Stochastische Simulation von Zufallsvariablen und Punktprozessen

Seminar "Stochastische Geometrie und ihre Anwendungen"

Christoph Englisch

12.11.2009

Inhalt

Monte-Carlo-Simulation

- stochastisches Verfahren zur numerischen Lösung analytisch unlösbarer oder aufwendiger Probleme
- Grundlage: Gesetz der großen Zahlen

Anwendungsbeispiele

- Monte-Carlo-Integration
  (Berechnung des Integrals einer Funktion auf $[0, 1]$)
- Buffonsches Nadelexperiment (Bestimmung von π)

Algorithmus zur Approximation von π

- betrachte Quadrat mit darin einbeschriebenem Kreis
- "werfe" Punkte willkürlich in das Quadrat
  1. Erzeuge $2n$ Pseudozufallszahlen $u_1, \ldots, u_{2n} \sim U(0,1)$ per Zufallszahlengenerator.
  2. Setze $x_i = 2u_i - 1$ und $y_i = 2u_{n+i} - 1$.
  3. Setze
  $$z_i = \begin{cases} 1, & \text{falls } x_i^2 + y_i^2 < 1, \\ 0, & \text{sonst.} \end{cases}$$
  4. Berechne $\frac{4(z_1 + \ldots + z_n)}{n}$.

Zufallszahlen

- "echte" Zufallszahlen manuell bzw. mechanisch erzeugt (z.B. durch Würfel oder Roulette)

  $\Rightarrow$ aufwendig, daher Zufallszahlengeneratoren

Zufallszahlen

- "echte" Zufallszahlen manuell bzw. mechanisch erzeugt (z.B. durch Würfel oder Roulette)
  $\Rightarrow$ aufwendig, daher Zufallszahlengeneratoren

Zufallszahlengeneratoren

- Algorithmen, die Realisierungen von Zufallsvariablen erzeugen können
- diese zwar deterministisch bestimmt, aber zufälliges Aussehen (Pseudozufallszahlen)
- Ausgangspunkt: Standardzufallszahlengeneratoren
  $\Rightarrow$ durch Transformation oder Verwerfung auch andere Verteilungen möglich

Linearer Kongruenzgenerator

- Rekursionsformel mit Startwert $z_0 \in \{0, 1, \ldots, m-1\}$:

$$z_k = (az_{k-1} + c)\ mod(m), \quad \forall k = 1, \ldots, n$$

- $m \in \mathbb{N},\ a, c \in \{0, 1, \ldots, m-1\}$
- nach Normierung $u_k = \frac{z_k}{m}$ erhält man die Standardpseudozufallszahlen $u_1, \ldots, u_n \in [0, 1)$
  $\Rightarrow$ höchstens m verschiedene $u_1, \ldots, u_n$
- es gibt gewisse Bedingungen, die man an a, c, m bzw. z_0 stellen kann, damit maximal mögliche Periodenlänge m erreicht wird

Tests von Güteeigenschaften von Zufallszahlengeneratoren

Mithilfe des χ^2- Anpassungstest von Pearson kann geprüft werden, ob die $u_1, \dots, u_n$ einer Realisierung von unabhängigen $U_1, \dots, U_n \sim U(0,1)$ entsprechen.

- Zerlege $[0,1)$ in r gleichlange Teilintervalle $[0, \frac{1}{r}), \dots, [\frac{r-1}{r}, 1)$.
- Betrachte $(r-1)$-dimensionalen Parametervektor $p_0 = (\frac{1}{r}, \dots, \frac{1}{r})$ und die Testgröße $T_n : \mathbb{R}^n \mapsto [0, \infty)$ mit

  $$T_n(u_1, \dots, u_n) = \sum_{j=1}^{r} \frac{(Z_j(u_1, \dots, u_n) - \frac{n}{r})^2}{\frac{n}{r}}$$

  wobei $Z_j(u_1, \dots, u_n) = \#\{i : 1 \leq i \leq n, j-1 < ru_i \leq j\}$.
- Da T_n asymptotisch χ^2_{r-1}-verteilt ist, falls $U_1, \dots, U_n \sim U(0,1)$ und unabhängig, wird $H_0 : p = p_0$ abgelehnt, falls

  $$T_n(u_1, \dots, u_n) > \chi^2_{r-1,1-\alpha}$$

Inversionsmethode

- Wiederholung: $F : \mathbb{R} \mapsto [0,1]$ Verteilungsfunktion, dann heißt $F^{-1} : [0,1] \mapsto \mathbb{R} \cup \{\infty\}$ verallgemeinerte Inverse von F mit

$$F^{-1}(y) = \inf\{x : F(x) \geq y\}$$

- Für $x \in \mathbb{R}$ und $y \in (0,1)$ gilt

$$y \leq F(x) \Leftrightarrow F^{-1}(y) \leq x$$

Inversionsmethode

- Wiederholung: $F : \mathbb{R} \mapsto [0,1]$ Verteilungsfunktion, dann heißt $F^{-1} : [0,1] \mapsto \mathbb{R} \cup \{\infty\}$ verallgemeinerte Inverse von F mit

$$F^{-1}(y) = \inf\{x : F(x) \geq y\}$$

- Für $x \in \mathbb{R}$ und $y \in (0,1)$ gilt

$$y \leq F(x) \Leftrightarrow F^{-1}(y) \leq x$$

Theorem

Sei $U_1, \ldots, U_n$ eine Folge von $U(0,1)$-verteilten und unabhängigen Zufallsvariablen und $F : \mathbb{R} \mapsto [0,1]$ eine beliebige Verteilungsfunktion. Dann sind die Zufallsvariablen $X_1, X_2, \ldots$ mit $X_i = F^{-1}(U_i)$ für jedes $i = 1, 2, \ldots$ unabhängig und besitzen die Verteilungsfunktion F.

Beispiel Exponentialverteilung

- Sei $\lambda > 0$ und $F : \mathbb{R} \mapsto [0,1]$ Verteilungsfunktion der Exp(λ)-Verteilung mit

$$F(x) = \begin{cases} 1 - e^{-\lambda x}, & \text{falls } x \geq 0, \\ 0, & \text{falls } x < 0. \end{cases}$$

  $\Rightarrow F^{-1}(u) = -\frac{log(1-u)}{\lambda}$ für jedes $u \in (0,1]$

- Nach Theorem: $X = -\frac{log(U)}{\lambda} \sim Exp(\lambda)$ für $U \sim U(0,1]$
  $\Rightarrow x_i = -\frac{log(u_i)}{\lambda}$ sind entsprechende Realisierungen

Akzeptanz- und Verwerfungsmethode

- Ziel: Generierung von x gemäß Dichte f
- geeignet, wenn es eine Hilfsdichte g gibt, die "ähnlich" aussieht wie f
- sei c eine Konstante, sodass $f(x) \leq cg(x) \quad \forall\, x$
  1. Generiere y gemäß Dichte g.
  2. Generiere $u \sim U(0,1]$.
  3. Wenn $u \leq \frac{f(y)}{cg(y)}$, setze $x = y$, ansonsten zurück zu 1.

$\Rightarrow$ generierte Zufallszahl x gemäß Dichte f

$\Rightarrow$ im Mittel c Pseudozufallszahlen notwendig bis zur Generierung von x

Inhalt

Wiederholung Poisson-Prozess

- Definition: $\{N_B, B \in \mathfrak{B}_0(\mathbb{R}^d)\}$ ist ein Poisson-Prozess mit Intensitätsmaß μ, wenn
  - $N_{B_1}, N_{B_2}, \dots$ unabhängige Zufallsvariablen sind für paarweise disjunkte $B_1, B_2, \dots \in \mathfrak{B}_0(\mathbb{R}^d)$
  - $N_B \sim Poi(\mu(B))$ für jedes $B \in \mathfrak{B}_0(\mathbb{R}^d)$
- $\{N_B\}$ heißt homogener Poisson-Prozess mit Intensität λ, falls

$$\mu(B) = \lambda \nu_d(B) \quad \forall B \in \mathfrak{B}_0(\mathbb{R}^d)$$

Theorem

Sei $\mu : \mathfrak{B}_0(\mathbb{R}^d) \mapsto [0, \infty)$ ein beliebiges Maß mit $0 < \mu(\mathbb{R}^d) < \infty$ und $N : \Omega \mapsto [0, \infty)$ bzw. $S_1, S_2, \ldots : \Omega \mapsto \mathbb{R}^d$ seien unabhängige Zufallsvariablen mit

$$N \sim Poi(\mu(\mathbb{R}^d)), \quad S_i \sim \frac{\mu(\cdot)}{\mu(\mathbb{R}^d)} \ \ \forall i \geq 1$$

Dann ist das zufällige Zählmaß $\{N_B, B \in \mathfrak{B}_0(\mathbb{R}^d)\}$ mit

$$N_B = \#\{i : 1 \leq i \leq N, S_i \in B\} \quad \forall B \in \mathfrak{B}_0(\mathbb{R}^d)$$

ein Poisson-Prozess mit dem Intensitätsmaß μ.

bedingte Gleichverteilungseigenschaft

- Sei $N_{\widetilde{C}}$ ein homogener Poisson-Prozess auf einem d-dimensionalen Quader der Form $\widetilde{C} = [a_1, b_1) \times \ldots \times [a_d, b_d)$, dann hat der Zufallsvektor $S_i = (S_{i1}, \ldots, S_{id})$ für jedes $i = 1, \ldots, n$ unabhängige Komponenten $S_{ij} \sim U(a_j, b_j)$ für jedes $j = 1, \ldots, d$

$\Rightarrow$ Algorithmus:

1. Generiere eine Realisierung von $N_{\widetilde{C}} \sim \mathrm{Poi}(\mu(\widetilde{C}))$.
2. Falls $N_{\widetilde{C}} = k$, generiere $S_1, \ldots, S_k$, wobei $S_i = (S_{i1}, \ldots, S_{id})$ und $S_{ij} \sim U(a_j, b_j)$.
3. Die Menge $\{S_i\}$ von Punkten im $\mathbb{R}^d$ ist dann eine Realisierung eines Poisson-Prozesses auf dem Quader $\widetilde{C}$.

Theorem

Sei $\{N_B, B \in \mathfrak{B}_0(\mathbb{R}^d)\}$ ein Poisson-Prozess mit dem Intensitätsmaß μ und sei $B_0 \in \mathfrak{B}_0(\mathbb{R}^d)$ eine beliebige Borel-Menge. Dann ist das zufällige Zählmaß $\{\widetilde{N}_B, B \in \mathfrak{B}_0(\mathbb{R}^d)\}$ mit $\widetilde{N}_B = N_{B \cap B_0}$ ein Poisson-Prozess auf B_0 mit Intensitätsmaß $\widetilde{\mu}$, wobei $\widetilde{\mu}(B) = \mu(B \cap B_0)$ für jedes $B \in \mathfrak{B}_0(\mathbb{R}^d)$.

Theorem

Sei $\{N_B, B \in \mathfrak{B}_0(\mathbb{R}^d)\}$ ein Poisson-Prozess mit dem Intensitätsmaß μ und sei $B_0 \in \mathfrak{B}_0(\mathbb{R}^d)$ eine beliebige Borel-Menge. Dann ist das zufällige Zählmaß $\{\widetilde{N}_B, B \in \mathfrak{B}_0(\mathbb{R}^d)\}$ mit $\widetilde{N}_B = N_{B \cap B_0}$ ein Poisson-Prozess auf B_0 mit Intensitätsmaß $\widetilde{\mu}$, wobei $\widetilde{\mu}(B) = \mu(B \cap B_0)$ für jedes $B \in \mathfrak{B}_0(\mathbb{R}^d)$.

Akzeptanz- und Verwerfungsmethode

1. Generiere eine Realisierung von $N_C \sim \mathrm{Poi}(\mu(C))$.
2. Falls $N_C = k$, generiere so lange Realisierungen $s_1, s_2, \ldots$ der unabhängigen Zufallsvektoren $S_1, S_2, \cdots \sim \mu(\cdot \cap \widetilde{C})/\mu(\widetilde{C})$, bis k der Pseudozufallszahlen $s_1, \ldots, s_n$ in der Menge C liegen
3. Die Menge $\{s_i : s_i \in C, 1 \leq i \leq n\}$ von Punkten im $\mathbb{R}^d$ ist dann eine Realisierung eines Poisson-Prozesses.

Theoreme

Sei $\{N_B, B \in \mathfrak{B}_0(E)\}$ ein Poisson-Prozess in E mit lokal endlichem Intensitätsmaß μ und $T : E \mapsto \widetilde{E}$ Borel-messbar, wobei die Urbilder von beschränkten Borel-Mengen beschränkt seien. Dann gelten folgende Theoreme:

- $\{\widetilde{N}_{\widetilde{B}}, \widetilde{B} \in \mathfrak{B}_0(\widetilde{E})\}$ mit $\widetilde{N}_{\widetilde{B}} = N_{T^{-1}(\widetilde{B})}$ ist ein Poisson-Prozess in $\widetilde{E}$ mit Intensitätsmaß $\widetilde{\mu}(\widetilde{B}) = \mu(T^{-1}(\widetilde{B}))$.
- Seien $\{S_i\}$ die Atome eines Poisson-Prozesses $\{N_B, B \in \mathfrak{B}_0(E)\}$ und $\{U_i\}$ eine Folge von unabhängigen und identischen Zufallsvektoren in $\mathbb{R}^d$, die von $\{S_i\}$ unabhängig sind, dann gilt: $N_{B\times C} = \#\{i : (S_i, U_i) \in B \times C\}, B \in \mathfrak{B}_0(E), C \in \mathfrak{B}_0(\mathbb{R}^m)$ ist ein Poisson-Prozess.

Radiale Simulation eines homogenen Poisson-Prozesses im $\mathbb{R}^2$

1. Generiere die $t_1, \ldots, t_{n(r)}$ gemäß der Exp(1)-Verteilung, wobei

$$n(r) = \max\left\{ i : \sqrt{\sum_{k=1}^{i} \frac{t_k}{\pi\lambda}} \leq r \right\}.$$

2. Generiere die $u_1, \ldots, u_{n(r)}$ gemäß der $\mathrm{U}(0, 2\pi)$-Verteilung.
3. Berechne die Vektoren $(s_1, u_1), \ldots, (s_{n(r)}, u_{n(r)})$, wobei

$$s_i = \sqrt{\sum_{k=1}^{i} \frac{t_k}{\pi\lambda}} \qquad \forall i \geq 1.$$

4. Transformiere $(s_1, u_1), \ldots, (s_{n(r)}, u_{n(r)})$ mit Hilfe der Abbildung $T : [0, \infty) \times [0, 2\pi) \mapsto \mathbb{R}^2$ mit $T(s, u) = (s \cos u, s \sin u) \;\; \forall s \geq 0, u \in [0, 2\pi)$.

$\Rightarrow$ Realisierung $T(s_1, u_1), \ldots, T(s_{n(r)}, u_{n(r)})$ eines homogenen Poisson-Prozesses im Kreis $B(o, r) = \{x \in \mathbb{R}^2 : |x| \leq r\}$

Theorem

Seien $\lambda_1, \lambda_2 : \mathbb{R}^d \mapsto [0, \infty)$ zwei Borel-messbare und lokal-integrierbare Funktionen, so dass

$$\lambda_1(x) \geq \lambda_2(x) \quad \forall x \in \mathbb{R}^d.$$

Sei $\{S_n\}$ Atome eines Poisson-Prozesses mit Intensitätsfunktion λ_1.
Außerdem sei $U_1, U_2, \ldots$ eine Folge von $U(0,1)$-verteilten und von $\{S_n\}$ unabhängigen Zufallsvariablen.
Dann ist das zufällige Zählmaß $\{\widetilde{N}_B, B \in \mathfrak{B}_0(\mathbb{R}^d)\}$ mit

$$\widetilde{N}_B = \#\{n : S_n \in B, U_n \leq \frac{\lambda_2(S_n)}{\lambda_1(S_n)}\} \quad \forall B \in \mathfrak{B}_0(\mathbb{R}^d)$$

ein inhomogener Poisson-Prozess mit der Intensitätsfunktion λ_2.

Ortsabhängige Verdünnung

1. Generiere Realisierung $s_1, s_2, \ldots, s_k \in C$ eines homogenen Poisson-Prozesses $\{S_n\}$ in C mit Intensität $\lambda_{max} = \sup_{x \in C} \lambda(x) < \infty$.
2. Generiere Realisierungen $u_1, u_2, \ldots, u_k$ von unabhängigen $U_1, U_2, \ldots, U_k \sim [0,1]$.
3. Eliminiere Punkte s_n, für die $u_n > \lambda(s_n)/\lambda_{max}$ gilt.
4. $\{s_{i_1}, \ldots, s_{i_m}\} \subset \{s_1, \ldots, s_k\}$ bilden eine Realisierung eines Poisson-Prozesses mit Intensitätsfunktion $\lambda : C \mapsto [0, \infty)$ in C.

Literatur

- V. Schmidt, *Vorlesungsskript Räumliche Statistik*, Universität Ulm (2008).
- V. Schmidt, *Vorlesungsskript Markov Chains and Monte-Carlo Simulation*, Universität Ulm (2006).
- S. Ross, *Simulation*, Academic Press, 4th edition (2006).
- J.Illian, A.Penttinen, H. Stoyan, *Statistical Analysis and Modelling of Spatial Point Patterns*, Wiley-Interscience (2008).