Angewandte Stochastik

Dr. C.J. Luchsinger

13 Allgemeine Theorie zu Markov-Prozessen (stetige Zeit, diskreter Zustandsraum)

Literatur Kapitel 13

* Grimmett & Stirzaker: Kapitel 6.9

Wie am Schluss von Kapitel 12 erwähnt, haben wir bereits Markov-Prozesse in stetiger Zeit kennengelernt - jedoch unsystematisch, exemplarisch. Der Poisson-Prozess, die stochastischen Epidemienmodelle in stetiger Zeit und die stochastischen Netzwerke sind solche Beispiele. Jetzt folgt die allgemeine Theorie dazu, ohne Beweise. Wir orientieren uns an Kapitel 6.9 von G. Grimmett und D. Stirzaker, "Probability and Random Processes". Dort findet man auch die Beweise, welche wir hier auslassen.

13.1 Grundlegende Definitionen

An dieser Stelle ist es sinnvoll, kurz eine simple Klassifikation der stochastischen Prozesse vorzunehmen. Im Wesentlichen hat man folgende Unterscheidungsmerkmale:

1. Ist die Zeit diskret ($n \in \{0, 1, 2, \ldots, N\}$) oder stetig ($t \in [0, T]$)?
2. Ist der Zustandsraum abzählbar (z.B. $\mathbb{N}$) oder kontinuierlich (z.B. $\mathbb{R}$)?

Je nach Situation haben dann Begriffe wie zum Beispiel Rekurrenz eine andere Bedeutung.

Es gibt also 4 wesentliche Möglichkeiten:

1. diskrete Zeit und abzählbarer Zustandsraum; Beispiele: Irrfahrt auf $\mathbb{Z}$, Markov-Ketten mit Übergangsmatrizen
2. diskrete Zeit und stetiger Zustandsraum; Beispiel: Zeitreihen

3. stetige Zeit und abzählbarer Zustandsraum; Beispiele: Poisson-Prozess, Markov-Prozesse mit infinitesimalen Generatoren (siehe dieses Kapitel)
4. stetige Zeit und stetiger Zustandsraum; Beispiel: Brownsche Bewegung (Kapitel 17)

Sei $X := \{X_t : t \geq 0\}$ eine Familie von Zufallsgrössen, welche Werte in $\mathbb{N}$ annehmen (allgemeiner ein abzählbarer Zustandsraum S) und mit dem halb-offenen Intervall $[0, \infty)$ (die Zeit) indexiert sind.

Definition 13.1 [Markov-Prozess (Stetige Zeit, diskreter Zustandsraum)] *Wir nennen den Prozess X einen Markov-Prozess in stetiger Zeit mit diskretem Zustandsraum, falls*

$$P[X_{t_n} = j | X_{t_1} = i_1, X_{t_2} = i_2, \ldots, X_{t_{n-1}} = i_{n-1}] = P[X_{t_n} = j | X_{t_{n-1}} = i_{n-1}]$$

für alle $j, i_1, i_2, \ldots, i_{n-1} \in \mathbb{N}$ und jede Folge $t_1 < t_2 < \ldots < t_n$ von Zeitpunkten.

Für gegebenes $\omega \in \Omega$ bezeichnet man die Abbildung $t \mapsto X_t(\omega)$ als Pfad des Prozesses.

Wie sieht

$$P[X_{t_n} = 5 | X_{t_{n-1}} = 3]$$

bei einem Poisson-Prozess der Rate λ aus und weshalb ist dort die Markov-Eigenschaft gemäss Definition 13.1 erfüllt (vgl. Definition 8.2)?

Bemerkung 13.2: Die Forderung nach einem diskreten Zustandsraum ist zentral. Kontinuierliche Zustandsräume sind mathematisch viel schwieriger zu handhaben, auch wenn wir ein sehr anschauliches Beispiel (die Brownsche Bewegung) *ohne Beweise* in Kapitel 17 einfach abhandeln können.

Bemerkung 13.3: In Definition 13.1 haben wir uns auf die Menge der natürlichen Zahlen als Wertebereich eingeschränkt. Diese Einschränkung ist irrelevant, solange der Zustandsraum diskret bleibt.

Bemerkung 13.4: Wir werden feststellen, dass viele Resultate in diskreter Zeit (Kapitel 3 bis 7) ein (fast) gleiches Analogon in stetiger Zeit haben. Die Unterschiede werden wir genau herausarbeiten. Es sei jedoch gleich gesagt, dass uns ein direktes Analogon zur Übergangsmatrix P fehlt, weil wir keine ausgezeichnete Zeiteinheit haben. Wir haben eine ganze Familie von $(P_t)_{t\geq 0}$, welche aber Gott sei Dank nicht sehr wichtig ist. Wir werden mit dem infinitesimalen Generator G einen würdigen Ersatz für P haben (in Teil 13.4 kommt ein P wie früher jedoch wieder vor).

Definition 13.5 [(homogene) Übergangswahrscheinlichkeiten, P_t] *Wir definieren die Übergangswahrscheinlichkeiten $p_{ij}(s,t)$ als*

$$p_{ij}(s,t) := P[X_t = j | X_s = i] \; \forall s \leq t.$$

Diese nennen wir homogen, falls $p_{ij}(s,t) = p_{ij}(0, t-s) \, \forall i,j,t,s$. Wir setzten in Zukunft Homogenität voraus und schreiben hierfür von jetzt an $p_{ij}(t-s)$. Die Matrix mit Einträgen $p_{ij}(t)$ nennen wir P_t.

Theorem 13.6 [$P_0 = I, P_t$ ist stochastisch, Chapman-Kolmogoroff]

a) $P_0 = \mathbf{I}$, die Identitätsmatrix,

b) P_t ist stochastisch, hat also nichtnegative Einträge, wobei jede Zeile auf 1 summiert.

c) $P_{s+t} = P_s P_t$ für $s, t \geq 0$.

Beweis von Theorem 13.6: a) ist offensichtlich.

b) Nichtnegativität folgt daraus, dass es sich um Wahrscheinlichkeiten handelt. Sei $\mathbf{1}$ der Spalten-Vektor mit Einträgen von lauter 1. Dann gilt:

$$(P_t \mathbf{1})_i = \sum_j p_{ij}(t) = P[\cup_j \{X_t = j\} | X_0 = i] = 1.$$

c) Mit FTW und der Markov-Eigenschaft folgt:

$$\begin{aligned} p_{ij}(s+t) &= P[X_{s+t} = j | X_0 = i] \\ &= \sum_k P[X_{s+t} = j | X_s = k, X_0 = i] P[X_s = k | X_0 = i] \\ &= \sum_k p_{ik}(s) p_{kj}(t). \end{aligned}$$

□

Wir wollen uns wegen Theorem 13.6 a) fragen, wie sich die Übergangswahrscheinlichkeiten $p_{ij}(t)$ verhalten, wenn $t \to 0$. Dazu definieren wir

Definition 13.7 [$(P_t)_{t \geq 0}$ **ist standard**] *Wir nennen die Familie $(P_t)_{t \geq 0}$ standard, falls komponentenweise $P_t \to I$, wenn $t \to 0$.*

Bemerkung 13.8: Es gilt (ohne Beweis): P_t ist genau dann standard, wenn die $p_{ij}(t)$ alle stetig in t sind (dies folgt übrigens aus den Chapman-Kolmogoroff-Gleichungen; kleine freiwillige Übung).

Bemerkung 13.9 [gleichmässig standard]: Die Theorie der Markov-Prozesse wird einfacher und genügt den meisten Anwendungen, wenn wir noch voraussetzen dürfen, dass obige Konvergenz in Definition 13.7. gleichmässig ist. Dies ist sicher erfüllt, wenn der Zustandsraum endlich ist. Wir setzen ab jetzt voraus, dass dies immer gilt und nennen diese Eigenschaft "P_t ist gleichmässig standard".

Setzen wir voraus, dass die Kette zur Zeit t in Zustand i ist: $X_t = i$. Was wird jetzt in einem kleinen Zeitintervall $(t, t+h)$ passieren (man kann zeigen, dass die Wahrscheinlichkeit für 2 Übergänge in einer kleinen Zeiteinheit h der Ordnung $o(h)$ ist)?

a) Nichts wird passieren mit Wahrscheinlichkeit:

(13.1)

b) Der Prozess wird genau einen Übergang von i nach $j, j \neq i$, machen mit Wahrscheinlichkeit:

$$ \tag{13.2} $$

Man kann wegen der Stetigkeit der $p_{ij}(t)$ in t zeigen, dass die $p_{ij}(h), j \neq i$, infinitesimal linear in h wachsen, genauer gilt: Es gibt Konstanten (g_{ij}), sodass

$$ p_{ij}(h) \simeq g_{ij}h \quad \forall i \neq j, \quad p_{ii}(h) \simeq 1 + g_{ii}h. \tag{13.3} $$

Es gilt: $g_{ij} \geq 0$ für $i \neq j$ und $g_{ii} \leq 0$ für alle i. Am besten denkt man an den Poisson-Prozess: die Wahrscheinlichkeit für ein Ereignis in einer sehr kleinen Zeiteinheit $2h$ ist etwa doppelt so gross wie in der Zeiteinheit h (sie ist übrigens $\lambda 2h$ bzw. λh).

Definition 13.10 [infinitesimaler Generator G (Q-Matrix)] *Wir nennen die Matrix G mit Einträgen g_{ij} den infinitesimalen Generator des Markov-Prozesses. In manchen Büchern heisst diese Matrix auch Q-Matrix.*

Von der Bedeutung her übernimmt in Markov-Prozessen in stetiger Zeit das G die Rolle der Übergangsmatrix P in Markov-Ketten in diskreter Zeit.

Es gilt wegen (13.1), (13.2) und (13.3), dass für alle i

$$ 1 = \sum_j p_{ij}(h) \simeq 1 + h \sum_j g_{ij}. $$

Damit folgt aber bei "gleichmässig standard" (vgl. die analoge Eigenschaft bei P):

$$ \sum_j g_{ij} = 0, \tag{13.4} $$

das heisst, die Zeilen in G summieren auf 0, oder nochmals anders: das Diagonalelement ist das negative der Summe der restlichen Elemente.

13.2 Beispiele und elementares Verhalten

Gesucht ist der infinitesimale Generator G_1 eines Poisson-Prozesses der Rate λ.

Geben Sie den infinitesimalen Generator G_2 des Markov-Prozesses von Modell (10.1) aus Kapitel 10 an.

Wir wissen aus Kapitel 10, wie sich der Poisson-Prozess und Modell (10.1) in der Zeit entwickeln. Fragen wir uns dort, wie lange man in einem Zustand i verweilt, stellt man fest, dass es eine exponentialverteilte Zeit ist. Der Parameter dieser Verteilung *ist genau* $(-1)i$*-te Diagonalelement von G, in beiden Beispielen!* Dies ist kein Zufall. Es gilt allgemein:

Satz 13.11 [Verweildauer in Zustand i**]** *Sei X ein Markov-Prozess mit der Eigenschaft "gleichmässig standard" und infinitesimalem Generator G. Gilt $X_t = i$, so ist die Zufallsgrösse*

$$\inf\{s \geq 0 | X_{t+s} \neq i\}$$

exponentialverteilt mit Parameter $-g_{ii}$.

Wir wissen jetzt, wie lange ein Markov-Prozess in einem gewissen Zustand bleibt. Aber wo geht er danach hin? Versuchen Sie anhand obiger beider Beispiele eine allgemeine Formel zu entwickeln, für welche man übrigens mit (13.1) und (13.2) eine einfache (approximative) Beweisskizze angeben kann:

(13.5)

13.3 Stationäre Verteilungen

Im Gegensatz zu den Markov-Ketten in diskreter Zeit haben wir bei den $p_{ij}(t)$ keine grossen Probleme, Irreduzibilität zu definieren; die Periodizität kommt hier in dieser Form gar nicht erst vor (weshalb?).

Es gilt für jedes Paar von Zuständen (i, j): entweder

$$p_{ij}(t) = 0 \quad \forall t > 0$$

oder

$$p_{ij}(t) > 0 \quad \forall t > 0.$$

Warum?

Damit können wir einfach definieren:

Definition 13.12 [Stationäre Verteilung] *Der Vektor π heisst stationäre Verteilung von X, falls $\pi_i \geq 0, \sum_j \pi_j = 1$ und $\pi^t = \pi^t P_t$ für alle $t \geq 0$.*

Mit diesen Verteilungen lässt sich im übrigen rechnen wie in diskreter Zeit: Hat X_0 die Verteilung $\mu^{(0)}$, so ist die Verteilung $\mu^{(t)}$ von X_t gegeben durch

$$\mu^{(t)} = \mu^{(0)} P_t.$$

Wenn man mit der stationären Verteilung von Anfang an beginnt, so bleibt die Verteilung immer gleich dieser stationären Verteilung.

In diskreter Zeit hat man die stationäre Verteilung gefunden, indem man Gleichungen der Art $\pi^t P = \pi^t$ gelöst hat. Das Pendant dazu ist

Theorem 13.13 [Berechnung der Stationären Verteilung] *Es gilt*

$$\pi^t = \pi^t P_t \quad \forall t \qquad \Leftrightarrow \qquad \pi^t G = 0.$$

Wir wollen zu Theorem 13.13 ein kleines Beispiel berechnen. Sei der Zustandsraum die Menge $\{1, 2\}$ und der Markov-Prozess hüpft jeweils immer zwischen Zustand 1 und 2 hin und her. Wir verbleiben in Zustand 1 für eine exponentialverteilte Zeit mit Parameter 1, in Zustand 2 für eine exponentialverteilte Zeit mit Parameter 2. Überlegen Sie sich zuerst heuristisch, wie π hier wohl sein wird und beweisen Sie dies danach.

(13.6)

Zum Schluss wollen wir noch das Langzeitverhalten untersuchen:

Theorem 13.14 [Ergodensatz] *Sei X ein irreduzibler Markov-Prozess mit der Eigenschaft "gleichmässig standard". Dann gilt:*

a) Falls es eine stationäre Verteilung π gibt, so ist sie eindeutig und es gilt $p_{ij}(t) \to \pi_j$ falls $t \to \infty$ für alle i, j.

b) Falls es keine stationäre Verteilung gibt, so gilt: $p_{ij}(t) \to 0$ für $t \to \infty$ für alle i, j.

Welches ist das Analogon zu Theorem 13.14 a) in diskreter Zeit und welche Voraussetzung mussten wir jetzt *nicht* treffen?

13.4 Der Sprungprozess

Wir wollen ganz zum Schluss noch eine alternative Sicht der Markov-Prozesse in stetiger Zeit antippen: wir schauen nur die Sprünge an und interessieren uns nicht dafür, wie lange der Prozess in den einzelnen Zuständen verweilt. Damit sind wir dann wieder bei den Markov-Ketten in diskreter Zeit angelangt (und müssen wieder die allfällige Periodizität beachten). Die Übergangsmatrix kann man von (13.5) ablesen. Anhand von Beispiel (13.6) kann man sich überlegen, wie die stationären Verteilungen des Markov-Prozesses und des dazugehörigen Sprungprozesses zusammenhängen müssen; es gilt nämlich:

13.15 Eine Warnung: Wenn wir die Eigenschaft "gleichmässig standard" fallen lassen, gibt es eventuell die "Explosionsgefahr" bei Markov-Prozessen - ein Phänomen, welches wir in diskreter Zeit so nicht gekannt haben. Dabei kann nicht nur die Zeit, welche man in einem Zustand bleibt, 0 sein. Es finden je nach dem auch unendlich viele Übergänge in endlicher Zeit statt.