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

Kapitel 1

EINFÜHRUNG

1.1 Zentrale Fragestellungen der Finanzmathematik

1.1.1 Bewertung von Derivaten und Abischerung 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
$\to$ 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 $(\Omega, \mathcal{F})$. 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 $I = [O, T]$ (stetiges Modell) Dabei wird T als **Zeithorizont** bezeichnet.

Definition 1.3 (stochastischer Prozess)
Ein stochastischer Prozess S ist eine messbare Abbildung

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

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

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

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 Zeipunkt 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 $(\Omega, \mathcal{F})$ heißt **adapiert** 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

$$S^i : \begin{cases} (\Omega \times I) & \to & \mathbb{R} \\ (\omega, t) & \mapsto & S^i_t(\omega) \end{cases} \qquad (i \in \{0, 1, \ldots, d\})$$

S^i_t beschreibt dabei den Preis des i-ten Anlageguts zum Zeitpunkt t. S^i $(i \in \{1, \ldots, d\})$ 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^i_t ist jederzeit möglich. Der "Numeraire" S^0 hat eine Sonderrolle und beschreibt die Verzinsung von nicht in $(S^1, \ldots, S^d)$ 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 stiochastischen Prozess $S_t = (S^0_t, S^1_t, \ldots, S^d_t)$ mit $t \in I$.

Beispiel 1.8 (Cox-Ross-Rubinstein-Modell)
Das CRR-Modell ist ein zeitdiskretes Modell beschrieben durch

- $S^0_n = (1 + r)^n$ ... Verzinsung mit konstanter Rate r
- $S^1_n = S^1_0 \prod_{k=1}^n (1+R_k)$, wobei $(R_1, R_2, \ldots)$ unabhängige Zufallsvariablen mit zwei möglichen Werten $a < b$ sind

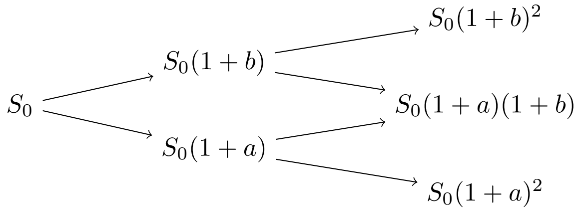

Abbildung 1.1: Cox-Ross-Rubinstein-Modell

$\hookrightarrow$ rekombinierender Baum,
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.

$$S_t^0 = e^{rt} \qquad \text{(Verzinsung mit konstanter Rate } r\text{)}$$
$$S_t^1 = S_0^1 \cdot \exp\left((\mu - \frac{\sigma^2}{2}) + \sigma B_t\right) \qquad \text{mit } \mu \in \mathbb{R}, \sigma > 0, S_0^1 > 0$$

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

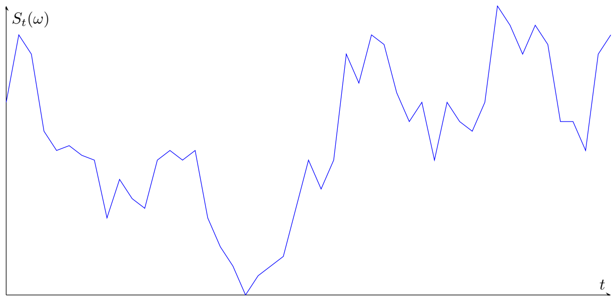

Abbildung 1.2: Black-Scholes-Modell

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 $t \le T$:	$B(t,T)$

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

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.

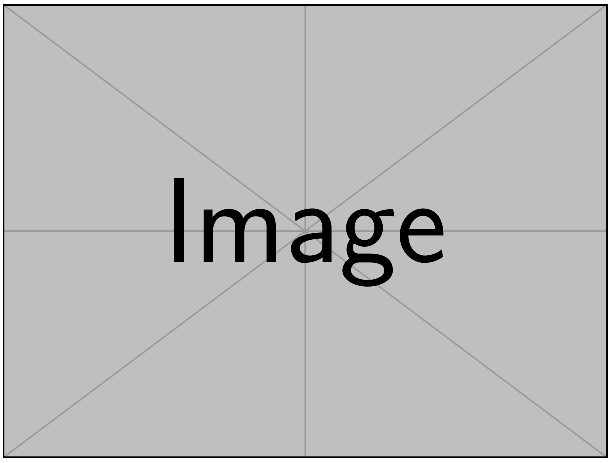

Abbildung 1.3: Zahlungsstrom einer Anleihe

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

aus Käufersicht: Vereinbarung zu bestimmtem zukünftigen Zeitpunkt T eine Einheit des Basisguts S zum Preis K zu kaufen (Kaufverpflichtung). Beliebt ist dieser bei Rohstoffen und Elektrizität.

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

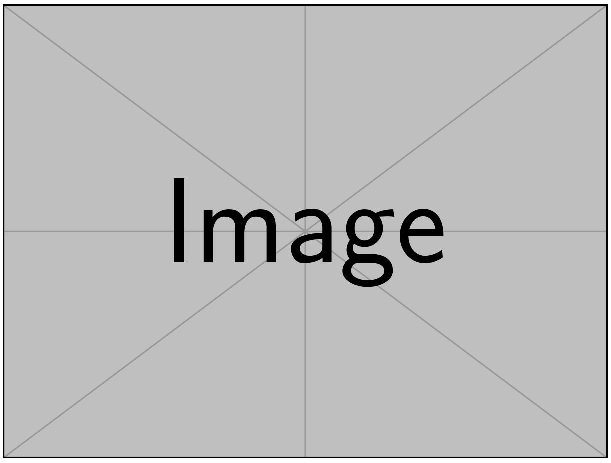

Abbildung 1.4: Auszahlungsprofil eines Terminvertrags

(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) $\rightarrow$ 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 1.10
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: $S_T \leq K$	Wert in T: $S_T > K$
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	$K - S_T + \frac{\varepsilon}{B(t,T)} > 0$ sicherer Gewinn	$\frac{\varepsilon}{B(t,T)} > 0$

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

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:

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

Definition 1.12
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 $g\colon \mathbb{R}^n \to \mathbb{R}^m, y \mapsto g(y)$.

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

Proposition 1.13

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

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

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

Allgemeiner gilt:

Theorem 1.14

eien (X, Y) Zufallsvariablen mit gemeinsamer Dichte auf $\mathbb{R}^m \times \mathbb{R}^n$ und $h\colon \mathbb{R}^m \times \mathbb{R}^n \to \mathbb{R}$ messbar mit $\mathbb{E}\left[h(X,Y)^2\right]$. Dann wird das Minimierungsproblem

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

gelöst durch

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

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

$$\begin{aligned}
\frac{\partial F}{\partial \varepsilon}(\varepsilon) &= 2\varepsilon\mathbb{E}\left[p(Y)^2\right] - 2\mathbb{E}\left[(X-g(Y))p(Y)\right] \quad\Rightarrow\quad \varepsilon_* := \frac{\mathbb{E}\left[(X-g(Y))p(Y)\right]}{\mathbb{E}\left[p(Y)^2\right]} = \frac{A}{B}\\
A &= \mathbb{E}\left[Xp(Y)\right] - \mathbb{E}\left[g(Y)p(Y)\right]\\
&= \int_{\mathbb{R}\times\mathbb{R}^n} xp(y) f_{XY}(x,y)\ \mathrm{d}x\ \mathrm{d}y - \int_{S_y} g(y)p(y)f_Y(y)\ \mathrm{d}y\\
&= \int_{\mathbb{R}\times\mathbb{R}^n} xp(y) f_{XY}(x,y)\ \mathrm{d}x\ \mathrm{d}y - \int_{\mathbb{R}\times S_y} xp(y)\underbrace{f_{X|Y}(x,y)f_Y(y)}_{=f_{XY}(x,y)}\ \mathrm{d}y = 0
\end{aligned}$$

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

Beispiel

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

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

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

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

$\rightarrow$ 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 $\{x_1, \ldots, x_m\}$ bzw. $\{y_1, \ldots, y_n\}$ annehmen, dann erhalten wir mit ähnlichen Überlegungen als Lösung von (min-1)

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

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 $X : \Omega \to \mathbb{R}$ und $p \in [1, \infty)$ definieren wir, doe L_p-Norm

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

und den L_p-Raum

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

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

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

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

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" ais dem letzten Abschnitt: Gegeben sei ein Zufallsvariale 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?

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

Aus Hilbertraumtheorie folgt, dass (min-2) besitzt eine eindeutige Lösung $G_* \in L_(\Omega, \mathcal{G}, \mathbb{P})$. 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 Erwartunswert von X bezüglich $\mathcal{G}$.

Theorem 1.15

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: $X \leq Y \quad \Rightarrow \quad \mathbb{E}[X|\mathcal{G}] \leq \mathbb{E}[Y|\mathcal{G}]$
- Dreiecksungleichung: $|\mathbb{E}[X|\mathcal{G}]| \leq \mathbb{E}[|X|\,|\mathcal{G}]$
- Unabhängigkeit: X unabhängig von $\mathcal{G} \quad \Rightarrow \quad \mathbb{E}[X|\mathcal{G}] = \mathbb{E}[X]$
- triviale σ-Algebra: $\mathcal{G} = \{\emptyset, \Omega\} \quad \Rightarrow \quad \mathbb{E}[X|\mathcal{G}] = \mathbb{E}[X]$

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 sag uns das Doob-Dynkin-Lemma, dass eine messbare Funktion $g\colon \mathbb{R}^n \to \mathbb{R}$ 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 $g\colon \mathbb{R}^n \to \mathbb{R}^m$ und falls eine bedingte Dichte existiert, dann gilt $\mathbb{E}[X|Y = y] = \int_{\mathbb{R}^m} x f_{X|Y}(x, y)\ \mathrm{d}x$.

(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 $\mathbb{E}[X|Y](\omega) = \int_{\mathbb{R}}^n x f_{X|Y}(x, Y(\omega))\ \mathrm{d}x$.

(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 1.16

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

$$\begin{aligned} \mathbb{E}[|X_n|] &< \infty \qquad \forall n \in \mathbb{N} \\ \mathbb{E}[X_{n+1}|X_1,\ldots,X_n] &= X_n \qquad \forall n \in \mathbb{N} \end{aligned} \tag{1.1}$$

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

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

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

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 $\sigma(X_1,\ldots,X_n)$ ist der momentane Wert X_n. Aus der Turmregel folgt:

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

denn

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

Man kann von $\left(\mathcal{F}_n^X\right)_{n\in\mathbb{N}}$ 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

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

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:

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

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

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

Beispiel 1.17

- 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

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

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

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

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

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

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 $\int h \ \mathrm{d}\rho$ 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

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

Satz 1.18

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

(1) X ist Martingal.

(2) $(H \bullet X)$ ist Martingal für alle lokal beschränkten, vorhersbaren $(H_n)_{n\in\mathbb{N}}$

Das heißt, dass das stochastische Integral die Martingaleigenschaft erhält.