

Numerische Methoden der Finanzmathematik

Lars Grüne
Mathematisches Institut
Fakultät für Mathematik und Physik
Universität Bayreuth
95440 Bayreuth
lars.gruene@uni-bayreuth.de
www.math.uni-bayreuth.de/~lgruene/

Vorlesungsskript
1. Auflage
Sommersemester 2009

Inhaltsverzeichnis

| | | |
|----------|---|-----------|
| 1 | Einführung und Grundlagen | 1 |
| 1.1 | Optionen | 2 |
| 1.2 | Das Marktmodell | 3 |
| 1.3 | Grundbegriffe aus der Stochastik | 5 |
| 2 | Optionen und ihre Bewertung | 11 |
| 2.1 | Grundidee der Optionsbewertung | 12 |
| 2.2 | Analytische Aussagen über Optionswerte | 13 |
| 2.3 | Folgerungen aus der risikoneutralen Bewertung | 17 |
| 3 | Die Binomialmethode | 19 |
| 3.1 | Zeitdiskrete stochastische dynamische Systeme | 19 |
| 3.2 | Das Binomialmodell eines Aktienkurses | 20 |
| 3.3 | Optionswertberechnung mit dem Binomialmodell | 21 |
| 3.4 | Die Parameter im Binomialmodell | 23 |
| 4 | Stochastische Differentialgleichungen | 25 |
| 4.1 | Der Wiener Prozess | 25 |
| 4.2 | Das Itô-Integral | 27 |
| 4.2.1 | Das Itô-Lemma | 30 |
| 4.2.2 | Die geometrische Brownsche Bewegung | 32 |
| 4.3 | Ein anspruchsvolleres Kursmodell | 34 |
| 5 | Die Monte-Carlo Methode | 35 |
| 5.1 | Grundidee | 35 |
| 5.2 | Konvergenzgeschwindigkeit | 37 |
| 5.3 | Antithetische Zufallsvariablen | 38 |

| | | |
|-----------|--|-----------|
| 6 | Numerische Lösung stochastischer DGLen | 41 |
| 6.1 | Stochastische Konvergenz- und Approximationsbegriffe | 41 |
| 6.2 | Approximationen des Wiener-Prozesses | 44 |
| 6.3 | Das stochastische Euler-Verfahren | 46 |
| 6.4 | Anwendung auf die Optionsbewertung | 48 |
| 6.5 | Das Binomialmodell als schwache Approximation | 50 |
| 7 | Die Black-Scholes Gleichung | 53 |
| 8 | Hedging und die “Griechen” | 59 |
| 8.1 | Zeitdiskretes Hedging | 59 |
| 8.2 | Monte-Carlo für das Δ | 63 |
| 9 | Einführung in die Finite Differenzen-Methode | 65 |
| 9.1 | Black-Scholes- und Wärmeleitungsgleichung | 66 |
| 9.2 | Grundidee der Finiten Differenzen | 67 |
| 9.3 | Lösung der Finiten Differenzgleichungen | 70 |
| 9.4 | Konsistenz, Stabilität und Konvergenz | 72 |
| 10 | Finite Differenzen für die Black-Scholes Gleichung | 77 |
| 10.1 | Wahl der numerischen Randbedingungen | 77 |
| 10.2 | Direkte Diskretisierung der Black-Scholes-Gleichung | 79 |
| 10.3 | Finite Differenzen und Binomialmethode | 81 |
| | Literaturverzeichnis | 84 |
| | Index | 86 |

Kapitel 1

Einführung und Grundlagen

Die Finanzmathematik ist ein relativ junges und im letzten Jahrzehnt stark gewachsenes Teilgebiet der angewandten Mathematik. Aufgrund der ständig steigenden Komplexität der weltweit gehandelten Finanzinstrumente sind diese — wenn überhaupt — oft nur noch mit mathematischen Methoden wirklich zu verstehen.

Der Bereich der Finanzmathematik, den wir hier näher betrachten wollen, beschäftigt sich weniger mit Börsenspekulationen (obwohl das nicht immer streng zu trennen ist, vgl. Abschnitt 1.1) sondern vielmehr mit Finanzinstrumenten, die vor allem zur Absicherung von tatsächlichen Geschäften dienen, wie z.B. zur Absicherung gegen Wechselkursschwankungen, die für international agierende Firmen zum Alltagsgeschäft gehört. Als Modellproblem betrachten wir dazu in dieser Vorlesung die Bewertung europäischer und amerikanischer Optionen. Diese Problemklasse hat den Vorteil, dass die mathematische Theorie und die zugehörige Numerik inzwischen relativ gut verstanden und ausgearbeitet ist. Zudem lassen sich die Methoden auf komplexere Finanzinstrumente erweitern, wenn man die Grundprinzipien einmal verstanden hat. Als ergänzende Literatur empfehle ich dabei die Bücher

- [2] M. Günther und A. Jüngel, Finanzderivate mit MATLAB. Vieweg, 2003
- [3] D.J. Higham, An introduction to financial option valuation. Mathematics, stochastics and computation. Cambridge University Press, 2004
- [5] R. Seydel, Tools for computational finance. Springer, 2009 (4. Auflage)¹,

die auch die Grundlage dieses Skripts bilden.

Im Rahmen der globalen Finanzkrise stellt sich natürlich die Frage, inwieweit gerade die in dieser Vorlesung behandelten Modelle und Finanzprodukte sowie die zugehörigen numerischen Methoden Ursache der derzeitigen Probleme sind. Da ich kein Experte für Finanzmärkte bin, möchte ich hier keine Spekulationen anstellen und nur einige offensichtliche Zusammenhänge — speziell in den folgenden beiden Abschnitten — aufzeigen. Was die derzeitige Krise für die Zukunft der Finanzmathematik bedeutet, ist aus meiner Sicht

¹Von [5] gibt es auch eine — allerdings ältere und weniger umfangreiche — deutsche Version: R. Seydel, Einführung in die numerische Berechnung von Finanz-Derivaten, Springer, 2000.

derzeit noch nicht absehbar. Unstrittig erscheint mir aber, dass das Verständnis komplexer Finanzinstrumente und ihres mathematischen Hintergrundes gerade in der Krise wichtig sind.

1.1 Optionen

Die Optionsbewertung ist eine der einfachsten Aufgabenstellungen der Finanzmathematik. Trotzdem ist sie alles andere als trivial, zudem können komplexere finanzmathematische Probleme oft mit einer Verallgemeinerung der Methoden für Optionen gelöst werden. Daher werden wir sie in dieser Vorlesung als beispielhafte finanzmathematische Anwendung betrachten. Wir definieren nun zunächst, was eine Option überhaupt ist und erläutern kurz das Problem der Optionsbewertung. In den nächsten Kapiteln werden wir dann Schritt für Schritt verschiedene Methoden zur numerischen Lösung des Problems kennenlernen.

Ganz allgemein ist eine Option ein Vertrag, der dem Inhaber die Möglichkeit (aber nicht die Verpflichtung) gibt, einen *Basiswert* (z.B. ein Aktienpaket oder einen festgelegten Betrag einer Fremdwährung) zu einem vorher vereinbarten *Ausübungspreis* K zu kaufen (*Call-Option*) oder zu verkaufen (*Put-Option*). Optionen und andere ähnliche Finanzinstrumente werden *Derivate* genannt, da sich ihr Wert von dem Basiswert ableitet.

Wir betrachten dabei sowohl die *europäischen Optionen* als auch die *amerikanischen Optionen*. Bei den europäischen Optionen muss der Kauf (bzw. Verkauf) zu einem von vornherein festgelegten *Ausübungszeitpunkt* T stattfinden. Ist dieser Ausübungszeitpunkt T erreicht, so kann der Inhaber die Option ausüben, also den Basiswert zu dem vorher festgelegten Ausübungspreis kaufen (bzw. verkaufen), oder er kann die Option verfallen lassen, den Basiswert also nicht kaufen (bzw. verkaufen).

Bei der amerikanischen Option kann die Ausübung zu jedem beliebigen Zeitpunkt $t \in [0, T]$ durchgeführt werden, wobei T ein vertraglich festgelegter Zeitpunkt ist. Der Startzeitpunkt der Option wird hier per Definition immer als 0 festgelegt, was durch einfache Verschiebung des tatsächlichen zeitlichen Arguments natürlich o.B.d.A. möglich ist. In der Praxis wird der tatsächliche Kauf oder Verkauf dabei übrigens häufig gar nicht ausgeführt, sondern der Differenzbetrag zwischen Ausübungspreis und Marktpreis ausgezahlt.

Optionen können vielfältig eingesetzt werden. Eine offensichtliche Anwendung ist die Absicherung gegenüber Währungsschwankungen: Erwartet z.B. ein europäisches Unternehmen in T Monaten eine Zahlung von 1 Mio. US-Dollar, so könnte es für diesen Zeitpunkt eine Put-Option auf den Verkauf von 1 Mio. US-Dollar für K Euro erwerben und hätte sich damit einen garantierten Wechselkurs gesichert. Liegt der Wert der 1 Mio. US-Dollar dann zum Ausübungszeitpunkt T unter dem Wert K Euro, so würde die Option ausgeübt um die garantierten Einnahmen in Euro zu realisieren. Liegt der Wert oberhalb der K Euro, so würde die Option nicht ausgeübt und das Geld am Markt umgetauscht. Analog kann man sich z.B. gegen schwankende Rohstoffpreise absichern.

Optionen eignen sich aber auch hervorragend zur Spekulation. Vermutet ein Spekulant, dass der Wert eines Aktienpakets eines Unternehmens vom heutigen Kurs K_1 bis zum Zeitpunkt T auf den Wert $K_2 < K_1$ fällt, so kann er eine Put-Option mit Ausübungspreis $K = K_1$ erwerben. Geht die Spekulation auf, so kann er zum Ausübungszeitpunkt das

Aktienpaket für den günstigeren Preis K_2 an der Börse kaufen und durch Ausübung der Option für den Preis $K = K_1$ teurer verkaufen.

Klar ist nun, dass der Ausgeber (oder Emittent) der Option (typischerweise eine Bank oder eine Versicherung) eine solche Option nicht umsonst ausgeben wird sondern einen Preis verlangen wird, der das mit der Option behaftete Risiko widerspiegelt. Das Problem der *Optionsbewertung* ist nun: Welchen Wert besitzt diese Option selbst zu einem Zeitpunkt $t < T$? Aus diesem Wert kann dann die Gebühr (typischerweise natürlich noch mit einem Aufschlag) errechnet werden. Der so ermittelte Wert der Option wird auch *fairer Preis* genannt, denn in der Praxis werden auch viele Optionen selbst wieder an der Börse gehandelt, so dass der tatsächliche Preis der Option wiederum durch den Markt bestimmt wird. Trotzdem ist die Berechnung des fairen Preises zur Einschätzung des Wertes einer Option auf jeden Fall theoretisch wichtig, im Falle spezialisierter Optionen, für die kein großer Markt besteht, aber auch ganz praktisch.

Ein eng verwandtes Problem ist die Frage, wie sich der Ausgeber der Option nun selbst wieder gegen das Risiko absichert. Je besser das gelingt, desto günstiger kann der Preis kalkuliert werden. Wir werden hier das Prinzip der *risikoneutralen Bewertung* zu Grunde legen. Dieses Prinzip hat seine Begründung in einer aus der Black-Scholes Theorie stammenden *Hedging-Strategie* (Absicherungsstrategie) und wird in dem entsprechenden Kapitel begründet. Anwenden lässt sich das Prinzip allerdings auch, ohne die entsprechenden Hintergründe zu kennen, was wir in den folgenden Kapiteln auch machen werden.

1.2 Das Marktmodell

Um den Finanzmarkt mathematisch modellieren zu können, müssen wir Annahmen treffen, die sich in mathematische Aussagen umsetzen lassen. Hierzu existieren eine ganze Reihe sehr ausgefeilter Modelle; hier werden wir uns aber auf die folgenden einfachen Grundannahmen beschränken, die für unsere Zwecke ausreichen:

- Zinsen werden kontinuierlich mit dem jährlichen Zinssatz r berechnet: für eine Anlage von x Euro für t Jahre ergibt sich der Zins zu $e^{rt}x$.
- Geld kann jederzeit zu dem zeitlich unveränderlichen Zinssatz r über einen beliebigen Zeitraum festverzinslich angelegt werden.
- Kredite können jederzeit zu dem gleichen Zinssatz r für einen beliebigen Zeitraum aufgenommen werden.
- Der Markt ist *arbitragefrei*, d.h. es gibt keine risikofreie Anlage, mit der ein Betrag x so angelegt werden kann, dass er nach einer Laufzeit T eine Auszahlung $y > e^{rT}x$ liefert.

Einige dieser Annahmen sind lediglich vereinfachend und können (allerdings mit z.T. nicht unerheblichem mathematischem Aufwand) gelockert werden, z.B. die Annahme, dass der Zinssatz r nicht mit der Zeit schwankt oder die Annahme, dass der Zinssatz für Anlagen und Kredite gleich ist. Andere Annahmen, wie z.B. die Verfügbarkeit von Krediten oder die Arbitragefreiheit sind wesentlich für die im Folgenden entwickelten Methoden. Beide sind

durchaus problematisch: gerade in der derzeitigen Finanzkrise ist die Annahme, dass Kredite in beliebigem Umfang verfügbar sind, sicherlich nicht gerechtfertigt und möglicherweise ist diese falsche Annahme eine der Ursachen für das Ausmaß der derzeitigen Krise. Auch die Arbitragefreiheit gilt in der Praxis nicht in dieser Reinform. Arbitragefreiheit verbietet z.B., dass Produkte auf verschiedenen Finanzmärkten zu verschiedenen Preisen gehandelt werden, da man sonst durch Kauf eines Produktes auf dem “billigeren” und Verkauf auf dem “teureren” Markt einen risikofreien Gewinn erzielen könnte. Tatsächlich gibt es aber solche Preisunterschiede, allerdings sind diese oft nur klein oder nur von kurzer Dauer, so dass wir die Annahme näherungsweise als erfüllt ansehen können.

Um zu zeigen, wie man aus diesen Annahmen präzise mathematische Aussagen ableiten kann, betrachten wir die folgende Frage der *Abzinsung* (oder *Diskontierung*), die später für die Optionsbewertung eine wichtige Rolle spielen wird: Angenommen, wir kennen den Wert $P(T) > 0$ eines Produktes P (Rohstoff, Anlage, Währung etc.) zu einem zukünftigen Zeitpunkt T . Was ist dann der Wert $P(t)$ dieses Produktes zu einem früheren Zeitpunkt $t < T$? Als konkretes Beispiel könnte $P(T)$ die vertraglich gesicherte Auszahlung einer Anlage zu einem festgelegten zukünftigen Termin T sein und die Frage wäre, welchen Betrag $P(t)$ wir sinnvollerweise zahlen sollten, wenn wir diese Anlage zum heutigen Zeitpunkt t einer anderen Person abkaufen wollten.

Um den Preis $P(t)$ zu ermitteln, vergleichen wir den erzielbaren Gewinn mit dem Gewinn aus einer festverzinslichen Anlage.

Nehmen wir an, wir besitzen das Produkt P und würden es zum Zeitpunkt t verkaufen und den Erlös $P(t)$ festverzinslich anlegen. Auf diese Weise erhielten wir gemäß unserer Annahmen zum Zeitpunkt T gerade $B(T) = e^{r(T-t)}P(t)$. Hätten wir P gehalten, so erhielten wir zum Zeitpunkt T gerade $P(T)$. Gälte für den Preis $P(t)$ nun

$$P(t) > e^{r(t-T)}P(T),$$

so hätten wir zum Zeitpunkt T einen risikofreien Gewinn von

$$B(T) - P(T) = e^{r(T-t)}P(t) - P(T) > e^{r(T-t)}e^{r(t-T)}P(T) - P(T) = 0$$

erzielt, was nach der Annahme der Arbitragefreiheit nicht möglich ist.

Andererseits könnten wir zum Zeitpunkt t einen Kredit der Höhe $P(t)$ aufnehmen und dafür das Produkt zum Preis $P(t)$ erwerben. Zum Zeitpunkt T erhalten wir dann vertragsgemäß $P(T)$ und zahlen davon den Kredit mit Zinsen, also $B(T) = e^{r(T-t)}P(t)$ zurück. Gälte für den Preis $P(t)$ nun

$$P(t) < e^{r(t-T)}P(T),$$

so hätten wir zum Zeitpunkt T einen risikofreien Gewinn von

$$P(T) - B(T) = P(T) - e^{r(T-t)}P(t) > P(T) - e^{r(T-t)}e^{r(t-T)}P(T) = 0$$

erzielt, was nach der Annahme der Arbitragefreiheit wiederum nicht möglich ist. Wir haben damit den folgenden Satz bewiesen.

Satz 1.1 Der aus den Marktannahmen folgende faire Preis des Produktes P zum Zeitpunkt $t < T$ beträgt

$$P(t) = e^{r(t-T)}P(T).$$

Bemerkung 1.2 Der Satz lässt sich wie folgt verallgemeinern (Übungsaufgabe): Wenn der Wert von P zum Zeitpunkt T garantiert die Ungleichung $P(T) \geq P_0$ für ein $P_0 \in \mathbb{R}$ erfüllt, so folgt $P(t) \geq e^{r(t-T)}P_0$. Analog gilt dies für “ \leq ”. Beachte, dass $P(T)$ hier negativ sein kann. P wäre in diesem Fall eine Anlage, bei der zum Zeitpunkt T eine Zahlungsverpflichtung eintritt, z.B. ein Kredit mit Fälligkeit zum Zeitpunkt T . \square

1.3 Grundbegriffe aus der Stochastik

Da der Wert einer Option von der (natürlich unbekannt) zukünftigen Entwicklung des zu Grunde liegenden Kurses abhängt, werden wir im Verlauf dieser Vorlesung verschiedene Modelle zur mathematischen Modellierung von Kursverläufen kennen lernen. Da niemand ernsthaft behaupten kann, dass man mit mathematischen Methoden den Verlauf von Aktienkursen, Wechselkursen etc. vorhersagen kann, verwendet man in der Finanzmathematik üblicherweise keine *deterministischen* Modelle, wie z.B. in den Vorlesungen “Numerik dynamischer Systeme” oder “Mathematische Kontrolltheorie”. Statt dessen benutzt man hier *stochastische Modelle*, mit denen ausgehend vom aktuellen Kurs viele zukünftige Kursverläufe simuliert werden können, welche gewisse aus der Vergangenheit des Kursverlaufs ermittelten statistische Werte, wie z.B. die Varianz (also die Streuung um den Mittelwert) des Kurses gut wiedergeben.

In diesem Abschnitt der Einführung wiederholen wir daher einige Grundbegriffe aus der Stochastik. Dies dient in erster Linie der Festlegung der im weiteren Verlauf der Vorlesung verwendeten Notation; wir gehen davon aus, dass die Grundbegriffe an sich bekannt sind (dieser Abschnitt ersetzt also keine Einführung in die Stochastik).

Ein *Wahrscheinlichkeitsraum* besteht aus einem Tripel $(\Omega, \Sigma, \mathbb{P})$ mit:

- Ω ist die Menge der Elementarereignisse $\omega \in \Omega$
- Σ ist eine σ -Algebra auf Ω , d.h. eine Menge von Teilmengen von Ω mit den Eigenschaften
 - (1) $\emptyset, \Omega \in \Sigma$
 - (2) wenn $A \in \Sigma$ so ist auch $A^c = \Omega \setminus A \in \Sigma$
 - (3) für $A_1, A_2, \dots \in \Sigma$ ist auch $\bigcup_{i=1}^{\infty} A_i \in \Sigma$ und $\bigcap_{i=1}^{\infty} A_i \in \Sigma$
- $\mathbb{P} : \Sigma \rightarrow [0, 1]$ ist ein Wahrscheinlichkeitsmaß auf Σ , d.h. es gilt $\mathbb{P}(\emptyset) = 0$, $\mathbb{P}(\Omega) = 1$ und

$$\mathbb{P}\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} \mathbb{P}(A_i) \text{ für alle Mengen } A_i \in \Sigma \text{ mit } A_i \cap A_j = \emptyset \text{ für alle } i \neq j.$$

Für eine Menge $\mathcal{A} \subseteq \mathcal{P}(\Omega)$ von Teilmengen von Ω (\mathcal{P} bezeichnet die Potenzmenge) bezeichnen wir mit $\sigma(\mathcal{A})$ die kleinste σ -Algebra, die \mathcal{A} enthält.

Um analytisch arbeiten zu können, werden wir Ω und \mathbb{P} wie üblich meist nicht direkt sondern durch (i.A. vektorwertige) Zufallsvariablen betrachten, also durch messbare Abbildungen

$$X : \Omega \rightarrow \mathbb{R}^n.$$

Hierbei heißt die Abbildung X *messbar*, wenn gilt

$$X^{-1}(B) := \{\omega \in \Omega \mid X(\omega) \in B\} \in \Sigma$$

für alle Mengen B aus der *Borel- σ -Algebra* \mathcal{B} .

$$\mathcal{B} = \sigma(\{(a_1, b_1] \times \dots \times (a_n, b_n] \mid a_i, b_i \in \mathbb{R}, a_i < b_i\})$$

auf \mathbb{R}^n . Statt $X^{-1}(B)$ schreiben wir alternativ auch “ $X \in B$ ”. Ebenso werden wir z.B. für eine zweidimensionale vektorwertige Zufallsvariable $X = (X_1, X_2)^T$ Schreibweisen wie “ $X_1 \leq X_2$ ” verwenden, was kurz für $X \in B = \{(x_1, x_2)^T \in \mathbb{R}^2 \mid x_1 \leq x_2\}$ steht.

Für messbare Zufallsvariablen definiert man ihre *Wahrscheinlichkeitsverteilung*

$$\mathbb{P}_X : \mathcal{B} \rightarrow [0, 1], \quad \mathbb{P}_X(B) := \mathbb{P}(X^{-1}(B)).$$

Beispiel 1.3 Es sei $\Omega = (0, 1]$, $\Sigma = \mathcal{B}|_{(0,1]}$, und \mathbb{P} die Gleichverteilung, also $\mathbb{P}((a, b]) = b - a$. Betrachte die Zufallsvariable

$$X(\omega) = \begin{cases} -1, & \omega \in (0, 1/2] \\ 1, & \omega \in (1/2, 1] \end{cases}$$

Dann sind die möglichen Urbilder $X^{-1}(B)$ gerade \emptyset , $(0, 1/2]$, $(1/2, 1]$ und Ω , die natürlich allesamt in $\Sigma = \mathcal{B}|_{(0,1]}$ liegen. X ist also messbar. Für \mathbb{P}_X ergibt sich

$$\mathbb{P}_X(B) = \begin{cases} 0, & -1 \notin B, 1 \notin B \\ 1/2, & -1 \in B, 1 \notin B \\ 1/2, & -1 \notin B, 1 \in B \\ 1, & -1 \in B, 1 \in B \end{cases}.$$

Tatsächlich ist die gewählte Borel’sche σ -Algebra $\Sigma = \mathcal{B}|_{(0,1]}$ viel größer als nötig, um Messbarkeit von X zu erhalten. Aus den oben angegebenen Urbilder sieht man leicht, dass die *minimale* σ -Algebra auf Ω , bezüglich der X messbar ist, durch

$$\{\emptyset, (0, 1/2], (1/2, 1], \Omega\}$$

gegeben ist. □

Die in diesem Beispiel angegebene minimale σ -Algebra bezeichnen wir mit $\sigma(X)$. Formal ist diese gegeben durch

$$\sigma(X) := \sigma(\{X^{-1}(B) \mid B \in \mathcal{B}\}).$$

Messbarkeit von X bzgl. einer σ -Algebra Σ ist dann äquivalent zu der Inklusion

$$\sigma(X) \subseteq \Sigma.$$

Wenn wir ein $\omega \in \Omega$ wählen und $X(\omega)$ auswerten (was dem Durchführen eines Zufallsexperiments mit nachfolgender Auswertung von X entspricht), so nennen wir den Wert $X(\omega)$ eine *Realisierung* von X .

Falls X nur endlich viele Werte x_1, \dots, x_m annehmen kann, ist ihr *Erwartungswert* als

$$E(X) = \sum_{i=1}^m x_i \mathbb{P}_X(\{x_i\}) \quad (1.1)$$

gegeben. Falls die Wertemenge von X kontinuierlich ist, werden wir stets annehmen, dass zu \mathbb{P}_X eine *Dichtefunktion* f_X existiert, so dass

$$\mathbb{P}_X(B) = \int_B f_X(x) dx \quad (1.2)$$

gilt. Der Erwartungswert von X ist dann als

$$E(X) = \int_{\mathbb{R}^n} x f_X(x) dx$$

definiert. Die *Varianz* ist für eine reellwertige Zufallsvariable $X : \Omega \rightarrow \mathbb{R}$ gegeben durch

$$\text{Var}(X) = E((X - E(X))^2) = E(X^2) - 2E(XE(X)) + E(X)^2 = E(X^2) - E(X)^2.$$

Für eine vektorwertige Zufallsvariable $X : \Omega \rightarrow \mathbb{R}^n$ definiert man die *Kovarianzmatrix*

$$\text{Cov}(X) = E\left((X - E(X))(X - E(X))^T\right) \in \mathbb{R}^{n \times n}.$$

Beachte, dass die Diagonale dieser Matrix gerade die Varianzen $\text{Var}(X_j)$ der Komponenten von $X = (X_1, \dots, X_n)^T$ enthält.

Wenn wir eine Zufallsvariable X betrachten und zusätzlich eine Menge $A \subset \Omega$ kennen, in der ω liegt, so können wir diese Information verwenden, wenn wir die Wahrscheinlichkeit von $X \in B$ berechnen wollen. Formal wird dies über die *bedingte Wahrscheinlichkeit*

$$\mathbb{P}_X(B | A) = \frac{\mathbb{P}(X^{-1}(B) \cap A)}{\mathbb{P}(A)}$$

gemessen. In dieser Definition muss $\mathbb{P}(A) \neq 0$ vorausgesetzt werden, allerdings gibt es allgemeinere Definitionen, die auch für $\mathbb{P}(A) = 0$ sinnvoll sind, allerdings weitergehende maßtheoretische Techniken benötigen, die wir hier nicht einführen wollen (vgl. [1, Definition 7.1.1]). Mittels der charakteristischen Funktion

$$\chi_A(x) = \begin{cases} 1, & x \in A \\ 0, & x \notin A \end{cases}$$

ist der *bedingte Erwartungswert* einer Zufallsvariablen durch

$$E(X | A) = \frac{E(X \cdot \chi_A)}{\mathbb{P}(A)}$$

definiert. Auch diese Definition lässt sich mit anspruchsvolleren maßtheoretischen Überlegungen auf den Fall $\mathbb{P}(A) = 0$ erweitern (vgl. [1, Definition 7.1.2] oder [4, Abschnitt 1.4.3]).

Ein wichtiger Spezialfall dieser bedingten Größen tritt auf, wenn wir zwei Zufallsvariablen X_1 und X_2 betrachten und die Menge A als $X_2^{-1}(B_2)$ gegeben ist, wenn wir also die Größen

$$\mathbb{P}_{X_1}(B_1 | X_2 \in B_2) \quad \text{bzw.} \quad E(X_1 | X_2 \in B_2)$$

betrachten.

Beispiel 1.4 Es sei $(\Omega, \Sigma, \mathbb{P})$ der Wahrscheinlichkeitsraum, der das (gleichzeitige) Würfeln mit zwei Würfeln beschreibt. Jedes Elementarereignis ist also ein Paar $\omega = (\omega_1, \omega_2) \in \Omega = \{1, \dots, 6\}^2$ von Würfelwerten, wobei die erste Komponente ω_1 das Ergebnis des ersten Würfels und ω_2 das Ergebnis des zweiten Würfels beschreibt (wir nehmen an, dass die zwei Würfel unterscheidbar sind). Als Σ wählen wir alle möglichen Teilmengen von Ω . Offenbar ist (für ideale Würfel) $\mathbb{P}(A) = |A|/36$ für jede Teilmenge $A \subseteq \Omega$ mit $|A|$ Elementen, da Ω gerade $36 = 6 \times 6$ Ereignisse enthält, die alle gleich wahrscheinlich sind. Wir betrachten nun drei Zufallsvariablen, die als Wert das Ergebnis des ersten Würfels, des zweiten Würfels und die Summe der Würfelwerte ausgeben, also $X_1(\omega) = \omega_1$, $X_2(\omega) = \omega_2$ und $X_3(\omega) = \omega_1 + \omega_2$. Wir betrachten die folgenden Ereignismengen für die Zufallsvariablen X_1 , X_2 bzw. X_3 :

$$\begin{aligned} X_1 \in B_1 &= \{1\} && \text{(erster Würfel würfelt 1)} \\ X_2 \in B_2 &= \{6\} && \text{(zweiter Würfel würfelt 6)} \\ X_3 \in B_3 &= \{10, 11, 12\} && \text{(Würfelsumme ist } \geq 10) \end{aligned}$$

Hierfür ergibt sich

$$\begin{aligned} X_1^{-1}(B_1) &= \{(1, 1), (1, 2), (1, 3), (1, 4), (1, 5), (1, 6)\} \\ X_2^{-1}(B_2) &= \{(1, 6), (2, 6), (3, 6), (4, 6), (5, 6), (6, 6)\} \\ X_3^{-1}(B_3) &= \{(4, 6), (5, 5), (5, 6), (6, 4), (6, 5), (6, 6)\} \end{aligned}$$

und wir erhalten die unbedingten Wahrscheinlichkeiten

$$\begin{aligned} \mathbb{P}_{X_1}(B_1) &= 1/6 \\ \mathbb{P}_{X_2}(B_2) &= 1/6 \\ \mathbb{P}_{X_3}(B_3) &= 1/6 \end{aligned}$$

und die bedingten Wahrscheinlichkeiten

$$\begin{aligned} \mathbb{P}_{X_2}(B_2 | X_1 \in B_1) &= 1/6 \\ \mathbb{P}_{X_3}(B_3 | X_1 \in B_1) &= 0 \\ \mathbb{P}_{X_3}(B_3 | X_2 \in B_2) &= 1/2 \end{aligned}$$

□

Der bedingte Erwartungswert $E(X_1 | X_2 \in B_2)$ liefert eine reelle Zahl. Man kann dieses Konzept so erweitern, dass man eine Zufallsvariable erhält, nämlich indem man die Abbildung

$$E(X_1 | X_2) : \omega \mapsto E(X_1 | X_2 = X_2(\omega))$$

definiert. Dies ist eine Zufallsvariable, die messbar bzgl. $\sigma(X_2)$ ist, und für deren Erwartungswert gerade

$$E(E(X_1 | X_2)) = E(X_1) \quad (1.3)$$

gilt. Gleichung (1.3) lässt sich für den Fall, dass X_2 nur abzählbar viele Werte annimmt, leicht beweisen: In diesem Fall existieren abzählbar viele Mengen der Form $X_2^{-1}(X_2(\omega))$, die wir mit A_1, A_2, \dots bezeichnen. Damit folgt

$$E(E(X_1 | X_2)) = \sum_{i=1}^{\infty} E(X_1 | A_i) \mathbb{P}_{X_1}(A_i) = \sum_{i=1}^{\infty} E(X_1 \chi_{A_i}) = E\left(\sum_{i=1}^{\infty} X_1 \chi_{A_i}\right) = E(X_1).$$

Im überabzählbaren Fall gilt (1.3) auch, benötigt aber wiederum tiefere maßtheoretische Techniken zum Beweis (vgl. [1, Theorem 7.1.1(3)] oder [4, Section 1.4.4]). Statt $E(X_1 | X_2)$ schreibt man auch

$$E(X_1 | \sigma(X_2)). \quad (1.4)$$

Ein wichtiges Konzept ist die stochastische Unabhängigkeit zweier (oder mehrerer) Zufallsvariablen. Anschaulich bedeutet dies, dass uns die Kenntnis der Werte einer Zufallsvariablen X_1 keinerlei Informationen über die Werte der anderen Zufallsvariablen X_2 liefert. Formal:

Definition 1.5 Die Zufallsvariablen X_1, X_2, \dots, X_k heißen *unabhängig*, falls

$$\mathbb{P}(X_1^{-1}(B_1) \cap \dots \cap X_k^{-1}(B_k)) = \mathbb{P}(X_1^{-1}(B_1)) \cdot \dots \cdot \mathbb{P}(X_k^{-1}(B_k)) \quad (1.5)$$

für alle $B_i \in \mathcal{B}$, $i = 1, \dots, k$, gilt. □

Für unabhängige Zufallsvariablen X_1, X_2, \dots, X_k gilt

$$E(X_1 \cdot \dots \cdot X_k) = E(X_1) \cdot \dots \cdot E(X_k). \quad (1.6)$$

Für eine Abbildung $h : (\mathbb{R}^n)^k \rightarrow \mathbb{R}^l$ und k Zufallsvariablen X_1, \dots, X_k ist $Y = h(X_1, \dots, X_k)$ wieder eine Zufallsvariable. Wenn dann X_1, X_2, \dots, X_k, Z unabhängig sind, so sind auch Y und Z unabhängig.

Beispiel 1.6 Betrachte die in Beispiel 1.4 definierten Zufallsvariablen. Für beliebige Mengen $B_1, B_2 \subseteq \{1, \dots, 6\}$ gilt hier

$$X_1^{-1}(B_1) = \{(\omega_1, \omega_2) \mid \omega_1 \in B_1, \omega_2 \in \{1, \dots, 6\}\},$$

$$X_2^{-1}(B_2) = \{(\omega_1, \omega_2) \mid \omega_2 \in B_2, \omega_1 \in \{1, \dots, 6\}\}$$

und

$$X_1^{-1}(B_1) \cap X_2^{-1}(B_2) = \{(\omega_1, \omega_2) \mid \omega_1 \in B_1, \omega_2 \in B_2\} = B_1 \times B_2.$$

Also gilt

$$|X_1^{-1}(B_1)| = 6|B_1|, \quad |X_2^{-1}(B_2)| = 6|B_2|, \quad |X_1^{-1}(B_1) \cap X_2^{-1}(B_2)| = |B_1| |B_2|$$

und damit

$$\begin{aligned}\mathbb{P}(X_1^{-1}(B_1) \cap X_2^{-1}(B_2)) &= |B_1||B_2|/36 \\ &= (6|B_1|/36)(6|B_2|/36) = \mathbb{P}(X_1^{-1}(B_1))\mathbb{P}(X_2^{-1}(B_2)).\end{aligned}$$

Also sind X_1 und X_2 unabhängig. Andererseits gilt für die in Beispiel 1.4 definierten Mengen B_1 und B_3

$$\mathbb{P}(X_1^{-1}(B_1) \cap X_3^{-1}(B_3)) = \mathbb{P}(\emptyset) = 0$$

aber

$$\mathbb{P}(X_1^{-1}(B_1))\mathbb{P}(X_3^{-1}(B_3)) = \mathbb{P}_{X_1}(B_1)\mathbb{P}_{X_3}(B_3) = (1/6)(1/6) = 1/36 \neq 0,$$

weswegen X_1 und X_3 nicht unabhängig sind. □

Kapitel 2

Optionen und ihre Bewertung

Wie bereits in Abschnitt 1.1 beschrieben, ist eine Option ein Vertrag, der dem Inhaber die Möglichkeit (aber nicht die Verpflichtung) gibt, einen *Basiswert* (z.B. ein Aktienpaket oder einen festgelegten Betrag einer Fremdwährung) zu einem vorher vereinbarten *Ausübungspreis* K zu kaufen (*Call-Option*) oder zu verkaufen (*Put-Option*). Optionen und andere ähnliche Finanzinstrumente werden *Derivate* genannt, da sich ihr Wert von dem Basiswert ableitet.

Wir betrachten dabei sowohl die *europäischen Optionen* als auch die *amerikanischen Optionen*. Bei den europäischen Optionen muss der Kauf (bzw. Verkauf) zu einem von vornherein festgelegten *Ausübungszeitpunkt* T stattfinden. Ist dieser Ausübungszeitpunkt T erreicht, so kann der Inhaber die Option ausüben, also den Basiswert zu dem vorher festgelegten Ausübungspreis kaufen (bzw. verkaufen), oder er kann die Option verfallen lassen, den Basiswert also nicht kaufen (bzw. verkaufen).

Bei der amerikanischen Option kann die Ausübung zu jedem beliebigen Zeitpunkt $t \in [0, T]$ durchgeführt werden, wobei T ein vertraglich festgelegter Zeitpunkt ist. Der Startzeitpunkt der Option wird hier per Definition immer als 0 festgelegt, was durch einfache Verschiebung des Zeitachse o.B.d.A. möglich ist.

Im Folgenden verwenden wir die folgenden Bezeichnungen:

$S(t)$: Basiswert zur Zeit t mit $S(t) \geq 0$ für alle $t \in [0, T]$

$V(t, S)$: Preis der Option zur Zeit t abhängig vom aktuellen Basiswert S

$K > 0$: vereinbarter Ausübungspreis

$T > 0$: vereinbarte Laufzeit

Falls wir genauer spezifizieren möchten, um welche Option es sich handelt, verwenden wir die Bezeichnungen V_C und V_P für Put bzw. Call oder noch genauer V_{Ce} , V_{Pe} , V_{Ca} oder V_{Pa} für den europäischen Call, europäischen Put, amerikanischen Call und amerikanischen Put.

2.1 Grundidee der Optionsbewertung

Die Grundidee der Bewertung europäischer Optionen, die in allen im Verlauf dieser Vorlesung behandelten numerischen Methoden gleich ist, besteht aus den folgenden drei Schritten:

Algorithmus 2.1 (Grundalgorithmus der Optionsbewertung)

- (1) Bestimme eine Formel für den Wert $V(T, S)$ zum Laufzeitende (in Abhängigkeit vom Kurs $S = S(T)$ am Laufzeitende)
- (2) Bestimme ausgehend vom Basiswert $S(t)$ zur Zeit $t < T$ mit Hilfe eines stochastischen Kursmodells die Zufallsvariable $S(T) = S(T, \omega)$
- (3) Berechne den Optionswert als den abgezinsten Erwartungswert

$$V(t, S(t)) = e^{r(t-T)} E(V(T, S(T)))$$

□

Schritt (3) ist dabei erläuterungsbedürftig: Die Formel in (3) ist offenbar durch Satz 1.1 motiviert. Wäre $E(V(T, S(T)))$ der sichere Wert der Option zum Zeitpunkt T , so wäre Satz 1.1 anwendbar. Hier ist $E(V(T, S(T)))$ allerdings der Erwartungswert, d.h. der tatsächlich eintretende Wert kann höher oder niedriger sein. Wendet man die Abzinsungsformel aus Satz 1.1 wie in (3) auf einen unsicheren zukünftigen Wert an, so spricht man von *risikoneutraler Bewertung*. Dabei ignoriert man das vorhandene Risiko und behandelt den unsicheren zukünftigen Wert wie einen sicheren Wert: man verhält sich also neutral gegenüber dem Risiko, daher der Name. Im Gegensatz dazu stehen die *risikoaverse Bewertung*, in der man zum Ausgleich des Risikos einen kleineren Faktor als $e^{r(t-T)}$ verwendet, und die *risikofreudige Bewertung*, in der man einen größeren Faktor benutzt.

Die risikoneutrale Bewertung ist bei Optionen üblich, ihre Rechtfertigung erhält sie durch die Black-Scholes-Theorie, die wir in einigen Kapiteln behandeln werden. Tatsächlich hat diese Bewertungsmethode Auswirkungen auf das Kursmodell in (2), die wir in Abschnitt 2.3 betrachten.

Um von Algorithmus 2.1 zu tatsächlich implementierbaren numerischen Algorithmen zu kommen, werden wir im Folgenden die einzelnen Schritte genauer untersuchen: Schritt (1) werden wir im folgenden Abschnitt behandeln. Für Schritt (2) werden wir in den folgenden Kapiteln verschiedene stochastische Modelle kennen lernen. Die numerische Umsetzung von Schritt (3) hängt dann wesentlich von dem in (2) verwendeten Modell ab. Je nach Modell kann der Erwartungswert entweder direkt (vgl. Kapitel 3), durch stochastische Simulation (vgl. Kapitel 5), oder durch analytische Umformung in eine partielle Differentialgleichung mit anschließender Diskretisierung (vgl. Kapitel 7 und 9) numerisch berechnet werden.

Für amerikanische Optionen wird Schritt (3) komplizierter: Hier muss zusätzlich für jeden Zeitpunkt zwischen t und T geprüft werden, ob eine vorzeitige Ausübung vorteilhaft ist. Dies funktioniert nicht für jedes der in dieser Vorlesung vorgestellten Verfahren. Aber auch wenn es prinzipiell funktioniert, ist eine Überprüfung für *jeden* Zeitpunkt dabei numerisch

aus offensichtlichen Gründen nicht möglich. Tatsächlich wird diese Überprüfung nur per Rückwärtsiteration zu diskreten Zeitpunkten durchgeführt. Welche Zeitpunkte dafür verwendet werden und wie dies im Detail implementiert wird, hängt dabei von der konkreten numerischen Ausgestaltung von Schritt (3) ab.

2.2 Analytische Aussagen über Optionswerte

In diesem Abschnitt berechnen wir zunächst die in Schritt (1) der Bewertung benötigten Formeln für den Optionswert am Laufzeitende T . Danach leiten wir daraus auf analytischem Wege einige Folgerungen für die Optionswerte zu beliebigen Zeiten $t \in [0, T]$ her.

Der Wert der Option $V(T, S)$ am Laufzeitende T ist bei Kenntnis des Basiswertes $S = S(T)$ leicht zu berechnen: Sei K der festgelegte Ausübungspreis einer Call-Option. Falls $S > K$ ist, so erzielt man durch Ausübung der Option offenbar den Gewinn $V(T, S) = S - K > 0$. Falls $S \leq K$, so würde man durch Ausübung der Option den Basiswert zu einem höheren Preis als dem aktuellen Marktpreis kaufen, folglich würde man die Option nicht ausüben, der Wert ist also $V(T, S) = 0$. Zusammen erhält man so den Wert

$$V_C(T, S) = \max\{S - K, 0\} =: (S - K)^+. \quad (2.1)$$

Gerade umgekehrt verhält es sich bei einer Put-Option. Hier wird man die Option nur dann ausüben, wenn der Marktpreis S unterhalb des Ausübungspreises K liegt. Man erhält so

$$V_P(T, S) = \max\{K - S, 0\} =: (K - S)^+. \quad (2.2)$$

Beachte, dass es am Laufzeitende keinen Unterschied zwischen europäischen und amerikanischen Optionen gibt.

Zusammen mit den Annahmen an unser Marktmodell können wir aus diesen einfachen Formeln Aussagen über den Zusammenhang der einzelnen Optionen sowie Abschätzungen für ihren Wert ableiten. In den Beweisen der folgenden Sätze spielen *Portfolios* eine wichtige Rolle. Ein Portfolio

$$\pi = \alpha_1 P_1 + \alpha_2 P_2 + \dots + \alpha_n P_n$$

ist eine Summe verschiedener Finanzanlagen P_i mit Gewichtungen $\alpha_i \in \mathbb{R}$ (in den folgenden Beweisen sind die α_i stets gleich $+1$ oder -1 und konstant in der Zeit, können aber allgemein beliebig und auch zeitvariant gewählt werden). Wichtig ist dabei, dass die Gewichtungen auch negativ sein können. Eine negativ gewichtete Anlage P_i kann in der Praxis durchaus eine sinnvolle Interpretation haben: Ist P_i z.B. eine Option, so ist die Halterin der negativen Option gerade die Emittentin der Option, die sich verpflichtet, den Basiswert vom Halter der Option (falls von diesem gewünscht) zum festgesetzten Ausübungspreis zu kaufen (Put) oder zu verkaufen (Call). In den folgenden Beweisen dieses Abschnitts gehen tatsächlich nur Optionen negativ in die Portfolios ein, so dass wir immer diese Interpretation verwenden können.

Prinzipiell ist es aber in der Finanzmathematik möglich, dass jede beliebige Anlage P_i mit negativem Gewicht $\alpha_i < 0$ in ein Portfolio eingehen kann — man nennt dies einen *Leerverkauf*. In der Praxis entspricht dies dem Verkauf von $-\alpha_i$ Anteilen von P_i beim Kauf des Portfolios und den Kauf von $-\alpha_i$ Anteilen von P_i beim Verkauf des Portfolios, natürlich

zum jeweils gültigen Preis. Mit Leerverkäufen von Aktien lässt sich also Gewinn machen, wenn die Aktie fällt — ein Grund dafür, warum Leerverkäufe in der derzeitigen Finanzkrise in Verruf geraten sind und in einigen Ländern für manche Aktien sogar zeitweise verboten wurden. Voraussetzung für Leerverkäufe ist natürlich, dass sich auf dem Markt ein Partner für das entsprechende Geschäft findet. Einige Hinweise zur technischen Abwicklung solcher Leerverkäufe finden sich z.B. auf Wikipedia.

Wir betrachten nun zunächst europäische Optionen.

Satz 2.2 Für alle $t \in [0, T]$ gilt

$$S(t) + V_{Pe}(t, S(t)) - V_{Ce}(t, S(t)) = Ke^{r(t-T)}.$$

Diese Gleichung wird auch als *Put-Call-Parität* bezeichnet und erlaubt es, den Wert des europäischen Put aus dem europäischen Call (und umgekehrt) zu berechnen.

Beweis: Wir betrachten ein Portfolio π , das aus dem Basiswert, einem europäischen Put und einem negativen europäischen Call besteht. Der Wert dieses Portfolios zur Zeit t beträgt

$$\pi(t) = S(t) + V_{Pe}(t, S(t)) - V_{Ce}(t, S(t))$$

und zur Zeit T gerade

$$\pi(T) = S(T) + (K - S(T))^+ - (S(T) - K)^+ = K.$$

Wir erhalten also zur Zeit T die garantierte Auszahlung K , weswegen der Wert des Portfolios zur Zeit t gemäß Satz 1.1 gerade $e^{r(t-T)}K$ ist, womit die Behauptung folgt.

Satz 2.3 Für alle $t \in [0, T]$ gelten die Ungleichungen

- (i) $(S(t) - Ke^{r(t-T)})^+ \leq V_{Ce}(t, S(t)) \leq S(t)$
- (ii) $(Ke^{r(t-T)} - S(t))^+ \leq V_{Pe}(t, S(t)) \leq Ke^{r(t-T)}$

Beweis: (i) Wir beweisen die Ungleichungen einzeln:

“ $0 \leq V_{Ce}(t, S(t))$ ”: Da sicher die Ungleichung $V_{Ce}(T, S(T)) \geq 0$ gilt, folgt die Aussage aus Bemerkung 1.2 mit $P(t) = V_{Ce}(t, S(t))$ und $P_0 = 0$.

“ $S(t) - Ke^{r(t-T)} \leq V_{Ce}(t, S(t))$ ”: Der Wert des Portfolios

$$\pi(t) = S(t) - V_{Ce}(t, S(t))$$

erfüllt zum Zeitpunkt T sicher die Ungleichung

$$\pi(T) = S(T) - (S(T) - K)^+ \leq K.$$

Aus Bemerkung 1.2 mit $P(t) = \pi(t)$ und $P_0 = K$ folgt daher

$$\pi(t) \leq e^{r(t-T)}K,$$

woraus die Ungleichung sofort folgt.

“ $V_{Ce}(t, S(t)) \leq S(t)$ ”: Das Portfolio

$$\pi(t) = S(t) - V_{Ce}(t, S(t))$$

erfüllt zum Zeitpunkt T sicher die Ungleichung

$$\pi(T) = S(T) - (S(T) - K)^+ \geq 0.$$

Die Ungleichung folgt daher aus Bemerkung 1.2 mit $P(t) = \pi(t)$ und $P_0 = 0$.

(ii) Folgt aus (i) mit Satz 2.2. □

Als nächstes betrachten wir die Beziehung zwischen den europäischen und amerikanischen Optionen. Klar ist, dass die beiden Ungleichungen

$$V_{Ca}(t, S) \geq V_{Ce}(t, S) \quad \text{und} \quad V_{Pa}(t, S) \geq V_{Pe}(t, S) \quad (2.3)$$

gelten, denn da die amerikanischen Optionen genau wie die europäischen Optionen zum Zeitpunkt T ausgeübt werden können, kann mindestens der gleiche Gewinn erzielt werden.

Durch die mögliche vorzeitige Ausübung könnte allerdings eventuell ein höherer Gewinn erzielt werden, der dann zu einem strikt größeren Wert führen würde. Der folgende Satz zeigt, dass die mögliche frühere Ausübung für den Call allerdings keinen solchen Vorteil bringt.

Satz 2.4 Der durch die vorzeitige Ausübung der amerikanischen Call-Option zu einem Zeitpunkt $t < T$ erzielbare Ertrag $(S(t) - K)^+$ ist immer kleiner oder gleich dem Wert der Option, d.h.

$$(S(t) - K)^+ \leq V_{Ca}(t, S(t)), \quad (2.4)$$

wobei die Ungleichung strikt ist, falls $(S(t) - K)^+ > 0$ gilt. Insbesondere ist eine vorzeitige Ausübung der Option also niemals sinnvoll und es folgt

$$V_{Ce}(t, S(t)) = V_{Ca}(t, S(t)).$$

Beweis: Angenommen, wir üben die amerikanische Option vorzeitig zum Zeitpunkt $t < T$ aus. Dann erhalten wir den Betrag $(S(t) - K)^+$. Aus (2.3) und Satz 2.3(i) folgt dann die Ungleichung

$$V_{Ca}(t, S(t)) \geq V_{Ce}(t, S(t)) \geq (S(t) - Ke^{r(t-T)})^+ \geq (S(t) - K)^+,$$

weil ja $e^{r(t-T)} < 1$ gilt wegen $t < T$. Die letzte Ungleichung dieser Kette ist dabei strikt, wenn $(S(t) - K)^+ > 0$ ist. Damit folgt (2.4).

Damit ist es also stets optimal, die Option entweder zum Zeitpunkt $t = T$ oder gar nicht auszuüben. Da dies genau die Möglichkeiten der europäischen Option sind, müssen die Werte übereinstimmen. □

Wir beenden den Abschnitt mit Ungleichungen für amerikanische Optionen, die analog zu Satz 2.2 und Satz 2.3(ii) sind.

Satz 2.5 Für alle $t \in [0, T]$ gelten die Ungleichungen

- (i) $Ke^{r(t-T)} \leq S(t) + V_{Pa}(t, S(t)) - V_{Ca}(t, S(t)) \leq K$
- (ii) $(Ke^{r(t-T)} - S(t))^+ \leq V_{Pa}(t, S(t)) \leq K$

Beweis: (i) Wir zeigen die beiden Ungleichungen separat.

“ $Ke^{r(t-T)} \leq S(t) + V_{Pa}(t, S(t)) - V_{Ca}(t, S(t))$ ”: Durch Anwendung von Satz 2.4, Ungleichung (2.3) und Satz 2.2 folgt

$$\begin{aligned} S(t) + V_{Pa}(t, S(t)) - V_{Ca}(t, S(t)) &= S(t) + V_{Pa}(t, S(t)) - V_{Ce}(t, S(t)) \\ &\geq S(t) + V_{Pe}(t, S(t)) - V_{Ce}(t, S(t)) \\ &= Ke^{r(t-T)}, \end{aligned}$$

also die gewünschte Ungleichung.

“ $S(t) + V_{Pa}(t, S(t)) - V_{Ca}(t, S(t)) \leq K$ ”:

Betrachte das Portfolio

$$\pi(t) = S(t) + V_{Pa}(t, S(t)) - V_{Ca}(t, S(t)).$$

Für $V_{Ca}(t, S(t))$ gilt wegen (2.4) und Satz 2.3(i) die Ungleichung $V_{Ca}(t, S(t)) \geq (S(t) - Ke^{r(t-T)})^+$ und damit

$$\pi(t) \leq S(t) + V_{Pa}(t, S(t)) - (S(t) - Ke^{r(t-T)})^+.$$

Bezeichne nun mit $t^* \in [t, T]$ den Ausübungszeitpunkt des amerikanischen Put.

Im Fall $K \geq S(t^*)$ erhalten wir dann

$$\pi(t^*) \leq S(t^*) + K - S(t^*) - (S(t^*) - Ke^{r(t^*-T)})^+ \leq K$$

und für $K \leq S(t^*)$ ergibt sich

$$\pi(t^*) \leq S(t^*) + 0 - (S(t^*) - Ke^{r(t^*-T)})^+ = Ke^{r(t^*-T)}.$$

Wegen $Ke^{r(t^*-T)} \leq K$ erhalten wir also die sichere obere Schranke

$$\pi(t^*) \leq K.$$

Damit folgt aus Bemerkung 1.2

$$\pi(t) \leq e^{r(t-t^*)} K \leq K$$

und damit die Behauptung.

(ii) Aus (i) folgt

$$Ke^{r(t-T)} - S(t) + V_{Ca}(t, S(t)) \leq V_{Pa}(t, S(t)) \leq K - S(t) + V_{Ca}(t, S(t)).$$

Damit und mit Satz 2.4 und Satz 2.3(i) folgt

$$\begin{aligned} V_{Pa}(t, S(t)) &\geq Ke^{r(t-T)} - S(t) + V_{Ca}(t, S(t)) \\ &= Ke^{r(t-T)} - S(t) + V_{Ce}(t, S(t)) \\ &\geq Ke^{r(t-T)} - S(t) + (S(t) - Ke^{r(t-T)})^+ = (Ke^{r(t-T)} - S(t))^+, \end{aligned}$$

also die erste Ungleichung in (ii). Mit den gleichen Sätzen folgt zudem

$$\begin{aligned} V_{Pa}(t, S(t)) &\leq V_{Pa}(t, S(t)) \leq K - S(t) + V_{Ca}(t, S(t)) \\ &= K - S(t) + V_{Ce}(t, S(t)) \leq K - S(t) + S(t) = K, \end{aligned}$$

also die zweite Ungleichung in (ii). \square

2.3 Folgerungen aus der risikoneutralen Bewertung

Wir wollen nun mit Hilfe der gerade bewiesenen Sätze eine Folgerung aus der risikoneutralen Bewertung für das Kursmodell $S(t)$ ableiten. Beachte, dass wir die risikoneutrale Bewertung im vorhergehenden Abschnitt gar nicht verwendet haben, denn alle Aussagen wurden aus sicheren Portfoliowerten zur Endzeit ermittelt. Die Sätze gelten also unabhängig von der Art der Bewertung.

Legen wir aber eine risikoneutrale Bewertung zu Grunde, so erhalten wir aus der ersten Ungleichung aus Satz (2.3)(i) und Algorithmus 2.1(3)

$$(S(t) - Ke^{r(t-T)})^+ \leq V_{Ce}(t, S(t)) = e^{r(t-T)} E(V_{Ce}(T, S(T))) = e^{r(t-T)} E((S(T) - K)^+).$$

Beachte dabei, dass beide Seiten der Ungleichung stetig in K sind und für alle $K > 0$ gelten sollen. Damit müssen die Ungleichungen also auch für $K = 0$ gelten, woraus wegen $S(t) \geq 0$ folgt

$$S(t) \leq e^{r(t-T)} E(S(T)).$$

Mit der zweiten Ungleichung aus Satz (2.3)(i) erhalten wir

$$e^{r(t-T)} E((S(T) - K)^+) = V_{Ce}(t, S(t)) \leq S(t)$$

und damit für $K = 0$

$$e^{r(t-T)} E(S(T)) \leq S(t).$$

Zusammen ergibt sich also

$$e^{r(t-T)} E(S(T)) = S(t) \quad \text{oder äquivalent} \quad E(S(T)) = e^{r(T-t)} S(t). \quad (2.5)$$

Die risikoneutrale Bewertung der Option verlangt also, dass auch auf $S(t)$ das Prinzip der risikoneutralen Bewertung angewendet werden muss. Oder, äquivalent formuliert, der Erwartungswert $E(S(T))$ des Basiswertes muss sich wie eine risikofreie Anleihe verhalten, d.h. mit dem gleichen Zinssatz wachsen.

Dies erscheint zunächst sehr einschränkend, denn man möchte ja eigentlich in der Lage sein, Basiswerte mit unterschiedlichen durchschnittlichen Wachstumsraten modellieren zu können. Auch hierfür wird uns die Black-Scholes Theorie später aber eine sinnvolle Begründung liefern.

Kapitel 3

Die Binomialmethode

In diesem Abschnitt beschreiben wir eine erste Methode zur Berechnung des fairen Optionspreises. Der Basiswert $S(t)$ wird dabei meist als Aktienpaket interpretiert und wir spezifizieren nun zunächst das stochastische Modell, mit dem wir aus dem derzeitigen Kurs $S(t)$ die möglichen zukünftigen Kurse $S(t')$, $t' > t$, berechnen können.

3.1 Zeitdiskrete stochastische dynamische Systeme

Das hier verwendete Modell ist zeitdiskret mit Zeitachse \mathbb{Z} , d.h. wir berechnen die möglichen Kurswerte für die zukünftigen Zeiten $S(t+1)$, $S(t+2)$, \dots , $S(T)$. Formal abstrakt ist das Modell dabei von der folgenden Form.

Definition 3.1 Ein zeitdiskretes stochastisches System ist gegeben durch eine Abbildung

$$f : \mathbb{R}^n \times \mathbb{R}^m \rightarrow \mathbb{R}^n, (S, Z) \mapsto f(S, Z).$$

Hierbei ist $S \in \mathbb{R}^n$ der Zustand und $Z \in \mathbb{R}^m$ der stochastische Einfluss. Für einen Anfangswert $S_0 \in \mathbb{R}^n$ und eine Folge von Zufallsvariablen Z_0, Z_1, \dots definieren wir die *Lösung* $S(t) = S(t, S_0)$ des Systems für $t \geq 0$ induktiv mittels

$$S(0) = S_0, \quad S(t+1) = f(S(t), Z_t). \tag{3.1}$$

□

Bemerkung 3.2 (i) Wir verwenden im Folgenden oft die Kurzschreibweise $S(t)$ und führen das zusätzliche Argument S_0 nur dann explizit auf, wenn ansonsten Uneindeutigkeiten entstünden.

(ii) Die Lösung $S(t)$ ist zu jeder Zeit $t > t_0$ eine Zufallsvariable, da sie von den Zufallsvariablen Z_{t_0}, \dots, Z_{t-1} abhängt. Für jeden Anfangswert S_0 ist $S(t, S_0)$ ein sogenannter *stochastischer Prozess*, eine zeitabhängige Zufallsvariable.

(iii) Wenn wir Realisierungen $Z_t(\omega)$ der Zufallsvariablen Z_t betrachten, erhalten wir eine Realisierung $S(t, \omega) = S(t, S_0, \omega)$. Diese Realisierung ist nun eine "normale" Funktion $t \mapsto S(t, S_0, \omega)$ von \mathbb{N}_0 nach \mathbb{R}^n und wird *Pfad* von $S(t)$ genannt.

(iv) Falls Z_0, \dots, Z_{t-1} und Z_t unabhängige Zufallsvariablen sind, sind auch $S(t)$ und Z_t unabhängig, da $S(t)$ nur von Z_0, \dots, Z_{t-1} abhängt.

(v) Die genaue Skalierung der Zeiteinheit “ t ” ist variabel. Die Zeit $t = 1$ könnte ein Jahr, einen Monat oder einen Tag bedeuten. Wichtig ist aber, dass die Parameter des Modells an die Zeitskala angepasst werden. Für den jährlichen Zinssatz r bedeutet dies z.B., dass sich die Zinsen im Intervall $[0, T]$ bei einer Zeitskala in Jahren zu e^{rT} , bei einer Zeitskala in Tagen hingegen zu $e^{\frac{r}{360}T}$ ergeben (wenn das Jahr wie üblich in 12 Monate mit je 30 Tagen eingeteilt wird). \square

3.2 Das Binomialmodell eines Aktienkurses

Das Modell, das der Binomialmethode seinen Namen gibt, ist das sogenannte *Binomialmodell* eines Aktienkurses. Ausgehend von dem Anfangskurs S_0 kann der Kurs dabei in jedem Zeitschritt entweder mit Wahrscheinlichkeit p um einen Faktor $\alpha_u > 1$ steigen (“up”) oder mit Wahrscheinlichkeit $1 - p$ um einen Faktor $\alpha_d < 1$ fallen (“down”). Hierbei soll die Bewegung im Zeitschritt $t + 1$ jeweils stochastisch unabhängig von der Bewegung im vorhergehenden Zeitschritt t sein.

Zur Modellierung als zeitdiskretes stochastisches System betrachten wir daher unabhängige Zufallsvariablen $Z_t : \Omega \rightarrow \{\alpha_u, \alpha_d\}$ mit

$$\mathbb{P}_{Z_t}(\{\alpha_u\}) = p \quad \text{und} \quad \mathbb{P}_{Z_t}(\{\alpha_d\}) = 1 - p$$

und definieren

$$f(S, Z) = S \cdot Z.$$

Die Lösung ergibt sich damit für $t \geq 1$ zu

$$S(t) = Z_{t-1}Z_{t-2} \dots Z_0 S_0.$$

Abbildung 3.1 zeigt einige Beispielpfade dieses Modells für $\alpha_u = 1.03$, $\alpha_d = 0.98$, $p = 1/2$ und $S_0 = 1$.

Das Binomialmodell hat eine wesentliche Eigenschaft, die es für die numerische Optionsbewertung günstig macht: Die Menge der möglichen zukünftigen Zustände $S(t)$ nimmt für wachsendes t relativ langsam zu:

Auf den ersten Blick sieht es so aus, als wenn sich die Anzahl der möglichen Zustände in jedem Schritt verdoppelt. Beginnen wir in $t = 0$ mit S_0 , so erhalten wir in $t = 1$ die zwei möglichen Zustände $\{\alpha_u S_0, \alpha_d S_0\}$, in $t = 2$ die vier Zustände $\{\alpha_u^2 S_0, \alpha_d \alpha_u S_0, \alpha_u \alpha_d S_0, \alpha_d^2 S_0\}$ usw. Da nun aber offensichtlich $\alpha_d \alpha_u = \alpha_u \alpha_d$ gilt, fallen hier zwei der Zustände zusammen. Ähnliches passiert in allen nachfolgenden Zeitschritten — man sagt, das Modell ist *rekombinierbar*. Daher lässt sich per Induktion leicht beweisen, dass es zum Zeitpunkt t gerade $t + 1$ verschiedene Zustände gibt. Abbildung 3.2 illustriert die entstehende Baumstruktur (mit den Wahrscheinlichkeiten jedes Teilpfades), die auch *Binomialbaum* genannt wird.

Die Bedeutung der Parameter p , α_u und α_d betrachten wir im übernächsten Abschnitt noch etwas genauer. Zunächst wollen wir aber erläutern, wie man mit Hilfe dieses Binomialmodells den Wert einer Option berechnet.

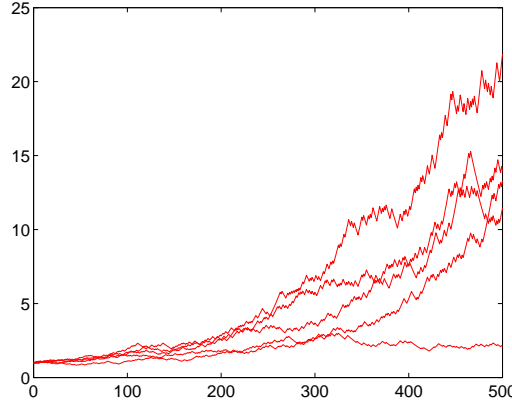


Abbildung 3.1: Beispielpfade des Binomialmodells

3.3 Optionswertberechnung mit dem Binomialmodell

Mit Hilfe des Binomialmodells können wir nun den Schritt (3) von Algorithmus 2.1 numerisch umsetzen. Wir beginnen mit europäischen Optionen und stellen dazu eine Methode vor, die iterativ vorgeht. Diese lässt sich danach leicht auf amerikanische Optionen verallgemeinern.

Wir bezeichnen die möglichen Zustände des Binomialmodells mit $S_k(t)$, $k = 0, \dots, t$, wobei wir aufsteigend nummerieren. Dann gilt für die $S(t)$ -Werte aus unserem Modell gerade $S(0) = S_0 = S_0(0)$ und

$$\begin{aligned}\mathbb{P}_{S(t+1)}\left(S_{k+1}(t+1) \mid S(t) = S_k(t)\right) &= p, \\ \mathbb{P}_{S(t+1)}\left(S_k(t+1) \mid S(t) = S_k(t)\right) &= 1 - p.\end{aligned}$$

Damit erhalten wir das folgende Lemma.

Lemma 3.3 Wenn der Basiswert $S(t)$ einer europäischen Option sich gemäß dem Binomialmodell verhält, so gilt für alle $t \in [0, T - 1]$ für den Wert $V(t, S)$ bei risikoneutraler Bewertung die Gleichung

$$V(t, S_k(t)) = e^{-r} \left(pV(t+1, S_{k+1}(t+1)) + (1-p)V(t+1, S_k(t+1)) \right). \quad (3.2)$$

Beweis: Bei risikoneutraler Bewertung gilt

$$V(t, S^1(t)) = e^{r(t-T)} E\left(V(T, S(T)) \mid S(t) = S^1(t)\right)$$

und

$$V(t+1, S^2(t+1)) = e^{r(t+1-T)} E\left(V(T, S(T)) \mid S(t+1) = S^2(t+1)\right),$$

setze $V_i := (K - S_i)^+$ für einen Put bzw. $V_i := (S_i - K)^+$ für einen Call
 Ende der Schleife

- (2) Für t von $T - 1$ bis 0
 Für i von 0 bis t
 setze $V_i := e^{-r}(pV_{i+1} + (1 - p)V_i)$
 Ende der Schleifen

Ausgabe: $V_0 = V(0, S_0)$ □

Dieser Algorithmus lässt sich nun leicht auf amerikanische Put-Optionen verallgemeinern. Hierzu maximieren wir in jedem Zeitschritt über den bei der vorzeitigen Ausübung erzielbaren und den durch Halten der Option für (mindestens) einen weiteren Zeitschritt erzielbaren Wert. Beachte, dass wir hier nicht kontinuierlich prüfen, ob die vorzeitige Ausübung sinnvoll ist sondern nur zu den durch das Binomialmodell vorgegebenen diskreten Zeitpunkten.

Algorithmus 3.5 (Iterative Bewertung einer amerikanischen Put-Option mit der Binomialmethode)

Eingabe: $S_0 \in \mathbb{R}^+$, $K \in \mathbb{R}^+$, $T \in \mathbb{N}$, $p \in (0, 1)$, $\alpha_u > \alpha_d \in \mathbb{R}_+$

- (1) Für i von 0 bis T
 setze $S_i = S_0 \alpha_u^i \alpha_d^{T-i}$
 setze $V_i := (K - S_i)^+$
 Ende der Schleife
- (2) Für t von $T - 1$ bis 0
 Für i von 0 bis t
 setze $S_i := S_0 \alpha_u^i \alpha_d^{t-i}$
 setze $V_i := \max\{e^{-r}(pV_{i+1} + (1 - p)V_i), K - S_i\}$
 Ende der Schleifen

Ausgabe: $V_0 = V_{Pa}(0, S_0)$ □

3.4 Die Parameter im Binomialmodell

Um eine reale Kursentwicklung brauchbar stochastisch zu simulieren, müssen die Parameter p , α_u und α_d geeignet angepasst werden. Das absolute Minimum an stochastischen Werten, das ein Aktienkursmodell widerspiegeln sollte, sind der Mittelwert des Kurses und seine Varianz, im Zusammenhang mit Kursen oft Volatilität genannt. Da wir es bei den Aktienkursmodellen mit stochastischen Prozessen zu tun haben, sind diese Werte hier nicht wie bei einer einfachen Zufallsvariablen statisch sondern abhängig von der Zeit und vom Anfangswert, d.h. wir erhalten

$$E(S(t, S_0)) \quad \text{und} \quad Var(S(t, S_0))$$

abhängig von t und S_0 . Aus Formel (2.5) folgt dabei, dass für unsere risikoneutrale Bewertung die Gleichung

$$E(S(t+1)) = e^r S(t) \quad (3.3)$$

gelten muss.

Im Binomialmodell gibt es nun gerade die drei Parameter p , α_u und α_d , die zunächst so angepasst werden, dass diese Bedingung erfüllt ist. Wegen

$$E(S(t+1)) = (p\alpha_u + (1-p)\alpha_d)S(t)$$

folgt aus (3.3) die Gleichung

$$p\alpha_u + (1-p)\alpha_d = e^r.$$

Dadurch ist einer der drei Parameter bereits eindeutig bestimmt, wenn die beiden anderen Parameter festgelegt sind. Die beiden anderen Parameter können nun auf unterschiedliche Weisen berechnet werden. Zum einen können sie so gewählt werden, dass die Varianz (und gegebenenfalls weitere stochastische Kenngrößen) des Modells mit den vergangenen Kursdaten gut übereinstimmen — man nennt dies *historische Parameter*. Die Hoffnung ist natürlich, dass die aus der Vergangenheit ermittelten Werte einen halbwegs zuverlässigen stochastischen Blick in die Zukunft erlauben. Eine andere Methode der Parameterbestimmung besteht darin, die Parameter so zu bestimmen, dass die ermittelten fairen Preise mit den beobachteten Marktpreisen für die Optionen übereinstimmen — man spricht von *impliziten Parametern*. Diese Parameterbestimmung ist ein eigenes, theoretisch und numerisch anspruchsvolles Problem, das wir hier nicht behandeln können (das aber ein schönes Thema für eine Diplomarbeit wäre).

Oft wird das Binomialmodell als Näherung der geometrischen Brownschen Bewegung angesehen, die dem Black-Scholes-Modell zu Grunde liegt. Wie die Parameter in diesem Fall gewählt werden müssen, werden wir in Abschnitt 4.2.2 genauer besprechen.

Kapitel 4

Stochastische Differentialgleichungen

Um eine stochastische Differentialgleichung zu formulieren, müssen wir zunächst erklären, auf welche Weise der “Zufall” in unsere Differentialgleichung eingeht.

Vereinfacht kann man eine stochastische Differentialgleichung zunächst als eine “übliche” gewöhnliche Differentialgleichung mit der speziellen Struktur

$$\dot{x}(t) = f(t, x(t)) := a(t, x(t)) + b(t, x(t))g(t) \quad (4.1)$$

schreiben¹. Hierbei sind $a, b : \mathbb{R} \times \mathbb{R}^n \rightarrow \mathbb{R}^n$ gewöhnliche Funktionen, der Zufall kommt durch die Funktion g ins Spiel, die nämlich keine “feste” Funktion ist, sondern eine Funktion, die abhängig vom Zufall verschiedene Gestalt annehmen kann. Für jedes mögliche g erhalten wir für einen festen Anfangswert weiterhin eine eindeutige Lösung. Da diese aber nun von g abhängt und g zufällig ist, ist in diesem Sinne auch die Lösungstrajektorie zufällig.

Natürlich ist g nicht “irgendwie” vom Zufall abhängig, sondern genügt gewissen statistischen Regeln. Um diese zu definieren, werden wir zuerst einen grundlegenden stochastischen Prozess einführen. Dabei spielt die Gauß- oder Normal-Verteilung eine wichtige Rolle. Wir erinnern hier daran, dass eine Zufallsvariable X normalverteilt mit Erwartungswert μ und Varianz σ^2 ist (Notation: $X \sim N(\mu, \sigma^2)$), falls sie die Dichtefunktion

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

besitzt.

4.1 Der Wiener Prozess

Zur Definition der stochastischen Differentialgleichung benötigen wir nun einen Begriff für zufällige Funktionen. Eine “zufällige Funktion” ist dabei nichts anderes als eine Funktion $X : \mathbb{R} \times \Omega \rightarrow \mathbb{R}$, die für jedes feste t gerade eine Zufallsvariable $X(t, \cdot) : \Omega \rightarrow \mathbb{R}$ ist,

¹Dies ist eine vorläufige Schreibweise, siehe (4.5) für die mathematisch gebräuchliche Formulierung.

wobei für alle t der gleiche Wahrscheinlichkeitsraum $(\Omega, \mathcal{A}, \mathbb{P})$ zu Grunde liegt. Solch eine Funktion nennt man *stochastischen Prozess*. Beachte, dass wir keinerlei Annahmen über den Wahrscheinlichkeitsraum $(\Omega, \mathcal{A}, \mathbb{P})$ voraussetzen; wichtig ist nur, dass alle Zufallsvariablen X auf dem selben Raum definiert sind. Für jedes feste $\omega \in \Omega$ ist die Realisierung $X(\cdot, \omega) : \mathbb{R} \rightarrow \mathbb{R}$ eine “normale” reelle Funktion. Wie bereits im vorhergehenden Kapitel erwähnt, nennt man die Realisierung bei stochastischen Prozessen *Pfad*.

Der stochastische Prozess, den wir hier näher betrachten wollen und aus dem das obige g später hervorgehen wird, ist der sogenannte *Wiener Prozess*, der üblicherweise mit W bezeichnet wird und für $t \geq 0$ definiert ist. Die Definition des Wiener Prozesses ergibt sich aus der von N. Wiener² eingeführten mathematischen Beschreibung der *Brownschen Bewegung*, die in der Physik die zufällige Bewegung eines auf einer Wasseroberfläche schwimmenden Teilchens beschreibt. Formal verlangt man die folgenden Bedingungen:

- (i) $W(t)$ ist eine Gauß-verteilte Zufallsvariable mit $E(W(t)) = 0$ und $\text{Var}(W(t)) = t$
- (ii) Für $t_1 \geq t_0 \geq 0$ sind die *Inkmente* $W(t_1) - W(t_0)$ Gauß-verteilte Zufallsvariablen mit $E(W(t_1) - W(t_0)) = 0$ und $\text{Var}(W(t_1) - W(t_0)) = t_1 - t_0$
- (iii) Für $s_1 \geq s_0 \geq t_1 \geq t_0 \geq 0$ sind die Inkmente $W(t_1) - W(t_0)$ und $W(s_1) - W(s_0)$ unabhängige Zufallsvariablen.

Beachte, dass wir oben die Gauß-Verteilung nur für $\sigma > 0$ definiert haben. Hier erhalten wir für $t = 0$ die Bedingung $\text{Var}(W(0)) = 0$, womit einfach $\mathbb{P}_{W(0)}(\mu) = 1$ gemeint ist. Mit anderen Worten nimmt die Zufallsvariable $W(0)$ also fast sicher ihren Erwartungswert μ (hier $\mu = 0$) an.

Ein Pfad $W(t, \omega)$ von W ist also gerade eine der vielen möglichen zufälligen Funktionen, die (in ihrer Gesamtheit) den obigen Bedingungen genügen. Für diese Pfade kann man beweisen, dass sie fast sicher stetig in t sind (zur Erinnerung: dies bedeutet, dass das Ereignis $A = \{\omega \in \Omega \mid X(t, \omega) \text{ ist stetig in } t\}$ die Wahrscheinlichkeit $\mathbb{P}(A) = 1$ besitzt) aber auch fast sicher nirgends differenzierbar. Abbildung 4.1 zeigt drei verschiedene Pfade des Wiener Prozesses.

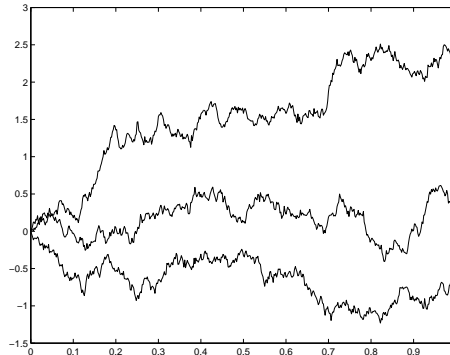


Abbildung 4.1: Verschiedene Pfade des Wiener Prozesses

²US-amerikanischer Mathematiker, 1894–1964

Die Bedingung (i) besagt gerade, dass die Streuung der Werte der Pfade $W(t, \omega)$ um so größer wird, je größer t wird. Der Mittelwert ist dabei zu jedem Zeitpunkt gleich 0. Die Bedingung (ii) bedeutet, dass der stochastische Prozess $\widetilde{W}(t) = W(t + t_0) - W(t_0)$ wieder ein Wiener Prozess ist. Wenn wir also ein “Endstück” eines Pfades abschneiden und dessen Anfangspunkt in den Nullpunkt verschieben, so erhalten wir wieder einen Pfad des Wiener Prozesses. Diese zwei Eigenschaften kann man anhand der Grafik gut nachvollziehen. Bedingung (iii) schließlich besagt, dass man aus der Kenntnis eines Teilstücks eines Pfades auf einem Teilintervall $[t_0, t_1]$ keinerlei Vorhersagen über den Verlauf auf einem anderen Teilintervall $[s_0, s_1]$ mit $s_0 \geq t_1$ machen kann, wenn man nur die Informationen über die Anfangswerte durch Abziehen der Werte $W(s_0)$ und $W(t_0)$ “löscht”. Der Pfad kann sich also zu jedem Zeitpunkt mit exakt der gleichen Wahrscheinlichkeit nach oben oder unten bewegen, egal welchen Verlauf er bis zu diesem Zeitpunkt genommen hat.

4.2 Das Itô-Integral

Betrachtet man die Pfade in Abbildung 4.1, so kann man durchaus eine vage Ähnlichkeit mit den Kursdiagrammen z.B. von Aktienkursen erkennen. Selbst für eine sehr grobe Modellierung von solchen Kurswerten ist der Wiener Prozess aber doch noch zu einfach, denn es gibt keine Parameter, an denen man “drehen” könnte, um den modellierten Verlauf einem echten Kursverlauf anzupassen: Der Wiener Prozess besitzt noch zu wenig Struktur.

Er ist aber bestens geeignet, um als “Zutat” in der Definition stochastischer Differentialgleichungen eine wichtige Rolle zu spielen; vereinfachend gesagt soll die zufällige Funktion g in (4.1) gerade die Ableitung des Wiener Prozesses sein. Um die dabei auftretenden mathematischen Probleme und ihre Lösung zu erläutern, wollen wir hier zunächst die denkbar einfachste stochastische Differentialgleichung betrachten und dann zur allgemeinen Formulierung übergehen.

Zunächst einige Anmerkungen zur Notation: Da die Lösung einer SDG über den eingehenden Wiener Prozess wieder eine zufällige Funktion — also ein stochastischer Prozess — ist, verwenden wir hier für den gesuchten unbekannt Prozess die groß geschriebene Bezeichnung $X(t)$, bzw. mit Anfangswert $X_0 \in \mathbb{R}^n$ und Anfangszeit $t_0 \in \mathbb{R}$ die Schreibweise $X(t; t_0, X_0)$. Wir erlauben hierbei, dass X vektorwertig, also aus dem \mathbb{R}^n ist, was einfach bedeutet, dass $X = (X_1, X_2, \dots, X_n)^T$ ist, wobei die X_i reellwertige stochastische Prozesse im oben eingeführten Sinne sind. Zu jedem Pfad $W(t, \omega)$ des eingehenden Wiener Prozesses gehört dann ein Lösungspfad des X -Prozesses, den wir mit $X(t; t_0, X_0, \omega)$ bezeichnen.

Die technische Hauptschwierigkeit in der mathematischen Formulierung stochastischer Differentialgleichungen zeigt sich nun bereits bei einer scheinbar trivialen Aufgabe, nämlich dem Problem, eine stochastische Differentialgleichung aufzustellen, deren Lösung gerade der Wiener Prozess ist. Scheinbar trivial ist die Aufgabe deswegen, weil wir ja den Wiener Prozess als gegeben voraussetzen und in der Formulierung verwenden dürfen, weswegen es nahe liegt, einfach die Differentialgleichung

$$\frac{d}{dt}X(t) = \frac{d}{dt}W(t) \quad (4.2)$$

mit Anfangsbedingung $X_0 = W(0)$ zur Anfangszeit $t_0 = 0$ zu verwenden. Das Problem ist jetzt aber: Was verstehen wir unter “ $\frac{d}{dt}W(t)$ ”? Man würde vielleicht versuchen, die

Ableitung pfadweise aufzufassen, d.h., wir berechnen die Ableitung für jeden Pfad $W(t, \omega)$. Nur ist ein typischer Pfad $W(t, \omega)$, wie oben erwähnt, nirgends differenzierbar.

Wie bei einer deterministischen gewöhnlichen Differentialgleichung, können wir nun Gleichung (4.2) zunächst in Integralform

$$X(t) = X_0 + \int_0^t \frac{d}{d\tau} W(\tau) d\tau$$

schreiben. Jetzt könnten wir formal integrieren, was uns aber bei der Frage “was ist $\frac{d}{dt} W(t)$?” nicht weiter bringt. Für das obige sogenannte *stochastische Integral* hat sich in der Literatur die kürzere Schreibweise

$$\int_0^t dW_\tau$$

eingebürgert, die wir hier übernehmen wollen. Dies zeigt die Richtung auf, die wir zur Lösung unseres Problems einschlagen wollen: Anstatt die Ableitung $\frac{d}{dt} W(t)$ zu betrachten, werden wir versuchen, diesem stochastischen Integral eine mathematische Definition zu geben, die

- (i) wohldefiniert ist, obwohl $\frac{d}{dt} W(t)$ nicht existiert
- (ii) das gewünschte Ergebnis, nämlich $X(t) = W(t)$, liefert
- (iii) sich auf allgemeinere Integrale der Form

$$I(F) := \int_{t_0}^{t_1} F(t) dW_t \tag{4.3}$$

verallgemeinern lässt, damit wir auch kompliziertere SDGs formulieren können. Hierbei ist F wiederum ein stochastischer Prozess, der auf dem selben Wahrscheinlichkeitsraum wie W definiert ist.

Wir wollen dieses Konzept nun für Integrale der Form (4.3) angeben. Die hier vorgestellte Lösung geht auf Kiyosi Itô³ zurück und wurde in den 1940er Jahren entwickelt. Die Idee besteht darin, das Integral (4.3) für jedes Paar von Pfaden $F(t, \omega)$ und $W(t, \omega)$ durch den Limes einer geeigneten Summe zu approximieren. Wir wählen dazu ein $N \in \mathbb{N}$ und eine Folge von Zeiten $\tau_i^{(N)}$, $i = 0, 1, \dots, N$ mit

$$t_0 = \tau_0^{(N)} < \tau_1^{(N)} < \dots < \tau_N^{(N)} = t_1$$

und definieren für jedes $\omega \in \Omega$

$$I^{(N)}(F)(\omega) := \sum_{i=0}^{N-1} F(\tau_i^{(N)}, \omega) (W(\tau_{i+1}^{(N)}, \omega) - W(\tau_i^{(N)}, \omega)).$$

³japanischer Mathematiker, 1915–2008

Das Integral (4.3) wird nun über den Limes dieser Summe definiert. Betrachte eine Familie von Folgen $\tau_i^{(N)}$ für $N \in \mathbb{N}$ mit $\lim_{N \rightarrow \infty} \delta(N) = 0$, wobei $\delta(N) := \max_{i=1, \dots, N} \tau_i^{(N)} - \tau_{i-1}^{(N)}$ ist. Dann definieren wir

$$I(F) := \lim_{N \rightarrow \infty} I^{(N)}(F). \quad (4.4)$$

Diese Definition wirft zunächst eine Reihe von Fragen auf, denn da die Pfade des Wiener Prozesses sehr unangenehme Funktionen sein können, ist nicht garantiert, dass dieser Limes für jedes ω überhaupt existiert. Tatsächlich lag der Haupttrick von Itô darin, zu definieren, was der Limes in (4.4) eigentlich bedeuten soll. Dieser Limes ist nämlich *nicht* pfadweise zu verstehen (in dem Sinne, dass wir $\lim_{N \rightarrow \infty} I^{(N)}(F)(\omega)$ für jedes $\omega \in \Omega$ bilden), sondern man muss die Werte $I^{(N)}(F)$ ebenso wie das Integral $I(F)$ wieder als Zufallsvariablen $I^{(N)}(F) : \Omega \rightarrow \mathbb{R}$ bzw. $I(F) : \Omega \rightarrow \mathbb{R}$ auffassen. Für Zufallsvariablen gibt es verschiedene Konvergenzbegriffe und der hier geeignete ist der Begriff der *Quadrat-Mittel-Konvergenz*, der wie folgt definiert ist: Eine Folge von Zufallsvariablen $X_N : \Omega \rightarrow \mathbb{R}$ konvergiert im Quadrat-Mittel-Sinne gegen eine Zufallsvariable $X : \Omega \rightarrow \mathbb{R}$, falls

$$\lim_{N \rightarrow \infty} E(|X_N - X|^2) = 0$$

gilt. Mit diesem Konvergenzbegriff kann man zeigen, dass die Folge $I^{(N)}(F)$ (unter geeigneten Bedingungen an F) tatsächlich konvergiert und (4.4) also wohldefiniert ist. Das resultierende Integral wird *Itô-Integral* genannt und es besitzt tatsächlich die oben aufgeführten gewünschten Eigenschaften (i)–(iii).

Mit Hilfe des Itô-Integrals können wir die informelle Schreibweise (4.1) mathematisch präzise formulieren. Statt der üblichen Differentialgleichungsschreibweise schreibt man *Itô-stochastische Differentialgleichungen* nämlich als

$$dX(t) = a(t, X(t))dt + b(t, X(t))dW_t. \quad (4.5)$$

Dies ist nur eine symbolische Schreibweise; was mit (4.5) tatsächlich gemeint ist, ist die längere Integralschreibweise

$$X(t) = X(t_0) + \int_{t_0}^t a(\tau, X(\tau))d\tau + \int_{t_0}^t b(\tau, X(\tau))dW_\tau,$$

bei der das zweite Integral gerade das Itô-Integral ist. Dies liefert eine mathematisch fundierte und brauchbare Definition stochastischer Differentialgleichungen. Falls $b(t, x) \equiv 0$ ist, also kein stochastischer Anteil vorhanden ist, reduziert sich (4.5) auf

$$X(t) = X(t_0) + \int_{t_0}^t a(\tau, X(\tau))dt \iff \frac{d}{dt}X(t) = a(t, X(t)),$$

also auf die wohlbekanntere deterministische gewöhnliche Differentialgleichung. Der deterministische Anteil der Gleichung $a(t, X(t))$ wird oft *Drift* genannt, der stochastische Anteil $b(t, X(t))$ oft als *Diffusion* bezeichnet.

Natürlich lässt sich (4.5) in vielfacher Hinsicht erweitern, z.B. kann man statt nur einem W mehrere unabhängige Wiener Prozesse W^1, \dots, W^m eingehen lassen, was zur Gleichung

$$dX(t) = a(t, X(t))dt + \sum_{j=1}^m b_j(t, X(t))dW_t^j$$

führt.

Für das Itô-Integral lassen sich eine Reihe von Eigenschaften beweisen; auf die meisten können wir hier nicht näher eingehen. In Lemma 4.2 werden wir eine wichtige Eigenschaft kennen lernen, eine weitere Eigenschaft ist die Tatsache, dass

$$E \left(\int_{t_0}^{t_1} F(t) dW_t \right) = 0 \quad (4.6)$$

ist, falls die Zufallsvariablen $F(t)$ und $W(s) - W(t)$ unabhängig sind für $s > t \geq 0$ (dies gilt immer, falls F die Lösung einer SDG bzgl. W ist). Eigenschaft (4.6) gilt zunächst für die Approximationen $I^{(N)}(F)$, weil nach Annahme an W und wegen der Unabhängigkeit und (1.6) die Gleichung

$$E(F(\tau_i)(W(\tau_{i+1}) - W(\tau_i))) = E(F(\tau_i)) \underbrace{E(W(\tau_{i+1}) - W(\tau_i))}_{=0} = 0 \quad (4.7)$$

gilt, und überträgt sich durch die Limesbildung auf $I(F)$.

Bemerkung 4.1 Es sollte hier erwähnt werden, dass es eine weitere sinnvolle stochastische Integraldefinition gibt, die auf R. Stratonovich⁴ zurück geht. Das Stratonovich-Integral

$$\int_{t_0}^{t_1} F(t) \circ dW_t$$

wird über eine ähnliche Limes-Bildung wie das Itô-Integral definiert und liefert ebenfalls eine mathematisch fundierte Definition stochastischer DGLs. Die beiden Integrale unterscheiden sich allerdings in den Rechenregeln ebenso wie in der Form der Lösungen. Die zugehörigen *Stratonovich-SDGs* werden in der Form

$$dX(t) = a(t, X(t))dt + b(t, X(t)) \circ dW_t$$

geschrieben. Aus Zeitgründen können wir auf diese zweite Definition und auf die Gemeinsamkeiten und Unterschiede zum Itô-Integral hier nicht näher eingehen. \square

4.2.1 Das Itô-Lemma

Zur Bewertung von Optionen, aber auch, um ganz allgemein mit dem Itô-Integral zu rechnen, benötigt man geeignete Rechenregeln. Falls $x(t)$ eine Lösung einer eindimensionalen deterministischen DGL $\dot{x}(t) = f(t, x(t))$ ist, so rechnet man für eine beliebige Funktion $g : \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}$ mittels der Kettenregel leicht nach, dass die Funktion $y(t) = g(t, x(t))$ die Gleichung

$$\dot{y}(t) = \frac{\partial g}{\partial t}(t, x(t)) + \frac{\partial g}{\partial x}(t, x(t))f(t, x(t))$$

erfüllt.

Das Itô-Lemma liefert nun eine Verallgemeinerung dieser Gleichung für die Lösung Itô-stochastischer DGL, und wird daher oft auch als stochastische Verallgemeinerung der Kettenregel bezeichnet.

⁴russischer Mathematiker, 1930–1997

Lemma 4.2 Sei $g : \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}$ eine zwei mal stetig differenzierbare Funktion und sei $X(t)$ die Lösung einer reellwertigen Itô-SDG vom Typ (4.5). Dann erfüllt $Y(t) = g(t, X(t))$ die Gleichung

$$\begin{aligned} dY(t) &= \left(\frac{\partial g}{\partial t}(t, X(t)) + \frac{\partial g}{\partial x}(t, X(t))a(t, X(t)) + \frac{1}{2} \frac{\partial^2 g}{\partial x^2} b(t, X(t))^2 \right) dt \\ &\quad + \frac{\partial g}{\partial x}(t, X(t))b(t, X(t))dW_t, \end{aligned}$$

wobei W hier gerade der Wiener Prozess aus der SDG ist, die $X(t)$ erfüllt. Diese Formel wird auch *Itô-Formel* genannt.

Beweisidee: Wir werden hier keinen vollständigen Beweis betrachten, wollen aber begründen, warum die (im Vergleich zur deterministischen Formel ungewöhnliche) zweite Ableitung von g nach x hier auftritt. Wir betrachten dazu zwei Zeitpunkte t und $t + \Delta t$. Dann gilt $X(t + \Delta t) = X(t) + \Delta X(t)$ mit

$$\begin{aligned} \Delta X(t) &= \int_t^{t+\Delta t} a(s, X(s))ds + \int_t^{t+\Delta t} b(s, X(s))dW_s \\ &\approx a(t, X(t))\Delta t + b(t, X(t))\Delta W_t, \end{aligned} \tag{4.8}$$

wobei $\Delta W_t = W(t + \Delta t) - W(t)$ ist. Die Approximation (4.8) folgt hierbei aus der Limes-Definition des Itô-Integrals, wenn wir $\tau_{i+1} - \tau_i = \Delta t$ setzen. Wir betrachten nun die Größe

$$\Delta Y := g(t + \Delta t, X(t) + \Delta X(t)) - g(t, X(t)).$$

Aus der Taylor-Entwicklung von g folgt

$$\begin{aligned} \Delta Y &\approx \frac{\partial g}{\partial x} \Delta X + \frac{\partial g}{\partial t} \Delta t + \frac{1}{2} \frac{\partial^2 g}{\partial x^2} (\Delta X)^2 \\ &\approx \frac{\partial g}{\partial x} \left(a(t, X(t))\Delta t + b(t, X(t))\Delta W_t \right) + \frac{\partial g}{\partial t} \Delta t \\ &\quad + \frac{1}{2} \frac{\partial^2 g}{\partial x^2} \left(a(t, X(t))^2 (\Delta t)^2 + 2a(t, X(t))b(t, X(t))\Delta t \Delta W_t + (b(t, X(t))\Delta W_t)^2 \right), \end{aligned}$$

wobei alle Ableitungen in $(t, X(t))$ ausgewertet werden. Um aus dieser Approximation die ‘‘Ableitung’’

$$dY(t) = \lim_{\Delta t \rightarrow 0} \Delta Y$$

zu berechnen, müssen wir alle Terme der Ordnung $O(\Delta t)$ berücksichtigen. Wären alle Größen deterministisch, so blieben hier gerade die Terme mit den ersten Ableitungen von g stehen. Im stochastischen Fall ist aber auch $(\Delta W_t)^2$ ein Term der Ordnung $O(\Delta t)$, denn es gilt

$$E((\Delta W_t)^2) = E(\underbrace{(E(\Delta W_t) - \Delta W_t)^2}_{=0}) = \text{Var}(\Delta W_t) = \text{Var}(W(t + \Delta t) - W(t)) = \Delta t.$$

Der Term

$$\frac{1}{2} \frac{\partial^2 g}{\partial x^2} b(t, X(t))^2 (\Delta W_t)^2$$

muss also mit berücksichtigt werden. Führt man nun den Grenzübergang für $\Delta t \rightarrow 0$ im richtigen stochastischen Sinne durch, so erhält man gerade die behauptete Formel. \square

Bemerkung 4.3 Es existiert auch eine Version des Itô-Lemmas für mehrdimensionale SDGs. In diesem Fall gehen — durch die im Beweis verwendete Taylor-Entwicklung für g — alle gemischten zweiten Ableitungen der Form $\frac{\partial^2 g}{\partial x_i \partial x_j}$ in die Gleichung ein. \square

4.2.2 Die geometrische Brownsche Bewegung

Zum Abschluss dieses Kapitels stellen wir hier ein einfaches aber sehr verbreitetes SDG-Modell zur Modellierung von Kursverläufen vor. Wie in den vergangenen Kapiteln schreiben wir hier nun wieder $S(t)$ statt $X(t)$.

Wie bereits in Abschnitt 3.4 erwähnt, möchte man in dieser Modellierung mindestens den Mittelwert und die Varianz des Kurses bestimmen. In der geometrischen Brownschen Bewegung stehen dazu die zwei Parameter μ und σ zur Verfügung, wobei die *Rücklaufquote* oder auch *Rendite* μ das mittlere Wachstum oder den “Trend” des Kurses modelliert und die *Volatilität* σ die Varianz, die hier von der aktuellen Höhe des Kurses abhängt.

Das zugehörige SDG-Modell ist gegeben durch

$$dS(t) = \mu S(t)dt + \sigma S(t)dW_t. \quad (4.9)$$

Trotz seiner Einfachheit ist dieses Modell Grundlage vieler finanzmathematischer Anwendungen in der Modellierung von Kurverläufen, auch wenn es heutzutage immer mehr durch komplexere Modelle ersetzt wird.

Das Schöne an dieser einfachen Gleichung ist die Tatsache, dass man die Lösungen noch per Hand angeben kann. Mit Hilfe des Itô-Lemmas rechnet man nach, dass

$$S(t; S_0) = S_0 \exp \left(\left(\mu - \frac{1}{2} \sigma^2 \right) t + \sigma W(t) \right) \quad (4.10)$$

die Lösung von (4.9) ist: Wir betrachten die SDG $dZ(t) = dW(t)$ (also $a \equiv 0$ und $b \equiv 1$) und wenden das Itô-Lemma an auf $Y(t) = g(t, Z(t))$ mit

$$g(t, x) = S_0 \exp \left(\left(\mu - \frac{1}{2} \sigma^2 \right) t + \sigma x \right).$$

Wegen

$$\frac{\partial g}{\partial t}(t, x) = g(t, x) \left(\mu - \frac{1}{2} \sigma^2 \right), \quad \frac{\partial g}{\partial x}(t, x) = g(t, x) \sigma \quad \text{und} \quad \frac{\partial^2 g}{\partial x^2}(t, x) = g(t, x) \sigma^2$$

und $Y(t) = g(t, Z(t)) = S(t)$ (weil $Z(t) = W(t)$ ist) folgt aus dem Itô-Lemma die SDG

$$dS(t) = dY(t) = \left(S(t) \left(\mu - \frac{1}{2} \sigma^2 \right) + \frac{1}{2} \sigma^2 S(t) \right) dt + S(t) \sigma dW_t = \mu S(t) dt + \sigma S(t) dW_t,$$

was zeigt, dass (4.10) tatsächlich eine Lösung der SDG (4.9) ist, wegen der Eindeutigkeit der Lösung also die einzige.

Für $\sigma = 0$ erhalten wir aus (4.9) gerade die bekannte lineare DGL $\dot{x}(t) = \mu x(t)$ und ihre zugehörige Lösung $x(t) = x(0)e^{\mu t}$. Durch diese Gleichung ist der Erwartungswert $E(S(t; S_0))$ bestimmt. Für die Lösungen von (4.9) gilt nämlich

$$\begin{aligned} E(S(t; S_0)) &= E(S(0)) + E\left(\int_0^t \mu S(\tau; S_0) d\tau\right) + \underbrace{E\left(\int_0^t \sigma S(\tau; S_0) dW_\tau\right)}_{=0, \text{ wegen (4.6)}} \\ &= E(S(0)) + \int_0^t \mu E(S(\tau; S_0)) d\tau. \end{aligned}$$

Die Funktion $e(t) = E(S(t, x_0))$ erfüllt also gerade die lineare DGL $\dot{e}(t) = \mu e(t)$ mit Anfangswert $e(0) = E(S_0) = S_0$, weswegen $E(S(t, x_0)) = S_0 e^{\mu t}$ ist.

Ebenso wie den Erwartungswert kann man auch die Varianz der Lösungen von (4.9) explizit berechnen: Es gilt

$$\begin{aligned} \text{Var}(S(t)) &= E\left(S_0^2 \exp\left(2\left(\mu - \frac{1}{2}\sigma^2\right)t + 2\sigma W(t)\right) - S_0^2 e^{2\mu t}\right) \\ &= E\left(S_0^2 e^{2\mu t} (\exp(-\sigma^2 t + 2\sigma W(t)) - 1)\right) \\ &= S_0^2 e^{2\mu t} (E(\exp(-\sigma^2 t + 2\sigma W(t))) - 1). \end{aligned}$$

Der Ausdruck unter dem Erwartungswert lässt sich nun schreiben als

$$\exp(-\sigma^2 t + 2\sigma W(t)) = \tilde{S}_0 \exp\left(\left(\tilde{\mu} - \frac{1}{2}\tilde{\sigma}^2\right)t + \tilde{\sigma}W(t)\right)$$

mit $\tilde{S}_0 = 1$, $\tilde{\mu} = \sigma^2$ und $\tilde{\sigma} = 2\sigma$. Er ist also gerade wieder eine geometrische Brownsche Bewegung, für die folglich

$$E\left(\tilde{S}_0 \exp\left(\left(\tilde{\mu} - \frac{1}{2}\tilde{\sigma}^2\right)t + \tilde{\sigma}W(t)\right)\right) = \tilde{S}_0 e^{\tilde{\mu}t} = e^{\sigma^2 t}$$

gilt. Also erhalten wir für die Varianz

$$\text{Var}(S(t)) = S_0^2 e^{2\mu t} (e^{\sigma^2 t} - 1).$$

Um die geometrische Brownsche Bewegung in der risikoneutralen Optionsbewertung verwenden zu können, muss gemäß (2.5) die Bedingung

$$E(S(T)) = e^{rT} S(0)$$

gelten, woraus sofort $\mu = r$ folgt. Es bleibt also nur ein einziger Parameter übrig, nämlich die Volatilität σ . Wie bereits in Abschnitt 3.4 erläutert, wird dieser Parameter entweder aus vergangenen Kursverläufen historisch oder aus Marktwerten implizit geschätzt. An diese Stelle soll noch einmal betont werden, dass natürlich auch die Kursmodellierung mit stochastischen Differentialgleichungen keine Vorhersagen über den Kurs zulässt; sie

erlaubt nur eine stochastische Modellierung des zukünftigen Verlaufes auf Basis geschätzter charakteristischer Werte.

Das im letzten Kapitel besprochene Binomialmodell wird oft als Approximation der geometrischen Brownschen Bewegung interpretiert. Warum dies gerechtfertigt ist, können wir im Detail erst dann verstehen, wenn wir uns mit der Numerik stochastischer Differentialgleichungen beschäftigen haben. Der Zusammenhang zwischen den Parametern in den beiden Modellen wird allerdings bereits auf dem aktuellen Übungsblatt untersucht.

4.3 Ein anspruchsvolleres Kursmodell

Um zu zeigen, dass man mit stochastischen Differentialgleichungen auch viel anspruchsvollere Modelle als nur die geometrische Brownsche Bewegung aufstellen kann, wollen wir hier noch kurz ein mehrdimensionales Kursmodell angeben. Dieses ist gegeben durch den dreidimensionalen Prozess $X(t) = (S(t), \sigma(t), \zeta(t))^T$, der durch die Gleichungen

$$\begin{aligned} dS(t) &= \mu S(t)dt + \sigma(t)S(t)dW_t^1 \\ d\sigma(t) &= -(\sigma(t) - \zeta(t))dt + \alpha\sigma(t)dW_t^2 \\ d\zeta(t) &= \beta(\sigma(t) - \zeta(t))dt \end{aligned}$$

mit Parametern $\mu \in \mathbb{R}$, $\alpha > 0$ und $\beta \geq 0$ gegeben ist. In der ersten Komponente erkennen wir (4.9) wieder, allerdings ist σ nun kein fester Parameter sondern ebenfalls ein stochastischer Prozess, was die Realität sicherlich deutlich angemessener widerspiegelt. Um zu verstehen, wie sich σ hier in der Zeit verändert, ist es sinnvoll, zunächst die Gleichung für ζ anzuschauen. Für jeden Pfad $\sigma(t, \omega)$ ist dies eine einfache inhomogene lineare DGL, für die man die exakte Lösung als

$$\zeta(t) = e^{-\beta t} \zeta(0) + \int_0^t e^{-\beta(t-s)} \sigma(s, \omega) ds$$

angeben kann. Der Wert ζ kann also als (durch $e^{-\beta(t-s)}$ gewichteter) Mittelwert für σ aufgefasst werden. Die Gleichung für σ bewirkt dann, dass die Volatilität σ ihrem Mittelwert folgt. Wäre dies eine deterministische Gleichung, so wäre das Verhalten langweilig da (für $\beta > 0$) sowohl ζ als auch σ gegen Null konvergieren würden. Da die Gleichung für σ aber durch einen zweiten Wiener Prozess W^2 wiederum stochastisch gestört wird, ergibt sich ein durchaus interessantes Verhalten der Gleichung.

Zur Erstellung noch realistischerer Kursmodelle geht man in der aktuellen Forschung inzwischen über Itô-SDGen hinaus. Diese sind nämlich nur schlecht geeignet, um plötzliche Kurssprünge modellieren zu können. Zu diesem Zweck erweitert man die Gleichungen um Sprungterme, die durch Poisson-Prozesse modelliert werden. Die Optionsbewertung auf Basis solcher Modelle wäre wiederum ein schönes Diplomarbeitsthema.

Kapitel 5

Die Monte-Carlo Methode

In diesem Kapitel werden wir einen Algorithmus zur Bewertung von Optionen kennen lernen, der auf stochastischen Simulationen beruht. Im Gegensatz zur Binomialmethode hat er den großen Vorteil, auch für komplexere Kursmodelle zu funktionieren. Es gibt aber auch Nachteile: die Methode konvergiert relativ langsam und sie eignet sich nicht für amerikanische Optionen (jedenfalls nicht für den amerikanischen Put).

5.1 Grundidee

Die Monte-Carlo Methode dient dazu, den Erwartungswert $\mu = E(X)$ einer Zufallsvariablen X zu schätzen. Sie beruht auf der wiederholten Durchführung von Zufallsexperimenten, d.h. auf simulierten Realisierungen von Zufallsvariablen X . In der Praxis geschieht dies durch den Zufallszahlengenerator im Computer. Dieser simuliert die Durchführung voneinander unabhängiger Zufallsexperimente. Je nach Programmiersprache bzw. Programmierumgebung stehen dabei verschiedene Verteilungen zur Verfügung. Die für uns oft wichtigen normalverteilten Zufallsvariablen werden in MATLAB z.B. mit der Anweisung `randn` erzeugt.

Das wiederholte Ziehen von Zufallszahlen x_1, x_2, x_3, \dots kann dabei auf zwei Arten stochastisch interpretiert werden: Zum einen kann man die x_i als verschiedene unabhängige Realisierungen ein und derselben Zufallsvariablen X interpretieren, also $x_i = X(\omega_i)$, zum anderen kann man die x_i als Realisierungen unabhängiger und identisch verteilter Zufallsvariablen $X_i \sim X$, also $x_i = X_i(\omega)$ auffassen. Beide Sichtweisen sind äquivalent und wir werden beide im Folgenden verwenden.

Die Funktion der Monte-Carlo Methode beruht auf dem zentralen Grenzwertsatz der Stochastik. Dieser Satz gilt für unabhängige und identisch verteilte Zufallsvariablen X_1, X_2, \dots mit Mittelwert $\mu = E(X_i)$ und Varianz $\sigma^2 = \text{Var}(X_i)$. Definiert man die neuen Zufallsvariablen

$$Y_n := \sum_{i=1}^n X_i,$$

so besagt der Satz, dass Y_n für große n approximativ $N(n\mu, n\sigma^2)$ -verteilt ist, also normal-

verteilt mit Mittelwert $n\mu$ und Varianz $n\sigma^2$. Genauer gilt die Konvergenz

$$\lim_{n \rightarrow \infty} P\left(\frac{Y_n - n\mu}{\sigma\sqrt{n}} \leq x\right) = F(x) \quad (5.1)$$

punktweise für alle $x \in \mathbb{R}$, wobei

$$F(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{y^2}{2}} dy$$

die Verteilungsfunktion der Standard-Normalverteilung ist. Ist das (zentrierte) dritte Moment $E((X_i - \mu)^3)$ endlich, so gilt darüberhinaus nach dem Satz von Berry-Esseen

$$\left| P\left(\frac{Y_n - n\mu}{\sigma\sqrt{n}} \leq x\right) - F(x) \right| = O\left(\frac{1}{\sqrt{n}}\right)$$

und die Konvergenz ist gleichmäßig in x .

Geben wir uns nun ein $\varepsilon > 0$ vor, so folgt

$$P\left(\frac{Y_n}{n} - \mu \leq \varepsilon\right) = P(Y_n - n\mu \leq n\varepsilon) = P\left(\frac{Y_n - n\mu}{\sigma\sqrt{n}} \leq \frac{\sqrt{n}\varepsilon}{\sigma}\right) \approx F\left(\frac{\sqrt{n}\varepsilon}{\sigma}\right) \rightarrow 1$$

und

$$P\left(\frac{Y_n}{n} - \mu \leq -\varepsilon\right) = P(Y_n - n\mu \leq -n\varepsilon) = P\left(\frac{Y_n - n\mu}{\sigma\sqrt{n}} \leq -\frac{\sqrt{n}\varepsilon}{\sigma}\right) \approx F\left(-\frac{\sqrt{n}\varepsilon}{\sigma}\right) \rightarrow 0$$

für $n \rightarrow \infty$. Damit folgt

$$P\left(\left|\frac{Y_n}{n} - \mu\right| \leq \varepsilon\right) \approx F\left(\frac{\sqrt{n}\varepsilon}{\sigma}\right) - F\left(-\frac{\sqrt{n}\varepsilon}{\sigma}\right) \rightarrow 1 \quad (5.2)$$

für $n \rightarrow \infty$. Betrachten wir nun Realisierungen $X_i(\omega)$ der Zufallsvariablen, oder — äquivalent — verschiedene unabhängige Realisierungen $X(\omega_i)$ unserer ursprünglichen Zufallsvariablen $X \sim X_i$, so gilt für große n mit hoher Wahrscheinlichkeit die Approximation

$$\frac{1}{n} \sum_{i=1}^n X(\omega_i) \approx \mu = E(X).$$

Dies ist die einfache Grundidee der Monte-Carlo-Simulation: Um den Erwartungswert $E(X)$ einer Zufallsvariablen zu schätzen, führen wir n (numerisch simulierte) Zufallsexperimente $X(\omega_1), \dots, X(\omega_n)$ durch und berechnen den Mittelwert.

Für eine Monte-Carlo-Simulation auf Basis der geometrischen Brownschen Bewegung (4.9) spezifizieren wir die Schritte (2) und (3) des Grundalgorithmus 2.1 für europäische Optionen daher wie folgt:

- (2) Setze $S(T) = S_0 \exp\left((\mu - \frac{1}{2}\sigma^2)T + \sigma W(T)\right)$ mit $W(T) \sim N(0, T)$

- (3) Simuliere n Kurswerte $S(T, \omega_1), \dots, S(T, \omega_n)$ numerisch und approximiere den Optionswert als den abgezinnten Mittelwert

$$V(t, S(t)) = e^{r(t-T)} \frac{1}{n} \sum_{i=1}^n V(T, S(T, \omega_i)).$$

Zur Berechnung von $S(T, \omega_1), \dots, S(T, \omega_n)$ werden dabei mittels des Zufallsgenerators die n $N(0, T)$ -verteilten Zufallszahlen $W(T, \omega_1), \dots, W(T, \omega_n)$ erzeugt und in die Formel in (2) eingesetzt. Die Formel für $V(T, S)$ in (3) ergibt sich gemäß (2.1) oder (2.2), je nachdem ob ein Call oder Put bewertet werden soll.

Der große Vorteil der Monte-Carlo Methode ist es, dass in Schritt (2) des Algorithmus im Prinzip jedes numerisch simulierbare Kursmodell eingesetzt werden kann. Dieses muss dabei gar keine geschlossene Lösungsformel zulassen, da man statt dessen numerische Lösungsverfahren verwenden kann. Näheres dazu behandeln wir in Kapitel 6.

Bemerkung 5.1 Im Gegensatz zur Binomialmethode lässt sich die Monte-Carlo Methode leider nicht auf amerikanische Put-Optionen verallgemeinern. Der Grund dafür ist der folgende: Um zu einem Zeitpunkt t auf Basis des Kurses $S(t)$ zu entscheiden, ob die vorzeitige Ausübung einen Vorteil bietet (und damit zu einem höheren Wert führt), muss man den Erwartungswert des Optionswertes über die von $S(t)$ ausgehenden zukünftigen Kurspfade kennen. In der Binomialmethode sind diese Werte vorhanden, in der Monte-Carlo-Methode hingegen nicht, weil hier nur einzelne Pfade $S(t)$, $t \in [0, T]$, aber keine Bäume betrachtet werden.

Zwar ließe sich die Methode so variieren, dass entlang jedes Pfades an einer Auswahl diskreter Zeitpunkte der Vorteil der vorzeitigen Ausübung überprüft wird. Dies liefert aber nicht den richtigen Wert (sondern im Allgemeinen einen deutlich größeren), da die Entscheidung dann nur auf Basis eines einzigen zukünftigen Pfades und nicht auf Basis des Erwartungswertes getroffen wird. Dies wäre auch in der Praxis nicht umsetzbar, da dies einer Ausübungsentscheidung mit "Blick in die Zukunft" entsprechen würde. \square

5.2 Konvergenzgeschwindigkeit

Numerische Tests (Übungsaufgabe) zeigen, dass die Monte-Carlo Methode in ihrer einfachen Form für wachsende Anzahl von Simulationen n relativ langsam konvergiert. Wir wollen die Konvergenzgeschwindigkeit hier zunächst theoretisch untersuchen um dann eine Strategie zu entwickeln, mit der man schnellere Konvergenz erhält.

Da die Monte-Carlo Methode nur im stochastischem Sinne konvergiert, müssen wir dazu zunächst ein geeignetes Maß finden. Dies ist das folgende:

Definition 5.2 Gegeben sei $p \in [0, 1]$, $n \in \mathbb{N}$. Ein Intervall der Form $I = [\mu - \varepsilon, \mu + \varepsilon]$ heißt p -Konfidenzintervall der Monte-Carlo Methode, wenn

$$P\left(\frac{1}{n} \sum_{i=1}^n X_i \in I\right) = p$$

ist. \square

Halten wir nun p fest, so können wir untersuchen, wie groß diese Konfidenzintervalle abhängig von n sind. Hierfür gilt der folgende Satz:

Satz 5.3 Gegeben sei $p \in (0, 1)$. Dann existiert $k > 0$ und eine Folge $p_n \rightarrow p$, so dass die p_n -Konfidenzintervalle I_n von der Form

$$I_n = \left[\mu - \frac{k\sigma}{\sqrt{n}}, \mu + \frac{k\sigma}{\sqrt{n}} \right]$$

sind.

Beweis: Wähle $x \in \mathbb{R}$ so, dass für die Verteilungsfunktion der Standard-Normalverteilung

$$F(x) - F(-x) = p$$

gilt. Auf Grund der Konvergenz (5.1) existiert für jedes $x \in \mathbb{R}$ eine Folge $\delta_n(x) \rightarrow 0$ mit

$$\left| P \left(\frac{Y_n - n\mu}{\sigma\sqrt{n}} \leq x \right) - F(x) \right| = \delta_n(x).$$

Analog zur Herleitung von (5.2) erhalten wir

$$P \left(\left| \frac{Y_n - n\mu}{\sigma\sqrt{n}} \right| \leq x \right) = F(x) - F(-x) + \delta_n(x) - \delta_n(-x) = p + \delta_n(x) - \delta_n(-x) =: p_n.$$

Sicherlich gilt mit dieser Definition $p_n \rightarrow p$. Zudem gilt

$$\left| \frac{Y_n - n\mu}{\sigma\sqrt{n}} \right| \leq x \Leftrightarrow \frac{Y_n}{n} \in \left[\mu - \frac{x\sigma}{\sqrt{n}}, \mu + \frac{x\sigma}{\sqrt{n}} \right].$$

Die Behauptung folgt also mit $k = x$. □

Wichtig an diesem Satz ist weniger die Folge $p_n \rightarrow p$, sondern die Aussage darüber, wie schnell die Konfidenzintervalle in Abhängigkeit von n schrumpfen. Tatsächlich hängt die Breite des Intervalls von \sqrt{n} ab, was bedeutet, dass man hundert mal so viele Simulationen benötigt um im Ergebnis eine korrekte Nachkommastelle hinzuzugewinnen. Dies erklärt die langsame Konvergenz der Methode.

Die Analyse zeigt aber auch, dass die Konfidenzintervalle kleiner werden, wenn die Varianz σ^2 der Zufallsvariablen X kleiner wird. Dies ist ein wichtiger Ansatzpunkt zur Erhöhung der Konvergenzgeschwindigkeit. Im folgenden Abschnitt betrachten wir daher eine Methode zur Reduktion der Varianz von X .

5.3 Antithetische Zufallsvariablen

Wir betrachten nun eine Monte-Carlo Simulation mit Zufallsvariable $X = f(Z)$, wobei $f : \mathbb{R} \rightarrow \mathbb{R}$ eine gegebene Funktion ist und $Z \sim N(0, 1)$ eine standard-normalverteilte Zufallsvariable ist.

Dies ist genau die Form unseres Optionsbewertungsproblems aus den obigen Schritten (2) und (3) mit

$$f(Z) = \left(S_0 \exp \left(\left(\mu - \frac{1}{2} \sigma^2 \right) T + \sigma \sqrt{T} Z \right) - K \right)^+ \quad (5.3)$$

(für einen Call, alles Folgende geht analog auch für den Put), wobei wir ausgenutzt haben, dass sich die Zufallsvariable $W(T) \sim N(0, T)$ auch als $\sqrt{T}Z$ mit $Z \sim N(0, 1)$ schreiben lässt.

Die Idee ist nun, die Abbildung f durch eine neue Zufallsvariable \hat{f} zu ersetzen, so dass $E(\hat{f}(Z)) = E(f(Z))$ und $\text{Var}(\hat{f}(Z)) < \text{Var}(f(Z))$ gilt.

Dazu verwenden wir die sogenannte *antithetische* Zufallsvariable zu Z , die gerade durch $-Z$ gegeben ist. Wir verwenden diese, in dem wir $f(Z)$ durch

$$\hat{f}(Z) := \frac{f(Z) + f(-Z)}{2}$$

ersetzen. In der Monte-Carlo Simulation führt dies zum Ersetzen von

$$\frac{1}{n} \sum_{i=1}^n f(Z(\omega_i)) \quad \text{durch} \quad \frac{1}{n} \sum_{i=1}^n \frac{f(Z(\omega_i)) + f(-Z(\omega_i))}{2}.$$

Jede gezogene Zufallszahl $Z(\omega_i)$ geht also zweimal — einmal positiv und einmal negativ — in f und damit in die Summe ein.

Der folgende Satz zeigt, dass wir damit unser Ziel der Varianzreduktion bei gleichbleibendem Erwartungswert erreichen.

Satz 5.4 Betrachte die Monte-Carlo Methode zur Optionsbewertung einer europäischen Call- oder Put-Option mit der geometrischen Brownschen Bewegung als Kursmodell. Dann gelten für die Zufallsvariable

$$\hat{f}(Z) = \frac{f(Z) + f(-Z)}{2}$$

die (Un-)Gleichungen

$$E(\hat{f}(Z)) = E(f(Z)) \quad \text{und} \quad \text{Var}(\hat{f}(Z)) \leq \text{Var}(f(Z))/2.$$

Beweis: Es gilt

$$E \left(\frac{f(Z) + f(-Z)}{2} \right) = \frac{E(f(Z)) + E(f(-Z))}{2} = \frac{E(f(Z)) + E(f(Z))}{2} = E(f(Z))$$

wegen $-Z \sim N(0, 1) \Leftrightarrow Z \sim N(0, 1)$. Dies zeigt die Gleichheit der Erwartungswerte.

Zur Untersuchung der Varianz verwenden wir die Formel

$$\text{Var} \left(\frac{f(Z) + f(-Z)}{2} \right) = \frac{1}{2} \left(\text{Var}(f(Z)) + \text{Cov}(f(Z), f(-Z)) \right), \quad (5.4)$$

die mit der üblichen Kovarianz

$$\text{Cov}(X, Y) = E[(X - E(X))(Y - E(Y))] = E[XY] - E[X]E[Y]$$

aus der Gleichung

$$\text{Var}(X + Y) = \text{Var}(X) + \text{Var}(Y) + 2\text{Cov}(X, Y)$$

und der Identität $\text{Var}(f(Z)) = \text{Var}(f(-Z))$ folgt.

Wir beweisen zunächst das folgende Hilfsresultat: Es seien $a, b : \mathbb{R} \rightarrow \mathbb{R}$ monotone Abbildungen mit umgekehrter Monotonie (also eine wachsend und eine fallend) Dann gilt für jede beliebige Zufallsvariable X die Ungleichung

$$\text{Cov}(a(X), b(X)) \leq 0. \quad (5.5)$$

Zum Beweis von (5.5) betrachten wir eine weitere beliebige Zufallsvariable Y . Dann gilt wegen der Monotonie von a und b fast sicher

$$(a(X) - a(Y))(b(X) - b(Y)) \leq 0$$

und daher auch $E[(a(X) - a(Y))(b(X) - b(Y))] \leq 0$. Damit erhalten wir

$$\begin{aligned} 0 &\geq E[(a(X) - a(Y))(b(X) - b(Y))] \\ &= E[a(X)b(X)] - E[a(X)b(Y)] - E[a(Y)b(X)] + E[a(Y)b(Y)]. \end{aligned}$$

Wählen wir nun $X \sim Y$ und Y unabhängig von X , so folgt

$$\begin{aligned} 0 &\geq E[a(X)b(X)] - E[a(X)b(Y)] - E[a(Y)b(X)] + E[a(Y)b(Y)] \\ &= 2E[a(X)b(X)] - 2E[a(X)]E[b(X)] \\ &= 2\text{Cov}(a(X), b(X)) \end{aligned}$$

und damit (5.5).

Unsere Funktion f aus (5.3) ist monoton wachsend (für einen Put wäre sie monoton fallend), weswegen die Funktion $x \mapsto f(-x)$ monoton fallend ist (für einen Put wäre sie monoton wachsend). Also können wir (5.5) mit $a(x) = f(x)$ und $b(x) = f(-x)$ anwenden und es folgt

$$\text{Cov}(f(Z), f(-Z)) \leq 0,$$

was eingesetzt in (5.4) die Behauptung zeigt. \square

In der Praxis beobachtet man oft, dass die Verwendung der antithetischen Variablen einen deutlich größeren Vorteil bringt, als es die bloße Halbierung der Varianz vermuten lässt.

Kapitel 6

Numerische Lösung stochastischer Differentialgleichungen

In diesem Kapitel geben wir eine kurze Einführung in die numerische Lösung stochastischer Itô-Differentialgleichungen (4.5). Grundwissen über die numerische Lösung deterministischer gewöhnlicher Differentialgleichungen wird dabei vorausgesetzt.

Ziel des Kapitels ist zum einen das Bereitstellen von Algorithmen, mit denen Approximationen $\tilde{X}(T) \approx X(T, X_0)$ der Lösung von (4.5) zum Anfangswert X_0 berechnet werden können. Zum anderen werden wir zeigen, dass — und in welchem Sinne — das in der Binomialmethode verwendete Modell eine Approximation der geometrischen Brownschen Bewegung darstellt.

6.1 Stochastische Konvergenz- und Approximationsbegriffe

Bei der numerischen Lösung stochastischer Differentialgleichungen kann man im Allgemeinen nicht erwarten, dass man für jeden möglichen Pfad $X(T, X_0, \omega)$ eine gute Approximation $\tilde{X}(T, \omega)$ berechnen kann. Dies ist aber auch nicht unbedingt notwendig: Für die allermeisten Anwendungen reicht es aus, wenn die berechneten Approximationen im Mittel hinreichend gut sind.

Es gibt allerdings verschiedene Möglichkeiten, wie man dieses “im Mittel” definieren kann und je nach dem, was man berechnen möchte, benötigt man verschiedene Konvergenzbegriffe. Tatsächlich gibt es in der stochastischen Numerik eine Vielzahl von Konvergenzbegriffen, von denen wir hier nur zwei herausheben wollen, die für unsere Zwecke besonders wichtig sind.

Wir betrachten dabei einen stochastischen Prozess X auf dem Intervall $[0, T]$. Dieser soll durch eine Folge numerischer Approximationen \tilde{X}_i auf Zeitgittern $\mathcal{T}_i = \{0, h_i, 2h_i, \dots, N_i h_i\}$ mit Schrittweiten $h_i = T/N_i$ (also $N_i h_i = T$) approximiert werden mit $N_i \rightarrow \infty$ für $i \rightarrow \infty$. Wir schreiben kurz $X(T) = X(T, \omega)$ und $\tilde{X}_i(T) = \tilde{X}_i(T, \omega)$.

Definition 6.1 (i) Die Folge \tilde{X}_i von stochastischen Prozessen heißt *starke Approximation* für X zur Zeit T bzgl. einer Funktion $g : \mathbb{R}^n \rightarrow \mathbb{R}^m$, wenn die Bedingung

$$\lim_{i \rightarrow \infty} E(\|g(X(T)) - g(\tilde{X}_i(T))\|) = 0$$

gilt. Sie heißt *starke Approximation mit Ordnung $\gamma > 0$* , falls für alle $i \geq i_0$ zusätzlich die Abschätzung

$$E(\|g(X(T)) - g(\tilde{X}_i(T))\|) \leq Ch_i^\gamma$$

für ein $C > 0$ gilt.

(ii) Die Folge \tilde{X}_i von stochastischen Prozessen heißt *schwache Approximation* für X zur Zeit T bzgl. einer Funktion $g : \mathbb{R}^n \rightarrow \mathbb{R}^m$, wenn die Bedingung

$$\lim_{i \rightarrow \infty} \|E(g(X(T))) - E(g(\tilde{X}_i(T)))\| = 0$$

gilt. Sie heißt *schwache Approximation mit Ordnung $\beta > 0$* , falls für alle $i \geq i_0$ zusätzlich die Abschätzung

$$\|E(g(X(T))) - E(g(\tilde{X}_i(T)))\| \leq Ch_i^\beta$$

für ein $C > 0$ gilt. □

Eine starke Approximation liefert also eine Approximation der Pfade, bei denen zwar u.U. nicht jeder Pfad gut approximiert wird, der Fehler aber zumindest im Mittel klein werden muss. Dies ist die natürliche Verallgemeinerung der deterministischen Approximation.

Bei einer schwachen Approximation hingegen (die sich, wie wir sehen werden, mit deutlich geringerem numerischen Aufwand berechnen lässt), können die einzelnen Pfade ganz anders aussehen, wichtig ist hier nur, dass die bezüglich g ermittelten statistischen Eigenschaften gleich sind.

Kurz gesagt können wir festhalten: Eine starke Approximation liefert eine *Approximation der Pfade* von X während eine schwache Approximation eine *Approximation der statistischen Eigenschaften* von X liefert, jeweils gemessen bzgl. g .

Das folgende Beispiel zeigt typische Anwendungen starker und schwacher Approximationen.

Beispiel 6.2 (i) Wenn wir den Erwartungswert $E(X(T))$ numerisch berechnen wollen, genügt eine schwache Approximation bzgl. $g(x) = x$, denn dafür gilt

$$\lim_{i \rightarrow \infty} |E(X(T)) - E(\tilde{X}_i(T))| = 0,$$

und damit

$$\lim_{i \rightarrow \infty} E(\tilde{X}_i(T)) = E(X(T)).$$

(ii) Wenn wir die Varianz $\text{Var}(X(T))$ numerisch berechnen wollen, genügt eine schwache Approximation bzgl. $g(x) = x$ und $g(x) = x^2$.

(iii) Für die Optionsbewertung genügt eine schwache Approximation bzgl. $g(x) = (x - K)^+$ bzw. $g(x) = (K - x)^+$.

(iv) Für kompliziertere Berechnungen reichen schwache Approximationen allerdings i.A. nicht mehr aus. Wenn wir z.B. für einen Wert $c \in \mathbb{R}$ die “minimale Überschreitungszeit”

$$t(\omega) := \inf\{t \geq 0 \mid X(t, \omega) \geq c\}$$

für ein vorgegebenes $c \in \mathbb{R}$ definieren, so erhalten wir eine weitere Zufallsvariable $t(\omega)$. Dies könnte z.B. die Überschreitung eines vorgegebenen Kursniveaus sein, was zur Bewertung von Barriere-Optionen wichtig ist. Bei diesen Optionstypen hängt der Wert der Option davon ab, ob der Wert des Kurses bis zum Laufzeitende ein vorgegebenes Kursniveau über- oder unterschreitet. Ziel einer numerischen Berechnung könnte es nun sein, die mittlere Zeit $E\{t(\omega)\}$ zu ermitteln. Da sich diese Größe nicht als $E\{g(X(T))\}$ schreiben lässt, genügt hier eine schwache Approximation nicht mehr. \square

Das die starke Approximation tatsächlich eine stärkere Eigenschaft als die schwache Approximation ist, zeigt das folgende Lemma, in dem wir uns der Einfachheit halber auf reellwertige g beschränken.

Lemma 6.3 Wenn $\tilde{X}_i(T)$ eine starke Approximation von $X(T)$ bzgl. $g : \mathbb{R}^n \rightarrow \mathbb{R}$ ist, so ist $\tilde{X}_i(T)$ auch eine schwache Approximation bzgl. g . Hierbei bleibt die Konvergenzordnung erhalten, d.h. es gilt $\beta \geq \gamma$.

Beweis: Da der Erwartungswert monoton ist, folgt

$$E(g(X(T)) - g(\tilde{X}(T))) \leq E(|g(X(T)) - g(\tilde{X}(T))|)$$

und

$$E(g(\tilde{X}(T)) - g(X(T))) \leq E(|g(\tilde{X}(T)) - g(X(T))|) = E(|g(X(T)) - g(\tilde{X}(T))|).$$

Damit folgt

$$|E(g(\tilde{X}(T)) - g(X(T)))| \leq E(|g(X(T)) - g(\tilde{X}(T))|)$$

und damit die Behauptung. \square

Eine nützliche Eigenschaft für starke Approximationen gibt das folgende Lemma.

Lemma 6.4 Es sei $h : \mathbb{R}^m \rightarrow \mathbb{R}^l$ eine global Lipschitz stetige Funktion. Falls \tilde{X}_i eine starke Approximation mit einer gegebenen Ordnung für eine Funktion $g : \mathbb{R}^n \rightarrow \mathbb{R}^m$ ist, so ist \tilde{X}_i auch eine starke Approximation mit einer gegebenen Ordnung für die Funktion $h \circ g$.

Beweis: Es sei $L > 0$ die globale Lipschitz-Konstante von h . Für den Ausdruck, der die starke Konvergenz beschreibt, gilt dann

$$E(\|h \circ g(X(T)) - h \circ g(\tilde{X}_i(T))\|) \leq E(L\|g(X(T)) - g(\tilde{X}_i(T))\|) = LE(\|g(X(T)) - g(\tilde{X}_i(T))\|),$$

weswegen sich sowohl die starke Konvergenz als auch die dazugehörige Ordnung von g auf $h \circ g$ übertragen. \square

Leider gilt diese praktische Eigenschaft für die schwache Approximation nicht.

6.2 Approximationen des Wiener-Prozesses

Bevor wir uns der Lösung stochastischer Differentialgleichungen widmen, wollen wir zuerst untersuchen, wie wir den Wiener-Prozess approximieren können. Diese sind für die im Folgenden entwickelten numerischen Verfahren wichtig: Da der Wiener-Prozess in die Definition der stochastischen Differentialgleichungen eingeht, werden diese approximativen Wiener-Prozesse in den numerischen Schemata benötigt.

Algorithmus 6.5 (Starke Approximation des Wiener-Prozesses durch Gitterfunktion)

Gegeben: Schrittweite h , Gitter $\mathcal{T} = \{t_0, \dots, t_N\}$ mit $t_i = ih$

Gesucht: M approximative Pfade $\widetilde{W}(t_i, \omega_1), \dots, \widetilde{W}(t_i, \omega_M)$ des Wiener Prozesses auf \mathcal{T} .

(1) Für $j = 1, \dots, M$:

(2a) Erzeuge unabhängige $N(0, h)$ -verteilte Zufallszahlen $\Delta W_i(\omega_j)$ für $i = 0, \dots, N - 1$

(2b) Erzeuge Gitterfunktionen $\widetilde{W}(t_i, \omega_j)$ mittels der Rekursion

$$\widetilde{W}(t_0, \omega_j) = 0, \quad \widetilde{W}(t_{i+1}, \omega_j) = \widetilde{W}(t_i, \omega_j) + \Delta W_i(\omega_j), \quad i = 0, \dots, N - 1$$

(3) Ende der j -Schleife □

Man kann beweisen, dass dieser approximative Wiener-Prozess die Bedingungen des exakten Wiener-Prozesses an allen Gitterpunkten t_i erfüllt ((i) und (ii) sind Übungsaufgabe, (iii) folgt aus der Unabhängigkeit der Zufallszahlen in Algorithmus 6.5). Daher kann man jedem Pfad $\widetilde{W}(\cdot, \omega)$ des approximierten Wiener-Prozesses gerade einen Pfad des exakten Prozesses $W(\cdot, \omega)$ mit $\widetilde{W}(t_i, \omega) = W(t_i, \omega)$ für jedes $t_i \in \mathcal{T}$ zuordnen, so dass die Approximation für jedes $g : \mathbb{R} \rightarrow \mathbb{R}$ und jeden Gitterpunkt $T \in \mathcal{T}$ die Bedingung

$$E\{|g(\widetilde{W}(T, \omega)) - g(W(T, \omega))|\} = 0$$

erfüllt. Es handelt sich also um eine “besonders starke” Form der starken Konvergenz.

Wir wollen nun untersuchen, wie man eine schwache Approximation des Wiener-Prozesses erhalten kann — in der Hoffnung, dass sich der Algorithmus dabei vereinfacht. Der folgende Algorithmus zeigt, wie dies geht. Wir benötigen eine spezielle Form von gleichverteilten diskreten Zufallsvariablen, nämlich die durch

$$X(\Omega) = \{x_1, x_2\}, \quad \mathbb{P}_X(\{x_1\}) = \mathbb{P}_X(\{x_2\}) = \frac{1}{2}$$

definierte *zweipunktverteilte Zufallsvariable* mit den zwei Werten $\{x_1, x_2\}$.

Algorithmus 6.6 (Schwache Approximation des Wiener-Prozesses durch Gitterfunktion)

Gegeben: Schrittweite h , Gitter $\mathcal{T} = \{t_0, \dots, t_N\}$ mit $t_i = ih$

Gesucht: M schwach approximierende Pfade $\widetilde{W}(T, \omega_1), \dots, \widetilde{W}(T, \omega_M)$ des Wiener Prozesses.

(1) Für $j = 1, \dots, M$:

(2a) Erzeuge unabhängige zweipunktverteilte Zufallszahlen $\Delta W_i(\omega_j)$ für $i = 0, \dots, N - 1$ mit

$$x_1 = -\sqrt{h}, x_2 = \sqrt{h}.$$

(2b) Erzeuge Gitterfunktionen $\widetilde{W}(t_i, \omega_j)$ mittels der Rekursion

$$\widetilde{W}(t_0, \omega_j) = 0, \quad \widetilde{W}(t_{i+1}, \omega_j) = \widetilde{W}(t_i, \omega_j) + \Delta W_i(\omega_j), i = 0, \dots, N - 1$$

(3) Ende der j -Schleife □

Der große Unterschied in der Konstruktion besteht darin, dass wir hier in jedem Schritt nur endlich viele (nämlich gerade 2) Möglichkeiten für $\Delta W_i(\omega_j)$ haben. Die Anzahl der möglichen Pfade ist demnach endlich, weswegen man hier tatsächlich alle möglichen Pfade mitsamt der Wahrscheinlichkeiten ihres Auftretens berechnet.

Abbildung 6.1 zeigt einen Pfad des Wienerprozesses (durchgezogen) sowie seine schwache Approximation (gestrichelt). Man sieht, dass die Pfade erheblich voneinander abweichen können.

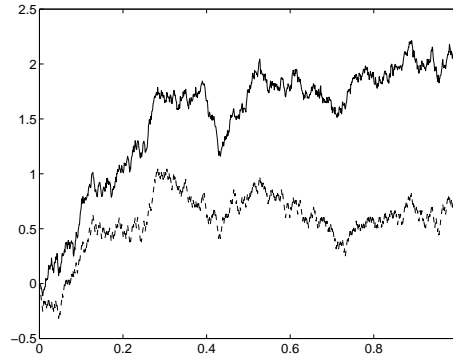


Abbildung 6.1: Schwache Approximation (- -) eines Pfades eines Wiener-Prozesses (—)

Dass dieser Algorithmus trotzdem die richtigen statistischen Eigenschaften besitzt, zeigt das folgende Lemma.

Lemma 6.7 Der Algorithmus 6.6 liefert eine schwache Approximation für den Wiener-Prozess bzgl. $g(x) = x$ und $g(x) = x^2$ für jeden Gitterpunkt $T = t_i$.

Beweis: Für die zweipunktverteilten Zufallszahlen ΔW_i berechnet man

$$E\{\Delta W_i(\cdot)\} = -\sqrt{h}/2 + \sqrt{h}/2 = 0 \quad \text{und} \quad E\{\Delta W_i(\cdot)^2\} = h/2 + h/2 = h.$$

Aus der Definition folgt

$$\widetilde{W}(t_i, \omega) = \sum_{k=0}^{i-1} \Delta W_k(\omega).$$

Also ergibt sich

$$E\{\widetilde{W}(t_i, \cdot)\} = \sum_{k=0}^{i-1} E\{\Delta W_k(\cdot)\} = 0 = E\{W(t_i, \cdot)\}$$

und, indem wir die Unabhängigkeit der Zufallszahlen und (1.6) ausnutzen, auch

$$\begin{aligned} E\{\widetilde{W}(t_i, \cdot)^2\} &= E\left\{\left(\sum_{k=0}^{i-1} \Delta W_k(\cdot)\right)^2\right\} \\ &= E\left\{\sum_{k=0}^{i-1} \Delta W_k(\cdot)^2 + \sum_{k \neq j} \Delta W_k(\cdot) \Delta W_j(\cdot)\right\} \\ &= \sum_{k=1}^i \underbrace{E\{\Delta W_k(\cdot)^2\}}_{=h} + \sum_{k \neq j} \underbrace{E\{\Delta W_k(\cdot)\} E\{\Delta W_j(\cdot)\}}_{=0} \\ &= hi = t_i = E\{W(t_i, \cdot)^2\}, \end{aligned}$$

denn es gilt

$$E\{W(t_i, \cdot)^2\} = E\{W(t_i, \cdot)\}^2 + \underbrace{E\{W(t_i, \cdot)^2\} - E\{W(t_i, \cdot)\}^2}_{=0} = \text{Var}(W(t_i, \cdot)) = t_i.$$

□

6.3 Das stochastische Euler–Verfahren

Das stochastische Euler–Verfahren lässt sich genau wie sein deterministisches Gegenstück heuristisch herleiten:

Für die Lösung zur Zeit $t + h$ gilt die Integralgleichung

$$X(t+h) = X(t) + \int_t^{t+h} a(\tau, X(\tau)) d\tau + \int_t^{t+h} b(\tau, X(\tau)) dW(\tau),$$

für die mit $a(\tau, X(\tau)) \approx a(t, X(t))$, $b(\tau, X(\tau)) \approx b(t, X(t))$ und der Definition des Itô-Integrals (vgl. auch (4.8)) die Approximation

$$\approx X(t) + ha(t, X(t)) + \Delta W(t)b(t, X(t))$$

mit $\Delta W(t) = W(t+h) - W(t)$ gilt.

Dies führt auf das stochastische Euler–Verfahren

$$\Phi(t, X, h, W, \omega) = X(\omega) + ha(t, X(\omega)) + \Delta W(t, \omega)b(t, X(\omega)), \quad (6.1)$$

das auch als Euler-Maruyama-Verfahren bezeichnet wird. Für die geometrische Brownsche Bewegung (4.9) ergibt sich die Verfahrensfunktion Φ gerade zu

$$\Phi(t, X, h, W, \omega) = X(\omega) + h\mu X(\omega) + \Delta W(t, \omega)\sigma X(\omega). \quad (6.2)$$

Dies ist das einfachste Beispiel eines *stochastischen Einschrittverfahrens*, deren Anwendung wir als Algorithmus formulieren wollen.

Algorithmus 6.8 (Lösung einer SDG mit stoch. Einschnitt-Verfahren)

Gegeben: Schrittweite h , Gitter $\mathcal{T} = \{t_0, \dots, t_N\}$ mit $t_i = ih$

Gesucht: M approximative Pfade $\tilde{X}(\cdot, \omega_1), \dots, \tilde{X}(\cdot, \omega_M)$ für die SDG (4.5) auf \mathcal{T} mit $X(t_0, \omega_j) = x_0$.

(0) Erzeuge M approximative Pfade $\tilde{W}(\cdot, \omega_j)$ des Wiener-Prozesses auf \mathcal{T}

(1) Für $j = 1, \dots, M$:

(2) Erzeuge Gitterfunktionen $\tilde{X}(t_i, \omega_j)$ mittels der Rekursion $\tilde{X}(t_0, \omega_j) = x_0$,

$$\tilde{X}(t_{i+1}, \omega_j) = \Phi(t_i, \tilde{X}(t_i, \omega_j), h, \tilde{W}, \omega_j)$$

für $i = 0, \dots, N - 1$

(3) Ende der j -Schleife □

Natürlich kann man die Implementierung noch optimieren, statt z.B. die Pfade $\tilde{W}(\cdot, \omega_j)$ im Voraus zu berechnen, kann man die im Euler-Verfahren benötigten Zufallszahlen ΔW auch direkt an der Stelle per Zufallsgenerator erzeugen, an der man sie benötigt. Der Algorithmus 6.8 soll in erster Linie den prinzipiellen Ablauf einer solchen Simulation verdeutlichen.

Die Konvergenzeigenschaften des stochastischen Euler-Verfahrens fasst der folgende Satz zusammen.

Satz 6.9 Betrachte eine stochastische Differentialgleichung (4.5) mit Funktionen $a(t, x)$ und $b(t, x)$.

(i) Angenommen, es gelten für geeignete Konstanten K_i und alle $t, s \in \mathbb{R}$, $x, y \in \mathbb{R}^n$ die folgenden Ungleichungen

$$\|a(t, x) - a(t, y)\| + \|b(t, x) - b(t, y)\| \leq K_1 \|x - y\| \quad (6.3)$$

$$\|a(t, x)\| + \|b(t, x)\| \leq K_2(1 + \|x\|) \quad (6.4)$$

$$\|a(t, x) - a(s, x)\| + \|b(t, x) - b(s, x)\| \leq K_3(1 + \|x\|)|t - s|^{1/2}. \quad (6.5)$$

Darüberhinaus seien die approximativen Pfade des Wiener-Prozesses in Algorithmus 6.8 die starken Approximationen des Wiener-Prozesses aus Algorithmus 6.5. Dann liefert ist das stochastische Euler-Verfahren eine starke Approximation der Ordnung $\gamma = 1/2$ bzgl. $g(x) = x$.

(ii) Angenommen, es gelten die Ungleichungen aus (i), zudem seien a und b unabhängig von t , von polynomialem Wachstum¹ und C^4 in x . Darüberhinaus seien die approximativen Pfade des Wiener-Prozesses in Algorithmus 6.8 die schwachen Approximationen des Wiener-Prozesses aus Algorithmus 6.6. Dann liefert ist das stochastische Euler-Verfahren eine schwache Approximation der Ordnung $\beta = 1$ bzgl. jeder C^4 -Funktion g mit polynomialem Wachstum, wobei die Konstante C in Definition 6.1 von g abhängt.

¹d.h. es gibt ein Polynom P mit $\|a(x)\| \leq P(\|x\|)$, $\|b(x)\| \leq P(\|x\|)$

Beweis: Der Beweis von (i) findet sich in Kloeden/Platen, Satz 10.2.2, der von (ii) in Kloeden/Platen, Satz 14.5.1. \square

Beachte, dass der Hauptunterschied zwischen (i) und (ii) in der Tatsache liegt, dass in (ii) nur eine schwache Approximation des Wiener-Prozesses verwendet wird.

Ebenso wie bei deterministischen Differentialgleichungen gilt, dass das Euler-Schema relativ langsam konvergiert und dass man bessere Ergebnisse für Schemata höherer Ordnung erhält. Aus Zeitgründen können wir darauf hier nicht eingehen. Eine knappe Einführung in das Thema findet man z.B. in meinem Skript zur Numerik II, eine sehr ausführliche Behandlung des Themas im Buch von Kloeden/Platen.

Bemerkung 6.10 Das Euler-Verfahren lässt sich leicht auf stochastische Differentialgleichungen mit mehreren Wiener-Prozessen verallgemeinern, indem man jedes der zugehörigen Itô-Integrale approximiert und eine entsprechende Anzahl approximativer Wiener-Prozesse verwendet. \square

6.4 Anwendung auf die Optionsbewertung

Aus Satz 6.9(i), Lemma 6.3 und Lemma 6.4 sowie der Lipschitz-Stetigkeit der Funktionen $g(x) = (x - K)^+$ bzw. $g(x) = (K - x)^+$ folgt sofort, dass der für die Optionsbewertung benötigte Erwartungswert durch eine starke Approximation korrekt approximiert wird.

Da wir für die Optionsbewertung aber nur den Erwartungswert $E(g(\tilde{X}(T)))$ für $g(x) = (x - K)^+$ oder $g(x) = (K - x)^+$ benötigen, reicht es im Prinzip aus, hier eine schwache Approximation zu verwenden. Das Problem bei der Anwendung von Satz 6.9(ii) liegt aber darin, dass diese Funktionen g nicht C^4 sind. Da sie einen Knick an der Stelle $x = K$ besitzen, sind sie nämlich noch nicht einmal einfach differenzierbar.

Wir wollen nun beweisen, dass wir für diese nichtdifferenzierbaren g trotzdem Konvergenz erhalten, d.h. dass

$$\lim_{h \rightarrow 0} |E(g(X(T))) - E(g(\tilde{X}(T)))| = 0$$

gilt, wobei wir hier Schrittweiten der Form $h = T/N$ betrachten, damit $t_N = T$ gilt. Dazu führen wir C^4 -Hilfsfunktionen g_ε für $\varepsilon > 0$ ein, die sich nur wenig von g unterscheidet, und für die wir Konvergenz erhalten. Diese verwenden wir allerdings nicht im Algorithmus — dort benutzen wir die “richtigen” g — sondern nur in der Konvergenzanalyse. Wir machen dies beispielhaft für $g(x) = (x - K)^+$, also für einen Call.

Die approximierende Funktion g_ε erhalten wir, indem wir für $\varepsilon > 0$ eine C^4 -Funktion $h_\varepsilon : [K - \varepsilon, K + \varepsilon] \rightarrow \mathbb{R}$ wählen mit den Eigenschaften

$$h_\varepsilon^{(k)}(K - \varepsilon) = 0, k = 0, 1, 2, 3, 4,$$

$$h_\varepsilon(K + \varepsilon) = \varepsilon, h_\varepsilon'(K + \varepsilon) = 1, h_\varepsilon^{(k)}(K + \varepsilon) = 0, k = 2, 3, 4$$

wobei $h_\varepsilon^{(k)}$ die k -te Ableitung bezeichnet. Zudem verlangen wir $h_\varepsilon(x) \in [0, \varepsilon]$ für $x \in [K - \varepsilon, K + \varepsilon]$. Eine solche Funktion existiert und kann z.B. durch Hermite-Interpolation erzeugt werden. Damit definieren wir

$$\tilde{g}_\varepsilon(x) = \begin{cases} g(x), & x \notin [K - \varepsilon, K + \varepsilon] \\ h_\varepsilon(x), & x \in [K - \varepsilon, K + \varepsilon] \end{cases}$$

Abbildung 6.2 zeigt die Originalfunktion g (dicke grüne Linie) und die mit Hermite-Interpolation erzeugte Funktion g_ε (dünne schwarze Linie) für $\varepsilon = 0.2$.

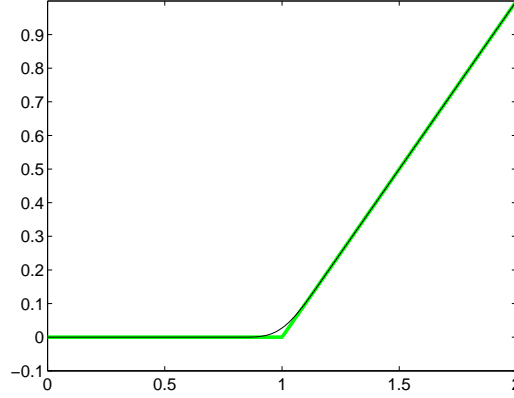


Abbildung 6.2: Die Funktionen $g(x) = (x - K)^+$ (dicke grüne Linie) und $g_\varepsilon(x)$ (dünne schwarze Linie) für $K = 1$ und $\varepsilon = 0.2$

Nun schätzen wir den Fehler mit der Dreiecksungleichung ab

$$\begin{aligned} |E(g(X(T))) - E(g(\tilde{X}(T)))| &\leq |E(g(X(T))) - E(g_\varepsilon(X(T)))| \\ &\quad + |E(g_\varepsilon(X(T))) - E(g_\varepsilon(\tilde{X}(T)))| \\ &\quad + |E(g_\varepsilon(\tilde{X}(T))) - E(g(\tilde{X}(T)))| \end{aligned}$$

Nun gilt einerseits nach Konstruktion $|g_\varepsilon(x) - g(x)| \leq \varepsilon$ für alle x , andererseits unterscheiden sich die beiden Funktionen nur auf einem Intervall der Länge 2ε . Da die Wahrscheinlichkeit, dass $X(T)$ bzw. $\tilde{X}(T)$ in diesem Intervall liegen proportional zu ε ist, folgt

$$|E(g(X(T))) - E(g_\varepsilon(X(T)))| \leq K\varepsilon^2 \quad \text{und} \quad |E(g(\tilde{X}(T))) - E(g_\varepsilon(\tilde{X}(T)))| \leq K\varepsilon^2$$

für eine geeignete Konstante $K > 0$. Den mittleren Term können wir mit Satz 6.9(ii) abschätzen, wobei die Konstante in der Konvergenz aber von ε abhängt, d.h. es gilt

$$|E(g_\varepsilon(X(T))) - E(g_\varepsilon(\tilde{X}(T)))| \leq C_\varepsilon h.$$

Insgesamt erhalten wir also

$$|E(g(X(T))) - E(g(\tilde{X}(T)))| \leq 2K\varepsilon^2 + C_\varepsilon h.$$

Die Konstante C_ε kann dabei für $\varepsilon \rightarrow 0$ möglicherweise gegen unendlich streben. Wir können nun aber zu gegebenem hinreichend kleinem $h > 0$ ein $\varepsilon(h) > 0$ so wählen, dass $C_\varepsilon \leq h^{-1/2}$ gilt. Wählen wir dieses $\varepsilon(h)$ minimal mit dieser Eigenschaft, so folgt $\varepsilon(h) \rightarrow 0$ für $h \rightarrow 0$. Damit erhalten wir

$$|E(g(X(T))) - E(g(\tilde{X}(T)))| \leq 2K\varepsilon(h)^2 + h^{1/2} \rightarrow 0$$

für $h \rightarrow 0$ und damit die gewünschte Konvergenz.

Beachte, dass wir mit diesem Beweis keine Aussage über die Konvergenzordnung treffen können. Dazu müssten wir wissen, wie schnell C_ε für $\varepsilon \rightarrow 0$ wächst, was im Prinzip berechenbar ist, aber nur mit sehr hohem Aufwand, weswegen wir uns hier mit der Konvergenzaussage ohne Abschätzung der Ordnung zufrieden geben.

6.5 Das Binomialmodell als schwache Approximation der geometrischen Brownschen Bewegung

In Abschnitt 3.2 haben wir das Binomialmodell für einen Aktienkurs eingeführt. Bezeichnen wir die Kursvariable in diesem Modell mit S_B , so ist das Modell durch die Iterationsvorschrift

$$S_B(t+1) = S_B(t)Z$$

mit der zweipunktverteilten Zufallsvariablen $Z_t : \Omega \rightarrow \{\alpha_u, \alpha_d\}$ mit

$$\mathbb{P}_{Z_t}(\{\alpha_u\}) = p \quad \text{und} \quad \mathbb{P}_{Z_t}(\{\alpha_d\}) = 1 - p$$

gegeben ist. Wir setzen hier $p = 1/2$.

Betrachten wir nun die Euler-Approximation (6.2) der geometrischen Brownschen Bewegung und verwenden in Algorithmus 6.8 eine schwache Approximation, so ergibt sich die Iterationsvorschrift zu

$$\tilde{X}(t_{i+1}) = \tilde{X}(t_i) + hr\tilde{X}(t_i) + \Delta\tilde{W}(t_i)\sigma\tilde{X}(t_i),$$

wobei $\Delta\tilde{W}(t_i)$ eine zweipunktverteilte Zufallsvariable mit Wert $\pm\sqrt{h}$ ist. Beachte, dass hier, wie in der risikoneutralen Optionsbewertung benötigt, $\mu = r$ gesetzt wurde.

Betrachten wir also das Binomialmodell mit

$$\alpha_d = 1 + hr - \sqrt{h}\sigma \quad \text{und} \quad \alpha_u = 1 + hr + \sqrt{h}\sigma, \quad (6.6)$$

so stimmen die beiden Iterationsvorschriften (bei gleichen Anfangswerten) überein, d.h. es gilt

$$S_B(i) \sim \tilde{X}(t_i).$$

Mit dieser Wahl von α_d und α_u erhält man für $h \rightarrow 0$ aus der Binomialmethode also gerade den Optionswert auf Basis der geometrischen Brownschen Bewegung, der — wie wir im kommenden Kapitel sehen werden — gerade der Wert des Black-Scholes Modells ist. Beachte, dass ein Zeitschritt “1” im Binomialmodell hier gerade einem Zeitschritt “ h ” im Euler-Verfahren entspricht. Soll also der Optionswert zum Zeitpunkt T mit $h = T/N$ mit dem Binomialmodell mit α_u und α_d aus (6.6) berechnet werden, so müssen wir die Binomialmethode mit N Schritten anwenden. Zudem muss die Zinsrate r in der Binomialmethode durch hr ersetzt werden.

Die Wahl (6.6) ist allerdings nicht die einzig mögliche. In Aufgabe 1 des 5. Übungsblatts haben wir gesehen, dass Erwartungswert und Varianz des Binomialmodells mit den Werten der geometrischen Brownschen Bewegung übereinstimmen, wenn wir

$$\alpha_d = e^{hr} - \sqrt{e^{2hr}(e^{h\sigma^2} - 1)} \quad \text{und} \quad \alpha_u = e^{hr} + \sqrt{e^{2hr}(e^{h\sigma^2} - 1)} \quad (6.7)$$

wählen (diese Werte sind gemäß Aufgabe 2 vom 5. Übungsblatt bereits auf einen Zeitschritt der Länge h umgerechnet). Auch mit dieser Wahl der Parameter ist das Binomialmodell eine schwache Approximation der geometrischen Brownschen Bewegung, allerdings eine, die nicht mit dem Euler-Verfahren übereinstimmt. Dies folgt aus der Tatsache, dass man mit Hilfe der Taylor-Entwicklung nachrechnen kann, dass die Erwartungswerte $E(g(S_B(h)))$ für die Parameter aus (6.6) und (6.7) für beliebige C^4 Funktionen g nur um $O(h^2)$ voneinander abweichen, was aufsummiert über das Verfahren eine Abweichung von $O(h)$ ergibt.

Welche der beiden Parameterwahlen genauere Ergebnisse liefert und wie schnell der Wert der Binomialmethode tatsächlich gegen den exakten Black-Scholes-Wert konvergiert, soll in einer Übungsaufgabe herausgefunden werden.

Kapitel 7

Die Black-Scholes Gleichung

Fisher Black¹ und Myron Scholes² sowie Robert Merton³ haben die nach den ersten beiden Autoren benannte partielle Differentialgleichung gleichzeitig in der ersten Hälfte der 1970er Jahre entwickelt. Merton und Scholes bekamen für ihre Arbeiten auf dem Gebiet der Finanzmathematik 1997 den Nobel-Preis für Wirtschaftswissenschaften (Black starb bereits 1995).

Wir wollen hier sowohl die Herleitung dieser Gleichung als auch ihre explizite Lösung angeben. Ziel ist die Berechnung der Funktion $V(t, S)$, die hier als Funktion in den zwei reellen Variablen S und t aufgefasst wird, also $V : \mathbb{R}_0^t \times [0, T] \rightarrow \mathbb{R}$. Wir erinnern daran, dass der Wert $V(T, S)$ dieser Funktion durch (2.1) oder (2.2) gegeben ist.

Wir fassen zunächst die Grundannahmen für die Black-Scholes Gleichung zusammen:

- (a) Es gibt keine Transaktionskosten und Gebühren, ebenso hat das eigene Kaufverhalten keinen Einfluss auf den Markt.
- (b) Der Markt ist Arbitrage-frei.
- (c) Der Kursverlauf ist durch (4.9) beschrieben.
- (d) Die risikofreie Zinsrate r ist konstant und im betrachteten Zeitraum werden keine Dividenden gezahlt.
- (e) Es wird eine risikoneutrale Optionsbewertung durchgeführt.

Beachte, dass wir diese Annahmen bereits bisher — explizit oder implizit — gemacht haben. Annahme (e) werden wir hier allerdings in anderer Form als bisher verwenden, vgl. Bemerkung 7.1(ii). Viele dieser Annahmen können abgeschwächt werden, führen dann aber auf (zum Teil deutlich) kompliziertere Gleichungen.

Die Idee der Black-Scholes Gleichung besteht darin, eine Gleichung für V herleiten, die nicht explizit von den Lösungen $S(t)$ von (4.9) abhängt. Diese Herleitung werden wir nun

¹US-amerikanischer Wirtschaftswissenschaftler, 1938–1995

²kanadischer Wirtschaftswissenschaftler, *1941

³US-amerikanischer Mathematiker und Wirtschaftswissenschaftler, *1944

durchführen wozu wir annehmen, dass $V(t, S)$ eine für $t < T$ zweimal stetig differenzierbare Funktion in t und S ist.

Unter dieser Annahme können wir das Itô-Lemma 4.2 auf $V(t, S(t))$ anwenden. Um die Schreibweise zu verkürzen, lassen wir die Argumente der Funktionen im Folgenden oft weg. Damit erhalten wir

$$dV = \left(\mu S \frac{\partial V}{\partial S} + \frac{\partial V}{\partial t} + \frac{1}{2} \sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} \right) dt + \sigma S \frac{\partial V}{\partial S} dW. \quad (7.1)$$

$V(t, S)$ fassen wir hier als deterministische Funktion in den unabhängigen Variablen S und t auf, W und S hingegen sind stochastische Prozesse.

Nun kommt die risikoneutrale Bewertung ins Spiel: Dazu stellt der Emittent der Option ein Portfolio Π zusammen, das aus einer ausgegebenen (also negativen) Option sowie zeitlich veränderlichen Anteilen des Basiswertes sowie einer festverzinslichen Anleihe besteht, also

$$\Pi(t) = -V(t, S(t)) + \Delta(t)S(t) + \beta(t)B(t). \quad (7.2)$$

Ziel der risikoneutralen Bewertung ist es nun, den Preis $V(t, S)$ so zu bestimmen, dass sich dieses Portfolio bei geeigneter Wahl von $\Delta(t)$ und $\beta(t)$ risikofrei entwickelt, also – gemäß der Arbitragefreiheit – wie eine festverzinsliche Anleihe verhält.

Dieses Portfolio soll dabei selbstfinanzierend sein, d.h. es soll während der Laufzeit nur umgeschichtet werden, es darf also kein Geld hineingesteckt oder herausgenommen werden. Was dies formal bedeutet, macht man sich am Besten für stückweise konstante Werte $\Delta(t)$ und $\beta(t)$ klar. Nehmen wir an, es gilt $\Delta|_{[t_0, t_1)} = \Delta_0$, $\Delta|_{[t_1, t_2)} = \Delta_1$, $\beta|_{[t_0, t_1)} = \beta_0$ und $\beta|_{[t_1, t_2)} = \beta_1$ für Zeiten $t_0 < t_1 < t_2$. Zum Zeitpunkt t_1 wird das Portfolio also umgeschichtet.

Dann beträgt der Wert des Portfolios unmittelbar vor der Umschichtung

$$\Pi_v(t_1) = -V(t_1, S(t_1)) + \Delta_0 S(t_1) + \beta_0 B(t_1)$$

und nach der Umschichtung

$$\Pi_n(t_1) = -V(t_1, S(t_1)) + \Delta_1 S(t_1) + \beta_1 B(t_1).$$

Das Portfolio ist nun gerade dann selbstfinanzierend, wenn diese beiden Werte übereinstimmen, also wenn

$$\Delta_0 S(t_1) + \beta_0 B(t_1) = \Delta_1 S(t_1) + \beta_1 B(t_1)$$

gilt. Diese Bedingung können wir auch anders ausdrücken: Betrachten wir den Wert

$$\Pi(t_2) = -V(t_2, S(t_2)) + \Delta_1 S(t_2) + \beta_1 B(t_2)$$

so ist das Portfolio genau dann selbstfinanzierend, wenn

$$\begin{aligned} \Pi(t_2) &= \Pi(t_0) + \Delta_0 S(t_1) + \beta_0 B(t_1) - \Delta_0 S(t_0) - \beta_0 B(t_0) + V(t_0, S(t_0)) \\ &\quad + \Delta_1 S(t_2) + \beta_1 B(t_2) - \Delta_1 S(t_1) - \beta_1 B(t_1) - V(t_2, S(t_2)) \end{aligned}$$

gilt. Für $p + 1$ Zeitpunkte t_0, \dots, t_p ergibt sich die Bedingung dann zu

$$\begin{aligned} \Pi(t_p) &= \Pi(t_0) + V(t_0, S(t_0)) - V(t_p, S(t_p)) \\ &\quad + \sum_{k=0}^{p-1} \Delta(t_k)(S(t_{k+1}) - S(t_k)) + \sum_{k=0}^{p-1} \beta(t_k)(B(t_{k+1}) - B(t_k)). \end{aligned}$$

Für kontinuierliche Umschichtung erhalten wir durch Grenzübergang $t_{k+1} - t_k \rightarrow 0$ daraus die Bedingung

$$\Pi(t) = \Pi(t_0) + V(t_0, S(t_0)) - V(t, S(t)) + \int_{t_0}^t \Delta(\tau) dS(\tau) + \int_{t_0}^t \beta(\tau) dB(\tau),$$

wobei die Integrale im Itô-Sinne mit S bzw. B an Stelle von W zu verstehen sind. Schreiben wir dies wieder in unserer symbolischen SDG-Schreibweise und nutzen aus, dass $dB = rBdt$ gilt, so erhalten wir die Gleichung

$$d\Pi(t) = -dV(t, S(t)) + \Delta(t)dS(t) + \beta(t)rB(t)dt. \quad (7.3)$$

Setzen wir die Gleichungen (4.9) und (7.1) in diese Gleichung ein, so erhalten wir

$$d\Pi = - \left(\mu S \left(\frac{\partial V}{\partial S} - \Delta \right) + \frac{\partial V}{\partial t} + \frac{1}{2} \sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} - \beta r B \right) dt - \sigma S \left(\frac{\partial V}{\partial S} - \Delta \right) dW.$$

Damit sich das Portfolio risikofrei entwickeln kann, müssen zunächst die stochastischen Terme in $d\Pi$ eliminiert werden. Dies erreichen wir, indem wir

$$\Delta(t) = \frac{\partial V}{\partial S}(t, S(t))$$

setzen. Damit erhalten wir

$$d\Pi = - \left(\frac{\partial V}{\partial t} + \frac{1}{2} \sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} - \beta r B \right) dt \quad (7.4)$$

also eine deterministische Differentialgleichung.

Schließlich verwenden wir die Arbitrage-Freiheit, um V zu bestimmen. Wir suchen einen Wert von V , für den sich Π risikofrei entwickelt. Wenn ein solcher Wert existiert, muss sich Π nach der Annahme der Arbitrage-Freiheit wie eine festverzinsliche Anleihe entwickeln, es muss also $\Pi(t) = e^{rt}\Pi(0)$ gelten. Daraus folgt die Differentialgleichung

$$d\Pi = r\Pi dt,$$

woraus mit (7.2) und $\Delta(t) = \frac{\partial V}{\partial S}(t, S(t))$ die Gleichung

$$d\Pi = r\Pi dt = \left(-rV + rS \frac{\partial V}{\partial S} + r\beta B \right) dt, \quad (7.5)$$

folgt.

Nun haben wir zwei Gleichungen für $d\Pi$, nämlich gerade (7.4) und (7.5). Gleichsetzen dieser Gleichungen liefert

$$\frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} - rV + rS \frac{\partial V}{\partial S} = 0. \quad (7.6)$$

Dies ist gerade die Black-Scholes Gleichung.

Die Black-Scholes Gleichung (7.6) ist also eine partielle Differentialgleichung, die von V erfüllt wird. Die Gleichung (7.6) hat (genau wie eine gewöhnliche Differentialgleichung ohne Anfangsbedingung) unendlich viele Lösungen, wenn wir aber geeignete *Randwerte* $V(T, S)$, für ein $T > 0$ und alle $S \in [0, \infty)$ festlegen, so erhalten wir eine eindeutige Lösung. Dies ist hier einfach, da sich die Randwerte in natürlicher Weise aus den Optionswerten zum Fälligkeitszeitpunkt $t = T$ ergeben.

Tatsächlich ist Gleichung (7.6) explizit lösbar. Für eine europäische Call-Option gilt

$$V_{Ce}(t, S) = SP(a) - Ke^{-r(T-t)}P(b),$$

während sich für die europäische Put-Option

$$V_{Pe}(t, S) = S(P(a) - 1) - Ke^{-r(T-t)}(P(b) - 1)$$

ergibt. Hierbei sind in beiden Fällen

$$a = \frac{\log \frac{S}{K} + \left(r + \frac{\sigma^2}{2}\right)(T-t)}{\sigma\sqrt{T-t}}, \quad b = a - \sigma\sqrt{T-t}$$

und die Funktion P ist die Verteilungsfunktion der Standard-Normalverteilung gegeben durch

$$P(y) = \int_{-\infty}^y \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right) dx.$$

Abbildung 7.1 zeigt Lösungen für den europäischen Call mit $r = 0.06$, $K = 1$, $T = 1$ und $\sigma = 0.3$ bzw. $\sigma = 1$.

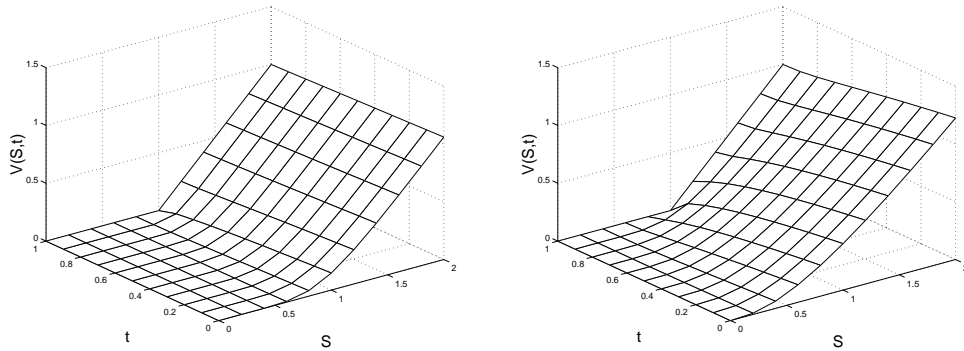


Abbildung 7.1: Lösung der Black-Scholes Gleichung für den europäischen Call mit $r = 0.06$, $K = 1$, $T = 1$ und $\sigma = 0.3$ (links) bzw. $\sigma = 1$ (rechts)

Wenn man die vereinfachenden Annahmen abschwächt oder wenn man amerikanische an Stelle von europäischen Optionen betrachtet, so erhält man deutlich kompliziertere partielle Differentialgleichungen von ähnlichem Typ wie (7.6), die dann nur noch numerisch gelöst werden können.

Bemerkung 7.1 (i) Die Annahme der risikoneutralen Bewertung wurde hier in die Modellannahme umgesetzt, dass der Emittent der Option sein Risiko durch das Vorhalten des selbstfinanzierenden Portfolios Π ausschalten kann. Diese Annahme ist wesentlich für den resultierenden Optionspreis. Sie bedeutet aber auch, dass der Emittent dazu sein Portfolio stets geeignet umschichten muss; tut er dies nicht, bleibt ein Risiko erhalten. Dieses Ausschalten des Risikos durch Umschichten nennt man “Hedging”. Dies bedeutet insbesondere, dass die Black-Scholes Gleichung tatsächlich viel mehr als nur den Optionspreis liefert: Sie liefert auch eine Regel für die risikoneutrale Anlagestrategie (auch “Hedging-Strategie” genannt) des Emittenten. Wir werden auf das Hedging im folgenden Kapitel näher eingehen.

(ii) Beachte, dass wir hier weder die Abzinsungsformel aus Algorithmus 2.1(3) noch die Wachstumsbedingung (2.5) an S verwendet haben. Insbesondere gilt der durch die Black-Scholes Gleichung ermittelte Wert tatsächlich für Basiswerte mit beliebigem Trend μ , da der Effekt des Trends durch das Hedging eliminiert wird.

Tatsächlich folgt die Abzinsungsformel in Algorithmus 2.1(3) (und damit auch (2.5)) aus der hier rein ökonomisch interpretierten Risikoneutralität, vgl. [3, Chapter 12]. \square

Kapitel 8

Hedging und die “Griechen”

Die Black-Scholes Gleichung liefert nicht nur eine — im Fall europäischer Optionen sogar explizit lösbare — Gleichung für den Optionswert. Aus den Annahmen in ihrer Herleitung ergibt sich darüberhinaus eine Regel, wie sich eine Emittentin gegen das durch die Ausgabe der Option entstehende Risiko absichern kann. Dies geschieht nämlich gerade durch das in der Herleitung konstruierte Portfolio mit den aus der Herleitung folgenden Umschichtungen Δ und β . Dieses Vorgehen nennt man *Hedging*.

Kontrolltheoretisch kann man Δ und β als Kontrollfunktionen betrachten, mit denen man das Portfolio Π steuert. In diesem Sinne ist die Black-Scholes Strategie

$$\Delta = \Delta(t, S(t)) = \frac{\partial V}{\partial S}(t, S(t))$$

dann ein (zeitabhängiges) Feedback, d.h. der gewünschte aktuelle Anteil des Basiswertes S im Portfolio hängt von der aktuellen Zeit t und vom aktuellen Kurs $S(t)$ ab. Man kann nachrechnen, dass

$$\Delta(t, S(t)) = P(a) \text{ (europ. Call)} \quad \text{und} \quad \Delta(t, S(t)) = P(a) - 1 \text{ (europ. Put)} \quad (8.1)$$

gilt mit den Bezeichnungen aus dem vorhergehenden Kapitel.

Die zweite Größe β ergibt sich dann aus der Selbstfinanzierungs-Bedingung und könnte daraus errechnet werden. Für die praktische Implementierung des Hedging werden wir im Folgenden aber etwas anders vorgehen, so dass wir die exakte Formel für β gar nicht benötigen.

8.1 Zeitdiskretes Hedging

Das kontinuierliche Umschichten des Portfolios Π ist in der Praxis nicht zu bewerkstelligen. Dies würde nämlich bedeuten, die Gewichte Δ und β zu jedem Zeitpunkt $t \in [0, T]$ an die aktuellen Kurse anzupassen.

Realistisch ist es hingegen nur möglich, die Gewichte zu festen diskreten Zeitpunkten $t_i \in [0, T]$ anzupassen, d.h. das Portfolio zu diesen Zeiten umzuschichten. Wir betrachten hier

den einfachsten Fall, nämlich dass das Portfolio alle τ Zeiteinheiten umgeschichtet wird, d.h. zu den Zeiten $t_i = i\tau$, $i = 1, \dots, N$, wobei $\tau > 0$ so gewählt ist, dass $N\tau = T$ gilt.

Das solchermaßen zeitdiskret gehedgte Portfolio ergibt sich damit für $t \in [t_i, t_{i+1}]$ zu

$$\Pi(t) = -V(t, S(t)) + \Delta(t_i, S(t_i))S(t) + \beta(t_i)B(t).$$

In der Sprache der Kontroll- bzw. Regelungstheorie werden Δ und β als sogenanntes *Abtastfeedback* implementiert. Zu jedem Zeitpunkt t_i , $i = 1, \dots, N - 1$ werden die Anteile dann umgeschichtet, indem $\Delta(t_{i-1}, S(t_{i-1}))$ durch $\Delta(t_i, S(t_i))$ und $\beta(t_{i-1})$ durch $\beta(t_i)$ ersetzt wird. Das $\Delta(t_i, S(t_i))$ wird dabei gemäß der Zeit t_i und dem Kurs $S(t_i)$ gemäß (8.1) berechnet und das $\beta(t_i)$ wird so gewählt, dass die Umschichtung selbstfinanzierend, also kostenneutral erfolgt.

Wir berechnen nun eine Formel für dieses $\beta(t_i)$. Zum Zeitpunkt t_i besitzt das Portfolio vor der Umschichtung den Wert

$$\Pi(t_i) = -V(t_i, S(t_i)) + \Delta(t_{i-1}, S(t_{i-1}))S(t_i) + \beta(t_{i-1})B(t_i). \quad (8.2)$$

Zum Zeitpunkt t_i wird nun der Anteil des Basiswertes auf $\Delta(t_i, S(t_i))S(t_i)$ verändert. Das Portfolio ist bei dieser Umschichtung genau dann selbstfinanzierend, wenn sich der Wert des Portfolios dabei nicht ändert. Mit $\Pi(t_i)$ aus (8.2) muss also nach der Umschichtung

$$\Pi(t_i) = -V(t_i, S(t_i)) + \Delta(t_i, S(t_i))S(t_i) + \beta(t_i)B(t_i) \quad (8.3)$$

gelten. Aus (8.3) lässt sich $\beta(t_i)$ nun eindeutig bestimmen. Es gilt

$$\beta(t_i) = \frac{\Pi(t_i) + V(t_i, S(t_i)) - \Delta(t_i, S(t_i))S(t_i)}{B(t_i)}.$$

Dies führt zu dem folgenden Algorithmus, in dem ein mittels der Black-Scholes Formeln gehedgtes Portfolio numerisch simuliert wird.

Algorithmus 8.1 (Simulation des Black-Scholes Portfolios mit zeitdiskretem Hedging)

Gegeben: $r, \sigma, T, S_0, B_0, \beta_0 > 0$ und ein Zeitgitter $\mathcal{T} = \{t_0, \dots, t_N\}$ mit $t_i = i\tau$, $\tau > 0$ und $t_N = T$.

- (1) Berechne Kurswerte $S(t_i)$, $i = 0, \dots, N$ mit $S(t_0) = S_0$ mittels der geometrischen Brownschen Bewegung auf Basis eines Pfades der starken Approximation des Wiener-Prozesses auf \mathcal{T}
- (2) Berechne $V(t_i, S(t_i))$, $\Delta(t_i, S(t_i))$ und $B(t_i) = e^{rt_i}B_0$, $i = 0, \dots, N$
- (3) Berechne das Portfolio induktiv mittels $\beta(0) := \beta_0$,

$$\Pi(t_0) := -V(t_0, S(t_0)) + \Delta(t_0, S(t_0))S(t_0) + \beta(t_0)B(t_0)$$

und

$$\Pi(t_i) := -V(t_i, S(t_i)) + \Delta(t_{i-1}, S(t_{i-1}))S(t_i) + \beta(t_{i-1})B(t_i),$$

$$\beta(t_i) := \frac{\Pi(t_i) + V(t_i, S(t_i)) - \Delta(t_i, S(t_i))S(t_i)}{B(t_i)}$$

für $i = 1, \dots, N$

□

Abbildung 8.1 zeigt Ergebnisse dieser Simulation für einen europäischen Call mit den Parametern $S_0 = B_0 = 80$, $T = 10$, $r = 0.05$, $\sigma = 0.3$, $K = 100$, $\beta_0 = 1$ sowie $\tau = 1, 0.1$ und 0.01 . Wie man sieht, erfüllt das Ergebnis voll und ganz die Erwartungen: Das Portfolio folgt fast genau der deterministischen Entwicklung $e^{rt}\Pi(0)$, die zum Vergleich gestrichelt eingezeichnet ist, und zwar um so genauer je kleiner τ ist.

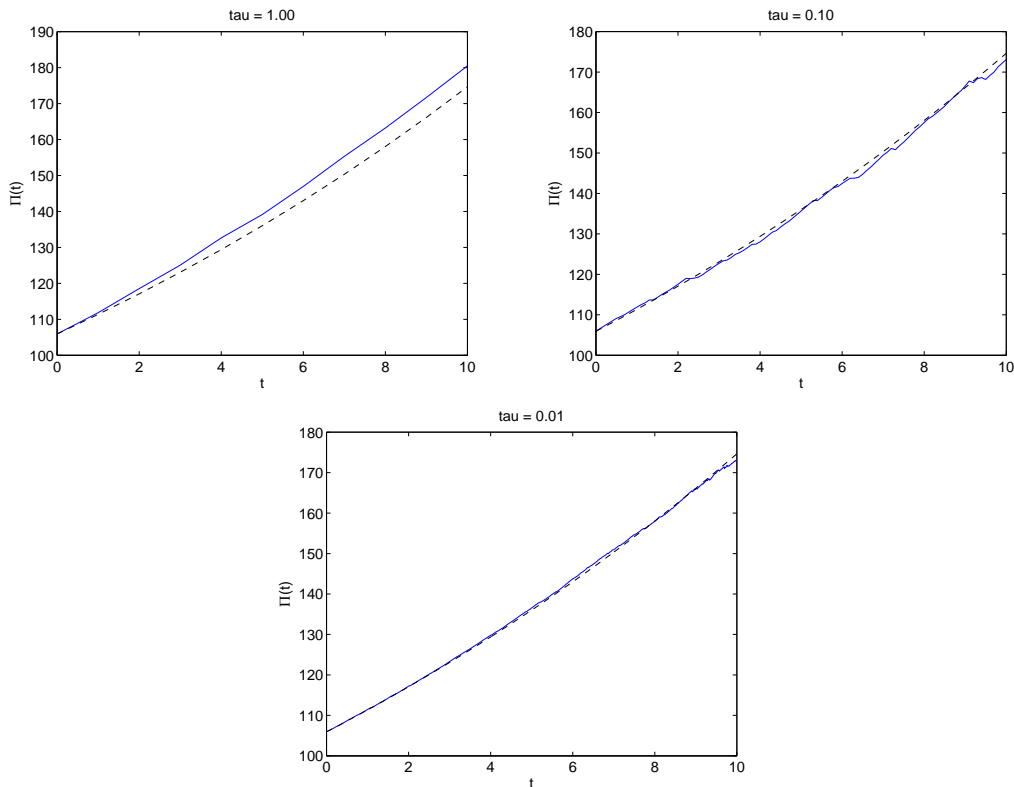


Abbildung 8.1: Simulierte Entwicklung des Portfolios Π bei diskretem Hedging mit $\tau = 1, 0.1, 0.01$ (blau durchgezogen). Zum Vergleich ist die theoretisch erwartete Entwicklung $e^{rt}\Pi(0)$ schwarz gestrichelt eingezeichnet.

Eine ganz wesentliche Annahme dafür, dass das Hedging so gut funktioniert, ist die Korrektheit (im stochastischen Sinne) des zu Grunde liegenden Kursmodells. Die Hedging-Strategie ist nämlich darauf angewiesen, dass sich der Kurs tatsächlich gemäß des Modells verhält, da sie durch die Herleitung aus den Modellannahmen implizit auf einer stochastischen Vorhersage des zukünftigen Verhaltens basiert.

Verhält sich der Kurs nicht entsprechend den Vorhersagen, so bricht die deterministische Entwicklung des Portfolios Π zusammen. Als Beispiel betrachten wir noch einmal die obige Simulation, in der der simulierte Kurswert mittels der Skalierung $S(t_i) \rightarrow S(t_i)/2$ für $t_i \geq 5$ verändert wurde. Dies entspricht einem ‘Kurssturz’ zur Zeit $t = 5$, bei dem sich der Kurs halbiert. Nach dem Sturz verläuft der Kurs wieder gemäß der geometrischen Brownschen

Bewegung, allerdings nun auf niedrigerem Niveau. Abbildung 8.2 zeigt, was dabei mit dem Portfolio Π passiert: der Wert bricht zur Zeit $T/2$ deutlich ein. Er wächst danach (wenn das Modell wieder stimmt) zwar wieder deterministisch, allerdings auf niedrigerem Niveau. Die Auswirkungen dieses Verhaltens auf die Bilanzen von Banken, Fondsgesellschaften etc. in Zeiten der Finanzkrise brauchen sicherlich nicht weiter kommentiert werden.

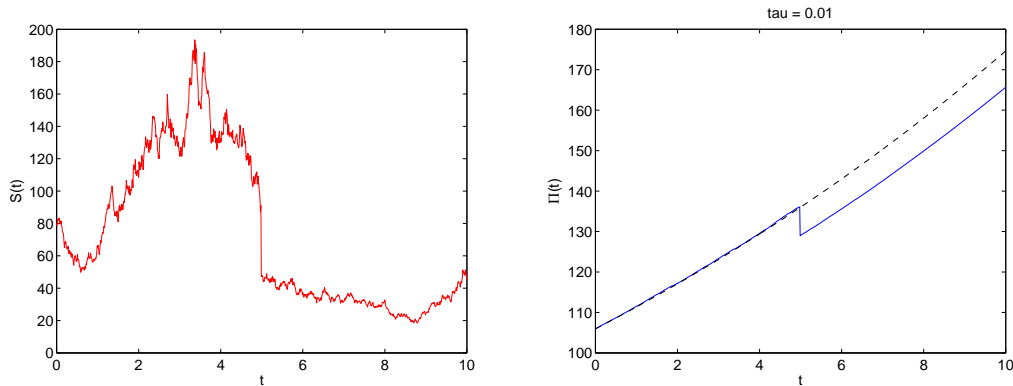


Abbildung 8.2: Simulierte Entwicklung des Kurses S (links) und des Portfolios Π (rechts) bei diskrettem Hedging mit $\tau = 0.01$ und simuliertem “Kurssturz” zur Zeit $t = 5$ (blau durchgezogen). Zum Vergleich ist die theoretisch erwartete Entwicklung $e^{rt}\Pi(0)$ schwarz gestrichelt eingezeichnet.

Mit diesem Kurssturz hängt das genaue Verhalten von Π nun wieder vom Zufall ab. Je nach dem sonstigen Verlauf des Kurses $S(t_i)$ wirkt sich der Sturz mehr oder weniger schlimm aus. Dieses Phänomen soll — ebenso wie die Frage was bei einem unerwarteten Kursanstieg oder einem vorübergehenden Kurssturz passiert — ebenfalls auf dem 10. Übungsblatt numerisch untersucht werden.

Bemerkung 8.2 Die Ableitung $\Delta = \frac{\partial V}{\partial S}$ spielt offenbar eine bedeutende Rolle beim Hedging. Neben diesem “Delta” werden auch alle anderen Ableitungen des Optionswerts mit griechischen Buchstaben versehen. Es bezeichnen

$$\begin{aligned} \Gamma &= \frac{\partial^2 V}{\partial S^2} \quad (\text{“Gamma”}) & \rho &= \frac{\partial V}{\partial r} \quad (\text{“Rho”}) \\ \Theta &= \frac{\partial V}{\partial t} \quad (\text{“Theta”}) & \text{vega} &= \frac{\partial V}{\partial \sigma} \quad (\text{“Vega”}). \end{aligned}$$

Im Englischen werden diese Werte als “Greeks”, also “Griechen” bezeichnet, im deutschen existiert dafür auch der Begriff “Kennzahlen”.

Neben dem für das praktische Hedging besonders wichtigen Δ geben auch die anderen “Griechen” gewisse Einsichten. Eine konkrete Berechnung der Werte für den Call (siehe Higham, Chapter 10) ergibt z.B. $\rho > 0$ und $\Theta < 0$, d.h. der Optionswert steigt, wenn die risikofreie Zinsrate steigt und der Optionswert fällt mit der Zeit. \square

8.2 Monte-Carlo für das Δ

Wenn wir komplexere Kursmodelle als die geometrische Brownsche Bewegung verwenden, ist die Monte-Carlo Simulation oft die einzige Methode, mit denen eine Optionsbewertung numerisch durchgeführt werden kann. Ebenso gibt es eine Reihe sogenannter pfadabhängiger Optionsarten, für die die Monte-Carlo Methode gut geeignet ist. Ein Beispiel ist die sogenannte asiatische Option, bei der die Auszahlung (im Falle eines Call) am Ende der Laufzeit durch

$$V(T, S(\cdot)) = \max\left(\frac{1}{T} \int_0^T S(\tau) d\tau - E, 0\right)$$

gegeben ist. Der Wert der Option hängt also vom Durchschnittswert und nicht nur vom Endwert des Basiskurses im Intervall $[0, T]$ ab.

Approximiert man das Integral durch eine Quadraturformel (z.B. durch die zusammengesetzte Trapez-Regel) mit Stützstellen $t_i = hi$, $i = 0, \dots, N$, $hN = T$, und erweitert man die Monte-Carlo-Methode so, dass nicht nur wie in Kapitel 5 beschrieben der Wert $S(T)$ sondern die Werte $S(t_i)$, $i = 0, \dots, N$ berechnet werden (dies kann man analog zur Aufgabe 2 auf dem 7. Übungsblatt machen), so kann man den Wert jedes Pfades errechnen. Bezeichnen wir den mit $e^{r(t-T)}$ abdiskontierten Wert des i -ten simulierten Pfades zur Zeit t mit $V(t, S, \omega_i)$, so kann man den Monte-Carlo-Wert der Option als

$$V_{Mc}(t, S) = \frac{1}{n} \sum_{i=1}^n V(t, S, \omega_i)$$

schreiben.

Unabhängig davon, welche Option man bewertet und welches Kursmodell man verwendet, benötigt man auch bei der Verwendung von Monte-Carlo-Methoden nun das Delta $\Delta = \frac{\partial V}{\partial S}$, um die Hedging-Strategie zu bestimmen. Zur Berechnung von Δ verwenden wir die aus der Taylor-Entwicklung stammende Approximation

$$\Delta(t, S) = \frac{\partial V}{\partial S}(t, S) = \frac{V(t, S + \Delta S) - V(t, S)}{\Delta S} + O(\Delta S)$$

für kleine $\Delta S \in \mathbb{R}$.

Bezeichnen wir die in der Monte-Carlo-Simulation erhaltenen Pfadwerte Eine naheliegende Anwendung dieser Formel wäre nun, das Δ mittels

$$\Delta(t, S) \approx \frac{V_{Mc}(t, S + \Delta S) - V_{Mc}(t, S)}{\Delta S} \quad (8.4)$$

zu approximieren. Da der Fehler in V_{Mc} von der Ordnung $O(\sqrt{n})$ ist und durch ΔS geteilt wird, beträgt der Fehler der Approximation (8.4) also

$$O(\sqrt{n}/\Delta S) + O(\Delta S).$$

Der sowieso schon recht große Fehler der Monte-Carlo-Simulation verstärkt sich also durch das Bilden des Differenzenquotienten noch weiter.

Obwohl sich dieses grundsätzliche Problem nicht vollständig beheben lässt, kann man eine bessere Approximation erhalten, wenn man das Δ pfadweise berechnet. Dazu setzt man zunächst

$$\Delta(t, S, \omega_i) := \frac{V(t, S + \Delta S, \omega_i) - V(t, S, \omega_i)}{\Delta S},$$

wobei gleiches ω_i bedeutet, dass wir die gleichen Pfade des Wiener-Prozesses verwenden. Dann setzen wir

$$\Delta_{Mc}(t, S) := \frac{1}{n} \sum_{i=1}^n \Delta(t, S, \omega_i).$$

Der Grund dafür, warum diese zweite Variante bessere Werte liefert, liegt in der niedrigeren Varianz der aufsummierten Zufallszahlen. Verwendet man die gleichen Pfade des Wiener-Prozesses zur Berechnung von $V(t, S + \Delta S)$ und von $V(t, S)$, so variieren die Werte der Differenzen deutlich weniger, als wenn man — wie implizit in (8.4) — unterschiedliche Pfade verwendet. Ganz analog zur Diskussion am Ende von Abschnitt 5.2 erhalten wir auch hier bei niedrigerer Varianz eine bessere Konvergenz der Monte-Carlo-Methode.

Kapitel 9

Einführung in die Finite Differenzen-Methode

Während die Black-Scholes Gleichung für europäische Optionen analytisch gelöst werden kann, ist dies für viele weitere Optionsarten nicht möglich.

Wenn andere Optionen als europäische Calls oder Puts betrachtet werden, kann es sein, dass die zugehörige Black-Scholes-Gleichung entsprechend verändert werden muss. Dies ist z.B. der Fall für die bereits erwähnten asiatischen Optionen oder für sogenannte Basket-Optionen, in denen die Option von mehreren Basiswerten abhängt, vgl. Günther und Jüngel, Abschnitt 6.1 bzw. Abschnitt 6.5. Im Falle amerikanischer Optionen ändert sich dabei sogar der Typ der zu lösenden Gleichung, vgl. Günther und Jüngel, Kapitel 7.

Es kann aber auch sein, dass sich die Black-Scholes Gleichung selbst gar nicht ändert, sondern dass nur andere Randbedingungen eingesetzt werden müssen. Dies ist z.B. der Fall für die sogenannten (zugegebenermaßen relativ exotischen) Power-Optionen, in denen die Auszahlungsfunktion z.B. (im Falle eines sogenannten europäischen capped symmetrischen Power-Calls) durch $V(T, S) = \min\{L, ((S - K)^+)^p\}$ für Parameter $L, p > 0$ gegeben ist. Hier bleibt die Gleichung bestehen und es ändern sich nur die Randbedingungen. Ähnlich ist dies bei sogenannten Barriere-Optionen, bei denen die Auszahlung nur dann fällig wird, wenn der Basiswert während der Laufzeit eine gegebene Schranke B unterschreitet (down-and-in) oder nicht unterschreitet (down-and-out), die sich im europäischen Fall allerdings wieder analytisch lösen lassen.

In diesem Kapitel werden wir ein numerisches Verfahren kennenlernen, mit dem sich die Black-Scholes-Gleichung numerisch lösen lässt, die sogenannte Finite Differenzen Methode. Allerdings werden wir das Verfahren nicht direkt für diese Gleichung sondern für eine einfachere partielle Differentialgleichung, die sogenannte Wärmeleitungsgleichung einführen. Ein Grund dafür ist, dass diese Gleichung viel einfacher strukturiert ist, weswegen sich das Verfahren daran leichter erklären lässt.

9.1 Black-Scholes- und Wärmeleitungsgleichung

Der zweite Grund für die Betrachtung der Wärmeleitungsgleichung ist, dass wir diese in gewissem Sinne “o.B.d.A.” verwenden können, da sich die Black-Scholes-Gleichung durch eine Variablentransformation in diese Gleichung umformen lässt. Diese Transformation wollen wir zunächst erläutern.

Wir verwenden in diesem Kapitel für Ableitungen die übliche Notation aus Theorie und Numerik partieller Differentialgleichungen, nämlich z.B.

$$\frac{\partial V}{\partial t} = V_t \quad \text{oder} \quad \frac{\partial^2 V}{\partial S^2} = V_{SS}.$$

Die Black-Scholes-Gleichung (7.6) schreibt sich damit als

$$V_t(t, S) + \frac{1}{2}\sigma^2 S^2 V_{SS}(t, S) - rV + rSV_S(t, S) = 0,$$

wobei $V : [0, T] \times \mathbb{R}_{\geq 0} \rightarrow \mathbb{R}$ die unbekannte Funktion ist, welche die Randbedingung (wir betrachten hier beispielhaft einen Call)

$$V(T, S) = (S - K)^+$$

zur Zeit $t = T$ für alle $S \in [0, \infty)$ erfüllt.

Die Wärmeleitungsgleichung hingegen lautet einfach

$$u_t(t, x) - u_{xx}(t, x) = 0 \tag{9.1}$$

für die unbekannte Funktion $u(t, x)$ mit $u : \mathbb{R}_{\geq 0} \times \mathbb{R} \rightarrow \mathbb{R}$. Die Randbedingung ist hierbei typischerweise zur Zeit $T = 0$ als $u(0, x) = u_0(x)$ für alle $x \in \mathbb{R}$ und eine vorgegebene Funktion $u_0 : \mathbb{R} \rightarrow \mathbb{R}$. Die Gleichung beschreibt die Temperatur $u(t, x)$ zur Zeit t im Punkt x in einem (näherungsweise) eindimensionalen Stab unendlicher Länge.

Die Transformation der Black-Scholes-Gleichung in die Wärmeleitungsgleichung erfolgt nun in zwei Schritten. Zunächst definieren wir die Variablentransformationen

$$x = \ln(S/K) \Leftrightarrow S = Ke^x, \quad \tau = \sigma^2(T - t)/2 \Leftrightarrow t = T - 2\tau/\sigma^2$$

und setzen

$$v(\tau, x) := V(t, S)/K.$$

Beachte, dass diese Transformation das Kursintervall $S \in (0, \infty)$ auf das Intervall $x \in (-\infty, \infty)$ abbildet und die Zeit $t = T$ auf $\tau = 0$, so dass insbesondere die Randbedingung für V in $t = T$ nun eine Randbedingung für v in $\tau = 0$ wird.

Man rechnet nach, dass v die Gleichung

$$v_\tau(\tau, x) - v_{xx}(\tau, x) + (1 - \kappa)v_x(\tau, x) + \kappa v(\tau, x) = 0$$

für alle $x \in \mathbb{R}$ und alle $\tau \in [0, T_0]$ mit $\kappa = 2r/\sigma^2$ und $T_0 = \sigma^2 T/2$ erfüllt. Zudem gilt die (Anfangs-)Randbedingung

$$v(0, x) = (e^x - 1)^+.$$

Setzen wir nun

$$u(\tau, x) = \exp\left(\frac{1}{2}(\kappa - 1)x + \frac{1}{4}(\kappa + 1)^2\tau\right)v(\tau, x),$$

so rechnet man nach, dass u die Wärmeleitungsgleichung

$$u_\tau(\tau, x) - u_{xx}(\tau, x) = 0$$

mit der Randbedingung

$$u(0, x) = u_0(x) = \left(e^{(\kappa+1)x/2} - e^{(\kappa-1)x/2}\right)^+$$

erfüllt.

Die Rücktransformation lautet

$$V(t, S) = K \exp\left(-\frac{1}{2}(\kappa - 1)x - \frac{1}{4}(\kappa + 1)^2\tau\right)u(\tau, x)$$

mit den oben definierten Beziehungen zwischen t und τ bzw. S und x .

Wenn wir also die Wärmeleitungsgleichung numerisch mit der transformierten Black-Scholes Randbedingung zur Zeit $\tau = 0$ auf dem Zeitintervall $[0, T_0]$ lösen, so können wir durch Rücktransformation die (approximative) Lösung der Black-Scholes-Gleichung erhalten.

9.2 Grundidee der Finiten Differenzen

Wir beschreiben nun die Methode der finiten Differenzen für die Wärmeleitungsgleichung (9.1). Dabei bezeichnen wir die Zeitvariable wieder mit t (statt τ) und das Zeitintervall mit $[0, T]$ (statt $[0, T_0]$).

Die Lösung $u(t, x)$ der Wärmeleitungsgleichung (9.1) ist durch Ihre Ableitungen an unendlich vielen Punkten (t, x) charakterisiert. Die Methode der finiten Differenzen besteht nun darin, diese unendlich vielen Ableitungen durch eine endliche (finite) Anzahl von Differenzenquotienten (Differenzen) zu ersetzen und so eine approximative Lösung der Gleichung zu erhalten.

Die verwendeten Differenzenquotienten sind dabei für $h, s > 0$

$$u_{xx}(t, x) \approx \frac{1}{h^2}(u(t, x + h) - 2u(t, x) + u(t, x - h))$$

für die 2. Ableitung nach x und entweder die Vorwärtsdifferenz

$$u_t(t, x) \approx \frac{1}{s}(u(t + s, x) - u(t, x))$$

oder die Rückwärtsdifferenz

$$u_t(t, x) \approx \frac{1}{s}(u(t, x) - u(t - s, x))$$

für die 1. Ableitung nach t .

Hierfür gilt das folgende Lemma. Darin (und allgemein in diesem Kapitel) verwenden wir das Landau-Symbol O im folgenden Sinne: Für einen beliebigen reellen Ausdruck $y(t, x, z)$ mit $t, x, z \in \mathbb{R}$, $t, z > 0$, schreiben wir $y(t, x) = O(z^p)$, falls für jede kompakte Menge $K \subset [0, \infty) \times \mathbb{R}$ eine Konstante $C_K > 0$ existiert mit $|y(t, x, z)| \leq Cz^p$ für alle hinreichend kleinen z und alle $(t, x) \in K$.

Lemma 9.1 Für $u \in C^4$ $h, s > 0$ gilt

$$u_{xx}(t, x) = \frac{1}{h^2}(u(t, x+h) - 2u(t, x) + u(t, x-h)) + O(h^2),$$

$$u_t(t, x) = \frac{1}{s}(u(t+s, x) - u(t, x)) + O(s)$$

und

$$u_t(t, x) = \frac{1}{s}(u(t, x) - u(t-s, x)) + O(s).$$

Beweis: Nach der Taylor-Entwicklung gelten die beiden Gleichungen

$$u(t, x+h) = u(t, x) + u_x(t, x)h + u_{xx}(t, x)\frac{h^2}{2} + u_{xxx}(t, x)\frac{h^3}{6} + O(h^4)$$

$$u(t, x-h) = u(t, x) - u_x(t, x)h + u_{xx}(t, x)\frac{h^2}{2} - u_{xxx}(t, x)\frac{h^3}{6} + O(h^4).$$

Addition der Gleichungen und Division durch h^2 ergibt

$$\frac{1}{h^2}u(t, x+h) + \frac{1}{h^2}u(t, x-h) = \frac{1}{h^2}2u(t, x) + u_{xx}(t, x) + \frac{1}{h^2}O(h^4)$$

und wegen $O(h^4)/h^2 = O(h^2)$ folgt die erste Gleichung.

Die zweite und dritte Gleichung folgen direkt aus der Taylor-Entwicklung in t . Die geforderte Existenz der Konstanten C_K folgt dabei in allen Fällen aus der Stetigkeit der Ableitungen, wodurch die entsprechenden Faktoren in den Taylor-Restgliedern stetig sind und damit auf kompakten Mengen beschränkt sind. \square

Wir erhalten damit das folgende Korollar.

Korollar 9.2 Für die Lösung $u(t, x)$ der Wärmeleitungsgleichung gilt

$$\frac{1}{s}(u(t+s, x) - u(t, x)) - \frac{1}{h^2}(u(t, x+h) - 2u(t, x) + u(t, x-h)) = O(s+h^2) \quad (9.2)$$

und

$$\frac{1}{s}(u(t, x) - u(t-s, x)) - \frac{1}{h^2}(u(t, x+h) - 2u(t, x) + u(t, x-h)) = O(s+h^2) \quad (9.3)$$

sowie für jedes $\theta \in [0, 1]$

$$\begin{aligned} & \frac{1}{s}(u(t+s, x) - u(t, x)) \\ & - \frac{1-\theta}{h^2}(u(t, x+h) - 2u(t, x) + u(t, x-h)) \\ & - \frac{\theta}{h^2}(u(t+s, x+h) - 2u(t+s, x) + u(t+s, x-h)) = O(s+h^2). \end{aligned} \quad (9.4)$$

Beweis: Die Gleichungen (9.2) und (9.3) folgen sofort durch Einsetzen der Gleichungen aus Lemma 9.1 in Gleichung (9.1) und $O(h^2) + O(s) = O(h^2 + s)$. Zum Beweis von Gleichung (9.4) betrachten wir (9.3) für $t + s$ an Stelle von t , also

$$\frac{1}{s}(u(t + s, x) - u(t, x)) - \frac{1}{h^2}(u(t + s, x + h) - 2u(t + s, x) + u(t + s, x - h)) = O(s + h^2).$$

Multiplizieren wir diese Gleichung mit θ und addieren (9.2) multipliziert mit $1 - \theta$, so erhalten wir gerade (9.4). \square

Für $\theta = 0$ ist (9.4) gerade (9.2), für $\theta = 1$ gerade (bis auf Verschiebung der Zeitvariablen t) (9.3). Wir werden später sehen, dass man für jedes $\theta \in [0, 1]$ aus (9.4) sinnvolle numerische Schemata ableiten kann, allerdings mit unterschiedlichen Eigenschaften.

Um nun zu einem implementierbaren Schema zu kommen, müssen wir zunächst das Definitionsgebiet $[0, T] \times \mathbb{R}$ von u auf eine kompakte Menge einschränken. Wir definieren dazu das Rechengebiet $\Omega = [0, T] \times [-a, a]$ mit $a > 0$. Weil das Intervall der x -Werte nun nicht mehr unendlich ist, müssen wir an den neuen Rändern $[0, T] \times \{a\}$ und $[0, T] \times \{-a\}$ Randbedingungen festlegen, um eine eindeutige Lösung zu erhalten. Wir setzen

$$u(t, -a) = g_1(t) \quad \text{und} \quad u(t, a) = g_2(t) \quad \text{für } t \in [0, T].$$

Für die physikalische Interpretation der Gleichung bedeutet dies, dass der eindimensionale Stab nun nicht mehr unendlich lang ist, sondern die Länge $2a$ besitzt. Die Funktionen g_1 und g_2 beschreiben dabei gerade die Temperatur an den freien Endpunkten $-a$ und a des Stabes.

Auf dem Gebiet Ω definieren wir nun ein regelmäßiges Gitter aus $(M + 1)(N + 1)$ Punkten (t_j, x_i) , $j = 0, \dots, M$, $i = 0, \dots, N$. Dazu setzen wir $s := T/M$ und $h := 2a/N$ und definieren

$$t_j := sj, \quad j = 0, \dots, M \quad \text{und} \quad s_i := -a + hi, \quad i = 0, \dots, N.$$

Beachte, dass damit $t_{j+1} = t_j + s$ und $x_{i+1} = x_i + h$ sowie $x_{i-1} = x_i - h$ gilt.

Bezeichnen wir nun die Werte der exakten Lösung u in den Gitterpunkten kurz mit

$$u_i^j := u(t_j, x_i),$$

so können wir (9.4) für $t = t_j$ und $x = x_i$ schreiben als

$$\frac{1}{s}(u_i^{j+1} - u_i^j) - \frac{1-\theta}{h^2}(u_{i+1}^j - 2u_i^j + u_{i-1}^j) - \frac{\theta}{h^2}(u_{i+1}^{j+1} - 2u_i^{j+1} + u_{i-1}^{j+1}) = O(s + h^2). \quad (9.5)$$

Die Finite Differenzenmethode beruht nun darauf, für gegebenes $\theta \in [0, 1]$ Werte $w_i^j \in \mathbb{R}$ zu berechnen, indem diese Gleichungen mit rechter Seite $= 0$ gelöst werden, wobei die entsprechenden Randbedingungen berücksichtigt werden.

Wir suchen also also Werte $w_i^j \in \mathbb{R}$, $i = 0, \dots, N$, $j = 0, \dots, M$, welche die Gleichungen

$$\frac{1}{s}(w_i^{j+1} - w_i^j) - \frac{1-\theta}{h^2}(w_{i+1}^j - 2w_i^j + w_{i-1}^j) - \frac{\theta}{h^2}(w_{i+1}^{j+1} - 2w_i^{j+1} + w_{i-1}^{j+1}) = 0 \quad (9.6)$$

für $i = 1, \dots, N - 1$ und $j = 0, \dots, M - 1$ sowie die Randbedingungen

$$w_i^0 = u_0(x_i), \quad i = 0, \dots, N \quad \text{und} \quad w_0^j = g_1(t_j), \quad w_N^j = g_2(t_j), \quad j = 0, \dots, M \quad (9.7)$$

erfüllen.

9.3 Lösung der Finiten Differenzgleichungen

Wie löst man nun das Gleichungssystem (9.6), (9.7) numerisch? Um (9.6) dafür noch etwas zu vereinfachen, multiplizieren wir die Gleichung mit s und setzen $\alpha := s/h^2$. Mit Umstellen der Terme (alle Terme mit $j+1$ nach links, den Rest nach rechts) ist (9.6) dann äquivalent zu

$$-\alpha\theta w_{i+1}^{j+1} + (2\alpha\theta + 1)w_i^{j+1} - \alpha\theta w_{i-1}^{j+1} = \alpha(1-\theta)w_{i+1}^j + (1 - 2\alpha(1-\theta))w_i^j + \alpha(1-\theta)w_{i-1}^j.$$

Diese Gleichung schreiben wir nun rekursiv in Matrixform und setzen die Randwerte gleich ein. Dies ergibt die implizite rekursive Vorschrift

$$Aw^{j+1} = Bw^j + d^j, \quad j = 0, \dots, M-1 \quad (9.8)$$

mit den Unbekannten

$$w^j = (w_1^j, \dots, w_{N-1}^j)^T,$$

der Matrix

$$A = \begin{pmatrix} 2\alpha\theta + 1 & -\alpha\theta & 0 & \cdots & 0 \\ -\alpha\theta & 2\alpha\theta + 1 & -\alpha\theta & & 0 \\ 0 & \ddots & \ddots & \ddots & 0 \\ \vdots & & & & 0 \\ 0 & \cdots & 0 & -\alpha\theta & 2\alpha\theta + 1 \end{pmatrix},$$

die wir kurz auch als $A = \text{diag}(-\alpha\theta, 2\alpha\theta + 1, \alpha\theta)$ schreiben, der Matrix

$$B = \text{diag}(\alpha(1-\theta), 1 - 2\alpha(1-\theta), \alpha(1-\theta))$$

und dem Vektor

$$d^j = \begin{pmatrix} \alpha(1-\theta)g_1(t_j) + \alpha\theta g_1(t_{j+1}) \\ 0 \\ \vdots \\ 0 \\ \alpha(1-\theta)g_2(t_j) + \alpha\theta g_2(t_{j+1}) \end{pmatrix},$$

der sich aus den Randbedingungen für w_0^j und w_N^j in (9.7) ergibt. Da durch (9.7) zudem die Werte w^0 eindeutig bestimmt sind, können wir (9.8) damit rekursiv für $j = 0, \dots, M-1$ lösen, indem wir in jedem Schritt das lineare Gleichungssystem mit der angegebenen Matrix A numerisch lösen, vorausgesetzt A ist invertierbar, was wir am Ende dieses Abschnitts untersuchen.

Formal könnte man die Iteration (9.8) dann auch explizit als

$$w^{j+1} = A^{-1}Bw^j + A^{-1}d^j$$

schreiben, aus der Numerik I ist aber bekannt, dass die explizite Verwendung von A^{-1} numerisch weniger effizient als die Lösung des zugehörigen linearen Gleichungssystems (9.8) ist.

Für $\theta = 0$ nennt man das Verfahren *Vorwärts-Differenzenverfahren*, für $\theta = 1$ *Rückwärts-Differenzenverfahren*. Für $\theta \in (0, 1)$ erhält man *gemischte Vorwärts-Rückwärts-Differenzenverfahren*, von denen besonders das Verfahren für $\theta = 1/2$ wichtig ist, das sogenannte *Crank-Nicolson-Verfahren*.

Für das Vorwärts-Differenzenverfahren mit $\theta = 0$ wird die Lösung von (9.8) besonders einfach, da A dann die Einheitsmatrix ist. Wir werden aber im nächsten Abschnitt sehen, dass dies aus anderen Gründen keine besonders gute Wahl ist.

Um die Lösbarkeit von (9.8) und allgemein die Existenz eindeutiger Lösungen w^j zu garantieren, reicht es aus, die Invertierbarkeit von A zu untersuchen. Dazu müssen wir nachweisen, dass A keinen Eigenwert $\lambda = 0$ besitzt. Dies kann man leicht mit dem folgenden Satz aus der linearen Algebra beweisen.

Satz 9.3 (Satz von Gerschgorin) Für alle Eigenwerte $\lambda \in \mathbb{C}$ einer Matrix $A \in \mathbb{R}^{n \times n}$ mit Einträgen a_{ij} gilt

$$\lambda \in \bigcup_{i=1}^n \left\{ z \in \mathbb{C} \mid |z - a_{ii}| \leq \sum_{\substack{j=1 \\ j \neq i}}^n |a_{ij}| \right\}.$$

Beweis: Sei $x = (x_1, \dots, x_n)^T \neq 0$ ein Eigenvektor von A zu λ und sei x_i der betragsmäßig größte Eintrag von x . Wegen

$$\lambda x_i = (Ax)_i = \sum_{j=1}^n a_{ij} x_j$$

folgt mit Division durch $x_i \neq 0$

$$|\lambda - a_{ii}| = \left| \sum_{j=1}^n a_{ij} \frac{x_j}{x_i} - a_{ii} \right| = \left| \sum_{\substack{j=1 \\ j \neq i}}^n a_{ij} \frac{x_j}{x_i} \right| \leq \sum_{\substack{j=1 \\ j \neq i}}^n |a_{ij}| \left| \frac{x_j}{x_i} \right| \leq \sum_{\substack{j=1 \\ j \neq i}}^n |a_{ij}|.$$

Daraus folgt die Behauptung. \square

Korollar 9.4 Für alle Eigenwerte λ der Matrix A aus (9.8) gilt $|\lambda| \geq 1$ für alle $\alpha > 0$ und alle $\theta \in [0, 1]$. Insbesondere ist A invertierbar.

Beweis: Sei λ ein Eigenwert von A . Nach Satz 9.3 gilt dann für ein $i = 1, \dots, N-1$

$$|\lambda - a_{ii}| \leq \sum_{\substack{j=1 \\ j \neq i}}^{N-1} |a_{ij}| \leq 2\alpha\theta.$$

Damit folgt mit der umgekehrten Dreiecksungleichung und wegen $a_{ii} = 2\alpha\theta + 1 > 0$

$$2\alpha\theta \geq |\lambda - a_{ii}| \geq 2\alpha\theta + 1 - |\lambda|,$$

woraus die Behauptung folgt. \square

9.4 Konsistenz, Stabilität und Konvergenz

Die Konvergenzanalyse des Verfahrens funktioniert ähnlich wie die Konvergenzanalyse von linearen Mehrschrittverfahren für gewöhnliche Differentialgleichungen, wie wir sie in der Numerik II betrachtet haben.

Die Konsistenzbedingung stellt dabei sicher, dass die Lösung in einem Schritt von (9.8) nur leicht von der exakten Lösung abweicht, die Stabilitätsbedingung garantiert, dass sich diese kleinen Abweichungen in der Iteration in (9.8) nicht zu sehr großen Fehlern aufschaukeln. Wir werden diese beiden Begriffe nun genau definieren und beweisen, dass daraus tatsächlich die Konvergenz — in geeignetem Sinne — folgt.

Konsistenz definieren wir völlig analog zu den Einschrittverfahren: Wir verlangen, dass die exakte Lösung der Gleichung (9.1) die Iterationsvorschrift (9.8) approximativ erfüllt.

Definition 9.5 Das Finite Differenzenschema (9.8) heißt konsistent mit Ordnung $a > 0$ in der Zeit und Ordnung $b > 0$ im Raum, falls für die mittels $u^j = (u(t_j, x_0), \dots, u(t_j, x_N))^T$ definierten Vektoren mit den exakten Lösungswerten die Ungleichung

$$\|Au^{j+1} - Bu^j - d^j\|_\infty \leq Cs(s^a + h^b) \quad (9.9)$$

für eine Konstante $C > 0$ und alle $j = 0, \dots, M - 1$ erfüllen. \square

Beachte, dass die Bedingung einen zusätzlichen Faktor s aber keinen Faktor h in der Abschätzung verlangt. Dies liegt daran, dass über die t -Variable $M = T/s$ Schritte iteriert wird, wodurch — ganz analog zu den numerischen Schemata für gewöhnliche Differentialgleichungen — in der Konvergenzanalyse eine Potenz “verlorengeht”. Da über x nicht iteriert wird, ist eine Multiplikation mit der Ortsschrittweite h nicht nötig.

Der folgende Satz zeigt, dass das Schema konsistent ist.

Satz 9.6 Die Lösung der Wärmeleitungsgleichung erfülle $u \in C^4$. Dann ist das Schema (9.8) für alle $\theta \in [0, 1]$ konsistent mit $a = 1$ und $b = 2$. Im Falle $\theta = 1/2$ erhöht sich die erste Konsistenzordnung auf $a = 2$.

Beweis: Die allgemeine Aussage folgt sofort aus (9.5) wenn wir beachten, dass die Komponenten der rechten Seite von (9.9) gerade durch Multiplikation von (9.5) mit s entstehen, weswegen wir nun $O(s(s + h^2))$ erhalten. Mit $K = \Omega$ und der vor Lemma 9.1 definierten Gleichmäßigkeit der Konstanten in den O -Termen folgt die Existenz von $C = C_K$.

Die spezielle Aussage für $\theta = 1/2$ folgt, wenn wir zusätzlich zu den Abschätzungen aus Lemma 9.1 noch die zusätzlichen Taylor-Entwicklungen

$$\frac{1}{s}(u_i^{j+1} - u_i^j) = u_t(t_j, x_i) + \frac{s}{2}u_{tt}(t_j, x_i) + O(s^2),$$

$$\frac{1}{s}(u_{xx}(t_{j+1}, x_i) - u_{xx}(t_j, x_i)) = u_{xxt}(t_j, x_i) + O(s)$$

sowie

$$\begin{aligned}
& \frac{1-\theta}{h^2}(u_{i+1}^j - 2u_i^j + u_{i-1}^j) - \frac{\theta}{h^2}(u_{i+1}^{j+1} - 2u_i^{j+1} + u_{i-1}^{j+1}) \\
&= (1-\theta)u_{xx}(t_j, x_i) + \theta u_{xx}(t_{j+1}, x_i) + O(h^2) \\
&= u_{xx}(t_j, x_i) + \theta(u_{xx}(t_{j+1}, x_i) - u_{xx}(t_j, x_i)) + O(h^2) \\
&= u_{xx}(t_j, x_i) + \theta s u_{xxt}(t_j, x_i) + O(s^2 + h^2)
\end{aligned}$$

betrachten. Damit lässt sich (9.5) verbessern zu

$$\begin{aligned}
& \frac{1}{s}(u_i^{j+1} - u_i^j) - \frac{1-\theta}{h^2}(u_{i+1}^j - 2u_i^j + u_{i-1}^j) - \frac{\theta}{h^2}(u_{i+1}^{j+1} - 2u_i^{j+1} + u_{i-1}^{j+1}) \\
&= \frac{s}{2}(u_{tt}(t_j, x_i) - 2\theta u_{xxt}(t_j, x_i)) + O(s^2 + h^2).
\end{aligned}$$

Wegen $u_{tt} - u_{xxt} = (u_t - u_{xx})_t = 0$ (denn u löst ja gerade (9.1), also $u_t - u_{xx} = 0$) fällt der erste Summand für $\theta = 1/2$ weg. Wie oben erhalten wir durch Multiplikation dieser Gleichung mit s gerade die Komponenten von (9.9) und damit die Behauptung. \square

Im Sinne der Konsistenz ist folglich $\theta = 1/2$ — also das Crank-Nicolson-Verfahren — am Besten.

Die Bedeutung der Stabilität eines numerischen Schemas haben wir in der Numerik II bei den Mehrschrittverfahren für gewöhnliche Differentialgleichungen bereits gesehen. Anschaulich kann die Notwendigkeit dieser Bedingung wie folgt erklärt werden:

Bezeichnen wir mit $e^j = Au^{j+1} - Bu^j - d^j$ den lokalen Konsistenzfehler und mit $\hat{e}^j = u^j - w^j$ den globalen Fehler der numerischen Lösung, so gilt für den Fehler die Rekursion

$$A\hat{e}^{j+1} = Au^{j+1} - Aw^{j+1} = Bu^j + d^j + e^j - Bw^j - d^j = B\hat{e}^j + e^j,$$

was wir auch als

$$\hat{e}^{j+1} = A^{-1}B\hat{e}^j + A^{-1}e^j$$

schreiben können. Schreiben wir kurz $\tilde{A} = A^{-1}B$, $\tilde{e}^j = A^{-1}e^j$ und verwenden $\hat{e}_0 = 0$ (da die Randbedingungen in $t_0 = 0$ für u und w ja identisch sind), so können wir per Induktion leicht die explizite Formel

$$\hat{e}^j = \sum_{k=0}^{j-1} \tilde{A}^k \tilde{e}^{j-k-1} \quad (9.10)$$

herleiten. Für immer kleiner werdende s und h passieren nun zwei Dinge: Die Anzahl der Schritte j wird immer größer und die Matrix A^{-1} , die den Vektor \tilde{e}^{j-k-1} definiert, verändert sich. Das zweite Problem untersuchen wir später im Beweis von Satz 9.9, das erste betrachten wir jetzt.

Da wir nicht davon ausgehen können, dass sich die einzelnen Summanden günstig gegeneinander aufheben, muss nun, damit \hat{e}^j klein wird, jeder der Summanden klein sein, insbesondere also auch $\tilde{A}^{j-1}\tilde{e}^0$. Da wir für immer kleinere Zeitschritte s auch bei festem Zeitintervall $[0, T]$ immer größere j erhalten, können wir keine obere Schranke für j angeben; ebenso haben wir keine Informationen über \tilde{e}^0 außer, dass wir annehmen können, dass der Wert klein ist. Wir brauchen also eine Eigenschaft der Matrix \tilde{A} , die sicherstellt, dass $\tilde{A}^j x$ für beliebige Vektoren x mit kleiner Norm und beliebigem j ebenfalls klein ist.

Dies liefert uns Satz 2.54 aus der Numerik II:

Satz 9.7 Für eine Matrix $A \in \mathbb{R}^{n \times n}$ existiert genau dann eine Konstante $C > 0$, so dass

$$\|A^j x\| \leq C \|x\|$$

gilt für alle $j \in \mathbb{N}$ und alle $x \in \mathbb{R}^n$, wenn alle Eigenwerte λ von A die Ungleichung $|\lambda| \leq 1$ erfüllen und alle Eigenwerte mit $|\lambda| = 1$ halbeinfach sind (d.h. dass λ eine einfache Nullstelle des Minimalpolynoms ist).

Der folgende Satz wendet dies Kriterium auf die Matrix $A^{-1}B$ an.

Satz 9.8 Die Matrix $A^{-1}B$ erfüllt die Stabilitätsbedingung aus Satz 9.7, wenn die Ungleichungen

$$\begin{aligned} 0 < \alpha \leq \frac{1}{2 - 4\theta}, & \quad \text{falls } 0 \leq \theta < \frac{1}{2} \\ 0 < \alpha, & \quad \text{falls } \frac{1}{2} \leq \theta \leq 1 \end{aligned}$$

gelten. Für die euklidische Norm $\|\cdot\|_2$ kann die Konstante C in Satz 9.7 sogar als $C = 1$ gewählt werden.

Beweis: Wegen

$$B = \frac{1}{\theta} \text{Id} - \frac{1 - \theta}{\theta} A$$

gilt

$$A^{-1}B = \frac{1}{\theta} A^{-1} - \frac{1 - \theta}{\theta} \text{Id}.$$

Jeder Eigenwert von $A^{-1}B$ ist also von der Form $\lambda = \frac{1}{\theta\mu} - \frac{1-\theta}{\theta}$, wobei μ ein Eigenwert von A ist. Wir können nun A schreiben als $A = \text{Id} + \alpha\theta G$ mit $G = \text{diag}(-1, 2, -1)$. Für G kann man nachrechnen, dass die Eigenwerte gerade durch

$$4 \sin^2 \left(\frac{k\pi}{2N} \right), \quad k = 1, \dots, N-1$$

gegeben sind. Die Eigenwerte von A lauten daher

$$\mu_k = 1 + 4\alpha\theta \sin^2 \left(\frac{k\pi}{2N} \right), \quad k = 1, \dots, N-1$$

und die von $A^{-1}B$ folglich

$$\lambda_k = \frac{1}{\theta} \frac{1}{1 + 4\alpha\theta \sin^2(k\pi/2N)} - \frac{1 - \theta}{\theta} = 1 - \frac{4\alpha \sin^2(k\pi/2N)}{1 + 4\alpha\theta \sin^2(k\pi/2N)}.$$

Weil $A^{-1}B$ symmetrisch und damit diagonalisierbar ist, sind alle Eigenwerte halbeinfach. Es reicht also, $|\lambda_k| \leq 1$ nachzuweisen.

Wegen $\alpha > 0$ ist $|\lambda_k| \leq 1$ äquivalent zu

$$\frac{4\alpha \sin^2(k\pi/2N)}{1 + 4\alpha\theta \sin^2(k\pi/2N)} \leq 2 \Leftrightarrow (2 - 4\theta)\alpha \sin^2(k\pi/2N) \leq 1.$$

Für $\theta \geq 1/2$ ist diese Ungleichung immer erfüllt, für $\theta \in [0, 1/2)$ wegen $\sin^2(k\pi/2N) \leq 1$ für alle $\alpha \leq 1/(2 - 4\theta)$.

Dass die Konstante als $C = 1$ gewählt werden kann, folgt wiederum aus der Symmetrie von $A^{-1}B$. Für symmetrische Matrizen A ist die induzierte Matrixnorm $\|A\|_2$ gerade gleich dem maximalen Betrag der Eigenwerte (dies folgt aus der Darstellung der 2-Norm für Matrizen über den Spektralradius, vgl. Satz 2.7 aus der Numerik I). Damit folgt für alle $x \in \mathbb{R}^n$ und alle $j \in \mathbb{N}$ die Ungleichung $\|A^j x\|_2 \leq \|A\|_2^j \|x\|_2 \leq \|x\|_2$, also $C = 1$. \square

Im Sinne der Stabilität sind also die Verfahren mit $\theta \geq 1/2$ besonders gut, weil sie für alle Kombinationen aus Zeit- und Raumschrittweite s und h stabil sind. Man sagt, diese Verfahren sind *unbedingt stabil*.

Nun können wir schließlich die Konvergenz beweisen.

Satz 9.9 Unter den Voraussetzungen der Sätze 9.6 und 9.8 existiert eine Konstante $C_0 > 0$, so dass für alle $\theta \in [0, 1]$ und alle hinreichend kleinen $s, h > 0$ gilt

$$\max_{j=1, \dots, M} \|w^j - u^j\|_{l^2} \leq C_0(s + h^2).$$

Im Falle $\theta = 1/2$ gilt diese Abschätzung sogar mit $C_0(s^2 + h^2)$.

Hierbei ist die *diskrete l^2 -Norm* $\|\cdot\|_{l^2}$ definiert durch

$$\|x\|_{l^2} = \|x\|_2 / (N - 1).$$

Beweis: Betrachte die Fehlergleichung (9.10)

$$\hat{e}^j = \sum_{k=0}^{j-1} \tilde{A}^k \tilde{e}^{j-k-1}$$

mit $\tilde{A} = A^{-1}B$, $\tilde{e}^j = A^{-1}e^j$ und $\hat{e}^j = u^j - w^j$. Wir müssen zeigen, dass $\|\hat{e}^j\|_{l^2} \leq C_0(s + h^2)$ bzw. $\leq C_0(s^2 + h^2)$ gilt für alle $j = 1, \dots, M$.

Für die lokalen Fehler e^j gilt nach Satz 9.6 und der Definition der diskreten l^2 -Norm die Ungleichung $\|e^j\|_{l^2} \leq \|e^j\|_\infty \leq Cs(s^p + h^2)$ mit $p = 1$ für allgemeines θ und $p = 2$ für $\theta = 1/2$. Nach Korollar 9.4 gilt $|\lambda| \geq 1$ für alle Eigenwerte λ von A . Also gilt $|\lambda| \leq 1$ für alle Eigenwerte von A^{-1} und weil A und damit auch A^{-1} symmetrisch sind folgt $\|A^{-1}\|_2 \leq 1$ (vgl. auch das Ende des Beweises von Satz 9.8). Damit folgt $\|\tilde{e}^j\|_{l^2} \leq Cs(s^p + h^2)$.

Nach Satz 9.8 gilt für die Summanden von \hat{e}^j nun $\|\tilde{A}^k \tilde{e}^{j-k-1}\|_{l^2} \leq \|\tilde{e}^{j-k-1}\|_{l^2}$. Damit folgt

$$\|\hat{e}^j\|_{l^2} \leq \sum_{k=0}^{j-1} \|\tilde{A}^k \tilde{e}^{j-k-1}\|_{l^2} \leq jCs(s^p + h^2).$$

Wegen $j \leq M \leq T/s$ folgt

$$\|\hat{e}^j\|_{l^2} \leq \frac{T}{s} Cs(s^p + h^2) = TC(s^p + h^2).$$

Damit ergibt sich die Behauptung mit $C_0 = TC$. \square

Dieser Satz zeigt sofort, warum das Crank-Nicolson-Verfahren, also $\theta = 1/2$, das bevorzugte Verfahren ist: Es hat die höchste Konvergenzordnung und ist unbedingt — d.h. für alle Kombinationen von s und h bzw. von N und M — stabil.

Kapitel 10

Finite Differenzen für die Black-Scholes Gleichung

10.1 Wahl der numerischen Randbedingungen

Im letzten Kapitel haben wir die Finite Differenzen-Methode für die Wärmeleitungsgleichung

$$u_t - u_{xx} = 0$$

betrachtet. Um nun die Black-Scholes-Gleichung mit dieser Methode zu lösen, können wir diese wie in Abschnitt 9.1 beschrieben in die Wärmeleitungsgleichung transformieren, die Wärmeleitungsgleichung mit den transformierten Randbedingungen lösen und die so erhaltene numerische Lösung mit der am Ende von Abschnitt 9.1 angegebenen Formel zurücktransformieren.

Wenn wir so vorgehen, müssen wir berücksichtigen, dass wir den Bereich $[0, \infty)$ der räumlichen Variablen S mittels $x = \ln(S/K)$ auf das Gebiet $(-\infty, \infty)$ für x transformieren, das wir für die numerische Approximation auf ein endliches Intervall $[-a, a]$ einschränken müssen. Dadurch muss an jedem der Ränder $[0, T_0] \times \{-a\}$ und $[0, T_0] \times \{a\}$ eine Randbedingung eingesetzt werden. Da die exakte Lösung nicht als bekannt vorausgesetzt werden soll (denn ansonsten bräuchten wir ja keine numerische Lösung), müssen hier Näherungswerte für die (transformierte) Lösung der Black-Scholes-Gleichung eingesetzt werden.

In den ursprünglichen Koordinaten entsprechen diese Ränder gerade

$$[0, T] \times \{Ke^{-a}\} \quad \text{und} \quad [0, T] \times \{Ke^a\}.$$

Der “untere” Rand liegt also für große a nahe an der Null, der “obere” für große a weit weg von der Null. Für die Wahl geeigneter Randbedingungen für europäische Optionen betrachten wir die a priori Abschätzungen aus Satz 2.3:

- (i) $(S - Ke^{r(t-T)})^+ \leq V_{Ce}(t, S) \leq S$
- (ii) $(Ke^{r(t-T)} - S)^+ \leq V_{Pe}(t, S) \leq Ke^{r(t-T)}.$

Für $S \rightarrow 0$ folgt hieraus $V_{Ce}(t, S) \rightarrow 0$ und $V_{Pe}(t, S) \rightarrow Ke^{r(t-T)}$. Für große S gilt $V_{Ce}(t, S) \approx S$ und $V_{Pe}(t, S) \approx 0$, wobei “ \approx ” hier im Sinne des relativen Fehlers zu verstehen

ist, d.h. es gilt

$$\frac{V_{Ce}(t, S) - S}{S} \rightarrow 0 \quad \text{und} \quad \frac{V_{Pe}(t, S) - 0}{S} \rightarrow 0$$

für $S \rightarrow \infty$. Dies führt zu den Randbedingungen

$$V_{Ce}(t, Ke^{-a}) = 0, \quad V_{Ce}(t, Ke^a) = Ke^a$$

und

$$V_{Pe}(t, Ke^{-a}) = Ke^{r(t-T)}, \quad V_{Pe}(t, Ke^a) = 0.$$

Beachte, dass wir durch die nur approximativ gültigen Randwerte hier einen Fehler erzeugen, der sich in der numerischen Lösung auswirkt. Man kann allerdings beweisen, dass der Fehler für festes S und t und immer größer werdendes a immer kleiner wird.

Für die numerische Lösung der Wärmeleitungsgleichung müssen wir diese Randbedingung nun noch entsprechend transformieren. Es gilt dann mit den Bezeichnungen des vorhergehenden Kapitels

$$g_1(t) = 0, \quad g_2(t) = \exp\left(-\frac{1}{2}(\kappa - 1)a + \frac{1}{4}(\kappa + 1)^2 t\right) e^a$$

für den europäischen Call und

$$g_1(t) = \exp\left(-\frac{1}{2}(\kappa - 1)a + \frac{1}{4}(\kappa + 1)^2 t\right) e^{r(-2t/\sigma^2)}, \quad g_2(t) = 0$$

für den europäischen Put.

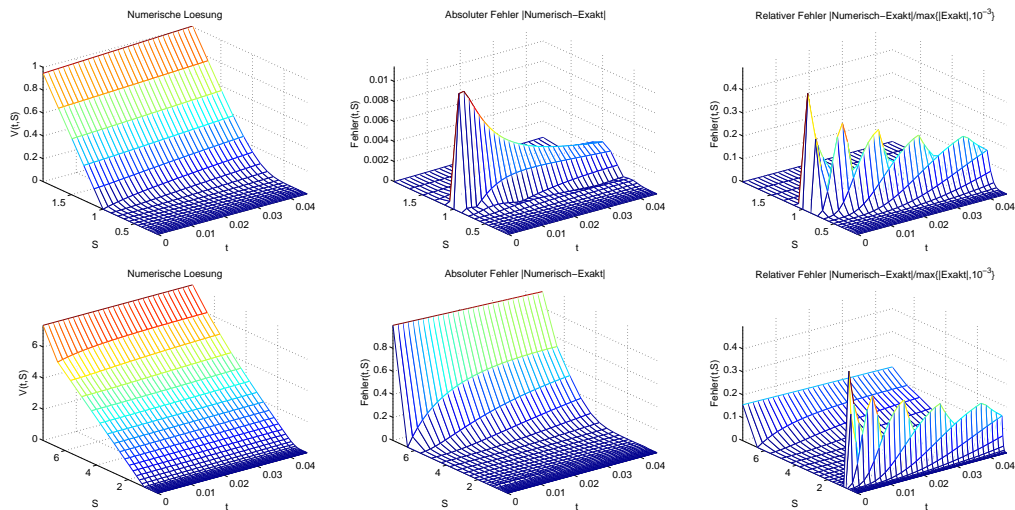


Abbildung 10.1: Numerische Ergebnisse und Fehler auf verschiedenen Gebieten

In Abbildung 10.1 sind die numerischen Lösungen sowie die — durch den Vergleich mit der exakten Lösung erhaltenen — Fehler für den Call abgebildet. Gerechnet wurde hier mit $a = 2$, also für $S \in [e^{-a}, e^a] \approx [0.1353, 7.3891]$. Das Gebiet wurde mit 30 Intervallen je Richtung diskretisiert und die Lösung mit dem Crank-Nicolson-Verfahren berechnet.

In der oberen Zeile sind die Lösungen auf $[0, T] \times [e^{-a}, 2]$ dargestellt. Man sieht, dass sich die nicht ganz korrekten Randwerte hier kaum auswirken. Anders ist dies, wenn die Lösung auf dem ganzen Rechengebiet $[0, T] \times [e^{-a}, e^a]$ dargestellt wird, was in der unteren Zeile des Bildes gemacht wurde. Hier zeigt sich, dass sich der Fehler deutlich stärker auswirkt.

Dass der Fehler im Inneren des Rechengebietes besonders in der Nähe des “Knicks” im Ausübungspreis groß ist, ist natürlich kein Zufall. Hier ist die Lösung am Rand nicht differenzierbar und auch wenn die Lösung im Inneren des Gebietes differenzierbar ist, besitzt sie doch große Ableitungen, was große Restglieder in der Taylor-Entwicklung und damit große numerische Fehler nach sich zieht.

10.2 Direkte Diskretisierung der Black-Scholes-Gleichung

Die numerische Lösung der Black-Scholes-Gleichung über die Transformation in die Wärmeleitungsgleichung ist nur ein möglicher Weg, um diese Gleichung numerisch zu lösen. Alternativ können wir die Gleichung nämlich auch direkt diskretisieren und die Transformation damit vermeiden.

Betrachten wir die Black-Scholes-Gleichung

$$V_t(t, S) + \frac{1}{2}\sigma^2 S^2 V_{SS}(t, S) - rV + rSV_S(t, S) = 0,$$

so haben wir bereits Differenzenquotienten für die Ableitungen V_t und V_{SS} eingeführt. Zusätzlich tauchen hier noch die Terme V und V_S auf. Für den ersten brauchen wir keinen Differenzenquotienten, weil es sich nicht um eine Ableitung handelt. Für den zweiten haben wir verschiedene Möglichkeiten: Wir könnten wieder Vorwärts- oder Rückwärtsdifferenzen verwenden, also

$$V_S(t, S) \approx \frac{1}{h}(V(t, S+h) - V(t, S)) \quad \text{oder} \quad V_S(t, S) \approx \frac{1}{h}(V(t, S) - V(t, S-h)).$$

Es gibt aber auch die Möglichkeit, den sogenannten *zentralen Differenzenquotienten*

$$V_S(t, S) \approx \frac{1}{2h}(V(t, S+h) - V(t, S-h))$$

zu verwenden. Aus der Taylor-Entwicklung folgt

$$V(t, S+h) = V(t, S) + hV_S(t, S) + \frac{h^2}{2}V_{SS}(t, S) + O(h^3)$$

sowie

$$V(t, S-h) = V(t, S) - hV_S(t, S) + \frac{h^2}{2}V_{SS}(t, S) + O(h^3).$$

Subtraktion der zweiten Gleichung von der ersten liefert

$$V(t, S+h) - V(t, S-h) = 2hV_S(t, S) + O(h^3),$$

woraus durch Multiplikation mit h und Umstellen der Terme

$$V_S(t, S) = \frac{1}{2h}(V(t, S+h) - V(t, S-h)) + O(h^2)$$

folgt. Der zentrale Differenzenquotient ist also eine Approximation zweiter Ordnung und liefert daher eine bessere Approximation als der Vorwärts- oder Rückwärts-Quotient, vgl. Lemma 9.1.

Das Finite Differenzschema für die Black-Scholes-Gleichung erhalten wir nun, indem wir die Ableitungen in der Gleichung durch die entsprechenden Differenzenquotienten ersetzen. Betrachten wir wieder ein Gitter mit Gitterpunkten (t_j, S_i) und schreiben kurz $V_i^j = V(t_j, S_i)$, so erhalten wir mit Vorwärtsdifferenzen in der Zeit

$$\frac{1}{s}(V_i^{j+1} - V_i^j) + \frac{\sigma^2 S_i^2}{2h^2}(V_{i+1}^j - 2V_i^j + V_{i-1}^j) - rV_i^j + \frac{rS_i}{2h}(V_{i+1}^j - V_{i-1}^j) = O(s + h^2)$$

und mit Rückwärtsdifferenzen in der Zeit (mit verschobenem Zeitargument)

$$\frac{1}{s}(V_i^{j+1} - V_i^j) + \frac{\sigma^2 S_i^2}{2h^2}(V_{i+1}^{j+1} - 2V_i^{j+1} + V_{i-1}^{j+1}) - rV_i^{j+1} + \frac{rS_i}{2h}(V_{i+1}^{j+1} - V_{i-1}^{j+1}) = O(s + h^2).$$

Mit $1 - \theta$ bzw. θ gewichtet und aufsummiert ergibt sich

$$\begin{aligned} & \frac{1}{s}(V_i^{j+1} - V_i^j) \\ & + (1 - \theta) \left(\frac{\sigma^2 S_i^2}{2h^2}(V_{i+1}^j - 2V_i^j + V_{i-1}^j) - rV_i^j + \frac{rS_i}{2h}(V_{i+1}^j - V_{i-1}^j) \right) \\ & + \theta \left(\frac{\sigma^2 S_i^2}{2h^2}(V_{i+1}^{j+1} - 2V_i^{j+1} + V_{i-1}^{j+1}) - rV_i^{j+1} + \frac{rS_i}{2h}(V_{i+1}^{j+1} - V_{i-1}^{j+1}) \right) = O(s + h^2). \end{aligned}$$

Wie im vorhergehenden Kapitel erhalten wir nun Approximationen $W_i^j \approx V_i^j$, indem wir den $O(s + h^2)$ -Term auf der rechten Seite dieser Gleichung vernachlässigen, also gleich Null setzen, und das entstehende Gleichungssystem

$$\begin{aligned} & \frac{1}{s}(W_i^{j+1} - W_i^j) \\ & + (1 - \theta) \left(\frac{\sigma^2 S_i^2}{2h^2}(W_{i+1}^j - 2W_i^j + W_{i-1}^j) - rW_i^j + \frac{rS_i}{2h}(W_{i+1}^j - W_{i-1}^j) \right) \\ & + \theta \left(\frac{\sigma^2 S_i^2}{2h^2}(W_{i+1}^{j+1} - 2W_i^{j+1} + W_{i-1}^{j+1}) - rW_i^{j+1} + \frac{rS_i}{2h}(W_{i+1}^{j+1} - W_{i-1}^{j+1}) \right) = 0 \end{aligned}$$

iterativ lösen. Auch hier liefert das Crank-Nicolson-Verfahren mit $\theta = 1/2$ die beste Approximation. Weil die Black-Scholes-Gleichung rückwärts von der Endbedingung zur Zeit $t_M = T$ aus berechnet werden muss, muss die Gleichung dafür nach W^j aufgelöst und in Vektorform geschrieben werden (während wir bei der Wärmeleitungsgleichung nach w^{j+1} aufgelöst hatten). Beachte, dass dadurch das Schema mit $\theta = 1$ explizit wird (bei der Wärmeleitungsgleichung war dies für $\theta = 0$ der Fall), während in allen Schemata für $\theta \in [0, 1)$ in jedem Schritt ein Gleichungssystem gelöst werden muss.

Da das räumliche Gebiet $S \in [0, \infty)$ hier nur einseitig unbeschränkt ist, müssen wir dabei nur auf einer Seite “abschneiden”, d.h. auf $[0, L]$ für ein $L > 0$ rechnen und den Randwert $V(t, L)$ analog zum vorhergehenden Abschnitt approximativ einsetzen. Die Details der vollständigen Herleitung der zu lösenden Matrix-Gleichung wollen wir hier aus Zeitgründen nicht betrachten.

10.3 Finite Differenzen und Binomialmethode

Wir haben diese Vorlesung in Abschnitt 3.2 mit der Binomialmethode begonnen und in Abschnitt 6.5 gezeigt, dass diese Methode als schwache Approximation der geometrischen Brownschen Bewegung verstanden werden kann. Zum Abschluss der Vorlesung zeigen wir nun, dass die Binomialmethode auch mit der Finiten Differenzenmethode zusammenhängt.

Wir betrachten dazu das Finite-Differenzenschema angewendet auf die Hilfsfunktion $v(t, x) = V(t, S)$ mit $x = \ln(S)$. Ähnlich wie in Abschnitt 9.1 bzw. Übungsaufgabe 1 von Blatt 11 rechnet man nach, dass diese Funktion die Gleichung

$$v_t + \frac{1}{2}\sigma^2 v_{xx} + (r - \frac{1}{2}\sigma^2)v_x - rv = 0$$

erfüllt. Setzen wir nun $w(t, x) = e^{-rt}v(t, x)$, so rechnet man nach, dass w die Gleichung

$$w_t + \frac{1}{2}\sigma^2 w_{xx} + (r - \frac{1}{2}\sigma^2)w_x = 0$$

erfüllt. Ersetzen wir nun die Ableitungen wie im vorherigen Abschnitt durch die Differenzenquotienten und verwenden die Rückwärtsdifferenzen in der Zeit, so erhalten wir Approximationen $\tilde{w}_i^j \approx w(t_j, x_i)$ mittels

$$\frac{1}{s}(\tilde{w}_i^{j+1} - \tilde{w}_i^j) + \frac{1}{2h^2}\sigma^2(\tilde{w}_{i+1}^{j+1} - 2\tilde{w}_i^{j+1} + \tilde{w}_{i-1}^{j+1}) + \frac{r - \sigma^2/2}{2h}(\tilde{w}_{i+1}^{j+1} - \tilde{w}_{i-1}^{j+1}) = 0.$$

Aufgelöst nach \tilde{w}_i^j ergibt sich

$$\begin{aligned} \tilde{w}_i^j &= \tilde{w}_i^{j+1} + \frac{s}{2h^2}\sigma^2(\tilde{w}_{i+1}^{j+1} - 2\tilde{w}_i^{j+1} + \tilde{w}_{i-1}^{j+1}) + \frac{s(r - \sigma^2/2)}{2h}(\tilde{w}_{i+1}^{j+1} - \tilde{w}_{i-1}^{j+1}) \\ &= \left(1 - \frac{s\sigma^2}{h^2}\right)\tilde{w}_i^{j+1} + \left(\frac{s\sigma^2}{2h^2} + \frac{s(r - \sigma^2/2)}{2h}\right)\tilde{w}_{i+1}^{j+1} + \left(\frac{s\sigma^2}{2h^2} - \frac{s(r - \sigma^2/2)}{2h}\right)\tilde{w}_{i-1}^{j+1}. \end{aligned}$$

Mit der Wahl $h = \sigma\sqrt{s}$ folgt

$$\begin{aligned} \tilde{w}_i^j &= \left(\frac{s\sigma^2}{2h^2} + \frac{s(r - \sigma^2/2)}{2h}\right)\tilde{w}_{i+1}^{j+1} + \left(\frac{s\sigma^2}{2h^2} - \frac{s(r - \sigma^2/2)}{2h}\right)\tilde{w}_{i-1}^{j+1} \\ &= \left(\frac{1}{2} + \frac{\sqrt{s}(r - \sigma^2/2)}{2\sigma}\right)\tilde{w}_{i+1}^{j+1} + \left(\frac{1}{2} - \frac{\sqrt{s}(r - \sigma^2/2)}{2\sigma}\right)\tilde{w}_{i-1}^{j+1} \\ &= p\tilde{w}_{i+1}^{j+1} + (1-p)\tilde{w}_{i-1}^{j+1} \end{aligned}$$

mit

$$p = \frac{1}{2} + \frac{\sqrt{s}(r - \sigma^2/2)}{2\sigma}.$$

Für $\tilde{V}_i^j \approx V(t_j, S_i)$ erhalten wir wegen $V(t_j, S_i) = e^{rt}w(t_j, x_i)$ mit $S_i = e^{x_i}$ die Vorschrift

$$\tilde{V}_i^j = e^{-rs} \left(p\tilde{V}_{i+1}^{j+1} + (1-p)\tilde{V}_{i-1}^{j+1} \right).$$

Dies ist genau die Iterationsvorschrift der Binomialmethode. Beachte, dass wir hier kein Binomialmodell mit $p = 1/2$ herausbekommen.

Zur Berechnung der zugehörigen Dynamik des Binomialmodells

$$f(S, Z) = S \cdot Z \quad \text{mit} \quad \mathbb{P}_Z(\alpha_u) = p \quad \text{und} \quad \mathbb{P}_Z(\alpha_d) = 1 - p,$$

müssen wir noch die Werte α_u und α_d bestimmen. Da sich die Werte für \tilde{V}_i^j aus V_{i+1}^{j+1} und V_{i-1}^{j+1} ergeben, muss $f(S_i, Z) = S_{i+1}$ mit Wahrscheinlichkeit p und $f(S_i, Z) = S_{i-1}$ mit Wahrscheinlichkeit $1 - p$ gelten. Wegen $x_i = hi$ gilt

$$S_i = e^{hi} = e^{\sigma i \sqrt{s}}$$

und damit

$$S_{i+1} = e^{\sigma \sqrt{s}} S_i \quad \text{und} \quad S_{i-1} = e^{-\sigma \sqrt{s}} S_i.$$

Folglich gilt $\alpha_u = e^{\sigma \sqrt{s}}$ und $\alpha_d = e^{-\sigma \sqrt{s}}$. Für den Erwartungswert dieses Modells gilt

$$\begin{aligned} E(f(S, Z)) &= pe^{\sigma \sqrt{s}} S + (1-p)e^{-\sigma \sqrt{s}} S \\ &= \underbrace{p\left(1 + \sigma \sqrt{s} + \frac{\sigma^2 s}{2}\right) S + (1-p)\left(1 - \sigma \sqrt{s} + \frac{\sigma^2 s}{2}\right) S}_{=(1+sr)S} + O(s^{3/2}) \\ &= (1 + sr)S + O(s^{3/2}) = e^{sr} S + O(s^{3/2}), \end{aligned}$$

d.h. bis auf einen Fehlerterm der Ordnung $s^{3/2}$ von der für die risikoneutralen Bewertung benötigten Form.

Der wesentliche Unterschied der Binomialmethode zur Finiten Differenzenmethode besteht darin, dass die Binomialmethode stets nur die Werte in den Knoten (t_j, S_i) berechnet, die zur Berechnung des gewünschten Wertes in $(0, S_{i_0})$ unbedingt benötigt werden, also diejenigen, die in dem entsprechenden Binomialbaum liegen. Im Gegensatz dazu liefert die Finite Differenzenmethode Werte für alle vorgegebenen Gitterknoten. Abbildung 10.2 veranschaulicht diese ‘‘Einbettung’’ des Binomialbaums in das Finite Differenzengitter schematisch.

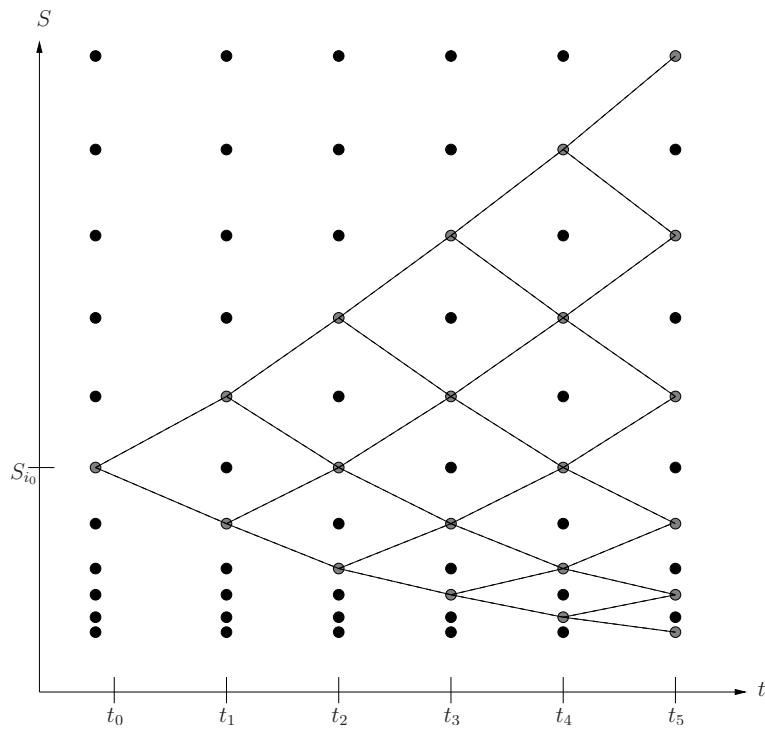


Abbildung 10.2: Binomialbaum und Finite Differenzengitter

Literaturverzeichnis

- [1] S. CYGANOWSKI, P. KLOEDEN, AND J. OMBACH, *From elementary probability to stochastic differential equations with MAPLE[®]*, Universitext, Springer-Verlag, Berlin, 2002.
- [2] M. GÜNTHER AND A. JÜNGEL, *Finanzderivate mit MATLAB[®]*, Vieweg-Verlag, Braunschweig, 2003.
- [3] D. J. HIGHAM, *An introduction to financial option valuation*, Cambridge University Press, Cambridge, 2004. Mathematics, stochastics and computation.
- [4] T. MIKOSCH, *Elementary stochastic calculus—with finance in view*, vol. 6 of Advanced Series on Statistical Science & Applied Probability, World Scientific Publishing Co. Inc., River Edge, NJ, 1998.
- [5] R. U. SEYDEL, *Tools for computational finance*, Universitext, Springer-Verlag, Berlin, fourth ed., 2009.

Index

- Abtastfeedback, 60
- Abzinsung, 4
- antithetische Zufallsvariable, 39
- Approximation
 - schwach, 42
 - stark, 41
- Arbitrage, 3
- Arbitragefreiheit, 3
- Ausübungspreis, 2, 11
- Ausübungszeitpunkt, 2, 11

- Basiswert, 2, 11
- bedingte Wahrscheinlichkeit, 7
- bedingter Erwartungswert, 7
- Binomialmethode
 - und Finite Differenzen, 81
- Binomialmodell, 20
 - als schwache Approximation, 50
 - Parameter, 23
- Binomialbaum, 20
- Black-Scholes Gleichung, 56
 - und Wärmeleitungsgleichung, 66
- Brownsche Bewegung, 26
 - geometrisch, 32

- Call-Option, 2, 11
- Crank-Nicolson-Verfahren, 70, 75

- Derivat, 11
- Dichtefunktion, 7
- Differenzenquotient, 67, 79
- Diffusion, 29
- Diskontierung, 4
- diskrete l^2 -Norm, 75
- Drift, 29

- Einschrittverfahren
 - stochastisch, 46
 - Grundalgorithmus, 46
- Elementarereignis, 5

- Erwartungswert, 7
 - bedingt, 7
- Euler-Maruyama, *siehe* Euler-Verfahren, stochastisch
- Euler-Verfahren
 - stochastisch, 46
 - Konvergenzeigenschaften, 47

- fairer Preis, 3
- Feedback, 59
- Finite Differenzen, 67
 - und Binomialmethode, 81

- geometrische Brownsche Bewegung, 32
- Gerschgorin
 - Satz von, 71
- Gleichverteilung, 6
- Griechen, 62

- Hedging, 3, 57, 59
 - Simulation, 60
 - zeitdiskret, 59
- historische Parameter, 24

- implizite Parameter, 24
- Inkrement des Wiener Prozesses, 26
- Itô-Integral, 29
- Itô-Lemma, 30

- Kennzahlen, 62
- Konfidenzintervall, 37
- Konsistenz, 72
- Konvergenz, 75
- Kovarianzmatrix, 7
- Kursverlaufsmodell, 32

- Leerverkauf, 13

- Martkmodell, 3
 - messbar, 6

- Monte-Carlo-Methode, 35, 63
 - Algorithmus, 36
 - Konvergenzgeschwindigkeit, 37
- Normalverteilung, 25
 - Verteilungsfunktion, 36
- Option, 2, 11
 - amerikanisch, 2, 11
 - asiatisch, 63
 - Barriere, 65
 - Call, 2, 11
 - europäisch, 2, 11
 - Power, 65
 - Put, 2, 11
- Optionsbewertung, 3
 - Grundalgorithmus, 12
 - risikoneutral, 3, 12, 17, 53, 57
- Pfad, 19, 26
- Portfolio, 13
 - Black-Scholes, 54
 - selbstfinanzierend, 54
- Put-Call-Parität, 14
- Put-Option, 2, 11
- Randbedingung
 - numerisch, 77
- Randwert, 56
- Realisierung, 7
- rekombinierbar, 20
- risikoavers, 12
- risikofreudig, 12
- risikoneutral, *siehe* Optionsbewertung, risikoneutral
- Rücklaufquote, 32
- Rückwärts-Differenzenverfahren, 70
- schwache Approximation, 42
- selbstfinanzierend, 54, 60
- σ -Algebra, 5
 - minimale, 6
- Stabilität, 73
 - unbedingt, 75
- Stabilitätsbedingung, 74
- starke Approximation, 41
- stochastisches Integral, 28
- stochastische Differentialgleichung, 25
- Itô, 29
 - Stratonovich, 30
- stochastischer Prozess, 19, 26
- stochastisches dynamisches System, 19
- Varianz, 7
- Varianzreduktion, 38
- Volatilität, 32
- Vorwärts-Differenzenverfahren, 70
- Wärmeleitungsgleichung, 66
- Wahrscheinlichkeit
 - bedingt, 7
- Wahrscheinlichkeitsmaß, 5
- Wahrscheinlichkeitsraum, 5
- Wahrscheinlichkeitsverteilung, 6
- Wiener Prozess, 26
 - schwache Approximation, 44
 - starke Approximation, 44
- Zufallsvariable, 6
 - antithetisch, 39
 - unabhängig, 9
- Zufallszahlen, 35