

Fakultät Statistik



Echtzeit-Extraktion relevanter Information aus multivariaten Zeitreihen basierend auf robuster Regression

Dissertation

zur Erlangung des akademischen Grades
'Doktor der Naturwissenschaften'
– Dr. rer. nat. –
der Technischen Universität Dortmund

Der Fakultät Statistik
der Technischen Universität Dortmund
vorgelegt von
Matthias Borowski
aus Heiden, Westfalen

Dortmund 2013

PRÜFUNGSKOMMISSION

1. Gutachterin:

Prof. Dr. Ursula Gather

2. Gutachter:

Prof. Dr. Roland Fried

Vorsitz der Prüfungskommission:

PD Dr. Sonja Kuhnt

TAG DER MÜNDLICHEN PRÜFUNG

08. Januar 2013

Kurzzusammenfassung

Diese Arbeit befasst sich mit der Echtzeit-Signalextraktion aus uni- und multivariaten Zeitreihen sowie mit der Echtzeit-Überwachung der Zusammenhänge zwischen den univariaten Komponenten multivariater Zeitreihen. Die in dieser Arbeit untersuchten und entwickelten Methoden eignen sich zur Echtzeit-Anwendung auf hochfrequent gemessene instationäre Zeitreihen, die Ausreißer und Fehler mit wechselnder Variabilität aufweisen.

Ein Verfahren zur Echtzeit-Signalextraktion aus univariaten Zeitreihen wird entwickelt, welches auf der Anpassung robuster Repeated Median-Regressionsgeraden in gleitenden Zeitfenstern gründet, deren Fensterbreite entsprechend der aktuell vorliegenden Datensituation gewählt wird. Eine umfassende Vergleichsstudie zeigt die Überlegenheit der neuen Methode gegenüber einem bereits bestehenden Signalfilter mit adaptiver Fensterbreitenwahl.

Auf Basis des neu entwickelten Signalfilters wird eine Methodik zur Echtzeit-Überwachung der Zusammenhänge zwischen den einzelnen Komponenten einer multivariaten Zeitreihe entwickelt. Dieses Verfahren bewertet zu jedem Zeitpunkt den Zusammenhang zwischen zwei univariaten Zeitreihen anhand der aktuell vorliegenden Trends. Bei diesem Ansatz resultiert ein Zusammenhang aus gleich bzw. ähnlich gerichteten Verläufen.

Das Verfahren zur Überwachung der Zusammenhänge wird mit dem neuen adaptiven Signalfilter kombiniert zu einer multivariaten Prozedur zur umfassenden Extraktion relevanter Information in Echtzeit. Neben der multivariaten Signalextraktion mit adaptiver Fensterbreitenwahl liefert dieses neue Verfahren für jede univariate Zeitreihenkomponente eine Schätzung der Fehlervariabilität, zeigt zu jedem Messzeitpunkt die aktuell bestehenden Zusammenhänge an und erkennt Sprünge und Trendwechsel in Echtzeit.

Inhaltsverzeichnis

1	Einleitung	1
2	Echtzeit-Signalextraktion aus uni- und multivariaten Zeitreihen	4
2.1	Univariate Signalextraktionsverfahren	4
2.1.1	Repeated Median-Regression	7
2.1.2	Der adaptive online Repeated Median	11
2.2	Multivariate Signalextraktionsverfahren	18
2.2.1	Der Trimmed Repeated Median-Least Squares Filter	20
2.2.2	Der adaptive online Trimmed Repeated Median-Least Squares Filter	22
3	Der Slope Comparing Adaptive Repeated Median	26
3.1	Der SCARM-Test	26
3.2	Die Schätzung der Varianz von RM-Steigungsdifferenzen	28
3.3	Der SCARM-Algorithmus	34
3.3.1	Wahl der Inputargumente	35
3.3.2	Recheneffiziente Update-Algorithmen des SCARM	36
3.3.3	Behandlung fehlender Werte	39
3.4	Die Verteilung der SCARM-Teststatistik	41
3.4.1	Normalverteilte Fehlerprozesse	42
3.4.2	Weitere Fehlerverteilungen	46
3.5	Diskussion	48
4	Eine Studie zum Vergleich des aoRM und SCARM	51
4.1	Robustheit der Tests	51
4.1.1	Test Resistances des SCARM-Tests	52
4.1.2	Test Resistances des aoRM-Tests	58
4.1.3	Vergleich der Test Resistances	64
4.2	Erkennung von Signalveränderungen	66
4.2.1	Güte der Tests	66
4.2.2	Zeitverzögerung bei der Erkennung von Signalveränderungen	70
4.3	Effizienz der Signalschätzungen	73
4.4	Anwendung auf echte Daten	78
4.4.1	Zeitreihen aus der intensivmedizinischen Patientenüberwachung	78
4.4.2	Zeitreihen von Aktienpreisen	80
4.5	Zusammenfassung der Vergleichsstudie und Diskussion	82

5	Echtzeit-Überwachung der Zusammenhänge in multivariaten Zeitreihen	84
5.1	Überblick über bestehende Methoden	85
5.2	Das SCARM-basierte Similar Slope Monitoring-Verfahren	86
5.2.1	Die SSM-Statistik zum Vergleich aktueller Trends	86
5.2.2	Schätzung von $\text{Var}[D_t(i, j)]$	88
5.2.3	Bewertung der SSM-Statistik	93
5.2.4	Diskussion	101
6	Der multivariate Slope Comparing Adaptive Repeated Median Filter	105
6.1	Der mSCARM-Algorithmus	105
6.1.1	Die Blockbildung des mSCARM	107
6.1.2	Die TRM-LS-Regression des mSCARM	109
6.1.3	Behandlung fehlender Werte	110
6.2	Anwendung auf echte Daten	111
6.3	Diskussion	115
7	Zusammenfassung	118
	Literaturverzeichnis	121
A	Definitionen	128
B	Simulationen	131
B.1	Varianz des RM-Steigungsschätzers für verschiedene Fehlerverteilungen . .	131
B.1.1	Normalverteilte Fehler	131
B.1.2	Weitere Fehlerverteilungen	131
B.1.3	Autokorrelierte Fehler	133
B.2	Effizienz des Q_δ^{adj} und $Q_\delta^{\text{adj}*}$ für fehlende Werte	133
C	Abbildungen und Tabellen	137

1 Einleitung

In vielen Bereichen werden Prozesse durch fortlaufende Messung einer oder mehrerer Variablen überwacht. So wird beispielsweise im intensivmedizinischen Online-Monitoring der Zustand eines Patienten kontrolliert, indem eine Vielzahl von Vitalparametern wie die Herzfrequenz oder die Sauerstoffsättigung des Blutes fortlaufend sekundlich gemessen wird. Weitere Beispiele sind in hochfrequenten Zeitabständen ausgegebene Aktienpreise oder industrielle Fertigungsprozesse, die durch fortlaufende Messung verschiedener Variablen kontrolliert werden.

Eine hohe Messfrequenz von z.B. einer Beobachtung oder einem Beobachtungsvektor pro Sekunde resultiert häufig in stark verrauschte und mit Ausreißern kontaminierte uni- oder multivariate Zeitreihen, für die gängige Annahmen wie die Stationarität nicht erfüllt sind. Vielmehr weisen die Zeitreihen neben stabilen Phasen auch längere oder wechselnde Trends, plötzliche und zum Teil starke Niveauänderungen sowie wechselnde Variabilitäten auf. Des Weiteren ist die Struktur der Abhängigkeiten zwischen den Variablen häufig sowohl unbekannt als auch zeitveränderlich. Aufgrund dieser Eigenschaften ist es selbst für den geübten Anwender oftmals schwierig die richtigen Schlüsse aus den gemessenen Zeitreihen zu ziehen, insbesondere wenn ausschließlich die aktuell vorliegende Situation zu bewerten ist. Das Ziel dieser Arbeit ist daher die Untersuchung und Entwicklung von Methoden zur Extraktion von relevanten Informationen aus derart komplexen multivariaten Zeitreihen in Echtzeit. In Echtzeit bedeutet in diesem Kontext, dass die extrahierte Information mit jeder neu gemessenen Beobachtung, also zu jedem Messzeitpunkt, aktualisiert wird. In dieser Arbeit werden drei Arten der Echtzeit-Informationsextraktion betrachtet:

1. Die Echtzeit-Extraktion von Signalen aus univariaten Zeitreihen.
2. Die Echtzeit-Extraktion von Signalen aus multivariaten Zeitreihen.
3. Die Echtzeit-Überwachung der Zusammenhänge zwischen den univariaten Komponenten einer multivariaten Zeitreihe.

Alle in dieser Arbeit behandelten Methoden sollen fortlaufend die zum gegenwärtigen Zeitpunkt t vorliegende relevante Information wiedergeben. Daher gründen die Verfahren auf dem Ansatz eines gleitenden Zeitfensters, das die letzten n Beobachtungen umfasst. Aufgrund möglicher Ausreißer in den Daten und der Notwendigkeit der Echtzeit-Fähigkeit kommen nur Verfahren in Frage, die robust gegenüber Ausreißern sind und kurze Rechenzeiten benötigen.

In Kapitel 2 werden bestehende Verfahren zur Signalextraktion aus Zeitreihen vorgestellt. Die Signalextraktion basiert auf der Annahme, dass die Daten aus einem unterliegenden wahren aber unbekanntem Signal bestehen, das überlagert ist von Rauschen und Ausreißern. Eine Extraktion des Signals bedeutet daher eine Glättung der Zeitreihe. Eine geeignete Vorgehensweise zur Echtzeit-Signalextraktion bietet die robuste Regression in gleitenden Zeitfenstern. In Vergleichsstudien hat sich die *Repeated Median*- (RM) Regression gegenüber anderen robusten Regressionsmethoden durchgesetzt. Daher existiert eine Reihe von uni- und multivariaten Verfahren zur Signalextraktion, die auf dem RM gründen. Da für die fensterbasierte Signalextraktion ein Bias-Varianz-Dilemma bzgl. der Wahl der Fensterbreite besteht, bietet die adaptive Fensterbreitenwahl einen viel versprechenden Ansatz zur Verbesserung der Signalextraktion. Diesen Ansatz verfolgt der univariate *adaptive online Repeated Median* (aoRM). Im Gegensatz zum aoRM bezieht der multivariate *Trimmed Repeated Median-Least Squares*- (TRM-LS) Filter die Abhängigkeiten zwischen den einzelnen Komponenten einer multivariaten Zeitreihe mit ein. Der *adaptive online TRM-LS*-Filter (aoTRM-LS) kombiniert den adaptiven aoRM mit dem multivariaten TRM-LS zu einem multivariaten Filter mit adaptiver Fensterbreitenwahl. Zwar vereint der aoTRM-LS in der Theorie die Vorteile des aoRM und TRM-LS, doch zeigen Anwendungen, dass die Fensterbreitenadaptation nicht befriedigend funktioniert. Einen Ansatz zur Verbesserung bietet die Wahl eines Testverfahrens, das im Vergleich zum Test des aoRM eine höhere Güte aufweist. In Kapitel 3 wird daher eine Alternative zum aoRM entwickelt. Ein weiterer Ansatz ist eine zeitabhängige Einteilung der K Komponenten in Blöcke abhängiger Variablen, so dass die Signalschätzung blockweise erfolgt. Dies motiviert die Entwicklung eines Verfahrens zur Echtzeit-Überwachung der Zusammenhänge zwischen den Komponenten einer multivariaten Zeitreihe in Kapitel 5.

Bei dem in Kapitel 3 entwickelten *Slope Comparing Adaptive Repeated Median* (SCARM) geschieht die Fensterbreitenadaptation mit Hilfe eines Tests, der auf der Differenz von RM-Steigungsschätzungen basiert, die in separaten aufeinander folgenden Zeitfenstern ermittelt werden. Die Schätzung der zur Standardisierung benötigten Varianz der RM-Steigungsdifferenzen ist von großer Bedeutung für die Leistung der Fensterbreitenadaptation. Daher wird eine Schätzmethodik entwickelt, die unempfindlich gegenüber starken und wiederholten Signalveränderungen ist und somit eine hohe Güte des Tests gewährleistet.

Der SCARM und der aoRM werden in einer umfangreichen Studie in Kapitel 4 gegenübergestellt. Als Vergleichskriterien werden die Robustheit der Tests gegenüber Ausreißern, die Güte der Tests, die Zeitverzögerung bzgl. der Erkennung von Signalveränderungen und

die Effizienz der Signalschätzungen betrachtet. Der Vergleich zeigt die Überlegenheit des SCARM, insbesondere hinsichtlich der Güte zur Erkennung von Signalveränderungen bei niedrigem Signifikanzniveau. Anwendungen der beiden Filter auf Zeitreihen aus der Intensivmedizin und auf Zeitreihen von Aktienpreisen bestätigen die Resultate.

In Kapitel 5 wird zunächst ein Überblick über bestehende Verfahren zur Überwachung der Zusammenhänge zwischen den Komponenten einer multivariaten Zeitreihe gegeben. Aufgrund der beschriebenen schwierigen Datenlage müssen die Methoden eine Reihe von Anforderungen erfüllen. Da in der Literatur meines Wissens kein Verfahren besteht, das allen gestellten Anforderungen genügt, wird das robuste *Similar Slope Monitoring*- (SSM) Verfahren entwickelt. Das SSM-Verfahren entdeckt einen Zusammenhang zwischen zwei univariaten Zeitreihen zu einem Zeitpunkt t , wenn die aktuell vorliegenden Trends ähnlich sind. Die komponentenweise Anwendung des SCARM liefert dabei die Schätzungen der aktuellen Trends, deren standardisierte Differenz die Vergleichsstatistik des SSM-Verfahrens bildet. Zur Schätzung der Varianz der Differenz wird eine Methodik entwickelt, die auf den vom SCARM gelieferten Schätzungen der Fehlervarianzen beruht. Die Anwendung der SSM-Methode in verschiedenen Szenarien ähnlicher und unähnlicher bivariater Zeitreihen liefert Erkenntnisse bzgl. der Bewertung der Zusammenhänge.

Auf Basis der neuen Verfahren SCARM und SSM wird in Kapitel 6 ein adaptiver Filter für multivariate Zeitreihen entwickelt, der *multivariate SCARM* (mSCARM). Dieser verwendet die Fensterbreitenadaptation des SCARM, bildet zu jedem Zeitpunkt mit Hilfe des SSM-Verfahrens Blöcke aus aktuell ähnlich verlaufenden Komponenten und schätzt das Signal innerhalb der Blöcke mit dem robusten multivariaten TRM-LS. Der mSCARM liefert neben der Signalextraktion zusätzlich Schätzungen der Fehlervarianz, zeigt Signalveränderungen an und gibt Aufschluss über aktuell bestehende Zusammenhänge zwischen den univariaten Komponenten. Der mSCARM stellt daher ein robustes RM-basiertes Verfahren zur umfassenden Echtzeit-Extraktion relevanter Information aus multivariaten Zeitreihen dar.

Eine Zusammenfassung der Ergebnisse sowie ein Überblick über mögliche Anwendungsgebiete der entwickelten Methoden liefert Kapitel 7.

2 Echtzeit-Signalextraktion aus uni- und multivariaten Zeitreihen

Durch die fortlaufende Messung einer oder mehrerer Variablen in kurzen Zeitabständen – z.B. zur Überwachung von industriellen Fertigungsprozessen oder zur Überwachung des Blutdrucks und der Herzfrequenz eines Intensivpatienten – ergeben sich aufgrund der hohen Messfrequenz von z.B. einer Beobachtung pro Sekunde oftmals stark verrauschte und mit Ausreißern kontaminierte uni- bzw. multivariate Zeitreihen, die auch als Datenströme bezeichnet werden. Dieses Kapitel befasst sich mit bestehenden Verfahren zur Echtzeit-Signalextraktion aus uni- und multivariaten Datenströmen. Dies meint die Schätzung des angenommenen wahren Signals, das den Daten unterliegt und die relevante Information trägt, jedoch unbekannt und von Rauschen und Ausreißern überlagert ist. Bei der Signalextraktion in Echtzeit erfolgt die Signalschätzung zum Zeitpunkt t , sobald die Beobachtung y_t gegeben ist und bevor der nächste Messwert y_{t+1} vorliegt. Da die Signale aus hochfrequent, z.B. sekundlich, gemessenen Zeitreihen extrahiert werden sollen, kommen nur Verfahren in Frage, die wenig Rechenzeit benötigen. Darüber hinaus soll das wahre unterliegende Signal möglichst exakt approximiert werden, ohne dass relevante Niveau- oder Trendänderungen in den Datenströmen übergangen werden. Zunächst werden Verfahren für univariate Zeitreihen behandelt.

2.1 Univariate Signalextraktionsverfahren

Eine zentrale Annahme in dieser Arbeit ist, dass sich die Beobachtungen einer univariaten Zeitreihe $(y_t)_{t \in \mathbb{Z}}$ als Realisationen von

$$Y_t = \mu_t + \varepsilon_t + \eta_t \quad (2.1)$$

ergeben. Dabei bezeichnet μ_t das unterliegende wahre, aber unbekanntes Signal zum Zeitpunkt t . Es wird angenommen, dass das Signal zumeist „glatt“ ist, d.h. einen stabilen Verlauf zeigt. Es kann jedoch auch längere Trends und plötzliche Niveau- und Trendwechsel aufweisen. Der Fehlerprozess $(\varepsilon_t)_{t \in \mathbb{Z}}$ stellt einen Prozess von Zufallsvariablen mit Erwartungswert $E(\varepsilon_t) = 0$ und zeitabhängiger Varianz $\text{Var}(\varepsilon_t) = \sigma_t^2$ dar. Der Ausreißer generierende Prozess $(\eta_t)_{t \in \mathbb{Z}}$ erzeugt zufällige stark abweichende Werte, ist aber für den Großteil der Zeitpunkte gleich null.

Einen intuitiven Ansatz zur Echtzeit-Signalextraktion bieten gleitende Zeitfenster. Der *gleitende Mittelwert* liefert für normalverteilte Daten sehr effiziente Signalschätzungen und gibt Trends gut wieder. Abrupte Niveauänderungen, so genannte *Level Shifts* (auch *Sprünge*, *Jumps* oder *Edges*), werden jedoch nicht exakt, sondern „verschmiert“ wiedergegeben. Darüber hinaus ist er nicht robust gegenüber Ausreißern. Im Gegensatz dazu ist der *gleitende Median* (Tukey, 1977) hoch-robust gegenüber Ausreißern und gibt Level Shifts gut wieder, approximiert jedoch das Signal in Trendphasen durch eine Treppenfunktion. Einen Kompromiss zwischen dem gleitenden Mittel und dem gleitenden Median bietet das *Modified-Trimmed-Mean* (MTM, Lee und Kassam, 1985). Dabei werden diejenigen Beobachtungen aus der Stichprobe entfernt, die weiter vom Median der Stichprobe entfernt sind als ein zuvor festgelegtes Vielfaches der Streuungsschätzung, die z.B. mit dem *Median der absoluten Abweichungen vom Median* (*Median Absolute Deviation*, MAD) erfolgen kann. Anschließend wird der Mittelwert der getrimmten Stichprobe berechnet. Eine Modifikation des MTM ist das *Double Window MTM* (DWMTM, Bernholt et al., 2006), bei dem der Median und der MAD in einem kürzeren Zeitfenster bestimmt werden, weshalb das DWMTM kürzere Rechenzeiten als der MTM benötigt.

Gleitende Mittelwerte und Mediane basieren auf Lokationsschätzern und implizieren daher ein lokal konstantes Niveau. Hinsichtlich des Prinzips der lokalen Anpassung von Polynomen an die Daten würden lokationsbasierte Schätzer demnach der Anpassung eines Polynoms vom Grad null entsprechen. Die Glättung von Daten durch lokale polynomiale Anpassung wurde bereits Ende des 19. bis Anfang des 20. Jahrhunderts praktiziert, vgl. z.B. Woolhouse (1870) und Spencer (1904). Einen robusten Ansatz der lokalen polynomialen Anpassung liefert Cleveland (1979), der in einer iterativen Prozedur großen Residuen kleine Gewichte zuordnet. Dabei empfiehlt er Polynome vom Grad eins, was der lokalen Anpassung einer Geraden entspricht. Weiterhin weisen McDonald und Owen (1986) darauf hin, dass Polynome der Ordnung ≥ 2 zwar die Nachzeichnung lokaler Extrema ermöglichen und somit Signalschätzungen mit geringen Verzerrungen liefern. Diese weisen jedoch hohe Varianzen auf.

Auch Davies et al. (2004) und Gather et al. (2006) empfehlen für die Echtzeit-Signalextraktion die lokale Anpassung von Regressionsgeraden. Davies et al. (2004) verwenden robuste Regressionsverfahren in gleitenden Zeitfenstern der Form $\{t - w, \dots, t, \dots, t + w\}$ mit ungerader Länge $n = 2w + 1$, um das Signal zum mittleren Fensterzeitpunkt t anhand des Niveaus der Regressionsgerade zu diesem Zeitpunkt t zu approximieren. Da bei einer Anwendung in Echtzeit der Zeitpunkt $t + w$ dem aktuellen Zeitpunkt entspräche, würde

dieser Ansatz zu Signalschätzungen mit einer Zeitverzögerung von w Zeiteinheiten führen. Daher lässt sich dieser Ansatz als *zeitverzögerte* Signalschätzung und das entsprechende Filterverfahren als *zeitverzögerter Filter* bezeichnen. Da bei der Echtzeit-Anwendung eine Zeitverzögerung in der Regel unerwünscht ist, schätzen Gather et al. (2006) das Signal zum rechten bzw. letzten Zeitpunkt $t + w$ des Zeitfensters. Dieser entspräche bei einer Anwendung in Echtzeit dem aktuellen Zeitpunkt, das Signal würde also ohne Zeitverzögerung extrahiert. Dieser Ansatz lässt sich daher als *Online-Signalschätzung* bezeichnen, die Filterverfahren entsprechend als *Online-Filter*. Um bei Online-Filtern auch die Wahl gerader Fensterbreiten zu ermöglichen, verwenden Schettlinger et al. (2010a) gleitende Zeitfenster $\{t - n + 1, \dots, t\}$ beliebiger Länge n . Hierbei wird das Signal zum aktuellen bzw. letzten Zeitpunkt t geschätzt.

Sowohl zeitverzögerte als auch Online-Regressionsfilter basieren auf der Annahme, dass sich das Signal „gut“ durch lokale Anpassung von Regressionsgeraden approximieren lässt. Beim zeitverzögerten Ansatz lässt sich diese Annahme wie folgt schreiben:

$$\mu_{t+i} \approx \mu_t + \beta_t \cdot i, \quad i = -w, \dots, w, \quad (2.2)$$

wobei μ_{t+i} das Signal zum Zeitpunkt $t + i$ und β_t die Steigung im Zeitfenster $\{t - w, \dots, t, \dots, t + w\}$ bezeichnet. Bei der Online-Signalschätzung lässt sich die Annahme der lokalen Linearität formulieren durch

$$\mu_{t-n+i} \approx \mu_t + \beta_t \cdot (i - n), \quad i = 1, \dots, n. \quad (2.3)$$

Dabei ist μ_{t-n+i} das Signal zum Zeitpunkt $t - n + i$ und β_t die Steigung im Zeitfenster $\{t - n + 1, \dots, t\}$. Beide Ansätze (2.2) und (2.3) zielen darauf ab, die Steigung β_t und das Niveau μ_t zu jedem Zeitpunkt t mit Hilfe von Regressionsverfahren zu schätzen. Dabei dient das Niveau $\hat{\mu}_t$ der Regressionsgeraden als Signalschätzung zum Designpunkt t .

Für den zeitverzögerten Ansatz (2.2) vergleichen Davies et al. (2004) die L_1 -, die *Repeated Median*- (RM, Siegel, 1982) und die *Least Median of Squares*- (LMS, Hampel, 1975; Rousseeuw, 1984) Regression hinsichtlich der Kriterien Robustheit, Rechenzeit, Effizienz bei normalverteiltem Rauschen und Verhalten bei Level Shifts und Trendwechseln. Dabei erweist sich die RM-Regression insgesamt als geeignetste Methode. Zum einen ist sie für äquidistante Designs robuster als die L_1 -Methode, zum anderen liefert sie glattere Signalschätzungen und benötigt weniger Rechenzeit als die LMS-Regression. Letztere hat jedoch

Vorteile bei der Wiedergabe von Level Shifts und Trendwechseln.

Gather et al. (2006) vergleichen die Methoden RM, LMS, *Least Trimmed Squares* (LTS, Rousseeuw, 1985) und *Deepest Regression* (DR, Rousseeuw und Hubert, 1999) zur Online-Signalextraktion. Dabei weisen der RM und die DR ebenso wie die LMS- und LTS-Regression ähnliche Eigenschaften auf. Die LMS- und LTS-Regression sind robuster gegenüber einer großen Anzahl an Ausreißern, was jedoch gleichzeitig eine trägere Reaktion auf Level Shifts bedeutet. Weiterhin liefern der RM und die DR stabilere Signalschätzungen bei kleinen und moderaten Ausreißern. Insgesamt zeigt der RM auch in dieser Vergleichsstudie die besten Eigenschaften, auch weil er von allen Schätzern die kürzeste Rechenzeit benötigt.

Einen Kompromiss zwischen der Glattheit des RM und der guten Wiedergabe von Level Shifts und Trendwechseln der LMS- bzw. LTS-Regression bieten die *Least Quartile Difference*- (LQD, Croux et al., 1994) und die *Least Trimmed Differences*- (LTD, Stromberg et al., 2000) Regression. Diese Regressionsschätzer sind jedoch instabil bei moderaten Variabilitäten in den Daten und benötigen lange Rechenzeiten, weshalb sie für die Echtzeit-Anwendung weniger geeignet sind als der RM.

Die hier wiedergegebenen Untersuchungsergebnisse charakterisieren den RM als geeignets-ten Kandidaten für die Echtzeit-Signalextraktion aus nichtstationären, stark verrauschten und mit Ausreißern kontaminierten Zeitreihen. Die im Rahmen dieser Arbeit entwickelten Verfahren basieren daher auf dem RM.

2.1.1 Repeated Median-Regression

Da bei der Signalextraktion in Echtzeit in der Regel keine Zeitverzögerung erwünscht ist, wird im Folgenden der Online-Ansatz (2.3) betrachtet. Sei $\mathbf{Y}_{t,n} := (Y_{t-n+1}, \dots, Y_t)$ ein Vektor von Zufallsvariablen in einem Zeitfenster der Länge n . Der RM-Schätzer der Steigung β_t (im Zeitfenster) und des Niveaus μ_t (zum Designpunkt t) ist gegeben durch $T_{\text{RM}} = (\hat{\beta}_t, \hat{\mu}_t)$ mit

$$\hat{\beta}_t := \hat{\beta}(\mathbf{Y}_{t,n}) = \operatorname{med}_{i=1,\dots,n} \left\{ \operatorname{med}_{j \neq i} \frac{Y_{t-n+i} - Y_{t-n+j}}{i - j} \right\}, \quad (2.4)$$

$$\hat{\mu}_t := \hat{\mu}(\mathbf{Y}_{t,n}) = \operatorname{med}_{i=1,\dots,n} \left\{ Y_{t-n+i} - \hat{\beta}_t \cdot (i - n) \right\}. \quad (2.5)$$

Für einen Vektor von Realisationen $\mathbf{y}_{t,n} := (y_{t-n+1}, \dots, y_t)$ wird die Notation $\hat{\beta}(\mathbf{y}_{t,n})$ und $\hat{\mu}(\mathbf{y}_{t,n})$ verwendet. In dieser Arbeit ist der Median $\operatorname{med}\{\cdot\}$ für gerade n definiert als das arithmetische Mittel der $(n/2)$ - und $(n/2 + 1)$ -ten Ordnungsstatistik.

Zur Schätzung des Parameters μ_t in (2.5) muss zunächst die Schätzung der Steigung β_t erfolgen. Alternativ kann μ_t auch direkt durch

$$\hat{\mu}(\mathbf{Y}_{t,n}) = \operatorname{med}_{i=1,\dots,n} \left\{ \operatorname{med}_{j \neq i} \frac{j Y_{t-n+i} - i Y_{t-n+j}}{i - j} \right\} \quad (2.6)$$

geschätzt werden. Die Schätzer (2.5) und (2.6) besitzen dieselbe beschränkte Influenzfunktion und asymptotische Effizienz für standardnormalverteilte Daten (Hössjer et al., 1995). Da der hierarchische Schätzer (2.5) jedoch weniger Rechenzeit benötigt und intuitiver ist, wird im Folgenden nur dieser verwendet.

Der RM-Steigungsschätzer ist im einfachen linearen Modell mit festen Kovariablen und symmetrischer Verteilung des Fehlers unverzerrt und Fisher-konsistent (Siegel, 1982). Zudem ist der RM regressions- und skalenäquivariant und äquivariant bzgl. affiner Transformationen der Zielvariablen, jedoch nicht bzgl. affiner Transformationen der Kovariablen (Schettlinger, 2004). Letzteres stellt jedoch im Zeitreihenkontext aufgrund des festen Designs keinen Nachteil dar. Die Definitionen der Äquivarianzeigenschaften von Schätzern sind in Anhang A gegeben.

Der RM besitzt einen Bruchpunkt¹ von $\lfloor n/2 \rfloor / n$ für Daten in allgemeiner Lage (Davies et al., 2004). Dies ist der größtmögliche Bruchpunkt für regressionsäquivalente Schätzer (Davies und Gather, 2005). Dabei weist der RM-Schätzer des Niveaus (2.5) gegenüber der Kleinsten Quadrate-Regression für endliche Stichproben aus einer Standardnormalverteilung eine recht hohe Effizienz von fast 70% auf (Gather et al., 2006). Die asymptotische Effizienz für normalverteilte Daten beträgt 63.7% (Hössjer et al., 1995). Der RM hat im Vergleich zu vielen anderen robusten Regressionsschätzern eine geringere Varianz, weshalb er für Zeitreihen eine glattere Signalextraktion liefert als andere Verfahren. Die asymptotische Effizienz des RM-Steigungsschätzers (2.4) für normalverteilte Daten beträgt lediglich 40.5% (Hössjer et al., 1994). Simulationen zeigen jedoch eine Effizienz von bis zu 61% für endliche normalverteilte Stichproben.

Weiterhin besitzt der RM die *exact-fit*-Eigenschaft (Rousseeuw und Leroy, 1987, S. 60f): Liegen mindestens $\lfloor n/2 \rfloor + 1$ der Beobachtungen einer Stichprobe vom Umfang n kollinear, so verläuft die RM Regressionsgerade exakt durch diese Datenpunkte. *Exact-fit*-Situationen sind jedoch auszuschließen, falls den Daten eine stetige Verteilung zugrunde liegt, da sich dann die Beobachtungen mit Wahrscheinlichkeit eins in allgemeiner Lage befinden.

Eine direkte Implementierung des RM benötigt eine Rechenzeit von $\mathcal{O}(n^2)$. Es bestehen

¹In dieser Arbeit ist mit Bruchpunkt der *finite sample replacement breakdown point* (fsbp) nach Donoho und Huber (1983) bezeichnet, vgl. Anhang A.

jedoch schnellere Algorithmen, die lediglich eine erwartete Rechenzeit von $\mathcal{O}(n \log^2 n)$ (Rousseeuw et al., 1993) bzw. $\mathcal{O}(n \log n)$ (Matoušek et al., 1998) aufweisen. Eine bzgl. der Echtzeit-Anwendung entscheidende Verbesserung stellt der von Bernholt und Fried (2003) entwickelte Update-Algorithmus für den RM im gleitenden Zeitfenster dar. Dieser benötigt eine Rechenzeit von nur $\mathcal{O}(n)$.

Für den RM besteht eine Reihe von Modifikationen und Erweiterungen, welche die Effizienz der Signalschätzung und/oder die Wiedergabe von Level Shifts verbessern. Bernholt et al. (2006) übertragen die Idee des lokationsbasierten MTM-Filters von S. 5 auf den Regressionsfall. Sie führen in einem Zeitfenster zunächst eine RM-Regression durch, woraufhin der MAD der RM-Residuen berechnet wird. Dieser dient dann als Schätzung für die Variabilität des Rauschens im Zeitfenster (Gather und Fried, 2003). Daraufhin werden diejenigen Beobachtungen aus der Fensterstichprobe entfernt, die weiter von der RM-Regressionsgeraden entfernt sind als ein Vielfaches der MAD-Skalenschätzung. Beim RM^2 -Filter wird dann auf Basis der getrimmten Stichprobe eine erneute RM-Regression durchgeführt, wohingegen der *Trimmed Repeated Median*- (TRM) Filter die Kleinste Quadrate-Regression verwendet. Der TRM verzeichnet gegenüber dem RM^2 nur sehr geringe Einbußen hinsichtlich der Robustheit, benötigt dafür jedoch weniger Rechenzeit und ist effizienter bei normalverteiltem Rauschen. Der RM^2 und TRM sind zeitverzögerte Filter, die sich jedoch problemlos für den Online-Ansatz modifizieren lassen. Dies ist bereits für die multivariate Variante des TRM (vgl. Kapitel 2.2) geschehen.

Die *Double Window* RM^2 - und *Double Window* TRM-Regression sind Varianten der RM^2 - und TRM-Regression, bei welchen die initiale RM-Regression in einem kürzeren Zeitfenster geschieht. Diese Varianten benötigen kürzere Rechenzeiten und geben Level Shifts genauer wieder. Sie wurden jedoch bisher lediglich für den zeitverzögerten Ansatz untersucht. Für eine Modifikation dieser Filter zur Online-Signalextraktion müsste die Wahl der unterschiedlichen Fensterbreiten bzw. deren Einfluss auf die Signalschätzung erforscht werden. Fried et al. (2007) entwickeln das Prinzip des *Weighted* RM (WRM) unter der Annahme, dass für nahe beieinander liegende Zeitpunkte die Steigung des Signals ähnlicher ist als für weiter entfernte. Für den zeitverzögerten Ansatz werden symmetrische und für die Online-Signalextraktion monotone Gewichtsprinzipien empfohlen. Der WRM ist effizienter als der RM und gibt nichtlineare Trends besser wieder. Für den Fall kurz aufeinander folgender Ausreißergruppen stellt der RM jedoch die bessere Wahl dar (Fried et al., 2007). Außerdem ist die optimale Wahl der Gewichte bzgl. eines geeigneten Kriteriums schwer zu ermitteln.

Der RM zeigt gute Robustheitseigenschaften hinsichtlich kleiner Ausreißer, weist jedoch Schwächen bei großen Ausreißern auf (Gather und Fried, 2003, 2004). Da große Ausreißer für gewöhnlich leichter zu erkennen sind als kleine, fügt Fried (2004) dem RM eine Regel zur Ausreißerererkennung und -ersetzung bei. Wie beim RM^2 und TRM gelten dabei Beobachtungen als Ausreißer, falls sie um mehr als ein Vielfaches der robust geschätzten Fehlervariabilität von der RM-Regressionsgeraden abweichen. Zusätzlich zu der Ausreißerererkennungsregel schlägt Fried (2004) eine Methode zur Erkennung von Level Shifts vor, die auf einer einfachen Mehrheitsregel gründet. Dadurch können Level Shifts besser und schneller nachgezeichnet werden, was den größten Nachteil des RM aufhebt.

Wie für alle lokal arbeitenden Signalextraktionsverfahren besteht auch für RM-basierte Filter ein Dilemma bzgl. der Wahl der Fensterbreite n . Große Fensterbreiten führen zu glatten Signalschätzungen mit geringer Variabilität, wohingegen kleine Fensterbreiten zu exakten Signalschätzungen führen, die näher an den Daten liegen. Solange die Zeitreihe einen stabilen Trend aufweist, der auch Null sein kann, ist n groß zu wählen um eine glatte Signalextraktion zu erhalten. Plötzliche Veränderungen des unterliegenden Signals sollen jedoch möglichst genau nachgezeichnet werden, da diese in der Regel eine wichtige Information darstellen. In derartigen Situationen wird also eine kleine Fensterbreite benötigt. Abbildung 2.1 verdeutlicht das beschriebene Dilemma. Hier führt die größere Fensterbreite $n = 50$ (blaue Signalextraktion) zwar zu guten Ergebnissen, wenn keine Veränderungen in den Daten auftreten; der vorliegende Level Shift wird jedoch stark verzögert und „verschmiert“ nachgezeichnet. Demgegenüber wird der Level Shift bei der kleineren Fensterbreite $n = 10$ (rote Signalextraktion) wesentlich exakter und mit geringerer Zeitverzögerung nachgezeichnet. Aufgrund des beschriebenen Dilemmas kann mit einer festen Fensterbreite nicht zu jedem Zeitpunkt eine (bzgl. eines geeigneten Kriteriums) optimale Signalextraktion erreicht werden. Vielmehr ist je nach aktueller Datenlage eine kleine oder große Fensterbreite von Vorteil, was eine automatische zeit- bzw. datenadaptive Fensterbreitenwahl motiviert. Für die retrospektive Signalextraktion schlagen Gather und Fried (2004) eine Methodik zur Fensterbreitenadaption vor, welche von Schettlinger et al. (2010a) zur Entwicklung eines adaptiven RM-basierten Online-Filters weiterentwickelt wird, dem *adaptiven online RM* (aoRM).

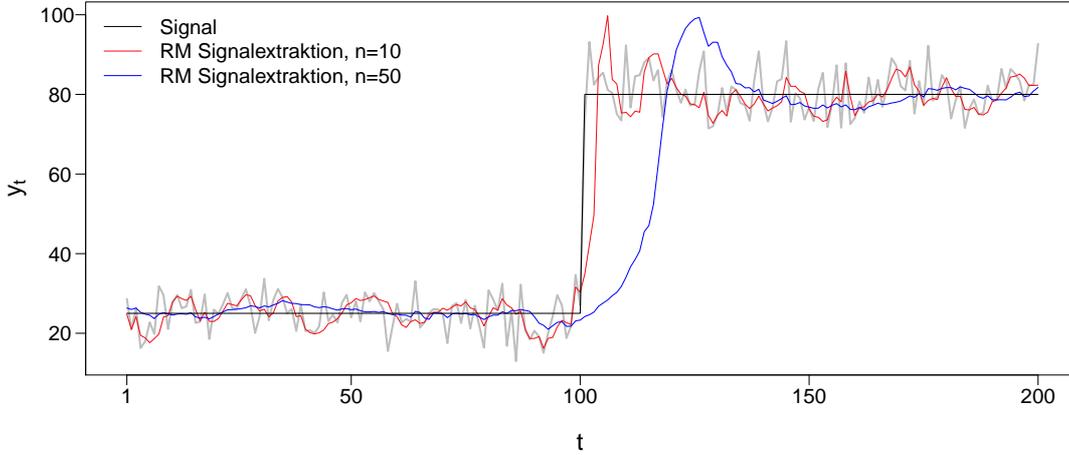


Abbildung 2.1: Vergleich der RM-Signalextraktion mit unterschiedlichen Fensterbreiten.

2.1.2 Der adaptive online Repeated Median

Um zu jedem Zeitpunkt eine passende Fensterbreite zu finden, verwendet der aoRM einen Test, der die Güte der Anpassung der RM-Regressionsgerade anhand der Vorzeichen der Residuen bewertet. Der Test des aoRM wird im Folgenden vorgestellt. Anschließend wird der Algorithmus des aoRM inklusive einiger Modifikationen beschrieben. Eine detaillierte Erklärung des aoRM findet sich in Schettlinger et al. (2010a) und in der Dissertation von Schettlinger (2009, Kap. 3.2). Die Dissertation liefert dabei umfangreiche Untersuchungen und Simulationen, u.a. hinsichtlich der Verteilung der vom aoRM verwendeten Teststatistik.

Der Test des aoRM

Da der aoRM zu jedem Zeitpunkt t die Fensterbreite an die aktuelle Datenlage adaptiert, sei die Fensterbreite im Folgenden mit n_t bezeichnet. Der Test des aoRM gründet auf der Tatsache, dass der Median der Residuen einer RM-Regression unter gewissen Bedingungen (s.u.) gleich Null ist. Daher gilt

$$\sum_{i=1}^{n_t} \text{sign}(\mathfrak{R}_{t,i}) = 0. \quad (2.7)$$

Dabei sind die Residuen $\mathfrak{R}_{t,i}$ im Zeitfenster $\{t - n_t + 1, \dots, t\}$ gegeben durch

$$\mathfrak{R}_{t,i} = Y_{t-n_t+i} - \hat{\mu}_{t-n_t+i}, \quad i = 1, \dots, n_t, \quad (2.8)$$

und $\hat{\mu}_{t-n_t+i} = \hat{\mu}_t + \hat{\beta}_t \cdot (i - n_t)$ bezeichnet die RM-Regressionsgerade mit $\hat{\mu}_t$ und $\hat{\beta}_t$ gemäß (2.4) und (2.5). Weiterhin ist

$$\text{sign}(x) = \begin{cases} 1, & \text{falls } x > 0 \\ 0, & \text{falls } x = 0 \\ -1, & \text{falls } x < 0 \end{cases}$$

die Signumfunktion. Für die Balance der Residuenvorzeichen wie in (2.7) müssen sowohl der Regressor als auch der Regressand einer stetigen Verteilung folgen (Hössjer et al., 1995). Aufgrund des festen Designs ist letztere Bedingung jedoch nicht erfüllt. Für stetig verteilte Zielvariablen gilt zwar die allgemeine Lage der Beobachtungen mit Wahrscheinlichkeit eins, doch ist diese Bedingung nicht hinreichend für (2.7). Nach Schettlinger (2009, Kap. 2.2) ist jedoch die Wahrscheinlichkeit, dass (2.7) bei stetig verteilten Daten mit festen Designpunkten nicht erfüllt ist, bereits bei kleinen Stichprobenumfängen vernachlässigbar gering.

Die Anpassung der RM-Geraden an die Daten wird als gut bewertet, wenn die Balance der RM-Residuen in einem bestimmten Teilfenster von $\{t - n_t + 1, \dots, t\}$ gegeben ist. Der aoRM testet daher anhand der Residuen die Lage der Verteilung der Fehler im Teilfenster $\{t - n_t + i\}$ mit $i \in I_t \subset \{1, \dots, n_t\}$ und Fensterbreite $|I_t| = n_{I_t}$. Als Lokationsmaß dient $\tilde{\mu}_\varepsilon^{I_t}$, der Median der Fehlerverteilung an den durch I_t spezifizierten Designpunkten. Das Testproblem lautet

$$H_0 : \tilde{\mu}_\varepsilon^{I_t} = 0 \quad \text{vs.} \quad H_1 : \tilde{\mu}_\varepsilon^{I_t} \neq 0, \quad (2.9)$$

und die Teststatistik ist gegeben durch

$$T_{\text{aoRM}}(\mathbf{Y}_{t,n_t}) = \sum_{i \in I_t} \text{sign}(\mathfrak{R}_{t,i}) \quad (2.10)$$

mit $\mathfrak{R}_{t,i}$ aus (2.8). Falls entweder die negativen oder positiven Residuen in der Teilstichprobe überwiegen, ist $|T_{\text{aoRM}}(\mathbf{Y}_{t,n_t})| \in \{0, 1, \dots, n_{I_t}\}$ entsprechend groß. In diesem Fall kann die Anpassung der Regressionsgeraden im Zeitfenster $\{t - n_t + 1, \dots, t\}$ nicht als adäquat bewertet werden. Die Nullhypothese wird daher verworfen, falls

$$|T_{\text{aoRM}}(\mathbf{Y}_{t,n_t})| > c_\alpha(n_t, n_{I_t}),$$

wobei der kritische Wert $c_\alpha(n_t, n_{I_t})$ vom Umfang n_t der gesamten Stichprobe und vom Umfang n_{I_t} der Teilstichprobe abhängt; α bezeichnet das gewählte Signifikanzniveau.

In einer umfangreichen Simulationsstudie vergleichen Schettlinger et al. (2010a) verschiedene Designpunktfolgen I_t , die aus den zentral gelegenen, aus den ersten und letzten und aus den letzten n_{I_t} Designpunkten im Zeitfenster $\{t - n_t + 1, \dots, t\}$ bestehen. Basierend auf den Ergebnissen der Studie empfehlen sie, die letzten n_{I_t} Designpunkte im Zeitfenster zu betrachten, d.h. $I_t = \{n_t - n_{I_t} + 1, \dots, n_t\}$ mit

$$n_{I_t} = \min\{n_I, \lfloor n_t/2 \rfloor\}, \quad n_I \in \mathbb{N}. \quad (2.11)$$

Für diese Wahl von I_t weisen die Signalschätzungen des aoRM eine geringe Verzerrung nach plötzlichen Niveau- und Trendveränderungen auf. Als kritischer Wert $c_\alpha(n_t, n_{I_t})$ dient das $(1 - \alpha/2)$ -Quantil der Verteilung von T_{aoRM} unter der Nullhypothese. Die wahre Verteilung von T_{aoRM} ist unbekannt, kann jedoch für große Stichprobenumfänge durch eine hypergeometrische Verteilung approximiert werden (Schettlinger, 2009, Kap. 3.2.3). Für kleine Stichprobenumfänge $n_t \in \{11, \dots, 121\}$ wird die Verteilung der Teststatistik mit Hilfe von Simulationen approximiert. Dabei werden die Beobachtungen aus einer Standardnormalverteilung generiert. Für schiefe Verteilungen oder für Verteilungen mit schweren Rändern ergibt sich kein wesentlicher Unterschied in der Verteilung der Teststatistik, da diese lediglich die Vorzeichen der RM-Residuen betrachtet.

Für den Anpassungstest empfehlen Schettlinger et al. (2010a) ein Signifikanzniveau von $\alpha = 0.1$, da der aoRM-Filter damit gute Ergebnisse in Simulationen erzielt. Abbildung C.1 in Anhang C zeigt die anhand der Simulationen gewonnenen kritischen Werte $c_\alpha(n_t, n_{I_t})$ für $n_{I_t} \in [10, 60]$, $n_t \in [2n_{I_t}, 120]$ und $\alpha = 0.1$. Hier ist $c_\alpha(n_t, n_{I_t})$ für jede Kombination von n_t und n_{I_t} nicht größer als sechs.

Der aoRM-Algorithmus

Der aoRM verwendet den beschriebenen Test zur Fensterbreitenadaptation: Lehnt der Test die Nullhypothese in 2.9 ab, wird die Anpassung der RM-Geraden im Zeitfenster der Länge n_t nicht als gut bewertet. In diesem Fall wird n_t verkleinert und die RM-Gerade im verkleinerten Zeitfenster neu geschätzt. Diese Entscheidungsregel bildet das Herzstück des aoRM-Algorithmus, für den die folgenden Inputargumente zu wählen sind:

- Minimal und maximal mögliche Fensterbreiten n_{\min} und n_{\max} , so dass $n_t \in \{n_{\min}, \dots, n_{\max}\} \subset \mathbb{N}$.
- Die maximale Länge n_I des Teilfensters I_t , so dass $I_t = \{n_t - n_{I_t} + 1, \dots, n_t\}$ mit n_{I_t} aus (2.11).

Die Wahl der maximalen Fensterbreite n_{\max} sollte unter Berücksichtigung der Annahme der lokalen Linearität (2.3) geschehen. Sehr große Fensterbreiten sind mit dieser Annahme oftmals schwer vereinbar. Des Weiteren sollte ein Zeitfenster nur aktuelle Information wiedergeben, also keine zu weit zurückliegenden Beobachtungen enthalten. Zudem ist zu berücksichtigen, dass größere Fensterbreiten zu längeren Rechenzeiten führen.

Die minimale Fensterbreite n_{\min} bestimmt die Anzahl an Ausreißern, denen der Filter in einem Zeitfenster widerstehen kann. Entscheidend für die Wahl von n_{\min} ist der Umfang der größten zu erwartenden Ausreißergruppe. So sollte n_{\min} beispielsweise mindestens den Wert 21 haben, damit der RM einer Gruppe von zehn aufeinander folgenden Ausreißern widerstehen kann. Diese Überlegung ist motiviert durch den Bruchpunkt des RM von $\lfloor n/2 \rfloor / n \approx 50\%$. Der RM wird jedoch bereits bei einem Ausreißeranteil von etwa $1/3$ stark verzerrt. Unter diesem Aspekt ist es unter Umständen sinnvoller, n_{\min} als das Dreifache des Umfangs der größten zu erwartenden Ausreißergruppe zu wählen.

Die Fensterbreitenadaption des aoRM wird maßgeblich von der Wahl von n_{I_t} beeinflusst. Je kleiner n_{I_t} , desto eher reagiert der aoRM auf plötzliche Signalveränderungen. Gleichzeitig sinkt jedoch die Resistenz gegenüber Ausreißer bedingten falschen Testentscheidungen. In Kapitel 4.1 werden die *Resistance to Rejection* und *Resistance to Acceptance* (RR und RA, Ylvisaker, 1977) des aoRM-Tests ermittelt. Grob gesagt beschreibt die RR (RA) den kleinsten Anteil m/n an Beobachtungen in einer Stichprobe vom Umfang n , der ersetzt werden muss, damit der Test die Nullhypothese ablehnt (nicht ablehnt) – unabhängig von den übrigen nicht-kontaminierten Daten. In Kapitel 4.1 wird gezeigt, dass für alle Fensterstichproben vom Umfang n_t durch Ersetzen von $m_{\text{RR}} = \lfloor (n_{I_t} + c_\alpha(n_t, n_{I_t})) / 2 \rfloor + 1$ Beobachtungen eine Ablehnung von H_0 erreicht werden kann. Auf der anderen Seite genügt es, $m_{\text{RA}} = \lceil (n_{I_t} - c_\alpha(n_t, n_{I_t})) / 2 \rceil$ Beobachtungen geeignet zu ersetzen um eine Nicht-Ablehnung von H_0 zu erwirken. Mit Rücksicht auf die Test Resistances sollten n_{I_t} und $c_\alpha(n_t, n_{I_t})$ derart auf den Umfang m^* der größten zu erwartenden Ausreißergruppe abgestimmt sein, dass $m_{\text{RR}} > m^*$ und $m_{\text{RA}} > m^*$. Dabei hängt der kritische Wert $c_\alpha(n_t, n_{I_t})$ zwar von α , n_t und n_{I_t} ab, der Wertebereich von $c_\alpha(n_t, n_{I_t})$ wird jedoch wegen (2.11) und $n_t \in \{n_{\min}, \dots, n_{\max}\}$ durch die gewählten Inputargumente α , n_I , n_{\min} und n_{\max} festgelegt. Damit z.B. $m_{\text{RR}} > m^* = 10$ erfüllt ist, muss für $\alpha = 0.1$ die Größe der Teilstichprobe $n_{I_t} \geq 16$ sein, da für $\alpha = 0.1$ und $n_{I_t} \geq 16$ gilt $c_\alpha(n_t, n_{I_t}) \in \{4, 5, \dots\}$. Dann ist $m_{\text{RR}} \geq \lfloor (16 + 4) / 2 \rfloor + 1 = 11 > 10 = m^*$. Damit $n_{I_t} \geq 16$, muss allerdings wegen (2.11) auch $n_I \geq 32$ und $n_{\min} \geq 32$ sein. Die gewählten Inputargumente α , n_I , n_{\min} und n_{\max} bestimmen somit die Robustheit des aoRM-Tests im Sinne der Test Resistances.

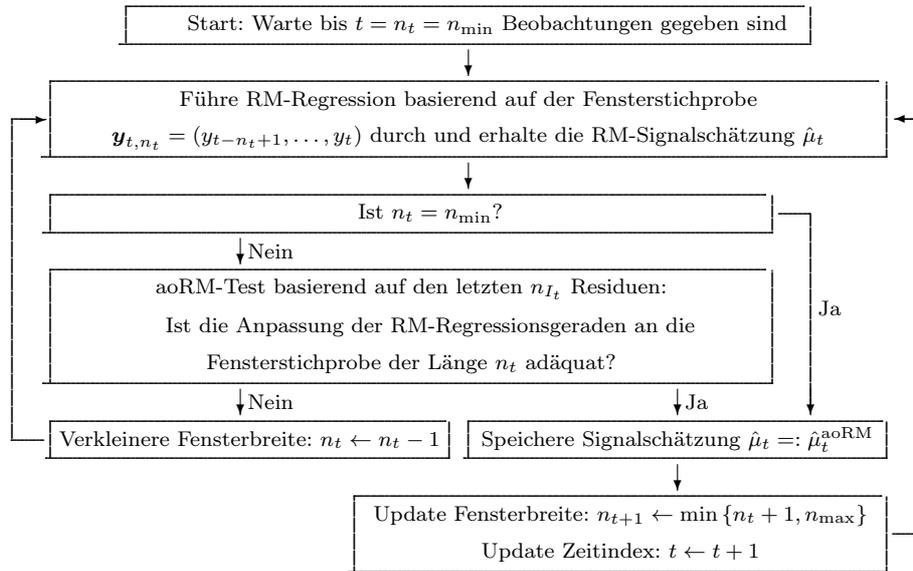


Abbildung 2.2: Flussdiagramm des aoRM-Algorithmus.

Der Algorithmus des aoRM ist in dem Flussdiagramm in Abbildung 2.2 dargestellt. Sobald n_{\min} Beobachtungen gegeben sind, d.h. zum Zeitpunkt $t = n_t = n_{\min}$, startet der erste Durchlauf, indem die RM-Regression auf Basis der Stichprobe $\mathbf{y}_{t,n_t} = (y_{t-n_t+1}, \dots, y_t)$ durchgeführt wird. Gilt zu einem Zeitpunkt t , dass $n_t = n_{\min}$ (was für den ersten Durchlauf immer der Fall ist), so gibt der Algorithmus das anhand der Stichprobe \mathbf{y}_{t,n_t} geschätzte Niveau der RM-Regressionsgeraden $\hat{\mu}_t$ als Signalschätzung $\hat{\mu}_t^{\text{aoRM}}$ aus. Falls jedoch $n_t > n_{\min}$, wird der beschriebene aoRM-Test durchgeführt. Kann der Test die Nullhypothese aus (2.9) nicht ablehnen, wird die Anpassung der RM-Regressionsgeraden im Zeitfenster der Länge n_t als gut bewertet und $\hat{\mu}_t$ als Signalschätzung $\hat{\mu}_t^{\text{aoRM}}$ ausgegeben. Lehnt der Test jedoch die Nullhypothese ab, so wird die Fensterbreite n_t verkleinert, indem die älteste Beobachtung aus der Fensterstichprobe entfernt wird. Daraufhin wird anhand der verkleinerten Stichprobe erneut eine RM-Regression durchgeführt. Diese Vorgänge werden so lange wiederholt, bis entweder $n_t = n_{\min}$ oder die Nullhypothese nicht abgelehnt werden kann. Dann gibt der Algorithmus das mit dem RM geschätzte $\hat{\mu}_t$ als Signalschätzung $\hat{\mu}_t^{\text{aoRM}}$ aus. Anschließend wird das Update des Zeitfensters für den nächsten Zeitpunkt $t + 1$ durchgeführt. Dazu wird die neu hinzukommende Beobachtung y_{t+1} in die Fensterstichprobe aufgenommen, wodurch sich deren Umfang auf $n_t + 1$ erhöht. Ist der Umfang dadurch größer als n_{\max} , wird die älteste Fensterbeobachtung aus der Stichprobe entfernt. Der Update-Vorgang wird abgeschlossen, indem der Algorithmus den Zeitindex auf $t \leftarrow t + 1$ setzt.

Der hier beschriebene aoRM nutzt einen sog. *linearen* Suchalgorithmus zur Bestimmung einer adäquaten Fensterbreite n_t : Diese wird so lange um den Wert eins verkleinert, bis der Vorzeichentest die Anpassung als gut bewertet oder bis $n_t = n_{\min}$ ist. Dabei wird in jeder Iteration eine RM-Regression durchgeführt, weshalb der lineare Suchalgorithmus mit einem hohen Rechenaufwand verbunden ist. So müssen zu einem Zeitpunkt t im Extremfall $n_{\max} - n_{\min}$ Iterationen gerechnet werden. Da die Rechenzeit bei hohen Messfrequenzen kritisch ist, schlägt Schettlinger (2009, Kap. 3.2.9) die *binäre* Suche als Alternative zur linearen Suche der Fensterbreite vor. Bezeichne $n_{t,i}$ die in der i -ten Iteration zum Zeitpunkt t gefundene Fensterbreite, wobei $n_{t,0} = \min\{n_{t-1}, n_{\max}\}$. Die binäre Suche nutzt das Prinzip der Intervallschachtelung. Wird die Nullhypothese in der ersten Iteration abgelehnt, so wird direkt $n_{t,1} \leftarrow n_{\min}$ gesetzt. Anschließend wird getestet, ob die Anpassung der RM-Geraden im Zeitfenster $\{t - n_{t,1} + 1, \dots, t\}$ gut ist. Falls der Test die Nullhypothese ablehnt, wird das RM-Niveau $\hat{\mu}_t$ als Signalschätzung ausgegeben. Lehnt der Test diese jedoch nicht ab, wird $n_{t,2} \leftarrow \lceil (n_{t,1} + n_{t,0})/2 \rceil$ gesetzt, eine erneute RM-Regression durchgeführt und die Anpassungsgüte getestet. Je nach Testentscheidung für $n_{t,2}$ ergibt sich als Menge der möglichen Fensterbreiten entweder $\{n_{t,1}, \dots, n_{t,2}\}$ oder $\{n_{t,2}, \dots, n_{t,0}\}$. Die Anzahl der möglichen Fensterbreiten wird also mit jeder Iteration halbiert, so lange bis eine Fensterbreite eindeutig feststeht. Dadurch ergeben sich maximal $\log(n_{\max} - n_{\min}) + 2$ Iterationen pro Zeitpunkt, weshalb die binäre Suche wesentlich geringere Rechenzeiten als die lineare Suche benötigt. Allerdings liefert die lineare Suche im Gegensatz zur binären Suche die bzgl. des Anpassungstest größtmögliche Fensterbreite, was glattere Signalschätzungen zur Folge hat. Als dritte Methode schlägt Schettlinger (2009) die *geometrische* Suche vor: Wird H_0 in der i -ten Iteration abgelehnt, so wird $n_{t,i} \leftarrow n_{t,0} - \sum_{j=1}^i 2^{j-1}$ gesetzt. Falls H_0 in der i -ten Iteration nicht verworfen werden kann, wird eine binäre Suche innerhalb der Grenzen $n_{t,i}$ und $n_{t,i-1}$ durchgeführt. Die Rechenzeit der geometrischen Suche ist von der Ordnung $\mathcal{O}(\log(n_{\max} - n_{\min}))$, also von der selben Ordnung wie die Rechenzeit der binären Suche. Auf Basis einer umfangreichen Simulationsstudie empfiehlt Schettlinger (2009, S. 71) die geometrische Suche, da diese einen Kompromiss zwischen einer schnellen Rechenzeit und der größtmöglichen Wahl der Fensterbreite bietet.

Nach Level Shifts und Trendwechseln sind die Signalschätzungen des aoRM oftmals derart stark verzerrt, dass sie außerhalb des Stichprobenintervalls $[\min(\mathbf{y}_{t,n_t}), \max(\mathbf{y}_{t,n_t})]$ liegen. Dies ist darin begründet, dass die Signalschätzungen mit dem hoch-robusten RM zum rechten äußeren Designpunkt des Zeitfensters erfolgen. Ein Trend wird demzufolge zu-

nächst extrapoliert, auch wenn im Zeitfenster bereits Beobachtungen enthalten sind, die einem entgegengesetzten Trend folgen. (Ein Level Shift kann in diesem Zusammenhang als zwei direkt aufeinander folgende, stark ausgeprägte Trendwechsel aufgefasst werden.) In der Folge kommt es zu einem „Überschießen“ der RM-Signalschätzungen, das auch in Abbildung 2.1 auf Seite 11 zu sehen ist. Solche deutlichen Verzerrungen sind offenbar auch für kleine Fensterbreiten nicht ausgeschlossen, wenn auch meist weniger ausgeprägt als für große Fensterbreiten. Die von Schettlinger et al. (2010a) vorgeschlagene *Restrict-to-Range-Regel* bietet eine einfache doch wirkungsvolle Lösung des Problems, indem sie die aoRM-Signalschätzungen auf das Stichprobenintervall beschränkt:

$$\hat{\mu}_t^{\text{aoRM}} = \begin{cases} \min(\mathbf{y}_{t,n_{\text{rtr}}}), & \text{falls } \hat{\mu}_t^{\text{aoRM}} < \min(\mathbf{y}_{t,n_{\text{rtr}}}) \\ \max(\mathbf{y}_{t,n_{\text{rtr}}}), & \text{falls } \hat{\mu}_t^{\text{aoRM}} > \max(\mathbf{y}_{t,n_{\text{rtr}}}) \\ \hat{\mu}_t^{\text{aoRM}}, & \text{sonst} \end{cases} \quad (2.12)$$

Dabei ist der Umfang n_{rtr} der Fensterstichproben $\mathbf{y}_{t,n_{\text{rtr}}}$ vom Anwender festzulegen. Schettlinger et al. (2010a) empfehlen n_{rtr} relativ klein zu wählen, z.B. $n_{\text{rtr}} = n_{I_t}$, damit weit zurück liegende Ausreißer die Grenzen des Stichprobenintervalls nicht beeinflussen.

Für die freie Software R (R Development Core Team, 2011) besteht eine Funktion des aoRM, bezeichnet mit `adore.filter` (**ad**aptive **on**line **re**peated median). Diese ist enthalten im R-Paket `robfilter` (Fried et al., 2012), welches auf dem CRAN-Server zur Verfügung steht:

<http://cran.r-project.org/web/packages/robfilter/>.

Eine Erweiterung des aoRM für multivariate Zeitreihen wird von Borowski et al. (2009) entwickelt. Sie kombinieren die Fensterbreitenadaption des aoRM mit der multivariaten *Trimmed Repeated Median-Least Squares*-Regression (TRM-LS) von Lanius und Gather (2010). Die daraus resultierende *adaptive online Trimmed Repeated Median-Least Squares*-Regression (aoTRM-LS) sowie die multivariate TRM-LS-Regression sind speziell für die Echtzeit-Signalextraktion aus multivariaten Zeitreihen entwickelt und werden in Kapitel 2.2.1 und 2.2.2 vorgestellt.

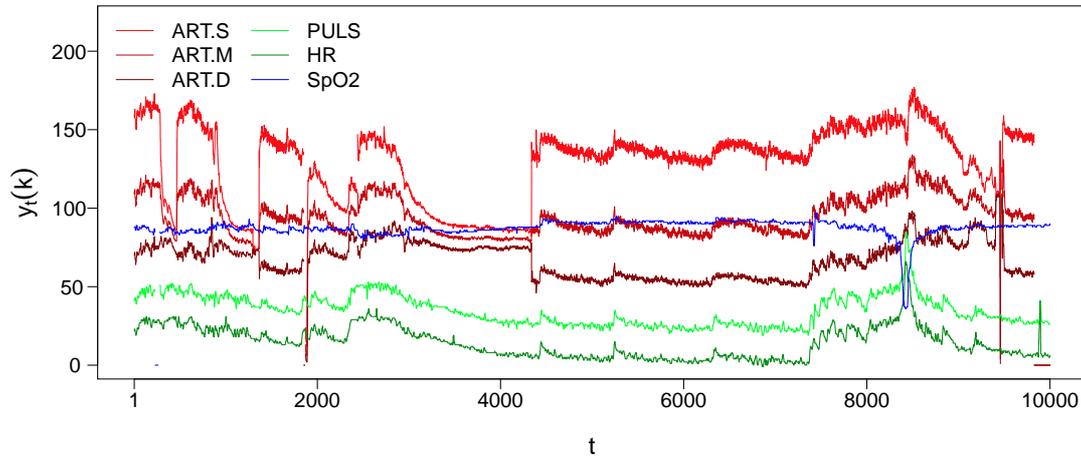


Abbildung 2.3: Multivariate Zeitreihe hämodynamischer Variablen aus der intensivmedizinischen Patientenüberwachung.

2.2 Multivariate Signalextraktionsverfahren

Multivariate Datenströme wie in Abbildung 2.3 ergeben sich z.B. bei der Überwachung des Patientenzustands in der Intensivmedizin, dem sog. *Online-Monitoring*. Diese Abbildung zeigt die sekundlich erfolgten Messungen der Vitalparameter Herzfrequenz (HR), Puls (jeweils in Schläge pro Minute), Sauerstoffsättigung des Blutes (SpO₂, in %) und systolischer, mittlerer und diastolischer Blutdruck (ART.S, ART.M, ART.D, in mmHg). Die Zeitreihen des Puls und der Herzfrequenz sind um 60 bzw. 80 Einheiten nach unten verschoben.

Zwischen den Variablen, insbesondere zwischen der Herzfrequenz und dem Puls sowie zwischen dem systolischen, mittleren und diastolischen Blutdruck, bestehen offensichtlich starke Abhängigkeiten. Die Sauerstoffsättigung des Blutes zeigt keine auf den ersten Blick ersichtlichen Abhängigkeiten zu den übrigen Variablen. Die Abhängigkeitsstruktur stellt möglicherweise eine wichtige Information dar, die im Zuge der parallelen Anwendung eines univariaten Filters auf die einzelnen univariaten Zeitreihen nicht beachtet würde. Im Gegensatz dazu beziehen die nachfolgend vorgestellten multivariaten Filter die Abhängigkeiten zwischen den univariaten Komponenten einer K -variaten Zeitreihe mit ein.

Allen in diesem Kapitel präsentierten multivariaten Filtern liegt dasselbe Modell für multivariate Zeitreihen zugrunde, das sich analog zum univariaten Modell (2.1) ergibt. Für eine K -variante Zeitreihe $(\mathbf{y}_t)_{t \in \mathbb{Z}}$ mit $\mathbf{y}_t = (y_t(1), \dots, y_t(K))' \in \mathbb{R}^K$ wird angenommen, dass sie

aus einem multivariaten Signal besteht, das von Rauschen und Ausreißern überlagert ist:

$$\mathbf{Y}_t = \boldsymbol{\mu}_t + \boldsymbol{\varepsilon}_t + \boldsymbol{\eta}_t. \quad (2.13)$$

Dabei ist $\boldsymbol{\mu}_t = (\mu_t(1), \dots, \mu_t(K))' \in \mathbb{R}^K$ der K -dimensionale Signalvektor zum Zeitpunkt t und $\boldsymbol{\eta}_t \in \mathbb{R}^K$ ein Ausreißer erzeugender Prozess. Die Fehler $\boldsymbol{\varepsilon}_t \in \mathbb{R}^K$ werden als Prozess unabhängiger Zufallsvariablen mit $E(\boldsymbol{\varepsilon}_t) = \mathbf{0}$ angenommen. Die Kovarianzmatrix der Fehler $\text{Cov}(\boldsymbol{\varepsilon}_t) = \boldsymbol{\Sigma}_t \in \mathbb{R}^{K \times K}$ ist zeitveränderlich und nicht notwendigerweise diagonal, d.h. Korrelationen zwischen den univariaten Fehlerkomponenten sind möglich. Die Fehlervarianz der k -ten Komponente sei mit $\text{Var}(\varepsilon_t(k)) = \text{Cov}(\varepsilon_t(k), \varepsilon_t(k)) = \sigma_t^2(k)$ bezeichnet, wobei $k = 1, \dots, K$.

Analog zur univariaten regressionsbasierten Signalextraktion wird angenommen, dass sich jede Komponente $\mu_t(k)$ des multivariaten Signals lokal „gut“ durch eine Gerade approximieren lässt:

$$\mu_{t+i}(k) \approx \mu_t(k) + \beta_t(k) \cdot i, \quad i \in I. \quad (2.14)$$

Dabei ist $\mu_{t+i}(k)$ das Signal der k -ten Komponente zum Zeitpunkt $t+i$ und $\beta_t(k)$ die Steigung der k -ten Komponente im Zeitfenster, das durch die Indexmenge I spezifiziert ist. Die komponentenweisen Signale und Steigungen werden zu Vektoren $\boldsymbol{\mu}_t = (\mu_t(1), \dots, \mu_t(K))' \in \mathbb{R}^K$ und $\boldsymbol{\beta}_t = (\beta_t(1), \dots, \beta_t(K))' \in \mathbb{R}^K$ zusammengefasst. Für die zeitverzögerte Signalschätzung, also die Schätzung des Signalvektors zum mittleren Zeitpunkt t des Zeitfensters $\{t-w, \dots, t, \dots, t+w\}$, ist $I = \{-w, \dots, w\}$. Für die Online-Signalschätzung, also für die Schätzung des Signalvektors zum letzten Zeitpunkt t des Zeitfensters $\{t-n+1, \dots, t\}$, ist $I = \{-n+1, \dots, 0\}$.

Lanius (2005, Kap. 5.2) und Lanius und Gather (2010) entwickeln die *Trimmed Repeated Median-Least Squares-Regression* (TRM-LS) als multivariaten Filter, der robuste und effiziente aber zeitverzögerte Signalschätzungen liefert. Der TRM-LS-Filter extrahiert das K -dimensionale Signal in einem gleitenden Zeitfenster fester Länge mittels zweier hintereinander ausgeführter Regressionsschätzungen. Dabei werden im Zeitfenster zunächst parallel K univariate RM-Regressionen durchgeführt; anhand der RM-Residuen werden dann mögliche Ausreißer detektiert und aus der Stichprobe entfernt; auf Basis der getrimmten Stichprobe wird schließlich eine multivariate LS-Regression zur Schätzung des Signalvektors vorgenommen. Der TRM-LS-Filter wird in Kapitel 2.2.1 vorgestellt.

Da es sich bei dem TRM-LS um einen zeitverzögerten und nicht-adaptiven Filter handelt, entwickeln Borowski et al. (2009) den *adaptive online TRM-LS* (aoTRM-LS), der sich aus

der Kombination des univariaten aoRM und des multivariaten TRM-LS ergibt. Dabei wird zu jedem Zeitpunkt t mit dem aoRM für jede der K univariaten Komponenten eine individuell passende Fensterbreite adaptiert; anhand der K individuellen Fensterbreiten wird dann eine gemeinsame Fensterbreite n_t ermittelt; im Zeitfenster der Länge n_t wird schließlich die TRM-LS Regression zur Schätzung des Signalvektors durchgeführt. Der aoTRM-LS ist somit ein multivariates, adaptives und nicht zeitverzögertes Signalextraktionsverfahren, das in Kapitel 2.2.2 präsentiert wird.

2.2.1 Der Trimmed Repeated Median-Least Squares Filter

Lanius (2005, Kap. 5.2) und Lanius und Gather (2010) entwickeln die TRM-LS-Regression als multivariaten Filter, der den Signalvektor $\boldsymbol{\mu}_t$ in einem gleitenden Zeitfenster $\{\mathbf{y}_{t+i} : i \in I\}$ mit $I = \{-w, \dots, w\}$ und Fensterbreite $|I| = 2w + 1 = n$ zum zentralen Designpunkt t schätzt und daher eher für die retrospektive Signalextraktion geeignet ist. Der TRM-LS lässt sich jedoch leicht zu einem Online-Filter modifizieren, der den Signalvektor in einem gleitenden Zeitfenster $\{\mathbf{y}_{t+i} : i \in I\}$ mit $I = \{-n + 1, \dots, 0\}$ und $|I| = n$ zum rechten Designpunkt t schätzt (Borowski et al., 2009). Der folgende Algorithmus des TRM-LS zur Berechnung der Signal- und Steigungsschätzung $\hat{\boldsymbol{\mu}}^{\text{TRM-LS}} \in \mathbb{R}^K$ und $\hat{\boldsymbol{\beta}}^{\text{TRM-LS}} \in \mathbb{R}^K$ ist sowohl für den zeitverzögerten Ansatz als auch für den Online-Ansatz gültig:

1. Führe für jede Komponente $k \in \{1, \dots, K\}$ anhand der Fensterstichprobe $\{y_{t+i}(k) : i \in I\}$ eine RM-Regression durch und kombiniere die K RM-Schätzungen $(\hat{\beta}_t(k), \hat{\mu}_t(k))$ zu Vektoren

$$\begin{aligned}\hat{\boldsymbol{\beta}}_t &= (\hat{\beta}_t(1), \dots, \hat{\beta}_t(K))', \\ \hat{\boldsymbol{\mu}}_t &= (\hat{\mu}_t(1), \dots, \hat{\mu}_t(K))'.\end{aligned}$$

2. Berechne die Residuenvektoren

$$\mathbf{r}_{t+i} = \mathbf{y}_{t+i} - (\hat{\boldsymbol{\mu}}_t + \hat{\boldsymbol{\beta}}_t \cdot i), \quad i \in I,$$

mit $\mathbf{r}_{t+i} = (\mathbf{r}_{t+i}(1), \dots, \mathbf{r}_{t+i}(K))'$.

3. Schätze die lokale Kovarianzmatrix der Fehler $\boldsymbol{\Sigma}_t \in \mathbb{R}^{K \times K}$ anhand der Residuenvektoren $\{\mathbf{r}_{t+i} : i \in I\}$.
4. Bestimme

$$S_t := \left\{ i \in I : \mathbf{r}'_{t+i} \hat{\boldsymbol{\Sigma}}_t^{-1} \mathbf{r}_{t+i} \leq d_n \right\}.$$

Die Indexmenge S_t umfasst die Designpunkte im Zeitfenster, deren Residuenvektoren bezüglich der lokalen Kovarianzstruktur eine Mahalanobis-Distanz aufweisen, die nicht größer ist als ein vorgegebener Wert d_n .

5. Führe anhand der getrimmten Fensterstichprobe $\{\mathbf{y}_{t+i} : i \in S_t\}$ eine LS-Regression durch und speichere die LS-Schätzung von $\boldsymbol{\mu}_t$ und $\boldsymbol{\beta}_t$ als $\hat{\boldsymbol{\mu}}_t^{\text{TRM-LS}}$ und $\hat{\boldsymbol{\beta}}_t^{\text{TRM-LS}}$.

In Schritt 4 wird anhand des Werts d_n zwischen Ausreißern und „guten“ Beobachtungen unterschieden. Dabei wird ein Beobachtungsvektor \mathbf{y}_{t+i} als Ausreißer angesehen, wenn der zugehörige Residuenvektor \mathbf{r}_{t+i} zu stark von der geschätzten lokalen Kovarianzstruktur der Fehler abweicht. Nach Lanius und Gather (2010) und Maronna und Zamar (2002) kann $d_n = \chi_K^2(\alpha)$ oder

$$d_n = \frac{\chi_K^2(\alpha) \cdot \text{med}\{\mathbf{r}'_{t+i} \hat{\boldsymbol{\Sigma}}_t^{-1} \mathbf{r}_{t+i} : i \in I\}}{\chi_K^2(0.5)}$$

mit z.B. $\alpha = 0.9$ oder $\alpha = 0.95$ gewählt werden, wobei $\chi_K^2(\omega)$ das ω -Quantil der χ^2 -Verteilung mit K Freiheitsgraden ist. Hintergrund ist, dass die quadrierten Mahalanobis-Distanzen $\mathbf{r}'_{t+i} \hat{\boldsymbol{\Sigma}}_t^{-1} \mathbf{r}_{t+i}$ asymptotisch einer χ^2 -Verteilung folgen, falls die Residuen unabhängig normalverteilt sind.

Zur Schätzung der lokalen Fehler-Kovarianzmatrix $\boldsymbol{\Sigma}_t$ in Schritt 3 empfehlen Lanius und Gather (2010) auf Basis einer umfangreichen Vergleichsstudie, den *orthogonalisierten Gnanadesikan-Kettenring-Schätzer* (OGK, Maronna und Zamar, 2002) auf die Residuen $\mathbf{r}_{t+i} \in \mathbb{R}^K$ im Zeitfenster anzuwenden. Der OGK ist eine Modifikation des Gnanadesikan-Kettenring-Kovarianzschätzers (GK, Gnanadesikan und Kettenring, 1972). Der GK nutzt die Gleichheit $\text{Cov}(X, Y) = (\text{Var}(X + Y) - \text{Var}(X - Y))/4$ zur Kovarianzschätzung auf Basis eines univariaten Skalenschätzers $\hat{\sigma}(\cdot)$. Der Skalenschätzer $\hat{\sigma}(\cdot)$ bestimmt dabei die Robustheit, da der maximal mögliche Explosionsbruchpunkt des OGK dem Bruchpunkt von $\hat{\sigma}(\cdot)$ entspricht. Beim OGK sind die paarweisen, robusten GK-Kovarianzschätzungen derart modifiziert, dass die OGK-Kovarianzmatrix positiv definit und approximativ affin äquivariant ist. Sei $\mathbf{X} \in \mathbb{R}^{n \times K}$ eine Stichprobenmatrix mit den Zeilen $\mathbf{x}_i = (x_{i1}, \dots, x_{in})$ und Spalten $X_k = (x_{1k}, \dots, x_{nk})'$, $i = 1, \dots, n$ und $k = 1, \dots, K$. Dann ergibt sich die Kovarianzschätzung $\text{OGK}_{\hat{\sigma}}(\mathbf{X})$ wie folgt:

1. Bestimme $\mathbf{D} = \text{diag}(\hat{\sigma}(X_1), \dots, \hat{\sigma}(X_K))$, so dass $\mathbf{Y} = \mathbf{X}\mathbf{D}^{-1}$.
2. Schätze die Korrelationsmatrix \mathbf{R} von \mathbf{X} mit

$$R_{ij} = (\hat{\sigma}(Y_i + Y_j)^2 - \hat{\sigma}(Y_i - Y_j)^2) / 4$$

und $R_{ii} = 1$, wobei Y_i die i -te Spalte von \mathbf{Y} bezeichnet.

3. Führe eine Eigenwertzerlegung $\mathbf{R} = \mathbf{E}\mathbf{\Lambda}\mathbf{E}'$ durch, wobei $\mathbf{\Lambda} = \text{diag}(\lambda_1, \dots, \lambda_K)$ die geordneten Eigenwerte und \mathbf{E} die zugehörigen Eigenvektoren von \mathbf{R} enthält.
4. Erhalte die OGK $_{\hat{\sigma}}$ -Kovarianzschätzung durch $\text{OGK}_{\hat{\sigma}}(\mathbf{X}) = \mathbf{A}\mathbf{\Gamma}\mathbf{A}'$ mit $\mathbf{A} = \mathbf{D}\mathbf{E}$ und $\mathbf{\Gamma} = \text{diag}(\hat{\sigma}(Z_1)^2, \dots, \hat{\sigma}(Z_K)^2)$, wobei Z_i die i -te Spalte von $\mathbf{Z} = \mathbf{X}(\mathbf{A}')^{-1}$ bezeichne.

Als Skalenschätzer $\hat{\sigma}(\cdot)$ empfehlen Lanius und Gather den robusten Q_n (Rousseeuw und Croux, 1993), der einen asymptotischen Bruchpunkt von 50% für Daten in allgemeiner Lage besitzt. Zudem beschränken sie den Q_n -Schätzer durch den Wert 0.02 nach unten:

$$\hat{\sigma}(\cdot) = \max \{Q_n(\cdot), 0.02\}. \quad (2.15)$$

Diese Regel soll die Nicht-Singularität und somit die Invertierbarkeit von $\hat{\Sigma}_t$ sicherstellen, vgl. Schritt 4 des TRM-LS-Algorithmus.

Der multivariate TRM-LS-Filter basiert auf der Idee des von Bernholt et al. (2006) entwickelten univariaten TRM, vgl. Kapitel 2.1.1. Wie beim univariaten TRM wird das Signal beim multivariaten TRM-LS nach der Ausreißerbereinigung mittels LS-Regression geschätzt, was die Effizienz der Signalschätzung verbessert. Im Gegensatz zum univariaten TRM ist der multivariate TRM-LS jedoch zudem in der Lage, Ausreißer bzgl. der lokalen Kovarianzstruktur zu detektieren. Diese würden durch die komponentenweise Anwendung einer univariaten Ausreißerererkennung womöglich nicht entdeckt werden.

2.2.2 Der adaptive online Trimmed Repeated Median-Least Squares Filter

Borowski et al. (2009) vereinen den multivariaten TRM-LS mit dem univariaten aoRM aus Kapitel 2.1.2 zu einem multivariaten Filter mit adaptiver Fensterbreitenwahl, dem *adaptive online Trimmed Repeated Median-Least Squares*-Filter (aoTRM-LS). Der aoTRM-LS adaptiert mittels aoRM zu jedem Zeitpunkt t für jede Komponente der multivariaten Zeitreihe eine individuelle Fensterbreite $n_t(k)$, $k = 1, \dots, K$. Anhand der K individuellen Fensterbreiten wird dann eine für die gesamte multivariate Zeitreihe zum Zeitpunkt t gültige Fensterbreite n_t bestimmt. Für den aoTRM-LS sind daher die Inputargumente des aoRM zu wählen, d.h. eine minimale und maximale Fensterbreite n_{\min} und n_{\max} sowie die Anzahl an Residuen vorzeichen n_I , die für den aoRM-Test betrachtet werden. Dabei sind die Residuen durch die Designpunktmenge $I_t = \{n_t - n_{I_t} + 1, \dots, t\}$ spezifiziert mit $n_{I_t} = \min\{n_I, \lfloor n_t/2 \rfloor\}$. Der Algorithmus des aoTRM-LS zur Signalschätzung ist wie folgt:

0. Start: Warte bis $t = n_t = n_{\min}$ Beobachtungen gegeben sind.
1. Bestimme im Zeitfenster $\{t - n_t + 1, \dots, t\}$ mit dem aoRM für jede univariate Komponente die individuelle Fensterbreite $n_t(k)$ mit $n_t(k) \in \{n_{\min}, \dots, n_t\}$.
2. Bestimme die gemeinsame Fensterbreite $n_t = \min_{k=1, \dots, K} \{n_t(k)\}$.
3. Führe den TRM-LS-Algorithmus mit $I = \{-n_t + 1, \dots, 0\}$ durch und speichere die Signalschätzung $\hat{\boldsymbol{\mu}}_t^{\text{TRM-LS}} =: \hat{\boldsymbol{\mu}}_t^{\text{aoTRM-LS}}$.
4. Update der Fensterbreite: Setze $n_{t+1} \leftarrow \min\{n_t + 1, n_{\max}\}$;
Update des Zeitindex: Setze $t \leftarrow t + 1$ und gehe zu Schritt 1.

Die Signalschätzung des aoTRM-LS erfolgt mittels multivariater TRM-LS-Regression im Zeitfenster der adaptierten Länge n_t . Dabei muss die gemeinsame Fensterbreite n_t dem Minimum der mit dem aoRM adaptierten individuellen Fensterbreiten $n_t(k)$ entsprechen, damit die Annahme der lokalen Linearität für alle Komponenten vertretbar ist.

Fehlende Werte in der k -ten Komponente eines Datenvektors \mathbf{y}_{t+i} bedeuten fehlende Werte in der k -ten Komponente des entsprechenden RM-Residuenvektors \mathbf{r}_{t+i} , so dass die lokale Kovarianzmatrix der Fehler $\boldsymbol{\Sigma}_t$ nicht geschätzt werden kann, vgl. Schritte 1–3 des TRM-LS-Algorithmus. Ausschließlich Residuenvektoren ohne fehlende Werte zur Schätzung von $\boldsymbol{\Sigma}_t$ zu nutzen, ist offensichtlich mit einem unnötigen Verlust an Information verbunden. Daher schlagen Borowski et al. (2009) eine Alternative zum Umgang mit fehlenden Werten vor, die einen Output in jeder Datensituation gewährleistet. Bei diesem kann es sich allerdings auch um einen fehlenden Wert handeln, wenn nicht genügend Daten für eine gesicherte Signalschätzung vorliegen. Borowski et al. (2009) fordern für jede Komponente k zu jedem Zeitpunkt t , dass zu den letzten n_q Positionen im Zeitfenster mindestens q Beobachtungen gegeben sind. Diese Regel soll sicherstellen, dass die Signalschätzung anhand genügend aktueller Beobachtungen erfolgt. Für diejenigen Komponenten, welche diese Bedingung nicht erfüllen, wird zum Zeitpunkt t keine Signalextraktion durchgeführt und an den entsprechenden Positionen im Vektor der Signalschätzung $\hat{\boldsymbol{\mu}}_t^{\text{aoTRM-LS}}$ ein fehlender Wert ausgegeben. Bei den übrigen Komponenten werden fehlende Werte durch das Niveau der RM-Geraden ersetzt, so dass die entsprechenden Residuen gleich null sind. Dadurch wird zwar die Schätzung von $\boldsymbol{\Sigma}_t$ mit dem Q_n -basierten OGK ermöglicht, allerdings sind die Q_n -Schätzungen offensichtlich in Richtung null verzerrt. Eine Implosion der Q_n -Schätzung wird jedoch durch die Vorschrift (2.15) verhindert.

Eine R-Funktion des aoTRM-LS, bezeichnet mit `madore.filter` (**m**ultivariate **a**daptive **o**nline **r**epeated median), ist im R-Paket `robfilter` (Fried et al., 2012) enthalten, welches auf dem CRAN-Server zur Verfügung steht:

<http://cran.r-project.org/web/packages/robfilter/>.

Der aoTRM-LS soll zu jedem Zeitpunkt eine gemeinsame Fensterbreite n_t wählen, die der aktuellen Datensituation angepasst ist. Borowski et al. (2009) weisen jedoch darauf hin, dass die vom aoTRM-LS adaptierte Fensterbreite n_t häufig kaum über die gewählte minimalen Fensterbreite n_{\min} hinaus wächst. Auch Schettlinger (2009, Kap. 3.3.3) untersucht die vom aoTRM-LS adaptierten Fensterbreiten anhand der retrospektiven Anwendung auf intensivmedizinische Zeitreihen der Dimension $K = 6$. Dabei sind die adaptierten Fensterbreiten in 91.25% der Fälle gleich der gewählten minimalen Fensterbreite $n_{\min} = 60$ und wachsen nicht über den Wert $n_t = 71$ hinaus. In der Folge unterscheidet sich die aoTRM-LS-Signalextraktion kaum von der Signalextraktion des Online-TRM-LS mit fester Fensterbreite $n = n_{\min}$. Die Fensterbreitenadaptation des aoTRM-LS funktioniert also bereits bei kleiner Dimension K nicht zufriedenstellend, so dass ein (positiver) Effekt der Fensterbreitenadaptation nicht oder kaum gegeben ist.

Dieser Umstand wird verschuldet durch die Annahme der lokalen Linearität und der hieraus folgenden Notwendigkeit, die gemeinsame Fensterbreite n_t als Minimum der individuellen Fensterbreiten $n_t(k)$ zu wählen. Wäre die gemeinsame Fensterbreite größer als das Minimum, wäre die Annahme der Linearität und somit die Signalschätzung mittels Geradenanpassung nicht für jede Komponente k vertretbar. Die gemeinsame Fensterbreite muss also gleich dem Minimum der individuellen Fensterbreiten sein. Folglich genügt eine einzelne kleine individuelle Fensterbreite $n_t(k)$, damit der Wert der gemeinsamen Fensterbreite n_t klein ist. Die Bedeutung einer geringen Rate an fälschlichen Verkleinerungen der individuellen Fensterbreiten, d.h. an Fehlern 1. Art des aoRM-Tests, steigt daher mit zunehmender Dimension der Zeitreihe. Das Signifikanzniveau α des aoRM-Tests sollte demnach entsprechend klein gewählt werden. Allerdings erfordert der Test des aoRM einen hohen Wert von α , um eine gute Güte für die Erkennung von Level Shifts und Trendwechseln aufzuweisen, wie die Vergleichsstudie in Kapitel 4.2 zeigt. Der aoRM bzw. dessen Test stellt demnach keine optimale Basis für einen multivariaten Filter mit adaptiver Fensterbreitenwahl dar. Im folgenden Kapitel 3 wird daher ein neuer adaptiver RM-basierter Filter entwickelt, der *Slope Comparing Adaptive Repeated Median* (SCARM). Der Test des SCARM bietet im Gegensatz zum Test des aoRM auch bei kleinem Signifikanzniveau eine hohe Güte.

Einen weiteren Ansatz zur Verbesserung der Fensterbreitenadaptation des aoTRM-LS bietet die *blockweise* Anwendung des aoTRM-LS (Borowski et al., 2009). Dabei bezeichnet ein *Block* bzw. eine *Gruppe* eine Menge von (stark) abhängigen Variablen. Beispielsweise ließe sich die multivariate intensivmedizinischen Zeitreihe in Abbildung 2.3 sinnvoll in die drei Blöcke $\{\text{ART.S, ART.M, ART.D}\}$, $\{\text{PULS, HR}\}$ und $\{\text{SpO}_2\}$ aufteilen. Der aoTRM-LS wird dann nicht auf die gesamte multivariate Zeitreihe der Dimension sechs angewendet, sondern separat auf drei Zeitreihen der Dimension drei, zwei und eins. Falls ein Block nur aus einer Variablen besteht, nutzen Borowski et al. den univariaten aoRM. Bei der blockweisen Anwendung des aoTRM-LS kann eine kleine individuelle Fensterbreite ausschließlich die Fensterbreitenadaptation und somit die Signalextraktion im eigenen Block beeinflussen. Daher werden durch die blockweise Anwendung des aoTRM-LS größere Fensterbreiten adaptiert, weshalb sich glatte Signalschätzungen ergeben – Signalveränderungen werden jedoch weiterhin exakt wiedergegeben (Borowski et al., 2009). Allerdings setzt die blockweise Anwendung des aoTRM-LS eine bekannte und feste Abhängigkeitsstruktur der multivariaten Zeitreihe voraus, was oftmals nicht der Fall ist. Es wird also ein Verfahren benötigt, welches die zeitveränderlichen Zusammenhänge zwischen den einzelnen Komponenten multivariater Datenströme in Echtzeit anzeigt, so dass eine automatische, zeitabhängige Blockbildung erfolgen kann. Hierzu wird in Kapitel 5 ein auf dem SCARM basierendes Verfahren entwickelt, das den hohen Anforderungen der Echtzeit-Überwachung „schwieriger“ multivariater Datenströme genügt.

3 Der Slope Comparing Adaptive Repeated Median

In Kapitel 3.1 wird zunächst der Test des *Slope Comparing Adaptive Repeated Median* (SCARM, Borowski und Fried, 2011, 2012) zur Erkennung von Signalveränderungen in Form von plötzlichen Niveau- und/oder Trendveränderungen vorgestellt. Die Teststatistik ergibt sich aus der Differenz zweier RM-Steigungen relativ zu der Standardabweichung dieser Differenz. Aus diesem Grund wird in Kapitel 3.2 eine Methodik zur Echtzeit-Schätzung der Varianz von Steigungsdifferenzen entwickelt. Der vollständige SCARM-Algorithmus wird dann in Kapitel 3.3 präsentiert. Um geeignete kritische Werte zur Testentscheidung zu erhalten, wird zudem in Kapitel 3.4 die Verteilung der SCARM-Teststatistik für verschiedene Fehlerverteilungen im Rahmen einer Simulationsstudie untersucht.

3.1 Der SCARM-Test

Der vom SCARM verwendete Test basiert auf der Annahme, dass sich die Beobachtungen als Realisationen von $Y_t = \mu_t + \xi_t$ ergeben, wobei μ_t das unterliegende Signal und $\xi_t = \sigma_t \varepsilon_t + \eta_t$ den Fehlerprozess mit Varianz σ_t^2 bezeichnet, vgl. (2.1). Veränderungen der Fehlervariabilität sind möglich, diese geschehen jedoch allmählich und nicht abrupt. Folglich kann die Fehlervariabilität als lokal konstant angenommen werden, d.h. $\sigma_{t-n+i} = \sigma_t$, $i = 1, \dots, n$. Im Zuge der Entwicklung des neuen Verfahrens wird zunächst angenommen, dass $\sigma_t \varepsilon_t \sim N(0, \sigma_t^2)$ und $\eta_t = 0$ ist.

Zu jedem Zeitpunkt t testet der SCARM die Nullhypothese eines linearen Signals im Zeitfenster $\{t - n + 1, \dots, t\}$ gegen die Alternative, dass zu einem Fensterzeitpunkt t_0 eine Signalveränderung in Form eines Level Shifts und/oder Trendwechsels vorliegt:

$$\begin{aligned} H_0 : \mu_{t-n+i} &= \mu_t + \beta_t \cdot (i - n), \quad i = 1, \dots, n, \\ H_1 : \mu_{t-n+i} &= \begin{cases} \mu_t^{\text{alt}} + \beta_t^{\text{alt}} \cdot (i - n), & i = 1, \dots, t_0 \\ \mu_t^{\text{neu}} + \beta_t^{\text{neu}} \cdot (i - n), & i = t_0 + 1, \dots, n \end{cases}, \end{aligned} \quad (3.1)$$

wobei $t_0 \in \{1, \dots, n - 1\}$ und $\mu_t^{\text{alt}} \neq \mu_t^{\text{neu}}$ und/oder $\beta_t^{\text{alt}} \neq \beta_t^{\text{neu}}$. Unter den obigen Voraussetzungen gilt $Y_{t-n+i} = \mu_{t-n+i} + \sigma_t \varepsilon_{t-n+i}$, so dass H_0 auch geschrieben werden kann als

$$Y_{t-n+i} = \mu_t + \beta_t \cdot (i - n) + \sigma_t \varepsilon_{t-n+i}. \quad (3.2)$$

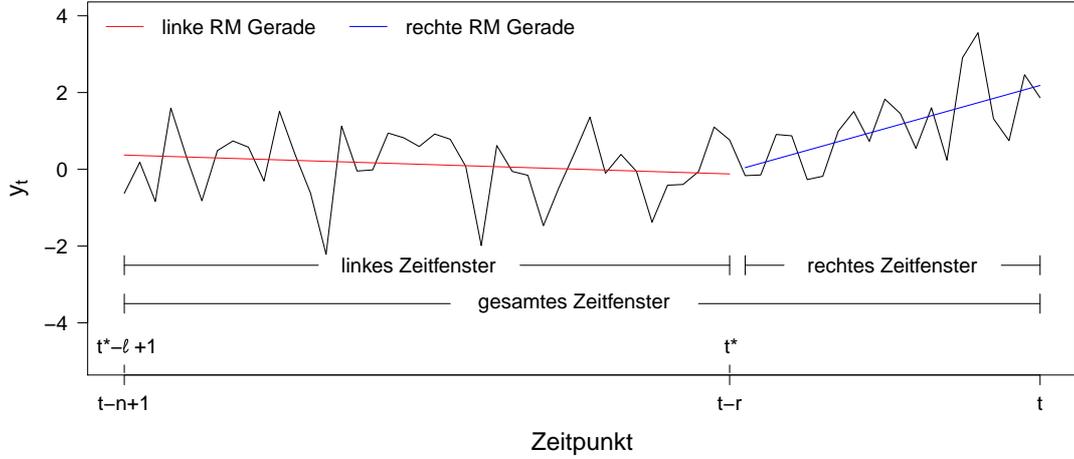


Abbildung 3.1: Aufteilung eines Zeitfensters in ein linkes und rechtes Zeitfenster mit linker und rechter RM Regressionsgerade.

Für den SCARM-Test wird das gesamte Zeitfenster $\{t - n + 1, \dots, t\}$ der Länge n in zwei separate Fenster geteilt, ein linkes Zeitfenster $\{t^* - \ell + 1, \dots, t^*\}$ der Länge ℓ und ein rechtes Zeitfenster $\{t - r + 1, \dots, t\}$ der Länge r . Dabei gilt $t^* = t - r$, $t^* - \ell + 1 = t - n + 1$ und $n = \ell + r$, vgl. Abbildung 3.1. Seien

$$\mathbf{y}_{t,n} = (y_{t-n+1}, \dots, y_t) \in \mathbb{R}^n, \quad \mathbf{y}_{t,r} = (y_{t-r+1}, \dots, y_t) \in \mathbb{R}^r,$$

$$\mathbf{y}_{t^*,\ell} = (y_{t^*-\ell+1}, \dots, y_{t^*}) = (y_{t-n+1}, \dots, y_{t-n+\ell}) = (y_{t-n+1}, \dots, y_{t-r}) \in \mathbb{R}^\ell$$

die gesamte, linke und rechte Fensterstichprobe und $\mathbf{Y}_{t,n}$, $\mathbf{Y}_{t,r}$ und $\mathbf{Y}_{t^*,\ell}$ die entsprechenden Zufallsvektoren. Des Weiteren seien $\hat{\mu}(\mathbf{y}_{t^*,\ell})$ und $\hat{\beta}(\mathbf{y}_{t^*,\ell})$ das RM-Niveau bzw. die RM-Steigung, welche anhand der linken Fensterstichprobe geschätzt werden und die linke RM-Gerade spezifizieren, vgl. Abbildung 3.1:

$$\hat{\mu}_{t^*-\ell+j}(\mathbf{y}_{t^*,\ell}) = \hat{\mu}(\mathbf{y}_{t^*,\ell}) + \hat{\beta}(\mathbf{y}_{t^*,\ell}) (j - \ell) \quad (3.3)$$

mit $j = 1, \dots, \ell$. Das RM-Niveau und die RM-Steigung im rechten Fenster seien bezeichnet mit $\hat{\mu}(\mathbf{y}_{t,r})$ und $\hat{\beta}(\mathbf{y}_{t,r})$. Sie spezifizieren die rechte RM-Gerade:

$$\hat{\mu}_{t-r+k}(\mathbf{y}_{t,r}) = \hat{\mu}(\mathbf{y}_{t,r}) + \hat{\beta}(\mathbf{y}_{t,r}) (k - r) \quad (3.4)$$

mit $k = 1, \dots, r$.

Nach Siegel (1982) ist der RM-Steigungsschätzer für symmetrisch verteilte Fehler unverzerrt. Falls H_0 wahr ist, sind daher auch die RM-Steigungsschätzer im gesamten, linken

und rechten Zeitfenster für symmetrische Fehler unverzerrt:

$$E(\hat{\beta}(\mathbf{Y}_{t,n})) = E(\hat{\beta}(\mathbf{Y}_{t^*,\ell})) = E(\hat{\beta}(\mathbf{Y}_{t,r})) = \beta_t. \quad (3.5)$$

Folglich ist der Erwartungswert der Differenz von $\hat{\beta}(\mathbf{Y}_{t^*,\ell})$ und $\hat{\beta}(\mathbf{Y}_{t,r})$ gleich Null. Aus diesen Überlegungen ergibt sich die SCARM Teststatistik als

$$T_t := T_t(\mathbf{Y}_{t,n}) = \frac{D_t}{\sqrt{\widehat{\text{Var}}(D_t)}}, \quad (3.6)$$

$$D_t := D_t(\mathbf{Y}_{t,n}) = \hat{\beta}(\mathbf{Y}_{t^*,\ell}) - \hat{\beta}(\mathbf{Y}_{t,r}),$$

wobei $\widehat{\text{Var}}(D_t)$ ein Schätzer von $\text{Var}(D_t)$ ist, der in Kapitel 3.2 vorgestellt wird. Ist $|T_t|$ „zu groß“, wird angenommen, dass sich die gesamte Stichprobe $\mathbf{y}_{t,n} := (y_{t-n+1}, \dots, y_t)$ nicht als Realisation von (3.2) ergibt, das Signal also nicht linear ist. Konkret wird die Nullhypothese verworfen, falls

$$T_t = \frac{D_t}{\sqrt{\widehat{\text{Var}}(D_t)}} < q_{\text{low}} \quad \text{oder} \quad T_t = \frac{D_t}{\sqrt{\widehat{\text{Var}}(D_t)}} > q_{\text{up}},$$

wobei als kritischer Wert q_{low} bzw. q_{up} das $(\alpha/2)$ - bzw. $(1 - \alpha/2)$ -Quantil einer t -Verteilung mit f Freiheitsgraden genutzt wird, vgl. Kapitel 3.4. Die Freiheitsgrade der t -Verteilung hängen dabei von der linken und rechten Fensterbreite ℓ und r ab. Für große ℓ und r werden Quantile der Standardnormalverteilung verwendet.

3.2 Die Schätzung der Varianz von RM-Steigungsdifferenzen

Die Varianz der Differenz D_t der RM-Steigungen $\hat{\beta}(\mathbf{Y}_{t^*,\ell})$ und $\hat{\beta}(\mathbf{Y}_{t,r})$ in (3.6) könnte anhand der letzten m Realisationen von D_t , also anhand d_{t-m+1}, \dots, d_t , mittels eines geeigneten Skalenschätzers, z.B. dem robusten MAD, geschätzt werden. Dieser intuitive Ansatz bringt jedoch Probleme mit sich. Abbildung 3.2 (a) zeigt eine simulierte Zeitreihe (y_t) mit mehreren Level Shifts. Darunter ist in (b) die Zeitreihe (d_t) der RM-Steigungsdifferenzen mit $\ell, r = 25$ abgebildet. Wegen $n = \ell + r$ beginnt diese Zeitreihe erst zum Zeitpunkt $t = n = 50$. Die untere Abbildung (c) zeigt die Schätzungen der Standardabweichung von D_t . Diese werden zu jedem Zeitpunkt t mit dem MAD anhand der vorangegangenen $m = 25, 50, 100$ realisierten RM-Steigungsdifferenzen d_{t-m+1}, \dots, d_t geschätzt. Aus diesem Grund beginnen die Zeitreihen für $m = 25, 50, 100$ jeweils zum Zeitpunkt $n + m$.

Kurz nach einem Level Shift in (y_t) zeigt die Zeitreihe (d_t) einen Peak in die entgegengesetzte Richtung. Dieses Verhalten von (d_t) resultiert aus der Tatsache, dass ein Level Shift

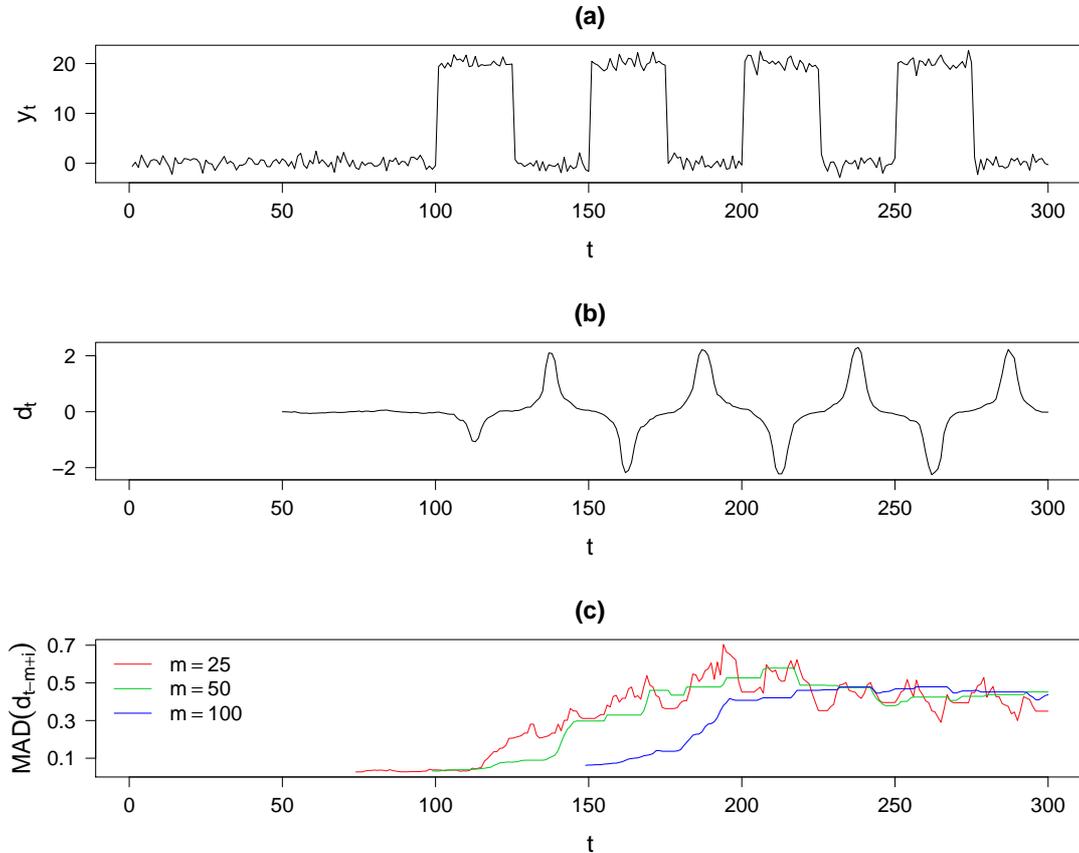


Abbildung 3.2: (a) Simulierte Zeitreihe (y_t); (b) Zeitreihe (d_t) der RM-Steigungsdifferenzen mit $\ell, r = 25$; (c) $\text{MAD}(d_{t-m+1}, \dots, d_t)$ für $m = 25, 50, 100$.

zunächst das rechte Zeitfenster durchläuft, bevor er das linke Zeitfenster erreicht. In dieser Phase weist die rechte RM-Steigungsschätzung einen großen absoluten Wert auf, während sich die linke weiterhin im Bereich um null bewegt. Folglich ist bei einem Level Shift in (y_t) die Differenz $d_t = \hat{\beta}(\mathbf{y}_{t^*,\ell}) - \hat{\beta}(\mathbf{y}_{t,r})$ kurzzeitig deutlich von null verschieden. Erfolgen in der Zeitreihe (y_t) mehrere Level Shifts in kurzer Abfolge hintereinander, ergeben sich entsprechend viele kurz aufeinander folgende Peaks in (d_t). Die Zeitreihe (d_t) weist folglich eine hohe Variabilität auf, so dass $\text{Var}(D_t)$ auch unter Verwendung eines hochrobusten Schätzers wie dem MAD groß geschätzt wird, wie in Abbildung 3.2 (c) zu sehen ist. Offenbar bietet dabei auch ein großes m keinen Schutz vor großen Skalenschätzungen. Große Skalenschätzungen sind problematisch, da sie zu *Masking-Effekten* führen können. Ein Masking-Effekt liegt vor, wenn unter der Alternative H_1 ein betragslich großer Wert d_t durch eine zu große Schätzung von $\text{Var}(D_t)$ geteilt wird, so dass die Teststatistik kleiner als

der kritische Wert ist und H_0 nicht verworfen wird. Der Test kann die Signalveränderung also nicht erkennen, da sie „maskiert“ ist. Die Verhinderung von Masking-Effekten steigert demnach die Güte des Tests bzgl. der Erkennung von Signalveränderungen. Daher wird im Folgenden eine alternative Methode zur Schätzung von $\text{Var}(D_t)$ unter H_0 entwickelt, mit der Masking-Effekte verhindert werden.

Im Zuge der Entwicklung des Schätzverfahrens für $\text{Var}(D_t)$ unter H_0 wird angenommen, dass $\xi_{t-n+i} = \sigma_t \varepsilon_{t-n+i} \sim N(0, \sigma_t^2)$ u.i.v. im gesamten Zeitfenster $\{t-n+1, \dots, t\}$. Verletzungen dieser Annahme werden in Kapitel 3.5 diskutiert. Unter H_0 gilt

$$\text{Var}(D_t) = \text{Var} \left[\hat{\beta}(\mathbf{Y}_{t^*,\ell}) - \hat{\beta}(\mathbf{Y}_{t,r}) \right] = \text{Var} \left[\hat{\beta}(\mathbf{Y}_{t^*,\ell}) \right] + \text{Var} \left[\hat{\beta}(\mathbf{Y}_{t,r}) \right], \quad (3.7)$$

d.h. $\text{Cov}[\hat{\beta}(\mathbf{Y}_{t^*,\ell}), \hat{\beta}(\mathbf{Y}_{t,r})] = 0$, da die ε_t unabhängig normalverteilt sind und sich die linke und rechte Fensterstichprobe nicht überschneiden. Zudem hängen $\text{Var}[\hat{\beta}(\mathbf{Y}_{t,r})]$ und $\text{Var}[\hat{\beta}(\mathbf{Y}_{t^*,\ell})]$ nicht von der wahren Steigung β_t ab:

Satz 3.1.

Unter H_0 ist die Varianz von $\hat{\beta}(\mathbf{Y}_{t,n})$ unabhängig von der wahren Steigung β_t . Sie ist abhängig von der Fehlervarianz σ_t^2 und der Fensterbreite n .

Beweis.

Unter H_0 kann der RM-Steigungsschätzer $\hat{\beta}(\mathbf{Y}_{t,n})$ geschrieben werden als

$$\hat{\beta}(\mathbf{Y}_{t,n}) = \underset{i=1,\dots,n}{\text{med}} \left\{ \underset{j \neq i}{\text{med}} \frac{Y_{t-n+i} - Y_{t-n+j}}{i - j} \right\} = \beta_t + \underset{i=1,\dots,n}{\text{med}} \left\{ \underset{j \neq i}{\text{med}} \frac{\sigma_t \varepsilon_{t-n+i} - \sigma_t \varepsilon_{t-n+j}}{i - j} \right\},$$

da $Y_{t-n+i} = \mu_t + \beta_t(i - n) + \sigma_t \varepsilon_{t-n+i}$ und $Y_{t-n+j} = \mu_t + \beta_t(j - n) + \sigma_t \varepsilon_{t-n+j}$ unter H_0 . Daher ist

$$\text{Var}(\hat{\beta}(\mathbf{Y}_{t,n})) = \sigma_t^2 \cdot \text{Var} \left(\underset{i=1,\dots,n}{\text{med}} \left\{ \underset{j \neq i}{\text{med}} \frac{\varepsilon_{t-n+i} - \varepsilon_{t-n+j}}{i - j} \right\} \right), \quad (3.8)$$

wobei $\text{Var}(\varepsilon_{t-n+i}) = 1$. □

Da die Zufallsvariablen im linken und rechten Zeitfenster unter H_0 dieselbe lineare Struktur aufweisen, folgt daraus

Korollar 3.1.

Unter H_0 sind $\text{Var}(\hat{\beta}(\mathbf{Y}_{t^,\ell}))$ und $\text{Var}(\hat{\beta}(\mathbf{Y}_{t,r}))$ unabhängig von der wahren Steigung β_t . Sie sind abhängig von der Fehlervarianz σ_t^2 und der jeweiligen Fensterbreite ℓ und r .*

An dieser Stelle sei auf die Bedeutung der Gleichung (3.8) im Beweis von Satz 3.1 hingewiesen:

Bemerkung 3.1.

Unter H_0 ist die Varianz von $\hat{\beta}(\mathbf{Y}_{t,n})$ – und somit auch von $\hat{\beta}(\mathbf{Y}_{t^*,\ell})$ und $\hat{\beta}(\mathbf{Y}_{t,r})$ – proportional zur Fehlervarianz σ_t^2 .

Aufgrund der obigen Aussagen lassen sich die Varianz des linken und des rechten RM-Steigungsschätzers als Funktionen $V : \mathbb{N} \times \mathbb{R}^+ \mapsto \mathbb{R}^+$ auffassen mit

$$V(\ell, \sigma_t) := \text{Var} \left(\hat{\beta}(\mathbf{Y}_{t^*,\ell}) \right) \quad \text{und} \quad V(r, \sigma_t) := \text{Var} \left(\hat{\beta}(\mathbf{Y}_{t,r}) \right). \quad (3.9)$$

Nach Bemerkung 3.1 gilt:

$$V(\ell, \sigma_t) = V(\ell, 1) \cdot \sigma_t^2 \quad \text{und} \quad V(r, \sigma_t) = V(r, 1) \cdot \sigma_t^2, \quad (3.10)$$

so dass sich (3.7) unter H_0 wegen (3.9) und (3.10) schreiben lässt als

$$\text{Var}(D_t) = \sigma_t^2 \cdot (V(\ell, 1) + V(r, 1)). \quad (3.11)$$

Dabei bezeichnet $V(n, 1)$ die Varianz der RM-Steigung für Stichproben der Länge n bei Fehlervarianz eins. Approximationen $\hat{V}(n, 1) =: v_n$ werden für $n \geq 5$ anhand des Modells $v(n) = 4.77 \cdot 10^{-7} + 17.71 \cdot n^{-3}$ in (B.1) gewonnen, welches in Anhang B.1 mit Hilfe von Simulationen erstellt wird. Zur Schätzung von σ_t in (3.11) kommen nur robuste Skalenschätzer in Frage. Gather und Fried (2003) vergleichen einige hochrobuste Skalenschätzer, u.a. den MAD und den S_n - und Q_n -Skalenschätzer (Rousseeuw und Croux, 1993), zur Schätzung der Variabilität in Zeitreihen. Sie zeigen auf, dass die Skalenschätzer bei einem nicht-konstanten Signal μ_t stark verzerrt sind, da bei einer direkten Anwendung der Skalenfunktionale auf die gegebenen Beobachtungen nicht nur die Variabilität des Fehlers, sondern auch die des Signals geschätzt wird. Es ist jedoch möglich, den Trend z.B. mit Hilfe des RM-Regressionsschätzers aus den Daten zu entfernen. Die Fehlervariabilität kann dann anhand der Residuen der RM-Regression geschätzt werden. Dieser regressionsbasierte Ansatz setzt jedoch ein lokal zumindest annähernd lineares Signal voraus. Eine Verletzung dieser Annahme, z.B. aufgrund eines Level Shifts, kann zu sehr großen absoluten Residuen führen. In diesem Fall können selbst hochrobuste Skalenschätzer unangemessen große Schätzungen von σ_t liefern. Es besteht also für diesen Ansatz die Gefahr von Masking-Effekten. Abhilfe schaffen die von Rousseeuw und Hubert (1996) vorgeschlagenen

regressions- und *modellfreien* Skalenschätzer. Mit regressionsfreien Skalenschätzern lässt sich σ_t direkt an den Daten schätzen, auch wenn diesen ein linearer Trend unterliegt. Es ist daher vorab keine lokale Regressionsanpassung nötig. Dies ist auch für die modellfreien Skalenschätzer der Fall. Diese benötigen jedoch darüber hinaus kein Modell für das Signal bzw. die Variabilität des Rauschens, so dass sie auch bei nichtlinearen Signalen verlässliche Schätzungen von σ_t liefern. In umfangreichen Studien vergleichen Gelper et al. (2009) und Schettlinger et al. (2010b) mehrere regressions- und modellfreie Echtzeit-Schätzer für σ_t in nichtstationären Zeitreihen und empfehlen den modellfreien Q_δ^{adj} -Skalenschätzer. Für einen Stichprobenvektor $\mathbf{y}_{t,n} := (y_{t-n+1}, \dots, y_t)$ ist der Q_δ^{adj} -Skalenschätzer gegeben durch

$$Q_\delta^{\text{adj}}(\mathbf{y}_{t,n}) = c_n \cdot h_{(\lfloor \delta(n-2) \rfloor)}. \quad (3.12)$$

Dabei ist $h_{(\lfloor \delta(n-2) \rfloor)}$ das δ -Quantil der $n - 2$ vertikalen Höhen

$$h_{t-n+i} = \left| y_{t-n+i} - \frac{y_{t-n+i-1} + y_{t-n+i+1}}{2} \right|, \quad i \in \{2, \dots, n-1\} \quad (3.13)$$

der Dreiecke, die sich aus den $n - 2$ Tripeln von aufeinander folgenden Beobachtungen $y_{t-n+i-1}, y_{t-n+i}, y_{t-n+i+1}$ ergeben. Das c_n in (3.12) ist eine Konstante, welche die Unverzerrtheit des Schätzers für normalverteilte Fehler und den Stichprobenumfang n sicherstellen soll. Mittels Simulationen lassen sich auch Konstanten für andere Fehlerverteilungen gewinnen. Gelper et al. (2009) empfehlen die Wahl von $\delta = 0.5$, da der Q_δ^{adj} -Skalenschätzer hierfür einen guten Kompromiss zwischen Robustheit und Effizienz liefert. Unter Verwendung des $Q_{0.5}^{\text{adj}}$ -Skalenschätzers und der Approximationen $\hat{V}(\ell, 1) := v_\ell$ und $\hat{V}(r, 1) := v_r$ wird $\text{Var}(D_t)$ dann gemäß (3.11) geschätzt durch

$$\widehat{\text{Var}}(d_t) = Q_{0.5}^{\text{adj}}(\mathbf{y}_{t,n})^2 \cdot (v_\ell + v_r) = c_n \cdot h_{(\lfloor n/2 - 1 \rfloor)} \cdot (v_\ell + v_r). \quad (3.14)$$

Der Q_δ^{adj} -Skalenschätzer schätzt σ_t anhand der vertikalen Höhen von Dreiecken, die sich aus jeweils drei aufeinander folgenden Beobachtungen ergeben, weshalb er invariant gegenüber linearen Trends ist. Darüber hinaus ist er unempfindlich gegenüber plötzlichen Signalveränderungen, was folglich auch für den vorgeschlagenen Varianzschätzer $\widehat{\text{Var}}(d_t)$ in (3.14) gilt. Es wird somit verhindert, dass plötzliche Signalveränderungen sich selbst oder nachfolgende Signalveränderungen maskieren.

Unter Verwendung des hier entwickelten Varianzschätzers $\widehat{\text{Var}}(D_t)$ ist die SCARM Teststatistik T_t aus (3.6) unter H_0 symmetrisch um null verteilt, wenn die Fehler symmetrisch um null verteilt sind:

Satz 3.2.

Für $\mathbf{Y}_{t,n} = (Y_{t-n+1}, \dots, Y_t)$ mit

$$Y_{t-n+i} = \mu_t + \beta_t \cdot (i - n) + \sigma_t \varepsilon_{t-n+i}, \quad i = 1, \dots, n,$$

und $\varepsilon_{t-n+i} \sim -\varepsilon_{t-n+i}$ gilt

$$T_t(\mathbf{Y}_{t,n}) \sim -T_t(\mathbf{Y}_{t,n}).$$

Beweis.

Im Folgenden sei der Zeitindex t vernachlässigt, d.h. statt $\mathbf{Y}_{t,n}$ und $T_t(\mathbf{Y}_{t,n})$ wird $\mathbf{Y} := (Y_1, \dots, Y_n)$ und $T(\mathbf{Y})$ betrachtet mit

$$Y_i = \mu + \beta \cdot (i - n) + \sigma \varepsilon_i, \quad i = 1, \dots, n. \quad (3.15)$$

Die SCARM-Teststatistik ist dann $T(\mathbf{Y}) = D(\mathbf{Y})/S(\mathbf{Y})$ mit

$$\begin{aligned} D(\mathbf{Y}) &:= \hat{\beta}(Y_1, \dots, Y_\ell) - \hat{\beta}(Y_{\ell+1}, \dots, Y_n), \\ S(\mathbf{Y}) &:= \sqrt{c_n \cdot (v_\ell + v_r) \cdot H_{(\lfloor n/2 - 1 \rfloor)}}, \\ H_j &:= H_j(\mathbf{Y}) = \left| Y_j - \frac{Y_{j-1} + Y_{j+1}}{2} \right|, \quad j = 2, \dots, n-1. \end{aligned}$$

Wenn (3.15) erfüllt ist, gilt

$$\begin{aligned} \hat{\beta}(Y_1, \dots, Y_\ell) &= \beta + \operatorname{med}_{i=1, \dots, \ell} \left\{ \operatorname{med}_{j \neq i, j=1, \dots, \ell} \frac{\sigma \varepsilon_i - \sigma \varepsilon_j}{i - j} \right\} \\ \text{und } \hat{\beta}(Y_{\ell+1}, \dots, Y_n) &= \beta + \operatorname{med}_{i=\ell+1, \dots, n} \left\{ \operatorname{med}_{j \neq i, j=\ell+1, \dots, n} \frac{\sigma \varepsilon_i - \sigma \varepsilon_j}{i - j} \right\}, \end{aligned}$$

vgl. den Beweis von Satz 3.1. In diesem Fall ist $D(\mathbf{Y})$ also unabhängig von μ und β . Wenn (3.15) erfüllt ist, lässt sich zudem durch einfaches Einsetzen und Kürzen zeigen, dass

$$H_j(\mathbf{Y}) = \sigma \cdot \left| \varepsilon_j - \frac{\varepsilon_{j-1} + \varepsilon_{j+1}}{2} \right|,$$

d.h. auch $H_j(\mathbf{Y})$ und somit $S(\mathbf{Y})$ sind unabhängig von μ und β . Daher kann in (3.15) o.B.d.A. $\mu = \beta = 0$ gesetzt werden, so dass $Y_i = \sigma \varepsilon_i$ und daher $Y_i \sim -Y_i$.

Weiterhin gilt offensichtlich

$$\hat{\beta}(Y_1, \dots, Y_\ell) = -\hat{\beta}(-Y_1, \dots, -Y_\ell) \quad \text{und} \quad \hat{\beta}(Y_{\ell+1}, \dots, Y_n) = -\hat{\beta}(-Y_{\ell+1}, \dots, -Y_n),$$

so dass

$$D(\mathbf{Y}) = \hat{\beta}(Y_1, \dots, Y_\ell) - \hat{\beta}(Y_{\ell+1}, \dots, Y_n) = -D(-\mathbf{Y}).$$

Außerdem gilt offensichtlich $H_j(\mathbf{Y}) = H_j(-\mathbf{Y})$ und daher auch $S(\mathbf{Y}) = S(-\mathbf{Y})$. Somit gilt für die SCARM-Teststatistik

$$T(\mathbf{Y}) = \frac{D(\mathbf{Y})}{S(\mathbf{Y})} = \frac{-D(-\mathbf{Y})}{S(-\mathbf{Y})} = -T(-\mathbf{Y}).$$

Wegen $Y_i \sim -Y_i$ gilt dann $T(\mathbf{Y}) \sim T(-\mathbf{Y}) = -T(\mathbf{Y})$. □

Der SCARM-Test inklusive der beschriebenen Varianzschätzung von RM-Steigungsdifferenzen bilden den Kern des SCARM-Filters, dessen Algorithmus im Folgenden präsentiert wird. Da der SCARM die Fensterbreite zu jedem Zeitpunkt t an die Daten adaptiert, ist diese nachfolgend mit n_t bezeichnet.

3.3 Der SCARM-Algorithmus

Der hier vorgestellte Algorithmus des SCARM bedient sich einiger Prinzipien des aoRM-Algorithmus aus Kapitel 2.1.2, weist jedoch entscheidende Unterschiede auf. Der SCARM benötigt die Wahl der folgenden Inputargumente, die vom Anwender mit Rücksicht auf die Daten zu wählen sind und weiter unten diskutiert werden:

- Das Signifikanzniveau α des Tests.
- Die feste Breite r des rechten Zeitfensters.
- Eine untere Grenze ℓ_{\min} für die linke Fensterbreite ℓ_t (wegen $\ell_t := n_t - r$ ist die linke Fensterbreite nicht fest, sondern zeitvariabel); die Untergrenze ℓ_{\min} garantiert, dass $\ell_t \geq \ell_{\min}$.
- Eine minimale und maximale Fensterbreite n_{\min} und n_{\max} , so dass gilt $n_t \in \{n_{\min}, \dots, n_{\max}\}$ und $n_{\min} < \ell_{\min} + r < n_{\max}$.

Der SCARM-Algorithmus ist in dem Flussdiagramm in Abbildung 3.3 dargestellt. Das Herzstück dieses Algorithmus ist der SCARM-Test auf Linearität des Signals im Zeitfenster der Länge n_t . Falls H_0 nicht verworfen werden kann, wird die Fensterbreite n_t als adäquat bewertet und das Signal μ_t anhand der Fensterstichprobe $\mathbf{y}_{t,n_t} = (y_{t-n_t+1}, \dots, y_t)$ mit dem RM-Niveau $\hat{\mu}_t$ geschätzt. Lehnt der Test die Nullhypothese jedoch ab, wird n_t auf n_{\min} gesetzt und das Signal anhand der verkleinerten Stichprobe geschätzt. Nach der

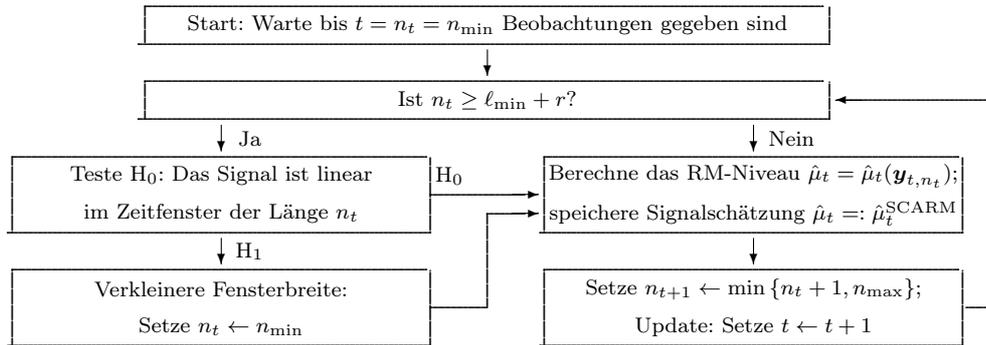


Abbildung 3.3: Flussdiagramm des SCARM-Algorithmus.

Signalschätzung wird das Update des Zeitfensters für den nächsten Zeitpunkt $t + 1$ durchgeführt. Dies geschieht wie beim aoRM, indem die Fensterstichprobe durch Hinzufügen der neu hinzukommenden Beobachtung y_{t+1} vergrößert wird. Wächst dadurch die Fensterbreite über den Wert n_{\max} hinaus, wird die älteste Beobachtung aus der Stichprobe entfernt. Anschließend setzt der Algorithmus $t \leftarrow t + 1$, und das Update ist abgeschlossen.

Der SCARM nutzt offenbar ein anderes Adaptionsprinzip als der aoRM. Bei letzterem geschieht die Verkleinerung des Zeitfensters sukzessiv über eine sequentielle Anwendung des Tests. Im Gegensatz dazu wird beim SCARM bei Ablehnung der Nullhypothese die Fensterbreite direkt auf n_{\min} gesetzt, da von einer Signalveränderung im Zeitfenster ausgegangen wird. Im Idealfall ist n_{\min} derart, dass das Zeitfenster der Länge $n_t = n_{\min}$ nur Beobachtungen *nach* der Signalveränderung enthält. Eine geeignete Wahl von n_{\min} wird unten diskutiert. Des Weiteren kann beim SCARM die Fensterbreite n_t nur dann verkleinert werden, wenn diese nicht kleiner ist als $\ell_{\min} + r$. Diese Regel garantiert, dass der SCARM-Test anhand von mindestens $\ell_{\min} + r$ Beobachtungen durchgeführt wird. Wird n_t zum Zeitpunkt t auf n_{\min} gesetzt (wobei gilt $n_{\min} < \ell_{\min} + r$), so wird der Test erst wieder zum Zeitpunkt $t + \ell_{\min} + r - n_{\min}$ durchgeführt.

3.3.1 Wahl der Inputargumente

Das Signifikanzniveau α des SCARM-Tests bestimmt die erwartete Rate an Fehlern 1. Art, also an fälschlicherweise entdeckten Signalveränderungen. Für z.B. $\alpha = 0.001$ ist auf 1000 Tests unter H_0 im Mittel in etwa eine falsch positive Entdeckung einer Signalveränderung zu erwarten. Die kritischen Werte zur Testentscheidung werden allerdings mittels Simulationen auf Basis von linearen Signalen mit unabhängig normalverteilten Fehlern ermittelt, vgl. Kapitel 3.4. Daher gilt die obige Aussage bzgl. der zu erwartenden Rate an falsch

positiven Entdeckungen lediglich annähernd und nicht für alle Arten von Fehlern. Lassen sich für die Daten andere Fehlerverteilungen vermuten, können die kritischen Werte bei Bedarf mit Hilfe von Simulationen oder auch mittels Bootstrapping bestimmt werden. Da ein größeres Signifikanzniveau eine höhere Güte des Tests bedeutet, kann das Signifikanzniveau auch als Glättungsparameter angesehen werden. Ein kleines α führt dabei zu glatten Signalextraktionen.

Die rechte Fensterbreite r ist von großer Bedeutung für die SCARM-Signalextraktion, da sie den „Cutoff“ definiert, der eine Gruppe von Ausreißern von einer Signalveränderung unterscheidet. Bezeichne mit r^* den Umfang der größten zu erwartenden Ausreißergruppe. Dann ist r aufgrund des Bruchpunkts des RM so zu wählen, dass $\lfloor r/2 \rfloor > r^*$. Die RM-Schätzung der Regressionsgeraden kann jedoch bereits bei einem Ausreißeranteil von $1/3$ beträchtlich verzerrt sein. So ist in Abbildung 2.1 auf Seite 11 zu erkennen, dass die Zeitreihen der RM-Signalschätzungen (rot und blau) den Level Shift in den Daten jeweils mit einer Verzögerung von etwa $n/3$ Zeitpunkten nachzeichnen. Aus diesem Grund ist es empfehlenswert, r größer als $3r^*$ zu wählen, so dass der RM im rechten Zeitfenster einer Gruppe von r^* Ausreißern widerstehen kann. Folglich sollte auch die Breite des linken Zeitfensters ℓ_t zu keinem Zeitpunkt t kleiner als $3r^*$ sein, d.h. es sollte $\ell_{\min} = r > 3r^*$ gelten, so dass $\ell_t > 3r^*$ für alle t .

Die rechte Fensterbreite r bestimmt nicht nur die maximale Größe einer Ausreißergruppe, welcher der RM im rechten Zeitfenster widerstehen kann, sondern auch die Zeit, die zur Erkennung einer Signalveränderung benötigt wird. Diese Verzögerung beträgt in etwa $r/3$, vgl. Abbildung 2.1 auf Seite 11 sowie Kapitel 4.2.2. Falls der Test zu einem Zeitpunkt t eine Signalveränderung entdeckt, lässt sich daher vermuten, dass diese etwa zum Zeitpunkt $t - r/3$ vorliegt. Es ist aus diesem Grund zu empfehlen die minimale Fensterbreite $n_{\min} \approx r/3$ zu wählen. Die Wahl der maximalen Fensterbreite n_{\max} ist in der Regel weniger kritisch und sollte mit Rücksicht auf die Rechenzeit und die Annahme der lokalen Linearität gewählt werden, also nicht zu groß sein.

3.3.2 Recheneffiziente Update-Algorithmen des SCARM

Der SCARM führt zu jedem Zeitpunkt eine RM-Regression in einem Zeitfenster der Länge n_t durch. Gilt zu einem Zeitpunkt t , dass $n_t \geq \ell_{\min} + r$, werden im Zuge des SCARM-Tests zusätzlich noch eine RM-Regression im linken und im rechten Teilfenster durchgeführt. Würden die RM-Regressionsschätzer in jeder Iteration von Grund auf neu bestimmt, wäre dieser Ansatz mit großem Rechenaufwand verbunden. Die erforderlichen Rechenoperatio-

nen lassen sich jedoch in Anlehnung an den von Bernholt und Fried (2003) entwickelten Update-Algorithmus des RM effizient gestalten.

Die erste Signalschätzung erfolgt zum Zeitpunkt $t = n_t = n_{\min}$. Dabei wird wegen $n_{\min} < \ell_{\min} + r$ nicht auf Linearität des Signals getestet. Für die Stichprobe $\mathbf{y}_{t,n_t} = (y_{t-n_t+1}, \dots, y_t)$ wird zunächst die Matrix B_{t,n_t} der paarweisen Differenzenquotienten

$$\hat{\beta}_{ij} := \frac{y_{t-n_t+i} - y_{t-n_t+j}}{t - n_t + i - (t - n_t + j)} = \frac{y_{t-n_t+i} - y_{t-n_t+j}}{i - j}$$

berechnet, wobei $\hat{\beta}_{ii} = \text{NA}$ und $\hat{\beta}_{ij} = \hat{\beta}_{ji}$:

$$B_{t,n_t} = \begin{pmatrix} \text{NA} & \hat{\beta}_{1,2} & \dots & \hat{\beta}_{1,n_t-1} & \hat{\beta}_{1,n_t} \\ \hat{\beta}_{2,1} & \text{NA} & \dots & \hat{\beta}_{2,n_t-1} & \hat{\beta}_{2,n_t} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \hat{\beta}_{n_t-1,1} & \hat{\beta}_{n_t-1,2} & \dots & \text{NA} & \hat{\beta}_{n_t-1,n_t} \\ \hat{\beta}_{n_t,1} & \hat{\beta}_{n_t,2} & \dots & \hat{\beta}_{n_t,n_t-1} & \text{NA} \end{pmatrix} \sim (n_t \times n_t).$$

Die RM-Steigung im gesamten Zeitfenster der Länge n_t lässt sich dann als Median der zeilenweisen Mediane berechnen:

$$\hat{\beta}_{t,n_t} = \text{med}_{i=1,\dots,n_t} \left\{ \text{med}_{j \neq i} \left\{ \hat{\beta}_{ij} \right\} \right\}.$$

Mit

$$\hat{\mu}_{t,n_t} = \text{med}_{i=1,\dots,n_t} \left\{ y_{t-n_t+i} - \hat{\beta}_{t,n_t} \cdot (i - n_t) \right\}$$

ergibt sich dann das Niveau der RM-Geraden zum Zeitpunkt t , d.h. die SCARM-Signalschätzung. Im Zuge des Update-Vorgangs wird die Beobachtung y_{t+1} in das Zeitfenster aufgenommen, so dass die Fensterbreite nunmehr $n_t + 1$ beträgt und die RM-Regression im vergrößerten Zeitfenster durchgeführt wird. Dazu müssen nicht alle $\hat{\beta}_{ij}$ mit $i, j = 1, \dots, n_t + 1$ und $i \neq j$ neu berechnet werden, sondern lediglich

$$\hat{\beta}_{n_t+1,1}, \hat{\beta}_{n_t+1,2}, \dots, \hat{\beta}_{n_t+1,n_t}, \quad (3.16)$$

da die übrigen $\hat{\beta}_{ij}$ bereits zum vorherigen Zeitpunkt berechnet worden sind. Die Matrix B_{t+1,n_t+1} ist dann gegeben durch

$$B_{t+1,n_t+1} = \begin{pmatrix} \text{NA} & \hat{\beta}_{1,2} & \dots & \hat{\beta}_{1,n_t} & \hat{\beta}_{1,n_t+1} \\ \hat{\beta}_{2,1} & \text{NA} & \dots & \hat{\beta}_{2,n_t} & \hat{\beta}_{2,n_t+1} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \hat{\beta}_{n_t,1} & \hat{\beta}_{n_t,2} & \dots & \text{NA} & \hat{\beta}_{n_t,n_t+1} \\ \hat{\beta}_{n_t+1,1} & \hat{\beta}_{n_t+1,2} & \dots & \hat{\beta}_{n_t+1,n_t} & \text{NA} \end{pmatrix} \sim (n_t + 1) \times (n_t + 1).$$

Die Teilmatrix der ersten n_t Zeilen und Spalten entspricht also der Matrix B_{t,n_t} des vorherigen Zeitpunkts. Der Update-Vorgang wird beendet, indem $B_{t,n_t} \leftarrow B_{t+1,n_t+1}$ gesetzt wird. Die beschriebenen Vorgänge werden so lange wiederholt, bis $n_t = \ell_{\min} + r$, was erstmalig zum Zeitpunkt $t = \ell_{\min} + r$ der Fall ist. Zu diesem Zeitpunkt hat die Matrix B_{t,n_t} die Dimension $(\ell_{\min} + r) \times (\ell_{\min} + r)$. Für den SCARM-Tests müssen dann die RM-Steigungen $\hat{\beta}_t^{\text{left}} := \hat{\beta}(\mathbf{y}_{t-r,\ell_t})$ und $\hat{\beta}_t^{\text{right}} := \hat{\beta}(\mathbf{y}_{t,r})$ anhand der linken und rechten Teilstichproben $\mathbf{y}_{t-r,\ell_t} = (y_{t-n_t+1}, \dots, y_{t-r})$ und $\mathbf{y}_{t,r} = (y_{t-r+1}, \dots, y_t)$ geschätzt werden, wobei $\ell_t = n_t - r$. Dabei wird die linke RM-Steigung $\hat{\beta}_t^{\text{left}}$ als Median der zeilenweisen Mediane der Teilmatrix der $\ell_t = \ell_{\min}$ ersten Zeilen und Spalten von B_{t,n_t} berechnet. Die rechte RM-Steigung $\hat{\beta}_t^{\text{right}}$ ergibt sich analog anhand der Teilmatrix der r letzten Zeilen und Spalten von B_{t,n_t} . Lehnt der Test die Nullhypothese nicht ab, wird die RM-Signalschätzung auf Basis der gesamten Matrix B_{t,n_t} durchgeführt. Wird die Nullhypothese jedoch abgelehnt, wird $n_t \leftarrow n_{\min}$ gesetzt, d.h. die RM-Signalschätzung erfolgt anhand der Teilmatrix der $n_t = n_{\min}$ letzten Zeilen und Spalten. Die übrigen $\hat{\beta}_{ij}$ werden verworfen. Anschließend werden wieder die $\hat{\beta}_{n_t+1,j}$ aus (3.16) ermittelt, so dass der Update-Vorgang vollzogen werden kann.

Für den Fall, dass beim Update-Vorgang durch das Hinzufügen der Beobachtung y_{t+1} die Fensterbreite $n_t + 1 > n_{\max}$ ist, wird die erste bzw. älteste Beobachtung aus der Stichprobe entfernt. Dadurch kann n_t nicht über den Maximalwert n_{\max} hinauswachsen. In diesem Fall ist die Matrix B_{t+1,n_t+1} gegeben durch

$$B_{t+1,n_t+1} = \begin{pmatrix} \text{NA} & \hat{\beta}_{2,3} & \dots & \hat{\beta}_{2,n_t} & \hat{\beta}_{2,n_t+1} \\ \hat{\beta}_{3,2} & \text{NA} & & \dots & \dots \\ \vdots & & \ddots & & \vdots \\ \hat{\beta}_{n_t,2} & \hat{\beta}_{n_t,3} & \dots & \text{NA} & \hat{\beta}_{n_t,n_t+1} \\ \hat{\beta}_{n_t+1,2} & \hat{\beta}_{n_t+1,3} & \dots & \hat{\beta}_{n_t+1,n_t} & \text{NA} \end{pmatrix} \sim (n_t \times n_t).$$

Falls also $n_t + 1 \leq n_{\max}$, ergibt sich B_{t+1,n_t+1} aus B_{t,n_t} durch Hinzufügen einer Zeile und Spalte am rechten und unteren Rand. Falls jedoch $n_t + 1 > n_{\max}$, wird gleichzeitig die obere Zeile und linke Spalte entfernt.

Die hier beschriebenen Operationen beschleunigen den SCARM-Algorithmus, da zu jedem Zeitpunkt lediglich die Steigungen in (3.16) neu bestimmt werden müssen; die zum vorangegangenen Zeitpunkt errechneten Steigungen $\hat{\beta}_{ij}$ werden wieder verwendet. Für die Berechnung der Q_δ^{adj} -Skalenschätzung (3.12) findet dasselbe Prinzip Anwendung: Zu jedem Zeitpunkt muss lediglich die Höhe des Dreiecks bestimmt werden, das sich aus der neuen Beobachtung y_{t+1} und den beiden vorangegangenen Werten y_{t-1} und y_t ergibt.

3.3.3 Behandlung fehlender Werte

In Datenströmen können aus verschiedenen Gründen, häufig technisch bedingt, fehlende Werte auftreten. Dabei sind vereinzelt fehlende Werte für das Verfahren bzw. für die Programmierung unproblematisch, da eine RM- und Q_δ^{adj} -Schätzung auch bei vereinzelt fehlenden Werten zu berechnen ist. Treten fehlende Werte jedoch gehäuft auf oder fehlen längere Abschnitte an Beobachtungen, kann dies zu einem Programmfehler und somit zu einem Abbruch der Signalextraktion führen. Daher sind Regeln erforderlich, die einen Output in jeder Datensituation und zu jedem Zeitpunkt garantieren. Im Folgenden sei ein fehlender Messwert mit $y_t = \text{NA}$ bezeichnet und ein nicht-fehlender bzw. gemessener Wert mit $y_t \neq \text{NA}$. Ein $y_t = \text{NA}$ wird jedoch weiterhin als Beobachtung aufgefasst.

Regel 1: Gilt zu einem Zeitpunkt t mit Fensterbreite n_t , dass

$$\# \{i \in \{1, \dots, n_{\min}\} : y_{t-n_{\min}+i} = \text{NA}\} = n_{\min}$$

und/oder $\# \{i \in \{1, \dots, m\} : y_{t-m+i} \neq \text{NA}, m = \min\{n_t, r\}\} < n_{\min}$,

so wird für diesen Zeitpunkt der SCARM-Algorithmus nicht ausgeführt. Stattdessen wird $\mu_t = \text{NA}$ ausgegeben und für den nachfolgenden Zeitpunkt die Fensterbreite auf $\min\{n_t + 1, \ell_{\min} + r\}$ gesetzt.

Regel 2: Gilt zu einem Zeitpunkt t mit Fensterbreite $n_t = \ell_t + r \geq \ell_{\min} + r$, dass

$$\# \{i \in \{1, \dots, \ell_t\} : y_{t-n_t+i} \neq \text{NA}\} < \ell_{\min}/2$$

und/oder $\# \{i \in \{1, \dots, r\} : y_{t-r+i} \neq \text{NA}\} < r/2$,

so wird der SCARM-Test nicht durchgeführt. Stattdessen wird das Signal anhand der rechten Fensterstichprobe $\{y_{t-r+1}, \dots, y_t\}$ geschätzt und für den nachfolgenden Zeitpunkt die Fensterbreite auf $\ell_{\min} + r$ gesetzt.

Die erste Regel stellt sicher, dass die Signalextraktion zu jedem Zeitpunkt anhand von mindestens n_{\min} Messwerten erfolgt. Außerdem wird das Signal nicht geschätzt, wenn die aktuellsten n_{\min} Werte fehlen. Die zweite Regel garantiert, dass für den SCARM-Test in beiden Teilfenstern genügend Messwerte gegeben sind. Die zweite Regel kann nur dann in Kraft treten, wenn Regel 1 nicht in Kraft tritt. Falls eine der beiden Regeln in Kraft tritt, wird die Fensterbreite für den nachfolgenden Zeitpunkt auf den Wert $\ell_{\min} + r$ verkleinert – bzw. auf $n_t + 1$ gesetzt, sofern $n_t < \ell_{\min} + r$. Somit kann der Test sofort wieder durchgeführt werden, sobald das Zeitfenster der Länge $\ell_{\min} + r$ genügend Beobachtungen enthält.

Die beiden Regeln stellen sicher, dass die Signalschätzung und der Test anhand genügend (aktueller) Messwerte durchgeführt werden. Allerdings ist es trotz dieser Regeln möglich, dass die Q_δ^{adj} -Skalenschätzung der Standardabweichung der Fehler σ_t aufgrund zu vieler fehlender Werte nicht erfolgen kann, was zu einem Programmabbruch führt. Gilt für eine Beobachtung $y_t = \text{NA}$, so sind die Dreieckshöhen h_{t-1}, h_t, h_{t+1} nicht zu berechnen, vgl. (3.13). Würde in einer Zeitreihe $(y_t)_{t \in T}$ exakt jeder dritte Messwert fehlen, wäre $h_t = \text{NA}$ für alle $t \in T$. Das heißt, selbst wenn 2/3 der Messwerte in einem Zeitfenster gegeben sind, ist möglicherweise die Q_δ^{adj} -Skalenschätzung von σ_t nicht zu berechnen. Zusätzlich zu den obigen zwei Regeln ließe sich fordern, dass in jedem Zeitfenster mehr als 2/3 der Messwerte vorliegen. Somit wäre eine Q_δ^{adj} -Skalenschätzung von σ_t gesichert. Bei diesem Ansatz kann jedoch die Anzahl der für die Schätzung betrachteten Dreieckshöhen sehr gering sein, da ein Messwert nur dann in die Skalenschätzung eingeht, wenn auch die beiden benachbarten Messwerte gegeben sind. Daher wird stattdessen eine leicht modifizierte Version des Q_δ^{adj} vorgeschlagen. Im Folgenden sei der Zeitindex vernachlässigt. Sei $\mathbf{y} = (y_1, \dots, y_n)$ ein Stichprobenvektor und $I = \{i = 1, \dots, n : y_i \neq \text{NA}\}$ mit $|I| = m$ die Menge der Designpunkte, an denen Messwerte gegeben sind. Mit $\mathbf{z} := (y_i)_{i \in I} = (z_j)_{j=1, \dots, m}$ und $\mathbf{x} := (i)_{i \in I} = (x_j)_{j=1, \dots, m}$ ist der modifizierte Q_δ^{adj} -Skalenschätzer dann gegeben durch

$$Q_\delta^{\text{adj}*}(\mathbf{y}) = c_m \cdot h_{\lfloor \delta(m-2) \rfloor}, \quad h_j = \left| z_j - z_{j-1} - (x_j - x_{j-1}) \cdot \frac{z_{j+1} - z_{j-1}}{x_{j+1} - x_{j-1}} \right|, \quad (3.17)$$

$j = 2, \dots, m-1$. Dabei ist h_j die vertikale Höhe des Dreiecks, das sich aus den drei aufeinander folgenden Beobachtungen z_{j-1}, z_j, z_{j+1} mit Designpunkten x_{j-1}, x_j, x_{j+1} ergibt. Der Korrekturfaktor c_m entspricht dem Korrekturfaktor c_n aus (3.12) und ist nach der Anzahl an Messwerten m zu wählen. Falls $m = n$, d.h. falls keine Messwerte fehlen, entspricht der $Q_\delta^{\text{adj}*}$ dem Q_δ^{adj} aus (3.12). Falls jedoch $m < n$, wird beim $Q_\delta^{\text{adj}*}$ im Gegensatz zum Q_δ^{adj} jeder Messwert zur Skalenschätzung verwendet. Zudem garantiert der $Q_\delta^{\text{adj}*}$ bereits für $m \geq 3$ eine Skalenschätzung. Wie der Q_δ^{adj} ist der $Q_\delta^{\text{adj}*}$ stabil gegenüber plötzlichen Signalveränderungen und invariant gegenüber linearen Trends:

Satz 3.3.

Seien $\mathbf{y} = (y_1, \dots, y_n)$ und $\mathbf{y}^* = (y_1^*, \dots, y_n^*)$ mit $y_i^* = y_i + i \cdot v$, $v \in \mathbb{R}$. Dann gilt

$$Q_\delta^{\text{adj}*}(\mathbf{y}) = Q_\delta^{\text{adj}*}(\mathbf{y}^*).$$

Beweis.

Seien I , \mathbf{x} und \mathbf{z} wie oben und $\mathbf{z}^* = (y_i^*)_{i \in I} = (z_j^*)_{j=1, \dots, m}$, so dass $z_j^* = z_j + x_j \cdot v$. Dann ist $Q_\delta^{\text{adj}^*}(\mathbf{y}) = c_m \cdot h_{\lfloor \delta(m-2) \rfloor}$ wie in (3.17) und $Q_\delta^{\text{adj}^*}(\mathbf{y}^*) = c_m \cdot h_{\lfloor \delta(m-2) \rfloor}^*$ mit

$$h_j^* = \left| z_j^* - z_{j-1}^* - (x_j - x_{j-1}) \cdot \frac{z_{j+1}^* - z_{j-1}^*}{x_{j+1} - x_{j-1}} \right|, \quad j = 2, \dots, m-1.$$

Durch einfache Rechenoperationen lässt sich mit $z_j^* = z_j + x_j \cdot v$ zeigen, dass $h_j = h_j^*$ für alle $j = 2, \dots, m-1$. \square

Neben der Trendinvarianz des $Q_\delta^{\text{adj}^*}$ lässt sich leicht zeigen, dass dieser bei fehlenden Werten einen nicht-kleineren Bruchpunkt als der Q_δ^{adj} besitzt. Sei dazu k die Anzahl an Tripeln y_{i-1}, y_i, y_{i+1} in einer Stichprobe \mathbf{y} mit $y_{i-1} \neq \text{NA}, y_i \neq \text{NA}, y_{i+1} \neq \text{NA}, i = 2, \dots, n-1$, d.h. k ist die Anzahl an Dreiecken, deren vertikale Höhen zur Q_δ^{adj} -Schätzung genutzt werden. Ist m die Anzahl an tatsächlich gegebenen Messwerten, so gilt offensichtlich $k \leq m-2$. Die Q_δ^{adj} -Schätzung erfolgt also anhand von $k+2 \leq m$ Beobachtungen, wohingegen für die $Q_\delta^{\text{adj}^*}$ -Schätzung alle m Beobachtungen genutzt werden. Folglich ist

$$\begin{aligned} \text{fsbp}^+(Q_\delta^{\text{adj}}, \mathbf{y}_n) &= \frac{1}{n} \cdot \left\lceil \frac{(k+2) - 1 - \lfloor \delta k \rfloor}{3} \right\rceil \\ &\leq \frac{1}{n} \cdot \left\lceil \frac{m-1 - \lfloor \delta(m-2) \rfloor}{3} \right\rceil = \text{fsbp}^+(Q_\delta^{\text{adj}^*}, \mathbf{y}_n) \\ \text{fsbp}^-(Q_\delta^{\text{adj}}, \mathbf{y}_n) &= \frac{1}{n} \cdot \lfloor \delta k \rfloor \leq \frac{1}{n} \cdot \lfloor \delta(m-2) \rfloor = \text{fsbp}^-(Q_\delta^{\text{adj}^*}, \mathbf{y}_n), \end{aligned}$$

vgl. (3.21).

Des Weiteren ist der $Q_\delta^{\text{adj}^*}$ bei fehlenden Werten ein effizienterer Schätzer für σ_t als der Q_δ^{adj} , wie Simulationen in Anhang B.2 anzeigen. Dies gilt sowohl für normalverteilte Fehler als auch für Fehler, die einer kontaminierten Normalverteilung folgen. Der $Q_\delta^{\text{adj}^*}$ ist dem Q_δ^{adj} also bzgl. der Robustheit und Effizienz überlegen und garantiert in Kombination mit Regel 1 auf S. 39, dass kein Programmabbruch aufgrund von fehlenden Messwerten geschieht, sofern $n_{\min} \geq 3$.

3.4 Die Verteilung der SCARM-Teststatistik

Im Folgenden wird die Verteilung der SCARM-Teststatistik T_t aus (3.6) unter H_0 untersucht. Das Ziel dieser Untersuchung ist es, für jedes beliebige Signifikanzniveau $\alpha \in (0, 1)$ und jede beliebige Kombination von Fensterbreiten (ℓ_t, r) einen passenden kritischen Wert

zu erhalten. Die Bestimmung der asymptotischen Verteilung von T_t unter H_0 wäre sicherlich aufschlussreich. Da es sich in der Anwendung bei den Fensterbreiten ℓ_t und insbesondere r in der Regel um relativ kleine Stichprobenumfänge handelt, ist die Kenntnis der asymptotischen Verteilung von T_t unter H_0 von eher geringem Nutzen. Daher wird hier die Verteilung der SCARM Teststatistik unter H_0 für endliche Stichproben mit unterschiedlich verteilten Fehlern mittels Simulationen untersucht. Für symmetrisch um null verteilte Fehlerprozesse weist T_t unter H_0 eine um null symmetrische Verteilung auf, vgl. Satz 3.2. Da die Variabilität der Steigungsdifferenz nicht bekannt ist sondern geschätzt wird, besitzt die Verteilung von T_t vermutlich schwere Ränder. Im Folgenden sei einfach $\ell = \ell_t$ und $n = n_t$. Im Rahmen der Simulationsstudie werden zunächst Stichproben $\mathbf{y}_{t,n}$ betrachtet, die entsprechend dem H_0 -Modell (3.2) generiert sind:

$$\begin{aligned} Y_{t-n+i} &= \mu_t + \beta_t \cdot (i - n) + \xi_{t-n+i}, \\ \xi_{t-n+i} &= \sigma_t \varepsilon_{t-n+i} + \eta_{t-n+i}, \quad i = 1, \dots, n, \end{aligned}$$

wobei aufgrund der Regressionsäquivarianz des RM o.B.d.A. $\mu_t = \beta_t = 0$ gewählt wird. Es werden unterschiedliche Arten von Fehlerprozessen ξ_{t-n+i} betrachtet. Dabei werden für jede Fehlerart $N = 10000$ Stichproben der Länge $n = \ell + r$ mit $r \in \{5, 10, \dots, 100\}$ und $\ell \in \{r, r + 5, \dots, 100\}$ generiert. Für jede Kombination von ℓ und r ergeben sich demnach $N = 10000$ Realisationen

$$T_t(\ell, r) := \frac{\hat{\beta}(\mathbf{y}_{t^*,\ell}) - \hat{\beta}(\mathbf{y}_{t,r})}{\sqrt{Q_{0.5}^{\text{adj}}(\mathbf{y}_{t,n})^2 \cdot (v_\ell + v_r)}}$$

mit $\mathbf{y}_{t^*,\ell} = (y_{t-n+1}, \dots, y_{t-r}) \in \mathbb{R}^\ell$ und $\mathbf{y}_{t,r} = (y_{t-r+1}, \dots, y_t) \in \mathbb{R}^r$, deren empirische Verteilung Aufschluss gibt über die wahre Verteilung der SCARM-Teststatistik unter H_0 für unterschiedliche Arten von Fehlern.

3.4.1 Normalverteilte Fehlerprozesse

Zunächst werden normalverteilte Fehler betrachtet, d.h. $\eta_{t-n+i} = 0$ und $\sigma_t \varepsilon_{t-n+i} \sim N(0, \sigma_t^2)$, wobei aufgrund der Skalenäquivarianz der RM-Steigung und des Q_δ^{adj} -Schätzers o.B.d.A. $\sigma_t = 1$ gesetzt wird. Die Abbildungen 3.4 und 3.5 zeigen QQ-Plots und Histogramme basierend auf den N Realisationen der Teststatistik $T_t(\ell, r)$ für $(\ell, r) = (10, 10), (50, 50), (100, 100), (50, 10), (100, 10)$ und $(100, 50)$. Die QQ-Plots vergleichen die empirischen Quantile der Realisationen $T_t(\ell, r)$ mit denen einer Standardnormalverteilung. Des Weiteren

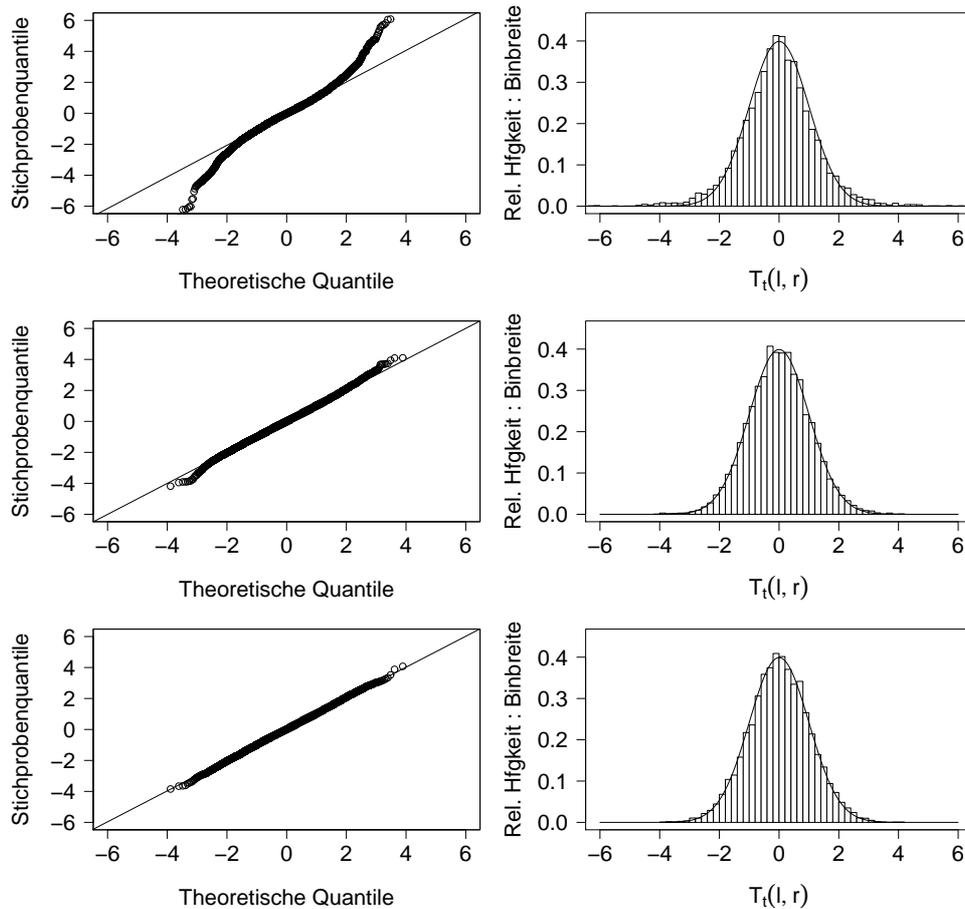


Abbildung 3.4: QQ-Plots und Histogramme zum Vergleich der Standardnormalverteilung mit der empirischen Verteilung der SCARM-Teststatistik für $(\ell, r) = (10, 10)$ (obere Zeile), $(\ell, r) = (50, 50)$ (mittlere Zeile) und $(\ell, r) = (100, 100)$ (untere Zeile).

ist zu den Histogrammen die Dichte einer Standardnormalverteilung eingezeichnet. Erwartungsgemäß deuten die Histogramme auf eine um null symmetrische Verteilung der SCARM-Teststatistik hin. Wenn ℓ und r nicht zu klein sind, folgt die SCARM-Teststatistik annähernd einer Standardnormalverteilung. Für kleine Fensterbreiten ℓ und r weisen die Histogramme jedoch eine leptokurtische Form mit schweren Rändern auf, und in den QQ-Plots weichen die empirischen Quantile in den Randbereichen deutlich von den Quantilen der Standardnormalverteilung ab, vgl. Abbildung 3.4 und 3.5, oben. Für kleine Fensterbreiten ℓ und r scheint es daher angemessener, die Verteilung der SCARM-Teststatistik über eine t -Verteilung zu approximieren, so dass deren Quantile als kritische Werte genutzt werden. Die Freiheitsgrade f der t -Verteilung sind dabei in Abhängigkeit von den Fenster-

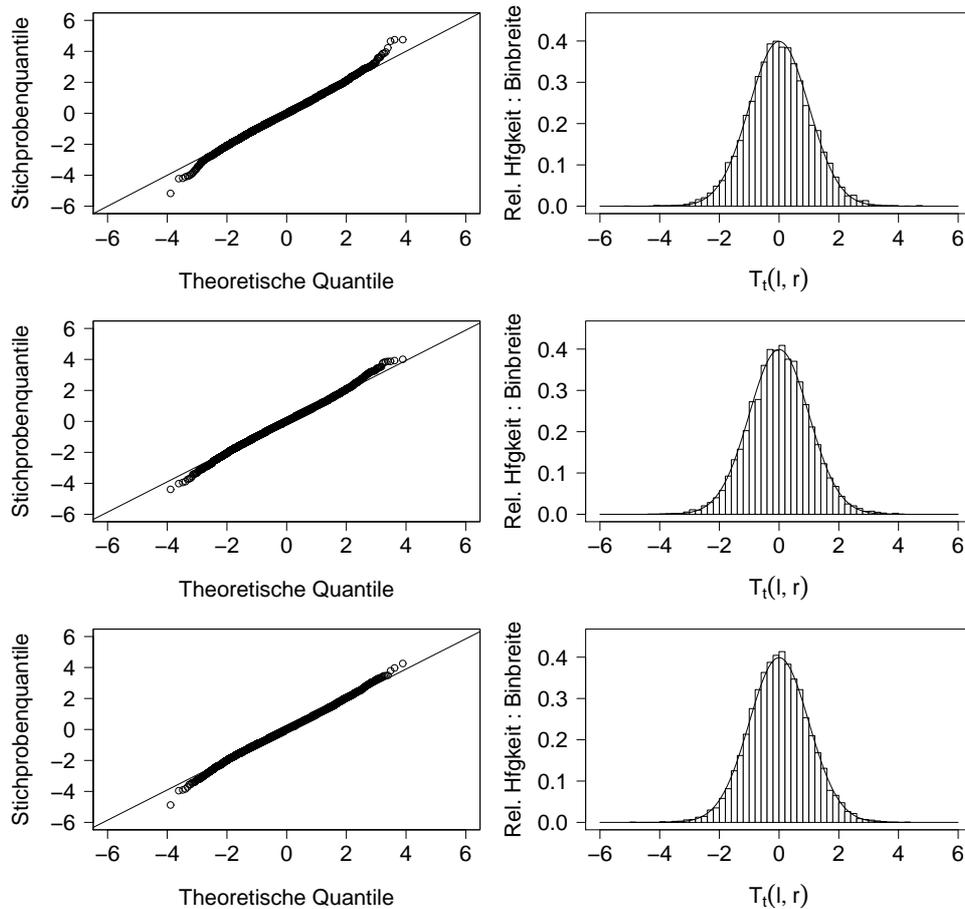


Abbildung 3.5: QQ-Plots und Histogramme zum Vergleich der Standardnormalverteilung mit der empirischen Verteilung der SCARM-Teststatistik für $(\ell, r) = (50, 10)$ (obere Zeile), $(\ell, r) = (100, 10)$ (mittlere Zeile) und $(\ell, r) = (100, 50)$ (untere Zeile).

breiten ℓ und r zu wählen, d.h. $f = f(\ell, r)$. Dabei lässt sich ein wachsendes Verhalten von f in ℓ und r vermuten – je kleiner die Fensterbreite, desto kleiner f , d.h. desto schwerer die Ränder der Verteilung von T_t .

Um für jede Kombination (ℓ, r) ein passendes f zu finden, werden die Stichprobenquantile q_α , $\alpha \in I_\alpha = \{1/N, 2/N, \dots, (N-1)/N\}$, mit den entsprechenden α -Quantilen einer t -Verteilung mit Freiheitsgraden $f \in I_f = \{1, 1.1, \dots, 100\}$ verglichen. Dabei wird jedem Tupel (ℓ, r) dasjenige f zugeordnet, für welches die mittlere quadratische Abweichung der N theoretischen Quantile $t_{\alpha, f}$ von den N Stichprobenquantilen q_α minimal ist:

$$f(\ell, r) = \arg \min_{f \in I_f} \sum_{\alpha \in I_\alpha} (t_{\alpha, f} - q_\alpha)^2.$$

		Rechte Fensterbreite r									
		5	10	15	20	25	30	35	40	45	50
Linke Fensterbreite ℓ	5	3.5	–	–	–	–	–	–	–	–	–
	10	4.8	6.1	–	–	–	–	–	–	–	–
	15	7.2	7.5	11.1	–	–	–	–	–	–	–
	20	8.1	9.4	12.1	12.3	–	–	–	–	–	–
	25	10.5	11.5	13.5	14.4	19.4	–	–	–	–	–
	30	10.8	12.1	14.8	19.4	19.4	21.4	–	–	–	–
	35	11.6	14.6	14.8	20.2	22.4	22.4	22.4	–	–	–
	40	14.4	14.6	14.8	21.7	22.4	22.4	22.4	22.6	–	–
	45	15.8	16.2	21.4	21.7	22.4	22.4	22.4	23.3	24.8	–
	50	15.8	20.0	21.4	21.7	22.4	22.4	22.4	23.3	24.8	25.3

Tabelle 3.1: Monotonisierte Freiheitsgrade $f(\ell, r)$ für die Approximation der Verteilung der SCARM-Teststatistik über eine t -Verteilung

Somit ist für alle Werte von α eine gewisse Nähe der theoretischen Quantile $t_{\alpha, f}$ zu den Stichprobenquantilen q_α gewährleistet. Die auf diese Weise bestimmten Freiheitsgrade für $r \in \{5, 10, \dots, 100\}$ und $\ell \in \{r, r+5, \dots, 100\}$ sind erwartungsgemäß tendenziell wachsend in ℓ und r , allerdings nicht monoton, vgl. Tabelle C.1 in Anhang C. Daher werden die Freiheitsgrade über die folgende Vorschrift monotonisiert:

$$f(\ell, r) = \min_{\ell' \geq \ell} \min_{r' \geq r} f(\ell', r'), \quad (3.18)$$

wobei $r' \in \{5, 10, \dots, 50\}$ und $\ell' \in \{r', r'+5, \dots, 50\}$. Die Regel (3.18) garantiert, dass das Signifikanzniveau eingehalten wird. Dabei werden einige $f(\ell, r)$ -Werte deutlich verkleinert, was größere kritische Werte zur Folge hat, wodurch der Test konservativ wird. Darüber hinaus sind durch die Regel (3.18) auch die Freiheitsgrade $f(\ell, r)$ für $\ell, r \in \{5, 6, \dots, 100\} \setminus \{5, 10, \dots, 100\}$ gegeben.

Tabelle 3.1 zeigt die monotonisierten Freiheitsgrade $f(\ell, r)$ für $r \in \{5, 10, \dots, 50\}$ und $\ell \in \{r, r+5, \dots, 50\}$. Offensichtlich ergeben sich bereits für moderate Fensterbreiten ℓ und r Freiheitsgrade $f(\ell, r) > 20$. Für große Fensterbreiten ließe sich die Verteilung der SCARM-Teststatistik also auch durch eine Standardnormalverteilung approximieren. Daher werden für $\ell > 100$ oder $r > 100$ die Quantile der Standardnormalverteilung als kritische Werte des SCARM-Tests genutzt.

3.4.2 Weitere Fehlerverteilungen

Die oben beschriebene Simulation mit normalverteilten Fehlern wird analog für vier weitere Fehlerarten durchgeführt. Dabei ist vor allem von Interesse, wie sehr sich die empirischen Quantile der SCARM-Teststatistik T_t von den Quantilen der zur Approximation der wahren Verteilung von T_t genutzten t -Verteilung unterscheiden. Die vier Fehlerverteilungen sind:

Typ 1: Die t -Verteilung mit drei Freiheitsgraden.

Typ 2: Die Weibull-Verteilung mit Formparameter zwei und Skalenparameter eins.

Typ 3: Eine Mischverteilung: $\xi_{t-n+i} \sim 0.9 N(0, 1) + 0.1 N(\mu_t = 10, 1)$.

Typ 4: Eine Mischverteilung: $\xi_{t-n+i} \sim 0.9 N(0, 1) + 0.1 N(0, \sigma_t^2 = 100)$.

Bei dem Verteilungstyp 1 handelt es sich um eine Verteilung mit schweren Rändern und beim Typ 2 um eine schiefe Verteilung. Dabei sind beide Verteilungen durch den jeweiligen Erwartungswert und die jeweilige Varianz standardisiert. Bei dem Verteilungstyp 3 werden lediglich Ausreißer in positiver Richtung generiert, wohingegen beim vierten Typ auch Ausreißer in negativer Richtung auftreten.

Für jeden der vier Fehlertypen und für jede der sechs Fensterbreitenkombinationen $(\ell, r) = (10, 10), (50, 50), (100, 100), (100, 50), (50, 10)$ und $(100, 10)$ ergeben die Simulationen $N = 10000$ Realisationen der SCARM-Teststatistik. Erkenntnisse über das Verhalten des SCARM-Tests für die jeweilige Fehlerart und Fensterbreitenkombination liefert dann der Vergleich der Stichprobenquantile mit den Quantilen der entsprechenden $t_{f(\ell, r)}$ -Verteilung. Dazu sind in den Abbildungen 3.6 und 3.7 die QQ-Plots der empirischen und theoretischen Quantile für $(\ell, r) = (50, 50)$ und $(100, 10)$ dargestellt. Für die übrigen vier (ℓ, r) -Kombinationen sind die QQ-Plots in den Abbildungen C.3 – C.6 in Anhang C gegeben.

Für den Fehlertyp 2 zeigen die empirischen und theoretischen Quantile große Übereinstimmungen. Die Quantile der t -Verteilung sind also auch für diesen Fehlertyp als kritische Werte geeignet. Für die übrigen drei Fehlertypen ergeben sich je nach Kombination (ℓ, r) unterschiedliche Resultate. Für $\ell, r = (10, 10), (50, 50), (100, 100)$ und $(100, 50)$ sind die Quantile der t -Verteilung in den Rändern extremer als die Stichprobenquantile, d.h. der SCARM-Test wäre in diesen Fällen konservativ. Im Gegensatz dazu wäre der SCARM-Test für die (ℓ, r) -Kombinationen $(50, 10)$ und $(100, 10)$ anti-konservativ. Allerdings sei darauf hingewiesen, dass der SCARM-Test zur Fensterbreitenadaptation genutzt wird. Das Einhalten des Signifikanzniveaus ist in der Regel eher von geringer Bedeutung. Dieses lässt sich vielmehr als eine Art Glättungsparameter ansehen.

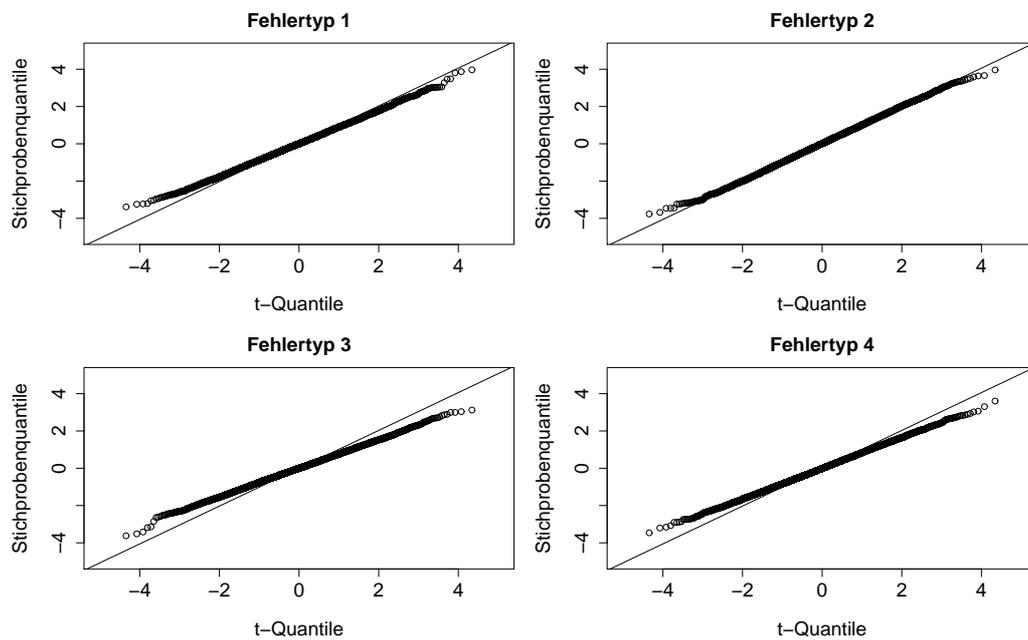


Abbildung 3.6: QQ-Plots der Quantile der $t_{f(\ell,r)}$ -Verteilung mit $(\ell, r) = (50, 50)$ gegen die Stichprobenquantile der SCARM-Teststatistik für die Fehlertypen 1-4.

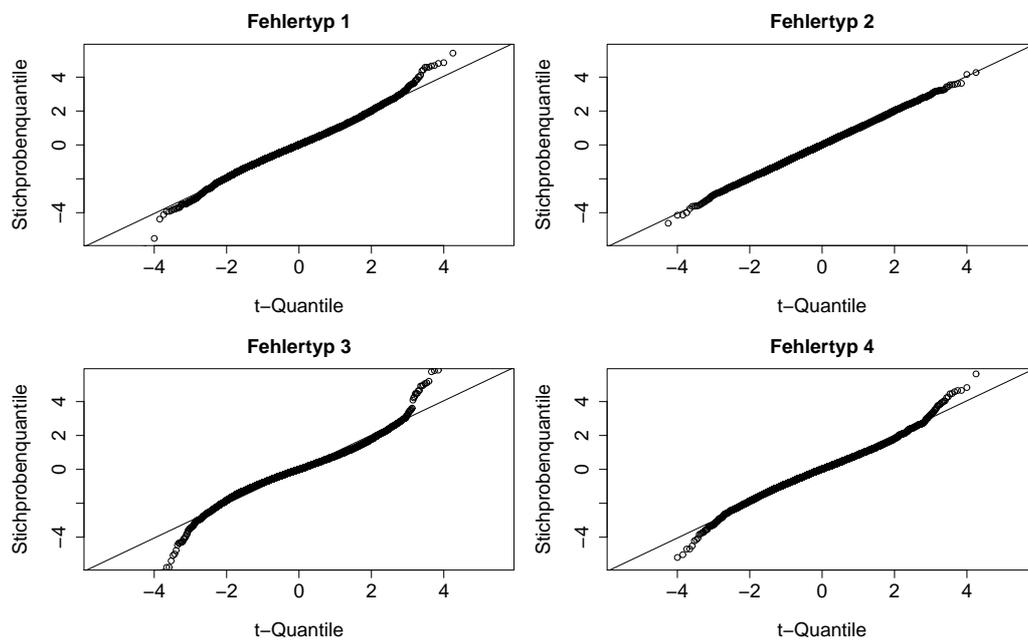


Abbildung 3.7: QQ-Plots der Quantile der $t_{f(\ell,r)}$ -Verteilung mit $(\ell, r) = (100, 10)$ gegen die Stichprobenquantile der SCARM-Teststatistik für die Fehlertypen 1-4.

3.5 Diskussion

Die für den SCARM-Test benötigte Schätzung der Varianz der RM-Steigungsdifferenzen (3.14) ist unter der Annahme normalverteilter Fehler entwickelt worden. Für andere Fehlerverteilungen ist dieser Schätzer im Allgemeinen verzerrt, vgl. Anhang B.1.2. Eine Verzerrung kann jedoch ausgeglichen werden, indem die Approximationen v_n (bzw. v_ℓ und v_r) der Varianz der RM-Steigung $V(n, 1)$ sowie die Korrekturkonstanten c_n des Q_δ^{adj} -Schätzers entsprechend der gegebenen Fehlerverteilung anhand von Simulationen neu bestimmt werden. Zudem ist der Schätzer in (3.14) verzerrt, wenn die Fehler nicht unabhängig sondern autokorreliert sind. Zum einen liefert der Q_δ^{adj} -Skalenschätzer bei Autokorrelationen verzerrte Schätzungen der Fehlervarianz, vgl. Schettlinger (2009, S. 113f.), da eine positive (negative) Autokorrelation kleinere (größere) Dreieckshöhen und somit negativ (positiv) verzerrte Schätzungen bewirkt. Zum anderen liefert das in Anhang B.1 erstellte Modell $v(n)$ für $V(n, 1)$ bei positiver (negativer) Autokorrelation zu geringe (zu hohe) Approximationen, d.h. die tatsächliche Varianz der RM-Steigungen $V(n, 1)$ ist größer (kleiner), vgl. Anhang B.1.3. Bei positiver Autokorrelation unterschätzt der SCARM demnach die Varianz der RM-Steigungsdifferenzen $\text{Var}(D_t)$. Folglich weist die SCARM-Teststatistik eine gesteigerte Variabilität auf, so dass vermehrt Fehler 1. Art auftreten. Der SCARM adaptiert daher tendenziell kleinere Fensterbreiten, was eine weniger glatte Signalextraktion bedeutet. Demgegenüber überschätzt der SCARM die Varianz der RM-Steigungsdifferenzen $\text{Var}(D_t)$ bei negativ autokorrelierten Fehlern. Es kommt daher vermehrt zu nicht erkannten Signalveränderungen, d.h. der SCARM-Test verliert an Güte. Der SCARM adaptiert folglich tendenziell zu große Fensterbreiten.

Um Verzerrungen des Schätzers $\widehat{\text{Var}}(d_t)$ in (3.14) aufgrund von Autokorrelationen entgegenzuwirken, werden die Konstanten v_n und c_n zusätzlich auf Basis von Daten ermittelt, die einem AR(1)-Prozess folgen. Dazu werden für alle $n \in \{5, \dots, 300\}$ und für alle $\varphi \in \{-0.9, -0.6, -0.3, 0, 0.3, 0.6, 0.9\}$ jeweils 10000 Stichproben (y_{t-n+1}, \dots, y_t) gemäß

$$\begin{aligned} Y_{t-n+i} &= \mu_t + \beta_t \cdot (i - n) + \varepsilon_{t-n+i}, \\ \varepsilon_{t-n+i} &= \xi_{t-n+i}^\varphi \sqrt{(1 - \varphi^2)}, \\ \xi_{t-n+i}^\varphi &= \varphi \xi_{t-n+i-1} + e_{t-n+i}, \quad e_{t-n+i} \sim N(0, 1), \end{aligned} \tag{3.19}$$

erzeugt, $i = 1, \dots, n$. Wegen $\varepsilon_{t-n+i} = \xi_{t-n+i}^\varphi \sqrt{(1 - \varphi^2)}$ gilt $\text{Var}(\varepsilon_{t-n+i}) = 1$. Aufgrund der Regressionsäquivarianz des RM ist o.B.d.A. $\mu_t = \beta_t = 0$ gewählt. Die c_n^φ und v_n^φ werden dann für alle φ und n bestimmt, indem der Q_δ^{adj} -Skalenschätzer mit $\delta = 0.5$ bzw. der RM-Steigungsschätzer auf die Stichproben (y_{t-n+1}, \dots, y_t) angewendet wird. Dann ist $1/c_n^\varphi$

gleich dem arithmetischen Mittel der 10000 Q_δ^{adj} -Schätzungen, und v_n^φ ist die empirische Varianz der 10000 RM-Steigungsschätzungen. Die so gewonnenen Konstanten c_n^φ und v_n^φ sind in der Abbildung C.2 in Anhang C zu sehen. Für $\varphi = 0$ entsprechen v_n^φ und c_n^φ den o.g. Konstanten v_n und c_n . Anhand der v_n^φ und c_n^φ werden analog zum Fall unabhängiger Fehler in Anhang B.1 Modelle erstellt, so dass die v_n^φ und c_n^φ für alle $n \geq 5$ entsprechend der gegebenen Autokorrelation gewählt werden können. Sind die Fehler beispielsweise schwach positiv (stark negativ) korreliert, wären die Konstanten v_n^φ und c_n^φ mit $\varphi = 0.3$ ($\varphi = -0.9$) zu empfehlen. Falls kein Vorwissen zu der Autokorrelationsstruktur besteht oder sich diese über die Zeit ändern kann, besteht die Möglichkeit der automatischen Wahl der Konstanten v_n^φ und c_n^φ in Abhängigkeit von der aktuellen Autokorrelation. Dazu wird zu jedem Zeitpunkt t die Autokorrelation der Fehler zum Lag eins anhand der letzten n_{\max} SCARM-Residuen $y_{t-n_{\max}+i} - \hat{\mu}_{t-n_{\max}+i}^{\text{SCARM}}$ mit $i = 1, \dots, n_{\max}$ geschätzt. Würde die Fehlerkorrelation anhand der Daten y_t geschätzt werden, ergäben sich in Trendphasen verzerrte Schätzungen. Bei diesem Ansatz ist die maximale Fensterbreite n_{\max} nicht zu klein zu wählen, so dass die Schätzung der Fehlerkorrelation, bezeichnet mit $\hat{\varphi}_t$, anhand von genügend Beobachtungen gewonnen wird. Die Korrelationsschätzung $\hat{\varphi}_t$ bestimmt dann die Wahl von φ und somit die Wahl der Konstanten v_n^φ und c_n^φ über die Regel

$$\varphi = \varphi(\hat{\varphi}_t) = \arg \min_{\phi \in \Psi} |\phi - \hat{\varphi}_t| \quad \text{mit} \quad \Psi = \{-0.9, -0.6, -0.3, 0, 0.3, 0.6, 0.9\}. \quad (3.20)$$

Zusätzlich zu der Annahme unabhängiger Fehler ist der Varianzschätzer $\widehat{\text{Var}}(d_t)$ in (3.14) unter der Voraussetzung entwickelt worden, dass der Ausreißerprozess $\eta_{t-n+i} = 0$ ist. Die Robustheit des Schätzers gegenüber Ausreißern ist jedoch durch die Robustheit des Q_δ^{adj} -Skalenschätzers gegeben. Für ein festes Design ist der Bruchpunkt von Q_δ^{adj} an einer Stichprobe $\mathbf{y}_n = (y_1, \dots, y_n)$ gegeben durch

$$\begin{aligned} \text{fsbp}(Q_\delta^{\text{adj}}, \mathbf{y}_n) &= \min \left\{ \text{fsbp}^+(Q_\delta^{\text{adj}}, \mathbf{y}_n), \text{fsbp}^-(Q_\delta^{\text{adj}}, \mathbf{y}_n) \right\}, \\ \text{fsbp}^+(Q_\delta^{\text{adj}}, \mathbf{y}_n) &= \frac{1}{n} \cdot \left\lceil \frac{n-1 - \lfloor \delta(n-2) \rfloor}{3} \right\rceil, \\ \text{fsbp}^-(Q_\delta^{\text{adj}}, \mathbf{y}_n) &= \frac{1}{n} \cdot \lfloor \delta(n-2) \rfloor \end{aligned} \quad (3.21)$$

(Gelper et al., 2009). Die Definition des Bruchpunkts für Skalenfunktionale ist in Anhang A zu finden. Für den von Gelper et al. empfohlen Wert $\delta = 0.5$ ergibt sich ein Explosions-Bruchpunkt von $\text{fsbp}^+(Q_\delta^{\text{adj}}, \mathbf{y}_n) \approx n/6$ und ein Implosions-Bruchpunkt von $\text{fsbp}^-(Q_\delta^{\text{adj}}, \mathbf{y}_n) \approx n/2$. Eine einfache Maßnahme zur Verhinderung einer Implosion von

$\widehat{\text{Var}}(d_t)$ – und einer daraus folgenden Ablehnung der Nullhypothese – bietet eine zuvor festgelegte untere Grenze $b > 0$ für die Q_δ^{adj} -Schätzung der Fehlerskala:

$$\widehat{\text{Var}}(d_t) = \left(\max \left\{ b, Q_\delta^{\text{adj}}(\mathbf{y}_{t,n}) \right\} \right)^2 \cdot (v_\ell + v_r). \quad (3.22)$$

Die Schranke b sollte unter Beachtung der tatsächlichen Fehlervariabilität gewählt werden.

Eine R-Funktion des SCARM, bezeichnet mit `scarm.filter`, ist enthalten im R-Paket `robfilter` (Fried et al., 2012), welches auf dem CRAN-Server zur Verfügung steht:

<http://cran.r-project.org/web/packages/robfilter/>.

Die Funktion `scarm.filter` ermöglicht die Wahl der auf Seite 34 aufgelisteten Inputargumente sowie der Schranke b für die Q_δ^{adj} -Schätzung der Fehlerskala in (3.22). Für die genannten Inputargumente gelten die folgenden Entsprechungen:

$$\begin{aligned} \ell_{\min} &\hat{=} \text{min.left.width}, & r &\hat{=} \text{right.width}, \\ n_{\min} &\hat{=} \text{min.width}, & n_{\max} &\hat{=} \text{max.width}, \\ \alpha &\hat{=} \text{sign.level}, & b &\hat{=} \text{bound.noise.sd}. \end{aligned}$$

Zudem kann mit dem Inputargument `rtr=TRUE/FALSE` die Restrict-to-Range-Regel (2.12) angewendet werden. Falls `rtr=TRUE`, werden die SCARM-Signalschätzungen auf das Stichprobenintervall der letzten n_{\min} Beobachtungen beschränkt. Weiterhin ermöglicht es das Inputargument `autocorrelations`, die Konstanten v_n^φ und c_n^φ für die Schätzung von $\text{Var}(D_t)$ gemäß (3.14) mit Rücksicht auf die vorliegende Autokorrelationsstruktur zu wählen, vgl. S. 48f. Dabei gelten die Entsprechungen

$$\begin{aligned} \text{autocorrelations} = \text{''high.negative''} &\hat{=} \varphi = -0.9, \\ \text{autocorrelations} = \text{''moderate.negative''} &\hat{=} \varphi = -0.6, \\ \text{autocorrelations} = \text{''small.negative''} &\hat{=} \varphi = -0.3, \\ \text{autocorrelations} = \text{''no''} &\hat{=} \varphi = 0, \\ \text{autocorrelations} = \text{''small.positive''} &\hat{=} \varphi = 0.3, \\ \text{autocorrelations} = \text{''moderate.positive''} &\hat{=} \varphi = 0.6, \\ \text{autocorrelations} = \text{''high.positive''} &\hat{=} \varphi = 0.9. \end{aligned}$$

Wird die Option `autocorrelations=''automatic''` gewählt, werden die Konstanten v_n^φ und c_n^φ zu jedem Zeitpunkt gemäß (3.20) in Abhängigkeit von der geschätzten Autokorrelation zum Lag eins bestimmt.

4 Eine Studie zum Vergleich des aoRM und SCARM

In diesem Abschnitt werden der aoRM- und SCARM-Filter hinsichtlich ihrer Eignung für die Online-Signalextraktion verglichen. Zunächst wird in 4.1 die Robustheit des aoRM- und SCARM-Tests gegenüber Ausreißern untersucht. Die Robustheitskriterien basieren auf den von Ylvisaker (1977) entwickelten *Test Resistances*. In 4.2 werden die Testprozeduren hinsichtlich der Erkennung von Signalveränderungen gegenübergestellt. Dabei interessieren zum einen die Entdeckungsraten der Tests, d.h. die relativen Häufigkeiten, mit der unterschiedliche Signalveränderungen entdeckt werden. Das zweite Vergleichskriterium sind die mittleren Zeitverzögerungen, welche die Tests zur Erkennung von Signalveränderungen benötigen. In Abschnitt 4.3 werden der SCARM und der aoRM hinsichtlich der Effizienz ihrer Signalschätzungen mittels Simulationen verglichen. Abschließend werden der aoRM und SCARM in 4.4 auf echte Daten angewendet um die Signalextraktionen und die adaptierten Fensterbreiten gegenüber zu stellen.

4.1 Robustheit der Tests

Ylvisaker (1977) führte die *Resistance to Rejection* (RR) und *Resistance to Acceptance* (RA) als Robustheitskriterien für Tests ein. Diese Kriterien wurden für Daten entwickelt, deren Ordnung unerheblich ist, was für Zeitreihen jedoch nicht zutrifft. Daher schlägt Fried (2007) für diesen Datentyp eine leicht modifizierte Definition der RA und RR vor: Die RR (RA) eines Tests ist der kleinste Anteil m/n an Beobachtungen einer Stichprobe $\mathbf{y} = (y_1, \dots, y_n)$, der ersetzt werden muss, damit der Test die Nullhypothese H_0 ablehnt (nicht ablehnen kann) – unabhängig von den nicht ersetzten „guten“ Beobachtungen in \mathbf{y} . Sei $U_m(\mathbf{y})$ eine Umgebung des Datenvektors \mathbf{y} , welche alle kontaminierten Stichproben $\mathbf{z} = (z_1, \dots, z_n)$ enthält mit $z_i \neq y_i$ für höchstens $0 \leq m \leq n$ Positionen. Weiterhin sei ϕ die Entscheidungsfunktion des Tests, wobei $\phi(\cdot) = 0$ eine Nicht-Ablehnung und $\phi(\cdot) = 1$ eine Ablehnung von H_0 bedeutet. Da die RR und RA des aoRM-Tests vom gewählten Signifikanzniveau α abhängen, werden an dieser Stelle leicht modifizierte Versionen der von Fried (2007) vorgeschlagenen Test Resistances eingeführt, die das Signifikanzniveau $\alpha \in (0, 1)$ mit einbeziehen:

$$\text{RR : } \varepsilon_R(\alpha) = \frac{1}{n} \min \left\{ m \geq 0 : \inf_{\mathbf{y} \in \mathbb{R}^n} \sup_{\mathbf{z} \in U_m(\mathbf{y})} \phi(\mathbf{z}) = 1 \right\}, \quad (4.1)$$

$$\text{RA : } \varepsilon_A(\alpha) = \frac{1}{n} \min \left\{ m \geq 0 : \sup_{\mathbf{y} \in \mathbb{R}^n} \inf_{\mathbf{z} \in U_m(\mathbf{y})} \phi(\mathbf{z}) = 0 \right\}. \quad (4.2)$$

4.1.1 Test Resistances des SCARM-Tests

Zur Bestimmung der Resistances des SCARM-Tests werden die folgenden Sätze und Bemerkungen verwendet.

Lemma 4.1.

Seien $\mathbf{y} = (y_1, \dots, y_n)$ und $\mathbf{z} = (z_1, \dots, z_n)$ Stichproben mit $z_i \neq y_i$ für $i \in I_z \subset \{1, \dots, n\}$, $|I_z| = m = \lfloor n/2 \rfloor - 1 \geq 1$, und $y_i = z_i$ für $i \in I_y = \{1, \dots, n\} \setminus I_z$, $|I_y| = n - m$. Dann ist die RM-Steigung für alle Stichproben \mathbf{z} wie folgt beschränkt:

$$\hat{\beta}_{\min} \leq \hat{\beta}(\mathbf{z}) \leq \hat{\beta}_{\max}$$

mit

$$\hat{\beta}_{\min} := \min_{i,j=1,\dots,n, j \neq i} \left\{ \hat{\beta}_{ij}(\mathbf{y}) \right\}, \quad \hat{\beta}_{\max} := \max_{i,j=1,\dots,n, j \neq i} \left\{ \hat{\beta}_{ij}(\mathbf{y}) \right\}$$

und $\hat{\beta}_{ij}(\mathbf{y}) = (y_i - y_j)/(i - j)$, $i \neq j$.

Beweis.

Sei $B_i := \{\hat{\beta}_{ij}(\mathbf{z})\}_{j \neq i}$, so dass $\hat{\beta}_i(\mathbf{z}) = \text{med} B_i$ und $\hat{\beta}(\mathbf{z}) = \text{med}_{i \in \{1, \dots, n\}} \{\hat{\beta}_i(\mathbf{z})\}$. Dann ist $\hat{\beta}_i(\mathbf{z}) \in (-\infty, \infty)$ für alle $i \in I_z$, d.h. $\hat{\beta}_i(\mathbf{z})$ ist beliebig, da z_i beliebig ist. Demgegenüber enthalten die Mengen B_i mit $i \in I_y$ insgesamt $n - 1$ Steigungen $\hat{\beta}_{ij}$, von denen $m = \lfloor n/2 \rfloor - 1$ beliebig sind, da die z_i beliebig sind, und von denen $n - 1 - m = \lceil n/2 \rceil$ fest sind bzw. durch die Beobachtungen y_i , $i \in I_y$, festgelegt sind. Daher gilt für $i \in I_y$:

$$\sup_{\mathbf{z}} \left\{ \hat{\beta}_{ij}(\mathbf{z}) \right\} = \begin{cases} \hat{\beta}_{ij}(\mathbf{y}), & \text{falls } j \in I_y, j \neq i \\ \infty, & \text{falls } j \in I_z. \end{cases}$$

Somit gilt für alle $i \in I_y$:

$$\begin{aligned} \hat{\beta}_i(\mathbf{z}) &\leq \text{med}_{j \neq i} \left\{ \sup_{\mathbf{z}} \left\{ \hat{\beta}_{ij}(\mathbf{z}) \right\} \right\} = \text{med} \left\{ \underbrace{\left\{ \hat{\beta}_{ij}(\mathbf{y}) \right\}_{j \in I_y, j \neq i}}_{\lceil n/2 \rceil \text{ Elemente}} \cup \underbrace{\left\{ \infty, \dots, \infty \right\}}_{\lfloor n/2 \rfloor - 1 \text{ Elemente}} \right\} \\ &\leq \max_{j \in I_y, j \neq i} \left\{ \hat{\beta}_{ij}(\mathbf{y}) \right\}. \end{aligned}$$

Damit ist

$$\begin{aligned} \hat{\beta}(\mathbf{z}) &= \operatorname{med}_{i=1,\dots,n} \left\{ \hat{\beta}_i(\mathbf{z}) \right\} \leq \operatorname{med}_{i=1,\dots,n} \left\{ \underbrace{\max_{j \in I_y, j \neq i} \left\{ \hat{\beta}_{ij}(\mathbf{y}) \right\}}_{|I_y|=\lceil n/2 \rceil+1 \text{ Elemente}}, \underbrace{\infty, \dots, \infty}_{|I_z|=\lfloor n/2 \rfloor-1 \text{ Elemente}} \right\} \\ &\leq \max_{i \in I_y} \left\{ \max_{j \in I_y, j \neq i} \left\{ \hat{\beta}_{ij}(\mathbf{y}) \right\} \right\} = \max_{i,j \in I_y, j \neq i} \left\{ \hat{\beta}_{ij}(\mathbf{y}) \right\} \leq \max_{i,j=1,\dots,n, j \neq i} \left\{ \hat{\beta}_{ij}(\mathbf{y}) \right\}. \end{aligned}$$

Der Beweis, dass $\min_{i,j=1,\dots,n, j \neq i} \left\{ \hat{\beta}_{ij}(\mathbf{y}) \right\} \leq \hat{\beta}(\mathbf{z})$, wird analog geführt. □

Bemerkung 4.1.

Offensichtlich gilt Lemma 4.1 auch für $m < \lfloor n/2 \rfloor - 1$.

Lemma 4.2.

Seien \mathbf{y} und \mathbf{z} Stichproben wie in Lemma 4.1, wobei jedoch $|I_z| = m = \lceil (n-1 - \lfloor \delta(n-2) \rfloor) / 3 \rceil - 1 \geq 1$. Dann ist der Skalenschätzer Q_δ^{adj} für alle Stichproben \mathbf{z} wie folgt beschränkt:

$$Q_\delta^{\text{adj}}(\mathbf{z}) \leq c_n \cdot \max_{w=1,\dots,n-2} \{h_w(\mathbf{y})\},$$

wobei $c_n > 0$ und $h_w(\mathbf{y}) = |y_{w+1} - (y_w + y_{w+2})/2|$, $w = 1, \dots, n-2$, vgl. (3.12).

Beweis.

Der Q_δ^{adj} -Schätzer lässt sich schreiben als

$$Q_\delta^{\text{adj}}(\mathbf{z}) = c_n \cdot q_\delta(\mathbf{h}(\mathbf{z})), \quad \mathbf{h}(\mathbf{z}) = (h_1(\mathbf{z}), \dots, h_{n-2}(\mathbf{z})),$$

wobei $q_\delta(\cdot)$ die δ -Quantilsfunktion bezeichnet mit $q_\delta(\mathbf{x}) = x_{(\lfloor \delta n \rfloor)}$ für Stichproben $\mathbf{x} \in \mathbb{R}^n$. Sind $m = \lceil (n-1 - \lfloor \delta(n-2) \rfloor) / 3 \rceil - 1$ der Beobachtungen y_i durch beliebige z_i -Werte ersetzbar, so sind höchstens $3m$ der Dreieckshöhen $h_w(\mathbf{z})$ beliebig. Genau $3m$ Dreieckshöhen sind beliebig, wenn $i \geq 3$ und $i \leq n-2$ für alle $i \in I_z$, und wenn $|i-j| \geq 3$ für alle $i, j \in I_z$, $i \neq j$. Im Folgenden sei dieser Fall gegeben.

Sei $W_z \subset \{1, \dots, n-2\}$ mit $|W_z| = 3m$ die Indexmenge der $3m$ beliebigen Dreieckshöhen, d.h. der Dreieckshöhen, die durch ein z_i mit $z_i \neq y_i$ kontaminiert sind. Die Indexmenge der $n-2-3m$ nicht-kontaminierten Dreieckshöhen sei bezeichnet mit $W_y \subset \{1, \dots, n-2\}$

mit $|W_y| = n - 2 - 3m$. Dann gilt:

$$\forall w \in W_z : \sup_z h_w(\mathbf{z}) = \infty,$$

$$\forall w \in W_y : h_w(\mathbf{z}) = h_w(\mathbf{y}) \implies \sup_z h_w(\mathbf{z}) = h_w(\mathbf{y}) = \left| y_{w+1} - \frac{y_w + y_{w+2}}{2} \right|.$$

Des Weiteren gilt

$$\begin{aligned} Q_\delta^{\text{adj}}(\mathbf{z}) &= c_n \cdot q_\delta(\mathbf{h}(\mathbf{z})) = c_n \cdot q_\delta(h_1(\mathbf{z}), \dots, h_{n-2}(\mathbf{z})) \\ &\leq c_n \cdot q_\delta\left(\sup_z \{h_1(\mathbf{z})\}, \dots, \sup_z \{h_{n-2}(\mathbf{z})\}\right) \\ &= c_n \cdot q_\delta\left(\underbrace{\{h_w(\mathbf{y})\}_{w \in W_y}}_{|W_y|=n-2-3m \text{ Elemente}}, \underbrace{\{\infty, \dots, \infty\}}_{|W_z|=3m \text{ Elemente}}\right), \end{aligned}$$

und

$$\begin{aligned} |W_y| &= n - 2 - 3m = n + 1 - 3 \cdot \left\lceil \frac{n - 1 - \lfloor \delta(n - 2) \rfloor}{3} \right\rceil \\ &\geq n + 1 - (n - 1 - \lfloor \delta(n - 2) \rfloor + 2) = \lfloor \delta(n - 2) \rfloor. \end{aligned}$$

Diese Ungleichung gilt, da $3 \cdot \lceil x/3 \rceil \leq x + 2$ für alle $x \geq 0$. Daher gilt

$$Q_\delta^{\text{adj}}(\mathbf{z}) \leq c_n \cdot \max_{w \in W_y} \{h_w(\mathbf{y})\} \leq c_n \cdot \max_{w=1, \dots, n-2} \{h_w(\mathbf{y})\}.$$

□

Bemerkung 4.2.

Offensichtlich gilt Lemma 4.2 auch für $m < \lceil (n - 1 - \lfloor \delta(n - 2) \rfloor) / 3 \rceil - 1$.

Satz 4.1.

Sei $\mathbf{y} \in \mathbb{R}^n$ und $\mathbf{z} \in U_m(\mathbf{y})$ mit $m \leq \lfloor r/2 \rfloor - 1$. Dann ist

$$|T_{\text{SCARM}}(\mathbf{z})| \leq \frac{d_{\max}(\mathbf{y})}{b \cdot \sqrt{v_\ell + v_r}}$$

mit b , v_ℓ und v_r aus (3.22) und

$$d_{\max}(\mathbf{y}) := \max \left\{ \left| \hat{\beta}_{\text{left}}^{\max} - \hat{\beta}_{\text{right}}^{\min} \right|, \left| \hat{\beta}_{\text{right}}^{\max} - \hat{\beta}_{\text{left}}^{\min} \right| \right\},$$

wobei

$$\begin{aligned} \hat{\beta}_{\text{left}}^{\min} &:= \min_{i,j=1, \dots, n, j \neq i} \left\{ \hat{\beta}_{ij}(\mathbf{y}^{\text{left}}) \right\}, & \hat{\beta}_{\text{left}}^{\max} &:= \max_{i,j=1, \dots, n, j \neq i} \left\{ \hat{\beta}_{ij}(\mathbf{y}^{\text{left}}) \right\}, \\ \hat{\beta}_{\text{right}}^{\min} &:= \min_{i,j=1, \dots, n, j \neq i} \left\{ \hat{\beta}_{ij}(\mathbf{y}^{\text{right}}) \right\}, & \hat{\beta}_{\text{right}}^{\max} &:= \max_{i,j=1, \dots, n, j \neq i} \left\{ \hat{\beta}_{ij}(\mathbf{y}^{\text{right}}) \right\}. \end{aligned}$$

Beweis.

Nach der Definition von $\widehat{\text{Var}}(d(\mathbf{z}))$ in (3.22) gilt $\sqrt{\widehat{\text{Var}}(d(\mathbf{z}))} \geq b \cdot \sqrt{v_\ell + v_r}$. Wegen Lemma 4.1 und Bemerkung 4.1 gilt zudem für alle $\mathbf{z} \in U_m(\mathbf{y})$ mit $m \leq \lfloor r/2 \rfloor - 1$:

$$\begin{aligned} \hat{\beta}_{\text{left}}^{\min} &\leq \hat{\beta}(\mathbf{z}^{\text{left}}) \leq \hat{\beta}_{\text{left}}^{\max}, \\ \hat{\beta}_{\text{right}}^{\min} &\leq \hat{\beta}(\mathbf{z}^{\text{right}}) \leq \hat{\beta}_{\text{right}}^{\max}. \end{aligned} \quad (4.3)$$

Mit $d_{\max}(\mathbf{y}) := \max \left\{ \left| \hat{\beta}_{\text{left}}^{\max} - \hat{\beta}_{\text{right}}^{\min} \right|, \left| \hat{\beta}_{\text{right}}^{\max} - \hat{\beta}_{\text{left}}^{\min} \right| \right\}$ gilt daher

$$|T_{\text{SCARM}}(\mathbf{z})| \leq \frac{d_{\max}(\mathbf{y})}{b \cdot \sqrt{v_\ell + v_r}}.$$

□

Satz 4.2.

Sei $\mathbf{y} \in \mathbb{R}^n$ und $\mathbf{z} \in U_m(\mathbf{y})$ mit

$$m \leq \min \left\{ \left\lfloor \frac{r}{2} \right\rfloor - 1, n_\delta - 1 \right\}, \quad n_\delta = \left\lceil \frac{(n-1 - \lfloor \delta(n-2) \rfloor)}{3} \right\rceil.$$

Dann ist

$$|T_{\text{SCARM}}(\mathbf{z})| \geq \frac{d_{\min}(\mathbf{y})}{\sqrt{v_{\max}(\mathbf{y})}},$$

wobei

$$v_{\max}(\mathbf{y}) := \max \left\{ b^2 \cdot (v_\ell + v_r), c_n \cdot \max_{w=1, \dots, n-2} \{h_w(\mathbf{y})\} \cdot (v_\ell + v_r) \right\}$$

mit c_n aus (3.12) und

$$d_{\min}(\mathbf{y}) := \begin{cases} 0, & \text{falls } [\hat{\beta}_{\text{right}}^{\min}, \hat{\beta}_{\text{right}}^{\max}] \cap [\hat{\beta}_{\text{left}}^{\min}, \hat{\beta}_{\text{left}}^{\max}] \neq \emptyset \\ \min \left\{ \left| \hat{\beta}_{\text{left}}^{\max} - \hat{\beta}_{\text{right}}^{\min} \right|, \left| \hat{\beta}_{\text{right}}^{\max} - \hat{\beta}_{\text{left}}^{\min} \right| \right\}, & \text{sonst} \end{cases}.$$

Beweis.

Wegen (4.3) gilt für alle $\mathbf{z} \in U_m(\mathbf{y})$ mit $m \leq \min\{\lfloor r/2 \rfloor - 1, n_\delta - 1\} \leq \lfloor r/2 \rfloor - 1$:

$$|d(\mathbf{z})| = \left| \hat{\beta}(\mathbf{z}^{\text{left}}) - \hat{\beta}(\mathbf{z}^{\text{right}}) \right| \geq d_{\min}(\mathbf{y}).$$

Wegen Lemma 4.2 und Bemerkung 4.2 gilt zudem

$$\widehat{\text{Var}}(d(\mathbf{z})) = \max\{b^2 \cdot (v_\ell + v_r), Q_\delta^{\text{adj}}(\mathbf{z})^2 \cdot (v_\ell + v_r)\} \leq v_{\max}(\mathbf{y}),$$

und daher ist

$$|T_{\text{SCARM}}(\mathbf{z})| = \frac{|d(\mathbf{z})|}{\sqrt{\widehat{\text{Var}}(d(\mathbf{z}))}} \geq \frac{d_{\min}(\mathbf{y})}{\sqrt{v_{\max}(\mathbf{y})}}.$$

□

Korollar 4.1.

Seien \mathbf{y} und \mathbf{z} wie in Satz 4.2. Dann existiert für alle $c \geq 0$ ein $\mathbf{y} \in \mathbb{R}^n$ mit

$$\frac{d_{\min}(\mathbf{y})}{\sqrt{v_{\max}(\mathbf{y})}} > c. \quad (4.4)$$

Beweis.

Sei \mathbf{y} derart, dass $\hat{\beta}_{ij}(\mathbf{y}) = 0$ für alle $i = 1, \dots, \ell - 1$ und $j = i + 1$, und $\hat{\beta}_{ij}(\mathbf{y}) = c^* \neq 0$ für alle $i = \ell, \dots, n - 1$ und $j = i + 1$. Dann gilt $\hat{\beta}_{\text{left}}^{\min} = \hat{\beta}_{\text{left}}^{\max} = 0$, $\hat{\beta}_{\text{right}}^{\min} = \hat{\beta}_{\text{right}}^{\max} = c^*$, und es gilt $h_w(\mathbf{y}) = 0$ für alle $w \in \{1, \dots, n - 2\} \setminus \{\ell - 1\}$ und $h_w(\mathbf{y}) = |c^*|/2$ für $w = \ell - 1$. Mit $c^{**} := |c^*|$ ist somit

$$\frac{d_{\min}(\mathbf{y})}{\sqrt{v_{\max}(\mathbf{y})}} = \frac{c^{**}}{\sqrt{\max\{b^2, c_n \cdot c^{**}/2\} \cdot (v_\ell + v_r)}} = \sqrt{\frac{2c^{**}}{c_n \cdot (v_\ell + v_r)}},$$

wenn $c^{**} > 2b^2/c_n$. Hieraus folgt

$$\frac{d_{\min}(\mathbf{y})}{\sqrt{v_{\max}(\mathbf{y})}} > c \Leftrightarrow c^{**} > \frac{c^2 \cdot c_n \cdot (v_\ell + v_r)}{2}.$$

Für die oben definierte Stichprobe \mathbf{y} mit

$$c^{**} > \max \left\{ \frac{2b^2}{c_n}, \frac{c^2 \cdot c_n \cdot (v_\ell + v_r)}{2} \right\}$$

gilt also die Ungleichung (4.4). □

Im Folgenden wird mit Hilfe der obigen Sätze und Bemerkungen zunächst die RR des SCARM-Tests hergeleitet. Der Varianzschätzer

$$\widehat{\text{Var}}(d(\mathbf{y})) = \max \left\{ b, Q_\delta^{\text{adj}}(\mathbf{y}) \right\}^2 \cdot (v_\ell + v_r)$$

bildet den Nenner der SCARM-Teststatistik in (3.6). Da der Varianzschätzer durch $b > 0$ nach unten beschränkt ist, kann eine kontaminierte Stichprobe $\mathbf{z} = (\mathbf{z}^{\text{left}}, \mathbf{z}^{\text{right}}) \in U_m(\mathbf{y})$ mit $m \leq n$ nicht zu einer Implosion von $\widehat{\text{Var}}(d(\mathbf{z}))$ und somit zu einer (falschen) Ablehnung von H_0 führen. Daher reicht es aus, den Effekt von Kontaminationen allein auf den Zähler der SCARM-Teststatistik, d.h. auf die Differenz der RM-Steigungen im linken und rechten

Zeitfenster, zu untersuchen. Darüber hinaus genügt es, lediglich Kontaminationen innerhalb der rechten Stichprobe $\mathbf{z}^{\text{right}} \in U_m(\mathbf{y}^{\text{right}}) \subset \mathbb{R}^r$, $m \leq r$, zu untersuchen, da $r \leq \lfloor n/2 \rfloor$.

Im Folgenden wird daher $|d(\mathbf{z} = (\mathbf{y}^{\text{left}}, \mathbf{z}^{\text{right}}))| = |\hat{\beta}(\mathbf{y}^{\text{left}}) - \hat{\beta}(\mathbf{z}^{\text{right}})|$ betrachtet.

Aufgrund des Bruchpunkts des RM genügt es, $\lfloor r/2 \rfloor$ der Beobachtungen in $\mathbf{y}^{\text{right}}$ zu ersetzen, damit $|d(\mathbf{z})|$ und somit $|T_{\text{SCARM}}(\mathbf{z})|$ beliebig groß wird. Daher gilt

$$\varepsilon_R^{\text{SCARM}}(\alpha) \leq \frac{1}{n} \cdot \left\lfloor \frac{r}{2} \right\rfloor. \quad (4.5)$$

Des Weiteren besagt Satz 4.1, dass

$$|T_{\text{SCARM}}(\mathbf{z})| \leq \frac{d_{\max}(\mathbf{y})}{b \cdot \sqrt{v_\ell + v_r}},$$

falls $m \leq \lfloor r/2 \rfloor - 1$ Beobachtungen beliebig ersetzbar sind. Dabei existiert offensichtlich immer eine Stichprobe $\mathbf{y} \in \mathbb{R}^n$, für die $d_{\max}(\mathbf{y}) = 0$. Für jeden kritischen Wert $c_\alpha(n, r) > 0$ bzw. für jedes Signifikanzniveau $\alpha \in (0, 1)$ existiert also eine Stichprobe $\mathbf{y} \in \mathbb{R}^n$, für die es nicht ausreicht, $m \leq \lfloor r/2 \rfloor - 1$ der n Beobachtungen zu ersetzen, damit $|T_{\text{SCARM}}(\mathbf{z})| > c_\alpha(n, r)$. Daher gilt

$$\varepsilon_R^{\text{SCARM}}(\alpha) > \frac{1}{n} \cdot \left(\left\lfloor \frac{r}{2} \right\rfloor - 1 \right), \quad (4.6)$$

und aus (4.5) und (4.6) folgt

$$\varepsilon_R^{\text{SCARM}}(\alpha) = \frac{1}{n} \cdot \left\lfloor \frac{r}{2} \right\rfloor.$$

Als nächstes erfolgt die Herleitung der RA des SCARM-Tests. Eine kontaminierte Stichprobe $\mathbf{z} \in U_m(\mathbf{y})$ mit $m \leq n$ bewirkt eine Nicht-Ablehnung von H_0 , wenn sie zu einem genügend kleinen Wert von $|d(\mathbf{z})|$ und/oder einem genügend großen Wert von $\widehat{\text{Var}}(d(\mathbf{z}))$ führt. Damit $|d(\mathbf{z})|$ beliebig klein wird, reicht es aus, $\lfloor r/2 \rfloor$ Beobachtungen in $\mathbf{y}^{\text{right}}$ zu ersetzen, unabhängig von der Differenz $|d(\mathbf{y})|$ der anhand der unkontaminierten Stichprobe \mathbf{y} geschätzten RM-Steigungen. Dies gilt wegen $r \leq \ell$ und aufgrund des Bruchpunkts und der Lipschitz-Stetigkeit des RM. Des Weiteren genügt es,

$$n_\delta := \left\lceil \frac{n - 1 - \lfloor \delta(n - 2) \rfloor}{3} \right\rceil \quad (4.7)$$

der Beobachtungen in \mathbf{y} zu ersetzen, damit $\widehat{\text{Var}}(d(\mathbf{z}))$ einen beliebig großen Wert annimmt, da n_δ/n der Explosionsbruchpunkt des Skalenschätzers Q_δ^{adj} ist (Gelper et al., 2009). Daher gilt

$$\varepsilon_A^{\text{SCARM}}(\alpha) \leq \frac{1}{n} \min \left\{ \left\lfloor \frac{r}{2} \right\rfloor, n_\delta \right\}. \quad (4.8)$$

Falls jedoch lediglich $m \leq \min \{ \lfloor r/2 \rfloor - 1, n_\delta - 1 \}$ Beobachtungen in \mathbf{y} beliebig ersetzt werden können, gilt nach Satz 4.2 und Korollar 4.1: Für jeden kritischen Wert $c_\alpha(n, r) \geq 0$ existiert ein $\mathbf{y} \in \mathbb{R}^n$, so dass

$$|T_{\text{SCARM}}(\mathbf{z})| \geq \frac{d_{\min}(\mathbf{y})}{\sqrt{v_{\max}(\mathbf{y})}} > c_\alpha(n, r).$$

Unabhängig vom kritischen Wert $c_\alpha(n, r)$ bzw. vom gewählten Signifikanzniveau α genügt es also nicht, lediglich $m \leq \min \{ \lfloor r/2 \rfloor - 1, n_\delta - 1 \}$ Beobachtungen in \mathbf{y} zu ersetzen, um eine Nicht-Ablehnung von H_0 zu erwirken. Folglich gilt

$$\varepsilon_A^{\text{SCARM}}(\alpha) > \frac{1}{n} \min \left\{ \left\lfloor \frac{r}{2} \right\rfloor, n_\delta \right\}, \quad (4.9)$$

und aus (4.8) und (4.9) folgt

$$\varepsilon_A^{\text{SCARM}}(\alpha) = \frac{1}{n} \min \left\{ \left\lfloor \frac{r}{2} \right\rfloor, n_\delta \right\}.$$

Als nächstes werden die Test Resistances des aoRM-Tests bestimmt.

4.1.2 Test Resistances des aoRM-Tests

Für eine gegebene Stichprobe \mathbf{y} kann die aoRM-Teststatistik aus (2.10) auch geschrieben werden als $T_{\text{aoRM}}(\mathbf{y}) = |n_{\text{pos}}(\mathbf{y}) - n_{\text{neg}}(\mathbf{y})|$. Dabei bezeichnen $n_{\text{pos}}(\cdot)$ und $n_{\text{neg}}(\cdot)$ die Anzahl an positiven und negativen Residuen unter den n_I letzten Residuen der RM-Anpassung im Zeitfenster der Länge n mit $n_I \leq \lfloor n/2 \rfloor$. An dieser Stelle sei darauf hingewiesen, dass $n_I = n_{\text{pos}}(\cdot) + n_{\text{neg}}(\cdot) + n_{\text{zero}}(\cdot)$, wobei die Anzahl an Nullresiduen $n_{\text{zero}}(\cdot)$ mit positiver Wahrscheinlichkeit größer als Null ist (Schettlinger, 2009, Kap. 2.2). Im Folgenden bezeichne

$$\begin{aligned} \hat{y}_i &= \hat{\mu}(\mathbf{y}) + \hat{\beta}(\mathbf{y}) \cdot (i - n), \\ \hat{z}_i &= \hat{\mu}(\mathbf{z}) + \hat{\beta}(\mathbf{z}) \cdot (i - n), \end{aligned}$$

$i = 1, \dots, n$, die anhand \mathbf{y} bzw. \mathbf{z} geschätzten RM-Geraden und $\mathbf{r}(y_i) = y_i - \hat{y}_i$ bzw. $\mathbf{r}(z_i) = z_i - \hat{z}_i$ die entsprechenden Residuen. Als nächstes werden die RA sowie Grenzen für die RR des aoRM-Tests bestimmt, wofür die folgenden Sätze und Bemerkungen benötigt werden.

Lemma 4.3.

Seien \mathbf{y} und \mathbf{z} wie in Lemma 4.1. Dann gilt

$$\hat{\mu}_{\min} \leq \hat{\mu}(\mathbf{z}) \leq \hat{\mu}_{\max},$$

wobei

$$\begin{aligned}\hat{\mu}_{\min} &:= \operatorname{med}_{i=1,\dots,n} \left\{ y_{\min} + \hat{\beta}_{\min} \cdot (n - i) \right\}, \\ \hat{\mu}_{\max} &:= \operatorname{med}_{i=1,\dots,n} \left\{ y_{\max} + \hat{\beta}_{\max} \cdot (n - i) \right\}\end{aligned}$$

mit $y_{\min} := \min_{i \in I_y} \{y_i\}$, $y_{\max} := \max_{i \in I_y} \{y_i\}$ und $\hat{\beta}_{\min}$ und $\hat{\beta}_{\max}$ aus Lemma 4.1.

Beweis.

Es gilt

$$\begin{aligned}\hat{\mu}(\mathbf{z}) = \operatorname{med}_{i=1,\dots,n} \left\{ z_i + \hat{\beta}(\mathbf{z}) \cdot (n - i) \right\} &\leq \operatorname{med}_{i=1,\dots,n} \left\{ z_i + \hat{\beta}_{\max} \cdot (n - i) \right\} \\ &\leq \operatorname{med}_{i=1,\dots,n} \left\{ y_{\max} + \hat{\beta}_{\max} \cdot (n - i) \right\} =: \hat{\mu}_{\max}.\end{aligned}$$

Die erste Ungleichung folgt aus Lemma 4.1, und die zweite Ungleichung gilt, da $m \leq \lfloor n/2 \rfloor - 1$. Analog lässt sich zeigen, dass $\hat{\mu}(\mathbf{z}) \geq \hat{\mu}_{\min}$. □

Bemerkung 4.3.

Offensichtlich gilt Lemma 4.3 auch für $m < \lfloor n/2 \rfloor - 1$.

Satz 4.3.

Sei $\mathbf{y} \in \mathbb{R}^n$ eine Stichprobe und $\mathbf{z} \in U_m(\mathbf{y})$ mit $m \leq \lfloor n/2 \rfloor - 1$. Dann gilt

$$z_i^{\min} \leq \hat{z}_i \leq z_i^{\max}$$

für alle $i = 1, \dots, n$, wobei

$$z_i^{\min} := \hat{\mu}_{\min} + \hat{\beta}_{\max} \cdot (i - n) \quad \text{und} \quad z_i^{\max} := \hat{\mu}_{\max} + \hat{\beta}_{\min} \cdot (i - n),$$

mit $\hat{\beta}_{\min}$, $\hat{\beta}_{\max}$, $\hat{\mu}_{\min}$, $\hat{\mu}_{\max}$ aus Lemma 4.1 bzw. 4.3.

Beweis.

Folgt direkt aus den Lemmata 4.1 und 4.3. □

Bemerkung 4.4.

Nach Satz 4.3 ist die anhand der kontaminierten Stichprobe $\mathbf{z} \in U_m(\mathbf{y})$, $m \leq \lfloor n/2 \rfloor - 1$, geschätzte RM-Gerade an allen Designpunkten $i = 1, \dots, n$ nach oben und unten beschränkt durch z_i^{\min} und z_i^{\max} . Dabei sind die Grenzen durch die unkontaminierte Stichprobe \mathbf{y} festgelegt, da $\hat{\beta}_{\min}$, $\hat{\beta}_{\max}$, $\hat{\mu}_{\min}$, und $\hat{\mu}_{\max}$ nur von \mathbf{y} abhängen. Folglich können die Residuen $\mathbf{r}(z_i)$ mit $z_i \neq y_i$ jede reelle Zahl annehmen, also beliebig groß oder klein sein. Sie können z.B. alle negativ sein.

Lemma 4.4.

Für jede Stichprobe $\mathbf{y} \in \mathbb{R}^n$ resultiert eine RM-Regression in jeweils höchstens $\lfloor n/2 \rfloor$ positive und negative Residuen, d.h.

$$\#\{\mathbf{r}(y_i) > 0\} \leq \lfloor n/2 \rfloor \quad \text{und} \quad \#\{\mathbf{r}(y_i) < 0\} \leq \lfloor n/2 \rfloor.$$

Beweis.

Sei $\hat{\beta} := \hat{\beta}(\mathbf{y})$ die anhand \mathbf{y} geschätzte RM-Steigung. Mit $y_i^* := y_i + \hat{\beta} \cdot (n - i)$ ist dann

$$\hat{\mu} := \hat{\mu}(\mathbf{y}) = \underset{i=1, \dots, n}{\text{med}} \{y_i^*\}.$$

Folglich sind jeweils höchstens $\lfloor n/2 \rfloor$ der y_i^* kleiner bzw. größer als $\hat{\mu}$, d.h.

$$\#\{y_i^* > \hat{\mu}\} \leq \lfloor n/2 \rfloor \quad \text{und} \quad \#\{y_i^* < \hat{\mu}\} \leq \lfloor n/2 \rfloor.$$

(Tatsächlich beträgt die Obergrenze $(n - 1)/2$ für n ungerade und $n/2$ für n gerade.) Dabei gilt $y_i^* > \hat{\mu} \Leftrightarrow \mathbf{r}(y_i) > 0$, da

$$y_i^* = y_i + \hat{\beta} \cdot (n - i) > \hat{\mu} \Leftrightarrow y_i > \hat{\mu} + \hat{\beta} \cdot (i - n) = \hat{y}_i \Leftrightarrow y_i - \hat{y}_i = \mathbf{r}(y_i) > 0.$$

Analog gilt $y_i^* < \hat{\mu} \Leftrightarrow \mathbf{r}(y_i) < 0$.

□

Satz 4.4.

Sei $\mathbf{y} \in \mathbb{R}^n$ eine Stichprobe derart, dass die RM-Residuen $\mathbf{r}(y_i) > 0$ für alle $i = n - n_I + 1, \dots, n$ mit $n_I \leq \lfloor n/2 \rfloor$. (Dann gilt $T_{\text{aoRM}}(\mathbf{y}) = n_I$.) Sei weiterhin $\mathbf{z} \in \mathbb{R}^n$ eine kontaminierte Stichprobe mit $z_i = y_i$ für alle $i = 1, \dots, n - n_I$ und $\mathbf{r}(z_i) < 0$ für alle $i \in I_z \subset \{n - n_I + 1, \dots, n\}$, $|I_z| = m < n_I$. (Nach Bemerkung 4.4 existiert immer eine derartige Stichprobe \mathbf{z} .) Dann besitzen die nicht-ersetzten Beobachtungen im rechten Zeitfenster $\{n - n_I + 1, \dots, n\}$ weiterhin positive Residuen, d.h.

$$\mathbf{r}(z_i) > 0 \quad \forall i \in \{n - n_I + 1, \dots, n\} \setminus I_z.$$

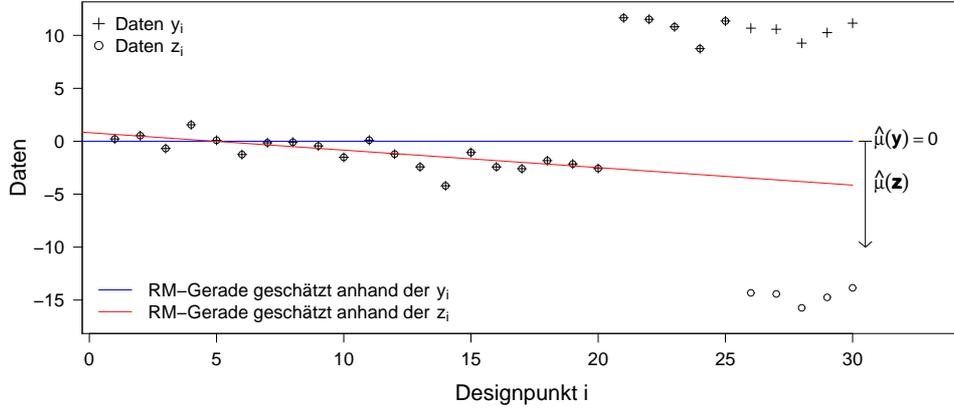


Abbildung 4.1: Der Effekt von Ersetzungen auf die RM-Gerade; hier ist $n = 30$, $n_I = 10$ und $m = 5$.

Beweis.

Die Abbildung 4.1 dient zur Veranschaulichung der in diesem Beweis dargestellten Sachverhalte. O.B.d.A. sei $\hat{\beta}(\mathbf{y}) = \hat{\mu}(\mathbf{y}) = 0$. Dann ist $y_i > 0$ für alle $i = n - n_I + 1, \dots, n$. Nach Lemma 4.4 kann die jeweilige Anzahl an positiven und negativen Residuen nicht größer sein als $\lfloor n/2 \rfloor$. Wenn sich alle n_I Beobachtungen y_i im rechten Zeitfenster $\{n - n_I + 1, \dots, n\}$ oberhalb der anhand \mathbf{y} geschätzten RM-Geraden befinden, liegen also höchstens $\lfloor n/2 \rfloor - n_I$ der y_i im linken Zeitfenster $\{1, \dots, n - n_I\}$ oberhalb der RM-Geraden. Die übrigen

$$n^* := n - n_I - \lfloor n/2 \rfloor + n_I = \lceil n/2 \rceil$$

Beobachtungen y_i liegen unterhalb oder auf der RM-Geraden. Für mindestens $n^* = \lceil n/2 \rceil$ der y_i im linken Zeitfenster gilt demnach $y_i \leq 0$.

Falls m der y_i im rechten Zeitfenster der Länge $n_I \leq \lfloor n/2 \rfloor$ derart durch z_i -Werte ersetzt werden, dass $\mathfrak{r}(z_i) < 0$, so folgt offensichtlich $\hat{\beta}(\mathbf{z}) \leq \hat{\beta}(\mathbf{y}) = 0$ und damit

$$\hat{\mu}(\mathbf{z}) \leq \hat{\mu}(\mathbf{y}) = 0. \quad (4.10)$$

Dies ist in Abbildung 4.1 am rechten Rand durch den senkrecht nach unten gerichteten Pfeil gekennzeichnet. Nun wird angenommen, dass mindestens ein z_i des rechten Fensters mit $i \in \{n - n_I + 1, \dots, n\} \setminus I_z$, d.h. $z_i = y_i \geq 0$, unterhalb der anhand \mathbf{z} geschätzten RM-Geraden liegt. Wegen $\hat{\mu}(\mathbf{z}) \leq 0$ lägen dann auch die n^* z_i -Werte des linken Zeitfensters mit $z_i = y_i \leq 0$ weiterhin unterhalb der RM-Geraden. Zudem würden auch die m z_i -Werte mit $i \in I_z$ unterhalb der RM-Geraden liegen, so dass insgesamt mindestens $n^* + m + 1 =$

$\lceil n/2 \rceil + m + 1 > \lfloor n/2 \rfloor$ der Residuen $\mathbf{r}(z_i)$ negativ wären. Dies steht jedoch im Widerspruch zu Lemma 4.4. Folglich können alle z_i mit $i \in \{n - n_I + 1, \dots, n\} \setminus I_z$ nur oberhalb der anhand \mathbf{z} geschätzten RM-Geraden liegen, so dass die zugehörigen Residuen positiv sind. \square

Bemerkung 4.5.

Offensichtlich gilt Satz 4.4 analog falls $\mathbf{r}(y_i) < 0$ für alle $i = n - n_I + 1, \dots, n$.

Im Folgenden wird zunächst die RR des aoRM-Tests betrachtet. Dazu wird vorausgesetzt, dass der kritische Wert $c_\alpha(n, n_I) \in \{0, 1, \dots, n_I - 4\}$. Die Untergrenze null ist sinnvoll, da H_0 für $c_\alpha(n, n_I) < 0$ immer verworfen werden würde. Die Gründe für die Obergrenze $n_I - 4$ werden im weiteren Verlauf deutlich. Diese Grenze stellt jedoch keine Einschränkung dar, falls n_I und α nicht beide sehr klein sind. Beispielsweise zeigt eine Simulation anhand von $N = 100000$ Stichproben unabhängig standardnormalverteilter Daten mit $n = 20$ und $n_I = 10$, dass das Maximum der aoRM-Teststatistik unter H_0 nicht größer ist als $6 = n_I - 4$. Eine Stichprobe \mathbf{y} spricht am meisten für H_0 , falls

$$T_{\text{aoRM}}(\mathbf{y}) = |n_{\text{pos}}(\mathbf{y}) - n_{\text{neg}}(\mathbf{y})| = 0.$$

Damit die m Ersetzungen in $\mathbf{z} \in U_m(\mathbf{y})$ eine Ablehnung von H_0 bewirken, muss gelten:

$$T_{\text{aoRM}}(\mathbf{z}) = |n_{\text{pos}}(\mathbf{z}) - n_{\text{neg}}(\mathbf{z})| > c_\alpha(n, n_I). \quad (4.11)$$

Da $n_I = n_{\text{neg}}(\mathbf{z}) + n_{\text{pos}}(\mathbf{z}) + n_{\text{zero}}(\mathbf{z})$ und $n_{\text{neg}}(\mathbf{z}), n_{\text{pos}}(\mathbf{z}), n_{\text{zero}}(\mathbf{z}) \geq 0$, folgt aus der Ungleichung (4.11), dass entweder $n_{\text{pos}}(\mathbf{z}) \geq m^*$ oder $n_{\text{neg}}(\mathbf{z}) \geq m^*$ mit

$$m^* = \left\lfloor \frac{c_\alpha(n, n_I) + n_I}{2} \right\rfloor + 1$$

gelten muss. Mindestens m^* der n_I Residuen müssen somit dasselbe Vorzeichen haben. Nach Satz 4.3 können m^* der y_i derart durch z_i ersetzt werden, dass die zugehörigen Residuen $\mathbf{r}(z_i)$ alle dasselbe Vorzeichen aufweisen, falls $m^* \leq \lfloor n/2 \rfloor - 1$. Letztere Bedingung ist unter der oben getroffenen Einschränkung $c_\alpha(n, n_I) \leq n_I - 4$ erfüllt. Folglich genügt es, m^* der Beobachtungen in \mathbf{y} zu ersetzen, um eine Ablehnung von H_0 zu erwirken, unabhängig von der unkontaminierten Stichprobe \mathbf{y} . Daher ist

$$\varepsilon_R^{\text{aoRM}}(\alpha) \leq \frac{1}{n} \cdot m^*.$$

Des Weiteren reicht es nicht für jede Stichprobe \mathbf{y} aus, lediglich $m = c_\alpha(n, n_I)$ Beobachtungen zu ersetzen, damit $T_{\text{aoRM}}(\mathbf{z}) > c_\alpha(n, n_I)$. Betrachte dazu eine Stichprobe \mathbf{y} von n kollinearen Beobachtungen. Aufgrund der Exact-fit-Eigenschaft des RM (vgl. S. 8) verläuft die Gerade exakt durch diese Beobachtungen, und alle Residuen sind null. Da $m = c_\alpha(n, n_I) \leq n_I - 1$ und $n_I \leq \lfloor n/2 \rfloor$, gilt $m \leq \lfloor n/2 \rfloor - 1$. Falls $m \leq \lfloor n/2 \rfloor - 1$ Beobachtungen in \mathbf{y} ersetzt werden, liegen weiterhin mindestens $n - m = n - \lfloor n/2 \rfloor + 1$ Beobachtungen in \mathbf{z} kollinear. Folglich verläuft die RM-Gerade weiterhin durch diese Beobachtungen. Daher gibt es mindestens $n_I - c_\alpha(n, n_I)$ Nullresiduen im rechten Zeitfenster $\{n - n_I + 1, \dots, n\}$, so dass $T_{\text{aoRM}}(\mathbf{z}) \leq c_\alpha(n, n_I)$. Die RR des aoRM-Tests ist somit wie folgt beschränkt:

$$\frac{1}{n} \cdot (c_\alpha(n, n_I) + 1) \leq \varepsilon_R^{\text{aoRM}}(\alpha) \leq \frac{1}{n} \cdot \left(\left\lfloor \frac{c_\alpha(n, n_I) + n_I}{2} \right\rfloor + 1 \right).$$

Als nächstes wird die RA des aoRM-Tests betrachtet. Eine Stichprobe \mathbf{y} spricht am meisten für H_1 , wenn $T_{\text{aoRM}}(\mathbf{y}) = n_I$. Dies ist der Fall, falls entweder $n_{\text{pos}}(\mathbf{y}) = n_I$ oder $n_{\text{neg}}(\mathbf{y}) = n_I$. O.B.d.A. wird angenommen, dass $n_{\text{pos}}(\mathbf{y}) = n_I$, d.h. alle Residuen $\mathbf{r}(y_i)$ im rechten Zeitfenster $\{n - n_I + 1, \dots, n\}$ seien positiv.

Für $T_{\text{aoRM}}(\mathbf{z}) \leq c_\alpha(n, n_I)$ müssen die m Ersetzungen in \mathbf{z} bewirken, dass eine gewisse Anzahl m^{**} der n_I Residuen im rechten Zeitfenster negativ wird. Dann ist $|n_{\text{pos}}(\mathbf{z}) - n_{\text{neg}}(\mathbf{z})| \leq c_\alpha(n, n_I)$ mit $n_{\text{pos}}(\mathbf{z}) = n_I - m^{**}$ und $n_{\text{neg}}(\mathbf{z}) = m^{**}$. Dies führt zu $|n_I - 2m^{**}| \leq c_\alpha(n, n_I)$ und somit zu

$$\left\lfloor \frac{n_I - c_\alpha(n, n_I)}{2} \right\rfloor \leq m^{**} \leq \left\lfloor \frac{n_I + c_\alpha(n, n_I)}{2} \right\rfloor,$$

d.h. mindestens $m^{**} = \lceil (n_I - c_\alpha(n, n_I))/2 \rceil$ der n_I Residuen im rechten Zeitfenster müssen negativ werden. Wegen $n_I \leq \lfloor n/2 \rfloor$ gilt $m^{**} = \lceil (n_I - c_\alpha(n, n_I))/2 \rceil \leq \lfloor n/2 \rfloor - 1$, falls $n \geq 4$. Nach Satz 4.3 können dann m^{**} der y_i derart durch hinreichend kleine z_i -Werte ersetzt werden, dass die zugehörigen m^{**} Residuen $\mathbf{r}_i(\mathbf{z})$ negativ sind. Darüber hinaus sind die zu den $n_I - m^{**}$ nicht ersetzten Beobachtungen gehörenden Residuen im rechten Zeitfenster nach Satz 4.4 weiterhin positiv. Daher kann für jede Stichprobe \mathbf{y} eine Nicht-Ablehnung von H_0 herbeigeführt werden, indem $m^{**} = \lceil (n_I - c_\alpha(n, n_I))/2 \rceil$ Beobachtungen in \mathbf{y} ersetzt werden. Es gilt somit

$$\varepsilon_A^{\text{aoRM}}(\alpha) \leq \frac{1}{n} \cdot \left\lfloor \frac{n_I - c_\alpha(n, n_I)}{2} \right\rfloor. \quad (4.12)$$

Als nächstes wird untersucht, ob es immer genügt, lediglich

$$m = m^{**} - 1 = \left\lceil \frac{n_I - c_\alpha(n, n_I)}{2} \right\rceil - 1$$

Beobachtungen in \mathbf{y} zu ersetzen, um eine Nicht-Ablehnung von H_0 zu bewirken. Sei \mathbf{y} eine Stichprobe, bei der die $n - n_I$ Beobachtungen im linken Zeitfenster $\{1, \dots, n - n_I\}$ kollinear sind, so dass die RM-Gerade durch diese Punkte verläuft. Des Weiteren seien die Beobachtungen im rechten Zeitfenster oberhalb dieser RM-Geraden, und es sei $n - n_I \geq n - \lfloor n/2 \rfloor + 1 + m$, was für $n \geq 3n_I$ erfüllt ist. Werden nun $m = m^{**} - 1$ Beobachtungen des linken Zeitfensters ersetzt, bleibt der Wert der Teststatistik unverändert, d.h. $T_{\text{aoRM}}(\mathbf{z}) = n_I$, da weiterhin mindestens $n - \lfloor n/2 \rfloor + 1$ Beobachtungen z_i kollinear liegen. Werden $m = m^{**} - 1$ Beobachtungen des rechten Zeitfensters durch beliebige z_i ersetzt, sind höchstens m der Residuen $\mathbf{r}(z_i)$ im rechten Zeitfenster negativ. Werden $m = m^{**} - 1$ Beobachtungen ersetzt, mindestens eine im rechten und linken Zeitfenster, ändert sich die RM-Gerade nicht, da weiterhin genügend Beobachtungen kollinear sind. Folglich werden weniger als m der n_I Residuen $\mathbf{r}(z_i)$ im rechten Zeitfenster negativ. In jedem Fall ist der kleinstmögliche Wert der Teststatistik

$$T_{\text{aoRM}}(\mathbf{z}) = |n_I - 2m| = c_\alpha(n, n_I) + 2 > c_\alpha(n, n_I).$$

Daher genügt es nicht für jede Stichprobe \mathbf{y} , lediglich $m = m^{**} - 1$ Beobachtungen zu ersetzen, damit H_0 nicht abgelehnt wird. Mit (4.12) folgt also

$$\varepsilon_A^{\text{aoRM}}(\alpha) \begin{cases} = m^{**}/n, & \text{falls } n \geq 3n_I \\ \leq m^{**}/n, & \text{sonst} \end{cases} \quad \text{mit } m^{**} = \left\lceil \frac{n_I - c_\alpha(n, n_I)}{2} \right\rceil.$$

Offenbar sind die Test Resistances des aoRM-Tests abhängig vom gewählten Signifikanzniveau, wobei ein kleines α zu einer hohen RR aber zu einer niedrigen RA führt.

4.1.3 Vergleich der Test Resistances

Abbildung 4.2 zeigt die RA und RR des SCARM- und aoRM-Tests exemplarisch für $n = 100$ und $r, n_I \in \{5, \dots, n/2\}$. Dabei sind für die RR des aoRM-Tests die Unter- und Obergrenze eingezeichnet. Hier ist $\alpha = 0.1$, was dem von Schettlinger et al. (2010a) empfohlenen Signifikanzniveau entspricht, und $\delta = 0.5$, wie von Gelper et al. (2009) empfohlen. Im Folgenden werden zunächst die RA des aoRM- und SCARM-Tests verglichen. Dazu sei der Einfachheit halber stets $n_I = r$.

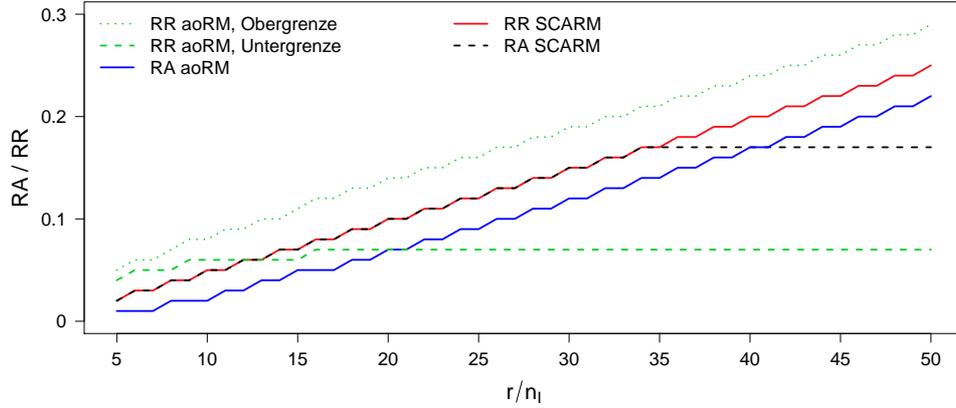


Abbildung 4.2: Test Resistances des aoRM- und SCARM-Tests für $n = 100$, $r, n_I \in \{5, \dots, n/2\}$, $\delta = 0.5$ und $\alpha = 0.1$.

Für $n = 100$ und $\delta = 0.5$ ist $\varepsilon_A^{\text{SCARM}} = 1/n \min\{\lfloor r/2 \rfloor, 0.17\}$. Folglich ist in diesem Beispiel die RA des SCARM-Tests gleich 0.17 falls $r \geq 34$. Für $r = 40, 41$ sind die RA des aoRM- und SCARM-Tests gleich. Für $r < 40$ weist der Test des SCARM eine höhere RA auf, für $r > 41$ ist das Gegenteil der Fall. Die RA des SCARM-Tests kann jedoch vergrößert werden, indem δ so gewählt wird, dass der Explosionsbruchpunkt des Q_δ^{adj} Skalenschätzers maximal ist, d.h.

$$\delta = \delta_{\text{opt}} = \frac{n+1}{4(n-2)},$$

vgl. Gelper et al. (2009). Falls $\delta = \delta_{\text{opt}}$, gilt $\varepsilon_A^{\text{SCARM}}(\alpha) > \varepsilon_A^{\text{aoRM}}(\alpha)$ für alle $r, n_I \in \{5, \dots, 50\}$, siehe Abbildung C.7 in Anhang C.

Ein Vergleich der RR des aoRM- und SCARM-Tests gestaltet sich schwieriger, da für $\varepsilon_R^{\text{aoRM}}(\alpha)$ lediglich Ober- und Untergrenzen bestehen. Für $r \leq 13$ ist die RR des SCARM-Tests nicht größer als die Untergrenze der RR des aoRM-Tests. Für $r \geq 14$ verläuft die RR des SCARM-Tests innerhalb der für die RR des aoRM-Tests bestehenden Grenzen.

Wird die Obergrenze der RR des aoRM-Tests als Maßstab genommen, lässt sich konstatieren, dass der SCARM-Test eine größere RA, aber eine kleinere RR als der aoRM-Test bietet. Dabei hängen die Resistances des aoRM-Tests vom Signifikanzniveau α ab; je kleiner α , desto kleiner die RA, aber desto größer die Grenzen für die RR, vgl. Abbildung C.7 und C.8 in Anhang C. Demgegenüber sind die Resistances des SCARM-Tests unabhängig von α .

4.2 Erkennung von Signalveränderungen

Im Folgenden werden der aoRM- und SCARM-Test in einer Simulationsstudie hinsichtlich der Erkennung von Signalveränderungen verglichen. Als Kriterien dienen die Güte der Tests sowie die mittlere Zeitverzögerung, mit der Signalveränderungen erkannt werden. Zur Vereinfachung der Notation bezeichne r im Folgenden für beide Verfahren die Breite des rechten Zeitfensters. Der aoRM führt demnach eine RM-Regression im Zeitfenster der Länge $n = \ell + r$ durch und testet anschließend auf Basis der letzten r Residuen. Die Güte der Tests wird approximiert über die Entdeckungsraten, d.h. über die relativen Häufigkeiten, mit der Signalveränderungen erkannt werden. Der Vergleich der Entdeckungsraten geschieht für verschieden ausgeprägte Level Shifts und Trendänderungen und unter Verwendung verschiedener Fensterbreiten (ℓ, r) und Signifikanzniveaus α . Die kritischen Werte des aoRM-Tests werden für jede Kombination von ℓ , r und α mittels Simulationen auf Basis von standardnormalverteiltem Rauschen gewonnen. Dabei ist der kritische Wert zum Signifikanzniveau α das $(1 - \alpha)$ -Quantil der absoluten realisierten aoRM-Teststatistiken.

4.2.1 Güte der Tests

Für den Vergleich der Entdeckungsraten werden für verschiedene (ℓ, r) -Kombinationen Zeitreihen (x_t) und (y_t) , $t = 1, \dots, n^*$, generiert, die einen Level Shift bzw. eine Trendänderung zum Zeitpunkt $t = n + 1 = \ell + r + 1 < n^*$ aufweisen. Die Zeitreihen (x_t) ergeben sich aus

$$x_t = \begin{cases} \varepsilon_t, & t = 1, \dots, n \\ a + \varepsilon_t, & t = n + 1, \dots, n^* \end{cases}.$$

Dabei ist $\varepsilon_t \sim N(0, 1)$, und a ist die Höhe des Level Shifts. Die Zeitreihen (y_t) ergeben sich aus

$$y_t = \begin{cases} \varepsilon_t, & t = 1, \dots, n \\ b \cdot (t - n) + \varepsilon_t, & t = n + 1, \dots, n^* \end{cases},$$

wobei b die Steigung des zum Zeitpunkt $t = n + 1$ beginnenden Trends ist. Für jede Kombination von ℓ und r sowie für $a = 1, 1.5, \dots, 4$ bzw. $b = 0.1, 0.15, \dots, 0.4$ werden $N = 1000$ Zeitreihen (x_t) bzw. (y_t) generiert. Anschließend werden der aoRM- und SCARM-Test jeweils mit dem Signifikanzniveau $\alpha = 0.001, 0.005, 0.01$ in einem gleitenden Zeitfenster angewendet, wobei die Fensterbreiten ℓ und r fest sind. Für jeden der N Durchläufe ergeben sich demnach $n^* - n + 1$ Testentscheidungen zu den Zeitpunkten $t = n, \dots, n^*$. Für einen Durchlauf gilt die Signalveränderung als erkannt, falls mindestens einer der $n^* - n + 1$

Tests die Nullhypothese ablehnt. Die Entdeckungsrate ist dann der Anteil an Durchläufen, bei denen die Signalveränderung erkannt wird.

Da die Entdeckungsrate für $n^* \rightarrow \infty$ sogar für $a = 0$ und $b = 0$ offenbar gegen eins konvergiert, darf n^* nicht zu groß gewählt sein. Andererseits sollte n^* nicht zu klein sein, um den Testverfahren genügend Zeitpunkte zur Entdeckung der Signalveränderungen einzuräumen. Zur Bestimmung eines geeigneten n^* werden zunächst für $(r, \ell) = (20, 20)$ und $(40, 40)$ jeweils $N = 1000$ Zeitreihen (x_t) bzw. (y_t) der Länge $n^* = 2n = 2(\ell + r)$ generiert, die einen extremen Level Shift bzw. Trendwechsel mit $a = 100$ bzw. $b = 100$ zum Zeitpunkt $t = n + 1$ aufweisen. Der aoRM- und SCARM-Test werden dann mit den entsprechenden festen Fensterbreiten $r = \ell$ und kleinem Signifikanzniveau $\alpha = 0.001$ zu den Zeitpunkten $t = n + 1, \dots, 2n$ angewendet. Aufgrund dieses Simulationsdesigns sind die Entdeckungsraten für beide Verfahren jeweils (annähernd) eins. Es zeigt sich, dass die Signalveränderungen erwartungsgemäß fast gänzlich im Bereich der Zeitpunkte $n + 1, \dots, n + r$ erkannt werden. Daher wird zur Bestimmung der Entdeckungsraten $n^* = n + r$ verwendet.

Die Abbildung 4.3 zeigt die Entdeckungsraten des aoRM- und SCARM-Tests für Level Shifts mit $a = 1, \dots, 4$, $\alpha = 0.001, 0.005, 0.01$ und $(\ell, r) = (20, 20), (40, 20), (200, 20)$. Für $(\ell, r) = (20, 20)$ (Abbildung 4.3, oben) weist der aoRM eine sehr geringe Güte auf, insbesondere wenn das Signifikanzniveau klein ist. Der SCARM zeigt deutlich höhere Entdeckungsraten. Für $(\ell, r) = (40, 20)$ (Abbildung 4.3, Mitte) erhöht sich die Güte des SCARM leicht und die des aoRM deutlich, wobei der SCARM weiterhin eine höhere Güte liefert. Für $(\ell, r) = (200, 20)$ (Abbildung 4.3, unten) zeigen beide Verfahren bei deutlichen Level Shifts Entdeckungsraten von annähernd eins. Bei geringfügigen Shifts liefert der aoRM bessere Entdeckungsraten. Vergleichbare Ergebnisse ergeben sich für $(\ell, r) = (40, 40), (80, 40), (200, 40)$, siehe Abbildung C.9 in Anhang C. Generell gilt: Falls ℓ und r ähnlich groß sind, weist der SCARM für Level Shifts deutlich höhere Entdeckungsraten als der aoRM auf. Ist jedoch ℓ groß im Vergleich zu r , zeigen beide Verfahren hohe Entdeckungsraten, wobei der aoRM für kleine Level Shifts etwas besser ist. Des Weiteren hängt die Güte beider Verfahren wie zu erwarten vom gewählten Signifikanzniveau ab, wobei der Effekt für den aoRM ausgeprägter ist. Der SCARM liefert bei kleinem Signifikanzniveau bessere Ergebnisse.

Abbildung 4.4 zeigt die Entdeckungsraten des aoRM- und SCARM-Tests bei Trendwechseln mit $b = 0.1, \dots, 0.4$, $\alpha = 0.001, 0.005, 0.01$ und $(\ell, r) = (20, 20), (40, 20), (200, 20)$. Auch für Trendwechsel weist der SCARM eine deutlich höhere Güte als der aoRM auf,

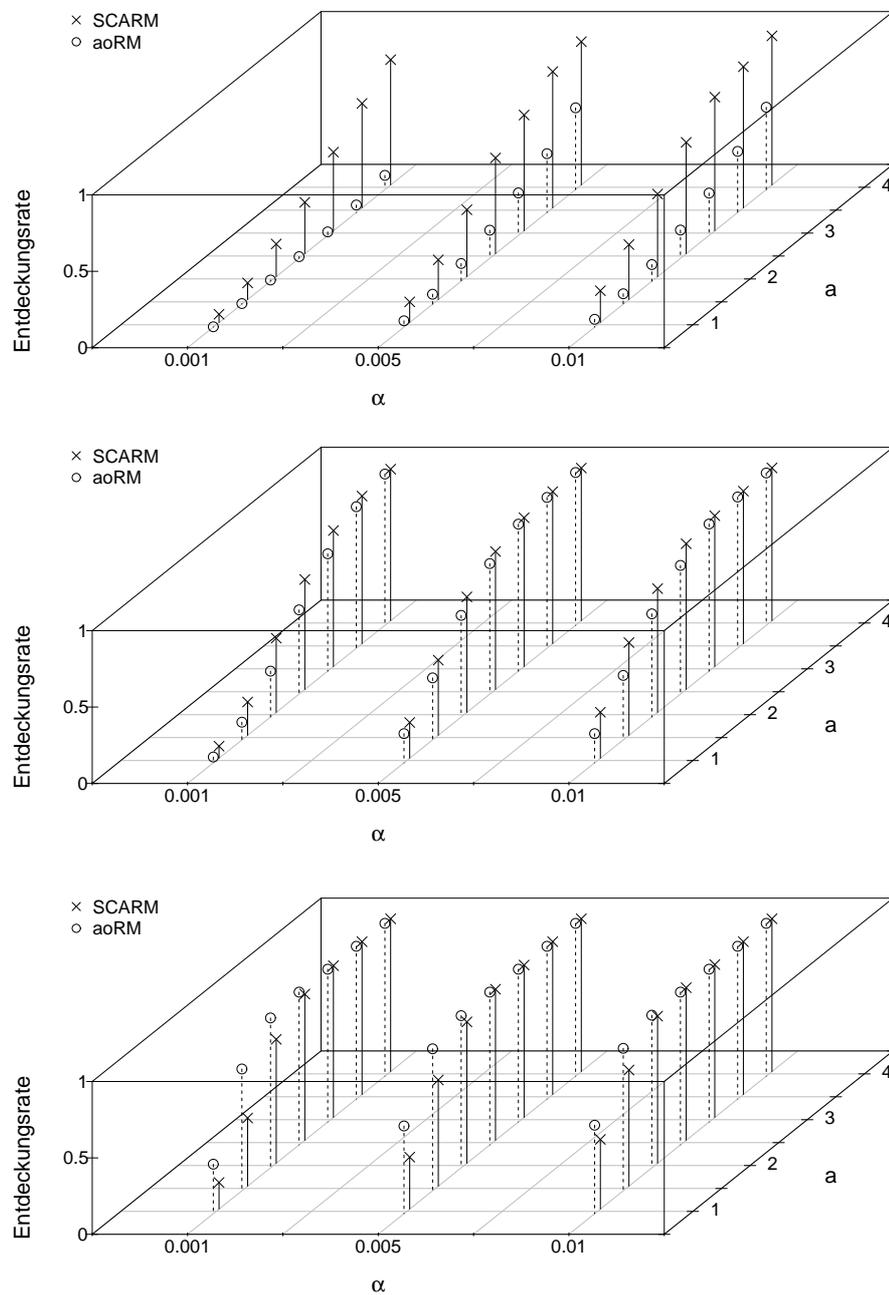


Abbildung 4.3: Entdeckungsraten des aoRM- und SCARM-Tests für verschieden hohe Level Shifts mit $(\ell, r) = (20, 20)$ (oben), $(\ell, r) = (40, 20)$ (Mitte) und $(\ell, r) = (200, 20)$ (unten).

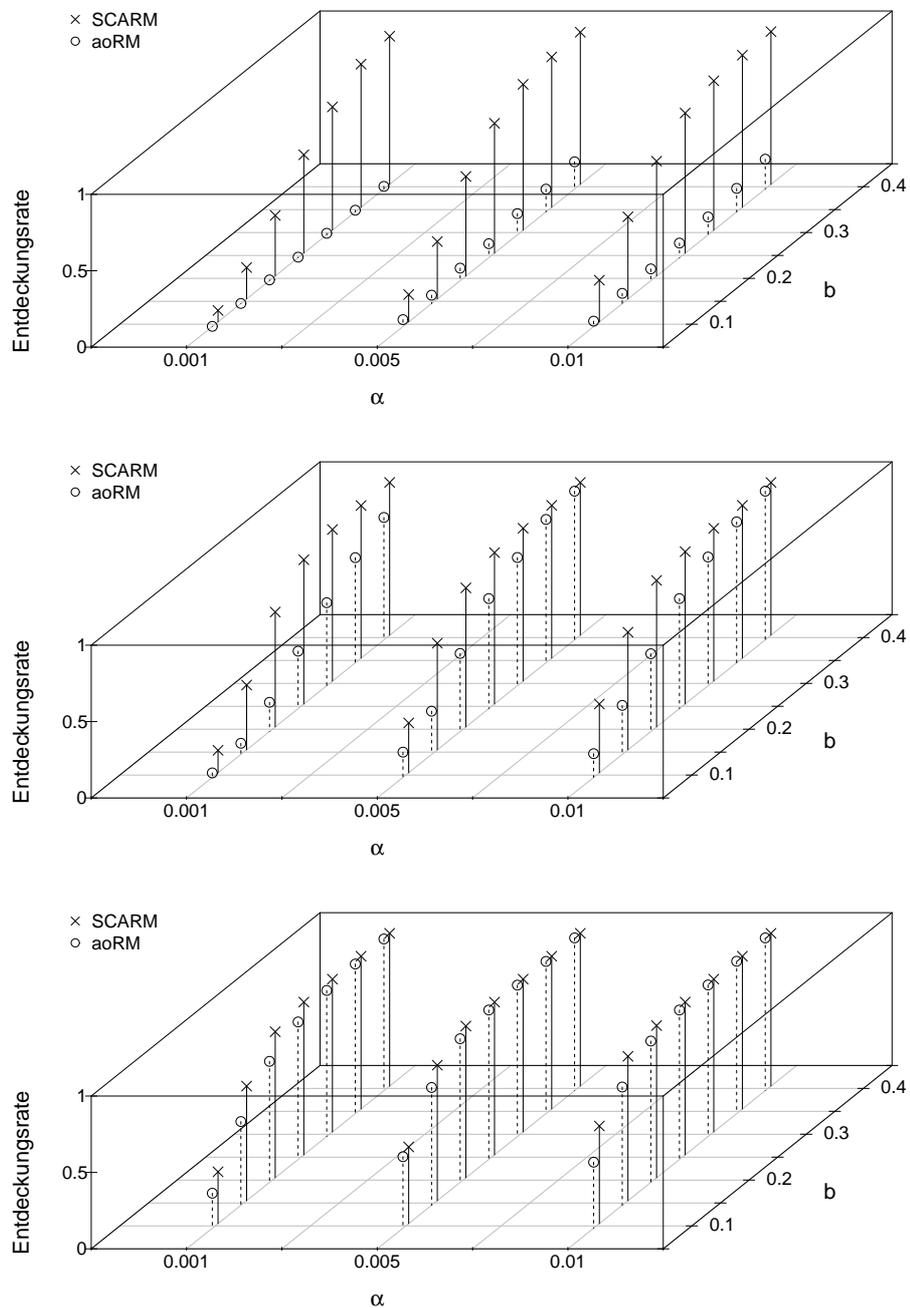


Abbildung 4.4: Entdeckungsraten des aoRM- und SCARM-Tests für verschieden hohe Trendwechsel mit $(\ell, r) = (20, 20)$ (oben), $(\ell, r) = (40, 20)$ (Mitte) und $(\ell, r) = (200, 20)$ (unten).

falls $\ell = r = 20$, vgl. Abbildung 4.4, oben. Selbst deutliche Trendwechsel werden vom aoRM selten entdeckt. Demgegenüber beträgt die Entdeckungsrate des SCARM für starke Änderungen auch bei einem kleinen Signifikanzniveau annähernd eins. Für $(\ell, r) = (40, 20)$ steigt die Güte beider Verfahren, siehe Abbildung 4.4, Mitte. Insbesondere der aoRM liefert dann wesentlich höhere Entdeckungsraten. Der SCARM ist jedoch weiterhin durchweg besser, insbesondere bei kleinem Signifikanzniveau und geringfügigen Trendänderungen. Für $\ell = 200$ und $r = 20$ erhöht sich die Güte beider Verfahren weiter, vgl. Abbildung 4.4, unten. Für deutlichere Trendwechsel ergeben sich vergleichbare Entdeckungsraten, für geringfügige Änderungen liefert der SCARM jedoch weiterhin bessere Resultate. Für größere Fensterbreiten, d.h. für $(\ell, r) = (40, 40), (80, 40), (200, 40)$, sind die Ergebnisse ähnlich, vgl. Abbildung C.10 in Anhang C. Dabei führt eine größere rechte Fensterbreite, hier $r = 40$, insbesondere für den aoRM erwartungsgemäß zu wesentlich höheren Entdeckungsraten. Der SCARM liefert jedoch weiterhin bessere Resultate, wenn ℓ und r ähnlich groß sind. Ist ℓ groß im Verhältnis zu r , zeigen beide Verfahren ähnlich gute Ergebnisse. Für $(\ell, r) = (200, 40)$ sind die Entdeckungsraten beider Verfahren sogar in jeder Konstellation annähernd eins, vgl. Abbildung C.10, unten. Es lässt sich festhalten, dass der SCARM in allen hier untersuchten Konstellationen von Trendwechseln nicht schlechter als der aoRM abschneidet. Vielmehr liefert der SCARM in einigen Situationen deutlich bessere Ergebnisse, vor allem wenn das Signifikanzniveau und die Differenz zwischen den Fensterbreiten ℓ und r klein sind. Erwartungsgemäß liefern beide Verfahren für ein kleineres Signifikanzniveau eine geringere Güte. Beim SCARM zeigt sich dieser Effekt jedoch in abgemilderter Form und hauptsächlich bei geringfügigen Trendwechseln.

4.2.2 Zeitverzögerung bei der Erkennung von Signalveränderungen

Neben der zuverlässigen Erkennung von Signalveränderungen stellt auch die möglichst schnelle Erkennung eine wichtige Anforderung an ein Fensterbreitenadaptionsverfahren dar. Um genauere Erkenntnisse über die Zeitverzögerung zu gewinnen, werden die Zeitpunkte festgehalten, an denen der SCARM und aoRM die Level Shifts und Trendwechsel erkennen. Die *mittlere Zeitverzögerung* bezeichnet dann das arithmetische Mittel der Zeitspannen vom Zeitpunkt der Signalveränderung $t = n + 1$ bis zum Zeitpunkt ihrer Entdeckung. Abbildung 4.5 zeigt die mittleren Zeitverzögerungen des aoRM und SCARM für Level Shifts (oben) und Trendwechsel (unten). Dabei werden das Signifikanzniveau $\alpha = 0.005$ und die linken Fensterbreiten $\ell = 40, 80, 200$ betrachtet; die rechte Fensterbreite beträgt in jedem Fall $r = 40$. Simulationen mit anderen Fensterbreiten r führen

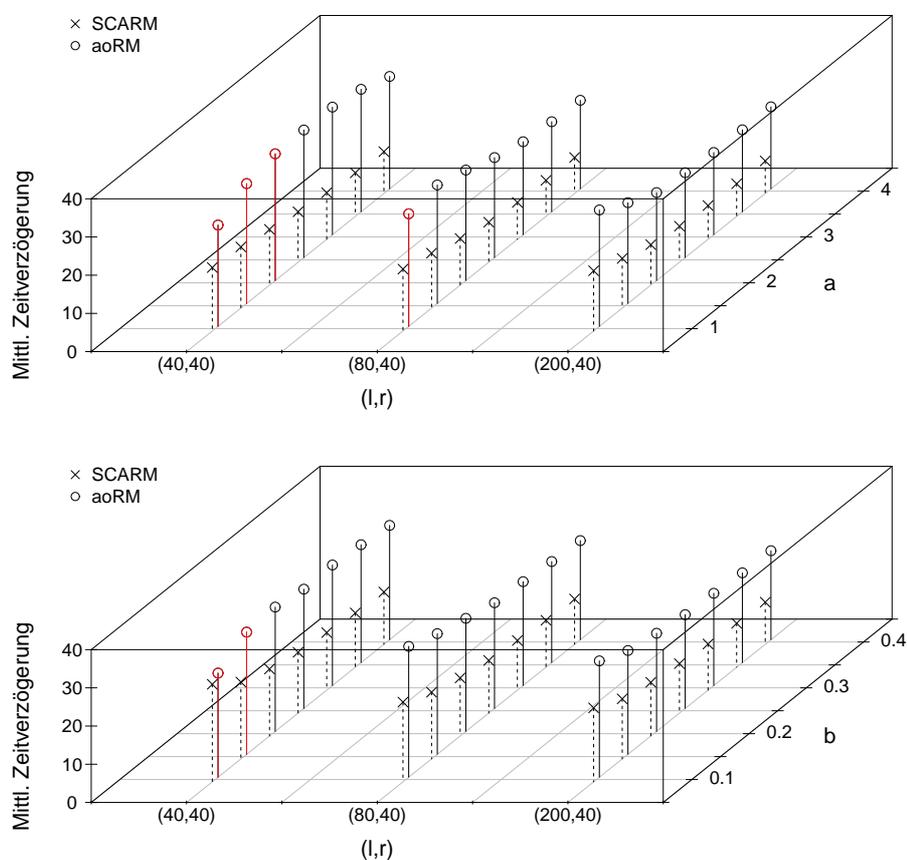


Abbildung 4.5: Mittlere Zeitverzögerungen des aoRM und SCARM für verschieden hohe Level Shifts (oben) und Trendwechsel (unten).

zu vergleichbaren Ergebnissen. In der Abbildung sind diejenigen Zeitverzögerungen rot gekennzeichnet, die mit einer Entdeckungsrate von unter 0.25 einhergehen und daher geringe Aussagekraft besitzen.

Bei Level Shifts beträgt die mittlere Zeitverzögerung des SCARM in etwa $1/4r$ bis $1/3r$ Zeitpunkte, wohingegen der aoRM diese im Mittel erst nach etwa $1/2r$ bis $3/4r$ Zeitpunkten erkennt. Beide Verfahren entdecken Level Shifts etwas schneller, wenn ℓ groß ist. Außerdem werden deutliche Level Shifts etwas eher erkannt als geringfügige. Auch bei Trendwechseln zeigt der SCARM deutlich geringere Zeitverzögerungen als der aoRM, vgl. Abbildung 4.5, unten. Der SCARM benötigt zur Erkennung von Trendwechseln etwa $1/4r$ bis $1/2r$ Zeitpunkte und der aoRM etwa $1/2r$ bis $3/4r$. Wiederum entdecken beide Ver-

fahren deutliche Änderungen schneller als geringfügige und benötigen weniger Zeit, falls ℓ groß ist.

Die Zeitverzögerung bzgl. der Erkennung von Signalveränderungen ist offenbar von den gewählten Inputargumenten abhängig. Der SCARM reagiert je nach Kombination der Inputargumente und je nach Form und Ausmaß der Signalveränderung im Mittel nach etwa $1/4r$ bis $1/2r$ Zeitpunkten auf eine Signalveränderung. Er reagiert also sobald etwa $1/4r$ bis $1/2r$ *signalveränderte Beobachtungen* (Beobachtungen, die dem neuen Signal folgen) in das rechte Zeitfenster aufgenommen worden sind. Sobald die signalveränderten Beobachtungen in das rechte Zeitfenster aufgenommen werden, verändert sich aufgrund der Lipschitz-Stetigkeit des RM die Steigung der rechten RM-Geraden entsprechend der neuen Beobachtungen. Diese Veränderung geschieht schrittweise mit jeder neu hinzukommenden Beobachtung. Daher zieht ein kleinerer kritischer Wert tendenziell eine geringere Zeitverzögerung nach sich. Da bei festem r eine größere linke Fensterbreite ℓ , d.h. eine größere Gesamtfensterbreite $n = \ell + r$, einen geringeren kritischen Wert bedeutet, führt ein größeres ℓ bei festem r zu einer schnelleren Erkennung von Signalveränderungen. Des Weiteren bewirken deutliche Signalveränderungen entsprechend deutliche Veränderungen der RM-Steigungsschätzung, weshalb sie schneller erkannt werden als geringfügige.

Auch beim aoRM wird die Zeitverzögerung durch die Inputargumente beeinflusst. Der aoRM erkennt Signalveränderungen, sobald in der rechten Teilstichprobe der Länge r die absolute Differenz der Anzahl an positiven und negativen Residuen größer als der kritische Wert ist. Bei der durchgeführten Simulationsstudie muss der aoRM je nach Kombination der Inputargumente und je nach Art und Ausmaß der Signalveränderung im Mittel $1/2r$ bis $3/4r$ signalveränderte Beobachtungen in das rechte Teilfenster aufnehmen bis er die Signalveränderung erkennt. Die Zeitverzögerung ist offenbar gering, falls der kritische Wert des aoRM-Tests klein ist. Dabei geht eine kleine Gesamtfensterbreite $n = \ell + r$ mit einem kleinen kritischen Wert einher, vgl. Abbildung C.1 in Anhang C. Es lässt sich also vermuten, dass bei einem festen r ein kleineres ℓ zu geringeren Zeitverzögerungen führt. Tatsächlich ist in Abbildung 4.5 das Gegenteil beobachtbar. Der Grund ist, dass signalveränderte Beobachtungen bei einer kleineren Gesamtfensterbreite n einen größeren Einfluss auf die Schätzung der RM-Geraden besitzen, auch weil diese zuerst am rechten Rand des Gesamtfensters auftreten. Folglich „kippen“ die signalveränderten Beobachtungen die RM-Gerade in ihre Richtung, so dass diese z.B. nach oben weist. Folglich liegen die übrigen Beobachtungen im rechten Teilfenster tendenziell unterhalb der RM-Geraden, d.h. die Anzahl an positiven und negativen Residuen wird durch das Kippen der Regressions-

geraden tendenziell ausgeglichener. Das Kippen der Regressionsgeraden erschwert also die Erkennung der Signalveränderung. Je kleiner die Gesamtfensterbreite n , desto höher der Zueffekt von signalveränderten Beobachtungen, d.h. desto eher kippt die Regressionsgerade. Daher erkennt der aoRM Signalveränderungen schneller, wenn n groß ist.

4.3 Effizienz der Signalschätzungen

Im Folgenden werden der SCARM, der aoRM sowie eine modifizierte Version des aoRM hinsichtlich der Effizienz ihrer Signalschätzungen mittels Simulationen verglichen. Die Effizienz eines Schätzers T_1 zu einem Schätzer T_2 sowie der mittlere quadratische Fehler (Mean Squared Error, MSE), die Verzerrung und die Varianz eines Schätzers sind in Anhang A definiert. Der modifizierte aoRM, nachfolgend als *mod.aoRM* bezeichnet, verwendet den Test des aoRM zur Fensterbreitenadaption, setzt jedoch wie der SCARM bei Ablehnung der Nullhypothese die Fensterbreite direkt auf den Minimalwert n_{\min} . Der Vergleich des SCARM mit dem mod.aoRM ermöglicht daher die Untersuchung des Einflusses der beiden Testverfahren auf die Güte der Signalextraktionen. Im Gegensatz dazu muss beim Vergleich des SCARM mit dem aoRM der Effekt der unterschiedlichen Mechanismen zur Fensterbreitenverkleinerung beachtet werden.

In der hier vorgestellten Simulationsstudie nutzen die Filter vergleichbare Inputargumente, so dass ein möglichst objektiver Vergleich erfolgt. Konkret verwenden der aoRM, mod.aoRM und SCARM mit $r = n_I = 20$ gleich viele Beobachtungen zur Testentscheidung, und die minimale und maximale Fensterbreite ist jeweils $n_{\min} = 7$ bzw. $n_{\max} = 200$. Als Signifikanzniveau ist jeweils $\alpha = 0.001$ gewählt. Da bei der retrospektiven Anwendung die Rechenzeit unkritisch ist, verwendet der aoRM in dieser Anwendung den linearen Suchalgorithmus zur Fensterbreitenadaption. Alle Verfahren nutzen die Restrict-to-Range-Regel (2.12).

Die Generierung der Zeitreihendaten und die Anwendung der beiden Filter erfolgt mit Hilfe der Software R (R Development Core Team, 2011). Es werden Zeitreihen betrachtet, die sich gemäß des Modells (2.1) als Realisationen von

$$X_t = \mu_t + \xi_t \quad (4.13)$$

mit $t = 1, \dots, 1000$ ergeben. Es werden drei verschiedene Signale μ_t betrachtet, die in den Abbildungen 4.6 – 4.8 auf den Seiten 74 – 76 zu sehen sind. Das erste Signal weist mehrere Level Shifts unterschiedlicher Ausprägung auf und das zweite Signal mehrere Trendwechsel. Das dritte Signal basiert auf dem Sinus, wobei sich die Amplitude nach jeder Periode

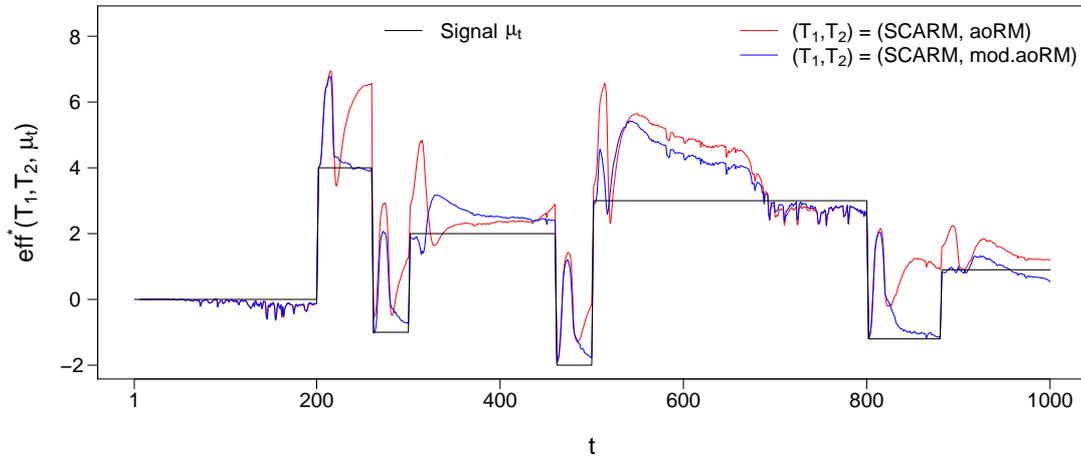


Abbildung 4.6: Signaltyp „Level Shift“ und $\text{eff}^*(T_1, T_2, \mu_t)$ für $(T_1, T_2) = (\text{SCARM}, \text{aoRM})$ und $(\text{SCARM}, \text{mod.aoRM})$.

vergrößert. Die drei Signaltypen werden nachfolgend als Typ „Level Shift“, Typ „Trendwechsel“ und Typ „Sinus“ bezeichnet. Der Rauschprozess $\xi_t = \varepsilon_t + \eta_t$ setzt sich in jedem Fall aus dem unabhängig standardnormalverteilten Fehlerprozess ε_t und einem Ausreißer generierenden Prozess η_t zusammen. Letzterer erzeugt 5% Ausreißer mit dem Wert fünf an zufälligen Zeitpunkten. Für jeden der drei Signaltypen werden jeweils $N = 1000$ Zeitreihen $(x_t)_{t=1, \dots, 1000}$ mit zufälligen Fehlern und Ausreißern erzeugt und der aoRM, mod.aoRM und SCARM auf jede Zeitreihe angewendet. Es ergeben sich demnach für jeden Filter für jeden Zeitreihentyp zu jedem Zeitpunkt $N = 1000$ Signalschätzungen. Die geschätzte Effizienz des SCARM zum aoRM für das Signal zum Zeitpunkt t ist dann

$$\text{eff}(\text{SCARM}, \text{aoRM}, \mu_t) = \frac{\text{mse}_{\mu_t}(\text{aoRM})}{\text{mse}_{\mu_t}(\text{SCARM})},$$

wobei $\text{mse}_{\mu_t}(\cdot)$ den geschätzten MSE für das Signal zum Zeitpunkt t bezeichnet. Gilt $\text{eff}(\text{SCARM}, \text{aoRM}, \mu_t) \geq 1$, so ist der SCARM effizienter als der aoRM. In den Abbildungen 4.6 – 4.8 werden die Effizienz des SCARM zum aoRM und des SCARM zum mod.aoRM für die drei Signaltypen „Level Shift“, „Trendwechsel“ und „Sinus“ in leicht abgewandelter Form präsentiert. Die modifizierte Effizienz eines Signalschätzers T_1 zu T_2 zum Zeitpunkt t ist dabei gegeben durch

$$\text{eff}^*(T_1, T_2, \mu_t) = \mu_t + \log [\text{eff}(T_1, T_2, \mu_t)].$$

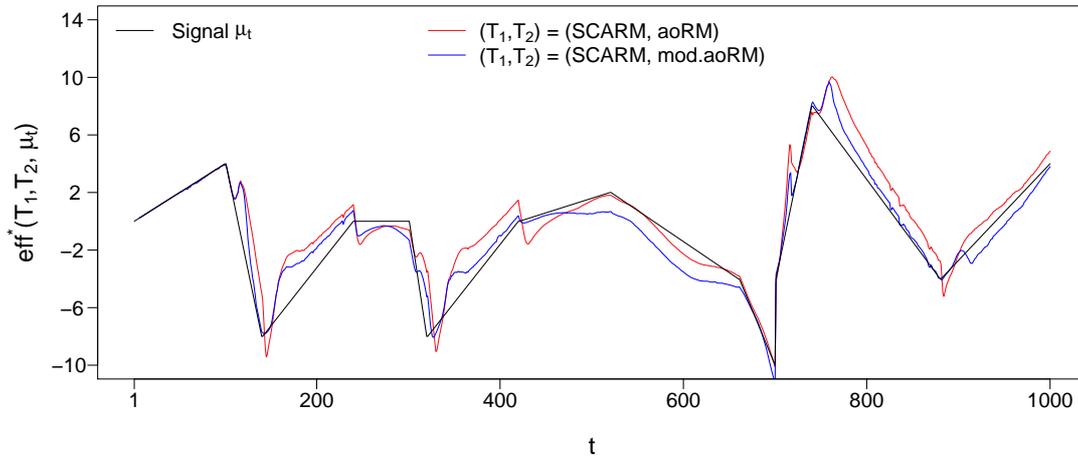


Abbildung 4.7: Signaltyp „Trendwechsel“ und $\text{eff}^*(T_1, T_2, \mu_t)$ für $(T_1, T_2) = (\text{SCARM}, \text{aoRM})$ und $(\text{SCARM}, \text{mod.aoRM})$.

Die Zeitreihe $\text{eff}^*(T_1, T_2, \mu_t)$ ermöglicht den Vergleich der Signalschätzer T_1 und T_2 zu jedem Zeitpunkt t . Zu den Zeitpunkten, an denen die Zeitreihe $\text{eff}^*(T_1, T_2, \mu_t)$ oberhalb des Signals μ_t verläuft, ist T_1 effizienter als T_2 und umgekehrt.

Bei dem Zeitreihentyp „Level Shift“ in Abbildung 4.6 verlaufen die Zeitreihen der Werte von $\text{eff}^*(\text{SCARM}, \text{aoRM}, \mu_t)$ und $\text{eff}^*(\text{SCARM}, \text{mod.aoRM}, \mu_t)$ für den Großteil der Zeitpunkte oberhalb der Signalzeitreihe, d.h. der SCARM ist hier effizienter im Vergleich zum aoRM und mod.aoRM. Besonders auffällig sind die Unterschiede nach den Level Shifts. An diesen Stellen sind der aoRM und mod.aoRM wesentlich verzerrter als der SCARM, da dieser die Signalveränderungen schneller und zuverlässiger erkennt. Dies bestätigt die Resultate aus Kapitel 4.2 bzgl. der Power und der mittleren Zeitverzögerung der Verfahren. Der aoRM und mod.aoRM sind jedoch etwas effizienter, wenn das Signal über einen längeren Zeitraum unverändert bleibt, was im Bereich der Zeitpunkte $t = 100, \dots, 200$ und $t = 700, \dots, 800$ zu erkennen ist. Der Grund ist, dass der aoRM und mod.aoRM im Bereich der Level Shifts im Mittel größere Fensterbreiten als der SCARM wählen. Folglich ergeben sich für diese Filter zu den Zeitpunkten nach den Level Shifts größere Fensterbreiten, weshalb ihre Signalschätzungen zu diesen Zeitpunkten eine geringere Varianz als die des SCARM aufweisen. Über alle Zeitpunkte hinweg gesehen liefert der SCARM dennoch die besten Resultate bzgl. des MSE. So beträgt der *Root MSE* (RMSE) des SCARM, gemittelt über alle 1000 Zeitpunkte, etwa 0.56, während der aoRM und mod.aoRM einen mittleren

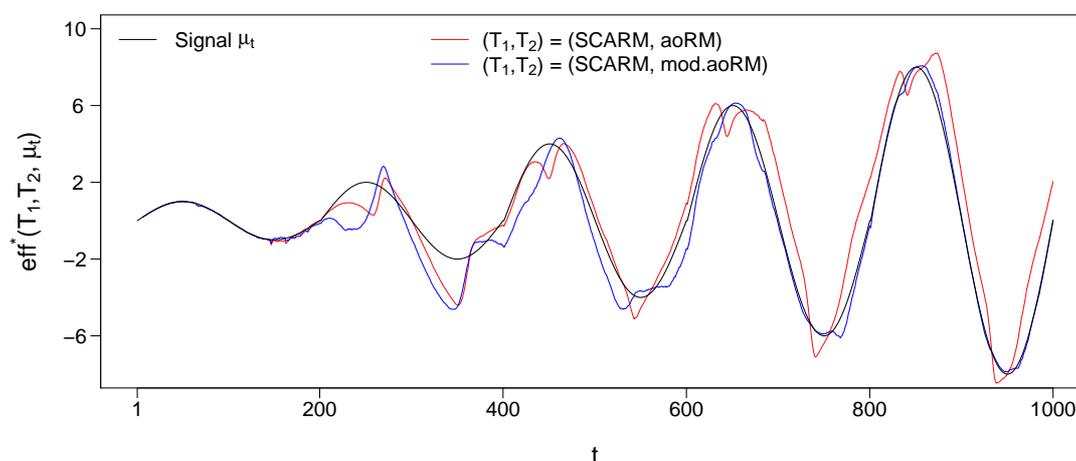


Abbildung 4.8: Signaltyp „Sinus“ und $\text{eff}^*(T_1, T_2, \mu_t)$ für $(T_1, T_2) = (\text{SCARM}, \text{aoRM})$ und $(\text{SCARM}, \text{mod.aoRM})$.

RMSE von 1.03 bzw. 0.80 aufweisen.

Für den Zeitreihentyp „Trendwechsel“ in Abbildung 4.7 sind die Resultate ähnlich. Nach deutlichen Trendwechseln liefert der SCARM effizientere Signalschätzungen, was wiederum auf die geringere Verzerrung zurückzuführen ist. Nach weniger deutlichen Trendwechseln, hier zu dem Zeitpunkten $t = 220$, $t = 420$ und $t = 520$, ist der SCARM jedoch weniger effizient als die beiden anderen Filter. Da der SCARM diese geringen Trendwechsel wesentlich häufiger erkennt als der aoRM und mod.aoRM, adaptiert er für die nachfolgenden Zeitpunkte im Mittel kleinere Fensterbreiten. Diese Trendwechsel sind jedoch weniger stark ausgeprägt, so dass das Ausmaß der Verzerrung bei größeren Fensterbreiten gering ist. Somit führen an diesen Stellen größere Fensterbreiten aufgrund der geringeren Varianz zu geringeren MSE-Werten. Über alle Zeitpunkte hinweg liefert der SCARM jedoch mit rund 0.94 den geringsten mittleren RMSE-Wert. Die entsprechenden mittleren RMSE-Werte des aoRM und mod.aoRM sind 1.41 und 0.98.

Im Gegensatz zu den Zeitreihentypen „Level Shift“ und „Trendwechsel“ verläuft beim Zeitreihentyp „Sinus“ in Abbildung 4.8 das Signal nicht stückweise linear. Die Annahme, dass sich das Signal lokal durch Geraden approximieren lässt, ist jedoch auch für diesen Signaltyp vertretbar. Die Zeitreihen $\text{eff}^*(\text{SCARM}, \text{aoRM}, \mu_t)$ und $\text{eff}^*(\text{SCARM}, \text{mod.aoRM}, \mu_t)$ verlaufen bei schwächeren Amplituden im Bereich der lokalen Extrema unterhalb des Signals, d.h. in diesen Bereichen sind der aoRM und mod.aoRM etwas effizienter als der SCARM. Bei stärkeren Amplituden ist jedoch der SCARM effizienter als der aoRM, wäh-

α	Typ „Level Shift“		
	SCARM	aoRM	mod.aoRM
0.001	0.56	1.03	0.80
0.01	0.56	0.97	0.70
0.1	0.66	0.84	0.66
α	Typ „Trendwechsel“		
	SCARM	aoRM	mod.aoRM
0.001	0.94	1.41	0.98
0.01	0.76	1.21	0.83
0.1	0.67	0.96	0.74
α	Typ „Sinus“		
	SCARM	aoRM	mod.aoRM
0.001	1.06	1.43	0.84
0.01	0.81	1.25	0.76
0.1	0.63	0.96	0.68

Tabelle 4.2: Mittlerer RMSE des SCARM, aoRM und mod.aoRM für die Signaltypen „Level Shift“, „Trendwechsel“ und „Sinus“ und für die Signifikanzniveaus $\alpha = 0.001, 0.01, 0.1$.

rend der SCARM und mod.aoRM etwa gleich effizient sind. Der mod.aoRM liefert über alle Zeitpunkte im Mittel den kleinsten RMSE von rund 0.84. Der mittlere RMSE des SCARM beträgt rund 1.06 und der mittlere RMSE des aoRM etwa 1.43.

Für die beschriebenen Effizienzvergleiche nutzen die Filter das Signifikanzniveau $\alpha = 0.001$. Da das Signifikanzniveau die Fensterbreitenadaptation und somit die Signalextraktion beeinflusst, ist der Effizienzvergleich der drei Signalextraktionsverfahren zusätzlich unter Verwendung von $\alpha = 0.01$ und $\alpha = 0.1$ durchgeführt worden. Letzteres ist das von Schettlinger et al. (2010a) empfohlene Signifikanzniveau. Die Zeitreihen $\text{eff}^*(\text{SCARM}, \text{aoRM}, \mu_t)$ und $\text{eff}^*(\text{SCARM}, \text{mod.aoRM}, \mu_t)$ für die drei Signaltypen und für $\alpha = 0.01$ und 0.1 sind in den Abbildungen C.11 – C.16 in Anhang C gezeigt. Hier ergeben sich ähnlich Resultate. Betrachte hierzu auch die Tabelle 4.2, welche die mittleren RMSE-Werte des SCARM, aoRM und mod.aoRM für alle Zeitreihentypen und alle Signifikanzniveaus auflistet. In dieser Tabelle sind für jedes α die kleinsten mittleren RMSE-Werte fett gekennzeichnet. Lediglich

für den Signaltyp „Sinus“ liefert der SCARM nicht für jeden Wert von α die besten Resultate; hier weist der mod.aoRM für $\alpha = 0.001$ und 0.01 den geringsten mittleren RMSE auf. Allerdings liefert der SCARM für jeden Signaltyp über alle α den kleinsten RMSE-Wert (unterstrichen). Dabei ergibt sich dieser für die Signaltypen „Trendwechsel“ und „Sinus“ unter Verwendung von $\alpha = 0.1$, wohingegen beim Typ „Level Shift“ der kleinste RMSE-Wert mit $\alpha = 0.001$ erreicht wird.

Insgesamt liefert der SCARM effizientere Signalschätzungen als der aoRM und mod.aoRM. Dabei ist für den SCARM ein kleines Signifikanzniveau empfehlenswert, wenn in den Zeitreihendaten häufige Signalveränderungen in Form von Level Shifts zu erwarten sind. Für Signalveränderungen in Form von Trendwechseln oder periodischen Schwingungen ist dagegen ein größeres Signifikanzniveau zu empfehlen.

4.4 Anwendung auf echte Daten

Der aoRM und SCARM werden im Folgenden retrospektiv unter Verwendung der R-Funktionen `adore.filter` und `scarm.filter` aus dem Paket `robfilter` auf eine Zeitreihe von Blutdruckmessungen und auf eine Zeitreihe von Aktienpreisen angewendet und bzgl. der Signalextraktion und Fensterbreitenadaptation verglichen.

4.4.1 Zeitreihen aus der intensivmedizinischen Patientenüberwachung

Bei der intensivmedizinischen Überwachung des Zustands des Patienten, dem sog. Online-Monitoring, werden mehrere hämodynamische Variablen wie die Herzfrequenz oder die Sauerstoffsättigung des Blutes fortlaufend in kurzen Zeitabständen gemessen. Abbildung 4.9 (a) zeigt die Zeitreihe (y_t) , $t = 1, \dots, 1000$, des sekundlich gemessenen systolischen Blutdrucks (in mmHg) eines Patienten. Die Zeitreihe (y_t) ist Teil einer umfangreichen Datenbank von aufgezeichneten Online-Monitoring-Zeitreihen hämodynamischer Variablen. Die Daten sind am Universitätsklinikum Regensburg im Rahmen des Teilprojekts C4 „Zeitreihenanalytische Methoden zur Behandlung von Online-Monitoring-Daten aus der Intensivmedizin“ des Sonderforschungsbereichs 475 „Komplexitätsreduktion in multivariaten Datenstrukturen“ der TU Dortmund erhoben worden.

Zu der Datenzeitreihe (y_t) in Abbildung 4.9 (a) sind die Signalextraktionen des aoRM (rot) und SCARM (blau) eingezeichnet. Die untere Abbildung (b) zeigt die adaptierten Fensterbreiten n_t . Für den SCARM und aoRM sind die Inputargumente $n_{\min} = 10$, $n_{\max} = 300$, $\ell_{\min} = r = 30 = n_{I_t}$ und $\alpha = 0.01$ gewählt. Des Weiteren nutzen beide Filter die Restrict-

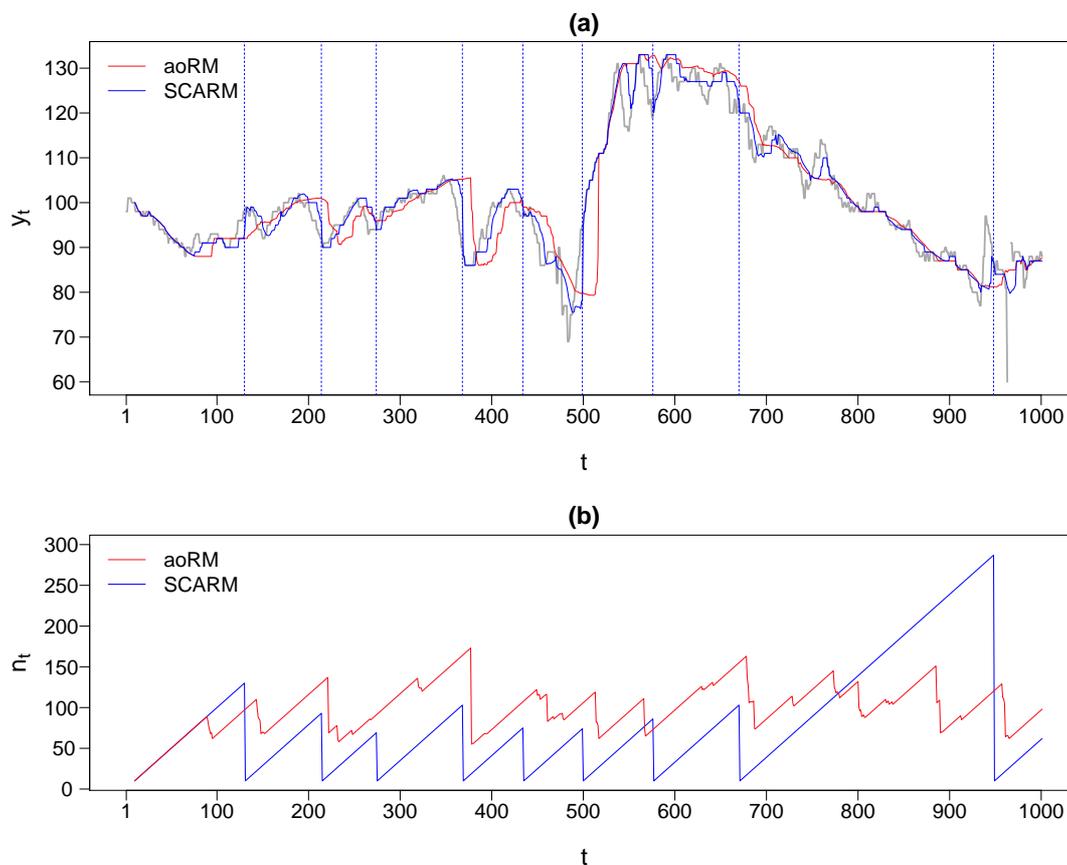


Abbildung 4.9: (a) Zeitreihe des sekundlich gemessenen systolischen Blutdrucks eines Patienten und Signalextraktion des aoRM und SCARM; (b) Adaptierte Fensterbreiten.

to-Range-Regel (2.12), und der aoRM verwendet den linearen Suchalgorithmus. Da es sich bei den Daten um ganzzahlige Beobachtungen handelt und somit Bindungen in den Fensterstichproben wahrscheinlich sind, wird für den SCARM $b = 0.1$ gesetzt um Implosionen der Q_δ^{adj} -Skalenschätzungen zu verhindern. Zudem wird für die R-Funktion `scarm.filter` die Option `autocorrelations='automatic'` gewählt, da die gemessenen Blutdruckwerte starke positive Autokorrelationen aufweisen.

Zunächst sei auf die adaptierten Fensterbreiten n_t in der unteren Abbildung 4.9 (b) hingewiesen. Die vom aoRM adaptierten Fensterbreiten (rot) schwanken größtenteils etwa im Bereich zwischen 50 und 150 und spiegeln den Verlauf der Datenzeitreihe nur unzureichend wieder. Auch Schettlinger (2009, Kap. 3.3.3) bemerkt im Zuge der Anwendung des aoRM auf intensivmedizinische Online-Monitoring-Daten, dass die adaptierten Fensterbreiten kaum Rückschlüsse auf die Daten zulassen. Im Gegensatz dazu lässt sich der Verlauf

der Datenzeitreihe anhand der vom SCARM adaptierten Fensterbreiten (blau) sehr wohl nachvollziehen. Hierzu sei auf die senkrechten blauen Linien in der oberen Abbildung (a) hingewiesen, die zu den Zeitpunkten $t^* - r/3 = t^* - 10$ eingezeichnet sind. Dabei bezeichnet t^* die Zeitpunkte, an denen der SCARM Signalveränderungen erkennt und die Fensterbreite demzufolge verkleinert. (Ein angemessener Schätzer für den Zeitpunkt der Signalveränderung ist $t^* - r/3$, vgl. S. 36.)

Anhand der vom aoRM und SCARM adaptierten Fensterbreiten lassen sich die jeweiligen Signalextraktionen in der oberen Abbildung 4.9 (a) erklären. Bis zum Zeitpunkt $t = 670$ verkleinert der SCARM mehrfach die Fensterbreite, was aufgrund des unruhigen Verlaufs der Datenzeitreihe angemessen erscheint. In diesem Zeitraum adaptiert der SCARM überwiegend kleinere Fensterbreiten als der aoRM, so dass die Signalveränderungen exakter und mit geringerer Zeitverzögerung wiedergegeben werden. Ab dem Zeitpunkt $t = 670$ wachsen dann die Fensterbreiten des SCARM bis auf den Wert $n_t = 287$ an, so dass sich glatte Signalextraktionen ergeben. Da die Datenzeitreihe in dem entsprechenden Zeitraum einen ruhigen Verlauf in Form eines Abwärtstrends aufweist, kann die Fensterbreitenadaptation als angemessen bewertet werden. Demgegenüber wachsen die Fensterbreiten des aoRM nicht an, sondern schwanken etwa um den Wert 100. Offensichtlich erfüllt der SCARM hier die Vorgabe, kleine Fensterbreiten bei Signalveränderungen und große Fensterbreiten in Phasen eines unveränderten Signalverlaufs zu wählen. Der aoRM erfüllt diese Vorgabe nicht.

4.4.2 Zeitreihen von Aktienpreisen

Ein weiteres Anwendungsbeispiel für die Signalfilter sind hochfrequent gemessene Zeitreihen von Aktienpreisen. Abbildung 4.10 (a) zeigt eine Zeitreihe (y_t), $t = 1, \dots, 800$, von Preisen verkaufter Aktien (in Euro) der Santander-Bank an der Börse von Madrid sowie die Signalextraktion des aoRM und SCARM; wie im vorherigen Kapitel 4.4.1 sind die vom SCARM erkannten Signalveränderungen durch senkrechte blaue Linien gekennzeichnet. Darunter sind in Abbildung 4.10 (b) die vom aoRM und SCARM adaptierten Fensterbreiten zu sehen. Die Daten sind der Webseite <http://www.tickdata.com/> am 21.03.2012 entnommen worden. Eine Beobachtung y_t entspricht dem mittleren Aktienpreis der Transaktionen der vergangenen fünf Sekunden, d.h. die Messfrequenz der Daten beträgt eine Beobachtung pro fünf Sekunden. In der Zeitreihe (y_t) fehlen 71 Werte, da in den entsprechenden fünf Sekunden keine Transaktion stattgefunden hat. Die fehlenden Werte treten jedoch vereinzelt auf, so dass sich keine längeren Lücken ergeben. Für die Filter sind hier

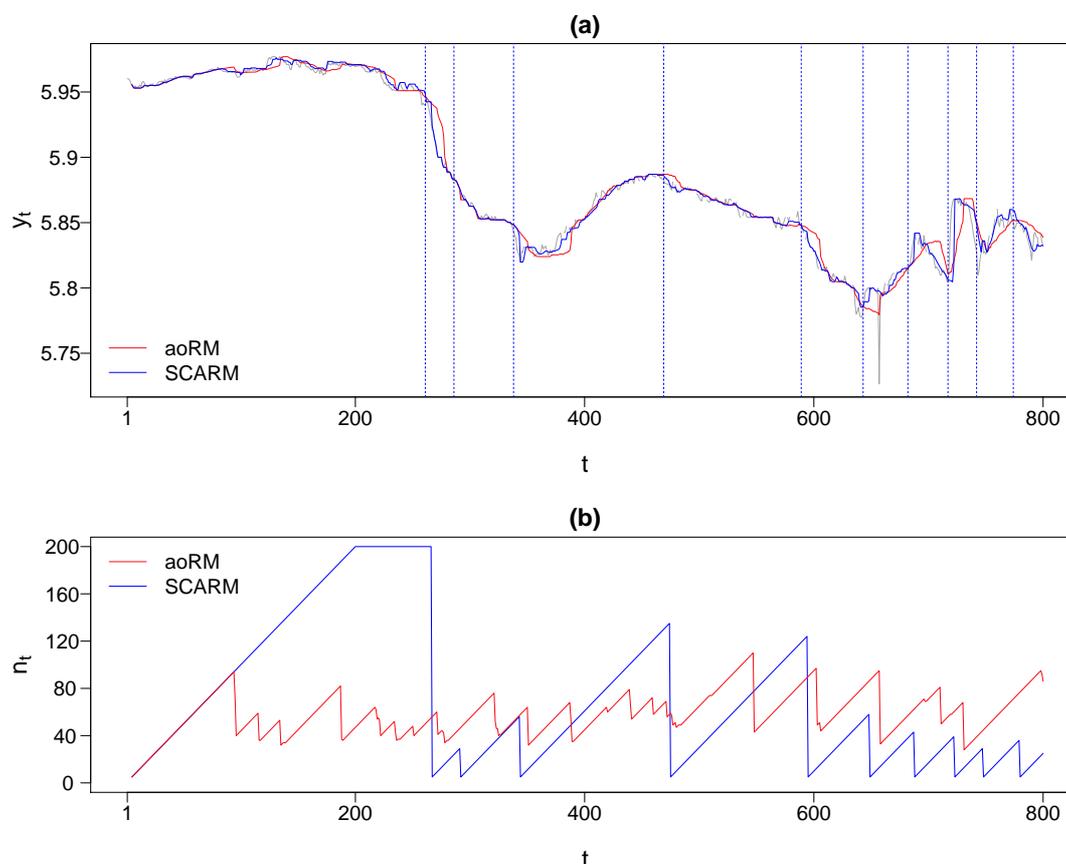


Abbildung 4.10: (a) Zeitreihe von Aktienpreisen und Signalextraktionen des aoRM und SCARM; (b) Adaptierte Fensterbreiten.

die Inputargumente $n_{\min} = 5$, $n_{\max} = 200$, $r = \ell_{\min} = 15 = n_I$ gewählt. Der aoRM- und SCARM-Test verwenden das Signifikanzniveau $\alpha = 0.01$ und beide Filter nutzen die Restrict-to-Range-Regel (2.12). Der aoRM verwendet den linearen Suchalgorithmus, und der SCARM nutzt die Option `autocorrelations='automatic'`.

Auch für diesen Datentyp lassen die vom aoRM adaptierten Fensterbreiten in Abbildung 4.10 (b) kaum Rückschlüsse auf den Verlauf der Datenzeitreihe zu. Weder wachsen die adaptierten Fensterbreiten bei unverändertem Verlauf der Datenzeitreihe kontinuierlich an, noch werden die Fensterbreiten bei Signalveränderungen, wie z.B. bei dem deutlichen Level Shift gegen $t = 250$, entsprechend klein gewählt. Vielmehr schwanken die vom aoRM adaptierten Fensterbreiten größtenteils zwischen etwa $n_t = 40$ und 80 . Im Gegensatz dazu lassen die vom SCARM adaptierten Fensterbreiten Rückschlüsse auf den Verlauf der Datenzeitreihe zu. So wachsen die Fensterbreiten des SCARM in ruhigen Phasen konti-

nuierlich an, so dass dieser nach einiger Zeit größere Fensterbreiten nutzt als der aoRM. Dennoch adaptiert der SCARM in Phasen eines unruhigen Signalverlaufs kleinere Fensterbreiten als der aoRM, vgl. den Zeitraum von $t = 600$ bis $t = 800$. Der SCARM liefert somit glatte Signalschätzungen in Phasen eines ruhigen Signalverlaufs und exakte Signalschätzungen bei Signalveränderungen. Im Gegensatz dazu kann der aoRM offenbar auch in dieser Anwendung nicht beide Zielvorgaben gleichzeitig erfüllen.

4.5 Zusammenfassung der Vergleichsstudie und Diskussion

Bezüglich der Test Resistances als Robustheitskriterien bieten der Test des aoRM und des SCARM ein vergleichbares Maß an Robustheit. Die Resistances des SCARM-Tests sind unabhängig vom Signifikanzniveau, wohingegen der Test des aoRM bei kleinem Signifikanzniveau eine hohe RR, aber eine geringe RA aufweist. Dies kann für Anwendungen von Vorteil sein, bei denen die RR wichtiger ist als die RA oder umgekehrt.

Hinsichtlich der Erkennung veränderter Signale ist der SCARM-Test gegenüber dem aoRM-Test im Vorteil. Der SCARM erkennt Signalveränderungen in Form von Level Shifts oder Trendwechselln durchweg schneller als der aoRM – grob gesagt etwa doppelt so schnell. Zudem weist der SCARM-Test insbesondere bei Trendwechselln und für kleine Signifikanzniveaus eine deutlich höhere Güte auf als der Test des aoRM. Bzgl. der Güte zur Erkennung von geringfügigen Level Shifts ist der aoRM-Test dem SCARM-Test allerdings überlegen, falls ℓ groß ist im Vergleich zu r . Ein Hybridverfahren, das je nach Fensterbreite den Test des aoRM oder SCARM zur Fensterbreitenadaption verwendet, stellt daher einen möglichen Ansatz zur Forschung dar.

Die Vorteile des SCARM bei der Erkennung von Signalveränderungen spiegeln sich im Vergleich der Effizienz der Signalschätzungen wieder. So liefert der SCARM für die drei betrachteten Signaltypen „Level Shift“, „Trendwechsel“ und „Sinus“ insgesamt die besten Resultate im Vergleich zum aoRM und zum mod.aoRM. Insbesondere für Zeitreihen mit Level Shifts und Trendwechselln ist der SCARM dem aoRM und dem mod.aoRM überlegen, da der SCARM das Signal nach einer Veränderung mit wesentlich geringerer Verzerrung schätzt. Die Resultate der Simulationsstudie sind allerdings streng genommen nur für normalverteilte Fehler gültig. Da der aoRM asymptotisch verteilungsfrei ist, liefert dieser für andere Fehlerverteilungen möglicherweise bessere Resultate. Ein umfassender Vergleich des SCARM und aoRM für verschiedene Fehlerarten steht jedoch noch aus.

Die Qualität der Signalextraktion der hier vorgestellten adaptiven RM-basierten Filter wird durch die Qualität der Fensterbreitenadaption bestimmt. Diese hängt wiederum maß-

geblich von dem Prinzip der Fensterbreitenverkleinerung sowie von der Güte des Tests ab. Dabei bewirkt das vom SCARM verwendete Adaptionsprinzip der direkten Verkleinerung der Fensterbreite auf den Minimalwert eine deutlich genauere Wiedergabe von Signalveränderungen. Darüber hinaus bietet der SCARM-Test im Gegensatz zum aoRM-Test auch bei einem kleinen Signifikanzniveau eine hohe Güte. Die Anwendung des SCARM auf Zeitreihen von Blutdruckmessungen und Aktienpreisen zeigt die Überlegenheit des SCARM gegenüber dem aoRM und bestätigt somit diese Resultate. So lässt sich anhand der vom SCARM adaptierten Fensterbreiten der Verlauf der Datenzeitreihe nachvollziehen – für die Fensterbreiten des aoRM ist dies nicht der Fall. Aufgrund der guten Fensterbreitenadaptation gibt der SCARM Signalveränderungen exakt und mit geringer Zeitverzögerung wieder und liefert glatte Signalextraktionen, wenn die Datenzeitreihe einen ruhigen Verlauf aufweist. Im Gegensatz dazu kann der aoRM diese beiden Ziele nicht gleichzeitig erfüllen.

Auf S. 24 wird die schlechte Fensterbreitenadaptation des aoRM-basierten multivariaten aoTRM-LS diskutiert. Bereits bei moderater Dimension der Zeitreihe adaptiert der aoTRM-LS überwiegend sehr kleine Fensterbreiten nahe dem Minimalwert. Ein Grund hierfür ist, dass der aoRM nur bei entsprechend hohem Signifikanzniveau eine zufriedenstellende Güte liefert. Ein hohes Signifikanzniveau führt allerdings häufig zu Fehlern 1. Art und somit zu unnötigen Verkleinerungen der Fensterbreite. Für Tests zur Fensterbreitenadaptation sollte das Signifikanzniveau demnach klein gewählt werden, weshalb eine hohe Güte bei kleinem Signifikanzniveau – wie sie der SCARM-Test bietet – von hoher Bedeutung ist. Daher stellt der SCARM-Test im Vergleich zum Test des aoRM eine geeignetere Basis für einen multivariaten Filter mit adaptiver Fensterbreitenwahl dar.

Zur Verbesserung der Fensterbreitenadaptation des aoTRM-LS wird von Borowski et al. (2009) zudem die Bildung von Blöcken abhängiger Variablen vorgeschlagen, vgl. S. 24. Dieser Ansatz erfordert allerdings Vorwissen über die Abhängigkeiten zwischen den Variablen sowie eine feste Abhängigkeitsstruktur. Wünschenswert ist daher eine automatische, zeitabhängige Blockbildung. Hierfür wird ein Verfahren benötigt, das die Zusammenhänge in multivariaten Datenströmen fortlaufend in Echtzeit anzeigt. Im folgenden Kapitel 5 wird die *Similar Slope Monitoring-* (SSM) Prozedur entwickelt, die auf dem SCARM basiert und zur Echtzeit-Überwachung der Zusammenhänge dient. Dabei liefert der SCARM zu jedem Zeitpunkt eine Trendschätzung in einem Zeitfenster, in dem die Annahme der lokalen Linearität gerechtfertigt ist. Ein Vergleich der aktuellen Trends zweier Zeitreihen gibt somit Aufschluss über den aktuellen Zusammenhang zwischen den beiden Variablen.

5 Echtzeit-Überwachung der Zusammenhänge in multivariaten Zeitreihen

Dieses Kapitel befasst sich mit Verfahren zur fortlaufenden Überwachung der aktuell vorliegenden Abhängigkeiten bzw. Zusammenhänge zwischen den Komponenten einer multivariaten Zeitreihe. Ein mögliches Anwendungsgebiet ist die intensivmedizinische Patientenüberwachung, bei der eine Reihe von Variablen wie dem systolischen Blutdruck und dem Puls fortlaufend sekundlich gemessen werden. Hier führen pathophysiologische Veränderungen des Patientenzustands in vielen Fällen zu veränderten Zusammenhängen zwischen den Vitalparametern (Gather et al., 2002). Auch bei der Überwachung von industriellen Produktionsprozessen können veränderte Zusammenhänge zwischen den überwachten Variablen auf Fehler in der Fertigung hinweisen. Des Weiteren ist bekannt, dass ökonomische Veränderungen, insbesondere Finanzkrisen, zu veränderten Abhängigkeitsstrukturen führen. In all diesen Fällen kann die Echtzeit-Überwachung der Zusammenhänge dazu dienen, Veränderungen rechtzeitig zu erkennen und entsprechende Maßnahmen oder Entscheidungen zu treffen.

Bei den in dieser Arbeit betrachteten multivariaten Datenströmen handelt es sich um hochfrequent gemessene, verrauschte und mit Ausreißern kontaminierte instationäre Zeitreihen. Aufgrund der schwierigen Datenlage und um die Anwendbarkeit in Echtzeit zu gewährleisten muss ein Verfahren zur Überwachung der Zusammenhänge eine Reihe von Anforderungen erfüllen:

1. Das Verfahren soll möglichst geringe Annahmen benötigen. Insbesondere die Stationarität der Zeitreihen stellt aufgrund möglicherweise wechselnder Trends und Variabilitäten keine vertretbare Annahme dar.
2. Das Verfahren soll einen geringen Rechenaufwand erfordern um in Echtzeit einsetzbar sein.
3. Das Verfahren soll robust gegenüber Ausreißern sein.
4. Der Verfahren soll mit fehlenden Messwerten umgehen können.

Im folgenden Kapitel 5.1 wird ein Überblick über bestehende Methoden gegeben, die zur Echtzeit-Überwachung der Abhängigkeiten bzw. Zusammenhänge in multivariaten Zeitreihen genutzt werden können. Anschließend wird in Kapitel 5.2 ein auf dem SCARM aus Kapitel 3 basierendes Verfahren vorgestellt, das alle genannten Anforderungen erfüllt.

5.1 Überblick über bestehende Methoden

Aktuell bestehen nur wenige Ansätze und Verfahren um Abhängigkeiten bzw. Zusammenhänge in multivariaten Zeitreihen in Echtzeit zu überwachen. Ein Test auf Änderungen der Kovarianzstruktur multivariater Zeitreihen wird von Aue et al. (2009) vorgeschlagen. Für eine Anwendung des Tests in Echtzeit müsste allerdings das globale Mittel der Zeitreihe bekannt sein. Zudem ist die Testprozedur nicht robust gegenüber Ausreißern. Wied (2009) und Wied und Galeano (2013) entwickeln Prozeduren, bei denen auf konstante Korrelationen zwischen zwei Zeitreihen getestet wird, siehe auch Arnold et al. (2010). Diese Testprozeduren ließen sich zwar in Echtzeit anwenden, sind jedoch nicht robust gegenüber Ausreißern.

Auf dem Gebiet der Kontrollkarten bestehen erst seit einigen Jahren Verfahren zur Überwachung der Kreuzkorrelationen bzw. -kovarianzen in multivariaten Zeitreihen, siehe z.B. Bodnar und Schmid (2007); Chan und Zhang (2001); Sliwa und Schmid (2005a,b). Diese Verfahren stellen jedoch häufig starke Annahmen, sind nicht robust gegenüber Ausreißern und verlangen die Spezifikation eines In-Control-Bereichs bzw. -Prozesses.

Lanius (2005) untersucht in ihrer Dissertation die Eignung der Faktoren- und Hauptkomponentenanalyse zur Echtzeit-Erkennung von Veränderungen der Abhängigkeitsstruktur. Für die lokale Durchführung einer Hauptkomponentenanalyse in gleitenden Zeitfenstern (Kano et al., 2001; Li et al., 2000) ist die Mittelwertstationarität der Zeitreihen erforderlich, so dass dieser Ansatz ausscheidet (Lanius, 2005, Kap. 4.2.2). Eine fensterbasierte dynamische Hauptkomponentenanalyse nach Brillinger (1981) ist ebenfalls nicht geeignet zur Echtzeit-Anwendung, da diese große Stichproben benötigt und nichtkausale Filter verwendet. Auch die Anpassung statischer oder dynamischer Faktormodelle (Peña und Box, 1987) in gleitenden Zeitfenstern ist aufgrund der Vielzahl an möglichen Modelltypen und -ordnungen sowie der schwierigen Interpretation ungeeignet für die Echtzeit-Überwachung der Abhängigkeitsstruktur instationärer multivariater Zeitreihen.

Im Bereich des Data Minings bestehen einige Methoden zum Auffinden von Zusammenhängen zwischen Komponenten multivariater Datenströme (Bulut und Singh, 2005; Cole et al., 2005; Jiang et al., 2009; Liu und Ferhatosmanoglu, 2003; Zhu und Shasha, 2002). Diese Verfahren dienen allerdings dem Auffinden lokaler, starker Zusammenhänge in hochdimensionalen Datenströmen und sind nicht robust gegenüber Ausreißern. Papadimitriou et al. (2006) und Idé et al. (2007) entwickeln Maßzahlen (*Scores*) für die lokale Ähnlichkeit von Zeitreihen, mit denen sich auch komplexere Zusammenhänge, z.B. durch lokal ähnliche Muster, erfassen lassen. Die Autoren bezeichnen die Scores als robust in dem Sinne, dass

diese über die Zeit einen stabilen Verlauf zeigen. Die Robustheit gegenüber Ausreißern ist jedoch nicht gegeben. Methoden zum Clustern von Datenströmen werden vorgeschlagen von Beringer und Hüllermeier (2006); Yeh et al. (2007); Dai et al. (2006); Rodrigues et al. (2008); Yang (2003). Diese Verfahren sind jedoch nicht robust gegenüber Ausreißern und stellen zumeist starke Annahmen an die Daten.

Lanius (2005, Kap. 5.3.1) schlägt eine lokale Clusteranalyse vor, bei der lokal ähnliche Komponenten der multivariaten Zeitreihe gruppiert werden. Die Bewertung der Ähnlichkeit zwischen den Komponenten erfolgt dabei anhand der Distanz zwischen den letzten vom robusten multivariaten TRM-LS-Filter (vgl. Kapitel 2.2.1) gelieferten Steigungsschätzungen. Veränderliche Variabilitäten in den multivariaten Zeitreihen bleiben bei dieser Methodik jedoch unberücksichtigt und können die Gruppierung erheblich beeinflussen.

Nach Sichtung der Literatur ist dem Autor keine bestehende Methodik bekannt, die dem Anforderungskatalog auf S. 84 vollständig genügt. Der von Lanius (2005) verfolgte Ansatz, die Ähnlichkeit von Zeitreihen anhand robust geschätzter Steigungen zu bewerten, ist allerdings vielversprechend. Im Folgenden wird daher eine neue auf dem SCARM basierende Methodik vorgestellt, welche diesen Ansatz aufgreift, jedoch im Gegensatz zu der von Lanius vorgeschlagenen Prozedur die lokalen Variabilitäten berücksichtigt.

5.2 Das SCARM-basierte Similar Slope Monitoring-Verfahren

Die hier vorgestellte Methode zur Echtzeit-Überwachung der Zusammenhänge ist in Zusammenarbeit mit Dipl.-Stat. Dennis Busse entwickelt worden, vgl. Busse (2012). Das Verfahren vergleicht zu jedem Zeitpunkt die aktuellen Steigungen in den einzelnen Komponenten der multivariaten Zeitreihe, weshalb es als *Similar Slope Monitoring* (SSM) bezeichnet ist.

5.2.1 Die SSM-Statistik zum Vergleich aktueller Trends

Wie die multivariaten Filter aus Kapitel 2.2 gründet das SSM-Verfahren auf der Annahme, dass der K -variaten Zeitreihe $(\mathbf{y}_t)_{t \in \mathbb{Z}}$ mit $\mathbf{y}_t = (y_t(1), \dots, y_t(K))' \in \mathbb{R}^K$ ein unbekanntes Signal $\boldsymbol{\mu}_t$ unterliegt, das von Rauschen und Ausreißern überlagert ist:

$$\mathbf{Y}_t = \boldsymbol{\mu}_t + \boldsymbol{\varepsilon}_t + \boldsymbol{\eta}_t, \quad (5.1)$$

vgl. (2.13). Dabei werden die Fehler $\boldsymbol{\varepsilon}_t \in \mathbb{R}^K$ als Prozess von zeitlich unabhängigen Zufallsvariablen mit $E(\boldsymbol{\varepsilon}_t) = \mathbf{0}$ und $\text{Cov}(\boldsymbol{\varepsilon}_t) = \boldsymbol{\Sigma}_t \in \mathbb{R}^{K \times K}$ angenommen. Kreuzkorrelationen zwischen den Fehlerkomponenten sind möglich, d.h. $\text{Cov}[\varepsilon_t(i), \varepsilon_t(j)] =: \sigma_t(i, j)$, $1 \leq i < j \leq K$,

ist nicht notwendigerweise gleich null. Die Fehlervarianz der k -ten Komponente sei mit $\sigma_t^2(k) = \sigma_t(k, k)$ bezeichnet, $k = 1, \dots, K$. Die Kovarianzmatrix Σ_t kann sich über die Zeit verändern; es wird jedoch angenommen, dass Veränderungen nicht abrupt auftreten, so dass Σ_t lokal als annähernd konstant angesehen werden kann. Weiterhin sei mit $\boldsymbol{\eta}_t$ ein Ausreißer generierender Prozess bezeichnet.

Das SSM-Verfahren stellt an jede univariate Komponente der multivariaten Zeitreihe zusätzlich die Annahme der lokalen Linearität des unterliegenden Signals:

$$\mu_{t-n_t(k)+s}(k) = \mu_t(k) + \beta_t(k) \cdot (s - n_t(k)) \quad (5.2)$$

$$\text{bzw. } Y_{t-n_t(k)+s}(k) = \mu_t(k) + \beta_t(k) \cdot (s - n_t(k)) + \varepsilon_{t-n_t(k)+s}(k) + \eta_{t-n_t(k)+s}(k)$$

mit $k = 1, \dots, K$ und $s = 1, \dots, n_t(k)$, wobei $n_t(k)$ eine zum Zeitpunkt t gültige Fensterbreite bezeichnet. Weiterhin ist $\mu_{t-n_t(k)+s}(k)$ das Signal der k -ten Komponente zum Zeitpunkt $t - n_t(k) + s$ und $\beta_t(k)$ die Steigung der k -ten Komponente im Zeitfenster, vgl. (2.14). Dabei gilt der Verlauf zweier Komponenten i und j , $i \neq j$, der multivariaten Zeitreihe zum Zeitpunkt t genau dann als gleich gerichtet, wenn $\beta_t(i) = \beta_t(j)$. Dies motiviert die Betrachtung der SSM-Statistik

$$S_t(i, j) := S_t(\mathbf{Y}_{t, n_t(i)}, \mathbf{Y}_{t, n_t(j)}) = \frac{D_t(i, j)}{\sqrt{\widehat{\text{Var}}[D_t(i, j)]}}, \quad 1 \leq i < j \leq K, \quad (5.3)$$

wobei $D_t(i, j) := \hat{\beta}(\mathbf{Y}_{t, n_t(i)}) - \hat{\beta}(\mathbf{Y}_{t, n_t(j)})$ mit

$$\mathbf{Y}_{t, n_t(i)} := (Y_{t-n_t(i)+1}(i), \dots, Y_t(i)), \quad \mathbf{Y}_{t, n_t(j)} := (Y_{t-n_t(j)+1}(j), \dots, Y_t(j)).$$

Dabei sind $\hat{\beta}(\mathbf{Y}_{t, n_t(i)})$ und $\hat{\beta}(\mathbf{Y}_{t, n_t(j)})$ Schätzungen von $\beta_t(i)$ und $\beta_t(j)$ anhand von $\mathbf{Y}_{t, n_t(i)}$ und $\mathbf{Y}_{t, n_t(j)}$, und $\widehat{\text{Var}}[D_t(i, j)]$ ist ein Schätzer der Varianz von $D_t(i, j)$.

Die Statistik $S_t(i, j)$ vergleicht die aktuell vorliegenden Trends zweier univariater Zeitreihen bzw. die Steigung der unterliegenden lokal linearen Signale und wird als Maß für den Zusammenhang im Sinne eines ähnlichen Verlaufs der beiden Zeitreihen genutzt. Je kleiner die Realisation $s_t(i, j)$ von $S_t(i, j)$, desto ähnlicher ist der Trend in den beiden Zeitreihen, d.h. desto größer ist der Zusammenhang. Dabei ist es für die Echtzeit-Anwendung aus Gründen der Einfachheit sinnvoll, eine geeignete Schranke c_{bound} festzulegen, so dass der Verlauf zweier Zeitreihen $(y_t(i))$ und $(y_t(j))$ zum Zeitpunkt t als gleich gerichtet angesehen wird, falls

$$|s_t(i, j)| = |s_t(j, i)| \leq c_{\text{bound}}. \quad (5.4)$$

Die Wahl von c_{bound} wird in Kapitel 5.2.3 behandelt.

Wird der SCARM auf die k -te Komponente angewendet, liefert dieser zu jedem Zeitpunkt neben einer passenden Fensterbreite $n_t(k)$ und der Signalschätzung $\hat{\mu}_t(k) := \hat{\mu}(\mathbf{y}_{t,n_t(k)})$ auch eine Steigungsschätzung $\hat{\beta}_t(k) := \hat{\beta}(\mathbf{y}_{t,n_t(k)})$ mit $\mathbf{y}_{t,n_t(k)} = (y_{t-n_t(k)+1}(k), \dots, y_t(k))$. Damit ist durch

$$\hat{\mu}_{t-n_t(k)+s}(k) = \hat{\mu}_t(k) + \hat{\beta}_t(k) \cdot (s - n_t(k)), \quad s = 1, \dots, n_t(k),$$

die an die letzten $n_t(k)$ Beobachtungen angepasste RM-Gerade gegeben. Dabei sei auf die Bedeutung der vom SCARM adaptierten Fensterbreite $n_t(k)$ hingewiesen: Da der SCARM die Linearität des unterliegenden Signals testet, kann angenommen werden, dass diese Bedingung für die Fensterstichprobe der Länge $n_t(k)$ erfüllt ist. Würde $\beta_t(k)$ in einem gleitenden Zeitfenster fester Länge geschätzt werden, wäre die Annahme eines im Zeitfenster linearen Signals womöglich nicht zu jedem Zeitpunkt vertretbar. Demgegenüber liefert die komponentenweise Anwendung des SCARM zu jedem Zeitpunkt geeignete Schätzungen $\hat{\beta}_t(1), \dots, \hat{\beta}_t(K)$ und somit geeignete Schätzungen $d_t(i, j)$ von $D_t(i, j)$ für alle Kombinationen (i, j) durch $d_t(i, j) = \hat{\beta}_t(i) - \hat{\beta}_t(j)$. Im Folgenden wird die Schätzung von $\text{Var}[D_t(i, j)]$ behandelt und eine auf dem SCARM basierende Schätzmethodik entwickelt.

5.2.2 Schätzung von $\text{Var}[D_t(i, j)]$

Sei nun $\hat{\beta}_t(k) := \hat{\beta}(\mathbf{Y}_{t,n_t(k)})$, wobei $\hat{\beta}$ den RM-Steigungsschätzer bezeichne. Zudem gelte stets o.B.d.A. $1 \leq i < j \leq K$. Es gilt

$$\text{Var}[D_t(i, j)] = \text{Var}[\hat{\beta}_t(i) - \hat{\beta}_t(j)] = \text{Var}[\hat{\beta}_t(i)] + \text{Var}[\hat{\beta}_t(j)] - 2 \cdot \text{Cov}[\hat{\beta}_t(i), \hat{\beta}_t(j)]. \quad (5.5)$$

Dabei hängen $\text{Var}[\hat{\beta}_t(i)]$, $\text{Var}[\hat{\beta}_t(j)]$ und $\text{Cov}[\hat{\beta}_t(i), \hat{\beta}_t(j)]$ von den Fensterbreiten $n_t(i)$ und $n_t(j)$ ab, vgl. Satz 3.1 auf S. 30 und Satz 5.1 auf S. 89. Da der SCARM die Steigungen $\beta_t(i)$ und $\beta_t(j)$ mit dem RM in Zeitfenstern von nicht festen Längen $n_t(i)$ und $n_t(j)$ schätzt, ist also

$$\text{Var}[D_t(i, j)] = \text{Var}[D_{t+r}(i, j)], \quad r \in \mathbb{Z},$$

im Allgemeinen nicht erfüllt. Vielmehr können sich die Fensterbreiten und somit die Varianzen von $D_t(i, j)$ abrupt und deutlich ändern. Eine Schätzung von $\text{Var}[D_t(i, j)]$ anhand der letzten m Realisationen $d_{t-m+1}(i, j), \dots, d_t(i, j)$ würde demnach mit möglicherweise deutlichen Verzerrungen einhergehen und ist daher nicht zu empfehlen. Aus diesem Grund wird eine alternative Methodik zur Schätzung von $\text{Var}[D_t(i, j)]$ entwickelt, die auf den Schätzungen der Fehlervarianzen $\sigma_t^2(i)$ und $\sigma_t^2(j)$ gründet.

Für die Varianzen $\text{Var}[\hat{\beta}_t(i)]$ und $\text{Var}[\hat{\beta}_t(j)]$ liefert der SCARM Schätzungen durch

$$\widehat{\text{Var}}[\hat{\beta}_t(i)] = \hat{\sigma}_t^2(i) \cdot v_{n_t(i)} \quad \text{und} \quad \widehat{\text{Var}}[\hat{\beta}_t(j)] = \hat{\sigma}_t^2(j) \cdot v_{n_t(j)}, \quad (5.6)$$

vgl. Kapitel 3.2. Dabei wird zur Schätzung der Fehlervarianzen $\sigma_t^2(i)$ und $\sigma_t^2(j)$ der Q^{adj} -Skalenschätzer verwendet, d.h. $\hat{\sigma}_t(i) = Q_{0.5}^{\text{adj}}[\mathbf{y}_{t,n_t(i)}]$ und $\hat{\sigma}_t(j) = Q_{0.5}^{\text{adj}}[\mathbf{y}_{t,n_t(j)}]$. Die Konstanten $v_{n_t(i)}$ und $v_{n_t(j)}$ sind Approximationen der Varianz des RM-Steigungsschätzers für Fensterstichproben der Länge $n_t(i)$ bzw. $n_t(j)$ mit einem linearem Signal und unabhängig standardnormalverteilten Fehlern. Die Approximationen ergeben sich aus dem in Anhang B.1 auf Basis von Simulationen erstellten Modell $v(n) = 4.77 \cdot 10^{-7} + 17.71 \cdot n^{-3}$.

Wenn die Fehler der i -ten und j -ten Komponente unabhängig sind, so sind auch $\hat{\beta}_t(i)$ und $\hat{\beta}_t(j)$ unabhängig, so dass $\text{Cov}[\hat{\beta}_t(i), \hat{\beta}_t(j)] = 0$. In diesem Fall ergibt sich ein geeigneter Schätzer für $\text{Var}[D_t(i, j)]$ gemäß (5.5) durch die vom SCARM gelieferten Schätzungen in (5.6). Da jedoch die Annahme unabhängiger Fehler oftmals nicht realistisch ist, wird im Folgenden ein Verfahren zur Schätzung von $\text{Cov}[\hat{\beta}_t(i), \hat{\beta}_t(j)]$ bzw. $\text{Corr}[\hat{\beta}_t(i), \hat{\beta}_t(j)]$ bei kreuzkorrelierten Fehlern entwickelt. Dabei ist nachfolgend mit Kovarianz bzw. Korrelation die Kreuzkovarianz bzw. -korrelation und nicht die Autokovarianz bzw. -korrelation gemeint.

Der SCARM schätzt die Steigungen $\beta_t(i)$ und $\beta_t(j)$ mit dem RM in Zeitfenstern der Länge $n_t(i)$ und $n_t(j)$. Dabei ist nicht zwangsläufig $n_t(i) = n_t(j)$. Im Folgenden sei aber o.B.d.A. stets $n_{\min} \leq n_t(i) \leq n_t(j) \leq n_{\max}$, wobei n_{\min} und n_{\max} die vom Anwender für den SCARM festgelegte minimale und maximale Fensterbreite ist. Da der SCARM die Fensterbreiten $n_t(i)$ und $n_t(j)$ entsprechend der aktuellen Datensituation wählt, kann in beiden Fällen die Linearität des Signals gemäß (5.2) angenommen werden. In diesem Fall gilt:

Satz 5.1.

Seien $\mathbf{Y}_{t,n_t(i)} := (Y_{t-n_t(i)+1}(i), \dots, Y_t(i))$ und $\mathbf{Y}_{t,n_t(j)} := (Y_{t-n_t(j)+1}(j), \dots, Y_t(j))$ Zufallsvektoren, für welche die Linearität des Signals gemäß (5.2) erfüllt ist, d.h.

$$\begin{aligned} Y_{t-n_t(i)+r}(i) &= \mu_t(i) + \beta_t(i) \cdot [r - n_t(i)] + \xi_{t-n_t(i)+r}(i), \\ Y_{t-n_t(j)+s}(j) &= \mu_t(j) + \beta_t(j) \cdot [s - n_t(j)] + \xi_{t-n_t(j)+s}(j), \end{aligned}$$

wobei

$$\begin{aligned} \xi_{t-n_t(i)+r}(i) &:= \varepsilon_{t-n_t(i)+r}(i) + \eta_{t-n_t(i)+r}(i), \\ \xi_{t-n_t(j)+s}(j) &:= \varepsilon_{t-n_t(j)+s}(j) + \eta_{t-n_t(j)+s}(j) \end{aligned}$$

mit $r = 1, \dots, n_t(i)$ und $s = 1, \dots, n_t(j)$. Weiterhin sei $\boldsymbol{\xi}_{t,n_t(i)} := (\xi_{t-n_t(i)+1}(i), \dots, \xi_t(i))$ und $\boldsymbol{\xi}_{t,n_t(j)} := (\xi_{t-n_t(j)+1}(j), \dots, \xi_t(j))$, und $\hat{\beta}$ bezeichne den RM-Steigungsschätzer. Dann gilt

$$\text{Cov}[\hat{\beta}_t(i), \hat{\beta}_t(j)] = \text{Cov}\left(\hat{\beta}[\boldsymbol{\xi}_{t,n_t(i)}], \hat{\beta}[\boldsymbol{\xi}_{t,n_t(j)}]\right),$$

wobei $\hat{\beta}_t(i) := \hat{\beta}[\mathbf{Y}_{t,n_t(i)}]$ und $\hat{\beta}_t(j) := \hat{\beta}[\mathbf{Y}_{t,n_t(j)}]$.

Beweis.

Nach dem Beweis von Satz 3.1 auf S. 30 gilt

$$\hat{\beta}_t(i) = \beta_t(i) + \hat{\beta}[\boldsymbol{\xi}_{t,n_t(i)}] \quad \text{und} \quad \hat{\beta}_t(j) = \beta_t(j) + \hat{\beta}[\boldsymbol{\xi}_{t,n_t(j)}].$$

Da $\beta_t(i)$ und $\beta_t(j)$ Konstanten sind, folgt hieraus unmittelbar die Behauptung. \square

Satz 5.1 besagt unter anderem, dass $\text{Cov}[\hat{\beta}_t(i), \hat{\beta}_t(j)]$ unabhängig von den wahren Steigungen $\beta_t(i)$ und $\beta_t(j)$ und den wahren Niveaus $\mu_t(i)$ und $\mu_t(j)$, d.h. unabhängig von den wahren unterliegenden Signalen ist, falls diese die Linearitätsannahme erfüllen. Da in diesem Fall nach Satz 3.1 auch $\text{Var}[\hat{\beta}_t(i)]$ und $\text{Var}[\hat{\beta}_t(j)]$ unabhängig von den Signalen sind, ist auch

$$\text{Corr}[\hat{\beta}_t(i), \hat{\beta}_t(j)] = \frac{\text{Cov}[\hat{\beta}_t(i), \hat{\beta}_t(j)]}{\sqrt{\text{Var}[\hat{\beta}_t(i)]\text{Var}[\hat{\beta}_t(j)]}}$$

unabhängig von den Signalen und hängt nur von den Fehlervektoren $\boldsymbol{\xi}_{t,n_t(i)}$ und $\boldsymbol{\xi}_{t,n_t(j)}$ ab. Dabei haben nicht nur die Fehlerverteilung und die Fehlerkorrelation Einfluss auf $\text{Corr}[\hat{\beta}_t(i), \hat{\beta}_t(j)]$, sondern auch die Fensterbreiten $n_t(i)$ und $n_t(j)$. Betrachte hierzu den Fall bivariat normalverteilter Fehler mit Kreuzkorrelation eins und $n_t(i) = n_t(j)$. In diesem Fall gilt auch $\text{Corr}[\hat{\beta}_t(i), \hat{\beta}_t(j)] = 1$. Für $n_t(i) \neq n_t(j)$ ist dies jedoch offensichtlich nicht der Fall. Vielmehr tendiert die Korrelation zwischen $\hat{\beta}_t(i)$ und $\hat{\beta}_t(j)$ mit zunehmender Differenz zwischen $n_t(i)$ und $n_t(j)$ gegen null. Im Zuge der Schätzung der Korrelation $\text{Corr}[\hat{\beta}_t(i), \hat{\beta}_t(j)]$ ist also zu beachten, dass diese abhängig ist von den adaptierten Fensterbreiten $n_t(i)$ und $n_t(j)$. Das heißt, selbst wenn die Signale $\mu_t(i)$ und $\mu_t(j)$ über den gesamten Zeitverlauf linear sind und sich die Fehlerkorrelation nicht ändert, gilt nicht, dass $\text{Corr}[\hat{\beta}_t(i), \hat{\beta}_t(j)]$ konstant ist über alle Zeitpunkte t , da sich die Fensterbreiten $n_t(i)$ und $n_t(j)$ über die Zeit verändern und zum Zeitpunkt t im Allgemeinen nicht gleich sind. Demnach ist

$$\text{Corr}[\hat{\beta}_t(i), \hat{\beta}_t(j)] = \text{Corr}[\hat{\beta}_{t+r}(i), \hat{\beta}_{t+r}(j)], \quad r \in \mathbb{Z},$$

im Allgemeinen nicht erfüllt. Würde $\text{Corr}[\hat{\beta}_t(i), \hat{\beta}_t(j)]$ anhand der letzten m RM-Steigungsschätzungen $\hat{\beta}[\mathbf{y}_{t-m+1, n_{t-m+1}(i)}], \dots, \hat{\beta}[\mathbf{y}_{t, n_t(i)}]$ und $\hat{\beta}[\mathbf{y}_{t-m+1, n_{t-m+1}(j)}], \dots, \hat{\beta}[\mathbf{y}_{t, n_t(j)}]$ geschätzt werden, ist daher mit verzerrten Schätzungen zu rechnen. Im Folgenden wird eine alternative Schätzmethodik für $\text{Corr}[\hat{\beta}_t(i), \hat{\beta}_t(j)]$ entwickelt, welche die Abhängigkeit der Korrelation von den Fensterbreiten $n_t(i)$ und $n_t(j)$ beachtet.

Für die Entwicklung der hier präsentierten Methodik zur Schätzung von $\text{Corr}[\hat{\beta}_t(i), \hat{\beta}_t(j)]$ wird neben den o.g. Annahmen zusätzlich vorausgesetzt, dass $\boldsymbol{\varepsilon}_t \sim N(\mathbf{0}, \boldsymbol{\Sigma}_t)$ unabhängig und $\boldsymbol{\eta}_t = \mathbf{0}$ für alle $t \in \mathbb{Z}$. Verletzungen dieser Annahmen, deren Auswirkungen auf die Methodik und mögliche Gegenmaßnahmen werden in Kapitel 5.2.4 diskutiert.

Nachfolgend sei die Korrelation der Fehler $\varepsilon_t(i)$ und $\varepsilon_t(j)$ bezeichnet mit

$$\rho_t(i, j) := \text{Corr}[\varepsilon_t(i), \varepsilon_t(j)] = \frac{\sigma_t(i, j)}{\sigma_t(i) \cdot \sigma_t(j)}.$$

Unter der Annahme lokal konstanter Fehlerkorrelationen, d.h. $\rho_{t-n_t(i)+1}(i, j) = \rho_{t-n_t(i)+2}(i, j) = \dots = \rho_t(i, j)$, wobei o.B.d.A. $n_t(i) \leq n_t(j)$, gilt

$$\begin{aligned} \rho_t(i, j) = 0 &\Rightarrow \text{Corr}[\hat{\beta}_t(i), \hat{\beta}_t(j)] = 0, \\ \rho_t(i, j) = 1 \text{ und } n_t(i) = n_t(j) &\Rightarrow \text{Corr}[\hat{\beta}_t(i), \hat{\beta}_t(j)] = 1, \\ \rho_t(i, j) = -1 \text{ und } n_t(i) = n_t(j) &\Rightarrow \text{Corr}[\hat{\beta}_t(i), \hat{\beta}_t(j)] = -1. \end{aligned}$$

Falls $\rho_t(i, j) \notin \{-1, 0, 1\}$ und/oder $n_t(i) \neq n_t(j)$, lässt sich jedoch nicht direkt von $\rho_t(i, j)$ auf $\text{Corr}[\hat{\beta}_t(i), \hat{\beta}_t(j)]$ schließen. Das Ziel ist daher, $\text{Corr}[\hat{\beta}_t(i), \hat{\beta}_t(j)]$ über eine Funktion der Fehlerkorrelation $\rho_t(i, j)$ und der Fensterbreiten $n_t(i)$ und $n_t(j)$ darzustellen. Dabei legen Simulationen in Busse (2012, Anhang B) nahe, dass $\text{Corr}[\hat{\beta}_t(i), \hat{\beta}_t(j)]$ nicht von den genauen Werten von $n_t(i)$ und $n_t(j)$ abhängt, sondern lediglich von dem Verhältnis $r_t(i, j) = n_t(i)/n_t(j) \in (0, 1]$. Demnach folgt $\text{Corr}[\hat{\beta}_t(i), \hat{\beta}_t(j)]$ einer unbekanntem Funktion C von $\rho_t(i, j)$ und $r_t(i, j)$:

$$\text{Corr}[\hat{\beta}_t(i), \hat{\beta}_t(j)] = C[\rho_t(i, j), r_t(i, j)]$$

mit $C : [-1, 1] \times (0, 1] \mapsto [-1, 1]$. Anhand der Ergebnisse der Simulationen erstellt Busse (2012, Anhang B) ein Modell \tilde{C} für die unbekanntem Funktion C (die Koeffizienten sind auf zwei Nachkommastellen gerundet):

$$\tilde{C}[\rho, r] = 0.09\rho + 0.01\rho^3 - 1.18\rho r + 3.58\rho r^2 - 1.73\rho r^3 - 0.18\rho^3 r + 0.58\rho^3 r^2 - 0.20\rho^3 r^3. \quad (5.7)$$

Bei bekannter Fehlerkorrelation $\rho_t(i, j)$ kann die Korrelation der RM-Steigungen also anhand dieses Modells \tilde{C} approximiert werden. Ist die Fehlerkorrelation $\rho_t(i, j)$ jedoch unbekannt oder zeitveränderlich, kann $\rho_t(i, j)$ in \tilde{C} durch eine geeignete Schätzung ersetzt werden. Lanius und Gather (2010) untersuchen und vergleichen verschiedene Methoden zur Schätzung der Fehlerkovarianzmatrix Σ_t in Echtzeit für das hier angenommene Modell (5.1) und empfehlen den *orthogonalisierten Gnanadesikan-Kettenring-Schätzer* (OGK, Maronna und Zamar, 2002) auf Basis des robusten Q_n -Skalenschätzers (Rousseeuw und Croux, 1993), der in Kapitel 2.2.1 vorgestellt wird und im Folgenden mit OGK- Q_n bezeichnet sei. Der Q_n wäre auf die Residuen der K RM-Regressionsgeraden anzuwenden, die sich im Zuge der komponentenweisen Anwendung des SCARM ergeben. Obwohl durch die adaptive Fensterbreitenwahl des SCARM die Annahme eines linearen Signals vertretbar ist, bleibt die Q_n -Schätzung der Fehlerskala abhängig von der Güte der Regressionsanpassung. Zudem geschieht bei der Anwendung des SCARM bereits eine Schätzung der Standardabweichung der Fehler mit dem modellfreien Q^{adj} -Skalenschätzer. Aus diesen Gründen erweitert Busse (2012, Kap. 4.2.2) die Vergleichsstudie von Lanius und Gather (2010) um den OGK-Schätzer auf Basis des modellfreien Q^{adj} -Skalenschätzers, nachfolgend mit OGK- Q^{adj} bezeichnet. Bei einem linearen Signal ist die Güte des OGK- Q_n und OGK- Q^{adj} , gemessen anhand der logarithmierten Konditionszahl, vergleichbar. Der OGK- Q_n besitzt eine höhere Güte bei unabhängigen Fehlern oder moderaten Ausreißern, wohingegen der OGK- Q^{adj} bei korrelierten Fehlern oder großen Ausreißern eine höhere Güte aufweist. Ist das Signal jedoch nicht linear, sondern weist z.B. einen Sprung auf, verschlechtern sich die OGK- Q_n -Schätzungen. Demgegenüber bleibt die Güte des OGK- Q^{adj} -Schätzers durch Sprünge unbeeinflusst. Darüber hinaus liegen die Q^{adj} -Skalenschätzungen der Fehler nach der komponentenweisen Anwendung des SCARM bereits vor, so dass keine zusätzliche Rechenzeit benötigt wird. Aus diesen Gründen ist der OGK- Q^{adj} dem OGK- Q_n für die hier entwickelte Schätzmethodik vorzuziehen. Unter Verwendung des OGK- Q^{adj} -Schätzers ergibt sich die Schätzung von $\text{Var}[D_t(i, j)]$ gemäß (5.5) durch

$$\widehat{\text{Var}}[d_t(i, j)] = \widehat{\text{Var}}[\hat{\beta}_t(i)] + \widehat{\text{Var}}[\hat{\beta}_t(j)] - 2\widehat{\text{Corr}}[\hat{\beta}_t(i), \hat{\beta}_t(j)] \cdot \sqrt{\widehat{\text{Var}}[\hat{\beta}_t(i)]\widehat{\text{Var}}[\hat{\beta}_t(j)]} \quad (5.8)$$

mit $\widehat{\text{Var}}[\hat{\beta}_t(i)]$ und $\widehat{\text{Var}}[\hat{\beta}_t(j)]$ aus (5.6) und $\widehat{\text{Corr}}[\hat{\beta}_t(i), \hat{\beta}_t(j)] = \tilde{C}[\hat{\rho}_t(i, j), r_t(i, j)]$. Dabei ist \tilde{C} die Funktion aus (5.7) und $\hat{\rho}_t(i, j)$ die OGK- Q^{adj} -Schätzung der Fehlerkorrelation. Ist die komponentenweise Anwendung des SCARM erfolgt, liegen die Schätzungen $\widehat{\text{Var}}[\hat{\beta}_t(i)]$ für alle $i = 1, \dots, K$ vor, so dass lediglich die Fehlerkorrelationen $\rho_t(i, j)$, $1 \leq i < j \leq K$, zu schätzen sind. Der OGK- Q^{adj} nutzt hierzu in Schritt 1 des Algorithmus (vgl. S. 21f.) die be-

reits vorliegenden Q^{adj} -Schätzungen $\hat{\sigma}_t(i)$ und $\hat{\sigma}_t(j)$ der Fehlerskala. Da die Fensterbreiten $n_t(i)$ und $n_t(j)$ im Allgemeinen nicht gleich sind, stellt sich die Frage nach der Anzahl an Beobachtungen n , die in Schritt 2 des OGK-Algorithmus verwendet werden. Hier empfiehlt Busse (2012, S. 35) die Wahl von $n = \max\{n_t(i), n_t(j)\}$, so dass kein Informationsverlust durch nicht beachtete Beobachtungen besteht. Falls o.B.d.A. $n_t(i) < n_t(j)$, ist zwar die Annahme der Linearität der Signale für die Komponente i nicht gerechtfertigt. Der Q^{adj} liefert jedoch auch bei Verletzung dieser Annahme geeignete Skalenschätzungen und somit auch geeignete OGK- Q^{adj} -Schätzungen der Fehlerkorrelationen.

Im Folgenden wird untersucht, wie sich die SSM-Statistik $S_t(i, j)$ in verschiedenen Szenarien von gleich und ungleich verlaufenden Zeitreihen verhält. Diese Untersuchung erfolgt mittels der R-Funktion $\text{SSM}()$, die in Kapitel 5.2.4 vorgestellt wird.

5.2.3 Bewertung der SSM-Statistik

Die SSM-Statistik $S_t(i, j)$ dient als Maß des zum Zeitpunkt t gegebenen Zusammenhangs zwischen zwei Zeitreihen. Je kleiner der Wert von $|s_t(i, j)|$, desto ähnlicher der aktuelle Verlauf, d.h. desto größer der aktuelle Zusammenhang. Dabei ist zu beachten, dass $S_t(i, j)$ nur den Rückschluss zulässt, wie sehr sich der Verlauf der beiden Zeitreihen gleicht. Zusammenhänge im Sinne entgegengesetzter Verläufe lassen sich anhand von $S_t(i, j)$ nicht erkennen.

Im Folgenden wird das Verfahren in sechs verschiedenen Szenarien bivariater Zeitreihen angewendet, um das Verhalten von $S_t(i, j)$ bei gleichen und ungleichen Verläufen zu untersuchen. Insbesondere sollen Erkenntnisse bzgl. der Wahl einer geeigneten Schranke c_{bound} für $S_t(i, j)$ gewonnen werden, so dass der Verlauf zweier Zeitreihen zum Zeitpunkt t als gleich gerichtet angesehen wird, wenn $|s_t(i, j)| \leq c_{\text{bound}}$, vgl. (5.4). Die sechs bivariaten Zeitreihen sind in Abbildung 5.1 dargestellt. In jedem Szenario ergeben sich die simulierten Zeitreihendaten $\mathbf{y}_t \in \mathbb{R}^2$, $t = 1, \dots, 500$, aus dem Modell $\mathbf{Y}_t = \boldsymbol{\mu}_t + \boldsymbol{\varepsilon}_t + \boldsymbol{\eta}_t$ mit Signal $\boldsymbol{\mu}_t = (\mu_t(1), \mu_t(2))' \in \mathbb{R}^2$ und Fehler $\boldsymbol{\varepsilon}_t = (\varepsilon_t(1), \varepsilon_t(2))' \in \mathbb{R}^2$. Dabei ist $\boldsymbol{\varepsilon}_t$ ein Prozess unabhängiger Zufallsvariablen $\boldsymbol{\varepsilon}_t \sim N(\mathbf{0}, \boldsymbol{\Sigma}_t)$ mit $\boldsymbol{\Sigma}_t = (\sigma_t(i, j))_{i,j=1,2}$, $\sigma_t(1, 1) = \sigma_t(2, 2) = 1$ und $\sigma_t(1, 2) = \sigma_t(2, 1) = \rho = 0$ für alle $t = 1, \dots, T$. Der Fall kreuzkorrelierter Fehler, also $\rho \neq 0$, wird hier nicht aufgeführt. Hierfür ergeben sich vergleichbare Resultate (Busse, 2012, Kap. 5.1), da bei der Schätzung von $\text{Var}[D_t(i, j)]$ die Kreuzkorrelationen berücksichtigt werden. Für den Ausreißer generierenden Prozess $\boldsymbol{\eta}_t = (\eta_t(1), \eta_t(2))'$ ist $\eta_t(k) \sim 0.9 \delta_0 + 0.1 \delta_5$, $k = 1, 2$, gewählt, wobei δ_ω die Einpunktverteilung in ω bezeichnet. In den Daten sind somit etwa 10% Ausreißer mit dem Wert 5 vorhanden.

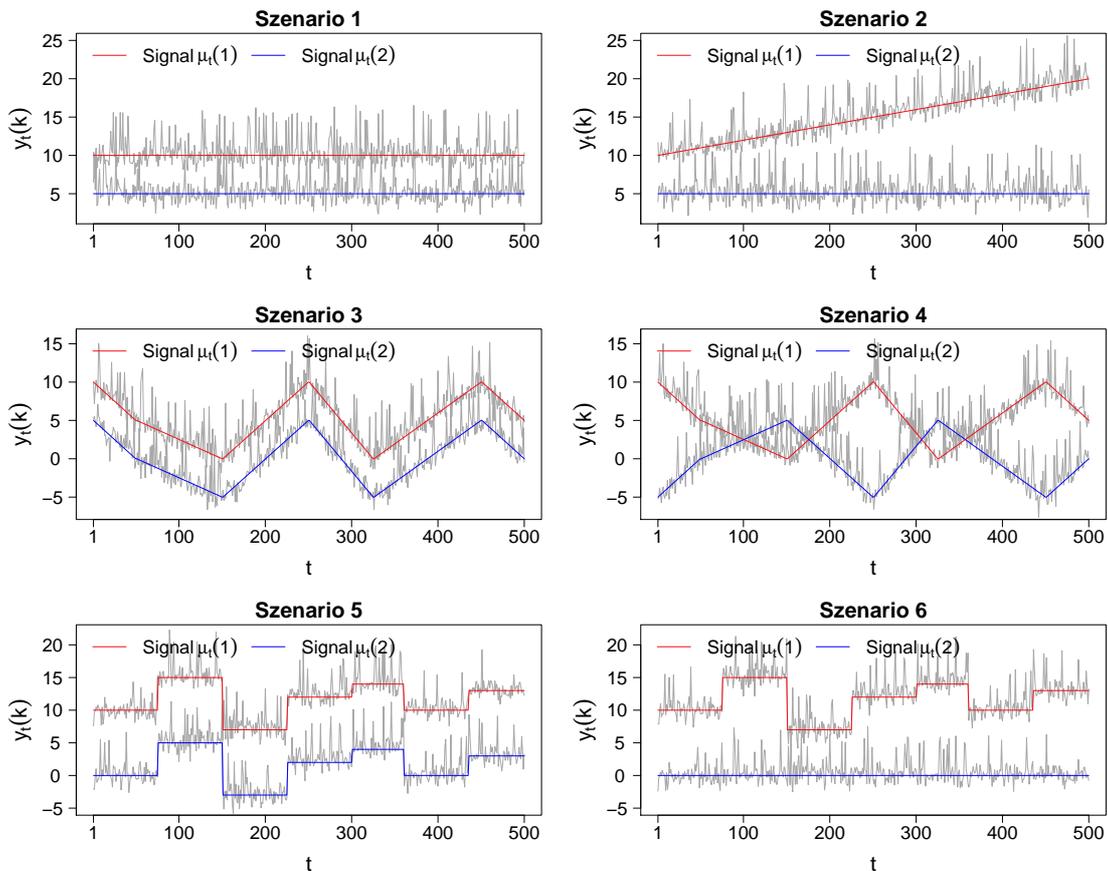


Abbildung 5.1: Sechs Szenarien von gleich und ungleich verlaufenden Zeitreihen.

In jedem der sechs Szenarien erfolgt die Anwendung des SCARM auf die beiden univariaten Zeitreihen unter Verwendung der Inputargumente $r = \ell_{\min} = 30$, $n_{\min} = 10$, $n_{\max} = 200$ und $\alpha = 0.001$. Wegen $n_{\min} = 10$ erfolgt in den sechs Anwendungen jeweils für die ersten Zeitpunkte $t = 1, \dots, 9$ keine SCARM-Signalextraktion und somit auch keine Berechnung von $s_t(i, j)$.

Szenario 1: Zwei konstante Signale

Abbildung 5.2 (a) zeigt zwei Zeitreihen, deren Signale konstant sind, d.h. $\mu_t(1) = 10$ und $\mu_t(2) = 5$ für alle $t = 1, \dots, 500$. Die Signale verlaufen somit über den gesamten Zeitraum gleich. In der unteren Abbildung (b) ist die Zeitreihe der realisierten Vergleichsstatistiken $s_t(1, 2)$ dargestellt. Hier ist die Schranke $c_{\text{bound}} = 3$ gewählt; der Bereich $[-c_{\text{bound}}, c_{\text{bound}}]$ ist durch die gestrichelten Linien gekennzeichnet. Für die Zeitpunkte, an denen $s_t(i, j) \in [-c_{\text{bound}}, c_{\text{bound}}]$ gilt, ist die bivariate Datenzeitreihe \mathbf{y}_t in der oberen Abbil-

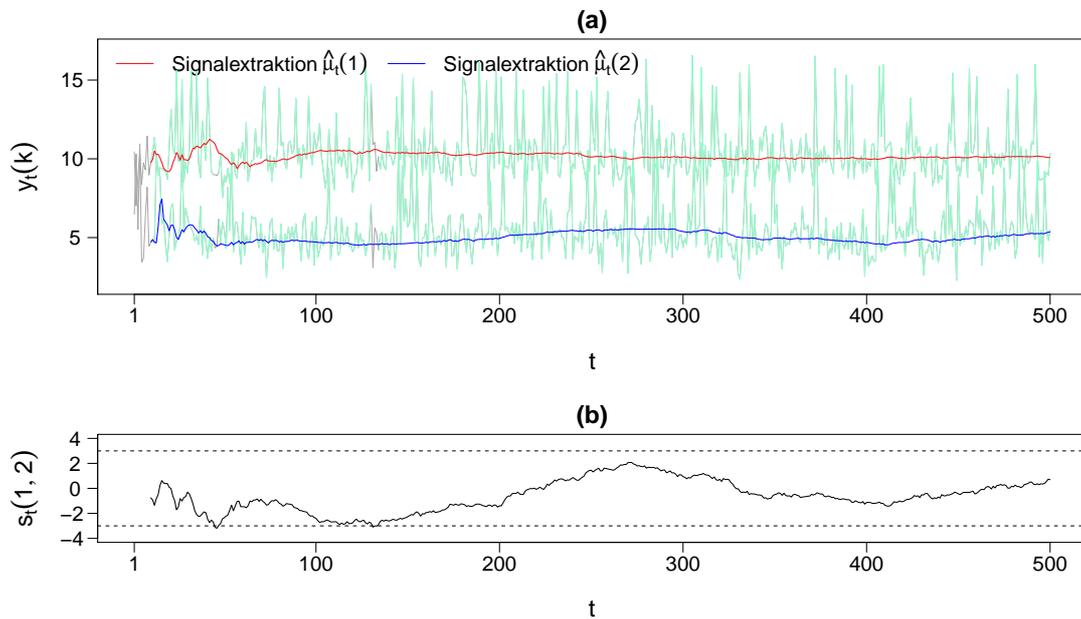


Abbildung 5.2: (a) Bivariate Zeitreihe (Szenario 1) mit SCARM-Signalextraktion und farblicher Kennzeichnung der Zusammenhänge auf Basis von $s_t(1,2)$; (b) Zeitreihe der Vergleichsstatistik $s_t(1,2)$.

Abbildung (a) grün gezeichnet, sonst grau. Diese farbliche Kennzeichnung ermöglicht eine schnelle visuelle Erfassung von (veränderten) Zusammenhängen und wäre daher auch für die Anwendung des Verfahrens in Echtzeit von Nutzen. Auch in den Abbildungen 5.3 – 5.7 zu den Szenarien 2 bis 6 wird diese Darstellungsweise verwendet. Die farbliche Kennzeichnung zeigt fast durchgehend einen Zusammenhang zwischen den beiden Zeitreihen $y_t(1)$ und $y_t(2)$ an. Ausnahmen bilden lediglich die Zeitpunkte $t = 45, 46$ und $t = 131, 132, 133$. Zu diesen Zeitpunkten liegt die Vergleichsstatistik unterhalb der unteren Schranke $-c_{\text{bound}} = -3$, wie in der unteren Abbildung (b) zu erkennen ist. Diese fehlerhaften Bewertungen stören jedoch nicht den durch die farbliche Kennzeichnung gegebenen Gesamteindruck, dass die beiden Zeitreihen einen ähnlichen Verlauf aufweisen. Es ist anzumerken, dass $y_t(1)$ und $y_t(2)$ in Abbildung 5.2 (a) zu jedem Zeitpunkt t unkorreliert sind, d.h. es besteht kein Zusammenhang im Sinne einer stochastischen Abhängigkeit. Die SSM-Methodik zeigt jedoch überwiegend einen Zusammenhang an, da die Signale durchweg gleich verlaufen.

Szenario 2: Zwei Signale mit unterschiedlicher Steigung

In Abbildung 5.3 (a) hat die Zeitreihe $y_t(2)$ ein konstantes Signal $\mu_t(2) = 5$, wohingegen die

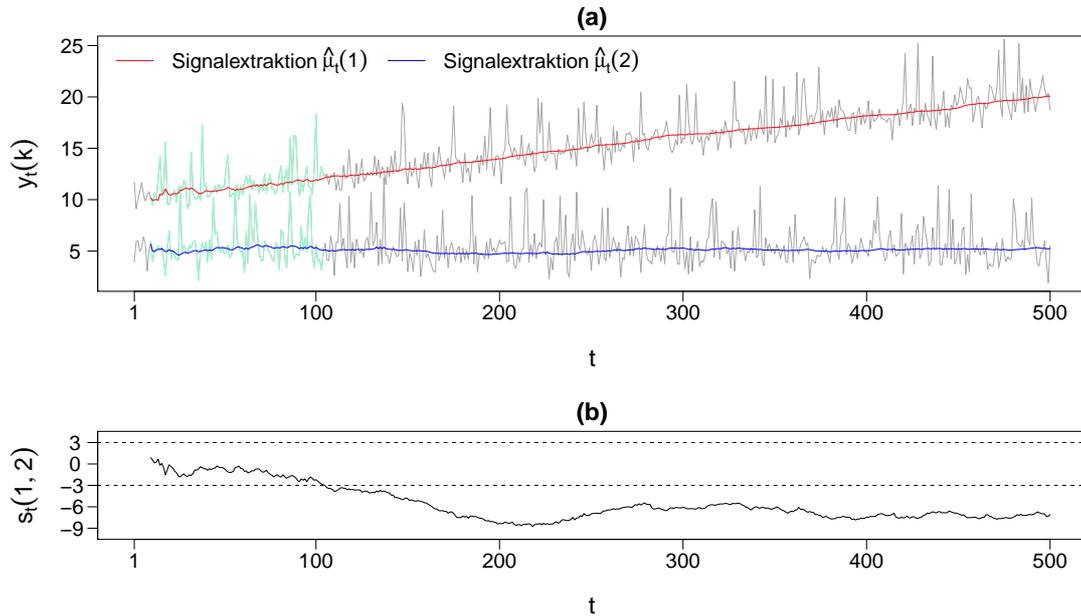


Abbildung 5.3: (a) Bivariate Zeitreihe (Szenario 2) mit SCARM-Signalextraktion und farblicher Kennzeichnung von Zusammenhängen auf Basis von $s_t(1,2)$; (b) Zeitreihe der Vergleichsstatistik $s_t(1,2)$.

Zeitreihe $y_t(1)$ einen positiven Trend von $\beta_t = 0.02$ für alle $t = 1, \dots, 500$ aufweist. Dieser leichte Trend wird ab dem Zeitpunkt $t = 105$ erkannt. Ab diesem Zeitpunkt liegen die Werte von $s_t(1,2)$ in Abbildung 5.3 (b) deutlich unterhalb der unteren Grenze $-c_{\text{bound}} = -3$. Offenbar kann das SSM-Verfahren auch geringe Trendunterschiede erkennen. Hintergrund ist die Tatsache, dass die Varianz von $D_t(i,j)$ mit zunehmenden Fensterbreiten $n_t(i)$ und $n_t(j)$ sinkt. Auch die Schätzung von $\text{Var}[D_t(i,j)]$ fällt entsprechend klein aus, da die Approximationen $v_{n_t(i)}$ und $v_{n_t(j)}$ monoton fallend in $n_t(i)$ und $n_t(j)$ sind, vgl. (5.6) und Anhang B.1. Folglich steigt die Güte des SSM-Verfahrens, d.h. die Wahrscheinlichkeit zur Erkennung ungleicher Steigungen, mit zunehmenden Fensterbreiten $n_t(i)$ und $n_t(j)$. In diesem Szenario erkennt der SCARM richtigerweise keine Signalveränderung, weshalb $n_t(1)$ und $n_t(2)$ auf den Maximalwert $n_{\text{max}} = 200$ anwachsen und das SSM-Verfahren eine hohe Güte erreicht. Würde der SCARM hier in einer oder beiden Zeitreihen fälschlicherweise eine Signalveränderung erkennen und somit eine oder beide Fensterbreiten auf den Minimalwert n_{min} setzen, würde das SSM-Verfahren den leichten Trendunterschied für die nachfolgenden Zeitpunkte vermutlich nicht erkennen. Eine solche Situation zeigt Abbildung C.17 in Anhang C. Hier werden die Fensterbreiten wegen $\alpha = 0.1$ mehrfach fälschlicherweise ver-

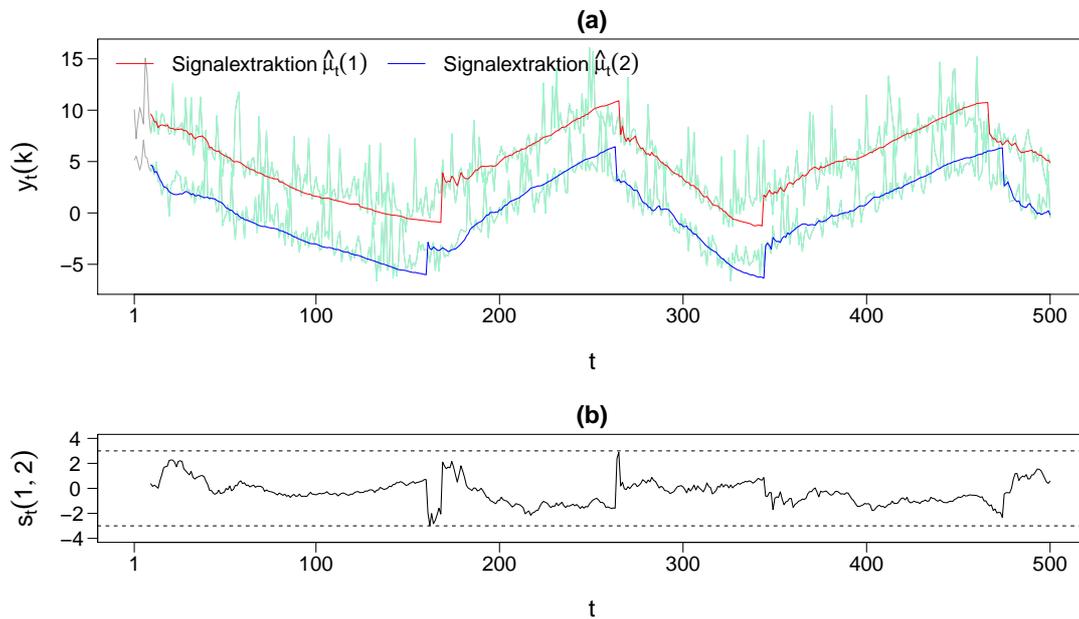


Abbildung 5.4: (a) Bivariate Zeitreihe (Szenario 3) mit SCARM-Signalextraktion und farblicher Kennzeichnung von Zusammenhängen auf Basis von $s_t(1,2)$; (b) Zeitreihe der Vergleichsstatistik $s_t(1,2)$.

kleinert, woraufhin $s_t(i,j)$ jeweils wieder in den Bereich $[-c_{\text{bound}}, c_{\text{bound}}]$ mit $c_{\text{bound}} = 3$ zurück fällt. Dies zeigt die Bedeutung des für den SCARM gewählten Signifikanzniveaus α . Damit das SSM-Verfahren eine hohe Güte aufweist, muss α entsprechend klein gewählt sein.

Szenario 3: Zwei gleich verlaufende Signale mit wechselnden Trends

Die Zeitreihen $y_t(1)$ und $y_t(2)$ in Abbildung 5.4 (a) bestehen aus stückweise linearen Signalen $\mu_t(1)$ und $\mu_t(2)$, wobei $\beta_t(1) = \beta_t(2) \neq 0$ und $\mu_t(1) = \mu_t(2) + 5$ für alle $t = 1, \dots, 500$. Die Signale $\mu_t(1)$ und $\mu_t(2)$ verlaufen demnach über den gesamten Zeitraum gleich. Da die SSM-Statistik $s_t(i,j)$ in der unteren Abbildung (b) zu jedem Zeitpunkt einen Zusammenhang erkennt – es gilt $s_t(i,j) \in [-c_{\text{bound}}, c_{\text{bound}}]$ mit $c_{\text{bound}} = 3$ für alle $t = 10, \dots, 500$ –, ist die Zeitreihe \mathbf{y}_t in Abbildung (a) durchgehend grün markiert. Allerdings liegen die $s_t(1,2)$ -Werte in der unteren Abbildung (b) gegen $t = 160$ und $t = 260$ nur knapp innerhalb des Grenzbereichs. Wäre hier $c_{\text{bound}} = 2$ gewählt, würden die beiden Zeitreihen $y_t(1)$ und $y_t(2)$ kurzzeitig als unähnlich bewertet werden. Der Gesamteindruck wäre dadurch jedoch nicht verändert.

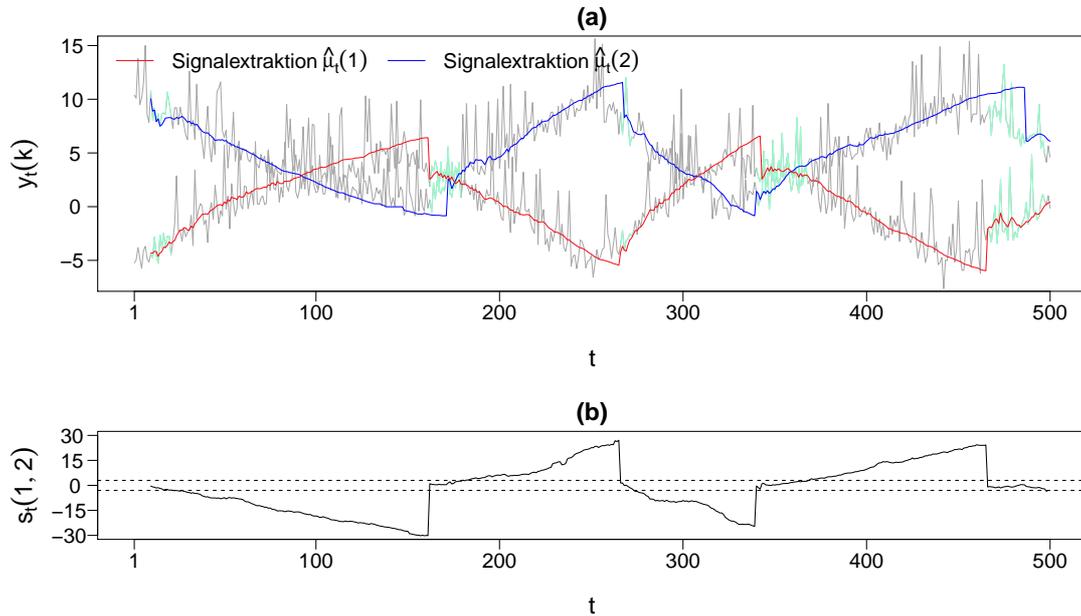


Abbildung 5.5: (a) Bivariate Zeitreihe (Szenario 4) mit SCARM-Signalextraktion und farblicher Kennzeichnung von Zusammenhängen auf Basis von $s_t(1,2)$; (b) Zeitreihe der Vergleichsstatistik $s_t(1,2)$.

Die kurzen Ausschläge in der Zeitreihe der $s_t(1,2)$ -Werte lassen sich wie folgt erklären: Weisen wie in Abbildung 5.4 (a) zwei ähnlich verlaufende Zeitreihen etwa zum selben Zeitpunkt eine Signalveränderung auf, so werden diese Veränderungen in der Regel nicht zeitgleich erkannt. In Abbildung 5.4 (a) wird z.B. der Trendwechsel zum Zeitpunkt $t = 150$ zuerst in der Zeitreihe $y_t(2)$ (blaues Signal) erkannt. Etwa zehn Zeitpunkte später wird auch der Trendwechsel in der Zeitreihe $y_t(1)$ (rotes Signal) erkannt. In diesem kurzen Zeitraum ergeben sich deutlich unterschiedliche Steigungsschätzungen $\hat{\beta}_t(1)$ und $\hat{\beta}_t(2)$ und somit deutlich von null verschiedene SSM-Statistiken $s_t(1,2)$.

Szenario 4: Zwei ungleich verlaufende Signale mit wechselnden Trends

In diesem Szenario verlaufen die Zeitreihen $y_t(1)$ und $y_t(2)$ entgegengesetzt, d.h. zu jedem Zeitpunkt gilt $\beta_t(1) = -\beta_t(2) \neq 0$. Ein Zusammenhang im Sinne eines ähnlichen Verlaufs ist somit zu keinem Zeitpunkt gegeben. Dieser Umstand wird durch die SSM-Methodik weitestgehend wiedergegeben, siehe Abbildung 5.5 (a). So liegen die Werte der SSM-Statistik $s_t(i,j)$ in der unteren Abbildung (b) zu den überwiegenden Zeitpunkten deutlich außerhalb des Bereichs $[-c_{\text{bound}}, c_{\text{bound}}]$ mit $c_{\text{bound}} = 3$. Lediglich nach einigen Trendwechseln

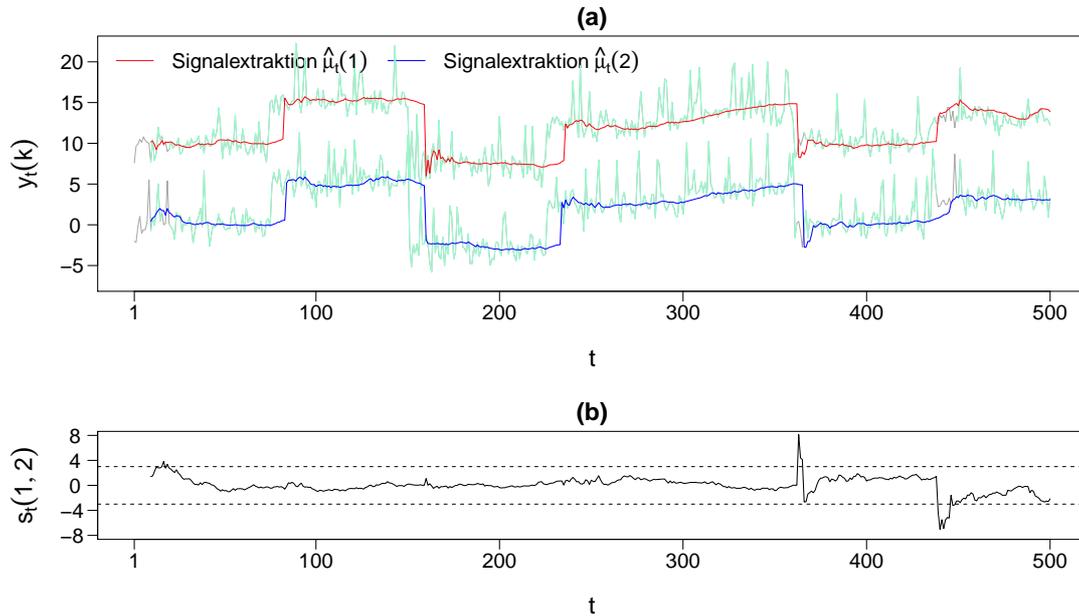


Abbildung 5.6: (a) Bivariate Zeitreihe (Szenario 5) mit SCARM-Signalextraktion und farblicher Kennzeichnung von Zusammenhängen auf Basis von $s_t(1,2)$; (b) Zeitreihe der Vergleichsstatistik $s_t(1,2)$.

gilt kurzzeitig $s_t(i,j) \in [-c_{\text{bound}}, c_{\text{bound}}]$. Dies ist dem Umstand geschuldet, dass bei den hier vorliegenden *gleichzeitigen* Trendwechseln sowohl $\beta_t(1)$ als auch $\beta_t(2)$ kurzzeitig nahe null geschätzt werden. In diesem Zeitraum ist dann auch $|s_t(i,j)| < c_{\text{bound}}$, so dass die Zeitreihe \mathbf{y}_t grün gezeichnet ist. Insgesamt wird durch die farbliche Kennzeichnung jedoch der richtige Eindruck vermittelt, dass $y_t(1)$ und $y_t(2)$ nicht ähnlich verlaufen.

Szenario 5: Zwei gleich verlaufende Signale mit Level Shifts

Im fünften Szenario in Abbildung 5.6 (a) werden Zeitreihen $y_t(1)$ und $y_t(2)$ mit stückweise konstanten Signalen betrachtet, wobei $\mu_t(1) = \mu_t(2) + 5$ für alle $t = 1, \dots, 500$. Die Signale zeigen somit durchgehend einen gleichen Verlauf, d.h. $\beta_t(1) = \beta_t(2) = 0$ für alle $t = 1, \dots, 500$, und die Level Shifts treten zeitgleich auf. Abgesehen von drei kurzen Zeiträumen verläuft die SSM-Statistik $s_t(i,j)$ in der unteren Abbildung (b) innerhalb des Bereichs $[-c_{\text{bound}}, c_{\text{bound}}]$ mit $c_{\text{bound}} = 3$. Das SSM-Verfahren zeigt somit richtigerweise fast durchgehend einen Zusammenhang im Sinne eines ähnlichen Verlaufs an. Lediglich zu den Zeitpunkten $t = 363, 364, 365$ und $t = 439, \dots, 448$ liegen die Werte von $s_t(i,j)$ kurzzeitig deutlich außerhalb des Bereichs $[-c_{\text{bound}}, c_{\text{bound}}]$. Der Grund hierfür ist, dass die Sprünge

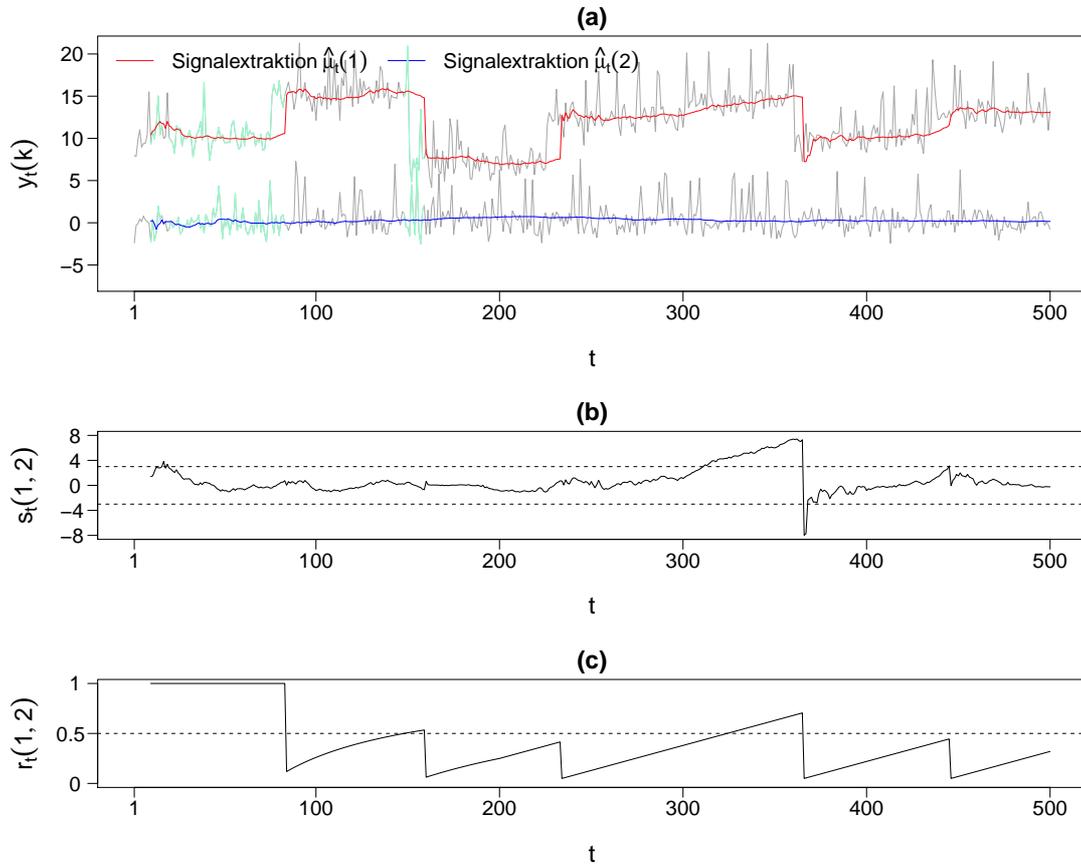


Abbildung 5.7: (a) Bivariate Zeitreihe (Szenario 6) mit SCARM-Signalextraktion und farblicher Kennzeichnung von Zusammenhängen auf Basis von $s_t(1,2)$; (b) Zeitreihe der Vergleichsstatistik $s_t(1,2)$; (c) Zeitreihe der Fensterbreiten-Verhältniszahlen $r_t(1,2)$.

nicht gleichzeitig erkannt werden. Folglich ist die Differenz $|d_t(i,j)|$ der Steigungsschätzungen $\hat{\beta}_t(1)$ und $\hat{\beta}_t(2)$ kurzzeitig groß. Sobald der Level Shift auch für die zweite Komponente erkannt ist, gilt wieder $s_t(i,j) \in [-c_{\text{bound}}, c_{\text{bound}}]$.

Szenario 6: Eine stationäre Zeitreihe und eine Zeitreihe mit Level Shifts

Beim sechsten Szenario, dargestellt in Abbildung 5.7 (a), weist die Zeitreihe $y_t(1)$ dasselbe Signal auf wie im Szenario 5. Die zweite Zeitreihe $y_t(2)$ besitzt jedoch ein konstantes Signal, d.h. $\mu_t(2) = 0$ für alle $t = 1, \dots, 500$. Wegen $\beta_t(1) = \beta_t(2) = 0$ für alle $t = 1, \dots, 500$ weisen die beiden Zeitreihen zwar zu jedem Zeitpunkt die selbe Steigung auf, sie verlaufen jedoch nicht ähnlich. Dieses Szenario stellt demnach eine Herausforderung an die SSM-Methodik dar. So gilt für die überwiegenden Zeitpunkte $s_t(i,j) \in [-c_{\text{bound}}, c_{\text{bound}}]$ mit $c_{\text{bound}} = 3$, vgl.

die mittlere Abbildung 5.7 (b). Das SSM-Verfahren erkennt also für die beiden Zeitreihen überwiegend einen Zusammenhang im Sinne einer ähnlichen Steigung. Dies entspricht zwar den Tatsachen, da $\beta_t(1) = \beta_t(2) = 0$ für alle $t = 1, \dots, 500$. Es wird jedoch ein falscher Eindruck vermittelt, da der Verlauf der beiden Zeitreihen $y_t(1)$ und $y_t(2)$ unterschiedlich ist. Für derartige Fälle können die vom SCARM adaptierten Fensterbreiten, bzw. deren Quotient $r_t(i, j) = \min\{n_t(i), n_t(j)\} / \max\{n_t(i), n_t(j)\} \in (0, 1]$, als zusätzliche Information genutzt werden. Ist $r_t(i, j)$ kleiner als eine festgelegte Schranke r_{bound} , ist die Differenz der adaptierten Fensterbreiten zu groß, als dass der Verlauf der Datenzeitreihen als ähnlich bewertet werden könnte. In der unteren Abbildung 5.7 (c) sind die Werte von $r_t(i, j)$ dargestellt, die Schranke $r_{\text{bound}} = 0.5$ ist hier durch die gestrichelte Linie gekennzeichnet. Zu den Zeitpunkten, an denen $r_t(i, j) < r_{\text{bound}} = 0.5$, ist für die Zeitreihen $y_t(1)$ und $y_t(2)$ ein Zusammenhang ausgeschlossen. Folglich sind die beiden Zeitreihen in der oberen Abbildung (a) überwiegend grau gezeichnet, und somit wird nicht der Eindruck eines Zusammenhangs vermittelt. Wird die auf $r_t(i, j)$ basierende Zusatzregel nicht angewendet, werden $y_t(1)$ und $y_t(2)$ überwiegend grün gekennzeichnet, vgl. Abbildung C.18 in Anhang C.

Nach der Anwendung der Prozedur in den hier betrachteten sechs Szenarien bivariater Zeitreihen erscheint es sinnvoll, die Schranke c_{bound} für die SSM-Statistik $S_t(i, j)$ etwa im Bereich $[2, 4]$ zu wählen. Des Weiteren ist zu empfehlen, für den SCARM-Test ein kleines Signifikanzniveau von z.B. $\alpha = 0.001$ festzulegen, damit die SSM-Prozedur auch geringe Trendunterschiede erkennen kann. Im Falle von Level Shifts kann die Wahl einer Schranke r_{bound} für den Quotienten der Fensterbreiten $r_t(i, j)$ zu besseren Resultaten führen. Diese sollte nicht zu groß gewählt werden, insbesondere wenn die Differenz zwischen der gewählten minimalen und maximalen Fensterbreite n_{min} und n_{max} groß ist. Sei z.B. $n_{\text{min}} = 10$ und $n_{\text{max}} = 400$ und $r_{\text{bound}} = 0.5$. Falls $n_t(i) = n_t(j) = n_{\text{max}} = 400$ und die Fensterbreite $n_{t+1}(i)$ (fälschlicherweise) auf den Wert $n_{\text{min}} = 10$ gesetzt würde, müsste diese erst wieder auf den Wert 200 anwachsen, bis die beiden Zeitreihen als ähnlich bewertet werden können.

5.2.4 Diskussion

Die entwickelte SSM-Methodik erfüllt alle der auf S. 84 geforderten Bedingungen und ist somit geeignet zur Echtzeit-Überwachung der Zusammenhänge in instationären, verrauschten und Ausreißer behafteten multivariaten Zeitreihen. So stellt das SSM-Verfahren lediglich geringe Annahmen an die Daten und benötigt nur geringe Rechenzeiten, da die vom recheneffizienten SCARM gelieferten Schätzungen $\hat{\beta}_t(i)$, $\hat{\beta}_t(j)$, $\widehat{\text{Var}}[\hat{\beta}_t(i)]$ und $\widehat{\text{Var}}[\hat{\beta}_t(j)]$

genutzt werden. Somit bleibt lediglich die Korrelation $\text{Corr}[\hat{\beta}_t(i), \hat{\beta}_t(j)]$ zu schätzen, vgl. (5.8), was mittels des schnell zu berechnenden OGK- Q^{adj} geschieht. Neben der schnellen Rechenzeit bietet das SSM-Verfahren die geforderte Robustheit gegenüber Ausreißern, da alle Komponenten der Vergleichsstatistik $S_t(i, j)$ robust geschätzt werden. Des Weiteren führen fehlende Werte nicht zu einem Abbruch des SSM-Verfahrens, da für den SCARM Regeln zur Behandlung fehlender Werte bestehen, vgl. Kapitel 3.3.3. Liegen für eine Zeitreihe $y_t(i)$ zu einem Zeitpunkt t nicht genügend aktuelle Messwerte vor, liefert der SCARM als Steigungsschätzung einen fehlenden Wert. In diesem Fall erfolgt für $y_t(i)$ zum Zeitpunkt t kein Vergleich mit den übrigen Zeitreihen.

Die SSM-Prozedur erfordert die Schätzung der Varianz der RM-Steigungsdifferenzen $\text{Var}[D_t(i, j)] = \text{Var}[\hat{\beta}(\mathbf{Y}_{t, n_t(i)}) - \hat{\beta}(\mathbf{Y}_{t, n_t(j)})]$. Hierzu wird in Kapitel 5.2.2 eine Methodik entwickelt, die auf den Annahmen gründet, dass der Fehlerprozess $\boldsymbol{\varepsilon}_t$ multivariat normalverteilt und der Ausreißerprozess $\boldsymbol{\eta}_t = 0$ ist für alle $t \in \mathbb{Z}$. In der Praxis sind diese Annahmen jedoch häufig nicht erfüllt. Die Robustheit gegenüber Ausreißern ist allerdings gegeben, da die Schätzungen von $\text{Var}[\hat{\beta}_t(i)]$ und $\text{Var}[\hat{\beta}_t(j)]$ in (5.8) mit dem robusten Q^{adj} -Skalenschätzer geschehen und zur Schätzung von $\rho_t(i, j)$ der robuste OGK- Q^{adj} verwendet wird. Zu beachten ist allerdings, dass sowohl das Modell zur Approximation von $v_{n_t(i)}$ und $v_{n_t(j)}$ in (5.6) als auch das Modell \tilde{C} aus (5.7) mittels Simulationen unter Verwendung nicht autokorrelierter normalverteilter Fehler erstellt worden ist. Für andere Fehlerverteilungen und für autokorrelierte Fehler ergeben sich im Allgemeinen verzerrte Schätzungen von $\text{Var}[D_t(i, j)]$.

Um Verzerrungen auszugleichen, können die Varianzschätzer $\widehat{\text{Var}}[\hat{\beta}_t(i)]$ und $\widehat{\text{Var}}[\hat{\beta}_t(j)]$ in (5.8) an die Fehlerverteilung angepasst werden. Dabei ist

$$\widehat{\text{Var}}[\hat{\beta}_t(k)] = \hat{\sigma}_t^2(k) \cdot v_{n_t(k)} = Q_\delta^{\text{adj}}[\mathbf{y}_{t, n_t(k)}] \cdot v_{n_t(k)} = c_{n_t(k)} \cdot h_{(\lfloor \delta \cdot (n_t(k) - 2) \rfloor)} \cdot v_{n_t(k)} \quad (5.9)$$

mit $\delta = 0.5$. Hier ist $h_{(\lfloor \delta \cdot (n_t(k) - 2) \rfloor)}$ das δ -Quantil der $n_t(k) - 2$ vertikalen Höhen der Dreiecke, die sich aus den $n_t(k) - 2$ Tripeln von jeweils drei aufeinander folgenden Beobachtungen aus $\mathbf{y}_{t, n_t(k)}$ ergeben, vgl. (3.12) und (3.13) auf S. 32. Des Weiteren ist $c_{n_t(k)}$ ein Korrekturfaktor, der die Unverzerrtheit des Q_δ^{adj} -Schätzers für den Stichprobenumfang $n_t(k)$ und Fehler aus einer bestimmten Verteilung sicherstellen soll. Das $v_{n_t(k)}$ ist eine Approximation der Varianz der RM-Steigung für Stichproben der Länge $n_t(k)$ mit linearem Signal und Fehlervarianz eins. Durch entsprechende Wahl der Konstanten c_n und v_n mit $n = n_t(k)$ lässt sich $\widehat{\text{Var}}[\hat{\beta}_t(k)]$ an nicht normalverteilte oder autokorrelierte Fehler anpassen. Für Fehler, die

sich durch einen AR(1)-Prozess mit Parameter φ beschreiben lassen, sind für den SCARM in Kapitel 3.5 bereits Konstanten c_n^φ und v_n^φ eingeführt worden. Ist der wahre Parameter φ zeitveränderlich oder nicht für alle Variablen gleich, kann dieser zu jedem Zeitpunkt t anhand der letzten n_{\max} Residuen der SCARM-Signalschätzung $y_{t-n_{\max}+i} - \hat{\mu}_{t-n_{\max}+i}^{\text{SCARM}}$ mit $i = 1, \dots, n_{\max}$ geschätzt werden. Zu jedem Zeitpunkt bestimmt dann die Schätzung $\hat{\varphi}_t$ die Wahl von φ und somit die Wahl der Konstanten v_n^φ und c_n^φ über die Regel (3.20).

Eine R-Funktion des SSM-Verfahrens zur Überwachung der Zusammenhänge in multivariaten Zeitreihen, bezeichnet mit `SSM()`, steht auf der folgenden Webseite zum Download zur Verfügung:

<http://www.statistik.tu-dortmund.de/1542.html>.

Für die Inputargumente der Funktion `SSM()` gelten die Entsprechungen `c.bound` $\hat{=}$ c_{bound} und `r.bound` $\hat{=}$ r_{bound} . Da die Funktion `SSM()` die in Kapitel 3.3.1 vorgestellte R-Funktion `scarm.filter` beinhaltet, ermöglicht sie die Wahl der Inputargumente für den SCARM, die auf S. 50 aufgeführt sind.

Das SSM-Verfahren liefert zu jedem Zeitpunkt Aufschluss über die aktuell bestehenden Zusammenhänge im Sinne aktuell ähnlich verlaufender Komponenten. In Hinblick auf die Anwendung in Echtzeit stellt sich die Frage, wie dem Anwender die Information auf leicht nachvollziehbare Weise angezeigt wird. Für den Fall $K = 2$ stellt die hier vorgenommene Kennzeichnung der gemessenen Zeitreihen durch wechselnde Farben eine Möglichkeit dar. Bei höher dimensionalen Zeitreihen ist die farbliche Kennzeichnung jedoch oftmals unübersichtlich und daher weniger geeignet. Hier könnte ein graphisches Modell angewendet werden, das fortlaufend aktualisiert wird. Ein graphisches Modell stellt die K Variablen als Knoten dar, und die Abhängigkeit zwischen zwei Variablen wird durch eine Kante zwischen den Knoten angezeigt. Damit ließe sich auch das Ausmaß des Zusammenhangs anzeigen, indem z.B. die Dicke der Kante den Wert der Vergleichsstatistik $s_t(i, j)$ wiedergibt. Da ein graphisches Modell jedoch fortlaufend zu aktualisieren wäre, kann dieses lediglich die aktuelle Situation wiedergeben. Der zeitliche Verlauf der Zusammenhänge wäre somit nicht nachvollziehbar.

Eine alternative Möglichkeit zur Wiedergabe der aktuell bestehenden Zusammenhänge wird im folgenden Kapitel aufgezeigt. In diesem wird ein Signalextraktionsverfahren für multivariate Zeitreihen mit adaptiver Fensterbreitenwahl entwickelt. Dieses nutzt das hier

vorgestellte SSM-Verfahren zur automatischen Blockbildung, wie sie im Zuge der Diskussion zu der inadäquaten Fensterbreitenadaption des aoTRM-LS auf S. 25 gefordert wird. Die gewählte Blockstruktur wird dann durch K farbige Rechtecke dargestellt, wobei den Variablen eines Blocks dieselbe Farbe zugeordnet ist. Die K farbigen Rechtecke werden übereinander angeordnet und in zeitlicher Abfolge gezeichnet, so dass sowohl die aktuellen Zusammenhänge als auch der zeitliche Verlauf der Zusammenhänge wiedergegeben wird, siehe Abbildung 6.2 (b) auf S. 113.

6 Der multivariate Slope Comparing Adaptive Repeated Median Filter

Der in Kapitel 2.2.2 vorgestellte aoTRM-LS ergibt sich aus der Kombination des aoRM aus Kapitel 2.1.2 und des TRM-LS aus Kapitel 2.2.1. Der aoTRM-LS soll die Vorteile des aoRM und des TRM-LS vereinen zu einem multivariaten Filter mit adaptiver Fensterbreitenwahl. Die Fensterbreitenadaptation des aoTRM-LS funktioniert jedoch nicht zufriedenstellend (Borowski et al., 2009; Schettlinger, 2009, Kap. 3.3.3). Ansätze zur Verbesserung bieten eine verbesserte Fensterbreitenadaptation sowie eine automatische zeitabhängige Einteilung der K Komponenten der K -variaten Zeitreihe in disjunkte Blöcke abhängiger Variablen. Eine verbesserte Fensterbreitenadaptation kann mit dem SCARM erreicht werden, da dieser im Gegensatz zum aoRM auch für kleine Signifikanzniveaus eine hohe Güte liefert, vgl. Kapitel 4.2. Die in Kapitel 5 entwickelte SCARM-basierte SSM-Prozedur zur Überwachung der Zusammenhänge kann dann zur automatischen zeitabhängigen Blockbildung verwendet werden. Aus diesen beiden Ansätzen wird in Kombination mit dem TRM-LS ein neuer Echtzeit-Filter für multivariate Zeitreihen entwickelt, der *multivariate Slope Comparing Adaptive Repeated Median Filter* (mSCARM). In Kapitel 6.1 wird zunächst der Algorithmus des mSCARM präsentiert. Anschließend werden die Schritte des Algorithmus detailliert erklärt sowie die Prinzipien zum Umgang mit fehlenden Messwerten vorgestellt. In Kapitel 6.2 erfolgt dann die Anwendung des mSCARM auf multivariate Zeitreihen aus dem intensivmedizinischen Online-Monitoring.

6.1 Der mSCARM-Algorithmus

Sei mit $(\mathbf{y}_t)_{t \in \mathbb{Z}}$ eine K -variate Zeitreihe gegeben mit $\mathbf{y}_t = (y_t(1), \dots, y_t(K))'$ gemäß (5.1), und $n_t(k)$ bezeichne die zum Zeitpunkt t für die k -te Komponente der Zeitreihe adaptierte individuelle Fensterbreite, $k = 1, \dots, K$. Des Weiteren sind α , r , ℓ_{\min} , n_{\min} und n_{\max} mit $n_{\min} < \ell_{\min} + r < n_{\max}$ sowie c_{bound} und r_{bound} die vom Anwender festzulegenden Inputargumente des SCARM und des SSM-Verfahrens.

Der mSCARM bildet zu jedem Zeitpunkt Blöcke aus aktuell ähnlich verlaufenden Komponenten. Es ergeben sich demnach $K_t \leq K$ Blöcke; der j -te Block sei dabei durch die Indexmenge $B_t(j) \subset \{1, \dots, K\}$ identifiziert, $j = 1, \dots, K_t$. Die Einteilung der K Komponenten auf die K_t Blöcke ist eindeutig und vollständig, d.h. es gilt $B_t(j) \cap B_t(j^*) = \emptyset$ für alle $j \neq j^*$ mit $j, j^* \in \{1, \dots, K_t\}$ und $\bigcup_{j=1}^{K_t} B_t(j) = \{1, \dots, K\}$.

Der mSCARM-Algorithmus lautet wie folgt:

0. Start: Warte bis n_{\min} Beobachtungen gegeben sind. Setze $t = n_{\min}$ und $n_t(k) = n_{\min}$ für alle $k = 1, \dots, K$.
1. Wende für jede der k Komponenten im Zeitfenster der Länge $n_t(k)$ den SCARM an:
 - (a) Schätze die Fehlervarianz $\sigma_t^2(k)$ mit dem $Q_{0.5}^{\text{adj}}$ -Skalenschätzer anhand der letzten $n_t(k)$ Beobachtungen.
 - (b) Falls $n_t(k) < \ell_{\min} + r$ ist, gehe zu Schritt (d).
 - (c) Teste H_0 : Das Signal ist linear im Zeitfenster der Länge $n_t(k)$. Falls H_0 abgelehnt wird, setze $n_t(k) \leftarrow n_{\min}$.
 - (d) Führe die RM-Regression im Zeitfenster $\{t - n_t(k) + 1, \dots, t\}$ durch und speichere das RM-Niveau zum Zeitpunkt t als Signalschätzung $\hat{\mu}_t^{\text{SCARM}}(k)$.

Speichere die K individuellen Fensterbreiten $n_t(k)$.

2. Bestimme die SSM-Statistiken $s_t(k, k^*)$ mit $k \neq k^*$ und $k, k^* \in \{1, \dots, K\}$. Bilde anhand der $s_t(k, k^*)$ Blöcke aus aktuell ähnlich verlaufenden Komponenten, die durch die Indexmengen $B_t(j)$, $j = 1, \dots, K_t$, identifiziert seien.
3. Blockweise Signalschätzung für alle $j = 1, \dots, K_t$:

Falls $|B_t(j)| = 1$: Speichere die Signalschätzung $\hat{\mu}_t^{\text{SCARM}}(k) =: \hat{\mu}_t^{\text{mSCARM}}(k)$ aus Schritt 1 (d) für $k \in B_t(j)$.

Falls $|B_t(j)| \geq 2$: Bestimme die Blockfensterbreite $n_t(j) = \min\{n_t(k) : k \in B_t(j)\}$. Führe die TRM-LS-Regression für die Komponenten $k \in B_t(j)$ im Zeitfenster der Länge $n_t(j)$ durch und speichere die Signalschätzungen $\hat{\mu}_t^{\text{TRM-LS}}(k) =: \hat{\mu}_t^{\text{mSCARM}}(k)$. Setze $n_t(k) \leftarrow n_t(j)$ für alle $k \in B_t(j)$.

Füge die Signalschätzungen zusammen zu $\hat{\boldsymbol{\mu}}_t^{\text{mSCARM}} = (\hat{\mu}_t^{\text{mSCARM}}(k))'_{k=1, \dots, K}$.

4. Update der Fensterbreiten: Setze $n_{t+1}(k) \leftarrow \min\{n_t(k) + 1, n_{\max}\}$ für $k = 1, \dots, K$.
Update des Zeitindex: Setze $t \leftarrow t + 1$ und gehe zu Schritt 1.

Sobald n_{\min} Beobachtungen vorliegen, startet der Algorithmus (Schritt 0). Zu diesem Zeitpunkt t gilt $n_t(k) = n_{\min} < \ell_{\min} + r$ für alle $k = 1, \dots, K$, so dass in Schritt 1 die Fensterbreiten nicht durch den SCARM adaptiert werden. Erst wenn zu späteren Zeitpunkten

$n_t(k) \geq \ell_{\min} + r$ gilt, erfolgt die Anwendung des SCARM-Tests und je nach Testentscheidung die Fensterbreitenadaptation.

In Schritt 2 werden dann die Blöcke auf Basis der $K(K-1)/2$ SSM-Statistiken $s_t(k, k^*) = d_t(k, k^*) / \widehat{\text{Var}}[d_t(k, k^*)]^{1/2}$ gebildet; der Algorithmus der Blockbildung wird in Kapitel 6.1.1 detailliert erklärt. Die Berechnung der Vergleichsstatistiken $s_t(k, k^*)$ geschieht wie in Kapitel 5.2 beschrieben. Dabei werden zur Bestimmung von $\widehat{\text{Var}}[d_t(k, k^*)]$ gemäß (5.8) die $Q_{0.5}^{\text{adj}}$ -Schätzungen der Fehlervarianzen $\sigma_t^2(k)$ aus Schritt 1 (a) genutzt. Die ebenfalls zur Berechnung von $\widehat{\text{Var}}[d_t(k, k^*)]$ benötigte Schätzung der lokalen Fehlerkovarianzmatrix $\Sigma_t \in \mathbb{R}^{K \times K}$ erfolgt anhand dieser $Q_{0.5}^{\text{adj}}$ -Schätzungen mit dem OGK- $Q_{0.5}^{\text{adj}}$.

Nach der Blockbildung erfolgt in Schritt 3 die Signalschätzung innerhalb der einzelnen Blöcke. Besteht ein Block aus lediglich einer Komponente k , wird die bereits in Schritt 1 (d) gewonnene SCARM-Signalschätzung als mSCARM-Signalschätzung für die Komponente k gespeichert. Enthält ein Block zwei oder mehr Komponenten, geschieht die Signalschätzung anhand einer Modifikation der multivariaten TRM-LS-Regression nach Lanius und Gather (2010), die in Kapitel 6.1.2 erklärt wird. Ist in jedem Block die Signalschätzung per SCARM bzw. TRM-LS erfolgt, werden die Signalschätzungen zusammengefasst zu dem Vektor $\hat{\mu}_t^{\text{mSCARM}}$ der mSCARM-Signalschätzungen zum Zeitpunkt t .

6.1.1 Die Blockbildung des mSCARM

Im Zuge der Blockbildung in Schritt 2 des mSCARM-Algorithmus sollen die Komponenten den Blöcken eindeutig zugewiesen werden und die Blöcke zudem disjunkt sein. Betrachte hierzu den durchaus möglichen Fall, dass $s_t(i, j) < c_{\text{bound}}$, $s_t(i, k) < c_{\text{bound}}$ und $s_t(j, k) > c_{\text{bound}}$. Hier stellt sich die Frage, ob die Komponenten i, j, k demselben Block zuzuordnen sind, obwohl j und k als unähnlich bewertet werden. Alternativ könnten entweder die Variablen i und j oder die Variablen i und k einen Block bilden und die Variable j oder k ein eigenständiger Block sein. Offenbar müssen Regeln definiert werden, welche die disjunkte vollständige Aufteilung der Komponenten auf die Blöcke gewährleisten. Busse (2012, Kap. 4.3) schlägt vor, eine Komponente einem Block zuzuordnen, sobald sie als ähnlich zu mindestens einer Komponente im Block bewertet wird. Dabei ist jedoch nicht gewährleistet, dass zwischen allen Komponentenpaaren innerhalb eines Blocks ein Zusammenhang besteht. Daher wird für den mSCARM ein Algorithmus entwickelt, der möglichst große Blöcke bildet unter der Bedingung, dass alle Komponenten innerhalb eines Blocks paarweise ähnlich sind.

Sei im Folgenden $s_t(k, k^*)$ der absolute Wert der SSM-Statistik zu den Komponenten k

und k^* mit $k, k^* \in \{1, \dots, K\}$, $k < k^*$ und $K \geq 2$. Dann lautet der Algorithmus der Blockbildung wie folgt:

0. Setze $j = 1$, und $R = \{1, \dots, K\}$ sei die Indexmenge der noch nicht eingeteilten Variablen.
1. Bestimme $B_t(j) = \left\{ \arg \min_{k, k^* \in R} \{s_t(k, k^*) : s_t(k, k^*) < c_{\text{bound}}\} \right\}$ und setze $R \leftarrow R \setminus B_t(j)$.
Falls $B_t(j) \neq \emptyset$: Gehe zu Schritt 3.
2. (a) Falls $R = \emptyset$: Setze $K_t = j - 1$ und breche ab.
(b) Falls $R \neq \emptyset$: Setze $B_t(j) \leftarrow \{\min\{k \in R\}\}$ und $R \leftarrow R \setminus B_t(j)$.
Setze $j \leftarrow j + 1$ und gehe zu (a).
3. Falls $R = \emptyset$: Setze $K_t = j$ und breche ab.
4. Bestimme $I = \{k \in R : s_t(k, k^*) < c_{\text{bound}} \forall k^* \in B_t(j)\}$.
Falls $I = \emptyset$: Setze $j \leftarrow j + 1$ und gehe zu Schritt 1.
5. Bestimme $k = \arg \min_{k \in I} \{s_t(k, k^*) : k^* \in B_t(j)\}$.
Setze $B_t(j) \leftarrow B_t(j) \cup \{k\}$ und $R \leftarrow R \setminus \{k\}$. Gehe zu Schritt 4.

Zunächst werden in Schritt 1 die beiden Komponenten k, k^* bestimmt, für welche die SSM-Statistik $s_t(k, k^*)$ minimal und kleiner als die Schranke c_{bound} ist; die Indizes der übrigen Komponenten bilden die Restmenge R . Gibt es kein derartiges Tupel (k, k^*) , so bildet jede der K Komponenten einen eigenständigen Block, und der Algorithmus ist beendet (Schritt 2). Existiert jedoch ein Paar (k, k^*) , so entspricht dieses dem ersten Block, d.h. $B_t(j) = \{k, k^*\}$ mit $j = 1$. Anschließend wird in Schritt 4 die Indexmenge I bestimmt. Diese enthält diejenigen Komponenten $k \in R$, die zu allen Komponenten $k^* \in B_t(j)$ einen Zusammenhang aufweisen. Ist I nicht leer, wird diejenige Komponente $k \in I$ dem Block $B_t(j)$ hinzugefügt, für die $s_t(k, k^*)$ mit $k^* \in B_t(j)$ minimal ist (Schritt 5). Anschließend werden die Indexmengen I und R neu bestimmt. Diese Vorgänge werden so lange wiederholt, bis $I = \emptyset$. Dann ist der Block $B_t(j)$ vollständig, und ein neuer Block kann gebildet werden, sofern nicht bereits alle Komponenten eingeteilt worden sind.

Dieser Algorithmus setzt voraus, dass die $K(K - 1)/2$ SSM-Statistiken $s_t(k, k^*)$ eindeutig zu ordnen sind. Andernfalls könnte die Menge $B_t(j)$ in Schritt 1 mehr als ein Tupel (k, k^*) enthalten, und die Komponente k in Schritt 5 wäre möglicherweise nicht eindeutig zu bestimmen. Da jedoch bei der Anwendung auf echte Daten in jedem Fall ein diskretes

Messniveau vorliegt, besteht eine positive Wahrscheinlichkeit, dass zwei Vergleichsstatistiken $s_t(k, k^*)$ und $s_t(k^{**}, k^{***})$ den gleichen Wert aufweisen. Folglich ist nicht garantiert, dass die Zuweisung der Komponenten auf die Blöcke eindeutig ist. Um eine eindeutige Zuweisung zu gewährleisten, wird das Tupel (k, k^*) in Schritt 1 bzw. die Komponente k in Schritt 5 ggf. aus den in Frage kommenden Kandidaten zufällig ausgewählt.

6.1.2 Die TRM-LS-Regression des mSCARM

Ist die Blockbildung erfolgt, geschieht in Schritt 3 des mSCARM-Algorithmus die Signalschätzung innerhalb der einzelnen Blöcke. Besteht ein Block j aus zwei oder mehr Komponenten, wird hierzu eine leicht modifizierte Version der multivariaten TRM-LS-Regression nach Lanius und Gather (2010) durchgeführt. Diese Version nutzt die RM-Regressionsschätzungen aus Schritt 1 (d) des mSCARM-Algorithmus sowie die OGK- $Q_{0.5}^{\text{adj}}$ -Schätzungen der lokalen Fehlerkovarianzen $\sigma_t(k, k^*)$ mit $k, k^* \in \{1, \dots, K\}$, die sich im Zuge der Schätzung von $\widehat{\text{Var}}[D_t(k, k^*)]$ in Schritt 2 des mSCARM-Algorithmus ergeben.

Für jeden Block j wird die TRM-LS-Regression im Zeitfenster der Blockfensterbreite $n_t(j) = \min\{n_t(k) : k \in B_t(j)\}$ wie folgt durchgeführt:

1. Bestimme für alle $k \in B_t(j)$ die letzten $n_t(j)$ Residuen der RM-Regression. Füge diese zusammen zu Residuenvektoren $\mathbf{r}_{t-n_t(j)+i} \in \mathbb{R}^{|B_t(j)|}$, $i = 1, \dots, n_t(j)$.
2. Füge die geschätzten lokalen Fehlerkovarianzen $\hat{\sigma}_t(k, k^*)$ mit $k, k^* \in B_t(j)$ zusammen zu der geschätzten *Block-Fehlerkovarianzmatrix* $\hat{\Sigma}_t(j) \in \mathbb{R}^{|B_t(j)| \times |B_t(j)|}$.
3. Bestimme

$$I_t(j) := \left\{ i = 1, \dots, n_t(j) : \mathbf{r}'_{t-n_t(j)+i} \hat{\Sigma}_t(j)^{-1} \mathbf{r}_{t-n_t(j)+i} \leq \chi_{|B_t(j)|}^2(0.95) \right\},$$

wobei $\chi_{|B_t(j)|}^2(0.95)$ das 0.95-Quantil der χ^2 -Verteilung mit $|B_t(j)|$ Freiheitsgraden ist.

4. Führe anhand der getrimmten Fensterstichprobe $\{(y_{t-n_t(j)+i}(k))_{k \in B_t(j)} : i \in I_t(j)\}$ eine multivariate LS-Regression durch und speichere die Niveaus der Regressionsgeraden zum Zeitpunkt t als Signalschätzungen $\hat{\mu}_t^{\text{TRM-LS}}(k) =: \hat{\mu}_t^{\text{mSCARM}}(k)$.

In Schritt 1 werden die Residuen der bereits geschätzten RM-Regressionsgeraden ermittelt. Die komponentenweisen RM-Regressionen erfolgen in Zeitfenstern individueller Länge $n_t(k)$. Für jede Komponente $k \in B_t(j)$ sind jedoch lediglich die Residuen zu den letzten $n_t(j)$, $n_t(j) \leq n_t(k)$, Zeitpunkten zu bestimmen. Dazu ließe sich für jede Komponente $k \in B_t(j)$ eine erneute RM-Regression im Zeitfenster der Länge $n_t(j)$ durchführen

und anschließend die Residuen bestimmen. Dies würde jedoch zum einen einen zusätzlichen Rechenaufwand bedeuten, und zum anderen können die bereits vorliegenden RM-Regressionsschätzungen in den Zeitfenstern der individuellen Längen $n_t(k)$ als adäquat angesehen werden, da diese mit dem SCARM gewonnen werden. Aus diesen Gründen wird keine erneute RM-Regression durchgeführt. Stattdessen werden lediglich die letzten $n_t(j)$ Residuen der bereits erfolgten individuellen RM-Regressionen bestimmt.

In Schritt 2 geschieht – im Gegensatz zu der in Kapitel 2.2.1 beschriebenen TRM-LS-Regression nach Lanius (2005) und Lanius und Gather (2010) – die Schätzung der Block-Kovarianzmatrix $\Sigma_t(j)$ nicht mit dem OGK- Q_n anhand der Residuen der RM-Regression, sondern direkt anhand der Daten mit dem OGK- Q_δ^{adj} mit $\delta = 0.5$. Dies spart Rechenzeit, da die Q_δ^{adj} -Schätzungen der Fehlervarianz bereits aus der Anwendung des SCARM vorliegen. Zudem ist der OGK- Q_δ^{adj} im Gegensatz zum OGK- Q_n nicht abhängig von der Regressionsanpassung und liefert bessere Schätzungen bei nichtlinearen Signalen, vgl. Kapitel 5.2.2.

Anhand der Residuenvektoren aus Schritt 1 und der geschätzten Block-Kovarianzmatrix aus Schritt 2 wird in Schritt 3 die Indexmenge $I_t(j)$ der „guten“ Beobachtungsvektoren $(y_{t-n_t(j)+1}(k))_{k \in B_t(j)}$ bestimmt. Dabei handelt es sich um diejenigen Beobachtungsvektoren, deren Residuenvektoren bzgl. der Block-Fehlerkovarianz eine nicht zu große Mahalanobis-Distanz aufweisen. Demnach ist der TRM-LS und somit der mSCARM zwar in der Lage, Ausreißer bzgl. der Block-Kovarianzstruktur zu erkennen – Ausreißer bzgl. der gesamten lokalen Kovarianzstruktur über alle K Dimensionen bleiben jedoch womöglich unerkannt.

6.1.3 Behandlung fehlender Werte

Zur Behandlung fehlender Werte bedient sich der mSCARM der in Kapitel 3.3.3 für den SCARM aufgestellten Regeln sowie des modifizierten Q_δ^{adj} -Skalenschätzers $Q_\delta^{\text{adj}*}$ aus (3.17) auf Seite 40. Bei der komponentenweisen Anwendung des SCARM in Schritt 1 des mSCARM-Algorithmus ist also die Schätzung der RM-Geraden und der Fehlervarianz garantiert, sofern zu den letzten n_{\min} Zeitpunkten mindestens ein Messwert gegeben ist und im Zeitfenster der Länge $\min\{n_t(k), r\}$ mindestens n_{\min} Messwerte vorliegen. Erfüllt eine Komponente k zu einem Zeitpunkt t diese beiden Bedingungen nicht, so bezieht der Algorithmus diese Komponente zu diesem Zeitpunkt nicht mit ein und liefert stattdessen als Signalschätzung $\hat{\mu}_t^{\text{mSCARM}}(k) = \text{NA}$. Auch die SCARM-basierten SSM-Statistiken werden auf $s_t(k, k^*) = \text{NA}$ gesetzt für alle $k^* \neq k$. Die Komponente k kann zum Zeitpunkt t also nicht als ähnlich zu einer anderen Komponente k^* bewertet werden.

Für alle Komponenten, welche die beiden Bedingungen zum Zeitpunkt t erfüllen, sind zwar genügend aktuelle Messwerte gegeben, fehlende Werte sind aber trotzdem möglich. Für diese Komponenten soll eine Signalschätzung erfolgen, wozu in den Schritten 2 und 3 des mSCARM-Algorithmus eine Schätzung der Fehlerkovarianzmatrix anhand des OGK- Q_δ^{adj} -Schätzers erfolgt. Fehlende Werte in den Beobachtungsvektoren bedeuten fehlende Werte in den Residuenvektoren, weshalb die OGK- Q_δ^{adj} -Schätzung der Fehlerkovarianzmatrix möglicherweise nicht erfolgen kann. Würden Residuenvektoren mit fehlenden Einträgen nicht in die OGK- Q_δ^{adj} -Schätzung einbezogen, ginge dies mit einem Informationsverlust einher. Daher werden fehlende Einträge in den Beobachtungsvektoren durch das entsprechende Niveau der in Schritt 1 (d) angepassten RM-Regressionsgeraden ersetzt. Die entsprechenden fehlenden Residuenwerte werden also durch den Wert null ersetzt. Dadurch ist die OGK- Q_δ^{adj} -Schätzung der Fehlerkovarianzmatrix gesichert; allerdings ist diese offensichtlich in Richtung null verzerrt. Eine Implosion wird jedoch durch die untere Schranke $b > 0$ für die Q_δ^{adj} -Skalenschätzung gemäß (3.22) auf S. 50 verhindert, und die Robustheit gegenüber Ausreißern ist durch diese Ersetzung fehlender Werte nicht vermindert.

6.2 Anwendung auf echte Daten

Im Folgenden wird der mSCARM auf multivariate Zeitreihen aus dem intensivmedizinischen Online-Monitoring angewendet, um den praktischen Nutzen dieser Filterprozedur zu vermitteln. Intensivmedizinische Online-Monitoring-Zeitserien sind bereits in den Kapiteln 2.2 und 4.4.1 betrachtet worden. Die Anwendung erfolgt hier retrospektiv mit Hilfe der R-Funktion `mscarm.filter`, die in Kapitel 6.3 vorgestellt wird. Die Anwendung des mSCARM in Echtzeit würde zu den selben Ergebnissen führen.

Abbildung 6.1 zeigt eine sechsdimensionale Zeitreihe $(\mathbf{y}_t)_{t=1,\dots,1500}$, des sekundlich gemessenen systolischen, mittleren und diastolischen arteriellen Blutdrucks (ART.S, ART.M und ART.D, in mmHg) sowie des Puls, der Herzfrequenz (PULS und HR, in Schläge pro Minute) und der Sauerstoffsättigung des Blutes (SpO2, in %). Die Zeitreihe der Herzfrequenz ist zur besseren Sichtbarkeit um 30 Einheiten nach oben verschoben. Die Zeitreihen sind offensichtlich nicht stationär, vielmehr sind Sprünge sowie wechselnde Trends und Variabilitäten zu erkennen. Zudem weisen die drei Blutdruckvariablen zu den Zeitpunkten $t = 694, \dots, 698$ deutliche Ausreißer nach unten auf. Des Weiteren sind aufgrund eines technischen Fehlers zu den Variablen SpO2 und PULS zu den Zeitpunkten $t = 819, \dots, 854$ keine Messwerte gegeben.

Da es sich bei den Messwerten um ganzzahlige Beobachtungen handelt, weisen die Fens-

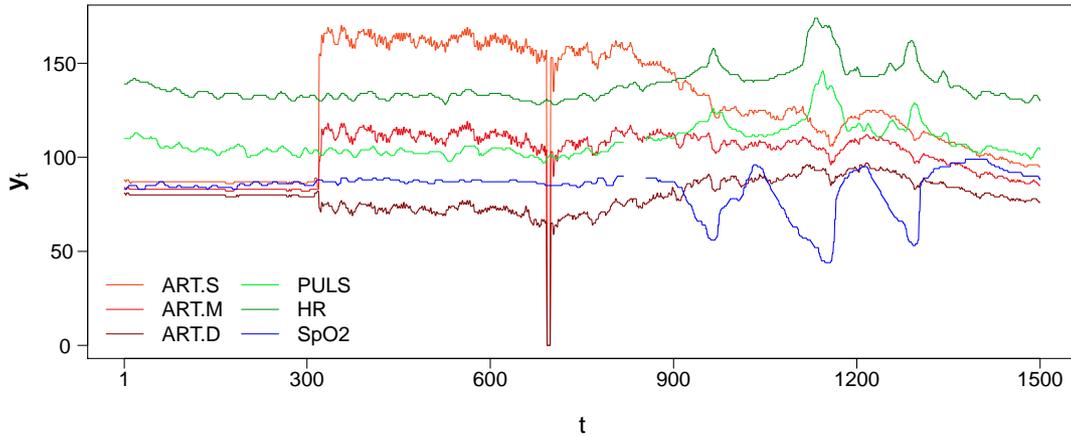


Abbildung 6.1: Multivariate Zeitreihe der sekundlich gemessenen Vitalparameter ART.S, ART.M, ART.D, PULS, HR und SpO2 (systolischer, mittlerer und diastolischer arterieller Blutdruck, Puls, Herzfrequenz und Sauerstoffsättigung des Blutes).

terstichproben häufig Bindungen auf, wodurch der Q_δ^{adj} -Skalenschätzer nach unten verzerrt ist. In der Folge weisen die Statistiken des SCARM-Tests sowie des SSM-Verfahrens stark erhöhte Variabilitäten auf. Dadurch kommt es zum einen vermehrt zu fälschlicherweise erkannten Signalveränderungen und somit zu unnötigen Verkleinerungen der Fensterbreiten. Zum anderen wird zwischen augenscheinlich ähnlich verlaufenden Komponenten der Zeitreihe oftmals kein Zusammenhang erkannt. Derartige Rundungseffekte könnten durch sog. *Verrauschen* oder *Verwackeln* (engl. *wobbling*, Fried und Gather, 2007) aufgehoben werden. Dabei wird zu jedem univariaten Messwert $y_t(k)$, $k = 1, \dots, K$, ein aus einer bestimmten Verteilung generierter Fehler addiert. Die Fehlervarianz sollte unter Berücksichtigung der Variabilität des Signals und der Rundungsgenauigkeit gewählt werden. In dieser Anwendung könnte jedem Messwert z.B. ein zufälliger Fehler aus einer Standardnormalverteilung hinzugefügt werden. Die Wahl der Untergrenze $b = 1$ für den Q_δ^{adj} -Skalenschätzer gemäß (3.22) auf S. 50 hat allerdings einen ähnlichen Effekt und bedeutet keine nachträgliche Datenmodifikation. In dieser Anwendung wird daher $b = 1$ verwendet.

Weiterhin erfolgt die Anwendung des mSCARM unter Verwendung der Inputargumente $n_{\min} = 10$, $n_{\max} = 200$, $r = \ell_{\min} = 30$ und $\alpha = 0.001$ für die Fensterbreitenadaptation des SCARM. Für die SSM-Prozedur zur automatischen Blockbildung bzw. zur Überwachung der Zusammenhänge sind $c_{\text{bound}} = 4$ und $r_{\text{bound}} = 0.25$ gewählt. Da die Zeitreihen hohe positive Autokorrelationen aufweisen, werden zu jedem Zeitpunkt die Konstanten c_n^φ und

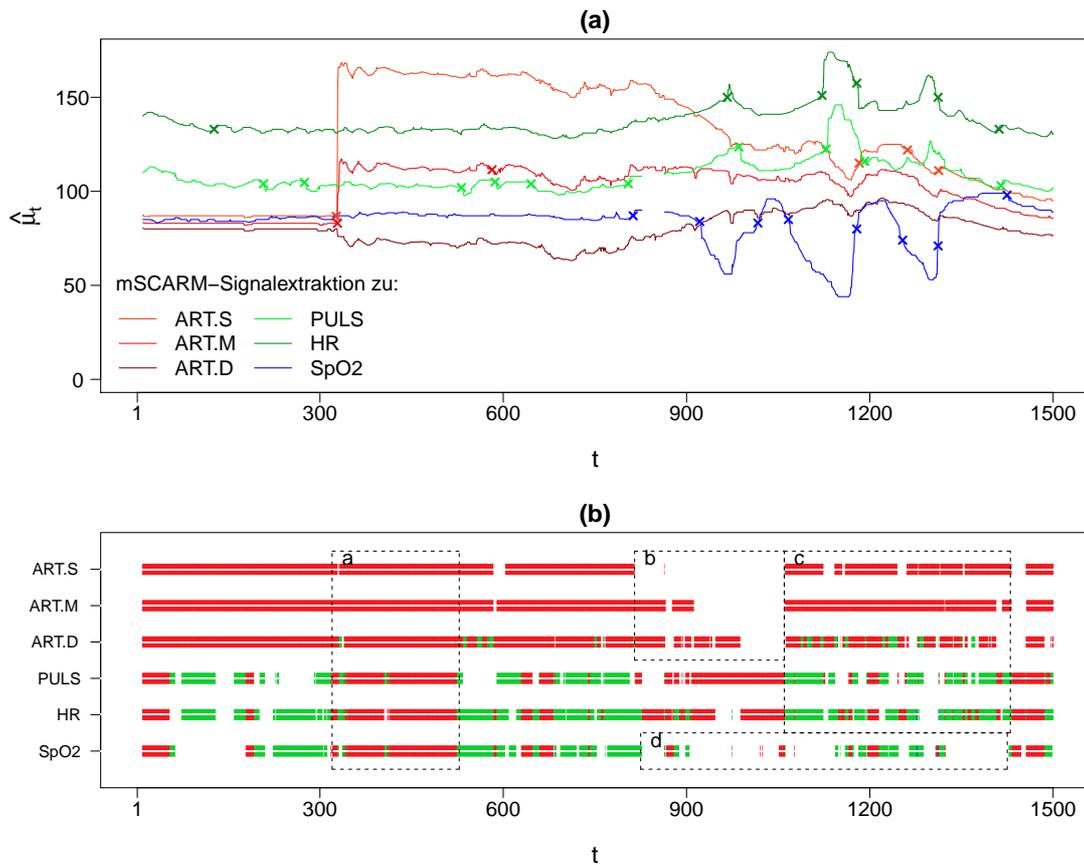


Abbildung 6.2: (a) mSCARM-Signalextraktion aus der multivariaten Zeitreihe aus Abbildung 6.1; (b) per SSM-Verfahren bestimmte Blöcke aus ähnlich verlaufenden Zeitreihenkomponenten; den Variablen eines Blockes wird dieselbe Farbe zugeordnet.

v_n^φ für die Schätzung der Varianz der RM-Steigungen gemäß Regel (3.20) anhand der aktuell geschätzten Autokorrelation der Fehler bestimmt, vgl. S. 48f. Dadurch werden die SCARM- und SSM-Statistiken unter Berücksichtigung der Autokorrelation der Fehler ermittelt.

Abbildung 6.2 (a) zeigt die multivariate Zeitreihe der mSCARM-Signalextraktion. Diese gibt die relevante Information der multivariaten Datenzeitreihe in Abbildung 6.1 in Form eines glatten multivariaten Signals wieder. Dabei werden die Signalschätzungen nicht durch die Ausreißer in den Blutdruckmessungen zu den Zeitpunkten $t = 694, \dots, 698$ gestört. Da für die Variablen PULS und SpO2 zu den Zeitpunkten $t = 819, \dots, 854$ keine Messwerte vorliegen, erfolgt nach dem Zeitpunkt $t = 819 + n_{\min} - 1 = 828$ für diese Variablen zunächst keine Signalschätzung mehr. Zum Zeitpunkt $t = 855 + n_{\min} - 1 = 864$ sind wieder n_{\min} ak-

tuelle Messwerte gegeben, weshalb der mSCARM zu diesem Zeitpunkt die Signalschätzung wieder aufnimmt.

Durch die Kreuze in Abbildung 6.2 (a) werden die erkannten Signalveränderungen gekennzeichnet. Diese können größtenteils als richtig positive Entscheidungen des SCARM-Tests bewertet werden. Allerdings werden auch einige Sprünge und Trendwechsel in den Datenzeitreihen nicht angezeigt, d.h. der SCARM-Test trifft einige falsch negative Entscheidungen. So wird z.B. für die Variable SpO2 gegen $t = 1000$ ein deutlicher Trendwechsel nicht angezeigt. Oftmals werden selbst deutliche Signalveränderungen nicht erkannt, weil kurz zuvor bereits eine Signalveränderung entdeckt und infolgedessen die Fensterbreite auf den Minimalwert n_{\min} gesetzt wurde. Dies ist auch für die Variable SpO2 gegen $t = 1000$ der Fall. Solange die Fensterbreite nicht wieder auf den Wert $r + \ell_{\min}$ angewachsen ist, wird kein Test durchgeführt. Der mSCARM und SCARM sind also nicht in der Lage, sehr kurz aufeinander folgende Signalveränderungen zu erkennen. Allerdings hat dieser Nachteil kaum einen Effekt auf die Genauigkeit der Signalextraktion, da die Fensterbreite in diesem Fall ohnehin klein ist, d.h. kleiner als $r + \ell_{\min}$.

Abbildung 6.2 (b) gibt die zu jedem Zeitpunkt anhand des SSM-Verfahrens ermittelten Blöcke aus ähnlich verlaufenden Zeitreihenkomponenten wieder. Dabei wird allen Variablen eines Blockes $B_t(j)$ die selbe Farbe zugeordnet. Weist eine Variable zu einem Zeitpunkt t keinen Farbwert auf, so wird für diese zu diesem Zeitpunkt kein Zusammenhang zu einer weiteren Variablen erkannt. Diese Darstellungsweise vermittelt dem Anwender auf einfache Weise die aktuell bestehenden Zusammenhänge zwischen den einzelnen Komponenten sowie den zeitlichen Verlauf der Zusammenhänge. In Abbildung 6.2 (b) sind vier Situationen bzw. Zeiträume a–d durch gestrichelte Rechtecke gekennzeichnet. Im Zeitraum a wird zu den meisten Zeitpunkten ein Block aus allen Variablen gebildet, d.h. in diesem Zeitraum werden alle Zeitreihen überwiegend als ähnlich bewertet. Tatsächlich zeigen alle Zeitreihen im Zeitraum a überwiegend einen stabilen Verlauf ohne erkennbare Trends.

Die Zeitreihen zu ART.S, ART.M und ART.D werden richtigerweise fast durchgehend demselben Block zugeordnet, d.h. bis auf wenige Ausnahmen wird für diese Zeitreihen ein Zusammenhang erkannt. Eine dieser Ausnahmen ist in Situation b gegeben. Hier werden die einzelnen Blutdruckzeitreihen kurzzeitig als unähnlich bewertet. Tatsächlich weisen die drei Datenzeitreihen der Blutdruckwerte in diesem kurzen Zeitraum ungleiche Trends auf, d.h. ART.S zeigt einen negativen und ART.D einen positiven Trend, während ART.M keinen erkennbaren Trend aufweist.

Im Zeitraum c werden überwiegend zwei Blöcke gebildet, wovon der eine die Herzfrequenz

und den Puls und der andere die drei Blutdruckarten umfasst. In diesem Zeitraum zeigen die Zeitreihen innerhalb der Blöcke tatsächlich sehr ähnliche Verläufe, ohne dass Ähnlichkeiten zu Zeitreihen außerhalb des Blocks zu erkennen ist.

In Situation d zeigt die Zeitreihe der Sauerstoffsättigung des Blutes SpO2 deutliche Niveau- und Trendwechsel, welche in den übrigen Zeitreihen in dieser Form nicht zu erkennen sind. Folglich bildet die Variable SpO2 in diesem Zeitraum d meistens einen eigenständigen Block, d.h. für diese Variable wird überwiegend kein Zusammenhang zu einer anderen Variablen angezeigt.

Die hier erfolgte Anwendung des mSCARM verdeutlicht den praktischen Nutzen dieser Methodik. Der mSCARM extrahiert die unterliegenden Signale und zeigt Signalveränderungen sowie die aktuell bestehenden Zusammenhänge zwischen den einzelnen Variablen an. Durch die gewählten Darstellungsweisen können dem Anwender die extrahierten Informationen schnell und leicht nachvollziehbar vermittelt werden.

6.3 Diskussion

Der mSCARM vereint die in dieser Arbeit entwickelten Verfahren SCARM und SSM und liefert für multivariate Zeitreihen bzw. Datenströme in Echtzeit

- die Extraktion der unterliegenden Signale,
- die Schätzung der Fehlervariabilität für jede univariate Zeitreihenkomponente,
- die Erkennung und Anzeige von Sprüngen und Trendwechseln in den univariaten Zeitreihenkomponenten,
- Aufschluss über aktuell bestehende Zusammenhänge im Sinne ähnlich verlaufender Zeitreihenkomponenten.

Der mSCARM stellt somit eine Prozedur zur umfassenden Echtzeit-Extraktion relevanter Information aus instationären multivariaten Zeitreihen dar. Eine R-Funktion des mSCARM, bezeichnet mit `mscarm.filter`, ist enthalten in dem R-Paket `robfilter` (Fried et al., 2012), welches auf dem CRAN-Server zur Verfügung steht:

<http://cran.r-project.org/web/packages/robfilter/>.

Die Funktion `mscarm.filter` ermöglicht die Wahl der Inputargumente des SCARM und des SSM-Verfahrens, für welche die folgenden Entsprechungen gelten:

$$\begin{aligned}
 \ell_{\min} &\hat{=} \text{min.left.width}, & r &\hat{=} \text{right.width}, \\
 n_{\min} &\hat{=} \text{min.width}, & n_{\max} &\hat{=} \text{max.width}, \\
 \alpha &\hat{=} \text{sign.level}, & b &\hat{=} \text{bound.noise.sd}, \\
 c_{\text{bound}} &\hat{=} \text{c.bound}, & r_{\text{bound}} &\hat{=} \text{r.bound}.
 \end{aligned}$$

Die R-Funktion `mscarm.filter` ermöglicht zudem die Anwendung der Restrict-to-Range-Regel mit der Option `rtr=TRUE`, vgl. (2.12) auf S. 17.

Die Funktion `mscarm.filter` erfordert lediglich geringe Rechenzeiten, da sie für den RM und den Q_{δ}^{adj} -Skalenschätzer den recheneffizienten Update-Algorithmus aus Kapitel 3.3.2 verwendet und die Q_{δ}^{adj} -Schätzung der Fehlervarianz zur Schätzung der Fehlerkovarianz mittels OGK- Q_{δ}^{adj} genutzt wird. So benötigt ein Computer mit AMD Athlon 64X2 Dual Core Processor 4400+ mit 2.3 GHz und 1.96 GB RAM rund 350 Sekunden für die Anwendung der R-Funktion `mscarm.filter` auf die sechsdimensionale Zeitreihe der Länge $T = 1500$ aus dem vorherigen Kapitel 6.2. In einer Echtzeit-Anwendung wären demnach zu jedem Messzeitpunkt durchschnittlich rund 0.23 Sekunden Rechenzeit erforderlich. Bei einer Messfrequenz von einem Beobachtungsvektor pro Sekunde wäre der mSCARM somit echtzeitfähig, auch weil Programmabbrüche aufgrund fehlender Beobachtungen ausgeschlossen sind, vgl. Kapitel 6.1.3.

Der mSCARM gründet auf dem SCARM und der SSM-Methodik und basiert daher auf der Annahme zeitlich unabhängiger normalverteilter Fehler, vgl. die Diskussionen zum SCARM und SSM in den Kapiteln 3.5 und 5.2.4. Dabei ist die Robustheit gegenüber Ausreißern gegeben, da die Verfahren auf der robusten RM-Regression sowie den robusten Schätzern Q_{δ}^{adj} und OGK- Q_{δ}^{adj} der Fehlervarianz bzw. -kovarianz basieren. Des Weiteren kann der mSCARM an autokorrelierte Fehler angepasst werden, indem die zur Schätzung der Varianz der RM-Steigungen benötigten Konstanten c_n und v_n entsprechend gewählt werden. Für Fehler, die sich durch einen AR(1)-Prozess mit Parameter φ beschreiben lassen, können die Konstanten c_n^{φ} und v_n^{φ} (automatisch) in Abhängigkeit vom (zum Zeitpunkt t geschätzten) Parameter φ gewählt werden, vgl. Kapitel 3.5. Hierzu bietet die R-Funktion `mscarm.filter` wie die R-Funktionen `scarm.filter` und `SSM` das Inputargument `autocorrelations`, vgl. S. 50 und S. 103.

Die Anwendung des mSCARM auf multivariate Zeitreihen aus der intensivmedizinischen Patientenüberwachung verdeutlicht den Nutzen dieser Methodik für die Praxis. Mittels übersichtlicher Darstellungsweisen kann die R-Funktion `mscarm.filter` die extrahierten relevanten Informationen schnell und einfach vermitteln und somit zur Entscheidungsfindung des Anwenders beitragen. Darüber hinaus kann insbesondere durch die Echtzeit-Signalextraktion des mSCARM die Leistung intensivmedizinischer Online-Monitoring-Systeme verbessert werden. Beim intensivmedizinischen Online-Monitoring werden für jeden Vitalparameter Ober- und Untergrenzen festgelegt; liegen die Messwerte außerhalb dieser Grenzen, ertönt ein Alarmsignal. Aufgrund von Ausreißern besteht eine hohe Rate an falsch positiven Alarmen (vgl. z.B. Chambrin et al., 1999; Tsien und Fackler, 1997), was zu einer Desensibilisierung des klinischen Personals führt. Eine Echtzeit-Anwendung des mSCARM kann zu einer Fehlalarmreduktion führen, wenn statt der Messwerte die extrahierten Signale mit den Alarmgrenzen verglichen werden (Borowski et al., 2011a). In einer Evaluationsstudie zeigen Borowski et al. (2011b), dass die Echtzeit-Anwendung des aoRM und aoTRM-LS bei einer Sensitivität von rund 95% eine Fehlalarmreduktion von etwa 25–30% bewirken würde. Aktuell führt cand. B.Sc. Anika Rottmann am Institut für Mathematische Statistik und industrielle Anwendungen der Technischen Universität Dortmund im Rahmen ihrer Bachelorarbeit eine vergleichbare Evaluationsstudie für den SCARM und mSCARM durch. Die Ergebnisse dieser Studie stehen jedoch noch aus.

7 Zusammenfassung

In dieser Arbeit werden neue uni- und multivariate robuste Verfahren zur Echtzeit-Extraktion von Signalen sowie zur Echtzeit-Überwachung von Zusammenhängen in multivariaten Zeitreihen entwickelt. Die Verfahren sind speziell für hochfrequent gemessene instationäre Datenströme konzipiert, die wechselnde Fehlervariabilitäten, Ausreißer, fehlende Messwerte, wechselnde Trends und Sprünge aufweisen. Alle in dieser Arbeit entwickelten Methoden gründen auf der robusten *Repeated Median*- (RM) Regression.

In Kapitel 2 werden bestehende uni- und multivariate Verfahren zur Echtzeit-Signalextraktion vorgestellt und diskutiert. Der Fokus liegt dabei auf Methoden, die auf der RM-Regression in gleitenden Zeitfenstern basieren. Der univariate *adaptive online Repeated Median* (aoRM) adaptiert die Fensterbreite zu jedem Zeitpunkt an die aktuelle Datenlage. Er verwendet dazu einen Test auf Anpassung der RM-Regressionsgeraden, der die Ausgewogenheit der Residuenvorzeichen bewertet. Der multivariate *adaptive online TRM-LS* (aoTRM-LS) ergibt sich aus der Kombination des aoRM mit dem multivariaten *Trimmed RM-Least Squares* (TRM-LS). Der aoTRM-LS verwendet den Test des aoRM, um zu jedem Zeitpunkt eine Fensterbreite zu finden, die für alle K univariaten Komponenten der multivariaten Zeitreihe gültig ist. Die Fensterbreitenadaptation des aoTRM-LS funktioniert jedoch nicht zufriedenstellend. Ansätze zur Verbesserung bieten ein alternatives Testverfahren zur Fensterbreitenadaptation, das bei niedrigem Signifikanzniveau eine hohe Güte aufweist, sowie die zeitabhängige, automatische Einteilung der K Komponenten in disjunkte Blöcke aus aktuell ähnlich verlaufenden Komponenten.

In Kapitel 3 wird der *Slope Comparing Adaptive RM* (SCARM) als Alternative zum aoRM entwickelt. Dieser testet die Linearität des Signals in der Fensterstichprobe. Ist diese nicht gegeben, wird eine Signalveränderung angenommen und die Fensterbreite verkleinert. Es wird eine Methodik zur Schätzung der Varianz der Steigungsdifferenz entwickelt, die auf dem robusten, modellfreien Q_δ^{adj} -Schätzer der Fehlervariabilität gründet und daher Masking-Effekte verhindert bzw. die Güte des SCARM-Tests erhöht. Weiterhin wird eine recheneffiziente Programmierung des SCARM präsentiert. Zudem werden Regeln für die Behandlung fehlender Werte in Echtzeit aufgestellt, die einen Output in jeder Datensituation garantieren. Hierzu wird eine modifizierte Version des Q_δ^{adj} -Schätzers entwickelt, die bei fehlenden Werten effizientere Signalschätzungen liefert und einen nicht kleineren Bruchpunkt aufweist. Mittels Simulationen wird die Verteilung der SCARM-Teststatistik unter

H_0 für verschiedene Fehlerarten untersucht. Auf Basis dieser Studie werden als kritische Werte zur Testentscheidung je nach Fensterbreite die Quantile einer t - bzw. Standardnormalverteilung verwendet.

Ein Vergleich des aoRM und SCARM erfolgt in einer umfassenden Studie in Kapitel 4. Die *Resistances to Rejection and Acceptance* (RR und RA) des aoRM- und SCARM-Tests werden bestimmt und verglichen. Die Tests bieten ein vergleichbares Maß an Robustheit gegenüber Ausreißern. Der Test des aoRM weist für ein kleines Signifikanzniveau eine hohe RR aber eine geringe RA auf, während die Resistances des SCARM-Tests unabhängig vom gewählten Signifikanzniveau sind. Ein Vergleich des aoRM und SCARM bzgl. der Erkennung von Signalveränderungen deckt die Vorteile des SCARM auf, da dieser in den meisten Fällen eine höhere Güte bietet und Signalveränderungen schneller erkennt. In einer Simulationsstudie wird die Effizienz der Signalschätzung des aoRM und SCARM verglichen. Hier liefert der SCARM in jedem Szenario effizientere Signalschätzungen. Die gewonnenen Erkenntnisse werden durch die Anwendung des aoRM und SCARM auf hochfrequent erfasste Zeitreihen von Blutdruckmessungen und Aktienpreisen bestätigt.

In Kapitel 5 wird zunächst ein Überblick über bestehende Verfahren gegeben, die zur Echtzeit-Überwachung der Zusammenhänge zwischen den einzelnen Komponenten einer multivariaten Zeitreihe in Frage kommen. Anschließend wird das neue *Similar Slope Monitoring*- (SSM) Verfahren entwickelt, das auf dem SCARM basiert und den Zusammenhang zwischen zwei univariaten Zeitreihen zum Zeitpunkt t anhand der aktuellen Steigungsschätzungen bewertet. Da die Varianz der Differenzen der Steigungsschätzungen selbst bei einem linearem Signal und konstanter Fehlervarianz über die Zeit nicht konstant ist, wird eine alternative Methodik zur Schätzung dieser Varianz entwickelt. Um Erkenntnisse bzgl. der Bewertung der SSM-Statistik zu erlangen, wird diese in sechs Szenarien ähnlicher und unähnlicher bivariater Zeitreihen angewendet.

In Kapitel 6 wird ein multivariater Filter mit adaptiver Fensterbreitenwahl, der *multivariate SCARM* (mSCARM) entwickelt. Dieser ergibt sich aus der Kombination der zuvor entwickelten Verfahren SCARM und SSM mit dem multivariaten TRM-LS. Der mSCARM nutzt den SCARM-Test zur Fensterbreitenadaptation, bildet zu jedem Zeitpunkt auf Basis des SSM Blöcke aus aktuell ähnlich verlaufenden Komponenten und schätzt das Signal in Blöcken mit zwei oder mehr Komponenten mit dem TRM-LS. Zur Blockbildung wird ein Algorithmus entwickelt, der möglichst große Blöcke unter der Bedingung der paarweisen Ähnlichkeit der Blockkomponenten erstellt. Um geringe Rechenzeiten zu gewährleisten wird zudem eine leicht modifizierte Version des TRM-LS vorgeschlagen, welche die vom

SCARM gelieferten Schätzungen der Fehlervarianz zur Schätzung der Fehlerkovarianz verwendet. Zur Behandlung fehlender Werte bedient sich der mSCARM der zuvor erstellten Prinzipien des SCARM und des modifizierten Q_δ^{adj} -Skalenschätzers $Q_\delta^{\text{adj}*}$. Die Anwendung des mSCARM auf multivariate Zeitreihen aus der intensivmedizinischen Patientenüberwachung zeigt den praktischen Nutzen der Methodik. So liefert der mSCARM in Echtzeit eine Extraktionen des multivariaten Signals und zeigt zudem Signalveränderungen sowie die aktuell bestehenden Zusammenhänge zwischen den einzelnen Variablen an. Darüber hinaus werden für jede Variable zu jedem Zeitpunkt Schätzungen der Fehlervariabilität ausgegeben. Der mSCARM stellt somit eine robuste Prozedur zur umfassenden Extraktion relevanter Information aus instationären multivariaten Zeitreihen in Echtzeit dar.

Im Rahmen dieser Arbeit sind R-Funktionen zu den entwickelten Verfahren SCARM, SSM und mSCARM erstellt worden. Der SCARM und mSCARM sind in dem R-Paket `robfilter` als Funktion `scarm.filter` bzw. `mscarm.filter` enthalten, und die R-Funktion `SSM` steht auf der Webseite

<http://www.statistik.tu-dortmund.de/1542.html>

des Instituts für Mathematische Statistik und industrielle Anwendungen der Technischen Universität Dortmund zum Download zur Verfügung. Die R-Funktionen eignen sich zwar lediglich für die retrospektive Anwendung auf abgeschlossene Datensätze, die Algorithmen können jedoch mit wenig Aufwand in echtzeitfähige Systeme implementiert werden.

Für die in dieser Arbeit entwickelten Verfahren der Echtzeit-Informationsextraktion bietet sich eine Fülle möglicher Anwendungen. So lässt sich der SCARM-Test auch zur Echtzeit-Erkennung von Strukturbrüchen in Form von Sprüngen und/oder Trendwechseln anwenden, z.B. bei der Überwachung thermischer Beschichtungsprozesse (Rudak et al., 2012). Bei der Beurteilung von hochfrequent gemessenen multivariaten Finanzzeitreihen kann das SSM-Verfahren aktuell bestehende Zusammenhänge und Veränderungen anzeigen. Des Weiteren können der SCARM und mSCARM zur Verbesserung der intensivmedizinischen Patientenüberwachung eingesetzt werden, insbesondere zur Reduktion von Fehlalarmen (Borowski et al., 2011a,b).

Literatur

- Arnold, M., Bissantz, N., Wied, D. und Ziggel, D. (2010). *A new online-test for changes in correlations between assets*. Discussion Paper 34/10, SFB 823, Technische Universität Dortmund.
- Aue, A., Hörmann, S., Horváth, L. und Reimherr, M. (2009). Break detection in the covariance structure of multivariate time series models. *The Annals of Statistics* **37**(6B), S. 4046–4087.
- Beringer, J. und Hüllermeier, E. (2006). Online clustering of parallel data streams. *Data and Knowledge Engineering* **58**, S. 180–204.
- Bernholt, T. und Fried, R. (2003). Computing the update of the repeated median regression line in linear time. *Information Processing Letters* **88**(3), S. 111–117.
- Bernholt, T., Fried, R., Gather, U. und Wegener, I. (2006). Modified repeated median filters. *Statistics and Computing* **16**, S. 177–192.
- Bodnar, O. und Schmid, W. (2007). Surveillance of the mean behavior of multivariate time series. *Statistica Neerlandica* **61**(4), S. 383–406.
- Borowski, M. und Fried, R. (2011). *Robust Repeated Median regression in moving windows with data-adaptive width selection*. Discussion Paper 28/11, SFB 823, Technische Universität Dortmund.
- Borowski, M. und Fried, R. (2012). *Online signal extraction by robust regression in moving windows with data-adaptive width selection*. In Revision.
- Borowski, M., Görges, M., Fried, R., Such, O., Wrede, C. und Imhoff, M. (2011a). Medical device alarms. *Biomedizinische Technik* **56**, S. 73–83.
- Borowski, M., Schettlinger, K. und Gather, U. (2009). Multivariate real time signal processing by a robust adaptive regression filter. *Communications in Statistics – Simulation and Computation* **38**(2), S. 426–440.
- Borowski, M., Siebig, S., Wrede, C. und Imhoff, M. (2011b). Reducing false alarms of intensive care online-monitoring systems: An evaluation of two signal extraction algorithms. *Computational and Mathematical Methods in Medicine* **2011**.

- Brillinger, D. (1981). *Time Series. Data Analysis and Theory*. Holden-Day, San Francisco.
- Bulut, A. und Singh, A. (2005). A unified framework for monitoring data streams in real time. In: *Proceedings of the 21st International Conference on Data Engineering (ICDE 2005)*.
- Busse, D. (2012). *Robuste Echtzeit-Überwachung der Abhängigkeiten multivariater nichtstationärer Zeitreihen*. Diplomarbeit, Fakultät Statistik, Technische Universität Dortmund.
- Chambrin, M., Calvelo-Aros, D., Jaborska, A., Chopin, C., Ravoux, P. und Boniface, B. (1999). Multicentric study at monitoring alarms in the adult intensive care unit (ICU): a descriptive analysis. *Intensive Care Medicine* **25**(12), S. 1360–1366.
- Chan, L. und Zhang, J. (2001). Cumulative sum control charts for the covariance matrix. *Statistica Sinica* **11**, S. 767–790.
- Cleveland, W. (1979). Robust locally weighted regression and smoothing scatterplots. *Journal of the American Statistical Association* **74**, S. 829–836.
- Cole, R., Shasha, D. und Zhao, X. (2005). Fast window correlations over uncooperative time series. In: *Proceedings of the eleventh ACM SIGKDD international conference on Knowledge discovery in data mining, KDD '05*, S. 743–749.
- Croux, C., Rousseeuw, P. und Hössjer, O. (1994). Generalized S-estimators. *Journal of the American Statistical Association* **89**(428), S. 1271–1281.
- Dai, B., Huang, J., Yeh, M. und Chen, M. (2006). Adaptive clustering for multiple evolving streams. *IEEE Transactions on Knowledge and Data Engineering* **18**(9), S. 1166–1180.
- Davies, P., Fried, R. und Gather, U. (2004). Robust signal extraction for on-line monitoring data. *Journal of Statistical Planning and Inference* **122**, S. 65–78.
- Davies, P. und Gather, U. (2005). Breakdown and groups. *The Annals of Statistics* **33**(3), S. 977–1035.
- Davies, P. und Gather, U. (2006). Addendum to the discussion of 'Breakdown and groups'. *The Annals of Statistics* **34**(3), S. 1577–1579.
- Davies, P. und Gather, U. (2007). The breakdown point – Examples and counterexamples. *Revstat – Statistical Journal* **5**, S. 1–17.

- Donoho, D. und Huber, P. (1983). The notion of breakdown point. In: Bickel, P., Doksum, K. und Hodges, J. (Hrsg.), *A Festschrift for Erich L. Lehmann*, S. 157–184. Wadsworth, Belmont, California.
- Fried, R. (2004). Robust filtering of time series with trends. *Nonparametric Statistics* **16**(3-4), S. 313–328.
- Fried, R. (2007). On the robust detection of edges in time series filtering. *Computational Statistics and Data Analysis* **52**, S. 1063–1074.
- Fried, R., Einbeck, J. und Gather, U. (2007). Weighted repeated median smoothing and filtering. *Journal of the American Statistical Association* **102**, S. 1300–1308.
- Fried, R. und Gather, U. (2007). On rank tests for shift detection in time series. *Computational Statistics and Data Analysis* **52**, S. 221–233.
- Fried, R., Schettlinger, K. und Borowski, M. (2012). *robfilter: Robust Time Series Filters*. R package version 4.0.
- Gather, U. und Fried, R. (2003). Robust estimation of scale for local linear temporal trends. *Tatra Mountains Mathematical Publications* **26**, S. 87–101.
- Gather, U. und Fried, R. (2004). Methods and algorithms for robust filtering. In: *Proceedings in Computational Statistics (COMPSTAT 2004)*, S. 159–170. Physica, Heidelberg.
- Gather, U., Imhoff, M. und Fried, R. (2002). Graphical models for multivariate time series from intensive care monitoring. *Statistics in Medicine* **21**, S. 2685–2701.
- Gather, U., Schettlinger, K. und Fried, R. (2006). Online signal extraction by robust linear regression. *Computational Statistics* **21**(1), S. 33–51.
- Gelper, S., Schettlinger, K., Croux, C. und Gather, U. (2009). Robust online scale estimation in time series: a model-free approach. *Journal of Statistical Planning and Inference* **139**, S. 335–349.
- Gnanadesikan, R. und Kettenring, J. (1972). Robust estimates, residuals, and outlier detection with multiresponse data. *Biometrics* **28**, S. 81–124.
- Hampel, F. (1971). A general qualitative definition of robustness. *The Annals of Mathematical Statistics* **42**(6), S. 1887–1896.

- Hampel, F. (1975). Beyond location parameters: Robust concepts and methods. *Bulletin of the International Statistical Institute* **46**, S. 375–382.
- Hodges, J. (1967). Efficiency in normal samples and tolerance of extreme values for some estimates of location. In: *Proceedings of the 5th Berkeley Symposium on Mathematical Statistics and Probability, Vol. 1*, S. 163–186.
- Hössjer, O., Rousseeuw, P. und Croux, C. (1994). Asymptotics of the repeated median slope estimator. *The Annals of Statistics* **22**(3), S. 1478–1501.
- Hössjer, O., Rousseeuw, P. und Ruts, I. (1995). The repeated median intercept estimator: Influence function and asymptotic normality. *Journal of Multivariate Analysis* **52**, S. 45–72.
- Idé, T., Papadimitriou, S. und Vlachos, M. (2007). Computing correlation anomaly scores using stochastic nearest neighbors. In: *Proceedings of the 7th IEEE International Conference on Data Mining 2007*, S. 523–528.
- Jiang, T., Feng, Y., Zhang, B., Cao, Z., Fu, G. und Shi, J. (2009). Monitoring correlative financial data streams by local pattern similarity. *Journal of Zhejiang University SCIENCE A* **10**(7), S. 937–951.
- Kano, M., Hasebe, S., Hashimoto, I. und Ohno, H. (2001). A new multivariate statistical process monitoring method using principal component analysis. *Computers and Chemical Engineering* **25**, S. 1103–1113.
- Lanius, V. (2005). *Statistische Extraktion relevanter Information aus multivariaten Online-Monitoring-Daten der Intensivmedizin*. Dissertation, Fakultät Statistik, Technische Universität Dortmund.
- Lanius, V. und Gather, U. (2010). Robust online signal extraction from multivariate time series. *Computational Statistics and Data Analysis* **54**, S. 966–975.
- Lee, Y. und Kassam, S. (1985). Generalized median filtering and related nonlinear filtering techniques. *IEEE Transactions on Acoustics, Speech, and Signal Processing* **33**, S. 672–683.
- Li, W., Yue, B., Valle-Cervantes, S. und Qin, S. (2000). Recursive PCA for adaptive process monitoring. *Journal of Process Control* **10**(5), S. 471–486.

- Liu, X. und Ferhatosmanoglu, H. (2003). Efficient k-NN search on streaming data series. In: Hadzilacos, T., Manolopoulos, Y., Roddick, J. und Theodoridis, Y. (Hrsg.), *Advances in Spatial and Temporal Databases*, Lecture Notes in Computer Science, S. 83–101. Springer, Berlin, Heidelberg.
- Maronna, A. und Zamar, R. (2002). Robust estimates of location and dispersion for high-dimensional data sets. *Technometrics* **44**(4), S. 307–317.
- Matoušek, J., Mount, D. und Netanyahu, N. (1998). Efficient randomized algorithms for the repeated median line estimator. *Algorithmica* **20**, S. 136–150.
- McDonald, J. und Owen, A. (1986). Smoothing with split linear fits. *Technometrics* **28**(3), S. 195–208.
- Papadimitriou, S., Sun, J. und Yu, P. (2006). Local correlation tracking in time series. In: *Proceedings of the Sixth International Conference on Data Mining (ICDM'06)*, S. 456–465.
- Peña, D. und Box, G. (1987). Identifying a simplifying structure in time series. *Journal of the American Statistical Association* **82**(399), S. 836–843.
- R Development Core Team (2011). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0.
- Rodrigues, P., Gama, J. und Pedroso, J. (2008). Hierarchical clustering of time-series data streams. *IEEE Transactions on Knowledge and Data Engineering* **20**(5), S. 615–627.
- Rousseeuw, P. (1984). Least median of squares regression. *Journal of the American Statistical Association* **79**(388), S. 871–880.
- Rousseeuw, P. (1985). Multivariate estimation with high breakdown point. In: Grossmann, W., Pflug, G., Vincze, I. und Wertz, W. (Hrsg.), *Proceedings of the 4th Pannonian Symposium on Mathematical Statistics and Probability*, volume B, Dordrecht. D. Reidel Publishing Company.
- Rousseeuw, P. und Croux, C. (1993). Alternatives to the median absolute deviation. *Journal of the American Statistical Association* **88**(424), S. 1273–1283.
- Rousseeuw, P. und Hubert, M. (1996). Regression-free and robust estimation of scale for bivariate data. *Computational Statistics and Data Analysis* **21**, S. 67–85.

- Rousseeuw, P. und Hubert, M. (1999). Regression Depth. *Journal of the American Statistical Association* **94**(446), S. 388–402.
- Rousseeuw, P. und Leroy, A. (1987). *Robust Regression and Outlier Detection*. Wiley, New York.
- Rousseeuw, P., Netanyahu, N. und Mount, D. (1993). *New statistical and computational results on the repeated median line*, S. 177–194. Birkhäuser, Basel.
- Rudak, N., Borowski, M., Wied, D., Hussong, B., Tillmann, W. und Kuhnt, S. (2012). *On- and off-line detection of structural breaks in thermal spraying processes (using recently developed methods)*. Preprint.
- Schettlinger, K. (2004). *Robust methods for signal extraction from time series*. Diplomarbeit, Fakultät Statistik, Technische Universität Dortmund.
- Schettlinger, K. (2009). *Signal and variability extraction for online monitoring in intensive care*. Dissertation, Fakultät Statistik, Technische Universität Dortmund.
- Schettlinger, K., Fried, R. und Gather, U. (2010a). Real time signal processing by adaptive repeated median filters. *International Journal of Adaptive Control and Signal Processing* **24**, S. 346–362.
- Schettlinger, K., Gelper, S., Gather, U. und Croux, C. (2010b). Regression-based, regression-free and model-free approaches for robust online scale estimation. *Journal of Statistical Computation and Simulation* **80**(9), S. 1023–1040.
- Siegel, A. (1982). Robust regression using repeated medians. *Biometrika* **69**(1), S. 242–244.
- Sliwa, P. und Schmid, W. (2005a). Monitoring the cross-covariances of a multivariate time series. *Metrika* **61**, S. 89–115.
- Sliwa, P. und Schmid, W. (2005b). Surveillance of the covariance matrix of multivariate nonlinear time series. *Statistics* **39**(3), S. 221–246.
- Spencer, J. (1904). On the graduation of the rates of sickness and mortality. *Journal of the Institute of Actuaries* **38**, S. 334–347.
- Stromberg, A., Hössjer, O. und Hawkins, D. (2000). The least trimmed differences regression estimator and alternatives. *Journal of the American Statistical Association* **95**(451), S. 853–864.

- Tsien, C. und Fackler, J. (1997). Poor prognosis for existing monitors in the intensive care unit. *Critical Care Medicine* **25**(4), S. 614–619.
- Tukey, J. (1977). *Exploratory data analysis*. Addison Wesley, Reading, Massachusetts.
- Wied, D. (2009). *Ein Fluktuationstest auf konstante Korrelation*. Dissertation, Fakultät Statistik, Technische Universität Dortmund.
- Wied, D. und Galeano, P. (2013). Monitoring correlation change in a sequence of random variables. *Journal of Statistical Planning and Inference* **143**(1), S. 186–196.
- Woolhouse, W. (1870). Explanation of a new method of adjusting mortality tables with some observations upon Mr. Makeham’s modification of Gompertz’s theory. *Journal of the Institute of Actuaries* **15**, S. 389–410.
- Yang, J. (2003). Dynamic clustering of evolving streams with a single pass. In: *Proceedings of the 19th International Conference on Data Engineering 2003*, S. 695–697.
- Yeh, M., Dai, B. und Chen, M. (2007). Clustering over multiple evolving streams by events and correlations. *IEEE Transactions on Knowledge and Data Engineering* **19**(10), S. 1349–1362.
- Ylvisaker, D. (1977). Test resistance. *Journal of the American Statistical Association* **72**(359), S. 551–556.
- Zhu, Y. und Shasha, D. (2002). Statstream: Statistical monitoring of thousands of data streams in real time. In: *Proceedings of the 28th international conference on Very Large Data Bases*, S. 358–369.

A Definitionen

Der Bruchpunkt

Der Bruchpunkt dient der Quantifizierung der globalen Robustheit von Funktionalen bzw. Schätzern. Die Bruchpunkt-Notation wurde erstmals von Hodges (1967) eingeführt und später von Hampel (1971) verallgemeinert. Weitere Diskussionen zum Konzept des Bruchpunkts sowie Verallgemeinerungen des Bruchpunkts liefern Davies und Gather (2005, 2006, 2007). In dieser Arbeit ist mit Bruchpunkt der *finite sample breakdown point* (fsbp) nach Donoho und Huber (1983) bezeichnet. Dieser stellt eine vereinfachte Version des Bruchpunkts für endliche Stichproben dar.

Sei $\mathbf{y}_n = (y_1, \dots, y_n)$ eine Stichprobe vom Umfang n und $T : \mathbb{R}^n \rightarrow \mathbb{R}^p$ mit $p \leq n$ ein Schätzer, der Werte innerhalb des Parameterraums Θ annimmt. Der Parameterraum sei mit einer Metrik D versehen, welche die Bedingung $\sup_{\theta_1, \theta_2 \in \Theta} D(\theta_1, \theta_2) = \infty$ erfüllt. Weiterhin sei \mathbf{y}_n^k eine Stichprobe, die sich durch Ersetzen von $k \in \{1, \dots, n\}$ Werten in \mathbf{y}_n ergibt. Der fsbp von T an der Stichprobe \mathbf{y}_n ist dann definiert als

$$\text{fsbp}(T, \mathbf{y}_n, D) = \min \frac{1}{n} \left\{ k \in \{1, \dots, n\} : \sup_{\mathbf{y}_n^k} D[T(\mathbf{y}_n), T(\mathbf{y}_n^k)] = \infty \right\}.$$

Der fsbp gibt also den kleinsten Anteil an Beobachtungen der Stichprobe \mathbf{y}_n an, der ausgetauscht werden muss, damit der Schätzer T einen Wert am Rand des Parameterraums Θ liefert. In diesem Fall wird der Schätzer T als zusammengebrochen bezeichnet.

Für die in dieser Arbeit betrachteten Regressionsschätzer zur univariaten Signalextraktion ist der Parameterraum $\Theta = \mathbb{R}^2$. Hier wird in der Regel die L_2 -Norm als Metrik D gewählt. Für einen Zusammenbruch des Regressionsschätzers sind die Beobachtungen aus \mathbf{y}_n dann durch k Werte zu ersetzen, die beliebig weit von der Stichprobe \mathbf{y}_n entfernt sind. Für Skalenschätzer S ist $\Theta = (0, \infty)$. Ein Skalenschätzer S ist also zusammengebrochen, wenn der Schätzer den Wert unendlich (Explosion) oder null (Implosion) annimmt. Für Skalenschätzer ist die Metrik $D(\theta_1, \theta_2) = |\log(\theta_1/\theta_2)|$ mit $\theta_1, \theta_2 \in (0, \infty)$ geeignet, da sie sowohl bei einer Explosion als auch bei einer Implosion den Wert ∞ liefert. Um eine Explosion herbeizuführen, wären die Beobachtungen aus \mathbf{y}_n durch k Werte zu ersetzen, die beliebig weit von der Stichprobe \mathbf{y}_n entfernt sind. Für eine Implosion sind k Beobachtungen aus \mathbf{y}_n durch kollineare Werte zu ersetzen. Der fsbp eines Skalenschätzers S an einer

Stichprobe \mathbf{y}_n kann daher definiert werden als

$$\begin{aligned} \text{fsbp}(S, \mathbf{y}_n) &= \min \{ \text{fsbp}^+(S, \mathbf{y}_n), \text{fsbp}^-(S, \mathbf{y}_n) \}, \\ \text{fsbp}^+(S, \mathbf{y}_n) &= \min \frac{1}{n} \left\{ k \in \{1, \dots, n\} : \sup_{\mathbf{y}_n^k} S(\mathbf{y}_n^k) = \infty \right\}, \\ \text{fsbp}^-(S, \mathbf{y}_n) &= \min \frac{1}{n} \left\{ k \in \{1, \dots, n\} : \inf_{\mathbf{y}_n^k} S(\mathbf{y}_n^k) = 0 \right\}. \end{aligned}$$

Dabei bezeichnen fsbp^+ und fsbp^- den Explosions- und Implosionsbruchpunkt.

Äquivarianzeigenschaften

Die Äquivarianz eines Schätzers T gegenüber bestimmten Transformationen von Stichproben beschreibt die Fähigkeit des Schätzers, sich entsprechend der Datentransformation zu ändern. Die folgenden Definitionen sind in Rousseeuw und Leroy (1987) zu finden.

Sei $\mathbf{Y} = \{(\mathbf{x}_i, y_i)'\}_{i=1, \dots, n}$ eine Stichprobe vom Umfang n , wobei $\mathbf{x}_i \in \mathbb{R}^d$ einen Spaltenvektor von Designpunkten und $y_i \in \mathbb{R}$ eine eindimensionale Zielvariable bezeichne, $i = 1, \dots, n$. Dann heißt ein Schätzer $T : \mathbb{R}^{(d+1) \times n} \rightarrow \mathbb{R}^d$ *skalenäquivalent*, wenn für jede Konstante $c \in \mathbb{R}$ gilt

$$T \left(\{(\mathbf{x}_i, c y_i)'\}_{i=1, \dots, n} \right) = c \cdot T \left(\{(\mathbf{x}_i, y_i)'\}_{i=1, \dots, n} \right) \quad \forall i = 1, \dots, n.$$

Ein Schätzer T heißt *affin äquivalent*, wenn für jede nicht-singuläre Matrix $\mathbf{A} \in \mathbb{R}^{d \times d}$ gilt

$$T \left(\{(\mathbf{x}_i \mathbf{A}, y_i)'\}_{i=1, \dots, n} \right) = \mathbf{A}^{-1} \cdot T \left(\{(\mathbf{x}_i, y_i)'\}_{i=1, \dots, n} \right) \quad \forall i = 1, \dots, n.$$

Die affine Äquivarianz von T bedeutet also, dass sich die durch T gewonnene Schätzung entsprechend der linearen Transformation der \mathbf{x}_i ändert.

Ein Schätzer T heißt *regressionsäquivalent*, wenn für und jeden Zeilenvektor $\mathbf{v} \in \mathbb{R}^d$ gilt

$$T \left(\{(\mathbf{x}_i, y_i + \mathbf{x}_i \mathbf{v})'\}_{i=1, \dots, n} \right) = T \left(\{(\mathbf{x}_i, y_i)'\}_{i=1, \dots, n} \right) + \mathbf{v} \quad \forall i = 1, \dots, n.$$

In dieser Arbeit handelt es sich bei den Stichproben um Ausschnitte aus Zeitreihen mit $\mathbf{x}_i = i \in \mathbb{N}$. Somit gilt $\mathbf{Y} = ((1, y_1)', \dots, (n, y_n)')$. In diesem Falle beschreibt die Regressionsäquivarianz also die Fähigkeit des Schätzers, äquivalent gegenüber Transformationen der Zielvariablen der Form $y_i + i v$, $v \in \mathbb{R}$, zu sein.

Mittlerer Quadratischer Fehler, Verzerrung, Varianz und Effizienz

Sei $\mathbf{Y} = (Y_1, \dots, Y_n)$ mit Zufallsvariablen $Y_1, \dots, Y_k \sim F$ und $T = T(\mathbf{Y})$ Schätzer für einen Parameter $\theta \in \Theta$ mit $T \sim G^F$. Dann ist der *mittlere quadratische Fehler* (Mean Squared Error, MSE) von T an der Verteilung F definiert als

$$\text{MSE}_F(T) = \text{Bias}_F^2(T) + \text{Var}_F(T).$$

Dabei ist

$$\text{Bias}_F(T) = E_{G^F}(T) - \theta$$

die *Verzerrung* oder der *Bias* von T an der Verteilung F . Gilt $\text{Bias}_F(T) = 0$, so ist der Schätzer *unverzerrt* bzw. *erwartungstreu*. Die *Varianz* von T an F ist gegeben durch

$$\text{Var}_F(T) = E_{G^F}([T - E_{G^F}(T)]^2).$$

Die *Effizienz* stellt ein Maß der Präzision eines Schätzers T_1 im Vergleich zu einem Schätzer T_2 an einer bestimmten Verteilung F dar. Die Effizienz von T_1 zu T_2 an F ist definiert als

$$\text{Eff}(T_1, T_2, F) = \frac{\text{MSE}_F(T_2)}{\text{MSE}_F(T_1)}.$$

Gilt $\text{Eff}(T_1, T_2, F) \geq 1 \Leftrightarrow \text{MSE}_F(T_1) \leq \text{MSE}_F(T_2)$, so heißt T_1 *effizienter* als T_2 an der Verteilung F . Häufig ist der Referenzschätzer T_2 ein unverzerrter Schätzer, dessen Varianz die Cramér-Rao-Schranke erreicht. In diesem Fall gibt es keinen Schätzer T mit geringerem MSE als T_2 . Der Schätzer T_2 heißt dann *effizient*.

Häufig wird der MSE mit Hilfe von Simulationen geschätzt. Dazu wird der Schätzer T ausgewertet an K Stichproben $\mathbf{y}_k \in \mathbb{R}^n$, $k = 1, \dots, K$, wobei sich jede Stichprobe \mathbf{y}_k als Realisation von Zufallsvariablen $X_1, \dots, X_n \sim F$ ergibt. Sei $T(\mathbf{y}_k) = T(k)$ die Schätzung von θ auf Basis der k -ten Stichprobe. Dann bezeichnet $\text{bias}_F(T)$ die geschätzte Verzerrung von T an F mit

$$\text{bias}_F(T) = \frac{1}{k} \sum_{k=1}^K T(k) - \theta,$$

und $\text{var}_F(T)$ ist die geschätzte Varianz von T an F mit

$$\text{var}_F(T) = \frac{1}{k-1} \sum_{k=1}^K \left(T(k) - \frac{1}{k} \sum_{k=1}^K T(k) \right)^2.$$

Der geschätzte MSE von T an F ist dann

$$\text{mse}_F(T) = \text{bias}_F^2(T) + \text{var}_F(T),$$

und $\text{eff}(T_1, T_2, F) = \text{mse}_F(T_2) / \text{mse}_F(T_1)$ ist die geschätzte Effizienz von T_1 zu T_2 an F .

B Simulationen

B.1 Varianz des RM-Steigungsschätzers für verschiedene Fehlerverteilungen

Im Folgenden wird die Varianz der RM-Steigung (2.4) aus Kapitel 2.1.1 untersucht. Nach Satz 3.1 in Kapitel 3.2 ist die Varianz des RM-Steigungsschätzers $\text{Var}[\hat{\beta}(\mathbf{Y})]$ unter H_0 für normalverteilte Fehler mit Erwartungswert null nur abhängig von der Länge n des Zufallsvektors \mathbf{Y} und von der Fehlervarianz σ^2 . Folglich lässt sich die Varianz der RM-Steigung als Funktion $V(n, \sigma)$ auffassen. Des Weiteren gilt $V(n, \sigma) = \sigma \cdot V(n, 1)$ nach Bemerkung 3.1.

B.1.1 Normalverteilte Fehler

Im Folgenden wird $V(n, 1)$ für $n = 5, \dots, 300$ mittels Simulationen approximiert. Dazu wird für jedes $n = 5, \dots, 300$ eine Zeitreihe der Länge $T = N + n - 1$ mit $N = 10000$ erzeugt, die gemäß der zentralen Annahme (2.1) dem Modell $Y_t = \mu_t + \varepsilon_t$ mit unabhängigen Fehlern $\varepsilon_t \sim N(0, 1)$ folgt. Die RM-Steigung wird dann im gleitenden Zeitfenster $\{t - n + 1, \dots, t\}$ ermittelt, so dass sich $N = 10000$ RM-Steigungsschätzungen $\hat{\beta}_{t,n}$ mit $t = n, \dots, T$ ergeben. Die empirische Varianz der $\hat{\beta}_{t,n}$ dient dann als Approximation $v_n := \hat{V}(n, 1)$:

$$v_n = s^2(\hat{\beta}_{t,n}) := \frac{1}{N-1} \sum_{t=n}^T \left(\hat{\beta}_{t,n} - \frac{1}{N} \sum_{t=n}^T \hat{\beta}_{t,n} \right)^2.$$

Die v_n sind wie zu erwarten monoton fallend in n und weisen einen konvexen Verlauf auf. Anhand der v_5, \dots, v_{300} wird das Modell

$$v(n) = 4.77 \cdot 10^{-7} + \frac{17.71}{n^3} \tag{B.1}$$

erstellt, das ein Bestimmtheitsmaß von 0.9983 aufweist. Die R-Funktion `scarm.filter` nutzt diese Funktion $v(n)$ zur Approximation von $V(n, 1)$.

B.1.2 Weitere Fehlerverteilungen

Das in B.1 erstellte Modell $v(n)$ liefert lediglich für unabhängig normalverteilte Fehler ε_t (nachfolgend als Fehlertyp 1 bezeichnet) geeignete Approximationen der wahren Varianz $V(n, 1)$ und ist nicht für alle Fehlerverteilungen gültig. Daher werden die Approximationen $v_n = s^2(\hat{\beta}_{t,n})$ für $n = 5, \dots, 150$ zusätzlich für unabhängige Fehler ε_t ermittelt, die den folgenden Verteilungen folgen:

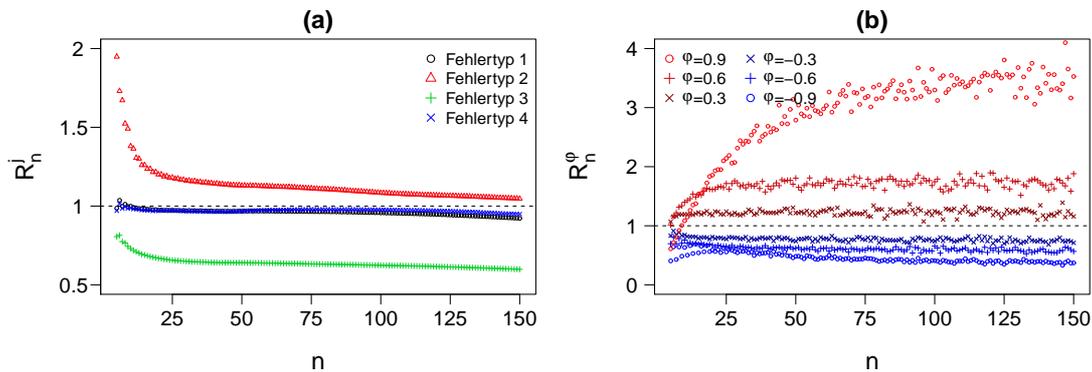


Abbildung B.1: (a) Vergleich der modellierten Varianzen der RM-Steigung mit den empirischen Varianzen der RM-Steigung für die Fehlertypen 1-4; (b) Vergleich der modellierten Varianzen der RM-Steigung mit den empirischen Varianzen der RM-Steigung für autokorrelierte Fehler.

Typ 2: Eine Mischverteilung: $\varepsilon_i \sim 0.9N(0, 1) + 0.1N(10, 1)$.

Typ 3: Die t -Verteilung mit drei Freiheitsgraden.

Typ 4: Die Weibull-Verteilung mit Formparameter zwei und Skalenparameter eins.

Bei den Typen 3 und 4 sind die Verteilungen durch den jeweiligen Erwartungswert und die jeweilige Varianz standardisiert. Für jeden Fehlertyp 2,3,4 werden die v_n analog zum Fehlertyp 1 anhand der empirischen Varianz der RM-Steigungsschätzungen gewonnen. Im Folgenden seien sie mit v_n^j bezeichnet, wobei der Index j den Fehlertyp kennzeichnet, $j = 1, \dots, 4$. Die v_n^j werden mit den von der R-Funktion `scarm.filter` genutzten $v(n)$ aus dem Modell (B.1) verglichen anhand

$$R_n^j := \sqrt{\frac{v_n^j}{v(n)}}.$$

Liegen für den Fehlertyp j die Werte von R_n^j nahe eins, stellen die $v(n)$ aus dem Modell (B.1) geeignete Approximationen der wahren Varianzen der RM-Steigungen unter diesem Fehlertyp dar. Gilt $R_n^j > 1$ (< 1), so streuen die RM-Steigungsschätzungen für den Fehlertyp j stärker (schwächer) als dies durch die Approximationen $v(n)$ beschrieben wird.

Abbildung B.1 (a) zeigt die R_n^j -Werte für die vier Fehlertypen. Für den Fehlertyp 1 liegen die R_n^j -Werte nahe eins, was aufgrund des hohen Bestimmtheitsmaßes von 0.9983 den Erwartungen entspricht. Darüber hinaus scheinen die $v(n)$ offenbar auch für die schiefen Fehler der Weibull-Verteilung (Fehlertyp 4) geeignet zu sein, da auch hier die R_n^j -Werte nahe

eins liegen. Für die kontaminierten Fehler (Fehlertyp 2) gilt $R_n^j > 1$ für alle $n = 5, \dots, 150$, d.h. die RM-Steigung weist unter Kontamination eine größere Variabilität auf als bei normalverteilten Fehlern. Mit zunehmender Fensterbreite streben die R_n^j jedoch gegen eins, d.h. die durch die Ausreißer bedingte Verzerrung verschwindet. Für die t_3 -verteilten Fehler (Fehlertyp 3) ergibt sich ein gegenteiliges Bild. Obwohl es sich hier um eine Verteilung mit schweren Rändern handelt, weisen die RM-Steigungen eine geringere Variabilität auf als bei normalverteilten Fehlern. Dabei ist die Standardabweichung der RM-Steigungen für t_3 -verteilte Fehler nur etwa halb so groß wie für standardnormalverteilte Fehler. Zudem tritt offenbar keine Annäherung für große n ein.

B.1.3 Autokorrelierte Fehler

Bei allen bisher betrachteten Fehlertypen sind die ε_t unabhängig, also nicht autokorreliert. Um die Auswirkung von Autokorrelationen auf die Varianz der RM-Steigung $V(n, 1)$ zu untersuchen, werden zusätzlich noch Fehler aus dem AR(1)-Modell

$$\xi_t^\varphi = \varphi \xi_{t-1} + e_t,$$

mit $e_t \sim N(0, 1)$ generiert. Dabei ist $\varphi \in \{-0.9, -0.6, -0.3, 0.3, 0.6, 0.9\}$, um negative und positive Korrelationen unterschiedlicher Stärke zu betrachten. Wegen $\text{Var}(\xi_t^\varphi) = (1 - \varphi^2)^{-1}$ werden die Fehler standardisiert, so dass

$$\varepsilon_t^\varphi = \xi_t^\varphi \sqrt{(1 - \varphi^2)}$$

mit $\text{Var}(\varepsilon_t^\varphi) = 1$ für alle φ .

Für alle φ und für $n = 5, \dots, 150$ werden die Approximationen von $V(n, 1)$ analog zu oben ermittelt; sie seien nachfolgend mit v_n^φ bezeichnet. Der Vergleich der v_n^φ mit den $v(n)$ aus Modell (B.1) erfolgt wie oben anhand $R_n^\varphi = \sqrt{v_n^\varphi / v(n)}$. Die R_n^φ -Werte für alle $\varphi \in \{-0.9, -0.6, -0.3, 0.3, 0.6, 0.9\}$ und $n = 5, \dots, 150$ sind in Abbildung B.1 (b) gegeben. Bei positiver (negativer) Autokorrelation sind die $v(n)$ offenbar zu klein (zu groß), d.h. die RM-Steigungen streuen stärker (schwächer) als dies durch die $v(n)$ beschrieben wird. (Eine Ausnahme bilden hier die Fälle $\varphi = 0.9$ und $n \leq 8$.) Dabei sind die Abweichungen der v_n^φ von dem Modell $v(n)$ umso größer, je stärker die Autokorrelation ist.

B.2 Effizienz des Q_δ^{adj} und $Q_\delta^{\text{adj}*}$ für fehlende Werte

Mit sinkendem Stichprobenumfang und steigendem Anteil an fehlenden Werten steigt das Risiko einer nicht zu berechnenden Q_δ^{adj} -Schätzung aufgrund ungünstig positionier-

	$n = 10$	$n = 20$	$n = 30$	$n = 40$
$p = 0.3$	0.0374	0.0002	0	0
$p = 0.4$	0.2116	0.0114	0.0011	0
$p = 0.5$	0.5033	0.1284	0.0330	0.0080
$p = 0.6$	0.7575	0.4145	0.2235	0.1171

Tabelle B.1: Rate an nicht berechenbaren Q_δ^{adj} -Schätzungen aufgrund von $\lfloor p \cdot n \rfloor$ fehlenden Werten in Stichproben vom Umfang n .

ter fehlender Werte, vgl. Tabelle B.1. Diese zeigt die Rate an nicht zu berechnenden Q_δ^{adj} -Skalenschätzungen aufgrund von $\lfloor p \cdot n \rfloor$ fehlenden Werten in Stichproben vom Umfang n bei $N = 10000$ Wiederholungen. Dies motiviert die Entwicklung des alternativen $Q_\delta^{\text{adj}*}$ -Skalenschätzers, der in (3.17) auf Seite 40 gegeben ist. In diesem Abschnitt erfolgt ein Vergleich des $Q_\delta^{\text{adj}*}$ und des Q_δ^{adj} bzgl. der Effizienz der Schätzer bei fehlenden Werten. Dazu werden Stichproben y_1, \dots, y_n gemäß dem Modell $Y_i = \mu + \beta \cdot i + \varepsilon_i$ mit $\varepsilon_i \stackrel{\text{u.i.v.}}{\sim} N(0, \sigma^2 = 1)$ generiert, wobei o.B.d.A. $\mu = \beta = 0$ ist. Anschließend werden $\lfloor p \cdot n \rfloor$ der n Werte durch $y_i = \text{NA}$ ersetzt. Für jede Kombination von n und p mit $n \in \{20, 60, 100\}$ und $p \in \{0.1, 0.3, 0.5, 0.7\}$ werden $N = 10000$ Stichproben generiert, jeweils die Q_δ^{adj} - und $Q_\delta^{\text{adj}*}$ -Schätzungen von $\sigma = 1$ berechnet und für jede (n, p) -Kombination die Verzerrung, die Standardabweichung und der *Root* MSE (RMSE) von Q_δ^{adj} und $Q_\delta^{\text{adj}*}$ geschätzt. Die Resultate sind in den Tabellen B.2 – B.4 verzeichnet. Mit zunehmendem Anteil an fehlenden Werten weisen der Q_δ^{adj} und $Q_\delta^{\text{adj}*}$ leichte Verzerrungen auf. Der Q_δ^{adj} unterschätzt die wahre Standardabweichung der Fehler σ , wohingegen der $Q_\delta^{\text{adj}*}$ diese überschätzt. Im Gegensatz zum Q_δ^{adj} ist der $Q_\delta^{\text{adj}*}$ jedoch nur leicht verzerrt. Darüber hinaus ist die Standardabweichung des $Q_\delta^{\text{adj}*}$ für jede Kombination von n und p geringer als die des Q_δ^{adj} . Folglich weist der $Q_\delta^{\text{adj}*}$ in jeder betrachteten Situation einen geringeren RMSE auf.

Die beschriebene Studie wird analog für Fehler durchgeführt, die einer Mischverteilung folgen, d.h. $\varepsilon_i \sim 0.9N(0, \sigma^2 = 1) + 0.1N(0, \sigma_*^2 = 100)$. Die Tabellen B.5 – B.7 zeigen die Schätzungen der Verzerrung, der Standardabweichung und des RMSE von Q_δ^{adj} und $Q_\delta^{\text{adj}*}$. Hier überschätzen beide Schätzer aufgrund der Ausreißer in den Daten die Standardabweichung $\sigma^2 = 1$. Dabei ist die Verzerrung des $Q_\delta^{\text{adj}*}$ geringer, was aufgrund des größeren Bruchpunkts des $Q_\delta^{\text{adj}*}$ plausibel ist. Darüber hinaus ist die geschätzte Standardabweichung des $Q_\delta^{\text{adj}*}$ für jede (n, p) -Kombination geringer. Der $Q_\delta^{\text{adj}*}$ weist folglich in jeder betrachteten Situation einen geringeren RMSE auf als der Q_δ^{adj} .

	$n = 20$		$n = 60$		$n = 100$	
	Q_δ^{adj}	$Q_\delta^{\text{adj}*}$	Q_δ^{adj}	$Q_\delta^{\text{adj}*}$	Q_δ^{adj}	$Q_\delta^{\text{adj}*}$
$p = 0.1$	-0.012	-0.016	0.005	0.001	0.010	0.004
$p = 0.3$	0.027	-0.021	0.009	-0.003	-0.003	-0.021
$p = 0.5$	0.177	-0.024	0.037	-0.020	0.021	-0.029
$p = 0.7$	0.020	-0.032	0.156	-0.032	0.138	-0.024

Tabelle B.2: Geschätzte Verzerrung von Q_δ^{adj} und $Q_\delta^{\text{adj}*}$ für Stichproben der Länge n mit $\lfloor n \cdot p \rfloor$ fehlenden Werten und normalverteilten Fehlern.

	$n = 20$		$n = 60$		$n = 100$	
	Q_δ^{adj}	$Q_\delta^{\text{adj}*}$	Q_δ^{adj}	$Q_\delta^{\text{adj}*}$	Q_δ^{adj}	$Q_\delta^{\text{adj}*}$
$p = 0.1$	0.381	0.339	0.214	0.192	0.164	0.152
$p = 0.3$	0.568	0.400	0.313	0.221	0.235	0.179
$p = 0.5$	0.650	0.476	0.501	0.270	0.387	0.212
$p = 0.7$	0.769	0.690	0.720	0.344	0.694	0.277

Tabelle B.3: Geschätzte Standardabweichung von Q_δ^{adj} und $Q_\delta^{\text{adj}*}$ für Stichproben der Länge n mit $\lfloor n \cdot p \rfloor$ fehlenden Werten und normalverteilten Fehlern.

	$n = 20$		$n = 60$		$n = 100$	
	Q_δ^{adj}	$Q_\delta^{\text{adj}*}$	Q_δ^{adj}	$Q_\delta^{\text{adj}*}$	Q_δ^{adj}	$Q_\delta^{\text{adj}*}$
$p = 0.1$	0.381	0.341	0.214	0.192	0.164	0.152
$p = 0.3$	0.569	0.401	0.313	0.221	0.235	0.179
$p = 0.5$	0.674	0.477	0.502	0.270	0.387	0.214
$p = 0.7$	0.769	0.691	0.737	0.345	0.708	0.277

Tabelle B.4: Geschätzter RMSE von Q_δ^{adj} und $Q_\delta^{\text{adj}*}$ für Stichproben der Länge n mit $\lfloor n \cdot p \rfloor$ fehlenden Werten und normalverteilten Fehlern.

	$n = 20$		$n = 60$		$n = 100$	
	Q_δ^{adj}	$Q_\delta^{\text{adj}*}$	Q_δ^{adj}	$Q_\delta^{\text{adj}*}$	Q_δ^{adj}	$Q_\delta^{\text{adj}*}$
$p = 0.1$	-0.449	-0.448	-0.369	-0.370	-0.355	-0.357
$p = 0.3$	-0.512	-0.492	-0.386	-0.382	-0.361	-0.365
$p = 0.5$	-0.661	-0.572	-0.444	-0.386	-0.373	-0.372
$p = 0.7$	-1.129	-0.917	-0.753	-0.446	-0.572	-0.400

Tabelle B.5: Geschätzte Verzerrung von Q_δ^{adj} und $Q_\delta^{\text{adj}*}$ für Stichproben der Länge n mit $\lfloor n \cdot p \rfloor$ fehlenden Werten und Fehlern aus einer kontaminierten Normalverteilung.

	$n = 20$		$n = 60$		$n = 100$	
	Q_δ^{adj}	$Q_\delta^{\text{adj}*}$	Q_δ^{adj}	$Q_\delta^{\text{adj}*}$	Q_δ^{adj}	$Q_\delta^{\text{adj}*}$
$p = 0.1$	0.887	0.789	0.382	0.351	0.281	0.257
$p = 0.3$	1.436	0.966	0.575	0.406	0.402	0.298
$p = 0.5$	2.506	1.349	1.228	0.504	0.768	0.352
$p = 0.7$	3.371	2.498	2.771	0.750	2.338	0.492

Tabelle B.6: Geschätzte Standardabweichung von Q_δ^{adj} und $Q_\delta^{\text{adj}*}$ für Stichproben der Länge n mit $\lfloor n \cdot p \rfloor$ fehlenden Werten und Fehlern aus einer kontaminierten Normalverteilung.

	$n = 20$		$n = 60$		$n = 100$	
	Q_δ^{adj}	$Q_\delta^{\text{adj}*}$	Q_δ^{adj}	$Q_\delta^{\text{adj}*}$	Q_δ^{adj}	$Q_\delta^{\text{adj}*}$
$p = 0.1$	0.993	0.907	0.531	0.510	0.453	0.440
$p = 0.3$	1.525	1.084	0.694	0.558	0.540	0.471
$p = 0.5$	2.592	1.466	1.305	0.635	0.854	0.512
$p = 0.7$	3.555	2.661	2.872	0.873	2.406	0.634

Tabelle B.7: Geschätzter RMSE von Q_δ^{adj} und $Q_\delta^{\text{adj}*}$ für Stichproben der Länge n mit $\lfloor n \cdot p \rfloor$ fehlenden Werten und Fehlern aus einer kontaminierten Normalverteilung.

C Abbildungen und Tabellen

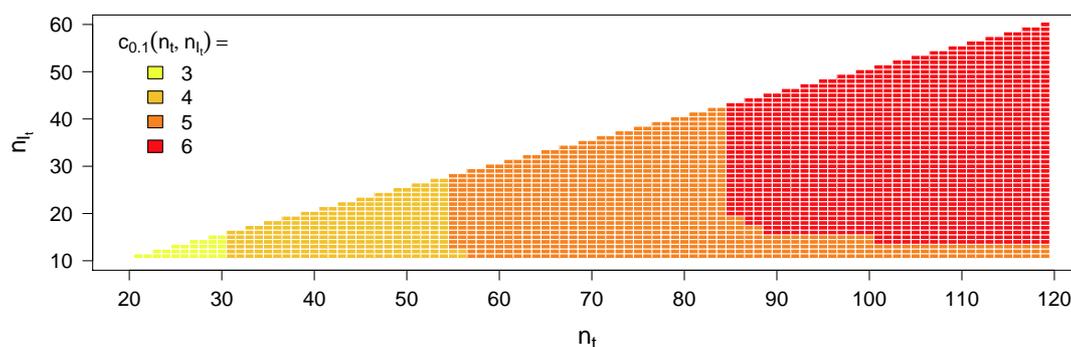


Abbildung C.1: Kritische Werte $c_{0,1}(n_t, n_{I_t})$ für $n_{I_t} \in [10, 60]$ und $n_t \in [2n_{I_t}, 120]$.

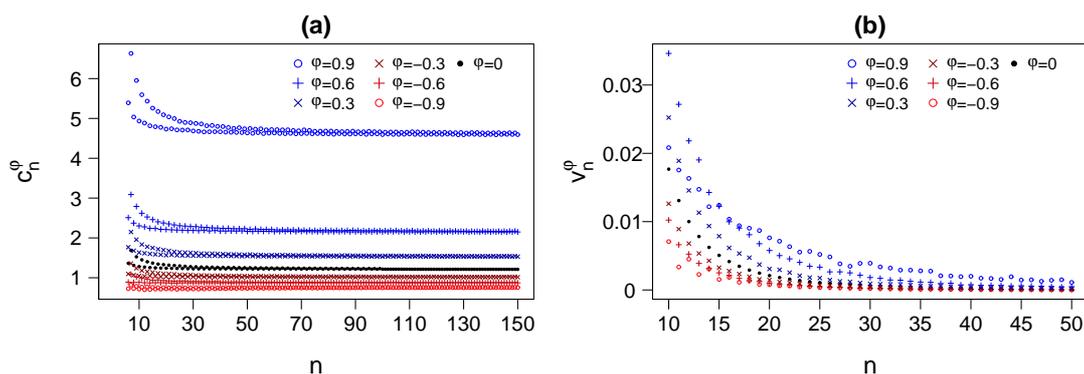


Abbildung C.2: (a) Korrekturfaktoren c_n^φ für den Q_δ^{adj} -Schätzer; (b) empirische Varianz v_n^φ der RM-Steigung; jeweils für Stichproben der Länge n mit AR(1)-Fehlern mit Parameter φ und Varianz eins.

		Rechte Fensterbreite r									
		5	10	15	20	25	30	35	40	45	50
Linke Fensterbreite ℓ	5	3.5	–	–	–	–	–	–	–	–	–
	10	4.8	6.1	–	–	–	–	–	–	–	–
	15	7.2	7.5	11.1	–	–	–	–	–	–	–
	20	8.1	9.4	12.1	12.3	–	–	–	–	–	–
	25	10.5	11.5	13.5	14.4	20.5	–	–	–	–	–
	30	10.8	12.1	15.5	20.0	19.4	21.4	–	–	–	–
	35	11.6	15.9	17.2	20.2	22.9	24.2	27.4	–	–	–
	40	14.4	14.6	14.8	25.9	24.1	23.0	31.6	22.6	–	–
	45	16.1	16.2	22.2	25.1	22.6	27.3	28.2	29.2	37.3	–
	50	20.8	22.9	22.6	23.0	29.0	26.2	25.2	53.9	38.1	34.9

Tabelle C.1: Freiheitsgrade $f(\ell, r)$ für die Approximation der Verteilung der SCARM-Teststatistik über eine t -Verteilung.

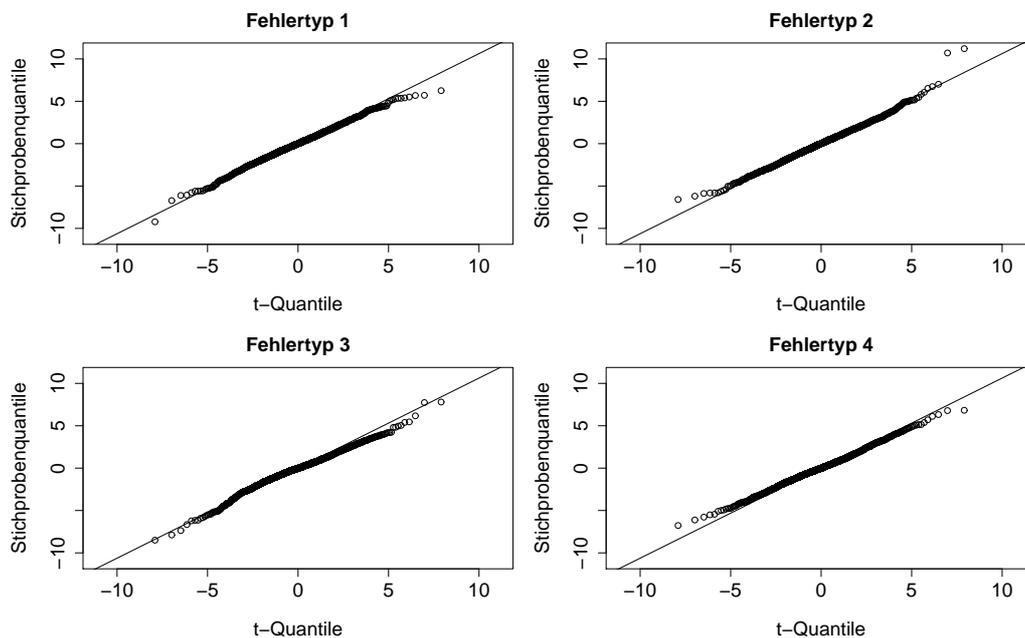


Abbildung C.3: QQ-Plots der Quantile der $t_{f(\ell,r)}$ -Verteilung mit $(\ell, r) = (10, 10)$ gegen die Stichprobenquantile der SCARM-Teststatistik für die Fehlertypen 1-4.

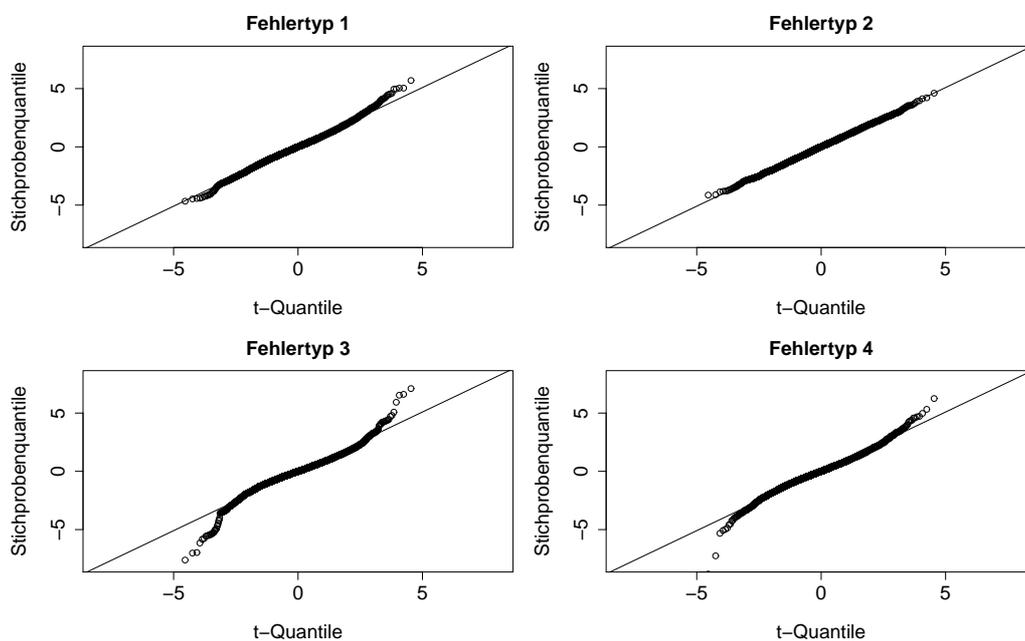


Abbildung C.4: QQ-Plots der Quantile der $t_{f(\ell,r)}$ -Verteilung mit $(\ell, r) = (50, 10)$ gegen die Stichprobenquantile der SCARM-Teststatistik für die Fehlertypen 1-4.

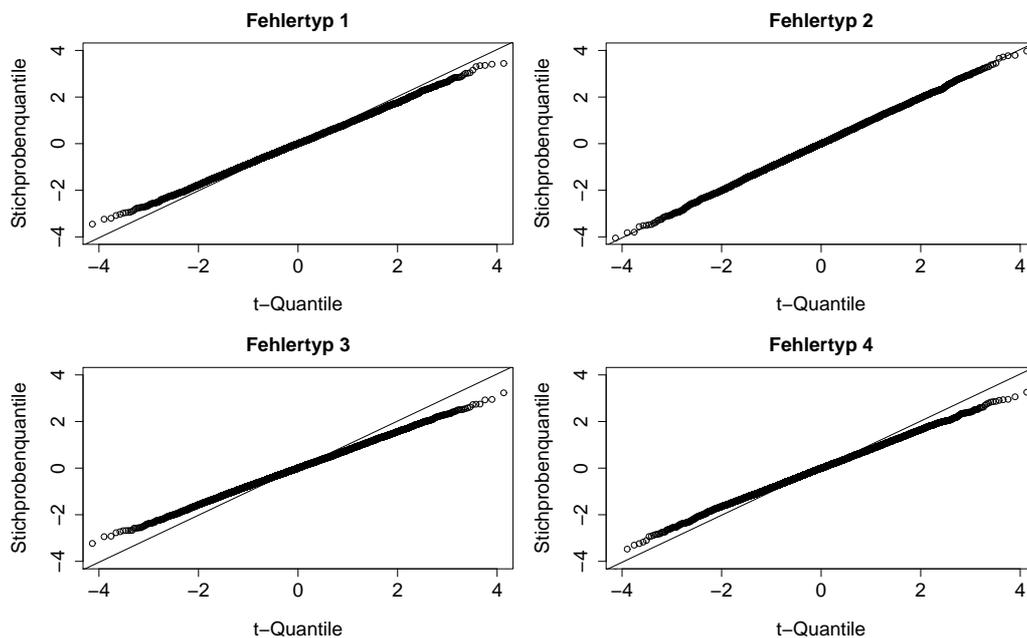


Abbildung C.5: QQ-Plots der Quantile der $t_{f(\ell,r)}$ -Verteilung mit $(\ell, r) = (100, 50)$ gegen die Stichprobenquantile der SCARM-Teststatistik für die Fehlertypen 1-4.

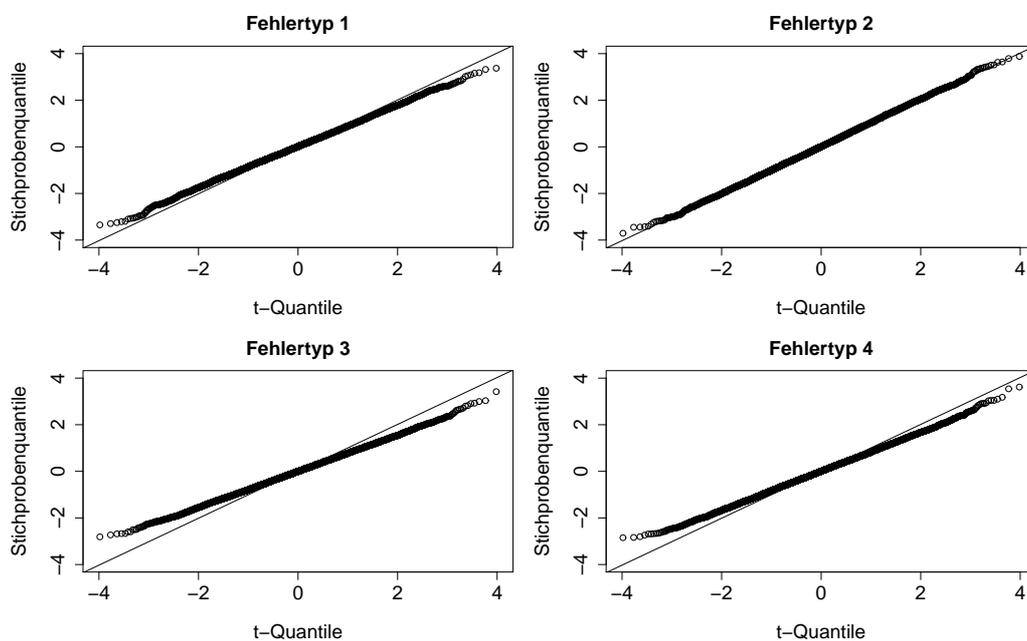


Abbildung C.6: QQ-Plots der Quantile der $t_{f(\ell,r)}$ -Verteilung mit $(\ell, r) = (100, 100)$ gegen die Stichprobenquantile der SCARM-Teststatistik für die Fehlertypen 1-4.

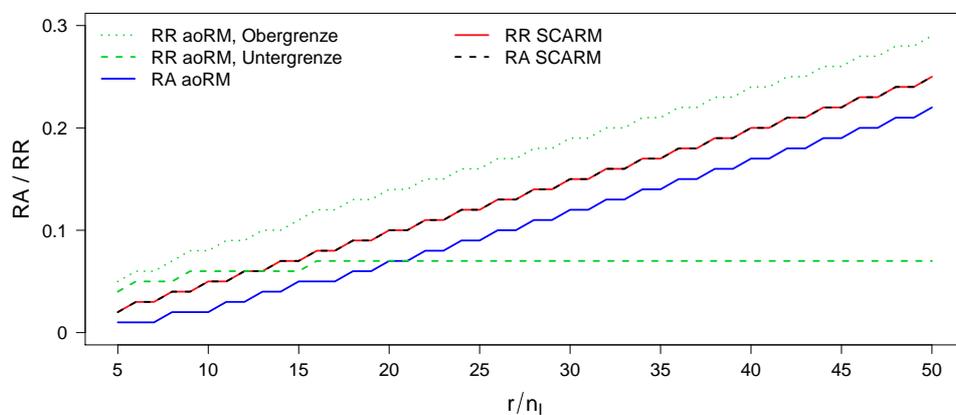


Abbildung C.7: Test Resistances des aoRM- und SCARM-Tests für $n = 100$, $r, n_I \in \{5, \dots, n/2\}$, $\alpha = 0.1$ und $\delta = \delta_{\text{opt}}$.

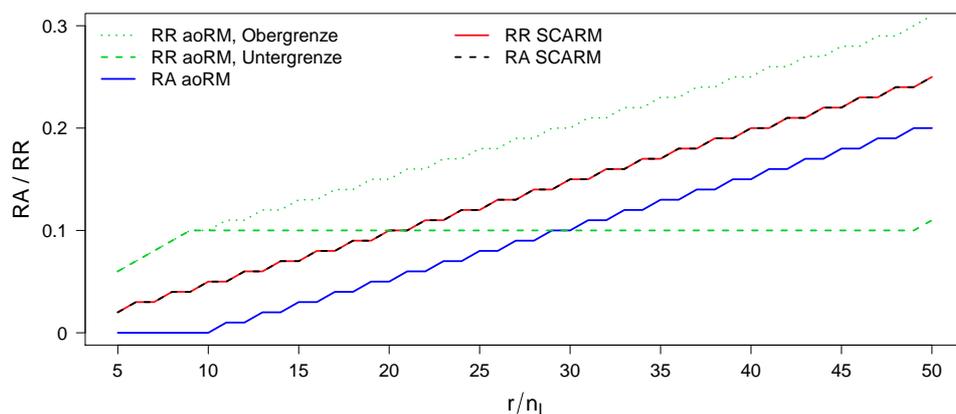


Abbildung C.8: Test Resistances des aoRM- und SCARM-Tests für $n = 100$, $r, n_I \in \{5, \dots, n/2\}$, $\alpha = 0.0001$ und $\delta = \delta_{\text{opt}}$.

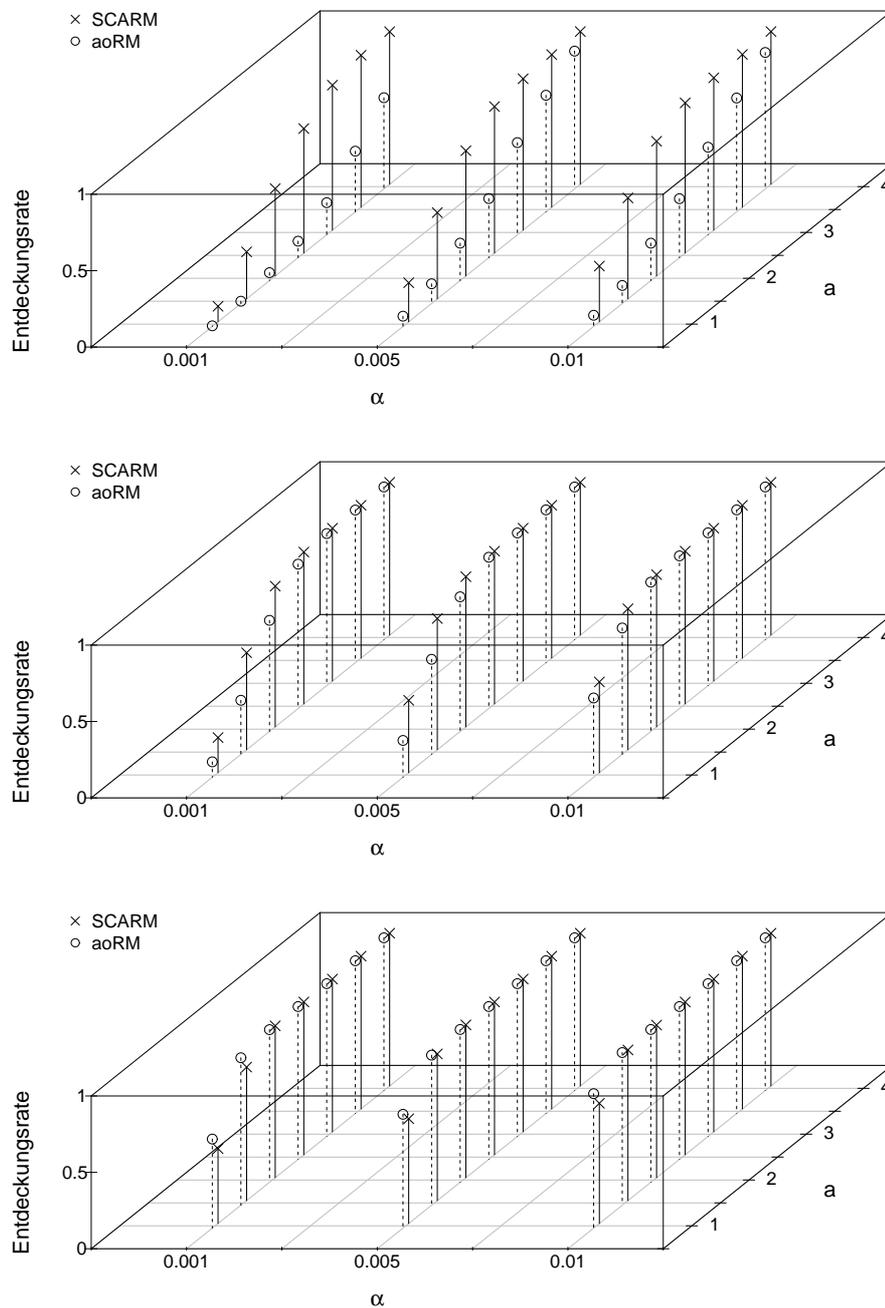


Abbildung C.9: Entdeckungsraten des aoRM- und SCARM-Tests für unterschiedliche Level Shifts mit $(\ell, r) = (40, 40)$ (oben), $(\ell, r) = (80, 40)$ (Mitte) und $(\ell, r) = (200, 40)$ (unten).

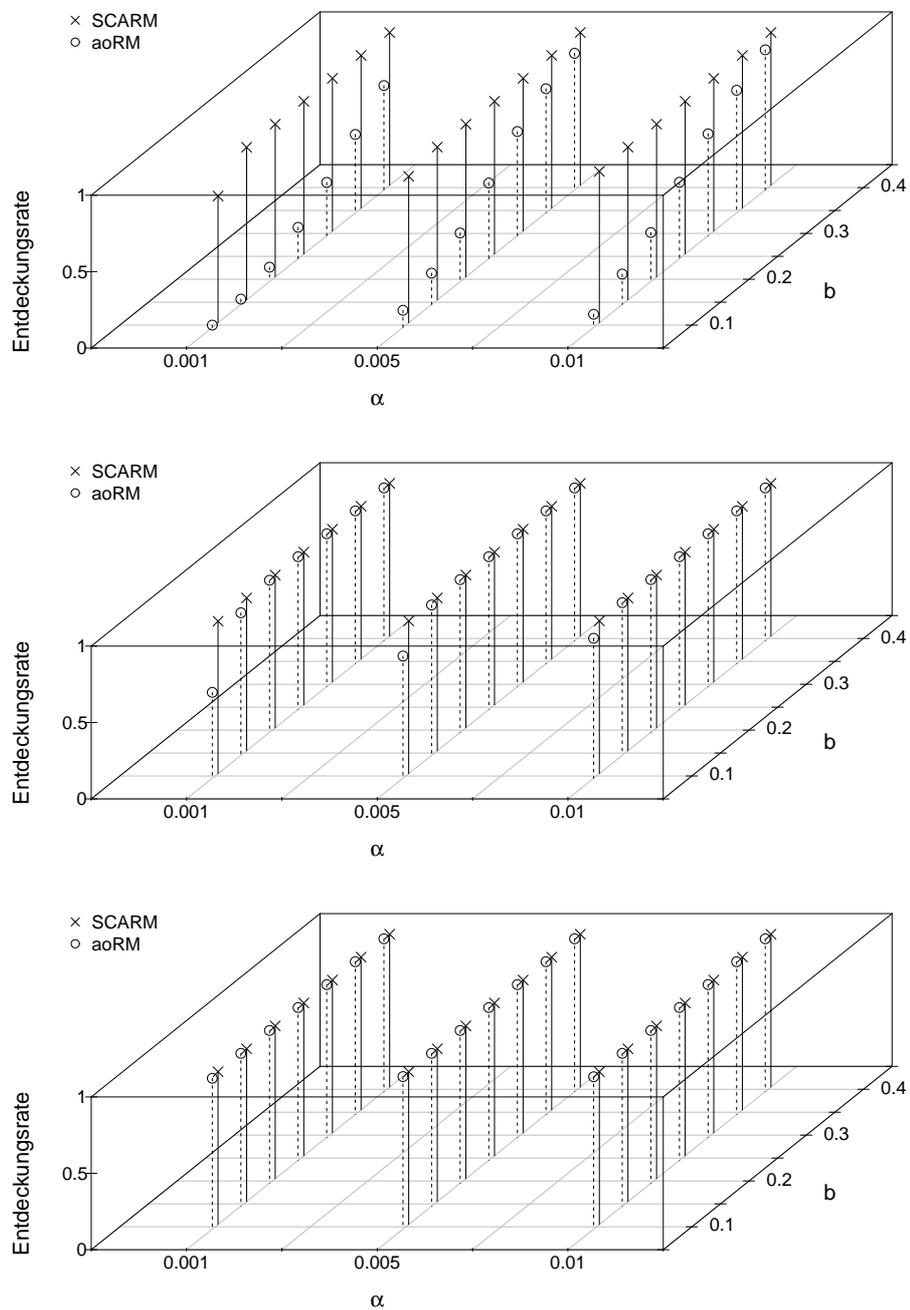


Abbildung C.10: Entdeckungsraten des aoRM- und SCARM-Tests für unterschiedliche Trendwechsel mit $(\ell, r) = (40, 40)$ (oben), $(\ell, r) = (80, 40)$ (Mitte) und $(\ell, r) = (200, 40)$ (unten).

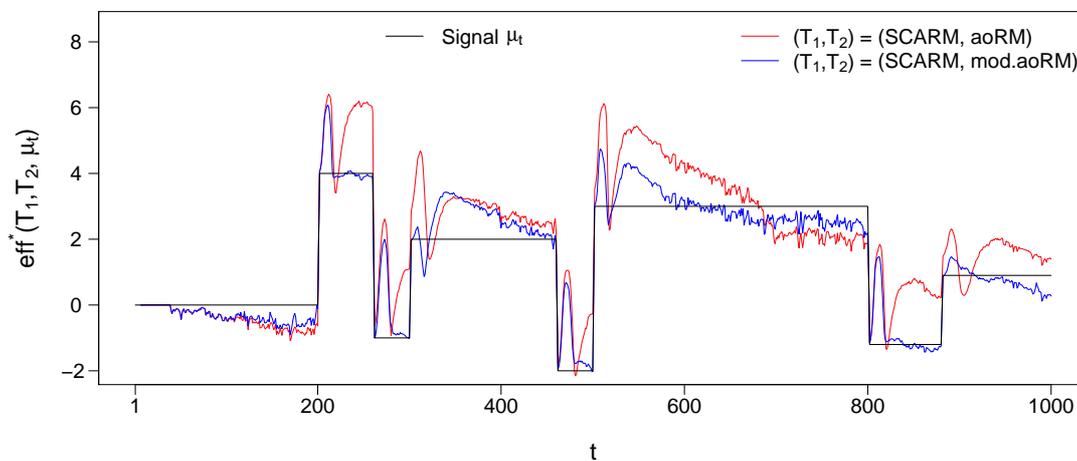


Abbildung C.11: Signaltyp „Level Shifts“ und $\text{eff}^*(T_1, T_2, \mu_t)$ für $(T_1, T_2) = (\text{SCARM}, \text{aoRM})$ und $(\text{SCARM}, \text{mod.aoRM})$ mit Signifikanzniveau $\alpha = 0.01$.

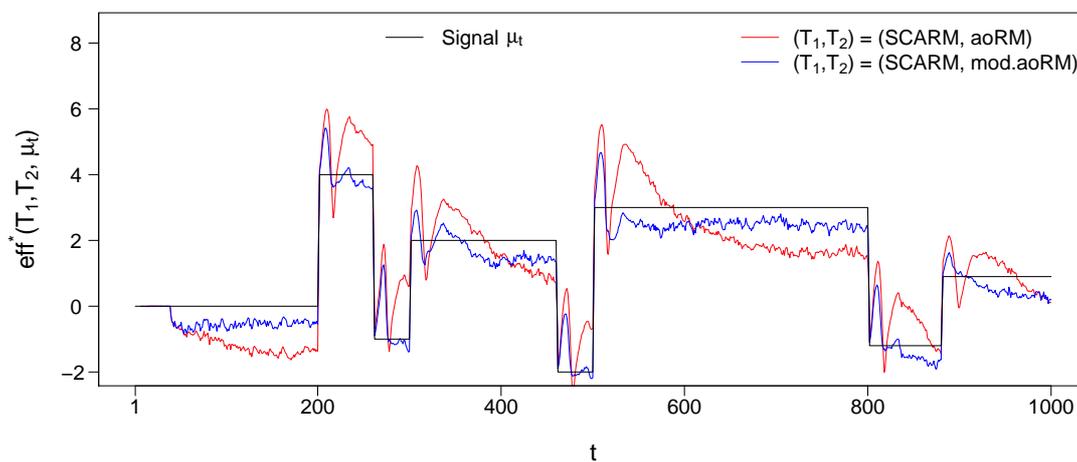


Abbildung C.12: Signaltyp „Level Shifts“ und $\text{eff}^*(T_1, T_2, \mu_t)$ für $(T_1, T_2) = (\text{SCARM}, \text{aoRM})$ und $(\text{SCARM}, \text{mod.aoRM})$ mit Signifikanzniveau $\alpha = 0.1$.

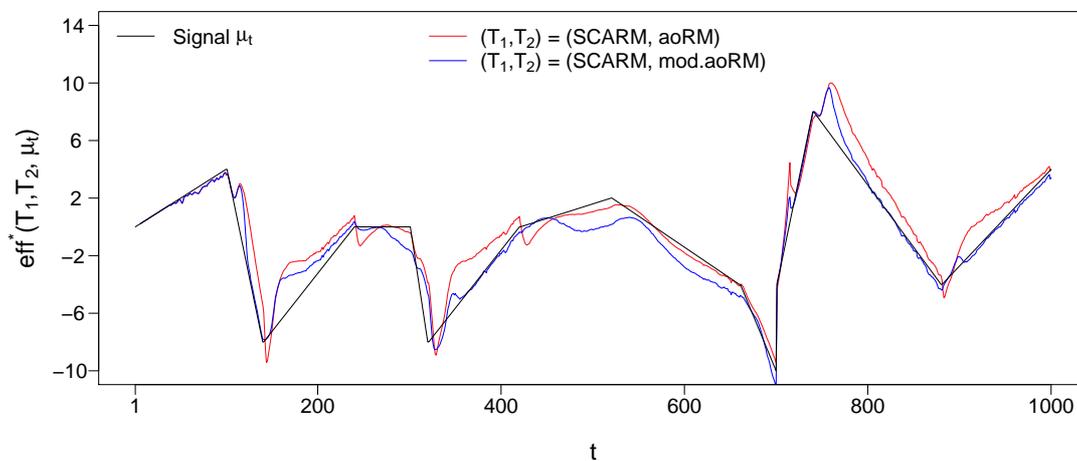


Abbildung C.13: Signaltyp „Trendwechsel“ und $\text{eff}^*(T_1, T_2, \mu_t)$ für $(T_1, T_2) = (\text{SCARM}, \text{aoRM})$ und $(\text{SCARM}, \text{mod.aoRM})$ mit Signifikanzniveau $\alpha = 0.01$.

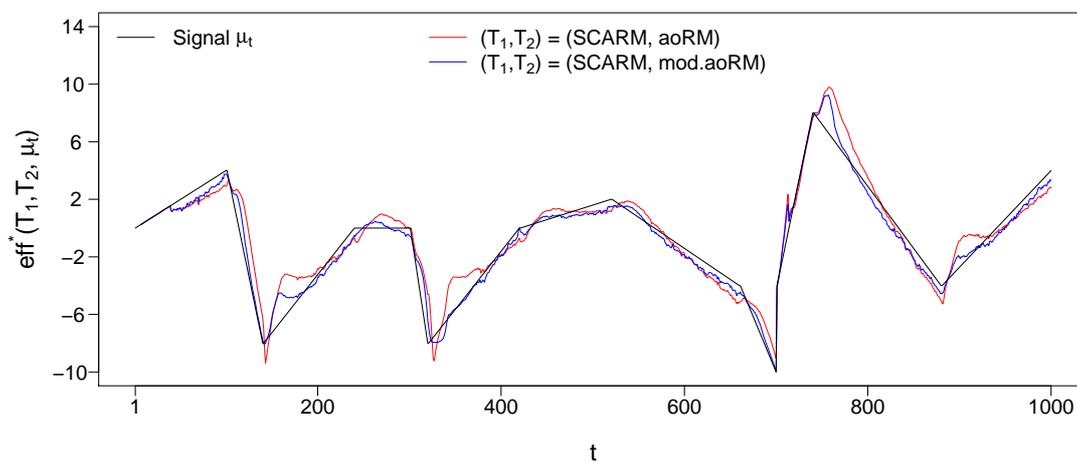


Abbildung C.14: Signaltyp „Trendwechsel“ und $\text{eff}^*(T_1, T_2, \mu_t)$ für $(T_1, T_2) = (\text{SCARM}, \text{aoRM})$ und $(\text{SCARM}, \text{mod.aoRM})$ mit Signifikanzniveau $\alpha = 0.1$.

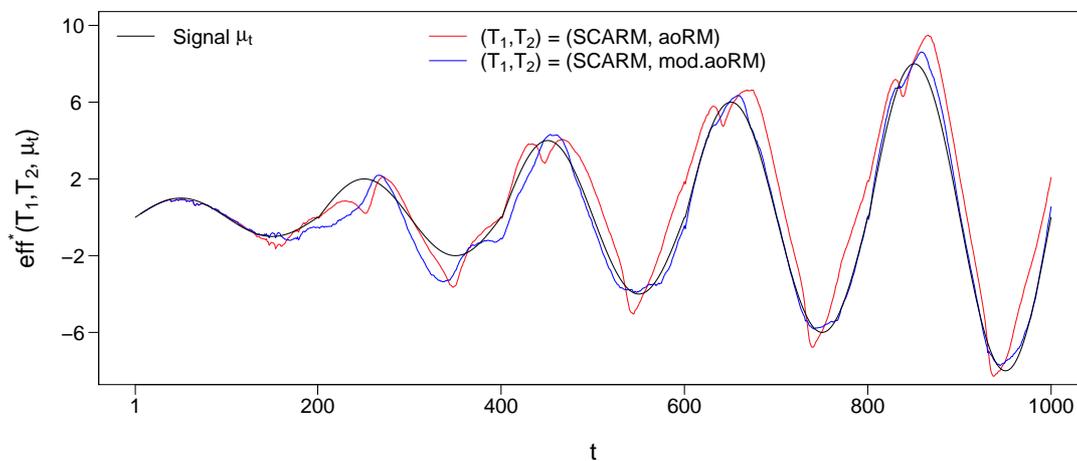


Abbildung C.15: Signaltyp „Sinus“ und $\text{eff}^*(T_1, T_2, \mu_t)$ für $(T_1, T_2) = (\text{SCARM}, \text{aoRM})$ und $(\text{SCARM}, \text{mod.aoRM})$ mit Signifikanzniveau $\alpha = 0.01$.

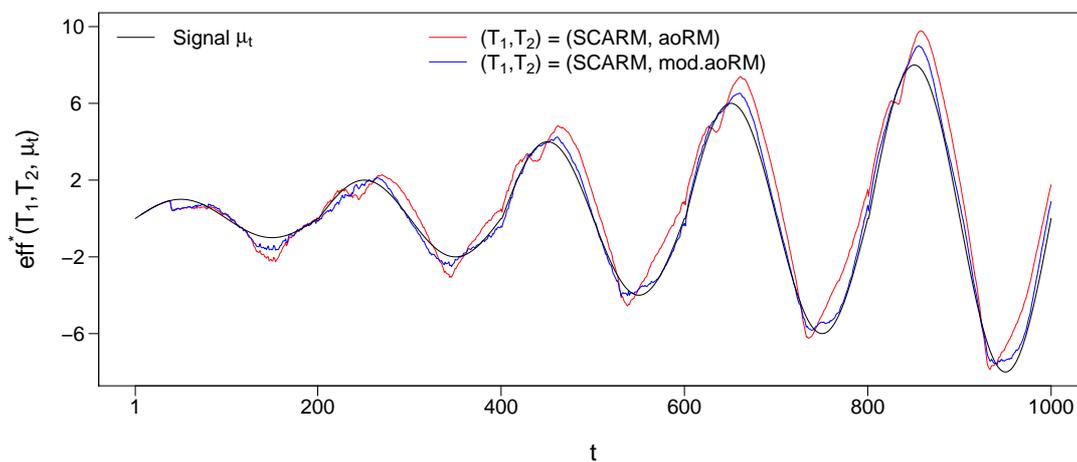


Abbildung C.16: Signaltyp „Sinus“ und $\text{eff}^*(T_1, T_2, \mu_t)$ für $(T_1, T_2) = (\text{SCARM}, \text{aoRM})$ und $(\text{SCARM}, \text{mod.aoRM})$ mit Signifikanzniveau $\alpha = 0.1$.

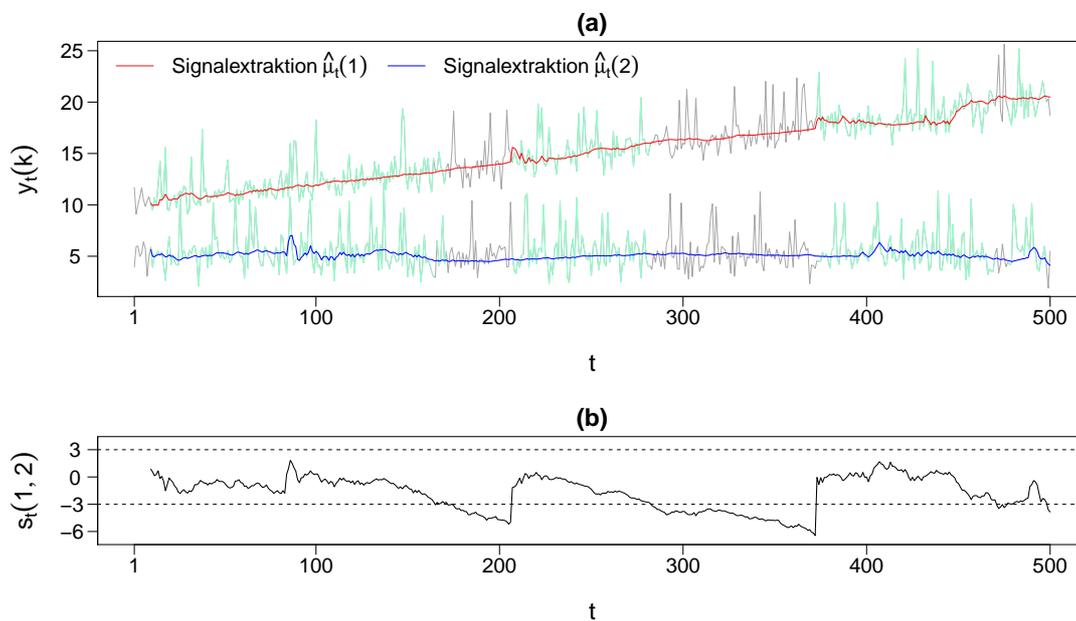


Abbildung C.17: (a) Bivariate Zeitreihe (Szenario 2) mit SCARM-Signalextraktion; das Signifikanzniveau ist mit $\alpha = 0.1$ hoch gewählt; (b) Zeitreihe der Vergleichsstatistik $s_t(1, 2)$.

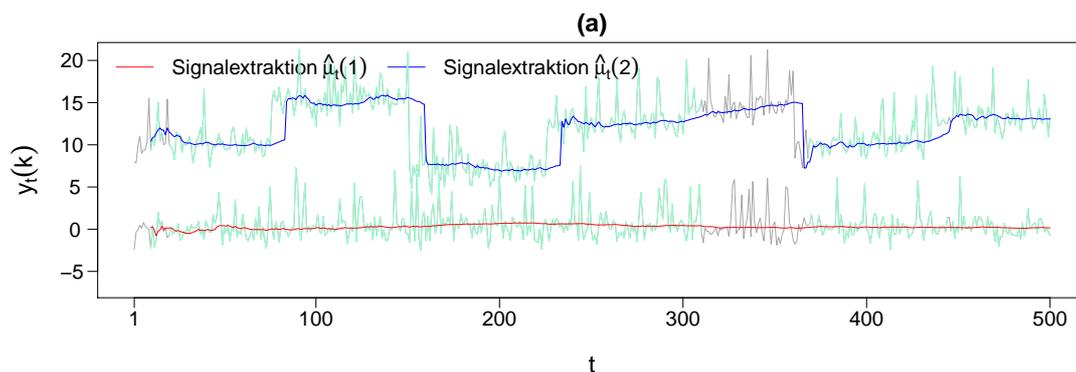


Abbildung C.18: Bivariate Zeitreihe (Szenario 6) mit SCARM-Signalextraktion und farblicher Kennzeichnung der Zusammenhänge auf Basis von $s_t(1, 2)$.