

RHEINISCHE FRIEDRICH-WILHELMS-UNIVERSITÄT BONN

Institut für Theoretische Geodäsie

Lehrstuhl für Theoretische Geodäsie

Untersuchungen zur optimalen Modellierung und Detektierung von Ausreißern in Zeitreihen

Diplomarbeit

vorgelegt von

cand. geod. Christian Siemes

29. März 2004

Aufgabensteller:
Univ. Prof. Dr. W.D. Schuh

Betreuer:
Dipl. Ing. B. Kargoll

Erklärung gemäß § 22 (7) DPO:

Hiermit versichere ich, dass ich die vorliegende Diplomarbeit selbstständig verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel benutzt sowie Zitate kenntlich gemacht habe.

Bonn, 29. März 2004

(Christian Siemes)

Inhaltsverzeichnis

1	Einführung	1
2	Modellerweiterung für Ausreißer	5
2.1	Erweiterung des Modells für Ausreißer	5
2.2	Rekursive Parameterschätzung	7
2.3	Rekursive Parameterschätzung mit Erweiterung des Modells für Ausreißer . .	8
2.4	Motivation für zeitreihenbezogene Methoden	9
3	Grundbegriffe für Zeitreihen	11
3.1	Zeitreihen und Prozesse	11
3.2	B -Operator	12
3.3	Wurzeln von Polynomen und Einheitskreis	12
3.4	Momente einer Zeitreihe	12
3.5	Autokovarianz und Autokorrelation	13
3.6	Stationarität	13
3.7	Ergodizität	14
4	Spezielle Prozesse	15
4.1	Weißes Rauschen	15
4.2	$AR(p)$ Prozesse	16
4.3	$MA(q)$ Prozesse	17

4.4	Invertierbare $MA(q)$ und $AR(p)$ Prozesse	17
4.5	$ARMA(p,q)$ Prozesse	18
4.6	$ARCH(m)$ Prozesse	19
5	Maximum-Likelihood Schätzung	21
5.1	Likelihoodfunktion	21
5.1.1	Likelihoodfunktion eines GAUSSschen $AR(p)$ Prozesses	22
5.1.2	Likelihoodfunktion eines GAUSSschen $MA(q)$ Prozesses	25
5.1.3	Likelihoodfunktion eines GAUSSschen $ARMA(p,q)$ Prozesses	26
5.1.4	Likelihoodfunktion eines Regressionsmodells mit GAUSSschen $ARCH(m)$ Störtermen	28
5.2	Numerische Optimierungsverfahren	29
5.2.1	Gittersuche	29
5.2.2	Steilster Anstieg	30
5.2.3	NEWTON-RAPHSON	31
6	Modellidentifikation mit Informationskriterien	33
7	Ausreißeranalyse in Zeitreihen	35
7.1	Ausreißertypen	35
7.1.1	Additiver Ausreißer	37
7.1.2	Level Shift	38
7.1.3	Zeitliche Änderung	39
7.1.4	Innovativer Ausreißer	40
7.2	Algorithmus zur Ausreißersuche	41
7.3	Modifizierungen	42
7.3.1	Schätzung eines $ARMA(p,q)$ Prozesses nach der Methode der kleinsten Quadrate	43

7.3.2	Update für die Schätzung eines $ARMA(p, q)$ Prozesses	45
7.3.3	Faltung eines Ausreißermusters mit den Residuen	45
7.3.4	Update für die Faltung	46
7.3.5	Gemeinsames Schätzen der Ausreißer	47
7.4	Modifizierter Algorithmus zur Ausreißersuche	48
8	Tests zur Ausreißeranalyse	53
8.1	Vorbemerkung zum Level-Shift	53
8.2	Vergleich der Algorithmen zur Ausreißersuche	55
8.3	Einzelner Ausreißer	57
8.4	Zwei Ausreißer	61
8.5	Viele Ausreißer	65
8.6	Qualität der Modellidentifikation	68
8.7	Zusammenfassung	70
9	Tests zu $ARCH(m)$ Zeitreihen	73
9.1	Vergleich der numerischen Optimierungsverfahren	73
9.2	Modellidentifikation für $ARCH(m)$ Prozesse	77
9.3	Zusammenfassung	81
10	Schlussbemerkungen und Ausblick	83
A	Tabellen und Abbildungen	87
A.1	Abbildungen von Ausreißermustern	87
A.2	Abbildungen von Ausschnitten aus Zeitreihen	91
A.3	Tabellen zu Abschnitt 8.2	92
A.4	Tabellen zu Abschnitt 8.3	98
A.5	Tabellen zu Abschnitt 8.4	108

A.6 Tabellen zu Abschnitt 8.5	119
A.7 Tabellen zu Abschnitt 8.6	125
B Funktionen und Variablen in MATLAB	141
B.1 Algorithmus von Chen und Liu	141
B.2 Modifizierter Algorithmus von Chen und Liu	146
B.3 $ARCH(m)$ Prozesse	151

Kapitel 1

Einführung

Zur Bestimmung des Erdschwerefeldes werden Satellitenmissionen unternommen. Ein aktuelles Beispiel hierfür ist die GOCE Mission. Bei solchen Missionen liefert der Satellit in regelmäßigen Zeitabständen Beobachtungen, weswegen man diese als Zeitreihe ansehen kann. Nach Abzug eines deterministischen Modells, wie beispielsweise Kugelfunktionen, sollten die übrigbleibenden Residuen weißes Rauschen ergeben. Wegen unberücksichtigten Systematiken, deren Ursachen in den Eigenschaften des Sensors oder bei Fehlern im Modell liegen, wird man jedoch farbiges Rauschen vorfinden. Um diese Systematiken aufzudecken, bedient man sich der Sichtweise der Signaltheorie, bei der die Beobachtungen l_t aus einem Trend $f_t(x)$, einem Signal s_t und weißem Rauschen n_t bestehen:

$$l_t = f_t(x) + s_t + n_t. \quad (1.1)$$

Dabei ist der Trend ein deterministisches Modell bei dem Basisfunktionen vorgegeben werden und das Signal ein stochastisches Modell, welches die Korrelationen in den Residuen modelliert. Der Index t gibt dabei den Zeitpunkt an, bei dem die Beobachtung vorliegt. Abbildung 1.1 stellt diese Zerlegung der Beobachtungen dar.

Diese Zerlegung erhält man, indem der Trend geschätzt und abgespalten wird, so dass zunächst das farbige Rauschen übrig bleibt. Durch die Modellierung des farbigen Rauschens als ARMA-Prozess ist es möglich, die Residuen in Signal und weißes Rauschen zu zerlegen. Mit Hilfe des ARMA-Prozesses kann man die Beobachtungen und die Design-Matrix filtern, so dass ein Ersatzproblem entsteht, welches mit einer Schätzung nach der Methode der kleinsten Quadrate gelöst werden kann. Damit kann auf das Speichern der Kovarianzmatrix der Beobachtungen verzichtet werden. Außerdem ist die rechenaufwendige und numerisch problematische Inversion dieser Matrix nach dem Filtern nicht mehr nötig. Die Schätzung von ARMA-Prozessen, die diese Vorteile mit sich bringt, wurde in der vorangegangenen Diplomarbeit [8] untersucht.

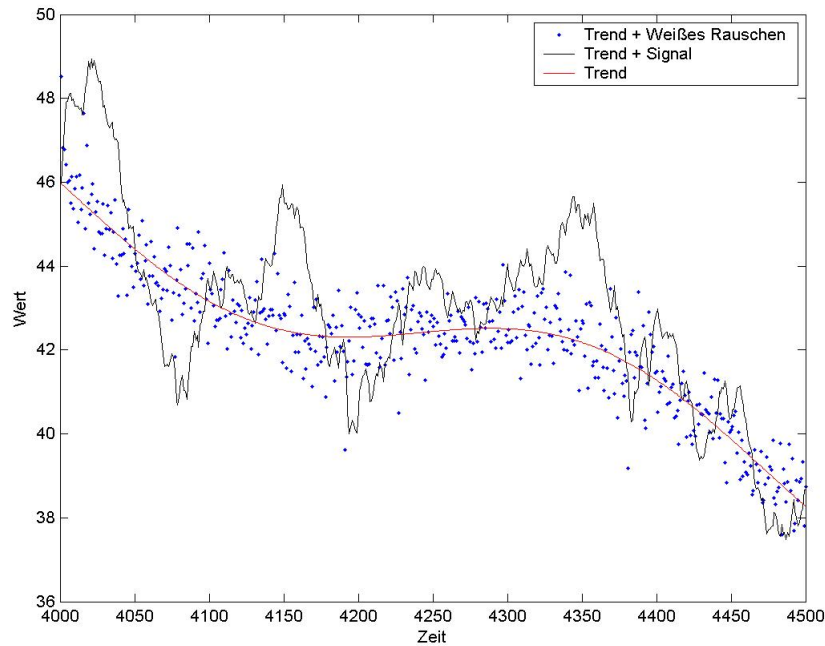


Abbildung 1.1: Zerlegung der Beobachtungen im Sinne der Signaltheorie

Es besteht die Möglichkeit, dass einige wenige Beobachtungen grob falsch sind. Solche Beobachtungen werden Ausreißer genannt. Sie können verschiedene Ursachen haben wie beispielsweise Defekte am Sensor oder Fehler bei der Datenübertragung. Die Gleichung (1.1) erhält dann einen zusätzlichen Term, der die Ausreißer der Größe ω_j mit Mustern $e_{t,j}$ modelliert, welche an die Zeitpunkte des Eintretens der Ausreißer gesetzt werden:

$$l_t = f_t(x) + s_t + n_t + \sum_j \omega_j e_{t,j}. \quad (1.2)$$

Sind Ausreißer in den Beobachtungen vorhanden, beeinträchtigen sie die Schätzung der Parameter des deterministischen Modells und des Signals. Um eine fehlerfreie Schätzung zu ermöglichen, müssen sie entfernt werden. Diese Diplomarbeit hat das Ziel, Mittel zu finden, mit denen die Ausreißer detektiert und deren Effekte geschätzt werden können. Dabei werden drei Varianten in Betracht gezogen:

1. Modellierung der Ausreißer mit Mustern direkt in den Beobachtungen,
2. Gemeinsames Schätzen von Signal und Ausreißer,
3. Modellierung der Residuen des deterministischen Modells als ARCH-Prozess.

Die erste Variante ist ein klassisches Konzept ohne direkten Zeitreihenbezug. Es wird nur in den Grundzügen vorgestellt, da es als Motivation für die zweite Variante dienen soll. Bei

der zweiten Variante werden verschiedene Ausreißertypen und ein Algorithmus vorgestellt, mit dem man die Ausreißer und das Signal gemeinsam Schätzen kann. Anschließend werden Modifikationen vorgeschlagen, die darauf zielen, die Rechenzeit wesentlich zu verringern. Die dritte Variante stellt eine alternative Modellierung der Residuen, welche sich aus Signal und weißem Rauschen zusammensetzen, dar.

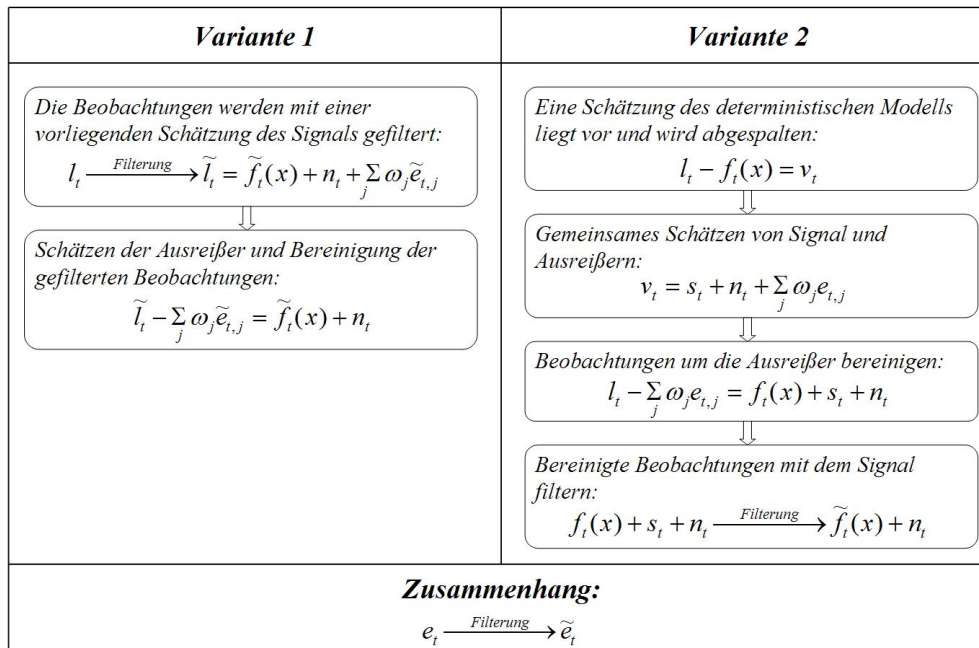


Abbildung 1.2: Vergleich der ersten beiden Varianten

Die ersten beiden Varianten haben das gleiche Ziel, die Beobachtungen um Ausreißer zu bereinigen und zu filtern, wie in Abbildung 1.2 zu erkennen ist. Setzt man bei der ersten Variante eine fehlerfreie Schätzung des Signals und bei der zweiten Variante eine fehlerfreie Schätzung des deterministischen Modells voraus, liefern sie identische Ergebnisse. In der Praxis werden weder eine fehlerfreie Schätzung des Signals noch eine fehlerfreie Schätzung des deterministischen Modells vorliegen. Eine Näherung für diese Schätzungen erhält man, indem die Ausreißer zuerst vernachlässigt werden. Die daraus folgenden, gefilterten und um Ausreißer bereinigten Beobachtungen lassen sich als eine zweite Näherung ansehen. Mit der zweiten Näherung lässt sich nach Abbildung 1.2 wiederum eine dritte berechnen, was letztlich zu einem iterativen Vorgehen bei der Berechnung der gefilterten und um Ausreißer bereinigten Beobachtungen führt.

Der Schwerpunkt dieser Diplomarbeit liegt auf der zweiten Variante, bei der ein Algorithmus¹ zum gemeinsamen Schätzen von Signal und Ausreißer für lange Zeitreihen optimiert wird.

¹Gemeint ist der Algorithmus von CHEN und LIU. Siehe [1].

Kapitel 2

Modellerweiterung für Ausreißer

Dieses Kapitel beschäftigt sich mit der Möglichkeit die Parameterschätzung mit einem GAUSS-MARKOFF Modell um eine Ausreißeranalyse zu ergänzen. Das geschieht durch eine Erweiterung des Modells, bei der die Ausreißer als zusätzliche Parameter angesetzt werden. Bei dieser klassischen Methode, welche in [5], Kapitel 4.4 dargestellt und in Abbildung 1.2 mit „Variante 1“ bezeichnet wird, ist kein Zeitreihenbezug gegeben. Anschließend wird kurz die rekursive Parameterschätzung vorgestellt, um sie schließlich mit der Ausreißeranalyse zu verbinden. Die Grundlagen hierfür sind in [5], Kapitel 3.2.8 wiederzufinden. Abschließend folgt ein Vergleich mit der zeitreihenbezogenen Methode der Ausreißeranalyse, welche in Abbildung 1.2 mit „Variante 2“ bezeichnet wird.

2.1 Erweiterung des Modells für Ausreißer

Zunächst soll davon ausgegangen werden, dass die Beobachtungen \mathbf{l} homogenisiert sind, so dass für die Kovarianzmatrix der Beobachtungen

$$\Sigma_{\mathbf{l}} = \mathbf{I}$$

gilt, wobei \mathbf{I} eine Einheitsmatrix ist. Die Homogenisierung der korelierten Beobachtungen $\tilde{\mathbf{l}}$ kann beispielsweise mit Hilfe der Cholesky Zerlegung $\Sigma_{\tilde{\mathbf{l}}}^{-1} = \tilde{\mathbf{R}}' \tilde{\mathbf{R}}$ vorgenommen werden. Dann wandelt sich das ursprüngliche GAUSS-MARKOFF Modell

$$\tilde{\mathbf{l}} + \tilde{\mathbf{v}} = \tilde{\mathbf{A}}\mathbf{x}, \quad \Sigma_{\tilde{\mathbf{l}}}$$

durch die Transformation $\mathbf{A} = \tilde{\mathbf{R}}\tilde{\mathbf{A}}$, $\mathbf{l} = \tilde{\mathbf{R}}\tilde{\mathbf{l}}$ und $\mathbf{v} = \tilde{\mathbf{R}}\tilde{\mathbf{v}}$ in das GAUSS-MARKOFF Modell mit homogenisierten Beobachtungen

$$\mathbf{l} + \mathbf{v} = \mathbf{A}\mathbf{x}, \quad \Sigma_{\mathbf{l}} = \mathbf{I}. \tag{2.1}$$

Dieses Modell wird nun für die Ausreißeranalyse um zusätzliche Parameter \mathbf{e} erweitert, die die Ausreißer in den Beobachtungen modellieren:

$$\mathbf{l} + \mathbf{v} = \mathbf{A}\mathbf{x} + \mathbf{E}\mathbf{e}, \quad \Sigma_{\mathbf{ll}} = \mathbf{I}. \quad (2.2)$$

Mit \mathbf{e} kann die Amplitude, beziehungsweise die Größe der Ausreißer variiert werden. Dabei stellt \mathbf{E} eine Art Basisfunktion oder Ausreißermuster dar. Im einfachsten Fall ist \mathbf{E} ein Einheitsvektor.

Nach dem Bilden und Invertieren der Normalgleichungen

$$\begin{pmatrix} \mathbf{A}'\mathbf{A} & \mathbf{A}'\mathbf{E} \\ \mathbf{E}'\mathbf{A} & \mathbf{E}'\mathbf{E} \end{pmatrix} \begin{pmatrix} \mathbf{x} \\ \mathbf{e} \end{pmatrix} = \begin{pmatrix} \mathbf{A}'\mathbf{l} \\ \mathbf{E}'\mathbf{l} \end{pmatrix}$$

erhält man die Formeln für die geschätzten Parameter der Ausreißer

$$\hat{\mathbf{e}} = -(\mathbf{E}'\Sigma_{\hat{\mathbf{v}}\hat{\mathbf{v}}}\mathbf{E})^{-1}\mathbf{E}'\hat{\mathbf{v}}$$

und die geschätzten Parameter des Modells

$$\hat{\hat{\mathbf{x}}} = \hat{\mathbf{x}} - (\mathbf{A}'\mathbf{A})^{-1}\mathbf{A}'\mathbf{E}\hat{\mathbf{e}}.$$

Dabei sind $\hat{\mathbf{x}}$ die geschätzten Parameter des Modells nach (2.1) und $\hat{\mathbf{v}}$ die entsprechenden Verbesserungen, sowie $\hat{\hat{\mathbf{x}}}$ die geschätzten Parameter des Modells nach (2.2). Man erkennt, dass die geschätzten Parameter der Ausreißer $\hat{\mathbf{e}}$ nur von den geschätzten Verbesserungen $\hat{\mathbf{v}}$ des Modells nach (2.1) abhängen. Die geschätzten Verbesserungen $\hat{\hat{\mathbf{v}}}$ des Modells nach (2.2) berechnet man mit

$$\hat{\hat{\mathbf{v}}} = \hat{\mathbf{v}} + \Sigma_{\hat{\mathbf{v}}\hat{\hat{\mathbf{v}}}}\mathbf{E}\hat{\mathbf{e}}.$$

Für die Kovarianzmatrizen der geschätzten Größen aus dem Modell nach (2.2) findet man

$$\begin{aligned} \Sigma_{\hat{\mathbf{e}}\hat{\mathbf{e}}} &= (\mathbf{E}'\Sigma_{\hat{\mathbf{v}}\hat{\mathbf{v}}}\mathbf{E})^{-1} \\ \Sigma_{\hat{\hat{\mathbf{x}}}\hat{\hat{\mathbf{x}}}} &= \Sigma_{\hat{\mathbf{x}}\hat{\mathbf{x}}} + \Sigma_{\hat{\mathbf{x}}\hat{\mathbf{x}}}\mathbf{A}'\mathbf{E}\Sigma_{\hat{\mathbf{e}}\hat{\mathbf{e}}}\mathbf{E}'\mathbf{A}\Sigma_{\hat{\mathbf{x}}\hat{\mathbf{x}}} \\ \Sigma_{\hat{\hat{\mathbf{v}}}\hat{\hat{\mathbf{v}}}} &= \Sigma_{\hat{\mathbf{v}}\hat{\mathbf{v}}} - \Sigma_{\hat{\mathbf{v}}\hat{\mathbf{v}}}\mathbf{E}\Sigma_{\hat{\mathbf{e}}\hat{\mathbf{e}}}\mathbf{E}'\Sigma_{\hat{\mathbf{v}}\hat{\mathbf{v}}}. \end{aligned}$$

Nun kann man alle geschätzten Größen mit Berücksichtigung von Ausreißern aus den geschätzten Größen ohne Berücksichtigung von Ausreißern berechnen.

Schließlich ist zu testen, ob die angesetzten Ausreißer signifikant sind. Die Nullhypothese H_0 ist folglich, dass der Erwartungswert der geschätzten Ausreißer $\hat{\mathbf{e}}$ gleich Null ist:

$$H_0 : E(\hat{\mathbf{e}}) = \mathbf{0}.$$

Dies kann mit einem FISHER-Test geschehen, wozu man die Quadratsumme $\Omega_{\hat{\mathbf{e}}}$ der geschätzten Ausreißer und die Quadratsumme der Verbesserungen $\Omega_{\hat{\hat{\mathbf{v}}}}$ des Modells nach (2.2) benötigt.

Dazu führt man die Nullhypothese wie eine Restriktion in das Modell ein. Die Quadratsummen lassen sich durch

$$\Omega_{\hat{\mathbf{v}}} = \Omega_{\hat{\mathbf{v}}} + \Omega_{\hat{\mathbf{e}}}$$

berechnen, wobei $\Omega_{\hat{\mathbf{v}}}$ die Quadratsumme der Verbesserungen des Modells nach (2.1) ist. Da homogenisierte Beobachtungen vorausgesetzt werden, ist $\Omega_{\hat{\mathbf{v}}} = \hat{\mathbf{v}}' \hat{\mathbf{v}}$ und $\Omega_{\hat{\mathbf{e}}} = \hat{\mathbf{v}}' \mathbf{E} \boldsymbol{\Sigma}_{\hat{\mathbf{e}}\hat{\mathbf{e}}} \mathbf{E}' \hat{\mathbf{v}}$. Ist n die Anzahl der Beobachtungen, m die Anzahl der Parameter des Modells nach (2.1) und q die Anzahl der zusätzlichen Parameter für die Ausreißer, so ergibt sich die Testgröße

$$T = \frac{\Omega_{\hat{\mathbf{e}}}/q}{\Omega_{\hat{\mathbf{v}}}/(n-m-q)} \sim F_{q,n-m-q}.$$

Setzt man nur einen einzelnen Ausreißer an, so dass \mathbf{E} ein Einheitsvektor ist, kann man auch einen BAARDA Test durchführen. Für die Testgröße erhält man

$$T = \frac{\hat{e}}{\sigma_{\hat{e}}} \sim N(0, 1) \quad \text{mit} \quad \hat{e} = \frac{-\hat{v}_k}{\sigma_{\hat{v}_k}^2} \quad \text{und} \quad \sigma_{\hat{e}}^2 = \frac{1}{\sigma_{\hat{v}_k}^2}.$$

2.2 Rekursive Parameterschätzung

Bei der rekursiven Parameterschätzung macht man den Ansatz, die Beobachtungen in Gruppen einzuteilen und dann gruppenweise nacheinander in den Ausgleich einfließen zu lassen. Die Beobachtungsgruppen sollen dabei untereinander unkorreliert sein. Auf Grund der Analogie zum Kalmanfilter spricht man hier auch von einem statischen Kalmanfilter, da die Parameter für jede Stufe des Ausgleichs als gleich angenommen werden. Für das GAUSS-MARKOFF Modell

$$\mathbf{l} + \mathbf{v} = \mathbf{A}\mathbf{x}, \quad \boldsymbol{\Sigma}_{\mathbf{ll}}$$

erhält man nach Unterteilung der Beobachtungen in M Gruppen

$$\begin{pmatrix} \mathbf{l}_1 \\ \mathbf{l}_2 \\ \vdots \\ \mathbf{l}_M \end{pmatrix} + \begin{pmatrix} \mathbf{v}_1 \\ \mathbf{v}_2 \\ \vdots \\ \mathbf{v}_M \end{pmatrix} = \begin{pmatrix} \mathbf{A}_1 \\ \mathbf{A}_2 \\ \vdots \\ \mathbf{A}_M \end{pmatrix} \mathbf{x}, \quad \boldsymbol{\Sigma}_{\mathbf{ll}} = \begin{pmatrix} \boldsymbol{\Sigma}_{\mathbf{l}_1\mathbf{l}_1} & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & \boldsymbol{\Sigma}_{\mathbf{l}_2\mathbf{l}_2} & \dots & \mathbf{0} \\ \vdots & \vdots & \dots & \vdots \\ \mathbf{0} & \mathbf{0} & \dots & \boldsymbol{\Sigma}_{\mathbf{l}_M\mathbf{l}_M} \end{pmatrix}.$$

Die erste Beobachtungsgruppe \mathbf{l}_1 wird wie im normalen GAUSS-MARKOFF Modell ausgeglichen:

$$\hat{\mathbf{x}}_1 = (\mathbf{A}'_1 \boldsymbol{\Sigma}_{\mathbf{l}_1\mathbf{l}_1}^{-1} \mathbf{A}_1)^{-1} \mathbf{A}'_1 \boldsymbol{\Sigma}_{\mathbf{l}_1\mathbf{l}_1}^{-1} \mathbf{l}_1, \quad \boldsymbol{\Sigma}_{\hat{\mathbf{x}}\hat{\mathbf{x}}} = (\mathbf{A}'_1 \boldsymbol{\Sigma}_{\mathbf{l}_1\mathbf{l}_1}^{-1} \mathbf{A}_1)^{-1}.$$

Danach startet die Rekursion, mit der die Beobachtungen $\mathbf{l}_2, \mathbf{l}_3, \dots, \mathbf{l}_M$ gruppenweise in den Ausgleich einfließen. Mit Hilfe der sogenannten Gain-Matrix

$$\mathbf{G}_m = \boldsymbol{\Sigma}_{\hat{\mathbf{x}}_{m-1}\hat{\mathbf{x}}_{m-1}} \mathbf{A}'_m (\boldsymbol{\Sigma}_{\mathbf{l}_m\mathbf{l}_m} + \mathbf{A}_m \boldsymbol{\Sigma}_{\hat{\mathbf{x}}_{m-1}\hat{\mathbf{x}}_{m-1}} \mathbf{A}'_m)^{-1}$$

kann man die Parameter der nächsten Ausgleichstufe berechnen. Die Rekursionsformeln lauten

$$\hat{\mathbf{x}}_m = \hat{\mathbf{x}}_{m-1} + \mathbf{G}_m(\mathbf{l}_m - \mathbf{A}_m \hat{\mathbf{x}}_{m-1}), \quad \Sigma_{\hat{\mathbf{x}}_m \hat{\mathbf{x}}_m} = \Sigma_{\hat{\mathbf{x}}_{m-1} \hat{\mathbf{x}}_{m-1}} - \mathbf{G}_m \mathbf{A}_m \Sigma_{\hat{\mathbf{x}}_{m-1} \hat{\mathbf{x}}_{m-1}},$$

wobei $m = 2, 3, \dots, M$ gilt.

2.3 Rekursive Parameterschätzung mit Erweiterung des Modells für Ausreißer

Die Idee ist nun, die Modellerweiterung für Ausreißer auf die rekursive Parameterschätzung anzuwenden. Dabei wird vorausgesetzt, dass homogenisierte Beobachtungen vorliegen, so dass

$$\Sigma_{\mathbf{1}\mathbf{1}} = \mathbf{I}$$

gilt. Außerdem wird angenommen, dass die Ausreißer sich immer innerhalb einer Beobachtungsgruppe befinden. Diese Einschränkung kann durch die Wahl der Größe einer Beobachtungsgruppe aufgeweicht werden. Findet man beispielsweise einen Ausreißer, der sich über zwei Beobachtungsgruppen \mathbf{l}_k und \mathbf{l}_{k+1} erstreckt, so könnte man diese Beobachtungsgruppen zu einer zusammenfassen.

Zuerst werden die Parameter in der ersten Stufe geschätzt:

$$\hat{\mathbf{x}}_1 = (\mathbf{A}'_1 \mathbf{A}_1)^{-1} \mathbf{A}'_1 \mathbf{l}_1, \quad \Sigma_{\hat{\mathbf{x}}_1 \hat{\mathbf{x}}_1} = (\mathbf{A}'_1 \mathbf{A}_1)^{-1}.$$

Die für die Ausreißeranalyse relevanten Verbesserungen und deren Kovarianzmatrix berechnet man mit

$$\hat{\mathbf{v}}_1 = \mathbf{A}_1 \hat{\mathbf{x}}_1 - \mathbf{l}_1, \quad \Sigma_{\hat{\mathbf{v}}_1 \hat{\mathbf{v}}_1} = \mathbf{I} - \mathbf{A}_1 \Sigma_{\hat{\mathbf{x}}_1 \hat{\mathbf{x}}_1} \mathbf{A}'_1.$$

Nun kann man die Ausreißer in der ersten Beobachtungsgruppe mit

$$\hat{\mathbf{e}}_1 = -(\mathbf{E}'_1 \Sigma_{\hat{\mathbf{v}}_1 \hat{\mathbf{v}}_1} \mathbf{E}_1)^{-1} \mathbf{E}'_1 \hat{\mathbf{v}}_1, \quad \Sigma_{\hat{\mathbf{e}}_1 \hat{\mathbf{e}}_1} = (\mathbf{E}'_1 \Sigma_{\hat{\mathbf{v}}_1 \hat{\mathbf{v}}_1} \mathbf{E}_1)^{-1}$$

schätzen, wodurch man in der Lage ist, die geschätzten Parameter der ersten Stufe ohne Berücksichtigung von Ausreißern zu korrigieren:

$$\hat{\hat{\mathbf{x}}}_1 = \hat{\mathbf{x}}_1 - \Sigma_{\hat{\mathbf{x}}_1 \hat{\mathbf{x}}_1} \mathbf{A}'_1 \mathbf{E}_1 \hat{\mathbf{e}}_1, \quad \Sigma_{\hat{\hat{\mathbf{x}}}_1 \hat{\hat{\mathbf{x}}}_1} = \Sigma_{\hat{\mathbf{x}}_1 \hat{\mathbf{x}}_1} + \Sigma_{\hat{\mathbf{x}}_1 \hat{\mathbf{x}}_1} \mathbf{A}'_1 \mathbf{E}_1 \Sigma_{\hat{\mathbf{e}}_1 \hat{\mathbf{e}}_1} \mathbf{E}'_1 \mathbf{A}_1 \Sigma_{\hat{\mathbf{x}}_1 \hat{\mathbf{x}}_1}.$$

Mit diesen geschätzten Parametern $\hat{\hat{\mathbf{x}}}_1$ der ersten Stufe, die nicht durch Ausreißer verfälscht sind, startet man die Rekursion. Als Erstes führt man die nächste Beobachtungsgruppe in den Ausgleich ein:

$$\hat{\mathbf{x}}_m = \hat{\hat{\mathbf{x}}}_{m-1} + \mathbf{G}_m(\mathbf{l}_m - \mathbf{A}_m \hat{\hat{\mathbf{x}}}_{m-1}), \quad \Sigma_{\hat{\mathbf{x}}_m \hat{\mathbf{x}}_m} = \Sigma_{\hat{\hat{\mathbf{x}}}_{m-1} \hat{\hat{\mathbf{x}}}_{m-1}} + \mathbf{G}_m \mathbf{A}_m \Sigma_{\hat{\hat{\mathbf{x}}}_{m-1} \hat{\hat{\mathbf{x}}}_{m-1}}.$$

Die Gain-Matrix berechnet man dabei mit

$$\mathbf{G}_m = \Sigma_{\hat{\mathbf{x}}_{m-1}\hat{\mathbf{x}}_{m-1}} \mathbf{A}'_m (\mathbf{I} + \mathbf{A}_m \Sigma_{\hat{\mathbf{x}}_{m-1}\hat{\mathbf{x}}_{m-1}} \mathbf{A}'_m)^{-1}.$$

Als Nächstes benötigt man die Verbesserungen mit deren Kovarianzmatrix:

$$\hat{\mathbf{v}}_m = \mathbf{A}_m \hat{\mathbf{x}}_m - \mathbf{l}_m, \quad \Sigma_{\hat{\mathbf{v}}_m \hat{\mathbf{v}}_m} = \mathbf{I} - \mathbf{A}_m \Sigma_{\hat{\mathbf{x}}_m \hat{\mathbf{x}}_m} \mathbf{A}'_m.$$

Nun kann man die Ausreißer, welche sich in dieser Beobachtungsgruppe befinden, schätzen:

$$\hat{\mathbf{e}}_m = -(\mathbf{E}'_m \Sigma_{\hat{\mathbf{v}}_m \hat{\mathbf{v}}_m} \mathbf{E}_m)^{-1} \mathbf{E}'_m \hat{\mathbf{v}}_m, \quad \Sigma_{\hat{\mathbf{e}}_m \hat{\mathbf{e}}_m} = (\mathbf{E}'_m \Sigma_{\hat{\mathbf{v}}_m \hat{\mathbf{v}}_m} \mathbf{E}_m)^{-1}.$$

Schließlich korrigiert man die geschätzten Parameter $\hat{\mathbf{x}}_m$ mit

$$\hat{\hat{\mathbf{x}}}_m = \hat{\mathbf{x}}_m - \Sigma_{\hat{\mathbf{x}}_m \hat{\mathbf{x}}_m} \mathbf{A}'_m \mathbf{E}_m \hat{\mathbf{e}}_m, \quad \Sigma_{\hat{\hat{\mathbf{x}}}_m \hat{\hat{\mathbf{x}}}_m} = \Sigma_{\hat{\mathbf{x}}_m \hat{\mathbf{x}}_m} + \Sigma_{\hat{\mathbf{x}}_m \hat{\mathbf{x}}_m} \mathbf{A}'_m \mathbf{E}_m \Sigma_{\hat{\mathbf{e}}_m \hat{\mathbf{e}}_m} \mathbf{E}'_m \mathbf{A}_m \Sigma_{\hat{\mathbf{x}}_m \hat{\mathbf{x}}_m},$$

so dass man mit geschätzten Parametern $\hat{\hat{\mathbf{x}}}_m$, die nicht durch Ausreißer verfälscht sind, in die nächste Stufe der Rekursion gehen kann. Die nicht durch Ausreißer verfälschten Verbesserungen werden für die Rekursion nicht benötigt. Außerdem fällt auf, dass die geschätzten Ausreißer $\hat{\mathbf{e}}_k$, die sich in einer Beobachtungsgruppe befinden, keinen Einfluss auf die nachfolgende Rechnung haben, was die Formeln vereinfacht.

2.4 Motivation für zeitreihenbezogene Methoden

Die rekursive Parameterschätzung mit Modellerweiterung für Ausreißer entspricht nach Gleichung (1.2) einer gemeinsamen Schätzung des deterministischen Modells mit den Ausreißern. Diese Variante steht nach Abbildung 1.2 in Konkurrenz zu der Variante, die Ausreißer gemeinsam mit dem Signal, welches als eine ARMA-Zeitreihe modelliert wird, zu schätzen.¹

In der vorangegangenen Diplomarbeit [8] wurde gezeigt, dass die Prozessparameter solcher ARMA-Zeitreihen nach der Methode der kleinsten Quadrate geschätzt werden können. In Kapitel 7.3.1 wird dargestellt, wie man die Normalgleichungen effizient aufstellt. Wegen der speziellen Struktur der Design-Matrix lassen sich viele Elemente der Normalgleichungsmatrix mit jeweils einer Multiplikation und Subtraktion aus den schon aufgestellten Elementen berechnen. Die Speicheranforderungen sind dabei gering, da im Wesentlichen nur doppelt so viel Speicherplatz verwendet wird, wie man zum Speichern der Zeitreihe benötigt.² Folglich ist die Schätzung des deterministischen Modells rechenaufwendiger und speicherintensiver als die Schätzung der Prozessparameter einer ARMA-Zeitreihe, wenn die Anzahl der Parameter für beide Schätzprobleme gleich ist und man für die Schätzung des deterministischen Modells

¹Voraussetzung dafür ist, dass die Beobachtungen eine Zeitreihe darstellen.

²Die Zeitreihe y_t sowie die mit Hilfe eines langen $AR(p)$ -Prozesses approximierten Residuen ε_t belegen den größten Teil des benötigten Speichers.

die Design-Matrix aufstellen muss. Im Kontext von Satellitenmissionen zur Schwerefeldbestimmung ist die Anzahl der Parameter des deterministischen Modells erheblich größer, als die Anzahl der Prozessparameter der ARMA-Zeitreihe,³ was den eben angestellten Vergleich bezüglich Rechenaufwand und Speicherintensität noch deutlicher ausfallen lässt.

Für beide Varianten in Abbildung 1.2 ist es sinnvoll, den signifikantesten Ausreißer zu finden, seine Größe zu schätzen und ihn aus den Residuen zu entfernen, bevor nach dem nächsten Ausreißer gesucht wird. Dies ist notwendig, um Maskierungseffekte verschiedener Ausreißertypen zu vermeiden. Bei der zweiten Variante wird die Größe der Ausreißer nach der Methode der kleinsten Quadrate geschätzt. Dies ist zulässig, da die Residuen bei der Schätzung der Prozessparameter der ARMA-Zeitreihe unkorreliert sind und einer Normalverteilung $N(0, \sigma^2)$ entsprechen. Daher ist die Suche nach dem signifikantesten Ausreißer sowie die Schätzung seiner Größe bei der zweiten Variante weniger rechenintensiv als bei der ersten.

Sind alle Ausreißer entfernt, wird bei beiden Varianten in Abbildung 1.2 die Schätzung der Parameter des deterministischen Modells beziehungsweise der Prozessparameter der ARMA-Zeitreihe korrigiert. Auch hier ergeben sich, speziell bei Satellitenmissionen wie die GOCE Mission, Vorteile beim Rechenaufwand für die zweite Variante.

Fazit: Im Bezug auf Rechenaufwand und Speicherintensität ist die zweite Variante in Abbildung 1.2 gegenüber der ersten wesentlich im Vorteil. Sie ist daher der ersten vorzuziehen.

³Bei der Satellitenmission GOCE wird das Erdschwerefeld durch Kugelfunktionen mit mehr als 100 000 Parametern modelliert. Als Signal erwartet man eine $ARMA(2,2)$ Zeitreihe.

Kapitel 3

Grundbegriffe für Zeitreihen

In diesem Kapitel sollen die grundlegenden Begriffe, die im Zusammenhang mit Zeitreihen auftauchen, erklärt werden. Für die in Kapitel 4 vorgestellten Prozesse werden sie benötigt. Sie sind beispielsweise in [2], Kapitel 3 und [10], Kapitel 5 wiederzufinden. Die hier verwendete Notation ist aus [2] weitestgehend übernommen.

3.1 Zeitreihen und Prozesse

Eine Zeitreihe Y_t ist eine Folge von zeitlich geordneten Beobachtungswerten

$$\{y_1, y_2, \dots, y_T\},$$

die in gleichen Zeitabständen Δt vorliegen. Jeder Wert y_t der Zeitreihe ist dabei durch einen stochastischen Prozess entstanden, welcher für die hier angestellten Untersuchungen immer als diskret und endlich anzunehmen ist. Er wird im Folgenden nur noch Prozess genannt. Formal hängt jeder Wert $y_t(\omega)$ von einem Elementarereignis ω ab. Man muss dabei folgende Fälle unterscheiden.¹

- Sind der Zeitpunkt t und das Elementarereignis ω fest, so handelt es sich bei $y_t(\omega)$ um eine feste Zahl.
- Hält man den Zeitpunkt t fest und läßt das Elementarereignis ω variabel, so ist $y_t(\omega)$ eine Zufallsvariable.
- Ist ω fest und t variabel, dann ist $y_t(\omega)$ eine Funktion der Zeit. Man spricht von einer Realisierung des Prozesses.

¹Siehe [10], Kapitel 5.2.

- Sind t und ω variabel, so fasst man $y_t(\omega)$ als eine Menge von Realisationen des Prozesses auf.

Die wichtigste Erkenntnis hierbei ist, dass man eine Zeitreihe Y_t immer als eine Realisierung eines Prozesses ansehen muss. Es können verschiedene Zeitreihen durch denselben Prozess erzeugt werden.

3.2 B -Operator

Eine der häufigsten Operationen, die bei der Analyse einer Zeitreihe verwendet wird, ist der Zugriff auf einen Wert der Zeitreihe, der $k \Delta t$ vor dem Wert t liegt. Um eine einfache Notation zu erhalten, führt man dazu den sogenannten Backward-Shift-, Backshift- oder Lag-Operator B ein. Für ihn gilt

$$B^k y_t = y_{t-k}.$$

Im Folgenden wird dieser Operator Backshift-Operator genannt.

3.3 Wurzeln von Polynomen und Einheitskreis

Sei $\varphi(z) = 1 + \varphi_1 z + \varphi_2 z^2 + \dots + \varphi_r z^r$ ein Polynom vom Grad r . Dieses Polynom besitzt r Nullstellen, die gegebenenfalls komplex sein können. Die Nullstellen des Polynoms $\varphi(z)$ werden auch als Wurzeln des Polynoms $\varphi(z)$ bezeichnet.

In die GAUSSsche Zahlenebene kann man komplexe Zahlen als Vektoren einzeichnen. Dabei wird der Realteil der komplexen Zahl entlang der Abszisse abgetragen und der Imaginärteil entlang der Ordinate. Als Einheitskreis wird der Kreis in der GAUSSschen Zahlenebene bezeichnet, dessen Mittelpunkt der Ursprung ist und dessen Radius gleich Eins ist.

3.4 Momente einer Zeitreihe

Ein Wert y_t einer Zeitreihe Y_t habe die unbedingte Dichtefunktion $f_{Y_t}(y_t)$. Dann ist das erste Moment der Zeitreihe

$$E(Y_t) = \int_{-\infty}^{\infty} y_t f_{Y_t}(y_t) dy_t$$

der Erwartungswert. Der Erwartungswert $E(y_t)$ wird auch als Mittelwert μ_t bezeichnet.

Das zweite zentrale Moment der Zeitreihe

$$E(Y_t - \mu_t)^2 = \int_{-\infty}^{\infty} (y_t - \mu_t)^2 f_{Y_t}(y_t) dy_t$$

ist die Varianz σ_t^2 .

3.5 Autokovarianz und Autokorrelation

Durch denselben Prozess seien verschiedene Realisationen $Y_t^{(1)}, Y_t^{(2)}, \dots, Y_t^{(n)}$ erzeugt worden. Da diese Realisationen auf denselben Prozess zurückzuführen sind, besitzen sie eine gemeinsame Dichtefunktion $f_{Y_t^{(1)}, Y_t^{(2)}, \dots, Y_t^{(n)}}(y_t^{(1)}, y_t^{(2)}, \dots, y_t^{(n)})$. Diese gemeinsame Dichtefunktion liegt nicht nur den verschiedenen Realisationen zu Grunde, sondern auch den Zeitreihen Y_{t-1}, \dots, Y_{t-j} , die zeitlich gegenüber Y_t verschoben sind und derselben Realisierung entstammen. In diesem Fall wird die gemeinsame Dichtefunktion mit $f_{Y_t, Y_{t-1}, \dots, Y_{t-j}}(y_t, y_{t-1}, \dots, y_{t-j})$ notiert.

Die Kovarianz γ der Zeitreihe Y_t mit der Zeitreihe Y_{t-j} erhält man aus

$$\gamma_{jt} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} (y_t - \mu_t)(y_{t-j} - \mu_{t-j}) f_{Y_t, Y_{t-1}, \dots, Y_{t-j}}(y_t, y_{t-1}, \dots, y_{t-j}) dy_t dy_{t-1} \dots dy_{t-j},$$

was dem Erwartungswert $E(Y_t - \mu_t)(Y_{t-j} - \mu_{t-j})$ entspricht. Betrachtet man die Analogie zur Formel für die Kovarianz der Variablen X und Y

$$\text{Cov}(X, Y) = E(X - \mu_X)(Y - \mu_Y),$$

wird deutlich warum man γ_{jt} die j -te Autokovarianz von Y_t nennt. Außerdem fällt auf, dass es sich bei der Kovarianz γ_{0t} um die Varianz σ_t^2 handelt.

Die Autokorrelation ρ_{jt} der Zeitreihe Y_t ergibt sich durch

$$\rho_{jt} = \frac{\gamma_{jt}}{\gamma_{0t}}.$$

3.6 Stationarität

Prinzipiell unterscheidet man zwischen schwacher und strenger Stationarität. Man spricht von strenger Stationarität, wenn die Verteilungsfunktion des Prozesses unabhängig vom Zeitpunkt t ist. Das bedeutet, dass alle Momente der Zeitreihe zu jedem Zeitpunkt t gleich sind.

Von schwacher Stationarität n -ter Ordnung spricht man, wenn nur die ersten n Momente des Prozesses gleich sind. Sind die ersten beiden Momente eines Prozesses gleich, so spricht man auch von einem Kovarianz-stationären Prozess. Für eine Zeitreihe Y_t , die durch einen solchen Prozess entstanden ist, gilt

$$\begin{aligned} E(Y_t) &= \mu && \text{für jedes } t \\ E(Y_t - \mu)(Y_{t-j} - \mu) &= \gamma_j && \text{für jedes } t \text{ für jedes } j. \end{aligned}$$

Der Mittelwert, die Varianz und die Kovarianz sind also konstant und damit unabhängig vom Zeitpunkt t . Außerdem gilt dann für die Autokorrelation

$$\gamma_j = \gamma_{-j}.$$

Ein einfaches Beispiel für einen solchen Prozess ist normalverteiltes Rauschen, so dass $y_t \sim N(0, \sigma^2)$ gilt. Im Folgenden soll mit Stationarität immer schwache Stationarität 2. Ordnung gemeint sein.

3.7 Ergodizität

Durch den selben Prozess kann man verschiedene Zeitreihen erzeugen. Da allerdings in der Regel nur eine einzelne Zeitreihe vorliegt, stellt sich die Frage, ob diese eine Realisierung des Prozesses ausreicht, um Rückschlüsse auf den Prozess machen zu können. Falls das nicht möglich ist, kann man die Parameter, die den Prozess steuern, nicht sicher bestimmen. Sind Rückschlüsse möglich, so nennt man den Prozess ergodisch.

Gilt für einen stationären Prozess

$$\sum_{j=0}^{\infty} |\gamma_j| < \infty,$$

so nennt man den Prozess Mittelwert-ergodisch. Einen stationären Prozess nennt man Kovarianz-ergodisch, wenn

$$\lim_{T \rightarrow \infty} \left(\frac{1}{T-j} \sum_{t=j+1}^T (Y_t - \mu)(Y_{t-j} - \mu) - \gamma_j \right) = 0$$

für alle j gilt.

Kapitel 4

Spezielle Prozesse

In Kapitel 1 wird auf die Möglichkeit hingewiesen, das farbige Rauschen, welches nach dem Abziehen des Trends von den Beobachtungen übrig bleibt, durch einen ARMA-Prozess zu modellieren. Die in Kapitel 6 dargestellte Modellidentifikation hat das Ziel die Modellordnung dieses ARMA-Prozesses zu bestimmen. Dabei kommen auch weißes Rauschen, AR-Prozesse und MA-Prozesse in Frage, da man sie als spezielle Fälle eines ARMA-Prozesses ansehen kann. Sie werden aus diesem Grund ebenfalls dargestellt. Die Grundlagen hierzu sind beispielsweise in [2], Kapitel 3 wiederzufinden.

Als alternative Modellierung des farbigen Rauschens wird in Kapitel 1 die Modellierung mit ARCH-Prozessen erwogen. Diese Prozesse werden ebenfalls in diesem Kapitel vorgestellt. Nachzulesen sind die Grundlagen zu ARCH-Prozessen beispielsweise in [2], Kapitel 21.

4.1 Weißes Rauschen

Der einfachste stationäre Prozess ist weißes Rauschen

$$Y_t = \{\varepsilon_1, \varepsilon_2, \dots, \varepsilon_T\}.$$

Wenn

$$\mu_t = 0, \quad \sigma_t^2 = \sigma^2, \quad \gamma_j = \begin{cases} \sigma^2 & \text{für } j = 0 \\ 0 & \text{für } j \neq 0 \end{cases} \quad \text{und} \quad \rho_j = \begin{cases} 1 & \text{für } j = 0 \\ 0 & \text{für } j \neq 0 \end{cases}$$

gilt, bezeichnet man diesen Prozess als weißes Rauschen. Sind die $\varepsilon_t \sim N(0, \sigma^2)$, so spricht man bei dem Prozess von einem GAUSSSchen weißen Rauschen.

4.2 $AR(p)$ Prozesse

Ein autoregressiver Prozess p -ter Ordnung, im Nachfolgenden $AR(p)$ notiert, ist definiert durch die Differenzgleichung

$$Y_t = c + \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \cdots + \phi_p Y_{t-p} + \varepsilon_t. \quad (4.1)$$

Dabei ist c eine Konstante und ε_t weißes Rauschen. Mit Hilfe des Backshift-Operators B kann der Prozess auch geschrieben werden als

$$Y_t = c + \phi_1 B Y_t + \phi_2 B^2 Y_t + \cdots + \phi_p B^p Y_t + \varepsilon_t.$$

Nach Ausklammern von Y_t erhält man

$$(1 - \phi_1 B - \phi_2 B^2 - \cdots - \phi_p B^p) Y_t = c + \varepsilon_t,$$

wobei man mit Hilfe des sogenannten Lag-Polynoms $\phi(B) = 1 - \phi_1 B - \phi_2 B^2 - \cdots - \phi_p B^p$ die verkürzte Schreibweise

$$\phi(B) Y_t = c + \varepsilon_t$$

für einen $AR(p)$ Prozess erhält. Liegen die Wurzeln des Lag-Polynoms $\phi(B)$ außerhalb des Einheitskreises, dann ist der Prozess stationär. Für einen stationären $AR(p)$ Prozess beträgt der Mittelwert

$$\mu = \frac{c}{1 - \phi_1 - \phi_2 - \cdots - \phi_p}.$$

Für die Autokovarianz erhält man

$$\gamma_j = \phi_1 \gamma_{j-1} + \phi_2 \gamma_{j-2} + \cdots + \phi_p \gamma_{j-p},$$

wobei $\gamma_j = \gamma_{-j}$ wegen der Stationarität des Prozesses gilt. Die Autokorrelation ρ_j erhält man, indem die Autokovarianz γ_j durch die Varianz des Prozesses γ_0 geteilt wird:

$$\rho_j = \phi_1 \rho_{j-1} + \phi_2 \rho_{j-2} + \cdots + \phi_p \rho_{j-p}.$$

Diese Gleichung wird als YULE-WALKER Gleichung bezeichnet, mit deren Hilfe man bei Kenntnis der Korrelationen $\rho_1, \rho_2, \dots, \rho_p$ die Parameter $\phi_1, \phi_2, \dots, \phi_p$ des Prozesses bestimmen kann. Umgekehrt ist es auch möglich die Korrelationen bei Kenntnis der Parameter zu bestimmen. Dazu stellt man unter Beachtung von $\rho_j = \rho_{-j}$ folgendes Gleichungssystem auf.

$$\begin{pmatrix} \rho_1 \\ \rho_2 \\ \vdots \\ \rho_p \end{pmatrix} = \begin{pmatrix} 1 & \rho_1 & \cdots & \rho_{p-1} \\ \rho_1 & 1 & \cdots & \rho_{p-2} \\ \vdots & \vdots & & \vdots \\ \rho_{p-1} & \rho_{p-2} & \cdots & 1 \end{pmatrix} \begin{pmatrix} \phi_1 \\ \phi_2 \\ \vdots \\ \phi_p \end{pmatrix}$$

4.3 $MA(q)$ Prozesse

Ein Moving Average Prozess q -ter Ordnung, im Nachfolgenden $MA(q)$ notiert, ist durch die Gleichung

$$Y_t = c + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \cdots + \theta_q \varepsilon_{t-q} \quad (4.2)$$

definiert. Dabei ist c eine Konstante und ε_t weißes Rauschen. Der Begriff Moving Average bedeutet übersetzt gleitendes Mittel. Dies kann dadurch erklärt werden, dass man den Prozess auch als ein gewichtetes gleitendes Mittel über das weiße Rauschen auffassen kann. Analog zu einem $AR(p)$ Prozess kann man den $MA(q)$ Prozess mit Hilfe des Lag-Polynoms $\theta(B) = 1 + \theta_1 B + \theta_2 B^2 + \cdots + \theta_q B^q$ auch verkürzt ausdrücken.

$$Y_t = c + \theta(B)\varepsilon_t$$

Der Prozess ist für beliebige Werte $\theta_1, \theta_2, \dots, \theta_q$ stationär. Da für das weiße Rauschen $\mu_{\varepsilon_t} = 0$ und $\sigma_{\varepsilon_t} = \sigma^2$ gilt, beträgt der Mittelwert der Zeitreihe

$$\mu = c$$

und die Varianz

$$\gamma_0 = (1 + \theta_1^2 + \theta_2^2 + \cdots + \theta_q^2)\sigma^2.$$

Für die Autokovarianz erhält man

$$\gamma_j = \begin{cases} (\theta_j + \theta_{j+1}\theta_1 + \theta_{j+2}\theta_2 + \cdots + \theta_q\theta_{q-j})\sigma^2 & \text{für } j = 1, 2, \dots, q \\ 0 & \text{für } j > q \end{cases}.$$

4.4 Invertierbare $MA(q)$ und $AR(p)$ Prozesse

Um die Stationarität des $MA(q)$ Prozesses zu wahren, sind keine Einschränkungen für die Parameter $\theta_1, \theta_2, \dots, \theta_q$ nötig. Es ist jedoch möglich einen $MA(q)$ Prozess in einen $AR(\infty)$ Prozess umzuwandeln, falls die Wurzeln des Lag-Polynoms $\theta(B)$ des $MA(q)$ Prozesses

$$Y_t = \theta(B)\varepsilon_t$$

außerhalb des Einheitskreises liegen. Man beachte, dass bei dem Prozess die Konstante c weggelassen wurde. Das ist zulässig, da man in der Gleichung (4.2) $Y_t - c$ durch \bar{Y}_t substituieren kann. Die Invertierung des $MA(q)$ Prozesses wird durch die Polynomdivision

$$\phi(B) = \frac{1}{\theta(B)}$$

vollzogen, wobei der Grad des Lag-Polynoms $\phi(B)$ unendlich ist.

Ist der $AR(p)$ Prozess

$$\phi(B) = c + \varepsilon_t$$

stationär, so kann man ihn ohne weitere Einschränkungen in einen $MA(\infty)$ Prozess umwandeln. Die Intertierung geschieht durch die Polynomdivision

$$\theta(B) = \frac{1}{\phi(B)},$$

wobei der Grad des Lag-Polynoms $\theta(B)$ unendlich ist.

4.5 $ARMA(p, q)$ Prozesse

Ein $ARMA(p, q)$ Prozess setzt sich aus einem $AR(p)$ Anteil und einem $MA(q)$ Anteil zusammen.

$$Y_t = c + \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \dots + \phi_p Y_{t-p} + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \dots + \theta_q \varepsilon_{t-q} \quad (4.3)$$

Mit Hilfe der Lag-Polynome

$$\begin{aligned} \phi(B) &= 1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p, \\ \theta(B) &= 1 + \theta_1 B + \theta_2 B^2 + \dots + \theta_q B^q \end{aligned}$$

kann man auch

$$\phi(B)Y_t = c + \theta(B)\varepsilon_t$$

schreiben. Der $ARMA(p, q)$ Prozess ist stationär, wenn die Wurzeln des Lag-Polynoms $\phi(B)$ außerhalb des Einheitskreises liegen. Der Mittelwert beträgt

$$\mu = \frac{c}{1 - \phi_1 - \phi_2 - \dots - \phi_p}.$$

Die Autokovarianz nimmt für $j \leq q$ eine komplizierte Gestalt an, da sich die Einflüsse für autoregressive und Moving Average Anteile mischen. Für $j > q$ hat der Moving Average Anteil keinen Einfluss mehr, so dass die Autokovarianz

$$\gamma_j = \phi_1 \gamma_{j-1} + \phi_2 \gamma_{j-2} + \dots + \phi_p \gamma_{j-p} \quad \text{für } j > q$$

beträgt. Falls die Wurzeln des Lag-Polynoms $\phi(B)$ außerhalb des Einheitskreises liegen, kann man den Prozess auch umschreiben in

$$Y_t = \mu + \psi(B)\varepsilon_t,$$

wobei

$$\psi(B) = \frac{\theta(B)}{\phi(B)} \quad \text{mit} \quad \sum_{j=0}^{\infty} |\psi_j| < \infty$$

ist. An dieser Gleichung erkennt man, dass eine redundante Parametrisierung wie beispielsweise $\theta_1 = -\phi_1$, $\theta_2 = -\phi_2$, \dots , $\theta_r = -\phi_r$ zu vermeiden ist, da sie in diesem Fall wegen

$\theta(B)/\phi(B) = 1$ zu einem Prozess führt, der nur den Mittelwert μ und weißes Rauschen ε_t enthält. Allgemein tritt eine Überparametrisierung dann ein, wenn einige oder alle Wurzeln der Lag-Polynome $\phi(B)$ und $\theta(B)$ gleich sind. Wenn $\lambda_1, \lambda_2, \dots, \lambda_p$ die Nullstellen des Lag-Polynoms $\phi(B)$ und $\eta_1, \eta_2, \dots, \eta_q$ die Nullstellen des Lag-Polynoms $\theta(B)$ sind, so muss

$$\lambda_j \neq \eta_k$$

für alle j und alle k gelten, um eine Überparametrisierung zu vermeiden.

4.6 ARCH(m) Prozesse

Die bis hierhin beschriebenen Prozesse sind stationär und besitzen daher eine konstante Varianz. Die Idee ist nun, dass die Varianz nicht konstant ist, sondern sich über die Zeit ändert. So könnten die Störterme u_t durch den AR(m) Prozess

$$u_t^2 = \zeta + \alpha_1 u_{t-1}^2 + \alpha_2 u_{t-2}^2 + \dots + \alpha_m u_{t-m}^2 + w_t \quad (4.4)$$

entstanden sein. Dabei ist ζ eine Konstante mit $\zeta > 0$ und w_t weißes Rauschen mit $-\zeta < w_t$. Solche Prozesse bezeichnet man als autoregressive bedingte heteroskedastische Prozesse der Ordnung m , welche im Folgenden nur noch ARCH(m) Prozesse genannt werden.

Solche ARCH(m) Prozesse kann man auch verkürzt notieren als

$$\alpha(B)u_t^2 = \zeta + w_t$$

mit dem Lag-Polynom $\alpha(B) = 1 - \alpha_1 B - \alpha_2 B^2 - \dots - \alpha_m B^m$. Damit sie stationär sind müssen die Wurzeln des Lag-Polynoms $\alpha(B)$ außerhalb des Einheitskreises liegen. Außerdem muss wegen der Nicht-Negativität der quadratischen Störterme u_t^2

$$\alpha_1, \alpha_2, \dots, \alpha_m \geq 0 \quad \text{und} \quad w_t > -\zeta \quad \text{für} \quad \zeta > 0$$

gelten. Wenn $\alpha_1, \alpha_2, \dots, \alpha_m \geq 0$ gilt, ist die Bedingung, dass alle Wurzeln des Lag-Polynoms $\alpha(B)$ außerhalb des Einheitskreises liegen, äquivalent zu der Bedingung

$$\alpha_1 + \alpha_2 + \dots + \alpha_m < 1.$$

Die unbedingte Varianz der Störterme u_t beträgt

$$E(u_t^2) = \frac{\zeta}{1 - \alpha_1 - \alpha_2 - \dots - \alpha_m}.$$

Es ist oft üblich den ARCH(m) Prozess durch die Gleichung

$$u_t = \sqrt{h_t} \cdot v_t$$

mit $h_t = \zeta + \alpha_1 u_{t-1}^2 + \alpha_2 u_{t-2}^2 + \cdots + \alpha_m u_{t-m}^2$ auszudrücken. Dabei ist v_t ein Folge unabhängiger und identisch verteilter Zufallsvariablen mit $E(v_t) = 0$ und $E(v_t^2) = 1$.

Üblicherweise wird $v_t \sim N(0, 1)$ angesetzt, so dass die *ARCH*(m) Prozesse einer Normalverteilung gehorchen. Das bedeutet allerdings nicht, dass die Störterme u_t normalverteilt sind. Die Verteilungsfunktion der Störterme u_t besitzt wesentlich mehr Masse in den Flanken im Vergleich zur Normalverteilung. Daher ist sie prinzipiell zur Modellierung von Ausreißern geeignet.¹

¹Siehe [10], Kapitel 20.1.2, Seite 351.

Kapitel 5

Maximum-Likelihood Schätzung

Die in Kapitel 4 vorgestellten Prozesse werden durch ihre Prozessparameter, die in der Gleichung des Prozesses enthalten sind, beschrieben. Eine häufig verwendete Methode zur Schätzung dieser Prozessparameter ist die Maximum-Likelihood Methode, welche in diesem Kapitel vorgestellt wird.

In Abschnitt 5.1 wird die Likelihoodfunktion vorgestellt und daraus die Log-Likelihoodfunktion abgeleitet, welche einfacher zu Handhaben ist. Die Grundlagen hierzu sind in [5], Kapitel 3.1.4 und 3.2.4 wiederzufinden. Anschließend wird die Log-Likelihoodfunktion für die in Kapitel 4 vorgestellten Prozesse hergeleitet, wobei auch der Unterschied zwischen exakter und bedingter Log-Likelihoodfunktion herausgestellt wird. Neben der Log-Likelihoodfunktion der verschiedenen Prozesse werden auch der Gradient und die HESSE-Matrix der jeweiligen Log-Likelihoodfunktion hergeleitet, welche bei den numerischen Optimierungsverfahren verwendet werden. Die Quelle hierfür ist [2], Kapitel 5.1 bis 5.6 und 21.1.

Bei der Bestimmung der besten Schätzung nach der Maximum-Likelihood Methode wird das Maximum der Log-Likelihoodfunktion ermittelt. Die verschiedenen numerischen Optimierungsverfahren, mit denen man dieses Maximum berechnen kann, werden in Abschnitt 5.2 vorgestellt. In [2], Kapitel 5.7 sind die Grundlagen dieser Verfahren dargestellt.

5.1 Likelihoodfunktion

Besitzt ein Zufallsvektor \mathbf{y} die von den unbekanntem, festen Parametern $\boldsymbol{\theta}$ abhängige Dichtefunktion $f(\boldsymbol{\theta})$, dann ist die Likelihoodfunktion $L(\mathbf{y}, \boldsymbol{\theta})$ definiert durch

$$L(\mathbf{y}, \boldsymbol{\theta}) = f(\boldsymbol{\theta}).$$

Die Dichtefunktion $f(\boldsymbol{\theta})$ des Zufallsvektors \mathbf{y} muss also für die Schätzung der Parameter $\boldsymbol{\theta}$ bekannt sein. Um die Parameter $\boldsymbol{\theta}$ zu schätzen, muss die Likelihoodfunktion maximiert werden. Folglich ist die beste Schätzung $\hat{\boldsymbol{\theta}}$ für die festen Parameter $\boldsymbol{\theta}$ gefunden, wenn

$$\hat{\boldsymbol{\theta}} = \arg \sup_{\boldsymbol{\theta}} L(\mathbf{y}, \boldsymbol{\theta})$$

gilt. Das Maximum der Likelihoodfunktion kann mit verschiedenen numerischen Optimierungsverfahren bestimmt werden. Oft ist es praktischer, nicht die Likelihoodfunktion, sondern die logarithmierte Likelihoodfunktion

$$\mathcal{L}(\mathbf{y}, \boldsymbol{\theta}) = \log L(\mathbf{y}, \boldsymbol{\theta})$$

zu verwenden. Diese Log-Likelihoodfunktion $\mathcal{L}(\mathbf{y}, \boldsymbol{\theta})$ an Stelle der Likelihoodfunktion $L(\mathbf{y}, \boldsymbol{\theta})$ zu maximieren ist zulässig, da die Dichtefunktion $f(\boldsymbol{\theta})$ positiv ist.

5.1.1 Likelihoodfunktion eines Gaußschen $AR(p)$ Prozesses

Ein $AR(p)$ Prozess

$$\phi(B)Y_t = c + \varepsilon_t$$

wird als GAUSSscher $AR(p)$ Prozess bezeichnet, wenn die Störterme ε_t unabhängig und identisch verteilt sind mit

$$\varepsilon_t \sim N(0, \sigma^2).$$

Zunächst benötigt man die Dichtefunktion der ersten p Beobachtungen y_1, y_2, \dots, y_p . Sie besitzen den Mittelwert

$$\mu = \frac{c}{1 - \phi_1 - \phi_2 - \dots - \phi_p}$$

und die Kovarianzmatrix

$$\sigma^2 \mathbf{V}_p = \begin{pmatrix} \gamma_0 & \gamma_1 & \gamma_2 & \dots & \gamma_{p-1} \\ \gamma_1 & \gamma_0 & \gamma_1 & \dots & \gamma_{p-2} \\ \gamma_2 & \gamma_1 & \gamma_0 & \dots & \gamma_{p-3} \\ \vdots & \vdots & \vdots & & \vdots \\ \gamma_{p-1} & \gamma_{p-2} & \gamma_{p-3} & \dots & \gamma_0 \end{pmatrix},$$

wobei γ_j die j -te Autokovarianz des $AR(p)$ Prozesses ist. Die ersten p Beobachtungen entsprechen wegen der Verteilung der Störterme ε_t einer Normalverteilung.

$$y_1, y_2, \dots, y_p \sim N(\boldsymbol{\mu}_p, \sigma^2 \mathbf{V}_p)$$

Folglich ergibt sich ihre gemeinsame Dichtefunktion durch

$$f_{Y_p, Y_{p-1}, \dots, Y_1}(y_p, y_{p-1}, \dots, y_1; \boldsymbol{\theta}) = \frac{1}{(2\pi\sigma^2)^{p/2}} \sqrt{|\mathbf{V}_p^{-1}|} \exp\left(-\frac{1}{2\sigma^2}(\mathbf{y}_p - \boldsymbol{\mu}_p)' \mathbf{V}_p^{-1}(\mathbf{y}_p - \boldsymbol{\mu}_p)\right),$$

wobei $|\mathbf{V}_p^{-1}|$ die Determinante der Inversen von \mathbf{V}_p bezeichnet und \mathbf{y}_p ein $p \times 1$ Vektor ist, der die ersten p Beobachtungen y_1, y_2, \dots, y_p enthält. Die Elemente des $p \times 1$ Vektors $\boldsymbol{\mu}_p$ sind gleich μ . Der Vektor $\boldsymbol{\theta} = (c, \phi_1, \phi_2, \dots, \phi_p, \sigma^2)'$ enthält die Parameter des $AR(p)$ Prozesses.

Mit Hilfe der Dichtefunktion der ersten p Beobachtungen kann man die bedingten Dichtefunktionen der nachfolgenden Beobachtungen $y_{p+1}, y_{p+2}, \dots, y_T$ angeben. Unter der Bedingung, dass y_1, y_2, \dots, y_p beobachtet wurde, erhält man wegen

$$y_{p+1} - c - \phi_1 y_p - \phi_2 y_{p-1} - \dots - \phi_p y_1 = \varepsilon_{p+1} \sim N(0, \sigma^2)$$

die bedingte Dichtefunktion

$$f_{Y_{p+1}|Y_p, Y_{p-1}, \dots, Y_1}(y_{p+1}|y_p, y_{p-1}, \dots, y_1; \boldsymbol{\theta}) = \frac{1}{2\pi\sigma^2} \exp\left(\frac{-\varepsilon_{p+1}^2}{2\sigma^2}\right)$$

für die Beobachtung y_{p+1} . Analog dazu kann man wegen der Prozessgleichung

$$y_t - c - \phi_1 y_{t-1} - \phi_2 y_{t-2} - \dots - \phi_p y_{t-p} = \varepsilon_t \sim N(0, \sigma^2)$$

die bedingte Dichtefunktion

$$f_{Y_t|Y_{t-1}, Y_{t-2}, \dots, Y_{t-p}}(y_t|y_{t-1}, y_{t-2}, \dots, y_{t-p}; \boldsymbol{\theta}) = \frac{1}{2\pi\sigma^2} \exp\left(\frac{-\varepsilon_t^2}{2\sigma^2}\right)$$

für jede Beobachtung y_t mit $t > p$ angeben. Die gemeinsame Dichtefunktion aller Beobachtungen $f_{Y_T, Y_{T-1}, \dots, Y_1}(y_T, y_{T-1}, \dots, y_1; \boldsymbol{\theta})$ ergibt sich aus dem Produkt der gemeinsamen Dichtefunktion der Beobachtungen y_1, y_2, \dots, y_p und der bedingten Dichtefunktionen der nachfolgenden Beobachtungen $y_{p+1}, y_{p+2}, \dots, y_T$.

$$\begin{aligned} f_{Y_T, Y_{T-1}, \dots, Y_1}(y_T, y_{T-1}, \dots, y_1; \boldsymbol{\theta}) &= f_{Y_p, Y_{p-1}, \dots, Y_1}(y_p, y_{p-1}, \dots, y_1; \boldsymbol{\theta}) \\ &\quad \times \prod_{t=p+1}^T f_{Y_t|Y_{t-1}, Y_{t-2}, \dots, Y_{t-p}}(y_t|y_{t-1}, y_{t-2}, \dots, y_{t-p}; \boldsymbol{\theta}) \end{aligned}$$

Diese gemeinsame Dichtefunktion aller Beobachtungen ist die exakte Likelihoodfunktion. Die exakte Log-Likelihoodfunktion erhält man durch logarithmieren der exakten Likelihoodfunktion, wodurch sich das Produkt der Dichtefunktionen wegen der Rechenregeln des Logarithmus in eine Summe umwandelt.

$$\begin{aligned} \mathcal{L}_e(\boldsymbol{\theta}) &= \log f_{Y_T, Y_{T-1}, \dots, Y_1}(y_T, y_{T-1}, \dots, y_1; \boldsymbol{\theta}) \\ &= \log f_{Y_p, Y_{p-1}, \dots, Y_1}(y_p, y_{p-1}, \dots, y_1; \boldsymbol{\theta}) \\ &\quad + \sum_{t=p+1}^T \log f_{Y_t|Y_{t-1}, Y_{t-2}, \dots, Y_{t-p}}(y_t|y_{t-1}, y_{t-2}, \dots, y_{t-p}; \boldsymbol{\theta}) \end{aligned}$$

Setzt man die Formeln für die einzelnen Dichtefunktionen ein, so erhält man wiederum unter Anwendung der Rechenregeln des Logarithmus die exakte Log-Likelihoodfunktion

$$\begin{aligned} \mathcal{L}_e(\boldsymbol{\theta}) &= -\frac{T}{2} \log(2\pi) - \frac{T}{2} \log(\sigma^2) - \frac{1}{2} \log |\mathbf{V}_p^{-1}| \\ &\quad - \frac{1}{2\sigma^2} (\mathbf{y}_p - \boldsymbol{\mu}_p)' \mathbf{V}_p^{-1} (\mathbf{y}_p - \boldsymbol{\mu}_p) - \sum_{t=p+1}^T \frac{\varepsilon_t^2}{2\sigma^2}. \end{aligned}$$

Vernachlässigt man den Anteil $\log(f_{Y_T, Y_{T-1}, \dots, Y_1}(y_T, y_{T-1}, \dots, y_1; \boldsymbol{\theta}))$ der ersten p Beobachtungen, so erhält man die bedingte Likelihoodfunktion

$$\begin{aligned} \mathcal{L}(\boldsymbol{\theta}) &= \sum_{t=p+1}^T \log f_{Y_t|Y_{t-1}, Y_{t-2}, \dots, Y_{t-p}}(y_t|y_{t-1}, y_{t-2}, \dots, y_{t-p}; \boldsymbol{\theta}) \\ &= -\frac{T-p}{2} \log(2\pi) - \frac{T-p}{2} \log(\sigma^2) - \sum_{t=p+1}^T \frac{\varepsilon_t^2}{2\sigma^2}, \end{aligned}$$

welche bei der numerischen Bestimmung des Maximums einfacher zu handhaben ist als die exakte Likelihoodfunktion. Für lange Zeitreihen ist diese Vereinfachung zulässig.¹ Der Gradient der bedingten Log-Likelihoodfunktion $\nabla \mathcal{L}(\boldsymbol{\theta})$ ergibt sich durch Ableiten der Log-Likelihoodfunktion $\mathcal{L}(\boldsymbol{\theta})$ nach den Parametern $\boldsymbol{\theta}$.

$$\nabla \mathcal{L}(\boldsymbol{\theta}) = \frac{\partial \mathcal{L}(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ -\frac{T-p}{2\sigma^2} \end{pmatrix} + \frac{1}{\sigma^2} \sum_{t=p+1}^T \begin{pmatrix} \varepsilon_t \\ \varepsilon_t y_{t-1} \\ \varepsilon_t y_{t-2} \\ \vdots \\ \varepsilon_t y_{t-p} \\ \frac{\varepsilon_t^2}{2\sigma^2} \end{pmatrix}$$

Die HESSE-Matrix der bedingten Log-Likelihoodfunktion $\mathbf{H}(\mathcal{L}(\boldsymbol{\theta}))$ ergibt sich durch zweifaches Ableiten der Log-Likelihoodfunktion $\mathcal{L}(\boldsymbol{\theta})$ nach den Parametern $\boldsymbol{\theta}$.

$$\begin{aligned} \mathbf{H}(\mathcal{L}(\boldsymbol{\theta})) &= \frac{\partial^2 \mathcal{L}(\boldsymbol{\theta})}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}'} \\ &= \begin{pmatrix} 0 & \dots & 0 & 0 \\ \vdots & & \vdots & \vdots \\ 0 & \dots & 0 & 0 \\ 0 & \dots & 0 & \frac{T-p}{2\sigma^4} \end{pmatrix} - \frac{1}{\sigma^2} \sum_{t=p+1}^T \begin{pmatrix} 1 & y_{t-1} & \dots & y_{t-p} & \frac{\varepsilon_t}{\sigma^2} \\ y_{t-1} & y_{t-1}^2 & \dots & y_{t-1}y_{t-p} & \frac{y_{t-1}\varepsilon_t}{\sigma^2} \\ y_{t-2} & y_{t-2}y_{t-1} & \dots & y_{t-2}y_{t-p} & \frac{y_{t-2}\varepsilon_t}{\sigma^2} \\ \vdots & \vdots & & \vdots & \vdots \\ y_{t-p} & y_{t-p}y_{t-1} & \dots & y_{t-p}^2 & \frac{y_{t-p}\varepsilon_t}{\sigma^2} \\ \frac{\varepsilon_t}{\sigma^2} & \frac{y_{t-1}\varepsilon_t}{\sigma^2} & \dots & \frac{y_{t-p}\varepsilon_t}{\sigma^2} & \frac{\varepsilon_t^2}{\sigma^4} \end{pmatrix} \end{aligned}$$

¹Wenn man σ^2 als Parameter in $\boldsymbol{\theta}$ vernachlässigt, dann sind die Terme $-\frac{T-p}{2} \log(2\pi)$ und $-\frac{T-p}{2} \log(\sigma^2)$ konstant und die Maximierung der bedingten Log-Likelihoodfunktion $\mathcal{L}(\boldsymbol{\theta})$ entspricht einer Minimierung nach der Methode der kleinsten Quadrate.

$$-\sum_{t=p+1}^T \frac{\varepsilon_t^2}{2\sigma^2} \longrightarrow \text{Max.} \quad \doteq \quad \sum_{t=p+1}^T \varepsilon_t^2 \longrightarrow \text{Min.}$$

5.1.2 Likelihoodfunktion eines Gaußschen $MA(q)$ Prozesses

Bei einem GAUSSschen $MA(q)$ Prozess sind die Störterme ε_t unabhängig und identisch verteilt mit $\varepsilon_t \sim N(0, \sigma^2)$. Für die Beobachtungen gilt daher

$$\mathbf{y} \sim N(\boldsymbol{\mu}, \sigma^2 \mathbf{V}),$$

wobei $\mathbf{y} = (y_1, y_2, \dots, y_T)'$, $\boldsymbol{\mu} = (\mu, \mu, \dots, \mu)'$ und

$$\sigma^2 \mathbf{V}_p = \begin{pmatrix} \gamma_0 & \gamma_1 & \gamma_2 & \dots & \gamma_{T-1} \\ \gamma_1 & \gamma_0 & \gamma_1 & \dots & \gamma_{T-2} \\ \gamma_2 & \gamma_1 & \gamma_2 & \dots & \gamma_{T-3} \\ \vdots & \vdots & \vdots & & \vdots \\ \gamma_{T-1} & \gamma_{T-2} & \gamma_{T-3} & \dots & \gamma_0 \end{pmatrix}$$

gilt. Die Kovarianzmatrix $\sigma^2 \mathbf{V}_p$ ist wegen $\gamma_j = 0$ für $j > q$ eine Bandmatrix. Mit der Abkürzung $\mathbf{Y} = Y_t, Y_{t-1}, \dots, Y_1$ lautet die exakte Likelihoodfunktion

$$f_{\mathbf{Y}}(\mathbf{y}; \boldsymbol{\theta}) = \frac{1}{(2\pi\sigma^2)^{T/2}} \sqrt{|\mathbf{V}^{-1}|} \exp\left(-\frac{1}{2\sigma^2}(\mathbf{y} - \boldsymbol{\mu})' \mathbf{V}^{-1}(\mathbf{y} - \boldsymbol{\mu})\right).$$

Die exakte Likelihoodfunktion kann mit Hilfe eines Kalmanfilters oder mit einer Dreieckszerlegung $\sigma^2 \mathbf{V} = \mathbf{L} \mathbf{D} \mathbf{L}'$, wobei \mathbf{L} eine linke untere Dreiecksmatrix bezeichnet, ausgewertet werden. Da die bedingte Likelihoodfunktion auch hier einfacher zu handhaben ist, soll nicht näher auf die exakte Likelihoodfunktion eines $MA(q)$ Prozesses eingegangen werden.

Bei der bedingten Likelihoodfunktion eines $MA(q)$ Prozesses wird die Annahme gemacht, dass $\varepsilon_0 = \varepsilon_{-1} = \dots = \varepsilon_{-q+1} = 0$ ist. Mit diesen Startwerten und der Vorgabe von $\boldsymbol{\theta}$ ergeben sich alle weiteren Störterme ε_t durch die Prozessgleichung

$$\varepsilon_t = y_t - c - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - \dots - \theta_q \varepsilon_{t-q}.$$

So lautet die gemeinsame bedingte Dichtefunktion

$$f_{Y_t, Y_{t-1}, \dots, Y_1}(y_T, y_{T-1}, \dots, y_1 | \varepsilon_0 = \varepsilon_{-1} = \dots = \varepsilon_{-q+1} = 0; \boldsymbol{\theta}) = \frac{1}{(2\pi\sigma^2)^{T/2}} \exp\left(-\sum_{t=1}^T \frac{\varepsilon_t^2}{2\sigma^2}\right),$$

woraus die bedingte Log-Likelihoodfunktion

$$\begin{aligned} \mathcal{L}(\boldsymbol{\theta}) &= \log f_{Y_t, Y_{t-1}, \dots, Y_1}(y_T, y_{T-1}, \dots, y_1 | \varepsilon_0 = \varepsilon_{-1} = \dots = \varepsilon_{-q+1} = 0; \boldsymbol{\theta}) \\ &= -\frac{T}{2} \log(2\pi) - \frac{T}{2} \log(\sigma^2) - \sum_{t=1}^T \frac{\varepsilon_t^2}{2\sigma^2} \end{aligned}$$

folgt mit $\boldsymbol{\theta} = (c, \theta_1, \theta_2, \dots, \theta_q, \sigma^2)'$. Dabei müssen die Wurzeln des Lag-Polynoms $\theta(B) = 1 + \theta_1 B + \theta_2 B^2 + \dots + \theta_q B^q$ außerhalb des Einheitskreises liegen.

Den Gradienten der bedingten Log-Likelihoodfunktion $\nabla \mathcal{L}(\boldsymbol{\theta})$ erhält man durch Differenzieren der Log-Likelihoodfunktion nach den Parametern $\boldsymbol{\theta}$.

$$\nabla \mathcal{L}(\boldsymbol{\theta}) = \frac{\partial \mathcal{L}(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ -\frac{T}{2\sigma^2} \end{pmatrix} + \frac{1}{\sigma^2} \sum_{t=1}^T \begin{pmatrix} \varepsilon_t \\ \varepsilon_t \varepsilon_{t-1} \\ \varepsilon_t \varepsilon_{t-2} \\ \vdots \\ \varepsilon_t \varepsilon_{t-q} \\ \frac{\varepsilon_t^2}{2\sigma^2} \end{pmatrix}$$

Durch zweifaches Differenzieren der Log-Likelihoodfunktion erhält man die HESSE-Matrix

$$\mathbf{H}(\mathcal{L}(\boldsymbol{\theta})) = \frac{\partial^2 \mathcal{L}(\boldsymbol{\theta})}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}'}$$

$$= \begin{pmatrix} 0 & \dots & 0 & 0 \\ \vdots & & \vdots & \vdots \\ 0 & \dots & 0 & 0 \\ 0 & \dots & 0 & \frac{T}{2\sigma^4} \end{pmatrix} - \frac{1}{\sigma^2} \sum_{t=1}^T \begin{pmatrix} 1 & \varepsilon_{t-1} & \dots & \varepsilon_{t-q} & \frac{\varepsilon_t}{\sigma^2} \\ \varepsilon_{t-1} & \varepsilon_{t-1}^2 & \dots & \varepsilon_{t-1} \varepsilon_{t-q} & \frac{\varepsilon_t \varepsilon_{t-1}}{\sigma^2} \\ \varepsilon_{t-2} & \varepsilon_{t-2} \varepsilon_{t-1} & \dots & \varepsilon_{t-2} \varepsilon_{t-q} & \frac{\varepsilon_t \varepsilon_{t-2}}{\sigma^2} \\ \vdots & \vdots & & \vdots & \vdots \\ \varepsilon_{t-q} & \varepsilon_{t-q} \varepsilon_{t-1} & \dots & \varepsilon_{t-q}^2 & \frac{\varepsilon_t \varepsilon_{t-q}}{\sigma^2} \\ \frac{\varepsilon_t}{\sigma^2} & \frac{\varepsilon_t \varepsilon_{t-1}}{\sigma^2} & \dots & \frac{\varepsilon_t \varepsilon_{t-q}}{\sigma^2} & \frac{\varepsilon_t^2}{\sigma^4} \end{pmatrix}.$$

5.1.3 Likelihoodfunktion eines Gaußschen $ARMA(p, q)$ Prozesses

Ein $ARMA(p, q)$ Prozess

$$\phi(B)Y_t = c + \theta(B)\varepsilon_t$$

wird als GAUSSscher $ARMA(p, q)$ Prozess bezeichnet, wenn die Störterme ε_t unabhängig und identisch verteilt sind mit

$$\varepsilon_t \sim N(0, \sigma^2).$$

Auch hier ist die bedingte Likelihoodfunktion einfacher zu handhaben, so dass nicht auf die exakte Likelihoodfunktion eingegangen wird. Bei der bedingten Likelihoodfunktion setzt man die Startwerte $\mathbf{y}_0 = (y_0, y_{-1}, \dots, y_{-p+1})'$ und $\boldsymbol{\varepsilon}_0 = (\varepsilon_0, \varepsilon_{-1}, \dots, \varepsilon_{-q+1})'$ als gegeben voraus, so dass man mit ihnen die nachfolgenden Störterme ε_t durch die Prozessgleichung

$$\varepsilon_t = y_t - c - \phi_1 y_{t-1} - \phi_2 y_{t-2} - \dots - \phi_p y_{t-p} - \theta_1 \varepsilon_t - \theta_2 \varepsilon_{t-2} - \dots - \theta_q \varepsilon_{t-q}$$

unter der Vorgabe von $\boldsymbol{\theta} = (c, \phi_1, \phi_2, \dots, \phi_p, \theta_1, \theta_2, \dots, \theta_q, \sigma^2)'$ berechnen kann. Die gemeinsame bedingte Dichtefunktion ist demnach

$$f_{Y_t, Y_{t-1}, \dots, Y_1 | \mathbf{Y}_0, \boldsymbol{\varepsilon}_0}(y_T, y_{T-1}, \dots, y_1 | \boldsymbol{\varepsilon}_0, \mathbf{y}_0; \boldsymbol{\theta}) = \frac{1}{(2\pi\sigma^2)^{T/2}} \exp\left(-\sum_{t=1}^T \frac{\varepsilon_t^2}{2\sigma^2}\right),$$

woraus die bedingte Log-Likelihoodfunktion

$$\begin{aligned}\mathcal{L}(\boldsymbol{\theta}) &= \log f_{Y_t, Y_{t-1}, \dots, Y_1 | \mathbf{Y}_0, \boldsymbol{\varepsilon}_0}(y_T, y_{T-1}, \dots, y_1 | \boldsymbol{\varepsilon}_0, \mathbf{y}_0; \boldsymbol{\theta}) \\ &= -\frac{T}{2} \log(2\pi) - \frac{T}{2} \log(\sigma^2) - \sum_{t=1}^T \frac{\varepsilon_t^2}{2\sigma^2},\end{aligned}$$

folgt. Für das Setzen der Startwerte \mathbf{y}_0 und $\boldsymbol{\varepsilon}_0$ sollen hier zwei Möglichkeiten vorgestellt werden. Die erste Möglichkeit ist, die Startwerte gleich ihrem Erwartungswert zu setzen. Das wäre $y_0 = y_{-1} = \dots = y_{-p+1} = c/(1 - \phi_1 - \phi_2 - \dots - \phi_p)$ und $\varepsilon_0 = \varepsilon_{-1} = \dots = \varepsilon_{-q+1} = 0$. Die zweite Möglichkeit ist, für die Startwerte \mathbf{y}_0 die beobachteten Werte am Anfang der Zeitreihe zu nehmen und $\varepsilon_0 = \varepsilon_{-1} = \dots = \varepsilon_{-q+1} = 0$ zu setzen. In diesem Fall werden die ersten p Beobachtungen aus der Berechnung der bedingten Log-Likelihoodfunktion herausgenommen.

$$\mathcal{L}(\boldsymbol{\theta}) = -\frac{T-p}{2} \log(2\pi) - \frac{T-p}{2} \log(\sigma^2) - \sum_{t=p+1}^T \frac{\varepsilon_t^2}{2\sigma^2}$$

Der Gradient der Log-Likelihoodfunktion lautet

$$\nabla \mathcal{L}(\boldsymbol{\theta}) = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ -\frac{T-p}{2\sigma^2} \end{pmatrix} + \frac{1}{\sigma^2} \sum_{t=p+1}^T \begin{pmatrix} \varepsilon_t \\ \varepsilon_t y_{t-1} \\ \varepsilon_t y_{t-2} \\ \vdots \\ \varepsilon_t y_{t-p} \\ \varepsilon_t \varepsilon_{t-1} \\ \varepsilon_t \varepsilon_{t-2} \\ \vdots \\ \varepsilon_t \varepsilon_{t-q} \\ \frac{\varepsilon_t^2}{2\sigma^2} \end{pmatrix}.$$

Die HESSE-Matrix der bedingten Log-Likelihoodfunktion $\mathbf{H}(\mathcal{L}(\boldsymbol{\theta}))$ ergibt sich durch zweifaches Ableiten der Log-Likelihoodfunktion $\mathcal{L}(\boldsymbol{\theta})$ nach den Parametern $\boldsymbol{\theta}$.

$$\mathbf{H}(\mathcal{L}(\boldsymbol{\theta})) = \begin{pmatrix} 0 & \dots & 0 & 0 \\ \vdots & & \vdots & \vdots \\ 0 & \dots & 0 & 0 \\ 0 & \dots & 0 & \frac{T-p}{2\sigma^4} \end{pmatrix}$$

$$-\frac{1}{\sigma^2} \sum_{t=p+1}^T \begin{pmatrix} 1 & y_{t-1} & \dots & y_{t-p} & \varepsilon_{t-1} & \dots & \varepsilon_{t-q} & \frac{\varepsilon_t}{\sigma^2} \\ y_{t-1} & y_{t-1}^2 & \dots & y_{t-1}y_{t-p} & y_{t-1}\varepsilon_{t-1} & \dots & y_{t-1}\varepsilon_{t-q} & \frac{y_{t-1}\varepsilon_t}{\sigma^2} \\ y_{t-2} & y_{t-2}y_{t-1} & \dots & y_{t-2}y_{t-p} & y_{t-2}\varepsilon_{t-1} & \dots & y_{t-2}\varepsilon_{t-q} & \frac{y_{t-2}\varepsilon_t}{\sigma^2} \\ \vdots & \vdots & & \vdots & \vdots & & \vdots & \vdots \\ y_{t-p} & y_{t-p}y_{t-1} & \dots & y_{t-p}^2 & y_{t-p}\varepsilon_{t-1} & \dots & y_{t-p}\varepsilon_{t-q} & \frac{y_{t-p}\varepsilon_t}{\sigma^2} \\ \varepsilon_{t-1} & \varepsilon_{t-1}^2 & \dots & \varepsilon_{t-1}\varepsilon_{t-p} & \varepsilon_{t-1}\varepsilon_{t-1} & \dots & \varepsilon_{t-1}\varepsilon_{t-q} & \frac{\varepsilon_{t-1}\varepsilon_t}{\sigma^2} \\ \varepsilon_{t-2} & \varepsilon_{t-2}\varepsilon_{t-1} & \dots & \varepsilon_{t-2}\varepsilon_{t-p} & \varepsilon_{t-2}\varepsilon_{t-1} & \dots & \varepsilon_{t-2}\varepsilon_{t-q} & \frac{\varepsilon_{t-2}\varepsilon_t}{\sigma^2} \\ \vdots & \vdots & & \vdots & \vdots & & \vdots & \vdots \\ \varepsilon_{t-q} & \varepsilon_{t-q}\varepsilon_{t-1} & \dots & \varepsilon_{t-q}^2 & \varepsilon_{t-q}\varepsilon_{t-1} & \dots & \varepsilon_{t-q}\varepsilon_{t-q} & \frac{\varepsilon_{t-q}\varepsilon_t}{\sigma^2} \\ \frac{\varepsilon_t}{\sigma^2} & \frac{\varepsilon_t y_{t-1}}{\sigma^2} & \dots & \frac{\varepsilon_t y_{t-p}}{\sigma^2} & \frac{\varepsilon_t \varepsilon_{t-1}}{\sigma^2} & \dots & \frac{\varepsilon_t \varepsilon_{t-q}}{\sigma^2} & \frac{\varepsilon_t^2}{\sigma^4} \end{pmatrix}$$

Die Summe beim Gradienten und bei der Hessematrix kann auch bei $t = 1$ beginnen. Dann ist $T - p$ durch T in dem Term außerhalb der Summe zu ersetzen. Welche der Varianten zu nehmen ist, entscheidet sich durch die Wahl der Startwerte.

5.1.4 Likelihoodfunktion eines Regressionsmodells mit Gaußschen $ARCH(m)$ Störtermen

Das Regressionsmodell

$$y_t = \mathbf{x}'_t \boldsymbol{\beta} + u_t$$

mit den GAUSSSchen $ARCH(m)$ Störtermen u_t sei gegeben. Der Vektor $\boldsymbol{\beta}$ enthält die Regressionsparameter. Ein $ARCH(m)$ Prozess wird als GAUSSScher $ARCH(m)$ Prozess bezeichnet, wenn die Störterme u_t durch die Prozessgleichung

$$u_t = \sqrt{h_t} \cdot v_t \quad \text{mit} \quad h_t = \zeta + \alpha_1 u_{t-1}^2 + \alpha_2 u_{t-2}^2 + \dots + \alpha_m u_{t-m}^2$$

entstehen und dabei die v_t unabhängige gleich verteilte Zufallsvariablen sind mit

$$v_t \sim N(0, 1).$$

Die Verteilung der einzelnen Störterme u_t gehorcht folglich einer Normalverteilung mit

$$u_t \sim N(0, h_t).$$

Mit den Abkürzungen $\mathbf{y}_0 = (y_0, y_{-1}, \dots, y_{-m+1})'$ und $\mathbf{y}_t = (y_t, y_{t-1}, \dots, y_1)'$ kann man daher die bedingte Verteilung einer einzelnen Beobachtung y_t angeben.

$$f(y_t | \mathbf{y}_{t-1}, \mathbf{y}_0, \mathbf{x}'_t, \mathbf{x}'_{t-1}, \dots, \mathbf{x}'_1) = \frac{1}{\sqrt{2\pi h_t}} \exp\left(-\frac{u_t^2}{2h_t}\right)$$

Die gemeinsame bedingte Dichtefunktion ergibt sich aus dem Produkt aller einzelnen Dichtefunktionen.

$$f(y_t, y_{t-1}, \dots, y_1 | \mathbf{y}_0, \mathbf{x}'_t, \mathbf{x}'_{t-1}, \dots, \mathbf{x}'_1) = \prod_{t=1}^T f(y_t | \mathbf{y}_{t-1}, \mathbf{y}_0, \mathbf{x}'_t, \mathbf{x}'_{t-1}, \dots, \mathbf{x}'_1)$$

Da sich durch Logarithmieren dieses Produkt in eine Summe umwandelt, erhält man die Log-Likelihoodfunktion

$$\begin{aligned}\mathcal{L}(\boldsymbol{\theta}) &= \sum_{t=1}^T \log f(y_t | \mathbf{y}_{t-1}, \mathbf{y}_0, \mathbf{x}'_t, \mathbf{x}'_{t-1}, \dots, \mathbf{x}'_1) \\ &= -\frac{T}{2} \log(2\pi) - \frac{1}{2} \sum_{t=1}^T \log(h_t) - \frac{1}{2} \sum_{t=1}^T \frac{u_t^2}{h_t}.\end{aligned}$$

Eine Möglichkeit für die Vorgabe der Startwerte ist, die ersten p Beobachtungen zu verwenden. Das entspricht dem Abtrennen der ersten p Werte der Zeitreihe. Nach dem Abtrennen habe die Zeitreihe dann die Länge T . Die Störterme u_t erhält man aus dem Regressionsmodell unter Vorgabe des Parametervektors $\boldsymbol{\theta} = (\beta_1, \beta_2, \dots, \beta_r, \zeta, \alpha_1, \alpha_2, \dots, \alpha_m)'$.

$$u_t = y_t - \mathbf{x}'_t \boldsymbol{\beta}$$

Den Gradienten erhält man durch Differenzieren der Log-Likelihoodfunktion. Mit den Abkürzungen $\tilde{\mathbf{u}} = (\sum_{j=1}^m -2\alpha_j u_{t-j} \mathbf{x}'_{t-j}, 1, u_{t-1}^2, u_{t-2}^2, \dots, u_{t-m}^2)'$ und $\tilde{\mathbf{x}} = (\mathbf{x}'_t, 0, \dots, 0)'$ kann man den Gradienten schreiben als

$$\nabla \mathcal{L}(\boldsymbol{\theta}) = \sum_{t=1}^T \left(\frac{u_t^2 - h_t}{2h_t^2} \tilde{\mathbf{u}} + \frac{u_t}{h_t} \tilde{\mathbf{x}} \right).$$

Für die HESSE-Matrix erhält man nach zweimaligen Differenzieren der Log-Likelihoodfunktion

$$\mathbf{H}(\mathcal{L}(\boldsymbol{\theta})) = \sum_{t=1}^T \left(-\frac{1}{2h_t^2} \tilde{\mathbf{u}} \tilde{\mathbf{u}}' - \frac{1}{h_t} \tilde{\mathbf{x}} \tilde{\mathbf{x}}' \right).$$

5.2 Numerische Optimierungsverfahren

In den vorigen Abschnitten wurde dargestellt, wie man die Log-Likelihoodfunktion $\mathcal{L}(\boldsymbol{\theta})$ für die in Kapitel 4 aufgeführten Prozesse erhält. Mit den beobachteten Daten \mathbf{y} und den gegebenen Formeln kann nun der Wert $\mathcal{L}(\boldsymbol{\theta})$ für jede Vorgabe von $\boldsymbol{\theta}$ berechnet werden. Dieser Abschnitt wird verschiedene Optimierungsverfahren vorstellen, mit denen das Maximum der Log-Likelihoodfunktion $\mathcal{L}(\boldsymbol{\theta})$ numerisch bestimmt werden kann. Diese Maximierung liefert die beste Schätzung für die Prozessparameter.

5.2.1 Gittersuche

Die zu Grunde liegende Idee bei der Gittersuche ist, die Parameter systematisch zu verändern und für jeden Wert der Parameter die Log-Likelihoodfunktion auszuwerten. Um das Vorgehen zu veranschaulichen, kann man sich einen ARMA(1,1) Prozess vorstellen:

$$y_t + \alpha_1 y_{t-1} = \epsilon_t + \beta_1 \epsilon_{t-1}.$$

Das Gitter ergibt sich dann durch Variieren der Parameter α_1 und β_1 , wie beispielsweise

$$-0.9 \leq \alpha_1 \leq +0.9 \quad \text{und} \quad -0.9 \leq \beta_1 \leq +0.9$$

jeweils in 0.1 Schritten. Wäre der Gitterpunkt mit den Werten $\alpha_1 = 0.3$ und $\beta_1 = 0.7$ derjenige mit dem größten Wert der Log-Likelihoodfunktion, so würde man das Gitter dort weiter verfeinern.

Dieses Vorgehen ist einfach umzusetzen, besitzt jedoch ein paar entscheidende Nachteile. Einerseits kann die Log-Likelihoodfunktion mehrere lokale Maxima besitzen. So muss das Gitter also eng genug sein, um das globale Maximum zu erfassen. Ist das Gitter nicht eng genug, könnte ein lokales Maximum als globales erkannt werden. Andererseits wird das Gitter für eine steigende Anzahl von Parametern schnell sehr groß. Dieses Optimierungsverfahren kann folglich sehr viel rechenintensiver als andere sein.

5.2.2 Steilster Anstieg

Hierbei beginnt die Suche nach dem Maximum der Log-Likelihoodfunktion an einem vorgegebenen Startpunkt $\boldsymbol{\theta}^{(0)}$. In der einfachsten Version dieses Optimierungsverfahren geht man einen Schritt mit vorgegebener Länge λ vom Startpunkt in Richtung des Gradienten vom Startpunkt $\boldsymbol{\theta}^{(0)}$ aus:

$$\boldsymbol{\theta}^{(1)} = \boldsymbol{\theta}^{(0)} + \lambda \frac{\nabla \mathcal{L}(\boldsymbol{\theta}^{(0)})}{|\nabla \mathcal{L}(\boldsymbol{\theta}^{(0)})|}.$$

Den Gradienten an der Stelle des Startpunktes $\nabla \mathcal{L}(\boldsymbol{\theta}^{(0)})$ erhält man durch Differenzieren der Log-Likelihoodfunktion $\mathcal{L}(\boldsymbol{\theta})$ nach den Parametern $\boldsymbol{\theta}$ an der Stelle des Startpunktes:

$$\nabla \mathcal{L}(\boldsymbol{\theta}^{(0)}) = \left. \frac{\partial \mathcal{L}(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \right|_{\boldsymbol{\theta}=\boldsymbol{\theta}^{(0)}}.$$

Falls die analytische Berechnung des Gradienten $\nabla \mathcal{L}(\boldsymbol{\theta})$ zu schwierig ist, kann man ihn auch numerisch bestimmen. Dazu betrachtet man die Veränderung des Wertes der Log-Likelihoodfunktion für eine Veränderung eines Elementes aus dem Parametervektor $\boldsymbol{\theta}$ beim Startpunkt. Ein Element aus dem Gradienten erhält man folglich durch

$$\begin{aligned} \nabla \mathcal{L}_i(\boldsymbol{\theta}^{(0)}) &= \frac{1}{\Delta} \left(\mathcal{L}(\theta_1^{(0)}, \theta_2^{(0)}, \dots, \theta_{i-1}^{(0)}, \theta_i^{(0)} + \Delta, \theta_{i+1}^{(0)}, \theta_{i+2}^{(0)}, \dots, \theta_a^{(0)}) \right. \\ &\quad \left. - \mathcal{L}(\theta_1^{(0)}, \theta_2^{(0)}, \dots, \theta_{i-1}^{(0)}, \theta_i^{(0)}, \theta_{i+1}^{(0)}, \theta_{i+2}^{(0)}, \dots, \theta_a^{(0)}) \right), \end{aligned}$$

wobei a die Anzahl der Parameter und Δ eine beliebige kleine Zahl wie beispielsweise $\Delta = 10^{-6}$ ist.

Da man in der Regel nach dem ersten Schritt das Maximum der Log-Likelihoodfunktion

noch nicht erreicht hat, schreitet man vom neu berechneten Punkt aus weiter in Richtung des Gradienten an der Stelle des neu berechneten Punktes. Dadurch ergibt sich die Rekursion

$$\boldsymbol{\theta}^{(k+1)} = \boldsymbol{\theta}^{(k)} + \lambda \frac{\nabla \mathcal{L}(\boldsymbol{\theta}^k)}{|\nabla \mathcal{L}(\boldsymbol{\theta}^k)|},$$

welche abgebrochen wird, wenn das Maximum der Log-Likelihoodfunktion erreicht ist. Das ist der Fall wenn der Wert der Log-Likelihoodfunktion an der neu berechneten Stelle nicht mehr größer ist als der Wert der Log-Likelihoodfunktion an der alten Stelle.

Da die Schrittweite λ fest gewählt wurde, muss sie klein genug gewählt werden, um das Maximum der Log-Likelihoodfunktion mit hinreichender Genauigkeit zu bestimmen. Je größer die Schrittweite λ gewählt wird, desto größer kann die Abweichung beim letzten Schritt vom Maximum der Log-Likelihoodfunktion sein. Ist der Startpunkt $\boldsymbol{\theta}^{(0)}$ weit von Maximum entfernt, sind folglich viele Schritte zu berechnen.

Die Idee ist nun, die Gittersuche für das Verfahren des steilsten Anstieg zu adaptieren, um den Rechenaufwand zu verringern. Es liegt auf der Hand, dass die feste Vorgabe der Schrittweite λ nicht die beste Wahl ist. Eine bessere Wahl wäre eine Schrittweite, für die der Wert der Log-Likelihoodfunktion entlang der Richtung des Gradienten maximal wird. Man könnte beispielsweise $\lambda = 2^{-5}, 2^{-4}, \dots, 2^5$ wählen und an jeder sich dadurch ergebenden neuen Stelle den Wert der Log-Likelihoodfunktion berechnen. Findet man an keiner neuen Stelle einen größeren Wert der Log-Likelihoodfunktion, so kann man das Gitter noch verfeinern, um näher an das Maximum heran zu schreiten.

Dieses numerische Verfahren sucht ein Maximum von einem Startpunkt aus in Richtung des steilsten Anstiegs der zu maximierenden Funktion. Es kann vorkommen, dass dabei nicht das globale Maximum gefunden wird, sondern nur ein lokales Maximum. Will man sicherstellen, dass man das globale Maximum findet, kann man die Suche von unterschiedlichen Startpunkten aus beginnen.

5.2.3 NEWTON-RAPHSON

Beim Verfahren des steilsten Anstiegs wird die erste Ableitung der Log-Likelihoodfunktion genutzt. Leider kann dies in bestimmten Fällen zu einer hohen Anzahl von Iterationen führen. Beim NEWTON-RAPHSON Verfahren nutzt man zusätzlich die zweite Ableitung der Log-Likelihoodfunktion. Die erste Ableitung ist der Gradient der Log-Likelihoodfunktion $\nabla \mathcal{L}(\boldsymbol{\theta})$. Die zweite Ableitung ist die HESSE-Matrix der Funktion $\mathbf{H}(\mathcal{L}(\boldsymbol{\theta}))$. Die Ableitungen erhält man durch Differenzieren der Log-Likelihoodfunktion $\mathcal{L}(\boldsymbol{\theta})$ nach den Parametern $\boldsymbol{\theta}$:

$$\nabla \mathcal{L}(\boldsymbol{\theta}) = \frac{\partial \mathcal{L}(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}}, \quad \mathbf{H}(\mathcal{L}(\boldsymbol{\theta})) = \frac{\partial^2 \mathcal{L}(\boldsymbol{\theta})}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}'}$$

Voraussetzung für die Anwendbarkeit des Verfahrens ist zum Einen die Existenz der zweiten Ableitung. Zum Anderen muss die Log-Likelihoodfunktion konkav sein, so dass $-\mathbf{H}(\mathcal{L}(\boldsymbol{\theta}))$ für alle $\boldsymbol{\theta}$ positiv definit ist. Auch die HESSE-Matrix kann man numerisch berechnen. Dies geschieht analog zur numerischen Berechnung des Gradienten.

Bei dem NEWTON-RAPHSON Verfahren wird die Log-Likelihoodfunktion durch eine TAYLOR-Reihe zweiter Ordnung approximiert:

$$\mathcal{L}(\boldsymbol{\theta}) \cong \mathcal{L}(\boldsymbol{\theta}^{(0)}) + \left(\nabla \mathcal{L}(\boldsymbol{\theta}^{(0)})\right)' \left(\boldsymbol{\theta} - \boldsymbol{\theta}^{(0)}\right) + \frac{1}{2} \left(\boldsymbol{\theta} - \boldsymbol{\theta}^{(0)}\right)' \mathbf{H}(\mathcal{L}(\boldsymbol{\theta}^{(0)})) \left(\boldsymbol{\theta} - \boldsymbol{\theta}^{(0)}\right).$$

Das Maximum der Log-Likelihoodfunktion erhält man, indem nach den Parametern $\boldsymbol{\theta}$ differenziert und das Ergebnis gleich Null gesetzt wird:

$$\nabla \mathcal{L}(\boldsymbol{\theta}^{(0)}) + \mathbf{H}(\mathcal{L}(\boldsymbol{\theta}^{(0)})) \left(\boldsymbol{\theta} - \boldsymbol{\theta}^{(0)}\right) = 0.$$

Diese Gleichung kann in wenigen Schritten nach den gesuchten Parametern aufgelöst werden:

$$\boldsymbol{\theta} = \boldsymbol{\theta}^{(0)} - \left(\mathbf{H}(\mathcal{L}(\boldsymbol{\theta}^{(0)}))\right)^{-1} \nabla \mathcal{L}(\boldsymbol{\theta}^{(0)}).$$

Diese Formel kann man zur iterativen Berechnung des Maximums der Log-Likelihoodfunktion verwenden. Die Rekursionsformel lautet:

$$\boldsymbol{\theta}^{(k+1)} = \boldsymbol{\theta}^{(k)} - \left(\mathbf{H}(\mathcal{L}(\boldsymbol{\theta}^{(k)}))\right)^{-1} \nabla \mathcal{L}(\boldsymbol{\theta}^{(k)}).$$

Wie gut oder schlecht das NEWTON-RAPHSON Verfahren funktioniert, hängt von der Log-Likelihoodfunktion ab. Ist sie eine quadratische Funktion, erhält man das Maximum schon nach der ersten Iteration:

$$\hat{\boldsymbol{\theta}} = \boldsymbol{\theta}^{(1)} = \boldsymbol{\theta}^{(0)} - \left(\mathbf{H}(\mathcal{L}(\boldsymbol{\theta}^{(0)}))\right)^{-1} \nabla \mathcal{L}(\boldsymbol{\theta}^{(0)}).$$

Ist die Log-Likelihoodfunktion nicht konkav, so wird das NEWTON-RAPHSON Verfahren nur sehr schlecht konvergieren. Dagegen ist das Verfahren des steilsten Anstiegs ist vergleichsweise langsamer, aber robuster bei der Suche nach dem Maximum der Log-Likelihoodfunktion.

Die Gittersuche kann man analog zum Verfahren des steilsten Anstiegs auch auf das NEWTON-RAPHSON Verfahren anwenden. Dabei wird ein Skalar λ eingeführt, der die Schrittweite kontrolliert:

$$\boldsymbol{\theta}^{(k+1)} = \boldsymbol{\theta}^{(k)} - \lambda \left(\mathbf{H}(\mathcal{L}(\boldsymbol{\theta}^{(k)}))\right)^{-1} \nabla \mathcal{L}(\boldsymbol{\theta}^{(k)}).$$

Auch hier wird λ variiert, um neue Stellen zu berechnen. Die Stelle mit dem größten Wert der Log-Likelihoodfunktion wird dann ausgewählt.

Kapitel 6

Modellidentifikation mit Informationskriterien

Das farbige Rauschen, welches nach dem Abziehen des Trends von den Beobachtungen übrig bleibt, kann nach Kapitel 1 mit ARMA-Prozessen oder alternativ mit ARCH-Prozessen modelliert werden. Dieses Kapitel beschäftigt sich mit der Bestimmung der Ordnung dieser Prozesse mit Hilfe von Informationskriterien, wobei nur auf das Wesentliche eingegangen wird. Für genauere Betrachtungen der verschiedenen Informationskriterien sei auf [7] und [10], Kapitel 8.1 hingewiesen.

Liegt keine Information über die Ordnung eines $ARMA(p, q)$ oder $ARCH(m)$ Prozesses vor, der einer Zeitreihe zu Grunde liegt, so benötigt man ein Kriterium für die Wahl von p und q beziehungsweise m . Im Zusammenhang mit Ausreißern stellen Informationskriterien ein einfaches Mittel zur Bestimmung der korrekten Modellordnung dar. Dabei berechnet man für alle in Erwägung gezogenen Modelle für einen $ARMA(p, q)$ Prozess das Informationskriterium

$$IC(i) = \log(\hat{\sigma}_i^2) + (p_i + q_i) \frac{C}{T} \quad (6.1)$$

beziehungsweise für einen $ARCH(m)$ Prozess

$$IC(i) = \log(\hat{\sigma}_i^2) + (m_i) \frac{C}{T}, \quad (6.2)$$

wobei i die Nummer des Modells ist. Es besteht aus der logarithmierten Varianz der Residuen und einem Strafterm, der mit steigender Anzahl von Parametern größer wird. Die Überlegung dabei ist, dass die Varianz der Residuen bei der korrekten Wahl der Prozessordnung minimal wird und durch Hinzunehmen zusätzlicher Parameter nicht weiter sinkt. Da der Strafterm jedoch beim Hinzunehmen zusätzlicher Parameter größer wird, wird auch das Informationskriterium größer. Wird das Informationskriterium minimal für eine Vorgabe von p und q , so hat man die beste Modellordnung gefunden.

Schätzt man nicht nur den $ARMA(p, q)$ Prozess, sondern auch in der Zeitreihe enthaltene Ausreißer, so muss die Anzahl k der detektierten Ausreißer bei der Wahl der Prozessordnung berücksichtigt werden. Dem Informationskriterium für die Wahl von p und q wird daher ein zweiter Strafterm hinzugefügt:

$$IC(i) = \log(\hat{\sigma}_i^2) + (p_i + q_i) \frac{C_1}{T} + k_i \frac{C_2}{T}. \quad (6.3)$$

Es werden hier vier verschiedene Informationskriterien vorgestellt, welche sich in der Wahl von C_1 und C_2 unterscheiden.

$$\begin{aligned} AIC(p, q, k) &= \log(\hat{\sigma}_{p,q}^2) + (p + q + k) \frac{2}{T} \\ BIC(p, q, k) &= \log(\hat{\sigma}_{p,q}^2) + (p + q + k) \frac{\log T}{T} \\ HQC(p, q, k) &= \log(\hat{\sigma}_{p,q}^2) + (p + q + k) \frac{\log \log T}{T} \\ MixedC(p, q, k) &= \log(\hat{\sigma}_{p,q}^2) + (p + q) \frac{\log T}{T} + k \frac{1}{T} \end{aligned}$$

Dabei bezeichnen AIC das AKAIKE-Informationskriterium, BIC das BAYESSche Informationskriterium, HQC das HANNAN-QUINN-Kriterium und $MixedC$ eine Mischung des AKAIKE-Informationskriteriums und des BAYESSchen Informationskriterium.

Kapitel 7

Ausreißeranalyse in Zeitreihen

Wie in Kapitel 1 dargestellt, können Beobachtungen Ausreißer enthalten. Im Kontext von Zeitreihen sind dies diejenigen Beobachtungen, die nicht dem Prozess entsprechen, welcher der beobachteten Zeitreihe zu Grunde liegt. Solche Ausreißer können verschiedenste Ursachen haben. Das Schätzen der Prozessparameter liefert möglicherweise falsche Werte, wenn Ausreißer in der Zeitreihe enthalten sind. Dieses Kapitel stellt einige Ausreißertypen vor und beschreibt einen Algorithmus, mit dem die Ausreißer in der Zeitreihe aufgespürt und bestimmt werden können, wodurch ein korrektes Schätzen der Modellparameter möglich wird. Dieses Kapitel stützt sich im Wesentlichen auf [1].

7.1 Ausreißertypen

Man kann sich viele verschiedene Ausreißertypen vorstellen. Die grundlegende Idee ist, für Ausreißer ein Muster anzunehmen, welches an den Zeitpunkt des Eintretens des Ausreißers gesetzt wird. Dann bleibt lediglich die Größe ω des Ausreißers zu schätzen. Das Muster eines Ausreißers kann man durch den Quotienten zweier Lag-Polynome $\alpha(B)$ und $\beta(B)$ beschreiben. Das Anbringen an einen Zeitpunkt gelingt durch Multiplikation des Quotienten mit einer Indikatorfunktion $I_t(t_a)$, wobei $I_t(t_a) = 1$ für $t = t_a$ gilt und $I_t(t_a) = 0$ für $t \neq t_a$ ist. Notiert man die mit einem Ausreißer verunreinigte Zeitreihe mit Y_t^* und die Zeitreihe ohne Ausreißer mit Y_t , so erhält man den Zusammenhang

$$Y_t^* = Y_t + \omega \frac{\alpha(B)}{\beta(B)} I_t(t_a), \quad (7.1)$$

wobei Y_t durch den $ARMA(p,q)$ Prozess

$$Y_t = \frac{\theta(B)}{\phi(B)} \varepsilon_t$$

entstanden sei.

Zum Schätzen der Größe ω der Ausreißer ist es wichtig, den Zusammenhang zwischen den mit Ausreißern verunreinigten Residuen ε_t^* und den nicht verunreinigten Residuen ε_t zu kennen. Man erhält ihn durch Multiplikation von (7.1) mit $\phi(B)/\theta(B)$. Definiert man das Lag-Polynom $\pi(B)$ als

$$\pi(B) = \phi(B)/\theta(B) = 1 - \pi_1 B - \pi_2 B^2 - \pi_3 B^3 - \dots,$$

so folgt

$$\varepsilon_t^* = \varepsilon_t + \omega \pi(B) \frac{\alpha(B)}{\beta(B)} I_t(t_a). \quad (7.2)$$

Wenn die Wurzeln der Lag-Polynome $\phi(B)$ und $\theta(B)$ außerhalb des Einheitskreises liegen, gehen die Koeffizienten π_k für steigende k genügend schnell gegen Null, so dass man das unendlich lange Lag-Polynom $\pi(B)$ durch ein endlich langes Lag-Polynom $\pi(B)$ approximieren kann. Für eine Schätzung nach der Methode der kleinsten Quadrate für die Größe ω eines Ausreißers schreibt man (7.2) um zu $\varepsilon_t^* = \omega x_t + \varepsilon_t$. Die Schätzung lautet dann

$$\hat{\omega} = \frac{\sum_{t=t_a}^T \varepsilon_t^* x_t}{\sum_{t=t_a}^T x_t^2},$$

wobei T die Länge der Zeitreihe ist. Wie x_t aussieht, wird im Folgenden gezeigt. Weiterhin benötigt man einen Parametertest für die geschätzte Größe $\hat{\omega}$ eines Ausreißers, um den Ausreißer auf Signifikanz überprüfen zu können. Ist σ_ε^2 die Varianz der Zufallsvariablen ε_t , so kann man die Testgröße $\hat{\tau}(t_a)$ aufstellen.

$$\hat{\tau}(t_a) = \frac{\hat{\omega}}{\sigma_\varepsilon} \left(\sum_{t=t_a}^T x_t^2 \right)^{\frac{1}{2}} \quad (7.3)$$

Die Testgröße folgt einer Normalverteilung, so dass $\hat{\tau}(t_a) \sim N(0, 1)$ unter der Nullhypothese $H_0: \hat{\omega} = 0$ gilt. In der Regel ist nicht nur ein einzelner Ausreißer in einer Zeitreihe enthalten, sondern mehrere Ausreißer verschiedenen Typs. Wenn M die Anzahl aller Ausreißer ist, wandeln sich die Gleichungen (7.1) und (7.2) zu

$$Y_t^* = Y_t + \sum_{j=1}^M \omega_j \frac{\alpha_j(B)}{\beta_j(B)} I_t(t_a) \quad (7.4)$$

und

$$\varepsilon_t^* = \varepsilon_t + \sum_{j=1}^M \omega_j \pi(B) \frac{\alpha_j(B)}{\beta_j(B)} I_t(t_a). \quad (7.5)$$

7.1.1 Additiver Ausreißer

Der additive Ausreißer ist von der Idee her der einfachste Ausreißer. Es wird auf die Beobachtung, die den Ausreißer enthalten soll, ein Wert addiert. So ergibt sich für den Quotienten der Lag-Polynome $\alpha(B)/\beta(B) = 1$, woraus die Zusammenhänge

$$Y_t^* = Y_t + \omega_A I_t(t_a) \quad \text{und} \quad \varepsilon_t^* = \varepsilon_t + \omega_A \pi(B) I_t(t_a)$$

folgen. Man erhält $x_{A,t} = 0$ für $t < t_a$, $x_{A,t} = 1$ für $t = t_a$ und $x_{A,t_a+k} = -\pi_k$ für $k \geq 1$.

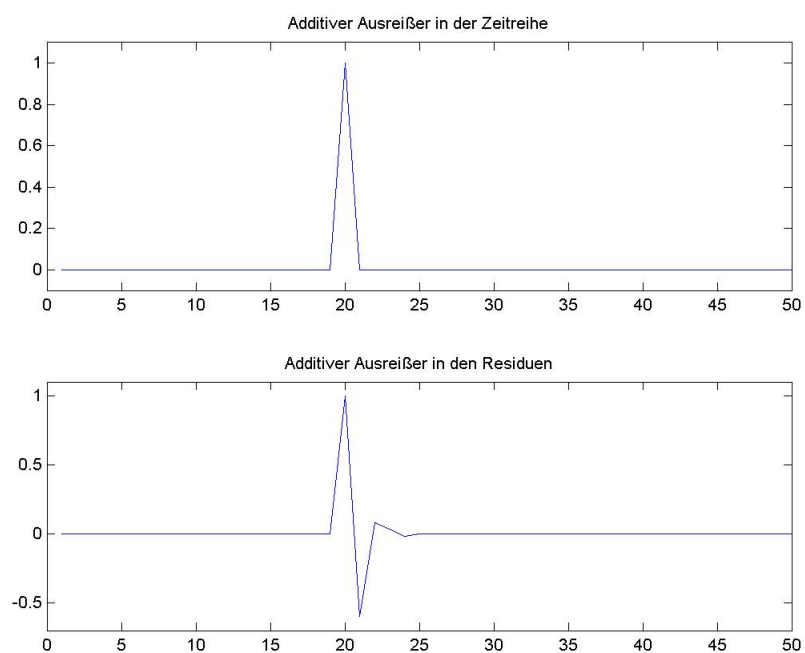


Abbildung 7.1: Dargestellt sind die Ausreißermuster eines additiven Ausreißers für den $AR-MA(1,2)$ Prozess $Y_t = \theta(B)/\phi(B)\varepsilon_t$ mit $\theta(B) = 1 + 0.3B + 0.1B^2$ und $\phi(B) = 1 - 0.3B$. Oben ist das Muster $\bar{x}_{A,t}$ des Ausreißers in der Zeitreihe und unten das Muster $x_{A,t}$ des Ausreißers in den Residuen zu sehen. Der Zeitpunkt des Eintretens des Ausreißers ist $t_a = 20$.

7.1.2 Level Shift

Der sogenannte Level Shift ist eine dauerhafte Änderung der Zeitreihe. Ab einem bestimmten Zeitpunkt wird auf jede Beobachtung ein konstanter Wert addiert. Mit dem Quotienten $\alpha(B)/\beta(B) = 1/(1-B)$ ergibt sich

$$Y_t^* = Y_t + \omega_L \frac{1}{1-B} I_t(t_a) \quad \text{und} \quad \varepsilon_t^* = \varepsilon_t + \omega_L \pi(B) \frac{1}{1-B} I_t(t_a),$$

wobei man durch Polynomdivision $1/(1-B) = 1 + B + B^2 + B^3 + \dots$ erhält. Man findet $x_{L,t} = 0$ für $t < t_a$, $x_{L,t} = 1$ für $t = t_a$ und $x_{L,t_a+k} = 1 - \sum_{j=1}^k \pi_j$ für $k \geq 1$.

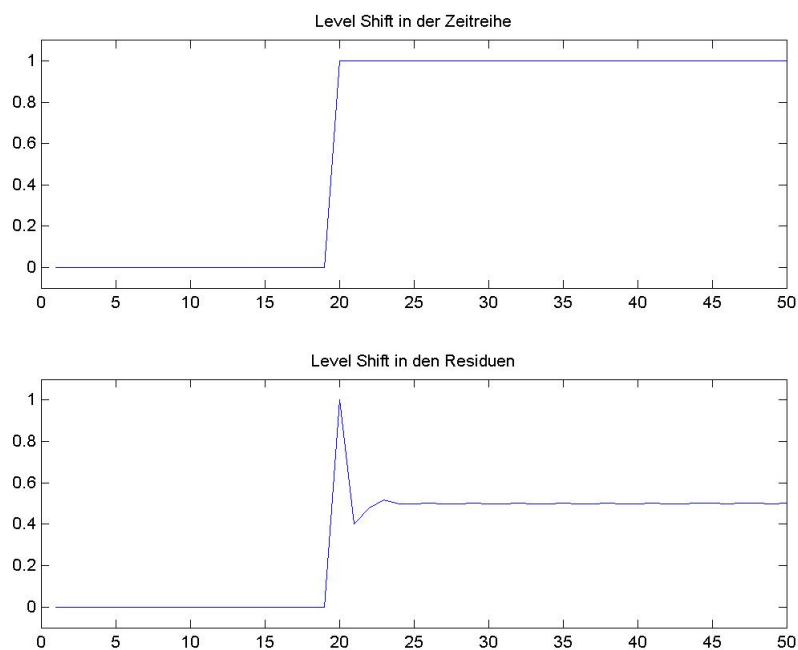


Abbildung 7.2: Dargestellt sind die Ausreißermuster eines Level Shifts für den $ARMA(1,2)$ Prozess $Y_t = \theta(B)/\phi(B)\varepsilon_t$ mit $\theta(B) = 1 + 0.3B + 0.1B^2$ und $\phi(B) = 1 - 0.3B$. Oben ist das Muster $\bar{x}_{L,t}$ des Ausreißers in der Zeitreihe und unten das Muster $x_{L,t}$ des Ausreißers in den Residuen zu sehen. Der Zeitpunkt des Eintretens des Ausreißers ist $t_a = 20$.

7.1.3 Zeitliche Änderung

Bei der zeitlichen Änderung handelt es sich um einen Effekt, der zu einem dem Ausreißer zugeordneten Zeitpunkt eintritt und dann wieder abklingt. Der Quotient lautet $\alpha(B)/\beta(B) = 1/(1-\delta B)$, wobei der Parameter δ zwischen Null und Eins liegt und die Länge des Abklingens steuert. Je größer δ ist, desto länger dauert das Abklingen. Es folgen die Zusammenhänge

$$Y_t^* = Y_t + \omega_Z \frac{1}{1-\delta B} I_t(t_a) \quad \text{und} \quad \varepsilon_t^* = \varepsilon_t + \omega_Z \pi(B) \frac{1}{1-\delta B} I_t(t_a).$$

Für $\delta = 0$ erhält man einen additiven Ausreißer und für $\delta = 1$ einen Level Shift. Man findet $1/(1-B) = 1 + \delta B + \delta^2 B^2 + \delta^3 B^3 + \dots$ durch Polynomdivision. Es ergibt sich $x_{Z,t} = 0$ für $t < t_a$, $x_{Z,t} = 1$ für $t = t_a$ und $x_{Z,t_a+k} = \delta^k - \pi_k - \sum_{j=1}^{k-1} \delta^{k-j} \pi_j$ für $k \geq 1$.

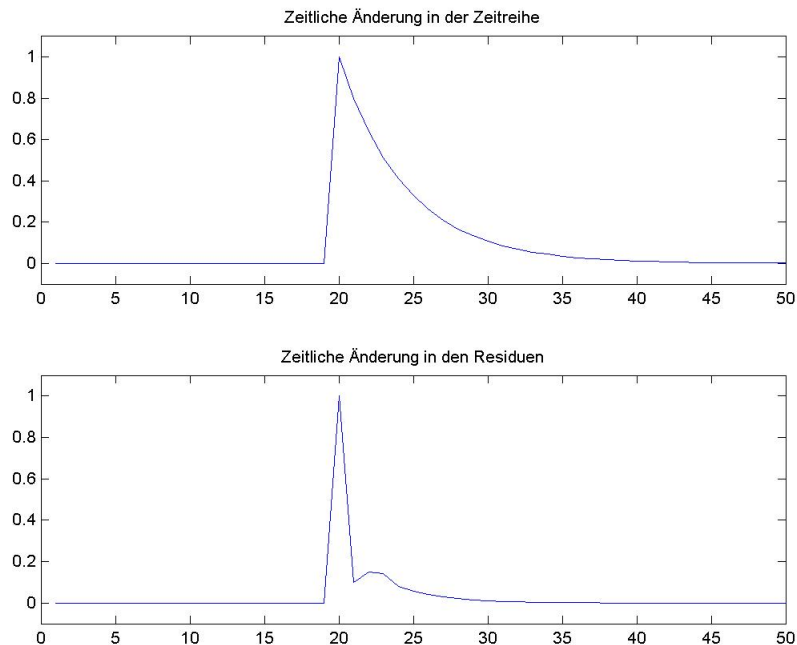


Abbildung 7.3: Dargestellt sind die Ausreißermuster einer zeitlichen Änderung für den $AR-MA(1,2)$ Prozess $Y_t = \theta(B)/\phi(B)\varepsilon_t$ mit $\theta(B) = 1 + 0.3B + 0.1B^2$ und $\phi(B) = 1 - 0.3B$. Oben ist das Muster $\bar{x}_{Z,t}$ des Ausreißers in der Zeitreihe und unten das Muster $x_{Z,t}$ des Ausreißers in den Residuen zu sehen. Der Zeitpunkt des Eintretens des Ausreißers ist $t_a = 20$.

7.1.4 Innovativer Ausreißer

Bei innovativen Ausreißern wird die Zeitreihe dadurch verfälscht, dass ein Ausreißer in den Prozess, welcher der beobachteten Zeitreihe zu Grunde liegt, einfließt. So entspricht der Quotient $\alpha(B)/\beta(B)$ dem Quotienten $\theta(B)/\phi(B)$ der Prozessgleichung, wodurch sich die Zusammenhänge

$$Y_t^* = Y_t + \omega_I \frac{\theta(B)}{\phi(B)} I_t(t_a) \quad \text{und} \quad \varepsilon_t^* = \varepsilon_t + \omega_I I_t(t_a).$$

ergeben. Man erhält $x_{I,t} = 0$ für $t < t_a$, $x_{I,t} = 1$ für $t = t_a$ und $x_{I,t_a+k} = 0$ für $k \geq 1$.

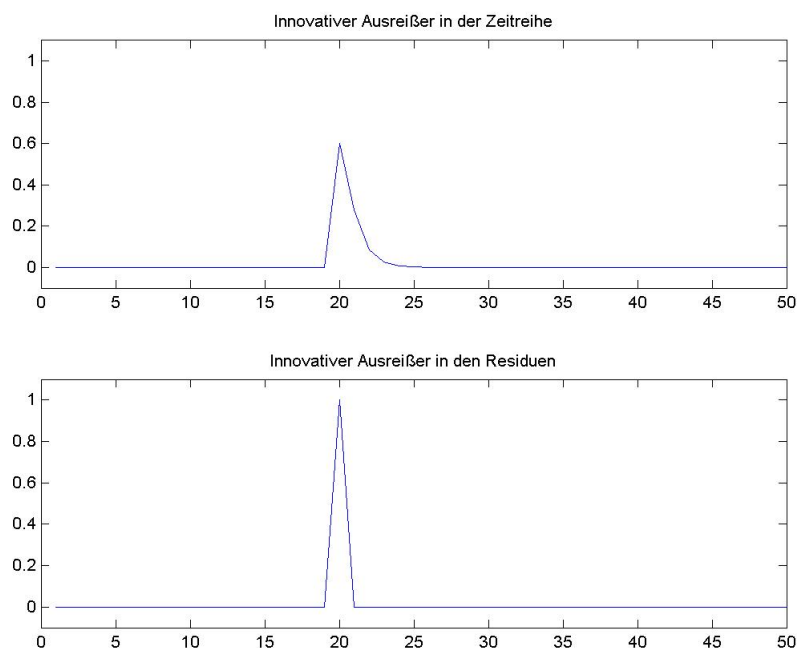


Abbildung 7.4: Dargestellt sind die Ausreißermuster eines innovativen Ausreißers für den $ARMA(1,2)$ Prozess $Y_t = \theta(B)/\phi(B)\varepsilon_t$ mit $\theta(B) = 1 + 0.3B + 0.1B^2$ und $\phi(B) = 1 - 0.3B$. Oben ist das Muster $\bar{x}_{I,t}$ des Ausreißers in der Zeitreihe und unten das Muster $x_{I,t}$ des Ausreißers in den Residuen zu sehen. Der Zeitpunkt des Eintretens des Ausreißers ist $t_a = 20$.

7.2 Algorithmus zur Ausreißersuche

In diesem Abschnitt wird ein Algorithmus zum gemeinsamen Schätzen der Prozessparameter und Ausreißer vorgestellt. Er wird dabei in der Form aufgeführt, wie ihn CHEN und LIU in [1] vorschlagen. Da die Ausreißermuster, wie im Abschnitt 7.1 dargestellt, zum Teil aus den Prozessparametern berechnet werden, kann die Schätzung der Prozessparameter und Ausreißer nicht gleichzeitig erfolgen. Der Algorithmus arbeitet daher bei der Schätzung der Prozessparameter und Ausreißer iterativ. Er gliedert sich in drei Stufen. Den prinzipiellen Ablauf des Algorithmus und die Verschachtelungstiefe kann man aus den Abbildungen 7.5 bis 7.7 entnehmen, welche die modifizierte Version des Algorithmus von CHEN und LIU darstellen.

Stufe I: Initiale Parameterschätzung und Ausreißerdetektion

- I.1. Bestimme die Maximum-Likelihood Schätzung für die Modellparameter. Dabei wird für die erste Iteration die ursprüngliche Zeitreihe verwendet und für alle weiteren Iterationen die angepasste Zeitreihe. Berechne mit den neuen Modellparametern die neuen Residuen ε_t .
- I.2. Berechne nach Gleichung (7.3) für $t = 1, \dots, T$ die Teststatistiken $\hat{\tau}_A(t)$, $\hat{\tau}_L(t)$, $\hat{\tau}_Z(t)$ und $\hat{\tau}_I(t)$, wobei die aus I.1 stammenden neuen Residuen verwendet werden. Wähle den signifikantesten Ausreißer durch $\eta = \max_t \{ |\hat{\tau}_A(t)|, |\hat{\tau}_L(t)|, |\hat{\tau}_Z(t)|, |\hat{\tau}_I(t)| \}$ aus. Falls $\eta = \hat{\tau}(t_a) > C$ gilt, liegt möglicherweise ein Ausreißer des entsprechenden Typs zum Zeitpunkt t_a vor. Die Konstante C ist dabei ein vorher festgesetzter kritischer Wert.
- I.3. Falls in Schritt I.2. kein Ausreißer mit $\eta = \hat{\tau}(t_a) > C$ gefunden werden konnte, fahre mit Schritt I.4 fort. Wurde jedoch ein Ausreißer gefunden, dann entferne seinen Effekt mit den Gleichungen (7.1) und (7.2) aus der Zeitreihe und den Residuen. Gehe anschließend wieder zu Schritt I.2 zurück.
- I.4. Wenn in den voran gehenden zwei Schritten in der ersten Iteration keine Ausreißer gefunden wurden, kann man annehmen, dass keine Ausreißer in der Zeitreihe vorhanden sind. Breche den Algorithmus dann an dieser Stelle ab. Sind jedoch in den voran gegangenen zwei Schritten Ausreißer detektiert worden, dann gehe zu Schritt I.1 zurück, um die Parameter mit der angepassten Zeitreihe neu zu schätzen. Erst wenn in der aktuellen Iteration in den Schritten I.2 und I.3 keine weiteren Ausreißer mehr gefunden werden, insgesamt jedoch die Anzahl der Ausreißer größer als Null ist, schreite zu Stufe II des Algorithmus fort.

Stufe II: Gemeinsames Schätzen der Ausreißer und Modellparameter

- II.1. Sei M die Anzahl der detektierten Ausreißer, die zu den Zeitpunkten t_1, t_2, \dots, t_M

vorliegen. Schätze unter Verwendung von Gleichung (7.5) die Größen $\omega_1, \omega_2, \dots, \omega_M$ der jeweiligen Ausreißer gemeinsam nach der Methode der kleinsten Quadrate.

- II.2. Berechne die Teststatistiken $\hat{\tau}_j = \hat{\omega}_j / \sigma_{\hat{\omega}_j}$, $j = 1, \dots, M$, um mit $\hat{\tau}_{\min} = \min_j \{|\hat{\tau}_j|\}$ den am wenigsten signifikanten Ausreißer zu identifizieren. Die Testgrößen folgen einer Normalverteilung $\hat{\tau}_j \sim N(0, 1)$ und die Nullhypothese lautet $H_0: \hat{\omega}_j = 0$. Falls $\hat{\tau}_{\min} \leq C$ gilt, lösche den entsprechenden Ausreißer und gehe zu Schritt II.1 zurück, wo die verbleibenden $M - 1$ Ausreißer gemeinsam geschätzt werden. Die Konstante C ist der gleiche kritische Wert wie in Stufe I. Wenn alle Ausreißer als signifikant eingestuft werden, fahre mit Schritt II.3 fort.
- II.3. Bestimme die angepasste Zeitreihe mit der neuesten Schätzung der Größen $\hat{\omega}_j$ der verbleibenden signifikanten Ausreißer unter Verwendung der ursprünglichen Zeitreihe und der neuesten Schätzung der Modellparameter. Die angepasste Zeitreihe enthält dann keine signifikanten Ausreißer mehr.
- II.4. Bestimme die Maximum-Likelihood Schätzung für die Modellparameter. Dabei wird die angepasste Zeitreihe aus Schritt II.3 verwendet. Ist die relative Änderung der Standardabweichung σ_ε der Residuen größer als eine vorher festgesetzte Konstante, gehe zu Schritt II.1 zurück, da die Änderung der Modellparameter die Teststatistiken in Schritt II.2 beeinflussen kann. Falls das nicht der Fall ist, gehe zu Stufe III über.

Stufe III: Auf der finalen Parameterschätzung basierende Detektion der Ausreißer

- III.1. Bestimme die Residuen ε_t mit den Modellparametern aus Schritt II.4.
- III.2. Iteriere noch einmal durch die Stufen I und II unter Verwendung der Residuen aus Schritt III.1. Dabei werden die beiden folgenden Modifikationen vorgenommen.
- Die Schätzung der Modellparameter wird nicht mehr geändert.
 - Schritt II.3 und Schritt II.4 werden weggelassen, so dass die neueste Schätzung der $\hat{\omega}_j$ in Schritt II.1 die endgültige Schätzung der Größe der Ausreißer ist.

7.3 Modifizierungen

Der Algorithmus zur Ausreißersuche arbeitet sehr rechenintensiv, was bei langen Zeitreihen zu Schwierigkeiten führt. Für lange Zeitreihen liefert eine Schätzung nach der Methode der kleinsten Quadrate gleichwertige Ergebnisse wie die Maximum-Likelihood Schätzung¹. Durch die Verwendung einer Schätzung nach der Methode der kleinsten Quadrate und weiteren Maßnahmen kann man die Rechenzeit stark reduzieren. Zunächst wird die Schätzung eines

¹Siehe [8], Kapitel 9.

$ARMA(p,q)$ Prozesses nach der Methode der kleinsten Quadrate vorgestellt. Dann wird der Algorithmus mit seinen Modifikationen dargestellt.

7.3.1 Schätzung eines $ARMA(p,q)$ Prozesses nach der Methode der kleinsten Quadrate

Es wird ein $ARMA(p,q)$ Prozess mit dem Mittelwert Null vorausgesetzt. Linearisiert man die Prozessgleichung (4.3) und setzt alle Größen mit negativen Indizes gleich ihrem Erwartungswert, welcher Null beträgt, dann erhält man folgendes Gleichungssystem.

$$\begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \\ \vdots \\ y_p \\ y_{p+1} \\ \vdots \\ y_T \end{pmatrix} = \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_3 \\ \varepsilon_4 \\ \vdots \\ \varepsilon_p \\ \varepsilon_{p+1} \\ \vdots \\ \varepsilon_T \end{pmatrix} + \underbrace{\begin{pmatrix} 0 & 0 & \dots & 0 & 0 & \dots & 0 \\ y_1 & 0 & \dots & 0 & \varepsilon_1 & \dots & 0 \\ y_2 & y_1 & \dots & 0 & \varepsilon_2 & \dots & 0 \\ y_3 & y_2 & \dots & 0 & \varepsilon_3 & \dots & 0 \\ \vdots & \vdots & & \vdots & \vdots & & \vdots \\ y_p & y_{p-1} & \dots & y_1 & \varepsilon_q & \dots & \varepsilon_1 \\ y_{p+1} & y_p & \dots & y_2 & \varepsilon_{q+1} & \dots & \varepsilon_2 \\ \vdots & \vdots & & \vdots & \vdots & & \vdots \\ y_{T-1} & y_{T-2} & \dots & y_{T-p} & \varepsilon_{T-1} & \dots & \varepsilon_{T-q} \end{pmatrix}}_{=\mathbf{A}_{ARMA}} \begin{pmatrix} \varphi_1 \\ \varphi_2 \\ \vdots \\ \varphi_p \\ \theta_1 \\ \vdots \\ \theta_q \end{pmatrix} \quad (7.6)$$

Darin sind die Residuen $\varepsilon_1, \varepsilon_2, \dots, \varepsilon_T$ zunächst unbekannt. Man bestimmt sie, indem man den Prozess der Zeitreihe mit einem langen $AR(r)$ Prozess approximiert. Durch die Verwendung der Prozessgleichung (4.1) erhält man folgendes Gleichungssystem. Alle Größen mit negativen Indizes sind gleich ihrem Erwartungswert gesetzt, welcher Null beträgt.

$$\begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ \vdots \\ y_r \\ y_{r+1} \\ \vdots \\ y_T \end{pmatrix} = \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_3 \\ \vdots \\ \varepsilon_r \\ \varepsilon_{r+1} \\ \vdots \\ \varepsilon_T \end{pmatrix} + \underbrace{\begin{pmatrix} 0 & 0 & \dots & 0 \\ y_1 & 0 & \dots & 0 \\ y_2 & y_1 & \dots & 0 \\ \vdots & \vdots & & \vdots \\ y_r & y_{r-1} & \dots & y_1 \\ y_{r+1} & y_r & \dots & y_2 \\ \vdots & \vdots & & \vdots \\ y_{T-1} & y_{T-2} & \dots & y_{T-r} \end{pmatrix}}_{=\mathbf{A}_{AR}} \begin{pmatrix} \varphi_1 \\ \varphi_2 \\ \vdots \\ \varphi_r \end{pmatrix} \quad (7.7)$$

Wegen der speziellen Struktur der Matrizen \mathbf{A}_{ARMA} und \mathbf{A}_{AR} ist es sinnvoll die Normalgleichungen $\mathbf{N}\mathbf{x} = \mathbf{n}$ direkt aufzustellen. Dies spart Speicherplatz und Rechenzeit. Für die Speicherung der beiden Matrizen benötigt man zusammen $(p+q+r)T$ Speicherplätze, was für sehr lange Zeitreihen speicherintensiv ist, insbesondere wegen des langen $AR(r)$ Prozesses. Die Rechenzeit wird dadurch verkürzt, dass bei den Matrizenprodukten $\mathbf{A}'_{ARMA}\mathbf{A}_{ARMA}$ und

$\mathbf{A}'_{AR}\mathbf{A}_{AR}$ zum Aufstellen der jeweiligen Normalgleichungen nicht alle Spalten der Matrizen \mathbf{A}_{ARMA} und \mathbf{A}_{AR} miteinander multipliziert werden müssen, wie nachfolgend erläutert wird.

Im Folgenden sei mit \mathbf{y}_t der Vektor $(y_1, y_2, \dots, y_t)'$ und mit $\boldsymbol{\varepsilon}_t$ der Vektor $(\varepsilon_1, \varepsilon_2, \dots, \varepsilon_t)'$ bezeichnet. Für die Schätzung des langen $AR(r)$ Prozesses muss man lediglich die erste Zeile der Normalgleichungen aufstellen.

$$\begin{pmatrix} N_{1,1} & N_{1,2} & \dots & N_{1,r} \end{pmatrix} = \begin{pmatrix} \mathbf{y}'_{T-1}\mathbf{y}_{T-1} & \mathbf{y}'_{T-1}\mathbf{y}_{T-2} & \dots & \mathbf{y}'_{T-1}\mathbf{y}_{T-r} \end{pmatrix}$$

Alle weiteren Elemente der rechten oberen Hälfte von \mathbf{N} lassen sich durch

$$N_{i+1,j+1} = N_{i,j} - y_{T-i}y_{T-j}$$

mit Hilfe der schon vorhandenen Elemente schnell berechnen. Der linke untere Teil der Matrix \mathbf{N} ist wegen $N_{i,j} = N_{j,i}$ symmetrisch zu ergänzen. Die rechte Seite \mathbf{n} der Normalgleichungen erhält man durch

$$\begin{pmatrix} n_1 & n_2 & \dots & n_r \end{pmatrix}' = \begin{pmatrix} \mathbf{y}'_T\mathbf{y}_{T-1} & \mathbf{y}'_T\mathbf{y}_{T-2} & \dots & \mathbf{y}'_T\mathbf{y}_{T-r} \end{pmatrix}'.$$

Für die Schätzung des $ARMA(p,q)$ Prozesses sind folgende Elemente der Matrix \mathbf{N} aufzustellen.

$$\begin{aligned} \begin{pmatrix} N_{1,1} & N_{1,2} & \dots & N_{1,p} \end{pmatrix} &= \begin{pmatrix} \mathbf{y}'_{T-1}\mathbf{y}_{T-1} & \mathbf{y}'_{T-1}\mathbf{y}_{T-2} & \dots & \mathbf{y}'_{T-1}\mathbf{y}_{T-p} \end{pmatrix} \\ \begin{pmatrix} N_{1,p+1} & N_{1,p+2} & \dots & N_{1,p+q} \end{pmatrix} &= \begin{pmatrix} \mathbf{y}'_{T-1}\boldsymbol{\varepsilon}_{T-1} & \mathbf{y}'_{T-1}\boldsymbol{\varepsilon}_{T-2} & \dots & \mathbf{y}'_{T-1}\boldsymbol{\varepsilon}_{T-p} \end{pmatrix} \\ \begin{pmatrix} N_{2,p+1} & N_{3,p+1} & \dots & N_{p,p+1} \end{pmatrix}' &= \begin{pmatrix} \mathbf{y}'_{T-2}\boldsymbol{\varepsilon}_{T-1} & \mathbf{y}'_{T-3}\boldsymbol{\varepsilon}_{T-1} & \dots & \mathbf{y}'_{T-p}\boldsymbol{\varepsilon}_{T-1} \end{pmatrix}' \\ \begin{pmatrix} N_{p+1,p+1} & N_{p+1,p+2} & \dots & N_{p+1,p+q} \end{pmatrix} &= \begin{pmatrix} \boldsymbol{\varepsilon}'_{T-1}\boldsymbol{\varepsilon}_{T-1} & \boldsymbol{\varepsilon}'_{T-1}\boldsymbol{\varepsilon}_{T-2} & \dots & \boldsymbol{\varepsilon}'_{T-1}\boldsymbol{\varepsilon}_{T-q} \end{pmatrix} \end{aligned}$$

Alle weiteren Elemente der rechten oberen Hälfte der Matrix \mathbf{N} lassen sich durch

$$\begin{aligned} N_{i+1,j+1} &= N_{i,j} - y_{T-i}y_{T-j}, & i &= 1, \dots, p-1, & j &= i, \dots, p-1 \\ N_{i+1,j+1} &= N_{i,j} - y_{T-i}\boldsymbol{\varepsilon}_{T-j}, & i &= 1, \dots, p-1, & j &= 1, \dots, q-1 \\ N_{i+1,j+1} &= N_{i,j} - \boldsymbol{\varepsilon}_{T-i}\boldsymbol{\varepsilon}_{T-j}, & i &= p+1, \dots, p+q-1, & j &= i, \dots, p+q-1 \end{aligned}$$

schnell berechnen. Der linke untere Teil der Matrix \mathbf{N} ist wegen $N_{i,j} = N_{j,i}$ symmetrisch zu ergänzen. Die rechte Seite \mathbf{n} der Normalgleichungen berechnet man mit

$$\begin{pmatrix} n_1 \\ \vdots \\ n_p \\ n_{p+1} \\ \vdots \\ n_{p+q} \end{pmatrix} = \begin{pmatrix} \mathbf{y}'_T\mathbf{y}_{T-1} \\ \dots \\ \mathbf{y}'_T\mathbf{y}_{T-p} \\ \mathbf{y}'_T\boldsymbol{\varepsilon}_{T-1} \\ \dots \\ \mathbf{y}'_T\boldsymbol{\varepsilon}_{T-p} \end{pmatrix}.$$

7.3.2 Update für die Schätzung eines $ARMA(p, q)$ Prozesses

Sind Ausreißer in der Zeitreihe vorhanden, so sind sie auch in den Matrizen \mathbf{A}_{ARMA} und \mathbf{A}_{AR} enthalten. Es reicht also nicht den Ausreißer aus der Zeitreihe zu entfernen. Man muss ihn auch aus den Matrizen \mathbf{A}_{ARMA} und \mathbf{A}_{AR} entfernen. Da die Normalgleichungen direkt aufgestellt werden, müssen sie folglich mit einem entsprechenden Update versehen werden. Die Residuen, die man durch den langen $AR(r)$ Prozess erhält, ändern sich nach dem Entfernen des Ausreißers. Das Update für einen Ausreißer mit dem Muster x_t in den Residuen zum Zeitpunkt t_a lautet für die Elemente der Normalgleichungen des langen $AR(r)$ Prozesses

$$N_{i,j} = N_{i,j} - \hat{\omega} \begin{pmatrix} \bar{x}_{t_a} & \bar{x}_{t_a+1} & \dots & \bar{x}_{t_a+\bar{l}_a-1} \end{pmatrix} \left(\begin{pmatrix} y_{t_a+j-i} \\ y_{t_a+j-i+1} \\ \vdots \\ y_{t_a+j-i+\bar{l}_a-1} \end{pmatrix} + \begin{pmatrix} y_{t_a+i-j} \\ y_{t_a+i-j+1} \\ \vdots \\ y_{t_a+i-j+\bar{l}_a-1} \end{pmatrix} \right) \\ + \hat{\omega}^2 \begin{pmatrix} \bar{x}_{t_a} & \bar{x}_{t_a+1} & \dots & \bar{x}_{t_a+\bar{l}_a-j+i-1} \end{pmatrix} \begin{pmatrix} \bar{x}_{t_a+j-i} \\ \bar{x}_{t_a+j-i+1} \\ \vdots \\ \bar{x}_{t_a+\bar{l}_a-1} \end{pmatrix},$$

wobei \bar{x}_t das Muster des Ausreißers in der Zeitreihe und \bar{l}_a die Länge des Musters des Ausreißers in der Zeitreihe bezeichnet. Die Normalgleichungen des $ARMA(p, q)$ Prozesses müssen nach der Berechnung der Residuen wie zuvor beschrieben aufgestellt werden. Es ist \bar{l}_a so zu wählen, dass das Muster des Ausreißers in der Zeitreihe beim Element $\bar{l}_a + 1$ unter einen Schwellwert fällt, welcher beispielsweise $0.001 \sigma_\varepsilon$ betragen kann.

7.3.3 Faltung eines Ausreißermusters mit den Residuen

Für die Berechnung der Teststatistiken $\hat{\tau}_A(t)$ und $\hat{\tau}_Z(t)$ mit $t = 1, 2, \dots, T$ ist es sinnvoll die Ausreißermuster mit den Residuen zu falten. Dazu wird zunächst die Reihenfolge der Elemente der Ausreißermuster umgekehrt und dann werden die Ausreißermuster und die Residuen mit Nullen verlängert, um eine zyklische Faltung im Frequenzraum zu ermöglichen. Sei $M = \max\{l_A, l_Z\}$, wobei l_A die Länge des Musters eines additiven Ausreißers und l_Z die Länge des Musters einer zeitlichen Änderung ist. Dann werden an die Residuen sowie an die Muster eines additiven Ausreißers und einer zeitlichen Änderung Nullen angehängt, so dass die Vektoren die Länge $T + M - 1$ besitzen. Dabei sind l_A und l_Z so zu wählen, dass das jeweilige Muster beim Element $l_A + 1$ und $l_Z + 1$ unter einen Schwellenwert fällt, welcher beispielsweise $0.001 \sigma_\varepsilon$ betragen kann. Hierbei wird vorausgesetzt, dass der $AR(p)$ Anteil

und der $MA(q)$ Anteil des $ARMA(p, q)$ Prozesses invertierbar sind.

$$\boldsymbol{\varepsilon}_T = \begin{pmatrix} \varepsilon_1 \\ \vdots \\ \varepsilon_T \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad \mathbf{x}_A = \begin{pmatrix} -\pi_{l_A} \\ \vdots \\ -\pi_2 \\ -\pi_1 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad \mathbf{x}_Z = \begin{pmatrix} \delta^{l_Z} - \sum_{j=1}^{l_Z-1} \delta^{l_Z-j} \pi_j - \pi_{l_Z} \\ \vdots \\ \delta^3 - \delta^2 \pi_1 - \delta \pi_2 - \pi_3 \\ \delta^2 - \delta \pi_1 - \pi_2 \\ \delta - \pi_1 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

Anschließend werden die Ausreißermuster und die Residuen mit einer FOURIER-Transformation \mathcal{F} in den Frequenzraum abgebildet, wo aus der zyklischen Faltung eine elementweise Multiplikation wird. Das Ergebnis wird dann mit der inversen FOURIER-Transformation \mathcal{F}^{-1} rücktransformiert.

$$\begin{aligned} \mathcal{C}_A &= \boldsymbol{\varepsilon}_T \circledast \mathbf{x}_A = \mathcal{F}^{-1}(\mathcal{F}(\boldsymbol{\varepsilon}_T) \circ \mathcal{F}(\mathbf{x}_A)) \\ \mathcal{C}_Z &= \boldsymbol{\varepsilon}_T \circledast \mathbf{x}_Z = \mathcal{F}^{-1}(\mathcal{F}(\boldsymbol{\varepsilon}_T) \circ \mathcal{F}(\mathbf{x}_Z)) \end{aligned}$$

Da die Reihenfolge der Elemente der Ausreißermuster vor der Faltung umgekehrt wurde, muss man noch bei \mathcal{C}_A die ersten $l_Z - 1$ Elemente und bei \mathcal{C}_Z die ersten $l_Z - 1$ Elemente abtrennen. Schließlich werden noch alle Elemente, die über die Länge T hinaus gehen, abgetrennt. Damit ist die obere Summe bei der Schätzung eines Ausreißers

$$\hat{\omega} = \frac{\sum_{t=t_a}^T \varepsilon_t x_t}{\sum_{t=t_a}^T x_t^2}$$

effizient für $t = 1, 2, \dots, T$ berechnet, wobei hier ε_t die mit Ausreißern verunreinigten Residuen bezeichnet.

7.3.4 Update für die Faltung

Findet man einen Ausreißer muss man nicht die Faltungen \mathcal{C}_A und \mathcal{C}_Z neu berechnen. Es genügt den durch den Ausreißer betroffenen Teil mit einem Update zu versehen. Die Muster werden mit Nullen ergänzt, um das Update für die Faltung als zyklische Faltung im Frequenzraum rechnen zu können. Für die Faltung \mathcal{C}_A werden vor und hinter dem Muster des gefundenen Ausreißers $l_A - 1$ Nullen angefügt. Dann muss man die Reihenfolge der Elemente des Musters eines additiven Ausreißers umkehren und $l_a + l_A - 2$ Nullen anhängen.

Entsprechendes gilt für die Faltung \mathcal{C}_Z .

$$\mathbf{x} = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ x_{t_a} \\ x_{t_a+1} \\ \vdots \\ x_{t_a+l_a-1} \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad \mathbf{x}_A = \begin{pmatrix} x_{A,t_a+l_A-1} \\ x_{A,t_a+l_A-2} \\ \vdots \\ x_{A,t_a} \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad \mathbf{x}_Z = \begin{pmatrix} x_{Z,t_a+l_Z-1} \\ x_{Z,t_a+l_Z-2} \\ \vdots \\ x_{Z,t_a} \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

Dabei wird das Muster des beim Zeitpunkt t_a gefundenen Ausreißers mit x_t bezeichnet und dessen Länge mit l_a . Anschließend wird die zyklische Faltung durch eine Multiplikation im Frequenzraum berechnet.

$$\begin{aligned} \mathbf{u}_A &= \mathbf{x} \otimes \mathbf{x}_A = \mathcal{F}^{-1}(\mathcal{F}(\mathbf{x}) \circ \mathcal{F}(\mathbf{x}_A)) \\ \mathbf{u}_Z &= \mathbf{x} \otimes \mathbf{x}_Z = \mathcal{F}^{-1}(\mathcal{F}(\mathbf{x}) \circ \mathcal{F}(\mathbf{x}_Z)) \end{aligned}$$

Für die Faltung \mathcal{C}_A werden die ersten $l_A - 1$ Elemente und entsprechend bei der Faltung \mathcal{C}_Z die ersten $l_Z - 1$ Elemente abgetrennt. Die Updates lauten dann

$$\begin{pmatrix} \mathcal{C}_{A,t_a-l_A+1} \\ \mathcal{C}_{A,t_a-l_A+2} \\ \vdots \\ \mathcal{C}_{A,t_a+l_a-1} \end{pmatrix} = \begin{pmatrix} \mathcal{C}_{A,t_a-l_A+1} \\ \mathcal{C}_{A,t_a-l_A+2} \\ \vdots \\ \mathcal{C}_{A,t_a+l_a-1} \end{pmatrix} - \hat{\omega} \mathbf{u}_A$$

und

$$\begin{pmatrix} \mathcal{C}_{Z,t_a-l_Z+1} \\ \mathcal{C}_{Z,t_a-l_Z+2} \\ \vdots \\ \mathcal{C}_{Z,t_a+l_z-1} \end{pmatrix} = \begin{pmatrix} \mathcal{C}_{Z,t_a-l_Z+1} \\ \mathcal{C}_{Z,t_a-l_Z+2} \\ \vdots \\ \mathcal{C}_{Z,t_a+l_z-1} \end{pmatrix} - \hat{\omega} \mathbf{u}_Z.$$

7.3.5 Gemeinsames Schätzen der Ausreißer

Auch hier ist es sinnvoll die Normalgleichungen direkt aufzustellen. Da sich viele Ausreißermuster der gefundenen Ausreißer nicht überschneiden werden, ist zu erwarten, dass viele Elemente der Normalgleichungsmatrix \mathbf{N} gleich Null sind.² Daher sollte beim Aufstellen der

²Sortiert man die Ausreißer nach ihren Zeitpunkten, so sollte aus der dünn besetzten Normalgleichungsmatrix eine bandartige Matrix werden, was bei deren Invertierung vorteilhaft ausgenutzt werden könnte. Diese Möglichkeit wurde jedoch nicht näher untersucht.

Normalgleichungen immer nur die von einem Ausreißer betroffene Zeitspanne betrachtet werden. Soll ein Ausreißer gelöscht werden, weil er sich als nicht signifikant herausstellt, so ist lediglich die entsprechende Zeile und Spalte in der Normalgleichungsmatrix \mathbf{N} und die entsprechende Zeile in der rechten Seite \mathbf{n} der Normalgleichungen zu löschen.

7.4 Modifizierter Algorithmus zur Ausreißersuche

Mit den in Abschnitt 7.3 beschriebenen Modifikationen lässt sich der in Abschnitt 7.2 dargestellte Algorithmus von CHEN und LIU umformulieren. Mit der inneren Schleife in Stufe I sind die Schritte I.2 und I.3 gemeint. Die Abbildungen 7.5 bis 7.7 sollen den Ablauf des Algorithmus und dessen Verschachtelungstiefe verdeutlichen.

Stufe I: Initiale Parameterschätzung und Ausreißerdetektion

- I.1. Bestimme die Schätzung nach der Methode der kleinsten Quadrate für die Modellparameter. Dabei wird für die erste Iteration die ursprüngliche Zeitreihe verwendet und für alle weiteren Iterationen die angepasste Zeitreihe sowie die mit den Updates versehenen Normalgleichungen des langen $AR(r)$ Prozesses. Berechne mit den neuen Modellparametern die neuen Residuen ε_t .
- I.2. Berechne nach Gleichung (7.3) für $t = 1, \dots, T$ die Teststatistiken $\hat{\tau}_A(t)$, $\hat{\tau}_L(t)$, $\hat{\tau}_Z(t)$ und $\hat{\tau}_I(t)$, wobei die aus I.1 stammenden neuen Residuen verwendet werden. Nach dem ersten Durchlauf in der inneren Schleife kann dabei auf die mit Updates versehenen Faltungen \mathcal{C}_A und \mathcal{C}_Z zurückgegriffen werden. Wähle den signifikantesten Ausreißer durch $\eta = \max_t \{ |\hat{\tau}_A(t)|, |\hat{\tau}_L(t)|, |\hat{\tau}_Z(t)|, |\hat{\tau}_I(t)| \}$ aus. Falls $\eta = \hat{\tau}(t_a) > C$ gilt, liegt möglicherweise ein Ausreißer des entsprechenden Typs zum Zeitpunkt t_a vor. Die Konstante C ist dabei ein vorher festgesetzter kritischer Wert.
- I.3. Falls in Schritt I.2. kein Ausreißer mit $\eta = \hat{\tau}(t_a) > C$ gefunden werden konnte, schreite mit Schritt I.4 fort. Wurde jedoch ein Ausreißer gefunden, dann entferne mit den Gleichungen (7.1) und (7.2) seinen Effekt aus der Zeitreihe und den Residuen. Außerdem sind die Faltungen \mathcal{C}_A und \mathcal{C}_Z , sowie die Normalgleichungen des langen $AR(r)$ Prozesses mit einem Update zu versehen. Gehe anschließend wieder zu Schritt I.2 zurück.
- I.4. Wenn in den voran gehenden zwei Schritten in der ersten Iteration keine Ausreißer gefunden wurden, kann man annehmen, dass keine Ausreißer in der Zeitreihe vorhanden sind. Breche den Algorithmus dann an dieser Stelle ab. Sind jedoch in den vorangegangenen zwei Schritten Ausreißer detektiert worden, dann gehe zu Schritt I.1 zurück, um die Parameter mit der angepassten Zeitreihe neu zu schätzen. Erst wenn in der aktuellen Iteration in den Schritten I.2 und I.3 keine weiteren Ausreißer mehr gefunden

werden, insgesamt jedoch die Anzahl der Ausreißer größer als Null ist, schreite zu Stufe II des Algorithmus fort.

Stufe II: Gemeinsames Schätzen der Ausreißer und Modellparameter

- II.1. Sei M die Anzahl der detektierten Ausreißer, die zu den Zeitpunkten t_1, t_2, \dots, t_M vorliegen. Schätze unter Verwendung von Gleichung (7.5) die Größen $\omega_1, \omega_2, \dots, \omega_M$ der jeweiligen Ausreißer gemeinsam nach der Methode der kleinsten Quadrate, wobei die Normalgleichungen direkt aufgestellt werden, falls sie noch nicht vorliegen.
- II.2. Berechne die Teststatistiken $\hat{\tau}_j = \hat{\omega}_j / \sigma_{\hat{\omega}_j}$, $j = 1, \dots, M$, um mit $\hat{\tau}_{\min} = \min_j \{|\hat{\tau}_j|\}$ den am wenigsten signifikanten Ausreißer zu identifizieren. Die Testgrößen folgen einer Normalverteilung $\hat{\tau}_j \sim N(0, 1)$ und die Nullhypothese lautet $H_0: \hat{\omega}_j = 0$. Falls $\hat{\tau}_{\min} \leq C$ gilt, lösche den entsprechenden Ausreißer durch Entfernen der zugehörigen Zeilen und Spalten in den Normalgleichungen und gehe zu Schritt II.1 zurück, wo die verbleibenden $M - 1$ Ausreißer gemeinsam geschätzt werden. Die Konstante C ist der gleiche kritische Wert wie in Stufe I. Wenn alle Ausreißer als signifikant eingestuft werden, fahre mit Schritt II.3 fort.
- II.3. Bestimme die angepasste Zeitreihe mit der neuesten Schätzung der Größen $\hat{\omega}_j$ der verbleibenden signifikanten Ausreißer unter Verwendung der ursprünglichen Zeitreihe und der neuesten Schätzung der Modellparameter. Die angepasste Zeitreihe enthält dann keine signifikanten Ausreißer mehr.
- II.4. Bestimme die Schätzung für die Modellparameter nach der Methode der kleinsten Quadrate. Dabei wird die angepasste Zeitreihe aus Schritt II.3 verwendet. Ist die relative Änderung der Standardabweichung σ_ε der Residuen größer als eine vorher festgesetzte Konstante, gehe zu Schritt II.1 zurück, da die Änderung der Modellparameter die Teststatistiken in Schritt II.2 beeinflussen kann. Falls das nicht der Fall ist, gehe zu Stufe III über.

Stufe III: Auf der finalen Parameterschätzung basierende Detektion der Ausreißer

- III.1. Bestimme die Residuen ε_t mit den Modellparametern aus Schritt II.4.
- III.2. Iteriere noch einmal durch die Stufen I und II unter Verwendung der Residuen aus Schritt III.1. Dabei werden die beiden folgenden Modifikationen vorgenommen.
- (a) Die Schätzung der Modellparameter wird nicht mehr geändert.
 - (b) Schritt II.3 und Schritt II.4 werden weggelassen, so dass die neueste Schätzung der $\hat{\omega}_j$ in Schritt II.1 die endgültige Schätzung der Größe der Ausreißer ist.

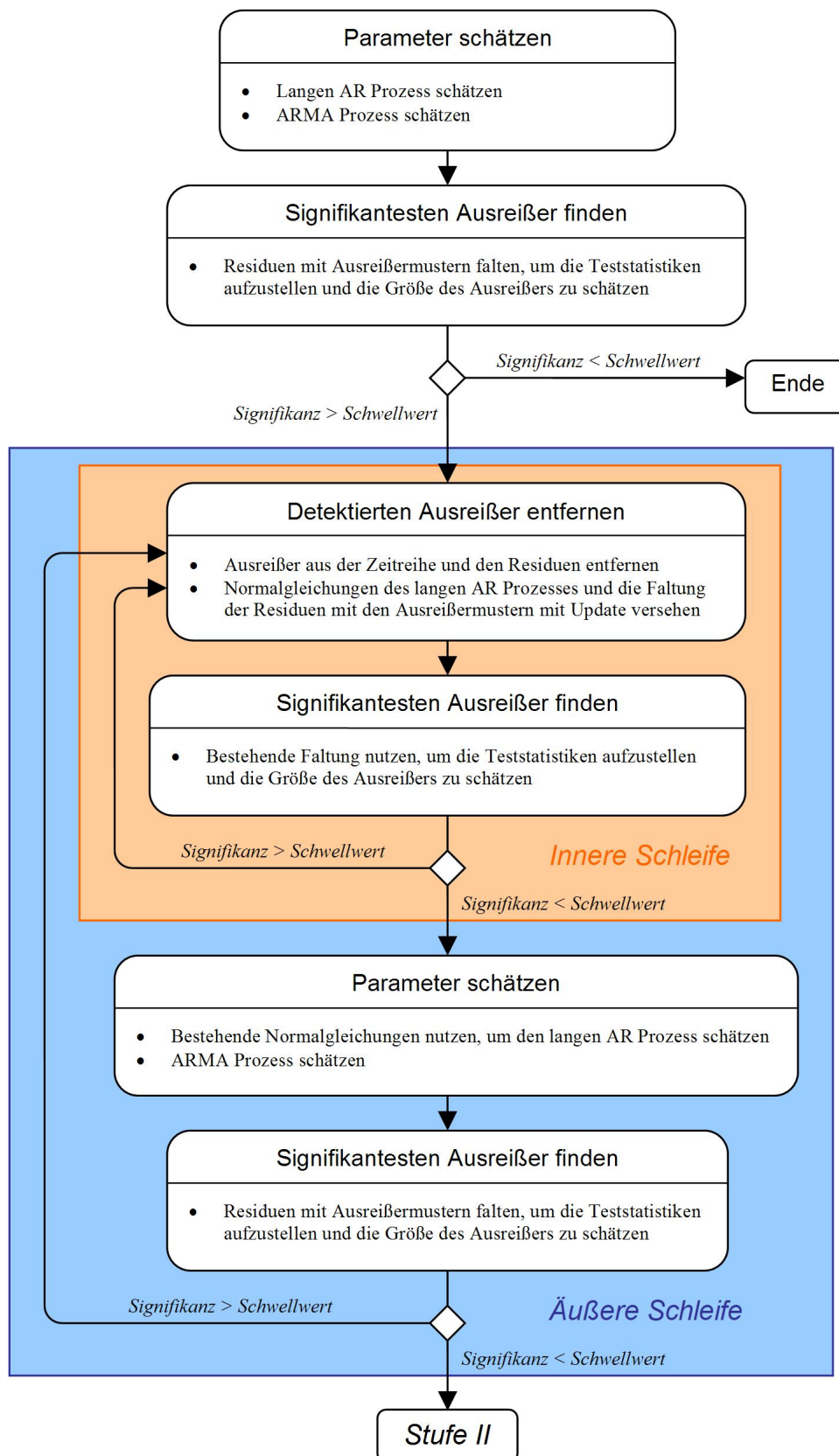


Abbildung 7.5: Stufe I des modifizierten Algorithmus von CHEN und LIU

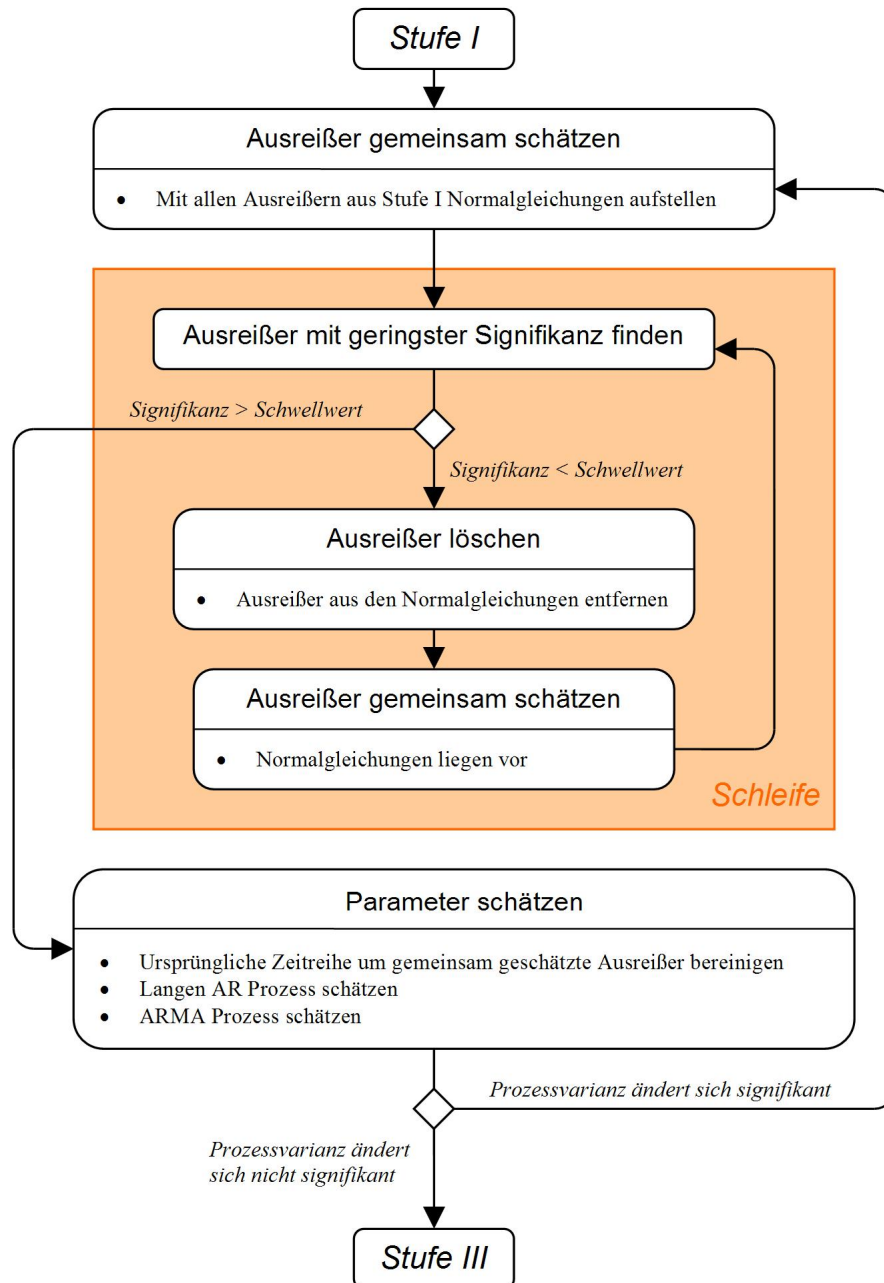


Abbildung 7.6: Stufe II des modifizierten Algorithmus von CHEN und LIU

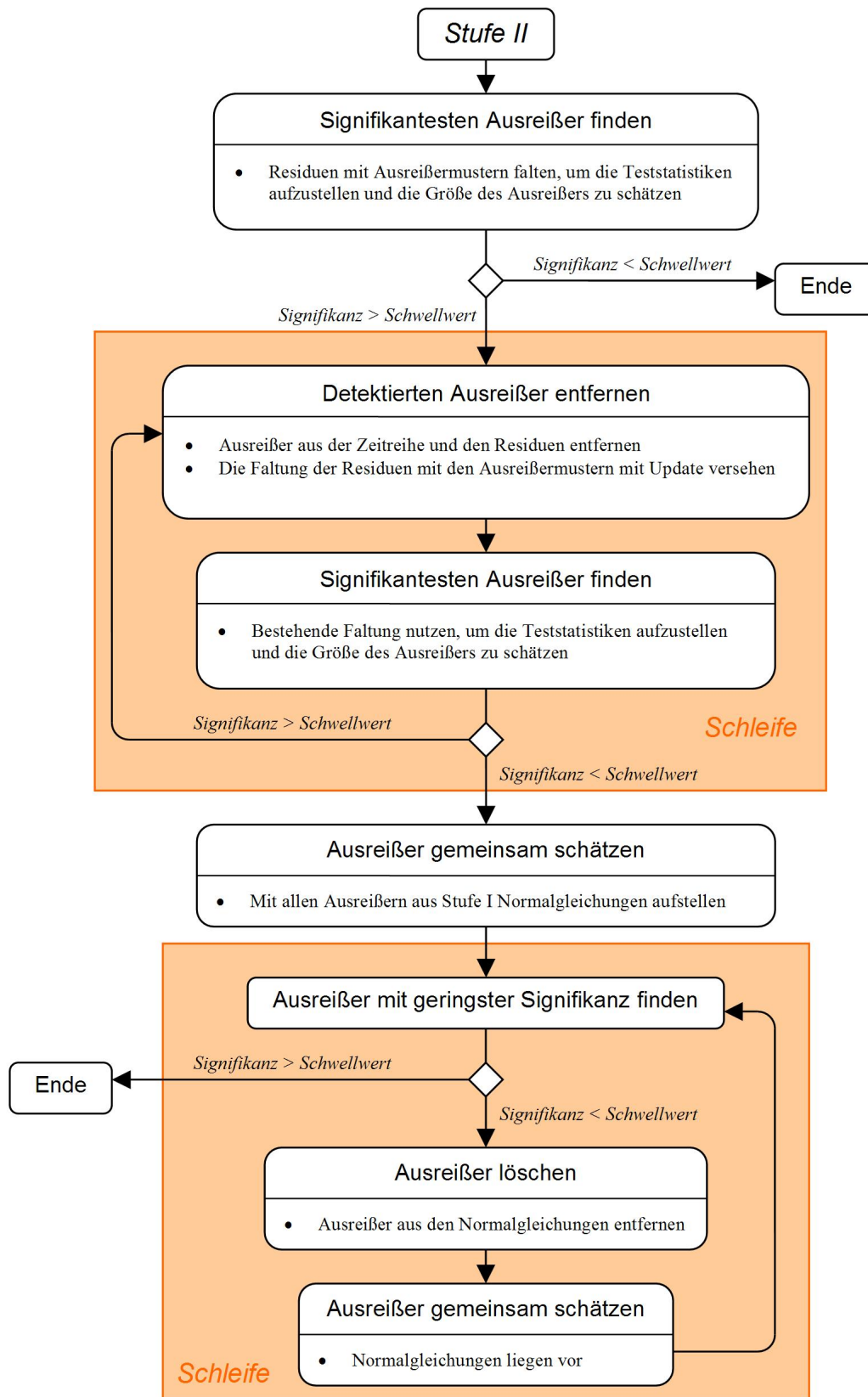


Abbildung 7.7: Stufe III des modifizierten Algorithmus von CHEN und LIU

Kapitel 8

Tests zur Ausreißeranalyse

Der in Abschnitt 7.2 dargestellte Algorithmus von CHEN und LIU und dessen in Abschnitt 7.4 aufgeführte modifizierte Version wurden mit Hilfe von MATLAB programmiert. Dieses Kapitel stellt verschiedene Tests dar, welche Aufschluss über die Qualität der Schätzung der Prozessparameter und Ausreißer geben sollen.

Zunächst wird ein Vergleich zwischen dem Algorithmus von CHEN und LIU und dessen modifizierter Version angestellt, um herauszustellen, ob die Algorithmen gleiche Ergebnisse liefern. Falls dies der Fall ist, kann der modifizierte Algorithmus von CHEN und LIU im Weiteren verwendet werden.

Anschließend werden die Fälle untersucht, bei denen die Zeitreihen einen, zwei oder viele Ausreißer enthalten. Ziel ist es, herauszufinden, welche Qualität die Schätzung der Prozessparameter und Ausreißer bei welchem Grad an Verunreinigung der Zeitreihen mit Ausreißern besitzt.

Abschließend wird die Qualität der Modellidentifikation für den Fall, dass Ausreißer in den Zeitreihen enthalten sind, untersucht. Dabei soll herausgestellt werden, ob die Schätzung von Ausreißern für die Modellidentifikation von Vorteil ist und welches Informationskriterium verwendet werden sollte.

8.1 Vorbemerkung zum Level-Shift

Beim Testen der programmierten MATLAB Funktionen stellte sich heraus, dass der Level-Shift ein schwierig zu schätzender Ausreißertyp ist. Die Ursache liegt in der Länge des Ausreißermusters. In der Zeitreihe und in den Residuen erstreckt sich das Muster eines Level-Shifts

Stufe I	
Startparameter	$\hat{\varphi}_1 = 0.00, \hat{\theta}_1 = 0.00, \hat{\sigma}_\varepsilon^2 = 1.0000$
Erste Parameterschätzung	$\hat{\varphi}_1 = 0.99, \hat{\theta}_1 = 0.04, \hat{\sigma}_\varepsilon^2 = 0.1253$
Detektion eines Ausreißers	L, $t_a = 500, \hat{\omega} = 2.86$
Erneute Parameterschätzung	$\hat{\varphi}_1 = 0.66, \hat{\theta}_1 = 0.24, \hat{\sigma}_\varepsilon^2 = 0.1004$
Detektion eines Ausreißers	A, $t_a = 343, \hat{\omega} = 0.82$
Erneute Parameterschätzung	$\hat{\varphi}_1 = 0.66, \hat{\theta}_1 = 0.25, \hat{\sigma}_\varepsilon^2 = 0.0992$
Stufe II	
Ausreißer	L, $t_a = 500, \hat{\omega} = 2.99$
Parameter	$\hat{\varphi}_1 = 0.65, \hat{\theta}_1 = 0.24, \hat{\sigma}_\varepsilon^2 = 0.1001$
Stufe III	
Ausreißer	L, $t_a = 500, \hat{\omega} = 2.99$

Tabelle 8.1: An Hand einer Zeitreihe, die dem Prozess nach Gleichung (8.1) folgt und die einen Level Shift der Größe $\omega = 3.0$ beim Zeitpunkt $t_a = 500$ enthält, werden mit dem Algorithmus von CHEN und LIU die Prozessparameter geschätzt. Es sind alle Zwischenergebnisse für die Schätzung der Prozessparameter, die der Algorithmus in den einzelnen Stufen liefert, dargestellt. Zusätzlich wird jede Detektion eines Ausreißers mit der Schätzung seiner Größe angegeben.

vom Zeitpunkt des Eintretens des Ausreißers bis zum Ende der Zeitreihe. Die Schätzung der Parameter wird daher durch einen Level-Shift stark verfälscht. Das hat zur Folge, dass auch das Muster des Ausreißers, welches mit den Parametern bestimmt wird, ebenfalls stark verfälscht wird. Folglich ist die Schätzung der Größe des Ausreißers nicht korrekt, wodurch der Ausreißer nicht korrekt aus der Zeitreihe entfernt werden kann. Einen einzelnen Level-Shift zu detektieren und korrekt zu schätzen gelingt mit dem Algorithmus. Sind mehrere Level-Shifts in der Zeitreihe vorhanden, schlägt deren Detektion und Schätzung jedoch oft fehl. Als Beispiel soll der $ARMA(1,1)$ Prozess

$$y_t = 0.6y_{t-1} + 0.3\varepsilon_{t-1} + \varepsilon_t, \quad \sigma_\varepsilon^2 = 0.1 \quad (8.1)$$

dienen. Es wird ein einzelner Level-Shift beim Zeitpunkt $t_a = 500$ hinzugefügt. Die Länge der Zeitreihe beträgt $T = 1000$. Als Schwellenwert für die Detektion wurde der Wert 3.5 gewählt, was einem Signifikanzniveau von 99.95 % entspricht. Tabelle 8.1 zeigt die Schätzung für die Parameter und den Ausreißer im Verlauf des Algorithmus. Es wurde die Maximum-Likelihoodmethode als Schätzverfahren verwendet. Die erste Parameterschätzung ist wegen des Ausreißers stark verfälscht. Es ist deutlich zu sehen, dass die Schätzung der Größe des Level-Shifts mit den ersten Parametern nicht korrekt ist. Erst nach der erneuten Parameterschätzung gelingt dies besser.

Sind mehrere Level-Shifts in der Zeitreihe enthalten, werden die Parameter nach der Detek-

tion und dem Entfernen eines Level-Shifts nicht wesentlich besser, so dass die Schätzung der Größe der weiteren Level-Shifts ebenfalls nicht gelingt. Ohne die korrekten Parameter ist man nicht in der Lage die Level-Shifts korrekt zu schätzen. Die korrekte Schätzung der Parameter setzt jedoch die korrekte Schätzung der Level-Shifts voraus.

Fazit: Level-Shifts sind bei der Schätzung der Parameter und der Größe der Ausreißer problematisch. Sie werden daher im Folgenden nicht mehr berücksichtigt.

8.2 Vergleich der Algorithmen zur Ausreißersuche

In diesem Abschnitt wird ein Vergleich des Algorithmus von CHEN und LIU mit dessen modifizierter Version angestellt. Zum Test sollen folgende Zeitreihen dienen, welche einmal mit und ein anderes Mal ohne Ausreißer versehen werden.

$$ARMA(1,1) \text{ Prozess: } y_t = 0.6y_{t-1} + 0.3\varepsilon_{t-1} + \varepsilon_t, \sigma_\varepsilon^2 = 2.0 \quad (8.2)$$

$$AR(2) \text{ Prozess: } y_t = 0.6y_{t-1} + 0.2y_{t-2} + \varepsilon_t, \sigma_\varepsilon^2 = 2.0 \quad (8.3)$$

$$MA(2) \text{ Prozess: } y_t = 0.6\varepsilon_{t-1} + 0.2\varepsilon_{t-2} + \varepsilon_t, \sigma_\varepsilon^2 = 2.0 \quad (8.4)$$

In den Abbildungen A.10 bis A.12 sind Ausschnitte aus den Zeitreihen dargestellt. Die Wurzeln der Lag-Polynome $\varphi(B)$ und $\theta(B)$ der Prozesse liegen alle weit genug außerhalb des Einheitskreises. Es sind von daher keine Probleme bei der Schätzung der Parameter und der Größe Ausreißer zu erwarten.

- Nullstellen des $ARMA(1,1)$ Prozesses: $B_1 \approx -3.3, B_2 \approx 1.7$
- Nullstellen des $AR(2)$ Prozesses: $B_1 \approx -4.2, B_2 \approx 1.2$
- Nullstellen des $MA(2)$ Prozesses: $B_1 \approx -1.5 - 1.7i, B_2 \approx -1.5 + 1.7i$

Der Schwellenwert für die Detektion von Ausreißern beträgt 3.5, was einem Signifikanzniveau von 99.95 % entspricht. Folglich kann man bei einer Länge der Zeitreihe von 10000 Werten ungefähr fünf detektierte Ausreißer erwarten, die nicht zu den hinzugefügten zählen. Die Modellordnung wird korrekt vorgegeben. Für die Qualität der Unterscheidung der einzelnen Ausreißertypen muss man die Ausreißermuster kennen. Die idealen Ausreißermuster für die drei angegebenen Prozesse sind in den Abbildungen A.1 bis A.9 dargestellt.

Zur Programmierung des modifizierten Algorithmus von CHEN und LIU ist Folgendes zu bemerken. Es werden bei der Schätzung nach der Methode der kleinsten Quadrate alle Beobachtungen verwendet. Alle Größen mit negativen Indizes in den Beobachtungsgleichungen werden gleich ihren Erwartungswerten gesetzt, welche Null betragen. Ausreißer werden in

allen Beobachtungen gesucht. Ragt ein Ausreißermuster über das Ende der Zeitreihe hinaus, wird es abgeschnitten. Im Gegensatz dazu ist der Algorithmus von CHEN und LIU so programmiert, dass die ersten Beobachtungen bei der Maximum-Likelihoodschätzung nicht genutzt werden, um Größen mit negativen Indizes zu vermeiden. Außerdem werden Ausreißer nicht bis an das Ende der Zeitreihe gesucht, damit die Ausreißermuster nicht über das Ende der Zeitreihe hinausragen.

Trotz dieser Abweichungen sind wegen der großen Länge der Zeitreihe ähnliche Ergebnisse für beide Algorithmen zu erwarten. Gelänge bei der Schätzung nach der Methode der kleinsten Quadrate die Approximation der Residuen perfekt und würde man bei der Maximum-Likelihood Schätzung alle Beobachtungen verwenden, sollten identische Ergebnisse für beide Algorithmen folgen, da dann die Minimierungsaufgabe der Schätzung nach der Methode der kleinsten Quadrate genau der Maximierungsaufgabe bei der Maximum-Likelihood Schätzung entspräche.

Betrachtet man die in den Tabellen A.1 bis A.6 dokumentierten Ergebnisse, so fällt auf, dass beide Algorithmen mehr Ausreißer finden als zuerst vermutet, insbesondere wenn Ausreißer der Zeitreihe hinzugefügt werden. Außerdem neigt der Algorithmus von CHEN und LIU dazu mehr Ausreißer zu detektieren als der modifizierte Algorithmus von CHEN und LIU. Die meisten detektierten Ausreißer werden jedoch von beiden Algorithmen gleichermaßen detektiert.

Den Ausreißertyp zu identifizieren scheint problematisch zu sein. Vergleicht man die idealen Ausreißermuster, wird deutlich, warum das so ist. Es wird nach den Ausreißermustern in den Residuen gesucht, so dass diese Muster zu vergleichen sind. Alle Muster sind einigermaßen ähnlich geformt und relativ kurz, so dass durch das Rauschen in der Zeitreihe leicht eine Verwechslung auftreten kann. Als deutlichstes Beispiel sei der Vergleich zwischen dem Muster eines additiven Ausreißers und dem Muster einer zeitlichen Änderung zu aufgeführt. Ist ein Ausreißer einmal falsch identifiziert, sehen beide Algorithmen keine Möglichkeit vor, diesen Fehler zu korrigieren. Die Größe der Ausreißer wird von beiden Algorithmen in etwa gleich geschätzt. Die hinzugefügten Ausreißer werden alle detektiert und deren Größe unter Berücksichtigung des Rauschens in der Zeitreihe auch korrekt geschätzt.

Die Schätzungen der Parameter sind bei beiden Algorithmen in etwa gleich. Die leichten Abweichungen lassen sich mit den Unterschieden in der Programmierung begründen. Auffällig ist jedoch die zu niedrig geschätzte Varianz der Residuen. Man kann hier einen Zusammenhang mit der zu großen Anzahl an detektierten Ausreißern vermuten. Durch das Ansetzen eines Ausreißers zum Zeitpunkt t_a wird das Residuum an diesem Zeitpunkt gegen Null gehen. Je mehr Ausreißer detektiert werden, desto mehr Residuen gehen gegen Null und desto niedriger ist die Varianz der Residuen. Da der Algorithmus von CHEN und LIU mehr Ausreißer findet

als der modifizierte Algorithmus von CHEN und LIU, ist auch die Varianz der Residuen beim Algorithmus von CHEN und LIU entsprechend niedriger als beim modifizierten Algorithmus von CHEN und LIU.

Fazit: Die beiden Algorithmen liefern gleichwertige Ergebnisse bei der Schätzung der Parameter und der Ausreißer, wenn man die Unterschiede bei der Programmierung berücksichtigt. Wegen der stark unterschiedlichen Rechenzeit ist jedoch der modifizierte Algorithmus von CHEN und LIU klar im Vorteil.

Die Identifikation des Typs der Ausreißer ist problematisch. Zum Einen sind sich die Muster ähnlich und zum Anderen sehen die Algorithmen keine Möglichkeit vor den Typ nachträglich zu wechseln. Geht man davon aus, dass wegen der Ausreißer die Parameter nicht ganz korrekt geschätzt werden, dann sind die Muster anfangs auch nicht ganz korrekt, da sie mit den Parametern berechnet werden. Auch daher ist eine Verwechslung der Typen denkbar.

8.3 Einzelner Ausreißer

Es stellt sich die Frage, mit welcher Qualität die Ausreißer detektiert werden. Um diese Frage beantworten zu können, werden die Zeitreihen

$$ARMA(1,1) \text{ Prozess: } y_t = 0.6y_{t-1} + 0.3\varepsilon_{t-1} + \varepsilon_t, \sigma_\varepsilon^2 = 1.0 \quad (8.5)$$

$$AR(2) \text{ Prozess: } y_t = 0.6y_{t-1} + 0.2y_{t-2} + \varepsilon_t, \sigma_\varepsilon^2 = 1.0 \quad (8.6)$$

$$MA(2) \text{ Prozess: } y_t = 0.6\varepsilon_{t-1} + 0.2\varepsilon_{t-2} + \varepsilon_t, \sigma_\varepsilon^2 = 1.0 \quad (8.7)$$

deren Länge $T = 1000$ beträgt, mit einem einzelnen Ausreißer zum Zeitpunkt $t_a = 500$ verunreinigt. Das Signifikanzniveau für die Detektion eines Ausreißers beträgt 99.95 %, wodurch sich ein kritischen Wert von 3.5 ergibt. Setzt man eine korrekte Parameterschätzung voraus, ergeben sich mit Gleichung (7.3) die in Tabelle 8.2 aufgeführten Mindestgrößen für die Detektion eines Ausreißers.

	Innovativer Ausreißer	Additiver Ausreißer	Zeitliche Änderung
<i>ARMA(1,1)</i> Prozess	$\hat{\omega}_I = 3.5000$	$\hat{\omega}_A = 2.5458$	$\hat{\omega}_Z = 3.4019$
<i>AR(2)</i> Prozess	$\hat{\omega}_I = 3.5000$	$\hat{\omega}_A = 2.9580$	$\hat{\omega}_Z = 3.4269$
<i>MA(2)</i> Prozess	$\hat{\omega}_I = 3.5000$	$\hat{\omega}_A = 2.9698$	$\hat{\omega}_Z = 3.3133$

Tabelle 8.2: Für den kritischen Wert $C = 3.5$ ergeben sich nach Gleichung (7.3) die aufgeführten Mindestgrößen $\hat{\omega}_I$, $\hat{\omega}_A$ und $\hat{\omega}_Z$ für die Detektion eines Ausreißers des entsprechenden Typs. Ist der Ausreißer kleiner, wird er nicht detektiert. Dabei sind die Ausreißermuster zu Grunde gelegt, die durch die Prozessgleichungen (8.5) bis (8.7) entstehen.

Die Wahrscheinlichkeit der Detektion eines Ausreißers und der Erwartungswert seiner Größe lassen sich wie folgt ableiten.

Die Verteilung der Größe eines Ausreißers entspricht einer Normalverteilung $N(\omega, \sigma_\omega^2)$, die wegen des kritischen Wertes abgeschnitten ist. Daher ist die Dichtefunktion der Verteilung nach einer Normierung zu $N(\omega, 1)$ zwischen -3.5 und 3.5 gleich Null. In diesem Bereich wird die Hypothese, dass ein Ausreißer vorliegt, abgelehnt. Folglich ist die Wahrscheinlichkeit einen Ausreißer der Größe ω zwischen $-3.5 \sigma_\omega$ und $3.5 \sigma_\omega$ zu detektieren gleich Null. Abbildung 8.1 soll dies verdeutlichen.

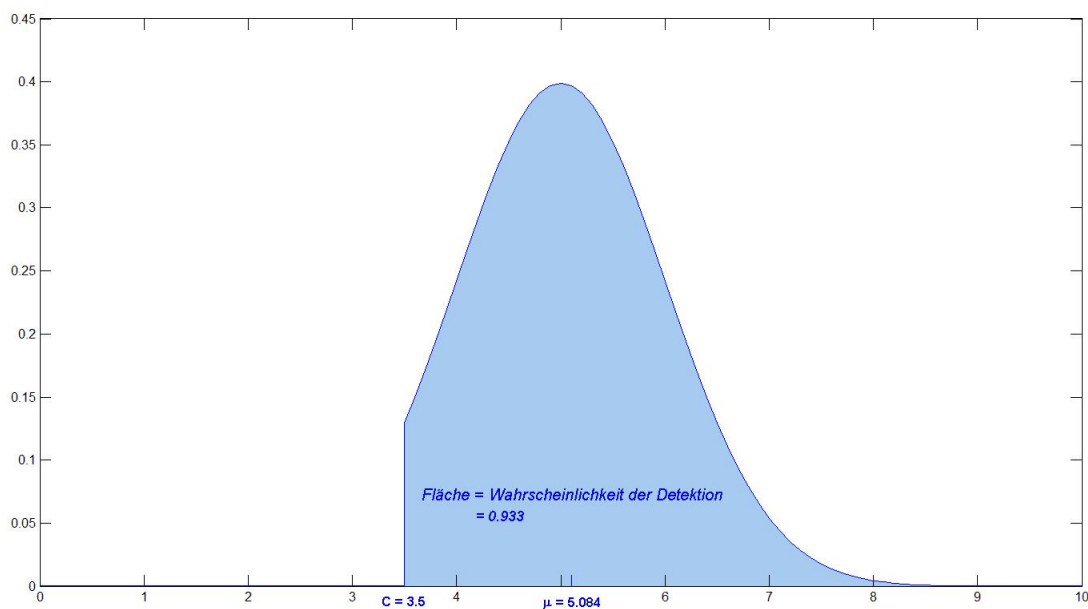


Abbildung 8.1: Dargestellt ist die abgeschnittene Dichtefunktion der Verteilung $N(5, 1)$, die zwischen -3.5 und 3.5 wegen des kritischen Wertes gleich Null ist. Man beachte, dass der Mittelwert nicht $\mu = 5.0$ sondern $\mu = 5.084$ ist.

Bei der Berechnung des Erwartungswertes μ der Größe eines detektierten Ausreißers muss man den Schwerpunkt der abgeschnittenen Dichtefunktion der Verteilung $N(\omega, \sigma_\omega^2)$ des Ausreißers finden. Dabei berechnet man σ_ω^2 mit den Ausreißermustern, die sich durch die Prozessgleichungen (8.5) bis (8.7) ergeben:

$$\sigma_\omega^2 = \left(\sum_{t=1}^T x_t^2 \right)^{-1}.$$

Zunächst bestimmt man die Teilflächen F_1 , F_2 , F_3 und F_4 unter der Dichtefunktion $f(\omega)$ der Verteilung $N(\omega, \sigma_\omega^2)$:

$$F_1 = \int_{-\infty}^{-3.5\sigma_\omega} f(\omega) d\omega, \quad F_2 = \int_{2\omega+3.5}^{\infty} f(\omega) d\omega,$$

$$F_3 = \int_{3.5\sigma_\omega}^{2\omega+3.5\sigma_\omega} f(\omega) d\omega, \quad F_4 = \int_{-\infty}^{3.5\sigma_\omega} f(\omega) d\omega.$$

Die gesamte Fläche unter der Dichtefunktion ist dann gleich $F_1 + F_2 + F_3$ und entspricht der Wahrscheinlichkeit einer Detektion des Ausreißers. Für die Prozessgleichungen (8.5) bis (8.7) ist diese Wahrscheinlichkeit in Abhängigkeit des Ausreißertyps sowie der Größe ω des Ausreißers in Tabelle 8.3 aufgeführt.

ω	ARMA(1,1)			AR(2)			MA(2)		
	I	A	Z	I	A	Z	I	A	Z
0	0.000	0.011	0.001	0.000	0.003	0.001	0.000	0.003	0.001
1	0.006	0.035	0.008	0.006	0.017	0.007	0.006	0.017	0.009
2	0.067	0.138	0.072	0.067	0.102	0.071	0.067	0.102	0.078
3	0.308	0.358	0.314	0.308	0.336	0.312	0.308	0.336	0.318
4	0.692	0.642	0.686	0.692	0.664	0.688	0.692	0.664	0.682
5	0.933	0.862	0.928	0.933	0.898	0.929	0.933	0.898	0.922
6	0.994	0.966	0.992	0.994	0.983	0.993	0.994	0.983	0.991
7	1.000	0.994	1.000	1.000	0.998	1.000	1.000	0.998	1.000

Tabelle 8.3: Dargestellt ist die Wahrscheinlichkeit der Detektion eines Ausreißers für die Prozesse nach den Gleichungen (8.5) bis (8.7) in Abhängigkeit der Größe ω des Ausreißers sowie des Ausreißertyps (I wie innovativer Ausreißer, A wie additiver Ausreißer und Z wie zeitliche Änderung). Sie ist kleiner Eins, da für ω zwischen $-3.5\sigma_\omega$ und $3.5\sigma_\omega$ die Hypothese, dass ein Ausreißer vorliegt, abgelehnt wird.

Der Schwerpunkt μ_{12} der Flächen F_1 und F_2 ist wegen deren Symmetrie gleich ω . Nun bestimmt man die Fläche $F = F_4 + 0.5 F_3$. Dann ermittelt den Wert der Umkehrfunktion F^{-1} der Verteilungsfunktion $F(\omega)$ für F und erhält somit den Schwerpunkt μ_3 der Fläche F_3 :

$$\mu_3 = F^{-1}(F).$$

Schließlich bildet man das gewichtete Mittel

$$\mu = \frac{F_1 + F_2}{F_1 + F_2 + F_3} \mu_{12} + \frac{F_3}{F_1 + F_2 + F_3} \mu_3,$$

um den Schwerpunkt μ der gesamten Fläche unter der Dichtefunktion $f(\omega)$ zu erhalten. Dieser Schwerpunkt entspricht dem Erwartungswert der Größe eines detektierten Ausreißers. Er

wird für die Prozessgleichungen (8.5) bis (8.7) in Abhängigkeit des Ausreißertyps und der Größe des Ausreißers in Tabelle 8.4 aufgeführt.

ω	ARMA(1,1)			AR(2)			MA(2)		
	I	A	Z	I	A	Z	I	A	Z
0	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
1	3.7346	4.6205	3.8022	3.7346	4.1654	3.7847	3.7346	4.1543	3.8668
2	3.8330	4.3340	3.8707	3.8330	4.0752	3.8609	3.8330	4.0690	3.9069
3	4.0183	4.1603	4.0290	4.0183	4.0870	4.0262	4.0183	4.0852	4.0392
4	4.3969	4.3133	4.3900	4.3969	4.3545	4.3918	4.3969	4.3556	4.3836
5	5.0838	5.0246	5.0770	5.0838	5.0476	5.0787	5.0838	5.0483	5.0709
6	6.0078	6.0004	6.0063	6.0078	6.0019	6.0067	6.0078	6.0020	6.0052
7	7.0003	7.0000	7.0002	7.0003	7.0000	7.0002	7.0003	7.0000	7.0001

Tabelle 8.4: Dargestellt ist der Erwartungswert der Größe eines detektierten Ausreißers für die Prozesse nach den Gleichungen (8.5) bis (8.7) in Abhängigkeit der Größe ω des Ausreißers sowie des Ausreißertyps (I wie innovativer Ausreißer, A wie additiver Ausreißer und Z wie zeitliche Änderung). Der Erwartungswert ist nicht gleich ω , was aus Abbildung 8.1 hervorgeht.

Den Tabellen 8.3 und 8.4 kann man entnehmen, dass ab einer Größe eines Ausreißers von 7.0 die Verteilungsfunktionen der Größe der Ausreißer kaum noch vom kritischen Wert beeinflusst werden. Da bei diesem Versuch der Stichprobenumfang nur 1000 Werte umfasst, ist es daher nicht sinnvoll über die Größe 7.0 hinauszugehen.

Betrachtet man Tabelle 8.2, so fällt auf, dass ein additiver Ausreißer schon früher als die anderen Ausreißertypen erkannt und als signifikant eingestuft wird. Der Grund dafür ist das Muster des Ausreißers in den Residuen. Die Summe $\sum_{t=t_a}^{t_a+l_a-1} x_t^2$ ist beim additiven Ausreißer am größten. Auch im Vergleich der Form der Muster hebt sich der additive Ausreißer heraus, wie in den Abbildungen A.1 bis A.9 zu sehen ist. Es ist also zu erwarten, dass der Typ eines additiven Ausreißer besser von den anderen unterschieden werden kann als ein innovativer Ausreißer von einer zeitlichen Änderung. Diese Erwartung wird durch den Versuch bestätigt. Betrachtet man beispielsweise die Tabellen A.7 bis A.9, so stellt man fest, dass der Typ eines innovativen Ausreißers bei einer Größe von $\omega_I = 7.0$ zu 77.5 % mit einer zeitlichen Änderung verwechselt wird, während der Typ eines additiven Ausreißers zu 98.9 % korrekt identifiziert wird. Es ist allerdings auch zu erkennen, dass durch die Verwechslung des Typs die Schätzung der Parameter nicht leidet, da zumindest ein ähnlicher Effekt aus der Zeitreihe entfernt wird, wie man in den Abbildungen A.2 und A.3 erkennen kann. Eine Ausnahme ist in Tabelle A.12 zu sehen. Die zeitliche Änderung wird praktisch nicht erkannt, da sie mit einem innovativen Ausreißer verwechselt wird. Die Ursache könnte in den durch den Ausreißer verfälschten Parametern liegen, aus denen die Muster der Ausreißer berechnet werden.

Will man die theoretische Wahrscheinlichkeit der Detektion eines Ausreißers bestätigt wissen, muss man die empirische Wahrscheinlichkeit der Detektion der drei Ausreißertypen aufsummieren, damit Verwechslungen des Typs der Ausreißer keinen Einfluss auf diese Statistik haben. In Anbetracht des Stichprobenumfangs passen die Ergebnisse für den Typ eines innovativen Ausreißers und einer zeitlichen Änderung sehr gut zu den Erwartungen. Additive Ausreißer werden dagegen deutlich häufiger detektiert, als zu erwarten ist.

Auch die Größe eines additiven Ausreißers wird im Grenzbereich der Detektion erkennbar besser geschätzt, als die Theorie vermuten lässt. Tabelle A.8 kann hierzu als Beispiel aufgeführt werden. Die Größe einer zeitlichen Änderung für den Fall des $AR(2)$ Prozesses ist wieder die Ausnahme. Sie wird deutlich zu klein geschätzt. Dies bekräftigt die Vermutung, dass das Muster der zeitlichen Änderung auf Grund der Parameter, die durch den Ausreißer verfälscht sind, nicht korrekt ist. Ansonsten passen die Ergebnisse gut zu den Erwartungen. Das gilt auch für die Varianz der geschätzten Größe der Ausreißer. Das einfachste Beispiel ist der innovative Ausreißer mit der Größe $\omega = 7.0$. Da die Verteilungsfunktion kaum mehr durch den kritischen Wert beeinflusst wird, sollte die Varianz der Größe des Ausreißers gleich der Varianz der Residuen sein, welche Eins beträgt.

Die Schätzung der Parameter gelingt in allen Fällen. Die geschätzten Parameter liegen alle sehr nah bei den Werten, mit denen die Zeitreihen erzeugt werden. Das bestätigt auch die Varianz der geschätzten Parameter. Dass mehr als die zu erwartenden 0.5 sonstigen Ausreißer pro Zeitreihe gefunden werden, kann mit den drei verwendeten Ausreißermustern begründet werden. Würde man nur ein Muster verwenden, sollten nur 0.5 sonstige Ausreißer pro Zeitreihe gefunden werden.

Fazit: Falls ein Ausreißer vorliegt, wird er auch detektiert. Problematisch ist die Identifikation des Ausreißertyps. Ursache dafür ist die Ähnlichkeit der Form der Ausreißermuster und die Berechnung der Ausreißermuster mit verfälschten Parametern. Sehr auffällig ist, dass die Identifikation einer zeitlichen Änderung bei der $AR(2)$ Zeitreihe völlig fehlschlägt.

8.4 Zwei Ausreißer

Falls zwei Ausreißer nahe beieinander liegen, stellt sich die Frage, ob die Detektion der Ausreißer, die Identifikation der Ausreißertypen und die Schätzung der Größe der Ausreißer davon beeinflusst werden. Um dies zu beantworten, muss man die Muster der Ausreißer in den Residuen betrachten. Da der Algorithmus von CHEN und LIU immer nur nach einem einzelnen Ausreißer sucht, diesen entfernt und dann weiter sucht, ist also die Frage, ob man zwei sich in den Residuen überschneidende Ausreißermuster noch auseinander halten kann oder ob sie

als ein Muster erkannt werden. Abbildung 8.2 illustriert dieses Problem. Für den $ARMA(1,1)$ Prozess nach Gleichung (8.8) werden die Ausreißermuster eines additiven Ausreißers und einer zeitlichen Änderung in den Residuen berechnet und zeitlich versetzt aufeinander aufaddiert. Einzeln sind sie in den Abbildungen A.1 und A.3 dargestellt. In dem Fall, dass zwischen den Ausreißern ein Zeitunterschied von $\Delta t = 7$ liegt, sind die Ausreißermuster noch gut zu unterscheiden. Liegt jedoch nur ein Zeitunterschied von $\Delta t = 1$ zwischen den Ausreißern, so verschmelzen die Ausreißermuster in der Art, dass das sich dadurch ergebende Muster am ehesten dem Muster eines innovativen Ausreißers ähnelt, wie in Abbildung A.2 zu dargestellt.

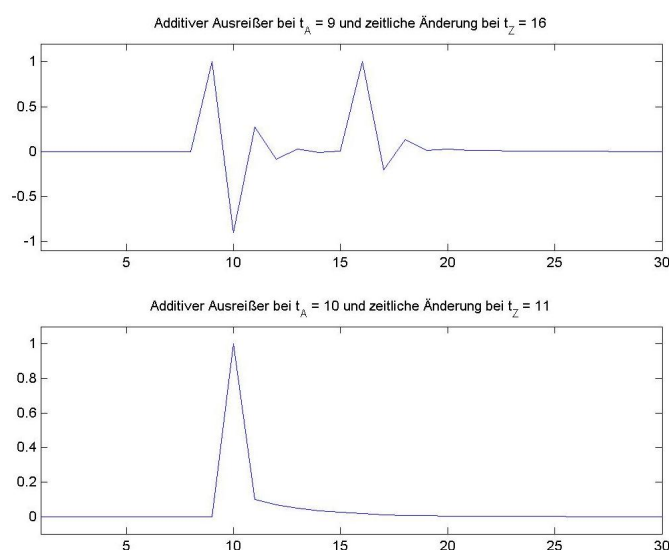


Abbildung 8.2: Oben sind die sich durch die Prozessgleichung (8.8) ergebenden Muster eines additiven Ausreißers x_A beim Zeitpunkt $t_a = 9$ und einer zeitlichen Änderung x_Z beim Zeitpunkt $t_a = 16$ dargestellt. Unten sind die selben Muster dargestellt, mit dem Unterschied dass der additive Ausreißers x_A beim Zeitpunkt $t_a = 10$ und die zeitliche Änderung x_Z beim Zeitpunkt $t_a = 11$ eintritt. Wegen der Überlagerung sind sie dem Muster eines innovativen Ausreißers x_I ähnlich.

Die Ausreißermuster werden aus den Parametern der $ARMA(p,q)$ Prozesse berechnet und sehen daher für unterschiedliche Prozesse unterschiedlich aus. Es kommt erschwerend hinzu, dass die geschätzten Parameter von den Ausreißern verfälscht sind und folglich die Ausreißermuster nicht korrekt bestimmt werden. Die oben gestellte Frage ist also nur schwer und nicht allgemeingültig zu beantworten. Es sollen die drei Prozesse

$$ARMA(1,1) \text{ Prozess: } y_t = 0.6y_{t-1} + 0.3\varepsilon_{t-1} + \varepsilon_t, \quad \sigma_\varepsilon^2 = 1.0 \quad (8.8)$$

$$AR(2) \text{ Prozess: } y_t = 0.6y_{t-1} + 0.2y_{t-2} + \varepsilon_t, \quad \sigma_\varepsilon^2 = 1.0 \quad (8.9)$$

$$MA(2) \text{ Prozess: } y_t = 0.6\varepsilon_{t-1} + 0.2\varepsilon_{t-2} + \varepsilon_t, \quad \sigma_\varepsilon^2 = 1.0 \quad (8.10)$$

als Beispiele dienen. Die Länge der damit erzeugten Zeitreihen beträgt 1000 Werte. Der erste Ausreißer wird beim Zeitpunkt $t_{a_1} = 500$ eingebaut. In den verschiedenen Durchläufen wird der zweite Ausreißer jeweils bei den Zeitpunkten $t_{a_2} = 501, \dots, 505, 510$ eingefügt. Die Größe der Ausreißer beträgt in allen Fällen $\omega = 7.0$, damit die Parameter möglichst wenig verfälscht werden und eine Detektion dennoch sichergestellt wird. Es werden alle möglichen Kombinationen von zwei Ausreißertypen getestet, welche in den ersten beiden Spalten angegeben sind. Für jede Kombination wurden 1000 Stichproben ausgewertet.

Im Folgenden soll nicht auf jede aus den Tabellen ablesbare Erkenntnis eingegangen werden, da diese für andere Prozesse zum Teil nicht gelten. Die Effekte resultieren aus der Überlagerung der verschiedenen Ausreißermuster und sind somit individuell für jeden Prozess. Das in Abbildung 8.2 gezeigte Beispiel kann man in den Tabellen A.16 und A.17 wiederfinden. Sinnvoller ist es, allgemeine Überlegungen anzustellen, die auf alle Fälle anwendbar sind und mit denen man Vorhersagen über die Qualität machen kann, die die Detektion und Identifikation der Ausreißer, sowie die Schätzung der Größe der Ausreißer betreffen.

Der Schätzung der Ausreißer geht immer eine Schätzung der Prozessparameter voraus, aus der die Ausreißermuster abgeleitet werden. Bevor man die Ausreißer sucht, sind also deren Muster bekannt und können untersucht werden. Wichtig sind auch die Wurzeln des Lag-Polynoms $\theta(B)$ des $ARMA(p, q)$ Prozesses. Es wird vorausgesetzt, dass die Wurzeln außerhalb des Einheitskreises liegen, damit sich die Ausreißermuster in den Residuen nur über einen begrenzten Zeitraum erstrecken. Je näher die Nullstellen am Einheitskreis liegen, desto länger werden die Muster und damit der benötigte Zeitunterschied zwischen den Ausreißern, um sie auseinander halten zu können. Als Beispiel sei der $ARMA(1, 1)$ Prozess

$$y_t = 0.6y_t + 0.8\varepsilon_{t-1} + \varepsilon_t \quad (8.11)$$

aufgeführt, bei dem die Wurzel des Lag-Polynoms $\theta(B) = 1 + 0.8B$ gleich -1.25 ist. Die Ausreißermuster des additiven Ausreißers und der zeitlichen Änderung sind in Abbildung 8.3 dargestellt.

Es fällt auf, dass diese Ausreißermuster deutlich länger andauern als die Ausreißermuster in den Abbildungen A.1 und A.3, die mit dem zuvor genannten $ARMA(1, 1)$ Prozess berechnet wurden. Eine einfache Formel zur Abschätzung der Länge der Muster herzuleiten, fällt schwer, da zuerst das Lag-Polynom $\pi(B) = \varphi(B)/\theta(B)$ berechnet werden muss, um anschließend mit dessen Koeffizienten π_1, π_2, \dots die Ausreißermuster aufzustellen. Als eine sehr grobe Abschätzung könnte man eine Formel wie

$$l_a = p - \frac{\log \epsilon}{\log |B_{0, \min}|} \quad (8.12)$$

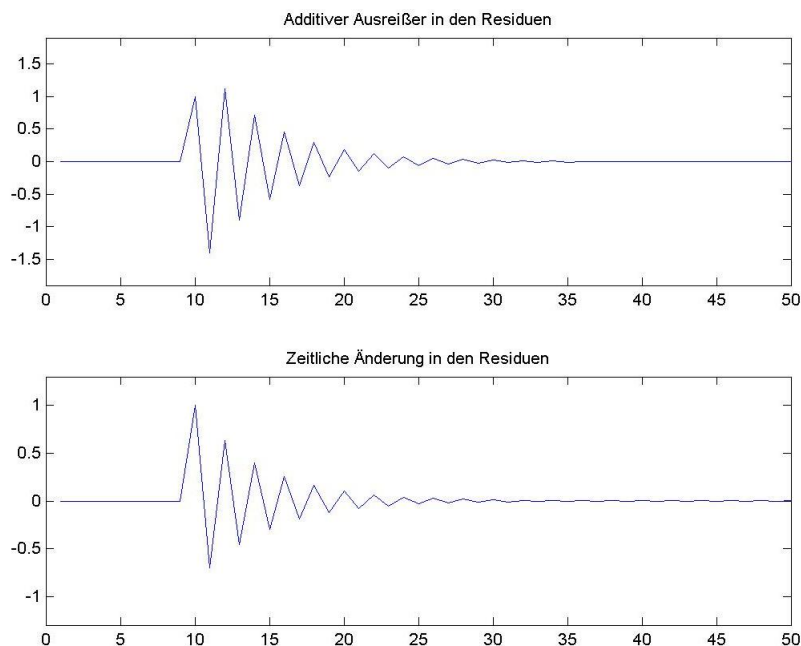


Abbildung 8.3: Dargestellt sind die sich durch die Prozessgleichung (8.11) ergebenden Muster eines additiven Ausreißers x_A und einer zeitlichen Änderung x_Z . Sie sind deutlich länger, als die sich durch die Prozessgleichungen (8.8) bis (8.10) ergebenden Ausreißermuster, welche in den Abbildungen A.1 bis A.9 dargestellt sind.

verwenden¹. Dabei ist l_a die Länge des Ausreißermusters, $B_{0,min}$ die betragsmäßig kleinste Nullstelle des Lag-Polynoms $\theta(B)$ und ϵ eine kleine Zahl, unter die der Wert des Musters beim Zeitpunkt $t_a + l_a$ fallen soll, so dass alle Werte des Musters vom Zeitpunkt t_a bis $t_a + l_a - 1$ größer als ϵ sind. Für die in Abbildung 8.3 dargestellten Muster ergibt sich nach Formel (8.12), dass sie nach dem Zeitpunkt $t_a + 22$ nur noch Werte unter $\epsilon = 0.01$ besitzen. Tatsächlich haben die Muster nach den Zeitpunkten $t_a + 24$ für einen additiven Ausreißer und $t_a + 21$ für eine zeitliche Änderung nur noch Werte unter $\epsilon = 0.01$.

Für die drei Prozesse, die hier als Beispiele dienen, ergeben sich nach der Formel (8.12) die Längen $l_{a,ARMA} = 10$, $l_{a,AR} = 2$ und $l_{a,MA} = 7$ für $\epsilon = 0.01$. Für diesen Wert von ϵ betragen die Längen der Ausreißermuster tatsächlich $l_{A,ARMA} = 5$, $l_{Z,ARMA} = 4$, $l_{A,AR} = 3$, $l_{Z,AR} = 10$, $l_{A,MA} = 6$ und $l_{Z,MA} = 11$. Daran erkennt man, dass es sich nur um eine grobe Abschätzung handelt. Der einfachste und genaueste Weg ist, die Ausreißermuster mit den geschätzten Parametern zu berechnen und dabei deren Längen in den Residuen zu speichern. Sie geben Auskunft über die Zeitdifferenz zwischen zwei Ausreißern, die benötigt wird, um eine eindeutige Identifikation zu erhalten. Falls zwei Ausreißer näher beieinander liegen, wird

¹Die Formel (8.12) wird mit der Betrachtung eines $MA(1)$ Prozesses und eines $AR(p)$ Prozesses hergeleitet. Für den $MA(1)$ Prozess ergibt sich $B_{0,min} = -1/\theta_1$ und es folgt $|\pi_k| = 1/|B_{0,min}|^k$. Der letzte Wert des Musters eines additiven Ausreißers ist $-\pi_k$, weshalb die Beziehung $\epsilon = 1/|B_{0,min}|^k$ nach k aufgelöst werden muss. Für den $AR(p)$ Prozess ergeben sich wegen $\pi(B) = \varphi(B)$ die Werte $\pi_p = \varphi_p$ und $\pi_{p+1} = 0$.

der Algorithmus von CHEN und LIU den Ausreißertyp auswählen, dessen Muster am ehesten zu den überlagerten Mustern passt.

Da der für die Identifikation entscheidende Teil des Ausreißermusters im vorderen Bereich des gesamten Musters liegt, ist die benötigte Zeitdifferenz zwischen den Ausreißern geringer als die Länge der Muster. Das liegt daran, dass für die drei Beispiele die ersten zwei bis drei Werte der Ausreißermuster wesentlich größer als die nachfolgenden Werte sind. Die Summe $\sum_{t=t_a}^T e_t x_t$ in der Berechnung der Testgröße $\hat{\tau}(t_a)$ nach Gleichung (7.3) wird demnach hauptsächlich durch die ersten Werte bestimmt, wie man auch mit Hilfe der Abbildungen A.1 bis A.9 nachvollziehen kann.

Den folgenden Tabellen kann man entnehmen, dass die benötigte Zeitdifferenz zur Identifikation der beiden Ausreißertypen für den $ARMA(1,1)$ Prozess bei $\Delta t = 5$, für den $AR(2)$ Prozess zwischen $\Delta t = 5$ und $\Delta t = 10$ und für den $MA(2)$ Prozess zwischen $\Delta t = 5$ und $\Delta t = 10$ liegt. Die Schätzung der Größe der Ausreißer gelingt schon bei etwas geringeren Zeitdifferenzen. Für den $ARMA(1,1)$ Prozess folgt $\Delta t = 3$, für den $AR(2)$ Prozess $5 < \Delta t < 10$ und für den $MA(2)$ Prozess $5 < \Delta t < 10$. Dabei ist zu beachten, dass beim $AR(2)$ Prozess die zeitliche Änderung immer mit einem innovativen Ausreißer verwechselt wird, was schon zuvor nachgewiesen wurde.

Fazit: Will man mehr über die Qualität der Identifikation und Schätzung der Größe der Ausreißer wissen, muss man die Ausreißermuster betrachten. Dabei ist besonders auf die Länge und den Anfang der Muster zu achten. Die Ausreißermuster sind für die verschiedenen Prozesse unterschiedlich, so dass man keine allgemeinen Regeln finden kann, welche Zeitdifferenz zwischen den Ausreißern ausreicht, um beide korrekt zu schätzen. Die Detektion der Ausreißer leidet nur dann, wenn sich die Ausreißermuster so überlagern, dass man sie nicht mehr auseinander halten kann, wie in Abbildung 8.2 unten zu sehen ist. Dabei ist zu beachten, dass ein neues Muster entsteht, dessen Effekt möglicherweise aus der Zeitreihe entfernt wird, so dass die Schätzung der Parameter des Prozesses wahrscheinlich nicht von den Ausreißern beeinflusst wird.

8.5 Viele Ausreißer

An dieser Stelle soll untersucht werden, wieviele Ausreißer in der Zeitreihe enthalten sein können, ohne dass die Schätzung der Parameter des Prozesses und der Ausreißer darunter leidet. Dazu wird ein Prozentsatz der Werte der Zeitreihe mit Ausreißern versehen. Die Ausreißer werden gleichmäßig über die ganze Zeitreihe verteilt. Beträgt der Prozentsatz beispielsweise 1% und ist die Länge der Zeitreihe $T = 1000$, dann ist die Zeitdifferenz zwischen zwei Ausreißern genau $\Delta t = 100$. Der Ausreißertyp wird zufällig ausgewählt. Die Größe der

Ausreißer ist ebenfalls ein zufälliger Wert, wobei eine Mindestgröße vorgegeben wird. Die Verteilungsfunktion der Größe $\omega = 5 \pm 1$ der Ausreißer ist in Abbildung 8.4 dargestellt. Es handelt sich um eine in der Mitte auseinander geschnittene GAUSSsche Glockenkurve.

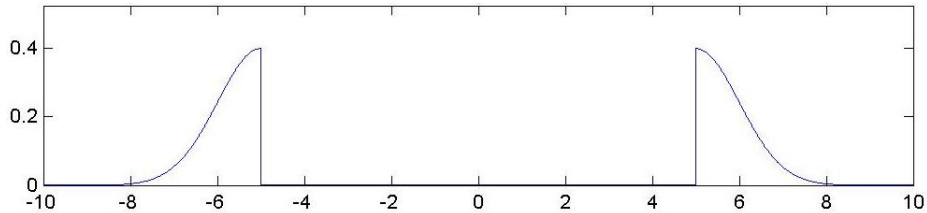


Abbildung 8.4: Dargestellt ist die Verteilungsfunktion der Größe der Ausreißer, die der Zeitreihe hinzugefügt werden. Es ist eine in der Mitte auseinander geschnittene Normalverteilung $N(0,1)$, was bewirkt, dass der Betrag der Größe des Ausreißers nicht kleiner als fünf sein kann.

Als Beispiele dienen wieder die drei Prozesse aus dem vorigen Kapitel, so dass sich wieder die Ausreißermuster ergeben, die in den Abbildungen A.1 bis A.9 zu sehen sind.

$$ARMA(1,1) \text{ Prozess: } y_t = 0.6y_{t-1} + 0.3\varepsilon_{t-1} + \varepsilon_t, \quad \sigma_\varepsilon^2 = 1.0 \quad (8.13)$$

$$AR(2) \text{ Prozess: } y_t = 0.6y_{t-1} + 0.2y_{t-2} + \varepsilon_t, \quad \sigma_\varepsilon^2 = 1.0 \quad (8.14)$$

$$MA(2) \text{ Prozess: } y_t = 0.6\varepsilon_{t-1} + 0.2\varepsilon_{t-2} + \varepsilon_t, \quad \sigma_\varepsilon^2 = 1.0 \quad (8.15)$$

Mit den Prozessen wurden für jede Wahl des Prozentsatzes und der Größe der Ausreißer 1000 Zeitreihen erzeugt. Von den geschätzten Parametern des Prozesses wurde jeweils der Mittelwert und dessen Standardabweichung bestimmt. Außerdem wurde die durchschnittliche Anzahl der detektierten Ausreißer bestimmt.

Betrachtet man die Tabellen A.25 bis A.33, so erkennt man, dass die Qualität der Parameterschätzung kontinuierlich schlechter wird, je mehr Ausreißer in der Zeitreihe enthalten sind. Dabei gibt es zwei Grenzen, die von Interesse sind. Bis zur ersten Grenze wird die Parameterschätzung nicht von Ausreißern beeinflusst. In diesem Fall werden alle Ausreißer detektiert und korrekt geschätzt. Die zweite Grenze ist der Punkt, bei dem die Detektion der Ausreißer nicht mehr gelingt und die Parameterschätzung folglich stark durch die Ausreißer verfälscht wird. Wichtigster Anhaltspunkt zum Finden der Grenzen ist die Varianz σ_ε^2 der Residuen. Der Freiheitsgrad der Varianz einer Stichprobe beträgt für alle drei Prozesse $f_1 = 998$, da die Zeitreihe 1000 Werte besitzt und jeweils zwei Prozessparameter zu schätzen sind. Der Mittelwert der Varianz wird aus 1000 Stichproben ermittelt, so dass hier ein FISHER-Test

$$\mu_{\sigma_\varepsilon^2} \sim F_{998 \times 1000 - 1, \infty}$$

durchgeführt werden kann. Bei einem Signifikanzniveau von 99 % beträgt der kritische Wert 1.0033 für die Hypothese $H_0: \mu_{\sigma_\varepsilon^2} = 1.0$. Daraus ergeben sich die in Tabelle 8.5 aufgeführten

Grenzen, die nicht überschritten werden dürfen, um eine korrekte Schätzung der Parameter des Prozesses zu gewährleisten.

	<i>ARMA(1,1)</i> Prozess	<i>AR(2)</i> Prozess	<i>MA(2)</i> Prozess
$\omega = 5 \pm 1$	1%	1%	< 0.5%
$\omega = 10 \pm 1$	0.5%	1%	< 0.5%
$\omega = 15 \pm 1$	< 0.5%	0.5%	< 0.5%

Tabelle 8.5: Dargestellt ist der Prozentsatz der mit Ausreißern verunreinigten Werte der Zeitreihe, der in der Zeitreihe enthalten sein kann, ohne dass die Schätzung der Prozessparameter durch die Ausreißer verfälscht wird. Er wird für verschiedene Größen ω der in den Zeitreihen enthaltenen Ausreißer aufgeführt. Die Zeitreihen folgen den Prozessgleichungen (8.13) bis (8.15).

Betrachtet man die Mittelwerte der geschätzten Prozessparameter, erkennt man, dass ab der ersten Grenze eine systematische Verfälschung eintritt. Als Beispiel kann man den geschätzten Mittelwert des Prozessparameters θ_1 des *ARMA(1,1)* Prozesses in Tabelle A.25 aufführen, welcher bei mehr als 1% Ausreißern stetig absinkt, je größer der Prozentsatz wird. Das bekräftigt die erste Grenze für den Prozentsatz an Ausreißern, der in der Zeitreihe enthalten sein darf.

Der eben angesprochene systematische Effekt wird von der zweiten Grenze unterbrochen. Sie markiert den Prozentsatz der Ausreißer, der enthalten sein darf, bevor der Algorithmus vollständig versagt, weil er nicht mehr in der Lage ist die Ausreißer zu detektieren. Die Ursache hierfür ist, dass die erste Schätzung der Prozessparameter durch die Ausreißer stark verfälscht wird. Da die Ausreißermuster mit den Prozessparametern berechnet werden, sind sie ebenfalls stark verfälscht, so dass sie in den Residuen kaum mehr wiederzufinden sind. Eine weitere Ursache liegt bei der durch die Ausreißer grob verfälschten Varianz σ_ε^2 der Residuen, welche in die Teststatistik nach Gleichung (7.3) einfließt. Da sie wesentlich zu groß geschätzt wird, werden die Teststatistiken entsprechend zu klein geschätzt. Folglich wird die Hypothese, dass ein Ausreißer vorliegt, fälschlicherweise häufig abgelehnt.

Den Punkt, bei dem dieser systematische Effekt eintritt, kann man leicht in den Tabellen wiederfinden. Er liegt dort, wo die durchschnittliche Anzahl der detektierten Ausreißer wieder zu sinken beginnt. Eine Begleiterscheinung ist die sprunghaft wachsende Varianz σ_ε^2 der Residuen. Diese ist damit zu begründen, dass die Ausreißer nicht mehr entfernt werden und daher in der Varianz σ_ε^2 der Residuen enthalten sind. Im Übergangsbereich besitzt der Mittelwert der Varianz $\mu_{\sigma_\varepsilon^2}$ der Residuen eine hohe Standardabweichung, welche dadurch erklärt werden kann, dass bei einem Teil der 1000 Stichproben das Detektieren der Ausreißer noch gelingt und beim Rest nicht. Genau in diesem Übergangsbereich ist die zweite Grenze anzu-

siedeln.

	<i>ARMA(1,1)</i> Prozess	<i>AR(2)</i> Prozess	<i>MA(2)</i> Prozess
$\omega = 5 \pm 1$	10%	10%	10%
$\omega = 10 \pm 1$	10%	11%	10%
$\omega = 15 \pm 1$	10%	10%	10%

Tabelle 8.6: Dargestellt ist der Prozentsatz der mit Ausreißern verunreinigten Werte der Zeitreihe, bei dem die Schätzung der Prozessparameter und Ausreißer mit dem modifizierten Algorithmus von CHEN und LIU nicht mehr gelingt. Er wird für verschiedene Größen ω der in den Zeitreihen enthaltenen Ausreißer aufgeführt. Die Zeitreihen folgen den Prozessgleichungen (8.13) bis (8.15).

Wenn mehr Ausreißer in der Zeitreihe enthalten sind, als in Tabelle 8.6 angegeben, ist der Algorithmus nicht mehr in der Lage, die Prozessparameter und Ausreißer korrekt zu schätzen. Diese Grenzen sind für verschiedene Prozesse unterschiedlich. Sie werden wesentlich von der Länge der Ausreißermuster beeinflusst und damit von der kleinsten Wurzel des Lag-Polynoms $\theta(B)$ der *ARMA(p,q)* Prozesse. Sind beispielsweise 10% Ausreißer in einer Zeitreihe der Länge $T = 1000$ enthalten, so beträgt die Zeitdifferenz zwischen den Ausreißern $\Delta t = 10$. Sind die Ausreißermuster länger als 10 Werte, so beginnen sie sich zu überlagern, was die Detektierung der Ausreißer erschwert. Außerdem bleiben auch bei Ausreißermustern, die kürzer als 10 Werte sind, kaum noch unverfälschte Residuen übrig, so dass die geschätzte Varianz der Residuen nach der ersten Schätzung der Prozessparameter wesentlich zu groß ist. Somit werden die Teststatistiken $\hat{\tau}$ in Gleichung (7.3), die für die Detektion der Ausreißer verantwortlich sind, häufig die Hypothese annehmen, dass kein Ausreißer vorliegt. Diese Effekte treten stärker auf, je länger die Ausreißermuster werden.

Fazit: Der Prozentsatz an Ausreißern, der in der Zeitreihe enthalten sein darf, hängt von der Länge der Ausreißermuster ab. Für die hier verwendeten Prozesse sollte der Prozentsatz nicht mehr als 1% betragen, damit die Parameter korrekt geschätzt werden. Kommt es dagegen nur auf die Detektierung der Ausreißer an, darf der Prozentsatz auch höher sein. Er sollte jedoch keinesfalls 10% überschreiten.

8.6 Qualität der Modellidentifikation

In der Regel wird man keine präzise Information über die Ordnung des *ARMA(p,q)* Prozesses haben, der einer Zeitreihe zu Grunde liegt. Informationskriterien sind Hilfsmittel für die Bestimmung der Prozessordnung. Es soll untersucht werden, welchen Einfluss Ausreißer darauf haben.

Dazu werden die Informationskriterien für $p = 0, 1, \dots, 9$ und $q = 0, 1, \dots, 9$ berechnet. Für jede Zeitreihe werden sie einmal mit dem modifizierten Algorithmus von CHEN und LIU bestimmt und ein anderes Mal ganz ohne Ausreißerschätzung. Letzteres ist die Schätzung der Prozessparameter nach der Methode der kleinsten Quadrate. Es werden 100 Zeitreihen herangezogen, die den Prozessen

$$ARMA(1,1) \text{ Prozess: } y_t = 0.6y_{t-1} + 0.3\varepsilon_{t-1} + \varepsilon_t, \sigma_\varepsilon^2 = 1.0 \quad (8.16)$$

$$AR(2) \text{ Prozess: } y_t = 0.6y_{t-1} + 0.2y_{t-2} + \varepsilon_t, \sigma_\varepsilon^2 = 1.0 \quad (8.17)$$

$$MA(2) \text{ Prozess: } y_t = 0.6\varepsilon_{t-1} + 0.2\varepsilon_{t-2} + \varepsilon_t, \sigma_\varepsilon^2 = 1.0 \quad (8.18)$$

folgen. Diese Prozesse werden auch in den vorigen Abschnitten als Beispiele verwendet, so dass die zuvor erzielten Ergebnisse auch hier gelten. Zusätzlich zu den Informationskriterien wird die Anzahl der detektierten Ausreißer bestimmt, um deren Einfluss auf die Informationskriterien herauszustellen. Die Anzahl der Werte pro Zeitreihe, die bei der Schätzung der Prozessparameter verwendet wird, ist gleich der Länge T der Zeitreihe. Dies entspricht der Empfehlung in [7], diese Größe konstant zu halten. Die in Tabelle 8.7 aufgelisteten Ausreißer werden der Zeitreihe hinzugefügt.

Zeitpunkt	100	200	300	400	500	600	700	800	900
Typ	I	A	A	Z	I	A	Z	Z	A
Größe	9.0	-9.0	-7.0	9.0	18.0	-7.0	-14.0	8.0	8.0

Tabelle 8.7: Dargestellt sind die Ausreißer, mit denen die zur Modellidentifikation verwendeten Zeitreihen verunreinigt wurden. Sie werden mit ihrem Zeitpunkt t_a des Eintretens, ihrem Typ und ihrer Größe ω angegeben. Dabei stehen die Abkürzungen I, A und Z für einen innovativen Ausreißer, einen additiven Ausreißer und eine zeitliche Änderung.

Der Strafterm für die detektierten Ausreißer wirkt sich lediglich auf eine zu niedrig gewählte Modellordnung aus. Nur in den Fällen $p = 0, q = 0$ und $p = 0, q = 1$ werden mehr Ausreißer als in allen anderen Fällen detektiert, wie in den Tabellen A.37 bis A.39 zu sehen ist. In den anderen Fällen ist der Strafterm gleich groß und kann daher keinen Unterschied bei der Wahl der Modellordnung ausmachen. Das bedeutet, dass die Ausreißer nicht durch ein Modell mit größerer Ordnung erklärt werden können.

Die Tabellen A.35 und A.36 zeigen, dass die Bestimmung der Modellordnung ohne Schätzung der Ausreißer generell dazu neigt, einen autoregressiven Prozess auszuwählen und Moving Average Anteile abzulehnen. Daher liegt sie im Fall der $AR(2)$ Zeitreihen oft richtig und für die $ARMA(1,1)$ und $MA(2)$ Zeitreihen entsprechend oft falsch. Die relativ guten Ergebnisse für die $AR(2)$ Zeitreihen, bei denen keine Schätzung von Ausreißern erfolgt ist, erhalten daher beim Vergleich zur Wahl der Modellordnung mit Schätzung der Ausreißer ein geringes Ge-

wicht. Da die Variante mit Schätzung von Ausreißern sonst vergleichsweise bessere Ergebnisse liefert, ist diese zu bevorzugen.

Die Ergebnisse in Tabelle A.34 zeigen, dass die Wahl der Modellordnung sehr oft fehlschlägt, wenn Ausreißer in der Zeitreihe enthalten sind. Die neun hinzugefügten Ausreißer entsprechen 0.9% verunreinigten Beobachtungen. Die vorangegangenen Untersuchungen haben ergeben, dass dieser Prozentsatz in dem Grenzbereich liegt, wo erste systematische Einflüsse auf die Schätzung der Prozessparameter zu erwarten sind. Davon ist besonders die geschätzte Varianz $\hat{\sigma}_\varepsilon^2$ der Residuen betroffen, die direkt in die Informationskriterien einfließt. Demnach ist die Bestimmung der Modellordnung mit Informationskriterien problematisch, wenn Ausreißer in der Zeitreihe enthalten sind.

Die Ergebnisse des *AIC* sind relativ zu den anderen Kriterien schlechter. Das hängt damit zusammen, dass dieses Kriterium nicht konsistent ist². Die Neigung des *AIC*, die Modellordnung zu überschätzen, ist auch hier zu erkennen. Das *MixedC* liegt etwas besser bei der Wahl der Modellordnung als die *BIC* und *HQC*. Diese leichten Vorteile resultieren daraus, dass der Strafterm für die Anzahl der Ausreißer beim *MixedC* etwas kleiner ist.

Fazit: Die Bestimmung der Modellordnung mit Informationskriterien ist sehr problematisch, wenn Ausreißer in der Zeitreihe enthalten sind. Man sollte sie nicht als einen genauen Wert ansehen, sondern als eine Schätzung, in deren Nähe die korrekte Modellordnung liegt. Diese Schätzungen sind besser, wenn Ausreißer berücksichtigt werden. Von den einzelnen Informationskriterien liefert das *MixedC* die besten Ergebnisse.

8.7 Zusammenfassung

In diesem Kapitel wurde herausgestellt, dass die gemeinsame Schätzung der Prozessparameter und Ausreißer sinnvoll ist. Durch sie wird die Schätzung der Prozessparameter genauer, was zur Folge hat, dass die Trennung von Signal und weißem Rauschen besser gelingt. Außerdem wird die Varianz σ_ε^2 der Residuen besser geschätzt, was sich positiv auf die Modellidentifikation auswirkt.

Der Vergleich des von CHEN und LIU vorgestellten Algorithmus mit dessen modifizierter Version hat gezeigt, dass beide Algorithmen gleichwertige Ergebnisse liefern. Da der modifizierte Algorithmus von CHEN und LIU jedoch wesentlich weniger Rechenzeit in Anspruch nimmt und wenig Speicherplatz verbraucht, da die Normalgleichungen direkt aufgestellt werden, lautet die Empfehlung diesen Algorithmus zu verwenden.

²Siehe [10], Kapitel 8.1, Seite 110.

Die Qualität der Schätzung von Ausreißern wird im Wesentlichen durch die Größe der Ausreißer, die Ausreißermuster und die Anzahl der Ausreißer in der Zeitreihe beeinflusst.

Die Ausreißer benötigen eine Mindestgröße, um detektiert zu werden, da sonst die Teststatistik $\hat{\tau}(t_a)$ in Gleichung (7.3) den kritischen Wert nicht überschreitet. Sind sehr große Ausreißer in der Zeitreihe, kann deren Detektion ebenfalls fehlschlagen, da durch sie die Varianz σ_ε^2 der Residuen, welche in die Berechnung der Teststatistiken einfließt, verfälscht wird.

Die Form der Ausreißermuster hat großen Einfluss auf die Identifizierung der Ausreißertypen. Sie wird mit den Prozessparametern berechnet und nimmt somit für jeden Prozess eine individuelle Gestalt an. Da sich die Ausreißermuster zum Teil ähneln, kann es zu Verwechslungen des Ausreißertyps kommen. Auch die Länge der Ausreißermuster spielt eine Rolle. Sie entscheidet darüber, wie nahe zwei Ausreißer beieinander liegen können, ohne dass die Detektion und Identifizierung dieser Ausreißer leidet. Extrem lange Ausreißermuster, wie beispielsweise das Muster eines Level Shifts, führen zu Problemen bei der Schätzung dieser Ausreißer.

Die Anzahl der Ausreißer, die in einer Zeitreihe enthalten sind, kann systematische Fehler bei der Schätzung der Prozessparameter bewirken, falls sie hoch ist. Ist diese Anzahl zu hoch, gelingt es nicht mehr die Ausreißer zu detektieren und die Schätzung der Prozessparameter und Ausreißer schlägt fehl. Wie hoch diese Anzahl sein darf, hängt von der Größe und den Mustern, insbesondere deren Länge, der in der Zeitreihe enthaltenen Ausreißer ab.

Kapitel 9

Tests zu $ARCH(m)$ Zeitreihen

In Kapitel 1 wird die Modellierung des farbigen Rauschens, welches nach dem Abziehen des Trends von den Beobachtungen übrig bleibt, durch einen ARCH-Prozess vorgeschlagen. In diesem Kapitel wird die Qualität der Schätzung der Prozessparameter untersucht. Dazu wird in der Gleichung

$$y_t = \mathbf{x}'\boldsymbol{\beta} + u_t$$

der Zeitreihe y_t mit dem ARCH-Rauschen u_t als Beispiel für das Regressionsmodell $\mathbf{x}'\boldsymbol{\beta}$ eine Gerade verwendet.

Die Schätzung der Modellparameter erfolgt nach der Maximum-Likelihood Methode. Bei der Maximierung der Log-Likelihoodfunktion werden vier verschiedene numerische Optimierungsverfahren getestet, um das geeignetste herauszufinden. Dabei werden auch verschiedene Probleme bei der Schätzung der Parameter herausgestellt.

Anschließend wird die Qualität der Modellidentifikation mit Hilfe von Informationskriterien untersucht. Dabei werden die Informationskriterien mit Schätzungen der Prozessparameter berechnet, welche durch verschiedene Schätzmethoden bestimmt werden. Das Ziel ist, die zur Modellidentifikation geeignetste Schätzmethode zu ermitteln.

9.1 Vergleich der numerischen Optimierungsverfahren

Dieser Abschnitt beschäftigt sich mit der Schätzung der Parameter von $ARCH(m)$ Prozessen und den Problemen, die dabei auftreten. Als Schätzmethode wird die Maximum-Likelihood Schätzung verwendet. Eine Minimierung nach der Methode der kleinsten Quadrate ist zur Schätzung der Parameter nicht geeignet, da die Residuen $u_t \sim N(0, h_t)$ einer Normalverteilung gehorchen, bei der die Varianz h_t selbst als Zeitreihe anzusehen ist und von den zu

schätzenden Prozessparametern abhängt. Für die Bestimmung von Näherungswerten hingegen kommt sie in Frage. Ein Ausschnitt aus einer Zeitreihe, die dem $ARCH(2)$ Prozess

$$u_t = \sqrt{10 + 0.5u_{t-1}^2 + 0.4u_{t-2}^2} \cdot v_t \quad (9.1)$$

folgt, ist in Abbildung 9.1 zu sehen. Man erkennt deutlich, dass die Varianz in der Zeitspanne von $t = 110$ bis $t = 135$ höher ist als im Rest des dargestellten Ausschnitts.

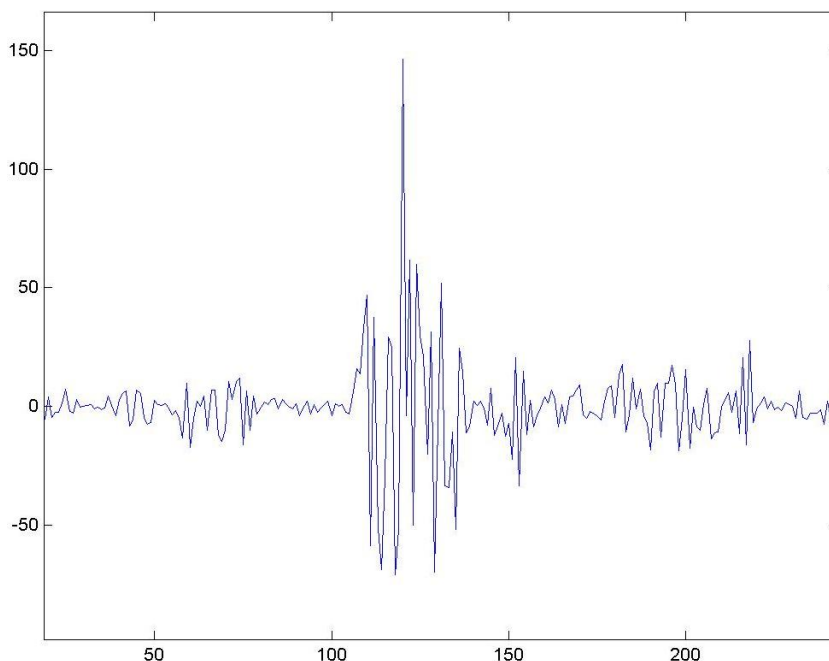


Abbildung 9.1: Dargestellt ist ein Ausschnitt aus einer $ARCH(2)$ Zeitreihe, die der Prozessgleichung (9.1) folgt. Es ist deutlich zu erkennen, dass sich die Varianz h_t über die Zeit ändert.

Um die Qualität der verschiedenen Optimierungsverfahren zu testen, wird als Beispiel für das Regressionsmodell in der Gleichung

$$y_t = \mathbf{x}'\boldsymbol{\beta} + u_t \quad (9.2)$$

eine Gerade mit der Steigung $\beta_1 = 0.3$ und dem Achsenabschnitt $\beta_2 = 10.0$ gewählt. Das Rauschen u_t , welches auf die Gerade addiert wird, folgt dem $ARCH(1)$ Prozess

$$u_t = \sqrt{10 + 0.5u_{t-1}^2} \cdot v_t. \quad (9.3)$$

Es werden insgesamt 1000 Zeitreihen erzeugt. Das Maximum der Log-Likelihood Funktion wird mit den folgenden Verfahren bestimmt, welche in den Abschnitten 5.2.2 und 5.2.3 dargestellt sind:

- Verfahren des steilsten Anstiegs
 - mit analytischer Berechnung des Gradienten (bezeichnet mit „SAA“),
 - mit numerischer Berechnung des Gradienten (bezeichnet mit „SAN“),
- NEWTON-RAPHSON Verfahren
 - mit analytischer Berechnung des Gradienten (bezeichnet mit „NRA“),
 - mit numerischer Berechnung des Gradienten (bezeichnet mit „NRN“).

Die HESSE-Matrix, welche die letzten beiden Verfahren verwenden, wird analytisch berechnet. Die ersten beiden Verfahren werden mit der Gittersuche kombiniert. Die Gitterpunkte liegen in Richtung des Gradienten und sind jeweils 2^{-k} , $k = 1, \dots, 20$ vom Ausgangspunkt der jeweiligen Iteration entfernt. Der größtmögliche Schritt beträgt folglich 0.5 und der kleinste ungefähr 10^{-8} in Richtung des Gradienten.

Bei der Bestimmung von Näherungswerten werden zuerst die Parameter der Geraden nach der Methode der kleinsten Quadrate geschätzt. Mit dieser Schätzung wird das $ARCH(m)$ Rauschen u_t bestimmt. Nach Quadrieren erhält man das quadratische Rauschen u_t^2 , welches einem $AR(m)$ Prozess nach Gleichung (4.4) folgt. Es ist sinnvoll, die Normalgleichungen für die Prozessparameter für eine Schätzung nach der Methode der kleinsten Quadrate direkt aufzustellen, da sich viele Elemente der Matrix aus den vorigen Elementen berechnen lassen.¹ Dass dieses Vorgehen gute Näherungswerte liefert, kann aus Tabelle 9.1 entnommen werden.

Aus den geschätzten Parametern der Gerade und den Prozessparametern werden die in Tabelle 9.1 aufgelisteten Mittelwerte sowie die zugehörigen Standardabweichungen abgeleitet. Führt man die Hypothesentests

$$H_0: \mu_{\beta_1} = 0.3, \quad H_0: \mu_{\beta_2} = 10.0, \quad H_0: \mu_{\zeta} = 10.0, \quad H_0: \mu_{\alpha_1} = 0.5$$

für die geschätzten Parameter durch, werden alle Nullhypothesen auch dann noch angenommen, wenn das Signifikanzniveau 90% beträgt. Von den vier zu Parametern wird die Geradensteigung β_1 am genauesten geschätzt. Die Bestimmung der Prozessparameter scheint dagegen sehr unsicher zu sein. Die Verfahren SAA und SAN liefern gleichwertige Ergebnisse, was an den Mittelwerten und Standardabweichungen zu erkennen ist. Das Verfahren NRA liefert dagegen eine etwas genauere Schätzung. Die Parameter der Gerade und die Prozessparameter werden vom Verfahren NRN am genauesten geschätzt, was man an den niedrigeren Standardabweichungen erkennen kann. Die Mittelwerte der geschätzten Prozessparameter liegen im Vergleich zu den anderen Verfahren wesentlich näher bei den Werten aus Gleichung

¹Ein ähnliches Vorgehen findet man bei der Schätzung der Parameter von $ARMA(p, q)$ Prozessen nach der Methode der kleinsten Quadrate.

(9.3). Die Ursachen für diese Feststellungen sollen im Folgenden erörtert werden. Es ist dabei zu beachten, dass für andere Beispiele, welche sich hauptsächlich im Regressionmodell $\mathbf{x}'\beta$ unterscheiden, andere Phänomene mit anderen Ursachen auftreten.

	Näherungswert	SAA	SAN	NRA	NRN
μ_{β_1}	0.2999984	0.2999963	0.2999949	0.2999968	0.2999843
μ_{β_2}	10.0044753	10.0044735	10.0044764	10.0036475	10.0093789
μ_{ζ}	11.5242836	11.5242603	11.5242645	11.3241764	9.9732332
μ_{α_1}	0.4142201	0.4144296	0.4145114	0.4343775	0.4986719
σ_{β_1}	0.0004910	0.0004693	0.0004684	0.0004316	0.0003812
σ_{β_2}	0.2835705	0.2835531	0.2835681	0.2505442	0.2227699
σ_{ζ}	1.7655174	1.7655038	1.7654773	1.6896786	0.7079097
σ_{α_1}	0.0997953	0.0996492	0.0994272	0.0876632	0.0626834

Tabelle 9.1: Es wurden 100 Zeitreihen erzeugt, die der Gleichung (9.2) folgen. Die Parameter jeder Zeitreihe wurden mit den Verfahren SAA, SAN, NRA und NRN geschätzt. Dargestellt sind die Mittelwerte und Standardabweichungen der mit den verschiedenen Verfahren geschätzten Parameter.

Dass die Schätzung der Geradensteigung am genauesten gelingt, liegt an der Beschaffenheit der Log-Likelihood Funktion dieses Beispiels. Die Abbildungen 9.2 und 9.3 zeigen Querschnitte durch die Log-Likelihood Funktion, deren Maximum zu bestimmen ist.

Für die Berechnung der Querschnitte werden die Prozessparameter konstant gehalten und die Parameter der Geraden variiert. Es ist deutlich zu erkennen, dass das Maximum der Log-Likelihood Funktion zu einer Kante ausartet. Verfahren, die nur den Gradienten nutzen, werden daher sehr schlecht konvergieren. Um die Rechenzeit für den Versuch in Grenzen zu halten, wird die Anzahl der Iterationen für jedes Verfahren auf 50 beschränkt, so dass die Ergebnisse für schlecht konvergierende Verfahren nicht viel besser als die Näherungswerte sind. Tabelle 9.1 bestätigt dies.

Die Verfahren SAA und SAN beenden jedoch oft schon nach wenigen Iterationen die Suche nach dem Maximum, da in Richtung des Gradienten kein größerer Wert der Log-Likelihood Funktion mehr gefunden wird. Der Grund dafür ist, dass der Gradient senkrecht zur Kante steht und die Kante schnell erreicht wird, da die Näherungswerte qualitativ gut sind. Dann springen die Verfahren von einer Seite der Kante zur anderen hinüber, ohne sich dem Maximum wesentlich zu nähern. Wird dabei ein Punkt gefunden, der sehr weit oben auf der Kante liegt, dann kann ein Punkt, der noch höher liegt, leicht zwischen den Gitterpunkten liegen. Damit der Gradient in Richtung des Maximums zeigt, müsste die Kante genau getroffen werden, was jedoch praktisch nicht passiert. Die Verfahren, die zusätzlich die HES-

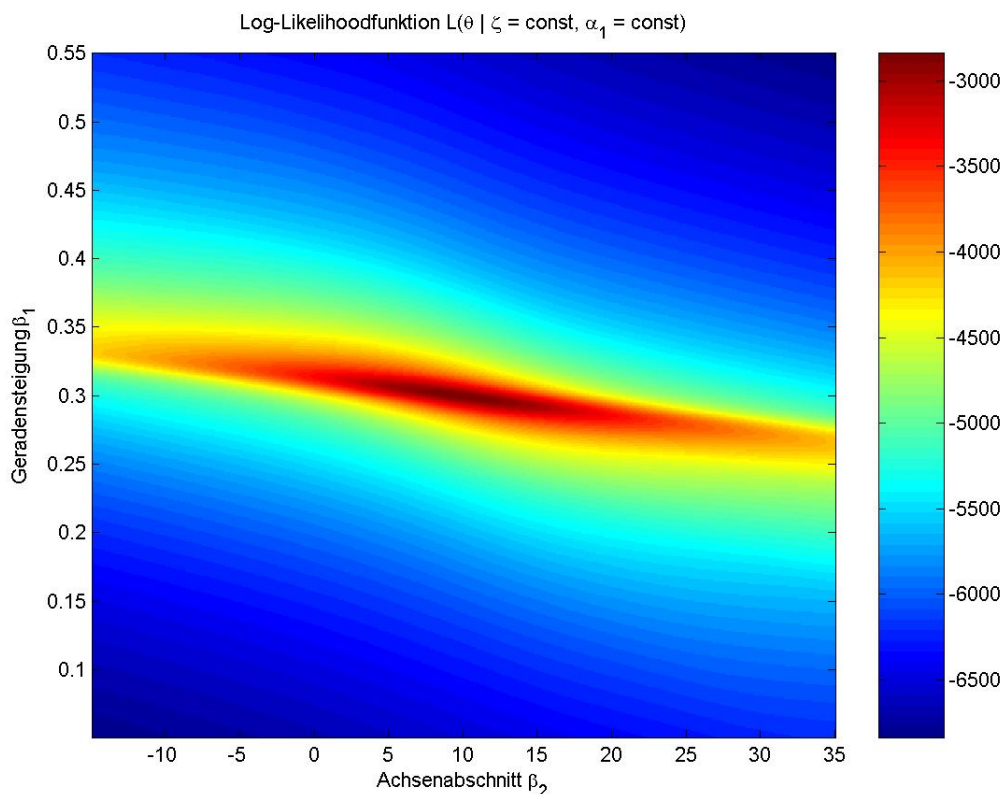


Abbildung 9.2: Dargestellt ist ein Querschnitt durch die Log-Likelihood Funktion einer Zeitreihe, die der Gleichung (9.2) folgt. Der Querschnitt läuft durch das geschätzte Maximum der Funktion. Für den Querschnitt sind die Prozessparameter des Rauschens u_t konstant, während die Parameter des Regressionsmodells variabel bleiben.

SE-Matrix nutzen, berücksichtigen nicht nur die Steigung, sondern auch die Krümmung der Log-Likelihood Funktion. Sie konvergieren daher besser.

Fazit: Das Maximum der Log-Likelihood Funktion ist für dieses Beispiel nur schwer zu bestimmen. Verfahren, die nur den Gradienten nutzen, scheitern hier. Die Tests haben herausgestellt, dass das NEWTON-RAPHSON Verfahren mit numerischer Bestimmung des Gradienten die beste Schätzung liefert, insbesondere was die Prozessparameter angeht.

9.2 Modellidentifikation für ARCH(m) Prozesse

In diesem Abschnitt wird die Qualität der Modellidentifikation für ARCH(m) Prozesse untersucht. Für die Modellidentifikation werden Informationskriterien genutzt. Ein besonderes Problem bei der Modellidentifikation bei ARCH(m) Prozessen stellen die Einschränkungen

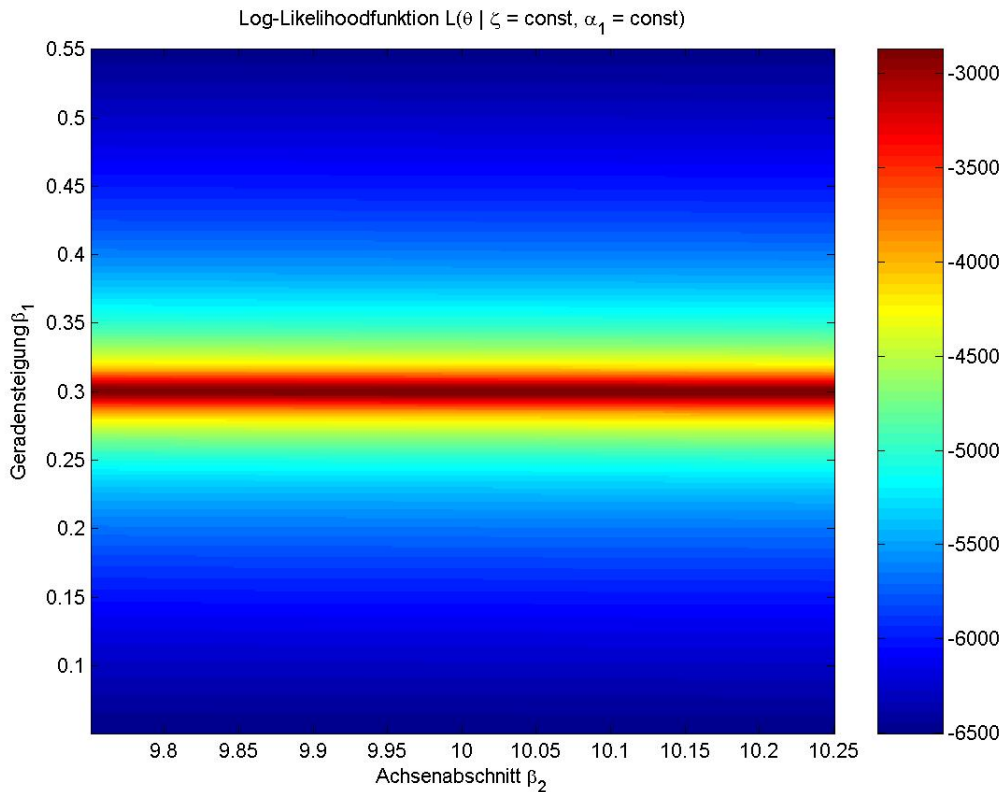


Abbildung 9.3: Dargestellt ist ein Querschnitt durch die Log-Likelihood Funktion einer Zeitreihe, die der Gleichung (9.2) folgt. Der Querschnitt läuft durch das geschätzte Maximum der Funktion. Für den Querschnitt sind die Prozessparameter des Rauschens u_t konstant, während die Parameter des Regressionsmodells variabel bleiben.

für die Prozessparameter dar:

$$\begin{aligned} \alpha_1 + \alpha_2 + \dots + \alpha_m &< 1, \\ \alpha_1, \alpha_2, \dots, \alpha_m &> 0. \end{aligned} \tag{9.4}$$

Berechnet man den Wert eines Informationskriterium für eine Modellordnung n , die größer als die wahre Modellordnung m ist, setzt man Parameter an, deren Erwartungswert gleich Null ist:

$$E(\alpha_{m+1}) = E(\alpha_{m+2}) = \dots = E(\alpha_n) = 0.$$

Da $ARCH(m)$ Prozesse stochastischer Natur sind, wird die Schätzung der Prozessparameter nie genau ihrem Erwartungswert entsprechen. Sie wird jedoch um den Erwartungswert streuen, so dass ein Teil der Parameter $\alpha_{m+1}, \alpha_{m+2}, \dots, \alpha_n$ größer Null und der Rest kleiner Null sein wird. Damit ist es wahrscheinlich, dass bei der Modellidentifikation die Einschränkungen nach Gleichung (9.4) verletzt werden. Aus diesem Grund werden folgende zwei Möglichkeiten in Erwägung gezogen.

Die erste Möglichkeit ist, die Schätzung der Parameter mit dem Verfahren NRN² mit numerischer Berechnung des Gradienten durchzuführen. Werden die Einschränkungen nach den Gleichungen (9.4) verletzt, wird der Wert des Informationskriteriums der betroffenen Modellordnung gleich Unendlich gesetzt.³

Bei der zweiten Möglichkeit wird eine Schätzung der Prozessparameter mit dem Simplexverfahren durchgeführt. Damit können negative α von vorne herein ausgeschlossen werden. Zunächst werden dabei die Parameter der Geraden nach der Methode der kleinsten Quadrate geschätzt. Mit dieser Schätzung kann man das Rauschen u_t bestimmen, welches nach einer elementweisen Quadrierung dem Prozess nach Gleichung (4.4) folgt:

$$u_t^2 = \zeta + \alpha_1 u_{t-1}^2 + \alpha_2 u_{t-2}^2 + \cdots + \alpha_m u_{t-m}^2 + w_t.$$

Mit Hilfe dieser Gleichung lässt sich ein lineares Gleichungssystem aufstellen, welches nach einer Aufspaltung der Prozessparameter α und den Residuen ω_t in ihre positiven und negativen Anteile

$$\begin{aligned}\alpha &= \alpha^+ - \alpha^- \\ \omega_t &= \omega_t^+ - \omega_t^-\end{aligned}$$

mit dem Simplexverfahren gelöst werden kann. Da die Prozessparameter α positiv sein sollen, werden die negativen Anteile α^- weggelassen. Für Genaueres zum Lösen solcher Gleichungssysteme nach dem Simplex-Verfahren sei auf [4] verwiesen.

Als Beispiel wird für das Regressionsmodell

$$y_t = \mathbf{x}'_t \boldsymbol{\beta} + u_t \tag{9.5}$$

wieder die Gerade mit der Steigung $\beta_1 = 0.3$ und dem Achsenabschnitt $\beta_2 = 10.0$ aus dem vorgigen Abschnitt gewählt. Das Rauschen folgt dieses Mal dem ARCH(2) Prozess:

$$u_t = \sqrt{10 + 0.5u_{t-1}^2 + 0.3u_{t-2}^2} \cdot v_t.$$

Die Länge der erzeugten Zeitreihen beträgt $T = 300$ Werte. Die Länge der Zeitreihen wird so kurz gewählt, da das Simplexverfahren sehr rechenintensiv ist. Zusätzlich werden die Informationskriterien auf den Näherungswerten basierend berechnet. Es werden insgesamt 100 Zeitreihen verwendet und dabei die Informationskriterien für die Ordnungen $m = 0, 1, \dots, 5$ bestimmt. Für die mit den Informationskriterien bestimmten Modellordnungen wird aus den 100 Stichproben jeweils der Mittelwert und dessen Standardabweichung abgeleitet sowie der

²Die Abkürzung „NRN“ steht für das NEWTON-RAPHSON Verfahren mit numerischer Berechnung des Gradienten und analytischer Bestimmung der HESSE-Matrix. Siehe Abschnitt 5.2.

³Dadurch wird diese Modellordnung abgelehnt. Problematisch wird dieses Vorgehen, wenn Parameter gleich Null sind wie beispielsweise beim Prozess $u_t = \sqrt{0.4u_{t-2}^2} \cdot v_t$.

Prozentsatz der korrekt bestimmten Modellordnung berechnet. Letzterer wird in den Tabellen 9.2 und 9.3 mit „Rate“ bezeichnet.

	Näherungswerte			Verfahren NRN			Simplexverfahren		
	<i>AIC</i>	<i>BIC</i>	<i>HQC</i>	<i>AIC</i>	<i>BIC</i>	<i>HQC</i>	<i>AIC</i>	<i>BIC</i>	<i>HQC</i>
μ_m	2.19	1.96	2.19	1.80	1.68	1.80	1.89	1.82	1.94
σ_m	0.82	0.68	0.82	0.70	0.51	0.70	1.23	1.21	1.26
Rate	58%	71%	58%	59%	64%	59%	15%	15%	16%

Tabelle 9.2: Für 100 $ARCH(m)$ Zeitreihen der Länge $T = 300$, die der Gleichung (9.5) folgen, wurde die Modellordnung m mit Gleichung (6.2) bestimmt. Als Grundlage der Berechnung wurden die nach der Methode der kleinsten Quadrate geschätzten Näherungswerte, die Schätzung nach dem Verfahren NRN sowie die Schätzung nach dem Simplexverfahren verwendet. Aus den 100 Modellordnungen wurde der Mittelwert und die Standardabweichung berechnet. Zusätzlich wurde von den 100 Modellordnungen der mit „Rate“ bezeichnete Prozentsatz bestimmt, bei dem die Modellordnung korrekt war.

Der Tabelle 9.2 kann man entnehmen, dass die Informationskriterien AIC und HQC , die mit den vom Verfahren NRN geschätzten Parametern berechnet werden, etwas besser sind, als die auf den Näherungswerten basierenden. Beim BIC , welches die besten Ergebnisse liefert, ist dies umgekehrt. Der Mittelwert des auf den Näherungswerten basierenden BIC liegt sehr nahe bei seinem Erwartungswert und die Rate der korrekt bestimmten Modellordnung ist mit 71% am höchsten. Die Mittelwerte der Informationskriterien, die mit dem Simplexverfahren bestimmt werden, liegen ebenfalls nahe bei ihrem Erwartungswert, jedoch ist die Standardabweichung vergleichsweise groß und die Rate der korrekt bestimmten Modellordnung niedrig. Deswegen und auf Grund der hohen Rechenzeit scheidet die Berechnung der Informationskriterien mit Hilfe des Simplexverfahrens aus.

In einem weiteren Test, bei dem lediglich die Länge der Zeitreihen zu $T = 10000$ geändert wird, soll geklärt werden, ob die Berechnung der Informationskriterien mit den Näherungswerten ausreicht. Da die Bestimmung der Näherungswerte sehr schnell ist, hätte dies Vorteile für die Rechenzeit bei langen Zeitreihen.

Betrachtet man die in Tabelle 9.3 dargestellten Ergebnisse, so stellt man fest, dass für lange Zeitreihen die Bestimmung der Informationskriterien mit den Näherungswerten für die Parameter nicht ausreicht. Der Mittelwert der Modellordnung, dessen Standardabweichung sowie der Prozentsatz der korrekt detektierten Modellordnungen sind bei Verwendung der nach dem Verfahren NRN geschätzten Parameter wesentlich besser.

	Näherungswerte			NEWTON-RAPHSON		
	<i>AIC</i>	<i>BIC</i>	<i>HQC</i>	<i>AIC</i>	<i>BIC</i>	<i>HQC</i>
μ_m	2.50	2.41	2.50	1.95	1.81	1.94
σ_m	0.99	0.94	0.99	0.74	0.54	0.74
Rate	51%	52%	51%	57%	67%	58%

Tabelle 9.3: Für 100 $ARCH(m)$ Zeitreihen der Länge $T = 10000$, die der Gleichung (9.5) folgen, wurde die Modellordnung m mit Gleichung (6.2) bestimmt. Als Grundlage der Berechnung wurden die nach der Methode der kleinsten Quadrate geschätzten Näherungswerte und die Schätzung nach dem Verfahren NRN verwendet. Aus den 100 Modellordnungen wurde der Mittelwert und die Standardabweichung berechnet. Zusätzlich wurde von den 100 Modellordnungen der mit „Rate“ bezeichnete Prozentsatz bestimmt, bei dem die Modellordnung korrekt war.

Fazit: Bei der Bestimmung der Modellordnung eines $ARCH(m)$ Prozesses eignet sich die Berechnung der Informationskriterien mit den Parametern, die nach dem Verfahren NRN geschätzt werden, am besten. Da der Prozentsatz der korrekt bestimmten Modellordnung beim BIC am höchsten liegt, ist die Verwendung dieses Informationskriteriums zu empfehlen.

9.3 Zusammenfassung

Auf Grund der Eigenschaften der Log-Likelihood Funktion des verwendeten Beispiels war es schwierig die Geradenparameter und Prozessparameter zu schätzen. Das NEWTON-RAPHSON Verfahren, bei dem der Gradient numerisch und die HESSE-Matrix analytisch berechnet werden, hat bei der Maximierung der Log-Likelihood Funktion die besten Ergebnisse geliefert. Die Modellidentifikation hat sich ebenfalls als schwierig entpuppt, was mit der unsicheren Schätzung der Prozessparameter zusammenhängen kann.

Kapitel 10

Schlussbemerkungen und Ausblick

Die Aufgabenstellung dieser Diplomarbeit ist Mittel zu finden, mit denen Ausreißer detektiert und geschätzt werden können. Eingebettet ist diese Aufgabenstellung in den Kontext sehr langer Zeitreihen, wie sie bei Satellitenmissionen vorkommen. Dabei wird die Sichtweise der Signaltheorie verwendet, bei der sich die Beobachtungen l_t in einen Trend $f_t(x)$, ein Signal s_t und weißes Rauschen n_t gliedern. Diese Aufspaltung der Beobachtungen wird um Ausreißer erweitert:

$$l_t = f_t(x) + s_t + n_t + \sum_j \omega_j e_{t,j}.$$

Nach dem Abspalten des Trends von den Beobachtungen bleibt farbiges Rauschen übrig, welches mit Hilfe von ARMA-Prozessen modelliert werden kann. Diese Prozesse ermöglichen die Trennung von Signal und weißem Rauschen.

In Kapitel 7 wird der Algorithmus von CHEN und LIU vorgestellt, bei dem die Schätzung von ARMA-Prozessen mit den Ausreißern gemeinsam erfolgt. Als Alternative hierzu wird die Modellierung der Ausreißer durch ARCH-Prozesse erwogen, da sie nach [10], Kapitel 20.1.2, Seite 351 dazu geeignet sind.

Ein bis hierhin nicht erwähnter Punkt ist, dass ARCH-Prozesse nicht dazu geeignet sind Korrelationen zu modellieren, da das ARCH-Rauschen u_t aus Gleichung (4.4) unkorreliert ist. ARCH-Prozesse sind jedoch im Gegensatz zu ARMA-Prozessen in der Lage eine sich über die Zeit ändernde Varianz zu modellieren, da die Varianz h_t des ARCH-Rauschens ein AR-Prozess ist. Denkbar ist eine Kombination der beiden Prozesse, so dass sich die Prozessgleichung $\varphi(B)y_t = \theta(B)\varepsilon_t$ eines ARMA-Prozesses mit dem weißen Rauschen ε_t zu $\varphi(B)y_t = \theta(B)u_t$ mit dem ARCH-Rauschen u_t wandelt. Wie sich solche Prozesse verhalten und wie man sie schätzen kann, müsste untersucht werden, da sie wegen der sich über die Zeit ändernden Varianz h_t nicht stationär sind.

ARCH-Prozesse stellen den einfachsten Fall heteroskedastischer Prozesse dar.¹ Demnach ist die Verwendung anderer Prozesse denkbar. Bei GARCH-Prozessen² wird beispielsweise die Varianz h_t des Rauschens durch einen ARMA-Prozess modelliert. In [10], Kapitel 20.2, Seite 363 wird insbesondere die Verwendung von bilinearen Prozessen zur Modellierung ausreißer-behafteter Zeitreihen als günstig angesehen. Allerdings wird in [10], Kapitel 20.2, Seite 366 auch auf die Schwierigkeiten bei der Identifikation und Parameterschätzung allgemeiner bilinearer Prozesse hingewiesen.

Kehrt man zur Betrachtung des in Kapitel 7 vorgestellten Algorithmus von CHEN und LIU zurück, stellt man fest, dass auch hier Erweiterungen denkbar sind. Beispielsweise werden nur die vier klassischen Ausreißertypen vorgestellt. Besitzt man Kenntnisse über die Muster von Ausreißern, die in Zeitreihen vorkommen können, so empfiehlt sich deren Verwendung anstatt der vier klassischen, falls sie von diesen abweichen.

Im Zusammenhang mit den Ausreißermustern besteht das Problem, dass diese aus den geschätzten Prozessparametern, welche durch die in der Zeitreihe enthaltenen Ausreißer verfälscht sein können, berechnet werden. Dies erschwert die Detektion und Identifizierung der Ausreißer. In den in Kapitel 7 vorgestellten Formen des Algorithmus von CHEN und LIU werden die Prozessparameter nach der Maximum-Likelihood Methode und der Methode der kleinsten Quadrate geschätzt. Es ist daher zu prüfen, ob robuste Schätzverfahren, wie beispielsweise eine L_1 -Norm Minimierung nach dem Simplexverfahren, bessere Ergebnisse liefern.

Ein nicht betrachteter Punkt ist, dass Zeitreihen Datenlücken enthalten können. Es stellt sich die Frage, wie man diese bei der Schätzung der Prozessparameter und Ausreißer überbrücken kann. Ein Werkzeug, welches sich dazu eignet, ist der KALMAN-Filter. In [2], Kapitel 13 wird detailliert dargestellt, wie ein GAUSSScher ARMA-Prozess mit dem KALMAN-Filter als Zustandsraummodell formuliert werden kann. Die sich daraus ergebenden Möglichkeiten im Zusammenhang mit der Suche nach Ausreißern sind in dieser Diplomarbeit nicht untersucht worden.

Es besteht die Möglichkeit, dass die Wurzeln von ARMA-Prozessen sehr nahe am Einheitskreis liegen. Wie sich dies auf die Schätzung der Prozessparameter und Ausreißer auswirkt, ist in Kapitel 8 nur teilweise betrachtet worden. Genauere Untersuchungen hierzu sind wünschenswert, da eine solche Problemstellung für die Satellitenmission GOCE vorliegt.

Bei der Satellitenmission GOCE werden die vier Tensorkomponenten xx , yy , zz und xz beobachtet. Für jede einzelne Komponente ist farbiges Rauschen nach dem Abziehen des

¹Siehe [10], Kapitel 20.1.5, Seite 353.

²GARCH-Prozesse sind generalisierte ARCH-Prozesse: „*Generalized Autoregressive Conditional Heteroscedasticity*“.

Trends zu erwarten. Die Komponenten sind auch untereinander korreliert. Eine Modellierung durch vektorielle ARMA-Prozesse³ kann also als zulässiger und sinnvoller Ansatz angesehen werden. Die Umsetzung kann beispielsweise mit Hilfe des KALMAN-Filters geschehen. Diese Möglichkeit ist auch in [8], Kapitel 4 erkannt worden. Im Zusammenhang mit der Suche nach Ausreißern sollte dieser Ansatz ebenfalls Vorteile bringen. Treten beispielsweise Fehler am Sensor auf, liegt die Vermutung nahe, dass man beim entsprechenden Zeitpunkt in allen beobachteten Komponenten Ausreißer vorfindet. An dieser Stelle öffnet sich ein großes Feld für weitere Untersuchungen.

³Solche Prozesse werden auch als multivariate Prozesse bezeichnet.

Anhang A

Tabellen und Abbildungen

A.1 Abbildungen von Ausreißermustern

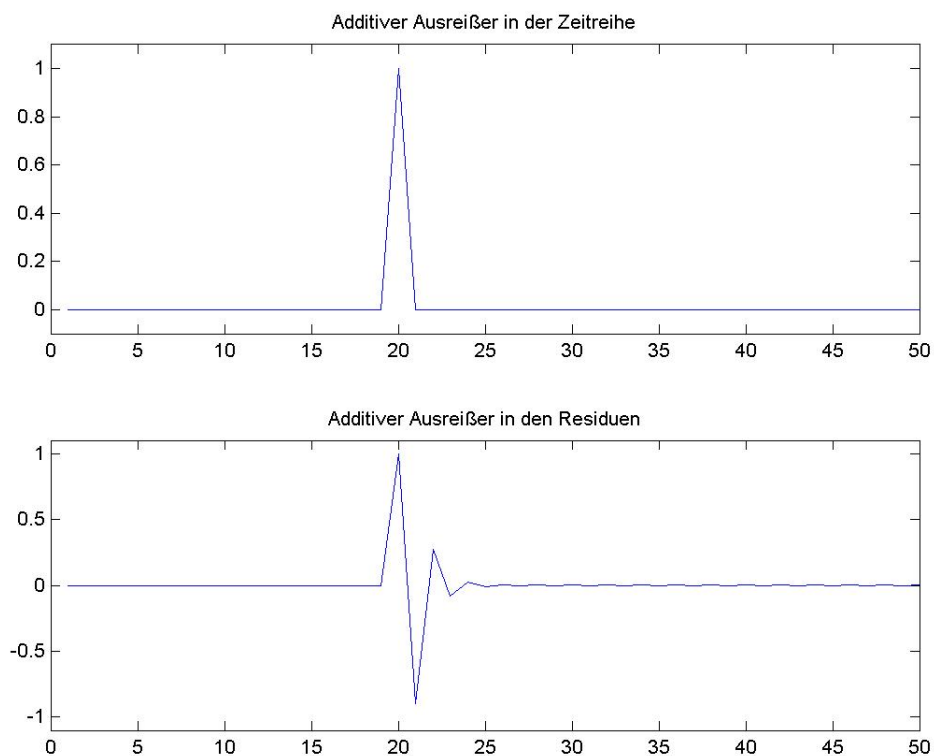


Abbildung A.1: Dargestellt sind die Muster eines additiven Ausreißers beim Zeitpunkt $t_a = 20$ für den $ARMA(1,1)$ Prozess nach Gleichung (8.2). Oben ist das Muster $\bar{x}_{A,t}$ des Ausreißers in der Zeitreihe und unten das Muster $x_{A,t}$ des Ausreißers in den Residuen zu sehen.

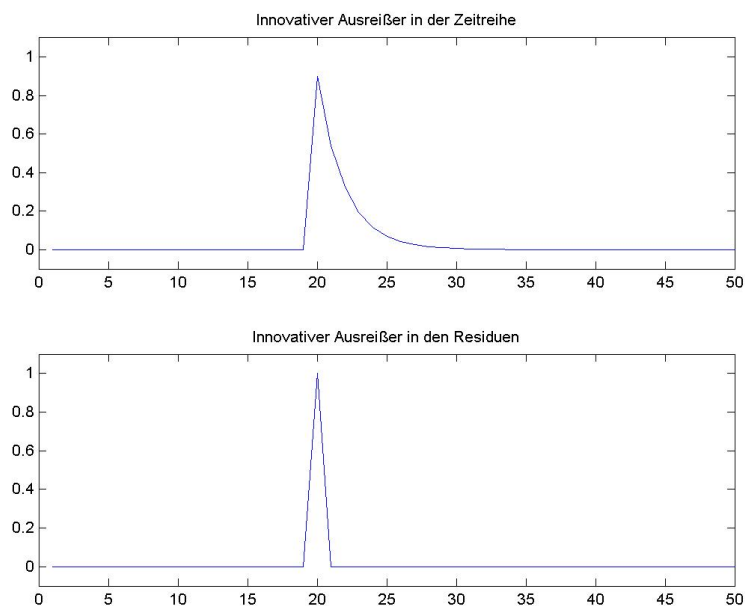


Abbildung A.2: Dargestellt sind die Muster eines innovativen Ausreißers beim Zeitpunkt $t_a = 20$ für den $ARMA(1,1)$ Prozess nach Gleichung (8.2). Oben ist das Muster $\bar{x}_{I,t}$ des Ausreißers in der Zeitreihe und unten das Muster $x_{I,t}$ des Ausreißers in den Residuen zu sehen.

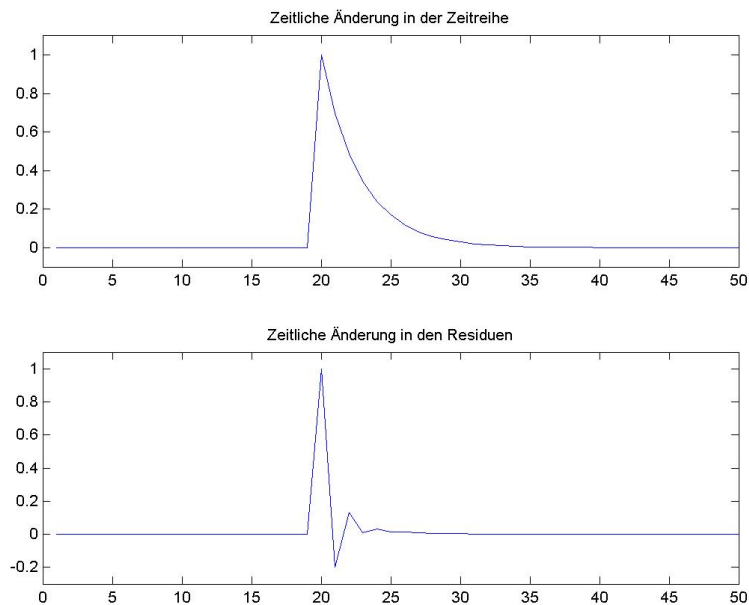


Abbildung A.3: Dargestellt sind die Muster einer zeitlichen Änderung beim Zeitpunkt $t_a = 20$ für den $ARMA(1,1)$ Prozess nach Gleichung (8.2). Oben ist das Muster $\bar{x}_{Z,t}$ des Ausreißers in der Zeitreihe und unten das Muster $x_{Z,t}$ des Ausreißers in den Residuen zu sehen.

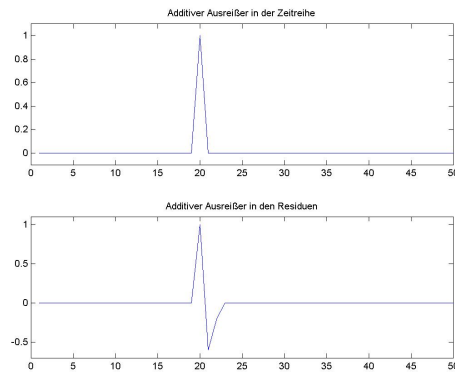


Abbildung A.4: Dargestellt sind die Muster eines additiven Ausreißers beim Zeitpunkt $t_a = 20$ für den $AR(2)$ Prozess nach Gleichung (8.3). Oben ist das Muster $\bar{x}_{A,t}$ des Ausreißers in der Zeitreihe und unten das Muster $x_{A,t}$ des Ausreißers in den Residuen zu sehen.

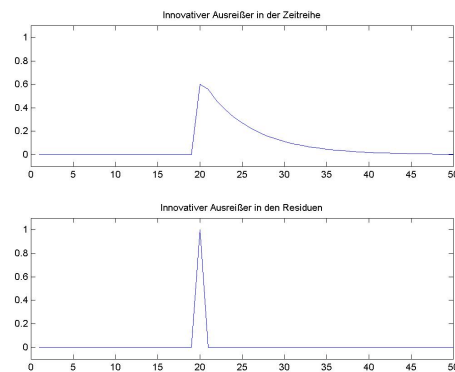


Abbildung A.5: Dargestellt sind die Muster eines innovativen Ausreißers beim Zeitpunkt $t_a = 20$ für den $AR(2)$ Prozess nach Gleichung (8.3). Oben ist das Muster $\bar{x}_{I,t}$ des Ausreißers in der Zeitreihe und unten das Muster $x_{I,t}$ des Ausreißers in den Residuen zu sehen.

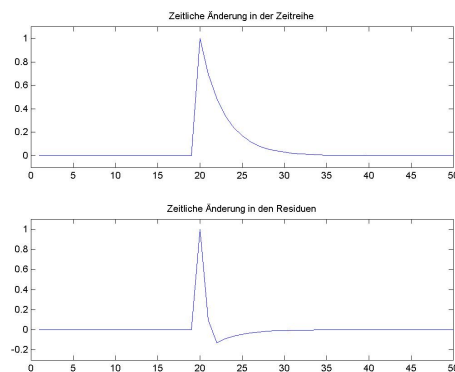


Abbildung A.6: Dargestellt sind die Muster einer zeitlichen Änderung beim Zeitpunkt $t_a = 20$ für den $AR(2)$ Prozess nach Gleichung (8.3). Oben ist das Muster $\bar{x}_{Z,t}$ des Ausreißers in der Zeitreihe und unten das Muster $x_{Z,t}$ des Ausreißers in den Residuen zu sehen.

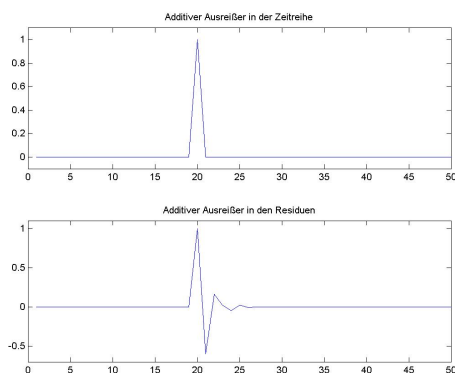


Abbildung A.7: Dargestellt sind die Muster eines additiven Ausreißers beim Zeitpunkt $t_a = 20$ für den $MA(2)$ Prozess nach Gleichung (8.4). Oben ist das Muster $\bar{x}_{A,t}$ des Ausreißers in der Zeitreihe und unten das Muster $x_{A,t}$ des Ausreißers in den Residuen zu sehen.

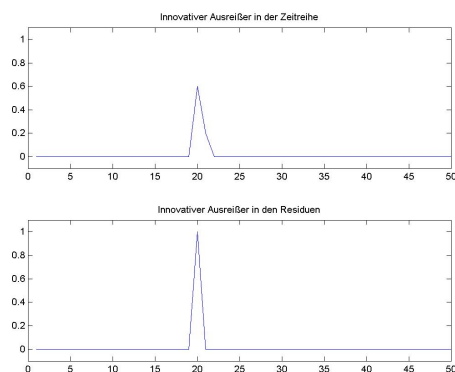


Abbildung A.8: Dargestellt sind die Muster eines innovativen Ausreißers beim Zeitpunkt $t_a = 20$ für den $MA(2)$ Prozess nach Gleichung (8.4). Oben ist das Muster $\bar{x}_{I,t}$ des Ausreißers in der Zeitreihe und unten das Muster $x_{I,t}$ des Ausreißers in den Residuen zu sehen.

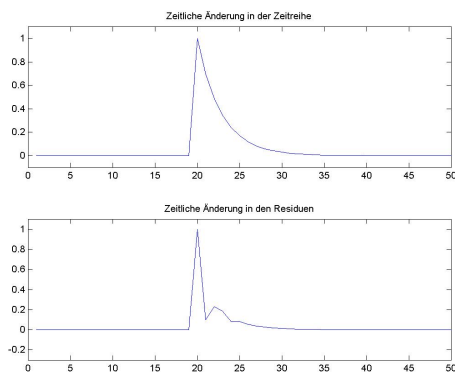


Abbildung A.9: Dargestellt sind die Muster einer zeitlichen Änderung beim Zeitpunkt $t_a = 20$ für den $MA(2)$ Prozess nach Gleichung (8.4). Oben ist das Muster $\bar{x}_{Z,t}$ des Ausreißers in der Zeitreihe und unten das Muster $x_{Z,t}$ des Ausreißers in den Residuen zu sehen.

A.2 Abbildungen von Ausschnitten aus Zeitreihen

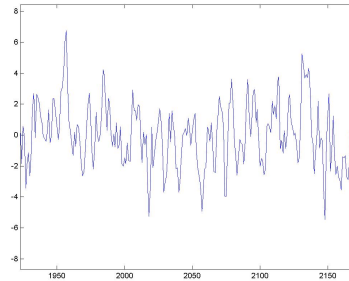


Abbildung A.10: Dargestellt ist ein Ausschnitt aus einer Zeitreihe, die dem $ARMA(1,1)$ Prozess nach Gleichung (8.2) folgt.

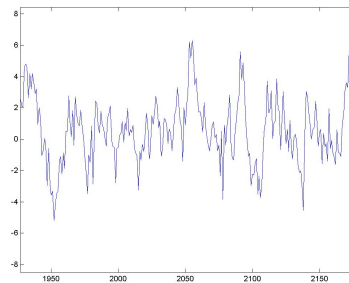


Abbildung A.11: Dargestellt ist ein Ausschnitt aus der Zeitreihe, die dem $AR(2)$ Prozess nach Gleichung (8.3) folgt.

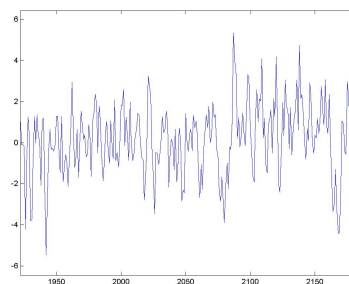


Abbildung A.12: Dargestellt ist ein Ausschnitt aus der Zeitreihe, die dem $MA(2)$ Prozess nach Gleichung (8.4) folgt.

A.3 Tabellen zu Abschnitt 8.2

In den Tabellen A.1 bis A.6 werden die Ergebnisse des Algorithmus von CHEN und LIU denen des modifizierten Algorithmus von CHEN und LIU spaltenweise gegenübergestellt. Grundlage für die Schätzung der Prozessparameter und Ausreißer sind verschiedene Zeitreihen, die den Prozessen nach den Gleichungen (8.2) bis (8.4) folgen. Dabei werden diese Zeitreihen mit den im oberen Teil der Tabelle angegebenen Ausreißern verunreinigt. Die detektierten Ausreißer werden mit ihrem Typ, dem Zeitpunkt ihres Eintretens t_a und ihrer geschätzten Größe $\hat{\omega}$ aufgeführt. Für den Ausreißertyp gelten die Abkürzungen:

- A ... additiver Ausreißer,
- I ... innovativer Ausreißer,
- Z ... zeitliche Änderung.

In der Liste der geschätzten Ausreißer sind jene Ausreißer unterstrichen, deren Zeitpunkt des Eintretens mit einem der Zeitpunkte des Eintretens der hinzugefügten Ausreißer übereinstimmt. Das bedeutet, dass ein vorher hinzugefügter Ausreißer von den Algorithmen detektiert wird. Die korrekte Identifikation des Ausreißertyps ist dabei unerheblich.

Hinzugefügte Ausreißer	keine	
	Algorithmus von CHEN und LIU	Modifizierter Algorithmus
Rechenzeit in Sekunden	52.2	0.8
Geschätzte Parameter	$\hat{\varphi}_1 = +0.60$ $\hat{\theta}_1 = +0.32$ $\hat{\sigma}_\varepsilon^2 = +2.0095$	$\hat{\varphi}_1 = +0.60$ $\hat{\theta}_1 = +0.32$ $\hat{\sigma}_\varepsilon^2 = +2.0151$
Geschätzte Ausreißer	A, $t_a = 1661$, $\hat{\omega} = +3.7$ A, $t_a = 2604$, $\hat{\omega} = +4.2$ Z, $t_a = 2695$, $\hat{\omega} = +4.9$ I, $t_a = 4640$, $\hat{\omega} = -5.2$ Z, $t_a = 5288$, $\hat{\omega} = +5.4$ A, $t_a = 5904$, $\hat{\omega} = +3.9$ Z, $t_a = 6603$, $\hat{\omega} = -5.0$ A, $t_a = 8716$, $\hat{\omega} = +4.2$ I, $t_a = 8762$, $\hat{\omega} = +5.2$	A, $t_a = 2604$, $\hat{\omega} = +4.2$ I, $t_a = 4640$, $\hat{\omega} = -5.2$ I, $t_a = 5288$, $\hat{\omega} = +5.2$ A, $t_a = 5904$, $\hat{\omega} = +3.9$ I, $t_a = 6603$, $\hat{\omega} = -5.1$ A, $t_a = 8716$, $\hat{\omega} = +4.2$ I, $t_a = 8762$, $\hat{\omega} = +5.2$

Tabelle A.1: Eine Zeitreihe, die dem $ARMA(1,1)$ Prozess nach Gleichung (8.2) folgt und die keine hinzugefügten Ausreißer enthält, ist die Grundlage für diesen Vergleich des Algorithmus von CHEN und LIU mit dem modifizierten Algorithmus von CHEN und LIU.

Hinzugefügte Ausreißer	keine	
	Algorithmus von CHEN und LIU	Modifizierter Algorithmus
Rechenzeit in Sekunden	56.2	0.4
Geschätzte Parameter	$\hat{\varphi}_1 = +0.61$ $\hat{\varphi}_2 = +0.20$ $\hat{\sigma}_\varepsilon^2 = +1.9606$	$\hat{\varphi}_1 = +0.61$ $\hat{\varphi}_2 = +0.20$ $\hat{\sigma}_\varepsilon^2 = +1.9711$
Geschätzte Ausreißer	I, $t_a = 1603$, $\hat{\omega} = -5.0$ A, $t_a = 3104$, $\hat{\omega} = +4.7$ A, $t_a = 3720$, $\hat{\omega} = -4.5$ A, $t_a = 5187$, $\hat{\omega} = -4.2$ Z, $t_a = 5331$, $\hat{\omega} = -6.1$ Z, $t_a = 6532$, $\hat{\omega} = -4.9$ Z, $t_a = 7024$, $\hat{\omega} = -4.9$ A, $t_a = 7708$, $\hat{\omega} = -5.1$	I, $t_a = 1603$, $\hat{\omega} = -5.0$ A, $t_a = 3104$, $\hat{\omega} = +4.9$ Z, $t_a = 5187$, $\hat{\omega} = -4.8$ A, $t_a = 5331$, $\hat{\omega} = -4.9$ Z, $t_a = 9784$, $\hat{\omega} = -5.5$

Tabelle A.2: Eine Zeitreihe, die dem $AR(2)$ Prozess nach Gleichung (8.3) folgt und die keine hinzugefügten Ausreißer enthält, ist die Grundlage für diesen Vergleich des Algorithmus von CHEN und LIU mit dem modifizierten Algorithmus von CHEN und LIU.

Hinzugefügte Ausreißer	keine	
	Algorithmus von CHEN und LIU	Modifizierter Algorithmus
Rechenzeit in Sekunden	87.5	0.4
Geschätzte Parameter	$\hat{\theta}_1 = +0.62$ $\hat{\theta}_2 = +0.21$ $\hat{\sigma}_\varepsilon^2 = +1.9751$	$\hat{\theta}_1 = +0.61$ $\hat{\theta}_2 = +0.21$ $\hat{\sigma}_\varepsilon^2 = +2.0006$
Geschätzte Ausreißer	I, $t_a = 53$, $\hat{\omega} = -5.0$ A, $t_a = 1292$, $\hat{\omega} = +4.4$ Z, $t_a = 2308$, $\hat{\omega} = -5.4$ I, $t_a = 2438$, $\hat{\omega} = -5.2$ A, $t_a = 2774$, $\hat{\omega} = -4.6$ Z, $t_a = 4310$, $\hat{\omega} = -4.7$ A, $t_a = 4496$, $\hat{\omega} = +4.6$ A, $t_a = 4745$, $\hat{\omega} = -5.1$ Z, $t_a = 6824$, $\hat{\omega} = +5.2$ I, $t_a = 6879$, $\hat{\omega} = -5.2$ A, $t_a = 8827$, $\hat{\omega} = +4.4$ A, $t_a = 8866$, $\hat{\omega} = +4.4$ I, $t_a = 9474$, $\hat{\omega} = -5.8$	I, $t_a = 2308$, $\hat{\omega} = -5.5$ I, $t_a = 2438$, $\hat{\omega} = -5.2$ I, $t_a = 4496$, $\hat{\omega} = +5.1$ A, $t_a = 4745$, $\hat{\omega} = -5.1$ Z, $t_a = 6824$, $\hat{\omega} = +5.3$ I, $t_a = 6879$, $\hat{\omega} = -5.2$ I, $t_a = 9474$, $\hat{\omega} = -5.8$ I, $t_a = 9877$, $\hat{\omega} = -5.2$

Tabelle A.3: Eine Zeitreihe, die dem $MA(2)$ Prozess nach Gleichung (8.4) folgt und die keine hinzugefügten Ausreißer enthält, ist die Grundlage für diesen Vergleich des Algorithmus von CHEN und LIU mit dem modifizierten Algorithmus von CHEN und LIU.

Hinzugefügte Ausreißer	$I, t_a = 1000, \hat{\omega} = +9.0$ $Z, t_a = 2000, \hat{\omega} = +10.0$ $A, t_a = 3000, \hat{\omega} = -8.0$ $Z, t_a = 5000, \hat{\omega} = -10.0$ $I, t_a = 7000, \hat{\omega} = +12.0$	
	Algorithmus von CHEN und LIU	Modifizierter Algorithmus
Rechenzeit in Sekunden	106.5	1.5
Geschätzte Parameter	$\hat{\varphi}_1 = +0.60$ $\hat{\theta}_1 = +0.28$ $\hat{\sigma}_\varepsilon^2 = +1.9885$	$\hat{\varphi}_1 = +0.60$ $\hat{\theta}_1 = +0.28$ $\hat{\sigma}_\varepsilon^2 = +1.9964$
Geschätzte Ausreißer	$A, t_a = 846, \hat{\omega} = -3.7$ $Z, t_a = 1000, \hat{\omega} = +10.3$ $Z, t_a = 2000, \hat{\omega} = +10.3$ $A, t_a = 3000, \hat{\omega} = -6.5$ $I, t_a = 3052, \hat{\omega} = -5.0$ $I, t_a = 3267, \hat{\omega} = +5.2$ $I, t_a = 5000, \hat{\omega} = -7.4$ $A, t_a = 5028, \hat{\omega} = -4.2$ $Z, t_a = 5916, \hat{\omega} = +5.2$ $I, t_a = 6091, \hat{\omega} = -6.0$ $I, t_a = 6281, \hat{\omega} = -5.6$ $I, t_a = 7000, \hat{\omega} = +11.6$ $Z, t_a = 7087, \hat{\omega} = +5.4$ $I, t_a = 7375, \hat{\omega} = -5.5$ $Z, t_a = 7395, \hat{\omega} = -4.8$ $A, t_a = 7591, \hat{\omega} = +3.9$ $I, t_a = 9620, \hat{\omega} = -5.1$	$Z, t_a = 1000, \hat{\omega} = +10.4$ $Z, t_a = 2000, \hat{\omega} = +10.3$ $Z, t_a = 3000, \hat{\omega} = -7.4$ $I, t_a = 3052, \hat{\omega} = -5.0$ $I, t_a = 3267, \hat{\omega} = +5.2$ $I, t_a = 5000, \hat{\omega} = -7.4$ $Z, t_a = 5028, \hat{\omega} = -5.5$ $Z, t_a = 5916, \hat{\omega} = -5.5$ $I, t_a = 6091, \hat{\omega} = -6.0$ $I, t_a = 6281, \hat{\omega} = -5.6$ $I, t_a = 7000, \hat{\omega} = +11.6$ $Z, t_a = 7087, \hat{\omega} = +5.3$ $I, t_a = 7375, \hat{\omega} = -5.5$ $I, t_a = 9620, \hat{\omega} = -5.1$

Tabelle A.4: Eine Zeitreihe, die dem $ARMA(1,1)$ Prozess nach Gleichung (8.2) folgt und der die aufgelisteten Ausreißer hinzugefügt wurden, ist die Grundlage für diesen Vergleich des Algorithmus von CHEN und LIU mit dem modifizierten Algorithmus von CHEN und LIU.

Hinzugefügte Ausreißer	I, $t_a = 1000$, $\hat{\omega} = +9.0$ Z, $t_a = 2000$, $\hat{\omega} = +10.0$ A, $t_a = 3000$, $\hat{\omega} = -8.0$ Z, $t_a = 5000$, $\hat{\omega} = -10.0$ I, $t_a = 7000$, $\hat{\omega} = +12.0$	
	Algorithmus von CHEN und LIU	Modifizierter Algorithmus
Rechenzeit in Sekunden	134.0	0.4
Geschätzte Parameter	$\hat{\varphi}_1 = +0.61$ $\hat{\varphi}_2 = +0.20$ $\hat{\sigma}_\varepsilon^2 = +1.9576$	$\hat{\varphi}_1 = +0.61$ $\hat{\varphi}_2 = +0.19$ $\hat{\sigma}_\varepsilon^2 = +1.9711$
Geschätzte Ausreißer	Z, $t_a = 137$, $\hat{\omega} = -5.4$ I, $t_a = 671$, $\hat{\omega} = +4.9$ Z, $t_a = 712$, $\hat{\omega} = +5.4$ Z, $t_a = 799$, $\hat{\omega} = +6.0$ <u>Z, $t_a = 1000$, $\hat{\omega} = +6.4$</u> A, $t_a = 1816$, $\hat{\omega} = +4.5$ <u>Z, $t_a = 2000$, $\hat{\omega} = +7.7$</u> Z, $t_a = 2337$, $\hat{\omega} = +5.0$ Z, $t_a = 2393$, $\hat{\omega} = +4.8$ A, $t_a = 2843$, $\hat{\omega} = +4.8$ <u>A, $t_a = 3000$, $\hat{\omega} = -9.2$</u> Z, $t_a = 3359$, $\hat{\omega} = -5.0$ A, $t_a = 3409$, $\hat{\omega} = -4.8$ A, $t_a = 4580$, $\hat{\omega} = -4.1$ <u>Z, $t_a = 5000$, $\hat{\omega} = -9.9$</u> I, $t_a = 5663$, $\hat{\omega} = +5.0$ <u>I, $t_a = 7000$, $\hat{\omega} = +10.6$</u> A, $t_a = 7645$, $\hat{\omega} = +4.4$ A, $t_a = 8171$, $\hat{\omega} = -4.8$ A, $t_a = 8717$, $\hat{\omega} = -4.7$ Z, $t_a = 8915$, $\hat{\omega} = +5.3$ A, $t_a = 9355$, $\hat{\omega} = -4.2$	I, $t_a = 137$, $\hat{\omega} = -5.3$ I, $t_a = 671$, $\hat{\omega} = +4.9$ I, $t_a = 712$, $\hat{\omega} = +5.5$ I, $t_a = 799$, $\hat{\omega} = +5.9$ <u>I, $t_a = 1000$, $\hat{\omega} = +6.4$</u> <u>I, $t_a = 2000$, $\hat{\omega} = +7.3$</u> I, $t_a = 2337$, $\hat{\omega} = +5.0$ I, $t_a = 2844$, $\hat{\omega} = -5.4$ <u>A, $t_a = 3000$, $\hat{\omega} = -9.2$</u> I, $t_a = 3359$, $\hat{\omega} = -4.9$ A, $t_a = 3409$, $\hat{\omega} = -4.8$ <u>I, $t_a = 5000$, $\hat{\omega} = -10.0$</u> I, $t_a = 5663$, $\hat{\omega} = +5.0$ <u>I, $t_a = 7000$, $\hat{\omega} = +10.6$</u> A, $t_a = 8171$, $\hat{\omega} = -4.8$ A, $t_a = 8717$, $\hat{\omega} = -4.7$ I, $t_a = 8915$, $\hat{\omega} = +5.2$

Tabelle A.5: Eine Zeitreihe, die dem $AR(2)$ Prozess nach Gleichung (8.3) folgt und der die aufgelisteten Ausreißer hinzugefügt wurden, ist die Grundlage für diesen Vergleich des Algorithmus von CHEN und LIU mit dem modifizierten Algorithmus von CHEN und LIU.

Hinzugefügte Ausreißer	$I, t_a = 1000, \hat{\omega} = +9.0$ $Z, t_a = 2000, \hat{\omega} = +10.0$ $A, t_a = 3000, \hat{\omega} = -8.0$ $Z, t_a = 5000, \hat{\omega} = -10.0$ $I, t_a = 7000, \hat{\omega} = +12.0$	
	Algorithmus von CHEN und LIU	Modifizierter Algorithmus
Rechenzeit in Sekunden	115.6	0.8
Geschätzte Parameter	$\hat{\theta}_1 = +0.61$ $\hat{\theta}_2 = +0.19$ $\hat{\sigma}_\varepsilon^2 = +1.9772$	$\hat{\theta}_1 = +0.61$ $\hat{\theta}_2 = +0.19$ $\hat{\sigma}_\varepsilon^2 = +1.9972$
Geschätzte Ausreißer	$Z, t_a = 1000, \hat{\omega} = +5.6$ $Z, t_a = 1078, \hat{\omega} = +4.9$ $I, t_a = 1234, \hat{\omega} = -6.2$ $A, t_a = 1985, \hat{\omega} = +4.5$ $I, t_a = 2000, \hat{\omega} = +10.0$ $A, t_a = 2110, \hat{\omega} = +4.3$ $A, t_a = 3000, \hat{\omega} = -7.7$ $I, t_a = 3632, \hat{\omega} = -4.9$ $A, t_a = 4222, \hat{\omega} = +4.3$ $A, t_a = 4352, \hat{\omega} = +4.4$ $Z, t_a = 4838, \hat{\omega} = +6.1$ $Z, t_a = 5000, \hat{\omega} = -11.1$ $Z, t_a = 5125, \hat{\omega} = -5.4$ $Z, t_a = 6022, \hat{\omega} = +5.1$ $Z, t_a = 6366, \hat{\omega} = +4.8$ $Z, t_a = 6646, \hat{\omega} = -4.9$ $I, t_a = 7000, \hat{\omega} = +6.7$ $I, t_a = 7957, \hat{\omega} = -6.0$ $A, t_a = 8079, \hat{\omega} = +4.4$ $Z, t_a = 8407, \hat{\omega} = +4.9$ $Z, t_a = 9337, \hat{\omega} = -5.7$ $A, t_a = 9414, \hat{\omega} = -4.5$	$I, t_a = 1000, \hat{\omega} = +5.8$ $Z, t_a = 1078, \hat{\omega} = +4.9$ $I, t_a = 1234, \hat{\omega} = -6.2$ $Z, t_a = 1985, \hat{\omega} = +4.8$ $I, t_a = 2000, \hat{\omega} = +10.1$ $I, t_a = 3000, \hat{\omega} = -7.0$ $I, t_a = 3001, \hat{\omega} = +6.0$ $I, t_a = 4838, \hat{\omega} = +6.3$ $Z, t_a = 5000, \hat{\omega} = -11.1$ $I, t_a = 5125, \hat{\omega} = -5.0$ $I, t_a = 6022, \hat{\omega} = +5.2$ $Z, t_a = 6366, \hat{\omega} = +4.8$ $Z, t_a = 6646, \hat{\omega} = -4.9$ $I, t_a = 7000, \hat{\omega} = +6.7$ $I, t_a = 7957, \hat{\omega} = -6.0$ $Z, t_a = 8407, \hat{\omega} = +4.9$ $I, t_a = 9337, \hat{\omega} = -5.7$

Tabelle A.6: Eine Zeitreihe, die dem $MA(2)$ Prozess nach Gleichung (8.4) folgt und der die aufgelisteten Ausreißer hinzugefügt wurden, ist die Grundlage für diesen Vergleich des Algorithmus von CHEN und LIU mit dem modifizierten Algorithmus von CHEN und LIU.

A.4 Tabellen zu Abschnitt 8.3

Die nachfolgenden Tabellen sollen Aufschluss über die Qualität der Detektion, Identifikation und Schätzung der Größe eines Ausreißers geben. Dazu sind Zeitreihen der Länge $T = 1000$ mit einem Ausreißer beim Zeitpunkt $t_a = 500$ verunreinigt worden. Es sind jeweils 1000 Zeitreihen für eine im Spaltenkopf der Tabellen angegebenen Größe des Ausreißers erzeugt worden. Für jede dieser 1000 Zeitreihen sind die Prozessparameter und Ausreißer geschätzt worden. Anschließend sind aus diesen 1000 Schätzungen die Mittelwerte und Standardabweichungen der folgenden Größen gebildet worden:

- Prozessparameter,
- Varianz der Residuen σ_ε^2 ,
- Größe $\hat{\omega}$ des hinzugefügten Ausreißers.

Dabei sind zur Bestimmung des Mittelwertes $\mu_{\hat{\omega}}$ und der Standardabweichung $\sigma_{\hat{\omega}}$ der Größe des hinzugefügten Ausreißers nur die Schätzungen verwendet worden, bei denen der Ausreißer detektiert und korrekt identifiziert wurde.

Neben den im oberen Teil der Tabelle aufgeführten Mittelwerten und Standardabweichungen der Schätzungen ist auch festgehalten worden, wie oft bei den 1000 Schätzungen ein Ausreißer beim Zeitpunkt $t_a = 500$ detektiert wurde und welcher Ausreißertyp dabei vorlag. Die daraus resultierende Wahrscheinlichkeit einer Detektion ist in den Zeilen „Gesamt“, „Typ I“, „Typ A“ und „Typ Z“ aufgeführt, wobei I für innovativer Ausreißer, A für additiver Ausreißer und Z für zeitliche Änderung steht. Sind außer dem beim Zeitpunkt $t_a = 500$ hinzugefügten Ausreißer zusätzliche Ausreißer detektiert worden, so wurde deren Anzahl gespeichert. Der Mittelwert der Anzahl dieser zusätzlich detektierten Ausreißer ist in der letzten Zeile der Tabellen aufgeführt.

ω_I	0.0	1.0	2.0	3.0	4.0	5.0	6.0	7.0
$\mu_{\hat{\varphi}_1}$	0.6082	0.6090	0.6085	0.6090	0.6085	0.6107	0.6101	0.6093
$\mu_{\hat{\theta}_1}$	0.2922	0.2914	0.2931	0.2920	0.2926	0.2874	0.2902	0.2892
$\mu_{\hat{\sigma}_\varepsilon^2}$	0.9968	0.9968	0.9981	1.0000	0.9960	0.9939	0.9945	0.9934
$\sigma_{\hat{\varphi}_1}$	0.0326	0.0331	0.0327	0.0327	0.0316	0.0328	0.0346	0.0323
$\sigma_{\hat{\theta}_1}$	0.0428	0.0428	0.0409	0.0412	0.0418	0.0423	0.0422	0.0416
$\sigma_{\hat{\sigma}_\varepsilon^2}$	0.0466	0.0461	0.0470	0.0449	0.0434	0.0459	0.0449	0.0456
$\mu_{\hat{\omega}_I}$	-	3.6863	3.8932	4.1655	4.5359	5.2159	6.1249	7.0694
$\sigma_{\hat{\omega}_I}$	-	0.0906	0.2945	0.5100	0.6711	0.8610	0.9480	0.9792
Detektionswahrscheinlichkeit								
Gesamt	0.001	0.004	0.065	0.297	0.686	0.923	0.993	1.000
Typ I	0.000	0.002	0.041	0.195	0.470	0.669	0.757	0.775
Typ A	0.000	0.002	0.006	0.011	0.008	0.008	0.004	0.001
Typ Z	0.001	0.000	0.018	0.091	0.208	0.246	0.232	0.224
Durchschnittliche Anzahl nicht hinzugefügter Ausreißer								
	0.722	0.665	0.689	0.686	0.643	0.690	0.725	0.777

Tabelle A.7: $ARMA(1,1)$ Zeitreihen der Länge $T = 1000$, die der Prozessgleichung (8.5) folgen, wurden mit einem innovativen Ausreißer beim Zeitpunkt $t_a = 500$ verunreinigt. Für jede im Spaltenkopf angegebene Größe des Ausreißers wurden 1000 Zeitreihen erzeugt und jeweils die Prozessparameter und Ausreißer geschätzt. Aus diesen 1000 Schätzungen für jede Größe des Ausreißers sind die aufgeführten Werte abgeleitet.

ω_A	0.0	1.0	2.0	3.0	4.0	5.0	6.0	7.0
$\mu_{\hat{\varphi}_1}$	0.6102	0.6089	0.6098	0.6087	0.6089	0.6087	0.6098	0.6090
$\mu_{\hat{\theta}_1}$	0.2909	0.2904	0.2893	0.2917	0.2908	0.2920	0.2904	0.2906
$\mu_{\hat{\sigma}_\varepsilon^2}$	0.9953	0.9977	0.9985	0.9992	0.9961	0.9970	0.9975	0.9974
$\sigma_{\hat{\varphi}_1}$	0.0332	0.0334	0.0328	0.0327	0.0339	0.0339	0.0331	0.0333
$\sigma_{\hat{\theta}_1}$	0.0439	0.0426	0.0437	0.0431	0.0427	0.0421	0.0430	0.0433
$\sigma_{\hat{\sigma}_\varepsilon^2}$	0.0450	0.0435	0.0459	0.0445	0.0465	0.0453	0.0448	0.0456
$\mu_{\hat{\omega}_A}$	-	3.0303	3.0198	3.3971	4.0750	5.0426	6.0369	6.9884
$\sigma_{\hat{\omega}_A}$	-	0.4563	0.3392	0.5295	0.6637	0.7154	0.7392	0.6997
Detektionswahrscheinlichkeit								
Gesamt	0.001	0.014	0.189	0.663	0.945	0.999	0.999	1.000
Typ I	0.001	0.004	0.006	0.004	0.000	0.000	0.000	0.000
Typ A	0.000	0.007	0.145	0.584	0.868	0.959	0.978	0.989
Typ Z	0.000	0.003	0.038	0.075	0.077	0.040	0.021	0.011
Durchschnittliche Anzahl nicht hinzugefügter Ausreißer								
	0.74	0.729	0.750	0.766	0.734	0.707	0.710	0.626

Tabelle A.8: $ARMA(1,1)$ Zeitreihen der Länge $T = 1000$, die der Prozessgleichung (8.5) folgen, wurden mit einem additiven Ausreißer beim Zeitpunkt $t_a = 500$ verunreinigt. Für jede im Spaltenkopf angegebene Größe des Ausreißers wurden 1000 Zeitreihen erzeugt und jeweils die Prozessparameter und Ausreißer geschätzt. Aus diesen 1000 Schätzungen für jede Größe des Ausreißers sind die aufgeführten Werte abgeleitet.

ω_Z	0.0	1.0	2.0	3.0	4.0	5.0	6.0	7.0
$\mu_{\hat{\varphi}_1}$	0.6103	0.6100	0.6081	0.6092	0.6091	0.6123	0.6109	0.6092
$\mu_{\hat{\theta}_1}$	0.2906	0.2905	0.2922	0.2922	0.2932	0.2896	0.2909	0.2906
$\mu_{\hat{\sigma}_\varepsilon^2}$	0.9953	0.9924	0.9990	0.9999	0.9977	0.9940	0.9953	0.9956
$\sigma_{\hat{\varphi}_1}$	0.0323	0.0332	0.0328	0.0335	0.0330	0.0328	0.0331	0.0332
$\sigma_{\hat{\theta}_1}$	0.0435	0.0426	0.0432	0.0429	0.0411	0.0436	0.0434	0.0427
$\sigma_{\hat{\sigma}_\varepsilon^2}$	0.0447	0.0455	0.0446	0.0471	0.0438	0.0467	0.0458	0.0441
$\mu_{\hat{\omega}_Z}$	-	3.5579	3.9366	4.2217	4.5258	5.3438	6.0892	7.1259
$\sigma_{\hat{\omega}_Z}$	-	0.1417	0.4144	0.5277	0.6757	0.8624	0.9234	0.9680
Detektionswahrscheinlichkeit								
Gesamt	0.000	0.009	0.077	0.325	0.712	0.931	0.988	0.999
Typ I	0.000	0.003	0.029	0.127	0.236	0.277	0.251	0.193
Typ A	0.000	0.002	0.013	0.048	0.102	0.074	0.045	0.031
Typ Z	0.000	0.004	0.035	0.150	0.374	0.580	0.692	0.775
Durchschnittliche Anzahl nicht hinzugefügter Ausreißer								
	0.719	0.684	0.741	0.690	0.742	0.696	0.717	0.705

Tabelle A.9: $ARMA(1,1)$ Zeitreihen der Länge $T = 1000$, die der Prozessgleichung (8.5) folgen, wurden mit einer zeitlichen Änderung beim Zeitpunkt $t_a = 500$ verunreinigt. Für jede im Spaltenkopf angegebene Größe des Ausreißers wurden 1000 Zeitreihen erzeugt und jeweils die Prozessparameter und Ausreißer geschätzt. Aus diesen 1000 Schätzungen für jede Größe des Ausreißers sind die aufgeführten Werte abgeleitet.

ω_I	0.0	1.0	2.0	3.0	4.0	5.0	6.0	7.0
$\mu_{\hat{\varphi}_1}$	0.6004	0.6017	0.6000	0.6007	0.5987	0.6001	0.6006	0.6017
$\mu_{\hat{\varphi}_2}$	0.1971	0.1971	0.1972	0.1969	0.2004	0.1974	0.1980	0.1975
$\mu_{\hat{\sigma}_\varepsilon^2}$	0.9899	0.9928	0.9951	0.9954	0.9940	0.9906	0.9909	0.9924
$\sigma_{\hat{\varphi}_1}$	0.0316	0.0313	0.0317	0.0303	0.0325	0.0325	0.0306	0.0323
$\sigma_{\hat{\varphi}_2}$	0.0317	0.0310	0.0309	0.0309	0.0323	0.0317	0.0317	0.0321
$\sigma_{\hat{\sigma}_\varepsilon^2}$	0.0453	0.0455	0.0454	0.0450	0.0443	0.0449	0.0454	0.0454
$\mu_{\hat{\omega}_I}$	-	4.0052	3.9818	4.1718	4.5325	5.2167	6.0688	7.0411
$\sigma_{\hat{\omega}_I}$	-	0.4307	0.3232	0.5137	0.7064	0.8732	0.9644	0.9807
Detektionswahrscheinlichkeit								
Gesamt	0.001	0.011	0.075	0.342	0.708	0.940	0.992	1.000
Typ I	0.000	0.005	0.045	0.217	0.552	0.777	0.878	0.917
Typ A	0.000	0.005	0.008	0.063	0.058	0.064	0.041	0.034
Typ Z	0.001	0.001	0.022	0.062	0.098	0.099	0.073	0.049
Durchschnittliche Anzahl nicht hinzugefügter Ausreißer								
	0.835	0.839	0.841	0.776	0.730	0.788	0.845	0.844

Tabelle A.10: $AR(2)$ Zeitreihen der Länge $T = 1000$, die der Prozessgleichung (8.6) folgen, wurden mit einem innovativen Ausreißer beim Zeitpunkt $t_a = 500$ verunreinigt. Für jede im Spaltenkopf angegebene Größe des Ausreißers wurden 1000 Zeitreihen erzeugt und jeweils die Prozessparameter und Ausreißer geschätzt. Aus diesen 1000 Schätzungen für jede Größe des Ausreißers sind die aufgeführten Werte abgeleitet.

ω_A	0.0	1.0	2.0	3.0	4.0	5.0	6.0	7.0
$\mu_{\hat{\varphi}_1}$	0.6006	0.6007	0.5999	0.5974	0.5996	0.6006	0.5993	0.6007
$\mu_{\hat{\varphi}_2}$	0.1980	0.1973	0.1991	0.1990	0.1978	0.1968	0.1977	0.1978
$\mu_{\hat{\sigma}_\varepsilon^2}$	0.9922	0.9918	0.9937	0.9943	0.9932	0.9890	0.9887	0.9918
$\sigma_{\hat{\varphi}_1}$	0.0329	0.0318	0.0316	0.0313	0.0304	0.0317	0.0318	0.0311
$\sigma_{\hat{\varphi}_2}$	0.0336	0.0316	0.0317	0.0320	0.0318	0.0312	0.0317	0.0313
$\sigma_{\hat{\sigma}_\varepsilon^2}$	0.0445	0.0465	0.0452	0.0450	0.0453	0.0458	0.0455	0.0451
$\mu_{\hat{\omega}_A}$	-	3.3538	3.4603	3.7618	4.2633	5.0700	5.9913	7.0120
$\sigma_{\hat{\omega}_A}$	-	0.1751	0.3964	0.5424	0.7136	0.8062	0.8445	0.8478
Detektionswahrscheinlichkeit								
Gesamt	0.001	0.005	0.121	0.479	0.868	0.987	0.999	1.000
Typ I	0.001	0.001	0.027	0.059	0.068	0.052	0.027	0.007
Typ A	0.000	0.004	0.083	0.397	0.779	0.930	0.969	0.993
Typ Z	0.000	0.000	0.011	0.023	0.021	0.005	0.003	0.000
Durchschnittliche Anzahl nicht hinzugefügter Ausreißer								
	0.876	0.836	0.860	0.787	0.742	0.816	0.791	0.795

Tabelle A.11: $AR(2)$ Zeitreihen der Länge $T = 1000$, die der Prozessgleichung (8.6) folgen, wurden mit einem additiven Ausreißer beim Zeitpunkt $t_a = 500$ verunreinigt. Für jede im Spaltenkopf angegebene Größe des Ausreißers wurden 1000 Zeitreihen erzeugt und jeweils die Prozessparameter und Ausreißer geschätzt. Aus diesen 1000 Schätzungen für jede Größe des Ausreißers sind die aufgeführten Werte abgeleitet.

ω_Z	0.0	1.0	2.0	3.0	4.0	5.0	6.0	7.0
$\mu_{\hat{\varphi}_1}$	0.6007	0.6003	0.6012	0.5999	0.6016	0.6008	0.6005	0.6002
$\mu_{\hat{\varphi}_2}$	0.1969	0.1974	0.1964	0.1974	0.1968	0.1968	0.1980	0.1972
$\mu_{\hat{\sigma}_\varepsilon^2}$	0.9917	0.9911	0.9933	0.9945	0.9956	0.9929	0.9935	0.9941
$\sigma_{\hat{\varphi}_1}$	0.0326	0.0306	0.0301	0.0317	0.0324	0.0325	0.0319	0.0317
$\sigma_{\hat{\varphi}_2}$	0.0324	0.0313	0.0305	0.0316	0.0318	0.0317	0.0314	0.0306
$\sigma_{\hat{\sigma}_\varepsilon^2}$	0.0454	0.0447	0.0450	0.0455	0.0438	0.0452	0.0449	0.0451
$\mu_{\hat{\omega}_Z}$	3.5003	-	3.5549	3.6908	4.1375	4.4257	5.2848	6.0030
$\sigma_{\hat{\omega}_Z}$	-	-	0.2896	0.4153	0.6318	0.6947	0.7963	0.9886
Detektionswahrscheinlichkeit								
Gesamt	0.001	0.003	0.076	0.311	0.714	0.924	0.994	0.999
Typ I	0.000	0.003	0.050	0.243	0.611	0.826	0.943	0.971
Typ A	0.000	0.000	0.008	0.023	0.045	0.038	0.017	0.014
Typ Z	0.001	0.000	0.018	0.045	0.058	0.060	0.034	0.014
Durchschnittliche Anzahl nicht hinzugefügter Ausreißer								
	0.828	0.868	0.821	0.789	0.795	0.739	0.809	0.854

Tabelle A.12: $AR(2)$ Zeitreihen der Länge $T = 1000$, die der Prozessgleichung (8.6) folgen, wurden mit einer zeitlichen Änderung beim Zeitpunkt $t_a = 500$ verunreinigt. Für jede im Spaltenkopf angegebene Größe des Ausreißers wurden 1000 Zeitreihen erzeugt und jeweils die Prozessparameter und Ausreißer geschätzt. Aus diesen 1000 Schätzungen für jede Größe des Ausreißers sind die aufgeführten Werte abgeleitet.

ω_I	0.0	1.0	2.0	3.0	4.0	5.0	6.0	7.0
$\mu_{\hat{\theta}_1}$	0.6001	0.5998	0.5998	0.6011	0.6013	0.6006	0.6005	0.5988
$\mu_{\hat{\theta}_2}$	0.1988	0.1990	0.1972	0.1990	0.1999	0.2000	0.2000	0.1979
$\mu_{\hat{\sigma}_\varepsilon^2}$	1.0095	1.0070	1.0138	1.0145	1.0130	1.0095	1.0101	1.0104
$\sigma_{\hat{\theta}_1}$	0.0338	0.0328	0.0319	0.0326	0.0318	0.0328	0.0318	0.0333
$\sigma_{\hat{\theta}_2}$	0.0372	0.0384	0.0381	0.0386	0.0388	0.0381	0.0373	0.0381
$\sigma_{\hat{\sigma}_\varepsilon^2}$	0.0460	0.0475	0.0461	0.0490	0.0454	0.0488	0.0470	0.0458
$\mu_{\hat{\omega}_I}$	-	3.5913	3.9646	4.3124	4.5578	5.2886	6.0894	7.1025
$\sigma_{\hat{\omega}_I}$	-	0.0193	0.3962	0.4910	0.6885	0.8906	0.9610	1.0129
Detektionswahrscheinlichkeit								
Gesamt	0.000	0.007	0.070	0.306	0.682	0.932	0.989	1.000
Typ I	0.000	0.002	0.038	0.182	0.460	0.674	0.788	0.850
Typ A	0.000	0.001	0.012	0.035	0.077	0.071	0.039	0.034
Typ Z	0.000	0.004	0.020	0.089	0.145	0.187	0.162	0.116
Durchschnittliche Anzahl nicht hinzugefügter Ausreißer								
	0.718	0.664	0.721	0.660	0.714	0.658	0.728	0.691

Tabelle A.13: $MA(2)$ Zeitreihen der Länge $T = 1000$, die der Prozessgleichung (8.7) folgen, wurden mit einem innovativen Ausreißer beim Zeitpunkt $t_a = 500$ verunreinigt. Für jede im Spaltenkopf angegebene Größe des Ausreißers wurden 1000 Zeitreihen erzeugt und jeweils die Prozessparameter und Ausreißer geschätzt. Aus diesen 1000 Schätzungen für jede Größe des Ausreißers sind die aufgeführten Werte abgeleitet.

ω_A	0.0	1.0	2.0	3.0	4.0	5.0	6.0	7.0
$\mu_{\hat{\theta}_1}$	0.6003	0.5988	0.5968	0.5972	0.5997	0.5998	0.6002	0.5993
$\mu_{\hat{\theta}_2}$	0.1979	0.1982	0.1981	0.2003	0.1993	0.1983	0.1994	0.1995
$\mu_{\hat{\sigma}_\varepsilon^2}$	1.0095	1.0136	1.0154	1.0129	1.0109	1.0114	1.0099	1.0133
$\sigma_{\hat{\theta}_1}$	0.0320	0.0327	0.0329	0.0346	0.0337	0.0326	0.0334	0.0343
$\sigma_{\hat{\theta}_2}$	0.0370	0.0381	0.0392	0.04010	0.0389	0.0381	0.0394	0.0407
$\sigma_{\hat{\sigma}_\varepsilon^2}$	0.0474	0.0478	0.0466	0.0461	0.0477	0.0477	0.0465	0.0470
$\mu_{\hat{\omega}_A}$	-	3.3726	3.5123	3.7734	4.2870	5.0743	5.9801	6.9903
$\sigma_{\hat{\omega}_A}$	-	0.2037	0.3620	0.5283	0.6664	0.7948	0.8810	0.8478
Detektionswahrscheinlichkeit								
Gesamt	0.001	0.008	0.112	0.463	0.848	0.982	0.999	1.000
Typ I	0.000	0.000	0.018	0.054	0.060	0.060	0.030	0.013
Typ A	0.000	0.007	0.082	0.388	0.764	0.907	0.963	0.985
Typ Z	0.001	0.001	0.012	0.021	0.024	0.015	0.006	0.002
Durchschnittliche Anzahl nicht hinzugefügter Ausreißer								
	0.708	0.719	0.676	0.687	0.742	0.713	0.785	0.762

Tabelle A.14: $MA(2)$ Zeitreihen der Länge $T = 1000$, die der Prozessgleichung (8.7) folgen, wurden mit einem additiven Ausreißer beim Zeitpunkt $t_a = 500$ verunreinigt. Für jede im Spaltenkopf angegebene Größe des Ausreißers wurden 1000 Zeitreihen erzeugt und jeweils die Prozessparameter und Ausreißer geschätzt. Aus diesen 1000 Schätzungen für jede Größe des Ausreißers sind die aufgeführten Werte abgeleitet.

ω_Z	0.0	1.0	2.0	3.0	4.0	5.0	6.0	7.0
$\mu_{\hat{\theta}_1}$	0.6007	0.5982	0.6004	0.5999	0.5976	0.5997	0.6000	0.5990
$\mu_{\hat{\theta}_2}$	0.1994	0.1988	0.1995	0.1998	0.1987	0.1985	0.1973	0.1998
$\mu_{\hat{\sigma}_\varepsilon^2}$	1.0108	1.0111	1.0146	1.0138	1.0148	1.0058	1.0081	1.0101
$\sigma_{\hat{\theta}_1}$	0.0344	0.0327	0.0328	0.0335	0.0334	0.0337	0.0306	0.0339
$\sigma_{\hat{\theta}_2}$	0.0376	0.0370	0.0387	0.0385	0.0369	0.0393	0.0385	0.0398
$\sigma_{\hat{\sigma}_\varepsilon^2}$	0.0478	0.0476	0.0485	0.0471	0.0464	0.0464	0.0455	0.0479
$\mu_{\hat{\omega}_Z}$	0.0157	3.9116	3.9060	4.1705	4.5017	5.1552	6.0367	7.0267
$\sigma_{\hat{\omega}_Z}$	5.1393	0.5276	0.4152	0.5624	0.7011	0.8617	0.9069	0.9567
Detektionswahrscheinlichkeit								
Gesamt	0.002	0.012	0.075	0.351	0.721	0.949	0.997	1.000
Typ I	0.000	0.003	0.027	0.101	0.161	0.175	0.130	0.102
Typ A	0.000	0.004	0.005	0.032	0.033	0.023	0.013	0.007
Typ Z	0.002	0.005	0.043	0.218	0.527	0.751	0.854	0.891
Durchschnittliche Anzahl nicht hinzugefügter Ausreißer								
	0.687	0.697	0.719	0.668	0.689	0.696	0.714	0.718

Tabelle A.15: $MA(2)$ Zeitreihen der Länge $T = 1000$, die der Prozessgleichung (8.7) folgen, wurden mit einer zeitlichen Änderung beim Zeitpunkt $t_a = 500$ verunreinigt. Für jede im Spaltenkopf angegebene Größe des Ausreißers wurden 1000 Zeitreihen erzeugt und jeweils die Prozessparameter und Ausreißer geschätzt. Aus diesen 1000 Schätzungen für jede Größe des Ausreißers sind die aufgeführten Werte abgeleitet.

A.5 Tabellen zu Abschnitt 8.4

Die in den Tabellen dargestellten Ergebnisse sind aus den Schätzungen von 1000 Zeitreihen der Länge $T = 1000$ ermittelt. Jede der Zeitreihen wurde mit zwei Ausreißern verunreinigt, wobei der erste Ausreißer immer beim Zeitpunkt $t_{a_1} = 500$ eintritt und der Zeitpunkt t_{a_2} des Eintretens des zweiten Ausreißers in der Tabelle angegeben wird. Eine Tabelle enthält Ergebnisse für mehrere Zeitpunkte t_{a_2} . Der Ausreißertyp der beiden Ausreißer wird variiert, so dass alle möglichen Kombinationen vorkommen. Jede Zeile steht für eine dieser Kombinationen. Welche betrachtet wird, ist in den ersten beiden Spalten der Tabellen angegeben. Die Abkürzungen I, A und Z stehen für innovativer Ausreißer, additiver Ausreißer und zeitliche Änderung.

Tabellen zur Qualität der Identifikation: Zunächst werden die Tabellen erläutert, mit deren Hilfe man Aussagen über die Qualität der Identifikation des Ausreißertyps machen kann. Hier werden pro Tabelle drei verschiedene Zeitpunkte t_{a_2} betrachtet. Für die in jeder Zeile der Tabelle dargestellten Ergebnisse wurden 1000 Zeitreihen mit den Ausreißern entsprechenden Typs erzeugt und mit dem modifizierten Algorithmus von CHEN und LIU ausgewertet. Ist dabei beim Zeitpunkt t_{a_1} oder t_{a_2} ein Ausreißer detektiert worden, so wurde der Typ dieses Ausreißers gespeichert. Daraus lässt sich die Wahrscheinlichkeit ableiten, einen bestimmten Ausreißertyp bei den Zeitpunkten t_{a_1} und t_{a_2} zu detektieren. In der vierten bis sechsten Spalte sind diese Wahrscheinlichkeiten für den Zeitpunkt t_{a_1} und in der achten bis zehnten Spalte für den Zeitpunkt t_{a_2} aufgeführt. So gibt beispielsweise die fünfte Spalte der Tabellen an, wie wahrscheinlich es ist, beim Zeitpunkt t_{a_1} einen additiven Ausreißer zu detektieren. Summiert man die Wahrscheinlichkeiten der verschiedenen Ausreißertypen für einen Zeitpunkt auf, erhält man die Wahrscheinlichkeit, dass ein Ausreißer beliebigen Typs vorliegt, was der Wahrscheinlichkeit einer Detektion eines Ausreißers entspricht. Sie ist für den ersten Zeitpunkt t_{a_1} in der dritten Spalte und für den zweiten Zeitpunkt t_{a_2} in der siebten Spalte aufgeführt.

Beispiel zum Lesen der Tabellen: Es wird die erste Zeile für den Zeitpunkt $t_{a_2} = 501$ in Tabelle A.16, in der die Kombination zweier innovativer Ausreißer betrachtet wird, erklärt. In der dritten bis sechsten Spalte stehen die Werte:

$$0.998, 0.926, 0.000, 0.072.$$

Der Wert 0.998 bedeutet, dass bei 998 der 1000 Zeitreihen ein Ausreißer beim Zeitpunkt t_{a_1} detektiert wurde. Der zweite Wert 0.926 gibt an, dass bei 926 der 1000 Zeitreihen ein innovativer Ausreißer beim Zeitpunkt t_{a_1} detektiert wurde. Der dritte Werte 0.000 zeigt an, dass bei keiner der 1000 Zeitreihen ein additiver Ausreißer beim Zeitpunkt t_{a_1} detektiert wurde. Und schließlich bedeutet der vierte Wert 0.072, dass bei 72 der 1000 Zeitreihen beim Zeitpunkt t_{a_1} eine zeitliche Änderung detektiert wurde.

Tabellen zur Qualität der Schätzung der Größe der Ausreißer: Nun werden die Tabellen erläutert, die Auskunft über die Qualität der Schätzung der Größe der Ausreißer geben. Hier werden sechs verschiedene Zeitpunkte t_{a_2} in jeder Tabelle betrachtet. Die dritte bis sechste Spalte führt die Ergebnisse für die Zeitpunkte $t_{a_2} = 501, 503$ und 505 auf, während die siebte bis zehnte Spalte die Ergebnisse für die Zeitpunkte $t_{a_2} = 502, 504$ und 510 enthält. Aus der geschätzten Größe der Ausreißer, die beim ersten Zeitpunkt t_{a_1} detektiert wurden, sind der Mittelwert $\mu_{\hat{\omega}_1}$ und die Standardabweichung $\sigma_{\hat{\omega}_1}$ ermittelt worden. Dabei wurden nur die Ausreißer verwendet, die korrekt identifiziert wurden. Die Mittelwerte $\mu_{\hat{\omega}_1}$ und Standardabweichungen $\sigma_{\hat{\omega}_1}$ ergeben sich analog für den Zeitpunkt t_{a_2} . Um zu erfahren, aus wie vielen Werten ein Mittelwert und die zugehörige Standardabweichung ermittelt wurden, muss man die Tabellen über die Qualität der Identifikation der Ausreißer zur Hilfe nehmen.

Beispiel zum Lesen der Tabellen: Es wird Tabelle A.18 betrachtet. Als Beispiel soll der Fall, bei dem der Zeitpunkt des zweiten Ausreißers $t_{a_2} = 501$ ist und zwei innovative Ausreißer hinzugefügt wurden, dienen. Der Mittelwert und die Standardabweichung des ersten Ausreißers beim Zeitpunkt t_{a_1} betragen hier $\mu_{\hat{\omega}_1} = 7.0463$ und $\sigma_{\hat{\omega}_1} = 0.9638$. Um zu erfahren, aus wie vielen Werten $\mu_{\hat{\omega}_1}$ und $\sigma_{\hat{\omega}_1}$ ermittelt wurden, muss man in Tabelle A.16 nachsehen. Dort findet man bei dem Fall, dass der zweite Ausreißer beim Zeitpunkt $t_{a_2} = 501$ eintritt, in der Zeile, die die Ergebnisse für die Kombination zweier innovativer Ausreißer enthält, in der vierten Spalte den Wert 0.926. Dies bedeutet, dass bei 926 der 1000 Zeitreihen beim Zeitpunkt t_{a_1} ein innovativer Ausreißer detektiert wurde. Der Ausreißer wurde also 926 mal korrekt identifiziert, so dass der Mittelwert $\mu_{\hat{\omega}_1} = 7.0463$ und die Standardabweichung $\sigma_{\hat{\omega}_1} = 0.9638$ aus 926 Schätzungen der Größe $\hat{\omega}_1$ eines innovativen Ausreißers beim Zeitpunkt t_{a_1} ermittelt wurden.

		Detektierter Typ bei t_{a_1}				Detektierter Typ bei t_{a_2}			
Typ 1	Typ 2	Gesamt	Typ I	Typ A	Typ Z	Gesamt	Typ I	Typ A	Typ Z
$t_{a_2} = 501$									
I	I	0.998	0.926	0.000	0.072	1.000	0.862	0.000	0.138
I	A	0.997	0.973	0.000	0.024	0.999	0.003	0.964	0.032
I	Z	0.999	0.973	0.000	0.026	1.000	0.376	0.013	0.611
A	I	0.999	0.682	0.000	0.317	0.004	0.004	0.000	0.000
A	A	0.999	0.995	0.000	0.004	0.043	0.003	0.037	0.003
A	Z	0.999	0.870	0.000	0.129	0.004	0.003	0.001	0.000
Z	I	0.999	0.932	0.000	0.067	0.975	0.893	0.000	0.082
Z	A	1.000	0.944	0.000	0.056	0.984	0.048	0.736	0.200
Z	Z	0.999	0.976	0.000	0.023	0.986	0.501	0.003	0.482
$t_{a_2} = 502$									
I	I	0.996	0.271	0.011	0.714	0.997	0.795	0.001	0.201
I	A	0.997	0.801	0.004	0.192	0.997	0.012	0.944	0.041
I	Z	0.999	0.335	0.007	0.657	0.996	0.187	0.037	0.772
A	I	0.997	0.219	0.762	0.016	0.732	0.640	0.002	0.090
A	A	0.996	0.451	0.531	0.014	0.662	0.009	0.653	0.000
A	Z	0.994	0.338	0.642	0.014	0.619	0.260	0.043	0.316
Z	I	0.997	0.189	0.080	0.728	0.997	0.811	0.000	0.186
Z	A	0.998	0.344	0.042	0.612	0.995	0.004	0.953	0.038
Z	Z	0.997	0.175	0.094	0.728	0.998	0.232	0.025	0.741
$t_{a_2} = 503$									
I	I	1.000	0.753	0.000	0.247	1.000	0.780	0.000	0.220
I	A	0.998	0.819	0.002	0.177	1.000	0.004	0.955	0.041
I	Z	0.999	0.774	0.000	0.225	0.999	0.229	0.024	0.746
A	I	1.000	0.001	0.914	0.085	0.999	0.817	0.000	0.182
A	A	0.999	0.000	0.988	0.011	1.000	0.001	0.991	0.008
A	Z	1.000	0.003	0.911	0.086	0.999	0.249	0.040	0.710
Z	I	0.999	0.213	0.006	0.780	1.000	0.817	0.000	0.183
Z	A	0.997	0.308	0.027	0.662	1.000	0.006	0.942	0.052
Z	Z	1.000	0.196	0.013	0.791	1.000	0.238	0.007	0.755

Tabelle A.16: Dargestellt sind die Wahrscheinlichkeiten einen bestimmten Ausreißertyp bei den Zeitpunkten t_{a_1} und t_{a_2} zu detektieren, woraus man Aussagen über die Qualität der Identifikation des Ausreißertyps ableiten kann. Dabei folgen die der Berechnung zu Grunde gelegten Zeitreihen dem $ARMA(1,1)$ Prozess nach Gleichung (8.8) und sind mit zwei Ausreißern der Größe $\omega = 7.0$ bei den Zeitpunkten t_{a_1} und t_{a_2} verunreinigt worden.

		Typ 1				Typ 2			
Typ 1	Typ 2	Gesamt	Typ I	Typ A	Typ Z	Gesamt	Typ I	Typ A	Typ Z
$t_{a_2} = 504$									
I	I	0.998	0.677	0.000	0.321	1.000	0.812	0.001	0.187
I	A	0.997	0.849	0.001	0.147	1.000	0.006	0.965	0.029
I	Z	0.999	0.666	0.000	0.333	1.000	0.224	0.014	0.762
A	I	0.999	0.006	0.942	0.051	0.997	0.823	0.001	0.173
A	A	1.000	0.000	0.994	0.006	0.999	0.000	0.996	0.003
A	Z	1.000	0.008	0.937	0.055	0.999	0.308	0.053	0.638
Z	I	0.999	0.176	0.008	0.815	0.998	0.792	0.000	0.206
Z	A	0.995	0.311	0.040	0.644	1.000	0.004	0.951	0.045
Z	Z	0.998	0.126	0.022	0.850	0.999	0.198	0.018	0.783
$t_{a_2} = 505$									
I	I	0.998	0.755	0.000	0.243	0.999	0.804	0.000	0.195
I	A	0.997	0.794	0.002	0.201	1.000	0.005	0.955	0.040
I	Z	0.999	0.740	0.001	0.258	0.999	0.215	0.021	0.763
A	I	1.000	0.003	0.955	0.042	0.999	0.815	0.000	0.184
A	A	0.999	0.000	0.995	0.004	1.000	0.000	0.996	0.004
A	Z	1.000	0.004	0.959	0.037	0.996	0.307	0.034	0.655
Z	I	0.998	0.196	0.015	0.787	0.997	0.810	0.000	0.187
Z	A	0.999	0.310	0.049	0.640	1.000	0.004	0.954	0.042
Z	Z	0.999	0.165	0.014	0.820	0.997	0.216	0.014	0.767
$t_{a_2} = 510$									
I	I	1.000	0.808	0.000	0.192	0.999	0.796	0.000	0.203
I	A	0.998	0.826	0.002	0.170	1.000	0.004	0.961	0.035
I	Z	1.000	0.766	0.003	0.231	0.999	0.233	0.013	0.753
A	I	1.000	0.005	0.956	0.039	0.998	0.817	0.001	0.180
A	A	1.000	0.000	0.996	0.004	1.000	0.001	0.996	0.003
A	Z	1.000	0.003	0.951	0.046	0.998	0.318	0.040	0.640
Z	I	0.999	0.209	0.015	0.775	0.998	0.784	0.000	0.214
Z	A	0.996	0.326	0.035	0.635	0.999	0.003	0.969	0.027
Z	Z	1.000	0.192	0.019	0.789	1.000	0.236	0.020	0.744

Tabelle A.17: Dargestellt sind die Wahrscheinlichkeiten einen bestimmten Ausreißertyp bei den Zeitpunkten t_{a_1} und t_{a_2} zu detektieren, woraus man Aussagen über die Qualität der Identifikation des Ausreißertyps ableiten kann. Dabei folgen die der Berechnung zu Grunde gelegten Zeitreihen dem $ARMA(1,1)$ Prozess nach Gleichung (8.8) und sind mit zwei Ausreißern der Größe $\omega = 7.0$ bei den Zeitpunkten t_{a_1} und t_{a_2} verunreinigt worden.

Typ 1	Typ 2	$\mu_{\hat{\omega}_1}$	$\sigma_{\hat{\omega}_1}$	$\mu_{\hat{\omega}_2}$	$\sigma_{\hat{\omega}_2}$	$\mu_{\hat{\omega}_1}$	$\sigma_{\hat{\omega}_1}$	$\mu_{\hat{\omega}_2}$	$\sigma_{\hat{\omega}_2}$
		$t_{a_2} = 501$				$t_{a_2} = 502$			
I	I	7.0463	0.9638	7.0848	1.0566	6.3300	0.7200	6.3810	1.1881
I	A	7.0456	1.0069	7.0860	0.7493	6.9740	0.9255	6.9602	0.7410
I	Z	7.0290	0.9930	7.2501	1.1291	6.3505	0.8087	6.5580	1.1963
A	I	-	-	3.9056	0.2021	7.1953	1.0875	6.9764	1.6487
A	A	-	-	3.3738	0.7963	7.1072	2.1374	6.7666	1.9733
A	Z	-	-	-	-	7.2636	1.3518	6.5911	0.7594
Z	I	6.1359	0.7343	5.7925	1.0848	7.1300	0.9765	7.0840	1.1902
Z	A	6.6192	0.8422	6.1356	0.7011	6.9542	0.9250	7.1461	0.7648
Z	Z	5.8589	0.9111	5.5249	1.0235	7.1942	0.9890	7.1387	1.1178
		$t_{a_2} = 503$				$t_{a_2} = 504$			
I	I	6.9959	0.9911	7.0464	0.9901	6.9450	0.9832	6.9569	1.0155
I	A	7.0217	0.9888	7.0374	0.6943	7.0596	1.0396	7.0417	0.6823
I	Z	7.0366	0.9921	7.0169	0.9804	6.9178	0.9905	7.0054	0.9669
A	I	7.0764	0.7197	6.9202	0.9798	7.0244	0.7108	7.1218	0.9770
A	A	6.9939	0.7670	7.0500	0.7698	6.9961	0.6976	7.0255	0.7117
A	Z	7.0175	0.6982	7.0090	0.9302	7.0350	0.6942	6.9392	0.8801
Z	I	7.0958	0.9719	7.0341	1.0551	7.0615	0.9816	7.1244	0.9767
Z	A	6.9389	0.9343	7.0495	0.7353	6.9375	0.9375	7.0678	0.6778
Z	Z	7.0402	0.9871	7.0694	0.9567	7.0579	0.9526	7.1005	0.9847
		$t_{a_2} = 505$				$t_{a_2} = 510$			
I	I	7.0139	0.9386	7.0093	1.0385	7.0211	0.9745	6.9601	1.0150
I	A	7.1092	1.0028	7.0438	0.6842	7.0987	1.0223	7.0353	0.7114
I	Z	7.0125	0.9776	7.0462	0.9932	7.0199	0.9744	6.9852	1.0196
A	I	7.0395	0.6925	7.0900	0.9945	7.0038	0.7258	7.0970	1.0039
A	A	6.9692	0.7346	7.0061	0.7245	7.0099	0.7259	6.9803	0.7400
A	Z	7.0526	0.7209	6.9829	0.8822	7.0414	0.6882	6.9300	0.8990
Z	I	7.0783	0.9578	7.0581	0.9747	7.0405	0.9767	7.0553	1.0072
Z	A	6.9897	0.9275	7.0629	0.7426	6.8806	0.8895	7.0307	0.7012
Z	Z	7.0424	0.9338	7.0376	0.9647	7.0661	0.9685	7.0640	0.9400

Tabelle A.18: Dargestellt sind die Mittelwerte und Standardabweichungen der geschätzten Größe der korrekt identifizierten Ausreißer bei den Zeitpunkten t_{a_1} und t_{a_2} , woraus man Aussagen über die Qualität der Schätzung der Größe der Ausreißer ableiten kann. Dabei folgen die der Berechnung zu Grunde gelegten Zeitreihen dem $ARMA(1,1)$ Prozess nach Gleichung (8.8) und sind mit zwei Ausreißern der Größe $\omega = 7.0$ bei den Zeitpunkten t_{a_1} und t_{a_2} verunreinigt worden.

		Typ 1				Typ 2			
Typ 1	Typ 2	Gesamt	Typ I	Typ A	Typ Z	Gesamt	Typ I	Typ A	Typ Z
$t_{a_2} = 501$									
I	I	0.999	0.784	0.000	0.215	0.999	0.945	0.011	0.043
I	A	0.999	0.893	0.002	0.104	1.000	0.210	0.783	0.007
I	Z	0.999	0.976	0.000	0.023	1.000	0.981	0.004	0.015
A	I	0.998	0.757	0.000	0.241	0.288	0.136	0.142	0.010
A	A	1.000	0.731	0.053	0.216	0.286	0.086	0.174	0.026
A	Z	1.000	0.980	0.000	0.020	0.292	0.182	0.099	0.011
Z	I	0.999	0.937	0.000	0.062	1.000	0.905	0.085	0.010
Z	A	0.999	0.957	0.024	0.018	1.000	0.042	0.957	0.001
Z	Z	0.998	0.995	0.000	0.003	1.000	0.944	0.056	0.000
$t_{a_2} = 502$									
I	I	1.000	0.953	0.020	0.027	1.000	0.955	0.007	0.038
I	A	1.000	0.951	0.029	0.020	1.000	0.185	0.814	0.001
I	Z	1.000	0.962	0.019	0.019	0.998	0.979	0.007	0.012
A	I	0.998	0.561	0.437	0.000	0.649	0.561	0.057	0.031
A	A	1.000	0.246	0.754	0.000	0.983	0.050	0.932	0.001
A	Z	0.999	0.582	0.417	0.000	0.622	0.583	0.028	0.011
Z	I	1.000	0.992	0.002	0.006	0.995	0.908	0.068	0.019
Z	A	0.998	0.988	0.008	0.002	0.997	0.213	0.782	0.002
Z	Z	1.000	0.998	0.002	0.000	0.997	0.948	0.045	0.004
$t_{a_2} = 503$									
I	I	1.000	0.113	0.010	0.877	0.920	0.584	0.330	0.006
I	A	0.998	0.795	0.038	0.165	0.983	0.157	0.818	0.008
I	Z	1.000	0.141	0.005	0.854	0.881	0.615	0.263	0.003
A	I	1.000	0.073	0.567	0.360	0.899	0.758	0.090	0.051
A	A	1.000	0.013	0.971	0.016	0.999	0.009	0.990	0.000
A	Z	1.000	0.070	0.583	0.347	0.906	0.835	0.054	0.017
Z	I	0.998	0.074	0.001	0.923	0.814	0.437	0.372	0.005
Z	A	1.000	0.848	0.013	0.139	0.957	0.161	0.784	0.012
Z	Z	1.000	0.084	0.002	0.914	0.793	0.486	0.304	0.003

Tabelle A.19: Dargestellt sind die Wahrscheinlichkeiten einen bestimmten Ausreißertyp bei den Zeitpunkten t_{a_1} und t_{a_2} zu detektieren, woraus man Aussagen über die Qualität der Identifikation des Ausreißertyps ableiten kann. Dabei folgen die der Berechnung zu Grunde gelegten Zeitreihen dem $AR(2)$ Prozess nach Gleichung (8.9) und sind mit zwei Ausreißern der Größe $\omega = 7.0$ bei den Zeitpunkten t_{a_1} und t_{a_2} verunreinigt worden.

		Typ 1				Typ 2			
Typ 1	Typ 2	Gesamt	Typ I	Typ A	Typ Z	Gesamt	Typ I	Typ A	Typ Z
$t_{a_2} = 504$									
I	I	0.999	0.335	0.015	0.649	0.888	0.719	0.147	0.022
I	A	0.999	0.882	0.023	0.094	0.998	0.147	0.847	0.004
I	Z	1.000	0.349	0.011	0.640	0.869	0.768	0.091	0.010
A	I	1.000	0.130	0.764	0.106	0.991	0.903	0.053	0.035
A	A	1.000	0.008	0.988	0.004	1.000	0.015	0.984	0.001
A	Z	1.000	0.113	0.789	0.098	0.993	0.956	0.029	0.008
Z	I	1.000	0.329	0.008	0.663	0.767	0.581	0.170	0.016
Z	A	1.000	0.933	0.010	0.057	0.992	0.186	0.797	0.009
Z	Z	0.999	0.331	0.004	0.664	0.749	0.620	0.126	0.003
$t_{a_2} = 505$									
I	I	0.999	0.551	0.012	0.436	0.990	0.883	0.066	0.041
I	A	1.000	0.909	0.025	0.066	1.000	0.146	0.850	0.004
I	Z	1.000	0.562	0.020	0.418	0.985	0.947	0.027	0.011
A	I	1.000	0.131	0.828	0.041	1.000	0.951	0.026	0.023
A	A	1.000	0.018	0.978	0.004	1.000	0.014	0.986	0.000
A	Z	1.000	0.144	0.836	0.020	1.000	0.980	0.005	0.015
Z	I	1.000	0.676	0.010	0.314	0.979	0.874	0.072	0.033
Z	A	0.999	0.955	0.008	0.036	0.996	0.184	0.804	0.008
Z	Z	0.999	0.659	0.008	0.332	0.978	0.923	0.037	0.018
$t_{a_2} = 510$									
I	I	0.999	0.932	0.012	0.055	0.999	0.945	0.018	0.036
I	A	1.000	0.951	0.027	0.022	1.000	0.140	0.858	0.002
I	Z	0.998	0.923	0.014	0.061	0.999	0.977	0.010	0.012
A	I	1.000	0.141	0.855	0.004	0.999	0.952	0.021	0.026
A	A	1.000	0.012	0.988	0.000	1.000	0.008	0.991	0.001
A	Z	1.000	0.128	0.868	0.004	0.996	0.981	0.007	0.008
Z	I	1.000	0.969	0.007	0.024	0.999	0.928	0.021	0.050
Z	A	0.999	0.975	0.013	0.011	0.999	0.145	0.849	0.005
Z	Z	1.000	0.976	0.007	0.017	0.999	0.967	0.012	0.020

Tabelle A.20: Dargestellt sind die Wahrscheinlichkeiten einen bestimmten Ausreißertyp bei den Zeitpunkten t_{a_1} und t_{a_2} zu detektieren, woraus man Aussagen über die Qualität der Identifikation des Ausreißertyps ableiten kann. Dabei folgen die der Berechnung zu Grunde gelegten Zeitreihen dem $AR(2)$ Prozess nach Gleichung (8.9) und sind mit zwei Ausreißern der Größe $\omega = 7.0$ bei den Zeitpunkten t_{a_1} und t_{a_2} verunreinigt worden.

Typ 1	Typ 2	$\mu_{\hat{\omega}_1}$	$\sigma_{\hat{\omega}_1}$	$\mu_{\hat{\omega}_2}$	$\sigma_{\hat{\omega}_2}$	$\mu_{\hat{\omega}_1}$	$\sigma_{\hat{\omega}_1}$	$\mu_{\hat{\omega}_2}$	$\sigma_{\hat{\omega}_2}$
		$t_{a_2} = 501$				$t_{a_2} = 502$			
I	I	6.8719	1.0140	6.8554	1.0912	7.0723	0.9567	7.0867	1.0110
I	A	6.9831	0.9771	7.2140	0.7439	7.0575	1.0014	7.2589	0.8501
I	Z	7.0318	1.0156	5.9843	0.8763	7.0320	0.9981	6.2322	0.8331
A	I	-	-	4.1861	0.5426	7.2370	0.9081	6.5534	1.0308
A	A	6.1393	0.6690	5.4321	1.7142	6.9369	0.8467	6.7957	1.0163
A	Z	-	-	3.4600	0.2740	7.2402	0.9050	5.3450	0.7271
Z	I	6.6706	0.8476	7.7142	1.0460	5.4025	0.2054	6.1327	0.9736
Z	A	6.3617	1.3402	8.0481	0.9344	6.8430	1.2042	6.8552	0.8193
Z	Z	5.7716	0.7129	-	-	-	-	5.6072	1.2773
		$t_{a_2} = 503$				$t_{a_2} = 504$			
I	I	5.8992	0.7899	5.8162	1.3638	6.3400	0.9793	6.4633	1.3213
I	A	6.8979	0.9421	7.1460	0.8342	6.9928	0.9593	7.1424	0.8064
I	Z	5.8290	0.7935	6.4775	1.9775	6.4593	0.9182	5.5592	1.2456
A	I	7.1270	0.8410	6.4107	1.3373	7.1101	0.8484	6.9932	1.0940
A	A	6.9879	0.8518	6.9989	0.8218	7.0321	0.8667	6.9923	0.8558
A	Z	7.2139	0.8530	4.3782	0.9888	7.1719	0.8022	5.8662	0.5568
Z	I	6.6887	1.0238	5.2625	1.1517	6.6717	1.0292	6.1611	1.2361
Z	A	7.4157	0.9862	6.9363	0.8658	6.9696	0.9128	6.9937	0.8187
Z	Z	6.5845	1.0297	5.1552	1.4806	6.6915	1.0876	5.6318	0.9331
		$t_{a_2} = 505$				$t_{a_2} = 510$			
I	I	6.7534	1.0249	6.6551	1.1852	7.0932	1.0082	7.0886	1.0119
I	A	6.9924	0.9560	7.0963	0.8391	7.0428	1.0289	7.0879	0.8334
I	Z	6.8352	1.0569	5.8515	1.0801	7.0282	0.9275	5.7393	0.7449
A	I	7.1146	0.7859	6.9581	1.0427	7.1146	0.8263	7.0444	1.0145
A	A	7.0565	0.8497	7.0320	0.8344	6.9761	0.8388	6.9808	0.8152
A	Z	7.1301	0.7925	5.6986	0.7993	7.1578	0.8302	5.2602	1.2251
Z	I	6.3401	0.9120	6.4820	1.1599	6.2230	0.8830	7.0188	0.9931
Z	A	6.5934	0.9432	7.0199	0.7799	5.5144	0.9992	7.1246	0.7987
Z	Z	6.4292	0.9505	5.1041	0.8781	6.1158	0.7276	6.2123	1.0208

Tabelle A.21: Dargestellt sind die Mittelwerte und Standardabweichungen der geschätzten Größe der korrekt identifizierten Ausreißer bei den Zeitpunkten t_{a_1} und t_{a_2} , woraus man Aussagen über die Qualität der Schätzung der Größe der Ausreißer ableiten kann. Dabei folgen die der Berechnung zu Grunde gelegten Zeitreihen dem $AR(2)$ Prozess nach Gleichung (8.9) und sind mit zwei Ausreißern der Größe $\omega = 7.0$ bei den Zeitpunkten t_{a_1} und t_{a_2} verunreinigt worden.

		Typ 1				Typ 2			
Typ 1	Typ 2	Gesamt	Typ I	Typ A	Typ Z	Gesamt	Typ I	Typ A	Typ Z
$t_{a_2} = 501$									
I	I	1.000	0.544	0.000	0.456	0.998	0.936	0.019	0.043
I	A	0.999	0.905	0.000	0.094	0.998	0.171	0.812	0.015
I	Z	0.999	0.316	0.000	0.683	0.992	0.544	0.003	0.445
A	I	1.000	0.319	0.000	0.681	0.124	0.112	0.012	0.000
A	A	1.000	0.938	0.000	0.062	0.404	0.167	0.203	0.034
A	Z	1.000	0.008	0.000	0.992	0.089	0.033	0.002	0.054
Z	I	1.000	0.357	0.000	0.643	0.998	0.621	0.002	0.375
Z	A	1.000	0.390	0.000	0.610	1.000	0.267	0.557	0.176
Z	Z	0.999	0.125	0.000	0.874	1.000	0.103	0.000	0.897
$t_{a_2} = 502$									
I	I	0.999	0.162	0.030	0.807	0.964	0.693	0.251	0.020
I	A	1.000	0.663	0.046	0.291	0.991	0.163	0.806	0.022
I	Z	1.000	0.142	0.010	0.848	0.953	0.292	0.166	0.495
A	I	1.000	0.137	0.797	0.066	0.953	0.845	0.019	0.089
A	A	1.000	0.027	0.915	0.058	0.947	0.032	0.914	0.001
A	Z	1.000	0.228	0.688	0.084	0.955	0.308	0.011	0.636
Z	I	0.999	0.338	0.003	0.658	1.000	0.749	0.020	0.231
Z	A	1.000	0.410	0.004	0.586	1.000	0.316	0.606	0.078
Z	Z	0.999	0.337	0.000	0.662	1.000	0.083	0.003	0.914
$t_{a_2} = 503$									
I	I	0.998	0.238	0.007	0.753	0.905	0.811	0.049	0.045
I	A	1.000	0.652	0.036	0.312	0.959	0.203	0.748	0.008
I	Z	0.999	0.279	0.004	0.716	0.945	0.237	0.021	0.687
A	I	1.000	0.119	0.774	0.107	1.000	0.855	0.037	0.108
A	A	1.000	0.008	0.976	0.016	1.000	0.015	0.985	0.000
A	Z	1.000	0.191	0.614	0.195	0.998	0.263	0.021	0.714
Z	I	1.000	0.180	0.005	0.815	1.000	0.839	0.007	0.154
Z	A	1.000	0.389	0.006	0.605	0.997	0.261	0.710	0.026
Z	Z	1.000	0.181	0.003	0.816	1.000	0.106	0.004	0.890

Tabelle A.22: Dargestellt sind die Wahrscheinlichkeiten einen bestimmten Ausreißertyp bei den Zeitpunkten t_{a_1} und t_{a_2} zu detektieren, woraus man Aussagen über die Qualität der Identifikation des Ausreißertyps ableiten kann. Dabei folgen die der Berechnung zu Grunde gelegten Zeitreihen dem $MA(2)$ Prozess nach Gleichung (8.10) und sind mit zwei Ausreißern der Größe $\omega = 7.0$ bei den Zeitpunkten t_{a_1} und t_{a_2} verunreinigt worden.

		Typ 1				Typ 2			
Typ 1	Typ 2	Gesamt	Typ I	Typ A	Typ Z	Gesamt	Typ I	Typ A	Typ Z
$t_{a_2} = 504$									
I	I	0.999	0.597	0.008	0.394	0.997	0.886	0.020	0.091
I	A	1.000	0.833	0.021	0.146	0.998	0.175	0.811	0.012
I	Z	1.000	0.516	0.004	0.480	0.995	0.156	0.007	0.832
A	I	1.000	0.116	0.837	0.047	1.000	0.875	0.032	0.093
A	A	1.000	0.018	0.961	0.021	0.999	0.022	0.976	0.001
A	Z	0.999	0.200	0.706	0.093	1.000	0.272	0.011	0.717
Z	I	1.000	0.143	0.006	0.851	1.000	0.833	0.010	0.157
Z	A	1.000	0.361	0.008	0.631	1.000	0.300	0.671	0.029
Z	Z	0.999	0.085	0.000	0.914	1.000	0.092	0.000	0.908
$t_{a_2} = 505$									
I	I	1.000	0.572	0.021	0.407	0.999	0.877	0.025	0.097
I	A	0.998	0.824	0.029	0.145	0.998	0.168	0.811	0.019
I	Z	0.999	0.550	0.018	0.431	0.997	0.154	0.006	0.837
A	I	0.999	0.105	0.867	0.027	1.000	0.870	0.029	0.101
A	A	1.000	0.012	0.985	0.003	1.000	0.013	0.986	0.001
A	Z	1.000	0.208	0.754	0.038	1.000	0.364	0.006	0.630
Z	I	0.999	0.133	0.002	0.864	1.000	0.850	0.008	0.142
Z	A	1.000	0.431	0.009	0.560	1.000	0.216	0.762	0.022
Z	Z	1.000	0.077	0.004	0.919	1.000	0.087	0.001	0.912
$t_{a_2} = 510$									
I	I	1.000	0.831	0.021	0.148	0.999	0.854	0.027	0.118
I	A	0.999	0.880	0.032	0.087	1.000	0.129	0.857	0.014
I	Z	1.000	0.828	0.010	0.162	1.000	0.164	0.006	0.830
A	I	1.000	0.118	0.864	0.018	1.000	0.898	0.029	0.073
A	A	1.000	0.011	0.986	0.003	1.000	0.008	0.991	0.001
A	Z	0.999	0.208	0.768	0.023	1.000	0.485	0.010	0.505
Z	I	1.000	0.161	0.004	0.835	1.000	0.872	0.015	0.113
Z	A	1.000	0.437	0.007	0.556	1.000	0.228	0.757	0.015
Z	Z	1.000	0.122	0.002	0.876	1.000	0.104	0.002	0.894

Tabelle A.23: Dargestellt sind die Wahrscheinlichkeiten einen bestimmten Ausreißertyp bei den Zeitpunkten t_{a_1} und t_{a_2} zu detektieren, woraus man Aussagen über die Qualität der Identifikation des Ausreißertyps ableiten kann. Dabei folgen die der Berechnung zu Grunde gelegten Zeitreihen dem $MA(2)$ Prozess nach Gleichung (8.10) und sind mit zwei Ausreißern der Größe $\omega = 7.0$ bei den Zeitpunkten t_{a_1} und t_{a_2} verunreinigt worden.

Typ 1	Typ 2	$\mu_{\hat{\omega}_1}$	$\sigma_{\hat{\omega}_1}$	$\mu_{\hat{\omega}_2}$	$\sigma_{\hat{\omega}_2}$	$\mu_{\hat{\omega}_1}$	$\sigma_{\hat{\omega}_1}$	$\mu_{\hat{\omega}_2}$	$\sigma_{\hat{\omega}_2}$
		$t_{a_2} = 501$				$t_{a_2} = 502$			
I	I	6.7429	1.0092	6.6939	1.1359	5.8842	0.7098	6.0505	1.2810
I	A	7.0179	0.9603	7.2337	0.7990	6.7594	0.8561	6.9759	1.0002
I	Z	6.2550	0.7753	6.9964	1.2299	5.8257	0.6905	5.9921	1.4016
A	I	-	-	4.0892	0.5030	7.1861	0.8555	7.1462	1.2232
A	A	-	-	3.7152	0.5964	6.9846	0.8551	6.9620	0.9164
A	Z	-	-	4.0808	0.5049	7.2250	0.7910	7.3071	1.2909
Z	I	7.1556	0.9651	6.8614	1.0311	7.2838	0.9781	7.3513	1.4340
Z	A	6.9306	0.9673	7.3132	0.7667	7.0393	0.9989	7.5228	0.9154
Z	Z	7.1416	0.9961	7.2584	1.1362	7.2768	0.9542	7.7146	1.5543
		$t_{a_2} = 503$				$t_{a_2} = 504$			
I	I	6.1229	0.7407	6.1976	1.2268	6.7919	0.9718	6.8301	1.1073
I	A	6.7084	0.8636	7.0482	0.9726	6.9972	0.9736	7.1432	0.8249
I	Z	6.1541	0.6899	6.3286	1.3263	6.6527	0.8718	6.7695	1.1072
A	I	7.1256	0.8074	6.9146	1.0297	7.0981	0.8241	6.9846	1.0108
A	A	7.0173	0.8269	7.0228	0.8451	7.0816	0.8491	6.9781	0.8954
A	Z	7.1367	0.8751	6.8490	1.1406	7.2064	0.8239	6.8826	0.9636
Z	I	7.1003	0.9738	7.1647	1.2483	7.0022	0.9308	7.0878	1.0286
Z	A	7.0077	1.0604	7.6794	0.8043	7.0503	1.0507	7.3860	0.7790
Z	Z	7.1632	0.9914	7.3334	1.2302	7.0723	0.9604	7.0874	1.0130
		$t_{a_2} = 505$				$t_{a_2} = 510$			
I	I	6.7693	0.9877	6.8131	1.0967	7.0718	1.0001	7.0399	1.0078
I	A	6.9506	0.9315	7.1121	0.7703	7.0225	0.9536	7.1135	0.8414
I	Z	6.7618	0.9242	6.8121	1.0474	7.0201	0.9723	7.0150	0.9667
A	I	7.1553	0.8482	7.0264	0.9933	7.1161	0.8081	7.0744	0.9711
A	A	6.9988	0.8230	6.9626	0.8354	7.0145	0.8529	7.0349	0.8340
A	Z	7.1972	0.7780	6.9883	1.0366	7.2130	0.7876	6.9869	1.0703
Z	I	7.0947	0.9705	7.0995	0.9985	7.0704	0.9462	7.0488	0.9494
Z	A	6.9941	1.0793	7.3507	0.795	6.9493	1.0754	7.2126	0.8491
Z	Z	7.0334	0.9482	7.0272	0.9962	7.0732	0.9354	7.0009	0.9100

Tabelle A.24: Dargestellt sind die Mittelwerte und Standardabweichungen der geschätzten Größe der korrekt identifizierten Ausreißer bei den Zeitpunkten t_{a_1} und t_{a_2} , woraus man Aussagen über die Qualität der Schätzung der Größe der Ausreißer ableiten kann. Dabei folgen die der Berechnung zu Grunde gelegten Zeitreihen dem $MA(2)$ Prozess nach Gleichung (8.10) und sind mit zwei Ausreißern der Größe $\omega = 7.0$ bei den Zeitpunkten t_{a_1} und t_{a_2} verunreinigt worden.

A.6 Tabellen zu Abschnitt 8.5

Es wurden Zeitreihen erzeugt, bei denen ein vorgegebener Prozentsatz der Werte mit Ausreißern verunreinigt ist. Dieser Prozentsatz ist in der ersten Spalte der Tabellen angegeben. Für jeden Prozentsatz wurden 1000 Zeitreihen erzeugt. Jede Zeitreihe wurde mit dem modifizierten Algorithmus von CHEN und LIU ausgewertet. Anschließend wurde aus den jeweils 1000 Schätzungen der Prozessparameter und der Varianz der Residuen deren Mittelwert μ und Standardabweichung σ ermittelt. Außerdem wurde die Anzahl der detektierten Ausreißer für jede der jeweils 1000 Zeitreihen gespeichert. Daraus wurde der Mittelwert gebildet. Dieser wurde in den Prozentsatz der mit Ausreißern verunreinigten Werte der Zeitreihen umgerechnet, welcher in der letzten Spalte der Tabellen aufgeführt ist.

Hinzugefügt	$\mu_{\hat{\varphi}_1}$	$\mu_{\hat{\theta}_1}$	$\mu_{\hat{\sigma}_\varepsilon^2}$	$\sigma_{\hat{\varphi}_1}$	$\sigma_{\hat{\theta}_1}$	$\sigma_{\hat{\sigma}_\varepsilon^2}$	Detektiert
0 %	0.610	0.290	0.995	0.032	0.04	0.045	0.07 %
0.5 %	0.611	0.289	0.995	0.032	0.043	0.045	0.37 %
1 %	0.610	0.290	0.997	0.033	0.043	0.045	0.88 %
2 %	0.609	0.284	1.006	0.034	0.044	0.050	1.92 %
3 %	0.606	0.277	1.025	0.036	0.044	0.052	2.98 %
4 %	0.607	0.266	1.044	0.036	0.045	0.061	4.07 %
5 %	0.609	0.254	1.062	0.039	0.045	0.067	5.13 %
6 %	0.611	0.240	1.082	0.041	0.047	0.078	6.20 %
7 %	0.611	0.229	1.114	0.043	0.048	0.089	7.25 %
8 %	0.610	0.219	1.135	0.042	0.050	0.125	8.29 %
9 %	0.614	0.180	1.405	0.044	0.087	0.690	8.54 %
10 %	0.625	0.041	3.164	0.036	0.104	1.236	4.08 %
11 %	0.632	-0.013	4.599	0.030	0.052	0.460	1.20 %

Tabelle A.25: Die der Berechnung zu Grunde liegenden Zeitreihen, bei denen ein vorgegebener Prozentsatz der Werte mit Ausreißern verunreinigt ist, folgen dem $ARMA(1,1)$ Prozess nach Gleichung (8.13). Die Größe der Ausreißer, die in den Zeitreihen enthalten sind, beträgt $\omega = 5 \pm 1$. Dargestellt sind die Mittelwerte und Standardabweichungen der Prozessparameter sowie der Prozentsatz der Werte der Zeitreihen, die mit detektierten Ausreißern verunreinigt sind.

Hinzugefügt	$\mu_{\hat{\varphi}_1}$	$\mu_{\hat{\varphi}_2}$	$\mu_{\hat{\sigma}_\varepsilon^2}$	$\sigma_{\hat{\varphi}_1}$	$\sigma_{\hat{\varphi}_2}$	$\sigma_{\hat{\sigma}_\varepsilon^2}$	Detektiert
0 %	0.601	0.198	0.991	0.030	0.032	0.046	0.08 %
0.5 %	0.602	0.197	0.991	0.032	0.032	0.046	0.37 %
1 %	0.602	0.198	1.000	0.032	0.032	0.046	0.87 %
2 %	0.608	0.193	1.012	0.033	0.033	0.049	1.93 %
3 %	0.617	0.185	1.030	0.033	0.033	0.054	3.01 %
4 %	0.624	0.181	1.049	0.034	0.033	0.060	4.09 %
5 %	0.631	0.175	1.070	0.035	0.034	0.069	5.13 %
6 %	0.637	0.169	1.093	0.036	0.034	0.074	6.21 %
7 %	0.644	0.163	1.116	0.036	0.034	0.082	7.24 %
8 %	0.649	0.156	1.145	0.037	0.033	0.095	8.27 %
9 %	0.655	0.150	1.209	0.042	0.035	0.299	9.17 %
10 %	0.595	0.166	2.293	0.078	0.038	1.279	6.58 %
11 %	0.525	0.185	4.197	0.045	0.028	0.828	1.80 %

Tabelle A.26: Die der Berechnung zu Grunde liegenden Zeitreihen, bei denen ein vorgegebener Prozentsatz der Werte mit Ausreißern verunreinigt ist, folgen dem $AR(2)$ Prozess nach Gleichung (8.14). Die Größe der Ausreißer, die in den Zeitreihen enthalten sind, beträgt $\omega = 5 \pm 1$. Dargestellt sind die Mittelwerte und Standardabweichungen der Prozessparameter sowie der Prozentsatz der Werte der Zeitreihen, die mit detektierten Ausreißern verunreinigt sind.

Hinzugefügt	$\mu_{\hat{\theta}_1}$	$\mu_{\hat{\theta}_2}$	$\mu_{\hat{\sigma}_\varepsilon^2}$	$\sigma_{\hat{\theta}_1}$	$\sigma_{\hat{\theta}_2}$	$\sigma_{\hat{\sigma}_\varepsilon^2}$	Detektiert
0 %	0.602	0.201	1.011	0.032	0.039	0.046	0.07 %
0.5 %	0.600	0.198	1.010	0.032	0.038	0.046	0.37 %
1 %	0.598	0.199	1.016	0.032	0.038	0.050	0.89 %
2 %	0.601	0.205	1.041	0.033	0.039	0.053	1.94 %
3 %	0.602	0.212	1.071	0.035	0.040	0.060	2.99 %
4 %	0.604	0.214	1.106	0.034	0.039	0.066	4.01 %
5 %	0.606	0.223	1.141	0.036	0.039	0.072	5.03 %
6 %	0.603	0.223	1.189	0.036	0.038	0.082	6.01 %
7 %	0.607	0.230	1.233	0.035	0.038	0.095	6.95 %
8 %	0.606	0.233	1.293	0.038	0.038	0.111	7.85 %
9 %	0.587	0.235	1.616	0.052	0.038	0.694	7.85 %
10 %	0.518	0.240	3.291	0.059	0.034	1.177	3.69 %
11 %	0.487	0.240	4.667	0.034	0.030	0.377	0.98 %

Tabelle A.27: Die der Berechnung zu Grunde liegenden Zeitreihen, bei denen ein vorgegebener Prozentsatz der Werte mit Ausreißern verunreinigt ist, folgen dem $MA(2)$ Prozess nach Gleichung (8.15). Die Größe der Ausreißer, die in den Zeitreihen enthalten sind, beträgt $\omega = 5 \pm 1$. Dargestellt sind die Mittelwerte und Standardabweichungen der Prozessparameter sowie der Prozentsatz der Werte der Zeitreihen, die mit detektierten Ausreißern verunreinigt sind.

Hinzugefügt	$\mu_{\hat{\varphi}_1}$	$\mu_{\hat{\theta}_1}$	$\mu_{\hat{\sigma}_\varepsilon^2}$	$\sigma_{\hat{\varphi}_1}$	$\sigma_{\hat{\theta}_1}$	$\sigma_{\hat{\sigma}_\varepsilon^2}$	Detektiert
0 %	0.610	0.291	0.992	0.033	0.042	0.045	0.07 %
0.5 %	0.609	0.291	0.998	0.033	0.041	0.046	0.38 %
1 %	0.608	0.286	1.007	0.033	0.042	0.048	0.96 %
2 %	0.606	0.277	1.073	0.034	0.041	0.075	2.02 %
3 %	0.603	0.268	1.203	0.037	0.043	0.104	3.03 %
4 %	0.608	0.254	1.246	0.036	0.043	0.118	4.09 %
5 %	0.610	0.256	1.172	0.037	0.044	0.104	5.22 %
6 %	0.610	0.243	1.232	0.037	0.045	0.124	6.24 %
7 %	0.610	0.232	1.320	0.037	0.045	0.137	7.29 %
8 %	0.612	0.231	1.297	0.036	0.047	0.133	8.39 %
9 %	0.614	0.215	1.369	0.035	0.047	0.148	9.47 %
10 %	0.615	0.184	2.125	0.035	0.084	2.740	9.98 %
11 %	0.640	-0.051	14.283	0.022	0.051	1.547	0.99 %

Tabelle A.28: Die der Berechnung zu Grunde liegenden Zeitreihen, bei denen ein vorgegebener Prozentsatz der Werte mit Ausreißern verunreinigt ist, folgen dem $ARMA(1,1)$ Prozess nach Gleichung (8.13). Die Größe der Ausreißer, die in den Zeitreihen enthalten sind, beträgt $\omega = 10 \pm 1$. Dargestellt sind die Mittelwerte und Standardabweichungen der Prozessparameter sowie der Prozentsatz der Werte der Zeitreihen, die mit detektierten Ausreißern verunreinigt sind.

Hinzugefügt	$\mu_{\hat{\varphi}_1}$	$\mu_{\hat{\varphi}_2}$	$\mu_{\hat{\sigma}_\varepsilon^2}$	$\sigma_{\hat{\varphi}_1}$	$\sigma_{\hat{\varphi}_2}$	$\sigma_{\hat{\sigma}_\varepsilon^2}$	Detektiert
0 %	0.601	0.197	0.991	0.033	0.032	0.044	0.08 %
0.5 %	0.600	0.200	0.994	0.033	0.031	0.044	0.39 %
1 %	0.603	0.200	1.002	0.033	0.032	0.046	0.95 %
2 %	0.607	0.196	1.028	0.035	0.033	0.052	2.11 %
3 %	0.615	0.192	1.074	0.037	0.033	0.058	3.26 %
4 %	0.622	0.189	1.116	0.039	0.034	0.070	4.43 %
5 %	0.633	0.185	1.117	0.040	0.036	0.072	5.56 %
6 %	0.639	0.183	1.157	0.042	0.035	0.080	6.71 %
7 %	0.643	0.180	1.222	0.043	0.037	0.087	7.85 %
8 %	0.653	0.172	1.215	0.042	0.034	0.099	8.98 %
9 %	0.663	0.168	1.272	0.045	0.038	0.112	10.09 %
10 %	0.665	0.163	1.513	0.048	0.036	1.302	11.05 %
11 %	0.514	0.182	13.061	0.051	0.020	3.149	1.61 %

Tabelle A.29: Die der Berechnung zu Grunde liegenden Zeitreihen, bei denen ein vorgegebener Prozentsatz der Werte mit Ausreißern verunreinigt ist, folgen dem $AR(2)$ Prozess nach Gleichung (8.14). Die Größe der Ausreißer, die in den Zeitreihen enthalten sind, beträgt $\omega = 10 \pm 1$. Dargestellt sind die Mittelwerte und Standardabweichungen der Prozessparameter sowie der Prozentsatz der Werte der Zeitreihen, die mit detektierten Ausreißern verunreinigt sind.

Hinzugefügt	$\mu_{\hat{\theta}_1}$	$\mu_{\hat{\theta}_2}$	$\mu_{\hat{\sigma}_\varepsilon^2}$	$\sigma_{\hat{\theta}_1}$	$\sigma_{\hat{\theta}_2}$	$\sigma_{\hat{\sigma}_\varepsilon^2}$	Detektiert
0 %	0.601	0.199	1.010	0.033	0.038	0.047	0.07 %
0.5 %	0.601	0.201	1.012	0.034	0.040	0.045	0.39 %
1 %	0.608	0.210	1.021	0.034	0.038	0.050	0.98 %
2 %	0.625	0.232	1.066	0.037	0.043	0.059	2.18 %
3 %	0.635	0.251	1.118	0.041	0.044	0.068	3.35 %
4 %	0.645	0.264	1.162	0.042	0.045	0.081	4.51 %
5 %	0.652	0.284	1.190	0.046	0.046	0.107	5.70 %
6 %	0.649	0.291	1.271	0.051	0.047	0.145	6.82 %
7 %	0.649	0.304	1.343	0.053	0.049	0.136	7.99 %
8 %	0.644	0.312	1.445	0.060	0.050	0.190	9.10 %
9 %	0.635	0.311	1.604	0.059	0.049	0.257	10.21 %
10 %	0.563	0.285	6.117	0.096	0.052	5.356	7.29 %
11 %	0.468	0.244	14.803	0.035	0.026	0.931	0.52 %

Tabelle A.30: Die der Berechnung zu Grunde liegenden Zeitreihen, bei denen ein vorgegebener Prozentsatz der Werte mit Ausreißern verunreinigt ist, folgen dem $MA(2)$ Prozess nach Gleichung (8.15). Die Größe der Ausreißer, die in den Zeitreihen enthalten sind, beträgt $\omega = 10 \pm 1$. Dargestellt sind die Mittelwerte und Standardabweichungen der Prozessparameter sowie der Prozentsatz der Werte der Zeitreihen, die mit detektierten Ausreißern verunreinigt sind.

Hinzugefügt	$\mu_{\hat{\varphi}_1}$	$\mu_{\hat{\theta}_1}$	$\mu_{\hat{\sigma}_\varepsilon^2}$	$\sigma_{\hat{\varphi}_1}$	$\sigma_{\hat{\theta}_1}$	$\sigma_{\hat{\sigma}_\varepsilon^2}$	Detektiert
0 %	0.608	0.293	0.996	0.033	0.043	0.046	0.07 %
0.5 %	0.610	0.283	1.014	0.033	0.044	0.058	0.40 %
1 %	0.608	0.276	1.046	0.034	0.045	0.076	1.07 %
2 %	0.604	0.254	1.242	0.036	0.045	0.149	2.28 %
3 %	0.604	0.236	1.455	0.040	0.047	0.195	3.35 %
4 %	0.611	0.219	1.611	0.041	0.046	0.217	4.36 %
5 %	0.615	0.218	1.692	0.040	0.048	0.239	5.46 %
6 %	0.616	0.229	1.446	0.041	0.045	0.226	6.90 %
7 %	0.622	0.197	1.878	0.042	0.048	0.285	7.43 %
8 %	0.627	0.199	1.937	0.042	0.052	0.287	8.52 %
9 %	0.630	0.175	2.300	0.039	0.047	0.380	9.26 %
10 %	0.634	0.135	5.143	0.036	0.096	7.831	9.26 %
11 %	0.639	-0.065	30.779	0.020	0.047	1.823	0.59 %

Tabelle A.31: Die der Berechnung zu Grunde liegenden Zeitreihen, bei denen ein vorgegebener Prozentsatz der Werte mit Ausreißern verunreinigt ist, folgen dem $ARMA(1,1)$ Prozess nach Gleichung (8.13). Die Größe der Ausreißer, die in den Zeitreihen enthalten sind, beträgt $\omega = 15 \pm 1$. Dargestellt sind die Mittelwerte und Standardabweichungen der Prozessparameter sowie der Prozentsatz der Werte der Zeitreihen, die mit detektierten Ausreißern verunreinigt sind.

Hinzugefügt	$\mu_{\hat{\varphi}_1}$	$\mu_{\hat{\varphi}_2}$	$\mu_{\hat{\sigma}_\varepsilon^2}$	$\sigma_{\hat{\varphi}_1}$	$\sigma_{\hat{\varphi}_2}$	$\sigma_{\hat{\sigma}_\varepsilon^2}$	Detektiert
0 %	0.600	0.198	0.990	0.031	0.031	0.045	0.09 %
0.5 %	0.604	0.197	0.999	0.033	0.032	0.047	0.38 %
1 %	0.611	0.196	1.017	0.036	0.033	0.050	0.94 %
2 %	0.626	0.187	1.117	0.037	0.033	0.078	2.07 %
3 %	0.644	0.179	1.279	0.040	0.034	0.126	3.16 %
4 %	0.665	0.168	1.417	0.041	0.034	0.130	4.22 %
5 %	0.680	0.166	1.511	0.041	0.033	0.153	5.31 %
6 %	0.690	0.167	1.415	0.046	0.035	0.146	6.59 %
7 %	0.698	0.158	1.675	0.043	0.032	0.222	7.46 %
8 %	0.705	0.149	1.815	0.046	0.034	0.219	8.46 %
9 %	0.704	0.146	2.122	0.044	0.035	0.279	9.40 %
10 %	0.713	0.141	2.609	0.057	0.034	3.414	10.30 %
11 %	0.500	0.181	29.637	0.035	0.016	2.878	0.63 %

Tabelle A.32: Die der Berechnung zu Grunde liegenden Zeitreihen, bei denen ein vorgegebener Prozentsatz der Werte mit Ausreißern verunreinigt ist, folgen dem $AR(2)$ Prozess nach Gleichung (8.14). Die Größe der Ausreißer, die in den Zeitreihen enthalten sind, beträgt $\omega = 15 \pm 1$. Dargestellt sind die Mittelwerte und Standardabweichungen der Prozessparameter sowie der Prozentsatz der Werte der Zeitreihen, die mit detektierten Ausreißern verunreinigt sind.

Hinzugefügt	$\mu_{\hat{\theta}_1}$	$\mu_{\hat{\theta}_2}$	$\mu_{\hat{\sigma}_\varepsilon^2}$	$\sigma_{\hat{\theta}_1}$	$\sigma_{\hat{\theta}_2}$	$\sigma_{\hat{\sigma}_\varepsilon^2}$	Detektiert
0 %	0.602	0.201	1.009	0.033	0.039	0.048	0.07 %
0.5 %	0.602	0.201	1.016	0.032	0.039	0.047	0.41 %
1 %	0.608	0.211	1.035	0.036	0.041	0.053	1.08 %
2 %	0.615	0.225	1.117	0.042	0.045	0.080	2.43 %
3 %	0.618	0.236	1.232	0.048	0.050	0.136	3.67 %
4 %	0.616	0.246	1.343	0.055	0.055	0.197	4.86 %
5 %	0.608	0.250	1.505	0.066	0.064	0.386	6.03 %
6 %	0.597	0.253	1.563	0.072	0.071	0.518	7.35 %
7 %	0.588	0.258	1.831	0.075	0.074	0.657	8.33 %
8 %	0.584	0.278	2.017	0.093	0.078	0.803	9.50 %
9 %	0.570	0.299	2.339	0.099	0.081	0.950	10.49 %
10 %	0.492	0.274	18.066	0.080	0.063	11.582	4.90 %
11 %	0.462	0.245	31.568	0.032	0.024	1.141	0.21 %

Tabelle A.33: Die der Berechnung zu Grunde liegenden Zeitreihen, bei denen ein vorgegebener Prozentsatz der Werte mit Ausreißern verunreinigt ist, folgen dem $MA(2)$ Prozess nach Gleichung (8.15). Die Größe der Ausreißer, die in den Zeitreihen enthalten sind, beträgt $\omega = 15 \pm 1$. Dargestellt sind die Mittelwerte und Standardabweichungen der Prozessparameter sowie der Prozentsatz der Werte der Zeitreihen, die mit detektierten Ausreißern verunreinigt sind.

A.7 Tabellen zu Abschnitt 8.6

Mit jeder der drei Prozessgleichungen (8.16) bis (8.18) wurden 100 Zeitreihen erzeugt, welche die in Tabelle 8.7 aufgeführten Ausreißer enthalten. Für jede der Zeitreihen wurde die Modellordnung mit den vier Informationskriterien *AIC*, *BIC*, *HQC* und *MixedC* bestimmt. Als Informationskriterium sei hier die Matrix bezeichnet, die man erhält, wenn man Gleichung (6.1) oder (6.3) für $p = 0, 1, \dots, 9$ und $q = 0, 1, \dots, 9$ mit einer Schätzung der Prozessparameter und Ausreißer für eine Zeitreihe berechnet.

Bei der Bestimmung der Informationskriterien wurden einmal die Parameter, die mit dem modifizierten Algorithmus von CHEN und LIU geschätzt wurden und ein anderes Mal die Parameter, die aus einer Schätzung nach der Methode der kleinsten Quadrate ohne Ausreißerschätzung hervorgehen, verwendet. Zusätzlich wurde für jede Zeitreihe die Anzahl der durch den modifizierten Algorithmus von CHEN und LIU detektierten Ausreißer festgehalten. Für jeden der drei Prozesse wurde daraus der Mittelwert berechnet. Diese Mittelwerte sind in den Tabellen A.37 bis A.39 dargestellt.

Zur Tabelle A.34: In dieser Tabelle ist der Prozentsatz der jeweils 100 Zeitreihen, bei dem die Bestimmung der Modellordnung mit dem Informationskriterium korrekt war, dargestellt. So findet man beispielsweise die Prozentzahl 35 beim *BIC*, welches mit Ausreißerschätzung für den *ARMA(1,1)* Prozess nach Gleichung (8.16) bestimmt wurde. Diese Zahl bedeutet, dass bei 35 der 100 Zeitreihen die Modellordnung $p = 1$ und $q = 1$ durch das *BIC*, welches mit den durch den modifizierten Algorithmus von CHEN und LIU geschätzten Parametern berechnet wurde, korrekt bestimmt wurde.

Zur Tabelle A.35: Aus den 100 Modellordnungen, die für jede Kombination der drei Prozesse und vier Informationskriterien bestimmt worden sind, wurde der Mittelwert und die Standardabweichung berechnet.

Zur Tabelle A.36: Aus den jeweils 100 Informationskriterien, die für jeden der drei Prozesse bestimmt worden sind, wurde der Mittelwert berechnet. Dargestellt ist die aus dem Mittelwert der Informationskriterien abgeleitete Modellordnung.

Zu den Tabellen A.40 bis A.63: Aus den jeweils 100 Informationskriterien, die für jeden der drei Prozesse bestimmt worden sind, wurde der Mittelwert, welcher in den Tabellen A.40 bis A.63 dargestellt ist, berechnet.

	Mit Ausreißerschätzung				Ohne Ausreißerschätzung			
	AIC	BIC	HQC	MixedC	AIC	BIC	HQC	MixedC
$ARMA(1,1)$	13%	35%	26%	40%	13%	10%	14%	10%
$AR(2)$	5%	14%	9%	14%	62%	63%	63%	63%
$MA(2)$	5%	9%	6%	11%	2%	2%	2%	2%

Tabelle A.34: Mit jeder der drei Prozessgleichungen (8.16) bis (8.18) wurden 100 Zeitreihen erzeugt, die die in Tabelle 8.7 aufgeführten Ausreißer enthalten. Dargestellt ist der Prozentsatz der jeweils 100 Zeitreihen, bei dem die Bestimmung der Modellordnung mit dem Informationskriterium korrekt war.

	$ARMA(1,1)$				$AR(2)$				$MA(2)$			
	μ_p	μ_q	σ_p	σ_q	μ_p	μ_q	σ_p	σ_q	μ_p	μ_q	σ_p	σ_q
Mit Ausreißerschätzung												
AIC	2.5	2.0	1.7	1.6	2.3	1.5	1.5	1.3	2.4	1.5	1.8	1.6
BIC	1.7	1.0	0.8	1.0	1.4	1.0	0.7	0.7	1.3	1.1	0.9	0.8
HQC	1.9	1.3	1.2	1.2	1.7	1.3	0.8	0.9	1.8	1.3	1.4	1.1
MixedC	1.5	1.1	0.7	0.8	1.4	1.0	0.7	0.6	1.3	1.0	0.8	0.7
Ohne Ausreißerschätzung												
AIC	1.7	0.2	0.6	0.4	1.7	0.4	0.5	0.5	1.7	0.2	0.8	0.5
BIC	1.3	0.1	2.3	0.3	1.6	0.4	2.6	0.5	1.2	0.1	2.2	0.4
HQC	1.5	0.1	0.5	0.3	1.6	0.4	0.5	0.5	1.4	0.2	0.5	0.4
MixedC	1.3	0.1	0.5	0.3	1.6	0.4	0.5	0.5	1.2	0.1	0.5	0.4

Tabelle A.35: Mit jeder der drei Prozessgleichungen (8.16) bis (8.18) wurden 100 Zeitreihen erzeugt, die die in Tabelle 8.7 aufgeführten Ausreißer enthalten. Dargestellt sind die Mittelwerte und Standardabweichungen der Modellordnungen, die von den einzelnen Informationskriterien bestimmt wurden.

	Mit Ausreißerschätzung				Ohne Ausreißerschätzung			
	AIC	BIC	HQC	MixedC	AIC	BIC	HQC	MixedC
$ARMA(1,1)$	2,2	1,1	2,1	1,1	2,0	1,0	2,0	1,0
$AR(2)$	2,1	2,1	1,1	1,1	2,0	2,0	2,0	2,0
$MA(2)$	1,1	1,1	1,1	1,1	2,0	1,0	2,0	1,0

Tabelle A.36: Mit jeder der drei Prozessgleichungen (8.16) bis (8.18) wurden 100 Zeitreihen erzeugt, die die in Tabelle 8.7 aufgeführten Ausreißer enthalten. Aus den jeweils 100 Informationskriterien AIC , BIC , HQC und $MixedC$, die für die Zeitreihen berechnet wurden, ist der Mittelwert gebildet worden. Diese Mittelwerte sind in den Tabellen A.40 bis A.63 dargestellt. Aus diesen Mittelwerten wurden die dargestellten Modellordnungen p , q bestimmt.

	q = 0	1	2	3	4	5	6	7	8	9
p = 0	25.5	29.6	13.0	12.3	11.8	11.9	11.6	11.7	11.8	11.8
1	12.1	11.9	11.9	11.9	11.9	11.9	11.9	12.0	11.9	12.0
2	11.9	11.8	12.0	11.9	11.9	11.9	11.9	11.9	11.9	12.0
3	11.9	11.9	11.9	11.9	11.9	11.9	12.0	11.9	11.9	12.0
4	11.9	12.0	11.9	11.9	11.9	12.0	12.0	12.0	11.9	12.0
5	11.9	11.9	11.9	11.9	11.9	12.0	12.0	11.9	11.9	12.0
6	12.0	11.9	11.9	11.9	11.9	12.0	11.9	11.9	11.8	11.9
7	11.9	11.9	11.9	11.9	11.9	11.9	11.9	11.9	11.9	11.9
8	11.9	11.9	11.9	11.9	11.9	11.9	12.0	11.9	11.9	11.8
9	11.9	12.0	11.9	11.9	12.0	12.0	12.0	11.9	11.9	11.8

Tabelle A.37: Mit der Prozessgleichung (8.16) wurden 100 $ARMA(1,1)$ Zeitreihen erzeugt, die die in Tabelle 8.7 aufgeführten Ausreißer enthalten. Für jede der Zeitreihen wurden mit dem modifizierten Algorithmus von CHEN und LIU die Prozessparameter und Ausreißer geschätzt. Dargestellt ist der Mittelwert der Anzahl der detektierten Ausreißer in Abhängigkeit von der durch die Informationskriterien vorgegebenen Modellordnung p, q .

	q = 0	1	2	3	4	5	6	7	8	9
p = 0	27.7	15.7	11.6	11.5	11.3	11.1	11.2	11.1	11.2	11.2
1	12.2	11.6	11.6	11.6	11.7	11.6	11.6	11.6	11.6	11.6
2	11.6	11.5	11.5	11.6	11.7	11.6	11.5	11.5	11.5	11.5
3	11.6	11.5	11.5	11.6	11.6	11.5	11.5	11.5	11.5	11.5
4	11.6	11.5	11.6	11.5	11.6	11.5	11.5	11.5	11.5	11.5
5	11.6	11.6	11.6	11.6	11.6	11.6	11.6	11.5	11.5	11.5
6	11.4	11.5	11.6	11.5	11.5	11.6	11.6	11.6	11.5	11.6
7	11.5	11.6	11.6	11.6	11.6	11.6	11.6	11.5	11.6	11.5
8	11.6	11.6	11.5	11.5	11.6	11.5	11.5	11.6	11.5	11.5
9	11.6	11.6	11.6	11.5	11.6	11.6	11.6	11.6	11.6	11.5

Tabelle A.38: Mit der Prozessgleichung (8.17) wurden 100 $AR(2)$ Zeitreihen erzeugt, die die in Tabelle 8.7 aufgeführten Ausreißer enthalten. Für jede der Zeitreihen wurden mit dem modifizierten Algorithmus von CHEN und LIU die Prozessparameter und Ausreißer geschätzt. Dargestellt ist der Mittelwert der Anzahl der detektierten Ausreißer in Abhängigkeit von der durch die Informationskriterien vorgegebenen Modellordnung p, q .

	q = 0	1	2	3	4	5	6	7	8	9
p = 0	14.9	13.5	11.8	11.8	11.7	11.7	11.7	11.6	11.7	11.6
1	11.4	11.8	11.8	11.8	11.9	11.9	11.8	11.7	11.7	11.8
2	11.9	11.8	11.8	11.8	11.8	11.9	11.9	11.8	11.8	11.8
3	11.9	11.8	11.8	11.8	11.9	11.9	11.8	11.7	11.7	11.8
4	11.9	11.8	11.8	11.8	11.9	11.9	11.8	11.8	11.7	11.8
5	11.8	11.8	11.8	11.8	11.9	11.8	11.7	11.7	11.7	11.7
6	11.8	11.8	11.8	11.8	11.9	11.8	11.7	11.7	11.7	11.7
7	11.7	11.7	11.7	11.8	11.8	11.8	11.7	11.7	11.7	11.8
8	11.7	11.7	11.7	11.8	11.8	11.8	11.7	11.8	11.8	11.7
9	11.7	11.8	11.7	11.7	11.7	11.7	11.7	11.7	11.8	11.7

Tabelle A.39: Mit der Prozessgleichung (8.18) wurden 100 $MA(2)$ Zeitreihen erzeugt, die die in Tabelle 8.7 aufgeführten Ausreißer enthalten. Für jede der Zeitreihen wurden mit dem modifizierten Algorithmus von CHEN und LIU die Prozessparameter und Ausreißer geschätzt. Dargestellt ist der Mittelwert der Anzahl der detektierten Ausreißer in Abhängigkeit von der durch die Informationskriterien vorgegebenen Modellordnung p, q .

	q = 0	1	2	3	4	5	6	7	8	9
p = 0	0.763	0.557	0.239	0.152	0.117	0.108	0.104	0.102	0.102	0.104
1	0.071	0.038	0.037	0.035	0.038	0.040	0.042	0.045	0.047	0.049
2	0.041	0.035	0.034	0.035	0.037	0.039	0.042	0.044	0.048	0.050
3	0.042	0.035	0.036	0.037	0.038	0.040	0.044	0.047	0.050	0.052
4	0.041	0.036	0.036	0.039	0.041	0.043	0.046	0.049	0.052	0.055
5	0.041	0.036	0.038	0.041	0.042	0.046	0.048	0.051	0.054	0.057
6	0.042	0.039	0.040	0.042	0.044	0.048	0.051	0.055	0.057	0.059
7	0.041	0.041	0.043	0.046	0.048	0.051	0.054	0.057	0.061	0.062
8	0.042	0.043	0.046	0.048	0.050	0.054	0.057	0.061	0.062	0.065
9	0.045	0.047	0.049	0.051	0.054	0.057	0.060	0.063	0.066	0.069

Tabelle A.40: Mit der Prozessgleichung (8.16) wurden 100 $ARMA(1,1)$ Zeitreihen erzeugt, welche die in Tabelle 8.7 aufgeführten Ausreißer enthalten. Für jede der Zeitreihen wurden mit dem modifizierten Algorithmus von CHEN und LIU die Prozessparameter und Ausreißer geschätzt. Dargestellt ist der Mittelwert der damit berechneten 100 AIC nach Gleichung (6.3).

	q = 0	1	2	3	4	5	6	7	8	9
p = 0	0.888	0.708	0.313	0.227	0.195	0.191	0.190	0.194	0.199	0.206
1	0.135	0.106	0.110	0.113	0.121	0.128	0.135	0.143	0.150	0.157
2	0.109	0.107	0.112	0.118	0.125	0.132	0.139	0.147	0.156	0.163
3	0.115	0.113	0.119	0.125	0.131	0.138	0.147	0.155	0.163	0.170
4	0.119	0.119	0.124	0.131	0.139	0.146	0.154	0.162	0.170	0.177
5	0.124	0.124	0.131	0.138	0.145	0.154	0.161	0.168	0.176	0.184
6	0.130	0.132	0.138	0.145	0.152	0.160	0.169	0.177	0.183	0.191
7	0.134	0.138	0.146	0.153	0.161	0.169	0.177	0.184	0.193	0.199
8	0.140	0.146	0.153	0.161	0.167	0.176	0.184	0.193	0.199	0.206
9	0.148	0.155	0.162	0.168	0.176	0.184	0.192	0.200	0.208	0.216

Tabelle A.41: Mit der Prozessgleichung (8.16) wurden 100 $ARMA(1,1)$ Zeitreihen erzeugt, welche die in Tabelle 8.7 aufgeführten Ausreißer enthalten. Für jede der Zeitreihen wurden mit dem modifizierten Algorithmus von CHEN und LIU die Prozessparameter und Ausreißer geschätzt. Dargestellt ist der Mittelwert der damit berechneten 100 BIC nach Gleichung (6.3).

	q = 0	1	2	3	4	5	6	7	8	9
p = 0	0.810	0.614	0.267	0.181	0.147	0.139	0.137	0.137	0.138	0.143
1	0.095	0.063	0.065	0.065	0.070	0.073	0.078	0.082	0.086	0.090
2	0.067	0.062	0.063	0.066	0.070	0.074	0.079	0.083	0.089	0.093
3	0.069	0.064	0.067	0.070	0.074	0.078	0.083	0.088	0.093	0.097
4	0.071	0.067	0.070	0.074	0.078	0.082	0.087	0.092	0.097	0.101
5	0.072	0.069	0.073	0.078	0.081	0.087	0.091	0.096	0.101	0.105
6	0.076	0.074	0.078	0.081	0.085	0.091	0.096	0.101	0.105	0.109
7	0.076	0.078	0.082	0.087	0.091	0.096	0.101	0.105	0.111	0.114
8	0.079	0.082	0.087	0.091	0.095	0.100	0.105	0.111	0.114	0.119
9	0.084	0.088	0.092	0.095	0.100	0.105	0.110	0.115	0.120	0.125

Tabelle A.42: Mit der Prozessgleichung (8.16) wurden 100 $ARMA(1,1)$ Zeitreihen erzeugt, welche die in Tabelle 8.7 aufgeführten Ausreißer enthalten. Für jede der Zeitreihen wurden mit dem modifizierten Algorithmus von CHEN und LIU die Prozessparameter und Ausreißer geschätzt. Dargestellt ist der Mittelwert der damit berechneten 100 HQC nach Gleichung (6.3).

	q = 0	1	2	3	4	5	6	7	8	9
p = 0	0.763	0.562	0.249	0.167	0.137	0.132	0.133	0.136	0.141	0.148
1	0.076	0.047	0.052	0.055	0.063	0.069	0.077	0.084	0.091	0.098
2	0.051	0.049	0.053	0.059	0.066	0.073	0.081	0.089	0.097	0.104
3	0.056	0.054	0.060	0.066	0.073	0.080	0.088	0.096	0.104	0.111
4	0.061	0.060	0.066	0.073	0.080	0.087	0.095	0.103	0.111	0.119
5	0.065	0.065	0.073	0.080	0.086	0.095	0.102	0.110	0.118	0.126
6	0.071	0.073	0.080	0.086	0.093	0.102	0.110	0.119	0.125	0.133
7	0.075	0.080	0.087	0.095	0.102	0.110	0.118	0.126	0.134	0.141
8	0.082	0.088	0.095	0.102	0.109	0.118	0.125	0.134	0.141	0.148
9	0.089	0.096	0.103	0.110	0.117	0.125	0.133	0.141	0.150	0.158

Tabelle A.43: Mit der Prozessgleichung (8.16) wurden 100 $ARMA(1,1)$ Zeitreihen erzeugt, welche die in Tabelle 8.7 aufgeführten Ausreißer enthalten. Für jede der Zeitreihen wurden mit dem modifizierten Algorithmus von CHEN und LIU die Prozessparameter und Ausreißer geschätzt. Dargestellt ist der Mittelwert der damit berechneten 100 $MixedC$ nach Gleichung (6.3).

	q = 0	1	2	3	4	5	6	7	8	9
p = 0	1.402	1.077	0.916	0.847	0.820	0.810	0.808	0.808	0.809	0.810
1	0.778	0.775	0.777	0.779	0.781	0.784	0.786	0.788	0.790	0.793
2	0.774	0.776	0.778	0.781	0.783	0.785	0.787	0.790	0.792	0.794
3	0.776	0.778	0.781	0.783	0.785	0.787	0.790	0.792	0.794	0.796
4	0.779	0.781	0.783	0.785	0.787	0.790	0.792	0.794	0.796	0.799
5	0.781	0.783	0.785	0.788	0.790	0.792	0.794	0.797	0.799	0.801
6	0.783	0.785	0.788	0.790	0.792	0.794	0.797	0.799	0.801	0.804
7	0.785	0.788	0.790	0.792	0.794	0.797	0.799	0.802	0.804	0.806
8	0.788	0.790	0.792	0.794	0.797	0.799	0.802	0.804	0.807	0.809
9	0.790	0.792	0.794	0.797	0.799	0.801	0.804	0.807	0.809	0.814

Tabelle A.44: Mit der Prozessgleichung (8.16) wurden 100 $ARMA(1,1)$ Zeitreihen erzeugt, welche die in Tabelle 8.7 aufgeführten Ausreißer enthalten. Für jede der Zeitreihen wurden die Prozessparameter nach der Methode der kleinsten Quadrate geschätzt. Dargestellt ist der Mittelwert der damit berechneten 100 AIC nach Gleichung (6.1).

	q = 0	1	2	3	4	5	6	7	8	9
p = 0	1.402	1.082	0.925	0.862	0.840	0.835	0.837	0.842	0.848	0.854
1	0.783	0.785	0.792	0.799	0.806	0.813	0.820	0.828	0.835	0.842
2	0.784	0.791	0.798	0.805	0.812	0.820	0.827	0.834	0.841	0.848
3	0.791	0.798	0.805	0.812	0.819	0.827	0.834	0.841	0.848	0.855
4	0.798	0.805	0.812	0.820	0.827	0.834	0.841	0.848	0.855	0.863
5	0.805	0.812	0.819	0.827	0.834	0.841	0.848	0.855	0.863	0.870
6	0.813	0.820	0.827	0.834	0.841	0.848	0.856	0.863	0.870	0.877
7	0.820	0.827	0.834	0.841	0.848	0.856	0.863	0.870	0.877	0.885
8	0.827	0.834	0.841	0.848	0.856	0.863	0.870	0.878	0.885	0.893
9	0.834	0.841	0.848	0.855	0.863	0.870	0.877	0.885	0.892	0.902

Tabelle A.45: Mit der Prozessgleichung (8.16) wurden 100 $ARMA(1,1)$ Zeitreihen erzeugt, welche die in Tabelle 8.7 aufgeführten Ausreißer enthalten. Für jede der Zeitreihen wurden die Prozessparameter nach der Methode der kleinsten Quadrate geschätzt. Dargestellt ist der Mittelwert der damit berechneten 100 BIC nach Gleichung (6.1).

	q = 0	1	2	3	4	5	6	7	8	9
p = 0	1.402	1.079	0.919	0.853	0.827	0.820	0.819	0.821	0.824	0.826
1	0.780	0.779	0.782	0.787	0.791	0.795	0.799	0.803	0.807	0.811
2	0.778	0.782	0.786	0.790	0.794	0.798	0.802	0.806	0.811	0.815
3	0.782	0.786	0.790	0.794	0.798	0.802	0.806	0.811	0.815	0.819
4	0.786	0.790	0.794	0.798	0.802	0.807	0.811	0.815	0.819	0.823
5	0.790	0.794	0.798	0.802	0.807	0.811	0.815	0.819	0.823	0.827
6	0.794	0.798	0.802	0.807	0.811	0.815	0.819	0.823	0.827	0.832
7	0.798	0.802	0.806	0.811	0.815	0.819	0.823	0.828	0.832	0.836
8	0.803	0.807	0.811	0.815	0.819	0.823	0.828	0.832	0.836	0.841
9	0.807	0.811	0.815	0.819	0.823	0.827	0.832	0.837	0.841	0.847

Tabelle A.46: Mit der Prozessgleichung (8.16) wurden 100 $ARMA(1,1)$ Zeitreihen erzeugt, welche die in Tabelle 8.7 aufgeführten Ausreißer enthalten. Für jede der Zeitreihen wurden die Prozessparameter nach der Methode der kleinsten Quadrate geschätzt. Dargestellt ist der Mittelwert der damit berechneten 100 HQC nach Gleichung (6.1).

	q = 0	1	2	3	4	5	6	7	8	9
p = 0	1.402	1.082	0.925	0.862	0.840	0.835	0.837	0.842	0.848	0.854
1	0.783	0.785	0.792	0.799	0.806	0.813	0.820	0.828	0.835	0.842
2	0.784	0.791	0.798	0.805	0.812	0.820	0.827	0.834	0.841	0.848
3	0.791	0.798	0.805	0.812	0.819	0.827	0.834	0.841	0.848	0.855
4	0.798	0.805	0.812	0.820	0.827	0.834	0.841	0.848	0.855	0.863
5	0.805	0.812	0.819	0.827	0.834	0.841	0.848	0.855	0.863	0.870
6	0.813	0.820	0.827	0.834	0.841	0.848	0.856	0.863	0.870	0.877
7	0.820	0.827	0.834	0.841	0.848	0.856	0.863	0.870	0.877	0.885
8	0.827	0.834	0.841	0.848	0.856	0.863	0.870	0.878	0.885	0.893
9	0.834	0.841	0.848	0.855	0.863	0.870	0.877	0.885	0.892	0.902

Tabelle A.47: Mit der Prozessgleichung (8.16) wurden 100 $ARMA(1,1)$ Zeitreihen erzeugt, welche die in Tabelle 8.7 aufgeführten Ausreißer enthalten. Für jede der Zeitreihen wurden die Prozessparameter nach der Methode der kleinsten Quadrate geschätzt. Dargestellt ist der Mittelwert der damit berechneten 100 *MixedC* nach Gleichung (6.1).

	q = 0	1	2	3	4	5	6	7	8	9
p = 0	0.949	0.532	0.540	0.461	0.381	0.314	0.262	0.222	0.193	0.174
1	0.072	0.033	0.033	0.035	0.038	0.039	0.042	0.045	0.047	0.050
2	0.040	0.032	0.032	0.034	0.036	0.038	0.042	0.045	0.048	0.050
3	0.042	0.033	0.033	0.036	0.039	0.041	0.044	0.047	0.049	0.052
4	0.043	0.035	0.036	0.038	0.041	0.043	0.046	0.049	0.052	0.054
5	0.042	0.037	0.038	0.041	0.043	0.045	0.048	0.051	0.054	0.056
6	0.041	0.039	0.041	0.043	0.045	0.048	0.051	0.054	0.057	0.059
7	0.041	0.042	0.043	0.046	0.048	0.050	0.053	0.057	0.060	0.061
8	0.042	0.044	0.046	0.048	0.051	0.054	0.057	0.059	0.062	0.063
9	0.045	0.048	0.049	0.052	0.054	0.057	0.059	0.062	0.064	0.066

Tabelle A.48: Mit der Prozessgleichung (8.17) wurden 100 $AR(2)$ Zeitreihen erzeugt, welche die in Tabelle 8.7 aufgeführten Ausreißer enthalten. Für jede der Zeitreihen wurden mit dem modifizierten Algorithmus von CHEN und LIU die Prozessparameter und Ausreißer geschätzt. Dargestellt ist der Mittelwert der damit berechneten 100 *AIC* nach Gleichung (6.3).

	q = 0	1	2	3	4	5	6	7	8	9
p = 0	1.085	0.614	0.607	0.532	0.456	0.393	0.347	0.311	0.287	0.273
1	0.137	0.099	0.104	0.112	0.119	0.125	0.133	0.142	0.148	0.156
2	0.106	0.103	0.108	0.115	0.123	0.129	0.137	0.145	0.154	0.161
3	0.113	0.109	0.114	0.122	0.130	0.137	0.144	0.152	0.160	0.167
4	0.119	0.117	0.122	0.129	0.137	0.144	0.151	0.160	0.167	0.174
5	0.124	0.123	0.130	0.137	0.144	0.151	0.159	0.167	0.174	0.182
6	0.126	0.130	0.137	0.144	0.151	0.159	0.166	0.174	0.182	0.189
7	0.132	0.138	0.144	0.152	0.159	0.166	0.174	0.182	0.190	0.196
8	0.138	0.145	0.152	0.159	0.167	0.174	0.182	0.190	0.197	0.203
9	0.146	0.154	0.160	0.167	0.175	0.182	0.190	0.197	0.204	0.211

Tabelle A.49: Mit der Prozessgleichung (8.17) wurden 100 $AR(2)$ Zeitreihen erzeugt, welche die in Tabelle 8.7 aufgeführten Ausreißer enthalten. Für jede der Zeitreihen wurden mit dem modifizierten Algorithmus von CHEN und LIU die Prozessparameter und Ausreißer geschätzt. Dargestellt ist der Mittelwert der damit berechneten 100 BIC nach Gleichung (6.3).

	q = 0	1	2	3	4	5	6	7	8	9
p = 0	1.001	0.563	0.565	0.488	0.410	0.344	0.294	0.256	0.228	0.211
1	0.097	0.058	0.060	0.064	0.069	0.072	0.077	0.082	0.085	0.090
2	0.065	0.059	0.061	0.065	0.069	0.073	0.078	0.083	0.088	0.092
3	0.069	0.062	0.064	0.069	0.074	0.077	0.082	0.087	0.091	0.096
4	0.072	0.066	0.068	0.073	0.078	0.081	0.086	0.091	0.096	0.100
5	0.073	0.069	0.073	0.077	0.081	0.086	0.090	0.095	0.100	0.104
6	0.073	0.074	0.077	0.082	0.086	0.090	0.095	0.100	0.104	0.108
7	0.075	0.078	0.081	0.086	0.090	0.094	0.099	0.105	0.110	0.112
8	0.079	0.083	0.086	0.090	0.095	0.099	0.104	0.109	0.113	0.116
9	0.084	0.088	0.091	0.096	0.100	0.104	0.109	0.113	0.117	0.121

Tabelle A.50: Mit der Prozessgleichung (8.17) wurden 100 $AR(2)$ Zeitreihen erzeugt, welche die in Tabelle 8.7 aufgeführten Ausreißer enthalten. Für jede der Zeitreihen wurden mit dem modifizierten Algorithmus von CHEN und LIU die Prozessparameter und Ausreißer geschätzt. Dargestellt ist der Mittelwert der damit berechneten 100 HQC nach Gleichung (6.3).

	q = 0	1	2	3	4	5	6	7	8	9
p = 0	0.949	0.537	0.550	0.476	0.401	0.339	0.292	0.256	0.232	0.218
1	0.077	0.043	0.047	0.055	0.062	0.068	0.076	0.085	0.091	0.099
2	0.049	0.047	0.052	0.058	0.066	0.073	0.081	0.089	0.097	0.104
3	0.057	0.053	0.058	0.065	0.073	0.080	0.088	0.096	0.103	0.111
4	0.062	0.060	0.065	0.073	0.080	0.087	0.095	0.103	0.111	0.118
5	0.067	0.066	0.073	0.080	0.087	0.094	0.102	0.110	0.118	0.125
6	0.070	0.074	0.080	0.087	0.094	0.102	0.110	0.118	0.126	0.132
7	0.075	0.081	0.087	0.095	0.102	0.109	0.117	0.126	0.134	0.140
8	0.082	0.089	0.095	0.102	0.110	0.117	0.125	0.133	0.141	0.147
9	0.090	0.097	0.103	0.110	0.118	0.125	0.133	0.140	0.147	0.154

Tabelle A.51: Mit der Prozessgleichung (8.17) wurden 100 $AR(2)$ Zeitreihen erzeugt, welche die in Tabelle 8.7 aufgeführten Ausreißer enthalten. Für jede der Zeitreihen wurden mit dem modifizierten Algorithmus von CHEN und LIU die Prozessparameter und Ausreißer geschätzt. Dargestellt ist der Mittelwert der damit berechneten 100 *MixedC* nach Gleichung (6.3).

	q = 0	1	2	3	4	5	6	7	8	9
p = 0	1.448	1.288	1.146	1.044	0.968	0.915	0.878	0.853	0.836	0.823
1	0.783	0.750	0.751	0.754	0.756	0.758	0.760	0.762	0.765	0.767
2	0.748	0.750	0.752	0.755	0.757	0.759	0.761	0.764	0.766	0.768
3	0.750	0.752	0.755	0.757	0.759	0.761	0.764	0.766	0.769	0.771
4	0.752	0.755	0.757	0.759	0.762	0.764	0.766	0.769	0.771	0.773
5	0.755	0.757	0.759	0.762	0.764	0.766	0.769	0.771	0.773	0.776
6	0.757	0.759	0.761	0.764	0.766	0.769	0.771	0.773	0.776	0.778
7	0.759	0.761	0.764	0.766	0.769	0.771	0.773	0.776	0.778	0.780
8	0.761	0.764	0.766	0.769	0.771	0.773	0.776	0.778	0.781	0.783
9	0.764	0.766	0.768	0.771	0.773	0.776	0.778	0.780	0.783	0.785

Tabelle A.52: Mit der Prozessgleichung (8.17) wurden 100 $AR(2)$ Zeitreihen erzeugt, welche die in Tabelle 8.7 aufgeführten Ausreißer enthalten. Für jede der Zeitreihen wurden die Prozessparameter nach der Methode der kleinsten Quadrate geschätzt. Dargestellt ist der Mittelwert der damit berechneten 100 *AIC* nach Gleichung (6.1).

	q = 0	1	2	3	4	5	6	7	8	9
p = 0	1.448	1.292	1.156	1.058	0.988	0.939	0.908	0.888	0.875	0.868
1	0.788	0.759	0.766	0.773	0.780	0.787	0.794	0.802	0.809	0.816
2	0.758	0.765	0.772	0.779	0.786	0.793	0.801	0.808	0.815	0.822
3	0.765	0.772	0.779	0.786	0.794	0.801	0.808	0.815	0.823	0.830
4	0.772	0.779	0.786	0.794	0.801	0.808	0.815	0.823	0.830	0.837
5	0.779	0.786	0.794	0.801	0.808	0.815	0.823	0.830	0.837	0.844
6	0.786	0.793	0.801	0.808	0.815	0.823	0.830	0.837	0.844	0.852
7	0.794	0.801	0.808	0.815	0.823	0.830	0.837	0.844	0.852	0.859
8	0.801	0.808	0.815	0.823	0.830	0.837	0.844	0.852	0.859	0.866
9	0.808	0.815	0.822	0.830	0.837	0.844	0.852	0.859	0.866	0.874

Tabelle A.53: Mit der Prozessgleichung (8.17) wurden 100 $AR(2)$ Zeitreihen erzeugt, welche die in Tabelle 8.7 aufgeführten Ausreißer enthalten. Für jede der Zeitreihen wurden die Prozessparameter nach der Methode der kleinsten Quadrate geschätzt. Dargestellt ist der Mittelwert der damit berechneten 100 BIC nach Gleichung (6.1).

	q = 0	1	2	3	4	5	6	7	8	9
p = 0	1.448	1.289	1.150	1.049	0.976	0.924	0.889	0.866	0.851	0.840
1	0.785	0.753	0.757	0.761	0.765	0.769	0.773	0.777	0.781	0.786
2	0.752	0.756	0.760	0.764	0.768	0.772	0.776	0.781	0.785	0.789
3	0.756	0.760	0.764	0.768	0.772	0.776	0.781	0.785	0.789	0.793
4	0.760	0.764	0.768	0.772	0.777	0.781	0.785	0.789	0.793	0.798
5	0.764	0.768	0.772	0.777	0.781	0.785	0.789	0.793	0.798	0.802
6	0.768	0.772	0.776	0.781	0.785	0.789	0.793	0.798	0.802	0.806
7	0.772	0.776	0.781	0.785	0.789	0.793	0.798	0.802	0.806	0.810
8	0.776	0.780	0.785	0.789	0.793	0.798	0.802	0.806	0.810	0.815
9	0.781	0.785	0.789	0.793	0.797	0.802	0.806	0.810	0.815	0.819

Tabelle A.54: Mit der Prozessgleichung (8.17) wurden 100 $AR(2)$ Zeitreihen erzeugt, welche die in Tabelle 8.7 aufgeführten Ausreißer enthalten. Für jede der Zeitreihen wurden die Prozessparameter nach der Methode der kleinsten Quadrate geschätzt. Dargestellt ist der Mittelwert der damit berechneten 100 BIC nach Gleichung (6.1).

	q = 0	1	2	3	4	5	6	7	8	9
p = 0	1.448	1.292	1.156	1.058	0.988	0.939	0.908	0.888	0.875	0.868
1	0.788	0.759	0.766	0.773	0.780	0.787	0.794	0.802	0.809	0.816
2	0.758	0.765	0.772	0.779	0.786	0.793	0.801	0.808	0.815	0.822
3	0.765	0.772	0.779	0.786	0.794	0.801	0.808	0.815	0.823	0.830
4	0.772	0.779	0.786	0.794	0.801	0.808	0.815	0.823	0.830	0.837
5	0.779	0.786	0.794	0.801	0.808	0.815	0.823	0.830	0.837	0.844
6	0.786	0.793	0.801	0.808	0.815	0.823	0.830	0.837	0.844	0.852
7	0.794	0.801	0.808	0.815	0.823	0.830	0.837	0.844	0.852	0.859
8	0.801	0.808	0.815	0.823	0.830	0.837	0.844	0.852	0.859	0.866
9	0.808	0.815	0.822	0.830	0.837	0.844	0.852	0.859	0.866	0.874

Tabelle A.55: Mit der Prozessgleichung (8.17) wurden 100 $AR(2)$ Zeitreihen erzeugt, welche die in Tabelle 8.7 aufgeführten Ausreißer enthalten. Für jede der Zeitreihen wurden die Prozessparameter nach der Methode der kleinsten Quadrate geschätzt. Dargestellt ist der Mittelwert der damit berechneten 100 $MixedC$ nach Gleichung (6.1).

	q = 0	1	2	3	4	5	6	7	8	9
p = 0	0.396	0.109	0.051	0.055	0.058	0.061	0.063	0.064	0.067	0.070
1	0.076	0.038	0.044	0.039	0.043	0.045	0.049	0.052	0.055	0.057
2	0.048	0.039	0.041	0.041	0.044	0.046	0.050	0.052	0.054	0.057
3	0.046	0.040	0.042	0.044	0.046	0.048	0.052	0.054	0.057	0.060
4	0.043	0.040	0.043	0.045	0.048	0.051	0.053	0.056	0.059	0.062
5	0.043	0.042	0.046	0.049	0.051	0.053	0.056	0.058	0.061	0.064
6	0.044	0.045	0.049	0.050	0.054	0.056	0.058	0.060	0.063	0.066
7	0.046	0.047	0.051	0.054	0.056	0.058	0.061	0.064	0.065	0.069
8	0.048	0.050	0.054	0.056	0.059	0.061	0.064	0.066	0.067	0.071
9	0.052	0.053	0.057	0.059	0.061	0.064	0.067	0.068	0.070	0.074

Tabelle A.56: Mit der Prozessgleichung (8.18) wurden 100 $MA(2)$ Zeitreihen erzeugt, welche die in Tabelle 8.7 aufgeführten Ausreißer enthalten. Für jede der Zeitreihen wurden mit dem modifizierten Algorithmus von CHEN und LIU die Prozessparameter und Ausreißer geschätzt. Dargestellt ist der Mittelwert der damit berechneten 100 AIC nach Gleichung (6.3).

	q = 0	1	2	3	4	5	6	7	8	9
p = 0	0.469	0.180	0.118	0.128	0.135	0.143	0.149	0.156	0.164	0.171
1	0.137	0.106	0.117	0.117	0.125	0.133	0.142	0.148	0.156	0.164
2	0.116	0.112	0.118	0.124	0.131	0.139	0.148	0.154	0.161	0.169
3	0.119	0.117	0.124	0.131	0.138	0.146	0.154	0.161	0.168	0.176
4	0.121	0.122	0.130	0.137	0.145	0.153	0.161	0.168	0.175	0.183
5	0.125	0.130	0.139	0.146	0.153	0.160	0.167	0.175	0.183	0.190
6	0.131	0.137	0.146	0.152	0.161	0.168	0.174	0.181	0.189	0.197
7	0.138	0.144	0.153	0.161	0.169	0.175	0.182	0.190	0.196	0.205
8	0.144	0.151	0.160	0.167	0.176	0.183	0.190	0.197	0.204	0.212
9	0.153	0.159	0.168	0.175	0.183	0.190	0.198	0.204	0.212	0.220

Tabelle A.57: Mit der Prozessgleichung (8.18) wurden 100 $MA(2)$ Zeitreihen erzeugt, welche die in Tabelle 8.7 aufgeführten Ausreißer enthalten. Für jede der Zeitreihen wurden mit dem modifizierten Algorithmus von CHEN und LIU die Prozessparameter und Ausreißer geschätzt. Dargestellt ist der Mittelwert der damit berechneten 100 BIC nach Gleichung (6.3).

	q = 0	1	2	3	4	5	6	7	8	9
p = 0	0.424	0.136	0.077	0.083	0.087	0.092	0.096	0.099	0.104	0.108
1	0.099	0.064	0.072	0.069	0.074	0.079	0.084	0.088	0.093	0.098
2	0.074	0.067	0.070	0.073	0.077	0.082	0.087	0.091	0.095	0.100
3	0.074	0.069	0.073	0.077	0.081	0.085	0.091	0.095	0.099	0.104
4	0.073	0.071	0.076	0.080	0.085	0.090	0.094	0.099	0.103	0.108
5	0.074	0.075	0.081	0.086	0.090	0.094	0.098	0.103	0.107	0.112
6	0.077	0.080	0.086	0.089	0.094	0.098	0.102	0.106	0.110	0.116
7	0.081	0.084	0.090	0.094	0.099	0.103	0.107	0.112	0.115	0.121
8	0.085	0.088	0.094	0.098	0.104	0.107	0.112	0.116	0.119	0.125
9	0.090	0.093	0.099	0.103	0.107	0.112	0.117	0.120	0.124	0.130

Tabelle A.58: Mit der Prozessgleichung (8.18) wurden 100 $MA(2)$ Zeitreihen erzeugt, welche die in Tabelle 8.7 aufgeführten Ausreißer enthalten. Für jede der Zeitreihen wurden mit dem modifizierten Algorithmus von CHEN und LIU die Prozessparameter und Ausreißer geschätzt. Dargestellt ist der Mittelwert der damit berechneten 100 HQC nach Gleichung (6.3).

	q = 0	1	2	3	4	5	6	7	8	9
p = 0	0.396	0.114	0.061	0.070	0.077	0.086	0.092	0.099	0.106	0.114
1	0.081	0.048	0.059	0.059	0.067	0.075	0.084	0.091	0.099	0.106
2	0.058	0.054	0.060	0.066	0.073	0.081	0.089	0.096	0.103	0.111
3	0.061	0.059	0.066	0.073	0.080	0.087	0.096	0.103	0.111	0.118
4	0.063	0.064	0.073	0.079	0.087	0.095	0.103	0.110	0.118	0.126
5	0.067	0.072	0.081	0.088	0.095	0.102	0.110	0.117	0.125	0.133
6	0.073	0.079	0.088	0.094	0.103	0.110	0.117	0.124	0.131	0.140
7	0.080	0.087	0.095	0.103	0.110	0.117	0.125	0.132	0.139	0.147
8	0.087	0.094	0.103	0.110	0.118	0.125	0.133	0.139	0.146	0.154
9	0.096	0.102	0.111	0.117	0.125	0.132	0.140	0.147	0.154	0.163

Tabelle A.59: Mit der Prozessgleichung (8.18) wurden 100 $MA(2)$ Zeitreihen erzeugt, welche die in Tabelle 8.7 aufgeführten Ausreißer enthalten. Für jede der Zeitreihen wurden mit dem modifizierten Algorithmus von CHEN und LIU die Prozessparameter und Ausreißer geschätzt. Dargestellt ist der Mittelwert der damit berechneten 100 *MixedC* nach Gleichung (6.3).

	q = 0	1	2	3	4	5	6	7	8	9
p = 0	1.067	0.823	0.774	0.771	0.771	0.772	0.773	0.775	0.777	0.779
1	0.761	0.758	0.761	0.762	0.764	0.766	0.768	0.771	0.773	0.775
2	0.758	0.760	0.762	0.764	0.766	0.768	0.770	0.773	0.775	0.777
3	0.760	0.762	0.764	0.766	0.768	0.771	0.773	0.775	0.777	0.780
4	0.761	0.764	0.766	0.768	0.771	0.773	0.775	0.778	0.780	0.782
5	0.763	0.766	0.768	0.771	0.773	0.775	0.778	0.780	0.783	0.785
6	0.766	0.768	0.770	0.773	0.775	0.778	0.780	0.782	0.785	0.787
7	0.768	0.771	0.773	0.775	0.778	0.780	0.782	0.785	0.800	0.790
8	0.770	0.773	0.775	0.778	0.780	0.782	0.785	0.787	0.798	0.792
9	0.773	0.775	0.777	0.780	0.783	0.788	0.796	0.793	0.795	0.794

Tabelle A.60: Mit der Prozessgleichung (8.18) wurden 100 $MA(2)$ Zeitreihen erzeugt, welche die in Tabelle 8.7 aufgeführten Ausreißer enthalten. Für jede der Zeitreihen wurden die Prozessparameter nach der Methode der kleinsten Quadrate geschätzt. Dargestellt ist der Mittelwert der damit berechneten 100 *AIC* nach Gleichung (6.1).

	q = 0	1	2	3	4	5	6	7	8	9
p = 0	1.067	0.828	0.784	0.786	0.790	0.796	0.802	0.809	0.816	0.823
1	0.766	0.768	0.775	0.782	0.789	0.796	0.803	0.810	0.817	0.824
2	0.767	0.775	0.782	0.789	0.796	0.803	0.810	0.817	0.824	0.831
3	0.775	0.782	0.789	0.795	0.803	0.810	0.817	0.824	0.831	0.839
4	0.781	0.788	0.795	0.803	0.810	0.817	0.824	0.832	0.839	0.846
5	0.788	0.795	0.803	0.810	0.817	0.824	0.832	0.839	0.846	0.854
6	0.795	0.802	0.810	0.817	0.824	0.832	0.839	0.846	0.854	0.861
7	0.802	0.810	0.817	0.824	0.832	0.839	0.846	0.853	0.874	0.868
8	0.810	0.817	0.824	0.832	0.839	0.846	0.854	0.861	0.876	0.875
9	0.817	0.824	0.831	0.839	0.846	0.856	0.870	0.872	0.879	0.883

Tabelle A.61: Mit der Prozessgleichung (8.18) wurden 100 $MA(2)$ Zeitreihen erzeugt, welche die in Tabelle 8.7 aufgeführten Ausreißer enthalten. Für jede der Zeitreihen wurden die Prozessparameter nach der Methode der kleinsten Quadrate geschätzt. Dargestellt ist der Mittelwert der damit berechneten 100 BIC nach Gleichung (6.1).

	q = 0	1	2	3	4	5	6	7	8	9
p = 0	1.067	0.825	0.778	0.777	0.778	0.781	0.784	0.788	0.792	0.796
1	0.763	0.762	0.766	0.769	0.773	0.777	0.782	0.786	0.790	0.794
2	0.761	0.766	0.770	0.774	0.777	0.781	0.785	0.790	0.794	0.798
3	0.766	0.770	0.773	0.777	0.781	0.786	0.790	0.794	0.798	0.802
4	0.769	0.773	0.777	0.781	0.786	0.790	0.794	0.798	0.802	0.806
5	0.773	0.777	0.781	0.786	0.790	0.794	0.798	0.802	0.807	0.811
6	0.777	0.781	0.785	0.790	0.794	0.798	0.802	0.807	0.811	0.815
7	0.781	0.785	0.790	0.794	0.798	0.802	0.807	0.811	0.828	0.819
8	0.785	0.790	0.794	0.798	0.802	0.807	0.811	0.815	0.828	0.824
9	0.789	0.794	0.798	0.802	0.807	0.814	0.824	0.823	0.827	0.828

Tabelle A.62: Mit der Prozessgleichung (8.18) wurden 100 $MA(2)$ Zeitreihen erzeugt, welche die in Tabelle 8.7 aufgeführten Ausreißer enthalten. Für jede der Zeitreihen wurden die Prozessparameter nach der Methode der kleinsten Quadrate geschätzt. Dargestellt ist der Mittelwert der damit berechneten 100 HQC nach Gleichung (6.1).

	q = 0	1	2	3	4	5	6	7	8	9
p = 0	1.067	0.828	0.784	0.786	0.790	0.796	0.802	0.809	0.816	0.823
1	0.766	0.768	0.775	0.782	0.789	0.796	0.803	0.810	0.817	0.824
2	0.767	0.775	0.782	0.789	0.796	0.803	0.810	0.817	0.824	0.831
3	0.775	0.782	0.789	0.795	0.803	0.810	0.817	0.824	0.831	0.839
4	0.781	0.788	0.795	0.803	0.810	0.817	0.824	0.832	0.839	0.846
5	0.788	0.795	0.803	0.810	0.817	0.824	0.832	0.839	0.846	0.854
6	0.795	0.802	0.810	0.817	0.824	0.832	0.839	0.846	0.854	0.861
7	0.802	0.810	0.817	0.824	0.832	0.839	0.846	0.853	0.874	0.868
8	0.810	0.817	0.824	0.832	0.839	0.846	0.854	0.861	0.876	0.875
9	0.817	0.824	0.831	0.839	0.846	0.856	0.870	0.872	0.879	0.883

Tabelle A.63: Mit der Prozessgleichung (8.18) wurden 100 $MA(2)$ Zeitreihen erzeugt, welche die in Tabelle 8.7 aufgeführten Ausreißer enthalten. Für jede der Zeitreihen wurden die Prozessparameter nach der Methode der kleinsten Quadrate geschätzt. Dargestellt ist der Mittelwert der damit berechneten 100 $MixedC$ nach Gleichung (6.1).

Anhang B

Funktionen und Variablen in MATLAB

Im Nachfolgenden sind die MATLAB Funktionen und MATLAB Variablen dokumentiert. Zum Teil sind in verschiedenen Verzeichnissen die gleichen Funktionen zu finden, die jedoch anders programmiert wurden. Daher ist zwischen den Funktionen in den verschiedenen Verzeichnissen zu unterscheiden. Die Überschriften der Abschnitte entsprechen den Namen der Verzeichnisse, in denen die Funktionen gespeichert sind. Es sollen nur die wesentlichen Funktionen und Variablen erläutert werden.

B.1 Algorithmus von Chen und Liu

Beispielcode

```
% Länge der Zeitreihe festsetzen:
```

```
T = 1000;
```

```
% Lag-Polynome festsetzen (AR(2) Prozess mit  $\varphi_1 = 0.6$  und  $\varphi_2 = 0.2$ ):
```

```
LagAR = [0.6 0.2]';
```

```
LagMA = []';
```

```
% Die Ordnung des Prozesses  $p = 2$  und  $q = 0$  sei bekannt:
```

```
lagsAR = 2;
```

```
lagsMA = 0;
```

```
% Varianz der Residuen beträgt  $\sigma_\varepsilon^2 = 1$ :
```

```
variance = 1.0;
```

```
% Zeitreihe erzeugen:
```

```
x = ARMA( LagAR, LagMA, variance, T );
```

```
% Innovativen Ausreißer beim Zeitpunkt t = 100 mit der Größe  $\omega = 9$  einfügen:
```

```
x = arma_addIO( x, LagAR, LagMA, 100, 9.0 );
```

```
% Additiver Ausreißer beim Zeitpunkt t = 300 mit der Größe  $\omega = -8$  einfügen:
```

```
x = arma_addAO( x, 300, -8.0 );
```

```
% Level-Shift beim Zeitpunkt t = 400 mit der Größe  $\omega = 3$  einfügen:
```

```
x = arma_addLS( x, 400, 3.0 );
```

```
% Zeitliche Änderung mit  $\delta = 0.7$  beim Zeitpunkt t = 500 mit der Größe  $\omega = -10$  einfügen:
```

```
x = arma_addTC( x, 500, 0.7, -10.0 );
```

```
% Algorithmus von Chen und Liu starten:
```

```
[ theta, Outlier ] = arma_detectionProcedure( x, lagsAR, lagsMA );
```

Liste der Variablen

- **D**
Design-Matrix zum gemeinsamen Schätzen der detektierten Ausreißer
- **delta**
Variable δ , die die Länge des Abklingens einer zeitlichen Änderung steuert
- **e**
Residuen ε_t des $ARMA(p,q)$ Prozesses
- **exists**
Boolsche Variable, die angibt, ob ein Ausreißer gefunden wurde oder nicht
- **gradient**
Gradient $\nabla \mathcal{L}(\boldsymbol{\theta})$ der Log-Likelihood Funktion $\mathcal{L}(\boldsymbol{\theta})$
- **LagAR**
Kofaktoren $\varphi_1, \varphi_2, \dots, \varphi_p$ des Lag-Polynoms $\varphi(B)$
- **LagMA**
Kofaktoren $\theta_1, \theta_2, \dots, \theta_q$ des Lag-Polynoms $\varphi(B)$

- **LagPi**
Kofaktoren $\pi_1, \pi_2, \dots, \pi_p$ des Lag-Polynoms $\pi(B) = \theta(B)/\varphi(B)$
- **LagPsi**
Kofaktoren $\psi_1, \psi_2, \dots, \psi_p$ des Lag-Polynoms $\psi(B) = \varphi(B)/\theta(B)$
- **lagsAR**
Ordnung p des $ARMA(p, q)$ Prozesses
- **lagsMA**
Ordnung q des $ARMA(p, q)$ Prozesses
- **magnitude**
Größe ω eines Ausreißers
- **Outlier**
Matrix, die Zeit, Typ, Größe und Signifikanz der detektierten Ausreißer in den vier Spalten speichert
- **outlierCount**
Variable, die die Anzahl der detektierten Ausreißer speichert
- **size**
Größe ω eines Ausreißers
- **T**
Länge der Zeitreihe y_t
- **theta**
Parametervektor $\theta = (\varphi_1, \varphi_2, \dots, \varphi_p, \theta_1, \theta_2, \dots, \theta_q, \sigma_\varepsilon^2)'$
- **time**
Zeitpunkt t_a eines Ausreißers
- **type**
Ausreißertyp
- **var**
Varianz σ_ε^2 der Residuen ε_t
- **vecA0**
Muster x_A eines additiven Ausreißers in den Residuen
- **vecLS**
Muster x_L eines Level-Shifts in den Residuen
- **vecTC**
Muster x_Z einer zeitlichen Änderung in den Residuen

- x
Zeitreihe y_t

Liste der Funktionen

- `ARMA` erzeugt eine Zeitreihe, die einem $ARMA(p,q)$ Prozess folgt:
[x] = `ARMA(LagAR, LagMA, var, T)`
- `arma_addAO` fügt der Zeitreihe einen additiven Ausreißer hinzu:
[x] = `arma_addAO(x, time, magnitude)`
- `arma_addIO` fügt der Zeitreihe einen innovativen Ausreißer hinzu:
[x] = `arma_addIO(x, LagAR, LagMA, time, magnitude)`
- `arma_addLS` fügt der Zeitreihe einen Level-Shift hinzu:
[x] = `arma_addLS(x, time, magnitude)`
- `arma_addTC` fügt der Zeitreihe eine zeitliche Änderung hinzu:
[x] = `arma_addTC(x, time, delta, magnitude)`
- `arma_calcGradientNum` berechnet den Gradienten $\nabla \mathcal{L}(\theta)$ der Log-Likelihood Funktion $\mathcal{L}(\theta)$ numerisch:
[`gradient`] = `arma_calcGradientNum(x, theta, lagsAR, lagsMA)`
- `arma_calcLagPi` berechnet die Polynomdivision $\theta(B)/\varphi(B)$:
[`LagPi`] = `arma_calcLagPi(LagAR, LagMA)`
- `arma_calcLagPsi` berechnet die Polynomdivision $\varphi(B)/\theta(B)$:
[`LagPsi`] = `arma_calcLagPsi(LagAR, LagMA)`
- `arma_calcLogL` berechnet den Wert der Log-Likelihood Funktion $\mathcal{L}(\theta)$ für ein vorgegebenes θ :
[`logL`] = `arma_calcLogL(x, theta, lagsAR, lagsMA)`
- `arma_calcMaxOutlier` findet den signifikantesten Ausreißer und schätzt seine Größe:
[`Outlier, outlierCount, exists`] = `arma_calcMaxOutlier(e, Outlier, outlierCount, theta, lagsAR, lagsMA)`
- `arma_calcResiduals` berechnet die Residuen für ein vorgegebenes θ :
[`e`] = `arma_calcResiduals(x, theta, lagsAR, lagsMA)`
- `arma_calcVecAO` berechnet das Muster eines additiven Ausreißers in den Residuen:
[`vecAO`] = `arma_calcVecAO(theta, lagsAR, lagsMA)`

- `arma_calcVecLS` berechnet das Muster eines Level-Shifts in den Residuen:
`[vecLS] = arma_calcVecLS(theta, lagsAR, lagsMA)`
- `arma_calcVecTC` berechnet das Muster einer zeitlichen Änderung in den Residuen:
`[vecTC] = arma_calcVecTC(theta, lagsAR, lagsMA)`
- `arma_checkForMultipleDetection` prüft, ob Ausreißer mehrfach detektiert wurden und behebt die mehrfache Detektion:
`[Outlier, outlierCount] = arma_checkForMultipleDetection(Outlier, outlierCount)`
- `arma_createOutlierDesignMatrix` stellt die Design-Matrix zum gemeinsamen Schätzen der Ausreißer auf:
`[D] = arma_createOutlierDesignMatrix(x, theta, lagsAR, lagsMA, Outlier)`
- `arma_detectionProcedure` führt den Algorithmus von CHEN und LIU durch:
`[theta, Outlier] = arma_detectionProcedure(x, lagsAR, lagsMA)`
- `arma_removeAllOutliers` entfernt alle detektierten Ausreißer aus der Zeitreihe:
`[x] = arma_removeAllOutliers(x, theta, lagsAR, lagsMA, Outlier)`
- `arma_removeAO` entfernt einen additiven Ausreißer aus der Zeitreihe und den Residuen:
`[x, e] = arma_removeAO(x, e, theta, lagsAR, lagsMA, time, size)`
- `arma_removeIO` entfernt einen innovativen Ausreißer aus der Zeitreihe und den Residuen:
`[x, e] = arma_removeIO(x, e, theta, lagsAR, lagsMA, time, size)`
- `arma_removeLS` entfernt einen Level-Shift aus der Zeitreihe und den Residuen:
`[x, e] = arma_removeLS(x, e, theta, lagsAR, lagsMA, time, size)`
- `arma_removeOutlier` entfernt einen Ausreißer aus der Zeitreihe und den Residuen:
`[x, e] = arma_removeOutlier(x, e, theta, lagsAR, lagsMA, latestOutlier)`
- `arma_removeTC` entfernt eine zeitliche Änderung aus der Zeitreihe und den Residuen:
`[x, e] = arma_removeTC(x, e, theta, lagsAR, lagsMA, time, delta, size)`
- `arma_steepestAscendGridNum` sucht das Maximum der Log-Likelihood Funktion $\mathcal{L}(\theta)$ nach dem Verfahren des steilsten Anstiegs mit numerischer Berechnung des Gradienten $\nabla \mathcal{L}(\theta)$ der Log-Likelihood Funktion $\mathcal{L}(\theta)$:
`[theta] = arma_steepestAscendGridNum(x, theta, lagsAR, lagsMA)`

B.2 Modifizierter Algorithmus von Chen und Liu

Beispielcode

```

% Länge der Zeitreihe festsetzen:
T = 1000;

% Lag-Polynome festsetzen (AR(2) Prozess mit  $\varphi_1 = 0.6$  und  $\varphi_2 = 0.2$ ):
LagAR = [0.6 0.2]';
LagMA = []';

% Die Ordnung des Prozesses  $p = 2$  und  $q = 0$  sei bekannt:
lagsAR = 2;
lagsMA = 0;

% Varianz der Residuen beträgt  $\sigma_\varepsilon^2 = 1$ :
variance = 1.0;

% Schwellenwert, bei dem die Berechnung der Ausreißermuster abgebrochen wird:
epsilon = 0.01;

% Zeitreihe erzeugen:
x = ARMA( LagAR, LagMA, variance, T );

% Innovativen Ausreißer beim Zeitpunkt  $t = 100$  mit der Größe  $\omega = 9$  einfügen:
% (Die Werte des Ausreißermusters sind größer als  $0.01\sigma_\varepsilon$ .)
x = arma_addOutlier( x, LagAR, LagMA, epsilon*sqrt(variance), 100, 1, 9.0 );

% Additiver Ausreißer beim Zeitpunkt  $t = 300$  mit der Größe  $\omega = -8$  einfügen:
x = arma_addOutlier( x, LagAR, LagMA, epsilon*sqrt(variance), 300, 2, -8.0 );

% Zeitliche Änderung mit  $\delta = 0.7$  beim Zeitpunkt  $t = 500$  mit der Größe  $\omega = -10$  einfügen:
x = arma_addOutlier( x, LagAR, LagMA, epsilon*sqrt(variance), 500, 3, -10.0 );

% Algorithmus von Chen und Liu starten:
[ theta, Outlier ] = arma_detectionProcedure( x, lagsAR, lagsMA );

```

Liste der Variablen

- **convAO**
Faltung des Musters eines additiven Ausreißers mit den Residuen
- **convTC**
Faltung des Musters einer zeitlichen Änderung mit den Residuen
- **criticalValue**
Kritischer Wert für die Signifikanz von Ausreißern
- **e**
Residuen ε_t des $ARMA(p,q)$ Prozesses
- **epsilon**
Schwellenwert für den Abbruch der Polynomdivisionen $\theta(B)/\varphi(B)$ und $\varphi(B)/\theta(B)$ sowie für den Abbruch bei der Berechnung der Ausreißermuster
- **exists**
Boolsche Variable, die angibt, ob ein Ausreißer gefunden wurde oder nicht
- **isInvertible**
Boolsche Variable, die angibt, ob der $AR(p)$ Anteil und der $MA(q)$ Anteil des $ARMA(p,q)$ Prozesses invertierbar sind oder nicht
- **LagAR**
Kofaktoren $\varphi_1, \varphi_2, \dots, \varphi_p$ des Lag-Polynoms $\varphi(B)$
- **LagMA**
Kofaktoren $\theta_1, \theta_2, \dots, \theta_q$ des Lag-Polynoms $\varphi(B)$
- **LagPi**
Kofaktoren $\pi_1, \pi_2, \dots, \pi_p$ des Lag-Polynoms $\pi(B) = \theta(B)/\varphi(B)$
- **LagPsi**
Kofaktoren $\psi_1, \psi_2, \dots, \psi_p$ des Lag-Polynoms $\psi(B) = \varphi(B)/\theta(B)$
- **lagsAR**
Ordnung p des $ARMA(p,q)$ Prozesses
- **lagsMA**
Ordnung q des $ARMA(p,q)$ Prozesses
- **magnitude**
Größe ω eines Ausreißers
- **minimum**
Minimale Größe ω eines Ausreißers, mit dem ein Wert der Zeitreihe verunreinigt wird

- `n`
Rechte Seite der Normalgleichungen der gemeinsamen Schätzung der Ausreißer
- `N`
Normalgleichungsmatrix der gemeinsamen Schätzung der Ausreißer
- `n_ar`
Rechte Seite der Normalgleichungen des langen $AR(p)$ Prozesses
- `N_ar`
Normalgleichungsmatrix des langen $AR(p)$ Prozesses
- `orderLongAR`
Ordnung des langen $AR(p)$ Prozesses
- `Outlier`
Matrix, die Zeit, Typ, Größe und Signifikanz der detektierten Ausreißer in den vier Spalten speichert
- `patternAO`
Muster x_A eines additiven Ausreißers in den Residuen
- `patternInTimeSeriesIO`
Muster \bar{x}_I eines innovativen Ausreißers in der Zeitreihe
- `patternInTimeSeriesTC`
Muster \bar{x}_Z einer zeitlichen Änderung in der Zeitreihe
- `patternTC`
Muster x_Z einer zeitlichen Änderung in den Residuen
- `percentage`
Prozentsatz der Werte, die mit Ausreißern verunreinigt werden
- `quadSumAO`
Quadratische L_2 -Norm des Ausreißermusters eines additiven Ausreißers in den Residuen
- `quadSumTC`
Quadratische L_2 -Norm des Ausreißermusters einer zeitlichen Änderung in den Residuen
- `stdDev`
Standardabweichung der Größe ω eines Ausreißers, mit dem ein Wert der Zeitreihe verunreinigt wird
- `T`
Länge der Zeitreihe y_t

- **theta**
Parametervektor $\theta = (\varphi_1, \varphi_2, \dots, \varphi_p, \theta_1, \theta_2, \dots, \theta_q, \sigma_\varepsilon^2)'$
- **time**
Zeitpunkt t_a eines Ausreißers
- **type**
Ausreißertyp
- **var**
Varianz σ_ε^2 der Residuen ε_t
- **x**
Zeitreihe y_t

Liste der Funktionen

- **ARMA** erzeugt eine Zeitreihe, die einem $ARMA(p,q)$ Prozess folgt:
[x] = ARMA(LagAR, LagMA, var, T)
- **arma_addOutlier** fügt der Zeitreihe einen Ausreißer hinzu:
[x] = arma_addOutlier(x, LagAR, LagMA, epsilon, time, type, magnitude)
- **arma_addOutlierPercentage** fügt einem Prozentsatz der Werte der Zeitreihe Ausreißer hinzu, so dass die Ausreißer gleichmäßig über die Zeitreihe verteilt sind:
[x] = arma_addOutlierPercentage(x, LagAR, LagMA, epsilon, percentage, minimum, stdDev)
- **arma_calcLagPi** berechnet die Polynomdivision $\theta(B)/\varphi(B)$:
[LagPi] = arma_calcLagPi(LagAR, LagMA, epsilon)
- **arma_calcLagPsi** berechnet die Polynomdivision $\varphi(B)/\theta(B)$:
[LagPsi] = arma_calcLagPsi(LagAR, LagMA, epsilon)
- **arma_calcMaxOutlier** findet und schätzt die Größe des signifikantesten Ausreißers und versieht die Faltungen mit dem entsprechenden Update:
[Outlier, exists, convA0, convTC] = arma_calcMaxOutlier(theta, Outlier, e, convA0, convTC, criticalValue, patternA0, patternTC, quadSumA0, quadSumTC)
- **arma_calcPatterns** berechnet die Ausreißermuster und die quadratische L_2 -Norm der Muster:

```
[ patternAO, patternTC, quadSumAO, quadSumTC, patternInTimeSeriesIO,
patternInTimeSeriesTC ] = arma_calcPatterns( theta, lagsAR, lagsMA,
epsilon )
```

- `arma_checkForMultipleDetection` prüft, ob zum gleichen Zeitpunkt mehrere Ausreißer detektiert wurden und behebt die mehrfache Detektion der Ausreißer:

```
[ Outlier ] = arma_checkForMultipleDetection( Outlier )
```
- `arma_checkInvertibility` prüft, ob der $AR(p)$ Anteil und der $MA(q)$ Anteil der geschätzten Parameter des $ARMA(p,q)$ Prozesses invertierbar sind:

```
[ isInvertible ] = arma_checkInvertibility( theta, lagsAR, lagsMA )
```
- `arma_convolute` faltet die Residuen mit den Ausreißermustern:

```
[ convAO, convTC ] = arma_convolute( e, patternAO, patternTC )
```
- `arma_detectionProcedure` führt den modifizierten Algorithmus von Chen und Liu durch:

```
[ theta, Outlier ] = arma_detectionProcedure( x, lagsAR, lagsMA )
```
- `arma_estimateOutlierJointly` schätzt die Ausreißer gemeinsam:

```
[ Outlier, N, n ] = arma_estimateOutlierJointly( e, Outlier, patternAO,
patternTC, quadSumAO, quadSumTC )
```
- `arma_estimation` schätzt den $ARMA(p,q)$ Prozess:

```
[ theta, e, N_ar, n_ar ] = arma_estimation( x, lagsAR, lagsMA,
orderLongAR )
```
- `arma_estimationLongAR` schätzt den $ARMA(p,q)$ Prozess, wobei die Normalgleichungen des langen $AR(p)$ Prozesses vorgegeben werden:

```
[ theta, e ] = arma_estimationLongAR( x, lagsAR, lagsMA, N_ar, n_ar )
```
- `arma_obtainResiduals` berechnet die Residuen mit vorgegebenen Parametern:

```
[ e ] = arma_obtainResiduals( x, theta, lagsAR, lagsMA )
```
- `arma_patternAO` berechnet das Muster eines additiven Ausreißers in den Residuen:

```
[ patternAO ] = arma_patternAO( LagPi )
```
- `arma_patternInTimeSeriesIO` berechnet das Muster eines innovativen Ausreißers in der Zeitreihe:

```
[ patternInTimeSeriesIO ] = arma_patternInTimeSeriesIO( LagPsi )
```
- `arma_patternInTimeSeriesTC` berechnet das Muster einer zeitlichen Änderung in der Zeitreihe:

```
[ patternInTimeSeriesTC ] = arma_patternInTimeSeriesTC( epsilon )
```

- `arma_patternTC` berechnet das Muster einer zeitlichen Änderung in den Residuen:
`[patternTC] = arma_patternTC(LagPi, epsilon)`
- `arma_removeOutlier` entfernt einen Ausreißer aus der Zeitreihe und den Residuen:
`[x, e] = arma_removeOutlier(x, e, time, type, magnitude, patternAO, patternTC, patternInTimeSeriesIO, patternInTimeSeriesTC)`
- `arma_removeOutlierFromResiduals` entfernt einen Ausreißer aus den Residuen:
`[e] = arma_removeOutlierFromResiduals(e, time, type, magnitude, patternAO, patternTC, patternInTimeSeriesIO, patternInTimeSeriesTC)`
- `arma_removeOutlierFromTimeSeries` entfernt einen Ausreißer aus der Zeitreihe:
`[x] = arma_removeOutlierFromTimeSeries(x, time, type, magnitude, patternAO, patternTC, patternInTimeSeriesIO, patternInTimeSeriesTC)`
- `arma_updateNormalEquationLongAR` versieht die Normalgleichungen mit einem Update:
`[N_ar, n_ar] = arma_updateNormalEquationLongAR(x, time, type, magnitude, N_ar, n_ar, patternInTimeSeriesIO, patternInTimeSeriesTC)`

B.3 ARCH(m) Prozesse

Beispielcode

```

% Länge der Zeitreihe festsetzen:
T = 1000;

% Lag-Polynom festsetzen (ARCH(1) Prozess mit  $\alpha_1 = 0.6$ ):
LagPol = [0.6]';

% Die Ordnung des Prozesses  $m = 2$  sei bekannt:
order = 2;

% Konstante  $\zeta$  des Prozesses festlegen:
const = 10.0;

% Geradensteigung festlegen:
slope = 0.3;

% Achsenabschnitt festlegen:
intercept = 10.0;

```

```

% ARCH-Rauschen erzeugen:
u = archNoise( LagPol, const, T );

% Gerade erzeugen:
g = createLine( slope, intercept, T );

% Zeitreihen aus ARCH-Rauschen und Gerade zusammensetzen:
x = u + g;

% Den Parametervektor initialisieren, wobei die Prozessordnung bekannt sei:
theta = zeros( 3 + order, 1 );

% Näherungswerte für die Gerade bestimmen:
theta( 1 : 2 ) = leastSquaresForLine( x );

% Mit den Näherungswerten das Rauschen bestimmen:
u = calcNoise( x, theta );

% Näherungswerte für den ARCH-Prozess bestimmen:
theta( 3 : 3 + order ) = leastSquaresForARCH( u, order );

% Parameter mit dem Newton-Raphson Verfahren schätzen:
% (Der zweite Rückgabeparameter der Funktion gibt die Anzahl der Iterationen an.)
[ thetaNRN, count ] = newtonRaphsonNumeric( x, theta );

```

Liste der Variablen

- **appNoise**
Näherung für das $ARCH(m)$ Rauschen u_t
- **beta**
Näherung für die Parameter β
- **const**
Konstante ζ des $ARCH(m)$ Prozesses
- **count**
Zählvariable, die die Anzahl der benötigten Iterationen des numerischen Optimierungsverfahrens angibt

- **intercept**
Achsenabschnitt der Gerade
- **h**
Varianz h_t des ARCH(m) Rauschens u_t
- **indexLeft**
Vektor, der die Sortierung der Parameter im Simplex-Tableau festhält
- **indexTop**
Vektor, der die Sortierung der Parameter im Simplex-Tableau festhält
- **LagPol**
Kofaktoren $\alpha_1, \alpha_2, \dots, \alpha_m$ des Lag-Polynoms $\alpha(B)$ des ARCH(m) Prozesses
- **M**
Ordnung m des ARCH(m) Prozesses
- **maxTimeUsed**
Länge des Zeitabschnitts, der bei der Berechnung der Näherungswerte für die Parameter des Prozesses mit dem Simplex-Algorithmus verwendet wird
- **slope**
Steigung der Gerade
- **stationary**
Boolesche Variable, die angibt, ob ein ARCH(m) Prozess stationär ist
- **Tableau**
Simplex-Tableau
- **theta**
Parametervektor $\theta' = (\beta', \zeta, \alpha_1, \alpha_2, \dots, \alpha_m)'$
- **thetaARCH**
Näherungswerte $\zeta_0, \alpha_{10}, \alpha_{20}, \dots, \alpha_{m0}$ für die Parameter eines ARCH(m) Prozesses
- **thetaStart**
Näherungswerte $\theta'_0 = (\beta'_0, \zeta_0, \alpha_{10}, \alpha_{20}, \dots, \alpha_{m0})'$ für die Parameter
- **u**
ARCH(m) Rauschen u_t
- **u2**
Elementweise quadriertes ARCH(m) Rauschen u_t^2
- **uTilde**
Hilfsvektor $\tilde{\mathbf{u}}$

- **T**
Länge T der Zeitreihe y_t
- **x**
Zeitreihe y_t
- **xTilde**
Hilfsvektor \tilde{x}

Liste der Funktionen

- **approximateValuesSimplex** berechnet Näherungswerte für die Parameter der Geraden nach der Methode der kleinsten Quadrate und für die Parameter des $ARCH(m)$ Prozesses mit dem Simplex-Algorithmus; außerdem wird eine Modellidentifikation mit Informationskriterien durchgeführt:
[theta] = approximateValuesSimplex(x, maxTimeUsed)
- **archNoise** erzeugt eine Zeitreihe u_t , die einem $ARCH(m)$ Prozess folgt:
[x] = archNoise(LagPol , const, T)
- **calcHelpVectors** berechnet die Hilfsvektoren \tilde{u} und \tilde{x} für ein vorgegebenes θ :
[uTilde, xTilde] = calcHelpVectors(u, u2, theta)
- **calcLogL** berechnet den Wert der Log-Likelihood Funktion $\mathcal{L}(\theta)$ für ein vorgegebenes θ :
[logL] = calcLogL(u2, h, theta)
- **calcNoise** bestimmt das $ARCH(m)$ Rauschen für ein vorgegebenes θ :
[u] = calcNoise(x, theta)
- **calcVariance** berechnet die Varianz h_t für ein vorgegebenes θ :
[h] = calcVariance(u2, theta)
- **checkTheta** prüft, ob die Parameter θ einem stationären Prozess entsprechen:
[stationary] = checkTheta(theta)
- **createLine** erzeugt eine Gerade:
[x] = createLine(slope, intercept, T)
- **leastSquaresForARCH** berechnet Näherungswerte für die Parameter des $ARCH(m)$ Prozesses mit einer Schätzung nach der Methode der kleinsten Quadrate:
[thetaARCH] = leastSquaresForARCH(appNoise, M)

- `leastSquaresForLine` berechnet Näherungswerte für die Parameter der Gerade mit einer Schätzung nach der Methode der kleinsten Quadrate:
`[beta] = leastSquaresForLine(x)`
- `newtonRaphsonAnalytic` berechnet das Maximum $\hat{\theta}$ der Log-Likelihood Funktion $\mathcal{L}(\theta)$ nach dem NEWTON-RAPHSON Verfahren mit analytischer Berechnung des Gradienten $\nabla \mathcal{L}(\theta)$ und analytischer Berechnung der HESSE-Matrix $\mathbf{H}(\mathcal{L}(\theta))$:
`[theta, count] = newtonRaphsonAnalytic(x, thetaStart)`
- `newtonRaphsonNumeric` berechnet das Maximum $\hat{\theta}$ der Log-Likelihood Funktion $\mathcal{L}(\theta)$ nach dem NEWTON-RAPHSON Verfahren mit numerischer Berechnung des Gradienten $\nabla \mathcal{L}(\theta)$ und analytischer Berechnung der HESSE-Matrix $\mathbf{H}(\mathcal{L}(\theta))$:
`[theta, count] = newtonRaphsonNumeric(x, thetaStart)`
- `simplexAlgorithm` führt den Simplex-Algorithmus durch:
`[Tableau,indexTop,indexLeft] = simplexAlgorithm(Tableau,indexTop,indexLeft)`
- `steepestAscendAnalytic` berechnet das Maximum $\hat{\theta}$ der Log-Likelihood Funktion $\mathcal{L}(\theta)$ nach dem Verfahren des steilsten Anstiegs mit analytischer Berechnung des Gradienten $\nabla \mathcal{L}(\theta)$:
`[theta, count] = steepestAscendAnalytic(x, thetaStart)`
- `steepestAscendNumeric` berechnet das Maximum $\hat{\theta}$ der Log-Likelihood Funktion $\mathcal{L}(\theta)$ nach dem Verfahren des steilsten Anstiegs mit numerischer Berechnung des Gradienten $\nabla \mathcal{L}(\theta)$:
`[theta, count] = steepestAscendNumeric(x, thetaStart)`

Literaturverzeichnis

- [1] Chen, C. u. L.-M. Liu (1993): *Joint Estimation of Model Parameters and Outlier Effects in Time Series*. In: Journal of the American Statistical Association, Vol. 88, No. 421, pp. 284-297, Theory and Methods.
- [2] Hamilton, J.D. (1994): *Time Series Analysis*. Princeton University Press, Princeton, New Jersey.
- [3] Kaiser, R. u. A. Maravall (2001): *Seasonal Outliers in Time Series*. Banco de Espana, Servicio de Estudios. Abrufbar unter:
<http://www.bde.es/servicio/software/tramo/slsoutlier.pdf> (Datum: 25. März 2004).
- [4] Kall, P. (2004): *Lineare Optimierung*. Skript zur Vorlesung im Wintersemester 2003/2004 an der Wirtschaftswissenschaftlichen Fakultät der Universität Zürich. Abrufbar unter:
<http://www.unizh.ch/ior/Pages/Deutsch/OR/Skripte/WS2003/LP/LP-ch01-web.pdf>
<http://www.unizh.ch/ior/Pages/Deutsch/OR/Skripte/WS2003/LP/LP-ch02-web.pdf>
(Datum: 24. März 2004).
- [5] Koch, K. R. (1997): *Parameterschätzung und Hypothesentests*. 3. Auflage, Ferd. Dümmers Verlag, Bonn.
- [6] Mautz, R. (2000): *Zur Lösung nichtlinearer Ausgleichungsprobleme bei der Bestimmung von Frequenzen in Zeitreihen*. Dissertation im Fachbereich 9 - Bauingenieurwesen und angewandte Geowissenschaften der Technischen Universität Berlin. Abrufbar unter:
http://edocs.tu-berlin.de/diss/2000/mautz_rainer.pdf (Datum: 25. März 2004).
- [7] Ng, S. u. P. Perron (2003): *A Note on Selection of Time Series Models*. In: Boston College Working Papers in Economics, No. 500, Boston College Department of Economics, Boston. Abrufbar unter:
<http://fmwww.bc.edu/ec-p/wp500.pdf> (Datum: 25. März 2004).
- [8] Ries, J. (2003): *Vergleichende Untersuchungen von Verfahren zur Schätzung von ARMA-Modellen zur Modellierung farbigen Rauschens*. Diplomarbeit am Institut für Theoretische Geodäsie der Rheinischen Friedrich-Wilhelms-Universität Bonn.

- [9] Schlittgen, R. u. B. Streitberg (2001): *Zeitreihenanalyse*. 9. Auflage, R. Oldenburg Verlag, München, Wien.
- [10] Stier, W. (2001): *Methoden der Zeitreihenanalyse*. Springer-Verlag, Berlin, Heidelberg.
- [11] Tsay, R. S. (1986): *Time Series Model Specification in the Presence of Outliers*. Journal of the American Statistical Association, Vol. 81, No. 393, pp. 132-141, Theory and Methods.