



TECHNISCHE
UNIVERSITÄT
WIEN

D I P L O M A R B E I T

Deep Learning in illiquiden Märkten

ausgeführt am

Institut für
Stochastik und Wirtschaftsmathematik
TU Wien

unter der Anleitung von

Univ.Prof. Dipl.-Math. Dr.rer.nat. Thorsten Rheinländer

durch

Valentin Linsbichler, BSc

Matrikelnummer: 11810331

Wien, 13.02.2025

Valentin Linsbichler

Thorsten Rheinländer

Kurzfassung

Viele herkömmliche Modelle, die in der Finanzmathematik zur Berechnung der Preisentwicklung diverser Finanzprodukte genutzt werden, setzen einen liquiden Markt voraus, in dem die Preisentwicklung nicht durch das Verhalten der Marktteilnehmer*innen beeinflusst werden kann. Die vorliegende Arbeit fokussiert sich auf illiquide Finanzmärkte und insbesondere auf den Einfluss, den Akteur*innen mit hinreichend großem Handelsvolumen auf den Markt haben. Um diese Effekte abzubilden, bedarf es einer Adaptierung beziehungsweise einer Erweiterung bestehender Modelle. Der entsprechende Einfluss kann durch einen Illiquiditätsparameter quantifiziert werden und unter Umständen zur Folge haben, dass die theoretisch optimale Handelsstrategie unbeschränkt wächst und eine Destabilisierung des Marktes zur Folge hätte. Auf Grund der Abwesenheit einer stabilen, optimalen Strategie bezeichnen wir dieses Szenario als instabiles Regime. Insbesondere, wenn zur Ermittlung des Illiquiditätsparameters ein neuronales Netz herangezogen wird, stellt sich die Frage, ob die Möglichkeit besteht, dass ein eingangs stabiles Szenario ausgelöst durch die auftretende Diffusion ins instabile Regime abdriftet. Die vorliegende Diplomarbeit beantwortet diese Frage mit Hilfe eines zu diesem Zweck implementierten neuronalen Netzes und zeigt, dass ein solcher Übergang mit positiver Wahrscheinlichkeit eintreten kann. Als anschauliches Beispiel dafür dient die Ausweitung des Bachelier Modells auf einen illiquiden Finanzmarkt. Anhand dieses konkreten Beispiels werden darüber hinaus auch die Verteilung des Wertprozesses sowie dessen Nutzen und die ersten beiden Momente dieser Verteilungen untersucht.

Abstract

Many of the conventional models used in financial mathematics to determine how asset values develop use the assumption that the market is liquid and therefore market participants actions have no bearing on asset prices. This thesis focuses on illiquid financial markets and particularly on the impact that traders with significant trading volumes have on the market. To map these effects, existing models need to be adapted or extended. The corresponding influence can be quantified by an illiquidity parameter and may result in the theoretically optimal trading strategy growing infinitely which could destabilize the market. Due to the absence of a stable, optimal strategy, we refer to this scenario as the unstable regime. In particular, if a neural network is used to determine the illiquidity parameter, the question arises whether there is a chance that an initially stable scenario drifts into the unstable regime caused by the occurring diffusion. This thesis provides an answer to this question using a neural network that shows that such a transition can occur with a positive probability. The extension of the Bachelier model to an illiquid financial market serves as an illustrative example. This specific example is also used to analyse the distribution of the value process as well as its utility and the first two moments of these distributions.

Danksagung

Ich möchte diese Gelegenheit nutzen, um mich bei allen Menschen zu bedanken, die mich im Verlauf meines Studiums unterstützt, begleitet und motiviert haben. Einen besonderen Dank möchte ich an dieser Stelle meinem Betreuer Univ.Prof. Dipl.-Math. Dr.rer.nat. Thorsten Rheinländer für seine Unterstützung und Begleitung aussprechen. Danken möchte ich außerdem meinen Eltern Petra und Andreas, ohne deren Unterstützung mein Studium nicht möglich gewesen wäre sowie meiner Partnerin Miriam für ihre emotionale Unterstützung.

Eidesstattliche Erklärung

Ich erkläre an Eides statt, dass ich die vorliegende Diplomarbeit selbstständig und ohne fremde Hilfe verfasst, andere als die angegebenen Quellen und Hilfsmittel nicht benutzt bzw. die wörtlich oder sinngemäß entnommenen Stellen als solche kenntlich gemacht habe.

Wien, am 13.02.2025

Valentin Linsbichler, BSc

Inhaltsverzeichnis

Abbildungsverzeichnis	ii
Quellcodeverzeichnis	iii
1 Einleitung	1
2 Illiquide Finanzmärkte	3
2.1 Stochastische Grundlagen	3
2.2 Liquide Finanzmärkte	6
2.3 Large Trader*innen	8
3 Mean-Variance Hedging	11
3.1 Liquider Fall	11
3.2 Illiquider Fall	13
4 Stabilität	19
4.1 Nutzenmaximierung	19
4.2 Illiquides Bachelier Modell	22
5 Verteilung im stabilen Regime	28
5.1 Verteilung des Wertprozesses	28
5.2 Verteilung des erwarteten Nutzens	30
6 Numerisches Experiment	40
6.1 Neuronale Netze	40
6.2 Implementierung	44
6.3 Ergebnisse	46
7 Code	51
8 Conclusio	55
Literatur	56

Abbildungsverzeichnis

3.1	Simulation liquides Bachelier Modell	14
3.2	Simulation illiquides Bachelier Modell	15
4.1	Polynom p im liquiden und im illiquiden Fall	25
4.2	Erwarteter Nutzen im liquiden und im illiquiden Fall	26
5.1	Dichtefunktion des Wertprozesses zur Maturität	30
5.2	Dichtefunktion des Nutzens des Wertprozesses zur Maturität	31
5.3	Verteilung des Schätzers Z_M für $M = 1.000$	37
5.4	Verteilung des Schätzers Z_M für $M = 100.000$	38
6.1	Struktur eines neuronalen Netzes [7, Abbildung 2.7]	41
6.2	Aufbau eines Neurons [7, Abbildung 2.5]	42
6.3	Vergleich von κ_i und $\hat{\kappa}_i$	46
6.4	Absolute Differenz zwischen κ_i und $\hat{\kappa}_i$	47
6.5	Beispiel eines Polynoms p vor Ausführung des neuronalen Netzes und dessen Maximum	48
6.6	Beispiel eines Polynoms p nach Ausführung des neuronalen Netzes	49
6.7	Realisationen des Wertprozesses eines stabilen Szenarios	50
6.8	Realisationen des Nutzens des Wertprozesses eines stabilen Szenarios	50

Quellcodeverzeichnis

7.1	Neuronales Netz	51
7.2	Vergleich der Illiquiditätsparameter	52
7.3	Regimewechsel	52
7.4	Polynom	53
7.5	Verteilung des Wertprozesses	53

1 Einleitung

Jedes Modell stellt einen Kompromiss zwischen Realität und Simplifizierung dar. Ob Vernachlässigung des Luftwiderstands, Unendlichkeit vorhandener Ressourcen oder idealisierte Verteilungseigenschaften, häufig werden bereits a priori vereinfachende Annahmen getroffen. Modelle können und sollen nicht die gesamte Komplexität unserer Welt abbilden. Eine besonders weitverbreitete Annahme in der Finanzmathematik, die in der Modellierung von Preisentwicklungen von Finanzprodukten oftmals vorausgesetzt wird, ist die Liquidität des betrachteten Finanzmarktes.

Insbesondere, wenn einzelne Akteur*innen mit überdurchschnittlich großem Handelsvolumen am Handel mit einem Finanzprodukt beteiligt sind, wird dessen Preisentwicklung stark vom Verhalten dieser Akteur*innen beeinflusst. Bei der Modellierung derartiger Handelsstrategien sowie den durch ebendiese beeinflussten Preisprozessen orientiert sich die vorliegende Arbeit an dem in [2] vorgestellten Zugang. In diesem Paper wird der in realen Märkten beobachtbare Rückkopplungseffekt durch einen sogenannten Illiquiditätsparameter quantifiziert. Schnell wird jedoch klar, dass die Existenz einer optimalen Handelsstrategie in diesem Fall bedroht ist. Wird der Rückkopplungseffekt überschätzt beziehungsweise der entsprechende Parameter zu hoch angesetzt, kann unter Umständen ein Kipppunkt überschritten werden. Ausgelöst durch die gegenseitige Beeinflussung von Preisprozess und Handelsstrategie führt dies zu einem nicht endenden Wachstum der optimalen Position. Die Folge dieser Strategie wäre eine Destabilisierung des Marktes.

Die beiden daraus resultierenden Szenarien bezeichnen wir einerseits als das stabile Regime, in dem eine optimale, stabile Handelsstrategie existiert und andererseits als das instabile Regime, in dem eine derartige Strategie nicht existiert. Insbesondere, wenn zum Auffinden des Illiquiditätsparameters ein neuronales Netz genutzt wird, stellt sich die Frage, ob die Stabilität des Regimes bereits vorab feststeht oder ob getrieben durch die entstehende Diffusion auch ein Übergang aus dem zu Beginn stabilen Regime ins instabile Regime möglich ist. Mit Hilfe des zu diesem Zweck implementierten neuronalen Netzes wird in der vorliegenden Arbeit nachgewiesen, dass mit positiver Wahrscheinlichkeit ein entsprechender Regimewechsel eintreten kann.

Ein weiterer integraler Bestandteil der vorliegenden Arbeit ist die Beschreibung der Verteilung des aus der Anwendung einer stabilen Handelsstrategie resultierenden Wertprozesses sowie des daraus erhaltenen Nutzens zur Maturität. Nach einer allgemeinen Beschreibung werden die beiden Verteilungen anhand des auf den illiquiden Fall ausgeweiteten Bachelier Modells konkretisiert und weiter untersucht. Zu diesem Zweck wird die exponentielle Nutzenfunktion herangezogen.

Das erste inhaltliche Kapitel dieser Arbeit widmet sich den illiquiden Finanzmärkten und insbesondere der Abgrenzung zum liquiden Fall. Neben später benötigten stochastischen Grundlagen wird darin auch der Begriff Large Trader*in diskutiert. Daran anschließend wird im nächsten Kapitel die Verallgemeinerung des klassischen Bachelier Modells auf den illiquiden Fall beschrieben. Das dadurch erhaltene illiquide Bachelier Modell zieht sich als anschauliches Beispiel durch die gesamte Arbeit. Darüber hinaus behandelt dieses Kapitel auch den Einfluss der Illiquidität auf das Auffinden von Hedgingstrategien unter dem Mean-Variance Ansatz. Das vierte Kapitel befasst sich mit dem Optimierungsproblem des aus einer Handelsstrategie resultierenden erwarteten Nutzens. Insbesondere wird im Zuge dessen diskutiert anhand welcher Kriterien die Existenz einer stabilen, optimalen Strategie überprüft werden kann. In Kapitel fünf werden die Verteilungen des Wertprozesses sowie dessen Nutzen und auch die Momente dieser Verteilungen beschrieben. Kapitel sechs beschreibt schließlich das implementierte neuronale Netz und die damit beobachteten Regimewechsel. Abschließend wird in Kapitel sieben der verwendete Code dargestellt.

2 Illiquide Finanzmärkte

In diesem Kapitel werden einige essenzielle Konzepte und Resultate der stochastischen Analysis und Finanzmathematik vorgestellt, denen bei der Modellierung und Analyse von Finanzmärkten und insbesondere Handelsstrategien eine zentrale Bedeutung zukommt. Zunächst werden in Abschnitt 2.1 einige wichtige Begriffe erläutert und für die folgenden Kapitel relevante Sätze der stochastischen Analysis vorgestellt. Anschließend werden in Kapitel 2.2 Annahmen für liquide Finanzmärkte sowie eine Differenzierung zwischen vollständigen und unvollständigen Märkten diskutiert. In Kapitel 2.3 wird schließlich das Konzept der Large Trader*innen eingeführt, deren Entscheidungen Einfluss auf den Preisprozess der zugrunde liegenden Finanzprodukte haben.

2.1 Stochastische Grundlagen

Es sei $(\Omega, \mathcal{F}, \mathbb{P})$ ein Wahrscheinlichkeitsraum, der die Grundlage für alle betrachteten stochastischen Prozesse bildet. Darüber hinaus sei $T \in \mathbb{R}_+$ ein endlicher Zeithorizont und es bezeichne $\mathbb{F} = (\mathcal{F}_t)_{t \in [0, T]}$ eine Filtration auf dem Zeitintervall $[0, T]$. Eine Schlüsselrolle bei der Bewertung und Modellierung von Finanzprodukten nehmen die Martingale ein.

Definition 2.1.1. Ein integrierbarer und $\mathbb{F}$ -adaptierter Prozess $M = (M_t)_{t \in [0, T]}$ mit Werten in $\mathbb{R}^d$, der die Eigenschaft

$$M_s \stackrel{f.s.}{=} \mathbb{E}[M_t \mid \mathcal{F}_s] \quad \forall s, t \in [0, T] : s \leq t$$

erfüllt, heißt *Martingal* bezüglich des Wahrscheinlichkeitsmaßes $\mathbb{P}$. Ist dieser Prozess zusätzlich rechtsstetig und existiert eine $\mathbb{R}_+ \cup \{\infty\}$ -wertige, punktweise steigende Folge von $\mathbb{F}^+$ -Stoppzeiten, die punktweise $\lim_{n \rightarrow \infty} \tau_n = T$ erfüllt, sodass der Prozess $M^{\tau_n} - M_0$ für alle $n \in \mathbb{N}$ ein $\mathbb{F}$ -Martingal bezüglich $\mathbb{P}$ ist, so handelt es sich dabei um ein *lokales Martingal* bezüglich $\mathbb{P}$. Dabei bezeichnet M^{τ_n} den entsprechenden gestoppten Prozess, welcher durch $M^{\tau_n} = M_{\tau_n \wedge t}$ gegeben ist.

Die Martingale sind also genau jene Prozesse, deren Erwartungswert für einen zukünftigen Zeitpunkt t dem Wert zum Beobachtungszeitpunkt s entspricht. Unabhängig von der Entwicklung der einzelnen realisierten Pfade erwarten wir mit dem zum Zeitpunkt s bestehenden Informationsstand daher weder Verlust noch Zuwachs. Dem gegenüber stehen die Prozesse lokal endlicher Variation (vgl. [18, Kapitel 5]).

Definition 2.1.2. Sei $A : \mathbb{R}_+ \times \Omega \rightarrow \mathbb{R}^d$ ein stochastischer Prozess und $p \in \mathbb{R}_+$. Dann definieren wir die p -Variation des Prozesses A auf dem Intervall $[0, t]$ als

$$\mathbb{V}_{A(\omega), p}([0, t]) := \lim_{\epsilon \rightarrow 0} \sup_{\pi_\epsilon} \sum_{i=1}^k \|A_{t_i}(\omega) - A_{t_{i-1}}(\omega)\|^p,$$

wobei sich das Supremum über alle Partitionen $\pi_\epsilon = (t_0, t_1, \dots, t_k)$ des Zeitintervalls $[0, t]$ erstreckt, bei denen der Abstand von je zwei Zeitpunkten t_i und t_j kleiner als ϵ ist. Der stochastische Prozess A ist von lokal endlicher p -Variation, wenn

$$\mathbb{V}_{A(\omega),p}([0, t]) < \infty \quad (2.1)$$

für alle $t \in [0, T]$ und $\omega \in \Omega$. Für den Spezialfall $p = 1$ bezeichnen wir $\mathbb{V}_{A(\omega),p}([0, t])$ als $\mathbb{V}_{A(\omega)}([0, t])$ und den Prozess A als von lokal endlicher Variation, sofern 2.1 erfüllt ist. Für den Spezialfall $p = 2$ bezeichnen wir $\mathbb{V}_{A(\omega),2}([0, T])$ als quadratische Variation von A .

Anschaulich sind die Prozesse lokal endlicher Variation genau jene Prozesse, deren Pfade auf beliebigen, kompakten Intervallen jeweils nur endliche Länge haben. Eine wichtige Verallgemeinerung der oben beschriebenen Prozesse stellt die Klasse der Semimartingale dar.

Definition 2.1.3. Ein $\mathbb{R}^d$ -wertiger Prozess X heißt stetiges Semimartingal, wenn er als Summe

$$X = A + M$$

geschrieben werden kann. Dabei bezeichnet A einen $\mathbb{R}^d$ -wertigen, $\mathbb{F}$ -adaptierten, stetigen Prozess von lokal endlicher Variation, der $A_0 = 0$ erfüllt und M ein $\mathbb{R}^d$ -wertiges, stetiges lokales Martingal.

Die in Definition 2.1.3 beschriebene Zerlegung eines Semimartingals ist laut [18, Bemerkung 5.106] bis auf Ununterscheidbarkeit eindeutig und wird im Nachfolgenden auch als kanonische Zerlegung des Semimartingals X bezeichnet.

Definition 2.1.4. Der quadratische Kovariationsprozess zweier Semimartingale X und Y ist definiert als

$$[X, Y] = \frac{1}{2} ([X + Y] - [X] - [Y]).$$

Dabei symbolisiert $[X]$ die quadratische Variation des Prozesses X [17, Kapitel 2.5].

Definition 2.1.5. Für ein stetiges Semimartingal X ist das zugehörige stochastische Exponential als

$$\mathcal{E}(X) = \exp \left(X - \frac{1}{2} [X, X] \right)$$

definiert [18, Definition 6.46].

Die stochastischen Exponentiale sind für unsere Zwecke von besonderem Interesse, da sie Kandidaten für Dichteprozesse sind, die bei Maßwechseln für die risikoneutrale Bewertung benötigt werden. Selbst wenn es sich bei X um ein Martingal handelt, kann im Allgemeinen jedoch nicht davon ausgegangen werden, dass $\mathcal{E}(X)$ ebenfalls ein Martingal ist (vgl. [17, Kapitel 2.5]).

Definition 2.1.6. Die vorhersehbare σ -Algebra $\mathcal{P}$ auf $\mathbb{R}_+ \times \Omega$ ist die von den $\mathbb{F}$ -adaptierten, linksstetigen Prozessen generierte σ -Algebra. Ein $\mathcal{P}$ -messbarer Prozess heißt vorhersehbar [17, Definition 2.7].

Definition 2.1.7. Für einen stetigen, adaptierten Prozess von lokal endlicher Variation mit $A_0 = 0$, sei $L(A)$ der Raum der progressiven Prozesse $V : \mathbb{R}_+ \times \Omega \rightarrow \mathbb{R}$, sodass

$$\int_0^t |V_s| d\mathbb{V}_A(s) < \infty, \quad t \in \mathbb{R}_+,$$

wobei $\mathbb{V}_A(s)$ die Totalvariation des Prozesses A auf $[0, s]$ ist. Für ein stetiges lokales Martingal bezeichne $L(M)$ den Raum aller progressiven Prozesse $V : \mathbb{R}_+ \times \Omega \rightarrow \mathbb{R}$, sodass

$$\int_0^t |V_s|^2 d[M]_s < \infty, \quad t \in \mathbb{R}_+.$$

Für ein stetiges Semimartingal $X = A + M$ definieren wir

$$L(X) = L(A) \cap L(M).$$

Ein zentraler Bestandteil der Modellierung von Handelsstrategien und ihren zugehörigen Wertprozessen ist das stochastische Integral. Die Konstruktion dieses Operators ist jedoch ein aufwendiger Prozess, der üblicherweise in mehreren aufbauenden Schritten erfolgt. Das für unsere Zwecke wichtigste Resultat in diesem Prozess ist die Definition eines stochastischen Integrals für Semimartingale. Aufbauend auf Definition 2.1.7 können wir einen solchen Prozess definieren.

Definition 2.1.8. Der stochastische Integralprozess $V \in L(X)$ bezüglich eines $\mathbb{R}$ -wertigen, stetigen Semimartingals X ist definiert als

$$V \cdot X = \int_0^\cdot V_s dX_s = \int_0^\cdot V_s dA_s + \int_0^\cdot V_s dM_s.$$

Für eine vollständige Herleitung des stochastischen Integrals sowie eine ausführliche Beschreibung der einzelnen Bestandteile der bisherigen Definitionen sei auf [18] verwiesen.

Der nachfolgende Satz gibt Auskunft darüber, wie sich die Dynamik eines stochastischen Prozesses bei einem Maßwechsel ändert. Insbesondere wird beschrieben, wie der zugrundeliegende Prozess modifiziert werden muss, um weiterhin Martingal zu bleiben [18, Satz 7.41].

Satz 2.1.9 (Satz von Girsanov). Es sei $\mathbb{Q}$ ein Wahrscheinlichkeitsmaß auf $(\Omega, \mathcal{F})$ und Z ein adaptierter, stetiger und strikt positiver Prozess, sodass $d\mathbb{Q} = Z_t d\mathbb{P}$ auf $\mathcal{F}_t$ für alle $t \in [0, T]$ gilt. Sei außerdem X ein $\mathbb{R}^d$ -wertiges $\mathbb{P}$ -Semimartingal mit der kanonischen $\mathbb{P}$ -Zerlegung $X = A + M$. Dann ist X ein $\mathbb{R}^d$ -wertiges, stetiges $\mathbb{Q}$ -Semimartingal auf $\mathbb{R}_+$ mit kanonischer $\mathbb{Q}$ -Zerlegung $X = \tilde{A} + \tilde{M}$. Dabei sind $\tilde{A}$ und $\tilde{M}$ gegeben durch

$$\tilde{A} = A + \int_0^\cdot \frac{1}{Z_s} d[M, Z]_s \quad \text{und} \quad \tilde{M} = M - \int_0^\cdot \frac{1}{Z_s} d[M, Z]_s.$$

Grundlage vieler Modelle in der Finanzmathematik ist eine Brownsche Bewegung. Zu diesen Modellen zählt beispielsweise auch das im späteren Verlauf dieser Arbeit behandelte illiquide Bachelier Modell. Die folgende Anwendung des Satzes von Girsanov auf eine Brownsche Bewegung W wird in [18, Korollar 7.44] beschrieben.

Korollar 2.1.10. *Sei W eine Brownsche Bewegung und $V \in L(W)$ ein Prozess, sodass $\mathcal{E}(V \cdot W)$ ein Martingal ist, welches durch eine Zufallsvariable $\mathcal{E}(V \cdot W)_\infty$ geschlossen werden kann. Dann handelt es sich bei $\mathbb{Q}(F) := \mathbb{E}[\mathcal{E}(V \cdot W)_\infty \mathbb{1}_F]$ für $F \in \mathcal{F}$ um ein Wahrscheinlichkeitsmaß auf $(\Omega, \mathcal{F})$ mit $\mathbb{Q} \ll \mathbb{P}$ auf $\mathcal{F}$. Außerdem ist $\tilde{W} := W - \int_0^\cdot V_s^T ds$ eine $(\mathbb{F}, \mathbb{Q})$ -Brownsche Bewegung.*

Ein weiteres sehr nützliches Resultat in diesem Kontext ist die Charakterisierung nach Levy. Diese ermöglicht es verhältnismäßig einfach zu überprüfen, ob ein gegebener Prozess eine Brownsche Bewegung ist. Nachdem Brownsche Bewegungen über eine Reihe nützlicher Eigenschaften verfügen, stellt sich diese alternative Charakterisierung als durchaus hilfreich dar [18, Korollar 7.2].

Satz 2.1.11 (Lévy Charakterisierung einer Brownschen Bewegungen). *Es sei W ein $\mathbb{R}^d$ -wertiger Prozess auf einem filtrierten Wahrscheinlichkeitsraum $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$. Dann handelt es sich bei W genau dann um eine in W_0 startende $(\mathbb{F}, \mathbb{P})$ -Brownsche Bewegung, wenn W ein stetiges, lokales $\mathbb{F}$ -Martingal ist, das bis auf Ununterscheidbarkeit für alle $t \in [0, T]$*

$$[W]_t = tI_d$$

erfüllt.

2.2 Liquide Finanzmärkte

Ein Finanzmarkt ist ein komplexes System. Wie in der Modellierung komplexer Systeme üblich, werden daher eine Reihe vereinfachender Annahmen getroffen. Unter diesen können Fragestellungen und Herausforderungen der realen Welt mit Hilfe theoretischer Konzepte aufgegriffen und modelliert werden. Da viele dieser Annahmen in realen Finanzmärkten nicht oder nur eingeschränkt zutreffen, ist bei der Interpretation und Anwendung theoretischer Resultate große Vorsicht geboten (vgl. [16]).

Zu den gängigen Voraussetzungen in Standardmodellen zählen beispielsweise das Ausbleiben von Transaktionskosten sowie die Annahme von Informationsgleichheit zwischen allen Akteur*innen (vgl. [6]). Letztere Annahme stellt unter anderem sicher, dass sich keine agierende Partei einen Vorteil durch zusätzliche, der breiten Öffentlichkeit unzugänglichen Informationen verschafft. Konkret wird damit sichergestellt, dass im vorhandenen Modell keine Insider Informationen kursieren beziehungsweise ausgenutzt werden. Wir gehen weiters davon aus, dass die treibende Motivation aller Marktteilnehmer*innen die Maximierung des erwarteten Nutzens ihres Endvermögens ist (vgl. [13, Kapitel 10]). Im Fall mehrerer unterschiedlicher Handlungsmöglichkeiten wird die Entscheidung daher stets auf jene fallen, die mit dem aktuellen Informationsstand zu maximalem Nutzen führt. Willkürliche Handlungen sind somit bis zu einem gewissen Grad ausgeschlossen.

Eine weitere durchaus verbreitete Annahme ist, dass Marktteilnehmer*innen und ihre Handelsstrategien keinen Einfluss auf die Preise der am Finanzmarkt gehandelten Güter haben. Insbesondere wird damit unterstellt, dass selbst wenn ein*e Marktteilnehmer*in beispielsweise große Mengen einer Aktie kauft oder verkauft, die Aktie dennoch zum selben Preis weiter gehandelt wird. Implizit wird dabei außerdem angenommen, dass stets eine ausreichend große Anzahl an Käufer*innen und Verkäufer*innen bereitsteht, um das entsprechende Produkt zum gegebenen Marktpreis zu kaufen beziehungsweise verkaufen. Erfüllt ein Markt diese Voraussetzungen, sprechen wir von einem liquiden Finanzmarkt.

In realen Finanzmärkten sind diese idealisierenden Annahmen nicht erfüllt. Für gewöhnlich haben große Käufe beziehungsweise Verkäufe einen sichtbaren Einfluss auf den Marktpreis und selbst kleinere Handlungen beeinflussen den entsprechenden Preis des gehandelten Produkts. Aus diesem Grund wird in Kapitel 2.3 die Modellierung des Preisprozesses behandelt, die entsteht, wenn wir diese Annahme zumindest teilweise aussetzen.

Unter den beschriebenen Annahmen können wir nun ein konkretes risikobehaftetes Finanzprodukt genauer betrachten. Wir modellieren die Preisentwicklung dieses Produkts mit Hilfe eines Semimartingals S . Dabei könnte es sich zum Beispiel um eine Aktie handeln.

Definition 2.2.1. Eine Handelsstrategie θ ist ein S -integrierbarer Prozess. Der zugehörige Wertprozess $V = V(c, \theta)$ mit Anfangskapital c und Handelsstrategie θ in dem Vermögensgegenstand S ist gegeben als

$$V = c + \int \theta dS.$$

Die Handelsstrategie θ_t kann interpretiert werden als Anzahl der Anteile des Finanzprodukts, die ein*e Investor*in zum Zeitpunkt t besitzt. Das stochastische Integral $\int_0^T \theta_t dS_t$ gibt folglich den Gewinn beziehungsweise Verlust im betrachteten Zeitintervall $[0, T]$ an, den der*die Investor*in erhält, wenn er*sie die Handelsstrategie θ verfolgt (vgl. [17, Kapitel 3.1]).

Definition 2.2.2. Eine Handelsstrategie θ , die für alle $t \in [0, T]$

$$V_t = V_0 + \int_0^t \theta_u dS_u$$

erfüllt, heißt selbstfinanzierende Strategie [17, Definition 3.1]

Das bereits erwähnte statistische Maß $\mathbb{P}$ spiegelt die Entwicklungen diverser Finanzprodukte in der realen Welt wider und bildet somit eine wichtige Grundlage für viele unserer Überlegungen. Zur Bewertung verschiedener Finanzprodukte sowie zum Modellieren von Handelsstrategien wird jedoch häufig auf das risikoneutrale Wahrscheinlichkeitsmaß $\mathbb{Q}$ zurückgegriffen (vgl. [17, Kapitel 2.4]). Im Zuge dessen ist auch die nachfolgende Definition von Bedeutung:

Definition 2.2.3. Ein bezüglich $\mathbb{P}$ absolutstetiges Wahrscheinlichkeitsmaß $\mathbb{Q}$ ist ein Martingalmaß für S , wenn S ein lokales $\mathbb{Q}$ -Martingal ist. Es handelt sich darüber hinaus um ein äquivalentes Martingalmaß, wenn $\mathbb{Q}$ äquivalent zu $\mathbb{P}$ ist [17, Definition 3.2].

Definition 2.2.4. Ein Claim ist eine $\mathcal{F}_T$ -messbare Zufallsvariable. Ein Claim H heißt erreichbar, wenn eine Konstante c und eine Handelsstrategie θ existieren, sodass

$$H = c + \int_0^T \theta_t dS_t.$$

Ein Finanzmarkt, in dem alle beschränkten Claims erreichbar sind, heißt vollständiger Markt. Ein Finanzmarkt, der nicht vollständig ist, heißt unvollständig [17, Kapitel 3.2].

Im Zuge dessen wird die Handelsstrategie θ oft auch replizierende Strategie oder Hedgingstrategie genannt, da durch Verfolgen von θ genau der Claim H repliziert wird. Neben den erreichbaren Claims können in unvollständigen Märkten auch Claims auftreten, die nicht erreichbar sind. Das sind also genau jene Claims, bei denen keine Hedgingstrategie existiert, mit der man sich restlos absichern kann. In diesen Fällen verbleibt ein gewisses Restrisiko, das es zu minimieren gilt. Ein vielversprechender Zugang zu dem daraus resultierenden Optimierungsproblem wird in Kapitel 3 beschrieben.

2.3 Large Trader*innen

Reale Marktpreise werden unter anderem durch Entscheidungen von Investor*innen beeinflusst. Je größer das Handelsvolumen, desto größer ist im Allgemeinen auch der damit einhergehende Einfluss. Eine Maßnahme, um Finanzmärkte realistischer zu modellieren, ist daher die Einführung sogenannter Large Trader*innen. Ein*e Large Trader*in ist ein*e Investor*in, dessen*deren Trades die Preise beeinflussen (vgl. [9]). Dabei wird davon ausgegangen, dass Marktpreise nur durch Handelsstrategien mit einem hinreichend großen Handelsvolumen beeinflusst werden. Marktteilnehmer*innen, die die Preise nicht mitbeeinflussen und somit nur zu vorgegebenen Preisen handeln können, werden als kleine Trader*innen bezeichnet. Da Marktteilnehmer*innen mit viel Kapital den größten Einfluss auf die Marktpreise haben, ist es durchaus gerechtfertigt, den Fokus vorerst auf diese zu legen. Dennoch sei betont, dass auch kleine Trader*innen einen Einfluss auf Marktpreise haben. Bei einem*einer Large Trader*in kann es sich beispielsweise um ein Versicherungsunternehmen oder ähnliches Finanzinstitut handeln. Finanzmärkte, bei deren Modellierung wir einen oder mehrere Large Trader*innen berücksichtigen, bezeichnen wir als illiquide Märkte.

Wir wollen nun die Preisentwicklung eines Finanzprodukts unter dem Einfluss eines*einer Large Traders*Traderin modellieren. In der Literatur finden sich mehrere Zugänge mit diesem Ziel. Im Folgenden arbeiten wir mit dem von Rheinländer und Blümmel in [2] vorgestellten Ansatz. Ausgangspunkt dafür sind die einfachen Handelsstrategien.

Definition 2.3.1. Es sei θ ein vorhersehbarer Prozess, der über eine Darstellung

$$\theta(t) = \theta_{-1} \mathbb{1}_{\{0\}}(t) + \sum_{i=0}^n \theta_i \mathbb{1}_{(\tau_i, \tau_{i+1}]}(t)$$

verfügt. Dabei sei $0 = \tau_0 \leq \tau_1 \leq \dots \leq \tau_{n+1} = T$ eine endliche Folge von $\mathbb{F}$ -Stoppzeiten, $\theta_{-1} \in \mathbb{R}^d$ und θ_i eine beschränkte, $\mathcal{F}_{\tau_i}$ -messbare, $\mathbb{R}^d$ -wertige Zufallsvariable für alle $i \in$

$\{0, \dots, n\}$. Sofern die Zufallsvariablen θ_i nur endlich viele Werte annehmen, nennen wir θ eine erweiterte, einfache Strategie. Nehmen darüber hinaus auch die Stoppzeiten τ_i nur endlich viele Werte an, nennen wir θ eine einfache Strategie. Den Raum aller $\mathbb{R}^d$ -wertigen, einfachen Strategien bezeichnen wir als $\mathcal{S}(\mathbb{R}^d)$.

Die Stoppzeiten $(\tau_i)_{i \leq n+1}$ sind dabei als zukünftige Handelszeitpunkte der Strategie θ zu interpretieren. Der topologische Abschluss des Raumes $\mathcal{S}(\mathbb{R})$ ist der Raum aller $\mathbb{R}$ -wertigen, adaptierten, càdlàg Prozesse [2, Kapitel 2].

Definition 2.3.2. Wir bezeichnen den Raum der $\mathbb{R}$ -wertigen, adaptierten, càdlàg Prozesse $\mathbb{L}(\mathbb{R})$ als Raum der Large Trader*innen Strategien und den Raum $\mathcal{S}(\mathbb{R})$ als Raum der einfachen Large Trader*innen Strategien.

Wie in Kapitel 2.2 beschrieben, modellieren wir den Preisprozess eines risikobehafteten Finanzprodukts in liquiden Finanzmärkten mit Hilfe eines Semimartingals. Diese Vorgehensweise möchten wir nun auf den illiquiden Fall ausweiten. Anstelle von einzelnen Semimartingalen benötigen wir zu diesem Zweck Familien von Semimartingalen.

Definition 2.3.3. Eine Familie $P := (P(\theta, \cdot))_{\theta \in \mathbb{R}^d}$ heißt primale Familie, wenn

$$P(\theta, \cdot) := (P(\theta, t))_{t \in [0, T]},$$

für jedes $\theta \in \mathbb{R}^d$ ein $\mathbb{R}$ -wertiges, stetiges, adaptiertes Semimartingal ist.

Mit dieser Verallgemeinerung können wir den durch eine Large Trader*innen Strategie θ beeinflussten Preisprozess S eines Finanzprodukts nun analog zum liquiden Fall definieren wie folgt:

Definition 2.3.4. Es sei $\theta \in \mathcal{S}(\mathbb{R})$ eine einfache Large Trader*innen Strategie und P eine primale Familie. Dann ist der durch θ beeinflusste Preisprozess gegeben als

$$S_t^\theta := \int_0^t P(\theta, ds) := P(\theta, 0) + \sum_{i=0}^n (P(\theta, \tau_{i+1} \wedge t) - P(\theta, \tau_i \wedge t)).$$

Der Einfluss der Handelsstrategie θ hängt daher von der primalen Familie $(P(\theta, \cdot))_{\theta \in \mathbb{R}^d}$ und insbesondere von der Regularität und Glattheit dieser ab. Eine ausführliche Diskussion, welche Voraussetzungen $(P(\theta, \cdot))_{\theta \in \mathbb{R}^d}$ erfüllen muss, um beispielsweise die Abwesenheit von Arbitragemöglichkeiten im Finanzmarkt zu bewahren, wird unter anderem in [2] und [1] genauer ausgeführt. Ein vielversprechender Zugang bedient sich dabei dem Konzept der stark nichtlinearen Integratoren, das auf [4] zurückgeht.

Annahme 2.3.5. Es gelte:

1. $(P(\theta, \cdot))_{\theta \in \mathbb{R}^d}$ ist ein stark nichtlinearer Integrator.
2. Für alle $\theta \in \mathbb{R}^d$ existiert ein äquivalentes, lokales Martingalmaß $\mathbb{Q}^\theta$ für $P(\theta, \cdot)$ bezüglich $\mathbb{F}$.

Bemerkung 2.3.6. *Aufbauend auf 2.3.4 kann auch ein durch eine Large Trader*innen Strategie $\theta \in \mathbb{L}(\mathbb{R}) \setminus \mathcal{S}(\mathbb{R})$ beeinflusster Preisprozess betrachtet werden. Sofern θ ein Element von $\mathbb{L}(\mathbb{R}) \setminus \mathcal{S}(\mathbb{R})$ ist, existiert eine Folge $(\theta^n)_{n \geq 1} \subset \mathcal{S}(\mathbb{R})$, die in der ucp-Topologie gegen θ konvergiert. Da es sich bei den Folgegliedern von $(\theta^n)_{n \geq 1}$ um einfache Large Trader*innen Strategien handelt, können wir eine entsprechende Folge von Preisprozessen wie in 2.3.4 konstruieren. Laut [2, Theorem 3.3] handelt es sich bei der Folge $(S^{\theta^n})_{n \in \mathbb{N}}$ unter Annahme 2.3.5 um eine Cauchyfolge und laut [2, Lemma 3.2] konvergiert diese Cauchyfolge in der Semimartingal Topologie gegen einen Prozess S^θ . Handelt es sich bei θ um eine Strategie aus $\mathbb{L}(\mathbb{R}) \setminus \mathcal{S}(\mathbb{R})$, bezeichnen wir S^θ als den zu θ gehörigen Preisprozess. Somit existiert für jede Large Trader*innen Strategie ein entsprechender beeinflusster Preisprozess.*

3 Mean-Variance Hedging

Während in vollständigen Finanzmärkten jeder beschränkte Claim durch eine Hedgingstrategie θ repliziert werden kann, existieren in unvollständigen Märkten auch Claims, bei denen diese Möglichkeit nicht besteht. Unabhängig von der gewählten Handelsstrategie verbleibt in diesem Fall ein nicht absicherbares Restrisiko, das es zu minimieren gilt. Beim Mean-Variance Hedging wird die Diskrepanz zwischen der Verpflichtung, die sich aus der Zufallsvariable H ergibt und dem aus der Hedgingstrategie θ resultierendem Wertprozess unter einem quadratischen Kriterium minimiert (vgl. [15, Kapitel 1]). In Kapitel 3.1 betrachten wir dieses Optimierungsproblem unter den üblichen Annahmen, die an einen liquiden Finanzmarkt gestellt werden. Im Zuge dessen wird das Bachelier Modell als anschauliches Beispiel dienen. Anschließend wird das Setting aus Kapitel 3.1 in Kapitel 3.2 um eine*n Large Trader*in erweitert. Dadurch ergibt sich als Erweiterung zum liquiden Fall das illiquide Bachelier Modell.

3.1 Liquider Fall

Das bereits erwähnte quadratische Kriterium kann theoretisch auch unter dem statistischen Maß $\mathbb{P}$ betrachtet werden. Für eine ausführliche Behandlung des Problems unter dem statistischen Maß sei beispielsweise auf [19] verwiesen. Für unsere Zwecke ist es jedoch ausreichend, das beschriebene Optimierungsproblem unter einem risikoneutralen Maß zu diskutieren. Folglich nehmen wir an, dass es sich bei dem Preisprozess S um ein quadratisch integrierbares, lokales Martingal handelt (vgl. [17, Kapitel 6]).

Bevor wir uns mit der Existenz und dem Auffinden einer optimalen Handelsstrategie beschäftigen, müssen wir uns der Frage widmen, welche Strategien wir als zulässig betrachten beziehungsweise über welche Grundmenge wir das Minimum bestimmen wollen. Dafür betrachten wir zunächst den Raum der für das Mean-Variance Hedging relevanten Martingale etwas genauer.

Definition 3.1.1. Der Raum $\mathcal{M}^2$ ist der Hilbertraum der L^2 -Martingale, die $\mathbb{E}[[M]_T] < \infty$ erfüllen. Dieser ist versehen mit dem Skalarprodukt

$$(M, N) := \mathbb{E}[M_T N_T].$$

Zwei Martingale $M, N \in \mathcal{M}^2$ heißen dabei orthogonal, wenn $(M, N) = 0$. Die Martingale M und N heißen stark orthogonal, sofern MN selbst wieder ein Martingal ist. Der Raum $\mathcal{M}_0^2$ besteht aus allen $M \in \mathcal{M}^2$, die $M_0 = 0$ erfüllen und der Raum $\mathcal{M}_{loc}^2$ besteht aus allen Prozessen, die lokal in $\mathcal{M}^2$ liegen [17, Definition 2.25].

Für Elemente aus diesem Raum lässt sich darüber hinaus auch das speziell im Hedgingkontext relevante stochastische Integral einer Handelsstrategie bezüglich dem entsprechenden Martingal in Form einer Norm definieren.

Definition 3.1.2. Sei $M \in \mathcal{M}_{loc}^2$. Der Raum $L^2(M)$ besteht aus allen vorhersehbaren Prozessen θ , die

$$\|\theta\|_{L^2(M)} := \mathbb{E} \left[\int_0^T \theta_t^2 d[M]_t \right]^{\frac{1}{2}} < \infty \quad (3.1)$$

erfüllen. Der Raum aller stochastischen Integralprozesse $\int \theta dM$, sodass $\theta \in L^2(M)$, wird als von M generierter, stabiler Subraum $\mathcal{S}(M)$ bezeichnet [17, Definition 2.28].

In Anlehnung an die in 3.1 erklärte Norm definieren wir schließlich den Raum der zulässigen Handelsstrategien [17, Kapitel 6].

Definition 3.1.3. Eine Handelsstrategie θ heißt zulässig, wenn $\theta \in L^2(S)$.

Das daraus resultierende Optimierungsproblem lautet daher

$$\min_{\theta \in L^2, c \in \mathbb{R}} \mathbb{E}_{\mathbb{Q}} \left[\left(H - c - \int_0^T \theta_t dS_t \right)^2 \right]. \quad (3.2)$$

Unter gewissen Bedingungen lässt sich 3.2 sogar auf eindeutige Weise lösen. Um diese Lösung für den allgemeinen Fall zu formalisieren, benötigen wir die Kunita-Watanabe Zerlegung des zugrundeliegenden Martingals [17, Satz 2.31].

Satz 3.1.4 (Kunita-Watanabe Zerlegung). Es sei $M \in \mathcal{M}_{loc}^2$ und $N \in \mathcal{M}^2$. Dann existiert eine eindeutige Zerlegung

$$N = N_0 + \int \theta dM + L,$$

wobei $\theta \in L^2(M)$ und $L \in \mathcal{M}_0^2$ stark orthogonal zu allen Elementen aus $\mathcal{S}(M)$ ist.

Gegeben sei ein Claim $H \in L^2(\mathcal{F}, \mathbb{P})$. Darauf aufbauend definieren wir nun einen Prozess V , durch den auf die Information zum jeweiligen Zeitpunkt t bedingten Erwartungswert des Claims:

$$V_t := \mathbb{E}[H \mid \mathcal{F}_t] \quad 0 \leq t \leq T.$$

Laut [18, Beispiel 4.4] handelt es sich bei dem Prozess V um ein Martingal. Daraus folgt, dass $V \in \mathcal{M}^2$. Unter den Annahmen eines liquiden Finanzmarktes gibt der folgende Satz aus [17, Kapitel 6] Auskunft über die Existenz, Form und Eindeutigkeit einer optimalen, zulässigen Handelsstrategie sowie eines optimalen Startkapitals.

Satz 3.1.5 (Optimale Mean-Variance Hedging Strategie). Es sei

$$V = \mathbb{E}(H) + \int \theta^H dS + L \quad (3.3)$$

die Kunita-Watanabe Zerlegung von V , wobei $\theta^H \in L^2$ und $L \in \mathcal{M}_0^2$. Dann sind die optimale Strategie θ^* und das optimale Startkapital c^* , die 3.2 minimieren, gegeben durch $\theta^* = \theta^H$

und $c^* = \mathbb{E}(H)$. Die optimale Strategie ist darüber hinaus eindeutig in dem Sinne, dass für je zwei optimale Strategien θ^* und ψ^* die resultierenden stochastischen Integralprozesse ununterscheidbar sind. Das ist äquivalent zu

$$\int (\theta^* - \psi^*)^2 d[S] = 0.$$

In Darstellung 3.3 kann $c^* + \int \theta^* dS$ nun als erreichbarer Anteil interpretiert werden, während L das nicht absicherbare, verbleibende Restrisiko beschreibt. Ist der Claim H erreichbar, verschwindet L somit. Wenn H jedoch nicht erreichbar ist, wird der Erwartungswert von L durch die genannte Strategie minimiert.

Das Bachelier Modell gehört zu den ersten Modellen, die die Preisentwicklung eines Finanzprodukts anhand einer Brownschen Bewegung dargestellt haben. Ausgangspunkt im klassischen Bachelier Modell ist ein deterministischer Drift μ und ein konstanter Volatilitätsparameter $\sigma > 0$. Der Prozess $(W_t)_{t \in [0, T]}$ beschreibe eine Brownsche Bewegung. Des Weiteren nehmen wir an, dass für den anzuwendenden Zinssatz $r = 0$ gilt. Dann ist die Preisentwicklung im Bachelier Modell durch

$$S_t = S_0 + \mu t + \sigma W_t \quad (3.4)$$

gegeben (vgl. [14]). Der Parameter S_0 beschreibt den ebenfalls vorab bekannten Anfangspreis des betrachteten Finanzprodukts. Zur Veranschaulichung ist im Folgenden ein einzelner realisierter Pfad des Prozesses S visualisiert. In diesem konkreten Beispiel startet das zugrunde liegende Finanzprodukt mit einem Preis von $S_0 = 100$ zum Zeitpunkt $t = 0$. Die rote Linie in Abbildung 3.1 veranschaulicht zum Vergleich die rein deterministische, hypothetische Preisentwicklung von S vorausgesetzt $\sigma = 0$. In diesem ersten Beispiel nehmen wir eine Maturität von einem Jahr an und folgen der Konvention, dass ein Jahr aus 252 Handelstagen besteht.

Äquivalent zu der Formulierung in 3.4 kann die Preisentwicklung auch als stochastische Differentialgleichung

$$dS_t = \mu dt + \sigma dW_t. \quad (3.5)$$

ausgedrückt werden.

3.2 Illiquider Fall

Wir wollen nun die in Kapitel 2.3 eingeführte Theorie der Large Trader*innen auf das Bachelier Modell anwenden. Die Handelsstrategie des*der Large Traders*Traderin soll also einen direkten Einfluss auf die Preisentwicklung haben. Im Folgenden gehen wir davon aus, dass sich dieser Einfluss auf den Drift bezieht. Die Volatilität bleibt unverändert konstant. Aus 3.5 wird somit

$$dS_t^\theta = \mu(\theta) dt + \sigma dW_t,$$

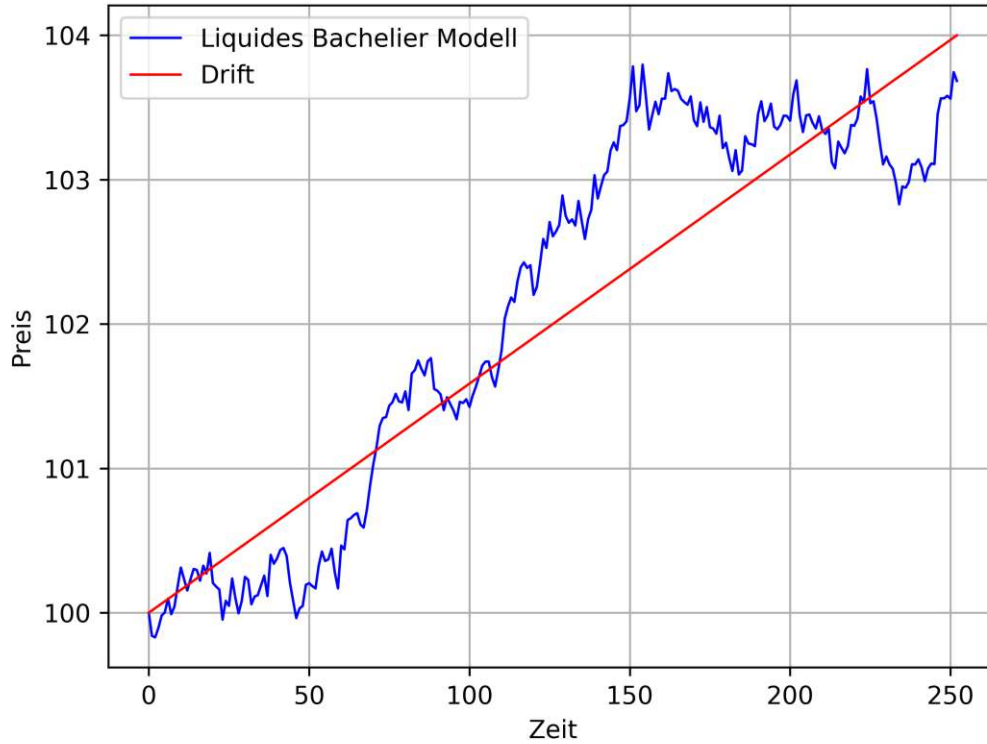


Abbildung 3.1: Simulation eines einzelnen Pfades in einem Bachelier Modell unter liquiden Marktannahmen mit Drift $\mu = 4$, $\sigma = 2$ und $S_0 = 100$

wobei $\mu(\theta)$ nun eine von der Strategie θ abhängige Driftfunktion ist. Den beeinflussten Drift modellieren wir als eine lineare Funktion, bestehend aus einem konstanten Driftterm und der gewichteten, im Allgemeinen zeitabhängigen Handelsstrategie. Somit erhalten wir

$$\mu(\theta) = \mu + \kappa\theta.$$

Der Parameter κ wird in diesem Zusammenhang Illiquiditätsparameter genannt. Er gibt Auskunft darüber, wie stark der Einfluss der Large Trader*innen Strategie ist. Durch Zusammenfügen obiger Gleichungen erhalten wir das illiquide Bachelier Modell (vgl. [2])

$$dS_t^\theta = (\mu + \kappa\theta) dt + \sigma dW_t.$$

In Abbildung 3.2 ist die Realisation aus Abbildung 3.1 im illiquiden Bachelier Modell dargestellt. Zu Demonstrationszwecken wurde hierbei angenommen, dass der*die Large Trader*in zu Beginn 10.000 Einheiten des Finanzproduktes besitzt und diese zum Zeitpunkt $t = 150$

verkauft. Anschließend besitzt der*die Large Trader*in bis zum Ende des Beobachtungszeitraumes keine Einheiten. In Abbildung 3.2 kann nun anhand der rot eingezeichneten Driftfunktion die Abhängigkeit des Drifts von der Handelsstrategie erkannt werden. Neben dem sprunghaften Einbruch bei $t = 150$ lässt sich außerdem erkennen, wie die Steigung des Driftterms nach Veräußerung flacher ist als zuvor. Ab diesem Zeitpunkt gilt $\theta_t = 0$ und die Driftfunktion ist somit konstant gleich μ . Das ruckartige Verkaufen spiegelt sich insbesondere auch im Preisprozess S wider. Wie im Driftterm ist auch hier ein Einbruch zu $t = 150$ erkennbar.

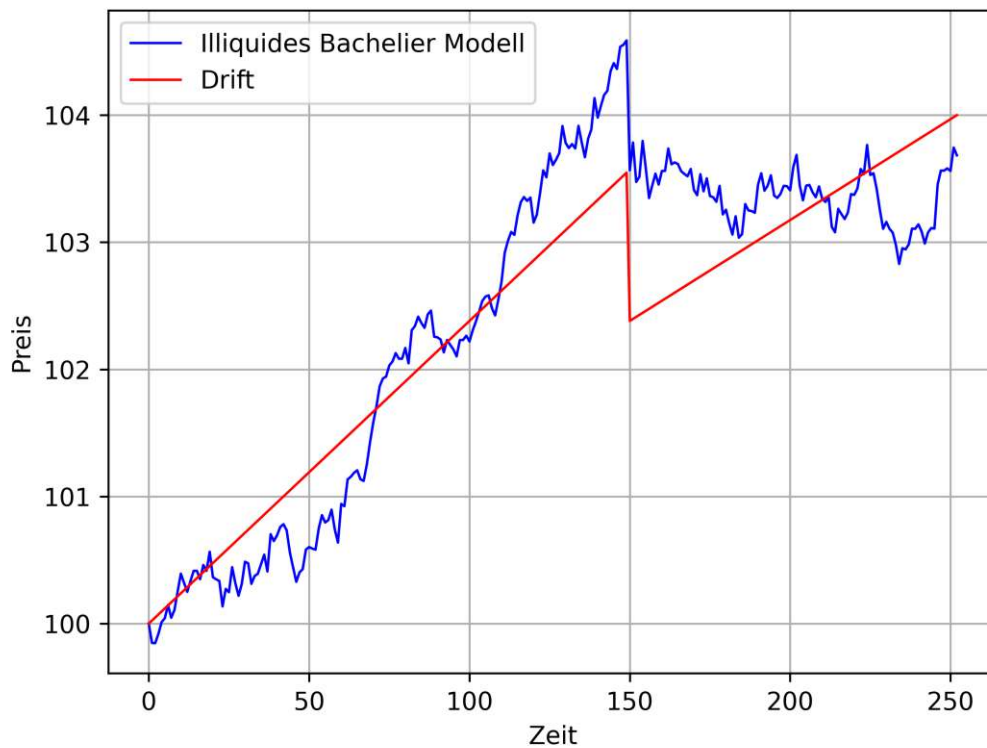


Abbildung 3.2: Simulation eines einzelnen Pfades in einem Bachelier Modell unter illiquiden Marktannahmen mit Drift $\mu = 4$, $\sigma = 2$, $\kappa = 0.0002$ und $S_0 = 100$

Wir betrachten nun einen Claim H^θ , der vom Endwert des Preisprozesses S^θ und in weiterer Folge auch von der Handelsstrategie θ abhängt. Diese Abhängigkeit sei beschreiben durch

$$H^\theta := h\left(S_T^\theta\right).$$

Dabei könnte es sich beispielsweise um eine europäische Call Option, der Form

$$H^\theta = \left(S_T^\theta - K\right)^+$$

handeln.

Ziel ist es nun eine Hedgingstrategie zu finden, durch deren Ausführung das Restrisiko minimiert wird. Zu diesem Zweck definieren wir ein von der Strategie θ abhängiges, risikoneutrales Wahrscheinlichkeitsmaß $\mathbb{Q}^\theta$. Laut [17, Kapitel 2.5] und 2.1.10 sind die stochastischen Exponentiale hierfür geeignete Kandidaten. Das Maß $\mathbb{Q}^\theta$ sei implizit durch

$$Z^\theta := \frac{d\mathbb{Q}^\theta}{d\mathbb{P}} := \mathcal{E}(V \cdot W)$$

festgelegt, wobei der Prozess V definiert sei als

$$V := -\frac{\mu(\theta)}{\sigma}. \quad (3.6)$$

Insgesamt ergibt sich somit

$$Z^\theta = \mathcal{E}\left(-\int \frac{\mu(\theta)}{\sigma} dW\right).$$

Als direkte Folgerung aus Definition 2.1.5 ergibt sich

$$Z^\theta = \exp\left(-\int \frac{\mu(\theta)}{\sigma} dW - \frac{1}{2}\left[-\int \frac{\mu(\theta)}{\sigma} dW, -\int \frac{\mu(\theta)}{\sigma} dW\right]\right). \quad (3.7)$$

Um diesen Prozess zu vereinfachen, gestaltet sich der folgende Satz aus [17, Kapitel 2.2] als hilfreich.

Satz 3.2.1 (Quadratische Kovariation eines stochastischen Integrals). *Es sei X ein Semimartingal und $\theta \in L(X)$. Dann gilt*

$$\left[\int \theta dX\right] = \int \theta^2 d[X].$$

Gemeinsam mit 2.1.11 vereinfacht sich 3.7 somit zu

$$Z^\theta = \exp\left(-\int \frac{\mu(\theta)}{\sigma} dW - \frac{1}{2} \int \frac{\mu(\theta)^2}{\sigma^2} dt\right)$$

und das zu lösende Optimierungsproblem lautet

$$\min_{\theta \in L^2(S)} \mathbb{E}_{\mathbb{Q}^\theta} \left[\left(H^\theta - \int_0^T \theta_t dS_t^\theta \right)^2 \right]. \quad (3.8)$$

Laut [18, Bemerkung 7.45] ist es für die Anwendung von 2.1.10 ausreichend zu zeigen, dass $V \cdot W$ das Novikov Kriterium erfüllt.

Satz 3.2.2 (Novikov Kriterium). *Es sei M ein reellwertiges, stetiges, lokales Martingal mit $\mathbb{E}[\exp(M_0)] < \infty$. Wenn*

$$C_N := \mathbb{E} \left[\exp \left(\frac{1}{2} [M]_T \right) \right] < \infty$$

gilt, dann ist das stochastische Exponential $\mathcal{E}(M)$ ein gleichmäßig integrierbares Martingal, dass fast sicher $\mathcal{E}(M) > 0$ erfüllt [18, Satz 7.110].

Gemäß [18, Bemerkung 5.110] handelt es sich bei $V \cdot W$ um einen reellwertigen, stetigen Prozess. Zum Beweis, dass es sich bei $V \cdot W$ um ein lokales Martingal handelt, zitieren wir folgenden Satz aus [17, Kapitel 2.2]:

Satz 3.2.3 (Martingaleigenschaft stochastischer Integrale). *Handelt es sich bei M um ein lokales Martingal und bei θ um einen adaptierten, càdlàg Prozess, so ist der zugehörige stochastische Integralprozess $\int \theta dM$ ebenfalls ein lokales Martingal.*

Laut [18, Beispiel 4.7 und 4.135] handelt es sich bei W um ein lokales Martingal und da θ eine Large Trader*innen Strategie symbolisiert, handelt es sich bei θ um einen adaptierten, càdlàg Prozess. Nachdem μ, κ und σ Konstanten sind, ist auch V ein adaptierter, càdlàg Prozess. Zusammen mit Satz 3.2.3 folgt daher, dass $V \cdot W$ ein lokales Martingal ist.

Darüber hinaus ist auch die Bedingung $\mathbb{E}[\exp((V \cdot W)_0)] < \infty$ erfüllt, wie aus Definition 2.1.8 hervorgeht. Schließlich bleibt noch $C_N < \infty$ zu zeigen. Auf Grund von Satz 3.2.1 gilt

$$\left[\int_0^T V_s dW_s \right] = \int_0^T V_s^2 d[W]_s.$$

Da es sich bei μ, κ und σ um Konstanten handelt und θ eine zulässige Handelsstrategie im Sinne von Definition 3.1.3 ist, folgt

$$\mathbb{E} \left[\exp \left(\frac{1}{2} [M]_T \right) \right] < \infty.$$

Somit ist Novikovs Kriterium erfüllt und Korollar 2.1.10 kann herangezogen werden. Demnach handelt es sich bei $\mathbb{Q}^\theta$ um ein zu $\mathbb{P}$ absolutstetiges Wahrscheinlichkeitsmaß und bei

$$\widetilde{W} := W - \int_0^\cdot V_s^T ds = W + \int_0^\cdot \frac{\mu(\theta)}{\sigma} ds$$

um eine $(\mathbb{F}, \mathbb{Q}^\theta)$ -Brownsche Bewegung. Durch Multiplikation mit σ und Anwendung des Hauptsatzes der Differential- und Integralrechnung [10, Satz 8.4.5] erhalten wir daraus

$$\sigma d\widetilde{W}_t = \mu(\theta) dt + \sigma dW_t = dS_t^\theta.$$

Da Vielfache einer Brownschen Bewegung selbst wieder Brownsche Bewegungen sind [18, Beispiel 4.7], folgt somit, dass auch S^θ eine Brownsche Bewegung unter $\mathbb{Q}^\theta$ ist. Hervorzuheben

ist an dieser Stelle jedenfalls, dass es sich sowohl bei $\widetilde{W}$ als auch bei W um Brownsche Bewegungen handelt und diese daher in Verteilung übereinstimmen. Es gilt somit

$$S^\theta = \sigma \widetilde{W} \sim \sigma W. \quad (3.9)$$

Auf Grund von 3.9 ergibt sich für den zugrundeliegenden Claim

$$H^\theta = h(S_T^\theta) = h(\sigma \widetilde{W}_T) \sim h(\sigma W_T) := H.$$

Insgesamt vereinfacht sich das vorhandene Optimierungsproblem 3.8 somit zu

$$\min_{\theta \in L^2(S)} \mathbb{E}_{\mathbb{Q}^\theta} \left[\left(H - \int_0^T \theta_t \sigma dW_t \right)^2 \right].$$

4 Stabilität

Nachdem wir in Kapitel 3 das aus dem Mean-Variance Hedging resultierende Optimierungsproblem formuliert und diskutiert haben, werden wir in diesem Kapitel dessen Verallgemeinerung betrachten. Das grundlegende Ziel des*der Large Traders*Traderin ist es, den Nutzen seines*ihres Vermögens nach Ablauf einer vorab definierten Zeitspanne zu maximieren. Da es sich dabei jedoch um eine Zufallsvariable handelt, betrachten wir für unser Optimierungsproblem den entsprechenden Erwartungswert.

Bei der Untersuchung des generellen Optimierungsproblems wird klar, dass sich nur unter bestimmten Bedingungen eine optimale Handelsstrategie findet, deren Realisierung den erwarteten Nutzen tatsächlich maximiert. Existiert eine entsprechende Strategie, bezeichnen wir das Szenario als stabiles Regime, andernfalls als instabiles Regime. In Kapitel 4.1 wird eine Differenzierung dieser beiden Fälle vorgenommen. Anschließend wird das verallgemeinerte Optimierungsproblem in Kapitel 4.2 auf das bereits diskutierte illiquide Bachelier Modell angewandt. Darüber hinaus werden auch konkrete Anforderungen an die Parameter beschrieben, deren Erfüllung Stabilität sicherstellt.

4.1 Nutzenmaximierung

Wie bereits in Kapitel 2.3 diskutiert, wird die Entwicklung des zugrundeliegenden Finanzprodukts mit Hilfe einer primalen Familie modelliert. Wir betrachten im Folgenden primale Familien, die sich als

$$P(\theta, t) := P_0 + \mu(\theta)[M]_t + \sigma(\theta)M_t$$

darstellen lassen, wobei M ein stetiges, quadratisch integrierbares, fast sicher in Null startendes Martingal mit deterministischer quadratischer Variation ist. Außerdem sei $P_0 \in \mathbb{R}$ und $\mu, \sigma \in C^2(\mathbb{R}, \mathbb{R})$ (vgl. [2, Kapitel 5]). Gemäß Definition 2.3.4 und Bemerkung 2.3.6 hat der durch diese primale Familie resultierende Preisprozess für eine beliebige Large Trader*innen Strategie im Sinne von Definition 2.3.2 die Form

$$S_t^\theta = \int_0^t P(\theta, ds) = P_0 + \int_0^t \mu(\theta) d[M]_s + \int_0^t \sigma(\theta) dM_s. \quad (4.1)$$

Unter der bereits getroffenen Annahme, dass es sich bei der Volatilität um einen konstanten, von der gewählten Strategie unbeeinflussten Parameter handelt, ist der zugehörige Vermögensprozess somit durch

$$V_t^\theta = V_0^\theta + \int_0^t \theta \mu(\theta) d[M]_s + \int_0^t \theta \sigma dM_s \quad (4.2)$$

gegeben. Dabei betrachten wir weiterhin zulässige Large Trader*innen Strategien und definieren zu Notationszwecken die Menge

$$\Theta := \{\theta \in \mathbb{L}(\mathbb{R}) \cap L^2(S)\}.$$

Zur Maximierung des erwarteten Nutzen des durch diesen Prozess beschriebenen Endvermögens folgen wir der Vorgehensweise in [2] und wählen die exponentielle Nutzenfunktion

$$u(x) := 1 - e^{-\alpha x} \quad (4.3)$$

mit Risikoavversionsparameter $\alpha > 0$. Analog zu Kapitel 3.2 können wir nun abermals ein risikoneutrales Wahrscheinlichkeitsmaß $\mathbb{Q}^\theta \sim \mathbb{P}$ definieren, unter dem wir den erwarteten Nutzen maximieren. Die Vorgehensweise lässt sich dabei übertragen, jedoch mit dem Unterschied, dass wir den zur Definition des Dichteprozesses benötigten Prozess V anders als in 3.6 als

$$V := -\alpha\theta\sigma$$

festlegen. Ähnlich wie in Kapitel 3.2 kann dann mit Hilfe des Kriteriums von Novikov argumentiert werden, dass durch den resultierenden Dichteprozess

$$\frac{d\mathbb{Q}_\alpha^\theta}{d\mathbb{P}} = \exp\left(-\alpha \int_0^T \theta_u \sigma dM_u - \frac{\alpha^2}{2} \int_0^T \theta_u^2 \sigma^2 d[M]_u\right) \quad (4.4)$$

tatsächlich ein risikoneutrales, äquivalentes Wahrscheinlichkeitsmaß $\mathbb{Q}_\alpha^\theta$ erklärt wird, das nicht nur von der Strategie des*der Large Traders*Traderin, sondern auch von der Risikoaversion abhängt.

Den Erwartungswert der Nutzenfunktion des Endvermögens können wir mit Hilfe von 4.3 und unter Anwendung der Linearität des Erwartungswertes zu

$$\mathbb{E}\left[u\left(V_T^\theta\right)\right] = 1 - \mathbb{E}\left[\exp\left(-\alpha V_T^\theta\right)\right]$$

umformen. Durch Einsetzen von 4.2 wird daraus

$$1 - \mathbb{E}\left[\exp\left(-\alpha\left(V_0^\theta + \int_0^T \theta_u \mu(\theta_u) d[M]_u + \int_0^T \theta_u \sigma dM_u\right)\right)\right].$$

Den daraus resultierenden Erwartungswert wollen wir nun einem Maßwechsel unterziehen. Zu diesem Zweck multiplizieren wir die innere Funktion mit dem Reziprokwert des zugehörigen Dichteprozesses 4.4 und erhalten somit

$$\begin{aligned} & 1 - \mathbb{E}_{\mathbb{Q}_\alpha^\theta}\left[\exp\left(-\alpha\left(V_0^\theta + \int_0^T \theta_u \mu(\theta_u) d[M]_u \right.\right.\right. \\ & \left.\left.\left. + \int_0^T \theta_u \sigma dM_u - \int_0^T \theta_u \sigma dM_u - \frac{\alpha}{2} \int_0^T \theta_u^2 \sigma^2 d[M]_u\right)\right)\right]. \end{aligned}$$

Dies lässt sich wiederum zu

$$1 - \mathbb{E}_{\mathbb{Q}_\alpha^\theta} \left[\exp \left(-\alpha \left(V_0^\theta + \int_0^T \theta_u \mu(\theta_u) - \frac{\alpha}{2} \theta_u^2 \sigma^2 d[M]_u \right) \right) \right]$$

vereinfachen. Definieren wir nun eine Hilfsfunktion

$$p(x) := x\mu(x) - \frac{\alpha}{2}x^2\sigma^2, \quad (4.5)$$

folgt insgesamt

$$\mathbb{E} \left[u \left(V_T^\theta \right) \right] = 1 - e^{-\alpha V_0^\theta} \mathbb{E}_{\mathbb{Q}_\alpha^\theta} \left[e^{-\alpha \int_0^T p(\theta_u) d[M]_u} \right]. \quad (4.6)$$

Um die Bedeutung der Funktion $p(x)$ aus 4.5 zu verdeutlichen, betrachten wir zunächst den Fall, dass diese zumindest ein Maximum besitzt. Wir nehmen also an, dass

$$\exists x^* \in \mathbb{R} : \forall x \in \mathbb{R} : p(x^*) \geq p(x) \quad (4.7)$$

erfüllt ist. Unter dieser Annahme können wir die Gleichung aus 4.6 folgendermaßen weiter umformen:

$$\begin{aligned} \mathbb{E} \left[u \left(V_T^\theta \right) \right] &= 1 - e^{-\alpha V_0^\theta} \mathbb{E}_{\mathbb{Q}_\alpha^\theta} \left[e^{-\alpha \int_0^T p(\theta_u) d[M]_u} \right] \\ &= 1 - e^{-\alpha V_0^\theta} \mathbb{E}_{\mathbb{Q}_\alpha^\theta} \left[e^{-\alpha \int_0^T p(\theta_u) + p(x^*) - p(x^*) d[M]_u} \right] \\ &\stackrel{(1)}{=} 1 - e^{-\alpha V_0^\theta} \mathbb{E}_{\mathbb{Q}_\alpha^\theta} \left[e^{-\alpha (p(x^*)[M]_T + \int_0^T p(\theta_u) - p(x^*) d[M]_u)} \right] \\ &\stackrel{(2)}{=} 1 - e^{-\alpha (V_0^\theta + p(x^*)[M]_T)} \mathbb{E}_{\mathbb{Q}_\alpha^\theta} \left[e^{-\alpha \int_0^T p(\theta_u) - p(x^*) d[M]_u} \right]. \end{aligned}$$

Dabei folgt Gleichheit (1) auf Grund der Linearität des stochastischen Integrals (vgl. [18, Kapitel 5]) und Gleichheit (2) auf Grund der Messbarkeit des konstanten Ausdrucks $p(x^*)[M]_T$. Da diese Gleichungskette für alle Large Trader*innen Strategien θ gilt, folgt aus 4.7:

$$\mathbb{P} (p(\theta) - p(x^*) \leq 0, \forall \theta \in \Theta) = 1.$$

Folglich wird der erwartete Nutzen des Endvermögens maximal, wenn die Differenz $p(\theta) - p(x^*)$ zu jedem Zeitpunkt verschwindet. Die optimale Handelsstrategie ist demnach $\theta^* = x^*$. Insbesondere folgt daraus, dass sich in diesem Szenario eine konstante Strategie optimal verhält. In [2, Kapitel 5] wird der Fall, dass die Funktion p ein Maximum annimmt und die optimale Handelsstrategie folglich konstant gleich diesem Maximum ist, als stabiles Regime bezeichnet. Durch Ausführen dieser optimalen Strategie erhalten wir in diesem Szenario schließlich

$$\sup_{\theta \in \Theta} \mathbb{E} [u(V_T)] = 1 - \exp \left(-\alpha \left(V_0^\theta + p(x^*)[M]_T \right) \right), \quad (4.8)$$

da der Exponent im Erwartungswert verschwindet und somit der gesamte Ausdruck im Erwartungswert den Wert eins annimmt.

Im Gegensatz dazu betrachten wir nun den Fall, dass die Funktion p zwar kein Maximum annimmt, aber dennoch von oben beschränkt ist. In diesem Fall können wir auf Grund der Stetigkeit der Funktion p eine Folge konstanter Strategien $(\theta_n)_{n \in \mathbb{N}}$ finden, sodass

$$\sup_{\theta \in \mathbb{R}} p(\theta) = \lim_{n \rightarrow \infty} p(\theta_n) =: p^*.$$

Für die Maximierung des erwarteten Nutzen des Endvermögens bedeutet das wiederum

$$\sup_{\theta \in \Theta} \mathbb{E}[u(V_T)] = 1 - \exp\left(-\alpha \left(V_0^\theta + p^*[M]_T\right)\right).$$

Falls p sogar unbeschränkt ist, gilt auf Grund von

$$\sup_{\theta \in \mathbb{R}} p(\theta) = \infty$$

schließlich

$$\sup_{\theta \in \Theta} \mathbb{E}[u(V_T)] = 1.$$

Da p eine stetige Funktion ist und wir wissen, dass stetige Funktionen ihre Maxima und Minima auf kompakten Mengen annehmen [10, Kapitel 6], folgt, dass in den letzten beiden Fällen $\theta_n \rightarrow \pm\infty$ gilt. Falls $\theta_n \rightarrow \infty$, ist dies so zu interpretieren, dass es für den*die Large Trader*in optimal ist, möglichst viele Einheiten des Finanzproduktes zu kaufen. Die Performance der Strategie wird also besser, je mehr Einheiten der*die Large Trader*in in seinen*ihren Besitz bringt. Umgekehrt impliziert $\theta_n \rightarrow -\infty$, dass der*die Large Trader*in optimal handelt, wenn er*sie möglichst viele Einheiten via shortselling verkauft. Beide Vorgehensweisen führen in der Praxis zu einer Destabilisierung des Marktes und sind insbesondere nicht zulässig. Es gibt somit keine optimale Handelsstrategie. Dieses Szenario wird daher in der Literatur auch als instabiles Regime bezeichnet (vgl. [2, Kapitel 5]).

Unter den genannten Annahmen, existiert also genau dann eine optimale Large Trader*innen Strategie, wenn die Funktion p aus 4.5 ein Maximum besitzt. In diesem Fall ist die resultierende Strategie sogar konstant.

4.2 Illiquides Bachelier Modell

Wir wollen nun die generelle Vorgehensweise aus Kapitel 4.1 so wie in [2, Kapitel 5] aufzeigt auf das in Kapitel 3.2 beschriebene illiquide Bachelier Modell anwenden. Für diesen Zweck gilt es zunächst das eingangs vorgestellte Martingal M , welches stetig, quadratisch integrierbar und von deterministischer quadratischer Variation sein muss sowie fast sicher in Null starten soll, zu wählen.

Um von der allgemeinen Form des Preisprozesses aus 4.1 auf das illiquide Bachelier Modell zu kommen, wählen wir für den Prozess M eine Standard Brownsche Bewegung W . Laut Satz

2.1.11 verfügt diese über deterministische quadratische Variation. Darüber hinaus resultiert Einsetzen der Brownschen Bewegung in 4.1 in

$$dS_t^\theta = \mu(\theta) dt + \sigma dW_t.$$

Wählen wir nun die Driftfunktion μ wie in Kapitel 3.2, ergibt sich der Prozess

$$dS_t^\theta = (\mu + \kappa\theta) dt + \sigma dW_t,$$

der das illiquide Bachlier Modell charakterisiert. Die übrigen Anforderungen werden durch eine Standard Brownsche Bewegung ebenfalls erfüllt (vgl. [18, Kapitel 2 und 5]).

Die Überlegungen aus Kapitel 4.3 lassen sich nun in analoger Weise auf diesen Spezialfall übertragen. Für die Funktion p aus 4.5 ergibt sich dann

$$\begin{aligned} p(x) &= x\mu(x) - \frac{\alpha}{2}x^2\sigma^2 \\ &= x(\mu + \kappa x) - \frac{\alpha}{2}x^2\sigma^2 \\ &= \mu x + \kappa x^2 - \frac{\alpha}{2}x^2\sigma^2 \\ &= \mu x + x^2\left(\kappa - \frac{\alpha}{2}\sigma^2\right) \\ &= \mu x + x^2\left(-\frac{\alpha\sigma^2 - 2\kappa}{2}\right). \end{aligned}$$

Um Wohldefiniertheit sicherzustellen, nehmen wir im Folgenden an, dass $\alpha\sigma^2 \neq 2\kappa$ gilt. Unter dieser Annahme können wir den obenstehenden Ausdruck auf ein vollständiges Quadrat ergänzen. Es folgt:

$$\begin{aligned} p(x) &= \mu x + x^2\left(-\frac{\alpha\sigma^2 - 2\kappa}{2}\right) \\ &= -\frac{\alpha\sigma^2 - 2\kappa}{2}\left(x^2 - \frac{2\mu x}{\alpha\sigma^2 - 2\kappa}\right) \\ &= -\frac{\alpha\sigma^2 - 2\kappa}{2}\left(x^2 - \frac{2\mu x}{\alpha\sigma^2 - 2\kappa}\right) - \frac{\mu^2}{2(\alpha\sigma^2 - 2\kappa)} + \frac{\mu^2}{2(\alpha\sigma^2 - 2\kappa)} \\ &= -\frac{\alpha\sigma^2 - 2\kappa}{2}\left(x^2 - \frac{2\mu x}{\alpha\sigma^2 - 2\kappa} + \frac{\mu^2}{(\alpha\sigma^2 - 2\kappa)^2}\right) + \frac{\mu^2}{2(\alpha\sigma^2 - 2\kappa)} \\ &= -\frac{\alpha\sigma^2 - 2\kappa}{2}\left(x - \frac{\mu}{\alpha\sigma^2 - 2\kappa}\right)^2 + \frac{\mu^2}{2(\alpha\sigma^2 - 2\kappa)}. \end{aligned}$$

Um festzustellen, ob es eine optimale Strategie gibt und wie diese aussieht, müssen wir im nächsten Schritt untersuchen, ob die Funktion p ein Maximum besitzt und wenn dies der Fall ist, für welches Argument das Maximum erreicht wird. Zu diesem Zweck betrachten wir zunächst die erste Ableitung der Funktion p :

$$p'(x) = -(\alpha\sigma^2 - 2\kappa)\left(x - \frac{\mu}{\alpha\sigma^2 - 2\kappa}\right).$$

Da wir bereits $\alpha\sigma^2 \neq 2\kappa$ vorausgesetzt haben, verschwindet die Ableitung genau dann, wenn

$$x = \frac{\mu}{\alpha\sigma^2 - 2\kappa}.$$

Zur Überprüfung, ob es sich bei diesem Argument um ein Maximum handelt, betrachten wir darüber hinaus auch die zweite Ableitung. Diese lautet

$$p''(x) = -(\alpha\sigma^2 - 2\kappa).$$

Die Art der Extremstelle und somit in weiterer Folge auch die Stabilität des Regimes hängt also von der Wahl der Parameter α, σ und κ ab. Sofern

$$\alpha\sigma^2 > 2\kappa \tag{4.9}$$

gilt, ist die zweite Ableitung negativ. In diesem Fall befinden wir uns somit im stabilen Regime und es existiert eine optimale Handelsstrategie, die einen konstanten Wert von

$$\theta^* = \frac{\mu}{\alpha\sigma^2 - 2\kappa} \tag{4.10}$$

hat. Gilt umgekehrt

$$\alpha\sigma^2 < 2\kappa,$$

ist die zweite Ableitung positiv. Folglich existiert kein Maximum und auch keine optimale Strategie.

In Abbildung 4.1 sind zwei mögliche Verläufe der Funktion p visualisiert. In beiden Fällen wurde der Risikoavversionsparameter $\alpha = 0.5$, ein konstanter Driftterm $\mu = 0.1$ und eine Volatilität $\sigma = 0.3$ gewählt. Für den ersten, rot dargestellten Fall wurde darüber hinaus für den Illiquiditätsparameter $\kappa = 0.02$ gewählt. In diesem Fall gilt

$$\alpha\sigma^2 = 0.045 > 0.04 = 2\kappa.$$

Somit befinden wir uns im stabilen Regime. Dies wird insbesondere anhand der rot dargestellten Funktion, die ein Maximum bei $x = 20$ aufweist, ersichtlich. Wählen wir bei gleichbleibenden sonstigen Annahmen den Illiquiditätsparameter $\kappa = 0.025$, folgt

$$\alpha\sigma^2 = 0.045 < 0.05 = 2\kappa.$$

Trotz der vermeintlich kleinen Änderung ist die Veränderung weitreichend. Die resultierende, blau dargestellte Funktion $p(x)$ besitzt kein Maximum, sondern lediglich ein Minimum bei $x = -20$. Wir befinden uns somit im instabilen Regime und es existiert keine optimale Strategie.

Gilt für den Vermögensprozess $V_0^\theta = 0$, so erhalten wir im stabilen Regime unter Anwendung von Formel 4.8 einen erwarteten Nutzen des Vermögensprozesses zur Maturität von

$$1 - \exp(-\alpha p(\theta^*)T) = 1 - \exp\left(-\frac{\alpha\mu^2 T}{2(\alpha\sigma^2 - 2\kappa)}\right).$$

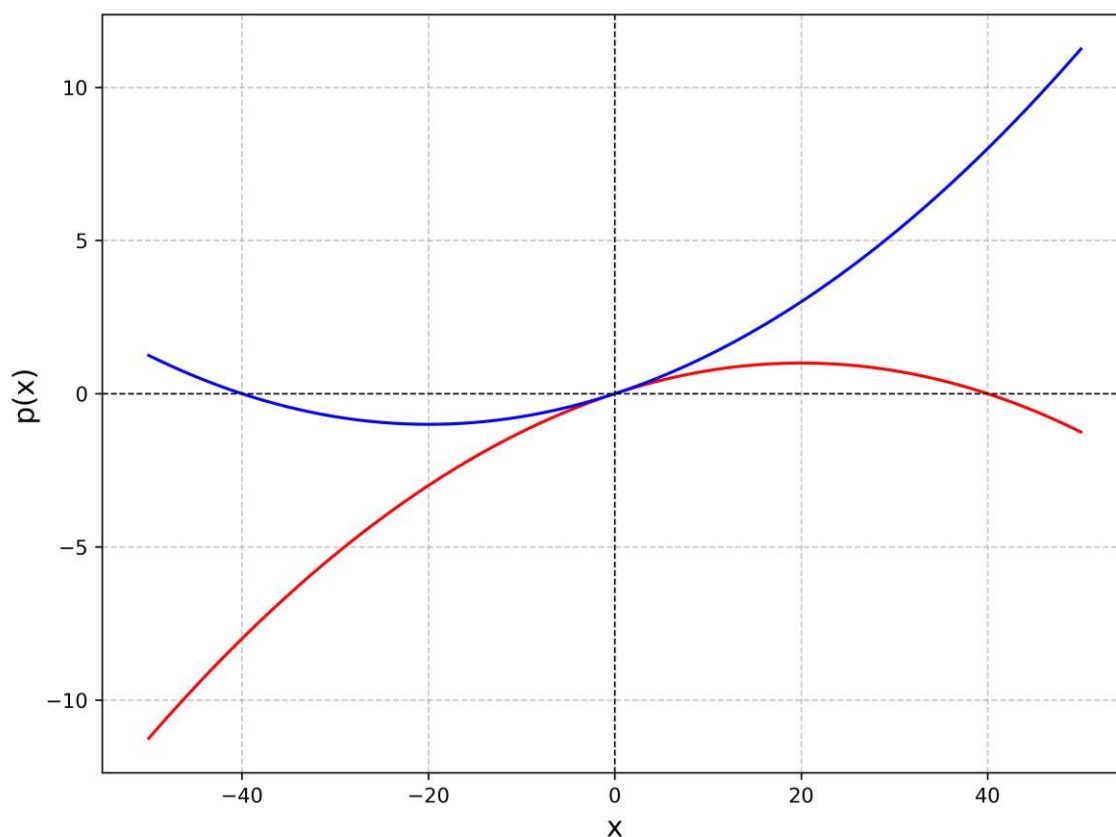


Abbildung 4.1: Verlauf des Polynoms $p(x)$ unter den Annahmen $\alpha = 0.5, \sigma = 0.3$ und $\mu = 0.1$, wobei die Entwicklung für $\kappa = 0.02$ rot und die Entwicklung für $\kappa = 0.025$ blau dargestellt ist.

Mit den beschriebenen exemplarischen Parameterwerten führt diese Formel im stabilen Regime auf

$$\sup_{\theta \in \Theta} \mathbb{E}[u(V_T)] \approx 0.39.$$

Das entsprechende Optimum wird auch in Abbildung 4.2 ersichtlich. Hier ist der erwartete Nutzen im stabile Regime abermals rot visualisiert. Darüber hinaus nimmt die auf dem gesamten Zeitintervall streng monoton wachsende, stetige, blaue Funktion, die den zweiten Fall darstellt und somit das instabile Regime symbolisiert, kein Maximum an.

Somit verbleibt noch der zuvor explizit ausgenommene Spezialfall

$$\alpha\sigma^2 = 2\kappa.$$

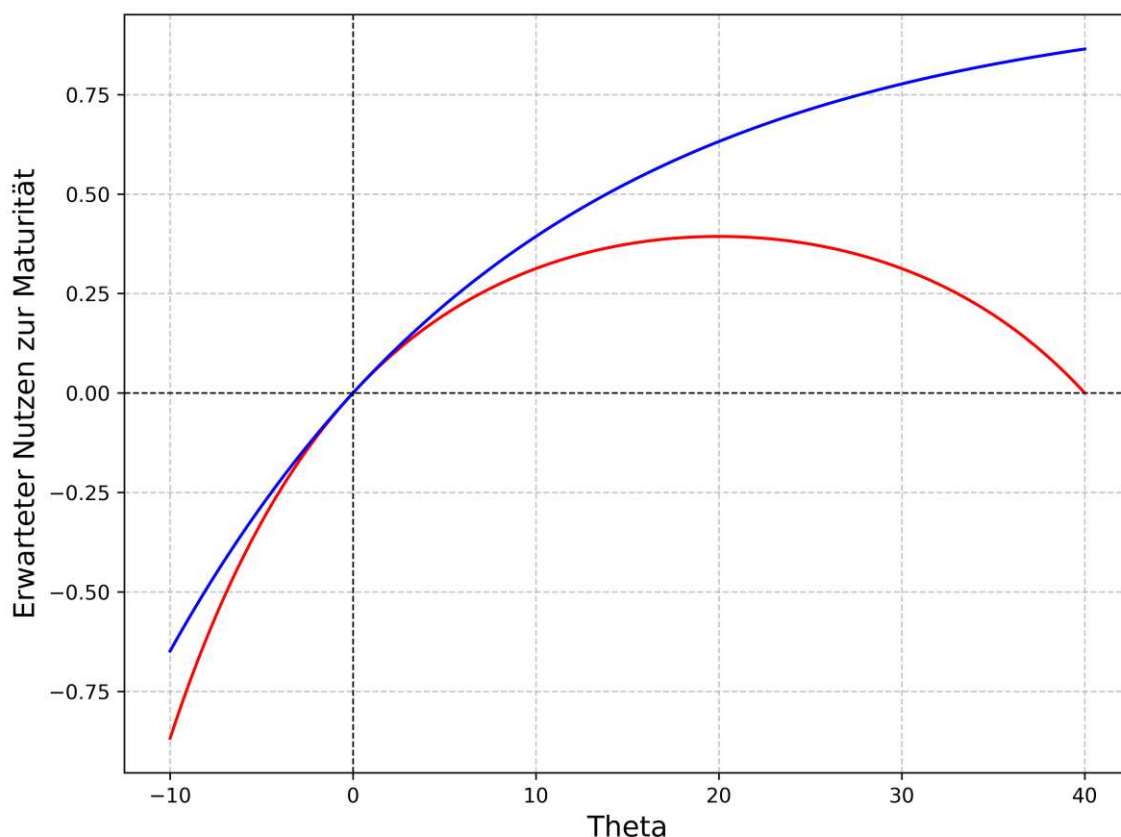


Abbildung 4.2: Erwartungswert des Nutzens des Vermögensprozesses zur Maturität $T = 1$ unter den Annahmen $\alpha = 0.5$, $\sigma = 0.3$ und $\mu = 0.1$ in Abhängigkeit einer konstanten Handelsstrategie θ , wobei der Verlauf für $\kappa = 0.02$ rot und der Verlauf für $\kappa = 0.025$ blau dargestellt ist.

In diesem Fall ergibt sich für die Funktion p aus 4.5:

$$\begin{aligned}
 p(x) &= x\mu(x) - \frac{\alpha}{2}x^2\sigma^2 \\
 &= x(\mu + \kappa x) - \frac{\alpha}{2}x^2\sigma^2 \\
 &= \mu x + \kappa x^2 - \kappa x^2 \\
 &= \mu x.
 \end{aligned}$$

Das Verhalten der Funktion p hängt somit allein vom Driftparameter μ ab. Verschwindet dieser, gibt es keine dezidierte optimale Strategie. In diesem Fall führen alle zulässigen Handelsstrategien zu denselben erwarteten Nutzen wie die Nullstrategie. Der*die Large Trader*in könnte also genauso gut nicht investieren. Ist $\mu \neq 0$, so handelt es sich bei p um eine nichtkonstante, lineare Funktion. Somit wird kein Maximum erreicht und wir befinden uns im instabilen Regime. Der Einfluss des Drifts der Preisfunktion überwiegt in diesem Fall die

anderen Parameter. Abhängig vom Vorzeichen des Parameters μ wird der*die Large Trader*in versuchen, möglichst viele Einheiten des Finanzproduktes zu kaufen beziehungsweise zu verkaufen. Dieses Verhalten lässt den Preis des zugrundeliegenden Produktes entweder gegen null konvergieren oder unlimitiert wachsen. Beide Fälle führen zu einer Destabilisierung des betrachteten Finanzmarktes.

5 Verteilung im stabilen Regime

Unter der Voraussetzung, dass wir uns im stabilen Regime befinden, lässt sich, wie in Kapitel 4 bereits hergeleitet, stets eine optimale Handelsstrategie finden. Aufbauend auf dieser Strategie lässt sich ein damit assoziierter Wertprozess bestimmen, der die Vermögensentwicklung unter Anwendung ebendieser Strategie darstellt. Das Ziel dieses Vorgehens ist die Maximierung des erwarteten Nutzens, der diesem Prozess entspringt. Da es sich bei dem bereits diskutierten Wertprozess im Allgemeinen um einen stochastischen Prozess handelt, ist der Wert, den er nach Ablauf der betrachteten Zeitspanne annimmt, eine Zufallsvariable. Natürlich ist die Verteilung sowie deren Momente dabei von höchstem Interesse für etwaige Large Trader*innen.

Im ersten Abschnitt dieses Kapitels steht die Verteilung des Wertprozesses selbst sowie die Verteilung des Wertprozesses unter Anwendung der exponentiellen Nutzenfunktion im Fokus. Aufbauend darauf wird im zweiten Abschnitt schließlich der erwartete Nutzen des Wertprozesses sowie dessen Varianz beschrieben. Außerdem wird mit Hilfe des starken Gesetzes der großen Zahl und des zentralen Grenzwertsatzes ein Schätzer für den erwarteten Nutzen hergeleitet. Zur Veranschaulichung wird das in Kapitel 4 begonnene Beispiel des illiquiden Bachelier Modell fortgesetzt.

5.1 Verteilung des Wertprozesses

Im Folgenden nehmen wir an, dass die Modellparameter μ, σ, α und κ so gewählt sind, dass 4.9 erfüllt ist, beziehungsweise äquivalent dazu die Stabilität des Regimes sichergestellt ist. Wie bereits in den vorherigen Kapiteln beschrieben, wird der Preisprozess des zugrundeliegenden Finanzprodukts durch

$$dS_t^{\theta^*} = (\mu + \kappa\theta^*) dt + \sigma dW_t$$

erklärt. Dabei bezeichnet θ^* die durch 4.10 gegebene konstante, optimale Strategie θ^* . Unter den vorausgesetzten Annahmen existiert diese. Demnach ist der zugehörige Vermögensprozess durch

$$V_T = \int_0^T \theta^* dS_t^{\theta^*} = \theta^* (\mu + \kappa\theta^*) T + \theta^* \sigma W_T$$

gegeben. Somit sind abgesehen von der Brownschen Bewegung W alle auftretenden Parameter deterministisch. Die Verteilung von V_T lässt sich daher auf Verteilungseigenschaften der Brownschen Bewegung zurückführen.

Laut [18, Kapitel 2] hat die Brownsche Bewegung normalverteilte Inkremente. Da affine Funktionen Normalität erhalten, folgt

$$V_T \sim \mathcal{N}(\theta^*(\mu + \kappa\theta^*)T, \sigma^2\theta^{*2}T). \quad (5.1)$$

Für unsere Zwecke betrachten wir stets den Nutzen des mit einer Strategie θ^* assoziierten Wertprozesses bezüglich der in 4.3 beschriebenen, exponentiellen Nutzenfunktion. Um die Verteilung des Nutzens des Wertprozesses zu bestimmen, gehen wir schrittweise vor und betrachten zunächst die Verteilung des mit einem Faktor $(-\alpha)$ multiplizierten Wertprozesses. Dabei bezeichnet α weiterhin die Risikoaversion und somit eine nichtnegative, reelle Zahl. Da es sich hierbei abermals um eine affine Funktion handelt, erhalten wir

$$-\alpha V_T \sim \mathcal{N}(-\alpha\theta^*(\mu + \kappa\theta^*)T, \alpha^2\sigma^2\theta^{*2}T).$$

Der nächste Schritt ist die Anwendung der Exponentialfunktion auf diesen Ausdruck. Dabei handelt es sich zwar um keine affine Transformation, die resultierende Zufallsvariable folgt jedoch einer durchaus bekannten Verteilung. Wie in [18, Definition 2.52] beschrieben, handelt es sich bei

$$\exp(-\alpha V_T) \sim \text{LogN}(-\alpha\theta^*(\mu + \kappa\theta^*)T, \alpha\sigma\theta^*\sqrt{T})$$

um eine lognormalverteilte Zufallsvariable. Folglich wird der Nutzen des mit der optimalen Strategie assoziierten Wertprozesses zur Maturität durch

$$u(V_T(\theta^*)) \sim 1 - \text{LogN}(-\alpha\theta^*(\mu + \kappa\theta^*)T, \alpha\sigma\theta^*\sqrt{T}) \quad (5.2)$$

beschrieben.

Zur Veranschaulichung setzen wir das bereits in Kapitel 4 eingeführte Beispiel mit den Werten $T = 1, \mu = 0.1, \sigma = 0.3, \kappa = 0.02$ und dem Risikoaversionsparameter $\alpha = 0.5$ fort. Wir befinden uns somit im stabilen Regime und gemäß Formel 4.10 ergibt sich für die optimale Strategie $\theta^* = 20$. Gemeinsam mit 5.1 wird nun ersichtlich, dass der zugehörige Wertprozess zur Maturität

$$V_T \sim \mathcal{N}(20(0.1 + 0.02 * 20), 0.3^2 * 20^2) = \mathcal{N}(10, 36)$$

erfüllt. In Abbildung 5.1 ist die Dichtefunktion dieser Verteilung rot visualisiert. Im blauen Histogramm sind 10.000 Realisationen der Zufallsvariable V_T zusammengefasst.

Bemerkenswert ist jedenfalls, dass auch unter Verwendung dieser optimalen Strategie 478 der 10.000 simulierten Szenarien einen negativen Wertprozess zum Endzeitpunkt prognostizieren. Das entspricht 4,78% der Fälle. In der überwiegenden Mehrheit der Fälle ist somit mit positivem Gewinn zu rechnen. Dabei handelt es sich jedoch lediglich um den tatsächlich erwirtschafteten Betrag und nicht um den für unsere Zwecke interessanteren Nutzen.

In Abbildung 5.2 ist die Verteilung des Nutzens des Wertprozesses dargestellt. Um dieses neue Histogramm zu erhalten, wurde die exponentielle Nutzenfunktion auf jede der

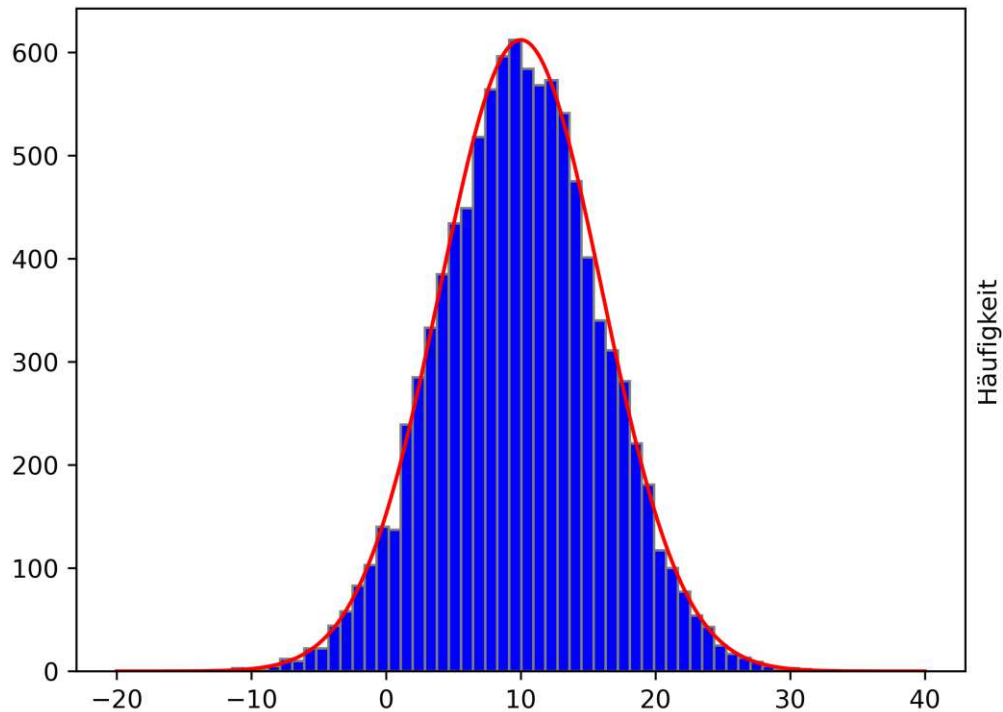


Abbildung 5.1: Dichtefunktion des Wertprozesses zur Maturität

10.000 Realisationen aus Abbildung 5.1 angewandt. Die in 5.2 beschriebene, theoretische Dichtefunktion lautet in unserem Beispiel

$$u(V_T) \sim 1 - \text{LogN}(-5, 9)$$

und ist abermals rot eingezeichnet.

5.2 Verteilung des erwarteten Nutzens

Als nächstes betrachten wir den erwarteten Nutzen des Vermögensprozesses unter Anwendung der optimalen Strategie zur Maturität. Bevor wir uns dem Erwartungswert im konkreten Fall des illiquiden Bachelier Modells widmen, betrachten wir eine stochastische Approximation des allgemeinen Falls. Ziel ist es, einen Schätzer für den Erwartungswert festzulegen und dessen Verteilung zu untersuchen. Dafür wenden wir die beispielsweise in [18, Kapitel 2] beschriebene Monte Carlo Methode für numerische Integration an. Ausgangspunkt für dieses Verfahren ist eine Folge von reellwertigen, unabhängigen und identisch verteilten Zufallsvariablen $(X_i)_{i \in \{1 \dots M\}}$ und eine messbare Funktion u , sodass es sich bei $u(X_1)$ um eine integrierbare Zufallsvariable handelt. In unserem Anwendungsfall wählen wir als Folge

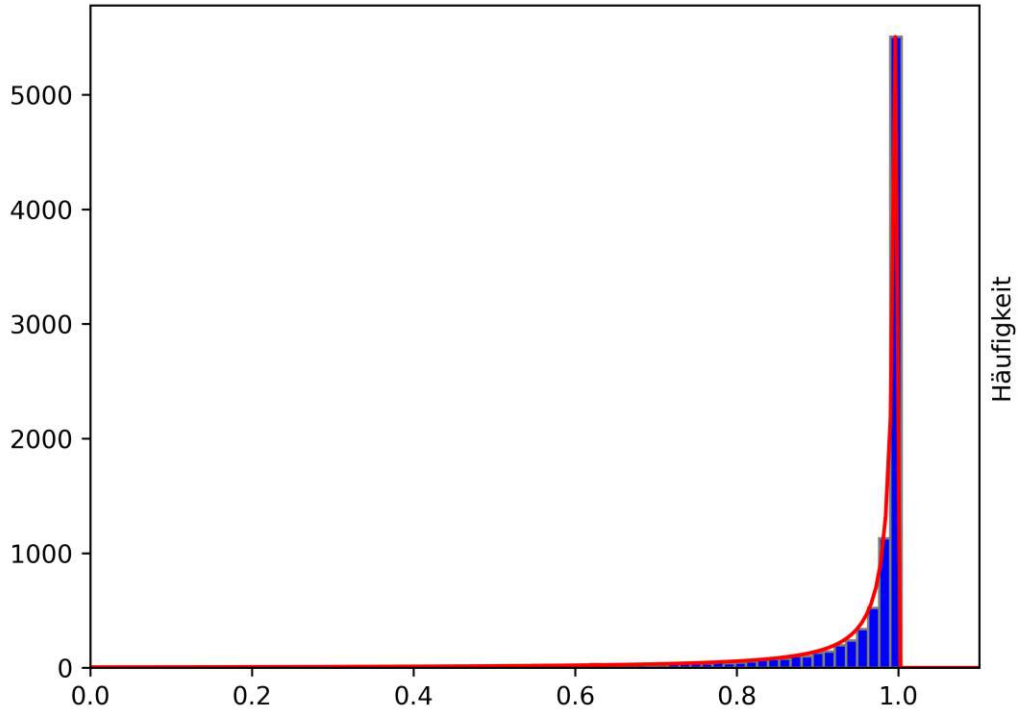


Abbildung 5.2: Dichtefunktion des Nutzens des Wertprozesses zur Maturität

von Zufallsvariablen die Folge der Realisationen des Wertprozesses:

$$(X_i)_{i \in \{1 \dots M\}} := \left(V_T^{\theta^*}(i) \right)_{i \in \{1 \dots M\}}.$$

Dabei bezeichnet M die Anzahl der betrachteten Realisationen und $V_T^{\theta^*}(i)$ die i -te Realisation der Zufallsvariable $V_T^{\theta^*}$. Da der Wertprozess neben deterministischen Konstanten nur aus einer Brownschen Bewegung besteht, ist klar, dass die Realisationen unabhängig voneinander sind. Darüber hinaus entspringen alle Realisationen der in Kapitel 5.1 beschriebenen Verteilung. Weiters wählen wir als messbare Funktion die exponentielle Nutzenfunktion.

Gemäß dem starken Gesetz der großen Zahl (vgl. [21, Kapitel 12.10]) konvergiert das arithmetische Mittel der durch Anwendung der Nutzenfunktion hervorgehenden Folgenglieder $(u(V_T^{\theta^*}(i)))_{i \in \{1 \dots M\}}$ gegen den erwarteten Nutzen der ersten Realisation des Wertprozesses. Für

$$Z_M := \frac{1}{M} \sum_{i=1}^M u(V_T^{\theta^*}(i))$$

gilt somit

$$\lim_{M \rightarrow \infty} Z_M \stackrel{f.s.}{=} \mathbb{E} \left[u(V_T^{\theta^*}) \right].$$

Wir ziehen daher die Zufallsvariable Z_M zur Approximation des erwarteten Nutzens des Wertprozesses heran. Für hinreichend große M gilt näherungsweise

$$\mathbb{E} \left[u(V_T^{\theta^*}) \right] \approx Z_M.$$

Darüber hinaus besagt der zentrale Grenzwertsatz (vgl. [21, Kapitel 18.4]), dass sogar

$$\frac{1}{\sqrt{M}} \sum_{i=1}^M \left(u(V_T^{\theta^*})(i) - \mathbb{E} \left[u(V_T^{\theta^*})(i) \right] \right) \rightarrow \mathcal{N} \left(0, \text{Var} \left[u(V_T^{\theta^*}) \right] \right)$$

für $M \rightarrow \infty$ gilt. Die dabei auftretende Summe lässt sich äquivalent dazu auch als

$$\frac{1}{\sqrt{M}} \sum_{i=1}^M \left(u(V_T^{\theta^*})(i) \right) - \frac{1}{\sqrt{M}} \sum_{i=1}^M \left(\mathbb{E} \left[u(V_T^{\theta^*})(i) \right] \right) \quad (5.3)$$

schreiben. Da alle Realisationen identisch verteilt sind und somit ihren Erwartungswert teilen, vereinfacht sich der zweite Teil von 5.3 zu

$$\frac{1}{\sqrt{M}} \sum_{i=1}^M \left(u(V_T^{\theta^*})(i) \right) - \frac{M}{\sqrt{M}} \left(\mathbb{E} \left[u(V_T^{\theta^*}) \right] \right)$$

beziehungsweise in weiterer Folge zu

$$\frac{1}{\sqrt{M}} \sum_{i=1}^M \left(u(V_T^{\theta^*})(i) \right) - \sqrt{M} \left(\mathbb{E} \left[u(V_T^{\theta^*}) \right] \right).$$

Anstatt den hier auftretenden Subtrahenden von der bestehenden Summe abzuziehen, können wir diesen auch zum Erwartungswert der oben stehenden Normalverteilung hinzuaddieren. Dadurch erhalten wir

$$\frac{1}{\sqrt{M}} \sum_{i=1}^M u(V_T^{\theta^*})(i) \rightarrow \mathcal{N} \left(\sqrt{M} \mathbb{E} \left[u(V_T^{\theta^*}) \right], \text{Var} \left[u(V_T^{\theta^*}) \right] \right).$$

Der Skalierungsfaktor $1/\sqrt{M}$ lässt sich äquivalent dazu auch als Faktor des Erwartungswerts und der Varianz der erhaltenen Normalverteilung darstellen (vgl. [12, Kapitel 9 und 14]). Auf diese Weise erhalten wir:

$$\sum_{i=1}^M u(V_T^{\theta^*})(i) \rightarrow \mathcal{N} \left(M \mathbb{E} \left[u(V_T^{\theta^*}) \right], M \text{Var} \left[u(V_T^{\theta^*}) \right] \right).$$

Um nun näherungsweise die Verteilung von Z_M zu bestimmen, müssen wir lediglich die entstandene Summe mit dem Faktor $1/M$ multiplizieren:

$$Z_M = \frac{1}{M} \sum_{i=1}^M u(V_T^{\theta^*})(i) \rightarrow \mathcal{N} \left(\mathbb{E} [u(V_T^{\theta^*})], \frac{\text{Var} [u(V_T^{\theta^*})]}{M} \right).$$

Für hinreichend große M gilt somit näherungsweise

$$Z_M \approx \mathcal{N} \left(\mathbb{E} [u(V_T^{\theta^*})], \frac{\text{Var} [u(V_T^{\theta^*})]}{M} \right). \quad (5.4)$$

Um ein tieferes Verständnis über die Verteilung des Schätzers Z_M zu erhalten, betrachten wir nun den Erwartungswert des Nutzens des Wertprozesses im illiquiden Bachelier Modell. Nachdem wir bereits wissen, dass es sich beim Nutzen des Wertprozesses abgesehen von einer Konstante um eine lognormalverteilte Zufallsvariable handelt, gilt für den Erwartungswert

$$\mathbb{E} [u(V_T^{\theta^*})] = 1 - \exp(a + \frac{1}{2}b^2), \quad (5.5)$$

wobei hierbei a den ersten Parameter der Lognormalverteilung aus 5.2 bezeichne und b den zweiten (vgl. [20]). Wie bereits in Kapitel 5.1 diskutiert, ist in unserem Fall

$$\begin{aligned} a &:= -\alpha\theta^* (\mu + \kappa\theta^*) T, \\ b &:= \alpha\sigma\theta^* \sqrt{T} \end{aligned}$$

vorgegeben. Setzen wir diese Variablen nun in 5.5 ein, erhalten wir

$$\begin{aligned} \mathbb{E} [u(V_T^{\theta^*})] &= 1 - \exp(-\alpha\theta^* (\mu + \kappa\theta^*) T + \frac{1}{2}\alpha^2\sigma^2\theta^{*2}T) \\ &= 1 - \exp(-\alpha\theta^* \mu T - \alpha\kappa\theta^{*2}T + \frac{1}{2}\alpha^2\sigma^2\theta^{*2}T) \\ &= 1 - \exp(-\alpha T\theta^* (\mu + (\kappa - \frac{1}{2}\alpha\sigma^2)\theta^*)). \end{aligned} \quad (5.6)$$

für den gesuchten Erwartungswert.

Neben dem Erwartungswert des Nutzens des Wertprozesses nimmt auch dessen Varianz in der Verteilung des Schätzers Z_M eine prominente Rolle ein. Der Steinersche Verschiebungssatz [12, Kapitel 15] besagt, dass für eine Zufallsvariable X

$$\text{Var}[X] = \mathbb{E}[X^2] - \mathbb{E}[X]^2$$

gilt, sofern alle in dieser Gleichung auftretenden Momente existieren. In unserem Fall wählen wir daher

$$X := 1 - u(V_T^{\theta^*}).$$

Um ein besseres Verständnis über die Verteilungseigenschaften von X^2 zu erhalten, definieren wir noch eine weitere Zufallsvariable Y . Diese dient jedoch nur als Hilfsvariable und sei durch

$$Y := \ln(X) \quad (5.7)$$

gegeben. An dieser Stelle sei hervorgehoben, dass es sich auf Grund der Verteilungseigenschaften von X bei Y um eine mit Mittelwert a und Varianz b normalverteilte Zufallsvariable handelt. Dank dieser neuen Hilfsvariable lässt sich nun erkennen, dass

$$\mathbb{E}[X^2] = \mathbb{E}[\exp(Y)^2] = \mathbb{E}[\exp(2Y)] \quad (5.8)$$

gilt. Die erste Gleichheit folgt dabei durch Anwendung von 5.7 und die zweite auf Grund der Eigenschaften der Exponentialfunktion. Als nächstes legen wir eine weitere Hilfsvariable durch

$$H := 2Y$$

fest. Da affine Funktionen Normalität erhalten, ist H immer noch normalverteilt. Insbesondere ist damit $\exp(H)$ lognormalverteilt mit den Parametern $2a$ und $4b$. Für den in 5.8 beschriebenen Erwartungswert ergibt sich somit

$$\mathbb{E}[X^2] = \mathbb{E}[\exp(H)] = \exp(2a + 2b^2),$$

wobei die letzte Gleichheit aus Formel 5.5 folgt.

Durch Einsetzen in den Steinerschen Verschiebungssatz erhalten wir für die Varianz von X :

$$\begin{aligned} \text{Var}[X] &= \exp(2a + 2b^2) - \left(\exp\left(a + \frac{1}{2}b^2\right) \right)^2 \\ &= \exp(2a + 2b^2) - \exp(2a + b^2) \\ &= \exp(2a) (\exp(2b^2) - \exp(b^2)). \end{aligned}$$

Setzen wir nun für a und b die in Kapitel 5.1 bestimmten Verteilungsparameter ein, erhalten wir schließlich

$$\text{Var}[u(V_T^{\theta^*})] = \exp(-2\alpha\theta^*(\mu + \kappa\theta^*)T) (\exp(2\alpha^2\sigma^2\theta^{*2}T) - \exp(\alpha^2\sigma^2\theta^{*2}T)). \quad (5.9)$$

Außerdem können wir nun auch die Standardabweichung des Nutzens des Wertprozesses angeben. Diese lautet

$$\text{SD}[X] = \sqrt{\text{Var}[X]} = \exp(a) \sqrt{\exp(2b^2) - \exp(b^2)}$$

beziehungsweise

$$\text{SD}[u(V_T^{\theta^*})] = \exp(-\alpha\theta^*(\mu + \kappa\theta^*)T) \sqrt{(\exp(2\alpha^2\sigma^2\theta^{*2}T) - \exp(\alpha^2\sigma^2\theta^{*2}T))}$$

im konkreten Anwendungsfall.

Um einen Einblick zu erhalten, inwieweit der Erwartungswert und die Standardabweichung zusammenhängen, betrachten wir schlussendlich noch das Verhältnis der Standardabweichung zum Erwartungswert. Um dieses Verhältnis einfach zu halten, werden wir jedoch nicht den Erwartungswert selbst heranziehen. Stattdessen betrachten wir den Erwartungswert der Zufallsvariable $\exp(-\alpha V_T^{\theta*})$. Anders formuliert handelt es sich dabei um den Erwartungswert des Nutzens des Wertprozesses, bevor der stochastische Anteil von Eins subtrahiert wird. Dafür greifen wir abermals auf den generellen Fall des lognormalverteilten Nutzens ohne weitere Konkretisierung der Verteilungsparameter zurück. Mit den in diesem Abschnitt bestimmten Momenten erhalten wir

$$\begin{aligned} \frac{SD[X]}{1 - \mathbb{E}[X]} &= \frac{\exp(a) \sqrt{\exp(2b^2) - \exp(b^2)}}{\exp(a + \frac{1}{2}b^2)} \\ &= \frac{\exp(a) \sqrt{\exp(2b^2) - \exp(b^2)}}{\exp(a) \exp(\frac{1}{2}b^2)} \\ &= \sqrt{\frac{\exp(2b^2) - \exp(b^2)}{\exp(b^2)}} \\ &= \sqrt{\frac{\exp(b^2)(\exp(b^2) - 1)}{\exp(b^2)}} \\ &= \sqrt{\exp(b^2) - 1}. \end{aligned}$$

Im Umkehrschluss kann die Standardabweichung durch die Formel

$$SD[X] = (1 - \mathbb{E}[X]) \sqrt{\exp(b^2) - 1} \quad (5.10)$$

somit auch als Transformation des Erwartungswertes angegeben werden.

Im Fall des illiquiden Bachelier Modells und der exponentiellen Nutzenfunktion berechnet sich der Parameter b wie bereits in Kapitel 5.1 hergeleitet zu $\alpha\sigma\theta^*\sqrt{T}$. Für die Standardabweichung des Nutzens des Wertprozesses unter Anwendung der optimalen Strategie zur Maturität erhalten wir somit

$$SD[u(V_T^{\theta*})] = \left(1 - \mathbb{E}[u(V_T^{\theta*})]\right) \sqrt{\exp\left((\alpha\sigma\theta^*)^2 T\right) - 1}.$$

Nachdem wir nun eine Vorgehensweise zur Bestimmung der Momente sowie der Verteilung des Schätzers für den erwarteten Nutzen hergeleitet haben, können wir das Beispiel aus Kapitel 5.1 fortsetzen. Dafür setzen wir die gegebenen Werte für $\alpha, \mu, \sigma, \kappa$ und T in Formel 5.6 ein und erhalten

$$\mathbb{E}[u(V_T^{\theta*})] = 1 - \exp\left(-\frac{1}{2}\theta^* \left(0.1 + \left(0.02 - \frac{1}{2}\frac{1}{2}0.3^2\right)\theta^*\right)\right).$$

Setzen wir zusätzlich dazu auch noch die in Kapitel 4 hergeleitete optimale Handelsstrategie ein, folgt:

$$\begin{aligned}\mathbb{E}[u(V_T^{\theta^*})] &= 1 - \exp\left(-\frac{1}{2}20\left(0.1 + \left(0.02 - \frac{1}{2}\frac{1}{2}0.3^2\right)20\right)\right) \\ &= 1 - \exp\left(-\frac{1}{2}\right) \approx 0.39.\end{aligned}$$

Ebenso können wir in Formel 5.9 einsetzen und erhalten in unserem konkreten Beispiel:

$$\begin{aligned}\text{SD}[u(V_T^{\theta^*})] &= \exp\left(-\frac{1}{2}\theta^*(0.1 + 0.02\theta^*)\right) \sqrt{\exp\left(\frac{1}{2}0.3^2\theta^{*2}\right) - \exp\left(\frac{1}{4}0.3^2\theta^{*2}\right)} \\ &= \exp\left(-\frac{1}{2}20(0.1 + 0.02 * 20)\right) \sqrt{\exp\left(\frac{1}{2}0.3^2 20^2\right) - \exp\left(\frac{1}{4}0.3^2 20^2\right)} \\ &= \exp(-5) \sqrt{\exp(18) - \exp(9)} \\ &\approx 54.59.\end{aligned}$$

Dieses Ergebnis ist insbesondere konsistent mit der alternativen Methode zur Bestimmung der Standardabweichung. Um dies zu überprüfen, müssen wir lediglich die Parameter unseres Beispiels in Formel 5.10 einsetzen. Dadurch erhalten wir abermals

$$\text{SD}[u(V_T^{\theta^*})] = \left(1 - \mathbb{E}[u(V_T^{\theta^*})]\right) \sqrt{\exp\left(\frac{1}{2}0.3^2 20^2\right) - 1} \approx 54.59.$$

Auffällig ist dabei unter anderem, dass die Standardabweichung relativ hoch und insbesondere deutlich höher als beispielsweise der Erwartungswert ist. Dies ist auf die charakteristischen schweren Enden der Lognormalverteilung zurückzuführen. Dieses Phänomen lässt sich auch an den Abbildungen 5.1 und 5.2 ablesen. Wie bereits diskutiert, weisen knapp 5% der Realisationen des Wertprozess einen negativen Wert auf. Vereinzelt werden dabei sogar Werte, die kleiner als -10, sind angenommen. Auf Grund der im Exponenten der Nutzenfunktion sowie vor der Exponentialfunktion auftretenden negativen Vorzeichen, nimmt der Nutzen dieser Realisationen sehr hohe, negative Werte an. Gleichzeitig liegt ein Großteil der Masse der Nutzenverteilung nahe der Eins. Insbesondere jene Realisationen, deren Wertprozess im zweistellig positiven Bereich liegt, nehmen nach Anwendung der Nutzenfunktion einen Wert, der minimal kleiner als eins ist, an. Bei der Bestimmung des Schätzers Z_M werden arithmetische Mittel aus jeweils einer großen Zahl an Realisationen gebildet. Sofern die Zahl der Summanden hinreichend groß ist, stoßen diese beiden Beobachtungen aufeinander. Dieses Verhalten erklärt einerseits die verhältnismäßig hohe Standardabweichung und andererseits weshalb der Erwartungswert überraschend niedrig ausfällt, gemessen daran, wie viele Realisationen einen Nutzen im Intervall $[0.9, 1]$ haben.

Schließlich können wir die soeben berechneten Momente in den Schätzer aus 5.4 einsetzen und erhalten in unserem Beispiel

$$Z_M \approx \mathcal{N}\left(0.39, \frac{54.59}{M}\right),$$

wobei M die Anzahl der Realisationen, die zur Berechnung des arithmetischen Mittels Z_M herangezogen werden, symbolisiert. In Abbildung 5.3 ist die Verteilung des Schätzers Z_M grün visualisiert. Ebenfalls grün ist der Erwartungswert der Verteilung des Schätzers eingezeichnet. Die beiden roten Linien kennzeichnen die Ränder des Intervalls, welches sich durch hinzuaddieren beziehungsweise subtrahieren einer Standardabweichung zum Erwartungswert ergibt. Das blaue Histogramm gibt Aufschluss über die tatsächliche Verteilung der simulierten Schätzer. Zu beachten ist hierbei jedenfalls, dass die Form der Verteilungskurve maßgeblich von der Anzahl der Summanden M abhängt. Für die vorliegende Graphik in Abbildung 5.3 wurden 100.000 Realisationen des Wertprozesses simuliert. Diese Realisationen wurden dann zu 100 Realisationen des Schätzers, bestehend aus jeweils 1.000 Summanden, zusammengefasst.

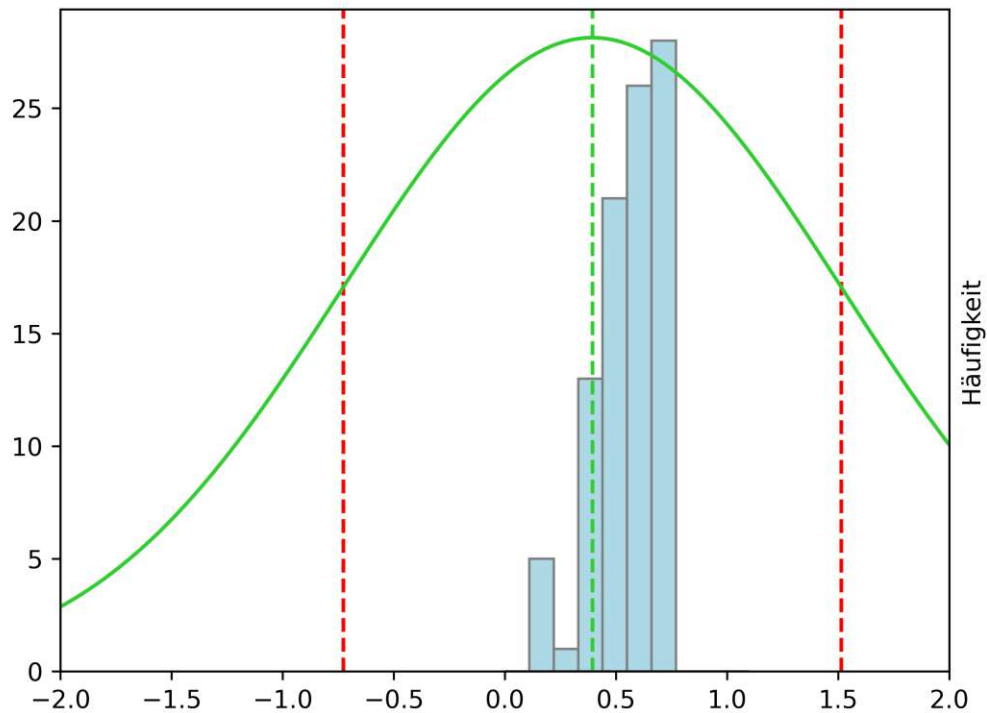


Abbildung 5.3: Verteilung des Schätzers Z_M für $M = 1.000$ inklusive Erwartungswert (grün) und davon abweichend eine Standardabweichung (rot)

Die Verteilung des Schätzers nähert sich mit wachsendem M immer weiter an den erwarteten Nutzen des Wertprozesses an. Da die Standardabweichung der Verteilung des Schätzers jedoch nur mit dem Faktor $1/\sqrt{M}$ sinkt, erfolgt diese Annäherung verhältnismäßig langsam.

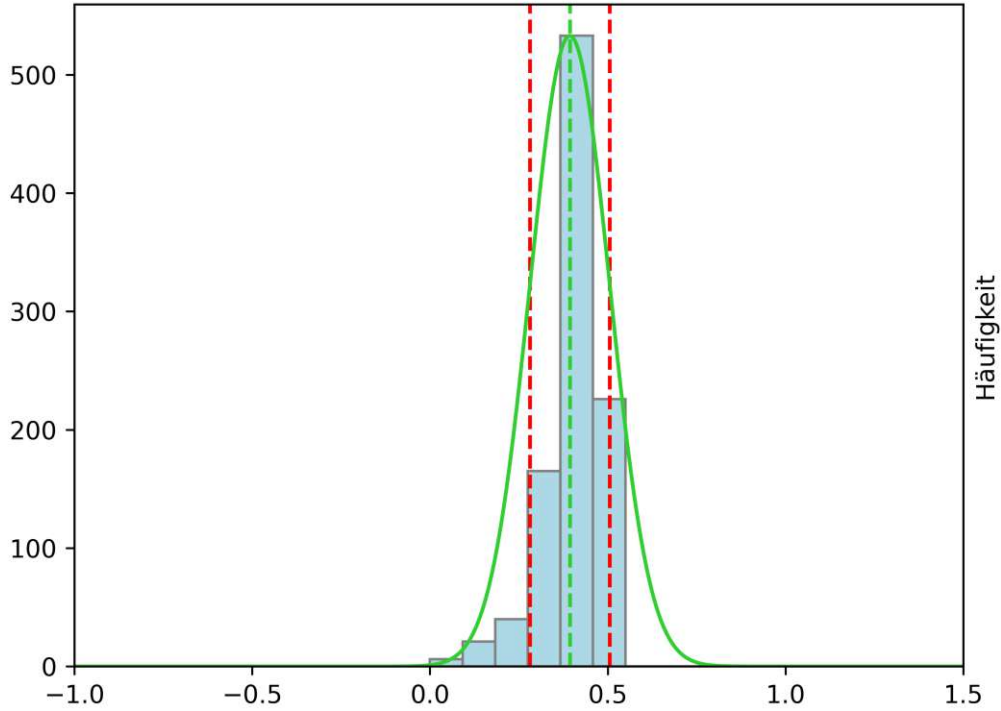


Abbildung 5.4: Verteilung des Schätzers Z_M für $M = 100.000$ inklusive Erwartungswert (grün) und davon abweichend eine Standardabweichung (rot)

Zum Vergleich findet sich in Abbildung 6.1 eine weitere Simulation. Diesmal bestehend aus 100 Millionen Realisationen des Wertprozess, wobei jeweils 1.000 derartige Realisationen zu einem arithmetischen Mittel zusammengefasst wurden. In diesem Fall ist bereits deutlich besser zu sehen, wie die Verteilungskurve enger wird und die Realisationen des Schätzers die angestrebte Normalverteilung besser annähern.

Dass es sich bei der in Kapitel 4 theoretisch hergeleiteten Handelsstrategie θ^* tatsächlich um jene Strategie handelt, die den erwarteten Nutzen maximiert, lässt sich nun auch analytisch verifizieren. Zu diesem Zweck bestimmen wir die erste Ableitung der Funktion aus 5.6 und setzen diese gleich null. Durch Differenzieren erhalten wir

$$\begin{aligned} \frac{d\mathbb{E}[u(V_T^\theta)]}{d\theta} &= \frac{d}{d\theta} \exp \left(-\alpha T \theta \left(\mu + \left(\kappa - \frac{1}{2} \alpha \sigma^2 \right) \theta \right) \right) \\ &= \exp \left(-\alpha T \theta \left(\mu + \left(\kappa - \frac{1}{2} \alpha \sigma^2 \right) \theta \right) \right) \left(-\alpha T \left(\mu + 2 \left(\kappa - \frac{1}{2} \alpha \sigma^2 \right) \theta \right) \right) \stackrel{!}{=} 0. \end{aligned}$$

Da wir bereits wissen, dass die Exponentialfunktion unabhängig von ihrem Argument niemals verschwindet (vgl. [10, Kapitel 6]), können wir äquivalent zu obiger Forderung auch lediglich die innere Ableitung null setzen. Die optimale Strategie muss somit

$$-\alpha T \left(\mu + 2 \left(\kappa - \frac{1}{2} \alpha \sigma^2 \right) \theta \right) \stackrel{!}{=} 0$$

erfüllen. Da wir bereits sowohl den Risikoaversionsparameter α als auch die Maturität T jeweils positiv vorausgesetzt haben, ist dies wiederum äquivalent zu

$$\mu + 2 \left(\kappa - \frac{1}{2} \alpha \sigma^2 \right) \theta \stackrel{!}{=} 0.$$

Durch Äquivalenzumformungen erhalten wir daraus

$$2 \left(\frac{1}{2} \alpha \sigma^2 - \kappa \right) \theta \stackrel{!}{=} \mu.$$

Schließlich folgt die Bedingung

$$\theta \stackrel{!}{=} \frac{\mu}{\alpha \sigma^2 - 2\kappa},$$

an die optimale Handelsstrategie θ^* , welche mit der Formel aus 4.10 übereinstimmt.

6 Numerisches Experiment

Im Folgenden betrachten wir das illiquide Bachelier Modell. Wie bereits in Kapitel 4 beschrieben, ist die Existenz einer stabilen, optimalen Handelsstrategie abhängig von der Risikoaversion α , der Volatilität σ und dem Illiquiditätsparameter κ . Je nachdem, ob Bedingung 4.10 erfüllt ist oder nicht, differenzieren wir zwischen dem stabilen und dem instabilen Regime. Sofern eine stabile, optimale Handelsstrategie existiert, lässt sich im Umkehrschluss der Illiquiditätsparameter κ auch aus Strategie, Drift, Volatilität und Risikoaversion berechnen.

Während beispielsweise der Driftparameter oder die Volatilität bereits in liquiden Modellen genutzt werden und je nach Kontext einigermaßen verlässlich bestimmt beziehungsweise geschätzt werden können, kann die Wahl des Illiquiditätsparameters durchaus eine Herausforderung darstellen. In diesem Kapitel wird die Fragestellung beleuchtet, welche Auswirkungen zu beobachten sind, wenn zur Bewältigung dieser Aufgabe ein neuronales Netz herangezogen wird. Nach einer generellen Einführung in die Funktionsweise neuronaler Netze in Kapitel 6.1 wird in Kapitel 6.2 auf die konkrete Implementierung und insbesondere auch auf die Wahl der Hyperparameter des neuronalen Netzes, das im Rahmen dieser Arbeit programmiert wurde, eingegangen. Abschließend werden in Kapitel 6.3 die resultierenden Ergebnisse präsentiert. Insbesondere wird auf die Erkenntnis eingegangen, dass getrieben durch die entstehende Diffusion mit positiver Wahrscheinlichkeit ein Wechsel aus dem eingangs stabilen Regime ins instabile Regime eintreten kann.

6.1 Neuronale Netze

Die erste Modellierung von neuronalen Netzen geht auf Warren McCulloch und Walter Pitts zurück und wurde schon Anfang der 1940er Jahre vorgestellt. Inspiriert von der Art und Weise, wie unser menschliches Gehirn Informationen verarbeitet, liegt die Grundidee eines neuronalen Netzes darin, die biologischen Neuronen sowie deren Interaktion nachzuahmen. Der Aufbau einer Modellzelle orientiert sich dabei stark an dem einer biologischen Nervenzelle. So wie unsere Nervenzelle aus einem Zellkern und einer Vielzahl an Synapsen besteht, werden auch die Modellzellen durch einen Kern, in dem das Ergebnis der Zelle berechnet wird und eine festgelegte Zahl an Verbindungen zu anderen Modellzellen charakterisiert. In Anlehnung an ihre Inspirationsquelle werden die Modellzellen häufig auch als Neuronen bezeichnet (vgl. [5, Kapitel 3]).

Die Neuronen des Modells sind in einer oder mehreren chronologisch anzuwendenden Schichten angeordnet. Die erste Schicht wird in der Regel als Eingabeschicht bezeichnet und die letzte als Ausgabeschicht. Alle weiteren, dazwischenliegenden Schichten werden durchnummeriert und als versteckte Schichten bezeichnet. Wird die Ausgabe eines Neurons an alle Zellen

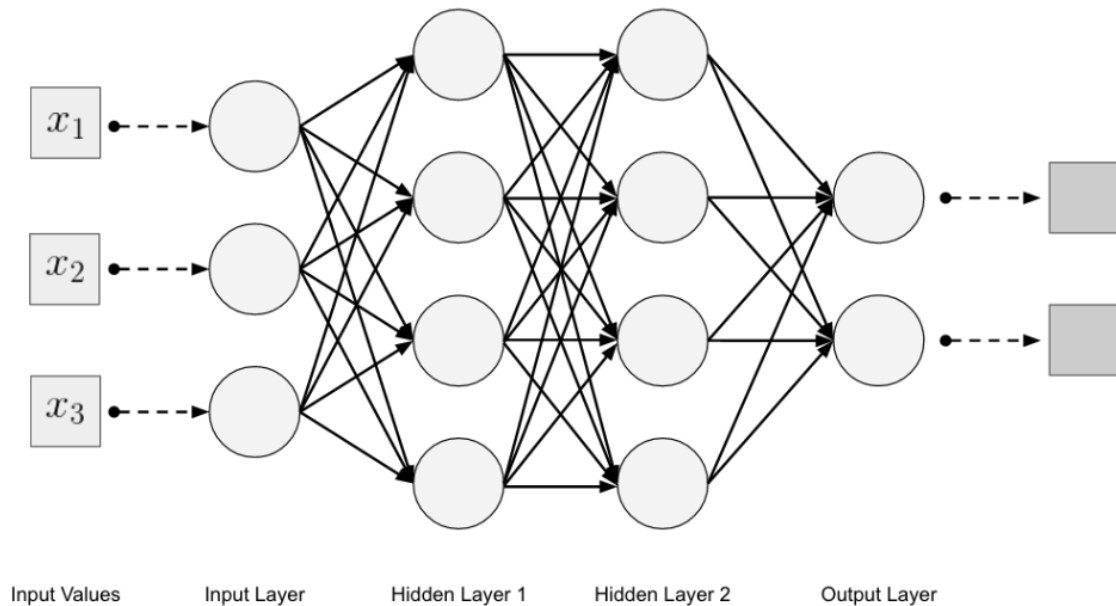


Abbildung 6.1: Struktur eines neuronalen Netzes [7, Abbildung 2.7]

der darauffolgenden Schicht übergeben, sprechen wir von einer vollständig verbundenen Zelle. Derartige Zellen sind beispielsweise in Abbildung 6.1 veranschaulicht. Grundsätzlich ist der Kern jeder Modellzelle ähnlich aufgebaut. Die Zelle übernimmt eine festgelegte Anzahl an Eingabeparametern $x_1, x_2, \dots, x_n$ und berechnet daraus in einem ersten Schritt eine gewichtete Linearkombination. Abhängig davon, ob es sich um ein Neuron der Eingabeschicht handelt oder nicht, werden diese Eingabeparameter entweder vom User als Inputs an das Modell übergeben oder entsprechen den Outputs der Zellen der vorherigen Schicht. In einem zweiten Schritt wird dann eine sogenannte Aktivierungsfunktion auf die Linearkombination angewandt (siehe Abbildung 6.2). Dieses Vorgehen bietet je nach Wahl der Aktivierungsfunktion unter anderem die Möglichkeit, die Eingabewerte nichtlinear zu transformieren. Außerdem werden die aus der Linearkombination erhaltenen Werte dadurch häufig auf den gewünschten Wertebereich skaliert. Die Ergebnisse dieses Prozesses werden anschließend als Eingabe an die nächste Schicht weitergereicht beziehungsweise im Fall der Ausgabeschicht ausgegeben. Sofern die Ausgabe eines Neurons von Null verschieden ist, bezeichnet man dieses Neuron als aktiviert (vgl. [7, Kapitel 2]).

Mathematisch betrachtet handelt es sich bei einem neuronalen Netz also vereinfacht gesagt um eine Verkettung von linearen Funktionen und Aktivierungsfunktionen. In [3, Abschnitt 4.1] wird ein neuronales Netz folgendermaßen definiert:

Definition 6.1.1. Es seien $L, N_0, N_1, \dots, N_L \in \mathbb{N}$ mit $L \geq 2$. Weiters sei $\sigma : \mathbb{R} \rightarrow \mathbb{R}$ und

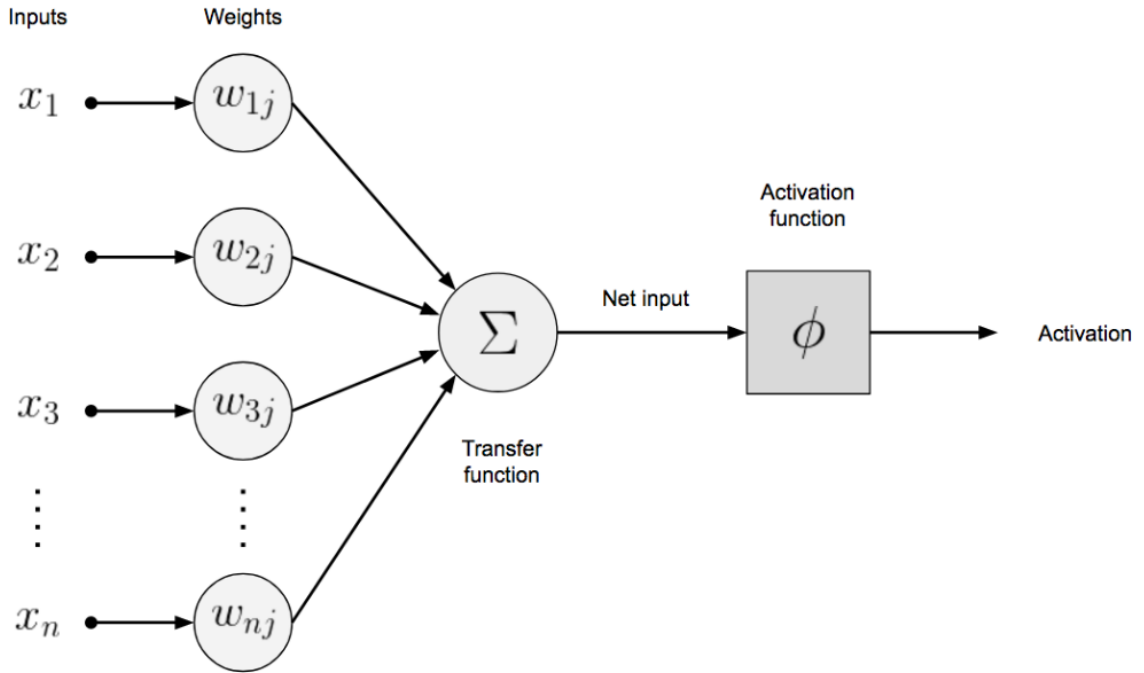


Abbildung 6.2: Aufbau eines Neurons [7, Abbildung 2.5]

für $l = 1, \dots, L$ sei $W_l : \mathbb{R}^{N_{l-1}} \rightarrow \mathbb{R}^{N_l}$ eine affine Funktion. Dann bezeichnen wir eine durch

$$\begin{aligned} F(x) &= W_L \circ F_{L-1} \circ \dots \circ F_1 \\ \text{mit } F_l &= \sigma \circ W_l \\ \text{für } l &= 1 \dots L-1 \end{aligned}$$

gegebene Funktion $F : \mathbb{R}^{N_0} \rightarrow \mathbb{R}^{N_L}$ als neuronales Netz.

Die mit σ bezeichnete Aktivierungsfunktion wird dabei komponentenweise angewandt. Die Anzahl der Schichten, aus denen das neuronale Netz aufgebaut ist, wird mit L bezeichnet. Folglich handelt es sich bei N_0 um die Eingangsschicht und bei N_L um die Ausgabeschicht. Je nach Wahl von L befinden sich dazwischen noch die $L-1$ versteckte Schichten $L_1, \dots, L_{N-1}$. Die affinen Funktionen W_l sind durch ihre Gewichte sowie einen Bias festgelegt. Der Bias beschreibt einen konstanten Term, der im Zuge der Linearkombination hinzuaddiert wird. Die Gewichte werden üblicherweise in Form einer Matrix $A^l : \mathbb{R}^{N_l \times N_{l-1}}$ zusammengefasst, sodass

$$W_l(x) = A^l(x) + b^l$$

gilt, wobei $b^l \in \mathbb{R}^{N_l}$ den Bias beschreibt. Die Einträge der Gewichtsmatrix A^l sind dabei so zu interpretieren, dass A_{ij}^l jenes Gewicht darstellt, mit dem die Ausgabe aus Zelle i der Schicht $l-1$ multipliziert wird, um die Linearkombination in Zelle j der Schicht l zu

berechnen. Darüber hinaus beschreibt der Eintrag b_i^l den Bias, der zur Berechnung der Linearkombination in Zelle i aus Schicht l hinzuaddiert wird (vgl. [3, Abschnitt 4.1]).

Die Gewichte des neuronalen Netzes werden in einem iterativen Prozess optimiert, dessen Ziel es ist, die Diskrepanz zwischen der finalen Ausgabe des Modells und der Zielgröße der Trainingsmenge zu minimieren. Zu diesem Zweck muss vorab eine Fehlerfunktion bestimmt werden, um die nach jedem Durchlauf entstandene Abweichung zwischen Ausgabe und Zielgröße zu quantifizieren. Typische Beispiele für eine geeignete Fehlerfunktion sind, abhängig vom Kontext des zugrundeliegenden Problems beispielsweise der mittlere, absolute Fehler (MAE), der mittlere, quadratische Fehler (MSE) oder die relative Entropie. Letztere eignet sich jedoch besser für Klassifikationsprobleme. Zur Minimierung des dadurch bestimmten Fehlerterms werden anschließend die Gewichte mit Hilfe von mathematischen Methoden angepasst. Häufig kommen dabei Verfahren, die auf einer Berechnung des Gradienten basieren, zum Einsatz (vgl. [5, Kapitel 3.2]).

Den vollständigen Durchlauf aller Trainingsdaten mit anschließender Fehlerevaluierung und Anpassung der Gewichte bezeichnen wir als eine Epoche. Üblicherweise werden mehrere Epochen durchlaufen, bis das Modell als final betrachtet wird. In Zuge einer Epoche werden jedoch nicht alle Daten auf einmal in das Modell gespeist. Dieser Vorgang wird aufgeteilt in mehrere sogenannte Batches. Dabei handelt es sich um eine kleinere Teilmenge der Trainingsmenge, die zeitgleich an das Modell übergeben wird. Größere Batches führen im Allgemeinen zu glatteren Gradienten, die dafür jedoch aufwendiger in ihrer Berechnung sind. Die Laufzeit des Modells wird unter anderem maßgeblich durch die Wahl der Größe der einzelnen Batches beeinflusst und ist von U -förmiger Gestalt. Mit steigender Größe reduziert sich die Laufzeit zunächst. Ab einer gewissen Größe steigt diese jedoch wieder (vgl. [7, Kapitel 6]).

Neben dem Optimierungsvorgang selbst stellt sich insbesondere auch die Frage der Initialisierung der Gewichte. Nachdem es sich um einen iterativen Prozess handelt, muss vorab der Ausgangspunkt spezifiziert werden. Dieser kann unter Umständen einen durchaus nicht zu unterschätzenden Einfluss auf das Verhalten des Netzes haben und im Extremfall sogar zu numerischer Instabilität führen. Eine geeignete Initialisierungsstrategie sollte jedenfalls eine gewisse Asymmetrie zwischen den einzelnen Zellen in das Netz einbringen. Diese Anforderung begründet sich dadurch, dass sich andernfalls im Fall gleicher Werte für zwei oder mehrere Eingabeparameter und gleicher Aktivierungsfunktion der Neuronen alle betroffenen Gewichte über den gesamten Prozess hinweg identisch entwickeln und somit auch in ihrer Endform übereinstimmen. Motiviert durch diese Beobachtung wird beispielsweise in [8] eine zufällige Initialisierung durch Realisationen einer Gleichverteilung oder einer Gaußschen Verteilung empfohlen. Die tatsächliche Wahl der Verteilung scheint dabei einen deutlich kleineren Einfluss auf das Verhalten des neuronalen Netzes zu haben als die Skalierung ebendieser. Einerseits besteht ein Interesse daran, die Gewichte groß zu wählen, da sich mit wachsenden initialen Gewichten der beschriebene asymmetrische Effekt erhöht und dadurch Redundanzen vermieden werden, andererseits können zu große Gewichte zu explodierenden Werten im Verlauf des iterativen Prozesses führen. Insbesondere kann in diesem Fall das Phänomen auftreten, dass selbst kleine Änderungen der Eingabeparameter zu einer überproportionalen Veränderung der Ausgabe führen, sodass tatsächliche Muster schwer erkennbar werden,

da diese durch auf natürliche Weise auftretende Schwankungen der Eingaben überschattet werden. In [8] wird als heuristischer Zugang beispielsweise vorgeschlagen, zufällige Werte der Gleichverteilung

$$U\left(-\frac{1}{\sqrt{m}}, \frac{1}{\sqrt{m}}\right) \quad (6.1)$$

zu generieren und diese als initiale Gewichte zu nutzen. Dabei bezeichnet m die Anzahl der Eingabeparameter des jeweiligen vollständig verbundenen Neurons (vgl. [8, Kapitel 8.4]).

6.2 Implementierung

Das dieser Arbeit zugrundeliegende neuronale Netz wurde in Python mit Hilfe des Tensorflow Packages implementiert. Konkret handelt es sich um Python Version 3.11.10 und Version 2.17.1 des Tensorflow Packages. Eine detaillierte Auflistung des verwendeten Codes befindet sich in Kapitel 7.

Das Modell selbst konzentriert sich auf die Ermittlung des Illiquiditätsparameters κ . Dieser wird auf Basis des Driftparameters μ , der Volatilität σ , der Risikoaversion α sowie der optimalen Handelsstrategie θ^* modelliert. Zu diesem Zweck wird daher in einem ersten Schritt eine Vielzahl an unterschiedlichen Szenarien simuliert. Jedes einzelne Szenario charakterisiert sich durch einen individuellen Drift, eine Volatilität und die Risikoaversion des*der Investors*Investorin. Um ausreichend verschiedene Fälle abzudecken, wurden 100.000 Datenpunkte, die jeweils aus den drei genannten Parametern bestehen, generiert. Um möglichst unterschiedliche, vielfältige Szenarien abzudecken, wurden μ , σ und α jeweils als Realisationen von drei voneinander unabhängigen Gleichverteilungen erstellt. Für den konstanten Driftanteil μ wurde als Untergrenze des zugrundeliegenden Intervalls der Fall 0.0 gewählt. Dieser Spezialfall repräsentiert ein Finanzprodukt, dessen deterministischer Trend, abgesehen von etwaigen durch den*die Large Trader*in verursachte Änderungen, verschwindet. Als Obergrenze wurde ein konstanter Driftanteil von $\mu = 0.2$ gewählt. Gleichzeitig wurde jedem Szenario eine zufällig generierte Volatilität zwischen 0.2 und 0.4 zugewiesen sowie ein Risikoaversionsparameter zwischen 0.4 und 0.6. Als Maturität wurde für alle Szenarien $T = 1$ gewählt.

Die Existenz einer optimalen Handelsstrategie und somit die Stabilität des Regimes hängt nach Feststellung der bisher diskutierten Parameter maßgeblich vom der Wahl des Illiquiditätsparameters ab. Für die iterative Ausführung des Modells müssen wir vorab noch die im späteren Verlauf von der Fehlerfunktion genutzte Zielgröße festlegen. Der Diffusionseffekt des neuronalen Netzes auf die Stabilität der optimalen Handelsstrategie wird deutlich ersichtlich, wenn der Illiquiditätsparameter eingangs nahe der Schwelle zum instabilen Regime gewählt wird. Gleichzeitig ist jedenfalls sicherzustellen, dass alle Datenpunkte eine stabile Lösung zulassen, da die zum Ausführen des neuronalen Netzes benötigte optimale Handelsstrategie andernfalls nicht existiert. Um diesen Anforderungen gerecht zu werden, bestimmen wir die

an dieser Stelle benötigten Zielwerte mit Hilfe der Formel

$$\kappa_i = 0.99 \cdot \frac{\alpha_i \cdot \sigma_i^2}{2}$$

für $i = 1 \dots 100.000$.

Nachdem mit dieser Wahl unter anderem sichergestellt ist, dass für jeden Datenpunkt Bedingung 4.9 erfüllt ist, können wir als nächstes zur Ermittlung der optimalen Handelsstrategie Formel 4.10 nutzen. Anschließend erweitern wir jeden der bestehenden Datenpunkte um die jeweils zugehörige, konstante Strategie θ_i^* . Zusammengefasst erhalten wir somit die Datenpunkte

$$(\mu_i, \sigma_i, \alpha_i, \theta_i^*)$$

für $i = 1 \dots 100.000$ und die zugehörigen Zielgrößen κ_i .

Damit sind Form und Struktur der Eingangs- sowie Ausgabeparameter geklärt und wir können uns der Architektur des neuronalen Netzes widmen. Die im folgenden vorgestellten Resultate wurden mit einem Netz bestehend aus zwei versteckten Schichten, die jeweils aus vier Neuronen bestehen, erreicht. Bei beiden Schichten handelt es sich um vollständig verbundene Schichten. Als Aktivierungsfunktion dient in der ersten versteckten Schicht die Identität und in der zweiten Schicht die ReLu Funktion. Wie in Kapitel 6.1 beschrieben, ziehen wir zur Generierung geeigneter Initialisierungswerte für die Gewichte des Netzes die in 6.1 erklärte Verteilung heran. Auf Grund der beschriebenen Netzwerkarchitektur bedeutet das, dass wir alle Gewichte eingangs mit einer Realisation der Verteilung

$$U\left(-\frac{1}{2}, \frac{1}{2}\right)$$

initialisieren.

Nach dem vollendeten ersten Vorwärtsschritt des neuronalen Netzes wird die mit den initialisierten Gewichten berechnete Ausgabe mit den Zielgrößen κ_i abgeglichen. Zur Bestimmung des Fehlerterms wird der mittlere, quadratische Fehler (MSE) herangezogen. Dieser ist definiert via

$$\text{MSE}(\kappa_i, \hat{\kappa}_i) = \frac{1}{100000} \sum_{i=1}^{100000} (\kappa_i - \hat{\kappa}_i)^2,$$

wobei $\hat{\kappa}_i$ den durch das neuronale Netz ermittelten Illiquiditätsparameter des i -ten Datenpunktes bezeichnet.

In dem anschließenden Optimierungsprozess der Gewichte wird das in [11] eingeführte ADAM (Adaptive Moment Estimator) Verfahren benutzt. Dieser gradientenbasierte Algorithmus zeichnet sich unter anderem durch seine hohe Effizienz und seine geringen Speicherkapazitätsanforderungen aus [11]. Durch die Aktualisierung der Gewichtsmatrizen ist die erste

Epoche abgeschlossen. Insgesamt werden in der hier zugrundeliegenden Implementierung 500 Epochen durchlaufen. Dieser iterative Prozess wird auf die Trainingsmenge angewandt. Dabei handelt es sich um die ersten 80% des oben konstruierten Datensatzes. Eine Durchmischung der einzelnen Datenpunkte ist auf Grund der zufälligen Generierung nicht notwendig. Die 80.000 Datenpunkte werden in Form von mehreren Batches, die aus jeweils 32 Datenpunkten bestehen, sukzessive verarbeitet.

6.3 Ergebnisse

Nach Ablauf der letzten Epoche werden mit Hilfe der finalen Gewichte die entsprechenden Illiquiditätsparameter der 20.000 Datenpunkte der Testmenge berechnet. Die Abbildungen 6.3 und 6.4 visualisieren, wie ähnlich die durch das neuronale Netz bestimmten Parameter den eingangs als Zielgröße festgelegten Parametern sind. Insbesondere wird in Abbildung 6.3 ersichtlich, von welcher Größenordnung die durch das Modell bestimmten Illiquiditätsparameter $\hat{\kappa}_i$ sind. Abhängig vom herangezogenem Drift, der Volatilität und der Risikoaversion reichen diese von etwa 0.007 bis zu 0.048.

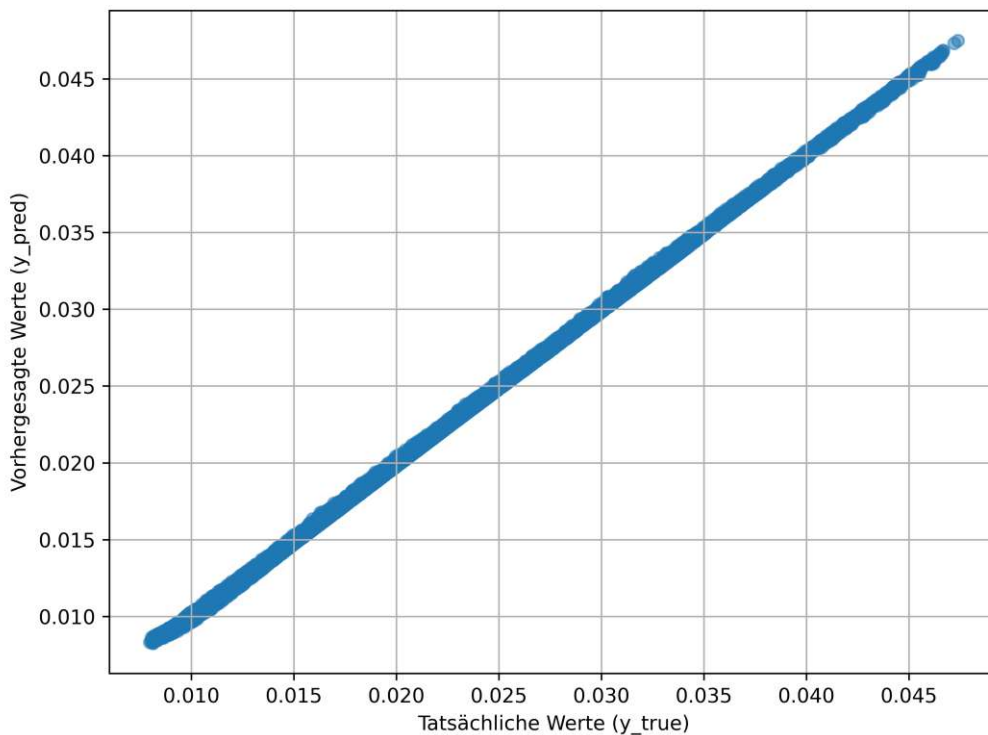


Abbildung 6.3: Gegenüberstellung der Zielgröße κ_i und der vom Modell bestimmten Illiquiditätsparameter $\hat{\kappa}_i$

Durch Vergleichen der durch das neuronale Netz bestimmten Illiquiditätsparameter $\hat{\kappa}_i$ mit der Volatilität σ und der Risikoaversion α wird nun jedoch sichtbar, dass Bedingung 4.9

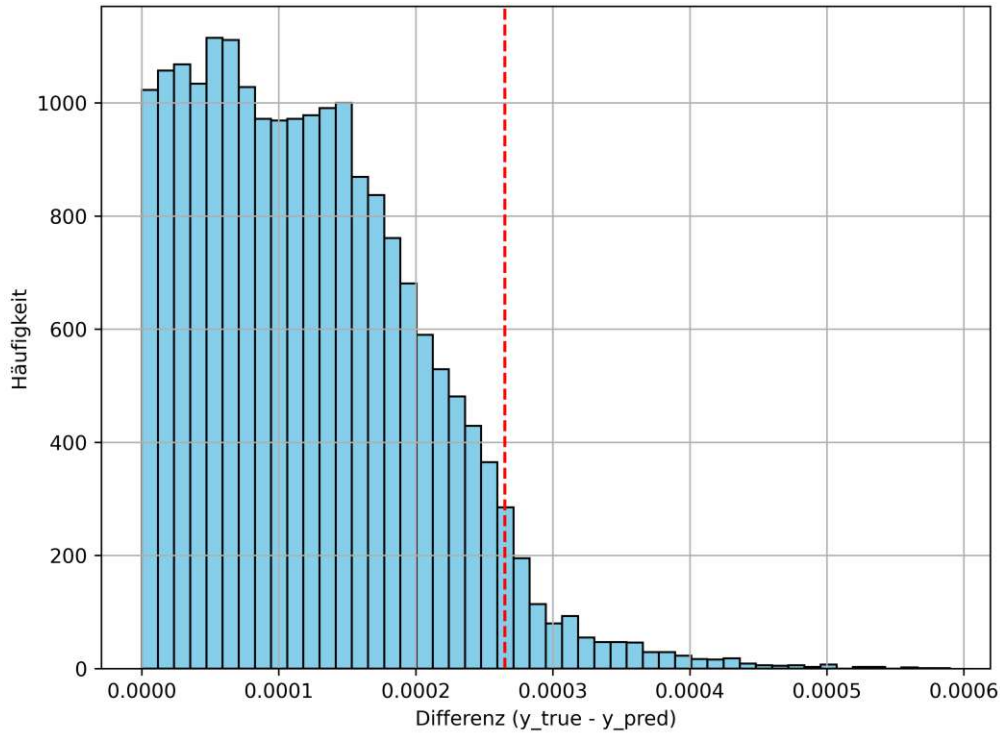


Abbildung 6.4: Absolute Differenz zwischen der Zielgröße κ_i und der vom Modell bestimmten Illiquiditätsparameter $\hat{\kappa}_i$. Die rote Linie symbolisiert dabei das 95%-Quantil.

nicht mehr in allen Szenarien erfüllt ist. Während in 19.243 der betrachteten Fälle nach wie vor ein Maximum zum entsprechenden Polynom 4.5 existiert und folglich auch eine stabile, optimale Handelsstrategie, ist das für die verbleibenden 757 der 20.000 Datenpunkte nicht der Fall. Das entspricht einem Anteil von rund 3,79%. Trotz Erfüllung aller diskutierten Voraussetzungen und Sicherstellung, dass das Modell ausschließlich von Daten, die die Einhaltung von 4.9 gewährleisten, gelernt hat, kann somit ein Regimewechsel vom eingangs stabilen Regime in das instabile Regime beobachtet werden. Insbesondere wird dadurch ersichtlich, wie die beiden Regime, bedingt durch den beschriebenen Diffusionsprozess, verschmelzen. Selbst unter Sicherstellung aller notwendigen Bedingungen existiert somit eine positive Wahrscheinlichkeit, dass unter Anwendung des neuronalen Netzes keine optimale Handelsstrategie gefunden werden kann.

Im Folgenden betrachten wir eines der Szenarien, in denen der diskutierte Übergang vom stabilen ins instabile Regime deutlich zu beobachten ist. In Abbildung 6.5 ist das zugehörige Polynom p aus 4.5 dargestellt. Dieses wird charakterisiert durch die in Kapitel 6.2 beschriebenen Eingangsparameter. Nachdem die Zielgröße κ gezielt so gewählt wurde, dass das Polynom p ein Maximum besitzt, existiert auch eine stabile, optimale Handelsstrategie θ^* , die den erwarteten Nutzen des Wertprozesses zur Maturität maximiert. In Abbildung 6.5 wird insbesondere visualisiert, dass das dafür benötigte Maximum existiert.

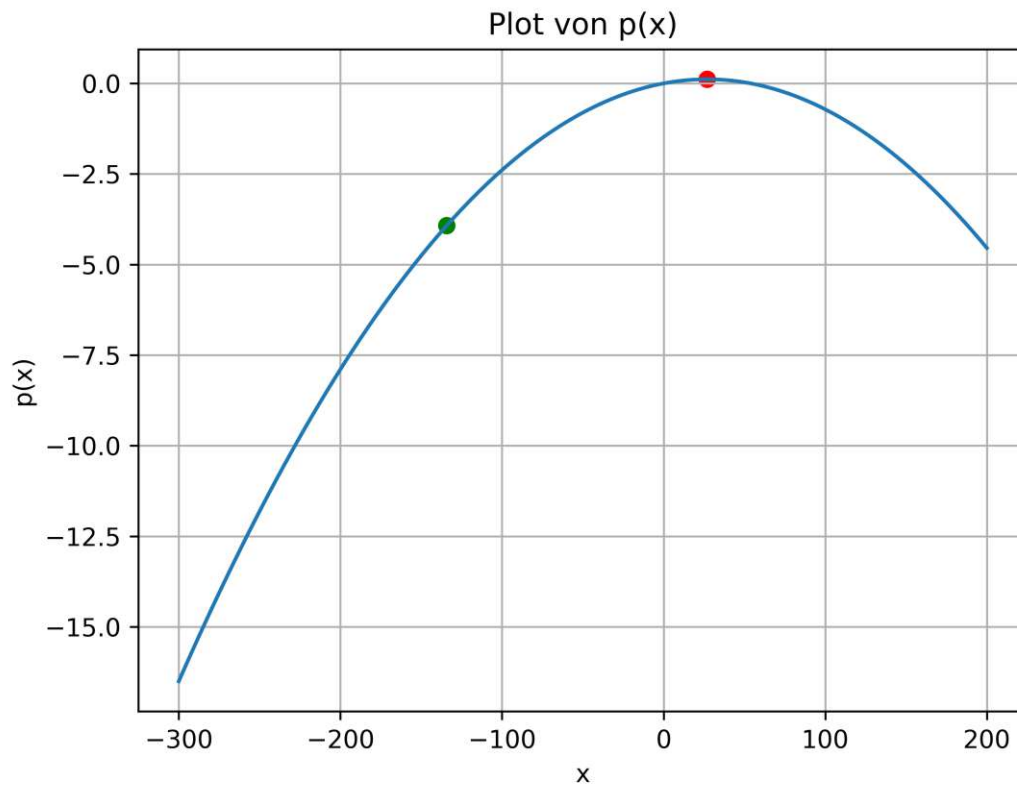


Abbildung 6.5: Beispiel eines Polynoms p im stabilen Regime vor Ausführung des neuronalen Netzes

Dem gegenüber steht nun das in Abbildung 6.6 veranschaulichte Polynom, das aus der Modellausgabe $\hat{\kappa}$ resultiert. Der Driftterm μ , die Volatilität σ und der Risikoavversionsparameter α blieben dabei unverändert. Aus Abbildung 6.6 wird schnell ersichtlich, dass im Zuge dieser Änderung ein Übergang ins instabile Regime stattfand. Das dort dargestellte Polynom verfügt über kein Maximum und folglich existiert keine stabile Strategie, die den erwarteten Nutzen des Wertprozesses maximiert.

Ebenso können wir exemplarisch einen jener Fälle, in denen die Stabilität der optimalen Strategie erhalten bleibt, heranziehen und die Ergebnisse visualisieren. Befinden wir uns nach Anwendung des neuronalen Netzes nach wie vor im stabilen Regime, existiert eine optimale Handelsstrategie $\hat{\theta}^*$. Diese lässt sich mit Hilfe von Formel 4.10 auch explizit bestimmen und ermöglicht somit eine genauere Betrachtung des daraus resultierenden Wertprozesses sowie dessen Nutzens. In Abbildung 6.7 ist die Verteilung des mit $\hat{\theta}^*$ assoziierten Wertprozesses dargestellt. Diese erinnert von ihrem Aufbau stark an die in Abbildung 5.1 visualisierte

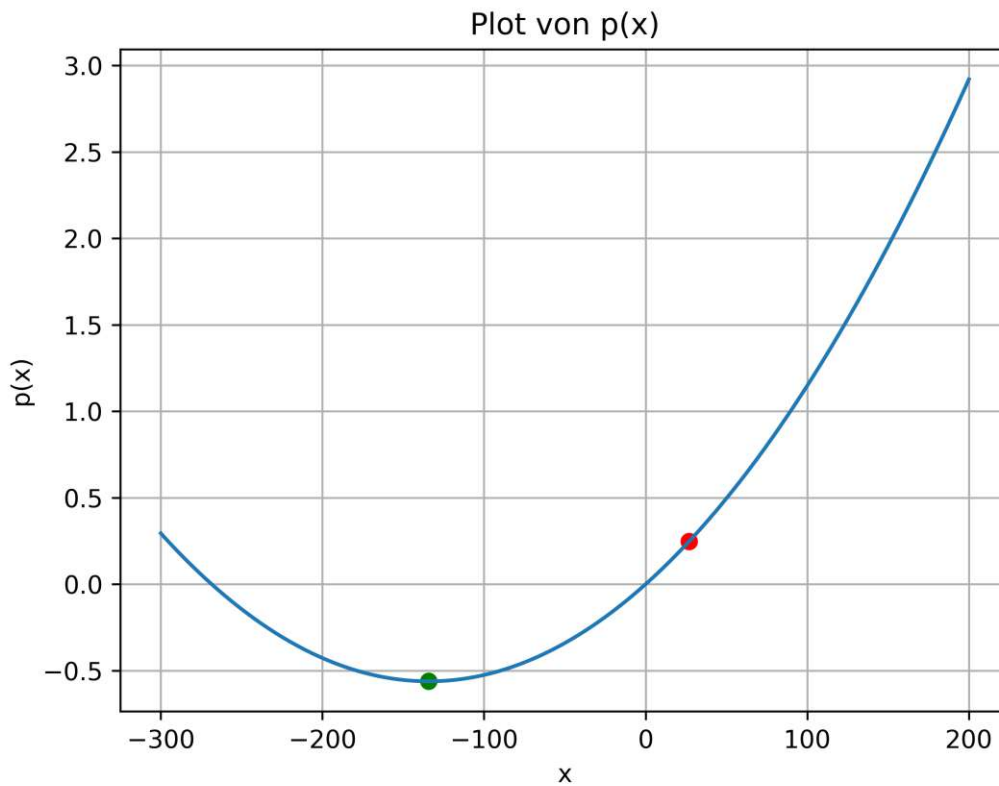


Abbildung 6.6: Beispiel eines Polynoms p im instabilen Regime nach Ausführung des neuronalen Netzes

theoretische Verteilung und verdeutlicht die durch das angewandte neuronale Netz erhaltene Normalverteilung des Wertprozesses.

Abschließend können wir den aus demselben Szenario resultierenden Nutzen des in Abbildung 6.7 dargestellten Wertprozesses betrachten. Abbildung 6.8 visualisiert diesen Nutzen. Abermals kann dabei beobachtet werden, dass die im Kapitel 5 beschriebene Verteilung erhalten bleibt und die dargestellten Realisationen die theoretische Dichtefunktion annähernd wiedergeben.

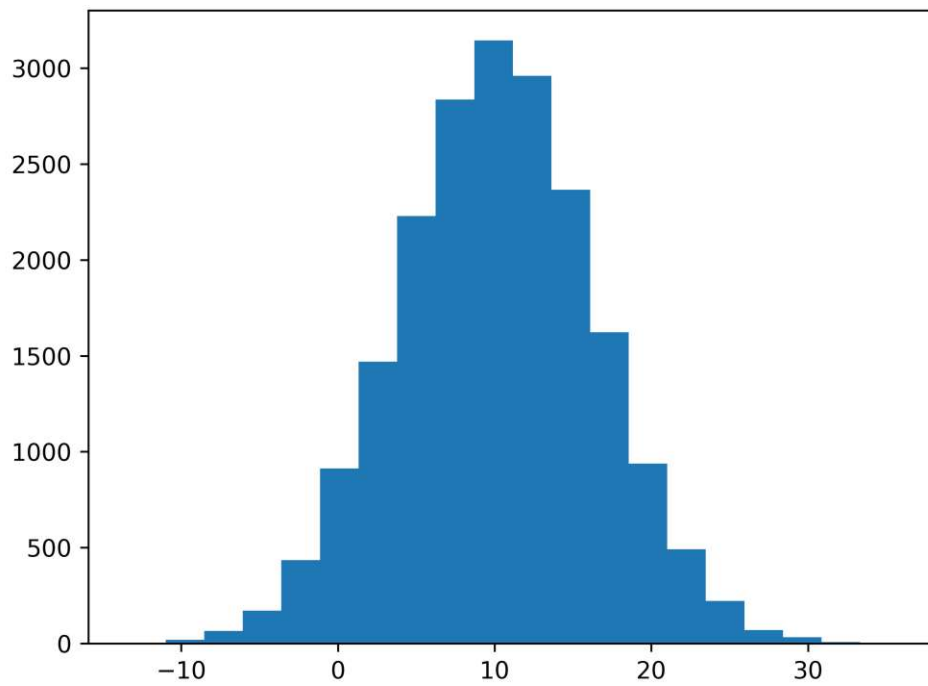


Abbildung 6.7: Realisationen des Wertprozesses eines stabilen Szenarios

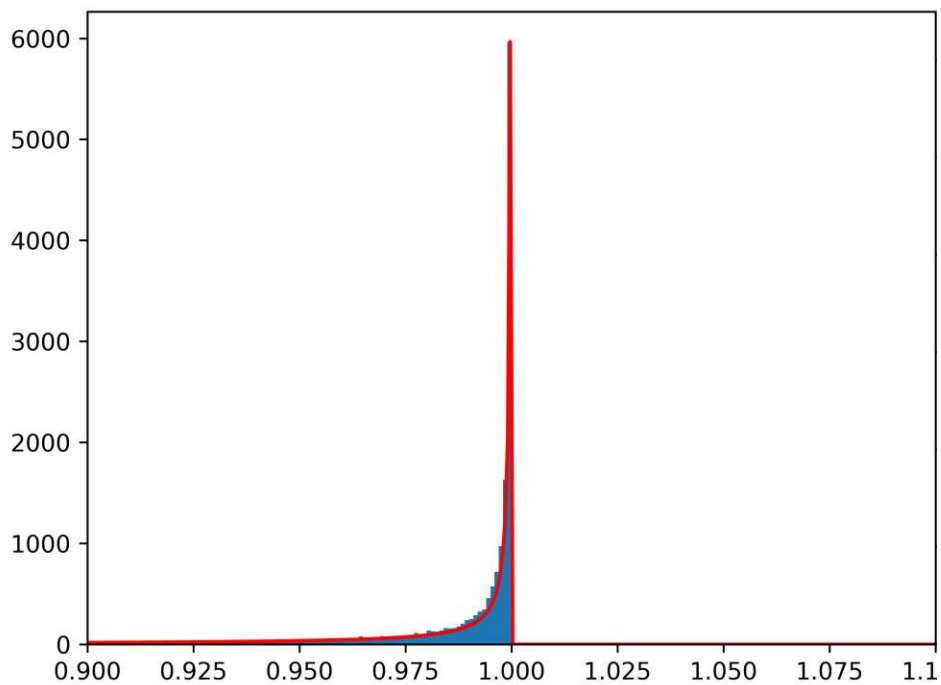


Abbildung 6.8: Realisationen des Nutzens des Wertprozesses eines stabilen Szenarios

7 Code

Die in diesem Kapitel dargestellten Listings beinhalten den Quellcode, der zur Ausführung des neuronalen Netzes benötigt wird sowie weitere Funktionen mit denen die in den Kapiteln 5 und 6 beschriebenen Resultate untersucht wurden.

Listing 7.1: Neuronales Netz

```

1 import numpy as np
2 import random
3 import tensorflow as tf
4 from tensorflow.keras.models import Sequential
5 from tensorflow.keras.layers import LSTM, Dense
6 from tensorflow.keras import initializers
7 from sklearn.model_selection import train_test_split
8
9
10 num_samples = 100000
11
12 def optimal_strategy(mu, sigma, alpha, kappa):
13     return mu / (alpha * sigma**2 - 2 * kappa)
14
15 mu = np.random.uniform(0, 0.2, num_samples)
16 sigma = np.random.uniform(0.2, 0.4, num_samples)
17 alpha = np.random.uniform(0.4, 0.6, num_samples)
18
19 kappa = (alpha*sigma**2)/2 * 0.99
20
21 theta = optimal_strategy(mu, sigma, alpha, kappa)
22
23 X = np.stack([mu, theta, sigma, alpha], axis=1)
24 y = kappa
25
26 X_train, X_test, y_train, y_test, mu_train, mu_test, sigma_train,
    sigma_test, alpha_train, alpha_test, kappa_train, kappa_test,
    theta_train, theta_test = train_test_split(X, y, mu, sigma, alpha
    , kappa, theta, test_size=0.2)
27
28 X_train = X_train.reshape(X_train.shape[0], 1, X_train.shape[1])
29 X_test = X_test.reshape(X_test.shape[0], 1, X_test.shape[1])
30
31 model = Sequential([
32     LSTM(4, input_shape=(X_train.shape[1], X_train.shape[2]),
        activation='linear', return_sequences=False,

```

```

        kernel_initializer=initializers.RandomUniform(minval
        =-0.5, maxval=0.5)),
33     Dense(4, activation='relu'),
34     Dense(1)
35 ])
36
37 model.compile(optimizer='adam', loss='mse')
38
39 model.fit(X_train, y_train, epochs=500, batch_size=32, verbose=1)
40
41 kappa_hat = model.predict(X_test)
42
43 def weights(model):
44     for layer in model.layers:
45         weights = layer.get_weights()
46         print("Layer:", layer.name)
47         if weights:
48             print("Weights:", weights[0])
49             print("Biases:", weights[1])
50         else:
51             print("Keine Gewichte")

```

Listing 7.2: Vergleich der Illiquiditätsparameter

```

1 import matplotlib.pyplot as plt
2
3 def compare_kappa(kappa, kappa_hat):
4     plt.figure(figsize=(8, 6))
5     plt.scatter(kappa, kappa_hat, alpha=0.5)
6     plt.xlabel("Kappa")
7     plt.ylabel("Kappa_Hat")
8     plt.grid(True)
9     plt.show()
10
11 def absolute_diff(kappa, kappa_hat):
12     diff = abs(kappa-kappa_hat)
13     quantile_95 = np.quantile(diff, 0.95)
14
15     plt.figure(figsize=(8, 6))
16     plt.hist(diff, bins=50, color='skyblue', edgecolor='black')
17     plt.xlabel("Absolute Differenz zwsichen kappa und kappa_hat")
18     plt.ylabel("Haeufigkeit")
19     plt.axvline(x=quantile_95, color='red', linestyle='--',
20                label='95% Quantile')
21     plt.grid(True)
22     plt.show()

```

Listing 7.3: Regimewechsel

```

1 def unstable_amount(sigma, alpha, kappa, amount):

```

```

2         counter = 0
3         for i in range(amount):
4             if alpha[i]*sigma[i]**2 < 2*kappa[i]:
5                 counter = counter + 1
6         return counter
7
8 def unstable_subset(sigma, alpha, kappa, amount):
9     unstable_sub = []
10    for i in range(amount):
11        if alpha[i]*sigma[i]**2 < 2*kappa[i]:
12            unstable_sub.append(i)
13    return unstable_sub
14
15 theta_hat = optimal_strategy(mu_test, sigma_test, alpha_test,
    kappa_hat)

```

Listing 7.4: Polynom

```

1 import scipy.stats as stats
2
3 def p(x, mu, sigma, alpha, kappa, index):
4     i = index
5     return x * (mu[i] + kappa[i] * x) - alpha[i]/2 * x**2 *
        sigma[i]**2
6
7 def p_plot(mu, sigma, alpha, kappa, index, theta_test, theta_hat):
8     x = np.linspace(-100, 100, 1000)
9     y = p(x, mu, sigma, alpha, kappa, index).tolist()
10    plt.plot(x, y)
11    plt.xlabel("x")
12    plt.ylabel("p(x)")
13    plt.title("Plot von p(x)")
14    plt.grid(True)
15    i = index
16    plt.scatter(theta_test[i], p(theta_test[i], mu, sigma, alpha,
        kappa, index), color='red')
17    plt.scatter(theta_hat[i], p(theta_hat[i], mu, sigma, alpha,
        kappa, index), color='green')
18    plt.show()

```

Listing 7.5: Verteilung des Wertprozesses

```

1 def value_process(mu, sigma, kappa, theta, amount):
2     B = np.random.normal(0, 1, amount)
3     return theta * (mu + kappa * theta) + theta * sigma * B
4
5 def u(x, alpha):
6     return 1 - np.exp(-x*alpha)
7
8 def generate_log_pdf(x, mu, sigma, alpha, kappa, theta):
9     a = (-1)*alpha[i]*theta[i]*(mu[i]+kappa[i]*theta[i])

```

```

10         b = (alpha[i]*sigma[i]*theta[i])**2
11         return stats.lognorm.pdf(1-x, s=b, loc=0, scale=np.exp(a))
12
13 def EV(mu, sigma, alpha, kappa, theta):
14     return theta * (mu + kappa * theta) * (-alpha)
15
16 def std(sigma, alpha, theta):
17     return sigma * theta * alpha
18
19 def utility_value_process(mu, sigma, kappa, alpha, theta):
20     B = np.random.normal(0, 1, 20000)
21     return 1 - np.exp((theta * (mu + kappa * theta) + sigma *
22                        theta * B) * (alpha))
23
24 def estimator(M, V, amount):
25     sum=0
26     Z=[]
27     for i in range(amount):
28         for j in range(M):
29             sum = sum + V[i*M+j]
30         Z.append(sum/M)
31     return Z
32
33 def EV_util(mu, sigma, alpha, kappa, theta):
34     return 1 - np.exp(-alpha * theta * (mu + (kappa - 0.5 *
35        alpha * sigma**2 ) * theta ))
36
37 def std_util(mu, sigma, alpha, kappa, theta):
38     return EV_util(mu, sigma, alpha, kappa, theta) * np.sqrt(np.
39        exp(alpha**2*sigma**2*theta**2)-1)

```

8 Conclusio

Die vorliegende Diplomarbeit beschäftigt sich mit dem Auffinden optimaler, stabiler Handelsstrategien in illiquiden Finanzmärkten unter Verwendung eines neuronalen Netzes. Die Auswirkungen der Handlungen eines*einer Large Traders*Traderin auf den Preisprozess eines Finanzprodukts werden sowohl im allgemeinen Fall als auch konkret anhand des illiquiden Bachelier Modells diskutiert. In diesem konkreten Modell lässt sich beobachten, dass auf Grund der Modellstruktur der erwartete Nutzen des Wertprozesses nach Ausübung der optimalen Handelsstrategie zur Maturität einer Lognormalverteilung folgt. Insbesondere gelingt eine konkrete Bestimmung der Parameter sowie der ersten Momente dieser Verteilung. Diese Folgerungen basieren auf der Erkenntnis in [2], dass die optimale Handelsstrategie unter den dort diskutierten Voraussetzungen und unter Einhaltung des stabilen Regimes konstant ist.

Nachdem die Herleitung der oben beschriebenen Verteilung auf der Existenz einer konstanten, stabilen, optimalen Strategie beruht, ist die Frage, inwieweit die Stabilität des Regimes veränderbar ist, von fundamentaler Bedeutung. Im Verlauf dieser Diplomarbeit wird ersichtlich, dass das stabile und das instabile Regime bei Anwendung eines neuronalen Netzes zur Bestimmung des Illiquiditätsparameters miteinander verschmelzen. Getrieben durch die daraus entstehende Diffusion ist die Wahrscheinlichkeit positiv, dass ein im stabilen Regime startender Preisprozess ins instabile Regime wechselt. Dieser Übergang hat insbesondere die Absenz einer stabilen, optimalen Handelsstrategie zur Folge.

Weiterführende Forschung in diesem Bereich könnte sich unter anderem der Untersuchung der Verteilung des Wertprozesses und dessen Nutzen in der illiquiden Adaptierung weiterer Modelle widmen. Gegebenenfalls müsste beziehungsweise könnte dann eine andere Nutzenfunktion herangezogen werden. Darüber hinaus könnte auch der zusätzliche Effekt einer stochastisch modellierten Volatilität weiter untersucht werden. Interessant gestalten könnte sich außerdem die Frage nach dem Einfluss der Erkenntnisse dieser Arbeit auf optimale Hedgingstrategien. Fest steht jedenfalls, dass die Erforschung illiquider Finanzmärkte und die Verknüpfung dieser mit Deep Learning Methoden insbesondere für Large Trader*innen auch zukünftig von zentraler Bedeutung sein wird.

Literatur

- [1] Bank, P. und Baum, D. 2004. „Hedging and portfolio optimization in financial markets with a large trader“. *Mathematical finance*. **14**.
- [2] Blümmel, T. und Rheinländer, T. 2017. „Financial markets with a large trader“. *The Annals of applied probability*. **27**.
- [3] Buehler, H., Gonon, L., Teichmann, J. und Wood, B. 2019. „Deep hedging“. *Quantitative finance*. **19**.
- [4] Carmona, R. A. und Nualart, D. 1990. *Nonlinear Stochastic Integrators, Equations and Flows*. Gordon und Breach, Science Publ.
- [5] Deru, M. und Ndiaye, A. 2020. *Deep Learning mit TensorFlow, Keras und TensorFlow.js*. Rheinwerk.
- [6] Fama, E. F. 1970. „Efficient capital markets: A review of theory and empirical work“. *The Journal of finance (New York)*. **25**.
- [7] Gibson, A. und Patterson, J. 2024. *Deep Learning*. O'Reilly Media, Inc.
- [8] Goodfellow, I., Bengio, Y. und Courville, A. 2016. *Deep Learning*. The MIT Press.
- [9] Jarrow, R. A. 1992. „Market manipulation, bubbles, corners, and short squeezes“. *Journal of financial and quantitative analysis*. **27**.
- [10] Kaltenböck, M. 2014. *Fundament Analysis*. Heldermann.
- [11] Kingma, D. P. und Ba, J. 2014. „Adam: A Method for Stochastic Optimization“. *CoRR*.
- [12] Kusolitsch, N. 2014. *Maß- und Wahrscheinlichkeitstheorie : Eine Einführung*. Springer Berlin Heidelberg.
- [13] Markowitz, H. 2000. *Portfolio Selection : Efficient Diversification of Investments*. Blackwell.
- [14] Melnikov, A. und Wan, H. 2021. „On modifications of the Bachelier model“. *Annals of finance*. **17**.
- [15] Pham, H., Rheinländer, T. und Schweizer, M. 1998. „Mean-variance hedging for continuous processes: New proofs and examples“. *Finance and stochastics*. **2**.
- [16] Pólya, G. 1973. *How to Solve it : A New Aspect of Mathematical Method*. Princeton Univ. Press.
- [17] Rheinländer, T. und Sexton, J. 2011. *Hedging Derivatives*. World Scientific Publ.
- [18] Schmock, U. 2022. *Lecture Notes in Stochastic Analysis for Financial and Actuarial Mathematics*.
- [19] Schweizer, M. 2001. *A Guided Tour through Quadratic Hedging Approaches*. Cambridge University Press.

- [20] Selvin, S. 2015. *The Log-Normal Distribution*. Cambridge University Press.
- [21] Williams, D. 2004. *Probability with Martingales*. Cambridge University Press.