УДК 519.21

Л.Б. ВОВК, Е.И. КАСИЦКАЯ, А.С. САМОСЁНОК

АСИМПТОТИЧЕСКИЕ СВОЙСТВА ЭМПИРИЧЕСКИХ ОЦЕНОК ПАРАМЕТРОВ МАРКОВСКИХ ПОСЛЕДОВАТЕЛЬНОСТЕЙ

Ключевые слова: *марковские случайные поля, стохастические процессы, оценка параметров.*

В теории стохастического программирования одним из наиболее используемых методов является метод эмпирических средних, основные идеи которого тесно переплетаются с идеями теории оценивания. Суть этого метода состоит в аппроксимации имеющейся критериальной функции ее эмпирической оценкой, для которой существует возможность решения соответствующей оптимизационной задачи. Естественно, что условия сходимости существенным образом зависят от функции критерия и вероятностных свойств случайных величин, от которых эта функция зависит. В терминологии теории статистических решений эти вопросы тесно связаны с асимптотическими свойствами оценок неизвестных параметров, а именно состоятельностью, асимптотическим распределением, скоростью сходимости оценок и т.д. В работах [1, 2] такие подходы тщательно проанализированы, в [3] приведен обзор основных работ и результатов, полученных в данной области. В этих публикациях основное внимание уделялось вопросам, когда наблюдаемые случайные величины были независимыми или слабо зависимыми. Однако при решении многих прикладных задач, возникающих в области биологии, социологии, теории коммуникационных сетей, появляется необходимость в решении оптимизационных задач для случая, когда последовательность наблюдаемых случайных величин обладает марковским свойством. Тогда при доказательстве утверждений о сходимости приближенных оценок в определенном вероятностном смысле или об их состоятельности возникает необходимость в доказательстве закона больших чисел или усиленного закона больших чисел для марковских последовательностей. В настоящей статье приводятся некоторые утверждения о выполнимости закона больших чисел для марковских последовательностей и их использовании при доказательстве теорем о состоятельности оценок неизвестных параметров.

При решении поставленной задачи рассматриваются два варианта: случай конечного фазового пространства значений марковской цепи и случай компактного множества состояний марковской последовательности. Предварительно сформулируем утверждения относительно выполнения закона больших чисел, а затем приведем доказательства утверждений о состоятельности оценок.

Определение 1. Пусть X — некоторое конечное множество. Семейство $(P(x,\cdot))_{x\in X}$ вероятностных распределений называется марковским ядром.

Определение 2. Цепь Маркова на конечном множестве X задается первоначальным распределением υ и марковскими ядрами $P_1, P_2, \ldots$ на X. Если $P_i = P \ \ \forall i$, то цепь Маркова называется однородной.

Это утверждение эквивалентно классическому определению цепи Маркова [4].

Определение 3. Марковское ядро P, имеющее строго положительную степень P^r, называется примитивным.

Определение 4. Пусть P — марковское ядро на X. Вероятностное распределение μ на X, удовлетворяющее условию $\mu P = \mu$, называется инвариантным для P.

Будем предполагать, что состояние процесса X_t до текущего момента времени τ определяется σ-алгебрами вида $\Theta_{-\infty}^{\tau}$, а в дальнейшем — вида Θ_{τ}^{∞}.

Определение 5. Стационарный процесс в узком смысле X_t считается удовлетворяющим условию сильного перемешивания, если

$$\sup_{A\in\Theta_{-\infty}^{0},\, B\in\Theta_{\tau}^{\infty}} |P(AB)-P(A)P(B)| = \alpha(\tau) \le \frac{c}{\tau^{1+\varepsilon}}, \quad \tau>0,\ \varepsilon>0, c>0.$$

Теорема 1 [4]. Пусть $\{\xi_i, i\in N\}$ — однородная цепь Маркова с примитивным марковским ядром P на конечном множестве X. Тогда для любого начального распределения υ и любой функции $f: X \to \Re$ по вероятности

$$\frac{1}{n}\sum_{i=1}^{n} f(\xi_i) \to E_{\mu}(f),\ n\to\infty,$$

где μ — инвариантное распределение марковского ядра P.

Теорема 2 [1]. Пусть (Ω, U, P) — вероятностное пространство, K — компактное подмножество некоторого банахова пространства с нормой $\|\cdot\|$. Пусть также $\{U_n, n\in N\}$ — последовательность σ-алгебр, $U_n \subset U$, $U_{n_1} \subset U_{n_2}$, $n_1 < n_2$, а $\{Q_n(s) = Q_n(s,w)\colon (s,w)\in K\times\Omega, n\in N\}$ — семейство действительных функций, удовлетворяющее следующим условиям:

1) для фиксированных n, w функция $Q_n(s,w)$, $s\in K$, непрерывна;

2) для любых n, s функция $Q_n(s,w)$, $w\in\Omega$, U_n-измерима;

3) для некоторого элемента $s_0\in K$ и любого $s\in K$ выполняется условие $Q_n(s,w)\to\Phi(s;s_0)$, $n\to\infty$ по вероятности, где $\Phi(s;s_0)$, $s\in K$, — действительная функция, непрерывная на K и удовлетворяющая условию $\Phi(s;s_0) > \Phi(s_0;s_0)$, $s\ne s_0$;

4) для любого $\delta>0$ существуют величина $\gamma_0>0$ и функция $c(\gamma)$, $\gamma>0$, $c(\gamma)\to 0$, $\gamma\to 0$, такие, что при всех $s'\in K$ и $\gamma\colon 0<\gamma<\gamma_0$, имеем

$$\sup_{\substack{\|s-s'\|<\gamma\\ \|s-s_0\|>\delta}} |Q_n(s)-Q_n(s')| < c_n(\gamma),$$

$c_n(\gamma) = c_n(\gamma, w)\to c(\gamma)$, $n\to\infty$ по вероятности.

Для всех n, w элемент $S_n = S_n(w)$ определяется соотношением $Q_n(s_n) = \min_{s\in K} Q_n(s)$. Такой элемент можно выбрать U_n-измеримым по w. Тогда по вероятности $s_n\to s_0$ при $n\to\infty$.

Теорема 3 [1]. Пусть (Ω, U, P) — вероятностное пространство, K — компактное подмножество некоторого банахова пространства с нормой $\|\cdot\|$. Пусть также $\{U_n, n\in N\}$ — последовательность σ-алгебр, $U_n \subset U$, $U_{n_1} \subset U_{n_2}$, $n_1 < n_2$, а $\{Q_n(s) = Q_n(s,w)\colon (s,w)\in K\times\Omega, n\in N\}$ — семейство действительных функций, удовлетворяющее аналогично теореме 2 следующим условиям:

1) непрерывность функции $Q_n(s,w)$;

2) измеримость функции $Q_n(s,w)$;

3) для некоторого элемента $s_0\in K$ и любого $s\in K$ выполняется условие $Q_n(s,w)\to\Phi(s;s_0)$, $n\to\infty$ с вероятностью 1, где $\Phi(s;s_0)$, $s\in K$, — действительная функция, непрерывная на K и удовлетворяющая условию $\Phi(s;s_0) > \Phi(s_0;s_0)$, $s\ne s_0$;

4) для любого $\delta>0$ существуют величина $\gamma_0>0$ и функция $c(\gamma)$, $\gamma>0$, $c(\gamma)\to 0$, $\gamma\to 0$, такие, что при всех $s'\in K$ и $\gamma\colon 0<\gamma<\gamma_0$, имеем

$$P\{\overline{\lim_{n\to\infty}} \sup_{\substack{\|s-s'\|<\gamma\\ \|s-s_0\|>\delta}} |Q_n(s)-Q_n(s')| < c(\gamma)\} = 1.$$

Для всех n, w элемент $s_n = s_n(w)$ определяется соотношением $Q_n(s_n) = \min_{s\in K} Q_n(s)$.

Такой элемент можно выбрать U_n-измеримым по w. Тогда $P\{s_n\to s_0,\ n\to\infty\} = 1$.

Перейдем к рассмотрению основных результатов.

Пусть K — компактное подмножество $\Re^l$, $l\ge 1$; $f\colon K\times X\to\Re$ — неотрицательная функция, непрерывная по первому аргументу и измеримая по второму. Имеются наблюдения $\{\xi_i, i=\overline{1,n}\}$, $n\in N$.

Задача состоит в поиске точки минимума функции

$$F(\vec{u}) = E\{f(\vec{u}, \xi_1)\},\ \vec{u}\in K. \tag{1}$$

Аппроксимируем функцию, заданную соотношением (1), следующей функцией:

$$F_n(\vec{u}) = \frac{1}{n}\sum_{i=1}^{n} f(\vec{u}, \xi_i), \ \vec{u} \in K, \tag{2}$$

и рассмотрим условия сходимости для случаев конечного и бесконечного множества состояний.

Теорема 4. Пусть множество состояний X случайной последовательности является конечным и функция (1) имеет единственную точку минимума $\vec{u}_0$. Тогда для любой точки минимума $\vec{u}_n$ функции (2) имеем $\vec{u}_n \to \vec{u}_0$ по вероятности при $n \to \infty$.

Доказательство. Используем теорему 2 и зададим $Q_n = F_n, U_n = \sigma\{\xi_i, i = \overline{1, n}\}$. Очевидно, что условия 1, 2 теоремы 2 выполнены. В силу теоремы 1 выполнено и условие 3 теоремы 2.

Обозначим

$$\psi(\gamma, z) = \sup_{\substack{\vec{u}, \vec{v} \in K: \\ \|\vec{u}-\vec{v}\| < \gamma}} |f(\vec{u}, z) - f(\vec{v}, z)|, \ \gamma > 0, \ z \in X.$$

Для всех n, $\vec{u}_1 \in K$, $\gamma > 0$

$$\zeta_n(\vec{u}_1, \gamma) = \sup_{\substack{\vec{u}_1 \in K: \\ \|\vec{u}-\vec{u}_1\| < \gamma}} |F_n(\vec{u}) - F_n(\vec{u}_1)| \le \frac{1}{n}\sum_{i=1}^{n} \psi(\gamma, \xi_i).$$

Согласно теореме 1 $\frac{1}{n}\sum_{i=1}^{n} \psi(\gamma, \xi_i) \to E_\mu(\psi(\gamma, \cdot))$ по вероятности при $n \to \infty$.

Обозначим $c(\gamma) = E\{\psi(\gamma, \cdot)\}$. При любом $z \in X$ имеем $\psi(\gamma, z) \to 0$, $\gamma \to 0$. Следовательно, $c(\gamma) \to 0, \gamma \to 0$. Отсюда имеем справедливость условия 4 теоремы 2. Теперь из теоремы 2 вытекает теорема 4. $\square$

Перейдем к случаю компактного множества состояний последовательности.

Пусть $\{\xi_i, i \in N\}$ — цепь Маркова с произвольным пространством состояний. Если на σ-алгебре U_X существуют конечная мера m с $m(U) > 0$, целое число $v > 1$ и число $\varepsilon > 0$ такие, что вероятность перехода за v шагов $p^{(v)}(\xi, A) \le 1 - \varepsilon$ при $m(A) \le \varepsilon$, то будем считать, что выполняется гипотеза Деблина.

Если выполняется гипотеза Деблина и существует только один эргодический класс, причем он не содержит подклассов, то существует такое распределение $p(A)$ на U_X, что

$$\sup_{x, A} |p^{(n)}(x, A) - p(A)| \le c\rho^n, \tag{3}$$

где $c > 0$, $0 < \rho < 1$ — константы [4].

Если справедливо соотношение (3) и $\pi(A) = p(A)$, то последовательность ξ_i стационарна и удовлетворяет условию сильного перемешивания с коэффициентом $\varphi(n) \le 2c\rho^n$ [4].

Теорема 5. Пусть функция (1) имеет единственную точку минимума $\vec{u}_0$, для функции $f(\vec{u}, \xi_1)$ выполняется условие $E\{\sup_{\vec{u} \in K} f(\vec{u}, \xi_1)\} < \infty$ и последовательность ξ_i удовлетворяет условию Деблина с коэффициентом $\rho^n \le \frac{c}{n^{1+\varepsilon}}$. Тогда для любой точки минимума $\vec{u}_n$ функции (2) справедливо равенство $P\{\vec{u}_n \to \vec{u}_0, n \to \infty\} = 1$.

Доказательство. Применим теорему 3 и зададим $Q_n = F_n, U_n = \sigma\{\xi_i, i = \overline{1, n}\}$. Очевидно, что условия 1, 2 теоремы 3 выполнены.

Рассмотрим условие 3 теоремы 3. Поскольку последовательность ξ_i удовлетворяет условию равномерного сильного перемешивания с коэффициентом — показательной функцией, можно выбрать ρ таким образом, что

$$|Ef(\vec{u}, \vec{\xi}_i) f(\vec{u}, \vec{\xi}_j) - Ef(\vec{u}, \vec{\xi}_i) Ef(\vec{u}, \vec{\xi}_j)| \le \frac{c}{1 + |i - j|^{1+\varepsilon'}}, \ \varepsilon' > 0.$$

Обозначим $\eta_n(\vec{u}) = F_n(\vec{u}) - EF_n(\vec{u})$ и оценим $E\eta_n^2(\vec{u})$:

$$E\eta_n^2(\vec{u}) = E\left\{\frac{1}{n}\sum_{i=1}^{n} f(\vec{u}, \vec{\xi}_i) - E\left(\frac{1}{n}\sum_{i=1}^{n} f(\vec{u}, \vec{\xi}_i)\right)\right\}^2 =$$

$$= E\left\{\frac{1}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n}[f(\vec{u}, \vec{\xi}_i) - Ef(\vec{u}, \vec{\xi}_i)][f(\vec{u}, \vec{\xi}_j) - Ef(\vec{u}, \vec{\xi}_j)]\right\} = \frac{1}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n} Ey_i y_j,$$

где $y_i = f(\vec{u}, \vec{\xi}_i) - Ef(\vec{u}, \vec{\xi}_i)$.

Учитывая условие сильного перемешивания, имеем $Ey_i y_j \le \dfrac{c}{1 + |i - j|^{1+\varepsilon'}}$, $\varepsilon' > 0$. Поэтому

$$\frac{1}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n} Ey_i y_j \le \frac{c}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n}\frac{c}{1 + |i - j|^{1+\varepsilon'}} \le \frac{c}{n}.$$

Пусть $n = m^2$. Тогда согласно лемме Бореля–Кантелли

$$P\{\lim_{m\to\infty} \eta_{m^2} = 0\} = 1.$$

Обозначим $\zeta_m = \sup\limits_{m^2 \le n \le (m+1)^2} |\eta_n - \eta_{m^2}|$.

Для $m^2 \le n \le (m+1)^2$ справедливо следующее неравенство:

$$|\eta_n| \le |\eta_{m^2}| + \sup_{m^2 \le n \le (m+1)^2} |\eta_n - \eta_{m^2}|.$$

Для ζ_m запишем

$$E(\zeta_m)^2 = E\sup_{m^2 \le n \le (m+1)^2} |\eta_n - \eta_{m^2}| \le \sup_{m^2 \le n \le (m+1)^2}\left|\frac{1}{m^2}\sum_{i=1}^{n}\sum_{j=1}^{n} Ey_i y_j\right| \le$$

$$\le \frac{1}{m^2}\sum_{i=m^2+1}^{(m+1)^2}\sum_{j=m^2+1}^{(m+1)^2} E|y_i y_j| \le c\left[\frac{(m+1)^2 - m^2 - 1}{m^2}\right]^2 = 2\frac{c}{m^2}.$$

Отсюда согласно лемме Бореля–Кантелли $P\{\lim\limits_{m\to\infty} \zeta_n = 0\} = 1$. Тогда $P\{\lim\limits_{m\to\infty} \eta_n = 0\} = 1$, а значит и $P\{\lim\limits_{n\to\infty} F_n(\vec{u}) = F(\vec{u})\} = 1$. Следовательно, и условие 3 теоремы 3 выполнено.

Перейдем к условию 4. Обозначим

$$\psi(\gamma, z) = \sup_{\substack{\vec{u}, \vec{v} \in K: \\ \|\vec{u} - \vec{v}\| < \gamma}} |f(\vec{u}, z) - f(\vec{v}, z)|, \gamma > 0, \ z \in X.$$

Для всех n, $\vec{u}_1 \in K$, $\gamma > 0$ имеем

$$\zeta_n(\vec{u}_1, \gamma) = \sup_{\substack{\vec{u}_1 \in K: \\ \|\vec{u} - \vec{u}_1\| < \gamma}} |F_n(\vec{u}) - F_n(\vec{u}_1)| \le \frac{1}{n} \sum_{i=1}^{n} \psi(\gamma, \xi_i).$$

Аналогично доказательству условия 3 теоремы 2 с учетом леммы Бореля–Кантелли несложно показать, что $\frac{1}{n} \sum_{i=1}^{n} \psi(\gamma, \xi_i) \to E_\mu(\psi(\gamma, \cdot))$, $n \to \infty$ с вероятностью 1.

Обозначим $c(\gamma) = E\{\psi(\gamma, \cdot)\}$. При любом $z \in X$ имеем $\psi(\gamma, z) \to 0$, $\gamma \to 0$. Следовательно, $c(\gamma) \to 0$ при $\gamma \to 0$, а значит справедливо условие 4 теоремы 3. Таким образом, из теоремы 3 вытекает теорема 5. $\square$

Итак, в статье проанализированы условия, при которых возможна аппроксимация критериальной функции марковского процесса ее эмпирической оценкой. Сформулированные в статье теоремы являются еще одним инструментом в исследовании сходимости оценок неизвестных параметров к их истинным значениям.

СПИСОК ЛИТЕРАТУРЫ

1. Dorogovtsev A.Ya. The theory of estimation of random processes parameters. — Kiev: Vishcha Shkola, 1982.
2. Knopov P.S., Kasitskaya E.J. Empirical estimates in stochastic optimization and identification. — Boston: Kluwer Acad. Publ. — 1995. — P. 70–74.
3. Кнопов П.С., Самосёнок А.С. О некоторых прикладных задачах марковских случайных процессов с локальным взаимодействием // Кибернетика и системный анализ. — 2011. — № 3. — С. 15–32.
4. Ибрагимов И.А., Линник Ю.В. Независимые и стационарно связанные величины. — М.: Наука, 1980. — С. 463–468.

Поступила 28.08.2011