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

Klausur „Algorithmische Mathematik II”

Musterlösung

---

1. (Zufallsvariablen und ihre Verteilung)

a) Wir nehmen an, dass die Würfel fair sind und beide Würfe einer Runde sowie die Ergebisse der Runden untereinander unabhängig sind.

Zur Konstruktion des W-Raums setze $\Omega_0 = \{1, \dots, 6\}^2$ und definiere

$$\Omega = \Omega_0^{24} , \quad \mathcal{A} = P(\Omega) , \quad P[\{\omega\}] = |\Omega|^{-1} = \frac{1}{6^{48}} \; \forall \omega \in \Omega ,$$

i.e. P ist die Gleichverteilung auf Ω. Wir schreiben $\omega = ((x_1, y_1), \dots, (x_{24}, y_{24}))$.

Für $i = 1, \dots, 24$ definiere die Ereignisse

$$\begin{aligned} A_i &:= \{\text{Spieler gewinnt Runde i}\} \\ &= \{\omega \in \Omega \mid x_i + y_i = 10\} . \end{aligned}$$

Da es in einer Runde 3 günstige Ausgänge gibt, $10 = 6 + 4 = 5 + 5 = 4 + 6$, folgt

$$p = P[A_i] = \frac{|A_i|}{|\Omega|} = \frac{3 \cdot 6^{46}}{6^{48}} = \frac{1}{12} .$$

b) Für $i = 1, \dots, 24$ definiere Zufallsvariable $Y_i = \mathbf{1}_{A_i}$. Dann gilt:

$$P[Y_i = 1] = p , \quad P[Y_i = 0] = 1 - p .$$

Die Gesamtzahl der gewonnenen Runden ist beschrieben durch $S = \sum_{i=1}^{24} Y_i$. Da die Ereignisse A_i unabhängig sind, ist S $\mathrm{Bin}(24, p)$-verteilt, i.e.

$$P[S = k] = \binom{24}{k} p^k (1 - p)^{24-k} , \quad \text{für } k = 0, \dots, 24 .$$

Begründung: Es gibt $\binom{24}{k}$ Möglichkeiten k Runden auszuwählen, in denen der Spieler gewinnt. Die Wkeit dass der Spieler in diesen Runden gewinnt und in den übrigen verliert beträgt $p^k (1 - p)^{24-k}$.

Der Erwartungswert errechnet sich wie folgt:

$$E[S] = \sum_{i=1}^{24} E[Y_i] = 24 \cdot E[Y_1] = 24p .$$

(Skizze der Massefunktion der $\mathrm{Bin}(24, p)$ Verteilung)

c) Der Zeitpunkt des ersten Gewinns ist beschrieben durch

$$T := \min\{i \in \{0, \dots, 24\} \mid Y_i = 1\} ,$$

wobei $\min \emptyset = \infty$. Die Verteilung von T berechnet sich durch

$$\begin{aligned} P[T = k] &= \begin{cases} P[Y_1 = 0, \dots, Y_{k-1} = 0, Y_k = 1] , & k = 1, \dots, 24 , \\ P[Y_1 = 0, \dots, Y_{24} = 0] , & k = \infty \end{cases} \\ &= \begin{cases} (1-p)^{k-1} p , & k = 1, \dots, 24 , \\ (1-p)^{24} , & k = \infty \end{cases} , \end{aligned}$$

wobei in der letzen Gleichheit Unabhängigkeit benutzt wurde. Die Verteilung der Wartezeit auf den dritten Erfolg ist gegeben durch

$$\begin{aligned} P[T_3 = k] &= \sum_{i<j\le k-1} P[Y_k = Y_i = Y_j = 1, Y_l = 0 \ \forall l \le k-1, l \ne i, j] \\ &= p^3 (1-p)^{k-3} \sum_{i<j\le k-1} 1 \\ &= p^3 (1-p)^{k-3} \binom{k-1}{2} , \end{aligned}$$

für $k = 3, \dots, 24$. Wobei wieder in der letzten Gleichheit Unabhängigkeit benutzt wurde. Alternativ kann man auch argumentieren:

$$\begin{aligned} P[T_3 = k] &= P[A_k \text{ und genau zwei der } A_1, \dots, A_{k-1} \text{ treten ein}] \\ &= P[\text{genau zwei der } A_1, \dots, A_{k-1} \text{ treten ein}] \cdot P[A_k] \\ &= p^2 (1-p)^{k-3} \binom{k-1}{2} p . \end{aligned}$$

d) Es gilt

$$T \cdot \mathbf{1}_{\{T<\infty\}} = \sum_{k=1}^{24} k \cdot \mathbf{1}_{\{T=k\}} = \sum_{k=1}^{24} \mathbf{1}_{\{T\ge k\}} - 24 \cdot \mathbf{1}_{\{T=\infty\}} .$$

Die Behauptung folgt dann durch nehmen des Erwartungswerts auf beiden Seiten. Wir berechnen nun zunächst:

$$\begin{aligned} P[T \ge k] &= P[A_1^c \cap \cdots \cap A_{k-1}^c] = (1-p)^{k-1} , \\ P[T = \infty] &= (1-p)^{24} . \end{aligned}$$

Damit ergibt sich

$$E[T; T < \infty] = \sum_{k=1}^{24} (1-p)^{k-1} - 24 \cdot (1-p)^{24} = \big[1 - (1+24p)(1-p)^{24}\big]/p .$$

2. (Mittelwerte und Varianz)

a) Der Erwartungswert und die Varianz einer Zufallsvariable $X : \Omega \to S$ auf einem W'Raum $(\Omega, \mathcal{A}, P)$ mit $S \subset \mathbb{R}$ abzählbar sind definiert durch

$$E[X] := \sum_{a \in S} a \cdot P[X = a] ,$$
$$Var[X] := E[(X - E[X])^2] = \sum_{a \in S} (a - E[X])^2 \cdot P[X = a] ,$$

falls die Summen wohldefiniert sind, d.h. unabhängig von der Reihenfolge der Abzählung.

Hier wählen wir den W'Raum

$$\Omega = \{1, \ldots, 6\}^n , \quad \mathcal{A} = P(\Omega) , \quad P[\{\omega\}] = |\Omega|^{-1} = \frac{1}{6^n} \ \forall \omega \in \Omega ,$$

i.e. P ist die Gleichverteilung auf Ω. Wir schreiben $\omega = (\omega_1, \ldots, \omega_n)$. Sei weiter $X_i(\omega) = \omega_i$. Dann gilt $Y_n = \sum_{i=1}^n X_i$ sowie

$$E[X_i] = \frac{1}{6}(1+2+3+4+5+6) = \frac{7}{2} , \quad E[X_i^2] = \frac{1}{6}(1+4+9+16+25+36) = \frac{91}{6} .$$

Daraus schliessen wir

$$E[Y_n] = \sum_{i=1}^n E[X_i] = \frac{7}{2} n ,$$
$$Var[Y_n] = \sum_{i=1}^n Var[X_i] = n(E[X_i^2] - E[X_i]^2) = n \frac{35}{12} .$$

b) **Chebycheff-Ungleichung:** Für $X \in \mathcal{L}^2$ and $c > 0$ gilt:

$$P[|X - E[X]| \geq c] \leq \frac{1}{c^2} Var[X] .$$

Beweis: Setze $A = \{|X - E[X]| \geq c\}$ und bemerke, dass $1 \leq c^{-2}(X - E[X])^2$ auf A. Damit folgt:

$$P[|X - E[X]| \geq c] = E[\mathbf{1}_A] \leq E[\mathbf{1}_A \cdot \frac{1}{c^2}(X - E[X])^2]$$
$$\leq E[\frac{1}{c^2}(X - E[X])^2] = \frac{1}{c^2} Var[X] .$$

Damit erhalten wir:

$$P[|\frac{Y_n}{n} - 3.5| \geq 0.1] \leq 100 \cdot Var[\frac{Y_n}{n}] = \frac{1}{n} 100 \cdot \frac{35}{12} .$$

Damit die rechte Seite ≤ 0.1 ist, muss gelten $n \geq 35.000/12 \approx 2917$.

c) Wir berechnen den Erwartungswert durch Bedingen auf N mit Hilfe der Formel von der totalen Wkeit.

$$\begin{aligned} E[Y_N] &= \sum_{n=1}^{\infty} E[Y_N | N = n] \cdot P[N = n] = \sum_{n=1}^{\infty} E[Y_n | N = n] \cdot P[N = n] \\ &= \sum_{n=1}^{\infty} E[Y_n] \cdot P[N = n] = \sum_{n=1}^{\infty} n \cdot \frac{7}{2} \cdot P[N = n] = \frac{7}{2} \cdot E[N] = 70 \, , \end{aligned}$$

wobei die dritte Gleichheit aus Unabhängigkeit von N und Y_n folgt.

3. (Gleichgewichte von Markovketten)

a) Eine Folge von ZV $X_0, X_1, \dots : \Omega \to S$ auf WRaum $(\Omega, \mathcal{A}, P)$ heißt **zeitlich homogene Markovkette** mit Startverteilung ν und Übergangswahrscheinlichkeit p, falls:

$$\begin{aligned} P[X_0 = x_0] &= \nu(x_0) \quad \forall x_0 \in S \, , \\ P[X_0 = x_0, \dots, X_n = x_n] &= \nu(x_0) \cdot p(x_0, x_1) \cdots p(x_{n-1}, x_n) \quad \forall n \in \mathbb{N}, x_i \in S \, . \end{aligned}$$

Hierbei ist $p = \Big(p(x,y)\Big)_{x,y \in S}$ eine stochastische Matrix, i.e.

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

Anstelle der zweiten Gleichung oben kann man auch fordern:

$$P[X_0 = x_0, \dots, X_{n+1} = x_{n+1} \mid X_0 = x_0, \dots, X_n = x_n] = p(x_n, x_{n+1}) \quad \forall n \in \mathbb{N}, x_i \in S \, ,$$

falls $P[X_0 = x_0, \dots, X_n = x_n] \neq 0$.

Ein Gleichgewicht der Markovkette ist eine Wahrscheinlichkeitsverteilung μ auf S, so dass

$$(\mu p)(y) := \sum_{x \in S} \mu(x) p(x,y) = \mu(y) \quad \forall y \in S \, .$$

Erfüllt eine W'keitsverteilung μ nun Detailed Balance, so folgt

$$\sum_{x \in S} \mu(x) p(x,y) = \sum_{x \in S} \mu(y) p(y,x) = \mu(y) \, ,$$

da p stochastische Matrix. Somit ist μ ein Gleichgewicht.

b) i) Die Übergangsw'keiten des RW auf (V, E) sind gegeben durch

$$p(x,y) = \begin{cases} \frac{1}{3}, & \{x,y\} \in E, \\ 0, & sonst. \end{cases}$$

Damit gilt $p(x,y) = p(y,x)$ für alle $x, y \in V$. Die Gleichverteilung ν auf V erfüllt $\nu(x) = \nu(y) = |V|^{-1}$ für alle $x, y \in V$. Daraus folgt sofort die Detailed Balance Bedingung, und somit ist ν ein Gleichgewicht von p nach a).

ii) Zur Simulation von μ verwenden wir den
Metropolis Algorithmus:
Input: x_0, p, μ;
Output: Stichproben $x_0, x_1, \cdots$ einer MK mit Gleichgewicht μ.
For $n := 1, 2, \ldots$ Do:

* Erzeuge Stichprobe $y_n \sim p(x_{n-1}, \cdot)$, erzeuge $u_n \sim Unif(0,1)$,
* If $u_n \leq \mu(y_n)/\mu(x_{n-1})$ then $x_n := y_n$, else $x_n := x_{n-1}$.

Die Übergangsmatrix dieser Markovkette ist gegeben durch

$$q(x,y) = \begin{cases} \alpha(x,y)p(x,y), & x \neq y, \\ 1 - \sum_{z \neq x} \alpha(x,z)p(x,z), & x = y. \end{cases}$$

mit $\alpha(x,y) = \min(1, \mu(y)/\mu(x))$. Es genügt zu zeigen, dass μ Detailed Balance erfüllt für q, dann ist es mit a) ein Gleichgewicht. DB muss nur für $x, y \in V$ mit $x \neq y$ geprüft werden. Dann folgt

$$\mu(x)q(x,y) = \mu(x) \min\Big(1, \frac{\mu(y)}{\mu(x)}\Big) p(x,y) = \min(\mu(x), \mu(y)) p(y,x) = \mu(y)q(y,x) ,$$

da der vorletzte Ausdruck symmetrisch in x, y ist.

c) Wir bestimmen zunächst Gleichgewichte der Kette. μ ist Gleichgewicht genau dann wenn:

$$\mu(r) = \frac{1}{2} \cdot \mu(r) + q \cdot \mu(s) ,$$
$$\mu(s) = (1-q) \cdot \mu(s) + \frac{1}{2} \cdot \mu(r) ,$$

sowie $\mu(r) + \mu(s) = 1$. Daraus folgt

$$\mu(r) = \frac{2q}{1+2q} , \qquad \mu(s) = \frac{1}{1+2q} ,$$

die Kette besitzt also ein eindeutiges Gleichgewicht. Nach Ausführen des Experiments für lange Zeit erwartet man ein Ergebnis verteilt nach μ nach dem Konvergenzsatz für Markovketten. Nach dem Gesetz der großen Zahlen sollte daher bei vielmaliger Wiederholung gelten:

$$\frac{\#\text{Ausgänge r}}{\#\text{Versuche}} \approx \mu(r) , \quad \frac{\#\text{Ausgänge s}}{\#\text{Versuche}} \approx \mu(s) .$$

Damit folgt $\mu(r) = (1+2q)^{-1} \approx 0.8$ und wir folgern $q = 1/8$.

4. (Iterationsverfahren für lineare Gleichungssysteme)

a) Der Jacobi-Algorithmus ist gegeben durch

Input: Matrix $A \in \mathbb{R}^{d\times d}$ mit $a_{ii} \neq 0$ für alle $i \leq d$, $b \in \mathbb{R}^d$, Startwert $x^{(0)} \in \mathbb{R}^d$,
Output: Iterationsfolge $x^{(k)} \in \mathbb{R}^d$

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

for $i = 1, 2, \ldots, d$ **do**

$$x_i^{(k+1)} = \frac{1}{a_{ii}}\Big(b_i - \sum_{j\neq i} a_{ij}\, x_j^{(k)}\Big)$$

end for

end for

Der Gauß-Seidel-Algorithmus ist gegeben durch

Input: Matrix $A \in \mathbb{R}^{d\times d}$ mit $a_{ii} \neq 0$ für alle $i \leq d$, $b \in \mathbb{R}^d$, Startwert $x^{(0)} \in \mathbb{R}^d$,
Output: Iterationsfolge $x^{(k)} \in \mathbb{R}^d$

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

for $i = 1, 2, \ldots, d$ **do**

$$x_i^{(k+1)} = \frac{1}{a_{ii}}\Big(b_i - \sum_{j<i} a_{ij}\, x_j^{(k+1)} - \sum_{j>i} a_{ij}\, x_j^{(k)}\Big)$$

end for

end for

b) Setze $A = D - L - U$ wobei

$$d_{ij} = a_{ij}\mathbf{1}_{\{i=j\}}\ , \qquad \ell_{ij} = -a_{ij}\mathbf{1}_{\{i>j\}}\ , \qquad u_{ij} = -a_{ij}\mathbf{1}_{\{i<j\}}\ .$$

Jacobi: Es gilt

$$x^{(k+1)} = D^{-1}\Big(b + (L+U)x^{(k)}\Big)\ ,$$

also $x^{(k+1)} = T_J x^{(k)} + f_J$ mit $T_J = D^{-1}(L+U)$ und $f_J = D^{-1}b$.

Gauß-Seidel:

$$\begin{aligned} x^{(k+1)} &= D^{-1}\Big(b + Lx^{(k+1)} + Ux^{(k)}\Big) && \Leftrightarrow \\ (I - D^{-1}L)x^{(k+1)} &= D^{-1}b + D^{-1}Ux^{(k)} && \Leftrightarrow \\ x^{(k+1)} &= (I - D^{-1}L)^{-1}D^{-1}b + (I - D^{-1}L)^{-1}D^{-1}Ux^{(k)} && \Leftrightarrow \\ x^{(k+1)} &= (D-L)^{-1}b + (D-L)^{-1}Ux^{(k)}\ , \end{aligned}$$

also $T_{GS} = (D-L)^{-1}U$ und $f_{GS} = (D-L)^{-1}b$.

- Der Spektralradius einer Matrix $T \in R^{d\times d}$ ist

$$\rho(T) := \max\{|\lambda|\ :\ \lambda \in \mathbf{C} \text{ Eigenwert von } T\}.$$

c) (3) impliziert mit Induktion dass

$$x^{(k+1)} - x^* = T^k(x^{(k)} - x^*) \ ,$$

also $x^{(k)} - x^* = T^k(x^{(0)} - x^*)$. Sei $v \neq 0$ so dass $Tv = \lambda v$ mit $|\lambda| \geq 1$ und setze $x^{(0)} = x^* + v$. Dann gilt für alle k,

$$\|x^{(k)} - x^*\| = \|T^k(x^{(0)} - x^*)\| = \|T^k v\| = |\lambda|^k \|v\| \geq \|v\|,$$

also konvergiert $\{x^{(k)}\}$ nicht gegen x^*.

d) (i) Da A strikt diagonaldominant ist, d.h., $|a_{ii}| > \sum_{j\neq i} |a_{ij}|$ für $i = 1, ..., d$, konvergieren das Jacobi- und das Gauß-Seidel-Verfahren monoton bzgl. der Maximumsnorm $\| \cdot \|_\infty$ auf $\mathbb{R}^d$. (Satz 5.11)

(ii) Es gilt

$$T_J = \begin{pmatrix} 0 & 2 \\ 1 & 0 \end{pmatrix} \ ,$$

Dieser Matrix hat Eigenwerte $\pm\sqrt{2}$. Also impliziert c) dass das Jacobi Verfahren nicht für alle Anfangswerte konvergiert.

Es gilt

$$T_{GS} = \begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix}^{-1} \begin{pmatrix} 0 & 2 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 2 \\ 0 & 2 \end{pmatrix} \ ,$$

Dieser Matrix hat Eigenwerte 0 und 2. Also impliziert c) dass das Gauß-Seidel-Verfahren nicht für alle Anfangswerte konvergiert.

5. (Interpolation)

a) Sei $f \in C^{m+1}([a,b])$ und $a \leq x_0 < \ldots < x_m \leq b$.

Zu zeigen: Eindeutigkeit des Interpolationspolynoms p zu $(x_i, f(x_i))$.

Angenommen es gibt zwei verschiedenen Polynome p, q vom Höchstgrad m mit der Eigenschaft, dass

$$p(x_k) \;=\; q(x_k) \;=\; f(x_k) \qquad \forall\, k = 0, \ldots, m.$$

Dann besitzt das Polynom $p - q$ aber $m+1$ Nullstellen. Nach dem Hauptsatz der Algebra (Beweis via Satz von Rolle) ist ein Polynom vom Grad m mit $m+1$ Nullstellen aber identisch 0, d.h. $p - q \equiv 0$. Dies steht aber im Widerspruch zur Annahme.

Zu zeigen: Existenz des Interpolationspolynoms.

Betrachte hierzu die Lagrange-Polynome vom Grad m, die gegeben sind durch

$$L_i(x) \;=\; \prod_{\substack{j=0\\ j\neq i}}^{m} \frac{x - x_j}{x_i - x_j}.$$

Diese besitzen nun die Eigenschaft, dass $L_i(x_k) = \delta_{i,k}$. Daher löst aber das Polynom

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

gerade das Interpolationsproblem zu $(x_k, f(x_k))$ für $k = 0, \ldots, m$. Alternativ Beweis der Existenz wie in Vorlesung: Die lineare Abbildung $p \mapsto (p(x_0), \ldots, p(x_m))$ von Π_m nach $\mathbb{R}^{m+1}$ ist injektiv, also aus Dimensionsgründen sogar bijektiv.

b) Man definiere eine Funktion p durch die rechte Seite:

$$p(x) := \frac{(x - x_i)p_{i+1,k}(x) - (x - x_k)p_{i,k-1}(x)}{x_k - x_i}$$

Durch Induktion nach $k - i$ folgt: Da $p_{i,k-1}$ und $p_{i+1,k}$ Polynome mit Höchstgrad $k - i - 1$ sind, ist p ein Polynom mit Höchstgrad $k - i$. Außerdem gilt $p(x_i) = y_i$, $p(x_k) = y_k$, und für alle $j = i+1, \ldots, k-1$ gilt

$$\begin{aligned} p(x_j) &= \frac{(x_j - x_i)p_{i+1,k}(x_j) - (x_j - x_k)p_{i,k-1}(x_j)}{x_k - x_i} \\ &= \frac{(x_j - x_i)y_j - (x_j - x_k)y_j}{x_k - x_i} = y_j. \end{aligned}$$

Teil a) impliziert dass nur ein Polynom von Höchstgrad $k - i$ existiert mit dieser Eigenschaft, also gilt $p = p_{i,k}$.

c) **Input:** Integer $0 \leq m$, Punkte $-\infty < x_1 < \ldots < x_m < \infty$, Werte $y_1, \ldots, y_m \in \mathbb{R}, x \in \mathbb{R}$
Output: Wert $p(x)$ des Interpolationspolynoms

for $k = 0, 1, \ldots, m$ **do**
  $p_{k,k}(x) = y_k$
end for
for $d = 1, \ldots, m$ **do**
  for $i = 0, \ldots, m-d$ **do**
    $$p_{i,i+d}(x) = \frac{(x - x_i)p_{i+1,i+d}(x) - (x - x_{i+d})p_{i,i+d-1}(x)}{x_{i+d} - x_i}$$
  end for
end for
$p(x) = p_{0,m}(x)$

d) (i) Das lineare Gleichungssystem

$$\begin{pmatrix} 0 & 0 & 1 \\ \frac{\pi^2}{16} & \frac{\pi}{4} & 1 \\ \frac{\pi^2}{4} & \frac{\pi}{2} & 1 \end{pmatrix} \begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} 1 \\ \frac{1}{2}\sqrt{2} \\ 0 \end{pmatrix}$$

hat die Lösung $a = \frac{8-8\sqrt{2}}{\pi^2}$, $b = \frac{4\sqrt{2}-6}{\pi}$, $c = 1$.

(ii) Setze $x_1 = 0$, $x_2 = \frac{\pi}{4}$, $x_3 = \frac{\pi}{2}$ und $y_1 = 1$, $y_2 = \frac{1}{2}\sqrt{2}$, $y_3 = 0$. Es gilt

$$\begin{aligned} L(x) &= \sum_{k=1}^{3} y_k \prod_{j \neq k} \frac{x - x_j}{x_k - x_j} \\ &= 1 \cdot \frac{(x - \frac{\pi}{4})(x - \frac{\pi}{2})}{(-\frac{\pi}{4})(-\frac{\pi}{2})} + \frac{1}{2}\sqrt{2} \cdot \frac{x(x - \frac{\pi}{2})}{\frac{\pi}{4}(-\frac{\pi}{2})} + 0 \cdot \{\ldots\} \\ &= \frac{8 - 8\sqrt{2}}{\pi^2} x^2 + \frac{4\sqrt{2} - 6}{\pi} x + 1 \,. \end{aligned}$$

(iii) Mit der Rekursionsformel zeigt man

$$p_{00}(x) = 1 \,, \qquad p_{11}(x) = \frac{1}{2}\sqrt{2} \,, \qquad p_{22}(x) = 0 \,,$$

$$p_{01}(x) = \frac{4}{\pi}\left(\frac{1}{2}\sqrt{2} - 1\right)x + 1 \,, \qquad p_{12}(x) = -\frac{2}{\pi}\sqrt{2}x + \sqrt{2} \,,$$

$$p_{02}(x) = \frac{8 - 8\sqrt{2}}{\pi^2} x^2 + \frac{4\sqrt{2} - 6}{\pi} x + 1 \,.$$