Лабораторная работа № 5

Скрытые марковские модели (СММ) с дискретным пространством наблюдений

Цель работы

Научиться моделировать наблюдаемые и скрытые последовательности, порождаемы СММ.

Указания к работе

Случайный процесс X_t, $t \in T$ называется марковским случайным процессом, если для всех $t_1 < t_2 < \ldots < t_n \in T$, $n > 1$ его условное распределение вероятностей в момент t_n не зависит от значений, которые процесс принимает в моменты $t_1, t_2, \ldots, t_{n-2}$, а определяется только значением процесса в момент t_{n-1}, т. е.:

$$f\left(x_n, t_n \mid x_{n-1}, t_{n-1}\right) = f\left(x_n, t_n \mid x_1, t_1, x_2, t_2, \ldots, x_{n-1}, t_{n-1}\right).$$

Данное свойство называется свойством марковости (или марковским свойством). Функции $f\left(x_n, t_n \mid x_{n-1}, t_{n-1}\right)$ называют вероятностями перехода из состояния x_{n-1} в состояние x_n. Вероятности перехода удовлетворяют двум основным соотношениям:

1) условию нормировки

$$\int f\left(x, t \mid x', t'\right) dx = 1;$$

2) уравнению Чепмена-Колмогорова

$$f\left(x_3, t_3 \mid x_1, t_1\right) = \int f\left(x_3, t_3 \mid x_2, t_2\right) f\left(x_2, t_2 \mid x_1, t_1\right) dx_2$$

для любых x_1, x_2, x_3, соответствующих моментам $t_1 < t_2 < t_3$.

Произвольный случайный процесс при соответствующем выборе фазового пространства может превращаться в марковский процесс, т. е. такой, для которого дальнейшая эволюция зависит от прошлого только через состояние в настоящий момент. Фазовое пространство (или пространство состояний) – это пространство, которому принадлежат все возможные значения, принимаемые случайным процессом.

Дискретная марковская цепь $\left\{x_n\right\}$ представляет собой марковский случайный процесс, пространство состояний которого счетно или конечно, и, кроме того, этот случайный процесс дискретен по $T = \left\{0, 1, 2, \ldots\right\}$. Часто пространство состояний процесса отождествляют с множеством неотрицательных целых чисел и говорят, что случайный процесс находится в состоянии i, если $x_n = i$. Вероятность того, что случайный процесс $\left\{x_n\right\}$ в момент t_{n+1} находится в состоянии j, если известно, что в момент t_n он находился в состоянии i, обозначается как:

$$P_{ij}^{n,n+1} = P\{x_{n+1} = j \mid x_n = i\}.$$

В этой формуле, таким образом, подчеркивается, что в общем случае переходные вероятности зависят не только от начального и конечного состояний, но и от момента осуществления перехода. Когда одношаговые переходные вероятности не зависят от переменной t, то такой марковский процесс обладает стационарными переходными вероятностями. В лабораторной работе будем рассматривать именно такой случай, т. е. когда вероятность $P_{ij} = P_{ij}^{n,n+1}$ не зависит от n и является вероятностью перехода из состояния i в состояние j за одно испытание. Обычно эти вероятности объединяют в матрицу, которая называется марковской матрицей, матрицей условных вероятностей переходов или матрицей переходных вероятностей марковской цепи.

Если число состояний конечно, то марковская матрица является конечной квадратной матрицей, порядок которой равен числу состояний. В общем случае, вероятности P_{ij} удовлетворяют следующим двум условиям:

1) условию неотрицательности $P_{ij} \ge 0,\ \forall i, j$;

2) условию нормировки $\sum_{j=0}^{\infty} P_{ij} = 1,\ \forall i$.

Второе условие отражает тот факт, что каждое испытание вызывает некоторый переход из состояния в состояние. Для удобства говорят о переходе и в том случае, когда состояние остается неизменным.

Марковский процесс полностью определен, если заданы переходные вероятности и распределение вероятностей начальных состояний:

$$\pi_i = P\{x_0 = i\}.$$

Большое число физических, биологических и экономических явлений описываются марковскими цепями.

Скрытый марковский процесс представляет собой математическую модель, являющуюся двухкомпонентным случайным процессом (X, Y) со скрытой компонентой X и наблюдаемой компонентой Y. Случайный процесс X является марковским случайным процессом. СММ является частным случаем скрытого марковского процесса в случае, когда процесс X – это марковская цепь с конечным множеством состояний, определяемая матрицей переходных вероятностей. Текущее состояние марковской цепи $x_t \in N^* = \{1, 2, ..., N\}$ интерпретируется как скрытое состояние источника данных (исследуемого явления), а моменты измерения состояния – как дискретные события в ходе развития исследуемого явления. Последовательность скрытых состояний, моделируемая такой цепью (т. е. реализация скрытого процесса), обозначается как $Q = \{q_1, q_2, ..., q_T\}$, где T – длина наблюдаемой последовательности. Случайный процесс Y является вещественнозначным ($y_t \in \mathbb{R}^L$, где $L \ge 1$ – размерность пространства, $t = \overline{1,T}$) марковским случайным векторным процессом конечного порядка (т. е. плотность вероятностей перехода в очередное его состояние зависит не от одного, а от нескольких предшествующих состояний) со скачкообразно изменяющимися вероятностными свойствами. Последовательность $O=\{o_1, o_2, ..., o_T\}$ является последовательностью наблюдений (т. е. это реализация наблюдаемого процесса). Случайный процесс Y связан со скрытым процессом X условной вероятностью: $P\{o_t | q_t = i\}$. Порождающий характер СММ определяет то обстоятельство, что вероятностные свойства модели задаются сначала маргинальными вероятностными свойствами первичного процесса X и только затем условными вероятностными свойствами второго процесса $(Y \mid X)$. Именно это обстоятельство позволяет исследователю выразить свое априорное представление о скрытом механизме исследуемого явления и его внешних проявлениях.

СММ полностью задают следующие параметры:

1) вектор вероятностей начальных состояний $\Pi=\{\pi_i\}$, $i=\overline{1,N}$, где $\pi_i = P\{q_1 = s_i\}$, множество скрытых состояний $S=\{s_1, s_2, ..., s_N\}$, N – количество скрытых состояний в модели;

2) матрица вероятностей переходов $A=\{a_{ij}\}$, $i, j=\overline{1,N}$, где $a_{ij} = P\{q_t = s_j | q_{t-1} = s_i\}$;

3) функции условной плотности распределений наблюдений $B=\{b_i(t)\}$, где $b_i(t)$ – это условные плотности вероятностей $P\{o_t | q_t = s_i\}$, o_t – наблюдение из последовательности O, фиксируемое в момент $t=\overline{1,T}$. В работе рассматривается случай, когда условные плотности распределений наблюдений являются смесями вероятностных распределений:

$$b_i(t) = \sum_{m=1}^{M_i} \tau_{im} f(o_t; \Theta_{im}),$$

где τ_{im} – это вес m-ой компоненты смеси в описании i-ого скрытого состояния, M_i – количество компонент смеси, $M_i < \infty$, $i=\overline{1,N}$, f – функции плотности вероятности некоторого закона распределения. Следовательно, распределение наблюдений в каждом скрытом состоянии описывается смесью своих M_i распределений с параметрами Θ_{im} Чаще всего полагают, что $M_i = M$, $i=\overline{1,N}$. СММ с такими характеристиками называются скрытыми марковскими моделями с непрерывным пространством наблюдений. Если функции плотности распределения наблюдений описываются смесью нормальных распределений, а сами наблюдения являются одномерными, то, функция плотности распределения имеет вид:

$$f(o_t; \Theta_{ij}) = \left(\sqrt{2\pi}\sigma_{ij}\right)^{-1} e^{-\left(o_t - \mu_{ij}\right)^2 / 2\sigma_{ij}^2},$$

где параметры μ_{ij} и σ_{ij}^2 являются, соответственно, математическим ожиданием и дисперсией j-ой компоненты смеси в описании i-ого скрытого состояния $i=\overline{1,N}$, $j=\overline{1,M}$.

В данной лабораторной работе мы будем полагать, что $B=\left\{b_{ij}\right\}$ является матрицей дискретных вероятностей, то есть: $b_{ij}=P\left\{o_t=v_j \middle| q_t=s_i\right\}$, где множество наблюдаемых состояний (наблюдений) $V=\{v_1,v_2,...,v_M\}$, M – количество наблюдаемых состояний в модели (алфавит). Такие СММ называются скрытыми марковскими моделями с дискретным пространством наблюдений.

Таким образом, СММ задается ненаблюдаемой (скрытой) марковской цепью, функциями плотности распределениями наблюдений и вероятностями начальных состояний: $\lambda=\left(A,B,\pi\right)$.

Для проведения вычислительных экспериментов необходимо смоделировать множество обучающих и тестовых последовательностей. Ниже приведена схема моделирования для случая, когда функция плотности условного распределения наблюдений описывается одним распределением, т. е. имеет смесь, состоящую из одной компоненты.

Моделируется случайное число $\xi_1 \succ RAV[0,1]$, далее выясняется, к какому из интервалов, на которые условно можно разбить вероятности из вектора π: $[0;\pi_1]\cup(\pi_1;\pi_1+\pi_2]\cup\ldots\left(\sum_{k=1}^{N-1}\pi_k;1\right]$, принадлежит значение ξ_1. Начальное состояние q_1 выбирается равным номеру интервала, в который попало значение ξ_1. Для остальных моментов t значения $\xi_t \succ RAV[0,1]$ сопоставляется с границами интервалов, на которые условно можно разбить вероятности из i-ой строки матрицы: $\left[0;a_{i1}\right]\cup\left(a_{i1};a_{i1}+a_{i2}\right]\cup\ldots\cup\left(\sum_{k=1}^{N-1}a_{ik};1\right]$, $i=\overline{1,N}$ при условии, что в предыдущий момент $t-1$ система находилась в i-ом состоянии $t=\overline{2,T}$.

После того, как таким образом была сформирована последовательность скрытых состояний $Q=\left\{q_1,q_2,\ldots,q_T\right\}$, моделируется последовательность наблюдений по матрице B таким же образом, как описано выше для матрицы A.

Задание

1. Смоделировать последовательность скрытых состояний длиной *T=100*
2. Смоделировать последовательность наблюдаемых состояний длиной *T=100 по* последовательности скрытых состояний
3. Представить полученную наблюдаемую последовательность в графическом виде.
4. Привести графики достигнутой точности по параметрам (ρ_A , ρ_B) в зависимости от *T* при *K*=100 и от *K* при *T*=100.

5. Подготовить по 2 набора (каждый набор включает по К=100) наблюдаемых последовательностей

Варианты заданий

Вектор начальных состояний выглядит следующим образом: $\pi = (\pi_1 \quad \pi_2 \dots \quad \pi_N)$, $\pi_1 = 1, \pi_i = 0, i = \overline{2, N}$.

Вариант	*Алфавит: V*	*Матрица переходных вероятностей A*	*Матрица эмиссей B*
1	$\{a, b\}$	$\begin{pmatrix} 0,5 & 0,5 & 0 \\ 0 & 0,5 & 0,5 \\ 0,5 & 0 & 0,5 \end{pmatrix}$	$\begin{pmatrix} 0 & 1 \\ 1 & 0 \\ 0,5 & 0,5 \end{pmatrix}$
2	$\{0,1\}$	$\begin{pmatrix} 0,3 & 0,3 & 0,4 \\ 0 & 0,5 & 0,5 \\ 0,5 & 0 & 0,5 \end{pmatrix}$	$\begin{pmatrix} 0,2 & 0,8 \\ 0,8 & 0,2 \\ 0,5 & 0,5 \end{pmatrix}$
3	$\{-1,1\}$	$\begin{pmatrix} 0,7 & 0,3 & 0 \\ 0 & 0,7 & 0,3 \\ 0,9 & 0 & 0,1 \end{pmatrix}$	$\begin{pmatrix} 0,2 & 0,8 \\ 0,9 & 0,1 \\ 0 & 1 \end{pmatrix}$
4	$\{a, b, c\}$	$\begin{pmatrix} 0,5 & 0,5 & 0 \\ 0 & 0,5 & 0,5 \\ 0,5 & 0 & 0,5 \end{pmatrix}$	$\begin{pmatrix} 0,5 & 0,4 & 0,1 \\ 0 & 0,5 & 0,5 \\ 0,4 & 0,1 & 0,5 \end{pmatrix}$
5	$\{-1,0,1\}$	$\begin{pmatrix} 0,3 & 0,3 & 0,4 \\ 0 & 0,5 & 0,5 \\ 0,5 & 0 & 0,5 \end{pmatrix}$	$\begin{pmatrix} 0,5 & 0,5 & 0 \\ 0 & 0,5 & 0,5 \\ 0,5 & 0 & 0,5 \end{pmatrix}$
6	$\{a, b, c\}$	$\begin{pmatrix} 0,5 & 0,4 & 0,1 \\ 0 & 0,5 & 0,5 \\ 0,4 & 0,1 & 0,5 \end{pmatrix}$	$\begin{pmatrix} 0,5 & 0,5 & 0 \\ 0 & 0,5 & 0,5 \\ 0,5 & 0 & 0,5 \end{pmatrix}$
7	$\{-1,0,1\}$	$\begin{pmatrix} 0,5 & 0,5 & 0 \\ 0 & 0,5 & 0,5 \\ 0,5 & 0 & 0,5 \end{pmatrix}$	$\begin{pmatrix} 0,3 & 0,3 & 0,4 \\ 0 & 0,5 & 0,5 \\ 0,5 & 0 & 0,5 \end{pmatrix}$
8	$\{a, b, c\}$	$\begin{pmatrix} 0,1 & 0,9 \\ 0,9 & 0,1 \end{pmatrix}$	$\begin{pmatrix} 0,1 & 0,1 & 0,8 \\ 0,8 & 0,1 & 0,1 \end{pmatrix}$
9	$\{-1,0,1\}$	$\begin{pmatrix} 0,5 & 0,5 \\ 0,9 & 0,1 \end{pmatrix}$	$\begin{pmatrix} 0,8 & 0,1 & 0,1 \\ 0,1 & 0,8 & 0,1 \end{pmatrix}$
10	$\{a, b, c, d\}$	$\begin{pmatrix} 0,5 & 0,5 \\ 0,9 & 0,1 \end{pmatrix}$	$\begin{pmatrix} 0,5 & 0,4 & 0,1 & 0 \\ 0 & 0,5 & 0,4 & 0,1 \end{pmatrix}$
11	$\{-1,0,1,2\}$	$\begin{pmatrix} 0,5 & 0,5 \\ 0,9 & 0,1 \end{pmatrix}$	$\begin{pmatrix} 0,1 & 0,4 & 0,1 & 0,5 \\ 0,1 & 0,4 & 0,4 & 0,1 \end{pmatrix}$

Контрольные вопросы

1. Дайте определение СММ.
2. Что такое СММ с непрерывным пространством наблюдений? Приведите пример.
3. Что такое СММ с дискретным пространством наблюдений? Приведите пример.
4. Как происходит процесс формирования последовательности скрытых состояний?
5. Как происходит процесс формирования последовательности наблюдаемых состояний?