1 Цепи Маркова

1.1 Марковская зависимость

Пусть G - эксперимент с конечным числом исходов $\{E_1, \ldots, E_n\}$. Предположим, что мы неограниченно повторяем эксперимент G. Рассмотрим, связанную с этим экспериментом, последовательность целочисленных случайных величин $\{X_m\}_{m=0}^{\infty}$. Будем считать что, если в m-м испытании произошло событие E_j, то $X_m = j$. Эта последовательность случайных величин образует цепь Маркова, если

$$\mathbf{P}(X_l = j | X_0 = k_0, \ldots, X_{l-2} = k_{l-2}, X_{l-1} = i) = \mathbf{P}(X_l = j | X_{l-1} = i) \doteq p_{ij}^{(l)}.$$

То есть, цепь Маркова можно представить себе как некоторую систему с возможными состояниями $\{E_1, \ldots, E_n\}$. Задается некое распределение начального состояния - величины X_0:

$$\mathbf{P}(X_0 = j) = p_j(0),\ j = 1, 2, \ldots, n, \quad \sum_{j=1}^{n} p_j(0) = 1. \tag{1.1}$$

В целочисленные моменты времени система меняет свое состояние, при этом вероятность в момент времени l попасть в состояние E_j (то есть, $\mathbf{P}(X_l = j)$) зависит только от того, в каком состоянии система находилась в момент $l-1$, не зависит от всех предыдущих состояний системы. Часто это свойство характеризуют словами: *при фиксированном настоящем будущее не зависит от прошлого.*

Определение 1.1 Цепь Маркова $\{X_m\}_{m=0}^{\infty}$ называется *однородной,* если $p_{ij}^{(l)}$ – вероятности перехода из состояния i в состояния j не зависят от l – момента времени перехода.

В дальнейшем мы будем рассматривать только однородные марковские цепи и обозначать вероятности перехода из состояния i в состояния j через p_{ij}.

1.2 Переходные вероятности

Переходные вероятности можно записать при помощи переходной матрицы

$$P = \begin{pmatrix} p_{11} & p_{12} & \cdots & p_{1n} \\ p_{21} & p_{22} & \cdots & p_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ p_{n1} & p_{n2} & \cdots & p_{nn} \end{pmatrix}.$$

Элементы переходной матрицы удовлетворяют двум свойствам:

$$1) \qquad p_{ij} \geq 0 \quad \text{для всех } i,\ j, \tag{1.2}$$

$$2) \qquad \sum_{j=1}^{n} p_{ij} = 1 \quad \text{для любого } i. \tag{1.3}$$

Свойство (1.2) следует из того, что величины p_{ij} являются вероятностями некоторых событий; свойство (1.3) вытекает из того, что на любом шаге с вероятностью единица из состояния E_i осуществляется переход в одно и только в одно из состояний $\{E_1, \ldots, E_n\}$.

Определение 1.2 Матрица P, удовлетворяющая свойствам (1.2), (1.3), называется *стохастической.*

Таким образом, мы получили, что переходная матрица P является *стохастической.* Ясно, что любая *стохастическая* матрица может быть переходной матрицей некоторой цепи Маркова. Матрица P полностью описывает изменение состояния системы за один шаг. Опишем изменение состояния системы за k шагов. Обозначим $p_{ij}(k) = \mathbf{P}(X_k = j | X_0 = i)$. Пусть $P(k) = ||p_{ij}(k)||$ – матрица переходов из состояния i в состояние j за k шагов.

Утверждение 1.1 *Матрица переходов за k шагов есть к-ая степень матрицы P, то есть, $P(k) = P^k$.*

Доказательство.
По формуле полной вероятности при $k > 1$ имеем

$$p_{ij}(k) = \sum_{l=1}^{n} \mathbf{P}(X_{k-1} = l | X_i = i) p_{lj} = \sum_{l=1}^{n} p_{il}(k-1) p_{lj}. \tag{1.4}$$

То есть,

$$P(k) = P(k-1) \cdot P.$$

Следовательно, по индукции получаем

$$P(k) = P^k.$$

Легко проверить, что любая степень стохастической матрицы снова является стохастической матрицей. $\square$

Заметим, что, если задано и начальное распределение вероятностей (1.1), то, аналогично, по формуле полной вероятности можно получить распределение $p_i(t),\ i = 1, \ldots, n$, вероятностей состояний в момент времени t:

$$\mathbf{P}(X_t = i) = p_i(t) = \sum_{l=1}^{n} p_l(0) p_{li}(t),\ j = 1, 2, \ldots, n. \tag{1.5}$$

1.3 Предельные вероятности

Определение 1.3 Цепь Маркова, для которой при любом $j = 1, \ldots, n$ существуют *предельные вероятности* $p_j = \lim_{t\to\infty} p_{ij}(t)$, не зависящие от начального состояния, называется *эргодической.*

Утверждение 1.2 *Если цепь Маркова является эргодической, то система уравнений*

$$x_j = \sum_{l=1}^{n} x_l p_{lj}, \quad j = 1, \ldots, n, \qquad \sum_{j=1}^{n} x_j = 1 \tag{1.6}$$

имеет единственное решение $x_j = p_j$, *где* p_j *– предельные вероятности из определения* (1.3).

Доказательство. Для того, чтобы доказать, что величины p_j являются решением системы уравнений (1.6) достаточно перейти к пределу $k \to \infty$ в равенстве (1.4) и в условии $\sum_{j=1}^{n} p_{ij}(k) = 1$.

Пусть теперь x_j, $j = 1, \ldots, n$ является решением системы (1.6). В этом случае, используя индукцию, легко видеть, что при любом натуральном m имеем

$$x_j = \sum_{k=1}^{n} x_k p_{kj}(m), \quad j = 1, \ldots, n. \tag{1.7}$$

При $m = 1$ это следует из (1.6), так как $p_{lj}(1) = p_{lj}$. Из предположения, что (1.7) справедливо при $m = l$ имеем

$$x_j = \sum_{k=1}^{n} \sum_{i=1}^{n} x_j p_{ik} p_{ki}(l) = \sum_{i=1}^{n} x_i \sum_{k=1}^{n} p_{ik} p_{kj}(l) = \sum_{i=1}^{n} x_i p_{ij}(l+1).$$

Следовательно, переходя в (1.7) к пределу $m \to \infty$ и, используя равенство $\sum_{j=1}^{n} x_j = 1$ получаем, что $x_j = p_j$. $\square$

Утверждение 1.3 *Если все переходные вероятности матрицы P положительны (строго больше нуля), то соответствующая цепь Маркова является эргодической.*

Доказательство. Обозначим

$$M_j(t) = \max_{1\le i\le n} p_{ij}(t), \quad m_j(t) = \min_{1\le i\le n} p_{ij}(t).$$

Имеем

$$p_{ij}(t+1) = \sum_{k} p_{ij} p_{kj}(t), \quad m_j(t) \le p_{kj}(t) \le M_j(t). \tag{1.8}$$

Следовательно, при всех i

$$m_j(t) = m_j(t) \sum_k p_{ij} \le p_{ij}(t+1) \le M_j(t) \sum_k p_{ij} = M_j(t).$$

При некотором $i = k$ достигается $\max_{1 \le i \le n} p_{ij}(t+1) = p_{kj}(t+1) = M_j(t+1)$ и при некотором $i = l$ достигается $\min_{1 \le i \le n} p_{ij}(t+1) = p_{lj}(t+1) = m_j(t+1)$. Отсюда получаем

$$m_j(t) \le m_j(t+1) \le M_j(t+1) \le M_j(t).$$

То есть последовательность $m_j(t)$ возрастает и ограничена сверху, а последовательность $_j(t)$ убывает и ограничена снизу и существуют пределы $\lim_{t\to\infty} M_j(t) = M_j$ и $\lim_{t\to\infty} m_j(t) = m_j$. Для доказательства утверждения достаточно показать, что при $t \to \infty$ существует $\lim(M_j(t) - m_j(t)) = 0$. Воспользуемся первым равенством (1.8), будем иметь

$$M_j(t+1) - m_j(t+1) = p_{kj}(t+1) - p_{lj}(t+1) = \sum_{m=1}^{n} (p_{km} - p_{lm}) p_{mj}(t) =$$

$$= \ \ (\sum\nolimits_m^+ p_{km} - p_{lm}) p_{mj}(t) + \sum\nolimits_m^- (p_{km} - p_{lm}) p_{mj}(t),$$

где $\sum^+$ - сумма неотрицательных слагаемых, а $\sum^-$ - сумма отрицательных слагаемых. При этом $m_{mj}(t) \le p_{mj}(t) \le M_j(t)$, следовательно

$$M_j(t+1) - m_j(t+1) \le M_j(t) \sum\nolimits_m^+ (p_{km} - p_{lm}) + m_j(t) \sum\nolimits_m^- p_{km} - p_{lm}). \qquad (1.9)$$