# Algorithmische Mathematik II

Dozent
Professor Dr. Patrik Ferrari

Mitschrift
Maximilian Keßler

Version
10. Mai 2021 11:57

**Zusammenfassung**

Bei folgenden Vorlesungsnotizen handelt es sich um (inoffizielle) Mitschriften zur Vorlesung 'Algorithmische Mathematik II', die im Sommersemester 2021 an der Universität Bonn gehalten wird. Ich garantiere weder für Korrektheit noch Vollständigkeit dieser Notizen, und bin dankbar für jegliche Art von Korrektur, sowohl inhaltlich, als auch Tippfehler.

Bemerkungen, die nicht zum eigentlichen Vorlesungsinhalte gehören, wurden mit einem * gekennzeichnet. Sie werden nach eigenem Ermessen hinzugefügt, um weitere Details oder evtl. mündliche Anmerkungen beizufügen.

Manche Umgebungen sind mit einem : versehen. Das ist dann der Fall, wenn ihr Inhalt so, oder zumindest in sehr ähnlicher Form, in der Vorlesung vorkam (unter Umständen auch mündlich), ich aber die Umgebung der Aussage geändert habe. Das ist z.B. dann der Fall, wenn ich aus Aussagen, die einfach erwähnt werden, ein Lemma: mache, um sie hervorzuheben.

Weitere Informationen finden sich bei GitHub oder auf der Vorlesungshomepage

# Inhaltsverzeichnis

# Übersicht der Vorlesungen

Vorlesung 1
Mi 05 Mai 2021 10:11

## 0.1 Mehrstufige Modelle

Sei eine Folge von $n$ Zufallsexperimenten in den Wahrscheinlichkeitsräumen $\Omega_1, \Omega_2, \dots, \Omega_n$ gegeben. Wir definieren ein $n$-stufiges Zufallsexperiment durch

- $\Omega = \Omega_1 \times \Omega_2 \times \dots \times \Omega_n = \{\omega = (\omega_1, \dots, \omega_n) \mid \omega_k \in \Omega_k, 1 \le k \le n\}$
- $\mathcal{F} = \mathcal{P}(\Omega)$.
- Definiere die Zufallsvariablen

$$X_k(\omega) = \omega_k \quad 1 \le k \le n.$$

  Den Index $k$ interpretieren wir hierbei als Zeit. $k \mapsto X_k$ ist eine Trajektion von $X = (X_1, \dots, X_n)$?
- $\mathbb{P} = ?$. Wir konstruieren $\mathbb{P}$ auf $(\Omega, \mathcal{F})$ mit
  (a) Der Anfangsverteilung $\mathbb{P}(X_1 = x_1) := p_1(x_1)$ für alle $x_1 \in \Omega_1$.
  (b) Den bedingten Verteilungen

$$\mathbb{P}(X_k = x_k \mid X_1 = x_1, X_2 = x_2, \dots, X_{k-1} = x_{k-1}) := p_k(x_k \mid x_1, \dots, x_{k-1}).$$

  für alle $x_l \in \Omega_l$, $1 \le l \le k-1$, sodass $\mathbb{P}(X_1 = x_1, \dots, X_{k-1} = x_{k-1}) > 0$.

i

> **Bemerkung*.** Man kann das Allgemeiner machen, indem wir $\mathcal{F}$ als die Produktsigmalalgebra der $\mathcal{F}_i$ wählen.

**Satz 0.1.** Sei $p_i(\cdot)$ die Massenfunktion einer Wahrscheinlichkeitsverteilung auf $\Omega_i$ und $p_k(\cdot \mid x_1, \dots, x_{k-1})$ für alle $1 \le k \le n$ mit $x_1 \in \Omega_1, \dots, x_{k-1} \in \Omega_{k-1}$ eine Massenfunktion auf $\Omega_k$.
Dann existiert eine eindeutige Wahrscheinlichkeitsverteilung $\mathbb{P}$ auf $(\Omega, \mathcal{F})$, sodass
(a) $\mathbb{P}(X_1 = x_1) = p_1(x_1) \quad \forall x_1 \in \Omega_1$
(b) $\mathbb{P}(X_k = x_k) \mid X_1 = x_1, \dots, X_{k-1} = x_{k-1}) = p_k(x_k \mid x_1, \dots, x_{k-1})$
Die Wahrscheinlichkeitsverteilung $\mathbb{P}$ hat die Massenfunktion

$$p(x_1, \dots, x_n) = p_1(x_1) p_2(x_2 \mid x_1) \cdot \ldots \cdot p_n(x_n \mid x_1, \dots, x_{n-1}).$$

*Beweis.* 1) Nimm zunächst an, dass solch ein Maß existiert, wir zeigen die letzte Aussage. Sei $\mathbb{P}$ sodass (a) und (b) erfüllt sind. Dann ist

$$\forall 1 \le k \le n \colon \mathbb{P}(X_1 = x_1, \dots, X_k = x_k) = p(x_1, \dots, x_k).$$

- Für $k = 1$ gilt das (aus (a)).
- Falls es für $k-1$ gilt, so haben wir die Fälle
- $p(x_1, \dots, x_{k-1}) = 0$, dann ist $0 = 0$ wahr.
- Falls $p(x_1, \dots, x_{k-1}) > 0$, so ist

$$\begin{aligned} \mathbb{P}(X_1 = x_1, \dots, X_k = x_k) &= \mathbb{P}(X_k = x_k \mid X_1 = x_1, \dots, X_{k-1} = x_{k-1}) \, \mathbb{P}(X_1 = x_1, \dots, X_{k-1} = x_{k-1}) \\ &= p_1(x_1) \dots p_{k-1}(x_{k-1} \mid x_1, \dots, x_{k-2}) \cdot p_k(x_k \mid x_1, \dots, x_{k-1}) = p(x_1, \dots, x_k) \end{aligned}$$

Normierung: $\forall x \in \Omega$, $x = (x_1, \dots, x_n)$ mit $x_k \in \Omega_k$ ist

$$\sum_{x\in\Omega} p(x) = \sum_{x_1\in\Omega_1} \dots \sum_{x_n\in\Omega_n} p(x_1, \dots, x_n)$$
$$= \sum_{x_1\in\Omega_1} p(x_1) \sum_{x_2\in\Omega_2} p(x_2 \mid x_1) \dots \sum_{x_n\in\Omega_n} p(x_n \mid x_1, \dots, x_{n-1}) = 1.$$

Für Eigenschaft (b) ist

$$P(X_1 = x_1, \dots, X_k = x_k) = \sum_{x_{k+1}\in\Omega_{k+1}} \dots \sum_{x_n\in\Omega_n} p(x_1, \dots, x_n)$$
$$= p_1(x_1) \dots p_{k-1}(x_{k-1} \mid x_1, \dots, x_{k-2}) p_k(x_k \mid x_1, \dots, x_{k-1})$$
$$= P(X_1 = x_1, \dots, X_{k-1} = x_{k-1}) p_k(x_k \mid x_1, \dots, x_{k-1})$$

Also erhalten wir

$$P(X_k = x_k \mid X_1 = x_1, \dots, X_{k-1} = x_{k-1}) = p_k(x_k \mid x_1, \dots, x_{k-1}).$$

□

Beweis nochmal sortieren

**Anmerkung.** Mir ist noch nicht klar, wo wir im Beweis des Satzes jetzt gezeigt haben wollen, dass solch ein Wahrscheinlichkeitsmaß existiert, das muss ich noch ausarbeiten.

**Bemerkung.** Falls $p_k(x_k \mid x_1, \dots, x_{k-1})$ nur eine Funktion von $x_{k-1}, \dots, x_{k-m-1}$, dann sagen wir, dass unser Modell ein Gedächtnis von $m$ Schritten hat.

### 0.1.1 Produktmodelle

Falls $p_k(x_k \mid x_1, \dots, x_{k-1}) = p_k(x_k)$, d.h. $x_k$ hängt nicht von den Werten $x_1, \dots, x_{k-1}$ ab. Dann erhalten wir aus Satz 0.1, dass

$$p(x_1, \dots, x_n) = \prod_{k=1}^{n} p_k(x_k).$$

**Definition 0.2 (Produktmodell).** Die Wahrscheinlichkeitsverteilung $P$ auf $\Omega = \Omega_1 \times \dots \times \Omega_n$ mit Massenfunktion

$$p(x_1, \dots, x_n) = \prod_{k=1}^{n} p_k(x_k).$$

heißt **Produkt von $P_1, \dots, P_n$**. ($P_k$ hat Massenfunktion $p_k$).

**Notation.** Wir schreiben $P = P_1 \otimes P_2 \otimes \dots \otimes P_n$, wenn $P$ das Produkt von $P_1, \dots, P_n$ ist.

**Beispiel.** Seien $n$ unabhängige 0-1-Experimente mit Erfolgswahrscheinlichkeit $p$ gegeben. Also

$$\Omega_1 = \dots = \Omega_n = \{0, 1\}.$$

und $p_k(1) = p = 1 - p_k(0)$ für $k = 1, \dots, n$. dann ist $p_k(x)$

$(1-p)\left(\frac{p}{1-p}\right)^x$ für $x \in \{0,1\}$. Die entstehende Verteilung

$$p(x_1,\dots,x_n) = (1-p)^n \prod_{k=1}^{n} \left(\frac{p}{1-p}\right)^{x_k}.$$

ist die **$n$-dimensionale Bernoulli-Verteilung mit Parameter $p$**
.

**Satz 0.3.** Sei $(\Omega, \mathcal{F}, P)$ ein Produktmodell. Dann ist für beliebige Ereignisse $A_k \subseteq \Omega_k$, $k = 1,\dots,n$:

$$P(A_1 \times \dots \times A_n) = \prod_{k=1}^{n} P_k(A_k).$$

und $P(\tilde{A}_k) = P_k(A_k)$, wobei

$$\tilde{A}_k := \Omega_1 \times \dots \times \Omega_{k-1} \times A_k \times \Omega_{k+1} \times \dots \times \Omega_n.$$

Deswegen sind $\tilde{A}_1,\dots,\tilde{A}_n$ unabhängige Ereignisse.

*Beweis.* Es ist

$$\begin{aligned} P(A_1 \times \dots \times A_n) &= P((X_1,\dots,X_n) \in A_1 \times \dots \times A_n) \\ &= \sum_{(x_1,\dots,x_n) \in A_1 \times \dots \times A_n} p(x_1,\dots,x_n) \\ &= \sum_{x_1 \in A_1} \dots \sum_{x_n \in A_n} p_1(x_1) \dots p_n(x_n) \\ &= \prod_{k=1}^{n} \sum_{x_k \in A_k} p_k(a_k) \\ &= \prod_{k=1}^{n} P(A_k) \end{aligned}$$

Es ergibt sich leicht

$$\begin{aligned} P(\tilde{A}_k) &= P(X_k \in A_k, X_l \in \Omega_l \ \forall l \neq k) \\ &= \prod_{l \neq k} \underbrace{P(X_l \in \Omega_l)}_{=1} P_k(X_k \in A_k) \\ &= P_k(A_k) \end{aligned}$$

Damit ergibt sich schlussendlich für beliebiges $I = \{i_1,\dots,i_l\}$:

$$\begin{aligned} P(\tilde{A}_{i_1} \cap \dots \cap \tilde{A}_{i_l}) &= P\Big(\bigcap_{i \in I} \{X_i \in A_i\} \cap \bigcap_{j \notin I} \{x_j \in \Omega_j\}\Big) \\ &= \prod_{i \in I} P_i(A_i) \prod_{j \notin I} \underbrace{P_j(\Omega_j)}_{=1} \\ &= \prod_{i \in I} P_i(A_i) \\ &\overset{P(A_i) = P(\tilde{A}_i)}{=} \prod_{k=1}^{l} P(\tilde{A}_{i_k}) \end{aligned}$$

□

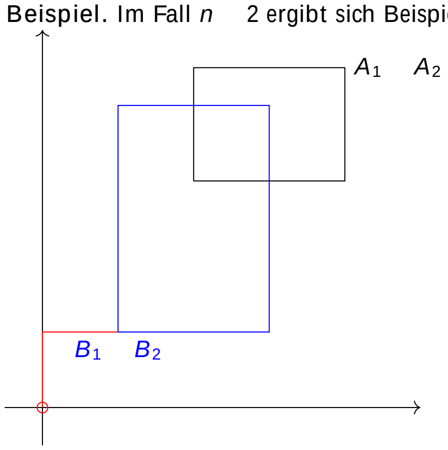
**Bemerkung.** • Es ist $\Omega = \Omega_1 \times \dots \times \Omega_n$
- Eigentlich müssen wir $(\Omega_i, \mathcal{F}_i, \mathbb{P}_i)$ als entsprechende Wahrscheinlichkeitsräume betrachten, wir unterdrücken aber oft die Notation $\mathcal{F}_i, \mathbb{P}_i$

Im Allgemeinen setzen wir

$$\mathcal{F} = \mathcal{F}_1 \otimes \dots \otimes \mathcal{F}_n.$$

wobei $\mathcal{F}$ dann die $\sigma$-Algebra ist, die von $A_1 \times \dots \times A_n$ mit $A_i \in \mathcal{F}_i$ erzeugt ist. Im Spezialfall $\mathcal{F}_i = \mathcal{P}(\Omega_i)$ ergibt sich insbesondere wieder der uns bekannte Fall $\mathcal{F} = \mathcal{P}(\Omega)$.
Für Produktmodelle erhalten wir also $\Omega = \Omega_1 \times \dots \times \Omega_n$, $\mathcal{F} = \mathcal{F}_1 \otimes \dots \otimes \mathcal{F}_n$ sowie $\mathbb{P} = \mathbb{P}_1 \otimes \dots \otimes \mathbb{P}_n$.
Beachte, dass $\mathcal{F}_1 \otimes \dots \otimes \mathcal{F}_n \neq \mathcal{F}_1 \times \dots \times \mathcal{F}_n$ im Allgemeinen, wie das folgende Beispiel zeigt.

**Beispiel.** Im Fall $n = 2$ ergibt sich Beispielsweise folgende Situation:




Vorlesung 2
Mo 10 Mai 2021 10:15

### 0.1.2 Markovketten (MK)

- Setze $X = (X_0, X_1, X_2, \dots, X_n)$. Die Zeit beginnt hier bei $k = 0$.
- Betrachte $\Omega_k = S$ für festes $S$, also

$$\Omega = S^{n+1} = \{(x_0, \dots, x_n) \mid x_i \in S, 0 \le i \le n\}.$$

**Definition 0.4** (Markovkette). Eine **Markovkette** (abgekürzt: MK) ist ein mehrstufiges Modell mit der Eigenschaft

$$p_k(x_k \mid x_0, \dots, x_{k-1}) = p_k(x_k \mid x_{k-1}).$$

**Frage.** Sei $S$ abzählbar. Wie beschreibt man die Übergänge von $X_k$ nach $X_{k+1}$?

**Definition 0.5** (Sotchastische Matrix). Eine Matrix $P = [P(x,y)]_{x,y\in S}$ mit den Eigenschaften

(a) $\forall x\, \forall y: P(x,y) \geq 0$

(b) $\forall x \in S: \sum_{y\in S} P(x,y) = 1$

heißt stochastische Matrix.

**Bemerkung*.** Beachte, dass die Matrix in obiger Definition nicht zwingend endlich sein muss, Definitionen verallgemeinern sich kanonisch. Wir fordern aber, dass $S$ abzählbar ist.

**Lemma 0.6.** Die Matrix $P_k$ mit den Einträgen

$$P_k(x,y) = p_k(Y \mid X) \quad \forall x,y \in S.$$

ist eine stochastische Matrix.

*Beweis.* Offenbar ist $p_k(Y \mid X) \geq 0$. Zudem

$$\begin{aligned}\sum_{y\in S} P_k(x,y) &= \sum_{y\in S} P(X_k = Y \mid X_{k-1} = x) \\ &= P\Big(\bigcup_{y\in S} \{X_k = y\} \mid X_{k-1} = x\Big) \\ &= P(\Omega_k \mid X_{k-1} = x) \\ &= 1\end{aligned}$$

weil es sich bei $P(\_ \mid X_{k-1} = x)$ um eine Wahrscheinlichkeitsverteilung handelt, und $S = \bigcup_{y\in S}\{y\}$ eine disjunkte Vereinigung ist. ☐

**Bemerkung.** $P_k$ ist eine sogenannte Übergangsmatrix. Sie beschreibt den Übergang der Markovkette von $\Omega_k$ nach $\Omega_{k+1}$

Die Massenfunktion einer Markovkette ist

$$p(x_0, x_1, \dots, x_n) = p_0(x_0)\, P_1(x_0, x_1) \cdots P_n(x_{n-1}, x_n).$$

wobei $p_0$ die sogenannte Anfangsverteilung ist.

**Bemerkung.** Falls $P_k = P$, d.h. die Übergangsmatrixk hängt nicht von $k$ ab, dann heißt die Markovkette (zeitlich) homogen.

**Bemerkung.** Seien $P, Q$ zwei stochastische Matrizen. Dann ist auch $P \cdot Q$ eine stochastische Matrix, wobei

$$(P \cdot Q)(x,y) = \sum_{z\in S} P(x,z)\, Q(z,y).$$

**Frage.** Was ist
1) $P(X_n = x)$
2) $\lim_{n\to\infty} P(X_n = x)$ (Existiert dieser überhaupt?)
3) Ist $\lim_{n\to\infty} P(X_n = x)$ von $x_0$ abhängig?

**Satz 0.7** (Massenfunktion in Markovketten)**.** Sei $\mu_0$ der Zeilenvektor mit Elementen $p_0(x)$, $x \in S$. Seien dazu $P_1, P_2, \dots, P_n$ die Übergangsmatrizen einer Markovkette $X = (X_0, X_1, \dots, X_n)$ auf $S$. Dann hat die Wahrscheinlichkeit von $X_n$ die Massenfunkiton

$$\mu_n(x) := P(X_n = x) = (\mu_0 \cdot P_1 \cdot \ldots \cdot P_n)(x) \quad \forall x \in S.$$

*Beweis.*

$$\begin{aligned} P(X_n = x) &= \sum_{x_0,\dots,x_{n-1}\in S} P(X_0 = x_0, \dots, X_{n-1} = x_{n-1}, X_n = x) \\ &= \sum_{x_0,\dots,x_{n-1}\in S} \mu_0(x_0) P_1(x_0, x_1) \dots P_n(x_{n-1}, x) \end{aligned}$$

□

**Beispiel.** (a) Produktmodelle sind Markovketten.

(b) Irrfahrten (eng: 'Random Walk') auf $\mathbb{Z}^d$ sind Markovketten:
Sei $S = \mathbb{Z}^d$ mit $d \in \mathbb{N}$ fest. Die (symmetrische) Irrfahrt ist eine homogene Markovkette mit

$$P_k(x, y) = P(x, y) = \begin{cases} \frac{1}{2d} & \text{falls } |x - y| = 1 \\ 0 & \text{sonst} \end{cases}.$$

In jedem Schritt bewegen wir uns also auf einen benachbarten Gitterpunkt.

(c) Urnenmodell von Ehrenfest. Wir haben ein System mit $N$ Teilchen, die auf zwei Urnen $A, B$ verteilt sind. Zu jedem Zeitpunkt $t \in \mathbb{N}$ welchselt eine zufällig ausgewählte Kugel die Urne.
In der Makroskopischen Modellierung sei $n_A := \#$Teilchen in A. Dann ist

$$\rho_A := \frac{n_A}{N} \in \left\{0, \frac{1}{N}, \dots, \frac{N-1}{N}, 1\right\} =: S.$$

und die Markovkette, die die Zeitentwicklung von $\rho_A$ beschreibt, hat die Übergangsmatrix

$$P(x, y) = \begin{cases} x & \text{falls } y = x - \frac{1}{N} \\ 1 - x & \text{falls } y = x + \frac{1}{N} \\ 0 & \text{sonst} \end{cases}.$$

Die Fälle spiegeln wieder, dass wir ein Teilchen aus $A$ bzw. $B$ gezogen haben, wobei der dritte Fall dijenigen $y$ abdeckt, die wir nicht erreichen können, weil sich $n_A$ immer nur um $\pm 1$ ändert.