Institut für angewandte Mathematik
Sommersemester 2009
Andreas Eberle, Martin Slowik

Klausur zur „Algorithmische Mathematik II"

Lösungen

---

1. (Urnenmodelle)

a) Sei $\Omega = \{1, \ldots, m\}^n$ mit $|\Omega| = m^n$ eine Nummerierung der m Kugeln, $\mathcal{A} = \mathcal{P}(\Omega)$ die Potenzmenge und $P[A] = |A|/|\Omega|$ die Gleichverteilung auf $(\Omega, \mathcal{A})$. Die Menge der gelben Kugeln sei hierbei gegeben durch $\{1, \ldots, g\}$. Weiterhin betrachte die Zufallsvariablen $X_1, \ldots, X_n$ mit [5 *Pkt*]

$$X_i : \Omega \;\to\; \{0, 1\}, \qquad \omega \;\mapsto\; X_i(\omega) := \mathbb{1}_{\{\omega_i \leq g\}} \quad \text{für alle } 1 \leq i \leq n.$$

Die Zufallsvariable X_k besitzt die Verteilung

$$P\big[X_i = 1\big] \;=\; P\big[\{\omega \in \Omega \mid \omega_i \leq g\}\big] \;=\; \frac{\big|\{\omega \in \Omega \mid \omega_i \leq g\}\big|}{|\Omega|} \;=\; \frac{g}{m} \;=:\; p.$$

Damit ist dann insbesondere $P[X_1 = 1] = p$. Weiterhin ergibt sich

$$P\big[X_1 = 1, X_2 = 1\big] \;=\; \frac{\big|\{\omega \in \Omega \mid \omega_1 \leq g,\, \omega_2 \leq g\}\big|}{|\Omega|} \;=\; \frac{g^2}{m^2} \;=\; p^2.$$

b) Die Anzahl der gelben unter den n gezogenen Kugeln wird beschrieben durch die Zufallsvariable $S_n : \Omega \to \mathbb{N}_0$, $S_n = \sum_{i=1}^n X_i$. Es soll nun der Erwartungswert und die Varianz dieser Zufallsvariable bestimmt werden. [4 *Pkt*]

Lösungsweg 1:

Aus der Linearität des Erwartungswertes ergibt sich

$$E\big[S_n\big] \;=\; E\Big[\sum_{i=1}^n X_i\Big] \;=\; \sum_{i=1}^n E\big[X_i\big] \;=\; n\, E\big[X_1\big] \;=\; n\, P\big[X_1 = 1\big] \;=\; n\, p,$$

wobei im dritten Schritt zudem benutzt wurde, dass die Zufallsvariablen X_i die gleiche Verteilung besitzten und dass $E[X_1] = 0 \cdot P[X_1 = 0] + 1 \cdot P[X_1 = 1] = P[X_1 = 1]$ ist. Zur Berechnung der Varianz betrachte zunächst einmal

$$\begin{aligned} E\big[S_n^2\big] \;&=\; E\Big[\sum_{i,j=1}^n X_i\, X_j\Big] \;=\; \sum_{i=1}^n E\big[X_i^2\big] + \sum_{i=1}^n \sum_{\substack{j=1 \\ j \neq i}}^n E\big[X_i\, X_j\big] \\ &=\; \sum_{i=1}^n P\big[X_i = 1\big] + \sum_{i=1}^n \sum_{\substack{j=1 \\ j \neq i}}^n P\big[X_i = 1, X_j = 1\big] \;=\; n\, p + n\, (n-1)\, p^2, \end{aligned}$$

wobei im zweiten Schritt die Linearität des Erwartungswertes und im dritten Schritt, dass $E[X_i^2] = P[X_i = 1]$ und $E[X_i X_j] = P[X_i = 1, X_j = 1]$ ist, benutzt wurde. Damit ergibt sich dann für die Varianz

$$\mathrm{Var}(S_n) \;=\; E\big[S_n^2\big] - E\big[S_n\big]^2 \;=\; n\,p + n\,(n-1)\,p^2 - n^2\,p^2 \;=\; n\,p\,(1-p).$$

Lösungsweg 2:
Da die Zufallsvariable S_n binomial-verteilt ist zu den Parametern n und p, d.h.

$$P\big[S_n = k\big] \;=\; \binom{n}{k}\, p^k\,(1-p)^{n-k} \qquad \text{für } k \in \{0, \ldots, n\}$$

ergibt sich somit für den Erwartungswert

$$\begin{aligned} E\big[S_n\big] \;&=\; \sum_{k=0}^{n} k \binom{n}{k} p^k\,(1-p)^{n-k} \;=\; n\,p \sum_{k=1}^{n} \binom{n-1}{k-1} p^{k-1}\,(1-p)^{n-k} \\ &=\; n\,p \sum_{k=0}^{n-1} \binom{n-1}{k} p^k\,(1-p)^{n-1-k} \;=\; n\,p, \end{aligned}$$

wobei im letzten Schritt die Normierungsbedingung der Binomialverteilung zu den Parametern $n-1$ und p benutzt wurde. Zur Berechnung der Varianz betrachte zunächst einmal

$$\begin{aligned} E\big[S_n^2\big] \;&=\; \sum_{k=0}^{n} k^2 \binom{n}{k} p^k\,(1-p)^{n-k} \;=\; n\,p \sum_{k=1}^{n} k \binom{n-1}{k-1} p^{k-1}\,(1-p)^{n-k} \\ &=\; n\,p \sum_{k=0}^{n-1} (k+1) \binom{n-1}{k} p^k\,(1-p)^{n-1-k} \\ &=\; n\,p + n\,(n-1)\,p^2 \sum_{k=1}^{n-1} \binom{n-2}{k-1} p^{k-1}\,(1-p)^{n-1-k} \\ &=\; n\,p + n\,(n-1)\,p^2 \sum_{k=0}^{n-2} \binom{n-2}{k} p^k\,(1-p)^{n-2-k} \;=\; n\,p + n\,(n-1)\,p^2, \end{aligned}$$

wobei im vierten bzw. sechsten Schritt die Normierungsbedingung der Binomialverteilung zu den Parametern $n-1$ und p bzw. $n-2$ und p benutzt wurde. Also,

$$\mathrm{Var}(S_n) \;=\; E\big[S_n^2\big] - E\big[S_n\big]^2 \;=\; n\,p + n\,(n-1)\,p^2 - n^2\,p^2 \;=\; n\,p\,(1-p).$$

c) ***Čebyšev-Ungleichung:*** Sei $(\Omega, \mathcal{A}, P)$ ein Wahrscheinlichkeitsraum, X ein Zufallsvariable mit $E\big[|X|^2\big] < \infty$. Dann gilt für alle $c > 0$ *[4 Pkt]*

$$P\big[|X - E[X]| \geq c\big] \;\leq\; \frac{1}{c^2}\,\mathrm{Var}(X).$$

Es soll das schwache Gesetz der großen Zahlen für $\frac{1}{n}S_n$ gezeigt werden, d.h.

$$P\Big[\Big|\frac{S_n}{n} - E\Big[\frac{S_n}{n}\Big]\Big| \geq \varepsilon\Big] \longrightarrow 0, \qquad \text{für } n \to \infty.$$

Aus der Čebyšev-Ungleichung ergibt sich nun für ein beliebiges $\varepsilon > 0$

$$P\Big[\Big|\frac{S_n}{n} - E\Big[\frac{S_n}{n}\Big]\Big| \geq \varepsilon\Big] \;\leq\; \frac{1}{\varepsilon^2}\,\mathrm{Var}\Big(\frac{S_n}{n}\Big) \;=\; \frac{1}{\varepsilon^2\, n^2}\,\mathrm{Var}(S_n) \;=\; \frac{p\,(1-p)}{\varepsilon^2\, n} \;\underset{n\to\infty}{\longrightarrow}\; 0.$$

d) Sei nun $\Omega = \big\{\omega \in \{1,\ldots,m\}^n \,\big|\, \omega_i \neq \omega_j \text{ für } i \neq j\big\}$ mit $|\Omega| = \frac{m!}{n!}$, $\mathcal{A} = \mathcal{P}(\Omega)$ die Potenzmenge und $P[A] = |A|/|\Omega|$ die Gleichverteilung auf $(\Omega, \mathcal{A})$. Weiterhin sei die Menge der gelben Kugeln wieder gegeben durch $\{1,\ldots,g\}$. Betrachte nun die Zufallsvariablen $X_1,\ldots,X_n$ mit [*7+1 Pkt*]

$$X_i : \Omega \;\to\; \{0,1\}, \qquad \omega \;\mapsto\; X_i(\omega) := \mathbb{1}_{\{\omega_i \leq g\}} \quad \text{für alle } 1 \leq i \leq n.$$

Dann ergibt sich für die Wahrscheinlichkeit

$$P\big[X_i = 1\big] \;=\; P\big[\{\omega \in \Omega \mid \omega_i \leq g\}\big] \;=\; \frac{\big|\{\omega \in \Omega \mid \omega_i \leq g\}\big|}{|\Omega|} \;=\; \frac{g\,\frac{(m-1)!}{n!}}{\frac{m!}{n!}} \;=\; \frac{g}{m}.$$

Desweiteren gilt, zusammen mit $p := \frac{g}{m}$, dass

$$P\big[X_1 = 1, X_2 = 1\big] \;=\; \frac{\{\omega \in \Omega \mid \omega_1 \leq g,\, \omega_2 \leq g\}}{|\Omega|} \;=\; \frac{g\,(g-1)\,\frac{(m-2)!}{n!}}{\frac{m!}{n!}} \;=\; p\,\frac{g-1}{m-1}.$$

Betrachte nun die Zufallsvariable $Z_n = \sum_{i=1}^n X_i$, die die Anzahl der gelben unter den n gezogenen Kugeln beschreibt. Es soll nun der Erwartungswert und die Varianz von Z_n berechnet werden.

Lösungsweg 1:

Aus der Linearität des Erwartungswertes ergibt sich nun

$$E\big[Z_n\big] \;=\; E\Big[\sum_{i=1}^n X_i\Big] \;=\; \sum_{i=1}^n E\big[X_i\big] \;=\; \sum_{i=1}^n P\big[X_i = 1\big] \;=\; n\,p,$$

wobei im letzten Schritt zudem benutzt wurde, dass $P[X_i = 1] = p$ für alle $i \leq n$ ist. Zur Berechnung der Varianz betrachte zunächst einmal

$$E\big[Z_n^2\big] \;=\; E\Big[\sum_{i,j=1}^n X_i\, X_j\Big] \;=\; \sum_{i=1}^n E\big[X_i^2\big] + \sum_{i=1}^n \sum_{\substack{j=1\\ j\neq i}}^n E\big[X_i\, X_j\big],$$

wobei im zweiten Schritt die Linearität des Erwartungswertes benutzt wurde. Zusammen mit

$$E\big[X_i^2\big] \;=\; P\big[X_i = 1\big] \;=\; p, \qquad E\big[X_i\, X_j\big] \;=\; P\big[X_i = 1, X_j = 1\big] \;=\; p\,\frac{g-1}{m-1}$$

ergibt sich dann zunächste einmal

$$E\big[Z_n^2\big] \;=\; \sum_{i=1}^{n} P\big[X_i = 1\big] + \sum_{i=1}^{n} \sum_{\substack{j=1 \\ j\neq i}}^{n} P\big[X_i = 1, X_j = 1\big] \;=\; n\,p + n\,(n-1)\,p\,\frac{g-1}{m-1}$$

Lösungsweg 2:
Diese Zufallsvariable ist nun aber hypergeometrisch verteilt zu den Parametern m, g und n, d.h.

$$P\big[Z_n = k\big] \;=\; \frac{\binom{g}{k}\binom{m-g}{n-k}}{\binom{m}{n}} \qquad \text{für } k \in \{0, \ldots, n\}.$$

Für den Erwartungswert von Z_n ergibt sich dann

$$\begin{aligned} E\big[Z_n\big] \;&=\; \sum_{k=0}^{n} k\,\frac{\binom{g}{k}\binom{m-g}{n-k}}{\binom{m}{n}} \;=\; \frac{n\,g}{m} \sum_{k=1}^{n} \frac{\binom{g-1}{k-1}\binom{m-1-(g-1)}{n-1-(k-1)}}{\binom{m-1}{n-1}} \\ &=\; n\,p \sum_{k=0}^{n-1} \frac{\binom{g-1}{k}\binom{m-1-(g-1)}{n-1-k}}{\binom{m-1}{n-1}} \;=\; n\,p, \end{aligned}$$

wobei im letzten Schritt die Normierungsbedingung der hypergeometrischen Verteilung zu den Parametern $m-1$, $g-1$, $n-1$ benutzt wurde.
Zur Berechnung der Varianz der Zufallsvariable Z_n betrachte zunächst einmal

$$\begin{aligned} E\big[Z_n^2\big] \;&=\; \sum_{k=0}^{n} k^2\,\frac{\binom{g}{k}\binom{m-g}{n-k}}{\binom{m}{n}} \;=\; \frac{n\,g}{m} \sum_{k=1}^{n} k\,\frac{\binom{g-1}{k-1}\binom{m-1-(g-1)}{n-1-(k-1)}}{\binom{m-1}{n-1}} \\ &=\; n\,p \sum_{k=0}^{n-1} (k+1) \frac{\binom{g-1}{k}\binom{m-1-(g-1)}{n-1-k}}{\binom{m-1}{n-1}} \;=\; n\,p + n\,p \sum_{k=1}^{n-1} k\,\frac{\binom{g-1}{k}\binom{m-1-(g-1)}{n-1-k}}{\binom{m-1}{n-1}} \\ &=\; n\,p + n\,(n-1)\,p\,\frac{g-1}{m-1} \sum_{k=1}^{n-1} \frac{\binom{g-2}{k-1}\binom{m-2-(g-2)}{n-2-(k-1)}}{\binom{m-2}{n-2}} \\ &=\; n\,p + n\,(n-1)\,p\,\frac{g-1}{m-1} \sum_{k=0}^{n-2} \frac{\binom{g-2}{k}\binom{m-2-(g-2)}{n-2-k}}{\binom{m-2}{n-2}} \;=\; n\,p + n\,(n-1)\,p\,\frac{g-1}{m-1}. \end{aligned}$$

Abschließend ergibt sich dann aber für die Varianz

$$\begin{aligned} \mathrm{Var}(Z_n) \;&=\; E\big[Z_n^2\big] - E\big[Z_n\big]^2 \;=\; n\,p\left(1 + (n-1)\,\frac{g-1}{m-1} - n\,\frac{g}{m}\right) \\ &=\; n\,p\,\frac{m^2 - m\,n - m\,g + n\,g}{m\,(m-1)} \;=\; n\,p\,\frac{m-g}{m}\,\frac{m-n}{m-1} \;=\; \frac{m-n}{m-1}\,n\,p\,(1-p) \\ &=\; \frac{m-n}{m-1}\,\mathrm{Var}(S_n). \end{aligned}$$

2. (Bedingte Wahrscheinlichkeiten)

a) Unter einem Wahrscheinlichkeitsraum versteht man ein Tripel $(\Omega, \mathcal{A}, P)$ bestehend aus einer Menge $\Omega \neq \emptyset$, einer σ-Algebra $\mathcal{A}$ und einer Wahrscheinlichkeitsverteilung P auf Ω. *[5 Pkt]*

Eine σ-Algebra ist hierbei eine Teilmenge der Potenzmenge $\mathcal{A} \subseteq \mathcal{P}(\Omega)$, die den folgenden Eigenschaften genügt

(i) $\Omega \in \mathcal{A}$

(ii) $A \in \mathcal{A} \quad \Rightarrow \quad A^c := \Omega \setminus A \in \mathcal{A}$

(iii) $A_1, A_2, \ldots \in \mathcal{A} \quad \Rightarrow \quad \bigcup_{i=1}^{\infty} A_i \in \mathcal{A}$.

Unter einer Wahrscheinlichkeitsverteilung P auf $(\Omega, \mathcal{A})$ versteht man eine Abbildung

$$P : \mathcal{A} \to [0,1], \qquad \mathcal{A} \ni A \mapsto P[A] \in [0,1]$$

mit den Eigenschaften, dass

(i) Normierung: $P[\Omega] = 1$

(ii) σ-Additivität: Für paarweise disjunkte Ereignisse $A_1, A_2, \ldots \in \mathcal{A}$, d.h. es sei $A_i \cap A_j = \emptyset$ für alle $i \neq j$, gilt

$$P\Big[\bigcup_{i=1}^{\infty} A_i\Big] = \sum_{i=1}^{\infty} P[A_i].$$

Es nun sei $(\Omega, \mathcal{A}, P)$ ein Wahrscheinlichkeitsraum und $A, B \in \mathcal{A}$ zwei Ereignisse, wobei $P[B] > 0$ ist. Dann ist die bedingte Wahrscheinlichkeit von A gegeben B definiert durch

$$P[A \mid B] := \frac{P[A \cap B]}{P[B]}.$$

b) Es sei $(\Omega, \mathcal{A}, P)$ ein Wahrscheinlichkeitsraum und $A, B \in \mathcal{A}$ zwei Ereignisse mit $P[B] \in (0,1)$. Dann gilt *[4 Pkt]*

$$\begin{aligned} P[A] &= P[A \cap \Omega] = P[A \cap (B \cup B^c)] = P[(A \cap B) \cup (A \cap B^c)] \\ &= P[A \cap B] + P[A \cap B^c] = P[B] \cdot \frac{P[A \cap B]}{P[B]} + P[B^c] \cdot \frac{P[A \cap B^c]}{P[B^c]} \\ &= P[B] \cdot P[A \mid B] + P[B^c] \cdot P[A \mid B^c], \end{aligned}$$

wobei im vierten Schritt die σ-Additivität und im sechsten Schritt die Definition der bedingten Wahrscheinlichkeit benutzt wurde, da $P[B] \neq 0$ und mit $P[B] \neq 1$ auch $P[B^c] = 1 - P[B] \neq 0$ ist. Betrachte nun den Fall, dass $P[B] = 0$ ist. Aus $0 \leq P[A \cap B] \leq P[B] = 0$ und $P[B^c] = 1 - P[B] = 1$ ergibt sich dann

$$P[A] = P[A \cap B^c] = P[B^c] \cdot P[A \mid B^c] = P[A \mid B^c],$$

d.h. das Ereignis A ist unabhängig vom Ereignis B^c. Analog ergibt sich im Fall $P[B] = 1$, dass

$$P[A] = P[A \cap B] = P[B] \cdot P[A \mid B] = P[A \mid B].$$

c) Es sei $(\Omega, \mathcal{A}, P)$ ein Wahrscheinlichkeitsraum und $X_0, X_1, X_2, \ldots$ ein Folge von Zufallsvariablen mit Werten in $\{0, 1\}$. Weiter sei $Z_1, Z_2, \ldots$ eine Folge von Zufallsvariablen mit $P[Z_n = 1] = P[Z_n = 0] = \frac{1}{2}$ für alle $n \in \mathbb{N}$, wobei Z_n unabhängig ist von den Zufallsvariablen $X_0, \ldots, X_n$ ist.

(i) Zu zeigen: $p_n = \frac{1}{2} - \frac{1}{2}\, p_{n-1}$ für alle $n \geq 1$, wobei $p_n := P[X_n = 1]$ ist. *[3 Pkt]*
Es gilt nun aber

$$\begin{aligned} p_n &= P\big[X_n = 1\big] \;=\; P\big[(X_n = 1, X_{n-1} = 1) \vee (X_n = 1, X_{n-1} = 0)\big] \\ &= P\big[X_n = 1, X_{n-1} = 1\big] + P\big[X_n = 1, X_{n-1} = 0\big] \\ &= 0 + P\big[Z_n = 1, X_{n-1} = 0\big] \;=\; P\big[Z_n = 1\big] \cdot P\big[X_{n-1} = 0\big] \\ &= (1 - p_{n-1}) \cdot P\big[Z_n = 1\big] \;=\; \frac{1}{2} - \frac{1}{2}\, p_{n-1} \end{aligned}$$

wobei im dritten Schritt die σ-Additivität, im vierten Schritt die Definition der Verteilung der Zufallsvariablen X_i und im fünften Schritt die Unabhängigkeit von Z_n bezüglich X_{n-1} benutzt wurde.

(ii) Zunächst einmal stellt die Abbildung $\phi(x) = \frac{1}{2} - \frac{1}{2}\, x$ eine Kontraktion dar, denn *[2+1 Pkt]*

$$|\phi(x) - \phi(y)| \;\leq\; \frac{1}{2}\, |x - y|.$$

Andererseits ist $p_n = \phi(p_{n-1})$. Aus dem Banachschen Fixpunktionsatz folgt nun, dass ϕ einen eindeutigen Fixpunkt besitzt und dass die Folge p_n gegen diesen konvergiert. Es gilt nun noch den Fixpunkt p zu bestimmen. Hierbei gilt aber

$$p \;=\; \phi(p) \;=\; \frac{1}{2} - \frac{1}{2}\, p \qquad \Longleftrightarrow \qquad p \;=\; \frac{1}{3}.$$

(iii) Zu zeigen: $P[X_1 = 1, X_3 = 1] = \frac{1}{4}$. *[2 Pkt]*
Es gilt nun aber, dass

$$\begin{aligned} &P\big[X_1 = 1, X_3 = 1\big] \\ &\quad = P\big[X_1 = 1, X_2 = 0, X_3 = 1\big] + P\big[X_1 = 1, X_2 = 1, X_3 = 1\big] \\ &\quad = P\big[X_1 = 1, X_2 = 0, X_3 = 1\big] \\ &\quad = P\big[X_1 = 1\big] \cdot P\big[X_2 = 0 \,|\, X_1 = 1\big] \cdot P\big[X_3 = 1 \,|\, X_1 = 1, X_2 = 0\big] \\ &\quad = P\big[Z_1 = 1\big] \cdot 1 \cdot P\big[Z_3 = 1 \,\big|\, X_1 = 1, X_2 = 0\big] \\ &\quad = P\big[Z_1 = 1\big] \cdot P\big[Z_3 = 1\big] \;=\; \frac{1}{4}, \end{aligned}$$

wobei im fünften Schritt benutzt wurde, dass Z_3 unabhängig von X_1, X_2 ist.

(iv) Aus der Definition der bedingten Wahrscheinlichkeit ergibt sich *[2 Pkt]*

$$\frac{1}{4} \;=\; P\big[X_1 = 1, X_3 = 1\big] \;=\; p_3 \cdot P\big[X_1 = 1 \,\big|\, X_3 = 1\big] \;=\; \frac{3}{8}\, P\big[X_1 = 1 \,\big|\, X_3 = 1\big],$$

wobei im dritten Schritt benutzt wurde, dass $p_3 = \frac{1}{2}\left(1 - \frac{1}{2} + \frac{1}{4}\right) = \frac{3}{8}$ ist. Dies ist aber äquivalent zu $P[X_1 = 1 \mid X_3 = 1] = \frac{2}{3}$.

(v) Da die Zufallsvariablen X_k nur Werte in $\{0, 1\}$ annehmnen, ergibt sich, zusammen mit $p_1 = \frac{1}{2}$ und $p_3 = \frac{3}{8}$, für die Kovarianz zwischen X_1 und X_3 [*2 Pkt*]

$$\begin{aligned} \mathrm{Cov}\big(X_1, X_3\big) &= E\big[X_1\, X_3\big] - E\big[X_1\big]\, E\big[X_3\big] \\ &= P\big[X_1 = 1, X_3 = 1\big] - P\big[X_1 = 1\big] \cdot P\big[X_3 = 1\big] \;=\; \frac{1}{16}. \end{aligned}$$

3. (Gleichgewichte von Markovketten)

a) Eine Folge $X_0, X_1, X_2, \dots$ von Zufallsvariablen auf einem Wahrscheinlichkeitsraum $(\Omega, \mathcal{A}, P)$ mit Werten in S heißt (zeithomogene) Markovkette mit Zustandsraum S und Übergangsmatrix p, wenn für alle $n \geq 0$ und alle $x_0, \dots, x_n \in S$ gilt [*6 Pkt*]

$$P[X_n = x_n \mid X_0 = x_0, \dots, X_{n-1} = x_{n-1}] \;=\; p(x_{n-1}, x_n)$$

sofern $P[X_0 = x_0, \dots, X_{n-1} = x_{n-1}] > 0$ ist. Hierbei ist p eine stochastische Matrix, d.h.

$$p(x, y) \;\geq\; 0 \quad \forall\, x, y \in S \qquad \text{und} \qquad \sum_{y \in S} p(x, y) = 1 \quad \forall x \in S.$$

Eine Wahrscheinlichkeitsverteilung μ auf S heißt Gleichgewichtsverteilung bezüglich der Übergangsmatrix p, falls

$$\mu(x) \;=\; \sum_{y \in S} \mu(y)\, p(y, x), \qquad \text{für alle } x \in S.$$

b) Es sei S eine endliche Menge, μ eine Wahrscheinlichkeitsverteilung auf S mit positiven Gewichten $\mu(x) > 0$ und $q(x, y)$ für $x, y \in S$ eine symmetrische, stochastische Matrix. [*6 Pkt*]

Der Metropolis-Algorithmus zur Simulation der Verteilung μ ist gegeben durch

Input: Zustand $x \in S$, symmetrische, stochastische Matrix q und Wahrscheinlichkeitsverteilung μ

Output: Zustand $y \in S$ bezüglich des Übergangs $x \to y$ einer Markovkette mit Gleichgewichtsverteilung μ

Erzeuge $z \in S$ bezüglich der Verteilung $q_x \equiv q(x, \cdot)$.

Erzeuge $u \in [0, 1]$ bezüglich der uniformen Verteilung auf $[0, 1]$.

$\alpha(x, z) := \min\left(1, \frac{\mu(z)}{\mu(x)}\right)$

if $u \leq \alpha(x, z)$ **then** $y = z$ **else** $y = x$ **end if**

Die Übergangsmatrix p der entsprechenden Metropoliskette ist für $x, y \in S$ mit $y \neq x$ gegeben durch

$$p(x, y) \;=\; q(x, y)\, \min\left(1, \frac{\mu(y)}{\mu(x)}\right), \qquad p(x, x) \;=\; 1 - \sum_{\substack{y \in S \\ y \neq x}} p(x, y).$$

c) Der Zustandsraum der Markovkette (X_n) ist $\mathbb{N}_0$. Weiterhin sind die von Null verschiedenen Matrixelemente der Übergangsmatrix p gegeben durch [8+1 Pkt]

$$\begin{aligned} p(0,1) &= 1, \\ p(k,0) &= p, \qquad \text{für alle } k>0, \\ p(k,k+1) &= 1-p, \quad \text{für alle } k>0. \end{aligned}$$

Lösungsweg 1:
Sei nun μ eine Gleichgewichtsverteilung. Dann gilt

$$\mu(1) \;=\; \sum_{n=0}^{\infty} \mu(n)\, p(n,1) \;=\; \mu(0)\, p(0,1) \;=\; \mu(0).$$

Andererseits ergibt sich für $k>1$, dass

$$\mu(k) \;=\; \sum_{n=0}^{\infty} \mu(n)\, p(n,k) \;=\; \mu(k-1)\, p(k-1,k) \;=\; \mu(k-1)\,(1-p).$$

Durch vollständige Induktion folgt daraus, dass $p(k) = \mu(0)\,(1-p)^{k-1}$ für $k \geq 1$ ist. Zusammen mit der Normierungsbedingung

$$1 = \sum_{k=0}^{\infty} \mu(k) = \mu(0) + \sum_{k=1}^{\infty} \mu(0)\,(1-p)^{k-1} = \mu(0)\left(1 + \sum_{k=0}^{\infty} (1-p)^k\right) = \mu(0)\,\frac{p+1}{p}$$

erhält man dann schließlich, dass jede Gleichgewichtsverteilung den folgenden beiden Gleichungen genügen muss:

$$\mu(0) \;=\; \frac{p}{p+1}, \qquad \mu(k) \;=\; \frac{p}{p+1}\,(1-p)^{k-1}, \quad \text{für } k \geq 1.$$

Also gibt es höchstens eine Gleichgewichtsverteilung. Andererseits läßt sich leicht nachrechnen, dass eine Verteilung μ, in der oben angegebenen Form, der Gleichung $\mu = \mu\, p$ genügt, d.h. μ ist tatsächlich eine Gleichgewichtsverteilung.

Lösungsweg 2:
Es gilt zu überprüfen, dass die gegebene Markovkette irreduzibel und aperiodisch ist.

Seien also $x, y \in \mathbb{N}_0$. Dann gilt im Falle $x < y$, dass $p^k(x,y) \geq (1-p)^k > 0$ für $k = y - x$, während im Falle $x \geq y$ gilt, dass $p^k(x,y) \geq p\,(1-p)^{k-2} > 0$ für $k = y+2$, d.h. die Markovkette ist irreduzibel.

Desweiteren ist $k, k+1 \in \{n \in \mathbb{N} \mid p^n(x,x) > 0\}$ für $k = x+2$, d.h. die Periode $d(x) = \mathrm{ggT}\{n \in \mathbb{N} \mid p^n(x,x) > 0\} = 1$. Da wegen der Irreduzibilität $d(x) = d(y)$ für alle $x, y \in \mathbb{N}_0$ ist, folgt die Aperiodizität der Markovkette.

Es nun $\nu \neq \mu$ eine weitere Gleichgewichtsverteilung. Dann ergibt sich aus dem Konvergenzsatz für Markovketten

$$d_{\mathrm{TV}}(\nu,\mu) \;=\; d_{\mathrm{TV}}(\nu\, p^n, \mu) \underset{n\to\infty}{\longrightarrow} 0 \qquad \Rightarrow \qquad \nu = \mu$$

ein Widerspruch, d.h. die Gleichgewichtsverteilung ist eindeutig.

Die Gleichgewichtsverteilung läßt sich nun durch eine analoge Rechnung wie im Lösungsweg 1 bestimmen, d.h.

$$\mu(0) \;=\; \frac{p}{p+1}, \qquad \mu(k) \;=\; \frac{p}{p+1}\,(1-p)^{k-1}, \quad \text{für } k \geq 1.$$

4. (Newton-Verfahren)

a) Der Algorithmus des Newton-Verfahrens ist gegeben durch [5 *Pkt*]

Input: Funktion $f : \mathbb{R} \to \mathbb{R}$, Startwert $x^{(0)} \in \mathbb{R}$, Fehlerschranke ε, Intervall $I = (-M, M)$, $M \in \mathbb{R}_+$.

Output: Nullstelle x von f oder Fehlermeldung

$tol = f(x^{(0)})$, $x = x^{(0)}$, $err = 0$

while $(err = 0\ \&\ tol > \varepsilon)$ **do**

  if $f'(x) = 0$ **then**

    $err = 1$, **print** "STOP: Ableitung verschwindet!"

  else

    $x = x - \dfrac{f(x)}{f'(x)}$, $tol = f(x)$

    if $x \notin I$ **then**

      $err = 1$, **print** "STOP: Intervall verlassen!"

    end if

  end if

end while

b) Die quadratische Konvergenz des Newtonverfahrens bedeutet, dass die Folge gegeben durch $x^{(k+1)} = \phi(x^{(k)})$ mit $\phi(x) = x - \frac{f(x)}{f'(x)}$ gegen x^*, eine Nullstelle von f, konvergiert und es gilt: Es existiert ein $c \in (0, \infty)$ derart, dass [5 *Pkt*]

$$|x^{(k+1)} - x^*| \ \le\ c\,|x^{(k)} - x^*|^2.$$

Betrachte die Folge $(a_k)_{k\in\mathbb{N}}$, die gegeben ist durch $a_k = \exp(-2^k)$. Offentsichtlich ist a_k eine Nullfolge nicht-negativer Zahlen. Weiterhin gilt

$$a_{k+1} \ =\ \frac{a_{k+1}}{a_k^p}\, a_k^p \ =\ \mathrm{e}^{-2^k(2-p)}\, a_k^p \ \le\ c\, a_k^p$$

d.h. wobei $c = \sup_k \mathrm{e}^{-2^k(2-p)}$ ist. Aber für $p \le 2$ ist $c = \mathrm{e}^{-(2-p)} \le 1$, während für $p > 2$ die Konstante $c = \infty$ ist, d.h. die maximale Konvergenzordnung ist 2.

c) Sei $x^* \in I$ mit $f(x^*) = 0$ und $f'(x) > 0$ für alle $x \in I$. [6 *Pkt*]

Zu zeigen: Für alle $k \ge 0$ mit $x^{(k)} \in I$ gilt

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

Aus dem Hauptsatz der Integral- und Differentialrechnung ergibt sich zunächst einmal

$$-f(x^{(k)}) \ =\ f(x^*) - f(x^{(k)}) \ =\ \int_{x^{(k)}}^{x^*} f'(y)\, dy.$$

Daraus folgt dann

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

Sei nun $c := \min_{x\in I}|f'(x)|$. Da die Funktion f stetig und $f'(x) > 0$ für alle $x \in I$ ist, folgt somit $c > 0$. Weiterhin gilt nach Voraussetzung, dass $|f'(y) - f'(x)| \le L\,|y - x|$ für alle $x, y \in I$. Dann ergibt sich aus der obigen Rechnung

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

Betrachte nun eine Umgebung $U_\varepsilon(x^*)$ von x^* mit $\varepsilon \in (0, \frac{2\,c}{L})$. Es gilt nun zu zeigen, dass $x^{(0)} \in U_\varepsilon(x^*)$ impliziert, dass $x^{(k)} \in U_\varepsilon(x^*)$ für alle k.

$\boxed{\text{IA}}$ $k = 0:$ trivial.

$\boxed{\text{IS}}$ $k \longrightarrow k+1$

Es gilt nun aber,

$$\big|x^{(k+1)} - x^*\big| \le \frac{L}{2\,c}\big|x^{(k)} - x^*\big|^2 \le \frac{L}{2\,c}\,\varepsilon\,\big|x^{(k)} - x^*\big| < \big|x^{(k)} - x^*\big|,$$

wobei im letzten Schritt die Definition von ε benutzt wurde. Zusammen mit der Induktionsvoraussetzung ergibt sich daher, dass $x^{(k+1)} \in U_\varepsilon(x^*)$ ist.

Da das Verfahren aber innerhalb dieser Umgebung kontraktiv ist, folgt schließlich die quadratische Konvergenz des Newton-Verfahrens.

d) Betrachte die Funktion $f(x) = x\,\mathrm{e}^{-x}$, wobei $f(x) = 0$ genau dann wenn $x = 0$ ist. Es sei nun angenommen, dass $x^{(k)} > 1$ ist. Dies impliziert dann [4+1 Pkt]

$$x^{(k+1)} = x^{(k)} - \frac{x^{(k)}}{1 - x^{(k)}} > x^{(k)} \qquad \text{für alle } k,$$

d.h. die Folge $x^{(k)}$ ist streng monoton wachsend und konvergiert daher für $x^{(0)} > 1$ bestimmt gegen $+\infty$.

5. (Abstiegsverfahren)

a) Sei $A \in \mathbb{R}^{d\times d}$ eine symmetrische, positiv definite Matrix, $b \in \mathbb{R}^d$ und [2 Pkt]

$$\phi(x) = \frac{1}{2}\big\langle x, A\,x\big\rangle - \big\langle x, b\big\rangle.$$

Zu zeigen: $\phi(x) = \phi(x^*) + \frac{1}{2}\left\langle x - x^*, A\,(x - x^*)\right\rangle$ für $x^* = A^{-1}b$.

Es gilt nun aber

$$\begin{aligned}
\phi(x) &= \phi(x^*) + \phi(x) - \phi(x^*) \\
&= \phi(x^*) + \tfrac{1}{2}\langle x, A\,x\rangle - \langle x, b\rangle - \tfrac{1}{2}\langle x^*, A\,x^*\rangle + \langle x^*, b\rangle \\
&= \phi(x^*) + \tfrac{1}{2}\langle x, A\,x\rangle - \langle x, A\,x^*\rangle - \tfrac{1}{2}\langle x^*, A\,x^*\rangle + \langle x^*, A\,x^*\rangle \\
&= \phi(x^*) + \tfrac{1}{2}\Big(\langle x, A\,x\rangle - \langle x, A\,x^*\rangle - \langle x^*, A\,x\rangle + \langle x^*, A\,x^*\rangle\Big) \\
&= \phi(x^*) + \tfrac{1}{2}\left\langle x - x^*, A\,(x - x^*)\right\rangle,
\end{aligned}$$

wobei im zweiten Schritt benutzt wurde, dass $b = A\,x^*$ ist, während im dritten Schritt ausgenutzt wurde, dass die Matrix A symmetrisch ist.

b) Betrachte nun zur näherungsweisen Berechnung des Minimums x^* von ϕ ein Abstiegsverfahren vom Typ $x^{(k+1)} = x^{(k)} + \alpha_k\, p^{(k)}$. Für das Gradientenverfahren ist dann die Suchrichtung gegeben durch [*8 Pkt*]

$$p^{(k)} \;=\; -\big(\nabla\phi\big)\big(x^{(k)}\big) \;=\; -A\,(x^{(k)} - x^*) \;=\; r^{(k)}$$

Es gilt nun die optimale Schrittweite α_k zu bestimmen.

Lösungsweg 1:

Zusammen mit Aufgabenteil a) ergibt sich, dass

$$\phi(x^{(k+1)}) \;=\; \phi(x^*) + \tfrac{1}{2}\,\|x^{(k+1)} - x^*\|_A^2$$

genau dann minimal ist, wenn $x^{(k+1)}$ gerade die A-orthogonale Projektion von $x^* - x^{(k)}$ auf $x^{(k)} + \operatorname{span}\{r^{(k)}\}$ ist, d.h.

$$x^{(k+1)} \;=\; x^{(k)} + \Big\langle A\,(x^* - x^{(k)}), \frac{r^{(k)}}{\|r^{(k)}\|_A}\Big\rangle \frac{r^{(k)}}{\|r^{(k)}\|_A} \;=\; x^{(k)} + \alpha_k\, r^{(k)}.$$

Daraus folgt dann aber schließlich

$$\alpha_k \;=\; \frac{\big\langle A\,(x^* - x^{(k)}), r^{(k)}\big\rangle}{\|r^{(k)}\|_A^2} \;=\; \frac{\big\langle b - A\,x^{(k)}, r^{(k)}\big\rangle}{\big\langle r^{(k)}, A\,r^{(k)}\big\rangle} \;=\; \frac{\|r^{(k)}\|^2}{\|r^{(k)}\|_A^2}.$$

Lösungsweg 2:

Aus der Darstellung in Aufgabenteil a) folgt zunächst einmal, dass $\phi(y) > \phi(x^*)$ für alle $y \neq x^*$. Weiterhin gilt

$$0 \;=\; \frac{\partial}{\partial\alpha_k}\,\phi\big(x^{(k)} + \alpha_k\, r^{(k)}\big) \;=\; \big\langle r^{(k)}, A\,(x^{(k)} - x^*)\big\rangle + \alpha_k\,\big\langle r^{(k)}, A\,r^{(k)}\big\rangle,$$

wobei im zweiten Schritt wiederum benutzt wurde, dass die Matrix A symmetrisch ist. Zusammen mit $r^{(k)} = b - A\,x^{(k)}$ ergibt sich daher dann

$$\alpha_k \;=\; -\frac{\big\langle A\,(x^{(k)} - x^*), r^{(k)}\big\rangle}{\big\langle r^{(k)}, A\,r^{(k)}\big\rangle} \;=\; \frac{\big\langle b - A\,x^{(k)}, r^{(k)}\big\rangle}{\big\langle r^{(k)}, A\,r^{(k)}\big\rangle} \;=\; \frac{\|r^{(k)}\|^2}{\|r^{(k)}\|_A^2}.$$

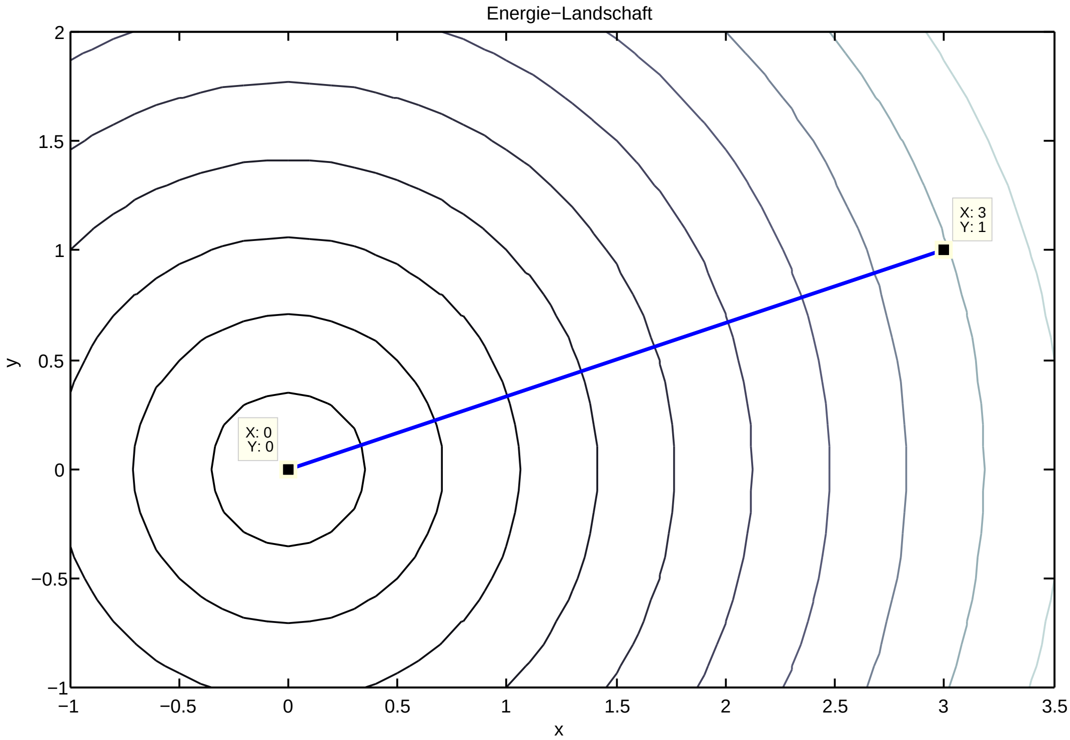

Abbildung 1: Höhenlinen der Funktion ϕ im Fall (i)

Für das Residium gilt weiterhin, dass

$$r^{(k+1)} \;=\; b - A\left(x^{(k)} + \alpha_k\, r^{(k)}\right) \;=\; r^{(k)} - \alpha_k\, A\, r^{(k)}.$$

Der Algorithmus des Gradienten-Verfahrens ist gegeben durch

Input: symmetrisch, positiv definite Matrix $A \in \mathbb{R}^{d\times d}$, $b \in \mathbb{R}^d$
Startwert $x^{(0)} \in \mathbb{R}^d$

Output: Iterationsfolge $x^{(k)} \in \mathbb{R}^d$ der Lösung x von $A\,x = b$
zugehörigen Residuen $r^{(k)} \in \mathbb{R}^d$.

Setze $r^{(0)} = b - A\,x^{(0)}$

for $k = 0, 1, 2, \ldots$ **do**

$\alpha_k = \frac{\|r^{(k)}\|^2}{\|r^{(k)}\|_A^2}$

$x^{(k+1)} = x^{(k)} + \alpha_k\, r^{(k)}$

$r^{(k+1)} = r^{(k)} - \alpha_k\, A\, r^{(k)}$

end for

c) Die Darstellung der Höhenlinen der Funktion ϕ am Fall (i) bzw. (ii), sowie die Iterationsschritte des Gradientenverfahrens sind in den Abbildungen 1 bzw. 2 dargestellt. Beachte, dass die Abstiegsrichtung stets senkrecht auf den Höhenlinien steht. *[6 Pkt]*

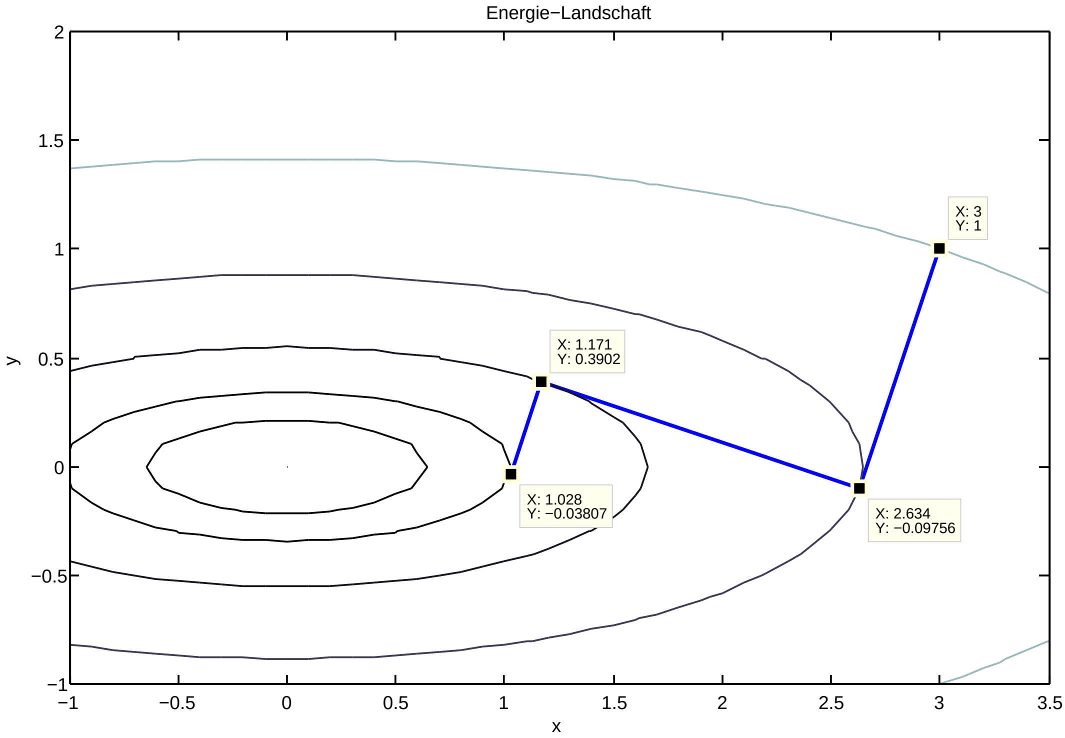

Abbildung 2: Höhenlinen der Funktion ϕ im Fall (ii)

d) Es gilt zunächst einmal, dass $\|A\|_2 = \sqrt{\rho\big(A\,A^T\big)}$, wobei $\rho(A)$ der Spektralradius der Matrix A ist, d.h. $\rho(A) = \max\big\{|\lambda| \;\big|\; \lambda \text{ Eigenwert von } A\big\}$. Ist die Matrix A symmetrisch und positiv definit, so ist $\|A\|_2 = \lambda_{\max}$ bzw. $\|A^{-1}\|_2 = 1/\lambda_{\min}$ der größte bzw. das Inverse des kleinsten Eigenwertes der Matrix A. *[4+1 Pkt]*

Für die Matrix im Fall (i) ergibt sich

$$A\,A^T = A^{-1}A^{-T} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad \Longrightarrow \quad \|A\|_2 = \|A^{-1}\|_2 = 1$$

Daraus folgt dann, dass $K_2(A) = 1$ ist und somit

$$\|x^{(k+1)} - x^*\|_A \leq \frac{1-1}{1+1}\,\|x^{(k)} - x^*\|_A = 0,$$

d.h. das Gradientenverfahren konvergiert in einem Schritt. Dies liegt daran, dass in diesem Fall die Höhenlinien $\phi(x) = c$ konzentrische Kreise bilden, weshalb dessen Mittelpunkt stets auf der Geraden liegt, die durch die Abstiegsrichtung im Gradienten-Verfahren gegeben ist.

Betrachte nun die Matrix aus Fall (ii). Dann ergibt sich

$$A\,A^T = \begin{pmatrix} 1 & 0 \\ 0 & 81 \end{pmatrix} \Longrightarrow \|A\|_2 = 9 \quad \text{und} \quad A^{-1}A^{-T} = \begin{pmatrix} 1 & 0 \\ 0 & \frac{1}{81} \end{pmatrix} \Longrightarrow \|A^{-1}\|_2 = 1$$

Daraus folgt dann, dass $K_2(A) = 9$ und somit

$$\|x^{(k+1)} - x^*\|_A \;\leq\; \frac{9-1}{9+1}\,\|x^{(k)} - x^*\|_A \;=\; \frac{4}{5}\,\|x^{(k)} - x^*\|_A,$$

d.h. das Gradientenverfahren konvergiert langsam, wenn das Verhältnis aus größten zu kleinsten Eigenwert $\lambda_{\max}/\lambda_{\min} \gg 1$ ist. In diesem Fall bilden die Höhenlinien zu $\phi(x) = c$ konzentrische Ellipsen mit Halbachsen proportional zu $\lambda_{\max}$ bzw. $\lambda_{\min}$. Da die Abstiegsrichtung im Gradienten-Verfahren senkrecht auf den Höhenlinien steht, führt dies zu einem "zig-zag" Verlauf der zugehörigen Abstiegskurve.

6. (Interpolation und Quadratur)
Sei $m \in \mathbb{N}$ und $f \in C^{m+1}([a,b])$ für $-\infty < a < b < \infty$.

a) Es bezeichne $\Pi_m = \{x \mapsto \sum_{i=0}^m a_i\, x^i \mid a_0, \dots, a_m \in \mathbb{R}\}$ den Raum der Polynome vom Grad $\leq m$. Betrachte nun die Abbildung $F : \Pi_m \to \mathbb{R}^{m+1}$, *[5 Pkt]*

$$p \;\mapsto F(p) := \big(p(x_0), \dots, p(x_m)\big)$$

Zu zeigen: F ist injektiv, d.h. $F(p) = 0 \;\Rightarrow\; p = 0$.

Es folgt nun aber aus $F(p) = 0$, dass das Polynom p mindestens die $m+1$ Nullstellen $x_0, \dots, x_m$ besitzt. Da aber $p \in \Pi_m$ ist, d.h. höchsten vom Grad m, impliziert der Hauptsatz der Algebra, dass $p \equiv 0$ ist.

Da weiter $\dim(\Pi_m) = m+1 = \dim(\mathbb{R}^{m+1})$, ist die Abbildung F folglich bijektiv, d.h. insbesondere zu $\big(f(x_0), \dots, f(x_m)\big) \in \mathbb{R}^{m+1}$ existiert genau ein Polynom $p \in \Pi_m$ mit $p(x_i) = f(x_i)$ für alle $0 \leq i \leq m$.

b) Die Lagrange-Interpolationspolynome sind gegeben durch *[6 Pkt]*

$$L_i(x) \;=\; \prod_{\substack{k=0\\k\neq i}}^{m} \frac{x - x_k}{x_i - x_k} \in \Pi_m.$$

Diese stellen die eindeutige Lösung des Interpolationsproblems $L_i(x_j) = \delta_{i,j}$ dar, denn für beliebige i, j mit $i \neq j$ gilt

$$L_i(x_j) \;=\; \prod_{\substack{k=0\\k\neq i}}^{m} \frac{x_j - x_k}{x_i - x_k} \;=\; 0 \cdot \prod_{\substack{k=0\\k\neq i,j}}^{m} \frac{x_j - x_k}{x_i - x_k} \;=\; 0, \qquad L_i(x_i) \;=\; \prod_{\substack{k=0\\k\neq i}}^{m} \frac{x_i - x_k}{x_i - x_k} \;=\; 1.$$

Zusammen mit Aufgabenteil a) folgt daraus die Eindeutigkeit der Darstellung.

Damit läßt sich dann das Polynom p, dass das Interpolationsproblem $(x_i, f(x_i))$ löst, eindeutig darstellen durch $p(x) = \sum_{i=0}^m f(x_i)\, L_i(x)$, denn es gilt

$$p(x_j) \;=\; \sum_{i=0}^{m} f(x_i)\, L_i(x_j) \;=\; \sum_{i=0}^{m} f(x_i)\, \delta_{i,j} \;=\; f(x_j),$$

wobei sich die Eindeutigkeit wiederum aus Aufgabenteil a) ergibt.

c) Die abgeschlossene Newton-Cotes-Formel mit $m+1$ Stützstellen ist gegeben durch [*5 Pkt*]

$$Q_{[a,b]}[f] \;=\; \sum_{i=0}^{m} w_i\, f(x_i) \qquad \text{mit} \qquad w_i = \int_a^b L_i(x)\,dx$$

und $x_i = a + i\,\frac{b-a}{m}$ für $i = 0, \ldots, m$.

Zu zeigen: Der Exaktheitsgrad der abgeschlossenen Newton-Cotes-Formel mit $m+1$ Stützstellen ist $\geq m$.

Es sei $p \in \Pi_m$. Da die Lagrange-Polynome aber eine Basis von Π_m bilden, gilt, dass $p(x) = \sum_{i=0}^{m} p(x_i)\, L_i(x)$ ist. Daraus folgt dann aber

$$Q_{[a,b]}[p] \;=\; \sum_{i=0}^{m} w_i\, p(x_i) \;=\; \int_a^b \sum_{i=0}^{m} p(x_i)\, L_i(x)\, dx \;=\; \int_a^b p(x)\, dx,$$

d.h. die abgeschlossenen Newton-Cotes-Formel mit $m+1$ Stützstellen besitzen midenstens den Exaktheitsgrad m.

d) Es sollen nun die Gewichte der Newton-Cotes-Formel für $m = 2$ berechnet werden. Hierzu betrachte zunächst einmal [*4 Pkt*]

$$w_i \;=\; \int_a^b L_i(x)\,dx \;=\; h\int_0^2 L_i(a + t\,h)\,dt \qquad \text{mit} \qquad h = \frac{b-a}{2}.$$

Damit ergibt sich dann

$$\begin{aligned}
w_0 \;&=\; h\int_0^2 \frac{t-1}{0-1}\cdot\frac{t-2}{0-2}\,dt \;=\; \frac{h}{2}\int_0^2 t^2 - 3\,t + 2\,dt \;=\; \frac{h}{2}\left(\frac{8}{3} - \frac{12}{2} + 4\right) \;=\; \frac{b-a}{6},\\
w_1 \;&=\; h\int_0^2 \frac{t-0}{1-0}\cdot\frac{t-2}{1-2}\,dt \;=\; -h\int_0^2 t^2 - 2\,t\,dt \;=\; -h\left(\frac{8}{3} - 4\right) \;=\; 4\,\frac{b-a}{6},\\
w_2 \;&=\; h\int_0^2 \frac{t-2}{2-0}\cdot\frac{t-1}{2-1}\,dt \;=\; \frac{h}{2}\int_0^2 t^2 - t\,dt \;=\; \frac{h}{2}\left(\frac{8}{3} - \frac{4}{2}\right) \;=\; \frac{b-a}{6}.
\end{aligned}$$

Man erhält daher die folgende Darstellung

$$Q_{[a,b]}[f] \;=\; \frac{b-a}{6}\left(f(a) + 4\,f\!\left(\frac{a+b}{2}\right) + f(b)\right).$$

Zu zeigen: $Q_{[a,b]}[p] - \int_a^b p(x)\,dx = 0$ für alle $p \in \Pi_3$.

Lösungsweg 1:

Da die Newton-Cotes-Formel für $m = 2$ mindestens den Exaktheitsgrad 2 besitzt, genügt es, nun die Aussage für $p(x) = x^3$ zu überprüfen. Hierbei ergibt sich nun aber

$$\begin{aligned}
\int_a^b p(x)\,dx \;&=\; \frac{b^4 - a^4}{4} \;=\; \frac{b-a}{4}\,(a^3 + a^2 b + a\,b^2 + b^3)\\
&=\; \frac{b-a}{6}\left(a^3 + \frac{4}{8}\left(a^3 + 3\,a^2 b + 3\,a\,b^2 + b^3\right) + b^3\right)\\
&=\; \frac{b-a}{6}\left(p(a) + 4\,p\!\left(\frac{a+b}{2}\right) + p(b)\right) \;=\; Q_{[a,b]}[p].
\end{aligned}$$

Lösungsweg 2:

Sei $q \in \Pi_2$ mit $q(x_0) = p(x_0)$, $q(x_1) = p(x_1)$ und $q(x_2) = p(x_2)$, wobei $x_i = a + i\,\frac{b-a}{2}$ für $i = 0, 1, 2$ ist. Weiter sei $s(x) := p(x) - q(x)$. Dann gilt zunächst einmal, dass

$$\begin{aligned} Q_{[a,b]}[s] - \int_a^b s(x)\,dx &= Q_{[a,b]}[p] - \int_a^b p(x)\,dx - Q_{[a,b]}[q] + \int_a^b q(x)\,dx \\ &= Q_{[a,b]}[p] - \int_a^b p(x)\,dx, \end{aligned}$$

wobei die Linearität des Integrals bzw. der Newton-Cotes-Formel und die Exaktheit der Newton-Cotes-Formel für Polynome mit Grad 2 benutzt wurde. Andererseits ist $s \in \Pi_3$ und besitzt die drei Nullstellen x_0, x_1, x_2, d.h. $s(x) = c\,(x-x_0)(x-x_1)(x-x_2)$. Daraus folgt aber

$$\begin{aligned} Q_{[a,b]}[s] - \int_a^b s(x)\,dx &= -c \int_a^b s(x)\,dx = -\frac{c}{2} \int_0^{b-a} x(2\,x + a - b)(x + a - b)\,dx \\ &= -\frac{c}{2} \int_0^{b-a} 2\,x^3 + 3\,x^2(a-b) + x\,(a-b)^2\,dx \\ &= -\frac{c}{2} \left[\frac{1}{2}\,x^4 - x^3(b-a) + \frac{1}{2}\,x^2(b-a)^2 \right]_0^{b-a} = 0, \end{aligned}$$

d.h. die Newton-Cotes-Formel für $m = 2$ besitzt einen Exaktheitsgrad ≥ 3.

Andererseits gilt für

$$\begin{aligned} Q_{[a,b]}\big[x^4\big] &= \frac{b-a}{6}\Big(a^4 + 4\left(\frac{a+b}{2}\right)^4 + b^4\Big) \\ &= \frac{5\,b^5 - a\,b^4 + 2\,a^2b^3 - 2\,a^3b^2 + a^4b - 5\,a^5}{24} \neq \frac{b^5 - a^5}{5} = \int_a^b x^4\,dx, \end{aligned}$$

d.h. der Exaktheitsgrad der Newton-Cotes-Formel für $m = 2$ ist genau 3.