Методы исследования скрытых марковских моделей

Багина Алёна Алексеевна, 522 -я группа

Санкт-Петербургский Государственный Университет
Математико-механический факультет
Кафедра статистического моделирования

Научный руководитель — д.ф.-м.н., доцент, Н.К. Кривулин
Рецензент: к.ф.-м.н., Шпилёв П.В.

Санкт-Петербург
2015г.

Постановка задачи.

Цель:
Рассмотреть и изучить методы исследования скрытых марковских моделей (СММ):

- Алгоритм Витерби
- Алгоритм прямого—обратного хода
- Алгоритм Баума—Уэлша

Задачи:
Для достижения поставленной цели, в работе рассмотрены следующие задачи:

- Исследование скрытых марковских моделей и их особенностей;
- Изучение трех основных задач, связанных с СММ;
- Изучение методов решения основных задач;
- Разработка программы для реализации методов исследования СММ.

Постановка задачи.

Цель:
Рассмотреть и изучить методы исследования скрытых марковских моделей (СММ):

- Алгоритм Витерби
- Алгоритм прямого—обратного хода
- Алгоритм Баума—Уэлша

Задачи:
Для достижения поставленной цели, в работе рассмотрены следующие задачи:

- Исследование скрытых марковских моделей и их особенностей;
- Изучение трех основных задач, связанных с СММ;
- Изучение методов решения основных задач;
- Разработка программы для реализации методов исследования СММ.

Постановка задачи.

Цель:
Рассмотреть и изучить методы исследования скрытых марковских моделей (СММ):

- Алгоритм Витерби
- Алгоритм прямого—обратного хода
- Алгоритм Баума—Уэлша

Задачи:
Для достижения поставленной цели, в работе рассмотрены следующие задачи:

- Исследование скрытых марковских моделей и их особенностей;
- Изучение трех основных задач, связанных с СММ;
- Изучение методов решения основных задач;
- Разработка программы для реализации методов исследования СММ.

Постановка задачи.

Цель:
Рассмотреть и изучить методы исследования скрытых марковских моделей (СММ):

- Алгоритм Витерби
- Алгоритм прямого—обратного хода
- Алгоритм Баума—Уэлша

Задачи:
Для достижения поставленной цели, в работе рассмотрены следующие задачи:

- Исследование скрытых марковских моделей и их особенностей;
- Изучение трех основных задач, связанных с СММ;
- Изучение методов решения основных задач;
- Разработка программы для реализации методов исследования СММ.

Предмет исследования

- Модель описывается двумя последовательностями сл.в. $\{\xi_k\}$ и $\{\eta_k\}$. Последовательность $\{\xi_k\}$ состоит из сл.в. со значениями $1, \ldots, n$ и образует конечную однородную цепь Маркова с множеством состояний $1, \ldots, n$. Известны вероятности
$$p_{ij} = P(\xi_k = j \mid \xi_{k-1} = i),\ p_i = P(\xi_0 = i),\ i, j = 1, \ldots, n,$$
составляющие матрицу переходных вероятностей и вектор начальных вероятностей
$$P = \begin{pmatrix} p_{11} & \ldots & p_{1n} \\ \vdots & \ddots & \vdots \\ p_{n1} & \ldots & p_{nn} \end{pmatrix},\ p = \begin{pmatrix} p_1 \\ \vdots \\ p_n \end{pmatrix}.$$
- Последовательность сл.в. η_k принимает значения $1, \ldots, n$ и связана с цепью Маркова ξ_k условными вероятностями
$$q_i(j) = P(\eta_k = j \mid \xi_k = i),\ i, j = 1, \ldots, n.$$
Эти вероятности можно записать в виде матрицы
$$Q = \begin{pmatrix} q_1(1) & \ldots & q_1(n) \\ \vdots & \ddots & \vdots \\ q_n(1) & \ldots & q_n(n) \end{pmatrix}.$$

Предполагается, что состояния цепи $\{\xi_k\}$ не доступны для наблюдения, однако, о них можно в определенной мере судить по наблюдениям известных значений последовательности $\{\eta_k\}$.
Пусть $X_k = \{\xi_k = i_k\}$, $Y_k = \{\eta_k = j_k\}$, $k = 1, 2, \ldots$.
Сделаем предположение о независимости случайных величин:

1. Случайная величина ξ_k (скрытое состояние в момент времени k) цепи при заданном значении величины ξ_{k-1} (состояние в момент времени $k-1$) не зависит от предыдущих случайных величин (событий),

$$P(X_k | X_1 Y_1 \cdots X_{k-1} Y_{k-1}) = P(X_k | X_{k-1}).$$

2. Случайная величина η_k (k-е наблюдение) при заданном значении ξ_k (скрытого состояния в момент времени k) не зависит от других случайных величин (событий),

$$P(Y_k | X_1 Y_1 \cdots X_{k-1} Y_{k-1} X_k Y_k X_{k+1} Y_{k+1} \cdots X_K Y_K) = P(Y_k | X_k).$$

Основные задачи

Основные задачи для скрытых марковских моделей:

- Задача оценки модели, которая заключается в вычислении вероятности того, что модель соответствует заданной наблюдаемой последовательности, то есть насколько выбранная СММ соответствует заданной наблюдаемой последовательности. (Алгоритм прямого—обратного хода)
- Задача восстановления состояний цепи, в которой необходимо подобрать последовательность состояний системы, которая лучше всего соответствует наблюдаемой последовательности, то есть "объясняет"наблюдаемую последовательность. (Алгоритм Витерби)
- Задача оценки модели, которая заключается в подборе параметров модели таким образом, чтобы она как можно лучше описывала реальную наблюдаемую последовательнсоть. (Алгоритм Баума—Уэлша)

Основные задачи

Основные задачи для скрытых марковских моделей:

- Задача оценки модели, которая заключается в вычислении вероятности того, что модель соответствует заданной наблюдаемой последовательности, то есть насколько выбранная СММ соответствует заданной наблюдаемой последовательности. (Алгоритм прямого—обратного хода)
- Задача восстановления состояний цепи, в которой необходимо подобрать последовательность состояний системы, которая лучше всего соответствует наблюдаемой последовательности, то есть "объясняет"наблюдаемую последовательность. (Алгоритм Витерби)
- Задача оценки модели, которая заключается в подборе параметров модели таким образом, чтобы она как можно лучше описывала реальную наблюдаемую последовательнность. (Алгоритм Баума—Уэлша)

Основные задачи

Основные задачи для скрытых марковских моделей:

- Задача оценки модели, которая заключается в вычислении вероятности того, что модель соответствует заданной наблюдаемой последовательности, то есть насколько выбранная СММ соответствует заданной наблюдаемой последовательности. (Алгоритм прямого—обратного хода)
- Задача восстановления состояний цепи, в которой необходимо подобрать последовательность состояний системы, которая лучше всего соответствует наблюдаемой последовательности, то есть "объясняет"наблюдаемую последовательность. (Алгоритм Витерби)
- Задача оценки модели, которая заключается в подборе параметров модели таким образом, чтобы она как можно лучше описывала реальную наблюдаемую последовательнность. (Алгоритм Баума—Уэлша)

Алгоритм Витерби (Forney G. D. 1973)

- **Дано:** последовательность наблюдений $j_1, \dots, j_k$ случайных величин $\eta_1, \dots, \eta_k$;
- **Найти:** последовательность значений $i_1, \dots, i_k$ случайных величин $\xi_1, \dots, \xi_k$, такую что

$$\max_{i_1,\dots,i_k} F_{j_1,\dots,j_k}(i_1, \dots, i_k),$$

где целевая функция имеет вид

$$F_{j_1,\dots,j_k}(i_1, \dots, i_k) = P(\xi_1 = i_1, \dots, \xi_k = i_k \mid \eta_1 = j_1, \dots, \eta_k = j_k).$$

Процедура последовательного расчета для каждого $k = 1, 2, \dots$ вероятностей состояний скрытой цепи Маркова по формулам

$$f_1(i) = p_i,$$

$$f_k(i) = q_i(j_k) \max_{i \le j \le n} p_{ji} f_{k-1}(j) \; i = 1, \dots, n, \; k > 1,$$

соответствует общей вычислительной схеме динамического программирования и называется алгоритмом Витерби.

Алгоритм прямого—обратного хода (Rabiner L. 1989)

- **Дано:** вероятности модели p, P, Q и последовательность наблюдений $j_1, \dots, j_k$;
- **Найти:** вероятность появления данной последовательности.

$$\alpha_k(i) = P(\eta_1 = j_1, \dots, \eta_k = j_k, \xi_k = i),$$

$$\alpha_1(i) = q_i(j_1)p_i,$$

$$\alpha_k(i) = q_i(j_k)\sum_{j=1}^{n} p_{ji}\alpha_{k-1}(j),\ i = 1, \dots, n,\ k = 2, 3, \dots$$

вероятность, что в момент наблюдения k система находится в состоянии i.

$$\beta_k(i) = P(\eta_{k+1} = j_{k+1}, \dots, \eta_K = j_K \mid \xi_k = i),$$

$$\beta_k(i) = \sum_{j=1}^{n} p_{ij}\beta_{k+1}q_j(j_{k+1}),\ k = K-1, K-2, \dots, 1,$$

$$\beta_K(i) = 1,\ i = 1, \dots, n.$$

$$\gamma_k(i) = P(\xi_k = i \mid \eta_1 = j_1, \ldots, \eta_K = j_K),$$

$$\gamma_k(i) = \frac{\alpha_k(i)\beta_k(i)}{\sum_{j=1}^{n} \alpha_k(j)\beta_k(j)}, \ i = 1, \ldots, n; \ k = 1, \ldots, K.$$

$$\delta_k(i,j) = P(\xi_k = i, \xi_{k+1} = j \mid \eta_1 = j_1, \ldots, \eta_K = j_K),$$

$$\delta_k(i,j) = \frac{q_j(j_{k+1}) p_{ij} \alpha_k(i) \beta_{k+1}(j)}{\sum_{i=1}^{n} \sum_{j=1}^{n} q_i(j_{k+1}) p_{ji} \alpha_k(j) \beta_{k+1}(i)}.$$

вероятность конечной заданной последовательности, при условии, что мы начали из исходного состояния i в момент наблюдения k.

Алгоритм Баума—Уэлша (Rabiner L. 2013)

- **Дано:** последовательность наблюдений $j_1, \ldots, j_K$;
- **Найти:** параметры модели, т.е. найти вектор начальных вероятностей p и матрицы переходных и условных вероятностей P и Q, $\theta = (p, P, Q)$. Для определения параметров модели требуется решить задачу

$$\max_{\theta} L(j_1, \ldots, j_K; \theta),$$

$$\sum_{i=1}^{n} p_i = 1,$$

$$\sum_{j=1}^{n} p_{ij} = 1, \ \sum_{j=1}^{n} q_i(j) = 1; \ i = 1, \ldots, n,$$

где

$$L(j_1, \ldots, j_K; \theta) = P(Y_1 \ldots Y_K) =$$
$$= \sum_{1 \leq i_1, \ldots, i_K \leq n} p_{i_1} q_{i_1}(j_1) p_{i_1 i_2} \ldots q_{i_{K-1}}(j_{K-1}) p_{i_{K-1} i_K} q_{i_K}(j_K).$$

- Алгоритм представляет собой итеративную процедуру уточнения параметров p'_i, p'_{ij} и $q'_i(j)$ набора $\theta' = (p, P, Q)$. Уточненные значения вычисляются по формулам

$$p_i = \gamma'_1(i),\ p_{ij} = \frac{\sum_{k=1}^{K-1} \delta'_k(i,j)}{\sum_{k=1}^{K-1} \gamma'_k(i)},\ q_i(j) = \frac{\sum_{k=1, j_k=j}^{K-1} \gamma'_k(i)}{\sum_{k=1}^{K} \gamma'_k(i)}. \qquad (1)$$

- На первом шаге алгоритма выбираются некоторые начальные значения параметров. На основе этих значений и результатов наблюдений параметры уточняются по формулам (1), а затем процедура повторяется для новых значений параметров. Вычисления завершаются при стабилизации значений параметров в пределах заданной точности.

Описание программы

Наша модель:

Размерность задачи:2

Вектор начальных вероятностей:

$p[1] = 0.7$ $p[2] = 0.3$

Матрица вероятностей переходов:

$P[1,1] = 0.8$ $P[1,2] = 0.2$

$P[2,1] = 0.9$ $P[2,2] = 0.1$

Матрица вероятностей наблюдений:

$Q_1(1) = 0.6$ $Q_1(2) = 0.4$

$Q_2(1) = 0.3$ $Q_2(2) = 0.7$

Вектор состояний:2121111112

Вектор наблюдений:2121211112

Вектор состояний по алгоритму Витерби:1111111112

Модель:

Вектор начальных вероятностей:

$p[1] = 0.7 \; p[2] = 0.3$

Матрица вероятностей переходов:

$P[1,1] = 0.8 \; P[1,2] = 0.2$

$P[2,1] = 0.9 \; P[2,2] = 0.1$

Матрица вероятностей наблюдений:

$Q_1(1) = 0.6 \; Q_1(2) = 0.4$

$Q_2(1) = 0.3 \; Q_2(2) = 0.7$

Алгоритм Баума-Уэлша:

$p[1] = 0.00069051 \; p[2] = 0.99931$

$P[1,1] = 0.72701 \; P[1,2] = 0.27299$

$P[2,1] = 0.984157 \; P[2,2] = 0.0158429$

$Q_1(1) = 0.551189 \; Q_1(2) = 0.448811$

$Q_2(1) = 0.00130969 \; Q_2(2) = 0.99869$

$$p[1] = 0.7 \; p[2] = 0.3$$
$$P[1,1] = 0.8 \; P[1,2] = 0.2$$
$$P[2,1] = 0.9 \; P[2,2] = 0.1$$
$$Q_1(1) = 0.6 \; Q_1(2) = 0.4$$
$$Q_2(1) = 0.3 \; Q_2(2) = 0.7$$

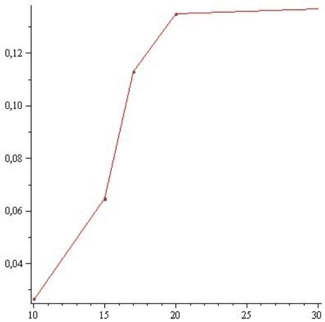

Рис.: Зависимость доли ошибок алгоритма Витерби от количества наблюдений в серии из 100 экспериментов.

$$p[1] = 0.7 \; p[2] = 0.3$$
$$P[1,1] = 0.8 \; P[1,2] = 0.2$$
$$P[2,1] = 0.9 \; P[2,2] = 0.1$$
$$Q_1(1) = 0.6 \; Q_1(2) = 0.4$$
$$Q_2(1) = 0.3 \; Q_2(2) = 0.7$$

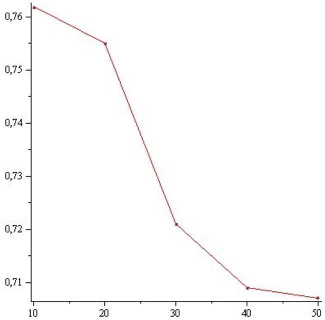

Рис.: Зависимость доли правильно угаданных состояний алгоритма Витерби от количества наблюдений в серии из 100 экспериментов.

Результаты

- Изучены скрытые марковские модели, их свойства и особенности;
- Поставлены и изучены основные задачи и пути их решения с помощью различных схем: алгоритмы Витерби, прямого—обратного хода, Баума—Уэлша;
- Составлена программа по реализации алгоритмов Витерби, прямого—обратного хода и Баума—Уэлша в среде MS Visual Studio C++;
- Представлены экспериментальные зависимости доли правильно угаданных состояний от количества наблюдений и доли ошибок в алгоритме Витерби.