Теория вероятностей. Лекция двадцать четвертая Марковские цепи

Дмитрий Валерьевич Хлопин
glukanat@mail.ru

Институт математики и механики им. Н.Н.Красовского

17.04.2019

Общее определение случайного процесса

Случайным процессом называют семейство случайных величин X_t $(t \in T)$, определенных на одном и том же вероятностном пространстве $(\Omega, \mathcal{F}, \mathbb{P})$, где $T \subset \mathbb{R}$; при этом предполагают, что измеримо отображение $(\omega, t) \to X_t(\omega)$ относительно $\mathcal{F} \otimes \mathcal{B}(T)$. Индекс t будем называть временем. В качестве множества T обычно выбирают либо $\mathbb{N} \cup \{0\}$, либо $\mathbb{Z}$, либо $\mathbb{R}^+ = [0, +\infty)$, либо $\mathbb{R}$.

Вообще говоря, T может быть, например, подмножеством многомерного пространства, тогда его называют случайным полем. Мы будем рассматривать только процессы: случай $T \subset \mathbb{R}$.

Подумать: даже для $T = \mathbb{R}$ множество $\{\omega \,|\, X_t \text{ ограничено}\}$ не обязано быть событием без дополнительных предположений на $\mathcal{F}$.

Рассмотрим для всякого набора $(t_1, \ldots, t_N) \in T^{\mathbb{N}}$ совместные распределения векторов $(X_{t_1}, \ldots, X_{t_N})$:

$$F_{t_1,\ldots,t_N}(x_1, \ldots, x_N) = \mathbb{P}(X_{t_1} \le x_1, \ldots, X_{t_N} \le x_N) \qquad \forall (x_1, \ldots, x_N) \in \mathbb{R}^{\mathbb{N}}.$$

1. $F_{t_1,\ldots,t_N}$ монотонна по каждой переменной, непрерывна справа, принимает значения из $[0,1]$, $F_{t_1,\ldots,t_N}(+\infty, \ldots, +\infty) = 1$ и $F_{t_1,\ldots,t_{i-1},t_i,t_{i+1},\ldots,t_N}(x_1, \ldots, x_{i-1}, -\infty, x_{i+1}, \ldots, x_N) = 0$;
2. выполнено условие согласованности:

$$F_{t_1,\ldots,t_{i-1},t_i,t_{i+1},\ldots,t_N}(x_1, \ldots, x_{i-1}, +\infty, x_{i+1}, \ldots, x_N) \\ = F_{t_1,\ldots,t_{i-1},t_{i+1},\ldots,t_N}(x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_N).$$

Теорема Колмогорова о построении случайных процессов [Коралов, Синай, Т.12.8] По любому семейству функций $F_{t_1,\ldots,t_N}$, удовлетворяющему этим двум пунктам, можно построить как минимум один случайный процесс с такими функциями распределения.

Марковские цепи

Будем говорить, что случайный процесс X_t — марковская цепь с дискретным временем, если

- (время дискретно) t принимает только целые неотрицательные значения;
- (счетное число состояний) множество значений X_t не более чем счетно;
- (марковость) вероятность состояния X_{t+1} зависит лишь от состояния X_t, т.е.

$$\mathbb{P}(X_{t+1}|X_t, X_{t-i_1}, \dots, X_{t-i_n}) = \mathbb{P}(X_{t+1}|X_t).$$

Отметим, что определение марковости верно почти всюду, в частности, нас не интересует значение $\mathbb{P}(X_{t+1} = x|X_t = y)$ при $\mathbb{P}(X_t = y) = 0$.

Стохастические матрицы

Под матрицей будем далее понимать отображение из $\mathbb{N} \times \mathbb{N}$ в $\mathbb{R}$. Операции над ними стандартные. Более того, никто не мешает считать, что обычные конечномерные матрицы также являются такими объектами.

Будем говорить, что матрица $Q = (q_{ij})_{i,j=1,2,\dots}$ является стохастической, если

- $q_{ij} \geq 0$;
- $\sum_{j\in\mathbb{N}} q_{ij} = 1$.

Будем говорить, что строка $(p_j)_{j=1,2,\dots}$ является распределением, если

- $p_j \geq 0$;
- $\sum_{j\in\mathbb{N}} p_j = 1$.

Легко проверить, что если p — распределение, Q', Q'' – стохастические матрицы, то $Q = Q'Q''$ — стохастическая матрица, pQ' — распределение.

Матрицы перехода

Теперь можно считать, что X_t отображается в $\mathbb{N}$ (состояния — натуральные числа). Тогда F_{X_t} можно задать бесконечной строкой $\mu_t \overset{\triangle}{=} (\mathbb{P}(X_t = 1), \ldots, \mathbb{P}(X_t = k), \ldots)$.

Подумать: каждая такая строка — распределение.

Матрицы $Q^{(k)} = (q_{ij}^{(k)})_{i,j=1,\ldots,n} \overset{\triangle}{=} \big(\mathbb{P}(X_k = j | X_{k-1} = i)\big)_{i,j=1,\ldots,n}$, $k \in \mathbb{N}$, называют матрицами переходов (матрицами вероятностей переходов, матрицами переходных вероятностей).

Подумать: каждая такая матрица — стохастическая.

Уравнение Колмогорова

Теорема. (Колмогоров) [С-но] В марковской цепи с матрицами переходов $Q^{(k)}$ для $\mu_t \overset{\triangle}{=} (\mathbb{P}(X_t = 1), \ldots, \mathbb{P}(X_t = k), \ldots)$ выполнено

$$\mu_1 = \mu_0 Q^{(1)},\ \mu_k = \mu_{k-1} Q^{(k)},\ \mu_k = \mu_0 Q^{(1)} \ldots Q^{(k)} \quad \forall k \in \mathbb{N}.$$

Подумать : матрицы переходов однозначно восстанавливают распределения X_k по μ_0. Однозначно ли определяются матрицы переходов (1 балл)?.

Подумать: верно ли, что всяким распределению μ_0 и последовательности стохастических матриц $Q^{(k)}$ соответствует марковская цепь с такими матрицами переходов? Такая марковская цепь единственна?

Стационарные марковские цепи. Стационарное распределение

Цепь Маркова называется стационарной (однородной по времени), если $Q^{(k)} \equiv Q$, то есть если соответствующие условные вероятности не зависят от времени.

Распределение μ называется стационарным (для цепи Маркова с матрицей переходов Q), если $\mu Q = \mu$.

Теорема. (еще раньше доказали) Всякая цепь Маркова с конечным числом состояний имеет хотя бы одно стационарное распределение.

Подумать: условие конечности в этой теореме существенно.

Подумать: единственность стационарного распределения, вообще говоря, не утверждалась.

Эргодичность

Всюду далее вновь считаем, что число состояний конечно, матрицы и распределения имеют конечное число элементов.

Определение. Стохастическая матрица $Q = (q_{ij})_{i,j=1,2,\ldots r}$ называется эргодической, если все её элементы положительны.

Теорема. Пусть матрица переходов Q эргодична. Тогда найдется такая строка μ_*, что $\mu_* Q = \mu_*$ и распределения $\mu_0 Q^n$ сходятся к μ_* для любого начального распределения μ_0; в частности, других стационарных распределений, помимо μ_*, у нее нет.

Для доказательства потребуется почти очевидный:

Принцип Банаха. Если в полном метрическом пространстве $\mathbb{Y}$ с метрикой d, оператор $A : \mathbb{Y} \to \mathbb{Y}$ сжимающий (для некоторого $\beta \in (0,1)$ $d(Ax_1, Ax_2) \leq \beta d(x_1, x_2)$), то существует единственный элемент $x_* \in \mathbb{Y}$ такой, что $Ax_* = x_*$, причем $d(A^k x, x_*) \to 0$ при $k \to \infty$, более того $d(A^k x, x_*) \leq \beta^k d(x, x_*)$.