

WAHRSCHEINLICHKEITSTHEORIE II

Maßtheoretische Vorbemerkungen

0.1. Das Faktorisierungslemma

(Ω, \mathcal{O}) , (S, \mathcal{F}) meßbare Räume (Menge mit σ -Algebra)

Def. Eine \mathcal{O}/\mathcal{F} -meßbare Abbildung $Y: \Omega \rightarrow S$ heißt S -wertige Zufallsvariable.

meßbar: $Y^{-1}(B) \in \mathcal{O} \quad \forall B \in \mathcal{F}$

Def. $\sigma(Y) := \{Y^{-1}(B) \mid B \in \mathcal{F}\}$ „die von Y erzeugte σ -Algebra“

Bem. $\sigma(Y) \subseteq \mathcal{O} \Leftrightarrow Y$ meßbar

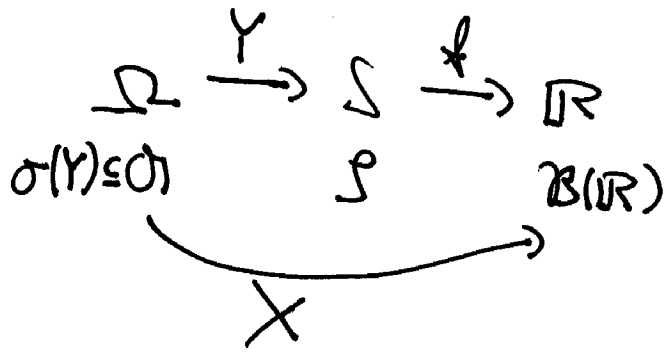
Satz (Faktorisierungslemma) Sei $Y: \Omega \rightarrow S$ meßbar. Eine Abbildung

$X: \Omega \rightarrow \mathbb{R}$ ist genau dann bzgl. $\sigma(Y)$ meßbar, wenn

$$X = f(Y) = f \circ Y$$

für eine meßbare Funktion $f: S \rightarrow \mathbb{R}$ gilt.

(2)



Beweis: $X = f \circ Y$, f messbar

$$\Rightarrow X^{-1}(C) = Y^{-1}(f^{-1}(C)) \in \sigma(Y) \quad \forall C \in \mathcal{B}(\mathbb{R})$$

$\underbrace{f^{-1}(C)}_{\in \mathcal{F}}$

d.h. X ist $\sigma(Y)$ -messbar.

Umkehrung durch maßtheoretische Induktion:

zz: X $\sigma(Y)$ -messbar $\Rightarrow X = f(Y)$

1) $X = I_A$ mit $A \in \sigma(Y)$

$$\Rightarrow A = Y^{-1}(B), \quad B \in \mathcal{F}$$

$$\Rightarrow X(\omega) = I_{Y^{-1}(B)}(\omega) = I_B(Y(\omega))$$

2) $X = \sum_{i=1}^n c_i I_{A_i}$, $A_i \in \sigma(Y)$, $c_i \in \mathbb{R}$

$$\Rightarrow X = \underbrace{\sum_{i=1}^n c_i I_{B_i}}_{=f}(Y), \quad A_i = Y^{-1}(B_i)$$

3) $X \geq 0$ $\sigma(Y)$ -meßbar

$\Rightarrow \exists X_n$ elementar $\sigma(Y)$ -meßbar: $X_n \nearrow X$

$X_n = f_n(Y) \Rightarrow X = f(Y)$ wobei $f = \sup f_n$

4) allgemein: $X = X^+ - X^-$ \square

Beispiel $(X_n)_{n \geq 0}$ stoch. Prozeß mit Werten in S

$\mathcal{F}_n := \sigma(X_0, X_1, \dots, X_n)$ „Vergangenheit bis zur Zeit n “

$Z: \Omega \rightarrow \mathbb{R}$ ist \mathcal{F}_n -meßbar $\Leftrightarrow Z = f(X_0, \dots, X_n)$ für eine meßbare Fkt. $f: S^{n+1} \rightarrow \mathbb{R}$

0.2. Erwartungswert (Ω, \mathcal{O}, P) Wahrscheinlichkeitsraum
(Menge + σ -Algebra + WV)

$X: \Omega \rightarrow \mathbb{R}$ ZV

$E[X] = \int X dP$ Lebesgueintegral

definiert durch:

$$E[X] = \sum_{i=1}^n c_i P[A_i] \text{ falls } X \text{ elementar, } X = \sum_{i=1}^n c_i I_{A_i}$$

$$E[X] = \sup_{n \in \mathbb{N}} E[X_n] \text{ falls } X \geq 0, X_n \text{ elementar mit } X_n \nearrow X$$

$$E[X] = E[X^+] - E[X^-] \text{ falls } E[X^+] < \infty \text{ oder } E[X^-] < \infty$$

$$E[X; A] = \int_A X dP := E[X \cdot I_A]$$

X durch Integrale über bestimmte Mengen eindeutig festgelegt?

Lemma $X, \tilde{X}: \Omega \rightarrow \mathbb{R}$ ZVn, ≥ 0 oder $\in \mathcal{L}^1$,

$\mathcal{A}_0 \subseteq \mathcal{A}$ \mathcal{A} -stabiles Erzeugendensystem. Dann gilt:

$$X = \tilde{X} \text{ P-f.s.} \iff E[X; A] = E[\tilde{X}; A] \forall A \in \mathcal{A}_0$$

Beweis: Übung

Beispiel $f: \mathbb{R}^n \rightarrow \mathbb{R}$ fast überall eindeutig festgelegt

$$\text{durch } \int_{[a_1, b_1] \times \dots \times [a_n, b_n]} f dx \quad (a_i, b_i \in \mathbb{R})$$

Schwerpunkt in dieser Semester:

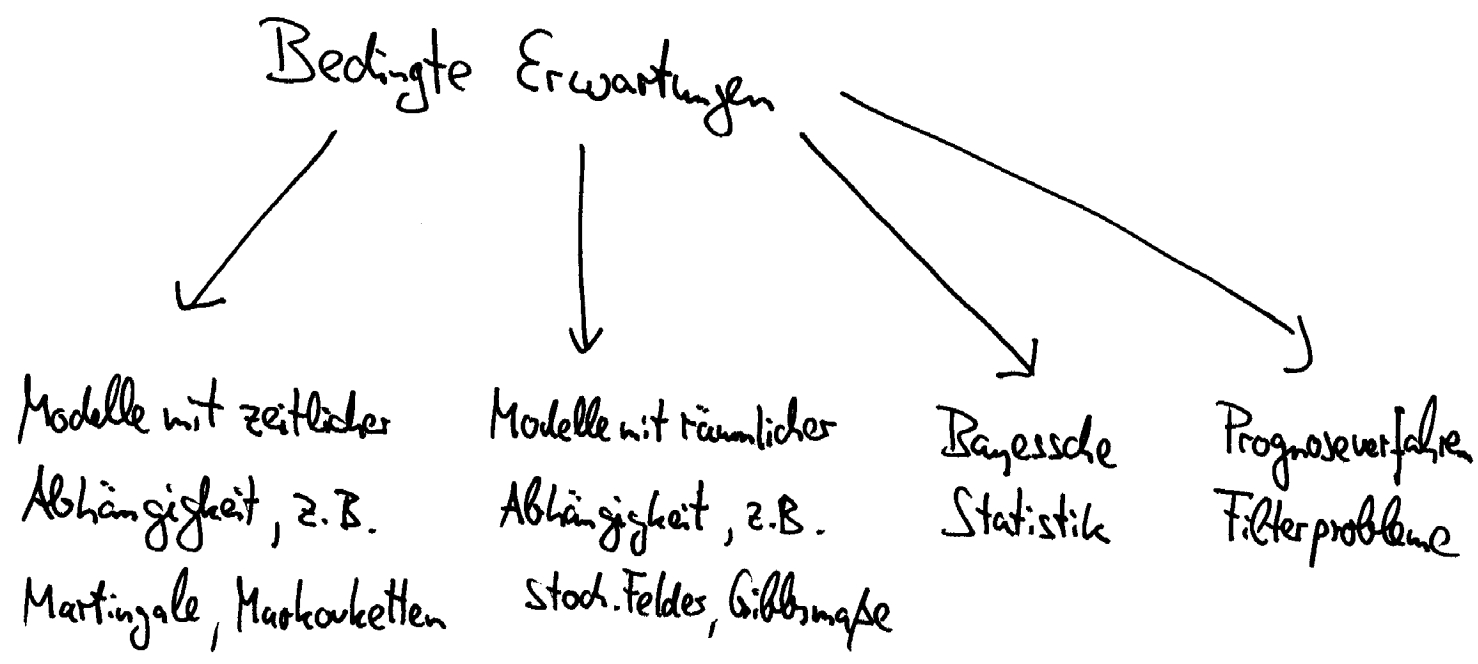
Modellierung von ABHÄNGIGEN Zufallsexperimenten
insbes. STOCHASTISCHE PROZESSE

Techniken u.a.:

- Bedingte Erwartungen
- Markoveigenschaft
- Martingale

I. BEDINGTE ERWARTUNGEN

~~UND MARTINGALE~~



X, Y Zufallsvariable, $Y(\omega)$ beobachteter Wert

Gesucht: Prognose für $X \rightarrow E[X|Y]$

1. $E[X|Y]$, Y diskrete Zufallsvariable (Ω, \mathcal{O}, P) W'raum

1.1. Wiederholung bedingte Wahrscheinlichkeiten $A, B \in \mathcal{O}$

$$P[A|B] = \frac{P[A \cap B]}{P[B]} \quad \text{bedingte W'heit gegeben } B \text{ tritt ein}$$

ist definiert falls $P[B] \neq 0$.

$P[\cdot | B]$ ist wieder WV auf (Ω, \mathcal{O}) (bedingte Verteilung)

$X: \Omega \rightarrow \mathbb{R}$ ZV, $X \geq 0$ oder $X \in \mathcal{L}^1(\Omega, \mathcal{O}, P)$:

$$E[X|B] = \int X \, dP[\cdot | B] = \frac{E[X; B]}{P[B]} = \frac{E[X \cdot I_B]}{P[B]}$$

Beste Prognose für X , wenn wir wissen daß $\omega \in B$.

Bem. X unabh. von $B \Rightarrow E[X|B] = \frac{E[X] \cdot E[I_B]}{P[B]} = E[X]$.

1.2. Wichtige Anwendung: Bayessche Statistik

$\Omega = \bigcup_{i \in I} H_i$ disjunkte Zerlegung in Hypothesen

Wie wahrscheinlich sind die Hypothesen?

Unbedingt: $P[H_i]$ a priori W'keiten

Bedingt: Sei $A \in \mathcal{O}$ festes Ereignis mit $P[A] \neq 0$.

$P[A|H_i]$ likelihood von A gegeben H_i tritt ein

$$P[H_i|A] = \frac{P[H_i] \cdot P[A|H_i]}{P[A]} = \frac{P[H_i] \cdot P[A|H_i]}{\sum_{j \in I} P[H_j] \cdot P[A|H_j]}$$

a posteriori
W'keit für H_i
gegeben A tritt ein

$$= \underbrace{\text{const.}}_{\text{unabh. von } i} \cdot \underbrace{P[H_i]}_{\text{a priori}} \cdot \underbrace{P[A|H_i]}_{\text{likelihood}}$$

BAYESSCHE FORMEL

Problem in der Praxis: Berechnung der Konstanten

- I sehr groß \rightarrow exakte Berechnung unmöglich
- \rightarrow Näherungsverfahren, Markov chain Monte Carlo

1.3. Bedingte Erwartung von X gegeben Y - diskreter Fall

$$X: \Omega \rightarrow \mathbb{R} \text{ ZV}$$

$$Y: \Omega \rightarrow S \text{ weitere ZV wobei } \underline{\text{abzählbar}} \text{ mit } \sigma\text{-Alg. } \mathcal{F} = \mathcal{P}(S)$$

Gesucht: beste Prognose für X wenn wir $Y(\omega)$ kennen

Falls $Y(\omega) = z$ mit $P[Y=z] > 0$:

$$\text{beste Prognose für } X = \underbrace{E[X|Y=z]}$$

nicht definiert falls $P[Y=z] = 0$!

Def. / Lemma Durch

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

wobei

$$g(z) := \begin{cases} E[X|Y=z] & \text{falls } P[Y=z] > 0 \\ \text{beliebig (z.B. } 0) & \text{sonst} \end{cases}$$

wird P-f.s. eindeutig eine Zufallsvariable $E[X|Y]$ definiert.

Jede so definierte ZV heißt Version der bedingten Erwartung von X gegeben Y .

Beweis: $E[X|Y]$ ist eindeutig festgelegt auf $\bigcup_{z \in S} \{Y=z\}$.

Diese Menge hat Maß 1. \square

$P\{Y=z\} > 0$

Also:

- $E[X|Y](\omega)$ ist eine Zufallsvariable.
- Diese ordnet jedem ω die beste Prognose für X unter der Voraussetzung, daß der Wert von Y gleich $Y(\omega)$ ist zu.
- $E[X|Y]$ ist i.a. nur P-f.s. eindeutig festgelegt (eindeutige Äquivalenzklasse von ZVn).

1.4. Bedingte Wahrscheinlichkeit $P[A|Y]$ Sei $A \in \mathcal{A}$.

Def. $P[A|Y] := E[I_A | Y]$

ist Zufallsvariable, P-f.s. eindeutig.

1.5. Beispiele1) X, Y unabhängig

$$\Rightarrow E[X|Y=z] = \frac{E[X \cdot \overset{\text{unabh.}}{I_{\{Y=z\}}}] }{P[Y=z]} = E[X]$$

für alle z mit $P[Y=z] \neq 0$.

$$\Rightarrow E[X|Y](\omega) = E[X] \quad P\text{-f.s.}$$

Kenntnis von Y liefert keine zusätzliche Information über X

→ beste Prognose ist weiterhin Erwartungswert

2) X_1, X_2, \dots iid mit $E[X_i] = m$, $S_n = X_1 + \dots + X_n$ (z.B. random walk)

$$E[S_{n+1} | S_n = z] \stackrel{S_n, X_{n+1} \text{ unabh.}}{=} E[z + X_{n+1}] = z + m \quad (\text{falls } P[S_n = z] \neq 0)$$

$$\underbrace{\quad}_{= S_n + X_{n+1} = z + X_{n+1}}$$

$$\Rightarrow E[S_{n+1} | S_n](\omega) = S_n(\omega) + m \quad P\text{-f.s.}$$

$$\underline{m=0}: E[S_{n+1} | S_n] = S_n \quad \text{„fares Spiel“ (Martingal)}$$

3) Markov-Eigenschaft:

$$P[X_{n+1}=y | X_0=x_0, \dots, X_n=x_n] = P[X_{n+1}=y | X_n=x_n]$$

für alle x_0, \dots, x_n mit $P[X_0=x_0, \dots, X_n=x_n] \neq 0$.

Effiziente Formulierung:

$$P[X_{n+1}=y | X_0, \dots, X_n] = P[X_{n+1}=y | X_n] \quad \text{P-f.s.}$$

wobei $P[X_{n+1}=y | X_0, \dots, X_n] := P[X_{n+1}=y | \underbrace{(X_0, \dots, X_n)}_{\text{Zufallsvariable mit Werten in } S^{n+1}}] = g(X_0, \dots, X_n)$

Zufallsvariable
mit Werten in S^{n+1}

1.6. Eigenschaften der bedingten Erwartung

Satz 1) $E[f(Y) \cdot X | Y] = f(Y) \cdot E[X | Y] \quad \text{P-f.s.}$

$\forall f: S \rightarrow \mathbb{R}$ meßbar mit $f(Y) \cdot X \geq 0$ oder $\in \mathcal{L}^1$.

„Heranziehen was bekannt ist“

2) $E[E[X | Y] | Y] = E[X | Y] \quad \text{P-f.s.}$

„Projektivität“

Bem. $Z: \Omega \rightarrow \mathbb{R}$ $\sigma(Y)$ -messbar bzgl. $\sigma(Y)$

Fakt. lemma
 \Rightarrow

$$Z = f(Y)$$

$\stackrel{!}{\Rightarrow}$

$$E[Z \cdot X | Y] = Z \cdot E[X | Y] \text{ P-f.s.}$$

(falls definiert)

Insbesondere:

$$E[Z | Y] = Z \text{ P-f.s.}$$

„Wenn wir Y kennen, dann kennen wir auch Z (da $\sigma(Y)$ -messbar)
- daher ist Z selbst die beste Prognose für Z gegeben Y .“

Beweis Satz: 1) Sei $z \in S$ mit $P\{Y=z\} > 0$. Auf $\{Y=z\}$ gilt:

$$\text{l.S.} = E[\underbrace{f(Y)}_{=f(z)} \cdot X | Y=z] = \underbrace{f(z)}_{=f(Y)} \cdot E[X | Y=z] = \text{r.S.}$$

2) $E[X | Y] = g(Y)$, also:

$$\text{l.S.} = E[g(Y) | Y] \stackrel{!}{=} g(Y) \cdot \underbrace{E[1 | Y]}_{=1} = E[X | Y].$$

□

1.7. Charakterisierende Eigenschaft der bedingten Erwartung

Satz Sei $X \geq 0$. Eine Zufallsvariable $\bar{X} \geq 0$ ist genau dann eine Version der bedingten Erwartung $E[X|Y]$ wenn gilt:

- 1) \bar{X} ist $\sigma(Y)$ -messbar.
- 2) $E[\bar{X}; A] = E[X; A] \quad \forall A \in \sigma(Y)$

(bew. 2') $E[\bar{X} \cdot Z] = E[X \cdot Z] \quad \forall \sigma(Y)$ -messbaren $Z \geq 0$)

Insbesondere hängt $E[X|Y]$ nur von $\sigma(Y)$ ab!

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

ERMÖGLICHT DEFINITION DER BEDINGTEN ERWARTUNG IM KONTINUIERLICHEN FALL!

Kor. $E[E[X|Y]] = E[X]$

Insbesondere: $E[\underbrace{E[X|Y]}_{\leq E[X|Y]}] \leq E[E[X|Y]] = E[X]$
Bedingte Erwartung ist Kontraktion auf L^1

Beweis: a) \bar{X} Version von $E[X|Y] \Rightarrow \bar{X} = g(Y)$, also $\sigma(Y)$ -messbar.

Zudem: $A \in \sigma(Y) \Rightarrow A = Y^{-1}(B) = \{Y \in B\}$ für ein $B \in \mathcal{F}$

$$\Rightarrow E[\bar{X}; A] = \sum_{\substack{z \in B \\ P\{Y=z\} > 0}} E[\bar{X}; Y=z] = \sum_{\substack{z \in B \\ P\{Y=z\} > 0}} E[X|Y=z] \cdot P\{Y=z\}$$

$$= E[X|Y=z]$$

$$= E[X; Y \in B] = E[X; A].$$

b) Umgekehrt: Erfüllt \bar{X} D und 2), dann gilt

$$E[\bar{X}; A] = E[X; A] = E[E[X|Y]; A] \quad \forall A \in \sigma(Y)$$

Da \bar{X} und $E[X|Y]$ beide $\sigma(Y)$ -messbar sind, folgt $\bar{X} = E[X|Y]$ P-f.s. \square

1.8. Beispiele Seien X_1, X_2, \dots iid mit $E[X_i] = \mu$, $S_n = X_1 + \dots + X_n$.

a) Beh.: $E[X_i | S_n] = \frac{1}{n} S_n$ P-f.s. $\forall 1 \leq i \leq n$

Beweis: z.z.: $\frac{1}{n} S_n$ erfüllt D und 2) aus Satz 1.7.

1) \checkmark $\frac{1}{n} S_n$ ist $\sigma(S_n)$ -messbar.

2) $Z \geq 0$ $\sigma(\mathcal{I}_n)$ -messbar $\xrightarrow{\text{Fakt. Lemma}}$ $Z = f(\mathcal{I}_n)$, $f \geq 0$ messbar (16)

$$\Rightarrow E[X_i \cdot Z] = E[X_i \cdot f(X_1 + \dots + X_n)] \stackrel{X_i \text{ i.i.d.}}{=} E[X_j \cdot f(X_1 + \dots + X_n)]$$

für alle $1 \leq j \leq n$

$$\Rightarrow E[X_i \cdot Z] = \frac{1}{n} \sum_{j=1}^n E[X_j \cdot Z] = E\left[\frac{1}{n} \mathcal{I}_n \cdot Z\right] \quad \square$$

b) Sei $N: \Omega \rightarrow \{0, 1, 2, \dots\}$ weitere ZV, unabh. von X_1, X_2, \dots

$$S_{N(\omega)}(\omega) = X_1(\omega) + X_2(\omega) + \dots + X_{N(\omega)}(\omega)$$

Zufällige Summe (mit zufälliger Summationsgrenze)

$$E[S_N] = ?$$

$$E[S_N | N=n] = E[X_1 + \dots + X_n | N=n] \stackrel{\text{unabh.}}{=} E[X_1 + \dots + X_n] = n \cdot m$$

$$\Rightarrow E[S_N | N] = N \cdot m \quad \text{P-f.s.}$$

$$\Rightarrow \underline{\underline{E[S_N]}} \stackrel{\text{Kor. 1.7}}{=} E[E[S_N | N]] = \underline{\underline{E[N] \cdot m}}$$

Spezialfall der „Waldschen Identität“

2. Bedingte Erwartungen: Der allgemeine Fall

(Ω, \mathcal{O}, P) W'raum, (S, \mathcal{F}) messbarer Raum

$X: \Omega \rightarrow \mathbb{R}$ ZV, $X \geq 0$ oder $\in \mathcal{L}^1$.

$Y: \Omega \rightarrow S$ ZV, z.B. $S = \mathbb{R}$.

$E[X|Y=z]$ ist i.a. nicht definiert, da $P[Y=z] = 0$ möglich.

Definition von $E[X|Y]$?

Im diskreten Fall ist $\bar{X} := E[X|Y]$ P.f.s. eindeutig festgelegt durch (1.7, 1.10d).

2.1. Definition von $E[X|Y]$ und $E[X|\mathcal{F}]$

Def. Sei $\mathcal{F} \subseteq \mathcal{O}$ eine σ -Algebra (z.B. $\mathcal{F} = \sigma(Y)$).

Eine ZV $\bar{X} \geq 0$ (bzw. $\bar{X} \in \mathcal{L}^1$) heißt Version der bedingten Erwartung

$E[X|\mathcal{F}]$ falls:

(i) \bar{X} ist \mathcal{F} -messbar und

(ii) $E[\bar{X}; A] = E[X; A] \quad \forall A \in \mathcal{F}$.

Insbesondere:

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

$$E[X|Y_1, \dots, Y_n] := E[X|\sigma(Y_1, \dots, Y_n)]$$

Satz (Existenz und Eindeutigkeit) Sei $X \geq 0$ oder $X \in \mathcal{L}^1$.

- 1) Es existiert eine Version der bedingten Erwartung $E[X|F]$.
- 2) Zwei Versionen stimmen \mathbb{P} -f.s. überein.

Beweis: 2) folgt aus Lemma 0.2. 1) s.u. \square

2.2. Diskreter Spezialfall

$\mathcal{F} = \sigma(B_i, i \in I)$ wobei I abzählbar, B_i disjunkt mit $\Omega = \bigcup_{i \in I} B_i$

$$\Rightarrow E[X|\mathcal{F}](\omega) = \sum_{\substack{i \in I \\ \mathbb{P}[B_i] \neq 0}} E[X|B_i] \cdot \frac{1}{\mathbb{P}[B_i]} \cdot \mathbb{I}_{B_i}(\omega) \quad \mathbb{P}\text{-f.s.}$$

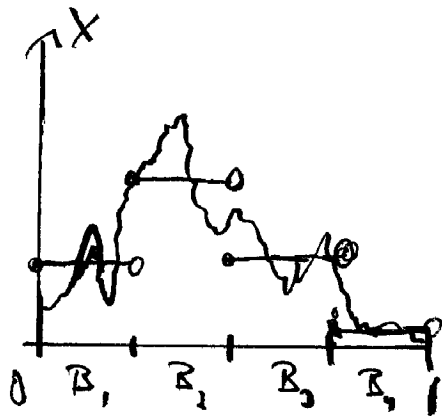
Bsp. $(\Omega, \mathcal{O}, P) = ([0,1), \mathcal{B}([0,1)), \lambda)$.

$$\mathcal{F} = \sigma\left(\left[\frac{i-1}{n}, \frac{i}{n}\right) \mid 1 \leq i \leq n\right)$$

\mathcal{F} -meßbare Funktionen sind konstant auf $[\frac{i-1}{n}, \frac{i}{n})$ $\forall i$

$X: [0,1) \rightarrow \mathbb{R}$ int. bar

$$E[X | \mathcal{F}] = \underbrace{n \cdot \int_{\frac{i-1}{n}}^{\frac{i}{n}} X(s) ds}_{\text{Mittelwert}} \text{ auf } \left[\frac{i-1}{n}, \frac{i}{n}\right)$$



2.3. Bedingte Wahrscheinlichkeit $A \in \mathcal{O}$

Def. $P[A | \mathcal{F}](\omega) := E[I_A | \mathcal{F}](\omega)$ P.f.s. eindeutig festgelegt

im Beispiel: $P\left[\left(\frac{3}{4}, 1\right) \mid \sigma\left(\left[0, \frac{1}{2}\right), \left[\frac{1}{2}, 1\right)\right)\right]_{(\omega)} = \begin{cases} 0 & \text{für } \omega \in [0, \frac{1}{2}) \\ \frac{1}{2} & \text{für } \omega \in [\frac{1}{2}, 1) \end{cases}$

2.4. Eigenschaften der bedingten Erwartung

Satz Seien X, Y, Z Zufallsvariablen, $\in \mathcal{L}^1(\mathcal{D}, \mathcal{O}, P)$ oder ≥ 0 .

- 1) $X=Y$ P-f.s. $\Rightarrow E[X|\mathcal{F}] = E[Y|\mathcal{F}]$ P-f.s.
- 2) $E[E[X|\mathcal{F}]] = E[X]$
- 3) Linearität: $E[\lambda X + \mu Y|\mathcal{F}] = \lambda E[X|\mathcal{F}] + \mu E[Y|\mathcal{F}]$ P-f.s. $\forall \lambda, \mu \in \mathbb{R}$
- 4) Positivität: $X \geq 0$ P-f.s. $\Rightarrow E[X|\mathcal{F}] \geq 0$ P-f.s.
- 5) Projektivität: $\mathcal{F} \subseteq \mathcal{G}$ σ -Algebren $\Rightarrow E[E[X|\mathcal{G}]|\mathcal{F}] = E[X|\mathcal{F}]$ P-f.s.
insbesondere: $E[E[X|Y, Z]|Y] = E[X|Y]$ P-f.s.
- 6) Herausziehen was bekannt ist: Z \mathcal{F} -messbar mit $Z \cdot X \in \mathcal{L}^1$ bzw. ≥ 0
 $\Rightarrow E[Z X|\mathcal{F}] = Z \cdot E[X|\mathcal{F}]$ P-f.s.
- 7) Unabhängigkeit: $\sigma(X)$ \mathcal{F} unabhängig
 $\Rightarrow E[X|\mathcal{F}] = E[X]$ P-f.s.
- 8) Y \mathcal{F} -messbar, $\sigma(X)$ unabhängig von \mathcal{F}
 $\Rightarrow E[f(X, Y)|\mathcal{F}](\omega) = E[f(X, Y(\omega))]$ $\forall f \geq 0$ messbar

Beweis:

1) z.z.: $X=Y$ P-f.s. \Rightarrow Jede Version von $E[Y|\mathcal{F}]$
ist auch eine Version von $E[X|\mathcal{F}]$

(i) \mathcal{F} -messbar \checkmark

(ii) $A \in \mathcal{F} \Rightarrow E[E[Y|\mathcal{F}]; A] = E[Y; A] = E[X; A] \checkmark$

2), 3) Übung

4) Sei \bar{X} Version von $E[X|\mathcal{F}]$.

$$X \geq 0 \text{ P-f.s.} \Rightarrow E[\underbrace{\bar{X}}_{\in \mathcal{F}}; \underbrace{\bar{X} < 0}_{\geq 0}] = E[X; \bar{X} < 0] \geq 0$$

$$\Rightarrow E[\bar{X}^-] = 0 \Rightarrow \bar{X}^- = 0 \text{ P-f.s.} \Rightarrow \bar{X} \geq 0 \text{ P-f.s.}$$

5) z.z.: Jede Version von $E[X|\mathcal{F}]$ ist auch Version von $E[E[X|\mathcal{G}]|\mathcal{F}]$:

(i) \mathcal{F} -messbar \checkmark

(ii) $A \in \mathcal{F} \subseteq \mathcal{G} \Rightarrow E[E[X|\mathcal{F}]; A] = E[X; A] = E[E[X|\mathcal{G}]; A]$

6), 7) Übung

8) a) ok falls $f(x, y) = g(x) \cdot h(y)$, $g, h \geq 0$ meßbar:

$$E[f(X, Y) | \mathcal{F}](\omega) = E[g(X) \cdot \underbrace{h(Y)}_{\mathcal{F}\text{-meßbar}} | \mathcal{F}](\omega)$$

$$\stackrel{6)}{=} \underbrace{h(Y(\omega))}_{\text{unabh. von } \mathcal{F}} \cdot E[g(X) | \mathcal{F}](\omega) \stackrel{7)}{=} \underbrace{h(Y(\omega)) \cdot E[g(X)]}_{\text{P-f.s.}}$$

$$= E[f(X, Y(\omega))] (\omega) \quad \text{P-f.s.}$$

b) auch ok falls $f(x, y) = I_B(x, y)$, $B \in \mathcal{B}(\mathbb{R}^2)$:

Sei $\mathcal{D} := \{ B \in \mathcal{B}(\mathbb{R}^2) \mid \text{Bel. gilt für } f(x, y) = I_B(x, y) \}$

ist Dynkinsystem, $\stackrel{a)}{\supseteq} \{ B_1 \times B_2 \mid B_1, B_2 \in \mathcal{B}(\mathbb{R}) \}$,

also $\supseteq \sigma(B_1 \times B_2 \mid B_1, B_2 \in \mathcal{B}(\mathbb{R})) = \mathcal{B}(\mathbb{R}^2)$.

c) maßtheoret. Induktion \rightarrow ok für alle $f \geq 0$ meßbar. \square

2.5. Beispiel

(23)

T, S unabhängig exponentialverteilt mit $\lambda > 0$

z.B. Ausfallzeiten von zwei Komponenten eines Systems

$$X := \min(T, S)$$

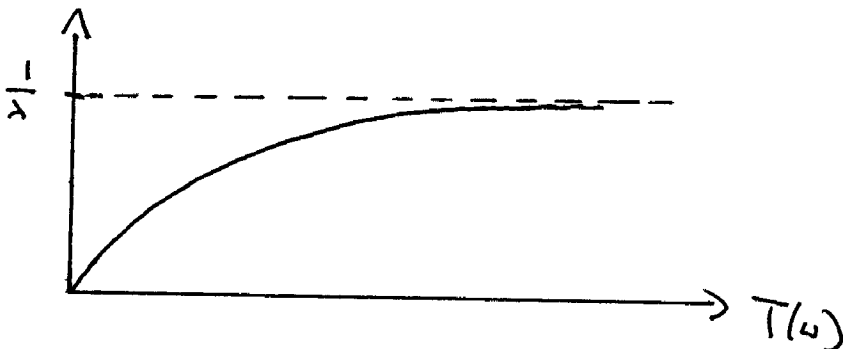
Annahme: nur T ist direkt beobachtbar - Prognose für X ?

$$\begin{aligned} \underline{\underline{E[X|T]}(\omega)} &\stackrel{2.4.8}{=} E[\min(T(\omega), S)] \\ &= \int_0^{\infty} \min(T(\omega), s) \lambda \cdot e^{-\lambda s} ds \\ &= \int_0^{T(\omega)} s \lambda e^{-\lambda s} ds + \int_{T(\omega)}^{\infty} T(\omega) \lambda e^{-\lambda s} ds = \underline{\underline{\frac{1}{\lambda} (1 - e^{-\lambda T(\omega)})}} \end{aligned}$$

P-f.s.

Beispiel für eine nichtlineare Prognose

(\neq beste lineare Prognose $\hat{X} = aT + b$)



2.6. Konvergenzsätze

(24)

Satz über monotone Konvergenz: $X_n: \Omega \rightarrow \mathbb{R} \geq 0$, $0 \leq X_1 \leq X_2 \leq \dots$

$$\Rightarrow E[\sup X_n | \mathcal{F}] = \sup E[X_n | \mathcal{F}] \quad \mathcal{F}\text{-f.s.}$$

Beweis: (i) $\sup E[X_n | \mathcal{F}]$ ist \mathcal{F} -messbar

(ii) $Y \geq 0$ \mathcal{F} -messbar

$$\begin{aligned} \Rightarrow E\left[Y \cdot \underbrace{\sup E[X_n | \mathcal{F}]}_{\substack{\geq 0 \text{ f.s.} \\ \text{mon. wachsend f.s.}}} \right] & \stackrel{\text{Mon.kov.}}{=} \sup E\left[Y \cdot E[X_n | \mathcal{F}] \right] \stackrel{\text{Mon.kov.}}{=} E\left[Y \cdot \sup X_n \right] \\ & = E\left[Y \cdot X_n \right] \quad \square \end{aligned}$$

Kor. Lemma von Fatou, Satz über dominierte Konvergenz (Übung)

2.7. Definition von $E[X|Y=z]$ im allgemeinen Fall

Sei $Y: \Omega \rightarrow S$ messbar.

$$E[X|Y] = \underbrace{E[X | \sigma(Y)]}_{\sigma(Y)\text{-messbar}} \stackrel{\text{Faktorisierungslemma}}{=} g(Y) \quad \text{für eine messbare Fkt. } g: S \rightarrow \mathbb{R}$$

Def. $E[X|Y=z] := g(z)$

WOHLDEFINIERT?

Nicht für jedes z , aber μ -fast sicher, wobei $\mu = P \circ Y^{-1}$:

Sei sowohl $g(Y)$ als auch $\tilde{g}(Y)$ Version von $E[X|Y]$

$$\Rightarrow \mu(g = \tilde{g}) = P[g(Y) = \tilde{g}(Y)] = 1$$

Alternative Notationen in der Literatur:

$$E[X|Y=z] = E[X | z \leq Y \leq z+dz] = E[X | Y \in dz]$$

Def. $P[A|Y=z] := E[I_A | Y=z]$ μ -f.s. eindeutig

2.8. Absolute Wahrscheinlichkeiten aus bedingten

Sei $Y: \mathcal{R} \rightarrow \mathcal{S}$ meßbar mit Verteilung $\mu = \mathbb{P} \circ Y^{-1}$.

Kor. $E[X] = \int E[X|Y=z] \mu(dz) \quad \forall X \geq 0$ meßbar

Identität

Insbesondere: $P[A] = \int P[A|Y=z] \mu(dz) \quad \forall A \in \mathcal{O}$

Beweis: $E[X] \stackrel{2.4.2}{=} E[\underbrace{E[X|Y]}] \stackrel{\text{Transf. Satz}}{=} \int g(z) \mu(dz).$
 $= g(Y)$ wobei $g(z) = E[X|Y=z]$

□

2.9. Der Poissonprozeß $\lambda > 0$

Seien S_1, S_2, \dots i.i.d. exponentialverteilt mit λ .

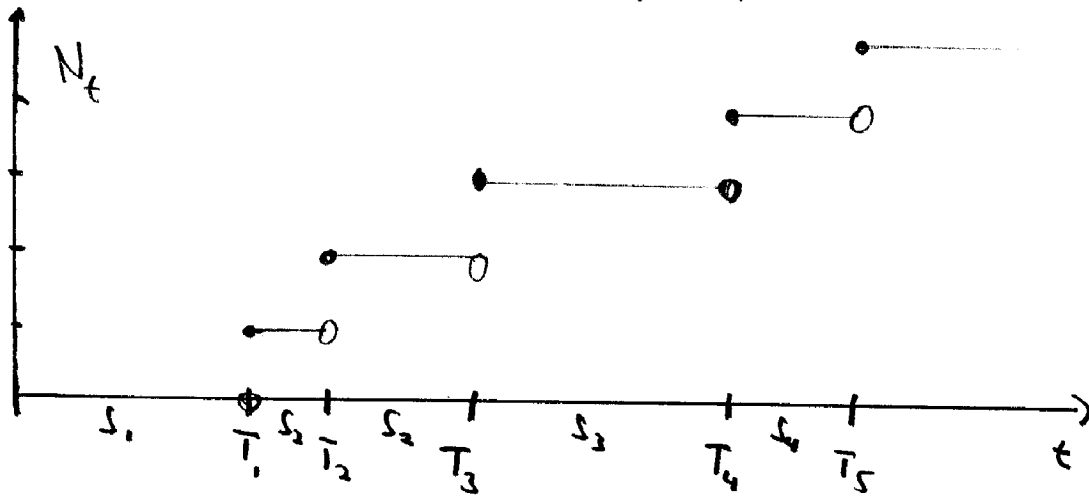
S_i = Wartezeit auf das Eintreten des i -ten Ereignisses
(z.B. Bedienung des i -ten Kunden in Warteschlange)

$T_n := S_1 + \dots + S_n$ Gesamt-wartezeit bis zum n -ten Ereignis

T_n ist $\Gamma(n, \lambda)$ verteilt (n-fache Faltung von $\text{Exp}(\lambda)$),

d. h. Dichte

$$f_{T_n}(t) = \frac{\lambda^n t^{n-1} e^{-\lambda t}}{(n-1)!}, \quad t \geq 0.$$



Sei $t \geq 0$.

$$N_t := \# \{n \mid T_n \leq t\} = \sum_{n=1}^{\infty} I_{\{T_n \leq t\}}$$

Anzahl der bis t eingetretenen Ereignisse.

$(N_t)_{t \geq 0}$ heißt Poissonprozess mit Parameter λ .

(stochastischer Prozess in kontinuierlicher Zeit)

Bem. Aus der „Gedächtnislosigkeit“ der Exponentialverteilung folgt, daß $(N_t)_{t \geq 0}$ ein Markovprozess ist; d.h.

$$s \leq t \Rightarrow P[N_t = n \mid \sigma(N_u \mid 0 \leq u \leq s)] \stackrel{\omega}{=} P[N_t = n \mid N_s] \stackrel{\omega}{=} \prod_{k=1}^n P[N_{t-s} = n - N_s(\omega)]$$

Berechnung der Verteilung von N_t , t fest:

$$\underline{\underline{P[N_t = n] = P[T_n \leq t < T_{n+1}]}}$$

$$\stackrel{2.8}{=} \int_0^{\infty} \underbrace{P[T_n \leq t < T_{n+1} | T_n = u]}_{\substack{= u \\ = u + S_{n+1} \leftarrow \text{unabh. von } T_n = S_1 + \dots + S_n}} f_{T_n}(u) du$$

$$\stackrel{(*)}{=} \int_0^{\infty} \underbrace{P[u \leq t \leq u + S_{n+1}]}_{= e^{-\lambda(t-u)} \cdot \mathbb{I}_{[0,t]}(u)} \underbrace{f_{T_n}(u)}_{= \frac{\lambda^n u^{n-1} e^{-\lambda u}}{(n-1)!}} du = \underline{\underline{\frac{(\lambda t)^n}{n!} e^{-\lambda t}}}$$

d.h. N_t ist Poisson (λt) -verteilt.

Bem. Die Ausföhrung von (*) ist so nicht ganz rigoros (aber richtig!).

Übung: Zeige ausgehend von den Definitionen, daß (*) gilt.

2.10. Bedingte Varianz Sei $X \in \mathcal{L}^2$, $\mathcal{F} \subseteq \mathcal{O}$ σ -Algebra

Def. $\text{var}(X|\mathcal{F}) := E[(X - E[X|\mathcal{F}])^2 | \mathcal{F}]$

heißt bedingte Varianz von X gegeben \mathcal{F} .

Bem. $\text{var}(X|\mathcal{F}) = E[X^2|\mathcal{F}] - 2E[X \cdot E[X|\mathcal{F}] | \mathcal{F}] + E[(E[X|\mathcal{F}])^2 | \mathcal{F}]$
 $= E[X^2|\mathcal{F}] - E[E[X|\mathcal{F}]^2 | \mathcal{F}]$ P-f.s.

Satz (Formel von der bedingte Varianz):

$$\text{var}(X) = E[\text{var}(X|\mathcal{F})] + \text{var}(E[X|\mathcal{F}]) \quad !$$

Beweis:

*.S. $\stackrel{\text{Be.}}{=} E[E[X^2|\mathcal{F}]] - E[E[X|\mathcal{F}]^2] + E[E[X|\mathcal{F}]^2] - E[E[X|\mathcal{F}]]^2$
 $= E[X^2] - E[X]^2 = \text{var}(X). \quad \square$

2.11. Beispiel: Varianz einer zufälligen Summe

$$S_N(\omega) := \sum_{i=1}^{N(\omega)} X_i(\omega) \quad \text{wobei } N, X_1, X_2, \dots \in \mathcal{L}^2(\mathcal{R}, \mathcal{O}, P) \text{ unabhängig,}$$

X_i ident. verteilt mit $E[X_i] = \mu$, $\text{var}(X_i) = \sigma^2$, $N: \mathcal{R} \rightarrow \{0, 1, 2, \dots\}$.

(z. B. nächste Generation im Verzweigungsprozeß).

$$E[S_N | N] = N \cdot m \Rightarrow E[S_N] = E[N] \cdot m, \text{ s.o. (1.8).}$$

$$\text{var}(S_N | N=n) = \text{var}(S_n | N=n) \stackrel{\text{unabh.}}{=} \text{var}(S_n) = n \cdot \sigma^2$$

$$\Rightarrow \text{var}(S_N | N) = N \cdot \sigma^2$$

$$\begin{aligned} \Rightarrow \underline{\underline{\text{var}(S_N)}} &= E[\text{var}(S_N | N)] + \text{var}(E[S_N | N]) \\ &= \underline{\underline{E[N] \cdot \sigma^2 + \text{var}(N) \cdot m^2}} \end{aligned}$$

3. Bedingte Verteilungen und bedingte Dichten

3.1. Absolutstetiger Fall

Sei (Ω, \mathcal{A}, P) W'raum, $X: \Omega \rightarrow \mathbb{R}^n$, $Y: \Omega \rightarrow \mathbb{R}^m$ Zufallsvariablen.

Annahme: Die gemeinsame Verteilung $P_{(X,Y)}$ von X und Y ist absolutstetig mit Dichte $f_{X,Y}: \mathbb{R}^{n+m} \rightarrow [0, \infty)$:

$$P[X \in A, Y \in B] = \int_A \int_B f_{X,Y}(x,y) dy dx$$

\Rightarrow Die Verteilung von Y ist absolutstetig mit Dichte

$$f_Y(y) = \int_{\mathbb{R}^n} f_{X,Y}(x,y) dx$$

Bedingte Verteilung von X gegeben Y ?

$$\begin{aligned} \text{Heuristik: } P[X \in (x, x+dx) \mid Y \in (y, y+dy)] &= \frac{P[X \in (x, x+dx), Y \in (y, y+dy)]}{P[Y \in (y, y+dy)]} \\ &= \frac{f_{X,Y}(x,y) dx dy}{f_Y(y) dy} = \frac{f_{X,Y}(x,y)}{f_Y(y)} dx \end{aligned}$$

Def. $f_{X|Y}(x,y) := \begin{cases} \frac{f_{X,Y}(x,y)}{f_Y(y)} & \text{falls } f_Y(y) \neq 0 \\ 0 & \text{sonst} \end{cases}$

heißt bedingte Dichte von X gegeben Y.

Satz $E[h(X,Y) | Y](\omega) = \int h(x, Y(\omega)) f_{X|Y}(x, Y(\omega)) dx$ P-f.s.

bzw. $E[h(X,Y) | Y=z] = \int h(x,z) f_{X|Y}(x,z) dx$ μ -f.s., $\mu = P \circ Y^{-1}$.

insbesondere: $P[X \in B | Y] = \int_B f_{X|Y}(x, Y) dx$ P-f.s.

↳ bedingte Verteilung von X gegeben Y

Beweis: (i) $f_{X|Y}$ ist meßbar auf \mathbb{R}^{n+m}

$\xRightarrow{\text{Fubini}} \bar{h}(y) = \int h(x,y) f_{X|Y}(x,y) dx$ meßbar auf \mathbb{R}^m

\Rightarrow r.s. = $\bar{h} \circ Y$ ist $\sigma(Y)$ -meßbar

(ii) $E[g(Y) \cdot \text{r.s.}] \stackrel{\text{Fubini}}{=} \int_{\mathbb{R}} E[g(Y) h(x,Y) f_{X|Y}(x,Y)] dx$

$= \int \int g(y) h(x,y) \underbrace{f_{X|Y}(x,y) f_Y(y)}_{f_{X,Y}(x,y)} dy dx = E[g(Y) \cdot h(X,Y)]$
 für alle $g \geq 0$ meßbar. \square

3.3. Bayessche Formel

Gegeben: - "a priori" Verteilung von X mit Dichte $\pi(x) = f_X(x)$
 - bedingte Dichte $\pi(y|x) = f_{Y|X}(y,x) > 0$

Gesucht: "a posteriori" Verteilung von X gegeben Y

$$\underline{\underline{f_{X|Y}(x,y) = \frac{f_X(x) \cdot f_{Y|X}(y,x)}{f_Y(y)} = \frac{f_X(x) \cdot f_{Y|X}(y,x)}{\int f_X(x) \cdot f_{Y|X}(y,x) dx}}}$$

$$\pi(x|y) = \text{const.} \cdot \pi(x) \cdot \pi(y|x)$$

a posteriori Dichte hängt nur von y ab a priori Dichte bedingte Dichte "likelihood"

3.4. Verrauschte Signale

X : un beobachtetes Signal, $X \sim N(\mu, \sigma^2)$ ($\mu \in \mathbb{R}, \sigma > 0$)

Y : Beobachtung - verrauschte Störung von X

$Y = X + Z$ wobei Z unabh. von X , $\sim N(0, \epsilon)$ ($\epsilon > 0$)

Bedingte Verteilung von Y gegeben $X(\omega)$: $N(X(\omega), \varepsilon)$

(denn: $P[X+Z \in B | X](\omega) \stackrel{2.48}{=} P[X(\omega)+Z \in B] = N(X(\omega), \varepsilon)(B)$)

$$\text{Also: } f_{Y|X}(y, x) = \frac{1}{\sqrt{2\pi\varepsilon}} e^{-\frac{(y-x)^2}{2\varepsilon}}$$

Gesucht: Bedingte Verteilung von X gegeben Beobachtung $Y(\omega)$

$$\begin{aligned} \text{Bayes: } \log f_{X|Y}(x, y) &= \log f_X(x) + \log f_{Y|X}(y, x) \Rightarrow \log f_Y(y) \\ &= \underbrace{\text{const.} - \frac{(x-\mu)^2}{2\mu}}_{\text{const.}} + \underbrace{\text{const.} - \frac{(y-x)^2}{2\varepsilon}}_{\text{const.}} \end{aligned}$$

$$= c_1(y) - \frac{x^2}{2\mu} + \frac{\mu x}{\mu} - \frac{x^2}{2\varepsilon} + \frac{yx}{\varepsilon}$$

$$= c_2(y) - \frac{(x-\hat{x})^2}{2\nu}$$

$$\text{wobei } \frac{1}{\nu} = \frac{1}{\mu} + \frac{1}{\varepsilon}, \quad \hat{x} = \frac{\mu}{\nu} + \frac{y}{\nu}$$

$$\Rightarrow \boxed{f_{X|Y}(x, y) = c_3(y) \cdot e^{-\frac{(x-\hat{x})^2}{2\nu}}}$$

A posteriori Verteilung von X geg. $Y(\omega)=y$ ist $N(\hat{x}, \nu)$.

Insbesondere:

1) \hat{X} definiert durch $\frac{\hat{X}}{v} = \frac{\mu}{u} + \frac{Y}{v}$ ist Prognosewert für X
 \rightarrow lineare Prognose!

2) mittlerer quadratischer Prognosefehler:

$$E[(X - \hat{X})^2] = E\left[\underbrace{E[(X - \hat{X})^2 | Y]}_{\sim N(0, v)}\right] = v$$

$\varepsilon \gg u \Rightarrow v \approx u$ Beobachtung verbessert Prognose nicht wesentlich

$\varepsilon \ll u \Rightarrow v \approx \varepsilon$

$\varepsilon = u \Rightarrow v = \frac{u}{2}$ Varianzreduktion um Faktor $\frac{1}{2}$ durch Beobachtung

Exkurs zu Gaußmatrizen:

3.5. Multivariate Normalverteilung

Sei $m \in \mathbb{R}^n$, V $n \times n$ -Matrix, symmetrisch und strikt positiv definit.

Def. Die Wahrscheinlichkeitsverteilung $N(\mu, V)$ auf $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$ mit Dichte

$$f(x) = \frac{1}{\sqrt{(2\pi)^n \det V}} e^{-\frac{1}{2} (x-\mu), V^{-1}(x-\mu)}_{\mathbb{R}^n} \quad (x \in \mathbb{R}^n)$$

heißt multivariate Normalverteilung (Gaußverteilung) mit Mittel $\mu = (\mu_1, \dots, \mu_n)$ und Kovarianzmatrix $V = (v_{ij})_{i,j=1}^n$.

Bem. $X \sim N(\mu, V) \Rightarrow E[X_i] = \mu_i, \text{cov}(X_i, X_j) = v_{ij}$

s.u. (3.7) oder auch s. W-Theorie I, 21.5.

„Philosophie“: • Viele Zufallsexperimente können durch Gaußverteilungen näherungsweise modelliert werden (Zentraler Grenzwertsatz!)

• Für Gaußverteilungen kann man vieles explizit berechnen.

Gaußsche Analysis \leftrightarrow Lineare/Multilineare Algebra

3.6. Beispiele, Berechnungen

(37)

$$1) \quad V = \begin{pmatrix} \sigma_1^2 & & 0 \\ & \sigma_2^2 & \\ 0 & & \ddots \\ & & & \sigma_n^2 \end{pmatrix} \quad \text{Diagonalmatrix}$$

$$\Rightarrow N(\mu, V) = \bigotimes_{i=1}^n \underbrace{N(\mu_i, \sigma_i^2)}_{\text{eindim. Normalverteilungen}}$$

Also: $X = (X_1, \dots, X_n)$ $N(\mu, V)$ -verteilt

$$\Leftrightarrow X_i \text{ unabh\u00e4ngig, } \sim N(\mu_i, \sigma_i^2)$$

2) Sei $A: \mathbb{R}^n \rightarrow \mathbb{R}^n$ linear mit $\det A \neq 0$.

Lemma $X \sim N(0, V) \Leftrightarrow AX \sim N(0, AVA^t)$

$$\Leftrightarrow AX + \mu \sim N(\mu, AVA^t)$$

Beweis: Transformationsatz

$$P[AX \in B] = P[X \in A^{-1}(B)] = \int_{A^{-1}(B)} f(x) dx = \int_B \underbrace{f(A^{-1}y)}_{= f_{AX}(y)} |\det A^{-1}| dy$$

$$f_{AX}(y) = \text{const.} \cdot e^{-\frac{1}{2} (A^{-1}y, V^{-1}A^{-1}y)} = \text{const.} \cdot e^{-\frac{1}{2} (y, (AVA^t)^{-1}y)} \quad \square$$

3) Allgemeiner Fall: $X \sim N(\mu, V)$, V symmetrisch, pos. definit

Diagonalisierung: $V = A \Lambda A^{tr}$

wobei $\Lambda = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}$ Diagonalmatrix aus Eigenwerten,

A orthogonal (A^{tr} ist Projektio. auf Eigenvektoren)

$\rightarrow X = AZ + \mu$ wobei $Z \sim N(0, \Lambda)$

(d.h. Z_1, \dots, Z_n sind unabh., $\sim N(0, \lambda_i)$)

3.7. Charakteristische Funktion Sei $X \sim N(\mu, V)$.

$$\varphi(p) = E[e^{i(p, X)}]_{\mathbb{R}^n} = e^{i(p, \mu)} \cdot e^{-\frac{1}{2}(p, Vp)} \quad (p \in \mathbb{R}^n)$$

Beweis: Sei A, Λ wie in 3.6.3), $V = A \Lambda A^{tr}$. $= e^{-\frac{1}{2} \lambda_k A_k^2}$

$X = AZ + \mu$ wobei $Z \sim N(0, \Lambda)$, $E[e^{i(p, Z)}] = \prod_{k=1}^n E[e^{i p_k Z_k}] = e^{-\frac{1}{2}(p, \Lambda p)}$

$\Rightarrow E[e^{i(p, X)}] = e^{i(p, \mu)} \cdot E[e^{i(A^{tr} p, Z)}] = e^{i(p, \mu)} \cdot e^{-\frac{1}{2}(A^{tr} p, \Lambda A^{tr} p)}$

□

Nützlich für viele konkrete Berechnungen mit Gaußschen Zufallsvariablen,
z.B.:

Kor. 1 $X \sim N(\mu, V) \Rightarrow E[X_j] = \mu_j, \text{Cov}(X_j, X_k) = v_{jk}$

Beweis: $E[X_j] = \left. \frac{1}{i} \frac{\partial}{\partial p_j} E[e^{i(p \cdot X)}] \right|_{p=0} = \mu_j$

$$E[X_j X_k] = \left. -\frac{\partial^2}{\partial p_j \partial p_k} E[e^{i(p \cdot X)}] \right|_{p=0}$$

$$= \left. -\frac{\partial}{\partial p_j} \left[(i\mu_k - (Vp)_k) \varphi(p) \right] \right|_{p=0}$$

$$= \left[-(i\mu_j - (Vp)_j) (i\mu_k - (Vp)_k) \varphi(p) + v_{jk} \varphi(p) \right] \Big|_{p=0}$$

$$= \mu_j \cdot \mu_k + v_{jk}$$

$$\Rightarrow \text{Cov}(X_j, X_k) = E[X_j X_k] - E[X_j] E[X_k] = v_{jk} \quad \square$$

Kor. 2 Ist $(X_1, \dots, X_n) \sim N(\mu, V)$ verteilt, dann ist jede
Linearkombination der Zufallsvariablen X_1, \dots, X_n normalverteilt.

Beweis: $Y = \sum_{i=1}^n a_i X_i \Rightarrow \varphi_Y(t) = E[e^{itY}] = E[e^{i(ta, X)}] \stackrel{\text{L.o.}}{=} e^{it(a, \mu) - \frac{1}{2} t^T (a, Va)}$
(Eck. der Fouriertf.) $\Rightarrow Y \sim N(a\mu, (a, Va)) \quad \square$

3.8. Lineare Prognose

Filter-/Prognoseproblem: $X_1, X_2: \Omega \rightarrow \mathbb{R}$ Zufallsvariablen

X_1 : unbeobachtetes Signal

X_2 : unsere Beobachtung

Gesucht: Prognose für X_1 gegeben X_2 .

Einfachster Fall: Gemeinsame Verteilung von X_1, X_2 ist zweidimensionale Normalverteilung $N(\mu, V)$,

$$V = \begin{pmatrix} \sigma_1^2 & \rho \sigma_1 \sigma_2 \\ \rho \sigma_1 \sigma_2 & \sigma_2^2 \end{pmatrix} \quad \begin{array}{l} \sigma_i = \text{Standardabweichung von } X_i \\ \rho = \text{Korrelation von } X_1 \text{ und } X_2 \\ \left(= \frac{\text{cov}(X_1, X_2)}{\sigma_1 \sigma_2} \right) \end{array}$$

(z. B. $X_1 \sim N(\mu, \sigma_1^2)$, $X_2 = X_1 + Z$, Z unabh. von X_1 , $Z \sim N(0, \epsilon)$, s.o. 3.4)

$$\det V = \sigma_1^2 \sigma_2^2 (1 - \rho^2), \quad V^{-1} = \frac{1}{1 - \rho^2} \begin{pmatrix} \frac{1}{\sigma_1^2} & -\frac{\rho}{\sigma_1 \sigma_2} \\ -\frac{\rho}{\sigma_1 \sigma_2} & \frac{1}{\sigma_2^2} \end{pmatrix}$$

Dichte:

$$f(x_1, x_2) = \frac{1}{2\pi \sqrt{\det V}} e^{-\frac{1}{2}(x - \mu)^T V^{-1}(x - \mu)}$$

$$= \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \cdot e^{-\frac{\xi_1^2 - 2\rho\xi_1\xi_2 + \xi_2^2}{2 \cdot (1-\rho^2)}}, \quad \xi_i = \frac{x_i - m_i}{\sigma_i}$$

$$= \text{const.} \cdot \underbrace{e^{-\frac{(x_1 - m_1)^2}{2\sigma_1^2}}}_{\text{proportional zu Dichte von } X_1} \cdot \underbrace{e^{-\frac{(x_2 - m(x_1))^2}{2\sigma_2^2(1-\rho^2)}}}_{\text{proportional zu } f_{X_2|X_1}(x_2, x_1)}, \quad m(x_1) = m_2 + \rho \frac{\sigma_2}{\sigma_1} (x_1 - m_1)$$

Also: Die bedingte Verteilung von X_2 gegeben $X_1(\omega)$ ist $N(m(x_1(\omega)), \sigma_2^2(1-\rho^2))$

Beste Prognose: $E[X_2|X_1] = m(X_1) = m_2 + \rho \frac{\sigma_2}{\sigma_1} (X_1 - m_1)$

hängt linear von X_1 ab!

~~GILT NUR FÜR~~
EIGENSCHAFT VON NORMALVERTEILTEN ZV!

Quadratischer Prognosefehler: $\text{var}(X_2|X_1) = \sigma_2^2 \cdot (1-\rho^2)$

Starke Korrelation von X_1 und $X_2 \Rightarrow |\rho| \approx 1 \Rightarrow$ kleine Prognosefehler
geringe " " " " " " $\Rightarrow \rho \approx 0 \Rightarrow$ Fehler $\approx \sigma_2^2 = \text{var}(X_2)$

Höherdimensionaler Fall geht analog:

$\left. \begin{array}{l} n \text{ dimensionale unbekannte Verteilung} \\ m \text{ Beobachtungen} \end{array} \right\}$ Zusammen Normalverteilung in \mathbb{R}^{n+m}

3.9. Reguläre bedingte Verteilungen

Sei (S, \mathcal{F}) meßbarer Raum, $X: \Omega \rightarrow S$ meßbar, $\mathcal{F} \subseteq \mathcal{O}$ σ -Algebra.

Bedingte Verteilung von X gegeben \mathcal{F} :

$$\mu(\omega, B) := P[X \in B | \mathcal{F}](\omega) \quad (\omega \in \Omega, B \in \mathcal{F}).$$

Für festes B P-f.s. eindeutig, ABER es gibt überabzählbar viele B (im allgemeinen).

Gibt es kanonische Wahl, so daß $\mu(\omega, \cdot)$ für jedes ω eine WV ist?

Definition Eine Abbildung $\mu: \Omega \times \mathcal{F} \rightarrow [0, 1]$ heißt

reguläre Version der bedingten Verteilung von X gegeben \mathcal{F} , falls:

(i) Für jedes $B \in \mathcal{F}$ ist $\omega \mapsto \mu(\omega, B)$ eine Versio. von $P[X \in B | \mathcal{F}]$.

(ii) Für jedes $\omega \in \Omega$ ist $\mu(\omega, \cdot)$ eine WV auf (S, \mathcal{F}) .

Bem. 1) Insbesondere ist μ ein stoch. Kern von (Ω, \mathcal{F}) nach (S, \mathcal{F}) .

2) Berechnung von bedingte Erwartungen:

$$E[h(X) | \mathcal{F}](\omega) = \int h(x) \mu(\omega, dx) \quad \text{P-f.s. } \forall h \geq 0 \text{ meßbar}$$

FAKT. 1) Eine reguläre Version der bedingten Verteilung existiert immer, falls $(\mathcal{S}, \mathcal{P})$ schwache Regularitätsannahme erfüllt, z.B. \mathcal{S} polnisch (vollständig metrisch separabel), $\mathcal{P} = \mathcal{B}(\mathcal{S})$, s. Billman: Probability, Ströck: Probability Theory

2) Sei (T, \mathcal{T}) weiterer messbarer Raum, und $Y: \Omega \rightarrow T$ messbar. Unter schwachen Voraussetzungen an \mathcal{S} und T existiert eine Funktion $f_{X|Y}: \mathcal{S} \times T \rightarrow [0,1]$, sodass

(i) $f_{X|Y}(\cdot, y)$ ist eine WV auf $(\mathcal{S}, \mathcal{P})$ für jedes $T \in \mathcal{T}$.

(ii) $\omega \mapsto f_{X|Y}(\cdot, Y(\omega))$ ist eine Version von $P[X \in B | Y]$ für jedes $B \in \mathcal{S}$.

Ist $f_{X|Y}(dx, y)$ für jedes y absolutstetig bzgl. eines Referenzmaßes ν auf $(\mathcal{S}, \mathcal{P})$, dann heißt die Dichte $f_{X|Y}(x, y)$ eine Version der bedingten Dichte bzgl. ν von X gegeben Y .

4. Bedingte Erwartungen als beste L^2 Approximation, Existenz

(44)

Sei (Ω, \mathcal{O}, P) W'raum, $\mathcal{F} \subseteq \mathcal{O}$ σ -Algebra.

4.1. Jensen'sche Ungleichung $X \in \mathcal{L}^1(\Omega, \mathcal{O}, P)$

Satz $u: \mathbb{R} \rightarrow \mathbb{R}$ konvex mit $u(X) \in \mathcal{L}^1$ oder $u \geq 0$

$$\Rightarrow E[u(X) | \mathcal{F}] \geq u(E[X | \mathcal{F}]) \quad P\text{-f.s.}$$

Beweis: Konvexe Funktionen sind abzählbare Suprema von linearen Funktionen (s. z. B. Williams: Prob. with martingales, 6.6), d.h.

$$\exists a_n, b_n \in \mathbb{R} : u(x) = \sup_{n \in \mathbb{N}} \underbrace{(a_n x + b_n)}_{=: \ell_n(x)} \quad \forall x \in \mathbb{R}$$

$$\Rightarrow E[u(X) | \mathcal{F}] \geq E[\ell_n(X) | \mathcal{F}] = \ell_n(E[X | \mathcal{F}]) \quad P\text{-f.s.}$$

$$\Rightarrow E[u(X) | \mathcal{F}] \geq \sup_n \ell_n(E[X | \mathcal{F}]) = u(E[X | \mathcal{F}]) \quad P\text{-f.s.}$$

□

Kor. $X \mapsto E[X|\mathcal{F}]$ ist Kontraktion auf L^p für alle $p \geq 1$:

$$\|E[X|\mathcal{F}]\|_p \leq \|X\|_p \quad \forall X \in L^p$$

Beweis: $|E[X|\mathcal{F}]|^p \stackrel{\text{Jensen}}{\leq} E[|X|^p|\mathcal{F}]$

$$\Rightarrow E[|E[X|\mathcal{F}]|^p] \leq E[|X|^p] \quad \square$$

Insbesondere: $X \in L^p \Rightarrow E[X|\mathcal{F}] \in L^p$

4.2. Bedingte Erwartung als beste Approximation mit quadratischem Fehler

$$L^2(\Omega, \mathcal{O}, P) = \mathcal{L}^2(\Omega, \mathcal{O}, P) / \sim \text{ ist Hilbertraum}$$

$$\text{Skalarprodukt } (X, Y)_{L^2} = E[XY]$$

$$L^2(\Omega, \mathcal{F}, P) \subseteq L^2(\Omega, \mathcal{O}, P) \text{ ist abgeschlossener Unterraum.}$$

(Grenzwerte von \mathcal{F} -meßbaren ZV sind wieder \mathcal{F} -meßbar)

$$X \in L^2(\Omega, \mathcal{O}, P) \stackrel{\text{Jensen}}{\Rightarrow} E[X|\mathcal{F}] \in L^2(\Omega, \mathcal{F}, P) \text{ (für jede Version)}$$

$X \mapsto E[X|\mathcal{F}]$ induziert lineare Abbildung von $L^2(\Omega, \mathcal{O}, P)$ nach $L^2(\Omega, \mathcal{F}, P)$. Welche?

Satz Sei $Y \in L^2(\Omega, \mathcal{F}, P)$. Es sind äquivalent:

(i) Y ist Version von $E[X|\mathcal{F}]$.

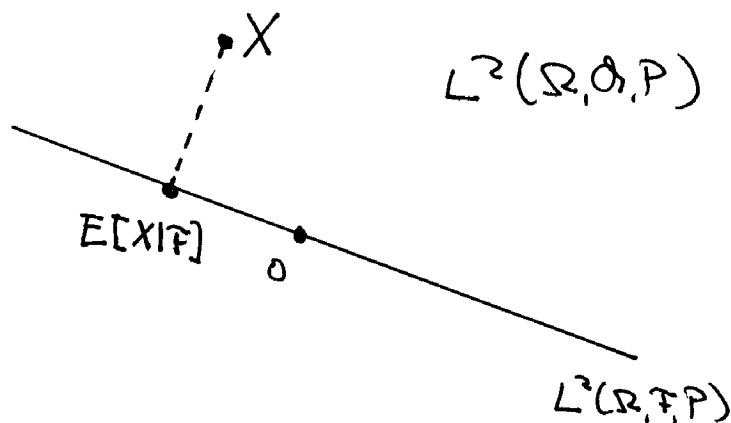
(ii) Y ist „beste Approximation“ von X in $L^2(\Omega, \mathcal{F}, P)$, d.h.

$$E[(X-Y)^2] \leq E[(X-Z)^2] \quad \forall Z \in L^2(\Omega, \mathcal{F}, P)$$

(iii) Y ist (eine Version der) orthogonale Projektion von X auf den Unterraum $L^2(\Omega, \mathcal{F}, P) \subseteq L^2(\Omega, \mathcal{O}, P)$, d.h.

$$X - Y \perp Z \quad \forall Z \in L^2(\Omega, \mathcal{F}, P)$$

$$\underbrace{\hspace{10em}}_{E[(X-Y) \cdot Z] = 0}$$



Rechtfertigt bedingte Erwartung als Prognoseverfahren, z.B.

$E[X|Y]$ = beste Prognose unter allen Funktionen $g(Y)$, g messbar.

Beweis:

(i) \Leftrightarrow (ii): Y Version von $E[X|F]$

$$\stackrel{\text{Def.}}{\Leftrightarrow} E[Y \cdot I_A] = E[X \cdot I_A] \quad \forall A \in \mathcal{F}$$

$$\stackrel{\substack{\text{M\ss}stheoret. \\ \text{h.d.}}}{\Leftrightarrow} E[Y \cdot Z] = E[X \cdot Z] \quad \forall Z \in \mathcal{L}^2(\Omega, \mathcal{F}, P)$$

$$\Leftrightarrow X - Y \perp Z \quad \forall Z \in \mathcal{L}^2(\Omega, \mathcal{F}, P)$$

(iii) \Rightarrow (ii): Sei Y orth. Proj. von X auf $\mathcal{L}^2(\Omega, \mathcal{F}, P)$, $Z \in \mathcal{L}^2(\Omega, \mathcal{F}, P)$

$$\begin{aligned} \Rightarrow E[(X-Z)^2] &= E[(X-Y)^2] + \overbrace{E[(Y-Z)^2]}{\geq 0} + 2 \overbrace{E[(X-Y)(Y-Z)]}_{\substack{= 0 \\ \in \mathcal{L}^2(\Omega, \mathcal{F}, P)}} \\ &\geq E[(X-Y)^2] \end{aligned}$$

(ii) \Rightarrow (iii): Sei Y beste Approx. und $Z \in \mathcal{L}^2(\Omega, \mathcal{F}, P)$

$$\Rightarrow E[(X-Y)^2] \leq E[(X-Y+tZ)^2] = E[(X-Y)^2] + 2tE[(X-Y)Z] + t^2E[Z^2]$$

für alle $t \in \mathbb{R}$

$$\Rightarrow X - Y \perp Z \quad \square$$

4.3. Bemerkung: Beste lineare Prognose $X, Y \in \mathcal{L}^2$

Die beste Prognose im quadratischen Mittel von X unter allen Funktionen $aY + b$ ($a, b \in \mathbb{R}$) ist

$$\hat{X} = aY + b, \quad a = \frac{\text{cov}(Y, X)}{\text{var}(Y)}, \quad b = E[X] - aE[Y],$$

s. W'theorie I, 8.3. In der Regel gilt $\hat{X} \neq E[X|Y]$.

Wichtige Ausnahme: Die gemeinsame Verteilung von X und Y ist eine Gaußverteilung, s. 3.8.

4.4. Existenz von $E[X|F]$

Satz $X \in \mathcal{L}^2(\Omega, \mathcal{O}, P) \Rightarrow \exists$ Version Y von $E[X|F]$ in $\mathcal{L}^2(\Omega, \mathcal{F}, P)$

Beweis: $\mathcal{L}^2(\Omega, \mathcal{F}, P) \subseteq \mathcal{L}^2(\Omega, \mathcal{O}, P)$ Unterraum, abgeschlossen

$\Rightarrow \exists$ orth. Projektion von X auf $\mathcal{L}^2(\Omega, \mathcal{F}, P)$. \square

Projektionsatz,
s. Funktionalanalysis

Existenz für allgemeine $X \geq 0$:

$$\exists X_n \in \mathcal{L}^2: 0 \leq X_n \nearrow X$$

$\sup E[X_n | \mathcal{F}]$ ist Version von $E[X | \mathcal{F}]$
 fast sicher monoton

Existenz für $X \in \mathcal{L}^1$:

$$X = X^+ - X^-, \quad E[X^+ | \mathcal{F}] - E[X^- | \mathcal{F}] \text{ ist Version von } E[X | \mathcal{F}].$$

Rem. Alternativer Existenzbeweis via Satz von Radon-Nikodym
 (Existenz einer Dichte für absolutstetige Wahrscheinlichkeitsverteilungen).

Wir beweisen umgekehrt diesen Satz mithilfe von bedingte Erwartungen, s.u.

II . MARTINGALE

(50)

5. Definitionen und Beispiele Sei (Ω, \mathcal{O}, P) W'raum.

5.1. Filtrationen

Def. Eine Filtration ist eine aufsteigende Folge

$$\mathcal{F}_0 \subseteq \mathcal{F}_1 \subseteq \mathcal{F}_2 \subseteq \dots$$

von σ -Algebren $\mathcal{F}_n \subseteq \mathcal{O}$.

Interpretation: \mathcal{F}_n = Information, die bis zur Zeit n vorliegt.

\mathcal{F}_n -meßbare Zufallsvariablen = ZV_n deren Wert wir zur Zeit n kennen

Wichtigstes Beispiel: $(X_n)_{n \geq 0}$ stoch. Prozeß auf (Ω, \mathcal{O}, P) ,

$$\mathcal{F}_n = \sigma(X_0, X_1, \dots, X_n) \quad \text{kanonische Filtration}$$

Alternativ kann \mathcal{F}_n noch zusätzliche Information enthalten,

z.B. $\mathcal{F}_n = \sigma(Y, X_0, X_1, \dots, X_n)$, Y unabh. von $\{X_n | n \geq 0\}$.

Def. 1) Ein stochastischer Prozeß $(X_n)_{n \geq 0}$ ist adaptiert an

die Filtration $(\mathcal{F}_n)_{n \geq 0} : \Leftrightarrow X_n$ ist \mathcal{F}_n -meßbar für jedes n

2) $(X_n)_{n \geq 0}$ ist previsibel : $\Leftrightarrow X_n$ ist \mathcal{F}_{n-1} -meßbar für jedes n

5.2. Martingaldefinition Sei $(\mathcal{F}_n)_{n \geq 0}$ Filtration.

Def. Ein reellwertiger stochastischer Prozeß $X_n: \Omega \rightarrow \mathbb{R}$ ($n=0,1,2,\dots$)

heißt Martingal bzgl. (\mathcal{F}_n) falls

1) adaptiert : X_n ist \mathcal{F}_n -meßbar und integrierbar $\forall n \geq 0$

2) faires Spiel : $E[X_{n+1} | \mathcal{F}_n] = X_n$ P-f.s.

(bzw. $E[X_{n+1} - X_n | \mathcal{F}_n] = 0$ P-f.s.)

Gilt statt 2) $E[X_{n+1} | \mathcal{F}_n] \stackrel{(\geq)}{\leq} X_n$, dann heißt $(X_n)_{n \geq 0}$

Supermartingal (Submartingal).

Bem. Insbesondere gilt $E[X_{n+1}] = E[X_n] \stackrel{\text{ind.}}{=} E[X_0]$ falls (X_n) Martingal

bzw. $E[X_{n+1}] \leq E[X_n]$ falls (X_n) Supermartingal.

Supermartingale : „im Mittel fallend“

Submartingale : „im Mittel wachsend“

Wichtige Anwendung : Martingalkonvergenzsätze (s.u.)

Analysis : mon. fallende, nach unten beschränkte Folgen konvergieren

Stochastik : nach unten beschränkte Supermartingale konvergieren fast sicher

Satz $(X_n)_{n \geq 0}$ Martingal (Supermartingal) bzgl. $(\mathcal{F}_n)_{n \geq 0}$

$$\Rightarrow E[X_n | \mathcal{F}_m] \stackrel{(\leq)}{=} X_m \text{ P-f.s. } \forall 0 \leq m \leq n$$

Beweis : $E[X_{m+1} | \mathcal{F}_m] = X_m$ P-f.s. nach Def.

$$E[X_{m+2} | \mathcal{F}_m] \stackrel{\text{Projektivität}}{=} E[E[X_{m+2} | \mathcal{F}_{m+1}] | \mathcal{F}_m] = X_m \text{ P-f.s.}$$

$= X_{m+1} \text{ P-f.s.}$

Induktion $\Rightarrow E[X_{m+k} | \mathcal{F}_m] = X_m$ P-f.s. $\forall k \geq 0$. \square

Übung: X_n Martingal, $u: \mathbb{R} \rightarrow \mathbb{R}$ konvex

$\Rightarrow u(X_n)$ ist Submartingal falls integrierbar

5.3. Beispiele

(53)

1) Summen von unabhängigen zentrierten ZVn:

$$S_n = Y_1 + Y_2 + \dots + Y_n, \quad Y_i \in \mathcal{L}^1 \text{ unabh. mit } E[Y_i] = 0$$

$\Rightarrow (S_n)_{n \geq 0}$ ist Martingal bzgl.

$$\mathcal{F}_n := \sigma(Y_1, \dots, Y_n) = \sigma(S_1, \dots, S_n)$$

Beweis: $E[S_{n+1} | \mathcal{F}_n] = E[S_n + Y_{n+1} | \mathcal{F}_n] = S_n + E[Y_{n+1}] = S_n.$ \square

\mathcal{F}_n -messbar \uparrow unabh. von \mathcal{F}_n

z.B. klassischer random walk: $Y_i = \pm 1$ mit $p = \frac{1}{2}$

„Balanzentwicklung bei fairem Glücksspiel“

2) wie oben, aber $E[Y_i] \leq 0 \quad \forall i$

$$\Rightarrow E[S_{n+1} | \mathcal{F}_n] \leq S_n \quad \forall n \geq 0$$

$\Rightarrow (S_n)_{n \geq 0}$ ist Supermartingal bzgl. (\mathcal{F}_n)

z.B. allgemeiner random walk: $S_n = Y_1 + \dots + Y_n$, Y_i iid mit $E[Y_i] = \mu$

• Martingal, falls $\mu = 0$ • Supermartingal, falls $\mu \leq 0$ • Submartingal, falls $\mu \geq 0$

Bem. Doob-Zerlegung: $S_n = \underbrace{\sum_{i=1}^n (Y_i - E[Y_i | \mathcal{F}_{i-1}])}_{=: M_n} + \underbrace{\sum_{i=1}^n E[Y_i | \mathcal{F}_{i-1}]}_{=: m \cdot n}$

ist Martingal

mon. fallend
bzw. wachsend
previsibel

3) Produkte von positiven unabhängigen ZVn

Y_1, Y_2, \dots unabh., ≥ 0 , $\in \mathcal{L}^1$, $m E[Y_i] = m \cdot \forall i$

$M_n = Y_1 \cdot Y_2 \cdot \dots \cdot Y_n$ (z.B. Populationswachstum)

$\mathcal{F}_n = \sigma(Y_1, \dots, Y_n)$

$E[M_{n+1} | \mathcal{F}_n] = E[M_n \cdot Y_{n+1} | \mathcal{F}_n] = M_n \cdot E[Y_{n+1}] = m \cdot M_n$
weil unabh.

$\Rightarrow (M_n)_{n \geq 0}$ ist Martingal falls $m=1$ (kritischer Fall)

Supermartingal falls $m < 1$ (subkritischer Fall)

Bem. Martingalkonvergenzsatz \Rightarrow Für $m \leq 1$ ex. $M_\infty = \lim M_n$ P-f.s.

$m=1$: $E[M_n] = E[M_0] = 1 \quad \forall n \geq 0$. Gilt auch $E[M_\infty] = 1$?

I.a. falsch, z.B. Y_i i.i.d. mit $P[Y_i=0] \neq 0 \rightarrow P[M_n=0 \text{ schließlich}] = 1 \Rightarrow M_\infty = 0$
P-f.s.

4) Subzessive Prognosen $X \in \mathcal{L}^1$, $(\mathcal{F}_n)_{n \geq 0}$ Filtration

$M_n := E[X | \mathcal{F}_n]$ ist Martingal:

$$E[M_{n+1} | \mathcal{F}_n] = E[E[X | \mathcal{F}_{n+1}] | \mathcal{F}_n] \stackrel{\text{Proj.}}{=} E[X | \mathcal{F}_n] = M_n$$

Bem. $(X_n)_{n \geq 0}$ beliebiges Martingal bzgl. (\mathcal{F}_n)

$$\Rightarrow X_n = E[X_\infty | \mathcal{F}_n] \quad \forall 0 \leq n < \infty$$

Frage: Existiert X_∞ mit

$$X_n = E[X_\infty | \mathcal{F}_n] \quad \forall n \geq 0 ?$$

n.g.V. ja ∇ (s.u.)

5.4. Harmonische Funktionen von Markovschen Ketten

$(X_n)_{n \geq 0}$ zeitlich homogene Markovkette auf (S, \mathcal{F})

mit Übergangskern $p(x, dy)$,

$$\mathcal{F}_n := \sigma(X_0, X_1, \dots, X_n)$$

$$P[X_{n+1} \in B | \mathcal{F}_n] = P[X_{n+1} \in B | X_n] = p(X_n, B) \quad P\text{-f.s.}$$

$$\Rightarrow E[f(X_{n+1}) | \mathcal{F}_n] = \int p(X_n, dy) f(y) = (pf)(X_n) \quad P\text{-f.s.}$$

für alle $f \geq 0$ meßbar, wobei:

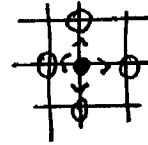
$$(pf)(x) := \int p(x, dy) f(y)$$

Satz: h harmonisch (d.h. $ph = h$) $\Rightarrow h(X_n)$ ist Martingal
falls integrierbar

h superharmonisch (d.h. $ph \leq h$) $\Rightarrow h(X_n)$ ist Supermartingal
falls integrierbar

Beweis: $E[h(X_{n+1}) | \mathcal{F}_n] = (ph)(X_n) \stackrel{(\leq)}{=} h(X_n) \quad \forall n \geq 0. \quad \square$

Beispiel $(X_n)_{n \geq 0}$ random walk auf \mathbb{Z}^d



$$p(x, y) = \begin{cases} \frac{1}{2d} & \text{falls } |x-y|=1 \\ 0 & \text{sonst} \end{cases}$$

$$p h \stackrel{(\leq)}{=} h \iff \frac{1}{2d} \sum_{i=1}^d (h(x+e_i) + h(x-e_i)) \stackrel{(\leq)}{=} h(x) \quad \forall x$$

Mittelwertesenschaft

$$\iff \frac{1}{2d} \sum_{i=1}^d \underbrace{\left[(h(x+e_i) - h(x)) - (h(x) - h(x-e_i)) \right]}_{\text{diskrete 2. Ableitung in Richtung } e_i} \stackrel{(\leq)}{=} 0$$

$$= \frac{1}{2d} \Delta h(x) \text{ diskreter Laplaceoperator}$$

Bem. Sei $(X_n)_{n \geq 0}$ Markovkette mit Start in x , h harmonisch

$$\implies h(x) = E[h(X_0)] \stackrel{\text{Markov}}{=} E[h(X_n)] \quad \forall n \geq 0$$

Sei $T(\omega) = \inf \{n \geq 0 \mid X_n(\omega) \in D^c\}$ wobei $D \subseteq S$.

Frage: $h(x) = E[h(X_T)]$? \rightarrow Darstellung harmonischer Funktionen über Randwerte

5.5. Martingale des random walk $x \in \mathbb{Z}$

$$X_n = x + S_n, \quad S_n = Y_1 + \dots + Y_n \text{ klassisches random walk,}$$

d.h. Y_i i.i.d., $Y_i = \pm 1$ mit $p = \frac{1}{2}$.

$$\mathcal{F}_n := \sigma(Y_1, \dots, Y_n)$$

Satz Folgende Prozesse sind Martingale bzgl. $(\mathcal{F}_n)_{n \geq 0}$:

(i) X_n

(ii) $X_n^2 - n$

(iii) $e^{\lambda X_n - a(\lambda) \cdot n}$ wobei $a(\lambda) := \log \cosh \lambda$

für jedes $\lambda \in \mathbb{R}$.

Beweis: Übung.

Verallgemeinerung von (iii):

5.6. Exponentielle Martingale $S_n = Y_1 + \dots + Y_n,$

Y_1, Y_2, \dots i.i.d., $M(\lambda) = E[e^{\lambda X_i}]$ ($\lambda \in \mathbb{R}$) momentenerzeugende Funktion

Satz Gilt $M(\lambda) < \infty$, dann ist

$$M_n^\lambda := \frac{e^{\lambda S_n}}{M(\lambda)^n} = e^{\lambda S_n - n \log M(\lambda)}$$

ein Martingal.

Beweis:

$$M_n^\lambda = \prod_{i=1}^n \frac{e^{\lambda Y_i}}{M(\lambda)}, \text{ also Martingal nach 5.33.}$$

unabh., Erw. wert 1

□

Konsequenz: Große Abweichungen $\mu = E[Y_i]$

$$P\left[\frac{S_n}{n} > \mu + \varepsilon\right] = P[S_n > n \cdot (\mu + \varepsilon)] \leq \frac{E[M_n^\lambda]}{e^{\lambda n \cdot (\mu + \varepsilon) - n \log M(\lambda)}}$$
$$= e^{-n (\lambda \cdot (\mu + \varepsilon) - \log M(\lambda))} \quad \forall \lambda > 0$$

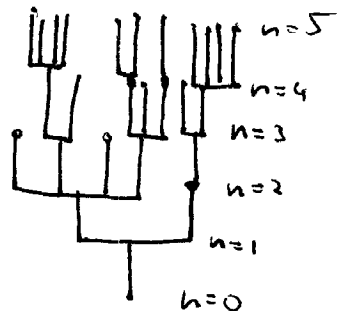
$$\Rightarrow P\left[\frac{S_n}{n} > \mu + \varepsilon\right] \leq e^{-n \cdot \sup_{\lambda > 0} (\lambda \cdot (\mu + \varepsilon) - \log M(\lambda))}$$

Legendretransf. von $\log M(\lambda)$

5.7. Martingale bei Verzweigungsprozessen

$$N_i^n: \Omega \rightarrow \{0, 1, 2, \dots\} \quad (i \geq 1, n \geq 0) \text{ i.i.d., } e \in \mathbb{Z}^1$$

Nachkommenzahl des i -ten Teilchens
in der $(n-1)$ -ten Generation.



Verteilung ν , Erwartungswert $m < \infty$, Annahme: $m > 0$ (sonst trivial),
Varianz σ^2 , erzeugende Funktion:

$$G(s) = E[s^{N_i^n}] = \sum_{k=0}^{\infty} s^k \cdot \nu(k) \quad (0 \leq s \leq 1)$$

Populationsgröße in n -ter Generation: $Z_0 = 1$,

$$Z_n = \sum_{i=0}^{Z_{n-1}} N_i^n$$

$(Z_n)_{n \geq 0}$ ist Markovkette auf $S = \{0, 1, 2, \dots\}$ mit Start in 1.

$$\mathcal{F}_n := \sigma(N_i^k \mid 0 \leq k \leq n, i \geq 1)$$

Lemma $E[Z_{n+1} | \mathcal{F}_n] = m \cdot Z_n, \text{ var}(Z_{n+1} | \mathcal{F}_n) = \sigma^2 \cdot Z_n.$

Beweis: $E\left[\sum_{i=0}^{Z_n} N_i^n | \mathcal{F}_n\right](\omega) = E\left[\sum_{i=0}^{Z_n(\omega)} N_i^n\right] = Z_n(\omega) \cdot m \text{ P-f.s.}$
 \uparrow
 unabh. etc. \square

Kor. 1 1) $E[Z_n] = m^n$

2) $\text{var}(Z_n) = \begin{cases} \sigma^2 m^{n-1} \frac{m^n - 1}{m - 1} & \text{falls } m \neq 1 \\ \sigma^2 m & \text{falls } m = 1 \end{cases}$

Beweis: 1) $E[Z_{n+1}] = E[E[Z_{n+1} | \mathcal{F}_n]] \stackrel{\text{Lem.}}{=} m \cdot E[Z_n]$

Induktion \Rightarrow Beh.

2) $\text{var}(Z_{n+1}) = E[\text{var}(Z_{n+1} | \mathcal{F}_n)] + \text{var}(E[Z_{n+1} | \mathcal{F}_n])$
 $\stackrel{\text{Lem.}}{=} \sigma^2 \cdot E[Z_n] + m^2 \cdot \text{var}(Z_n)$

Induktion \Rightarrow Beh. \square

Kor. 2 $M_n := \frac{Z_n}{m^n} (n \geq 0)$ ist ein Martingal bzgl. (\mathcal{F}_n) .

Beweis: $E[M_{n+1} | \mathcal{F}_n] = \frac{E[Z_{n+1} | \mathcal{F}_n]}{m^{n+1}} = \frac{m Z_n}{m^{n+1}} = M_n \quad \square$

Bem. $\text{var}(M_n) = \frac{\text{var}(Z_n)}{m^{2n}} \stackrel{m+1}{=} \frac{\sigma^2}{m} \cdot \frac{1-m^{-n}}{m-1}$

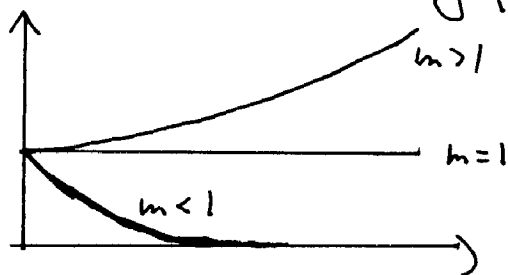
→ Für $m > 1$ ist M_n beschränkt in L^2 falls $\sigma^2 < \infty$!

insbes. gleichmäßig integrierbar für $m > 1, \sigma^2 < \infty$.

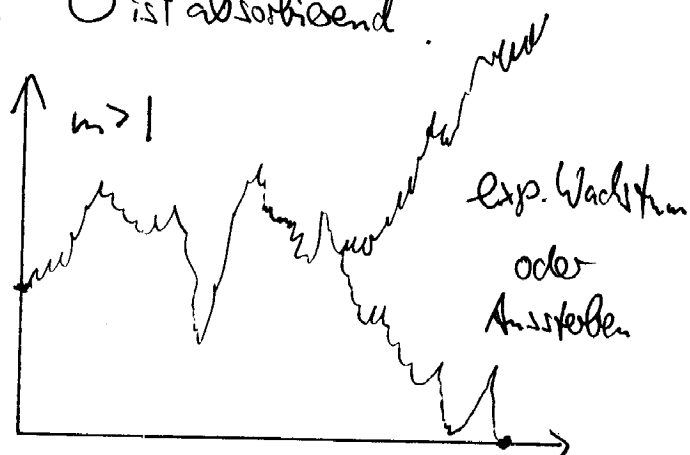
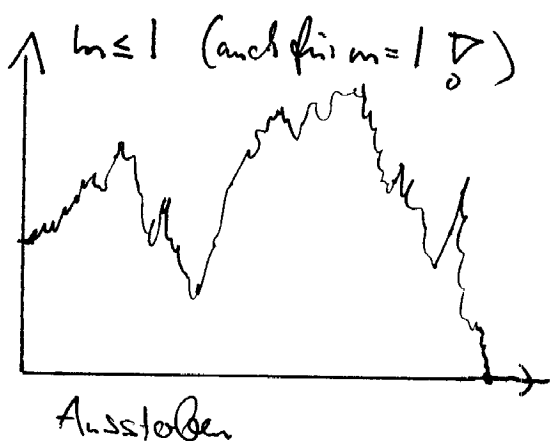
Übung: $X_n := s Z_n$ ist Martingal bzgl. (\mathcal{F}_n) $\Leftrightarrow s = G(s)$

5.8. Konsequenzen für die Asymptotik von Verzweigungsprozessen

deterministische Betrachtung (Erwartungswerte)



Was machen die Trajektorien? 0 ist absorbierend.



a) $m < 1$: subkritischer Fall

$$P[Z_n \neq 0] = P[Z_n \geq 1] \leq E[Z_n] = m^n \rightarrow 0$$

$$\Rightarrow P[Z_n = 0 \text{ schließlich}] = P\left[\bigcup_n \{Z_n = 0\}\right] = \lim P[Z_n = 0] = 1$$

$\underbrace{\qquad\qquad\qquad}_{\text{monoton}}$

Prozess stirbt P-f.s. aus.

Bem. $Z_n = 0$ schl. $\Rightarrow M_n = 0$ schl. $\Rightarrow M_\infty = 0$ P-f.s.,
 obwohl $E[M_n] = 1 \forall n$
 (Lebesgue gibt nicht!)

b) $m = 1$: kritischer Fall $M_n = Z_n \Rightarrow Z_n$ ist Martingal, ≥ 0

Martingalkonvergenzsatz: $\exists Z_\infty = \lim Z_n$ P-f.s.

$$E[Z_\infty] \stackrel{\text{Fatou}}{\leq} \liminf E[Z_n] = 1 \Rightarrow Z_\infty < \infty \text{ P-f.s.}$$

$$\stackrel{\text{Übung}}{\Rightarrow} Z_\infty = 0 \text{ P-f.s. falls } \nu(1) \neq 1.$$

Prozess stirbt P-f.s. aus außer im trivialen Fall (konstante Population).

c) $m > 1$: Superkritischer Fall

M_n Martingal, ≥ 0

Martingalkonvergenzsatz (s.c.)
 $\implies \exists M_\infty = \lim M_n = \lim \frac{Z_n}{m^n} \in [0, \infty) \text{ P-f.s.}$

\implies (*) $Z_n \sim M_\infty \cdot m^n$ exponentielles Wachstum auf $\{M_\infty > 0\}$

$P[M_\infty > 0] = ?$ Was passiert auf $\{M_\infty = 0\}$?

Bem. $\sigma^2 < \infty \xRightarrow{\text{s. 57}} \sup_{n \in \mathbb{N}} \text{var}(M_n) < \infty \implies M_n \text{ beschr. in } \mathcal{L}^2$

$\xRightarrow{\text{s. W44 Koroll. I}} M_n \text{ gleichm. int. Bar}$

$\implies E[M_\infty] = \lim_{n \rightarrow \infty} \underbrace{E[M_n]}_{=1} = 1$

$\implies P[M_\infty > 0] \neq 0$

Satz Sei $\sigma^2 < \infty$. Dann gilt

$$P[M_\infty = 0] = P[Z_n = 0 \text{ schließlich}] = \pi$$

wobei π die kleinste nichtnegative Lösung von $G(s) = s$ ist.

Bem. I.a. ist $P[M_\infty=0]=1$ möglich, obwohl $P[Z_1=0 \text{ schl.}] < 1$.

(Kesten / St. gem: $P[M_\infty=0]=1 \Leftrightarrow \sum_k k \log k \nu(k) < \infty$, s. Atkrega / Ney Ziff.)

Aus dem Satz und aus (*) folgt sofort:

Korollar Ist $\sigma^2 < \infty$, dann gilt für P-f.a. ω :

entweder: (i) $M_\infty(\omega) > 0$, d.h. $Z_n(\omega)$ wächst exponentiell (W'keit $1-\pi$)

oder: (ii) $Z_n(\omega) = 0$ schließlich, d.h. Prozess stirbt aus (W'keit π)

Beweis des Kor.: $\{Z_n=0 \text{ schließlich}\} = \{M_n=0 \text{ schließlich}\} \subseteq \{M_\infty=0\}$

W'keiten gleich nach Satz $\Rightarrow P[M_\infty=0 \text{ und kein Aussterben}] = 0$. □

Beweis des Satzes via Markoveigenschaft (first step analysis):

① Z_n ist eine zeitlich homogene Markovkette mit

$$P[Z_{n+1}=y | Z_0, \dots, Z_n](\omega) = P\left[\sum_{i=1}^{Z_n(\omega)} N_i^{\text{tot}} = y\right] = p(Z_n(\omega), y)$$

i.i.d. $\sim \nu$

wobei $p(x, y) = \underbrace{\nu * \dots * \nu}_x(y)$
x-fache Faltung

② Start mit k Teilchen :

Seien $(Z_n^{(1)})_{n \geq 0}, (Z_n^{(2)})_{n \geq 0}, \dots, (Z_n^{(k)})_{n \geq 0}$ unabhängige Kopien von (Z_n) ,

$$\tilde{Z}_n := Z_n^{(1)} + \dots + Z_n^{(k)}$$

$\Rightarrow (\tilde{Z}_n)_{n \geq 0}$ ist Markovkette mit demselben Kern und $\tilde{Z}_0 = k$,
d.h. (\tilde{Z}_n) ist Realisierung derselben (in Verteilung eindeutigen)

Markovkette mit Startverteilung δ_k

③ Markoveigenschaft + ②

\Rightarrow Bedingte Verteilung von $(Z_{n+1})_{n \geq 0}$ gegeben $Z_i = k$

$$= \text{Verteilung von } (\tilde{Z}_n)_{n \geq 0}$$

④ Konsequenzen: Sei $\pi := P[Z_n = 0 \text{ schl.}]$, $\tilde{\pi} := P[M_n = 0]$

Wir zeigen $\pi = G(\pi)$ und $\tilde{\pi} = G(\tilde{\pi})$:

$$\text{a) } P[\underbrace{Z_n = 0 \text{ schließlich}}_{\Leftrightarrow Z_{n+1} = 0 \text{ schließlich}} | Z_i = k] \stackrel{\text{③}}{=} P[\underbrace{\tilde{Z}_n = 0 \text{ schließlich}}_{\substack{= Z_n^{(1)} + \dots + Z_n^{(k)}}}]$$

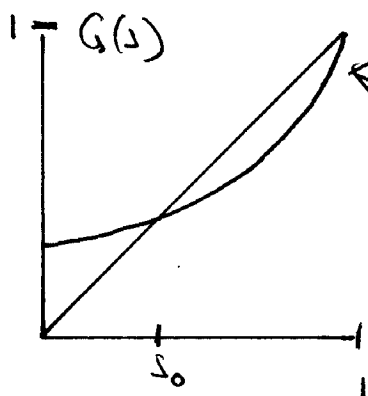
$$= P[\bigcap_{i=1}^k \{Z_n^{(i)} = 0 \text{ schl.}\}] \stackrel{\text{unabl.}}{=} P[Z_n = 0 \text{ schl.}]^k = \pi^k$$

$$\Rightarrow \underline{\underline{\pi}} = P[Z_n=0 \text{ schl.}] = \sum_{k=0}^{\infty} \underline{\underline{\pi}}^k \underbrace{P[Z_1=k]}_{=v(k)} = \underline{\underline{G(\pi)}}$$

b) $\tilde{\pi} = G(\tilde{\pi})$ folgt analog (Übung!)

⑤ $G(s)$, $0 \leq s \leq 1$, ist mon. wachsend und konvex, $G(1)=1$,

$$\left(\frac{d}{ds}\right)^- G(1) = E[N_1^-] = m > 1 \quad (\text{s. W' Theorie I, 26.3 a. H.})$$



genau ein Fixpunkt
 $s_0 \in [0, 1)$

$$\sigma^2 < \infty \stackrel{\text{s. Bem. oben}}{\Rightarrow} \tilde{\pi} = P[M_{\infty}=0] < 1$$

$$\Rightarrow \pi = P[Z_n=0 \text{ schl.}] < 1$$

$$\text{Also } \underline{\underline{\pi}} = \tilde{\pi} = s_0 \quad \square$$

Bem. Ohne Voraussetzung $\sigma^2 < \infty$ kann auch $\tilde{\pi} = 1$ und $\underline{\underline{\pi}} = s_0 < 1$ gelten.

6. Der Stoppsatz

(68)

(Ω, \mathcal{F}, P) W'raum, $(\mathcal{F}_n)_{n \geq 0}$ Filtration.

6.1. Spielstrategien / diskrete stochastische Integrale

Sei $(X_n)_{n \geq 0}$ Martingal bzgl. $(\mathcal{F}_n)_{n \geq 0}$.

Def. Ein stoch. Prozeß $C_n: \Omega \rightarrow \mathbb{R}$ ($n=1, 2, 3, \dots$) heißt previsibel falls C_n \mathcal{F}_{n-1} -meßbar ist $\forall n$.

Interpretation: C_n = Einsatz in der n -ten Runde im Glücksspiel
- hängt ab von Information aus vorherigen
Runden, also von \mathcal{F}_{n-1} .

Resultierende Bilanzentwicklung:

$$M_n = M_{n-1} + C_n \cdot \underbrace{(X_n - X_{n-1})}_{\text{Gewinn/Volust bei Einsatz 1}}$$

d.h.

$$M_n = M_0 + \sum_{k=1}^n C_k \cdot (X_k - X_{k-1})$$

Def. Der stochastische Prozeß $C \cdot X$,

$$(C \cdot X)_n := \sum_{k=1}^n C_k \cdot (X_k - X_{k-1})$$

heißt Martingaltransformation von X bzgl. C

oder diskretes stochastisches Integral von C bzgl. X .

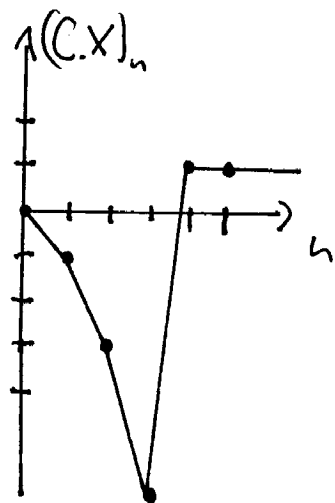
$${}^{\prime\prime} (C \cdot X)_n = \int_0^n C dX {}^{\prime\prime}$$

Beispiel „Martingal-System“, Setzen mit Verdoppeln bei Verlust

$X_n = Y_1 + \dots + Y_n$ Standard random walk

$$C_n = \begin{cases} 2^{n-1} & \text{falls } Y_1 = \dots = Y_{n-1} = -1 \\ 0 & \text{sonst} \end{cases}$$

$$P[(C \cdot X)_n = +1 \text{ schließlich}] = 1$$



6.2. You can't beat the system!

Satz $(X_n)_{n \geq 0}$ Martingal bzgl. (\mathcal{F}_n) , $(C_n)_{n \geq 1}$ previsibel mit $C_n \cdot (X_n - X_{n-1}) \in \mathcal{V}_n$

$\implies C \cdot X$ ist Martingal

- Bem. 1) Insbesondere $E[(C \cdot X)_n] = E[(C \cdot X)_0] = 0 \quad \forall n \geq 0$,
 obwohl beim „Martingalsystem“ $(C \cdot X)_n \rightarrow 1$ P-f.s.
 (Gegenbeispiel zu Lebesgue)
- 2) Integrabilitätsbedingung ist z.B. erfüllt falls C_n beschr. $\forall n$,
 oder falls $C_n, X_n \in \mathcal{L}^2 \forall n$.
- 3) Analoge Aussage gilt für Supermartingale, falls zusätzlich
 $C_n \geq 0 \quad \forall n$:
 X_n Supermart., $C_n \geq 0$ previsibel $\Rightarrow C \cdot X$ Supermart. falls int. bar

Beweis Satz: $E[(C \cdot X)_n - (C \cdot X)_{n-1} \mid \mathcal{F}_{n-1}]$

$$= E[\underbrace{C_n}_{\mathcal{F}_{n-1}\text{-messbar}} \cdot (X_n - X_{n-1}) \mid \mathcal{F}_{n-1}] = C_n \cdot \underbrace{E[X_n - X_{n-1} \mid \mathcal{F}_{n-1}]}_{= 0} = 0 \text{ P-f.s.} \quad \square$$

6.3. Stoppzeiten

Def. $T: \Omega \rightarrow \{0, 1, 2, \dots\} \cup \{\infty\}$ heißt Stoppzeit bzgl. (\mathcal{F}_n) falls $\{T \leq n\} \in \mathcal{F}_n$ für alle $n \geq 0$.

Bem. T ist Stoppzeit $\Leftrightarrow \{T \leq n\} \in \mathcal{F}_n$ für alle $n \geq 0$.

Interpretation: $T(\omega)$ ist z.B. $\left\{ \begin{array}{l} \text{Zeitpunkt, zu dem wir Spiel beenden,} \\ \text{Verkaufszeitpunkt für Aktie.} \end{array} \right.$

Dieser hängt ab von der Information, die wir bis zur momentanen Zeit haben, d.h. $\{T \leq n\}$ muß \mathcal{F}_n -meßbar sein!

Beispiel $(Z_n)_{n \geq 0}$ stoch. Prozeß, adaptiert bzgl. (\mathcal{F}_n) ,

$T_A(\omega) = \inf \{n \geq 0 \mid Z_n(\omega) \in A\}$ (bzw. $= \infty$ falls Menge leer)
erste Trefferzeit von A .

T_A ist Stoppzeit: $\{T_A \leq n\} = \bigcup_{k=0}^n \{Z_k \in A\} \in \mathcal{F}_n$

z. B. Börse: Verkauf, wenn Aktie den Wert λ überschreitet:

$$T = T_{(\lambda, \infty)} = \inf_{n \geq 0} \{Z_n > \lambda\}$$

Bem. Die letzte Besuchszeit

$$L_A(\omega) = \sup \{n \geq 0 \mid X_n(\omega) \in A\}, \quad \sup \emptyset := 0,$$

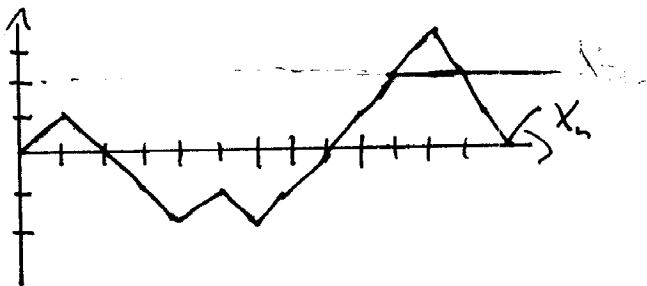
ist in der Regel keine Stoppzeit (Übung).

6.4. Der Stoppsatz

Sei $(X_n)_{n \geq 0}$ Martingal bzgl. (\mathcal{F}_n) , T Stoppzeit bzgl. (\mathcal{F}_n) .

Gestoppter Prozeß:

$$X_{T \wedge n}(\omega) = X_{T(\omega) \wedge n}(\omega) = \begin{cases} X_n(\omega) & \text{für } n \leq T(\omega) \\ X_{T(\omega)}(\omega) & \text{für } n > T(\omega) \end{cases}$$



Satz $(X_n)_{n \geq 0}$ Martingal (bzw. Supermartingal), T Stoppzeit

$$\implies (X_{T \wedge n})_{n \geq 0} \text{ Martingal (Supermartingal)}$$

Insbesondere: $E[X_{T \wedge n}] \stackrel{(\leq)}{=} E[X_0] \quad \forall n \geq 0$

Beweis: Betrachte folgende Spielstrategie:

$$C_n = \begin{cases} 1 & \text{falls } T \geq n \\ 0 & \text{falls } T \leq n-1 \end{cases} = 1 - \mathbb{I}_{\{T \leq n-1\}} \quad \tilde{\mathcal{F}}_{n-1}\text{-messbar}$$

also previsibel.

$$X_{T \wedge n} - X_0 = (C \cdot X)_n,$$

$$\left(\begin{aligned} \text{denn: } X_{T \wedge n} - X_{T \wedge (n-1)} &= \begin{cases} X_n - X_{n-1} & \text{falls } T \geq n \\ 0 & \text{falls } T \leq n-1 \end{cases} = C_n \cdot (X_n - X_{n-1}) \\ \rightarrow X_{T \wedge n} - X_0 &= \sum_{k=1}^n C_k (X_k - X_{k-1}) = (C \cdot X)_n \end{aligned} \right)$$

6.2. \Rightarrow $C \cdot X$ Martingal (Supermartingal) $\Rightarrow (X_{T \wedge n})$ (Super)Martingal.
 C previsibel
 beschränkt ≥ 0 □

6.5. Wann gilt $E[X_T] = E[X_0]$?

- I.a. falsch, z.B. klassischer random walk, $T = \inf\{n \geq 0 \mid X_n = 1\}$:

$X_T = 1$ P-f.s., ~~obwohl~~ ^{also} $E[X_T] = 1$ obwohl $E[X_0] = 0$.

- Ok, falls T beschränkt: $T \leq n$ P-f.s. $\Rightarrow E[X_T] = E[X_{T \wedge n}] = E[X_0]$.

Allgemeines Kriterium:

Satz $(X_n)_{n \geq 0}$ Martingal, T Stoppzeit, $T < \infty$ P-f.s.,

$\{X_{T \wedge n} | n \in \mathbb{N}\}$ gleichmäßig integrierbar

$$\Rightarrow E[X_T] = E[X_0]$$

Beweis: $E[X_0] \stackrel{\text{Stoppzeit}}{=} E[X_{T \wedge n}] \xrightarrow[n \uparrow \infty]{\text{gleichmäßig int. Bsp.}} E[X_T]$. \square
 $\Rightarrow X_T$ P-f.s. (auf $\{T < \infty\}$)

Analog für Supermartingale mit " \leq " statt " $=$ ".

Kor. X_n Martingal mit beschränkte Inkremente, d.h. $\exists k < \infty$ mit

$$|X_{n+1} - X_n| \leq k \quad \forall n \geq 0,$$

T Stoppzeit mit $E[T] < \infty$

$$\Rightarrow E[X_T] = E[X_0]$$

Beweis: $|X_n| \leq |X_0| + k \cdot n \quad \forall n \Rightarrow |X_{T \wedge n}| \leq |X_0| + k \cdot T$. \square
integrierbare Majorante

Übung Sei T Stoppzeit. Existiert $\varepsilon > 0$ und $k \in \mathbb{N}$ mit

$$P[T \leq n+k | \mathcal{F}_n] > \varepsilon \quad \mathbb{P}\text{-f.s.} \quad \forall n \in \mathbb{N},$$

dann gilt $E[T] < \infty$.

Satz X_n Supermartingal, ≥ 0

$$\Rightarrow E[X_0] \geq E[X_T; T < \infty]$$

Beweis: $E[X_0] \stackrel{\text{Stopp-Satz}}{\geq} \lim_{n \rightarrow \infty} E[X_{T \wedge n}] \stackrel{\text{Fatou}}{\geq} E[\liminf_{n \rightarrow \infty} X_{T \wedge n}] \geq E[X_T; T < \infty].$

$\underbrace{\liminf_{n \rightarrow \infty} X_{T \wedge n}}_{= X_T \text{ auf } \{T < \infty\}}$

$\geq 0 \text{ sonst}$ □

6.6. Anwendung auf das klassische Ruinproblem $a, b, x \in \mathbb{Z}, a < x < b$.

$$X_n = x + S_n, \quad S_n = Y_1 + \dots + Y_n, \quad Y_i = \pm 1 \text{ mit } p \in (0, 1).$$

$$T(\omega) = \min \{n \geq 0 \mid X_n(\omega) \notin (a, b)\}$$



$$P = \frac{1}{2} :$$

1) Ruinwahrscheinlichkeit $r(x) = P[X_T = a] = ?$

X_n ist Martingal bzgl. $\mathcal{F}_n = \sigma(Y_1, \dots, Y_n)$, $T < \infty$ P-f.s. (Übung)

$$\Rightarrow x = E[X_0] = E[X_{T \wedge n}] \xrightarrow[n \rightarrow \infty]{\text{Lebesgue}} E[X_T]$$

$\underbrace{\hspace{10em}}_{\rightarrow X_T \text{ P-f.s.}} \quad \underbrace{\hspace{10em}}_{\in \{a, b\}}$
 beschränkt ($\in [a, b]$)

$$= a \cdot r(x) + b \cdot (1 - r(x))$$

$$\Rightarrow \underline{\underline{r(x) = \frac{b-x}{b-a}}}$$

2) Mittlere Austrittszeit aus (a,b) $E[T] = ?$

$M_n = X_n^2 - n$ ist ein Martingal bzgl. \mathcal{F}_n , s. 4.5

$$\Rightarrow x^2 = E[M_0] = E[M_{T_{1/n}}] = E[X_{T_{1/n}}^2] - E[T_{1/n}]$$

$\underbrace{\hspace{10em}}_{\text{beschränkt}} \quad \underbrace{\hspace{10em}}_{\text{monoton}}$

$$\xrightarrow{n \uparrow \infty} E[X_T^2] - E[T]$$

dominierte
und mon. konv.

$$\Rightarrow \underline{E[T]} = E[X_T^2] - x^2 = a^2 \cdot r(x) + b^2 \cdot (1-r(x)) - x^2$$

$$= \dots = \underline{(b-x) \cdot (x-a)}$$

Mittlere Trefferzeit von b

$T_b := \min \{ n \geq 0 \mid X_n = b \}$ \geq Austrittszeit aus (a,b) $\forall a < x$

$$\Rightarrow E[T_b] \geq \underbrace{(b-x) \cdot (x-a)}_{\rightarrow \infty \text{ für } a \rightarrow -\infty} \Rightarrow \underline{\underline{E[T_b] = \infty}}$$

Beispiel für nicht-integrierbare Stoppzeit ∇

3) Verteilung von T_b ? Berechnung im Fall $x=0, b=1$:

Sei $T := T_1$.

$M_n^\lambda := \frac{e^{\lambda X_n}}{(\cosh \lambda)^n}$ ist Martingal für jedes $\lambda \in \mathbb{R}$.

Sei $\lambda > 0$.

$$\begin{aligned}
1 = E[M_0^\lambda] &= E[M_{T_m}^\lambda] = E\left[\frac{e^{\lambda X_{T_m}}}{(\cosh \lambda)^{T_m}}\right] \xrightarrow[n \uparrow \infty]{\text{Lebesgue}} E\left[\frac{e^\lambda}{(\cosh \lambda)^T}; T < \infty\right] \\
&\rightarrow \frac{e^\lambda}{(\cosh \lambda)^T} \cdot \mathbb{I}_{\{T < \infty\}} \\
&\leq e^\lambda \text{ (glm. beschränkt) }
\end{aligned}$$

$$\Rightarrow E\left[\left(\frac{1}{\cosh \lambda}\right)^T; T < \infty\right] = e^{-\lambda}$$

Für $\lambda \downarrow 0$ folgt $P[T < \infty] = 1$, also $(s = \frac{1}{\cosh(\lambda)})$:

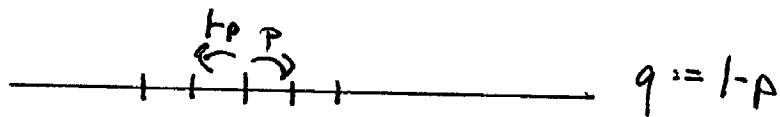
$$\begin{aligned}
E[s^T] &= e^{-\lambda} = \frac{1}{s} (1 - \sqrt{1-s^2}) \quad \forall s \in (0,1) \\
&= \sum_{k=0}^{\infty} s^k \cdot P[T=k] = \sum_{k=0}^{\infty} (-1)^{k+1} \binom{\frac{1}{2}}{k} s^{2k-1}
\end{aligned}$$

Es ist jede Funktion der Verteilung von T

$$\Rightarrow \underline{\underline{P[T=2m-1] = (-1)^{m+1} \binom{\frac{1}{2}}{m} = (-1)^{m+1} \frac{\frac{1}{2} \cdot (-\frac{1}{2}) \cdot \dots \cdot (\frac{1}{2}-m+1)}{m!}}}$$

Eindeutigkeit der Potenzreihendarst.

$$\boxed{p \neq \frac{1}{2}}$$



79

T = Austrittszeit aus (a, b)

$$r(x) = P_x[X_T = a] = ?$$

$$h(x) = \frac{q^x}{p^x} \text{ erfüllt } h(x) = p \cdot h(x+1) + q \cdot h(x-1) \text{ (harmonisch)}$$

$\Rightarrow h(X_n)$ Martingal

$$\Rightarrow h(x) = E_x[h(X_0)] = E_x[\overbrace{h(X_{T \wedge n})}^{\text{beschr.}}] \rightarrow E_x[h(X_T)]$$

$$= h(b) \cdot (1-r(x)) + h(a) \cdot r(x)$$

$$\Rightarrow r(x) = \frac{h(b) - h(x)}{h(b) - h(a)} = \frac{1 - \left(\frac{p}{q}\right)^{b-x}}{1 - \left(\frac{p}{q}\right)^{b-a}}$$

Übung: Berechne $E_x[T]$

(Stoppsetz auf $X_n - n \cdot (2p-1)$ anwenden)

6.7. Waldsche Identität (Zufällige Summen)

(80)

$$S_n = Y_1 + \dots + Y_n, \quad Y_i \in \mathcal{L}^1 \text{ i.i.d.}, \quad E[Y_i] = m.$$

$$M_n := S_n - n \cdot m \text{ ist Martingal bzgl. } \mathcal{F}_n = \sigma(Y_1, \dots, Y_n).$$

Satz T Stopzeit bzgl. (\mathcal{F}_n) mit $E[T] < \infty$

$$\Rightarrow E[S_T] = m \cdot E[T]$$

Beweis: 1) für $Y_i \geq 0$:

$$0 = E[M_0] = E[M_{T \wedge n}] = \underbrace{E[S_{T \wedge n}]}_{\substack{\text{non.wachsend} \\ \text{da } Y_i \geq 0}} - m \cdot \underbrace{E[T \wedge n]}_{\text{mon.wachsend}}$$

$$\xrightarrow{E[T] < \infty} E[S_T] - m \cdot E[T]$$

$$\Rightarrow E[S_T] = m \cdot E[T] < \infty$$

2) allgemeiner Fall: Sei $\bar{S}_n := \sum_{i=1}^n |Y_i|$, $\bar{m} := E[|Y_i|] < \infty$.

$$|S_{T \wedge n}| \leq \sum_{i=1}^{T \wedge n} |Y_i| \leq \sum_{i=1}^T |Y_i| = \bar{S}_T \quad \text{für alle } n \geq 0$$

Nach 1) gilt $E[\bar{S}_T] = \bar{m} \cdot E[T] < \infty$, also $\bar{S}_T \in \mathcal{L}^1$
integrierbare Majorante!

Lebesgue $\Rightarrow 0 = E[S_{T \wedge n}] - \bar{m} \cdot E[T \wedge n] \xrightarrow{n \uparrow \infty} E[S_T] - \bar{m} \cdot E[T]$. \square

7. Das Dirichletproblem für Markovketten

7.1. Kanonisches Modell für (zeitlich homogene) Markovkette

(S, \mathcal{F}) meßbarer Raum

$p(x, dy)$ stochastischer Kern auf (S, \mathcal{F})

$$\Omega := \prod_{i=0,1,2,\dots} S = \{ \omega = (x_0, x_1, \dots) \mid x_i \in S \}$$

$X_n(\omega) := x_n$ n-te Komponente

$$\mathcal{O}_1 := \sigma(X_i \mid i \geq 0), \quad \mathcal{F}_n := \sigma(X_i \mid 0 \leq i \leq n)$$

Fakt Zu jeder WV μ auf (S, \mathcal{S}) existiert genau eine WV P_μ auf (Ω, \mathcal{O}) , so daß $(X_n)_{n \geq 0}$ bzgl. P_μ eine (zeitl. homogene) Markovkette mit Startverteilung μ und Übergangskern p ist, d.h.

① $P_\mu [X_0 \in B] = \mu(B) \quad \forall B \in \mathcal{S}$

② $P_\mu [X_{n+1} \in B | \mathcal{F}_n] = P_\mu [X_{n+1} \in B | X_n] = p(X_n, B) \quad P_\mu$ -f.s.

Notation: $P_x := P_{\delta_x} \quad (x \in S)$, E_μ, E_x : Erw.wert bzgl. P_μ, P_x

Bem. Aus ② folgt für alle $f \geq 0$ meßbar:

$$E_\mu [f(X_{n+1}) | \mathcal{F}_n] = \int p(X_n, dy) f(y) = E_{X_n} [f(X_1)] =: Pf(X_n)$$

bzw. allgemein:

Satz (Markoveigenschaft)

$$E_\mu [f(X_n, X_{n+1}, \dots) | \mathcal{F}_n] = E_{X_n} [f(X_0, X_1, \dots)] \quad P_\mu$$
-f.s.

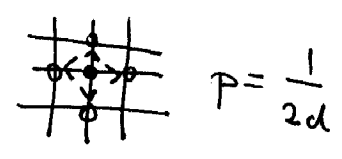
für alle \mathcal{O} -meßbaren $f: \Omega \rightarrow [0, \infty)$. Beweis: s. W' Theorie I

Kompakte Formulierung $\Theta: \mathbb{R} \rightarrow \mathbb{R}$ Shift

$$\Theta((x_0, x_1, x_2, \dots)) = (x_1, x_2, \dots), \quad X_n \circ \Theta = X_{n+1}$$

! $E_\mu[f \circ \Theta^n | \mathcal{F}_n] = E_{X_n}[f]$ P_μ -f.s. $\forall f \geq 0$ meßbar

Standardbeispiel: Random walk auf \mathbb{Z}^d



7.2. Stochastischer Zugang zum Dirichletproblem

Sei $D \in \mathcal{S}$, $f: D^c \rightarrow \mathbb{R}$ meßbar und beschränkt.

Gesucht: $h: S \rightarrow \mathbb{R}$ meßbar mit

1) $h(x) = \underbrace{(ph)}(x) \quad \forall x \in D$ "harmonisch bzgl. p "

(*) $= E_x[h(X_1)]$

2) $h(x) = f(x) \quad \forall x \in D^c$

$T_{D^c} := \inf \{n \geq 0 \mid X_n \in D^c\}$ ist Stopzeit.

Satz Gilt $P_x [T_{D^c} < \infty] = 1 \quad \forall x \in D$, dann hat das Dirichletproblem (*) genau eine beschränkte Lösung, und zwar

$$h(x) = E_x [f(X_{T_{D^c}})]$$

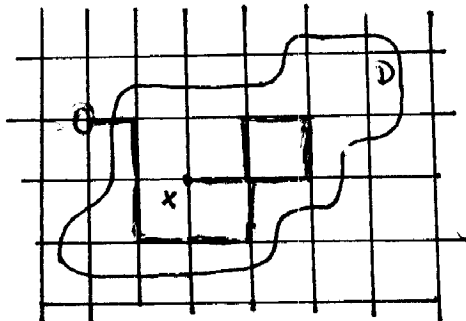
Bem. 1) Darstellung der Lösung als „Pfadintegral“:

$$h(x) = \int_{\{\omega \in \Omega\}} f(X_{T_{D^c}}(\omega)) P_x(d\omega)$$

2) Integralformel für harmonische Funktionen:

$$h(x) = \int_{D^c} f(y) \mu_x(dy) \quad \text{wobei } \mu_x := P_x \circ X_{T_{D^c}}^{-1}$$

Austrittsverteilung der MK bei Start in x (harmonisches Maß)



Beweis a) Einzigkeit:

Sei h beschränkte Lösung von $(*)$, $T := T_{\infty}$.

$(X_{T+n})_{n \geq 0}$ ist bzgl. P_x eine Markovkette mit Kern

$$\tilde{p}(x, \cdot) = \begin{cases} p(x, \cdot) & \forall x \in D \\ \delta_x & \forall x \in D^c \end{cases}$$

$$(ph)(x) = h(x) \quad \forall x \in D \Rightarrow (\tilde{p}h)(x) = \begin{cases} (ph)(x) & (x \in D) \\ h(x) & (x \in D^c) \end{cases} = h(x) \quad \forall x \in S$$

Also ist $h(X_{T+n})$ ($n \geq 0$) ein Martingal. Insbesondere:

$$\underline{h(x)} = E_x[h(X_{T+n})] = E_x[\underbrace{h(X_{T+n})}] \rightarrow E_x[h(X_T)] = \underline{E_x[h(X_T)]}$$

$\rightarrow h(X_T)$ P_x -f.s. da $P_x[T < \infty] = 1$
beschränkt

Beweis: a) Einzigkeit via Stoppzeit:

Sei h beschr. Lösung von $(*)$, $T := T_{D^c}$

$\Rightarrow h(X_n)$ Martingal bzgl. $\mathbb{P}_x \forall x \in S$ (so., 5.4)

$\rightarrow h(x) = E_x[h(X_{T_{D^c}})] \xrightarrow{\text{Lebesgue}} E_x[h(X_T)] \stackrel{x \in D^c}{=} E_x[f(X_T)]$
 $\rightarrow h(X_T)$ \mathbb{P}_x -f.s. da $\mathbb{P}_x[T < \infty] = 1$
 beschränkt

b) Existenz via "first step analysis":

$h(x) := E_x[f(X_T)]$ ist beschränkt, da f beschränkt.

(i) $x \in D \Rightarrow T \geq 1$ \mathbb{P}_x -f.s. $\Rightarrow X_T(\omega) = X_T(\Theta(\omega))$ für \mathbb{P}_x -f.a.w

$$\Rightarrow h(x) = E_x[E_x[f(X_T) | \mathcal{F}_1]] \stackrel{\text{Markovaj.}}{=} E_x[E_x[f(X_T)]] = E_x[f(X_T \circ \Theta)] \mathbb{P}_x\text{-f.s.}$$

$$= E_x[h(X_1)] = (Ph)(x)$$

(ii) $x \in D^c \Rightarrow T = 0$ \mathbb{P}_x -f.s. $\rightarrow h(x) = E_x[f(X_0)] = f(x)$. \square

Bem. I.a. existieren mehrere unbeschr. Lösungen: z.B. random walk auf \mathbb{Z} ,

$D = \mathbb{Z} \setminus \{0\}$, $D^c = \{0\}$: $h(x) = a \cdot x$ ($a \in \mathbb{R}$) sind alle harmonisch obwohl $T_{D^c} < \infty$ \mathbb{P}_x -f.s. $\forall x$.

7.3. Verhalten bei Nicht-eindeutigkeit

Was gilt bei $P_x [T_{D^c} < \infty] \neq 1$?

Für jedes $a \in \mathbb{R}$ ist $h_a(x) := a \cdot P_x [T_{D^c} = \infty]$ eine beschränkte Lösung des Dirichletproblems mit Nullrandbedingung (Übung).

Also:

$P_x [T_{D^c} = \infty] \neq 0$ für ein $x \Rightarrow$ Dirichletproblem ist nicht eindeutig



Satz Sei $f \geq 0$. Dann ist

$$h(x) := E_x [f(X_{T_{D^c}}); T_{D^c} < \infty]$$

die minimale positive Lösung des Dirichletproblems.

Beweis: via Lemma von Fatou, Übung. \square

Bem. Neben h_a existieren i.a. weitere Lösungen des homogenen Dirichletproblems. Zum Teil gibt es stochastische Darstellung über Asymptotik von X_n

\rightarrow Anwendungen z.B. in der Differentialgeometrie, harmonische Funktionen auf Mannigfaltigkeiten.

7.4. Superharmonische Funktionen und Asymptotik von Markovketten

S abzählbar, $p(x,y)$ ($x,y \in S$) stoch. Kern (Matrix).

Sei $(X_n)_{n \geq 0}$ Markovkette mit Kern p in kanonisches Modell (bzgl. P_x),

$T_y := \inf \{n \geq 1 \mid X_n = y\}$ erste Trefferzeit falls $X_0 \neq y$,
erste Rückkehrzeit falls $X_0 = y$.

Satz Es sind äquivalent:

1) $P_x [T_y < \infty] = 1 \quad \forall x,y \in S$

(d.h. die Markovkette ist irreduzibel und rekurrent, s.u.)

2) Jede nicht-negative superharmonische Funktion h ist konstant.

Beweis: 1) \Rightarrow 2): $h \geq 0$ superharm. $\Rightarrow h(X_n)$ Supermartingal

$\Rightarrow h(x) = E_x [h(X_0)] \geq E_x [h(X_{T_y, n})] \quad \forall n$

$\Rightarrow h(x) \geq \lim_{n \rightarrow \infty} E_x [h(X_{T_y, n})] \geq E_x [\lim_{n \rightarrow \infty} h(X_{T_y, n})] = h(y)$

für alle $x, y \in S$

$\Rightarrow h$ ist konstant.

$= h(X_{T_y}) = h(y)$

da $T_y < \infty$ P_x -f.s.

$$2) \Rightarrow 1) \quad h(x) := P_x [T_y < \infty]$$

First step analysis:

$$h(x) = \sum_{z \in S} \underbrace{P_x [T_y < \infty | X_1 = z]}_{= 1 \text{ falls } z=y} \cdot \underbrace{P_x [X_1 = z]}_{= p(x,z)}$$

($P_z [T_y < \infty] = h(z)$ sonst)

$$= \sum_{z \neq y} p(x,z) h(z) + p(x,y) \geq \sum_{z \in S} p(x,z) h(z) \quad \forall x$$

$\Rightarrow h$ superharmonisch $\Rightarrow \exists \alpha \in \mathbb{R} : h(x) \equiv \alpha$

$$\alpha = \sum_{z \neq y} p(x,z) \alpha + p(x,y) = \alpha \cdot \underbrace{\sum_{z \in S} p(x,z)}_{=1} + (1-\alpha) \cdot p(x,y) \quad \forall x$$

Angenommen $\alpha \neq 1$. Dann folgt $p(x,y) = 0 \quad \forall x$.

In diesem Fall ist $\tilde{h}(x) := \delta_{x,y}$ superharmonisch und nicht konstant, denn $\sum_z p(x,z) \tilde{h}(z) = p(x,y) = 0 \leq \tilde{h}(x)$. Widerspruch. Also $\alpha = 1$,

d. h. $P_x [T_y < \infty] = 1 \quad \forall x$. \square

8. Der Konvergenzsatz

X_n ($n=0,1,2,\dots$) Supermartingal bzgl. $(\mathcal{F}_n)_{n \geq 0}$.

(Super-)Martingalkonvergenzsatz:

Gilt $\sup_{n \geq 0} E[X_n^-] < \infty$, dann existiert $X_\infty \in \mathcal{L}^1$ mit $X_n \rightarrow X_\infty$ P-f.s.

Bem. 1) Insbesondere: Nach unten beschr. Supermartingale bzw. (analog) nach oben beschr. Submartingale konvergieren P-f.s. gegen $X_\infty \in \mathcal{L}^1$.

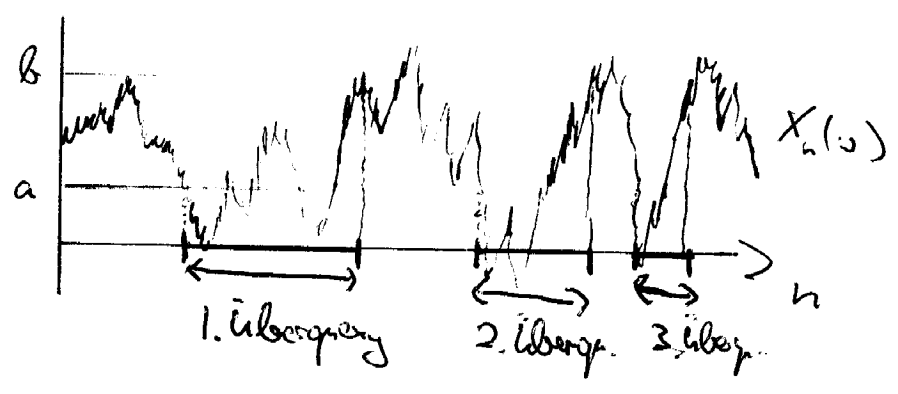
2) I.a. gilt nicht $X_n \rightarrow X_\infty$ in \mathcal{L}^1 .

3) $\sup E[X_n^-] < \infty \iff X_n$ beschränkt in \mathcal{L}^1

denn:
$$E[|X_n|] = \underbrace{E[X_n]}_{\leq E[X_0]} + 2E[X_n^-] \quad (\text{da } |X_n| = X_n + 2X_n^-)$$

Beweisstrategie:

Für $a < b$ sei $U_{a,b}(\omega)$ die Anzahl der Aufwärtssübergänge (Upcrossings) des Intervalls $[a,b]$ durch $X_n(\omega)$.



$$U_{a,b} < \infty \quad \forall a < b \quad \Rightarrow \quad \underline{\lim} X_n = \overline{\lim} X_n$$

$$\Rightarrow X_n \text{ konvergiert in } [-\infty, \infty]$$

Brander Abschätzung für $U_{a,b}$!

8.1. Doob's upcrossing Ungleichung

Für $N \in \mathbb{N}$ und $a < b$ sei

$$U_{a,b}^N(\omega) := \# \text{ upcrossings über } [a,b] \text{ von } X_n(\omega) \text{ bis zur Zeit } N$$

$$= \max \{ k \geq 0 \mid \exists 0 \leq t_0 < t_1 < t_2 < \dots < t_k \leq N : X_{t_i}(\omega) \leq a, X_{t_{i+1}}(\omega) \geq b \}$$

Lemma (Doob) $(b-a) \cdot E[U_{a,b}^N] \leq E[(X_N - a)^-]$

Beweis: Betrachte folgendes Spielstrategie:

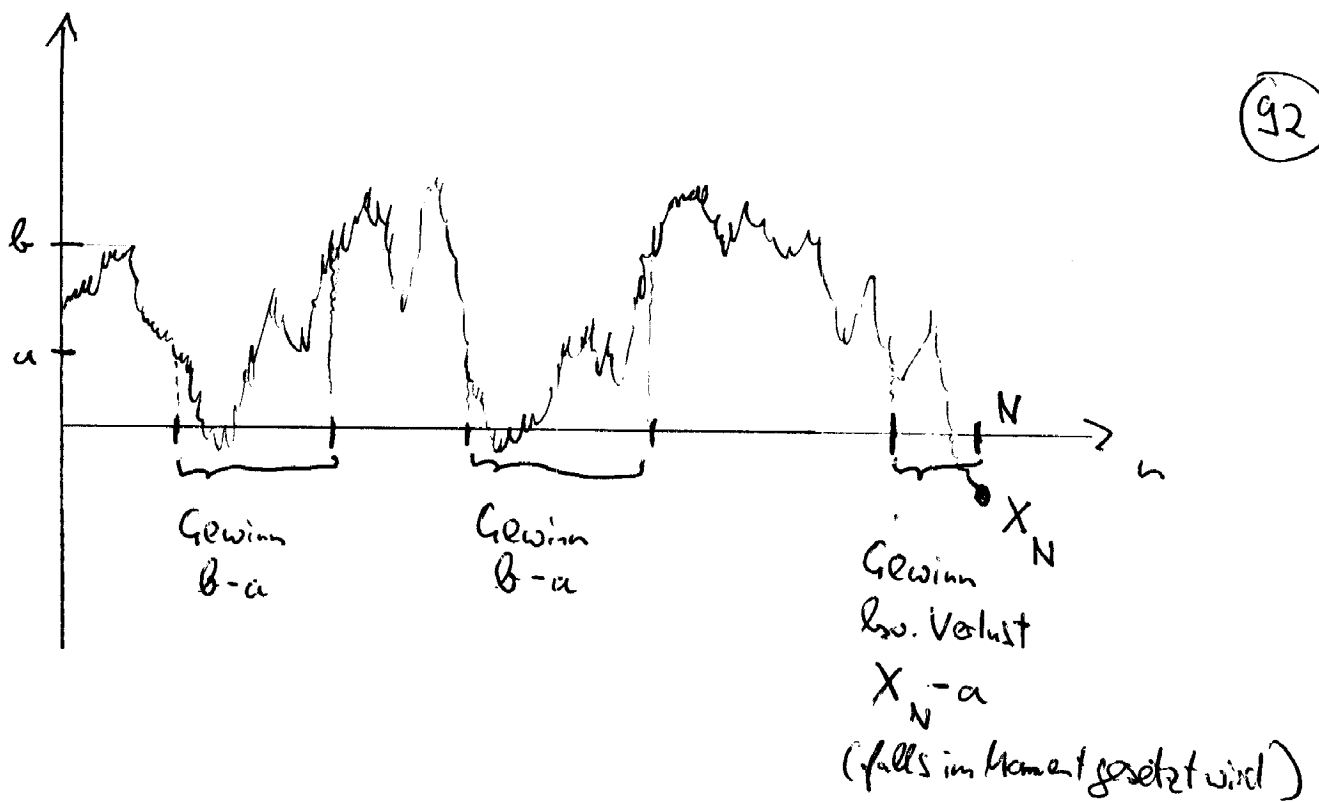
- Warte bis $X_n \leq a$
 - Setze in jeder der folgenden Runden Einsatz 1, solange bis $X_n \geq b$, dann setze keine Einsätze mehr.
- Wiederhole

Einsatz: $C_1 = I_{\{X_0 \leq a\}}$

$$C_n = \begin{cases} 1 & \text{falls } C_{n-1} = 1 \text{ und } X_{n-1} < b \\ & \text{oder falls } C_{n-1} = 0 \text{ und } X_{n-1} \geq a \\ 0 & \text{sonst} \end{cases}$$

$$= I_{\{C_{n-1} = 1, X_{n-1} < b\}} + I_{\{C_{n-1} = 0, X_{n-1} \geq a\}}$$

C_n prädizierbar, $\geq 0 \Rightarrow M_n := (C \cdot X)_n$ ist Supermartingal
(Gewinn zur Zeit n)



$$M_N \geq (b-a) \cdot U_{a,b}^N + X_N - a$$

$$M_N \geq (b-a) \cdot U_{a,b}^N - (X_N - a)^-$$

$$\Rightarrow (b-a) \cdot U_{a,b}^N \leq M_N + (X_N - a)^-$$

M_n Separat, $M_0 = 0$

$$\Rightarrow (b-a) \cdot E[U_{a,b}^N] \leq E[(X_N - a)^-] \quad \square$$

8.2. Beweis des Konvergenzsatzes

Voraussetzung: $\sup_{n \in \mathbb{N}} E[X_n^-] < \infty$

Dann folgt für $a < b$:

$$E[U_{a,b}] \stackrel{\text{Mon. u. v.}}{=} \sum_{N \geq 0} E[U_{a,b}^N] \leq \frac{1}{b-a} \cdot \sup_N E[(X_n - a)^-] \leq |a| + E[X_n^-] < \infty$$

$\Rightarrow U_{a,b} < \infty$ P-f.s.

$\{ \lim X_n \neq \overline{\lim} X_n \} = \bigcup_{\substack{a, b \in \mathbb{Q} \\ a < b}} \{ U_{a,b} = \infty \}$ ist Nullmenge,

also $\exists X_\infty = \lim X_n \in [-\infty, \infty]$ P-f.s.

noch z.z.: $X_\infty \in \mathcal{L}^1$ (also insbes. $X_\infty \in \mathbb{R}$ P-f.s.)

$$\sup E[X_n^-] < \infty \stackrel{\text{Ban. ob.}}{\Rightarrow} \sup E[|X_n|] < \infty$$

$$\Rightarrow E[|X_\infty|] = E[\lim |X_n|] \stackrel{\text{Fatou}}{\leq} \liminf E[|X_n|] < \infty.$$

□

8.3. Beispiel: Randverhalten harmonischer Funktionen auf $D \subset \mathbb{R}^d$

(Kreiskettenverfahren)

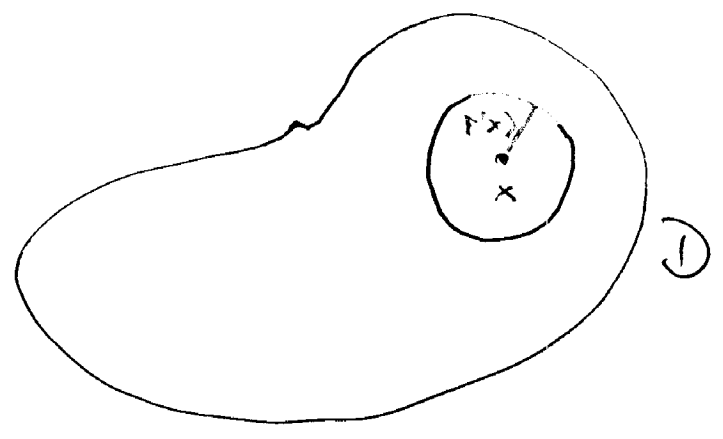
Sei $D \subset \mathbb{R}^d$ beschr. Gebiet, $h: D \rightarrow \mathbb{R}^k$ mit $\Delta h = 0$.
wert nicht beschränkt

$r: D \rightarrow (0, \infty)$ Radiusfunktion

$r(x)$ stetig mit $0 < r(x) < d(x, \partial D) \forall x \in D$

Sei $(X_n)_{n \geq 0}$ bzgl. P_x kanonische Markovkette auf D
mit Übergangskern

$p(x, \cdot) :=$ Gleichverteilung auf $\partial B(x, r(x))$
 $= \{y \in \mathbb{R}^d \mid |y-x| = r(x)\}$



$\Delta h = 0 \Rightarrow h = ph$ auf $D \Rightarrow h(X_n)$ Martingal bzgl. P_x
klassische
Mittelwerteigenschaft

$h \geq \text{const.} \Rightarrow h(X_n)$ P_x -f. konvergiert

Insb. $h(x) = x_i$ ($1 \leq i \leq d$) ist harmonisch, nach unten beschr. auf \bar{D}

$$\Rightarrow P_x\text{-f.s. } \exists X_\infty = \lim X_n, X_\infty \in \partial D$$

Also:

① $\Delta h = 0$, $h \geq 0$ (bzw. nach unten beschr.)

$\Rightarrow \exists \lim h(X_n)$ für fast jede Trajektorie zum Rand
(Regularität!)

② Für $h \in C(\bar{D})$ folgt

$$\lim h(X_n) = h(X_\infty),$$

$$\text{also } \underline{h(x)} = E_x[h(X_n)] \rightarrow \underline{E_x[h(X_\infty)]}$$

Darstellung von h über Randwerte,

$\mu_x = P_x \circ X_\infty^{-1}$ ist harmonisches Maß.

③ Satz von Littlewood: $h \in C(\bar{D})$. Für jedes $x \in D$ gelte die Mittelwertegenschaft nur für die Kugel $B(x, r(x))$.

Dann gilt $\Delta h = 0$.

8.4. Asymptotik von Martingalen mit L^1 -beschränkten Inkrementen

Satz Sei X_n ($n=0,1,2,\dots$) ein Martingal mit in L^1 beschränkten Inkrementen, d. h. $\exists Y \in L^1$ mit

$$|X_{n+1} - X_n| \leq Y \quad \forall n \in \mathbb{N}$$

Dann gilt für \mathbb{P} -f. a. ω :

entweder: $\exists \lim_{n \rightarrow \infty} X_n(\omega) \in \mathbb{R}$ (konvergent)

oder: $\overline{\lim} X_n(\omega) = +\infty$ und $\underline{\lim} X_n = -\infty$ (unbeschr. Oszillationen)

Beweis: Für $c \in \mathbb{R}$ sei $T_c := \min\{n \geq 0 \mid X_n \leq c\}$.

$(X_{T_c \wedge n})_{n \geq 0}$ ist Martingal, $X_{T_c \wedge n} \geq \begin{cases} X_0 & \text{falls } X_0 \leq c \\ c - Y & \text{falls } X_0 > c \end{cases}$
invariant

konv. Satz
 $\implies X_{T_c \wedge n}$ konvergiert \mathbb{P} -f. a.

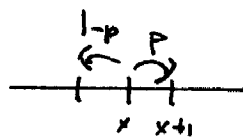
$\implies X_n$ konvergiert \mathbb{P} -f. a. auf $\{T_c = \infty\} \supseteq \{\inf X_n > c\}$

Also konvergiert X_n \mathbb{P} -f. a. auf

$$\{\underline{\lim} X_n > -\infty\} = \{\inf X_n > -\infty\} = \bigcup_{\substack{c < 0 \\ c \in \mathbb{Q}}} \{\inf X_n > c\}$$

Analog: X_n konvergiert \mathbb{P} -f. a. auf $\{\overline{\lim} X_n < \infty\}$. \square

Beispiel Klassischer random walk



$p = \frac{1}{2}$: Martingal mit beschr. Inkrementen

\Rightarrow P-f.s. konvergent oder unbeschr. Oszillationen
 unmöglich da $|X_{n+1} - X_n| = 1 \forall n$

$$\Rightarrow P[\underline{\lim} X_n = -\infty, \overline{\lim} X_n = +\infty] = 1$$

$p > \frac{1}{2}$: Nach Gesetze der großen Zahlen gilt $\lim X_n = +\infty$ P-f.s.

Grund: X_n ist Submartingal, daher fixiert obiges Argument
 nur noch in eine Richtung ($\lim X_n = \infty$).

8.5. Verallgemeinertes Borel-Cantelli Lemma

Sei $(\mathcal{F}_n)_{n \geq 0}$ Filtration, $A_n \in \mathcal{F}_n$ ($n = 0, 1, 2, \dots$).

Satz P-f.s. gilt:

$$A_n \text{ oft} \Leftrightarrow \sum_{n=1}^{\infty} P[A_n | \mathcal{F}_{n-1}] = \infty$$

Klassischer Spezialfall: A_{n+1} unabhängig von \mathcal{F}_n

$$\Rightarrow \sum_{k=1}^{\infty} P[A_k | \mathcal{F}_{k-1}] = \sum_{k=1}^{\infty} P[A_k] \text{ P-f.s.}$$

$$\text{Also: } \sum P[A_n] = \infty \Rightarrow P[A_n \text{ oft}] = 1$$

$$\sum P[A_n] < \infty \Rightarrow P[A_n \text{ oft}] = 0$$

Beweis Satz: $A_n \text{ oft} \Leftrightarrow \sum I_{A_n} = \infty \stackrel{\text{Beh. P-f.s.}}{\Leftrightarrow} \sum E[I_{A_n} | \mathcal{F}_{n-1}] = \infty$

$$\text{Sei } S_n := \sum_{k=0}^n I_{A_k}, \quad Z_n := \sum_{k=0}^n E[I_{A_k} | \mathcal{F}_{k-1}].$$

S_n und Z_n sind P-f.s. monoton wachsend, also konvergent in $[0, \infty]$.

$S_n - Z_n$ ist Martingal mit beschr. Inkrementen, also gilt P-f.s.:

$$\text{entweder } \exists \lim (S_n - Z_n) \in \mathbb{R} \Rightarrow S_{\infty} = Z_{\infty}$$

$$\text{oder } S_n - Z_n \text{ oszilliert unbeschr.} \Rightarrow S_{\infty} = \infty \text{ und } Z_{\infty} = \infty$$

In beiden Fällen gilt $S_{\infty} = Z_{\infty}$. \square

8.6. Gleichmäßig integrierbare Martingale sind sukzessive Projektionen

$X \in \mathcal{L}^1 \Rightarrow X_n := E[X | \mathcal{F}_n]$ ist Martingal bzgl. (\mathcal{F}_n)

Läßt sich jedes Martingal so darstellen?

Ja, falls gleichmäßig integrierbar:

Satz Sei $(X_n)_{n \geq 0}$ Martingal bzgl. (\mathcal{F}_n) . Dann sind äquivalent:

(i) $\{X_n | n \geq 0\}$ ist gleichmäßig integrierbar.

(ii) X_n konvergiert bzgl. der L^1 -Norm

(iii) $\exists X \in \mathcal{L}^1 : X_n = E[X | \mathcal{F}_n]$

Erinnerung: $\{X_n | n \geq 0\}$ gleichmäßig int. bar $\iff \sup_{n \geq 0} E[|X_n|; |X_n| \geq c] \rightarrow 0$
für $c \rightarrow \infty$

Hinreichende Bedingungen: • integrierbare Majorante

• $\sup_n E[|X_n|^p] < \infty$ für ein $p > 1$

oder • $\sup_n E[|X_n| \log^+ |X_n|] < \infty$

Beweis Satz: (iii) \Rightarrow (i): Übung

(i) \Rightarrow (ii): $\{X_n | n \geq 0\}$ glm. int. bar \Rightarrow beschränkt in L^1

$$(E|X_n|) \leq E[|X_n|; |X_n| \geq c] + c \leq 1 \text{ für } c \text{ groß}$$

Mart. konv. Satz $\Rightarrow \exists X_\infty = \lim X_n$ P-f.s.
glm. int. bar $\Rightarrow X_n \rightarrow X_\infty$ in L^1

(ii) \Rightarrow (iii): $X_n \rightarrow X_\infty$ in $L^1 \Rightarrow X_n = E[X_\infty | \mathcal{F}_n]$

denn: • X_n ist \mathcal{F}_n -messbar

$$\bullet A \in \mathcal{F}_n \rightarrow E[X_\infty; A] \stackrel{L^1 \text{ konv.}}{=} \lim_{k \rightarrow \infty} E[X_k; A] = E[X_n; A] = E[X_n; A] \quad \square$$

Kor. $X \in L^1 \Rightarrow E[X | \mathcal{F}_n] \rightarrow E[X | \mathcal{F}_\infty]$ P-f.s. und in L^1
wobei $\mathcal{F}_\infty = \sigma(\cup_{n \geq 0} \mathcal{F}_n)$

Beweis: Sei $X_n = E[X | \mathcal{F}_n]$. Nach Satz $\exists X_\infty = \lim X_n$ P-f.s. und in L^1
zz: $X_\infty = E[X | \mathcal{F}_\infty]$.

1) X_∞ ist \mathcal{F}_∞ -messbar ✓

2) $A \in \mathcal{F}_n, n \geq 0$

$$\Rightarrow E[X_\infty; A] \stackrel{\text{s. Beweis}}{\stackrel{\text{Satz}}{=}} E[X_n; A] \stackrel{E[X_n; A]}{=} E[X; A]$$

$$\text{Also: } \int_A X_\infty dP = \int_A X dP \quad \forall A \in \underbrace{\bigcup_{n \geq 0} \mathcal{F}_n}_{\text{durchschnittsstabil}}$$

und damit auch $\forall A \in \sigma(\bigcup_{n \geq 0} \mathcal{F}_n)$. □

8.7. Existenz der bedingten Erwartung (Alternativer Existenzbeweis)

Sei $X \in \mathcal{L}^1, \mathcal{F} \subseteq \mathcal{O}$ σ -Algebra.

Existenz von $E[X|\mathcal{F}]$?

Annahme: \mathcal{F} ist separabel, d.h. $\exists A_1, A_2, \dots \in \mathcal{O}: \mathcal{F} = \sigma(A_i; i \in \mathbb{N})$

Sei $\mathcal{F}_n := \sigma(A_1, \dots, A_n)$.

ist atomar, d.h. $\mathcal{F}_n = \sigma(B_1, B_2, \dots, B_k)$ mit endlich vielen disjunkten Mengen $B_i, \bigcup_{i=1}^k B_i = \Omega$.

$$\Rightarrow E[X|\mathcal{F}_n] \text{ ist elementar definiert: } = \sum_{P(B_i) > 0} E[X|B_i] \cdot \mathbb{I}_{B_i}$$

$$X_\infty := \lim_{n \rightarrow \infty} E[X | \mathcal{F}_n] \quad (\text{konvergiert } P\text{-f.s. und in } L^1)$$

ist Version von $E[X | \mathcal{F}]$.

8.8. 0-1-Gesetz von P. Lévy Sei $\mathcal{F}_\infty := \sigma(\cup_{n \geq 0} \mathcal{F}_n)$.

Kor. $A \in \mathcal{F}_\infty \Rightarrow P[A | \mathcal{F}_n] \rightarrow \mathbb{I}_A$ P -f.s.

Hieraus folgt das 0-1-Gesetz von Kolmogorov:

$\sigma_1, \sigma_2, \dots$ unabhängige σ -Algebren,

$\mathcal{F}_n := \sigma(\sigma_1, \dots, \sigma_n)$, $\hat{\mathcal{F}}_n := \sigma(\sigma_{n+1}, \sigma_{n+2}, \dots)$ unabhängig

Teil field: $\mathcal{T} := \bigcap_n \hat{\mathcal{F}}_n$

$A \in \mathcal{T} \Rightarrow A$ unabhängig von \mathcal{F}_n für alle n

$$\Rightarrow P[A | \mathcal{F}_n] = P[A] \quad \forall n$$

$$\rightarrow \mathbb{I}_A \text{ } P\text{-f.s.}$$

$$\Rightarrow P[A] \in \{0, 1\}$$

9. Doobzerlegung und Varianzprozeß

103

$(\mathcal{F}_n)_{n \geq 0}$ Filtration auf (Ω, \mathcal{O}, P)

9.1. Doobzerlegung

Sei X_n ($n=0,1,2,\dots$) adaptiert bzgl. (\mathcal{F}_n) , $\in \mathcal{L}^1$.

Satz $\exists!$ Zerlegung $X_n = M_n + A_n$ (eindeutige Äquivalenzklasse bzgl. P)
in ein Martingal $(M_n)_{n \geq 0}$ und einen previsiblen Prozeß $(A_n)_{n \geq 0}$
mit $A_0 = 0$, und zwar

$$(*) \quad A_n = \sum_{k=0}^{n-1} E[X_{k+1} - X_k | \mathcal{F}_k], \quad M_n = X_n - A_n.$$

Beweis: 1) Eindeutigkeit: Für jede Zerlegung wie oben gilt:

$$X_{k+1} - X_k = M_{k+1} - M_k + A_{k+1} - A_k \quad \forall k \geq 0$$

$$\begin{aligned} \Rightarrow E[X_{k+1} - X_k | \mathcal{F}_k] &= \underbrace{E[M_{k+1} - M_k | \mathcal{F}_k]}_{= 0 \text{ da } M_n \text{ Martingal}} + \underbrace{E[A_{k+1} - A_k | \mathcal{F}_k]}_{= A_{k+1} - A_k \text{ da previsibel}} \quad \forall k \geq 0 \\ &= A_{k+1} - A_k \end{aligned}$$

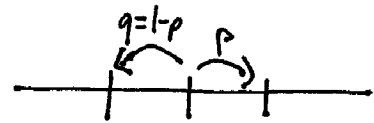
$\Rightarrow (*)$

2) Existenz: Definiere A_n, M_n via (*)

$\Rightarrow A_n$ previsibel, M_n Martingal □

Bem. X_n Supermartingal $\Leftrightarrow A_n$ mon. fallend
Submartingal $\Leftrightarrow A_n$ mon. wachsend

dem: $A_{n+1} - A_n = E[X_{n+1} - X_n | \mathcal{F}_n]$.

Beispiel X_n klass. random walk 

Doobzerlegung für X_n : $X_n = M_n + (p-q) \cdot n$

Doobzerlegung für X_n^2 im Fall $p = \frac{1}{2}$: $M_n := X_n^2 - n$ ist Martingal

9.2. Doobzerlegung für Markovketten - Martingalproblem

Sei $(X_n)_{n \geq 0}$ Markovkette auf (S, \mathcal{S}) mit Übergangskern $p(x, dy)$,

$f: S \rightarrow \mathbb{R}$ meßbar mit $f(X_n) \in \mathcal{L}^1 \forall n$. Doob-Zerlegung:

$$f(X_n) = f(X_0) + M_n^f + \sum_{k=0}^{n-1} (p - I)f(X_k)$$

$I f = f$
Identität

Basis: $A_{n+1} - A_n = E[f(X_{n+1}) - f(X_n) | \mathcal{F}_n] = (p f)(X_n) - f(X_n)$. \square

Insbesondere:

f (super-) harmonisch, d.h. $(p-I)f \stackrel{(\leq)}{=} 0 \Rightarrow f(X_n)$ (Super-)Martingal

Ferguson (Stroock/Varadhan): X_n löst das Martingalproblem mit Generator $p-I$,

falls $M_n^f := f(X_n) - f(X_0) - \sum_{k=0}^{n-1} (p-I)f(X_k)$ für alle meßb.

f mit $f(X_n) \in \mathcal{L}^1$ ein Martingal ist.

Satz X_n löst Martingalproblem zu $p-I \Leftrightarrow X_n$ Markovkette mit Übergangskern p

J.3. Varianzprozess M_n Martingal, $\in \mathcal{L}^2 \xRightarrow{\text{Dense}} M_n^2$ Submartingal

Doobzerlegung: $M_{n+1}^2 - M_n^2 = (M_{n+1} - M_n)(M_{n+1} + M_n) = 2(M_{n+1} - M_n) \cdot M_n + (M_{n+1} - M_n)^2$

$$\Rightarrow E[M_{n+1}^2 - M_n^2 | \mathcal{F}_n] = E[(M_{n+1} - M_n)^2 | \mathcal{F}_n]$$

Also $M_n^2 = \text{Martingal} + \langle M \rangle_n$

wobei

$$\langle M \rangle_n := \sum_{k=0}^{n-1} E[(M_{k+1} - M_k)^2 | \mathcal{F}_k] = \sum_{k=0}^{n-1} \text{var}(M_{k+1} - M_k | \mathcal{F}_k)$$

Definition $\langle M \rangle_n$ ($n=0,1,2,\dots$) heißt Varianzprozess von M .

Beispiel $M_n = \sum_{i=1}^n Y_i$, Y_i unabh. und zentriert, $\mathcal{F}_n = \sigma(Y_1, \dots, Y_n)$

$$\Rightarrow \langle M \rangle_n = \sum_{k=1}^n \text{var}(Y_k) = \text{var}(M_n)$$

Bem. 1) $\langle M \rangle_n \nearrow \langle M \rangle_\infty = \sum_{k=0}^{\infty} E[(M_{k+1} - M_k)^2 | \mathcal{F}_k] \in [0, \infty]$

$$2) E[M_n^2] = E[\langle M \rangle_n] \quad \forall n$$

10. Quadratintegrierbare Martingale

Sei $(M_n)_{n \geq 0}$ Martingal bzgl. $(\mathcal{F}_n)_{n \geq 0}$ auf (Ω, \mathcal{O}, P) .

Annahme: $M_n \in \mathcal{L}^2 \quad \forall n \geq 0$

10.1. Orthogonalität der Inkremente

$$M_n = M_0 + \sum_{k=1}^n Y_k, \quad Y_k := M_k - M_{k-1}$$

$$M_n \text{ Martingal} \Leftrightarrow E[Y_k | \mathcal{F}_{k-1}] = 0 \quad \forall k$$

Konsequenz: Die Inkremente Y_k sind zentriert und orthogonal in \mathcal{L}^2
(unkorrelliert):

$$E[Y_k] = 0$$

$$\text{cov}(Y_k, Y_\ell) = E[Y_k Y_\ell] = E[Y_k \cdot \underbrace{E[Y_\ell | \mathcal{F}_k]}_{=0}] = 0 \quad \forall k \neq \ell$$

Insbesondere:

$$\underline{E[M_n^2]} = \left(M_0 + \sum_{k=1}^n Y_k, M_0 + \sum_{k=1}^n Y_k \right)_{\mathcal{L}^2} \stackrel{\text{alles orthog.}}{=} (M_0, M_0)_{\mathcal{L}^2} + \sum_{k=1}^n (Y_k, Y_k)_{\mathcal{L}^2}$$

$$(*) \quad \underline{=} E[M_0^2] + \sum_{k=1}^n E[Y_k^2]$$

10.2. L^2 Konvergenz von Martingalen

(108)

Satz Äquivalent:

(i) M_n konvergiert in L^2 .

(ii) $\sup E[M_n^2] < \infty$

(iii) $\sum_{k=1}^{\infty} E[(M_k - M_{k-1})^2] < \infty$

In diesem Fall gilt auch P-f.s. Konvergenz.

Beweis: (i) \Rightarrow (ii) klar, (ii) \Leftrightarrow (iii) wegen (*).

(ii) + (iii) \Rightarrow (i): M_n beschr. in $L^2 \Rightarrow$ beschr. in $L^1 \xrightarrow[\text{Fatou}]{\text{Konv.}} \exists M_{\infty} = \lim M_n$ P-f.s.

$$\begin{aligned} \text{Zudem: } E[(M_{\infty} - M_n)^2] &\stackrel{\text{Fatou}}{\leq} \lim_{n \rightarrow \infty} E[(M_n - M_n)^2] = \lim_{n \rightarrow \infty} \sum_{k=n+1}^{\infty} E[Y_k^2] \\ &= \sum_{k=n+1}^{\infty} E[Y_k^2] \xrightarrow{n \uparrow \infty} 0 \text{ nach (iii).} \quad \square \end{aligned}$$

Bem. (iii) $\Leftrightarrow E[\langle M \rangle_{\infty}] < \infty$.

Zusammenfassung Konvergenzsätze:

1) M_n beschränkt in \mathcal{L}^1 (bzw. $\sup E[M_n^+] < \infty$)

$\Rightarrow M_n \rightarrow M_\infty$ P-f.s., $M_\infty \in \mathcal{L}^1$

2) $M_n \rightarrow M_\infty$ in $L^1 \Leftrightarrow \{M_n | n \in \mathbb{N}\}$ gleichmäßig integrierbar

3) $M_n \rightarrow M_\infty$ in $L^2 \Leftrightarrow M_n$ beschränkt in L^2

$\Leftrightarrow \sum_1^\infty E[(M_k - M_{k-1})^2] < \infty$

Genauere Analyse via Varianzprozess, s. 10.4.

10.3. Random Signs

$$\sum_{n=1}^{\infty} \frac{1}{n^{\alpha}} \text{ konvergiert } \Leftrightarrow \alpha > 1$$

$$\sum_{n=1}^{\infty} (-1)^n \frac{1}{n^{\alpha}} \text{ konvergiert } (\Leftrightarrow) \alpha > 0$$

Was passiert bei zufälligen Vorzeichen $\varepsilon_n = \begin{cases} +1 & \text{mit } p = \frac{1}{2} \\ -1 & \text{mit } p = \frac{1}{2} \end{cases}$, unabh.?

Kor. 1 Sei $a_n \in \mathbb{R}$ eine Folge. Dann gilt:

$$\sum_{n=1}^{\infty} \varepsilon_n a_n \text{ konv. in } L^2(\mathbb{R}, \mathcal{O}, P) \Leftrightarrow \sum_{n=1}^{\infty} a_n^2 < \infty$$

Beweis: $M_n = \sum_{k=1}^n \varepsilon_k a_k$ ist Martingal, $\sum_{k=1}^{\infty} E[(M_k - M_{k-1})^2] = \sum_{k=1}^{\infty} a_k^2$. \square

Kor. 2 $\sum a_n^2 < \infty \Rightarrow \sum_{k=1}^{\infty} \varepsilon_k a_k$ konv. P-f.s.

$\sum a_n^2 = \infty \Rightarrow M_n = \sum_{k=1}^n \varepsilon_k a_k$ hat P-f.s. unbeschr. Oszillationen

Beweis: Übung \square

Beispiel $\sum_{n=1}^{\infty} \frac{\varepsilon_n}{n^{\alpha}}$ konv. in L^2 bzw. P-f.s. $\Leftrightarrow \alpha > \frac{1}{2}$

10.4. Varianzprozess und P-f.s. Konvergenz

M_n Martingal, $\in L^2$, $\langle M \rangle_n = \sum_{k=1}^n E[(M_k - M_{k-1})^2 | \mathcal{F}_{k-1}]$ Varianzprozess
previsibel

$$M_n^2 = \text{Martingal} + \langle M \rangle_n$$

Bisher: Nach 10.2. gilt: $E[\langle M \rangle_\infty] < \infty \implies M_n$ konv. in L^2 und P-f.s.

Verfeinerung:

Satz $M_n \rightarrow$ endl. Limes P-f.s. auf $\{\langle M \rangle_\infty < \infty\}$.

Beweis: z.z.: Für jedes $c > 0$ konvergiert M_n P-f.s. auf $\{\langle M \rangle_\infty \leq c\}$.

$T := \min \{n \geq 0 \mid \underbrace{\langle M \rangle_{n+1}}_{\in \mathcal{F}_n} \geq c\}$ ist Stoppzeit bzgl. (\mathcal{F}_n)

$\xrightarrow{\text{Stoppzeit}} M_n^T := M_{T \wedge n} \ (n \geq 0)$ ist Martingal

Varianzprozess: $\langle M^T \rangle_n = \langle M \rangle_{T \wedge n} \leq c \ \forall n$

$\implies \langle M^T \rangle_\infty \leq c \implies M_n^T$ konv. (in L^2 und) P-f.s.

$\implies M_n$ konv. P-f.s. auf $\{T = \infty\} \supseteq \{\langle M \rangle_\infty \leq c\}$.

„Lokalisieren durch Stoppen“



Kor. $P[\langle M \rangle_\infty < \infty] = 1 \Rightarrow M_n$ konvergiert P-f.s.

Was passiert auf $\{\langle M \rangle_\infty = \infty\}$?

Gesetz der großen Zahlen: $\frac{M_n}{\langle M \rangle_n} \rightarrow 0$ P-f.s. (s. 10.5)

Satz vom iterierten Logarithmus: $\overline{\lim} \frac{M_n}{\sqrt{2 \langle M \rangle_n \log \log \langle M \rangle_n}} = +1, \underline{\lim} = -1$

(Beweis im Spezialfall s. 22.5, Kor. 4)

10.5. Gesetz der großen Zahlen für Martingale

Satz Sei $f: \mathbb{R}^+ \rightarrow \mathbb{R}^+$ mon. wachsend mit $\int_0^\infty \frac{1}{(1+f(x))^2} dx < \infty$.

Dann gilt:

$$\frac{M_n}{f(\langle M \rangle_n)} \rightarrow 0 \text{ P-f.s. auf } \langle M \rangle_\infty = \infty.$$

Beispiele: 1) $f(x) = x^\alpha, \alpha > \frac{1}{2}$: $\frac{M_n}{\langle M \rangle_n^\alpha} \rightarrow 0$ P-f.s. auf $\langle M \rangle_\infty = \infty$.

2) $f(x) = \sqrt{x} \cdot (\log x)^\beta, \beta > \frac{1}{2}$.

Beis. 1) Klassisches GdGZ: $Y_k \in L^2$ unabhängig, $\sup E[Y_k^2] < \infty$ (112)

$$\Rightarrow \frac{M_n}{n^\alpha} \rightarrow 0 \text{ P-f.s. } \forall \alpha > \frac{1}{2}$$

$$\text{Basis: } \langle M \rangle_n = \sum_{k=1}^n \text{var}(Y_k | \mathcal{F}_{k-1}) = \sum_{k=1}^n \text{var}(Y_k) \leq \text{const.} \cdot n$$

$$\Rightarrow \frac{M_n}{n^\alpha} \leq \text{const.} \cdot \frac{M_n}{\langle M \rangle_n^\alpha} \rightarrow 0 \quad \square$$

Aussage gilt auch ohne Unabhängigkeit! \rightarrow Übung

2) Die Aussage in Beispiel 1 ist scharf. Für $\alpha = \frac{1}{2}$ gilt im klass. iid Fall:

$$\overline{\lim} \frac{M_n}{\sqrt{n}} = +\infty, \quad \underline{\lim} \frac{M_n}{\sqrt{n}} = -\infty$$

(folgt aus zentralen Grenzwertsatz, s. Übung zu WY Theorie I)

3) Genauere Asymptotik bei beschränkten Inkrementen:

Satz vom iterierten Logarithmus: Auf $\{\langle M \rangle_n = \infty\}$ gilt P-f.s.:

$$\overline{\lim} \frac{M_n}{\sqrt{2 \langle M \rangle_n \log \log \langle M \rangle_n}} = +1, \quad \underline{\lim} \frac{M_n}{\sqrt{2 \langle M \rangle_n \log \log \langle M \rangle_n}} = -1.$$

(Spezialfall s. 22.5)

" Variationsprozess mit Wachstumsgeschwindigkeit von $M_n \sim \sqrt{n}$

Beweis Satz: $\tilde{M} := \underbrace{(1 + f(\langle M \rangle))}_{\text{previsibel}} \cdot M$ Martingaltransformation 113

d.h. $\tilde{M}_n = \sum_{k=1}^n \tilde{Y}_k$ mit $\tilde{Y}_k = \frac{Y_k}{1 + f(\langle M \rangle_k)}$ "Renormierung der Inkremente"

Beh. 1 \tilde{M}_n konvergiert P-f.s.:

$$\begin{aligned} \langle \tilde{M} \rangle_n &= \sum_{k=1}^n E[\tilde{Y}_k^2 | \mathcal{F}_{k-1}] = \sum_{k=1}^n \frac{E[Y_k^2 | \mathcal{F}_{k-1}]}{(1 + f(\langle M \rangle_k))^2} \\ &= \sum_{k=1}^n \frac{\langle M \rangle_k - \langle M \rangle_{k-1}}{(1 + f(\langle M \rangle_k))^2} \stackrel{f \uparrow}{\leq} \sum_{k=1}^n \int_{\langle M \rangle_{k-1}}^{\langle M \rangle_k} \frac{dx}{(1 + f(x))^2} \\ &= \int_0^{\langle M \rangle_n} \frac{dx}{(1 + f(x))^2} \leq \int_0^\infty \frac{dx}{(1 + f(x))^2} \end{aligned}$$

Voraus. $\Rightarrow \langle \tilde{M} \rangle_\infty < \infty$ P-f.s. $\Rightarrow \tilde{M}_n$ konvergiert P-f.s.

Beh. 2 Hieraus folgt $\frac{M_n}{f(\langle M \rangle_n)} \rightarrow 0$ P-f.s. auf $\langle M \rangle_\infty = \infty$:

via Kronecker's Lemma (s.u.): $\tilde{M}_n = \sum_{k=0}^n \frac{Y_k}{1 + f(\langle M \rangle_k)}$ konvergiert

Kronecker $\Rightarrow \frac{\sum_{k=0}^n Y_k}{1 + f(\langle M \rangle_n)} \rightarrow 0$ falls $f(\langle M \rangle_n) \rightarrow \infty$
 Ok auf $\{\langle M \rangle_\infty = \infty\}$ \square

10.6. Kronecker's Lemma $b_n > 0$ mon. wachsend, $\nearrow \infty$.

Dann gilt:

$$\sum_{k=1}^{\infty} \frac{x_k}{b_k} \text{ konvergent} \Rightarrow \frac{1}{b_n} \sum_{k=1}^n x_k \rightarrow 0$$

Beispiel: $a_n \downarrow 0$, $\sum a_k < \infty \Rightarrow n \cdot a_n \rightarrow 0$ ($a_n = \frac{1}{b_n}$, $x_k = 1$)

Beweis: $S_n := \sum_{k=1}^n \frac{x_k}{b_k}$ konvergent.

$$\sum_{k=1}^n x_k = \sum_{k=1}^n b_k \cdot (s_k - s_{k-1}) = b_1 \cdot (s_1 - s_0) + b_2 \cdot (s_2 - s_1) + \dots + b_n \cdot (s_n - s_{n-1})$$

partielle
Summation

$$\begin{aligned} &= s_1 \cdot (b_1 - b_2) + s_2 \cdot (b_2 - b_3) + \dots + s_{n-1} \cdot (b_{n-1} - b_n) + s_n \cdot b_n \\ &= \sum_{k=0}^{n-1} -s_k \cdot (b_{k+1} - b_k) + s_n \cdot b_n \\ &= \sum_{k=0}^{n-1} (s_n - s_k) (b_{k+1} - b_k) \\ &< \varepsilon \text{ für } n, k \geq N(\varepsilon), \text{ da konvergent} \end{aligned}$$

$$\Rightarrow \left| \sum_{k=1}^n x_k \right| \leq \underbrace{\left| \sum_{k=0}^{N(\varepsilon)} (s_n - s_k) \cdot (b_{k+1} - b_k) \right|}_{\text{beschränkt in } n} + \varepsilon \cdot b_n \leq 2\varepsilon b_n$$

für n groß genug

$$\Rightarrow \lim_{n \rightarrow \infty} \left| \frac{1}{b_n} \sum_{k=1}^n x_k \right| \leq 2\varepsilon \quad \forall \varepsilon > 0. \quad \square$$

11. Differentiation von Maßen

(Ω, \mathcal{O}) meßbarer Raum, P, Q Wahrscheinlichkeitsverteilungen auf (Ω, \mathcal{O}) .

11.1. Absolutstetigkeit Sei $\mathcal{F} \subseteq \mathcal{O}$ σ -Algebra.

Definition 1) Q heißt absolutstetig bzgl. P auf \mathcal{F} ($Q \ll P$), falls:

$$P[A] = 0 \Rightarrow Q[A] = 0 \quad \forall A \in \mathcal{F}$$

2) Q, P äquivalent ($Q \approx P$ auf \mathcal{F}) $\Leftrightarrow Q \ll P$ und $P \ll Q$

3) Q, P singular ($Q \perp P$ auf \mathcal{F}) $\Leftrightarrow \exists A \in \mathcal{F}: P[A] = 0 \wedge Q[A] = 1$.

Beispiel $\delta_{\frac{1}{2}} \perp \lambda_{(0,1)}$

Lemma $Q \ll P$ auf \mathcal{F}

$$\Leftrightarrow \forall \varepsilon > 0 \exists \delta > 0 \forall A \in \mathcal{F}: P[A] < \delta \Rightarrow Q[A] < \varepsilon \quad (*)$$

Beweis: " \Leftarrow ": $P[A] = 0 \stackrel{(*)}{\Rightarrow} Q[A] < \varepsilon \forall \varepsilon > 0 \Rightarrow Q[A] = 0$.

" \Rightarrow ": Angenommen $(*)$ gilt nicht.

$$\Rightarrow \exists \varepsilon > 0, A_n \in \mathcal{F} (n=1,2,3,\dots): Q[A_n] \geq \varepsilon, P[A_n] \leq 2^{-n}$$

Sei $A := \bigcap_n \bigcup_{m \geq n} A_m = \{A_n \text{ unendlich oft}\} \in \mathcal{F}$.

$P[A] = 0$ nach Borel-Cantelli, aber:

$$Q[A] = \lim_{n \rightarrow \infty} Q\left[\bigcup_{m \geq n} A_m\right] \geq \varepsilon. \quad \text{Also } Q \not\ll P. \quad \square$$

11.2. Der Satz von Radon-Nikodym

Satz $Q \ll P$ auf \mathcal{F}

$$\Leftrightarrow \exists X \in \mathcal{L}^1(\mathcal{R}, \mathcal{F}, P): Q[A] = \int_A X dP \quad \forall A \in \mathcal{F} \quad (*)$$

Beweis s.u., 11.4.

Bem. X ist durch (*) Q -f.s. eindeutig festgelegt.

Notation: $X = \frac{dQ}{dP} \Big|_{\mathcal{F}}$ Radon-Nikodym Dichte, likelihood ratio⁴

Einfachster Spezialfall: \mathcal{F} atomar,

d.h. $\mathcal{F} = \sigma(B_1, B_2, \dots, B_n)$, $\Omega = \bigcup_{i=1}^n B_i$ disj. Zerlegung in Atome.

$$Q \ll P \Leftrightarrow [P[B_i] = 0 \Rightarrow Q[B_i] = 0]$$

In diesem Fall ist

$$\frac{dQ}{dP} = \sum_{i: P(B_i) > 0} \frac{Q(B_i)}{P(B_i)} \cdot \mathbb{I}_{B_i} \quad \text{Dichte}$$

11.3. Lokale Absolutstetigkeit

Sei $\mathcal{F}_0 \subseteq \mathcal{F}_1 \subseteq \mathcal{F}_2 \subseteq \dots \subseteq \mathcal{F}$ Filtration.

Definition $Q \ll_{loc} P \Leftrightarrow Q \ll P$ auf \mathcal{F}_n für alle n

Beispiel 1 $P = \lambda_{[0,1]}$, $Q = \delta_0$

$$\mathcal{F}_n = \sigma \left(\left[\frac{k}{2^n}, \frac{k+1}{2^n} \right) \mid k=0,1,\dots,2^n-1 \right)$$

$Q \perp P$ aber $Q \ll P$ auf $\mathcal{F}_n \forall n$

$$\left. \frac{dQ}{dP} \right|_{\mathcal{F}_n} = 2^n \cdot \mathbb{I}_{[0, 2^{-n})}$$

Beispiel 2: Produktdichten

Seien X_1, X_2, \dots iid unter P und unter Q mit Verteilung μ bzw. ν ,

$\nu \ll \mu$ mit Dichte f (z.B. $\Omega = \{(\omega_1, \omega_2, \dots) \mid \omega_i \in \mathbb{R}\}$, $X_i(\omega) = \omega_i$, $P = \bigotimes_{i=1}^{\infty} \mu$, $Q = \bigotimes_{i=1}^{\infty} \nu$)

$$\mathcal{F}_n := \sigma(X_1, \dots, X_n)$$

$$\Rightarrow Q \ll P \text{ auf } \mathcal{F}_n \text{ mit Dichte } \frac{dQ}{dP} \Big|_{\mathcal{F}_n} = \prod_{i=1}^n f(X_i)$$

$$\begin{aligned} \Gamma \text{ denn: } E_Q[g(X_1, \dots, X_n)] &= \int \int g(x_1, \dots, x_n) \nu(dx_1) \dots \nu(dx_n) \\ &= \int \int g(x_1, \dots, x_n) \prod_{i=1}^n f(x_i) \mu(dx_1) \dots \mu(dx_n) \\ &= E_P \left[g(X_1, \dots, X_n) \cdot \prod_{i=1}^n f(X_i) \right] \end{aligned}$$

Aber $Q \perp P$ auf \mathcal{F}_∞ falls $\mu \neq \nu$:

$$\sum_{i=1}^n \mathbb{I}_A(X_i) \rightarrow \begin{cases} \nu(A) & Q\text{-f.s.} \\ \mu(A) & P\text{-f.s.} \end{cases}$$

11.4. Von lokalen zu globalen Dichten

Sei $Q \ll P$ mit sukzessiven Dichten

$$X_n := \frac{dQ}{dP} \Big|_{\mathcal{F}_n}$$

Existiert Dichte auf $\mathcal{F}_\infty = \sigma(\cup \mathcal{F}_n)$?

Satz 1) X_n ist ein (\mathcal{F}_n) -Martingel bzgl. P .

2) $\exists X_\infty = \lim X_n$ P -f.s., $X_\infty \in \mathcal{L}'(\Omega, \mathcal{O}, P)$

3) X_n ist gleichm. integrierbar bzgl. $P \Leftrightarrow Q \ll P$ auf \mathcal{F}_∞
In diesem Fall ist X_∞ Dichte von Q bzgl. P .

Konsequenz: Satz von Radon-Nikodym auf separabler σ -Algebra:

$Q \ll P$ auf $\mathcal{F} = \sigma(\cup \mathcal{F}_n)$, \mathcal{F}_n atomar

$\Rightarrow Q \ll P$ auf $\mathcal{F}_n \quad \forall n \Rightarrow \exists \frac{dQ}{dP} \Big|_{\mathcal{F}_n} \quad \forall n$ (s.o.)

Satz
 $\Rightarrow \exists \frac{dQ}{dP} \Big|_{\mathcal{F}} = \lim_{n \rightarrow \infty} \frac{dQ}{dP} \Big|_{\mathcal{F}_n} \quad P$ -f.s.

Beweis Satz: 1) $X_n = E_P[X_{n+1} | \mathcal{F}_n]$:

- \mathcal{F}_n -meßbar

- $A \in \mathcal{F}_n \Rightarrow \int_A X_n dP = Q[A] = \int_A X_{n+1} dP$

2) $X_n \geq 0 \Rightarrow \exists X_\infty = \lim X_n$ P -f.s. nach Martingalkonvergenzatz.

3) a) X_n glm. intbar bzgl. $P \Rightarrow X_n = E_P[X_\infty | \mathcal{F}_n]$

$\Rightarrow Q[A] = \int_A X_n dP = \int_A X_\infty dP \quad \forall A \in \mathcal{F}_n \quad \forall n$

Eindeigkeitsatz

\Rightarrow

$$Q[A] = \int_A X_\infty dP \quad \forall A \in \sigma(\cup F_n) = \mathcal{F}_\infty \quad (120)$$

b) $Q \ll P$ auf $\mathcal{F}_\infty \Rightarrow X_n$ gleichm. intbar bzgl. P : $c > 0$

$$\int_{\{|X_n| \geq c\}} |X_n| dP \stackrel{X_n \geq 0}{=} \int_{\{X_n \geq c\}} X_n dP = Q[X_n \geq c] \leq Q[\sup X_n \geq c] \xrightarrow[c \uparrow \infty]{?} 0$$

In der Tat:

$$X_n \text{ konv. } P\text{-f.s. in } \mathbb{R} \Rightarrow \sup X_n < \infty \text{ } P\text{-f.s.} \stackrel{Q \ll P}{\Rightarrow} \sup X_n < \infty \text{ } Q\text{-f.s.}$$

□

11.5. Absolutstetige Funktionen

Sei $F: [0,1] \rightarrow \mathbb{R}$ monoton wachsend, nicht konstant.

OBdA $F(0)=0, F(1)=1$ (sonst lineare Transformation).

Ableitung von F ?

$$\mathcal{F} := \mathcal{B}([0,1]), \quad \mathcal{F}_n := \sigma\left(\left[\frac{k}{2^n}, \frac{k+1}{2^n}\right] \mid 0 \leq k < 2^n\right).$$

$P := \lambda_{[0,1]}$, Q sei WV auf $[0,1]$ mit Verteilungsfunktion $F(x)$.

$$\left. \frac{dQ}{dP} \right|_{\mathcal{F}_n}(x) = \frac{Q\left[\left(\frac{k}{2^n}, \frac{k+1}{2^n}\right]\right]}{P\left[\left(\frac{k}{2^n}, \frac{k+1}{2^n}\right]\right)} = \frac{F\left(\frac{k+1}{2^n}\right) - F\left(\frac{k}{2^n}\right)}{2^{-n}} \quad \text{falls } x \in \left(\frac{k}{2^n}, \frac{k+1}{2^n}\right]$$

dyadischer Differenzenquotient

Korollar (Satz von Lebesgue)

1) F monoton \Rightarrow Für f.a. $x \in [0,1]$ existiert die dyadische

$$\text{Ableitung } F'(x) = \lim_{n \rightarrow \infty} \left. \frac{dQ}{dP} \right|_{\mathcal{F}_n}(x).$$

2) Ist F absolutstetig (d.h. $Q \ll \lambda_{[0,1]}$), dann folgt

$$Q[A] = \int_A F'(x) dx \quad \forall A \in \mathcal{B}([0,1]),$$

$$\text{also insbesondere } F(y) = \int_0^y F'(x) dx.$$

11.6. Lebesgue Zerlegung

Satz Seien P, Q Wahrscheinlichkeitsvorstellungen auf (Ω, \mathcal{O}) . Dann existiert eine Zufallsvariable $X: \Omega \rightarrow [0, \infty]$ mit $X \in \mathcal{L}^1(\Omega, \mathcal{O}, P)$, sodass

$$\begin{aligned} Q[A] &= \int_A X dP + Q[A \cap \{X = \infty\}] \\ &=: Q_a[A] + Q_s[A] \end{aligned}$$

Q_a : absolutstetiger Anteil von Q (bzgl. P)

Q_s : singulärer Anteil von Q (bzgl. P)

Notation $X = \frac{dQ}{dP} \Big|_{\mathcal{O}}$ Lebesguedichte

Also: $Q \ll P$ auf $\{X < \infty\}$ mit Dichte X

$Q \perp P$ auf $\{X = \infty\}$

Beweis: ~~Existenz~~ Existenz der Lebesguedichte s. 11.7

Ben. Maßtheoret. Induktion $\Rightarrow \int Y dQ = \int YX dP + \int Y dQ \Big|_{\{X = \infty\}} \forall Y \geq 0$ u.b.

Lemma X Lebesguedichte von Q bzgl. P

$\Rightarrow \frac{1}{X}$ Lebesguedichte von P bzgl. Q ($\frac{1}{0} := +\infty, \frac{1}{\infty} = 0$)

Beweis: $Q[X=0] = 0$

$$\int_A \frac{1}{X} dQ + P[A \cap \{X = \infty\}] = \int_A \frac{1}{X} dQ + P[A \cap \{X = 0\}] = P[A]$$

$\underbrace{\int_A \frac{1}{X} dQ}_{\substack{\text{Ben.} \\ \equiv \int_A \frac{1}{X} X dP \\ A \cap \{X > 0\}}}$

□

Eindeutigkeit von X : ist sowohl P -f.s. als auch Q -f.s. eind. festgelegt:

$$P\text{-f.s.}: P[X < \infty] = 1 \text{ und } \int_A X dP = Q[A] \quad \forall A \subset \{X < \infty\}$$

$$Q\text{-f.s.}: \frac{1}{X} = \frac{dP}{dQ} \Rightarrow \frac{1}{X} \text{ } Q\text{-f.s. eindeutig} \Rightarrow X \text{ } Q\text{-f.s. eindeutig}$$

11.7. Existenz der Lebesgedichte

a) Elementarer Fall $\mathcal{G} = \sigma(B_i; i \in \mathbb{N})$, $\mathcal{R} = \cup B_i$

$$\frac{dQ}{dP}(w) := \frac{Q[B_i]}{P[B_i]} \quad \text{für } w \in B_i$$

$$:= \begin{cases} +\infty & \text{falls } P[B_i] = 0, Q[B_i] \neq 0 \\ \text{irgendeine Konstante} & \text{falls } P[B_i] = Q[B_i] = 0 \end{cases}$$

b) Separabler Fall $\mathcal{G} = \mathcal{F}_\infty = \sigma(\cup \mathcal{F}_n)$, \mathcal{F}_n atomar.

$$X_n = \frac{dQ}{dP} \Big|_{\mathcal{F}_n}$$

Satz 1) X_n ist ein Supermartingal bzgl. P ,
 $\frac{1}{X_n}$ ist ein Supermartingal bzgl. Q .

insbesondere $\exists X_\infty = \lim X_n \begin{cases} \in [0, \infty) & P\text{-f.s.} \\ \in (0, \infty] & Q\text{-f.s.} \end{cases}$

$$X_\infty \in \mathcal{L}^1(P), \frac{1}{X_\infty} \in \mathcal{L}^1(Q).$$

2) X_∞ ist Lebesguedichte von Q bzgl. P auf \mathcal{M} :

$$Q[A] = \int_A X_\infty dP + Q[A \cap \{X_\infty = \infty\}]$$

$\frac{1}{X_\infty}$ ist Lebesguedichte von P bzgl. Q auf \mathcal{M} .

Bemerkung: X_n Martingal $\Leftrightarrow Q \ll_{loc} P$

Beweis: 1) $Q[\{X_n = \infty\} \cap \{X_{n+1} = \infty\}^c] = 0$

d.h. $\{X_n = \infty\} \subseteq \{X_{n+1} = \infty\}$ Q -f.s. ^u

denn: $Q[\underbrace{X_n = \infty}_{\in \mathcal{F}_n \subseteq \mathcal{F}_{n+1}}] \stackrel{\text{Leb. rel. auf } \mathcal{F}_{n+1}}{=} \int_{\underbrace{\{X_n = \infty\}}_{\in \mathcal{F}_n}} X_{n+1} dP + Q[\{X_n = \infty\} \cap \{X_{n+1} = \infty\}]$
 $= 0$ da $X_n < \infty$ P -f.s.

2) X_n ist Supermartingal bzgl. P : Sei $A \in \mathcal{F}_n$.

$$\int_A X_{n+1} dP = Q[A] - Q[A \cap \{X_{n+1} = \infty\}]$$

$$\stackrel{1)}{\leq} Q[A] - Q[A \cap \{X_n = \infty\}] = \int_A X_n dP$$

Analogy: $\frac{1}{X_n}$ ist Supermartingal bzgl. Q .

3) Martingalkonvergenzsatz:

$$X_n \rightarrow X_\infty \in [0, \infty) \text{ P-f.s.}$$

$$\frac{1}{X_n} \rightarrow \frac{1}{X_\infty} \in [0, \infty) \text{ Q-f.s.}$$

$$(\text{also } X_n \rightarrow X_\infty \in (0, \infty] \text{ Q-f.s.}).$$

$$4) Q_a[A] := \int_A X_\infty dP, \quad Q_s[A] := Q[A] - Q_a[A]$$

Sind positive Maße auf \mathcal{F}_∞ .

Positivität von Q_s : Sei $A \in \mathcal{F}_n$.

$$Q_a[A] = \int_A X_\infty dP \stackrel{\text{Fatou}}{\leq} \liminf_{k \rightarrow \infty} \int_A X_k dP \leq \int_A X_n dP \stackrel{\text{Leb. vert. auf } \mathcal{F}_n}{\leq} Q[A].$$

Also $Q_s[A] \geq 0 \quad \forall A \in \cup \mathcal{F}_n$, also auch $\forall A \in \mathcal{F}_\infty$.

$$5) \text{ wach zu zeigen: } Q_s[A] = Q[A \cap \{X_\infty = \infty\}]$$

Dies folgt aus:

$$\text{Beh. } Q[X_\infty < \infty] \subseteq Q_a[X_\infty < \infty],$$

denn dann $Q[X_\infty = \infty] \stackrel{(*)}{=} Q_a[X_\infty = \infty]$, $Q_s[X_\infty = \infty] = 0$, also

$$Q[A \cap \{X_\infty = \infty\}] = Q_s[A \cap \{X_\infty = \infty\}] = Q_s[A].$$

Beweis der Beh.:

$$Q[X_\infty < \infty] = Q[\sup X_n < \infty] = \sup_{c > 0} Q[\sup X_n \leq c]$$

$$Q[\sup X_n \leq c] \subseteq Q[\bigcap_n \{X_n \leq c\}] \subseteq \inf Q[X_n \leq c]$$

$$\leq \overline{\lim}_{\{X_n \leq c\}} \int X_n dP \stackrel{\text{Fatou}}{\leq} \int \overline{\lim} (X_n \cdot I_{\{X_n \leq c\}}) dP$$

$$\leq \int X_\infty dP = Q_a[\mathbb{R}]$$

$$\Rightarrow Q[X_\infty < \infty] \subseteq Q_a[\mathbb{R}] = Q_a[X_\infty < \infty] \quad \square$$

12. Absolutstetigkeit von unendlichen Produktmaßen

Seien $Y_1, Y_2, \dots: \Omega \rightarrow S$ unabh. unter P sowie unter Q
mit Verteilungen $\mu_i = P \circ Y_i^{-1}$ bzw. $\nu_i = Q \circ Y_i^{-1}$,

$$\nu_i \ll \mu_i \quad \text{mit Dichte } f_i.$$

$$\text{z.B.: } \Omega = S^{\mathbb{N}} = \{(\omega_1, \omega_2, \dots) \mid \omega_i \in S\}$$

$$Y_i(\omega) = \omega_i$$

$$P = \bigotimes_{i=1}^{\infty} \mu_i, \quad Q = \bigotimes_{i=1}^{\infty} \nu_i$$

$$\mathcal{F}_n := \sigma(Y_1, \dots, Y_n), \quad \mathcal{F} = \sigma(Y_1, Y_2, \dots)$$

Lokale Radon-Nikodym Dichte von Q bzgl. P :

$$\frac{dQ}{dP} \Big|_{\mathcal{F}_n} = \prod_{i=1}^n f(Y_i) =: X_n$$

Globale Lebesguedichte:

$$\frac{dQ}{dP} \Big|_{\mathcal{F}} = X_{\infty} = \lim X_n \in \begin{cases} [0, \infty) & P\text{-f.s.} \\ (0, \infty] & Q\text{-f.s.} \end{cases}$$

Lebesguezerlegung: $Q[A] = Q_a[A] + Q_s[A]$

(131)

$$Q_a[A] = \int_A X_\infty dP, \quad Q_s[A] = Q[A \cap \{X_\infty = \infty\}]$$

12.1. Asymptotik der Dichte im i.i.d. Fall

- Satz von Shannon-McMillan

$$\mu_i \equiv \mu, \quad \nu_i \equiv \nu, \quad \mu \approx \nu, \quad \mu \neq \nu$$

Dann ist $Q \perp P$ nach Gesetz der großen Zahlen:

$$\frac{1}{n} \sum_{i=1}^n \mathbb{I}_A(Y_i) \rightarrow \begin{cases} \mu(A) & P\text{-f.s.} \\ \nu(A) & Q\text{-f.s.} \end{cases}$$

Also $Q = Q_s, Q_a = 0 \Rightarrow X_\infty = \infty$ Q -f.s., $X_\infty = 0$ P -f.s.

Genauere Asymptotik von X_n :

(i) Q -fast sicher: $\quad \quad \quad =: H(\nu|\mu)$

$$\frac{1}{n} \log X_n = \frac{1}{n} \sum_{i=1}^n \log \frac{d\nu}{d\mu}(Y_i) \xrightarrow[Q\text{-f.s.}]{n \rightarrow \infty} \int \log \frac{d\nu}{d\mu} d\nu = \int \frac{d\nu}{d\mu} \log \frac{d\nu}{d\mu} d\mu$$

„Satz von Shannon-McMillan“

(ii) P-fast sicher:

$$\frac{1}{n} \log X_n = -\frac{1}{n} \log \frac{1}{X_n} = -\frac{1}{n} \log \left. \frac{dP}{dQ} \right|_{\mathcal{F}_n} \xrightarrow[n \uparrow \infty]{P\text{-f.s.}} -H(\mu|v)$$

$H(v|\mu)$ =: relative Entropie von v bzgl. μ

$$H(v|\mu) \underset{\text{ Jensen }}{>} \int \frac{d\nu}{d\mu} d\mu \log \underbrace{\int \frac{d\nu}{d\mu} d\mu}_{=1} = 0$$

da $x \log x$ strikt konvex und $\frac{d\nu}{d\mu}$ nicht konstant ($v \neq \mu$).

Also: $X_n \sim e^{n H(v|\mu)}$ Q-f.s. (exp. Wachstum)
 $X_n \sim e^{-n H(\mu|v)}$ P-f.s. (exp. Abfall)

Bem. Das Gesetz der großen Zahlen ist im Beweis des Satzes von Shannon-McMillan anwendbar, denn:

$$x \log x \geq x - 1 \geq -1 \quad \forall x \geq 0 \Rightarrow \log \frac{d\nu}{d\mu} \geq -\frac{1}{\frac{d\nu}{d\mu}} = -\frac{d\mu}{d\nu}$$

$$\Rightarrow \int \left(\log \frac{d\nu}{d\mu} \right)^- d\nu \leq \int \frac{d\mu}{d\nu} d\nu = 1$$

12.2. Anwendung auf Testen von Hypothesen

Angenommen, wir wissen nicht, ob die Y_i unabhängige Stichproben von der Verteilung μ (Hypothese A) oder von der Verteilung ν (Hyp. B) sind.

$$X_n = \frac{dQ}{dP} \Big|_{\mathcal{F}_n} = \prod_{i=1}^n \frac{d\nu}{d\mu}(Y_i) = \frac{\text{"W'heit der Stichprobenfolge } (Y_1, \dots, Y_n) \text{ unter } \nu \text{"}}{\text{"W'heit der Stichprobenfolge } (Y_1, \dots, Y_n) \text{ unter } \mu \text{"}}$$

Likelihood-Quotient

Neyman-Pearson-Test (Spezialfall eines Likelihood-Quotienten-Tests):

Sei $c \in (0, \infty)$. Entscheide für A falls $X_n \geq c$, für B falls $X_n < c$.

12.1 \Rightarrow Test ist konsistent, d.h.

$$\mathbb{I}_{\{X_n \geq c\}} \rightarrow \begin{cases} 1 & Q\text{-f.s.} \\ 0 & P\text{-f.s.} \end{cases}$$

$H(\nu|\mu)$ bestimmt, wie schnell der "Testfehler" gegen 0 geht

relative Entropie = Maß für "statistische Unterscheidbarkeit" der W_n v. ν

12.3. Kakutani's Dichotomie

Y_i i.i.d. unter P und Q mit μ_i bzw. ν_i , $\nu_i \ll \mu_i$ mit Dichte f_i .

$$\mathcal{F} = \sigma(Y_1, Y_2, \dots)$$

Kor. Entweder $Q \ll P$ oder $Q \perp P$ auf \mathcal{F} .

$$Q \ll P \Leftrightarrow \prod_{i=1}^{\infty} f_i(Y_i) < \infty \text{ } Q\text{-f.s.} \Leftrightarrow \sum_{i=1}^{\infty} \log f_i(Y_i) < \infty \text{ } Q\text{-f.s.}$$

$$Q \perp P \Leftrightarrow \text{ " } = \infty \text{ " } \Leftrightarrow \text{ " } = \infty \text{ "}$$

Beweis: $\prod_{i=1}^n f_i(Y_i) = X_n \rightarrow X_{\infty} \in (0, \infty] \text{ } Q\text{-f.s.}$, s. 11.7.

$\log \sum_{i=1}^n f_i(Y_i) \rightarrow \log X_{\infty} \in (-\infty, \infty] \text{ } Q\text{-f.s.}$

Kolmogorov's 0-1-Gesetz: $Q[X_{\infty} = \infty] \in \{0, 1\} \Rightarrow Q \ll P$ bzw. $Q \perp P$. □

Rem. Sei $\nu_i \approx \mu_i$, d.h. $f_i > 0 \text{ } \mu_i\text{-f.s.}$. Dann folgt:

$$Q \ll P \Leftrightarrow Q \not\ll P \Leftrightarrow P \not\ll Q \Leftrightarrow P \ll Q$$

Also entweder $Q \perp P$ oder $Q \approx P$ ▽

12.4. Beispiel: Unendlichdimensionale Normalverteilungen

$$\mu_i = N(0, 1), \nu_i = N(a_i, 1) \quad (a_i > 0)$$

$$f_i(x) = \frac{dN(a_i, 1)}{dN(0, 1)} = e^{a_i x - \frac{1}{2} a_i^2}$$

$$X_n = \frac{dQ}{dP} \Big|_{\mathcal{F}_n} = \prod_{i=1}^n f_i(Y_i) = \exp\left(\sum_{i=1}^n a_i Y_i - \frac{1}{2} \sum_{i=1}^n a_i^2\right)$$

Satz 1) $\sum a_i^2 < \infty \Rightarrow Q \ll P$ auf \mathcal{F}

2) $\sum a_i^2 = \infty \Rightarrow Q \perp P$ auf \mathcal{F}

Bem. z.B. $Q = \bigotimes_{i=1}^{\infty} N(a_i, 1) = N(a, I), \quad a = (a_1, a_2, \dots), \quad I = \begin{pmatrix} 1 & & 0 \\ & \ddots & \\ 0 & & \ddots \end{pmatrix}$

$$P = \bigotimes_{i=1}^{\infty} N(0, 1) = N(0, I)$$

unendlichdimensionale Normalverteilungen

$$Q \ll P \Leftrightarrow Q \approx P \Leftrightarrow a \in \ell^2$$

Beweis Satz: $M_n := \sum_{i=1}^n a_i \underbrace{(Y_i - a_i)}_{\sim N(0, a_i^2) \text{ bzgl. } Q, \text{ unabh.}}$ ist Martingal bzgl. Q

1) $\sum a_i^2 < \infty \Rightarrow E[M_n^2] = \sum_{i=1}^n a_i^2$ beschr. $\xrightarrow{L^2\text{-Konvergenz}} M_n \rightarrow$ all. limes Q -f.s.

$$\Rightarrow X_n = e^{M_n + \frac{1}{2} \sum_{i=1}^n a_i^2} \rightarrow \text{endl. Lines Q-f.s.}$$

$$\Rightarrow Q[X_\infty < \infty] = 1 \Rightarrow Q \ll P \stackrel{\text{Dichtewirk}}{\Rightarrow} Q \approx P$$

$$2) M_n \sim N(0, \underbrace{\sum_{i=1}^n a_i^2}_{\rightarrow \infty}) \text{ bzgl } Q$$

$$\Rightarrow Q[M_n > c \text{ oft}] = \lim Q[M_n > c] = \lim Q\left[\frac{M_n}{\sqrt{\sum_{i=1}^n a_i^2}} > \frac{c}{\sqrt{\sum_{i=1}^n a_i^2}}\right] = \frac{1}{2}$$

für alle $c > 0$

$$\rightarrow Q[\overline{\lim} M_n = \infty] \geq \frac{1}{2} \stackrel{0-1}{\Rightarrow} = 1 \Rightarrow X_\infty = \infty \text{ Q-f.s.} \quad \square$$

Allgemeiner gilt:

12.5. Absolutstetigkeitskriterium von Kakutani (Situation wie in 12.3)

Satz $Q \perp P$ falls $\prod_{i=1}^{\infty} \int \sqrt{\frac{d\mu_i}{d\nu_i}} d\nu_i = 0$

$Q \ll P$ falls $\prod_{i=1}^{\infty} \int \sqrt{\frac{d\mu_i}{d\nu_i}} d\nu_i > 0$

Beweis: s. Übung.

Bem. Sei $\nu_i \approx \mu_i$. Dann gilt:

$$\int \sqrt{\frac{d\mu_i}{d\nu_i}} d\nu_i = \int \sqrt{\frac{d\nu_i}{d\mu_i}} d\mu_i \quad \left(= \int \sqrt{d\nu_i d\mu_i} \right)$$

13. Martingale bzgl. absteigender σ -Algebren

$(\mathbb{R}, \mathcal{O}, P)$ W'raum

$\dots \subseteq \mathcal{F}_{-2} \subseteq \mathcal{F}_{-1} \subseteq \mathcal{F}_0 \subseteq \mathcal{O}$ absteigende Filtration

M_n ($n \leq 0$) Martingal bzgl. $(\mathcal{F}_n)_{n \leq 0}$

$\Rightarrow M_n = E[M_0 | \mathcal{F}_n] \quad \forall n \leq 0$

$\Rightarrow \{M_n | n \leq 0\}$ gleichm. int. bar

$\mathcal{F}_{-\infty} := \bigcap_{n \leq 0} \mathcal{F}_n$

13.1. Rückwärtsmartingalkonvergenzsatz

$M_n \xrightarrow{n \downarrow -\infty} M_{-\infty} = E[M_0 | \mathcal{F}_{-\infty}]$ \mathcal{F} -f.s. und in L^1

Beweis: via Doob's Upcrossing Ungleichung - Übung.

13.2. Martingalbeweis von Kolmogorovs Gesetz der großen Zahlen

$$Y_1, Y_2, \dots \text{ i.i.d.}, \in \mathcal{L}^1, S_n = Y_1 + \dots + Y_n$$

Beh. $\frac{S_n}{n} \rightarrow E[Y_1]$ P-f.s.

$$\text{Beweis: } \frac{S_n}{n} = E[Y_1 | \mathcal{F}_n] = E[Y_1 | S_n, Y_{n+1}, Y_{n+2}, \dots] = E[Y_1 | \underbrace{\sigma(S_n, S_{n+1}, \dots)}_{=: \mathcal{F}_n}]$$

$$\left[\begin{array}{l} \text{denn: } Z \text{ } \sigma(S_n)\text{-meßbar} \Rightarrow Z = f(S_n) \Rightarrow E[Y_1, Z] \stackrel{\text{Symmetrie}}{=} E[Y_{i_2}, Z] \forall i_2 \\ \Rightarrow E[Y_1, Z] = \frac{1}{n} \sum_{k=1}^n E[Y_k, Z] = E\left[\frac{S_n}{n}, Z\right]. \end{array} \right.$$

→ Interpretation von $\frac{S_n}{n}$ als Rückwärtsmartingal

$$13.1 \Rightarrow \frac{S_n}{n} \rightarrow E[Y_1 | \mathcal{F}_\infty] \text{ P-f.s.}$$

Identifikation des Limes via 0-1-Gesetz:

$$\overline{\lim} \frac{S_n}{n} \text{ ist } \mathcal{F}_\infty\text{-meßbar} \Rightarrow \lim \frac{S_n}{n} = \text{const. P-f.s.}$$

$$E\left[\lim \frac{S_n}{n}\right] \stackrel{\text{gl. int.}}{=} \lim E\left[\frac{S_n}{n}\right] = E[Y_1] \Rightarrow \text{const.} = E[Y_1].$$

□

14. Maximalungleichungen

(139)

= Abschätzungen für $\max(X_0, X_1, \dots, X_n)$ ($n \in \mathbb{N}$),
wobei X_n Sub- bzw. Supermartingal

Konsequenz des Stoppsatzes

14.1. Doob's Ungleichung

Satz 1) X_n ($n=0,1,2,\dots$) Supermartingal, ≥ 0

$$\Rightarrow P\left[\sup_k X_k \geq c\right] \leq \frac{E[X_0]}{c} \quad \forall c > 0$$

2) X_n Submartingal, ≥ 0

$$\Rightarrow P\left[\max_{0 \leq k \leq n} X_k \geq c\right] \leq \frac{1}{c} \cdot E[X_n; \max_{0 \leq k \leq n} X_k \geq c] \leq \frac{E[X_n]}{c}$$

(Doob's Submartingalungleichung)

Beweis: 1) $c \cdot P[\sup X_k > c] \leq c \cdot P[T_c < \infty]$, $T_c := \min\{k \geq 0 \mid X_k \geq c\}$

$$= E[c \cdot \mathbb{I}_{\{T_c < \infty\}}] \leq E[X_{T_c}; T_c < \infty] \leq E[X_0]$$

↑ Stoppsatz für positive Supermart.

$$\Rightarrow P[\sup X_k \geq c] = \lim_{\varepsilon \downarrow 0} \underbrace{P[\sup X_k > c - \varepsilon]}_{\leq E[X_0]/(c-\varepsilon)} \leq \frac{E[X_0]}{c}$$

$$\begin{aligned}
 2) \quad c \cdot P[\max_{0 \leq k \leq n} X_k \geq c] &= c \cdot P[T_c \leq n] \leq E[X_{T_c}; T_c \leq n] \\
 &= \sum_{k=0}^n \underbrace{E[X_k; T_c = k]}_{\in \mathcal{F}_k} \leq E[X_n; T_c \leq n] \\
 &\leq E[X_n; T_c = k] \quad \square
 \end{aligned}$$

Anwendungen: Große Abweichungen (Übung),
Satz vom iterierten Logarithmus (s.u.)

Kor. 1 X_n Submartingal (nicht notwendig ≥ 0)

$$\Rightarrow P[\max_{k \leq n} X_k \geq c] \leq \frac{1}{c} E[X_n^+; \max_{k \leq n} X_k \geq c] \quad \forall c > 0$$

Beweis: Jede Doobungleichung auf X_n^+ an
(ist wieder Submartingal, da $E[X_{n+1}^+ | \mathcal{F}_n] \geq E[X_n^+ | \mathcal{F}_n] \geq X_n^+$). \square
($x \mapsto x^+$ konv.)

Bem. Erweiterung der Markovungleichung.

Kor. 2 M_n Martingal

$$\begin{aligned}
 \Rightarrow 1) \quad P[\max_{k \leq n} |M_k| \geq c] &\leq \frac{1}{c} E[|M_n|; \max_{k \leq n} |M_k| \geq c] \\
 2) \quad P[\max_{k \leq n} |M_k| \geq c] &\leq \frac{1}{c^2} E[M_n^2] \quad (\text{Kolmogorovungl.})
 \end{aligned}$$

Beweis: Doob für $|M_n|$ bzw. M_n^2 \square Anwendung von 2): Gesetz der großen Zahlen, ...

14.2. Doob's L^p -Ungleichung, L^p -Martingalb. verzerzte

(141)

Satz Sei X_n Martingal, $X_n^* := \max_{k \leq n} |X_k|$, $X^* := \sup |X_k|$

1) Für $p > 1$ gilt:

$$\|X_n^*\|_p \leq q \cdot \|X_n^+\|_p, \quad \frac{1}{p} + \frac{1}{q} = 1$$

$$\|X^*\|_p \leq q \cdot \sup \|X_n^+\|_p$$

Also: X_n beschränkt in $\mathcal{L}^p \Rightarrow \sup |X_n| \in \mathcal{L}^p$

$$2) \|X_n^*\|_1 \leq \frac{e}{e-1} (1 + E[|X_n| \log^+ |X_n|])$$

$$\|X^*\|_1 \leq \frac{e}{e-1} (1 + \sup_n E[|X_n| \log^+ |X_n|])$$

Korollar X_n Martingal, beschr. in \mathcal{L}^p , $p > 1 \Rightarrow X_n \rightarrow X_\infty$ in \mathcal{L}^p

Beweis: X_n beschr. in $\mathcal{L}^p \Rightarrow$ beschr. in $\mathcal{L}^1 \Rightarrow X_n \rightarrow X_\infty$ P-f.s.

$$E[|X_n - X_\infty|^p] \xrightarrow{\text{Lebesgue}} 0$$

$$\leq 2 \cdot \sup |X_n|^p \text{ integrierbar \& Majorante, da } \sup |X_n| \in \mathcal{L}^p$$

□

Beweis des Satzes basiert auf:

Lemma Sei $Y \geq 0$ ZV, $G(y) = \int_0^y g(x) dx$, $g \geq 0$. Dann:

$$E[G(Y)] = \int_0^\infty g(c) P[Y \geq c] dc$$

Beweis: $E[G(Y)] = E[\int_0^Y g(c) dc] = E[\int_0^\infty I_{[0, Y]}(c) g(c) dc]$

Fubini $= \int_0^\infty g(c) P[Y \geq c] dc. \square$

Beweis Satz: 1) $E[(X_n^*)^p] = \int_0^\infty p c^{p-1} \overbrace{P[X_n^* \geq c]}^{\leq \frac{1}{c} E[X_n^+; X_n^* \geq c] \text{ nach Max. wgl.}}$ dc

Fubini $\leq E[X_n^+ \int_0^{X_n^*} p c^{p-2} dc] = \frac{p}{p-1} E[X_n^+ \cdot (X_n^*)^{p-1}]$

Hölder $\leq \frac{p}{p-1} \cdot \|X_n^+\|_p \cdot \underbrace{\|(X_n^*)^{p-1}\|_q}_q$ $\frac{1}{p} + \frac{1}{q} = 1, \text{ d.h. } q = \frac{p}{p-1}$
 $= E[(X_n^*)^p]^{\frac{1}{q}}$

$$\Rightarrow \underbrace{E[(X_n^*)^p]^{\frac{1}{p}}}_{= \|X_n^*\|_p} \leq q \cdot \|X_n^+\|_p \quad \forall n \geq 0$$

$$\Rightarrow \|X^*\|_p \stackrel{\text{Fatou}}{\leq} \liminf \|X_n^*\|_p \leq q \cdot \sup \|X_n^+\|_p$$

2) Übung

□

III. Entropie und große Abweichungen

Lit. Ellis: Entropy, large deviations and statistical mechanics

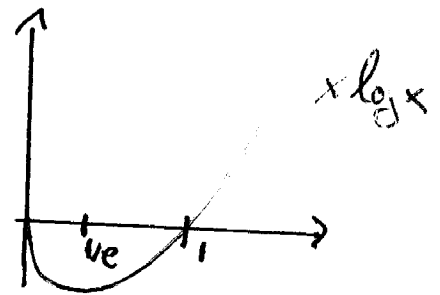
Duges, Rost: Prinzipien der Stochastik (Entropie ↔ Physik, Informationstheorie)

Georgii: Stochastik, Teil II (Entropie ↔ Statistik)

Dembo, Zeitouni: Large deviation techniques and applications
(Standardwerk über große Abweichungen)

15. Absolute und relative Entropie

Vorbem.: $u(x) := x \log x \geq x - 1$, $u(0) := 0$,
streikt konvex, Minimum $u(\frac{1}{e}) = -\frac{1}{e}$

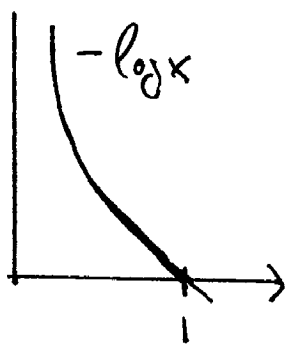


15.1. Absolute Entropie

Salzählbar, μ WV auf S , $\mu(x) \in [0, 1]$

Def. $H(\mu) := - \sum_{\substack{x \in S \\ \mu(x) > 0}} \mu(x) \log \mu(x) \in [0, \infty]$ Entropie von μ

Interpretation: $-\log p(x) =$ Maß für "Überraschung", "Informationsgewinn"
falls $x \in S$ eintritt



$H(\mu) =$ mittlere Überraschung

Extremfälle: 1) minimale Entropie:

$$H(\mu) = 0 \Leftrightarrow p(x) \in \{0, 1\} \forall x \in S \Leftrightarrow \mu \text{ Diracmaß "deterministisch"}$$

2) maximale Entropie: S endlich

$$H(\text{Gleichverteilung}) = -\log\left(\frac{1}{|S|}\right) = \log |S| \geq H(\mu) \quad \forall \text{ WV } \mu \text{ auf } S, \text{ s.u.}$$

Diracmaß \leftrightarrow geordneter Zustand \leftrightarrow niedrige Entropie

Gleichverteilung \leftrightarrow ungeordneter Zustand \leftrightarrow hohe Entropie

Bem. Entropie von absolutstetigen Verteilungen auf \mathbb{R} :

$$H(\mu) = -\int f(x) \log f(x) dx, \quad f = \frac{d\mu}{dx}$$

nicht notwendig ≥ 0 , da $f \leq 1$ möglich ∇

15.2. Entropie von Produktmaßen

Satz $H(\mu \otimes \nu) = H(\mu) + H(\nu)$

Beweis: $\mu \otimes \nu(x, y) = \mu(x) \cdot \nu(y)$

$$\Rightarrow H(\mu \otimes \nu) = - \sum_{x, y} \mu(x) \nu(y) (\log \mu(x) + \log \nu(y))$$

$$= - \sum_x \mu(x) \log \mu(x) - \sum_y \nu(y) \log \nu(y) \quad \square$$

Bem. Wegen der Faktorisierungseigenschaft verwendet man in der Definition der Entropie den Logarithmus!

15.3. Relative Entropie

Seien $\mu \approx \nu$ Wk'n auf Ω (abzählbar, zunächst). Wenn ich μ als Modell annehme, aber tatsächlich ν die zugrundeliegende Wk ist, dann ist meine mittlere Überraschung $-\sum \nu(x) \log \mu(x)$ im Gegensatz zu $-\sum \mu(x) \log \nu(x)$ bei korrektem Modell.

Zuwachs an Überraschung durch falsches Modell:

$$H(\nu | \mu) = - \sum \nu(x) (\log \mu(x) - \log \nu(x)) = \sum \nu(x) \log \frac{\nu(x)}{\mu(x)} = \int \log \frac{d\nu}{d\mu} d\nu$$

Definition Seien μ, ν Wk_n auf meßb. Raum (S, \mathcal{F}) .

$$H(\nu | \mu) := \begin{cases} \int \log \frac{d\nu}{d\mu} d\nu = \int \frac{d\nu}{d\mu} \log \frac{d\nu}{d\mu} d\mu & \text{falls } \nu \ll \mu \\ +\infty & \text{sonst} \end{cases}$$

heißt relative Entropie (oder Kullback-Leibler-Information) von ν bezgl. μ .

(Nichtsymmetrisches) Maß für "statistische Unterscheidbarkeit" der Wk_n ν und μ , s. 12.1, 12.2.

Satz $H(\nu | \mu) \geq 0$, $= 0 \Leftrightarrow \nu = \mu$

Beweis. $\int \underbrace{f \log f}_{\geq f-1} d\mu \geq \int f d\mu - 1 = 0$

"=" $\Leftrightarrow f \log f = f - 1$ μ -f.s. $\Leftrightarrow f = 1$ μ -f.s. \square

Bem. 1) Zählbar $\Rightarrow H(\nu) = -\sum \nu(x) \log \underbrace{\nu(x)}_{\substack{\text{Dichte w.} \\ \nu \text{ bzgl. Zählmaß}}} = -H(\nu | \text{Zählmaß})$
 ≤ 0 da Zählmaß keine Wk, diese nicht anwendbar.

2) Analog: $S \subset \mathbb{R}^d$: $H(\nu) = -H(\nu | \lambda)$

15.4. Beispiele

1) Sendekod, $\mu = \text{Gleichverteilung} = \frac{\text{Zählmaß}}{|S|}$ Beachte Vorzeichen!

$\Rightarrow H(\nu | \mu) = \sum_{\nu(x)} \log \frac{\nu(x)}{|S|} = \log |S| - H(\nu) \quad \forall \text{WV } \nu$

Insbesondere: $H(\nu | \mu) \geq 0 \Rightarrow H(\nu) \leq \log |S| = H(\mu) \quad \forall \text{WV } \nu$

Gleichverteilung hat maximale Entropie!

2) $S = \{1, 2, 3, \dots\}$, μ geom. Verteilung mit Parameter $p \in (0, 1)$

$\xrightarrow{\text{Übung}} H(\nu) = H(\mu) - H(\nu | \mu) \leq H(\mu)$

für alle Wk ν auf S mit Mittel $m = \frac{1}{p}$.

Geometrische Verteilung maximiert Entropie bei vorgegebenem Erwartungswert

3) $H(N(m, \sigma^2) | N(\bar{m}, \bar{\sigma}^2)) = \log \frac{\bar{\sigma}}{\sigma} - \frac{1}{2} + \frac{1}{2\bar{\sigma}^2} (\sigma^2 + (\bar{m} - m)^2)$

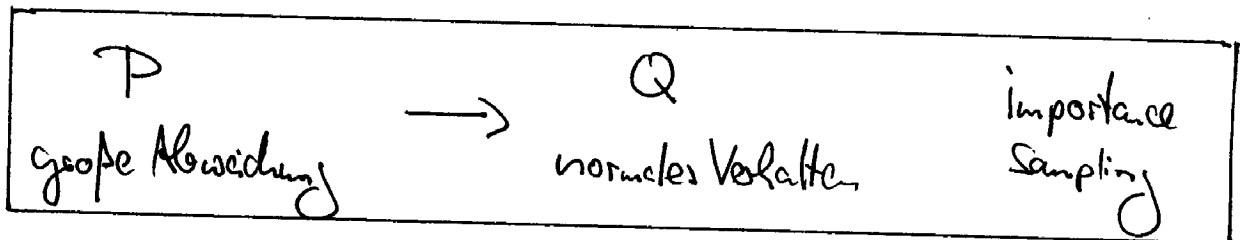
z.B. = $\begin{cases} \frac{(\bar{m} - m)^2}{2\sigma^2} & \text{falls } \sigma = \bar{\sigma} \\ \log \frac{\bar{\sigma}}{\sigma} - \frac{1}{2} + \frac{\sigma^2}{2\bar{\sigma}^2} & \text{falls } m = \bar{m} \end{cases}$

15.5. Faktorisierungseigenschaft $H(\nu_1 \otimes \dots \otimes \nu_n | \mu_1 \otimes \dots \otimes \mu_n) = \sum_{i=1}^n H(\nu_i | \mu_i)$ (Übung)

16. Importance Sampling

Ziel: Untere Schranke für die Wahrscheinlichkeit großer Abweichungen bzgl. einer Wahrscheinlichkeitsverteilung P .

Grundidee: Wähle WV $Q \stackrel{\text{loc}}{\approx} P$, so daß große Abweichung bzgl. Q normal ist
Anssage über Q -heit bzgl. P via Dichte $\frac{dQ}{dP}$.



$$\Omega = S^N = \{(\omega_1, \omega_2, \dots) \mid \omega_i \in S\}, \quad (S, \mathcal{F}) \text{ meßb. Raum.}$$

$$X_i(\omega) = \omega_i, \quad \mathcal{G} = \sigma(X_i, i \in \mathbb{N}), \quad \mathcal{F}_n = \sigma(X_1, \dots, X_n)$$

P, Q WVN auf (Ω, \mathcal{G}) mit $Q \stackrel{\text{loc}}{\approx} P$,

$$Z_n := \left. \frac{dQ}{dP} \right|_{\mathcal{F}_n} \in (0, \infty) \quad P\text{-f.s.}$$

Def. Eine Folge A_n ($n \geq 1$) von Mengen $A_n \in \mathcal{F}_n$ heißt

wesentlich für Q , falls $\liminf Q[A_n] > 0$

(bzw., je nach Strenge der Definition, falls $Q[A_n] \rightarrow 1$).

Frage: Wie klein kann man die Anwarte, gemessen durch P?

16.1. Beispiel: Monte Carlo Berechnung von kleinen Wahrscheinlichkeiten im Münzwurfmmodell

$S = \{0, 1\}$, $P = \bigotimes_{i=1}^{\infty} \mu$, $\mu(0) = \mu(1) = \frac{1}{2}$

$A_n := \left\{ \frac{1}{n} \sum_{i=1}^n X_i \geq m \right\}$, z.B. $m = 0,7$, $n = 1000$

$P[A_n] = ?$

Oberer Schranke nach Satz von Cramér:

$P[A_n] \leq e^{-n \Lambda^*(m)}$ wobei $\Lambda^*(m) = \sup_{t \in \mathbb{R}} (tm - \log E[e^{tX_1}])$
Größenordnung 10^{-28}
S.W.theorie 19.5
 $= m \log \frac{m}{1/2} + (1-m) \log \frac{1-m}{1/2} = H(p_m | \mu)$
wobei $p_m(1) = m, p_m(0) = 1-m$.

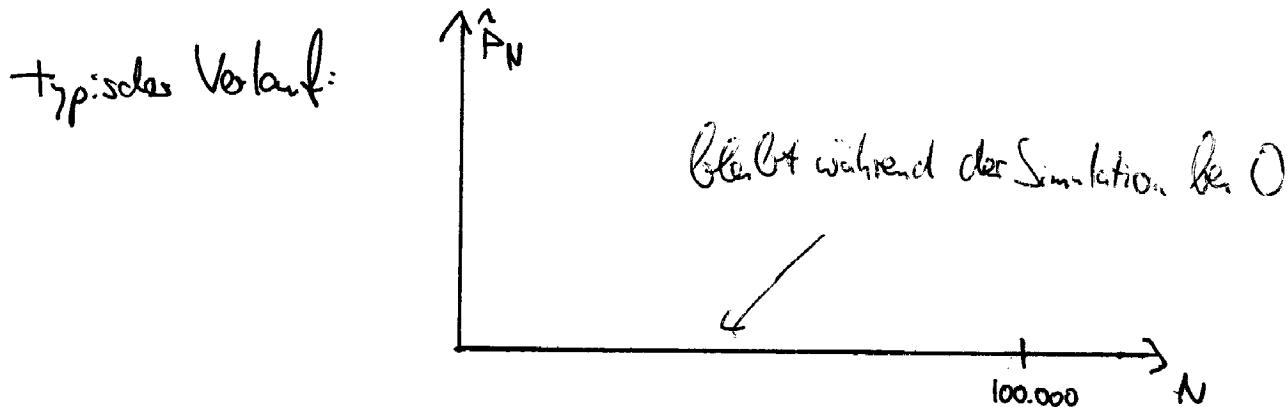
Numerische Berechnung von $P[A_n]$:

a) Klassisches Monte Carlo Verfahren

Seien $W^{(1)}, W^{(2)}, W^{(3)}, \dots$ ~~(i.i.d. Zufallsvektoren)~~ (Werte in Ω bzw. in S^n)
i.i.d. Münzwurffolgen mit Verteilung P. Dann:

$$\hat{P}_N := \frac{1}{N} \sum_{i=1}^N I_{A_n}(W^{(i)}) \xrightarrow{N \uparrow \infty} P[A_n] \quad \text{f.s.}$$

Problem: $P[A_n]$ ist so klein, daß $I_{A_n}(W^{(i)})$ praktisch immer 0 ist.



Bem. \hat{P}_N ist Schätzer für $P[A_n]$

- erwartungstreu (d.h. Erwartungswert $(\hat{P}_N) = P[A_n]$)
- konsistent (d.h. $\hat{P}_N \rightarrow P[A_n]$ f.s.)

ABER:

$$\text{var}(\hat{P}_N) = \frac{N}{N^2} \text{var}_P[I_{A_n}] = \frac{1}{N} (P[A_n] - P[A_n]^2) \approx \frac{P[A_n]}{N}$$

$$\text{Standardabweichung: } \sigma = \sqrt{\frac{P[A_n]}{N}} \gg P[A_n] \quad (\sim 10^{-38})$$

Der mittlere quadratische Schätzfehler ist viel größer als der zu prognostizierende Wert. Eine gute Prognose würde sich erst nach $N \gg \frac{1}{P[A_n]} (\sim 10^{38})$ Schritten ergeben!

ZIEL: Varianzreduktion

6) Monte Carlo mit importance sampling

Sei Q das Münzwurfmodell mit $p \geq 0,7$: $Q = \bigotimes_{i=1}^{\infty} p, 1-p$

$$Q[A_n] \xrightarrow{n \rightarrow \infty} \begin{cases} 1 & \text{für } p > 0,7 \text{ (Gesetz der großen Zahlen)} \\ \frac{1}{2} & \text{für } p = 0,7 \text{ (Zentraler Grenzwertsatz)} \end{cases}$$

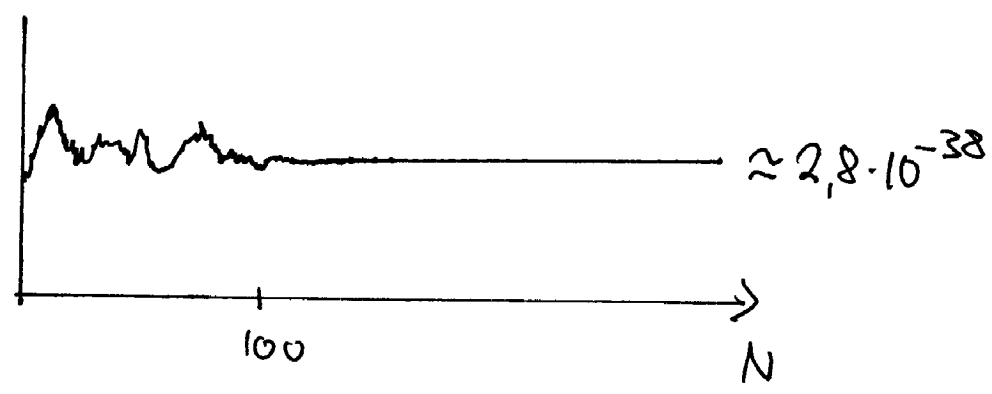
d.h. A_n ist wesentlich für Q .

$$Z_n = \frac{dQ}{dP} \Big|_{\mathcal{F}_n} = \frac{p^{\sum_{i=1}^n X_i} (1-p)^{n - \sum_{i=1}^n X_i}}{2^{-n}}$$

$$P[A_n] = \int_{A_n} \frac{1}{Z_n} dQ = \int \frac{I_{A_n}}{Z_n} dQ$$

Monte Carlo bzgl. Q : $\tilde{\omega}^{(1)}, \tilde{\omega}^{(2)}, \dots$ i.i.d. mit Verteilung Q

$$\Rightarrow \hat{I}_N = \frac{1}{N} \sum_{i=1}^N \frac{I_{A_n}(\tilde{\omega}^{(i)})}{Z_n(\tilde{\omega}^{(i)})} \xrightarrow{NT \rightarrow \infty} \int \frac{I_{A_n}}{Z_n} dQ = P[A_n] \text{ f.s.}$$



Auch \hat{I}_N ist erwartungstreu, konsistenter Schätzer.

Abschätzen der Varianz:

$$\text{var}(\hat{I}_N) = \frac{1}{N} \text{var}_Q \left(\frac{I_{A_n}}{Z_n} \right) \leq \frac{1}{N} \int_{A_n} \frac{1}{z_n^2} dQ \stackrel{?}{\leq}$$

Auf A_n gilt:

$$\begin{aligned} Z_n &= \left(\frac{p}{1-p} \right)^{\sum_{i=1}^n X_i} \cdot \left(\frac{1-p}{1/2} \right)^n \geq \left(\frac{p}{1/2} \right)^{0.7n} \cdot \left(\frac{1-p}{1/2} \right)^{0.3n} \\ &= e^{n \cdot \underbrace{(0.7 \log \frac{p}{1/2} + 0.3 \log \frac{1-p}{1/2})}_{\text{natural für } p=0.7}} = e^{n H(p_{0.7}|p)} \end{aligned}$$

Varianz soll möglichst klein sein \rightarrow wähle $p=0.7$

Damit:

$$\text{var}(\hat{I}_N) \leq \frac{1}{N} e^{-2n H(p_{0.7}|p)}$$

$$\sigma \leq \frac{1}{\sqrt{N}} e^{-n H(p_{0.7}|p)}$$

Nach Cramér:

$$P[A_n] \leq e^{-n H(p_{0.7}|p)}$$

← obere Schranken haben
jetzt dieselbe
Größenordnung!

Ben. 1) Ähnliche Simulationsprobleme treten in der statistischen Mechanik auf, aber die X_i sind in der Regel nicht unabhängig.

2) Um zu zeigen, dass $P[A_n]$ in einer festen Anzahl von Schritten hinreichend gut approximiert wird, bräuhete man eine möglichst genaue untere Schranke für $P[A_n]$.

- untere Schranke auf exp. Skala: s.u.

- genauere untere Schranke: nur in sehr expliziten Modellen

16.2. Satz von Shannon Mc Millan, spezifische relative Entropie

Situation wie am Anfang des Abschnitts

$$Z_n = \frac{dQ}{dP} \Big|_{\mathcal{F}_n}, \quad \mathcal{F}_n = \sigma(X_1, \dots, X_n)$$

Def. Es gilt ein Satz von Shannon Mc Millan, falls eine Konstante $h(Q|P) \in [0, \infty)$ existiert mit

$$\frac{1}{n} \log Z_n \rightarrow h(Q|P) \quad Q\text{-f.s.}$$

$$(Z_n \sim e^{n h(Q|P)} \quad Q\text{-f.s.})$$

Beispiele 1) klassischer Spezialfall: X_i i.i.d. unter P und Q mit p bzw. v

$$\Rightarrow \frac{1}{n} \log Z_n = \frac{1}{n} \sum_{i=1}^n \log \frac{dQ}{dP}(X_i) \rightarrow H(v|p) \quad Q\text{-f.s.}$$

$$h(Q|P) = H(v|p)$$

2) X_i unabh. unter P und Q mit p_i bzw. v_i . Es existiere

$$h(Q|P) := \lim_{n \uparrow \infty} \frac{1}{n} H_{\mathcal{F}_n}(Q|P) \quad \text{„spezifische relative Entropie“}$$

$$\text{wobei } H_{\mathcal{F}_n}(Q|P) := \int \log Z_n dQ \stackrel{\text{Faktorisierung}}{=} \sum_{i=1}^n H(v_i|p_i)$$

(relative Entropie auf \mathcal{F}_n)

2) X_i unabh. unter P und Q mit p_i bzw. v_i

~~$\Rightarrow h(Q|P) = \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n H(v_i/p_i)$ mittlere Entropie~~

Existiert $h(Q|P)$, dann:

Shannon McMillan $\Leftrightarrow \underbrace{\frac{1}{n} \log Z_n - \frac{1}{n} H_{F_n}(Q|P)}_{= \frac{1}{n} \sum_{i=1}^n Y_i} \rightarrow 0 \quad Q\text{-f.s.}$
 $Y_i := \log \frac{dv_i}{dp_i}(X_i) - H(v_i/p_i)$

$E_Q[Y_i] = 0$

\Leftrightarrow Gesetz der großen Zahlen für das Martingal $M_n := \sum_{i=1}^n Y_i$

16.3. Untere Schranke für große Abweichungen

Wie klein können Q -wesentliche Mengen bzgl. P sein?

Satz Gibt es Satz von Shannon McMillan, dann folgt:

1) $\lim_{n \rightarrow \infty} \frac{1}{n} \log P[A_n] \geq -h(Q|P)$

für jede Folge $A_n \in F_n$ mit $\lim Q[A_n] > 0$.

2) Diese Schranke ist optimal: Für jedes $\epsilon > 0$ gibt es $A_n \in F_n$ mit $Q[A_n] \rightarrow 1$, aber $\frac{1}{n} \log P[A_n] \leq -h(Q|P) + \epsilon$.

16.4. Beispiel: Große Abweichungen im Münzwurfmodell

$$P = \bigotimes_{i=1}^{\infty} \mu_p, \quad \mu_p(0) = 1-p, \quad \mu_p(1) = p, \quad p \in (0,1)$$

$$A_n := \left\{ \frac{1}{n} \sum_{i=1}^n X_i \geq m \right\}, \quad m > p$$

$P[A_n] \rightarrow 0$ nach Gesetz der großen Zahlen.

$$Q := \bigotimes_{i=1}^{\infty} \mu_m \xrightarrow{\text{ZGS}} \lim Q[A_n] \approx 1/2$$

$$\begin{aligned} \stackrel{16.3}{\Rightarrow} \lim \frac{1}{n} \log P[A_n] &\geq -h(Q|P) = -H(\mu_m | \mu_p) \\ &= m \log \frac{m}{p} + (1-m) \log \frac{1-m}{p} \end{aligned}$$

obere Schranke
Satz von Granér: $\frac{1}{n} \log P[A_n] \leq -\Lambda^*(m) = -\sup_{t \in \mathbb{R}} (tm - \log E[e^{tX_1}])$ // s. W'K. I, 19.5

Korollar $\frac{1}{n} \log P[\frac{1}{n} \sum_{i=1}^n X_i \geq m] \rightarrow -H(\mu_m | \mu_p)$

$$P[\frac{1}{n} \sum_{i=1}^n X_i \geq m] \sim e^{-n H(\mu_m | \mu_p)}$$

(auf exponentieller Skala)

16.5. Beispiel Scndlich

unter P: X_1, X_2, \dots i.i.d. mit $\mu = \text{Gleichverteilung auf } S$

unter Q: X_1, X_2, \dots i.i.d. mit ν

$$h(Q|P) = H(\nu|\mu) = \log |S| - H(\nu)$$

$$C_n \subseteq S^n: \frac{1}{n} \log P[(X_1, \dots, X_n) \in C_n] = \frac{1}{n} \log \frac{|C_n|}{|S|^n} = \frac{1}{n} \log |C_n| - \log |S|$$

Korollar 1) $C_n \subseteq S^n$ wesentlich für Q (d.h. $A_n = \{(x_1, \dots, x_n) \in C_n\}$ wesentlich für Q)

$$\Rightarrow \liminf \frac{1}{n} \log |C_n| \geq H(\nu)$$

$$2) C_{n,\epsilon} := \{(x_1, \dots, x_n) \mid \frac{1}{n} \sum_{i=1}^n \log \nu(x_i) \geq -H(\nu) - \epsilon\}$$

ist wesentlich für Q und

$$\frac{1}{n} \log |C_{n,\epsilon}| \leq H(\nu) + \epsilon \quad \forall n$$

16.6. Anwendung auf Codierung

Eine Nachrichtenquelle sendet Texte (x_1, \dots, x_n) aus dem Alphabet S (d.h. $x_i \in S$).

Binäre Codierung = Bijektion $\varphi: S^n \rightarrow \{0,1\}^k$

ermöglicht eindeutige Codierung von 2^k verschiedenen Texten.

$k(n)$ = minimale Anzahl Bits für vollständige Codierung:

$$2^{k(n)-1} < |S^n| \leq 2^{k(n)} \Rightarrow \frac{k(n)}{n} \geq \log_2 |S| > \frac{k(n)-1}{n}$$

$$\Rightarrow \lim_{n \rightarrow \infty} \frac{k(n)}{n} = \log_2 |S| = -\sum p(x) \log_2 p(x) = \frac{H(\text{Gleichverteilung})}{\log 2}$$

Benötigte Anzahl Bits pro Zeichen

"Entropie zur Basis 2"

Wie viele Bits werden benötigt, um richtige Übertragung mit $W'_{\text{kor}} \geq 1-\epsilon$ zu gewährleisten?

Ansatz von Shannon/Wiener: Wenn nicht bekannt ist, setze zufällige

Verteilung für X_1, X_2, \dots, a_n , z. B.:

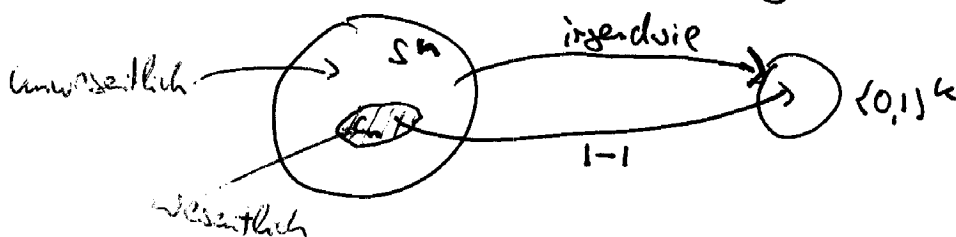
Einfachstes Modell: X_n i.i.d. bzgl. Q mit Verteilung ν ,

$\nu(x)$ = relative Häufigkeit des Buchstaben x in der verwendeten Sprache

Ziel: Finde $C_n \subseteq S^n$ mit $DQ[(X_1, \dots, X_n) \in C_n] \geq 1-\epsilon$

2) $|C_n|$ möglichst klein

Dann kodiere 1-1 auf C_n , beliebig auf C_n^c



Wie klein können wir C_n wählen, damit 1) gilt?

16.4. $\Rightarrow \lim_{n \rightarrow \infty} \frac{1}{n} \log |C_n| \geq H(\nu)$ stets

$\frac{1}{n} \log |C_n| \leq H(\nu) + \delta$ bei geeigneter Wahl
 $\approx 2^{k-1}$ minimales Wahl von k

Korollar (Quellencodierungssatz, Shannon/Wiener)

$\lim_{n \rightarrow \infty} \frac{k(n, \epsilon)}{n} = \frac{H(\nu)}{\log 2} \quad \forall \epsilon > 0$

Entropie von ν statt Entropie der Gleichverteilung!

17. Minima der relativen Entropie und der Satz von Cramér

(S, \mathcal{F}) meßb. Raum, $\mathcal{M}_1(S) :=$ alle WV auf (S, \mathcal{F})

$X_1, X_2, \dots : \Omega \rightarrow S$ i.i.d. bzgl. P mit Verteilung $\mu \in \mathcal{M}_1(S)$

$U: S \rightarrow \mathbb{R}$, $U \in \mathcal{L}^1(\mu)$, $m_0 := \int U d\mu$

$\Rightarrow U(X_i) \in \mathcal{L}^1(P)$ i.i.d. mit $E[U(X_i)] = m_0$

Gesetz der großen Zahlen: $P[\frac{1}{n} \sum_{i=1}^n U(X_i) \geq a] \rightarrow 0 \quad \forall a > m_0$

Ziel: Satz von Cramér: $\frac{1}{n} \log P[\frac{1}{n} \sum_{i=1}^n U(X_i) \geq a] \xrightarrow{n \rightarrow \infty} -I(a)$

wobei $I(a) = \sup_{t \in \mathbb{R}} (at - \log \int e^{tU} d\mu)$

$\left[\begin{array}{l} = \min_{\nu \in \mathcal{M}_1(S)} H(\nu | \mu) \\ \int U d\nu = a \end{array} \right] = \min_{\nu \in \mathcal{M}_1(S)} H(\nu | \mu) \quad \text{falls } \int U d\mu < \infty \quad \forall t \in \mathbb{R}$

Bem. $H(\nu | \mu) =$ nichtsymmetrischer Abstand ~~der~~ WV ν und μ
 exp. Konvergenzrate = kleinster Abstand einer WV mit Mittel $\geq a$
 (d.h. große Abweichung ist normal) von μ

17.1. Die untere Schranke

Lemma $\lim_{n \rightarrow \infty} \frac{1}{n} \log P\left[\frac{1}{n} \sum_{i=1}^n U(X_i) \geq a\right] \geq -\inf_{\nu \in M(S), \int U d\nu \geq a} H(\nu | \mu) \quad \forall a \in \mathbb{R}$

Bem. Für $a \leq m_\mu$ ist rechte Seite = 0 ($\nu = \mu$)

Beweis: Betrachte oBdA kanonisches Modell:

$\Omega = S^{\mathbb{N}}, X_i(\omega) = \omega_i, P = \bigotimes_{i=1}^{\infty} \mu$

Sei $\nu \in M(S)$ mit $\int U d\nu \geq a, Q := \bigotimes_{i=1}^{\infty} \nu$

$Q\left[\frac{1}{n} \sum_{i=1}^n U(X_i) \geq a\right] \rightarrow 1$ nach GdJZ

$\stackrel{16.3.}{\implies} \lim_{n \rightarrow \infty} \frac{1}{n} \log P\left[\frac{1}{n} \sum_{i=1}^n U(X_i) \geq a\right] \geq -H(\nu | \mu)$

gilt für alle ν mit $\int U d\nu \geq a$, also auch $\geq -\inf_{\nu: \int U d\nu \geq a} H(\nu | \mu)$. □

Oberer Schranke nach Cramér: $-\sup_{t \in \mathbb{R}} \left(at - \log \overbrace{E[e^{tU(X_1)}]}^{= \int e^{tu} d\mu} \right)$

Identifikation mit unterer Schranke?

→ Berechnung des Minimums von $H(\nu | \mu)$ unter Nebenbed. $\int U d\nu = a$

Variationsproblem! Lösungswege: 1) via Lagrange-Multiplikatoren
2) direkt:

17.2. Variationsformel (S, \mathcal{F}, μ) W'raum

Lemma Sei $\bar{\nu} \in \mathcal{M}_+(S)$, $\bar{\nu} \approx \mu$ mit Dichte $\bar{g} = \frac{d\bar{\nu}}{d\mu} > 0$.

Ist $H(\bar{\nu} | \mu) < \infty$, dann gilt:

$$H(\nu | \mu) = H(\nu | \bar{\nu}) + \int \log \bar{g} \, d\nu \quad \forall \nu \ll \mu \text{ mit } \log \bar{g} \in L^1(\nu)$$

Inbesondere hat $\bar{\nu}$ minimale relative Entropie $H(\cdot | \mu)$ unter allen $\nu \in \mathcal{M}_+(S)$ mit Erwartungswert $\int \log \bar{g} \, d\nu \geq \bar{m} := \int \log \bar{g} \, d\bar{\nu}$.

Beweis: $\nu \ll \mu \Rightarrow \nu \ll \bar{\nu}$, $\frac{d\nu}{d\mu} = \frac{d\nu}{d\bar{\nu}} \cdot \frac{d\bar{\nu}}{d\mu} = \frac{d\nu}{d\bar{\nu}} \cdot \bar{g}$

$$\Rightarrow H(\nu | \mu) = \underbrace{\int \log \frac{d\nu}{d\mu} \, d\nu}_{\in [0, \infty]} = \underbrace{\int \log \frac{d\nu}{d\bar{\nu}} \, d\nu}_{= H(\nu | \bar{\nu}) \in [0, \infty]} + \underbrace{\int \log \bar{g} \, d\nu}_{\text{Integral existiert n.V.}}$$

Also: $H(\nu | \mu) \geq \int \log \bar{g} \, d\nu$, " $=$ " $\Leftrightarrow \nu = \bar{\nu}$. \square

17.3. Exponentielle Familien

Gegeben: $U: S \rightarrow \mathbb{R}$, $\mu \in \mathcal{L}^1(\mu)$, $m \in \mathbb{R}$.

Gesucht: Minimum $\bar{\mu}$ der relativen Entropie $H(\nu|\mu)$ unter Nebenbed. $\int U d\nu \stackrel{(\geq)}{=} m$.

Nach 17.2. ist $\bar{\mu}$ Minimum unter allen WV mit fester Erw.wert von U (bzw., äquivalent, von $\alpha + \beta U$), falls

$$\exists \alpha, \beta \in \mathbb{R}: \log \frac{d\bar{\mu}}{d\mu} = \alpha + \beta U, \quad \frac{d\bar{\mu}}{d\mu} = \text{const.} \cdot e^{\beta U}$$

Def. Für $\beta \in \mathbb{R}$ sei

$$Z(\beta) = \int e^{\beta U} d\mu \in [0, \infty] \quad (\text{momentengenz. Funktion}),$$

$$\mathcal{D} := \{ \beta \in \mathbb{R} \mid Z(\beta) < \infty \}.$$

Die Familie der Wahrscheinlichkeitsverteilungen

$$\mu_{\beta}(dx) := \frac{1}{Z(\beta)} e^{\beta U(x)} \mu(dx), \quad \beta \in \mathcal{D},$$

heißt exponentielle Familie zu μ und U .

$$m(\beta) := \int U d\mu_{\beta} = \frac{1}{Z(\beta)} \int U e^{\beta U} d\mu \in [-\infty, \infty]$$

nach unten (bzw. oben) beschr.
für $\beta < 0$ (bzw. $\beta > 0$).

Bem. 1) Z ist konvex

2) $\Rightarrow \mathcal{J}$ konvex, also Intervall, $0 \in \mathcal{J}$

3) $[-c, c] \subseteq \mathcal{J} \Rightarrow Z$ analytisch auf $(-c, c)$,

$$Z^{(n)}(\beta) = \int U^n e^{\beta U} d\mu \quad (\text{s. W' Theorie I, Abschnitt 19})$$

$$\eta(\beta) = \frac{Z'(\beta)}{Z(\beta)} = (\log Z)'(\beta) \text{ auch analytisch auf } (-c, c)$$

Beispiele 1) $\mu = \text{Exp}(1), U(x) = x \Rightarrow \mathcal{J} = (-\infty, 1), \mu_\beta = \text{Exp}(1-\beta)$

exp. Familie = alle Exp. verteilungen

2) $\mu = N(0, \sigma^2), U(x) = x \Rightarrow \mathcal{J} = \mathbb{R}, \mu_\beta = N(\sigma^2\beta, \sigma^2)$

exp. Familie = alle Normalverteilungen mit fester Varianz

Def.

$$\text{ess sup } U := \min \{c \in \mathbb{R} \mid U \leq c \mu\text{-f.s.}\} \quad (\min \emptyset := \infty)$$

$$\text{ess inf } U := \max \{c \in \mathbb{R} \mid U \geq c \mu\text{-f.s.}\} \quad (\max \emptyset := -\infty)$$

Korollar 1) Sei $\beta \in \mathcal{J}$ mit $\eta(\beta) \in \mathbb{R}$. Dann

$$H(\mu_\beta \mid \mu) \stackrel{(\ast)}{=} \beta \eta(\beta) - \log Z(\beta) = \min_{\nu \in \mathcal{M}(S)} H(\nu \mid \mu) \quad \text{falls } \beta > 0 \text{ (} \beta < 0$$

insbes. hat μ_β min. rel. Entropie $H(\cdot \mid \mu)$ unter allen ν mit $\int \nu d\mu = \eta(\beta)$

2) Gilt $\mathcal{F} = \mathbb{R}$ (z.B. oh falls $|U| \leq c$ μ -f.s.), und ist U nicht μ -f.s. konstant, dann ist

$m: \mathbb{R} \rightarrow (\text{ess inf } U, \text{ess sup } U)$, $m(\beta) = \int U d\mu_\beta$
streng mon. wachsend, C^∞ , und surjektiv. Ist $\beta(m)$ die Umkehrfunktion, dann gilt

$$\min_{\int U d\nu > a} H(\nu | \mu) = \min_{\int U d\nu = a} H(\nu | \mu) = \inf_{\int U d\nu > a} H(\nu | \mu) = H(\mu_{\beta(a)} | \mu)$$

für alle $a \in (\text{ess inf } U, \text{ess sup } U)$.

Beweis: 1) $\log \frac{d\mu_\beta}{d\mu} = \beta \cdot U - \log Z(\beta) \Rightarrow$ (i) (*)

(ii) μ_β hat min. rel Entropie unter allen ν mit

$$\int (\beta U - \log Z(\beta)) d\nu \geq \int (\beta U - \log Z(\beta)) d\mu_\beta$$

$$\Leftrightarrow \begin{cases} \int U d\nu \geq \int U d\mu_\beta & \text{falls } \beta > 0 \\ \int U d\nu \leq \int U d\mu_\beta & \text{falls } \beta < 0 \end{cases}$$

2) $Z(\beta) < \infty \forall \beta \in \mathbb{R} \xrightarrow{\text{Ban.}} Z, m \in C^\infty, m = \frac{Z'}{Z}$

$$\Rightarrow m'(\beta) = \frac{Z''(\beta)}{Z(\beta)} - \left(\frac{Z'(\beta)}{Z(\beta)} \right)^2 = \text{var}_{\mu_\beta}(U) > 0$$

$\underbrace{\hspace{10em}}_{= \int U^2 d\mu_\beta} \quad \underbrace{\hspace{10em}}_{= \int U d\mu_\beta}$

\uparrow
nicht μ -f.s. konstant

(163a)

Zudem gilt $\lim_{\beta \uparrow \infty} m(\beta) = \text{esssup } U$, $\lim_{\beta \downarrow -\infty} m(\beta) = \text{essinf } U$ (Übung).

also bijektiv auf $(\text{essinf } U, \text{esssup } U)$.

$$a \in (\underbrace{m_0}_{=m(0)}, \underbrace{\text{esssup } U}_{=m(\infty)}) \Rightarrow \beta(a) \in (0, \infty)$$

$$\stackrel{1)}{\Rightarrow} H(\mu_{\beta(a)} | \mu) = \min_{\mathcal{S} \text{udo} = a} H(\nu | \mu) = \min_{\mathcal{S} \text{udo} \geq a} H(\nu | \mu)$$

$$\inf_{\mathcal{S} \text{udo} \geq a} H(\nu | \mu) = \lim_{\bar{a} \uparrow a} \min_{\mathcal{S} \text{udo} \geq \bar{a}} H(\nu | \mu) = H(\mu_{\beta(\bar{a})} | \mu)$$

$= H(\mu_{\beta(\bar{a})} | \mu)$ stetig in \bar{a} nach (*) und
da Z in stetig.

□

17.4. Beispiel: Thermodynamisches Gleichgewicht

Endlich, μ Gleichverteilung, $U: S \rightarrow \mathbb{R}$

z.B. statistische Mechanik:

$S =$ Zustandsraum

$U(x) = -E(x)$, $E(x)$ Energie des Zustands x

$$\mu_{\beta}(x) = \frac{1}{\tilde{Z}(\beta)} e^{-\beta E(x)}, \quad \tilde{Z}(\beta) = \sum_{x \in S} e^{-\beta E(x)}$$

Verteilung im thermodynamischen Gleichgewicht bei Temperatur $T = \frac{1}{\beta}$.

Asymptotiken: $T \downarrow 0$ ($\beta \uparrow \infty$): $\mu_{\beta} \xrightarrow{w} \text{Gleichver. auf absoluten Minima von } E$

$T \uparrow \infty$ ($\beta \downarrow 0$): $\mu_{\beta} \xrightarrow{w} \mu$ Gleichver. auf S

$$H(\nu | \mu) = \log |S| - H(\nu)$$

Kor. $H(\mu_{\beta}) = \max_{\nu \in \mathcal{M}_1(S)} H(\nu)$
 $\int U d\nu = m(\beta)$

Vorgabe der Temperatur \Leftrightarrow Vorgabe der mittleren Energie $-m(\beta)$

μ_{β} maximiert die Entropie unter allen bei $T = \frac{1}{\beta}$ möglichen Verteilungen!

17.5. Minimum der relativen Entropie als Legendetransformation

Lemma Sei $\beta \in \mathcal{J}$ mit $m(\beta) \in \mathbb{R}$. Dann..

$$H(\mu_\beta | \mu) = \beta \cdot m(\beta) - \log Z(\beta) = \max_{t \in \mathbb{R}} (t \cdot m(\beta) - \log Z(t))$$

Legendetransf. von $\log Z$ an der Stelle $m(\beta)$

Beweis: $H(\mu_\beta | \mu) \stackrel{17.2}{=} \underbrace{H(\mu_\beta | \mu_t)}_{\geq 0} + \underbrace{\int \log \frac{d\mu_t}{d\mu} d\mu_\beta}_{= tU - \log Z(t)} \geq t \cdot m(\beta) - \log Z(t)$

für alle $t \in \mathcal{J}$, " " für $t = \beta$. \square

17.6. Satz von Cramér X_1, X_2, \dots i.i.d. bzgl. P mit Verteilung μ

Satz $\frac{1}{n} \log P[\frac{1}{n} \sum_{i=1}^n U(X_i) \in [a, a+\epsilon]] \xrightarrow{n \rightarrow \infty} -I(a)$

für alle $a \stackrel{(*)}{>} m_0 = \int U d\mu$, $\epsilon \in (0, \infty]$,

wobei $I(a) = \sup_{t \in \mathbb{R}} (at - \log \int e^{tU} d\mu)$ "Kosten der großen Abw."

Ist U μ -f.s. beschränkt und $a \in (\text{ess inf } U, \text{ess sup } U)$, dann gilt

$$I(a) = -\min_{\nu \in \mathcal{M}(U)} H(\nu | \mu) = -H(\mu_{\beta(a)} | \mu)$$

(U und $\mathcal{M}(U)$ / l. 17.5.1)

1) obere Schranke: via exponentielle Markovungleichung, s. W-Theorie I

$$\frac{1}{n} \log P\left[\frac{1}{n} \sum_{i=1}^n U(X_i) \geq a\right] \leq -\left(at - \log E[e^{tU(X_1)}]\right) \quad \forall t > 0$$

$$\Rightarrow \text{auch} \leq -I(a) = -\sup_{t > 0} \dots$$

Aussage für ALLE n - nicht nur asymptotisch!

2) untere Schranke falls U μ -f.s. beschränkt: via importance sampling

$$\liminf \frac{1}{n} \log P\left[\frac{1}{n} \sum_{i=1}^n U(X_i) \in [a, a+\varepsilon]\right] \stackrel{P.1}{\geq} -H(\nu) \quad \forall \nu: \int U d\nu \in (a, a+\varepsilon)$$

$$\Rightarrow \text{auch} \geq -\inf_{\nu} H(\nu) \stackrel{P.3}{=} -\min_{\nu} H(\nu) \stackrel{P.3}{=} -H(\mu_{P(a)} | \mu) \stackrel{P.5}{=} -I(a)$$

3) untere Schranke allgemein: via Abschneideargument

Betrachte zunächst Wahrscheinlichkeit bzgl. $P_c = \bigotimes_{i=1}^{\infty} \mu(\cdot | |U| \leq c)$,
dann $c \uparrow \infty$.

Sei $c > 0$ mit $\mu(|U| \leq c) \neq 0$.

$$P\left[\frac{1}{n} \sum_{i=1}^n U(X_i) \in [a, a+\varepsilon]\right] \geq P\left[\frac{1}{n} \sum_{i=1}^n U(X_i) \in [a, a+\varepsilon]; |U(X_i)| \leq c \forall 1 \leq i \leq n\right]$$

$$\begin{aligned} (*) &= P\left[\frac{1}{n} \sum_{i=1}^n U(X_i) \in [a, a+\varepsilon] \mid |U(X_i)| \leq c \forall 1 \leq i \leq n\right] \cdot P[|U(X_i)| \leq c]^n \\ &= P^c\left[\frac{1}{n} \sum_{i=1}^n U(X_i) \in [a, a+\varepsilon]\right] \cdot \mu(|U| \leq c)^n \end{aligned}$$

wobei $P^c = \bigotimes_{i=1}^{\infty} \mu^c$, $\mu^c := \mu(\cdot \mid |U| \leq c)$.

Bzgl. μ^c ist U essentiell beschränkt ($|U| \leq c$ μ^c -f.s.)

$$\begin{aligned} \stackrel{2)}{\implies} \liminf \frac{1}{n} \log P^c\left[\frac{1}{n} \sum_{i=1}^n U(X_i) \in [a, a+\varepsilon]\right] &\geq -\max_{t \in \mathbb{R}} (at - \log \int e^{tU} d\mu^c) \\ &= \frac{\int_{\{|U| \leq c\}} e^{tU} d\mu}{\mu(|U| \leq c)} \end{aligned}$$

$$\begin{aligned} (*) \implies \liminf \frac{1}{n} \log P[\dots] &\geq \underbrace{-\max_{t \in \mathbb{R}} (at - \log \int_{\{|U| \leq c\}} e^{tU} d\mu)}_{\substack{\text{s. 174} \\ = at_c - \log \int_{\{|U| \leq c\}} e^{t_c U} d\mu \text{ für ein } t_c \in \mathbb{R}}} - \log \mu(|U| \leq c) + \log \mu(|U| \leq c) \end{aligned}$$

$$\implies \liminf \frac{1}{n} \log P[\dots] \geq \lim_{c \uparrow \infty} (at_c - \log \int_{\{|U| \leq c\}} e^{t_c U}) \geq a\bar{t} - \log \int e^{\bar{t}U} \geq -I(a)$$

falls die Folge $\{t_c \mid c \in \mathbb{N}\}$ einen Häufungspunkt $\bar{t} \in \mathbb{R}$ hat.

Existenz eines Häufungspunkts: t_c ist Minimum von

$$-at + \log \int e^{tU} d\rho_c = \log \int e^{t(U-a)} d\rho_c =: f_c(t)$$

$\rightarrow \infty$ für $t \rightarrow \pm \infty$

falls $a \in (\text{essinf } U, \text{esssup } U)$, c groß genug

Man kann zeigen:

- $\exists K \subset \mathbb{R}$ kompakt: $\{t_c \mid c \in \mathbb{N}\} \subseteq K$ $\forall c$ groß genug
 t_c da $f_c(0) = 1$

also \exists Häufungspunkt von $\{t_c \mid c \in \mathbb{N}\}$. \square

- 17.7. Beispiel: Boltzmanns Entropie

S endlich, μ Gleichverteilung, $U: S \rightarrow \mathbb{R}$

z. B. statistische Mechanik:

S = Zustandsraum für ein Teilchen

S^n = Zustandsraum für n -Teilchen-System (n sehr groß)

Energieschale zum Energieniveau $-a$:

$$\Gamma_{a,\epsilon}^{(n)} = \left\{ x = (x_1, \dots, x_n) \mid \underbrace{\frac{1}{n} \sum_{i=1}^n U(x_i)}_{= \text{mittlere Energie}} \in [a, a+\epsilon] \right\}$$

$$P \left[\frac{1}{n} \sum_{i=1}^n U(x_i) \in [a, a+\epsilon] \right] = \frac{|\Gamma_{a,\epsilon}^{(n)}|}{|S|^n}$$

$$I(a) = H(\mu_{\beta(a)} | \mu) = \log |S| - H(\mu_{\beta(a)})$$

Aus Crooner folgt:

$$\text{Kon. } \underbrace{\frac{1}{n} \log |\Gamma_{a,\epsilon}^{(n)}|}_{\substack{\text{Boltzmanns Entropie} \\ \text{zum Energieniveau } -a}} \xrightarrow{n \uparrow \infty} H(\mu_{\beta(a)})$$

17.8. Eigenschaften der Ratefunktion

$$I(a) = \sup_{t \in \mathbb{R}} (at - \log \int e^{tu} d\mu)$$

1) $I: \mathbb{R} \rightarrow [0, \infty]$ konvex und unterhalbstetig

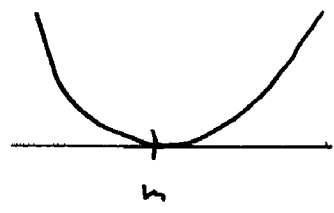
(da $I = \Lambda^*$ Legendretrf. der konvexen, unterhalbstetigen Fkt. $\Lambda(t) = \log \int e^{tu} d\mu$)

2) $I(m_0) = 0$ (nach Jensen: $\log \int e^{tu} d\mu \geq \int \log e^{tu} d\mu = m_0 \cdot t$)

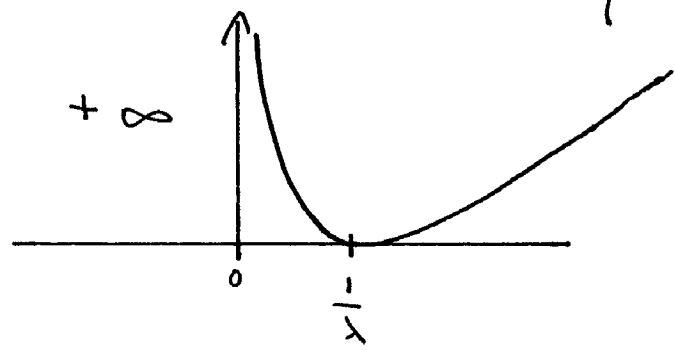
3) $I(a) = \sup_{\substack{t > 0 \\ (t \leq 0)}} (at - \log \int e^{tu} d\mu)$ falls $a > m_0$ ($a \leq m_0$)
 (nach Jensen, s. W' Theorie, Kap. 19)

insbesondere: $I \nearrow$ auf $[m_0, \infty)$, $I \searrow$ auf $(-\infty, m_0]$.

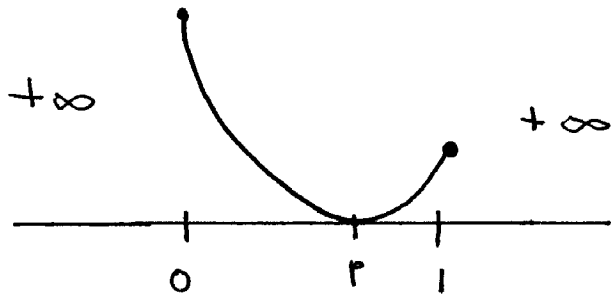
Beispiele 1) $U(X_i) \sim N(m, \sigma^2) \Rightarrow I(a) = \frac{(a-m)^2}{2\sigma^2}$



2) $U(X_i) \sim \text{Exp}(\lambda) \Rightarrow I(a) = \begin{cases} \lambda a - 1 - \log(\lambda a) & \text{für } a > 0 \\ +\infty & \text{für } a \leq 0 \end{cases}$



3) $U(X_i) \sim \text{Bernoulli}(p) \Rightarrow I(a) = \begin{cases} a \log \frac{a}{p} + (1-a) \log \frac{1-a}{1-p}, & a \in [0, 1] \\ +\infty & \text{sonst} \end{cases}$



Beachte: Die Werte 0 und 1 treten mit positiver W'keit als Mittelwert

$\frac{1}{n} \sum_{i=1}^n U(X_i)$ auf diese W'keiten fallen mit einer endlichen exp. Rate ab: $I(0) \neq \infty, I(1) \neq \infty$

18. Prinzipien der großen Abweichungen (Large deviation principles)

(170)

Betrachte Situation wie in Abschnitt 17.

18.1. Level 1

$$\mu_n(B) := \mathbb{P}\left[\frac{1}{n} \sum_{i=1}^n U(X_i) \in B\right], \quad B \in \mathcal{B}(\mathbb{R}).$$

- Asymptotische Darstellung von μ_n für $n \rightarrow \infty$?

$$\text{Cramér: } \mu_n([a, a+\varepsilon]) \sim e^{-nI(a)} \quad \forall a > m_0 \\ (\text{auf expo-enteller Skala})$$

Allgemeiner:

- Satz μ_n erfüllt ein Prinzip der großen Abweichungen mit Ratefunktion

$$I(a) = \sup_{t \in \mathbb{R}} (at - \log \int e^{tu} d\mu) \in [0, \infty]$$

d.h.

$$(i) \quad \overline{\lim}_{n \rightarrow \infty} \frac{1}{n} \log \mu_n(A) \leq - \inf_{x \in A} I(x) \quad \forall A \subset \mathbb{R} \text{ abg.}$$

$$(ii) \quad \underline{\lim}_{n \rightarrow \infty} \frac{1}{n} \log \mu_n(O) \geq - \inf_{x \in O} I(x) \quad \forall O \subset \mathbb{R} \text{ offen}$$

Für beliebige $B \in \mathcal{B}(\mathbb{R})$ gilt also:

$$-\inf_{x \in \mathbb{R}} I(x) \leq \liminf_{n \rightarrow \infty} \frac{1}{n} \log \mu_n(B) \leq \overline{\lim}_{n \rightarrow \infty} \frac{1}{n} \log \mu_n(B) \leq -\inf_{x \in \mathbb{R}} I(x)$$

(d.h. die Wert der großen Abweichung ist bestimmt durch das $x \in \mathbb{R}$ mit minimaler Kosten $I(x)$ (falls \exists))

Beweis: Übung

Bem. 1) $A = [a, a + \varepsilon] \Rightarrow I(a) = \inf_{x \in A} I(x)$ (da $I \nearrow$ auf $[n_0, \infty)$)

daher tritt in der Aussage von Gärtner kein Infimum auf

2) Heuristische Asymptotik (auf exp. Skala, d.h. modulo logarithmischer wachsende Faktoren):

$$\mu_n(dx) \sim e^{-n I(x)} dx \quad \text{für } n \rightarrow \infty$$

$$\mu_n(B) \sim \int_B e^{-n I(x)} dx \sim e^{-n \inf_{x \in B} I(x)}$$

Jede große Abweichung geschieht auf die wahrscheinlichste aller unwahrscheinlichen Arten

18.2. Level 2: Der Satz von Sanov

S polnischer Raum (metrisch, vollständig, separabel), $\mathcal{I} = \mathcal{B}(S)$

$X_1, X_2, \dots : \Omega \rightarrow S$ i.i.d. bzgl. P mit Verteilung μ

$\mathcal{P}_n(\omega) = \frac{1}{n} \sum_{i=1}^n \delta_{X_i(\omega)}$ empirische Verteilung von X_1, \dots, X_n

$\mathcal{P}_n : \Omega \rightarrow \mathcal{M}_1(S)$ n -wertige Zufallsvariable

Gesetz der großen Zahlen:

P -f.s. gilt: $\mathcal{P}_n(\omega) \xrightarrow{w} \mu$ (schwache Konvergenz)

d.h. $\int f d\mathcal{P}_n(\omega) = \frac{1}{n} \sum_{i=1}^n f(X_i(\omega)) \rightarrow \int f d\mu \quad \forall f \in C_b(S)$

Satz von Sanov: Die Verteilung von \mathcal{P}_n unter P erfüllt ein Prinzip der großen Abweichungen mit Ratefunktion $I^{(2)} : \mathcal{M}_1(S) \rightarrow [0, \infty]$,

$$I^{(2)}(\nu) = H(\nu/\mu),$$

d.h.

(i) $\overline{\lim} \frac{1}{n} \log P[\mathcal{P}_n \in A] \leq -\inf_{\nu \in A} H(\nu/\mu) \quad \forall A \subset \mathcal{M}_1(S) \text{ abgeschlossen}$

(ii) $\underline{\lim} \frac{1}{n} \log P[\mathcal{P}_n \in \Theta] \leq -\inf_{\nu \in \Theta} H(\nu/\mu) \quad \forall \Theta \subset \mathcal{M}_1(S) \text{ offen}$

wobei $\mathcal{M}_1(S)$ versehen mit Topologie der schwachen Konvergenz.

IV. BROWNSCHE BEWEGUNG

Zusätzliche Literatur:

Rogers, Williams: Diffusions, Markov processes and martingales, Vol. 1, Ch. I

Karatzas / Shreve: Brownian motion and stochastic calculus

Steele: Stochastic calculus and financial applications

19. Brownsche Bewegung und Wienermaß (Grundlagen)

Brown (Botaniker, 19. Jhdt.): beobachtete Molekularbewegung

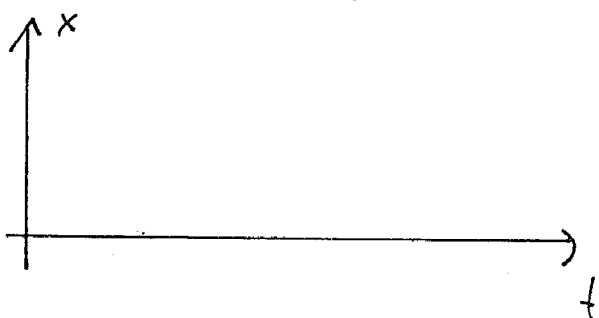
Einstein (1905): Modell, Berechnung der Avogadrozahl

Bachelier (1900): Modell für Finanzmärkte

„Sur la théorie de la spéculation“

N. Wiener (1923): Rigorose mathematische Formulierung

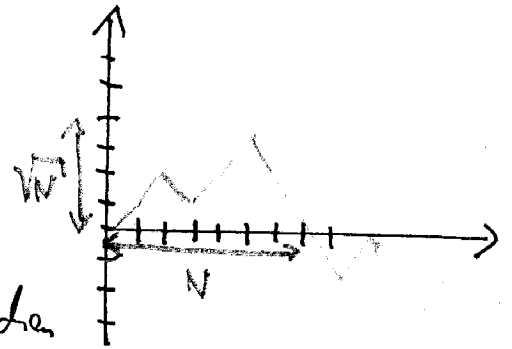
19.1. Brownsche Bewegung in Dimension 1 - Heristile



Random walk $S_k = \sum_{i=1}^k X_i$, X_i i.i.d., ± 1 mit $p = \frac{1}{2}$. (175)

Reskaliere und interpoliere:

$$S_t^N(\omega) = \begin{cases} \frac{S_k(\omega)}{\sqrt{N}} & \text{falls } t = \frac{k}{N} \\ \text{lin. Interpolation dazwischen} & \end{cases}$$



$$E[S_t^N] = 0$$

$$\text{Var}(S_t^N) = \frac{1}{N} E[S_k^2] = \frac{k}{N} = t \quad \text{für } t = \frac{k}{N}$$

Erwartungswerte und Varianzen sind konstant in N ∇

Für jede Unterteilung $0 \leq t_0 < t_1 < \dots < t_n$ mit $t_i = \frac{k_i}{N}$

Sind die Inkremente $S_{t_1}^N - S_{t_0}^N, S_{t_2}^N - S_{t_1}^N, \dots, S_{t_n}^N - S_{t_{n-1}}^N$

unabhängige, zentrierte ZV mit Varianz

$$E[(S_{t_{i+1}}^N - S_{t_i}^N)^2] = \frac{1}{N} E[(S_{k_{i+1}} - S_{k_i})^2] = \frac{k_{i+1} - k_i}{N} = t_{i+1} - t_i$$

Zentraler Grenzwertsatz

\Rightarrow Die Verteilung von $S_t^N - S_s^N$ ($0 \leq s < t$) konv. schwach gegen $N(0, t-s)$

19.2. Definition Sei $a \in \mathbb{R}$.

(176)

Ein stochastischer Prozeß $B_t: \Omega \rightarrow \mathbb{R}$, $t \in [0, \infty)$, auf (Ω, \mathcal{O}, P) heißt Brownsche Bewegung (Wienerprozeß) mit Start in a , falls

(i) $B_0 = a$ P-f.s.

(ii) Für alle $0 \leq t_0 < t_1 < t_2 < \dots < t_n$ sind die Inkremente $B_{t_{i+1}} - B_{t_i}$ unabhängig und $N(0, t_{i+1} - t_i)$ verteilt.

(iii) $t \mapsto B_t(\omega)$ ist P-f.s. stetig.

(d.h. $\exists N \in \mathcal{O}: P[N] = 0$, $t \mapsto B_t(\omega)$ stetig $\forall \omega \in \Omega \setminus N$)

Kein Startpunkt gegeben: Betrachte B.B. mit Start in 0.

Existenz, Eindeutigkeit in Verteilung: s.u. (21, 19.8)

Bemerkungen: 1) Stetigkeit muß vorausgesetzt werden:

$$(B_t)_{t \geq 0} \text{ B.B.}, \quad \tilde{B}_t(\omega) := \begin{cases} B_t(\omega) & \text{falls } B_t(\omega) \notin \mathbb{Q} \\ 0 & \text{sonst} \end{cases}$$

$\Rightarrow \tilde{B}_t$ erfüllt (i) und (ii), aber nicht (iii).

2) $(B_t)_{t \geq 0}$ B.B. mit Start in 0

(177)

$\Rightarrow (a+B_t)_{t \geq 0}$ B.B. mit Start in a „räumlich homogen“

3) Brownsche Bewegungen sind die einzigen stochastischen Prozesse

$(B_t)_{t \geq 0}$ mit (i), (iii), $\text{Var}(B_t) = t$, und unabhängigen

stationären Inkrementen (d.h. $B_{t+h} - B_{s+h} \sim B_t - B_s \quad \forall 0 \leq s \leq t, h \geq 0$)

Normalverteilung der Inkremente folgt aus diesen Voraussetzungen
mithilfe von erweitertem ZGS, s. Breiman: Probability.

19.3. Brownsche Bewegung im \mathbb{R}^d

Ein stoch. Prozeß $(B_t)_{t \geq 0}$ mit Werten in \mathbb{R}^d heißt Br. Bewegung,

falls die Komponenten $(B_t^1)_{t \geq 0}, (B_t^2)_{t \geq 0}, \dots, (B_t^d)_{t \geq 0}$

unabhängige Brownsche Bewegungen sind.

19.4. Brownsche Filtration I Sei $(B_t)_{t \geq 0}$ B.B. im \mathbb{R}^1 mit $B_0 = a$ ~~Start in a~~ .

$$\mathcal{F}_t = \sigma(B_s \mid 0 \leq s \leq t) \quad (t \geq 0)$$

Lemma $B_t - B_s \perp \mathcal{F}_s \quad \forall 0 \leq s \leq t$

(178)

Beweis: $0 \leq t_0 \leq t_1 \leq \dots \leq t_n \leq s$

$$\Rightarrow B_t - B_s \perp \sigma(B_{t_0} - B_{t_0}, \underbrace{B_{t_1} - B_{t_0}}_{=a}, \dots, B_{t_n} - B_{t_{n-1}})$$

$$= \sigma(B_{t_0}, B_{t_1}, \dots, B_{t_n})$$

erzeugen \mathcal{F}_s

\square

19.5. Markoveigenschaft

Satz $(B_t)_{t \geq 0}$ ist ein (zeitstetiger) Markovprozess mit Übergangskern

$$P_t(x, y) = \underbrace{\frac{1}{\sqrt{2\pi t}} e^{-\frac{(x-y)^2}{2t}}}_{=: P_t(x, y)} dy$$

d.h.

$$P[B_t \in A \mid \mathcal{F}_s] = P_{t-s}(B_s, A) = \int_A P_{t-s}(B_s, y) dy$$

P-f.s.

Beweis: $B_t = \underbrace{B_s}_{\mathcal{F}_s\text{-mb.}} + \underbrace{B_t - B_s}_{\perp \mathcal{F}_s}$

$\Rightarrow P[B_t \in A | \mathcal{F}_s](\omega) = P[B_s(\omega) + \overbrace{B_t - B_s}^{\sim N(0, t-s)} \in A]$
 $= N(B_s(\omega), t-s)(A) = \int_A \frac{1}{\sqrt{2\pi(t-s)}} e^{-\frac{(y-B_s(\omega))^2}{2 \cdot (t-s)}} dy$
 $\quad \quad \quad \text{P-f.s.} \quad \quad \quad = P_{t-s}(B_s(\omega), y) \quad \square$

Bem. 1) Analog: BB im \mathbb{R}^d ist Markovprozess mit Kern

$$P_t(x, y) = \frac{1}{\sqrt{2\pi t}^d} e^{-\frac{|x-y|_{\mathbb{R}^d}^2}{2t}}$$

2) $P_t(x, y)$ ist Fundamentallösung der Wärmeleitgleichung

$$\frac{\partial}{\partial t} P_t(x, y) = \frac{1}{2} \Delta_x P_t(x, y), \quad \Delta = \sum_{i=1}^d \frac{\partial^2}{\partial x_i^2}$$

(analog zu Kolmogorovs Rückwärtsgl. bei zeitstetigen Markovketten)

$$\lim_{t \downarrow 0} \int P_t(x, y) f(y) dy = f(x) \quad \forall f \in C_b(\mathbb{R}^d)$$

(d.h. $P_t(x, y) = \delta_t(x) = \delta_t(y-x)$)

Mikroskopisch: Brownsche Bewegung - Stoch. Prozeß

Makroskopisch: Wärmeleitung - PDgl.

19.6. Randverteilungen

Kor. $E[f(B_{t_1}, B_{t_2}, \dots, B_{t_n})]$

$$= \int \dots \int f(x_1, \dots, x_n) p_{t_1}(0, x_1) p_{t_2-t_1}(x_1, x_2) p_{t_3-t_2}(x_2, x_3) \dots p_{t_n-t_{n-1}}(x_{n-1}, x_n) dx_n \dots dx_1$$

für alle $f \geq 0$ meßbar, $0 \leq t_1 < t_2 < \dots < t_n$ (Übung)

Bsp. B_t B.B. mit Start in 0, f beschr. und stetig

$$\Rightarrow u(t, x) = E[f(\underbrace{x + B_t}_{\text{BB mit Start in } x})] = \int p_t(x, y) f(y) dy$$

löst $\frac{\partial u}{\partial t} = \frac{1}{2} \Delta u$ mit Anfangsbedingung $\lim_{t \downarrow 0} u(t, x) = f(x)$.

Kanonisches Modell für Markovkette:

$$\Omega = S^{\mathbb{N}_0} = \{(\omega_0, \omega_1, \dots) \mid \omega_i \in S\}$$

Raum aller möglichen Verläufe

Kanonisches Modell für Brownsche Bewegung in \mathbb{R}^d :

$$\Omega = C([0, \infty), \mathbb{R}^d)$$

Raum aller möglichen Trajektorien

19.7. Zufällige Pfade

Sei $(B_t)_{t \geq 0}$ B.B. im \mathbb{R}^d mit Start in x_0 .

O.B.d.A. $t \mapsto B_t(\omega)$ stetig für alle ω (sonst modifiziere auf Nullmenge)

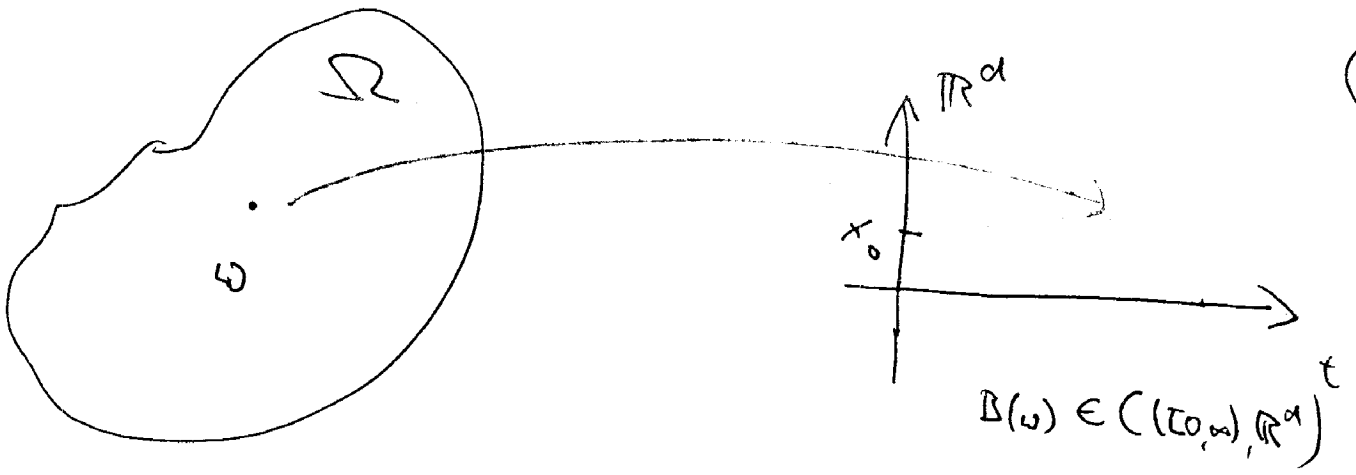
Dann können wir $(B_t)_{t \geq 0}$ auch als eine einzige Zufallsvariable

$$B: \Omega \rightarrow C([0, \infty), \mathbb{R}^d)$$

$$\omega \mapsto (B_t(\omega))_{t \geq 0}$$

„Zufälliger Pfad“

anfassen.



Sei \$\mathcal{B}\$ die von den Abbildungen \$X_t(\bar{\omega}) = \bar{\omega}(t)\$, \$t \ge 0\$, erzeugte \$\sigma\$-Algebra auf \$\widehat{\Omega} = C([0, \infty), \mathbb{R}^d)\$.

Bem. \$\mathcal{B}\$ = Borelsche \$\sigma\$-Algebra auf \$C([0, \infty), \mathbb{R}^d)\$ bzgl. üblicher Topologie

Lemma \$B: (\Omega, \mathcal{A}) \to (\widehat{\Omega}, \mathcal{B})\$ ist meßbar

Beweis: \$B^{-1}(\{X_{t_1} \in A_1, \dots, X_{t_n} \in A_n\}) = \{X_{t_1} \circ B \in A_1, \dots, X_{t_n} \circ B \in A_n\} \in \mathcal{A}\$.
 (Note: \$X_{t_i} \circ B = B_{t_i}\$)

Die Verteilung \$\tilde{P}_{x_0} = P_0 \circ B^{-1}\$ auf dem Pfadraum \$\widehat{\Omega} = C([0, \infty), \mathbb{R}^d)\$
 (Note: \$P_0\$ ist eine WV)

19.8. Eindeutigkeit in Verteilung, Wienermaß

Satz \$\tilde{P}_{x_0}\$ ist eindeutig festgelegt durch die Randverteilungen

*) \$\tilde{P}_{x_0}[X_{t_1} \in A_1, \dots, X_{t_n} \in A_n] = \int_{A_1} \dots \int_{A_n} \prod_{i=1}^n P_{t_i - t_{i-1}}(x_{i-1}, x_i) dx_n \dots dx_1 = \int_{A_1} \dots \int_{A_n} \prod_{i=1}^n \frac{1}{(2\pi(t_i - t_{i-1}))^{d/2}} e^{-\frac{1}{2} \sum_{i=1}^n \frac{(x_i - x_{i-1})^2}{2(t_i - t_{i-1})}} dx\$

$$0 = t_0 < t_1 < t_2 < \dots < t_n, \quad A_1, \dots, A_n \in \mathcal{B}(\mathbb{R}).$$

(183)

Beweis: $\tilde{P}_{x_0} [X_{t_i} \in A_i, \forall 1 \leq i \leq n] = P [\underbrace{X_{t_i} \circ B \in A_i}_{= B_{t_i}}, \forall 1 \leq i \leq n] = \text{r. S.}$

Diese Ereignisse bilden einen \mathcal{P} -stabilen Erzeuger von \mathcal{B} . \square

Def. Die durch (*) festgelegte WV \tilde{P}_{x_0} auf $C([0, \infty), \mathbb{R}^d)$ heißt Wienermaß (mit Start in x_0).

Bem. Die Randverteilungen sind multivariante Normalverteilungen.

\rightarrow Interpretation von \tilde{P}_{x_0} als Gaußmaß auf dem ∞ -dimensionalen Raum $C([0, \infty), \mathbb{R}^d)$.

19.9. Kanonisches Modell der Brownschen Bewegung

Der kanonische Prozeß $X_t(\tilde{\omega}) = \tilde{\omega}(t)$ auf $\tilde{\Omega} = C([0, \infty), \mathbb{R}^d)$ ist unter \tilde{P}_{x_0} eine Brownsche Bewegung mit Start in x_0 .

(denn: $X = (X_t)_{t \geq 0} \Rightarrow \tilde{P} \circ X^{-1} = \tilde{P} = P \circ B^{-1}$)

Bem. Analog kanon. Modell für andere zeitstetige Markovprozesse.

19.10. Heuristische Darstellung des Wienermaßes - Pfadintegrale in der Physik

\tilde{P} := Wienermaß auf $C([0,1])$ bei Start in 0

= Verteilung von $(B_t)_{0 \leq t \leq 1}$, B_t Br. Bewegung, $B_0 = 0$

$\tau^{(n)}$: $0 = t_0^{(n)} < t_1^{(n)} < \dots < t_{k(n)}^{(n)} = 1$ Folge von Partitionen von $[0,1]$

$\tilde{P}^{(n)}$ = Verteilung von $(X_{t_1^{(n)}}^{(n)}, \dots, X_{t_{k(n)}^{(n)}}^{(n)})$ unter \tilde{P} .

$$\tilde{P}^{(n)}(dx_{t_1}, \dots, dx_{t_n}) = \frac{1}{Z_n} e^{-\frac{1}{2} \sum_{i=1}^n \frac{(x_{t_i} - x_{t_{i-1}})^2}{t_i - t_{i-1}}} \prod_{i=1}^n dx_{t_i} \quad (x_0 = 0)$$

↑
Normierungskonstante

- Heuristik für $n \rightarrow \infty$, $\tau^{(n)}$ immer feiner.

$$\tilde{P}(dx) = \frac{1}{Z_\infty} e^{-\frac{1}{2} \int_0^1 |\dot{x}(t)|^2 dt} \prod_{t \in [0,1]} dx_t$$

Feynman-Darstellung des Wienermaßes

NICHT RIGOROS, denn: $Z_n = \infty$, $\int_0^1 |\dot{x}(t)|^2 dt = \infty$ (Br. Pfade sind nicht diffbar)

\mathcal{H} = dimensionales Lebesguemaß

Feynman: Pfadintegraldarstellung von Lösungen der Wärmeleitungsgleichung, bzw. (nach formaler Substitution $t \rightarrow it$) von Lösungen der Schrödingergleichung, s. Feynman lecture notes.

19.11. Brownsche Bewegung als Gaußprozess

185

Erinnerung: $V = (v_{ij})_{i,j=1}^n$ symmetrisch, nicht-negativ definit ($x \cdot V x \geq 0 \quad \forall x \in \mathbb{R}^n$)

1) $X = (X_1, \dots, X_n): \mathbb{R} \rightarrow \mathbb{R}^n$ hat multivariate Normalverteilung $N(\mu, V)$
mit Mittel μ und Kovarianzmatrix V

$$\Leftrightarrow E[e^{ip \cdot X}] = e^{ip \cdot \mu - \frac{1}{2} p \cdot V p} \quad \forall p \in \mathbb{R}^n$$

(In dieser Sinne gilt $\delta_n = N(\mu, 0)$)

2) X multinomial verteilt $\Leftrightarrow p \cdot X = \sum_{i=1}^n p_i X_i$ normalverteilt $\forall p \in \mathbb{R}^n$

3) X multinomial verteilt, $\text{cov}(X_i, X_j) = 0 \quad \forall i \neq j \Rightarrow X_1, \dots, X_n$ unabh.

Satz Ein stetiger stoch. Prozess $(B_t)_{t \geq 0}$ auf $(\mathbb{R}, \mathcal{O}, P)$ ist eine Brownsche Bewegung mit Start in 0 g.d.w.

(i) $(B_t)_{t \geq 0}$ ist Gaußprozess, d.h. $(B_{t_1}, \dots, B_{t_n})$ hat multivariate Normalverteilung für alle $n \in \mathbb{N}$ und $0 \leq t_1 < t_2 < \dots < t_n$.

(ii) $E[B_t] = 0 \quad \forall t \geq 0$.

(iii) $\text{cov}(B_t, B_s) = t \wedge s \quad \forall t, s \geq 0$

Beweis: " \Rightarrow ": nachrechnen

" \Leftarrow ": via Eindeutigkeit der Fouriertransformation in \mathbb{R}^n (Übung)

19.12. Symmetrien der Brownschen Bewegung

Satz $(B_t)_{t \geq 0}$ Brownsche Bewegung mit Start in 0.

Dann sind auch die folgende Prozesse Brownsche Bewegungen:

(i) $(-B_t)_{t \geq 0}$ Reflexionsinvarianz

(ii) $(B_{t+h} - B_h)_{t \geq 0}$ für jedes $h \geq 0$ Zeitliche Homogenität

(iii) $(\frac{1}{\sqrt{a}} B_{at})_{t \geq 0}$ für jedes $a > 0$ Skalainvarianz

(iv) $\hat{B}_t := \begin{cases} t \cdot B_{1/t} & \text{für } t > 0 \\ 0 & \text{für } t = 0 \end{cases}$ Zeitinversion

M.a.W.: Das Wienermaß ist invariant unter den genannten Operationen auf $C([0, \infty))$.

Beweis: (i)-(iii) Übung

(iv) $0 = t_0 < t_1 < \dots < t_n \Rightarrow (\hat{B}_{t_0}, \hat{B}_{t_1}, \dots, \hat{B}_{t_n})$ multinomial verteilt

denn: $p \in \mathbb{R}^{n+1} \Rightarrow \sum_{i=0}^n p_i \hat{B}_{t_i} = \sum_{i=0}^n p_i t_i B_{\frac{1}{t_i}}$ normalverteilt

$$E[\hat{B}_t] = 0 \quad \forall t \geq 0$$

$$\text{cov}(\hat{B}_t, \hat{B}_s) \stackrel{t,s > 0}{=} t \cdot s \cdot \text{cov}(B_{\frac{1}{t}}, B_{\frac{1}{s}}) = t \cdot s \cdot \left(\frac{1}{t} \wedge \frac{1}{s}\right) = (1 \wedge s)$$

$$\text{cov}(\hat{B}_0, \hat{B}_s) = 0 = 0_{1 \times 1} \quad \forall s \geq 0$$

(186)

noch zu zeigen: $t \mapsto \hat{B}_t(\omega)$ ist f.s. stetig

- ok für $t > 0$ da $t \mapsto B_t(\omega)$ stetig
- Damit: Verteilung von $(\hat{B}_t)_{t \geq 0}$ auf $C([0, \infty))$
 $=$ Verteilung von $(B_t)_{t \geq 0}$

(da Randverteilungen gleich: multivariat mit demselben Mittel und Kovarianz)

Insbesondere: $P[\lim_{\substack{t \downarrow 0 \\ t \in \mathbb{Q}}} \hat{B}_t = 0] = P[\lim_{\substack{t \downarrow 0 \\ t \in \mathbb{Q}}} B_t = 0] = 1$

P-f.s.: $\hat{B}_t \xrightarrow[\substack{t \downarrow 0 \\ t \in \mathbb{Q}}]{\text{f.s.}} 0$ und $(\hat{B}_t)_{t \geq 0}$ stetig $\Rightarrow \lim_{\substack{t \downarrow 0 \\ t \in \mathbb{R}}} \hat{B}_t = 0 = \hat{B}_0$

□

Bem. (iv): Langzeitasymptotik der BB \leftrightarrow Lokale Regularität Brownscher Pf. te

Z. B.: Aus Stetigkeit in 0 folgt Gesetz der großen Zahlen:

$$\lim_{t \uparrow \infty} \frac{1}{t} B_t = \lim_{s \downarrow 0} s B_{\frac{1}{s}} = 0 \quad \text{P-f.s.}$$

$\underbrace{\hspace{10em}}_{= \hat{B}_s}$

ABER: Existenzbeweis für stetige BB ist schwieriger als direkter Beweis des GdGZ.

20. Martingale der Brownschen Bewegung

$(B_t)_{t \geq 0}$ eindim. Brownsche Bewegung auf (Ω, \mathcal{O}, P) .

20.1. Rechtsstetige Filtration

$\mathcal{F}_t^B = \sigma(B_s | 0 \leq s \leq t)$ ist Filtration

- Problem: $T := \inf \{t \geq 0 | B_t > c\}$ ist keine (\mathcal{F}_t^B) -Stoppzeit, da $\{T \leq t\} \notin \mathcal{F}_t^B$.

→ verwende stattdessen die rechtsstetige Filtration

$$\tilde{\mathcal{F}}_t := \mathcal{F}_{t+}^B = \bigcap_{\varepsilon > 0} \mathcal{F}_{t+\varepsilon}^B \quad (t \geq 0)$$

- Lemma $B_t - B_s \perp \mathcal{F}_s \quad \forall 0 \leq s < t$

Beweis: $B_t - B_s = \lim_{\substack{\varepsilon \downarrow 0 \\ c \in \mathbb{Q}}} (B_t - B_{s+c}) \quad P\text{-f.s.} \quad \square$
 $\perp \mathcal{F}_{s+c}^B \supseteq \mathcal{F}_s$

Hieraus folgt u.a. die Markoveigenschaft bzgl. $(\tilde{\mathcal{F}}_t)_{t \geq 0}$.

20.2. Brownsche Martingale

Def. $(M_t)_{t \geq 0}$ (Super-) Martingal bzgl. (\mathcal{F}_t)

- \Leftrightarrow (i) $M_t \in \mathcal{L}^1 \quad \forall t \geq 0$
- (ii) M_t \mathcal{F}_t -meßbar
- (iii) $E[M_t | \mathcal{F}_s] \stackrel{(\Leftrightarrow)}{=} M_s \quad \forall 0 \leq s \leq t$

Satz Folgende Prozesse sind Martingale bzgl. (\mathcal{F}_t^B) und bzgl. (\mathcal{F}_t) :

- 1) B_t
- 2) $B_t^2 - t$
- 3) $M_t^\lambda := \exp(\lambda B_t - \frac{\lambda^2 t}{2})$ für jedes $\lambda \in \mathbb{R}$
(exponentielles Martingal, Girsanovmartingal)

Beweis: 1), 3) Übung

- 2) $B_t^2 - t \in \mathcal{L}^1$ (da $B_t \sim N(0, t)$), \mathcal{F}_t -meßbar

zz: $E[B_t^2 - t | \mathcal{F}_s] = B_s^2 - s \quad \forall 0 \leq s \leq t$

d.h. $E[B_t^2 - B_s^2 | \mathcal{F}_s] = t - s$

$$= \underbrace{E[(B_t - B_s)^2 | \mathcal{F}_s]}_{\substack{\text{unabh.} \\ = E[(B_t - B_s)^2] = t - s}} + 2 B_s \cdot \underbrace{E[B_t - B_s | \mathcal{F}_s]}_{= 0} = t - s \quad \checkmark$$

□

Bem. 1) Spezielle Doob-Meyer-Zerlegung:

$$B_t^2 = M_t + t$$

Submartingal Martingal no. wachsend, adaptiert

$$2) M_t = B_t^2 - t = 2 \int_0^t B_s dB_s = 2 \lim_{n \rightarrow \infty} \sum_{t_i, t_{i+1}} B_{t_i} (B_{t_{i+1}} - B_{t_i})$$

Stochastisches Integral \uparrow previsibel

Speziellfall der Itô-Formel: $dB_t^2 = 2B_t dB_t + dt$

$$\text{Allgemein: } df(B_t) = \underbrace{f'(B_t) dB_t}_{\text{Martingalteil}} + \underbrace{\frac{1}{2} f''(B_t) dt}_{\text{Rest}}$$

$$df(B_t, t) = \frac{\partial f}{\partial x}(B_t) dB_t + \left(\frac{\partial f}{\partial t} + \frac{1}{2} \frac{\partial^2 f}{\partial x^2} \right) (B_t) dt$$

20.3. Harmonische Funktionen $(B_t)_{t \geq 0}$ Br. Bew. im \mathbb{R}^d

Satz $h: \mathbb{R}^d \rightarrow \mathbb{R}$ harmonisch (d.h. $\Delta h = 0$)

$\Rightarrow h(B_t)$ ist (\mathcal{F}_t) -Martingal, falls $e \in \mathcal{L}^1$

Beweis via Mittelwertigkeit (Übung) oder via Itô Formel (stoch. Analysis)

Allgemein: $h(x, t)$ Lösung von $\frac{1}{2} \Delta h + \frac{\partial h}{\partial t} = 0$ (durch Wärmeleitungsgl.)

$\xRightarrow{\text{u.g.v.}}$ $h(B_t, t)$ ist (\mathcal{F}_t) -Martingal. Beweis via Itô

Beispiel: $h(x,t) = x, |x|^2 dt, e^{\lambda \cdot x - \frac{1}{2}|\lambda|^2 t}$ ($\lambda \in \mathbb{R}^n$)

- Konsequenz:
- Stochastische Darstellung von Lösungen von PDgl
 - Analytische Berechnung von Erwartungswerten der B.B.

20.4. Stoppzeiten

- Def. 1) $T: \Omega \rightarrow [0, \infty]$ heißt (\mathcal{F}_t) -Stoppzeit
 falls $\{T \leq t\} \in \mathcal{F}_t \quad \forall t \geq 0$.

2) $\mathcal{F}_T := \{A \in \mathcal{O} \mid A \cap \{T \leq t\} \in \mathcal{F}_t\}$
 σ -Algebra der Informatio. bis zur Stoppzeit $T(\cdot)$

Eine ZV X ist \mathcal{F}_T -meßbar $\Leftrightarrow X \cdot I_{\{T \leq t\}}$ \mathcal{F}_t -meßbar $\forall t \geq 0$

20.5. Stoppzeit

Satz $(M_t)_{t \geq 0}$ Martingal bzgl. $(\mathcal{F}_t)_{t \geq 0}$, $t \mapsto M_t(\omega)$ ~~stetig~~ ^{P-f.s.} stetig,
 S, T beschränkte Stoppzeiten, $S \leq T$ P-f.s.

$\Rightarrow E[M_T | \mathcal{F}_S] = M_S$ P-f.s.

Beh. Insbesondere $E[M_T] = E[E[M_T | \mathcal{F}_0]] = E[M_0]$ für alle beschränkten Stoppzeit T . Für unbeschr. Stoppzeit gilt $E[M_{T \wedge n}] = E[M_0]$ (da $T \wedge n$ beschr. Stoppzeit). Daraus folgt $E[M_T] = E[M_0]$ falls $T < \infty$ P-f. und $M_{T \wedge n}$ glim.intbar.

Beweis Stoppzeit: Durch Reduktion auf diskrete Fall.

o B d A $t \mapsto M_t(\omega)$ stetig für alle ω (sonst Modifikation auf Nullmenge)

1) M_S ist \mathcal{F}_S -messbar:

$$\tilde{S}_n(\omega) = \sum_{k=0}^{\infty} \frac{k}{2^n} \cdot \mathbb{I}_{[\frac{k}{2^n}, \frac{k+1}{2^n})}(S(\omega)) \nearrow S(\omega)$$

↑
(i.a. hes. Stoppzeit \triangleright)

$\xRightarrow{M_t \text{ stetig}} M_{\tilde{S}_n} \rightarrow M_S$ ~~punktweise~~ punktweise

$M_{\tilde{S}_n}$ ist \mathcal{F}_S -messbar: $M_{\tilde{S}_n} \cdot \mathbb{I}_{\{S \leq t\}} = \sum_{k: \frac{k}{2^n} \leq t} \underbrace{M_{\frac{k}{2^n}}}_{\text{messbar bzgl. } \mathcal{F}_t} \cdot \underbrace{\mathbb{I}_{\{\frac{k}{2^n} \leq S < \frac{k+1}{2^n} \text{ und } S \leq t\}}}_{\in \mathcal{F}_t \text{ da Stoppzeit}}$

ist \mathcal{F}_t messbar für jedes $t \geq 0$.

2) Beh. $E[M_T; A] = E[M_S; A] \quad \forall A \in \mathcal{F}_S$

$$T_n = \sum_{k=1}^{\infty} \frac{k}{2^n} \cdot \mathbb{I}_{[\frac{k-1}{2^n}, \frac{k}{2^n})}(T) \quad \text{Sind beschränkte}$$

$$S_n = \sum_{k=1}^{\infty} \frac{k}{2^n} \cdot \mathbb{I}_{[\frac{k-1}{2^n}, \frac{k}{2^n})}(S) \quad \text{Stoppzeit } \triangleright \text{ (Übung)}$$

$S_n \leq T_n$ f.s., $T_n \downarrow T$, $S_n \downarrow S$

$A \in \mathcal{F}_S \stackrel{S \leq T_n \leq T}{\subseteq} \mathcal{F}_{T_n}$ (Übung: $S \leq T \Rightarrow \mathcal{F}_S \subseteq \mathcal{F}_T$)

$$\begin{aligned} \Rightarrow_{S_n, T_n \leq c} E[M_{T_n}; A] &= \sum_{k: \frac{k}{2^n} \leq c} E[M_{\frac{k}{2^n}}; \underbrace{A \cap \{T_n = \frac{k}{2^n}\}}_{\in \mathcal{F}_{\frac{k}{2^n}} \text{ da } A \in \mathcal{F}_{T_n}}] \\ &\stackrel{M_n \text{ bzgl. } \mathcal{F}_n}{=} \sum_k E[M_c; A \cap \{T_n = \frac{k}{2^n}\}] = E[M_c; A] = \dots = E[M_{S_n}; A] \end{aligned}$$

Für $n \rightarrow \infty$ folgt Beh. falls M_{T_n}, M_{S_n} gleichmäßig intbar.

Ok, da $M_{T_n} = E[M_c | \mathcal{F}_{T_n}]$, $M_{S_n} = E[M_c | \mathcal{F}_{S_n}]$ (siehe oben!). \square

20.6. Ruinproblem B_t B.B. im \mathbb{R}^1 mit Start 0 , $a, b \geq 0$

$T := \inf \{t > 0 \mid B_t \notin (-b, a)\}$

Satz 1) $P[T < \infty] = 1$ a-

2) $P[B_T = a] = \frac{b}{a+b}$, $P[B_T = -b] = \frac{a}{a+b}$ 0-

3) $E[T] = a \cdot b$ -b-

Beweis: 1) $P[T > k] \leq P[\forall n \leq k: |B_{nt} - B_n| < at] = \prod_{n=1}^k \underbrace{P[|B_{nt} - B_n| < at]}_{= \text{const.} < 1}$

$\rightarrow 0$ für $k \rightarrow \infty$
2) und 3) folgt aus Stokast. (für ...)

Bem. Dieselbe Aussage kann man auch mithilfe des Reflexionsprinzips (s.u.) beweisen.



20.8. Brownsche Bewegung mit Drift

$\Omega = C([0, \infty))$, $X_t(\omega) = \omega(t)$, $P =$ Wiener-Maß auf Ω (Start in 0)

$(X_t)_{t \geq 0}$ ist Brownsche Bewegung bzgl. P .

$Y_t = X_t + \lambda t$ Brownsche Bewegung mit Drift λ

$dY_t = dX_t + \lambda dt$ stochastische Differentialgleichung

- z.B. Kursentwicklung in der Börse, ...

Gesucht: $E[F(Y)]$, wobei $F: C([0, \infty)) \rightarrow \mathbb{R}$.

(analytische oder numerische Berechnung)

z.B. $P[\max_{0 \leq s \leq t} Y_s \geq c] \rightarrow F = I_{\{\max_{0 \leq s \leq t} Y_s \geq c\}}$

Die Wk. berechnbarkeit kann man exakt analytisch berechnen, s.u.

Trotzdem zur Illustration: Wie könnte man die Wk. numerisch berechnen?

Klassische Monte Carlo Methode versagt:

$$\begin{array}{l}
 \text{PI mit } (X_s + \lambda s) \geq c \text{] } \xrightarrow{t \rightarrow \infty} 0 \text{ exp. schnell falls } \lambda < 0 \\
 \text{0 \leq s \leq t} \qquad \qquad \qquad \searrow \text{ exp. schnell falls } \lambda > 0
 \end{array}$$

Idee: Importance Sampling bzgl. $P = W$ -Maß ∇

Satz Die Verteilung Q von $(Y_t)_{t \geq 0}$ auf $C([0, \infty))$ ist lokal äquivalent zu P , d.h. $Q \approx P$ auf

$$\mathcal{F}_t := \sigma(X_s \mid 0 \leq s \leq t) \text{ für alle } t \geq 0,$$

mit Dichte

$$\frac{dQ}{dP} \Big|_{\mathcal{F}_t} = e^{-\lambda X_t - \frac{\lambda^2}{2} t} = M_t^\lambda$$

Bem. Einfachstes Beispiel einer Gittertransformation. Allgemein:

$$dY_t = dX_t + B(t, Y_t) dt \text{ stoch. Differenzgleichung}$$

\leftrightarrow Maßtransformation auf $C([0, \infty))$

Beweis (für $t \leq 1$):

Sei $\mathcal{D}_n := \{t_k = \frac{k}{2^n} \mid 0 = k = 2^n\}$ n -te dyadische Partition von $[0, 1]$

Inkremente $\Delta_i Y = Y_{t_{i+1}} - Y_{t_i} = \Delta_i X + \lambda \Delta t$ unabh., $N(\lambda \Delta t, \Delta t)$

$$\begin{aligned} \Rightarrow P_0(Y_{t_1}, \dots, Y_{t_{2^n}})^{-1} &= \text{const.} \prod_{i=1}^{2^n-1} e^{-\frac{(\Delta_i X - \lambda \Delta t)^2}{2 \Delta t}} dx_i \dots dx_{2^n} \\ &= \text{const.} \cdot e^{-\sum \frac{(\Delta_i X - \lambda \Delta t)^2}{2 \Delta t}} dx_1 \dots dx_{2^n} \\ &= \underbrace{e^{\sum (\lambda \Delta_i X - \frac{\lambda^2}{2} \Delta t)}}_{= e^{\lambda X_{2^n} - \frac{\lambda^2}{2}}} \cdot \underbrace{\text{const.} \cdot e^{-\sum \frac{(\Delta_i X)^2}{2 \Delta t}}}_{= P_0(X_{t_1}, \dots, X_{t_{2^n}})^{-1}} dx_1 \dots dx_{2^n} \end{aligned}$$

$\Rightarrow Q \ll P$ auf $\mathcal{F}_{\mathcal{D}_n} = \sigma(X_t \mid t \in \mathcal{D}_n)$ mit Dichte

$$\frac{dQ}{dP} \Big|_{\mathcal{F}_{\mathcal{D}_n}} = e^{\lambda X_{t_{2^n}} - \frac{\lambda^2}{2}} = e^{\lambda X_1 - \frac{\lambda^2}{2}} = M_1$$

unabhängig von n , insbes. gleich. integr.

$\Rightarrow Q \ll P$ auf $\sigma(\cup_{n \in \mathbb{N}} \mathcal{F}_{\mathcal{D}_n}) = \mathcal{F}_1$ mit derselb. Dichte

$\Rightarrow Q \ll P$ auf \mathcal{F}_t ($t \leq 1$) mit Dichte $E[M_1^{\lambda} | \mathcal{F}_t] = M_t^{\lambda}$. \square

Bem. Heuristike à la Feynman:

$$dQ = \frac{1}{Z_\infty} e^{-\frac{1}{2} \int_0^1 |\dot{x}(t) - \lambda|^2 dt} \prod_{t \in [0,1]} dx(t) \quad \text{auf } C[0,1]$$

$x(0) = 0$

Damit informeller Beweis:

$$dQ = e^{\int_0^1 \lambda \dot{x}(t) dt - \frac{\lambda^2}{2}} \cdot \underbrace{\frac{1}{Z_\infty} e^{-\frac{1}{2} \int_0^1 |\dot{x}(t)|^2 dt} \prod_{t \in [0,1]} dx(t)}_{= dP}$$

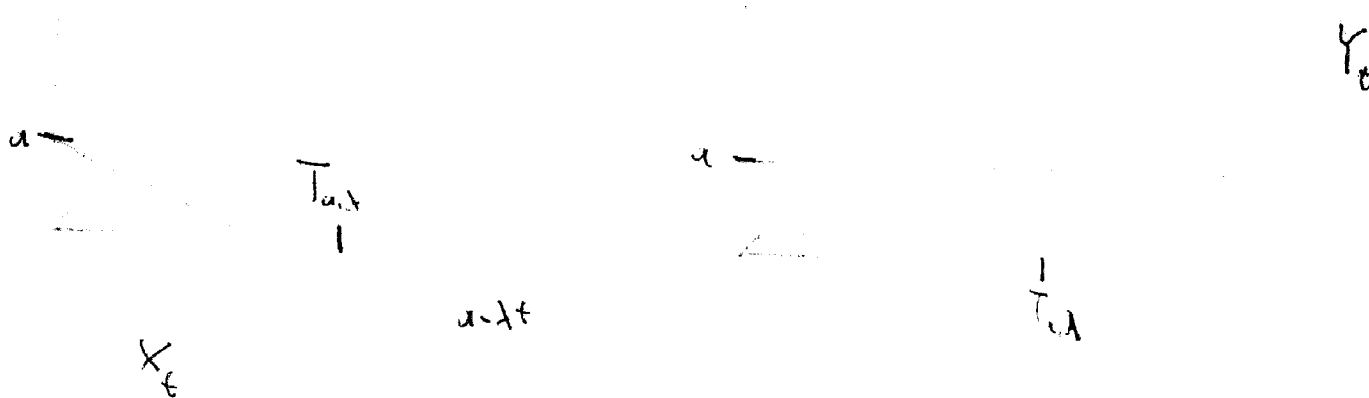
Rigorous Version: $dQ = e^{\int_0^1 \lambda dx(t) - \frac{\lambda^2}{2}} dP$, s.o.

stoch. Integral

20.9 Beispiel: Passierzeiten $Y_t = X_t + \lambda t, \quad a > 0$

Gesucht: Verteilung von

$$T_{a,\lambda} := \inf \{t > 0 \mid X_t > a - \lambda t\} = \inf \{t > 0 \mid Y_t > a\} \in [0, \infty)$$



$$P[T_{a,\lambda} \leq t] = Q[T_a \leq t] = E_P[M_t^\lambda; T_a = t]$$

Stoppzeit

$$= E_P[M_{T_a}^\lambda; T_a \leq t] = E_P[e^{\lambda T_a - \frac{1}{2} \lambda^2 T_a}; T_a = t]$$

$$= \int_0^t e^{\lambda a - \frac{1}{2} \lambda^2 s} f_{a,\lambda}(s) ds = \int_0^t f_{a,\lambda}(s) ds$$

Wahrscheinlichkeit
dass T_a <= t

Substituieren
mit T_a, etc.

wobei $f_{a,\lambda}(t) = \frac{a}{\sqrt{2\pi t^3}} e^{-\frac{(a-\lambda t)^2}{2t}}$

Insbesondere ergibt sich:

$$P[T_{a,\lambda} < \infty] = \int_0^\infty f_{a,\lambda}(t) dt \begin{cases} = 1 & \text{für } \lambda \geq 0 \\ < 1 & \text{für } \lambda < 0 \end{cases}$$

$\lambda > 0$

$\lambda < 0$

20.10. Satz von Cameron-Martin $h \in C([0,1])$

$$Y_t(\omega) = X_t(\omega) + h(t) \quad (0 \leq t \leq 1)$$

$$Q := P \circ Y^{-1} \text{ W.V. auf } C([0,1])$$

Satz $Q \ll P \iff h$ absolutstetige Fkt. mit $h' \in L^2([0,1], dt)$ und $h(0) = 0$

- In diesem Fall gilt:

$$\left\| \frac{dQ}{dP}(\omega) = \exp \left(\int_0^1 h'(t) dX_t(\omega) - \frac{1}{2} \int_0^1 |h'(t)|^2 dt \right) \right.$$

wobei $\int_0^1 h'(t) dX_t := \lim_{n \rightarrow \infty} \sum_{t_i \in \mathcal{D}_n} h'(t_i) \cdot (X_{t_{i+1}} - X_{t_i})$ in $L^2(\Omega, \mathcal{O}, P)$.

(stochastisches Integral)

Beweisskizze "⇐": Inkremente $\Delta Y = \Delta X + \Delta h \sim N(\Delta h, \Delta t)$

$$\left. \frac{dQ}{dP} \right|_{\mathcal{F}_{\mathcal{D}_n}} \stackrel{\text{wie zuvor}}{=} e^{\sum_{t_i \in \mathcal{D}_n} \frac{\Delta h}{\Delta t} \Delta X - \frac{1}{2} \sum_{t_i \in \mathcal{D}_n} \left(\frac{\Delta h}{\Delta t} \right)^2 \Delta t}$$

dyadische Ableitung

$n \rightarrow \infty$: h absolutstetig mit $h' \in L^2([0,1], dt) \implies \overbrace{\frac{dh}{dt}}^{\text{dyadische Ableitung}} \Big|_{\mathcal{D}_n} \xrightarrow{L^2([0,1])} h'$

$$\Rightarrow \sum_{t_i \in \mathcal{D}_n} \left(\frac{\Delta h}{\Delta t} \right)^2 \Delta t \rightarrow \int_0^1 |h'(t)|^2 dt$$

~~///~~

$M_n := \sum_{t_i \in \mathcal{D}_n} \frac{\Delta h}{\Delta t} \cdot \Delta X$ ist Cauchyfolge in $L^2(\Omega, \mathcal{O}_t, P)$, denn:

(i) Martingal bzgl. $(\mathcal{F}_{\mathcal{D}_n})_{n \in \mathbb{N}}$

(ii) L^2 -beschränkt:

$$\|M_n\|_2^2 = \sum \left(\frac{\Delta h}{\Delta t} \right)^2 \cdot \underbrace{E[(\Delta X)^2]}_{=\Delta t} \rightarrow \int_0^1 |h'(t)|^2 dt$$

□

20.11. Doob's L^p -Ungleichung

Satz $(M_t)_{t \geq 0}$ Martingal oder nicht-negatives Submartingal bzgl. $(\mathcal{F}_t)_{t \geq 0}$, rechtsstetig: $\lim_{t \downarrow s} M_t = M_s$.

Dann gilt für $p \geq 1$ und $\lambda > 0$:

1) $\lambda^p \cdot P[\sup_{0 \leq s \leq t} |M_s| \geq \lambda] \leq E[|M_t|^p] \quad \forall t \geq 0$

2) $\| \sup_{0 \leq s \leq t} |M_s| \|_{L^p} \leq \frac{p}{p-1} \cdot \|M_t\|_{L^p} \quad \forall t \geq 0$

Beweis: 1) $\mathcal{D}_n = \{t_n = \frac{n}{2^k} \cdot t \mid 0 \leq n \leq 2^k\}$

M_{t_n} ($n=0,1,2,\dots$) (Sub-) Martingal bzgl. (\mathcal{F}_{t_n})

\Rightarrow Diskrete L^p -Ungl. $\lambda^p \cdot P[\max_{s \in \mathcal{D}_n} |M_s| \geq \lambda] \leq E[|M_t|^p]$

$\rightarrow \sup_{s \in [0,t]} |M_s|$ da rechtsstetig

$\Rightarrow \lambda^p \cdot P[\sup_{s \in [0,t]} |M_s| \geq \lambda] \leq \lambda^p \cdot P[\exists n: \max_{s \in \mathcal{D}_n} |M_s| \geq \lambda]$

$= \lambda^p \cdot \lim_{n \rightarrow \infty} P[\max_{s \in \mathcal{D}_n} |M_s| \geq \lambda] \leq E[|M_t|^p]$

Folgt $\lambda \rightarrow \lambda - \epsilon, \epsilon \downarrow 0$

2) folgt aus 1) wie im diskreten Fall.

20.12. Submartingalkonvergenzsatz

Satz $(M_t)_{t \geq 0}$ rechtsstetiges Submartingal.

1) P-f.s. $\exists M_{t-} := \lim_{s \uparrow t} M_s \quad \forall t > 0$

(d.h. M_t ist "càdlàg" = continue à droite, limites à gauche)

2) $\sup_{t \geq 0} E[M_t^+] < \infty \Rightarrow$ P-f.s. $\exists \lim_{t \rightarrow \infty} M_t$

Beweis: Erweitere Upcrossing (oder downcrossing) Ungleichung vom diskreten auf den zeitstetigen Fall, dann Beweis wie im diskreten Fall (s. Übung). \square

21. Konstruktionen der Brownschen Bewegung

Drei Konstruktionen:

- Als Skalierungslimes von Random walks
- via Fortsetzungssatz von Kolmogorov
- Wiener-Lévy: Zufällige Überlagerung deterministischer Pfade

[Zufällige Fourierreihe (Wiener), Zufällige wavelet Entwicklung (Lévy)]

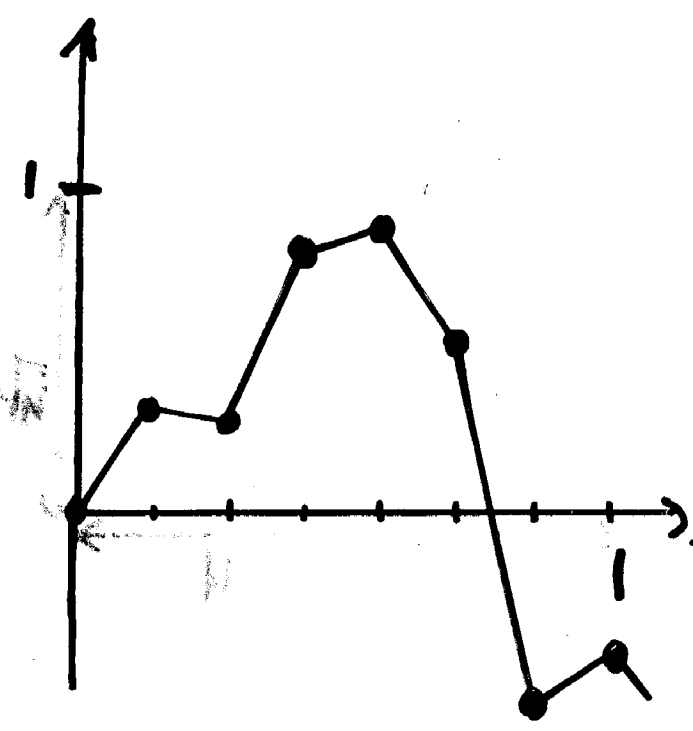
21.1. B.B. als Skalierungstimes von random walks

-Funktionaler Zentraler Grenzwertsatz

$$S_k = \sum_{i=1}^k X_i,$$

allgemeiner random walk
(z.B. $X_i = \pm 1$ mit $p = \frac{1}{2}$)

$$S_t^N(\omega) = \begin{cases} \frac{S_k(\omega)}{N} & \text{für } t = \frac{k}{N} \\ \text{linear dazwischen} \end{cases}$$



$\mu_N :=$ Verteilung von $(S_t^N)_{0 \leq t \leq 1}$ auf $C([0,1])$.

Satz (Invarianzprinzip von Donsker, Funktionaler ZGS)

μ_N konvergiert schwach gegen das Wienermaß P_0 .

d.h.

(205)

$$\int F(\omega) \mu_N(d\omega) \xrightarrow{NT_{\infty}} \int F(\omega) P_0(d\omega)$$

für alle stetigen, beschränkten $F: C([0,1]) \rightarrow \mathbb{R}$.

Bem. 1) „ZGS für die gesamten Trajektorien“

Klassischer ZGS:

Verteilung eines reskalierten r.w.
zu fester Zeit $\xrightarrow{w} N(0,1)$

Funktionale ZGS:

Verteilung der Trajektorien \xrightarrow{w} Gaußmaß auf (\mathbb{R}^{∞})

2) Universalität / Invarianzprinzip:

Jeder reskalierte r.w. verhält sich asymptotisch
wie eine B.B. (unabhängig von Verteilung der X_i) ∇

\uparrow
 \mathcal{L}^2

Beweis hat zwei Teile:

1) Konvergenz der Randverteilungen:

Verteilung von $(\xi_1^N, \dots, \xi_n^N) \xrightarrow{w} \text{multivar. Normalverteilung}$
 $N(\mu, \Sigma)$

ähnlich zum Beweis des ZGS, s. W'theorie I.

2) Relative Kompaktheit: (Jede Teilfolge von)

$(\mu_N)_{N \in \mathbb{N}}$ hat eine schwach konvergente Teilfolge.

Folgt aus

Satz von Prokhorov: Spohnisch (s. W'theorie I), $M \subseteq M_+(S)$.

M relativ kompakt $\iff M$ straff

(d.h. jede Folge in M hat eine in $M_+(S)$ schwach konvergente Teilfolge)

(d.h. $\forall \epsilon > 0 \exists K \subseteq S$ kompakt:
 $P(S \setminus K) < \epsilon$ für alle $P \in M$)
„Gleichmäßige Konzentration auf Kompakta“

Beweis von Polkorov s.e.B.:

Billingsley, Convergence of prob. measures, p. 35 ff.

Nachweis der Straffheit von $\{\mu_N | N \in \mathbb{N}\}$ via:

Satz von Arzela-Ascoli: Sei $K \subseteq C([0,1])$ abgeschl. Dann:

K kompakt \Leftrightarrow Die Funktionen in K sind gleichmäßig beschränkt und gleichmäßig gleichmäßig stetig.

(d.h. $\sup_{w \in K} \sup_{|t-s| < \delta} |w(t) - w(s)| \xrightarrow{\delta \rightarrow 0} 0$)

s. Funktionalanalysis.

ANWENDUNG von Donsker:

Aussagen für random walk \Rightarrow Aussagen für B.B.

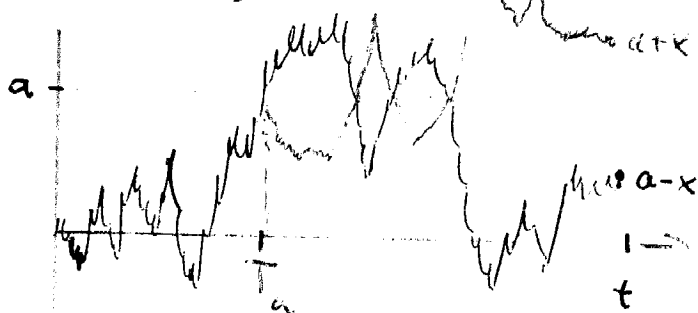
- z.B. - Verteilung des Maximums und der ersten Passierzeit
- Verteilung der letzten Passierzeit etc.

21.2. Maximum eines Brownschen Pfades

$(B_t)_{t \geq 0}$ Br. Bew. in \mathbb{R}^1 mit Start in 0, $a > 0$, $T_a = \min\{t \geq 0 \mid B_t \geq a\}$

Idee: Reflexionsprinzip

Reflektierter Pfad ist
"gleich wahrscheinlich"



Damit sollte gelten:

$$\underline{\underline{P[\max_{0 \leq s \leq t} B_s \geq a] = P[T_a \leq t]}}$$

← gleichwahrscheinlich →

$$= P[T_a \leq t \text{ und } B_t \geq a] + P[T_a \leq t \text{ und } B_t \leq a]$$

$$= 2 P[T_a \leq t \text{ und } B_t \geq a] = 2 P[B_t \geq a] = \underline{\underline{P[|B_t| \geq a]}}$$

Also:

Satz 1) $\max_{0 \leq s \leq t} B_s \sim |B_t|$

2) Die Verteilung von T_a ist absolutstetig mit Dichte

$$f_a(t) = \frac{a}{\sqrt{2\pi t^3}} e^{-\frac{a^2}{2t}} \quad (t > 0)$$

Beweis von 2): $P[T_a \leq t] = 2 P[B_t \geq a] = 2 \cdot P[B_t \geq \frac{a}{\sqrt{t}}] = \frac{2}{\sqrt{\pi}} \int_{\frac{a}{\sqrt{t}}}^{\infty} e^{-\frac{x^2}{2}} dx$

$$\Rightarrow f_u(t) = \frac{d}{dt} P[T_u \leq t] = \frac{a}{\sqrt{2\pi} t^3} e^{-\frac{a^2}{2t}} \quad \square$$

(20)

Rigorous Beweis des Reflektionsprinzips: 2 Möglichkeiten.

a) via starke Markoveigenschaft, s.u.

b) verwende Reflektionsprinzip für random walk + Donsker

2. KONSTRUKTION:

(210)

2.1.3. Der Fortsetzungssatz von Kolmogorov

$$\Omega = \{\omega: [0, \infty) \rightarrow \mathbb{R}\} = \mathbb{R}^{[0, \infty)}, \quad X_t(\omega) = \omega(t)$$

$$\mathcal{O} = \sigma(X_t \mid t \geq 0)$$

Ist P_0 das Wienermaß auf (Ω, \mathcal{O}) , dann gilt für $0 < t_1 < \dots < t_n$:

$$\begin{aligned} P_0(X_{t_1}, \dots, X_{t_n})^{-1} &= \prod_{i=1}^n P_{t_i - t_{i-1}}(x_i, x_{i-1}) dx_1 \dots dx_n \quad (t_0 = 0) \\ &= \prod_{i=1}^n \frac{1}{\sqrt{2\pi(t_i - t_{i-1})}} e^{-\frac{(x_i - x_{i-1})^2}{2(t_i - t_{i-1})}} dx_1 \dots dx_n \\ &=: \mu_{t_1, \dots, t_n} \quad \text{Randverteilungen} \end{aligned}$$

Umgekehrt: Konstruktion von P_0 aus Randverteilungen:

Satz von Kolmogorov Für jede konsistente Familie μ_{t_1, \dots, t_n} ($n \in \mathbb{N}$, $0 < t_1 < t_2 < \dots < t_n$) von Randverteilungen existiert genau eine WV P auf (Ω, \mathcal{O}) mit

$$P_0(X_{t_1}, \dots, X_{t_n})^{-1} = \mu_{t_1, \dots, t_n}$$

Beweis: rein mathematisch, s.z.B. Bauer: Theorie

Vorteil: Allgemeine Existenzsätze - Funktioniert für stochastische Prozesse (z.B. beliebige zeitstetige Markovprozesse) auf polnischen Räumen, stochastische Felder $(X_t, t \in \mathbb{R}^d)$, etc.

Problem: Keine unmittelbare Konstruktion der Trajektorien.

Regularität der Pfade?

21.4. Der Satz von Kolmogorov-Čentsov

$$D := \{k \cdot 2^{-n} \mid n \in \mathbb{N}, k = 0, 1, 2, \dots, 2^n\} \subseteq [0, 1]$$

Satz Gilt $E[|X_t - X_s|^\alpha] \leq C \cdot |t-s|^{1+\beta} \quad \forall t, s \in [0, 1]$ (*)

mit $\alpha, \beta, C \in (0, \infty)$, dann

$$\sup_{\substack{s, t \in D \\ s \neq t}} \frac{|X_t - X_s|}{|t-s|^\gamma} < \infty \quad \text{P-f.s.} \quad \forall \gamma < \frac{\beta}{\alpha} \quad (**)$$

(and $\mathbb{Q} \cap N, \mathbb{P} \cap N = \emptyset$).

Beweis: s.z.B. Bauer 39.3 oder Karatzas / Shreve 2.2.

Damit Konstruktion einer (Hölder-)stetigen Modifikation von $(X_t)_{t \geq 0}$:

$$\tilde{X}_t(\omega) := \begin{cases} \lim_{\substack{s \rightarrow t \\ s \in D}} X_s(\omega) & \text{für } \omega \in \Omega \cap N \text{ (existiert wg. (*))} \\ 0 & \text{für } \omega \in N \end{cases}$$

ist Hölderstetig mit Parameter α , $P[\tilde{X}_t = X_t] = 1 \quad \forall t \geq 0$.

(denn: $t_n \in D, t_n \rightarrow t \Rightarrow X_{t_n} \rightarrow X_t$ stochastisch wg. (*))
 $X_{t_n} \rightarrow \tilde{X}_t$ P-f.s. nach Def. von \tilde{X}_t
 $\Rightarrow X_t = \tilde{X}_t$ P-f.s.)

Anwendung auf Brownsche Bewegung: P via Kolmogorov

$$E[|X_t - X_s|^\alpha] = \text{const.} \cdot |t-s|^{\frac{\alpha}{2}} \quad \forall \alpha > 0, s \leq t$$

$\underbrace{\hspace{1cm}}_{\sim N(0, t-s)}$

d.h. (*) gilt mit $1 + \beta = \frac{\alpha}{2} \Leftrightarrow \frac{\beta}{\alpha} = \frac{\alpha-2}{2\alpha} \rightarrow \frac{1}{2}$ für $\alpha \rightarrow \infty$

21.5. Brownsche Bewegung als zufällige Überlagerung deterministischer Pfade (213)

$P_0 =$ Wienermaß auf $C([0,1])$ (Start in 0)

Heuristische Darstellung:

$$P_0(dw) = \frac{1}{Z_\infty} e^{-\frac{1}{2} \int_0^1 |\dot{w}(t)|^2 dt} \prod_{t \in (0,1)} dw(t), \quad w(0)=0$$

- d.h. P_0 ist „Standardnormalverteilung“ auf dem H -Raum

$$H_{ch} = \{h \in C([0,1]) \mid h(0)=0, \text{ absolutstetig mit } h' \in L^2(0,1)\}$$

$$\langle h, g \rangle_{ch} = \int_0^1 h'(t) g'(t) dt$$

(Cameron-Martin-Raum)

- Damit Ansatz zur Konstruktion einer Zufallsvariable

$$B: \Omega \rightarrow C([0,1])$$

mit Verteilung P_0 (d.h. einer Brownschen Bewegung):

$$B(\omega) = \sum_{n=1}^{\infty} Y_n(\omega) e_n$$

wobei $\{e_n \mid n \in \mathbb{N}\}$ vollständige Orthonormalbasis von H_{ch} ,

$Y_n: \Omega \rightarrow \mathbb{R}$ ($n \in \mathbb{N}$) i.i.d. $N(0,1)$ -verteilte ZV auf dem Wiener-Wahrsch. (Ω, \mathcal{F}, P) .

$$(*) \quad B_t(\omega) = \sum_{n=1}^{\infty} Y_n(\omega) e_n(t) \quad (0 \leq t \leq 1, \omega \in \Omega) \quad (214)$$

PROBLEM:

Satz H unendlichdimensionaler Hilbertraum mit Orthonormalbasis $\{e_n | n \in \mathbb{N}\}$

$$\Rightarrow P \left[\sum_{n=1}^{\infty} Y_n(\omega) e_n \text{ konvergiert in } H \right] = 0$$

Beweis: $\left\| \sum_{n=1}^{\infty} Y_n(\omega) e_n \right\|_H^2 = \sum_{n=1}^{\infty} |Y_n(\omega)|^2 \xrightarrow[n \rightarrow \infty]{\text{a. S.}} \infty \quad P\text{-f.s.} \quad \square$

AUSWEG:

Die Reihe in $(*)$ konvergiert nicht bzgl. $\|\cdot\|_{CH}$, aber bzgl. der schwächeren Supremumsnorm!

Satz (Wiener, Lévy, Hô-Nisio)

Die Reihe in $(*)$ konvergiert P -f.s. im Banachraum $C([0,1]) \cong H_{CH}$, und der Grenzwert $B = (B_t)_{t \in [0,1]}$ ist eine Br. Bewegung auf $(\mathbb{R}, \mathcal{O}, P)$.

Zwei wichtige Spezialfälle:

Wiener: Zufällige Funktion $e_0(t) = t, e_n(t) = \frac{\sqrt{2}}{\pi n} \sin(\pi n t)$

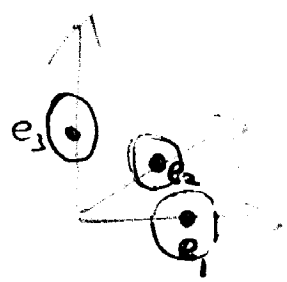
Lévy: Zufällige Wavelentwicklung, s.u.

Beweis im 2. Fall s.u.

Bem. Eine Stochastikbewertung auf einem ∞ -dimensionalen Hilbertraum H kann man stets nur auf einen echt größeren Banach- oder Hilbertraum $E \supseteq H$ realisieren!

(Beweisstrategie: $1 = \mu(H) \geq \sum_{k=1}^{\infty} \mu(B(e_k, \frac{1}{2})) \stackrel{\text{Symmetrie}}{=} \infty \cdot \mu(B(e_1, \frac{1}{2}))$
disjunkte Bälle

$\Rightarrow \mu(B(e_1, \frac{1}{2})) = 0$

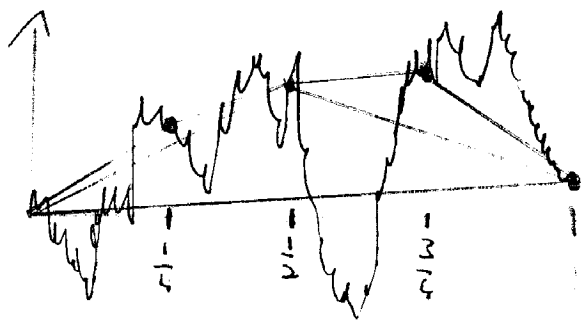


Ähnlich folgt $\mu(B(h, \frac{1}{2} \|h\|)) = 0$ für alle $h \in H$,
und hieraus kann man $\mu \equiv 0$ schließen \leftarrow

21.6. Wiener-Lévy-Darstellung der Br. Bewegung

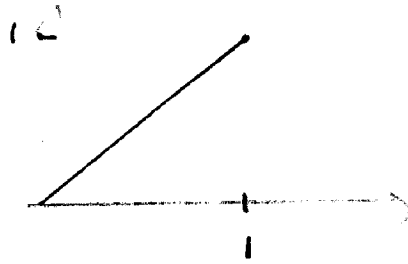
$P_0 =$ Wienermaß auf $([0,1])$

Idee: polygonale Approx. v. Brownsche Pfaden bzgl. dyad. Partitionen

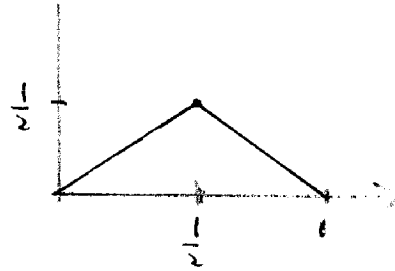


Schäuderfunktionen ("wavelets"):

$$e(t) = t$$

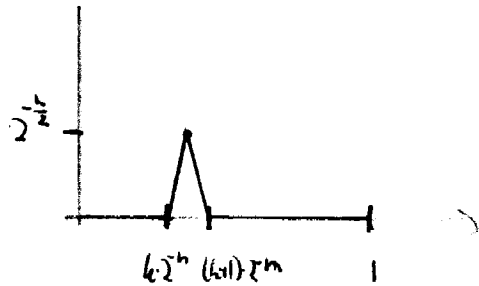


$$e_{0,0}(t) = t \wedge (1-t) = \begin{cases} t & \text{für } t \leq \frac{1}{2} \\ 1-t & \text{für } t > \frac{1}{2} \end{cases}$$



$$e_{n,k}(t) = 2^{-\frac{n}{2}} e_{0,0}(2^n t - k)$$

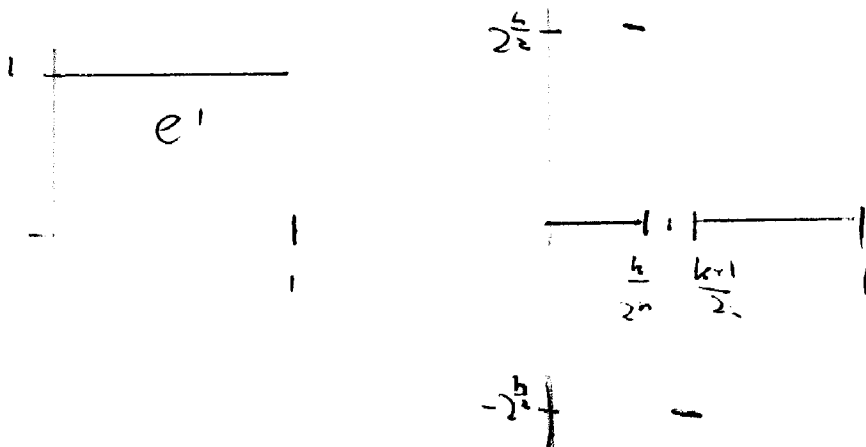
$$(n \in \mathbb{N}, k = 0, 1, \dots, 2^n - 1)$$



Lemma Die Schäuderfunktionen bilden eine vollständige Orthonormalbasis im Cameron-Martin-Raum H_{CM} .

Beweis: $\mathcal{J}: H_{CM} \rightarrow L^2(0,1)$ ist Isometrie.
 $h \mapsto h'$

Die Haarfunktionen $e'(t), e'_{n,k}(t)$ bilden eine vollst. ONB in $L^2(0,1)$. \square



Satz Sei $x \in C([0,1])$, $\Delta_{n,k} x := \left(x\left(\frac{k+1}{2^n}\right) - x\left(\frac{k+\frac{1}{2}}{2^n}\right) \right) - \left(x\left(\frac{k+\frac{1}{2}}{2^n}\right) - x\left(\frac{k}{2^n}\right) \right)$ (217)

Dann gilt:

$$(*) \quad x(t) = x(1) \cdot e(t) + \sum_{n=0}^{\infty} \sum_{k=0}^{2^n-1} 2^{\frac{n}{2}} \Delta_{n,k} x \cdot e_{n,k}(t)$$

mit

a) gleichmäßiger Konvergenz für alle $x \in C([0,1])$

b) Konvergenz bzgl. $\|\cdot\|_{CH} \quad \forall x \in H_{CH}$

- Beweis: a) $x(1) \cdot e(t) + \sum_{n=0}^m \sum_{k=0}^{2^n-1} 2^{\frac{n}{2}} \Delta_{n,k} x \cdot e_{n,k}(t)$ ist die polygonale Approximation von x bzgl. der m -ten dyad. Zerlegung von $[0,1]$

b) $(*) \Leftrightarrow x = (x, e) e + \sum_{n,k} (x, e_{n,k}) e_{n,k}$ Entwicklung in ONB. \square

Kor. $(B_t)_{t \in [0,1]}$ Brownsche Bewegung

$$\Rightarrow \boxed{B_t(\omega) = Y(\omega) \cdot e(t) + \sum_{n=0}^{\infty} \sum_{k=0}^{2^n-1} Y_{n,k}(\omega) e_{n,k}(t) \quad (**)}$$

mit gleichmäßiger Konvergenz in t für P-f.a. ω ,

wobei $Y = B_1$, $Y_{n,k} = 2^{\frac{n}{2}} \Delta_{n,k} B$ i.i.d., $\sim N(0,1)$.

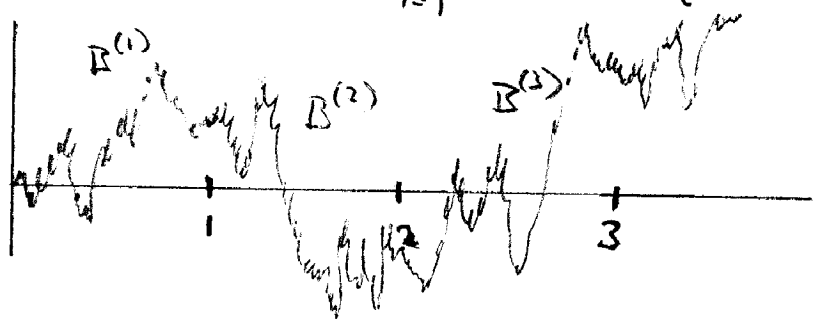
Umgekehrt kann man $(**)$ zur Konstruktion einer B.B. nutzen:

21.7. Lévy's Konstruktion der Brownschen Bewegung

Satz (Lévy 1948) Seien $Y, Y_{n,k}$ ($n \geq 0, k = 0, 1, \dots, 2^n - 1$) i.i.d. $\sim N(0,1)$ auf einem W'raum (Ω, \mathcal{O}, P) . Dann konvergiert die Reihe in (**)
P-f.s. gleichmäßig, und der Grenzwert $(B_t)_{t \in [0,1]}$ ist eine B.B.

Bem. $B_t^{(1)}, B_t^{(2)}, \dots$ unabh. Br. Bewegungen mit Start in 0 für $t \in [0,1]$

$\Rightarrow B_{ns+t} = \sum_{i=1}^n B_t^{(i)} + B_t^{(n+1)}$ definiert BB auf $[0, \infty)$



Lévy's Lévy: ① Gleichmäßige Konvergenz für P-f.a.w

via Borel-Cantelli + Weierstrass

$$\left| \sum_{k=0}^{2^n-1} Y_{n,k}(\omega) e_{n,k}(t) \right| \leq 2^{-\frac{n}{2}} \cdot \underbrace{\max_{0 \leq k < 2^n} |Y_{n,k}(\omega)|}_{=: M_n(\omega)}$$

$\neq 0$ auf disjunkten Intervallen, $\leq 2^{-n/2}$

Also glm. Konvergenz der Reihe in (**) falls $\sum 2^{-\frac{n}{2}} M_n(\omega) < \infty$.

ok falls $M_n(\omega) \leq n$ schließt sich.

219

$$\begin{aligned} P[M_n > n] &\leq 2^n \cdot P[|Y_{n,0}| > n] \leq \sqrt{\frac{2}{\pi}} \frac{2^n}{n} e^{-\frac{n^2}{2}} \\ &\leq \frac{2}{n} E[|Y_{n,0}| \mathbb{1}_{|Y_{n,0}| > n}] \quad \text{summierbare Reihe} \\ &= \frac{2}{n} \frac{1}{\sqrt{2\pi}} \int_n^\infty x e^{-\frac{x^2}{2}} dx = \frac{1}{n} \sqrt{\frac{2}{\pi}} e^{-\frac{n^2}{2}} \end{aligned}$$

Borel-Cantelli:

$$\Rightarrow P[M_n \leq n \text{ schließt sich}] = 1$$

$$\Rightarrow B_t^{(m)}(\omega) = Y(\omega)e(t) + \sum_{n=0}^m \sum_{k=0}^{2^n-1} Y_{n,k}(\omega) e_{n,k}(t), \quad m \in \mathbb{N},$$

P-f.s. gleichmäßig konvergent

$$B_t(\omega) := \lim_{m \rightarrow \infty} B_t^{(m)}(\omega)$$

② L^2 Konvergenz für t fest $l < m$

$$\|B_t^{(m)} - B_t^{(l)}\|_{L^2}^2 = E \left[\left(\sum_{n=l+1}^m \sum_{k=0}^{2^n-1} Y_{n,k} e_{n,k}(t) \right)^2 \right]$$

$$\underbrace{Y_{n,k} \text{ i.i.d.}} \sum_{n=l+1}^m \sum_k e_{n,k}(t)^2 \rightarrow 0 \quad \text{da } \sum_{n=1}^{\infty} \sum_k e_{n,k}(t)^2 < \infty \text{ s.u.}$$

$$\Rightarrow B_t^{(m)} \text{ Cauchy in } L^2 \stackrel{\text{P-f.s. konv.}}{\Rightarrow} B_t^{(m)} \rightarrow B_t \text{ in } L^2$$

③ Erwartungswert und Kovarianz $0 \leq s, t \leq 1$

$$\underline{E[B_t]} \stackrel{L^2\text{-Konv.}}{=} \lim_{m \rightarrow \infty} E[B_t^{(m)}] = \underline{0}$$

Linearkomb. von $N(0,1)$ -verteilten ZV

$$\underline{\text{Cov}(B_t, B_s)} = E[B_t B_s] \stackrel{L^2\text{-Konv.}}{=} \lim_{m \rightarrow \infty} E[B_t^{(m)} B_s^{(m)}]$$

$$= e(t)e(s) + \sum_{h=0}^{\infty} \sum_{k=0}^{2^n-1} e_{n,h}(t) e_{n,h}(s)$$

$$Y_{n,h} \text{ iid } \sim N(0,1)$$

$$\Gamma_{e_{n,h}}(s) = \int_0^s e_{n,h}' \cdot I_{(0,s)} = (e_{n,h}, h_s), \quad h_s(t) := \int_0^t I_{(0,s)} = t \wedge s$$

$$e \cdot (s) = (e \cdot, h_s)$$

$$= (e, h_t)(e, h_s) + \sum_{h,k} (e_{n,h}, h_t)(e_{n,h}, h_s)$$

$$\stackrel{\text{Parseval}}{=} (h_t, h_s) = \int_0^t h_t' h_s' = \int_0^t I_{(0,t)} I_{(0,s)} = \underline{t \wedge s}$$

$$e_{n,h} \text{ ONB}$$

④ $(B_t)_{t \in [0,1]}$ ist ein Gaußprozess $0 \leq t_1 < \dots < t_\ell \leq 1, p_1, \dots, p_\ell \in \mathbb{R}$

$$\sum_{j=1}^{\ell} p_j B_{t_j} = \lim_{m \rightarrow \infty} \sum_{j=1}^{\ell} p_j B_{t_j}^{(m)}$$
 normalverteilt

Linearkomb. von unabh. $N(0,1)$ verteilten ZV \rightarrow also normalverteilt

denn f.s. Limiten von normalverteilten ZV sind wieder normalverteilt falls Limes ex. und $\in \mathcal{L}^2$ (Übung, via char. Fktn.)

19.11
 $\implies (B_{t_1}, \dots, B_{t_n})$ multivariat normalverteilt

Nach ③ und ④ ist $(B_t)_{t \in [0,1]}$ B.B., s. 19.11.

□

Bem. $X_n \sim N(\mu_n, \sigma_n^2)$, $X_n \rightarrow X$ P-f.s.

$$\implies E[e^{ipX}] \stackrel{\text{Lebesgue}}{=} \lim_n E[e^{ipX_n}] = \lim_n e^{ip\mu_n} e^{-\frac{1}{2}\sigma_n^2 p^2}$$

Der Limes rechts existiert nur dann für alle p , wenn $\sigma_n \rightarrow \infty$, oder $\sigma = \lim \sigma_n$ und $\mu = \lim \mu_n$ in \mathbb{R} existieren. Der Limes ist genau dann die Fouriertransf. einer L^2 -Zufallsvariable wenn $\sigma < \infty$. In diesem Fall gilt $X \sim N(\mu, \sigma^2)$.

22. (I-) Regularität Brownscher Pfade

$(B_t)_{0 \leq t \leq 1}$ indim. B.B. auf $(\mathbb{R}, \mathcal{O}, P)$.

22.1. Typische Brownsche Pfade sind nirgends differenzierbar

Satz (Paley, Wiener, Zygmund 1933)

P-f.s. ist $t \mapsto B_t$ nirgends diffbar. Sogar:

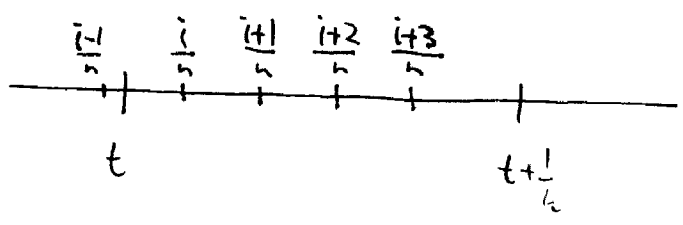
$$\overline{\lim}_{s \downarrow t} \frac{|B_s - B_t|}{|s - t|} = \infty \quad \forall t \in [0, 1) \text{ P-f.s.}$$

Beweis: zz:

$N = \{ \omega \in \Omega \mid \exists t \in [0, 1], k, L \in \mathbb{N} : |B_s - B_t| \leq L \cdot |s - t| \quad \forall s \in (t, t + \frac{1}{k}) \}$
ist Nullmenge.

Angenommen $|B_s(\omega) - B_t(\omega)| \leq L \cdot |s - t|$ für $s \in (t, t + \frac{1}{k})$.

Sei $n \in \mathbb{N}, n > 4k$, und $i \in \mathbb{N}$ sodaf:



⇒ Für $j=i, i+1, i+2$:

$$\begin{aligned} |B_{\frac{j+1}{n}} - B_{\frac{j}{n}}| &\leq |B_{\frac{j+1}{n}} - B_t| + |B_t - B_{\frac{j}{n}}| \\ &\leq L \cdot \left(\underbrace{\left(\frac{j+1}{n} - t \right)}_{\leq \frac{4}{n}} + \underbrace{\left| t - \frac{j}{n} \right|}_{\leq \frac{4}{n}} \right) \leq \frac{8L}{n} \end{aligned}$$

$$\begin{aligned} P\left[\forall j=i, i+1, i+2: |B_{\frac{j+1}{n}} - B_{\frac{j}{n}}| \leq \frac{8L}{n} \right] &\stackrel{\text{unabh. Inkrem.}}{=} P\left[|B_t| \leq \frac{8L}{\sqrt{n}} \right] \stackrel{\sim N(0; \text{Diffus } \frac{1}{n})}{=} \\ &=: A_{i,n,L} \leq \left(\frac{1}{\sqrt{2\pi}} \frac{16L}{\sqrt{n}} \right)^2 = \text{const.} \frac{L^2}{n^{3/2}} \end{aligned}$$

$\omega \in N \Rightarrow \exists L \in \mathbb{N} \exists k \in \mathbb{N} \forall n > 4k \exists i=0,1,\dots,n-1: \omega \in A_{i,n,L}$

$$P\left[\bigcap_{n > 4k} \bigcup_{i=0}^{n-1} A_{i,n,L} \right] \leq \lim_{n \rightarrow \infty} P\left[\bigcup_{i=0}^{n-1} A_{i,n,L} \right] = 0$$

$\leq c \cdot \frac{nL^2}{n^{3/2}}$

Also ist N Nullmenge. \square

22.2. Hölderstetigkeit

Satz \Rightarrow Es gilt P-f.s.:

1) Für alle $\alpha > \frac{1}{2}$:

$$\overline{\lim}_{s \downarrow t} \frac{|B_s - B_t|}{|s-t|^\alpha} = \infty \quad \forall t \in [0,1] \quad \text{"nirgends h\u00f6lderstetig"}$$

- 2) Für alle $\alpha < \frac{1}{2}$:

$$\sup_{\substack{s, t \in [0,1] \\ s \neq t}} \frac{|B_s - B_t|}{|s-t|^\alpha} < \infty \quad \text{"h\u00f6lderstetig auf [0,1]"}$$

Beweis: 1) Analog zu 22.1 - \u00dcbung.

2) Wiener-L\u00e9vy:

$$B_t = Y_0 \cdot t + \sum_{n=0}^{\infty} \sum_{k=0}^{2^n-1} Y_{n,k} \underbrace{e_{n,k}}_{\text{Schauderfkt.}}(t) \quad \forall t \in [0,1] \text{ P-f.s.}$$

mit $Y_0, Y_{n,k}$ i.i.d., $\sim N(0,1)$

$$\Rightarrow |B_t - B_s| \leq |Y_0| \cdot |t-s| + \sum_n \max_k |Y_{n,k}| \cdot \sum_k |e_{n,k}(t) - e_{n,k}(s)|$$

$= M_n(\omega) \leq C(\omega) \cdot n,$ höchstens zwei Summande $\neq 0$
 s. 21.7, $C < \infty$ P-f.s.

- $|e_{n,k}(t) - e_{n,k}(s)| \leq 2 \cdot \frac{1}{2} 2^{-\frac{n}{2}} \quad (\text{da } |e_{n,k}(t)| \leq \frac{1}{2} \cdot 2^{-\frac{n}{2}})$
- $|e_{n,k}(t) - e_{n,k}(s)| \leq \int_s^t |e'_{n,k}| \leq 2^{\frac{n}{2}} \cdot |t-s| \quad (\text{da } |e'_{n,k}| \leq 2^{\frac{n}{2}})$

Wähle N mit $2^{-N} \leq |t-s| < 2^{-(N+1)}$

$$\Rightarrow |B_t - B_s| \leq |Y_0| \cdot |t-s| + C \cdot \left(\sum_{n=1}^N n 2^{\frac{n}{2}} \cdot |t-s| + \sum_{n=N+1}^{\infty} n 2^{-\frac{n}{2}} \right)$$

$\leq |t-s|^\alpha \quad \leq N^2 \cdot 2^{\frac{N}{2}} \cdot |t-s| \quad \leq \text{const.} \cdot 2^{-\alpha N} \quad (\alpha > \frac{1}{2})$
 $\leq \text{const.} \cdot |t-s|^\alpha \quad \leq 2^\alpha \text{const.} \cdot |t-s|^\alpha$

□

Hölderstetig für $\alpha = \frac{1}{2}$? Nein, s. Satz von it. Log., s.u. (22.4)

22.3. Quadratische Variation

226

$$\tau = \{0 = t_0 < t_1 < \dots < t_k = 1\} \text{ Partition von } [0,1]$$

$$|\tau| = \max_i |t_{i+1} - t_i| \text{ Feinheitsgrad}$$

Sei $x: [0,1] \rightarrow \mathbb{R}$.

Definition $V_p(x) := \sup_{\substack{\tau \text{ Part.} \\ \text{von } [0,1]}} \sum_i |x(t_{i+1}) - x(t_i)|^p$

heißt klassische p -Variation von x ($1 \leq p < \infty$).

Beispiel: $x \in C^1 \Rightarrow x$ hat beschr. Variation $V_1(x) = \int_0^1 |x'(u)| du$

Ist x eine beliebige Funktion mit beschränkter Variation (d.h. $V_1(x) < \infty$),
und τ_n eine Folge von Partitionen mit $|\tau_n| \rightarrow 0$, dann gilt:

$$\sum_{t_i \in \tau_n} |x(t_{i+1}) - x(t_i)| \xrightarrow{n \rightarrow \infty} V_1(x)$$

1-Variation entlang τ_n

ABER:

$$\sum_{t_i \in \tau_n} |x(t_{i+1}) - x(t_i)|^2 \leq \max_{t_i \in \tau_n} |x(t_{i+1}) - x(t_i)| \cdot V_1(x) \rightarrow 0 \neq V_2(x)$$

quadr. Variation entlang τ_n

Satz (P. Lévy) $(B_t)_{t \in [0,1]}$ Brownsche Bewegung im \mathbb{R}^1

227

Dann gilt für alle $t \in [0,1]$:

$$S_n := \sum_{\substack{t_i \in \tau_n \\ t_i \leq t}} (B_{t_{i+1}} - B_{t_i})^2 \xrightarrow{n \uparrow \infty} t \text{ in } \mathcal{L}^2(\mathbb{R}^d, \mathcal{P}) \text{ und P-f.s.}$$

- quadrat. Variation von $(B_s)_{0 \leq s \leq t}$ bzgl. τ_n

Insbesondere ist ein typischer Brownscher Pfad nicht von beschr. Variation.
(Dies folgt auch aus 22.1).

WARNUNG: klassische quadrat. Variation: $V_2(B) = \infty$ P-f.s.

Hingegen gilt $V_p(B) < \infty$ P-f.s. für alle $p > 2$.
(folgt aus Hölderstetigkeit für $\alpha < \frac{1}{2}$)

Beweis Satz: 1) \mathcal{L}^2 Konvergenz: analog zu schwachem Gdgt

$$E[S_n] = \sum_{\substack{t_i \leq t \\ t_i \in \tau_n}} \underbrace{E[(B_{t_{i+1}} - B_{t_i})^2]}_{= t_{i+1} - t_i} \rightarrow t \text{ für } n \rightarrow \infty$$

$$\text{var}(S_n) \stackrel{\text{unabh. Inkrem.}}{=} \sum_{\substack{t_i \in \mathcal{T}_n \\ t_i \leq t}} \text{var}((B_{t_{i+1}} - B_{t_i})^2)$$

$$= \sum \left(\underbrace{E[(\Delta B)^4]} - (\Delta t)^2 \right)$$

$$= \frac{1}{\sqrt{2\pi}\Delta t} \int x^4 e^{-\frac{x^2}{2\Delta t}} dx \stackrel{\text{p.w. Wt.}}{=} \frac{\Delta t}{\sqrt{2\pi}\Delta t} \int x^2 e^{-\frac{x^2}{2\Delta t}} dx$$

$$= 3(\Delta t)^2$$

$$= 2 \sum_{\substack{t_i \in \mathcal{T}_n \\ t_i \leq t}} (t_{i+1} - t_i)^2 \rightarrow 0$$

(da $x(t) = t$ von beschr. Variation)

$$\Rightarrow S_n - E[S_n] \rightarrow 0 \text{ in } \mathcal{L}^2 \Rightarrow S_n \rightarrow t \text{ in } \mathcal{L}^2$$

2) P-f.s. Konvergenz: analog zu starker GdgtZ

$$\mathcal{O}_n := \sigma((B_{t_{i+1}} - B_{t_i})^2 \mid t_i \in \mathcal{T}_n \text{ für } a_n \leq n) \text{ absteigende Filtration}$$

$$\text{oBdA } t \in \mathcal{T}_n \Rightarrow E[B_t^2 \mid \mathcal{O}_n] = \sum_{\substack{t_i \in \mathcal{T}_n \\ t_i \leq t}} (B_{t_{i+1}} - B_{t_i})^2 \text{ (Übung)}$$

Rückwärtsmartingalkonvergenzsatz \Rightarrow P-f.s. konvergent. \square

22.4. Satz vom iterierten Logarithmus

$(B_t)_{t \geq 0}$ B.B. in \mathbb{R}^1

$$E[(B_{s+t} - B_s)^2] = t$$

also im L^2 Sinn $B_{s+t} - B_s \sim t^{1/2}$

Präzise pfadweise Asymptotik für $t \downarrow 0$?

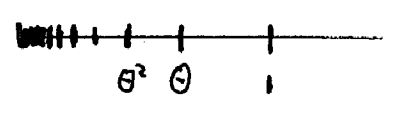
o BdA $s=0$, $B_0=0$ (da zeitlich und räumlich homogen)

Ziel: Finde $h(t)$ mit

$$\overline{\lim}_{t \downarrow 0} \frac{B_t(\omega)}{h(t)} = 1 \text{ für P-f.a. } \omega$$



Untere Schranke: $\theta \in (0,1)$, $t_n := \theta^n \downarrow 0$



Wenn

$$(*) \quad B_{t_n} - B_{\underbrace{t_{n+1}}_{=\theta t_n}} \geq \sqrt{1-\theta} h(t_n) \text{ unendlich oft}$$

dann

$$\overline{\lim}_{t \downarrow 0} \frac{B_t - B_{\theta t}}{h(t)} \geq \sqrt{1-\theta}$$

Wann gilt $(*)$?

$$P[B_{t_n} - B_{t_{n+1}} \geq \sqrt{1-\theta} \cdot h(t_n)]$$

$$= P \left[\underbrace{\frac{B_{t_n} - B_{t_{n+1}}}{\sqrt{t_n - t_{n+1}}}}_{\sim N(0,1)} \geq \sqrt{1-\theta} \underbrace{\frac{h(t_n)}{\sqrt{t_n - t_{n+1}}}}_{\frac{\theta^n}{\theta^{n+1}}} \right]$$

$$= \frac{h(t_n)}{\sqrt{t_n}} =: a_n$$

$$= \frac{1}{\sqrt{2\pi}} \int_{a_n}^{\infty} e^{-\frac{x^2}{2}} dx \geq \frac{1}{\sqrt{2\pi}} \frac{a_n}{1+a_n^2} e^{-\frac{a_n^2}{2}}$$

N.B. $\frac{d}{dx} \left(-\frac{x e^{-\frac{x^2}{2}}}{1+x^2} \right) = \left(\frac{2x^2}{(1+x^2)^2} - \frac{1-x^2}{1+x^2} \right) e^{-\frac{x^2}{2}} < e^{-\frac{x^2}{2}}$
 $= \frac{x^4 + 2x - 1}{(1+x^2)^2} < 1$

also $\int_a^{\infty} e^{-\frac{x^2}{2}} dx > \int_a^{\infty} \frac{d}{dx} (\dots) dx = \frac{a e^{-\frac{a^2}{2}}}{1+a^2}$

Nach Borel-Cantelli II gibt (*) falls

$$\sum_{n=1}^{\infty} \frac{a_n e^{-\frac{a_n^2}{2}}}{1+a_n^2} = \infty$$

ok falls $a_n = \sqrt{2(\log n + c)}$ für ein $c \in \mathbb{R}$ (denn $\sum \frac{1}{n \log n} = \infty$ nach Integraltest)

$$a_n = \frac{h(\Theta^n)}{\sqrt{\Theta^n}} \Rightarrow h(\Theta^n) \stackrel{!}{=} \sqrt{\Theta^n} \cdot \sqrt{2 \cdot (\log n + c)}$$

$$h(t) = ?$$

$$t = \Theta^n \Rightarrow n = \frac{\log t}{\log \Theta} = \frac{\overbrace{\log t}^{-1}}{\underbrace{\log \Theta}^{-1}}$$

$$\Rightarrow \log n = \log \log t^{-1} - \underbrace{\log \log \Theta^{-1}}_{=: c}$$

Also Bedingung erfüllt für

$$h(t) = \sqrt{2t \log \log \frac{1}{t}}$$

Satz (Khintchine 1933, zuvor Gesetze vom it. Log. für random walk: Kolmogorov 1929)

Für jedes fixe $s \geq 0$ gilt:

$$(i) \quad \overline{\lim}_{t \rightarrow 0} \frac{B_{s+t} - B_s}{h(t)} = +1 \quad \text{P-f.s.}$$

$$(ii) \quad \underline{\lim}_{t \rightarrow 0} \frac{B_{s+t} - B_s}{h(t)} = -1 \quad \text{P-f.s.}$$

Beweis: s.u. (22.6.)

Bem. ABER: Für P-f.a. ω ist das Gesetz vom iterierten Log. für ein $s > 0$ verletzt! Stattdessen gilt global:

Lévy's Stetigkeitsmodul:

$$(*) \quad \overline{\lim}_{s \downarrow 0} \frac{\sup_{\substack{t \in [0, s] \\ s \in [0, 1]}} |B_{s+t} - B_s|}{\sqrt{2s \log s^{-1}}} = +\infty \quad \text{P-f.s.}$$

Einfaches Logarithmus !!

Beweis: s. Revuz/Yor: Cont. martingales and B.M., (2.7)

oder Karatzas/Shreve: B.M. and stoch. calc., 2.9.F.

22.5. Folgerungen

Kor. 1 Ein typischer Brownscher Pfad ist nirgends Hölderstetig mit $\alpha = \frac{1}{2}$.

Beweis: Nach ~~dem~~ Satz von it. Log. gilt

$$\overline{\lim}_{t \downarrow 0} \frac{|B_{s+t} - B_s|}{\sqrt{t}} = +\infty \quad \text{P-f.s. für jedes } s \geq 0,$$

also ist $(B_t)_{t \geq 0}$ auf keinem Intervall $[a, b]$, $a < b$, Hölderstetig mit $\frac{1}{2}$.

Kor. 2 $T_0 := \inf \{t > 0 \mid B_t = 0\} = 0$ P-f.s.

„Die eindimensionale Brownsche Bewegung kehrt sofort zum Startpunkt zurück“

Allgemeiner: Für P-f.a. ω ist jeder Punkt in

$$Z(\omega) = \{t \in [0, \infty) \mid B_t(\omega) = 0\}$$

ein Häufungspunkt von $Z(\omega)$, s. Karstas / Sheve 9.6.

Kor. 3 (Globaler Satz vom iterierten Logarithmus) P-f.s. gilt:

$$\overline{\lim}_{t \rightarrow \infty} \frac{B_t}{\sqrt{2t \log \log t}} = +1, \quad \underline{\lim}_{t \rightarrow \infty} \frac{B_t}{\sqrt{2t \log \log t}} = -1$$

Beweis: $\overline{\lim}_{t \rightarrow \infty} \frac{B_t}{\sqrt{2t \log \log t}} \stackrel{s = \frac{1}{\varepsilon}}{=} \overline{\lim}_{s \rightarrow 0} \frac{s B_{1/s}}{\sqrt{2s \log \log \frac{1}{s}}}$ ist wieder R.B. mit Start in 0 $\stackrel{\text{kollekt. it. Log.}}{=} +1$ P-f.s. \square

Kor. 4 (Satz vom it. Log. für Random walk, Kolmogorov 1929)

$$S_n = \sum_{i=1}^n X_i, \quad X_i \in \mathcal{L}^2 \text{ i.i.d., } E[X_i] = 0, \text{ var}(X_i) = 1$$

$$\Rightarrow \overline{\lim}_{n \rightarrow \infty} \frac{S_n}{\sqrt{2n \log \log n}} = +1, \quad \underline{\lim}_{n \rightarrow \infty} \frac{S_n}{\sqrt{2n \log \log n}} = -1.$$

Beweis via Skorokhod-Einbettung:

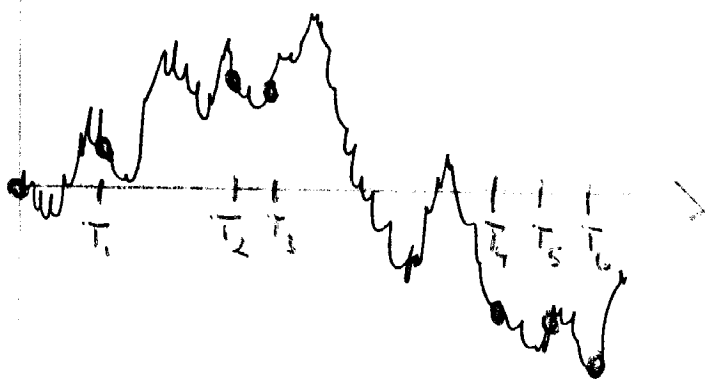
$(B_t)_{t \geq 0}$ B.B. mit Start in 0

$\Rightarrow \exists$ Stoppzeiten $T_1 < T_2 < \dots \nearrow \infty$:

$$S_n := B_{T_n} \quad (n \in \mathbb{N})$$

ist random walk mit derselben Verteilung wie $(S_n)_{n \in \mathbb{N}}$
(s. z. B. Rogers/Williams I.7)

it. Log. für $(B_t)_{t \geq 0} \Rightarrow$ it. Log. für $(S_n)_{n \in \mathbb{N}}$. □



22.6. Beweis des Satzes vom it. Log. (22.4)

① $\circ B d A \quad s=0 \quad (\widehat{B}_t := B_{s+t} - B_s \text{ ist wieder B.B.})$

② (i) \Rightarrow (ii) (denn $-B_t$ ist auch B.B.)

③ Untere Schranke in (i): Nach Borel-Cantelli gilt für $\theta \in (0,1)$:

$$P [B_{\theta^n} - B_{\theta^{n+1}} \geq \sqrt{1-\theta} h(\theta^n) \text{ unendlich oft}] = 1,$$

s.o. Also:

$$\overline{\lim}_{t \downarrow 0} \frac{B_t - B_{\theta t}}{h(t)} \geq \sqrt{1-\theta} \quad \text{P-f.s.}$$

$$\leq \overline{\lim}_{t \downarrow 0} \frac{B_t}{h(t)} + \overline{\lim}_{t \downarrow 0} \frac{-B_{\theta t}}{h(t)}$$

$$\Rightarrow \overline{\lim}_{t \downarrow 0} \frac{B_t}{h(t)} \geq \sqrt{1-\theta} - \overline{\lim}_{t \downarrow 0} \frac{-B_{\theta t}}{h(t)} \geq \sqrt{1-\theta} - \sqrt{\theta} \quad \text{P-f.}$$

$$\leq \overline{\lim}_{t \downarrow 0} \frac{-B_{\theta t}}{h(\theta t)} \cdot \overline{\lim}_{t \downarrow 0} \frac{h(\theta t)}{h(t)}$$

$$\leq -1 \cdot \sqrt{\theta} = -\sqrt{\theta}$$

nach obere Schranke!

($-B_t$ ist wieder eine B.B. Wenn wir die obere Schranke zunächst

Voraussetzen (Beweis s.u.), dann gilt also $\overline{\lim}_{t \downarrow 0} \frac{-B_{\theta t}}{h(\theta t)} = \overline{\lim}_{s \downarrow 0} \frac{-B_s}{h(s)} \leq 1$)

$\ominus \downarrow 0 \Rightarrow$ untere Schranke

④ Obere Schranke in (i) Basiert auf:

236

Lemma $P[B_t \geq \frac{\alpha t}{2} + \beta \text{ für ein } t \in [0,1]] \leq e^{-\alpha\beta} \quad \forall \alpha, \beta > 0$

Beweis: folgt aus 20.9 (Passierzeit einer Gerade) oder direkt so:

$$M_t^\alpha = e^{\alpha B_t - \frac{\alpha^2}{2}t} \text{ exp. Martingal}$$

$$P[B_t \geq \frac{\alpha t}{2} + \beta \text{ für ein } t \in [0,1]] = P[\max_{t \in [0,1]} (B_t - \frac{\alpha t}{2}) \geq \beta]$$

$$= P[\max_{t \in [0,1]} M_t^\alpha \geq e^{\alpha\beta}] \leq \underbrace{\frac{1}{e^{\alpha\beta}}}_{\text{Maximal-}} \underbrace{E[M_t^\alpha]}_{\text{ungleichung}} = e^{-\alpha\beta} \cdot \underbrace{1}_{= E[M_0^\alpha]} \quad \square$$

- Sei nun $\delta > 0$. zz: P-f.s. gilt

$$B_t \leq (1+\delta)h(t) \text{ für hinreichend kleine } t > 0.$$

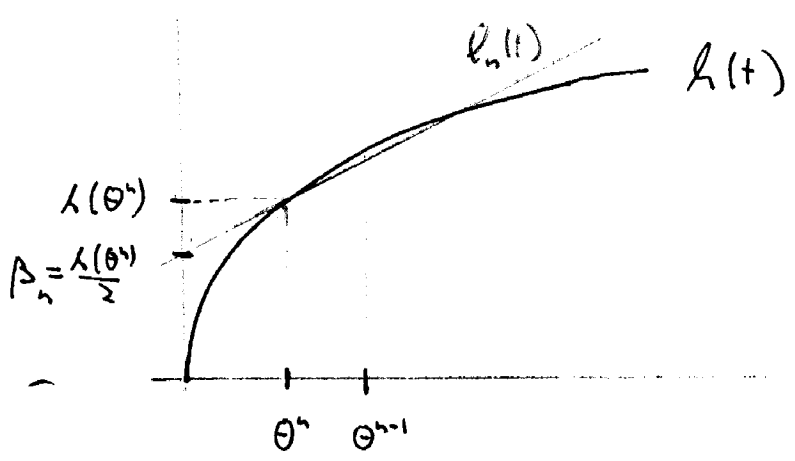
Idee: $\Theta \in (0,1)$. Approximiere $h(t)$ durch lineare Fkt. auf (Θ^n, Θ^{n+1}) , dann benutze Lemma.

$$f_n(t) := \frac{\alpha_n}{2}t + \beta_n, \text{ wobei } \beta_n := \frac{h(\Theta^n)}{2}, \alpha_n := \frac{h(\Theta^n)}{\Theta^n}$$

(vgl. a. Skizze auf S.238)

$$l_n(\theta^n) = h(\theta^n)$$

$$l_n(t) \leq \frac{l_n(\theta^n)}{\theta} = \frac{h(\theta^n)}{\theta} \leq \frac{h(t)}{\theta} \quad \text{für alle } t \in [\theta^n, \theta^{n+1}]$$



z.z.: $B_t \leq (1+\delta) \cdot l_n(t)$
 Schließlich (für $n \geq N(\omega)$)

via Borel-Cantelli:

$$P[B_t > (1+\delta) l_n(t) \text{ für ein } t \in [0,1]] \leq e^{-\alpha_n \beta_n \cdot (1+\delta)^2}$$

$$= e^{-\frac{h(\theta^n)^2}{2\theta^n} \cdot (1+\delta)^2} = e^{-(1+\delta)^2 \log \log \theta^{-n}}$$

\uparrow
 $h(t) = \sqrt{2 + \log \log t^{-1}}$

$$= (n \cdot \log \theta^{-1})^{-(1+\delta)^2} = \text{const.} \cdot n^{-(1+\delta)^2}$$

$\underbrace{\hspace{10em}}_{\text{summierbare Reihe}}$

Also nach Borel-Cantelli: P-f.s.:

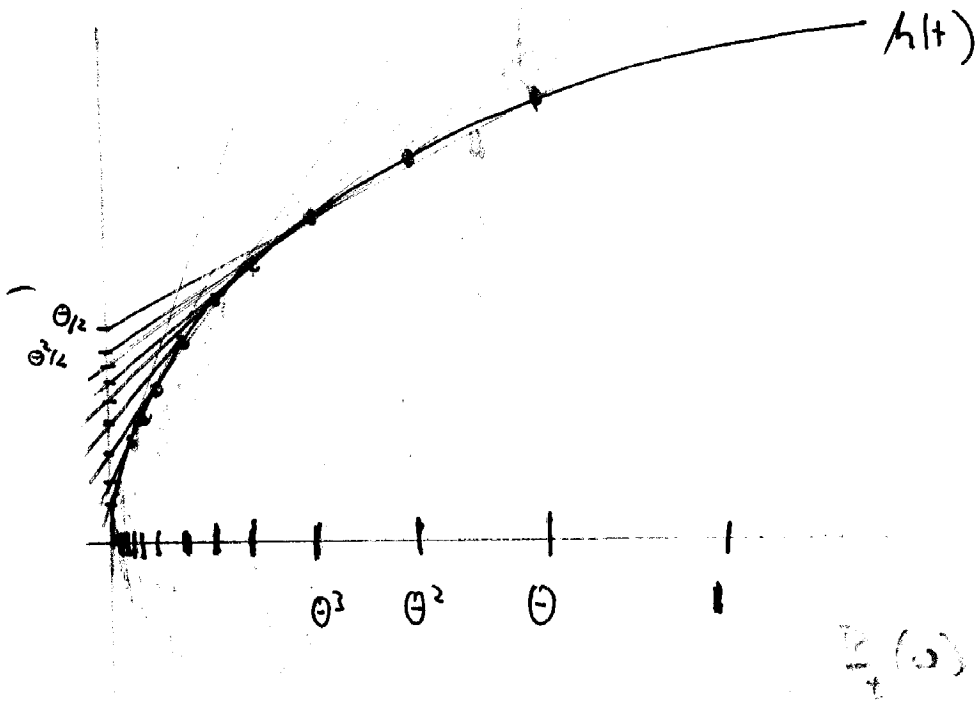
$$B_t \leq (1+\delta) l_n(t) \quad \forall t \in [0,1] \quad \forall n \geq N(\omega)$$

$$\Rightarrow B_t \leq (1+\delta) \inf_{n \geq N} l_n(t) \leq (1+\delta) \frac{h(t)}{\theta} \quad \text{für } t \in \bigcup_{n \geq N} [\theta^n, \theta^{n+1}] = (0, \theta^{N-1})$$

$\leq \frac{h(t)}{\theta} \text{ für } t \in [0, \theta^{N-1}]$

$$\Rightarrow \overline{\lim}_{t \downarrow 0} \frac{B_t}{h(t)} \leq \frac{1+\delta}{\Theta}$$

Jetzt $\Theta \nearrow 1, \delta \downarrow 0, \Theta, \delta \in \mathbb{Q}$. \square



V. ASYMPTOTIK VON ZEITLICH HOMOGENEN

MARKOVKETTEN

23. Grundlagen, starke Markoveigenschaft (S, \mathcal{S}) meßb. Raum

23.1. Definition (Ω, \mathcal{O}, P) W'raum, $(\mathcal{F}_n)_{n \geq 0}$ Filtration

Ein stoch. Prozeß $(X_n)_{n \geq 0}$ auf (Ω, \mathcal{O}, P) mit Zustandsraum S heißt Markovsche Kette bzgl. $(\mathcal{F}_n)_{n \geq 0}$ falls

(i) X_n ist \mathcal{F}_n -meßbar $\forall n \geq 0$

(ii) $P[X_{n+1} \in B | \mathcal{F}_n] = P[X_{n+1} \in B | X_n]$ P-f.s. $\forall B \in \mathcal{S}$

23.2. Übergangskerne

^{Fakt-Lemma}
(ii) $\Rightarrow P[X_{n+1} \in B | \mathcal{F}_n] = p^{(n)}(X_n, B)$ P-f.s.

für eine meßbare Funktion $p^{(n)}(\cdot, B): S \rightarrow [0, 1]$.

Fakt: (S, \mathcal{S}) „gut“ (z.B. Sprechisch, $\mathcal{S} = \mathcal{B}(S)$)

$\Rightarrow \exists$ reguläre Version von $p^{(n)}$, d.h.

(i) $p^{(n)}$ ist stoch. Kern

(ii) $p^{(n)}(X_n, B)$ ist Version von $P[X_{n+1} \in B | \mathcal{F}_n]$ $\forall B \in \mathcal{S}$

Mit maßtheoretischer Induktion folgt dann:

(240)

$$(*) \quad E [f(X_{n+1}) | \mathcal{F}_n] = (p^{(n)} f)(X_n) := \int p^{(n)}(X_n, dy) f(y)$$

für alle $f \geq 0$ messbar.

Def. X_n heißt zeitlich homogene MK mit Übergangskern $p(x, dy)$

(bezgl. (\mathcal{F}_n)) $\Leftrightarrow X_n$ ist (\mathcal{F}_n) -adaptiert und (*) gilt mit $p^{(n)} = p \forall n$

n -Schritt-Übergangswahrscheinlichkeit (im zeitl. hom. Fall):

$$P_n(x, dy) = \int \dots \int p(x, dx_1) p(x_1, dx_2) \dots p(x_{n-1}, dy)$$

$$P_n f = p^n f \quad (\text{n-faches Anwenden des li. Operators } p: f \mapsto pf)$$

insbesondere: $P_{n+m} = P_n P_m$ (Chapman-Kolmogorov)

23.3. Zeitlich homogene Markovprozesse in kontinuierlicher Zeit

Seien $P_t, t \in [0, \infty)$, stoch. Kerne. (analog zu p_n in diskreter Zeit)

Def. $(X_t)_{t \in [0, \infty)}$ heißt zeitl. homogener Markovprozess bezgl. $(\mathcal{F}_t)_{t \geq 0}$
mit Übergangskernen $P_t (t \geq 0)$ falls

(i) adaptiert (ii) $P[X_{t+s} \in B | \mathcal{F}_t] = P_s(X_t, B) \forall s, t \geq 0, B \in \mathcal{S}$.

Existenz eines Markovprozesses mit Kernen $(p_t)_{t \geq 0}$ und Start in x

$$\Rightarrow P_{t+s}(x, dz) = \int P_t(x, dy) P_s(y, dz)$$

Chapman-Kolmogorov-Gleichung

$$P_{t+s} f = P_t P_s f \quad \forall f \text{ m\ss} \text{bar, beschr.}$$

d.h. $(P_t)_{t \geq 0}$ ist Halbgruppe von linearen Operatoren.

Charakterisierung durch einen Kern wie in diskreter Zeit nicht m\oglich.

Stattdessen: Betrachte „infinitesimalen Generator“:

$$\mathcal{L} f = \lim_{\varepsilon \downarrow 0} \frac{P_\varepsilon f - f}{\varepsilon} \quad (\text{f\ur} \text{ so da\ss} \text{ Limes in geeigneter Sinn ex.)}$$

z. B.: zeitstetige Markovkette $\rightarrow \mathcal{L} = Q$ -Matrix

Brownsche Bewegung $\rightarrow \mathcal{L} = \frac{1}{2} \Delta$

(denn $P_t f(x) = E_x[f(X_t)]$ l\ost u.g.V. $\frac{\partial}{\partial t} P_t f(x) = \frac{1}{2} \Delta P_t f(x)$)

Bem. 1) $\mathcal{L} f(x) = \lim_{\varepsilon \downarrow 0} \frac{E_x[f(X_\varepsilon)] - f(x)}{\varepsilon} = \lim_{\varepsilon \downarrow 0} E_x \left[\frac{f(X_\varepsilon) - f(X_0)}{\varepsilon} \right]$

„mittlere \u00c4nderungsrate von f entlang X_t “

$$2) \text{ Chapman-Kolmogorov} \Rightarrow \underline{\left(\frac{d}{dt}\right)^+ P_t f} = \lim_{\varepsilon \downarrow 0} \frac{\overset{= P_t P_{t+\varepsilon} f}{P_{t+\varepsilon} f - P_t f}}{\varepsilon} = \underline{L P_t f} \quad (242)$$

(Kolmogorovs Rückwärtsgleichung)

u.g.V. \Rightarrow Rekonstruktion von $(P_t)_{t \geq 0}$ aus L

(z.B. falls Spolnisch, P_t Feller, d.h. $f \in C_0(S) \Rightarrow P_t f \in C_0(S)$)

stetige Fkt., die
im unendlichen verschwindet

Zeitlich inhomogene Markovprozesse:

$$P[X_{t+s} \in B | \mathcal{F}_t] = P_{t,t+s}(X_t, B)$$

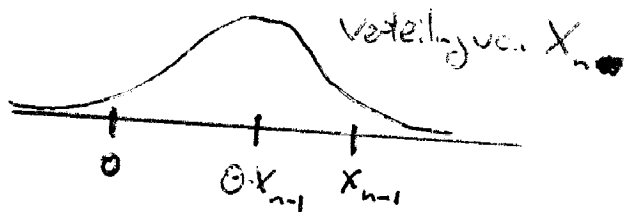
Chapman-Kolmogorov: $P_{t,u} = P_{t,s} P_{s,u}$

Generatoren: $L_t f = \lim_{\varepsilon \downarrow 0} \frac{P_{t,t+\varepsilon} f - f}{\varepsilon}, \quad t \geq 0$

23.4. Beispiel: Ornstein-Uhlenbeck Prozess $S = \mathbb{R}$

in diskreter Zeit: $U_n (n \in \mathbb{N})$ i.i.d., $\sim N(0,1)$, $\Theta, \sigma > 0$

$$X_n = \Theta X_{n-1} + \sigma \cdot U_n$$



Übergangskern $p(x, \cdot) = N(\Theta x, \sigma^2)$

[allgemein: Zufälliges dynamisches System: $X_n = g(X_{n-1}, U_n)$]
 (noise $\leftarrow U_n$)
 (Bewegungsgesetz $\leftarrow g$)

Stochastische Differenzgleichung: $X_n - X_{n-1} = (\Theta - 1) X_{n-1} + \sigma U_n$

$\sim B_n - B_{n-1}$
 wobei $(B_t)_{t \geq 0}$ Br. Bew.

\Rightarrow Analogon in kontinuierlicher Zeit:

$$dX_t = \underbrace{(\Theta - 1) X_t}_{\text{deterministische Drift}} + \underbrace{\sigma dB_t}_{\text{Rauschen / Br. Bewegung}}$$

Stoch. Differentialgleichung

$$\text{(d.h. } X_t - X_0 = \int_0^t (\Theta - 1) X_s ds + \sigma \cdot B_t \text{)}$$

ist Markovprozess mit Generator $\mathcal{L}f = \frac{\sigma^2}{2} f'' + (\Theta - 1) \cdot f'$

23.5. Starke Markov-Eigenschaft in diskreter Zeit

Betrachte kanonisches Modell einer zeitl. homogenen MK mit Übergangskern $p(x, dx)$ und Startverteilung $\nu \in \mathcal{M}_1(S)$:

$$\Omega = S^{\{0,1,2,\dots\}}, X_n(\omega) = \omega_n, \mathcal{O} = \sigma(X_n | n \geq 0), \mathcal{F}_n = \sigma(X_0, \dots, X_n)$$

$$P_\nu = \nu \otimes p \otimes p \otimes \dots, \text{ d.h. } P_\nu \circ (X_0, \dots, X_n)^{-1} = \nu(dx_0) p(x_0, dx_1) p(x_1, dx_2) \dots p(x_{n-1}, dx_n)$$

Markov-Eigenschaft: $F: \Omega \rightarrow \mathbb{R}$ meßbar, ≥ 0

$$\Rightarrow E_\nu [F(X_n, X_{n+1}, \dots) | \mathcal{F}_n] = E_{X_n} [F(X_0, X_1, \dots)]$$

$$E_\nu [F \circ \Theta_n | \mathcal{F}_n] = E_{X_n} [F], \Theta_n(x_0, x_1, \dots) = (x_n, x_{n+1}, \dots)$$

Verallgemeinerung: Sei $T: \Omega \rightarrow [0, \infty]$ Stoppzeit bzgl. (\mathcal{F}_n) ,

$$\mathcal{F}_T := \{A \in \mathcal{O} \mid A \cap \{T=n\} \in \mathcal{F}_n \forall n \geq 0\}$$

Konvention: $\Theta_\infty(\omega) = \Delta$, wobei Δ Zusatzpunkt ($\Delta \notin \Omega$)

Satz (Starke Markov-Eigenschaft) $E_\nu [F \circ \Theta_T | \mathcal{F}_T](\omega) = E_{X_{T(\omega)}} [F] P_\nu$ f.s. auf $\{T < \infty\}$

für alle $F: \Omega \rightarrow \mathbb{R}$ meßbar, ≥ 0 , $F(\Delta)$ beliebig definiert.

$$E_0 [F(X_T, X_{T+1}, \dots) | \mathcal{F}_T] = E_{X_T} [F(X_0, X_1, \dots)] \quad P_0\text{-f.s. auf } \langle T < \infty \rangle$$

Beweis: zz: $E_0 [F \circ \Theta_T | \mathcal{F}_T] \cdot I_{\langle T < \infty \rangle} = E_{X_T} [F] \cdot I_{\langle T < \infty \rangle} \quad P_0\text{-f.s.}$

1) beide Seiten sind \mathcal{F}_T -messbar

2) Also ok falls $\int_A \text{l.s.} dP_0 = \int_A \text{r.s.} dP_0 \quad \forall A \in \mathcal{F}_T$

d.h. $E_0 [\underbrace{E_0 [F \circ \Theta_T | \mathcal{F}_T]}_{\in \mathcal{F}_T}; A \cap \langle T < \infty \rangle] = E_0 [E_{X_T} [F]; A \cap \langle T < \infty \rangle]$

In der Tat:

$$E_0 [\underbrace{F \circ \Theta_T}_{\equiv \Theta_n}; \underbrace{A \cap \langle T = n \rangle}_{\in \mathcal{F}_n}] \stackrel{ME}{=} E_0 [\underbrace{E_{X_n} [F]}_{\equiv X_T}; A \cap \langle T = n \rangle]$$

für alle $n \geq 0$. □

23.6. Starke Markoveigenschaft der Brownschen Bewegung

$\Omega = C([0, \infty))$, $X_t(\omega) = \omega(t)$, $\mathcal{O}_t = \sigma(X_s | t \geq 0)$, $P_x =$ Wienermaß (Start in x)

$\mathcal{F}_t = \mathcal{F}_t^X = \sigma(X_s | 0 \leq s \leq t)$ oder $\mathcal{F}_t = \mathcal{F}_{t+}^X$,

$\bar{T}: \Omega \rightarrow [0, \infty)$ Stoppzeit bzgl. $(\mathcal{F}_t)_{t \geq 0}$,

$(\Theta_t \omega)(s) = \omega(s+t)$, $\bar{T}: \Omega \rightarrow \mathbb{R}$ messbar, ≥ 0 oder beschr.

Satz $E_x [F \circ \Theta_T | \mathcal{F}_T] = E_{X_T} [F]$ \mathbb{P}_x -f.s. auf $\{T < \infty\}$

Beweisskizze: 1) T diskret, d.h. $T(\omega) \in \{t_1, t_2, \dots\} \cup \{\infty\}$ $t_n \uparrow \infty$:
verfahre wie bei Markovketten

2) T beliebig, $F = f(X_{s_1}, \dots, X_{s_n})$ mit $f \in C_b(\mathbb{R}^n)$:

Wähle T_n diskret, $T_n \downarrow T$, mit $\{T_n = \infty\} = \{T = \infty\}$

Beh. für $T_n \xrightarrow{n \rightarrow \infty} \Rightarrow$ Beh. für T (benutze Stetigkeit von $t \mapsto X_t$)

3) T beliebig, $F \geq 0$ neßbar: via maßtheoret. Induktion / monotone Klassen □

Kor. $(\mathcal{B}_t)_{t \geq 0}$ B.B., T Stoppzeit bzgl. $\mathcal{F}_t = \mathcal{F}_t^B$ oder $\mathcal{F}_t = \mathcal{F}_{t+}^B$, $T < \infty$ P.f.s.

$\Rightarrow (\mathcal{B}_{T+t} - \mathcal{B}_T)_{t \geq 0}$ ist wieder B.B., unabhängig von \mathcal{F}_T

Beweis: oBdA harmonisches Modell, $F: \Omega \rightarrow \mathbb{R}$ neßb. beschr.

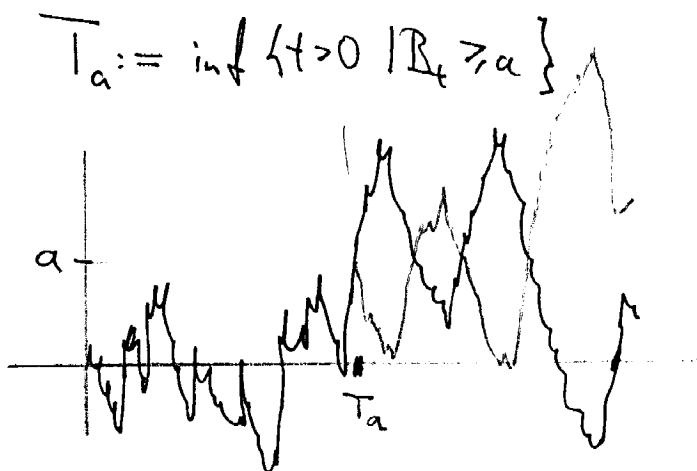
$$\begin{aligned} \Rightarrow E[F((X_{T+t} - X_T)_{t \geq 0}) | \mathcal{F}_T](\omega) & \stackrel{SME}{=} E_{X_T(\omega)} [F((X_t - X_0)_{t \geq 0})] \\ & = E_0[F] \quad \mathbb{P}\text{-f.s.} \end{aligned}$$

z.B. mit Start in 0 unter \mathbb{P}_x für alle x

Also: 1) r.S. unabh. von $\omega \Rightarrow (X_{T+t} - X_T)_{t \geq 0} \perp \mathcal{F}_T$

2) $E[F((X_{T+t} - X_T)_{t \geq 0})] = E[F]$ \Rightarrow Wert von $X_{T+t} - X_T = 0$ (Wiederwahl ...)

23.7. Reflektionsprinzip $(B_t)_{t \geq 0}$ B.B. im \mathbb{R}^1 mit Start in 0, $a > 0$



$B_t^{T_a} := B_{T_a+t} \quad (t \geq 0)$

$\tilde{B}_t := B_{T_a+t} - B_{T_a} \quad (t \geq 0)$

Satz $(T_a, B^{T_a}, \tilde{B}) \sim (T_a, B^{T_a}, -\tilde{B})$ gleiche gemeinsame Verteilung

Beweis: \tilde{B} Br. Bew., $\perp \mathcal{F}_{T_a}$ also $\perp \sigma(B^{T_a}, T_a)$

$\Rightarrow P_0(T_a, B^{T_a}, \tilde{B})^{-1} = P_0(T_a, B^{T_a})^{-1} \otimes \text{Wienerma\ss} = P_0(T_a, B^{T_a}, -\tilde{B})^{-1}$

Kor. $P[T_a \leq t \text{ und } B_t \leq a-x] = P[T_a \leq t \text{ und } B_t \geq a+x] = P[B_t \geq a+x]$
für alle $t \geq 0, x \geq 0$.

Analyse zeigt ~~man~~ ^{man} explizite Formel für gemeinsame Vert. von T_a und B_t ∇

Beweis: $B_t = B_{T_a} + (B_t - B_{T_a}) = a + \tilde{B}_{t-T_a}$ auf $\{T_a \leq t\}$

\Rightarrow l.S. $= P[T_a \leq t, \tilde{B}_{t-T_a} \leq -x] \stackrel{\text{Reflex. p.}}{=} P[T_a \leq t, \tilde{B}_{t-T_a} \geq x]$ r.S. \square

24. Rekurrenz und Transienz

S abzählbar, $P(x,y)$ stoch. Matrix auf S ,

$(\mathbb{R}, \mathcal{O}, (P_x)_{x \in S}, (X_n)_{n \geq 0})$ kanon. Modell der Markovkette
mit Übergangsmatrix $P(x,y)$

$B_y(\omega) := \sum_{n=0}^{\infty} I_{\{y\}}(X_n(\omega))$ Anzahl der Besuche in y
Aufenthaltszeit in y

24.1. Definition

$x \in S$ heißt transient $\Leftrightarrow P_x[B_x < \infty] = 1$

rekurrent $\Leftrightarrow P_x[B_x = \infty] = 1$

Die Markovkette heißt transient bzw. rekurrent falls alle $x \in S$
transient bzw. rekurrent sind.

24.2. Verteilung von B_y $f_{x,y} := P_x[T_y < \infty], T_y := \min\{n \geq 1 \mid X_n = y\}$

Satz $P_x[B_y \geq n] = \begin{cases} f_{x,y} \cdot f_{y,y}^{n-1} & \text{wenn } x \neq y \\ f_{y,y}^{n-1} & \text{wenn } x = y \end{cases}$

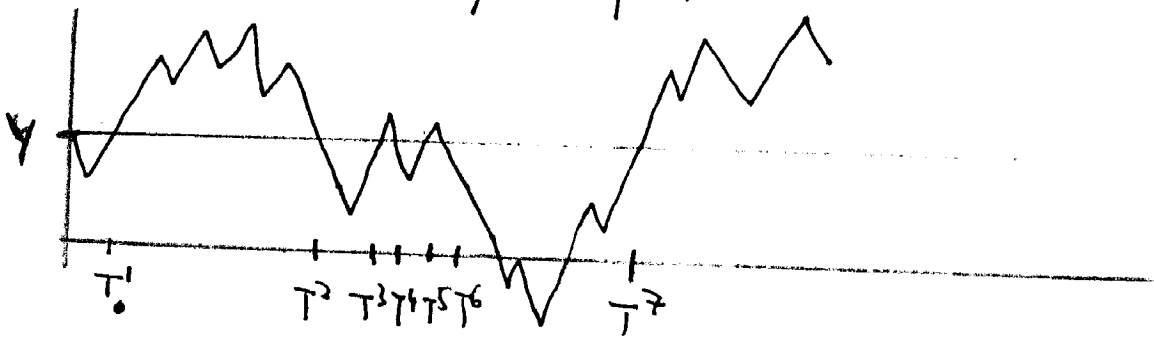
Inbesondere: Jedes $y \in S$ ist entweder rekurrent oder transient,
 y rekurrent $\Leftrightarrow f_{y,y} = 1$, y transient $\Leftrightarrow f_{y,y} < 1$.

Beweis: via starke Markov-Eigenschaft

$T_y := \min \{n \geq 1 \mid X_n = y\}$ erste Besuchszeit / Rückkehrzeit

$$T^0 := 0$$

$T^n := T^{n-1} + T_y \circ \Theta_{T^{n-1}}$ n-te Besuchs- / Rückkehrzeit nach y



$$P_x [T^n < \infty] = P_x [T^{n-1} < \infty \text{ und } T_y \circ \Theta_{T^{n-1}} < \infty]$$

$$= E_x \left[\underbrace{P_x [T_y \circ \Theta_{T^{n-1}} < \infty \mid \mathcal{F}_{T^{n-1}}]}_{\stackrel{\text{SME}}{=} P_x [T_y < \infty]} ; T^{n-1} < \infty \right]$$

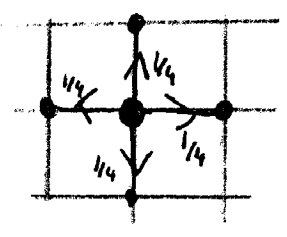
$$= P_y [T_y < \infty] \cdot P_x [T^{n-1} < \infty] = \dots = P_x [T_y < \infty] \cdot P_y [T_y < \infty]^{n-1}$$

$$= f_{x,y} \cdot f_{y,y}^{n-1}$$

$$P_x [R_y \geq n] = \begin{cases} P_x [T^n < \infty] & \text{falls } x \neq y \\ P_x [T^{n-1} < \infty] & \text{falls } x = y \end{cases}$$

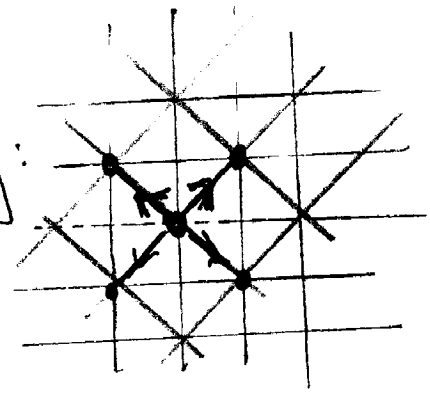
(Stirling: $n! \sim \sqrt{2\pi n} \left(\frac{n}{e}\right)^n$, d.h. $\frac{n!}{\sqrt{2\pi n} \left(\frac{n}{e}\right)^n} \rightarrow 1$)

b) d=2



$X_n = (X_n^1, X_n^2)$ Komponenten sind nicht unabhängig!

- 45°-Drehung:



$Y_n = (X_n^1 + X_n^2, X_n^1 - X_n^2)$

Jetzt sind die Komponenten unabhängige 1-D random walks!

X_n rekurrent $\Leftrightarrow Y_n$ rekurrent

Übergangsw'keiten für Y_n :

$$P^{2n}(x,x) = P_x [Y_{2n}^1 = x_1, Y_{2n}^2 = x_2] \stackrel{\text{unabl.}}{=} P_x [Y_{2n}^1 = x_1] \cdot P_{x_2} [Y_{2n}^2 = x_2]$$
$$= \left(\binom{2n}{n} 2^{-2n} \right)^2 \stackrel{\text{s.o.}}{\sim} \frac{1}{\pi n}$$

$\Rightarrow Q(x,x) = \infty \Rightarrow$ Jedes x ist rekurrent.

c) $d=3$ $X_n = (X_n^1, X_n^2, X_n^3)$ mit unabhängiger r.v. X_n^i auf \mathbb{Z}^1

$$\Rightarrow P^{2n}(k, x) = \left(\binom{2n}{n} 2^{-2n} \right)^3 \sim \frac{1}{(\pi n)^{3/2}}$$

$$\Rightarrow G(k, x) = \sum P^{2n}(k, x) < \infty \quad \underline{\text{transient!}}$$

Dies gilt auch für den Standard random walk auf \mathbb{Z}^3

- - aber diesmal kann man die beiden random walks nicht durch eine Doobing ineinander überführen, und muß RECHNEN

→ Übung

24.5. Kommunikationsklassen

Definition $x \rightsquigarrow y \Leftrightarrow P_x[T_y < \infty] > 0$

Bem. 1) $x \rightsquigarrow y \Rightarrow \exists n: P^n(x, y) > 0 \Rightarrow G(x, y) > 0$

2) $x \rightsquigarrow y \wedge y \rightsquigarrow z \Rightarrow x \rightsquigarrow z$

Lemma $G(x, y) = P_x[T_y < \infty] \cdot G(y, y) \leq G(y, y)$

a) $d \in \mathbb{Z}$ ($d \geq 3$)

(254)

$X_n = (X_n^1, X_n^2, X_n^3)$, X_n^i i.i.d. $N(0, 1)$

$$\Rightarrow p^{2n}(x, x) = \left(\binom{2n}{n} 2^{-2n} \right)^3$$

$$\approx \frac{1}{(\pi n)^{3/2}}$$

$\Rightarrow G(x, x) < \infty$ transient

Dasselbe gilt für klass. r.w. on \mathbb{Z}^3 (Jury)

2.5 Kommunikationssysteme

Def: $x \sim y \Leftrightarrow P_x[T_y < \infty] > 0$

- Bem:
- 1) $x \sim y \Leftrightarrow \exists n: p^n(x, y) > 0 \Leftrightarrow G(x, y) > 0$
 - 2) $x \sim y, y \sim z \Leftrightarrow x \sim z$

Lemma

$$G(x, y) = P_x[T_y < \infty] \cdot G(y, y) \leq G(y, y)$$

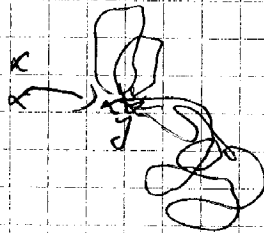
Beweis:

$$G(x, y) = E_x[B_y]$$

$$= E_x[B_y; T_y < \infty]$$

$$B_y \circ \theta_{T_y}$$

$$= E_y[B_y] \cdot P_x[T_y < \infty]$$



Satz:

x rekurrent $\Leftrightarrow y$ rek., $y \sim x$, $R_y = \infty$ P_x -f.s.
 $x \rightsquigarrow y$
 $R_x = \infty$ P_x -f.s.

Beweis:

① $y \rightarrow x$:

$$P_x[T_y < \infty; T_x = \theta_{T_y} < \infty] \\ \stackrel{\text{rech.}}{=} P_x[T_y < \infty] > 0 \stackrel{\text{SMB}}{=} P_x[T_y < \infty] P_y[T_x < \infty]$$

② $y \rightarrow x$:

Sei $m \geq 0$ mit $p^n(x, y) > 0$

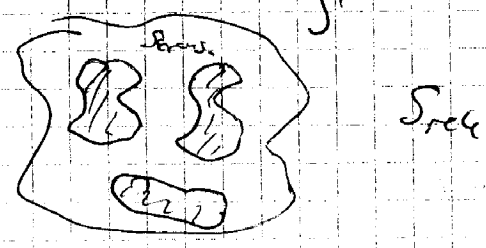
$$G(x, y) \stackrel{\text{Lemma}}{\geq} G(x, y) \geq \sum_{n=0}^{\infty} p^{n+m}(x, y) \\ \geq \sum_{n=0}^{\infty} p^n(x, x) p^m(x, y) \\ = \underbrace{G(x, x)}_{> 0} \underbrace{p^m(x, y)}_{> 0}$$

③ $P_x[B_y = \infty] = \underbrace{P_x[T_y < \infty]}_{> 0} \cdot \underbrace{P_y[B_y = \infty]}_{= 1}$

$> 0 \cdot 1 = 1$

Kor: $x \sim y$ ist Äquivalenzrelation auf $S_{rech} = \{x \in S \mid x_{rech}\}$

Darüber Zerlegung $S_{rech} = \cup_{i \in I} S_i$ Äquivalenzklassen



Def: $P(x,y)$ rekurrenz auf $\Omega \subset S \Leftrightarrow$
 $\exists n \geq 1 \forall x,y \in \Omega$
 $\exists \{x_0, \dots, x_n\} \subset \Omega$ $P(x_0, x_1) > 0$

Korollar

Eine irreduzible MK auf S ist entweder rekurrent oder transient.

Genauer:

Entweder:

- 1) Alle $x \in S$ sind rek. und $P_x[B_y = \infty] = 1 \forall x,y \in S$
- 2) Alle $x \in S$ sind transient und $P_x[B_y < \infty] = 1 \forall x,y \in S$
 sogar $E_x[B_y] < \infty$

Beweis

$\exists x \in S$ rek. \Rightarrow Alle $y \in S$ rekurrenz \rightarrow 1)
 sonst $G(x,x) < \infty \forall x \in S \Rightarrow G(x,y) < \infty \forall x,y$

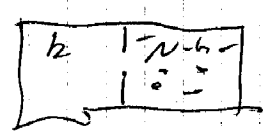
24.7 Beispiele

1) S endl. $\Rightarrow \exists y \in S$ rek.

Sei $x \in S, E[\sum_{j \in S} B_j] = \infty \Rightarrow \exists y: E_x[B_y] = \infty$
 $G(x,y) = \infty$
 $G(y,y) = \infty$

insbesondere: S endl. \Rightarrow p. rec.
 \Rightarrow Alle $x \in S$ sind rekurrent

z.B. Ehrenfestmodell



\leftarrow alle rechts und links erreicht,
 1) aber Europa kommt nur zu ??

2) Wie häufig wird ein $x \in S$ im Schnitt durchlaufen? \rightarrow Ergodensatz

2) Verzweigungsprozess

$$Z_n = \sum_{i=0}^{n-1} N_i$$

$N_i: \mathbb{Z} \rightarrow \{0, 1, 2, \dots\}$ iid $\sim \nu$

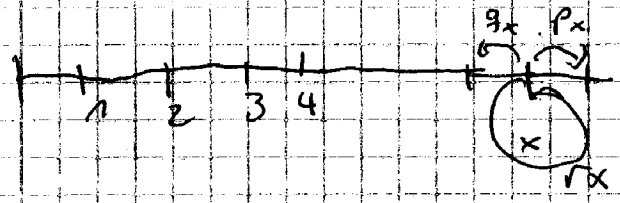
0 ist absorbierende $\forall x \neq 0$

$$\nu(0) = 0 \Rightarrow x \sim 0 \quad \forall x$$

\rightarrow Alle $x \neq 0$ sind transient

$\Rightarrow P_x$ -f.s. $Z_n = 0$ ultrophisch oder $Z_n \rightarrow \infty$

3) Birth and death processes



$$h(x) = \sum_{y=0}^{x-1} \prod_{z=y+1}^x \frac{q_z}{p_z}$$

$$h(0) = 0$$

$$h(1) - h(0) = 1$$

$$\Rightarrow P_x[\text{first } X_{T_{0,0}} = a] = \frac{h(a) - h(x)}{h(0) - h(x)}$$

\Rightarrow [transient $\Leftrightarrow h(\infty) < \infty$ (Übung)]

" Phasenübergang von [red = 0]

Rekurrenz zu Transienz

$$\text{bei } p_x = \frac{1}{2} + \frac{1}{4x} \quad (\text{Übung})$$

Bem Br. Bew:

(258)

$U \subseteq \mathbb{R}^d$ offn. (oder abg.)

Mittlere Aufenthaltzeit in U bei Start in x

$$G(x, U) = \mathbb{E}_x \left[\int_0^{\infty} \mathbb{1}_U(B_s) ds \right]$$

$$\stackrel{\text{Fubini}}{=} \int_0^{\infty} \underbrace{\mathbb{P}_x [B_s \in U]}_{p_s(x, U)} ds$$

$$\stackrel{\text{Fubini}}{=} \int_U \int_0^{\infty} p_s(x, y) ds dy =: G(x, y)$$

$$G(x, y) = \int_0^{\infty} (2\pi s)^{-\frac{d}{2}} e^{-\frac{|x-y|^2}{2s}} ds$$

$$= \begin{cases} +\infty & \text{für } d=2 \\ \frac{\text{const.}}{|x-y|^{d-2}} & \text{für } d \geq 3 \end{cases}$$

24.8 Rekurrenz- und Transienzskriterium

S abz. (X_n, P_x)

a) Effiziente Methode um Rekurrenz zu zeigen:

Satz: Sei S irreduzibel. Existiert $h: S \rightarrow \mathbb{R}$ und FES endl.

mit

i) $ph \leq h$ auf $S \setminus F$ (Superharmonizität)

ii) $\{h \leq c\}$ endlich $\forall c \in \mathbb{R}$

dann ist die MK rekurrent.

Beweis

- 1) h ist stochastische Lyapunovfunktion
- 2) Bedingung ist auch notwendig (ohne Beweis)

Beweis $\{h \leq 0\}$ endl. $\Leftrightarrow \inf h > -\infty$

o.B.d.A. $h \geq 0$ (sonst $h \rightarrow h+c$)

i) $M_n = h(X_n) \cdot \mathbb{1}_{\{n < T_F\}}$ ist Supermartingal unter $P_x \forall x \in F^c$

Konvergenz $\Rightarrow \exists M_\infty = \lim M_n$ P_x -f.s. $\forall x \in F^c$

ii) X_n transient

$\Rightarrow X_n \in \{h \leq c\}$ nur endl. oft. P_x -f.s. $\forall x$

$\Rightarrow h(X_n) \rightarrow \infty$ P_x -f.s. $\forall x$

M_n konv. $\Rightarrow T_F < \infty$ P_x -f.s. $\forall x \in F^c$

SMB $\Rightarrow X_n \in F$ ∞ oft. P_x -f.s. $\forall x$

Parad. $\Rightarrow \exists x \in F$ rekurrenz \downarrow

Nachweis von Transienz

Satz S i-rod und rek. \Leftrightarrow Jede nichtnegative superharmonische Funktion ist konstant

Beweis siehe 7.4 bzw. Satz 5.4 (via Stoppzeit)

Bsp: r.w. in \mathbb{Z} mit $p = \frac{1}{2}$ $h(x) = (\frac{q}{p})^x$ harmonisch

\Rightarrow transient

25. Invariante Maße und stationäre Verteilungen

(260)

(S, \mathcal{F}) meßb. Raum, (X, d_X) stochast. Vers.,
 \rightarrow WV auf (S, \mathcal{F}) , $(X_n)_{n \geq 0}$ auf $(\mathcal{R}, \mathcal{O}, P)$
kanonische MK', $\Theta: \mathcal{R} \rightarrow \mathcal{R}$ shift

Fragen:

- ① Existiert eine „Gleichgewichtsverteilung“ μ auf (S, \mathcal{F}) mit Verteilung $(X_n) = P_n \circ X_0^{-1} = \nu P^n \xrightarrow{n \rightarrow \infty} \mu$?
- ② (Zunächst) allgemein: F WV P auf $(\mathcal{R}, \mathcal{O})$
Vert $(X_n, X_{n+1}, \dots) = P_n \circ \Theta^n \rightarrow P$?
„Asymptotische Stationarität“

Konvergenzbesonderheit:

z.B. Konvergenz in Variationsnorm:

$$\|P - Q\|_{\text{var}} = \sup_A |P(A) - Q(A)|$$

$$\stackrel{\text{Jens}}{=} \frac{1}{2} \sup_{f \text{ nichtneg.}} \left| \int f dP - \int f dQ \right|$$

$$= \frac{1}{2} \sum_{\omega} |p(\omega) - q(\omega)| \quad \text{falls } \mathcal{F} \text{ diskret}$$

Satz:

- 1) ② \Rightarrow ① mit $\mu = P \circ X_0^{-1}$, Umkehrung gilt auch s.u.
- 2) Gilt ①, dann folgt $\mu P = \lim_{n \rightarrow \infty} \nu P^{n+1} = \lim_{n \rightarrow \infty} \nu P^n = \mu$

3) gilt \textcircled{D} , dann folgt auch analog $P \circ \theta^{-1} = P$

25.1 Stationarität

$$\Omega = S^{\mathbb{Z}}, \alpha = \sigma(X_n | n \in \mathbb{Z}), X_n(\omega) = \omega_n$$

Definition: Eine WV P auf (Ω, α) heißt stationär, falls $P \circ \theta^{-1} = P$

• Ein stochastischer Prozess heißt stationär, falls die Verteilung auf (Ω, α) stationär ist.

- Bsp: X_0, X_1, X_2, \dots iid wv $P \Rightarrow P$ stationär

Bem: P stationär \Leftrightarrow Vert $(X_n, X_{n+1}, \dots) = P \circ \theta^{-n} = P$ unabhängig von n

$(\Omega, \alpha, P, \theta)$ ist dann ein dynamisches System, d.h.

(Ω, α, P) endl. Maßraum, $\theta: \Omega \rightarrow \Omega$ maßerhaltend

\rightarrow Ergodentheorie

25.2: Stationäre Markovketten

$(\Omega, \alpha, X_n, P_\mu)$ mit Kern $p(x, dy)$

Satz: $P_\mu \circ \theta^{-n} = P_{\mu^n} \quad \forall \mu \in M_1(S)$

Insbesondere

1) P_μ stationär $\Leftrightarrow \mu = \mu^p$

2) $\|P_n \circ \theta^{-n} - P_\mu\|_{var} = \|\nu^n - \mu\|_{var}$

Also P_n asymptotisch stationär falls $\nu^n \rightarrow \mu$

Beweis $P_{\mu \circ \theta^{-n}}(A) = E_{\mu} [I_A \circ \theta^n] = E_{\mu} [E_{X_n} [I_A]]$ (262)

$$= \int P_{\mu} [A] (\mu \circ \theta^n)(dx)$$

$$= P_{\mu \circ \theta^n} [A]$$

⇒ 1)

$$2) \text{ L.S.} = \sup_{|F| \leq 1} \left| \int F d(P_{\mu \circ \theta^{-n}}) - \int F dP_{\mu} \right|$$

$$= \sup_{|F| \leq 1} \left| \int \underbrace{E_{X_n} [F]}_{\leq 1} \mu \circ \theta^n - \int \underbrace{E_{X_n} [F]}_{\leq 1} \mu(dx) \right|$$

↳ $\mu \circ \theta^n$ -nullbar

oder andersherum ist

Def: Ein (positives) Maß μ auf (S, \mathcal{F}) mit $\mu \circ \theta = \mu$ heißt invariantes Maß

Ein invariantes Maß heißt auch stationäre (Start-) Verteilung oder Gleichgewicht

Existenz? Eindeutigkeit? Darstellung?

25.3. Beispiele:

1) Random walk auf \mathbb{Z} $\begin{matrix} q & p \\ \leftarrow & \rightarrow \end{matrix}$ $q = 1 - p > 0$

$$\mu = \mu \circ \theta \Leftrightarrow \mu(x) = \mu(x+1) \cdot p + \mu(x-1) \cdot q \quad \text{CEV!}$$

$$\Leftrightarrow \begin{cases} \mu(x) = a + bx & p = \frac{1}{2} \\ \mu(x) = a + b \left(\frac{p}{q}\right)^x & p \neq \frac{1}{2} \end{cases}$$

$p = \frac{1}{2}$: $\mu(x) = 1$ ist $\chi_{\mathbb{Z}}$ auf \mathbb{Z} ein ~~endliches~~ endliches invariantes Maß

$\mu(\mathbb{R} \setminus \{0\}) \Rightarrow$ nicht normierbar
 \Rightarrow \mathbb{R} invariante WV.

$p + \frac{1}{2}$ $\mu(x) = 1, \mu(x) = \left(\frac{q}{p}\right)^x$ sind invariant, aber
 \mathbb{R} inv. WV

2) R. w. mit Reflexion \Rightarrow $\mathbb{R} \setminus \frac{q}{p}$

$$\mu(x) = \begin{cases} \left(\frac{p}{q}\right)^x & \text{für } x \geq 1 \\ p & \text{für } x = 0 \end{cases}$$

invar.

normierbar $\Leftrightarrow p < q$ (prüft nach check!)

3) Ornstein-Uhlenbeck auf \mathbb{R}^1

$$X_n = \theta X_{n-1} + \sigma U_n, \quad \theta, \sigma > 0, \quad U_n \text{ i.i.d. } \mathcal{N}(0,1)$$

Ab $X_{n-1} \sim \mathcal{N}(m, v) \Rightarrow X_n \sim \mathcal{N}(\theta m, \theta^2 v + \sigma^2)$

invariantes ~~WV~~ Verteilung falls $m = \theta m, v = \theta^2 v + \sigma^2$

$$\theta < 1: \mathcal{N}\left(0, \frac{\sigma^2}{1-\theta^2}\right) \text{ invariante WV}$$

$$\theta > 1 \Rightarrow \mu(dx) = e^{-\frac{(\theta^2-1)x^2}{2\sigma^2}} dx \text{ invariante Maß, nicht normierbar}$$

25.4. Der Wiederkehrsatz von Lac

Sei $(X_n)_{n \geq 0}$ stationärer stoch. Prozess auf (Ω, \mathcal{A}, P) .

$T_B := \inf \{ n \geq 1 \mid X_n \in B \}$ erste Eintrittszeit zum Zustand B

$$\Rightarrow E[T_B; X_0 \in B] = P[T_B < \infty]$$

Beweis

$$E[T_B; X_0 \in B] = \sum_{k=0}^{n-1} P[T_B > k, X_0 \in B]$$

$$P[X_0 \in B, X_1 \notin B, \dots, X_k \notin B]$$

$\underbrace{\hspace{10em}}_{X_{k+1}}$

$$= \sum_{k=0}^{n-1} P[X_{k+1} \notin B, X_{k+2} \notin B, \dots, X_n \notin B]$$

wieder Beweis zu Zeit $n-k$
 da stationär

$$= P[T_B \leq n] \quad n \rightarrow \infty \Rightarrow P[B] \neq \emptyset$$

Bem.

1) Insbesondere $T_B < \infty$ P.f. auf $\{X_0 \in B\}$

BS gilt sogar $X_n \in B \Rightarrow$ oft P.f. auf $\{X_0 \in B\}$

Beweis: Wende Aussage auf $(X_{n+k})_{n \geq 0}$ an, wobei stationär.

REKURRENZ!

2) Allgemein folgt für $A \in \mathcal{A}$:

von $\omega \in \Omega$ $\Theta^n(\omega) = (X_n(\omega), X_{n+1}(\omega), \dots) \in A$ \Leftrightarrow oft für P.f. $\omega \in A$

$$\text{Lac } E[\tau_A; A] = P[\tau_A < \infty]$$

$$\tau_A^{(\omega)} := \inf \{ n \geq 1 \mid \Theta^n(\omega) \in A \}$$

Zusatz:

$\leftarrow = 1$ (falls ergodisch)

(265)

$$E[T_A | A] = \frac{P[T_A < \infty]}{P(A)} \quad \text{falls } P(A) \neq 0$$

ZS.5: Anwendung auf invariante Verteilungen von MK

S abz., (X_n, P_x) MK, $T_x := \min\{n \geq 1 \mid X_n = x\}$

Wsp. μ invariante WV

$$\Rightarrow \mu(x) \cdot E_x[T_x] = P_\mu[T_x < \infty]$$

Index:

1) Alle $x \in S$ mit $\mu(x) > 0$ sind positiv rekurrent, d.h. $E_x[T_x] < \infty$

2) S irreduzibel und rekurrent

$$\Rightarrow \mu(x) = \frac{1}{E_x[T_x]} \quad \forall x \in S$$

Zusatz: Eindeutigkeit der invar. WV bei Irreduzibilität

Beweis:

$$\underbrace{E_\mu[T_x \mid X_0 = x]}_{\mu(x) \cdot E_x[T_x]} \stackrel{Wsp.}{=} \underbrace{P_\mu[T_x < \infty]}_{=1}$$

Bem: S transient (oder nullrekurrent, d.h. $E_x[T_x] = \infty$)

$\Rightarrow \nexists$ invariante WV

Invariante Maße können existieren

z.B. Zählmaß bei r.w. auf \mathbb{Z}

Bemerkung Reversibilität und Stationarität

P_μ stationär (shift-invariant) $\Leftrightarrow \mu$ invariantes Maß $\Leftrightarrow \mu P = \mu$

diskret $\sum_x \mu(x) p(x,y) = \mu(y) \quad \forall y$

P_μ invariant unter Zeitumkehr \Leftrightarrow

$P_\mu \circ (X_0, \dots, X_n)^{-1} = P_\mu \circ (X_n, \dots, X_0)^{-1} \quad \forall n$

$\hookrightarrow \mu \circ P \circ P^{-1}$

$\Leftrightarrow \mu \circ P$ invariant unter $(y,x) \mapsto (x,y)$

$\Leftrightarrow \mu$ reversibel (Shift-) Verteilung

diskret

$\mu(x) p(x,y) = \mu(y) p(y,x) \quad \forall x,y$

"detailed balance"

P invariant unter Zeitumkehr

$\sim (X_n, X_{n-1}, \dots, X_0)$

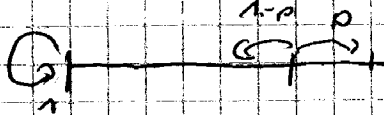
$\sim (X_n, X_{n-1}, \dots, X_0)$

$\sim (X_0, X_1, \dots, X_n)$

μ reversibel \Rightarrow invariant

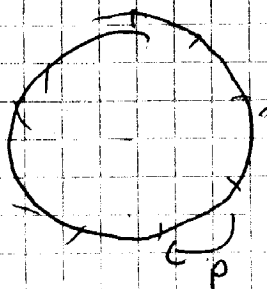
Bsp:

1) RW mit Absorption in 0



Mit δ_0 ist invariantes Maß

2)



$S = \mathbb{Z}/n\mathbb{Z}$

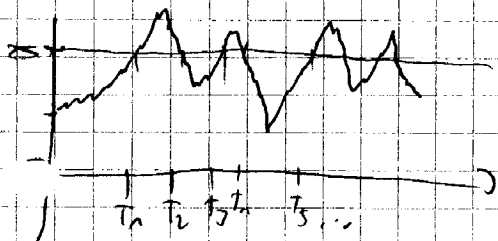
Stationäres μ invariant \Leftrightarrow reversibel $\Leftrightarrow p = \frac{1}{2}$

$T_x < \infty$ P_x -f.s. da x rekurrent

$$\rightarrow \sum_{n=0}^{T_x-1} I_A(X_n) = \sum_{n=0}^{T_x} I_A(X_n) = \sum_{n=0}^{T_x-1} I_A(X_{n+1}) \quad \text{da } x_0 = x_{T_x} = x \text{ } P_x\text{-f.s.}$$

$$\begin{aligned} \Rightarrow \mu(A) &= E_x \left[\sum_{n=0}^{T_x-1} I_A(X_{n+1}) \right] = \sum_{n=0}^{\infty} P_x [X_{n+1} \in A; \overbrace{T_x > n}^{< \infty}] \\ &= \sum_{n=0}^{\infty} E_x [P(X_{n+1} \in A); T_x > n] \\ &= \sum_{z \in S} \underbrace{\sum_{n=0}^{\infty} P_x [X_n = z; T_x > n]}_{\mu(z)} \cdot P(z, A) \\ &= \sum P(A) \end{aligned}$$

Beweis 26.1 (Blatt 1)



$$\sum f(X_n)$$

$T^k := k$ -te Begegnung mit x $\rightarrow \infty$ P_x -f.s.

$$T^1 = T_x, \quad T^{k+1} = T^k + T_{x \circ \Theta_{T^k}}$$

$$\sum_{j=0}^{T^k} f(X_j) = \sum_{i=0}^{T^1} f(X_i) + \sum_{k=0}^{k-1} \underbrace{\sum_{T^k \leq j < T^{k+1}} f(X_j)}_{z^i = y_k}$$

263
25.6. Invariante Maße bei Rekurrenz - negative Form

Sabzählbar

Satz 1) Sei $x \in S$ rekurrenz. Dann ist

$$\mu(A) := E_x \left[\sum_{n=0}^{T_A-1} I_A(X_n) \right] = \sum_{n=0}^{\infty} P_x[X_n \in A; T_A > n]$$

(Mittlere Zahl der Besuche in A während einer Exkursion von x)

ein invariantes Maß.

2) Ist x positiv rekurrenz, dann gilt $\mu(S) = E_x[T_x] < \infty$,

und

$$\bar{\mu}(A) := \frac{\mu(A)}{E_x[T_x]} = \frac{\text{mittlere Zeit in A}}{\text{mittlere Exkursionsdauer}}$$

ist eine invariante WV.

Kor. Für $x \in S$ sind äquivalent:

- (i) \exists invariante WV μ mit $\mu(x) > 0$
- (ii) x ist positiv rekurrenz.
- (iii) Alle $y \sim x$ sind positiv rekurrenz.

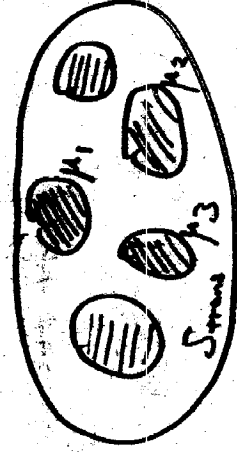
In diesem Fall ex. genau eine invar. WV μ mit $\mu(y) > 0 \forall y \sim x$

Es gilt
$$\mu(y) = \frac{E_x \left[\sum_{n=0}^{T_y-1} I_{\{y\}}(X_n) \right]}{E_x[T_y]}$$

$\forall y, z \sim x$

$$= \frac{1}{E_x[T_y]}$$
 falls $z = y$

Alle invariante WV μ erhält man durch Bilden aller Konvergenz-kombinationen der eindeutige invar. WV μ zu allen positiv rekurrenz. Rekurrenzklassen.



positiv rekurrenz nullrekurrenz

$$\mu = \sum \alpha_i \mu_i, \quad \alpha_i \geq 0, \quad \sum \alpha_i = 1$$

263

25.7. Existenz von invarianten Verteilungen auf allg. Zustandsräume
 Polische Raum

Erinnerung: $P(x, dy)$ Feller: $\Leftrightarrow P \in C_0(S) \forall f \in C_0(S)$

$(\Leftrightarrow P \in C(S) \forall f \in C(S) \text{ falls Sp.})$

Satz 1 Skompakt, $P(x, dy)$ Feller $\Rightarrow \exists$ i.var. WV

Bem. Skompakt $\Rightarrow M_1(S)$ kompakt bzgl. schwaches Konvergenz

denn: $M_1(S) \cong$ positive lineare Funktionale $\mu: C(S) \rightarrow \mathbb{R}$ mit $\mu(1) = 1$

$\mu \in (M_1(S)) \Leftrightarrow \mu \in \Pi [\| \mu \|_\infty, \| \mu \|_1]$ kompakt

$M_1(S)$ abgeschlossen \Rightarrow auch kompakt

Beweis Satz: Sei $\mu \in M_1(S)$, $\mu_n = \frac{1}{n} \sum_{i=0}^{n-1} \nu_i$

Skp. $\Rightarrow M_1(S)$ kp. $\Rightarrow \exists$ Häufungspunkt $\mu = \lim \mu_{n_k}$

Jedes solche Häufungspunkt ist invariant: Sei $f \in C(S)$.

$\int f d(\mu) = \int f d(\mu) = \lim_{n \rightarrow \infty} \int f d(\nu_n) = \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=0}^{n-1} \int f d(\nu_i)$
Folgen
schon konst.
 stetig wegen Feller

27.1

$$= \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n \int f d(\nu_i) = \int f d\mu$$

Also auch $\mu P = \mu$ \square

Nichtkompakter Fall:

$\{ \mu_n | n \in \mathbb{N} \} \subseteq M_1(S)$ kompakt \Leftrightarrow streckt

(d.h. $\forall \epsilon > 0 \exists K \subseteq S$ kompakt: $\sup_n \mu_n(K^c) < \epsilon$)

Satz 2 Ist $P(x, dy)$ Feller und existiert $k: S \rightarrow \mathbb{R}$ mit

(i) $Pk \leq k$ (superharmonisch)

(ii) $\{k \leq c\}$ kompakt $\forall c \in \mathbb{R}$

dann existiert eine invariante WV.

Bem. Vergleiche analoges Rekurrenz-kriterium in 24.8. Dieses notwendige

Kriterium liefert Existenz einer invar. WV - also im abzählbaren

Fall positive Rekurrenz $\nabla \varphi$ $h =$ stochastische Lyapunovfunktion

Beweis: $\nu^n (h > c) \leq \frac{1}{c} \int h d(\nu^n) = \frac{1}{c} \int P^n h d\nu \leq \frac{1}{c} \int h d\nu$

Also gilt für $\mu = \frac{1}{n} \sum_{i=1}^n \nu_i$: $\sup_n \mu_n (h > c) \leq \frac{1}{c} \int h d\nu \xrightarrow{c \rightarrow \infty} 0$.

$\Rightarrow \{h \leq n\}$ streckt $\Rightarrow \square$ T. 0.01

27.2

273

Nach viel angestrengt gilt:

26.3. Birkhoff's individueller Ergodensatz

Sei $(S, \mathcal{O}, P, \theta)$ dynamisches System (d.h. $P \circ \theta^{-1} = P$),
z.B. $P = P_n$ stationäre Verteilung einer MK

Satz $\frac{1}{n} \sum_{i=1}^n F(\theta^i(\omega)) \rightarrow E[F|I_\theta](\omega)$ P-f.s.

für alle $F \in \mathcal{Z}(S, \mathcal{O}, P)$

wobei $\mathcal{Z} = \{A \in \mathcal{O} \mid \theta^{-1}(A) = A\}$

σ -Algebra der shift-invarianten Ereignisse

insbesondere ist θ ergodisch, d.h. $P[A] \in \{0, 1\} \forall A \in \mathcal{Z}$,
d.h. folgt

$$\frac{1}{n} \sum_{i=1}^n F \circ \theta^i \rightarrow E[F] \quad P\text{-f.s.}$$

Beispiel P stat. Vert. einer MK, $F(\omega) = f(x_0, x_1, \dots, x_T) \Rightarrow$ 26.2

Beweis: s. B. Durrett: Probability

272

26.2. Vollgitterierung

Satz: S irreduzibel, P invariante WV

$$\Rightarrow \frac{1}{n} \sum_{i=1}^n f(x_i, x_{i+1}, \dots, x_{i+r}) \xrightarrow{n \rightarrow \infty} \int f(x_0, x_1, \dots, x_r) P(x_0, dx_1) \dots P(x_{r-1}, dx_r)$$

P_0 -f.s. $\forall v_0$

Beweis: Wende 26.1. auf $\tilde{X}_i := (X_i, X_{i+1}, \dots, X_{i+r}) \in S^{r+1}$ an.

Ist messbare MK mit invar. WV

$$P_n(x_0, \dots, x_r) = p(x_0) p(x_0, x_1) \dots p(x_{r-1}, x_r)$$

□

KONSEQUENZ: Statistik der Markovkette

$$\frac{1}{n} \sum_{i=1}^n \mathbb{I}_{\{X_i = x, X_{i+1} = y\}} \xrightarrow{n \rightarrow \infty} P(x, y)$$

Rekonstruktion des Übergangskerns aus Beobachtungsdaten!

26. DER ERGODENSATZ

26.1. Ergodensatz für diskrete Markovketten Subäquivalenz

(S. 07, $(P_n)_{n \in \mathbb{N}}$, $(X_n)_{n \geq 0}$) kanon. MK. mit Zustandsraum S

$$B_x^{(n)} := \sum_{i=1}^n I_{\{X_i = x\}} \quad \text{Besuche in } x \text{ bis zur Zeit } n$$

$x \in S$ fest

SATZ Sei S irreduzibel.

1) Ist die MK relaxiert, dann gilt für $f: S \rightarrow \mathbb{R}$, $f > 0$:

$$\boxed{\frac{1}{B_x^{(n)}} \sum_{i=1}^n f(X_i) \xrightarrow{n \rightarrow \infty} \int f d\mu \quad P_0\text{-f.s. } \forall 0}$$

wobei $\mu(A) = E_x \left[\sum_{i=1}^{\tau_A} I_A(X_i) \right]$ (invar. Maß)

2) Existiert eine invariante WV $\bar{\mu}$, dann gilt

$$\boxed{\frac{1}{n} \sum_{i=1}^n f(X_i) \xrightarrow{n \rightarrow \infty} \int f d\bar{\mu} \quad P_0\text{-f.s. } \forall 0}$$

Bem.

1) Spezieller **ERGODENSATZ**:

Zeitmittel \rightarrow Raummittel

$$2) f = I_A \Rightarrow \underbrace{\frac{1}{n} \sum_{i=1}^n I_A(X_i)}_{\text{pos. rel.}} \rightarrow \bar{\mu}(A)$$

relative Häufigkeit
der Besuche in A

3) Anwendung: Markov chain Monte Carlo Methode
Zur Berechnung (numerisch) von
Erwartungswerten bezgl. $\bar{\mu}$.

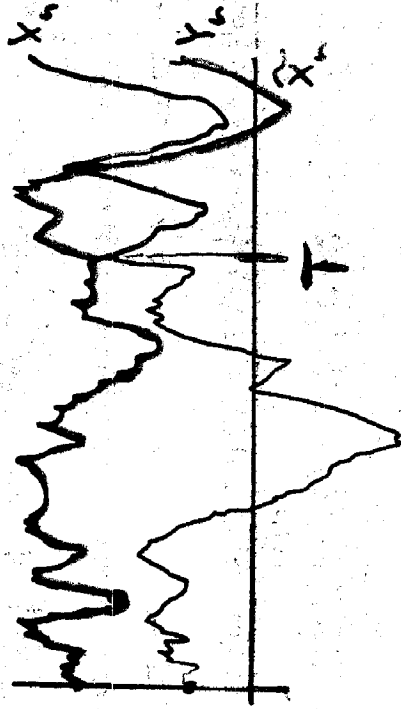
276

27.2. Unabhängige Koppeln von Markovketten

Satz Seien $(X_n)_{n \geq 0}$ und $(Y_n)_{n \geq 0}$ unabhängige M.K. auf (Ω, \mathcal{F}, P) mit demselben Übergangskern, und sei

$$X_n := \begin{cases} X_n & \text{für } n \leq T \\ Y_n & \text{für } n > T \end{cases} \quad \text{"Kopplungszeit"}$$

Dann ist $((X_n)_{n \geq 0}, (Y_n)_{n \geq 0})$ eine Koppeln von P_{X_0} und P_{Y_0} mit $X_n = Y_n \forall n \geq T$.



Beweis: (X_n, Y_n) MK auf $\mathcal{S} \times \mathcal{S}$, T Stoppzeit

$$\begin{aligned} \text{Starke ME.} &\Rightarrow (Y_{T+n})_{n \geq 0} \sim (X_{T+n})_{n \geq 0} \text{ unabh. von } \mathcal{F}_T \\ &\Rightarrow (\bar{X})_{n \geq 0} \sim (X_n)_{n \geq 0} \end{aligned}$$

277

Nachtrag: Periodizität von Markovketten

Def. $x \in S$ heißt aperiodisch, falls

$$\text{Periode}(x) := \text{ggT} \{n \geq 1 \mid p^n(x,x) > 0\} = 1$$



Bem. 1) Recht der Zeitdiskretisierung: Zeitstetige MK sind immer aperiodisch

2) Periode $(x) = d \Rightarrow p^{n+d}(x,x) = 0 \forall n \notin d\mathbb{Z}$

3) x aperiodisch $\Rightarrow \exists n_0 \forall n \geq n_0 : p^n(x,x) > 0$

Beweis: $p^n(x,x) > 0, p^m(x,x) > 0 \Rightarrow p^{n+m}(x,x) > 0$

jetzt Übung in Zähltheorie

Dann Vollständigkeit des Borel's in 27.3

$x, y \in S$

$$\exists n, p^n(x,y) > 0 \quad \text{aperiod.}$$

$$\exists n, p^{n+k}(x,y) > 0 \Rightarrow p^n(x,y) \geq p^n(x,y) p^k(y,y) > 0$$

für alle $k \geq k_0$

$$\Rightarrow \bar{P}^n((y_1, y_2), (y_1, y_2)) = p^n(x_1, y_1) \cdot p^n(x_1, y_2) > 0$$

Also p imed. + aperiod. $\Rightarrow \bar{P}$ imed. für alle n groß genug

Ben. L^2 -Konvergenz ist von funktionalanalytischer Sicht (Neuman 1992):

Θ induziert Isometrie $UF = F \circ \Theta$ auf Hilbertraum L^2

$$\frac{1}{n} \sum_{i=1}^n F \circ \Theta^i = \frac{1}{n} \sum_{i=1}^n U^i F$$

$\xrightarrow{n \rightarrow \infty}$ Projektion von F auf $\langle G \in L^2 | UG = G \rangle$
 $= E[F | \mathcal{G}]$

F -f.s. Konvergenz ist schwieriger zu zeigen

~~27~~ ~~28~~

27. KONVERGENZ UND GLEICHGEWICHT ASYMPTOTISCHE STATIONARITÄT

27.1. Kopplung von WVN

X auf (Ω, \mathcal{F}, P) , Y auf $(\Omega', \mathcal{F}', P')$ Zufallsvariablen
(Verteilung μ bzw. ν)

Def. Eine Kopplung von X und Y (bzw. von μ und ν) ist ein Paar (\bar{X}, \bar{Y}) von auf einem gemeinsamen Wahrraum $(\bar{\Omega}, \bar{\mathcal{F}}, \bar{P})$ definierten Zufallsvariablen mit

$$\bar{P} \circ \bar{X}^{-1} = \mu, \quad \bar{P} \circ \bar{Y}^{-1} = \nu$$

Satz Für jede Kopplung von μ und ν gilt:

$$(*) \quad \|\mu - \nu\|_{\text{var}} \leq P[\bar{X} \neq \bar{Y}]$$

Beweis: $\mu(A) - \nu(A) = P[\bar{X} \in A] - P[\bar{Y} \in A]$

$$\leq P[\bar{X} \in A, \bar{Y} \notin A] \leq P[\bar{X} \neq \bar{Y}] \quad \forall A \subseteq \Omega$$

Ben. Es existiert eine Kopplung mit $\|\mu - \nu\|_{\text{var}}$ in $(*)$ [maximale Kopplung]

Bearbeitung:

$$Y_n = Y_0 + \theta T^2, \quad Y = \sum_{i=1}^{T_n} f(x_i)$$

Insbesondere Y_n iid, $P_{Y_n} \circ Y_n^{-1} = P_X \circ Y^{-1}$

Revers: $Y_n = \sum_{j=1}^{T_n \circ \theta_n} f(x_i) = \sum_{j=1}^{T_n \circ \theta_n} f(x_{T_n \circ \theta_n} \circ j) = Y_0 \circ \theta_n$

SMB \Rightarrow iid Verteilung

~~Wdh~~
 $E[Y_n] = E_X \left[\sum_{i=1}^{T_n} f(x_i) \right] = \int f d\mu$

oder falls $f = I_A$
siehe math. Induktion

Gdgl: $\Rightarrow \frac{1}{l} \sum_{i=1}^l Y_n \rightarrow \int f d\mu$ RJ-G.S.

$$\Rightarrow \frac{1}{l} \sum_{i=1}^l f(x_i) \xrightarrow{l \rightarrow \infty} \int f d\mu \quad \text{RJ-G.S.}$$

$T_n \in n \subset T_m \Rightarrow P_X(m) = l$

$$\underbrace{\frac{1}{l} \sum_{i=1}^l f(x_i)}_{\rightarrow \int f} \in \frac{1}{P_X(m)} \sum_{i=1}^m f(x_i) \in \underbrace{\frac{1}{l} \sum_{i=1}^{T_m} f(x_i)}_{\int f}$$

2) Erhaltungssatz $\Rightarrow \mu(A) = \frac{\mu(S)}{\mu(S)}$

nach μ mit $f \equiv 1$ gilt

$$\frac{1}{P_X(m)} \rightarrow \int 1 d\mu = \mu(S)$$

$$\Rightarrow \frac{1}{n} \sum_{i=1}^n f(x_i) = \frac{B_n}{n} \cdot \frac{1}{B_n} \sum_{i=1}^n f(x_i) = \int f d\mu$$

27.2

$$\|P \circ X_n^{-1} - P \circ Y_n^{-1}\|_{\text{var}} \stackrel{27.1}{\leq} P[\tilde{X}_n + Y_n] \\ = P[T > n]$$

insbes. $\rightarrow 0$ falls $T < \infty$ A.f.s.

~~Satz~~ ~~ist S irreduzibel~~

27.3 Konv. vs. Gleichheit

S abz. p-erg., μ invariant

Satz: Ist S irreduzibel und p-aperiodisch
(\Rightarrow \exists f, g Funktionen $p^n(x, y) > 0$)

dann $\|P^n - \mu\|_{\text{var}} \rightarrow 0$

Konsequenzen

- 1) MCMC Methode zur Simulation von μ
- 2) Asymptotische Stabilität

Beweis: X_n, Y_n auf (Ω, \mathcal{A}, P) unabh. MK mit Lawa p und stationär \sim Gew μ .

$$T := \inf \{k \geq 0 \mid X_k = Y_k\}$$

$$\|P^n - \mu\|_{\text{var}} \stackrel{27.2}{\leq} P[T > n] \rightarrow 0 \quad \text{falls } T < \infty \text{ A.f.s.}$$

$\int \mu^n$

In der Tat

(x_1, y_1) MUE auf $S \times S$ mit

2872

$$\bar{p}((x_1, y_1), (x_2, y_2)) = p(x_1, y_1) p(x_2, y_2)$$

- Irreduzibel
- Invariante WV

} \Rightarrow pos. rekurrent

$\rightarrow T$ (Trajektor der Dinsomale) $\in \infty$ Pf.-S.

irreduzibel, da p irreduzibel und aperiodisch

WV: $\bar{\mu} = \mu \otimes \mu$

22.4: Irreversibilität und Rekurrenz

Boltzmanns Thermodynamik Irreversibilität

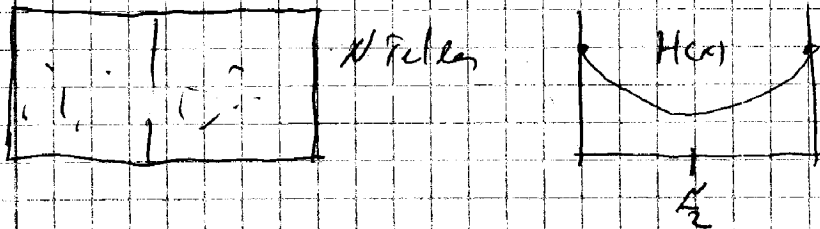
Entropie fällt ($H(x) = -\log p(x)$)

Zermelo: Widerspruch zu Reversibilität in der stat. Mechanik?

\Rightarrow Rekurrenz (Wiederkehrsatz von Poincaré)

Boltzmanns soll solange warten!

Kollision im Ehrenfestmodell



mittlere Zeit für

$$\left[\begin{array}{c} 1 \\ \vdots \\ 1 \end{array} \right] \rightarrow \left[\begin{array}{c|c} \frac{1}{2} & \frac{1}{2} \\ \hline \frac{1}{2} & \frac{1}{2} \end{array} \right] \sim N \log N$$

$$\leftarrow \sim \frac{1}{2N} 2^{2N}$$

Rekurrenzzeit von $\frac{1}{2}$ nicht beobachtbar

Irreversibilität im Substitutionsprozess

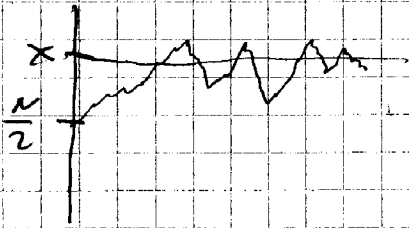
(28/3)

Loschmidt's reversibel

Wenn es immer fällt, muss es ^{auch} immer steigen

→ Modifikation: für $H(x) \gg H_{min}$ geht es fast immer nach unten

Beispiel:



Wie kann es sein, dass es bei Level x oder nach unten geht?
↳ zur Reversibilität?

Antwort:

Für große x ist bei Erreichen des Levels x ^{benach} immer ein
Min. Maximum

M. Kac: Random walks & the theory of Brownian motion
Ann. Math. Monthly 54, 369 ff