Vorlesung Sprachverarbeitung

Übung 6

1. Hidden Markov Modelle

(a) Der verrückte Getränkeautomat
Gegeben sei ein Getränkeautomat, der auf Knopfdruck eines von drei Getränken ausgibt. Stellen Sie diesen Automaten mittels eines HMMs mit zwei Zuständen (s1, s2) und der folgenden Übergangsmatrix dar:

$$A = \begin{pmatrix} 0.7 & 0.3 \\ 0.5 & 0.5 \end{pmatrix} \tag{1}$$

Die Ausgabewahrscheinlichkeit für die einzelnen Getränke lautet:

$$\begin{array}{ll} \text{Cola}_{s1} = 0.6 & \text{Cola}_{s2} = 0.1 \\ \text{Tee}_{s1} = 0.1 & \text{Tee}_{s2} = 0.7 \\ \text{Limo}_{s1} = & \text{Limo}_{s2} = \end{array}$$

i. Wie lauten die Ausgabewahrscheinlichkeiten für Limo in Zustand $s1$ und $s2$?
ii. Geben Sie den Wahrscheinlichkeitsausdruck für die Ausgabe b_{ijk} an (mit den Zuständen s_i und s_j und dem Getränk k) unter der Berücksichtigung, dass zuerst ein Übergang und dann eine Ausgabe erfolgt?
iii. Der Automat befindet sich in Zustand s1. Mit welcher Wahrscheinlichkeit wird zuerst eine Limo und dann ein Tee ausgegeben?

(b) Gegeben seien zwei HMMs mit den Übergangswahrscheinlichkeiten A_1 und A_2:

$$A_1 = \begin{pmatrix} 0.2 & 0.8 & 0.0 \\ 0.0 & 0.6 & 0.4 \\ 0.0 & 0.0 & 1.0 \end{pmatrix} \quad A_2 = \begin{pmatrix} 0.6 & 0.4 & 0.0 \\ 0.0 & 0.8 & 0.2 \\ 0.0 & 0.0 & 1.0 \end{pmatrix}$$

sowie den Produktionswahrscheinlichkeitsmatrizen B_1 und B_2:

$$B_1 = \begin{pmatrix} 0.8 & 0.0 & 0.2 \\ 0.2 & 0.7 & 0.3 \\ 0.0 & 0.3 & 0.5 \end{pmatrix} \quad B_2 = \begin{pmatrix} 0.7 & 0.4 & 0.5 \\ 0.2 & 0.6 & 0.0 \\ 0.1 & 0.0 & 0.5 \end{pmatrix}$$

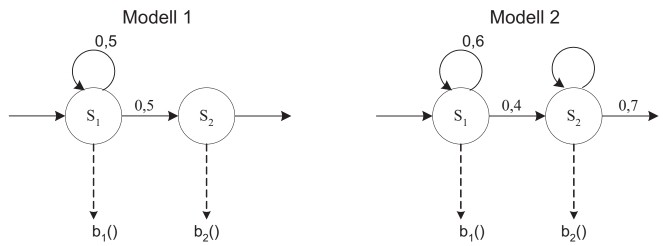
die Spalten von B_1 und B_2 repräsentieren die Zustände, die Zeilen die Werte a, b, c. Die Startwahrscheinlichkeit lautet für beide Modelle:

$$\pi = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$$

i. Stellen Sie unter Berücksichtigung der Übergangswahrscheinlichkeiten alle mögliche Pfade für 5 Zeitschritte durch die beiden Modelle dar (der letzte Zeitschritt muss sich im letzten Zustand befinden!).

ii. Gesucht sind alle möglichen Pfade zur Produktion der Zeichenfolge $O = \{aabcc\}$. Reduzieren Sie die möglichen Pfade unter Berücksichtigung der Produktionswahrscheinlichkeiten.
iii. Berechnen Sie die Gesamtwahrscheinlichkeiten für beide Modelle.
iv. Berechnen Sie die Viterbi-Wahrscheinlichkeiten und stellen Sie den wahrscheinlichsten Pfad im Diagramm dar.

(c) Schätzung von HMM Parametern
Gegeben seien die beiden folgenden HMMs sowie eine Beobachtungsfolge



$O_1 = (o_1 = 1, o_2 = 2, o_3 = 4)$ mit der die Parameter des ersten Modells geschätzt werden sollen und die Beobachtungsfolge $O_2 = (o_1 = 3, o_2 = 3, o_3 = 3)$ mit der das zweite HMM trainiert werden soll.
Skizzieren Sie alle erlaubten Wege der Beobachtungen durch die jeweiligen Modelle.

(d) Schätzen Sie mit Hilfe des Baum-Welch Verfahrens die Produktionswahrscheinlichkeiten b_1 und b_2 sowie die Übergangswahrscheinlichkeiten.
Die Startparameter lauten für Modell 1:
$\mu_1 = 1$ und $\sigma_1 = 1$ und $\mu_2 = 2$ und $\sigma_2 = 1$
Die Startparameter lauten für Modell 2:
$\mu_1 = 1$ und $\sigma_1 = 1$ und $\mu_2 = 4$ und $\sigma_2 = 1$

(e) Von welchem Modell wurde die Beobachtungsfolge $O_{Test}(o_1 = 2, o_2 = 3, o_3 = 4)$ am wahrscheinlichsten produziert (verwenden Sie den Vorwärtsalgorithmus)?

Anhang

gaussverteilte Produktionswahrscheinlichkeiten:

$$b_i(o_t) = \frac{1}{\sigma_i\sqrt{2\pi}} \cdot exp[-\frac{1}{2}(\frac{o_t - \mu_i}{\sigma_i})^2]$$

Berechnung der Alpha-Matrix:

$$\alpha_j(t) = \left(\sum_{i=1}^{N} \alpha_i(t-1)a_{ij}\right) b_j(o_t) \quad \text{mit} \quad 1 < j < N, 1 < t \leq T$$

Randbedingung: $\alpha_j(1) = a_{1j}b_j(o_1)$

Berechnung der Beta-Matrix:

$$\beta_i(t) = \sum_{j=1}^{N} a_{ij}b_j(o_{t+1})\beta_j(t+1) \quad \text{mit} \quad 1 < i < N, 1 \leq t < T$$

Randbedingung: $\beta_i(T) = 1$

Neuschätzung der Übergangswahrscheinlichkeiten $\hat{a}_{ij}$:

$$\hat{a}_{ij} = \frac{\sum_{t=1}^{T-1} \alpha_i(t) a_{ij} b_j(o_{t+1}) \beta_j(t+1)}{\sum_{t=1}^{T} \alpha_i(t) \beta_i(t)} \quad \text{mit:} \quad 1 < i < N, 1 < j < N$$

mit: $\hat{a}_{1j} = 1/P \alpha_j(1) \beta_j(1)$

$$\hat{a}_{iN} = \frac{\alpha_i(T) \beta_i(T)}{\sum_{t=1}^{T} 1/P \alpha_i(t) \beta_i(t)}$$

Neuschätzung der Mittelwerte $\hat{\mu}_j$:

$$\hat{\mu}_j = \frac{\sum_{t=1}^{T} \frac{1}{P} \alpha_j(t) \beta_j(t) o_t}{\sum_{t=1}^{T} \frac{1}{P} \alpha_j(t) \beta_j(t)} \quad \text{mit:} \quad P = P(O|\lambda) = \alpha_N(T)$$