Оглавление

Глава 2. Представление СМО марковским случайным процессом

I. Классификация СМО по Кендалл

Обозначения, введенные Кендаллом для СМО:

D – детерминированное или регулярное;

M – показательное (пуассоновский поток, exp интервал между двумя заявками);

G – никаких специальных предположений о виде;

E_n – χ^2_{2n};

B(ξ) – функция распределения времени обслуживания ($0 \le \xi \le \infty$).

$B(\xi) = \begin{cases} 1, \xi \ge \frac{1}{\mu} \\ 0, \xi < \frac{1}{\mu}, \end{cases}$ длительность обслуживания – детерминирован. или регулярн.

$B(\xi) = \begin{cases} 1 - e^{-\mu\xi}, \xi \ge 0 \\ 0, \xi < 0 \end{cases}$, длительность имеет показательное распределение,

где $\frac{1}{\mu}$ – среднее время обслуживания.

Функция распределения, введенная Эрлангом (обощает два случая сверху):

$dB(\xi) = \frac{(\mu n)^n}{n!}\xi^{n-1}e^{-\mu n\xi}d\xi$, т.е. время обслуживания распределено по λ_{2n}^2. Здесь $\frac{(\mu n)^n}{n!}\xi^{n-1}e^{-\mu n\xi}$ – нормированное распределение Эрланга n-ного порядка.

В этом случае можно считать, что заявка проходит через n фаз или стадий обслуживания. При n = 1 вычисления происходят по формуле (3), при n→∞ по формуле (2).

G – никаких специальных предположений о виде B(ξ) не сделано.

GI – входной поток является рекуррентным, т.е. случайные величины τ_n (интервал между наступлениями событий) статистически независимы и имеют одинаковое распределение: $A(\tau) = \begin{cases} 1 - e^{-\lambda\tau}, \tau \geq 0 \\ 0, \tau < 0 \end{cases}$.

С помощью этих символов можно задать СМО, располагая их в определенном порядке.

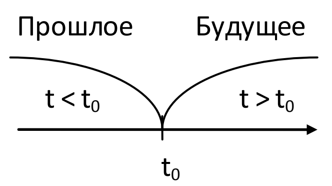
Если в классификации Кендалла на первых двух местах стоят буквы M или E_n, то эта СМО относится к так называемым марковским системам.

Характерным отличием марковских систем является то, что их функционирование может быть описано марковским процессом с непрерывным временем и дискретным множеством состояний, а иногда даже процессом гибели-размножения. Это позволяет использовать для их изучения хорошо разработанные математические методы.

II. Марковский случайный процесс

Пусть X(t) – случаный процесс с непрерывным t (t ≥ 0) и дискретным множеством состояний $S = \{s_1, s_2, \dots, s_m\}$ – конечное множество, $S = \{s_1, s_2, \dots\}$ – счетное множество. (S тоже непрерывно в общем случае.)

Процесс X(t) называется марковским, если для любых наборов моментов времени: $0 \le t_1 < t_2 < \dots < t_{n+1}$ и состояний $s_1, s_2, \dots, s_{n+1} \in S$ выполняется равенство: $P\{x(t_{n+1}) = s_{n+1} | x(t_1) = s_1, x(t_2) = s_2, \dots, x(t_n) = s_n\} = P\{x(t_{n+1}) = s_{n+1} | x(t_n) = s_n\}$ (марковаость первого порядка). Т.е. марковский случайный процесс – это такой случайный процесс, эволюция которого (т.е. развитие, в каких состояниях он пребывает в моменты времени t) после любого фиксированного момента времени t (т.е. в будущем) и до момента времени t (т.е. в прошлом) является независимой при известном состоянии в момент времени t (в настоящем). Это свойство называется марковским.



Марковский процесс (МП) называется однородным (по времени), если $P\{x(t + \tau) = j | x(\tau) = i\}, i, j \in S$ не зависит от τ. В дальнейшем всегда будем рассматривать однородные процессы.

МП (по характеру реализации случайных процессов):

1. цепи Маркова (дискретная последовательность) (чаще всего используется в СМО);
2. случайная (марковская) последовательность;
3. дискретный случайный процесс (дискретная МП) (Марковские цепи с непрерывным t);
4. непрерывнозначный случайный процесс (непрерывнозначный Марковский процесс).

III. Марковские цепи и дискретные марковские процессы

Цепи Маркова – это Марковские случайные процессы с дискретным множеством значений (состояний цепи) $S_1, S_2, \dots, S_m$, и значений аргумента t_0, t_1, ... , t ... Если множество m конечно, то это конечная цепь. Такая цепь Маркова задается множеством значений $\{S_1, \dots, S_m\}$ и следующими вероятностями: (вместо t_0, t_1, ... будем писать 0,1, ...)

1. Начальная вероятность: $\pi_i(0) = P(x(0) = S_i)$, $\sum_i \pi_i = 1$.

2. Вероятность перехода из одного состояния в другое за один шаг: $a_{ij} = P(x(n+1) = S_j | x(n) = S_i)$,), $\sum_j a_{ij} = 1$.

   Вероятность перехода за k шагов: $a_{ij} = P(x(n+k) = S_j | x(n) = S_i)$, $\sum_j a_{ij} = 1$.
3. Вероятность состояния на k-том шаге ($t = t_k$): $\pi_i(k) = P(x(k) = S_i)$, $\sum_i \pi_i = 1$.

В случае, если множество временных отсчетов T – непрерывное, а множество состояний S – счетное, то Марковский процесс называют Марковским процессом с дискретным состоянием или дискретным Марковским процессом. Рассматривая Марковский процесс с дискретным множеством состояний и непрерывным временем, подразумевается, что **все переходы системы из состояния в состояние происходят под действием простейших потоков событий** (потоков вызовов, отказов и т.д.). Переход из одного состояния в другое происходит скачком в момент времени, когда реализуется случайное событие: поступление новой заявки, окончание обслуживания заявки и т.д., вызывающее переход системы в новое состояние.

Если все потоки событий, переводящие СМО из состояние в состояние, простейшие, время обслуживания распределено по экспонециальному закону, то процесс, протекающий в СМО, будет Марковским СП с дискретным состоянием.

Это значит, что развитие таких процесов определяется текущим состоянием и не зависит от его предыстории.

Обознчим за $A_к$ событие, состоящее в том, что в момент времени $t \geq 0$ система находится в состоянии S_k, $k = \overline{0,m}$, $P(A_k) = p_k(t)$, где $p_k(t)$ – это вероятность нахождения в k-ом состоянии. Очевидно, что события A_k образуют полную группу событий, следовательно : $\sum_{k=1}^{m} P_k(t) = 1$.

Одна из задач ТМО сводится к нахождению вектора p(t) = ($p_1(t)$, $p_2(t)$,... $p_m(t)$) как функции времени.

- **Пример**

Техническое устройство S состоит из двух узлов, каждый из которых в случайный момент времени может выйти из строя (отказать), после чего мгновенно начинается ремонт узла, продолжающийся заранее неизвестное время.

Состояния системы:

S_0 – оба узла исправны;

S_1 – первый узел ремонтируется, второй исправен;

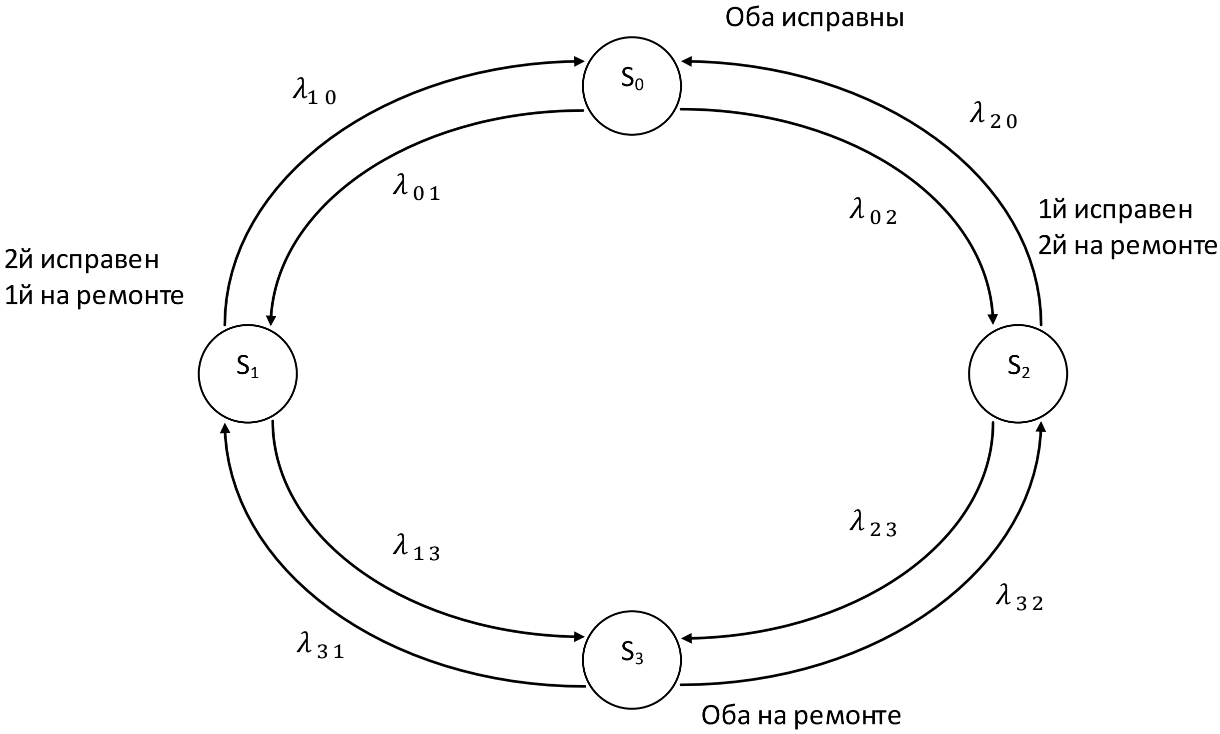
S_2 – второй узел ремонтируется, первый исправен;

S_4 – оба узла ремонтируются.

Будем полагать, что все переходы системы из состояния S_i в S_j происходят под воздействие простейших потоков событий с интенсивностью λ_{ij} $(i,j = \overline{0,3})$.

Переход из S_0 в S_1 будет происходить под воздействием потока отказов первого узла; из S_1 в S_0 - под воздействием потока «окончаний ремонтов» первого узла и т.п.

Граф состояний системы с проставленными у стрелок интенсивностями называется размеченным.



На этом графе отсутствуют стрелкии из S_0 в S_3 и из S_1 в S_2. Это обьясняется тем, что выходы узлов из строя независимыми друг от друга, и, например, вероятность одновременного выхода из строя двух узлов ($S_0 \rightarrow S_3$) или одновременного окончания ремонтов двух узлов ($S_3 \rightarrow S_0$) можно пренебречь.

Рассмотрим систему в момент времени t и, задав малый промежуток Δt, найдем вероятность $P_0(t + \Delta t)$ того, что система в момент времени $(t + \Delta t)$ будет находиться в состояниии S_0. Это достигается разными способами:

1. Либо система в момент времени t с вероятностью $P_0(t)$ находится в состоянии S_0 и за время Δt не вышла из него;
2. Либо система в момент времени t с вероятностью $P_1(t)$ (или $P_2(t)$) находилась в состоянии S_1 (или S_2) и за время Δt перешла в S_0.

(1) Найдем вероятность первого варианта.

Заметим, что для простейшего потока с интенсивностью λ вероятность попадания на элементраный (малый) отрезок времени Δt хотя бы одного события потока равна: $P_{\Delta t} = P(T < \Delta t) = 1 - e^{-\lambda \Delta t} \approx \lambda \Delta t$; $\lambda \Delta t$ – два первых члена ряда Тейлора по степеням Δt.

Вывести систему из сотсяния S_0 можно либо потоком отказов первого узла, либо потоком отказов второго узла, т.е. *суммарным простейшим потоком* (заметим, что при наложении двух простейших потоков получается опять простейший поток) с интенсивностью $\lambda_{01} + \lambda_{02}$, т.е. с вероятностью $P \approx (\lambda_{01} + \lambda_{02})\Delta t$. Тогда вероятность того, что система не выйдет из состояния S_0: $1 - (\lambda_{01} + \lambda_{02})\Delta t$.

Вероятность того, что система будет находиться в состоянии S_0 и не выйдет из него за время Δt, равна: $P_0(t) \cdot (1 - (\lambda_{01} + \lambda_{02})\Delta t)$.

(2) Найдем вероятность второго варианта.

Под действием потока интенсивности λ_{10} (или λ_{20}) система перейдет в состояние S_0 с вероятностью, приближенно равной $\lambda_{10}\Delta t$ (или $\lambda_{20}\Delta t$).

Вероятность того, что система будет находиться в S_0 по этому способу равна $P_1(t)\lambda_{10}\Delta t$ или $P_2(t)\lambda_{20}\Delta t$.

Применяя теорему сложения вероятностей (для попарно несовместимых событий), получим:

$$\underbrace{P_0(t+\Delta t)}_{\text{Вероятность того, что через }\Delta t\text{ система будет в состоянии }S_0} = \underbrace{P_1(t)\lambda_{10}\Delta t}_{\text{Система была в }S_1\text{ и перешла в }S_0} + \underbrace{P_2(t)\lambda_{20}\Delta t}_{\text{Система была в }S_2\text{ и перешла}} + \underbrace{P_0(t)(1-(\lambda_{01}+\lambda_{02})\Delta t)}_{\text{Была в }S_0\text{ и осталась в }S_0}$$

Тогда:

$$\frac{P_0(t+\Delta t)-P_0(t)}{\Delta t} = P_1(t)\lambda_{10} + P_2(t)\lambda_{20} - P_0(t)(\lambda_{01}+\lambda_{02})$$

Переходя к пределу по времени при $\Delta t \to 0$: получим в левой части уравнения производную $P_0'(t)$:

$$P_0'(t) = \lambda_{10}P_1(t) + \lambda_{20}P_2(t) - P_0(t)(\lambda_{01}+\lambda_{02}).$$

Получено дифференциальное уравнение первого порядка.

Далее букву t для обозначения времени будем опускать.

Рассуждая аналогично для других состояний СМО, можно получить систему дифференциальных уравнений Колмогорова для вероятностей состояний:

$$\begin{cases} P_0' = \lambda_{10}P_1 + \lambda_{20}P_2 - (\lambda_{01}+\lambda_{02})P_0 \\ P_1' = \lambda_{01}P_0 + \lambda_{31}P_3 - (\lambda_{10}+\lambda_{13})P_1 \\ P_2' = \lambda_{02}P_0 + \lambda_{32}P_3 - (\lambda_{20}+\lambda_{23})P_2 \\ P_3' = \lambda_{13}P_1 + \lambda_{23}P_2 - (\lambda_{31}+\lambda_{32})P_3 \end{cases} \qquad (1)$$

Сформулируем правило составления уравнений Колмогорова. В левой части стоит производная вероятности i-ого состояния. В правой – сумма произведений вероятности всех состояний, из которых идут стрелки в данное состояние, умноженных на интенсивности соответсвующих потоков событий, минус суммарная интенсивность всех потоков, выводящих систему из данного состояния, умноженная на вероятность i-ого состояния.

В системе (1) независимых уравнений на 1 меньше общего числа уравнений. Поэтому для решения (1) необходимо добавить еще одно уравнение: $\sum_{i=0}^{3} P_i = 1$. Особенность решения дифферинциальных уравнений состоит в том, что требуется задавать так называемые начальные условия, т.е. в данном случае – вероятность состояний системы в начальный момент времени t = 0.

Так, например, (1) естественно решать при условии, что в начальный момент времени оба узла были исправны и система находилась в состоянии S_0, т.е. при начальных условиях $P_0(0) = 1, P_1(0) = P_2(0) = P_3(0) = 0$.

Как решать подобные уравнения? Вообще говоря, СЛДУ с постоянными коэффициентами можно решать аналитически, но это удобно, когда число уравнений не превосходит двух (иногда трех). Если уравнений больше, уравнения Колмогорова дают возможность найти все вероятности состояний как функции времени. Особый интерес представляют вероятности системы $P_i(t)$ в предельном стационарном режиме, т.е. при $t \to \infty$, которые называются предельными (финальными) вероятностями состояний.

В теории случайных процессов доказывается тот факт, что если число состояний системы конечно и из каждого из них можно (за конечное число шагов) перейти в любое другое состояние, то предельные вероятности существуют.

Предельная вероятность состояния S_i имеет четкий смысл: она показывает среднее относительное время пребывания системы в этом состоянии.

Если, например, $P_0 = ½$, то это означает, что в среднем ½ времени система находится в состоянии S_0.

Как же вычислить предельные вероятности?

Очень просто: т.к. предельные вероятности постоянные (не зависят от времени), то, заменяя в уравнениях Колмогорова их производные нулевыми значениями, получим СЛАУ, описывающую стационарный режим. Для системы с графом состояний, приведенным на рисунке выше, такая система имеет вид:

$$\begin{cases} (\lambda_{01} + \lambda_{02})P_0 = \lambda_{10}P_1 + \lambda_{20}P_2 \\ (\lambda_{10} + \lambda_{13})P_1 = \lambda_{01}P_0 + \lambda_{31}P_3 \\ (\lambda_{20} + \lambda_{23})P_2 = \lambda_{02}P_0 + \lambda_{32}P_3 \\ (\lambda_{31} + \lambda_{32})P_3 = \lambda_{13}P_1 + \lambda_{23}P_2 \end{cases} \qquad (2)$$

Эти уравнения однородны (т.е. не имеют свободного члена), и, значит, определяют неизвестные только точностью до произвольного множителя.

Добавим условия нормирования: $\sum_{i=0}^{3} P_i = 1$ и решим систему (2). При этом одно (любое) из уравнений можно отбросить (оно вытекает как следствие из остальных).

<u>Задача №1</u> (к этому примеру).

Найти предельные вероятности для системы при:

$$\begin{cases} \lambda_{01} = 1 \\ \lambda_{02} = 2 \\ \lambda_{10} = 2 \\ \lambda_{13} = 2 \\ \lambda_{20} = 3 \\ \lambda_{23} = 1 \\ \lambda_{31} = 3 \\ \lambda_{32} = 2 \end{cases}$$

Решение:

Из (2) => $\begin{cases} 3P_0 = 2P_1 + 3P_2 \\ 4P_1 = P_0 + 3P_3 \\ 4P_2 = 2P_0 + 2P_3 \\ P_0 + P_1 + P_2 + P_3 = 1 \end{cases}$ => $\begin{cases} P_0 = 0{,}4 \\ P_1 = 0{,}2 \\ P_2 = 0{,}27 \\ P_3 = 0{,}13 \end{cases}$ =>

Т.е. в предельном стационарном режиме система в среднем 40% времени будет находиться в состоянии S_0, 20% в S_1 и т.д.

<u>Задача №2</u>

Найти средний чистый доход от эксплуатации в стационарном режиме системы из примера выше, если известно, что в единицу времени исправная работа первого узла приносит доход в 10$ (денежных единиц), второго узла – в 6$, а их ремонт требует 4$ и 2$ соответственно.

Оценить экономическую эффективность имеющейся возможности уменьшения вдвое среднего времени ремонта каждого из двух узлов, если при этом придется вдвое увеличить затраты на ремонт каждого узла (в единицу времени).

Решение:

Из предыдущей задачи известно, что в среднем первый узел исправно работает долю времени, равную $P_0 + P_2 = 0{,}67$, второй: $P_0 + P_1 = 0{,}6$. В то же время первый узел находится в ремонте в среднем долю времени, равную $P_1 + P_3 = 0{,}33$, а второй: $P_2 + P_3 = 0{,}4$. Поэтому средний чистый доход в единицу времени от эксплуатации системы, т.е. разность между доходом и расходом: $D = 0{,}67 * 10 + 0{,}6 * 6 - 0{,}33 * 4 - 0{,}4 * 2 = 8{,}18\$$.

Уменьшение вдвое среднего времени ремонта каждого из узлов будет означать увеличение вдвое интенсивности потока «окончаний ремонтов» каждого узла.

Таким образом, теперь интенсивность потоков событий будет равна: $\lambda_{10} = 4, \lambda_{20} = 6, \lambda_{31} = 6, \lambda_{32} = 4$. Это следует из того, что $\frac{1}{\lambda} = \mu$. Напомним, что временной интервал между двумя последовательными заявками распределены экспоненциально. $F_\tau(t) = \begin{cases} 1 - e^{-\lambda t}, t \geq 0 \\ 0, t < 0 \end{cases}, \mu = \frac{1}{\lambda}, \sigma^2 = \frac{1}{\lambda^2}$, где μ – математическое ожидание случайной величины, τ – промежуток времени между произвольными двумя событиями в простейшем потоке.

При этом (2) преобразуется следующим образом:

$$\begin{cases} 3P_0 = 4P_1 + 6P_2 \\ 6P_1 = P_0 + 6P_3 \\ 7P_2 = 2P_0 + 4P_3 \\ P_0 + P_1 + P_2 + P_3 = 1 \end{cases} => P_0 = 0{,}6; P_1 = 0{,}15; P_2 = 0{,}2; P_3 = 0{,}05.$$

Теперь затраты на ремонт составляют 8$ и 4$ => $D^* = 0{,}8 * 10 + 0{,}75 * 6 - 0{,}2 * 8 - 0{,}25 * 4 = 9{,}9\$$, т.е. $D^* > D$ на 21%. Таким образом экономическая целесообразность ускорения ремонтов узлов очевидна.