Finanzmathematik
(Bachelorkurs)

9. November 2024

Zusammenfassung

Dieses Skriptum entsteht parallel zur Vorlesung Finanzmathematik (Universität Wien, WS 2024). Inhaltlich orientieren wir uns wesentlich an den Texten [1, 2] sowie an früheren Vorlesungen von Christa Cuchiero.

Fokus der Vorlesung werden insbesondere Modelle in diskreter Zeit sein, erst gegen Ende der Vorlesung werden wir auch kurz die Modelle von Bachelier und Black–Scholes betrachten.

Wichtige Themen der Vorlesung sind: Arbitragetheorie, Vollständigkeit von Finanzmärkten, Derivate, Preistheorie, Nutzenmaximierung.

1 Grundlegende Konzepte aus der Wahrscheinlichkeitstheorie

1.1 Filtrierte Wahrscheinlichkeitsräume, Zufallsvariablen, stochastische Prozesse

Ein Wahrscheinlichkeitsraum besteht aus den Komponenten:

1. nicht leere Menge Ω, der Ereignisraum,
2. eine nicht leere σ-Algebra $\mathcal{F}$ die in der Potenzmenge $\mathcal{P}(\Omega)$ enthalten ist, Elemente der Menge $\mathcal{F}$ repräsentieren die Ereignisse die eintreten können,
3. ein Wahrscheinlichkeitsmaß $\mathbb{P}$ das den Ereignissen Wahrscheinlichkeiten zuordnet.

Genauer gilt: $\mathcal{F} \subseteq \mathcal{P}(\Omega)$ ist eine σ-Algebra genau dann wenn

- $\Omega \in \mathcal{F}$.
- Falls $A \in \mathcal{F}$ dann gilt auch $A^c \in \mathcal{F}$.

- Falls $A_1, A_2, \ldots \in F$ dann gilt auch $\bigcup_{n=1}^{\infty} A_n \in \mathcal{F}$.

Wir verwenden meist die Buchstaben $\mathcal{F}, \mathcal{G}$ um σ-Algebren zu bezeichnen.

Gegeben eine σ-Algebra $\mathcal{F}$ auf Ω ist ein Wahrscheinlichkeitsmaß ist eine Funktion $\mathbb{P} : \mathcal{F} \to [0,1]$ die folgende Eigenschaften erfüllt:

- $\mathbb{P}(\Omega) = 1$.
- Sind $A_1, A_2, \ldots \in \mathcal{F}$ paarweise disjunkt, dann gilt $\mathbb{P}(\bigcup_{n=1}^{\infty} A_n) = \sum_{n=1}^{\infty} \mathbb{P}(A_n)$.

Ein Wahrscheinlichkeitsraum besteht also aus einem Tripel $(\Omega, \mathcal{F}, \mathbb{P})$ mit obigen Eigenschaften. Das Paar $(\Omega, \mathcal{F})$ wird als messbarer Raum oder Meßraum bezeichnet.

Für uns spielen σ-Algebren vor allem deswegen eine wichtige Rolle, weil sie ein technisch sehr praktisches Hilfsmittel sind, um den zeitlichen Fluß von Information darzustellen. Dies geschieht über den Begriff der Filtration.

Definition 1.1. *Sei $T \in \mathcal{N}$. Unter einer Filtration auf einem Meßraum $(\Omega, \mathcal{F})$ verstehen wir eine aufsteigende Folge von σ-Algebren $\mathcal{F}_0 \subseteq \mathcal{F}_1 \subseteq \ldots \subseteq \mathcal{F}_T \subseteq \mathcal{F}$.*

T spielt für uns die Rolle eines Zeithorizonts. Wir stellen uns vor, das wir heute, d.h. zum Zeitpunkt $t = 0$ über die Information verfügen, die von $\mathcal{F}_0$ kodiert wird, morgen am Tag $t = 1$ werden wir über die Information in $\mathcal{F}_1$ verfügen usw.

Das Quadtrupel $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t=0}^{T}, \mathbb{P})$ heißt *filtrierter Wahrscheinlichkeitsraum.* Wie in der Wahrscheinlichkeitstheorie üblich, nehmen wir an, dass ein fester filtrierter Wahrscheinlichkeitsraum stets im Hintergrund gegeben ist, nur bei manchen Beispielen werden wir angeben wie dieser genau gewählt wird.

In dieser Vorlesung werden wir (falls nicht anders gesagt) unter der folgenden Voraussetzung arbeiten:

Annahme 1.2. *Wir nehmen an, dass $|\Omega| < \infty$, $\mathcal{F}_0 = \{\emptyset, \Omega\}$, $\mathcal{F}_T = \mathcal{F} = \mathcal{P}(\Omega)$. Wir schreiben $\Omega =: \{\omega_1, \ldots, \omega_N\}$ wobei $\omega_i \neq \omega_j$ für $i \neq j$. Weiters nehmen wir an, dass $\mathbb{P}(\omega_i) > 0$ für alle i.*

Offenbar handelt es sich hier um eine starke Voraussetzung. Wir treffen sie, weil sie unsere Beweise sehr erleichtern wird und in allen interessanten Modellen für den Fall diskreter Zeit erfüllt ist. Unter geeigneten Zusatzvoraussetzungen lassen sich die meisten Resultate dieser Vorlesung auch ohne Annahme 1.2 (sowie in stetiger statt diskreter Zeit) beweisen.

Im Fall $|\Omega| < \infty$ sind natürlich auch alle σ-Algebren auf Ω endlich und haben damit eine besonders einfache Gestalt: sie entsprechen gerade Partitionen von Ω. Unter einer Partition von Ω verstehen wir eine endlich Folge $A_1, \ldots, A_n$ paarweise disjunkter, nichtleerer Teilmengen von Ω für die gilt das $A_1 \cup \ldots \cup A_n = \Omega$. Wir werden zwei Partitionen als ident betrachten, sobald sie bis auf die Nummerierung der einzelnen Zellen überein stimmen. Man kann dann leicht folgendes zeigen:

1. Sei $A_1, \ldots, A_n$ eine Partition von Ω. Dann gilt für die erzeugte σ-Algebra
$$\sigma(A_1, \ldots, A_n) = \left\{ \bigcup_{i \in F} : F \subseteq \{1, \ldots, n\} \right\}.$$
2. Umgekehrt kann man jeder σ-Algebra $\mathcal{F}$ eindeutig eine Partition zuordnen, indem man für $\omega \in \Omega$ die Zellen
$$\bigcap_{A \in \{B \in \mathcal{F} : \omega \in B\}} A$$
betrachtet.

Aus der Maßtheorie bekannte Eigenschaften lassen sich oft leicht mithilfe der einer σ-Algebra zugehörigen Partition ausdrücken. Wir betrachten die Eigenschaft bezüglich einer σ-Algebra messbar zu sein. Wir erinnern uns, dass eine Funktion (Zufallsvariable) $X : \Omega \to \mathbb{R}$ genau dann $\mathcal{G}$-messbar ist, wenn für alle $a \in \mathbb{R}$
$$\{X \leq a\} = X^{-1}((-\infty, a]) \in \mathcal{G}.$$
Wir schreiben dafür auch kurz
$$X \in \mathcal{G}.$$
Mithilfe von Partitionen lässt sich das wie folgt ausdrücken:

Theorem 1.3. *Sei $A_1, \ldots, A_n$ eine Partition von Ω. Dann ist X genau dann $\mathcal{G}$-messbar wenn es $a_1, \ldots, a_n \in \mathbb{R}$ gibt sodass*
$$X = \sum_{i \leq n} a_i I_{A_i}.$$

Beweis. Übung. □

Ein stochastischer Prozess ist eine Folge von Zufallsvariablen $X_0, X_1, \ldots, X_T$. Wir betrachten ausschließlich Prozesse bei denen X_t jeweils aus der am Tag t vorhanden Information bestimmt werden kann:

Definition 1.4. *Ein* adaptierter *Prozess ist eine Folge von Zufallsvariablen $X_0, X_1, \ldots, X_T$ mit $X_t \in \mathcal{F}_t$ für $t \leq T$.*

Jeder Prozess X ist adaptiert bezüglich der sogenannten *natürlichen Filtration*, das ist die von X erzeugte Filtration $(\mathcal{F}_t^0)_t$ wobei $\mathcal{F}_t^0 := \sigma(X_0, \ldots, X_t)$ für $t \leq T$. In konkreten Beispielen definieren wie oft nur den zugrundeliegenden Prozess und erwähnen die entsprechende Filtration gar nicht. In diesem Fall ist implizit gemeint, dass die betrachtete Filtration gerade die natürliche Filtration ist.

Wenn wir von einem Modell für den Aktienmarkt sprechen, meinen wir in dieser Vorlesung schlicht einen adaptierten Prozess $X = (X_t)_{t=0}^T$ auf einem filtrierten Wahrscheinlichkeitsraum.

Aktien werden für uns dadurch interessant, dass wir auch damit handeln können. Wenn wir vom Zeitpunkt $t-1$ bis zum Zeitpunkt t eine Aktie besitzen, so ändert sich unser Vermögen gerade um den Gewinn oder Verlust den die Aktie in dieser Periode macht, d.h. um

$$X_t - X_{t-1}.$$

Etwas idealisierend nehmen wir an, dass wir beliebige, auch negative und nicht ganzzahlige Anteile von Aktien kaufen dürfen. Die wesentliche Einschränkung die wir machen, ist dass die Anzahl an Aktien die wir während einer Periode besitzen, durch die am Anfang der Periode gegebene Information bestimmt ist. Gegeben eine $\mathcal{F}_{t-1}$-messbare Zufallsvariable H können wir uns entschließen, von $t-1$ bis t gerade H Aktien zu halten. Der entsprechende Gewinn / Verlust ist dann gerade

$$H(X_t - X_{t-1}).$$

Natürlich wollen wir nicht nur in dieser einen Periode Aktien kaufen sondern zu beliebigen Zeitpunkten. Das führt auf folgende Definition.

Definition 1.5. *Ein stochastischer Prozess $H_1, H_2, \ldots, H_T$ heißt* previsibel *oder* vorhersehbar *falls $H_t \in \mathcal{F}_{t-1}$ für $t = 1, 2, \ldots, T$. Unter einer* Handelsstrategie *verstehen wir gerade einen previsiblen Prozess . Wir schreiben $\mathcal{H}$ für die Menge aller Handelsstrategien.*

Für eine Handelsstrategie H ist der entsprechende Gewinn-/Verlustprozeß gegeben durch

$$(H \cdot X)_t := \sum_{k=1}^{t} H_k(X_k - X_{k-1}),$$

wobei $t = 0, 1, \ldots, T$. Die Menge alle Gewinne/Verluste die erzielt werden können, indem man von über den gesamten Zeitraum handelt, bezeichnen wir mit

$$K := \{(H \cdot X)_T : H \in \mathcal{H}\}.$$

1.2 Bedingte Erwartung und Martingale

Ein zentraler Begriff der Finanzmathematik ist der des Martingals, der wiederum auf der bedingten Erwartung aufbaut.

Wir erinnern uns, dass der Erwartungswert einer Zufallsvariable bezüglich des Wahrscheinlichkeitsmaßes $\mathbb{P}$ durch

$$\mathbb{E}[X] := \mathbb{E}_{\mathbb{P}}[X] := \sum_{i=1}^{N} X(\omega_i)\mathbb{P}(\{\Omega_i\})$$

gegeben ist.

Definition 1.6. *Seien $A, B \in \mathcal{F}, \mathbb{P}(B) > 0$. Dann ist die* bedingte Wahrscheinlichkeit *von A gegeben B definiert durch*

$$\mathbb{P}(A|B) := \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)}.$$

Sei $B \in \mathcal{F}, \mathbb{P}(B) > 0$ und X eine Zufallsvariable. Dann ist die bedingte Erwartung *von X gegeben B definiert über*

$$\mathbb{E}[X|B] := \mathbb{E}_{\mathbb{P}(.|B)}[X] = \frac{\mathbb{E}[XI_B]}{\mathbb{P}(B)}.$$

Sei allgemeiner $\mathcal{G}$ eine σ-Algebra mit entsprechender Partition $B_1, \dots, B_k$. Dann ist die bedingte Erwartung von X gegeben $\mathcal{G}$ *definiert durch*

$$\mathbb{E}[X|\mathcal{G}] := \sum_{i=1}^{k} I_{B_i} \mathbb{E}[X|B_i].$$

Die folgende Charaktersierung der bedingten Erwartung wird insbesondere verwendet, um das Konzept auch im Fall von allgemeinen Wahrscheinlichkeitsräumen einzuführen.

Theorem 1.7. *Seien X, Y Zufallsvariablen und $\mathcal{G}$ eine σ-Algebra. Dann ist gilt $Y = \mathbb{E}[X|\mathcal{G}]$ genau dann wenn folgende Eigenschaften erfüllt sind:*

1. *Y ist $\mathcal{G}$-messbar.*
2. *Für alle $B \in \mathcal{G}$ gilt $\mathbb{E}[XI_B] = \mathbb{E}[YI_B]$.*

Beweis. Übung. □

Theorem 1.8 (Eigenschaften der bedingten Erwartung)**.** *Sei X eine Zufallsvariable und $\mathcal{G} \subseteq \mathcal{F}$ eine σ-Algebra.*

1. *$\mathbb{E}[\mathbb{E}[X|\mathcal{G}]] = \mathbb{E}[X]$.*
2. *Die Abbildung $X \mapsto \mathbb{E}[X|\mathcal{G}]$ ist linear.*
3. *Falls $X \geq 0$ dann gilt auch $\mathbb{E}[X|\mathcal{G}] \geq 0$.*
4. *Sei $\mathcal{H} \subseteq \mathcal{G}$ eine weitere σ-Algebra. Dann gilt die* Turmeigenschaft

$$\mathbb{E}[\mathbb{E}[X|\mathcal{G}]|\mathcal{H}] = \mathbb{E}[X|\mathcal{H}].$$

5. *Ist Y $\mathcal{G}$-messbar, dann gilt $Y\mathbb{E}[X|\mathcal{G}] = \mathbb{E}[XY|\mathcal{G}]$.*
6. *Man nennt $\mathcal{G}$ und X* unabhängig *falls für $A \in \mathcal{G}$ und $B \subseteq \mathbb{R}$ gilt, dass $\mathbb{P}(A \cap \{X \in B\}) = \mathbb{P}(A)\mathbb{P}(X \in B)$. In diesem Fall gilt*

$$\mathbb{E}[X|\mathcal{G}] = \mathbb{E}[X].$$

Mithilfe der bedingten Erwartung können wir nun den für die Finanzmathematik zentralen Begriff der bedingten Erwartung definieren.

Definition 1.9. *Sei $X_0, \ldots, X_T$ ein (adaptierter) stochastischer Prozess .*

1. *X ist ein* Martingal *falls für $t = 0, \ldots, T-1$ gilt dass*
$$\mathbb{E}[X_{t+1}|\mathcal{F}_t] = X_t.$$

2. *X ist ein* Submartingal *falls für $t = 0, \ldots, T-1$ gilt dass*
$$\mathbb{E}[X_{t+1}|\mathcal{F}_t] \geq X_t.$$

3. *X ist ein* Supermartingal *falls für $t = 0, \ldots, T-1$ gilt dass*
$$\mathbb{E}[X_{t+1}|\mathcal{F}_t] \leq X_t.$$

Bemerkung 1.10. *Falls X ein Martingal bezüglich einer gegebenen Filtration ist, dann ist es auch ein Martingal bezüglich der X entsprechenden natürlichen Filtration. Wenn wir sagen, dass X ein Martingale ist ohne über die betrachtete Filtration zu reden, dann meinen wir das X ein Martingal bezüglich der natürlichen Filtration ist. Analoges gilt auch für Sub- und Supermartingale.*

Beispiel 1.11.

1. *'Einfache Irrfahrt / random walk': Seien $Y_1, \ldots, Y_T$ unabhängig und gleich verteilt mit $\mathbb{P}(Y_t = +1) = \mathbb{P}(Y_t = -1) = 1/2$. Dann ist durch $X_0 = x_0 \in \mathbb{R}$,*
$$X_{t+1} := X_t + Y_{t+1},$$
für $t = 0, \ldots, T-1$ ein Martingal gegeben.

2. *Sei X eine Irrfahrt wie im letzten Beispiel und definiere Z durch $Z_t := X_t^2 - t$. Dann ist Z auch ein Martingal.*

Im nächsten, einfachen aber tatsächlich sehr wichtigen Satz sehen wir zum ersten Mal wie der Begriff des Martingals im Zusammenhang mit der Finanzmathematik auftritt.

Theorem 1.12. *Sei X ein adaptierter Prozess. Dann sind äquivalent:*

1. *X ist ein Martingal.*

2. *Für alle $H \in \mathcal{H}$ gilt $\mathbb{E}[(H \cdot X)_T] = 0$.*

Beweis. übung. □

Etwas stärker als die Implikation (1) ⇒ (2) im obigen Satz gilt auch:

Theorem 1.13. *Sei X ein Martingal und $H \in \mathcal{H}$. Dann ist der Prozess $((H \cdot X)_t)_t$ auch ein Martingal.*

Wir schließen dieses Kapitel mit einem einfachen Satz der zeigt, dass sich jeder adaptierte stochastische Prozess in ein Martingal und einen vorhersehbaren Prozess zerlegen lässt.

Theorem 1.14 (Doob'scher Zerlegungssatz)**.** *Sei X ein adaptierter Prozess. Dann gibt es genau eine Zerlegung $X = M + A$ mit:*

1. *M ist ein Martingal.*
2. *A ist vorhersehbar.*
3. *$A_0 = 0$.*

Beweis. Falls es so eine Zerlegung gibt, dann gilt jedenfalls

$$\mathbb{E}[X_{t+1} - X_t | F_t] = \mathbb{E}[M_{t+1} - M_t + A_{t+1} - A_t | F_t] = A_{t+1} - A_t. \tag{1.1}$$

Es folgt, dass der vorhersehbare Anteil jedenfalls die Rekursion

$$A_0 = 0, A_{t+1} = A_t + \mathbb{E}[X_{t+1} - X_t | F_t].$$

Diese bestimmt offenbar A und damit auch M eindeutig.

Umgekehrt, zeigt sich, dass wenn man A über diese Rekursion definiert, die im Satz geforderten Eigenschaften automatisch erfüllt sind. □

Zum Doob'schen Zerlegungssatz analoge Aussagen gelten auch für stochastische Prozesse in stetiger Zeit (Satz von Doob-Meyer bzw. Satz von Bichteler-Dellacherie). Diese Sätze sind jedoch deutlich tiefliegender.

Im Zusammenhang mit dem Doob'schen Zerlegung beobachten wir noch: Der Prozess $X = M + A$ ist genau dann ein Submartingal wenn der vorhersehbare Anteil monoton wachsend ist.

2 Arbitrage und Martingalmaße

Wenn wir von einem Modell für den Finanzmarkt sprechen, meinen wir damit einen filtrierten Wahrscheinlichkeitsraum $(\Omega, \mathcal{F}, \mathbb{P}, (\mathcal{F}_t)_{t=0}^T)$ sowie einen adaptierten stochastischen Prozess $X = (X_t)_{t=0}^T$ (der die Entwicklung des Aktienkurses beschreibt). Oft spricht man, etwas verkürzt. nur vom Prozess X. Das sollte aber nicht zu Verwirrungen führen.

Ziel dieses Kapitels ist es, zu verstehen welche Modelle wir prinzipiell für geeignet halten. Zentral ist dabei der Begriff "Arbitrage". Gemeint ist damit im allgemeinen einen Möglichkeit risikolos Gewinn zu machen. Aus ökonomischen Überlegungen, nimmt man an, dass dies in der Realität, und daher auch in unseren Modellen, nicht vorkommen sollte.

Definition 2.1. *Gegeben sei ein Modell $(\Omega, \mathcal{F}, \mathbb{P}, (\mathcal{F}_t)_{t=0}^T, X)$. Eine Handelsstrategie H heißt* Arbitrage *oder* Arbitragestrategie *falls*[1]

$$\mathbb{P}((H \cdot X)_T \geq 0) = 1, \quad \mathbb{P}((H \cdot X)_T > 0) > 0. \tag{2.1}$$

Falls es keinen solchen Strategien gibt, sagen wir, dass der Markt No Arbitrage (NA) *erfüllt.*

1. Angenommen X ist ein Martingal. Dann erfüllt X NA. Um das zu sehen, beachten wir, dass für jedes H, gilt, dass $\mathbb{E}[(H \cdot X)_T] = 0$. Falls $(H \cdot X)_T \geq 0$ ($\mathbb{P}$-a.s.), gilt dann automatisch $(H \cdot X)_T = 0$ ($\mathbb{P}$-a.s.). Somit ist H keine Arbitragestrategie.

2. Seien $\mathbb{P}_1, \mathbb{P}_2$ zwei Wahscheinlichkeismaße auf $(\Omega, \mathcal{F}, \mathbb{P})$. $\mathbb{P}_1$ ist absolut stetig bezüglich $\mathbb{P}_2$ ($\mathbb{P}_1 \ll \mathbb{P}_2$) Falls alle $\mathbb{P}_2$-Nullmengen auch $\mathbb{P}_1$-Nullmengen sind. Haben $\mathbb{P}_1$ und $\mathbb{P}_2$ ($\mathbb{P}_1 \sim \mathbb{P}_2$) die selben Nullmengen, so nennen wir sie *äquivalent.*

   Sei nun $(\Omega, \mathcal{F}, \mathbb{P}, (\mathcal{F}_t)_{t=0}^T, X)$ ein Modell eines Aktienmarktes das NA erfüllt. Falls $\mathbb{Q} \sim \mathbb{P}$, dann erfüllt auch $(\Omega, \mathcal{F}, \mathbb{Q}, (\mathcal{F}_t)_{t=0}^T, X)$ NA.

Aus obigen Beobachtungen folgt:

Proposition 2.2. *Gegeben sei ein Modell $(\Omega, \mathcal{F}, \mathbb{P}, (\mathcal{F}_t)_{t=0}^T, X)$ des Aktienmarktes. Sei $\mathbb{Q} \sim \mathbb{P}$ sodass X ein Martingal bezüglich $\mathbb{Q}$ ist. Dann erfüllt $(\Omega, \mathcal{F}, \mathbb{P}, (\mathcal{F}_t)_{t=0}^T, X)$ (NA).*

Dies motiviert folgende Defintion:

Definition 2.3. *Ein Maß $\mathbb{Q}$ (auf dem stets fix gegeben filtrierten Wahrscheinlickeitsraum) heißt Martingalmaß falls X ein $\mathbb{Q}$-Martingal ist, d.h. falls für $t < T$ stets*

$$\mathbb{E}_{\mathbb{Q}}[X_{t+1} | \mathcal{F}_t] = X_t.$$

Die Menge aller Martingalmaße bezeichnen wir mit

$$M := M(X) := M(\mathbb{P}).$$

Falls zusätzlich $\mathbb{Q} \sim \mathbb{P}$, nennen wir $\mathbb{Q}$ ein äquivalentes Martingalmaß. Die entsprechende Menge bezeichnen wir mit $M^e = M^e(X) = M^e(\mathbb{P})$.

[1] Gemäß unserer Annahme 1.2 liegt also eine Arbitragestrategie vor falls $(H \cdot X)_T \geq 0, (H \cdot X)_T \neq 0$.

Die obige Proposition können wir dann wie folgt formulieren:

Proposition 2.4. *Falls $M^e(X) \neq \emptyset$ so erfüllt X (NA).*

Einer der zentralen Sätze der Finanzmathematik besagt, dass hier auch die Umkehrung gilt. Unter unserer Annahme 1.2 ist der Beweis nicht allzu schwer. Wesentlicher Teil unserer Beweisstrategie wird sein, die bisherigen Begriffe in einem geometrischen Sinn neu zu interpretieren.

Definition 2.5. *Mit $L = L(\mathcal{F})$ bezeichnen wir den Vektorraum aller ($\mathcal{F}$-messbaren) Zufallsvariablen. Mit L^+ bezeichnen wir den positiven Orthanden, d.h. die Teilmenge aller nicht negativen Zufallsvariablen. Weiters sei die Abbildung ϕ definiert durch*

$$\phi : L \to \mathbb{R}^N \tag{2.2}$$

$$X \mapsto (X(\omega_1), \ldots, X(\omega_N)). \tag{2.3}$$

Nach kurzem Überlegen sieht man dann:

1. ϕ ist eine lineare Bijektion, also ein Vektorraumisomorphismus.
2. $K = \{(H \cdot X)_T : H \in \mathcal{H}\}$ ist ein Unterraum von L.
3. X erfüllt (NA) genau dann wenn $K \cap L^+ = \{0\}$.

Analog zum Raum der Zufallsvariablen, wollen wir auch Wahrscheinlichkeitsmaße als Elemente eines linearen Raumes betrachten. Dazu betrachten wir folgende Erweiterung des Maßbegriffes.

Definition 2.6. *Ein* signiertes Maß *auf $\mathcal{F}$ ist eine Abbildung*

$$\sigma : \mathcal{F} \to \mathbb{R}$$

die σ-additiv ist, d.h. für disjunkte Mengen $A_1, A_2, \ldots \in \mathcal{F}$ gilt stets

$$\sum_n \sigma(A_n) = \sigma\left(\bigcup_n A_n\right).$$

Offenbar ist jedes (Wahrscheinlichkeits-) Maß ein signiertes Maß. Umgekehrt ist ein signiertes Maß σ offenbar genau dann ein Wahrscheinlichkeitsmaß wenn es folgende beiden Bedingungen erfüllt:

$$\forall A \in \mathcal{F}, \sigma(A) \geq 0 \tag{2.4}$$

$$\sigma(\Omega) = 1. \tag{2.5}$$

Wir interessieren uns für signierte Maße, da sie (mit punktweise) definierten Operation einen linearen Raum bilden. Da wir unter der Annahme 1.2 arbeiten, ist der Raum der signierten Maße recht übersichtlich. Jedes signierte Maß ist eindeutig durch die Werte

$$\sigma(\{\omega_1\}), \ldots, \sigma(\{\omega_N\})$$

festgelegt. Es folgt, dass die Abbildung ψ,

$$\psi(\sigma) := \sigma(\{\omega_1\}), \ldots, \sigma(\{\omega_N\})$$

eine lineare Bijektion auf $\mathbb{R}^N$ ist.

Gegeben eine Zufallsvariable X und ein signiertes Maß σ, ist das Integral definiert als

$$\int_\Omega X \, d\sigma := \sum_{\omega \in \Omega} X(\{\omega\})\sigma(\{\omega\}) = \langle \phi(X), \psi(\sigma) \rangle.$$

Damit induziert σ eine lineare Abbildung

$$\begin{aligned} \langle \cdot, \sigma \rangle &: L \to \mathbb{R} \\ X \mapsto \langle X, \sigma \rangle &:= \int_\Omega X \, d\sigma. \end{aligned} \tag{2.6}$$

Da $\mathbb{R}^N$ selbstdual ist, folgt dass jedes lineare Funktional auf L von dieser Form ist. D.h. die signierten Maße bilden gerade den Dualraum L^d des Raumes L aller Zufallsvariablen auf Ω.

Im folgenden schreiben wir für $A \subseteq L$ und $B \subseteq L^d$

$$A^\perp := \{\sigma \in L^d : \forall X \in A, \langle X, \sigma \rangle = 0\} \tag{2.7}$$

$$B^\perp := \{X \in L : \forall \sigma \in B, \langle X, \sigma \rangle = 0\}. \tag{2.8}$$

Da L und L^d über die Abbildungen ϕ, ψ mit $\mathbb{R}^N$ identifiziert sind, überträgt sich auch die Topologie (Konvergenz) von $\mathbb{R}^N$ natürlich auf L bzw. L^d: Gegeben $X, X_1, X_2, \ldots \in L$ und $\sigma, \sigma_1, \sigma_2, \ldots \in L^d$ schreiben wir

$$\lim_n X_n = X \quad :\Leftrightarrow \quad \lim_n \phi(X_n) = \phi(X), \tag{2.9}$$

$$\lim_n \sigma_n = \sigma \quad :\Leftrightarrow \quad \lim_n \psi(\sigma_n) = \psi(\sigma). \tag{2.10}$$

Wie wir in den Übungen sehen werden, gilt dann

$$\lim_n X_n = X \quad \Leftrightarrow \quad \lim_n \forall \omega \in \Omega, X_n(\omega) = X(\omega), \tag{2.11}$$

$$\lim_n \sigma_n = \sigma \quad \Leftrightarrow \quad \forall A \in \mathcal{F}, \ \lim_n \sigma_n(A) = \sigma(A). \tag{2.12}$$

Weiters gilt:

Proposition 2.7. *Bezeichne $W \subseteq L^d$ die Menge aller Wahrscheinlichkeitsmaße. Dann ist W konvex und abgeschlossen.*

Beweis. Übung. $\square$

Unsere bekannte Charaktersierung der Martingaleigenschaft über Handelsstrategien können wir nun wie folgt ausdrücken:

Proposition 2.8. *Sei $W \subseteq L^d$ die Menge aller Wahrscheinlichkeitsmaße, M die Menge aller Martingalmaße und $K = \{(H \cdot X)_T : H \in \mathcal{H}\}$ die Menge aller Auszahlungen, die sich durch handeln erzielen lassen. Dann gilt*

$$M = K^\perp \cap W. \tag{2.13}$$

Die wesentliche Zutat um auch die Umkehrung von Proposition 2.2 zu beweisen ist folgende Version des Satzes von Hahn Banach.

Theorem 2.9. *Seien $A, B \subseteq \mathbb{R}^N$ disjunkt, beide konvex, A abgeschlossen und B kompakt. Dann können A und B durch Hyperebenen getrennt werden, d.h. es gibt $v \in \mathbb{R}^N$, $\alpha, \beta \in \mathbb{R}$ sodass für alle $a \in A, b \in B$*

$$\langle a, v\rangle \leq \alpha < \beta \leq \langle b, v\rangle.$$

Ist A ein Unterraum, so gilt $v \in A^\perp$ und können wir $\alpha = 0$ wählen.

Beweis. Wir geben nur einen Beweisskizze. Wegen A abgeschlossen, B kompakt gibt es $a_0 \in A, b_0 \in B$ mit

$$|a_0 - b_0| = \inf\{|a - b| : a \in A, b \in B\} > 0.$$

Wir setzten dann $v := b_0 - a_0$, $\alpha := \langle a_0, v\rangle$, $\beta := \langle b_0, v\rangle$. Dann folgt

$$\beta - \alpha = \langle b_0 - a_0, v\rangle = |a_0 - b_0|^2 > 0.$$

Weiters gilt für $b \in B$ laut Definition von a_0, b_0 und Konvexität von B, dass $\langle b - b_0, v\rangle \geq 0$. Daher ist $\langle b, v\rangle \geq \beta$. Das $\langle a, v\rangle \leq \alpha$ zeigt man genauso.

Sei schließlich A ein Unterraum. Wegen $0 \in A$ gilt dann jedenfalls $\alpha \geq 0$. Falls $\langle a, v\rangle \neq 0$ für ein $a \in A$, dann würde für ein geeignetes Vielfaches von a auch $\langle n \cdot a, v\rangle > \alpha$ gelten. Daher gilt $v \in A^\perp$ und wir können $\alpha = 0$ wählen. $\square$

Theorem 2.10 (FTAP). *Gegeben sei eine Modell X eines Finanzmarktes in diskreter Zeit, d.h. genauer gegeben ist ein filtrierter Wahrscheinlichkeitsraum $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t=0}^T, \mathbb{P})$ und darauf ein adaptierter Prozess X und wir nehmen an, dass Annahme 1.2 erfüllt ist. Dann sind äquivalent:*

1. *Es existiert ein äquivalentes Martingalmaß, d.h. $M^e \neq \emptyset$.*

2. *Das Modell erfüllt (NA).*

Beweis. Wir haben oben schon gesehen, das eine Richtung einfach ist. Wir nehmen nun umgekehrt an, dass (NA) erfüllt ist. Das bedeutet, dass $L^+ \cap K = \{0\}$. Wir betrachten nun die Menge

$$B := \operatorname{conv}\{\mathbb{1}_\omega : \omega \in \Omega\} \subseteq L^+.$$

Dann gilt $K \cap B = \emptyset$, daher können wir Theorem 2.9 auf $\phi(K), \phi(B)$ anwenden und erhalten $v = \psi(\sigma)$ mit

$$\sigma \in K^\perp, \quad \langle b, \sigma \rangle > 0 \text{ für } b \in B.$$

Es folgt, dass σ ein positives Maß ist, tatsächlich gilt $0 < \langle \mathbb{1}_\omega, \sigma \rangle = \sigma(\omega)$ für $\omega \in \Omega$. Daher ist

$$\mathbb{Q} := \frac{\sigma}{\sigma(\Omega)}$$

ein Wahrscheinlichkeitsmaß mit $\mathbb{Q} \sim \mathbb{P}$. Schließlich gilt auch $\mathbb{Q} \in K^\perp$, somit ist $\mathbb{Q}$ ein äquivalentes Wahrscheinlichkeitsmaß, wie gefordert. □

Bemerkung 2.11. *Wir betrachten bis jetzt nur 1-dimensionale Prozesse. Dies hat jedoch nur notationelle Gründe. Tatsächlich gelten die Resultate dieses sowie der folgenden Kapitel auch wenn man einen mehrdimensionalen Prozess* $(X_t)_{t=0}^N$ *mit* $X_t \in R^d$ *betrachtet. Die (wichtige) Interpretation ist, dass es Finanzgüter* $X^1, \dots, X^d$ *gibt, deren Preise wir gemeinsam beobachten wollen.*

3 Arbitragefreie Preise

3.1 Verallgemeinerung von NA und Erreichbarkeit

Grundlegendes Ziel dieser Vorlesung ist es, vernünftige Preise für *Derivate* oder *Optionen* zu bestimmen.

Definition 3.1. *Ein* Derivat *oder eine* Optionen *bezeichnet schlicht eine Zufallsvariable Z.*

In den allermeisten Fällen wird Z von der Form $Z = f(X_0, \ldots, X_T)$ oder $Z = f(X_T)$ sein. Insbesondere interessieren wir uns für *Europäische call-* bzw. *put-Optionen*

$$Z = (X_T - k)_+, \quad Z = (k - X_T)_+.$$

(Hierbei heißt der Zeitpunkt T Maturität und k wird strike oder strike-Preis genannt.)

Ökonomische Interpretation: Zum Zeitpunkt 0 kaufen wir (bzw. ein Akteur am Finanzmarkt) ein Derivat zu einem Preis p um dafür zum Zeitpunkt T vom Verkäufer eine Auszahlung Z zu erhalten.

Gegeben ein arbitragefreies Modell interessieren wir uns welche Preise für ein festes Derivat Z angemessen sind.

Dazu erweitern wir zunächst wie wir am Markt handeln: Wie nehmen wie bisher an, das wir gemäß einer Strategie H Aktien halten. Zusätzlich nehmen wir an, dass wir zum Startzeitpunkt $a \in R$ Einheiten des Derivats Z zum Preis p kaufen um schließlich zum Endzeitpunkt die Auszahlung T zu lukrieren. Für eine entsprechende Strategie $(a, H) \in \mathbb{R} \times \mathcal{H}$ ändert sich unser Vermögen dann um

$$a(Z - p) + (H \cdot X)_T.$$

Definition 3.2. *Gegeben ein Derivat Z, dass zum Preis p gehandelt wird, nennen wir (a, H) Arbitragestrategie falls*

$$a(Z - p) + (H \cdot X)_T \geq 0, \quad a(Z - p) + (H \cdot X)_T \neq 0.$$

Wir nennen p einen arbitragefreien Preis für Z falls keine Arbitragestrategie existiert.

Wenn unser Modell eine Arbitrage zulässt, hat es auch keinen Sinn nach arbitragefreien Preisen für bestimmte Derivate zu fragen. Deswegen nehmen wir im Folgenden stets an, dass wir mit einem arbitragefreien Modell arbeiten.

Für manche Derivate gibt es genau einen arbitragefreien Preis:

Definition 3.3. *Ein Derivat Z heißt erreichbar oder replizierbar falls $p \in \mathbb{R}$ und $H \in \mathcal{H}$ gibt, sodass*

$$Z = p + (H \cdot X)_T.$$

Proposition 3.4. *Angenommen Z ist replizierbar, $Z = p + (H \cdot X)_T$, dann ist p der einzige arbitragefreie Preis von Z.*

Beweis. Wir müssen zeigen:

1. p ist ein arbitragefreier Preis von Z. Sei (a, G) eine erweiterte Strategie. Dann gilt für die entsprechende Vermögensänderung:

$$a(Z - p) + (G \cdot X)_T = a(H \cdot X)_T + (G \cdot X)_T = ((aH + G) \cdot X)_T,$$

das kann keine Arbitrage sein.

2. $q \neq p$ ist kein arbitragefreier Preis. Sei $q < p$ Betrachte die erweiterte Strategie $(a, G) = (1, -H)$. Dann gilt für die entsprechende Änderung im Vermögen:

$$a(Z - q) + (G \cdot X)_T = Z - q - (H \cdot X)_T = (p + (H \cdot X)_T) - q - (H \cdot X)_T = p - q > 0.$$

Im Fall $q > p$ betrachten wir $(a, G) = (-1, H)$...

□

Im Fall $Z = p + (H \cdot X)_T$ können wir nun sagen: Z ist zum Preis p erreichbar. Wir schreiben

$$\pi(Z)$$

für diesen eindeutig bestimmten Preis von Z.

Fragen die wir uns nun natürlich stellen sind: Welche Derivate sind erreichbar? Falls ein Derivat nicht erreichbar ist, wie sieht dann die Menge aller arbitragefreien Preise aus? Gibt es Modelle in denen jedes Derivat erreichbar ist?

Proposition 3.5. *Sei $Z = p + (H \cdot X)_T$. Dann gilt für alle $\mathbb{Q} \in M$*

$$p = \mathbb{E}_{\mathbb{Q}} Z.$$

Beweis. Übung. □

Allgemeiner gilt

Proposition 3.6. *Das Modell X erfülle NA. Sei Z ein Derivat, $\mathbb{Q} \in M^e$. Dann ist*

$$\mathbb{E}_{\mathbb{Q}} Z$$

ein arbitragefreier Preis von Z.

Beweis. Übung. □

Mit den obigen Definitionen ist K gerade die Menge aller Derivate, die man zum Preis 0 erreichen kann. Wir wollen diese Menge nochmals besser verstehen.

Als Gegenstück zu (2.13), d.h. der Gleichung $K^\perp \cap W = M$, liefert das folgende Lemma eine wichtige Charakterisierung der Menge K.

Lemma 3.7. *Unser Modell erfülle NA. Sei Z ein Derivat mit $\mathbb{E}_{\mathbb{Q}}[Z] = 0$ für alle Martingalmaße $\mathbb{Q}$. Dann gilt $Z \in K$, d.h. es gibt eine Handelsstrategie H mit $(H \cdot X)_T = Z$.*

Knapp können wir dies auch wie folgt ausdrücken:

$$M^{\perp} = K. \tag{3.1}$$

Beweis. Wir schreiben $(L^d)^{++}$ für die Menge aller 'strikt positiven' Maße, d.h. Maße die äquivalent zu $\mathbb{P}$ sind.

Wir zeigen zunächst dass

$$\operatorname{span}(K^{\perp} \cap (L^d)^{++}) = K^{\perp}. \tag{3.2}$$

Da $K^{\perp}$ ein Unterraum ist, gilt jedenfalls die Inklusion $\operatorname{span}(K^{\perp} \cap (L^d)^{++}) \subseteq K^{\perp}$.

Um die andere Richtung zu beweisen wählen wir $\sigma \in K^{\perp}$. Sei $\mathbb{Q}$ ein äquivalentes Martingalmaß, $\mathbb{Q} \in K^{\perp}$. Wähle $\alpha \in \mathbb{R}_+$ so dass $\sigma + \alpha\mathbb{Q} \in (L^d)^{++}$. Dann gilt auch $\sigma + \alpha\mathbb{Q} \in K^{\perp} \cap (L^d)^{++}$. Da $\mathbb{Q}$ selbst ein Element von $K^{\perp} \cap (L^d)^{++}$ ist, erhalten wir $\sigma \in \operatorname{span}(K^{\perp} \cap (L^d)^{++})$, und damit (3.2).

Wir nehmen nun, dass $\mathbb{E}_Q[Z] = 0$ für alle $\mathbb{Q} \in M^e$. Dann gilt auch $\mathbb{E}_\sigma[Z] = 0$ für alle $\sigma \in K^{\perp} \cap (L^d)^{++}$ und somit

$$Z \in (K^{\perp} \cap (L^d)^{++})^{\perp} = (K^{\perp})^{\perp} = K.$$

□

Unter der Annahme (NA) liegt M^e dicht in M. Daher gilt

$$K = M^{\perp} = (M^e)^{\perp}. \tag{3.3}$$

Als Folgerung des Satzes erhalten wir nun folgende Charakterisierung erreichbarer Optionen:

Corollary 3.8. *Es gelte NA und $Z \in L$ sei ein Derivat. Dann sind äquivalent:*

1. *Für alle $\mathbb{Q}, \mathbb{Q}' \in M^e$ gilt $\mathbb{E}_{\mathbb{Q}}[Z] = \mathbb{E}_{\mathbb{Q}'}[Z]$.*

2. *Z ist erreichbar,*

3.2 Vollständige Modelle

Definition 3.9. *Ein Modell S das NA erfüllt heißt* vollständig*, wenn jedes Derivat erreichbar ist.*

Wir erhalten nun das sogenannte "Second Fundamental Theorem of Asset Pricing":

Theorem 3.10. *[FTAP2] X erfülle NA. Dann sind äquivalent:*

1. M^e besteht nur aus einem Element.

2. Das Modell ist vollständig.

Viele einfache Modelle sind vollständig. In diskreter Zeit gilt das insbesondere für das Binomial- / Cox-Ross-Rubinstein-Modell. In stetiger Zeit für das Black-Scholes-Modell.

Theorem 3.11. *Für $\mathbb{Q} \in M^e$ sind äquivalent:*

1. $M^e = \{\mathbb{Q}\}$.

2. Für jedes $\mathbb{Q}$-martingal Z gibt es $a \in \mathbb{R}$ und $H \in \mathcal{H}$ sodass

$$Z_t = a + (H \cdot X)_t. \tag{3.4}$$

Beweis. Folgt leicht aus Theorem 3.10. □

Die zweite Eigenschaft in obigem Satz heißt 'predictable representation property'. In stetiger Zeit besagt ein zentraler Satz ('Martingale representation theorem') der stochastischen Analysis, dass das *Wiener Maß* / die *Brown'sche Bewegung* diese Eigenschaft hat.

3.3 Arbitragefreie Preise

Wir nehmen weiterhin an, dass X NA erfüllt. Gegeben $p \in \mathbb{R}$ und ein Derivat Z schreiben wir

$$\mathcal{H}^{(1)} := \{(a, H) : a \in \mathbb{R}, H \in \mathcal{H}\}, \tag{3.5}$$
$$K_{(Z,p)} := \{a(Z - p) + (H \cdot X)_T : (a, H) \in \mathcal{H}^{(1)}\} \tag{3.6}$$

für die Menge der erweiterten Strategien bzw. der induzierten Auszahlungen. Mit diesen Notationen ist p offenbar genau dann ein NA-Preis wenn

$$K_{(Z,p)} \cap L^+ = \{0\}.$$

Wir erhalten nun leicht:

Theorem 3.12. *Das Modell X erfülle NA und Z sei ein Derivat. Dann sind für $p \in \mathbb{R}$ äquivalent:*

1. Es existiert ein äquivalentes Martingalmaß $\mathbb{Q} \in M^e$ mit $p = \mathbb{E}_Q[Z]$, d.h. p ist ein Martingalpreis.

2. p ist NA-Preis von Z.

Beweis. Die einfache Richtung '(1) $\Rightarrow$ (2)' haben wir oben schon besprochen.

Um die andere Richtung '(2) $\Rightarrow$ (1)' gehen wir genau wie im Beweis des FTAP vor. D.h. wir verwenden den Satz von Hahn-Banach um den Unterraum $K_{(Z,p)}$ von der Menge B zu trennen. Durch Normieren erhalten wir dann ein Wahrscheinlichkeitsmaß $\mathbb{Q} \in K_{(Z,p)}^{\perp}, \mathbb{Q} \sim \mathbb{P}$. Dieses ist offenbar ein äquivalentes Martingalmaß und erfüllt überdies $\mathbb{E}_Q[Z - p] = 0$. $\square$

Theorem 3.12 besagt gerade

$$\{p : p \text{ ist ein NA-Preis von } Z\} = \{\ \mathbb{E}_{\mathbb{Q}} Z : \mathbb{Q} \in M^e\}. \tag{3.7}$$

Da M^e eine konvexe Menge ist, bildet die Menge der möglichen Preise von Z ein Intervall.

Als nächstes betrachten wir den sogenannten Superreplikationssatz:

Theorem 3.13. *X erfülle NA und Z sei ein Derivat. Dann gilt*

$$\bar{p} := \sup\{\ \mathbb{E}_{\mathbb{Q}} Z : \mathbb{Q} \in M^e\} = \inf\{p : \exists H \in \mathcal{H}, Z \leq p + (H \cdot X)_T\}. \tag{3.8}$$

Beweis. Es ist leicht zu sehen, dass $\sup\{\ \mathbb{E}_{\mathbb{Q}} Z : \mathbb{Q} \in M^e\} \leq \inf\{p : \exists H \in \mathcal{H}, Z \leq p + (H \cdot X)_T\}$: Dazu nimmt man für die linke und die rechte Seite jeweils einen Kandidaten und zeigt die Ungleichung für diese.

Um zu sehen, dass $\sup\{\ \mathbb{E}_{\mathbb{Q}} Z : \mathbb{Q} \in M^e\} \geq \inf\{p : \exists H \in \mathcal{H}, p \leq (H \cdot X)_T\}$ und dass die rechte Seite angenommen wird, gehen wir indirekt vor: Wir nehmen an, dass es für $p > \bar{p}$ kein passendes H gibt. D.h. für jedes feste H existiert ein ω^* mit

$$Z(\omega^*) > p + (H \cdot X)_T(\omega^*). \tag{3.9}$$

Für $\mathbb{Q} \in M^e$ gilt $\mathbb{E}_{\mathbb{Q}}[Z] \leq \bar{p} < p$. Daher gilt

$$\mathbb{E}_{\mathbb{Q}}[Z - p] < 0 = \mathbb{E}_{\mathbb{Q}}[(H \cdot X)_T].$$

Daher gibt es ein ω_* sodass

$$Z(\omega_*) < p + (H \cdot X)_T(\omega_*). \tag{3.10}$$

Indem wir (3.9) und (3.10) kombinieren, erhalten wir, dass $((-1), H)$ und $(1, (-H))$ beides keine Arbitrage-Strategien sind. Da H beliebig war, folgt das p ein NA-Preis von Z ist, im Widerspruch zur Wahl von p.

Da $p > \bar{p}$ beliebig war, folgt die Aussage. $\square$

Wenn man sich noch ein bisschen mehr Mühe gibt, kann man zeigen, dass in Theorem 3.15 die rechte Seite angenommen wird. Falls Z nicht replizierbar ist, wird die linke Seite nicht angenommen, d.h. das Intervall der möglichen Preise ist dann offen.

3.4 Arbitragetheorie mit Marktdaten

In der Praxis gibt es typischer Weise eine Reihe von Derivaten $Z_i, i \in I$ die zu Preisen $p_i, i \in I$ liquide gehandelt werden; I ist hier eine endliche Indexmenge. Wenn wir am Markt investieren, können wir daher nicht nur in X sondern auch in all diesen Derivaten handeln. Um das zu formalisieren schreiben wir

$$\mathcal{H}^{(I)} := \{(a, H) : a = (a_i)_{i\in I} \in \mathbb{R}^I, H \in \mathcal{H}\}, \tag{3.11}$$

$$K_{(Z_i,p_i)_{i\in I}} := \left\{ \sum_{i\in I} a_i(Z_i - p_i) + (H \cdot X)_T : (a, H) \in \mathcal{H}^{(I)} \right\} \tag{3.12}$$

für die Menge der erweiterten Strategien bzw. der induzierten Auszahlungen. Die hier betrachteten Strategien werden auch *semi-statisch* genannt, weil Derivate nur zu Beginn gehandelt werden können.

Wie zuvor sagen wir, dass unser Modell $(\Omega, \mathcal{F}, \mathbb{P}, (F_t)_{t=1}^T, (X_t)_{t=1}^T, (Z_i, p_i)_{i\in I})$ NA erfüllt, falls wir keinen risikolosen Gewinn machen können, d.h. falls

$$K_{(Z_i,p_i)_{i\in I}} \cap L^+ = \{0\}.$$

Analog zu vorher können wir die NA-Bedingung mit Hilfe von Martingalmaßen charakterisieren, d.h. wir erhalten folgende Version des FTAP:

Theorem 3.14 (FTAP mit Marktdaten)**.** *Gegeben sei ein Modell X eines Finanzmarktes in diskreter Zeit, sowie Derivate $Z_i, i \in I$ die zu Preisen $p_i, i \in I$ gehandelt werden. Wir nehmen an, dass Annahme 1.2 erfüllt ist. Dann sind äquivalent:*

1. *Es existiert ein äquivalentes Martingalmaß $\mathbb{Q}$ mit*

$$\mathbb{E}_\mathbb{Q}[Z_i] = p_i, i \in I.$$

2. *Das Modell erfüllt (NA).*

Der Beweis erfolgt analog zum FTAP bzw. Theorem 3.12.

Gegeben ein weiteres Derivat Z, nennen wir $p \in \mathbb{R}$ einen NA-Preis, falls das um (Z, p) erweiterte Modell wiederum NA erfüllt. Indem wir das FTAP mit Marktdaten auf die Menge $\{(Z_i, p_i)_i \in I\} \cup \{(Z, p)\}$ anwenden, erhalten wir: Die Menge der NA-Preise von Z ist gegeben durch

$$\{\mathbb{E}_\mathbb{Q} Z : \mathbb{Q} \in M^e, \mathbb{E}_\mathbb{Q} Z_i = p_i, i \in I\}.$$

Der Superreplikationssatz lautet nun:

Theorem 3.15. $(\Omega, \mathcal{F}, \mathbb{P}, (F_t)_{t=1}^T, (X_t)_{t=1}^T, (Z_i, p_i)_{i\in I})$ *erfülle NA und Z sei ein Derivat. Dann gilt*

$$\bar{p} := \sup\{ \mathbb{E}_\mathbb{Q} Z : \mathbb{Q} \in M^e, \mathbb{E}_\mathbb{Q} Z_i = p_i, i \in I\} = \inf\{p : \exists Y \in K_{(Z_i,p_i)_{i\in I}}, Z \leq p + Y\}. \tag{3.13}$$

Der Beweis ist analog zu oben. Wiederum wird das Infimum auf der rechten Seite angenommen. Wenn Z nicht erreichbar ist, wird das supremum auf der linken Seite nicht angenommen, d.h. die Menge der möglichen Preise von Z ist offen.

Genau wie oben erhalten wir wiederum, dass ein Derviat Z genau dann mit von $(Z_i, p_i)_{i\in I}$ und handeln im Markt repliziert werden kann, wenn $\{\mathbb{E}_{\mathbb{Q}} Z : \mathbb{Q} \in M^e, \mathbb{E}_{\mathbb{Q}} Z_i = p_i, i \in I\}$ nur aus einem Element besteht.

4 Das Binomialmodell (Cox-Ross-Rubinstein-Modell)

Das Binomialmodell oder CRR-Modell ist ein vollständiges Modell in T Zeitschritten. Es ist im Wesentlichen dadurch charakterisiert, dass die *relativen* Zuwächse in jedem Zeitschritt die selbe Verteilung haben. Tatsächlich wird der relative Zuwachs in jedem Zeitschritt durch einen Faktor beschrieben der entweder $(1+b)$ oder $(1+a)$, mit $-1 < a < b$, ist.

Genauer wählen wir

$$\Omega = \{-1, +1\}^T \tag{4.1}$$

$$Y_t((\omega_1, \dots, \omega_T))) = \omega_t \tag{4.2}$$

$$R_t = \begin{cases} b & \text{falls } Y_t = +1 \\ a & \text{falls } Y_t = -1 \end{cases} \tag{4.3}$$

$$X_t = X_{t-1}(1 + R_t) = X_0 \prod_{k=1}^{t} (1 + R_k), \tag{4.4}$$

wobei jeweils $1 \leq t \leq T$.

Als Filtration wählen wir die kleinste die diese Prozesse adaptiert macht, d.h.

$$\mathcal{F}_t = \sigma(Y_1, \dots, Y_t) = \sigma(R_1, \dots, R_t) = \sigma(X_1, \dots, X_t).$$

$\mathbb{P}$ wählen wir als ein beliebiges Wahrscheinlichkeitsmaß, dass vollen Support hat.

Theorem 4.1. *Das CRR ist arbitragefrei genau dann wenn $a < 0 < b$. In diesem Fall ist es vollständig und das einzige Martingalmaß wird charakterisiert durch*

1. $Y_1, \dots, Y_t$ *unabhängig.*

2. $\mathbb{P}(Y_t = +1) = \frac{-a}{b-a} =: p, \mathbb{P}(Y_t = -1) = \frac{b}{b-a} =: q.$

Bemerkung: Die erste Bedingung ist äquivalent dazu, dass Y_t unabhängig von $\mathcal{F}_{t-1}$ für $t \in \{1, \dots, T\}$.

Im Beweis brauchen wir folgendes Lemma über bedingte Wahrscheinlichkeiten. (Zur Erinnerung: $\mathbb{P}(A|\mathcal{G}) := \mathbb{E}[\mathbb{1}_A|\mathcal{G}]$.)

Lemma 4.2. *Ein Ereignis A ist von der σ-Algebra $\mathcal{G}$ genau dann unabhängig wenn*

$$\mathbb{P}[A|\mathcal{G}] = \mathbb{P}[A].$$

Beweis. Übung. □

Beweis von Satz 4.1. NA gilt genau dann wenn es $\mathbb{Q} \in \mathcal{W}$ gibt, sodass für $t \in \{1, \dots, T\}$.

$$\begin{aligned}
&\mathbb{E}_{\mathbb{Q}}[X_t|\mathcal{F}_{t-1}] = X_{t-1} \\
\iff \quad &\mathbb{E}_{\mathbb{Q}}[X_{t-1}(1+R_t)|\mathcal{F}_{t-1}] = X_{t-1} \\
\iff \quad &\mathbb{E}_{\mathbb{Q}}[R_t|\mathcal{F}_{t-1}] = 0 \\
\iff \quad &b\mathbb{Q}[Y_t = +1|\mathcal{F}_{t-1}] + a\mathbb{Q}[Y_t = -1|\mathcal{F}_{t-1}] = 0 \\
\iff \quad &b\mathbb{Q}[Y_t = +1|\mathcal{F}_{t-1}] + a(1 - \mathbb{Q}[Y_t = +1|\mathcal{F}_{t-1}]) = 0 \\
\iff \quad &(b-a)\mathbb{Q}[Y_t = +1|\mathcal{F}_{t-1}] = -a \\
\iff \quad &\mathbb{Q}[Y_t = +1|\mathcal{F}_{t-1}] = \frac{-a}{b-a} \\
\iff \quad &\mathbb{Q}[Y_t = +1] = \frac{-a}{b-a} \text{ und } Y_t \text{ unabhängig von } \mathcal{F}_{t-1}.
\end{aligned}$$

Wir müssen uns noch überlegen, dass es wirklich genau ein Wahrscheinlichkeitsmaß mit diesen Eigenschaften gibt.

Das ist nicht schwer, es ist gegeben durch

$$\mathbb{Q}(\{(\omega_1, \dots, \omega_T)\}) = p^{|\{k:\omega_k=+1\}|} q^{|\{k:\omega_k=-1\}|}. \tag{4.5}$$

□

Wir betrachten von nun an das Maß $\mathbb{Q}$ aus obigem Satz und fixieren ein Derivat $Z = f(X_1, \dots, X_T)$ unser Ziel ist es, explizit ein $a \in \mathbb{R}$ und $H \in \mathcal{H}$ mit

$$a + (H \cdot X)_T = Z.$$

zu bestimmen. Dazu betrachten wir das Martingal V ('value process'),

$$V_t := \mathbb{E}[Z|\mathcal{F}_t] = \sum_{x_1,\dots,x_t} \mathbb{E}[Z|\{X_1 = x_1, \dots, X_t = x_t\}] \mathbb{1}_{\{X_1=x_1,\dots,X_t=x_t\}}. \tag{4.6}$$

Aus obiger Darstellung erkennen wir, dass es jeweils eine Funktion v_t mit

$$v_t(X_1, \dots, X_t) = V_t$$

gibt. Aus $V_t = \mathbb{E}[V_{t+1}|\mathcal{F}_t]$ erhalten wir die Rekursion

$$v_T = f \tag{4.7}$$

$$v_t(x_1, \dots, x_t) = \mathbb{E}[v_{t+1}(X_1, \dots, X_{t+1})|X_1 = x_1, \dots, X_t = x_t] \tag{4.8}$$

$$= pv_{t+1}(x_1, \dots, x_t, x_t(1+b)) + qv_{t+1}(x_1, \dots, x_t, x_t(1+a)). \tag{4.9}$$

Im nächsten Schritt bestimmen wir H aus der Gleichung

$$H_{t+1}(X_{t+1} - X_t) = V_{t+1} - V_t.$$

Durch Einsetzen ergibt sich auf der Menge $\{X_1 = x_1, \dots, X_t = x_t, X_{t+1} = x_t(1+b)\}$ ebenso wie auf der Menge $\{X_1 = x_1, \dots, X_t = x_t, X_{t+1} = x_t(1+a)\}$

$$\frac{V_{t+1} - V_t}{X_{t+1} - X_t} = \frac{v_{t+1}(x_1, \dots, x_t, x_t(1+b)) - v_{t+1}(x_1, \dots, x_t, x_t(1+a))}{(-a+b)x_t}.$$

Daraus folgt

$$H_{t+1} = h_{t+1}(X_1, \dots, X_t),$$

wobei

$$h_{t+1}(x_1, \dots, x_t) = \frac{v_{t+1}(x_1, \dots, x_t, x_t(1+b)) - v_{t+1}(x_1, \dots, x_t, x_t(1+a))}{(-a+b)x_t}.$$

Wir betrachten noch den wichtigen Spezialfall

$$Z = f(X_t).$$

Induktiv erhalten wir hier, das $V_t = v_t(X)$ und dann auch $H_{t+1} = h_{t+1}(x_t)$ mit

$$v_t(x_t) = p v_{t+1}(x_t(1+b)) + q v_{t+1}(x_t(1+a)) \tag{4.10}$$

$$h_{t+1}(x_t) = \frac{v_{t+1}(x_t(1+b)) - v_{t+1}(x_t(1+a))}{(-a+b)x_t}. \tag{4.11}$$

Aus diesen Gleichungen kann man bei speziellen Derivaten einiges über die Struktur der Replikationsstrategie ablesen. Wir betrachten z.B. den Fall einer europäischen Call-option, d.h.

$$Z = f(X_T) = (X_T - k)_+.$$

Hier ist f wachsend, konvex und 1-Lipschitz. Die Konvexkombination in (4.10) erhält alle diese Eigenschaften. Daher ist für jedes t die Funktion $x \mapsto v_t(x)$ monoton, konvex und 1-Lip. Daraus folgt wiederum leicht, dass die Funktion $x \mapsto h_{t+1}(x)$ monoton ist und nur Werte in $[0, 1]$ annimmt.

5 Nutzenmaximierung

5.1 Grundlegendes Problem und Nutzenfunktionen

Ein typisches Lehrbuchproblem in der Finanzmathematik ist es, den erwarteten Nutzen des Endvermögens zu maximieren. Angenommen, unsere Nutzenfunktion ist durch U gegeben und wir haben ein Anfangsvermögen von X_0, dann lautet das Problem, dem wir gegenüberstehen:

$$\sup E[U(X_0 + (H \cdot X)_T)].$$

Klassische Annahmen sind, dass $U : \mathbb{R} \to [-\infty, \infty)$ streng monoton wachsend, stetig auf $\mathbb{R}$, differenzierbar und streng konkav im Inneren seines Definitionsbereichs $\{x \in \mathbb{R} : U(x) \in \mathbb{R}\}$ ist und dass der Grenznutzen gegen Null strebt, wenn das Vermögen gegen Unendlich geht, also:

$$U'(x) \to 0 \quad \text{für} \quad x \to \infty.$$

Wichtige Beispiele sind:

- $u(x) = \log(x)$, $x > 0$ (negatives Vermögen ist nicht erlaubt),
- $u(x) = \frac{x^\alpha}{\alpha}$, $\alpha \in (0, 1), x > 0$ (negatives Vermögen ist nicht erlaubt),
- $u(x) = -e^{-\gamma x}$, $\gamma > 0, x \in \mathbb{R}$ (negatives Vermögen ist erlaubt).

Obwohl die meisten Theorien für allgemeine Nutzenfunktionen entwickelt werden können, konzentrieren wir uns hier auf $u(x) = -e^{-x}$. In diesem Fall reicht es natürlich aus, das Nutzenmaximierungsproblem für $X_0 = 0$ zu lösen, d. h., wir betrachten

$$\sup_{H \in \mathcal{H}} \mathbb{E}[-\exp(-(H \cdot X)_T)].$$

Wichtiger ist, dass im Fall des exponentiellen Nutzens die Lösung des Nutzenmaximierungsproblems sehr elegant durch das Konzept der relativen Entropie beschrieben werden kann.

5.2 Ein kurzer Exkurs über Entropie und Informationstheorie

Um etwas Kontext zu bieten, führen wir zuerst die klassische Shannon-Entropie ein: Gegeben eine Wahrscheinlichkeitsverteilung $p = (p_1, \dots, p_n)$, ist deren Entropie definiert als

$$H(p) := -\sum_{i=1}^{n} p_i \log_2 p_i,$$

was die Menge der Unsicherheit misst, die enthüllt wird, wenn ein (p_i)-Experiment beobachtet wird.

Wir versuchen, dies etwas konkreter zu machen:

1. Angenommen, $n = 2, p_1 = p_2 = \frac{1}{2}$. Wenn wir das Ergebnis beobachten, wird genau $H((p_i)) = 1$ Bit enthüllt.

2. Betrachten wir nun $k \in \mathbb{N}, p_i = \frac{1}{k}$ für alle i. Wenn k keine Zweierpotenz ist, ist nicht ganz klar, wie viele Bits enthüllt werden. Man kann diese Frage folgendermaßen beantworten: Angenommen, das Experiment wird N-mal für ein großes N wiederholt. Dann wird ein Wort der Länge N beobachtet. Es gibt k^N solche Wörter, und jedes tritt mit gleicher Wahrscheinlichkeit $\frac{1}{k^N}$ auf. Um zu codieren, welches Wort aufgetreten ist, benötigen wir $\lceil N \log_2(k) \rceil \approx H((p_i))N$ Bits. In diesem Sinne enthüllt das Ergebnis eines einzelnen Experiments $\log_2(k) = H((p_i))$ Bits.

3. Sei nun $p_1, \ldots, p_k$ beliebig. Wenn wir das Experiment N mal wiederholen, gibt es k^N mögliche Wörter, aber abhängig von (p_i) sind einige sehr unwahrscheinlich. Bezeichne X_i als das Ergebnis des i-ten Experiments und wende das Gesetz der großen Zahlen auf die i.i.d. Zufallsvariablen $\log_2(p_{X_i})$ an, um zu erhalten

$$-H((p_i)) = \mathbb{E}[\log_2 p_{X_1}] = \lim_{N\to\infty} 1/N \sum_{i=1}^{N} \log_2(p_{X_i}).$$

Anders ausgedrückt, für $\delta > 0$ gilt für N hinreichend groß

$$\mathbb{P}(\{\omega : |H((p_i)) + 1/N \log_2(\text{Wahrscheinlichkeit}(X_1(\omega) \ldots X_N(\omega)))| < \delta\}) > 1 - \delta. \quad (5.1)$$

Ein typisches Wort tritt also mit einer Wahrscheinlichkeit von ungefähr $2^{-NH((p_i))}$ auf es gibt $2^{H((p_i))N}$ solche Wörter.

Um das rigoros auszudrücken, sei $\delta = 0.01$ und N hinreichend groß. Dann folgt aus (5.1): Bezüglich (p_i) decken die $2^{1.01*H((p_i))N}$ häufigsten Wörter 99% aller Wörter der Länge N ab.

Wie zuvor reichen $\approx H(p)N$ Bits um wenigstens diese 'typischen' Wörter zu speichern.

Stärker und schwieriger zu beweisen als das Obige ist der *Shannon-Quellcodierungssatz*: Dies ist ein grundlegendes Resultat in der Informationstheorie, das die Grenzen der effizienten Datenkompression für eine diskrete Gedächtnislose Quelle beschreibt. Gegeben eine Quelle mit einer Wahrscheinlichkeitsverteilung $p_1, p_2, \ldots, p_k$ über k möglichen Symbolen, besagt der Satz, dass es möglich ist, den Ausgang der Quelle mit einer durchschnittlichen Codewortlänge pro Symbol nahe der Entropie $H((p_i))$ der Quelle zu codieren. Grob gesagt, benötigt der optimale Code $-\log_2 p_i$ Bits, um das Symbol i zu codieren, sodass $H((p_i)) = -\sum_{i=1}^{N} p_i \log_{p_i}$ die erwartete Anzahl von Bits ist, die notwendig ist, um ein Symbol zu codieren.

Formal hat der Satz zwei Teile:

1. Erreichbarkeit: Für jedes $\varepsilon > 0$ ist es möglich, einen eindeutig dekodierbaren Code mit einer durchschnittlichen Länge pro Symbol kleiner als $h(X) + \varepsilon$ zu konstruieren. In der Praxis

bedeutet dies, dass wir uns dem Entropielimit für effiziente Codierung annähern, es aber nicht unterschreiten können.

2. Umkehrung: Kein eindeutig dekodierbarer Code kann eine durchschnittliche Länge pro Symbol haben, die kleiner als die Entropie $H(X)$ ist. Dieser Teil des Satzes legt fest, dass $H(X)$ die theoretische Untergrenze der durchschnittlichen Länge pro Symbol für jede verlustfreie Kompressionsmethode ist.

Der Shannon-Quellcodierungssatz garantiert daher, dass die Entropie $H(X)$ die bestmögliche Kompressionsrate für eine Quelle mit der gegebenen Verteilung $p_1, \ldots, p_k$ liefert und eine optimale Codierungseffizienz sicherstellt.

Die *Kreuzentropie* zwischen zwei diskreten Wahrscheinlichkeitsverteilungen $(q_1, \ldots, q_k)$ und $(p_1, \ldots, p_k)$ wird definiert als die erwartete Anzahl von Bits, die benötigt werden, um ein Ereignis der Verteilung (q_i) mit einem Code zu codieren, der für die Verteilung (p_i) optimiert wurde. Mathematisch ausgedrückt ist sie:

$$H((q_i), (p_i)) = -\sum_{i=1}^{k} q_i \log_2 p_i.$$

Die *relative Entropie*

$$H((q_i)|(p_i)) = H((q_i), (p_i)) - H((q_i)) = \sum_{i=1}^{k} q_i \log_2 \frac{q_i}{p_i}.$$

gibt die *überschüssige* Kapazität an, die benötigt wird, wenn eine Quelle (q_i) mit einem Code codiert wird, der für (p_i) optimiert wurde.

Als Nächstes beginnen wir, die relative Entropie ernsthaft zu diskutieren.

5.3 Relative Entropie

Ab jetzt wechseln wir zur Verwendung des natürlichen Logarithmus, da dies für unsere Zwecke bequemer ist.

Definition 5.1. *Gegeben Wahrscheinlichkeiten P, Q, wird die relative Entropie definiert als*

$$H(\mathbb{Q}|\mathbb{P}) = \mathbb{E}_{\mathbb{Q}}\left[\log \frac{d\mathbb{Q}}{d\mathbb{P}}\right] = \mathbb{E}_{\mathbb{P}}\left[\frac{d\mathbb{Q}}{d\mathbb{P}} \log \frac{d\mathbb{Q}}{d\mathbb{P}}\right]$$

falls $\mathbb{Q} \ll \mathbb{P}$, und $H(\mathbb{Q}|\mathbb{P}) = +\infty$ andernfalls.

Da die Funktion $x \mapsto x \log x$ streng konvex ist, folgt (Übung!), dass

$$\mathbb{Q} \mapsto H(\mathbb{Q}|\mathbb{P})$$

streng konvex in $\mathbb{Q}$ ist.

Folglich existiert ein eindeutiger Minimierer von

$$\inf\{H(\mathbb{Q}|\mathbb{P}) : \mathbb{Q} \in M\},$$

das sogenannte entropieminimale Martingalmaß. Später werden wir sehen, dass es uns erlaubt, das Nutzenmaximierungsproblem für exponentiellen Nutzen zu lösen. Dazu benötigen wir einige weitere Vorbereitungen.

Theorem 5.2. *[Dualdarstellung der Entropie] Seien P, Q Wahrscheinlichkeitsmaße. Dann gilt:*

1. $$H(\mathbb{Q}\|\mathbb{P}) = \sup_X \left(\mathbb{E}_\mathbb{P}\left[\frac{d\mathbb{Q}}{d\mathbb{P}}X\right] - \log \mathbb{E}_\mathbb{P}\left[e^X\right]\right)$$

2. $$\log \mathbb{E}\left[e^X\right] = \sup_{\frac{d\mathbb{Q}}{d\mathbb{P}} \in \{prob.\ Dichten\}} \left(\mathbb{E}_\mathbb{P}\left[\frac{d\mathbb{Q}}{d\mathbb{P}}X\right] - H(\mathbb{Q}|\mathbb{P})\right).$$

Beide Supremums werden erreicht, das zweite eindeutig, das erste bis auf Verschiebungen eindeutig. In beiden Fällen wird der Optimierer durch

$$\frac{d\mathbb{Q}}{d\mathbb{P}} = \frac{e^X}{\mathbb{E}_\mathbb{P}\left[e^X\right]} \quad \Longleftrightarrow \quad \exists c,\ X + c = \log \frac{d\mathbb{Q}}{d\mathbb{P}}.$$

charakterisiert.

In Bezug auf Legendre-Transformationen besagt das Theorem, dass die Abbildungen

$$D \mapsto \begin{cases} H(\mathbb{Q}|\mathbb{P}) & \text{falls } D \text{ eine Wahrscheinlichkeitsdichte ist} \\ \infty & \text{sonst} \end{cases}$$

$$X \mapsto \log \mathbb{E}\left[e^X\right]$$

konvexe Konjugierte im Hilbertraum $L^2(P)$ sind. [2]

Beweis. Beide sind konkave Maximierungsprobleme, daher müssen wir nur die erste Bedingung prüfen.

Da wir auf endlichem Ω arbeiten, nehmen wir an, $\Omega = \{1, \ldots, N\}, P(\{i\}) = p_i > 0$, und bezeichnen $X(i) = x_i, Q(\{i\}) = q_i$.

[2] Das konvexe Konjugierte einer Funktion $f : X \to (-\infty, \infty]$ ist definiert als $f^*(x^*) = \sup_x (x^*, x) - f(x)$, was immer eine untere Halbstetige konvexe Funktion ist. Der Satz von Fenchel-Moreau besagt, dass f^{**} der untere halbstetige konvexe Abschluss von f ist. Für hinreichend reguläre konvexe Funktionen f, g auf $\mathbb{R}^N$ mit dem üblichen euklidischen Skalarprodukt ist $f = g^*$ äquivalent zu $\nabla f = (\nabla g)^{-1}$.

Daher lautet 1):

$$H(\mathbb{Q}|\mathbb{P}) = \sup_{x\in\mathbb{R}^N} \underbrace{\sum_{i=1}^N x_i q_i - \log \sum_{i=1}^N e^{x_i} p_i}_{=:F(x).}$$

Die erste Bedingung $\frac{dF}{dx} = 0$ ergibt

$$q_i = p_i \frac{e^{x_i}}{\sum_{j=1}^N e^{x_j} p_j}$$

wie behauptet. Durch Einsetzen erhält man nun das Gewünschte.

Ähnlich ergibt sich 2):

$$\log \mathbb{E}_{\mathbb{P}}\left[e^X\right] = \sup_{q\in\mathbb{R}^N, q_i\geq 0, \sum_i q_i = 1} \underbrace{\sum_{i=1}^N x_i q_i - \sum_{i=1}^N q_i \log \frac{q_i}{p_i}}_{=:G(q)}.$$

Im ersten Schritt führen wir einen Lagrange-Multiplikator λ ein und betrachten $\frac{d(G(q)-\lambda(1-\sum_i q_i)}{dq} = 0$, was zu

$$x_i = \frac{d}{dq_i}\left(q_i \log q_i - q_i \log p_i + \lambda q_i\right) \tag{5.2}$$

$$= \log q_i + 1 + \lambda - \log p_i = \log \frac{q_i}{p_i} + (1-\lambda). \tag{5.3}$$

Daraus folgt, dass $x_i + c = \log \frac{q_i}{p_i}$ für ein $c \in \mathbb{R}$. Wiederum liefert Einsetzen das Gewünschte. □

5.4 Minimierung der relativen Entropie und exponentielle Nutzenmaximierung

Theorem 5.3. *Nehmen wir an, das Modell $(\Omega, \mathcal{F}, P, (\mathcal{F}_t), (X_t))$ erfüllt (NA).*
Sei $\mathbb{Q}^*$ das entropieminimale Martingalmaß. Dann gilt

$$\sup_{Y\in K} \mathbb{E}_{\mathbb{P}}\left[-e^{-Y}\right] = -e^{-H(\mathbb{Q}^*|\mathbb{P})}.$$

Darüber hinaus hat das Nutzenmaximierungsproblem einen eindeutigen Maximierer Y^*; Y^* ist das eindeutig bestimmte Element in $\mathcal{K}$, das die Form $-\log \frac{d\mathbb{Q}^*}{d\mathbb{P}}$ + Konstante hat.

Dieses Ergebnis kann sowohl als Charakterisierung von Y^* in Bezug auf Q^* als auch als Charakterisierung von Q^* in Bezug auf Y^* gelesen werden.

Wir bemerken auch, das Theorem 5.3 einen neuen Beweis für das FTAP liefert.

Beweis von Theorem 5.3. Wir haben

$$
\begin{aligned}
H(\mathbb{Q}^*|\mathbb{P}) &= \min_{\mathbb{Q}\in M} H(\mathbb{Q}|\mathbb{P}) && (5.4)\\
&= \min_{\mathbb{Q}\in M} \sup_{Z\in K} \left(-E_{\mathbb{P}}\left[\frac{d\mathbb{Q}}{d\mathbb{P}}Z\right] + H(\mathbb{Q}|\mathbb{P})\right) && (5.5)\\
&= \sup_{Z\in K} \min_{\mathbb{Q}\in W} \left(-\mathbb{E}_{\mathbb{P}}\left[\frac{d\mathbb{Q}}{d\mathbb{P}}Z\right] + H(\mathbb{Q}|\mathbb{P})\right) && (5.6)\\
&= \sup_{Z\in K} - \max_{\mathbb{Q}\in W} \left(\mathbb{E}_{\mathbb{P}}\left[\frac{d\mathbb{Q}}{d\mathbb{P}}Z\right] - H(\mathbb{Q}|\mathbb{P})\right) && (5.7)\\
&= \sup_{Z\in K} - \log \mathbb{E}_{\mathbb{P}}\left[e^Z\right] && (5.8)\\
&= -\log\left(\inf_{Z\in K} \mathbb{E}_{\mathbb{P}}\left[e^Z\right]\right). && (5.9)\\
&= -\log\left(-\sup_{Z\in K} \mathbb{E}_{\mathbb{P}}\left[-e^Z\right]\right). && (5.10)
\end{aligned}
$$

Hier haben wir den Min-Max-Satz von Sion[3] verwendet, um den Tausch von min und sup zu rechtfertigen. Es bleibt als Übung zu zeigen, dass $\sup_{Z\in K} \mathbb{E}_{\mathbb{P}}\left[-e^Z\right]$ erreicht wird. Der Rest folgt aus der Charakterisierung des Optimierers in der dualen Darstellung von $\log \mathbb{E}_{\mathbb{P}}\left[e^Z\right]$. $\square$

Offenbar könnten wir den obigen Beweis auch für einen anderen Vektorraum als K führen, so lange die NA-Bedingung erfüllt ist. Insbesondere überträgt sich das Resultat auf den Fall in dem mehrere Assets oder auch Derivate gehandelt werden.

In allgemeinen Wahrscheinlichkeitsräumen gilt folgende Version des obigen Theorems, die insbesondere in der Statistik wichtig ist:

Theorem 5.4 (Gibbs-Jaynes principle)**.** *Seien $\phi_1, \ldots, \phi_n : \Omega \to \mathbb{R}$ beschränkte messbare*[4] *Funktionen und setze*

$$\mathcal{Q} = \{\mathbb{Q} \in \Omega : \mathbb{E}_{\mathbb{Q}}[\phi_i] = 0,\, i \leq n\}.$$

Falls es ein $\mathbb{Q} \in \mathcal{Q}$ gibt mit $H(\mathbb{Q}|\mathbb{P}) < \infty$, dann existiert ein eindeutiger Minimierer $\mathbb{Q}^*$ von

$$\inf\{H(\mathbb{Q}|\mathbb{P}) : \mathbb{Q} \in \mathcal{Q}\}.$$

[3]Der Min-Max-Satz von Sion besagt Folgendes: Seien A und B nichtleere konvexe Mengen in topologischen Vektorräumen, A kompakt, und $f : A \times B \to \mathbb{R}$ eine Funktion derart, dass: 1. $f(x, \cdot)$ für jedes feste $x \in A$ konkav und obenhalbstetig in y ist, 2. $f(\cdot, y)$ für jedes feste $y \in B$ konvex und untenhalbstetig in x ist. Dann gilt:

$$\min_{x\in X} \sup_{y\in Y} f(x, y) = \sup_{y\in Y} \min_{x\in X} f(x, y)$$

[4]Etwas allgemeiner reicht $\mathbb{E}_{\mathbb{P}}[\exp(r|\phi_i|)] < \infty$ für alle $r > 0$.

$\mathbb{Q}*$ *ist das einzige Element von $\mathcal{Q}$ mit einer Dichte der Form*

$$\frac{d\mathbb{Q}}{d\mathbb{P}} = a \exp(b_1\phi_1 + \ldots + b_n\phi_n). \tag{5.11}$$

Weiters gilt

$$e^{-H(\mathbb{Q}^*|\mathbb{P})} = \min\{\mathbb{E}_{\mathbb{P}} \exp(b_1\phi_1 + \ldots + b_n\phi_n) : b_1, \ldots, b_n \in \mathbb{R}\}.$$

Optimierer der rechten Seite sind dadurch charakterisiert, dass für ein $a \in \mathbb{R}$, (5.11) *gilt.*

6 Das Black-Scholes Modell

Bei Grenzübergang denken wir daran, dass wir ein Modell betrachten in dem sehr viele, jeweils sehr kleine Änderungen des Aktienkurses erfolgen.

Es gibt verschiedene Arten zu einem Grenzwert des CRR-Modells überzugehen, wir versuchen eine zu finden bei der Grenzübergang möglichst einfach wird. Dazu werden wir die Parameter des CRR-Modells umbenennen (leider mehr als einmal). Es gibt eindeutige Zahlen $\alpha, \beta > 0$, sodass

$$1 + b = e^{\alpha - \beta} \tag{6.1}$$
$$1 + a = e^{-\alpha - \beta}. \tag{6.2}$$

Mit diesen Parametern gilt dann

$$X_n = X_0 e^{\alpha \sum_{k=1}^{n} Y_k - n\beta}.$$

Weiters wollen wir unsere Parameter so wählen, das $\mathbb{Q}(Y_k = +1) = \mathbb{Q}(Y_k = -1) = 1/2$. Das ist gleichbedeutend mit

$$1 = \mathbb{E}[e^{\alpha Y_t - \beta}] = 1/2 \left[e^{\alpha - \beta} + e^{-\alpha - \beta}\right] \tag{6.3}$$
$$\Leftrightarrow e^{\alpha} + e^{-\alpha} = 2e^{\beta}. \tag{6.4}$$

Wir interessieren uns, wie gesagt, vor allem für den Fall, dass alle Sprünge sehr klein sind. Dann gilt natürlich, $\alpha \approx \beta \approx 0$. Wir erhalten dann aus obiger Gleichung

$$(1 + \alpha + 1/2\alpha^2 + \ldots) + (1 - \alpha + 1/2\alpha^2 + \ldots) = 2(1 + \beta + \beta^2/2\ldots) \tag{6.5}$$
$$\Leftrightarrow 1/2\alpha^2 \ldots = \beta + \beta^2/2 \ldots \tag{6.6}$$
$$\Leftrightarrow 1/2\alpha^2 \approx \beta. \tag{6.7}$$

Wir setzen nun $\beta =: \tilde{\alpha}^2/2$, sodass für kleines α ungefähr $\alpha \approx \tilde{\alpha}$ (in dem Sinn, dass $\lim_{\alpha \to 0} \frac{\alpha}{\tilde{\alpha}} = 1$). Dass CRR-Modell mit unserer neuen Parameterisierung wird dann beschrieben durch

$$X_n = X_0 e^{\alpha \sum_{k=1}^{n} Y_k - n\tilde{\alpha}^2/2},$$

wobei $e^{\alpha} - e^{-\alpha} = 2e^{\tilde{\alpha}^2/2}$ und $\mathbb{Q}(Y_k = +1) = \mathbb{Q}(Y_k = -1) = 1/2$. Weiters beobachten wir noch, dass

$$\mathrm{Var}(\log(X_n)) = \mathrm{Var}\left(\alpha \sum_{k=1}^{n} Y_k\right) = n\alpha^2.$$

Um nun eine Art Grenzwert zu bilden, nehmen wir an, dass im N-ten Schritt das Zeitintervall $[0, T]$ in N gleiche Teile geteilt wird und unser Modell die Form

$$X^N_{n/N} := X_0 e^{\alpha_N \sum_{k=1}^{n} Y_k - n\tilde{\alpha}_N^2/2}, \tag{6.8}$$

wobei die Parameter $\alpha_N, \tilde{\alpha}_N$ nun von N abhängen dürfen.

Für $t \in \{n/N : n \leq NT\}$ erhalten wir dann

$$X_t^N := X_0 e^{\alpha_N \sum_{k=1}^{tN} Y_k - tN\tilde{\alpha}_N^2/2}.$$

Aus der Überlegung oben folgt, dass

$$\mathrm{Var}(\log X_t^N) = N\alpha_N^2 t.$$

Da wir die Varianz gerne unabhängig vom Diskretisierungschritt hätten, setzen wir $\alpha_N := \sigma/\sqrt{N}$, $\tilde{\alpha}_N =: \tilde{\sigma}_N/\sqrt{N}$, wobei $\sigma > 0$ konstant ist. Dann hängt zwar $\tilde{\sigma}_N$ von N ab, aber wenigstens wissen wir, dass $\lim_{N\to\infty} \tilde{\sigma}_N = \sigma$.

Mit unseren neuen Parametern erhalten wir

$$\mathrm{Var}(\log X_t^N) = \sigma^2 t.$$

und

$$X_t^N = X_0 e^{\sigma \sum_{k=1}^{tN} Y_k/\sqrt{N} - t\tilde{\sigma}_N^2/2} = X_0 e^{\sigma B_t^N - t\tilde{\sigma}_N^2/2},$$

wobei der Prozess B^N durch

$$B_t^N := \sum_{k=1}^{tN} Y_k/\sqrt{N}$$

gegeben ist. Wir bemerken nun, dass der Prozess B^N die folgenden Eigenschaften hat:

1. $\mathbb{E}[B_t^N] = 0$
2. $\mathbb{E}[(B^N)_t^2] = t.$
3. Für $0 \leq t_1 \leq t_2 \leq \ldots \leq t_m \leq T$ sind die Inkremente $B_{t_2}^N - B_{t_1}^N, \ldots B_{t_m}^N - B_{t_{m-1}}^N$ unabhängig.

Bis jetzt nehmen wir stets implizit an, dass $t \in \{n/N : n \leq NT\}$, d.h. B^N ist weiterhin ein Prozess in diskreter Zeit. Wir erweitern B^N nun zu einem Prozess auf dem ganzen Intervall $[0, T]$, in dem wir

$$B_t^N := B_{\lfloor tN \rfloor/N}^N$$

setzen. Entsprechend setzten wir auch

$$X_t^N = X_0 e^{\sigma B_t^N - t\sigma_N^2/2}$$

zu einem Prozeß auf $[0, T]$ fort.

Offenbar vermuten wir, dass B_t^N für $N \to \infty$ in einem bestimmten Sinn konvergiert. Wir erinnern uns dafür zunächst an den zentralen Grenzwertsatz:

Theorem 6.1. *Seien $Z_1, Z_2, \dots$ unabhängig und gleich verteilt, $\mathbb{E}Z_i = 0$, $Var(Z_i) = \sigma^2 < \infty$. Sei weiters $Z \sim N(0, \sigma^2)$. Dann gilt*

$$\lim_{n\to\infty} \frac{1}{\sqrt{n}} \sum_{k=1}^{n} Z_k = Z \quad \textit{in Verteilung.}$$

Konkret bedeutet das, dass die Verteilungsfunktionen von $\frac{1}{\sqrt{n}} \sum_{k=1}^{n} Z_k$ punktweise gegen die Verteilungsfunktion von $N(0, \sigma^2)$ konvergieren.

Dieser Satz liefert uns, dass für jedes $t \in [0, T]$ die Folge B_t^N gegen eine $N(0, T)$-verteilte Zufallsvariable konvergiert. Tatsächlich wollen wir aber noch ein Stück mehr: wir hätten gerne, dass die gesamten Prozesse B^N wiederum gegen einen Prozeß konvergieren.

Das wird der sogenannte Satz von Donsker leisten. Zuerst brauchen wir die Definition der Brown'schen Bewegung.

Definition 6.2. *Sei $B = (B_t)_{t\in[0,T]}$ ein stochastischer Prozess mit den Eigenschaften*

1. $\mathbb{E}[B_t] = 0$
2. $\mathbb{E}[B_t^2] = t$.
3. *Für $0 \le t_1 \le t_2 \le \dots \le t_m \le T$ sind die Inkremente $B_{t_2} - B_{t_1}, \dots B_{t_m} - B_{t_{m-1}}$ unabhängig und normalverteilt.*
4. *Für fast alle ω ist die Abbildung*
$$t \mapsto B_t(\omega)$$
stetig.

Dann heißt B Brown'sche Bewegung*.*

Theorem 6.3 (Satz von Donsker)**.** *Für $N \to \infty$ konvergiert B^N in Verteilung gegen die Brown'sche Bewegung B.*

Wir werden an dieser Stelle nicht genau ausführen, was schwache Konvergenz für stochastische Prozesse bedeutet, wir begnügen uns damit festzuhalten, dass für all $0 \le t_1 \le \dots \le t_n \le T$

$$\lim_{N\to\infty} (B_{t_1}^N, \dots, B_{t_n}^N) = (B_{t_1}, \dots, B_{t_n})$$

in Verteilung.

Da wir schon wissen, dass $\tilde{\sigma}_N \to \sigma$, erhalten wir auch, dass die Prozesse X^N gegen den Prozess X,

$$X_t = X_0 e^{\sigma B_t - t\sigma^2/2} \tag{6.9}$$

konvergieren.

Definition 6.4. *Der Prozess X in* (6.9) *heisst* geometrische Brown'sche Bewegung *oder* Black-Scholes Modell*.*

Da X ein Grenzwert von Martingalen ist, können wir hoffen, dass X selbst wieder ein Martingal ist. Dem ist tatsächlich so: mit $\mathcal{F}_t := \sigma(X_s : s \leq t)$ gilt für $0 \leq u \leq t \leq T$

$$\mathbb{E}[X_t|\mathcal{F}_u] = X_u. \tag{6.10}$$

Als nächstes untersuchen wir, was wir für den Wert v_0 einer Europäischen Option

$$Z = f(X_T)$$

erwarten würden. Da $X_T^N \to X_t$ sollte (für hinreichend beschränkte) f gelten, dass

$$v_0 = \mathbb{E}f(X_T) = \lim_{N\to\infty} \mathbb{E}f(X_T^N).$$

Da X_T log-normalverteilt ist, gilt konkret

$$v_0 = \mathbb{E}f\left(X_0 e^{\sigma\sqrt{T}Z - T\sigma^2/2}\right)$$

wobei $Z \sim N(0,1)$, oder

$$v_0 = \int_z f\left(X_0 e^{\sigma\sqrt{T}z - T\sigma^2/2}\right) \frac{1}{\sqrt{2\pi}} e^{-z^2/2}\, dz. \tag{6.11}$$

(6.11) ist eine Version der *Black-Scholes Formel.*

Unter der Bedingung, dass der Aktienkurs zum Zeitpunkt t den Wert x_t hat, erwarten wir für den Wert von $f(X_T)$

$$v_t(x_t) = \mathbb{E}f\left(x_t e^{\sigma\sqrt{T-t}Z-(T-t)\sigma^2/2}\right) = \int_z f\left(x_t e^{\sigma\sqrt{T-t}z-(T-t)\sigma^2/2}\right) \frac{1}{\sqrt{2\pi}} e^{-z^2/2}\, dz. \tag{6.12}$$

Wenn wir $f(X_T)$ replizieren wollen, dann sollten wir für $X_t = x_t$ stets

$$h_t(x_t) = \frac{d}{dx}(v)(x_t) = \int_z f'\left(x_t e^{\sigma\sqrt{T-t}z-(T-t)\sigma^2/2}\right) \frac{e^{\sigma\sqrt{T-t}z-(T-t)\sigma^2/2}}{\sqrt{2\pi}} e^{-z^2/2}\, dz \tag{6.13}$$

Aktien halten.

Dass ein Portfolio mit obiger Handelsstrategie das gegebene Derivat repliziert, erscheint (in Anbetracht der Approximation oben) plausibel, im Moment können wir das allerdings nicht beweisen. Eigentlich können wir noch nicht einmal genau sagen, was wir in stetiger Zeit mit Handel meinen.

6.1 Handeln in stetiger Zeit - Stochastisches Integral

Ein adaptierter Prozess $H = (H_t)_{0 \le t \le T}$ in stetiger Zeit heißt *simpel* falls

$$H = \sum_{i=0}^{N-1} \hat{H}_i \mathbb{1}_{(t_i, t_{i+1}]},$$

für $\hat{H}_i \in \mathcal{F}_{t_i}$, und $0 = t_0 < t_1 < \ldots < t_N = T$. Gegeben so einen Prozess und einen adaptierten Prozess X schreiben wir

$$(H \cdot X)_T := \int_0^T H_u \, dX_u = \sum_{i=1}^{N} \hat{H}_i (X_{t_{i+1}} - X_{t_i}) = \sum_{i=1}^{N} H_{t_{i+1}} (X_{t_{i+1}} - X_{t_i}).$$

Die endliche Folge $\{t_0 < t_1 < \ldots < t_N = T\} =: \pi$ bezeichnen wir als zugehörige Partition. $|\pi| := \max_{t_i \in \pi} |t_i - t_{i+1}|$ bezeichnet die Gitterweite oder Norm der Partition.

Für die mathematische Theorie ist es deutlich praktischer, den Prozess $(H \cdot X)_t$ noch allgemeiner zu definieren. Wir beschränken uns der Einfachheit halber auf stetige Prozesse H. Gegeben eine Partition π, definieren wir einen simplen Prozess der H approximiert durch

$$H^\pi := \sum_{t_i \in \pi} H_{t_i} \mathbb{1}_{(t_i, t_{i+1}]}.$$

Wir sagen, dass das stochastische Integral $(H \cdot X)_T = \int_0^T H \, dX$ von H bezüglich X existiert, falls der Grenzwert

$$(H \cdot X)_T := \lim_{n \to \infty} (H^{\pi_n} \cdot X)_T = \lim_{n \to \infty} \sum_{t_i \in \pi_n} H_{t_i} (X_{t_{i+1}} - X_{t_i}) \tag{6.14}$$

für jede Folge von Partitionen $\pi_n, |\pi_n| \to 0$ im Sinne der Konvergenz in Wahrscheinlichkeit existiert.

Etwas allgemeiner definieren wir für $t \in [0, T]$ und H simpel

$$(H \cdot X)_t := \sum_{i=1}^{N} \hat{H}_i (X_{t_{i+1} \wedge t} - X_{t_i \wedge t})$$

und dann auch für allgemeines H durch $(H \cdot X)_t := \lim_{n \to \infty} (H^{\pi_n} \cdot X)_t$.

Analog zum entsprechenden Satz in diskreter Zeit gilt:

Theorem 6.5. *Sei H simpel oder H adaptiert stetig und beschränkt und sei X ein stetiges Martingal. Dann existiert $(H \cdot X)_t$ und ist ein Martingal.*

Wie jedes vernünftige Integral ist auch das stochastische Integral linear, wir werden das in den Übungen besprechen.

Wir würden nun gerne zeigen, das für die oben definierten Prozesse gilt, dass

$$f(X_T) = v(X_0) + (H \cdot X)_T = v(X_0) + \int_0^T h(X_t)\, dX_t.$$

Der wichtigste Baustein, der uns dafür fehlt ist die sogenannte Ito-Formel.

Theorem 6.6 (Ito-Formel, 1. Version)**.** *Sei B eine Brown'sche Bewegung und $f : \mathbb{R} \to \mathbb{R}$ zwei mal stetig differenzierbar. Dann gilt*

$$f(B_t) - f(B_0) = \int_0^t 1\, df(B_t) = \int_0^t f'(B_s)\, dB_s + \frac{1}{2}\int_0^t f''(B_s)\, ds, \tag{6.15}$$

bzw. in Kurzform

$$df(B_t) = f'(B_t)\, dB_t + \frac{1}{2} f''(B_t)\, dt. \tag{6.16}$$

Diese Form stellt eine Art Kettenregel dar und ist als solche ziemlich bemerkenswert. Zum Vergleich: Sind Funktionen f, g stetig differenzierbar, so gilt

$$df(g(t)) = f'(g(t))g'(t)\, dt = f'(g(t))\, dg(t). \tag{6.17}$$

Beweis. Wir zeigen nur eine grobe Skizze. Die wichtigste Heuristik ist, dass ungefähr gilt, dass

$$dB_t = \pm\sqrt{dt},$$

dass hatten wir auch in unserer Approximation der Brown'schen Bewegung durch den random walk. Wenn man daran glaubt, gilt ungefähr

$$dB_t^2 = dt. \tag{6.18}$$

Um $df(B_t)$ abzuschätzen, verwenden wir dann die Taylorformel und erhalten

$$\begin{aligned} df(B_t) &= f(B_{t+dt}) - f(B_t) = f((B_t) + dB_t) - f(B_t) && (6.19)\\ &= f'(B_t)dB_t + \frac{1}{2}f''(B_t)dB_t^2 + \frac{1}{6}f'''(B_t)dB_t^3 + \ldots && (6.20)\\ &= f'(B_t)dB_t + \frac{1}{2}f''(B_t)dB_t^2 = f'(B_t)\, dB_t + \frac{1}{2}f''(B_t)\, dt. && (6.21) \end{aligned}$$

□

Theorem 6.7 (Ito-Formel, 2. Version)**.** *Sei B eine Brown'sche Bewegung und $f : \mathbb{R}_+ \times \mathbb{R} :\to \mathbb{R}$ zwei mal stetig differenzierbar. Dann gilt*

$$f(t, B_t) - f(0, B_0) = \int_0^t 1\, df(t, B_t) = \int_0^t f'(s, B_s)\, dB_s + \int_0^t \frac{1}{2} f''(s, B_s) + \frac{d}{ds} f(s, B_s)\, ds, \tag{6.22}$$

bzw. in Kurzform

$$df(t, B_t) = df(t, B_t) = f'(t, B_t)\, dB_t + \frac{1}{2} f''(t, B_t)\, dt + \frac{d}{dt} f(t, B_t)\, dt, \tag{6.23}$$

Beispiele zur Ito-Formel:

1. $d(B_t^2) = 2B_t + dt$ bzw. $\int B_t\, dB_t = B_t^2/2 - t/2$.

2. Bezeichne X die geometrische Brown'sche Bewegung bzw. das Black-Scholes-Modell. Dann gilt

$$dX_t = \sigma(X_t)\, dB_t, \tag{6.24}$$

bzw.

$$X_t - X_0 = \int_0^t \sigma(X_u)\, dB_u. \tag{6.25}$$

Allgemeiner lassen sich stochastische Integrale für sogenannte Semi-martingale definieren, dass sind Prozesse die als Summe eines Martingals und eines Prozesses mit beschränkter Totalvariation darstellbar sind. Man kann zeigen dass jede zweimal stetig differenzierbare Funktion eines Semi-martingals wiederum ein Semi-martingal ist und auch in diesem Fall gilt eine Version der Ito-Formel. Wir betrachten noch einen weiteren Spezialfall einer entsprechenden Ito-formel.

Theorem 6.8 (Ito-Formel, 3. Version)**.** *Seien X, Y stetige Semi-martingale und $f : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ zwei mal stetig differenzierbar. Dann gilt*

$$df(X_t, Y_t) = f_x\, dX_t + f_y\, dY_t + \frac{1}{2} f_{xx}\, (dX_t)^2 + f_{xy}\, dX_t dY_t + \frac{1}{2} f_{yy}\, (dY_t)^2. \tag{6.26}$$

Zum Beispiel folgt daraus folgende Produktregel für stetige Semimartingale:

$$d(X_t Y_t) = X_t\, dY_t + Y_t\, dX_t + dX_t dY_t. \tag{6.27}$$

Um die beiden letzten Formeln praktisch anwenden zu können, muß man noch wissen, wie Produkte der Form $dX_t dY_t$ in konkreten Fällen auszuwerten sind. Gegeben eine Brown'sche Bewegung B gilt

$$dB_t^2 = dt, dB_t dt = dt dB_t = 0.$$

Ist $\tilde{B}$ eine von B unabhängige Brown'sche Bewegung so gilt

$$dB_t d\tilde{B}_t = 0.$$

Zum Beispiel gilt dann für Prozesse $dX_t^i = \sigma_t^i \, dB_t + \mu_t^i \, dt$, $i = 1, 2$

$$dX_t^1 dX_t^2 = \sigma_t^1 \sigma_t^2 \, dt.$$

Schließlich brauchen wir noch eine (letzte) weitere Rechenregel für stochastische Integrale.

Proposition 6.9. *Seien H, K, X stetige Prozesse. Dann gilt*

$$\int_0^t H_s \, d\left(\int_0^s K_u \, dX_u\right) = \int_0^t H_s K_s \, dX_u \quad \textit{bzw.} \tag{6.28}$$

$$(H \cdot (K \cdot X))_T = ((HK) \cdot X)_T, \tag{6.29}$$

sobald die linke oder die rechte Seite definiert ist (und dann ist auch die andere Seite definiert).

Beweis. Eine Möglichkeit, das zu beweisen, ist es für den Fall von simplen H, K die Gleichung nachzurechnen und dann zum Grenzwert überzugehen. Jedenfalls illustrativ ist es den diskreten Fall anzusehen.

Man kann sich die Gleichung auch über eine Finanzmathematische Interpretation plausibel machen: Bezeichne K die Anzahl der Aktien die ein Aktienfonds kauft (und nehmen wir an, dass der Fonds sonst nichts tut). Dann sind dessen Gewinne / Verluste des Fonds durch $(K \cdot X)$ beschrieben und dies gibt auch den Wert des funds zum jeweiligen Zeitpunkt wieder. Nehmen wir nun an wir handeln selbst Anteile des Funds und wir beschreiben diese Strategie durch H. Dann sind unsere Gewinne / Verluste durch $(H \cdot (K \cdot X))$ beschrieben. Natürlich läuft das auf das selbe hinaus wenn wir gleich HK Aktien kaufen und demnach $(HK \cdot X)$ Gewinn / Verlust haben. □

Mit Hilfe dieser Vorbereitungen folgt nun:

Theorem 6.10. *Sei X eine geometrische Browon'sche Bewegung und sei f zweimal stetig differenzierbar. Dann gilt mit den Formeln in* (6.12) *und* (6.13)

$$v_0(X_0) + \int_0^T h_t(X_t) \, dX_t = f(X_T), \quad v_t(X_t) = \mathbb{E}[f(X_T)|\mathcal{F}_t]. \tag{6.30}$$

Beweis. Ito-Formel auf $f(X_t)$ anwenden... □

6.2 Brown'sche Bewegung mit drift

Das Black-Scholes-Modell mit drift hat die Form

$$dX_t = X_t \mu \, dt + X_t \sigma \, dB_t. \tag{6.31}$$

bzw. äquivalent dazu

$$X_t = X_0 \exp\left(\sigma B_t - \frac{1}{2}\sigma^2 t + \mu t\right). \tag{6.32}$$

Theorem 6.11. *Es gibt genau ein äquivalentes Martingalmaß und unter diesem ist $W_t := B_t - \frac{\mu}{\sigma} t$ eine Brown'sche Bewegung und X erfüllt die Differentialgleichung*

$$dX_t = X_t \sigma \, dW_t. \tag{6.33}$$

6.3 Implizite Volatilität

Die Black-Scholes Formel führt für den Preis $p_{K,T}$ einer call-Option $(X_T - K)_+$ zu einem funktionalen Zusammenhang

$$p_{K,T} = \mathrm{BS}(\sigma, X_0, T, K). \tag{6.34}$$

In der Praxis werden call-Optionen liquide am Markt gehandelt, der Preis wird vom Markt bestimmt. Da X_0, T, K offenbar fest vorgegeben sind, kann man (6.34) verwenden um $\sigma = \sigma_{T,K}$ in Abhängigkeit von T und K zu bestimmen.

Würde der Markt einem Black-Scholes-Modell folgen, dann wäre σ schlicht konstant (und würde nicht von K, T abhängen). In der Praxis ist dies nicht erfüllt und wir können folgern, dass der Finanzmarkt tatsächlich wohl einem komplexeren Modell folgt.

In den folgenden Abschnitten werden wir zwei solche Modelle kurz vorstellen.

6.4 local volatility model

$$dX_t = X_t \sigma(t, X_t) \, dB_t.$$

6.5 Heston model

$$dX_t = X_t \sqrt{v_t} \, dB_t \tag{6.35}$$

$$dv_t = \kappa(\theta - v_t) \, dt + \xi \sqrt{v_t} \, dW_t, \tag{6.36}$$

where κ, θ, ξ are real non-negative parameters. (Sidenote: if $\theta = 0, \kappa < 0$ we obtain a reasonable model for the evolution of infectious diseases.)

Der Prozeß in (6.36) heißt Cox-Ingersoll-Ross-Prozeß oder Feller-square-root-diffusion...

6.6 Diskontieren

7 Weitere Themen

7.1 Risikomaße

7.2 Sharpe ratio

7.3 Amerikanische Optionen

8 Fragen auf die man vorbereitet sein sollte

1. Was ist eine Option, was ist ihr ökonomischer Nutzen?
2. Was ist ein Modell?
3. Grundbegriffe der Maßtheorie: Maß, σ-Algebra,
4. Wie ist der bedingte Erwartungswert definiert? Eigenschaften der bedingten Erwartung.
5. Was ist ein Martingal?
6. Doob'scher Zerlegungssatz, Beweisidee.
7. Was ist eine Handelsstrategie?
8. Wie ist $(H \cdot X)$ definiert, was ist die ökonomische Interpretation.
9. Was ist Arbitrage? (mathematisch)
10. Geometrische Interpretation von Zufallsvariablen, Maßen, NA-Bedingung
11. Trennungssatz von Hahn-Banach mit Beweisidee.
12. Was sagt FTAP 1? Beweis der einfachen Richtung, Beweisidee der schwierigen Richtung.
13. Was sagt FTAP 2? grobe Struktur des Beweises.
14. Angenommen es es gibt zwei verschiedene Martingalemaße. Warum kann das Modell dann nicht vollständig sein?
15. Geben sie ein Beispiel für ein einfaches Modell das nicht vollständig ist und ein Beispiel für eine Option die in diesem Modell replizierbar ist (o.ä.).
16. Wann nennen wir einen Preis arbitragefrei?
17. Was bedeutet, dass ein Derivat replizierbar ist?
18. Wie läßt sich die Menge der arbitragefreien Preise mithilfe der Martingalmaße beschreiben?
19. Definition des Binomialmodells.
20. Für welche Parameter ist es arbitragefrei / vollständig? Beweisidee.
21. Gegeben ein Derivat $f(X_T)$ wie findet man die Replikationsstrategie im Binomialmodell? (Strategie.)

22. Gegeben eine Call-Option, wie hängen v_t, h_t vom momentanen Aktienkurs ab? (monoton wachsend, ...).

23. Übergang von CRR zu Black-Scholes (Strategie)?

24. Zentraler Grenzwertsatz.

25. Definition Brown'sche Bewegung und einfache Eigenschaften.

26. Black-Scholes Modell, Definition.

27. Definition stochastisches Integral, für einfache Funktionen / stetige Funktionen. Finanzmathematische Interpretation.

28. Martingaleigenschaft und stochastisches Integral ...

29. Itoformel, wozu brauchen wir sie, Beweisidee, eine Anwendung?

30. Was ist implizite Volatilität? Inwiefern widersprechen Marktdaten dem Black-Scholes Modell?

31. Kennen Sie Alternativen zum Black-Scholes Modell, was ist das 'local-volatility model'?

Literatur

[1] F. Delbaen and W. Schachermayer. *The mathematics of arbitrage.* Springer Finance. Springer-Verlag, Berlin, 2006.

[2] H. Föllmer and A. Schied. *Stochastic finance.* De Gruyter Graduate. De Gruyter, Berlin, 2016. An introduction in discrete time, Fourth revised and extended edition of [MR1925197].