Zeitreihen in der Finanzmathematik

Jessica Aigner

Sommersemester 2009

Inhaltsverzeichnis

Kapitel 1

Einführung

1.1 Definition

Bereits aus der Statistik kennen wir Beobachtungen einer Größe, die unter identischen Bedingungen gewonnen wurden. Für die Statistische Analyse von Stichproben spielt die Reihenfolge der gemessenen Werte keine Rolle.
Wir betrachten nun ein Teilbegebiet der Statistik in dem die zeitliche Anordung sehr wohl eine Rolle spielt. Eine zeitlich geordnete Abfolge von Messwerten $(x_t)_{t \in T}$ bezeichnen wir als Zeitreihe. Für jeden Zeitpunkt t einer Menge T von Beobachtungszeitpunkten liegt dabei genau eine Beobachtung vor. Aus dem zeitlichen Ablauf der Veränderungen der untersuchten Größe versucht man in der Zeitreihenanalyse nun Trends und Veränderungen in der Vergangenheit zu bestimmen und diese Daten für Vorhersagen in der Zukunft zu nutzen.
Zeitreihen treten in den verschiedensten Bereichen auf:

- METEOROLOGIE Temperatur (Zeiteinheit: z.B. Tag, Jahrhundert)
- BIOMETRIE Herzrhythmus (Zeiteinheit: Minuten)
- ÖKONOMETRIE Arbeitslosenquote, Bruttosozialprodukt (Zeiteinheit: z.B. Jahr, Monat)
- SOZIOLOGIE Bevölkerungsentwicklung (Zeiteinheit: z.B. Jahrhundert)
- FINANZMATHEMATIK Der Einsatz von Zeitreihen in der Finanzmathematik ist sehr vielfältig.
  - *Vorhersage von Preisen* Man hat beispielsweise verderbliche, nicht lagerbare Güter, welche nur zu einer gewissen Jahreszeit geerntet werden, so kann man auf Grund von früheren Daten unter Berücksichtigung von Witterung und saisonalen Einflüssen versuchen, die künftigen Preise vorherzusagen bzw. zu schätzen.
  - *Schätzung von Zinsstrukturen* Wenn keine monetäre Extremsituation vorhanden ist, so sind die Zinsen von langfristig angelegtem Geld höher als von kurzfristig angelegtem Geld. Der Grund liegt in der längeren Ungewissheit. Der Zins als Funktion der Laufzeit wird als Zinsstruktur bezeichnet. Diesen Verlauf versucht man auch mit Zeitreihenanalyse anhand von Faktoren zu erklären.
  - *Länder- und Brancheneffekte* Mit der Einführung des Euros hat man viel über Länder- und Brancheneffekte gelesen. Gab es in der EU früher noch eindeutige Ländereffekte (die Aktienkurse von Firmen im gleichen Land (mit gleicher Währung) reagierten bei gewissen relevanten Änderungen des wirtschaftlichen Umfeldes (Zinsentscheide der Notenbank) gleich), so fallen diese zumindest bei Währungsfragen jetzt weg. In den Vordergrund treten jetzt Brancheneffekte (die Aktienkurse von

Firmen der gleichen Branche reagieren bei gewissen relevanten Änderungen des wirtschaftlichen Umfeldes (höhere Erdölpreise zum Beispiel) gleich). Früher waren diese beiden Effekte vermischt. Dies hat Folgen für das Portfoliomanagement, wo man die Aktienanlagen gezielt diversifizieren möchte. Dazu muss man Zeitreihenanalysen machen.

1.1.1 Grafische Darstellung

Die grafische Darstellung der Zeitreihe - der Plot - ist der erste Schritt bei der Analyse von Zeitreihen. Auch wenn wir uns mit diskreten Zeitreihen beschäftigen, können wir den Grafen stetig darstellen. Bei Betrachten eines Plots können wir zum Beispiel bereits saisonale Schwankungen erkennen. Neben den saisonalen Schwankungen und Perioden wie zum Beispiel Jahreszeiten sind ebenso externe Einflüsse, geschichtlichte Ereignisse oder politische Entwicklungen erkennbar, die es zu filtern gilt, um den Trend zu bestimmen. Natürlich ist es nicht möglich in der Finanzmathematik, wie in jedem anderen Bereich, mit Gewissheit die Zukunft zu bestimmen, jedoch ist es möglich nach Analyse vergangener Daten Trends herauszulesen, die mit positiver Wahrscheinlichkeit eintreten. Ein wichtiges Beispiel ist die Finanzkrise der vergangenen Jahre. Die Überbewertung von Geldanlagen (insbesondere Immobilien) führte in den USA zu erhöhtem Konsum und gleichzeitig zu erhöhten Investitionen und Überproduktion. Eine lange Preissteigerungsphase im Immobilienmarkt hatte sich in den USA zu einer Immobilienblase (deutlichen Überbewertung von Immobilien, die sich durch einen dramatischen Preisverfall innerhalb kurzer Zeit normalisiert) entwickelt. Gleichzeitig konnten immer mehr Kreditnehmer ihre Kreditraten nicht mehr bedienen, teils wegen steigender Zinsen, teils wegen fehlender Einkommen. Die Situation wurde riskant als amerikanische Banken begannen, durch die überbewerteten Immobilien gegen finanzierte Kredite auszugeben. Die Spekulationsblase platzte. Der Wert der Immobilien sank innerhalb kurzer Zeit. Die Banken blieben auf ihren Krediten sitzen. Mehrere große amerikanische Banken wie Lehman Brothers und Versicherer wie AIG mussten Konkurs anmelden oder von der Regierung gerettet werden.

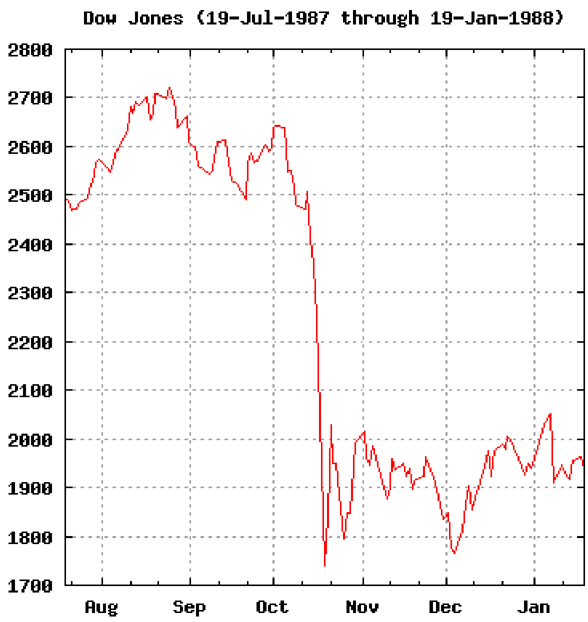

Abbildung 1.1: *Der Schwarze Montag am 19. Oktober 1987 war der erste Börsenkrach nach dem Zweiten Weltkrieg. Der Dow Jones fiel innerhalb eines Tages um 22,6 % (508 Punkte), was den größten prozentualen Abrutsch innerhalb eines Tages in dessen Geschichte darstellt.*

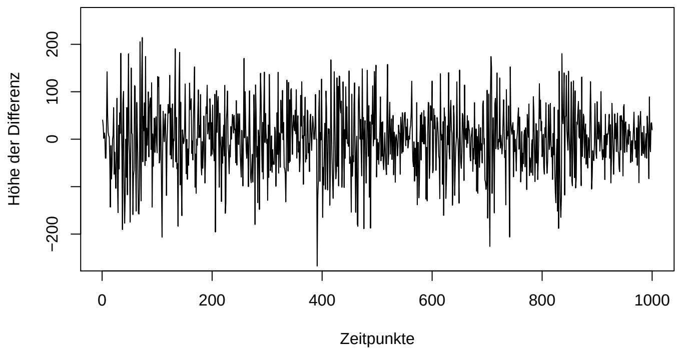

Abbildung 1.2: *Tages-Kurs-Differenzen des FTSE100-Indexes zu 1000 Tageswerten, 01.07.1998-15.05.2002*

1.1.2 Empirische Analyse

Die grundlegenden Kenngrößen zur statistischen Analyse von Zeitreihen sind Erwartungswert und Varianz bzw. Standardabweichung.
Die Berechnung für diese Größen ist allerdings nur für stationäre Reihen sinnvoll. Stationarität bedeutet, dass Momente, die aus verschiedenen Teilreihen berechnet werden, nur gering voneinander abweichen.

- Ein stochastischer Prozess ist schwach stationär, wenn:
  - $\mathbb{E}Y_t = \mu$, für alle $t \in \mathbb{Z}$
  - $\mathbb{V}Y_t = \gamma(0) < \infty$, für alle $t \in \mathbb{Z}$
  - $\mathbb{C}(Y_{t+k}, Y_t) = \gamma(k)$, für alle $t, k \in \mathbb{Z}$
  - Die ersten und zweiten Momente müssen also existieren.
  - Der Erwartungswert ist zeitunabhängig.
  - Die Kovarianzen sind nur vom Zeitabstand (lag) abhängig.
- Ein stochastischer Prozess ist strikt stationär, wenn die gemeinsame Verteilung von $(Y_{t_1+k}, \ldots, Y_{t_n+k})$ für alle endlichen Teilmengen $\{t_1, \ldots t_n\} \subset \mathbb{Z}$ und für alle $k \in \mathbb{Z}$ unabhängig von k ist.
- Es gibt schwach stationäre Prozesse, die nicht strikt stationär und es gibt strikt stationäre Prozesse, die nicht schwach stationär sind.

Die zentrale Lage der Werte einer Zeitreihe $(x_t)_{t=1,\ldots,N}$ wird durch das *arithmetische Mittel* beschrieben.

$$\bar{x} = \frac{1}{N} \sum_{t=1}^{N} x_t$$

Um $\bar{x}$ schwanken die Beobachtungen. Die Stärke der Schwankung wird durch die *empirische Varianz* gemessen.

$$s^2 = \frac{1}{N} \sum_{t=1}^{N} (x_t - \bar{x})^2$$

Die Schwankung kann auch über die Standardabweichung $s = \sqrt{s^2}$gemessen werden, die dieselbe Maßeinheit hat wie das arithmetische Mittel. Seien $(x_t, y_t), t \in \{1, \ldots N\}$ Beobachtungspaare. Die *empirische Kovarianz* ist ein Maß für die Stärke des linearen Zusammenhangs zwischen den $x-$ und $y-$Werten.

$$c = \frac{1}{N} \sum_{t=1}^{N} (x_t - \bar{x})(y_t - \bar{y})$$

Durch Normieren der Kovarianz mit den einzelnen Standardabweichungen erhält man den *Korelationskoeffizienten.*

$$r = \frac{c}{s_x \cdot s_y}$$

Nach Anwendung der Cauchy-Schwarz'schen Ungleichung erkennt man unmittelbar, dass $-1 \leq r \leq 1$ gilt. Liegt r nahe bei 1 bzw. -1 bedeutet dies einen starken positiven bzw. starken negativen linearen Zusammenhang. Gilt $r = 0$ so sind x und y unkorreliert - das ist genau dann der Fall, wenn $c = 0$. Die *Autokovarianzfunktion* c_τ einer Zeitreihe ist definiert wie folgt, wobei $\tau = -(N-1), \ldots, -1, 0, 1, \ldots, (N-1)$ der Lag (Zeitabstand) ist.

$$c_\tau = \frac{1}{N} \sum_t (x_t - \bar{x})(x_{t+\tau} - \bar{x})$$

Die *Autokorrelationsfunktion* ist durch $r_\tau = \frac{c_\tau}{c_0}$ gegeben. Ein Beispiel für einen schwach stationären Prozess ist das weiße Rauschen (white noise) ϵ_t. Es gilt:

- $\mathbb{E}\epsilon_t = 0$
- $\mathbb{V}\epsilon_t = \mathbb{E}\epsilon_t^2 < \infty$
- $\mathbb{E}(\epsilon_{t+k}\epsilon_t) = 0$, für alle $k \neq 0$

Eine Serie von unabhäbhängig, identisch verteilten Zufallsvariablen $(X_t)_{t\in\mathbb{Z}}$ mit endlicher Varianz σ^2 wird als strenger White Noise Prozess bezeichnet. Der strenge White Noise Prozess ist ein komplett zufälliger Prozess, man kann aus der Vergangenheit keine Informationen für die Zukunft beziehen. Im Unterschied zum strengen White Noise Prozess ist beim „normalen“ WhiteNoise Prozess nur die lineare Unabhängigkeit gefordert, andere Abhängigkeiten dürfen existieren.
Das weiße Rauschen ist die Basis um kompliziertere Prozesse zu definieren.
Man betrachte einen Aktienmarkt als ein „faires Spiel“. Das bedeutet, die Information, die man hat, setzt sich aus den vergangenen und den aktuellen Daten zusammen. Man bezeichnet diese Information, die durch den heutigen Zeitpunkt $t \in \mathbb{Z}$ begrenzt ist, als Filtration $\mathcal{F}_{t\in T} = \sigma(\{X_s : s \leq\})$. Zusammengefasst ergibt sich die Eigenschaft der Martingaldifferenz, die eine Zeitreihe $(X_t)_{t\in\mathbb{Z}}$ dann erfüllt, wenn folgendes gilt:

$$\begin{aligned} &E|X_t| < \infty \\ &X_t \text{ ist } \mathcal{F}_t\text{-messbar} \\ &E(X_t|\mathcal{F}_{t-1}) = 0 \quad \forall t \in \mathbb{Z} \end{aligned}$$

Kapitel 2

Modelle

2.1 lineare Modelle

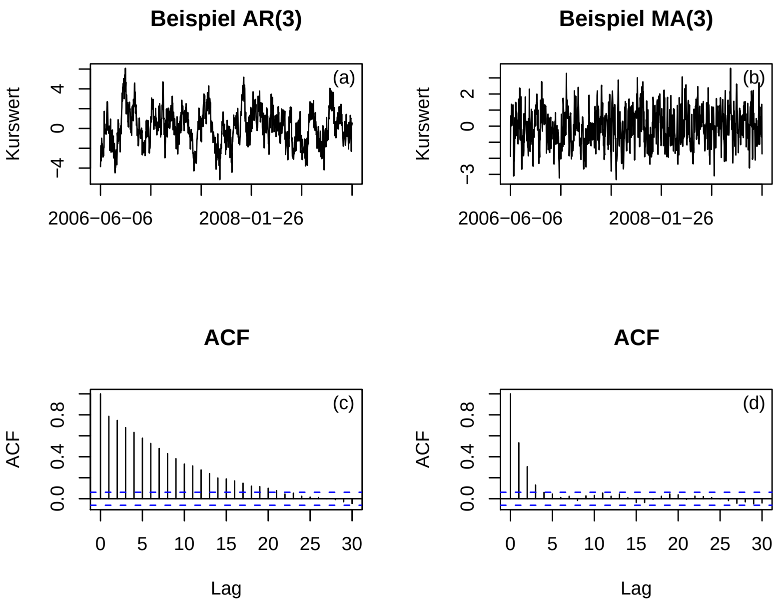

Abbildung 2.1: *Vergleich eines simulierten AR(3)-Prozesses mit einem simulierten MA(3)-Prozess. Anhand der ACF-Darstellungen sind die Unterschiede besonders gut zu sehen. Beim MA(3)-Prozess (rechts) erfolgt ein Schnitt nach Lag 3. Der AR(3)-Prozess (links) dagegen fällt kontinuierlich ab. Werte der Prozesse: MA(3) mit* $\beta_1 = 0.5, \beta_2 = 0.3, \beta_3 = 0.1$*; AR(3) mit* $\alpha_1 = 0.5, \alpha_2 = 0.3, \alpha_3 = 0.1$

2.1.1 Moving Average Process - MA(q) Prozess

Sei (ϵ_t) $WN(\sigma^2)$ Weißes Rauschen. Man nennt

$$y_t = \sum_{k=0}^{q} b_k \epsilon_{t-k}$$

einen moving average (MA) Prozess. Gilt $b_0 \neq 0, b_q \neq 0$ dann ist (y_t) ein MA Prozess der Ordnung q (MA(q)).
MA Prozesse sind stationär und haben ein endliches lineares Gedächtnis. Es gilt:

- $\mathbb{E}y_t = 0$
- $\mathbb{V}y_t = \sigma^2 \sum_{i=0}^{q} b_i^2$

2.1.2 Autoregressiver Prozess - AR(p)

Bei einem AR Prozess fließt im Gegensatz zu einem MA Prozess stärker die Vergangenheit des Prozesses ein und weniger der zufallsbehaftete White Noise Prozess. Dieser Prozess lässt sich darstellen als:

$$y_t = a_1 y_{t-1} + \ldots + a_p y_{t-p} + \epsilon_t$$

für $(\epsilon_t \; WN(\sigma^2), \sigma^2 > 0$ und $a_p \neq 0$. Es gilt:

- $\mathbb{E}X_t = 0$
- $\mathbb{V}X_t = \sigma^2 \sum_{i=0}^{\infty} a^{2i}$.

Beide Modelle betrachten die Zeitreihe nur einseitig. Dies führt uns zu folgender Kombination:

2.1.3 Autoregressiver Moving Average Process - ARMA(p,q)

Sei $(\epsilon_t)_{t \in \mathbb{Z}}$ ein White Noise Prozess mit Erwartungswert 0 und Varianz σ^2. Ein ARMA(p,q) Prozess mit $p, q \in \mathbb{N}$ und Erwartungswert 0 ist ein schwach stationärer Prozess $(X_t)_{t \in \mathbb{Z}}$, welcher folgene Differenzengleichung erfüllt.

$$X_t - \phi_1 X_{t-1} - ... - \phi_p X_{t-p} = \epsilon_t + \theta_1 \epsilon_{t-1} + ... + \theta_q \epsilon_{t-q}, \quad \forall t \in \mathbb{Z}$$

Ein Prozess mit Erwarungswert μ ist dann ein ARMA(p,q) - Prozess, wenn die zentrierte Folge $(X_t - \mu)$ die Eigenschaften eines ARMA(p,q)-Prozesses mit Erwartungswert 0 erfüllt.

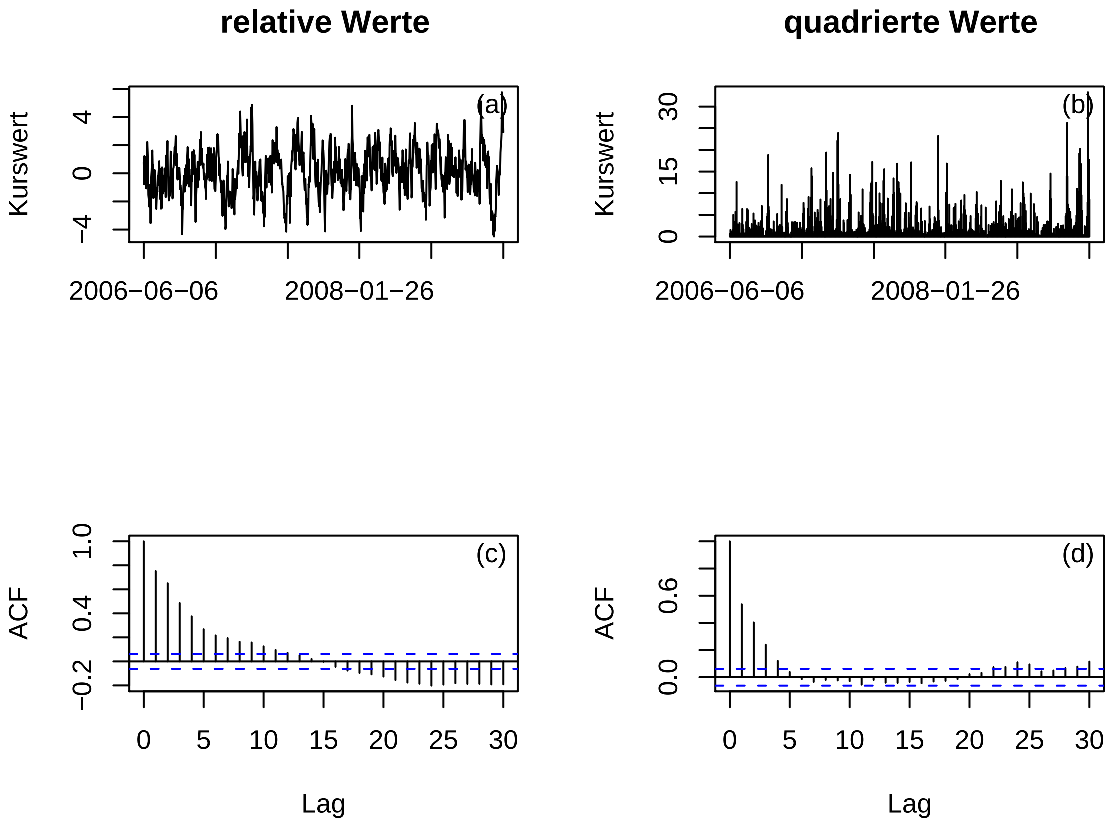

Abbildung 2.2: *Simulation eines ARMA(2,1)-Prozesses. Die Kombination eines MA(1)-Teiles mit einem AR(2)-Teil ergibt einen gemischteren Prozess. Simuliert wurde mit den Werten* $\alpha_1 = 0.5, \alpha_2 = 0.3$ *und* $\beta_1 = 0.1$.

Gehen in ein Modell nur Differenzen von Zeitreihen ein, wie es zum Beispiel bei Aktienkursen oft der Fall ist, so bezeichnet man den ARMA(p,q)-Prozess als ARIMA(p,q,d)-Prozess. Das „I“ steht für „Integrated“ , die Werte nach einer Prognose müssen integriert werden.

2.1.4 ARIMA-Prozess ARIMA(p,q,d)

Ein stochastischer Prozess $(X_t)_{t\in\mathbb{Z}}$ wird als ARIMA(p,q,d)-Prozess mit Schrittweite $d \in \mathbb{N}$ und Ordnungen $p, q \in \mathbb{N}$ bezeichnet, wenn der aus der Differenz gebildete Prozess

$$Y_t = X_t - X_{t-d} \quad \forall t \in \mathbb{Z}$$

die Bedingungen eines ARMA(p,q)-Prozesses erfüllt.
Oft liefern Daten kein brauchbares Ergebnis, wenn man sie mit einer linearen Approximation betrachtet, weil die Abhängigkeiten der vergangenen Zeitwerte nicht zwingend linear sein müssen.
Dies führt zur Betrachtung nicht linearer Modelle. Häufig benutze Modelle verwenden eine nicht konstante, bedingte Varianz.

2.2 nicht lineare Modelle

2.2.1 Autoregressive Conditional Heteroscedasticity - ARCH Prozess

$(Z_t)_{t\in\mathbb{Z}}$ sei ein strenger White Noise Prozess SWN(0,1). Ein strikt stationärer Prozess $(X_t)_{t\in\mathbb{Z}}$ wird als ARCH-Prozess bezeichnet, wenn für alle $t \in \mathbb{Z}$ und einige strikt positive Werte $(\sigma_t)_{t\in\mathbb{Z}}$ folgende Gleichungen erfüllt werden:

$$X_t = \sigma_t Z_t,$$
$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^{p} \alpha_i X_{t-i}^2$$

Damit die Varianz immer positiv ist, wird zusätzlich $\alpha_0 > 0$ und $\alpha_i \geq 0$, $i = 1, ..., p$ gefordert. Der ARCH(p)-Prozess ist ein weit ausschweifender Prozess, das heißt Werte weit abseits des Erwartungswertes (Ausreißer) können mit positiver Wahrscheinlichkeit auftreten, was zum Beispiel dem Aktienmarkt entspricht. Außerdem verschwinden die Autokorrelationswerte für einen Lag $k \neq 0$. Für quadratische Werte allerdings kommen positive Autokorrelationswerte vor. Wenn bei den quadrierten Werten große, positive Autokorrelationswerte für kleine Lags auftreten, so treten Extremwerte in Gruppen auf. Ist der Prozess zusätzlich als schwach stationär gegeben, handelt es sich um einen White Noise Prozess.

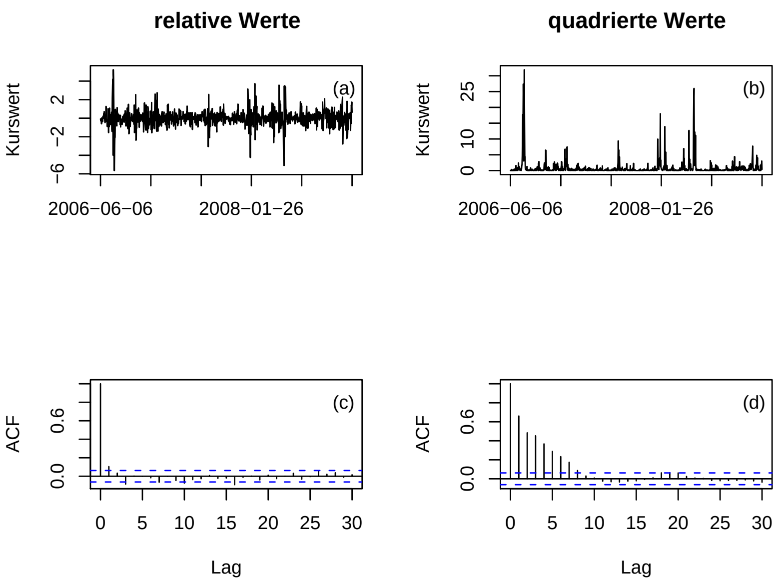

Abbildung 2.3: *Simulation eines ARCH(3)-Prozesses. Die Autokorrelationsfunktion der relativen Werte zeigt wie gering die linearen Abhängigkeiten sind. Die Autokorrelationswerte sind mit denen eines White Noise Prozesses vergleichbar. Erst die quadratischen Werte zeigen Abhängigkeiten auf. Diese positiven Werte zeigen, dass Volatilitätscluster existieren. Diese stehen dafür, dass Extremwerte in Gruppen auftreten. Für diese Simulation wurden die Werte* $\alpha_0 = 0.1, \alpha_1 = 0.5, \alpha_2 = 0.2,$ *und* $\alpha_3 = 0.2$ *verwendet.*

2.2.2 GARCH-Prozess GARCH(p,q)

Der GARCH Prozess ist eine Verallgemeinerung des ARCH Prozesses. Man ergänzt einen ARCH(p) Prozess um einen MA(q)-Teil und erhält damit einen GARCH Prozess.

$(Z_t)_{t\in\mathbb{Z}}$ sei ein strenger White Noise Prozess SWN(0,1). Ein strikt stationärer Prozess $(X_t)_{t\in\mathbb{Z}}$ wird als GARCH(p,q)-Prozess, mit $p, q \in \mathbb{N}$, bezeichnet, wenn für alle $t \in \mathbb{Z}$ und einige strikt positive Werte $(\sigma_t)_{t\in\mathbb{Z}}$ folgende Gleichungen erfüllt werden:

$$X_t = \sigma_t Z_t,$$
$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^{p} \alpha_i X_{t-i}^2 + \sum_{j=1}^{q} \beta_j \sigma_{t-j}^2$$

für $\alpha_0 > 0, \alpha_i \geq 0, i = 1, ..., p$ und $\beta_j \geq 0, j = 1, \dots, q$.
Im Gegensatz zum ARCH-Prozess ist die Varianz eines GARCH-Prozesses abhängig von früheren Varianzen und früheren Werten des Prozesses. Die linearen Abhängigkeiten zwischen den verschiedenen Zeitwerten in diesem Prozess sind Null - er ist unkorelliert. Die quadratischen Werte können aber positive Abhängigkeiten aufweisen.

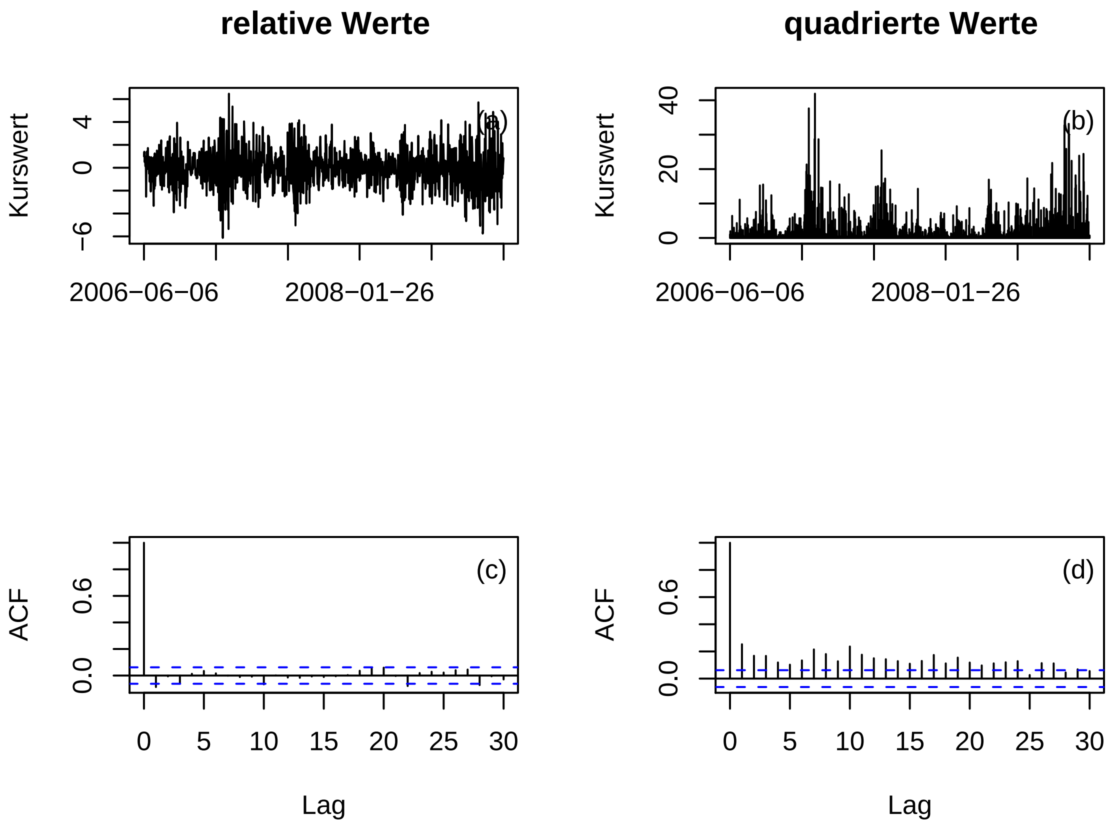

Abbildung 2.4: *Simulation eines GARCH(3,2)-Prozesses. Die linearen Abhängigkeiten sind verschwindend gering. Bei den quadratischen Werten zeigen sich Abhängigkeiten, welche auf Volatilitätscluster hindeuten. Anders als beim ARCH-Prozess sind diese nicht linear fallend.*

2.2.3 Modellanalyse

ARCH(1)

Die Werte eines ARCH-Prozesses sind unabhängig von den vorangegangenen Werten. Eine Lösung ist nur abhängig von den vergangenen Zufallswerten und somit von der Form $X_t =$

$f(Z_t, Z_{t-1}, \ldots)$. Aus der Definition eines ARCH-Prozesses der Ordnung p folgt unmittelbar $X_t^2 = \sigma^2 Z_t^2$. Sei $p = 1$, so gilt:

$$X_t^2 = \alpha_0 Z_t^2 + \alpha_1 Z_t^2 X_{t-1}^2$$

Dies ist eine rekursive Gleichung der Form:

$$Y_t = A_t Y_{t-1} + B_t$$

mit (A_t) und (B_t) als Folgen von unabhängigen, identisch verteilten stochastischen Größen. Voraussetzungen für eine Lösung sind:

$$E(\max\{0, \ln|B_t|\}) < \infty \text{ und } E(\ln|A_t|) < 0$$

Die eindeutige Lösung ist gegeben durch:

$$Y_t = B_t + \sum_{i=1}^{\infty} B_{t-i} \prod_{j=0}^{i-1} A_{t-j}$$

wobei $\sum_{i=1}^{\infty} B_{t-i} < \infty$ fast sicher.
Betrachte den quadrierten ARCH(1) Prozess in rekursiver Form mit $A_t = \alpha_1 Z_t^2$ und $B_t = \alpha_0 T_t^2$. So ergibt sich folgende Lösung:

$$X_t^2 = \alpha_0 \sum_{i=0}^{\infty} \alpha_1^i \prod_{j=0}^{i} Z_{t-j}^2$$

Sei (Z_t) standardnormalverteilt, so wäre die Bedingung für eine strikt stationäre Lösung:

$$\alpha_1 < 3.5$$

Die strikte Stationarität steht also in Abhängigkeit der Verteilung von (Z_t). Aber die notwendige und hinreichende Bedingung für die schwache Stationarität ist:

$$\alpha_1 < 1$$

Der ARCH(1)-Prozess ist ein schwach stationärer White Noise Prozess genau dann wenn $\alpha_1 < 1$. *Die Varianz des schwach stationären Prozesses ist* $\alpha_0/(1-\alpha_1)$.

GARCH(1,1)

Wie oben setze man in die allgemeine Form ein, $p = q = 1$.

$$\sigma_t^2 = \alpha_0 + (\alpha_1 Z_{t-1}^2 + \beta)\sigma_{t-1}^2$$

Analog zum ARCH(1) Prozess ist diese Gleichung wieder der Form $Y_t = A_t Y_{t-1} + B_t$. Hier ist $A_t = \alpha_1 Z_{t-1}^2 + \beta$ und $B_t = \alpha_0$. Die Bedingung für eine strikt stationäre Lösung ist hier $E(\ln(\alpha_1 Z_t^2 + \beta)) < 0$ und es ergibt sich die Lösung:

$$\sigma_t^2 = \alpha_0 + \alpha_0 \sum_{t=1}^{\infty} \prod_{j=1}^{i} (\alpha_1 Z_{t-j}^2 + \beta)$$

Ist $(\sigma_t^2)_{t\in\mathbb{Z}}$ ein strikt stationärer Prozess, so auch $(X_t)_{t\in\mathbb{Z}}$ und $(Z_t)_{t\in\mathbb{Z}}$ ein strenger White Noise Prozess. Die Lösung des GARCH(1,1)-Prozesses ist dann:

$$X_t = Z_t \sqrt{\alpha_0 (1 + \sum_{i=1}^{\infty} \prod_{j=1}^{i} (\alpha_1 Z_{t-j}^2 + \beta)}$$

Der GARCH(1,1)-Prozess ist ein schwach stationärer Prozess genau dann, wenn $\alpha_1 + \beta < 1$. *Die Varianz des schwach stationären Prozesses ist dann* $\alpha_0/(1-\alpha_1-\beta)$.

2.3 Modellanpassung

Um Vorhersagen bezüglich der Entwicklung einer Zeitreihe treffen zu können, muss man zuerst die vorhandenen Daten analysieren. Dazu überprüft man jede Form der Abhängigkeit, falls welche vorhanden sind.
Die allgemeine Annahme ist immer, dass die Daten einem White Noise Prozess entsprechen. Diese Hypothese kann man zum Beispiel mittels Ljung-Box Test prüfen. Dieser Test bezieht sich auf die Autokorrelationsfunktion (ACF) für alle Lags anstatt jeden einzelnen Lag zu testen. Der Ljung-Box Test gehört zur Gruppe der Portmanteau Tests, die für mehrere Autokorrelationskoeffizienten die Signifikanz zu Null testet. Portmanteau-Tests sind reine Signifikanztests, sie testen also nicht gegen eine Gegenhypothese.
Wird die Hypothese mit diesem Test nicht widerlegt, so sind die Daten zufällig. Nach Berechnung von Erwartungswert und Varianz wäre die Analyse somit fertig.
Der zweite und wahrscheinlichere Fall ist, dass die Hypothese verworfen wird. Wesentlich ist dabei die Abhängigkeit zwischen den absoluten Werten. In der Finanzmathematik sind zum Beispiel Risikofaktoren sehr stark voneinander abhängig. Allerdings ist die Abhängigkeit zwischen der relativen, täglichen Risikofaktorenänderung eher gering. Unter Annahme eines „normalen“ White Noise Prozessen ließe sich die Hypothese nicht so einfach wiederlegen. In diesem Fall setze man einen strengen White Noise Prozess voraus, da dieser zusätzlich zu Unkorreliertheit noch Unabhängigkeit fordert.
Sei die ursprüngliche Hypothese nun verworfen, dann unterscheide man zwischen linearer und nicht linearer Korrelation der Zeitwerte und wähle nach dieser Unterscheidung ein passendes Modell. Im linearen Fall ist ein ARMA Modell am geeignetsten, dessen Ordnung noch zu bestimmen ist. Im anderen Fall eignet sich ein GARCH oder ARCH Modell.
Sei also nun bereits ein Modell gewählt, dieses muss nun noch angepasst werden und die Parameter müssen geschätzt werden.

2.3.1 Modellanpassung an ARMA

Die Ordnung eines ARMA-Prozesses wird anhand der Korrelationen und der partiellen Korrelationen bestimmt. Nachdem sich ein ARMA Prozess ja aus einem AR(p) und MA(q) Prozess zusammensetzt, erhalten wir zum Beispiel bei Betrachtung bis zum Lag q einen MA(q) Prozess. Das partielle Korrelogramm hingegen ist allerdings einem AR(p) Prozess sehr ähnlich.
Ein Kriterium zur Auswahl eines Modelles bezeichnet man als Informationskriterium. Für MA, AR und ARMA Prozesse erweisen sich das Akaikes Informationskriterium als günstig. Es wird das ARMA(p,q) Modell gesucht, welches das folgende Kriterium minimiert.

$$AIC(p,q) := \ln \hat{\sigma}^2_{p,q} + 2\frac{(p+q)}{N}$$

Eine weitere Möglchkeit ist das Bayessche Informationskriterium.

$$BIC(p,q) := \ln \hat{\sigma}^2_{p,q} + \frac{(p+q)\ln N}{N}$$

Oder das Hannan-Quinn Informationskriterium.

$$HQ(p,q) := \ln \hat{\sigma}^2_{p,q} + \frac{2(p+q)*c*\ln(\ln N)}{N} \quad \text{mit } c > 1.$$

Je größer der Stichprobenumfang, desto besser werden die Vorhersagen.

2.3.2 Modellanpassung an ARCH und GARCH

Deuten die Daten auf ein ARCH(p) oder GARCH(p,q) Modell hin, greift man vor allem in der Praxis gerne auf die Maximum Likelihood Methode zurück um die Parameter zu schätzen. Man betrachte im folgenden die Modelle für $p = q = 1$.

Maximum Likelihood Methode

Gegeben seien $n+1$ Werte $X_0, \ldots, X_n$. Die gemeinsame Dichte lässt sich anschreiben wie folgt.

$$f_{X_0,\ldots,X_n}(x_0, \ldots, x_n) = f_{X_0}(x_0) \prod_{t=1}^{n} f_{X_t|X_{t-1},\ldots,X_0}(x_t|x_{t-1}, \ldots, x_0).$$

Bei einem ARCH(1) Prozess ist die bedingte Dichte $f_{X_t|X_{t-1},\ldots,X_0}$ nur von X_{t-1} oder äquivalent dazu σ_t abhängig. Somit ergibt sich:

$$f_{X_t|X_{t-1},\ldots,X_0}(x_t|x_{t-1}, \ldots, x_0) = f_{X_t|X_{t-1}}(x_t|x_{t-1}) = \frac{1}{\sigma_t} g(\frac{x_t}{\sigma_t})$$

wobei $\sigma_t = (\alpha_0 + \alpha_1 x_{t-1}^2)$ und $g(z)$ die Dichte vom strengen White Noise Prozess $(Z_t)_{t \in T}$, mit $\mathbb{E}Z_t = 0$ und $\mathbb{V}Z_t = 1$.
Die σ_t Werte sind alle unabhängig voneinander und so kann man alle Werte auf den Anfangswert X_0 zurückführen.

$$f_{X_1,\ldots,X_n|X_0}(x_1, \ldots, x_n|x_0) = \prod_{t=1}^{n} f_{X_t|X_{t-1},\ldots,X_0}(x_t|x_{t-1}, \ldots, x_0).$$

Damit bestimmt man die bedingte Likelihoodfunktion. α_0, α_1 sind die zu schätzenden Parameter, $\sigma_t = (\alpha_0 + \alpha_1 X_{t-1}^2)$ folgt aus der Definition des ARCH Models mit $p = 1$.

$$L(\alpha_0, \alpha_1; \mathbf{X}) = f_{X_1,\ldots,X_n|X_0}(x_1, \ldots, x_n|x_0) = \prod_{t=1}^{n} \frac{1}{\sigma_t} g(\frac{X_t}{\sigma_t}),$$

Für ein ARCH(p) Modell mit $p > 1$ geht man analog vor, man muss die bedingte Dichte für p Anfangswerte berechnen.
Beim GARCH(1,1) Modell wird σ_t durch σ_{t-1} bestimmt. Somit ist der Prozess von den Anfangswerten X_0 und σ_0 abhängig. Die bedingte Dichte für den GARCH Prozess ist dann:

$$f_{X_1,\ldots,X_n|X_0,\sigma_0}(x_1, \ldots, x_n|x_0, \sigma_0) = \prod_{t=1}^{n} f_{X_t|X_{t-1},\ldots,X_0,\sigma_0}(x_t|x_{t-1}, \ldots, x_0, \sigma_0)$$

Die bedingte Dichte $f_{X_t|X_{t-1},\ldots,X_0,\sigma_0}$ ist also nur vom Wert σ_t abhängig, der sich rekursiv aus den vorangegangenen Werten $X_{t-1}, \ldots, X_0$ und σ_0 berechnen lässt.

$$\sigma_t^2 = \alpha_0 + \alpha_1 X_{t-1}^2 + \beta \sigma_{t-1}^2$$

Daraus ergibt sich die bedingte Likelihoodfunktion:

$$L(\alpha_0, \alpha_1, \beta; \mathbf{X}) = \prod_{t=1}^{n} \frac{1}{\sigma_t} g(\frac{X_t}{\sigma_t}) \text{ mit } \sigma_t = \sqrt{\alpha_0 + \alpha_1 X_{t-1}^2 + \beta \sigma_{t-1}^2}.$$

Ein Problem ergibt sich, wenn der Wert σ_0^2 nicht bekannt ist. Ist dies der Fall, schätzt man den Startwert, oder setzt ihn einfach gleich Null. Für großes t nimmt der Einfluss des Startwerts bei $t = 0$ ohnehin ab.
Für einen GARCH(p,q) Prozess mit $p, q > 1$ braucht man $n+p$ ermittelte Daten $X_{-p+1}, \ldots, X_0, X_1, \ldots, X_n$. Die Likelihoodfunktion ist dann bedingt auf die bekannten Daten $X_{-p+1}, \ldots, X_0$ und die unbekannten Werte $\sigma_{-q+1}, \ldots, \sigma_0$, wobei man die Startwerte analog zu vorher schätzen muss.

Kapitel 3

Multivariate Zeitreihen

3.1 Definition

Man betrachte nun mehrere Zeitreihen gemeinsam. Zeitreihen können sich gegenseitig beeinflussen und so entstehen Abhängigkeiten, die sich analysieren lassen.

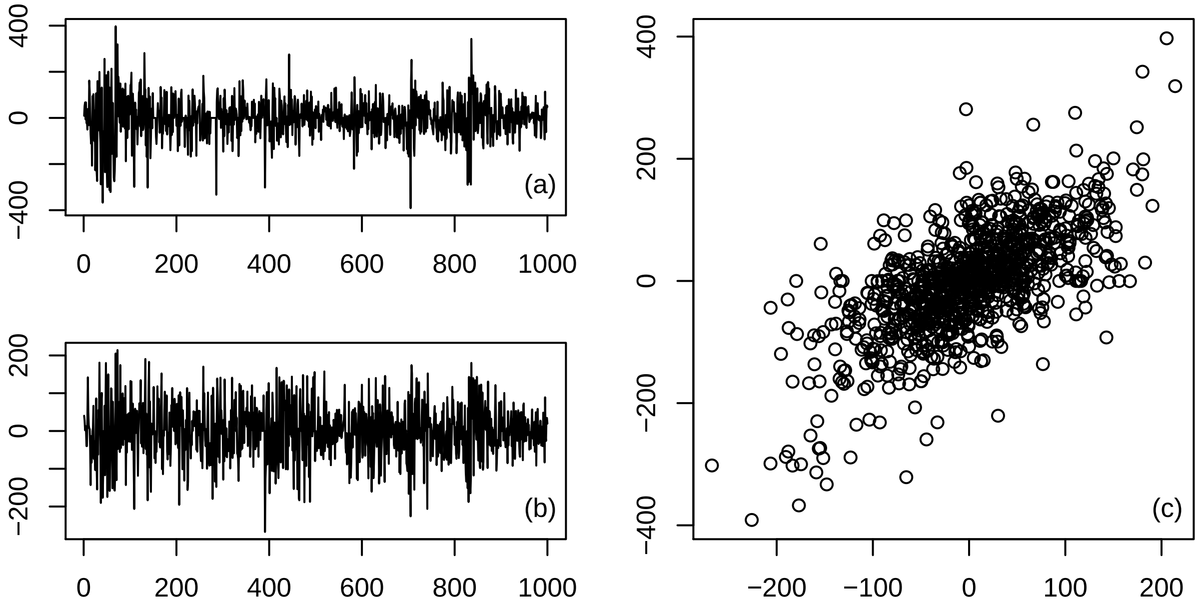

Abbildung 3.1: *Gegenüberstellung der täglichen FTSE100 Kursschwankungen zu den täglichen SMI Kursschwankungen zwischen dem 1. Juli 1998 und dem 15. Mai 2002. (a)zeigt den SMI-Kurs und (b)den FTSE100-Kurs. Abbildung (c) zeigt eine Gegenüberstellung der Werte in einem QQ-Plot. Gut ersichtlich ist, dass die Werte sich in Richtung einer Geraden zueinander ordnen, was eine Abhängigkeit der beiden zueinander aufzeigt.*

Bei der Zusammensetzung mehrerer Zeitreihen hat man - anders als bei univariaten Zeitreihen - nun mehrere Beobachtungswerte pro Zeitpunkt. Während man bei einer Zeitreihe nur die Korrelation zwischen den Zeitpunkten betrachtet, so stehen hier auch die Werte an einem Zeitpunkt in Relation. Somit ergibt sich die

- Definition der Varianz-Kovarianz-Matrix:

$$\Gamma(t,t) = cov(\mathbf{X}_t)$$

Die Komponenten der Matrix sind folgendermaßen definiert:

$$\gamma_{i,j}(t,s) = cov(X_{t,i}, X_{s,j}) = cov(X_{s,j}, X_{t,i}) = \gamma_{j,i}(s,t)$$

$X_{t,i}$ ist der Wert der Zeitreihe i zum Zeitpunkt t.

- Eine multivariate Zeitreihe $(\mathbf{X}_t)_{t\in\ \mathbb{Z}}$ ist strikt stationär, wenn jedes endliche Teilsystem in der Verteilung gleich einem um s Zeitpunkte verschobenen System ist.

$$(\mathbf{X}_t, ..., \mathbf{X}_{t_n}) \stackrel{d}{=} (\mathbf{X}_{t_1+k}, ..., \mathbf{X}_{t_n+k}), \quad \forall t_1, ..., t_{t_n}, k \in \mathbb{Z}, \forall n \in \mathbb{N}.$$

- Eine multivariate Zeitreihe $(\mathbf{X}_t)_{t\in\mathbb{Z}}$ ist schwach stationär, wenn das erste und zweite Moment existieren, die Erwartungswerte konstant sind und die Kovarianzfunktion nur vom zeitlichen Abstand abhängt:

$$\begin{aligned} \mu(t) &= \mu && t \in \mathbb{Z}, \\ \Gamma(t,s) &= \Gamma(t+k, s+k) && t,s,k \in \mathbb{Z}. \end{aligned}$$

- Eine strikt stationäre multivariate Zeitreihe mit endlicher Kovarianzmatrix ist auch schwach stationär. Die Umkehrung gilt nicht.

- Aus der Definition der schwachen Stationärität folgt:

$$\Gamma(t-s, 0) = \Gamma(t,s) \qquad \forall t,s$$

Die Kovarianz von X_t und X_s ist also nur vom zeitlichen Abstand (Lag) der Länge $t-s$ abhängig. Für einen schwach stationären Prozess ist die Varianz- Kovarianz-Matrix also eine Funktion mit nur einem Parameter $\Gamma(h) := \Gamma(h,0)\forall h \in \mathbb{Z}$, wobei $\Gamma(0) = cov(X_t)\forall t$

- Das multivariate Analogon zur Autokorrelationsfunktion ist die Korrelationsmatrixfunktion von $(X_t)_{t\in\mathbb{Z}}$.

$$P(h) := diag(\Gamma(0))^{-1}\Gamma(h)diag(\Gamma(0)) \qquad \forall h \in \mathbb{Z}$$

Ein schwach stationärer Prozess $(X_t)_{t\in\mathbb{Z}}$ ist ein multivariater White Noise Prozess, wenn die Korrelationsmatrix-Funktion aussieht wie folgt.

$$P(h) = \begin{cases} P, & h = 0 \\ 0, & h \neq 0 \end{cases}$$

Ein multivariater White Noise Prozess mit Erwartungswert gleich Null und Varianz-Kovarianzmatrix $\Sigma = cov(X_t)$ wird formal angeschrieben mit $WN(\mathbf{0}, \Sigma)$. Analog zum univariaten White Noise Prozess bestehen lineare Zusammenhänge nur zum aktuellen Zeitpunkt. Es gilt nun für alle Zeitreihen, dass keine vergangenen Daten linearen Einfluss auf die jetzigen Werte haben.
$(X_t)_{t\in\mathbb{Z}}$ ist ein multivariater strenger White Noise Prozess $SWN(\mathbf{0}, \Sigma)$, falls (X_t) eine Folge unabhängig identisch verteilter Zufallsvariablen mit der Varianz-Kovarianz-Matrix ist.
Erfüllt eine multivariate Zeitreihe $(X_t)_{t\in\mathbb{Z}}$ für $\forall t \in \mathbb{Z}$ und eine beliebige Filtrierung $(\mathcal{F}_t)$:

$$E|X_t| < \infty$$
$$E(X_t|\mathcal{F}_{t-1}) = \mathbf{0}$$

so erfüllt die Zeitreihe die Eigenschaft der multivariaten Martingal-Differenz. Es folgt, dass der unbedingte Erwartungswert ebenfalls Null ist und die Kovarianz-Matrix-Funktion $\Gamma(t,s) = 0$ für $t \neq s$ erfüllt.

3.2 Multivariate Prozesse

Die multivariaten Prozesse sind lediglich eine höherdimensionale Verallgemeinerung der bereits kennengelernten Prozesse. Die Parameter und ermittelten Daten, werden im Mehrdimensionalen als Vektoren aufgefasst.

3.2.1 vector ARMA Modell (VARMA)

Sei $(\epsilon_t)_{t\in\mathbb{Z}}$ $WN(\mathbf{0}, \Sigma_\epsilon)$. Der Prozess $(\mathbf{X}_t)_{t\in\mathbb{Z}}$ ist ein VARMA(p,q) Prozess mit Mittelwert Null, falls er die folgende Differenzengleichung löst.

$$\mathbf{X}_t - \Phi_1\mathbf{X}_{t-1} - ... - \Phi_p\mathbf{X}_{t-p} = \epsilon_t + \Theta_1\epsilon_{t-1} + ... + \Theta_q\epsilon_{t-q} \quad \forall t \in \mathbb{Z}$$

Hier sind Φ_i und Θ_j Matrizen der Form $\mathbb{R}^{d\times d}$.
(X_t) ist ein VARMA Prozess mit Mittelwert μ, falls die zentrierte Folge $(X_t - \mu)_{t\in\mathbb{Z}}$ ein VARMA Prozess mit Mittelwert Null ist.
Betrachte nun auch im Mehrdimensionalen den nicht linearen Fall.

3.2.2 Multivariater GARCH Prozess

Sei $(Z_t)_{t\in\mathbb{Z}}$ ein strenger White Noise Prozess $SWN(\mathbf{0}, I_d)$. Der Prozess $(X_t)_{t\in\mathbb{Z}}$ ist ein multivariater GARCH Prozess, falls er strikt stationär ist und folgende Gleichung erfüllt.

$$\mathbf{X}_t = \Sigma_t^{\frac{1}{2}}\mathbf{Z}_t, \quad t \in \mathbb{Z} \tag{3.1}$$

Wobei $\Sigma_t^{\frac{1}{2}} \in \mathbb{R}^{d\times d}$ der Cholesky-Faktor der positiv-definiten Matrix Σ_t ist - dieser Faktor ist messbar bezüglich der Filtrierung $\mathcal{F}_{t-1} = \sigma(\{\mathbf{X}_s : s \leq t-1\})$ der Vergangenheit des Prozesses bis zur Zeit $t-1$.
Ein schwach stationärer Prozess dieses Typs erfüllt die multivariate Martingal-Differenz Eigenschaft.

$$E(\mathbf{X}_t|\mathcal{F}_{t-1}) = E(\Sigma_t^{\frac{1}{2}}\mathbf{Z}_t|\mathcal{F}_{t-1}) = \Sigma_t^{\frac{1}{2}}E(\mathbf{Z}_t) = 0$$

Es folgt, dass ein schwach stationärer GARCH Prozess einem multivariaten White Noise Prozess entspricht. Weiters ergibt sich Σ_t als die bedingte Kovarianz-Matrix.

$$cov(\mathbf{X}_t|\mathcal{F}_{t-1}) = E(\mathbf{X}_t\mathbf{X}_t'|\mathcal{F}_{t-1}) = \Sigma_t^{\frac{1}{2}}E(\mathbf{Z}_t\mathbf{Z}_t')(\Sigma_t^{\frac{1}{2}}(\Sigma_t^{\frac{1}{2}})') = \Sigma_t$$

Die unbedingte Kovarianz Matrix ist dann:

$$\Sigma = cov(\mathbf{X}_t) = E(cov(\mathbf{X}_t|\mathcal{F}_{t-1})) + cov(E(\mathbf{X}_t|\mathcal{F}_{t-1})) = E(\Sigma_t).$$

Daraus ergeben sich die Komponenten der unbedingten Korrelationsmatrix P:

$$\rho_{ij} = \frac{E(\sigma_{t,ij})}{\sqrt{E(\sigma_{t,ii})E(\sigma_{t,jj})}} = \frac{E(\rho_{t,ij}\sigma_{t,i}\sigma_{t,j})}{\sqrt{E(\sigma_{t,i}^2)E(\sigma_{t,j}^2)}},$$

Für die bedingte Korrelationsmatrix P_t werden im Folgenden Modelle vorgestellt, wobei die Schwankung von univariaten Zeitreihen beschrieben wird. Im folgenden einfachsten Modell wird P_t als konstant angenommen ($\forall t$).

3.2.3 (constant conditional correlation) CCC GARCH Prozess

Der Prozess $(X_t)_{t\in\mathbb{Z}}$ ist ein CCC GARCH Prozess falls die bedingte Kovarianz-Matrix die Form $\Sigma_t = \Delta_t P_c \Delta_t$ hat und diese 1. und 2. erfüllt.

1. P_c ist eine konstante, positiv-definite Korrelationsmatrix.

2. Δ_t ist eine diagonale Varianzmatrix, Einträge $\sigma_{t,k}$

$$\sigma^2_{t,k} = \alpha_{k0} + \sum_{i=1}^{p_k} \alpha_{ki} X^2_{t-i,k} + \sum_{j=1}^{q_k} \beta_{kj}\sigma^2_{t-j,k} \text{ für } k = 1, ..., d$$
$$\text{mit } \alpha_{k0} > 0, \alpha_{ki} \geq 0, i = 1, \ldots, q_k, \beta_{kj} \geq 0, j = 1, \ldots, p_k$$

In diesem Modell kann man die Beobachtungen schreiben als $X_t = \Delta_t P_c^{\frac{1}{2}}$. Dies kann umgeschrieben werden auf $\mathbf{X}_t = \Delta_t \widetilde{\mathbf{Z}}_t$ wobei $\widetilde{\mathbf{Z}}_t$ einen strengen White Noise Prozess SWN($\mathbf{0}$,P_c) beschreibt.
Betrachte man nun die dynamische Verallgemeinerung des CCC GARCH Modells.

3.2.4 DCC GARCH Prozess

Ein GARCH-Prozess $(\mathbf{X}_t)_{t\in\mathbb{Z}}$, mit Varianz-Kovarianz-Matrix der Form $\Sigma_t = \Delta_t P_c \Delta_t$, wird als DCC-GARCH-Prozess bezeichnet, wenn Δ_t eine Form, wie in der Definition des CCC GARCH Prozesses hat und die konstante Korrelationsmatrix P_t für alle $t \in \mathbb{Z}$ Folgendes erfüllt.

$$P_t = \wp((1 - \sum_{i=1}^{p} \alpha_i - \sum_{j=1}^{q} \beta_j)P_c + \sum_{i=1}^{p} \alpha_i \mathbf{Y}_{t-i}\mathbf{Y}'_{t-i} + \sum_{j=1}^{q} \beta_j P_{t-j})$$

P_c ist eine positiv-definite Korrelationsmatrix und $\wp(\Sigma) := (\Delta(\Sigma))^{-1}\Sigma(\Delta(\Sigma))^{-1}$. $\mathbf{Y}_t = \Delta_t^{-1}\mathbf{X}_t$ ist wieder ein Prozess und für die Koeffizienten gilt $\alpha_i \geq 0, \beta_j \geq 0$ und $\sum_{i=1}^{p} \alpha_i - \sum_{j=1}^{q} \beta_j < 1$.

3.3 Modellanpassung

Analog zum univariaten Fall entscheidet man auch im multivariaten Fall nach Testen der vorhandenen Daten für ein Modell. Für eine bestmögliche Vorhersage muss man die Parameter ebenso wie bei univariater Betrachtung anpassen.
Deuten unsere Daten beispielsweise auf ein multivariates GARCH Modell hin, kann man ebenso die Maximum Likelihood Methode anwenden.
Wieder geht man von einem Modell mit $p = q = 1$ und $n+1$ vorhandenen Werten $\mathbf{X}_0, \mathbf{X}_1, \ldots, \mathbf{X}_n$ aus. Die aktuellen Werte $\mathbf{X}_t$ hängen wieder vom vorherigen Wert $\mathbf{X}_{t-1}$ und der Varianz-Kovarianz-Matrix Σ_t ab. Weil Σ_t aus Σ_{t-1} bestimmt wird, ist die geimensame bedingte Dichte von $\mathbf{X}_0, ..., \mathbf{X}_n$ letztendlich nur mehr von den Anfangswerten $\mathbf{X}_0$ und Σ_0 abhängig. Die bedingte Dichte ist somit:

$$f_{\mathbf{X}_1,...,\mathbf{X}_n|\mathbf{X}_0,\Sigma_0}(\mathbf{x}_1, ..., \mathbf{x}_n|\mathbf{x}_0, \Sigma_0) = \prod_{t=1}^{n} f_{\mathbf{X}_t|\mathbf{X}_{t-1},...,\mathbf{X}_0,\Sigma_0}(\mathbf{x}_t|\mathbf{x}_{t-1}, ..., \mathbf{x}_0, \Sigma_0).$$

Die multivariate Dichte von $\mathbf{Z}_t$ mit g($\mathbf{z}$) ist:

$$f_{\mathbf{X}_t|\mathbf{X}_{t-1},...,\mathbf{X}_0,\Sigma_0}(\mathbf{x}_t|\mathbf{x}_{t-1}, ..., \mathbf{x}_0, \Sigma_0) = |\Sigma_t|^{-1/2} g(\Sigma_t^{-1/2} x_t).$$

Wobei Σ_t dabei die Abbildungsmatrix einer Funktion mit den Parametern $\mathbf{x}_{t-1}, ..., \mathbf{x}_0$ und Σ_0 ist. $g(\mathbf{z})$ wird üblicherweise so gewählt, dass wir $g(\mathbf{z}) = h(\mathbf{z}'\mathbf{z})$ setzen können, für eine skalare Funktion h.
Einsetzen in die bedingte Likelihood-Funktion ergibt:

$$L(\theta; \mathbf{X}_1, ..., \mathbf{X}_n) = \prod_{t=1}^{n} |\Sigma_t|^{-1/2} h(\mathbf{X}_t' \Sigma_t^{-1} \mathbf{X}_t),$$

wobei θ die Parameter der multivariaten GARCH-Funktion beinhaltet. Bei geschickter Wahl von dem unbekannten Anfangswert Σ_0 ist es möglich mit der Likelihood-Funktion Wahrscheinlichkeiten für die Vektoren der Parameter zu bestimmen. Bei der Auswertung erhält man sowohl einen Schätzwert für die Parameter als auch ein Konfidenzintervall. Aus diesen Ergebnissen kann man schließlich die Werte für das nächste Zeitintervall schätzen.

Kapitel 4

Literaturverzeichnis

- *Alexander J. McNeil, Rüdiger Frey, Paul Embrechts*: Quantitative Risk Management, Kapitel 4
- *Rainer Schlittgen, Bernd Streitberg*: Zeitreihenanalyse, 8. Auflage
- *Wolfgang Scherrer*: Folien zur LVA 105.078, Einführung in Stochastische Prozesse und Zeitreihenanalyse
- *http://de.wikipedia.org/Zeitreihenanalyse*
- *http://www.r-project.org*