



**TECHNISCHE
UNIVERSITÄT
DRESDEN**

Fakultät Mathematik Institut für Stochastik, Professur für Stoch. Analysis und Finanzmathematik

VERTIEFUNG IN DER STOCHASTIK

Prof. Dr. Martin Keller-Ressel

Wintersemester 2019/20

Autor : Eric Kunze
E-Mail : eric.kunze@mailbox.tu-dresden.de

Inhaltsverzeichnis

1	Einführung	2
1.1	Zentrale Fragestellungen der Finanzmathematik	2
1.2	Mathematisches Finanzmodell	3
1.3	Anleihen und grundlegende Beispiele für Derivate	5
1.4	Elementare Replikations- und Arbitrageargumente	7
1.5	Bedingte Erwartungswerte und Martingale	8
2	Cox-Ross-Rubinstein-Modell	16
2.1	Anlagestrategien im CRR-Modell	17
2.2	Replikation bzw. Hedging von Derivaten im CRR-Modell	18
2.3	Martingalmaß und Arbitrage im CRR-Modell	22
3	Das Black-Scholes-Modell	25
3.1	Asymptotik von Put- und Call-Preisen	28
3.2	Implizite Volatilität & Grenzen des BS-Modells	32
4	Optimale Investition	34
4.1	Das Anlageproblem	34
4.2	Exkurs: Optimierung mit Nebenbedingungen	35
4.3	Die Markowitz-Modelle	38
4.4	Capital Asset Pricing Model (CAPM)	41
4.5	Präferenzordnungen und Erwartungsnutzen	43
5	Risikomaße	48
5.1	Value-at-Risk	48
5.2	Expected Shortfall	50
5.3	Risikomaße: Axiomatischer Zugang	52

Kapitel 1

EINFÜHRUNG

1.1 Zentrale Fragestellungen der Finanzmathematik

1.1.1 Bewertung von Derivaten und Absicherung gegen aus deren Kauf/Verkauf entstehende Risiken

Definition 1.1 (Derivat)

Ein **Derivat** ist ein Finanzprodukt, dessen Auszahlung sich vom Preis eines oder mehrerer Basisgüter [underlying] ableitet.

Beispiel 1.2

- Recht in drei Monaten 100.000 GBP gegen 125.000 EUR zu erhalten (Call-Option; underlying: Wechselkurs GBP in EUR)
- Recht innerhalb des nächsten Jahres 100.000 MWh elektrische Energie zum Preis von 30 EUR/MWh zu konsumieren mit Mindestabnahme 50.000 MWh (Swing-Option; underlying: Strompreis)
- Kauf- und Verkaufsoptionen auf Aktien (underlying: Aktienkurs)

Fragestellungen:

- Was ist der "faire" Preis für solch ein Derivat? ("Pricing" / Bewertung)
- Wie kann sich der Verkäufer gegen die eingegangenen Risiken absichern? ("Hedging" / Absicherung)

1.1.2 Optimale Investition: Zusammenstellen von nach Risiko-/Ertragsgesichtspunkten optimalen Portfolios

- Wie wäge ich Risiko gegen Ertrag ab?
- Was bedeutet optimal?
- Lösung des resultierenden Optimierungsproblems

1.1.3 Risikomanagement und Risikomessung

gesetzliche Vorschriften (Basel und Solvency) sollen Stabilität des Banken- und Versicherungssystems angesichts verschiedener Risiken sicherstellen

→ mathematische Theorie der konvexen und kohärenten Risikomaße

Mathematische Werkzeuge: Wahrscheinlichkeitstheorie und stochastische Prozesse (Dynamik in der Zeit), zusätzlich etwas lineare Algebra, Optimierung, Maßtheorie

1.2 Mathematisches Finanzmodell

Wir betrachten

- (1) einen Wahrscheinlichkeitsraum $(\Omega, \mathcal{F}, \mathbb{P})$, später auch weitere Maße \mathbb{Q}, \dots auf demselben Maßraum (Ω, \mathcal{F}) . Die $\omega \in \Omega$ werden als **Elementarereignisse** oder **”Szenarien“** bezeichnet.
- (2) Zeitachse I entweder $I = \{t_1, t_2, \dots, t_N = T\}$ (N -Perioden-Modell; diskretes Modell) oder $I = [0, T]$ (stetiges Modell) Dabei wird T als **Zeithorizont** bezeichnet.

Definition 1.3 (stochastischer Prozess)

Ein stochastischer Prozess S ist eine messbare Abbildung

$$S: \begin{cases} (\Omega \times I) & \rightarrow \mathbb{R}^d \\ (\omega, t) & \mapsto S_t(\omega) \end{cases}$$

Insbesondere ist

- $t \mapsto S_t(\omega)$ eine Funktion $I \rightarrow \mathbb{R}^d$ für jedes $\omega \in \Omega$
- $\omega \mapsto S_t(\omega)$ eine Zufallsvariable $\Omega \rightarrow \mathbb{R}^d$ für jedes $t \in I$

(3) Definition 1.4 (Filtration)

Eine Filtration ist eine Folge von σ -Algebren $(\mathcal{F}_t)_{t \in I}$ mit der Eigenschaft

$$\mathcal{F}_s \subseteq \mathcal{F}_t \quad \forall s, t \in I, s \leq t \quad \text{und} \quad \mathcal{F}_t \subseteq \mathcal{F} \quad \forall t \in I$$

Interpretation. \mathcal{F}_t beschreibt die den Marktteilnehmern zum Zeitpunkt t bekannte bzw. verfügbare Information. Ein Ereignis $A \in \mathcal{F}_t$ gilt als ”zum Zeitpunkt t bekannt“.

Erinnerung. Eine \mathbb{R}^d -wertige Zufallsvariable X heißt \mathcal{F}_t -messbar, wenn

$$X^{-1}(B) \in \mathcal{F}_t \quad \forall \text{ Borelmengen } B \subseteq \mathbb{R}^d$$

Beispiel 1.5

Sei S ein stochastischer Prozess. Dann heißt

$$\mathcal{F}_t^S = \sigma(\{(S_r) : r \in I, r \leq t\})$$

von S erzeugte Filtration.

Definition 1.6 (adaptierter Prozess)

Ein stochastischer Prozess $(S_t)_{t \in I}$ auf (Ω, \mathcal{F}) heißt **adaptiert** bezüglich einer Filtration $(\mathcal{F}_t)_{t \in I}$, wenn gilt S_t ist \mathcal{F}_t -messbar für alle $t \in I$.

Interpretation: Der Wert S_t ist zum Zeitpunkt t "bekannt".

Warum Filtrationen in der Finanzmathematik?

- Unterscheidung Zukunft/Vergangenheit
- Unterscheidung Informationen (Insider/Outsider) Unterscheidung Filtration $(\mathcal{F}_t)_{t \in I}$ bzw. $(\mathcal{G}_t)_{t \in I}$

(4) Anlagegüter [assets]: \mathbb{R}^{d+1} -wertiger stochastischer Prozess mit Komponenten

$$S^i: \begin{cases} (\Omega \times I) & \rightarrow \mathbb{R} \\ (\omega, t) & \mapsto S_t^i(\omega) \end{cases} \quad (i \in \{0, 1, \dots, d\})$$

S_t^i beschreibt dabei den Preis des i -ten Anlageguts zum Zeitpunkt t . S^i ($i \in \{1, \dots, d\}$) ist typischerweise

- Aktien [stock], Unternehmensanteil
- Währung [currency] bzw. Wechselkurs
- Rohstoff [commodity] wie z.B. Öl, Edelmetall, Elektrizität
- Anleihe [bond] ... Schuldverschreibung

Hauptannahme: S^i ist liquide gehandelt (z.B. Börse), d.h. der Kauf und Verkauf zum Preis S_t^i ist jederzeit möglich. Der "Numeraire" S^0 hat eine Sonderrolle und beschreibt die Verzinsung von nicht in (S^1, \dots, S^d) angelegtem Kapital. Er wird als risikolos betrachtet.

Definition 1.7

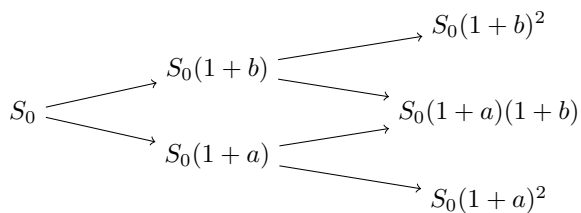
Ein Finanzmarktmodell (FFM) mit Zeitachse I ist gegeben durch

- (1) einen Wahrscheinlichkeitsraum $(\Omega, \mathcal{F}, \mathbb{P})$ mit Filtration $(\mathcal{F}_t)_{t \in I}$
- (2) einem an $(\mathcal{F}_t)_{t \in I}$ adaptierten, \mathbb{R}^{d+1} -wertigen stochastischen Prozess $S_t = (S_t^0, S_t^1, \dots, S_t^d)$ mit $t \in I$.

Beispiel 1.8 (Cox-Ross-Rubinstein-Modell)

Das CRR-Modell ist ein zeitdiskretes Modell beschrieben durch

- $S_n^0 = (1 + r)^n$... Verzinsung mit konstanter Rate r
- $S_n^1 = S_0^1 \prod_{k=1}^n (1 + R_k)$, wobei (R_1, R_2, \dots) unabhängige Zufallsvariablen mit zwei möglichen Werten $a < b$ sind



\hookrightarrow rekombinierender Baum,

Abbildung 1.1: Cox-Ross-Rubinstein-Modell

Ereignisse ω entsprechen Pfaden im Baum

Beispiel 1.9 (Black-Scholes-Modell, zeitstetig)

Beim Black-Scholes-Modell handelt es sich um ein zeitstetiges Modell auf einem unendlichen Wahrscheinlichkeitsraum.

$$S_t^0 = e^{rt} \quad (\text{Verzinsung mit konstanter Rate } r)$$

$$S_t^1 = S_0^1 \cdot \exp\left(\left(\mu - \frac{\sigma^2}{2}\right)t + \sigma B_t\right) \quad \text{mit } \mu \in \mathbb{R}, \sigma > 0, S_0^1 > 0$$

Der Term $\mu - \frac{\sigma^2}{2}$ beschreibt dabei eine Trendkomponenten, B_t eine "Brownsche Bewegung" (zeitstetiger Prozess).

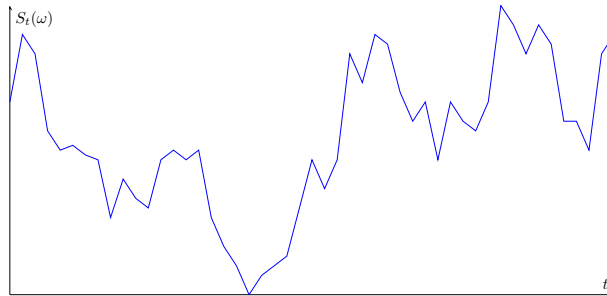


Abbildung 1.2: Black-Scholes-Modell

1.3 Anleihen und grundlegende Beispiele für Derivate

Hier betrachten wir immer nur ein Basisgut $S_t = S_t^1$.

(a) **Anleihe** [bond] (genauer: Null-Kupon-Anleihe [zero-coupon bond])

Der Emittent (Herausgeber) einer Anleihe mit Endfälligkeit [maturity] T garantiert dem Käufer zum Zeitpunkt T den Betrag N (EUR/USD/...) zu zahlen. Typische Emittenten sind z.B. Staaten [government bond] oder Unternehmen (als Alternative zur Kreditaufnahme). Nach Emission werden Anleihen auf dem Sekundärmarkt weiterverkauft, d.h. liquide gehandelte Wertpapiere.

Preis bei Emission: $B(0, T)$

Preis bei Weiterverkauf zum Zeitpunkt $t \leq T$: $B(t, T)$

Es ist $B(T, T) = N$ und wir normieren stets $N = 1 \Rightarrow B(T, T) = 1$.

Anleihen von West-/ Nord-/ Mitteleuropäischen Staaten und den USA sowie Kanada werden als risikolos betrachtet (sichere Zahlung). Sonst: Kreditrisiko

Risikofreie Anleihen können als Numeraire $S_t^0 = B(t, T)$ genutzt werden.

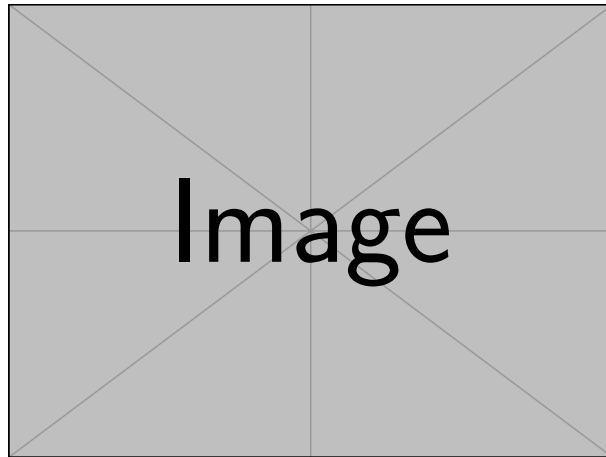


Abbildung 1.3: Zahlungsstrom einer Anleihe

(b) **Terminvertrag** [forward contract]

aus Käufersicht: Vereinbarung zu bestimmtem zukünftigen Zeitpunkt T eine Einheit des Basisguts S zum Preis K zu kaufen (Kaufverpflichtung). Beliebte Beispiele sind Rohstoffe und Elektrizität.

Auszahlungsprofil: $F_T = S_T - K$ Preis zum Zeitpunkt t : F_t

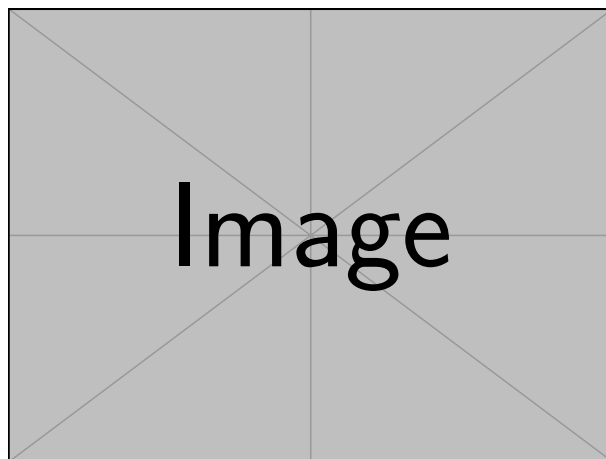


Abbildung 1.4: Auszahlungsprofil eines Terminvertrags

(c) **(Europäische) Put- bzw. Call-Option**

Recht zu einem zukünftigen Zeitpunkt T eine Einheit des Basisguts S zum Preis K zu verkaufen (put) bzw. zu kaufen (call) → keine Kaufverpflichtung !

Auszahlungsprofil:

- Call: $C_T = \begin{cases} S_T - K & S_T \geq K \\ 0 & S_T < K \end{cases} = (S_T - K)_+$

- Put: $P_T = \begin{cases} 0 & S_T \geq K \\ K - S_T & S_T < K \end{cases} = (K - S_T)_+$

(d) **Amerikanische Put- bzw. Call-Option**

wie Put/Call, aber mit Ausübung zu beliebigem Zeitpunkt $\tau \in [0, T]$.

Preis zum Zeitpunkt t : P_t^{AM}, C_t^{AM}

Auszahlungsprofil zum Zeitpunkt τ : $(S_\tau - K)_+, (K - S_\tau)_+$

Der Zeitpunkt τ muss im Allgemeinen als Lösung eines stochastischen Optimierungsproblems bestimmt werden (optimales Stopp-Problem).

1.4 Elementare Replikations- und Arbitrageargumente

Was können wir (mit elementaren Mitteln) über die "fairen" Preise $B(t, T), F_t, C_t, P_t$ aussagen?

Wir verwenden:

- **Replikationsprinzip**: zwei identische, zukünftige Zahlungsströme haben auch heute denselben Wert (ein Zahlungsstrom "repliziert" den anderen)
- **No-Arbitrage-Prinzip**: "Ohne Kapitaleinsatz kann kein sicherer Gewinn ohne Verlustrisiko erzielt werden." (Arbitrage = risikofreier Gewinn)
- **Superreplikationsprinzip** (schwächere Form des Replikationsprinzips): Ist ein Zahlungsstrom in jedem Fall größer als ein anderer, so hat er auch heute den größeren Wert.

stark	Replikationsprinzip	eingeschränkt anwendbar
↓	Superreplikationsprinzip	↑
schwach	No-Arbitrage-Prinzip	immer anwendbar

Lemma 1.1

Für den Preis C_T des europäischen Calls gilt:

$$(S_t - KB(t, T))_+ \leq C_t \leq S_t$$

Beweis. untere Schranke: Für Widerspruch nehme an, dass $S_t - KB(t, T) - C_t = \varepsilon > 0$.

Portfolio	Wert in t	Wert in T	
		$S_T \leq K$	$S_T > K$
Kaufe Call	C_t	0	$S_T - K$
Verkaufe Basisgut	$-S_t$	$-S_T$	$-S_T$
Kaufe Anleihe	$\varepsilon + KB(t, T)$	$\frac{\varepsilon}{B(t, T)} + K$	$\frac{\varepsilon}{B(t, T)} + K$
Σ	0 kein Anfangskapital	$K - S_T + \frac{\varepsilon}{B(t, T)} > 0$	$\frac{\varepsilon}{B(t, T)} > 0$
		sicherer Gewinn	

Dies steht jedoch im Widerspruch zum No-Arbitrage-Prinzip. Somit ist $S_t - KB(t, T) \leq C_t$. Außerdem ist $C_t \geq 0$, d.h. $C_t \geq (S_t - KB(t, T))_+$.

obere Schranke: ↗ Übung

□

Lemma 1.2 (Put-Call-Parität)

Für Put P_t , Call C_t mit selbem Ausübungspreis K und Basisgut S_t gilt

$$C_t - P_t = S_t - B(t, T) \cdot K$$

Beweis. Mit Replikationsprinzip:

Portfolio 1	Wert in t	Wert in T	
		$S_T \leq K$	$S_T > K$
Kaufe Call	C_t	0	$S_T - K$
Kaufe Anleihe	$K \cdot B(t, T)$	K	K
Wert Portfolio 1	$C_t + K \cdot B(t, T)$	K	S_T

Portfolio 2	Wert in t	Wert in T	
		$S_T \leq K$	$S_T > K$
Kaufe Put	P_t	$K - S_T$	0
Kaufe Basisgut	S_t	S_T	S_T
Wert Portfolio 2	$P_t + S_t$	K	S_T

Replikationsprinzip: $C_t + K \cdot B(t, T) = P_t + S_t \Rightarrow C_t + P_t = S_t - K \cdot B(t, T)$

□

1.5 Bedingte Erwartungswerte und Martingale

1.5.1 Bedingte Dichte und bedingter Erwartungswert

Motivation: Gegeben seien zwei Zufallsvariablen (X, Y) mit Werten in $\mathbb{R}^m \times \mathbb{R}^n$ und gemeinsamer Dichte $f_{XY}(x, y)$.

Aus Dichte f_{XY} können wir ableiten:

- $f_Y(y) := \int_{\mathbb{R}^m} f_{XY}(x, y) \, dx$, die Randverteilung von Y
- $S_y := \{y \in \mathbb{R}^n : f_Y(y) > 0\}$, der Träger von Y

Definition 1.10

Die bedingte Dichte von X bzgl. Y ist definiert als

$$f_{X|Y}(x, y) = \begin{cases} \frac{f_{XY}(x, y)}{f_Y(y)} & y \in S_y \\ 0 & \notin S_y \end{cases}$$

Betrachte folgende Problemstellung: Was ist die beste Vorhersage von X gegeben eine Beobachtung $Y = y$?

Kriterium: Minimiere quadratischen Abstand bzw. das zweite Moment bzw. die L_2 -Norm.

Vorhersage: messbare Funktion $g: \mathbb{R}^n \rightarrow \mathbb{R}^m, y \mapsto g(y)$.

$$\min \left\{ \mathbb{E} \left[(X - g(y))^2 \right] : g \text{ messbar } \mathbb{R}^n \rightarrow \mathbb{R}^m \right\} \quad (\text{min-1})$$

Proposition 1.3

Wenn (X, Y) eine gemeinsame Dichte besitzen und $\mathbb{E} [|X|^2] < \infty$ gilt, dann wird (min-1) minimiert durch die bedingte Erwartung

$$g(y) = \mathbb{E} [X|Y = y] := \int_{\mathbb{R}^m} x f_{X|Y}(x, y) \, dx$$

Wir bezeichnen $\mathbb{E} [X|Y = y]$ als Erwartungswert von X bedingt auf $Y = y$.

Allgemeiner gilt:

Theorem 1.4

Seien (X, Y) Zufallsvariablen mit gemeinsamer Dichte auf $\mathbb{R}^m \times \mathbb{R}^n$ und $h: \mathbb{R}^m \times \mathbb{R}^n \rightarrow \mathbb{R}$ messbar mit $\mathbb{E} [h(X, Y)^2] < \infty$. Dann wird das Minimierungsproblem

$$\min \left\{ \mathbb{E} \left[(h(X, Y) - g(Y))^2 \right] : g \text{ messbar } \mathbb{R}^n \rightarrow \mathbb{R} \right\}$$

gelöst durch

$$g(y) = \mathbb{E} [h(X, Y)|Y = y] = \int_{\mathbb{R}^m} h(x, y) \cdot f_{X|Y}(x, y) \, dx$$

Beweis (nur Proposition für $m = 1$, Theorem analog). Setze $g(y) = \int_{\mathbb{R}} x f_{X|Y}(x, y) \, dx$. Sei $p: \mathbb{R}^n \rightarrow \mathbb{R}$ eine beliebige messbare Funktion mit $\mathbb{E} [p(Y)^2] < \infty$. Setze weiter $g_\varepsilon(y) = g(y) + \varepsilon p(y)$. Minimiere $F(\varepsilon) := \mathbb{E} [(X - g_\varepsilon(y))^2] = \mathbb{E} [(X - g(y) - \varepsilon p(y))^2] = \mathbb{E} [(X - g(Y))^2] - 2\varepsilon \mathbb{E} [(X - g(Y))p(Y)] + \varepsilon^2 \mathbb{E} [p(Y)^2]$.

$$\frac{\partial F}{\partial \varepsilon}(\varepsilon) = 2\varepsilon \mathbb{E} [p(Y)^2] - 2\mathbb{E} [(X - g(Y))p(Y)] \Rightarrow \varepsilon_* := \frac{\mathbb{E} [(X - g(Y))p(Y)]}{\mathbb{E} [p(Y)^2]} = \frac{A}{B}$$

$$\begin{aligned} A &= \mathbb{E} [Xp(Y)] - \mathbb{E} [g(Y)p(Y)] \\ &= \int_{\mathbb{R} \times \mathbb{R}^n} xp(y) f_{XY}(x, y) \, dx \, dy - \int_{S_y} g(y)p(y) f_Y(y) \, dy \\ &= \int_{\mathbb{R} \times \mathbb{R}^n} xp(y) f_{XY}(x, y) \, dx \, dy - \int_{\mathbb{R} \times S_y} \underbrace{xp(y) f_{X|Y}(x, y) f_Y(y)}_{=f_{XY}(x, y)} \, dy = 0 \end{aligned}$$

Damit ist $\varepsilon^* = 0$ unabhängig von p und $g(y)$ minimiert (min-1). \square

Beispiel

Seien (X, Y) normalverteilt auf $\mathbb{R} \times \mathbb{R}$ mit

$$\mu = \begin{pmatrix} \mu_X \\ \mu_Y \end{pmatrix} \quad \Sigma = \begin{pmatrix} \text{Var}[X] & \text{Cov}(X, Y) \\ \text{Cov}(X, Y) & \text{Var}(Y) \end{pmatrix} = \begin{pmatrix} \sigma_x^2 & \rho\sigma_x\sigma_y \\ \rho\sigma_x\sigma_y & \sigma_y^2 \end{pmatrix} \quad \rho \in [-1, 1]$$

Dann ist die bedingte Dichte $f_{X|Y}(x, y)$ wieder die Dichte einer Normalverteilung mit

$$\begin{aligned}\mathbb{E}[X|Y = y] &= \mu_X + \rho \frac{\sigma_X}{\sigma_Y} (y - \mu_Y) \\ \text{Var}(X|Y = y) &= \sigma_X^2 (1 - \rho^2)\end{aligned}$$

→ siehe Übung.

Die Abbildung $y \mapsto \mu_X + \rho \frac{\sigma_X}{\sigma_Y} (Y - \mu_Y)$ heißt Regressionsgerade für X gegeben $Y = y$.

Die Steigung wird im Wesentlichen durch ρ bestimmt.

Für diskrete Zufallsvariablen, d.h. wenn X, Y nur endliche viele Werte $\{x_1, \dots, x_m\}$ bzw. $\{y_1, \dots, y_n\}$ annehmen, dann erhalten wir mit ähnlichen Überlegungen als Lösung von (min-1)

$$\mathbb{E}[X|Y = y_j] = \sum_{i=1}^m x_i \mathbb{P}(X = x_i | Y = y_j)$$

wobei direkt die bedingten Wahrscheinlichkeiten

$$\mathbb{P}(X = x_i | Y = y_j) = \begin{cases} \frac{\mathbb{P}(X=x_i \wedge Y=y_j)}{\mathbb{P}(Y=y_j)} & \text{wenn } \mathbb{P}(Y = y_j) > 0 \\ 0 & \text{wenn } \mathbb{P}(Y = y_j) = 0 \end{cases}$$

folgen.

1.5.2 Bedingte Erwartung: Maßtheoretischer Zugang

Wir betrachten den Wahrscheinlichkeitsraum $(\Omega, \mathcal{F}, \mathbb{P})$. Für eine Zufallsvariable $X: \Omega \rightarrow \mathbb{R}$ und $p \in [1, \infty)$ definieren wir, die L_p -Norm

$$\|X\|_p := \mathbb{E}[|X|^p]^{\frac{1}{p}} = \left(\int_{\Omega} |X(\omega)|^p \, d\mathbb{P}(\omega) \right)^{\frac{1}{p}}$$

und den L_p -Raum

$$L_p(\Omega, \mathcal{F}, \mathbb{P}) := \left\{ X: \Omega \rightarrow \mathbb{R} \text{ } \mathcal{F}\text{-messbar, } \|X\|_p < \infty \right\}$$

Dabei identifizieren wir Zufallsvariablen, die sich nur auf \mathbb{P} -Nullmengen unterscheiden miteinander, d.h. $\mathbb{P}(X \neq X') = 0 \Rightarrow X = X'$ in L_p . Aus der Maßtheorie bekannt: Die Räume $L_p(\Omega, \mathcal{F}, \mathbb{P})$ mit Norm $\|\dots\|_p$ mit $p \in [1, \infty)$ sind

- Banachräume, d.h. vollständige, normierte Vektorräume.
- für $p = 2$ auch Hilbertraum mit inneren Produkt

$$\langle X, Y \rangle = \mathbb{E}[XY] = \int_{\Omega} X(\omega)Y(\omega) \, d\mathbb{P}(\omega)$$

Für $\mathcal{G} \subseteq \mathcal{F}$ Unter- σ -Algebra ist $L_p(\Omega, \mathcal{G}, \mathbb{P}) \subseteq L_p(\Omega, \mathcal{F}, \mathbb{P})$ ein abgeschlossener Unterraum.

Wir verallgemeinern das "Vorhersageproblem" aus dem letzten Abschnitt: Gegeben sei eine Zufallsvariable X aus $L_2(\Omega, \mathcal{F}, \mathbb{P})$ und $\mathcal{G} \subseteq \mathcal{F}$ eine Unter- σ -Algebra. Was ist die beste \mathcal{G} -messbare Vorhersage für X ?

$$\min \left\{ \mathbb{E} \left[(X - G)^2 \right] : G \in L_2(\Omega, \mathcal{G}, \mathbb{P}) \right\} \quad (\text{min-2})$$

Aus Hilbertraumtheorie folgt, dass (min-2) besitzt eine eindeutige Lösung $G_* \in L_2(\Omega, \mathcal{G}, \mathbb{P})$. G_* ist die Orthogonalprojektion (bzgl. $\langle \cdot, \cdot \rangle$) von $X \in L_2(\Omega, \mathcal{F}, \mathbb{P})$ auf den abgeschlossenen Unterraum $L_2(\Omega, \mathcal{G}, \mathbb{P})$.

Wir bezeichnen G_* mit $\mathbb{E}[X|\mathcal{G}]$ als bedingten Erwartungswert von X bezüglich \mathcal{G} .

Theorem 1.5

Seien $X, Y \in L_2(\Omega, \mathcal{F}, \mathbb{P})$ und $\mathcal{G} \subseteq \mathcal{F}$ eine Unter- σ -Algebra. Dann gilt

- Linearität: $\mathbb{E}[aX + bY|\mathcal{G}] = a\mathbb{E}[X|\mathcal{G}] + b\mathbb{E}[Y|\mathcal{G}]$
- Turmregel: Für jede weitere σ -Algebra $\mathcal{H} \subseteq \mathcal{G}$ gilt $\mathbb{E}[\mathbb{E}[X|\mathcal{G}]|\mathcal{H}] = \mathbb{E}[X|\mathcal{H}]$
- Pull-out-Property: $\mathbb{E}[XZ|\mathcal{G}] = Z \cdot \mathbb{E}[X|\mathcal{G}]$ für alle beschränkten und \mathcal{G} -messbaren Zufallsvariablen Z . Für Z \mathcal{G} -messbar mit $\mathbb{E}[|XZ|] < \infty$ gilt $\mathbb{E}[XZ|\mathcal{G}] = Z \cdot \mathbb{E}[X|\mathcal{G}]$. Insbesondere gilt für \mathcal{G} -messbare X schon $\mathbb{E}[X|\mathcal{G}] = X$.
- Monotonie: $X \leq Y \Rightarrow \mathbb{E}[X|\mathcal{G}] \leq \mathbb{E}[Y|\mathcal{G}]$
- Dreiecksungleichung: $|\mathbb{E}[X|\mathcal{G}]| \leq \mathbb{E}[|X||\mathcal{G}]$
- Unabhängigkeit: X unabhängig von $\mathcal{G} \Rightarrow \mathbb{E}[X|\mathcal{G}] = \mathbb{E}[X]$
- triviale σ -Algebra: $\mathcal{G} = \{\emptyset, \Omega\} \Rightarrow \mathbb{E}[X|\mathcal{G}] = \mathbb{E}[X]$

Beweis. siehe VL "Wahrscheinlichkeitstheorie mit Martingalen" □

Die für $X \in L_2(\Omega, \mathcal{F}, \mathbb{P})$ definierte bedingte Erwartung $\mathbb{E}[X|\mathcal{G}]$ lässt sich durch Approximation auf alle $X \in L_1(\Omega, \mathcal{F}, \mathbb{P})$ erweitern. Alle Eigenschaften aus Theorem 1.5 bleiben erhalten.

Sei Y eine Zufallsvariable und $\mathcal{G} = \sigma(Y)$ die von Y erzeugte σ -Algebra. Wir schreiben $\mathbb{E}[X|Y] = \mathbb{E}[X|\sigma(Y)]$; dies ist eine \mathcal{G} -messbare Zufallsvariable.

Aus der Maßtheorie sagt uns das Doob-Dynkin-Lemma, dass eine messbare Funktion $g: \mathbb{R}^n \rightarrow \mathbb{R}$ existiert, sodass $\mathbb{E}[X|Y] = g(Y)$. Dabei ist g genau die Funktion aus (min-1).

Zusammenfassung

Sei X, Y aus $L_1(\Omega, \mathcal{F}, \mathbb{P})$ und $\mathcal{G} \subseteq \mathcal{F}$ eine Unter- σ -Algebra.

- (a) $\mathbb{E}[X|Y = y]$ ist eine messbare Funktion $g: \mathbb{R}^n \rightarrow \mathbb{R}^m$ und falls eine bedingte Dichte existiert, dann gilt $\mathbb{E}[X|Y = y] = \int_{\mathbb{R}^m} x f_{X|Y}(x, y) dx$.
- (b) $\mathbb{E}[X|Y]$ ist eine $\sigma(Y)$ -messbare Zufallsvariable und kann als $g(Y)$ dargestellt werden. Falls eine bedingte Dichte existiert, dann gilt $\mathbb{E}[X|Y](\omega) = \int_{\mathbb{R}^n} x f_{X|Y}(x, Y(\omega)) dx$.
- (c) $\mathbb{E}[X|\mathcal{G}]$ ist eine \mathcal{G} -messbare Zufallsvariable. Falls $\mathcal{G} = \sigma(Y)$ tritt Fall (b) ein.

In allen Fällen kann $\mathbb{E}[X|\cdot]$ interpretiert werden als *beste Vorhersage* für X gegeben

- (a) eine punktweise Betrachtung $Y = y$
- (b) die Beobachtung Y
- (c) die Information \mathcal{G}

1.5.3 Martingale

Prototyp eines "neutralen" stochastischen Prozesses, der weder Aufwärts- noch Abwärtstrend besitzt. Wir betrachten hier den Prozess nur in diskreter Zeit $I = \mathbb{N}_0$.

Definition 1.11

Sei $(X)_{n \in \mathbb{N}_0}$ ein stochastischer Prozess. Wenn gilt

$$\begin{aligned} \mathbb{E}[|X_n|] &< \infty & \forall n \in \mathbb{N} \\ \mathbb{E}[X_{n+1}|X_1, \dots, X_n] &= X_n & \forall n \in \mathbb{N} \end{aligned} \quad (1.1)$$

dann heißt $(X_n)_{n \in \mathbb{N}}$ **Martingal**.

Wenn wir $\mathcal{F}_n^X = \sigma(X_1, \dots, X_n)$ definieren, können wir (b) schreiben als

$$\mathbb{E}[X_{n+1}|\mathcal{F}_n^X] = X_n \quad \forall n \in \mathbb{N}$$

Konvention: Alle stochastischen Prozesse $(X_n)_{n \in \mathbb{N}_0}$ haben deterministischen Startwert X_0 .

Interpretation: Beste Vorhersage für zukünftigen Wert X_{n+1} basierend auf Vergangenheit $\sigma(X_1, \dots, X_n)$ ist der momentane Wert X_n . Aus der Turmregel folgt:

$$\mathbb{E}[X_{n+k}|\mathcal{F}_n^X] = X_n \quad \forall n, k \in \mathbb{N}_0$$

denn

$$\mathbb{E}[X_{n+k}|\mathcal{F}_n^X] = \mathbb{E}\left[\underbrace{\mathbb{E}[X_{n+k}|\mathcal{F}_{n+k-1}^X]}_{=X_{n+k-1}}|\mathcal{F}_n^X\right] = \mathbb{E}[X_{n+k-1}|\mathcal{F}_n^X] \stackrel{k \text{ mal}}{=} X_n$$

Man kann von $(\mathcal{F}_n^X)_{n \in \mathbb{N}}$ auf beliebige Filtrationen $(\mathcal{F}_n)_{n \in \mathbb{N}_0}$ erweitert werden.

Definition. Sei $(X_n)_{n \in \mathbb{N}_0}$ ein stochastischer Prozess, adaptiert an eine Filtration $(\mathcal{F}_n)_{n \in \mathbb{N}_0}$. Wenn gilt

$$\begin{aligned} \mathbb{E}[|X_n|] &< \infty \quad \forall n \in \mathbb{N}_0 \\ \mathbb{E}[X_{n+1}|\mathcal{F}_n] &= X_n \quad \forall n \in \mathbb{N}_0 \end{aligned}$$

dann heißt $(X_N)_{n \in \mathbb{N}}$ Martingal bezüglich der Filtration $(\mathcal{F}_n)_{n \in \mathbb{N}}$.

Interpretation: Beste Vorhersage für zukünftigen Wert X_{n+1} , basierend auf verfügbarer Information \mathcal{F}_n ist der momentane Wert X_n .

Definition. Falls in Punkt (b) statt " $=$ " die Ungleichung " \leq " oder " \geq " gilt, so heißt $(X_n)_{n \in \mathbb{N}}$ **Super- bzw. Submartingal**.

- Wenn $X = (X_n)_{n \in \mathbb{N}}$ ein Martingal ist, dann gilt $\mathbb{E}[X_n] = X_0$, d.h. $n \mapsto \mathbb{E}[X_n]$ ist konstant.
Begründung:

$$\mathbb{E}[X_{n+1} | \mathcal{F}_n] = X_n \Rightarrow \underbrace{\mathbb{E}[\mathbb{E}[X_{n+1} | \mathcal{F}_n]]}_{=\mathbb{E}[X_{n+1}]} = \mathbb{E}[X_n] \stackrel{n \text{ mal}}{\Rightarrow} \mathbb{E}[X_n] = X_0$$

- X Submartingal $\Rightarrow n \mapsto \mathbb{E}[X_n]$ ist monoton steigend
- X Supermartingal $\Rightarrow n \mapsto \mathbb{E}[X_n]$ ist monoton fallend

"Das Leben ist ein Supermartingal – die Erwartungen fallen mit der Zeit" ☺

Beispiel 1.12

- Seien $(Y_n)_{n \in \mathbb{N}}$ unabhängige Zufallsvariablen in $L_1(\Omega, \mathcal{F}, \mathbb{P})$ mit $\mathbb{E}[Y_n] = 0$. Betrachten wir die Partialsummen $X_n := \sum_{k=1}^n Y_k$ und $X_0 = 0$. Dann ist $(X_n)_{n \in \mathbb{N}_0}$ ein Martingal, denn

$$\begin{aligned} \mathbb{E}[|X_n|] &\leq \sum_{k=1}^n \mathbb{E}[|Y_k|] < \infty \quad \forall n \in \mathbb{N} \\ \mathbb{E}[X_{n+1} | \mathcal{F}_n^X] &= \mathbb{E}[Y_{n+1} + X_n | \mathcal{F}_n^X] \\ &= \mathbb{E}[Y_{n+1} | \mathcal{F}_n^X] + \mathbb{E}[X_n | \mathcal{F}_n^X] = \underbrace{\mathbb{E}[Y_{n+1}]}_{=0} + X_n \\ &= X_n \end{aligned}$$

Definition. Sei $(\mathcal{F}_n)_{n \in \mathbb{N}_0}$ eine Filtration. Ein stochastischer Prozess $(H_n)_{n \in \mathbb{N}}$ heißt **vorhersehbar** [predictable] bezüglich $(\mathcal{F}_n)_{n \in \mathbb{N}_0}$ wenn gilt

$$H_n \text{ ist } \mathcal{F}_{n-1}\text{-messbar} \quad \forall n \in \mathbb{N}$$

Bemerkung

Vorhersehbarkeit ist eine stärkere Eigenschaft als Adaptiertheit.

Definition. Sei X ein adaptierter und H ein vorhersehbarer stochastischer Prozess bezüglich $(\mathcal{F}_n)_{n \in \mathbb{N}}$. Dann heißt

$$(H \bullet X)_n := \sum_{k=1}^n H_k (X_k - X_{k-1}) \quad (\star)$$

diskretes stochastisches Integral von H bezüglich X .

Bemerkung

Summen (\star) heißen in der Analysis Riemann-Stieltjes-Summen und werden für die Konstruktion

des Riemann-Stieltjes-Integrals $\int h \, d\rho$ verwendet.

Definition. Ein stochastischer Prozess $(H_n)_{n \in \mathbb{N}}$ heißt lokal beschränkt, wenn eine deterministische Folge $c_n \in \mathbb{R}_{\geq 0}$ existiert, sodass

$$|H_n| \leq c_n \quad \text{fast sicher} \quad \forall n \in \mathbb{N}$$

Satz 1.6

Sei X adaptierter stochastischer Prozess (bezüglich Filtration $(\mathcal{F}_n)_{n \in \mathbb{N}}$). Dann sind äquivalent:

- (1) X ist Martingal bezüglich $(\mathcal{F}_n)_{n \in \mathbb{N}}$.
- (2) $(H \bullet X)$ ist Martingal für alle lokal beschränkten, vorhersbaren Prozesse H

Das heißt, dass das stochastische Integral die Martingaleigenschaft erhält.

Beweis. (\Rightarrow) $(H \bullet X)_n = \sum_{k=1}^n H_k(X_k - X_{k-1})$.

- Adaptiertheit: klar
- Integrierbarkeit: H lokal beschränkt, d.h. $|H_k| \leq c_k < \infty$ für alle k .

$$\mathbb{E}[|H_k(X_k - X_{k-1})|] \leq c_k \cdot (\mathbb{E}[|x_k|] + \mathbb{E}[|X_{k+1}|]) < \infty$$

Mit der Dreiecksungleichung folgt daraus $\mathbb{E}[|(H \bullet X)_n|] < \infty$.

- Martingaleigenschaft:

$$\mathbb{E}[(H \bullet X)_n \mid \mathcal{F}_{n-1}] = (H \bullet X)_{n-1} + \mathbb{E}[H_n(X_n - X_{n-1}) \mid \mathcal{F}_{n-1}] = (H \bullet X)_{n-1} + H_n \cdot \underbrace{(\mathbb{E}[X_n \mid \mathcal{F}_{n-1}] - X_{n-1})}_{=0}$$

Damit ist also auch $(H \bullet X)$ ein Martingal.

(\Leftarrow) Fixiere $N \in \mathbb{N}$. Setze $H_n := \mathbb{1}_{n=N}$, dieser ist lokal beschränkt und deterministisch (also auch vorhersehbar). Man stellt fest, dass $(H \bullet X)_n = 0$ für alle $n \leq N-1$. Für alle $n \geq N$ gilt dagegen $(H \bullet X)_n = X_N - X_{N-1}$. Wir überprüfen nur die Martingaleigenschaft (Integrierbarkeit folgt aus Dreiecksungleichung). Wir wissen, dass $(H \bullet X)$ ein Martingal ist.

$$0 = (H \bullet X)_{N-1} = \mathbb{E}[(H \bullet X)_N \mid \mathcal{F}_{N-1}] = \mathbb{E}[X_N - X_{N-1} \mid \mathcal{F}_{N-1}] = \mathbb{E}[X_N \mid \mathcal{F}_{N-1}] - X_{N-1} \Rightarrow \mathbb{E}[X_N \mid \mathcal{F}_{N-1}] = X_{N-1}$$

Somit ist X ein Martingal. □

Korollar 1.7

Sei $X = (X_n)_{n=1, \dots, N}$ ein adaptierter stochastischer Prozess bezüglich einer Filtration $(\mathcal{F}_n)_{n=1, \dots, N}$. Wenn $\mathbb{E}[(H \bullet X)_N] = 0$ für alle lokal beschränkten vorhersehbaren Prozesse H , dann ist X ein Martingal bezüglich (\mathcal{F}_n) .

Beweis. Fixiere $K \in [N] := \{1, 2, \dots, N\}$ und eine Menge $A \in \mathcal{F}_{K-1}$. Definiere $H_n(\omega) =$

$\mathbb{1}_A(\omega) \cdot \mathbb{1}_{\{n=K\}}$, dieser ist lokal beschränkt und vorhersehbar. Es ist $(H \bullet X)_n = 0$ für alle $n \leq K - 1$. Für alle $n \geq K$ gilt $(H \bullet X)_n = \mathbb{1}_A \cdot (X_K - X_{K-1})$.

für beliebige K . Somit ist X ein Martingal. □

Bemerkung 1.13

Wir schreiben $[N] := \{1, 2, \dots, N\}$ und $[N]_0 := \{0, 1, 2, \dots, N\}$.

Kapitel 2

COX-ROSS-RUBINSTEIN-MODELL

Das Cox-Ross-Rubinstein-Modell (kurz: CRR-Modell) wird auch Binomialmodell genannt und wurde 1979 von Cox, Ross und Rubinstein entwickelt.

Es handelt sich dabei um ein Modell für die Preisentwicklung eines Wertpapiers plus ein Verrechnungskonto mit konstanter Verzinsung (Numeraire) in diskreter Zeit.

Parameter:

r	Zinsrate
b	Rendite des Wertpapiers bei Aufwärtsbewegung ("up")
a	Rendite des Wertpapiers bei Abwärtsbewegung ("down")
$p \in (0, 1)$	Wahrscheinlichkeit für "up"
$S_0 > 0$	Preis Wertpapier zum Zeitpunkt Null
$N \in \mathbb{N}$	Anzahl der Zeitschritte

Annahmen: $r > -1$, $b > a > -1$

Wir modellieren Wertpapiere $(S_k)_{k \in [N]}$ und Verrechnungskonto $(S_k^0)_{k \in \mathbb{N}}$ als stochastische Prozesse auf einem Wahrscheinlichkeitsraum $(\Omega, \mathcal{F}, \mathbb{P})$.

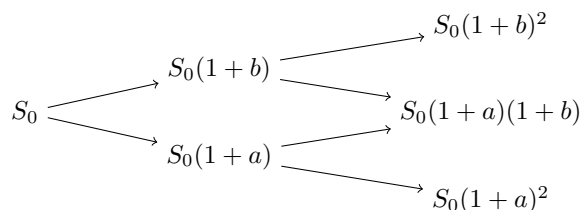
- $S_0^0 = 1$ und $S_n^0 = (1 + r)^n$
- Wir definieren die **Rendite** $R_n(\omega)$ in der n -ten Marktperiode durch

$$R_n = \begin{cases} b & \text{mit } p \\ a & \text{mit } 1 - p \end{cases}$$

Die Renditen (R_1, \dots, R_N) sind unabhängig.

$$S_n = S_0 \cdot \prod_{k=1}^n (1 + R_k)$$

Der Verlauf von S lässt sich grafisch als Binomialbaum darstellen:



Man nennt dies auch ein "rekombinierendes Baummmodell". Es hat den Vorteil, dass die Anzahl der Knoten nur linear mit n wächst.

- Abgezinster Preisprozess $\tilde{S}_n := \frac{S_n}{S_0} = S_0 \cdot \prod_{k=1}^n \frac{1+R_k}{1+r}$.
- Filtration: natürliche Filtration $\mathcal{F}_n = \sigma(S_1, \dots, S_n)$.

Proposition 2.1

Im CRR-Modell gilt:

- (a) Die Anzahl der Aufwärtsbewegungen $U_n := \#\{k \in [n] : R_k = b\}$ ist binomialverteilt, d.h. $U_n \sim \text{Bin}(n, p)$.

- (b) Es gilt

$$\log \left(\frac{\tilde{S}_n}{S_0} \right) = U_n \log \left(\frac{1+b}{1+a} \right) + n \log \left(\frac{1+a}{1+r} \right)$$

d.h. $\log \left(\frac{\tilde{S}_n}{S_0} \right)$ ist nach Skalen-Lagen-Transformation binomialverteilt.

- (c) Die Verteilung von S_n ist gegeben durch

$$\mathbb{P}(S_n = S_0(1+b)^k(1+a)^{n-k}) = \binom{n}{k} p^k (1-p)^{n-k}$$

Beweis. (zu a) klar

$$\text{(zu b)} \quad \frac{\tilde{S}_n}{S_0} = \left(\frac{1+b}{1+a} \right)^{U_n} \cdot \left(\frac{1+a}{1+r} \right)^n \Rightarrow \log \left(\frac{\tilde{S}_n}{S_0} \right) = U_n \log \left(\frac{1+b}{1+a} \right) + n \log \left(\frac{1+a}{1+r} \right)$$

(zu c) Es ist $S_n = S_0(1+b)^{U_1}(1+a)^{n-U_1}$. Also

$$\mathbb{P}(S_n = S_0(1+b)^k(1+a)^{n-k}) = \mathbb{P}(U_n = k) \stackrel{(a)}{=} \binom{n}{k} p^k (1-p)^{n-k} \quad \square$$

Bemerkung. Teil (b) suggeriert Konvergenz von $\log \left(\frac{\tilde{S}_n}{S_0} \right)$ gegen Normalverteilung für $n \rightarrow \infty$ (nach Skalierung) \leadsto Black-Scholes-Modell (\nearrow Kapitel 3).

2.1 Anlagestrategien im CRR-Modell

Eine Anlagestrategie soll durch ein Anfangskapital w und einen stochastischen Prozess $(\eta_n, \xi_n)_{n \in [N]}$ dargestellt werden.

- Dabei steht ξ_n für die "Anzahl" der Wertpapiere, die im Zeitintervall $(n-1, n]$ gehalten werden. Negative Werte von ξ_n sind sogenannte Leerverkäufe. Wir erlauben beliebige Anteile $\xi_n \in \mathbb{R}$.
- Weiter beschreibt η_n den Stand des Verrechnungskontos in Geldeinheiten zum Zeitpunkt n . Negative Werte von η_n entsprechend einer Kreditaufnahme. Außerdem sei das Anfangskapital $w \in \mathbb{R}$.

Gesamtwert des Portfolios

$$\begin{aligned}\Pi_n &= \eta_n \cdot S_n^0 + \xi_n \cdot S_n \\ \Pi_0 &= w\end{aligned}\tag{Port}$$

Man nennt (Port) die **Portfoliogleichung**.

Annahmen an die Strategie:

- Die Strategie darf nur von Beobachtungen der Vergangenheit abhängen, d.h. (η_n, ξ_n) ist ein *vorhersehbarer* Prozess.
- Portfolio wird nur zwischen S und S_0 umgeschichtet, d.h. es wird kein Kapital zugeschossen oder abgezogen:

Lemma 2.2

Eine selbstfinanzierende Anlagestrategie $(\eta_n, \xi_n)_{n \in \mathbb{N}}$ mit Anfangskapital $w \in \mathbb{R}$ und ihr Werteprozess Π_n sind durch w und $(\xi_n)_{n \in \mathbb{N}}$ vollständig definiert.

(a) Der diskontierte Werteprozess lässt sich darstellen als

$$\widetilde{\Pi}_n = w + \sum_{k=1}^n \xi_k (\widetilde{S}_k - \widetilde{S}_{k-1}) = w + (\xi \bullet \widetilde{S})_n$$

(b) Der Anteil η_n ist eindeutig gegeben durch

$$\eta_n = \widetilde{\Pi}_n - \xi_n \widetilde{S}_n$$

Beweis. siehe Herleitung oben □

2.2 Replikation bzw. Hedging von Derivaten im CRR-Modell

Betrachten wir ein Derivat C mit Auszahlung $h(S_1, S_2, \dots, S_N)$ zum Zeitpunkt N , d.h. $C = h(S_1, S_2, \dots, S_N)$ mit h messbar.

Gesucht ist nun eine replizierende Strategie $(\xi_n)_{n \in \mathbb{N}}$ und eine Anfangskapital w , d.h.

- $(\xi_n)_{n \in \mathbb{N}}$ soll vorhersehbar sein mit diskontiertem Werteprozess $\widetilde{\Pi}_n = w + (\xi \bullet \widetilde{S})_n$ und
- die Replikationsbedingung soll gelten

$$C = h(S_1, \dots, S_N) = \Pi_N \quad \text{fast sicher} \tag{Rep}$$

Definition. ▪ Ein Derivat C heißt **erreichbar**, wenn eine Replikationsstrategie existiert.

- Ein Finanzmarktmodell heißt **vollständig**, wenn jedes Derivat erreichbar ist.

Satz 2.3

Sei $C = h(S_1, \dots, S_N)$ ein Derivat im CRR-Modell. Dann ist C erreichbar, d.h. es existieren $w \in \mathbb{R}$ und $(\xi_n)_{n \in \mathbb{N}}$ mit (Rep). Es gilt:

- Es existieren messbare Funktionen $f_n: \mathbb{R}^n \rightarrow \mathbb{R}$ für $n = 1, \dots, N$, sodass

$$\Pi_n = f_n(S_1, \dots, S_n)$$

und die Werte entlang der Pfade des Binomialbaums sind rekursiv bestimmt:

$$\begin{aligned} f_N(S_1, \dots, S_N) &= h(S_1, \dots, S_N) = C \\ f_n(S_1, \dots, S_n) &= \frac{1}{1+r} \left(\frac{r-a}{b-r} f_{n+1}^b + \frac{b-r}{b-a} f_{n+1}^a \right) \end{aligned} \quad (\text{Rek})$$

wobei

$$\begin{aligned} f_{n+1}^b &= f_{n+1}(S_1, \dots, S_n, S_n(1+b)) \\ f_{n+1}^a &= f_{n+1}(S_1, \dots, S_n, S_n(1+a)) \quad n \in [N]_0 \end{aligned}$$

- Die replizierende Strategie ist gegeben durch

$$\xi_n = \frac{f_n^b - f_n^a}{S_{n-1}(b-a)} \quad (\Delta\text{-Hedge})$$

Korollar 2.4

Das CRR-Modell ist vollständig.

Korollar 2.5

Ist C ein europäisches Derivat, d.h. $C = h(S_N)$ mit $h: \mathbb{R} \rightarrow \mathbb{R}$ messbar, dann gelten folgende Vereinfachungen: Es reicht $f_n: \mathbb{R} \rightarrow \mathbb{R}$ zu wählen und es gilt $\Pi_n = f_n(S_n)$ sowie

$$f_{n+1}^b = f_{n+1}(S_n(1+b)) \quad \text{und} \quad f_{n+1}^a = f_{n+1}(S_n(1+a))$$

Bemerkung

- Die Rekursion (Rek) entspricht einem Rückwärtsdurchlauf des Baumdiagramms. f_n wird als diskontierter, gewichteter Mittelwert von f_{n+1}^b und f_{n+1}^a bestimmt. Die Gewichte sind dabei $q_b = \frac{r-a}{b-a}$ und $q_a = \frac{b-r}{b-a}$. Es gilt $q_a + q_b = 1$.
- Die ursprünglichen Übergangswahrscheinlichkeiten p spielen für Bewertung von C keine Rolle. Sie werden durch die "risikoneutralen" Wahrscheinlichkeiten q_b und $q_a = 1 - q_b$ ersetzt.
- Diese Rekursion lässt sich auf dem Computer auch für große Bäume effizient implementieren.
- Die Formel (Δ -Hedge) für ξ_n wird auch als Delta-Hedge bezeichnet.

$$\xi_n = \frac{\text{Preisänderung Derivat}}{\text{Preisänderung Basisgut}} \quad \dots \quad \text{Differenzenquotient}$$

- Weitere Interpretation von ξ_n :

- ▷ $\xi_n > 0$: Preisänderung Derivat hat selbes Vorzeichen wie Preisänderung Basisgut ; keine Leerverkäufe notwendig.
- ▷ $\xi_n < 0$: Preisänderung Derivat hat entgegengesetztes Vorzeichen wie Preisänderung Basisgut ; Leerverkäufe sind notwendig.
- ▷ $\xi_n \approx 0$: Preisänderung Derivat kaum von Preisänderung Basisgut abhängig

Beweis (Satz 2.3). Mittels Rückwärtsinduktion über $n \in [N]_0$.

(IA) Für $n = N$ gilt $\Pi_N = C = h(S_1, \dots, S_N)$ nach (Rep), also $\Pi_N = f_N(S_1, \dots, S_N)$ mit $f_N = h$.

(IS) Aus Selbstfinanzierungsbedingung-Schlange folgt

$$\begin{aligned} \tilde{\Pi}_{n+1} - \tilde{\Pi}_n &= \xi_{n+1} (\tilde{S}_{n+1} - \tilde{S}_n) \quad | \cdot (1+r)^{n+1} \\ \Rightarrow \quad \Pi_{n+1} - (1+r)\Pi_n &= \xi_{n+1} (S_{n+1} - (1+r)S_n) \end{aligned} \quad (\star)$$

Nach Induktionsvoraussetzung gilt also

$$\Pi_{n+1} = f_{n+1}(S_1, \dots, S_{n+1}) = f_{n+1}(S_1, \dots, S_n, S_n(1+R_{n+1}))$$

Die zwei Fälle $R_{n+1} = b$ und $R_{n+1} = a$ können jeweils mit strikt positiver Wahrscheinlichkeit eintreten.

- Fall 1: $S_{n+1} = S_n(1+b)$ und $\Pi_{n+1} = f_{n+1}(S_1, \dots, S_n, S_n(1+b)) = f_{n+1}^b$. Setzen wir dies in (\star) ein, so erhalten wir $f_{n+1}^b - (1+r)\Pi_n = \xi_{n+1}S_n(b-r)$ (I).
- Fall 2: $S_{n+1} = S_n(1+a)$ und $\Pi_{n+1} = f_{n+1}(S_1, \dots, S_n, S_n(1+a)) = f_{n+1}^a$. Setzen wir dies wieder in (\star) ein, so erhalten wir $f_{n+1}^a - (1+r)\Pi_n = \xi_{n+1}S_n(a-r)$ (II).

Damit erhalten wir ein lineares Gleichungssystem $[I, II]$ in den Unbekannten Π_n und ξ_n . Π_n und ξ_n sind \mathcal{F}_n -messbar, also unabhängig von R_{n+1} . Damit müssen (I) und (II) gleichzeitig erfüllt sein: Subtrahieren wir I - II, so erhalten wir $f_{n+1}^b - f_{n+1}^a = \xi_{n+1}S_n(b-a)$ und somit folgt (Δ -Hedge)

$$\Rightarrow \quad \xi_{n+1} = \frac{f_{n+1}^b - f_{n+1}^a}{S_n(b-a)}$$

Setzen wir dies wieder in I ein, so erhalten wir

$$f_{n+1}^b - (1+r)\Pi_n = \frac{b-r}{b-a} (f_{n+1}^b - f_{n+1}^a) \Rightarrow \Pi_n = \frac{1}{1+r} \left(\frac{r-a}{b-a} f_{n+1}^b + \frac{b-r}{b-a} f_{n+1}^a \right)$$

Dies ist die Rekursionsgleichung (Rek). □

Bemerkung. Das Lineare Gleichungssystem [I, II] können wir schreiben als

$$\begin{pmatrix} 1+r & b-r \\ 1+r & a-r \end{pmatrix} \cdot \begin{pmatrix} \Pi_n \\ \xi_{n+1}S_n \end{pmatrix} = \begin{pmatrix} f_{n+1}^b \\ f_{n+1}^a \end{pmatrix} \quad (\text{LGS-1})$$

Wir können auch ein Trinomialmodell (up, down, gleichbleibender Preis) betrachten. Dabei

bekommen wir eine zusätzliche Zeile in obiger Matrix, jedoch ist das Gleichungssystem dann überbestimmt und i.A. nicht mehr eindeutig auflösbar.

Beispiel (Asiatische Call-Option)

Die Auszahlung ist gegeben durch $(\bar{S}_N - K)_+$ mit $\bar{S}_N = \frac{1}{1+N} \sum_{k=0}^N S_k$. Somit handelt es sich um ein pfadabhängiges Derivat (Summe hängt von allen S_k ab). Bewertung im CRR-Modell mit zwei Perioden ($N = 2$) und Parametern $b = 0.3$, $a = -0.3$, $r = 0.1$, $S_0 = 100$, $K = 100$. Es gilt $C = h(S_1, S_2)$, also

$$\begin{aligned} h(120, 169) &= \left(\frac{399}{3} - 100 \right)_+ = 33 \\ h(130, 91) &= \left(\frac{321}{3} - 100 \right)_+ = 7 \\ h(70, 91) &= \left(\frac{261}{3} - 100 \right)_+ = 0 \\ h(70, 49) &= \left(\frac{219}{3} - 100 \right)_+ = 0 \end{aligned}$$

Außerdem gilt $h = f_2$. Wir berechnen die Übergangswahrscheinlichkeit

$$q = \frac{r - a}{b - a} = \frac{0.4}{0.6} = \frac{2}{3} \Rightarrow 1 - q = \frac{1}{3}$$

Rekursion:

$$\begin{aligned} f_1(130) &= \frac{1}{1+r} \left(q \cdot f^b + (1-q)f^a \right) \\ &= \frac{1}{1.1} \left(\frac{2}{3} \cdot 33 + \frac{1}{3} \cdot 7 \right) \\ &= \frac{1}{1.1} \cdot \frac{73}{3} \\ &\approx 22.12 \\ f_1(70) &= \frac{1}{1.1} \left(\frac{2}{3} \cdot 0 + \frac{1}{3} \cdot 0 \right) = 0 \\ f_0 &= \frac{1}{1.1} \left(\frac{2}{3} \cdot \frac{1}{1.1} \cdot \frac{73}{3} + \frac{1}{3} \cdot 0 \right) = \frac{1}{(1.1)^2} \frac{146}{9} \approx 13.41 \end{aligned}$$

Strategie

$$\begin{aligned} \xi_2(130) &= \frac{f_2^b - f_2^a}{S_1(b-a)} = \frac{33-7}{130 \cdot 0.6} = \frac{26}{13 \cdot 6} = \frac{1}{3} \\ \xi_2(70) &= \frac{0-0}{70 \cdot 0.6} = 0 \\ \xi_1 &= \frac{f_1^b - f_1^a}{S_0(b-a)} = \frac{\frac{1}{1.1} \cdot \frac{73}{3} - 0}{100 \cdot 0.6} = \frac{73}{2 \cdot 11 \cdot 6} = \frac{73}{196} \approx 0.37 \end{aligned}$$

2.3 Martingalmaß und Arbitrage im CRR-Modell

bisher: CRR-Modell auf einem Wahrscheinlichkeitsraum $(\Omega, \mathcal{F}, \mathbb{P})$

jetzt: weiteres Wahrscheinlichkeitsmaß \mathbb{Q} auf (Ω, \mathcal{F}) .

Das bedeutet nun, dass wir die Baumstruktur unverändert lassen, jedoch die Übergangswahrscheinlichkeiten ändern von $p = \mathbb{P}(R_n = b)$ zu $q = \mathbb{Q}(R_n = b)$. Dabei bleibt die Unabhängigkeit der R_n erhalten.

Notation: Wir bezeichnen mit $\mathbb{E}^{\mathbb{Q}}[\cdot]$ den Erwartungswert unter \mathbb{Q} .

Definition. Ein Wahrscheinlichkeitsmaß \mathbb{Q} auf (Ω, \mathcal{F}) heißt äquivalentes Martingalmaß (EMM) für das CRR-Modell, wenn gilt

- (1) $\mathbb{Q} \sim \mathbb{P}$ (\mathbb{Q} äquivalent zu \mathbb{P})
- (2) diskontierter Preisprozess $(\tilde{S}_n)_{n \in [N]}$ ist ein \mathbb{Q} -Martingal, d.h.

$$\mathbb{E}^{\mathbb{Q}}[\tilde{S}_{n+1} \mid \mathcal{F}_n] = \tilde{S}_n \quad \forall n \in [N-1]_0$$

Bemerkung. Seien \mathbb{P}, \mathbb{Q} Wahrscheinlichkeitsmaß auf (Ω, \mathcal{F}) .

- Äquivalenz: $\mathbb{P} \sim \mathbb{Q} : \Leftrightarrow (\mathbb{P}(A) = 0 \Leftrightarrow \mathbb{Q}(A) = 0 \quad \forall A \in \mathcal{F})$
- Absolutstetigkeit: $\mathbb{Q} \ll \mathbb{P} : \Leftrightarrow (\mathbb{P}(A) = 0 \Rightarrow \mathbb{Q}(A) = 0 \quad \forall A \in \mathcal{F})$
- Es gilt $\mathbb{Q} \sim \mathbb{P} \Leftrightarrow (\mathbb{Q} \ll \mathbb{P} \wedge \mathbb{P} \ll \mathbb{Q})$

Theorem 2.6

- (a) Im CRR-Modell existiert ein EMM genau dann, wenn $a < r < b$ gilt.
- (b) Das EMM \mathbb{Q} ist eindeutig und es gilt

$$q = \mathbb{Q}(R_n = b) = \frac{r - a}{b - a} \quad \text{und} \quad 1 - q = \mathbb{Q}(R_n = a) = \frac{b - r}{b - a} \quad \forall n \in [N]$$

Bemerkung. q und $1 - q$ sind genau die risikoneutralen Gewichte, die in (Rek) auftauchen.

Beweis. Sei \mathbb{Q} ein beliebiges Wahrscheinlichkeitsmaß auf (Ω, \mathcal{F}) . Setze $q_n := \mathbb{Q}(R_n = a \mid \mathcal{F}_{n-1}) \in [0, 1]$.

$$\begin{aligned} \mathbb{E}^{\mathbb{Q}}[\tilde{S}_n \mid \mathcal{F}_{n-1}] &= \mathbb{E}^{\mathbb{Q}}\left[\tilde{S}_{n-1} \cdot \left(\frac{1 + R_n}{1 + r}\right) \mid \mathcal{F}_{n-1}\right] \\ &= \tilde{S}_{n-1} \cdot \frac{1}{1 + r} \cdot \mathbb{E}^{\mathbb{Q}}[1 + R_n \mid \mathcal{F}_{n-1}] \\ &= \tilde{S}_{n-1} \cdot \frac{1}{1 + r} (q_n(1 + b) + (1 - q_n)(1 + a)) \end{aligned}$$

Somit gilt

$$\begin{aligned}
 (\tilde{S}_n)_{n \in [N]} \text{ ist } \mathbb{Q}\text{-Martingal} &\Leftrightarrow \frac{1}{q+r} (q_n(1+b) + (1-q_n)(1+a)) = 1 \\
 &\Leftrightarrow q_nb + (1-q_n)a = r \\
 &\Leftrightarrow q_n(b-a) = r-a \\
 &\Leftrightarrow q_n = \frac{r-a}{b-a}
 \end{aligned}$$

Es gilt $q_n \in [0, 1]$ genau dann, wenn $a \leq r \leq b$.

$\mathbb{Q} \sim \mathbb{P}$: $q_n \in (0, 1) \Leftrightarrow a < r < b$. Somit folgt schließlich: \mathbb{Q} ist EMM $\Leftrightarrow a < r < b$. \square

Theorem 2.7 (Risikoneutrale Bewertungsformel)

Sei $C = h(S_1, \dots, S_N)$ ein Derivat im CRR-Modell mit EMM \mathbb{Q} . Für den Preisprozess $(\Pi_n)_{n \in [N]}$ von C gilt

$$\Pi_n = (1+r)^{-(N-n)} \cdot \mathbb{E}^{\mathbb{Q}}[C \mid \mathcal{F}_n]$$

Insbesondere gilt

$$w = \Pi_0 = (1+r)^{-N} \cdot \mathbb{E}^{\mathbb{Q}}[C]$$

In Worten: Der faire Preis von C ist eindeutig und gegeben durch den diskontierten Erwartungswert von C unter dem Martingalmaß \mathbb{Q} .

Beweis. Der Wahrscheinlichkeitsraum für das CRR-Modell ist endlich, d.h. $|\Omega| = 2^N < \infty$. Somit ist jede Zufallsvariable beschränkt, insbesondere C und $(\xi_n)_{n \in [N]}$. Sei $(\xi_n)_{n \in [N]}$ eine Replikationsstrategie für C mit diskontiertem Werteprozess $(\tilde{\Pi}_n)$, d.h.

$$\begin{aligned}
 \tilde{\Pi}_n &= w + \sum_{k=1}^n \xi_k (\tilde{S}_k - \tilde{S}_{k-1}) = w + (\xi \bullet \tilde{S})_n \\
 \tilde{\Pi}_N &= (1+r)^{-N} C
 \end{aligned}$$

\mathbb{Q} ist EMM, also ist (\tilde{S}_n) ein \mathbb{Q} -Martingal. Mit Theorem 1.6 ist demnach auch $(\xi \bullet \tilde{S})_n$ ein \mathbb{Q} -Martingal, also ist ebenso $\tilde{\Pi}_n$ ein \mathbb{Q} -Martingal.

$$\Pi_n = (1+r)^n \cdot \tilde{\Pi}_n = (1+r)^n \cdot \mathbb{E}^{\mathbb{Q}}[\tilde{\Pi}_N \mid \mathcal{F}_n] = (1+r)^{-(N-n)} \cdot \mathbb{E}^{\mathbb{Q}}[C \mid \mathcal{F}_n] \quad \square$$

Bemerkung (Martingalbedingung für \mathbb{Q}). Wir schreiben (etwas umständlich) $q_b = \mathbb{Q}(R_n = b)$ und $q_a = \mathbb{Q}(R_n = a)$. Da \mathbb{Q} ein Wahrscheinlichkeitsmaß sein soll, muss $q_a + q_b = 1$, oder äquivalent dazu $q_b(1+r) + q_a(1+r) = 1+r$ gelten. Die Martingalbedingung liefert $(1+b)q_b + (1+a)q_a = 1+r$. Beide Gleichungen zusammen sind äquivalent zu $q_b(b-r) + q_a(a-r) = 0$. Als lineares Gleichungssystem geschrieben ergibt sich als

$$\begin{pmatrix} 1+r & 1+r \\ b-r & a-r \end{pmatrix} \cdot \begin{pmatrix} q_b \\ q_a \end{pmatrix} = \begin{pmatrix} 1+r \\ 0 \end{pmatrix} \quad (\text{LGS-2})$$

Vergleicht man dies mit (LGS-1), dann stellt man fest, dass die Matrizen gleiche Einträge haben, aber transponiert zueinander sind. (LGS-1) ist also eine Bedingung für die Replikationsstrategie, (LGS-2) dagegen eine Bedingung für äquivalente Martingalmaße. Multinomial würden Spalten hinzukommen, also wäre das LGS unterbestimmt und es gäbe mehrere EMM.

Bemerkung. Das äquivalente Martingalmaß ist genau dann eindeutig, wenn ein vollständiges Modell vorliegt.

2.3.1 Arbitrage im CRR-Modell

Definition. Eine Anlagestrategie $(\xi_n)_{n \in [N]}$ mit Zeithorizont N und diskretem Werteprozess $(\tilde{\Pi}_n)_{n \in [N]}$ heißt Arbitrage, wenn gilt

- (1) $\tilde{\Pi}_0 = 0$ (kein Anfangskapital)
 - (2) $\mathbb{P}(\tilde{\Pi}_N \geq 0) = 1$ (kein Verlustrisiko)
 - (3) $\mathbb{P}(\tilde{\Pi}_N > 0) > 0$ (positiver Gewinn mit positiver Wahrscheinlichkeit)
- (Arb)

Theorem 2.8

Im CRR-Modell sind äquivalent:

- (1) Es existiert keine Arbitrage (NA = "No Arbitrage").
- (2) Es existiert ein EMM \mathbb{Q} .

Bemerkung. Dieser Satz gilt im Wesentlichen in allen Finanzmarktmodellen und heißt dann auch "Erster Hauptsatz der Preistheorie".

Beweis. (\Leftarrow) via Widerspruch. Sei \mathbb{Q} ein EMM und es existiere eine Arbitrage (ξ_n) . Wegen $\mathbb{Q} \sim \mathbb{P}$ folgt (Arb):

$$\begin{aligned}\mathbb{Q}(\tilde{\Pi}_N \geq 0) &= 1 \\ \mathbb{Q}(\tilde{\Pi}_N > 0) &> 0\end{aligned}$$

Aus beiden Bedingungen folgt $\mathbb{E}^{\mathbb{Q}}[\tilde{\Pi}_N] > 0$ (\star). Andererseits gilt $\tilde{\Pi}_N = 0 + (\xi \bullet \tilde{S})_N$. Da \tilde{S} ein \mathbb{Q} -Martingal ist, ist auch $(\xi \bullet \tilde{S})$ ein \mathbb{Q} -Martingal. Dann gilt $\mathbb{E}^{\mathbb{Q}}[\tilde{\Pi}_N] = \mathbb{E}^{\mathbb{Q}}[(\xi \bullet \tilde{S})_N] = 0$ im Widerspruch zu (\star). \square

Kapitel 3

DAS BLACK-SCHOLES-MODELL

- Ziel: Übergang vom CRR-Modell (zeitdiskret) zum Black-Scholes-(BS-) Modell (zeitstetig) durch Grenzwertbildung.
- Herleitung der Black-Scholes-Formel für Preise von europäischen Put- und Call-Optionen.

Wir betrachten ein Zeitintervall $[0, T]$. Für jedes $n \in \mathbb{N}$ geteilt in Schritte der Länge $\Delta_N := \frac{T}{N}$. Wähle Parameter $r \in \mathbb{R}$, $\mu \in \mathbb{R}$ (Trendparameter), $\sigma > 0$ (Volatilität). Definiere eine Folge von CRR-Modellen $(S^N)_{n \in \mathbb{N}}$ eingebettet in $[0, T]$ mit Parametern $r_N = r \cdot \Delta_N$

$$b_N := \mu \Delta_N + \sigma \sqrt{\Delta_N} \quad a_N := \mu \Delta_N - \sigma \sqrt{\Delta_N} \quad p \in (0, 1) \quad s > 0$$

d.h.

$$S_0^N = s, \quad S_{t_k}^N = s \cdot \prod_{i=1}^k (1 + R_i^N) \quad \text{mit } t_k = k \cdot \Delta_N \quad (\text{CRR-N})$$

bzw.

$$\tilde{S}_0^N = s, \quad \tilde{S}_{t_k}^N = s \cdot \prod_{i=1}^k \frac{1 + R_i^N}{1 + r_N}$$

wobei

$$\mathbb{P}(R_i^N = b_N) = p \quad \text{und} \quad \mathbb{P}(R_i^N = a_N) = 1 - p$$

Falls notwendig, interpolieren wir zwischen den Gitterpunkten mit $S_t^N = S_{t_k}^N$ für $t \in [t_k, t_{k+1}]$ (d.h. konstante Interpolation).

Berechne risikoneutrale Wahrscheinlichkeiten

$$q_N = \mathbb{Q}_N(R_i^N = b_N) = \frac{r_N - a_N}{b_N - a_N} = \frac{(r - \mu)\Delta_N + \sigma\sqrt{\Delta_N}}{2\sigma\sqrt{\Delta_N}} = \frac{1}{2} - \frac{\lambda}{2}\sqrt{\Delta_N}$$

mit $\lambda := \frac{\mu - r}{\sigma}$.

Bemerkung

- Für $\mu = r$ gilt $q_N = \frac{1}{2}$, im Allgemeinen $\lim_{N \rightarrow \infty} q_N = \frac{1}{2}$.
- $\lambda = \frac{\mu - r}{\sigma}$ heißt "Sharpe-Ratio" oder Marktrisikopreis. (großes λ heißt entweder hoher Ertrag oder geringes Risiko)

Frage: Konvergenz der Verteilung von S_T^N unter \mathbb{Q}_N für $N \rightarrow \infty$?

Übergang zum Logarithmus:

$$Z_N := \log \left(\frac{S_T^N}{S_0} \right) = \sum_{k=1}^N \underbrace{\log(1 + R_k^N)}_{=: L_k^N}$$

Summe von unabhängigen, identisch verteilten Zufallsvariablen \Rightarrow Zentraler Grenzwertsatz?

Es liegt ein sogenanntes Dreiecksschema vor:

$$\begin{aligned} Z_1 &= L_1^1 \\ Z_2 &= L_1^2 + L_2^2 \\ Z_3 &= L_1^3 + L_2^3 + L_3^3 \\ &\vdots \end{aligned}$$

Die Zufallsvariablen in einer Zeile sind unabhängige Zufallsvariablen.

Theorem 3.1 (CLT für Dreiecksschemata)

Sei für jedes $N \in \mathbb{N}$ ein Vektor $L^N := (L_1^N, L_2^N, \dots, L_N^N)$ von Zufallsvariablen gegeben ("Dreiecksschema") mit folgenden Eigenschaften:

- (a) für alle $N \in \mathbb{N}$ sind die $(L_1^N, L_2^N, \dots, L_N^N)$ unabhängig mit identischer Verteilung
- (b) es existiert eine Folge von (deterministischen) Konstanten $c_N \rightarrow 0$, sodass $|L_k^N| \leq c_N$ für alle $k \in [N]$.
- (c) Mit $Z_N := L_1^N + \dots + L_N^N$ gilt $\mathbb{E}[Z_N] \rightarrow m \in \mathbb{R}$ sowie $\text{Var}(Z_N) \rightarrow s^2 > 0$ jeweils für $N \rightarrow \infty$.

Dann konvergiert $(Z_N)_{N \in \mathbb{N}}$ in Verteilung gegen eine normalverteilte Zufallsvariable Z mit $\mathbb{E}[Z] = m$ und $\text{Var}(Z) = s^2$.

Beweis. siehe z.B. "Wahrscheinlichkeitstheorie mit Martingalen" □

Bemerkung. Die Dichte der Standardnormalverteilung ist gegeben durch

$$\varphi(x) := \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$$

und für die Verteilungsfunktion gilt

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

Eine Normalverteilung mit Erwartungswert m und Varianz s^2 hat Verteilungsfunktion $\Phi\left(\frac{x-m}{s}\right)$.

Definition. Eine strikt positive Zufallsvariable X heißt **lognormalverteilt** mit Parametern m und s^2 , wenn gilt $\log(X) \sim \mathcal{N}(m, s^2)$.

Theorem 3.2

Betrachte eine Folge $(S^N)_{N \in \mathbb{N}}$ von CRR-Modellen wie in (CRR-N) beschrieben. Dann konvergiert S_T^N unter \mathbb{Q} in Verteilung gegen eine Zufallsvariable S_T und $\frac{S_T}{S_0}$ ist lognormalverteilt mit Parametern $m = T \left(r - \frac{\sigma^2}{2} \right)$ und $s^2 = T\sigma^2$. Äquivalent dazu gilt mit $Z_N = \log \left(\frac{S_T^N}{S_0} \right)$

$$\mathbb{Q}_N(Z_N \leq x) \xrightarrow{N \rightarrow \infty} \Phi \left(\frac{x - T \left(r - \frac{\sigma^2}{2} \right)}{\sigma \sqrt{T}} \right)$$

Beweis. Das Dreiecksschema $L^N = (L_1^N, \dots, L_N^N)$ mit $L_k^N = \log(1 + R_k^N)$ erfüllt (unter \mathbb{Q}_N) offensichtlich Bedingungen (a) und (b) aus Theorem 3.1. Für (b) wähle z.B.

$$c_N = \max \left\{ \left| \log(1 + \mu \Delta_N + \sigma \sqrt{\Delta_N}) \right|, \left| \log(1 + \mu \Delta_N - \sigma \sqrt{\Delta_N}) \right| \right\}$$

Wir berechnen Erwartungswert und Varianz von L_k^N bzw. Z_N . Verwende Taylorentwicklung des Logarithmus um die Entwicklungsstelle 1:

$$\log(1 + x) = x - \frac{x^2}{2} + \frac{x^3}{3} + \mathcal{O}(x^4) \quad (x \rightarrow 0)$$

Das heißt

$$\log(1 + \mu \Delta_N \pm \sigma \sqrt{\Delta_N}) = \pm \sigma \sqrt{\Delta_N} + \mu \Delta_N - \frac{\sigma^2}{2} \Delta_N + \mathcal{O}(\Delta_N^{3/2})$$

Für die risikoneutralen Wahrscheinlichkeiten gilt $q_N = \frac{1}{2} + \frac{\lambda}{2} \sqrt{\Delta_N}$ und $1 - q_N = \frac{1}{2} - \frac{\lambda}{2} \sqrt{\Delta_N}$.

$$\begin{aligned} \mathbb{E}^{\mathbb{Q}}[L_k^N] &= \mathbb{E}^{\mathbb{Q}_N}[\log(1 + R_k^N)] = q_N \cdot \log(1 + b_N) + (1 - q_N) \log(1 + a_N) \\ &= \left(\mu - \frac{\sigma^2}{2} \right) \Delta_N - \lambda \sigma \Delta_N + \mathcal{O}(\Delta_N^{3/2}) \\ &= \left(\mu - (\mu - r) - \frac{\sigma^2}{2} \right) \Delta_N + \mathcal{O}(\Delta_N^{3/2}) \\ &= \left(r - \frac{\sigma^2}{2} \right) \Delta_N + \mathcal{O}(\Delta_N^{3/2}) \\ \mathbb{E}^{\mathbb{Q}_N}[(L_k^N)^2] &= q_N \log^2(1 + b_N) + (1 - q_N) \log^2(1 + a_N) \\ &= \sigma^2 \Delta_N + \mathcal{O}(\Delta_N^{3/2}) \\ \text{Var}^{\mathbb{Q}_N}(L_k^N) &= \mathbb{E}^{\mathbb{Q}_N}[(L_k^N)^2] - \mathbb{E}^{\mathbb{Q}_N}[L_k^N]^2 \\ &= \sigma^2 \Delta_N + \mathcal{O}(\Delta_N^{3/2}) \end{aligned}$$

Also gilt

$$\begin{aligned} \mathbb{E}^{\mathbb{Q}_N}[Z_N] &= N \cdot \mathbb{E}^{\mathbb{Q}_N}[L_k^N] = \left(r - \frac{\sigma^2}{2} \right) T + \mathcal{O}(N^{-1/2}) \xrightarrow{N \rightarrow \infty} \left(r - \frac{\sigma^2}{2} \right) T = m \\ \text{Var}^{\mathbb{Q}_N}[Z_N] &= N \cdot \text{Var}^{\mathbb{Q}_N}[L_k^N] = \sigma^2 T + \mathcal{O}(N^{-1/2}) \xrightarrow{N \rightarrow \infty} \sigma^2 T = s^2 \end{aligned}$$

Das Resultat folgt nun aus dem Zentralen Grenzwertsatz (Theorem 3.1). \square

3.1 Asymptotik von Put- und Call-Preisen

Wir fixieren die Laufzeit T und den Ausübungspreis K und schreiben $C_N(t, S_t^N)$ für den Preis einer europäischen Call-Option im (CRR-N)-Modell in Abhängigkeit von Zeit t und Basisgut S_t^N . Analog gilt dies auch für den Put-Preis $P_N(t, S_t^N)$.

Theorem 3.3 (Black-Scholes-Formel)

Die Preise C_N und P_N konvergieren für $N \rightarrow \infty$ gegen den Black-Scholes-Preis

$$C_{BS}(t, S_t) := \lim_{N \rightarrow \infty} C_N(t, S_t^N)$$

$$P_{BS}(t, S_t) := \lim_{N \rightarrow \infty} P_N(t, S_t^N)$$

und es gilt die Black-Scholes-Formel

$$C_{BS}(t, S_t) = S_t \Phi(d_1) - e^{-r(T-t)} K \Phi(d_2)$$

$$P_{BS}(t, S_t) = -S_t \Phi(-d_1) + e^{-r(T-t)} K \Phi(-d_2)$$

wobei

$$d_1 = d_1(t, S_t) = \frac{\log\left(\frac{S_t}{K}\right) + \left(r + \frac{\sigma^2}{2}\right)(T-t)}{\sigma\sqrt{T-t}}$$

$$d_2 = d_2(t, S_t) = \frac{\log\left(\frac{S_t}{K}\right) + \left(r - \frac{\sigma^2}{2}\right)(T-t)}{\sigma\sqrt{T-t}} = d_1 - \sigma\sqrt{T-t}$$

Bemerkung

- geschlossener Ausdruck für Bewertung von europäischen Put- und Call-Optionen
- Herleitung als Grenzwert aus dem CRR-Modell entspricht nicht der ursprünglichen Herleitung von Black und Scholes mittels stochastischer Analysis (siehe Vorlesung "Stochastic Calculus")
- Für Entwicklung der BS-Formel und des BS-Modells erhielten **SCHOLES** und **Merton** den Wirtschaftsnobel(gedenk)preis 1997
- Der Parameter σ heißt **Volatilität** und entspricht der Schwankungsbreite der Preisänderungen

Bemerkung

- innerer Wert $(S_t - K)_+$ konvergiert gegen Auszahlungsprofil $(S_T - K)_+$ für $t \rightarrow T$
- Zeitwert $C_{BS}(t, S_t) - (S_t - K)_+ \geq 0$ konvergiert gegen Null für $t \rightarrow T$
- "out-of-the-money" (OTM): innerer Wert = 0 bzw. $S_t < K$
- "in-the-money" (ITM): innerer Wert > 0 bzw. $S_t > K$
- "at-the-money" (ATM): Grenzfall $S_t = K$
- Zeitwert ist am größten für ATM-Optionen
- $t \mapsto C_{BS}(t, S_t)$ ist streng monoton fallend bzw. $\frac{\partial}{\partial t} C_{BS}(t, S_t) < 0$
- $S_t \mapsto C_{BS}(t, S_t)$ ist streng monoton wachsend und konvex bzw. $\frac{\partial}{\partial S} C_{BS}(t, S_t) > 0$ und $\frac{\partial^2}{\partial S^2} C_{BS}(t, S_t) > 0$

Bemerkung

- innerer Wert $(K - S_t)_+$ konvergiert gegen Auszahlungsprofil $(K - S_T)_+$ für $t \rightarrow T$

- Zeitwert $P_{BS}(t, S_t) - (K - S_t)_+ \geq 0$ konvergiert gegen Null für $t \rightarrow T$
- "out-of-the-money" (OTM): innerer Wert = 0 bzw. $S_t > K$
- "in-the-money" (ITM): innerer Wert > 0 bzw. $S_t < K$
- "at-the-money" (ATM): Grenzfall $S_t = K$
- Zeitwert ist am größten für ATM-Optionen
- $t \mapsto P_{BS}(t, S_t)$ ist streng monoton fallend bzw. $\frac{\partial}{\partial t} P_{BS}(t, S_t) < 0$
- $S_t \mapsto P_{BS}(t, S_t)$ ist streng monoton fallend und konvex bzw. $\frac{\partial}{\partial S} P_{BS}(t, S_t) < 0$ und $\frac{\partial^2}{\partial S^2} P_{BS}(t, S_t) > 0$

Beweis (Theorem 3.3). Wir beweisen das Resultat für $t = 0$, andere Zeitpunkte $t \in [0, T]$ können analog mittels konstanter Interpolation behandelt werden. Nach ?? gilt für den Preis der Put-Option im CRR_N-Modell

$$\begin{aligned} P^N(0, S_0^N) &= (1 + r\Delta_N)^{-N} \cdot \mathbb{E}^{\mathbb{Q}_N} [(K - S_T^N)_+] = (1 + r\Delta_N)^{-N} \cdot \mathbb{E}^{\mathbb{Q}_N} [(K - S_0 e^{Z_N})_+] \\ &= (1 + r\Delta_N)^{-N} \cdot \mathbb{E}^{\mathbb{Q}_N} [f(Z_N)] \text{ mit } f(z)(K - S_0 e^z)_+ \text{ stetig und beschränkt} \end{aligned}$$

wobei $Z_N = \log\left(\frac{S_T^N}{S_0}\right)$ gilt. Aus der Stochastik ist bekannt, dass aus $Z_N \xrightarrow{d} Z$ in Verteilung bereits $\mathbb{E}[f(Z_N)] = \mathbb{E}[f(Z)]$ folgt für alle $f \in C_b(\mathbb{R})$.

Weiterhin gilt $\lim_{N \rightarrow \infty} (1 + r\Delta_N)^{-N} = \lim_{N \rightarrow \infty} (1 + rT/N)^{-N} = e^{-rt}$.

$\lim_{N \rightarrow \infty} \mathbb{E}^{\mathbb{Q}_N} [f(Z_N)] = \mathbb{E} [f(Z)]$ nach Theorem 3.1 mit $Z \sim \mathcal{N}((r - \frac{\sigma^2}{2})T, \sigma^2 T)$. Im Folgenden schreiben wir $m := (r - \frac{\sigma^2}{2})T$.

$$\begin{aligned} \mathbb{E} [f(Z)] &= \frac{1}{\sqrt{2\pi}} \frac{1}{6\sqrt{T}} \int_{-\infty}^{\infty} (K - S_0 e^z)_+ \cdot \exp\left(-\frac{(z - mT)^2}{2\sigma^2 T}\right) dz \\ &= \frac{1}{\sqrt{2\pi}} \frac{1}{6\sqrt{T}} \int_{-\infty}^{\log\left(\frac{K}{S_0}\right)} (K - S_0 e^z) \cdot \exp\left(-\frac{1}{2} \left(\frac{z - mT}{\sigma\sqrt{T}}\right)^2\right) dz \\ &= \left[\begin{array}{l} y = \frac{z - mT}{\sigma\sqrt{T}} \\ dy = \frac{dz}{\sigma\sqrt{T}} \end{array} \right] \\ &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{-d_2} (K - S_0 \cdot \exp(y\sigma\sqrt{T} + mT)) \cdot e^{-\frac{y^2}{2}} dy \\ &= K\Phi(-d_2) - S_0 \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{-d_2} \exp\left(-\frac{y^2}{2} + y\sigma\sqrt{T} + mT\right) dy \end{aligned}$$

Nebenrechnung liefert

$$-\frac{y^2}{2} + y\sigma\sqrt{T} + mT = rT - \frac{1}{2}(y^2 - 2y\sigma\sqrt{T} + \sigma^2 T) = rT - \frac{1}{2}(y - \sigma\sqrt{T})^2$$

$$\begin{aligned}
 K\Phi(-d_2) - S_0 e^{rT} \underbrace{\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{-d_2} e^{-\frac{(y-\sigma\sqrt{T})^2}{2}} dy}_{=\Phi(-d_2-\sigma\sqrt{T})} &= K\Phi(-d_2) - S_0 e^{rT} \Phi(-d_1) \\
 \Rightarrow \lim_{N \rightarrow \infty} P_N(0, S_0^N) &= e^{-rT} K\Phi(-d_2) - S_0 \Phi(-d_1)
 \end{aligned}$$

Für den Call bedienen wir uns der Put-Call-Parität:

$$C^N(0, S_0^N) - \underbrace{P^N(0, S_0^N)}_{\rightarrow P_{BS}(0, S_0)} = \underbrace{S_0^N}_{=S_0} - \underbrace{(1+r\Delta_N)^{-N} \cdot K}_{\rightarrow e^{-rT} \cdot K}$$

Somit gilt mit Grenzübergang

$$\begin{aligned}
 C_{BS}(0, S_0) &= \lim_{N \rightarrow \infty} C^N(0, S_0) \\
 &= P_{BS}(0, S_0) + S_0 - e^{-rT} \cdot K = e^{-rT} \cdot K \cdot \underbrace{(\Phi(-d_2) - 1)}_{=-\Phi(d_2)} - S_0 \underbrace{(\Phi(-d_1) - 1)}_{=-\Phi(d_1)} \\
 &= S_0 \Phi(d_1) - e^{-rT} \cdot K \cdot \Phi(d_2)
 \end{aligned}$$

was genau der Black-Scholes-Formel für den Call entspricht. □

Wir haben gezeigt, dass die CRR_N -Preise gegen BS-Preise konvergieren.

Frage: Was gilt für die Replikationsstrategie? Konvergiert auch diese?

Theorem 3.4

Für die Replikationsstrategie $\xi_{t_N}^N$ der Put- bzw. Call-Option im CRR_N -Modell gilt:

$$\begin{aligned}
 \lim_{N \rightarrow \infty} \xi_{t_N}^N &= \frac{\partial}{\partial S} P_{BS}(t, S_t) = -\Phi(-d_1) \\
 \lim_{N \rightarrow \infty} \xi_{t_N}^N &= \frac{\partial}{\partial S} C_{BS}(t, S_t) = \Phi(d_1)
 \end{aligned}$$

Diese partiellen Ableitungen heißen auch "Delta" des Put- bzw. Call-Preises.

Beweis. Wir betrachten den Zeitpunkt $t = 0$, $t \in [0, T]$ kann analog betrachtet werden. Nach ?? ist ξ_0^N für den Put gegeben durch

$$\begin{aligned}
 \xi_0^N &= \frac{P_N(\Delta_N, S_0(1+b_N)) - P_N(\Delta_N, S_0(1+a_N))}{S_0(b_N - a_N)} \\
 &= \frac{P_N(\Delta_N, S_0(1+\mu\Delta_N + \sigma\sqrt{\Delta_N})) - P_N(\Delta_N, S_0(1+\mu\Delta_N - \sigma\sqrt{\Delta_N}))}{2 \cdot S_0 \sigma \sqrt{\Delta_N}}
 \end{aligned}$$

Es gilt $\lim_{N \rightarrow \infty} P_N(\Delta_N, S_0(1+\mu\Delta_N)) = P_{BS}(0, S_0)$. Unter geeigneten Annahmen an gleichmäßige Konvergenz folgt

$$\lim_{N \rightarrow \infty} \xi_0^N = \frac{\partial}{\partial S} P_{BS}(t, S_t)$$

und analog zeigt man dies auch für den Call. Wir berechnen $\frac{\partial}{\partial S} C_{BS}$ explizit:

$$\begin{aligned}\frac{\partial}{\partial S} C_{BS}(t, S) &= \Phi(d_1) + S\varphi(d_1) \cdot \frac{\partial}{\partial S} d_1 - e^{-r(T-t)} K\varphi(d_2) \cdot \frac{\partial}{\partial d_1} d_2 \\ &= \Phi(d_1) + \frac{\partial}{\partial S} d_1 \left(S\varphi(d_1) - e^{-r(T-t)} K\varphi(d_2) \right)\end{aligned}$$

Nebenrechnung: Setze $\tau = T - t$.

$$\begin{aligned}e^{-r\tau} \frac{K}{S} \varphi(d_2) &= \frac{1}{\sqrt{2\pi}} e^{-r\tau} \frac{K}{S} \exp \left(-\frac{1}{2} \left(\frac{\log \left(\frac{S}{K} \right) + r\tau - \frac{\sigma^2 \tau}{2}}{\sigma \tau} \right)^2 \right) \\ &= \frac{1}{\sqrt{2\pi}} e^{-r\tau} \frac{K}{S} \exp \left(-\frac{1}{2} \left(\frac{(\log \left(\frac{S}{K} \right) + r\tau)^2}{\sigma^2 \tau} \right) - 2 \frac{1}{2} \left(\log \left(\frac{S}{K} \right) - r\tau \right) + \frac{\sigma^2 \tau}{4} \right) \\ &= \frac{1}{\sqrt{2\pi}} \exp \left(-\frac{1}{2} \left(\frac{(\log \left(\frac{S}{K} \right) + r\tau)^2}{\sigma^2 \tau} \right) + \left(\log \left(\frac{S}{K} \right) + r\tau \right) + \frac{\sigma^2 \tau}{4} \right) \\ &= \varphi(d_1)\end{aligned}$$

d.h. $e^{-r(T-t)} K\varphi(d_2) = S\varphi(d_1)$. Somit gilt

$$\frac{\partial}{\partial S} C_{BS}(t, S) = \Phi(d_1)$$

Die Gleichung für den Put folgt analog oder mithilfe der Put-Call-Parität. \square

Bemerkung

- Die Ableitungen $\frac{\partial}{\partial S} C_{BS}$ bzw. $\frac{\partial}{\partial S} P_{BS}$ lassen sich auch interpretieren als Sensitivität des Call- bzw. Put-Preises gegenüber Preisänderungen des Basisguts.

Analog lassen sich die Sensitivitäten ("Greeks") nach den weiteren Parametern berechnen.

Definition. Die Greeks des BS-Preises sind folgende partielle Ableitungen:

Bezeichnung	Def.	Wert Call	Wert Put	Bemerkungen
Delta	$\frac{\partial}{\partial S}$	$\Phi(d_1)$	$-\Phi(-d_1)$	bestimmt die Replikations- bzw.
Gamma	$\frac{\partial^2}{\partial S^2}$	$\frac{\varphi(d_1)}{S_t \sigma \sqrt{T-t}}$		Sensitivität von Delta gegenüber
Vega	$\frac{\partial}{\partial \sigma}$	$S_t \sqrt{T-t} \varphi(d_1)$		Sensitivität gegenüber Änderung
Theta	$\frac{\partial}{\partial t}$	siehe Übung		Änderung in der Zeit
Rho	$\frac{\partial}{\partial r}$	$K(T-t)e^{-r(T-t)}\Phi(d_2)$	$-K(T-t)e^{-r(T-t)}\Phi(-d_2)$	Sensitivität gegenüber Änderung

Korollar 3.5

Der BS-Preis $C_{BS}(t, S_t)$ erfüllt die folgende partielle Differentialgleichung

$$\frac{\partial}{\partial t} C_{BS} + rS \frac{\partial}{\partial S} C_{BS} + \frac{\sigma^2}{2} S^2 \frac{\partial^2}{\partial S^2} C_{BS} - rC_{BS} = 0 \quad (\text{BS-PDE})$$

auf $(t, s) \in [0, T) \times \mathbb{R}_{\geq 0}$ mit der Endwertbedingung $\lim_{t \rightarrow T} C_{BS}(t, S) = (S - K)_+$. Für den Put P_{BS} gilt die gleiche PDE mit Endwertbedingung $\lim_{t \rightarrow T} P_{BS}(t, S) = (K - S)_+$.

Beweis. siehe Übung □

Bemerkung. In Erweiterungen des Black-Scholes-Modells gibt es keine geschlossenen Ausdrücke für Put- bzw. Call-Preise, aber eine partielle Differentialgleichung ähnlich zur (BS-PDE) gilt weiterhin.

3.2 Implizite Volatilität & Grenzen des BS-Modells

Wir schreiben etwas ausführlicher

$$C_{BS}(t, S_t; T, K, \sigma) := C_{BS}(t, S_t)$$

um die Abhängigkeit von T , K und σ zu verdeutlichen.

Theorem 3.6 (Implizite Volatilität)

Sei $C_*(0, S_0, T, K)$ ein vorgegebener (beobachteter) Preis einer Call-Option mit Fälligkeit T , Ausübungspreis K , welcher innerhalb der Arbitragegrenzen liegt, d.h. $(S_0 - e^{-rT}K)_+ < C_*(0, S_0, T, K) < S_0$. Dann existiert ein eindeutiges $\sigma_*(T, K) \in (0, \infty)$, die implizite Volatilität von C_* , sodass

$$C_*(0, S_0; T, K) = C_{BS}(0, S_0; T, K, \sigma_*(T, K))$$

gilt.

Bemerkung. $\sigma_*(T, K)$ ist Lösung eines inversen Problems:

- Vorwärtsproblem: Parameter \rightarrow Call-Preis
- inverses Problem: Call-Preis \rightarrow Parameter

Es kann zur empirischen Überprüfung des BS-Modells verwendet werden:

- BS-Modell passt gut zu Daten: $(T, K) \mapsto \sigma_*(T, K)$ ist annähernd konstant
- BS-Modell passt nicht gut zu Daten: $(T, K) \mapsto \sigma_*(T, K)$ variiert stark mit (T, K)

Typische tatsächliche Beobachtung

- konvex
- asymmetrisch (höher für große K)
- Minimum bei at-the-money
- flacher für lange Laufzeiten, steiler für kurze Laufzeiten

Die Form weist darauf hin, dass das Black-Scholes-Modell große Preissprünge des Basisguts unterschätzt.

Die Form des Vola-Smiles in Modellen jenseits des Black-Scholes-Modell ist ein aktuelles Forschungsthema.

Kapitel 4

OPTIMALE INVESTITION

4.1 Das Anlageproblem

Gegeben seien ein Vermögen w sowie Anlagegüter S^1, \dots, S^n (Aktien, Anleihen, ...). Gesucht ist nun eine optimale Verteilung $w = w_1 + \dots + w_n$ auf S^1, \dots, S^n .

Die Anlagegüter S^1, \dots, S^n weisen unterschiedliche Erträge, Risiken und typischerweise Korrelationen auf.

Wir unterscheiden:

- Einperiodenproblem: Aufteilung wird heute ($t = 0$) festgelegt und bis zum Zeithorizont $t = T$ beibehalten.
- Mehrperiodenproblem: Umschichten zu mehreren Zeitpunkten $\{t_0, t_1, \dots, t_N\}$ möglich.

Das einfachste Optimalitätsprinzip ist die **Pareto-Optimalität**:

- Bei gleichem Risiko wird Anlage mit größerem Ertrag bevorzugt.
- Bei gleichem Ertrag wird Anlage mit kleinerem Risiko bevorzugt.

d.h. **Pareto-optimal**: Es gibt keine Anlagestrategie mit größerem Ertrag und kleinerem Risiko.

Zum Aufwärmen betrachten wir zwei Toy-Models, also stark vereinfachte Beispiele:

Toy-Model 1: Einperiodenmodell, eine risikofreie und eine risikobehaftete Anlagemöglichkeit

Wir setzen den Zeithorizont auf $T = 1$. Für die risikofreie Anlage sei $S_0^0 = 1$ und $S_T^0 = (1+r)$. Für die risikobehaftete Anlage gelte $S_0^1 = 1$ und $S_T^1 = (1+R)$, wobei R stochastisch ist mit Erwartungswert $\mu = \mathbb{E}[R]$ (Ertrag) und Standardabweichung $\sigma = \sqrt{\text{Var}(R)} > 0$ (Risiko). Mit $s := \mu - r$ bezeichnen wir die Überrendite [excess return]. Ist $s \leq 0$, dann sollte alles in S^0 investiert werden (Pareto-optimal). Sei $s > 0$: teile das Vermögen w auf in zwei Teile $(w_0, w - w_0)$ auf (S^0, S^1) ¹. Dabei entspricht $w - w_0 < 0$ einem Leerverkauf und $w_0 < 0$ einer Kreditaufnahme. Gelte $w = 1$.

- Endvermögen: $P_T = w_0(1+r) + (1-w_0)(1+R)$
- zu erwartende Rendite²: $\mu_P = \mathbb{E}[P_T - 1] = w_0(1+r) + (1-w_0)(1+\mu) - 1 = w_0r + (1-w_0)\mu + 1$
- Risiko: $\sigma_P = (1-w_0) \cdot \sigma$
- Überrendite: $s_P = (1-w_0)(\mu - r)$

Somit ist jede Strategie Pareto-optimal und das Pareto-Prinzip hilft nicht bei der Auswahl.

¹investiere also w_0 in S^0 und den Rest in w_1

²Rendite = $\frac{P_T - P_0}{P_0}$

Insbesondere ist **Sharpe-Ratio**:

$$SR(w_0) = \frac{\text{„Überrendite“}}{\text{„Risiko“}} = \frac{s_p}{\sigma_p} = \frac{\mu - r}{\sigma} = \text{const.}$$

Alternative zum Pareto-Prinzip: Festlegen von individueller Risikoaversion (mehr dazu später)

Toy-Model II: Einperiodenproblem, zwei risikobehaftete Anlagemöglichkeiten, Zeithorizont $T = 1$, Vermögen $w = 1$ mit

$$\begin{aligned} S_0^1 &= 1 & S_T^1 &= (1 + R_1) \text{ mit } \mathbb{E}[R_1] = \mu, \text{Var}(R_1) = \sigma_1^2 \\ S_0^2 &= 1 & S_T^2 &= (1 + R_2) \text{ mit } \mathbb{E}[R_2] = \mu, \text{Var}(R_2) = \sigma_2^2 \end{aligned}$$

- Portfoliowert: $P_T = w_1(1 + R_1) + (1 - w_1)(1 + R_2)$
- Rendite: $\mu_P = \mathbb{E}[P_T - 1] = w_1\mathbb{E}[R_1] + (1 - w_1)\mathbb{E}[R_2] = \mu$
- Risiko: $\sigma_P^2 = \text{Var}(w_1R_1 + (1 - w_1)R_2) = w_1^2\sigma_1^2 + (1 - w_1)^2\sigma_2^2$
- Pareto-optimales Portfolio:

$$\begin{aligned} \frac{\partial}{\partial w_1} \sigma_P^2 &= 2w_1 \cdot \sigma_1^2 - 2(1 - w_1)\sigma_2^2 = 0 \\ \Rightarrow w_1(\sigma_1^2 + \sigma_2^2) &= \sigma_2^2 \\ \Rightarrow w_* &= \frac{\sigma_2^2}{\sigma_1^2 + \sigma_2^2} \in (0, 1) \end{aligned}$$

Somit existiert also genau eine Pareto-optimale Strategie.

Das Vermögen wird proportional zum Verhältnis der Risiken aufgeteilt. Außerdem wird das Vermögen *nicht* vollständig in risikoärmere Anlage gesteckt. Man nennt dies auch das **Diversifikationsprinzip**. w_* ist auch die Strategie mit maximaler Sharpe-Ratio.

4.2 Exkurs: Optimierung mit Nebenbedingungen

Betrachte das Optimierungsproblem

$$\min f_0(x) \quad (x \in \mathbb{R}^n) \tag{OPT}$$

unter Nebenbedingungen

$$\begin{cases} f_i(x) \leq 0 & i = 1, \dots, m \\ h_i(x) = 0 & i = 1, \dots, p \end{cases} \tag{NB}$$

Ein $x \in \mathbb{R}^n$, welches (NB) erfüllt, heißt **zulässig**, ein $x_* \in \mathbb{R}^n$, welches (OPT) minimiert, heißt (Optimal-)Lösung mit $p_* = f_0(x_*)$ als Minimalwert.

Definition. Die Funktion

$$\mathcal{L}(x, \lambda, \nu) = f_0(x) + \sum_{i=1}^m f_i(x) \lambda_i + \sum_{i=1}^p h_i(x) \nu_i$$

mit $\lambda \in \mathbb{R}_{\geq 0}^m$ und $\nu \in \mathbb{R}^p$ heißt **Lagrange-Zielfunktion** für (OPT).

Die Funktion

$$g(\lambda, \nu) := \inf_{x \in \mathbb{R}^n} \mathcal{L}(x, \lambda, \nu)$$

heißt (Lagrange-)duale Funktion für (OPT)

Bemerkung. Als Infimum von (in λ, ν) linearen Funktionen ist g konkav³. Die duale Funktion $g(\lambda, \nu)$ erzeugt eine untere Schranke für p_* . Begründung: Sei $\bar{x} \in \mathbb{R}^n$ zulässig für (OPT), d.h. $f_i(\bar{x}) \leq 0$ für alle $i \in [m]$ und $h_i(\bar{x}) = 0$ für alle $i \in [p]$. Somit ist

$$\mathcal{L} = f_0(\bar{x}) + \underbrace{\sum_{i=1}^m f_i(\bar{x}) \lambda_i + \sum_{i=1}^p h_i(\bar{x}) \nu_i}_{\leq 0} \leq f_0(\bar{x})$$

Also ist $g(\lambda, \nu) = \inf_{x \in \mathbb{R}^n} \mathcal{L}(x, \lambda, \nu) \leq \mathcal{L}(\bar{x}, \lambda, \nu) \leq f_0(\bar{x})$ für alle zulässigen \bar{x} . Somit ist $g(\lambda, \nu) \leq p_*$ für alle $\lambda \in \mathbb{R}_{\geq 0}^m$ und $\nu \in \mathbb{R}^p$. Die beste untere Schranke erhalten wir durch Maximieren über λ und ν .

Definition. Das duale Optimierungsproblem zu (OPT) ist

$$\max g(\lambda, \nu) \quad \lambda \in \mathbb{R}_{\geq 0}^m, \nu \in \mathbb{R}^p \quad (\text{D})$$

unter der Nebenbedingung $\lambda_i \geq 0$ für alle $i \in [m]$. Den Maximalwert bezeichnen wir mit d_* .

Zwischen (OPT) und (D) gilt **schwache Dualität**, d.h. $d_* \leq p_*$. Unter bestimmten Voraussetzungen gilt auch die **starke Dualität**, d.h. $d_* = p_*$.

Lemma 4.1 (Schwache Dualität)

Zwischen (OPT) und (D) gilt die schwache Dualität, d.h.

$$d_* \leq p_* \quad (\text{WD})$$

Bemerkung

- Die Differenz $p_* - d_* \geq 0$ heißt **Dualitätslücke** [duality gap].
- Wenn die Dualitätslücke verschwindet, dann spricht man von starker Dualität, d.h.

$$d_* = p_*$$

- Hinreichende Bedingungen für starke Dualität existieren vor allem für *konvexe* Probleme.

³Wir wissen, dass jede konvexe Funktion sich darstellen lässt als Supremum von (affin) linearen Funktionen.

Definition. Das Optimierungsproblem (OPT) heißt **konvex**, wenn f_0 konvex ist und die Menge der zulässigen Werte konvex ist. In diesem Fall kann (OPT) in folgende Form gebracht werden:

$$\min f_0(x) \text{ über } x \in \mathbb{R}^n \text{ unter NB } \begin{cases} f_i(x) \leq 0 & i = 1, \dots, m \\ Ax = b & i = 1, \dots, p \end{cases} \quad (\text{K-OPT})$$

mit f_0, f_1, \dots, f_m konvex, $A \in \mathbb{R}^{p \times m}$, $b \in \mathbb{R}^p$.

Theorem 4.2 (Slatters Bedingung)

Betrachte das konvexe Optimierungsproblem (K-OPT). Wenn $x \in \mathbb{R}^n$ existiert mit

$$f_i(x) < 0 \quad \text{für alle } i \in [m] \quad \text{und} \quad Ax = b$$

dann gilt starke Dualität.

Für den Beweis verwenden wir den folgenden Trennungssatz für konvexe Mengen:

Theorem 4.3 (Trennungssatz für konvexe Mengen)

Seien $A, B \subseteq \mathbb{R}^n$ konvex, nichtleer und disjunkt, d.h. $A \cap B = \emptyset$. Dann existieren $a \in \mathbb{R}^n \setminus \{0\}$ und $b \in \mathbb{R}$, sodass

$$\begin{aligned} a^\top x &\geq b & \forall x \in A \\ a^\top x &\leq b & \forall x \in B \end{aligned}$$

Die Hyperebene $h = \{x \in \mathbb{R}^n : a^\top x = b\}$ heißt trennende Hyperebene für A und B .

Beweis. siehe Funktionalanalysis □

Beweis (Theorem 4.2). Betrachte folgende Teilmengen von $\mathbb{R}^N = \mathbb{R}^{m+p+1}$

$$\begin{aligned} \mathcal{G} &:= \{(u, v, t) \in \mathbb{R}^N : \exists x \in \mathbb{R}^n \text{ mit } f_i(x) = 0 \ \forall i \in [m], Ax - b = v, f_0(x) = t\} \\ \mathcal{A} &:= \{(u, v, t) \in \mathbb{R}^N : \exists x \in \mathbb{R}^n \text{ mit } f_i(x) \leq u_i \ \forall i \in [m], Ax - b = v, f_0(x) \leq t\} \\ &= \mathcal{G} \oplus \mathbb{R}_{\geq 0}^m \times \{0\}^p \times \mathbb{R}_{\geq 0} \\ \mathcal{B} &:= \{(0, 0, t) \in \mathbb{R}^N : t \leq p_*\} \end{aligned}$$

Es gilt: \mathcal{A} und \mathcal{B} sind konvex.

Behauptung: $A \cap B = \emptyset$ — Beweis mit Widerspruch. Angenommen es existiert $(u, v, t) \in \mathcal{A} \cap \mathcal{B}$, dann gilt (wegen \mathcal{B}) $u = 0$, $v = 0$ und $t < p_*$ sowie (wegen \mathcal{A}), dass ein $x \in \mathbb{R}^n$ existiert mit

- $f_i(x) \leq u_i = 0$ für alle $i \in [m]$
- $h_i(x) \leq v_i = 0$ für alle $i \in [p]$
- $f(x) \leq t < p_*$

d.h. x ist zulässig für (K-OPT) und besser als optimal. Somit ist $\mathcal{A} \cap \mathcal{B} = \emptyset$.

Wir wenden den Trennungssatz an und erhalten, dass ein $(\lambda, \nu, \mu) \in \mathbb{R}^N \setminus \{0\}$ und $\alpha \in \mathbb{R}$ mit

- I. $\lambda^\top u + \nu^\top v + \mu t \geq \alpha \quad \forall (u, v, t) \in \mathcal{A}$
 II. $\lambda^\top u + \nu^\top v + \mu t \leq \alpha \quad \forall (u, v, t) \in \mathcal{B}$

Aus (II) erhalten wir $\mu t \leq \alpha$ für alle $t < p_*$ (da $u = 0 = v$) und somit auch (wegen Stetigkeit linearer Funktionen) $\mu p_* \leq \alpha$. Aus (I) erhält man $\lambda_i \geq 0$ für alle $i \in [m]$ und $\mu \geq 0$ (sonst Widerspruch, da linke Seite mit negativer Komponente beliebig klein werden kann). Mit (I) und (II) folgt nun für alle $x \in \mathbb{R}^n$

$$\sum_i \lambda_i f_i(x) + \sum_i \nu_i (Ax - b) \sum_i \mu f_0(x) = \lambda^\top u + \nu^\top v + \mu t \stackrel{(I)}{\geq} \alpha \geq \mu p_* \quad (\star)$$

Falls $\mu > 0$ ist, setze $\tilde{\lambda} := \frac{\lambda}{\mu}$ und $\tilde{\nu} := \frac{\nu}{\mu}$. Dann gilt für alle $x \in \mathbb{R}^n$

$$\sum_i \tilde{\lambda}_i f_i(x) + \sum_i \tilde{\nu}_i (Ax - b) + f_0(x) \geq p_*$$

Daraus folgt nun $g(\tilde{\lambda}, \tilde{\nu}) \geq p_*$, d.h. $d_* = \max_{(\lambda, \nu) \in \mathbb{R}_{\geq 0}^m \times \mathbb{R}^n} g(\lambda, \nu) \geq p_*$. Aber mit schwacher Dualität gilt $d_* \leq p_*$ und somit $p_* = d_*$.

Falls $\mu = 0$, dann folgt aus (\star)

$$\sum_i \lambda_i f_i(x) + \sum_i \nu_i (Ax - b) \geq 0 \quad \forall x \in \mathbb{R}^n$$

Mit Slaters Bedingung existiert ein $\bar{x} \in \mathbb{R}^n$ mit $f_i(\bar{x}) < 0$ für alle $i \in [m]$ und $A\bar{x} - b = 0$.

$$\Rightarrow \sum_i \underbrace{\lambda_i}_{\geq 0} \underbrace{f_i(\bar{x})}_{< 0} \geq 0 \quad \Rightarrow \quad \lambda = 0$$

$(\lambda, \nu, \mu) = (0, \nu, 0) \in \mathbb{R}^N \setminus \{0\}$, also $\nu \neq 0$ und $\nu^\top (A\bar{x} - b) = 0$, dann existiert auch \tilde{x} mit $\nu^\top (A\tilde{x} - b) < 0$ im Widerspruch zur Annahme. Also tritt der Fall $\mu = 0$ nicht ein. \square

4.3 Die Markowitz-Modelle

4.3.1 Markowitz-Modell 1: Portfoliooptimierung ohne risikofreie Anlage

Wir betrachten Anlagegüter $S = (S^1, \dots, S^n)$ mit stochastischen Einperioden-Renditen $R = (R^1, \dots, R^n)^\top$, d.h. $S_T^i = S_0^i(1 + R^i)$ ($i \in [n]$), mit

- Erwartungswert $\mu = \mathbb{E}[R] = (\mu_1, \dots, \mu_n)^\top \in \mathbb{R}^n$
- Kovarianzmatrix $\Sigma = \mathbb{E}[(R - \mu) \cdot (R - \mu)^\top]^4$
 Annahme: Σ ist regulär (und ist positiv definit), d.h. Σ^{-1} existiert

Das **Ziel** ist nun, das Anlagevermögen⁵ $w = 1$ auf Anlagegüter S^1, \dots, S^n aufzuteilen. Mit p_i

⁴ $\Sigma_{ii} = \text{Var}(R^i)$ und $\Sigma_{ij} = \text{Cov}(R^i, R^j)$

⁵Wir skalieren das Vermögen stets aus 1, daraus ergeben sich keine Einschränkungen.

bezeichnen wir die **Investition** in S^i , d.h. es muss stets $p_1 + \dots + p_n = w = 1$ gelten.

Erwartete Rendite:

$$\mu_p = \mathbb{E} [p^\top R] = p^\top \mu$$

Risiko (Standardabweichung):

$$\sigma_p = \sqrt{\text{Var}(p^\top R)} = \sqrt{\mathbb{E} [(p^\top (R - \mu))^2]} = \sqrt{\mathbb{E} [p^\top (R - \mu) (R - \mu)^\top p]} = \sqrt{p^\top \Sigma p}$$

Optimales Anlageproblem: minimiere das Risiko gegeben einer Zielrendite μ_*

$\min \frac{1}{2} p^\top \Sigma p$ über \mathbb{R}^n unter Nebenbedingungen $p^\top \mu = \mu_*$ (Zielrendite μ_*) und $p^\top \mathbb{1} = 1$ mit $\mathbb{1} = (1, \dots, 1)^\top \in \mathbb{R}^n$
(Mark-1)

Die Lagrange-Zielfunktion ergibt sich zu

$$\mathcal{L}(p, \lambda_1, \lambda_2) = \frac{1}{2} p^\top \Sigma p + \lambda_1 (\mu_* - p^\top \mu) + \lambda_2 (1 - p^\top \mathbb{1})$$

mit dualer Zielfunktion

$$g(\lambda_1, \lambda_2) = \inf_{p \in \mathbb{R}^n} \mathcal{L}(p, \lambda_1, \lambda_2)$$

Für den Gradienten der Zielfunktion gilt

$$\begin{aligned} \nabla_p \mathcal{L}(p, \lambda_1, \lambda_2) &= \Sigma p - \lambda_1 \mu - \lambda_2 \mathbb{1} \stackrel{!}{=} 0 \\ \Rightarrow p_* &= \Sigma^{-1} (\lambda_1 \mu + \lambda_2 \mathbb{1}) \end{aligned}$$

d.h

$$\begin{aligned} g(\lambda_1, \lambda_2) &= \mathcal{L}(p_*, \lambda_1, \lambda_2) \\ &= \frac{1}{2} (\lambda_1 \mu + \lambda_2 \mathbb{1})^\top \Sigma^{-1} \Sigma \Sigma^{-1} (\lambda_1 \mu + \lambda_2 \mathbb{1}) - (\lambda_1 \mu + \lambda_2 \mathbb{1})^\top \Sigma^{-1} (\lambda_1 \mu + \lambda_2 \mathbb{1}) + \lambda_1 \mu_* + \lambda_2 \\ &= -\frac{1}{2} (\lambda_1^2 a + 2\lambda_1 \lambda_2 b + \lambda_2^2 c) + \lambda_1 \mu_* + \lambda_2 \end{aligned}$$

mit

$$a = \mu^\top \Sigma^{-1} \mu \quad b = \mu^\top \Sigma^{-1} \mathbb{1} \quad c = \mathbb{1}^\top \Sigma^{-1} \mathbb{1}$$

Es gilt $a \geq 0$, $c \geq 0$ und mit Cauchy-Schwarz auch $ac \geq b^2$.

Maximiere g .

$$\begin{aligned} \frac{\partial}{\partial \lambda_1} g &= -a\lambda_1 - b\lambda_2 + \mu_* \stackrel{!}{=} 0 \Leftrightarrow a\lambda_1 + b\lambda_2 = \mu_* \\ \frac{\partial}{\partial \lambda_2} g &= -b\lambda_1 - c\lambda_2 + 1 \stackrel{!}{=} 0 \Leftrightarrow b\lambda_1 + c\lambda_2 = 1 \end{aligned}$$

Wir nehmen an, dass $ac > b^2$.
-bI +aII: $(ac - b^2)\lambda_2 = a - b\mu_* \Rightarrow \lambda_2^* = \frac{a - b\mu_*}{ac - b^2}$.
cI -bII: $(ac - b^2)\lambda_1 = c\mu_* - b \Rightarrow \lambda_1^* = \frac{c\mu_* - b}{ac - b^2}$

Minimierer von (Mark-1):

$$p_* = \lambda_1^* \Sigma^{-1} \mu + \lambda_2^* \Sigma^{-1} \mathbb{1}$$

Korollar 4.4 (Tobin's Two-Fund-Separation)

Jedes Pareto-optimale Portfolio für (Mark-1) kann (unabhängig von n) als Linearkombination der zwei Portfolios

$$p^* = \sigma^{-1} \mu \quad \text{und} \quad p_2^* = \Sigma^{-1} \mathbb{1}$$

dargestellt werden.

Dabei kann man p_1^* als das renditeorientierte und p_2^* als das sicherheitsorientierte (das Risiko minimierende) Portfolio beschreiben.

Bemerkung

- Die Gewichtung der zwei Portfolios p_1^* und p_2^* orientiert sich am Renditeziel μ_* .
- Die Portfolios p_1^* und p_2^* sind breit diversifiziert⁶, d.h. sie nutzen alle Anlagegüter $S = (S^1, \dots, S^n)$.
- p_1^* und p_2^* kann man auch als Anlagefonds interpretieren, welche Vermögen entsprechend den Portfolios p_1^* , p_2^* anlegen. Diese zwei Fonds sind ausreichend um (unabhängig von μ_*) Vermögen Pareto-optimal zu investieren.

Zuletzt wollen wir noch das Risiko der optimalen Strategie p_* berechnen.

$$\begin{aligned} \sigma_*^2 &= \text{Var} \left(p_*^\top R \right) = \mathbb{E} \left[(p_*^\top (R - \mu))^2 \right] = p_*^\top \Sigma p_* \\ &= (\lambda_1^* \mu + \lambda_2^* \mathbb{1})^\top \Sigma^{-1} \Sigma \Sigma^{-1} (\lambda_1^* \mu + \lambda_2^* \mathbb{1}) \\ &= (\lambda_1^*)^2 a + 2 \lambda_1^* \lambda_2^* b + (\lambda_2^*)^2 c \\ &= \frac{1}{(ac - b^2)^2} \left((c\mu_* - b)^2 a + 2(c\mu_* - b)(a - b\mu_*)b + (a - b\mu_*)^2 c \right) \\ &= \frac{1}{ac - b^2} \left(c\mu_*^2 - 2b\mu_* + a^2 \right) \end{aligned}$$

Der Graph von (σ_*, μ_*) ist ein Hyperbelast.

4.3.2 Markowitz-Modell II: optimale Anlage mit risikofreier Anlage

Wir betrachten Anlagegüter $S = (S^1, \dots, S^n)$ mit stochastischen Einperioden-Renditen $R = (R^1, \dots, R^n)^\top$ und zusätzlich einer *risikofreien* Anlage S^0 mit Verzinsung r .

Vermögen $w = 1$ aufgeteilt auf $1 = p_0 + p_1 + \dots + p_n$. Wir setzen $p = (p_1, \dots, p_n)^\top \in \mathbb{R}^n$.

Erwartete Rendite

$$\mu_p = \mathbb{E} \left[p^\top R + (1 - p^\top \mathbb{1})r \right] = p^\top (\mu - r\mathbb{1}) + r$$

⁶Die Vektoren p_1^* und p_2^* sind also nicht dünnbesetzt und haben in der Regel keine Null-Einträge.

Risiko

$$\sigma_p = \sqrt{\text{Var}(p^\top R)} = \sqrt{p^\top \Sigma p}$$

Anlageproblem

$$\min \frac{1}{2} p^\top \Sigma p \quad p \in \mathbb{R}^n \text{ unter } NB : p^\top (\mu - r \mathbb{1}) = \mu_* - r(\text{Zielrendite}) \quad (\text{Mark-2})$$

Lösung: UE

optimales Portfolio

$$p_* = \underbrace{\lambda_*}_{\in \mathbb{R}} \cdot \underbrace{\Sigma^{-1}(\mu - r \mathbb{1})}_{\in \mathbb{R}^n} \text{ mit } \lambda_* = \frac{\mu_* - r}{a^2 - 2br + cr^2}$$

Korollar 4.5 (Tobin's One-Fund-Theorem)

Jedes Pareto-optimale Portfolio für (Mark-2) kann als Linearkombination der risikofreien Anlage und des Portfolios

$$\Sigma^{-1}(\mu - r \mathbb{1})$$

dargestellt werden.

Graph von minimalem Risiko σ_* und Zielrendite μ_* :

Bemerkung

nominales Portfolio: $\theta = (\theta_1, \dots, \theta_n) \in \mathbb{R}^n$ Stückzahl von Anlagegut S^i . Portfoliowert: $V_0 = \theta^\top S_0 = \sum_{i=1}^n \theta_i S_0^i =: w \dots$ Anfangskapital $V_T = \theta^\top S_T = \sum_{i=1}^n \theta_i S_T^i$

relatives Portfolio: $p = (p_1, \dots, p_n) \in \mathbb{R}^n$ mit $p_i = \frac{\theta_i S_0^i}{w} \dots$ Vermögensanteil in S^i . $\sum_{i=1}^n p_i = \frac{1}{w} \sum_{i=1}^n \theta_i S_0^i = \frac{w}{w} = 1$

Renditen: Einzelnes Anlagegut: $R_i = \frac{S_T^i - S_0^i}{S_0^i}$ Gesamtes Portfolio: $R_p = \frac{V_T - V_0}{V_0} = \frac{1}{w} (\sum_{i=1}^n \theta_i S_T^i - \theta_i S_0^i)$

$$\frac{1}{w} \sum_{i=1}^n \theta_i (S_T^i - S_0^i) = \sum_{i=1}^n \underbrace{\frac{\theta_i S_0^i}{w}}_{=p_i} \cdot R_i = \sum_{i=1}^n p_i \cdot R_i = p^\top R \text{ ist linear in } p.$$

4.4 Capital Asset Pricing Model (CAPM)

Ausgangspunkt: Optimalportfolio im zweiten Markowitz-Problem $p_* = \lambda \cdot \Sigma^{-1}(\mu - r \cdot \mathbb{1})$. Wir normieren dies so, dass $p_*^\top \mathbb{1} = \mathbb{1}$. Damit brauchen wir $\lambda_* = \frac{1}{\mathbb{1}^\top \Sigma^{-1}(\mu - r \mathbb{1})} = \frac{1}{b - cr}$.

Wert des Marktportfolios: $M_0 = 1$ und $M_T = (1 + p_*^\top \cdot R)$, Rendite $R_M = p_*^\top \cdot R$

Zentrale Idee des CAPM:

- Betrachte M als beobachtbare Größe (im Gegensatz zum Markowitz-Modell, wo M der Output war)
- Aktienindex wie DAX oder S&P500 sollte gute Näherung für M ergeben.

Wir betrachten folgende Kennzahlen.

- Überschussrendite [excess return] (alpha):

$$\alpha_i = \mathbb{E}[R_i] - r \text{ für Wertpapier } S^i$$

$$\alpha_M = \mathbb{E}[R_M] - r = p_*^\top \mu - r = \frac{\mu^\top \Sigma^{-1}(\mu - r\mathbb{1})}{\mathbb{1}^\top \Sigma^{-1}(\mu - r\mathbb{1})} - r = \frac{a - rb}{b - rc} - r = \frac{a - 2rb + r^2c}{b - rc}$$

- Beta-Koeffizient:

$$\beta_i = \frac{\text{Cov}(R_i, R_M)}{\text{Var}(R_M)} \text{ skalierte Kovarianz zwischen Erträgen von } S^i \text{ und } M$$

- ▷ Maß für Korrelation der Wertpapiere S^i und Marktportfolio
- ▷ volle Kovarianzmatrix wird nicht benötigt

Wir berechnen:

$$\begin{aligned} \beta_i &= \frac{\mathbb{E}[(R_i - \mu_i)(R_M - \mu_M)]}{\mathbb{E}[(R_M - \mu_M)^2]} \\ &= \frac{\mathbb{E}[e_i^\top (R - \mu)(R - \mu)^\top p_*]}{\mathbb{E}[p_*^\top (R - \mu)(R - \mu)^\top p_*]} \\ &= \frac{e_i^\top \Sigma p_*}{p_*^\top \Sigma p_*} \\ &= \frac{\lambda_* e_i^\top \Sigma \Sigma^{-1}(\mu - r\mathbb{1})}{\lambda_*^2 (\mu - r\mathbb{1})^\top \Sigma^{-1} \Sigma \Sigma^{-1}(\mu - r\mathbb{1})} \\ &= \frac{\mu_i - r}{\lambda_*(a - 2rb - r^2c)} \\ &= \frac{\mu_i - r}{\mu_M - r} \end{aligned}$$

Durch Umstellen erhalten wir die CAPM-Gleichung

$$\beta_i(\mu_M - r) = (\mu_i - r) \quad \forall i \in [n]$$

Dabei bezeichnet β_i den Beta-Koeffizienten von S^i , $(\mu_M - r)$ die Überschussrendite des Marktportfolios und $(\mu_i - r)$ als Überschussrendite von S^i (alpha).

- ▷ Das kann als Regressionsgleichung für $(\alpha_i, \beta_i)_{i \in [n]}$ interpretiert werden.
- ▷ Entscheidend für die Attraktivität eines Wertpapiers S^i ist nicht die Überschussrendite $\alpha_i = \mu_i - r$ alleine, sondern in Relation zu β_i .
- ▷ CAPM kann empirisch überprüft werden durch Schätzung $(\hat{\alpha}_i, \hat{\beta}_i)$ und Regression

$$\hat{\beta}_i(\mu_M - r) = \hat{\alpha}_i + \varepsilon_i \quad (\star)$$

wobei ideal bedeutet, dass $\sum_{i=1}^n \varepsilon_i^2$ klein ist.

Kritik am CAPM:

- Regrression (\star) empirisch im Allgemeinen nicht besonders gut (Fehler $\sum \varepsilon_i^2$ groß)
- Schätzung von μ_i und μ_M schwierig

Erweiterungen:

- Ergänze Schätzer $\hat{\mu}_i$ und $\hat{\mu}_M$ um Expertenmeinungen und Kofidenzaussagen
 \Rightarrow Black-Littermann-Modell
- Erweiterung der Regressionsgleichung (\star) um weitere Variablen
 \Rightarrow Fama-French-Modell

4.5 Präferenzordnungen und Erwartungsnutzen

Kritik an Markowitz:

- Standardabweichung $\sqrt{\text{Var}(R)}$ nicht unbedingt gutes Risikomaß
- Entwicklungen unter Unsicherheit meist komplexer als durch Erwartungswert-Varianz-Prinzip beschrieben.

Axiomatischer Zugang: Präferenzordnungen (PO) Sei $\Omega, \mathcal{F}, \mathbb{P}$ ein Wahrscheinlichkeitsraum und wie üblich $L_1(\Omega, \mathcal{F}, \mathbb{P})$ der Raum der integrierbaren Zufallsvariablen. Setze

$$\mathcal{M} := \text{Menge der Verteilungsfunktionen } F_X \text{ von } X \in L_1(\Omega, \mathcal{F}, \mathbb{P})$$

Seien $X, Y \in L_1(\Omega)$ mit der Interpretation als risikobehaftete Auszahlungen ("Lotterie") versehen. Wir wollen eine Ordnungsrelation \preceq mit Bedeutung

$$X \preceq Y \Leftrightarrow "Y \text{ wird bevorzugt gegenüber } X"$$

Wir beschränken uns auf "verteilungsinvariante" Präferenzordnungen (unabhängig von der Verteilung), welche durch Relation auf \mathcal{M} erklärt werden, d.h.

$$X \preceq Y \Leftrightarrow F_X \preceq F_Y$$

Definition. Eine Relation \preceq auf \mathcal{M} heißt **Präferenzordnung**, wenn gilt

$$\begin{aligned} F &\preceq F \quad \forall F \in \mathcal{M} && \text{(Reflexivität)} \\ (F \preceq G) \wedge (G \preceq H) &\Rightarrow (F \preceq H) \quad \forall F, G, H \in \mathcal{M} && \text{(Transitivität)} \\ (F \preceq G) \vee (G \preceq F) &\quad \forall F, G \in \mathcal{M} && \text{(Vollständigkeit)} \end{aligned}$$

Bemerkung

- Menge \mathcal{M} ist konvex, d.h. für alle $F, G \in \mathcal{M}$ und $\alpha \in [0, 1]$ gilt

$$H := (1 - \alpha)F + \alpha G \in \mathcal{M} \tag{\star}$$

- (\star) lässt sich als "Mischen" von F und G interpretieren.
- Sei $X \sim F_X$ und $Y \sim F_Y$ sowie $A \perp\!\!\!\perp (X, Y)$ mit $\mathbb{P}(A = 0) = \alpha$ und $\mathbb{P}(A = 1) = (1 - \alpha)$. Dann gilt

$$(1 - A)X + AY \sim (1 - \alpha)F_X + \alpha F_Y$$

- Aus gegebener Präferenzordnung können wir ableiten:

- ▷ Äquivalenzrelation $F \sim G \Leftrightarrow (F \preceq G) \wedge (G \preceq F)$ – "Indifferenz zwischen F und G "
- ▷ strikte Relation: $F \prec G \Leftrightarrow (F \preceq G) \wedge \neg(G \preceq F)$ – " G wird strikt gegenüber F bevorzugt"
- Für "deterministische Zufallsvariable" $a \in \mathbb{R}$ ist die Verteilungsfunktion $F_a = \mathbb{1}_{[a, \infty)}$.

Eine Präferenzordnung kann folgende Eigenschaften besitzen:

- (I) **Monotonie:** Für alle $a, b \in \mathbb{R}$ mit $a \leq b$ gilt $F_a \preceq F_b$ ("mehr besser als weniger")
- (II) **Risikoaversion:** Für alle $X \in L_1(\Omega)$ gilt $F_X \preceq F_{\mathbb{E}[X]}$ ("sicher besser als unsicher")
- (III) **Mittelwertseigenschaft:** Seien $F, G, H \in \mathcal{M}$ mit $F \preceq G \preceq H$. Dann existiert ein $\alpha \in [0, 1]$ mit $(1 - \alpha)F + \alpha H \sim G$.
- (IV) **Unabhängigkeitsaxiom:** Für alle $F, G, H \in \mathcal{M}$ und alle $\alpha \in [0, 1]$ gilt

$$F \preceq G \Rightarrow (1 - \alpha)F + \alpha H \preceq (1 - \alpha)G + \alpha H$$

Definition (Erwartungsnutzen). Sei $U: \mathbb{R} \rightarrow [-\infty, \infty]$ monoton steigend und konkav ("Bernoulli'sche) Nutzenfunktion"). Dann definiert

$$F \preceq_U G \Leftrightarrow \int U dF \leq \int U dG \quad (*)$$

eine Präferenzordnung auf \mathcal{M} . Wir sagen " \preceq_U " folgt dem Erwartungsnutzenprinzip (ENP).

Bemerkung

Wenn wir $(*)$ für Zufallsvariablen $X, Y \in L_1$ formulieren, erhalten wir

$$X \preceq_U Y \Leftrightarrow \mathbb{E}[U(X)] \leq \mathbb{E}[U(Y)]$$

Beispiel

Wichtige Beispiele für Nutzenfunktionen:

- **logarithmischer Nutzen:** $U(x) = \begin{cases} \log(x) & \text{für } x > 0 \\ -\infty & \text{für } x \leq 0 \end{cases}$
- **Potenznutzen:** $U(x) = \begin{cases} \frac{x^{1-\alpha}}{1-\alpha} & \text{für } x \geq 0 \\ -\infty & \text{für } x \leq 0 \end{cases}$ mit $\alpha \in (0, \infty) \setminus \{1\}$
- **Exponentialnutzen:** $U(x) = \gamma \exp(-\gamma x)$ mit $\gamma \in \mathbb{R}$.

Theorem 4.6 (Satz von von NEUMANN & MORGENTERN, 1953)

Sei $|\Omega| < \infty$ und \preceq eine Präferenzordnung auf \mathcal{M}_Ω . Dann sind äquivalent:

- (1) Die Präferenzordnung erfüllt (I) bis (IV)
- (2) Die Präferenzordnung folgt dem Erwartungsnutzenprinzip.

Beweis. Wir zeigen hier nur die Rückrichtung, d.h. \preceq erfüllt die Eigenschaften (I) bis (IV).

(I) Sei $a \leq b$. Dann gilt $\int U dF_a = U(a) \stackrel{U \text{ steigend}}{\leq} U(b) = \int U dF_b$, also $F_a \preceq F_b$.

(II) Sei $X \in L_1$. Es gilt

$$\int U dF_X = \int_{-\infty}^{\infty} U(x) dF_X \stackrel{\text{Jensen}}{\leq} U\left(\int_{-\infty}^{\infty} x dF_X\right) = U(\mathbb{E}[X]) = \int U dF_{\mathbb{E}[X]}$$

und somit $F_X \preceq F_{\mathbb{E}[X]}$ (oder alternativ $\mathbb{E}[U(X)] \leq U(\mathbb{E}[X])$ mit der Jensen-Ungleichung).

(III) Sei $F \preceq G \preceq H$. Zu zeigen: es existiert ein $\alpha \in [0, 1]$ mit $(1 - \alpha)F + \alpha H \sim G$. Setze

$$\alpha_* := \frac{\int U dG - \int U dF}{\int U dH - \int U dF}$$

Es gilt

$$\begin{aligned} \int U((1 - \alpha)dF + \alpha_*dH) &= (1 - \alpha_*) \int U dF + \alpha_* \int U dH \\ &= \int U dF + \alpha_* \left(\int U dH - \int U dF \right) \\ &= \int U dG \end{aligned}$$

(IV) zu zeigen: $F \preceq G \Rightarrow (1 - \alpha)F + \alpha H \preceq (1 - \alpha)G + \alpha H$

$$\begin{aligned} \int U((1 - \alpha)dF + \alpha dH) &= (1 - \alpha) \int U dF + \alpha \int U dH \\ &\stackrel{F \preceq G}{\leq} (1 - \alpha) \int U dG + \alpha \int U dH \\ &= \int U((1 - \alpha)dG + \alpha dH) \end{aligned} \quad \square$$

Bemerkung

- Theorem 4.6 wurde in den 50ern als starke Rechtfertigung für das Erwartungsnutzenprinzip wahrgenommen.
- (Empirische) Kritik Ende der 70er durch die Psychologen KAHNEMANN und TVERSKY
 \Rightarrow "Prospect Theory"

Definition. Sei U eine Bernoullische Nutzenfunktion und \preceq_U die zugehörige Präferenzordnung.

- (a) Für $X \in L_1$ heißt $c := c_*(X, U) \in \mathbb{R}$ mit $c \sim_U X$ das **certainty equivalent** von X .
- (b) Für $U \in C^2$ heißen

$$A_U(x) := -\frac{U''(x)}{U'(x)} \quad R_U(x) := -\frac{xU''(x)}{U'(x)}$$

die **Arrow-Pratt-Koeffizienten** der absoluten bzw. relativen Risikoaversion.

Bemerkung

- Für U streng monoton steigen gilt $c_*(X, U) = U^{-1}(\mathbb{E}[U(X)])$, denn

$$c \sim_U X \Leftrightarrow U(c) = \mathbb{E}[U(X)]$$

- Motivation für $A_U(x)$: Einer Person mit Vermögen $x \in \mathbb{R}$ werde eine "Lotterie" εY angeboten (mit ε klein). Nach dem Erwartungsnutzenprinzip sollte die Person diese annehmen, wenn $\mathbb{E}[U(x + \varepsilon Y)] \geq U(x)$. Taylor-Entwicklung liefert

$$\mathbb{E}[U(x + \varepsilon Y)] - U(x) = \mathbb{E}[\varepsilon Y] U'(x) + \frac{1}{2} \mathbb{E}[\varepsilon^2 Y^2] U''(x) + \dots$$

Die Person sollte also annehmen, wenn

$$2 \cdot \frac{\mathbb{E}[\varepsilon Y]}{\mathbb{E}[\varepsilon^2 Y^2]} \geq -\frac{U''(x)}{U'(x)} = A_U(x) \quad \frac{\text{erw. Gewinn}}{\text{Risiko}}$$

Klassisches Beispiel zum "certainty equivalent"

St. Petersburg Paradoxon (NICHOLAS BERNOULLI, 1713)

Spiel: Es werde eine faire Münze solange geworfen bis in der N -ten Runde das erste Mal "Zahl" fällt. Der Gewinn beträgt 2^{N-1} Euro.

Frage: Wie hoch soll der Einsatz sein um am Spiel teilzunehmen?

Antwort 1: Einsatz = erwarteter Gewinn: N ist geometrisch verteilt, d.h. $\mathbb{P}(N = k) = \left(\frac{1}{2}\right)^k$ für $k \in \mathbb{N}$.

$$\mathbb{E}[2^{N-1}] = \sum_{k=1}^{\infty} \left(\frac{1}{2}\right)^k 2^{k-1} = \sum_{k=1}^{\infty} \left(\frac{1}{2}\right) = +\infty$$

Die Antworten 2 und 3 von Gabriel Cramer und Daniel Bernoulli entsprechen dem certainty equivalent mit Nutzenfunktionen

$$U_1(x) = \sqrt{x} \quad (\text{Cramer})$$

$$U_2(x) = \log(x) \quad (\text{Bernoulli})$$

Antwort 2 von Cramer: Es ist $c_1 = c_*(2^{N-1}, U_1) = U_1^{-1}(\mathbb{E}[U_1(2^{N-1})])$ und

$$\begin{aligned} \mathbb{E}[U_1(2^{N-1})] &= \sum_{k=1}^{\infty} \left(\frac{1}{2}\right)^k \sqrt{2^{k-1}} = \sum_{k=1}^{\infty} 2^{-k+\frac{k-1}{2}} = \frac{1}{\sqrt{2}} \sum_{k=1}^{\infty} 2^{-\frac{k}{2}} \\ &= \frac{1}{\sqrt{2}} \sum_{k=1}^{\infty} \left(\frac{1}{\sqrt{2}}\right)^k \\ &= \frac{1}{\sqrt{2}} \cdot \frac{\frac{1}{\sqrt{2}}}{1 - \frac{1}{\sqrt{2}}} = \frac{1}{2-\sqrt{2}} \end{aligned}$$

Somit also $c_1 = \left(\frac{1}{2-\sqrt{2}}\right)^2 \approx 2.914$, d.h. der faire Einsatz betrüge 2.91 Euro.

Antwort 3 von Bernoulli: $c_2 = c_*(2^{N-1}, U_2)$.

$$\begin{aligned}
 \mathbb{E} \left[U_2(2^{N-1}) \right] &= \sum_{k=1}^{\infty} \left(\frac{1}{2} \right)^k \log(2^{k-1}) = \log(2) \cdot \sum_{k=1}^{\infty} 2^{-k} (k-1) \\
 &= \frac{\log(2)}{2} \cdot \sum_{k=0}^{\infty} k \cdot 2^{-k} \\
 &\stackrel{(\star)}{=} \frac{\log(2)}{4} \cdot \frac{1}{(1 - \frac{1}{2})^2} \\
 &= \log(2)
 \end{aligned}$$

wobei wir in (\star) verwendet haben, dass

$$\sum_{k=0}^{\infty} z^k = \frac{1}{1-z} \quad \xrightarrow{\frac{\partial}{\partial z}} \quad \sum_{k=0}^{\infty} k \cdot z^{k-1} = \frac{1}{(1-z)^2}$$

Somit gilt $c_2 = e^{\log(2)} = 2$, d.h. der faire Einsatz beträgt 2 Euro.

Kapitel 5

RISIKOMASSE

- Risikomessung aus der Perspektive einer Bank / Versicherung / Aufsichtsbehörde
- Fokus auf ungünstigste (d.h. auf verlustreichste) Szenarien, d.h. Finanzkrise, Kreditausfälle
- Wie können wir verhindern, dass Banken / Versicherungen in solchen Szenarien pleite gehen?
- aus gesellschaftlicher Sicht wichtig, da „Bail-out“-Kosten vom Staat / Steuerzahler übernommen werden müssen

Prinzip: Risiko muss mit *angemessenem* Risikokapital hinterlegt werden, welches mögliche Verluste abfedern kann

Analogie: Damm / Sturmflut — Wie hoch muss der Damm sein um auch ein Jahrhunderthochwasser standhalten zu können?

Rechtliches Rahmenwerk: Basel I, II und III für Banken sowie Solveny I und II für Versicherungen.

In dieser Vorlesung betrachten wir die mathematischen Aspekte.

Betrachte Zufallsvariable $L: \Omega \rightarrow \mathbb{R}$ („loss“ / Verlust), d.h. hohe Werte entsprechen hohen Verlusten.

5.1 Value-at-Risk

Definition. Der Value-at-Risk von L mit Verteilungsfunktion F_L zum Level $\alpha \in (0, 1)$ ist gegeben durch

$$\text{VaR}_\alpha(L) := \inf \{ \ell \in \mathbb{R} : F_L(\ell) \geq \alpha \} = \inf \{ \ell \in \mathbb{R} : \mathbb{P}(L > \ell) \leq 1 - \alpha \}$$

- In Worten: $\text{VaR}_\alpha(L)$ ist die kleinste Zahl ℓ mit der Eigenschaft, dass die Wahrscheinlichkeit einen Verlust L zu erleiden, der ℓ übersteigt, kleiner als $1 - \alpha$ ist.
- Value-at-Risk entspricht genau dem α -Quantil $q_\alpha(L)$ von F_L .
- Typische Werte für α sind beispielsweise 0.95 oder 0.99.

Skizze: Dichte von L , Verteilung von L

Die Quantilfunktion

$$q_\alpha(L) = \inf \{ \ell \in \mathbb{R} : F_L(\ell) \geq \alpha \} \quad (\star)$$

wird auch als verallgemeinerte Inverse von F_L bezeichnet.

Lemma 5.1

- (a) Wenn F_L stetig und streng monoton ist, dann gilt $q_\alpha(L) = F_L^{-1}(\alpha)$ für alle $\alpha \in (0, 1)$.
- (b) Wenn F_L eine Dichte f_L besitzt, dann gilt $\int_{-\infty}^{q_\alpha(L)} f_L(u) du = \alpha$ für alle $\alpha \in (0, 1)$.
- (c) Ein Wert $x_0 \in \mathbb{R}$ ist das α -Quantil einer Verteilung genau dann, wenn gilt

$$F_L(x_0) \geq \alpha \text{ und } F_L(x) < \alpha \text{ für alle } x < x_0$$

Bemerkung (Problematische Fälle)

- F_L nicht stetig (Sprung über Schwelle α). In diesem Fall ist $F_L(q_\alpha(L)) \neq \alpha$.
- F_L ist nicht streng monoton (flache Stelle). Dann hat $F_L(x_0) = \alpha$ mehrere Lösungen, d.h. F_L^{-1} ist nicht wohldefiniert.

Beweis. Sei F_L stetig, dann ist $F_L: \mathbb{R} \rightarrow (0, 1)$ surjektiv und das Infimum in (\star) wird angenommen, d.h. $F_L(q_\alpha(L)) = \alpha$.

(zu a) F_L ist streng monoton, d.h. auch injektiv und somit mit der Eingangsfeststellung bijektiv. Somit existiert F_L^{-1} und es gilt $q_\alpha(L) = F_L^{-1}(\alpha)$.

(zu b) F_L besitze eine Dichte f_L . Dann ist $F_L(x) = \int_{-\infty}^x f_L(u) du$ stetig. Mit $x = q_\alpha(L)$ gilt $\alpha = \int_{-\infty}^{q_\alpha(L)} f_L(u) du$.

(zu c) entspricht der Definition des Infimums als größte untere Schranke □

Beispiel

(1) Sei $L = \mathcal{N}(\mu, \sigma^2)$, d.h. $F_L(x) = \Phi\left(\frac{x-\mu}{\sigma}\right)$.

$$\alpha = F_L(x_0) \Rightarrow \alpha = \Phi\left(\frac{x_0 - \mu}{\sigma}\right) \Rightarrow \Phi^{-1}(\alpha) = \frac{x_0 - \mu}{\sigma} \Rightarrow x_0 = \mu + \sigma \Phi^{-1}(\alpha)$$

$$\text{d.h. } \text{VaR}_\alpha(L) = \mu + \sigma \Phi^{-1}(\alpha).$$

(2) Verallgemeinerung: Sei $L \sim F_{\mu, \sigma}$ mit $F_{\mu, \sigma}(x) = F\left(\frac{x-\mu}{\sigma}\right)$ und F ist stetig und streng monoton („Lage-Skalen-Familie“ von Verteilungen). Dann gilt

$$\text{VaR}_\alpha(L) = \mu + \sigma F^{-1}(\alpha)$$

Aber im Allgemeinen muss μ nicht dem Erwartungswert von L entsprechen und σ^2 nicht der Varianz.

(3) Betrachte die Student-t-Verteilung mit $\nu > 0$ Freiheitsgraden und Dichte

$$f(x) = c_\nu \cdot \left(1 + \frac{x^2}{\nu}\right)^{-\left(\frac{\nu+1}{2}\right)}$$

Es sei $L \sim F_{\mu,\sigma}$ mit $F_{\mu,\sigma}(x) = F\left(\frac{x-\mu}{\sigma}\right)$ als Verteilungsfunktion von L . Es gilt:

$$\mathbb{E}[L] = \begin{cases} \text{n. def.} & \text{für } \nu \in [0, 1] \\ \mu & \text{für } \nu > 1 \end{cases} \quad \text{und} \quad \text{Var}(L) = \begin{cases} \text{n. def.} & \text{für } \nu \in [0, 2] \\ \sigma \sqrt{\frac{\nu}{\nu-1}} & \text{für } \nu > 2 \end{cases}$$

- (4) Sei $L = \sigma B_t$ mit einer Brownschen Bewegung B (d.h. $B_t \sim \mathcal{N}(0, t)$). Es gilt $\text{VaR}_\alpha(L) = \sigma \sqrt{t} \cdot \Phi^{-1}(\alpha)$. Intuition: B_t beschreibt die Summe vieler kleiner Verluste über ein Zeitintervall $[0, t] \rightarrow \text{VaR}$ skaliert mit \sqrt{t} („square-root-law“)

Vorteile des VaR:

- für jede Verteilung F_L definiert
- einfach zu berechnen
- einfach zu verstehen / zu erklären

Nachteile des VaR:

- berücksichtigt *Wahrscheinlichkeit* hoher Verluste, aber nicht deren *Ausmaß*
- nicht subadditiv, d.h. bestraft unter Umständen Risikostreuung bzw. Diversifikation (mehr dazu im Kapitel 5.3)

5.2 Expected Shortfall

Definition. Für einen Verlust L mit $\mathbb{E}[L^+] < \infty$ ist der **Expected Shortfall** (ES) zum Level $\alpha \in (0, 1)$ definiert als

$$\text{ES}_\alpha(L) := \frac{1}{1-\alpha} \int_\alpha^1 \text{VaR}_u(L) du$$

Bemerkung. Der Expected Shortfall wird auch „Average Value-at-Risk“ oder „Conditional Value-at-Risk“ genannt.

Skizze:

Theorem 5.2

Sei $\mathbb{E}[L^+] < \infty$.

- (a) $\text{ES}_\alpha(L) \geq \text{VaR}_\alpha(L)$ für alle $\alpha \in [0, 1]$
- (b) Besitzt L eine stetige, streng monotone Verteilungsfunktion, so gilt

$$\text{ES}_\alpha(L) = \mathbb{E}[L | L \geq \text{VaR}_\alpha(L)]$$

- (c) Besitzt L eine Dichte f_L , so gilt

$$\text{ES}_\alpha(L) = \frac{1}{1-\alpha} \int_{\text{VaR}_\alpha(L)}^{\infty} u f_L(u) du$$

Interpretation von (b): Der ES_α ist der erwartete Verlust gegeben, dass der Verlust die Schwelle VaR_α übersteigt.

Beweis. (zu a) Die Abbildung $u \mapsto \text{VaR}_u(L)$ ist monoton steigend.

$$\Rightarrow \text{ES}_\alpha(L) = \frac{1}{1-\alpha} \int_\alpha^1 \text{VaR}_u(L) du \geq \frac{1}{1-\alpha} \int_\alpha^1 \text{VaR}_\alpha(L) du = \text{VaR}_\alpha(L)$$

(zu b) Sei F Verteilungsfunktion von L . Aus der Wahrscheinlichkeitstheorie ist bekannt, dass $U := F(L)$ ist gleichverteilt¹ auf $[0, 1]$ und $\text{VaR}_\alpha(L) = F^{-1}(\alpha)$. Also ist

$$\begin{aligned} \mathbb{E}[L | L \geq \text{VaR}_\alpha(L)] &= \frac{\mathbb{E}[L \cdot \mathbb{1}_{L \geq \text{VaR}_\alpha(L)}]}{\mathbb{P}(L \geq \text{VaR}_\alpha(L))} = \frac{1}{1-\alpha} \mathbb{E}[L \cdot \mathbb{1}_{L \geq F^{-1}(\alpha)}] \\ &= \frac{1}{1-\alpha} \mathbb{E}[F^{-1}(U) \mathbb{1}_{U \geq \alpha}] \\ &= \frac{1}{1-\alpha} \int_\alpha^1 F^{-1}(u) du \\ &= \frac{1}{1-\alpha} \int_\alpha^1 \text{VaR}_u(L) du \\ &= \text{ES}_\alpha(L) \end{aligned}$$

$$(zu c) \quad \mathbb{E}[L | L \geq \text{VaR}_\alpha(L)] = \frac{1}{1-\alpha} \mathbb{E}[L \cdot \mathbb{1}_{L \geq \text{VaR}_\alpha(L)}] = \frac{1}{1-\alpha} \int_{\text{VaR}_\alpha(L)}^{\infty} u f_L(u) du \quad \square$$

Beispiel

(1) $L \sim \mathcal{N}(\mu, \sigma^2)$, d.h. $F_L(x) = \Phi\left(\frac{x-\mu}{\sigma}\right)$ und Dichte $f_L(x) = \frac{1}{\sigma} \varphi\left(\frac{x-\mu}{\sigma}\right)$ sowie $\varphi(x) =$

¹ $\mathbb{P}(U \leq x) = \mathbb{P}(F(L) \leq x) = \mathbb{P}(L \leq F^{-1}(x)) = F(F^{-1}(x)) = x \quad \forall x \in [0, 1] \Rightarrow U \sim U[0, 1]$

$\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}$. Diese hat die Eigenschaft $\varphi'(x) = -x\varphi(x)$. Für den Expected Shortfall geht

$$\begin{aligned} \text{ES}_\alpha(L) &= \frac{1}{1-\alpha} \int_\alpha^1 \text{VaR}_u(L) du = \frac{1}{1-\alpha} \int_\alpha^1 (\mu + \sigma \Phi^{-1}(u)) du \\ &= \mu + \frac{\sigma}{1-\alpha} \int_\alpha^1 \Phi^{-1}(u) du \\ &\quad \left[\begin{array}{l} x = \Phi^{-1}(u) \\ \Phi(x) = u \\ \varphi(x) dx = du \end{array} \right] \\ &= \mu + \frac{\sigma}{1-\alpha} \int_{\Phi^{-1}(\alpha)}^\infty x \varphi(x) dx \\ &= \mu - \frac{\sigma}{1-\alpha} \varphi(x) \Big|_{x=\Phi^{-1}(\alpha)}^\alpha \\ &= \mu + \frac{\sigma}{1-\alpha} \varphi(\Phi^{-1}(\alpha)) \end{aligned}$$

vergleiche dazu $\text{VaR}_\alpha(L) = \mu + \sigma \Phi^{-1}(\alpha)$

(2) Student-t-Verteilung: siehe Übung

Vorteile des ES:

- berücksichtigt die Wahrscheinlichkeit *und* das Ausmaß hoher Verluste (d.h. Form des Verteilungsende wird erfasst)
- Subadditivität, d.h. er berücksichtigt das Diversifikationsprinzip (mehr dazu später)

Nachteile des ES:

- nicht immer definiert ($\mathbb{E}[L^+] < \infty$)
- im Allgemeinen schwieriger zu berechnen und zu erklären

5.3 Risikomaße: Axiomatischer Zugang

Sei \mathcal{M} die Menge der \mathbb{R} -wertigen Zufallsvariablen auf $(\Omega, \mathcal{F}, \mathbb{P})$. Wir wollen untersuchen, welche Abbildungen

$$\rho: \mathcal{M} \rightarrow \mathbb{R}, L \mapsto \rho(L)$$

als Risikomaße geeignet sind.

Definition (Eigenschaften von Risikomaßen). Seien $L_1, L_2 \in \mathcal{M}$.

- (a) Monotonie: $L_1 \leq L_2$ fast sicher $\Rightarrow \rho(L_1) \leq \rho(L_2)$
- (b) Translationsinvarianz: $\rho(L + m) = \rho(L) + m$ für alle $m \in \mathbb{R}$ (Kapitalanforderung)
- (c) Subadditivität: $\rho(L_1 + L_2) \leq \rho(L_1) + \rho(L_2)$ (belohnt Risikostreuung/Diversifikation)
- (d) positive Homogenität: $\rho(c \cdot L) = c \cdot \rho(L)$ für alle $c \geq 0$
- (e) Konvexität: $\rho(\gamma L_1 + (1 - \gamma)L_2) \leq \gamma \rho(L_1) + (1 - \gamma)\rho(L_2)$ für alle $\gamma \in [0, 1]$
- (f) Verteilungsinvarianz: $F_{L_1} = F_{L_2} \Rightarrow \rho(L_1) = \rho(L_2)$

Bemerkung. Translationsinvarianz beschreibt eine Kapitalanforderung; Subadditivität belohnt

Risikostreuung/Diversifikation

- Es gilt (c) + (d) \Rightarrow (e), denn

$$\rho(\gamma L_1 + (1 - \gamma)L_2) \stackrel{(c)}{\leq} \rho(\gamma L_1) + \rho((1 - \gamma)L_2) \stackrel{(d)}{=} \gamma \rho(L_1) + (1 - \gamma)\rho(L_2)$$

- Es gilt (d) + (e) \Rightarrow (c), denn

$$\rho(L_1 + L_2) \stackrel{(d)}{=} 2 \cdot \rho\left(\frac{L_1}{2} + \frac{L_2}{2}\right) \stackrel{(e)}{=} \rho(L_1) + \rho(L_2)$$

Definition. (1) Ein Risikomaß heißt **monetär**, wenn (a) und (b) gelten.
 (2) Ein Risikomaß heißt **konvex**, wenn (a), (b) und (e) gilt.
 (3) Ein Risikomaß heißt **kohärent**, wenn (a) bis (d) gilt (und damit auch (e)).

Bemerkung. Offensichtlich gilt:

$$\text{kohärente Risikomaße} \subseteq \text{konvexe Risikomaße} \subseteq \text{monetäre Risikomaße}$$

Beispiel

- Der Value-at-Risk VaR_α ist monetär, positiv homogen und verteilungsinvariant, aber nicht subadditiv oder konvex.

Seien X_1, X_2 unabhängig und identisch verteilt mit

$$X_i = \begin{cases} +1 & \text{mit Wahrscheinlichkeit } p = \frac{1}{3} \\ -1 & \text{mit Wahrscheinlichkeit } 1 - p = \frac{2}{3} \end{cases}$$

Somit ist $\text{VaR}_{0.5}(X_i) = -1$. Also gilt

$$X_1 + X_2 = \begin{cases} +2 & \text{mit Wahrscheinlichkeit } p^2 = \frac{1}{9} \\ 0 & \text{mit Wahrscheinlichkeit } p(1 - p) = \frac{2}{9} \\ -2 & \text{mit Wahrscheinlichkeit } (1 - p)^2 = \frac{4}{9} \end{cases}$$

d.h. $\text{VaR}_{0.5}(X_1 + X_2) = 0$.

$$\Rightarrow \text{VaR}_{0.5}(X_1 + X_2) > \text{VaR}_{0.5}(X_1) + \text{VaR}_{0.5}(X_2) \Rightarrow \text{nicht subadditiv}$$

- Der Expected-Shortfall ist ein kohärentes, verteilungsinvariantes Risikomaß (siehe Übung).