Institut für angewandte Mathematik
Sommersemester 2017
Andreas Eberle, Matthias Erbar / Behrend Heeren

Klausur „Algorithmische Mathematik II"

Musterlösung

---

1. (Unabhängige Zufallsvariablen)

a) Wir bezeichnen mit S_i den Wertebereich von X_i.

Die Zufallsvariablen $X_1, \ldots, X_n$ heißen *unabhängig*, falls gilt

$$P[X_1 \in A_1, \ldots, X_n \in A_n] = \prod_{i=1}^{n} P[X_i \in A_i] \qquad \forall A_i \subseteq S_i \; .$$

Sie heißen *paarweise unabhängig*, falls X_i und X_j unabhängig sind für alle $i, j = 1, \ldots, n$ mit $i \neq j$.

Die Zufallsvariablen $X_1, \ldots, X_n$ heißen *unkorreliert*, falls gilt $X_i \in L^2(\Omega, \mathcal{A}, P)$ für alle i und

$$\mathrm{Cov}(X_i, X_j) = 0 \qquad \forall i, j = 1, \ldots, n \text{ mit } i \neq j \; .$$

b) (i) Es sei $\Omega = \{1, 2, 3, 4\}, \mathcal{A} = \mathcal{P}(\Omega)$ und P die Gleichverteilung auf Ω. Weiter seien Zufallsvariablen X, Y definiert durch

$$X(1) = 1,\; X(2) = -1,\; X(3) = 0,\; X(4) = 0 \; ,$$
$$Y(1) = 0,\; Y(2) = 0,\; Y(3) = -1,\; Y(4) = 1 \; .$$

Es gilt $E[X] = E[Y] = 0$ und $E[XY] = 0$. Somit sind X und Y unkorreliert. Sie sind jedoch nicht unabhängig, da z.B. gilt

$$P[X = 0, Y = 0] = 0 \neq \frac{1}{4} = P[X = 0]P[Y = 0] \; .$$

(ii) Auf dem selben Wahrscheinlichkeitsraum definieren wir Zufallsvariablen X, Y, Z durch:

$$X(1) = 1,\; X(2) = 1,\; X(3) = 0,\; X(4) = 0 \; ,$$
$$Y(1) = 1,\; Y(2) = 0,\; Y(3) = 1,\; Y(4) = 0 \; ,$$
$$Z(1) = 1,\; Z(2) = 0,\; Z(3) = 0,\; Z(4) = 1 \; .$$

Man überprüft leicht, dass für alle $a, b \in \{0, 1\}$ gilt:

$$P[X = a, Y = b] = \frac{1}{4} = P[X = a]P[Y = b] \; ,$$

Dasselbe gilt für X, Z sowie Y, Z. Somit sind X, Y, Z paarweise unabhängig. Sie sind jedoch nicht unabhängig, da z.B. gilt:

$$P[X = 1, Y = 1, Z = 1] = \frac{1}{4} \neq \frac{1}{8} = P[X = 1]P[Y = 1]P[Z = 1] \; .$$

c) (i) Es gilt

$$P[X \le k] = 1 - P[X > k] = 1 - \sum_{i=k+1}^{\infty} 2^{-i} = 1 - 2^{-k}$$

(ii) Da X und Y unabhängig sind gilt

$$\begin{aligned} P[\min(X,Y) \le k] &= 1 - P[\min(X,Y) > k] = 1 - P[X > k, Y > k] \\ &= 1 - P[X > k]\, P[Y > k] = 1 - 2^{-k}2^{-k} = 1 - 4^{-k}. \end{aligned}$$

(iii) Es gilt

$$P[X = Y] = \sum_{k=1}^{\infty} P[X = k, Y = k] = \sum_{k=1}^{\infty} P[X = k]P[Y = k] = \sum_{k=1}^{\infty} 4^{-k} = \frac{1}{3}\ .$$

(iv) Es gilt:

$$P[X \ne Y] = 1 - P[X = Y] = 2/3,$$

und damit wegen $(X,Y) \sim (Y,X)$:

$$P[X < Y] = (P[X < Y] + P[X > Y])/2 = P[X \ne Y]/2 = 1/3.$$

d)

(i) Setze $x := k$ falls $u \in (2^{-k}, 2^{1-k}]$. Dann ist x eine Stichprobe von einer ZV X mit

$$P[X = k] = P[2^{-k} < U \le 2^{1-k}] = P[U \le 2^{1-k}] - P[U \le 2^{-k}] = 2^{1-k} - 2^{-k} = 2^{-k}.$$

(ii) Es kann das *direkte Verfahren* verwendet werden. Zu einer gegebenen Wahrscheinlichkeitsverteilung μ auf $S = \{a_1, a_2, \ldots\}$ definieren wir $s : \mathbb{N} \cup \{0\} \to \mathbb{R}$ via

$$s(0) = 0\ ,\ \ s(n) = \sum_{i=1}^{n} \mu(a_i)\ .$$

Wir definieren eine Zufallsvariable $Z : \Omega \to S$ durch

$$Z(\omega) = a_n\ , \qquad \text{falls } s(n-1) < U(\omega) \le s(n)\ .$$

Dann hat Z die Verteilung μ. Denn für alle $n \in \mathbb{N}$ gilt:

$$\begin{aligned} P[Z = a_n] &= P[s(n-1) < U \le s(n)] = P[U \le s(n)] - P[U \le s(n-1)] \\ &= s(n) - s(n-1) = \mu(a_n)\ . \end{aligned}$$

Um das direkte Verfahren hier anwenden zu können, wählen wir eine Abzählung $a_i = (n_i, m_i)$, $i \in \mathbb{N}$, von $S = \mathbb{N} \times \mathbb{N}$ und betrachten die WV $\mu(n,m) = 2^{-n}2^{-m}$ auf $\mathbb{N} \times \mathbb{N}$. Dann erhalten wir eine Zufallsvariable $Z = (X,Y)$ mit Werten in $\mathbb{N} \times \mathbb{N}$, so dass

$$P[X = n, Y = m] = P[Z = (n,m)] = 2^{-n}2^{-m}\ .$$

Damit sind X und Y unabhängig mit Massenfuntion p.

2. (Bedingte Wahrscheinlichkeiten)

a) Die bedingte Wahrscheinlichkeit von A gegeben B ist definiert als

$$\mu[A|B] \ := \ \frac{\mu[A \cap B]}{\mu[B]} \ .$$

b) **Bayessche Formel**: Für $A \subset S$ mit $\mu[A] > 0$ gilt:

$$\mu[H_i|A] = \frac{1}{c} \cdot \mu[H_i] \cdot \mu[A|H_i] \ , \text{ wobei}$$
$$c = \sum_{j,\mu[H_j]\neq 0} \mu[A|H_j] \cdot \mu[H_j] \ .$$

Beweis: Aus der Additivität von μ und da die H_i disjunkt sind mit $\cup_i H_i = S$, folgt mit a):

$$c = \sum_{i,\mu[H_i]\neq 0} \mu[A|H_i] \cdot \mu[H_i] = \sum_i \mu[A \cap H_i] = \mu[A \cap \bigcup_i H_i] = \mu[A] \ .$$

Damit erhält man nach Teil a):

$$\mu[H_i|A] = \mu[A]^{-1} \cdot \mu[H_i \cap A] = \mu[A]^{-1} \cdot \mu[H_i] \cdot \mu[A|H_i] = \frac{1}{c} \cdot \mu[H_i] \cdot \mu[A|H_i] \ .$$

c)

(i) Wir modellieren das Experiment auf folgendem Wahrscheinlichkeitsraum: Setze $p_1 = \frac{1}{2}$, $p_2 = \frac{3}{5}$, $p_3 = \frac{1}{10}$ und definiere

$$\Omega = \{1,2,3\} \times \{0,1\}^3 \ , \quad \mathcal{A} = P(\Omega) \ ,$$
$$P[(i,k_1,k_2,k_3)] = \frac{1}{3} p_i^{k_1+k_2+k_3} (1-p_i)^{3-k_1-k_2-k_3} \ .$$

Hierbei steht 1 für Kopf und 0 für Zahl. Weiter betrachten wir die Ereignisse

$$A_i = \text{Münze i wird ausgewählt} \ , \quad i = 1,2,3 \ .$$
$$B = \text{Es fallen Z,Z,K} \ .$$

Damit ergibt sich

$$P[B] = P[B \cap A_1] + P[B \cap A_2] + P[B \cap A_3]$$
$$= \frac{1}{3} \left[\left(\frac{1}{2}\right)^3 + \frac{3}{5} \left(\frac{2}{5}\right)^2 + \frac{1}{10} \left(\frac{9}{10}\right)^2 \right] = \frac{151}{1500} \ .$$

(ii) Gesucht ist die bedingte Wahrscheinlichkeit $P[A_1|B]$. Nach Definition ergibt sich hierfür:

$$P[A_1|B] = \frac{P[A_1 \cap B]}{P[B]} = \frac{1}{3} \cdot \frac{1}{8} \cdot \frac{1500}{151} = \frac{125}{302} \ .$$

d)

(i) Die Zufallsvariable T ist geometrisch verteilt mit Parameter $p = \frac{1}{2}$. Zum Beweis definieren wir Zufallsvariablen X_n, die den Wert der Münze in der n-ten Runde angeben. Dann sind die X_n unabhängig und Bernoulli verteilt mit Parameter $\frac{1}{2}$. Wir berechnen $P[T = 1] = P[X_1 = 0] = \frac{1}{2}$ und für $k > 1$:

$$P[T = k] = P[X_1 = \cdots = X_{k-1} = 1, X_k = 0] = 2^{-k} .$$

Damit ist T geometrisch verteilt mit Parameter $\frac{1}{2}$.

(ii) Es sei A_n die Augenzahl des Würfels in der n-ten Runde. Die Zufallsvariablen $A_1, A_2, \ldots$ sind unabhängig und uniform verteilt auf $\{1, \ldots, 6\}$ sowie unabhängig von T. Es gilt

$$G = \sum_{n=1}^{T} A_n .$$

Wir berechnen

$$\begin{aligned} E[G] &= \sum_{k=1}^{\infty} E[G|T = k]P[T = k] = \sum_{k=1}^{\infty} \sum_{n=1}^{k} E[A_n]P[T = k] = \sum_{k=1}^{\infty} kE[A_1]P[T = k] \\ &= E[A_1]E[T] = \frac{1}{6}(1 + 2 + 3 + 4 + 5 + 6) \cdot 2 = 7. \end{aligned}$$

(iii) Wir erhalten

$$\begin{aligned} E[G^2] &= \sum_{k=1}^{\infty} E[G^2|T = k]P[T = k] = \sum_{k=1}^{\infty} \sum_{n,m=1}^{k} E[A_n A_m]P[T = k] \\ &= \sum_{k=1}^{\infty} \left[kE[A_1^2] + k(k-1)E[A_1 A_2]\right] P[T = k] \\ &= \sum_{k=1}^{\infty} \left[kE[A_1^2] + k(k-1)E[A_1]^2\right] P[T = k] \\ &= E[A_1^2]E[T] + E[A_1]^2 \sum_{k=1}^{\infty} k(k-1)2^{-k} . \\ &= \frac{1}{6}(1 + 4 + 9 + 16 + 25 + 36) \cdot 2 \, + \left(\frac{7}{2}\right)^2 \cdot 4 \\ Var[G] = E[G^2] - E[G]^2 &= \frac{1}{6}(1 + 4 + 9 + 16 + 25 + 36) \cdot 2 = \frac{91}{3}. \end{aligned}$$

3. (Interpolation und Quadratur)

a) Die dividierten Differenzen berechnen sich rekursiv via

$$f[x_i,\dots,x_k] = \frac{f[x_{i+1},\dots,x_k] - f[x_i,\dots,x_{k-1}]}{x_k - x_i}, \quad 0 \le i < k \le n,$$

wobei $f[x_i] = f_i$ für $i = 0,\dots,n$. Es ergibt sich das Dreiecksschema

$$\begin{array}{r|lll}
-2 & f[x_0]=2 & & \\
-1 & f[x_1]=1 & f[x_0,x_1]=\frac{f[x_1]-f[x_0]}{x_1-x_0}=-1 & \\
1 & f[x_2]=1 & f[x_1,x_2]=\frac{f[x_2]-f[x_1]}{x_2-x_1}=0 & f[x_0,x_1,x_2]=\frac{f[x_1,x_2]-f[x_0,x_1]}{x_2-x_0}=1/3 \\
2 & f[x_3]=1 & f[x_2,x_3]=\frac{f[x_3]-f[x_2]}{x_3-x_2}=0 & f[x_1,x_2,x_3]=\frac{f[x_2,x_3]-f[x_1,x_2]}{x_3-x_1}=0
\end{array}$$

und

$$f[x_0,x_1,x_2,x_3] = \frac{f[x_1,x_2,x_3] - f[x_0,x_1,x_2]}{x_3 - x_0} = \frac{0 - 1/3}{4} = -1/12\,.$$

Das Interpolationspolynom in Newtonscher Darstellung lautet daher

$$\begin{aligned}
p_{0,1,2,3}(x) &= f[x_0] + f[x_0,x_1]\,(x-x_0) + f[x_0,x_1,x_2]\,(x-x_0)\,(x-x_1) \\
&\qquad + f[x_0,x_1,x_2,x_3]\,(x-x_0)\,(x-x_1)\,(x-x_2) \\
&= 2 - (x+2) + \frac{1}{3}(x+2)(x+1) - \frac{1}{12}(x+2)(x+1)(x-1).
\end{aligned}$$

b) Da $n+1$ verschiedene Stützstellen ausreichen, um ein Polynom n-ten Grades exakt zu interpolieren, gilt dies insbesondere für das Monom $f(x) = x^n$. Das Interpolationspolynom $p \in \Pi_n$ ist in der Newton-Darstellung gegeben durch

$$p(x) = f[x_0] + \sum_{j=1}^{n} f[x_0,\dots,x_j] \prod_{i=0}^{j-1} (x - x_i)\,.$$

Da in der Summe der n-te Summand der einzige ist, der einen Term x^n beinhaltet und zudem $\prod_{i=0}^{n-1}(x-x_i) = x^n + q_{n-1}(x)$ gilt mit $q_{n-1} \in \Pi_{n-1}$, muss für den n-ten Vorfaktor $f[x_0,\dots,x_n] = 1$ gelten.

c) (i) Es sei $\|\cdot\|_\infty$ die Supremumsnorm auf besagtem Intervall. Dann gilt nach VL allgemein für Polynominterpolation n-ten Grades:

$$|f_\alpha(x) - p_\alpha(x)| \le \frac{\|f_\alpha^{(n)}\|_\infty}{(n+1)!} \prod_{i=0}^{n} |x - x_i|.$$

Hier also mit $n = 1$ und gegebenen Stützstellen

$$|f_\alpha(x) - p_\alpha(x)| \le \frac{\|f_\alpha''\|_\infty}{2} |x+2||x+1|$$

Ein einfache Symmetrieüberlegung ergibt, dass der Ausdruck $|x+2||x+1|$ im Intervall $[-2,-1]$ im Punkt $x=-1.5$ maximal ist. Somit

$$|f_\alpha(x) - p_\alpha(x)| \leq \frac{\|f''_\alpha\|_\infty}{2}\frac{1}{2}\frac{1}{2} = \frac{\|f''_\alpha\|_\infty}{8}$$

Es ist $f'_\alpha(x) = -\frac{1}{(\alpha^2-x)^2}$, $\quad f''_\alpha(x) = \frac{2}{(\alpha^2-x)^3}$, also

$$\|f''_\alpha\|_\infty = \frac{2}{(\alpha^2-(-1))^3} = \frac{2}{(\alpha^2+1)^3}.$$

Insgesamt erhält man also auf dem Intervall $[-2,-1]$ die Abschätzung

$$|f_\alpha(x) - p_\alpha(x)| \leq \frac{1}{4(\alpha^2+1)^3}.$$

(ii) Damit diese Schranke den Wert 2^{-5} nicht überschreitet, muss $(\alpha^2+1) \geq 2$ gelten, also

$$|\alpha| \geq 1.$$

d) (i) Bestimme $p_2 \in \Pi_2$ mit Hilfe der Dreitermrekursion (Satz 5.26):

$$\begin{aligned}
p_0(x) &= 1\,,\\
p_1(x) &= (x-\alpha_0)p_0(x)\,, \quad \alpha_0 = \frac{(xp_0,p_0)_w}{(p_0,p_0)_w} = \frac{\int_{-1}^1 x^3\,\mathrm{d}x}{\int_{-1}^1 x^2\,\mathrm{d}x} = 0\\
&= x\,,\\
p_2(x) &= (x-\alpha_1)p_1(x) - \beta_1 p_0(x)\,, \quad \alpha_1 = \frac{(xp_1,p_1)_w}{(p_1,p_1)_w} = \frac{\int_{-1}^1 x^5\,\mathrm{d}x}{\int_{-1}^1 x^4\,\mathrm{d}x} = 0\\
&\qquad\qquad\qquad\qquad\qquad\qquad\qquad \beta_1 = \frac{(p_1,p_1)_w}{(p_0,p_0)_w} = \frac{\int_{-1}^1 x^4\,\mathrm{d}x}{\int_{-1}^1 x^2\,\mathrm{d}x} = \frac{3}{5}\\
&= x^2 - \frac{3}{5}
\end{aligned}$$

(ii) Die Nullstellen $x_1 = -\sqrt{3/5}$ und $x_2 = \sqrt{3/5}$ von p_2 liefern die Stützstellen der Gauß-Quadratur. Nach Satz 5.20 der VL berechnen sich a und b als (gewichtete) Integration der ersten beiden Lagrange-Polynome zu den Stützstellen x_1 und x_2, d.h.

$$a = \int_{-1}^1 L_0(x)x^2\,\mathrm{d}x = -\sqrt{\frac{5}{12}}\int_{-1}^1\left(x-\sqrt{\frac{3}{5}}\right)x^2\,\mathrm{d}x = -\sqrt{\frac{5}{12}}\cdot\left(-2\sqrt{\frac{3}{5}}\frac{1}{3}\right) = \frac{1}{3}$$

$$b = \int_{-1}^1 L_1(x)x^2\,\mathrm{d}x = \sqrt{\frac{5}{12}}\int_{-1}^1\left(x+\sqrt{\frac{3}{5}}\right)x^2\,\mathrm{d}x = \sqrt{\frac{5}{12}}\cdot\left(2\sqrt{\frac{3}{5}}\frac{1}{3}\right) = \frac{1}{3}$$

4. (Matrixnormen und Iterationsverfahren)

a) Für die Spektralnorm gilt:

$$\|T_\alpha\|_2 = \sqrt{\rho(T_\alpha^T T_\alpha)}.$$

Es ist

$$T_\alpha^T T_\alpha = \begin{pmatrix} 2\alpha^2 & -\alpha^2 & 0 \\ -\alpha^2 & 2\alpha^2 & 0 \\ 0 & 0 & 3\alpha^2 \end{pmatrix} = \alpha^2 \begin{pmatrix} 2 & -1 & 0 \\ -1 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix}$$

und somit

$$\det(T_\alpha^T T_\alpha - \lambda I) = ((2\alpha^2 - \lambda)^2 - \alpha^4)(3\alpha^2 - \lambda/\alpha^2).$$

Die Nullstellen dieses Polynoms in λ sind die Eigenwerte von $T_\alpha^T T_\alpha$:

$$\lambda = 3\alpha^2 \quad \text{(doppelt)}, \quad \lambda = \alpha^2$$

Es gilt daher genau dann $\|T_\alpha\|_2 < 1$, wenn

$$\rho(T_\alpha^T T_\alpha) = 3\alpha^2 < 1 \quad \Leftrightarrow \quad |\alpha| < \frac{1}{\sqrt{3}}.$$

b) (i) Sei $\|x\|_\infty = |x_i|$. Dann gilt, unter Verwendung von Dreiecksungleichungen,

$$\begin{aligned} \|Ax\|_\infty \geq |Ax|_i = \left|\sum_j a_{ij}x_j\right| &\geq |a_{ii}x_i| - \left|\sum_{j\neq i} a_{ij}x_j\right| \\ &\geq |a_{ii}||x_i| - \sum_{j\neq i} |a_{ij}||x_i| \\ &= \left(|a_{ii}| - \sum_{j\neq i} |a_{ij}|\right) \|x\|_\infty \geq c\|x\|_\infty. \end{aligned}$$

(ii) Zunächst ist A invertierbar, da die Ungleichung aus (i) eine nichttriviale Lösung der Gleichung $Ax = 0$ ausschließt.

Es ist $\kappa_\infty(A) = \|A\|_\infty \|A^{-1}\|_\infty$, und nach (i)

$$\|A^{-1}\|_\infty = \max_{x\neq 0} \frac{\|A^{-1}x\|_\infty}{\|x_\infty\|} \overset{x=Ay}{=} \max_{y\neq 0} \frac{\|y\|_\infty}{\|Ay\|_\infty} \leq \max_{y\neq 0} \frac{\|y\|_\infty}{c\|y\|_\infty} = \frac{1}{c}.$$

Daraus folgt die Behauptung.

c) (i) Es gilt, unter Verwendung der Symmetrie von A,

$$f(\alpha) = \phi(x + \alpha p) = \frac{1}{2}\alpha^2(p, Ap) + \alpha(Ax - b, p) + \phi(x)$$

Da A positiv definit ist und $p \neq 0$, d.h. $(p, Ap) > 0$, ist dies eine strikt konvexe Parabel, deren globales Minimum die Nullstelle der ersten Ableitung ist:

$$0 \overset{!}{=} f'(\alpha_*) = \alpha_*(p, Ap) + (Ax - b, p) \quad \Rightarrow \quad \alpha_* = -\frac{(Ax - b, p)}{(p, Ap)}.$$

(ii) Aus (i) ergeben sich durch Einsetzen von $p = e_i$ die Schrittweiten

$$\alpha_i^{(k)} = -\frac{\sum_j a_{ij} x_j^{(k)} - b_i}{a_{ii}}.$$

(iii) Das hier beschriebene Verfahren lautet in Koordinatenschreibweise (mit (ii)):

$$x_i^{(k+1)} = x_i^{(k)} - \frac{1}{a_{ii}} \left(\sum_{j=1}^{n} a_{ij} x_j^{(k)} - b_i \right), \quad i = 1, \dots, n,$$

welches identisch mit dem Jacobi-Verfahren ist.

d) Man nehme zunächst an, dass $|f'(x^*)| < 1$. Da f' stetig ist, gibt es ein $\delta > 0$, so dass $f'(x) \leq q < 1$ für alle x mit $|x - x^*| \leq \delta$. Sei nun $|x^n - x^*| \leq \delta$. Dann gilt mit Mittelwertsatz

$$|x_{k+1} - x^*| = |f(x_k) - f(x^*)| = |f'(\xi)(x_k - x^*)| \leq q|x_k - x^*|.$$

Ist nun $|x_0 - x^*| \leq \delta$, ergibt sich durch Induktion

$$|x_k - x^*| \leq q^k |x_0 - x^*| \to 0,$$

also lokale Konvergenz der esten Fixpunktiteration.

Alternative: lokale Version des Banachschen Fixpunktsatzes (dann sind die Voraussetzungen zu prüfen, was im Wesentlichen auf die selben Argumente hinausläuft).

Falls $|f'(x^*)| > 1$, so ist nach Umkehrsatz $|(f^{-1})'(x^*)| = |1/f'(x^*)| < 1$ und die Argumente lassen sich für die zweite Fixpunktiteration wiederholen.