Практика 9. Как формализован переход от эмпирических данных двумерной таблицы к матрице переходных вероятностей?

студентка 2 курса магистратуры Кобзева В.М.

21 декабря 2020 г.

Как формализован переход от эмпирических данных двумерной таблицы к матрице переходных вероятностей?

Определение.

Последовательность дискретных случайных величин $\{X_n\}_{n\geq 0}$ называется **простой** (то есть имеет конечное число возможных состояний – S) **цепью Маркова** (с дискретным временем), если

$$\mathbb{P}(X_{n+1} = i_{n+1}|X_n = i_n, X_{n-1} =$$

$$= i_{n-1}, ..., X_0 = i_0) = \mathbb{P}(X_{n+1} = i_{n+1}|X_n = i_n).$$

Таким образом, в простейшем случае условное распределение последовательности состояний цепи Маркова зависит только от текущего состояния и не зависит от всех предыдущих состояний.

Матрица $P(n)$, где $P_{ij}(n) = \mathbb{P}(X_{n+1} = j|X_n = i)$ называется матрицей **переходных вероятностей** на $n-$м шаге, а вектор $p = (p_1, p_2, ..., p_s)^T$, где $p_i = \mathbb{P}(X_0 = i)$ – начальное распределение цепи Маркова.

Как формализован переход от эмпирических данных двумерной таблицы к матрице переходных вероятностей?

Матрица переходных состояний является стохастической, то есть

$$\sum_{j\in S} P_{ij} = 1, \quad \forall n \in \mathbb{N}, \quad \text{сумма элементов матрицы по строке равна 1.}$$

Определение.

Цепь Маркова называется **однородной**, если матрица переходных вероятностей не зависит от номера шага (матрица перехода постоянная), то есть $P_{ij}(n) = P_{ij}, \quad \forall n \in \mathbb{N}$.

Определение.

Цепь Маркова является **не поглощающей**, если в ней нет поглощающих состояний.
Состояние i называется **поглощающим**, если $P_{ii} = 1$.

Как формализован переход от эмпирических данных двумерной таблицы к матрице переходных вероятностей?

Основные требования к матрице переходных вероятностей:

1. Как уже писалось выше $\sum_{j \in S} P_{ij} = 1, \quad \forall n \in \mathbb{N}$;
2. Так как элементами матрицы являются условные вероятности перехода из одного состояния в другое, то $0 \leq p_{ij} \leq 1, \forall i, j \in S$.
3. Вероятность перехода в состояние j зависит только от текущего состояния i.

Как формализован переход от эмпирических данных двумерной таблицы к матрице переходных вероятностей?

Алгоритм перехода от эмпирических данных двумерной таблицы к матрице переходных вероятностей.

Пусть $t_0, t_1, ..., t_n$ – точки дискретного времени, такие что $\Delta t_k = t_{k+1} - t_k = const$.

Пусть $N(\Delta t_k)$ – двумерная таблица с эмпирическими данными, где каждый элемент $n_{ij}(\Delta t_k)$ – число наблюдений, в которых произошел переход из состояния i в состояние j за время Δt_k. Обозначим $n_i(t_k)$ – число наблюдений, где объект находился в состоянии i в момент времени t_k.

Тогда элементы $p_{ij}(t_k)$ матрицы переходных вероятностей P имеют вид:

$$p_{ij}(t_k) = \frac{n_{ij}(\Delta t_k)}{n_i(t_k)}.$$

Из свойств матрицы переходных вероятностей знаем, что

$$p_{ii} = 1 - \sum_{j \neq i}, \quad \forall i$$