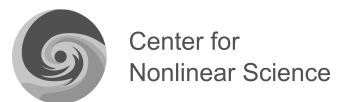


Empirische Bestimmung stochastischer Differentialgleichungssysteme für hochfrequente Aktienkursdaten



vorgelegt als Diplomarbeit von
Jan Schuppert
am
09. Dezember 2009



Inhaltsverzeichnis

Einleitung	1
1 Theorie	3
1.1 Grundlagen der Wahrscheinlichkeitstheorie	3
1.1.1 Bedingte Wahrscheinlichkeit	4
1.1.2 Momente und Kumulanten	4
1.1.3 Zentraler Grenzwertsatz und Normalverteilung	5
1.1.4 Korrelation	6
1.2 Einführung in die Theorie stochastischer Prozesse	6
1.2.1 Markov-Prozess	7
1.2.2 Chapman-Kolmogorov-Gleichung	7
1.2.3 Langevin-Gleichungen	7
1.2.4 Fokker-Planck-Gleichung	9
2 Die statistische Beschreibung von Finanzdaten	13
2.1 Eine Einführung	13
2.2 Der Datensatz	15
2.3 Stilisierte Fakten	16
2.3.1 Die Wahrscheinlichkeitsverteilung	16
2.3.2 Skalierungsverhalten	18
2.3.3 Abhängigkeiten	19
3 Drift- und Diffusionschätzung von Finanzdaten	25
3.1 Ergebnisse für den DAX	27
3.1.1 Markovanalyse für die Kursdifferenzen des DAX	27
3.1.2 Drift- und Diffusionsschätzung	31
3.2 Ergebnisse für einen GARCH(1,1)-Prozess	33
3.2.1 Die Definition des Prozesses	33
3.2.2 Parametrisierung und statistische Eigenschaften	35
4 Die Dynamik zweier Aktienkurse	41
4.1 Das Modell	43
4.2 Der Datensatz	45
4.2.1 Datenpräparation	45
4.3 Markovanalyse der interpolierten Datensätze	52

4.4	Schätzen der Drift- und Diffusionskomponenten	54
4.4.1	Der Driftvektor	54
4.4.2	Die Diffusionsmatrix	57
4.4.3	Die Koeffizienten vierter Ordnung	61
4.4.4	Der Prozess zur Beschreibung der Renditedynamik . .	61
5	Anwendung der Ergebnisse	65
5.1	Moderne Portfoliotheorie	65
5.1.1	Bedingte optimale Portfolios	67
6	Zusammenfassung und Ausblick	73
A	Anhang	75
A.1	Koordinatentransformation	75
A.2	Bestimmung der Parameter eines kubischen Polynoms	76
A.3	Aktienkursmodellierung	77
A.3.1	Geometrische Brownsche Bewegung	77
A.3.2	Jump-Diffusion-Modell	77
A.3.3	Stochastische Volatilität	78

Einleitung

Komplexe, dynamische Systeme spielen in der Physik und in vielen anderen Wissenschaften eine wichtige Rolle. Allgemein bestehen komplexe Systeme aus einer sehr großen Anzahl von Elementen. Die mikroskopischen Zustände der Elemente bestimmen den makroskopischen Zustand des Gesamtsystems. Dynamiken auf mikroskopischer Skala unterliegen gewissen Fluktuationen. Diese werden daher häufig mittels stochastischer Differentialgleichungen beschrieben.

Die Weltwirtschaft lässt sich ebenfalls als ein komplexes System im gerade beschriebenen Kontext auffassen. Dieses Obersystem setzt sich aus vielen gekoppelten Untersystemen zusammen, welche man auch als Märkte bezeichnen kann. Ein Markt ist durch das Aufeinandertreffen von Angebot und Nachfrage gekennzeichnet. Beispiele sind der globale Weltmarkt für Waren und Dienstleistungen, der europäische Geldmarkt oder ein Kapitalmarkt, an dem deutsche Aktien gehandelt werden. Die zeitliche Wertentwicklung des deutschen Aktienmarktes oder der Aktien einzelnen Unternehmen lassen sich durch die Bewegungen der Aktienkurse beschreiben. Die Dynamik der Kurse lässt sich nicht deterministisch mit den einzelnen Transaktionen, die von den vielen Teilnehmern auf der Angebots- und Nachfrageseite des Marktes getätigt werden, in Verbindung bringen. Vielmehr weisen diese Entwicklungen der Preise einen stochastischen Charakter auf, weshalb Kursentwicklungen mit Hilfe von stochastischen Prozessen beschrieben werden [40].

Eine Begründung, sich auf eine spezielle Klasse von gedächtnislosen Prozessen, den Markov-Prozessen, zu beschränken, liefert die *Effizienzmarkthypothese*. Diese besagt, dass bereits alle marktrelevanten Informationen in einem Kurs enthalten sind. Damit besteht kein Vorteil darin Kurshistorien zu berücksichtigen [40]. Markov-Prozesse lassen sich vollständig durch einen deterministischen Teil (Drift) und einen stochastischen Teil (Diffusion) beschreiben. Dass sich diese Komponenten mittels eines einfachen Verfahrens aus experimentellen Daten schätzen lassen, wurde vor einigen Jahren von Friedrich et al. gezeigt [16]. In Bezug auf Finanzdaten wurden bereits einzelne Prozesse analysiert und so Kenntnisse über die Drift- und Diffusionsstruktur gewonnen [15]. Gegenstand dieser Arbeit wird es sein, zusätzlich zu dem Verhalten isolierter Kursentwicklungen auch Interpendenzen zwischen zwei Wertpapieren zu betrachten.

Die Dynamiken des Wirtschaftssystems und seiner Untersysteme besser zu verstehen ist von zentraler Bedeutung. Jüngste Beispiele, wie die Finanzkrise von 2007/08 haben gezeigt wie wichtig es ist Kenntnisse über die komplexen Abhängigkeitsstrukturen der Weltwirtschaft zu erlangen, um mit geeigneten Mitteln Krisen dieser Art entgegenwirken zu können. Als Auslöser solcher Krisen werden häufig Märkte im Ungleichgewicht genannt. Im Falle der Finanzkrise spricht man von einer Immobilienblase als Ursache für die nachfolgenden negativen Entwicklungen. Auch die Möglichkeit der Früherkennung solcher Spekulationsblasen und Marktzusammenbrüche könnte einen wertvollen Beitrag zur Abschwächung von Wirtschaftskrisen leisten. Für kleinere Wirtschaftseinheiten wie Banken oder Unternehmen ist es ebenfalls wichtig die Methoden des Risikomanagements ständig zu verbessern. Gegenstand dieser Arbeit wird es sein die statistischen Charakteristika eines komplexen System zu untersuchen und dabei auf mögliche universelle Eigenschaften zu anderen dynamischen Systemen (z. B. Turbulenz) aufmerksam zu machen. Auf der anderen Seite handelt es sich bei den Finanzmärkten um ein praxisrelevantes System, und ein besseres Verständnis der Marktdynamiken ist sowohl für wirtschaftspolitische, als auch für unternehmerische Entscheidungen relevant.

In Kapitel 1 wird es eine kurze Einführung in die Theorie der stochastischen Prozesse geben. Zentral sind dabei die Zusammenhänge der mehrdimensionalen Fokker-Planck-Gleichung und stochastischen Differentialgleichungssystemen in der Notation nach Langevin. Das Kapitel 2 wird dem Leser eine geschichtliche Einordnung der wichtigsten Forschungsergebnisse der Kapitalmarkttheorie ermöglichen. Des Weiteren werden prominente statistische Eigenschaften von Finanzmarktdaten vorgestellt und an einem empirischen Datensatz geprüft. Zur erweiterten Charakterisierung kommt es in Kapitel 3 zu einer Drift- und Diffusionsanalyse der vorgestellten Daten. Das Herzstück dieser Arbeit bildet die zweidimensionale Drift- und Diffusionsschätzung der Preisdifferenzen zweier deutscher Aktien wie sie in Kapitel 4 durchgeführt wird. Anhand dieser Ergebnisse sollen Rückschlüsse auf mögliche Abhängigkeitsstrukturen im deterministischen oder stochastischen Teil der Dynamik aufgezeigt werden. Mögliche Anwendungsfelder, die sich aus den Resultaten ergeben, werden in Kapitel 5 vorgeschlagen. Die Beobachtungen dieser Arbeit werden anschließend in Kapitel 6 zusammenfassend wiedergegeben.

Kapitel 1

Theorie

Empirische Untersuchungen von Finanzzeitreihen lassen auf einen stochastischen Charakter von Kursdynamiken schließen. Statistik und Stochastik liefern eine Fülle von Methoden und Modellen, um Zeitreihen dieser Art zu analysieren und zu modellieren. In diesem Kapitel werden die wichtigsten Formeln und Gleichungen vorgestellt, die in dieser Arbeit eine Anwendung finden. Auf ausführliche Herleitungen wurde dabei an dieser Stelle verzichtet. Des Weiteren werden Kenntnisse aus der Wahrscheinlichkeitstheorie vorausgesetzt. Umfassende Beschreibungen der hier verwendeten Theorie sind der Standardliteratur von z. B. Haken [22], Gardiner [18] oder Risken [43] zu entnehmen.

1.1 Grundlagen der Wahrscheinlichkeitstheorie

Sei X_t eine Zufallsvariable mit Zeitparameter t . Diese kann sowohl diskrete (z. B. Farben, Münzwurf) als auch kontinuierliche Werte (z. B. Gewicht, Preis) annehmen. Die Ein-Punkt-Wahrscheinlichkeitsdichtefunktion (pdf; Abk. für engl. probability density function) ist definiert durch:

$$p(x, t) = \int_{\Omega} d\mu(\omega) \delta(x - X_t(\omega)), \quad (1.1)$$

wobei $\omega \in \Omega$, μ das Maß und Ω den Ergebnisraum darstellt. Die Dichtefunktion ist auf Eins normiert:

$$\int_{-\infty}^{\infty} dx p(x, t) = 1. \quad (1.2)$$

Für das gemeinsame Auftreten von N Ereignissen definiert man eine N -Punkt Verteilungsfunktion (*Verbundwahrscheinlichkeit*, jpdf; Abk. für engl. joint pro-

bability density function):

$$f_N(x_1, t_1; \dots; x_N, t_N) = \int_{\Omega} d\mu(\omega) \delta(x_1 - X_{t_1}(\omega)) \cdots \delta(x_N - X_{t_N}(\omega)), \quad (1.3)$$

mit $N \geq 2$.

1.1.1 Bedingte Wahrscheinlichkeit

Die einfach bedingte Wahrscheinlichkeitsdichte ist definiert als der Quotient aus Verbundwahrscheinlichkeitsdichte und der pdf:

$$p(x, t|x', t') = \frac{f_2(x, t; x', t')}{p(x', t')}, \quad (1.4)$$

wobei $p(x', t') \neq 0$ angenommen wird. Der Zusammenhang lässt sich erweitern zur allgemeinen bedingten Wahrscheinlichkeitsdichtefunktion (cpdf; Abk. für engl. conditional probability density function):

$$p(x_N, t_N|x_{N-1}, t_{N-1}; x_{N-2}, t_{N-2}; \dots; x_1, t_1) = \frac{f_N(x_1, t_1; x_2, t_2; x_3, t_3; \dots; x_N, t_N)}{f_{N-1}(x_1, t_1; x_2, t_2; \dots; x_{N-1}, t_{N-1})}. \quad (1.5)$$

1.1.2 Momente und Kumulanten

Momente n -ter Ordnung sind im Kontinuierlichen gegeben durch:

$$m_n = \langle x^n(t) \rangle = \int dx x^n p(x, t). \quad (1.6)$$

Voraussetzung für die Definition (1.6) ist die Konvergenz des Integrals. Diese ist gegeben, wenn die pdf $p(x, t)$ für große $|x|$ verschwindet. Die in dieser Arbeit auftretenden Wahrscheinlichkeitsdichtefunktionen haben alle die entsprechende Eigenschaft. Das erste Moment $\langle x \rangle$ bezeichnet man als *Mittelwert*. Für diskrete Zeitreihen lässt sich ein statistischer Schätzer definieren:

$$\tilde{m}_1 = \langle x \rangle = \mu = \frac{1}{N} \sum_{i=1}^N x_i. \quad (1.7)$$

Die Differenz aus dem zweiten und dem quadrierten ersten Moment $\sigma_2 = m_2 - m_1^2$ wird als *Varianz* bezeichnet. Ein Schätzer für diskrete Werte folgt mit:

$$\tilde{m}_2 = \tilde{\sigma}^2 = \frac{1}{N-1} \sum_{i=1}^N (x_i - \mu)^2. \quad (1.8)$$

Die Fouriertransformierte der pdf nennt man *charakteristische Funktion* von $p(x, t)$:

$$\hat{p}(k, t) = \int_{-\infty}^{\infty} dx e^{ikx} p(x, t), \quad (1.9)$$

welche man über die Reihendefinition der Exponentialfunktion und (1.6) auch wie folgt schreiben kann:

$$\hat{p}(k, t) = \sum_n \frac{m_n}{n!} (ik)^n.$$

Die Kumulanten einer pdf sind Kombinationen der Momente. Sie sind definiert als die Ableitung der logarithmierten charakteristischen Funktion:

$$c_n = (-i)^n \frac{d^n}{dk^n} \ln \hat{p}(k, t) \Big|_{k=0}.$$

Die dritte und vierte Kumulante errechnen sich somit wie folgt:

$$\begin{aligned} c_3 &= m_3 - 3m_2m_1 + 2m_1^3 \\ c_4 &= m_4 - 4m_3m_1 - 3m_2^2 + 12m_2m_1^2 - 6m_1^4. \end{aligned}$$

Für symmetrische Verteilungen verschwinden die Terme mit ungeraden Momenten. Die normalisierte dritte Kumulante bezeichnet man als *Schiefe* ϵ , die vierte Kumulante als *Kurtosis* κ :

$$\epsilon = \frac{c_3}{\sigma^3} \quad (1.10)$$

$$\kappa = \frac{c_4}{\sigma^4} - 3. \quad (1.11)$$

Die Schiefe beschreibt die Neigungsstärke der Verteilung, während die Kurtosis eine Aussage über die Spitze und die Ausläufer der Verteilung macht. Für die Standardnormalverteilung sind beide Größen Null.

1.1.3 Zentraler Grenzwertsatz und Normalverteilung

Der zentrale Grenzwertsatz beschreibt das Verhalten einer Summe von Zufallszahlen. Sei $(X_i)_{i \geq 1}$ eine Folge von unabhängigen, identisch verteilten Zufallsvariablen mit existierendem Mittelwert und endlicher Varianz größer Null. Die Verteilung dieser Summe strebt gegen die Standardnormalverteilung

$$\frac{1}{\sqrt{N}} \sum_{i=1}^N \frac{X_i - \mu}{\sqrt{\sigma^2}} \xrightarrow{N \rightarrow \infty} \Phi_{0,1}. \quad (1.12)$$

Die Normalverteilung oder Gauß-Verteilung nach ihrem Entdecker Carl Friedrich Gauß ist eine wichtige kontinuierliche Verteilung. Die *Gauß'sche Glockenkurve* hat folgende Form:

$$\Phi_{\mu,\sigma}(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}. \quad (1.13)$$

Die Kurzschreibweise $X \sim \mathcal{N}(\mu, \sigma)$ bedeutet im Folgenden, dass die Zufallsvariable X normalverteilt mit Mittelwert μ und Varianz σ ist.

1.1.4 Korrelation

Für zwei quadratintegrale Zufallsvariablen X_1, X_2 ist die *Kovarianz* wie folgt definiert:

$$\rho(X_1, X_2) = \sigma_{12} = \langle (X_1 - \langle X_1 \rangle)(X_2 - \langle X_2 \rangle) \rangle. \quad (1.14)$$

Mit Hilfe dieser Definition lässt sich nun der *Korrelationskoeffizient* definieren:

$$\gamma(X_1, X_2) = \frac{\rho(X_1, X_2)}{\sqrt{\sigma^2(X_1)}\sqrt{\sigma^2(X_2)}} = \frac{\sigma_{12}}{\sigma_1\sigma_2}. \quad (1.15)$$

Der Korrelationskoeffizient kann Werte zwischen $\gamma = -1$ für perfekt negativ korrelierte Zufallsvariablen und $\gamma = 1$ für perfekt positiv korrelierte Variablen annehmen. Bei $\gamma = 0$ liegt keinerlei Korrelation vor. Aus einer Unabhängigkeit von X_1 und X_2 folgt Unkorreliertheit, der Umkehrschluss ist i. A. nicht zulässig.

1.2 Einführung in die Theorie stochastischer Prozesse

Ein stochastischer Prozess ist eine Abfolge von Zufallsvariablen $(X_t)_{t \in I}$. Wenn $I = \mathbb{N}$ ist, spricht man von einem diskreten, stochastischen Prozess, für $I = \mathbb{R}$ von einem kontinuierlichen, stochastischen Prozess. Eine N -Punkt-Verteilung gem. Definition (1.3) beschreibt den stochastischen Prozess vollständig. Ein Prozess gilt als *stationär* für alle N , falls gilt:

$$f_N(x_1, t_1; x_2, t_2; \dots; x_N, t_N) = f_N(x_1, t_1 + T; x_2, t_2 + T; \dots; x_N, t_N + T) \quad (1.16)$$

mit $T \in I$.

1.2.1 Markov-Prozess

Eine sehr wichtige Gruppe von stochastischen Prozessen stellen die sogenannten Markov-Prozesse dar. Allgemein spricht man dabei von Prozessen ohne Gedächtnis, das heißt, dass der zukünftige Zustand eines Systems nur vom aktuellen Zustand abhängt und nicht etwa von der gesamten Historie aller Zustände. Ein Markov-Prozess erfüllt somit die folgende Gleichung:

$$p(x_N, t_N | x_{N-1}, t_{N-1}) = p(x_N, t_N | x_{N-1}, t_{N-1}; \dots; x_2, t_2; x_1, t_1), \quad (1.17)$$

wobei $t_1 < t_2 < \dots < t_N$ o.B.d.A. Eine wichtige Konsequenz aus dieser Eigenschaft ist, dass sich nun der gesamte Markov-Prozess mit Hilfe eines Produkts der Übergangswahrscheinlichkeiten und der Ein-Punkt-Wahrscheinlichkeit beschreiben lässt:

$$\begin{aligned} p(x_N, t_N; \dots; x_2, t_2; x_1, t_1) = \\ p(x_N, t_N | x_{N-1}, t_{N-1}) p(x_{N-1}, t_{N-1} | x_{N-2}, t_{N-2}) \cdots p(x_2, t_2 | x_1, t_1) p(x_1, t_1). \end{aligned} \quad (1.18)$$

1.2.2 Chapman-Kolmogorov-Gleichung

Beschränkt man Gleichung (1.18) mit $N = 3$ so erhält man folgenden Zusammenhang:

$$f(x_3, t_3; x_2, t_2; x_1, t_1) = p(x_3, t_3 | x_2, t_2) p(x_2, t_2 | x_1, t_1) p(x_1, t_1). \quad (1.19)$$

Integriert man nun über x_2 und dividiert dann mit $p(x_1, t_1)$ erhält man eine Gleichung die als Chapman-Kolmogorov-Gleichung bekannt ist:

$$p(x_3, t_3 | x_1, t_1) = \int dx_2 p(x_3, t_3 | x_2, t_2) p(x_2, t_2 | x_1, t_1). \quad (1.20)$$

Sie bildet den Ausgangspunkt für weitere Beschreibungen der zeitlichen Entwicklung von Wahrscheinlichkeitsdichteverteilungen.

1.2.3 Langevin-Gleichungen

Mit stochastischen Differentialgleichungen (SDE), welche in der Physik häufig als Langevin-Gleichungen bezeichnet werden, ist es möglich, eine ganze Reihe von Diffusionsprozessen zu modellieren, mit denen verschiedenste dynamische Systeme beschrieben werden können. Ursprünglich wurde die Gleichung von

Paul Langevin 1908 entwickelt, um die Dynamik von Brownschen Partikeln in einem System zu untersuchen, auf das eine äußere Kraft wirkt. Eine allgemeine Form für die mehrdimensionale Langevin-Gleichung ist gegeben durch:

$$\frac{d}{dt}X_i(t) = h_i(\mathbf{X}, t) + \sum_j g_{ij}(\mathbf{X}, t)\Gamma_j(t) \quad (1.21)$$

mit $\mathbf{X} \in \mathbb{R}^N$. Der deterministische Teil $h_i(\mathbf{X}, t)$ wird auch als Driftterm bezeichnet. Der stochastische Teil der Gleichung besteht aus den Diffusionskoeffizienten g_{ij} und den Fluktuationstermen $\Gamma_j(t)$. $\Gamma(t)$ wird auch als δ -korreliertes, weißes Rauschen bezeichnet. Es hat folgende Eigenschaften:

$$\langle \Gamma_i(t) \rangle = 0, \quad (1.22)$$

$$\langle \Gamma_i(t)\Gamma_j(t') \rangle = 2\delta_{ij}\delta(t-t'). \quad (1.23)$$

Für eine Dimension läßt sich folgende allgemeine Langevin-Gleichung angeben:

$$\frac{d}{dt}X(t) = h(X, t) + g(X, t)\Gamma(t). \quad (1.24)$$

Wiener-Prozess

Eine Langevin-Gleichung, in der ausschließlich eine fluktuierende Kraft auftritt, wird auch *Wiener-Prozess* genannt:

$$\frac{d}{dt}W(t) = \Gamma(t). \quad (1.25)$$

Wegstücke auf einem Zufallspfad sind stochastisch unabhängig, stationär und normalverteilt. Betrachtet man ein kleines Wegstück des stochastischen Pfades so muss man über die Fluktuationen integrieren:

$$dW(t) = W(t+dt) - W(t) = \int_t^{t+dt} dt' \Gamma(t'). \quad (1.26)$$

Das letztgenannte Integral bezeichnet man auch als *Stieltjes-Integral*. Im Folgenden soll kurz auf die Behandlung solcher stochastischen Integrale eingegangen werden.

Stochastische Integralgleichungen

Es gibt unterschiedliche Interpretationen von stochastischer Integration. Betrachtet man stochastische Integrale als Grenzwert einer Summe und ist, wie

in Gleichung (1.26) der Integrant unstetig, macht es einen Unterschied, ob man mittige (Interpretation nach Stratonovich) oder linksseitige Stützstellen (Interpretation nach Itô) bei der Integration verwendet. Für detaillierte Ausführungen sei auf die Literatur, z. B. [18] verwiesen. In dieser Arbeit beschränkt man sich ausschließlich auf Itô's Definition. Die Langevin-Gleichung (1.24) lässt sich mit (1.26) als Integralgleichung der Form

$$x(t) = x(0) + \int_0^t dt' h(x, t') + \int_0^t dW(t') g(x, t') \quad (1.27)$$

schreiben. Eine verkürzte Schreibweise, die man in der Finanzmathematik (siehe Kapitel 2 und A.3) oft findet, lautet:

$$dx(t) = h(x, t)dt + g(x, t)dW(t). \quad (1.28)$$

1.2.4 Fokker-Planck-Gleichung

Eine fundamentale Gleichung, um die Dynamik von Markov-Prozessen zu beschreiben, ist die Fokker-Planck-Gleichung. Sie gibt die zeitliche Entwicklung von Wahrscheinlichkeitsdichtefunktionen wieder.

Die Wahrscheinlichkeit, ein System im Zustand x zum Zeitpunkt t und zu einem späteren Zeitpunkt $t + \tau$ im Zustand x' zu finden, ist in folgender Weise miteinander verknüpft:

$$p(x, t + \tau) = \int dx p(x, t + \tau | x', t) p(x', t). \quad (1.29)$$

Um nun einen Ausdruck für das Differential $\frac{\partial}{\partial t} p(x, t)$ zu erhalten, benötigt man die Übergangswahrscheinlichkeit $p(x, t + \tau | x', t)$. Diese soll im Folgenden kurz hergeleitet werden. Dabei wird eine Herleitung aus [43] verwendet, welche insbesondere die Eigenschaften der charakteristischen Funktion ausnutzt. Startpunkt ist die Definition der *bedingten Momente* M_n :

$$\begin{aligned} M_n(x', t, \tau) &= \langle (X(t + \tau) - X(t))^n | X(t) = x' \rangle \\ &= \int dx (x - x')^n p(x, t + \tau | x', t). \end{aligned} \quad (1.30)$$

Durch die Transformation in die charakteristische Funktion gemäß (1.9) und durch Ausnutzen der Reihendefinition der Exponentialfunktion ergibt sich fol-

gender Zusammenhang:

$$\begin{aligned}\hat{p}(k, x', t, \tau) &= \int_{-\infty}^{\infty} dx e^{ik(x-x')} p(x, t + \tau | x', t) \\ &= 1 + \sum_{n=1}^{\infty} \frac{(ik)^n}{n!} \underbrace{\int dx (x - x')^n p(x, t + \tau | x', t)}_{M_n(x', t, \tau)}.\end{aligned}$$

Dieser Ausdruck in die Rücktransformation eingesetzt ergibt folgendes:

$$p(x, t + \tau | x', t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} dk e^{-ik(x-x')} \left(1 + \sum_{n=1}^{\infty} \frac{(ik)^n}{n!} M_n(x', t, \tau) \right).$$

Mit Hilfe einer Fourier-Transformation läßt sich folgender Zusammenhang herstellen:

$$\frac{1}{2\pi} \int_{-\infty}^{\infty} dk (ik)^n e^{-ik(x-x')} = \left(-\frac{\partial}{\partial x} \right)^n \delta(x - x').$$

Unter Ausnutzung dieser Gleichung folgt:

$$p(x, t + \tau | x', t) = \left[1 + \sum_{n=1}^{\infty} \frac{1}{n!} \left(-\frac{\partial}{\partial x} \right)^n M_n(x, t, \tau) \right] \delta(x - x').$$

Berücksichtigung der Symmetrieeigenschaften der δ -Funktion

$$\delta(x - x') = \delta(x' - x)$$

und das Einsetzen in Gleichung (1.29) liefert die sogenannte *Kramers-Moyal-Entwicklung*:

$$\frac{\partial}{\partial t} p(x, t) = \sum_{n=1}^{\infty} \left(-\frac{\partial}{\partial x} \right)^n D^{(n)}(x, t) p(x, t), \quad (1.31)$$

mit den Kramers-Moyal-Koeffizienten:

$$D^{(n)}(x, t) = \lim_{\tau \rightarrow 0} \frac{M_n(x, t, \tau)}{n! \tau} = \lim_{\tau \rightarrow 0} \frac{\langle [X(t + \tau) - X(t)]^n | X(t) = x \rangle}{n! \tau}. \quad (1.32)$$

Pawulas-Theorem

Pawulas-Theorem macht eine Aussage darüber, wie viele Terme der Kramers-Moyal-Entwicklung berücksichtigt werden müssen. Falls der Prozess durch eine Langevin-Gleichung (1.24) mit δ -korreliertem weißen Rauschen beschrieben werden kann, verschwinden alle Koeffizienten mit $n \geq 3$. In diesem Fall ergibt

sich eine Gleichung die als *Fokker-Planck-Gleichung* oder *Vorwärts-Kolmogorov-Gleichung* bekannt ist:

$$\frac{\partial}{\partial t} p(x, t) = \left[-\frac{\partial}{\partial x} D^{(1)}(x, t) + \frac{1}{2} \frac{\partial^2}{\partial x^2} D^{(2)}(x, t) \right] p(x, t). \quad (1.33)$$

Der Koeffizient $D^{(1)}$ wird als Drift bezeichnet, $D^{(2)}$ als Diffusion. Sie ist die korrespondierende Fokker-Planck-Gleichung zur Langevin-Gleichung (1.24). Nach Itô's Definition stehen die Koeffizienten in folgender Beziehung zueinander:

$$D^{(1)}(x, t) = h(x, t), \quad (1.34)$$

$$D^{(2)}(x, t) = g^2(x, t). \quad (1.35)$$

Die mehrdimensionale Fokker-Planck Gleichung

In analoger Weise läßt sich die allgemeine Form der Gleichung für N Dimensionen angeben:

$$\frac{\partial}{\partial t} p(\mathbf{x}, t) = \left[-\sum_i \frac{\partial}{\partial x_i} D_i^{(1)}(\mathbf{x}, t) + \frac{1}{2} \sum_{ij} \frac{\partial^2}{\partial x_i \partial x_j} D_{ij}^{(2)}(\mathbf{x}, t) \right] p(\mathbf{x}, t) \quad (1.36)$$

mit $\mathbf{x} \in \mathbb{R}^N$. Dabei bezeichnet man $D_i^{(1)}$ nun als Driftvektor und $D_{ij}^{(2)}$ als Diffusionsmatrix. Sie sind definiert durch:

$$D_i^{(1)}(\mathbf{x}, t) = \lim_{\tau \rightarrow 0} \frac{1}{\tau} \langle X_i(t + \tau) - X_i(t) | \mathbf{X}(t) = \mathbf{x} \rangle, \quad (1.37)$$

$$D_{ij}^{(2)}(\mathbf{x}, t) = \lim_{\tau \rightarrow 0} \frac{1}{2\tau} \langle [X_i(t + \tau) - X_i(t)] [X_j(t + \tau) - X_j(t)] | \mathbf{X}(t) = \mathbf{x} \rangle. \quad (1.38)$$

Der Zusammenhang zu den Drift- und Diffusionstermen der mehrdimensionalen Langevin-Gleichung (1.21) ergibt sich wie folgt (Definition nach Itô):

$$D_i^{(1)} = h_i, \quad (1.39)$$

$$D_{ij}^{(2)} = g_{ij}^2, \quad (1.40)$$

für eine positiv definite Matrix $D^{(2)}$.

Kapitel 2

Die statistische Beschreibung von Finanzdaten

Das Teilgebiet der wirtschaftswissenschaftlichen Forschung, welches sich mit der mathematischen Beschreibung ökonomischer Zusammenhänge beschäftigt, wird als Ökonometrie bezeichnet. Seit den neunziger Jahren ergänzen Physiker mit Hilfe von Methoden und Modellen aus vorwiegend nichtlinearer oder statistischer Physik die Forschungsergebnisse. Dieses interdisziplinäre Forschungsfeld wird als Ökonophysik bezeichnet. Eine Einführung in die Schwerpunktthemen findet sich in den Standardwerken von Mantegna und Stanley [35] bzw. Paul und Baschnagel [40]. Viele der in dieser Arbeit angewendeten Verfahren und Analysen sind aus physikalischen Problemen motiviert worden.

Im Folgenden werden die wichtigsten Ergebnisse aus empirischen Untersuchungen vorgestellt, die in Bezug auf die Dynamik von Aktienkursen gemacht wurden. Diese *stilisierten Fakten* ¹ werden anschließend anhand eines DAX-Datensatzes überprüft, und die Resultate mit Ergebnissen aus der Literatur verglichen. Zunächst soll es jedoch eine geschichtliche Einordnung einiger wichtiger wissenschaftlicher Errungenschaften der Finanzmarkttheorie und Finanzmathematik geben. Diese ist jedoch keinesfalls als vollständig anzusehen.

2.1 Eine Einführung

Als Begründer der Finanzmathematik wird heute der französische Mathematiker Louis Bachelier gehandelt. Bereits im Jahre 1900 beschrieb er in seiner Dissertation „Théorie de la Spéculation“ [4] die Dynamik von Aktienkursen mit Hilfe der Brownschen Bewegung. Neue Impulse bekam die Disziplin durch die Erfindung von programmierbaren Rechenmaschinen. Die Möglichkeit größere

¹Ökonomische Zusammenhänge die durch viele Beobachtungen gestützt werden. (Gem. [48])

Datenmengen zu untersuchen ließ um 1960 einige neue Erkenntnisse über das statistische Verhalten von Finanzdaten aufkommen. So konnten neben anderen Mandelbrot [34] und Fama [14] festhalten, dass extreme Preisänderungen von Aktienkursen wesentlich wahrscheinlicher sind als von der Normalverteilung beschrieben. Im Jahr 1952 begründet Harry Markowitz die moderne Portfoliotheorie [36]. Dass Diversifikation in einem Wertpapierportfolio zu einer Risikominderung führen kann, wurde hier zum ersten Mal wissenschaftlich hergeleitet und bestätigt. 1964 wurde das Modell von Sharpe [45] erweitert. Im Kern des Preismodells für Kapitalgüter (Capital Asset Pricing Model, CAPM) steht das Ziel, Gleichgewichtspreise für Wertpapiere unter Risiko im Portfoliozusammenhang zu ermitteln.

Die bis heute wohl bekannteste Errungenschaft der Finanzmathematik ist das 1973 veröffentlichte Black-Scholes-Modell [5] zur Bewertung von Aktienoptionen. Der Verlauf (die Trajektorie) eines Aktienkurses S_t wird dabei durch eine geometrische Brownsche Bewegung (GBM) beschrieben, welche die Lösung der folgenden stochastischen Differentialgleichung ist:

$$dS_t = \mu S_t dt + \sigma S_t dW_t.$$

Dabei bezeichnet man μ als Drift und σ als Volatilität. W_t ist ein Wiener Prozess gemäß der Definition (1.25). Die Bekanntheit und Akzeptanz hat dieses Modell vor allem seiner Einfachheit zu verdanken. Restriktionen des Grundmodells sind zum Beispiel die als konstant angenommene Volatilität und der konstante Zins, sowie die normalverteilten Aktienkursrenditen (siehe Anhang A.7), welche nicht der Empirie entsprechen. In den Folgejahren wurde durch Modellerweiterungen versucht dieses Grundmodell entsprechend der empirischen Ergebnisse besser anzupassen. Eine wichtige Modellerweiterung lieferten Hull und White [25] und Heston [23] mit der Einführung der stochastischen Volatilität. Im Basismodell wird die konstante Volatilität durch einen weiteren stochastischen Prozess ersetzt (siehe Anhang A.3.3). Ein Jump-Diffusion-Modell (SJ) wurde 1976 von Merton [37] eingeführt. Dabei wird der Grundgleichung additiv ein Poisson-Prozess angehängt (siehe Anhang A.3.2). Von Seiten der Zeitreihenanalyse sind im Laufe der Jahre auch sehr viele Modelle und Prozesse konstruiert worden, um das Verhalten von Finanzdaten adäquat zu beschreiben. Hervorzuheben ist dabei die Klasse der 1986 veröffentlichten GARCH-Prozesse ² [6], welche zur Volatilitätsprognose verwendet werden.

Für die Untersuchung der statistischen Eigenschaften werden im Folgenden die logarithmierten Preisdifferenzen $x_{\Delta t}$ mit dem zeitlichen Abstand Δt betrach-

²Ist die Abkürzung für „Generalized AutoRegressive Conditionally Heteroscedastic“.

tet ³:

$$x_{\Delta t} = \log \left[\frac{s(t + \Delta t)}{s(t)} \right] = \log s(t + \Delta t) - \log s(t) = y(t + \Delta t) - y(t), \quad (2.1)$$

wobei $s(t)$ für den Kurs eines Wertpapiers zum Zeitpunkt t steht. Der Grund für diese Art der Behandlung liegt vor allem in der Forderung einen möglichst stationären Prozess zu erhalten. Des Weiteren können bei Aktien verschiedene Effekte auftreten, die zu einer Preisänderung der Aktien führen, aber nichts an den Beteiligungsverhältnissen ändern. Beispiele sind der Aktiensplit oder die Aktienzusammenlegung ⁴. Es wäre auch denkbar die Gewinne oder Verluste direkt als Renditen

$$r_t = \frac{s(t + \Delta t) - s(t)}{s(t)} \quad (2.2)$$

prozentual zu betrachten. Für hochfrequente Daten gilt jedoch $r_{\Delta t} \approx x_{\Delta t}$ (siehe z. B. [35], S. 38).

2.2 Der Datensatz

In der vorliegenden Arbeit werden insgesamt drei Datensätze untersucht [52]. Es handelt sich dabei zum einen um die XETRA-Kursnotierungen des Deutschen Aktienindex (DAX) und zum anderen um zwei Einzelwerte aus diesem Index mit tick-by-tick ⁵ Aufzeichnungen aus den Jahren 1993 bis 2003 ⁶. Details sind Tabelle 2.1 zu entnehmen. Die minimale Zeitskala liegt beim DAX bei 15

Tabelle 2.1: Details der verwendeten Finanzzeitreihen.

WKN	Titel	Datenpunkte	Zeitraum
846900	DAX	4.294.221	Jan 1993 – Dez 2003
575200	Bayer AG	2.936.719	Jan 1993 – Dez 2003
766400	Volkswagen AG	2.814.917	Jan 1993 – Dez 2003

Sekunden. Das ist der zeitliche Abstand, mit dem der Datensatz ursprünglich aufgezeichnet wurde. Die Auflösung bei den Einzeltiteln liegt bei 0.01 Sekunden. Es werden im Folgenden nur Preisdifferenzen betrachtet, die in der Handelszeit liegen, d.h., dass keine Übernacht-Sprünge berücksichtigt werden. Üblicherweise

³Je nach Anwendung werden im Folgenden noch weitere Bezeichnungen $x_{\Delta t} = X(t, \Delta t) = x_i(t)$ für die logarithmierten Preisdifferenzen verwendet.

⁴Kapitalmaßnahmen, die es einer Aktiengesellschaft erlaubt den Nennwert oder die Stückzahl ihre ausgegebenen Aktien zu ändern.

⁵Immer dann, wenn der Börsenmakler einen Handel ausführt, wird der Börsenkurs und der Zeitpunkt aufgezeichnet.

⁶XETRA wurde erst im November 1997 eingeführt, zuvor existierte das IBIS System.

werden die Handelszeiten aneinander „geklebt“, um so die Statistik auf größere Skalen ausweiten zu können. Man könnte jedoch auch argumentieren, dass Übernacht-Wertsprünge besondere Wertänderungen darstellen, die nicht mit den üblichen Sprüngen gleichgesetzt werden sollten.

Nawroth identifiziert in [39] unterschiedliche Zeitskalen in Aktienkursen. Für Zeiträume im Sekundenbereich spricht man von kleinen Skalen. Wenn im Folgenden von einem mittleren Skalenbereich die Rede ist, sind Minuten-Zeiträume gemeint. Der große Skalenbereich beinhaltet Tages-, Wochen- oder Monatsintervalle und alles, was darüber hinaus geht.

2.3 Stilisierte Fakten

Über Finanzzeitreihen im Allgemeinen wurden in den letzten Jahrzehnten viele verschiedene stilisierte Fakten zusammengetragen. Eine sehr gute Übersicht dazu erstellte Sewell in [44]. Für den hier vorgestellten Prozess, also Preisdifferenzen von Aktien auf verschiedenen Zeitskalen, sind ebenfalls bestimmte Merkmale beobachtet worden. Die Wichtigsten werden im Folgenden kurz vorgestellt und auf den vorliegenden DAX-Datensatz geprüft.

2.3.1 Die Wahrscheinlichkeitsverteilung

Betrachtet man die Verteilungen der Kursdifferenzen auf verschiedenen Skalen, so kann man folgende Beobachtungen machen:

- Die Verteilungen sind symmetrisch und es ist keine Schiefe zu beobachten.
- Die Verteilung ist *leptokurtisch*, das heißt, diese ist „spitzer“ als die Normalverteilung.
- Die Ausläufer der Verteilung flachen langsamer ab als bei der Normalverteilung. Sie werden deshalb auch als *heavy/fat tails* bezeichnet.
- Je größer die Zeitdifferenzen Δt werden, desto mehr nähert sich die beobachtete Verteilung der Normalverteilung an. Dies lässt sich gut an der Kurtosis erkennen, die in Abb. 2.2 aufgetragen ist.

Die Abbildungen 2.1 zeigen die ersten beiden Kumulanten der Verteilungen. Dass der Mittelwert verschwindet, kann später bei der Berechnung der Korrelationsfunktionen ausgenutzt werden. Bei der Varianz kann man einen nahezu

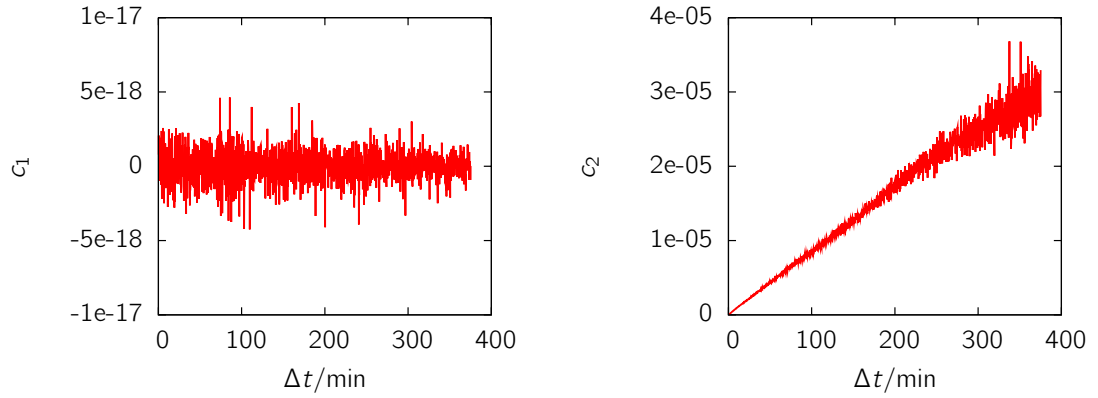


Abbildung 2.1: Mittelwert $c_1(\Delta t)$ und Varianz $c_2(\Delta t)$ der DAX-Kursdifferenzen auf unterschiedlichen Zeitskalen Δt in Minuten.

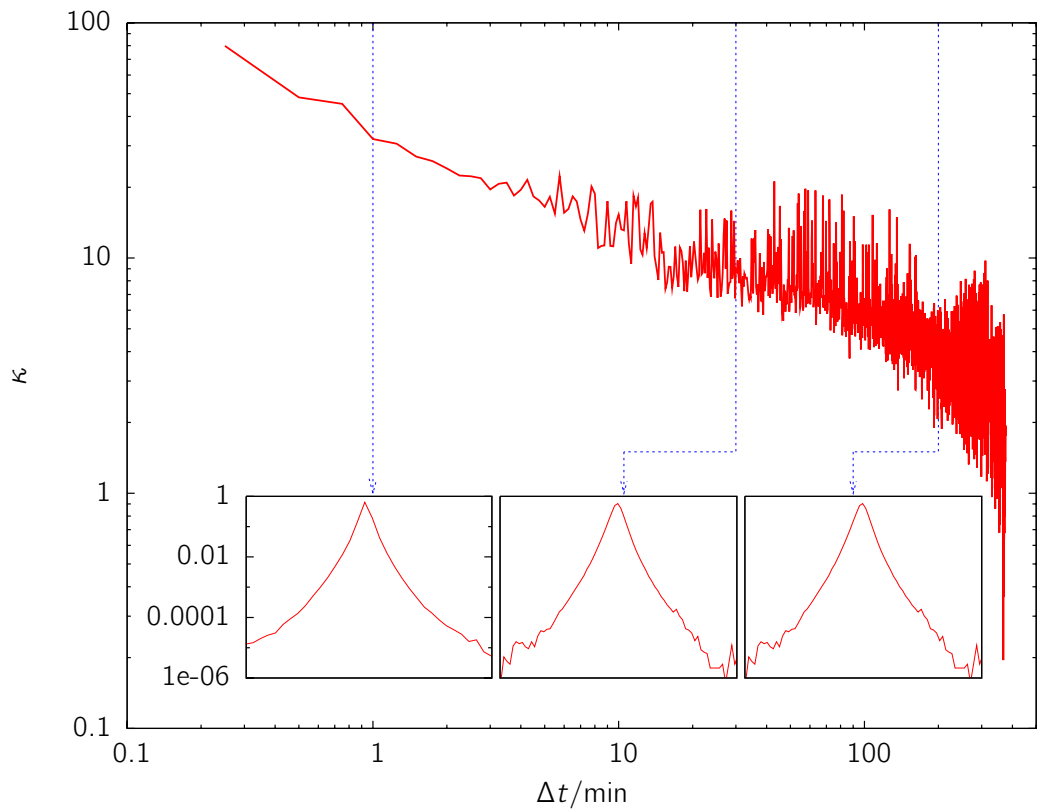


Abbildung 2.2: Kurtosis $\kappa(\Delta t)$ der Wahrscheinlichkeitsdichtefunktionen der DAX-Kursdifferenzen auf unterschiedlichen Zeitskalen Δt in Minuten. Für die Skalen $\Delta t_1 = 60\text{s}$, $\Delta t_2 = 30\text{ min}$ und $\Delta t_3 = 200\text{ min}$ (v.l.n.r.) sind die pdfs im Intervall $\frac{x}{\sigma} \in [-15 : 15]$ aufgetragen.

linearen Anstieg zu größeren Skalen hin beobachten. Dass sich die Verteilungen, wie in Abbildung 2.2 zu sehen, leptokurtisch verhalten, ist eine wichtige Entdeckung. Anerkannte Kapitalmarkttheorien (z. B. Black-Scholes-Merton Optionspreistheorie [5], Capital Asset Pricing Method [45], Value at Risk ⁷) gehen von normalverteilten Kursen aus. Dass die Kurse jedoch wesentlich wildere Fluktuationen als die Normalverteilung aufweisen, hat massiven Einfluss auf entscheidungsrelevante Berechnungen, die aus diesen Theorien resultieren. Über die Schiefe von Renditeverteilungen gibt es in der Literatur verschiedene Ergebnisse. Sowohl links- als auch rechtsschiefe Verteilungen bei Indizes und Aktien wurden beobachtet [38]. Vermutlich sind diese Ergebnisse jedoch stark von der Wahl der Zeitskalen und der beobachteten Größe abhängig. Eine Schiefe in den Wahrscheinlichkeitsdichteverteilungen der hier verwendeten Daten auf den in Abbildung 2.2 betrachteten Skalen konnte nicht beobachtet werden.

2.3.2 Skalierungsverhalten

Eine Definition, die in diesem Zusammenhang passend ist, lautet gem. [20]: „Selbstähnlich ist [...] eine Form, wenn ein Teil bzw. ein Ausschnitt in verkleinerter Weise dem Ganzen ähnlich ist und wenn dieses Phänomen bei weiteren verkleinerten Ausschnitten wieder auftritt.“ Anschaulich bezeichnet man einen Prozess als *selbstähnlich*, wenn man Trajektorien eines Prozesses auf unterschiedlichen Zeitskalen qualitativ nicht voneinander unterscheiden kann, falls die Achsenbeschriftung fehlt. In Abbildung 2.3 soll für den DAX exemplarisch dargestellt werden wie ähnlich Kursverläufe für unterschiedliche Zeitskalen aussehen können. Bei strikter Selbstähnlichkeit liegt Skaleninvarianz vor. Diese lässt sich mit dem charakteristischen Exponenten eines Skalengesetzes beschreiben. Um einen ersten quantitativen Test für die Selbstähnlichkeit in einem Datensatz zu bekommen, trägt man den Mittelwert der absoluten Wertänderung gegen die Zeitskalen auf [13]:

$$\langle |x_{\Delta t}| \rangle_T \sim \Delta t^H. \quad (2.3)$$

Der Index T steht hier für die Größe der Zeitfenster über die gemittelt wird. H lässt sich als der *Hurst-Exponent* interpretieren. Der britische Hydrologe H. E. Hurst veröffentlichte 1951 in [27] Untersuchungen über die Abhängigkeiten in den Wasserständen des Flusses Nil. Später greift Mandelbrot diese Theorie auf und wendet sie auf die Fluktuationen von Baumwollpreisen an. $H > 0.5$ deutet Persistenz an, was im ökonomischen Kontext für ein trendverstärkendes System steht. $H < 0.5$ steht für ein mittelwertstationäres System welches als Antipersistent bezeichnet wird. Bei Brownscher Bewegung liegt $H = 0.5$ vor.

In der Literatur findet man bei Evertsz [13] ein $H \approx 0.54$ für den DAX im

⁷Risikomaß das von JPMorgan Chase & Co. entwickelt wurde.

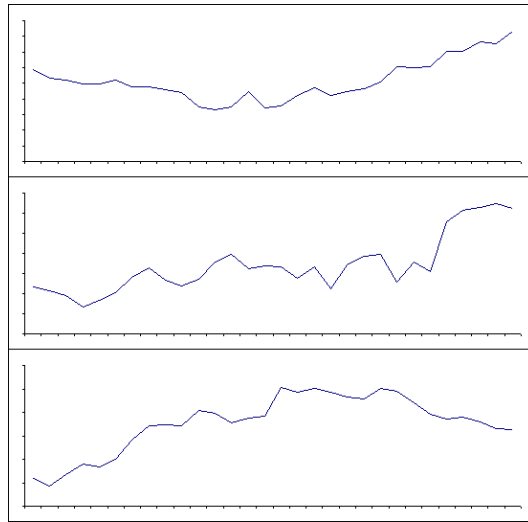


Abbildung 2.3: Zwanzig aufeinanderfolgende Tages-, Wochen- und Monatschlusskursnotierungen (v.o.n.u.) des DAX aus den Jahren 2008/09 (Datenquelle [53]). Ohne Beschriftung der Achsen fällt eine Abschätzung über die Größenordnung der Zeitskala schwer.

mittleren Skalenbereich. Ähnliche Ergebnisse ergeben sich in [21]. Für den vorliegenden Datensatz konnte ein Hurst-Exponent von

$$H = 0.5156 \pm 0.0014 \quad (2.4)$$

ermittelt werden. Somit wird hier mit einem einfachen Test ein selbstähnliches Verhalten für die Kursdifferenzen des DAX bestätigt. Die Koeffizienten des Potenzgesetzes lassen sich gut an die errechneten Daten anpassen, und der Hurst-Exponent stimmt mit den Werten aus der Literatur weitestgehend überein. Die kleinen Unterschiede erklären sich möglicherweise aus der Art der Vorbehandlung der Daten und über die Unterschiede bei der Wahl der Zeitintervalle T .

2.3.3 Abhängigkeiten

Eine weitere wohlbekannte Eigenschaft von Finanzdaten ist, dass die Abhängigkeiten zwischen den Renditen zeitlich sehr schnell verschwinden. Mit anderen Worten, die Preisdifferenzen sind schon nach einer sehr kurzen Zeitspanne vollständig unkorreliert. Bei quadrierten $x_{\Delta t}^2(t)$ oder absoluten $|x_{\Delta t}(t)|$ Preisdifferenzen liegen dagegen sehr lange Abhängigkeiten vor. Diese Eigenschaft ist auch als *Volatility Clustering* bekannt. Es bedeutet, dass es Perioden mit großen Fluktuationen und Perioden mit kleinen Fluktuationen gibt. Sobald der

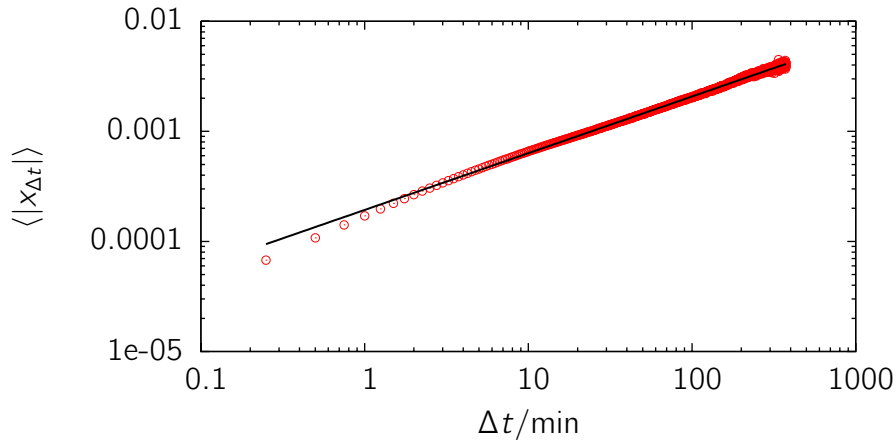


Abbildung 2.4: Der Mittelwert der absoluten Preisänderungen gegen die Zeitskala Δt in Minuten aufgetragen. Die schwarze Linie entspricht einem Skalengesetz.

Kurs also einmal stark ausgebrochen ist, bleibt die Volatilität über einen gewissen Zeitraum groß. Das Vorzeichen spielt dabei keine Rolle.

Die Berechnung einer Korrelationsfunktion setzt Stationarität der Zeitreihe voraus. Diese ist für empirische Daten, und damit auch für die vorliegenden Finanzdaten nicht unbedingt gegeben. Diese Eigenschaft wird in diesem Zusammenhang auch *Heterogenität* genannt. Dennoch lassen sich zumindest qualitative Aussagen treffen. Für die Berechnung einer Autokorrelationsfunktion (AKF) zwischen den Kursdifferenzen $x_{\Delta t}(t)$ lässt sich folgende Formel verwenden:

$$\gamma(x_{\Delta t}(t), x_{\Delta t}(t+T)) = \frac{\langle (x_{\Delta t}(t) - \langle x_{\Delta t}(t) \rangle) (x_{\Delta t}(t+T) - \langle x_{\Delta t}(t+T) \rangle) \rangle}{\sqrt{\langle (x_{\Delta t}(t) - \langle x_{\Delta t}(t) \rangle)^2 \rangle} \sqrt{\langle (x_{\Delta t}(t+T) - \langle x_{\Delta t}(t+T) \rangle)^2 \rangle}}.$$

Auf Grund der symmetrischen Verteilungsfunktionen 2.1 kann man die Mittelwerte vernachlässigen:

$$\gamma(x_{\Delta t}(t), x_{\Delta t}(t+T)) = \frac{\langle x_{\Delta t}(t) x_{\Delta t}(t+T) \rangle}{\sqrt{\langle x_{\Delta t}^2(t) \rangle} \sqrt{\langle x_{\Delta t}^2(t+T) \rangle}}. \quad (2.5)$$

Des Weiteren lässt sich die AKF des quadrierten Prozesses mit

$$\begin{aligned} & \gamma(x_{\Delta t}^2(t), x_{\Delta t}^2(t+T)) \\ &= \frac{\langle (x_{\Delta t}^2(t) - \langle x_{\Delta t}^2(t) \rangle) (x_{\Delta t}^2(t+T) - \langle x_{\Delta t}^2(t+T) \rangle) \rangle}{\sqrt{\langle (x_{\Delta t}^2(t) - \langle x_{\Delta t}^2(t) \rangle)^2 \rangle} \sqrt{\langle (x_{\Delta t}^2(t+T) - \langle x_{\Delta t}^2(t+T) \rangle)^2 \rangle}} \end{aligned} \quad (2.6)$$

darstellen. Die Mittelwerte dürfen hier nicht vernachlässigt werden. Es stellt sich die Frage, wie das T_0 , also der Startwert der Verzögerung T , zu wählen ist. Um

Abhängigkeiten aus Überschneidungen zu verhindern müsste man $T_0 > \Delta t$ wählen. Welche Auswirkungen die Wahl des T_0 hat, kann man der Abbildung 2.5 entnehmen. Wie zu erwarten, ergeben sich für $T_0 < \Delta t$ künstliche Korrelatio-

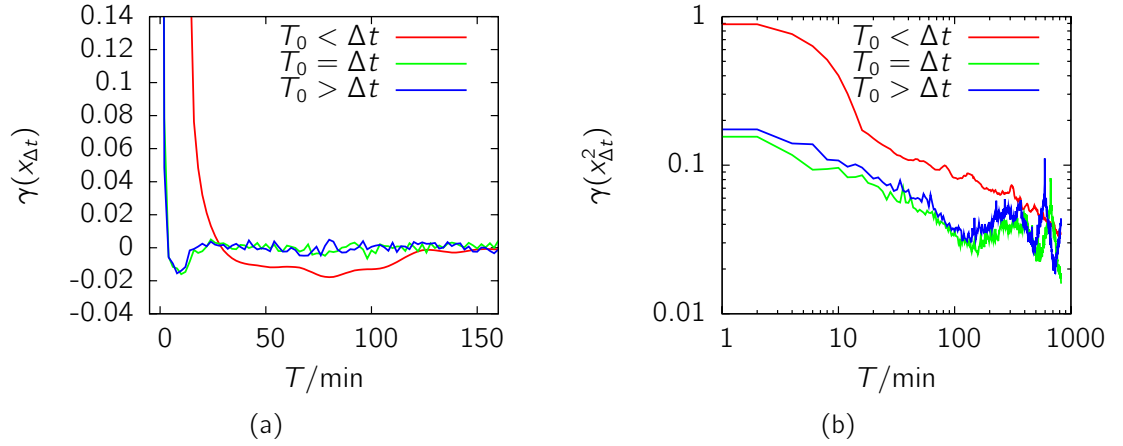


Abbildung 2.5: Die Auswirkungen der Wahl des T_0 auf die Form der Auto-korrelationsfunktionen. Auf der linken Seite ist die AKF bei $\Delta t = 1$ min aufgetragen. Rechts die Autokorrelation des quadrierten Prozesses bei $\Delta t = 1$ min.

nen aus den Überschneidungen der Kursdifferenzen. Es gibt jedoch auch einen kleinen Unterschied zwischen der Wahl $T_0 = \Delta t$ und $T_0 > \Delta t$. In den folgenden Berechnungen wird das letztgenannte Setting gewählt. Im kleinen Skalenbereich $T = 15$ s bis $T = 2$ min ist ein sehr schneller Abfall der AKF auf Null zu beobachten. Im Minutenbereich ist für Verzögerungen $5 \text{ min} < T < 15 \text{ min}$ eine Art Unterschwinger ins Negative zu sehen. Als Ursache für diesen Effekt werden in der Literatur die Angebots-Nachfrage-Preisspanne und die Eingriffe des Market-Makers genannt⁸. Nach einer Verzögerung von wenigen Minuten ist keine Korrelation mehr vorhanden. Für den quadrierten Prozess fällt die AKF zwar sehr schnell auf einen niedrigen Wert, bleibt dann jedoch langsam abfallend über Null.

Auf größeren Skalen fällt die AKF ebenfalls sofort auf Null. Auch hier ist zu sehen, dass die AKF der quadrierten Inkremente nur sehr langsam verschwindet. Die gestrichelte Linie beschreibt einen Fit mit der Potenzfunktion:

$$p(x) = \frac{1}{1 + x^\alpha}. \quad (2.7)$$

Das hier gefittete

$$\alpha = 0.47 \pm 0.01 \quad (2.8)$$

entspricht größenordnungsmäßig den Werten in der Literatur [31].

⁸Für genauere Ausführungen siehe [8].

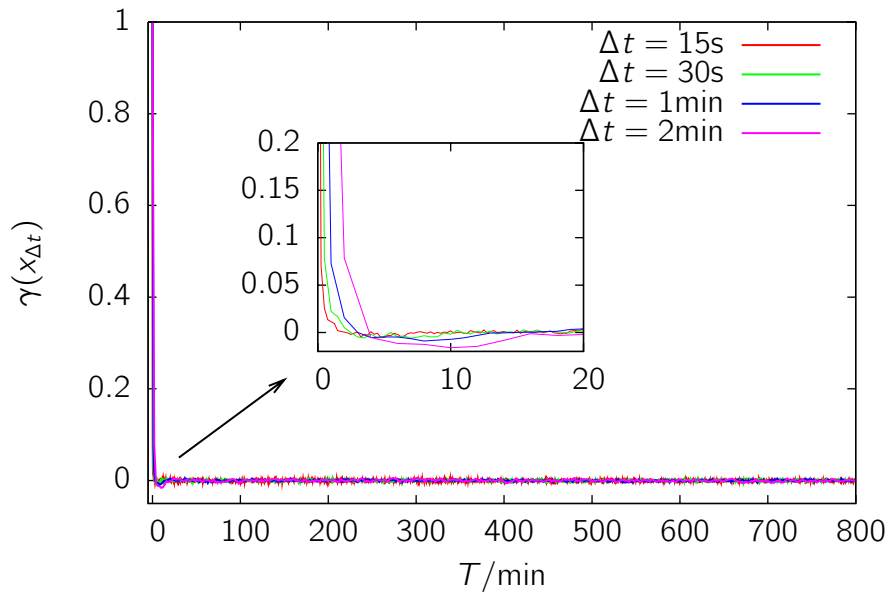


Abbildung 2.6: Autokorrelationsfunktion der Preisdifferenzen $x_{\Delta t}$ für unterschiedliche $\Delta t = 15\text{ s}, 30\text{ s}, 1\text{ min}, 2\text{ min}$ gegen die Zeitspanne T in Minuten. Der kleinere Graph stellt eine Vergrößerung der ersten zwanzig Minuten dar.

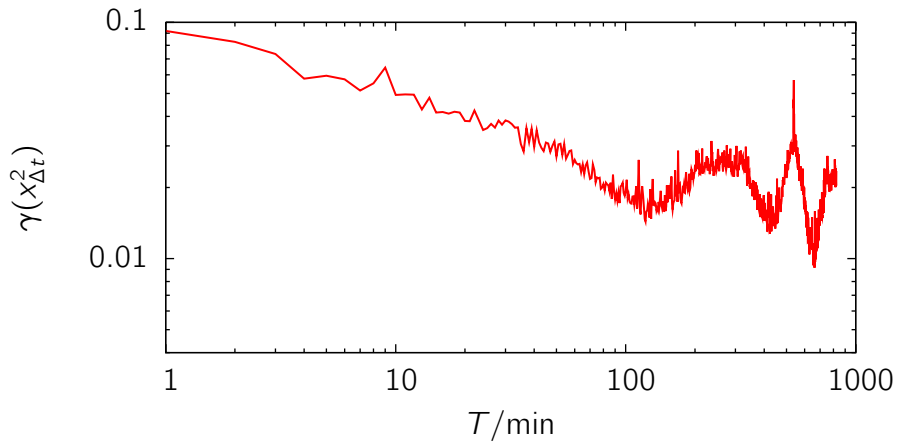
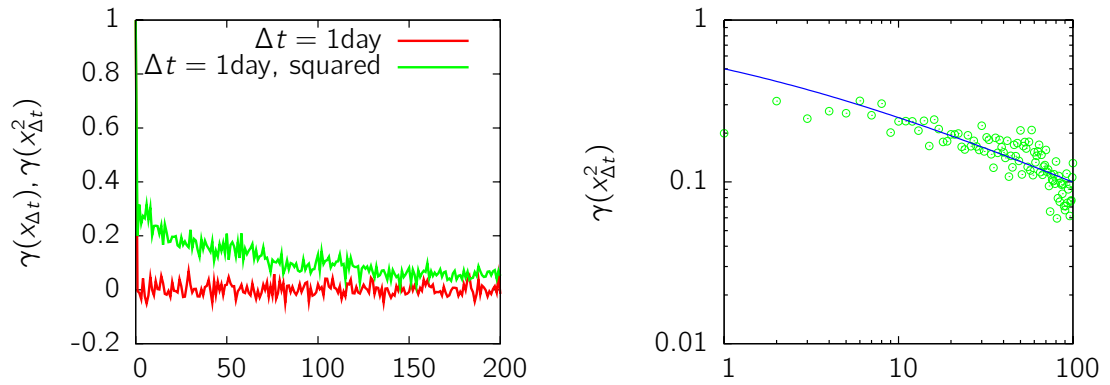


Abbildung 2.7: Autokorrelationsfunktion des quadrierten Prozesses $x_{\Delta t}^2$ für $\Delta t = 1\text{ min}$. Die Kurve ist mit doppelt logarithmierter Skala aufgetragen.



(a) AKF der normalen und quadrierten Kursdifferenzen des DAX auf Tagesschlusskursbasis. (b) Quadrierte Kursdifferenzen in einem doppelt logarithmierten Diagramm. Die durchgezogene Linie repräsentiert das Potenzgesetz (vgl. (2.7) und (2.8)).

Abbildung 2.8: Autokorrelation der DAX-Tagesschlusskurse.

Die empirischen Untersuchungen zu Korrelationen in Finanzzeitreihen konnten anhand des vorliegenden Datensatzes bestätigt werden. Insbesondere für die großen Zeitskalen sind die Effekte sehr deutlich. Dementsprechend ist das Volatility Clustering auch nicht auf kleinere Skalen beschränkt, sondern tritt beispielsweise auch bei Tagesschlusskursen auf. Die Mechanismen, die dieses Phänomen hervorrufen, lassen sich mit den hier verwendeten statistischen Methoden nicht erklären. Bessere Erfolgsaussichten die Gründe für Volatility Clustering zu identifizieren verspricht man sich von sogenannten *agentenbasierten Modellen* [51, 9].

Kapitel 3

Drift- und Diffusionschätzung von Finanzdaten

Die Möglichkeit, Drift und Diffusion direkt aus empirischen Daten zu schätzen, wurde schon in vielen Disziplinen angewendet. Als physikalische Beispiele kann man z. B. turbulente Felder [32] und Synchronisation [46] nennen. Aber auch Schätzungen mit medizinischen Daten oder Finanzdaten wurden durchgeführt (vgl. z. B. Friedrich et al. [16, 15]). Als einführendes Beispiel wird an dieser Stelle kurz auf die Anwendung bei Verkehrsdaten [29] eingegangen. Es findet eine zweidimensionale Betrachtung statt, wie sie auch in Kapitel 4 angewendet wird. Auf einer Autobahn in der Nähe von Köln wurden der Verkehrsfluss q und die Geschwindigkeit v der Fahrzeuge auf drei Spuren gemessen. Diese beiden Prozesse wurden im Verkehrszustandsvektor \mathbf{x} zusammengefasst. Per Driftschätzung konnten Erkenntnisse über die deterministische Flussdynamik des Verkehrs gewonnen werden. In Abbildung 3.1 sind Drift von Verkehrsfluss und Geschwindigkeit in einem Phasenraumdiagramm aufgetragen. Man kann zwei stabile Fixpunkte erkennen. Der eine stabile Fixpunkt bei $v_2 \approx 110$ km/h steht dabei für einen freien Verkehrsfluss knapp oberhalb der erlaubten Höchstgeschwindigkeit von 100 km/h. Der andere stabile Fixpunkt beschreibt zäh fließenden Verkehr mit einer Geschwindigkeit bei $v_1 \approx 30$ km/h.

In Kapitel 4 der vorliegenden Arbeit wird ebenfalls versucht werden, durch mehrdimensionale Drift- und Diffusionsschätzung die Dynamik eines Systems zu erklären. Dabei spielt jedoch der stochastische Teil der Gleichungen eine wichtigere Rolle als der deterministische. Im jetzigen Kapitel wird man sich auf die eindimensionale Betrachtung der in Kapitel 2 charakterisierten DAX-Daten beschränken. Zunächst wird kurz auf die Besonderheiten bei der Drift- und Diffusionsschätzung von empirischen Daten eingegangen. Danach wird der Datensatz auf die Markov-Eigenschaft geprüft. Die Koeffizientenschätzung erfolgt anschließend für die DAX-Daten und einen GARCH-Prozess, der auf diese Daten angepasst wurde.

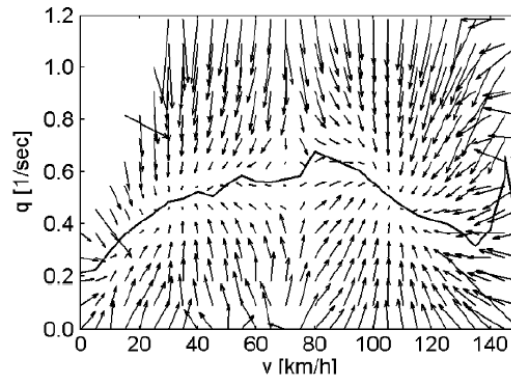


Abbildung 3.1: Die deterministische Verkehrsdynamik einer dreispurigen Autobahn in der Nähe von Köln mit einer Geschwindigkeitsbegrenzung von 100 km/h (entnommen aus [29]).

Zunächst ein kurzer Kommentar dazu, welche Probleme bei der Anwendung der Koeffizientenschätzung bei empirischen Daten Einfluss auf die Qualität der Ergebnisse haben. Grundsätzlich kann es für empirische Daten mit endlicher Samplingrate und damit endlichen Zeitinkrementen Qualitätsverluste bei der Abschätzung von Drift und Diffusion geben. Zum einen sind die Koeffizienten für den Übergang $\tau \rightarrow 0$ definiert, so dass es bei endlichem τ zu Einschränkungen kommt. Des Weiteren entstehen statistische Fehler durch die endliche Anzahl an Datenpunkten und die Diskretisierung des Ortsraums, das sogenannte *Binning* [47]. Die Auswirkungen all dieser Effekte werden ausführlich in [28] diskutiert. Da es sich in diesem Fall um diskret abgetastete Finanzdaten handelt, ist es nicht möglich, die Samplingrate zu erhöhen, wie es möglicherweise in einem Laborexperiment realisierbar wäre. Beim Binning ist vor allem darauf geachtet worden, dass ein gutes Verhältnis aus geringer Binbreite und einer möglichst hohen Anzahl an Datenpunkten in den einzelnen Bins vorliegt.

Eine wichtige Forderung an die zu untersuchenden Daten ist das Vorliegen der Markov-Eigenschaft, da sie die Voraussetzung für Vereinfachungen in den Ausführungen dieser Arbeit ist. Gemäß der Definition wäre es eigentlich notwendig Gleichung (1.17) für alle N und alle Zeitinkremente $\tau_1, \tau_2, \dots, \tau_N$ zu bestätigen. Da dies nicht möglich ist, greift man auf praxisnahe Methoden zurück, die im Folgenden beschrieben werden.

Zu prüfen wäre, ob die einfach bedingte pdf mit der zweifach bedingten pdf übereinstimmt. In diesem Fall würde man die Gleichung (1.17) auf $N = 3$ vereinfachen. Üblicherweise wird die zweite Bedingung als konstant angenommen und Schnitte (z.B. $x_3 = 0$) der Verteilung geprüft:

$$p(x_1, \tau_1 | x_2, \tau_2) \stackrel{!}{=} p(x_1, \tau_1 | x_2, \tau_2; x_3 = 0, \tau_3). \quad (3.1)$$

Dieses Verfahren wird z. B. in [49] und [15] angewendet. Das Erzeugen der

zweifach bedingten pdf verlangt eine relativ große Anzahl an Datenpunkten. Dies schränkt die Anwendbarkeit teilweise ein.

Natürlich ist es auch möglich, die Gültigkeit der Chapman-Kolmogorov-Gleichung gemäß (1.20) nachzuweisen. Dabei werden die direkt geschätzte pdf und die per numerischer Integration errechnete rechte Seite der Gleichung miteinander verglichen. Dieses Verfahren wird zum Beispiel in [19] verwendet. Eine gute Übereinstimmung lässt auf einen Markov-Prozess schließen. Neben dem visuellen Vergleich lassen sich die Beobachtungen auch quantifizieren. Distanzmaße wie zum Beispiel die Kullback-Leibner-Entropie oder die χ^2 -Distanz liefern Kennzahlen, die eine Aussage zum Grad der Übereinstimmung zwischen zwei Verteilungen zulässt. In dieser Arbeit soll stattdessen der gewichtete, mittlere quadrierte Fehler im logarithmischen Raum d_M gemessen werden. Die folgende Formel ist an die in [39] verwendete Definition angelehnt:

$$d_M(p_l(x_1, \tau_1 | x_3, \tau_3), p_r(x_1, \tau_1 | x_3, \tau_3)) = \frac{\int dx_1 dx_3 (p_l + p_r)(\ln p_l - \ln p_r)^2}{\int dx_1 dx_3 (p_l + p_r)(\ln^2 p_l - \ln^2 p_r)}. \quad (3.2)$$

Im rechten Teil der Gleichung wurde aus Gründen der Übersichtlichkeit auf die Argumente verzichtet.

3.1 Ergebnisse für den DAX

Aus den in Kapitel 2 charakterisierten DAX-Kursdifferenzen sollen nun die Drift- und Diffusionskoeffizienten des Prozesses geschätzt werden. Vorher wird überprüft ob dieser Prozess die Markov-Eigenschaft erfüllt.

3.1.1 Markovanalyse für die Kursdifferenzen des DAX

Zunächst wird die Gültigkeit der Chapman-Kolmogorov-Gleichung für den Datensatz überprüft und die oben angegebenen Distanzen berechnet. Danach findet ein Vergleich zwischen zweifach und dreifach bedingten Verteilungen statt. In beiden Fällen werden dazu so genannte *Contourplots* erstellt. In den Graphen sind die Höhenlinien der pdfs dargestellt. Das vereinfacht die Beurteilung über den Grad der Übereinstimmung.

Chapman-Kolmogorov-Gleichung

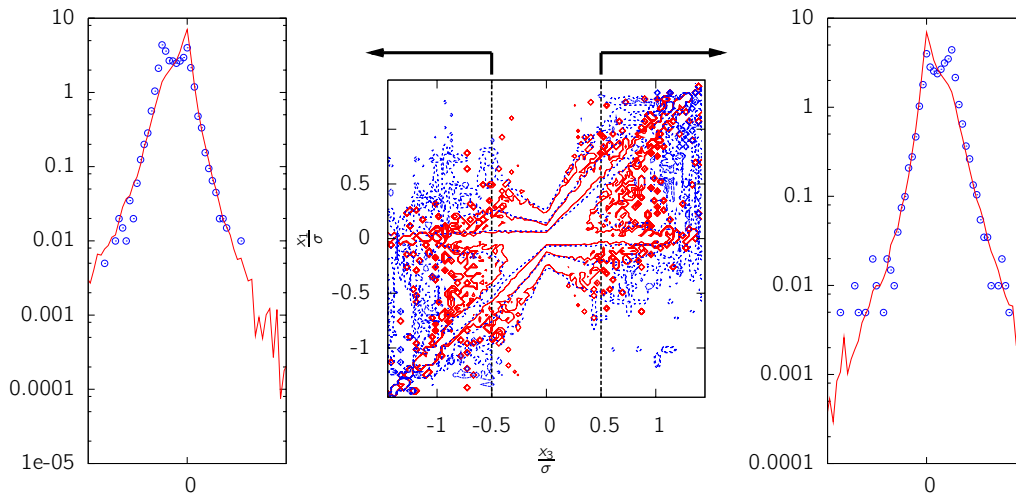
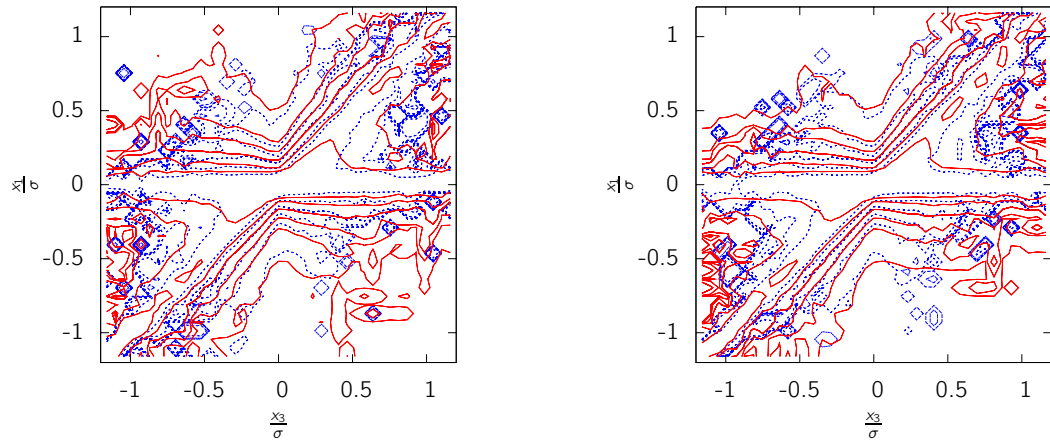


Abbildung 3.2: Chapman-Kolmogorov-Gleichung mit $\tau_1 = 30$ s, $\tau_2 = 45$ s und $\tau_3 = 60$ s. Die durchgezogenen Linien (rot) repräsentieren die direkt geschätzte cpdf $p(x_3, \tau_3 | x_1, \tau_1)$, während die gestrichelten Linien (blau) für die errechneten Verteilungen $\int dx_2 p(x_3, \tau_3 | x_2, \tau_2) p(x_2, \tau_2 | x_1, \tau_1)$ stehen. Der linke und rechte Graph zeigen Schnitte durch die bedingten Wahrscheinlichkeiten bei $x_1 \approx \pm 0.5\sigma$. Die Standardabweichung bezieht sich auf die kleinste Zeitskala.

Die beiden Seiten der Chapman-Kolmogorov Gleichung stimmen in Abbildung 3.2 im inneren Wertebereich um den Nullpunkt sehr gut überein. Nach außen hin stehen deutlich weniger Datenpunkte zur Verfügung und die Statistik wird schlechter. Die beiden Schnitte durch die bedingten Verteilungen zeigen ebenfalls eine akzeptable Übereinstimmung. In Abbildung 3.3(a) und 3.3(b) kann man erkennen, dass auch für den mittleren Skalenbereich annähernd deckungsgleiche Verteilungen vorliegen gegeben sind. Die Ergebnisse der Berechnungen des mittleren, quadratischen Fehlers sind Tabelle 3.1 zu entnehmen.

Tabelle 3.1: Mittlerer, quadratischer Fehler im logarithmischen Raum als Maß für die Übereinstimmung der Chapman-Kolmogorov Gleichung.

Zeitskalen τ	d_M
30,45,60 s	0.021
60,120,180 s	0.014
5,10,15 min	0.013



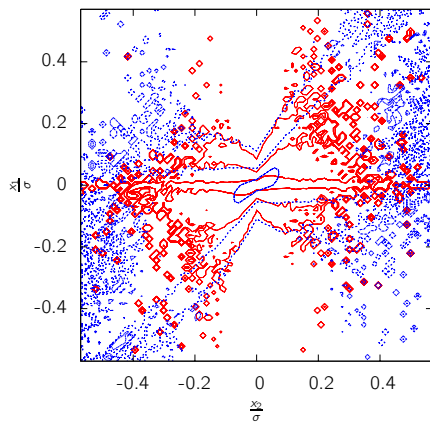
(a) $\tau_1 = 1 \text{ min}$, $\tau_2 = 2 \text{ min}$ und $\tau_3 = 3 \text{ min}$.

(b) $\tau_1 = 5 \text{ min}$, $\tau_2 = 10 \text{ min}$ und $\tau_3 = 15 \text{ min}$.

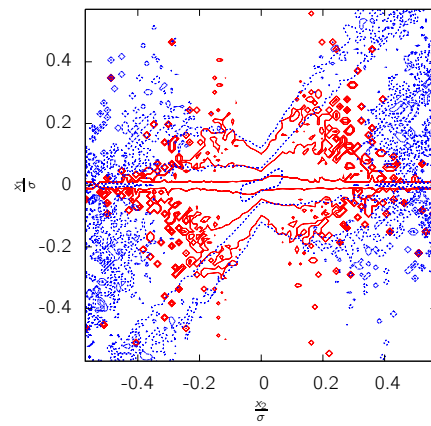
Abbildung 3.3: Methodisch identisch wie in Abb. 3.2 werden die beiden Seiten der Chapman-Kolmogorov-Gleichung verglichen. Die verwendeten Zeitskalen sind unter den Graphen angegeben.

Vergleich der einfach und zweifach bedingten Wahrscheinlichkeitsdichteverteilungen

Die Wahrscheinlichkeitsdichteverteilungen aus Gleichung (1.17) sind für den vereinfachten Fall $N = 3$ in Abbildung 3.4 aufgetragen. Es liegt in diesem Fall keine perfekte Deckungsgleichheit vor, jedoch kann man von einer guten Übereinstimmung im Rahmen der Fehlertoleranzen reden. Im Folgenden wird für die betrachteten Skalen das Vorliegen der Markov-Eigenschaft angenommen.



(a) $\tau_1 = 30$ s, $\tau_2 = 45$ s und $\tau_3 = 60$ s



(b) $\tau_1 = 1$ min, $\tau_2 = 2$ min und $\tau_3 = 3$ min

Abbildung 3.4: Vergleich der Kontourlinien der bedingten Wahrscheinlichkeiten $p(x_1, \tau_1 | x_2, \tau_2)$ (rote, durchgezogene Linien) und $p(x_1, \tau_1 | x_2, \tau_2; x_3 = 0, \tau_3)$ (blaue, gestrichelte Linien). Die verwendeten Zeitskalen sind unter den Graphen angegeben, und die Standardabweichungen beziehen sich auf die jeweiligen kleinsten Skalen τ_1 .

3.1.2 Drift- und Diffusionsschätzungen für die DAX-Kursdifferenzen

Gem. Definition (1.32) werden die Drift- und Diffusionskoeffizienten aus den bedingten Momenten berechnet:

$$D_D^{(n)}(x, t) = \lim_{\tau \rightarrow 0} \frac{1}{\tau} \frac{\langle [X(t + \tau) - X(t)]^n | X(t) = x_D \rangle}{n!} \quad (3.3)$$

mit

$$x_D = y(t + \Delta t) - y(t). \quad (3.4)$$

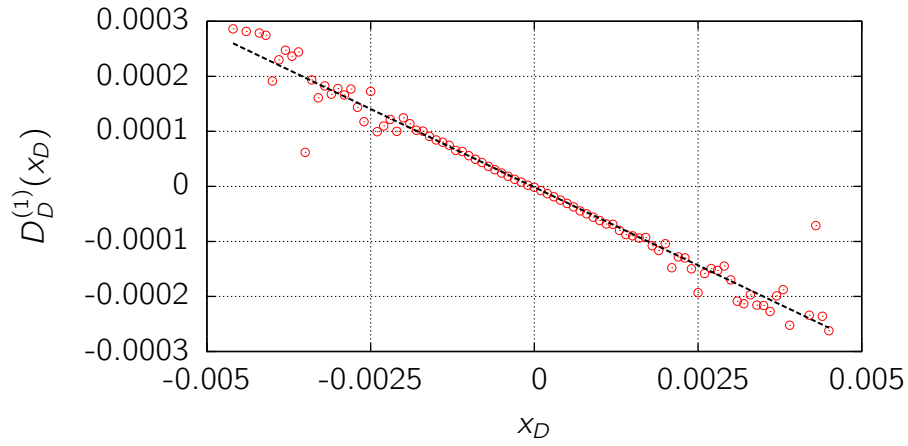
Es wird im Folgenden nicht die gesamte Kaskade, also der Prozess auf unterschiedlichen Skalen, sondern lediglich die kleinstmögliche Zeitskala τ_{min} betrachtet. Diese sollte einerseits möglichst klein gewählt werden um Schätzfehler zu minimieren, aber auch größer als die Zeitdifferenz zwischen den Kursnotierungen Δt um künstliche Korrelationen auszuschließen. Für die Schätzung wird $\Delta t = 15$ s gewählt und damit ergibt sich die optimale Zeitskala zu $\tau_{min} = 30$ s. In Abbildung 3.5 sind Drift und Diffusion gegen die logarithmierten Kursdifferenzen $x_{\tau_{min}}$ aufgetragen. Wie man in Abb. 3.5(a) erkennen kann, hat der Drift einen deutlich linearen Charakter mit negativer Steigung. Folgende Funktion lässt sich gut anpassen:

$$I(x_D) = D_D^{(1)} = (-0.057 \pm 0.001) x_D. \quad (3.5)$$

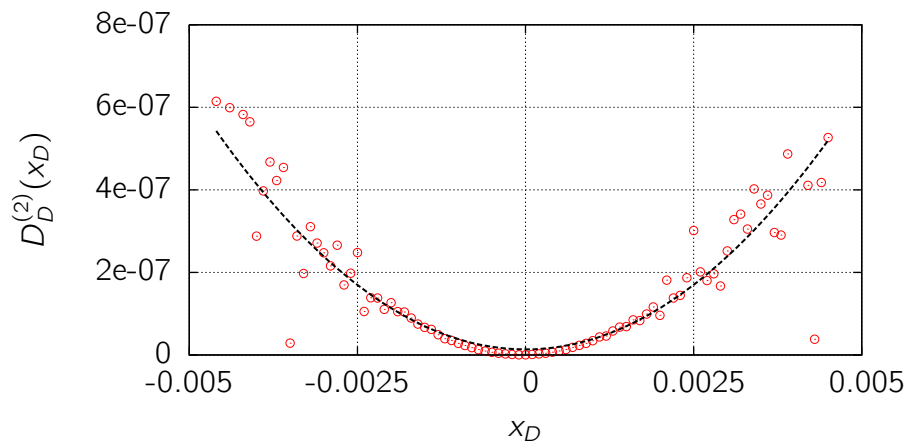
Wenn man den Drift als erwartete Momentanrendite interpretiert, kann man den Versuch unternehmen die Form der Driftfunktion ökonomisch zu interpretieren. Je weiter der Kurs in eine Richtung (positiv oder negativ) ausgelenkt ist, desto mehr geht die Erwartung der folgenden Rendite in die entgegengesetzte Richtung. Mit anderen Worten bewegen sich die Kursdifferenzen ständig um den langfristigen Erwartungswert herum und werden immer wieder zu diesem „zurückgezogen“. Die Form der Diffusion in Abb. 3.5(b) entspricht einer Parabel. Die Varianz steigt mit größerer Auslenkung quadratisch an. Entsprechend lässt sich die Diffusion durch eine quadratische Funktion darstellen:

$$\sqrt{q(x_D)} = D_D^{(2)} = (0.025 \pm 0.001) x_D^2. \quad (3.6)$$

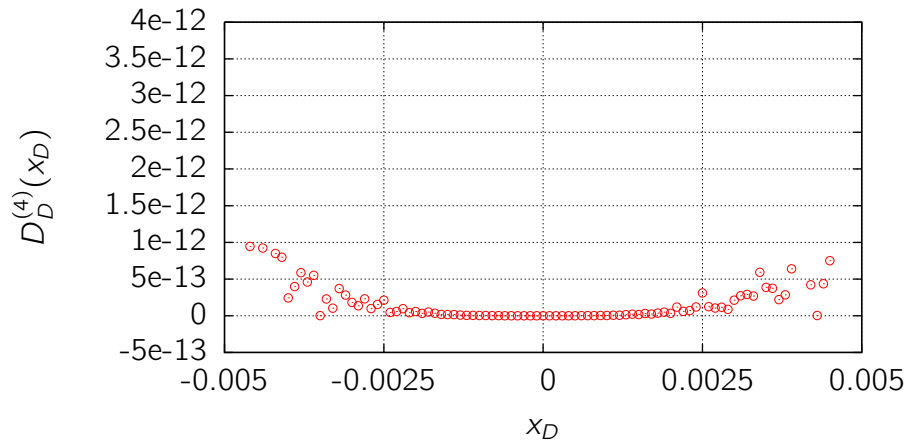
Es sei angemerkt, dass für die Drift- und Diffusionsfunktionen ein endlicher Offset existiert, dieser jedoch um Größenordnungen unter den multiplikativen Parametern liegt. Der vierte Koeffizient $D^{(4)}$ liegt ebenfalls einige Ordnungen unter dem Diffusionsterm (vgl. Abbildungen 3.5(b) und 3.5(c)). In erster Näherung kann dieser vernachlässigt werden. Gemäß Pawulas-Theorem (siehe Abschnitt 1.2.4) bricht damit die Kramers-Moyal-Entwicklung nach dem zweiten Term ab. Der Prozess lässt sich also vollständig durch die ersten beiden Momente charakterisieren. Diese Ergebnisse decken sich mit Untersuchungen,



(a)



(b)



(c)

Abbildung 3.5: Drift- (a), Diffusions- (b) und der vierte Kramers-Moyal-Koeffizient (c) als Funktion der logarithmierten Kursdifferenzen des DAX bei $\tau_{min} = 30$ s. Die gestrichelten Kurven stellen die angepassten Funktionen (3.5) und (3.6) dar.

die im Zusammenhang mit FOREX-Kursen in [15] gemacht wurden. Es wurden ebenfalls eine lineare Driftstruktur und eine quadratische Diffusionsfunktion entdeckt. Die entsprechende Langevin-Gleichung, die zu einer Simulation der Trajektorien herangezogen werden könnte, hat gemäß den Gleichungen (1.34) folgende Form:

$$\dot{x} = l(x) + \sqrt{q(x)}\Gamma(t). \quad (3.7)$$

Es wäre nun möglich, einen Selbstkonsistenztest durchzuführen. Dazu würde man mit Hilfe der Gleichung (3.7) einen künstlichen Prozess simulieren ¹ und die Statistik mit dem Ursprungsprozess vergleichen. Im Zusammenhang mit FOREX-Kursen [15] wurden diese Schritte durchgeführt und man ist zu einem guten Ergebnis gekommen. So konnte gezeigt werden, dass die numerischen Lösungen der Fokker-Planck-Gleichung und die empirischen Beobachtungen sowohl bezüglich der Ein-Punkt Wahrscheinlichkeitsdichten auf verschiedenen Skalen, als auch der bedingten Dichten in guter Übereinstimmung zueinander stehen. An dieser Stelle soll auf einen Vergleich dieser Art verzichtet werden.

3.2 Ergebnisse für einen GARCH(1,1)-Prozess

ARCH²- und GARCH-Modelle werden in der Finanzmathematik zur Zeitreihenanalyse und -prognose genutzt. Sie sind in der Lage, charakteristische Eigenschaften von Finanzdaten wie Leptokurtosis (siehe Abbildung 2.2) und Heteroskedastie (siehe Abbildung 2.1(e)) nachzubilden [6]. Skalierungseigenschaften können nicht abgebildet werden. In Abbildung 3.6 werden die pdfs eines GARCH(1,1)-Prozesses mit denen eines hochfrequenten S&P500 Datensatzes verglichen. Die Parameter des GARCH-Prozesses wurden aus der Zeitreihe des empirischen Datensatzes für eine feste Zeitskala $\Delta t = 1$ min geschätzt. Der Vergleich auf unterschiedlichen Skalen zeigt, dass die pdfs auseinanderlaufen. Es ist also notwendig, für jede Skala die Parameter des Prozesses neu zu schätzen. Parameterschätzungen dieser Art werden häufig mit Hilfe einer Quasi Maximum Likelihood (QML) Schätzung ermittelt (siehe z.B. [11]).

3.2.1 Die Definition des Prozesses

Allgemein ist ein GARCH(p,q)-Prozess definiert als:

$$\sigma_t^2 = \alpha_0 + \alpha_1 X_{t-1}^2 + \dots + \alpha_q X_{t-q}^2 + \beta_1 \sigma_{t-1}^2 + \dots + \beta_p \sigma_{t-p}^2,$$

¹Wie eine solche Computersimulation durchzuführen wäre, ist z.B. in [43] auf Seite 60ff. beschrieben.

²Ist Abkürzung für „AutoRegressive Conditionally Heteroscedastic“.

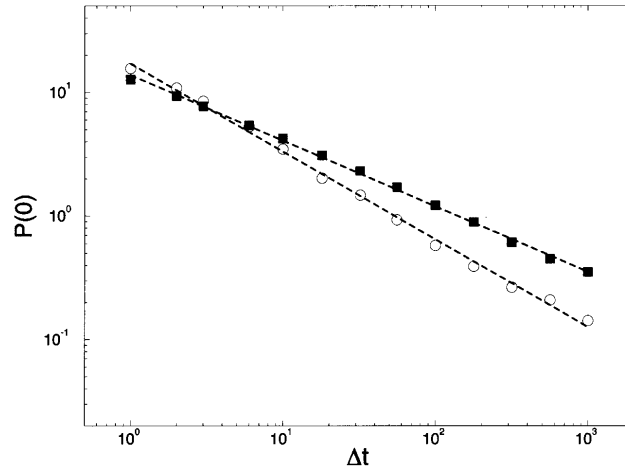


Abbildung 3.6: Skalierungseigenschaften eines GARCH(1,1)-Prozesses (schwarze Quadrate) mit konstanten Parametern und eines S&P500-Datensatzes (weiße Kreise) auf unterschiedlichen Skalen Δt . $P(0)$ steht hier für das Maximum der Dichten (entnommen aus [35]).

wobei $\alpha_0, \alpha_1, \dots, \alpha_q, \beta_1, \dots, \beta_p$ Kontrollparameter sind. X_t ist eine Zufallsvariable mit $\langle X_t \rangle = 0$ und Varianz σ_t^2 die beliebig verteilt ist. In dieser Arbeit soll der einfachste Prozess, der GARCH(1,1)-Prozess mit normalverteilten Zufallsvariablen Anwendung finden:

$$X_t = \Gamma_t \sigma_t \quad (3.8)$$

$$\sigma_t = \sqrt{\alpha_0 + \alpha_1 X_{t-1}^2 + \beta_1 \sigma_{t-1}^2}, \quad (3.9)$$

wobei Γ_t gaußverteilte Zufallsvariablen mit verschwindendem Mittelwert und Einheitsvarianz sind. Es kann gezeigt werden, dass für diesen Prozess gemäß [35] folgende Relationen gelten:

$$\sigma^2 = \frac{\alpha_0}{1 - \alpha_1 - \beta_1} \quad (3.10)$$

$$\kappa = 3 + \frac{6\alpha_1^2}{1 - 3\alpha_1^2 - 2\alpha_1\beta_1 - \beta_1^2}. \quad (3.11)$$

σ^2 wird als *bedingte Varianz*³ bezeichnet und κ ist die Kurtosis gemäß der Definition (1.10). Über diese Gleichungen lassen sich die Parameter α_0, α_1 direkt aus den empirischen Daten bestimmen.

³Die Varianz einer empirischen Zeitreihe auf einem großen Zeitintervall (Gemäß [35]).

3.2.2 Parametrisierung und statistische Eigenschaften

Für den dritten Parameter wird gemäß [35] ein

$$\beta_1 = 0.9 \quad (3.12)$$

vorgegeben. Für den DAX-Datensatz lässt sich eine Kurtosis von $\kappa \approx 79.78$ bei $\tau = 15$ s ablesen (siehe Abbildung 2.2). Für die bedingte Varianz wurde ein $\sigma \approx 2.3 \cdot 10^{-3}$ bestimmt. Daraus lassen sich die ersten beiden Parameter

$$\alpha_0 \approx 2 \cdot 10^5 \quad (3.13)$$

und

$$\alpha_1 \approx 0.0913 \quad (3.14)$$

gemäß (3.10) berechnen. Mit Hilfe dieser Parameter lassen sich nun Trajektorien des Prozesses simulieren. Eine entsprechende Simulation wurde mit den Anfangsbedingungen $x_0 = 0$ und $\sigma_0 = 0$ für $N = 4 \cdot 10^6$ Datenpunkte durchgeführt. Die unabhängigen, gaußverteilten Zufallsvariablen Γ_t wurden mit Hilfe des Box-Muller-Verfahrens [41] generiert. Die Eigenschaften dieses generierten Prozesses auf Basis der geschätzten Parameter sollen nun mit den Charakteristika des empirischen Prozesses verglichen werden.

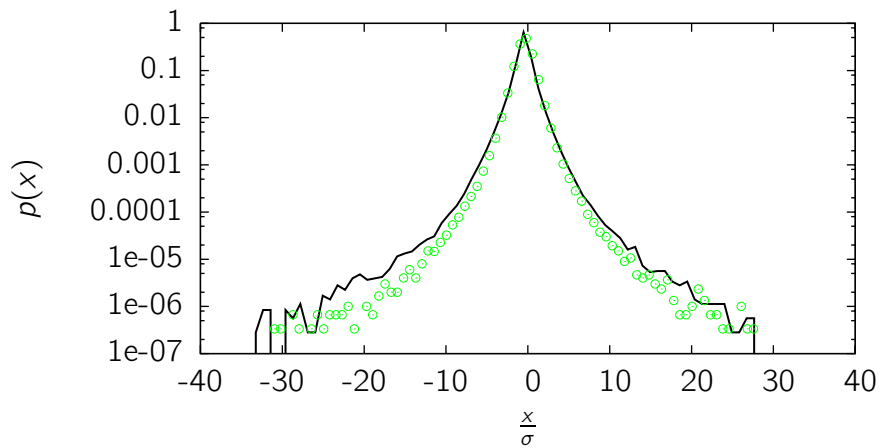


Abbildung 3.7: Vergleich der empirischen pdf des DAX mit $\tau = 60$ s (durchgezogene Kurve) mit der pdf eines GARCH(1,1)-Prozesses (Kreise) mit einer Parametereinstellung gem. (3.12), (3.13) und (3.14).

Der Vergleich der Wahrscheinlichkeitsdichten macht deutlich, dass sich schon der einfachste GARCH-Prozess recht gut an den DAX-Datensatz anpassen lässt. Leichte Abweichungen treten in den Außenbereichen der Verteilung auf. Die logarithmierte Darstellung in Abb. 3.7 macht dies deutlich. Eine verbesserte

Anpassung könnte durch die Hinzunahme von zusätzlichen Parametern α_i, β_i erfolgen.

Eine Übereinstimmung der linken und rechten Seite der Chapman-Kolmogorov-Gleichung bei $\tau_1 = 30$ s, $\tau_2 = 45$ s und $\tau_3 = 60$ s ist nicht gegeben (Abbildung 3.8). Dies ist ein Hinweis darauf, dass der Prozess nicht die Markov-

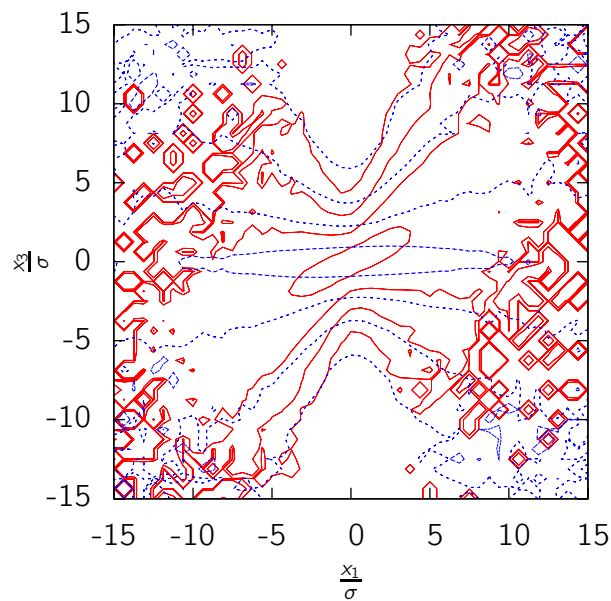


Abbildung 3.8: Chapman-Kolmogorov-Gleichung für $\tau_1 = 30$ s, $\tau_2 = 45$ s und $\tau_3 = 60$ s. Datengrundlage ist der simulierte GARCH(1,1)-Prozess mit der oben genannten Parametereinstellung.

Eigenschaft erfüllt. Die Tabelle 3.2 stellt eine Erweiterung von Tabelle 3.1 dar. Zusätzlich ist der mittlere, quadratische Fehler für den GARCH(1,1)-Prozess auf den angegebenen Skalen eingetragen. Der Wert ist um zwei Größenordnungen höher als die für den DAX errechneten Distanzen. Dies ist ebenfalls ein Hinweis darauf, dass die bedingten Wahrscheinlichkeitsdichteverteilungen des angepassten GARCH-Prozesses nicht in dem Maß übereinstimmen, wie sie es für die DAX-Daten tun.

Die Autokorrelationsfunktionen des GARCH(1,1)-Prozesses geben die Eigenschaften wieder, die auch für die empirischen DAX-Kursdifferenzen in Kapitel 2.3.3 beobachtet wurden. Allerdings sind die hier beobachteten Korrelationen für den quadrierten Prozess deutlich kürzer als im DAX-Datensatz, wie sie in Abbildung 2.7 aufgetragen sind. Die Drift- und Diffusionsanalyse des simu-

Tabelle 3.2: Mittlerer, quadratischer Fehler im logarithmischen Raum (vgl. Definition (3.2)) als Maß für die Übereinstimmung der Chapman-Kolmogorov-Gleichung.

Zeitskalen τ	d_M
DAX 30,45,60 s	0.021
DAX 60,120,180 s	0.014
DAX 5,10,15 min	0.013
GARCH 15,30,45 s	1.331

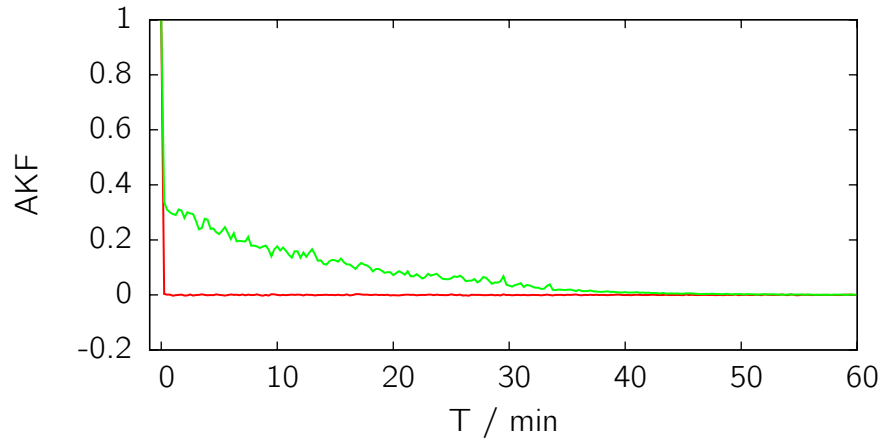


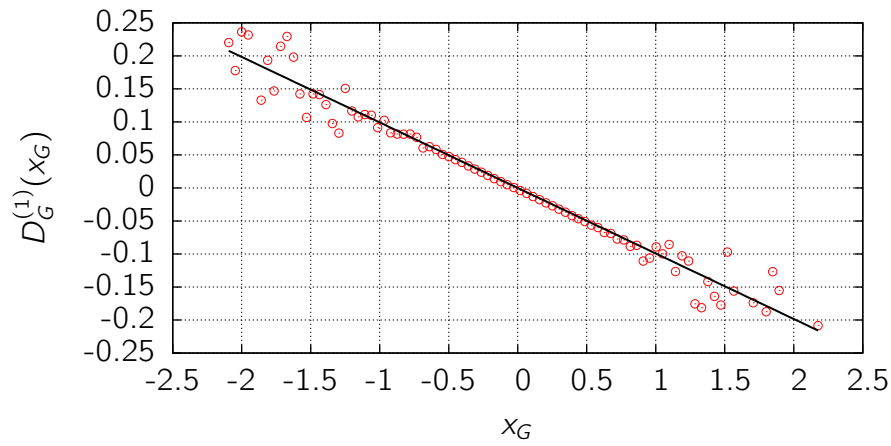
Abbildung 3.9: Gemäß Gleichung (2.5) und (2.6) der errechneten Autokorrelationsfunktion der normalen (rote Kurve) und quadrierten (grüne Kurve) Kursdifferenzen des GARCH(1,1)-Prozesses gegen die Zeitspanne T in Minuten.

lierten Prozesses auf Basis von Definition (3.8) lieferte folgende Ergebnisse:

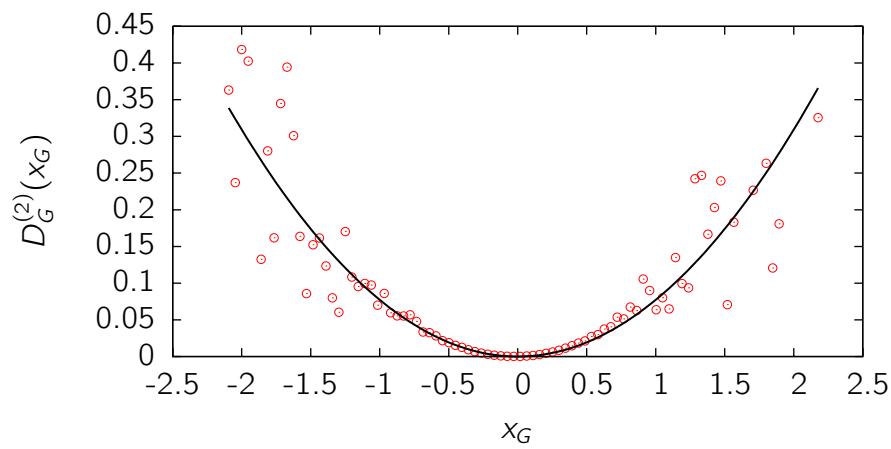
$$D_G^{(1)} = (-0.099 \pm 0.002) \cdot x_G, \quad (3.15)$$

$$D_G^{(2)} = (0.077 \pm 0.003) \cdot x_G^2. \quad (3.16)$$

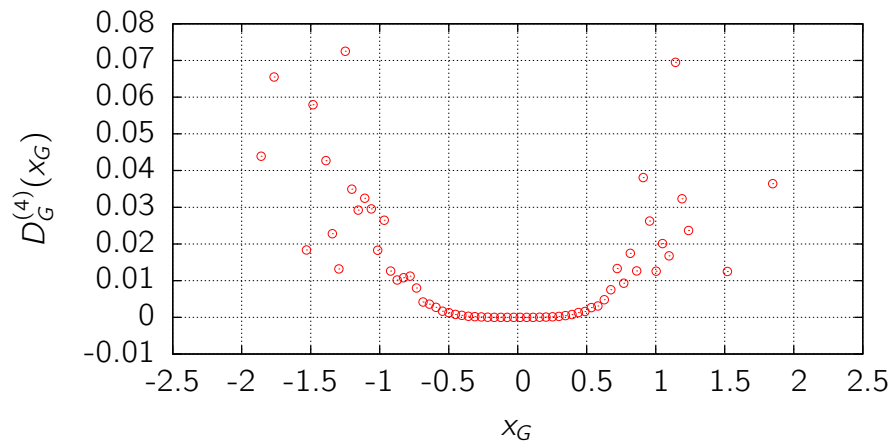
Auch hier wurde auf die explizite Darstellung des endlichen, sehr kleinen Offsets verzichtet.



(a)



(b)



(c)

Abbildung 3.10: Geschätzter Drift- (a), Diffusions- (b) und vierter Kramers-Moyal-Koeffizient (c) des GARCH(1,1)-Prozesses mit den oben genannten Parametereinstellungen.

Abschließende Betrachtung

Zur erweiterten Charakterisierung der in Kapitel 2 vorgestellten DAX-Daten wurde in diesem Abschnitt eine Drift- und Diffusionsschätzung durchgeführt. Damit sollten Informationen über die Struktur des deterministischen und stochastischen Anteils des Prozesses gewonnen werden. Ebenfalls wurden die Annahmen dieser Schätzung geprüft. Dazu zählt das Vorliegen der Markov-Eigenschaft und das Verschwinden der Kramers-Moyal-Koeffizienten der Ordnung drei und höher.

Zur Markovanalyse sind zwei unterschiedliche Methoden verwendet worden. Sowohl das Prüfen der Gültigkeit der Chapman-Kolmogorov-Gleichung als auch das direkte Prüfen der Definition eines Markov-Prozesses im einfachen Fall $N = 3$ führten zum gleichen Ergebnis. Im Rahmen der Schätzfehler zeigen die Ergebnisse, dass auf den untersuchten Skalen die Markov-Eigenschaft erfüllt ist. Unter Zuhilfenahme eines Distanzmaßes für Wahrscheinlichkeitsdichtefunktionen konnte dieses Ergebnis auch quantifiziert werden. Eine weiterhin interessante Frage ist, speziell für Finanzdaten, ob die Markov-Eigenschaft auf sehr kleinen Skalen erfüllt ist, oder ob möglicherweise eine Art endliche Markovlänge wie beispielsweise bei Turbulenzdaten zu beobachten ist.

Die eigentliche Drift- und Diffusionsanalyse liefert ähnliche Ergebnisse, wie sie zuvor schon für andere Finanzdaten gemacht wurden (vgl. [15]). Eine lineare Drift- und eine quadratische Diffusionsstruktur scheint eine universelle Eigenschaft für logarithmierte Preisdifferenzen zu sein. Das Verschwinden des vierten Kramers-Moyal-Koeffizienten lässt darauf schließen, dass sich dieser Prozess, zumindest auf dieser untersuchten Zeitskala, mit Hilfe einer Langevin-Gleichung vollständig beschreiben lässt.

Ein GARCH(1,1)-Prozess, angepasst auf statistischen Eigenschaften der DAX-Daten, wurde ebenfalls den oben genannten Analysen unterzogen. Hier konnte gezeigt werden, welche Ergebnisse die Markovanalyse liefert, wenn die Eigenschaft nicht vorliegt. Im Contourplot ist deutlich zu sehen, wie sich die direkt geschätzte und die numerisch berechnete Seite der Chapman-Kolmogorov-Gleichung unterscheiden. Auch das verwendete Distanzmaß unterstützt diese Beobachtung. Die Drift- und Diffusionsanalyse des GARCH-Prozesses lieferte Koeffizienten die zwar von denen des DAX-Prozesses abweichen, jedoch in der gleichen Größenordnung liegen. Insgesamt ist zu beobachten, dass sich ein sehr einfacher GARCH-Prozess schon gut auf einen Finanzdatensatz anpassen lässt. Allerdings ist diese Aussage nur bei einer oberflächlichen, statistischen Analyse zulässig.

Kapitel 4

Die Dynamik zweier Aktienkurse

Wenn man sich unterschiedliche Aktienkurse auf großen Zeitskalen anschaut, kann man beobachten, dass die Trajektorien teilweise einen sehr ähnlichen Verlauf haben. In Abbildung 4.1 sind logarithmierte Preisdifferenzen vom DAX und von der Aktie der Commerzbank aufgetragen. Mit wenigen Ausnahmen gehen die Ausschläge in die gleiche Richtung. Fällt also ein Titel, z.B. wegen einer negativen Konjunkturmeldung, dann sinkt ein anderer Titel meist auch im Wert und analog gilt dies für einen Kursanstieg. Besonders bei Unternehmen der gleichen Branche ist diese Entwicklung häufig zu beobachten. Ein Maß für

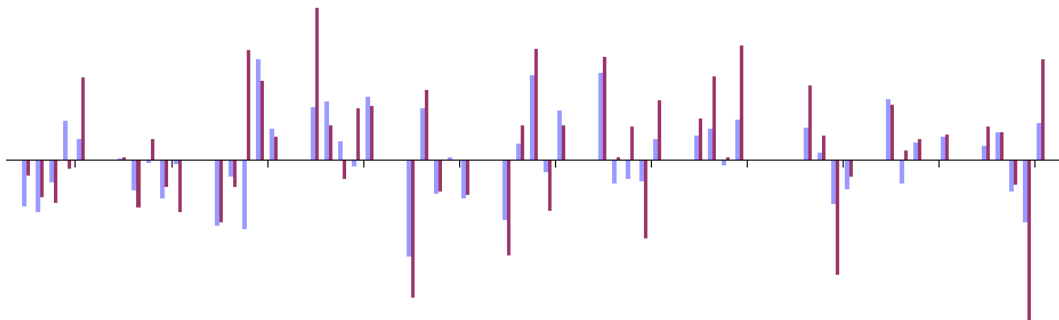


Abbildung 4.1: Logarithmierte Kursdifferenzen des DAX und der Commerzbankaktie. Es handelt sich um einen zufällig ausgewählten Zeitraum des Jahres 2003. Jeder Balken steht für einen Tageschlusskurs (Datenquelle [53]).

die Synchronität von zwei unterschiedlichen Zeitreihen ist die Kreuzkorrelation. In der folgenden Tabelle 4.1 sind einige Kreuzkorrelationskoeffizienten ρ_{ij} verschiedener Kurse aufgelistet. Als Datengrundlage wurden Tagesschlusskurse aus dem Jahr 2003 von [53] verwendet. Eine relativ starke Korrelation ist zwischen den Banken und dem DAX zu beobachten, während zwischen Bayer und der Commerzbank eine eher geringe positive Korrelation zu sehen ist. Generell ist es jedoch nicht verwunderlich, dass zwischen allen Werten eine positive Abhän-

gigkeit besteht, da alle großen Unternehmen einer Volkswirtschaft letztendlich von der Konjunktur abhängig sind.

Tabelle 4.1: Korrelationskoeffizientenmatrix ausgewählter Finanztitel. Bedeutung der Abkürzungen: CB = Commerzbank, DB = Deutsche Bank, BAY = Bayer, VW = Volkswagen.

ρ_{ij}	DAX	CB	DB	BAY	VW
DAX	1	0,6761	0,8552	0,5577	0,6167
CB		1	0,6485	0,3349	0,3946
DB			1	0,5330	0,5175
BAY				1	0,4430
VW					1

Mantegna und Stanley benutzen in [35] einen speziellen Graphen, genannt *indexed hierarchical tree (IHT)*, um die *Taxonomy* eines Aktienportfolios sichtbar zu machen. Die Anordnung und Verlinkung der Objekte folgt dem *Kruskal'schen Algorithmus*, wie er ebenfalls in [35] beschrieben ist. Die Distanz

$$d_{ij} = \sqrt{2(1 - \rho_{ij})} \quad (4.1)$$

ist dabei gerade so definiert, dass sie verschwindet, wenn eine vollständige positive Korrelation $\rho_{ij} = 1$ vorliegt.

Tabelle 4.2: Distanzen ausgewählter Finanztitel.

d_{ij}	DAX	CB	DB	BAY	VW
DAX	0	0,8048	0,5382	0,9406	0,8755
CB		0	0,8385	1,1533	1,1004
DB			0	0,9665	0,9823
BAY				0	1,06
VW					0

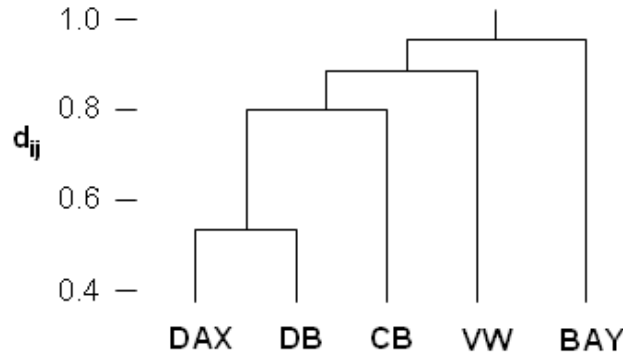


Abbildung 4.2: In diesem Indexed hierarchical tree sind die Distanzen aus Tabelle 4.2 für ausgewählte deutsche Finanztitel aufgetragen. Die kerzenförmigen Verbindungen, die nach dem Kruskal'schen Algorithmus gezogen wurden, machen die Korrelationsstruktur des Portfolios sichtbar.

4.1 Das Modell

Wie bereits erwähnt ist es üblich, die Dynamik eines Aktienkurses mit Hilfe von stochastischen Differenzialgleichungen zu beschreiben. Die allgemeine Langevin-Gleichung beschreibt die Entwicklung durch einen deterministischen Teil (Drift) und einen stochastischen Teil (Diffusion).

Gemäß Definition (2.1) ist der Kurs der Aktie i zum Zeitpunkt t weiterhin gegeben durch

$$s_i(t). \quad (4.2)$$

Logarithmiert ergibt sich

$$y_i(t) = \log s_i(t). \quad (4.3)$$

Man betrachte nun Kursdifferenzen mit einem zeitlichen Abstand von Δt :

$$x_i(t, \Delta t) = y_i(t + \Delta t) - y_i(t). \quad (4.4)$$

Zur Bestimmung der Drift- und Diffusionskoeffizienten wird gemäß Gleichung (1.32) folgender Schätzer verwendet:

$$D_{i,\Delta t,\tau}^{(n)}(x_{i,\Delta t}) = \lim_{\tau \rightarrow 0} \frac{1}{\tau} \frac{\langle [X(t + \tau, \Delta t) - X(t, \Delta t)]^n | X(t, \Delta t) = x_{i,\Delta t} \rangle}{n!}. \quad (4.5)$$

Mit der Beziehung

$$h_{i,\Delta t,\tau} = D_{i,\Delta t,\tau}^{(1)} \quad (4.6)$$

$$g_{i,\Delta t,\tau} = \sqrt{D_{i,\Delta t,\tau}^{(2)}} \quad (4.7)$$

folgt ein Prozess für die zeitliche Entwicklung von Kursdifferenzen (zeitliche Differenz bei Δt) auf der Skala τ für die Aktie i :

$$\dot{X}_{i,\Delta t,\tau}(t) = h_{i,\Delta t,\tau} + g_{i,\Delta t,\tau}\Gamma(t). \quad (4.8)$$

Für den Zweiaktienfall ($i = 1, 2$) betrachtet man einen zweidimensionalen Prozess. Im Folgenden wird auf die Indizes Δt und τ verzichtet.

Im vorherigen Kapitel wurde auf die Synchronität der Aktienkursentwicklungen von zwei oder mehr Kursen verwiesen. Will man also N unterschiedliche Wertpapiere beschreiben, deren Kursentwicklungen voneinander abhängen, würde ein N -dimensionales, gekoppeltes, stochastisches Differenzialgleichungssystem erster Ordnung in Frage kommen. In dieser Arbeit beschränkt man sich auf $N = 2$ Aktien. Gemäß Gleichung (1.21) würde sich ein zweidimensionales System folgendermaßen darstellen:

$$\dot{X}_i(t) = h_i(\mathbf{X}, t) + \sum_j g_{ij}(\mathbf{X}, t)\Gamma_j(t), \quad (4.9)$$

mit $i = 1, 2$ und $\mathbf{X} \in \mathbb{R}^2$. Es werden hierbei zunächst unkorrelierte Rauschkräfte $\langle \Gamma_1 \Gamma_2 \rangle = 0$ mit Einheitsvarianz angenommen. In Vektordarstellung folgt:

$$\begin{pmatrix} \dot{X}_1 \\ \dot{X}_2 \end{pmatrix} = \begin{pmatrix} h_1(\mathbf{X}) \\ h_2(\mathbf{X}) \end{pmatrix} + G \cdot \begin{pmatrix} \Gamma_1 \\ \Gamma_2 \end{pmatrix}, \quad (4.10)$$

mit der Diffusionsmatrix:

$$G = \begin{pmatrix} g_{11}(\mathbf{X}) & g_{12}(\mathbf{X}) \\ g_{21}(\mathbf{X}) & g_{22}(\mathbf{X}) \end{pmatrix}. \quad (4.11)$$

Die korrespondierende, zweidimensionale Fokker-Planck-Gleichung hat folgende Form:

$$\frac{\partial}{\partial t} p(\mathbf{x}, t) = \left[- \sum_{i=1}^2 \frac{\partial}{\partial x_i} D_i^{(1)}(\mathbf{x}, t) + \frac{1}{2} \sum_{i,j=1}^2 \frac{\partial^2}{\partial x_i \partial x_j} D_{ij}^{(2)}(\mathbf{x}, t) \right] p(\mathbf{x}, t). \quad (4.12)$$

Der Driftvektor und die Diffusionsmatrix lassen sich aus den Aktienzeitreihen schätzen. Dazu betrachtet man die bedingten Momente:

$$D_i^{(1)}(\mathbf{x}, t) = \frac{1}{\tau} \langle x_i(t + \tau) - x_i(t) | \mathbf{x}(t) = \mathbf{x} \rangle, \quad (4.13)$$

$$D_{ij}^{(2)}(\mathbf{x}, t) = \frac{1}{2\tau} \langle [x_i(t + \tau) - x_i(t)] [x_j(t + \tau) - x_j(t)] | \mathbf{x}(t) = \mathbf{x} \rangle. \quad (4.14)$$

Dabei sei angemerkt, dass es sich hierbei für endliches τ und verschwindende Mittelwerte der Preisdifferenzen auch um eine Schätzung der bedingten

Kovarianz gemäß Gleichung (1.14) handelt. Über die Zusammenhänge gemäß (1.39):

$$D_i^{(1)} = h_i \quad (4.15)$$

$$(\sqrt{D^{(2)}})_{ij} = g_{ij} \quad (4.16)$$

lässt sich so aus den empirischen Daten auf die Struktur der Drift- und Diffusionsterme der Langevin-Gleichungen (4.9) schließen. Eine Voraussetzung für die Berechnung der bedingten Momente gemäß 4.14 ist die Fähigkeit, die Inkremente der Aktienpreisdifferenzen ¹ $x_i(t + \tau) - x_i(t)$ und $x_j(t + \tau) - x_j(t)$ zu gleichen Zeiten t abtasten zu können. Bei zeitlich inhomogenen, empirischen Datensätzen liegt somit ein Problem vor. Dieses wird im nächsten Kapitel ausführlich diskutiert.

4.2 Der Datensatz

Die hier untersuchten Aktienkursdaten von Volkswagen und Bayer sind Aufzeichnungen aus den Jahren 1993 bis 2003. Diese beiden Titel weisen im genannten Zeitraum die größte Handelsaktivität im DAX und damit die meisten Datenpunkte auf. Die Aufzeichnung der Daten erfolgte zeitlich inhomogen, d.h. die Datenpunkte liegen in nicht-äquidistanten Zeitabständen vor. Außerdem liegen zwischen den Aufzeichnungsperioden längere Handelspausen in Form von Nächten und Wochenenden vor (vgl. Abbildung 4.3). Die Auflösung der Zeitstempel liegt bei einer Millisekunde. Um die Datensätze gleichzeitig abtastbar zu machen (vgl. Abschnitt 4.1), müssen die Daten zunächst transformiert werden.

4.2.1 Datenpräparation

Eine zielführende Auswertung ohne Datenpräparation wäre möglich, wenn es ausreichend viele übereinstimmende Zeitstempel in den beiden Datensätzen gäbe und somit eine große Anzahl an gleichzeitig abfragbaren Datenpunkten existierte. In Abbildung 4.4 kann man erkennen, dass auf den einzelnen Zeitskalen nur eine sehr begrenzte Anzahl an Datenpunkten zur Verfügung steht.

¹In [50] wird darauf hingewiesen, dass sich, je nach Definition der Inkremente, unterschiedliche Ergebnisse für skalenabhängige Analysen von stochastischen Prozessen ergeben können. Für die Drift- und Diffusionsanalyse konnten in diesem Fall keine abweichenden Ergebnisse mit den sogenannten zentrierten Inkrementen gemacht werden.

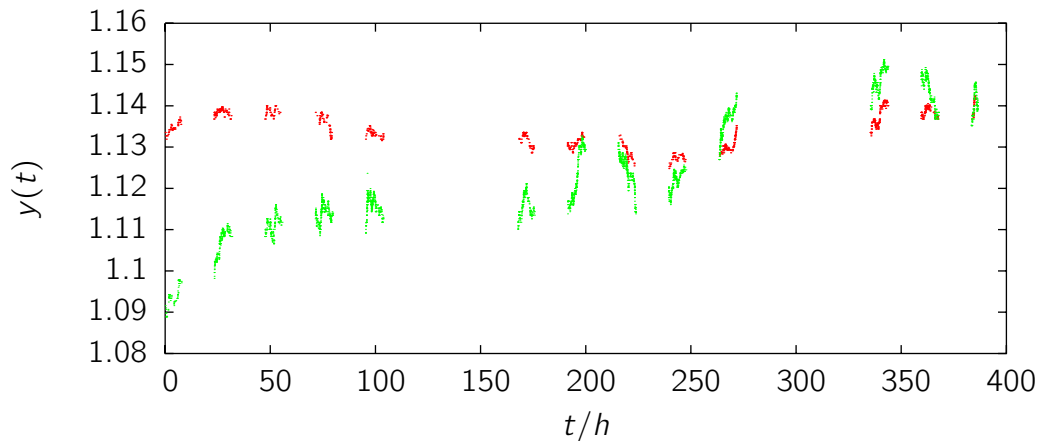


Abbildung 4.3: Exemplarisch sind hier die logarithmierten Aktienkurse von Bayer (rot) und VW (grün) dargestellt. Es handelt sich um die ersten 13 Tage aus dem Datensatz. Gut zu erkennen sind die Pausen zwischen den Handelszeiten.

Außerdem stimmen die Zeitstempel der verschiedenen Datensätze auf den entsprechenden Skalen nur in den seltensten Fällen überein. Also trifft diese Bedingung für die vorliegenden Datensätze von Bayer und VW nicht zu. Es ist somit notwendig, die Daten so zu transformieren, dass sie zeitlich äquidistant bzw. zu gleichen Zeiten abtastbar sind. Um eine vernünftige Datengrundlage für die Drift- und Diffusionschätzung zu bekommen, fordert man eine Anzahl an Datenpunkten in der Größenordnung 10^6 . Die Daten werden später klassifiziert in ungefähr 50×50 Bins. Um ein valides Ergebnis zu erhalten, sollte jedes Bin eine gewisse Anzahl an Datenpunkten in der Größenordnung 10^3 oder mehr aufweisen. Für nähere Ausführungen siehe [28]. Die Breite des maximalen Zeitintervalls bzw. der maximale Abstand zwischen den Stützstellen lässt sich approximieren. Bei neun Handelsstunden pro Tag und ca. 220 Handelstagen pro Jahr ergibt sich die Gesamthandelszeit über 10 Jahre in Sekunden zu:

$$9 \frac{h}{d} \cdot 220 \frac{d}{y} \cdot 10 y \cdot 3.600 = 71.280.000 \text{ s.}$$

Das maximale Zeitintervall sollte demnach bei ungefähr einer Minute liegen. Das Minimum ist zunächst einmal nur durch die Auflösung beschränkt. Jedoch würde die sich ergebene Datengröße bei zu kleinen Intervallen ebenfalls ein Problem darstellen. Hier wird das minimale Intervall auf 15 Sekunden festgelegt.

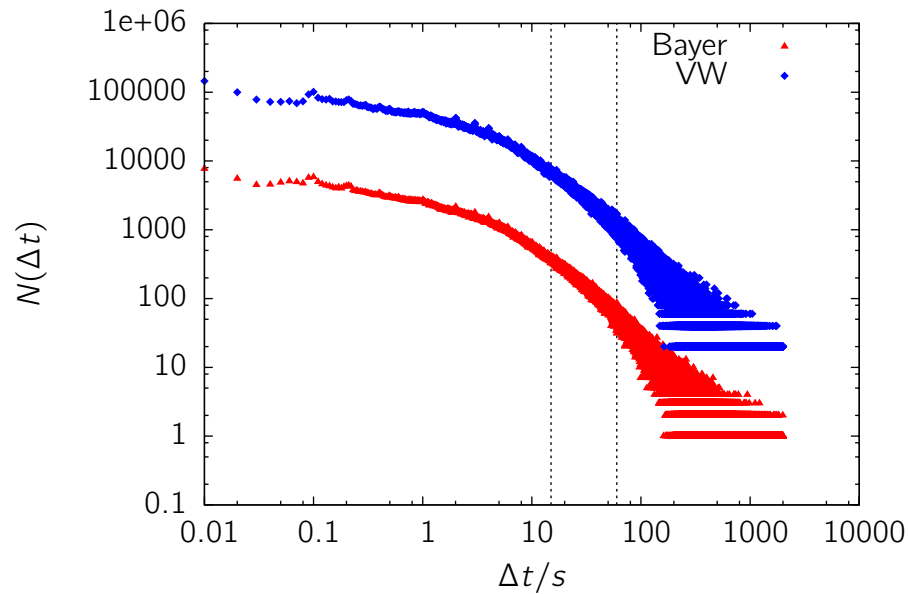


Abbildung 4.4: Anzahl der Preisdifferenzen gegen die Zeitskalen für den Bayer- (unten) und VW-Datensatz (oben) in seiner ursprünglichen Form. Die VW-Kurve ist aus Gründen der Übersichtlichkeit vertikal nach oben verschoben. Die gestrichelten Linien entsprechen den Zeitsprüngen $\Delta t_1 = 15\text{s}$ und $\Delta t_2 = 60\text{s}$.

Zusammenfassend werden also zwei Forderungen an die Datenpräparation gestellt:

1. Die Zeitstempel der einzelnen Kursnotierungen der beiden Datensätze müssen übereinstimmen.
2. Es müssen ausreichend viele Datenpunkte (Größenordnung 10^6) vorhanden sein oder generiert werden.

Als Verfahren, die diese Forderungen erfüllen, kommen verschiedene Arten von Interpolationen in Betracht. Dazu zählt das Einführen einer *Stufenfunktion*, die *lineare Interpolation* sowie eine einfache Form der *Spline-Interpolation*. Zunächst muss jedoch festgelegt werden, wie der zeitliche Abstand der Stützstellen $\delta_i - \delta_{i-1}$ und der Interpolationspunkte $t_i - t_{i-1}$ zu wählen ist. Bei der Wahl der Stützstellen gibt es zwei zweckmäßige Möglichkeiten:

1. Man greift auf den gesamten ursprünglichen Datensatz zurück.
2. Zunächst wird eine Datenreduzierung durchgeführt. Man verwendet anschließend weniger Datenpunkte, die in einem gewissen zeitlichen Abstand zueinander liegen (5 bis 30 Minuten ², vgl. Abbildung 4.5).

²Zeitabstände dieser Größenordnung werden im Zusammenhang mit Volatilitätsschätzungen

Die zweite Alternative ergibt sich durch die Überlegung, dass Kursveränderungen auf einer größeren Skala eher in einen wirtschaftlichen Kontext gebracht werden können als Returns auf sehr kleinen Zeitskalen.

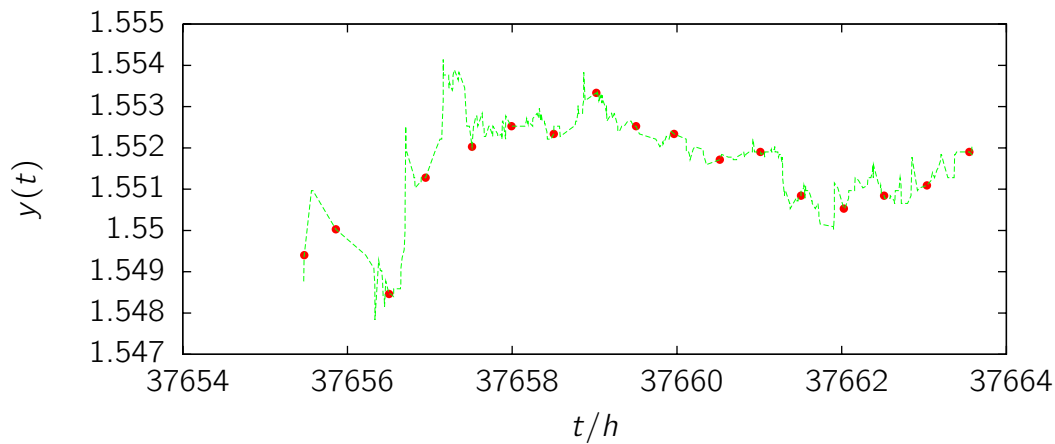


Abbildung 4.5: Bei der Datenreduzierung werden von dem ursprünglichen Datensatz (durchgezogene Linie) nur bestimmte Datenpunkte in einem möglichst gleichmäßigen Abstand herausgefiltert (Punkte). Diese dienen als Stützstellen (\tilde{y}_i, δ_i) für die Interpolation. Hier ist ein Handelstag aufgetragen, und die zeitlichen Abstände im reduzierten Datensatz betragen etwa 30 Minuten.

Interpolationsverfahren

1. Die einfachste und naheliegendste Lösung ist das Einführen einer **Stufenfunktion**. Dabei wird der aktuelle Kurswert solange als konstant angenommen bis ein neuer Kurssprung einsetzt. Auch aus dem ökonomischen Blickwinkel scheint diese Methode am naheliegendsten, denn es gilt solange der aktuelle Preis für eine Aktie bis ein neuer Preis festgelegt wird.
2. Die **lineare Interpolation** wird angewandt, um aus den Datenpunkten des ursprünglichen (oder reduzierten) Datensatzes (\tilde{y}_i, δ_i) mit $i = 1, 2, \dots, N$ einen neuen, äquidistanten Datensatz (y_j, t_j) zu generieren. Es gilt:

$$(y_1, t_1) := (\tilde{y}_1, \delta_1).$$

Die weiteren Interpolationspunkte werden nach folgender Vorschrift berechnet:

$$y_j = \tilde{y}_{i-1} + \frac{\tilde{y}_i - \tilde{y}_{i-1}}{\delta_i - \delta_{i-1}}(t_j - \delta_{i-1}),$$

aus Hochfrequenzdaten üblicherweise verwendet (vgl. [1] und [3]).

mit $j > 1$ falls $t_j \in [\delta_{i-1}, \delta_i]$ mit $\delta_i - \delta_{i-1} < T_n$ wobei T_n die nicht-Handelszeit darstellt. Das bedeutet, dass nur innerhalb der Handelszeit interpoliert wird.

3. Bei der **Spline-Interpolation** handelt es sich um ein Verfahren bei dem ein Datensatz stückweise mit Hilfe eines Polynoms N -ten Grades interpoliert wird. Bei der hier verwendeten Vereinfachung wird die Forderung nach Stetigkeit und Differenzierbarkeit an den Rändern vernachlässigt. Es wird ein kubisches Polynom ($N = 3$) verwendet:

$$y_j = at_j^3 + bt_j^2 + ct_j + d.$$

Die benötigten Stützstellen $(\tilde{y}_{i-1}, \delta_{i-1})$, (\tilde{y}_i, δ_i) , $(\tilde{y}_{i+1}, \delta_{i+1})$ und $(\tilde{y}_{i+2}, \delta_{i+2})$ mit $\delta_{i-1} < \delta_i < \delta_{i+1} < \delta_{i+2}$ o. B. d. A. müssen dabei wieder innerhalb der Handelszeit liegen, und es wird gefordert, dass $t_j \in [\delta_i, \delta_{i+1}]$ gilt. Die Parameter a , b , c und d werden mit Hilfe eines linearen Gleichungssystems bestimmt (vgl. Anhang A.2).

Durch die Interpolation werden viele künstliche Datenpunkte erzeugt, die Auswirkungen auf die Statistik haben. Welche Auswirkungen die Datenpräparation hat, wird nun diskutiert.

Vergleich der Methoden

Im Folgenden soll ein Vergleich der Interpolationsmethoden stattfinden. Ein erstes Vergleichsmerkmal stellt dabei die Form der pdf vor und nach der Datentransformation dar. Ein Vergleich mit dem ursprünglichen Datensatz von Bayer und VW ist jedoch nicht konsistent möglich, da diese im mittleren Zeitskalenbereich über zu wenig Datenpunkte verfügen. So ist in Abbildung 4.4 zu sehen, dass im Beispiel Bayer für $\Delta t_1 = 15$ s nur noch etwa $N_{\Delta t_1} \approx 300$ Preisdifferenzen und für $\Delta t_2 = 60$ s sogar nur noch $N_{\Delta t_2} \approx 50$ Differenzen vorliegen. Es wäre nicht sinnvoll, bei dieser Datenlage einen Vergleich der Wahrscheinlichkeitsdichtefunktionen zu machen. Stattdessen kann man sich mit einem Vergleich mit den return-pdf's des DAX-Datensatzes behelfen, der in diesem Skalenbereich über ausreichend Datenpunkte verfügt.

Dieser Vergleich findet unter der Prämisse statt, dass sich die Inkrementverteilungen der Einzelaktien nicht maßgeblich von denen des Index unterscheiden. Zum einen weisen die genannten Einzelaktien einen relativ hohen positiven Korrelationskoeffizienten zum Index auf (vgl. Tabelle 4.1), und zum anderen wird hier ein Gesamtzeitraum betrachtet bei dem sich kurzweilige Extremvolatilitäten in den Einzeltiteln relativieren. So sollten sich zumindest sehr starke Abweichungen leicht identifizieren lassen, wie es für die im Folgenden diskutierte Stufenfunktion der Fall ist. Im Zuge der Differenzenbildung (vgl. Definition (2.1)) und

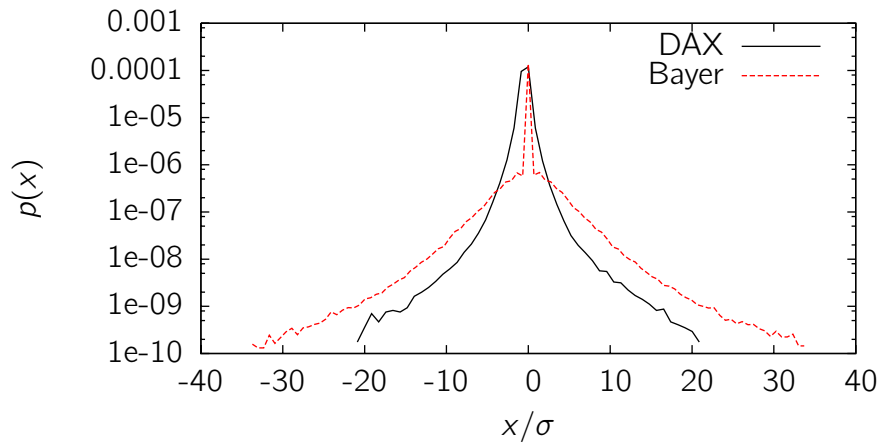


Abbildung 4.6: Vergleich von Wahrscheinlichkeitsdichteverteilungen der Renditen vom DAX und der Bayer-Aktie. Die pdf der DAX-Renditen im Abstand von $\Delta t = 15s$ (durchgezogene Linie) unterscheidet sich stark von dem mit Stufenfunktion interpolierten Datensatz von Bayer für einen zeitlichen Abstand von 15s zwischen den Stützstellen. Man beachte vor allem den Peak der Bayer-Renditen bei $x = 0$ (logarithmische Darstellung).

der vielen konstanten Abschnitte im interpolierten Datensatz erzeugt man sehr viele Renditen $x(\Delta t) = 0$ (vgl. Abbildung 4.6). Die Verteilung der Renditen hat sich in diesem Fall so stark verändert, dass diese Art der Datentransformation in diesem Fall nicht geeignet zu sein scheint, um einen äquidistanten Datensatz mit ungefähr gleichbleibenden statistischen Eigenschaften zu generieren. Aus diesem Grund wird diese Methode nicht weiter betrachtet.

Ein weiterer Ansatz um eine zielführende Datentransformation durchzuführen, stellt die zuvor beschriebene Datenreduktion dar. In Abbildung 4.8 sind die resultierenden pdf's gegen die des DAX aufgetragen. Sowohl für die lineare Interpolation, als auch für die Spline-Interpolation führt die Datenfilterung dazu, dass mehr große Inkremente erzeugt werden und so die Verteilung breitflügeliger wird. Bei der Spline-Interpolation ist dieser Effekt noch deutlicher. Das liegt an der Tatsache, dass die Spline-Interpolation im Zweifel immer zu extremeren Interpolationsergebnissen zwischen den Stützstellen führt als die lineare, aber niemals vice versa (vgl. Abbildung 4.7). Daraus folgt der Schluss, sowohl die Datenreduktion als auch die Spline-Interpolation für die weitere Betrachtung auszuschließen.

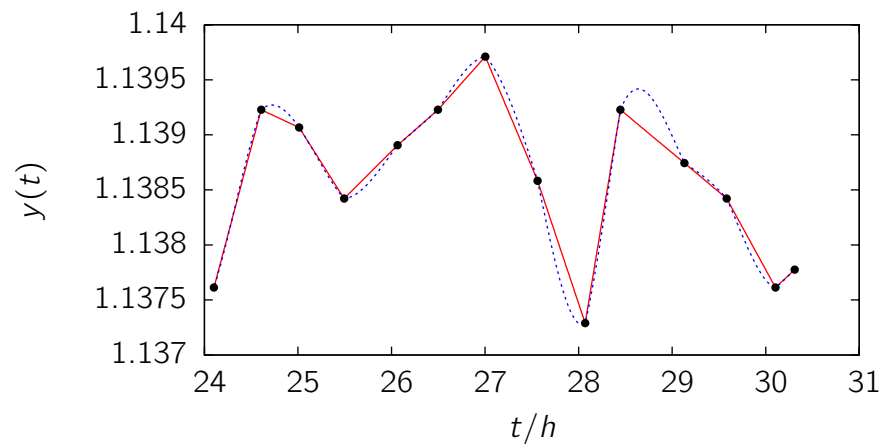


Abbildung 4.7: Vergleich der linearen Interpolation (durchgezogene, rote Kurve) mit der Spline-Interpolation (gestrichelte, blaue Kurve). Exemplarisch ist der zweite Handelstag im Datensatz der Bayer-Aktie aufgetragen. Im Zuge der Datenreduktion liegen die Stützstellen im Abstand von etwa 30 Minuten (schwarze Punkte), die Interpolationspunkte liegen in 30 Sekunden Abständen. Deutlich sind leichte Abweichungen zu erkennen.

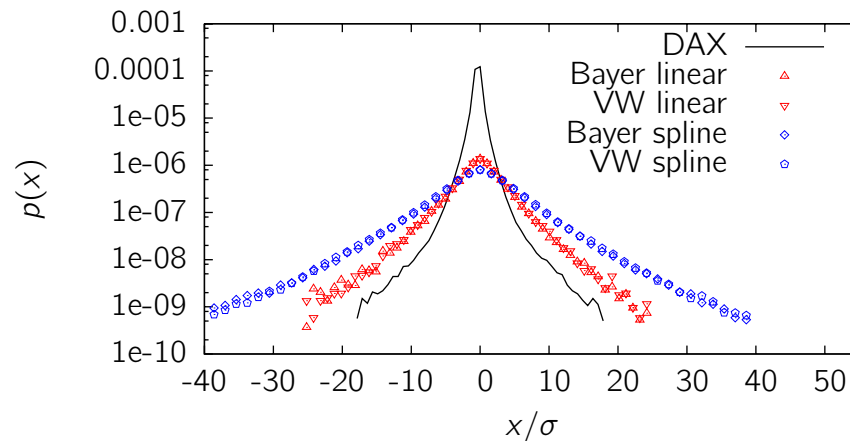


Abbildung 4.8: Wahrscheinlichkeitsdichtefunktionen der Renditen aus den Datensätzen von Bayer und VW, die durch lineare Interpolation und Spline-Interpolation erzeugt wurden. Zuvor wurde eine Datenreduktion durchgeführt, so dass die Stützstellen mit einem Abstand von 30 Minuten zueinander liegen. Die Interpolationspunkte wurden mit einem Abstand von 60 Sekunden gewählt.

Demnach ergibt sich aus den oben aufgezeigten Möglichkeiten nur noch die Alternative der linearen Interpolation ohne vorher den Datensatz künstlich zu reduzieren. In Abbildung 4.9 wird durch den erneuten Vergleich der Wahrscheinlichkeitsdichteverteilungen auf kleinster Skala deutlich, dass diese Methode die geringsten Abweichungen erzeugt. Zwar werden die größeren Kursdifferenzen auch hier etwas überschätzt, aber die pdf bleibt weitestgehend erhalten. Im

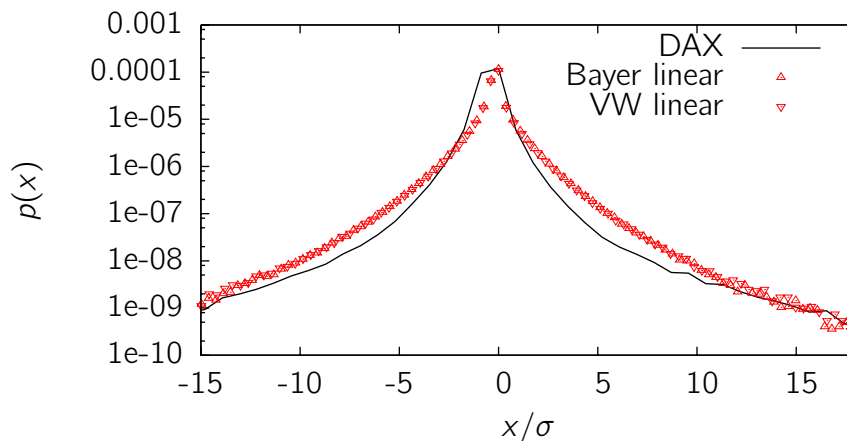


Abbildung 4.9: Wahrscheinlichkeitsdichtefunktionen der Renditen aus den Datensätzen von Bayer und VW, die durch lineare Interpolation erzeugt wurden. Die Interpolationspunkte wurden mit einem Abstand von 15 Sekunden gewählt. Es wurde keine Datenreduktion durchgeführt.

Folgenden wird dieser interpolierte Datensatz mit einem zeitlichen Abstand von 15 s zwischen den Interpolationsstellen verwendet.

4.3 Markovanalyse der interpolierten Datensätze

Für die transformierten Datensätze wird nun eine Markovanalyse durchgeführt. Es soll dabei gezeigt werden, dass trotz des Einsatzes der linearen Interpolation die Markoveigenschaft erhalten geblieben ist. Dabei wird erneut auf die zwei bereits vorgestellten Analysemethoden zurückgegriffen (vgl. Kapitel 3.1.1). Auch an dieser Stelle zeigt sich eine gute Übereinstimmung der Wahrscheinlichkeitsdichten. Deshalb wird davon ausgegangen, dass die Preisinkremente einem Markovprozess folgen.

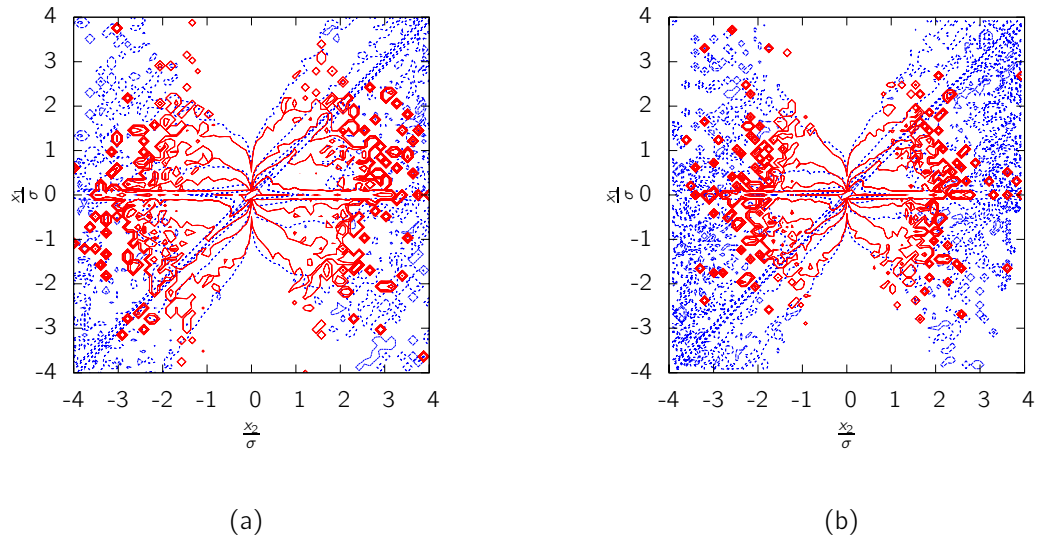


Abbildung 4.10: Bedingte Wahrscheinlichkeitsdichten von Bayer (links) und VW (rechts). Es sind die Höhenlinien von $p(x_1, \tau_1 | x_2, \tau_2)$ (blaue, gestrichelte Linien) und $p(x_1, \tau_1 | x_2, \tau_2; x_3 = 0, \tau_3)$ (rote, durchgezogene Linien) mit $\tau_1 = 30$ s, $\tau_2 = 45$ s und $\tau_3 = 60$ s aufgetragen. Die Normierung mit der Standardabweichung σ bezieht sich auf die kleinste Skala τ_1 .

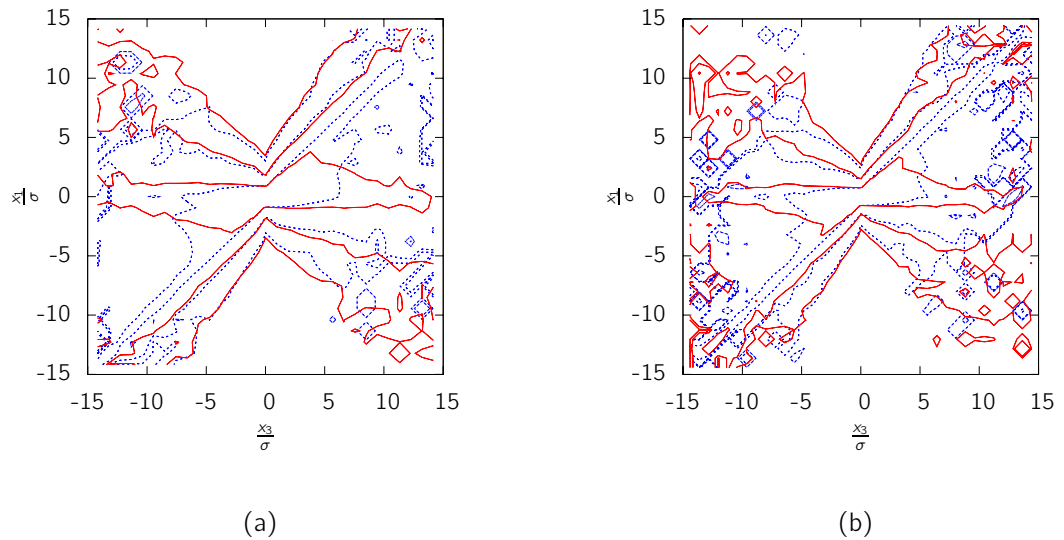


Abbildung 4.11: Chapman-Kolmogorov-Gleichung mit $\tau_1 = 30$ s, $\tau_2 = 45$ s und $\tau_3 = 60$ s für die Aktienkursdifferenzen von Bayer (links) und VW (rechts).

4.4 Schätzen der Drift- und Diffusionskomponenten

Im Folgenden werden die Koeffizienten der Fokker-Planck-Gleichung aus den bedingten Momenten gemäß Gleichung (4.13) und (4.14) geschätzt. Auch an dieser Stelle sei noch einmal darauf hingewiesen, dass lediglich eine Zeitskala betrachtet wird. So wurde erneut $\Delta t = 15$ s und $\tau = 30$ s gewählt.

4.4.1 Der Driftvektor

Wie in Abbildung 4.12 ersichtlich, konnten die Komponenten des Driftvektors $D_B^{(1)}(x_B)$ und $D_V^{(1)}(x_V)$ linear angefitzt werden:

$$I_B(x_B) = a_B x_B \quad (4.17)$$

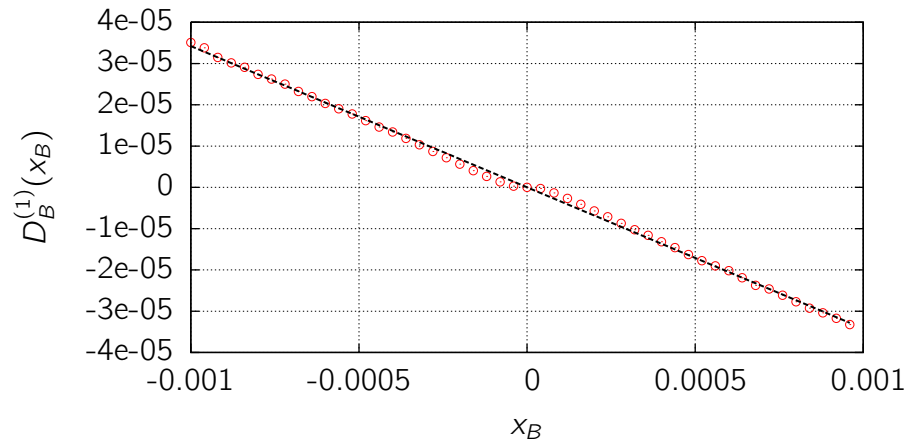
$$I_V(x_V) = a_V x_V. \quad (4.18)$$

Dabei existiert zwar ein endlicher Offset, dieser ist jedoch um einige Größenordnungen kleiner und wird daher vernachlässigt. Die Parameter ergeben sich zu:

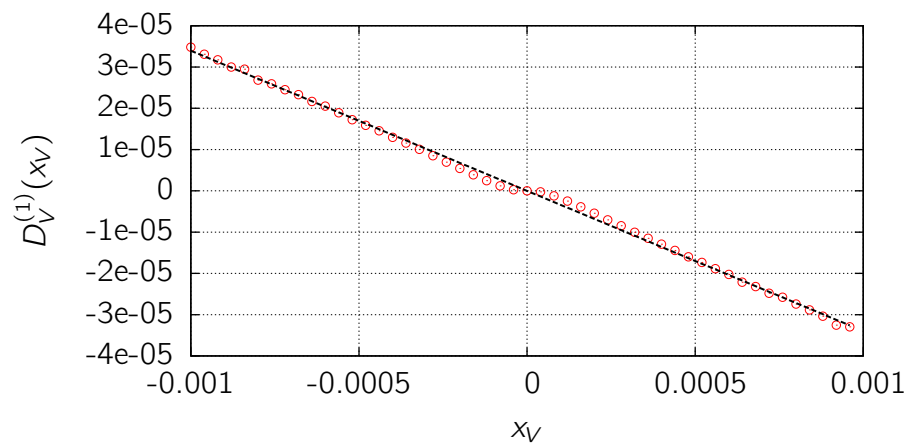
$$a_B = -0.0342 \pm 0.0002 \quad (4.19)$$

$$a_V = -0.0340 \pm 0.0002. \quad (4.20)$$

Die Drifts sind unabhängig von der Bewegung der jeweils anderen Aktie. Es liegt keine deterministische Kopplung der Preisdifferenzen von Bayer und VW vor, oder diese ist zu klein, um sie hier aufzulösen. Deshalb sind die Driftkomponenten in Abbildung 4.12(a) und 4.12(b) nur gegen die abhängige Variable aufgetragen.



(a) Bayer



(b) VW

Abbildung 4.12: Driftkoeffizienten der logarithmierten Aktienkursrenditen bei $\tau = 30$ s. Da näherungsweise keine Abhängigkeit zur jeweils anderen Aktie besteht, wurde die Fläche auf die Achse der abhängigen Variable projiziert. Die gestrichelten Geraden stellen die linearen Fitfunktionen dar.

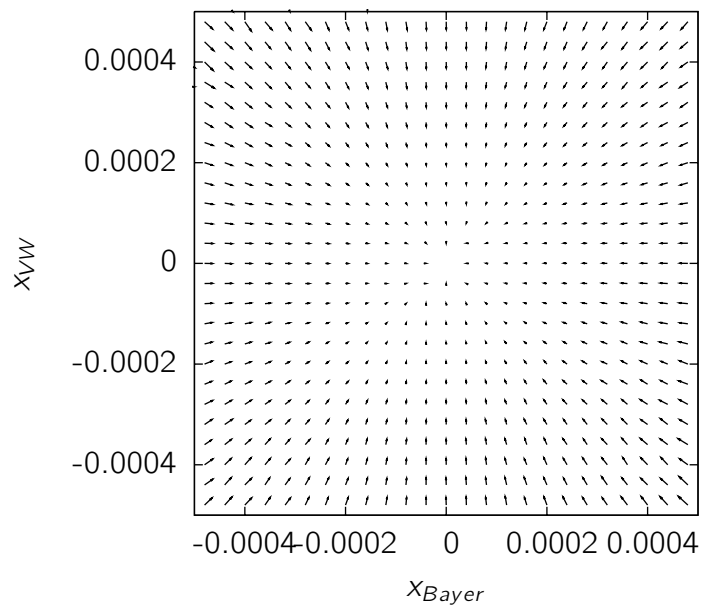


Abbildung 4.13: Deterministische Dynamik von Bayer und Volkswagen in einem Flussdiagramm. Aufgetragen sind die Driftvektoren $\mathbf{D}^{(1)}(\mathbf{x}) = (D_B^{(1)}(x_B, x_V), D_V^{(1)}(x_B, x_V))^T$ für die Auslenkungen der Preisdifferenzen x_B und x_V .

4.4.2 Die Diffusionsmatrix

Die Diagonalelemente der Diffusionsmatrix $D_{B,B}^{(2)}$ und $D_{V,V}^{(2)}$ (vgl. Abbildung 4.14) lassen sich durch Parabeln der Form:

$$p_B(x_B) = \alpha_0 + \alpha x_B^2, \quad (4.21)$$

$$p_V(x_V) = \beta_0 + \beta x_V^2 \quad (4.22)$$

anfitten. Auch hier existieren endliche Offsets α_0 und β_0 ; diese liegen in der Größenordnung 10^{-8} , und die angepassten Werte sind mit großen Fehlern behaftet. Die übrigen Parameter ergeben sich zu

$$\alpha = 0.0216 \pm 0.0002, \quad (4.23)$$

$$\beta = 0.0216 \pm 0.0002. \quad (4.24)$$

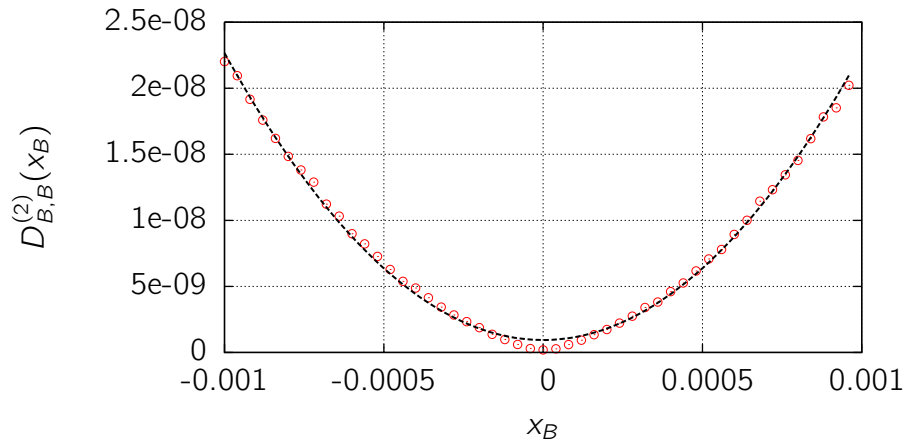
Der Mischterm $D_{B,V}^{(2)} = D_{V,B}^{(2)}$ erinnert an eine Sattelfunktion der Form $s(x, y) = x^2 - y^2$, welche aber im Winkel von $\phi_0 = 45^\circ$ um die z-Achse gedreht werden muss. Dies lässt sich mittels einer Rotationsmatrix sehr einfach durchführen (siehe Anhang A.1). Es folgt eine Sattelfunktion gemäß:

$$f(x_B, x_V) = \gamma_0 + \gamma x_B x_V. \quad (4.25)$$

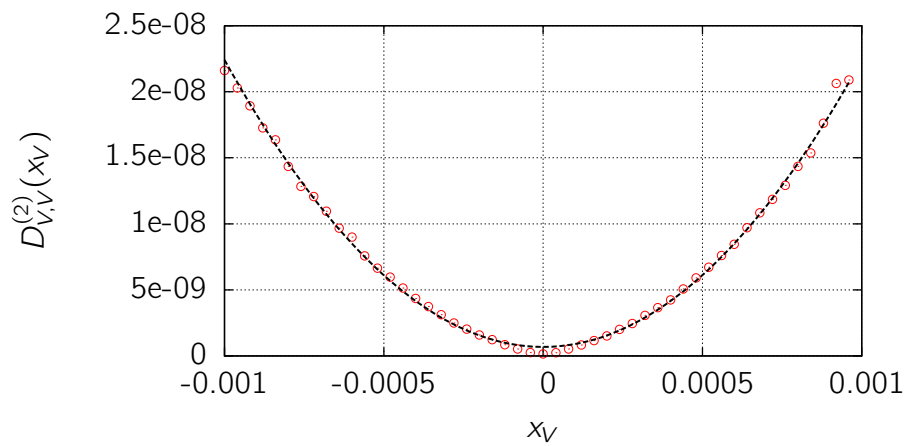
Das Anpassen der Sattelfunktion (4.25) hat ein

$$\gamma = 0.0175 \pm 0.0001 \quad (4.26)$$

ergeben (vgl. Abbildung 4.15). Der Offset γ_0 ist um viele Größenordnungen kleiner und wird auch hier vernachlässigt.



(a) Bayer



(b) VW

Abbildung 4.14: Diffusionskoeffizient der logarithmierten Aktienkursrenditen bei $\tau = 30$ s. Da näherungsweise keine Abhängigkeit zur jeweils anderen Aktie besteht, wurde die Fläche auf die Achse der abhängigen Variable projiziert. Die gestrichelten Kurven stellen die quadratischen Fitfunktionen dar.

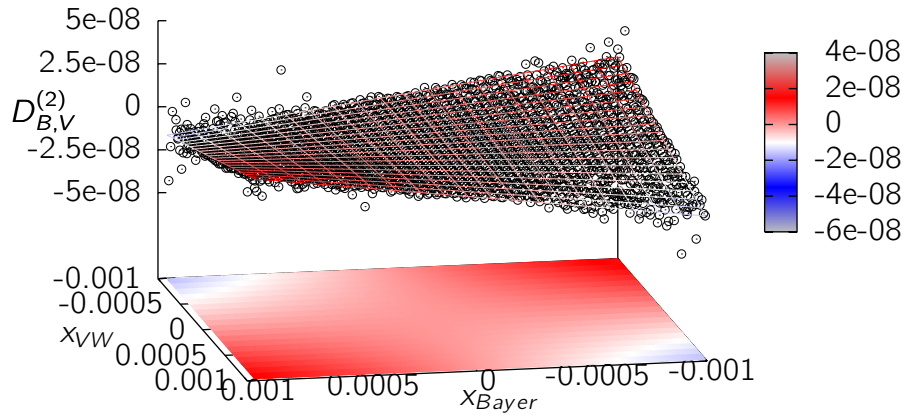


Abbildung 4.15: Der Mischterm der Diffusionsmatrix $D_{B,V}^{(2)} = D_{V,B}^{(2)}$ gegen die Preisdifferenzen von Bayer und VW. Das gestrichelte Gitter repräsentiert die angepasste Funktion 4.25.

Die Diffusionmatrix der zweidimensionalen Fokker-Planck-Gleichung lautet demnach wie folgt:

$$D = \begin{pmatrix} \alpha_0 + \alpha x_B^2 & \gamma_0 + \gamma x_B x_V \\ \gamma_0 + \gamma x_B x_V & \beta_0 + \beta x_V^2 \end{pmatrix}. \quad (4.27)$$

Die Diffusionsmatrix ist eine positiv definite ($\det D_{ij} > 0, \forall \mathbf{x}$), symmetrische Matrix und hat ausschließlich reelle Komponenten. Für die Berechnung der Eigenwerte und Eigenvektoren werden im Folgenden die Vereinfachungen:

$$\alpha_0 = \beta_0 = \gamma_0 = 0$$

und

$$\alpha = \beta$$

eingeführt. Es ergibt sich die vereinfachte Matrix:

$$\tilde{D} = \begin{pmatrix} \alpha x_B^2 & \gamma x_B x_V \\ \gamma x_B x_V & \alpha x_V^2 \end{pmatrix}. \quad (4.28)$$

Diese Matrix besitzt die reellen Eigenwerte:

$$\lambda_{1,2} = \frac{1}{2} (\alpha (x_B^2 + x_V^2) \pm \nu), \quad (4.29)$$

wobei

$$\nu = \sqrt{4\gamma^2 x_B^2 x_V^2 + \alpha^2 (x_B^2 - x_V^2)^2}.$$

Die Eigenvektoren zu den Eigenwerten stehen orthogonal zueinander:

$$\xi_{1,2} = \begin{pmatrix} \frac{\alpha(x_B^2 - x_V^2) \pm \nu}{2\gamma x_B x_V} \\ 1 \end{pmatrix}, \quad (4.30)$$

Für verschwindende Erwartungswerte der Preisdifferenzen kann die Diffusionsmatrix D als bedingte Kovarianzmatrix der Form

$$C = \begin{pmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{21} & \sigma_2^2 \end{pmatrix} \quad (4.31)$$

aufgefasst werden. Kovarianzmatrizen spielen in vielen Anwendungen eine Rolle. Beispiele dazu sind in Kapitel 5 angegeben.

4.4.3 Die Koeffizienten vierter Ordnung

Wie auch für die DAX-Daten beobachtet, liegen die Koeffizienten vierter Ordnung der Kramers-Moyal-Entwicklung um Größenordnungen unter den Diffusionstermen. Es wird also erneut angenommen, dass sich die Dynamiken mit Hilfe der ersten beiden Momente vollständig beschreiben lassen.

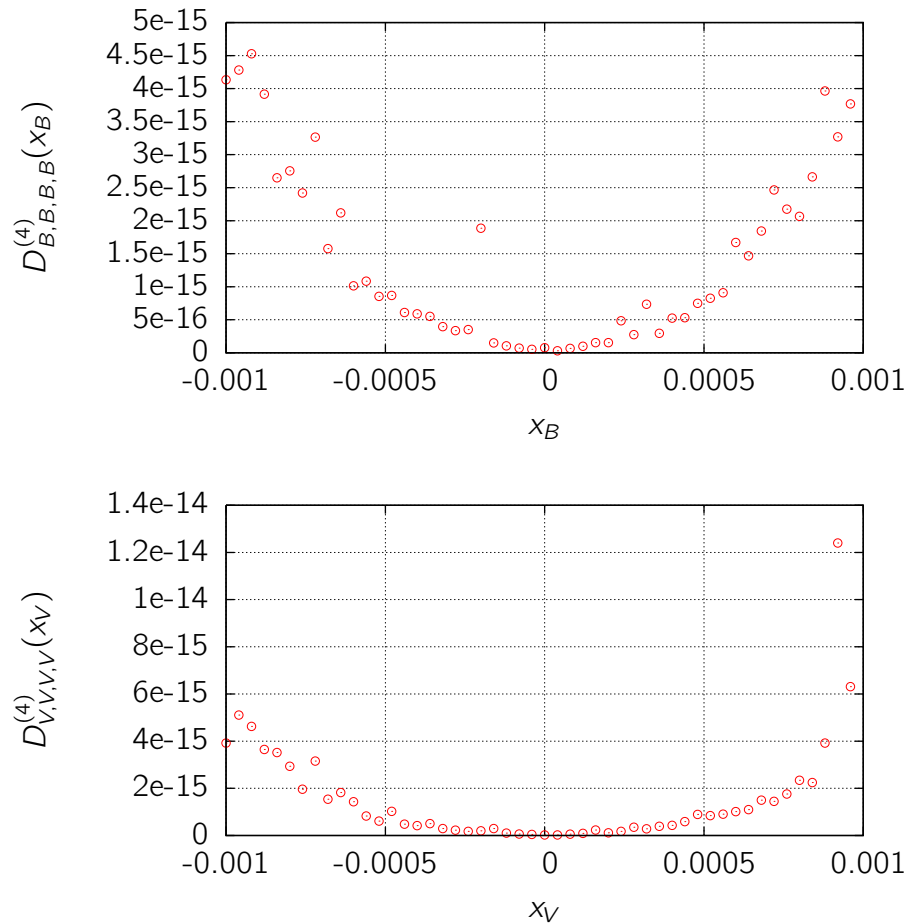


Abbildung 4.16: Exemplarisch sind hier Schätzungen für den vierten Koeffizienten der Kramers-Moyal-Erweiterung für die logarithmierten Renditen der beiden Aktien aufgetragen.

4.4.4 Der Prozess zur Beschreibung der Renditedynamik

Die gerade bestimmten Drift- und Diffusionskoeffizienten der Fokker-Planck-Gleichung (4.12) sollen nun über die Zusammenhänge (4.15) und (4.16) in

die Koeffizienten der korrespondierenden Langevin-Gleichung überführt werden. Dafür wird zunächst die „Wurzel der Diffusionsmatrix“ gesucht. Danach wird ein wesentlich schlankeres Differenzialgleichungssystem vorgestellt welches ebenfalls die Dynamik der Aktienkursdifferenzen wiedergibt.

Jede positiv definite, symmetrische Matrix D besitzt eine eindeutig bestimmte, positiv definite, symmetrische Wurzel $D^{1/2}$. Um diese zu bestimmen, diagonalisiert man die Matrix D und ersetzt die Diagonalelemente durch ihre positive Wurzel.

$$D^{1/2} = \begin{pmatrix} d_{11} & d_{12} \\ d_{21} & d_{22} \end{pmatrix} = T \cdot \begin{pmatrix} \sqrt{\lambda_1} & 0 \\ 0 & \sqrt{\lambda_2} \end{pmatrix} \cdot T^{-1}$$

Dabei sind $\lambda_{1,2}$ die Eigenwerte von D , und in den Spalten von T stehen die normierten Eigenvektoren der Diffusionsmatrix. Das Einsetzen der Größen führt auf ein Differentialgleichungssystem zur Beschreibung einzelner Trajektorien der logarithmierten Preisdifferenzen der Aktien von Bayer und VW gem. (4.10):

$$\begin{pmatrix} \dot{X}_B \\ \dot{X}_V \end{pmatrix} = \begin{pmatrix} a_B X_B \\ a_V X_V \end{pmatrix} + \begin{pmatrix} d_{11}(X_B, X_V) & d_{12}(X_B, X_V) \\ d_{21}(X_B, X_V) & d_{22}(X_B, X_V) \end{pmatrix} \cdot \begin{pmatrix} \Gamma_1 \\ \Gamma_2 \end{pmatrix}. \quad (4.32)$$

Auf das explizite Einsetzen der Größen wird verzichtet, da sich ein sehr komplexer Ausdruck ergibt.

Dieser intuitive Zugang die Langevin-Gleichung aus den geschätzten Kramers-Moyal-Koeffizienten zu rekonstruieren, führt auf einen komplizierten Ausdruck für den diffusiven Teil der Gleichung (4.32). Im Folgenden soll aufgezeigt werden wie man ein deutlich schlankeres Differentialgleichungssystem erhält. Dazu betrachtet man nun zunächst noch einmal ein einfaches System von Langevin-Gleichungen:

$$\dot{X}_1 = h_1(\mathbf{X}) + X_1 F_1 \quad (4.33)$$

$$\dot{X}_2 = h_2(\mathbf{X}) + X_2 F_2. \quad (4.34)$$

Es stellt sich die Frage wie sich die komplizierte Struktur des Diffusionsterms in Gleichung (4.32) vereinfachen lässt. Zunächst soll folgendes gelten:

$$F_1 := a\Gamma_1 + b\Gamma_2$$

$$F_2 := c\Gamma_1 + d\Gamma_2.$$

Die Rauschkräfte Γ_1 und Γ_2 werden weiterhin als unkorreliert angenommen. Die jetzt definierten Fluktuationen F_1 und F_2 sind damit korreliert. Es ergeben sich die Langevin-Gleichungen (4.33) und (4.34) in Vektorschreibweise zu:

$$\begin{pmatrix} \dot{X}_1 \\ \dot{X}_2 \end{pmatrix} = \begin{pmatrix} h_1(\mathbf{X}) \\ h_2(\mathbf{X}) \end{pmatrix} + \underbrace{\begin{pmatrix} X_1 a & X_1 b \\ X_2 c & X_2 d \end{pmatrix}}_M \cdot \begin{pmatrix} \Gamma_1 \\ \Gamma_2 \end{pmatrix}. \quad (4.35)$$

4.4 Schätzen der Drift- und Diffusionskomponenten

Die Elemente der Matrix M lassen sich aus der korrespondierenden Fokker-Planck-Gleichung mit folgender Beziehung herleiten:

$$(MM^t)_{ij} = D_{ij}^{(2)}.$$

Der Term $(MM^t)_{ij}$ steht dabei für die Elemente der entstehenden Matrix aus der Multiplikation von M mit ihrer transponierten M^t . Die empirisch geschätzte Fokker-Planck-Gleichung hat folgende Form:

$$\begin{aligned} \frac{\partial}{\partial t} p(\mathbf{x}, t) = & \left[-\frac{\partial}{\partial x_B} D_B^{(1)} - \frac{\partial}{\partial x_V} D_V^{(1)} + \frac{1}{2} \frac{\partial^2}{\partial x_B^2} D_{B,B}^{(2)} \right. \\ & \left. + \frac{1}{2} \frac{\partial^2}{\partial x_V^2} D_{V,V}^{(2)} + \frac{\partial^2}{\partial x_B \partial x_V} D_{B,V}^{(2)} \right] p(\mathbf{x}, t), \end{aligned}$$

mit

$$\begin{aligned} D_B^{(1)} &= a_B x_B \\ D_V^{(1)} &= a_V x_V \\ D_{B,B}^{(2)} &= \alpha x_B^2 \\ D_{V,V}^{(2)} &= \beta x_V^2 \\ D_{B,V}^{(2)} &= \gamma x_B x_V. \end{aligned}$$

Die Parameter der Diffusionsfunktionen (4.23), (4.24) und (4.26) werden als Varianz der Fluktuationen F_1 und F_2 berücksichtigt. Die Offsets werden an dieser Stelle vernachlässigt. Damit lässt sich die vereinfachte Form der Langevin-Gleichungen wie folgt angeben:

$$\dot{X}_B = a_B X_B + X_B F_1 \quad (4.36)$$

$$\dot{X}_V = a_V X_V + X_V F_2, \quad (4.37)$$

mit

$$\begin{aligned} \langle F_1^2 \rangle &= a^2 + b^2 = \alpha \\ \langle F_2^2 \rangle &= c^2 + d^2 = \beta \\ \langle F_1 F_2 \rangle &= ac + bd = \gamma. \end{aligned}$$

Die Koeffizienten a , b , c und d lassen sich im Prinzip über die Äquivalenz der Matrizen $D^{1/2}$ und M berechnen. Allerdings sind die Komponenten der Wurzel der Diffusionsmatrix zu komplex um diese hier explizit anzugeben. Dies ist auch der Fall wenn die vereinfachte Diffusionsmatrix \tilde{D} betrachtet wird. Nichtlineare Rechenoperationen (Charakteristisches Polynom, Invertierung der Eigenvektormatrix T) führen auf eine nichttriviale Struktur der Diffusionsmatrix in dem erstgenannten Differenzialgleichungssystem nach Langevin gemäß Gleichung (4.32).

Abschließende Betrachtung

Mit Hilfe bestimmter Datenaufbereitungsmaßnahmen (vgl. Kapitel 4.2.1) ist es gelungen die vorliegenden hochfrequenten Aktienkursdatensätze von Bayer und VW in der Weise aufzubereiten, dass eine zweidimensionale Drift- und Diffusionsschätzung durchgeführt werden kann. Die Dynamik der Preisdifferenzen kann so in einem zweidimensionalen Prozess zusammengefasst werden. Außerdem sind gewisse Abhängigkeitsstrukturen dieses Systems identifiziert worden.

Zunächst hat man auf die Synchronität von Kursverläufen hingewiesen und diese unter Zuhilfenahme von Korrelationsmessungen bestätigt. In diesem Zusammenhang wird mit dem IHT auf eine Methode aufmerksam gemacht welche geeignet ist um Korrelationshierarchien in Portfolien zu visualisieren. Die zweidimensionale Drift- und Diffusionsanalyse liefert weitere Hinweise auf die Abhängigkeitsstruktur der beiden Wertpapiere. So kann auf der einen Seite keine deterministische Kopplung aufgelöst werden, jedoch lässt sich die Diffusionsmatrix als bedingte Kovarianzmatrix identifizieren. Bedingt ist diese auf die jeweilige Auslenkung x_B und x_V der beiden Aktien. Außerdem ist eine zweidimensionale Langevin-Gleichung zur Beschreibung der Aktienkursdynamik bestimmt worden. Die Diagonalisierung der Diffusionsmatrix führte dabei zunächst auf einen komplizierten Ausdruck für den diffusiven Anteil der Langevin-Gleichung gemäß (4.32). Durch das Einführen von korrelierten Rauschkräften lassen sich die Prozesse zur Beschreibung der Preistrajektorien mit (4.36) und (4.37) deutlich vereinfacht darstellen.

Kapitel 5

Anwendung der Ergebnisse

Die in Kapitel 4 geschätzte Diffusionsmatrix im Zwei-Aktien-Fall kann als eine bedingte Kovarianzmatrix interpretiert werden. Diese beschreibt die Korrelation der beiden Aktienrenditen, bedingt auf deren Auslenkungen. Im Folgenden werden mit Hilfe dieser bedingten Kovarianzmatrix und den bedingten erwarteten Renditen in Form der geschätzten Driftfunktionen bedingte Minimal-Varianz-Portfolien (i. F. bedingte optimale Portfolien) berechnet.

5.1 Moderne Portfoliotheorie

Bei der Portfoliodiversifikation nach Markowitz geht es um die gezielte Kombination von Wertpapieren unter Berücksichtigung der Korrelationsstruktur. Eine Maximierung der Diversifikation führt dabei zu einer Minimierung des Konzentrationsrisikos (unsystematisches Risiko), es bleibt jedoch das Marktrisiko (systematisches Risiko). Die Notation richtet sich nach der Vorlesung „Finanzmathematik“ [30] die Privatdozent Jörg C. Lemm im Wintersemester 2008/09 an der Westfälischen-Wilhelms Universität Münster gehalten hat. Für eine quantitative Einführung in das Thema wird außerdem [12] empfohlen.

Der Gesamtwert eines Portfolios sei gegeben durch:

$$X = \sum_{i=1}^n W_i X_i, \quad (5.1)$$

wobei X_i den absoluten Wert und W_i die Anteile der Anlage i darstellen. Die Relativgewichte ergeben sich zu:

$$w_i = \frac{W_i X_i}{\sum_{i=1}^n W_i X_i}, \quad (5.2)$$

dabei ergibt sich die Summe aller Relativgewichte zu Eins. Mit σ_{ij} sei im Folgenden die Kovarianz zwischen Anlage i und Anlage j bezeichnet. Für das nun folgende Optimierungsproblem wird angenommen, dass die Kovarianzmatrix C und die Erwartungswerte der Rendite $\bar{r}_i = \langle r_i \rangle$ zur Verfügung stehen. Es gilt, die Portfoliovarianz σ_P^2 unter zwei Nebenbedingungen zu minimieren. In dieser allgemeinen Betrachtung sind Leerverkäufe ¹ ($w_i < 0$) zugelassen.

- Minimiere Portfoliovarianz mit variablem \mathbf{w} :

$$\sigma_P^2 = \mathbf{w}^T C \mathbf{w} \quad (5.3)$$

- 1. NB: Normierung der Gewichte:

$$\sum_{i=1}^n w_i = \mathbf{w}^T \mathbf{1}. \quad (5.4)$$

- 2. NB: Erwartete Gesamtrendite:

$$\sum_{i=1}^n w_i \bar{r}_i = \mathbf{w}^T \bar{\mathbf{r}} = \mu_P. \quad (5.5)$$

Dabei steht $\mathbf{1}$ für einen mit Einsen gefüllten Vektor. Mit Hilfe eines Lagrange-Ansatzes lässt sich das Problem lösen. Verzichtet man auf die Vorgabe einer Portfoliorendite μ_P , ergibt sich folgendes Optimum:

$$\mathbf{w} = \frac{1}{a} C^{-1} \mathbf{1}, \quad (5.6)$$

$$\sigma_{\min} = \frac{1}{\sqrt{a}}, \quad (5.7)$$

$$\mu_{\min} = b \sigma_{\min}^2, \quad (5.8)$$

mit:

$$a = \mathbf{1}^T C^{-1} \mathbf{1}, \quad (5.9)$$

$$b = \mathbf{1}^T C^{-1} \bar{\mathbf{r}}. \quad (5.10)$$

Im Fall von nur zwei möglichen Anlagealternativen ($n = 2$) ergibt sich die optimale Portfoliovarianz zu:

$$\sigma_P = \sqrt{w^2 \sigma_1^2 + (1-w)^2 \sigma_2^2 + 2w(1-w) \sigma_{12}}. \quad (5.11)$$

Anlagekombinationen unter variierender Korrelation ρ (Zusammenhang zu Kovarianz gemäß Gleichung (1.15).) lassen sich in einem Risiko-Rendite-Diagramm

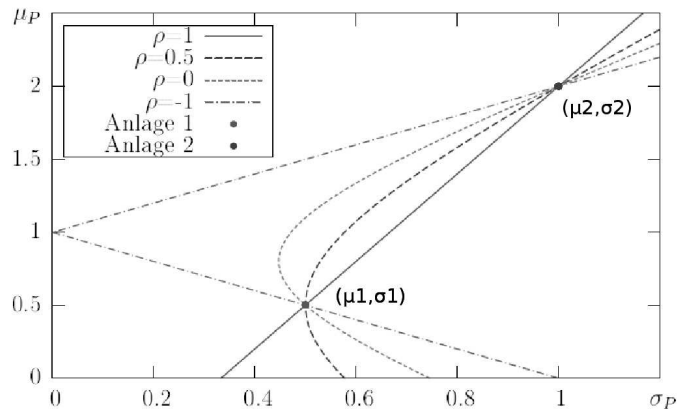


Abbildung 5.1: Das Risiko-Rendite-Diagramm für zwei Anlagen mit festen erwarteten Renditen μ_1, μ_2 und Varianzen σ_1, σ_2 . Für vier unterschiedliche Korrelationskoeffizienten sind mögliche Portfoliokombinationen eingezeichnet (entnommen aus [30]).

(siehe Abbildung 5.1) darstellen. Kombiniert man in einem Portfolio eine Anlage mit dem Risiko σ_m und der erwarteten Rendite r_m (Marktportfolio) mit einer risikolosen Anlage mit $\sigma_0 = 0$ und erwarteter Rendite r_0 , lässt sich eine folgende Gerade in der μ_P - σ_P -Ebene konstruieren:

$$\bar{r}_P = r_0 + \theta \sigma_P. \quad (5.12)$$

Diese Gerade nennt man *Kapitalmarktklinie* (KML) und sie stellt die erwarteten Rendite-Risiko-Kombinationen von effizienten Portfolios im Gleichgewicht dar. Der „Preis des Risikos“, die sogenannte *Sharpe-Ratio*, ergibt sich wie folgt:

$$\theta = \frac{\bar{r}_m - r_0}{\sigma_m}. \quad (5.13)$$

Diese Grundlagen legen das Fundament für das *Capital Asset Pricing Model* (CAPM), welches von Sharpe [45], Lintner und Mossin in den sechziger Jahren unabhängig voneinander entwickelt wurde.

5.1.1 Bedingte optimale Portfolios für die Aktien von Bayer und VW

Mit Hilfe der aus den Datensätzen geschätzten Diffusionsmatrizen und Driftfunktionen und der oben hergeleiteten Portfoliotheorie lassen sich nun optimale Portfolios bedingt auf die Auslenkung der Preisdifferenzen der Aktien

¹Verkauf von Anteilen, die von anderen Anlegern geliehen wurden, am Markt (nach [26]). Der Verkäufer profitiert dabei i. A. von fallenden Kursen.

berechnen. Die bedingte Kovarianzmatrix ist gegeben durch die Diffusionsmatrix (4.27):

$$C(x_B, x_V) = \begin{pmatrix} \alpha_0 + \alpha x_B^2 & \gamma_0 + \gamma x_B x_V \\ \gamma_0 + \gamma x_B x_V & \beta_0 + \beta x_V^2 \end{pmatrix}.$$

Die inverse Matrix ergibt sich zu:

$$C^{-1} = \frac{1}{N} \begin{pmatrix} \beta_0 + \beta x_V^2 & -(\gamma_0 + \gamma x_B x_V) \\ -(\gamma_0 + \gamma x_B x_V) & \alpha_0 + \alpha x_B^2 \end{pmatrix},$$

mit

$$N = ((\alpha_0 + \alpha x_B^2) (\beta_0 + \beta x_V^2)) - (\gamma_0 + \gamma x_B x_V)^2.$$

Die erwartete Rendite folgt aus den Komponenten des Driftvektors:

$$\begin{aligned} \mu_1(x_B) &= a_B x_B = D_B^{(1)}(x_B) \\ \mu_2(x_V) &= a_V x_V = D_V^{(1)}(x_V). \end{aligned}$$

Die Parameter dieser Eingangsgrößen wurden in Kapitel 4 aus den empirischen Daten geschätzt. Die Risiko-Rendite-Profile für die unterschiedlichen Auslenkungen der beiden Aktien sind in Abbildung 5.2 aufgetragen.

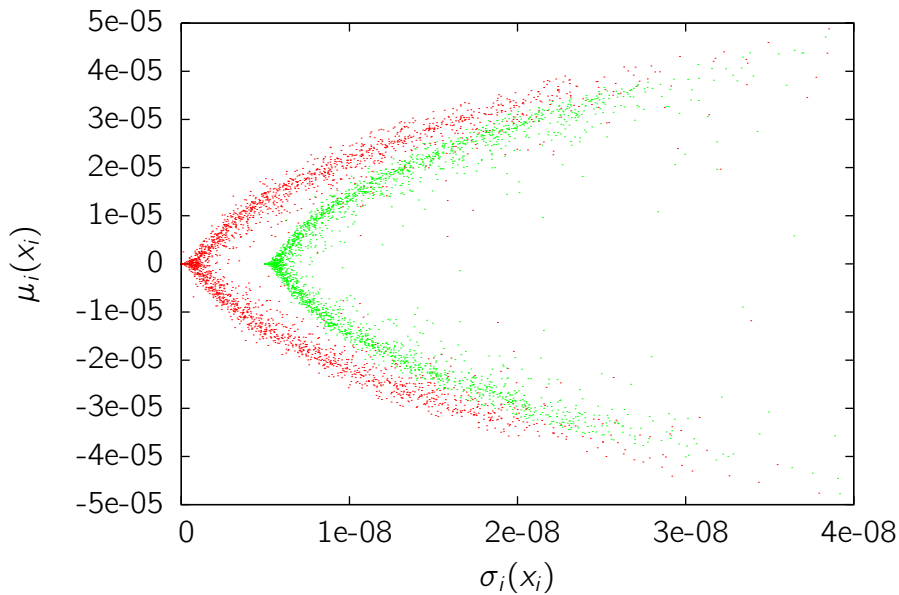
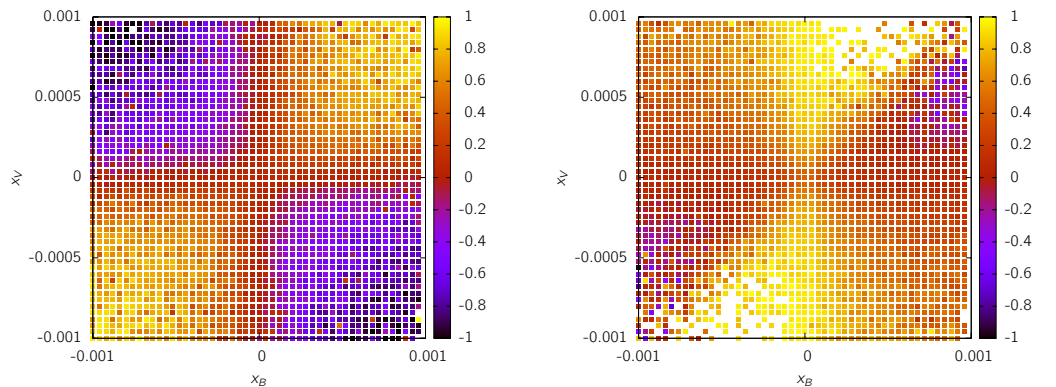


Abbildung 5.2: Risiko-Rendite-Profil für die unterschiedlichen Auslenkungenkombinationen von Bayer und VW. Aus Gründen der Übersichtlichkeit sind die Kombinationen der VW-Aktie (rechte Kurve) nach rechts verschoben. Für jedes Paar von Risiko-Rendite-Kombinationen der Aktien ließe sich ein Diagramm wie in Abb. 5.1 erstellen. Dabei wird die Korrelation durch die bedingte Kovarianzmatrix vorgegeben.

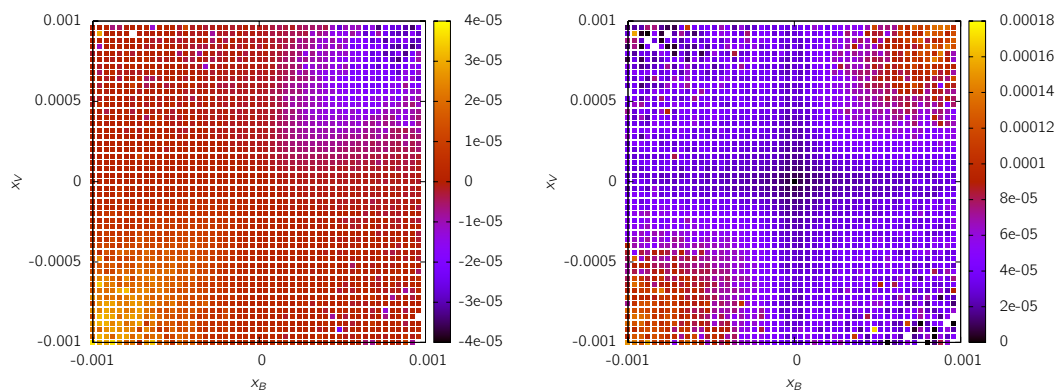
Der positive Teil der Kurven lässt sich wie folgt interpretieren: Je höher die erwartete Rendite, desto größer die Unsicherheit (das Risiko ²). Diese Beobachtung entspricht in den meisten Fällen der ökonomischen Realität.

Im Folgenden werden nun für die verschiedenen Kombinationen von x_B und x_V mit Hilfe der Diffusionsmatrix als bedingte Kovarianzmatrix und dem Drift als Erwartungswert der Rendite die optimale Gewichtung der beiden Aktien im Portfolio ausgerechnet. Außerdem lässt sich für jede Auslenkungscombination eine optimale Portfoliovarianz σ_P und eine optimale Portfoliorendite μ_P ausrechnen. Zusätzlich wird ein Korrelationsprofil aufgetragen.

²Als Risikomaß wird an dieser Stelle die Varianz verwendet.



- (a) In diesem Korrelationsprofil ist für jede Kombination der Preisdifferenzen zwischen Bayer und VW der Korrelationskoeffizient gem. Gleichung (1.15) angegeben.
- (b) Die optimale relative Gewichtung der Bayer-Aktie w_B ist hier aufgetragen. Die relativen Gewichte der Volkswagen-Aktie w_V ergeben sich über $w_V = 1 - w_B$.



- (c) Aufgetragen sind die optimalen Renditen gem. Gleichung (5.5). Die freien Stellen in der Nähe des Nullpunkts liegen unterhalb (4. Quadrant) und überhalb (2. Quadrant) des aufgetragenen Bereichs.
- (d) Die optimalen Varianzen für alle Auslenungskombinationen gem. Gleichung (5.11).

Abbildung 5.3: Diese Graphen zeigen Kenngrößen zur Charakterisierung der optimalen Portfolios im Zwei-Aktien-Fall. Auf den Achsen sind jeweils die Auslenkungen der Preisdifferenzen von Bayer x_B und Volkswagen x_V aufgetragen. Die Farbkodierung gibt die Werte der unter den Abbildungen genannten Größen wieder.

Für das Korrelationsprofil in Abbildung 5.3(a) ergibt sich das erwartete Ergebnis. Für gegenläufige Auslenkungen der Preisdifferenzen liegt negative Korrelation vor, für Auslenkungen in die gleiche Richtung liegt positive Korrelation vor.

Zur Interpretation der optimalen Gewichte in Abbildung 5.3(b) kann man anmerken, dass die optimalen Gewichte von der inversen Kovarianzmatrix gem. (5.6) abhängen und mittels der Optimierung (5.3) so gewählt werden, dass sich eine minimale Portfoliovarianz σ_P^2 ergibt.

Bei den optimalen Renditen in Abbildung 5.3(c) ist folgende Tendenz zu erkennen: Für negative Auslenkungen, also positive erwartete Renditen (die linearen Driftfunktionen haben eine negative Steigung), ergeben sich tendenziell positive Portfoliorenditen. Analoges gilt für den Fall der positiven Auslenkungen der Preisdifferenzen.

Für die optimalen Portfoliovarianzen in Abbildung 5.3(d) lässt sich ablesen, dass sich das Risiko durch die Wertpapiermischung verringert. Lediglich für große positive Korrelationen bleibt das Risiko auf einem hohen Niveau und die Varianz der Portfolios liegt deutlich über der Varianz der Einzelwerte.

Bewertung der Anwendungsmöglichkeit

In diesem Kapitel wurde der Versuch unternommen, für unterschiedliche Kombinationen von Preisdifferenzen zweier Aktien optimale Portfolios zu bestimmen. Dies stellt eine neue, mikroskopische Sicht auf die Portfoliostruktur im Zwei-Wertpapier-Fall dar.

Grundsätzlich kann angemerkt werden, dass einige qualitative Beobachtungen gemacht werden konnten, die sich mit ihrer ökonomischen Bedeutung vereinbaren lassen. So wurden für die Form der optimalen erwarteten Renditen und optimalen Varianzen Zusammenhänge gezeigt, die sich aus der Korrelationsstruktur ableiten lassen.

Für eine Anwendung für das Management von Aktienportfolios scheint diese Betrachtung jedoch keine Relevanz zu haben. Eine Neustrukturierung eines Aktienportfolios je nach erwarteter Momentanrendite ist u. a. durch Transferkosten unwirtschaftlich. Möglicherweise lässt sich diese Art der Portfoliosteuerung jedoch auf andere Bereiche anwenden, in denen insbesondere Transferpreise oder Transferzeiten eine untergeordnete Rolle spielen. Allerdings ist auch anzumerken, dass diese Methode mit steigender Anzahl an betrachteten Prozessen (im vorliegenden Fall nur zwei Aktien) komplexer wird.

Kapitel 6

Zusammenfassung und Ausblick

Zunächst ist eine umfangreiche statistische Charakterisierung von Kursdifferenzen des Deutschen Aktien Index auf unterschiedlichen Zeitskalen durchgeführt worden. Dabei konnten prominente Ergebnisse aus empirischen Untersuchungen bestätigt werden. Insgesamt kann festgestellt werden, dass auf die unterschiedlichen Arten von Wirtschaftsdaten (z. B. Aktienkurse, Rohstoffpreise, Managergehälter, Arbeitslosenzahlen, usw.) zahlreiche statistische Tests und Messungen angewendet werden. Viele Effekte unterschiedlicher Art sind dabei bereits identifiziert und verstanden. Allerdings ist es nach wie vor sehr schwer auch nur in Teilen die Entwicklung eines komplexen Systems wie der Weltwirtschaft adäquat zu beschreiben oder mit zufriedenstellender Güte vorausszusagen. Insbesondere wenn es sich weit außerhalb des Gleichgewichts befindet wie es zum Beispiel bei einer Wirtschaftskrise der Fall ist.

Die eindimensionale Drift- und Diffusionsanalyse lässt sich erfolgreich auf hochfrequente Finanzdaten anwenden. Für die Kursdifferenzen des DAX konnten ähnliche Koeffizientenstrukturen wie z. B. für FOREX-Kurse beobachtet werden. Diese anscheinend allgemeingültigen Beobachtungen für den Trend und die Volatilität von Renditen lassen sich ökonomisch begründen und finden bereits in Modellen zur Kursmodellierung Anwendung.

Interessant könnte eine detailliertere Betrachtung der Markoveigenschaften von Finanzdaten auf kleinen Skalen ($\tau < 1$ min) sein. Wie bereits von Nawroth in [39] angedeutet, lassen sich unterschiedliche Zeitskalenregime nachweisen. Die in dieser Arbeit vorgestellten Markovanalysen erfordern eine gewisse Anzahl an Datenpunkten auf den Skalen die betrachtet werden. Zwar sind die hier verwendeten Datensätze teilweise ausreichend fein aufgelöst, liefern jedoch insgesamt zu wenig Datenpunkte um Analysen dieser Art anzuwenden. Erfolgsversprechender könnten aktuelle FOREX-Datensätze sein, welche eine sehr hohe Datendichte ¹ liefern.

¹Zum Beispiel sind für den Wechselkurs EUR/USD für das Jahr 2008 ≈ 78 Mio. Datenpunkte

Die zweidimensionale Drift- und Diffusionsanalyse von hochfrequenten Aktienkursen stellt besondere Herausforderungen an die Datenpräparation. Mit dem Einsatz von Interpolationsverfahren riskiert man i. A. eine „statistische Verfälschung“ der Ursprungsdaten. In dieser Arbeit konnte gezeigt werden, dass bei einer Datenaufbereitung mit Hilfe einer einfachen linearen Interpolation sowohl die Form der Wahrscheinlichkeitsdichten als auch die Markoveigenschaften weitestgehend erhalten geblieben sind. Eine Erkenntnis dieser Arbeit ist, dass für die präparierten Aktienrenditen keine deterministische Kopplung vorliegt, wie man sie zum Beispiel für Verkehrsdaten zwischen der Geschwindigkeit und dem Verkehrsfluss finden konnte (siehe Kapitel 4 oder [29]). Die Diffusionsmatrix zeigt jedoch das Kopplungsterme in dem fluktuierenden Teil des Prozesses auftreten. Es konnte schließlich ein stochastisches Differentialgleichungssystem in der Notation nach Langevin angegeben werden welches die Dynamiken der Aktienkursrenditen beschreibt.

Ausgehend von diesen Ergebnissen könnte es interessant sein noch größere Portfolios bzw. vollständige Indizes zu betrachten, also die Anzahl der zu untersuchenden Prozesse zu erhöhen. Denkbar wäre es zum Beispiel eine hochdimensionale Drift- und Diffusionsanalyse der Kursverläufe wichtiger Wechselkurspaare durchzuführen.

Mit der Betrachtung der Diffusionsmatrix als eine bedingte Kovarianzmatrix eröffnete sich eine Anwendung auf die Portfoliotheorie nach Markowitz. So konnte für die unterschiedlichen Auslenkungskombinationen der beiden Aktien Mischungsverhältnisse ausgerechnet werden, welche die Gesamtvarianz des Portfolios minimieren. Auch wenn diese Betrachtung der „mikroskopische Portfoliostruktur“ nicht sofort eine Praxisanwendung zur Portfoliosteuerung verspricht, so könnte es dennoch interessant sein Zusammenhänge dieser Art weiter zu studieren.

verfügbar. (Quelle: Olsen Data AG)

Anhang A

Anhang

A.1 Koordinatentransformation

Gegeben ist folgende Rotationsmatrix:

$$R(\phi) = \begin{pmatrix} \cos \phi & -\sin \phi \\ \sin \phi & \cos \phi \end{pmatrix}. \quad (\text{A.1})$$

Mit $\phi_0 = 45^\circ$ folgt:

$$R(\phi_0) = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}, \quad (\text{A.2})$$

welche auf die Ausgangskordinaten x und y angewendet wird:

$$\begin{pmatrix} \tilde{x} \\ \tilde{y} \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}. \quad (\text{A.3})$$

Es folgt:

$$\begin{aligned} \tilde{x} &= \frac{1}{\sqrt{2}}(x - y) \\ \tilde{y} &= \frac{1}{\sqrt{2}}(x + y) \end{aligned}$$

und aus $f(x, y) = x^2 - y^2$ folgt mit \tilde{x} und \tilde{y} :

$$\begin{aligned} f(\tilde{x}, \tilde{y}) &= \frac{1}{\sqrt{2}} [(x - y)^2 - (x + y)^2] \\ &= -2xy. \end{aligned}$$

A.2 Lösung des linearen Gleichungssystems zur Bestimmung der Parameter eines kubischen Polynoms

Zur Bestimmung der Parameter eines kubischen Polynoms der Form

$$x = at^3 + bt^2 + ct + d$$

benötigt man vier Koordinaten (x_1, t_1) , (x_2, t_2) , (x_3, t_3) und (x_4, t_4) . Durch Lösen des linearen Gleichungssystems

$$\begin{aligned} x_1 &= at_1^3 + bt_1^2 + ct_1 + d, \\ x_2 &= at_2^3 + bt_2^2 + ct_2 + d, \\ x_3 &= at_3^3 + bt_3^2 + ct_3 + d, \\ x_4 &= at_4^3 + bt_4^2 + ct_4 + d, \end{aligned}$$

erhält man folgende Parameter:

$$\begin{aligned} a &= \left[t_2 (t_1^2(x_3 - x_4) + t_3^2(x_4 - x_1) + t_4^2(x_1 - x_3)) + t_3 (t_1^2(x_4 - x_2) + t_4^2(x_2 - x_1)) \right. \\ &\quad + t_2^2(t_1(x_4 - x_3) + t_3(x_1 - x_4) + t_4(x_3 - x_1)) + t_3^2(t_1(x_2 - x_4) + t_4(x_1 - x_2)) \\ &\quad \left. + t_1 t_4(t_1 - t_4)(x_2 - x_3) \right] \\ &\quad / [(t_1 - t_2)(t_1 - t_3)(t_1 - t_4)(t_2 - t_3)(t_2 - t_4)(t_3 - t_4)], \\ b &= \left[t_2 (t_1^3(x_4 - x_3) + t_3^3(x_1 - x_4) + t_4^3(x_3 - x_1)) + t_3 (t_1^3(x_2 - x_4) + t_4^3(x_1 - x_2)) \right. \\ &\quad - t_1 t_4 (t_1^2 - t_4^2)(x_2 - x_3) + t_2^2(t_1(x_3 - x_4) + t_3(x_4 - x_1) \\ &\quad \left. + t_4(x_1 - x_3)) + t_3^2(t_1(x_4 - x_2) + t_4(x_2 - x_1)) \right] \\ &\quad / [(t_1 - t_2)(t_1 - t_3)(t_1 - t_4)(t_2 - t_3)(t_2 - t_4)(t_3 - t_4)], \\ c &= \left[t_2^2 (t_1^3(x_3 - x_4) + t_3^3(x_4 - x_1) + t_4^3(x_1 - x_3)) + t_3^2 (t_1^3(x_4 - x_2) + t_4^3(x_2 - x_1)) \right. \\ &\quad + t_2^3 (t_1^2(x_4 - x_3) + t_3^2(x_1 - x_4) + t_4^2(x_3 - x_1)) \\ &\quad \left. + t_3^3 (t_1^2(x_2 - x_4) + t_4^2(x_1 - x_2)) + t_1^2 t_4^2(t_1 - t_4)(x_2 - x_3) \right] \\ &\quad / [(t_1 - t_2)(t_1 - t_3)(t_1 - t_4)(t_2 - t_3)(t_2 - t_4)(t_3 - t_4)], \\ d &= \left[t_2^3 (t_3 (t_4^2 x_1 - t_1^2 x_4) + t_3^2(t_1 x_4 - t_4 x_1) + t_1 t_4 x_3(t_1 - t_4)) \right. \\ &\quad + t_2^2 (t_3 (t_1^3 x_4 - t_4^3 x_1) + t_1 t_4 x_3 (t_4^2 - t_1^2) \\ &\quad + t_3^3(t_4 x_1 - t_1 x_4)) + t_2 (t_3^2 (t_4^3 x_1 - t_1^3 x_4) \\ &\quad + t_3^3 (t_1^2 x_4 - t_4^2 x_1) + t_1^2 t_4^2 x_3(t_1 - t_4)) \\ &\quad \left. + t_1 t_3 t_4 x_2(t_1 - t_3)(t_1 - t_4)(t_3 - t_4) \right] \\ &\quad / [(t_1 - t_2)(t_1 - t_3)(t_1 - t_4)(t_2 - t_3)(t_2 - t_4)(t_3 - t_4)]. \end{aligned}$$

A.3 Aktienkursmodellierung

Im Folgenden werden drei bekannte Modelle zur Beschreibung von Aktienkursdynamiken vorgestellt.

A.3.1 Geometrische Brownsche Bewegung

Im Optionspreismodell von Black und Scholes [5] wird angenommen, dass der Aktienkurs S_t einer geometrischen brownischen Bewegung der Form:

$$dS_t = \mu S_t dt + \sigma S_t dW_t \quad (\text{A.4})$$

folgt. Die erwartete Momentanrendite μ und die Momentanvolatilität σ werden dabei als konstant angenommen [38]. Durch Anwendung der Itô-Formel erhält man folgende Lösung der SDE:

$$S_t = S_0 \exp \left\{ \underbrace{\left(\mu - \frac{\sigma^2}{2} \right) t + \sigma W_t}_A \right\}. \quad (\text{A.5})$$

Es gilt:

$$A \sim \mathcal{N} \left(\left(\mu - \frac{\sigma^2}{2} \right) t, \sigma^2 t \right)$$

und dementsprechend ist S_t log-normalverteilt. Die Renditen über ein Zeitintervall $[t, t+h]$ mit $h > 0$

$$Z_t(h) = \ln \left(\frac{S_{t+h}}{S_t} \right), \quad (\text{A.6})$$

sind ebenfalls normalverteilt:

$$Z_t(h) \sim \mathcal{N}(\mu h, \sigma^2 h). \quad (\text{A.7})$$

A.3.2 Jump-Diffusion-Modell

In einem Jump-Diffusion-Modell erweitert man den Renditeprozess einer GBM mit einem Poisson-Prozess [10]:

$$Z_t(h) = \begin{cases} x & \text{für } K = 0 \\ x + y_1 + \dots + y_K & \text{für } K > 1 \end{cases}$$

Dabei sind die Variablen wie folgt verteilt:

$$x \sim \mathcal{N}(\alpha h, \sigma^2 h) \quad (\text{A.8})$$

$$y_1, y_2, \dots, y_K \sim \mathcal{G} \quad (\text{A.9})$$

$$K \sim P(k = K) = e^{-\lambda h} \frac{\lambda h^k}{k!} \quad (\text{A.10})$$

mit $k = 0, 1, 2, \dots$. Der Parameter λ bezeichnet die Sprungintensität mit $\lambda > 0$. In einem bekannten Modell von Merton gilt $\mathcal{G} = \mathcal{N}$.

A.3.3 Stochastische Volatilität

In Modellen mit stochastischer Volatilität (SV) wird der Varianz $\sigma^2 = V$ ein eigener stochastischer Prozess unterstellt. Dabei gilt nun $x_t = \ln\left(\frac{S_t}{S_0}\right)$:

$$dx_t = \alpha dt + \sqrt{V} dW_t \quad (\text{A.11})$$

$$dV_t = \xi(t, V_t) dt + \beta(t, V_t) dB_t. \quad (\text{A.12})$$

In einem bekannten Modell von Heston [23] entspricht die Volatilität einem sogenannten *Mean-Reverting Wurzel Prozess*:

$$dV_t = \kappa(\theta - V_t) dt + \eta \sqrt{V_t} dB_t. \quad (\text{A.13})$$

Der Parameter κ bezeichnet die Geschwindigkeit, mit der die Momentvarianz V_t zu ihrem langfristigen Zielwert θ zurückgezogen wird. η ist die Volatilität der Momentvarianz [38]. Dieses Modell berücksichtigt die empirisch beobachtete Korrelation ρ zwischen Rendite und Volatilität (Leverage-Effekt ¹) über die folgende Beziehung:

$$\frac{dB_t dW_t}{dt} = \rho. \quad (\text{A.14})$$

¹Durch empirische Untersuchungen wurde festgestellt das die Volatilität abnimmt wenn der Kurs steigt, dieser Effekt wird in diesem Zusammenhang Leverage-Effekt genannt (siehe [38], S.11f).

Literaturverzeichnis

- [1] AïT-SAHALIA, Y.; MYKLAND, P.A.; ZHANG, Z.: *Ultra High Frequency Volatility Estimation with Dependent Microstructure Noise*. NBER Working Paper No. W11380 (2005)
- [2] ALBRECHT, P.; RAIMOND, M.: *Investment- und Risikomanagement*. Stuttgart: Schäffer Poeschel, 2008.
- [3] ANDERSEN, T.G.; BOLLERSLEV, T.; MEDDAHI, N.: *Correcting the Errors: Volatility Forecast Evaluation Using High-Frequency Data and Realized Volatilities*. *Econometrica* **73** 1, 279-296 (2005)
- [4] BACHELIER, L.: *Théorie de la Spéculation*. Annales scientifiques de l'École Normale Supérieure Sér. 3, **17**, 21-86 (1900)
- [5] BLACK, F.; SCHOLES, M.: *The Pricing of Options and Corporate Liabilities*. *Journal of Political Economy* **81**, 637-654.(1973)
- [6] BOLLERSLEV, T.: *Generalized Autoregressive Conditional Heteroskedasticity*. *Journal of Econometrics* **31** No.: 3, 307-327 (1986)
- [7] BRONSTEIN, I.N.; SEMENDJAJEW, K.A., MUSIOL, G.; MÜHLIG, H.: *Taschenbuch der Mathematik*. Frankfurt am Main: Harri Deutsch, 1999.
- [8] CONT, R.: *Empirical properties of asset returns: stylized facts and statistical issues*. *Quantitative Finance* **1** 223-236 (2001)
- [9] CONT, R.: *Volatility Clustering in Financial Markets: Empirical Facts and Agent-Based Models*. erschienen in KIRMAN, A.; TEYSSIERE, G.: *Long memory in economics*. Berlin: Springer, 2005.
- [10] DAS, S.R.; SUNDARAM, R.K.: *Of Smiles and Smirks: A Term Structure Perspective*. *Journal of Financial and Quantitative Analysis* **34** 2, 211-239 (1999)
- [11] DETTLING, M.: *Bessere Volatilitäts- und Risikoschätzungen mit Hochfrequenzdaten*. Diplomarbeit, Eidgenössische Technische Hochschule, Zürich (2000)
- [12] DEUTSCH, H.-P.: *Quantitative Portfoliosteuerung*. Stuttgart: Schäffer-Poeschel, 2005.
- [13] EVERTSZ, C.J.G.: *Fractal Geometry of Financial Time Series*. *Fractals*. **3**, 609-616 (1995)

- [14] FAMA, E.F.: *The Behavior of Stock Market Prices*. Journal of Business **38**, 34-105 (1965)
- [15] FRIEDRICH, R.; PEINKE, J.; RENNER, CH.: *Markov properties of high frequency exchange rate data*. International Journal of Theoretical and Applied Finance **3** 3, S. 415-416 (2000)
- [16] FRIEDRICH, R.; SIEGERT, S.; PEINKE, J.; LÜCK, S.; SIEFERT, M.; LINDEMANN, M.; RAETHJEN, J.; DEUSCHL, G.; PFISTER, G.: *Extracting model equations from experimental data*. Phys. Lett. A **271**, 217-222 (2000)
- [17] FUNG, W.; HSIEH, D.: *Empirical characteristics of dynamic trading strategies: The case of Hedge Fonds*. The Review of Financial Studies **10** 2, 275-302 (1997)
- [18] GARDINER, C.W.: *Handbook of Stochastic Methods*. Heidelberg: Springer, 1983.
- [19] GHASEMI, F.; SAHINI, M.; PEINKE, J.; REZA, M.; TABAR, R.: *Analysis of Non-stationary Data for Heart-Rate Fluctuations in Terms of Drift and Diffusion Coefficients*. Journal of Biological Physics **32** 2, 117-128 (2006)
- [20] GILGEN, K.: *Planungsmethodik in der kommunalen Raumplanung: Vom Praxisbeispiel zur Theorie*. Zürich: vdf Hochschulverlag AG ETH Zürich, 2006.
- [21] GORSKI, A.Z.; PROZDZ, S.; SPETH, J.: *Financial multifractality and its subtleties: an example of DAX*. Physica A. **316**, 496-510 (2002)
- [22] HAKEN, H.: *Synergetics: Introduction and Advanced Topics*. Stuttgart: Springer, 2004.
- [23] HESTON, S.L.: *A Closed-Form Solution for Options with Stochastic Volatility wocations to Bond and Currency Options*. The Review of Financial Studies **6** 2, 327-343 (1993)
- [24] HONERKAMP, J.: *Statistical Physics: An Advanced Approach with Applications*. Heidelberg: Springer, 1998.
- [25] HULL, J.; WHITE, A.: *The Pricing of Options on Assets with Stochastic Volatilities*. The Journal of Finance **42** 2, 281-300 (1987)
- [26] HULL, J.C.: *Optionen, Futures and andere Derivate*. München: Pearson Studium, 2006.
- [27] HURST, H.E.: *Long-term storage capacity of reservoirs*. Trans. Amer. Soc. Civil Engineers **116**, 770-799 (1951)
- [28] KLEINHANS, D.: *Analyse stochastischer Prozesse: Eine iterative Methode zur Bestimmung von Drift und Diffusion*. Diplomarbeit, Westfälische-Wilhelms Universität, Münster (2005)

- [29] KRISO, S.; FRIEDRICH, R.; PEINKE, J.; WAGNER, P.: *Reconstruction of dynamical equations for traffic flow*. Phys. Lett. A **299**, 287 (2002)
- [30] LEMM, J.: *Vorlesung Finanzmathematik: 2. Portfoliooptimierung*. [On-line] Verfügbar unter HTTP: Hostname: http://pauli.uni-muenster.de/lemm/vorl_WS08/Folien_Vorlesung_WS08_09_Portfolio.pdf [12.04.2009]
- [31] LIU, Y.; GOPIKRISHNAN, P.; CIZEAU, P.; MEYER, M.; PENG, C.K.; STANLEY, H.E.: *Statistical properties of the volatility of price fluctuations*. Phys. Rev. E **60**, 1390 (1999)
- [32] LÜCK, St.; PEINKE, J.: *Uniform Statistical Description of the Transition between Near and Far Field Turbulence in a Wake Flow*. Phys. Rev. Lett. **83** 26, 5495-5498 (1999)
- [33] MAHNKE, R.; KAUPUŽS, J.; LUBASHEVSKY, I.: *Physics of Stochastic Processes: How Randomness Acts in Time*. Weinheim: WILEY-VCH, 2009.
- [34] MANDELBROT, B.: *The Variation of Certain Speculative Prices*. Journal of Business **36**, 392-417 (1963)
- [35] MANTEGNA, R. N.; STANLEY, H. E.: *An Introduction to Econophysics: Correlations and Complexity in Finance*. Cambridge: Cambridge University Press, 2000.
- [36] MARKOWITZ, H.: *Portfolio Selection* The Journal of Finance **7** 1, (1952)
- [37] MERTON, R. C.: *Continuous-Time Finance*. Basil Blackwell, Cambridge, MA (1990)
- [38] NAGEL, H.: *Optionsbewertung bei stochastischer Volatilität*. Wiesbaden: Gabler, 2001.
- [39] NAWROTH, A.P.; PEINKE, J.: *Small scale behavior of financial data*. Eur. Phys. J. B **50**, 147-151 (2006)
- [40] PAUL, W.; BASCHNAGEL, J.: *Stochastic Processes: From Physics to Finance*. Heidelberg: Springer, 1999.
- [41] PRESS, W.H.; TEUKOLSKY, S.A.; VETTERLING, W.T.; FLAUNERY, B.P.: *Numerical Recipes: The Art of Scientific Computing*. Cambridge: Cambridge University Press, 3. Edition, 2007.
- [42] REMER, R.: *Theorie und Simulation von Zeitreihen mit Anwendungen auf die Aktienkursdynamik*. Dissertation, Universität Rostock, Rostock (2005)
- [43] RISKEN, H.: *The Fokker-Planck Equation: Methods of Solution and Applications*. Heidelberg, 1989.

- [44] SEWELL, M.: *Characterization of Financial Time Series*. [On-line] Verfügbar unter HTTP: Hostname: finance.martinsewell.com Verzeichnis: /stylized-facts/characterization.pdf [25.02.2009]
- [45] SHARPE, W.F.: *Capital Asset Prices: A Theory of Market Equilibrium under Conditions of Risk*. The Journal of Finance **19** 3, 425-442 (1964)
- [46] SIEGERT, S.; FRIEDRICH, R.; PEINKE, J.: *Analysis of datasets of stochastic systems*. Physics Letters A **243**, 275-280 (1998)
- [47] STURGES, H.A.: *The choice of a class interval*. J. Amer. Statist. Assoc. **21**, S. 65-66 (1926)
- [48] TICHY, G.: *Konjunktur: Stilisierte Fakten, Theorie, Prognose*. Heidelberg: Springer-Verlag, 2. Auflage, 1993.
- [49] WAECHTER, M.; RIESS, F.; SCHIMMEL, TH.; WENDT, U.; PEINKE, J.: *Stochastic analysis of different rough surfaces*. Eur. Phys. J. B **41** 259-277 (2004)
- [50] WAECHTER, M.; KOUZMITCHEV, A.; PEINKE, J.: *Increment definitions for scale dependent analysis of stochastic data*. Phys. Rev. E **70**, (2004)
- [51] WAGNER, F.: *Volatility Cluster and Herding*. Physica A: Statistical Mechanics and its Applications **322** 1, 607-619 (2003)
- [52] Die Daten wurden freundlicherweise von Prof. Dr. Joachim Peinke vom Institut für Physik & ForWind der Universität Oldenburg zur Verfügung gestellt.
- [53] <http://finance.yahoo.com>

Danksagung

Mein besonderer Dank gilt ...

- ... Prof. Dr. R. Friedrich für die interessante Themenwahl und die freundliche Unterstützung.
- ... Prof. Dr. J. Peinke für die Bereitstellung der Datensätze.
- ... der gesamten Arbeitsgruppe „Selforganization and Complexity“. Insbesondere seien erwähnt Oliver Kamps, Michel Voßkuhle, Michael Wilczek, Max Theodor Kuchel und Holger Angenent für fachliche und technische Unterstützung und das Korrekturlesen der Arbeit.
- ... meinen Eltern, Friederike und Friedhelm Schuppert, für den familiären Rückhalt.
- ... den Jungs vom CLUB, echtefreunde, allen MoFFs und ganz besonders Katharina Kipp für die gemeinsame Zeit neben dem Studium.

Erklärung zur Diplomarbeit

Hiermit versichere ich, diese Arbeit selbstständig angefertigt und keine anderen als die angegebenen Hilfsmittel verwendet zu haben.

Münster, den 09. Dezember 2009