BZQ II: Stochastikpraktikum

Block 4: Simulation stochastischer Prozesse

Randolf Altmeyer

January 26, 2017

Überblick

1. Monte-Carlo-Methoden, Zufallszahlen, statistische Tests
2. Nichtparametrische Methoden
3. Lineares Modell, Klassifikation, PCA
4. **Simulation stochastischer Prozesse**
5. Markov-Chain-Monte-Carlo-Verfahren

Überblick

1. Stochastische Prozesse und wie man sie simuliert
2. Markovketten
3. Poisson-Prozess
4. Brownsche Bewegung
5. Geometrische Brownsche Bewegung
6. Anwendung: Option Pricing

Literatur:

- Robert Dobrow: *Robert P. Dobrow-Introduction to stochastic processes with R*, John Wiley & Sons, 2016

Stochastische Prozesse

- stochastischer Prozess = Familie von Zufallsvariablen $X = (X_t)_{t \in I}$, beliebige Indexmenge I
- **Beispiele:**
  - $X = (X_k)_{k \in \mathbb{N}}$, X_k iid
  - random walk: $X_k = \sum_{i=1}^{k} Y_i$, Y_i iid
  - Poisson-Prozess: $N_t = \sum_{k=1}^{\infty} \mathbf{1}(T_k \leq t)$
  - (compound) Poisson-Prozess: $X_t = \sum_{i=1}^{N_t} Y_i$, Y_i iid
  - Gaussprozesse, Martingale, etc.
- Modellierung von (zeitlichen oder räumlichen) Abhängigkeiten
- **Anwendungen:**
  - Simulation von Partikelsystemen
  - Google page rank
  - Modellierung von Versicherungsfällen
  - Verkehrs- oder Wettersimulationen
  - Verteilung der Sterne am Himmel

Simulation von stochastischen Prozessen

- einzelne X_t eher uninteressant, wichtiger sind Pfade $t \mapsto X_t$
- interessante Fragen:
  - Eintrittszeiten in bestimmte Mengen
  - Wahrscheinlichkeiten für bestimmte Ereignisse
  - stabile Zustände/Verteilungen
- Computer kann nur endliche viele $X_{t_1}, \ldots, X_{t_n}$ bestimmen
  $\Rightarrow$ simuliere von endlichdimensionalen Verteilungen
- Art der Simulation hängt stark von den Abhängigkeitsstrukturen der X_t ab

Markovketten

- zur Modellierung von diskreten Prozessen bei denen der nächste Schritt nur vom letzten Schritt abhängt (z.B. Brettspiel)
- **Definition:**
  - $X = (X_k)_{k\in\mathbb{N}_0}$ heißt (zeitlich homogene) Markovkette mit diskretem Zustandsraum S, wenn für alle n und alle Zustände $x_1, \dots, x_n \in S$ gilt:

$$\mathbb{P}\left(X_n = x_n \middle| X_0 = x_0, \dots, X_{n-1} = x_{n-1}\right) = \mathbb{P}\left(X_n = x_n \middle| X_{n-1} = x_{n-1}\right)$$

- Beispiel: random walk, Sprache
- Verteilung von X ist eindeutig charakterisiert durch Anfangsverteilung $X_0 \sim \mu$ und Übergangsmatrix $P = (P_{ij})_{i,j\in S}$, wobei $P_{ij} = \mathbb{P}(X_1 = j \mid X_0 = i)$, denn

$$\begin{aligned}
\mathbb{P}\left(X_0 = x_0, \dots, X_n = x_n\right) &= \mathbb{P}\left(X_n = x_n \middle| X_0 = x_0, \dots, X_{n-1} = x_{n-1}\right) \\
&\quad \cdot \mathbb{P}\left(X_0 = x_0, \dots, X_{n-1} = x_{n-1}\right) \\
&= P_{x_{n-1},x_n}\mathbb{P}\left(X_0 = x_0, \dots, X_{n-1} = x_{n-1}\right) \\
&= \left(\prod_{k=2}^{n} P_{x_{k-1},x_k}\right)\mu(x_0)
\end{aligned}$$

- Markovketten sind oft gute erste Näherungen von komplizierten Systemen
- **interessante Größen:**
  - Eintrittszeiten: z.B. $\tau_x = \inf\{k \geq 0 : X_k = x\}$ für Zustand $x \in S$
  - mittlere Eintrittszeit: $\mathbb{E}[\tau_x] = \sum_{k=0}^{\infty} k\mathbb{P}(\tau_x = k)$
  - Verteilung von X_k für große k (vgl. Stationarität)
- **Simulation:**
  - erzeuge $x_0 = X_0 \sim \mu$,
  - erzeuge $x_1 = X_1$ gemäß der Verteilung $\mathbb{P}(X_1 = \cdot \,|\, X_0 = x_0)$ mit Zähldichte $x \mapsto P_{x_0,x}$
  - erzeuge $x_2 = X_2$ gemäß der Verteilung $\mathbb{P}(X_2 = \cdot \,|\, X_1 = x_1)$ mit Zähldichte $x \mapsto P_{x_1,x}$
  - usw.

Poisson-Prozess

- Modell für zufällige Zeitpunkte (Anrufe im Callcenter, Anzahl Geburten pro Tag, etc.)
- $(N_t)_{t\geq 0}$ = zählt Anzahl der eingetretenen Ereignisse bis t $\Rightarrow N_t \in \mathbb{N}$, $N_t \geq N_s$ für $t \geq s$
- **Formale Definition:**
  - seien $X_i \overset{iid}{\sim} \text{Exp}(\lambda)$ (=Wartezeiten, $0 < \lambda$=Intensität=mittlere Eintrittsrate)
  - $T_k = \sum_{i=1}^{k} X_i$ (= Eintrittszeiten)
  - $N_t := \sum_{k=1}^{\infty} \mathbf{1}\,(T_k \leq t)$
- **äquivalente Definition:**
  - $N_0 = 0$
  - für $0 \leq s < t$ ist $N_t - N_s \sim \text{Poi}(\lambda(t-s))$ (stationäre Inkremente)
  - $N_t - N_s \perp N_s$ (unabhängige Inkremente)

Inhomogener Poisson-Prozess

- Intensität ist häufig zeitabhängig (z.B. Menschen in der Mensa)
- **Definition:**
  - $N_0 = 0$
  - für $0 \leq s < t$ ist $N_t - N_s \sim \mathrm{Poi}(\int_s^t \lambda(r)dr)$ ($\lambda =$ Intensitätsfunktion)
  - N_t hat unabhängige Inkremente
- es gilt immer noch $N_t = \sum_{k=1}^{\infty} \mathbf{1}(T_k \leq t)$ für Eintrittszeiten T_k, aber Wartezeiten $T_k - T_{k-1}$ nicht mehr unbedingt exponentialverteilt
- **Simulation:**
  - wähle festen Zeithorizont T
  - simuliere $(\tilde{N}_t)_{t \geq 0}$ Poisson-Prozess der Intensität 1 mit Eintrittszeiten $\tilde{T}_k$ bis $\Lambda(T)$ für $\Lambda(t) := \int_0^t \lambda(r)dr$
  - $T_k = \Lambda^{-1}(\tilde{T}_k)$, wobei $\Lambda^{-1}(t) = \inf\{x \geq 0 : \Lambda(x) \geq t\}$, $N_t := \tilde{N}_{\Lambda(t)}$
    $\Rightarrow N_0 = \tilde{N}_{\Lambda(0)} = 0$, $N_t - N_s = \tilde{N}_{\Lambda(t)} - \tilde{N}_{\Lambda(s)} \sim \mathrm{Poi}(\Lambda(t) - \Lambda(s))$, Unabhängigkeit klar

Brownsche Bewegung

- **Motivation:**
  - X_k iid, $\mathbb{E}[X_k] = 0$, $\mathrm{Var} X_k = 1$, $S_n = \sum_{k=1}^n X_k$
  - $Y_t = \frac{1}{\sqrt{n}} S_{\lfloor nt \rfloor} +$ lineare Interpolation für $0 \leq t \leq 1$
  - $(Y_t)_{0 \leq t \leq 1} \xrightarrow{d} (B_t)_{0 \leq t \leq 1}$
- Modell für zufällige Bewegungen von Partikeln (ursprüngliche Motivation des Botanikers Robert Brown waren Pollen, die auf Wasseroberfläche schwimmen)
- **Definition:**
  - $B_0 = 0$
  - für $0 \leq s < t$ ist $B_t - B_s \sim N(0, t-s)$
  - $B_t - B_s \perp B_s$
  - Pfade $t \mapsto B_t$ sind fast sicher stetig

Brownsche Bewegung

- **Eigenschaften:**
  - fast jeder Pfad ist *nirgendwo* differenzierbar
  - Selbstähnlichkeit: $\frac{1}{a} B_{a^2 t} \sim B_t$, $a > 0$
  - Gesetz des iterierten Logarithmus: $\limsup_{t \to \infty} \frac{|B_t|}{\sqrt{2t \log \log t}} = 1$
- **Simulation über Approximation durch random walk:**
  - wähle maximale Länge T, *Schrittweite* $\frac{T}{N}$
  - erzeuge $X_k \overset{iid}{\sim} N(0, \frac{T}{N})$, $k = 1, \dots, N$
  - $\tilde{B}_t = \sum_{k=1}^{\lfloor tN/T \rfloor} X_k \sim N(0, \frac{\lfloor tN/T \rfloor T}{N}) \approx N(0, t)$

Stochastische Differentialgleichungen (SDEs)

- Brownsche Bewegung modelliert zufällige Bewegungen ohne Muster
- ODE für Modellierung von linearem Wachstum: $dX_t = \mu dt$
- Wachstum mit zufälligen Störungen:

$$
\begin{aligned}
&"dX_t = \mu dt + \sigma dB_t" \\
\Leftrightarrow\ &X_t = \mu t + \sigma B_t
\end{aligned}
$$

  Achtung: B_t fast sicher nirgendwo differenzierbar
- allgemeiner:

$$
\begin{aligned}
&dX_t = \mu(t, X_t)dt + \sigma(t, X_t)dB_t \\
\Leftrightarrow &X_t = X_0 + \int_0^t \mu\left(r, X_r\right) dr + \int_0^t \sigma\left(r, X_r\right) dB_r
\end{aligned}
$$

- X_0 ist Anfangswert, μ beschreibt Trend, σ beschreibt Störungen
- Anwendung:
  - Modellierung (komplizierter) Partikelbewegungen
  - Finanzmathematik (siehe geometrische Brownsche Bewegung weiter unten)

Simulation von SDEs

- **Euler-Schema:**
  - wähle maximal Länge T, Schrittweite $\frac{T}{N}$
  - erzeuge $Z_k \overset{iid}{\sim} N(0,1)$
  - erzeuge $\tilde{X}_0 := X_0$ (wenn bekannt, oft deterministisch)
  - setze

$$\tilde{X}_{\frac{kT}{N}} = \tilde{X}_{\frac{(k-1)T}{N}} + \mu\left(\frac{(k-1)T}{N}, \tilde{X}_{\frac{(k-1)T}{N}}\right) \cdot \frac{T}{N} + \sigma\left(\frac{(k-1)T}{N}, \tilde{X}_{\frac{(k-1)T}{N}}\right)\sqrt{\frac{T}{N}} Z_k$$

- Euler-Schema konvergiert gegen X “unter schwachen Voraussetzungen”

Geometrische Brownsche Bewegung

- **Ziel:** Modellierung von Preisprozessen $(X_t)_{t\geq 0}$ (z.B. Aktienpreise)
- warum nicht $X_t = B_t$?
  $\Rightarrow$ Probleme: negative Preise? unabhängige Inkremente? Wachstum ohne klares Muster?
- **besser:**
  - typischer (deterministischer) Wachstumsprozess erfüllt $\frac{dX_t}{X_t} = \mu dt$ mit Lösung $X_t = e^{\mu t}$
  - beachte Risiko im Preis bzw. Störungen durch andere Marktteilnehmer:
    $$\frac{dX_t}{X_t} = \mu dt + \sigma dB_t$$
    $$\Rightarrow dX_t = \mu X_t dt + \sigma X_t dB_t$$
  - explizite/exakte Lösung: $X_t = X_0 \exp\left(\left(\mu - \frac{\sigma^2}{2}\right) t + \sigma B_t\right)$
- "geometrische" Brownsche Bewegung, da

$$X_0 \to X_0 \left(\left(\mu - \frac{\sigma^2}{2}\right)\frac{1}{N} + \sigma\sqrt{\frac{1}{N}} Z_1\right) \to \cdots \to X_0 \prod_{k=1}^{N} \left(\left(\mu - \frac{\sigma^2}{2}\right)\frac{1}{N} + \sigma\sqrt{\frac{1}{N}} Z_k\right)$$

$\to X_t$ für $N \to \infty$

Anwendung: Option Pricing

- Wertpapier hat Preis X_t, folgt geometrischer Brownscher Bewegung
- risikolose Zinsrate r
- Europäische Call-Option mit *Ausübungszeit T und Ausübungspreis K*
  - gibt das Recht (nicht die Pflicht!) zum Zeitpunkt T eine Einheit des Wertpapiers zu kaufen zum Preis K
  - die Option hat zum Zeitpunkt T den Wert $(X_T - K)_+$
  - wieviel sollte die Option zum Zeitpunkt 0 kosten?
  - Annahme: es gibt keine Handelsstrategie, die ohne Risiko zu Gewinn führt (no *arbitrage*)
  - man kann zeigen, dass der *faire Preis* ist $C = e^{-rT}\mathbb{E}^*[(X_T - K)_+]$, wobei X_T unter dem Maß $\mathbb{P}^*$ eine geometrische Brownsche Bewegung mit den Parametern $\mu = r$ und σ ist
  - Black und Scholes haben 1973 gezeigt:

$$
\begin{aligned}
C &= X_0\Phi\left(d_1\right) - Ke^{-rT}\Phi\left(d_2\right), \\
d_1 &= \frac{1}{\sigma\sqrt{T}}\left(\log\frac{X_0}{K} + \left(r + \frac{\sigma^2}{2}\right)T\right), \\
d_2 &= d_1 - \sigma\sqrt{T},
\end{aligned}
$$

    und Φ ist die Verteilungsfunktion von $N(0,1)$