УДК 519.217

А. А. Ибрагимов (Ин-т кибернетики НПО „Кибернетика", Ташкент, Узбекистан)

ОБ ОДНОМ СВОЙСТВЕ РЕГУЛЯРНОЙ ЦЕПИ МАРКОВА

For the transition probability matrix of a regular Markov chain, it is proved that if one subtracts along the row a certain matrix row and then eliminates this row and the corresponding column, then the spectral radius of the matrix obtained should be less than one. This property of the regular Markov chain is used in constructing of iterative process for the solution of the Howard system of equations encountered in controllable Markov chains with one ergodic class and, possibly, transient states.

Доведено, що якщо в матриці ймовірностей переходу регулярного ланцюга Маркова від рядка відняти деякий рядок, а потім видалити цей рядок і відповідний стовпець, то спектральний радіус отриманої таким способом матриці менший за одиницю. Цю властивість регулярного ланцюга Маркова використано при побудові ітераційного процесу для розв'язання системи рівнянь Ховарда, що зустрічається в керованих марковських ланцюгах з одним ергодичним класом і, можливо, з незворотними станами.

1. Постановка задачи и основная теорема. Рассмотрим конечную однородную марковскую цепь $\{\xi_n, n \geq 0\}$ с пространством состояний $S = \{1, 2, \ldots, N\}$ [1, 2].

Цепь Маркова в рассматриваемом случае вполне определяется заданием начальных вероятностей $p_i^0 = \mathrm{P}(\xi_0 = i)$, $i \in S$, и матрицы вероятностей перехода

$$P = \| p_{ij} \|, \quad i, j \in S,$$

где $p_{ij} = \mathrm{P}(\xi_{n+1} = j \mid \xi_n = i)$ — переходная вероятность из i в j за один шаг.

Элементы матрицы P удовлетворяют обычным условиям стохастичности:

$$p_{ij} \geq 0, \qquad \sum_{j \in S} p_{ij} = 1 \qquad \forall\ i, j \in S.$$

Вероятности перехода за n шагов случайной величины ξ_n для любого $\nu \geq 0$ определяются следующим образом:

$$p_{ij}^{(n)} = \mathrm{P}(\xi_{\nu+n} = j \mid \xi_\nu = i), \quad i, j \in S.$$

Заметим, что матрица $\| p_{ij}^{(n)} \|$, $i, j \in S$, является n-й степенью матрицы P, т. е.

$$P^n = \| p_{ij}^{(n)} \|, \quad i, j \in S.$$

Цепь Маркова называется регулярной, если при любых i и j существует предел $\alpha_j = \lim_{n \to \infty} p_{ij}^{(n)}$, который не зависит от первого индекса.

Различают два вида регулярности [2] : цепь Маркова неотрицательно регулярна, если предельные вероятности удовлетворяют условию $\alpha_j \geq 0$, $j \in S$; если $\alpha_j > 0$ для всех $j \in S$, то цепь Маркова положительно регулярна.

Обозначим через $\mathcal{M}_N$ множество всех квадратных матриц порядка N.

Спектральным радиусом $\rho(A)$ матрицы $A \in \mathcal{M}_N$ называют число

$$\rho(A) = \max \{ |\lambda|;\ \lambda \text{ — собственное значение матрицы } A \}.$$

Напомним следующее предложение.

Предложение [1, 2]. *Спектральный радиус $\rho(P)$ стохастической матрицы $P \in \mathcal{M}_N$ равен единице.*

В настоящей работе доказывается следующая теорема.

Теорема 1. *Пусть марковская цепь положительно или неотрицательно регулярна, B_s — матрица размера $(N-1)\times(N-1)$, полученная из матрицы P вычитанием произвольной s-й строки из оставшихся $N-1$ строк и вычеркиванием s-й строки и s-го столбца. Тогда спектральный радиус $\rho(B_s)$ матрицы B_s меньше единицы.*

Идея теоремы 1 возникла из попыток решить систему уравнений Ховарда, встречающуюся в управляемых марковских цепях [3 – 5], итерационным методом, что вполне оправданно, так как эту систему приходится решать многократно в процессе нахождения оптимальной стратегии управления и в качестве начального приближения каждый раз можно использовать точки предыдущих решений.

В настоящей работе на основании теоремы 1 рассматривается возможность применения метода простой итерации [6] при решении системы уравнений Ховарда, предназначенной для управляемых марковских цепей с одним эргодическим классом и, возможно, с невозвратными состояниями.

Приведем два результата из матричного анализа.

Лемма 1. *Пусть $A \in \mathcal{M}_N$. Тогда $A^n \to 0$ при $n \to \infty$ в том и только в том случае, когда $\rho(A) < 1$.*

Доказательство см. в [7, с. 360].

Лемма 2. *Пусть $A \in \mathcal{M}_N$. Тогда если $A^n \to 0$ при $n \to \infty$, то матрица $I - A$ имеет обратную, причем*

$$(I-A)^{-1} = \sum_{n=0}^{\infty} A^n.$$

Доказательство см. в [8, с. 37].

2. Доказательство теоремы 1. Пусть $p_s^{(n)}$ обозначает s-ю строку матрицы $P^{(n)}$. Введем матрицу $p_s^{(n)}$, состоящую из одинаковых строк

$$p_s^{(n)} = \mathbf{1} p_s^{(n)},$$

где $\mathbf{1}$ — N-мерный вектор-столбец, все элементы которого равны 1.

Сначала покажем, что для любого $s \in S$

$$\lim_{n\to\infty} (P - P_s)^n = 0, \tag{1}$$

где $P_s = P_s^{(1)}$.

Имеем

$$PP_s = P_s P_s = P_s,$$

$$P_s P^{n-1} = P_s^{(n)}, \quad n \ge 1.$$

Отсюда следует

$$(P-P_s)^2 = (P-P_s)(P-P_s) = P^2 - P_s^{(2)},$$

$$(P-P_s)^3 = (P-P_s)\left(P^2 - P_s^{(2)}\right) = P^3 - P_s^{(3)}$$

и, вообще, для любого $n \ge 1$

$$(P-P_s)^n = (P-P_s)\left(P^{n-1} - P_s^{(n-1)}\right) = P^n - P_s^{(n)}.$$

Поскольку цепь регулярна, то

$$P^n - P_s^{(n)} \to 0 \quad \text{при} \quad n \to \infty.$$

Итак, соотношение (1) доказано.

Далее для наглядности положим $s = N$. Тогда $B_N = \| p_{ij} - p_{Nj} \|$, $1 \le i, j \le N-1$. Определим вектор-столбец b_N следующим образом:

$$b_N = [p_{iN} - p_{NN}], \quad 1 \le i \le N-1.$$

Матрицу $P - P_N$ представим в блочном виде

$$P - P_N = \left[\begin{array}{c|c} B_N & b_N \\ \hline 0 \ldots 0 & 0 \end{array}\right].$$

Очевидно, что

$$(P - P_N)^n = \left[\begin{array}{c|c} B_N^n & b_N^n \\ \hline 0 \ldots 0 & 0 \end{array}\right], \qquad n \ge 1,$$

где $b_N^1 = b_N$, $b_N^n = B_N b_N^{n-1}$ для $n > 1$. Отсюда в силу (1) следует

$$\lim_{n \to \infty} B_N^n = 0. \tag{2}$$

Тогда согласно лемме 1 $\rho(B_N) < 1$. Теорема доказана.

3. Приложение к управляемым марковским цепям. Пусть задана управляемая марковская цепь с множеством состояний $S = \{1, 2, \ldots, N\}$ и множествами управлений $U_i = \{1, 2, \ldots, k_i\}$, $i \in S$ [3 – 5].

Элемент $f = [f(i)]$, $i \in S$, из пространства $F = U_1 \times U_2 \times \ldots \times U_N$, где $f(i) \in U_i$, называется решающей функцией, последовательность решающих функций $\pi = (f_1, f_2, \ldots, f_n, \ldots)$ — стратегией. Стратегия вида $f^\infty = (f, f, \ldots, f, \ldots)$ называется стационарной. Для каждой решающей функции $f \in F$ заданы матрица вероятностей перехода $P(f) = \| p_{ij}^k \|$, $i, j \in S$, размера $N \times N$ и $(N \times 1)$-мерный вектор доходов $r(f) = [r_i^k]$, $i \in S$, где $k = f(i) \in U_i$.

Прибыль, приносимая стратегией π, оценивается вектор-функцией

$$\Gamma(\pi) = \lim_{T \to \infty} \inf \frac{1}{T} \sum_{n=0}^{T-1} P_n(\pi) r(f_{n+1}),$$

где $P_n(\pi) = P(f_1) P(f_2) \ldots P(f_n)$, $P_0(\pi) = I$ — единичная матрица размера $N \times N$; i-й элемент $\gamma_i(\pi)$ вектора $\Gamma(\pi)$ представляет собой нижний предел среднего дохода за один шаг, соответствующий начальному состоянию $i \in S$.

Стратегия π^* называется оптимальной, если $\Gamma(\pi^*) \ge \Gamma(\pi)$ для всех π. В [9, 10] доказано существование стационарной оптимальной стратегии в предположении, что множества состояний и управлений конечны.

В дальнейшем предполагаем выполненным следующее условие.

Условие А. *Для любого $f \in F$ марковская цепь, задаваемая матрицей $P(f)$, положительно или неотрицательно регулярна.*

Этот случай, по существу, является основным, заслуживающим внимания с точки зрения задач, встречающихся на практике.

При соблюдении условия А

$$\Gamma(f^\infty) = \gamma(f) \mathbf{1},$$

где $\gamma(f) = \alpha(f)\, r(f) = \sum_{j\in S}\alpha_j(f)\, r_j^{f(j)}$ — предельный средний доход за один шаг; $\alpha(f) = [\alpha_1(f), \alpha_2(f), \ldots, \alpha_N(f)]$ — стационарное распределение вероятностей.

Пусть величина $V^n(f)$ представляет собой вектор суммарных средних доходов, полученных к моменту n при стационарной стратегии f^∞. Тогда он удовлетворяет рекуррентному соотношению

$$V^n(f) = r(f) + P(f)V^{n-1}(f), \quad n = 1, 2, \ldots, \tag{3}$$

где $V^0(f) = 0$.

Следующее предложение является матричным аналогом результата Ховарда [3].

Лемма 3. *Пусть выполнено условие* A. *Тогда*

$$V^n(f) = n\gamma(f)\mathbf{1} + v(f) + \varepsilon_n(f), \quad n > 0,$$

где $\varepsilon_n(f) \to 0$ *при* $n \to \infty$ *и*

$$v(f) = \left([I - P(f) + \mathbf{1}\alpha(f)]^{-1} - \mathbf{1}\alpha(f)\right) r(f).$$

Доказательство см. в [4, с. 42].

Элементы $v_i(f)$ вектора-столбца $v(f)$ называються весами, величины $v_i(f) - v_s(f)$, $i \in S$, — относительными весами.

Если n достаточно велико, то, подставляя в соотношение (3) вместо $V^n(f)$ значение $n\gamma(f)\mathbf{1} + v(f)$, получаем систему уравнений Ховарда

$$v(f) + \gamma(f)\mathbf{1} = r(f) + P(f)v(f). \tag{4}$$

Эта система используется для определения предельного среднего дохода $\gamma(f)$ и N весов $v_i(f)$. Поскольку в (4) число неизвестных больше, чем число уравнений, и вектор $v(f) + c\mathbf{1}$ $(c = \text{const})$ также удовлетворяет (4), один из весов приравнивают нулю. Так, в случае положительно регулярной цепи любое из весов, например $v_N(f)$, полагают равным нулю. А в случае неотрицательно регулярной цепи, исходя из общего случая (с несколькими эргодическими классами), полагают $v_s(f) = 0$ для некоторого эргодического (существенного) состояния s (см. [4, с. 82]). После этого уравнения (4) можно решать относительно предельного среднего дохода $\gamma(f)$ и оставшихся $N-1$ весов $v_i(f)$, $i \neq s$. А это соответствует процедуре определения весов в итерационном алгоритме Ховарда, состоящем из следующих вычислительных процедур:

1. Для $f \in F$ решаем систему уравнений

$$v_i + \gamma = r_i^k + \sum_{j\in S} p_{ij}^k v_j, \quad i \in S, \tag{5}$$

относительно весов v_i, $i \in S$, и прибыли $\gamma = \gamma(f)$, полагая $v_s = 0$ для некоторого эргодического состояния $s \in S$, где $k = f(i)$.

2. Используя найденные значения γ, v_i, $i \in S_s = S \setminus \{s\}$, при каждом $i \in S$ находим решение

$$k' = \arg\max_{k\in U_i}\left[r_i^k + \sum_{j\in S_s} p_{ij}^k v_j \right]$$

и принимаем его за новое решение в состоянии i. При этом следует соблюдать правило выбора решения: если старое решение в i-м состоянии придает величине критерия столь же большое значение, как и любое другое решение, необходимо оставить старое решение неизменным. Процедура заканчивается, когда в двух последовательных итерациях будут получены одинаковые решающие функции, в противном случае нужно перейти к п. 1.

Следует отметить, что использование метода Ховарда для процессов с большим числом состояний (этим, как правило, характеризуются практически интересные случаи) даже с помощью современных вычислительных средств весьма затруднительно, а зачастую и невозможно. Здесь возникает так называемая проблема размерности (или „проклятие размерности", по выражению Р. Беллмана). В связи с этим представляется важным специальное изучение вопроса решения системы уравнений Ховарда с помощью итерационных методов [6]. Заметим, что итерационные методы находят многочисленные приложения даже в случае, когда число уравнений системы велико и применение метода Гаусса или метода Крамера невозможно. Рассмотрим известный итерационный метод решения систем линейных алгебраических уравнений — метод простой итерации.

Сначала покажем, что система уравнений Ховарда (5) при условии $v_s = 0$ для любого $s \in S$ имеет единственное решение.

Теорема 2. *Пусть выполнено условие* A. *Если для любого* $s \in S$ *положить* $v_s = 0$, *то система уравнений Ховарда* (5) *имеет единственное решение* $v_1, \ldots, v_{s-1}, v_{s+1}, \ldots, v_N$, γ *такое, что* $v_i = v_i(f) - v_s(f)$, $i \in S_s$ *и* $\gamma = \gamma(f)$.

Доказательство. Запишем s-е уравнение системы (5) в виде

$$\gamma = r_s^l + \sum_{j \in S_s} p_{sj}^l \left(v_j - v_s\right), \tag{6}$$

где $l = f(s) \in U_s$.

Вычитая (6) из оставшихся уравнений системы (5), получаем соотношения

$$v_i - v_s = r_i^k - r_s^l + \sum_{j \in S_s} \left(p_{ij}^k - p_{sj}^l\right)\left(v_j - v_s\right), \quad i \in S_s. \tag{7}$$

Эти соотношения можно представить в матричном виде

$$w = d(f) + B_s(f)w, \tag{8}$$

где $w = [v_i - v_s]$ $(i \in S_s)$ и $d(f) = \left[r_i^k - r_s^l\right]$ $(i \in S_s)$ — $(N-1)$-мерные векторы-столбцы.

В силу леммы 2 и соотношения (2) матрица $I - B_s(f)$ невырожденная, что и завершает доказательство теоремы.

Следствие. *Пусть выполнено условие* A. *Тогда при любой решающей функции* $f \in F$, *прежде чем решать систему уравнений Ховарда* (5), *можно приравнивать нулю любое из весов* v_i, $i \in S$.

Таким образом, в случае неотрицательно регулярной цепи, вопреки утверждению в [4], нет необходимости искать (при большом N это достаточно сложно) некоторое существенное состояние s, чтобы положить $v_s = 0$. В этом случае, как и для положительно регулярной цепи, можно положить $v_N = 0$.

Пусть задано начальное приближение v_j^0, $i \in S_s$. Все последующие приближения определим правилом

$$\gamma_s^n = r_s^l + \sum_{j \in S_s} p_{sj}^l v_j^{n-1},$$

$$v_i^n = r_i^k + \sum_{j \in S_s} p_{ij}^k v_j^{n-1} - \gamma_s^n, \quad i \in S_s, \ n \geq 1, \tag{9}$$

где s — произвольный элемент S, причем $l = f(s) \in U_s$, $k = f(i) \in U_i$.

Теорема 3. *Пусть выполнено условие А. При любом начальном приближении* v_i^0, $i \in S_s$, *последовательность приближений* $\left(v_i^n, i \in S_s, \gamma_s^n\right)$ *в итерационном процессе* (9) *сходится к* $\left(v_i^*, i \in S_s, \gamma_s^*\right)$ *— решению системы Ховарда с* $v_s = 0$ *т. е.* $v_i^* = v_i(f) - v_s(f)$, $i \in S_s$, *и* $\gamma_s^* = \gamma(f)$.

Доказательство. Итерационный процесс (9) можно представить в виде (см. формулу (8))

$$w^n = d(f) + B_s(f) w^{n-1}. \tag{10}$$

Рекуррентное соотношение (10) представляет собой метод простой итерации [10], который сходится к $w^* = \left[v_i^*\right]$ $(i \in S_s)$ при любом начальном приближении w^0 тогда и только тогда, когда $\rho\left[B_s(f)\right] < 1$. Согласно теореме 1 это неравенство имеет место. Теорема доказана.

Таким образом, доказано, что если из матрицы вероятностей перехода регулярной цепи Маркова построчно вычесть некоторую строку, а затем удалить эту строку и соответствующий столбец, то спектральный радиус полученной таким способом матрицы меньше единицы. Это свойство регулярной цепи Маркова использовано при построении итерационного процесса для решения системы уравнений Ховарда, встречающейся в управляемых марковских цепях с одним эргодическим классом и, возможно, с невозвратными состояниями.

1. *Романовский В. И.* Дискретные цепи Маркова. – М.; Л.: Гостехтеоретиздат, 1949. – 436 с.
2. *Сарымсаков Т. А.* Основы теории процессов Маркова. – Ташкент: Фан, 1988. – 248 с.
3. *Ховард Р.* Динамическое программирование и марковские процессы. – М.: Сов. радио, 1964. – 192 с.
4. *Майн Х., Осаки С.* Марковские процессы принятия решений. – М.: Наука, 1977. – 176 с.
5. *Дынкин Е. Б., Юшкевич А. А.* Управляемые марковские процессы и их приложения. – М.: Наука, 1975. – 338 с.
6. *Крылов В. И., Бобков В. В., Монастырный П. И.* Вычислительные методы. – М.: Наука, 1976. – Том 1. – 304 с.
7. *Хорн Р. , Джонсон Ч.* Матричный анализ. – М.: Мир, 1989. – 655 с.
8. *Кемени Дж., Снелл Дж.* Конечные цепи Маркова. – М.: Наука, 1970. – 272 с.
9. *Derman C.* On sequential decisions and Markov chains // Manag. Sci. – 1962 – **9.** – P. 16 – 24.
10. *Висков О. В., Ширяев А. Н.* Об управлениях, приводящих к оптимальным стационарным режимам // Тр. Мат. ин-та АН СССР. – 1964. – **71.** – С. 35 – 45.

Получено 11.05.98,
после доработки — 25.09.2001