Seminar Stochastische Geometrie und räumliche Statistik, SS 2002

Endliche Markov Ketten und Monte-Carlo-Simulation

Universität Ulm
Abteilung Stochastik

Eleonore Gräf
Mai, 2002

1 Einleitung

1.1 Monte-Carlo-Methode

- 1949, J.v.Neumann und S.Ulam
- theor. Grundlagen schon lange vorher bekannt
- breite Anwendung erst mit dem Einsatz von DV-Anlagen in den 90-er Jahre(Bildanalyse, -verarbeitung)
- Monte-Carlo, Monaco; Spielkasino, Roulette
- universelle Methode, weil sie viele Vorgänge nachzubilden vermag, auf deren Ablauf zufällige Faktoren Einfluß nehmen

1.2 Markov-Ketten

Definition 1

Sei $(\Omega, \mathcal{F}, P)$ ein Wahrscheinlichkeitsraum, I eine beliebige nicht leere Indexmenge und $(S, \mathcal{S})$ ein meßb. Raum. Eine Familie $\{X_i,\ i \in I\}$ von Zufallsvariablen mit Werten in S heißt *stochastischer Prozeß* mit Parameterbereich I und Zustandsraum S.

Definition 2
Eine *Markov-Kette* ist ein stochastischer Prozeß $\{X_n,\ n \in I := \mathbb{Z}^+\}$ mit abzähl. Zustandsraum S, der die folg. Markovsche Eigenschaft besitzt: $\forall\ n, i_0, \ldots, i_{n-1}, i, j \in \mathbb{Z}^+$ und $\forall\ s_{i_0}, \ldots, s_{i_{n-1}}, s_i, s_j \in S$ mit $P(X_0 = s_{i_0}, \ldots, X_n = s_{i_n}) > 0$ ist

$$P(X_{n+1} = s_j \mid X_0 = s_{i_0}, \ldots, X_{n-1} = s_{i_{n-1}}, X_n = s_i)$$
$$= P(X_{n+1} = s_j \mid X_n = s_i)$$
$$= p_{ij}.$$

Mit $p_{ij} \geq 0$ und $\sum\limits_{j \in I} p_{ij} = 1\ \forall\ i \in I$.

Bemerkung 1 *Wir betrachten hier nur den endl. Zustandsraum* $S = \{s_1, \ldots, s_k\}$.

Definition 3

(a) Eine MK heißt *homogen* oder Kette mit *stationären Übergangswahrscheinlichkeiten* p_{ij}, wenn $\forall\ i, j \in I$ $p_{ij} := P(X_{n+1} = s_j \mid X_n = s_i)$ unabh. von n ist.

(b) $p_{ij}^{(m)} := P(X_{n+m} = s_j \mid X_n = s_i)$ ist die *m-Schritt Übergangswahrscheinlichkeit.*

(c) Eine WFkt $\alpha = (\alpha_i,\ i \in I)$ mit $\alpha_i = P(X_0 = s_i)$ heißt *Start-/Anfangsverteilung*.

(d) Anfangsverteilung heißt *stationär*, wenn $\alpha^T = \alpha^T \mathbf{P}$, d.h. $\forall\ j \in I \quad \alpha_j = \sum_{i \in I} \alpha_i p_{ij}$ gilt, wobei $\mathbf{P}$ die Übergangsmatrix ist.

Definition 4

(a) $s_i \to s_j[n]$ heißt der Zustand s_i führt in n Schritten zum Zustand s_j, wenn $p_{ij}^{(n)} > 0$.

(b) Gibt es ein $n \geq 1$, mit $s_i \to s_j[n]$, so sagen wir s_i führt zu s_j und schreiben $s_i \to s_j$.

(c) $s_i \leftrightarrow s_j$ bedeutet s_i *kommuniziert* mit s_j.

(d) s_i heißt *wesentlich*, wenn jeder Zustand s_j, zu dem s_i führt, auch zu s_i führt.

Definition 5
MK $\{X_n\}$ oder ihre Übergangsmatrix $\mathbf{P} = (p_{ij})$ heißt *irreduzibel*, wenn S nur aus einer Klasse besteht, d.h. $s_i \leftrightarrow s_j \ \forall \ s_i, s_j \in S$.

Definition 6

(a) Für einen Zustand s_i mit $s_i \to s_i$ heißt der ggT der potenziellen Rückkehrzeiten $d(s_i) = ggT\{n \geq 1 : s_i \to s_i[n]\}$ *die Periode von* s_i.

(b) Gilt nicht $s_i \to s_i$, so sei $d(s_i) = \infty$.

(c) Zustände mit $d(s_i) = 1$ heißen *aperiodisch*.

(d) Die Kette heißt aperiodisch, wenn alle Zustände aperiodisch sind, und periodisch mit Periode d, wenn alle $d(s_i) = d \geq 2$ sind.

1.3 Ergodizität

Definition 7

Ein Zustand s_i heißt *rekurrent*, wenn
$P(X_n = s_i\ f\ddot{u}r\ \infty - viele\ n \mid X_0 = s_i) = 1$ gilt, sonst *transient*.
MK heißt rekurrent(transient), wenn jeder Zustand rekurrent(transient) ist.

Definition 8

Eine MK $\{X_n\}$ mit Übergangsmatrix $\mathbf{P} = (p_{ij})$ heißt *ergodisch*, wenn
(a) $\forall\ j \in I: \quad \pi_j := \lim_{n\to\infty} p_{ij}^{(n)}$ ex.,
(b) $\pi_j \geq 0$ und von i unabh.,
(c) (π_j) ist WFunktion, d.h. $\sum_{j\in I} \pi_j = 1$.

Theorem 1 *Eine MK* $\{X_n\}$ *mit Übergangsmatrix* $\mathbf{P} = (p_{ij})$ *ist ergodisch* $\Leftrightarrow$ $\mathbf{P} = (p_{ij})$ *regulär ist (d.h.* $\exists\, n_0 \geq 1$*, so daß alle* $p_{ij}^{(n_0)} > 0$ *sind).*

Beweis: (s. Literatur (4))

Theorem 2 *Sei $\{X_n\}$ ergodische MK, dann ist $\pi = (\pi_1, \ldots, \pi_k)$ die eindeutige WLösung von $\pi_j = \sum_{i \in I} \pi_i p_{ij}, \forall\, j \in I$.*

Beweis: (s. Literatur (4))

Theorem 3 *Sei $\{X_n\}$ eine MK mit endl. Zustandsraum. Dann gilt: MK ist ergodisch $\Leftrightarrow$ MK irreduzibel und aperiodisch ist.*

Beweis: (s. Literatur (4))

Definition 9
Sei $\{X_n\}$ eine MK mit einem Zustandsraum $S = \{s_1, \ldots, s_k\}$ und Übergangsmatrix $\mathbf{P} = (p_{ij})$. Die Anfangsverteilung α heißt *reversibel*, wenn $\forall\ i, j \in I$ gilt $\alpha_i p_{ij} = \alpha_j p_{ji}$.
Eine MK heißt reversibel, wenn sie eine reversible Anfangsverteilung besitzt.

Theorem 4 *Sei* $\{X_n\}$ *eine MK mit dem Zustandsraum* $S = \{s_1, \ldots, s_k\}$ *und Übergangsmatrix* $\mathbf{P}$. *Wenn die Anfangsverteilung der Kette* α *reversibel ist, dann ist sie auch stationär.*

Beweis: (s. Literatur(2))

2 MCMC-Methode

2.1 Problemstellung

Gegeben ist eine Verteilung π auf dem Zustandsraum $S = \{s_1, \ldots, s_k\}$.

Frage: Wie generiert man eine Zufallsvariable X mit der Verteilung π ?

2.2 Beispiel: Das hard-core-Modell

$G = (V, E)$ - Graph mit einer Menge der Pixel V und einer Menge der Kanten E. Den Pixeln wird zuf. der Wert 0 bzw. 1 zugewiesen, so daß 2 benachbarte Pixel nicht gleichzeitig den Wert 1 annehmen können. Konfigurationen aus der Menge $\{0, 1\}^V$, die dieses erfüllen, nennen wir *zulässig*.

μ_G - WMaß auf $\{0,1\}^V$, so daß für $\xi \in \{0,1\}^V$ mit
$Z_G := \{\#$ der zulässigen Konfigurationen$\}$ gilt:

$$(\pi_\xi) = \mu_G(\xi) = \begin{cases} \frac{1}{Z_G} & f\ddot{u}r \quad \xi \quad zul\ddot{a}ssig \\ 0 & sonst \end{cases}$$

$n(\xi) := \{\#$ von “1“ in $\xi\}$,
X - zuf. Konfig. mit Vert. μ_G,
$\mathbb{E}[n(X)] = \sum_{\xi \in \{0,1\}^V} n(\xi)\mu_G(\xi) = \frac{1}{Z_G} \sum_{\xi \in \{0,1\}^V} n(\xi) 1_{\{\xi \ zul\ddot{a}ssig\}} = ?$

2.3 Ausweg: Simulation

- Generierung von ZV $X_i \sim \mu_G$ mit $i \in \{1, \ldots, N\}$
- $\hat{n} = \sum_{i=1}^{N} \frac{n(X_i)}{N} \rightarrow \mathbb{E}[n(X)] \quad f.s. \quad für\ N \rightarrow \infty$ (SGGZ)

2.4 Erster Versuch(allg.):

$X := \psi(U)$, mit $U \sim$ U $[0,1]$ und $\psi : [0,1] \to S$

$$\psi(x) = \begin{cases} s_1 & f\ddot{u}r \quad x \in [0, \pi_{s_1}) \\ \vdots & \\ s_i & f\ddot{u}r \quad x \in [\sum_{j=1}^{i-1} \pi_{s_j}, \sum_{j=1}^{i} \pi_{s_j}) \\ \vdots & \\ s_k & f\ddot{u}r \quad x \in [\sum_{j=1}^{k-1} \pi_{s_j}, 1] \end{cases}$$

$\Rightarrow X \sim \pi$

2.5 Verbesserung: MCMC-Methode

Idee:
Wir konstruieren eine MK $(X_0, X_1, \ldots)$ auf einem WRaum $(\Omega, \mathcal{F}, P)$ mit Zustandsraum $S = \{s_1, \ldots, s_k\}$. Die Kette soll:

- irreduzibel und aperiodisch sein
- und deren stationäre Verteilung soll die gewünschte Verteilung π sein.

2.6 Beispiel: s.o.

$V = \{v_1, \ldots, v_k\}$, $S = \{\xi \in \{0,1\}^V : \ \xi \ ist \ zul\ddot{a}ssig\}$, μ_G.
Für die Konstruktion der MK benötigen wir den folg.
Übergangsmechanismus: zu jedem Zeitpunkt $(n+1) \in \mathbb{N}$:

1. Generiere $v \in V$ (gleichvert.)
2. Werfe eine faire Münze(Bernoulli)
3. (Kopf) $\wedge$ (alle Nachbarn von v$= 0$): $X_{n+1}(v) := 1$
4. Sonst: $X_{n+1}(v) := 0$
5. $\forall \ w \in V \setminus \{v\} : \ X_{n+1}(w) := X_n(w)$

Diese MK ist irreduzibel und aperiodisch, μ_G ist die stat. Anfangsverteilung dieser Kette.

Beweis der Stationarität:

Sei $\xi, \xi' \in S$. Mit Theorem 4 noch z.z.: μ_G ist reversibel, d.h.

$\mu_G(\xi) p_{\xi\xi'} = \mu_G(\xi') p_{\xi'\xi}$.

$d := d(\xi, \xi') =\{$ # von Pixeln, in denen ξ und ξ' sich unterscheiden$\}$.

Fall 1: $d = 0$: $\xi = \xi'$.

Fall 2: $d \geq 2$: $\mu_G(\xi) p_{\xi\xi'} = \mu_G(\xi') p_{\xi'\xi} = 0$.

Fall 3: $d = 1$: $\mu_G(\xi) p_{\xi\xi'} = \frac{1}{Z_G} \frac{1}{2k} = \mu_G(\xi') p_{\xi'\xi}$. □

Der obige Algorithmus ist ein typischer MCMC-Algorithmus

i) erfüllt sogar die strengere Eigenschaft: die Reversibilität,

ii) gehört in eine spezielle Klasse der MCMC-Algorithmen: *Gibbs Sampler*.

2.7 Hastings-Metropolis-Algorithmus:

Wir generieren eine MK mit einer stationären Verteilung π. Sei $\pi_j = \frac{b(j)}{B}$ mit $b(j) \geq 0, \quad j = 1, \ldots, k$, k groß und $B = \sum_{j=1}^{k} b_j < \infty$ und sei $Q = (q_{ij})$ eine irreduzible Übergangsmatrix.
Für die Konstruktion der MK benötigen wir folg.
Übergangsmechanismus: zu jedem Zeitpunkt $(n+1) \in \mathbb{N}$ mit $X_n = i$:

1. Generiere eine ZVar. Z mit

   $P(Z=j)=q_{ij}$, $j=1,\ldots,k$

2. Wenn $Z=j$, dann

   $$X_{n+1} := \begin{cases} j & mit \quad W.\ \phi_{ij} \\ i & mit \quad W.\ (1-\phi_{ij}) \end{cases}$$

Diese MK hat $\mathbf{P}$ als Übergangsmatrix mit $p_{ij}=q_{ij}\phi_{ij}$, wenn $j\neq i$
$p_{ii}=1-\sum\limits_{m\neq i} q_{im}\phi_{im}$.

Diese MK ist reversibel und damit stat., falls ϕ_{ij} wie folgt gewählt werden:

$\phi_{ij} = min\left\{\frac{\pi_j q_{ji}}{\pi_i q_{ij}}, 1\right\} = min\left\{\frac{b_j q_{ji}}{b_i q_{ij}}, 1\right\}.$

π ist die stat. Verteilung der Kette.

Beweis der Stationarität:

Mit Theorem 4 noch z.z.:

$\pi_i p_{ij} = \pi_j p_{ji} \quad \forall\ j \neq i$

$\Leftrightarrow \pi_i q_{ij} \phi_{ij} = \pi_j q_{ji} \phi_{ji} \quad (*).$

$(*)$ ist erfüllt.

Denn wenn $\phi_{ij} = \frac{\pi_j q_{ji}}{\pi_i q_{ij}}$ ist, dann ist $\phi_{ji} = 1$ und umgekehrt. $\square$

$\phi_{ij} = min\left\{\frac{b(j) q_{ji}}{b(i) q_{ij}}, 1\right\}$ ist jetzt von B unabh. und π_j ist gleichzeitig auch die Grenzverteilung (wenn $p_{ii} > 0\ \forall\ i$).

2.8 Eine Version des HM-Algorithmus:

Sei $\pi = (\pi_1, \ldots, \pi_k)$ WVerteilung auf $S = (s_1, \ldots, s_k)$.

- Konstruiere G (Graph) mit Pixeln aus S und einer bel. Menge E, die jedoch folgende Kriterien erfüllen sollte:
  - die hervorg. Kette ist irreduzibel,
  - jedes Pixel sollte nicht der Endpkt. von zu vielen Kanten sein .
- $d_i := \{\ \#$ von Nachb. von $s_i\}$ und $s_j \sim s_i$ heißt: s_j ist Nachb. von s_i. $\mathbf{P}$ ist die Übergangsmatrix mit

$$p_{ij} = \begin{cases} \frac{1}{d_i} \min\left\{\frac{\pi_j d_i}{\pi_i d_j}, 1\right\} & wenn \quad s_i \sim s_j \\ 0 & wenn \quad s_i \not\sim s_j \\ 1 - \sum\limits_{l, s_l \sim s_i} \frac{1}{d_i} \min\left\{\frac{\pi_l d_i}{\pi_i d_l}, 1\right\} & wenn \quad s_i = s_j \end{cases}$$

- Übergangsmechanismus: zu jedem Zeitpunkt $(n+1) \in \mathbb{N}$ mit $X_n = s_i$:

  1. Generiere s_j (gleichvert. auf der Menge der Nachb. von s_i)
  2. $X_{n+1} := \begin{cases} s_j & mit \quad W.\ \min\left\{\frac{\pi_j d_i}{\pi_i d_j}, 1\right\} \\ s_i & mit \quad W.\ \left(1 - \min\left\{\frac{\pi_j d_i}{\pi_i d_j}, 1\right\}\right) \end{cases}$

Diese MK besitzt die stat. Verteilung π.

2.9 Gibbs-Sampler:

Die meistgenutzte Version des HM-Algorithmus ist der Gibbs-Sampler.

Sei $\mathbf{X} = (X(1), \ldots, X(k))$ ein ZVektor mit WFkt $p(\mathbf{x})$.

Wir wollen einen ZVektor $\mathbf{X}$ generieren mit WFkt

$\pi_{\mathbf{x}} = \frac{p(\mathbf{x})}{P(\mathbf{X} \in \mathcal{A})}\ \forall\ \mathbf{x} \in \mathcal{A}$.

Ann: für $i,\ i = 1, \ldots, k$ und die Werte $x_j,\ j \neq i$, können wir eine ZVar. Z generieren mit W.

$$P(Z = z) = P(X(i) = z | X(j) = x_j,\ j \neq i) = \frac{p(\mathbf{y})}{\sum\limits_{\mathbf{x} \in \mathcal{A}, X(j) = x_j, j \neq i} p(\mathbf{x})} \quad (**)$$

mit $\mathbf{y} = (x_1, \ldots, x_{i-1}, z, x_{i+1}, \ldots, x_k)$ und

$\mathbf{x} = (x_1, \ldots, x_{i-1}, x_i, x_{i+1}, \ldots, x_k)$.

Gibbs-Sampler ist eine MK, die zum Zeitpunkt $(n+1) \in \mathbb{N}$ folg. Anweisungen folgt mit $\mathbf{X_n} = \mathbf{x} = (x_1, \ldots, x_i, \ldots, x_k) \in \mathcal{A}$:

1. Generiere eine Koord. i (gleichv.)
2. $\forall\, j \neq i:\; X_{n+1}(j) := x_j$
3. Generiere eine ZVar. Z mit W. (**)
4. Wenn $Z = z$, dann ist

   $\mathbf{y} = (x_1, \ldots, x_{i-1}, z, x_{i+1}, \ldots, x_k)$ und

   $$X_{n+1} := \begin{cases} \mathbf{y} & mit \quad W.\ \phi_{\mathbf{xy}} \\ \mathbf{x} & mit \quad W.\ (1 - \phi_{\mathbf{xy}}) \end{cases}$$

wobei $\phi_{\mathbf{xy}} = min\left\{\frac{\pi_{\mathbf{y}} q_{\mathbf{yx}}}{\pi_{\mathbf{x}} q_{\mathbf{xy}}}, 1\right\}$ und $q_{\mathbf{xy}} = \frac{1}{k} \frac{p(\mathbf{y})}{\sum\limits_{\mathbf{x} \in \mathcal{A}, X(j) = x_j, j \neq i} p(\mathbf{x})}$.

Für $\mathbf{x} \in \mathcal{A}$ und $\mathbf{y} \in \mathcal{A}$ gilt:
$\frac{\pi_{\mathbf{y}} q_{\mathbf{yx}}}{\pi_{\mathbf{x}} q_{\mathbf{xy}}} = \frac{\pi_{\mathbf{y}} p(\mathbf{x})}{\pi_{\mathbf{x}} p(\mathbf{y})} = 1$
und für $\mathbf{x} \in \mathcal{A}$ und $\mathbf{y} \notin \mathcal{A}$(da $\pi_{\mathbf{y}} = 0$):
$\frac{\pi_{\mathbf{y}} q_{\mathbf{yx}}}{\pi_{\mathbf{x}} q_{\mathbf{xy}}} = 0.$
$\pi_{\mathbf{x}}$ ist die stat. Grenzverteilung dieser MK.

2.10 Beispiel: ein MCMC-Algorithmus für q-Colourings

Sei $G = (V, E)$, $q \geq 4$, ganzzahlig.
Zuweisung von q versch. Farben/Werten, so dass 2 benachbarte Pixel versch. Farben besitzen.
Zuf. q-Colouring bedeutet, daßein q-Colouring auf der Menge der für G zulässigen q-Colourings generiert wird (gleichvert).
p_G - W-Maß auf $\{1, \ldots, q\}^V$.

Ein Gibbs-Sampler für zuf. q-Colouring ist eine MK, die zu jedem Zeitpunkt $(n+1) \in \mathbb{N}$ folg. Vorschriften folgt:

1. Generiere $v \in V$ (gleichvert.)
2. Generiere eine Farbe für $X_{n+1}(v)$ (gleichvert. auf der Menge der für v zul. Farben)
3. $\forall\, w \in V \setminus \{v\}$: $X_{n+1}(w) := X_n(w)$

Diese Kette ist aperiodisch und hat die stat. Verteilung p_G.

2.11 Eine Gibbs-Sampler-Variante:

Anstatt v zuf. zu ermitteln, können wir uns system. durch V bewegen. Wenn $V = \{v_1, \ldots, v_k\}$, dann entscheiden wir uns für das Pixel

$$\begin{cases} v_1 & zum\ Zeitpunkt\ 1, k+1, \ldots \\ \vdots & \\ v_i & zum\ Zeitpunkt\ i, k+i, \ldots \\ \vdots & \\ v_k & zum\ Zeitpunkt\ k, 2k, \ldots \end{cases}$$

Diese MK ist aperiodisch und besitzt gewünschte reversible Verteilung. Sie ist irreduzibel, wenn das ursprüngliche „zuf.-Punkte-Gibbs-Sampler“ irreduzibel ist.

3 Literatur

(1) P.Brémaud(1998): Markov Chains: Gibbs Fields, Monte Carlo Simulation and Queues, Springer.

(2) O.Häggström(2001): Finite Markov chains and Algorithmic Applications, http//dimacs.rutgers.edu/~dbwilson/exact.html#haggstrom:course

(3) U.Krengel(1998): Einführung in die Wahrscheinlichkeitstheorie und Statistik, Vieweg.

(4) T.Rolski, H.Schmidli, V.Schmidt, J.Teugels(1999): Stochastic Processes for Insurance and Finance, John Wiley & Sons Ltd.

(5) S.M.Ross(1997): Introduction to Probability Models, Academic Press.

(6) J.M.Sobol(1991): Die Monte-Carlo-Methode, Deutscher Verlag der Wissenschaften Berlin.