Kapitel IV Stochastische Prozesse

1. Einführung

Wir betrachten zeitliche Folgen von Zufallsexperimenten. Mathematisch beschreibt man diese durch einen so genannten stochastischen Prozess. Darunter versteht man eine Folge von Zufallsvariablen $(X_t)_{t \in T}$, die das Verhalten des Systems zu verschiedenen Zeitpunkten t angeben.

Wenn wir $T = \mathbb{N}_0$ annehmen, sprechen wir von einem stochastischen Prozess mit diskreter Zeit. Lässt man andererseits $T = \mathbb{R}_0^+$ zu, so spricht man von stochastischen Prozessen mit kontinuierlicher Zeit.

Eine besonders einfache Art von stochastischen Prozessen sind so genannte Markov-Ketten. Diese haben die Eigenschaft, dass der nächste Zustand des Prozesses zwar vom aktuellen Zustand abhängen darf, nicht aber von der Historie, d.h. davon, wie der aktuelle Zustand erreicht wurde.

2. Prozesse mit diskreter Zeit

2.1 Einführung

Definition 131

Eine (endliche) Markov-Kette (mit diskreter Zeit) über der Zustandsmenge $S = \{0, \ldots, n-1\}$ besteht aus einer unendlichen Folge von Zufallsvariablen $(X_t)_{t \in \mathbb{N}_0}$ mit Wertemenge S sowie einer Startverteilung q_0 mit $q_0^T \in \mathbb{R}^n$. Die Komponenten von q_0 sind hierbei ≥ 0 und addieren sich zu 1. Für jede Indexmenge $I \subseteq \{0, \ldots, t-1\}$ und beliebige Zustände i, j, s_k $(k \in I)$ gilt

$$\begin{aligned} &\Pr[X_{t+1} = j \mid X_t = i,\ \forall k \in I:\ X_k = s_k] = \\ &\qquad \Pr[X_{t+1} = j \mid X_t = i]\,. \end{aligned} \tag{9}$$

Sind die Werte

$$p_{ij} := \Pr[X_{t+1} = j \mid X_t = i]$$

von t unabhängig, so nennt man die Markov-Kette (zeit)homogen. In diesem Fall definiert man die Übergangsmatrix durch $P = (p_{ij})_{0 \le i,j < n}$. Wenn man $S = \mathbb{N}_0$ zulässt, so spricht man von einer unendlichen Markov-Kette.

Markov-Ketten sind nach Andrey Andreyevich Markov (1856–1922) benannt.

Bedingung (9) heißt Markov-Bedingung und besagt:
Wenn wir den Zustand i zum Zeitpunkt t kennen, so hängt die Übergangswahrscheinlichkeit zum Folgezustand j nur von i und j ab. Die Vergangenheit (Zustände zu Zeitpunkten $< t$) der Markov-Kette spielt keine Rolle. Das „Gedächtnis“ der Markov-Kette besteht also nur aus ihrem aktuellen Zustand und sie „weiß“ nicht, wie sie dorthin gekommen ist.

Bei einer zeithomogenen Markov-Kette hat die (absolute) Zeit t keinen Einfluss auf die Übergangswahrscheinlichkeiten p_{ij}, d.h. das Systemverhalten wird nur durch den aktuellen Zustand bestimmt und nicht durch eine absolute Uhr.

Wahrscheinlichkeitsraum einer Markov-Kette

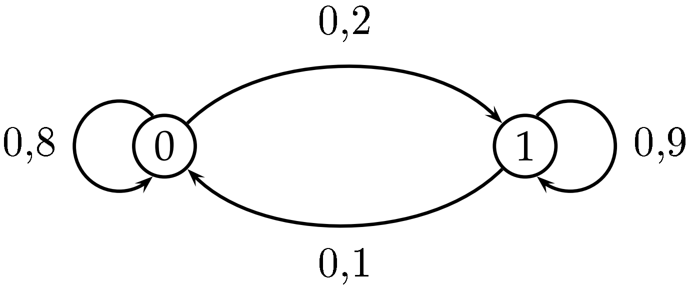
Nehmen wir an, dass wir die Kette von der Zeit 0 bis zur Zeit t_0 beobachten wollen. Wir bezeichnen die Folge von Zuständen, die von der Kette in dieser Zeit durchlaufen wurde, mit $\vec{x} = (x_0, x_1, \ldots, x_{t_0})$. $\Omega \subseteq S^{t_0+1}$ sei die Menge möglicher Zustandsfolgen. Einer beliebigen Folge $\omega := (x_0, x_1, \ldots, x_{t_0}) \in \Omega$ ordnen wir die Wahrscheinlichkeit

$$\Pr[\omega] = (q_0)_{x_0} \cdot \prod_{i=1}^{t_0} \Pr[X_i = x_i \mid X_{i-1} = x_{i-1}]$$

zu. Dadurch erhalten wir einen diskreten Wahrscheinlichkeitsraum im Sinne der Definition.

Beispiel 132

$$\Pr[X_{t+1} = 1 \mid X_t = 1] = 0{,}9,\ \Pr[X_{t+1} = 1 \mid X_t = 0] = 0{,}2$$
$$\Pr[X_{t+1} = 0 \mid X_t = 1] = 0{,}1,\ \Pr[X_{t+1} = 0 \mid X_t = 0] = 0{,}8$$



Einen bestimmten Ablauf des Systems kann man sich als so genannten Random Walk vorstellen.

Wenn wir beispielsweise uns zum Zeitpunkt $t = 0$ im Knoten 1 (also $X_0 = 1$), dann führen von dort zwei Kanten weiter, nämlich zu den Knoten 0 und 1. Diese Kanten sind mit Wahrscheinlichkeiten beschriftet, die sich zu Eins addieren. Gemäß dieser Wahrscheinlichkeiten entscheiden wir zufällig, wohin wir uns im nächsten Schritt begeben.

Wir können auch die Frage beantworten, mit welcher Wahrscheinlichkeit wir uns zum Zeitpunkt $t = 2$ im Knoten 1 befinden. Da wir vereinbarungsgemäß beim Knoten 1 starten, gibt es zwei mögliche Wege der Länge zwei durch den Graphen mit Endknoten 1, nämlich „111“ und „101“. Die Wahrscheinlichkeiten für diese Wege lauten $0{,}9 \cdot 0{,}9 = 0{,}9^2$ bzw. $0{,}1 \cdot 0{,}2$. Insgesamt erhalten wir also eine Wahrscheinlichkeit von $0{,}81 + 0{,}02 = 0{,}83$.

Auch eine Aussage über die erwartete Anzahl Schritte, die wir im Knoten 1 bis zum ersten Übergang zu Knoten 0 verbleiben, ist schnell getroffen. Die Wahrscheinlichkeit, dass man genau k Schritte verbleibt, ist $(0{,}9)^k \cdot 0{,}1$. Die Anzahl Schritte ist also geometrisch verteilt mit Erfolgswahrscheinlichkeit $0{,}1$. Der Erwartungswert ist daher $1/0{,}1 = 10$.

2.2 Berechnung von Übergangswahrscheinlichkeiten

Wir beschreiben die Situation zum Zeitpunkt t durch einen Zustandsvektor q_t (den wir als Zeilenvektor schreiben). Die i-te Komponente $(q_t)_i$ bezeichnet dabei die Wahrscheinlichkeit, mit der sich die Kette nach t Schritten im Zustand i aufhält.

Es gilt

$$\Pr[X_{t+1} = k] = \sum_{i=0}^{n-1} \Pr[X_{t+1} = k \mid X_t = i] \cdot \Pr[X_t = i],$$

also

$$(q_{t+1})_k = \sum_{i=0}^{n-1} p_{ik} \cdot (q_t)_i,$$

bzw. in Matrixschreibweise

$$q_{t+1} = q_t \cdot P.$$

Mit der Matrixschreibweise können wir q_t einfach durch die Startverteilung q_0 ausdrücken:

$$q_t = q_0 \cdot P^t \ .$$

Ebenso gilt wegen der Zeithomogenität allgemein für alle $t, k \in \mathbb{N}$:

$$q_{t+k} = q_t \cdot P^k.$$

Die Einträge von P^k geben an, mit welcher Wahrscheinlichkeit ein Übergang vom Zustand i zum Zustand j in genau k Schritten erfolgt.

$$p_{ij}^{(k)} := \Pr[X_{t+k} = j \mid X_t = i] = (P^k)_{ij}.$$

Exponentiation von Matrizen

Wenn P diagonalisierbar ist, so existiert eine Diagonalmatrix D und eine invertierbare Matrix B, so dass $P = B \cdot D \cdot B^{-1}$ gilt. Diese erhalten wir durch Berechnung der Eigenwerte und Eigenvektoren von P und durch Transformation von P in den Raum der Eigenvektoren.

Dann gilt

$$P^k = B \cdot D^k \cdot B^{-1} \; .$$

Beispiel 133

$$P = \begin{pmatrix} 0{,}8 & 0{,}2 \\ 0{,}1 & 0{,}9 \end{pmatrix}$$

Durch Bestimmung der Nullstellen des charakteristischen Polynoms der Matrix $(P - \lambda \cdot I)$ erhalten wir die Eigenwerte $0{,}7$ und 1, sowie die zugehörigen (rechten) Eigenvektoren

$$\nu_1 = \begin{pmatrix} -2 \\ 1 \end{pmatrix} \text{ und } \nu_2 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

Beispiel 133

Damit

$$D = \begin{pmatrix} 0{,}7 & 0 \\ 0 & 1 \end{pmatrix} \text{ und } B = \begin{pmatrix} -2 & 1 \\ 1 & 1 \end{pmatrix}$$

und

$$B^{-1} = \begin{pmatrix} -\frac{1}{3} & \frac{1}{3} \\ \frac{1}{3} & \frac{2}{3} \end{pmatrix}.$$

Damit ergibt sich beispielsweise

$$P^3 = \begin{pmatrix} -2 & 1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 0{,}7^3 & 0 \\ 0 & 1^3 \end{pmatrix} \begin{pmatrix} -\frac{1}{3} & \frac{1}{3} \\ \frac{1}{3} & \frac{2}{3} \end{pmatrix} \approx \begin{pmatrix} 0{,}562 & 0{,}438 \\ 0{,}219 & 0{,}781 \end{pmatrix}$$