Institut für angewandte Mathematik
Sommersemester 2023
Andreas Eberle, Stefan Oberdörster

Klausur „Algorithmische Mathematik II"

Bitte diese Felder in Druckschrift ausfüllen

Name:		Vorname:	
Matrikelnr.:		Studiengang:	

Wichtige Hinweise:

- **Es sind keine eigenen Unterlagen, Handys, Taschenrechner u.ä. zugelassen!**
- Die Klausur enthält 3 Aufgaben, die Sie *alle* bearbeiten sollten.
- Pro Aufgabe können maximal 32 Punkte erreicht werden.
- Dieses Deckblatt ist vollständig ausgefüllt zusammen mit den Lösungen abzugeben. Jedes abgegebene Blatt ist zudem mit Namen und Matrikelnummer zu versehen.
- Bitte den Studierendenausweis und einen amtlichen Lichtbildausweis bereithalten.
- Abgabe bis 11.15 Uhr.

Viel Erfolg!

Diese Felder NICHT ausfüllen:

Aufgabe	1	2	3	**Summe**	**Note**
Punkte					

1. (Zufallsvariablen und Gesetz der großen Zahlen; 32 Punkte)

Sei $S \subseteq \mathbb{R}$ eine abzählbare Teilmenge der reellen Zahlen.

a) Geben Sie eine kurze, aber vollständige Definition der folgenden Begriffe: [*11 Punkte*]

(i) Wahrscheinlichkeitsraum,

(ii) Zufallsvariable X mit Werten in S,

(iii) Erwartungswert von X,

(iv) Varianz von X.

b) Seien $n \in \mathbb{N}$ und $p \in [0, 1]$. [*9 Punkte*]

(i) Berechnen Sie den Erwartungswert und die Varianz für eine Zufallsvariable X mit $\mathbb{P}[X = 1] = p$ und $\mathbb{P}[X = 0] = 1 - p$.

(ii) Zeigen Sie: Für eine Zufallsvariable $Y \sim \mathrm{Bin}(n, p)$ gilt

$$\mathbb{E}[Y] = np \quad \text{und} \quad \mathrm{Var}[Y] = np(1-p)\,.$$

c) Sei $Z : \Omega \to S$ eine Zufallsvariable auf einem Wahrscheinlichkeitsraum $(\Omega, \mathcal{A}, \mathbb{P})$ mit Erwartungswert $m \in \mathbb{R}$ und Varianz $v \in (0, \infty)$. Beweisen Sie die Ungleichung [*4 Punkte*]

$$\mathbb{P}\left[|Z - m| \geq c\right] \;\leq\; \frac{v}{c^2} \qquad \text{für alle } c \in (0, \infty).$$

d) Seien $X_1, X_2, \ldots : \Omega \to S$ unkorrelierte Zufallsvariablen auf $(\Omega, \mathcal{A}, \mathbb{P})$ mit Erwartungswert $m \in \mathbb{R}$ und Varianz $v \in (0, \infty)$. Formulieren und beweisen Sie in diesem Rahmen eine Version des Gesetzes der großen Zahlen. [*8 Punkte*]

Lösung:

a) (i) Ein Wahrscheinlichkeitsraum $(\Omega, \mathcal{A}, \mathbb{P})$ besteht aus einer nicht-leeren Menge Ω, einer σ-Algebra $\mathcal{A} \subseteq \mathcal{P}(\Omega)$, und einer Wahrscheinlichkeitsverteilung $\mathbb{P}$ auf $(\Omega, \mathcal{A})$. Dabei hat eine σ-Algebra $\mathcal{A}$ die folgenden Eigenschaften:

1. $\Omega \in \mathcal{A}$
2. Für alle $A \in \mathcal{A}$ ist $A^c \in \mathcal{A}$.
3. Für alle $A_1, A_2, \ldots \in \mathcal{A}$ ist $\bigcup_{n\in\mathbb{N}} A_n \in \mathcal{A}$.

Eine Wahrscheinlichkeitsverteilung auf $(\Omega, \mathcal{A})$ ist eine Abbildung $\mathbb{P} : \mathcal{A} \to [0, 1]$ mit den folgenden Eigenschaften:

1. $\mathbb{P}[\Omega] = 1$
2. Für alle $A_1, A_2, \ldots$ mit $A_i \cap A_j = \emptyset$ für alle $i \neq j$ gilt

$$\mathbb{P}\left[\bigcup_{n\in\mathbb{N}} A_n\right] = \sum_{n=1}^{\infty} \mathbb{P}[A_n]$$

(ii) Sei $(\Omega, \mathcal{A}, \mathbb{P})$ ein Wahrscheinlichkeitsraum. Eine Zufallsvariable mit Werten in der abzählbaren Menge S ist eine Funktion $X : \Omega \to S$ mit der Eigenschaft, dass

$$X^{-1}(s) \;=\; \{X = s\} \;=\; \{\omega \in \Omega : X(\omega) = s\} \;\in\; \mathcal{A}$$

für alle $s \in S$ gilt.

(iii) Sei X eine nicht-negative Zufallsvariable auf dem Wahrscheinlichkeitsraum $(\Omega, \mathcal{A}, \mathbb{P})$ mit Werten in S. Dann ist der Erwartungswert von X definiert als

$$\mathbb{E}[X] \;:=\; \sum_{s\in S} s \cdot \mathbb{P}[X = s].$$

Für eine allgemeine Zufallsvariable mit Werten in S gilt dieselbe Definition, sofern die Summe wohldefiniert ist. Dies ist der Fall, wenn mindestens einer der beiden Erwartungswerte $E[X^+]$ oder $E[X^-]$ endlich ist.

(iv) Sei X eine Zufallsvariable auf dem Wahrscheinlichkeitsraum $(\Omega, \mathcal{A}, \mathbb{P})$ mit Werten in S mit $\mathbb{E}[|X|] < \infty$. Dann ist die Varianz von X definiert als die mittlere quadratische Abweichung vom Erwartungswert, d.h.

$$\mathrm{Var}[X] := \mathbb{E}\left[(X - \mathbb{E}[X])^2\right].$$

b) (i) Der Erwartungswert von X ist gegeben durch

$$\mathbb{E}[X] = 1 \cdot \mathbb{P}[X = 1] + 0 \cdot \mathbb{P}[X = 0] = p$$

Die Varianz berechnen wir zum Beispiel über die Identität $\mathrm{Var}[X] = \mathbb{E}[X^2] - \mathbb{E}[X]^2$:

$$\mathrm{Var}[X] = 1^2 \cdot p - p^2 = p(1 - p)$$

(ii) Seien $X_1, \ldots, X_n$ unabhängig und $X_i \sim \text{Bernoulli}(p)$ verteilt wie in (i). Dann hat die Summe $X_1 + \ldots + X_n$ dieselbe Verteilung wie Y. Daher folgt mit der Linearität des Erwarungswertes:

$$\mathbb{E}[Y] = \mathbb{E}[X_1 + \ldots + X_n] = \mathbb{E}[X_1] + \ldots + \mathbb{E}[X_n] = n\mathbb{E}[X_1] = np.$$

Die Unabhängigkeit der X_i impliziert deren Unkorreliertheit, sodass gilt

$$\text{Var}[Y] = \text{Var}[X_1 + \ldots + X_n] = \text{Var}[X_1] + \ldots + \text{Var}[X_n] = n\text{Var}[X_1] = np(1-p).$$

c) Es gilt $\mathbb{1}_{\{|Z-m|\geq c\}} \leq |Z-m|^2/c^2$. Zusammen mit der Monotonie und Linearität des Erwartungswertes folgt

$$\mathbb{P}[|Z-m| \geq c] = \mathbb{E}\left[\mathbb{1}_{|Z-m|\geq c}\right] \leq \mathbb{E}[|Z-m|^2]/c^2 = \text{Var}[Z]/c^2 = v/c^2.$$

d) *Schwaches Gesetz der großen Zahlen*: Sei $S_n = \sum_{k=1}^n X_k$. Dann konvergiert $S_n/n \to m$ für $n \to \infty$ in Wahrscheinlichkeit, d.h. für alle $\varepsilon > 0$ ist

$$\lim_{n\to\infty} \mathbb{P}\left[\left|\frac{S_n}{n} - m\right| \geq \varepsilon\right] = 0$$

<u>Beweis:</u> Zunächst berechnen wir mit Hilfe der Linearität des Erwartungswertes

$$\mathbb{E}\left[\frac{S_n}{n}\right] = \frac{1}{n}\sum_{k=1}^n \mathbb{E}[X_k] = \frac{1}{n}nm = m.$$

Da die Zufallsvariablen $X_1, \ldots, X_n$ unkorreliert sind, folgt weiter

$$\text{Var}\left[\frac{S_n}{n}\right] = \frac{1}{n^2}\sum_{k=1}^n \text{Var}[X_k] = \frac{1}{n^2}nv = \frac{v}{n}.$$

Mit Hilfe der Ungleichung aus (c) folgt für alle $\varepsilon > 0$

$$\mathbb{P}\left[\left|\frac{S_n}{n} - m\right| \geq \varepsilon\right] = \mathbb{P}\left[\left|\frac{S_n}{n} - \mathbb{E}\left[\frac{S_n}{n}\right]\right| \geq \varepsilon\right] \leq \frac{1}{\varepsilon^2}\text{Var}\left[\frac{S_n}{n}\right] = \frac{v}{\varepsilon^2 n} \xrightarrow{n\to\infty} 0.$$

■

2. (Stochastische Modelle und Simulationsverfahren; 32 Punkte)

a) 90% der in einer Radarstation eintreffenden Signale sind mit einer Störung überlagerte Nutzsignale, und 10% sind reine Störungen. Wird ein gestörtes Nutzsignal empfangen, zeigt die Anlage mit Wahrscheinlichkeit 0,98 die Ankunft eines Nutzsignals an, und ansonsten die Ankunft eines Störsignals. Beim Empfang einer reinen Störung wird mit Wahrscheinlichkeit 0,1 fälschlicherweise die Ankunft eines Nutzsignals angezeigt. [*16 Punkte*]

(i) Formulieren Sie ein geeignetes stochastisches Modell.

(ii) Mit welcher Wahrscheinlichkeit ist ein als Störsignal angezeigtes Signal tatsächlich ein (störungsüberlagertes) Nutzsignal?

b) Sei u eine Stichprobe von der Gleichverteilung auf dem reellen Intervall $(0, 1)$. Wie können Sie aus u Stichproben von den folgenden Wahrscheinlichkeitsverteilungen erzeugen? (ohne Beweis) [*8 Punkte*]

(i) Gleichverteilung auf einer endlichen Menge S mit n Elementen $x_1, \ldots, x_n$.

(ii) Bernoulli-Verteilung mit Parameter $p = 1/3$.

(iii) Geometrische Verteilung mit Parameter $p = 1/7$.

c) Sei $u_1, u_2, u_3, \ldots$ eine Folge von unabhängigen Stichproben von der Gleichverteilung auf dem reellen Intervall $(0, 1)$. Geben Sie ohne Beweis einen (natürlich möglichst effizienten) Algorithmus an, der daraus eine Stichprobe von der Wahrscheinlichkeitsverteilung μ auf $\mathbb{N}$ mit den Gewichten [*8 Punkte*]

$$\mu\left[\{n\}\right] \; = \; \frac{1}{\mathcal{Z}} |\cos(\sqrt{n})| \cdot \left(\frac{6}{7}\right)^n$$

erzeugt. Dabei ist $\mathcal{Z}$ eine unbekannte Normierungskonstante.

Lösung:

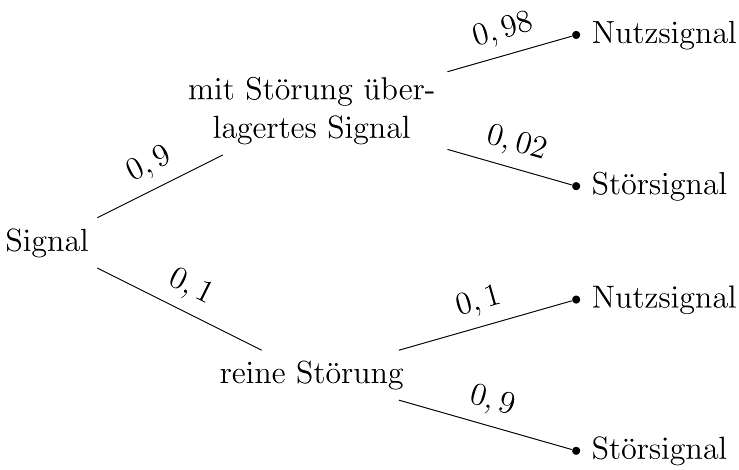
a) (i) Seien $(\Omega, \mathcal{A}, \mathbb{P})$ ein Wahrscheinlichkeitsraum und $X, Y : \Omega \to \{0, 1\}$ zwei Zufallsvariablen. Dabei entspreche das Ereignis $\{X = 1\}$ dem Eintreffen eines (überlagerten) Nutzsignals und $\{X = 0\}$ dem Eintreffen einer reinen Störung. Weiter sei $\{Y = 1\}$ das Ereignis, dass das eingetroffene Signal als Nutzsignal angezeigt wird und $\{Y = 0\}$, dass es als Störung angezeigt wird. Dann ist die Verteilung von X gegeben durch

$$\mathbb{P}[X = 0] = 0,1 \quad \text{und} \quad P[X = 1] = 0,9.$$

und die bedingte Verteilung von Y gegeben X ist gegeben durch

$$\begin{aligned} \mathbb{P}[Y = 0|X = 0] = 0,9 \quad &\text{und} \quad \mathbb{P}[Y = 0|X = 1] = 0,1\,, \\ \mathbb{P}[Y = 1|X = 0] = 0,02 \quad &\text{und} \quad \mathbb{P}[Y = 1|X = 1] = 0,98\,. \end{aligned}$$

Das folgende Pfaddiagramm veranschaulicht das beschriebene mehrstufige Modell:



(ii) Das Ereignis, dass ein als Störsignal angezeigtes Signal tatsächlich ein (störungsüberlagertes) Nutzsignal ist, ist gegeben durch $X = 1$ bedingt auf $Y = 0$. Also ist mit der Bayesschen Regel

$$\begin{aligned} \mathbb{P}[X = 1|Y = 0] &= \frac{P[Y = 0|X = 1]\mathbb{P}[X = 1]}{P[Y = 0|X = 1]\mathbb{P}[X = 1] + P[Y = 0|X = 0]\mathbb{P}[X = 0]} \\ &= \frac{0,02 \cdot 0,9}{0,02 \cdot 0,9 + 0,9 \cdot 0,1} = \frac{0,02}{0,12} = \frac{1}{6} \end{aligned}$$

b) (i) $y := x_i$ für $u \in \left(\frac{i-1}{n}, \frac{i}{n}\right]$ mit $i \in \{1, \dots, n\}$.

(ii) $x := \mathbb{1}_{(0,1/3]}(u)$.

(iii) $t := k$ für $u \in \left((6/7)^k, (6/7)^{k-1}\right]$ mit $k \in \mathbb{N}$, oder $t := \lceil \log u / \log(6/7) \rceil$.

c) Wir können das Acceptance-Rejection-Verfahren verwenden, da

$$\mu[\{n\}] \;=\; \frac{1}{\mathcal{Z}} |\cos(\sqrt{n})| \cdot \left(\frac{6}{7}\right)^n \;\propto\; |\cos(\sqrt{n})| \cdot \mathrm{Geom}(1/7)[\{n\}]\,,$$

und $|\cos(\sqrt{n})| \leq 1$. Dazu benötigen wir unabhängige Stichproben von Geom(1/7), die wir wie in b) erzeugen können.

Algorithmus 1: Acceptance-Rejection-Verfahren (AR)

repeat
    $u, v \leftarrow$`Stichprobe(`Unif$(0,1)$`)`;
    $t \leftarrow \lceil \log u / \log(6/7) \rceil$ Stichprobe von Geom(1/7) ;
until $v \leq |\cos(\sqrt{n})|$;
return t;

3. (Newton-Verfahren; 32 Punkte)

Sei $f : \mathbb{R} \to \mathbb{R}$ eine stetig differenzierbare Funktion.

a) Geben Sie den Algorithmus des Newton-Verfahrens zur Approximation einer Nullstelle x^* der Funktion f an (nur schematisch in Pseudocode - es wird kein lauffähiges Programm erwartet). Der Algorithmus sollte abbrechen, wenn die Iterationsfolge $x^{(k)}$ ein vorgegebenes Intervall $I = (-M, M)$, $M \in (0, \infty)$, verlässt (*overflow*), oder wenn $|f(x^{(k)})|$ eine vorgegebene Schranke $\varepsilon > 0$ unterschreitet. *[8 Punkte]*

b) (i) Was bedeutet *"quadratische Konvergenz des Newton-Verfahrens"* (Definition)? *[8 Punkte]*

(ii) Geben Sie eine Folge reeller Zahlen an, die quadratisch, aber nicht mit höherer Ordnung gegen Null konvergiert (mit Beweis).

c) Sei $x^* \in I$ mit $f(x^*) = 0$. Es gelte $f'(x) > 0$ für alle $x \in [-M, M]$. *[10 Punkte]*

(i) Zeigen Sie: Für $k \geq 0$ mit $x^{(k)} \in I$ gilt

$$x^{(k+1)} - x^* \;=\; \frac{1}{f'(x^{(k)})} \cdot \int_{x^{(k)}}^{x^*} \left(f'(y) - f'(x^{(k)})\right) \, dy.$$

(ii) Folgern Sie: Existiert ein $L \in (0, \infty)$ mit

$$|f'(y) - f'(x)| \leq L \cdot |y - x| \qquad \text{für alle } x, y \in I,$$

dann konvergiert die Iterationsfolge $x^{(k)}$ für $x^{(0)}$ nahe x^* quadratisch gegen x^*.

d) Zeigen Sie anhand eines Beispiels, dass das Newton-Verfahren im Allgemeinen nicht global konvergiert. *[6 Punkte]*

Lösung:

a)

Algorithmus 2: Newton-Verfahren

Input: $f : I \to \mathbb{R}$, Anfangsnäherung $x^{(0)} \in I$, Fehlertoleranz $\epsilon > 0$.
Output: Iterationsfolge $x^{(k)}$, $k = 0, 1, 2, \ldots$

$k \leftarrow 0$;
repeat
  $x^{(k+1)} \leftarrow x^{(k)} - f(x^{(k)})/f'(x^{(k)})$;
  if $x^{(k+1)} \notin I$ **then**
    error („overflow")
until $|f(x^{(k)})| \leq \epsilon$;

b) Die Folge $x^{(k)}$ konvergiert quadratisch gegen x^*, falls $x^{(k)} \to x^*$ und ein $C \in (0, \infty)$ und ein $n \in \mathbb{N}$ existieren, so dass $|x^{(k+1)} - x^*| \leq C|x^{(k)} - x^*|^2$ für alle $k \geq n$.

Die Nullfolge $a_k = e^{-2^k}$ konvergiert quadratisch gegen 0, da $a_{k+1} = a_k^2$ für alle $k \geq 0$. Für jedes $p > 2$ gilt dagegen $a_{k+1} = a_k^{2-p} a_k^p$ mit $a_k^{2-p} \to \infty$ wenn $k \to \infty$.

c) (i) Es gilt

$$\begin{aligned}
x^{(k+1)} - x^* &= x^{(k)} - x^* - \frac{f(x^{(k)})}{f'(x^{(k)})} = x^{(k)} - x^* - \frac{f(x^{(k)}) - f(x^*)}{f'(x^{(k)})} \\
&= \frac{1}{f'(x^{(k)})}\Big(f(x^*) - f(x^{(k)}) - f'(x^{(k)})(x^* - x^{(k)})\Big) \\
&= \frac{1}{f'(x^{(k)})}\int_{x^{(k)}}^{x^*}\big(f'(y) - f'(x^{(k)})\big)\,dy\ .
\end{aligned}$$

(ii) Für $x^{(k)} \in I$ gilt

$$\begin{aligned}
|x^{(k+1)} - x^*| &\leq \frac{1}{|f'(x^{(k)})|}\int_{x^{(k)}}^{x^*}\big|f'(y) - f'(x^{(k)})\big|\,dy \\
&\leq \frac{L}{|f'(x^{(k)})|}\int_{x^{(k)}}^{x^*}|y - x^{(k)}|\,dy \\
&\leq \frac{L}{2\inf_I |f'|}|x^{(k)} - x^*|^2\ .
\end{aligned}$$

Da f' auf dem kompakten Intervall $[-M, M]$ stetig ist, ist das Infimum strikt positiv. Hieraus folgt, dass die Fixpunktiteration eine Kontraktion auf einem Intervall $(x^* - \varepsilon, x^* + \varepsilon) \subset I$ mit $\varepsilon > 0$ ist. Insbesondere wird dieses Intervall nicht verlassen, und für Startwerte $x^{(0)} \in (x^* - \varepsilon, x^* + \varepsilon)$ folgt quadratische Konvergenz.

d) Sei $f(x) = \arctan(x)$ mit Nullstelle in $x^* = 0$. Die Iteration des Newton-Verfahrens ist $x^{(k+1)} = x^{(k)} - (1 + (x^{(k)})^2)\arctan(x^{(k)})$. Falls

$$\arctan|x^{(0)}| > \frac{2|x^{(0)}|}{1 + |x^{(0)}|^2},$$

divergiert die Iterationsfolge.