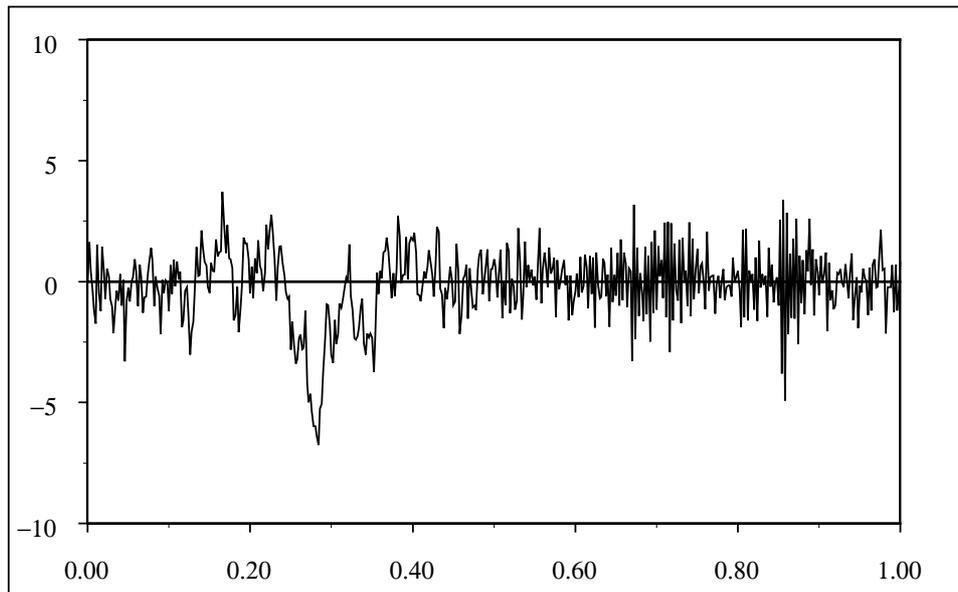


Abbildung 3.6: *tvAR*(1)-Reihe mit  $n=500$ ,  $\alpha(u) = 0,9 \sin(2\pi u)$  und  $\epsilon_t \sim \mathcal{N}(0, 1)$ .

In Abbildung 3.6 fällt einerseits auf, dass die Kurve in der ersten Hälfte des Intervalls  $[0, 1]$  ruhiger verläuft als in der zweiten, was daran liegt, dass  $\alpha(u)$  für  $u \in [0, 1/2]$  positiv und für  $u \in [1/2, 1]$  negativ ist (siehe auch Abschnitt 3.2). Andererseits verhält sich die Amplitude über die Zeit entsprechend der Höhe der Koeffizientenfunktion. In Abbildung 3.7 ist eine immer ruhiger werdende Kurve zu sehen, deren Amplitude im Intervall  $[0, 1/2]$  kontinuierlich abnimmt und danach wieder zunimmt, was auf die lineare Gestalt von  $\alpha(\cdot)$  zurückzuführen ist.

Insgesamt kommt man zu der Überzeugung, dass die Verallgemeinerung des stationären autoregressiven Modells durch Zulassen von sich langsam ändernden Koeffizienten ein gutes Werkzeug darstellt um nicht-stationäres Verhalten von Zeitreihen modellieren zu können. Im *tvAR*-Modell ist es wie in den Abbildungen 3.6 und 3.7 zu sehen möglich, die verschiedenen Ausprägungen der Struktur einer Realisierung, etwa den Verlauf über die Zeit, variabel zu gestalten. Beispielsweise lässt sich der Oszillationsgrad zeitabhängig modellieren.

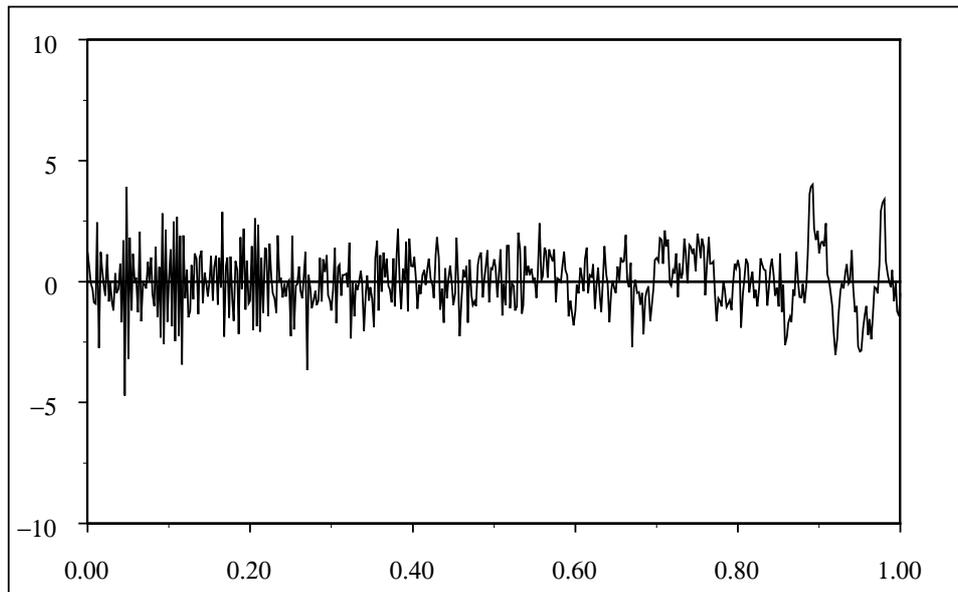


Abbildung 3.7:  $tvAR(1)$ -Reihe mit  $n = 500$ ,  $\alpha(u) = 1,8u - 0,9$  und  $\epsilon_t \sim \mathcal{N}(0, 1)$ .

### 3.4 Eigenschaften von $tvAR$ -Prozessen

Wie bereits in Kapitel 2 erwähnt, interessiert man sich in der klassischen Zeitreihenanalyse besonders für kausale Lösungen der zugrunde liegenden Modellgleichungen (siehe Definition 2.11). Dies liegt vor allem daran, dass sie leicht zu handhaben sind und an der natürlichen Vorstellung, dass der Wert einer Zeitreihe zu jedem Zeitpunkt  $t$  nur von der Vergangenheit abhängen sollte. Es stellt sich also die Frage ob bzw. unter welchen Bedingungen die  $tvAR(p)$ -Modellgleichungen

$$\begin{aligned}
 X_{1,n} &= \sum_{k=1}^p \alpha_k\left(\frac{1}{n}\right) X_{1-k,n} + \epsilon_1, \\
 X_{2,n} &= \sum_{k=1}^p \alpha_k\left(\frac{2}{n}\right) X_{2-k,n} + \epsilon_2, \\
 &\vdots \\
 X_{n,n} &= \sum_{k=1}^p \alpha_k\left(\frac{n}{n}\right) X_{n-k,n} + \epsilon_n
 \end{aligned} \tag{3.12}$$

solche Lösungen  $X_{t,n}$  besitzen, welche nur von  $\epsilon_s$ ,  $s \leq t$  abhängen. Im zeitunabhängigen Fall fordert man außerdem die absolute Summierbarkeit der Koeffizienten  $\{\psi_k\}_{k \in \mathbb{N}}$  für die kausale  $MA(\infty)$ -Darstellung. Als natürliche Verallgemeinerung von (2.12) ist es

hier sinnvoll, eine von  $\epsilon_s$ ,  $s \leq t$  abhängige Lösung der Differenzgleichungen (3.12) als kausal zu bezeichnen, wenn die Summe der Absolutbeträge der von  $t$  und  $n$  abhängigen Koeffizienten  $\{\psi_{j,t,n}\}_{j \in \mathbb{N}}$  der zeitabhängigen  $MA(\infty)$ -Darstellung gleichmäßig beschränkt bleibt in  $t$  und  $n$ . Statt  $\sum_{k=0}^{\infty} |\psi_k| < \infty$  im zeitunabhängigen Fall, fordert man also für eine kausale Lösung im  $tvAR(p)$ -Modell  $\sup_{t \leq n} \sum_{k=0}^{\infty} |\psi_{j,t,n}| < \infty$  (siehe auch (3.15)).

Unter ähnlichen Voraussetzungen wie in Satz 3.2 gelang es Künsch in [16] die Existenz einer kausalen Lösung von (3.12) mit oben geforderten Eigenschaften nachzuweisen.

**Satz 3.4 (Existenz einer kausalen Lösung).**

Für  $k = 1, \dots, p$  seien die Koeffizientenfunktionen  $\alpha_k(\cdot)$  in (3.12) stetig auf  $[0, 1]$  und auf  $(-\infty, 1]$  durch  $\alpha_k(u) = \alpha_k(0)$  für  $u < 0$  fortgesetzt. Weiter gelte  $1 - \sum_{k=1}^p \alpha_k(u)z^k \neq 0$  für alle  $z \in \mathbb{C}$  mit  $|z| \leq 1 + c$  mit  $c > 0$  gleichmäßig in  $u$ . Dann existiert eine Folge  $\{\varphi_{j,t,n}\}_{j \in \mathbb{N}}$ , so dass das System von Differenzgleichungen (3.12) eine Lösung der Form

$$X_{t,n} = \sum_{j=0}^{\infty} \varphi_{j,t,n} \epsilon_{t-j}, \quad 1 \leq t \leq n, \quad (3.13)$$

mit

$$\sup_{t \leq n} \sum_{j=0}^{\infty} |\varphi_{j,t,n}| < \infty \quad (3.14)$$

hat. Die Schreibweise in (3.14) steht hier (siehe auch (3.32)) abkürzend für

$$\exists C < \infty \forall n \in \mathbb{N} : \sup_{1 \leq t \leq n} \sum_{j=0}^{\infty} |\varphi_{j,t,n}| < C. \quad (3.15)$$

**Bemerkung 3.5.**

Künsch stützt seinen Beweis im Wesentlichen auf die aus der Theorie über lineare Differenzgleichungen bekannte Funktion von Green (siehe auch Miller [19], Hallin [11] und [12] sowie Mélard [18]) und ein Resultat über die Beschränktheit der Norm einer Matrix unter gewissen Bedingungen an ihre Eigenwerte (siehe Householder [13]). Definiert man für  $u \in (-\infty, 1]$  die Matrix von Green durch

$$\mathbf{A}(u) := \begin{pmatrix} \alpha_1(u) & \alpha_2(u) & \cdots & \cdots & \alpha_p(u) \\ 1 & 0 & \cdots & \cdots & 0 \\ 0 & 1 & \ddots & \cdots & \vdots \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & 1 & 0 \end{pmatrix}, \quad (3.16)$$

so gilt für die Koeffizienten

$$\varphi_{j,t,n} = \left( \prod_{k=0}^{j-1} \mathbf{A}\left(\frac{t-k}{n}\right) \right)_{1,1}. \quad (3.17)$$

Künsch zeigt in seiner Arbeit eine noch stärkere Eigenschaft als (3.14). Er beweist nämlich, dass für gewisse positive Konstanten  $C < \infty$  und  $\rho < 1$  gilt

$$\sup_{t \leq n} |\varphi_{j,t,n}| = \sup_{t \leq n} \left| \left( \prod_{k=0}^{j-1} \mathbf{A}\left(\frac{t-k}{n}\right) \right)_{1,1} \right| \leq C \cdot \rho^j. \quad (3.18)$$

□

Die von Künsch bewiesene Existenz jener kausalen Lösung hilft zunächst nur insofern weiter, dass man eine Darstellung gewonnen hat, welche explizit nur noch vom weißen Rauschen  $\{\epsilon_t\}_{t \in \mathbb{Z}}$  abhängt. Noch ist es aber weiterhin nicht möglich die bestehende Theorie für stationäre Zeitreihen anzuwenden.

Aufgrund der in Satz 3.2 nachgewiesenen lokalen Stationarität des  $tvAR$ -Modells können wir aber zeigen, dass sich die global nicht-stationäre zeitabhängige  $MA(\infty)$ -Reihe (3.13) lokal durch stationäre  $MA(\infty)$ -Prozesse approximieren lässt und somit die Theorie stationärer Zeitreihen auf diese anwenden. Außerdem werden wir zeigen, dass der Approximationsfehler von der Größenordnung  $O_P(\frac{1}{n})$  ist und sich die gleichmäßige Beschränktheit aus (3.14) auf die Koeffizienten der approximierenden  $MA(\infty)$ -Reihen überträgt.

### Lemma 3.6.

Es gelten die Voraussetzungen von Satz 3.2. Dann existiert für  $\{X_{t,n}\}_{t=1}^n$  aus (3.12) eine Folge auf  $(0, 1)$  differenzierbarer Funktionen  $\{\varphi_j(\cdot) | \varphi_j : [0, 1] \rightarrow \mathbb{R}\}_{j=0}^{\infty}$  und ein Prozess  $\{\tilde{X}_{t,n}\}_{t=1}^n$  mit  $\tilde{X}_{t,n} := \sum_{j=0}^{\infty} \varphi_j(\frac{t}{n}) \epsilon_{t-j}$ , so dass

$$(i) \quad \sup_{t \leq n} |X_{t,n} - \tilde{X}_{t,n}| = O_P\left(\frac{1}{n}\right), \quad (3.19)$$

$$(ii) \quad \sup_{t \leq n} \sum_{j=0}^{\infty} |\varphi_j(\frac{t}{n})| < \infty \quad (3.20)$$

gilt.

*Beweis.*

Für jedes feste  $u \in [0, 1]$  ist wegen der gleichmäßigen Beschränktheit weg vom Einheitskreis der Nullstellen von  $1 - \sum_{k=1}^p \alpha_k(u) z^k$  einerseits

$$A(u, \lambda) := \frac{\sigma_\epsilon}{\sqrt{2\pi}} \left[ 1 - \sum_{k=1}^p \alpha_k(u) e^{-i\lambda k} \right]^{-1} \quad (3.21)$$

die Transferfunktion bzw.  $f(u, \lambda) := |A(u, \lambda)|^2$  die Spektraldichte eines stationären  $AR(p)$ -Modells (siehe Satz 2.16), andererseits existiert aus dem selben Grund eine kausale Lösung dieses  $AR(p)$ -Prozesses (siehe Satz 2.13). Es bezeichne  $\{\varphi_j(u)\}_{j=0}^{\infty}$  die Koeffizienten dieser zu  $u$  gehörigen kausalen Lösung. Aus der kausalen  $MA(\infty)$ -Darstellung dieser  $AR(p)$ -Reihe folgt ebenfalls mit Satz 2.16, dass für ihre Transferfunktion auch

$$A(u, \lambda) = \frac{\sigma_\epsilon}{\sqrt{2\pi}} \sum_{j=0}^{\infty} \varphi_j(u) e^{-i\lambda j} \quad (3.22)$$

gilt. Wegen der beiden Darstellungen übertragen sich durch Koeffizientenvergleich in

$$1 = \left( \sum_{j=0}^{\infty} \varphi_j(u) e^{-i\lambda j} \right) \left( 1 - \sum_{k=1}^p \alpha_k(u) e^{-i\lambda k} \right) \quad (3.23)$$

die Glattheitseigenschaften der  $\alpha_k(\cdot)$ 's auf die  $\varphi_j(\cdot)$ 's.

Weil die Voraussetzungen von Satz 3.4 erfüllt sind, existiert wegen der kausalen Darstellung (3.13) des Prozesses die Spektraldarstellung

$$X_{t,n} = \int_{(-\pi, \pi]} e^{i\lambda t} A_{t,n}^0(\lambda) dZ_\epsilon(\lambda) \quad (3.24)$$

mit zeitabhängiger Transferfunktion

$$A_{t,n}^0(\lambda) := \frac{\sigma_\epsilon}{\sqrt{2\pi}} \sum_{j=0}^{\infty} \varphi_{j,t,n} e^{-i\lambda j}, \quad (3.25)$$

wobei  $Z_\epsilon(\lambda)$  ein stochastischer Prozess mit orthogonalen Zuwächsen (siehe Definition 2.14) auf  $(-\pi, \pi]$  ist und  $\bar{Z}_\epsilon(\lambda) = Z_\epsilon(-\lambda)$  gilt. Dahlhaus hat in Satz 3.2 gezeigt, dass  $\{X_{t,n}\}$  aus (3.12) lokal stationär ist mit zeitabhängiger Spektraldichte  $f(u, \lambda)$ . Damit folgt aus Definition 3.1 (ii), dass für eine Konstante  $K_1$

$$\sup_{t, \lambda} |A_{t,n}^0(\lambda) - A(\frac{t}{n}, \lambda)| \leq K_1 n^{-1} \quad (3.26)$$

gilt. Da für jedes  $u \in [0, 1]$  der approximierende stationäre  $MA(\infty)$ -Prozess eine Spektraldarstellung

$$\sum_{j=0}^{\infty} \varphi_j(u) \epsilon_{t-j} = \int_{(-\pi, \pi]} e^{i\lambda t} A(u, \lambda) dZ_\epsilon(\lambda) \quad (3.27)$$

besitzt (siehe Satz 2.16), folgt mit einer veränderlichen Konstanten  $C$  die Rechnung

$$\begin{aligned}
\sup_t |X_{t,n} - \tilde{X}_{t,n}| &= \sup_t \frac{\sigma_\epsilon}{\sqrt{2\pi}} \left| \int_{(-\pi, \pi]} e^{i\lambda t} [A_{t,n}^0(\lambda) - A(\frac{t}{n}, \lambda)] dZ_\epsilon(\lambda) \right| \\
&\leq \sup_t \frac{\sigma_\epsilon}{\sqrt{2\pi}} \int_{(-\pi, \pi]} |e^{i\lambda t}| \cdot |A_{t,n}^0(\lambda) - A(\frac{t}{n}, \lambda)| dZ_\epsilon(\lambda) \\
&\leq C \sup_{t, \lambda} |A_{t,n}^0(\lambda) - A(\frac{t}{n}, \lambda)| \int_{(-\pi, \pi]} dZ_\epsilon(\lambda) \\
&\leq C \cdot Z((-\pi, \pi]) \cdot n^{-1},
\end{aligned} \tag{3.28}$$

wobei  $Z((-\pi, \pi])$  eine Zufallsvariable ist, so dass die Behauptung (3.19) folgt.

Abschließend folgt noch die gleichmäßige Beschränktheit in (3.20) einfach aus der Tatsache, dass für alle  $u \in [0, 1]$  der kausale Prozess  $\sum_{j=0}^{\infty} \varphi_j(u) \epsilon_{t-j}$  stationär ist und daher sogar  $\sum_{j=0}^{\infty} |\varphi_j(u)| < \infty$  für alle  $u \in [0, 1]$  gilt.  $\square$

Bei den Aussagen in Lemma 3.6 wird jeweils das Supremum über alle  $t$  mit  $1 \leq t \leq n$  gebildet. Wegen der konstanten Fortsetzung der Koeffizientenfunktionen durch die Vorschrift  $\alpha_k(u) = \alpha_k(0)$  für  $u < 0$  gelten die obigen Resultate aber auch für Supremumbildung über alle  $t \in \mathbb{Z}$  mit  $t \leq n$ . Beachtet man nämlich, dass für  $t \leq 0$  die  $tvAR(p)$ -Modellgleichung

$$\begin{aligned}
X_{t,n} &= \sum_{k=1}^p \alpha_k(\frac{t}{n}) X_{t-k,n} + \epsilon_t \\
&= \sum_{k=1}^p \alpha_k(0) X_{t-k,n} + \epsilon_t
\end{aligned} \tag{3.29}$$

zu einer stationären  $AR(p)$ -Gleichung wird, wo die Koeffizienten konstant sind, so existiert hierfür eine kausale Lösung

$$X_{t,n} = \sum_{j=0}^{\infty} \varphi_j \epsilon_{t-j}, \quad t \leq 0 \tag{3.30}$$

mit von  $t$  und  $n$  unabhängigen Koeffizienten  $\varphi_j = \varphi_j(0)$ ,  $j \in \mathbb{N}_0$ . Der Prozess  $\{X_{t,n}\}$  ist also für  $t \leq 0$  stationär und daher ist auch die Darstellung  $X_{t,n} = \sum_{j=0}^{\infty} \varphi_j \epsilon_{t-j}$  für  $t \leq 0$  exakt und nicht nur approximativ. Es gilt folgendes Lemma.

**Lemma 3.7.**

Es gelten die Voraussetzungen von Satz 3.2. Dann existiert für  $\{X_{t,n}\}$  aus (3.12) eine Folge auf  $(0, 1)$  differenzierbarer Funktionen  $\{\varphi_j(\cdot) | \varphi_j : (-\infty, 1] \rightarrow \mathbb{R}\}_{j=0}^{\infty}$  und ein Prozess  $\{\tilde{X}_{t,n}\}_{t=-\infty}^n$  mit  $\tilde{X}_{t,n} := \sum_{j=0}^{\infty} \varphi_j(\frac{t}{n}) \epsilon_{t-j}$ , so dass

$$(i) \quad \sup_{t \in \mathbb{Z}, t \leq n} |X_{t,n} - \tilde{X}_{t,n}| = O_p\left(\frac{1}{n}\right), \quad (3.31)$$

$$(ii) \quad \sup_{t \in \mathbb{Z}, t \leq n} \sum_{j=0}^{\infty} |\varphi_j(\frac{t}{n})| < \infty \quad (3.32)$$

gilt. Man erhält diese Folge durch konstantes stetiges Fortsetzen  $\varphi_j(u) := \varphi_j(0)$  für  $u < 0$  der Funktionenfolge aus Lemma 3.6. Im Folgenden werden wir die Schreibweise  $t \in \mathbb{Z}, t \leq n$  in (3.31) und (3.32) unterdrücken und stattdessen wieder kurz  $t \leq n$  schreiben (siehe auch (3.15)).

Um die soeben bewiesenen Aussagen zu veranschaulichen, betrachten wir als Spezialfall das *tvAR*-Modell erster Ordnung etwas genauer. Im *tvAR*(1)-Modell hat die Matrix  $\mathbf{A}(u)$  aus Bemerkung 3.5 die Dimension eins und es gilt  $\mathbf{A}(u) = \alpha_1(u)$ , so dass für die Koeffizienten der kausalen Lösung  $X_{t,n} = \sum_{j=0}^{\infty} \varphi_{j,t,n} \epsilon_{t-j}$  aus Satz 3.4 die Darstellung

$$\varphi_{j,t,n} = \prod_{k=0}^{j-1} \alpha_1\left(\frac{t-k}{n}\right) \quad (3.33)$$

folgt. Andererseits haben die von  $t$  abhängigen Koeffizienten der stationären *MA*( $\infty$ )-Reihen  $\tilde{X}_{t,n} := \sum_{j=0}^{\infty} \varphi_j(\frac{t}{n}) \epsilon_{t-j}$  aus Lemma 3.7 die Gestalt

$$\varphi_j\left(\frac{t}{n}\right) = \alpha_1\left(\frac{t}{n}\right)^j. \quad (3.34)$$

Die erste Aussage von Lemma 3.6 bzw. von Lemma 3.7 bedeutet daher, dass die Gleichheit

$$\sum_{j=0}^{\infty} [\varphi_{j,t,n} - \varphi_j(\frac{t}{n})] \epsilon_{t-j} = 0 \quad (3.35)$$

für kein endliches  $n$  erfüllt ist, jedoch asymptotisch gilt.

# Kapitel 4

## Schätztheorie

Im vorangegangenen Kapitel haben wir ein zeitabhängiges  $AR$ -Modell motiviert und die Klasse der  $tvAR$ -Prozesse eingeführt, wobei wir in Abschnitt 3.4 bereits einige Eigenschaften bzgl. ihrer kausalen Lösbarkeit und ihrer approximativen Darstellbarkeit beweisen konnten.

In diesem Kapitel werden wir uns nun mit der Frage beschäftigen wie man aus einer gegebenen bzw. beobachteten Menge von Daten  $(\frac{1}{n}, X_1), \dots, (\frac{n}{n}, X_n)$ , denen unterstellt wird, dass sie aus einem  $tvAR$ -Modell der Ordnung  $p$  stammen, die Koeffizientenfunktionen  $\alpha_k : [0, 1] \rightarrow \mathbb{R}$ ,  $k = 1, \dots, p$  schätzen kann. Dazu werden wir zunächst auf die parametrische Kurvenschätzung (insbesondere lineare Regression) eingehen, bevor wir an deren Nachteilen die Idee der nicht-parametrischen Kurvenschätzung (*local linear smoothing*) erläutern und einen Kleinste-Quadrate-Kernschätzer für die  $p$ -dimensionale Funktion  $\underline{\alpha}(\cdot) := (\alpha_1(\cdot), \dots, \alpha_p(\cdot))$  herleiten werden. In den beiden Abschnitten 4.4 und 4.5 werden wir schließlich alternative Darstellungen für den Schätzer bzw. für den Schätzfehler beweisen.

### 4.1 Parametrische Kurvenschätzung

Die Kurvenschätzung allgemein befasst sich mit Daten der Form  $(X_1, Y_1), \dots, (X_n, Y_n)$ , also mit Paaren von Beobachtungen  $(X_k, Y_k)$ , und dem Schätzen eines *unbekannten* funktionalen Zusammenhangs zwischen  $X_k$  und  $Y_k$ . Beim parametrischen Ansatz unterstellt man diesen Beobachtungen eine funktionale Abhängigkeitsstruktur, wobei man *a-priori* annimmt, dass die tatsächliche Funktion  $m$ , welche die Abhängigkeit beschreibt, aus einer parametrischen Familie von Funktionen

$$\{m_\theta | \theta = (\theta_1, \dots, \theta_d) \in \Theta\} \tag{4.1}$$

mit Parameterraum  $\Theta \subset \mathbb{R}^d$  stammt. Bezeichnet man mit  $\hat{\theta} := (\hat{\theta}_1, \dots, \hat{\theta}_d)$  einen Schätzer für  $\theta = (\theta_1, \dots, \theta_d)$ , so definiert man durch  $\hat{m} := m_{\hat{\theta}}$  einen Schätzer für die gesuchte

Funktion  $m$ .

Beim wohlbekanntem Spezialfall der linearen Regression beispielsweise unterstellt man eine lineare Abhängigkeitsstruktur, indem man annimmt, dass die Beobachtungspaare  $(X_k, Y_k)$  das lineare Regressionsmodell

$$Y_k = \alpha_0 + \alpha_1 X_k + \epsilon_k, \quad k = 1, \dots, n \quad (4.2)$$

mit weißem Rauschen  $\epsilon_k \sim (0, \sigma_\epsilon^2)$  erfüllen. Mit der Methode der kleinsten Fehlerquadrate nach Gauß lässt sich für die beiden Parameter des Modells  $\alpha_0$  und  $\alpha_1$  der Schätzer

$$(\hat{\alpha}_0, \hat{\alpha}_1) := \min_{\alpha_0, \alpha_1} \sum_{k=1}^n (Y_k - \alpha_0 - \alpha_1 X_k)^2 \quad (4.3)$$

definieren, wobei das Minimierungsproblem die Lösung  $(\hat{\alpha}_0, \hat{\alpha}_1)$  mit

$$\hat{\alpha}_0 = \bar{Y} - \hat{\alpha}_1 \bar{X}, \quad (4.4)$$

$$\hat{\alpha}_1 = \frac{\sum_{k=1}^n (X_k - \bar{X})(Y_k - \bar{Y})}{\sum_{k=1}^n (X_k - \bar{X})^2} \quad (4.5)$$

besitzt und wie üblich  $\bar{X} := \frac{1}{n} \sum_{k=1}^n X_k$  bzw.  $\bar{Y} := \frac{1}{n} \sum_{k=1}^n Y_k$  die jeweiligen arithmetischen Mittel bezeichnen. Durch

$$\hat{m}(x) := \hat{\alpha}_0 + \hat{\alpha}_1 x \quad (4.6)$$

erhält man einen Schätzer für die Funktion  $m$ , welche den Zusammenhang von  $X_k$  und  $Y_k$  bestimmt.

In Abbildung 4.1 ist die Sinusfunktion  $\alpha(u) = 0,9 \sin(2\pi u)$ , welche bereits in Kapitel 3 zur Simulation eines  $tvAR(1)$ -Prozesses verwendet wurde (siehe Abbildung 3.6), aufgetragen. Würde man sich nun einen Datensatz der Form  $(\frac{1}{n}, \alpha(\frac{1}{n})), \dots, (\frac{n}{n}, \alpha(\frac{n}{n}))$  erzeugen und versuchen in soeben beschriebenem linearen Regressionsmodell (4.2) aus diesen die Sinusfunktion zu schätzen, so kann dieses Vorgehen offensichtlich kein zufriedenstellendes Ergebnis liefern, da sich eine Gerade niemals der deutlich nicht-linearen Funktion  $\alpha(u) = 0,9 \sin(2\pi u)$  anpassen kann.

Abhilfe könnte man sich verschaffen, indem man das lineare Regressionsmodell (4.2) verallgemeinert und anstatt einer Geraden ein Polynom  $q$ -ter Ordnung ansetzt, um die funktionale Abhängigkeitsstruktur zu schätzen. Man könnte etwa für ein  $q \in \mathbb{N}$ ,  $q \geq 2$  das polynomielle Regressionsmodell

$$Y_k = \alpha_0 + \alpha_1 X_k + \dots + \alpha_q X_k^q + \epsilon_k, \quad k = 1, \dots, n, \quad \epsilon_k \sim (0, \sigma_\epsilon) \quad (4.7)$$

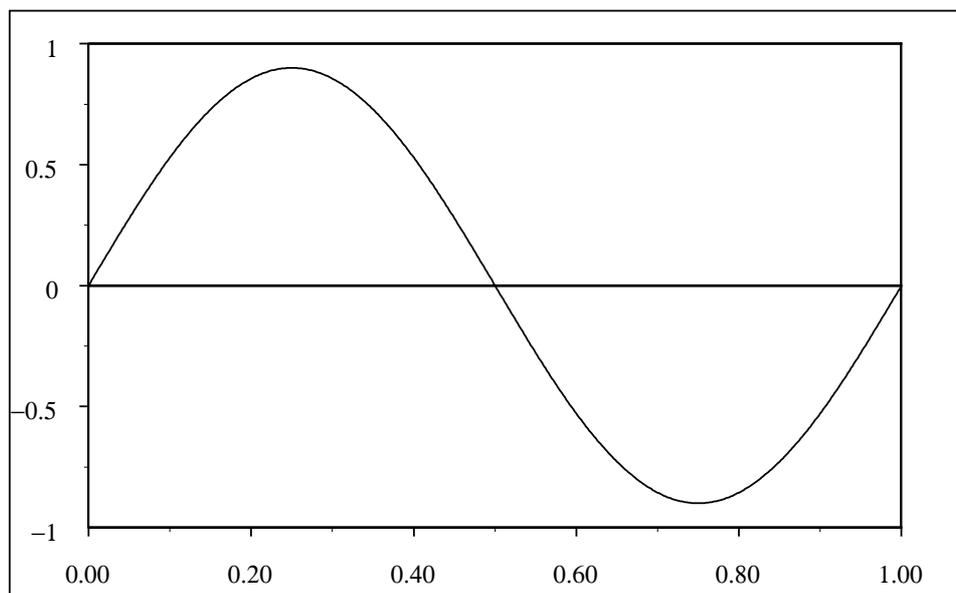


Abbildung 4.1: Verlauf der Sinusfunktion  $\alpha(u) = 0,9 \sin(2\pi u)$  im Intervall  $[0, 1]$ .

benutzen, um dem Kurvenverlauf der geschätzten Funktion eine größere Variabilität zu erlauben.

In realen Anwendungen wird jedoch der Fall oft so sein, dass man eine Menge von Beobachtungspaaren gegeben hat und einem keine oder nur sehr wenige Informationen über die zu wählende parametrische Familie  $\{m_\theta(\cdot) | \theta \in \Theta\}$  zur Verfügung stehen. Man kann also im Vorhinein selten eine begründete Annahme über die parametrische Struktur der anzusetzenden Funktionenfamilie treffen. Tatsächlich ist für Ökonomen die Art und Weise der veränderlichen Struktur an sich von direktem Interesse. Um daher das unterstellte Modell variabler zu gestalten, wie etwa beim Übergang vom linearen (4.2) zum polynomiellen Regressionsmodell (4.7), erkaufte man sich die höhere Flexibilität durch eine größere Komplexität des Modells und durch einen größeren Rechenaufwand. In (4.7) etwa müssen  $q + 1$  Parameter geschätzt werden und in (4.2) nur zwei, wobei die Frage nach der optimalen Ordnung  $q$  in der Regel unbeantwortet bleibt. Sind  $n$  Beobachtungspaare  $(X_1, Y_1), \dots, (X_n, Y_n)$  gegeben, so wäre etwa denkbar die Ordnung des Modells durch  $q := n - 1$  festzulegen, da man immer ein Polynom  $(n - 1)$ -ter Ordnung finden kann, auf welchem alle  $n$  Datenpaare liegen (insofern die  $X_k$ 's paarweise verschieden sind), wo also die Summe der quadratischen Abweichungen an diesen Stellen null ist. Dieses Vorgehen führt aber in der Regel zu einer Überparametrisierung, bei der eine glatte ruhig verlaufende Kurve als stark schwankende, um den tatsächlichen Kurvenverlauf oszillierende Funktion geschätzt wird.

In der parametrischen Kurvenschätzung steht man daher immer vor dem Problem, welche Funktionenfamilie man im Vorhinein ansetzen muss, damit man eine möglichst gute Regression erhält. Denn wählt man a-priori ein schlechtes Modell, dann können die Schätzer noch so gut sein und man bekommt für die Abhängigkeit von  $X_k$  und  $Y_k$  falsche bzw. ungenaue Resultate. Dieser Probleme zum Trotz findet der parametrische Ansatz vielerorts Anwendung, da er eine einfache Struktur hat und es meist leicht fällt die benötigten Parameter zu schätzen.

## 4.2 Nicht-parametrische Kurvenschätzung

Im vorangegangenen Abschnitt konnten wir sehen wo die Probleme und Nachteile, aber auch die Vorzüge eines parametrischen Ansatzes liegen. Dahlhaus etwa wählt in seinen Arbeiten (siehe [3], [6] und [7]) einen parametrischen Zugang, indem er beispielsweise in [7] an die zeitabhängige Spektraldichte  $f(u, \lambda)$  zur Zeit  $t$  ein endlich-dimensionales Modell mit Spektraldichte  $f_{\theta(t/n)}(\lambda)$  anpasst und die Parameterfunktion  $\theta(u)$ ,  $u \in [0, 1]$  schätzt.

Um die oben erläuterten Schwierigkeiten beim parametrischen Schätzen zu umgehen, führen wir in diesem Abschnitt das allgemeine nicht-parametrische Regressionsmodell ein und behandeln das Verfahren des *local linear smoothing* (lokal lineare Glättung). Das nicht-parametrische Regressionsmodell lässt sich als Verallgemeinerung von (4.2) bzw. (4.7) in der Form

$$Y_k = m(X_k) + \epsilon_k, \quad k = 1, \dots, n \quad (4.8)$$

mit  $\epsilon_k \sim (0, \sigma_\epsilon^2)$  darstellen, wobei die Zufallsvariablen  $X_k$  auf dem kompakten Intervall  $[0, 1]$  konzentriert seien und die Funktion  $m : [0, 1] \rightarrow \mathbb{R}$  als hinreichend glatt vorausgesetzt wird.

Das bereits angesprochene nicht-parametrische Verfahren des *local linear smoothing* (siehe auch Efromovich [9], S.334ff oder Fan und Gijbels [10], S.57ff) ermöglicht das Schätzen der unbekanntenen Funktion  $m$  aus (4.8) ohne vorher eine Annahme über die Gestalt der Funktionenfamilie zu machen aus der diese stammt. Die Idee dabei ist, unter der Annahme einer glatten Funktion  $m$ , die Daten *lokal* durch eine Gerade anzupassen. Denn ist  $m$  hinreichend glatt in  $x_0$ , so folgt aus der Taylorentwicklung von  $m$  um  $x_0$  die approximative Darstellung

$$m(x) \simeq m(x_0) + m'(x_0)(x - x_0) \quad (4.9)$$

für alle  $x$  aus einer kleinen Umgebung von  $x_0$  und rechtfertigt somit das Vorgehen lokal eine Gerade für die unbekanntene Kurve anzusetzen.

Im linearen Regressionsmodell (4.2) setzt man über die Zeit konstante Parameter  $\alpha_0$  und  $\alpha_1$  für die Steigung bzw. für den Ordinatenabschnitt an und schätzt diese dann *global* aus den verfügbaren Daten, d.h. man gewährt jedem Beobachtungspaar  $(X_k, Y_k)$  denselben Einfluss auf den Schätzer  $(\hat{\alpha}_0, \hat{\alpha}_1)$ . Beim *local linear smoothing* hingegen erlaubt man den Parametern  $\alpha_0$  und  $\alpha_1$  eine Abhängigkeit von der Zeit, indem man sie als veränderliche Funktionen  $\alpha_0 : [0, 1] \rightarrow \mathbb{R}$  und  $\alpha_1 : [0, 1] \rightarrow \mathbb{R}$  modelliert und die Gleichung

$$Y_k = \alpha_0(x) + \alpha_1(x)X_k + \epsilon_k, \quad k = 1, \dots, n \quad (4.10)$$

betrachtet. Die Funktionen  $\alpha_0$  und  $\alpha_1$  schätzt man dann an jeder Stelle  $x_0 \in [0, 1]$ , indem man *lokal* eine Gerade ansetzt und  $\alpha_0(x_0)$  und  $\alpha_1(x_0)$  nur aus den Beobachtungspaaren  $(X_k, Y_k)$  schätzt, für welche  $X_k$  *in der Nähe* von  $x_0$  liegt bzw. man gewährt diesen Daten mehr Einfluss auf den Schätzer  $(\hat{\alpha}_0(x_0), \hat{\alpha}_1(x_0))$  als den weiter entfernt liegenden Beobachtungen. Formal geschieht dies mit Hilfe einer nicht-negativen, glatten Gewichtsfunktion  $K$ , welche man als *Kern* oder *Kernfunktion* bezeichnet, falls sie über  $\mathbb{R}$  integriert eins ergibt, also falls

$$(K1) \quad \int K(x)dx = 1$$

gilt. Diese Annahme erlaubt zunächst, dass jede beliebige Wahrscheinlichkeitsdichte als Kernfunktion verwendet werden kann. Bevor wir zusätzliche Forderungen an die hier zugelassenen Kernfunktionen stellen, betrachten wir im folgenden Beispiel den Uniform-, den Epanechnikov- und den Biweight-Kern, wobei die beiden letzteren in den Abbildungen 4.2 und 4.3 dargestellt sind.

**Beispiel 4.1 (Kernfunktionen).**

Es bezeichne  $I_{[a,b]}(x)$  die Indikatorvariable, welche für  $x \in [a, b]$  gleich eins und sonst null ist.

- (i) Uniform-Kern  $K(x) = I_{[-1/2, 1/2]}(x)$
- (ii) Epanechnikov-Kern  $K(x) = \frac{3}{4}(1 - x^2)I_{[-1, 1]}(x)$  (Abbildung 4.2)
- (iii) Biweight-Kern  $K(x) = \frac{15}{16}(1 - x^2)^2 I_{[-1, 1]}(x)$  (Abbildung 4.3) □

Die Klasse der Funktionen, welche wir im Weiteren als Kernfunktionen zulassen wollen, schränken wir nun ein, indem wir von ihnen zusätzlich zu (K1) das Erfüllen der folgenden Annahmen fordern:

- (K2) Die nicht-negative Kernfunktion  $K$  sei stetig auf  $\mathbb{R}$  und symmetrisch. Weiter habe  $K$  einen kompakten Träger (d.h.  $K$  ist auf einer kompakten Menge größer null und sonst gleich null).

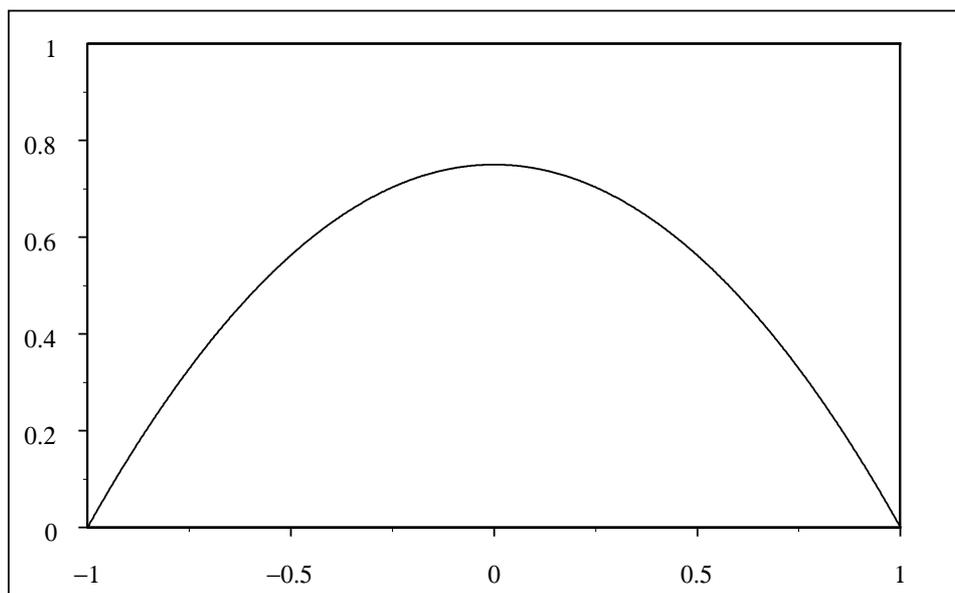


Abbildung 4.2: Epanechnikov-Kern  $K(x) = \frac{3}{4}(1 - x^2)I_{[-1,1]}(x)$ .

(K3) Der Kern  $K$  besitze auf  $\mathbb{R}$  beschränkte erste Ableitungen, d.h. es gelte

$$\|K\|_{\infty}^{(1)} := \sup_{x \in \mathbb{R}} |K'(x)| < \infty. \quad (4.11)$$

(K4) Es existieren die folgenden Integrale

$$\begin{aligned} \mu_K^2 &:= \int K(x)x^2 dx < \infty, \\ \|K\|_2^2 &:= \int K^2(x) dx < \infty, \\ \int K^2(x)x^{2l} dx &< \infty, \quad l = 1, 2, 3, \\ \int K^4(x) dx &< \infty. \end{aligned} \quad (4.12)$$

Im Vergleich der beiden Abbildungen 4.2 und 4.3 liegt der Vorteil des Biweight-Kerns darin, dass er an den Rändern *glatter* an null anschließt. Weil er auch die Bedingungen (K2) - (K4) erfüllt, werden wir ihn in Abschnitt 4.3 zum Schätzen der Koeffizientenfunktionen verwenden.

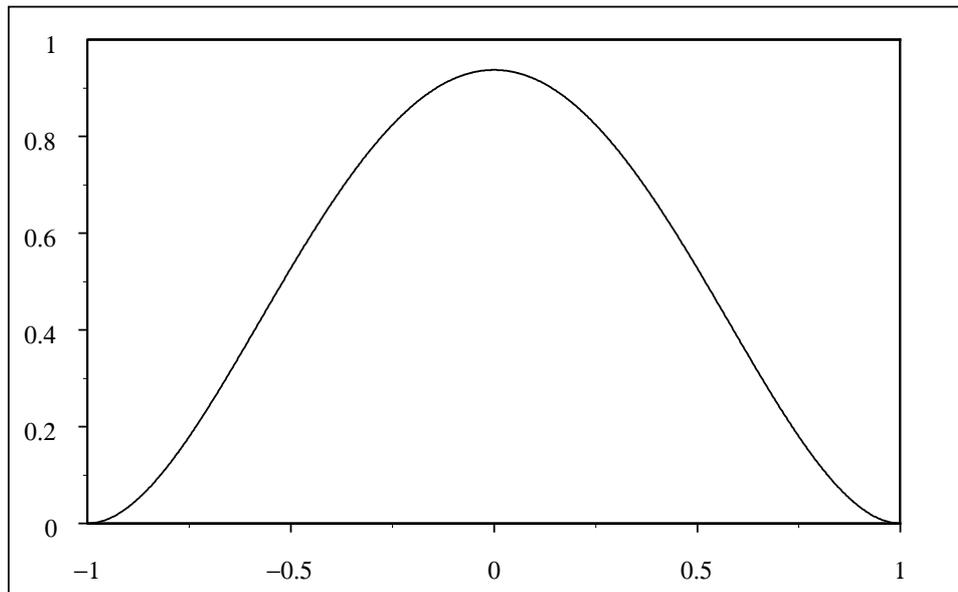


Abbildung 4.3: Biweight-Kern  $K(x) = \frac{15}{16}(1 - x^2)^2 I_{[-1,1]}(x)$ .

Mit Hilfe einer Kernfunktion  $K$  kann man dann den Schätzer für  $(\alpha_0(x), \alpha_1(x))$  als Minimalstelle der Summe der lokal gewichteten quadrierten Fehler gemäß

$$(\hat{\alpha}_0(x), \hat{\alpha}_1(x)) := \min_{\alpha_0(x), \alpha_1(x)} \sum_{k=1}^n (Y_k - \alpha_0(x) - \alpha_1(x)X_k)^2 K_h(x - X_k) \quad (4.13)$$

definieren (vgl. (4.3)), wobei  $K_h(\cdot) := \frac{1}{h}K(\frac{\cdot}{h})$  gesetzt wird. Dabei gewährleistet der hier eingefügte Parameter  $h$ , dass zur Berechnung des Schätzers  $(\hat{\alpha}_0(x), \hat{\alpha}_1(x))$  nur die Datenpaare  $(X_k, Y_k)$  benutzt werden, bei denen  $X_k$  nicht weiter als  $h$  von  $x$  entfernt liegt (falls  $K$  den kompakten Träger  $[-1, 1]$  hat). Daher nennt man  $h$  *Bandweite* oder auch *smoothing parameter*.

Um die Wirkungsweise von  $h$  in (4.13) besser zu verstehen, hilft es  $K$  als Wahrscheinlichkeitsdichte einer Zufallsvariablen  $Z$  zu betrachten, so dass  $K_h(\cdot) := \frac{1}{h}K(\frac{\cdot}{h})$  die Dichte der Zufallsvariablen  $hZ$  ist. Für  $h = 0,5$  ist in Abbildung 4.4 die Funktion  $K_h$  abgebildet, wobei  $K$  der Biweight-Kern aus Abbildung 4.3 ist und man deutlich sieht, wie die Bandweite  $h$  das Streuverhalten der zugehörigen Zufallsvariablen  $hZ$  beeinflusst. Darüber hinaus sei im Weiteren  $h = h(n)$  abhängig vom Stichprobenumfang  $n$  und konvergiere mit wachsendem  $n$  gegen null.

In praktischen Anwendungen spielt die Wahl der Bandweite eine entscheidende Rolle, da

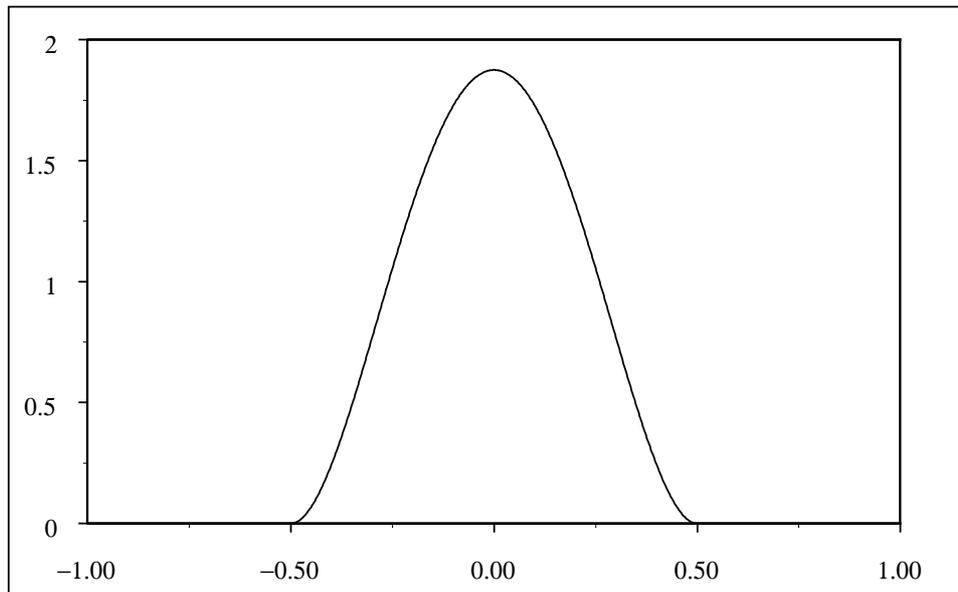


Abbildung 4.4: Biweight-Kern  $K_h(x) := \frac{1}{h}K(\frac{x}{h})$  mit Bandweite  $h = 0,5$  und  $K(x) = \frac{15}{16}(1 - x^2)^2 I_{[-1,1]}(x)$ .

sie einen großen Einfluss auf die Gestalt der geschätzten Kurve hat. Eine zu große Bandweite bedeutet eine Unterparametrisierung, was einen hohen Modellfehler zur Folge hat und eine zu klein gewählte Bandweite verursacht eine Überparametrisierung, was sich in einem stark oszillierenden Kurvenverlauf widerspiegelt. Wir wollen uns diesem Problem aber nicht weiter widmen und auf das Buch von Fan und Gijbels ([10], S.59f und Kapitel 4) verweisen, wo Verfahren zur Bandweitenwahl vorgestellt werden.

Fordert man von der zu schätzenden Funktion  $m$  aus (4.8), dass sie nicht nur einmal, sondern  $q$  mal stetig differenzierbar sei, so wäre auch ein lokal polynomieller Ansatz in Anlehnung an das parametrische Modell (4.7) gerechtfertigt und man könnte analog zu (4.10) bzw. (4.13) das Modell

$$Y_k = \alpha_0(x) + \alpha_1(x)X_k + \cdots + \alpha_q(x)X_k^q + \epsilon_k, \quad k = 1, \dots, n \quad (4.14)$$

sowie den Schätzer

$$\begin{aligned} & (\hat{\alpha}_0(x), \dots, \hat{\alpha}_q(x)) \\ & := \min_{\alpha_0(x), \dots, \alpha_q(x)} \sum_{k=1}^n (Y_k - \alpha_0(x) - \alpha_1(x)X_k \cdots - \alpha_q(x)X_k^q)^2 K_h(x - X_k) \end{aligned} \quad (4.15)$$

definieren.

Obwohl wir uns auf die lokal lineare Modellierung beschränken werden, da sie hinreichend gute Ergebnisse für die asymptotischen Betrachtungen in Kapitel 6 liefern wird und die Wahl der Ordnung  $q$  des nicht-parametrischen Regressionsmodells eine weitaus geringere Rolle spielt als etwa die Wahl der Bandweite  $h$  (vgl. Fan und Gijbels [10], S.59 und S.76ff), wollen wir trotzdem noch ein Resultat über die Matrixdarstellung des Schätzers (4.15) angeben. Dazu bezeichne

$$\mathbf{X} := \begin{pmatrix} 1 & (x - X_1) & \cdots & (x - X_1)^q \\ \vdots & \vdots & & \vdots \\ 1 & (x - X_n) & \cdots & (x - X_n)^q \end{pmatrix} \in \text{Mat}_{n,q+1}(\mathbb{R}) \quad (4.16)$$

die *Designmatrix* des Minimierungsproblems (4.15) sowie

$$\mathbf{W} := \begin{pmatrix} K_h(X_1 - x) & & 0 \\ & \ddots & \\ 0 & & K_h(X_n - x) \end{pmatrix} \in \text{Mat}_{n,n}(\mathbb{R}) \quad (4.17)$$

die *Gewichtsmatrix*. Setzt man weiter

$$\underline{Y} := \begin{pmatrix} Y_1 \\ \vdots \\ Y_n \end{pmatrix} \in \text{Mat}_{n,1}(\mathbb{R}), \quad \hat{\underline{\alpha}}(x) := \begin{pmatrix} \hat{\alpha}_0(x) \\ \vdots \\ \hat{\alpha}_q(x) \end{pmatrix} \in \text{Mat}_{q+1,1}(\mathbb{R}), \quad (4.18)$$

so ist das Minimierungsproblem in (4.15) äquivalent zu

$$\min_{\underline{\alpha}} (\underline{Y} - \mathbf{X}\underline{\alpha})^T \mathbf{W} (\underline{Y} - \mathbf{X}\underline{\alpha}) \quad (4.19)$$

mit  $\underline{\alpha} = (\alpha_0, \dots, \alpha_q)^T$ .

Aus der Theorie über lineare Ausgleichsprobleme (siehe etwa Deuffhard und Homann [8], S.66ff) ergibt sich, dass der Lösungsvektor des Minimierungsproblems (4.19) die Gestalt

$$\hat{\underline{\alpha}} = (\mathbf{X}^T \mathbf{W} \mathbf{X})^{-1} \mathbf{X}^T \mathbf{W} \underline{Y} \quad (4.20)$$

besitzt.

### 4.3 Kleinste-Quadrate-Kernschätzer im *tvAR*-Modell

Basierend auf der oben erläuterten Methode des *local linear smoothing* werden wir in diesem Abschnitt den Schätzer (4.13) auf das *tvAR*-Modell (3.5) zuschneiden und somit einen Kernschätzer für den  $p$ -dimensionalen Vektor der Koeffizientenfunktionen  $\underline{\alpha}(\cdot) :=$

$(\alpha_1(\cdot), \dots, \alpha_p(\cdot))^T$  herleiten.

Setzt man voraus, dass  $\alpha_k(\cdot)$  stetig differenzierbar in  $u \in (0, 1)$  ist und entwickelt  $\alpha_k(\frac{t}{n})$  nach Taylor um  $u$ , so erhält man ähnlich zu (4.9) die approximative Darstellung

$$\alpha_k\left(\frac{t}{n}\right) \simeq \alpha_k(u) + \alpha'_k(u)\left(\frac{t}{n} - u\right). \quad (4.21)$$

Die asymptotische Gleichheit in (4.21) ergibt sich zum einen aus der vorausgesetzten Stetigkeit von  $\alpha'_k(\cdot)$  in  $u$  und zum anderen, wenn man  $t := [nu]$  als den ganzzahligen Anteil von  $nu$  betrachtet und  $n \rightarrow \infty$  strebt, da dann  $[nu]/n \rightarrow u$  gilt. Genauso wie in Gleichung (4.9), wo die Approximation für  $x$  aus einer kleinen Umgebung von  $x_0$  verwendet wird, betrachtet man die Beziehung (4.21) für  $\frac{t}{n}$  nahe bei  $u$ .

Im Abschnitt 4.2 haben wir die nicht-parametrische Schätztheorie für den allgemeinen Fall vorliegender Beobachtungsdaten der Form  $(X_1, Y_1), \dots, (X_n, Y_n)$  behandelt. Aufgrund des durch die *tvAR*-Modellgleichung

$$X_{t,n} = \alpha_1\left(\frac{t}{n}\right)X_{t-1,n} + \dots + \alpha_p\left(\frac{t}{n}\right)X_{t-p,n} + \epsilon_t, \quad t = 1, \dots, n \quad (4.22)$$

hergestellten Zusammenhangs der  $X_{t,n}$  untereinander und der Skalierung des *tvAR*-Modells auf  $[0, 1]$  sowie dessen festen Designs (d.h. die Beobachtungen sind gleichmäßig und äquidistant im Intervall  $[0, 1]$  verteilt) sind dort Daten in Form eines Dreiecksschemas  $(\frac{1}{n}, X_{1,n}), \dots, (\frac{n}{n}, X_{n,n})$  gegeben, welche wir im Folgenden kurz mit  $\{X_{t,n}\}_{t=1}^n$  bezeichnen werden.

Man fährt nun ähnlich zur Herleitung von (4.13) fort, indem man die lokale Approximation (4.21) auf die Modellgleichung (4.22) anwendet und einen Kleinste-Quadrate-Kernschätzer für die Funktionen  $\alpha_k(\cdot)$  sowie deren erste Ableitungen  $\alpha'_k(\cdot)$  gemäß

$$\begin{aligned} (\widehat{\underline{\alpha}}(u), \widehat{\underline{\alpha}}'(u)) &:= (\widehat{\alpha}_1(u), \dots, \widehat{\alpha}_p(u), \widehat{\alpha}'_1(u), \dots, \widehat{\alpha}'_p(u)) \\ &:= \min_{a_{k,0}, a_{k,1}} \sum_{t=1}^n \left( X_{t,n} - \sum_{k=1}^p (a_{k,0} + a_{k,1}\left(\frac{t}{n} - u\right)) X_{t-k,n} \right)^2 K_h\left(\frac{t}{n} - u\right), \end{aligned} \quad (4.23)$$

definiert, wobei dann die Minimalstellen  $a_{k,0}$  und  $a_{k,1}$  Schätzer für  $\alpha_k(u)$  bzw. für  $\alpha'_k(u)$  sind.

Der Unterschied zu (4.13) besteht hier darin, dass man nicht nur eine, sondern  $p$  Funktionen simultan schätzt. Obwohl sich aus diesem Grund die Situation komplexer darstellt, ist es möglich wie am Ende von Abschnitt 4.2 auch den Kernschätzer für das *tvAR*-Modell (4.23) mit Hilfe von Matrizen anzugeben. Dazu führt man wieder zunächst einige

Bezeichnungen ein, indem man die Vektoren

$$\underline{X} := \begin{pmatrix} X_{1,n} \\ \vdots \\ X_{n,n} \end{pmatrix} \in \text{Mat}_{n,1}(\mathbb{R}), \quad \underline{X}_{t-1} := \begin{pmatrix} X_{t-1,n} \\ \vdots \\ X_{t-p,n} \end{pmatrix} \in \text{Mat}_{p,1}(\mathbb{R}) \quad (4.24)$$

sowie die Matrizen

$$\mathbf{X} := \begin{pmatrix} \underline{X}_0^T \\ \vdots \\ \underline{X}_{n-1}^T \end{pmatrix} = \begin{pmatrix} X_{0,n} & \cdots & X_{1-p,n} \\ \vdots & \ddots & \vdots \\ X_{n-1,n} & \cdots & X_{n-p,n} \end{pmatrix} \in \text{Mat}_{n,p}(\mathbb{R}), \quad (4.25)$$

$$\mathbf{D} := \begin{pmatrix} \frac{1}{n} - u & & 0 \\ & \ddots & \\ 0 & & \frac{n}{n} - u \end{pmatrix} \in \text{Mat}_{n,n}(\mathbb{R}), \quad (4.26)$$

$$\mathbf{W} := \begin{pmatrix} K_h(\frac{1}{n} - u) & & 0 \\ & \ddots & \\ 0 & & K_h(\frac{n}{n} - u) \end{pmatrix} \in \text{Mat}_{n,n}(\mathbb{R}) \quad (4.27)$$

definiert. Wie üblich bezeichnet hier und im Folgenden  $\mathbf{I}_{n \times n}$  die  $n \times n$ -Einheitsmatrix bzw.  $\mathbf{O}_{p \times p}$  die  $p \times p$ -Nullmatrix sowie  $I_n$  den  $n$ -dimensionalen Einsvektor und  $0_p$  den  $p$ -dimensionalen Nullvektor. Setzt man schließlich noch

$$\mathbf{Z} := (\mathbf{I}_{n \times n}, \mathbf{D}) \mathbf{X} := (\mathbf{X}, \mathbf{D}\mathbf{X}) \in \text{Mat}_{n,2p}(\mathbb{R}), \quad (4.28)$$

so erhält man das zu (4.23) äquivalente Minimierungsproblem

$$\min_{\underline{a}} (\underline{X} - \mathbf{Z}\underline{a})^T \mathbf{W} (\underline{X} - \mathbf{Z}\underline{a}) \quad (4.29)$$

mit  $\underline{a} := (a_{1,0}, \dots, a_{p,0}, a_{1,1}, \dots, a_{p,1})^T$ , dessen Lösung ähnlich zu (4.20) durch

$$(\hat{\underline{a}}(u), \hat{\underline{a}}'(u)) = (\mathbf{Z}^T \mathbf{W} \mathbf{Z})^{-1} (\mathbf{Z}^T \mathbf{W} \underline{X}) \quad (4.30)$$

gegeben ist.

Das Interesse konzentriert sich im Folgenden auf die Funktionen  $\alpha_k(u)$  selbst und nicht auf ihre ersten Ableitungen  $\alpha'_k(u)$ , so dass man  $(\hat{\underline{a}}(u), \hat{\underline{a}}'(u))$  noch mit der Matrix  $\mathbf{E}_0 := (\mathbf{I}_{p \times p}, \mathbf{O}_{p \times p}) \in \text{Mat}_{p,2p}(\mathbb{R})$  multipliziert und von nun an den Schätzer

$$\begin{aligned} \hat{\underline{a}}(u) &= (\hat{\alpha}_1(u), \dots, \hat{\alpha}_p(u))^T \\ &= \mathbf{E}_0 (\mathbf{Z}^T \mathbf{W} \mathbf{Z})^{-1} \mathbf{Z}^T \mathbf{W} \underline{X} \end{aligned} \quad (4.31)$$

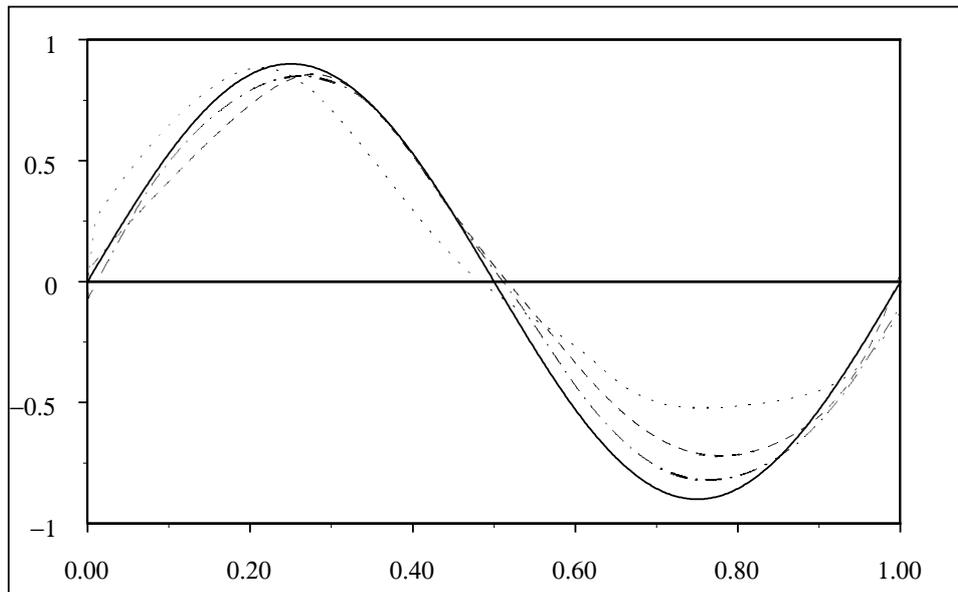


Abbildung 4.5: Vergleich der tatsächlichen Koeffizientenfunktion  $\alpha(u) = 0,9 \sin(2\pi u)$  (—) mit Schätzer (4.31) für  $n = 100$  (···),  $n = 500$  (- - -) und  $n = 1000$  (- · -).

betrachtet.

In den Abbildungen 4.5 und 4.6 sind jeweils die tatsächlichen Koeffizientenfunktionen  $\alpha(u) = 0,9 \sin(2\pi u)$  und  $\alpha(u) = 1,8u - 0,9$  der bereits in Kapitel 3 betrachteten  $tvAR(1)$ -Modelle (siehe auch die Abbildungen 3.6 und 3.7) sowie die mit der soeben hergeleiteten Formel (4.31) geschätzten Kurven für  $n = 100$ ,  $n = 500$  und  $n = 1000$  dargestellt, wobei als Bandweite  $h = n^{-1/5}$  benutzt wurde (vgl. auch die Voraussetzungen zu Theorem 6.11). In beiden Abbildung sieht man deutlich, dass die geschätzten Kurven die tatsächlichen Verläufe der Koeffizientenfunktionen  $\alpha(\cdot)$  sinnvoll annähern. Darüber hinaus ist eine Verbesserung des Schätzergebnisses mit wachsendem  $n$  erkennbar, so dass die hier angewandte nicht-parametrische Methode des *local linear smoothing* eine geeignete Methode darstellt, um eine Kurve zu schätzen ohne a-priori eine Annahme über ihre Gestalt treffen zu müssen.

Wie bei der Simulation von  $tvAR$ -Prozessen in Abschnitt 3.3 treten aber auch hier Probleme in der Nähe der linken Intervallgrenze auf (siehe Abbildung 4.6). Dies liegt daran, dass der Kernschätzer (4.31) aufgrund der Abhängigkeitsstruktur des  $tvAR$ -Modells nach links aus dem Intervall hinausreicht und Beobachtungen  $X_{1-p,n}, \dots, X_{0,n}$  benutzt, welche nicht gegeben sind. Diese Anfangswerte setzt man gleich null, da man keine Kenntnis über das Verhalten der zu schätzenden Funktion an der linken Intervallgrenze hat,

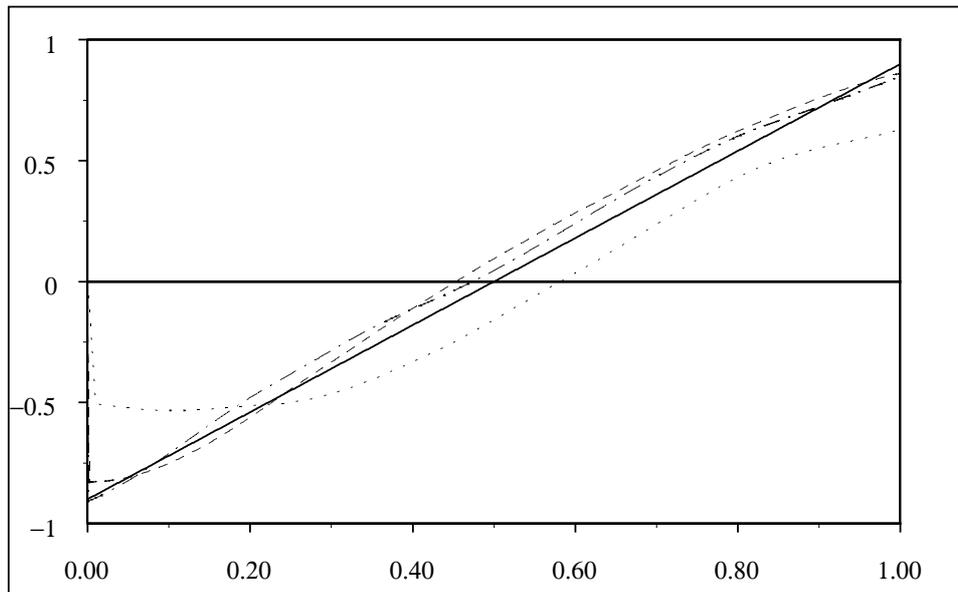


Abbildung 4.6: Vergleich der tatsächlichen Koeffizientenfunktion  $\alpha(u) = 1,8u - 0,9$  (—) mit Schätzer (4.31) für  $n = 100$  (···),  $n = 500$  (- - -) und  $n = 1000$  (- · -).

was jedoch in Abbildung 4.6 Probleme bereitet. Im Gegensatz zu Abbildung 4.5 ist dort  $\alpha(0) = -0,9 \neq 0$ , so dass sich die Schätzkurven mit wachsendem  $n$  nur auf dem links offenem Intervall  $(0, 1]$  der tatsächlichen Kurve nähern.

## 4.4 Darstellung des Kernschätzers als Funktion empirischer Momente

Die Berechnung des Kernschätzers (4.31) beinhaltet neben der Invertierung einer  $2p \times 2p$ -Matrix mehrfache Matrixmultiplikationen, was den Umgang mit ihm recht schwierig und unhandlich gestaltet. Aus diesem Grund ist es auch fast unmöglich asymptotische Untersuchungen, welche in Kapitel 6 folgen werden, für den Schätzer durchzuführen, ohne seine Asymptotik auf die einfacheren Ausdrücke zurückzuführen. Das Ziel dieses Abschnitts ist es daher eine Darstellung von  $\hat{\alpha}(u)$  anzugeben, welche vollständig mit empirischen Momenten der Art

$$\frac{1}{n} \sum_{t=1}^n K_h\left(\frac{t}{n} - u\right) \left[\frac{1}{h} \left(\frac{t}{n} - u\right)\right]^l X_{t-m,n} X_{t-q,n}, \quad l = 0, 1, 2, 3 \quad (4.32)$$

und

$$\frac{1}{n} \sum_{t=1}^n K_h\left(\frac{t}{n} - u\right) \left[ \frac{1}{h} \left( \frac{t}{n} - u \right) \right]^l X_{t-m,n} X_{t,n}, \quad l = 0, 1 \quad (4.33)$$

ausgedrückt werden kann, wobei  $m, q \in \{1, \dots, p\}$  gilt. Diese Summen werden sich als recht handlich erweisen und asymptotische Betrachtungen des Schätzfehlers  $\widehat{\underline{\alpha}}(u) - \underline{\alpha}(u)$  in Kapitel 6 ermöglichen.

**Lemma 4.2 (Alternative Darstellung von  $\widehat{\underline{\alpha}}(u)$ ).**

Für den Schätzer (4.31) existiert eine Darstellung, welche nur von empirischen Momenten der Form (4.32) und (4.33) abhängt, d.h. mit den Bezeichnungen

$$\mathbf{S}_n(u) := \begin{pmatrix} \mathbf{S}_{n,0}(u) & \mathbf{S}_{n,1}(u) \\ \mathbf{S}_{n,1}(u) & \mathbf{S}_{n,2}(u) \end{pmatrix} \in \text{Mat}_{2p,2p}(\mathbb{R}) \quad (4.34)$$

und

$$\underline{t}_n(u) := \begin{pmatrix} \underline{t}_{n,0}(u) \\ \underline{t}_{n,1}(u) \end{pmatrix} \in \text{Mat}_{2p,1}(\mathbb{R}), \quad (4.35)$$

wobei die Einträge der Matrix  $\mathbf{S}_n(u)$  für  $l = 0, 1, 2$  und die des Vektors  $\underline{t}_n(u)$  für  $l = 0, 1$  gemäß

$$\mathbf{S}_{n,l}(u) := \frac{1}{n} \sum_{t=1}^n K_h\left(\frac{t}{n} - u\right) \left[ \frac{1}{h} \left( \frac{t}{n} - u \right) \right]^l \underline{X}_{t-1} \underline{X}_{t-1}^T \in \text{Mat}_{p,p}(\mathbb{R}), \quad (4.36)$$

$$\underline{t}_{n,l}(u) := \frac{1}{n} \sum_{t=1}^n K_h\left(\frac{t}{n} - u\right) \left[ \frac{1}{h} \left( \frac{t}{n} - u \right) \right]^l \underline{X}_{t-1} X_{t,n} \in \text{Mat}_{p,1}(\mathbb{R}) \quad (4.37)$$

definiert sind, gilt die Gleichheit

$$\widehat{\underline{\alpha}}(u) = \mathbf{E}_0 \mathbf{S}_n^{-1}(u) \underline{t}_n(u). \quad (4.38)$$

*Beweis.*

Man führt den Beweis, indem man eine invertierbare Diagonalmatrix  $\mathbf{D}_h$  einführt und dann  $\mathbf{D}_h \cdot \mathbf{D}_h^{-1} = \mathbf{I}_{2p \times 2p}$  in die Darstellung des Schätzers  $\widehat{\underline{\alpha}}(u)$  aus (4.31) einfügt. Für  $h > 0$  bezeichne

$$\mathbf{D}_h := \begin{pmatrix} \mathbf{I}_{p \times p} & \mathbf{O}_{p \times p} \\ \mathbf{O}_{p \times p} & \mathbf{H} \end{pmatrix} \in \text{Mat}_{2p,2p}(\mathbb{R}) \quad (4.39)$$

mit

$$\mathbf{H} := \begin{pmatrix} \frac{1}{h} & & 0 \\ & \ddots & \\ 0 & & \frac{1}{h} \end{pmatrix} \in \text{Mat}_{p,p}(\mathbb{R}) \quad (4.40)$$

diese Diagonalmatrix. Fügt man nun die Einheitsmatrix in Form von  $\mathbf{D}_h \cdot \mathbf{D}_h^{-1}$  und  $(\mathbf{D}_h^T)^{-1} \cdot \mathbf{D}_h^T$  in den Schätzer  $\hat{\underline{\alpha}}(u)$  aus (4.31) ein und formt um, so erhält man

$$\begin{aligned}
\hat{\underline{\alpha}}(u) &= \mathbf{E}_0 (\mathbf{Z}^T \mathbf{W} \mathbf{Z})^{-1} \mathbf{Z}^T \mathbf{W} \underline{X} \\
&= \mathbf{E}_0 \mathbf{D}_h \mathbf{D}_h^{-1} (\mathbf{Z}^T \mathbf{W} \mathbf{Z})^{-1} (\mathbf{D}_h^T)^{-1} \mathbf{D}_h^T \mathbf{Z}^T \mathbf{W} \underline{X} \\
&= \mathbf{E}_0 \mathbf{D}_h (\mathbf{D}_h^T \mathbf{Z}^T \mathbf{W} \mathbf{Z} \mathbf{D}_h)^{-1} \mathbf{D}_h^T \mathbf{Z}^T \mathbf{W} \underline{X} \\
&= \mathbf{E}_0 \mathbf{D}_h ((\mathbf{Z} \mathbf{D}_h)^T \mathbf{W} (\mathbf{Z} \mathbf{D}_h))^{-1} (\mathbf{Z} \mathbf{D}_h)^T \mathbf{W} \underline{X}.
\end{aligned} \tag{4.41}$$

Multipliziert man noch mit  $(\frac{1}{n})^{-1} \cdot \frac{1}{n} = 1$  und beachtet, dass  $\mathbf{E}_0 \mathbf{D}_h = \mathbf{E}_0$  gilt, so ergibt sich

$$\hat{\underline{\alpha}}(u) = \mathbf{E}_0 \left( \frac{1}{n} (\mathbf{Z} \mathbf{D}_h)^T \mathbf{W} (\mathbf{Z} \mathbf{D}_h) \right)^{-1} \left( \frac{1}{n} (\mathbf{Z} \mathbf{D}_h)^T \mathbf{W} \underline{X} \right). \tag{4.42}$$

Ausmultiplizieren der beiden Ausdrücke in Klammern ergibt dann

$$\frac{1}{n} (\mathbf{Z} \mathbf{D}_h)^T \mathbf{W} (\mathbf{Z} \mathbf{D}_h) = \mathbf{S}_n(u) \tag{4.43}$$

und

$$\frac{1}{n} (\mathbf{Z} \mathbf{D}_h)^T \mathbf{W} \underline{X} = \underline{t}_n(u), \tag{4.44}$$

woraus die Behauptung folgt. □

## 4.5 Zerlegung des Schätzfehlers

Da es das Hauptanliegen dieser Arbeit ist, die asymptotische Verteilung des Schätzfehlers  $\hat{\underline{\alpha}}(u) - \underline{\alpha}(u)$  in Kapitel 6 herzuleiten, befassen wir uns hier näher mit seiner Darstellbarkeit. In Lemma 4.3 werden wir zeigen, dass sich der Schätzfehler in eine Summe bestehend aus einem Bias-, einem stochastischen und einem Restterm zerlegen lässt. In Kapitel 6 werden wir schließlich die Asymptotik dieser Ausdrücke untersuchen, wobei sich herausstellen wird, dass der Restterm nach Wahrscheinlichkeit gegen null und der stochastische Term in Verteilung gegen eine Normalverteilung konvergiert, wobei der Grenzwert nach Wahrscheinlichkeit des Bias-Terms den asymptotischen Bias des Schätzfehlers darstellt.

Der in Lemma 4.3 vorkommende Ausdruck  $\mathbf{S}_{n,3}(u)$  ist analog zur Definition von  $\mathbf{S}_{n,l}(u)$ ,  $l = 0, 1, 2$  in Lemma 4.2 zu verstehen und mit  $\underline{\alpha}''(u)$  ist der  $p$ -dimensional Vektor mit Einträgen  $\alpha_k''(u)$ ,  $k = 1, \dots, p$  gemeint.

**Lemma 4.3 (Zerlegung des Schätzfehlers  $\widehat{\underline{\alpha}}(u) - \underline{\alpha}(u)$ ).**

Für  $k = 1, \dots, p$  seien die Koeffizientenfunktionen  $\alpha_k(\cdot)$  zweimal stetig differenzierbar auf  $(0, 1)$  mit gleichmäßig beschränkten zweiten Ableitungen auf  $[0, 1]$ . Mit den Bezeichnungen

$$\underline{B}_n(u) := \frac{h^2}{2} \mathbf{E}_0 \mathbf{S}_n^{-1}(u) \begin{pmatrix} \mathbf{S}_{n,2}(u) \\ \mathbf{S}_{n,3}(u) \end{pmatrix} \underline{\alpha}''(u) \in \text{Mat}_{p,1}(\mathbb{R}), \quad (4.45)$$

$$\underline{T}_n(u) := \mathbf{E}_0 \mathbf{S}_n^{-1}(u) \underline{\tau}_n(u) \in \text{Mat}_{p,1}(\mathbb{R}), \quad (4.46)$$

$$\underline{R}_n(u) := \frac{h^2}{2} \mathbf{E}_0 \mathbf{S}_n^{-1}(u) \begin{pmatrix} \mathbf{S}_{n,2}(u) \\ \mathbf{S}_{n,3}(u) \end{pmatrix} o(1) \in \text{Mat}_{p,1}(\mathbb{R}) \quad (4.47)$$

für den Bias-, den stochastischen und den Restterm, wobei man

$$\underline{\tau}_n(u) := \begin{pmatrix} \underline{\tau}_{n,0}(u) \\ \underline{\tau}_{n,1}(u) \end{pmatrix} \in \text{Mat}_{2p,1}(\mathbb{R}) \quad (4.48)$$

und

$$\underline{\tau}_{n,l}(u) := \frac{1}{n} \sum_{t=1}^n K_h\left(\frac{t}{n} - u\right) \left[ \frac{1}{h} \left( \frac{t}{n} - u \right) \right]^l \underline{X}_{t-1} \epsilon_t \in \text{Mat}_{p,1}(\mathbb{R}), \quad l = 0, 1 \quad (4.49)$$

definiert, ergibt sich für den Schätzfehler

$$\widehat{\underline{\alpha}}(u) - \underline{\alpha}(u) = \underline{B}_n(u) + \underline{T}_n(u) + \underline{R}_n(u). \quad (4.50)$$

*Beweis.*

Es sei  $\mathbf{E}_0 := (\mathbf{I}_{p \times p}, \mathbf{O}_{p \times p})$  und  $\mathbf{E}_1 := (\mathbf{O}_{p \times p}, \mathbf{I}_{p \times p})$ . Mit den Bezeichnungen für  $\mathbf{W}$ ,  $\mathbf{D}$  und  $\mathbf{Z}$  aus (4.24) - (4.28) gelten die Beziehungen

$$\mathbf{E}_0 (\mathbf{Z}^T \mathbf{W} \mathbf{Z})^{-1} (\mathbf{Z}^T \mathbf{W} \mathbf{Z}) \mathbf{E}_0^T = \mathbf{I}_{p \times p}, \quad \mathbf{Z} \mathbf{E}_0^T = \mathbf{X}, \quad (4.51)$$

$$\mathbf{E}_0 (\mathbf{Z}^T \mathbf{W} \mathbf{Z})^{-1} (\mathbf{Z}^T \mathbf{W} \mathbf{Z}) \mathbf{E}_1^T = \mathbf{O}_{p \times p}, \quad \mathbf{Z} \mathbf{E}_1^T = \mathbf{D} \mathbf{X}. \quad (4.52)$$

Setzt man nun in (4.51) und (4.52) jeweils den zweiten Ausdruck in den ersten ein und addiert die beiden Gleichungen, so folgt

$$\mathbf{E}_0 (\mathbf{Z}^T \mathbf{W} \mathbf{Z})^{-1} \mathbf{Z}^T \mathbf{W} \mathbf{X} + \mathbf{E}_0 (\mathbf{Z}^T \mathbf{W} \mathbf{Z})^{-1} \mathbf{Z}^T \mathbf{W} \mathbf{D} \mathbf{X} = \mathbf{I}_{p \times p}. \quad (4.53)$$

Multipliziert man diese Gleichung mit  $\underline{\alpha}(u)$  und zieht diesen Ausdruck dann von der Darstellung (4.31) von  $\widehat{\underline{\alpha}}(u)$  ab, so erhält man für den Schätzfehler

$$\begin{aligned} & \widehat{\underline{\alpha}}(u) - \underline{\alpha}(u) \\ &= \mathbf{E}_0 (\mathbf{Z}^T \mathbf{W} \mathbf{Z})^{-1} \mathbf{Z}^T \mathbf{W} (\underline{X} - \mathbf{X} \underline{\alpha}(u) - \mathbf{D} \mathbf{X} \underline{\alpha}'(u)) \\ &= \mathbf{E}_0 \left( \frac{1}{n} (\mathbf{Z} \mathbf{D}_h)^T \mathbf{W} \mathbf{Z} \mathbf{D}_h \right)^{-1} \frac{1}{n} (\mathbf{Z} \mathbf{D}_h)^T \mathbf{W} (\underline{X} - \mathbf{X} \underline{\alpha}(u) - \mathbf{D} \mathbf{X} \underline{\alpha}'(u)), \quad (4.54) \end{aligned}$$

wobei wie im Beweis von Lemma 4.2 die Diagonalmatrix  $\mathbf{D}_h$  und der Bruch  $\frac{1}{n}$  eingefügt wurden. Betrachtet man nun den Ausdruck in Klammern ganz rechts aus (4.54), schreibt die  $tvAR$ -Modellgleichung (3.9) als  $X_{t,n} = \underline{X}_{t-1}^T \cdot \underline{\alpha}(\frac{t}{n}) + \epsilon_t$  und setzt diese Darstellung in die Komponenten des Vektors  $\underline{X}$  ein, so folgt

$$\begin{aligned} & \underline{X} - \mathbf{X}\underline{\alpha}(u) - \mathbf{D}\mathbf{X}\underline{\alpha}'(u) \\ = & \begin{pmatrix} \underline{X}_0^T \underline{\alpha}(\frac{1}{n}) \\ \vdots \\ \underline{X}_{n-1}^T \underline{\alpha}(\frac{n}{n}) \end{pmatrix} + \begin{pmatrix} \epsilon_1 \\ \vdots \\ \epsilon_n \end{pmatrix} - \begin{pmatrix} \underline{X}_0^T \underline{\alpha}(u) \\ \vdots \\ \underline{X}_{n-1}^T \underline{\alpha}(u) \end{pmatrix} - \mathbf{D} \begin{pmatrix} \underline{X}_0^T \underline{\alpha}'(u) \\ \vdots \\ \underline{X}_{n-1}^T \underline{\alpha}'(u) \end{pmatrix} \\ = & \begin{pmatrix} \underline{X}_0^T (\underline{\alpha}(\frac{1}{n}) - \underline{\alpha}(u) - (\frac{1}{n} - u)\underline{\alpha}'(u)) \\ \vdots \\ \underline{X}_{n-1}^T (\underline{\alpha}(\frac{n}{n}) - \underline{\alpha}(u) - (\frac{n}{n} - u)\underline{\alpha}'(u)) \end{pmatrix} + \begin{pmatrix} \epsilon_1 \\ \vdots \\ \epsilon_n \end{pmatrix}. \end{aligned} \quad (4.55)$$

Taylorentwicklung der Komponenten von  $\underline{\alpha}(\frac{1}{n}), \dots, \underline{\alpha}(\frac{n}{n})$  um  $u$  bis zur zweiten Ableitung ergibt

$$\underline{\alpha}(\frac{t}{n}) = \underline{\alpha}(u) + \frac{(t/n - u)}{1} \underline{\alpha}'(u) + \frac{(t/n - u)^2}{2} \underline{\alpha}''(u) + \frac{(t/n - u)^2}{2} (\underline{\alpha}''(\tilde{u}) - \underline{\alpha}''(u)), \quad (4.56)$$

wobei  $\tilde{u}$  eine Zwischenstelle von  $\frac{t}{n}$  und  $u$  bezeichnet. Beachtet man, dass wegen der vorausgesetzten Stetigkeit von  $\alpha_k''(\cdot)$  für  $k = 1, \dots, p$

$$\underline{\alpha}''(\tilde{u}) - \underline{\alpha}''(u) = o(1) \quad (4.57)$$

gilt, und setzt (4.56) in (4.55) ein, so erhält man insgesamt für den Schätzfehler  $\widehat{\underline{\alpha}}(u) - \underline{\alpha}(u)$  aus (4.54) die Darstellung

$$\begin{aligned} \widehat{\underline{\alpha}}(u) - \underline{\alpha}(u) &= \frac{h^2}{2} \mathbf{E}_0 \left( \frac{1}{n} (\mathbf{ZD}_h)^T \mathbf{WZD}_h \right)^{-1} \frac{1}{n} (\mathbf{ZD}_h)^T \mathbf{W} \frac{1}{h^2} \mathbf{D}^2 \mathbf{X} \underline{\alpha}''(u) \\ &+ \frac{h^2}{2} \mathbf{E}_0 \left( \frac{1}{n} (\mathbf{ZD}_h)^T \mathbf{WZD}_h \right)^{-1} \frac{1}{n} (\mathbf{ZD}_h)^T \mathbf{W} \frac{1}{h^2} \mathbf{D}^2 \mathbf{X} o(1) \\ &+ \mathbf{E}_0 \left( \frac{1}{n} (\mathbf{ZD}_h)^T \mathbf{WZD}_h \right)^{-1} \frac{1}{n} (\mathbf{ZD}_h)^T \mathbf{W} \underline{\epsilon}, \end{aligned} \quad (4.58)$$

wobei noch die beiden ersten Terme mit  $\frac{h^2}{h^2}$  multipliziert wurden und  $\underline{\epsilon} := (\epsilon_1, \dots, \epsilon_n)^T$  bezeichnet. Ausmultiplizieren ergibt schließlich

$$\frac{h^2}{2} \mathbf{E}_0 \left( \frac{1}{n} (\mathbf{ZD}_h)^T \mathbf{WZD}_h \right)^{-1} \frac{1}{n} (\mathbf{ZD}_h)^T \mathbf{W} \frac{1}{h^2} \mathbf{D}^2 \mathbf{X} \underline{\alpha}''(u) = \underline{B}_n(u), \quad (4.59)$$

$$\frac{h^2}{2} \mathbf{E}_0 \left( \frac{1}{n} (\mathbf{ZD}_h)^T \mathbf{WZD}_h \right)^{-1} \frac{1}{n} (\mathbf{ZD}_h)^T \mathbf{W} \frac{1}{h^2} \mathbf{D}^2 \mathbf{X} o(1) = \underline{R}_n(u) \quad (4.60)$$

sowie

$$\mathbf{E}_0 \left( \frac{1}{n} (\mathbf{ZD}_h)^T \mathbf{WZD}_h \right)^{-1} \frac{1}{n} (\mathbf{ZD}_h)^T \mathbf{W} \underline{\boldsymbol{\varepsilon}} = \underline{T}_n(u), \quad (4.61)$$

womit die Behauptung gezeigt ist. □

# Kapitel 5

## Die Beveridge-Nelson-Zerlegung

Vorbereitend für die asymptotischen Untersuchungen des Kernschätzers  $\hat{\alpha}(u)$  bzw. des Schätzfehlers  $\hat{\alpha}(u) - \underline{\alpha}(u)$  im nächsten Kapitel wollen wir hier eine Zerlegung linearer Prozesse ( $MA(\infty)$ -Reihen) der Form  $Y_t = \sum_{j=0}^{\infty} c_j \epsilon_{t-j}$  mit einem weißen Rauschen  $\epsilon_t \sim (0, \sigma_\epsilon^2)$ , basierend auf der Idee von Beveridge und Nelson (siehe [1]), vorstellen. Dazu werden wir unter gewissen Summierbarkeitsbedingungen an die reellen Koeffizienten  $\{c_j\}_{j \in \mathbb{N}_0}$  die *Beveridge-Nelson-Zerlegung* zweiter Ordnung ( $BN_2$ -Zerlegung) aus der Arbeit über die Asymptotik linearer Prozesse von Phillips und Solo (siehe [20]) angeben und diese auf den zeitabhängigen Fall verallgemeinern. Der Grundgedanke dabei ist eine rein algebraische, additive Zerlegung des linearen Filters in eine *long-run* (langfristige) und eine *transitory* (vergängliche) Komponente. Dabei sichern die Voraussetzungen an die Summierbarkeit der Folge  $\{c_j\}_{j \in \mathbb{N}_0}$  die Existenz der  $BN_2$ -Zerlegung.

In Abschnitt 5.1 befassen wir uns mit der Beveridge-Nelson-Zerlegung erster Ordnung ( $BN$ -Zerlegung) und zeigen ihre Vorzüge am Beispiel 5.2 auf, bevor wir ihre Anwendbarkeit auf den Fall zeitabhängiger Koeffizienten ausdehnen ( $tvBN$ -Zerlegung). Danach verallgemeinern wir in Abschnitt 5.2 die  $BN$ -Zerlegung zweiter Ordnung für  $Y_t Y_{t+h}$  von Phillips und Solo auf Ausdrücke der Form  $\tilde{X}_{t,n} \tilde{X}_{t+h,n}$ , wobei  $\tilde{X}_{t,n}$  gemäß Lemma 3.7 definiert ist ( $tvBN_2$ -Zerlegung).

### 5.1 $BN/tvBN$ -Zerlegung

Der Ansatz von Beveridge und Nelson basiert auf einer gänzlich algebraischen Zerlegung der Potenzreihe  $C(x) := \sum_{j=0}^{\infty} c_j x^j$ . Um die Zerlegung eines linearen Prozesses der Form

$$Y_t = \sum_{j=0}^{\infty} c_j \epsilon_{t-j} \tag{5.1}$$

auf die Zerlegung einer unendlichen Reihe zurückführen zu können, schreibt man mit Hilfe des *Lag-Operators*  $L$ , definiert durch  $L^j(X_t) := X_{t-j}$  für einen stochastischen Prozess  $\{X_t\}$ , obige Gleichung als

$$Y_t = C(L)\epsilon_t, \quad (5.2)$$

wobei  $C(L) = \sum_{j=0}^{\infty} c_j L^j$  bedeutet. In folgendem Lemma geben wir die besagte Beveridge-Nelson-Zerlegung erster Ordnung sowie eine Bedingung für die Folge  $\{c_j\}_{j \in \mathbb{N}_0}$  an, unter welcher die Koeffizienten  $\tilde{c}_j$  (siehe (5.5)) der vergänglichen Komponente der *BN-Zerlegung* absolut summierbar sind. Diese Summierbarkeit sichert dabei die Existenz der Zerlegung von  $Y_t$  (siehe Gleichung (5.10) in Beispiel 5.2).

**Lemma 5.1 (BN-Zerlegung).**

Sei  $L$  der *Lag-Operator* und bezeichne  $C(L) := \sum_{j=0}^{\infty} c_j L^j$  eine Potenzreihe in  $L$ . Erfüllen die Koeffizienten  $\{c_j\}_{j \in \mathbb{N}_0}$  die Bedingung

$$\sum_{j=0}^{\infty} j|c_j| < \infty, \quad (5.3)$$

dann existiert die *Beveridge-Nelson-Zerlegung* erster Ordnung

$$C(L) = C(1) - (1 - L)\tilde{C}(L) \quad (5.4)$$

mit

$$\tilde{C}(L) := \sum_{j=0}^{\infty} \tilde{c}_j L^j, \quad \tilde{c}_j := \sum_{k=j+1}^{\infty} c_k. \quad (5.5)$$

*Beweis.*

Startet man mit der rechten Seite von (5.4) und formt diese um, so ergibt sich die Zerlegung aus folgender Rechnung

$$\begin{aligned} C(1) - (1 - L)\tilde{C}(L) &= \sum_{j=0}^{\infty} c_j - (1 - L) \sum_{j=0}^{\infty} \tilde{c}_j L^j \\ &= \sum_{j=0}^{\infty} c_j - \sum_{j=0}^{\infty} \tilde{c}_j L^j + \sum_{j=0}^{\infty} \tilde{c}_j L^{j+1} \\ &= \sum_{j=0}^{\infty} c_j - \sum_{k=1}^{\infty} c_k - \sum_{j=1}^{\infty} \tilde{c}_j L^j + \sum_{j=1}^{\infty} \tilde{c}_{j-1} L^j \\ &= c_0 + \sum_{j=1}^{\infty} c_j L^j \\ &= C(L). \end{aligned} \quad (5.6)$$

Aufgrund der vorausgesetzten Summierbarkeit (5.3) folgt die Existenz schließlich aus

$$\sum_{j=0}^{\infty} |\tilde{c}_j| \leq \sum_{j=0}^{\infty} \sum_{k=j+1}^{\infty} |c_k| = \sum_{j=0}^{\infty} j|c_j| < \infty, \quad (5.7)$$

wobei hier Lemma 7.2 (i) benutzt wurde.  $\square$

Bevor wir die Aussage des obigen Lemmas auf den Fall linearer Prozesse  $\{\tilde{Y}_{t,n}\}_{t=1}^n$  mit zeitabhängigen Koeffizienten  $c_j(\frac{t}{n})$  der Form  $\tilde{Y}_{t,n} := \sum_{j=0}^{\infty} c_j(\frac{t}{n})\epsilon_{t-j}$  ausdehnen und somit erst die Anwendung der *BN*-Zerlegung auf die approximierenden *MA*( $\infty$ )-Reihen  $\tilde{X}_{t,n} = \sum_{j=0}^{\infty} \varphi_j(\frac{t}{n})\epsilon_{t-j}$  aus Kapitel 3, Lemma 3.7 ermöglichen, werden wir in einem Beispiel die Zerlegung (5.4) auf den Prozess  $Y_t = \sum_{j=0}^{\infty} c_j\epsilon_{t-j}$  anwenden, um daran ihre Vorzüge aufzuzeigen.

**Beispiel 5.2 (Anwendung der *BN*-Zerlegung).**

Sei  $\{Y_t\}$  ein linearer Prozess der Form (5.1) bzw. (5.2), welcher die Bedingungen

$$0 < |C(1)| < \infty \quad (5.8)$$

und

$$\sum_{j=1}^{\infty} j|c_j| < \infty \quad (5.9)$$

erfüllt. Wendet man nun die *BN*-Zerlegung (5.4) auf diesen Prozess an, so erhält man für ihn die Darstellung

$$Y_t = C(1)\epsilon_t - (1 - L)\tilde{C}(L)\epsilon_t. \quad (5.10)$$

Dabei kann der erste Term  $C(1)\epsilon_t$  als langfristig bezeichnet werden, da zu jeder Zeit  $t$  die unabhängig identisch verteilten Zufallsvariablen  $\epsilon_t$  denselben Faktor  $C(1)$  haben, und der zweite Term  $(1 - L)\tilde{C}(L)\epsilon_t$  stellt dann die vergängliche Komponente dar, welche eine Auslöschung, hervorgerufen durch den Faktor  $(1 - L)$ , beinhaltet. Die Gestalt (5.10) des linearen Prozesses ermöglicht es nun, seine Eigenschaften auf die des weißen Rauschens zurückzuführen. Beispielsweise lässt sich auf recht einfache und besonders direkte Art und Weise ein starkes Gesetz der großen Zahlen für  $Y_t$  aus einem für  $\epsilon_t$  folgern (siehe auch [20], Theorem 3.1), indem man die Gleichung (5.10) zu

$$\frac{1}{n} \sum_{t=1}^n Y_t = C(1) \frac{1}{n} \sum_{t=1}^n \epsilon_t + \frac{1}{n} \left( \tilde{C}(L)\epsilon_0 - \tilde{C}(L)\epsilon_n \right) \quad (5.11)$$

aufsummiert und beachtet, dass (5.9) die fast sichere Konvergenz der Terme  $\frac{1}{n}\tilde{C}(L)\epsilon_0$  und  $\frac{1}{n}\tilde{C}(L)\epsilon_n$  gegen null impliziert. Besonders hervorzuheben ist dabei die Vereinfachung des vergänglichen Anteils in (5.11), welche gerade durch die bereits erwähnte Auslöschung hervorgerufen wird. Mit Voraussetzung (5.8) folgt dann aus

$$\frac{1}{n} \sum_{t=1}^n \epsilon_t \xrightarrow{f.s.} 0 \quad (5.12)$$

ein starkes Gesetz der großen Zahlen für  $Y_t$ , also

$$\frac{1}{n} \sum_{t=1}^n Y_t \xrightarrow{f.s.} 0. \quad (5.13)$$

□

Wie weiter oben schon angekündigt, werden wir jetzt die Anwendbarkeit der  $BN$ -Zerlegung (5.4) aus Lemma 5.1 auf Prozesse der Form

$$\tilde{Y}_{t,n} = \sum_{j=0}^{\infty} c_j(\frac{t}{n}) \epsilon_{t-j} \quad (5.14)$$

erweitern. In Anlehnung an die  $tvAR$ -Modelle werden wir die Verallgemeinerung mit  $tvBN$ -Zerlegung (time varying  $BN$ -Zerlegung) bezeichnen.

**Lemma 5.3 ( $tvBN$ -Zerlegung).**

Sei  $L$  der Lag-Operator und für  $t \in \mathbb{Z}$ ,  $t \leq n$  bezeichne  $C(\frac{t}{n}; L) := \sum_{j=0}^{\infty} c_j(\frac{t}{n}) L^j$  eine Potenzreihe in  $L$ . Erfüllen die von  $t$  abhängigen Koeffizienten  $c_j(\frac{t}{n})$  die Bedingung

$$\sup_{t \leq n} \sum_{j=0}^{\infty} j |c_j(\frac{t}{n})| < \infty, \quad (5.15)$$

dann existiert die  $tvBN$ -Zerlegung erster Ordnung

$$\begin{aligned} C(\frac{t}{n}; L) &= C(\frac{t}{n}; 1) - \tilde{C}(\frac{t}{n}; L)(1 - L) \\ &= C(\frac{t}{n}; 1) - (1 - L)\tilde{C}(\frac{t}{n}; L) + [\tilde{C}(\frac{t}{n}; L) - \tilde{C}(\frac{t-1}{n}; L)]L \end{aligned} \quad (5.16)$$

mit

$$\tilde{C}(\frac{t}{n}; L) := \sum_{j=0}^{\infty} \tilde{c}_j(\frac{t}{n}) L^j, \quad \tilde{c}_j(\frac{t}{n}) := \sum_{k=j+1}^{\infty} c_k(\frac{t}{n}). \quad (5.17)$$

*Beweis.*

Zu beachten ist hierbei, dass der Lag-Operator  $L$  jetzt auch auf die Koeffizienten  $c_j(\frac{t}{n})$  wirkt und deshalb die erste Gleichheit nur gilt, weil man von rechts mit  $(1 - L)$  multipliziert. Sie ergibt sich jedoch völlig analog zu der von (5.4). Die zweite Darstellung der Potenzreihe  $C(\frac{t}{n}; L)$ , bei welcher der Faktor  $(1 - L)$  links von  $\tilde{C}(\frac{t}{n}; L)$  steht und der Restterm  $[\tilde{C}(\frac{t}{n}; L) - \tilde{C}(\frac{t-1}{n}; L)]L$  vorkommt, folgt aus der Rechnung

$$\begin{aligned}
& C(\frac{t}{n}; 1) - (1 - L)\tilde{C}(\frac{t}{n}; L) + [\tilde{C}(\frac{t}{n}; L) - \tilde{C}(\frac{t-1}{n}; L)]L \\
&= \sum_{j=0}^{\infty} c_j(\frac{t}{n}) - (1 - L) \sum_{j=0}^{\infty} \tilde{c}_j(\frac{t}{n})L^j + [\sum_{j=0}^{\infty} \tilde{c}_j(\frac{t}{n})L^j - \sum_{j=0}^{\infty} \tilde{c}_j(\frac{t-1}{n})L^j]L \quad (5.18) \\
&= \sum_{j=0}^{\infty} c_j(\frac{t}{n}) - \sum_{j=0}^{\infty} \tilde{c}_j(\frac{t}{n})L^j + \sum_{j=0}^{\infty} \tilde{c}_j(\frac{t-1}{n})L^{j+1} + \sum_{j=0}^{\infty} \tilde{c}_j(\frac{t}{n})L^{j+1} - \sum_{j=0}^{\infty} \tilde{c}_j(\frac{t-1}{n})L^{j+1} \\
&= C(\frac{t}{n}; 1) - \tilde{C}(\frac{t}{n}; L)(1 - L).
\end{aligned}$$

Die Existenz der Zerlegung folgt analog zu der aus Lemma 5.1, wobei man hier jedoch die gleichmäßige Beschränktheit (5.15) fordern muss.  $\square$

## 5.2 $BN/tvBN$ -Zerlegung zweiter Ordnung

Phillips und Solo haben in [20] basierend auf der in Lemma 5.1 eingeführten Zerlegung für lineare Prozesse  $Y_t = \sum_{j=0}^{\infty} c_j \epsilon_{t-j}$  eine  $BN$ -Zerlegung zweiter Ordnung ( $BN_2$ -Zerlegung) für Ausdrücke der Form  $Y_t Y_{t+h}$  hergeleitet. Mit  $c_j := 0$  für  $j < 0$  und den Bezeichnungen

$$f_j(L) := \sum_{k=0}^{\infty} c_k c_{k+j} L^k, \quad j \in \mathbb{Z}, \quad (5.19)$$

$$\tilde{f}_j(L) := \sum_{k=0}^{\infty} \left( \sum_{s=k+1}^{\infty} c_s c_{s+j} \right) L^k, \quad j \in \mathbb{Z} \quad (5.20)$$

lautet für  $Y_t Y_{t+h}$  die  $BN_2$ -Zerlegung

$$\begin{aligned}
Y_t Y_{t+h} &= f_h(1) \epsilon_t^2 + \sum_{r=1}^{\infty} (f_{h+r}(1) \epsilon_{t-r} \epsilon_t + f_{h-r}(1) \epsilon_{t+r} \epsilon_t) \\
&\quad - (1 - L) \tilde{f}_h(L) \epsilon_t^2 - (1 - L) \sum_{r=1}^{\infty} \left( \tilde{f}_{h+r}(L) \epsilon_{t-r} \epsilon_t + \tilde{f}_{h-r}(L) \epsilon_{t+r} \epsilon_t \right), \quad (5.21)
\end{aligned}$$

wobei hier wieder der erste Teil mit von  $t$  unabhängigen Vorfaktoren  $f_h(1)$ ,  $f_{h+r}(1)$  und  $f_{h-r}(1)$  als langfristig und der zweite als vergänglich bezeichnet werden kann. Analog zu

Beispiel 5.2 ermöglicht auch hier die recht einfache additive Struktur dieser Darstellung von  $Y_t Y_{t+h}$  einen direkten asymptotischen Zugang.

Das Ziel dieses Abschnitts ist es nun, diese Darstellung für  $Y_t Y_{t+h}$  mit Hilfe der *tvBN*-Zerlegung aus Lemma 5.3 auf Ausdrücke der Form  $\tilde{X}_{t,n} \tilde{X}_{t+h,n}$  mit  $\{\tilde{X}_{t,n}\}_{t=-\infty}^n$  aus Lemma 3.7 zu verallgemeinern. Dieses geschieht in folgendem Satz.

**Satz 5.4 (*tvBN-Zerlegung zweiter Ordnung*).**

Sei  $\{X_{t,n}\}$  ein *tvAR*-Prozess der Ordnung  $p$  gemäß (3.9), für welche die Voraussetzungen von Satz 3.2 erfüllt seien sowie  $\{\tilde{X}_{t,n}\}$  mit  $\tilde{X}_{t,n} = \sum_{j=0}^{\infty} \varphi_j(\frac{t}{n}) \epsilon_{t-j}$  der Prozess gleichmäßig approximierender *MA*( $\infty$ )-Reihen aus Lemma 3.7 mit

$$\sup_{t \leq n} |X_{t,n} - \tilde{X}_{t,n}| = O_p\left(\frac{1}{n}\right). \quad (5.22)$$

Weiter gelte für ein  $\delta > 0$  die Bedingung

$$(S1) \quad \sup_{t \leq n} \sum_{j=0}^{\infty} j^{2+\delta} |\varphi_j(\frac{t}{n})| < \infty. \quad (5.23)$$

Definiert man für  $h \in \{1-p, \dots, 0\}$ ,  $t \in \mathbb{Z}$  mit  $t \leq n$  und  $\varphi_j(\cdot) := 0$  für  $j < 0$  die *Lag*-Reihen

$$\phi_d\left(\frac{t}{n}, \frac{t+h}{n}; L\right) := \sum_{j=0}^{\infty} \varphi_j\left(\frac{t}{n}\right) \varphi_{j+d}\left(\frac{t+h}{n}\right) L^j, \quad (5.24)$$

$$\tilde{\phi}_d\left(\frac{t}{n}, \frac{t+h}{n}; L\right) := \sum_{j=0}^{\infty} \left( \sum_{s=j+1}^{\infty} \varphi_s\left(\frac{t}{n}\right) \varphi_{s+d}\left(\frac{t+h}{n}\right) \right) L^j \quad (5.25)$$

sowie den *Restterm*

$$\begin{aligned} R_d\left(\frac{t}{n}, \frac{t+h}{n}; L\right) \epsilon_t^2 &:= \left[ \tilde{\phi}_d\left(\frac{t}{n}, \frac{t+h}{n}; L\right) - \tilde{\phi}_d\left(\frac{t-1}{n}, \frac{t+h-1}{n}; L\right) \right] \epsilon_{t-1}^2 \\ &+ \sum_{r=1}^{\infty} \left[ \tilde{\phi}_{d+r}\left(\frac{t}{n}, \frac{t+h}{n}; L\right) - \tilde{\phi}_{d+r}\left(\frac{t-1}{n}, \frac{t+h-1}{n}; L\right) \right] \epsilon_{t-1} \epsilon_{t-r-1} \\ &+ \sum_{r=1}^{\infty} \left[ \tilde{\phi}_{d-r}\left(\frac{t}{n}, \frac{t+h}{n}; L\right) - \tilde{\phi}_{d-r}\left(\frac{t-1}{n}, \frac{t+h-1}{n}; L\right) \right] \epsilon_{t-1} \epsilon_{t+r-1}, \end{aligned} \quad (5.26)$$

so gilt die  $tvBN_2$ -Zerlegung für  $\tilde{X}_{t,n}\tilde{X}_{t+h,n}$  gemäß

$$\begin{aligned}
\tilde{X}_{t,n}\tilde{X}_{t+h,n} &= \phi_h\left(\frac{t}{n}, \frac{t+h}{n}; 1\right)\epsilon_t^2 \\
&+ \sum_{r=1}^{\infty} \left(\phi_{h+r}\left(\frac{t}{n}, \frac{t+h}{n}; 1\right)\epsilon_t\epsilon_{t-r} + \phi_{h-r}\left(\frac{t}{n}, \frac{t+h}{n}; 1\right)\epsilon_t\epsilon_{t+r}\right) \\
&- (1-L)\tilde{\phi}_h\left(\frac{t}{n}, \frac{t+h}{n}; L\right)\epsilon_t^2 \\
&- (1-L)\sum_{r=1}^{\infty} \left(\tilde{\phi}_{h+r}\left(\frac{t}{n}, \frac{t+h}{n}; L\right)\epsilon_t\epsilon_{t-r} + \tilde{\phi}_{h-r}\left(\frac{t}{n}, \frac{t+h}{n}; L\right)\epsilon_t\epsilon_{t+r}\right) \\
&+ R_h\left(\frac{t}{n}, \frac{t+h}{n}; L\right)\epsilon_t^2.
\end{aligned} \tag{5.27}$$

*Beweis.*

Es sei  $h \in \{1-p, \dots, 0\}$  und  $t \in \mathbb{Z}$  mit  $t \leq n$ . Beachtet man, dass das Supremum in (5.22) über  $t \in \mathbb{Z}$ ,  $t \leq n$  gebildet wird (siehe Lemma 3.7), so beinhaltet die dortige gleichmäßige Approximation auch  $\tilde{X}_{t+h,n}$ . Setzt man für  $\tilde{X}_{t,n}$  und  $\tilde{X}_{t+h,n}$  ein und ordnet die entstandene Doppelsomme nach den Indizes des weißen Rauschens, so erhält man zunächst

$$\begin{aligned}
\tilde{X}_{t,n}\tilde{X}_{t+h,n} &= \sum_{j=0}^{\infty} \varphi_j\left(\frac{t}{n}\right)\epsilon_{t-j} \sum_{r=0}^{\infty} \varphi_r\left(\frac{t+h}{n}\right)\epsilon_{t+h-r} \\
&= \sum_{j=0}^{\infty} \varphi_j\left(\frac{t}{n}\right)\varphi_{j+h}\left(\frac{t+h}{n}\right)\epsilon_{t-j}^2 + \sum_{j=0}^{\infty} \sum_{\substack{r=0 \\ r \neq h+j}}^{\infty} \varphi_j\left(\frac{t}{n}\right)\varphi_r\left(\frac{t+h}{n}\right)\epsilon_{t-j}\epsilon_{t+h-r}.
\end{aligned} \tag{5.28}$$

Eine Indexverschiebung der Summe über  $r$  und Vertauschen der Summationsreihenfolge ergibt dann

$$\begin{aligned}
\tilde{X}_{t,n}\tilde{X}_{t+h,n} &= \sum_{j=0}^{\infty} \varphi_j\left(\frac{t}{n}\right)\varphi_{j+h}\left(\frac{t+h}{n}\right)\epsilon_{t-j}^2 + \sum_{j=0}^{\infty} \sum_{\substack{r=-h-j \\ r \neq 0}}^{\infty} \varphi_j\left(\frac{t}{n}\right)\varphi_{j+h+r}\left(\frac{t+h}{n}\right)\epsilon_{t-j}\epsilon_{t-j-r} \\
&= \sum_{j=0}^{\infty} \varphi_j\left(\frac{t}{n}\right)\varphi_{j+h}\left(\frac{t+h}{n}\right)\epsilon_{t-j}^2 + \sum_{\substack{r=-\infty \\ r \neq 0}}^{\infty} \sum_{j=0}^{\infty} \varphi_j\left(\frac{t}{n}\right)\varphi_{j+h+r}\left(\frac{t+h}{n}\right)\epsilon_{t-j}\epsilon_{t-j-r} \\
&= \phi_h\left(\frac{t}{n}, \frac{t+h}{n}; L\right)\epsilon_t^2 + \sum_{r=1}^{\infty} \left(\phi_{h+r}\left(\frac{t}{n}, \frac{t+h}{n}; L\right)\epsilon_t\epsilon_{t-r} + \phi_{h-r}\left(\frac{t}{n}, \frac{t+h}{n}; L\right)\epsilon_t\epsilon_{t+r}\right),
\end{aligned} \tag{5.29}$$

wobei benutzt wurde, dass  $\varphi_j(\cdot) = 0$  für  $j < 0$  gilt. Anwenden der  $tvBN$ -Zerlegung (5.16) aus Lemma 5.3 auf die zeitabhängigen Lag-Reihen  $\phi_d\left(\frac{t}{n}, \frac{t+h}{n}; L\right)$  und Einsetzen von

$$\begin{aligned}
\phi_d\left(\frac{t}{n}, \frac{t+h}{n}; L\right) &= \phi_d\left(\frac{t}{n}, \frac{t+h}{n}; 1\right) - (1-L)\tilde{\phi}_d\left(\frac{t}{n}, \frac{t+h}{n}; L\right) \\
&+ \left[\tilde{\phi}_d\left(\frac{t}{n}, \frac{t+h}{n}; L\right) - \tilde{\phi}_d\left(\frac{t-1}{n}, \frac{t+h-1}{n}; L\right)\right] L,
\end{aligned} \tag{5.30}$$

für  $\phi_h(\frac{t}{n}, \frac{t+h}{n}; L)$ ,  $\phi_{h+r}(\frac{t}{n}, \frac{t+h}{n}; L)$  und  $\phi_{h-r}(\frac{t}{n}, \frac{t+h}{n}; L)$  ergibt dann die behauptete  $tvBN_2$ -Zerlegung, deren Existenz aus der Annahme (5.23) mit Lemma 7.3 aus Abschnitt 7.2 folgt.  $\square$

# Kapitel 6

## Asymptotische Betrachtungen

Das Ziel dieses Kapitels ist es, für  $u \in (0, 1)$  das asymptotische Verhalten des Schätzfehlers  $\widehat{\alpha}(u) - \underline{\alpha}(u)$  näher zu untersuchen und seine asymptotische Normalität zu zeigen. Dabei werden wir der unter *Phillips-Solo-device* (siehe Phillips und Solo [20]) bekannten Methode folgen, welche mit Hilfe der Beveridge-Nelson-Zerlegung asymptotische Resultate für lineare Prozesse liefert.

In Kapitel 4, Lemma 4.3 hatten wir bereits unter der Annahme, dass die Koeffizientenfunktionen  $\alpha_k(\cdot)$  zweimal stetig differenzierbar sind und gleichmäßig beschränkte zweite Ableitungen besitzen, eine Darstellung für den Schätzfehler als Summe aus Bias-Term, stochastischen Term und Restterm der Art

$$\widehat{\alpha}(u) - \underline{\alpha}(u) = \underline{B}_n(u) + \underline{T}_n(u) + \underline{R}_n(u) \quad (6.1)$$

bewiesen, welche nur noch von Matrizen bzw. Vektoren der Form

$$\mathbf{S}_{n,l}(u) = \frac{1}{n} \sum_{t=1}^n K_h\left(\frac{t}{n} - u\right) \left[ \frac{1}{h} \left( \frac{t}{n} - u \right) \right]^l \underline{X}_{t-1} \underline{X}_{t-1}^T \in \text{Mat}_{p,p}(\mathbb{R}), \quad l = 0, 1, 2, 3 \quad (6.2)$$

und

$$\underline{\tau}_{n,l} = \frac{1}{n} \sum_{t=1}^n K_h\left(\frac{t}{n} - u\right) \left[ \frac{1}{h} \left( \frac{t}{n} - u \right) \right]^l \underline{X}_{t-1} \epsilon_t \in \text{Mat}_{p,1}(\mathbb{R}), \quad l = 0, 1 \quad (6.3)$$

abhängt. Im weiteren Verlauf dieses Kapitels betrachten wir den Ausdruck

$$\sqrt{nh} (\widehat{\alpha}(u) - \underline{\alpha}(u)) = \sqrt{nh} \underline{B}_n(u) + \sqrt{nh} \underline{T}_n(u) + \sqrt{nh} \underline{R}_n(u) \quad (6.4)$$

und weisen nach, dass dieser in Verteilung gegen eine Normalverteilung konvergiert. Da alle drei Terme  $\underline{B}_n(u)$ ,  $\underline{T}_n(u)$  und  $\underline{R}_n(u)$  die Matrix  $\mathbf{S}_{n,l}(u)$  in irgendeiner Form enthalten, widmen wir uns als erstes in Abschnitt 6.1 der Asymptotik dieser Matrix, bevor wir in den

Abschnitten 6.2 und 6.3 auf das Grenzverhalten von  $\sqrt{nh}\underline{B}_n(u)$  und  $\sqrt{nh}\underline{R}_n(u)$  bzw. auf die asymptotische Normalität von  $\sqrt{nh}\underline{T}_n(u)$  eingehen. Um die Herleitung der asymptotischen Verteilung des Schätzfehlers  $\hat{\underline{\alpha}}(u) - \underline{\alpha}(u)$  zu vervollständigen, fassen wir schließlich in Theorem 6.11 die erzielten Ergebnisse zusammen und bestimmen die asymptotische Normalverteilung des Schätzfehlers  $\hat{\underline{\alpha}}(u) - \underline{\alpha}(u)$ .

Aufgrund der Übersichtlichkeit der in diesem Kapitel geführten Beweise sind einige benötigte Aussagen in den Anhang (siehe Kapitel 7) ausgelagert.

## 6.1 Asymptotik der empirischen Momente

Um das asymptotische Verhalten der empirischen Momente  $\mathbf{S}_{n,l}(u)$  zu berechnen, genügt es einen repräsentativen Eintrag der Matrix  $\mathbf{S}_{n,l}(u)$  herauszugreifen und die Untersuchungen für diesen fortzuführen. Die symmetrische  $p \times p$ -Matrix  $\underline{X}_{t-1}\underline{X}_{t-1}^T$  besteht aus Einträgen der Form  $X_{t-m,n}X_{t-q,n}$  mit  $m, q \in \{1, \dots, p\}$ . Ohne Einschränkungen gelte im Folgenden  $m \leq q$  (ansonsten vertauscht man  $X_{t-m,n}$  und  $X_{t-q,n}$ ). Mit  $d := m - q \leq 0$  bezeichne dann

$$\begin{aligned} S_{n,l,m,q}(u) &:= \frac{1}{n} \sum_{t=1}^n K_h\left(\frac{t}{n} - u\right) \left[\frac{1}{h} \left(\frac{t}{n} - u\right)\right]^l X_{t-m,n} X_{t-q,n} \\ &= \frac{1}{n} \sum_{t=1-m}^{n-m} K_h\left(\frac{t+m}{n} - u\right) \left[\frac{1}{h} \left(\frac{t+m}{n} - u\right)\right]^l X_{t,n} X_{t+d,n}. \end{aligned} \quad (6.5)$$

diesen Eintrag. Das nun folgende Lemma stellt die Beziehung von  $S_{n,l,m,q}(u)$  zu den approximierenden  $MA(\infty)$ -Reihen aus Lemma 3.7 her.

### Lemma 6.1.

Neben den Annahmen von Satz 3.2 gelten die Voraussetzungen (K1) - (K4) für den Kern  $K$  sowie  $h \rightarrow 0$  und  $nh^2 \rightarrow \infty$ . Sind für  $k = 1, \dots, p$  die Koeffizientenfunktionen  $\alpha_k(\cdot)$  zweimal stetig differenzierbar auf  $(0, 1)$  mit gleichmäßig beschränkten Ableitungen auf  $[0, 1]$ , dann gilt für  $l = 0, 1, 2, 3$

$$S_{n,l,m,q}(u) = \frac{1}{n} \sum_{t=1-m}^{n-m} K_h\left(\frac{t+m}{n} - u\right) \left[\frac{1}{h} \left(\frac{t+m}{n} - u\right)\right]^l \tilde{X}_{t,n} \tilde{X}_{t+d,n} + o_P(1), \quad (6.6)$$

wobei  $\{\tilde{X}_{t,n}\}_{t=-\infty}^n$  den linearen Prozess aus Lemma 3.7 bezeichnet.

*Beweis.*

Setzt man in (6.5) für  $X_{t,n}X_{t+d,n}$  den Ausdruck

$$\tilde{X}_{t,n} \tilde{X}_{t+d,n} + (X_{t,n} X_{t+d,n} - \tilde{X}_{t,n} \tilde{X}_{t+d,n}) \quad (6.7)$$

ein, schätzt die Differenz durch ihr Supremum über  $t \leq n$  ab und zieht es nach vorne, so ergibt sich

$$\begin{aligned} S_{n,l,m,q}(u) &= \frac{1}{n} \sum_{t=1-m}^{n-m} K_h\left(\frac{t+m}{n} - u\right) \left[ \frac{1}{h} \left( \frac{t+m}{n} - u \right) \right]^l \tilde{X}_{t,n} \tilde{X}_{t+d,n} \\ &+ \sup_{t \leq n} \left| X_{t,n} X_{t+d,n} - \tilde{X}_{t,n} \tilde{X}_{t+d,n} \right| \frac{1}{n} \sum_{t=1-m}^{n-m} K_h\left(\frac{t+m}{n} - u\right) \left[ \frac{1}{h} \left( \frac{t+m}{n} - u \right) \right]^l. \end{aligned} \quad (6.8)$$

Die Voraussetzungen von Lemma 3.7 sind erfüllt, so dass

$$\sup_{t \leq n} \left| X_{t,n} - \tilde{X}_{t,n} \right| = O_P\left(\frac{1}{n}\right) \quad (6.9)$$

gilt, woraus

$$\sup_{t \leq n} \left| X_{t,n} X_{t+d,n} - \tilde{X}_{t,n} \tilde{X}_{t+d,n} \right| = O_P\left(\frac{1}{n}\right) \quad (6.10)$$

folgt. Weiter ergibt sich mit  $h \rightarrow 0$  und  $nh^2 \rightarrow \infty$  die Konvergenz (siehe Beweisführung von Lemma 6.2)

$$\frac{1}{n} \sum_{t=1-m}^{n-m} K_h\left(\frac{t+m}{n} - u\right) \left[ \frac{1}{h} \left( \frac{t+m}{n} - u \right) \right]^l \xrightarrow{n \rightarrow \infty} \int K(s) s^l ds, \quad (6.11)$$

wobei die Integrale für  $l \in \{0, 2\}$  nach Voraussetzung (K4) endlich und für  $l \in \{1, 3\}$  gleich null sind. Die letztere Aussage folgt dabei aus der nach (K2) vorausgesetzten Symmetrie von  $K$ , welche impliziert, dass die Integranden  $K(s)s$  und  $K(s)s^3$  ungerade Funktionen sind und daher ihre Integrale über  $\mathbb{R}$  verschwinden. Damit ist die Behauptung des Lemmas gezeigt.  $\square$

Da es gelungen ist das Problem auf den eindimensionalen Fall zu reduzieren sowie den Eintrag  $S_{n,l,m,q}(u)$  mit Hilfe des Prozesses  $\{\tilde{X}_{t,n}\}_{t=-\infty}^n$  auszudrücken, ist es nun möglich unter der Voraussetzung

$$(S1) \quad \sup_{t \leq n} \sum_{j=0}^{\infty} j^{2+\delta} |\varphi_j\left(\frac{t}{n}\right)| < \infty \quad (6.12)$$

die  $tvBN_2$ -Zerlegung aus Kapitel 5, Satz 5.4 auf die Terme  $\tilde{X}_{t,n} \tilde{X}_{t+d,n}$  anzuwenden. Setzt man (5.27) in (6.6) ein und führt ähnlich zu Phillips und Solo (siehe [20], S.979) die Bezeichnungen

$$\epsilon_{t-1,d}^{\phi} := \sum_{r=1}^{\infty} \phi_{d+r}\left(\frac{t}{n}, \frac{t+d}{n}; 1\right) \epsilon_{t-r} \quad (6.13)$$

und

$$\epsilon_{t+1,d}^\phi := \sum_{r=1}^{\infty} \phi_{d-r}\left(\frac{t}{n}, \frac{t+d}{n}; 1\right) \epsilon_{t+r} \quad (6.14)$$

ein, so erhält man für  $S_{n,l,m,q}(u)$  die Zerlegung

$$S_{n,l,m,q}(u) = \frac{1}{n} \sum_{t=1-m}^{n-m} K_h\left(\frac{t+m}{n} - u\right) \left[\frac{1}{h} \left(\frac{t+m}{n} - u\right)\right]^l \phi_d\left(\frac{t}{n}, \frac{t+d}{n}; 1\right) \epsilon_t^2 \quad (6.15)$$

$$+ \frac{1}{n} \sum_{t=1-m}^{n-m} K_h\left(\frac{t+m}{n} - u\right) \left[\frac{1}{h} \left(\frac{t+m}{n} - u\right)\right]^l \epsilon_t \epsilon_{t-1,d}^\phi \quad (6.16)$$

$$+ \frac{1}{n} \sum_{t=1-m}^{n-m} K_h\left(\frac{t+m}{n} - u\right) \left[\frac{1}{h} \left(\frac{t+m}{n} - u\right)\right]^l \epsilon_t \epsilon_{t+1,d}^\phi \quad (6.17)$$

$$- \frac{1}{n} \sum_{t=1-m}^{n-m} K_h\left(\frac{t+m}{n} - u\right) \left[\frac{1}{h} \left(\frac{t+m}{n} - u\right)\right]^l (1-L) \tilde{\phi}_d\left(\frac{t}{n}, \frac{t+d}{n}; L\right) \epsilon_t^2 \quad (6.18)$$

$$- \frac{1}{n} \sum_{t=1-m}^{n-m} K_h\left(\frac{t+m}{n} - u\right) \left[\frac{1}{h} \left(\frac{t+m}{n} - u\right)\right]^l (1-L) \sum_{r=1}^{\infty} \tilde{\phi}_{d+r}\left(\frac{t}{n}, \frac{t+d}{n}; L\right) \epsilon_t \epsilon_{t-r} \quad (6.19)$$

$$- \frac{1}{n} \sum_{t=1-m}^{n-m} K_h\left(\frac{t+m}{n} - u\right) \left[\frac{1}{h} \left(\frac{t+m}{n} - u\right)\right]^l (1-L) \sum_{r=1}^{\infty} \tilde{\phi}_{d-r}\left(\frac{t}{n}, \frac{t+d}{n}; L\right) \epsilon_t \epsilon_{t+r} \quad (6.20)$$

$$+ \frac{1}{n} \sum_{t=1-m}^{n-m} K_h\left(\frac{t+m}{n} - u\right) \left[\frac{1}{h} \left(\frac{t+m}{n} - u\right)\right]^l R_d\left(\frac{t}{n}, \frac{t+d}{n}; L\right) \epsilon_t^2 \quad (6.21)$$

$$+ o_P(1). \quad (6.22)$$

Mit Hilfe dieser Zerlegung sind wir nun in der Lage, das asymptotische Verhalten von  $S_{n,l,m,q}(u)$  und damit auch von  $\mathbf{S}_{n,l}(u)$  herzuleiten. Dazu untersucht man jede Zeile, also jeden Summanden, obiger Zerlegung einzeln auf seine Asymptotik. Dabei ergibt sich, dass ausschließlich der erste Term

$$\frac{1}{n} \sum_{t=1-m}^{n-m} K_h\left(\frac{t+m}{n} - u\right) \left[\frac{1}{h} \left(\frac{t+m}{n} - u\right)\right]^l \phi_d\left(\frac{t}{n}, \frac{t+d}{n}; 1\right) \epsilon_t^2 \quad (6.23)$$

asymptotisch einen Beitrag ungleich null zum Grenzwert nach Wahrscheinlichkeit von  $S_{n,l,m,q}(u)$  liefert und alle anderen Summanden stochastisch gegen null konvergieren und damit verschwinden. Weil die Beweisführung beim Nachweis der soeben behaupteten Asymptotik einiger Summanden sehr ähnlich ist, fassen wir diese jeweils in einem Lemma zusammen, bevor wir die Aussagen zusammentragen und den Grenzwert nach Wahrscheinlichkeit der  $p \times p$ -Matrix  $\mathbf{S}_{n,l}(u)$  in Satz 6.6 angeben.

**Lemma 6.2.**

Neben den Annahmen von Satz 3.2 gelten die Voraussetzungen (K1) - (K4) für den Kern  $K$  und für  $k = 1, \dots, p$  seien die Koeffizientenfunktionen  $\alpha_k(\cdot)$  zweimal stetig differenzierbar auf  $(0, 1)$  mit gleichmäßig beschränkten Ableitungen auf  $[0, 1]$ . Falls für ein  $\delta > 0$  die Bedingung (S1) gilt sowie  $h \rightarrow 0$  und  $nh^2 \rightarrow \infty$ , dann gilt für  $l = 0, 1, 2, 3$  die Konvergenz

$$\begin{aligned} & \frac{1}{n} \sum_{t=1-m}^{n-m} K_h\left(\frac{t+m}{n} - u\right) \left[ \frac{1}{h} \left( \frac{t+m}{n} - u \right) \right]^l \phi_d\left(\frac{t}{n}, \frac{t+d}{n}; 1\right) \epsilon_t^2 \\ & \xrightarrow{P} \sigma_\epsilon^2 \phi_d(u, u; 1) \int K(s) s^l ds. \end{aligned} \quad (6.24)$$

*Beweis.*

Der Kern  $K$  hat nach Voraussetzung einen kompakten und damit auch einen beschränkten Träger. Daher existiert eine Konstante  $T$  mit  $K(s) = 0$  für alle  $s \in [-T, T]^c$ . Sei nun  $n$  so groß, dass

$$\frac{1}{h} \left( \frac{1}{n} - u \right) < -T \quad \text{und} \quad T < \frac{1}{h} \left( \frac{n}{n} - u \right) \quad (6.25)$$

gilt. Dies ist möglich, da wegen  $u \in (0, 1)$  der Term  $\frac{1}{h}(\frac{1}{n} - u)$  gegen  $-\infty$  und  $\frac{1}{h}(\frac{n}{n} - u)$  gegen  $+\infty$  strebt. Man braucht dann nur noch über alle  $t \in \mathbb{T}_n$  mit

$$\mathbb{T}_n := \{t \in \mathbb{Z} \mid -T < \frac{1}{h} \left( \frac{t}{n} - u \right) < T\} \quad (6.26)$$

$$= \{t \in \mathbb{Z} \mid n(-Th + u) < t < n(Th + u)\} \quad (6.27)$$

summieren, d.h. es gilt

$$\frac{1}{n} \sum_{t=1-m}^{n-m} K_h\left(\frac{t+m}{n} - u\right) \left[ \frac{1}{h} \left( \frac{t+m}{n} - u \right) \right]^l \phi_d\left(\frac{t}{n}, \frac{t+d}{n}; 1\right) \epsilon_t^2 \quad (6.28)$$

$$= \frac{1}{nh} \sum_{t=1}^n K\left(\frac{1}{h} \left( \frac{t}{n} - u \right)\right) \left[ \frac{1}{h} \left( \frac{t}{n} - u \right) \right]^l \phi_d\left(\frac{t-m}{n}, \frac{t-q}{n}; 1\right) \epsilon_{t-m}^2 \quad (6.29)$$

$$= \frac{1}{nh} \sum_{t \in \mathbb{T}_n} K\left(\frac{1}{h} \left( \frac{t}{n} - u \right)\right) \left[ \frac{1}{h} \left( \frac{t}{n} - u \right) \right]^l \phi_d\left(\frac{t-m}{n}, \frac{t-q}{n}; 1\right) \epsilon_{t-m}^2. \quad (6.30)$$

Zum einen ist  $K(s)s^l$  als Komposition stetiger Funktionen wieder stetig, zum anderen ist das Intervall  $[-T, T]$  äquidistant in Teilintervalle zerlegt, deren Breite sich aus

$$\frac{1}{h} \left( \frac{t+1}{n} - u \right) - \frac{1}{h} \left( \frac{t}{n} - u \right) = \frac{1}{nh} \quad (6.31)$$

ergibt und nach Voraussetzung gegen null strebt. Damit folgt die Konvergenz

$$\begin{aligned} \frac{1}{n} \sum_{t \in \mathbb{T}} K_h\left(\frac{t}{n} - u\right) \left[ \frac{1}{h} \left( \frac{t}{n} - u \right) \right]^l \phi_d\left(\frac{t-m}{n}, \frac{t-q}{n}; 1\right) &\xrightarrow{n \rightarrow \infty} \phi_d(u, u; 1) \int_{-T}^T K(s) s^l ds \\ &= \phi_d(u, u; 1) \int K(s) s^l ds. \end{aligned} \quad (6.32)$$

Dabei sorgt der kompakte Träger von  $K$  und der Faktor  $\frac{1}{h}$  in  $K\left(\frac{1}{h}\left(\frac{t}{n} - u\right)\right)$  dafür, dass  $K\left(\frac{1}{h}\left(\frac{t}{n} - u\right)\right)$  nur für  $t \in \mathbb{T}_n$  echt größer null ist, also asymptotisch nur für die  $t$ , welche für alle  $h \geq 0$  die Ungleichungen

$$-Th < \frac{t}{n} - u < Th \quad (6.33)$$

erfüllen, womit sich der Faktor  $\phi_d(u, u; 1)$  erklärt. Mit einem Gesetz der großen Zahlen folgt aus  $E[\epsilon_t^2] = \sigma_\epsilon^2$  insgesamt

$$\begin{aligned} &\frac{1}{n} \sum_{t=1-m}^{n-m} K_h\left(\frac{t+m}{n} - u\right) \left[ \frac{1}{h} \left( \frac{t+m}{n} - u \right) \right]^l \phi_d\left(\frac{t}{n}, \frac{t+d}{n}; 1\right) \epsilon_t^2 \\ &\xrightarrow{P} \sigma_\epsilon^2 \phi_d(u, u; 1) \int K(s) s^l ds. \end{aligned} \quad (6.34)$$

□

**Lemma 6.3.**

Neben den Annahmen von Satz 3.2 gelten die Voraussetzungen (K1) - (K4) für den Kern  $K$  und für  $k = 1, \dots, p$  seien die Koeffizientenfunktionen  $\alpha_k(\cdot)$  zweimal stetig differenzierbar auf  $(0, 1)$  mit gleichmäßig beschränkten Ableitungen auf  $[0, 1]$ . Falls für ein  $\delta > 0$  die Bedingung (S1) gilt sowie  $h \rightarrow 0$  und  $nh^2 \rightarrow \infty$ , dann konvergieren die Summanden (6.16) und (6.17) stochastisch gegen null, d.h für  $l = 0, 1, 2, 3$  gilt

$$(i) \quad \frac{1}{n} \sum_{t=1-m}^{n-m} K_h\left(\frac{t+m}{n} - u\right) \left[ \frac{1}{h} \left( \frac{t+m}{n} - u \right) \right]^l \epsilon_t \epsilon_{t-1,d}^\phi = o_P(1), \quad (6.35)$$

$$(ii) \quad \frac{1}{n} \sum_{t=1-m}^{n-m} K_h\left(\frac{t+m}{n} - u\right) \left[ \frac{1}{h} \left( \frac{t+m}{n} - u \right) \right]^l \epsilon_t \epsilon_{t+1,d}^\phi = o_P(1). \quad (6.36)$$

*Beweis.*

zu (i): Nach Lemma 2.26 (ii) genügt es für den ersten Ausdruck, wenn man zeigt, dass sein zweites Moment gegen null konvergiert. Per Definition fängt die Summe  $\epsilon_{t-1,d}^\phi$  erst

bei eins an (siehe (6.13)), weshalb  $E[\epsilon_t \epsilon_{t-1,d}^\phi \epsilon_s \epsilon_{s-1,d}^\phi] = 0$  für  $s \neq t$  gilt. Damit ergibt sich für das zweite Moment

$$\begin{aligned} & E \left[ \left( \frac{1}{n} \sum_{t=1-m}^{n-m} K_h \left( \frac{t+m}{n} - u \right) \left[ \frac{1}{h} \left( \frac{t+m}{n} - u \right) \right]^l \epsilon_t \epsilon_{t-1,d}^\phi \right)^2 \right] \\ &= \frac{1}{n^2 h^2} \sum_{t=1-m}^{n-m} K^2 \left( \frac{1}{h} \left( \frac{t+m}{n} - u \right) \right) \left[ \frac{1}{h} \left( \frac{t+m}{n} - u \right) \right]^{2l} E \left[ \epsilon_t^2 (\epsilon_{t-1,d}^\phi)^2 \right]. \end{aligned} \quad (6.37)$$

Der letzte Erwartungswert lässt sich aufgrund der stochastischen Unabhängigkeit von  $\epsilon_t^2$  und  $(\epsilon_{t-1,d}^\phi)^2$  multiplikativ auseinanderziehen. Für den ersten Anteil gilt  $E[\epsilon_t^2] = \sigma_\epsilon^2 < \infty$  und wegen  $E[\epsilon_{t-r} \epsilon_{t-s}] = 0$  für  $t \neq s$  folgt für den zweiten

$$E \left[ (\epsilon_{t-1,d}^\phi)^2 \right] = \sigma_\epsilon^2 \sum_{r=1}^{\infty} \phi_{d+r}^2 \left( \frac{t}{n}, \frac{t+d}{n}; 1 \right). \quad (6.38)$$

Abschätzen der Summe durch ihr Supremum und Hervorziehen ergibt dann für das zweite Moment

$$\begin{aligned} & E \left[ \left( \frac{1}{n} \sum_{t=1-m}^{n-m} K_h \left( \frac{t+m}{n} - u \right) \left[ \frac{1}{h} \left( \frac{t+m}{n} - u \right) \right]^l \epsilon_t \epsilon_{t-1,d}^\phi \right)^2 \right] \\ & \leq \sup_{t \leq n} \sum_{r=1}^{\infty} \phi_{d+r}^2 \left( \frac{t}{n}, \frac{t+d}{n}; 1 \right) \frac{\sigma_\epsilon^4}{nh} \left( \frac{1}{nh} \sum_{t=1-m}^{n-m} K^2 \left( \frac{1}{h} \left( \frac{t+m}{n} - u \right) \right) \left[ \frac{1}{h} \left( \frac{t+m}{n} - u \right) \right]^{2l} \right). \end{aligned} \quad (6.39)$$

Aus  $h \rightarrow 0$  und  $nh^2 \rightarrow \infty$  folgt zum einen die Konvergenz

$$\frac{1}{nh} \sum_{t=1-m}^{n-m} K^2 \left( \frac{1}{h} \left( \frac{t+m}{n} - u \right) \right) \left[ \frac{1}{h} \left( \frac{t+m}{n} - u \right) \right]^{2l} \xrightarrow{n \rightarrow \infty} \int K^2(s) s^{2l} ds, \quad (6.40)$$

wobei die Integrale für  $l \in \{0, 1, 2, 3\}$  nach Voraussetzung (K1) und (K4) existieren, und zum anderen, dass  $nh \rightarrow \infty$ , also  $\frac{1}{nh} \rightarrow 0$  gilt. Zusammen folgt

$$E \left[ \left( \frac{1}{n} \sum_{t=1-m}^{n-m} K_h \left( \frac{t+m}{n} - u \right) \left[ \frac{1}{h} \left( \frac{t+m}{n} - u \right) \right]^l \epsilon_t \epsilon_{t-1,d}^\phi \right)^2 \right] \xrightarrow{n \rightarrow \infty} 0, \quad (6.41)$$

wobei benutzt wurde, dass die gleichmäßige Beschränktheit von  $\sum_{r=1}^{\infty} \phi_{d+r}^2 \left( \frac{t}{n}, \frac{t+d}{n}; 1 \right)$  nach Lemma 7.4 (i) erfüllt ist.

zu (ii): Die Behauptung ergibt sich hier völlig analog zum Beweis der ersten Aussage aus Lemma 7.4 (ii).  $\square$

**Lemma 6.4.**

Neben den Annahmen von Satz 3.2 gelten die Voraussetzungen (K1) - (K4) für den Kern  $K$  und für  $k = 1, \dots, p$  seien die Koeffizientenfunktionen  $\alpha_k(\cdot)$  zweimal stetig differenzierbar auf  $(0, 1)$  mit gleichmäßig beschränkten Ableitungen auf  $[0, 1]$ . Falls für ein  $\delta > 0$  die Bedingung (S1) gilt sowie  $h \rightarrow 0$  und  $nh^2 \rightarrow \infty$ , dann konvergieren die Summanden (6.18), (6.19) und (6.20) stochastisch gegen null, d.h für  $l = 0, 1, 2, 3$  gilt

$$\begin{aligned}
(i) \quad & \frac{1}{n} \sum_{t=1-m}^{n-m} K_h\left(\frac{t+m}{n} - u\right) \left[ \frac{1}{h} \left( \frac{t+m}{n} - u \right) \right]^l (1-L) \tilde{\phi}_d\left(\frac{t}{n}, \frac{t+d}{n}; L\right) \epsilon_t^2 = o_P(1), \\
(ii) \quad & \frac{1}{n} \sum_{t=1-m}^{n-m} K_h\left(\frac{t+m}{n} - u\right) \left[ \frac{1}{h} \left( \frac{t+m}{n} - u \right) \right]^l (1-L) \sum_{r=1}^{\infty} \tilde{\phi}_{d+r}\left(\frac{t}{n}, \frac{t+d}{n}; L\right) \epsilon_t \epsilon_{t-r} = o_P(1), \\
(iii) \quad & \frac{1}{n} \sum_{t=1-m}^{n-m} K_h\left(\frac{t+m}{n} - u\right) \left[ \frac{1}{h} \left( \frac{t+m}{n} - u \right) \right]^l (1-L) \sum_{r=1}^{\infty} \tilde{\phi}_{d-r}\left(\frac{t}{n}, \frac{t+d}{n}; L\right) \epsilon_t \epsilon_{t+r} = o_P(1).
\end{aligned}$$

*Beweis.*

zu (i): Als erstes multipliziert man den Term  $(1-L) \tilde{\phi}_d\left(\frac{t}{n}, \frac{t+d}{n}; L\right) \epsilon_t^2$  aus und addiert dann eine null, indem man den Ausdruck

$$\frac{1}{n} \sum_{t=1-m}^{n-m} K_h\left(\frac{t+m-1}{n} - u\right) \left[ \frac{1}{h} \left( \frac{t+m-1}{n} - u \right) \right]^l \tilde{\phi}_d\left(\frac{t-1}{n}, \frac{t+d-1}{n}; L\right) \epsilon_{t-1}^2 \quad (6.42)$$

addiert und subtrahiert. Umordnen und Zusammenfassen ergibt dann

$$\begin{aligned}
& \frac{1}{n} \sum_{t=1-m}^{n-m} K_h\left(\frac{t+m}{n} - u\right) \left[ \frac{1}{h} \left( \frac{t+m}{n} - u \right) \right]^l (1-L) \tilde{\phi}_d\left(\frac{t}{n}, \frac{t+d}{n}; L\right) \epsilon_t^2 \\
= & \frac{1}{n} \sum_{t=1-m}^{n-m} \left( K_h\left(\frac{t+m}{n} - u\right) \left[ \frac{1}{h} \left( \frac{t+m}{n} - u \right) \right]^l \tilde{\phi}_d\left(\frac{t}{n}, \frac{t+d}{n}; L\right) \epsilon_t^2 \right. \\
& \left. - K_h\left(\frac{t+m-1}{n} - u\right) \left[ \frac{1}{h} \left( \frac{t+m-1}{n} - u \right) \right]^l \tilde{\phi}_d\left(\frac{t-1}{n}, \frac{t+d-1}{n}; L\right) \epsilon_{t-1}^2 \right) \quad (6.43)
\end{aligned}$$

$$\begin{aligned}
+ & \frac{1}{n} \sum_{t=1-m}^{n-m} \left( K_h\left(\frac{t+m-1}{n} - u\right) \left[ \frac{1}{h} \left( \frac{t+m-1}{n} - u \right) \right]^l \right. \\
& \left. - K_h\left(\frac{t+m}{n} - u\right) \left[ \frac{1}{h} \left( \frac{t+m}{n} - u \right) \right]^l \right) \tilde{\phi}_d\left(\frac{t-1}{n}, \frac{t+d-1}{n}; L\right) \epsilon_{t-1}^2. \quad (6.44)
\end{aligned}$$

Als nächstes betrachtet man die beiden Summen oben einzeln und weist ihr asymptotisches Verschwinden nach. Die erste Summe (6.43) ist per Konstruktion gerade so zusam-

mengefasst, dass sie eine Teleskopsumme darstellt, welche zu

$$\begin{aligned}
& \frac{1}{n} \left( K_h\left(\frac{n}{n} - u\right) \left[ \frac{1}{h} \left( \frac{n}{n} - u \right) \right]^l \tilde{\phi}_d\left(\frac{n-m}{n}, \frac{n-m+d}{n}; L\right) \epsilon_{n-m}^2 \right. \\
& \quad \left. - K_h\left(\frac{0}{n} - u\right) \left[ \frac{1}{h} \left( \frac{0}{n} - u \right) \right]^l \tilde{\phi}_d\left(\frac{-m}{n}, \frac{-m+d}{n}; L\right) \epsilon_{-m}^2 \right) \\
&= \frac{1}{nh} \left( K\left(\frac{n-nu}{nh}\right) \left[ \frac{n-nu}{nh} \right]^l \tilde{\phi}_d\left(\frac{n-m}{n}, \frac{n-q}{n}; L\right) \epsilon_{n-m}^2 \right. \\
& \quad \left. - K\left(\frac{0-nu}{nh}\right) \left[ \frac{0-nu}{nh} \right]^l \tilde{\phi}_d\left(\frac{-m}{n}, \frac{-q}{n}; L\right) \epsilon_{-m}^2 \right) \tag{6.45}
\end{aligned}$$

kollabiert. Nimmt man nun den Erwartungswert vom Betrag der rechten Seite und schätzt ab, so folgt

$$\begin{aligned}
& E \left[ \frac{1}{nh} \left| K\left(\frac{n-nu}{nh}\right) \left[ \frac{n-nu}{nh} \right]^l \tilde{\phi}_d\left(\frac{n-m}{n}, \frac{n-q}{n}; L\right) \epsilon_{n-m}^2 \right. \right. \\
& \quad \left. \left. - K\left(\frac{0-nu}{nh}\right) \left[ \frac{0-nu}{nh} \right]^l \tilde{\phi}_d\left(\frac{-m}{n}, \frac{-q}{n}; L\right) \epsilon_{-m}^2 \right| \right] \\
& \leq \frac{\sigma_\epsilon^2}{nh} K\left(\frac{n-nu}{nh}\right) \left| \frac{n-nu}{nh} \right|^l \sum_{j=0}^{\infty} \sum_{s=j+1}^{\infty} |\varphi_s\left(\frac{n-m}{n}\right) \varphi_{s+d}\left(\frac{n-q}{n}\right)| \\
& \quad + \frac{\sigma_\epsilon^2}{nh} K\left(\frac{0-nu}{nh}\right) \left| \frac{0-nu}{nh} \right|^l \sum_{j=0}^{\infty} \sum_{s=j+1}^{\infty} |\varphi_s\left(\frac{-m}{n}\right) \varphi_{s+d}\left(\frac{-q}{n}\right)|. \tag{6.46}
\end{aligned}$$

Einerseits sind die beiden Terme  $K\left(\frac{n-nu}{nh}\right) \left| \frac{n-nu}{nh} \right|^l$  und  $K\left(\frac{0-nu}{nh}\right) \left| \frac{0-nu}{nh} \right|^l$  beschränkt, weil  $K$  einen kompakten Träger hat und dort stetig ist, andererseits sind die beiden vorkommenden Doppelsummen gleichmäßig beschränkt nach Lemma 7.3 (iv), so dass wegen  $\frac{1}{nh} \rightarrow 0$  mit Lemma 2.26 (i) die stochastische Konvergenz der Summe (6.43) gegen null folgt.

Um den Beweis der Aussage (i) zu komplettieren, muss man noch nachweisen, dass auch die zweite Summe (6.44) asymptotisch verschwindet. Dazu bildet man wieder den Erwar-

tungswert vom Betrag der Summe und schätzt ab. Es gilt

$$\begin{aligned}
 & E \left[ \left| \frac{1}{n} \sum_{t=1-m}^{n-m} \left( K_h\left(\frac{t+m-1}{n} - u\right) \left[ \frac{1}{h} \left( \frac{t+m-1}{n} - u \right) \right]^l \right. \right. \right. \\
 & \quad \left. \left. \left. - K_h\left(\frac{t+m}{n} - u\right) \left[ \frac{1}{h} \left( \frac{t+m}{n} - u \right) \right]^l \right) \tilde{\varphi}_d\left(\frac{t-1}{n}, \frac{t+d-1}{n}; L\right) \epsilon_{t-1}^2 \right| \right] \\
 & \leq \frac{\sigma_\epsilon^2}{n} \sum_{t=1-m}^{n-m} \left| K_h\left(\frac{t+m-1}{n} - u\right) \left[ \frac{1}{h} \left( \frac{t+m-1}{n} - u \right) \right]^l \right. \\
 & \quad \left. - K_h\left(\frac{t+m}{n} - u\right) \left[ \frac{1}{h} \left( \frac{t+m}{n} - u \right) \right]^l \right| \sum_{j=0}^{\infty} \sum_{s=j+1}^{\infty} |\varphi_s\left(\frac{t-1}{n}\right) \varphi_{s+d}\left(\frac{t+d-1}{n}\right)|.
 \end{aligned} \tag{6.47}$$

Als nächstes betrachtet man die Differenz in Betragstrichen etwas genauer. Dazu definiert man zunächst für  $l \in \{0, 1, 2, 3\}$  die Funktion  $\tilde{K}_l$  durch

$$\tilde{K}_l(s) := K(s)s^l \tag{6.48}$$

für  $s \in \mathbb{R}$  und entwickelt dann  $\tilde{K}_l\left(s - \frac{1}{nh}\right)$  nach Taylor. Mit Voraussetzung (K3) hat der Kern  $K$  gleichmäßig beschränkte erste Ableitungen auf  $\mathbb{R}$  und damit auch  $\tilde{K}_l$ , so dass für eine geeignete Zwischenstelle  $\tilde{s}$  die Gleichheit

$$\tilde{K}_l\left(s - \frac{1}{nh}\right) = \tilde{K}_l(s) - \frac{\tilde{K}'_l(\tilde{s})}{nh} \tag{6.49}$$

gilt. Schreibt man nun obige Differenz mit Hilfe von  $\tilde{K}_l$  um, wendet danach die Taylorentwicklung von  $\tilde{K}_l\left(s - \frac{1}{nh}\right)$  für  $s = \frac{t+m-nu}{nh}$  an und schätzt die Funktion  $\tilde{K}'_l$  an der Zwischenstelle  $\tilde{s}$  gegen ihr Supremum über  $\mathbb{R}$  ab, so folgt

$$\begin{aligned}
 & \left| K_h\left(\frac{t+m-1}{n} - u\right) \left[ \frac{1}{h} \left( \frac{t+m-1}{n} - u \right) \right]^l - K_h\left(\frac{t+m}{n} - u\right) \left[ \frac{1}{h} \left( \frac{t+m}{n} - u \right) \right]^l \right| \\
 & = \frac{1}{h} \left| \tilde{K}_l\left(\frac{t+m-nu}{nh} - \frac{1}{nh}\right) - \tilde{K}_l\left(\frac{t+m-nu}{nh}\right) \right| \\
 & = \frac{1}{h} \left| -\frac{\tilde{K}'_l(\tilde{s})}{nh} \right| \\
 & \leq \frac{1}{nh^2} \sup_{x \in \mathbb{R}} |\tilde{K}'_l(x)|.
 \end{aligned} \tag{6.50}$$

Setzt man diese Ungleichung in (6.47) ein, schätzt die Doppelsumme gegen ihr Supremum über  $t \leq n$  ab und beachtet, dass sich dann die Summe über  $t$  gegen den Bruch  $\frac{1}{n}$  weghebt,

so erhält man

$$\begin{aligned}
 & E \left[ \left| \frac{1}{n} \sum_{t=1-m}^{n-m} \left( K_h\left(\frac{t+m-1}{n} - u\right) \left[ \frac{1}{h} \left( \frac{t+m-1}{n} - u \right) \right]^l \right. \right. \right. \\
 & \quad \left. \left. \left. - K_h\left(\frac{t+m}{n} - u\right) \left[ \frac{1}{h} \left( \frac{t+m}{n} - u \right) \right]^l \right) \tilde{\phi}_d\left(\frac{t-1}{n}, \frac{t+d-1}{n}; L\right) \epsilon_{t-1}^2 \right| \right] \quad (6.51) \\
 & \leq \frac{\sigma_\epsilon^2}{nh^2} \sup_{x \in \mathbb{R}} |\tilde{K}'_l(x)| \sup_{t \leq n} \sum_{j=0}^{\infty} \sum_{s=j+1}^{\infty} |\varphi_s\left(\frac{t-1}{n}\right) \varphi_{s+d}\left(\frac{t+d-1}{n}\right)|.
 \end{aligned}$$

Weil das erste Supremum wegen Voraussetzung (K3) beschränkt ist und das zweite wieder nach Lemma 7.3 (iv), so folgt die Konvergenz nach Wahrscheinlichkeit gegen null von (6.18) aus  $\frac{1}{nh^2} \rightarrow 0$  mit Lemma 2.26 (i).

zu (ii) und (iii): Man führt die Beweise der zweiten und dritten Aussage völlig analog, jedoch schätzt man die hier vorkommenden Erwartungswerte  $E[|\epsilon_{t-j}\epsilon_{t-r-j}|]$  bzw.  $E[|\epsilon_{t-j}\epsilon_{t+r-j}|]$  mit Cauchy-Schwarz (2.39) nach oben durch  $\sigma_\epsilon^2 < \infty$  ab und die gleichmäßige Beschränktheit der jeweiligen Mehrfachsumme folgt aus Lemma 7.3 (v) bzw. aus Lemma 7.3 (vi).  $\square$

### Lemma 6.5 (Verschwinden des Restterms).

Neben den Annahmen von Satz 3.2 gelten die Voraussetzungen (K1) - (K4) für den Kern  $K$  und für  $k = 1, \dots, p$  seien die Koeffizientenfunktionen  $\alpha_k(\cdot)$  zweimal stetig differenzierbar auf  $(0, 1)$  mit gleichmäßig beschränkten Ableitungen auf  $[0, 1]$ . Falls für ein  $\delta > 0$  zusätzlich zu (S1) die Bedingung

$$(S2) \quad \sup_{t \leq n} \sum_{s=1}^{\infty} s^{2+\delta} |\Delta_t[\varphi_s(\frac{t}{n})]| = o(1) \quad (6.52)$$

gilt, wobei  $\Delta_t$  den Differenzoperator bzgl.  $t$  gemäß  $\Delta_t[f(t)] := f(t) - f(t-1)$  bezeichnet sowie  $h \rightarrow 0$  und  $nh^2 \rightarrow \infty$ , dann konvergiert der Summand (6.21) stochastisch gegen null, d.h für  $l = 0, 1, 2, 3$  gilt

$$\frac{1}{n} \sum_{t=1-m}^{n-m} K_h\left(\frac{t+m}{n} - u\right) \left[ \frac{1}{h} \left( \frac{t+m}{n} - u \right) \right]^l R_d\left(\frac{t}{n}, \frac{t+d}{n}; L\right) \epsilon_t^2 = o_P(1). \quad (6.53)$$

*Beweis.*

Per Definition in Satz 5.4 besteht der Restterm  $R_d(\frac{t}{n}, \frac{t+d}{n}; L)\epsilon_t^2$  aus den drei additiven

Komponenten

$$R_d^{(1)}\left(\frac{t}{n}, \frac{t+d}{n}; L\right) \epsilon_t^2 := \left[ \tilde{\phi}_d\left(\frac{t}{n}, \frac{t+d}{n}; L\right) - \tilde{\phi}_d\left(\frac{t-1}{n}, \frac{t+d-1}{n}; L\right) \right] \epsilon_{t-1}^2, \quad (6.54)$$

$$R_d^{(2)}\left(\frac{t}{n}, \frac{t+d}{n}; L\right) \epsilon_t^2 := \sum_{r=1}^{\infty} \left[ \tilde{\phi}_{d+r}\left(\frac{t}{n}, \frac{t+d}{n}; L\right) - \tilde{\phi}_{d+r}\left(\frac{t-1}{n}, \frac{t+d-1}{n}; L\right) \right] \epsilon_{t-1} \epsilon_{t-r-1}, \quad (6.55)$$

$$R_d^{(3)}\left(\frac{t}{n}, \frac{t+d}{n}; L\right) \epsilon_t^2 := \sum_{r=1}^{\infty} \left[ \tilde{\phi}_{d-r}\left(\frac{t}{n}, \frac{t+d}{n}; L\right) - \tilde{\phi}_{d-r}\left(\frac{t-1}{n}, \frac{t+d-1}{n}; L\right) \right] \epsilon_{t-1} \epsilon_{t+r-1}, \quad (6.56)$$

so dass die Behauptung des Lemmas folgt, wenn man

$$(i) \quad \frac{1}{n} \sum_{t=1-m}^{n-m} K_h\left(\frac{t+m}{n} - u\right) \left[ \frac{1}{h} \left( \frac{t+m}{n} - u \right) \right]^l R_d^{(1)}\left(\frac{t}{n}, \frac{t+d}{n}; L\right) \epsilon_t^2 = o_P(1), \quad (6.57)$$

$$(ii) \quad \frac{1}{n} \sum_{t=1-m}^{n-m} K_h\left(\frac{t+m}{n} - u\right) \left[ \frac{1}{h} \left( \frac{t+m}{n} - u \right) \right]^l R_d^{(2)}\left(\frac{t}{n}, \frac{t+d}{n}; L\right) \epsilon_t^2 = o_P(1), \quad (6.58)$$

$$(iii) \quad \frac{1}{n} \sum_{t=1-m}^{n-m} K_h\left(\frac{t+m}{n} - u\right) \left[ \frac{1}{h} \left( \frac{t+m}{n} - u \right) \right]^l R_d^{(3)}\left(\frac{t}{n}, \frac{t+d}{n}; L\right) \epsilon_t^2 = o_P(1) \quad (6.59)$$

zeigt. Dazu nimmt man jeweils den Erwartungswert vom Absolutbetrag und zeigt, dass dieser dann gegen null konvergiert, was nach Lemma 2.26 (i) eine hinreichende Bedingung für Konvergenz nach Wahrscheinlichkeit gegen null darstellt.

zu (i): Beachtet man, dass

$$|\varphi_s\left(\frac{t}{n}\right) \varphi_{s+d}\left(\frac{t+d}{n}\right) - \varphi_s\left(\frac{t-1}{n}\right) \varphi_{s+d}\left(\frac{t+d-1}{n}\right)| = |\Delta_t[\varphi_s\left(\frac{t}{n}\right) \varphi_{s+d}\left(\frac{t+d}{n}\right)]| \quad (6.60)$$

gilt, so folgt durch Abschätzen und Hereinziehen des Erwartungswerts

$$\begin{aligned} & E \left[ \left| \frac{1}{n} \sum_{t=1-m}^{n-m} K_h\left(\frac{t+m}{n} - u\right) \left[ \frac{1}{h} \left( \frac{t+m}{n} - u \right) \right]^l R_d^{(1)}\left(\frac{t}{n}, \frac{t+d}{n}; L\right) \epsilon_t^2 \right| \right] \\ & \leq \frac{\sigma_\epsilon^2}{n} \sum_{t=1-m}^{n-m} K_h\left(\frac{t+m}{n} - u\right) \left[ \frac{1}{h} \left( \frac{t+m}{n} - u \right) \right]^l \sum_{j=0}^{\infty} \sum_{s=j+1}^{\infty} |\Delta_t[\varphi_s\left(\frac{t}{n}\right) \varphi_{s+d}\left(\frac{t+d}{n}\right)]|. \end{aligned} \quad (6.61)$$

Bildet man noch das Supremum der Doppelsumme über  $t \leq n$  und zieht es nach vorne, so folgt die Konvergenz

$$E \left[ \left| \frac{1}{n} \sum_{t=1-m}^{n-m} K_h\left(\frac{t+m}{n} - u\right) \left[ \frac{1}{h} \left( \frac{t+m}{n} - u \right) \right]^l R_d^{(1)}\left(\frac{t}{n}, \frac{t+d}{n}; L\right) \epsilon_t^2 \right| \right] \xrightarrow{n \rightarrow \infty} 0 \quad (6.62)$$

aus der gleichmäßigen Konvergenz gegen null des Supremums nach Lemma 7.5 (i) bzw. aus

$$\frac{1}{n} \sum_{t=1-m}^{n-m} K_h\left(\frac{t+m}{n} - u\right) \left[ \frac{1}{h} \left( \frac{t+m}{n} - u \right) \right]^l \xrightarrow{n \rightarrow \infty} \int K(s) s^l ds < \infty, \quad (6.63)$$

womit die stochastische Konvergenz gegen null des ersten Ausdrucks (i) gezeigt ist.

zu (ii) und (iii): Der Nachweis der Konvergenz nach Wahrscheinlichkeit gegen null des zweiten und des dritten Ausdrucks folgt analog. Man schätzt jedoch wieder  $E[|\epsilon_{t-j}\epsilon_{t-r-j}|]$  und  $E[|\epsilon_{t-j}\epsilon_{t+r-j}|]$  mit der Cauchy-Schwarz-Ungleichung (2.39) durch  $\sigma_\epsilon^2 < \infty$  ab und das Verschwinden der vorkommenden Mehrfachsummen folgt aus Lemma 7.5 (ii) bzw. Lemma 7.5 (iii).  $\square$

Am Anfang dieses Abschnitts konnten wir das Problem der Asymptotik von  $\mathbf{S}_{n,l}(u)$  auf den eindimensionalen Fall reduzieren, indem wir einen repräsentativen Eintrag  $S_{n,l,m,q}(u)$  betrachtet haben, und wegen Lemma 6.1 die in Kapitel 5 hergeleitete  $tvBN_2$ -Zerlegung darauf anwenden. Im nun folgenden Satz fassen wir die Aussagen der vorangegangenen Lemmata 6.2 bis 6.5 über das asymptotische Verhalten der Summanden (6.15) bis (6.21) jener Zerlegung zusammen.

**Satz 6.6 (Asymptotik der empirischen Momente  $\mathbf{S}_{n,l}(u)$ ).**

*Neben den Annahmen von Satz 3.2 gelten die Voraussetzungen (K1) - (K4) für den Kern  $K$  und für  $k = 1, \dots, p$  seien die Koeffizientenfunktionen  $\alpha_k(\cdot)$  zweimal stetig differenzierbar auf  $(0, 1)$  mit gleichmäßig beschränkten Ableitungen auf  $[0, 1]$ . Falls für ein  $\delta > 0$  die Bedingungen (S1) und (S2) gelten sowie  $h \rightarrow 0$  und  $nh^2 \rightarrow \infty$ , dann folgt für  $l = 0, 1, 2, 3$  die Konvergenz nach Wahrscheinlichkeit*

$$S_{n,l,m,q}(u) \xrightarrow{P} \sigma_\epsilon^2 \phi_d(u, u; 1) \int K(s) s^l ds. \quad (6.64)$$

Mit den Bezeichnungen  $\phi_k(u) := \phi_k(u, u; 1)$  für  $k \in \{0, \dots, p-1\}$  und  $u \in (0, 1)$  sowie

$$\mathbf{\Gamma}(u) := \begin{pmatrix} \phi_0(u) & \cdots & \cdots & \phi_{p-1}(u) \\ \phi_1(u) & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \vdots \\ \phi_{p-1}(u) & \cdots & \phi_1(u) & \phi_0(u) \end{pmatrix} \quad (6.65)$$

folgt dann für die  $p \times p$ -Matrix  $\mathbf{S}_{n,l}(u) = (S_{n,l,m,q}(u))_{m,q=1,\dots,p}$  die Konvergenz

$$\mathbf{S}_{n,l}(u) \xrightarrow{P} \left( \sigma_\epsilon^2 \int K(s) s^l ds \right) \mathbf{\Gamma}(u). \quad (6.66)$$

*Beweis.*

Da die Voraussetzungen der Lemmata 6.1 bis 6.5 erfüllt sind, folgt die Behauptung sofort aus der  $tvBN_2$ -Zerlegung von  $S_{n,l,m,q}(u)$  auf Seite 63 und der Definition von  $\mathbf{S}_n(u)$  aus Lemma 4.2.  $\square$

## 6.2 Asymptotik des Bias-Terms und des Restterms

Da wir im letzten Abschnitt den Grenzwert nach Wahrscheinlichkeit von  $S_{n,l,m,q}(u)$  bzw. von  $\mathbf{S}_{n,l}(u)$  für  $u \in (0, 1)$  bestimmen konnten, sind wir jetzt auch in der Lage die Grenzwerte des Bias-Terms  $\underline{B}_n(u)$  und des Restterms  $\underline{R}_n(u)$  aus Lemma 4.3 anzugeben.

**Satz 6.7 (Asymptotik des Bias-Terms  $\underline{B}_n(u)$ ).**

Neben den Annahmen von Satz 3.2 gelten die Voraussetzungen (K1) - (K4) für den Kern  $K$  und für  $k = 1, \dots, p$  seien die Koeffizientenfunktionen  $\alpha_k(\cdot)$  zweimal stetig differenzierbar auf  $(0, 1)$  mit gleichmäßig beschränkten Ableitungen auf  $[0, 1]$ . Für ein  $\delta > 0$  gelten die Bedingungen (S1) und (S2) sowie  $h \rightarrow 0$ ,  $nh^2 \rightarrow \infty$  und  $n^{1/2}h^{5/2} \rightarrow C$  für eine Konstante  $C < \infty$ . Ist  $\Gamma(u)$  aus Satz 6.6 invertierbar, dann folgt für den Bias-Term die Konvergenz

$$\sqrt{nh}\underline{B}_n(u) := \frac{n^{1/2}h^{5/2}}{2} \mathbf{E}_0 \mathbf{S}_n^{-1}(u) \begin{pmatrix} \mathbf{S}_{n,2}(u) \\ \mathbf{S}_{n,3}(u) \end{pmatrix} \underline{\alpha}''(u) \xrightarrow{P} \frac{C}{2} \mu_K^2 \underline{\alpha}''(u), \quad (6.67)$$

wobei  $\mu_K^2 := \int K(r)r^2 dr$  definiert ist.

*Beweis.*

Zum einen sind die Voraussetzungen von Lemma 4.3 erfüllt, so dass die Zerlegung des Schätzfehlers  $\widehat{\alpha}(u) - \alpha(u)$  Gültigkeit besitzt, und zum anderen ist Satz 6.6 anwendbar, da auch seine Bedingungen erfüllt sind. Wegen  $\int K(r)dr = 1$  nach Voraussetzung (K1) gilt deshalb

$$\mathbf{S}_{n,0}(u) \xrightarrow{P} \sigma_\epsilon^2 \int K(r)dr \Gamma(u) = \sigma_\epsilon^2 \Gamma(u), \quad (6.68)$$

$$\mathbf{S}_{n,2}(u) \xrightarrow{P} \sigma_\epsilon^2 \int K(r)r^2 dr \Gamma(u) = \sigma_\epsilon^2 \mu_K^2 \Gamma(u). \quad (6.69)$$

Nach (K2) ist der Kern  $K$  als symmetrisch bzw. als gerade vorausgesetzt und die Ausdrücke  $r$  und  $r^3$  sind bekanntermaßen ungerade, so dass  $K(r)r$  und  $K(r)r^3$  wieder ungerade Funktionen sind, deren Integrale über die reelle Achse null ergeben. Nochmals mit Satz 6.6 ergibt sich daher

$$\mathbf{S}_{n,1}(u) \xrightarrow{P} \sigma_\epsilon^2 \int K(r)r dr \Gamma(u) = \mathbf{O}_{p \times p}, \quad (6.70)$$

$$\mathbf{S}_{n,3}(u) \xrightarrow{P} \sigma_\epsilon^2 \int K(r)r^3 dr \Gamma(u) = \mathbf{O}_{p \times p}. \quad (6.71)$$

Zusammen folgt damit die Konvergenz der  $2p \times 2p$ -Matrix  $\mathbf{S}_n(u)$  gemäß

$$\mathbf{S}_n(u) \xrightarrow{P} \sigma_\epsilon^2 \begin{pmatrix} \mathbf{\Gamma}(u) & \mathbf{O}_{p \times p} \\ \mathbf{O}_{p \times p} & \mu_K^2 \mathbf{\Gamma}(u) \end{pmatrix} =: \mathbf{S}(u). \quad (6.72)$$

Weil aufgrund der Invertierbarkeit von  $\mathbf{\Gamma}(u)$  auch die Inverse von  $\mathbf{S}(u)$  existiert, gilt ebenso

$$\mathbf{S}_n^{-1}(u) \xrightarrow{P} \sigma_\epsilon^{-2} \begin{pmatrix} \mathbf{\Gamma}^{-1}(u) & \mathbf{O}_{p \times p} \\ \mathbf{O}_{p \times p} & \mu_K^{-2} \mathbf{\Gamma}^{-1}(u) \end{pmatrix} = \mathbf{S}^{-1}(u). \quad (6.73)$$

Mit Satz 2.30 (Continuous Mapping Theorem) folgt dann die Behauptung aus

$$\begin{aligned} \sqrt{nh} \underline{B}_n &= \frac{n^{1/2} h^{5/2}}{2} \mathbf{E}_0 \mathbf{S}_n^{-1}(u) \begin{pmatrix} \mathbf{S}_{n,2}(u) \\ \mathbf{S}_{n,3}(u) \end{pmatrix} \underline{\alpha}''(u) \\ &\xrightarrow{P} \frac{C}{2} (\mathbf{I}_p, \mathbf{O}_{p \times p}) \sigma_\epsilon^{-2} \begin{pmatrix} \mathbf{\Gamma}^{-1}(u) & \mathbf{O}_{p \times p} \\ \mathbf{O}_{p \times p} & \mu_K^{-2} \mathbf{\Gamma}^{-1}(u) \end{pmatrix} \begin{pmatrix} \sigma_\epsilon^2 \mu_K^2 \mathbf{\Gamma}(u) \\ \mathbf{O}_{p \times p} \end{pmatrix} \underline{\alpha}''(u) \\ &= \frac{C}{2} \mu_K^2 \underline{\alpha}''(u). \end{aligned} \quad (6.74)$$

□

**Satz 6.8 (Asymptotik des Restterms  $\underline{R}_n(u)$ ).**

Neben den Annahmen von Satz 3.2 gelten die Voraussetzungen (K1) - (K4) für den Kern  $K$  und für  $k = 1, \dots, p$  seien die Koeffizientenfunktionen  $\alpha_k(\cdot)$  zweimal stetig differenzierbar auf  $(0, 1)$  mit gleichmäßig beschränkten Ableitungen auf  $[0, 1]$ . Für ein  $\delta > 0$  gelten die Bedingungen (S1) und (S2) sowie  $h \rightarrow 0$ ,  $nh^2 \rightarrow \infty$  und  $n^{1/2} h^{5/2} \rightarrow C$  für eine Konstante  $C < \infty$ . Ist  $\mathbf{\Gamma}(u)$  aus Satz 6.6 invertierbar, dann konvergiert der Restterm  $\sqrt{nh} \underline{R}_n(u)$  stochastisch gegen null, d.h. es gilt

$$\sqrt{nh} \underline{R}_n(u) := \frac{n^{1/2} h^{5/2}}{2} \mathbf{E}_0 \mathbf{S}_n^{-1}(u) \begin{pmatrix} \mathbf{S}_{n,2}(u) \\ \mathbf{S}_{n,3}(u) \end{pmatrix} o(1) = o_P(1). \quad (6.75)$$

*Beweis.*

Wie im Beweis von Satz 6.7 gelten hier die Konvergenzen

$$\mathbf{S}_{n,2}(u) \xrightarrow{P} \sigma_\epsilon^2 \int K(r) r^2 dr \mathbf{\Gamma}(u) = \sigma_\epsilon^2 \mu_K^2 \mathbf{\Gamma}(u), \quad (6.76)$$

$$\mathbf{S}_{n,3}(u) \xrightarrow{P} \sigma_\epsilon^2 \int K(r) r^3 dr \mathbf{\Gamma}(u) = \mathbf{O}_{p \times p}, \quad (6.77)$$

und

$$\mathbf{S}_n^{-1}(u) \xrightarrow{P} \sigma_\epsilon^{-2} \begin{pmatrix} \mathbf{\Gamma}^{-1}(u) & \mathbf{O}_{p \times p} \\ \mathbf{O}_{p \times p} & \mu_K^{-2} \mathbf{\Gamma}^{-1}(u) \end{pmatrix} = \mathbf{S}^{-1}(u), \quad (6.78)$$

so dass sich aus der Annahme  $n^{1/2}h^{5/2} \rightarrow C$  mit  $C < \infty$  für den ersten Teil des Restterms

$$\frac{n^{1/2}h^{5/2}}{2} \mathbf{E}_0 \mathbf{S}_n^{-1}(u) \begin{pmatrix} \mathbf{S}_{n,2}(u) \\ \mathbf{S}_{n,3}(u) \end{pmatrix} = O_P(1) \quad (6.79)$$

ergibt. Mit Satz 2.29 (Slutsky) folgt dann die Behauptung.  $\square$

### 6.3 Asymptotische Normalität des Schätzfehlers

An dieser Stelle befassen wir uns mit dem führenden stochastischen Term  $\underline{T}_n(u)$  der Zerlegung des Schätzfehlers aus Lemma 4.3 und dessen Asymptotik. Für diesen gilt

$$\underline{T}_n(u) = \mathbf{E}_0 \mathbf{S}_n^{-1}(u) \underline{\tau}_n(u), \quad (6.80)$$

wobei die Einträge des  $2p$ -dimensionalen Vektors  $\underline{\tau}_n(u)$  aus

$$\tau_{n,l} := \frac{1}{n} \sum_{t=1}^n K_h\left(\frac{t}{n} - u\right) \left[ \frac{1}{h} \left( \frac{t}{n} - u \right) \right]^l \underline{X}_{t-1} \epsilon_t \in \text{Mat}_{p,1}(\mathbb{R}) \quad (6.81)$$

mit  $l = 0$  und  $l = 1$  bestehen. Da die Asymptotik von  $\mathbf{S}_n^{-1}(u)$  bereits aus Satz 6.6 bekannt ist, bleibt hier nur noch das asymptotische Verhalten von  $\underline{\tau}_n(u)$  zu untersuchen. Ähnlich zu Lemma 6.1 stellt das nun folgende Lemma den Zusammenhang von  $\underline{\tau}_n(u)$  mit dem Prozess  $\{\tilde{X}_{t,n}\}_{t=-\infty}^n$  her.

#### Lemma 6.9.

*Neben den Annahmen von Satz 3.2 gelten die Voraussetzungen (K1) - (K4) für den Kern  $K$  sowie  $h \rightarrow 0$  und  $nh^2 \rightarrow \infty$ . Sind für  $k = 1, \dots, p$  die Koeffizientenfunktionen  $\alpha_k(\cdot)$  zweimal stetig differenzierbar auf  $(0, 1)$  mit gleichmäßig beschränkten Ableitungen auf  $[0, 1]$ , dann gilt für  $l = 0, 1$*

$$\sqrt{nh} \tau_{n,l}(u) = \frac{1}{\sqrt{nh}} \sum_{t=1}^n K\left(\frac{1}{h}\left(\frac{t}{n} - u\right)\right) \left[ \frac{1}{h} \left( \frac{t}{n} - u \right) \right]^l \tilde{X}_{t-1} \epsilon_t + o_P(1), \quad (6.82)$$

wobei  $\tilde{X}_{t-1} := (\tilde{X}_{t-1,n}, \dots, \tilde{X}_{t-p,n})^T$  definiert ist mit  $\{\tilde{X}_{t,n}\}_{t=-\infty}^n$  aus Lemma 3.7.

*Beweis.*

Ähnlich zum Beweis von Lemma 6.1 setzt man für  $\underline{X}_{t-1}$  den Ausdruck

$$\tilde{X}_{t-1} + (\underline{X}_{t-1} - \tilde{X}_{t-1}) \quad (6.83)$$

ein, schätzt die Differenz durch ihr Supremum über  $t \leq n$  ab und zieht es nach vorne, so folgt

$$\begin{aligned} \sqrt{nh}\underline{\tau}_{n,l} &= \frac{1}{\sqrt{nh}} \sum_{t=1}^n K\left(\frac{1}{h}\left(\frac{t}{n} - u\right)\right) \left[\frac{1}{h}\left(\frac{t}{n} - u\right)\right]^l \tilde{X}_{t-1} \epsilon_t \\ &+ \sup_{t \leq n} \left| \underline{X}_{t-1} - \tilde{X}_{t-1} \right| \frac{1}{\sqrt{nh}} \sum_{t=1}^n K\left(\frac{1}{h}\left(\frac{t}{n} - u\right)\right) \left[\frac{1}{h}\left(\frac{t}{n} - u\right)\right]^l \epsilon_t, \end{aligned} \quad (6.84)$$

wobei  $|\underline{X}_{t-1} - \tilde{X}_{t-1}| := \sum_{k=1}^p |X_{t-k,n} - \tilde{X}_{t-k,n}|$  bezeichnet. Weil die Voraussetzungen von Lemma 3.7 erfüllt sind, gilt

$$\sup_{t \leq n} |X_{t,n} - \tilde{X}_{t,n}| = O_P\left(\frac{1}{n}\right), \quad (6.85)$$

woraus auch

$$\sup_{t \leq n} |\underline{X}_{t-1} - \tilde{X}_{t-1}| = O_P\left(\frac{1}{n}\right) \quad (6.86)$$

folgt. Die Behauptung des Lemmas ergibt sich schließlich aus Satz 2.29 (Slutsky), wenn man die Beschränktheit nach Wahrscheinlichkeit von

$$\frac{1}{\sqrt{nh}} \sum_{t=1}^n K\left(\frac{1}{h}\left(\frac{t}{n} - u\right)\right) \left[\frac{1}{h}\left(\frac{t}{n} - u\right)\right]^l \epsilon_t \quad (6.87)$$

zeigt. Dazu bildet man das zweite Moment dieses Ausdrucks und erhält wegen  $E[\epsilon_t \epsilon_s] = 0$  für  $s \neq t$  die Gleichheit

$$\begin{aligned} &E \left[ \left( \frac{1}{\sqrt{nh}} \sum_{t=1}^n K\left(\frac{1}{h}\left(\frac{t}{n} - u\right)\right) \left[\frac{1}{h}\left(\frac{t}{n} - u\right)\right]^l \epsilon_t \right)^2 \right] \\ &= \sigma_\epsilon^2 \frac{1}{nh} \sum_{t=1}^n K^2\left(\frac{1}{h}\left(\frac{t}{n} - u\right)\right) \left[\frac{1}{h}\left(\frac{t}{n} - u\right)\right]^{2l}, \end{aligned} \quad (6.88)$$

wobei die nach Lemma 2.27 (ii) hinreichende Bedingung der gleichmäßigen Beschränktheit des zweiten Moments dann aus der Konvergenz

$$\frac{1}{nh} \sum_{t=1}^n K^2\left(\frac{1}{h}\left(\frac{t}{n} - u\right)\right) \left[\frac{1}{h}\left(\frac{t}{n} - u\right)\right]^{2l} \xrightarrow{n \rightarrow \infty} \int K^2(s) s^{2l} ds \quad (6.89)$$

folgt, weil die beiden Integrale für  $l \in \{0, 1\}$  nach (K4) endlich sind.  $\square$

Die spezielle Blockstruktur des Grenzwerts nach Wahrscheinlichkeit der Matrix  $\mathbf{S}_n^{-1}(u)$  (siehe (6.73)) sowie die Multiplikation mit  $\mathbf{E}_0 = (\mathbf{I}_{p \times p}, \mathbf{O}_{p \times p})$  bei der Berechnung von  $\underline{T}_n(u)$  hat zur Folge, dass man nur die Verteilungskonvergenz von  $\underline{T}_{n,0}(u)$  untersuchen muss. Für den zweiten Anteil  $\underline{T}_{n,1}(u)$  genügt es zu zeigen, dass

$$\sqrt{nh}\underline{T}_{n,1}(u) \quad (6.90)$$

beschränkt nach Wahrscheinlichkeit ist (siehe auch Theorem 6.11). Beides geschieht in folgendem Satz.

**Satz 6.10 (Asymptotik des stochastischen Terms  $\underline{T}_n(u)$ ).**

Neben den Annahmen von Satz 3.2 gelten die Voraussetzungen (K1) - (K4) für den Kern  $K$  und für  $k = 1, \dots, p$  seien die Koeffizientenfunktionen  $\alpha_k(\cdot)$  zweimal stetig differenzierbar auf  $(0, 1)$  mit gleichmäßig beschränkten Ableitungen auf  $[0, 1]$ . Für ein  $\delta > 0$  gelten die Bedingungen (S1) und (S2) sowie  $h \rightarrow 0$ ,  $nh^2 \rightarrow \infty$  und  $n^{1/2}h^{5/2} \rightarrow C$  für eine Konstante  $C < \infty$ . Dann folgt für den stochastischen Term

$$(i) \quad \sqrt{nh}\underline{T}_{n,1}(u) = O_P(1) \quad (6.91)$$

sowie unter der Annahme  $\kappa_4 := E[\epsilon_t^4] < \infty$  die Verteilungskonvergenz

$$(ii) \quad \sqrt{nh}\underline{T}_{n,0}(u) \xrightarrow{D} \mathcal{N}(0_p, \Sigma(u)), \quad (6.92)$$

wobei  $\Sigma(u) := \sigma_\epsilon^4 \left( \int K^2(r) dr \right) \Gamma(u)$  definiert ist mit  $\Gamma(u)$  aus Satz 6.6.

*Beweis.*

zu (i): Wegen Lemma 2.27 (ii) ist es hinreichend, wenn man die gleichmäßige Beschränktheit der zweiten Momente der Komponenten des Vektors  $\sqrt{nh}\underline{T}_{n,1}(u)$  zeigt. Nach Lemma 6.9 gilt aber

$$\sqrt{nh}\underline{T}_{n,1}(u) = \frac{1}{\sqrt{nh}} \sum_{t=1}^n K\left(\frac{1}{h}\left(\frac{t}{n} - u\right)\right) \left[ \frac{1}{h} \left(\frac{t}{n} - u\right) \right] \tilde{X}_{t-1} \epsilon_t + o_P(1), \quad (6.93)$$

so dass es genügt, für alle  $m \in \{1, \dots, p\}$  die gleichmäßige Beschränktheit des zweiten Moments von

$$\frac{1}{\sqrt{nh}} \sum_{t=1}^n K\left(\frac{1}{h}\left(\frac{t}{n} - u\right)\right) \left[ \frac{1}{h} \left(\frac{t}{n} - u\right) \right] \tilde{X}_{t-m,n} \epsilon_t \quad (6.94)$$

zu zeigen. Setzt man  $\tilde{X}_{t-m,n} = \sum_{j=0}^{\infty} \varphi_j\left(\frac{t-m}{n}\right) \epsilon_{t-m-j}$  ein und beachtet, dass der Erwartungswert  $E[\epsilon_{t-m-j} \epsilon_{s-m-k} \epsilon_t \epsilon_s]$  nur ungleich null ist, wenn gleichzeitig  $t = s$  und  $k = j$

gilt, dann folgt

$$\begin{aligned} & E \left[ \left( \frac{1}{\sqrt{nh}} \sum_{t=1}^n K\left(\frac{1}{h}\left(\frac{t}{n} - u\right)\right) \left[ \frac{1}{h} \left( \frac{t}{n} - u \right) \right] \tilde{X}_{t-m,n}\epsilon_t \right)^2 \right] \\ &= \frac{\sigma_\epsilon^4}{nh} \sum_{t=1}^n K^2\left(\frac{1}{h}\left(\frac{t}{n} - u\right)\right) \left[ \frac{1}{h} \left( \frac{t}{n} - u \right) \right]^2 \sum_{j=0}^{\infty} \varphi_j^2\left(\frac{t-m}{n}\right). \end{aligned} \quad (6.95)$$

Die Behauptung ergibt sich dann aus der gleichmäßigen Beschränktheit von  $\sum_{j=0}^{\infty} \varphi_j^2\left(\frac{t-m}{n}\right)$  (siehe Abschnitt 7.2) sowie aus der Konvergenz

$$\frac{1}{nh} \sum_{t=1}^n K^2\left(\frac{1}{h}\left(\frac{t}{n} - u\right)\right) \left[ \frac{1}{h} \left( \frac{t}{n} - u \right) \right]^2 \xrightarrow{n \rightarrow \infty} \int K^2(s) s^2 ds, \quad (6.96)$$

wobei das Integral nach (K4) existiert.

zu (ii): Zunächst nutzt man die Cramér-Wold-Technik (siehe Satz 2.31), um das Problem der Asymptotik von  $\sqrt{nh}\underline{\tau}_{n,0}(u)$  auf Dimension eins zu reduzieren. Sie besagt nämlich, dass es hinreichend (und notwendig) ist, wenn für alle  $\underline{a} = (a_1, \dots, a_p)^T \in \mathbb{R}^p$  die (eindimensionale) Verteilungskonvergenz

$$\sqrt{nh}\underline{a}^T \underline{\tau}_{n,0}(u) \xrightarrow{D} \mathcal{N}(0, \underline{a}^T \underline{\Sigma}(u) \underline{a}) \quad (6.97)$$

erfüllt ist. Dabei genügt es, diese Konvergenz für alle Vektoren  $\underline{a} \in \mathbb{R}^p$  mit  $\|\underline{a}\| = 1$  nachzuweisen, weil sich jeder Vektor  $\underline{b} \in \mathbb{R}^p$ ,  $\underline{b} \neq 0$  durch Multiplikation mit  $\frac{1}{\|\underline{b}\|}$  auf eins normieren lässt und die Aussage trivialerweise gilt, wenn  $\underline{b}$  der Nullvektor ist. Beachtet man noch für  $l = 0$  die Aussage von Lemma 6.9, so folgt die Behauptung (ii) mit Satz 2.22 (Slutsky), wenn man für alle  $\underline{a} \in \mathbb{R}^p$  mit  $\|\underline{a}\| = 1$  die Verteilungskonvergenz

$$\frac{1}{\sqrt{nh}} \sum_{t=1}^n K\left(\frac{1}{h}\left(\frac{t}{n} - u\right)\right) \underline{a}^T \tilde{X}_{t-1} \epsilon_t \xrightarrow{D} \mathcal{N}(0, \underline{a}^T \underline{\Sigma}(u) \underline{a}) \quad (6.98)$$

zeigt. Dazu bestimmt man als erstes die Varianz des Terms auf der linken Seite von (6.98), welche im Folgenden mit  $V_n(u)$  bezeichnet wird. Wegen

$$\begin{aligned} E \left[ \underline{a}^T \tilde{X}_{t-1} \epsilon_t \right] &= \sum_{k=1}^p a_k E \left[ \tilde{X}_{t-k,n} \epsilon_t \right] \\ &= \sum_{k=1}^p a_k \sum_{j=0}^{\infty} \varphi_j\left(\frac{t-k}{n}\right) E \left[ \epsilon_{t-k-j} \epsilon_t \right] \end{aligned} \quad (6.99)$$

und  $E[\epsilon_{t-k-j}\epsilon_t] = 0$  für  $k \in \{1, \dots, p\}$  und  $j \in \mathbb{N}_0$  ist  $V_n(u)$  gleich dem zweiten Moment der linken Seite von (6.98), d.h. es gilt

$$\begin{aligned} V_n(u) &= E \left[ \left( \frac{1}{\sqrt{nh}} \sum_{t=1}^n K\left(\frac{1}{h}\left(\frac{t}{n} - u\right)\right) \underline{a}^T \tilde{X}_{t-1} \epsilon_t \right)^2 \right] \\ &= \frac{1}{nh} \sum_{t=1}^n K^2\left(\frac{1}{h}\left(\frac{t}{n} - u\right)\right) E \left[ \underline{a}^T \tilde{X}_{t-1} \tilde{X}_{t-1}^T \underline{a} \epsilon_t^2 \right], \end{aligned} \quad (6.100)$$

wobei benutzt wurde, dass  $E[\underline{a}^T \tilde{X}_{t-1} \tilde{X}_{s-1}^T \underline{a} \epsilon_t \epsilon_s] = 0$  für  $s \neq t$  gilt. Im Folgenden sei mit  $\{\mathcal{F}_t^n\}_{t=-\infty}^n$  für  $n \in \mathbb{N}$  die natürliche Filtration von  $\{\tilde{X}_{t,n}\}_{t=-\infty}^n$  bezeichnet, d.h.

$$\mathcal{F}_t^n := \sigma \left( \tilde{X}_{t,n} : t \in \mathbb{Z}, t \leq n \right). \quad (6.101)$$

Um den Erwartungswert in (6.100) weiter ausrechnen zu können, fügt man in ihm einen bedingten Erwartungswert bzgl.  $\mathcal{F}_{t-1}^n$  ein. Man benutzt dann die üblichen Rechenregeln für bedingte Erwartungen, indem man den  $\mathcal{F}_{t-1}^n$ -messbaren Anteil  $\underline{a}^T \tilde{X}_{t-1} \tilde{X}_{t-1}^T \underline{a}$  hervorzieht und der bedingte Erwartungswert  $E[\epsilon_t^2 | \mathcal{F}_{t-1}^n]$  gleich seinem unbedingten ist, weil  $\epsilon_t^2$  unabhängig von  $\mathcal{F}_{t-1}^n$  ist. Es gilt also

$$\begin{aligned} E \left[ \underline{a}^T \tilde{X}_{t-1} \tilde{X}_{t-1}^T \underline{a} \epsilon_t^2 \right] &= E \left[ E \left[ \underline{a}^T \tilde{X}_{t-1} \tilde{X}_{t-1}^T \underline{a} \epsilon_t^2 | \mathcal{F}_{t-1}^n \right] \right] \\ &= E \left[ \underline{a}^T \tilde{X}_{t-1} \tilde{X}_{t-1}^T \underline{a} E \left[ \epsilon_t^2 | \mathcal{F}_{t-1}^n \right] \right] \\ &= \sigma_\epsilon^2 E \left[ \underline{a}^T \tilde{X}_{t-1} \tilde{X}_{t-1}^T \underline{a} \right]. \end{aligned} \quad (6.102)$$

Zieht man nun den Erwartungswert in die Einträge der Matrix  $\tilde{X}_{t-1} \tilde{X}_{t-1}^T$  herein und bedenkt, dass für  $m, q \in \{1, \dots, p\}$  der Erwartungswert  $E[\epsilon_{t-m-j}\epsilon_{t-q-k}]$  nur für  $m + j = q + k$  ungleich null ist, so folgt mit der Bezeichnung

$$\mathbf{\Gamma}_{t,n} := \left( \begin{array}{c} \sum_{j=0}^{\infty} \varphi_j\left(\frac{t-m}{n}\right) \varphi_{j+m-q}\left(\frac{t-q}{n}\right) \\ m, q = 1, \dots, p \end{array} \right) \in \text{Mat}_{p,p}(\mathbb{R}) \quad (6.103)$$

für obigen Erwartungswert die Gleichheit

$$E \left[ \underline{a}^T \tilde{X}_{t-1} \tilde{X}_{t-1}^T \underline{a} \right] = \sigma_\epsilon^4 \underline{a}^T \mathbf{\Gamma}_{t,n} \underline{a}. \quad (6.104)$$

Insgesamt erhält man für die Varianz der linken Seite von (6.98) die Darstellung

$$V_n(u) = \frac{\sigma_\epsilon^4}{nh} \sum_{t=1}^n K^2\left(\frac{1}{h}\left(\frac{t}{n} - u\right)\right) \underline{a}^T \mathbf{\Gamma}_{t,n} \underline{a}, \quad (6.105)$$

welche mit wachsendem  $n$  gegen

$$\sigma_\epsilon^4 \underline{a}^T \left( \int K^2(s) ds \right) \mathbf{\Gamma}(u) \underline{a} = \underline{a}^T \mathbf{\Sigma}(u) \underline{a} \quad (6.106)$$

konvergiert. Weil die Matrix  $\mathbf{\Sigma}(u)$  als Vielfaches der positiv definiten Matrix  $\mathbf{\Gamma}(u)$  selbst ebenfalls positiv definit ist und der Vektor  $\underline{a}$  wegen  $\|\underline{a}\| = 1$  nicht der Nullvektor ist, gilt

$$\underline{a}^T \mathbf{\Sigma}(u) \underline{a} > 0. \quad (6.107)$$

Daher existiert ein  $n_0 \in \mathbb{N}$ , so dass für alle  $n \geq n_0$  die Ungleichheiten

$$V_n(u) > \frac{1}{2} \underline{a}^T \mathbf{\Sigma}(u) \underline{a} > 0 \quad (6.108)$$

erfüllt sind. Im Folgenden sei  $n \geq n_0$  angenommen, was die strikte Positivität der Varianz  $V_n(u)$  sicher stellt, welche es wiederum erlaubt durch  $V_n(u)$  zu teilen.

Als nächstes wollen wir den Zentralen Grenzwertsatz für Martingaldifferenzschemata (siehe Satz 2.32 oder [17], Korollar 6) benutzen, um die zu zeigende Verteilungskonvergenz aus (6.98) nachzuweisen. Dazu normiert man die Summanden des Ausdrucks auf der linken Seite von (6.98), indem man sie durch die Wurzel der Varianz  $\sqrt{V_n(u)}$  teilt und gemäß

$$\xi_{t,n}(u) := \frac{1}{\sqrt{V_n(u)}} \frac{1}{\sqrt{nh}} K\left(\frac{1}{h}\left(\frac{t}{n} - u\right)\right) \underline{a}^T \tilde{X}_{t-1} \epsilon_t \quad (6.109)$$

definiert. Man betrachtet nun für festes  $u \in (0, 1)$  und  $n \in \mathbb{N}$  das Dreiecksschema

$$\begin{array}{cccc} \xi_{1,1}(u) & & & \\ \xi_{1,2}(u) & \xi_{2,2}(u) & & \\ \vdots & \vdots & \ddots & \\ \xi_{1,n}(u) & \xi_{2,n}(u) & \cdots & \xi_{n,n}(u) \end{array} \quad (6.110)$$

und weist nacheinander die Bedingungen des Zentralen Grenzwertsatzes nach. Als erstes zeigt man, dass  $\{\xi_{t,n}(u), \mathcal{F}_t^n\}_{t=1}^n$  für alle  $n \in \mathbb{N}$  eine quadratisch integrierbare Martingaldifferenz ist. Dazu bildet man zunächst für  $t \in \{1, \dots, n\}$  den bedingten Erwartungswert von  $\xi_{t,n}(u)$  bzgl.  $\mathcal{F}_{t-1}^n$  und erhält

$$\begin{aligned} E[\xi_{t,n}(u) | \mathcal{F}_{t-1}^n] &= \frac{1}{\sqrt{V_n(u)}} \frac{1}{\sqrt{nh}} K\left(\frac{1}{h}\left(\frac{t}{n} - u\right)\right) E\left[\underline{a}^T \tilde{X}_{t-1} \epsilon_t | \mathcal{F}_{t-1}^n\right] \\ &= \frac{1}{\sqrt{V_n(u)}} \frac{1}{\sqrt{nh}} K\left(\frac{1}{h}\left(\frac{t}{n} - u\right)\right) \underline{a}^T \tilde{X}_{t-1} E[\epsilon_t] \\ &= 0, \end{aligned} \quad (6.111)$$

wobei die  $\mathcal{F}_{t-1}^n$ -Messbarkeit von  $\tilde{X}_{t-1}$  und die Zentriertheit des weißen Rauschens benutzt wurde. Die quadratische Integrierbarkeit von  $\xi_{t,n}(u)$  folgt dann mit (6.104) und Voraussetzung (S1) aus

$$\begin{aligned} E [\xi_{t,n}^2(u)] &= \frac{1}{V_n(u)} \frac{1}{nh} K^2\left(\frac{1}{h}\left(\frac{t}{n} - u\right)\right) E \left[ \underline{a}^T \tilde{X}_{t-1} \tilde{X}_{t-1}^T \underline{a} \epsilon_t^2 \right] \\ &= \frac{1}{V_n(u)} \frac{1}{nh} K^2\left(\frac{1}{h}\left(\frac{t}{n} - u\right)\right) \sigma_\epsilon^4 \underline{a}^T \mathbf{\Gamma}_{t,n} \underline{a} \\ &< \infty, \end{aligned} \quad (6.112)$$

so dass nachgewiesen ist, dass es sich für  $n \in \mathbb{N}$  bei  $\{\xi_{t,n}(u), \mathcal{F}_t^n\}_{t=1}^n$  um ein quadratisch integrierbare Martingaldifferenzen handelt. Für die nächste zu zeigende Eigenschaft genügt es, wenn es ein  $n_0 \in \mathbb{N}$  gibt, so dass sie für alle  $n \geq n_0$  gilt. Sei  $n_0 \in \mathbb{N}$  so groß, dass es die Ungleichung (6.108) erfüllt. Dann folgt nach Konstruktion von  $\xi_{t,n}(u)$  für  $n \geq n_0$  die geforderte Gleichheit aus

$$\begin{aligned} \sum_{t=1}^n E [\xi_{t,n}^2(u)] &= \sum_{t=1}^n \frac{1}{V_n(u)} \frac{1}{nh} K^2\left(\frac{1}{h}\left(\frac{t}{n} - u\right)\right) \sigma_\epsilon^4 \underline{a}^T \mathbf{\Gamma}_{t,n} \underline{a} \\ &= \frac{V_n(u)}{V_n(u)} \\ &= 1. \end{aligned} \quad (6.113)$$

Um die stochastische Konvergenz

$$\sum_{t=1}^n E [\xi_{t,n}^2(u) | \mathcal{F}_{t-1}^n] \xrightarrow{P} 1 \quad (6.114)$$

zu zeigen, setzt man für  $\xi_{t,n}^2(u)$  ein, rechnet ähnlich zur Gleichung (6.102) die bedingten Erwartungswerte aus und erhält

$$\begin{aligned} \sum_{t=1}^n E [\xi_{t,n}^2(u) | \mathcal{F}_{t-1}^n] &= \frac{1}{V_n(u)} \frac{1}{nh} \sum_{t=1}^n K^2\left(\frac{1}{h}\left(\frac{t}{n} - u\right)\right) E \left[ \underline{a}^T \tilde{X}_{t-1} \tilde{X}_{t-1}^T \underline{a} \epsilon_t^2 \mid \mathcal{F}_{t-1}^n \right] \\ &= \frac{1}{V_n(u)} \underline{a}^T \left( \sigma_\epsilon^2 \frac{1}{nh} \sum_{t=1}^n K^2\left(\frac{1}{h}\left(\frac{t}{n} - u\right)\right) \tilde{X}_{t-1} \tilde{X}_{t-1}^T \right) \underline{a}. \end{aligned} \quad (6.115)$$

Die nachzuweisende Konvergenz in (6.114) folgt dann zum einen aus dem asymptotischen Verhalten von  $V_n(u)$  gemäß

$$V_n(u) \xrightarrow{n \rightarrow \infty} \sigma_\epsilon^4 \underline{a}^T \left( \int K^2(s) ds \right) \mathbf{\Gamma}(u) \underline{a}, \quad (6.116)$$

und zum anderen aus der stochastischen Konvergenz

$$\sigma_\epsilon^2 \frac{1}{nh} \sum_{t=1}^n K^2\left(\frac{1}{h}\left(\frac{t}{n} - u\right)\right) \underline{\tilde{X}}_{t-1} \underline{\tilde{X}}_{t-1}^T \xrightarrow{P} \sigma_\epsilon^4 \left( \int K^2(s) ds \right) \Gamma(u), \quad (6.117)$$

welche sich völlig analog zu Satz 6.6 beweisen lässt, man jedoch hier die Bedingung  $\int K^4(x) dx < \infty$  aus (K4) benötigt. Abschließend ist noch der Nachweis der bedingten Lindeberg-Bedingung zu erbringen. Dafür ist zu zeigen, dass für alle  $\epsilon > 0$  die stochastische Konvergenz

$$\sum_{j=0}^n E \left[ \xi_{t,n}^2(u) I(|\xi_{t,n}(u)| > \delta) | \mathcal{F}_{t-1}^n \right] \xrightarrow{P} 0 \quad (6.118)$$

erfüllt ist. Dazu betrachtet man zunächst  $\xi_{t,n}^2(u)$  einzeln und schätzt ab. Es gilt

$$\begin{aligned} \xi_{t,n}^2(u) &= \frac{1}{V_n(u)} \frac{1}{nh} K^2\left(\frac{1}{h}\left(\frac{t}{n} - u\right)\right) \underline{a}^T \underline{\tilde{X}}_{t-1} \underline{\tilde{X}}_{t-1}^T \underline{a} \epsilon_t^2 \\ &\leq \frac{2}{\underline{a}^T \underline{\Sigma}(u) \underline{a}} \frac{\|K\|_\infty}{nh} K\left(\frac{1}{h}\left(\frac{t}{n} - u\right)\right) \underline{a}^T \underline{\tilde{X}}_{t-1} \underline{\tilde{X}}_{t-1}^T \underline{a} \epsilon_t^2, \end{aligned} \quad (6.119)$$

wobei der Bruch  $\frac{1}{V_n(u)}$  mit Ungleichung (6.108) und einer der Faktoren  $K\left(\frac{1}{h}\left(\frac{t}{n} - u\right)\right)$  durch seine Supremumsnorm  $\|K\|_\infty$  nach oben abgeschätzt wurden. Interpretiert man  $\underline{a}^T \underline{\tilde{X}}_{t-1}$  und  $\underline{\tilde{X}}_{t-1}^T \underline{a}$  als Skalarprodukte, so folgt mit der Cauchy-Schwarz-Ungleichung (2.39)

$$\begin{aligned} \underline{a}^T \underline{\tilde{X}}_{t-1} \underline{\tilde{X}}_{t-1}^T \underline{a} &\leq \|\underline{a}\|^2 \|\underline{\tilde{X}}_{t-1}\|^2 \\ &= \|\underline{\tilde{X}}_{t-1}\|^2, \end{aligned} \quad (6.120)$$

weil  $\|\underline{a}\| = 1$  gilt. Für eine geeignete Konstante  $D < \infty$  ergibt sich daher

$$\xi_{t,n}^2(u) \leq D \frac{1}{nh} K\left(\frac{1}{h}\left(\frac{t}{n} - u\right)\right) \|\underline{\tilde{X}}_{t-1}\|^2 \epsilon_t^2. \quad (6.121)$$

Als nächstes untersucht man die bedingten Erwartungswerte in der Lindeberg-Bedingung (6.118) näher und schätzt diese weiter ab. Einsetzen von Ungleichung (6.121) für  $\xi_{t,n}^2(u)$  und  $|\xi_{t,n}(u)|$  ergibt

$$\begin{aligned} &E \left[ \xi_{t,n}^2(u) I(|\xi_{t,n}(u)| > \delta) | \mathcal{F}_{t-1}^n \right] \\ &\leq E \left[ D \frac{1}{nh} K\left(\frac{1}{h}\left(\frac{t}{n} - u\right)\right) \|\underline{\tilde{X}}_{t-1}\|^2 \epsilon_t^2 I \left( D^{1/2} (nh)^{-1/2} \|K\|_\infty^{1/2} \|\underline{\tilde{X}}_{t-1}\| |\epsilon_t| > \delta \right) \middle| \mathcal{F}_{t-1}^n \right], \end{aligned} \quad (6.122)$$

wobei in  $I(\cdot)$  nochmals  $K\left(\frac{1}{h}\left(\frac{t}{n} - u\right)\right)$  durch  $\|K\|_\infty$  abgeschätzt wurde. Mit nicht-negativen reellen Zahlen  $a$ ,  $b$  und  $c$  gilt für Indikatorvariablen die Ungleichung

$$I(ab > c) \leq I(a > \sqrt{c}) + I(b > \sqrt{c}), \quad (6.123)$$

woraus unter Beachtung der  $\mathcal{F}_{t-1}^n$ -Messbarkeit von  $\|\tilde{\underline{X}}_{t-1}\|^2 I(\|\tilde{\underline{X}}_{t-1}\| > c)$  die Abschätzung

$$\begin{aligned} & E [\xi_{t,n}^2(u) I(|\xi_{t,n}(u)| > \delta) | \mathcal{F}_{t-1}^n] \\ & \leq D \frac{1}{nh} K\left(\frac{1}{h}\left(\frac{t}{n} - u\right)\right) \|\tilde{\underline{X}}_{t-1}\|^2 E [\epsilon_t^2 I(|\epsilon_t| > \delta^{1/2} D^{-1/4} (nh)^{1/4} \|K\|_\infty^{-1/4}) | \mathcal{F}_{t-1}^n] \quad (6.124) \\ & + D \frac{1}{nh} K\left(\frac{1}{h}\left(\frac{t}{n} - u\right)\right) \|\tilde{\underline{X}}_{t-1}\|^2 I\left(\|\tilde{\underline{X}}_{t-1}\| > \delta^{1/2} D^{-1/4} (nh)^{1/4} \|K\|_\infty^{-1/4}\right) E [\epsilon_t^2 | \mathcal{F}_{t-1}^n] \end{aligned}$$

folgt. Aufsummieren ergibt dann

$$\begin{aligned} & \sum_{t=1}^n E [\xi_{t,n}^2(u) I(|\xi_{t,n}(u)| > \delta) | \mathcal{F}_{t-1}^n] \\ & \leq \frac{D}{nh} \sum_{t=1}^n K\left(\frac{1}{h}\left(\frac{t}{n} - u\right)\right) \|\tilde{\underline{X}}_{t-1}\|^2 E [\epsilon_t^2 I(|\epsilon_t| > \delta^{1/2} D^{-1/4} (nh)^{1/4} \|K\|_\infty^{-1/4})] \quad (6.125) \end{aligned}$$

$$+ \frac{\sigma_\epsilon^2 D}{nh} \sum_{t=1}^n K\left(\frac{1}{h}\left(\frac{t}{n} - u\right)\right) \|\tilde{\underline{X}}_{t-1}\|^2 I\left(\|\tilde{\underline{X}}_{t-1}\| > \delta^{1/2} D^{-1/4} (nh)^{1/4} \|K\|_\infty^{-1/4}\right) \quad (6.126)$$

wobei benutzt wurde, dass  $\epsilon_t^2$  und  $\epsilon_t^2 I(|\epsilon_t| > c)$  unabhängig von  $\mathcal{F}_{t-1}^n$  sind, womit die bedingten Erwartungswerte zu unbedingten werden. Die Richtigkeit der bedingten Lindeberg-Bedingung (6.118) ergibt sich also nun aus der stochastischen Konvergenz der beiden Summenterme (6.125) und (6.126) gegen null. Dazu betrachtet man als erstes die Erwartungswerte in (6.125). Wegen  $I^2(\cdot) = I(\cdot)$  und  $E[I(\cdot)] = P(\cdot)$  folgt mit der Cauchy-Schwarz-Ungleichung (2.39)

$$\begin{aligned} & E [\epsilon_t^2 I(|\epsilon_t| > \delta^{1/2} D^{-1/4} (nh)^{1/4} \|K\|_\infty^{-1/4})] \\ & \leq \sqrt{E[\epsilon_t^4]} \sqrt{P\left(|\epsilon_t| > \delta^{1/2} D^{-1/4} (nh)^{1/4} \|K\|_\infty^{-1/4}\right)} \quad (6.127) \end{aligned}$$

und mit der Markov-Ungleichung (2.37), dass

$$P\left(|\epsilon_t| > \delta^{1/2} D^{-1/4} (nh)^{1/4} \|K\|_\infty^{-1/4}\right) \leq \frac{E[|\epsilon_t|]}{\delta^{1/2} D^{-1/4} (nh)^{1/4} \|K\|_\infty^{-1/4}} \quad (6.128)$$

gilt. Weil der Zähler  $E[|\epsilon_t|]$  wieder mit Cauchy-Schwarz nach oben durch  $\sigma_\epsilon < \infty$  beschränkt ist und der Nenner mit wachsendem  $n$  gegen  $\infty$  strebt, konvergiert der Bruch gegen null. Die Konvergenz

$$E [\epsilon_t^2 I(|\epsilon_t| > \delta^{1/2} D^{-1/4} (nh)^{1/4} \|K\|_\infty^{-1/4})] \xrightarrow[n \rightarrow \infty]{} 0 \quad (6.129)$$

folgt dann unabhängig von  $t$  aus der Annahme, dass die vierten Momente des weißen Rauschen existieren. Beachtet man noch, dass  $\|\tilde{\underline{X}}_{t-1}\|^2$  die Summe der Hauptdiagonalelemente

der Matrix  $\mathbf{\Gamma}_{t,n}$  ist, so folgt mit Satz 6.6 die Konvergenz nach Wahrscheinlichkeit

$$\frac{1}{nh} \sum_{t=1}^n K\left(\frac{1}{h}\left(\frac{t}{n} - u\right)\right) \|\tilde{\underline{X}}_{t-1}\|^2 \xrightarrow{P} \sigma_\epsilon^2 \left( \int K(s) ds \right) \text{tr}(\mathbf{\Gamma}(u)), \quad (6.130)$$

wobei  $\text{tr}(A)$  die Spur einer Matrix  $A$  bezeichne und das Integral nach (K1) endlich ist. Die stochastische Konvergenz von (6.125) folgt schließlich aus der von  $t$  unabhängigen Konvergenz (6.129) und aus (6.130). Für den zweiten Summenterm (6.126) nimmt man den Erwartungswert seines Absolutbetrags und zeigt, dass dieser gegen null konvergiert. Da diese Summe aber ausschließlich aus nicht-negativen Anteilen besteht, genügt es hier seinen Erwartungswert zu betrachten. Es gilt

$$\begin{aligned} & E \left[ \frac{\sigma_\epsilon^2 D}{nh} \sum_{t=1}^n K\left(\frac{1}{h}\left(\frac{t}{n} - u\right)\right) \|\tilde{\underline{X}}_{t-1}\|^2 I \left( \|\tilde{\underline{X}}_{t-1}\| > \delta^{1/2} D^{-1/4} (nh)^{1/4} \|K\|_\infty^{-1/4} \right) \right] \\ &= \frac{\sigma_\epsilon^2 D}{nh} \sum_{t=1}^n K\left(\frac{1}{h}\left(\frac{t}{n} - u\right)\right) E \left[ \|\tilde{\underline{X}}_{t-1}\|^2 I \left( \|\tilde{\underline{X}}_{t-1}\| > \delta^{1/2} D^{-1/4} (nh)^{1/4} \|K\|_\infty^{-1/4} \right) \right] \end{aligned}$$

und wie oben folgt mit Cauchy-Schwarz (2.39)

$$\begin{aligned} & E \left[ \|\tilde{\underline{X}}_{t-1}\|^2 I \left( \|\tilde{\underline{X}}_{t-1}\| > \delta^{1/2} D^{-1/4} (nh)^{1/4} \|K\|_\infty^{-1/4} \right) \right] \\ & \leq \sqrt{E \left[ \|\tilde{\underline{X}}_{t-1}\|^4 \right]} \sqrt{P \left( \|\tilde{\underline{X}}_{t-1}\| > \delta^{1/2} D^{-1/4} (nh)^{1/4} \|K\|_\infty^{-1/4} \right)} \end{aligned} \quad (6.131)$$

und mit Markov (2.37)

$$P \left( \|\tilde{\underline{X}}_{t-1}\| \geq \delta^{1/2} D^{-1/4} (nh)^{1/4} \|K\|_\infty^{-1/4} \right) \leq \frac{E \left[ \|\tilde{\underline{X}}_{t-1}\| \right]}{\delta^{1/2} D^{-1/4} (nh)^{1/4} \|K\|_\infty^{-1/4}}. \quad (6.132)$$

Lemma 7.6 liefert die gleichmäßige Beschränktheit von  $E[\|\tilde{\underline{X}}_{t-1}\|^4]$ , woraus auch die von  $E[\|\tilde{\underline{X}}_{t-1}\|]$  folgt. Weil der Nenner in (6.132) gegen unendlich strebt und

$$\frac{1}{nh} \sum_{t=1}^n K\left(\frac{1}{h}\left(\frac{t}{n} - u\right)\right) \xrightarrow{n \rightarrow \infty} \int K(s) ds < \infty \quad (6.133)$$

gilt, ergibt sich schließlich die Konvergenz nach Wahrscheinlichkeit gegen null von (6.126), so dass die Gültigkeit der bedingten Lindeberg-Bedingung (6.118) gezeigt ist.

Da nun alle Voraussetzungen von Satz 2.32 nachgewiesen sind, folgt die Konvergenz in Verteilung

$$\sum_{t=1}^n \xi_{t,n}(u) \xrightarrow{D} \mathcal{N}(0, 1), \quad (6.134)$$

welche äquivalent zu

$$\sqrt{nh}\underline{a}^T \underline{T}_{n,0}(u) \xrightarrow{D} \mathcal{N}(0, \underline{a}^T \underline{\Sigma}(u) \underline{a}) \quad (6.135)$$

ist, was aber wiederum nach Cramér-Wold eine notwendige und hinreichende Bedingung für die behauptete Verteilungskonvergenz

$$\sqrt{nh}\underline{T}_{n,0}(u) \xrightarrow{D} \mathcal{N}(0_p, \underline{\Sigma}(u)) \quad (6.136)$$

darstellt.  $\square$

Abschließend tragen wir im nun folgenden Satz die erzielten Ergebnisse dieses Kapitels über das asymptotische Verhalten zeitvariabler autoregressiver Prozesse  $p$ -ter Ordnung mit hinreichend glatten Koeffizientenfunktionen zusammen.

**Theorem 6.11 (Asymptotische Normalität des Schätzfehlers  $\widehat{\underline{\alpha}}(u) - \underline{\alpha}(u)$ ).**

Neben den Annahmen von Satz 3.2 gelten die Voraussetzungen (K1) - (K4) für den Kern  $K$  und für  $k = 1, \dots, p$  seien die Koeffizientenfunktionen  $\alpha_k(\cdot)$  zweimal stetig differenzierbar auf  $(0, 1)$  mit gleichmäßig beschränkten Ableitungen auf  $[0, 1]$ . Für ein  $\delta > 0$  gelten die Bedingungen (S1) und (S2) sowie  $h \rightarrow 0$ ,  $nh^2 \rightarrow \infty$  und  $n^{1/2}h^{5/2} \rightarrow C$  für eine Konstante  $C < \infty$ . Ist  $\Gamma(u)$  invertierbar, dann folgt unter der Annahme  $\kappa_4 := E[\epsilon_t^4] < \infty$  mit der Bezeichnung

$$\underline{\Sigma}_\alpha(u) := \left( \int K^2(r) dr \right) \Gamma^{-1}(u) \quad (6.137)$$

für den Schätzfehler  $\widehat{\underline{\alpha}}(u) - \underline{\alpha}(u)$  die Konvergenz in Verteilung

$$\sqrt{nh} [\widehat{\underline{\alpha}}(u) - \underline{\alpha}(u) - \underline{B}_n(u)] \xrightarrow{D} \mathcal{N}(0_p, \underline{\Sigma}_\alpha(u)). \quad (6.138)$$

*Beweis.*

Da die Voraussetzungen von Lemma 4.3 erfüllt sind, gilt für den Schätzfehler die Zerlegung

$$\widehat{\underline{\alpha}}(u) - \underline{\alpha}(u) = \underline{B}_n(u) + \underline{T}_n(u) + \underline{R}_n(u), \quad (6.139)$$

woraus zusammen mit der stochastischen Konvergenz des Restterms gegen null (Satz 6.8) und der Konvergenz des Bias-Terms (Satz 6.7)

$$\begin{aligned} \sqrt{nh} [\widehat{\underline{\alpha}}(u) - \underline{\alpha}(u) - \underline{B}_n(u)] &= \sqrt{nh}\underline{T}_n(u) + o_P(1) \\ &= \mathbf{E}_0 \mathbf{S}_n^{-1}(u) \sqrt{nh}\underline{T}_n(u) + o_P(1) \end{aligned} \quad (6.140)$$

folgt. Wie im Beweis von Satz 6.7 gilt

$$\mathbf{S}_n^{-1}(u) \xrightarrow{P} \sigma_\epsilon^{-2} \begin{pmatrix} \Gamma^{-1}(u) & \mathbf{O}_{p \times p} \\ \mathbf{O}_{p \times p} & \mu_K^{-2} \Gamma^{-1}(u) \end{pmatrix}, \quad (6.141)$$

so dass

$$\begin{aligned} \mathbf{E}_0 \mathbf{S}_n^{-1}(u) &\xrightarrow{P} (\mathbf{I}_{p \times p}, \mathbf{O}_{p \times p}) \sigma_\epsilon^{-2} \begin{pmatrix} \mathbf{\Gamma}^{-1}(u) & \mathbf{O}_{p \times p} \\ \mathbf{O}_{p \times p} & \mu_K^{-2} \mathbf{\Gamma}^{-1}(u) \end{pmatrix} \\ &= (\sigma_\epsilon^{-2} \mathbf{\Gamma}^{-1}(u), \mathbf{O}_{p \times p}) \end{aligned} \quad (6.142)$$

folgt. Die  $p \times p$ -Nullmatrix im Grenzwert nach Wahrscheinlichkeit von  $\mathbf{E}_0 \mathbf{S}_n^{-1}(u)$  sorgt dafür, dass die zweite Hälfte  $\underline{\mathcal{I}}_{n,1}(u)$  des  $2p$ -dimensionalen Vektors  $\underline{\mathcal{I}}_n(u)$  keinen Einfluss auf die Verteilungskonvergenz von  $\mathbf{E}_0 \mathbf{S}_n^{-1}(u) \sqrt{nh} \underline{\mathcal{I}}_n(u)$  hat, weil  $\sqrt{nh} \underline{\mathcal{I}}_{n,1}(u)$  nach Satz 6.10 (i) stochastisch beschränkt bleibt. Satz 6.10 (ii) liefert

$$\sqrt{nh} \underline{\mathcal{I}}_{n,0}(u) \xrightarrow{D} \mathcal{N}(0_p, \mathbf{\Sigma}(u)) \quad (6.143)$$

mit  $\mathbf{\Sigma}(u) = \sigma_\epsilon^4 (\int K^2(r) dr) \mathbf{\Gamma}(u)$ , woraus schließlich mit (6.142) und Satz 2.22 (Slutsky) die Konvergenz in Verteilung

$$\begin{aligned} \sqrt{nh} [\hat{\alpha}(u) - \underline{\alpha}(u) - \underline{B}_n(u)] &\xrightarrow{D} \mathcal{N}(0_p, \sigma_\epsilon^{-2} \mathbf{\Gamma}^{-1}(u) \mathbf{\Sigma}(u) \sigma_\epsilon^{-2} \mathbf{\Gamma}^{-1}(u)) \\ &= \mathcal{N}(0_p, \mathbf{\Sigma}_\alpha(u)) \end{aligned} \quad (6.144)$$

folgt, womit die Behauptung gezeigt ist.  $\square$

# Kapitel 7

## Anhang

### 7.1 Ungleichungen und Doppelsummen

An dieser Stelle beweisen wir einige Aussagen, welche wir bei den Beweisen im folgenden Abschnitt 7.2 häufig benutzen werden.

**Hilfssatz 7.1 (Ungleichungen).**

Es seien  $a, b \in \mathbb{R}$  sowie  $\{a_n\}_{n \in \mathbb{N}_0}$  und  $\{b_n\}_{n \in \mathbb{N}_0}$  absolut summierbare reelle Zahlenfolgen. Dann gelten die Ungleichungen

$$(i) \quad \sum_{j=0}^{\infty} |a_j b_j| \leq \left( \sum_{j=0}^{\infty} |a_j| \right) \left( \sum_{j=0}^{\infty} |b_j| \right), \quad (7.1)$$

$$(ii) \quad \left( \sum_{j=0}^{\infty} a_j b_j \right)^2 \leq \left( \sum_{j=0}^{\infty} a_j^2 \right) \left( \sum_{j=0}^{\infty} b_j^2 \right). \quad (7.2)$$

*Beweis.*

zu (i): Auf der rechten Seite der Ungleichung kommen zusätzlich zur linken noch die gemischten Terme hinzu. Weil alle vorkommenden Summanden nicht-negativ sind, folgt die Behauptung.

zu (ii): Spezialfall der Cauchy-Schwarz-Ungleichung (2.39). □

**Hilfssatz 7.2 (Doppelsummen).**

Sei  $\{a_n\}_{n \in \mathbb{N}_0}$  eine Folge nicht-negativer reeller Zahlen, welche

$$\sum_{s=1}^{\infty} s a_s < \infty \quad (7.3)$$

erfüllt. Weiter sei  $\delta > 0$  und  $a_j := 0$  für  $j < 0$ . Dann gilt für ein  $h \in \mathbb{Z}$  mit  $h \leq 0$

$$(i) \quad \sum_{r=1}^{\infty} \sum_{j=0}^{\infty} a_{j+h+r} = \sum_{j=0}^{\infty} (j-h)a_j, \quad (7.4)$$

$$(ii) \quad \sum_{r=1}^{\infty} \sum_{j=1}^{\infty} a_{j+h-r} \frac{1}{j^{1+\delta}} = \sum_{j=0}^{\infty} \left( a_j \sum_{r=1}^{\infty} \frac{1}{(j-h+r)^{1+\delta}} \right) \quad (7.5)$$

sowie für  $h \in \mathbb{Z}$  und  $r \in \mathbb{N}$

$$(iii) \quad \sum_{j=0}^{\infty} \sum_{s=j+1}^{\infty} a_{s+h+r} = \sum_{j=1}^{\infty} j a_{j+h+r}. \quad (7.6)$$

*Beweis.*

zu (i): Beachtet man, dass wegen  $a_j := 0$  für  $j < 0$  im Fall  $h+r < 0$  auch  $a_{j+h+r} = 0$  für  $j < -h-r$  gilt, so folgt mit der Bezeichnung  $a \wedge b := \max\{a, b\}$  die Behauptung aus der Rechnung

$$\begin{aligned} \sum_{r=1}^{\infty} \sum_{j=0}^{\infty} a_{j+h+r} &= \sum_{r=1}^{\infty} \sum_{j=(-h-r) \wedge 0}^{\infty} a_{j+h+r} \\ &= \sum_{r=1}^{-h} \sum_{j=-h-r}^{\infty} a_{j+h+r} + \sum_{r=-h+1}^{\infty} \sum_{j=0}^{\infty} a_{j+h+r} \\ &= \sum_{r=1}^{-h} \sum_{j=0}^{\infty} a_j + \sum_{r=1}^{\infty} \sum_{j=0}^{\infty} a_{j+r} \\ &= -h \sum_{j=0}^{\infty} a_j + \sum_{j=0}^{\infty} j a_j \\ &= \sum_{j=0}^{\infty} (j-h)a_j. \end{aligned} \quad (7.7)$$

zu (ii): Mit  $a_{j+h-r} = 0$  für  $j < -h+r$  und einer Indexverschiebung folgt

$$\begin{aligned} \sum_{r=1}^{\infty} \sum_{j=1}^{\infty} a_{j+h-r} \frac{1}{j^{1+\delta}} &= \sum_{r=1}^{\infty} \sum_{j=-h+r}^{\infty} a_{j+h-r} \frac{1}{j^{1+\delta}} \\ &= \sum_{r=1}^{\infty} \left( \sum_{j=0}^{\infty} a_j \frac{1}{(j-h+r)^{1+\delta}} \right). \end{aligned} \quad (7.8)$$

zu (iii): Folgt sofort, wenn man beachtet, dass die Summe über  $s$  mit wachsendem  $j$  immer einen Summanden später anfängt.  $\square$

## 7.2 Gleichmäßige Beschränktheit

Hier beweisen wir einige Aussagen, welche wir aufgrund der Übersichtlichkeit der Beweise in den Kapiteln 5 und 6 in den Anhang ausgelagert haben. Dabei weisen wir die gleichmäßige Beschränktheit gewisser Ausdrücke nach, welche wir dort gebraucht haben. In den folgenden Lemmata benötigt man dazu die Summierbarkeitsbedingungen (S1) und (S2) (siehe unten), welche für  $0 \leq k \leq 2 + \delta$  nach dem Majorantenkriterium

$$\sup_{t \leq n} \sum_{j=0}^{\infty} j^k |\varphi_j(\frac{t}{n})| < \infty, \quad \sup_{t \leq n} \sum_{j=0}^{\infty} j^k |\varphi_j^2(\frac{t}{n})| < \infty \quad (7.9)$$

bzw.

$$\sup_{t \leq n} \sum_{j=0}^{\infty} j^k |\Delta_t[\varphi_j(\frac{t}{n})]| = o(1) \quad (7.10)$$

implizieren. Jene Eigenschaften werden in den Beweisen der nun folgenden Hilfssätze benutzt, ohne dass darauf besonders hingewiesen wird.

### Lemma 7.3 (Zu Lemma 5.4 und Lemma 6.4).

Die Folge auf  $(0, 1)$  differenzierbarer Funktionen  $\{\varphi_j(\cdot) | \varphi_j : (-\infty, 1] \rightarrow \mathbb{R}\}_{j=0}^{\infty}$  aus Lemma 3.7 erfülle die Bedingung

$$(S1) \quad \sup_{t \leq n} \sum_{j=0}^{\infty} j^{2+\delta} |\varphi_j(\frac{t}{n})| < \infty. \quad (7.11)$$

Setzt man  $\varphi_j(\cdot) := 0$  für  $j < 0$ , dann folgt für  $t \in \mathbb{Z}$  mit  $t \leq n$  und  $h \in \{1 - p, \dots, 0\}$  die gleichmäßige Beschränktheit der folgenden Ausdrücke

$$(i) \quad \sum_{j=0}^{\infty} |\varphi_j(\frac{t}{n}) \varphi_{j+h}(\frac{t+h}{n})|, \quad (7.12)$$

$$(ii) \quad \sum_{r=1}^{\infty} \sum_{j=0}^{\infty} |\varphi_j(\frac{t}{n}) \varphi_{j+h+r}(\frac{t+h}{n})|, \quad (7.13)$$

$$(iii) \quad \sum_{r=1}^{\infty} \sum_{j=0}^{\infty} |\varphi_j(\frac{t}{n}) \varphi_{j+h-r}(\frac{t+h}{n})|, \quad (7.14)$$

$$(iv) \quad \sum_{j=0}^{\infty} \sum_{s=j+1}^{\infty} |\varphi_s(\frac{t}{n}) \varphi_{s+h}(\frac{t+h}{n})|, \quad (7.15)$$

$$(v) \quad \sum_{r=1}^{\infty} \sum_{j=0}^{\infty} \sum_{s=j+1}^{\infty} |\varphi_s(\frac{t}{n}) \varphi_{s+h+r}(\frac{t+h}{n})|, \quad (7.16)$$

$$(vi) \quad \sum_{r=1}^{\infty} \sum_{j=0}^{\infty} \sum_{s=j+1}^{\infty} |\varphi_s(\frac{t}{n}) \varphi_{s+h-r}(\frac{t+h}{n})|. \quad (7.17)$$

*Beweis.*

zu (i): Mit Hilfssatz 7.1 (i) folgt wegen  $\varphi_j(\cdot) = 0$  für  $j < 0$  und  $h \leq n$

$$\begin{aligned} \sum_{j=0}^{\infty} |\varphi_j(\frac{t}{n})\varphi_{j+h}(\frac{t+h}{n})| &\leq \left( \sum_{j=0}^{\infty} |\varphi_j(\frac{t}{n})| \right) \left( \sum_{j=0}^{\infty} |\varphi_{j+h}(\frac{t+h}{n})| \right) \\ &\leq \left( \sup_{t \leq n} \sum_{j=0}^{\infty} |\varphi_j(\frac{t}{n})| \right)^2 \\ &< \infty. \end{aligned} \quad (7.18)$$

zu (ii): Wendet man Hilfssatz 7.1 (i) auf die Summe über  $j$  an, zieht die erste Summe hervor und wendet dann Hilfssatz 7.2 (i) auf die Doppelsumme an, so ergibt sich die Behauptung aus der Rechnung

$$\begin{aligned} \sum_{r=1}^{\infty} \sum_{j=0}^{\infty} |\varphi_j(\frac{t}{n})\varphi_{j+h+r}(\frac{t+h}{n})| &\leq \left( \sum_{j=0}^{\infty} |\varphi_j(\frac{t}{n})| \right) \left( \sum_{r=1}^{\infty} \sum_{j=0}^{\infty} |\varphi_{j+h+r}(\frac{t+h}{n})| \right) \\ &= \left( \sum_{j=0}^{\infty} |\varphi_j(\frac{t}{n})| \right) \left( \sum_{j=0}^{\infty} (j-h) |\varphi_j(\frac{t+h}{n})| \right) \\ &\leq \sup_{t \leq n} \sum_{j=0}^{\infty} |\varphi_j(\frac{t}{n})| \left( \sup_{t \leq n} \sum_{j=0}^{\infty} j |\varphi_j(\frac{t}{n})| - h \cdot \sup_{t \leq n} \sum_{j=0}^{\infty} |\varphi_j(\frac{t}{n})| \right) \\ &< \infty. \end{aligned} \quad (7.19)$$

zu (iii): Wegen  $h \leq 0$  und  $r \geq 1$  folgt  $0 + h - r \leq 0$ , weshalb

$$\sum_{r=1}^{\infty} \sum_{j=0}^{\infty} |\varphi_j(\frac{t}{n})\varphi_{j+h-r}(\frac{t+h}{n})| = \sum_{r=1}^{\infty} \sum_{j=1}^{\infty} |\varphi_j(\frac{t}{n})\varphi_{j+h-r}(\frac{t+h}{n})| \quad (7.20)$$

gilt. Multipliziert man  $|\varphi_j(\frac{t}{n})\varphi_{j+h-r}(\frac{t+h}{n})|$  von links mit  $j^{1+\delta}$  und von rechts mit  $j^{-1-\delta}$ , benutzt danach Hilfssatz 7.1 (i) und zieht wieder die erste Summe nach vorne, so ergibt sich

$$\sum_{r=1}^{\infty} \sum_{j=1}^{\infty} |\varphi_j(\frac{t}{n})\varphi_{j+h-r}(\frac{t+h}{n})| \leq \left( \sum_{j=1}^{\infty} j^{1+\delta} |\varphi_j(\frac{t}{n})| \right) \left( \sum_{r=1}^{\infty} \sum_{j=1}^{\infty} |\varphi_{j+h-r}(\frac{t+h}{n})| j^{-1-\delta} \right). \quad (7.21)$$

Die erste Summe über  $j$  auf der rechten Seite ist wegen (S1) gleichmäßig beschränkt und nach Hilfssatz 7.2 (ii) gilt für die Doppelsumme

$$\sum_{r=1}^{\infty} \sum_{j=1}^{\infty} |\varphi_{j+h-r}(\frac{t+h}{n})| j^{-1-\delta} = \sum_{j=0}^{\infty} \left( |\varphi_j(\frac{t+h}{n})| \sum_{r=1}^{\infty} \frac{1}{(j-h+r)^{1+\delta}} \right). \quad (7.22)$$

Wegen  $j \geq 0$  und  $-h \geq 0$  lässt sich die Summe über  $r$  gegen die konvergente und von  $j$  unabhängige Reihe  $\sum_{r=1}^{\infty} \frac{1}{r^{1+\delta}} < \infty$  abschätzen und  $\sum_{j=0}^{\infty} |\varphi_j(\frac{t+h}{n})|$  ist wegen (S1) gleichmäßig beschränkt.

zu (iv): Benutzt man hintereinander den Hilfssatz 7.1 (i), schätzt  $\sum_{s=j+1}^{\infty} |\varphi_s(\frac{t}{n})|$  durch  $\sum_{s=1}^{\infty} |\varphi_s(\frac{t}{n})|$  ab und wendet Hilfssatz 7.2 (iii) auf die Doppelsumme an, so folgt

$$\begin{aligned}
\sum_{j=0}^{\infty} \sum_{s=j+1}^{\infty} |\varphi_s(\frac{t}{n}) \varphi_{s+h}(\frac{t+h}{n})| &\leq \left( \sum_{s=1}^{\infty} |\varphi_s(\frac{t}{n})| \right) \left( \sum_{j=0}^{\infty} \sum_{s=j+1}^{\infty} |\varphi_{s+h}(\frac{t+h}{n})| \right) \\
&= \left( \sum_{s=1}^{\infty} |\varphi_s(\frac{t}{n})| \right) \left( \sum_{j=1}^{\infty} j |\varphi_{j+h}(\frac{t+h}{n})| \right) \tag{7.23} \\
&= \sum_{s=1}^{\infty} |\varphi_s(\frac{t}{n})| \left( \sum_{j=1}^{\infty} (j+h) |\varphi_{j+h}(\frac{t+h}{n})| - h \sum_{j=1}^{\infty} |\varphi_{j+h}(\frac{t+h}{n})| \right) \\
&\leq \sup_{t \leq n} \sum_{s=1}^{\infty} |\varphi_s(\frac{t}{n})| \left( \sup_{t \leq n} \sum_{j=0}^{\infty} j |\varphi_j(\frac{t}{n})| - h \cdot \sup_{t \leq n} \sum_{j=0}^{\infty} |\varphi_j(\frac{t}{n})| \right) \\
&< \infty.
\end{aligned}$$

zu (v): Multipliziert man  $|\varphi_s(\frac{t}{n}) \varphi_{s+h+r}(\frac{t+h}{n})|$  von links mit  $s$  und von rechts mit  $\frac{1}{s}$ , wendet dann Hilfssatz 7.1 (i) an und zieht die erste Summe abgeschätzt nach vorne, so folgt

$$\sum_{r=1}^{\infty} \sum_{j=0}^{\infty} \sum_{s=j+1}^{\infty} |\varphi_s(\frac{t}{n}) \varphi_{s+h+r}(\frac{t+h}{n})| \leq \sum_{s=1}^{\infty} s |\varphi_s(\frac{t}{n})| \left( \sum_{r=1}^{\infty} \sum_{j=0}^{\infty} \sum_{s=j+1}^{\infty} |\varphi_{s+h+r}(\frac{t+h}{n})| \frac{1}{s} \right). \tag{7.24}$$

Die erste Summe über  $s$  ist wieder nach (S1) gleichmäßig beschränkt, so dass noch die der Dreifachsumme zu zeigen bleibt. Benutzt man nacheinander Hilfssatz 7.2 (iii) und (i), so ergibt sich

$$\begin{aligned}
\sum_{r=1}^{\infty} \sum_{j=0}^{\infty} \sum_{s=j+1}^{\infty} |\varphi_{s+h+r}(\frac{t+h}{n})| \frac{1}{s} &= \sum_{r=1}^{\infty} \sum_{j=1}^{\infty} |\varphi_{j+h+r}(\frac{t+h}{n})| \\
&\leq \sum_{r=1}^{\infty} \sum_{j=0}^{\infty} |\varphi_{j+h+r}(\frac{t+h}{n})| \\
&\leq \sum_{j=0}^{\infty} (j-h) |\varphi_j(\frac{t+h}{n})| \tag{7.25} \\
&\leq \sup_{t \leq n} \sum_{j=0}^{\infty} j |\varphi_j(\frac{t}{n})| - h \cdot \sup_{t \leq n} \sum_{j=0}^{\infty} |\varphi_j(\frac{t}{n})| \\
&< \infty.
\end{aligned}$$

zu (vi): Multipliziert man  $|\varphi_s(\frac{t}{n})\varphi_{s+h-r}(\frac{t+h}{n})|$  von links mit  $s^{2+\delta}$  und von rechts mit  $s^{-2-\delta}$ , wendet wieder Hilfssatz 7.1 (i) an und zieht die erste Summe abgeschätzt nach vorne, so folgt

$$\sum_{r=1}^{\infty} \sum_{j=0}^{\infty} \sum_{s=j+1}^{\infty} |\varphi_s(\frac{t}{n})\varphi_{s+h-r}(\frac{t+h}{n})| \leq \sum_{s=1}^{\infty} s^{2+\delta} |\varphi_s(\frac{t}{n})| \sum_{r=1}^{\infty} \sum_{j=0}^{\infty} \sum_{s=j+1}^{\infty} |\varphi_{s+h-r}(\frac{t+h}{n})| \frac{1}{s^{2+\delta}}. \quad (7.26)$$

Die erste Summe über  $s$  ist wieder nach Voraussetzung (S1) gleichmäßig beschränkt. Mit Hilfssatz 7.2 (iii) und (ii) ergibt sich für die Dreifachsumme

$$\begin{aligned} \sum_{r=1}^{\infty} \sum_{j=0}^{\infty} \sum_{s=j+1}^{\infty} |\varphi_{s+h-r}(\frac{t+h}{n})| \frac{1}{s^{2+\delta}} &= \sum_{r=1}^{\infty} \sum_{j=1}^{\infty} |\varphi_{j+h-r}(\frac{t+h}{n})| \frac{1}{j^{1+\delta}} \\ &= \sum_{j=0}^{\infty} \left( |\varphi_j(\frac{t+h}{n})| \sum_{r=1}^{\infty} \frac{1}{(j-h+r)^{1+\delta}} \right), \quad (7.27) \end{aligned}$$

wobei sich die Summe über  $r$  wieder wegen  $j \geq 0$  und  $-h \geq 0$  gegen die konvergente und von  $j$  unabhängige Reihe  $\sum_{r=1}^{\infty} \frac{1}{r^{1+\delta}} < \infty$  abschätzen lässt und  $\sum_{j=0}^{\infty} |\varphi_j(\frac{t+h}{n})|$  ebenfalls wegen (S1) gleichmäßig beschränkt bleibt.  $\square$

**Lemma 7.4 (Zu Lemma 6.3).**

Die Folge auf  $(0, 1)$  differenzierbarer Funktionen  $\{\varphi_j(\cdot) | \varphi_j : (-\infty, 1] \rightarrow \mathbb{R}\}_{j=0}^{\infty}$  aus Lemma 3.7 erfülle die Bedingung

$$(S1) \quad \sup_{t \leq n} \sum_{j=0}^{\infty} j^{2+\delta} |\varphi_j(\frac{t}{n})| < \infty. \quad (7.28)$$

Setzt man  $\varphi_j(\cdot) := 0$  für  $j < 0$ , dann folgt für  $t \in \mathbb{Z}$  mit  $t \leq n$  und  $h \in \{1-p, \dots, 0\}$  die gleichmäßige Beschränktheit der folgenden Ausdrücke

$$(i) \quad \sum_{r=1}^{\infty} \left( \sum_{j=0}^{\infty} \varphi_j(\frac{t}{n}) \varphi_{j+h+r}(\frac{t+h}{n}) \right)^2, \quad (7.29)$$

$$(ii) \quad \sum_{r=1}^{\infty} \left( \sum_{j=0}^{\infty} \varphi_j(\frac{t}{n}) \varphi_{j+h-r}(\frac{t+h}{n}) \right)^2. \quad (7.30)$$

*Beweis.*

zu (i): Anwenden von Hilfssatz 7.1 (ii), Hervorziehen der ersten Summe über  $j$  und an-

schließendes Benutzen von Hilfssatz 7.2 (i) ergibt

$$\begin{aligned}
\sum_{r=1}^{\infty} \left( \sum_{j=0}^{\infty} \varphi_j\left(\frac{t}{n}\right) \varphi_{j+h+r}\left(\frac{t+h}{n}\right) \right)^2 &\leq \left( \sum_{j=0}^{\infty} \varphi_j^2\left(\frac{t}{n}\right) \right) \sum_{r=1}^{\infty} \sum_{j=0}^{\infty} \varphi_{j+h+r}^2\left(\frac{t+h}{n}\right) \\
&= \left( \sum_{j=0}^{\infty} \varphi_j^2\left(\frac{t}{n}\right) \right) \left( \sum_{j=0}^{\infty} (j-h) \varphi_j^2\left(\frac{t+h}{n}\right) \right) \\
&\leq \sup_{t \leq n} \sum_{j=0}^{\infty} \varphi_j^2\left(\frac{t}{n}\right) \left( \sup_{t \leq n} \sum_{j=0}^{\infty} j \varphi_j^2\left(\frac{t}{n}\right) - h \cdot \sup_{t \leq n} \sum_{j=0}^{\infty} \varphi_j^2\left(\frac{t}{n}\right) \right) \\
&< \infty.
\end{aligned} \tag{7.31}$$

zu (ii): Wegen  $h - r < 0$  gilt zunächst

$$\sum_{r=1}^{\infty} \left( \sum_{j=0}^{\infty} \varphi_j\left(\frac{t}{n}\right) \varphi_{j+h-r}\left(\frac{t+h}{n}\right) \right)^2 = \sum_{r=1}^{\infty} \left( \sum_{j=1}^{\infty} \varphi_j\left(\frac{t}{n}\right) \varphi_{j+h-r}\left(\frac{t+h}{n}\right) \right)^2. \tag{7.32}$$

Multipliziert man  $\varphi_j\left(\frac{t}{n}\right) \varphi_{j+h-r}\left(\frac{t+h}{n}\right)$  von links mit  $j^{\frac{1+\delta}{2}}$  und von rechts mit  $j^{-\frac{1+\delta}{2}}$  und wendet danach Hilfssatz 7.1 (ii) an, so folgt

$$\sum_{r=1}^{\infty} \left( \sum_{j=0}^{\infty} \varphi_j\left(\frac{t}{n}\right) \varphi_{j+h-r}\left(\frac{t+h}{n}\right) \right)^2 \leq \left( \sum_{j=0}^{\infty} j^{1+\delta} \varphi_j^2\left(\frac{t}{n}\right) \right) \left( \sum_{r=1}^{\infty} \sum_{j=1}^{\infty} \varphi_{j+h-r}^2\left(\frac{t+h}{n}\right) j^{-1-\delta} \right). \tag{7.33}$$

Die erste Summe über  $j$  auf der rechten Seite ist wegen (S1) gleichmäßig beschränkt und Hilfssatz 7.2 (ii) liefert für die Doppelsumme

$$\sum_{r=1}^{\infty} \sum_{j=1}^{\infty} \varphi_{j+h-r}^2\left(\frac{t+h}{n}\right) j^{-1-\delta} = \sum_{j=0}^{\infty} \left( \varphi_j^2\left(\frac{t+h}{n}\right) \sum_{r=1}^{\infty} \frac{1}{(j-h+r)^{1+\delta}} \right), \tag{7.34}$$

wobei wieder die Summe über  $r$  auf der rechten Seite nach oben durch die konvergente und von  $j$  unabhängige Reihe  $\sum_{r=1}^{\infty} \frac{1}{r^{1+\delta}} < \infty$  abgeschätzt werden kann und  $\sum_{j=0}^{\infty} \varphi_j^2\left(\frac{t+h}{n}\right)$  wegen (S1) gleichmäßig beschränkt ist.  $\square$

### Lemma 7.5 (Zu Lemma 6.5).

Die Folge auf  $(0, 1)$  differenzierbarer Funktionen  $\{\varphi_j(\cdot) | \varphi_j : (-\infty, 1] \rightarrow \mathbb{R}\}_{j=0}^{\infty}$  aus Lemma 3.7 erfülle die Bedingungen

$$(S1) \quad \sup_{t \leq n} \sum_{j=0}^{\infty} j^{2+\delta} |\varphi_j\left(\frac{t}{n}\right)| < \infty, \tag{7.35}$$

$$(S2) \quad \sup_{t \leq n} \sum_{j=0}^{\infty} j^{2+\delta} |\Delta_t[\varphi_j\left(\frac{t}{n}\right)]| = o(1). \tag{7.36}$$

Setzt man  $\varphi_j(\cdot) := 0$  für  $j < 0$ , dann folgt für  $t \in \mathbb{Z}$  mit  $t \leq n$  und  $h \in \{1-p, \dots, 0\}$  die gleichmäßige Konvergenz gegen null der folgenden Ausdrücke

$$(i) \quad \sum_{j=0}^{\infty} \sum_{s=j+1}^{\infty} |\Delta_t[\varphi_s(\frac{t}{n})\varphi_{s+h}(\frac{t+h}{n})]|, \quad (7.37)$$

$$(ii) \quad \sum_{r=1}^{\infty} \sum_{j=0}^{\infty} \sum_{s=j+1}^{\infty} |\Delta_t[\varphi_s(\frac{t}{n})\varphi_{s+h+r}(\frac{t+h}{n})]|, \quad (7.38)$$

$$(iii) \quad \sum_{r=1}^{\infty} \sum_{j=0}^{\infty} \sum_{s=j+1}^{\infty} |\Delta_t[\varphi_s(\frac{t}{n})\varphi_{s+h-r}(\frac{t+h}{n})]|. \quad (7.39)$$

*Beweis.*

Zunächst betrachtet man für  $r \in \mathbb{Z}$  den  $\Delta_t$ -Term etwas genauer. Es lässt sich leicht nachrechnen, dass für ihn die Gleichheit

$$\Delta_t[\varphi_s(\frac{t}{n})\varphi_{s+h+r}(\frac{t+h}{n})] = \Delta_t[\varphi_s(\frac{t}{n})]\varphi_{s+h+r}(\frac{t+h}{n}) + \varphi_s(\frac{t-1}{n})\Delta_t[\varphi_{s+h+r}(\frac{t+h}{n})] \quad (7.40)$$

gilt, welche im Folgenden benötigt wird.

zu (i): Setzt man (7.40) für  $r = 0$  ein und zerlegt mit Hilfe der Dreiecksungleichung den Ausdruck, so folgt

$$\begin{aligned} \sum_{j=0}^{\infty} \sum_{s=j+1}^{\infty} |\Delta_t[\varphi_s(\frac{t}{n})\varphi_{s+h}(\frac{t+h}{n})]| &\leq \sum_{j=0}^{\infty} \sum_{s=j+1}^{\infty} |\Delta_t[\varphi_s(\frac{t}{n})]\varphi_{s+h}(\frac{t+h}{n})| \\ &+ \sum_{j=0}^{\infty} \sum_{s=j+1}^{\infty} |\varphi_s(\frac{t-1}{n})\Delta_t[\varphi_{s+h}(\frac{t+h}{n})]| \quad (7.41) \\ &\leq \left( \sum_{s=1}^{\infty} |\Delta_t[\varphi_s(\frac{t}{n})]| \right) \left( \sum_{j=0}^{\infty} \sum_{s=j+1}^{\infty} |\varphi_{s+h}(\frac{t+h}{n})| \right) \\ &+ \left( \sum_{s=1}^{\infty} |\varphi_s(\frac{t-1}{n})| \right) \left( \sum_{j=0}^{\infty} \sum_{s=j+1}^{\infty} |\Delta_t[\varphi_{s+h}(\frac{t+h}{n})]| \right). \end{aligned}$$

Wendet man noch jeweils nacheinander Hilfssatz 7.1 (i) und Hilfssatz 7.2 (iii) an, so ergibt sich die Abschätzung

$$\begin{aligned} \sum_{j=0}^{\infty} \sum_{s=j+1}^{\infty} |\Delta_t[\varphi_s(\frac{t}{n})\varphi_{s+h}(\frac{t+h}{n})]| &\leq \left( \sum_{s=1}^{\infty} |\Delta_t[\varphi_s(\frac{t}{n})]| \right) \left( \sum_{j=1}^{\infty} j |\varphi_{j+h}(\frac{t+h}{n})| \right) \\ &+ \left( \sum_{s=1}^{\infty} |\varphi_s(\frac{t-1}{n})| \right) \left( \sum_{j=1}^{\infty} j |\Delta_t[\varphi_{j+h}(\frac{t+h}{n})]| \right). \quad (7.42) \end{aligned}$$

Dabei sind analog zu vorherigen Überlegungen die Ausdrücke in der zweiten und dritten Klammer auf der rechten Seite nach (S1) gleichmäßig beschränkt und die in der ersten und vierten streben nach (S2) gleichmäßig gegen null, woraus insgesamt die behauptete gleichmäßige Konvergenz gegen null von (i) folgt.

zu (ii): Analoges Vorgehen wie im Beweis von (ii) ergibt zunächst

$$\begin{aligned} \sum_{r=1}^{\infty} \sum_{j=0}^{\infty} \sum_{s=j+1}^{\infty} |\Delta_t[\varphi_s(\frac{t}{n})\varphi_{s+h+r}(\frac{t+h}{n})]| &\leq \sum_{r=1}^{\infty} \sum_{j=0}^{\infty} \sum_{s=j+1}^{\infty} |\Delta_t[\varphi_s(\frac{t}{n})]\varphi_{s+h+r}(\frac{t+h}{n})| \quad (7.43) \\ &+ \sum_{r=1}^{\infty} \sum_{j=0}^{\infty} \sum_{s=j+1}^{\infty} |\varphi_s(\frac{t-1}{n})\Delta_t[\varphi_{s+h+r}(\frac{t+h}{n})]|. \end{aligned}$$

Multipliziert man die Ausdrücke in Betragstrichen von links mit  $s$  und von rechts mit  $\frac{1}{s}$  und wendet anschließend Hilfssatz 7.1 (i) und hintereinander Hilfssatz 7.2 (iii) und (i) an, so folgt

$$\begin{aligned} \sum_{r=1}^{\infty} \sum_{j=0}^{\infty} \sum_{s=j+1}^{\infty} |\Delta_t[\varphi_s(\frac{t}{n})\varphi_{s+h+r}(\frac{t+h}{n})]| &\leq \sum_{s=1}^{\infty} s |\Delta_t[\varphi_s(\frac{t}{n})]| \left( \sum_{j=0}^{\infty} (j-h) |\varphi_j(\frac{t+h}{n})| \right) \quad (7.44) \\ &+ \sum_{s=1}^{\infty} s |\varphi_s(\frac{t-1}{n})| \left( \sum_{j=0}^{\infty} (j-h) |\Delta_t[\varphi_j(\frac{t+h}{n})]| \right). \end{aligned}$$

Wieder sind wegen (S1) und (S2) die Ausdrücke in der zweiten und der dritten Klammer gleichmäßig beschränkt und die in der ersten und vierten konvergieren gleichmäßig gegen null.

zu (iii): Wie im Beweis von (i) und (ii) gilt

$$\begin{aligned} \sum_{r=1}^{\infty} \sum_{j=0}^{\infty} \sum_{s=j+1}^{\infty} |\Delta_t[\varphi_s(\frac{t}{n})\varphi_{s+h-r}(\frac{t+h}{n})]| &\leq \sum_{r=1}^{\infty} \sum_{j=0}^{\infty} \sum_{s=j+1}^{\infty} |\Delta_t[\varphi_s(\frac{t}{n})]\varphi_{s+h-r}(\frac{t+h}{n})| \quad (7.45) \\ &+ \sum_{r=1}^{\infty} \sum_{j=0}^{\infty} \sum_{s=j+1}^{\infty} |\varphi_s(\frac{t-1}{n})\Delta_t[\varphi_{s+h-r}(\frac{t+h}{n})]|. \end{aligned}$$

Multiplizieren der Ausdrücke in Betragstrichen von links mit  $s^{2+\delta}$  und von rechts mit  $s^{-2-\delta}$  und anschließendes Anwenden von Hilfssatz 7.1 (i) bzw. nacheinander Hilfssatz 7.2

(iii) und (ii) ergibt dann

$$\begin{aligned}
& \sum_{r=1}^{\infty} \sum_{j=0}^{\infty} \sum_{s=j+1}^{\infty} |\Delta_t[\varphi_s(\frac{t}{n})\varphi_{s+h-r}(\frac{t+h}{n})]| \\
& \leq \left( \sum_{s=1}^{\infty} s^{2+\delta} |\Delta_t[\varphi_s(\frac{t}{n})]| \right) \left( \sum_{j=0}^{\infty} |\varphi_j(\frac{t+h}{n})| \sum_{r=1}^{\infty} \frac{1}{(j-h+r)^{1+\delta}} \right) \\
& + \left( \sum_{s=1}^{\infty} s^{2+\delta} |\varphi_s(\frac{t-1}{n})| \right) \left( \sum_{j=0}^{\infty} |\Delta_t[\varphi_j(\frac{t+h}{n})]| \sum_{r=1}^{\infty} \frac{1}{(j-h+r)^{1+\delta}} \right).
\end{aligned} \tag{7.46}$$

Die Summe  $\sum_{r=1}^{\infty} \frac{1}{(j-h+r)^{1+\delta}}$  kann zweimal gegen  $\sum_{r=1}^{\infty} \frac{1}{r^{1+\delta}} < \infty$  abgeschätzt werden, woraus mit (S1) die gleichmäßige Beschränktheit der zweiten und dritten Klammer folgt sowie mit (S2) die gleichmäßige Konvergenz gegen null der ersten und vierten Klammer.  $\square$

**Lemma 7.6 (Zu Satz 6.10).**

Die Folge auf  $(0, 1)$  differenzierbarer Funktionen  $\{\varphi_j(\cdot) | \varphi_j : (-\infty, 1] \rightarrow \mathbb{R}\}_{j=0}^{\infty}$  aus Lemma 3.7 erfülle die Bedingung

$$(S1) \quad \sup_{t \leq n} \sum_{j=0}^{\infty} j^{2+\delta} |\varphi_j(\frac{t}{n})| < \infty. \tag{7.47}$$

Weiter setzt man  $\varphi_j(\cdot) := 0$  für  $j < 0$  und  $\tilde{X}_{t-1} := (\tilde{X}_{t-1,n}, \dots, \tilde{X}_{t-p,n})^T$  für  $t \in \{1, \dots, n\}$ . Falls  $\kappa_4 := E[\epsilon_t^4] < \infty$  gilt, dann sind für  $r \in \{1, \dots, p\}$  die vierten Momente von  $\tilde{X}_{t-r,n}$  gleichmäßig beschränkt und damit auch  $E[\|\tilde{X}_{t-1}\|^4]$ .

*Beweis.*

Zunächst betrachtet man  $E[\tilde{X}_{t-r,n}^4]$  und dabei zuerst die vorkommenden Erwartungswerte des weißen Rauschens. Aus dessen stochastischer Unabhängigkeit folgt für  $i, j, k, l \in \mathbb{N}$ , dass

$$E[\epsilon_{t-r-i}\epsilon_{t-r-j}\epsilon_{t-r-k}\epsilon_{t-r-l}] = \begin{cases} \kappa_4, & i = j = k = l \\ \sigma_\epsilon^4, & \text{zwei Paare} \\ 0, & \text{sonst} \end{cases} \tag{7.48}$$

gilt. Damit ergibt sich für das vierte Moment von  $\tilde{X}_{t-r,n}$  die Gleichheit

$$E[\tilde{X}_{t-r,n}^4] = (\kappa_4 - 3\sigma_\epsilon^4) \sum_{j=0}^{\infty} \varphi_j^4(\frac{t-r}{n}) + \sigma_\epsilon^4 \left( \sum_{j=0}^{\infty} \varphi_j^2(\frac{t-r}{n}) \right)^2. \tag{7.49}$$

Supremumbildung über  $t \leq n$  auf beiden Seiten und Abschätzen auf der rechten Seite ergibt die Ungleichung

$$\begin{aligned} \sup_{t \leq n} E[\tilde{X}_{t-r,n}^4] &\leq (\kappa_4 - 3\sigma_\epsilon^4) \sup_{t \leq n} \sum_{j=0}^{\infty} \varphi_j^4\left(\frac{t-r}{n}\right) + \sigma_\epsilon^4 \left( \sup_{t \leq n} \sum_{j=0}^{\infty} \varphi_j^2\left(\frac{t-r}{n}\right) \right)^2 \\ &\leq (\kappa_4 - 3\sigma_\epsilon^4) \sup_{t \leq n} \sum_{j=0}^{\infty} \varphi_j^4\left(\frac{t}{n}\right) + \sigma_\epsilon^4 \left( \sup_{t \leq n} \sum_{j=0}^{\infty} \varphi_j^2\left(\frac{t}{n}\right) \right)^2. \end{aligned} \quad (7.50)$$

Das zweite Supremum auf der rechten Seite ist nach (S1) beschränkt und das Supremum über die Summe vierter Potenzen lässt sich mit Hilfssatz 7.1 (i) auf das Supremum über eine Summe mit quadratischen Termen gemäß

$$\sup_{t \leq n} \sum_{j=0}^{\infty} \varphi_j^4\left(\frac{t}{n}\right) \leq \left( \sup_{t \leq n} \sum_{j=0}^{\infty} \varphi_j^2\left(\frac{t}{n}\right) \right)^2 \quad (7.51)$$

zurückführen, dessen Beschränktheit ebenfalls aus (S1) folgt. Für alle  $r \in \{1, \dots, p\}$  gilt also

$$\sup_{t \leq n} E[\tilde{X}_{t-r,n}^4] < \infty. \quad (7.52)$$

Für die zweite Aussage multipliziert man  $E[\|\tilde{\underline{X}}_{t-1}\|^4]$  aus und ordnet nach den Indizes. Es gilt

$$\begin{aligned} E[\|\tilde{\underline{X}}_{t-1}\|^4] &= E \left[ \sum_{r=1}^p \sum_{s=1}^p \tilde{X}_{t-r,n}^2 \tilde{X}_{t-s,n}^2 \right] \\ &= \sum_{r=1}^p E \left[ \tilde{X}_{t-r,n}^4 \right] + \sum_{\substack{r,s=1 \\ r \neq s}}^p E \left[ \tilde{X}_{t-r,n}^2 \tilde{X}_{t-s,n}^2 \right] \end{aligned} \quad (7.53)$$

und damit auch

$$\sup_{t \leq n} E[\|\tilde{\underline{X}}_{t-1}\|^4] \leq \sum_{r=1}^p \sup_{t \leq n} E \left[ \tilde{X}_{t-r,n}^4 \right] + \sum_{\substack{r,s=1 \\ r \neq s}}^p \sup_{t \leq n} E \left[ \tilde{X}_{t-r,n}^2 \tilde{X}_{t-s,n}^2 \right]. \quad (7.54)$$

Die erste Summe auf der rechten Seite ist endlich und daher wegen (7.52) gleichmäßig beschränkt. Mit der Cauchy-Schwarz-Ungleichung (2.39) folgt ebenfalls mit (7.52) auch die der Doppelsumme aus

$$E \left[ \tilde{X}_{t-r,n}^2 \tilde{X}_{t-s,n}^2 \right] \leq \sqrt{E \left[ \tilde{X}_{t-r,n}^4 \right]} \sqrt{E \left[ \tilde{X}_{t-s,n}^4 \right]}. \quad (7.55)$$

□

# Literaturverzeichnis

- [1] Beveridge, S.; Nelson, C.R.: A new approach to decomposition of economic time series into permanent and transitory components with particular attention to measurement of the „Business Cycle“. *Journal of Monetary Economics*, 7: 151-174 (1981)
- [2] Brockwell, P.J.; Davis, R.A.: Time Series: Theory and Methods. *New York: Springer* (1991)
- [3] Dahlhaus, R.: On the Kullback-Leibler information divergence of locally stationary processes. *Stochastic Processes and their Applications*, 62: 139-168 (1996)
- [4] Dahlhaus, R.: Maximum likelihood estimation and modelselection for locally stationary processes. *Journal of Nonparametric Statistics*, 6: 171-191 (1996)
- [5] Dahlhaus, R.: Asymptotic statistical inference for nonstationary processes with evolutionary spectra. In *Athens Conference on Applied Probability and Time Series*, Vol. II (eds. P.M. Robinson and M. Rosenblatt), Lecture Notes in Statistics 115. *New York: Springer*, 145-159 (1996)
- [6] Dahlhaus, R.: Fitting time series models to nonstationary processes. *Annals of Statistics*, 25: 1-37 (1997)
- [7] Dahlhaus, R. ; Giraitis, L.: On the optimal segment length for parameter estimates for locally stationary time series. *J. of Time Series Analysis*, 19: 629-655 (1998)
- [8] Deuffhard, P.; Homann, A.: Numerische Mathematik 1, 3. Auflage. *Berlin: de Gruyter* (2002)
- [9] Efromovich, S.: Nonparametric Curve Estimation. *New York: Springer* (1999)
- [10] Fan, J.; Gijbels, I.: Local Polynomial Modelling and its Applications. (Monographs on Statistics and Applied Probability ; 66) *London: Chapman and Hall* (1996)
- [11] Hallin, M.: Mixed autoregressive-moving average multivariate processes with time-dependent coefficients. *J. of Multivariate Analysis*, 8: 567-572 (1978)

- 
- [12] Hallin, M.: Spectral factorization of nonstationary moving average processes. *The Annals of Statistics*, 12: 172-192 (1984)
- [13] Householder, A.S.: The Theory of Matrices in Numerical Analysis. *New York: Blaisdell* (1964)
- [14] Kim, W.: Nonparametric kernel estimation of evolutionary autoregressive processes. *SFB 373 Discussion paper 103*, Berlin (2001)
- [15] Kreiß, J.-P.; Neuhaus, G.: Einführung in die Zeitreihenanalyse *Berlin: Springer* (2006)
- [16] Künsch, H.R.: A note on causal solutions for locally stationary AR-processes. *ETH Zürich, Preprint* (1995)
- [17] Liptser, R.S.; Shiryaev, A.N.: A functional central limit theorem for semimartingales. *Theory of Probability and its Applications*, 25: 667-688 (1981)
- [18] Mélard, G.: An example of the evolutionary spectral theory. *J. of Time Series Analysis*, 6: 81-90 (1985)
- [19] Miller, K.S.: Linear Difference Equations. *New York: Benjamin* (1968)
- [20] Phillips, P.C.B. ; Solo, V.: Asymptotics for linear processes. *The Annals of Statistics*, 20: 971-1001 (1992)
- [21] Priestley, M.B.: Evolutionary spectra and non-stationary processes. *J. of the Royal Statistical Society. Series B*, 27: 204-237 (1965)
- [22] Schlittgen, R.; Streitberg, B.H.J.: Zeitreihenanalyse, 2. Auflage. *Wien: Oldenbourg* (1987)