Zeitreihen in der Finanzmathematik

von ASANI Silvia

Inhaltsverzeichnis

1 Einleitung

1.1 Was sind Zeitreihen?

Zeitreihen sind Sammlungen von Messwerten, die zu bestimmten Zeitpunkten aufgenommen wurden.

Definition:
Eine **Zeitreihe** ist eine (endliche) Folge von zeitlich geordneten Beobachtungen/Messwerten.

$$((t_k, y_k)|k = 1, 2, ..., T)$$

$t_k \in \mathbb{R}$: Zeitpunkt, $t_1 < t_2 < ... < t_T$.
$y_k \in \mathbb{R}^n$: Messwert(e) zum Zeitpunkt t_k.

Eigenschaften von Zeitreihen:

- Die Werte der untersuchten Größe sind in einem (geringen) Maß abhängig von einander.
- Die absoluten oder quadrierten Werte dagegen zeigen greifbarere Abhängigkeiten.
- Die auf die Information bedingten Erwartungswerte sind nahe Null.
- Die Varianzen tendieren dazu, sich über die Zeit zu verändern.
- Die Serie von Werten scheint leptokurtisch zu sein. Das heißt die Zeitreihe dürfte eine große Wölbung haben.
- Extreme Werte scheinen in Gruppen aufzutreten.

Definition:
Die **Zeitreihenanalyse** ist die Disziplin die sich mit der mathematisch - statistischen Analyse von Zeitreihen und der Vorhersage ihrer künftigen Entwicklung beschäftigt.

Anwendungen der Zeitreihenanalyse:

- besseres Verständnis des zugrunde liegenden Systems charakteristische Eigenschaften von Finanzmärkten
- Prognose:
  Prognose der Arbeitslosenzahlen für das kommende Jahr

- Erkennen von Veränderungen:
  online Monitoring von Intensivstations-Patienten
- Extraktion von Kennzahlen:
  automatische Sprechererkennung
- Signalverarbeitung
  Rauschunterdrückung

1.2 Zusammenhang von Zeitreihen mit der Finanzmathematik

Die Modelle werden einmal berechnet, jedoch kommen in diesen Modellen Parameter vor, welche häufig nicht bekannt sind. Sie müssen dem zufolge aus historischen Daten "geschätzt"werden. Was nun die Wirtschaftswissenschaften, insbesondere die Finanzwelt, auszeichnet ist, dass diese Daten meist in Form von Zeitreihen vorkommen.
Es werden zufällige Veränderungen in diskreter oder kontinuierlicher Zeit betrachtet. Klassische Zeitreihen sind gleitende Mittel (MA), Auto-Regressive-Prozess (AR) oder in der Kombination beider Fälle ARMA-Prozesse (näheres im 2. Abschnitt: Modelle für Zeitreihen).

In der **Finanzmathematik** kann man mit ihnen unter anderem Liquiditätsentwicklungen oder Wertpapierkurse berechnet. Es werden gewisse Variationen dieser Zeitreihen zur Modellierung von Renditen verwendet, nämlich ARCH-, GARCH-Zeitreihen. Typische Abhängigkeitsfunktionen für Zeitreihen sind die Autokovarianz- und Autokorrelationsfunktionen.

Vorhersage von Preisen:
Mit vielfältigsten Methoden (auch aus der Zeitreihenanalyse) versucht man, Aktienkurse vorherzusagen. (ein Beispiel für eine Vorhersage siehe *Abbbildung 1*)

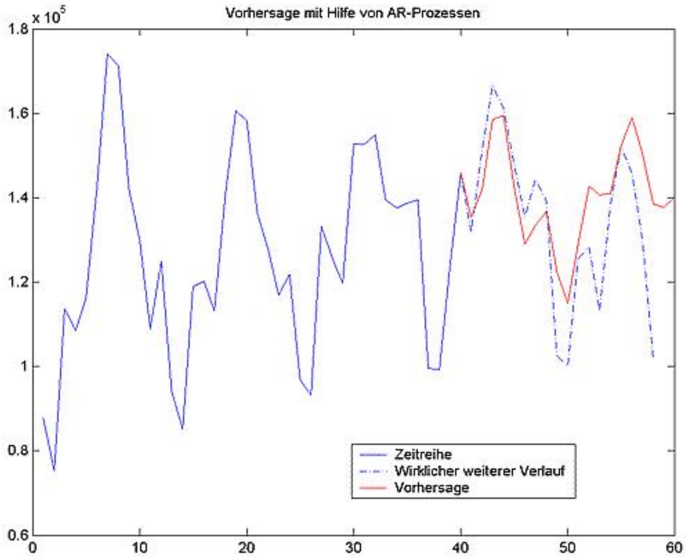

Abbildung 1:
Beispiel einer Vorhersage mit einem AR-Prozess

Einsatz im Riskmanagement:
Man kann realistischere Modelle für die Logreturns aufstellen (ARCH und GARCH), mit denen man die Wahrscheinlichkeiten für große Verluste (im Modell!) berechnen kann.

Schätzung von Zinsstrukturen:
Zinsen von langfristig angelegtem Geld sind 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.

Je nach Verwendung dient eine Zeitreihenanalyse also zum Beispiel dazu um historische Zeitreihen verkürzt zu beschreiben, Prognosen zu formulieren und Veränderungen zu erkennen.

Hier ein Beispiel für einen Aktienindex der Dow Jones Industrial Average:

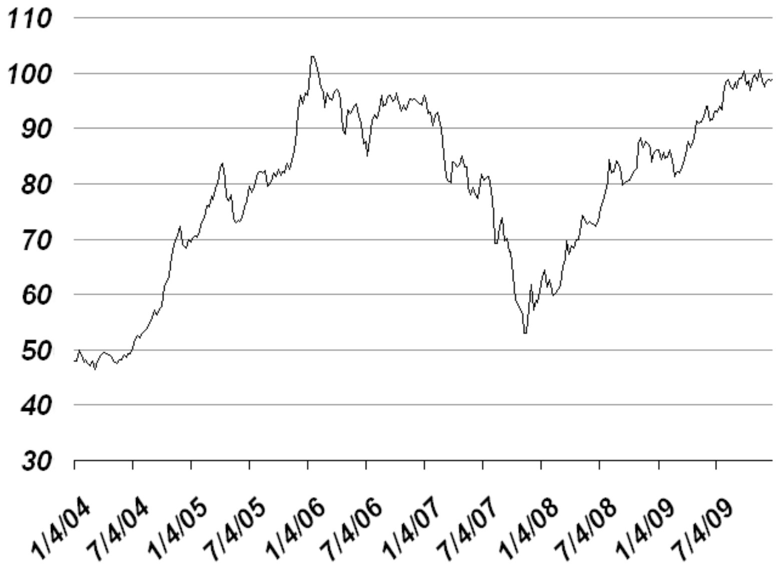

Abbildung 2:
Dow-Jones-Index von Januar 1904 bis Juli 1909

Zeitreihen sind auch in anderen Branchen begehrt:
In der **Ökonometrie** ergeben sich aus ihnen zum Beispiel Daten über Arbeitslosenquoten oder das Bruttosozialprodukt.

In der **Meteorologie** gibt man mit Zeitreihen Temperaturwechsel, Windgeschwindigkeiten und Kriterien für den Klimawandel an.

1.3 Historischer Rückblick

Prinzipiell können die Methoden der Zeitreihenanalyse überall dort eingesetzt werden, wo Daten als zeitabhängige Größen erfasst werden.

Erste Ansätze zu einer Erfassung und Aufbereitung größerer Datenmengen finden sich bereits bei den antiken Hochkulturen. Die Statistik diente zur Erfassung von Bevölkerungszahlen und damit dem Steueraufkommen sowie dem Außenhandel, der Ermittlung der Wehrfähigkeit und der Analyse der Sternenbewegungen.
Im mittelalterlichen Europa wurden nur wenige Daten erfasst. Erst mit dem Aufkommen der absolutistisch-bürokratischen Staaten gewann die Statistik wieder mehr an Bedeutung. Während auf dem europäischen Festland zunächst diverse Daten nur aufgenommen wurden, um den aktuellen Zustand

quantitativ zu beschreiben, wurden im England des 17. Jahrhunderts eine Reihe von Analysetechniken entwickelt, zur Gewinnung von Aufschlüssen über die Funktionsweise von Staat, Wirtschaft und Gesellschaft. Hier wurden die Grundideen für die heutige Zeitreihenanalyse gewonnen. Auf dem europäischen Kontinent begannen Wirtschaftswissenschaftler etwa um 1850 damit, ökonomische Zeitreihen zu erfassen und die darin beobachteten Schwankungen zu analysieren. Der zunehmenden Bedeutung dieser Zeitreihenanalyse trägt die Tatsache Rechnung, dass seit 1880 jährlich diverse, Deutschland betreffende Zahlen in dem "statistischen Jahrbuch" veröffentlicht werden.
Mit der zunehmenden Verbreitung des marktwirtschaftlichen Modells und dem damit verbundenen Anstieg der Bedeutung von Handelsbörsen, wurde die Analyse und Vorhersage von Börsenkursen äußerst interessant. Dieser Trend wurde in der zweiten Hälfte des 20.Jahrhunders nochmals dadurch verstärkt, dass immer mehr Rechenleistung zur Auswertung der aufgenommenen Daten zur Verfügung stand.
Heute gibt es komplexe Systeme zur Echtzeitanalyse von Börsenkursen. Zeitreihen werden für Klimamodelle und Erdbebenvorhersage genauso verwendet wie für die Optimierung verschiedenster Parameter in Produktionsprozessen.

2 Modelle für Zeitreihen

Definition:
Ein **stochastischer Prozess** ist eine Familie von reellen Zufallsvektoren, die auf gemeinsamen Wahrscheinlichkeitsraum $(\Omega, \mathcal{F}, \mathbb{P})$ definiert sind:

$$X_t : (\Omega \times \mathbb{T}) \to \mathbb{R}^n$$

$$(\omega, t) \longmapsto X_t(\omega)$$

Bemerkung:
Die Indexmenge $\mathbb{T}$ wird als Zeit interpetiert.

2.1 Stationäre Prozesse

Eine Zeitreihe wird stationär bezeichnet, wenn ...
...sich der Mittelwert nicht ändert,
...sich die Varianz nicht ändert,
...periodische Variationen nicht vorkommen.

Die Zeitreihenanalyse im engeren Sinne befasst sich mit solchen stationären Zeitreihen. Man muss zuerst eine nicht-stationäre Zeitreihe stationär machen. Dann kann man versuchen die Modelle (siehe Abschnitte 2.2 - 2.6) auf die Residuen anwenden.

Definition:
Ein stochastischer Prozess ist **schwach stationär**, wenn

- $\mathbb{E}X_t = \mu$ für alle $t \in \mathbb{Z}$
- $\mathbb{V}X_t = \gamma(0) < \infty$, für alle $t \in \mathbb{Z}$
- $\mathbb{C}(X_{t+k}, X_t) = \gamma(k)$, für alle $t, k \in \mathbb{Z}$

Bemerkung:

Die ersten und zweiten Momente müssen also existieren.

Der Erwartungswert ist zeitunabhängig.

Die Kovarianzen sind nur vom Zeitabstand (lag) abhängig.

Definition:

Ein stochastischer Prozess ist **strikt stationär**, wenn die gemeinsame Verteilung von $(X_{t_1+k}, \ldots, X_{t_n+k})$ für alle endlichen Teilmengen $\{t_1, \ldots t_n\} \subset \mathbb{Z}$ und für alle $k \in \mathbb{Z}$ unabhängig von k ist.

Bemerkung:
Es gibt schwach stationäre Prozesse, die nicht strikt stationär und es gibt strikt stationäre Prozesse, die nicht schwach stationär sind.

Hier werden Zeitreihen als Trajektorie eines stochastischen Prozesse interpretiert, d.h. stochastische Prozesse werden als Modell für die beobachtete Zeitreihen verwendet.

Definition:
Die **Autokovarianzfunktion** *(Autokorrelationsfunktion)* eines *(schwach)* stationären Prozesses ist definiert durch

$\gamma(k) = \mathbb{C}(X_{t+k}, X_t) = \mathbb{E}(X_{t+k} - \mathbb{E}X_t)(X_t - \mathbb{E}X_t)$
$\rho(k) = \frac{\gamma(k)}{\gamma(0)}$

Eigenschaften:

- Symmetrie: $\gamma(k) = \gamma(-k)$
- Positivität: $(\gamma(k)|k \in \mathbb{Z})$ ist positiv semidefinit.
- $-1 \leq \rho(k) \leq 1$. ($\rho(0) = 1$.)

Interpretation:

Die Autokovarianz- (Autokorrelations-) Funktion zeigt die zeitliche (lineare) Abhängigkeitsstruktur des Prozesses.

Definition:(White-Noise $(\epsilon_t) \sim WN(\mu, \sigma^2)$)
Ein **Weißes Rauschen** oder auch White-Noise, ist ein schwach stationärer Prozess $(X_t)_{t\in\mathbb{Z}}$ mit der Autokorrelationsfunktion:

$$\rho(k) = \begin{cases} 1, & k = 0 \\ 0, & k \neq 0 \end{cases}$$

Einfachstes Beispiel für einen White-Noise-Prozess ist eine Serie von unabhängig, identisch verteilen Zufallsvariablen. Aus einem weißen Rauschen (White-Noise) lassen sich drei wichtige Klassen linearer Zeitreihenmodelle ableiten, nähreres zu den Modellen in den Abschnitten (2.2, 2.3, 2.4).

Definition:
Eine Serie von unabhängig, identisch verteilten Zufallsvariablen $(Z_t)_{t\in\mathbb{Z}}$ mit endlicher Varianz σ^2 wird als **strikter White Noise Prozess** $(Z_t)_{t\in\mathbb{Z}} \sim SWN(\mu, \sigma^2)$ bezeichnet.

Mit dem strikten White-Noise-Prozess haben wir einen komplett zufälligen Prozess definiert. Wir können aus der Vergangenheit keine Informationen für die Zukunft beziehen. Wir werden immer hoffen, dass vorliegende Daten nicht einem solchen Prozess entsprechen. Bei einem generellen White-Noise-Prozess dagegen haben wir nur keine linearen Abhängigkeiten. Weitere Abhängigkeiten können jedoch existieren. Wir werden spätere Modelle immer auf einen dieser beiden Prozesse aufbauen um unseren Zufallswert einzubeziehen.

Generell werden wir in diesem Dokument unter einem Prozess (ϵ_t) immer einen White-Noise-Prozess und unter einem Prozess (Z_t) immer einen strikten White-Noise-Prozess verstehen.

Wir gehen davon aus, dass Aktienmärkte und Risikofaktoränderungen nichts anderes als faire Spiele sind. Das bedeutet, wir haben zum aktuellen Zeitpunkt, außer den Informationen aus der Vergangenheit und die aktuellen Werte, keine weiteren Informationen. Diese Eigenschaft fassen wir in der Martingal-Differenz zusammen.

Definition:
Erfüllt eine Zeitreihe $(X_t)_{t\in\mathbb{Z}}$ folgende Eigenschaften:

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

sagt man sie erfüllt die Eigenschaft der **Martingal-Differenz**.

2.2 MA(q)-Prozess

Definition:
Sei $(\epsilon_t) \sim WN(\sigma^2)$ weißes Rauschen. Einen Prozess

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

nennt man **moving average (MA)-Prozess**. Gilt $b_0 \neq 0, b_q \neq 0$ dann ist (X_t) ein MA-Prozess der Ordnung q (MA(q)).

Eigenschaften:

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

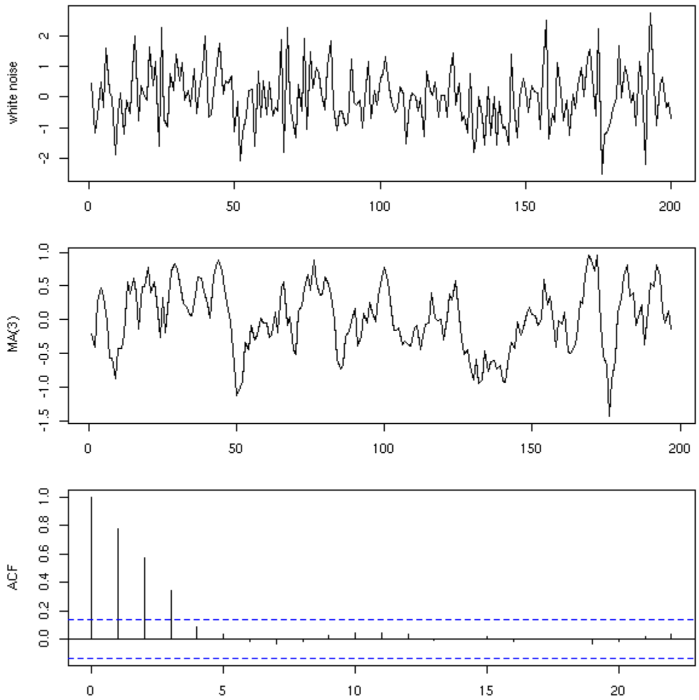

Abbildung 3: Simulation eines MA(3)-Prozesses ACF-Darstellung

2.3 AR(p)-Prozess

Definition:
Sei $(\epsilon_t) \sim WN(\sigma^2)$ ein White-Noise-Prozess. Dann nennen wir einen Prozess X_t **Autoregressiver Prozess der Ordnung p** (AR(p)), wenn er eine Darstellung der Form

$$X_t = a_1 X_{t-1} + \ldots + a_p X_{t-p} + \epsilon_t$$

für $a_p \neq 0$ hat.

Eigenschaften:

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

Dies ist in der Tat etwas wie eine Regression. Nur sind die Regressoren vergangene Werte der Zeitreihe; X wird auf sich selbst regressiert, deshalb Autoregressiver Prozess. Wir werden immer p für die Ordnung der autoregressiven Prozesse benutzen und q für die MA-Prozesse.

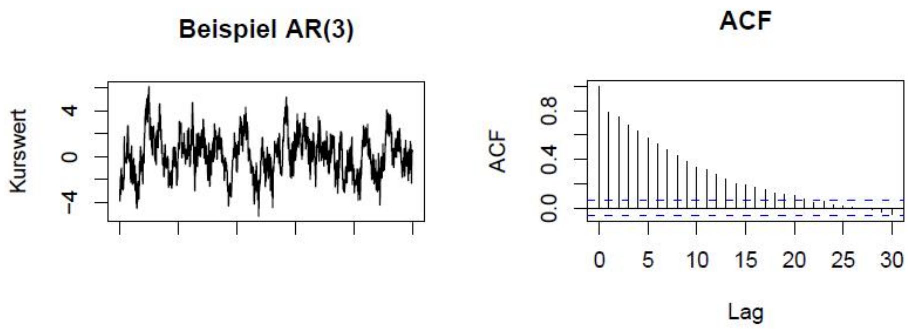

Abbildung 4:
Simulation eines AR(3) Prozesses ACF-Darstellung

2.4 ARMA(p,q)-Prozess

Die MA- und AR-Prozesse kann man als die Grundbausteine der Zeitreihenanalyse bezeichnen. Die Ordnung kann frei gewählt werden, je grösser die Ordnung, desto mehr Parameter stehen uns zur Verfügung, um ein Modell an eine gegebene empirische Reihe anzupassen. Wenn viele Parameter zur Verfügung stehen, kann man fast jeden endlichen Datensatz an ein Modell anpassen. Man möchte aber ein Modell, bei dem möglichst wenig Parameter vorkommen, welche zudem auch eine reale Bedeutung (Interpretation) haben.
Zum Beispiel mit einem ARMA(1,1) (Autoregressiver Moving Average Process) Prozess (mit 2 Parameter) kann man viel umfassender Eigenschaften von empirschen Zeitreihen einfangen als mit dem „einfachen“ MA- oder AR-Prozessen.

Definition:
Sei $(Z_t)_{t\in\mathbb{Z}} \sim SWN(0,1)$ ein White-Noise-Prozess.
Ein **Autoregressiver Moving Average Process** (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:

$$X_t - a_1 X_{t-1} - ... - a_p X_{t-p} = \epsilon_t + b_1\epsilon_{t-1} + ... + b_q\epsilon_{t-q}, \quad \forall t \in \mathbb{Z}$$

Ein ARMA(p, q)-Prozess besteht...
...aus dem AR(p)-Polynom: $a(L) = 1 - a_1 L - ... - a_p L^p$
(L ist der Lag-Operator)
...aus dem MA(q)-Polynom: $b(L) = 1 + b_1 L + ... + b_q L^q$p.

2.5 ARIMA(p,q,d)-Prozess

Gehen in ein Modell nur Differenzen von Zeitreihen ein, wie es bei Aktienkursen oder Risikoschwankungen oft der Fall ist, so bezeichnet man den ARMA(p,q)-Prozess als **ARIMA(p,q,d)-Prozess**. Das **I** steht für „Integrated“ und gibt einen Hinweis darauf, dass die Werte nach einer Prognose wieder integriert werden müssen.

Definition:
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.

2.6 Nicht-lineare Zeitreihen

2.6.1 ARCH-Prozess

Das Modell der autoregressiven bedingt heteroskedastischen Zeitreihenmodelle (autoregressive with conditional hetoeroscedasticity, kurz ARCH-Modelle) kann, trotz ihrer recht einfachen Struktur, die real auftretende Phänomene bei Finanzzeitreihen einigermaßen „gut" beschreiben.

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

$$X_t = \sigma_t Z_t, \tag{1}$$

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

Damit die Varianz immer positiv ist, wird zusätzlich $\alpha_0 > 0$ und $\alpha_i \geq 0$, $i = 1, ..., p$ gefordert.

2.6.2 GARCH-Prozess

Es gibt viele Verallgemeinerungen von ARCH. Die bekannteste und in der Finanzwelt wichtigste ist der GARCH-Prozess. Das G kommt von Generalized. AR-Polynom mit einem MA-Polynom wird ja zu einem ARMA. Genauso erhält man aus ARCH mit einem MA-Polynom den GARCH. Es überrascht deshalb auch nicht besonders, dass während ein ARCH im Quadrat ein AR ist, ein GARCH im Quadrat ein ARMA ist. Also der GARCH-Prozess bildet sich aus der Ergänzung eines ARCH(p)-Prozesses um einem MA(q)-Teil.

Definition:
Sei $(Z_t)_{t\in\mathbb{Z}} \sim SWN(0,1)$ ein strikter White-Noise-Prozess.
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 streng positive Werte $(\sigma_t)_{t\in\mathbb{Z}}$ folgende Gleichungen erfüllt werden:

$$X_t = \sigma_t Z_t, \tag{3}$$

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^{p} \alpha_i X_{t-i}^2 + \sum_{j=1}^{q} \beta_j \sigma_{t-j}^2 \tag{4}$$

für $\alpha_0 > 0$, $\alpha_i \geq 0$, $i = 1, ..., p$ und $\beta_j \geq 0$, $j = 1, ..., q$.

Bemerkung:
Im Gegensatz zum ARCH-Prozess ist die Varianz eines GARCH-Prozesses abhängig von früheren Varianzen und früheren Werten des Prozesses.

Mit dem GARCH(p,q)-Modell können viele Möglichkeiten modelliert werden. Der Prozess selbst ist unkorreliert, dass bedeutet, dass die linearen Abhängigkeiten zwischen den verschiedenen Zeitwerten null sind. Die quadratischen Werte dagegen können positive Abhängigkeiten aufweisen.

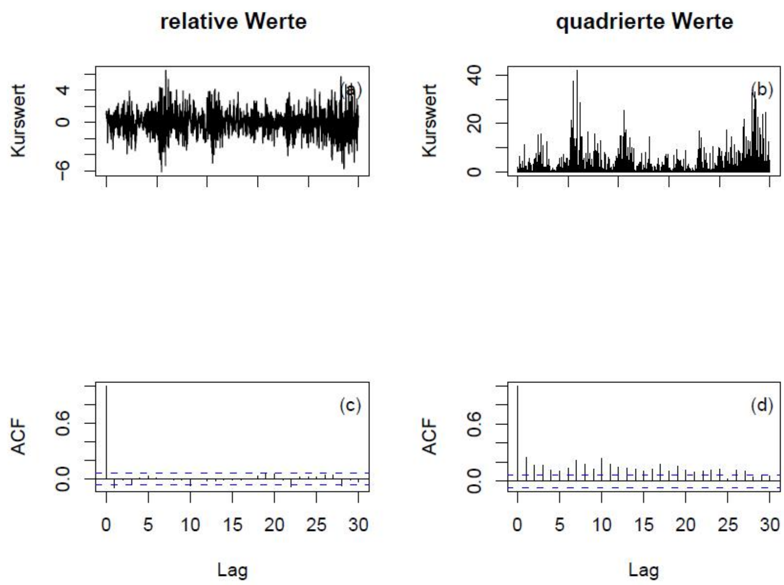

Abbildung 5: Simulation eines GARCH-Modells

3 Vorgehensweise von Zeitreihenanalyse

In diesem Abschnitt wird eine Analyse der genannten Schätzverfahren und die Überprüfung ihrer Anwendbarkeit am Markt vorgestellt. Anhand der Ergebnisse statistischer Analysen soll eine Entscheidung zugunsten eines der Modelle getroffen werden.

3.1 Modellauswahl

Bei der Modellauswahl, ist die am einfachste Methode, die graphische Darstellung der empirischen Zeitreihenwerte zu analysiert. Im Rahmen der graphischen Analyse lassen sich erste Schlüsse über das Vorliegen von Trends, Saisonalitäten, Ausreißern, Varianzinstationarität sowie sonstiger Auffälligkeiten ziehen. Also werden den Zeitreihen zunächst verschiedene statistische Tests unterzogen. Dadurch kann überprüft werden, ob ARCH-Effekte in Form von Autokorrelationen in den Zeitreihen enthalten sind. Wenn eine lineare Korrelation zwischen den Zeitwerten vorhanden ist, dann handelt es sich eher um ein ARMA-Modell. Jedoch bei nicht-linearen Korrelationen zwischen Zeitwerten kann man auf ein ARCH- oder GARCH-Modell schließen.
Es kann jedoch sinnvoll sein, einen ARCH(p)-Prozess als Modell mit weniger als p Parametern zu modellieren, wenn es Parameter gibt, die einen vernachlässigbar kleinen Beitrag zum Gesamtprozess leisten. Aus diesem Grund gibt es die folgenden Kriterien, die zur Modellordnungswahl verwendet werden.

3.1.1 Informationskriterium

Die Auswahl der Ordnung p des ARCH-Modells ist sehr wichtig, denn je größer p ist, desto allgemeiner ist das Modell und kann daher potentiell die beobachteten Renditen genauer beschreiben. Andererseits ist es bei einem großen Wert von p nur schwer möglich, die Parameter hinreichend genau zu schätzen.
Es gibt verschiedene Kriterien für die Wahl von p. Da wäre das **Akaikes Informations Kriterium** (AIC) ist ein allgemeines Kriterium, um ein Modell für einen kausalen AR(p)-Prozess auszuwählen. Es gibt auch eine Variante des AIC zur Schätzung der Ordnungen eines ARMA(p,q) Prozesses.

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

Das **Bayesians Informations Kriterium** (BIC) ist ein Kriterium, um die Ordnung eines Modells zu schätzen. Im Vergleich zum AIC tendiert das BIC jedoch nicht dazu, die Ordnung des Modells zu überschätzen. Für einen kausalen invertierbaren ARMA(p,q) ist das BIC wie folgt definiert:

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

Das letzte Kriterium nennt man **Hannan-Quinn Informationskriterium** (HQ) und ist definiert als:

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

3.2 Modellschätzung

In der Schätzphase werden die Modellparameter und -koeffizienten mit Hilfe unterschiedlicher Techniken geschätzt. Für das Trendmodell bieten sich die OLS-Methode und die Maximum-Likelihood-Methode für diese Schätzung an.
In den nächsten Abschnitten wird eine Modellschätzung an einem ARCH- und GARCH-Modell durchgeführt, wobei die Parameter mit Hilfe der Maximum-Likelihood-Methode geschätzt werden.

3.2.1 Maximum-Likelihood-Schätzer

Die zu schätzenden Parameter $\alpha_0, \alpha_1, ..., \alpha_n$ werden so gewählt, dass die Dichte an der Stelle der tatsächlich beobachteten Zeitreihe maximal wird, wobei eine Verteilung der Innovationen als bekannt angenommen wird. Oft ist dies die Standardnormalverteilung; setzt man die Dichte des Vektors $(X_1, ..., X_n)$ so an, als wären die $N(0,1)$-verteilt, obwohl man dies nicht wirklich als realistisch ansieht, so nennt man den resultierenden Schätzer einen **Maximum-Likelihood-Schätzer**.

Wir nehmen an, wir hätten $n+1$ Werte einer Zeitreihe gemessen: $X_0, X_1, ..., X_n$. Aus der Wahrscheinlichkeitstheorie wissen wir, dass wir die gemeinsame Dichte der beobachteten Werte anschreiben können als:

$$f_{X_0,...,X_n}(x_0, ..., x_n) = f_{X_0}(x_0) \prod_{t=1}^{n} f_{X_t|X_{t-1},...,X_0}(x_t|x_{t-1}, ..., x_0). \tag{5}$$

Dabei gibt $g(z)$ die Dichte von $(Z_t)_{t\in\mathbb{Z}}$ an. Wir erinnern uns, dass wir Z_t als strikt White-Noise-Prozess mit Erwartungswert $E(Z_t) = 0$ und $Z_t = 1$ definiert haben. Es ist üblich die Dichte des strikten White-Noise-Prozesses mit der Standardnormalverteilung oder einer t-Verteilung zu schätzen.

3.2.2 Bestimmung der Parameter eines ARCH(1)-Modells

Das ARCH(p)-Modelle wird meist durch einen modifizierten Maximum-Likelihood-Schätzer an beobachtete Renditezeitreihen angepasst.

Aus der Definition eines ARCH(p)-Prozesses lassen sich die Gleichungen (1),(2) aufstellen, wobei $Z_t = \frac{X_t}{\sigma_t}$ N(0,1) - verteilt ist.
Nun bedingt man die Berechnung auf $X_1, ..., X_p$ (da die Formel für die Dichte von X_t nicht bekannt ist). Zur Berechnung eines ML-Schätzers benötigt man:

$$f_{X_0,...,X_n}(x_0,...,x_n) = f_{X_0}(x_0)\prod_{t=1}^{n} f_{X_t|X_{t-1},...,X_0}(x_t|x_{t-1},...,x_0) \tag{6}$$

Wir schätzen die Parameter an einem ARCH(1)-Modell ab, die zu schätzenden Parameter sind α_0, α_1. Bei einem ARCH(1)-Prozess hängt X_t nur von σ_t und X_{t-1} ab. Daher können wir die bedingte Dichte aus (5) kürzer anschreiben mit

$$f_{X_t|X_{t-1},...,X_0}(x_t|x_{t-1},...,x_0) = f_{X_t|X_{t-1}}(x_t|x_{t-1}) = \frac{1}{\sigma_t}g(\frac{x_t}{\sigma_t}) \tag{7}$$

Wir können daraus die Likelihoodfunktion zur Plausibilitätsbestimmung der aus der Stichprobe zu schätzenden Parameter α_0 und α_1 bestimmen mit

$$L(\alpha_0, \alpha_1; \mathbf{X}) = f_{X_1,...,X_n|X_0}(x_1,...,x_n|x_0) = \prod_{t=1}^{n}\frac{1}{\sigma_t}g(\frac{X_t}{\sigma_t}), \tag{8}$$

wobei aus der Definition des ARCH-Prozesses für einen ARCH(1)-Prozess $\sigma_t = \sqrt{\alpha_0 + \alpha_1 x_{t-1}^2}$ folgt.

Wollen wir nun Daten statt an einen ARCH(1)-Prozess an einen ARCH(p)-Prozess anpassen, so müssen wir die Likelihoodfunktion nur auf p Anfangswerte $X_{-p+1}, ..., X_0$ umschreiben, um so die Werte $\alpha_0, \alpha_1, ...\alpha_p$ zu schätzen.

3.2.3 Schätzung der Parameter eines GARCH(1,1)-Prozesses

Für ein allgemeines GARCH(p,q)-Modell lassen sich die folgende Gleichungen aufstellen:

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

Wir setzen nun für p,q = 1 und erhalten ein GARCH(1,1)-Modell und die zu schätzenden Parameter sind $\alpha_0, \alpha_1, \beta$

Im Gegensatz zum ARCH(1)-Prozess hängt beim GARCH(1,1)-Prozess σ_t von σ_{t-1} ab, wodurch wir diesen Einfluss ebenso in der Funktion beachten müssen. Der gesamte Prozess ist in diesem Fall abhängig von den Anfangswerten X_0 und σ_0. Daher müssen wir für den GARCH(1,1)-Prozess eine eigene bedingte Dichte angeben:

$$f_{X_1,...,X_n|X_0,\sigma_0}(x_1,...,x_n|x_0,\sigma_0) = \prod_{t=1}^{n} f_{X_t|X_{t-1},...,X_0,\sigma_0}(x_t|x_{t-1},...,x_0,\sigma_0) \quad (9)$$

Die bedingte Dichte $f_{X_t|X_{t-1},...,X_0,\sigma_0}$ hängt von der Vergangenheit nur durch σ_t ab. Die Varianzen für die Zeitwerte lassen sich rekursiv aus den vorangegangen Werten $X_{t-1},...,X_0$ und σ_0 berechnen, durch $\sigma_t^2 = \alpha_0 + \alpha_1 X_{t-1}^2 + \beta\sigma_{t-1}^2$. Dadurch ergibt sich die bedingte Likelihoodfunktion als

$$L(\alpha_0,\alpha_1,\beta;\mathbf{X}) = \prod_{t=1}^{n} \frac{1}{\sigma_t} g(\frac{X_t}{\sigma_t}) \text{ mit } \sigma_t = \sqrt{\alpha_0 + \alpha_1 X_{t-1}^2 + \beta\sigma_{t-1}^2}. \quad (10)$$

Ein Problempunkt ist, dass wir den Wert σ_0 nicht beobachten konnten. Wir können diesen nur schätzen oder mit null annehmen. Mit fortschreitendem t sollte der Startwert immer weniger Einfluss auf den Wert X_t und sollte mit der Zeit immer unwichtiger werden. Aus diesem Grund können wir auch für σ_0 den Werte null annehmen.

Für einen beliebigen GARCH(p,q)-Prozess benötigen wir Beobachtungen von $n+p$ Werten $X_{-p+1},...,X_0,X_1,...,X_n$. Wir entwickeln die Likelihood-Bedingung für die beobachten Werte $X_{-p+1},...,X_0$ genauso wie für die nicht beobachten Werte $\sigma_{-q+1},...,\sigma_0$ für welche wir wiederum Starwerte annehmen oder festlegen müssen.

3.3 Modelldiagnose

In der Diagnosephase wird das Modell (oder ggf. mehrere ausgewählte Modelle) nochmal anhand ihrer geschätzten Parameter beurteilt. Dabei bietet sich folgende Vorgehensweise an:

1. Prüfen, ob die geschätzten Koeffizienten signifikant von Null verschieden sind. Bei einzelnen Koeffizienten erfolgt dies mit Hilfe eines t-Tests und bei mehrere Koeffizienten werden die mit einem F-Test untersucht.

2. Schließlich erfolgt eine sorgfältige Analyse der Residuen. Die Residuen sollten keine Struktur mehr aufweisen. Dabei kann man die Zentriertheit der Residuen mit einem t-Test kontrollieren.

3.4 Einsatzphase - Prognose erstellen

Wurde das Modell für passend befunden, kann es zur Erstellung von ex-ante Prognosen, also in die Zukunft gerichtete Prognosen, verwendet werden. Beim Eintreffen von aktuellen Beobachtungen können Prognosefehler ermittelt werden. Mit Hilfe dieser kann analysiert werden, ob systematische Abweichungen vorhanden sind. Wenn dies zutrifft, kann es angebracht sein, ein neues Modell auszuwählen und die Modellierungsphasen erneut zu durchlaufen.

4 Multivariate Zeitreihen

Viele Zeitreihen kann man auch in Zusammenhang zu einander betrachten. Von der aktuellen Finanzkrise beispielsweise sind nahezu alle Aktienkurse betroffen. Bei multivariaten Zeitreihen versucht man dieses Wissen mathematisch umzusetzen. Man betrachtet mehrere Zeitreihen gemeinsam und versucht aus den Zusammenhängen zwischen ihnen zusätzliche Informationen für die Vorhersage zu beziehen. Im Gegensatz zu den vorherigen Berechnungen haben wir in diesem Fall nicht nur einen Wert pro Zeiteinheit, sondern mehrere. Es ist vergleichbar mit einer Familie von Zufallsvariablen auf einem gemeinsamen Wahrscheinlichkeitsraum. Wir können uns nun nicht mehr auf die Korrelation zwischen den einzelnen Zeiten beschränken, sondern müssen uns auch mit der Korrelation zwischen den verschiedenen Zeitreihen beschäftigen.

Modelle für multivariaten Zeitreihen für mehrfache Risikofaktoren sind stochastische Prozesse, die ja als eine Familie von reellen Zufallsvektoren auf gemeinsamen Wahrscheinlichkeitsraum $(\Omega, \mathcal{F}, \mathbb{P})$ definiert sind.

4.1 Stationäre multivariate Zeitreihen

Die Begriffe der Stationärität müssen wir neu definieren:

Definition:
Die multivariate Zeitreihe $(\mathbf{X}_t)_{t\in\mathbb{Z}}$ ist **schwach stationär**, wenn die ersten zwei Momente existieren, die Erwartungswerte konstant sind und die Kovarianzfunktion nur vom zeitlichen Abstand abhängt:

$$\mathbb{E}X_t = \mu \ t \in \mathbb{Z},$$
$$\Gamma(t,s) = \Gamma(t+k, s+k) \ t,s,k \in \mathbb{Z}.$$

Definition:
Eine multivariate Zeitreihe $(\mathbf{X}_t)_{t\in\ \mathbb{Z}}$ nennt man **strikt stationär**, wenn jedes endliche Teilsystem in der Verteilung mit einem um s Zeitpunkte verschobenen System ident 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 strikt stationäre multivariate Zeitreihe mit endlicher Kovarianzmatrix ist auch schwach stationär. Die Umkehrung ist jedoch nicht sicher gegeben. Aus der Definition der schwachen Stationärität folgt ebenfalls, dass $\Gamma(t-s,0) = \Gamma(t,s)\forall t,s$. Die Kovarianz zwischen $\mathbf{X}_t$ und $\mathbf{X}_s$ hängt also nur von dem Abstand zwischen t und s ab. Daher können wir wiederum $\Gamma(h) := \Gamma(h,0)$ definieren. $\Gamma(0)$ entspricht dabei der Varianz-Kovarianz-Matrix.

Ähnlich wie bei der Autokorrelationsfunktion für die normale Zeitreihe können wir nun die Korrelationsmatrix-Funktion als $P(h)$ definieren. Wobei die Elemente $p_{i,i}(h)$ die ACF der i-ten Zeitreihe mit Lag h festlegen. Die Werte außerhalb der Diagonale geben Beziehungen zwischen den einzelnen Zeitreihen zu einem bestimmten Lag h an.

Betrachten wir nun einen multivariaten White-Noise-Prozess als Grundlage für weitere Prozesse.

Definition:
Ein schwach stationärer Prozess $(X_t)_{t\in\mathbb{Z}}$ ist ein **multivariater White-Noise-Prozess**, wenn seine Korrelationsmatrix-Funktion wie folgt aufgebaut ist:

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

Es darf also, wie schon im univariaten Fall, ein linearer Zusammenhang nur bei den aktuellen Werten bestehen. Bei einem multivariaten White-Noise-Prozess dürfen keine vergangen Daten, egal von welcher Zeitreihe, linear in die aktuelle Gegenwart einfließen.

4.2 VARMA(p,q)-Prozess

Definition:
Ein Vektor von Serien von identisch, unabhängig verteilten Zufallsvariablen mit endlicher Kovarianz-Matrix ist ein **strikter multivariater White-Noise-Prozess**.

Definition:
Sei $(\epsilon_t)_{t\in\mathbb{Z}}$ ein multivariater White-Noise-Prozess $WN(\mathbf{0}, \Sigma_\epsilon)$.
Die schwach stationäre Lösung $(\mathbf{X}_t)_{t\in\mathbb{Z}}$ der Differenzengleichung

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

wobei die Parameter Φ_i und Θ_j Matrizen der Form $\mathbb{R}^{dxd}$ sind, ist ein **VARMA(p,q)-Prozess** mit Erwartungswert null.

Ein VARMA(p,q)-Prozess mit Erwartungswert μ ergibt sich, wenn der Prozess $(\mathbf{X}_t - \mu)_{t\in\mathbb{Z}}$ die obige Differenzengleichung erfüllt.

4.3 multivariater GARCH-Prozess

Definition:
Der Prozess $(\mathbf{X}_t)_{t\in\mathbb{Z}}$ ist ein **multivariater GARCH-Prozess**, wenn er strikt stationär und die Gleichung:

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

wobei $\Sigma_t^{\frac{1}{2}} \in \mathbb{R}^{dxd}$ der Cholesky-Faktor der positiv-definiten Matrix Σ_t ist, erfüllt.
Auch die Eigenschaft der Martingal-Differenz können wir in diesem Zusammenspiel von mehreren Zeitreihen neu definieren.

Definition:
Erfüllt eine multivariate Zeitreihe $(X_t)_{t\in\mathbb{Z}}$ für $\forall t \in \mathbb{Z}$ und eine beliebige Filtrierung $(\mathcal{F}_t)$:

$$\begin{aligned}\mathbb{E}|X_t| &< \infty \\ \mathbb{E}(X_t|\mathcal{F}_{t-1}) &= 0\end{aligned}$$

so sagt man, die Zeitreihe erfüllt die **multivariate Martingal-Differenz**.

Es folgt daraus, dass der unbedingte Erwartungswert ebenfalls null ist und die Kovarianz-Matrix-Funktion $\Gamma(t,s) = 0$ für $t \neq s$ erfüllt.

Literaturverzeichnis

- Quantitative Risk Management: Concepts, Techniques, and Tools *von Alexander J. McNeil, Rüdiger Frey, & Paul Embrechts, Kapitel 4, Financial Time Series*

- Einführung in die stochastischen Prozesse und Zeitreihenanalyse *von Prof. Scherrer, SS09*

- http://www.math.uni-hamburg.de/home/drees/VM3_05_06/Kapitel5.pdf