ETH Zürich
Dept. Informatik

Prof. Markus Gross
WS 05/06

Informationstheorie

Lösung 2

2.1 Über die Güte von Wetterprognosen

a) $P[W \neq V] = P_{VW}(regen, schön) + P_{VW}(schön, regen) = 3/16 + 3/16 = 6/16$, während $P[W \neq V'] = P_W(regen) = 2/16 + 3/16 = 5/16$.
Somit scheint Kachelmanns Vorhersage schlechter zu sein.

b) Richtet man sich nach Kachelmanns Vorhersage und wählt einen Tag mit der Vorhersage $V = schön$, dann ist die Wahrscheinlichkeit, dass es wirklich *schön* ist gleich

$$P_{W|V}(schön, schön) = \frac{P_{WV}(schön, schön)}{P_V(schön)} = \frac{8/16}{3/16 + 8/16} = \frac{8}{11} \approx 0.73$$

im Vergleich zu der Wahrscheinlichkeit

$$P_W(schön) = \frac{3}{16} + \frac{8}{16} = \frac{11}{16} \approx 0.69,$$

dass das Grillfest nicht verregnet wird, wenn man irgendeinen Tag wählt (was man bei der Vorhersage V' wohl oder übel machen muss).
Auch wenn man einen regnerischen Tag wünscht, fährt man mit Kachelmanns Vorhersage besser: Man hat dann $P_{W|V}(regen, regen) = 2/5 = 0.4$ im Vergleich zu $P_W(regen) = 5/16 \approx 0.31$.
Von dieser Seite her gesehen ist also Kachelmanns Vorhersage besser.

c) Offensichtlich kann man mit Hilfe von V'' das Wetter ohne Fehler vorhersagen: Man prognostiziert einfach *schön*, falls $V'' = regen$, und *regen*, falls $V'' = schön$.

d) Nach obiger Überlegung gibt V'' die volle Information über das Wetter, während V' gar keine Information liefert und V irgendwo dazwischen liegt.

2.2 Stochastischer Prozess

a)

$$E(X_0) = \sum_{i \in \{0,1,2,3,4\}} P(X_0 = i) \cdot i = 2$$

$$Var(X_0) = E(X_0^2) - E(X_0)^2 = \sum_{i \in \{0,1,2,3,4\}} P(X_0 = i) \cdot i^2 - 4 = 6 - 4 = 2$$

b) Ja, denn X_i haengt nur von X_{i-1} ab.

c) Wir zeigen per Induktion, dass alle Verteilung Gleichverteilungen über den gesamten Wertebereich sind.

Induktionsanfang: $P(X_0)$ ist eine Gleichverteilung auf $\{0, 1, 2, 3, 4\}$.

Induktionsschluss: Zu zeigen ist: Wenn $P(X_i)$ eine Gleichverteilung auf $\{0, 1, 2, 3, 4\}$ ist, so ist auch $P(X_{i+1})$ eine.

Für alle $k \in \{0, 1, 2, 3, 4\}$ gilt laut der Definition von $P(X_{i+1})$:

$$P(X_{i+1} = k) = P(X_i \equiv (k-1) \bmod 5) \cdot 0.5 + P(X_i \equiv (k+1) \bmod 5) \cdot 0.5$$

Da X_i gleichverteilt ist, ist $P(X_i = k) = p = \frac{1}{5} \quad \forall k \in \{0, 1, 2, 3, 4\}$. Daraus folgt:

$$P(X_{i+1} = k) = p \cdot 0.5 + p \cdot 0.5 = p = \frac{1}{5}$$

Also ist X_{i+1} gleichverteilt.

d) Zu zeigen ist, dass die gemeinsamen Verteilungen $P(X_t, ..., X_{t+n})$ gleich sind für alle n (Für $n = 0$ haben wir das gerade gezeigt). Wir zeigen dies mittels Induktion über n.

Induktionsanfang: Alle X_i sind gleichverteilt, wie gerade gezeigt. Also ist die Aussage für $n = 0$ bewiesen. Weiterhin ist

$$P(X_t = x_t, X_{t+1} = x_{t+1}, ..., X_{t+n} = x_{t+n}) = \begin{cases} 0.2 \cdot 0.5^n & \text{wenn } (x_t, x_{t+1}, \ldots, x_{t+n}) \\ & \text{erlaubt, also } \forall k \in \{0, \ldots, n-1\} \\ & (x_{t+k} - x_{t+k+1}) \bmod 5 \equiv \pm 1 \\ 0 & \text{sonst} \end{cases}$$

$P(X_t, ..., X_{t+n})$ ist also eine Gleichverteilung auf allen möglichen Tupeln $(x_t, x_{t+1}, ..., x_{t+n})$.

Induktionsschluss: Zu zeigen ist: Sind die Verteilungen $P(X_t, ..., X_{t+n})$ $\forall t$ identisch, so sind es auch die Verteilungen $P(X_t, ..., X_{t+n+1})$, und $P(X_t, ..., X_{t+n+1})$ ist wieder eine Gleichverteilung aller erlaubten Tupel $(x_t, x_{t+1}, ..., x_{t+n+1})$.

Jedes erlaubte Tupel $(x_t, x_{t+1}, ..., x_{t+n})$ kann man auf zwei Arten erweitern. Mit jeweils Wahrscheinlichkeit 0.5 wird entweder $x_{t+n+1} = (x_{t+n} + 1) \bmod 5$ oder $x_{t+n+1} = (x_{t+n} + 1) \bmod 5$ hinzugenommen. Das sich ergebende Tupel $(x_t, x_{t+1}, ..., x_{t+n+1})$ kann auf keinem anderen Weg erreicht werden. Also ist die Wahrscheinlichkeit für ein beliebiges erlaubtes Tupel:

$$P(X_t = x_t, \ldots, X_{t+n+1} = x_{t+n+1}) = P(X_t = x_t, \ldots, X_{t+n} = x_{t+n}) \cdot 0.5$$

Wenn $P(X_t = x_t, \ldots, X_{t+n} = x_{t+n})$ eine Gleichverteilung ist, ist es auch $P(X_t = x_t, \ldots, X_{t+n+1} = x_{t+n+1})$. Da die erlaubten Tupel für alle t gleich sind, sind die Gleichverteilungen $P(X_t = x_t, \ldots, X_{t+n+1} = x_{t+n+1})$ gleich. $\square$

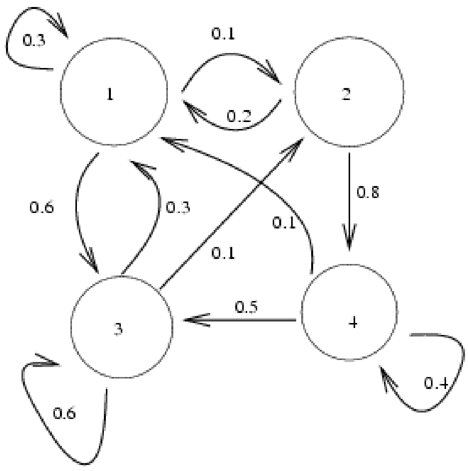
2.3 Markov-Kette

Wir müssen Y so definieren, dass Z nicht mehr von X abhängt, wenn Y gegeben ist. Wenn X gerade ist, dann ist Z gleichverteilt über die ungeraden Werte. Wenn X ungerade ist, ist Z gleichverteilt über die geraden Werte. Die Abhängigkeit von Z von X steckt also in der Parität von X.

Wenn wir nun $Y = X \pmod 2$ wählen, ist $P_{Z|XY} = P_{Z|Y}$.

2.4 Markov-Zustandsautomaten

a) Der Zustandsgraph sieht so aus:



b) Jeder Zustandsübergang kann betrachtet werden als eine Multiplikation mit M. Die Wahrscheinlichkeitsverteilung erhält man also am einfachsten durch

$$P_3(X) = M^3 \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 0.275 \\ 0.083 \\ 0.538 \\ 0.104 \end{bmatrix}$$

c) Die totalen Wahrscheinlichkeiten stehen im Eigenvektor v_1 zum Eigenwert 1, positiv und normiert nach der 1-norm ($\|v_1\|_1 = 1$, die Summe der Elemente ist 1), i.e. die Werte sind tatsaechlich eine Wahrscheinlichkeitsverteilung. Matlab sez:

$$P(X) = v_1 = \begin{bmatrix} 0.2703 \\ 0.0811 \\ 0.5405 \\ 0.1081 \end{bmatrix}$$

d) Ist der momentane Zustand der Automaten bekannt, so ist die totale Wahrscheinlichkeit im allgemeinen wenig Aussagekräftig. Die bedingten Wahrscheinlichkeiten in der Übergangsmatrix, insbesondere die zum aktuellen Zustand gehörige Spalte ist interassanter. Diese Spalte gibt die Verteilung des Zustands nach einem Zustandsübergang an.

Für die Wahrscheinlichkeitsverteilung $P_n(X|x_i)$, die die Wahrscheinlichkeiten der Zustände nach n Zustandsübergängen angibt (bei ursprünglichem Zustand x_i), gilt, dass

$$\lim_{n\to\infty} P_n(X|x_i) = P(X)$$

Egal, von welchem Zustand mal startet, sofern $P(x_i) \neq 0$. [Für einen Beweis siehe die Herleitung der Potenzmethode zur Eigenwertberechnung.]

Man kann also bei unserem kleinen Markov-Automaten erwarten, dass die Wahrscheinlichkeitsverteilung $P_{100}(X)$ recht nahe an $P(X)$ liegt.