

Тема 6: ОСНОВИ НА ТЕОРИЯТА НА СКРИТИТЕ МАРКОВСКИ МОДЕЛИ ПРИ РАЗПОЗНАВАНЕ НА РЕЧ

Скритите марковски модели (The hidden Markov model – HMM) е много мощен статистически метод за определяне на наблюдаваните данни в дискретни редици (discrete-time series).

1. Марковски вериги

От гледна точка на теорията на вероятностите един случаен процес притежава **Марковско свойство**, ако условното разпределение на вероятностите на бъдещи състояния на наблюдавания процес, при предварително известни текущи и минали състояния, зависи само от състоянието към този момент, и не зависи от миналите такива. С други думи бъдещите състояния на един процес са условно независими от миналите състояния, при развитие на наблюдавания процес.

Процес, който притежава това свойство се нарича **Марковски процес**. Най-популярните Марковски процеси са Марковските вериги.

Марковската верига е Марковски процес, който приема стойности от дискретното множество, наречено пространство на състоянията, като стойността му се променя в точно определени моменти от време. Тези вериги са широко разпространени и използвани в теорията на разпознаване на реч, тъй като заемат много малко памет за моделиране на динамични процеси и освен това са лесни за реализация.

2. Дефиниране на скрити Марковски модели (Hidden Markov Model – HMM)

В терминологията на български език освен *скрити Марковски модели* се използва и *скрити модели на Марков*. За целите на настоящият курс от тук напред ще използваме само първата нотация.

Скритите Марковски модели е статистически модел, при който системата, която е била моделирана, се приема да бъде Марковски процес с неизвестни параметри, с цел определяне на скритите параметри на изследваните процеси. Извлечените параметри по модела могат после да бъдат използвани за по-нататъшни анализи. Всеки определен (наблюдаван) параметър е функция на вероятностите на дадено състояние. Всяко състояние може да се възпроизведе, което не позволява да се определи поредицата от състояния до момента на наблюдение. С други думи поредицата от състояния не може да бъде наблюдавана и остава скрита. Скритите Марковски модели могат да се разглеждат и като най-обикновена динамична мрежа на Бейс.

В общия модел на Марков, състоянието е директно видимо от наблюдателя и за това вероятните състояния на промяна са единствените параметри. В скрития Марковски модел,

състоянието не е пряко видимо, но променливите повлияни от състоянието, са видими. Всяко състояние има вероятност да бъде разпределено върху възможните изходи признаци.

Скритите Марковски модели са особено известни с приложенията си в разпознаването на реч, ръкопис, жестове, и др.

2.1 Динамично програмиране (Dynamic Programming) и Динамично времево изкривяване (Dynamic Time Warping - DTW)

Разпознаването на реч основано на DTW е прост и ефективен метод при разпознаващи системи с малък речник от думи.

Алгоритъм за динамично програмиране [1]

Стъпка 1: Първоначална инициализация

$$D(1, 1) = d(1,1), B(1, 1) = 1, \text{ for } j = 2, \dots, M \text{ и се изчислява } D(1, j) = \infty$$

Стъпка 2: Итерации

при $i = 2, \dots, N$

{

за $j = 1, \dots, M$ се изчислява

{

$$D(i, j) = \min_{1 \leq p \leq M} [D(i-1, p) + d(p, j)]$$

$$B(i, j) = \arg \min_{1 \leq p \leq M} [D(i-1, p) + d(p, j)]$$

}

}

Стъпка 3: Връщане назад и преустановяване на процеса

Оптималното (минималното) разстояние е $D(N, M)$ и оптималния път е (s_1, s_2, \dots, s_N) , където

$$s_N = M \text{ и } s_i = B(i+1, s_{i+1}), \quad i = N-1, N-2, \dots, 1$$

2.2 Оценка на скрити Марковски модели с алгоритъма Forward

The Forward Algorithm [1] with complexity – $O(N^2T)$

Стъпка 1: Инициализация

$$\alpha_1 = \pi_1 b_1(X_1), \quad 1 \leq i \leq N$$

Стъпка 2: Индукция

$$\alpha_t(j) = \left[\sum_{i=1}^N \alpha_{t-1}(i) a_{ij} \right] b_j(X_t), \quad 2 \leq t \leq T; \quad 1 \leq j \leq N$$

Стъпка 3: Преустановяване на процеса

$$P(X|\Phi) = \sum_{i=1}^N \alpha_T(i)$$

Ако е необходимо да се завърши в крайно състояние,

$$P(X|\Phi) = \alpha_T(s_F)$$

2.3 Декодиране на скрити Марковски модели с алгоритъм на Витерби (Viterbi Algorithm)

Стъпка 1: Инициализация

$$V_1(i) = \pi_i b_i(X_1)$$

$$B_1(i) = 0$$

Стъпка 2: Въведение

$$V_t(j) = \max_{1 \leq i \leq N} [V_{t-1}(i) a_{ij}] b_j(X_t), \quad 2 \leq t \leq T; \quad 1 \leq j \leq N$$

$$B_t(j) = \arg \max_{1 \leq i \leq N} [V_{t-1}(i) a_{ij}], \quad 2 \leq t \leq T; \quad 1 \leq j \leq N$$

Стъпка 3: Прекратяване

$$\text{The best score} = \max_{1 \leq j \leq N} [V_T(j)]$$

$$S_T^* = \arg \max_{1 \leq i \leq N} [B_T(i)]$$

Стъпка 4: Връщане назад

$$s_t^* = B_{t+1}(s_{t+1}^*), \quad t = T-1, T-2, \dots, 1$$

$S^* = (s_1^*, s_2^*, \dots, s_T^*)$ е най-добрата последователност

Сложността на алгоритъма на Витерби е $O(N^2T)$.

2.4 Оценяване на параметрите на скрит Марковски модел с алгоритъма на Баум-Уелч (Baum-Welch algorithm)

Итеративния алгоритъм на Баум-Уелч е известен още като алгоритъм *forward-backward*.

Стъпки на алгоритъма Forward-backward:

Стъпка 1: Инициализация

Избира се първоначалната оценка Φ .

Стъпка 2: E-Стъпка

Изчисляване на спомагателната функция $Q(\Phi, \hat{\Phi})$ въз основа на Φ .

Стъпка 3: M-Стъпка

Изчисляване на $\hat{\Phi}$ според преоценката за максимизиране на спомагателната Q-функция.

Стъпка 4: Итерация

Нека $\Phi = \hat{\Phi}$, алгоритъмът се повтаря от стъпка 2 до приближаване към определена стойност (конвергенция).

3. Непрекъснати и полунепрекъснати скрити Марковски модели

- *Непрекъснати смесени плътности на скрити Марковски модели (Continuous Mixture Density HMMs)*
- *Полунепрекъснати скрити Марковски модели (Semi-continuous HMMs)*

4. Практически аспекти при използването на скрити Марковски модели

- Първоначални оценки
- Топология на модела: Използване на *нулев преход*, е подходящи в случаите, при които е необходимо опростено преминаване през скрит Марковски модел, без използването на всички наблюдавани символи.
- Критерии за обучение
- Изтрита интерполации
- Замъгляване на параметрите

- Представяне на вероятностите

5. Ограничения на скрития Марковски модел

Има редица ограничения в стандартните скрити Марковски модели, като:

- Скритите Марковски модели допускат продължителност следвана от експоненциално разпределение;
- Вероятността на прехода зависи само от източника и предназначението;
- Всички наблюдавани рамки са зависими само от състоянието, което ги поражда, без да се интересува от наблюдаваните съседни рамки.

Изследователите са предложили редица техники за справяне с тези ограничения, макар че тези решения не осигуряват значителни подобрения в точността на разпознаване на реч.

Тези подходи са:

- Продължително моделиране
- Предположения от първи ред
- Условно независими предположения

Задачи за работа по време на упражненията и самостоятелна работа

Задача 1: Разработете програмна реализация на алгоритъма forward.

Литература

- [1] Xuedong Huang, Alex Acero, Hsiao-Wuen Hon, *Spoken Language processing – A Guide to Theory, Algorithm, and System Development*, Prentice Hall PTR, 2001, pp. 375-407
- [2] Joseph Keshet, Samy Bengio, *Automatic Speech and Speaker Recognition – Large Margin and Kernel Method*, John Wiley & Sons, 2009
- [3] Gerosa, M., Giuliani, D., Narayanan, S. (2006) *Acoustic analysis and automatic recognition of spontaneous children's speech*, In proc. of InterSpeech-2006
- [4] Д. Вълчев (2002) *Теория на вероятностите - Тема 16: Марковски вериги*, <http://www.fmi.uni-sofia.bg/fmi/statist/Personal/Vandev/lectures/prob/Markov.htm>