

Differentialgleichungen in der Wirtschaftsmathematik

Skript zur Vorlesung im Wintersemester 2010/11
an der TU Dortmund

PD Dr. Flavius Guiaş

2. Februar 2011

Inhaltsverzeichnis

1	Bedingter Erwartungswert	3
2	Stochastische Prozesse in stetiger Zeit	5
2.1	Stochastische Prozesse	5
2.2	Filtration	6
2.3	Stopzeiten	7
2.4	Martingale	8
2.5	Brownsche Bewegung	8
2.6	Poisson-Prozess	10
3	Markov-Prozesse	12
3.1	Definition von Markov-Prozessen	12
3.2	Markov-Prozesse mit abzählbarem Zustandsraum	13
3.3	Kolmogorovsche Differentialgleichung	14
4	Riemann-Stieltjes Integral	17
5	Thiele'sche Differentialgleichungen	21
5.1	Reguläres Versicherungsmodell	21
5.2	Deckungskapital	24
6	Stochastisches Integral	28
6.1	Konstruktion des stochastischen Integrals	28
6.2	Die Itô-Formel	32
7	Stochastische Differentialgleichungen	36
7.1	Problemformulierung	36
7.2	Existenz und Eindeutigkeit	37
7.3	Starke Markoveigenschaft	37
7.4	Generator	38
7.5	Fokker-Planck-Gleichung	40
7.6	Feynman-Kac-Formel	40

8	Black-Scholes Differentialgleichung	42
8.1	Optionen	42
8.2	Herleitung der Black-Scholes Gleichung	46
8.2.1	Herleitung mit Hilfe des Duplikationsprinzips	46
8.2.2	Herleitung mit Hilfe der risikoneutralen Bewertung	48
8.3	Eigenschaften der Black-Scholes Gleichung	50
8.3.1	Exkurs: Die eindimensionale Diffusionsgleichung	53
8.4	Amerikanische Optionen und freie Randwertprobleme	57
8.4.1	Exkurs: Das Hindernisproblem	60
8.5	Asiatische Optionen	63
8.6	Die Monte-Carlo-Methode	65
8.6.1	Das Euler-Maruyama Verfahren	66
8.6.2	Das Milstein-Verfahren	67
9	Stochastische Steuerung und Hamilton-Jacobi-Bellman Differentialgleichung	69
9.1	Problemformulierung	69
9.2	Konzept der dynamischen Programmierung	72
9.3	Ein Verifikationstheorem	74
9.4	Portfolio-Optimierung: Optimaler Konsum und optimales Endvermögen bei endlichem Zeithorizont	78
9.5	Unendlicher Zeithorizont	81
9.6	Portfolio-Optimierung: Optimaler Konsum bei unendlichem Zeithorizont	83
9.7	Stoppen des gesteuerten Prozesses	84
9.8	Portfolio-Optimierung: Portfolio-Versicherung	86
	Literaturverzeichnis	88

Kapitel 1

Bedingter Erwartungswert

Ein wichtiges Konzept in der Finanz- und Versicherungsmathematik ist der erwartete Wert einer Zufallsvariablen X , gegeben die aktuell verfügbare Information, die durch eine σ -Algebra \mathcal{G} beschrieben wird. Dieser bedingte Erwartungswert ist wiederum eine Zufallsvariable (beste "Darstellung" von X durch eine \mathcal{G} -messbare Zufallsvariable):

Definition 1.1 Sei (Ω, \mathcal{A}, P) ein Wahrscheinlichkeitsraum, $\mathcal{G} \subseteq \mathcal{A}$ eine Unter- σ -Algebra und $X \in L_1(\Omega, \mathcal{A}, P)$ (d.h. X ist eine \mathbb{R} -wertige Zufallsvariable mit $E[|X|] < \infty$). Eine \mathbb{R} -wertige Zufallsvariable Z nennen wir bedingten Erwartungswert von X unter der Bedingung \mathcal{G} , falls:

- (i) Z ist \mathcal{G} -messbar und
- (ii) $\int_G Z dP = \int_G X dP$ für alle $G \in \mathcal{G}$.

Formal schreiben wir $E[X|\mathcal{G}] := Z$.

Bemerkungen:

- (i) Man kann zeigen, dass eine Zufallsvariable Z wie in Definition 1.1 existiert. Diese ist P -fast eindeutig. Genauer gesagt ist $E[X|\mathcal{G}]$ also eine Äquivalenzklasse in $L_1(\Omega, \mathcal{A}, P)$ unter der Äquivalenzrelation „fast sicher gleich“. Wir verstehen $E[X|\mathcal{G}]$ aber stets als einen Vertreter dieser Klasse.
- (ii) Es gilt: X ist \mathcal{G} -messbar genau dann, wenn $E[X|\mathcal{G}] = X$.

Definition 1.2 Sei (Ω, \mathcal{A}, P) ein Wahrscheinlichkeitsraum, $\mathcal{G} \subseteq \mathcal{A}$ eine Unter- σ -Algebra und $X \in L_1(\Omega, \mathcal{A}, P)$.

(i) Sei nun Y eine weitere Zufallsvariable auf (Ω, \mathcal{A}, P) und $\sigma(Y)$ die von Y erzeugte σ -Algebra. Dann bezeichnet

$$E[X|Y] := E[X|\sigma(Y)]$$

den bedingten Erwartungswert von X unter der Bedingung Y .

(ii) Für $A \in \mathcal{A}$ ist die bedingte Wahrscheinlichkeit unter der Bedingung \mathcal{G} gegeben durch

$$P(A|\mathcal{G}) := E[\mathbb{1}_A|\mathcal{G}].$$

Für den bedingten Erwartungswert gelten die folgenden Rechenregeln:

Satz 1.3 Sei (Ω, \mathcal{A}, P) ein Wahrscheinlichkeitsraum, $\mathcal{G} \subseteq \mathcal{A}$ eine Unter- σ -Algebra und $X, Y \in L_1(\Omega, \mathcal{A}, P)$. Dann gilt:

- (i) Mittelwert: $E[E[X|\mathcal{G}]] = E[X]$.
- (ii) Linearität: $E[\alpha X + \beta Y|\mathcal{G}] = \alpha E[X|\mathcal{G}] + \beta E[Y|\mathcal{G}]$ für $\alpha, \beta \in \mathbb{R}$.
- (iii) Monotonie: $E[X|\mathcal{G}] \leq E[Y|\mathcal{G}]$ für $X \leq Y$.
- (iv) $E[X|\mathcal{G}] = E[X]$, falls X unabhängig von \mathcal{G} ist (d.h. wenn gilt $P(A \cap A') = P(A)P(A')$ für beliebige $A \in \mathcal{G}$ und $A' \in \sigma(X)$).
- (v) $E[XY|\mathcal{G}] = YE[X|\mathcal{G}]$, falls Y \mathcal{G} -messbar ist.
- (vi) Tower property: $E[E[X|\mathcal{G}_2]|\mathcal{G}_1] = E[X|\mathcal{G}_1]$ bzw. $E[E[X|\mathcal{G}_1]|\mathcal{G}_2] = E[X|\mathcal{G}_1]$ für $\mathcal{G}_1 \subseteq \mathcal{G}_2 \subseteq \mathcal{A}$.
- (vii) Jensensche Ungleichung: Falls f konvex ist mit $E[|f(X)|] < \infty$, so gilt $f(E[X|\mathcal{G}]) \leq E[f(X)|\mathcal{G}]$.

Beweis: Siehe z.B. [Ir01] oder [B93]. □

Da der bedingte Erwartungswert eine Zufallsvariable ist, sind die obigen Gleichungen und Ungleichungen stets nur im P -f.s. Sinne zu verstehen.

Kapitel 2

Stochastische Prozesse in stetiger Zeit

In diesem Abschnitt sammeln wir einige grundlegende Begriffe und Resultate zu stochastischen Prozessen. Dazu legen wir im Folgenden einen Wahrscheinlichkeitsraum (Ω, \mathcal{A}, P) und eine Indexmenge I zugrunde.

2.1 Stochastische Prozesse

Die Grundlage für eine allgemeine mathematische Modellierung für zufällige Phänomene bildet die folgende Definition:

Definition 2.1 Sei (S, \mathcal{B}) ein Messraum und $X = (X_t)_{t \in I}$ eine Familie von Zufallsvariablen $X_t : \Omega \rightarrow S$. Dann heißt X ein stochastischer Prozess mit Zustandsraum (S, \mathcal{B}) , Indexmenge I und Wahrscheinlichkeitsraum (Ω, \mathcal{A}, P) . Sei $\omega \in \Omega$. Dann heißt $X(\cdot, \omega) : I \rightarrow S$, $X(t, \omega) = X_t(\omega)$ der Pfad von X zu $\omega \in \Omega$.

Deutung:

- (i) Meistens wird $I = [0, \infty)$ oder $I = [0, T]$ gewählt und entspricht einem Zeitbereich.
- (ii) Für jedes $\omega \in \Omega$ beschreibt der Pfad $t \mapsto X_t(\omega)$ die zugehörige Entwicklung des Prozesses.

Bemerkungen:

- (i) Sind die Pfade von X (P -f.s.) stetig (linksstetig, rechtsstetig), so nennen wir X entsprechend stetig (linkstetig, rechtsstetig).

- (ii) Sind alle X_t integrierbar (d.h. liegen in L_1) oder quadratintegrierbar (d.h. liegen in L_2), so sprechen wir von einem integrierbaren bzw. quadratintegrierbaren Prozess (oder entsprechend von einem L_p -Prozess, $p = 1, 2$).

Um zu entscheiden, ob zwei Prozesse als „gleich“ anzusehen sind, benötigen wir die folgende Definition:

Definition 2.2 Seien $X = (X_t)_{t \in I}$ und $Y = (Y_t)_{t \in I}$ zwei stochastische Prozesse. Dann heißt Y eine Modifikation von X , falls

$$P(\{\omega \in \Omega : X_t(\omega) = Y_t(\omega)\}) = 1 \text{ für alle } t \in I.$$

X und Y heißen weiterhin ununterscheidbar, falls sie mit Wahrscheinlichkeit 1 die gleichen Pfade haben:

$$P(\{\omega \in \Omega : X_t(\omega) = Y_t(\omega) \forall t \in I\}) = 1.$$

2.2 Filtration

Definition 2.3 Eine Familie $\mathcal{F} = (\mathcal{F}_t)_{t \in I}$ von Unter- σ -Algebren von \mathcal{A} mit $\mathcal{F}_s \subseteq \mathcal{F}_t$ für alle $s \leq t \in I$ heißt Filtration.

Deutung: Die Menge \mathcal{F}_t wird als die Information interpretiert, die der Beobachter zur Zeit t hat, d.h. für ein Ereignis $A \in \mathcal{F}_t$ kann er zur Zeit t entscheiden ob A eingetreten ist oder nicht.

Definition 2.4 Ein stochastischer Prozess ist adaptiert zur Filtration \mathcal{F} , falls X_t für alle $t \in I$ \mathcal{F}_t -messbar ist.

Beispiel: Sei $X = (X_t)_{t \in I}$ ein stochastischer Prozess. Dann ist $\mathcal{F}^X = (\mathcal{F}_t^X)_{t \in I}$ mit $\mathcal{F}_t^X = \sigma(X(s) : s \leq t)$ eine Filtration und heißt die von X erzeugte *kanonische Filtration*. Dies ist die kleinste σ -Algebra, bezüglich der alle $X_t, t \in I$, messbar sind. Insbesondere ist X adaptiert zu \mathcal{F}^X .

Sei von nun an $I = [0, \infty)$. Durch geeignete Festlegung von X_t für $t > T$ gelten sämtliche Aussagen auch für Prozesse $(X_t)_{t \in [0, T]}$ mit endlichem Zeitintervall T . Statt $t \in I$ schreiben wir kurz $t \geq 0$.

Aus technischen Gründen benötigen wir die folgenden Definitionen:

Definition 2.5 (i) Eine Filtration \mathcal{F} heißt rechtsstetig, falls $\mathcal{F}_t = \mathcal{F}_t^+ := \bigcap_{s > t} \mathcal{F}_s$.

(ii) Eine Filtration \mathcal{F} heißt vollständig, falls \mathcal{F}_0 bereits alle P -Nullmengen aus \mathcal{A} enthält.

(iii) Wir sagen, dass \mathcal{F} die üblichen Bedingungen erfüllt, wenn \mathcal{F} rechtsstetig und vollständig ist.

Definition 2.6 Ein reellwertiger stochastischer Prozess X ist progressiv messbar bezüglich einer Filtration \mathcal{F} , falls für alle $t \geq 0$ die Abbildungen $(s, \omega) \mapsto X_s(\omega)$ auf $[0, t] \times \Omega$ messbar sind bezüglich $\mathcal{B}([0, t]) \otimes \mathcal{F}_t$. Hierbei bezeichnet $\mathcal{B}(U)$ die Borel- σ -Algebra über U .

Bemerkung 2.7 (i) Ein progressiv messbarer Prozess ist stets adaptiert.

(ii) Ist X ein progressiv messbarer stochastischer Prozess bezüglich einer vollständigen Filtration \mathcal{F} , so ist auch jede Modifikation von X progressiv messbar bezüglich \mathcal{F} .

Satz 2.8 Ist X links- oder rechtsstetig, so ist X progressiv messbar.

Beweis: Proposition 1.1.13 in [KS99] oder Satz 28 in [KK01]. \square

2.3 Stoppzeiten

Definition 2.9 Eine Zufallsvariable $\tau : \Omega \rightarrow [0, \infty]$ ist eine Stoppzeit bezüglich einer Filtration \mathcal{F} , falls $\{\tau \leq t\} \in \mathcal{F}_t$ für alle $t \in [0, \infty)$.

Deutung: Die Stoppzeit entspricht einer zulässigen Abbruchbedingung. Die Forderung $\{\tau \leq t\} \in \mathcal{F}_t$ besagt gerade, dass wir zum Zeitpunkt t in der Lage sein müssen, zu entscheiden, ob gestoppt wird oder nicht.

Bemerkung 2.10 Sind τ_1, τ_2 Stoppzeiten, so auch $\tau_1 \wedge \tau_2 := \min\{\tau_1, \tau_2\}$.

Definition 2.11 Zu einer \mathcal{F} -Stoppzeit τ ist die σ -Algebra der Ereignisse bis zur Zeit τ definiert durch

$$\mathcal{F}_\tau = \{A \in \mathcal{A} : A \cap \{\tau \leq t\} \in \mathcal{F}_t \text{ für alle } t \geq 0\}.$$

Definition 2.12 Ist X adaptiert bezüglich \mathcal{F} und τ eine \mathcal{F} -Stoppzeit, so wird der Wert des gestoppten Prozesses definiert durch

$$X_{t \wedge \tau}(\omega) = \begin{cases} X_t(\omega) & , \text{ falls } t \leq \tau(\omega), \\ X_{\tau(\omega)}(\omega) & , \text{ falls } t > \tau(\omega). \end{cases}$$

Satz 2.13 *Ist X progressiv messbar bezüglich \mathcal{F} und τ eine \mathcal{F} -Stoppzeit, so ist $X_{t \wedge \tau}$ selbst $\mathcal{F}_{t \wedge \tau}$ -messbar.*

Beweis: Siehe Satz 6.10 (i) in [Ir03]. □

Satz 2.14 *Sei \mathcal{F} eine rechtstetige Filtration, X ein progressiv messbarer Prozess, und für ein $B \in \mathcal{B}(\mathbb{R})$*

$$\tau_B = \inf\{t \geq 0 : X_t \in B\}.$$

Dann gilt

- (i) *Ist X rechtsstetig und B offen, so ist τ_B eine Stoppzeit.*
- (ii) *Ist X stetig und B abgeschlossen, so ist τ_B eine Stoppzeit.*

Beweis: Siehe Korollar 6.7 und Satz 6.8 in [Ir03].

2.4 Martingale

Definition 2.15 *Sei X ein reellwertiger \mathcal{F} -adaptierter Prozess mit $E[|X_t|] < \infty$ für alle $t \geq 0$.*

- (a) *X heißt \mathcal{F} -Martingal, falls für alle $s, t \in I$ mit $s \leq t$ gilt:*

$$E[X_t | \mathcal{F}_s] = X_s \quad (P - f.s.).$$

- (b) *X heißt \mathcal{F} -Submartingal, falls für alle $s, t \in I$ mit $s \leq t$ gilt:*

$$E[X_t | \mathcal{F}_s] \geq X_s \quad (P - f.s.).$$

- (c) *X heißt \mathcal{F} -Supermartingal, falls für alle $s, t \in I$ mit $s \leq t$ gilt:*

$$E[X_t | \mathcal{F}_s] \leq X_s \quad (P - f.s.).$$

Ist die Filtration klar, so sagt man kurz „Martingal“ statt \mathcal{F} -Martingal.

2.5 Brownsche Bewegung

Ein wichtiges Beispiel eines Martingals für unsere Anwendungen in der Finanzmathematik ist die Brownsche Bewegung:

Definition 2.16 *Eine (eindimensionale) Brownsche Bewegung ist ein \mathbb{R} -wertiger Prozess $W = (W_t)_{t \geq 0}$ mit*

(W1) $W_0 = 0$ (P-f.s.),

(W2) W ist stetig,

(W3) $W_t - W_s$ ist unabhängig von \mathcal{F}_s , $t > s \geq 0$,

(W4) $W_t - W_s$ ist normalverteilt mit Mittelwert 0 und Varianz $t-s$, $t > s \geq 0$.

Seien W^1, \dots, W^n n unabhängige eindimensionale Brownsche Bewegungen. Dann heißt der \mathbb{R}^n -wertige stochastische Prozess $W = (W^1, \dots, W^n)$ eine n -dimensionale Brownsche Bewegung.

Der Name der Brownschen Bewegung geht zurück auf den Botaniker Brown (1928), der eine Zick-Zack-Bewegung eines Samenkorns in Flüssigkeit unter einem Mikroskop beobachtete (Deutung: Flüssigkeitsmoleküle stoßen das Samenkorn an; Idealisierung: viele, kleine, unabhängige, identisch verteilte Stöße ergeben als Summe die beobachtete Bewegung, die nach dem zentralen Grenzwertsatz damit näherungsweise normalverteilt ist). Der erste, der die Existenz der Brownschen Bewegung als stochastischen Prozess zeigte, war Wiener (1923), weshalb man auch von einem *Wiener-Prozess* spricht.

Die Brownsche Bewegung kann somit auch als Grenzwert diskreter Irrfahrten charakterisiert werden. Man betrachte ein Teilchen, das sich zur Zeit $t = 0$ an der Position $x = 0$ befindet und das sich in konstanten Zeitintervallen der Länge Δt mit gleicher Wahrscheinlichkeit um die Weglänge Δx nach links oder nach rechts bewegt. Man bezeichne mit X_k die Position des Teilchens nach k Sprüngen. Dann gilt $X_k = \sum_{i=1}^k Z_i$, wobei die unabhängigen Zufallsvariablen Z_i die Werte $\pm \Delta x$ jeweils mit Wahrscheinlichkeit $1/2$ annehmen. Eine elementare Rechnung liefert $E[X_k] = 0$ und $Var[X_k] = k\Delta t$. Unter der Bedingung $\Delta t = (\Delta x)^2 \rightarrow 0$ erhält man als Grenzwert die Brownsche Bewegung: Für $t = t_k$ liefert der Zentrale Grenzwertsatz

$$\lim_{k \rightarrow \infty} P(X_k \leq x) = \int_{-\infty}^x \frac{1}{\sqrt{2\pi t}} \exp\left(-\frac{y^2}{2t}\right) dy,$$

also die Normalverteilung der Zuwächse. (Für detailliertere Rechnungen siehe Vorlesung.)

Bemerkung: Die Pfade der Brownschen Bewegung sind (f.s.) an keiner Stelle differenzierbar und besitzen eine unendliche Variation (vgl. Kap. 4), somit kann man ohne Weiteres kein Integral im Riemann-Stieltjes Sinn bezüglich der Brownschen Bewegung definieren. Das stochastische Integral (Itô-Integral) wird in Kap. 6 eingeführt.

Satz 2.17 Die eindimensionale Brownsche Bewegung W ist ein Martingal bezüglich ihrer kanonischen Filtration \mathcal{F}^W .

Beweis: Siehe Vorlesung. □

Bemerkung 2.18 Sei W eine eindimensionale Brownsche Bewegung mit der kanonischer Filtration \mathcal{F}^W . Nach Satz 2.17 ist die Brownsche Bewegung mit Drift $\mu \in \mathbb{R}$ und Volatilität $\sigma \in \mathbb{R}$,

$$X_t = \mu t + \sigma W_t, \quad t \geq 0,$$

ein Martingal, falls $\mu = 0$, ein Submartingal, falls $\mu \geq 0$, und ein Supermartingal, falls $\mu \leq 0$.

Satz 2.19 Sei $W = (W_t)_{t \geq 0}$ eine eindimensionale Brownsche Bewegung mit der kanonischer Filtration \mathcal{F}^W . Dann ist die geometrische Brownsche Bewegung $X = (X_t)_{t \geq 0}$, definiert durch

$$X_t = \exp\left(\sigma W_t - \frac{1}{2}\sigma^2 t\right), \quad t \geq 0,$$

ein \mathcal{F}^W -Martingal.

Beweis: Siehe Übung. □

2.6 Poisson-Prozess

Definition 2.20 Ein Poisson-Prozess mit Intensität $\lambda > 0$ ist ein \mathbb{N}_0 -wertiger Prozess $N = (N_t)_{t \geq 0}$ mit

(N1) $N_0 = 0$ (P -f.s.),

(N2) N ist rechtsstetig und monoton wachsend auf \mathbb{R}_+ ,

(N3) $N_t - N_s$ ist unabhängig von \mathcal{F}_s , $t > s \geq 0$,

(N4) $N_t - N_s$ ist Poisson-verteilt mit Parameter $\lambda(t - s)$, $t > s \geq 0$.

Ein Poisson-Prozess N mit Intensität λ entspricht einem Zählprozess, der eine zufällige Anzahl von Ereignissen modelliert, die mit einer Intensität λ eintreten. Verändert sich diese Eintrittsintensität im Zeitverlauf, so hat man es mit einem *inhomogenen* Poisson-Prozess zu tun, siehe Übungen.

Eigenschaften von Poisson-Prozessen:

Satz 2.21 Für jeden Poisson Prozess gilt:

(i) $P(X_t - \lim_{s \nearrow t} X_s \in \{0, 1\} \text{ für alle } t > 0) = 1$, d.h. die Sprunghöhen sind f.s. gleich 1.

(ii) Für alle $t \geq 0$ gilt:

$$P(s \mapsto X_s \text{ stetig in } s = t) = 1.$$

(iii) $P(\lim_{t \rightarrow \infty} X_t = \infty) = 1$.

Beweis: Siehe z.B. [B93].

□

Satz 2.22 Sei N ein Poisson-Prozess mit der kanonischen Filtration \mathcal{F}^N . Dann ist der compound Poisson-Prozess $\tilde{N} = (\tilde{N}_t)_{t \geq 0}$, definiert durch

$$\tilde{N}_t = N_t - \lambda t, \quad t \geq 0,$$

ein \mathcal{F}^N -Martingal.

Beweis: Siehe Übung.

□

Kapitel 3

Markov-Prozesse

Markov-Prozesse bilden aufgrund ihrer Charakteristik eine wichtige Klasse von stochastischen Prozessen. Zum einen können durch Markov-Prozesse viele Phänomene modelliert werden. Zum anderen sind sie relativ einfach zu handhaben.

In diesem Kapitel sei (Ω, \mathcal{A}, P) ein Wahrscheinlichkeitsraum und I eine Indexmenge.

3.1 Definition von Markov-Prozessen

Definition 3.1 Sei \mathcal{F} eine Filtration. Ein stochastischer Prozess $X = (X_t)_{t \in I}$ mit Zustandsraum (S, \mathcal{B}) heißt \mathcal{F} -Markov-Prozess, falls X \mathcal{F} -adaptiert ist und für alle $s, t \in I$ mit $s \leq t$ gilt:

$$P(X_t \in B | \mathcal{F}_s) = P(X_t \in B | X_s) \text{ für alle } B \in \mathcal{B}.$$

Wir sprechen abkürzend von einem Markov-Prozess ist, wenn X ein \mathcal{F}^X -Markov-Prozess ist.

Deutung: Für die zukünftige Entwicklung eines Markov-Prozesses spielt bei Kenntnis des aktuellen Zustands die Vergangenheit keine Rolle.

Definition 3.2 Ein Markov-Prozess X mit Zustandsraum (S, \mathcal{B}) und Indexmenge $I = [0, \infty)$ heißt homogen, falls die Übergangsfunktion (von Zustand $x \in S$ in die Zustandsmenge $B \in \mathcal{B}$)

$$p_{x,B}(s, t) := P(X_t \in B | X_s = x), \quad 0 \leq s \leq t,$$

translationsinvariant ist, d.h.

$$p_{x,B}(s+h, t+h) = p_{x,B}(s, t) \text{ für alle } 0 \leq s \leq t, h \geq 0.$$

Deutung: Ein homogener Markov-Prozess zeichnet sich dadurch aus, dass die Wahrscheinlichkeit für den Übergang von Zustand x zum Zeitpunkt s in die Zustandsmenge $B \in \mathcal{B}$ zum Zeitpunkt t nur von der Zeitdifferenz $t - s$ abhängt.

Beispiele: Die Brownsche Bewegung und ein Poisson-Prozess mit konstanter Intensität sind homogene Markov-Prozesse.

3.2 Markov-Prozesse mit abzählbarem Zustandsraum

Im Folgenden sei S eine abzählbare Menge.

Satz 3.3 *Ein stochastischer Prozess $(X_t)_{t \in I}$ mit Zustandsraum (S, \mathcal{B}) ist genau dann ein Markov-Prozess, falls für alle $n > 1$, $t_1 < t_2 < \dots < t_{n+1} \in I$, $i_1, \dots, i_{n+1} \in S$ mit $P(X_{t_1} = i_1, \dots, X_{t_n} = i_n) > 0$ gilt:*

$$P(X_{t_{n+1}} = i_{n+1} | X_{t_n} = i_n, \dots, X_{t_1} = i_1) = P(X_{t_{n+1}} = i_{n+1} | X_{t_n} = i_n).$$

Beweis: Siehe Vorlesung. □

Für einen Markov-Prozess $X = (X_t)_{t \in I}$ mit Zustandsraum (S, \mathcal{B}) bezeichnen wir mit

$$p_{ij}(s, t) := P(X_t = j | X_s = i), \quad s \leq t \in I, \quad i, j \in S$$

die Übergangsfunktion von Zustand i in den Zustand j und mit

$$P(s, t) := (p_{ij}(s, t))_{i, j \in S}, \quad s \leq t \in I,$$

die (eventuell unendlichdimensionale) Übergangsmatrix. Ist X homogen (mit $I = [0, \infty)$), so schreiben wir

$$p_{ij}(h) = p_{ij}(0, h) \quad \text{und} \quad P(h) = P(0, h) \quad \text{für } h > 0.$$

Bemerkung: Für alle $s \leq t \in I$ ist die Übergangsmatrix $P(s, t)$ eine sogenannte stochastische Matrix, d.h. alle Einträge sind nicht-negativ und die Zeilensummen sind stets gleich eins.

Satz 3.4 (Chapman-Kolmogorov-Gleichung) *Sei $(X_t)_{t \in I}$ ein Markov-Prozess mit Zustandsraum (S, \mathcal{B}) und seien $s \leq t \leq u \in I$ und $i, k \in S$ mit $P(X_s = i) > 0$. Dann gilt*

$$p_{ik}(s, u) = \sum_{j \in S} p_{ij}(s, t) p_{jk}(t, u) \quad \text{bzw. in Matrix-Form: } P(s, u) = P(s, t) \cdot P(t, u).$$

Beweis: Siehe Vorlesung. □

Satz 3.5 Ein stochastischer Prozess X mit Zustandsraum (S, \mathcal{B}) ist genau dann ein Markov-Prozess, falls

$$P(X_{t_1} = i_1, \dots, X_{t_n} = i_n) = P(X_{t_1} = i_1) \prod_{k=1}^{n-1} p_{i_k i_{k+1}}(t_k, t_{k+1})$$

für alle $n > 1$, $t_1 < t_2 < \dots < t_n \in I$, $i_1, \dots, i_n \in S$.

Beweis: Siehe Vorlesung. □

Bemerkung: Nach Satz 3.5 ist die Verteilung eines Markov-Prozesses gegeben durch seine Startverteilung und seine Übergangswahrscheinlichkeiten.

3.3 Kolmogorovsche Differentialgleichung

In diesem Abschnitt sei $I = [0, \infty)$ und S eine endliche Menge.

Definition 3.6 Sei $X = (X_t)_{t \geq 0}$ ein Markov-Prozess in stetiger Zeit mit endlichem Zustandsraum (S, \mathcal{B}) . Dann heißt X regulär, falls die folgenden Grenzwerte existieren und stetig in den entsprechenden Variablen sind:

$$\begin{aligned} \mu_i(t) &= \lim_{h \searrow 0} \frac{1 - p_{ii}(t, t+h)}{h}, \quad i \in S, \\ \mu_{ij}(t) &= \lim_{h \searrow 0} \frac{p_{ij}(t, t+h)}{h}, \quad i \neq j \in S. \end{aligned}$$

Die Funktionen μ_i, μ_{ij} heißen Übergangsintensitäten von X . Ferner definieren wir μ_{ii} durch

$$\mu_{ii}(t) = -\mu_i(t), \quad i \in S.$$

Bemerkung 3.7 (i) Die Übergangsintensitäten können als Ableitungen der Übergangswahrscheinlichkeiten aufgefasst werden. So gilt z.B. für $i \neq j$:

$$\mu_{ij}(t) = \lim_{h \searrow 0} \frac{p_{ij}(t, t+h)}{h} = \lim_{h \searrow 0} \frac{p_{ij}(t, t+h) - p_{ij}(t, t)}{h} = \left. \frac{\partial}{\partial s} p_{ij}(t, s) \right|_{s=t}.$$

(ii) $\mu_{ij}(t)dt$ kann als infinitesimale Übergangswahrscheinlichkeit $i \rightarrow j$ im Zeitintervall $[t, t+dt]$ interpretiert werden, während $\mu_i(t)dt$ der infinitesimalen Wahrscheinlichkeit entspricht, im gegebenen Zeitintervall den Zustand i zu verlassen.

(iii) Es bezeichne $\Lambda(t) = (\mu_{ij}(t))_{i,j \in S}$. Für homogene Markov-Prozesse mit $\Lambda(t) = \Lambda$ gilt dann $P(t) = \exp(t\Lambda) = \sum_{n=0}^{\infty} \frac{t^n}{n!} \Lambda^n$. Die Matrix Λ bezeichnet man auch als infinitesimalen Generator der Halbgruppe $(P(t))_{t \geq 0}$, wobei die Halbgruppeneigenschaft durch die Chapman-Kolmogorov Gleichung $P(s+t) = P(s)P(t)$ gegeben ist..

Satz 3.8 (Kolmogorov) Sei $(X_t)_{t \geq 0}$ ein regulärer Markov-Prozess mit endlichem Zustandsraum (S, \mathcal{B}) . Dann gelten die folgenden Aussagen:

(i) (Rückwärts-Differentialgleichung)

$$\frac{\partial}{\partial s} p_{ij}(s, t) = \mu_i(s) p_{ij}(s, t) - \sum_{k \neq i} \mu_{ik}(s) p_{kj}(s, t),$$

oder in Matrix-Form

$$\frac{d}{ds} P(s, t) = -\Lambda(s) P(s, t).$$

(ii) (Vorwärts-Differentialgleichung)

$$\frac{\partial}{\partial t} p_{ij}(s, t) = -p_{ij}(s, t) \mu_j(t) + \sum_{k \neq j} p_{ik}(s, t) \mu_{kj}(t),$$

oder in Matrix-Form

$$\frac{d}{dt} P(s, t) = P(s, t) \Lambda(t).$$

Beweis: Siehe Vorlesung. □

Bemerkung: Die Kolmogorovschen Differentialgleichungen dienen primär dazu, die Übergangswahrscheinlichkeiten p_{ij} ausgehend von den Intensitäten μ_{ij} zu berechnen.

Definition 3.9 Sei $(X_t)_{t \geq 0}$ ein regulärer Markov-Prozess mit endlichem Zustandsraum (S, \mathcal{B}) . Für $0 \leq s \leq t$, $i \in S$ bezeichnen wir mit

$$\bar{p}_{ii}(s, t) := P \left(\bigcap_{u \in [s, t]} \{X_u = i\} \mid X_s = i \right)$$

die bedingte Wahrscheinlichkeit, im Zeitintervall $[s, t]$ immer im Zustand i zu bleiben.

Satz 3.10 Sei $(X_t)_{t \geq 0}$ ein regulärer Markov-Prozess mit endlichem Zustandsraum (S, \mathcal{B}) . Für $0 \leq s \leq t$ und $i \in S$ mit $P(X_s = i) > 0$ gilt

$$\bar{p}_{ii}(s, t) = \exp \left(- \sum_{j \neq i} \int_s^t \mu_{ij}(u) du \right) = \exp \left(- \int_s^t \mu_i(u) du \right).$$

Beweis: Siehe Vorlesung. □

Beispiele: Konstruktion von homogenen Markov-Prozessen mit endlichem Zustandsraum

(i) Sei Y_n eine Markov-Kette in diskreter Zeit mit Übergangswahrscheinlichkeiten $p(i, j)$ und $N(t)$ ein Poisson-Prozess mit Intensität λ . Dann ist $X_t := Y_{N(t)}$ ein Markov-Prozess in stetiger Zeit, welcher bei den Sprungzeiten von $N(t)$ einen Übergang gemäß $p(i, j)$ durchführt.

(ii) Gegeben sei der Zustand $X_t = i$.

- falls $\mu_i = 0$, so bleibt X für immer im Zustand i .
- falls $\mu_i > 0$, so ist nach Satz 3.10 die Wartezeit im Zustand i exponentialverteilt mit Parameter μ_i . Anschließend springt der Prozess im Zustand $j \neq i$ mit Wahrscheinlichkeit μ_{ij}/μ_i .

Kapitel 4

Riemann-Stieltjes Integral

Sei $I = [a, b]$ ein Intervall mit $-\infty < a < b < \infty$.

Definition 4.1 Eine endliche Folge $Z = \{a = x_0 < x_1 < \dots < x_n = b\}$ heißt Zerlegung von I mit Feinheit $\delta(Z) := \max\{x_i - x_{i-1} : i = 1, \dots, n\}$. Mit \mathcal{Z} bezeichnen wir die Menge aller Zerlegungen von I .

Definition 4.2 Sei $f : I \rightarrow \mathbb{R}$. Die totale Variation von f auf I wird definiert als

$$\|f\|_{TV,I} := \sup_{Z \in \mathcal{Z}} V_Z(f) := \sup_{Z \in \mathcal{Z}} \sum_{i=1}^n |f(x_i) - f(x_{i-1})|.$$

Falls $\|f\|_{TV,I} < \infty$ gilt, so schreiben wir $f \in BV(I)$ und nennen f von beschränkter Variation auf I .

Bemerkung 4.3 (i) Monotone Funktionen liegen in $BV(I)$.

(ii) Eine differenzierbare Funktion $f : I \rightarrow \mathbb{R}$ mit $f' \in L_1(I)$ (d.h. f' ist Lebesgue-integrierbar) liegt in $BV(I)$, und es gilt

$$\|f\|_{TV,I} = \int_a^b |f'(x)| dx.$$

(iii) Es gibt stetige Funktionen, die nicht von beschränkter Variation sind, z.B. liegt die stetige Funktion $f : x \mapsto x \cos(1/x)$ (mit $f(0) = 0$) nicht in $BV([0, 1])$.

(iv) $BV(I)$ ist ein linearer Vektorraum. Wegen $\|1\|_{TV,I} = 0$ ist $\|\cdot\|_{TV,I}$ im Allgemeinen keine Norm. Auf dem Raum aller Funktionen $f \in BV(I)$ mit $f(a) = 0$ dagegen ist $\|\cdot\|_{TV,I}$ eine Norm.

(v) Für $a < c < b$ gilt $\|f\|_{TV,[a,b]} = \|f\|_{TV,[a,c]} + \|f\|_{TV,[c,b]}$.

Satz 4.4 Eine Funktion $f : I \rightarrow \mathbb{R}$ ist genau dann von beschränkter Variation, wenn es monoton wachsende Funktionen $g, h : I \rightarrow \mathbb{R}$ gibt mit $f = g - h$.
Beweis: Siehe Vorlesung. \square

Definition 4.5 Sei $f \in \mathcal{C}([a, b])$ und $g \in BV([a, b])$. Dann ist das Riemann-Stieltjes Integral definiert durch

$$\int_a^b f(x) dg(x) := \lim_{\substack{Z \in \mathcal{Z} \\ \delta(Z) \rightarrow 0}} \sum_{i=1}^n f(\xi_i)(g(x_i) - g(x_{i-1})),$$

wobei $\xi_i \in [x_{i-1}, x_i]$ beliebige Zwischenpunkte seien.

Bemerkung 4.6 (i) Das Riemann-Stieltjes Integral ist wohldefiniert: Wegen Satz 4.4 genügt es, dies für einen monoton wachsenden Integrator g zu zeigen. Dann gilt

$$\begin{aligned} \sum_{i=1}^n \inf_{\xi \in [x_{i-1}, x_i]} f(\xi)(g(x_i) - g(x_{i-1})) &\leq \sum_{i=1}^n f(\xi_i)(g(x_i) - g(x_{i-1})) \\ &\leq \sum_{i=1}^n \sup_{\xi \in [x_{i-1}, x_i]} f(\xi)(g(x_i) - g(x_{i-1})). \end{aligned}$$

Ferner gilt

$$\begin{aligned} 0 &\leq \sum_{i=1}^n \left(\sup_{\xi \in [x_{i-1}, x_i]} f(\xi) - \inf_{\xi \in [x_{i-1}, x_i]} f(\xi) \right) (g(x_i) - g(x_{i-1})) \\ &\leq \sup_{|x-y| \leq \delta(Z)} |f(x) - f(y)| \sum_{i=1}^n (g(x_i) - g(x_{i-1})) \xrightarrow{\delta(Z) \rightarrow 0} 0. \end{aligned}$$

Damit folgt die Konvergenz der Riemann-Stieltjes Summen.

(ii) Für $g(x) = x$ erhält man das Riemann Integral.

(iii) Sei g differenzierbar mit $g' \in L_1([a, b])$. Dann gilt

$$\int_a^b f(x) dg(x) = \int_a^b f(x) g'(x) dx,$$

formal: $dg(x) = g'(x) dx$ (siehe Übungen).

(iv) Sei g eine Treppenfunktion, d.h. g habe in höchstens abzählbar vielen Punkten $\dots < x_i < x_{i+1} < \dots \in [a, b]$ Sprünge der Höhe $\Delta g_i \in \mathbb{R}$ und sei ansonsten konstant. Dann gilt

$$\int_a^b f(x) dg(x) = \sum_i f(x_i) \Delta g_i.$$

(v) Für $f, g \in BV([a, b]) \cap \mathcal{C}([a, b])$ gilt die Formel für die partielle Integration

$$\int_a^b f(x) dg(x) + \int_a^b g(x) df(x) = f(b)g(b) - f(a)g(a)$$

(siehe Übungen).

(vi) Das Riemann-Stieltjes Integral ist linear in f und g , d.h. für $\alpha \in \mathbb{R}$ gilt:

$$\begin{aligned} \int_a^b (f_1 + \alpha f_2) dg &= \int_a^b f_1 dg + \alpha \int_a^b f_2 dg, \\ \int_a^b f d(g_1 + \alpha g_2) &= \int_a^b f dg_1 + \alpha \int_a^b f dg_2, \end{aligned}$$

(vii) Wie üblich lassen sich uneigentliche Integrale der Form $\int_{-\infty}^{\infty} f(x) dg(x)$ durch den Übergang von $a \downarrow -\infty$ und $b \uparrow \infty$ erklären.

Definition 4.7 (i) Eine Funktion $f : [a, b] \rightarrow \mathbb{R}$ heißt absolut stetig, wenn für jedes $\varepsilon > 0$ ein $\delta > 0$ existiert, so dass für jedes disjunkte, höchstens abzählbare System von offenen Intervallen $(a_k, b_k) \subset [a, b]$ gilt

$$\sum_k (b_k - a_k) < \delta \Rightarrow \sum_k |f(b_k) - f(a_k)| < \varepsilon.$$

(ii) Eine Funktion $g : [a, b] \rightarrow \mathbb{R}$ heißt singular, falls $g'(x) = 0$ f.ü. gilt.

Bemerkung: Absolut stetige Funktionen sind speziell gleichmäßig stetig. Wie im nächsten Satz gezeigt wird, stellen die absolut stetigen Funktionen eine allgemeinere Klasse von Funktionen dar, für welche der Hauptsatz der Differential- und Integralrechnung gültig ist. Weiterhin sind Treppenfunktionen Beispiele singularer Funktionen. Die *Cantor-Funktion* (kumulative Verteilungsfunktion einer gleichverteilten ZV auf der Cantor-Menge) ist eine stetige, singuläre Funktion.

Satz 4.8 (Hauptsatz)

- (i) Für eine Funktion $f \in L_1([a, b])$ ist die Integralfunktion $F(x) := \int_0^x f(t)dt$, $x \in [a, b]$, absolut stetig und es gilt $F'(x) = f(x)$ f.ü.
- (ii) Eine absolut stetige Funktion f ist differenzierbar f.ü. mit $f' \in L_1([a, b])$ und man hat $f(x) - f(a) = \int_0^x f'(t)dt$.

Satz 4.9 Jede Funktion $f \in BV(I)$ besitzt eine eindeutige Zerlegung der Form $f = f_c + f_d$, wobei f_c absolut stetig ist mit $f_c(a) = f(a)$ und f_d eine singuläre Funktion ist.

Beweis: Siehe Vorlesung. □

Korollar 4.10 Sei $f \in C([a, b])$ und $g \in BV([a, b])$, so dass g_d eine Treppenfunktion ist. Das Riemann-Stieltjes Integral kann eindeutig zerlegt werden in

$$\int_a^b f(x)dg(x) = \int_a^b f(x)g'_c(x)dx + \sum_i f(x_i)\Delta g_i,$$

wobei $g'_c \in L_1([a, b])$ und die Summe über die Unstetigkeitsstellen x_i von g mit den Sprunghöhen Δg_i läuft. □

Mit dieser Darstellung können wir das Riemann-Stieltjes Integral für eine größere Klasse von Integranden verallgemeinern:

Definition 4.11 Sei $f \in L_1([a, b])$ beschränkt und $g \in BV([a, b])$, so dass g_d eine Treppenfunktion ist. Dann definieren wir

$$\int_a^b f(x)dg(x) := \int_a^b f(x)g'_c(x)dx + \sum_i f(x_i)\Delta g_i$$

mit den Notationen aus Korollar 4.10.

Bemerkung Für die Zwecke dieser Vorlesung reicht diese Definition aus, denn in die für den nächsten Kapitel relevanten Versicherungsmodellen ist dieser Wert gleich dem Wert des Riemann-Stieltjes Integrals $\int_a^b f dg$. Eine notwendige Bedingung zur Existenz dieses Integrals im Riemann-Stieltjes Sinn (vgl. Definition 4.5) ist, dass f und g keine gemeinsamen Unstetigkeitsstellen besitzen.

Kapitel 5

Thiele'sche Differentialgleichungen

Ziel des Kapitels ist die Modellierung einer Lebensversicherung durch zeitstetige stochastische Prozesse. Weiterhin wollen wir den Wert und den Preis einer Versicherung ermitteln. Dafür legen wir im Folgenden einen Wahrscheinlichkeitsraum (Ω, \mathcal{A}, P) zugrunde.

5.1 Reguläres Versicherungsmodell

Ausgangspunkt des Versicherungsmodells ist eine endliche Menge S von Zuständen, dessen Elemente die verschiedenen möglichen Zustände einer versicherten Person darstellen.

Beispiel 5.1 (i) Todesfallversicherung/Erlebensfallversicherung/Gemischte Versicherung: $S = \{*, \dagger\}$. Dabei steht $*$ für „Versicherungsnehmer lebt“, \dagger für „Versicherungsnehmer tot“.

(ii) Invaliditätsversicherung: $S = \{*, \dagger, \diamond\}$. Neben den schon eingeführten Bezeichnungen steht \diamond für „Versicherungsnehmer ist invalide“.

(iii) Versicherung auf zwei Leben: $S = \{(*, *), (*, \dagger), (\dagger, \dagger)\}$, wobei etwa $(*, \dagger)$ als „Versicherungsnehmer lebt und Ehefrau tot“ interpretiert werden kann.

Die Versicherung werde zum Zeitpunkt 0 abgeschlossen und laufe bis $T > 0$. Wir bezeichnen mit X_t den Zustand der Versicherung zum Zeitpunkt $t \in [0, T]$ und nehmen an, dass $X = (X_t)_{t \in [0, T]}$ ein regulärer Markov-Prozess mit Übergangintensitäten $\mu_{ij}(t)$ ist.

Die vertraglich vereinbarten Zahlungen werden wie folgt definiert:

Definition 5.2 Ein Zahlungsstrom ist ein adaptierter stochastischer Prozess $(A(t))_{t \in [0, T]}$, für welchen fast alle Pfade von beschränkter Variation sind.

Interpretativ ist $A(t)$ als die bis zum Zeitpunkt t ausgezahlte Geldmenge zu sehen.

Satz 5.3 Sei $(A(t))_{t \in [0, T]}$ ein Zahlungsstrom und sei $f : [0, T] \times \Omega \rightarrow \mathbb{R}$ eine $\mathcal{B}([0, T]) \otimes \mathcal{A}$ -messbare Funktion. Dann ist auch durch

$$B(t, \omega) = \int_0^t f(s, \omega) dA(s, \omega)$$

ein Zahlungsstrom gegeben (falls das R-S Integral für fast alle Pfade existiert). Formal schreiben wir

$$dB = f dA.$$

Beweis: Siehe Vorlesung. □

Für $i, j \in S$ führen wir die folgenden Hilfsfunktionen ein:

$$\begin{aligned} \mathbb{1}_i(t) &:= \mathbb{1}_{\{X_t=i\}}, \\ N_{i,j}(t) &:= \#\{s \in (0, t] : X_{s-} = i, X_s = j\}, \end{aligned}$$

d.h. $\mathbb{1}_i(t)$ ist die Indikatorfunktion für den Zustand i zum Zeitpunkt t und $N_{i,j}(t)$ misst die Anzahl der Sprünge von i nach j , die bis zum Zeitpunkt t aufgetreten sind.

Damit können wir die zufälligen Zahlungsströme unseres Lebensversicherungsmodells exakt formulieren: Für $i, j \in S, i \neq j$ seien $a_i, a_{i,j} : [0, T] \rightarrow \mathbb{R}$ Funktionen von beschränkter Variation. Ferner sei $\mathcal{D} = \{t_1, \dots, t_m\}$ mit $0 \leq t_1 < t_2 < \dots < t_m \leq T, m \in \mathbb{N}_0$, und $\Delta A_i : \mathcal{D} \rightarrow \mathbb{R}$ (Konvention: Setze $\Delta A_i(t) = 0$ für $t \notin \mathcal{D}$). Dann lauten die Lebensversicherungszahlungsströme:

$$\begin{aligned} dA_i(t, \omega) &= \mathbb{1}_i(t, \omega) (a_i(t) dt + \Delta A_i(t)), \\ dA_{ij}(t, \omega) &= a_{ij}(t) dN_{ij}(t, \omega), \\ dA &= \sum_{i \in S} dA_i + \sum_{i, j \in S: i \neq j} dA_{ij}. \end{aligned}$$

Dabei ist:

A_i : Der kumulierte Teilzahlungsstrom, der durch Zahlungen bei Verbleiben in Zustand i induziert wird. Hierbei unterscheiden wir zwischen kontinuierlichen Zahlungen mit Intensität a_i und diskreten Zahlungen $\Delta A_i(t)$ in den Zeitpunkten $t \in \mathcal{D}$.

A_{ij} : Der kumulierte Teilzahlungsstrom, der durch Zahlungen der Größe $a_{ij}(t)$ infolge eines Zustandswechsels von i nach j zum Zeitpunkt t induziert wird. Auf $[0, T]$ sind mehrere solche Übergänge möglich (genau zu den Zeitpunkten wo sich $N_{ij}(t)$ um 1 erhöht). Falls a_{ij} konstant ist, ergibt sich somit $A_{ij}(t) = a_{ij}N_{ij}(t)$.

A : Der kumulierte Gesamtzahlungsstrom.

Bemerkung: Die obigen Integrale sind nach Definition 4.11 wohldefiniert.

Beispiel 5.4 (Invaliditätsversicherung) *Der Versicherungsnehmer sei bei Vertragsabschluss 40 Jahre. Ist der Versicherungsnehmer aktiv, so erfolge eine Prämienzahlung der Intensität c , solange er das 65. Lebensjahr nicht erreicht. Ist der Versicherungsnehmer invalide, so erhält er eine Rente mit Intensität b . Stirbt er, so erfolgt eine Zahlung der Höhe a an seine Erben. Ist der Versicherungsschutz lebenslanglich und sind keine weiteren Leistungen oder Prämien vereinbart, so erhält man*

$$\begin{aligned} a_*(t) &= -c\mathbb{1}_{[0,25)}(t), a_\diamond(t) = b, a_{*\dagger}(t) = a_{\diamond\dagger}(t) = a, \\ a_\dagger(t) &= a_{*\diamond}(t) = a_{\diamond*}(t) = a_{\dagger*}(t) = a_{\dagger\diamond}(t) = 0, \end{aligned}$$

also

$$\begin{aligned} A(t) &= -c \int_0^{25 \wedge t} \mathbb{1}_*(s) ds + b \int_0^t \mathbb{1}_\diamond(s) ds + a \int_0^t (dN_{*\dagger}(s) + dN_{\diamond\dagger}(s)). \\ &= -c \int_0^{25 \wedge t} \mathbb{1}_*(s) ds + b \int_0^t \mathbb{1}_\diamond(s) ds + a \mathbb{1}_{\{\tau \leq t\}}(t), \end{aligned}$$

wobei τ der (zufällige) Todeszeitpunkt der versicherten Person ist (aus einem der Zustände $*$ oder \diamond).

Den Zins modellieren wir durch eine rechtsstetige Zinsintensität $r : [0, T] \rightarrow \mathbb{R}_+$, d.h.

$$Y(t+s) = \exp\left(\int_t^{t+s} r(u) du\right) Y(t),$$

wobei $Y(t)$ den Wert eines Bankkontos zur Zeit t bezeichnet. Der entsprechende Diskontierungsfaktor (von $t+s$ nach t) ist dann gegeben durch

$$B(t, t+s) = \exp\left(-\int_t^{t+s} r(u) du\right).$$

Mit dem Diskontierungsfaktor kann der Barwert zukünftiger Leistungen berechnet werden: $Y(t) = B(t, t+s)Y(t+s)$. Falls der Zins stochastisch modelliert wird, ist die Zinsintensität r ein stochastischer Prozess $(r(t))_{t \in [0, T]}$. Der Einfachheit halber gehen wir im Folgenden von einer konstanten Zinsintensität r aus. Bei einem Jahreszins von i lässt sich diese wie folgt berechnen:

$$e^r = 1 + i \Leftrightarrow r = \ln(1 + i).$$

Definition 5.5 *Ein solches Modell, bestehend aus einem Wahrscheinlichkeitsraum (Ω, \mathcal{A}, P) mit*

1. *einer regulären Markovkette $(X_t)_{t \in [0, T]}$ mit endlichem Zustandsraum S ,*
2. *Auszahlungsfunktionen a_i, a_{ij} von beschränkter Variation und diskrete Auszahlungsfunktionen ΔA_i auf der Menge $\mathcal{D} = \{t_1, \dots, t_m\}$,*
3. *einer rechtsstetigen Zinsintensität r ,*

nennen wir reguläres Versicherungsmodell.

5.2 Deckungskapital

Eine zentrale Aufgabe in der Lebensversicherung ist die Berechnung des Deckungskapitals, d.h. des Geldbetrages, der für jede Police reserviert werden muss, um allen zukünftigen Leistungen gerecht zu werden.

Definition 5.6 *Für ein reguläres Versicherungsmodell ist*

$$V_i(t) := E \left[\int_t^T B(t, s) dA(s) \middle| X_t = i \right]$$

das Deckungskapital, wenn sich die Versicherung zum Zeitpunkt $t \in [0, T]$ im Zustand $i \in S$ befindet.

Bemerkungen: Das Deckungskapital entspricht dem erwarteten Barwert aller zukünftigen Zahlungen und ist von der Versicherungsgesellschaft auf der Passivseite der Bilanz als Reserve auszuweisen. Das Deckungskapital ist ferner für den Versicherungsnehmer relevant, wenn er seinen Vertrag vorzeitig kündigen möchte. Für eine konstante Zinsintensität r ergibt sich die Formel

$$V_i(t) := E \left[\int_t^T e^{-r(s-t)} dA(s) \middle| X_t = i \right].$$

Darüber hinaus müssen nach dem sog. *Äquivalenzprinzip* Versicherungsleistungen und Prämien so kalkuliert werden, dass $V_{X_0}(0) = 0$ gilt, d.h. auf $[0, T]$ soll der erwartete Wert der Einnahmen gleich dem erwarteten Wert der Ausgaben sein.

Um die Prämien und das Deckungskapital der Versicherung zu bestimmen, benötigen wir die folgenden Formeln für Erwartungswerte. Dazu schreiben wir für die Übergangswahrscheinlichkeiten

$$p_{ij}(s, t) = P(X_t = j | X_s = i), \quad s \leq t, \quad i, j \in S.$$

Satz 5.7 Sei $t \in [0, T]$ und $i, j, k \in S$. Dann gelten die folgenden Aussagen:

(i) Für $a \in L_1([0, T])$ gilt

$$E \left[\int_t^T a(s) dN_{jk}(s) \middle| X_t = i \right] = \int_t^T a(s) p_{ij}(t, s) \mu_{jk}(s) ds.$$

(ii) Für $A \in BV([0, T])$ gilt

$$E \left[\int_t^T \mathbf{1}_j(s) dA(s) \middle| X_t = i \right] = \int_t^T p_{ij}(t, s) dA(s).$$

Beweis: Siehe Vorlesung. □

Beispiel 5.8 (Fortsetzung von Beispiel 5.4) Wir wollen die Prämie c der Invaliditätsversicherung aus Beispiel 5.4 bestimmen. Wir unterstellen ein Maximalalter von 120 Jahren (Ende der Sterbetafel), d.h. $T = 120 - 40 = 80$. Das Äquivalenzprinzip liefert

$$\begin{aligned} 0 &\stackrel{!}{=} E \left[\int_0^T e^{-rt} dA(t) \middle| X_0 = * \right] \\ &= -cE \left[\int_0^{25} e^{-rt} \mathbf{1}_*(t) dt \middle| X_0 = * \right] + bE \left[\int_0^{80} e^{-rt} \mathbf{1}_\diamond(t) dt \middle| X_0 = * \right] \\ &\quad + aE \left[\int_0^{80} e^{-rt} (dN_{*\dagger}(t) + dN_{\diamond\dagger}(t)) \middle| X_0 = * \right]. \end{aligned}$$

Damit erhält man mit Satz 5.7

$$c = \frac{b \int_0^{80} e^{-rt} p_{*\diamond}(0, t) dt + a \int_0^{80} e^{-rt} (p_{**}(0, t) \mu_{*\dagger}(t) + p_{*\diamond}(0, t) \mu_{\diamond\dagger}(t)) dt}{\int_0^{25} e^{-rt} p_{**}(0, t) dt}.$$

Wir können nun eine Differentialgleichung für das Deckungskapital herleiten:

Satz 5.9 (Thiele'sche Differentialgleichungen) *Die Funktionen V_i erfüllen das folgende gekoppelte System von Differentialgleichungen:*

$$\begin{aligned}\frac{\partial}{\partial t}V_i(t) &= rV_i(t) - a_i(t) - \sum_{j \neq i} \mu_{ij}(t)(a_{ij}(t) + V_j(t) - V_i(t)), \quad t \notin \mathcal{D}, \\ V_i(t) &= V_i(t-) - \Delta A_i(t), \quad t \in \mathcal{D}, \\ V_i(T) &= 0,\end{aligned}$$

wobei die Zinsintensität r als konstant angenommen wird. Beweis: Siehe Vorlesung. \square

Diese Differentialgleichungen wurden 1875 von Thiele für eine einfache Versicherungsart bewiesen.

Beispiel 5.10 (Invaliditätsversicherung mit endlicher Laufzeit) *Wir betrachten ein Invaliditätsmodell mit endlicher Laufzeit T und den folgenden Vertragsbedingungen:*

- Der Versicherungsnehmer zahlt zur Zeit t eine Prämie der Intensität c , falls $t \leq T$ und $X_t = *$.
- Der Versicherungsnehmer bekommt zur Zeit t eine Rente der Intensität b , falls $t \leq T$ und $X_t = \diamond$.
- Stirbt der Versicherungsnehmer vor T , so bekommen seine Erben ein Kapital der Höhe a .
- Nach dem Tod oder nach Eintreten des Zeitpunktes T erlöschen alle Zahlungsverpflichtungen.

Wegen $\mu_{\dagger*} = \mu_{\dagger\diamond} = 0$ erhalten wir unmittelbar

$$V_{\dagger}(t) = 0.$$

Für V_*, V_{\diamond} lauten die Thiele'schen DGL

$$\begin{aligned}\frac{\partial}{\partial t}V_*(t) &= rV_*(t) + c - \mu_{*\diamond}(t)(V_{\diamond}(t) - V_*(t)) - \mu_{*\dagger}(t)(a - V_*(t)), \\ \frac{\partial}{\partial t}V_{\diamond}(t) &= rV_{\diamond}(t) - b - \mu_{\diamond*}(t)(V_*(t) - V_{\diamond}(t)) - \mu_{\diamond\dagger}(t)(a - V_{\diamond}(t))\end{aligned}$$

mit den Endbedingungen

$$V_*(T) = V_{\diamond}(T) = 0.$$

Im Prinzip kann man nun numerische Methoden verwenden, um zu einer Approximation der Lösung zu gelangen. Offenbar sind die Gleichungen jedoch linear. Daher lässt sich theoretisch die Lösung mit der Variation der Konstanten bestimmen. Wir illustrieren dies im Folgenden für den Fall, dass es keine Reaktivierung gibt, d.h. $\mu_{\diamond*} = 0$. Dann entkoppelt sich das System und mit $\mu_{\diamond} = \mu_{\diamond*} + \mu_{\diamond\dagger} = \mu_{\diamond\dagger}$ erhalten wir für V_{\diamond} die DGL

$$\frac{\partial}{\partial t} V_{\diamond}(t) = (r + \mu_{\diamond}(t))V_{\diamond}(t) - b - \mu_{\diamond}(t)a.$$

Mit der Methode der Variation der Konstanten können wir nun eine Lösung ausrechnen:

Homogene Gleichung:

$$\frac{\partial}{\partial t} \varphi(t) = (r + \mu_{\diamond}(t))\varphi(t).$$

Lösung der homogenen Gleichung:

$$\varphi(t) = \exp\left(\int_T^t (r + \mu_{\diamond}(s))ds\right).$$

Ansatz für V_{\diamond} :

$$V_{\diamond}(t) = \varphi(t) \int_T^t \frac{-b - \mu_{\diamond}(s)a}{\varphi(s)} ds + C\varphi(t).$$

Die Endbedingung $V_{\diamond}(T) = 0$ liefert $C = 0$. Wegen $\bar{p}_{\diamond\diamond}(t, s) = \exp(-\int_t^s \mu_{\diamond}(u)du)$ (vgl. Satz 3.10) folgt

$$V_{\diamond}(t) = \int_t^T e^{-r(s-t)} \bar{p}_{\diamond\diamond}(t, s) (b + \mu_{\diamond}(s)a) ds.$$

Analog:

$$V_*(t) = \int_t^T e^{-r(s-t)} \bar{p}_{**}(t, s) (-c + \mu_{*\diamond}(s)V_{\diamond}(s) + \mu_{*\dagger}(s)a) ds.$$

Kapitel 6

Stochastisches Integral

6.1 Konstruktion des stochastischen Integrals

Gegeben sei ein Wahrscheinlichkeitsraum (Ω, \mathcal{A}, P) mit einer Filtration $\mathcal{F} = (\mathcal{F}_t)_{t \in [0, T]}$, die die üblichen Bedingungen erfüllt. Ferner sei $W = (W_t)_{t \in [0, T]}$ eine (zunächst eindimensionale) \mathcal{F} -adaptierte Brownsche Bewegung.

Definition 6.1 Ein stochastischer Prozess $X = (X_t)_{t \in [0, T]}$ heißt einfacher Prozess, falls

$$X_t(\omega) = f_0(\omega) \cdot \mathbf{1}_{\{0\}}(t) + \sum_{i=1}^n f_i(\omega) \cdot \mathbf{1}_{(t_{i-1}, t_i]}(t)$$

für $0 = t_0 < t_1 < t_2 < \dots < t_n = T$, wobei f_i für $i = 1, \dots, n$ beschränkte $\mathcal{F}_{t_{i-1}}$ -messbare Zufallsvariablen sind und f_0 \mathcal{F}_0 -messbar ist.

Bemerkungen

- (i) X_t ist somit $\mathcal{F}_{t_{i-1}}$ -messbar für $t \in (t_{i-1}, t_i]$. Diese Forderung an X (*Vorhersehbarkeit*) begründet sich z.B. aus der Modellierung einer Handelsstrategie mit Aktien, deren Stückpreis gemäß W_t verläuft. Zu Beginn des Zeitintervalls $(t_{i-1}, t_i]$ kauft man f_i Aktien, welche man am Ende dieser Periode wieder verkauft. Der entstandene Gewinn ist somit $f_i(W_{t_i} - W_{t_{i-1}})$. Die „Strategie“ f_i muss $\mathcal{F}_{t_{i-1}}$ -messbar sein, da sie nicht vom späteren Preisverlauf der Aktien ($t > t_{i-1}$) beeinflusst werden darf.
- (ii) Die Pfade des Prozesses X sind linksstetig. Diese Forderung spielt eine Rolle bei der Konstruktion des stochastischen Integrals bezüglich rechtsstetigen Integratoren. Die Begründung ist ähnlich wie oben: Man kann/darf nicht die Strategie *gleichzeitig* mit dem Sprung des Aktienpreises modifizieren um sie daran anzupassen.

Es wird sich herausstellen, dass die obigen Annahmen u.a. die Martingaleigenschaft für das stochastische Integral, welches im Folgenden definiert wird, implizieren.

Definition 6.2 Für einen einfachen Prozess $X = (X_t)_{t \in [0, T]}$ definiert man das stochastische Integral durch:

$$I_t(X) := \int_0^t X_s dW_s := \sum_{i=1}^k f_i(W_{t_i} - W_{t_{i-1}}) + f_{k+1}(W_t - W_{t_k}) \quad (6.1)$$

für $t \in (t_k, t_{k+1}]$.

Satz 6.3 Es gelten folgende Eigenschaften:

(i) $(I_t(X))_{t \in [0, T]}$ ist ein stetiges \mathcal{F} -Martingal. Es gilt somit $E[I_t(X)] = 0$ für $t \in [0, T]$.

(ii) $E[I_t(X)^2] = E\left[\int_0^t X_s^2 ds\right]$ für $t \in [0, T]$.

(iii) Linearität: $I_t(aX + bY) = aI_t(X) + bI_t(Y)$.

Beweis: Siehe Vorlesung. □

Das stochastische Integral wird nun auf geeignete Grenzwerte stochastischer Prozesse fortgesetzt.

Definiere:

$$L^2[0, T] := \{(X_t)_{t \in [0, T]} : X \text{ progressiv messbar bzgl. } \mathcal{F} \text{ mit } \|X\|_T < \infty\}, \quad (6.2)$$

wobei

$$\|X\|_T^2 := E\left[\int_0^T X_t^2 dt\right]. \quad (6.3)$$

$L^2[0, T]$ entspricht dem Raum $([0, T] \times \Omega, \mathcal{B}([0, T]) \otimes \mathcal{A}, \lambda \otimes P)$ mit der üblichen L^2 -Norm (λ bezeichnet das Lebesgue-Maß). Dabei identifizieren wir einen Prozesse X als einen Prozess Y , wenn $X = Y$ $\lambda \otimes P$ -f.s. gilt. (X ist äquivalent zu Y).

Sei M_c^2 der Raum der stetigen, quadratintegrierbaren Martingale bezüglich \mathcal{F} . Für einen einfachen Prozess X ist nach Satz 6.3 das Integral $I_t(X) \in M_c^2$ und wegen

$$\|I_t(X)\|_{L^2} := E[I_t^2(X)] = E\left[\int_0^t X_s^2 ds\right] = \|X\|_t^2$$

ist die Abbildung $X \mapsto I(x)$ eine Isometrie vom $L^2[0, T]$ -Unterraum der einfachen Prozesse nach M_c^2 (Itô-Isometrie).

Im Folgenden wird gezeigt, dass sich die Itô-Isometrie auf $L^2[0, T]$ fortsetzen lässt: Approximiere $X \in L^2[0, T]$ durch einfache Prozesse X_n und betrachte die Folge der stochastischen Integrale $I(X_n)$. Diese erweist sich als eine Cauchy-Folge und den entsprechenden Grenzwert bezeichnet man als $I_t(X) := \int_0^t X_s dW_s$. Die genaue Formulierung dieser Konstruktion ergibt sich aus den folgenden Sätzen.

Satz 6.4 *Sei $X \in L^2[0, T]$. Dann existiert eine Folge X_n einfacher Prozesse mit $\|X_n - X\|_T \rightarrow 0$ für $n \rightarrow \infty$.*

Beweis: Siehe Vorlesung. □

Satz 6.5 *Es existiert eine lineare Abbildung $J : L^2[0, T] \rightarrow M_c^2$ mit den Eigenschaften:*

- (i) *Für einfache Prozesse X sind die Prozesse $I_t(X)$ und $J_t(X)$ ununterscheidbar.*
- (ii) *$E[J_t(X)^2] = E[\int_0^t X_s^2 ds]$ (Itô-Isometrie).*
- (iii) *J ist eindeutig: Falls J' ebenfalls (i) und (ii) erfüllt, so sind die Prozesse $J'(X)$ und $J(X)$ ununterscheidbar.*

Beweis: Siehe Vorlesung. □

Definition 6.6 *Für $X \in L^2[0, T]$ und J aus Satz 6.5 definiere durch*

$$\int_0^t X_s dW(s) := J_t(X) \tag{6.4}$$

das stochastische Integral (Itô-Integral) von X (bzgl. W).

Satz 6.7 *(Doob'sche Ungleichung)*

- (i) *Sei $(M_t)_{t \in [0, T]}$ ein Martingal mit rechtsstetigen Pfaden und $E[M(t)^2] < \infty$ für alle $t \in [0, T]$. Dann gilt:*

$$E\left[\sup_{t \in [0, T]} |M_t|\right]^2 \leq 4E[M_T^2].$$

(ii) Für $X \in L^2[0, T]$ gilt:

$$E \left[\left(\sup_{t \in [0, T]} \left| \int_0^t X_s dW_s \right| \right)^2 \right] \leq E \left[\int_0^T X_s^2 ds \right].$$

Beweis: Siehe Vorlesung. □

Definition 6.8 Sei $(W(t))_{t \in [0, T]}$ eine m -dimensionale Brownsche Bewegung bezüglich der Filtration \mathcal{F} und $(X(t))_{t \in [0, T]}$ ein progressiv messbarer, $\mathbb{R}^{n \times m}$ -wertiger Prozess mit $X_{ij} \in L^2[0, T]$ für $i = 1, \dots, n$, $j = 1, \dots, m$. Definiere

$$\int_0^t X(s) dW(s) := \begin{pmatrix} \sum_{j=1}^m \int_0^t X_{1j}(s) dW_j(s) \\ \vdots \\ \sum_{j=1}^m \int_0^t X_{nj}(s) dW_j(s) \end{pmatrix}.$$

Das stochastische Integral wird nun auf eine größere Klasse von Prozessen fortgesetzt.

Definiere

$$H^2[0, T] := \{(X_t)_{t \in [0, T]} : X \text{ progr. messbar bzgl. } \mathcal{F} \text{ mit } \int_0^T X_t^2 dt < \infty \text{ P-f.s.}\}.$$

Da (im Unterschied zu $L^2[0, T]$) für $X \in H^2[0, T]$ möglicherweise $E[\int_0^T X_t^2 dt] = \infty$ gelten kann, sind diese Prozesse nicht durch einfache Prozesse im L^2 -Sinne approximierbar. Eine Erweiterung des stochastischen Integrals erhält man jedoch durch *Lokalisierung*:

Betrachte die Stoppzeiten

$$\tau_n(\omega) := T \wedge \inf\{t \in [0, T] : \int_0^t X_s^2 ds \geq n\}$$

und definiere den gestoppten Prozess $X_n = (X_n(t))_{t \in [0, T]}$,

$$(X_n)(t, \omega) := X(t, \omega) \cdot \mathbb{1}_{\{t \leq \tau_n(\omega)\}}.$$

Damit folgt $X_n \in L^2[0, T]$ und somit existiert das stochastische Integral $J_t(X_n)$.

Definiere für $X \in H^2[0, T]$: $J_t(X) := J_t(X_n)$ für $0 \leq t \leq \tau_n$. Die Definition ist konsistent, denn aufgrund der Eigenschaften von X_n gilt $J_t(X_m) = J_t(X_n)$ für $m \geq n$ und $0 \leq t \leq \tau_n \leq \tau_m$. Außerdem gilt $\tau_n \rightarrow T$ P-f.s. für $n \rightarrow \infty$.

Der Prozess $J_t(X)$ ist für $X \in H^2[0, T]$ ein sogenanntes *lokales Martingal* mit *lokalisierender Folge* (τ_n) . Die Linearität und Stetigkeit der Pfade bleiben erhalten, aber wegen möglicher unendlicher Erwartungswerte kann man von einer Itô-Isometrie nicht mehr sprechen.

6.2 Die Itô-Formel

Definition 6.9 (Itô-Prozess)

Sei $(W(t))_{t \geq 0}$ eine m -dimensionale, \mathcal{F} -adaptierte Brownsche Bewegung.

(i) Ein Prozess der Form

$$\begin{aligned} X(t) &= X(0) + \int_0^t K(s)ds + \int_0^t H(s)dW(s) \\ &= X(0) + \int_0^t K(s)ds + \sum_{j=1}^m \int_0^t H_j(s)dW_j(s), \end{aligned} \quad (6.5)$$

wobei $X(0)$ \mathcal{F}_0 -messbar ist und K, H progressiv messbare Prozesse sind mit $\int_0^t |K(s)|ds < \infty$ und $\int_0^t H_i^2(s)ds < \infty$ P -f.s. für alle $t \geq 0, i = 1, \dots, m$, heißt eindimensionaler Itô-Prozess. Symbolische Schreibweise:

$$dX(t) = K(t)dt + H(t)dW(t).$$

(ii) Ein Prozess der Form $X = (X_1, \dots, X_n)$, wobei X_i eindimensionale Itô-Prozesse sind, heißt n -dimensionaler Itô-Prozess.

Bemerkungen

- (i) Es gilt $H_j \in H^2[0, T]$ für alle $T > 0, j = 1, \dots, m$.
- (ii) Die Darstellung in Definition 6.9 (i) ist eindeutig. Falls diese auch mit K', H' gilt, so sind K', K bzw. H', H ununterscheidbar.

Definition 6.10 (i) Für zwei Itô-Prozesse bezüglich der gleichen Brownschen Bewegung

$$\begin{aligned} dX(t) &= K(t)dt + H(t)dW(t), \\ dY(t) &= L(t)dt + M(t)dW(t) \end{aligned}$$

heißt

$$\langle X, Y \rangle_t := \sum_{i=1}^m \int_0^t H_i(s)M_i(s)ds$$

die quadratische Kovariation von X und Y . Insbesondere heißt $\langle X \rangle_t := \langle X, X \rangle_t$ die quadratische Variation von X .

(ii) Sei X ein Itô-Prozess mit $dX(t) = K(t)dt + H(t)dW(t)$ und Y ein progressiv messbarer, eindimensionaler Prozess. Definiere

$$\int_0^t Y(s)dX(s) := \int_0^t Y(s)K(s)ds + \int_0^t Y(s)H(s)dW(s)$$

falls die Integrale auf der rechten Seite existieren. Symbolisch schreibt man:

$$Y(t)dX(t) = Y(t)K(t)dt + Y(t)H(t)dW(t).$$

Beispiel: Für die eindimensionale Brownsche Bewegung $W(t)$ gilt $\langle W \rangle_t = \int_0^t 1^2 ds = t$.

Satz 6.11 (Eindimensionale Itô-Formel)

Sei $(W(t))_{t \geq 0}$ eine eindimensionale Brownsche Bewegung und X ein reellwertiger Itô-Prozess mit $dX(t) = K(t)dt + H(t)dW(t)$. Für eine Funktion $f \in C^2(\mathbb{R})$ und alle $t \geq 0$ gilt

$$df(X(t)) = f'(X(t))dX(t) + \frac{1}{2}f''(X(t))d\langle X \rangle_t.$$

Genauer: Es gilt

$$\begin{aligned} f(X(t)) &= f(X(0)) + \int_0^t f'(X(s))K(s)ds + \int_0^t f'(X(s))H(s)dW(s) \\ &\quad + \frac{1}{2} \int_0^t f''(X(s))H^2(s)ds \end{aligned}$$

P-f.s.

Beweis: Siehe Vorlesung. □

Satz 6.12 (Mehrdimensionale Itô-Formel)

Sei $(W(t))_{t \geq 0}$ eine m -dimensionale Brownsche Bewegung und X ein n -dimensionaler Itô-Prozess mit der Darstellung $dX_i(t) = K_i(t)dt + \sum_{j=1}^m H_{ij}(t)dW_j(t)$, $i = 1, \dots, n$. Für eine $C^{1,2}$ -Funktion $f : [0, \infty) \times \mathbb{R}^n \rightarrow \mathbb{R}$ (C^1 in t und C^2 in x) gilt dann:

$$\begin{aligned} f(t, X(t)) &= f(0, X(0)) + \int_0^t \frac{\partial f}{\partial t}(s, X(s))ds + \sum_{i=1}^n \int_0^t \frac{\partial f}{\partial x_i}(s, X(s))dX_i(s) \\ &\quad + \frac{1}{2} \sum_{i,j=1}^n \int_0^t \frac{\partial^2 f}{\partial x_i \partial x_j}(s, X(s))d\langle X_i, X_j \rangle_s. \end{aligned}$$

Korollar 6.13 (*Produktregel, partielle Integration*)

Für eindimensionale Itô-Prozesse bezüglich der gleichen Brownschen Bewegung

$$\begin{aligned}dX(t) &= K(t)dt + H(t)dW(t), \\dY(t) &= L(t)dt + M(t)dW(t)\end{aligned}$$

gilt

$$\begin{aligned}dX(t)Y(t) &= X(t)dY(t) + Y(t)dX(t) + \langle X, Y \rangle_t \\ &= (X(t)L(t) + Y(t)K(t) + H(t)M(t))dt \\ &\quad + (X(t)M(t) + Y(t)H(t))dW(t).\end{aligned}$$

Beispiele: (Anwendungen der Itô-Formel)

- (i) Für einen deterministischen Prozess $dX(t) = K(t)dt$ und $f \in C^2$ gilt $df(X(t)) = f'(X(t))dX(t)$, also die bekannte Kettenregel.
- (ii) Für die Brownsche Bewegung $W(t)$ und $f(x) = x^2$ gilt:

$$d(W^2(t)) = 2W(t)dW(t) + \frac{1}{2}2d\langle W \rangle_t = 2W(t)dW(t) + dt,$$

d.h.:

$$\int_0^t W(s)dW(s) = \frac{1}{2}(W^2(t) - t).$$

- (iii) Man betrachte eine Aktie deren Preisprozess durch

$$P(t) = P(0) \exp\left(\left(b - \frac{1}{2}\sigma^2\right)t + \sigma W(t)\right)$$

(geometrische Brownsche Bewegung) gegeben ist. Man hat somit

$$P(t) = P(0) \cdot \underbrace{\exp(bt)}_{\substack{\text{determ. Trend} \\ \text{(Verzinsung)}}} \cdot \underbrace{\exp\left(-\frac{1}{2}\sigma^2 t + \sigma W(t)\right)}_{\substack{\text{trendloser Term} \\ \text{(Martingal)}}}.$$

Hintergrund: Der Preisprozess muss stets positive Werte annehmen. Daher ist die Brownsche Bewegung für eine solche Modellierung nicht geeignet. Deswegen betrachtet man die Exponentialfunktion eines Itô-Prozesses. Die Martingaleigenschaft der zufälligen Schwankung um den mittleren Kurs $\exp(bt)$ erreicht man durch Hinzunahme des Terms $-\frac{1}{2}\sigma^2 t$.

Mit dem Itô-Prozess $dX(t) = (b - \frac{1}{2}\sigma^2)dt + \sigma dW(t)$ gilt also $P(t) = f(X(t))$ für $f(x) = P(0) \exp(x)$. Wegen $f' = f = f''$ liefert die Itô-Formel:

$$\begin{aligned} dP(t) &= f(X(t))dX(t) + \frac{1}{2}f(X(t))d\langle X \rangle_t \\ &= P(t)(b - \frac{1}{2}\sigma^2)dt + P(t)\sigma dW(t) + \frac{1}{2}P(t)\sigma^2 dt \\ &= b \cdot P(t)dt + \sigma \cdot P(t)dW(t). \end{aligned} \tag{6.6}$$

Diese Gleichung für P ist ein Beispiel einer sogenannten *stochastischen Differentialgleichung*, welche im nächsten Kapitel untersucht werden. Die durchgeführte Rechnung zeigt also, dass der Prozess

$$P(t) = P(0) \exp((b - \frac{1}{2}\sigma^2)t + \sigma W(t))$$

eine Lösung der stochastischen Differentialgleichung

$$dP(t) = b \cdot P(t)dt + \sigma \cdot P(t)dW(t)$$

ist.

Kapitel 7

Stochastische Differentialgleichungen

7.1 Problemformulierung

Sei (Ω, \mathcal{A}, P) ein vorgegebener Wahrscheinlichkeitsraum und $T > 0$. Zu lösen ist die *stochastische Differentialgleichung* (SDGL)

$$X(t) = X(0) + \int_0^t b(s, X(s))ds + \int_0^t \sigma(s, X(s))dW(s),$$

wobei

$$b : [0, T] \times \mathbb{R}^n \rightarrow \mathbb{R}^n, \quad \sigma : [0, T] \times \mathbb{R}^n \rightarrow \mathbb{R}^{n \times m}$$

messbare Funktionen und W eine m -dimensionale Brownsche Bewegung bezüglich einer Filtration \mathcal{F} seien. In der symbolischen Differentialschreibweise lautet die Gleichung

$$dX(t) = b(t, X(t))dt + \sigma(t, X(t))dW(t). \quad (7.1)$$

Eine Lösung $X = (X(t))_{t \in [0, T]}$ von (7.1) ist also ein n -dimensionaler stochastischer Prozess, der

$$X(t) = X(0) + \int_0^t b(s, X(s))ds + \int_0^t \sigma(s, X(s))dW(s) \quad \text{für alle } t \in [0, T]$$

fast sicher erfüllt. Wir sprechen von einer eindeutigen Lösung, wenn für zwei Lösungen X, \tilde{X} pfadweise Eindeutigkeit vorliegt, d.h. $P(X(t) = \tilde{X}(t) \text{ für alle } t \in [0, T]) = 1$.

Üblicherweise wünschen wir uns eine Lösung zu einer gegebenen Anfangsbedingung $X(0) = \xi$, wobei ξ eine \mathcal{F} -messbare, von W unabhängige Zufallsvariable bezeichne.

7.2 Existenz und Eindeutigkeit

Satz 7.1 Seien b, σ Lipschitz-stetig in x , d.h. es existiere eine Konstante $C > 0$ mit

$$|b(t, x) - b(t, y)| + |\sigma(t, x) - \sigma(t, y)| \leq C|x - y|, \quad t \in [0, T], x, y \in \mathbb{R}^n$$

(wobei $|\sigma|^2 := \sum |\sigma_{ij}|^2$), und erfüllen zusätzlich die Wachstumsbedingung

$$|b(t, x)| + |\sigma(t, x)| \leq D(1 + |x|), \quad t \in [0, T], x \in \mathbb{R}^n$$

für eine Konstante $D > 0$. Ferner sei ξ eine von W unabhängige Zufallsvariable mit $E[|\xi|^2] < \infty$. Dann hat die stochastische Differentialgleichung (7.1) mit der Anfangsbedingung $X(0) = \xi$ eine eindeutige stetige Lösung $X = (X(t))_{t \in [0, T]}$ mit den Eigenschaften:

X ist adaptiert bezüglich der von ξ und W erzeugten Filtration (erweitert um Nullmengen)

und

$$E \left[\int_0^T |X(t)|^2 dt \right] < \infty.$$

Beweis: Siehe Theorem 5.2.1 in [Ok00]. □

7.3 Starke Markoveigenschaft

Im Folgenden betrachten wir die zeitlich homogene SDGL

$$dX(t) = b(X(t))dt + \sigma(X(t))dW(t), \quad (7.2)$$

zu einer m -dimensionalen Brownschen Bewegung W und messbaren Funktionen $b : \mathbb{R}^n \rightarrow \mathbb{R}^n, \sigma : \mathbb{R}^n \rightarrow \mathbb{R}^{n \times m}$, die die Bedingungen von Satz 7.1 erfüllen, d.h. für Konstanten $C, D > 0$ gilt

$$\begin{aligned} |b(x) - b(y)| + |\sigma(x) - \sigma(y)| &\leq C|x - y|, \quad x, y \in \mathbb{R}^n, \\ |b(x)| + |\sigma(x)| &\leq D(1 + |x|), \quad x \in \mathbb{R}^n. \end{aligned}$$

Sei X eine Lösung von (7.2). Der Prozess X wird auch als *Itô-Diffusion* mit Driftkoeffizientem $b(x)$ und Diffusionskoeffizientem $\sigma(x)$ bezeichnet. (In der Literatur wird manchmal auch $a = \sigma\sigma^T(x)$ oder $\frac{1}{2}\sigma\sigma^T(x)$ als Diffusionskoeffizient bezeichnet.)

Notation: Im Folgenden schreiben wir abkürzend $E_x[\cdot]$ für den bedingten Erwartungswert $E[\cdot | X(0) = x]$.

Satz 7.2 Sei $f : \mathbb{R}^n \rightarrow \mathbb{R}$ beschränkt und messbar.

(i) Dann gilt für fast alle $\omega \in \Omega$

$$E_x[f(X(t+s))|\mathcal{F}_t](\omega) = E_{X(t,\omega)}[f(X(s))], \quad t, s \geq 0.$$

(ii) Ist τ eine Stoppzeit mit $\tau < \infty$ f.s., so gilt für fast alle $\omega \in \Omega$

$$E_x[f(X(\tau+s))|\mathcal{F}_\tau](\omega) = E_{X(\tau,\omega)}[f(X(s))], \quad s \geq 0.$$

Beweis: Siehe Theoreme 7.12 und 7.2.4 in [Ok00]. □

Teil (i) wird als *Markoveigenschaft*, Teil (ii) als *starke Markoveigenschaft* bezeichnet. Die Abhängigkeit von ω wird üblicherweise nicht notiert. So ist unter $E_{X(t,\omega)}[f(X(s))]$ die Zufallsvariable $g(X(t))$ mit

$$g : \mathbb{R}^n \rightarrow \mathbb{R}, \quad g(x) = E_x[f(X(s))]$$

zu verstehen.

7.4 Generator

Sei $dX(t) = b(X(t))dt + \sigma(X(t))dW(t)$ eine Itô-Diffusion.

Definition 7.3 Definiere den Differentialoperator A durch

$$(Af)(x) = \sum_{i=1}^n b_i(x) \frac{\partial f}{\partial x_i} + \frac{1}{2} \sum_{i,j=1}^n (\sigma\sigma^T)_{ij}(x) \frac{\partial^2 f}{\partial x_i \partial x_j}.$$

für Funktionen $f \in C^2(\mathbb{R}^n)$.

Satz 7.4 (Dynkinsche Formel) Sei X eine Itô-Diffusion und $f \in \mathcal{C}_b^2(\mathbb{R}^n)$. Dann gilt:

(i) Der Prozess

$$M(t) = f(X(t)) - \int_0^t (Af)(X(s))ds$$

ist ein Martingal bezüglich der Filtration \mathcal{F} .

(ii) Für eine Stoppzeit τ mit $E_x[\tau] < \infty$. gilt:

$$E_x[f(X(\tau))] = f(X(0)) + E_x\left[\int_0^\tau (Af)(X(s))ds\right].$$

Falls τ die Erstaustrittszeit aus einer beschränkten Menge B ist, so gilt die Formel in (ii) für beliebiges $f \in \mathcal{C}^2(\mathbb{R}^n)$, da $f|_B$ außerhalb von B geeignet fortgesetzt werden kann.

Satz 7.5 *Sei X eine Itô-Diffusion. Dann gilt*

$$(Af)(x) = \lim_{t \searrow 0} \frac{E_x[f(X(t))] - f(x)}{t}$$

für alle Funktionen $f \in \mathcal{C}_0^2(\mathbb{R}^n)$.

Bemerkung: Für Markov-Prozesse $X(t)$ ist die Familie von Operatoren $(T(t))_{t \geq 0}$ definiert durch $T(t)f(x) := E_x[f(X(t))]$, für beschränkte, messbare Funktionen f , eine *Halbgruppe*, d.h. es gilt $T(s+t) = T(s)T(t)$. In der Tat, die Markov-Eigenschaft impliziert

$$E_x[f(X(s+t)) | \mathcal{F}_s] = E_{X(s)}[f(X(t))] = T(t)f(X(s)).$$

Damit sind die Erwartungswerte $E_x[\cdot]$ beider Seiten gleich:

$$E_x[f(X(s+t))] = E_x[(T(t)f)(X(s))],$$

also

$$T(s+t)f(x) = T(s)(T(t)f(x))$$

d.h. es gilt die Halbgruppeneigenschaft. Es handelt sich um eine *Kontraktionshalbgruppe*: $\|T(t)f\|_\infty \leq \|f\|_\infty$, denn $|E_x[f(X(t))]| \leq E_x[|f(X(t))|] \leq \|f\|_\infty$.

Den Limes

$$(Af)(x) = \lim_{t \searrow 0} \frac{T(t)f(x) - T(0)f(x)}{t}$$

bezeichnet man auch als *infinitesimalen Generator* der Halbgruppe $T(t)$ (und somit des Markov-Prozesses X). Dessen Definitionsbereich \mathcal{D}_A besteht aus allen Funktionen f , für welche der obige Grenzwert existiert. Satz 7.5 besagt also, dass die \mathcal{C}_0^2 -Funktionen im Definitionsbereich des infinitesimalen Generators der Itô-Diffusion liegen, und für diese Funktionen stimmt der Generator mit dem Operator A aus Definition 7.3 überein.

7.5 Fokker-Planck-Gleichung

Sei X eine Itô-Diffusion und μ_t die Verteilung von $X(t)$, d.h. $\mu_t(f) := \int_{\Omega} f d\mu_t = E[f(X(t))]$ für $f \in C_b$.

Satz 7.6 Für $f \in C_b$ gilt

$$\mu_t(f) = \mu_0(f) + \int_0^t \mu_s(Af) ds.$$

Beweis: Siehe Vorlesung. □

Definiere den adjungierten Operator A^* durch

$$(A^*f)(x) := - \sum_{i=1}^n \frac{\partial}{\partial x_i} (b_i(x)f(x)) + \frac{1}{2} \sum_{i,j=1}^n \frac{\partial^2}{\partial x_i \partial x_j} (a_{ij}(x)f(x))$$

für $a_{ij}(x) = \sum_{k=1}^n \sigma_{ik}(x)\sigma_{jk}(x)$ (also $a = \sigma\sigma^T$).

Satz 7.7 (Fokker-Planck-Gleichung)

Falls die Verteilung der Itô-Diffusion eine Dichte $p(t, x) \in C^{1,2}$ besitzt (C^1 in t , C^2 in x), so gilt die Gleichung:

$$\frac{\partial}{\partial t} p(t, x) = (A^*p)(t, x), \quad t \geq 0, x \in \mathbb{R}^n$$

mit der Anfangsbedingung $p(0, x) = p_0(x)$ f.s., wobei p_0 die Verteilungsdichte von $X(0)$ ist.

Beweis: Siehe Vorlesung. □

7.6 Feynman-Kac-Formel

Gegeben sei die partielle Differentialgleichung

$$\begin{cases} \frac{\partial u}{\partial t}(t, x) = \frac{1}{2} \Delta u(t, x) + c(x)u(t, x) + f(x), & t > 0, x \in \mathbb{R}^n, \\ u(0, x) = g(x), & x \in \mathbb{R}^n \end{cases}$$

mit $c, f, g \in C_b(\mathbb{R}^n)$. Die Lösung dieser Gleichung besitzt eine Darstellung mit Hilfe der n -dimensionalen Itô-Diffusion $dX(t) = dW(t)$ (X unterscheidet sich von der Standard-Brownschen Bewegung nur durch die Anfangsbedingung $X(0) = x$):

Satz 7.8 (Feynman-Kac-Formel)

Es gilt

$$u(t, x) = E_x \left[g(X(t)) \exp\left(\int_0^t c(X(s)) ds\right) + \int_0^t f(X(s)) \exp\left(\int_0^s c(X(r)) dr\right) ds \right].$$

Beweis: Siehe Vorlesung. \square

Mit Hilfe von allgemeinen Itô-Diffusionen der Form $dX(s) = b(X(s))ds + \sigma(X(s))dW(s)$ für $s \geq t$ mit $X(t) = x$ kann man die Lösung folgender partiellen Differentialgleichung mit Endbedingung darstellen:

$$\begin{cases} \left(\frac{\partial u}{\partial t} + Au - cu\right)(t, x) = f(x), & t \in [0, T], x \in \mathbb{R}^n, \\ u(T, x) = g(x), & x \in \mathbb{R}^n \end{cases}$$

mit $f, g, c \in C_b$ und dem infinitesimalen Generator A aus Definition 7.3.

Satz 7.9 (Feynman-Kac-Formel)

Es gilt

$$u(t, x) = E_x \left[\beta_{t,T} \cdot g(X(T)) - \int_t^T \beta_{t,s} \cdot f(X(s)) ds \right]$$

mit

$$\beta_{t,s} = \exp\left(-\int_t^s c(X(r)) dr\right).$$

Beweis: Siehe Vorlesung. \square

Bemerkung: Aufgrund der zeitlichen Homogenität des Prozesses X kann man die Lösung auch durch

$$u(t, x) = E_x \left[\beta_{0, T-t} \cdot g(X(T-t)) - \int_0^{T-t} \beta_{0, s-t} \cdot f(X(s)) ds \right]$$

darstellen.

Kapitel 8

Black-Scholes Differentialgleichung

8.1 Optionen

Definition 8.1 *Eine europäische Call-Option (bzw. Put-Option) ist ein Vertrag mit den folgenden Bedingungen: Zu einem bestimmten Zeitpunkt T hat der Käufer der Option das Recht, aber nicht die Verpflichtung, einen Basiswert zu einem vorher festgelegten Ausübungspreis vom Verkäufer der Option zu erhalten (bzw. an dem Verkäufer der Option zu verkaufen).*

Beispiele

- (i) Europäischer Standard-Call auf eine Aktie mit Preisprozess $S(t)$ und Ausübungspreis K :
- Falls $S(T) > K$ gilt, so kann der Käufer der Option die Aktie zum Preis K erwerben und sie sofort zum Marktpreis $S(T)$ verkaufen. Der entstandene Gewinn beträgt $S(T) - K$.
 - Falls $S(T) < K$ gilt, so verfällt die Option, denn die Aktie kann man zum günstigeren Marktpreis $S(T)$ erwerben.

Die *Auszahlungsfunktion* zum Zeitpunkt T für den Käufer der Option lautet also $\max(S(T) - K, 0) =: (S(T) - K)^+$.

- (ii) Für einen europäischen Standard-Put auf eine Aktie mit Preisprozess $S(t)$ und Ausübungspreis K beträgt zum Fälligkeitstermin T die Auszahlungsfunktion $(K - S(T))^+$.

Bemerkung: Die Optionen in (i) und (ii) werden auch *plain vanilla*-Optionen genannt. Allgemeinere europäische Optionen dürfen nur am vorher festgelegten Zeitpunkt T ausgeübt werden und besitzen Auszahlungsfunktionen der Form $\Lambda(S(T))$.

Definition 8.2 Eine amerikanische Call-Option (bzw. Put-Option) ist ein Vertrag, welcher dem Käufer der Option das Recht verbrieft, zu einem beliebigen Zeitpunkt $t \in [0, T]$ einen Basiswert zu einem vorher festgelegten Ausübungspreis vom Verkäufer der Option zu erhalten (bzw. an dem Verkäufer der Option zu verkaufen).

Bemerkung: Bei den amerikanischen Standard-Optionen welche zum Zeitpunkt $t \in [0, T]$ ausgeübt werden beträgt die Auszahlungsfunktion

$$\begin{aligned} \text{Call: } \Lambda(S(t)) &= (S(t) - K)^+, \\ \text{Put: } \Lambda(S(t)) &= (K - S(t))^+. \end{aligned}$$

Ein zentrales Problem in diesem Kapitel ist die Bestimmung eines angemessenen Preises für die Standard-Optionen, so dass keine Arbitrage-Möglichkeiten (risikoloser Gewinn ohne Eigenkapital) entstehen.

Zur Illustration betrachte man folgendes

Beispiel: Bewertung einer europäischen Call-Option in einem Ein-Perioden-Modell.

Betrachte ein festverzinsliches Wertpapier mit Preis $B(t)$ (Bond) und eine Aktie mit Preis $S(t)$, wobei $t \in \{0, T\}$ nur zwei Werte annimmt. Man betrachte die Startwerte $B(0) = S(0) = 100$. Mit der vereinfachenden Annahme, dass für den Zinssatz $r = 0$ gilt, folgt $B(T) = 100$. Für $S(T)$ kommen die Werte $S(T) = 120$ (Aktienkurs steigt) bzw. $S(T) = 80$ (Aktienkurs fällt) in Frage. Die entsprechenden Wahrscheinlichkeiten seien unbekannt.

Betrachte einen europäischen Standard-Call mit Ausübungspreis $K = 100$. Der Wert der Option zum Zeitpunkt T beträgt dann:

$$C(T) = (S(T) - 100)^+ = \begin{cases} 20, & \text{falls } S(T) = 120 \\ 0, & \text{falls } S(T) = 80. \end{cases}$$

Der Wert $C(0)$ der Option zum Zeitpunkt $t = 0$ wird mit Hilfe des *Duplikationsprinzips* ermittelt: Man konstruiert ein Portfolio bestehend aus c_1 Bond-Anteile und c_2 Aktienanteile, welches zu jedem Zeitpunkt t den Wert der Option widerspiegelt: $C(t) = c_1 B(t) + c_2 S(t)$ für $t \in \{0, T\}$.

Für $t = T$ gilt also das System:

$$\begin{cases} 100c_1 + 120c_2 &= 20 \\ 100c_1 + 80c_2 &= 0 \end{cases}$$

mit Lösung $c_1 = -\frac{2}{5}$ und $c_2 = \frac{1}{2}$ (negative Anteile bedeuten Verkauf).

Damit ergibt sich $C(0) = -\frac{2}{5}B(0) + \frac{1}{2}S(0) = 10$.

Es wird nun gezeigt, dass ein von $c_1B(0) + c_2S(0)$ verschiedener Preis für die Option zu Arbitragemöglichkeiten führt.

Sei der Preis der Option $\tilde{P} \neq c_1B(0) + c_2S(0)$.

- Falls $\tilde{P} < c_1B(0) + c_2S(0)$ (Preis zu niedrig) ergibt sich eine Arbitragemöglichkeit für den Käufer der Option:
 - $t = 0$: Der Kauf der Option für den Preis \tilde{P} wird durch Leerverkauf eines Portfolios bestehend aus c_1 Bonds und c_2 Aktien finanziert. Es ergibt sich ein positiver Geldfluss der Höhe $c_1B(0) + c_2S(0) - \tilde{P} > 0$.
 - $t = T$: Durch Ausübung der Option erhält der Optionskäufer den Betrag $C(T)$. Aufgrund der Eigenschaft $C(T) = c_1B(T) + c_2S(T)$ kann er mit dieser Einnahme die geschuldeten c_1 Bonds und c_2 Aktien finanzieren, d.h. der kumulierte Zahlungsstrom ist gleich 0.

Dem Käufer der Option bleibt somit der in $t = 0$ erzielte risikolose Gewinn.

- Falls $\tilde{P} > c_1B(0) + c_2S(0)$ (Preis zu hoch) ergibt sich eine Arbitragemöglichkeit für den Verkäufer der Option:
 - $t = 0$: Verkaufe die Option für \tilde{P} und kaufe die Position (c_1, c_2) , die weniger kostet. Es bleibt der Betrag $\tilde{P} - c_1B(0) - c_2S(0)$ als Gewinn übrig.
 - $t = T$: Die Einzahlung durch den Verkauf des Portfolios beträgt $c_1B(T) + c_2S(T) = (S(T) - K)^+$ und egalisiert somit die mit der verkauften Option verbundene Auszahlung.

Es werden folgende Eigenschaften des Finanzmarktes vorausgesetzt:

- Arbitragefreiheit,
- keine Dividendenzahlungen auf den Basiswert
- einen konstanten Zinssatz r sowohl für Geldanlagen, als auch für Kredite,
- Liquidität (Handel ist zu jedem Zeitpunkt möglich),

- die Wertpapiere sind beliebig teilbar,
- alle betrachteten stochastischen Prozesse sind stetig (keine Börsen-crashes),

Betrachte europäische und amerikanische Standard-Optionen mit Fälligkeit T und Ausübungspreis K auf eine Aktie mit Preisprozess $S(t)$. Es werden folgende Bezeichnungen eingeführt:

$$\begin{aligned}
 C_E(t) &:= C_E(t, S(t)) && \text{Wert des europ. Calls zur Zeit } t \in [0, T], \\
 P_E(t) &:= P_E(t, S(t)) && \text{Wert des europ. Puts zur Zeit } t \in [0, T], \\
 C_A(t) &:= C_A(t, S(t)) && \text{Wert des amerik. Calls zur Zeit } t \in [0, T], \\
 P_A(t) &:= P_A(t, S(t)) && \text{Wert des amerik. Puts zur Zeit } t \in [0, T].
 \end{aligned}$$

Aus den Eigenschaften des Finanzmarktes, insbesondere aus der Annahme der Arbitrage-Freiheit, lassen sich dann folgende Beziehungen herleiten:

Satz 8.3 (*Put-Call Parität*) *Es gilt:*

$$P_E(t) = C_E(t) + Ke^{-r(T-t)} - S(t).$$

Beweis: Siehe Vorlesung. □

Satz 8.4 *Es gelten die Schranken:*

$$\begin{aligned}
 (i) \quad & (S(t) - Ke^{-r(T-t)})^+ \leq C_E(t) \leq S(t). \\
 (ii) \quad & (Ke^{-r(T-t)} - S(t))^+ \leq P_E(t) \leq Ke^{-r(T-t)}.
 \end{aligned}$$

Beweis: Siehe Vorlesung. □

Satz 8.5 *Es gilt:*

$$\begin{aligned}
 (i) \quad & P_A(t) \geq P_E(t). \\
 (ii) \quad & C_A(t) = C_E(t). \\
 (iii) \quad & Ke^{-r(T-t)} \leq S(t) + P_A(t) - C_A(t) \leq K. \\
 (iv) \quad & (Ke^{-r(T-t)} - S(t))^+ \leq P_A(t) \leq K.
 \end{aligned}$$

Beweis: Siehe Übung. □

Bemerkung: Falls Dividendenzahlungen auf den Basiswert zugelassen werden, so stimmen der europäische und amerikanische Call nicht mehr überein. Aufgrund der größeren Flexibilität in deren Ausübung sind amerikanische Optionen generell teurer als entsprechende europäische Optionen.

8.2 Herleitung der Black-Scholes Gleichung

In diesem Abschnitt wird eine partielle Differentialgleichung für den Wert einer europäischen Option auf eine Aktie hergeleitet. Es werden die Voraussetzungen an dem Finanzmarkt aus dem vorherigen Abschnitt angenommen. Man betrachtet folgende Prozesse:

- ein Bond (risikolose Geldanlage) mit Dynamik $dB(t) = rB(t)dt$.
- eine Aktie mit Preisprozess $S(t)$ mit der Dynamik

$$dS(t) = bS(t)dt + \sigma S(t)dW(t)$$

(vgl. Beispiel im Abschnitt 6.2) mit den Konstanten $b \in \mathbb{R}$, $\sigma \geq 0$ und der eindimensionalen Brownschen Bewegung W .

Die Konstante b kann als die *subjektive* mittlere Wachstumsrate des Aktienkurses betrachtet werden. Dieser Wert ist unbekannt und entspricht den Einschätzungen der Investoren auf den Markt. Es gilt $b > r$, sonst würde man ausschließlich in den Bond investieren. Die Konstante σ wird auch als *Volatilität* bezeichnet. Je höher die Volatilität, desto größere Kursschwankungen sind zu erwarten. Der Wert von σ kann anhand der Marktdaten geschätzt werden. Der Wert von b ist unbekannt, trotzdem ist eine Bewertung der Option durch eine Duplikationsstrategie möglich, analog zum Beispiel des Ein-Perioden-Modells aus dem vorherigen Abschnitt.

Sei $V(t, S(t))$ der Wert der Option zum Zeitpunkt t . Die Laufzeit betrage T und die Auszahlungsfunktion $\Lambda(S)$. Diese bestimmt dann den Typ der Option, z.B. liegt für $\Lambda(S) = (S - K)^+$ ein europäischer Standard-Call vor, während für einen europäischen Standard-Put $\Lambda(S) = (K - S)^+$ gilt.

8.2.1 Herleitung mit Hilfe des Duplikationsprinzips

Man betrachte das Portfolio

$$Y(t) := c_1(t)B(t) + c_2(t)S(t) - V(t, S(t))$$

welches aus $c_1(t)$ Bond-Anteile, $c_2(t)$ Aktienanteile und einer verkauften Option besteht.

Die *Handelsstrategie* $(c_1(t), c_2(t))$ ist so zu bestimmen, dass folgende Eigenschaften gelten:

- 1) Das Portfolio Y ist *risikolos*, d.h. es unterliegt keinen zufälligen Schwankungen. Es gilt daher

$$dY(t) = rY(t)dt, \quad (8.1)$$

d.h. das Portfolio kann nur soviel wie der risikolose Bond erwirtschaften (sonst entstehen Arbitragemöglichkeiten).

- 2) Das Portfolio ist *selbstfinanzierend*, d.h. alle Umschichtungen werden nur aus Käufen und Verkäufen von Positionen des Portfolios finanziert. Dies wird durch die Gleichung

$$dY(t) = c_1(t)dB(t) + c_2(t)dS(t) - dV(t, S(t)) \quad (8.2)$$

ausgedrückt.

Die Anwendung der Itô-Formel für $V(t, S(t))$ liefert

$$dV(t, S(t)) = \left(\frac{\partial V}{\partial t} + bS \frac{\partial V}{\partial S} + \frac{1}{2} \sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} \right) dt + \sigma S \frac{\partial V}{\partial S} dW(t)$$

(wobei wir der Übersichtlichkeit halber auf die Abhängigkeit der Ableitungen von $(t, S(t))$ verzichten).

Nun können wir die Dynamiken $dB(t)$, $dS(t)$ und $dV(t, S(t))$ in die Gleichung (8.2) einsetzen und erhalten:

$$\begin{aligned} dY(t) = & \left[c_1 r B + c_2 b S - \left(\frac{\partial V}{\partial t} + bS \frac{\partial V}{\partial S} + \frac{1}{2} \sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} \right) \right] dt \\ & + \left(c_2 \sigma S - \sigma S \frac{\partial V}{\partial S} \right) dW(t). \end{aligned}$$

Gleichsetzen mit (8.1) liefert

$$c_1 r B + c_2 b S - \left(\frac{\partial V}{\partial t} + bS \frac{\partial V}{\partial S} + \frac{1}{2} \sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} \right) = r(c_1 B + c_2 S - V)$$

und

$$c_2 \sigma S - \sigma S \frac{\partial V}{\partial S} = 0.$$

Die Terme mit c_1 aus der ersten Gleichung verschwinden und aus der zweiten Gleichung ergibt sich

$$c_2(t) = \frac{\partial V}{\partial S}(t, S(t)).$$

Einsetzen in die erste Gleichung ergibt

$$bS \frac{\partial V}{\partial S} - \left(\frac{\partial V}{\partial t} + bS \frac{\partial V}{\partial S} + \frac{1}{2} \sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} \right) = rS \frac{\partial V}{\partial S} - rV$$

und nach Umformen folgt die *Black-Scholes Gleichung* für die Funktion $V(t, S)$:

$$\frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} + rS \frac{\partial V}{\partial S} - rV = 0 \quad (8.3)$$

mit der Endbedingung

$$V(T, S) = \Lambda(S).$$

$V(t, S)$ ist Optionspreis zur Zeit $t \in [0, T]$ wenn für den Basiswert $S(t) = S$ gilt. Man stellt fest, dass diese Gleichung und somit der Optionspreis nicht von der unbekanntem (subjektiven) mittleren Wachstumsrate b des Aktienkurses abhängt.

Die Black-Scholes Gleichung (8.3) ist für $S \in [0, \infty)$ zu lösen. Um die Eindeutigkeit der Lösung zu gewährleisten, sind zusätzliche Randbedingungen in $S = 0$ und $S \rightarrow \infty$ erforderlich.

Sei zunächst $V = C$ der Wert einer Call-Option. Falls $S(t) = 0$ (Basiswert ist wertlos), so ist das Recht, diesen Basiswert zu kaufen, ebenfalls wertlos. Es gilt also $C(t, 0) = 0$. Für $S(t) \rightarrow \infty$ wird die Option mit hoher Wahrscheinlichkeit ausgeübt, d.h. der Wert des Calls ist näherungsweise gleich $C(t, S) \approx S - Ke^{-r(T-t)}$.

Sei nun $V = P$ der Wert einer Put-Option. Für $S(t) \rightarrow \infty$ wird die Option mit hoher Wahrscheinlichkeit nicht ausgeübt, d.h. $P(S, t) \rightarrow 0$ für $S \rightarrow \infty$. Für $S = 0$ folgt aus der Put-Call-Parität (Satz 8.3) $P(t, 0) = Ke^{-r(T-t)}$.

8.2.2 Herleitung mit Hilfe der risikoneutralen Bewertung

Zur Veranschaulichung dieses Zugangs betrachten wir zunächst wieder ein Ein-Perioden-Modell.

Es sei $t \in \{0, T\}$ und man betrachtet eine Aktie mit $S(0) = S_0$. $S(T)$ nimmt (mit unbekanntem Wahrscheinlichkeiten) einen der Werte $S_0 \cdot u$ oder $S_0 \cdot d$ mit $d < u$ ein. Sei $V(t)$ der Preis einer Option mit Auszahlungsfunktion $\Lambda(S)$. Zum Zeitpunkt T gilt also $V(T) = \Lambda(S(T))$. Die Betrachtung eines Bonds ist in diesem Fall nicht notwendig. Unser Ziel ist die Bestimmung eines sogenannten *risikoneutralen Wahrscheinlichkeitsmaßes* Q , so dass sich für den Optionspreis $V(0)$ die Formel

$$V(0) = E_Q[e^{-rT}V(T)] \quad (8.4)$$

ergibt, wobei E_Q den Erwartungswert unter Q bezeichnet. Nach (8.4) ist der Optionspreis genau der Erwartungswert (unter Q) der diskontierten Auszahlungsfunktion. Es reicht aus, Q nur auf den Ereignissen zu bestimmen, die

für dieses Modell von Interesse sind, und zwar $q := Q(S(T) = S_0 \cdot u)$ (und somit $1 - q = Q(S(T) = S_0 \cdot d)$).

Man konstruiert sich wieder ein risikoloses (deterministisches) Portfolio der Form

$$Y(t) = cS(t) - V(t).$$

Somit muss der Wert $Y(T)$ konstant sein, unabhängig vom Kurs der Aktie. Es gilt also

$$cS_0 \cdot u - \Lambda(S_0 \cdot u) = cS_0 \cdot d - \Lambda(S_0 \cdot d)$$

woraus

$$c = \frac{\Lambda(S_0 \cdot u) - \Lambda(S_0 \cdot d)}{S_0 \cdot u - S_0 \cdot d}$$

folgt.

Aus Arbitragegründen folgt, dass dieses Portfolio den gleichen Ertrag wie ein Bond unter dem Zinssatz r erwirtschaftet, es gilt also $Y(0) = e^{-rT}Y(T)$, oder

$$cS_0 - V(0) = e^{-rT}(cS(T) - \Lambda(S(T))) = e^{-rT}(cS_0 \cdot u - \Lambda(S_0 \cdot u)).$$

Damit ergibt sich der Optionspreis zum Zeitpunkt $t = 0$:

$$\begin{aligned} V(0) &= cS_0 - e^{-rT}(cS_0 \cdot u - \Lambda(S_0 \cdot u)) \\ &= e^{-rT}[c(S_0 e^{rT} - S_0 \cdot u) + \Lambda(S_0 \cdot u)] \\ &= \frac{e^{-rT}}{u - d}[(\Lambda(S_0 \cdot u) - \Lambda(S_0 \cdot d))(e^{rT} - u) + \Lambda(S_0 \cdot u)(u - d)] \\ &= \frac{e^{-rT}}{u - d}[\Lambda(S_0 \cdot u)(e^{rT} - d) + \Lambda(S_0 \cdot d)(u - e^{rT})] \\ &= e^{-rT}\left[\Lambda(S_0 \cdot u)\frac{e^{rT} - d}{u - d} + \Lambda(S_0 \cdot d)\frac{u - e^{rT}}{u - d}\right] \\ &= e^{-rT}[\Lambda(S_0 \cdot u) \cdot q + \Lambda(S_0 \cdot d) \cdot (1 - q)] \end{aligned}$$

für

$$q = \frac{e^{rT} - d}{u - d}.$$

Es gilt also

$$V(0) = E_Q[e^{-rT}V(T)].$$

Analog zeigt man

$$S(0) = E_Q[e^{-rT}S(T)].$$

Unter dem Wahrscheinlichkeitsmaß Q ist damit der Prozess $e^{-rt}S(t)$ ein Martingal. (Beachte, dass es nur 2 mögliche Zeitpunkte gibt.)

Im kontinuierlichen Fall lässt sich zeigen, dass ein eindeutiges Wahrscheinlichkeitsmaß Q (äquivalentes Martingalmaß oder risikoneutrales Maß) existiert, so dass der Prozess $e^{-rt}S(t)$ ein Martingal bezüglich Q ist und dass für den Optionspreis

$$V(t) = E_Q[e^{-r(T-t)}V(T) | \mathcal{F}_t], \quad t \in [0, T]$$

gilt (also besitzt $e^{-rt}V(t)$ ebenfalls die Martingaleigenschaft).

Unter dem risikoneutralen Maß lautet die Gleichung für den Aktienkurs

$$dS(t) = rS(t)dt + \sigma S(t)d\tilde{W}(t),$$

d.h. die subjektive mittlere Wachstumsrate b wird durch den risikofreien Zinssatz r ersetzt und \tilde{W} ist eine Brownsche Bewegung bezüglich Q . Für Details siehe [KK01].

Nach der Feynman-Kac Formel (Satz 7.9) erfüllt

$$V(t, S) = E_Q[e^{-r(T-t)}\Lambda(S(T)) | S(t) = S]$$

die partielle Differentialgleichung (Black-Scholes Gleichung)

$$\begin{cases} \frac{\partial V}{\partial t} + AV - rV = 0, & t \in [0, T), \\ V(T, S) = \Lambda(S), \end{cases}$$

wobei der Operator A , gegeben durch

$$AV = \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} + rS \frac{\partial V}{\partial S},$$

der infinitesimale Generator von $S(t)$ unter dem Wahrscheinlichkeitsmaß Q ist.

8.3 Eigenschaften der Black-Scholes Gleichung

Satz 8.6 (Black-Scholes Formel)

Gegeben sei die Black-Scholes Gleichung

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

mit der Endbedingung $V(T, S) = \Lambda(S)$. Sei

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

die Verteilungsfunktion der Standard-Normalverteilung und

$$d_{1/2} = \frac{\ln(S/K) + (r \pm \sigma^2)(T - t)}{\sigma\sqrt{T - t}}.$$

(i) Für eine europäische Call-Option, d.h. $\Lambda(S) = (S - K)^+$ und Randbedingungen $V(t, 0) = 0$, $V(t, S)/S \rightarrow 1$ für $S \rightarrow \infty$ ist die Lösung gegeben durch

$$V(t, S) = S \cdot \Phi(d_1) - Ke^{-r(T-t)}\Phi(d_2),$$

$$S \in (0, \infty), t \in [0, T].$$

(ii) Für eine europäische Put-Option, d.h. $\Lambda(S) = (K - S)^+$ und Randbedingungen $V(t, 0) = Ke^{-r(T-t)}$, $V(t, S) \rightarrow 0$ für $S \rightarrow \infty$ ist die Lösung gegeben durch

$$V(t, S) = -S \cdot \Phi(-d_1) + Ke^{-r(T-t)} \cdot \Phi(-d_2),$$

$$S \in (0, \infty), t \in [0, T].$$

Beweis: (Skizze)

Man führt die Substitutionen

$$x = \ln(S/K), \tau = \frac{1}{2}\sigma^2(T-t), v(\tau, x) = V(t, S)/K$$

und anschließend

$$u(\tau, x) = \exp(-\alpha x - \beta\tau)v(\tau, x)$$

mit

$$\alpha = -\frac{1}{2}(k-1), \beta = -\frac{1}{4}(k+1)^2 \text{ für } k = \frac{2r}{\sigma^2}$$

durch. Dann erfüllt u die Diffusionsgleichung

$$u_\tau = u_{xx}, \quad x \in \mathbb{R}, \tau \in (0, \sigma^2 T/2]$$

mit der Anfangsbedingung (im Falle eines Calls)

$$u(0, x) = u_0(x) = (e^{(k+1)x/2} - e^{(k-1)x/2})^+.$$

Die Lösung dieser Gleichung ergibt sich aus der Feynman-Kac Formel in Satz 7.8:

$$u(\tau, x) = E[u_0(X(\tau))]$$

für die Itô-Diffusion $dX(t) = \sqrt{2}dW(t)$ mit $X(0) = x$. Da X eine reskalierte Brownsche Bewegung mit Startpunkt x ist, besitzt $X(\tau)$ die Dichtefunktion $\phi(s) = \frac{1}{\sqrt{4\pi\tau}}e^{-(s-x)^2/4\tau}$, es folgt also

$$u(\tau, x) = \frac{1}{\sqrt{4\pi\tau}} \int_{\mathbb{R}} u_0(s)e^{-(s-x)^2/4\tau} ds. \quad (8.5)$$

Eine Rücktransformation in den ursprünglichen Variablen t, S liefert die Black-Scholes Formel. Man überprüft auch, dass die so erhaltene Lösung auch die Randbedingungen erfüllt.

Für weitere Details siehe Vorlesung. □

Bemerkung: Wenn man in der Formel (8.5) die Substitutionen $y = (s-x)/\sqrt{2\tau}$ und anschließend $S' = \exp(\sqrt{2\tau}y)S$ (für festes S) durchführt, erhält man

$$V(t, S) = K e^{\alpha x + \beta \tau} u(\tau, x) = e^{-r(T-t)} \int_0^\infty S' f(S'; S, t) \Lambda(S') dS'$$

mit

$$f(S'; S, t) = \frac{1}{S' \sigma \sqrt{2\pi(T-t)}} \exp\left(-\frac{[\ln(S'/S) - (r - \sigma^2/2)(T-t)]^2}{2\sigma^2(T-t)}\right),$$

also die Form

$$V(t, S) = E_Q[e^{-r(T-t)} \Lambda(S)],$$

wobei f die Dichtefunktion des risikoneutralen Wahrscheinlichkeitsmaßes Q darstellt.

8.3.1 Exkurs: Die eindimensionale Diffusionsgleichung

Man betrachte die partielle Differentialgleichung

$$\begin{cases} u_t = ku_{xx}, & x \in I, t > 0, \\ u(0, x) = g(x), & x \in I \end{cases}$$

für $k > 0$ und $I = \mathbb{R}$ oder $I = [a, b]$, mit geeigneten Randbedingungen.

Diese Gleichung, auch genannt *Wärmeleitungsgleichung*, beschreibt die Temperaturentwicklung in einem isolierten Stab, der von außen nur durch die beiden Enden beeinflusst werden kann. Man kann z.B. die Temperatur der beiden Enden konstant halten: $u(t, a) = M_1, u(t, b) = M_2$ (Dirichlet-Randbedingungen), oder die beiden Enden isolieren, d.h. der Wärmefluss durch den Rand verschwindet. Dies wird durch die Neumann-Randbedingung $u_x = 0$ ausgedrückt.

Die folgende Rechnung zeigt die Regularitätseigenschaften der Lösung (unter der Annahme, dass eine hinreichend glatte Lösung existiert).

Es gilt:

$$0 = \int_I (u_t - ku_{xx})^2 dx = \int_I u_t^2 - 2ku_t u_{xx} + u_{xx}^2 dx.$$

Falls in den Randpunkten von I entweder $u_t(a, t) = u_t(b, t) = 0$ gilt (z.B. für konstante Dirichlet-Randbedingungen $u(t, a) = M_1, u(t, b) = M_2$ oder für Neumann-Randbedingungen $u_x(t, a) = u_x(t, b) = 0$ bzw. für $I = \mathbb{R}$ soll u samt Ableitungen bei $\pm\infty$ hinreichend schnell gegen 0 konvergieren), so liefert die partielle Integration

$$- \int_I u_t u_{xx} dx = \int_I u_{tx} u_x dx = \int_I \frac{1}{2} \frac{d}{dt} (u_x^2) dx,$$

denn die zusätzlichen Randterme verschwinden.

Somit erhält man

$$\int_I u_t^2 + u_{xx}^2 + k \frac{d}{dt} (u_x^2) dx = 0.$$

Sei $T > 0$ beliebig. Integration über $[0, T_1]$ mit $T_1 \in (0, T]$ und Einsetzen der Anfangsbedingung liefert

$$\int_0^{T_1} \int_I u_t^2 + u_{xx}^2 dx dt + k \int_I u_x^2(x, T_1) - g_x^2(x) dx = 0,$$

also

$$\frac{1}{k} \int_0^{T_1} \int_I u_t^2 + u_{xx}^2 dx dt + \int_I u_x^2(x, T_1) dx = \int_I g_x^2 dx.$$

Daraus folgt

$$\sup_{t \in [0, T]} \|u_x\|_{L^2(I)}^2 \leq \|g_x\|_{L^2(I)}^2,$$

d.h. die L^2 -Norm der Ableitung der Lösung ist zu allen Zeiten kleiner als die L^2 -Norm der Ableitung der Anfangsbedingung. Dabei war die Tatsache $k > 0$ entscheidend. Für $k < 0$ (bzw. für eine Gleichung der Form $u_t + ku_{xx} = 0$ mit $k > 0$ (Rückwärts-Diffusionsgleichung)) gilt gerade das Gegenteil und dieses Problem erweist sich als schlecht gestellt. Dies wird später durch ein Beispiel erläutert.

Dabei heißt eine partielle Differentialgleichung *wohlgestellt* (im Sinne von Hadamard), wenn eine eindeutige Lösung existiert, welche stetig von den Anfangs- und Randbedingungen abhängt, d.h. eine kleine Störung in diesen Werten soll eine kleine Störung der Lösung implizieren (Stabilität).

Betrachte nun für $k > 0$ die Gleichung

$$\begin{cases} u_t = ku_{xx}, & x \in (0, a), t > 0, \\ u(0, x) = g(x), & x \in (0, a), \\ u(t, 0) = u(t, a) = 0, & t > 0. \end{cases} \quad (8.6)$$

Wir suchen Lösungen der Form $u(t, x) = T(t)X(x)$ (Trennung der Variablen). Damit folgt $u_t = T'X$ und $u_{xx} = TX''$. Einsetzen in die Gleichung liefert $T'X = kTX''$ oder

$$\frac{T'}{T} = k \frac{X''}{X} =: -\lambda.$$

Es ergeben sich getrennte gewöhnliche Differentialgleichungen für T und X :

$$\begin{cases} T' + \lambda T = 0, \\ kX'' + \lambda X = 0. \end{cases}$$

Die Gleichung für T hat die allgemeine Lösung $T(t) = Ce^{-\lambda t}$. Für $C = 0$ erhält man $u(t, x) \equiv 0$, welche die Lösung zur Anfangsbedingung $g \equiv 0$ ist. Nehme also an, $C \neq 0$.

Die Gleichung für X ist eine lineare, homogene Differentialgleichung zweiter Ordnung. Setze $\mu := \lambda/k$.

- Für $\lambda < 0$ ist die allgemeine Lösung gegeben durch $X(x) = Ae^{\sqrt{-\mu}x} + Be^{-\sqrt{-\mu}x}$. Diese muss zusätzlich die Randbedingungen $X(0) = X(a) = 0$ erfüllen (da $T(t) \neq 0$ im Fall $C \neq 0$ gilt). Das entstandene lineare Gleichungssystem liefert die Lösung $A = B = 0$, also wieder $u(t, x) \equiv 0$, d.h. die Triviale Lösung zur Anfangsbedingung $g \equiv 0$.
- Falls $\lambda = 0$ reduziert sich die Gleichung auf $X'' = 0$, welche die allgemeine Lösung $X(x) = A + Bx$ besitzt. Eine analoge Überlegung liefert auch in diesem Fall nur die triviale Lösung.

- Falls $\lambda > 0$ ist die allgemeine Lösung gegeben durch $X(x) = A \sin(\sqrt{\mu}x) + B \cos(\sqrt{\mu}x)$. Die Randbedingungen $X(0) = X(a) = 0$ liefern das System

$$\begin{cases} B = 0, \\ A \sin(\sqrt{\mu}a) = 0. \end{cases}$$

Für $A \neq 0$ erhält man nichttriviale Lösungen, falls $\sin(\sqrt{\mu}a) = 0$ gilt, also für $\sqrt{\mu}a = n\pi$, $n \in \mathbb{N}$:

$$X_n(x) = A_n \sin\left(\frac{n\pi}{a}x\right).$$

Mit $\lambda = \mu k = \frac{n^2\pi^2}{a^2}k$ und dem Ansatz $u(t, x) = T(t)X(x)$ ergibt sich, dass die Funktionenschar

$$u_n(t, x) = c_n e^{-\frac{n^2\pi^2}{a^2}kt} \cdot \sin\left(\frac{n\pi}{a}x\right)$$

die Gleichung

$$u_t = k u_{xx}, \quad u(t, 0) = u(t, a) = 0 \quad (8.7)$$

erfüllt. Für $t = 0$ erhält man aber nur eine spezielle Klasse von möglichen Anfangswerten.

Man bemerkt jedoch, dass aufgrund der Linearität der Differentialoperatoren und der 0-Randbedingungen die Gleichung (8.7) *linear* ist, d.h. Linearkombinationen von Lösungen liefern weitere Lösungen. Somit sind auch alle Funktionen der Form

$$u(t, x) = \sum_{n=0}^N c_n e^{-\frac{n^2\pi^2}{a^2}kt} \cdot \sin\left(\frac{n\pi}{a}x\right)$$

Lösungen von (8.7) mit der Anfangsbedingung

$$u(0, x) = g(x) := \sum_{n=0}^N c_n \sin\left(\frac{n\pi}{a}x\right).$$

Bemerkung: Für allgemeine Randbedingungen der Form $u(t, 0) = M_1$, $u(t, a) = M_2$ ist die Diffusionsgleichung nicht mehr linear. Die Funktion $U(x) = M_1 + (M_2 - M_1)\frac{x}{a}$ (unabhängig von t) ist jedoch Lösung der Diffusionsgleichung mit diesen Randbedingungen und für eine Lösung v von (8.7) ist $u := v + U$ eine Lösung der Diffusionsgleichung mit Randbedingungen $u(t, 0) = M_1$, $u(t, a) = M_2$.

Für beliebige Anfangswerte g (z.B. stückweise stetig) muss die Darstellung durch endliche trigonometrische Summen durch eine Fourierreihendarstellung ersetzt werden. Sei

$$D := \{(t, x) \in \mathbb{R}^2 : x \in (0, a), t > 0\},$$

$$\bar{D} := \{(t, x) \in \mathbb{R}^2 : x \in [0, a], t \geq 0\}.$$

Satz 8.7 (Existenz) Sei $g : [0, a] \rightarrow \mathbb{R}$ stückweise stetig, mit $g(0) = g(a) = 0$, so dass die seitlichen Ableitungen in jedem Punkt existieren, und mit der Fourier-Sinusreihendarstellung

$$g(x) = \sum_{n=0}^{\infty} c_n \sin\left(\frac{n\pi}{a}x\right)$$

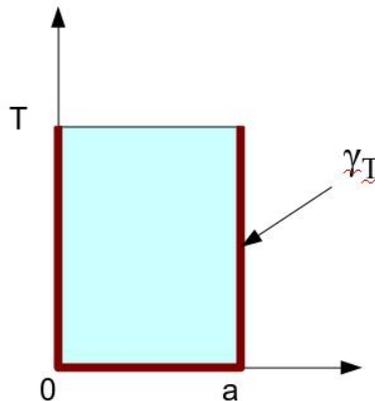
(d.h. $c_n = \frac{2}{a} \int_0^a g(x) \sin\left(\frac{n\pi}{a}x\right) dx$).

Dann konvergiert die Reihe

$$\sum_{n=0}^{\infty} c_n e^{-\frac{n^2\pi^2}{a^2}kt} \cdot \sin\left(\frac{n\pi}{a}x\right)$$

gleichmäßig in t gegen eine Funktion $u \in C^2(D) \cap C(\bar{D})$ welche eine Lösung der Gleichung (8.6) ist.

Für die Wohlgestellttheit des Anfang-Randwertproblems muss noch die Eindeutigkeit und die Stabilität der Lösung nachgewiesen werden. Diese ergeben sich aus dem *Maximumprinzip*. Für $T > 0$ definiere dazu $D_T := \{(t, x) \in D : t \leq T\}$ bzw. $\bar{D}_T := \{(t, x) \in \bar{D} : t \leq T\}$ und $\gamma_T := \bar{D}_T \setminus D_T$.



Satz 8.8 (*Maximumprinzip*)

Sei u auf D_T eine Lösung der Gleichung $u_t = ku_{xx}$, $u(t, 0) = u(t, a) = 0$. Dann wird das Maximum und das Minimum von u auf γ_T erreicht.

Beweis: Siehe Vorlesung. □

Satz 8.9 Das Problem (8.6) ist wohlgestellt und es gilt $u \geq 0$ für $g \geq 0$.

Beweis: Siehe Vorlesung. □

Bemerkung: Die Rückwärts-Diffusionsgleichung $u_t + ku_{xx} = 0$ in $(0, a)$ mit $u(t, 0) = u(t, a) = 0$, $u(0, x) = g(x)$ und $k > 0$ ist jedoch nicht wohlgestellt. Zu jedem $T > 0$, $\epsilon > 0$ (beliebig klein) und $M > 0$ beliebig groß kann man (für eine passende Anfangsbedingung $g = g_\epsilon$) eine Lösung u finden mit $\|u(0, \cdot)\|_\infty \leq \epsilon$ und $\|u(T, \cdot)\|_\infty \geq M$, d.h. beliebig kleine Störungen der Anfangswerte führen zu beliebig großen Störungen in der Lösung auf jedem gegebenen Zeitintervall.

Wähle dazu als Anfangsbedingung $u(0, x) = g(x) = \epsilon \sin(\frac{n\pi}{a}x)$. Man rechnet nach, dass

$$u(t, x) = \epsilon e^{\frac{n^2\pi^2}{a^2}kt} \cdot \sin(\frac{n\pi}{a}x)$$

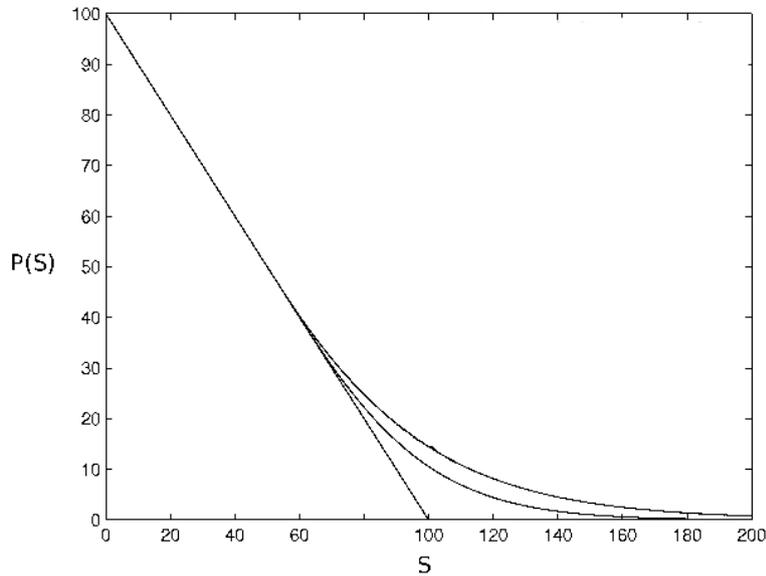
eine Lösung der Rückwärts-Diffusionsgleichung ist, mit $\|u(T, \cdot)\|_\infty \geq M$ für hinreichend große Werte von n .

8.4 Amerikanische Optionen und freie Randwertprobleme

Betrachte eine amerikanische Put-Option mit Wert $P(t, S)$, die entsprechende Auszahlungsfunktion ist somit $(K - S)^+$. Im folgenden Bild für $K = 100$ sind die Preiskurven der Option für $t = T$, $t = 0$ und ein beliebiges $t \in (0, T)$ dargestellt.

Es gilt $P(t, S) \geq (K - S)^+$, d.h. die Preiskurve der Option liegt stets über der Kurve der Auszahlungsfunktion (entspricht der Preiskurve für $t = T$, s. auch Bild). Für $S \geq K$ ist die Ungleichung klar, da $(K - S)^+ = 0$. Falls $S < K$ und $P(t, S) < K - S$ gilt, so kann man folgende Arbitragemöglichkeit finden: Kaufe die Option für P (auf Kredit). Die Ausübung der Option bringt den Gewinn $K - S > P$ ein, nach Zurückzahlung der Kreditschulden bleibt der positive Gesamtgewinn von $K - S - P > 0$.

Behauptung: Es gibt einen kritischen Wert S_f , so dass die vorzeitige Ausübung der Option für $S < S_f$ lohnenswert ist, aber für $S \geq S_f$ nicht lohnenswert ist (s. auch Bild).



Beweis der Behauptung:

Für $S \geq K$ ist die Ausübung der Option auf jeden Fall nicht lohnenswert, denn der Marktpreis ist höher als der Ausübungspreis.

Wenn nun S von K nach 0 fällt, so steigt $P(t, S)$. Man betrachte das Portfolio aus einem Put und einer Aktie mit dem Wert $\pi = P + S$.

Falls $P > (K - S)^+$ gilt, ist die Ausübung der Option nicht lohnenswert. Vor der Ausübung hat das Portfolio den Wert $\pi = P + S > (K - S)^+ + S \geq K$ und nach der Ausübung gilt $\pi = (K - S)^+ + S = K - S + S = K$. Sobald $P = (K - S)^+ = K - S$ gilt, lohnt sich die Ausübung der Option: der Wert von π vor und nach der Ausübung beträgt K . Ferner gilt für $S = 0$ die Randbedingung $P(t, 0) = K$, d.h. die Ausübung lohnt sich. Also existiert ein größter Wert $S_f < K$, so dass $P(S_f, t) = K - S_f$. Somit gilt:

$$P(t, S) \begin{cases} = (K - S)^+ = K - S & \text{für } S \leq S_f(t), \\ > (K - S)^+ & \text{für } S > S_f(t). \end{cases}$$

Genauer: Es gilt

$$P(t, S) \begin{cases} = K - S & \text{für } S \leq S_f(t), \\ \text{erfüllt die Black-Scholes-Gleichung} & \text{für } S > S_f(t), \end{cases}$$

denn in dem letzten Fall wird die Option gehalten und die Herleitung der Gleichung erfolgt analog zu der für den Preis einer europäischen Option.

Die Randbedingungen der Black-Scholes-Gleichung für eine amerikanische Put-Option sind: $P(t, S) \rightarrow 0$ für $S \rightarrow \infty$ und $P(t, S_f(t)) = K - S_f(t)$,

wobei der *freie Randwert* $S_f(t)$ zunächst unbekannt ist. Um diesen eindeutig bestimmen zu können, fordert man zusätzlich die Stetigkeit der Ableitung $\frac{\partial P(t,S)}{\partial S}$ an der Stelle $S = S_f$ (die Begründung geht auch über eine Arbitrage-Argumentation). Wegen $P(t,S) = K - S$ für $S < S_f$ muss daher $\frac{\partial P(t,S)}{\partial S} = -1$ gelten.

Zusammenfassung: Das *freie Randwertproblem* für den Wert eine amerikanischen Put-Option lautet:

$$\begin{aligned} P(t,S) &= K - S \quad \text{für } S \leq S_f(t), \\ \frac{\partial P}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 P}{\partial S^2} + rS \frac{\partial P}{\partial S} - rP &= 0 \quad \text{für } S > S_f(t), \\ P(S,T) &= (K - S)^+, \quad (\text{Endbedingung}) \\ \left. \begin{aligned} P(t,S) &\rightarrow 0 \quad \text{für } S \rightarrow \infty, \\ P(t,S_f(t)) &= K - S_f(t), \\ \frac{\partial P}{\partial S}(t,S_f(t)) &= -1, \end{aligned} \right\} & (\text{Randbedingungen}) \end{aligned}$$

Die Bestimmung des freien Randwertes S_f ist problematisch, man sucht daher alternative Formulierungen, in denen dieser Wert nicht explizit vorkommt. Analog zur Herleitung der Black-Scholes-Gleichung für europäische Optionen betrachtet man ein risikoloses, selbstfinanzierendes Portfolio $Y = c_1 B + c_2 S - V$. Die Wahl $c_2 = \partial V / \partial S$ annulliert den stochastischen Anteil und man erhält

$$dY(t) = (c_1 r B - \frac{\partial V}{\partial t} - \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2}) dt.$$

Der Besitzer von Y hat eine amerikanische Put-Option verkauft (ausgedrückt durch den Term $-V$ in Y). Wenn der Käufer der Option diese nicht optimal einlöst, so kann der Verkäufer einen größeren Gewinn erzielen als durch eine risikolose Anlage. Es gilt also $dY \geq rY dt$ und somit erhält man die *Black-Scholes-Ungleichung*:

$$\frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} + rS \frac{\partial V}{\partial S} - rV \leq 0.$$

Nach den vorherigen Überlegungen gilt Gleichheit für $S > S_f$. Für $S \leq S_f$ gilt $V = K - S$ und ein Einsetzen in die Black-Scholes-Ungleichung zeigt, dass diese strikt erfüllt wird.

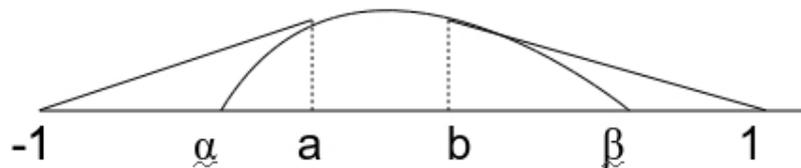
Man kann somit folgendes *lineares Komplementaritätsproblem* aufstellen:

$$\begin{aligned} (V - \Lambda(S)) \left(\frac{\partial V}{\partial t} + \frac{1}{2} \sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} + rS \frac{\partial V}{\partial S} - rV \right) &= 0 \\ - \left(\frac{\partial V}{\partial t} + \frac{1}{2} \sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} + rS \frac{\partial V}{\partial S} - rV \right) &\geq 0 \\ V - \Lambda(S) = 0 &\geq 0, \end{aligned}$$

wobei Λ eine beliebige Auszahlungsfunktion bezeichnet, d.h. das obige Komplementaritätsproblem ergibt sich für allgemeine amerikanische Optionen. Je nach der speziellen Option müssen noch End- und Randbedingungen formuliert werden. Für einen Standard-Put, d.h. $\Lambda(S) = (K - S)^+$, gilt die Endbedingung $V(T, S) = \Lambda(S)$ und die Randbedingungen $V(t, 0) = K, V(t, S) \rightarrow 0$ für $S \rightarrow \infty$.

8.4.1 Exkurs: Das Hindernisproblem

Durch eine passende Transformation kann man das freie Randwertproblem bzw. das lineare Komplementaritätsproblem für amerikanische Optionen auf folgendes Problem zurückführen: Fixiere ein Seil mit minimaler Länge in den Punkten $x = -1$ und $x = 1$ über den Graphen einer Funktion $f \geq 0$.



Gegeben:

$$\begin{aligned} f &\geq 0 \text{ in } [-1, 1], \\ f &= 0 \text{ in } [-1, \alpha] \cup [\beta, 1] \\ f &> 0 \text{ in } (\alpha, \beta) \\ f &\in C^2(\alpha, \beta), f'' < 0 \end{aligned}$$

Gesucht:

$$\begin{aligned}
 u &\in C^1(-1, 1) \cap C[-1, 1] \\
 u &\in C^2(-1, a) \cap C^2(a, b) \cap C^2(b, 1) \\
 u(-1) &= u(1) = 0 \\
 u &\geq f \text{ in } (-1, 1) \\
 u'' &= 0 \text{ in } (-1, a) \cup (b, 1) \\
 u &= f \text{ in } (a, b)
 \end{aligned}$$

Die freien Randwerte a, b sind unbekannt und somit Teil des Problems. In der alternativen Formulierung als lineares Komplementaritätsproblem treten diese nicht mehr auf:

Gesucht wird $u \in C^1(-1, 1)$ mit $u(-1) = u(1) = 0$ und:

$$\begin{aligned}
 -u'' &\geq 0 & (8.8) \\
 u - f &\geq 0 \\
 u''(u - f) &= 0
 \end{aligned}$$

Da man i.A. nicht erwarten kann, dass $u \in C^2(-1, 1)$ (an den Stellen a, b ist u möglicherweise nicht 2 Mal differenzierbar), wird weiterhin eine Formulierung vorgestellt, in der die zweiten Ableitungen von u nicht auftreten.

Definiere $\mathcal{K} := \{v \in C_{st}^1(-1, 1) : v(-1) = v(1) = 0, v \geq f\}$ als Raum der zulässigen Funktionen (Testfunktionen). Die Bezeichnung C_{st}^1 bedeutet stetig auf $[-1, 1]$ und stückweise C^1 .

Satz 8.10 (i) Sei $u \in C^2(-1, 1)$ eine Lösung von (8.8). Dann löst u die Variationsungleichung

$$\text{Suche } u \in \mathcal{K} \text{ mit } \int_{-1}^1 u'(v - u)' dx \geq 0 \quad \forall v \in \mathcal{K}. \quad (8.9)$$

(ii) Sei u eine Lösung von (8.9) mit $u \in C^2(-1, 1)$ Dann ist u auch Lösung von (8.8).

Beweis: Siehe Vorlesung. □

Den natürlichen Rahmen für solche Probleme stellen jedoch die *Sobolev-Räume* dar, auf denen die Ableitungen im verallgemeinerten Sinn zu verstehen sind. Im Folgenden werden diese im eindimensionalen Fall kurz erläutert. Ausgangspunkt sind die Räume

$$L^p(-1, 1) = \{v : (-1, 1) \rightarrow \mathbb{R} : \int_{-1}^1 |v(x)|^p dx < \infty\}$$

für $p \geq 1$ (genauer: man fasst Funktionen, die fast überall gleich sind, in Äquivalenzklassen zusammen). Da das Intervall beschränkt ist, folgt sofort dass $L^p(-1, 1) \subseteq L^1(-1, 1)$ für alle p (in diesem kurzen Exkurs betrachten wir nur endliche p).

Eine Funktion $w \in L^1(-1, 1)$ heißt *schwache Ableitung (verallgemeinerte Ableitung)* der Funktion $v \in L^1(-1, 1)$, wenn die partielle Integration mit glatten Testfunktionen mit kompaktem Träger gilt:

$$\int_{-1}^1 w\phi dx = - \int_{-1}^1 v\phi' dx \quad \forall \phi \in C_0^\infty(-1, 1).$$

Man bezeichnet $v' := w$. Falls v im klassischen Sinne differenzierbar ist, stimmt die schwache Ableitung mit der klassischen überein.

Die *Sobolev-Räume* $H^{m,p}(-1, 1)$ sind wie folgt definiert:

$$\begin{aligned} H^{1,p}(-1, 1) &:= \{v \in L^p(-1, 1) : \exists v' \text{ mit } v' \in L^p(-1, 1)\} \\ H^{m,p}(-1, 1) &:= \{v \in H^{m-1,p}(-1, 1) : \exists v^{(m-1)} \text{ mit } v^{(m-1)} \in L^p(-1, 1)\} \\ &\text{für } m \geq 2 \end{aligned}$$

wobei v' und $v^{(m-1)}$ schwache Ableitungen erster Ordnung bzw. der Ordnung $m-1$ sind. Für $p = 2$ benutzt man für die Hilbert-Räume $H^{m,2}$ die abgekürzte Notation $H^m(-1, 1)$.

Die Ergebnisse aus Satz 8.10 bleiben gültig, wenn man C_{st}^1 durch H^1 und C^2 durch H^2 ersetzt. Es ist noch das Problem der Randwerte zu klären, da bekanntlich den L^p -Funktionen keine eindeutigen Randwerte zugeordnet werden können (eine Veränderung auf Nullmengen liefert die gleiche L^p -Funktion). Im allgemeinen Fall definiert man Randwerte von H^1 -Funktionen über einen Spuroperator, im eindimensionalen Fall kann dies auch über ein klassisches Argument erfolgen, denn die H^1 -Funktionen sind in diesem Fall stetig. Es gilt genauer

$$H^1(-1, 1) = \{v \in L^2(-1, 1) : v(x) = c + \int_{-1}^x w(y)dy \text{ für } c \in \mathbb{R}, w \in L^2\}.$$

Nach bekannten Resultaten aus der Analysis sind die Funktionen v mit der obigen Eigenschaft stetig, fast überall differenzierbar und es gilt $w = v'$ f.ü.. w ist auch die schwache Ableitung von v .

Eine weitere Formulierung des Hindernisproblems lautet wie folgt: Betrachte als Raum der zulässigen Funktionen

$$\mathcal{K} := \{v \in H^1(-1, 1) : v(-1) = v(1) = 0, v \geq f\}$$

und das Funktional

$$J : \mathcal{K} \rightarrow [0, \infty), \quad J(v) = \frac{1}{2} \int_{-1}^1 (v')^2 dx.$$

Dann gilt:

Satz 8.11 *Das Problem (8.9) ist äquivalent zu*

$$\text{Suche } u \in \mathcal{K} \text{ mit } J(u) = \min_{v \in \mathcal{K}} J(v). \quad (8.10)$$

Beweis: Siehe Vorlesung. □

Die Existenz und Eindeutigkeit der Lösung von (8.10) zeigt man mit Argumenten aus der Funktionalanalysis.

8.5 Asiatische Optionen

Bei *asiatischen Optionen* hängt die Auszahlungsfunktion von den gemittelten Aktienkursen \hat{S} über die Laufzeit T der Option ab.

Mögliche Variationen:

- Der Mittelwert \hat{S} bezieht sich auf diskrete Zeitpunkte $S(t_1), \dots, S(t_n)$ oder auf alle Zeitpunkte $S(t)$, $t \in [0, T]$.

- Der Mittelwert \hat{S} ist arithmetisch oder geometrisch:

- $\hat{S} = n^{-1} \sum_{i=1}^n S(t_i)$
- $\hat{S} = (\prod_{i=1}^n S(t_i))^{1/n}$
- $\hat{S} = \frac{1}{T} \int_0^T S(t) dt$
- $\hat{S} = \exp\left(\frac{1}{T} \int_0^T \ln S(t) dt\right)$

- Typen der Auszahlungsfunktion:

- *Strike-Optionen*: $(S - \hat{S})^+$ (Call) bzw. $(\hat{S} - S)^+$ (Put)
- *Rate-Optionen*: $(\hat{S} - K)^+$ (Call) bzw. $(K - \hat{S})$ (Put)

- Asiatische Optionen können vom europäischen oder amerikanischen Typ sein

Bemerkung: Bei Strike-Optionen ersetzt der Mittelwert \hat{S} den Ausübungsbetrag (engl. strike) K , während er bei Rate-Optionen den momentanen Kurs des Basiswertes zum Ausübungszeitpunkt ersetzt.

Wir betrachten nun asiatische Optionen vom europäischen Typ mit Auszahlungsfunktion $\Lambda = \Lambda(S, I)$ für $I = \int_0^t f(\tau, S(\tau))d\tau$. Im Falle eines *Arithmetic-Average-Strike Call* ist etwa $f(t, S) = S$ und $\Lambda = (S - I/T)^+$.

Sei $V = V(t, S, I)$ der Preis dieser Option zum Zeitpunkt t . Es wird eine partielle Differentialgleichung für V hergeleitet, analog zur Herleitung der Black-Scholes-Gleichung. Man betrachtet ebenfalls ein risikoloses, selbstfinanzierendes Portfolio der Form $Y = c_1B + c_2S - V(t, S, I)$ und man verwendet die üblichen Gleichungen für dB und dS . Es gilt zusätzlich $dI = f(t, S)dt$. Der Unterschied hier ist, dass V zusätzlich von $I(S)$ abhängt. Daher wird für die Bestimmung von dV zunächst die Itô-Formel angewendet, wobei sich weitere von $S(t)$ abhängige Terme ergeben. Für diese wendet man erneut die Itô-Formel an und man erhält eine sogenannte *Itô-Taylor Entwicklung*, in der mehrfache stochastische Integrale auftreten. Unter Vernachlässigung der Terme höherer Ordnung erhält man ähnlich zum Vorgehen bei europäischen Optionen folgende partielle Differentialgleichung:

$$\frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} + rS \frac{\partial V}{\partial S} - rV + f(t, S) \frac{\partial V}{\partial I} = 0, \quad S, I > 0, t \in (0, T).$$

Die Endbedingung lautet $V(T, S, I) = \Lambda(S, I)$ und die Randbedingungen hängen vom Optionstyp ab.

Unter zusätzlichen Annahmen lässt sich die obige Gleichung weiter vereinfachen. Es gelte $f(t, S) = S$ und $\Lambda(t, S, I) = S^\alpha F(t, I/S)$ für eine Funktion $F = F(t, R)$. Zum Beispiel sei $\Lambda = (S - I/T)^+ = S(1 - I/(ST))^+$. Dann ist $\alpha = 1$ und $F(t, R) = (1 - R/t)$.

Unter diesen Voraussetzungen löst $H(t, R) = S^{-\alpha}V$ die Gleichung

$$H_t + \frac{1}{2}\sigma^2 R^2 H_{RR} + (1 - \sigma^2(\alpha - 1)R - rR)H_R + (\alpha - 1)\left(\frac{\alpha}{2}\sigma^2 + r\right)H = 0, \\ R > 0, t \in (0, T)$$

mit der Endbedingung $H(T, R) = F(T, R)$, $R > 0$.

Man erhält also eine Gleichung in zwei statt in drei Variablen. Für den Arithmetic-Average-Strike Call mit $\alpha = 1$, $F(t, R) = (1 - R/t)^+$ gilt

$$H_t + \frac{1}{2}\sigma^2 R^2 H_{RR} + (1 - rR)H_R = 0, \quad R > 0, t \in (0, T)$$

mit der Endbedingung $H(T, R) = (1 - R/T)^+$ und den Randbedingungen: $H(R, t) \rightarrow 0$ für $R \rightarrow \infty$ bzw. $H_t + H_R = 0$ für $R = 0$. Diese letzte Gleichung

leitet man unter der Annahme her, dass die partielle Differentialgleichung auch in $R = 0$ gilt.

Für den Optionspreis V erhält man somit

$$V(t, S, I) = S \cdot H\left(t, \frac{I}{S}\right).$$

Der Kaufpreis der Option bei $t = 0$ beträgt also

$$V(0, S, 0) = S \cdot H(0, 0),$$

da $I(0) = 0$.

8.6 Die Monte-Carlo-Methode

Im Folgenden wird eine numerische Methode für die approximative Berechnung der Preise von Optionen europäischen Typs vorgestellt, welche auf die Simulation der Pfade der Aktienkurse basiert.

Sei $V(t, S(t))$ der Preis der Option zum Zeitpunkt t und Λ die Auszahlungsfunktion. Unter dem risikoneutralen W'Maß Q ist die Dynamik des Aktienkurses gegeben durch:

$$dS(t) = rS(t)dt + \sigma S(t)dW(t). \quad (8.11)$$

Der Optionspreis berechnet sich in diesem Fall durch

$$V(0, S(0)) = e^{-rT} E_Q[V(T, S(T))] = e^{-rT} E_Q[\Lambda(S(T))].$$

Monte-Carlo-Algorithmus

1. Simulation von M Basiswert-Pfaden $S^{(k)}(t)$, $k = 1 \dots M$, $t \in [0, T]$, mit Anfangswert $S(0)$. Zum Zeitpunkt $t = T$ gilt dann $V(T, S^{(k)}(T)) = \Lambda(S^{(k)}(T))$.
2. Nach dem Gesetz der großen Zahlen ist ein Schätzer für $E_Q[\Lambda(S(T))]$ gegeben durch

$$\hat{E}(V_T) = \frac{1}{M} \sum_{k=1}^M V(T, S^{(k)}(T)),$$

wobei $V_T := (V(T, S^{(1)}(T)), \dots, V(T, S^{(M)}(T)))^T$.

3. Der Optionspreis in $t = 0$ wird approximiert durch $\hat{V} = e^{-rT} \hat{E}(V_T)$.

Es sind somit numerische Simulationen erforderlich:

- für die Realisierungen der Brownschen Bewegung
- für die Approximation der Lösung der SDGL (8.11) zum jeweiligen Pfad der Brownschen Bewegung.

Details zur Erzeugung von Zufallszahlen werden in den Übungen besprochen. Es sei nur erwähnt, dass die Computer sogenannte *Pseudo-Zufallszahlen* erzeugen, welche unabhängige, gleichverteilte Zufallsvariablen (z.B. auf $[0,1]$) approximieren. Mithilfe dieses Zufallszahlengenerators kann man dann auch andere Verteilungen simulieren.

8.6.1 Das Euler-Maruyama Verfahren

Betrachte die SDGL

$$dX(t) = a(X(t), t)dt + b(X(t), t)dW(t) \quad (8.12)$$

mit $X(0) = X_0$ für $t \in [0, T]$.

Eine direkte Verallgemeinerung des Euler-Verfahrens für gewöhnliche Differentialgleichungen lautet wie folgt:

Man betrachte N äquidistante Zeitschritte mit Länge $h = T/N$. Es sei $t_i = i \cdot h$, $i = 0 \dots N$, und $\tilde{X}_0 = X_0$. Die Approximation \tilde{X}_i für die Werte $X(t_i)$ berechnet sich wie folgt:

Für $i = 0 \dots N - 1$:

$$\begin{aligned} \Delta W_i &:= W(t_{i+1}) - W(t_i) \sim \mathcal{N}(0, h) \sim \sqrt{h}\mathcal{N}(0, 1), \\ \tilde{X}_{i+1} &= \tilde{X}_i + a(\tilde{X}_i, t_i) \cdot h + b(\tilde{X}_i, t_i) \cdot \Delta W_i. \end{aligned}$$

Für die Güte der Approximation benötigen wir folgende Konvergenzbegriffe:

Definition 8.12 Sei $X(t)$ eine Lösung der SDGL (8.12) und $X^h(t)$ eine Approximation von $X(t)$.

- (i) $X^h(T)$ konvergiert stark mit Ordnung $\gamma > 0$ gegen $X(T)$ zur Zeit T , falls

$$E[|X(T) - X^h(T)|] \leq C \cdot h^\gamma$$

für eine Konstante $C > 0$ und für alle hinreichend kleine $h > 0$ gilt.

- (ii) $X^h(T)$ konvergiert schwach mit Ordnung $\gamma > 0$ gegen $X(T)$ zur Zeit T , falls

$$|E[X(T) - X^h(T)]| \leq C \cdot h^\gamma$$

für eine Konstante $C > 0$ und für alle hinreichend kleine $h > 0$ gilt.

Bemerkung: Die starke Konvergenzordnung gibt die Güte der Approximation der Pfadwerte $X(T)$ durch $X^h(T)$ an, während die schwache Konvergenzordnung die Güte der Approximation des Mittelwertes $E[X(T)]$ durch den Mittelwert $E[X^h(T)]$ angibt.

Satz 8.13 *Das Euler-Maruyama Verfahren hat eine starke Konvergenzordnung von $\gamma = 0.5$ und eine schwache Konvergenzordnung von $\gamma = 1$.*

8.6.2 Das Milstein-Verfahren

Zur Vereinfachung wird zunächst die autonome SDGL

$$dX(t) = a(X(t))dt + b(X(t))dW(t) \quad (8.13)$$

mit Anfangswert $X(t_0)$ betrachtet. In Integralform bedeutet dies

$$X(t) = X(t_0) + \int_{t_0}^t a(X(s))ds + \int_{t_0}^t b(X(s))dW(s).$$

Eine Anwendung der Itô-Formel für $a(X(s))$ und $b(X(s))$ liefert die Itô-Taylor Entwicklung:

$$\begin{aligned} X(t) &= X(t_0) + \int_{t_0}^t \left(a(X(t_0)) + \int_{t_0}^s (a'(X(u))a(X(u)) + \frac{1}{2}a''(X(u))b^2(X(u)))du \right. \\ &\quad \left. + \int_{t_0}^s a'(X(u))b(X(u))dW(u) \right) ds \\ &\quad + \int_{t_0}^t \left(b(X(t_0)) + \int_{t_0}^s (b'(X(u))a(X(u)) + \frac{1}{2}b''(X(u))b^2(X(u)))du \right. \\ &\quad \left. + \int_{t_0}^s b'(X(u))b(X(u))dW(u) \right) dW(s) \\ &= X(t_0) + a(X(t_0))(t - t_0) + b(X(t_0)) \int_{t_0}^t dW(s) + R \end{aligned} \quad (8.14)$$

Durch Vernachlässigen des Restterms R erhält man das bereits vorgestellte Euler-Maruyama Verfahren. Ein Verfahren höherer Ordnung erhält man durch zusätzliches Berücksichtigen des Doppelintegrals bezüglich $dW(u)dW(s)$ und Ersetzen von $b'(X(u))b(X(u))$ durch $b'(X(t_0))b(X(t_0))$:

$$\begin{aligned} X(t) &= X(t_0) + a(X(t_0))(t - t_0) + b(X(t_0)) \int_{t_0}^t dW(s) \\ &\quad + b'(X(t_0))b(X(t_0)) \int_{t_0}^t \int_{t_0}^s dW(u)dW(s) + R_1. \end{aligned}$$

Mit $h = t - t_0$ ist das Doppelintegral bezüglich $dW(u)dW(s)$ von der Ordnung $O(h)$, wie folgende Rechnung zeigt, während der Restterm R_1 von der Ordnung $O(h^{3/2})$ ist:

$$\begin{aligned}
\int_{t_0}^t \int_{t_0}^s dW(u)dW(s) &= \int_{t_0}^t (W(s) - W(t_0))dW(s) \\
&= \int_{t_0}^t W(s)dW(s) - W(t_0) \int_{t_0}^t dW(s) \\
&= \frac{1}{2}(W^2(t) - W^2(t_0)) - \frac{1}{2}(t - t_0) - W(t_0)(W(t) - W(t_0)) \\
&= \frac{1}{2}((W(t) - W(t_0))^2 - (t - t_0)) \\
&= O(t - t_0)
\end{aligned}$$

Die wiederholte Anwendung dieser Vorgehensweise auf $[0, T]$ mit $h = T/N$ und $t_i = i \cdot h$, $i = 0 \dots N$ liefert das *Milstein-Verfahren* (auch allgemein für die nichtautonome SDGL (8.12)):

Für $i = 0 \dots N - 1$:

$$\begin{aligned}
\Delta W_i &:= W(t_{i+1}) - W(t_i) \sim \mathcal{N}(0, h) \sim \sqrt{h}\mathcal{N}(0, 1), \\
\tilde{X}_{i+1} &= \tilde{X}_i + a(\tilde{X}_i, t_i) \cdot h + b(\tilde{X}_i, t_i) \cdot \Delta W_i + \frac{1}{2}b'(\tilde{X}_i, t_i)b(\tilde{X}_i, t_i)((\Delta W_i)^2 - h).
\end{aligned}$$

Satz 8.14 *Das Milstein-Verfahren hat eine starke Konvergenzordnung von $\gamma = 1$ und eine schwache Konvergenzordnung ebenfalls von $\gamma = 1$.*

Kapitel 9

Stochastische Steuerung und Hamilton-Jacobi-Bellman Differentialgleichung

9.1 Problemformulierung

Gegeben sei ein Wahrscheinlichkeitsraum (Ω, \mathcal{A}, P) , der Endzeitpunkt $T > 0$, eine Filtration $\mathcal{F} = (\mathcal{F}_t)_{t \in [0, T]}$, welche die üblichen Bedingungen erfüllt und eine m -dimensionale Brownsche Bewegung $W = (W_t)_{t \in [0, T]}$ (bzgl. \mathcal{F}).

- Ein *Steuerungsprozess* (kurz: *Steuerung*) ist ein progressiv messbarer Prozess $u = (u_t)_{t \in [0, T]}$ mit Werten in einer gegebenen Menge $U \subset \mathbb{R}^d$.
- Der n -dimensionale *gesteuerte Prozess* $X = (X_t)_{t \in [0, T]}$ ist gegeben durch

$$dX_t = b(t, X_t, u_t)dt + \sigma(t, X_t, u_t)dW_t, \quad (9.1)$$

wobei $b : [0, T] \times \mathbb{R}^n \times U \rightarrow \mathbb{R}^n$, $\sigma : [0, T] \times \mathbb{R}^n \times U \rightarrow \mathbb{R}^{n \times m}$ messbare Funktionen sind. Um die Abhängigkeit von X von der Steuerung u zu betonen, schreiben wir X_t^u .

- Als Optimierungskriterium betrachten wir das *Nutzenfunktional*

$$J(t, x, u) = E_{t,x} \left[\int_t^T \psi(s, X_s^u, u_s)ds + \Psi(X_T^u) \right], \quad (9.2)$$

wobei $\psi : [0, T] \times \mathbb{R}^n \times U \rightarrow \mathbb{R}$, $\Psi : \mathbb{R}^n \rightarrow \mathbb{R}$ stetige Funktionen sind. Wir können $\psi(s, X_s^u, u_s)$ als laufenden Nutzen und $\Psi(X_T^u)$ als Endnutzen interpretieren.

- Eine Steuerung $(u_s)_{s \in [t, T]}$ heißt *zulässig*, falls für alle Werte $x \in \mathbb{R}^n$ die gesteuerte SDGL (9.1) mit der Anfangsbedingung $X_t^u = x$ eine eindeutige Lösung auf $[t, T]$ besitzt und das Nutzenfunktional (9.2) wohldefiniert ist. Die Menge aller zulässigen Steuerungen bezeichnen wir mit $\mathcal{A}(t, x)$.
- Unser Ziel ist es, zu gegebenen Startwerten (t, x) durch Wahl einer zulässigen Steuerung $u \in \mathcal{A}(t, x)$ das Nutzenfunktional zu maximieren. Wir erhalten somit das *stochastische Steuerungsproblem*

$$\sup_{u \in \mathcal{A}(t, x)} J(t, x, u).$$

- Die *Wertfunktion* unseres Steuerungsproblems ist damit definiert durch

$$V(t, x) = \sup_{u \in \mathcal{A}(t, x)} J(t, x, u).$$

Gesucht wird eine Steuerung $u^* \in \mathcal{A}(t, x)$, die diesen optimalen Wert realisiert, d.h. $V(t, x) = J(t, x, u^*)$. In diesem Fall heißt u^* *optimal*.

Portfolio-Optimierung: Ein erstes Beispiel

Wir betrachten einen Finanzmarkt, der aus einem Bond mit Preisprozess

$$dB_t = rB_t dt, \quad B_0 = 1, \quad \text{d.h. } B_t = e^{rt}$$

und einer Aktie mit Preisprozess

$$dS_t = \mu S_t dt + \sigma S_t dW_t, \quad S_0 = s > 0$$

mit Drift $\mu \in \mathbb{R}$ und Volatilität $\sigma > 0$ besteht. Die eindeutige Lösung dieser SDGL ist bekanntermaßen

$$S_t = s \exp \left\{ \left(\mu - \frac{1}{2} \sigma^2 \right) t + \sigma W_t \right\}.$$

Das Vermögen eines Investors mit Anfangskapital $x > 0$ ist gegeben durch

$$dX_t = n_t^B dB_t + n_t^S dS_t, \quad X_0 = x,$$

wobei n_t^B und n_t^S die Anzahl der Bonds bzw. Aktien im Portfolio des Investors zur Zeit t bezeichnen. Diese Definition entspricht einem selbstfinanzierendem Portfolio, da Änderungen im Portfoliowert einzig auf Schwankungen

der Bond- und Aktienpreise zurückzuführen sind (Einzahlungen und Entnahmen sind ausgeschlossen).

Als Steuerung zur Zeit t betrachten wir den in die Aktie investierten Anteil u_t am Gesamtvermögen. Dann folgt

$$n_t^B = \frac{(1 - u_t)X_t}{B_t}, \quad n_t^S = \frac{u_t X_t}{S_t},$$

und damit

$$\begin{aligned} dX_t &= (1 - u_t)X_t r dt + u_t X_t (\mu dt + \sigma dW_t) \\ &= X_t ((r + (\mu - r)u_t) dt + \sigma u_t dW_t). \end{aligned}$$

Mit der Itô-Formel verifiziert man (vgl. das Beispiel für den Preisprozess der Aktie aus Abschnitt 6.2)

$$X_t = x \exp \left\{ \int_0^t (r + (\mu - r)u_s - \frac{1}{2}\sigma^2 u_s^2) ds + \int_0^t \sigma u_s dW_s \right\}.$$

Wir wollen nun (9.2) mit $\psi \equiv 0$ und $\Psi(x) = \ln(x)$ zum Zeitpunkt $t = 0$ maximieren, d.h.

$$J(0, x, u) = E [\ln(X_T^u) | X_0 = x],$$

über alle Strategien in

$$\mathcal{A}(0, x) = \left\{ u : E \left[\int_0^T (|\mu u_t| + |\sigma u_t|^2) dt \right] < \infty, X_t^u > 0, E[(\ln(X_T^u))^-] < \infty \right\}$$

Diese Zulässigkeitsbedingung impliziert, dass $\int_0^t \sigma u_s dW_s$ ein Martingal ist. Insbesondere gilt also

$$E \left[\int_0^T \sigma u_t dW_t \right] = 0.$$

Daher erhalten wir für $u \in \mathcal{A}(0, x)$

$$J(0, x, u) = \ln(x) + E \left[\int_0^T (r + (\mu - r)u_t - \frac{1}{2}\sigma^2 u_t^2) dt \right].$$

Maximieren in u_t unter dem Integranden führt wegen

$$\begin{aligned} \frac{\partial}{\partial u} \left(r + (\mu - r)u - \frac{1}{2}\sigma^2 u^2 \right) &= \mu - r - \sigma^2 u, \\ \frac{\partial^2}{\partial u^2} \left(r + (\mu - r)u - \frac{1}{2}\sigma^2 u^2 \right) &= -\sigma^2 < 0. \end{aligned}$$

auf die optimale Wahl

$$u_t^* = \pi^* := \frac{\mu - r}{\sigma^2} \quad \text{für alle } t \in [0, T] \quad (9.3)$$

Die Wertfunktion lautet

$$V(0, x) = \sup_{u \in \mathcal{A}(t, x)} J(t, x, u) = J(t, x, u^*) = \ln(x) + \left(r + \frac{1}{2} \left(\frac{\mu - r}{\sigma} \right)^2 \right) T.$$

Die Strategie π^* in (9.3) heißt *Merton-Strategie* (für logarithmischen Nutzen). Gilt z.B. $\pi^* = 0,4$, so bedeutet dies, dass der Investor 40% seines Geldes in die Aktie investieren sollte. Beachte, dass diese Strategie ständiges Handeln erfordert!

9.2 Konzept der dynamischen Programmierung

Um das stochastische Steuerungsproblem aus Abschnitt 9.1 zu lösen, gehen wir wie folgt vor:

- Betrachte die folgende Beziehung, das sog. *Bellman-Prinzip*:

$$\begin{aligned} V(t, x) &= \sup_{u \in \mathcal{A}(t, x)} E_{t, x} \left[\int_t^\theta \psi(s, X_s^u, u_s) ds + V(\theta, X_\theta^u) \right], \\ V(T, x) &= \Psi(T, x) \end{aligned}$$

für $t \in [0, T]$, $\theta \in [t, T]$ und $x \in \mathbb{R}^n$. Dieses Prinzip besagt, dass sich der maximale Nutzen $V(t, x)$ als Supremum über die folgenden zusammengesetzten Strategien ergibt:

- „Wähle die Steuerung u_s auf $[t, \theta]$.“
Hierdurch entsteht der Nutzen

$$E_{t, x} \left[\int_t^\theta \psi(s, X_s^u, u_s) ds \right].$$

- „Verhalte dich auf $[\theta, T]$ ausgehend von X_θ^u optimal.“
Dies führt in Zustand (θ, X_θ^u) zum Nutzen $V(\theta, X_\theta^u)$.

Beachte: Das Bellman-Prinzip muss noch bewiesen werden, siehe später.

- Wir nehmen nun an, dass das Bellman-Prinzip gültig ist. Anwendung der mehrdimensionalen Itô-Formel auf $V(\theta, X_\theta^u)$ (falls V hinreichend regulär ist, d.h. $V \in \mathcal{C}^{1,2}$) liefert:

$$\begin{aligned}
V(t, x) = \sup_{u \in \mathcal{A}(t, x)} E_{t, x} & \left[\int_t^\theta \psi(s, X_s^u, u_s) ds + V(t, x) \right. \\
& + \int_t^\theta V_t(s, X_s^u) + (V_x(s, X_s^u))^T b(s, X_s^u, u_s) ds \\
& + \int_t^\theta \frac{1}{2} \text{Spur}((V_{xx}(s, X_s^u))a(s, X_s^u, u_s)) ds \\
& \left. + \int_t^\theta (V_x(s, X_s^u))^T \sigma(s, X_s^u, u_s) dW_s \right],
\end{aligned}$$

wobei wir die abkürzenden Schreibweisen $a = \sigma\sigma^T$ und $V_t = \frac{\partial V}{\partial t}$, $V_x = \frac{\partial V}{\partial x}$ und $V_{xx} = \frac{\partial^2 V}{\partial x^2}$ für die partiellen Ableitungen benutzen. Falls $\int_t^\theta (V_x(s, X_s^u))^T \sigma(s, X_s^u, u_s) dW_s$ ein Martingal ist und somit Erwartungswert 0 besitzt, folgt

$$\begin{aligned}
V(t, x) = \sup_{u \in \mathcal{A}(t, x)} E_{t, x} & \left[\int_t^\theta \psi(s, X_s^u, u_s) ds + V(t, x) \right. \\
& + \int_t^\theta V_t(s, X_s^u) + (V_x(s, X_s^u))^T b(s, X_s^u, u_s) ds \\
& \left. + \int_t^\theta \frac{1}{2} \text{Spur}((V_{xx}(s, X_s^u))a(s, X_s^u, u_s)) ds \right].
\end{aligned}$$

- Durch Subtraktion von $V(t, x)$ auf beiden Seiten, Division durch $(\theta - t)$ und Grenzwertübergang $\theta \searrow t$ erhalten wir:

$$\begin{aligned}
0 = \sup_{u \in \mathcal{A}(t, x)} E_{t, x} & \left[\psi(t, X_t^u, u_t) + V_t(t, X_t^u) \right. \\
& \left. + (V_x(t, X_t^u))^T b(t, X_t^u, u_t) + \frac{1}{2} \text{Spur}((V_{xx}(t, X_t^u))a(t, X_t^u, u_t)) \right].
\end{aligned}$$

Da zur Zeit t sowohl der Wert von X_t als auch der von u_t bekannt ist, kann man in der letzten Gleichung den Erwartungswert weglassen und erhält

$$\begin{aligned}
0 = \sup_{u \in U} & \left\{ \psi(t, x, u) + V_t(t, x) + (V_x(t, x))^T b(t, x, u) \right. \\
& \left. + \frac{1}{2} \text{Spur}((V_{xx}(t, x))a(t, x, u)) \right\}. \tag{9.4}
\end{aligned}$$

Mit dem von u abhängigen Differentialoperator

$$A^u f(t, x) = f_t(t, x) + (f_x(t, x))^T b(t, x, u) + \frac{1}{2} \text{Spur}((f_{xx}(t, x))a(t, x, u))$$

lässt sich (9.4) schreiben als

$$0 = \sup_{u \in U} \{ \psi(t, x, u) + A^u V(t, x) \}. \quad (9.5)$$

Die Gleichung (9.4) bzw. (9.5) heißt *Hamilton-Jacobi-Bellman-Gleichung* (kurz: *HJB-Gleichung*). Die obige Argumentation zeigt, dass die Wertfunktion unter gewissen Bedingungen die HJB-Gleichung mit der Endbedingung

$$V(T, x) = \Psi(x), \quad x \in \mathbb{R}^n,$$

löst. Beachte, dass als Konsequenz dieser Umformung hier nun kein zufälliger Ausdruck mehr auftritt. Zudem ist nur noch das Supremum über alle möglichen Startwerte $u \in U$ von zulässigen Steuerungen zu bilden, da nur diese in der partiellen DGL auftreten.

9.3 Ein Verifikationstheorem

In Abschnitt 9.2 haben wir die HJB-Gleichung als ein notwendiges Kriterium für die Wertfunktion des Steuerungsproblems hergeleitet. Es stellt sich die Frage, ob wir die HJB-Gleichung auch als hinreichendes Kriterium heranziehen können. Die Antwort liefert das folgende Hauptresultat der stochastischen Steuerung, ein sog. *Verifikationstheorem*. Dieses stellt ein System von Bedingungen an V, ψ, Ψ, b, σ und u bereit, unter welchem es sich bei einer Lösung V der HJB-Gleichung tatsächlich um die gesuchte Wertfunktion handelt und ein Maximierer von $u \mapsto \psi(t, x, u) + A^u V(t, x)$ eine optimale Steuerung definiert.

Satz 9.1 (*Verifikationstheorem*) Sei $\Phi \in \mathcal{C}^{1,2}([0, T] \times \mathbb{R}^n) \cap \mathcal{C}([0, T] \times \mathbb{R}^n)$ eine Lösung der HJB-Gleichung

$$\sup_{u \in U} \{ \psi(t, x, u) + A^u \Phi(t, x) \} = 0, \quad (t, x) \in [0, T] \times \mathbb{R}^n,$$

$$\Phi(T, x) = \Psi(x), \quad x \in \mathbb{R}^n.$$

Ferner sei Φ „hinreichend integrierbar“ im Sinne von

$$E_{t,x} \left[\int_t^T (\Phi_x(s, X_s^u))^T \sigma(s, X_s^u, u_s) dW_s \right] = 0$$

für alle $(t, x) \in [0, T] \times \mathbb{R}^n, u \in \mathcal{A}(t, x)$. Dann gilt:

(i) $\Phi(t, x) \geq V(t, x)$ für alle $(t, x) \in [0, T] \times \mathbb{R}^n$.

(ii) Angenommen, es existiere für alle $(t, x) \in [0, T] \times \mathbb{R}^n$ ein $\hat{u}(t, x) \in U$ mit

$$\hat{u}(t, x) \in \arg \max_{u \in U} \{\psi(t, x, u) + A^u \Phi(t, x)\},$$

so dass $u^* = (u_t^*)_{t \in [0, T]}$, $u_t^* = \hat{u}(t, X_t^*)$, eine zulässige Steuerung definiert, wobei X_t^* die Lösung der gesteuerten SDGL (9.1) unter der Steuerung u_s^* auf $[0, t]$ ist. Dann gilt

$$\Phi(t, x) = V(t, x) = J(t, x, u^*) \quad \text{für alle } (t, x) \in [0, T] \times \mathbb{R}^n,$$

d.h. Φ ist die Wertfunktion und u^* eine optimale Steuerung.

Beweis: Siehe Vorlesung. □

Beachte: Der Wert X_t^* hängt nicht von u_t^* ab, sondern nur von u_s^* , $s \in [0, t]$. Es handelt sich deshalb in Teil (ii) des Verifikationstheorems nicht um eine implizite, sondern um eine explizite Definition von u_t^* .

Bemerkung 9.2 Analog zum Beweis des Verifikationstheorems kann gezeigt werden, dass unter den dortigen Annahmen sowie der Annahme der Existenz einer $C^{1,2}$ -Lösung der HJB-Gleichung und einer optimalen Steuerung u^* das folgende Bellman-Prinzip gilt:

Für alle Stoppzeiten θ mit Werten in $[t, T]$ gilt

$$V(t, x) = \sup_{u \in \mathcal{A}(t, x)} E_{t, x} \left[\int_t^\theta \psi(s, X_s^u, u_s) ds + V(\theta, X_\theta^u) \right].$$

Bemerkung 9.3 (i) Man kann gewisse technische Bedingungen für eine zulässige Steuerung, für die Koeffizientenfunktionen b und σ , für die Nutzenfunktionen ψ und Ψ und für die Testfunktion Φ aus Satz 9.1 spezifizieren, so dass für alle Startwerte $(t, x) \in [0, T] \times \mathbb{R}^d$ und für alle zulässigen Steuerungen $u \in \mathcal{A}(t, x)$ die gesteuerte SDGL eine eindeutige Lösung X_s^u besitzt, das Nutzenfunktional $J(t, x, u)$ wohldefiniert ist und Φ den Integrationsanforderungen aus Satz 9.1 genügt. Ein mögliches System von Annahmen ist etwa (vgl. [KK01]):

– Es gilt $u \in \mathcal{A}(t, x)$ genau dann, wenn $u = (u_s)_{s \in [t, T]}$ progressiv messbar und U -wertig ist mit

$$E \left[\int_t^T |u_s|^k ds \right] < \infty \quad \text{für alle } k \in \mathbb{N}$$

und die zugehörige gesteuerte SDGL (9.1) eine eindeutige Lösung $(X_s)_{s \in [t, T]}$ mit $X_t = x$ und

$$E_{t,x} \left[\sup_{s \in [t, T]} |X_s|^k \right] < \infty \quad \text{für alle } k \in \mathbb{N}$$

besitzt.

- b, σ sind stetig mit $b(\cdot, \cdot, u), \sigma(\cdot, \cdot, u)$ in $\mathcal{C}^1([0, T] \times \mathbb{R}^n)$ für $u \in U$.
Zusätzlich gelte mit einer Konstanten $C > 0$

$$\begin{aligned} |b_t| + |b_x| &\leq C, & |\sigma_t| + |\sigma_x| &\leq C, \\ |b(t, x, u)| + |\sigma(t, x, u)| &\leq C(1 + |x| + |u|) \end{aligned}$$

auf $[0, T] \times \mathbb{R}^n \times U$.

- ψ, Ψ sind stetig und genügen den polynomialen Wachstumsbedingungen

$$|\psi(t, x, u)| \leq C(1 + |x|^k + |u|^k), \quad \Psi(x) \leq C(1 + |x|^k)$$

für ein $C > 0$ und $k \in \mathbb{N}$ auf $[0, T] \times \mathbb{R}^n \times U$.

- $\Phi \in \mathcal{C}^{1,2}([0, T] \times \mathbb{R}^n) \cap \mathcal{C}([0, T] \times \mathbb{R}^n)$ genügt der polynomialen Wachstumsbedingung

$$|\Phi(t, x)| \leq C(1 + |x|^k)$$

für ein $C > 0$ und $k \in \mathbb{N}$ auf $[0, T] \times \mathbb{R}^n$.

(ii) Da die Wertfunktion eines stochastischen Steuerungsproblems eindeutig ist, kann es auch nur eine Lösung der HJB-Gleichung (mit entsprechender Endbedingung) geben, die die Annahmen von Satz 9.1 erfüllt. D.h. gelten die Bedingungen aus (i), dann existiert höchstens eine $C^{1,2}$ -Lösung der HJB-Gleichung von polynomialen Wachstum. Beachte aber, dass eine optimale Steuerung nicht eindeutig sein muss.

(iii) Die Existenz einer Lösung ist im Allgemeinen schwer nachzuweisen. Ein typisches System von sehr starken Bedingungen ist:

- U ist kompakt.
- Ψ liegt in $\mathcal{C}^3(\mathbb{R}^n)$ und ist beschränkt.
- b, σ, ψ liegen in $\mathcal{C}^{1,2}$ und sind beschränkt.

– Es herrscht gleichmäßige Parabolizität, d.h.

$$y^T a(t, x, u)y \geq c|y|^2 \quad \text{für alle } (t, x) \in [0, T] \times \mathbb{R}^n, u \in U, y \in \mathbb{R}^n.$$

Dann besitzt die HJB-Gleichung eine eindeutige beschränkte Lösung $V \in \mathcal{C}^{1,2}$.

Das Verifikationstheorem 9.1 legt das folgende Vorgehen zur Lösung des Steuerungsproblems nahe:

Algorithmus 9.4 1. Löse das Maximierungsproblem in der HJB-Gleichung (9.5) in Abhängigkeit von der noch unbekanntem Wertfunktion V (und ihrer partiellen Ableitungen).

2. Es sei

$$\hat{u}(t, x) = \hat{u}(t, x, V(t, x), V_t(t, x), V_x(t, x), V_{xx}(t, x))$$

eine Lösung des Maximierungsproblems. Löse dann die partielle DGL

$$\begin{aligned} \psi(t, x, \hat{u}(t, x)) + A^{\hat{u}(t, x)}V(t, x) &= 0, \quad (t, x) \in [0, T) \times \mathbb{R}^n, \\ V(T, x) &= \Psi(x), \quad x \in \mathbb{R}^n. \end{aligned}$$

3. Überprüfe die Gültigkeit der getroffenen Annahmen.

Bemerkung: Die Überprüfung der gemachten Annahmen im 3. Schritt von Algorithmus 9.4 bereitet oft die größten Schwierigkeiten, insbesondere wenn man keine explizite Lösung der HJB-Gleichung finden kann. Leider existieren in den meisten Fällen jedoch keine expliziten Lösungen.

Typen von Steuerungen

Im Allgemeinen ist eine Steuerung zur Zeit t gegeben durch eine Zufallsvariable $u_t(\omega)$. Man kann aber verschiedene Spezialfälle unterscheiden:

- *Deterministische* oder *Open-loop*-Steuerungen: $u_t(\omega) = f(t)$ ist nur eine Funktion in t (nicht zufällig).
- *Feedback*- oder *Closed-loop*-Steuerungen: u ist adaptiert bzgl. der Filtration, die vom gesteuerten Prozess erzeugt wird, d.h. u_t ist $\sigma(X_s^u : s \leq t)$ -messbar.
- Ein Spezialfall von Feedback Steuerungen sind *Markov*-Steuerungen, die von der Form $u_t(\omega) = f(t, X_t(\omega))$ sind.

Insbesondere handelt es sich bei der in Satz 9.1(ii) generierten Steuerung um eine Markov-Steuerung. Das bedeutet, dass man - sofern diese Steuerung existiert - mit einer Markov-Steuerung mindestens genauso gut abschneiden kann wie mit jeder anderen beliebigen Steuerung.

Erweiterungen des Optimierungsproblems

- Interpretieren wir $J(t, x, \cdot)$ als Kostenfunktional, so haben wir es mit einem Minimierungsproblem anstelle eines Maximierungsproblems zu tun. Auch in diesem Fall können wir auf das Verifikationstheorem 9.1 als Lösungsansatz zurückgreifen. Hierzu müssen wir in der HJB-Gleichung lediglich das Supremum durch das Infimum austauschen.
- Auch für den allgemeineren Fall, dass der Zustandsraum der zulässigen Steuerungen von der Zeit und dem Zustand des gesteuerten Prozesses abhängt, bleiben die hergeleiteten Resultate bestehen. In der HJB-Gleichung wird das Supremum dann über dem Raum $U(t, x)$ ermittelt.
- Erweiterungsmöglichkeiten bzgl. der Steuerungsdauer betrachten wir in den folgenden Abschnitten.

9.4 Portfolio-Optimierung: Optimaler Konsum und optimales Endvermögen bei endlichem Zeithorizont

Wir betrachten das Finanzmarktmodell aus dem Beispiel in Abschnitt 9.1:

$$\begin{aligned} \text{Bond:} \quad dB_t &= rB_t dt, \quad B_0 = 1, \\ \text{Aktie:} \quad dS_t &= \mu S_t dt + \sigma S_t dW_t, \quad S_0 = s > 0. \end{aligned}$$

Hier erlauben wir dem Investor auch einen kontinuierlichen Konsum. Als Steuerung betrachten wir daher den 2-dimensionalen Prozess $u = (\pi, c)$, wobei π_t den in die Aktie investierte Anteil am Vermögen und c_t die Konsumrate zur Zeit t angibt. Der zugehörige Vermögensprozess genügt dann der SDGL

$$dX_t^u = ((r + (\mu - r)\pi_t)X_t^u - c_t)dt + \sigma\pi_t X_t^u dW_t, \quad X_0^u = x > 0.$$

Der Investor erzielt den Nutzen $U_1(c_t)$ aus dem Konsum und den Endnutzen $U_2(X_T^u)$. Wir wollen nun für die Nutzenfunktionen

$$U_1(x) = U_2(x) = U(x) := \frac{x^\gamma}{\gamma}, \quad \gamma \in (0, 1),$$

und einen Diskontierungsfaktor $e^{-\beta t}$, $\beta \geq 0$, das Funktional

$$J(t, x, u) = E_{t,x} \left[\int_t^T e^{-\beta s} U(c_s) ds + e^{-\beta T} U(X_T^u) \right]$$

maximieren. Mit dem Differentialoperator

$$A^{(\pi,c)} f(t, x) = f_t(t, x) + ((r + (\mu - r)\pi)x - c)f_x(t, x) + \frac{1}{2}\sigma^2\pi^2x^2f_{xx}(t, x)$$

lautet die assoziierte HJB-Gleichung

$$\sup_{(\pi,c) \in \mathbb{R}^2} \left\{ e^{-\beta t} \frac{c^\gamma}{\gamma} + A^{(\pi,c)} V(t, x) \right\} = 0.$$

Wende nun Algorithmus 9.4 an:

1. Schritt: Löse das Maximierungsproblem in der HJB-Gleichung.

Unter der Annahme, dass $V(t, x)$ strikt monoton wachsend und strikt konkav in x ist und der zugehörige Vermögensprozess X_t^u strikt positiv ist (diese Annahmen müssen später überprüft werden!), erhalten wir die Maximierer

$$\hat{c}(t, x) = (e^{\beta t} V_x(t, x))^{\frac{1}{\gamma-1}}, \quad (9.6)$$

$$\hat{\pi}(t, x) = -\frac{\mu - r}{\sigma^2} \frac{V_x(t, x)}{x V_{xx}(t, x)}. \quad (9.7)$$

2. Schritt: Löse die partielle DGL.

Einsetzen von (9.6)-(9.7) in die HJB-Gleichung führt zu der partiellen DGL

$$\frac{1-\gamma}{\gamma} e^{-\frac{\beta t}{1-\gamma}} V_x^{\frac{\gamma}{\gamma-1}} + V_t + rxV_x - \frac{1}{2} \left(\frac{\mu-r}{\sigma} \right)^2 \frac{V_x^2}{V_{xx}} = 0 \quad (9.8)$$

mit der Endbedingung

$$V(T, x) = e^{-\beta T} \frac{x^\gamma}{\gamma}.$$

Diese Endbedingung legt den Ansatz

$$V(t, x) = f(t) \frac{x^\gamma}{\gamma}$$

nahe. Bildet man hiervon alle in (9.8) vorkommenden Ableitungen und setzt deren Form in (9.8) ein, so erhält man nach Division durch x^γ/γ die gewöhnliche DGL

$$0 = (1-\gamma)e^{-\frac{\beta t}{1-\gamma}} f(t)^{-\frac{\gamma}{1-\gamma}} + f'(t) + \gamma \left(r + \frac{\gamma}{2(1-\gamma)} \left(\frac{\mu-r}{\sigma} \right)^2 \right) f(t).$$

mit der Endbedingung $f(T) = e^{-\beta T}$. (Beachte, dass es hierfür wesentlich ist, dass der Quotient V_x^2/V_{xx} durch den gewählten Ansatz eine sehr einfache Form erhält.) Mit den Abkürzungen

$$\begin{aligned} a_1 &:= \gamma \left(r + \frac{\gamma}{2(1-\gamma)} \left(\frac{\mu-r}{\sigma} \right)^2 \right), \\ a_2(t) &:= (1-\gamma)e^{-\frac{\beta t}{1-\gamma}} \end{aligned}$$

ergibt sich

$$f'(t) = -a_1 f(t) - a_2(t) f(t)^{-\frac{\gamma}{1-\gamma}}, \quad f(T) = e^{-\beta T}.$$

Die Substitution

$$g(t) = f(t)^{\frac{1}{1-\gamma}}$$

führt dann zu

$$\begin{aligned} g'(t) &= \frac{1}{1-\gamma} f(t)^{\frac{\gamma}{1-\gamma}} f'(t) \\ &= \frac{1}{1-\gamma} f(t)^{\frac{\gamma}{1-\gamma}} \left(-a_1 f(t) - a_2(t) f(t)^{-\frac{\gamma}{1-\gamma}} \right) \\ &= -\frac{a_1}{1-\gamma} f(t)^{\frac{1}{1-\gamma}} - \frac{a_2(t)}{1-\gamma}. \end{aligned}$$

Also:

$$g'(t) = -\frac{a_1}{1-\gamma} g(t) - e^{-\frac{\beta t}{1-\gamma}}, \quad g(T) = e^{-\frac{\beta T}{1-\gamma}}. \quad (9.9)$$

Explizites Lösen von (9.9) mittels Variation der Konstanten ergibt (für $\beta \neq a_1$)

$$g(t) = e^{-\frac{\beta T}{1-\gamma}} e^{\frac{a_1}{1-\gamma}(T-t)} + \frac{(1-\gamma)e^{-\frac{a_1}{1-\gamma}t}}{a_1 - \beta} \left(e^{\frac{a_1 - \beta}{1-\gamma}T} - e^{\frac{a_1 - \beta}{1-\gamma}t} \right) \quad (9.10)$$

(Übungsaufgabe). Hieraus ergibt sich wegen $V(t, x) = g(t)^{1-\gamma} \frac{x^\gamma}{\gamma}$ die explizite Darstellung der Wertfunktion. Ferner erhalten wir für die Feedback-Funktionen (9.6)-(9.7)

$$\begin{aligned} \hat{c}(t, x) &= \frac{e^{-\frac{\beta t}{1-\gamma}}}{g(t)} x, \\ \hat{\pi}(t, x) &= \frac{1}{1-\gamma} \frac{\mu-r}{\sigma^2}. \end{aligned}$$

Die optimale Steuerung ist somit gegeben durch

$$c_t^* = \hat{c}(t, X_t^*) = \frac{e^{-\frac{\beta t}{1-\gamma}}}{g(t)} X_t^*, \quad (9.11)$$

$$\pi_t^* = \hat{\pi}(t, X_t^*) = \frac{1}{1-\gamma} \frac{\mu - r}{\sigma^2} \quad (9.12)$$

(mit $X^* := X^{(c^*, \pi^*)}$).

3. Schritt: Überprüfung der gemachten Annahmen.

$g(t)$ ist strikt positiv. (Beachte hierzu, dass a_1, β beide positiv sind und dass das Vorzeichen des Klammerausdrucks in (9.10) gleich dem seines Vorfaktors ist.) Weiter liegt V in $\mathcal{C}^{1,2}$ und genügt einer polynomialen Wachstumsbedingung, so dass die Integrationsbedingung aus Satz 9.1 erfüllt ist. Für ein strikt positives Vermögen $x > 0$ ist V wegen

$$V_x(t, x) = g(t)^{1-\gamma} x^{\gamma-1} > 0, \quad V_{xx}(t, x) = -(1-\gamma)g(t)^{1-\gamma} x^{\gamma-2} < 0.$$

strikt monoton wachsend und strikt konkav. Der mit (9.11)-(9.12) gesteuerte Prozess X^* genügt der SDGL

$$\begin{aligned} dX_t^* &= X_t^* \left\{ \left(r + \frac{1}{1-\gamma} \left(\frac{\mu - r}{\sigma} \right)^2 - \frac{e^{-\frac{\beta t}{1-\gamma}}}{g(t)} \right) dt + \frac{1}{1-\gamma} \frac{\mu - r}{\sigma} dW_t \right\} \\ &=: X_t^* (\tilde{b} dt + \tilde{\sigma} dW_t). \end{aligned}$$

Es folgt

$$X^*(t) = x \exp \left\{ \left(\tilde{b} - \frac{1}{2} \tilde{\sigma}^2 \right) t + \tilde{\sigma} W_t \right\} > 0,$$

d.h. der optimal gesteuerte Prozess ist strikt positiv. Da π^* konstant und $g(t)^{-1}$ beschränkt ist, kann man die Zulässigkeitsbedingung für die optimale Steuerung nachweisen, indem man einen passenden Zulässigkeitsbereich $[0, \infty) \times [\underline{\pi}, \bar{\pi}]$ für (c, π) wählt.

9.5 Unendlicher Zeithorizont

In diesem Abschnitt nehmen wir an, dass b, σ, ψ zeitunabhängig sind, d.h. der gesteuerte Prozess ist gegeben durch

$$dX_t = b(X_t, u_t) dt + \sigma(X_t, u_t) dW_t, \quad X_0 = x. \quad (9.13)$$

Wir definieren den zugehörigen Differentialoperator, in Abhängigkeit von u , durch

$$Af(x) = (f_x(x))^T b(x, u) + \frac{1}{2} \text{Spur}((f_{xx}(x))^T a(x, u)).$$

Als Optimierungskriterium betrachten wir

$$J(x, u) = E_x \left[\int_0^\infty e^{-\beta s} \psi(X_s, u_s) ds \right]$$

mit Diskontierungsfaktor $\beta > 0$. Die Bedingungen für b, σ, ψ und für die Zulässigkeit einer Steuerung - hier dargestellt durch die Klasse $\mathcal{A}(x)$ - werden analog zum Fall mit endlichem Zeithorizont definiert. Die Wertfunktion lautet

$$V(x) = \sup_{u \in \mathcal{A}(x)} J(x, u).$$

Satz 9.5 (Verifikationstheorem) Sei $\Phi \in \mathcal{C}^2(\mathbb{R}^n)$ eine Lösung der HJB-Gleichung

$$\sup_{u \in U} \{ \psi(x, u) + A^u \Phi(x) - \beta \Phi(x) \} = 0, \quad x \in \mathbb{R}^n.$$

Ferner sei Φ „hinreichend integrierbar“ im Sinne von

$$E_x \left[\int_0^\infty (\Phi_x(X_s^u))^T \sigma(X_s^u, u_s) dW_s \right] = 0$$

für alle $x \in \mathbb{R}^n$, $u \in \mathcal{A}(x)$. Dann gilt:

(i) $\Phi(x) \geq V(x)$ für alle $x \in \mathbb{R}^n$.

(ii) Angenommen, es existiere für alle $x \in \mathbb{R}^n$ ein $\hat{u}(x) \in U$ mit

$$\hat{u}(x) \in \arg \max_{u \in U} \{ \psi(x, u) + A^u \Phi(x) - \beta \Phi(x) \},$$

so dass $u^* = (u_t^*)_{t \geq 0}$, $u_t^* = \hat{u}(X_t^*)$, eine zulässige Steuerung definiert, wobei X_t^* die Lösung der gesteuerten SDGL (9.13) unter der Steuerung u_s^* auf $[0, t)$ ist. Dann gilt

$$\Phi(x) = V(x) = J(x, u^*) \quad \text{für alle } x \in \mathbb{R}^n,$$

d.h. Φ ist die Wertfunktion und u^* eine optimale Steuerung.

Beweis: Analog zum Beweis von Satz 9.1. □

9.6 Portfolio-Optimierung: Optimaler Konsum bei unendlichem Zeithorizont

Betrachte das Modell aus Abschnitt 9.4. Wir wollen nun das folgende Lebenszeit-Konsum-Problem lösen:

$$\sup_{(\pi, c) \in \mathcal{A}(x)} E_x \left[\int_0^\infty e^{-\beta t} \frac{1}{\gamma} c_t^\gamma dt \right]$$

mit $\beta > 0$, $\gamma \in (0, 1)$, $x > 0$. Als zusätzliche Zulässigkeitsbedingung an eine Steuerung $u = (\pi, c)$ fordern wir, dass der gesteuerte Vermögensprozess

$$dX_t^u = ((r + (\mu - r)\pi_t)X_t^u - c_t)dt + \sigma\pi_t X_t^u dW_t, \quad X_0^u = x,$$

f.s. positiv ist, d.h. $P(X_t > 0) = 1$ für alle $t \geq 0$. Die HJB-Gleichung für die zugehörige Wertfunktion $V(x)$ lautet

$$\sup_{(\pi, c) \in \mathbb{R}^2} \left\{ (r + (\mu - r)\pi)x - c \right\} V'(x) - \beta V + \frac{1}{2} \sigma^2 \pi^2 x^2 V''(x) + \frac{1}{\gamma} c^\gamma \Big\} = 0.$$

Es ergeben sich wieder die folgenden Schritte:

1. Schritt: Maximierung:

Angenommen, es gilt $V'(x) > 0$ und $V''(x) < 0$. Dann erhalten wir die folgenden Maximierer:

$$\begin{aligned} \hat{\pi}(x) &= -\frac{\mu - r}{\sigma^2} \frac{V'(x)}{xV''(x)}, \\ \hat{c}(x) &= V'(x)^{\frac{1}{\gamma-1}}. \end{aligned}$$

2. Schritt: Lösen der DGL:

Einsetzen von $(\hat{\pi}, \hat{c})$ in die HJB-Gleichung liefert die DGL

$$-\frac{1}{2} \left(\frac{\mu - r}{\sigma} \right)^2 \frac{V'(x)^2}{V''(x)} + rxV'(x) - \beta V + \frac{1 - \gamma}{\gamma} V'(x)^{\frac{\gamma}{\gamma-1}} = 0.$$

Wir versuchen den Ansatz $V(x) = \frac{1}{\gamma} Ax^\gamma$ für ein $A > 0$. Einsetzen der Ableitungen $V'(x) = Ax^{\gamma-1}$, $V''(x) = -(1 - \gamma)Ax^{\gamma-2}$ in die obige DGL und Division durch den Faktor Ax^γ liefert dann die Gleichung

$$\frac{1}{2} \frac{1}{1 - \gamma} \left(\frac{\mu - r}{\sigma} \right)^2 + r - \frac{\beta}{\gamma} + \frac{1 - \gamma}{\gamma} A^{\frac{1}{\gamma-1}} = 0.$$

Diese Gleichung hat genau dann eine positive Lösung A , wenn

$$\beta > \frac{1}{2} \frac{\gamma}{1-\gamma} \left(\frac{\mu-r}{\sigma} \right)^2 + \gamma r$$

gilt, d.h. falls der Abzinsungsfaktor β hinreichend groß gewählt ist. In diesem Fall gilt

$$A = \left(\frac{\gamma}{1-\gamma} \left(\frac{\beta}{\gamma} - r - \frac{1}{2} \frac{1}{1-\gamma} \left(\frac{\mu-r}{\sigma} \right)^2 \right) \right)^{\gamma-1}.$$

3. Schritt: Überprüfen der Annahmen:

Als Kandidaten für eine optimale Steuerung erhalten wir somit

$$\begin{aligned} \pi_t^* &= \frac{1}{1-\gamma} \frac{\mu-r}{\sigma^2}, \\ c_t^* &= A^{\frac{1}{\gamma-1}} X_t^*, \end{aligned}$$

wobei der zugehörige Vermögensprozess X^* die SDGL

$$dX_t^* = X_t^* \left\{ \left(r + \frac{1}{1-\gamma} \left(\frac{\mu-r}{\sigma} \right)^2 - A^{\frac{1}{\gamma-1}} \right) dt + \frac{1}{1-\gamma} \frac{\mu-r}{\sigma} dW_t \right\}, \quad X_0 = x,$$

erfüllt. Diese SDGL hat die eindeutige Lösung

$$X_t^* = x \cdot \exp \left\{ \left(r + \frac{1}{1-\gamma} \left(\frac{\mu-r}{\sigma} \right)^2 - A^{\frac{1}{\gamma-1}} - \frac{1}{2} \left(\frac{1}{1-\gamma} \right)^2 \left(\frac{\mu-r}{\sigma} \right)^2 \right) t + \frac{1}{1-\gamma} \frac{\mu-r}{\sigma} W_t \right\}$$

die strikt positiv ist. Somit schließen wir $V'(x) > 0$ und $V''(x) < 0$. (Klar: $V \in \mathcal{C}^2((0, \infty))$.) Die Integrierbarkeitsbedingung folgt aus der polynomialen Wachstumseigenschaft von V . Damit sind die Voraussetzungen von Satz 9.5 erfüllt, so dass die gefundenen Kandidaten die gesuchte Optimallösung darstellen.

9.7 Stoppen des gesteuerten Prozesses

Wir betrachten erneut das stochastische Steuerungsproblem aus Abschnitt 9.1. Allerdings soll jetzt nur so lange gesteuert werden, wie sich der gesteuerte Prozess X_t innerhalb einer vorgegebenen offenen Teilmenge $\mathcal{O} \subset \mathbb{R}^n$ bewegt. Hierzu setzen wir $\mathcal{S} := [0, T) \times \mathcal{O}$ und definieren durch

$$\tau := \inf \{s \in [t, T) \mid (s, X_s) \notin \mathcal{S}\}$$

einen Abbruchzeitpunkt. Ferner bezeichnen wir mit $\partial\mathcal{S} := ([0, T] \times \partial\mathcal{O}) \cup (\{T\} \times \bar{\mathcal{O}})$ den parabolischen Rand von \mathcal{S} (wobei $\partial\mathcal{O}$ den Rand und $\bar{\mathcal{O}}$ den Abschluss von \mathcal{O} bezeichne). Als Optimierungskriterium betrachten wir das Nutzenfunktional

$$J(t, x, u) = E_{t,x} \left[\int_t^\tau \psi(s, X_s, u_s) ds + \Psi(\tau, X_\tau) \right].$$

Die Wertfunktion lautet wie gewohnt

$$V(t, x) = \sup_{u \in \mathcal{A}(t,x)} J(t, x, u),$$

und die Bedingungen für b, σ, ψ, Ψ und für die Zulässigkeit einer Steuerung entsprechen denen aus dem ursprünglichen Problem mit $\mathcal{O} = \mathbb{R}^n$.

Satz 9.6 (*Verifikationstheorem*) Sei $\Phi \in \mathcal{C}^{1,2}(\mathcal{S}) \cap \mathcal{C}(\bar{\mathcal{S}})$ eine Lösung der HJB-Gleichung

$$\begin{aligned} \sup_{u \in U} \{ \psi(t, x, u) + A^u \Phi(t, x) \} &= 0, \quad (t, x) \in \mathcal{S}, \\ \Phi(t, x) &= \Psi(t, x), \quad (t, x) \in \partial\mathcal{S}. \end{aligned}$$

Ferner sei Φ „hinreichend integrierbar“ im Sinne von

$$E_{t,x} \left[\int_t^\tau (\Phi_x(s, X_s^u))^T \sigma(s, X_s^u, u_s) dW_s \right] = 0$$

für alle $(t, x) \in \mathcal{S}$, $u \in \mathcal{A}(t, x)$. Dann gilt:

(i) $\Phi(t, x) \geq V(t, x)$ für alle $(t, x) \in \mathcal{S}$.

(ii) Angenommen, es existiere für alle $(t, x) \in \bar{\mathcal{S}}$ ein $\hat{u}(t, x) \in U$ mit

$$\hat{u}(t, x) \in \arg \max_{u \in U} \{ \psi(t, x, u) + A^u \Phi(t, x) \},$$

so dass $u^* = (u_t^*)_{t \in [0, \tau]}$, $u_t^* = \hat{u}(t, X_t^*)$, eine zulässige Steuerung definiert, wobei X_t^* die Lösung der gesteuerten SDGL (9.1) unter der Steuerung u_s^* auf $[0, t]$ ist. Dann gilt

$$\Phi(t, x) = V(t, x) = J(t, x, u^*) \quad \text{für alle } (t, x) \in \bar{\mathcal{S}},$$

d.h. Φ ist die Wertfunktion und u^* eine optimale Steuerung.

Beweis: Siehe Vorlesung. □

9.8 Portfolio-Optimierung: Portfolio-Versicherung

Betrachte weiterhin den Finanzmarkt aus Abschnitt 9.4, jetzt jedoch ohne Konsumkomponente, d.h. wir benutzen die Steuerung $u_t = \pi_t$ (wobei π_t der in die Aktie investierte Vermögensanteil ist). Für $\gamma \in (0, 1)$, $G > 0$ und einen Startwert $(t, x) \in [0, T] \times \mathbb{R}$ wollen wir das folgende restringierte stochastische Steuerungsproblem lösen:

$$\sup_{\pi} E_{t,x} \left[\frac{1}{\gamma} X_T^\gamma \right].$$

NB : $P(X_T \geq G) = 1$

Dieses Problem wird *Portfolio-Versicherungs-Problem* genannt, da das Vermögen durch den garantierten Betrag G nach unten abgesichert ist. Diese Erweiterung des ursprünglichen unrestringierten Problems ist durchaus sinnvoll, da die Verteilung des optimalen Endvermögens im unrestringierten Problem hohe Verluste mit hoher Wahrscheinlichkeit erlaubt (sog. *Fat-Tail* Problematik).

Wir unterscheiden 3 Fälle:

- Gilt $x < e^{-r(T-t)}G$, so können wir die garantierte Auszahlung G zum Zeitpunkt T nicht mit W'keit 1 erzielen, da die reine Bondstrategie in diesem Fall nur den Betrag $xe^{r(T-t)} < G$ generiert.
- Für $x = e^{-r(T-t)}G$, liefert die reine Bondstrategie exakt den Betrag G . Somit können wir nicht in die Aktie investieren, da wir sonst mit positiver W'keit Verluste machen würden, so dass wir unser Ziel G mit positiver W'keit verfehlen würden.
- Für $x > e^{-r(T-t)}G$ sind sowohl Investitionen in den Bond als auch in die Aktie möglich.

Im Folgenden wollen wir daher von $x > e^{-r(T-t)}G$ ausgehen. Dazu definieren wir

$$\mathcal{S} = \{(t, x) \in [0, T] \times \mathbb{R} : x > e^{-r(T-t)}G\}$$

und damit

$$\partial\mathcal{S} = ([0, T] \times \{Ge^{-r(T-t)}\}) \cup (\{T\} \times [G, \infty)).$$

Dann folgt:

- Randbedingungen:

$$\Psi(t, x) = \begin{cases} \frac{1}{\gamma} G^\gamma, & (t, x) \in \partial\mathcal{S}, t < T, \\ \frac{1}{\gamma} x^\gamma, & t = T. \end{cases}$$

- Abbruchzeitpunkt:

$$\tau = \inf\{s \geq t : (s, X_s) \notin \mathcal{S}\}.$$

- Zielfunktional:

$$J(t, x, \pi) = E_{t,x} \left[\frac{1}{\gamma} X_\tau^\gamma \right], \quad (t, x) \in \mathcal{S}.$$

- Wertfunktion:

$$V(t, x) = \sup_{\pi \in \mathcal{A}(t,x)} J(t, x, \pi)$$

für eine „geeignete“ Menge $\mathcal{A}(t, x)$ zulässiger Steuerungen.

- HJB-Gleichung:

$$\sup_{\pi \in \mathbb{R}} \left\{ V_t + (r + (\mu - r)\pi)xV_x + \frac{1}{2}\sigma^2\pi^2x^2V_{xx} \right\} = 0, \quad (t, x) \in \mathcal{S},$$

$$V(t, x) = \Psi(t, x), \quad (t, x) \in \partial\mathcal{S}.$$

Optimalitätsbedingungen 1. Ordnung liefern für eine optimale Portfolio-Strategie den Kandidaten

$$\hat{\pi}(t, x) = -\frac{\mu - r}{\sigma^2} \frac{V_x(t, x)}{xV_{xx}(t, x)}.$$

Dies führt zu der partiellen DGL

$$V_t + rxV_x - \frac{1}{2} \left(\frac{\mu - r}{\sigma} \right)^2 \frac{V_x^2}{V_{xx}} = 0, \quad (t, x) \in \mathcal{S},$$

mit Randbedingungen

$$V(t, x) = \Psi(t, x), \quad (t, x) \in \partial\mathcal{S}.$$

Beachte, dass wir im unrestringierten Fall dieselbe partielle DGL lösen müssen, jedoch ausschließlich mit der Endbedingung

$$V(T, x) = \frac{1}{\gamma} x^\gamma.$$

Wie in Abschnitt 9.4 verifiziert man, dass die Wertfunktion V^0 für das unrestringierte Portfolio-Problem ohne Konsum gegeben ist durch

$$V^0(t, x) = \exp \left\{ \left(\gamma r + \frac{\gamma}{2(1-\gamma)} \left(\frac{\mu - r}{\sigma} \right)^2 \right) (T - t) \right\} \frac{x^\gamma}{\gamma}, \quad (t, x) \in [0, T] \times [0, \infty).$$

Da V^0 jedoch die Randbedingung für $t < T$ nicht erfüllt, können wir keine explizite Lösung von V durch Faktorisierung in Form von $f(t)x^\gamma$ herleiten. Daher wendet man in der Praxis ein *finites Differenzenverfahren* an, um eine numerische Näherung von V zu erhalten.

Literaturverzeichnis

- [B93] H. Bauer: *Wahrscheinlichkeitstheorie*. deGruyter, 1993.
- [GJ03] M. Günther, A. Jüngel: *Finanzderivate mit MATLAB : Mathematische Modellierung und numerische Simulation*, Vieweg, 2003
- [Ir01] A. Irle: *Wahrscheinlichkeitstheorie und Statistik: Grundlagen - Resultate - Anwendungen*. Teubner, 2001.
- [Ir03] A. Irle: *Finanzmathematik: Die Bewertung von Derivaten*. Teubner, 2003.
- [KS99] I. Karatzas, S. E. Shreve: *Brownian motion and stochastic calculus*. Springer, 1999.
- [Kol00] M. Koller: *Stochastische Modelle in der Lebensversicherung*. Springer, 2000.
- [KK01] E. Korn, R. Korn: *Optionsbewertung und Portfolio-Optimierung*. Vieweg, 2001.
- [Ok00] B. Øksendal: *Stochastic Differential Equations*. Springer, 2000.