Konvergenzraten des Robbins-Monro-Algorithmus

Alexander Schlüter

13. Juli 2016

Der Robbins-Monro-Algorithmus ist ein iteratives Verfahren zur Bestimmung von Nullstellen einer Funktion, deren Wert nur stochastisch gestört gemessen werden kann. Im letzten Vortrag wurde die fast sichere Konvergenz gegen eine Nullstelle bewiesen. Ziel dieses Vortrages ist es, eine Aussage über die Konvergenzgeschwindigkeit und die Form der Verteilung im Limes ähnlich dem zentralen Grenzwertsatz zu beweisen.

Inhaltsverzeichnis

1 Erinnerungen

Im Folgenden sei

- $(\Omega, \mathcal{F}, P)$ ein Wahrscheinlichkeitsraum und $(V, \mathcal{V})$ ein messbarer Raum,
- $(Z_n)_{n\geq 1}$ eine Folge u.i.v $(V, \mathcal{V})$-wertiger Zufallsvariablen,
- $I \subset \mathbb{R}$ ein abgeschlossenes Intervall oder $I = \mathbb{R}$,
- X_0 eine von $(Z_n)_{n\geq 1}$ unabhängige I-wertige Zufallsvariable,
- $\mathcal{F}_n = \sigma(X_0, Z_1, \ldots, Z_n)$ und $\mathbb{F} = (\mathcal{F}_n)_{n\geq 0}$ eine Filtration,
- $(\gamma_n)_{n\geq 1}$ eine Folge in $(0, \infty)$ und
- $H : (I \times V, \mathcal{B}(I) \otimes \mathcal{V}) \to (\mathbb{R}, \mathcal{B}(\mathbb{R}))$ eine messbare Abbildung mit
$$H(x, Z_1) \in \mathcal{L}^1 \quad \text{für alle } x \in I.$$

Der **Robbins-Monro-Algorithmus** ist der $\mathbb{F}$-adaptierte reelle Prozess $X = (X_n)_{n\geq 0}$ definiert durch die Rekursion

$$X_{n+1} = X_n + \gamma_{n+1} H(X_n, Z_{n+1})\,, \tag{1}$$

wobei wir annehmen, dass die Werte immer in I liegen.

Die **Erwartungswertfunktion** des Algorithmus ist definiert durch

$$h : I \to \mathbb{R}, \quad h(x) \coloneqq \mathbb{E}[H(x, Z_1)]$$

und außerdem sei

$$g : I \to \overline{\mathbb{R}}_+, \quad g(x) \coloneqq \mathbb{E}[H(x, Z_1)^2].$$

Bemerkung 1.1. Wir haben gesehen, dass der Robbins-Monro-Algorithmus unter zusätzlichen Annahmen fast sicher gegen eine Nullstelle von h konvergiert. Diese Annahmen waren

(a) (Downcrossing-Bedingung) h ist stetig und es gibt ein $x_0 \in h^{-1}(0)$ mit

$$\sup_{x\in I}(x - x_0)h(x) \leq 0\,,$$

(b) (Beschränkte Störung) $X_0 \in \mathcal{L}^2$ und $g(x) \leq C(1 + x^2)$ für eine Konstante $C \in \mathbb{R}_+$ und alle $x \in I$,

(c) (Abnehmende Schrittweiten) $\sum_{n=1}^\infty \gamma_n = \infty$ und $\sum_{n=1}^\infty \gamma_n^2 < \infty$.

Wir benötigen noch folgende Resultate vom letzten Mal:

Lemma 1.2 (Doob-Zerlegung)**.** Unter den o.g. Bedingungen ist X ein $\mathcal{L}^2$-Prozess und für die Doob-Zerlegung $X = N + A$ in ein Martinal $N = (N_n)_{n\geq 0}$ und einen previsiblen Prozess $A = (A_n)_{n\geq 0}$ gilt

$$\begin{aligned} N_n &= X_0 + \sum_{j=1}^n \gamma_j (H(X_{j-1}, Z_j) - h(X_{j-1})) \\ \langle N\rangle_n &= \sum_{j=1}^n \gamma_j^2 (g(X_{j-1}) - h(X_{j-1})^2) \end{aligned}$$

für $n \geq 0$. Außerdem ist N ein $\mathcal{L}^2$-Martingal.

Lemma 1.3 (Substitution bei Unabhängigkeit)**.** Für $n \geq 1$ und jede messbare Funktion $f : (I \times V, \mathcal{B}(I) \otimes \mathcal{V}) \to (\mathbb{R}, \mathcal{B}(\mathbb{R}))$ mit wohldefiniertem Integral bezüglich $P^{(X_{n-1},Z_n)}$ gilt

$$\mathbb{E}[f(X_{n-1}, Z_n)\,|\,\mathcal{F}_{n-1}](\cdot) = \int f(X(\cdot), z)\,\mathrm{d}P^{Z_1}(z) \quad \text{fast sicher.}$$

Beweis. Z_n ist unabhängig von $\mathcal{F}_{n-1}$, also ist die Aussage vom letzten Mal anwendbar. $\square$

2 Vorbereitungen

Definition 2.1. Es sei $(Y_n)_{n\geq 1}$ eine Folge von Zufallsvariablen von $(\Omega, \mathcal{F})$ nach $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$ und ν ein Wahrscheinlichkeitsmaß auf $\mathcal{B}(\mathbb{R}^d)$. $(Y_n)_{n\geq 1}$ heißt **mischend konvergent** gegen ν und wir schreiben

$$Y_n \longrightarrow \nu \quad \textit{mischend},$$

falls für alle $f \in \mathcal{L}^1(P)$ und $h \in C_b(\mathbb{R}^d)$ gilt

$$\lim_{n\to\infty} \mathbb{E}[f \cdot h(Y_n)] = \int f \,\mathrm{d}P \int h \,\mathrm{d}\nu \,.$$

Bemerkung 2.2. Durch Wahl von $f = 1$ folgt aus der mischenden Konvergenz die bekannte Konvergenz in Verteilung.

Lemma 2.3. Sei $(Y_n)_{n\geq 1}$ eine Folge von reellen Zufallsvariablen und ν ein Wahrscheinlichkeitsmaß auf $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ mit $Y_n \to \nu$ *mischend* für $n \to \infty$. Außerdem sei $(U_n)_{n\geq 1}$ eine Folge von reellen Zufallsvariablen mit $U_n \to 0$ *in Wahrscheinlichkeit.* Dann gilt

$$(Y_n, U_n) \to \nu \otimes \delta_0 \quad \textit{mischend.}$$

Beweis. Siehe [LPW08, Korollar 5.29(a)]. □

Korollar 2.4. Sei $(Y_n)_{n\geq 1}$ eine Folge von reellen Zufallsvariablen und ν ein Wahrscheinlichkeitsmaß auf $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ mit $Y_n \to \nu$ *mischend* für $n \to \infty$. Außerdem sei $(U_n)_{n\geq 1}$ eine Folge von reellen Zufallsvariablen mit $U_n \to 0$ *in Wahrscheinlichkeit.* Dann gilt

$$Y_n + U_n \to \nu \quad \textit{mischend.}$$

Beweis. Es seien $f \in \mathcal{L}^1(P)$ und $h \in C_b(\mathbb{R})$ beliebig. Dann ist $h' : \mathbb{R}^2 \to \mathbb{R}$, $h'(y,z) := h(y+z)$ in $C_b(\mathbb{R}^2)$. Nach Lemma 2.3 gilt

$$(Y_n, U_n) \to \nu \otimes \delta_0 \quad \textit{mischend.}$$

Nach Definition der mischenden Konvergenz folgt

$$\begin{aligned}\mathbb{E}[f \cdot h(Y_n + U_n)] = \mathbb{E}[f \cdot h'(Y_n, U_n] &\to \int f \,\mathrm{d}P \int h'(y,u) \,\mathrm{d}\nu \otimes \delta_0(y,u) \\ &= \int f \,\mathrm{d}P \int h(y+0) \,\mathrm{d}\nu(y)\,.\end{aligned}$$

□

Satz 2.5 (Stabiler CLT)**.** Es sei $Y = (Y_n)_{n\geq 0}$ ein $\mathcal{L}^2$-Martingal bezüglich der Filtration $\mathbb{F}$ und $(\kappa_n)_{n\geq 1}$ eine Folge in $(0,\infty)$ mit $a_n \to 0$. Die folgenden Bedingungen seien erfüllt:

(a) Es existiert eine Konstante $v > 0$ mit

$$\kappa_n^2 \langle Y \rangle_n \longrightarrow v \quad \textit{in Wahrscheinlichkeit}$$

für $n \to \infty$,

(b) ($\mathbb{F}$-bedingte Lyapunov-Bedingung) Es existiert ein $\delta > 0$, sodass

$$\kappa_n^{2+\delta} \sum_{j=1}^{n} \mathbb{E}[|\Delta Y_j|^{2+\delta} \,|\, \mathcal{G}_{j-1}] \to 0 \quad \textit{in Wahrscheinlichkeit}$$

für $n \to \infty$.

Dann gilt für $n \to \infty$:

$$\kappa_n Y_n \to \mathcal{N}(0, v) \quad \textit{mischend.}$$

Beweis. Siehe [LPW08, Satz 5.31 und Bemerkung 5.32(a)]. □

Lemma 2.6 (Diskrete Regel von de l'Hospital)**.** Sei $(a_n)_{n\geq 1}$ eine Folge in $\mathbb{R}$ und $(b_n)_{n\geq 1}$ eine Folge positiver reeller Zahlen, sodass

(a) $\sum_{n=1}^{\infty} b_n = \infty$ und
(b) es existiert $c \in \mathbb{R}$ mit $\lim_{n\to\infty} a_n/b_n = c$

Dann gilt

$$\lim_{n\to\infty} \frac{\sum_{j=1}^{n} a_j}{\sum_{j=1}^{n} b_j} = c\,.$$

Beweis. Seien $\varepsilon > 0$ und $n_0 \in \mathbb{N}$ mit $|a_n/b_n - c| \leq \epsilon$ für alle $n > n_0$. Dann gilt für $n > n_0$

$$\begin{aligned}
\left| \frac{\sum_{j=1}^{n} a_j}{\sum_{j=1}^{n} b_j} - c \right| &\leq \frac{\sum_{j=1}^{n} b_j |a_j/b_j - c|}{\sum_{j=1}^{n} b_j} \\
&= \frac{\sum_{j=1}^{n_0} b_j |a_j/b_j - c|}{\sum_{j=1}^{n} b_j} + \frac{\sum_{j=n_0+1}^{n} b_j |a_j/b_j - c|}{\sum_{j=1}^{n} b_j} \\
&\leq \frac{\sum_{j=1}^{n_0} b_j |a_j/b_j - c|}{\sum_{j=1}^{n} b_j} + \varepsilon\,.
\end{aligned}$$

Es folgt

$$\limsup_{n\to\infty} \left| \frac{\sum_{j=1}^{n} a_j}{\sum_{j=1}^{n} b_j} - c \right| \leq \varepsilon$$

und da ε beliebig war folgt die Behauptung. □

Lemma 2.7 (Kronecker, WT Lem. 4.28)**.** Sei $(a_n)_{n\geq 1}$ eine monoton wachsende Folge positiver Zahlen mit $\lim_{n\to\infty} a_n = \infty$ und sei $(b_n)_{n\geq 1}$ eine Folge reeller Zahlen. Konvergiert $\sum_{j=1}^{\infty} b_j/a_j$ in $\mathbb{R}$, so folgt

$$\lim_{n\to\infty} \frac{1}{a_n} \sum_{j=1}^{n} b_j = 0\,.$$

Lemma 2.8. (i) Für reelle Konstanten $b \geq 0$ und $a > -b - 1$ existiert $L \in (0, \infty)$, sodass

$$\prod_{j=1}^{n} \left(1 + \frac{a}{b+j}\right) \sim L n^a \quad \text{für } n \to \infty.$$

(ii) Für $b > -1$ gilt

$$\sum_{j=1}^{n} j^b \sim \frac{n^{b+1}}{b+1} \quad \text{für } n \to \infty$$

und für $b = -1$

$$\sum_{j=1}^{n} j^{-1} \sim \log n \quad \text{für } n \to \infty.$$

Beweis. (i) Die Gammafunktion hat die Eigenschaft

$$\Gamma(t+1) = t\Gamma(t)$$

für eine positive reelle Zahl t. Deshalb gilt

$$\prod_{j=1}^{n} \left(1 + \frac{a}{b+j}\right) = \frac{\prod_{j=1}^{n}(a+b+j)}{\prod_{j=1}^{n}(b+j)} = \frac{\Gamma(a+b+n+1)}{\Gamma(a+b+1)} \frac{\Gamma(b+1)}{\Gamma(b+n+1)}$$

Wir identifizieren den von n unabhängigen Faktor

$$L := \frac{\Gamma(b+1)}{\Gamma(a+b+1)}$$

und untersuchen den übrigen Term. Stirlings Formel für die Gammafunktion

$$\Gamma(t) \sim \sqrt{2\pi} t^{t-1/2} e^{-t}, \quad t \to \infty$$

liefert

$$\begin{aligned} \frac{\Gamma(a+b+n+1)}{\Gamma(b+n+1)} &\sim (a+b+n+1)^a \left(\frac{a+b+n+1}{b+n+1}\right)^{b+n+1/2} e^{-a} \\ &\sim n^a \left(1 + \frac{a}{b+n+1}\right)^{b+n+1} e^{-a}. \end{aligned}$$

Im letzten Schritt wurde genutzt, dass

$$\lim_{n\to\infty} \left(\frac{a+b+n+1}{n}\right)^a = 1 \quad \text{und} \quad \lim_{n\to\infty} \sqrt{\frac{b+n+1}{a+b+n+1}} = 1\,.$$

Schließlich folgt mit

$$\lim_{n\to\infty} \left(1 + \frac{a}{b+n+1}\right)^{b+n+1} = e^a$$

die Behauptung.

(ii) Falls $-1 < b \leq 0$, so ist die Abbildung $(0, \infty) \to \mathbb{R}$, $x \mapsto x^b$ monoton fallend. Also gilt für $j \in \mathbb{N}$

$$(j+1) \leq \int_j^{j+1} x^b \,\mathrm{d}x \leq j^b$$

und deshalb

$$\frac{(n+1)^{b+1}-1}{b+1} = \int_1^{n+1} x^b \,\mathrm{d}x \le \sum_{j=1}^{n} j^b \le \int_0^n x^b \,\mathrm{d}x = \frac{n^{b+1}}{b+1}\,.$$

Die Behauptung folgt, da

$$\lim_{n\to\infty} \frac{(n+1)^{b+1}-1}{b+1} \cdot \frac{b+1}{n^{b+1}} = 1\,.$$

Mit demselben Argument bekommt man den Fall $b = -1$:

$$\log(n+1) = \int_1^{n+1} x^b \,\mathrm{d}x \le \sum_{j=1}^{n} j^b \le 1 + \int_1^n x^b \,\mathrm{d}x = 1 + \log(n)$$

und

$$\lim_{n\to\infty} \frac{\log(n+1)}{\log(n)} = 1 = \lim_{n\to\infty} \frac{1+\log(n)}{\log(n)}\,.$$

Falls $b > 0$, so ist $x \mapsto x^b$ monoton wachsend, also

$$j^b \le \int_j^{j+1} x^b \,\mathrm{d}x \le (j+1)^b$$

und die Abschätzung ist analog zu oben.

□

3 Konvergenzraten

Satz 3.1 (Konvergenzraten, stabiler CLT)**.** Seien $\gamma_n = C_1/(C_2+n)$ für $n \ge 1$ mit reellen Konstanten $C_1 > 0$, $C_2 \ge 0$ und $x_0 \in h^{-1}(0)$. Die folgenden Bedingungen seien erfüllt:

(a) $X_n \to x_0$ f.s. für $n \to \infty$.
(b) $X_0 \in \mathcal{L}^2$ und $g(x) \le C(1+x^2)$ für eine Konstante $C \in \mathbb{R}_+$ und alle $x \in I$,
(c) g sei stetig in x_0, h sei differenzierbar in x_0, $h'(x_0) < 0$ und

$$h(x) = h'(x_0)(x-x_0) + \mathcal{O}((x-x_0)^2) \quad \text{für } x \to x_0\,,$$

(d) Es existieren $\varepsilon, \delta > 0$ sodass $H(x, Z_1) \in \mathcal{L}^{2+\delta}$ für alle $x \in I$ und

$$\sup_{|x-x_0|\le\varepsilon} \mathbb{E}[|H(x,Z_1)|^{2+\delta}] < \infty\,.$$

Dann gilt für $n \to \infty$:

(i) Falls $|h'(x_0)|C_1 > 1/2$:

$$\sqrt{n}(X_n - x_0) \to \mathcal{N}\left(0, \frac{g(x_0)C_1^2}{2|h'(x_0)|C_1 - 1}\right) \quad \textit{mischend,}$$

(ii) Falls $|h'(x_0)|C_1 = 1/2$:

$$\sqrt{\frac{n}{\log n}}(X_n - x_0) \to \mathcal{N}\left(0, g(x_0)C_1^2\right) \quad \textit{mischend,}$$

(iii) Falls $|h'(x_0)|C_1 < 1/2$:

$$n^{|h'(x_o)|C_1}(X_n - x_0) \to \xi \quad \text{f.s.}$$

für eine reelle Zufallsvariable ξ, die von X_0 abhängt.

Bemerkung 3.2. Downcrossing etc.

Beweis. Hier wird nur ein Teil vom Fall $|h'(x_0)|C_1 > 1/2$ bewiesen. Für den vollen Beweis siehe [LPW08, Satz 11.4].

Wie wir letzte Woche gesehen haben, folgt aus Voraussetzung (b), dass X ein $\mathcal{L}^2$-Prozess ist. Sei

$$a := -h'(x_0) = |h'(x_o)|\,.$$

Nach (c) gilt $a > 0$. Man wähle $n_0 \in \mathbb{N}$ mit $a\gamma_{n_0} < 1$ und definiere für $n \geq 0$

$$\beta_n := \prod_{j=n_0}^{n}(1 - a\gamma_j) = \prod_{j=1}^{n-n_0+1}(1 - a\gamma_{j+n_0-1})\,.$$

Insbesondere gilt also $\beta_0 = \cdots = \beta_{n_0-1} = 1$ und nach Lemma 2.8(i)

$$\beta_n \sim Ln^{-aC_1} \tag{2}$$

für $n \to \infty$ mit einer Konstanten $L \in (0, \infty)$. Wegen $\gamma_n \sim C_1/n$ folgt

$$\frac{\gamma_n}{\beta_n} \sim \frac{C_1}{L}n^{aC_1-1} \tag{3}$$

Mit der Doob-Zerlegung $X = N + A$ aus Lemma 1.2 bekommen wir

$$\begin{aligned}
\sqrt{n}(X_n - x_0) &= \sqrt{n}\beta_n\left(X_0 - x_0 + \sum_{j=1}^{n}\frac{1}{\beta_j}(X_j - X_{j-1} + X_{j-1}) - \frac{1}{\beta_{j-1}}X_{j-1} - \frac{x_0}{\beta_j} + \frac{x_0}{\beta_{j-1}}\right)\\
&= \quad \sqrt{n}\beta_n(X_0 - x_0)\\
&\quad + \sqrt{n}\beta_n\sum_{j=1}^{n}\frac{1}{\beta_j}\Delta N_j\\
&\quad + \sqrt{n}\beta_n\sum_{j=1}^{n}\left(\frac{1}{\beta_j}\Delta A_j + (X_{j-1} - x_0)(\frac{1}{\beta_j} - \frac{1}{\beta_{j-1}})\right)\,.
\end{aligned} \tag{4}$$

Wir wollen die drei Summanden einzeln auf ihr Verhalten für $n \to \infty$ untersuchen. Definiere dazu für $n \geq 0$

$$\begin{aligned} M_n &:= \sum_{j=1}^{n} \frac{1}{\beta_j} \Delta N_j \\ B_n &:= \sum_{j=1}^{n} \left(\frac{1}{\beta_j} \Delta A_j + (X_{j-1} - x_0)(\frac{1}{\beta_j} - \frac{1}{\beta_{j-1}}) \right) . \end{aligned}$$

Da $\sqrt{n}\beta_n \sim Ln^{-aC_1+1/2} \to 0$ gilt auch $\sqrt{n}\beta_n(X_0 - x_0) \to 0$ punktweise.

Es gilt auch $\sqrt{n}\beta_n B_n \to 0$ fast sicher, allerdings ist der Beweis nicht trivial und kann nachgelesen werden in [LPW08, Satz 11.4]. Er basiert auf der Idee, die Zuwüchse ΔB_n für große n gegen $(X_n - x_0)^2$ abzuschätzen. Dazu wird Voraussetzung (c) genutzt.

Im Folgenden wird mit dem stabile CLT 2.5 gezeigt, dass

$$\sqrt{n}\beta_n M_n \to \mathcal{N}\left(0, \frac{g(x_0)C_1^2}{2aC_1 - 1}\right) \quad \textit{mischend.} \tag{5}$$

Wähle also $\kappa_n = \sqrt{n}\beta_n$, dann gilt nach Gleichung (2) und wegen $aC_1 > 1/2$

$$\kappa_n \sim Ln^{-aC_1+1/2} \to 0 \quad \text{für } n \to \infty.$$

$M = (M_n)_{n \geq 0}$ ist ein $\mathcal{L}^2$-Martingal, da N eins ist.

Wir wollen nun Voraussetzung (a) des CLT überprüfen. Es gilt nach der Darstellung für $\langle N \rangle$ aus Lemma 1.2

$$\begin{aligned} \Delta\langle M \rangle_n = \mathbb{E}[(\Delta M_n)^2 \,|\, \mathcal{F}_{n-1}] &= \frac{1}{\beta_j^2} \mathbb{E}[(\Delta N_n)^2 \,|\, \mathcal{F}_{n-1}] \\ &= \frac{\gamma_j^2}{\beta_j^2} (g(X_{j-1}) - h(X_{j-1})^2) \end{aligned}$$

für $n \geq 0$. Wegen der Stetigkeit von h und g in x_0 gilt

$$g(X_{n-1}) - h(X_{n-1})^2 \to g(x_0) - h(x_0)^2 = g(x_0) \quad \text{f.s.},$$

und zusammen mit Gleichung (3) ergibt sich

$$\frac{\Delta\langle M \rangle_n}{n^{2aC_1-2}} \to \frac{g(x_0)C_1^2}{L^2} \quad \text{f.s.}$$

für $n \to \infty$. Da wir $aC_1 > 1/2$ annehmen, ist $2aC_1 - 2 > -1$ und deshalb

$$\sum_{j=1}^{\infty} j^{2aC_1-2} = \infty \, .$$

Somit können wir die diskrete Regel von l'Hospital 2.6 anwenden und bekommen

$$\frac{\langle M \rangle_n}{\sum_{j=1}^{n} j^{2aC_1-2}} \to \frac{g(x_0)C_1^2}{L^2} \quad \text{f.s.}$$

für $n \to \infty$. Aus Lemma 2.8(ii) und Gleichung (2) kennen wir die asymptotischen Äquivalenzen

$$\frac{1}{\sum_{j=1}^{n} j^{2aC_1-2}} \overset{2.8}{\sim} \frac{2aC_1-1}{n^{2aC_1-1}} \overset{2}{\sim} \frac{2aC_1-1}{L^2} n\beta_n^2$$

d.h. insgesamt

$$\kappa_n^2 \langle M \rangle_n \to \frac{g(x_0)C_1^2}{2aC_1-1} \quad \text{f.s.}$$

Zur Anwendung des CLT fehlt uns noch die bedingte Lyapunov-Bedingung. Es gilt für $j \geq 1$

$$\Delta M_j = \frac{\gamma_j}{\beta_j}(H(X_{j-1}, Z_j) - h(X_{j-1}))\,,$$

also mit dem δ aus Voraussetzung (d)

$$\sum_{j=1}^{n} \mathbb{E}[|\Delta M_j|^{2+\delta} \,|\, \mathcal{F}_{j-1}] \leq \sum_{j=1}^{n} \left(\frac{\gamma_j}{\beta_j}\right)^{2+\delta} \sup_{k\geq 1} \mathbb{E}[|H(X_{k-1}, Z_k) - h(X_{k-1})|^{2+\delta} \,|\, \mathcal{F}_{k-1}]\,. \tag{6}$$

Das Supremum über die bedingten Erwartungswerte kann gebildet werden, weil die Indexmenge abzählbar ist. Für $n \to \infty$ gilt wegen $\gamma_n \sim C_1/n$

$$(\sqrt{n}\beta_n)^{2+\delta} \left(\frac{\gamma_n}{\beta_n}\right)^{2+\delta} = n^{1+\delta/2}\gamma_n^{2+\delta} \sim C_1^{2+\delta} n^{-1-\delta/2}$$

und deshalb

$$\sum_{n=1}^{\infty} \kappa_n^{2+\delta} \left(\frac{\gamma_n}{\beta_n}\right)^{2+\delta} < \infty\,.$$

Mit dem Kronecker-Lemma 2.7 folgt

$$\kappa_n^{2+\delta} \sum_{j=1}^{n} \left(\frac{\gamma_n}{\beta_n}\right)^{2+\delta} \to 0\,.$$

Es soll nun das Supremum in Gleichung (6) mithilfe von Voraussetzung (d) abgeschätzt werden. Mittels Substitution bei Unabhängigkeit 1.3 gilt für $n \geq 1$

$$\begin{aligned}
\mathbb{E}[|H(X_{n-1}, Z_n) - h(X_{n-1})|^{2+\delta} \,|\, \mathcal{F}_{n-1}] &= \int |H(X_{n-1}, z) - h(X_{n-1})|^{2+\delta} \,\mathrm{d}P^{Z_1}(z) \\
&= \|H(X_{n-1}, Z_1(\cdot)) - h(X_{n-1})\|_{2+\delta}^{2+\delta}
\end{aligned}$$

mit der Norm auf $\mathcal{L}^{2+\delta}(P)$. Definiere $\varphi(x) \coloneqq \mathbb{E}[|H(x, Z_1)|^{2+\delta}]$. Durch die Dreiecksungleichung und die Jensensche Ungleichung kann für $x \in I$ man abschätzen

$$\begin{aligned}
\|H(x, Z_1) - h(x)\|_{2+\delta} &\leq \|H(x, Z_1)\|_{2+\delta} + |h(x)| \\
&= \varphi(x)^{1/(2+\delta)} + |\mathbb{E}[H(x, Z_1)]| \\
&\leq \varphi(x)^{1/(2+\delta)} + \mathbb{E}[|H(x, Z_1)|^{2+\delta}]^{1/(2+\delta)} \\
&= 2\varphi(x)^{1/(2+\delta)}\,,
\end{aligned}$$

also gilt

$$\sup_{n\geq 1}\mathbb{E}[|H(X_{n-1},Z_n)-h(X_{n-1})|^{2+\delta}\,|\,\mathcal{F}_{n-1}] \leq 2^{2+\delta}\sup_{n\geq 1}\varphi(X_{n-1}) < \infty \text{ f.s.}$$

Die rechte Seite ist fast sicher endlich nach Voraussetzung (d) zusammen mit der fast sicheren Konvergenz von X_n gegen x_0.

Insgesamt folgt die Lyapunov-Bedingung

$$\kappa_n^{2+\delta}\sum_{j=1}\mathbb{E}[|\Delta M|^{2+\delta}\,|\,\mathcal{F}_{j-1}] \to 0 \quad \text{f.s.}$$

und der stabile CLT liefert

$$\sqrt{n}\beta_n M_n \to \mathcal{N}\left(0, \frac{g(x_0)C_1^2}{2aC_1-1}\right) \quad \textit{mischend.}$$

Wir kehren zur Zerlegung aus Gleichung (4) zurück:

$$\sqrt{n}(X_n - x_0) = \sqrt{n}\beta_n(X_0 - x_0) + \sqrt{n}\beta_n M_n + \sqrt{n}\beta_n B_n$$

Mit der Konvergenz der beiden anderen Terme gegen 0 folgt die Behauptung mittels Korollar 2.4. □

Bemerkung 3.3. Optimale Schrittweite

Beispiel 3.4. Mittelwert und Quantil

Literatur

[LPW08] David A. Levin, Yuval Peres und Elizabeth L. Wilmer. *Markov Chains and Mixing Times*. American Mathematical Society, 2008. URL: `http://pages.uoregon.edu/dlevin/MARKOV/`.