

Домашнее задание : Program НММ

Красильникова Виктория

9 марта 2019 г.

В ходе работы будут рассмотрены основные алгоритмы скрытых Марковских моделей.

1 Определение скрытых марковских моделей

Скрытая Марковская модель или кратко НММ формально может быть описана следующим образом.

- Множество дискретных состояний $S = (S_1, S_2, \dots, S_M)$
- Соостояние в момент времени $t : q_t$
- Множество возможных сигналов $V = (V_1, V_2, \dots, V_K)$
- Матрица переходов A из состояния в состояние на мн-ве S , $A_{ij} = P(q_{t+1} = S_j | q_t = S_i)$
- Матрица B , которая содержит вероятности выдать сигнал в определенном состоянии $B_{ij} = b_i(j) = P(v_j | q_t = S_i)$
- Наблюдаемая последовательность сигналов $O = (O_1, O_2, \dots, O_L)$
- Вектор начального состояния π , $\pi_i = P(q_1 = S_i)$
- НММ Model $\lambda = (A, B, \pi)$

2 НММ program

2.1 Алгоритм Витерби

Реализуем алгоритм Viterbi для нахождения наиболее вероятной последовательности состояний данной НММ по последовательности наблюдаемых сигналов. Дана последовательность $O = (O_1, O_2, \dots, O_T)$ и модель $\lambda = (A, B, \pi)$. Необходимо подобрать последовательность состояний системы $Q = q_1, q_2, \dots, q_T$, которая лучше всего соответствует наблюдаемой последовательности, т.е. объясняет последовательность наблюдений O .

2.2 Back-forward propagation

Необходимо реализовать алгоритм Back-forward для нахождения вероятности последовательности наблюдаемых сигналов для данной НММ.

Таким образом дана последовательность $O = (O_1, O_2, \dots, O_T)$ и модель $\lambda = (A, B, \pi)$ Необходимо вычислить вероятность $P(O, \lambda)$ что данная последовательность построена именно для нашей модели.

Алгоритм состоит из 2-х частей

1). Вычисление Forward значений $\alpha_t(i) = P(O_1, O_2, \dots, O_t, q_t = S_i | \lambda)$

- $\alpha_1(i) = \pi_i b_i(O_1)$
- $\alpha_{t+1}(j) = (\sum_{i=1}^N \alpha_t(i) a_{ij}) b_j(O_{t+1})$
- $P(O | \lambda) = \sum_{i=1}^N \alpha_T(i)$

2). Вычисление Backward значений $\beta_t(i) = P(O_{t+1}, \dots, O_T, q_t = S_i | \lambda)$

- $\beta_T(i) = 1$
- $\beta_t(i) = \sum_{j=1}^N a_{ij} b_j(O_{t+1}) \beta_{t+1}(j)$
- $P(O | \lambda) = \sum_{i=1}^N \beta_1(i) \pi_i b_i(O_1)$

3). Вычисление постериорных вероятностей

- $\gamma_t(i) = \frac{\alpha_t(i) \beta_t(i)}{P(O | \lambda)}$
- $P(O | \lambda) = \sum_{i=1}^N \alpha_t(i) \beta_t(i)$

3 Выводы

В ходе выполнения работы были реализованы основные алгоритмы модели скрытых Марковских цепей. Также были вычислены наиболее вероятная последовательность состояний и параметры НММ на основе наблюдаемых сигналов в предположении, что модель описывается эргодической Марковской цепью.

Код исследования представлен в файле *hmm.ipynb*

Полученные значения параметров записаны в следующих файлах:

- Начальное распределение вероятностей pi.csv
- Матрица переходов A.csv
- Матрица эмиссий B.csv