Institut für angewandte Mathematik
Sommersemester 2013
Andreas Eberle, Matthias Erbar / Jan Maas

2. Klausur „Algorithmische Mathematik II"

Musterlösung

---

1. (Bedingte Wahrscheinlichkeiten)

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

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

Zunächst gilt $\mu_B[A] = \mu[A|B] \geq 0$ für alle $A \subseteq S$, und

$$\mu_B[S] = \frac{\mu[B\cap S]}{\mu[B]} = 1 \ .$$

Weiter sei $(A_i)_i$ eine abzählbare Kollektion von disjunkten Teilmengen von S. Dann gilt:

$$\begin{aligned}\mu_B[\bigcup_i A_i] &= \frac{\mu[B\cap(\bigcup_i A_i)]}{\mu[B]} = \frac{\mu[\bigcup_i(A_i\cap B)]}{\mu[B]}\\ &= \sum_i \frac{\mu[B\cap A_i]}{\mu[B]} = \sum_i \mu_B[A_i] \ ,\end{aligned}$$

da die Ereignisse $A_i \cap B$ wieder disjunkt sind. Damit ist auch μ_B eine Wahrscheinlichkeitsverteilung auf S.

b) (i) Wir behaupten, dass gilt:

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

In der Tat folgt aus der Additivität von μ und da die H_i disjunkt sind und $\cup_i H_i = S$:

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

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

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

Beweis: Dies folgt sofort aus

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

und der Formel für $\mu[A]$ aus (i).

c)

(i) Wir modellieren das Experiment entweder als mehrstufiges Modell auf $\Omega = \{1,2,3\} \times \{S,W\}$, oder alternativ auf folgendem Wahrscheinlichkeitsraum: Setze $n_1 = 6, n_2 = 8, n_3 = 4$ und definiere

$$\Omega = \{(i,j) \mid i \in \{1,2,3\}, j \in \{1,\dots,n_i\}\} , \quad \mathcal{A} = P(\Omega) , \quad P[(i,j)] = \frac{1}{3}\frac{1}{n_i} .$$

Weiter betrachten wir die Ereignisse

$$\begin{aligned} A_i &= \text{Urne i wird ausgewählt} , \quad i = 1,2,3 . \\ B &= \text{Eine weiße Kugel wird gezogen} . \end{aligned}$$

Damit ergibt sich (direkt oder mithilfe von bedingten W'keiten):

$$P[B] = P[B \cap A_1] + P[B \cap A_2] + P[B \cap A_3] = \frac{1}{3}\Big(\frac{4}{6} + \frac{5}{8} + \frac{3}{4}\Big) = \frac{49}{72} .$$

(ii) Gesucht ist die bedingte Wahrscheinlichkeit $P[A_2|B]$. Nach Definition oder durch Anwenden der Bayesschen Formel ergibt sich hierfür:

$$P[A_2|B] = \frac{P[A_2 \cap B]}{P[B]} = \frac{1}{3} \cdot \frac{5}{8} \cdot \frac{72}{49} = \frac{15}{49} .$$

d)

(i) Wir berechnen $P[T=1] = P[X_1 \in B] = \mu[B]$ und für $k > 1$:

$$P[T=k] = P[X_1,\dots,X_{k-1} \notin B, X_k \in B] = P[X_1 \notin B]^{k-1} P[X_1 \in B] = (1-\mu[B])^{k-1}\mu[B] ,$$

wobei wir benutzt haben, dass die X_i unabhängig und nach μ verteilt sind. Damit ist T geometrisch verteilt mit Parameter $\mu[B]$.

(ii) Wir berechnen

$$\begin{aligned} P[X_T \in A] &= \sum_{k=1}^{\infty} P[X_T \in A, T=k] = \sum_{k=1}^{\infty} P[X_1,\dots,X_{k-1} \notin B, X_k \in A \cap B] \\ &= \sum_{k=1}^{\infty} (1-\mu[B])^{k-1}\mu[A \cap B] = \frac{\mu[A \cap B]}{\mu[B]} = \mu_B[A] . \end{aligned}$$

Hierbei haben wir erneut die Unabhängigkeit der X_i benutzt sowie die Tatsache, dass für alle $0 < p < 1$ gilt:

$$\sum_{k=1}^{\infty} p^{k-1} = \frac{1}{1-p} .$$

e) Nach Aufgabenteil d) leistet der folgende Algorithmus das Gewünschte:
Input: Ereignis B mit $\mu[B] > 0$, **Output:** Stichprobe y verteilt nach μ_B

- Erzeuge Stichprobe $x \sim \mu$; $y := x$;
- **While** $y \notin B$ **Do**:
  - Erzeuge Stichprobe $x \sim \mu$; $y := x$;

Die mittlere Anzahl Schritte des Algorithmus ist gegeben durch den Erwartungswert der Zufallsvariable T aus Teil d). Wir setzen $p := \mu[B]$ und berechnen

$$E[T] = \sum_{k=1}^{\infty} k \cdot (1-p)^{k-1} p = -pf'(p) \ ,$$

wobei f definiert ist durch die Potenzreihe

$$f(p) = \sum_{k=1}^{\infty} (1-p)^k = \frac{1}{p} - 1$$

und wir gliedweise Differentiation verwendet haben. Damit ergibt sich

$$E[T] = -pf'(p) = \frac{1}{p} \ .$$

2. (Varianz, Kovarianz und Gesetz der großen Zahlen)

a) Die Varianz einer diskreten reellwertigen Zufallsvariable X mit $E[|X|] < \infty$ ist definiert durch

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

Seien X und Y Zufallsvariablen mit $\mathbb{E}[X^2], \mathbb{E}[Y^2] < \infty$. Die Kovarianz von X und Y ist definiert als

$$\mathrm{Cov}[X,Y] := \mathbb{E}[(X - \mathbb{E}[X])(Y - \mathbb{E}[Y])].$$

Es gilt

$$\begin{aligned} \mathrm{Var}[X+Y] &= \mathrm{Cov}[X+Y, X+Y]] \\ &= \mathrm{Var}[X] + \mathrm{Var}[Y] + 2\,\mathrm{Cov}[X,Y]. \end{aligned}$$

wegen Symmetrie und Bilinearität der Kovarianz.

b) Es gilt

$$\mathrm{Var}[X] = \mathbb{E}[X^2] - (\mathbb{E}[X])^2 = \mathbb{E}[X] - (\mathbb{E}[X])^2 = \frac{1}{3} - \left(\frac{1}{3}\right)^2 = \frac{2}{9}.$$

Analog gilt $\mathbb{E}[Y] = \frac{2}{9}$. Ausserdem,

$$\begin{aligned} \mathrm{Var}[X+Y] &= \mathbb{E}[(X+Y)^2] - (\mathbb{E}[(X+Y)])^2 \\ &= \mathbb{E}[X+Y] - (\mathbb{E}[X+Y])^2 = \frac{2}{3} - \left(\frac{2}{3}\right)^2 = \frac{2}{9}. \end{aligned}$$

Die Formel von a) liefert jetzt

$$\mathrm{Cov}[X,Y] = \frac{1}{2}\Big(\mathrm{Var}[X+Y] - \mathrm{Var}[X] - \mathrm{Var}[Y]\Big) = \frac{1}{2}\Big(\frac{2}{9} - \frac{2}{9} - \frac{2}{9}\Big) = -\frac{1}{9}.$$

Alternativ direkt via $\mathrm{Cov}[X,Y] = \mathbb{E}[XY] - E[X]E[Y]$ (diese Darstellung ist zu zeigen).

c) Es gilt, wegen Unabhängigkeit,

$$\begin{aligned} p_n &= \mathbb{P}[\exists j \le 2^n : Y_j = 4^n] \\ &= 1 - \mathbb{P}[\forall j \le 2^n : Y_j \neq 4^n] \\ &= 1 - \prod_{j=1}^{2^n} \mathbb{P}[Y_j \neq 4^n] \\ &= 1 - \prod_{j=1}^{2^n} (1 - 2^{-n}) \\ &= 1 - (1-2^{-n})^{2^n}, \end{aligned}$$

also gilt $\lim_{n\to\infty} 1 - p_n = \lim_{n\to\infty}(1-2^{-n})^{2^n} = e^{-1}$, und deshalb $\lim_{n\to\infty} p_n = 1 - e^{-1}$. (Alternativ:

$$\log(1-p_n) = \log((1-2^{-n})^{2^n}) = 2^n \log(1-2^{-n}) = 2^n(-2^{-n} + o(2^{-n})) = 1 + o(1),$$

also $\lim_{n\to\infty} p_n = 1 - e^{-1}$.)

Zudem:

$$\mathbb{P}\left[2^{-n} \sum_{1}^{2^n} Y_i \ge 2^n\right] \;=\; \mathbb{P}\left[\sum_{1}^{2^n} Y_i \ge 4^n\right] \;\ge\; \mathbb{P}\left[\bigcup_{i=1}^{2^n} \{Y_i \ge 4^n\}\right],$$

wobei die rechte Seite gegen $1 - e^{-1}$ konvergiert. Dies ist kein Widerspruch zum GGZ, da die Varianzen nicht beschränkt sind.

3. (Konvergenz von Fixpunktiterationen)

a) ***Banachscher Fixpunktsatz:*** Sei (X,d) ein vollständiger, metrischer Raum und $\phi : X \to X$. Ferner sei ϕ eine Kontraktion auf X, d.h. es gibt ein $L \in [0,1)$, so dass

$$d\big(\phi(x), \phi(y)\big) \;\le\; L\, d(x,y), \qquad \forall\, x,y \in X.$$

Dann gelten die folgenden Aussagen:

1. Es existiert genau ein $x^* \in X$ mit $\phi(x^*) = x^*$.
2. Für beliebiges $x^{(0)} \in X$ konvergiert

$$x^{(n+1)} = \phi(x^{(n)}) \qquad n = 0, 1, \ldots$$

gegen x^*.

3. Es gilt die Monotonie

$$d\big(x^{(n+1)}, x^*\big) \leq L d\big(x^{(n)}, x^*\big).$$

4. Es gilt die a priori Fehlerabschätzung

$$d\big(x^{(n)}, x^*\big) \leq \frac{L^n}{1-L} d\big(x^{(1)}, x^{(0)}\big).$$

5. Es gilt die a posteriori Fehlerabschätzung

$$d\big(x^{(n)}, x^*\big) \leq \frac{L}{1-L} d\big(x^{(n)}, x^{(n-1)}\big).$$

(Aussagen 3. bis 5. sind optional).

b) Es gilt

$$\|D\phi(x)\|_2 = \sup_{v \in \mathbb{R}^n,\, \|v\|_2 = 1} \|D\phi(x)v\|_2 \,,$$

wobei

$$D\phi(x)v = \lim_{h \to 0} \frac{\phi(x + hv) - \phi(x)}{h} \,.$$

Nach Dreiecksungleichung und Definition von $D\phi$ gilt

$$\begin{aligned} \|\phi(y) - \phi(x)\|_2 &= \left\| \int_0^1 D\phi((1-t)x + ty)(y-x)\, dt \right\|_2 \\ &\leq \int_0^1 \|D\phi((1-t)x + ty)\|_2 \, \|y - x\|_2 \, dt \\ &\leq L\|y - x\|_2 \,. \end{aligned}$$

Da $L < 1$, erfüllt ϕ die Voraussetzungen von dem Banachschen Fixpunktsatz. Da $(\mathbb{R}^n, \|\cdot\|_2)$ ein vollständiger metrischer Raum ist, folgt die Behauptung aus 1. und 2.

c) Sei

$$\phi(x) = \begin{pmatrix} \frac{1}{11}\sin x_1 + \frac{2}{11}\cos x_2 \\ \frac{3}{11}\cos x_1 + \frac{4}{11}\sin x_2 \end{pmatrix} \,.$$

Die Lösungen des Gleichungssystems sind dann genau die Fixpunkte von ϕ. Wegen b) reicht es also zu zeigen dass $L < 1$. Es gilt

$$D\phi(x) = \begin{pmatrix} \frac{1}{11}\cos x_1 & -\frac{2}{11}\sin x_2 \\ -\frac{3}{11}\sin x_1 & \frac{4}{11}\cos x_2 \end{pmatrix} =: A(x)$$

Mit Cauchy-Schwarz findet man

$$\begin{aligned}\|A(x)v\|_2^2 = \sum_i \left(\sum_j a_{ij}(x)v \right)^2 &\leq \|v\|^2 \sum_{ij} a_{ij}(x)^2 \\ &\leq \|v\|^2 \left[(\tfrac{3}{11})^2 + (\tfrac{4}{11})^2 \right] = (\tfrac{5}{11})^2 \|v\|^2 \ ,\end{aligned}$$

also gilt $L \leq \frac{5}{11} < 1$.

Iterationsverfahren $x^{(n+1)} = \phi(x^{(n)}$; eine explizite Fehlerabschätzung erhält man über die Aussagen 4 oder 5 aus Teil a).

d) Sei

$$\psi(x) = \begin{pmatrix} \frac{1}{11} \sin x_1 + \frac{2}{11} \cos x_2 - x_1 \\ \frac{3}{11} \cos x_1 + \frac{4}{11} \sin x_2 - x_2 \end{pmatrix} .$$

Die Lösung des Gleichungssystems ist dann genau die Nullstellen von ψ. Eine Iterationsschritt in dem Newton-Verfahren ist gegeben durch

$$D\psi(x^{(k)})h = -\psi(x^{(k)}) \ , \tag{1}$$

$$x^{(k+1)} = x^{(k)} + h \ . \tag{2}$$

Die explizite form von (1) ist gegeben durch

$$\begin{pmatrix} \frac{1}{11} \cos x_1 - 1 & -\frac{2}{11} \sin x_2 \\ -\frac{3}{11} \sin x_1 & \frac{4}{11} \cos x_2 - 1 \end{pmatrix} \begin{pmatrix} h_1 \\ h_2 \end{pmatrix} = \begin{pmatrix} \frac{1}{11} - \sin x_1 - \frac{2}{11} \cos x_2 + x_1 \\ -\frac{3}{11} \cos x_1 - \frac{4}{11} \sin x_2 + x_2 \end{pmatrix}$$

Es gilt:

- $x \mapsto D\psi(x)$ ist Lipschitz-stetig in einer Umgebung von x^* (da $\psi \in C^1$);
- $D\psi(x^*)$ ist invertierbar, da $D\psi(x^*) = D\phi(x^*) - I$ und $\|D\phi(x^*)\|_2 < 1$.

Satz 5.5 impliziert, dass eine Umgebung U von x^* existiert, so dass die Folge $\{x^{(k)}\}_k$ quadratisch gegen x^* konvergiert für alle Startwerte $x^{(0)} \in U$.

4. (Interpolation und Quadratur)

a) Sei $-\infty < x_0 < x_1 < \ldots < x_n < \infty$. Das Lagrange-Interpolationsproblem ist folgendes: finde ein Polynom p vom Grad $\leq n$, mit der Eigenschaft dass $p(x_i) = y_i$ für $i = 0, \ldots, n$, wobei $y_0, \ldots, y_n \in \mathbb{R}$ gegeben sind.

b) Die Lagrange Darstellung des Interpolationspolynom ist

$$p(x) = \sum_{i=0}^{n} y_i L_i(x),$$

wobei

$$L_i(x) = \prod_{j \neq i} \frac{x - x_j}{x_i - x_j}.$$

c) Die Newton-Interpolationsformel ist

$$\begin{aligned} p(x) = f[x_0] + f[x_0, x_1](x - x_0) + f[x_0, x_1, x_2](x - x_0)(x - x_1) + \ldots \\ + f[x_0, \ldots, x_n](x - x_0)(x - x_1) \cdots (x - x_{n-1}), \end{aligned}$$

wobei

$$\begin{aligned} f[x_i] &:= y_i, \\ f[x_i, x_{i+1}, \ldots, x_k] &:= \frac{f[x_{i+1}, x_{i+2}, \ldots, x_k] - f[x_i, x_{i+1}, \ldots, x_{k-1}]}{x_k - x_i}, \qquad 0 \leq i < k \leq n. \end{aligned}$$

d) Die Quadraturformel mit äquidistanten Knoten

$$x_i = a + i\frac{b - a}{m}, \qquad i = 0, 1, \ldots, m,$$

und Gewichten

$$w_i = \int_a^b L_i(x) w(x) dx$$

heißt abgeschlossene Newton-Cotes-Formel auf $[a, b]$.

e)

$$\begin{aligned} T_n[f] := \int_a^b s_n dx &= \sum_{i=0}^{n-1} \int_{x_i}^{x_{i+1}} s_n dx \\ &= h \cdot \sum_{i=0}^{n-1} \frac{f(x_{i+1}) + f(x_i)}{2} \end{aligned}$$

heißt n-te (zusammengesetzte) Trapezformel zur Approximation von $I[f] := \int_a^b f(x)dx$.

f) Das Quadraturverfahren Q_n hat Konsistenzordnung $s \geq 0$, falls

$$|Q_n[f] - I_w[f]| = O(n^{-s}) = O(h^s)$$

für alle hinreichend glatten Funktionen $f : [a, b] \to \mathbb{R}$ gilt, wobei $I_w[f] := \int_a^b f(x)w(x)dx$.