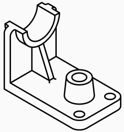
Statistik

Vorlesung 12 - Stochastische Prozesse

Prof. Dr. Sandra Eisenreich

Hochschule Landshut

Motivation Stochastische Prozesse



Meistens hat man nicht ein einzelnes Zufallsexperiment, sondern eine lange Produktionskette von Experimenten, deren Ergebnisse davon von den vorigen Zufallsexperimenten abhängen = stochastischer Prozess

Zwei Arten von stochastischen Prozessen:

- diskrete: z.B. Produktionsprozesse, Warteschlagen, Text-Generierung (siehe Motivation Markov-Ketten), ...
- kontinuierliche: z.B. Atomarer Zerfall, Eintreffen eines Blitzes in der Zeit, Anrufe in einem Callcenter, ...

Stochastische Prozesse

Folgen von Zufallsvariablen, wie zum Beispiel $(X_n)_{n\in\mathbb{N}}$ oder $(X_t)_{t\in\mathbb{R}^+}$ heißen stochastische Prozesse.

- Kontinuierlicher Index: oft die zeitliche Entwicklung einer Zufallsvariablen, t=Beobachtungszeitpunkt
- Diskreter Index: oft Zustände mit Zustandübergängen.

diskrete Prozesse: Markov-Ketten

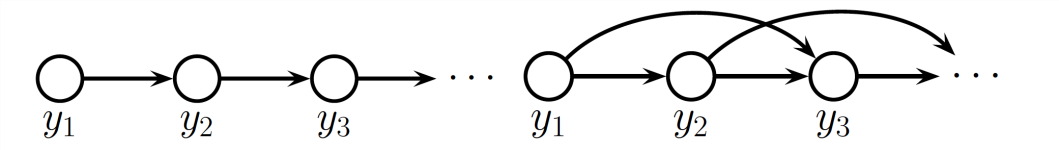
Motivation Markov-Ketten

NLP (Natural Language Processing): Jedes Wort im Text hängt ab von den vorhergehenden $P(X_n|x_1, \ldots, x_{n-1})$

Solch einen diskreten stochastischen Prozess nennt man Markov-Kette n-ter Ordnung, weil jeder Schritt des Prozesses von den letzten n Ereignissen abhängt.

Motivation Markov-Ketten

Die einfachsten Beispiele sind Markov-Ketten erster und zweiter Ordnung:



Wir betrachten hier nur: Markov-Kette (ohne Ordnung bedeutet erster Ordnung)

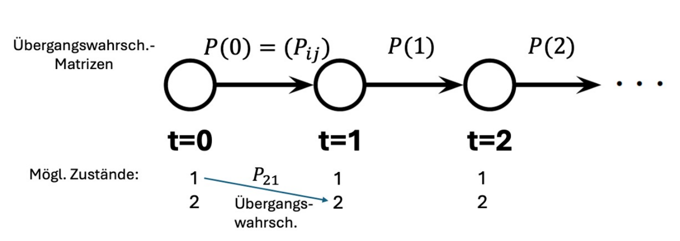
$$P(X_j = i | X_{j-1} = k) = P_{ik} = \text{"Übergangswahrscheinlichkeit von k nach i"}$$

also die Wahrscheinlichkeit, dass vom Zustand k zum Zeitpunkt $j-1$ übergegangen wird in Zustand i zum Zeitpunkt j $\Rightarrow$ Matrix für alle i, j!

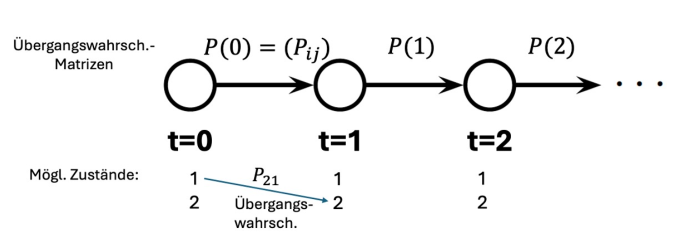
Markov-Ketten

Sei $(X_t)_{t\in\mathbb{N}_0}$ für alle $t \in \mathbb{N}_0$ eine diskrete Zufallsvariable mit Wertebereich W.

- Der stochastische Prozess $(X_t)_{t\in\mathbb{N}_0}$ heißt Markov-Kette, wenn $P(X_{t+1} = k)$ nur von X_t abhängig ist. Davon gehen wir ab jetzt aus.
- Übergangswahrscheinlichkeit von i nach k: $P_{ik}(t) := P(X_{t+1} = k | X_t = i)$.
- Ist $P_{ik}(t) =: P_{ik}$ konstant für alle $t \in \mathbb{N}_0$, so heißt $(X_t)_{t\in\mathbb{N}_0}$ homogen.



Markov-Ketten



Sei $(X_t)_{t \in \mathbb{N}_0}$ eine Markov-Kette wie auf der Vorseite.

- Der Vektor $(a_0, a_1, a_2, \ldots)$ mit $a_i = P(X_0 = i)$ heißt Anfangsverteilung der Kette.
- Ist der Wertebereich W einer homogenen Markov-Kette endlich, so heißt die Matrix $P = (P_{ij})$ Übergangsmatrix der Markov-Kette.
- Die Elemente der Wertemenge W heißen Zustände.

Beispiel: Wetter

Wir betrachten das heutige Wetter als Zufallsvariable mit 3 Zuständen:

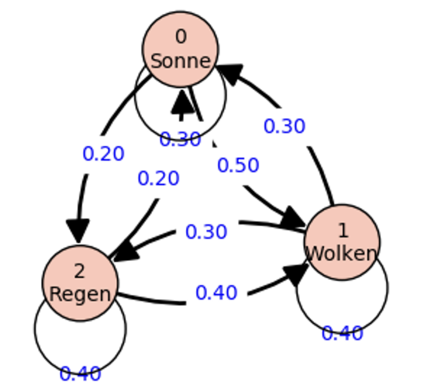
- 0: Sonne den ganzen Tag
- 1: Wolken, aber trocken
- 2: Regen

Markov-Eigenschaft: Das morgige Wetter hängt nur vom Wetter von heute ab, nicht von der Vergangenheit.

Homogen: Die Abhängigkeiten von einem Tag auf den anderen sind für alle Zeiten die selben, z.B. $P(X_{i+1} = \text{Sonne} | X_i = \text{Regen}) = 0.2$ für alle i.

gerichteter Graph

Homogene Markov-Ketten mit endlicher Wertemenge lassen sich in gerichteten Graphen darstellen:



Was ist die Matrix der Übergangswahrscheinlichkeiten?

$$\begin{pmatrix} 0.3 & 0.5 & 0.2 \\ 0.3 & 0.4 & 0.3 \\ 0.2 & 0.4 & 0.4 \end{pmatrix}$$

Zeilensumme muss immer 1 sein!

Wie sieht die Markov-Kette also aus?

Anfangswahrscheinlichkeit für Tag 0:

- $a_0 = P(X_0 = \text{Sonne}) = 0.7$,
- $a_1 = P(X_0 = \text{Wolken}) = 0.2$,
- $a_2 = P(X_0)(X_0 = \text{Regen}) = 0.1$

für Tag 1:

- $P(X_1 = \text{Sonne}|X_0 = \text{Sonne}) = 0.3 \cdot 0.7 = 0.21$,
  $P(X_1 = \text{Sonne}|X_0 = \text{Wolken}) = 0.3 \cdot 0.2 = 0.06$,
  $P(X_1 = \text{Sonne}|X_0 = Regen) = 0.2 \cdot 0.1 = 0.02$
- $P(X_1 = \text{Wolken}|X_0 = \text{Sonne}) = 0.5 \cdot 0.7 = 0.35$,
  $P(X_1 = \text{Wolken}|X_0 = \text{Wolken}) = 0.4 \cdot 0.2 = 0.08$,
  $P(X_1 = \text{ Wolken } |X_0 = \text{Regen}) = 0.4 \cdot 0.1 = 0.04$
- $P(X_1 = \text{Regen}|X_0 = \text{Sonne}) = 0.2 \cdot 0.7 = 0.14$,
  $P(X_1 = \text{Regen}|X_0 = \text{Wolken}) = 0.3 \cdot 0.2 = 0.06$,
  $P(X_1 = \text{Regen}|X_0 = \text{Regen}) = 0.4 \cdot 0.1 = 0.04$

Wegen dem Satz der totalen Wahrscheinlichkeiten gilt also:

$$
\begin{aligned}
P(X_1 = \text{Sonne}) &= P(X_1 = \text{Sonne}|X_0 = \text{Sonne}) \cdot P(X_0 = \text{Sonne}) \\
&+ P(X_1 = \text{Sonne}|X_0 = \text{Wolken}) \cdot P(X_0 = \text{Wolken}) \\
&+ P(X_1 = \text{Sonne}|X_0 = \text{Regen}) \cdot P(X_0 = \text{Regen}) \\
&= 0.21 + 0.06 + 0.02 = 0.29 \\
P(X_1 = \text{Wolken}) &= P(X_1 = \text{Wolken}|X_0 = \text{Sonne}) \cdot P(X_0 = \text{Sonne}) \\
&+ P(X_1 = \text{Wolken}|X_0 = \text{Wolken}) \cdot P(X_0 = \text{Wolken}) \\
&+ P(X_1 = \text{Wolken}|X_0 = \text{Regen}) \cdot P(X_0 = \text{Regen}) \\
&= 0.35 + 0.08 + 0.04 = 0.47 \\
P(X_1 = \text{Regen}) &= P(X_1 = \text{Regen}|X_0 = \text{Sonne}) \cdot P(X_0 = \text{Sonne}) \\
&+ P(X_1 = \text{Regen}|X_0 = \text{Wolken}) \cdot P(X_0 = \text{Wolken}) \\
&+ P(X_1 = \text{Regen}|X_0 = \text{Regen}) \cdot P(X_0 = \text{Regen}) \\
&= 0.14 + 0.06 + 0.04 = 0.24
\end{aligned}
$$

Für den nächsten Tag nehmen wir die obigen Wahrscheinlichkeiten als Ausgangswahrscheinlichkeit, und führen genau die gleiche Rechnung durch wie oben um zu bestimmen, mit welcher Wahrscheinlichkeit es an jedem Tag sonnig, bewölkt oder regnerisch ist.

Hinweis: Nehmen Sie die Übergangsmatrix P und multiplizieren Sie ihn mit der Wahrscheinlichkeitsverteilung $a = (0.7, 0.2, 0.1)$ von Tag 0 von links. Was erhalten Sie?

$$(0.7, 0.2, 0.1) \cdot \begin{pmatrix} 0.3 & 0.5 & 0.2 \\ 0.3 & 0.4 & 0.3 \\ 0.2 & 0.4 & 0.4 \end{pmatrix} = (0.29, 0.47, 0.24)$$

Die Wahrscheinlichkeitsverteilung von Tag 1!

Allgemeiner Fall

Problemstellung: Berechne aus einer gegebenen Anfangsverteilung $a(0) = (a_0, \ldots, a_n)$ die Verteilung zu einem späteren Zeitpunkt.

Dazu nutzen wir wie im Beispiel den Satz der totalen Wahrscheinlichkeit: t sei fest.

- $B_i = \{X_t = i\}$ Die B_i sind disjunkt und $\bigcup B_i = \Omega$
- $A_i = \{X_{t+1} = k\}$

$$P(A) = P(X_{t+1} = k) = \sum_{i=0}^{n} P(X_t = i) P(X_{t+1} = k | X_t = i) = \sum_{i=0}^{n} P(X_t = i) P_{ik}$$

z.B. gilt für $t = 0$: $P(X_1 = k) = \sum_{i=0}^{n} a_i \cdot P_{ik} = ((a_0, \ldots, a_n) \cdot P)_k$

Mit Matrizen-Multiplikation ausgedrückt: Ist $P = (P_{ij})$ die Übergangsmatrix, und $a = (a_0, \ldots, a_n)$, dann bekommen wir die Wahrscheinlichkeitsverteilung von Schritt 1 als folgendes Matrizenprodukt:

$$\begin{aligned} a(1) &:= (P(X_1 = 0), \ldots, P(X_1 = n)) = (\sum_{i=0}^{n} a_i \cdot P_{i0}, \ldots, \sum_{i=0}^{n} a_i \cdot P_{in}) \\ &= (a_0, \ldots, a_n) \cdot P \end{aligned}$$

Und so geht es weiter:

$$\begin{aligned} a(2) &= a(1) \cdot P, \\ a(3) &= a(2) \cdot P, \text{ etc.} \end{aligned}$$

also allgemein:

$$a(t) = a(t-1) \cdot P = a(t-2) \cdot P^2 = \ldots = a(0) \cdot P^t.$$

$\Rightarrow$ Wahrscheinlichkeiten der Zustände zum Zeitpunkt t lassen sich aus der Anfangsverteilung und der Übergangsmatrix berechnen!

Berechnen wir die Potenzen von P und die Verteilungen für unseren Fall mit Anfangsverteilung $a = (0.7, 0.2, 0.1)$:

$$
\begin{aligned}
P = \begin{pmatrix} 0.3 & 0.5 & 0.2 \\ 0.3 & 0.4 & 0.3 \\ 0.2 & 0.4 & 0.4 \end{pmatrix}, \quad a(1) &= a \cdot P = (0.29, 0.47, 0.24) \\
P^2 = \begin{pmatrix} 0.28 & 0.43 & 0.29 \\ 0.27 & 0.43 & 0.3 \\ 0.26 & 0.42 & 0.32 \end{pmatrix}, \quad a(2) &= a \cdot P^2 = (0.2760.4290.295) \\
\vdots \quad & \quad \vdots \\
P^{365} \simeq \begin{pmatrix} 0.270 & 0.427 & 0.303 \\ 0.270 & 0.427 & 0.303 \\ 0.270 & 0.427 & 0.303 \end{pmatrix}, \quad a(365) &= a \cdot P^{365} = (0.270, 0.427, 0.303) \\
P^{366} \simeq \begin{pmatrix} 0.270 & 0.427 & 0.303 \\ 0.270 & 0.427 & 0.303 \\ 0.270 & 0.427 & 0.303 \end{pmatrix}, \quad a(366) &= a \cdot P^{366} = (0.270, 0.427, 0.303)
\end{aligned}
$$

Berechnen wir dasselbe mit anderer Anfangsverteilung $a = (0, 0, 1)$:

$$
\begin{aligned}
P &= \begin{pmatrix} 0.3 & 0.5 & 0.2 \\ 0.3 & 0.4 & 0.3 \\ 0.2 & 0.4 & 0.4 \end{pmatrix}, & a(1) &= a \cdot P = (0.2, 0.4, 0.4) \\
P^2 &= \begin{pmatrix} 0.28 & 0.43 & 0.29 \\ 0.27 & 0.43 & 0.3 \\ 0.26 & 0.42 & 0.32 \end{pmatrix}, & a(2) &= a \cdot P^2 = (0.26, 0.42, 0.32) \\
& & &\vdots \qquad \vdots \\
P^{365} &\simeq \begin{pmatrix} 0.270 & 0.427 & 0.303 \\ 0.270 & 0.427 & 0.303 \\ 0.270 & 0.427 & 0.303 \end{pmatrix}, & a(365) &= a \cdot P^{365} = (0.270, 0.427, 0.303)
\end{aligned}
$$

$\Rightarrow$ Das Ergebnis nach langer Zeit ist dasselbe! Und es gilt:
$(0.270, 0.427, 0.303) \cdot P = (0.270, 0.427, 0.303)$.

Beobachtungen:

- P^n konvergiert für $n \to \infty$ gegen eine feste Matrix P^∞,
- Die Zeilen dieser Matrix P^∞ sind alle gleich
- Diese eine Zeile sind die Zustandswahrscheinlichkeiten nach langer Zeit
- Der Grenzzustand ist unabhängig von der Anfangsverteilung
- Ist p der Endzustand nach langer Zeit, dann gilt $P \cdot p = p$

Der letzte Punkt ist wichtig zum Berechnen des Endzustands p: P ist linker Eigenvektor zum Eigenwert 1 der Matrix P!

Wichtige Eigenschaften von Markov-Ketten

Theorem

Ist $P = (P_{ij})$ die Übergangsmatrix einer homogenen Markov-Kette mit endlichem Wertebereich $W = \{0, \ldots, n\}$ und gibt es ein $m \in \mathbb{N}$, so dass P^m nur Elemente > 0 hat, so gilt:

- *Es existiert der Grenzwert $P^\infty = \lim_{n\to\infty} P^n$. Dabei sind alle Zeilen der Matrix P^x gleich. Die Summe der Elemente dieser Zeile ist 1.*

Für die Zeile $p = (p_0, p_1, \ldots, p_n)$ der Matrix P^∞ gilt:

c *Für jede Anfangsverteilung $a = (a_0, a_1, \ldots, a_n)$ ist $a \cdot P^\infty = p$.*

d *p ist die einzige Anfangsverteilung mit der Eigenschaft $p \cdot P = p$*

Verständnisfragen

- Was bedeutet in der Übergangsmatrix einer Markov-Kette das Element P_{ij}?
- Wenn die Matrixpotenzen P^m einer Übergangsmatrix konvergieren, was steht dann in den Zeilen der Grenzwertmatrix?

Verständnisfragen - Ergebnis

- $P_{ij} = P(X_{t+1=j}|X_t = i)$ ist die Wahrscheinlichkeit, dass das System in Zustand j übergeht, wenn es in Zustand i war.
- Die Zeilen sind gleich, darin steht die konstante Grenzverteilung, das heißt die Summe der Zeileneinträge ist 1.

Kontinuierlicher Index - Poisson-Prozess

Beispiele

Untersuche Ereignisse, die im zeitlichen Verlauf zufällig und unabhängig voneinander eintreten. (wie im Wartezimmer)

Beispiele:

- Taxi kommt vorbeigefahren
- Blitze im Gewitter
- Eintreffen von Bedienwünschen an einem Server
- Anrufe in einem Callcenter
- Atomarer Zerfall

Beispiel: atomarer Zerfall

Wir sehen uns den Prototypen eines Poisson-Prozesses an, den atomaren Zerfall.

Bekannt: Zerfallsrate (wieviele Teilchen zerfallen in einem festen Zeitraum)
Unbekannt: wie verteilen sich die Zerfälle in diesem Zeitraum.

Zahlenbeispiel: In einer Materialprobe sollen 7200 Teilchen pro Stunde zerfallen (Zerfallsrate).

Frage: Wie groß ist die Wahrscheinlichkeit für 0,1,2,3,... Zerfälle pro Sekunde?

Ergebnis: atomarer Zerfall ist Bernoulliexperiment mit 7200 Stufen

Ereignis A: Teilchen zerfällt in der ersten Sekunde.

Zufallsvariable X: Anzahl von Zerfällen in der ersten Sekunde.

- $P(A) = \frac{1}{3600}$ (1 Stunde hat 3600 Sekunden)
- $P(X = k)$ ist binomialverteilt, d.h.

$$P(X = k) = b_{7200, \frac{1}{3600}}(k)$$

Erinnerung: Ist $n \geq 1500p, np \leq 10 \Rightarrow$ Binomialverteilung kann durch Poissonverteilung ersetzt werden:

$$p(X = k) = \frac{\lambda^k}{k!} e^{-\lambda} \text{ wobei } \lambda = n \cdot p$$

Hier: $n = 7200 \geq 1500 \cdot \frac{1500}{3600} \checkmark \;\; n \cdot p = \frac{7200}{3600} = 2 \leq 10 \checkmark$

Es gilt also:

$$p(X = k) = \frac{2^k}{k!} e^{-2}$$

Jetzt sei X_t der stochastische Prozess, der die Anzahl der Zerfälle bis zum Zeitpunkt t beschreibt. (Bis jetzt war $t = 1$, wir haben $P(X_1 = k)$ berechnet.)

Wieder ist $n = 7200$, aber $p = \frac{t}{3600}$ bei t Sekunden $\Rightarrow b_{7200,\frac{t}{3600}}$-verteilt, bzw. mit $\lambda = n \cdot p = 2 \cdot t$ wie oben ist X_t also poisson-verteilt mit Parameter $2t$ für alle t:

$$p(X_t = k) = \frac{(2t)^k}{k!} e^{-2t}$$

Der atomare Zerfall ist der Prototyp einer Klasse von stochastischen Prozessen, der Poisson-Prozesse:

Definition (Poisson-Prozess)

Für $t \in \mathbb{R}^+$ sei X_t die Zufallsvariable, welche die Anzahl des Auftretens eines Ereignisses im Zeitraum von 0 bis t beschreibt. Für diese Ereignisse soll gelten:

- Die Wahrscheinlichkeit für das Eintreten von k Ereignissen in einem Intervall I der Länge t hängt nur von k und und t ab, nicht von der Lage des Intervalls auf der Zeitachser, das heißt $p(X_{s+t} - X_s = k) = p(X_t = k)$.
- Die Anzahlen der Ereignisse in disjukten Zeitintervallen sind voneinander unabhängige Zufallsvariable.
- Es treten niemals zwei Ereignisse zur gleichen Zeit ein.

Dann heißt $(X_t)_{t\in\mathbb{R}^+}$ Poisson-Prozess.

Zusammenfassung Poisson-Prozess

Theorem

- *In einem Poisson-Prozess* $(X_t)_{t\in\mathbb{R}^+}$ *ist die Zufallsvariable* X_t *für alle* t *Poisson-verteilt mit Parameter* λt, *es ist also*

$$P(X_t = k) = \frac{(\lambda t)^k}{k!} e^{-\lambda t}.$$

- *Im Poisson-Prozess ist* λ *der Erwartungswert pro Zeiteinheit (Rate).*

Begründung: $\lambda t =$ Erwartungswert in Poisson-Verteilung, d.h. λt ist der Erwartungswert für $X_t =$ Anzahl der Ereignisse in $[0, t]$.

Beispiel: Call-Center

Call-Center erhält im Mittel 120 Anrufe pro Stunde (Rate), $\lambda = 120/h$. Es ist überlastet, wenn es mehr als 4 Anrufe pro Minute erhält. Wie hoch ist die Wahrscheinlichkeit dafür?

Antwort:

- $\lambda t = \frac{120/60min}{1min} = 2 \Rightarrow p(X_1 = k) = \frac{2^k}{k!}e^{-2}$
- Damit berechnet man: $P(X_1 < 5) = 0.9473 \Rightarrow P(X_1 \geq 5) \simeq 5,3\%$

Wartezeit zwischen Ereignissen

Manchmal interessiert man sich jedoch für die Wartezeit zwischen zwei Ereignissen in einem Poisson-Prozess. Ohne Beweis bringen wir dafür noch folgenden Satz:

Theorem

Es sei $(X_t)_{t \in \mathbb{R}^+}$ ein Poisson-Prozess, also X_t sei Poisson-verteilt mit Parameter λt. Es sei W_i die Zeit zwischen dem i-ten und dem $(i+1)$-ten Eintreten des Ereignisses. Dann ist W_i exponentialverteilt mit Erwartungswert $\frac{1}{\lambda}$.

Beispiel: Kernkraftwerke

Ein Kernkraftwerksbetreiber gibt als mittlere Lebensdauer seines Kernkraftwerks bis zu einem katastrophalen Ausfall des Kühlsystems 10000 Jahre an (das ist eine Phantasie-Zahl). Nehmen wir an, diese Aussage gilt für alle 420 Kernkraftwerksblöcke auf der Welt und nehmen wir weiter an, dass diese Zahl in den nächsten 30 Jahren konstant bleibt. Mit welcher Wahrscheinlichkeit werden wir in den nächsten 30 Jahren noch einen GAU erleben?

Ergebnis: Kernkraftwerke

Die Wartezeit bis zum Ausfall ist exponentialverteilt. Für ein Kraftwerk ist nach Vorlesung $P(X \leq 30) = 1 - e^{-30/10000}$ und $P(X > 30) = e^{-30/10000}$. Die Wahrscheinlichkeit, dass dies für alle KKW gilt ist $(e^{-30/10000})^{420} = 0,28$. Die gesuchte Wahrscheinlichkeit beträgt also etwa 72%.

Literatur

- Hartmann, Peter; Mathematik für Informatiker, Springer-Vieweg; 7. Auflage; 2019
- Georgii, Hans-Otto; Stochastik; de Gruyter, 5. Auflage; 2015
- Krengel, Ulrich; Einführung in die Wahrscheinlichkeitstheorie und Statistik; Springer-Vieweg; 8. Auflage; 2005
- Henze, Norbert; Stochastik für Einsteiger; Springer; 10. Auflage; 2013
- Meintrup, David; Schäfer, Stefan; Stochastik; Springer; 1. Auflage; 2005
- Behrends, Ehrhard; Elementare Stochastik; Springer-Vieweg; 2013