А. В. Назин, А. С. Позняк, О сходимости автоматных алгоритмов Льюиса и Варшавского-Воронцовой, *Автомат. и телемех.*, 1983, выпуск 9, 165–169

УДК 65.01

О СХОДИМОСТИ АВТОМАТНЫХ АЛГОРИТМОВ ЛЬЮИСА И ВАРШАВСКОГО — ВОРОНЦОВОЙ

НАЗИН А. В., ПОЗНЯК А. С.

(Москва)

Рассматривается задача адаптивного выбора вариантов с бинарными потерями. Исследуется сходимость и оценивается скорость сходимости известных автоматных алгоритмов Льюиса и Варшавского – Воронцовой, осуществляющих перестройку вероятностей выбора вариантов с целью минимизации средних потерь.

Пусть имеется N возможных вариантов $x(1),\ldots,x(N)$ и в каждый момент времени t_n ($n=0, 1,\ldots$; $t_{n+1}>t_n$) необходимо выбрать один из них. Качество выбранного в момент времени t_n варианта $x_n\in X=\{x(1),\ldots$ $\ldots,x(N)\}$ будем оценивать случайной величиной

(1) $$\xi_n=\xi_n(x_n, \omega),$$

которая имеет смысл потерь, соответствующих выбранному варианту x_n (ω – элементарный исход). Реализация величины ξ_n наблюдается до наступления момента времени t_{n+1}.

Будем предполагать, что последовательности вариантов $x_n\in X$ и случайных функций $\xi_n(x, \omega)$, $x\in X$ удовлетворяют следующим условиям:

А) $$\xi_n(x(i), \omega)=\begin{cases}1 & \text{с вероятностью } v_i,\\ 0 & \text{с вероятностью } 1-v_i,\end{cases}$$

при этом $\xi_n=1$ интерпретируется как «штраф» за выбор соответствующего варианта $x_n=x(i)$ и $\xi_n=0$ – как «поощрение» ($i=\overline{1, N}$, $n=0, 1,\ldots$);

Б) для любых $n=0, 1,\ldots$ и $x\in X$ случайная величина $\xi_n(x, \omega)$ и совокупность $\{x_k, \xi_t(x, \omega)\,|\,k=\overline{0, n};\ t=\overline{0, n-1}\}$ независимы.

Приведенные условия являются обычными в теории обучающихся систем [1]: набор априори неизвестных чисел v_i в условии А) характеризует свойства «цетлиновской внешней среды», а условие Б) является условием независимости статистических свойств внешней среды в момент времени t_n от предыстории процесса $\{\xi_t(x_t, \omega), t=\overline{0, n-1}\}$.

Пусть существует единственный вариант $x(\alpha)\in X$, соответствующий минимальным средним потерям

(2) $$M\{\xi_n(x(\alpha),\omega)\}=v_\alpha<v_-=\min_{i\neq\alpha} v_i,$$

который будем называть оптимальным. Задача адаптивного выбора вариантов состоит в организации такого перебора элементов из X, использующего текущую информацию $\{x_t, \xi_t(x, \omega)\,|\,t=\overline{0, n-1}\}$, при котором вероятность выбора оптимального варианта стремилась бы к единице при $n\to\infty$.

Для решения этой задачи были предложены различные алгоритмы, которые на каждом шаге $n=0, 1,\ldots$ осуществляют рандомизированный выбор варианта $x_n=x(i)$ с условной вероятностью

$$p_n(i)=P\{x_n=x(i)/x_k, \xi_k(k=\overline{0, n-1})\}$$

и перестраивают вектор $p_n^T=(p_n(1),\ldots,p_n(N))$ вероятностей выбора на основе получаемой информации $\xi_n(x_n,\omega)$ о потерях [1—7]. Алгоритмы такого типа изучались в теории поведения автоматов в случайных средах, и в связи с этим их часто называют автоматными, хотя они обладают непрерывными, а не дискретными множествами состояний (в смысле теории автоматов).

Рассмотрим алгоритм Льюиса [5, 6]

$$
\begin{aligned}
&x_n=x(i),\\
(3)\qquad &p_{n+1}=p_n+\gamma p_n(i)(e_i-p_n)\left[\frac{1-\xi_n}{a+(1-a)p_n(i)}-\frac{\xi_n}{1-(1-a)p_n(i)}\right],
\end{aligned}
$$

где $e_i^T=(\underbrace{0,\ldots,0,1}_{i},0,\ldots,0)\in R^N$, $\gamma\in(0,\ 1]$ — длина шага, $a\in(0,\ 1]$ — параметр алгоритма.

Утверждение. Пусть выполнены условия А), Б) и

$$
(4)\qquad v_\alpha<\frac{a}{1+a},\quad v_-\geqslant\frac{1}{1+a},\quad p_0(\alpha)>0.
$$

Тогда при любом $a\in(0,\ 1]$ существует число $\bar\gamma=\bar\gamma(v_\alpha,\ v_-,\ a)$, такое, что для всякого $\gamma\in(0,\ \bar\gamma)$ последовательность векторов $\{p_n\}$, порождаемая алгоритмом Льюиса (3), сходится с вероятностью 1 и в среднеквадратическом к оптимальной чистой стратегии $p^*=e_\alpha$ и, следовательно, при $n\to\infty$ вероятность выбора оптимального варианта $x(\alpha)$ стремится к единице. При этом имеет место экспоненциальная скорость сходимости в среднеквадратическом:

$$
(5)\qquad M\{(1-p_n(\alpha))^2\}\leqslant K(\gamma,\ a,\ v_\alpha,\ v_-,\ p_0(\alpha))\lambda^n,
$$

где $\lambda=\lambda(\gamma,\ a,\ v_\alpha,\ v_-)\in(0,\ 1)$.

Следствие. Величина $\bar\gamma$, определяющая границу области сходимости алгоритма Льюиса (3), является корнем уравнения

$$
(6)\qquad \max_{q\in[0,1]} G(q,a,\gamma,v_\alpha,v_-)=2\gamma^{-1},
$$

где функция G задается выражением (П.3); величина λ в (5), характеризующая порядок скорости сходимости, задается выражениями (П.1)—(П.5) (см. приложение).

Сформулированное утверждение гарантирует сходимость алгоритма Льюиса к оптимальной стратегии $p^*=e_\alpha$ лишь для таких задач, в которых средние потери v_α и v_- достаточно разделены между собой (4). При $a=1$ условие (4) выделяет наибольший класс задач. Соответствующий этому значению параметра a алгоритм (3) упрощается и переходит в алгоритм Варшавского — Воронцовой [3]:

$$
\begin{aligned}
&x_n=x(i),\\
(7)\qquad &p_{n+1}=p_n+\gamma(1-2\xi_n)p_n(i)(e_i-p_n).
\end{aligned}
$$

В результате проведенных на ЭВМ вычислений для различных величин v_α, v_- и a были определены значения $\bar\gamma$, $\gamma^*\overset{\Delta}{=}\arg\min\limits_{\gamma\in(0,\bar\gamma)}\lambda(\gamma,a,v_\alpha,v_-)$ и $\lambda(\gamma^*)\overset{\Delta}{=}\lambda(\gamma^*,\ a,\ v_\alpha,\ v_-)$. Полученные результаты показали, что во всех случаях $\lambda(\gamma^*,\ 1,\ v_\alpha,\ v_-)<\lambda(\gamma^*,\ a,\ v_\alpha,\ v_-)$, т. е. значение параметра $a=1$ в алгоритме Льюиса является оптимальным. Таким образом, в классе алгоритмов Льюиса (3) оптимальным по скорости сходимости в смысле (5) является алгоритм Варшавского — Воронцовой (7).

Приведем в заключение некоторые значения $\bar{\gamma}$, γ^* и $\lambda(\gamma^*)$ при $a=1$, $N=2$:

а) для $v_\alpha=0{,}1$, $v_-=0{,}9$ – $\bar{\gamma}=0{,}331$, $\gamma^*=0{,}198$, $\lambda(\gamma^*)=0{,}904$;

б) для $v_\alpha=0{,}3$, $v_-=0{,}7$ – $\bar{\gamma}=0{,}169$, $\gamma^*=0{,}101$, $\lambda(\gamma^*)=0{,}977$.

ПРИЛОЖЕНИЕ

Доказательство утверждения. Введем последовательности величин $q_n=1-p_n(\alpha)$ и σ-алгебр $\mathfrak{F}_n=\sigma\{x_0, \xi_0; \ldots; x_{n-1}, \xi_{n-1}\}$. Рассмотрим выпуклую монотонно возрастающую дифференцируемую на интервале $(0, 1)$ функцию $W(q)$, $q\in[0, 1)$, $W(0)=0$. В силу (3) для любого $n=0, 1, \ldots$ с вероятностью 1

$$M\{W(q_{n+1})/\mathfrak{F}_n\}=(1-q_n)\left[v_\alpha W\left(q_n+\frac{\gamma q_n(1-q_n)}{a+(1-a)q_n}\right)+\right.$$
$$\left.+(1-v_\alpha)W\left(q_n-\frac{\gamma(1-q_n)q_n}{1-(1-a)q_n}\right)\right]+$$
$$+\sum_{i\neq\alpha}p_n(i)\left[v_iW\left(q_n-\frac{\gamma p_n(i)(1-q_n)}{1-(1-a)p_n(i)}\right)+\right.$$
$$\left.+(1-v_i)W\left(q_n+\frac{\gamma p_n(i)(1-q_n)}{a+(1-a)p_n(i)}\right)\right].$$

На основании (2) и с учетом монотонности $W(q)$ получим

$$M\{W(q_{n+1})/\mathfrak{F}_n\}\leqslant(1-q_n)\left[v_\alpha W\left(q_n+\frac{\gamma}{a}q_n(1-q_n)\right)+\right.$$
$$\left.+(1-v_\alpha)W(q_n-\gamma q_n(1-q_n))\right]+\sum_{i\neq\alpha}p(i)\left[v_-W(q_n-\gamma p_n(i)(1-q_n))+\right.$$
$$\left.+(1-v_-)W\left(q_n+\frac{\gamma}{a}p_n(i)(1-q_n)\right)\right].$$

Используя формулу Тейлора и учитывая монотонность $W''(q)$, получим

$$M\{W(q_{n+1})/\mathfrak{F}_n\}\leqslant W(q_n)+\gamma W'(q_n)(1-q_n)\left\{q_n(1-q_n)\left[\left(\frac{1}{a}+1\right)v_\alpha-1\right]+\right.$$
$$\left.+\left[\frac{1}{a}-\left(1+\frac{1}{a}\right)v_-\right]\min\sum_{i\neq\alpha}p^2(i)\right\}+$$
$$+\frac{\gamma^2}{2}(1-q_n)^2\left\{q_n^2(1-q_n)\left[(1-v_\alpha)W''(q_n)+\right.\right.$$
$$\left.+\frac{v_\alpha}{a^2}W''\left(q_n+\frac{\gamma}{a}q_n(1-q_n)\right)\right]+$$
$$\left.+\left[v_-W''(q_n)+\frac{1-v_-}{a^2}W''\left(q_n+\frac{\gamma}{a}q_n(1-q_n)\right)\right]\max\sum_{i\neq\alpha}p^3(i)\right\},$$

где максимум и минимум берутся по всем $p(i)$, $i\neq\alpha$ на симплексе $\sum_{i\neq\alpha}p(i)=q_n$, $\forall p(i)\geqslant0$ и равны $q_n^2/(N-1)$ и q_n^3 соответственно. В качестве $W(q)$, как видно из полученного неравенства, удобно выбрать функцию, удовлетворяющую дифференциальному уравнению

$$q(1-q)[A(1-q)+Bq]W'(q)=\mu ABW(q),\quad q\in(0, 1),$$

где μ – некоторый параметр, т. е. функцию

$$W(q)=q^{B\mu}(1-q)^{-A\mu}[A+(B-A)q]^{(A-B)\mu},$$

(П.1) $$A=1-\left(1+\frac{1}{a}\right)v_\alpha,\qquad B=\frac{1}{N-1}\left[v_- \left(1+\frac{1}{a}\right)-\frac{1}{a}\right].$$

Поскольку в силу условий теоремы $A>0$ и $B>0$, то при

(П.2) $$\mu=B^{-1}\max\left\{2;\, 2+\frac{|B-2A|}{A|B-A|}(B-2A)\right\}$$

эта функция обладает перечисленными выше свойствами (т. е. $W(0)=0$, $W'(q)>0$, $W''(q)>0$, $q\in(0,1)$) и, кроме того, $W'''(q)>0$, $q\in(0,1)$. Заменив далее $W''(q)$ явным выражением, окончательно приходим к неравенству

$$M\{W(q_{n+1})/\mathfrak{F}_n\}\leqslant\left[1-\gamma\mu AB+\frac{\gamma^2}{2}\mu ABG(q_n, a, \gamma, v_-, v_\alpha)\right]W(q_n),$$

где

$$G(q, a, \gamma, v_-, v_\alpha)=\frac{(1-v_\alpha)(1-q)+v_-q}{[A+(B-A)q]^2}\psi(q)+$$

$$+a^{-2}\frac{v_\alpha(1-q)+(1-v_-)q}{\left[1-\frac{\gamma}{a}q\right]^{A\mu+2}}\left[1+\frac{\gamma}{a}(1-q)\right]^{B\mu-2}\times$$

$$\times\frac{\left[A+(B-A)q\left(1+\frac{\gamma}{a}(1-q)\right)\right]^{(A-B)\mu-2}}{[A+(B-A)q]^{(A-B)\mu}}\psi\left(q+\frac{\gamma}{a}q(1-q)\right),$$

(П.3) $$\psi(q)=3(B-A)q^2+2(2A-B)q+\mu AB-A.$$

При $\gamma\in[0, a)$ и $q\in[0, 1]$ функция $G(q, a, \gamma, v_-, v_\alpha)$ является непрерывной по γ и q. Поэтому найдется $\overline{\gamma}=\overline{\gamma}(a, v_-, v_\alpha)$, такое, что для всех $\gamma\in(0\ \overline{\gamma})$

(П.4) $$\lambda_0=1-\gamma\mu AB+\frac{\gamma^2}{2}\mu AB\max_{q\in[0,1]}G(q, a, \gamma, v_-, v_\alpha)<1$$

и, следовательно,

$$M\{W(q_{n+1})/\mathfrak{F}_n\}\leqslant\lambda_0 W(q_n)\leqslant W(q_n),$$

т. е. $\{W(q_n)\}$ – супермартингал. Из этого вытекает, что $q_n\xrightarrow[n\to\infty]{}q^*(\omega)$ с вероятностью 1. Кроме того, при всех $n=0, 1, \ldots$

$$M\{W(q_n)\}\leqslant M\{W(q_0)\}\lambda_0^n\xrightarrow[n\to\infty]{}0,$$

откуда следует, что $q^*(\omega)=0$ с вероятностью 1. Значит, $\lim\limits_{n\to\infty}p_n(\alpha)=1$ почти наверное и в среднеквадратическом.

Поскольку

$$W(q)\geqslant q^{B\mu}C^\mu,$$

где

$$C=\begin{cases}A^{A-B}, & B\geqslant 2A,\\ A^{-A}(B-A)^{2A-B}, & B>2A,\end{cases}$$

то, используя неравенство Иенсена (так как $B\mu\geqslant 2$), получаем

$$M\{q_n^2\}\leqslant C^{-2/B}M\{W(q_n)\}^{2/B\mu}\leqslant C^{-2/B}[W(q_0)]^{2/B\mu}\lambda_0^{\frac{2}{B\mu}n},$$

что соответствует утверждению теоремы, если принять $\lambda_0^{2/B\mu}=\lambda$,

(П.5) $$K(\gamma, a, v_-, v_\alpha, p_0, (\alpha))=C^{-2/B}[W(1-p_0(\alpha))]^{2/B\mu}.$$

Теорема доказана.

ЛИТЕРАТУРА

1. *Цетлин М. Л.* Исследования по теории автоматов и моделированию биологических систем. М.: Наука, 1969.
2. *Luce R. D.* Individual Choice Behavior, N. Y.: Willey, 1959.
3. *Варшавский В. И., Воронцова И. П.* О поведении стохастических автоматов с переменной структурой.— Автоматика и телемеханика, 1963, № 3, с. 353—360.
4. *Варшавский В. И.* Коллективное поведение автоматов. М.: Наука, 1973.
5. *Цыпкин Я. З., Позняк А. С.* Обучающиеся конечные автоматы.— Техническая кибернетика, 1972, № 3, с. 127—140.
6. *Tsypkin Ya. Z., Poznyak A. S.* Lerning Automata.— J. Cybernetics and Inform. Sci. Special Issue on Lerning Automata. ASC, 1977, v. 1, № 2, THRU 4, Spring-Summer-Fall, p. 128—160.
7. *Narendra K. S., Lakshmivarahan S.* Lerning Automata — A Critique.— J. Cybernetics and Inform. Sci. Special Issue on Lerning Automata. ASC, 1977, v. 1, № 2, THRU 4, Spring-Summer-Fall, p. 53—65.

Поступила в редакцию
14.VII.1981

ON CONVERGENCE OF LEWIS AND VARSHAVSKIY — VORONTSOVA AUTOMATON ALGORITHMS

NAZIN A. V., POZNYAK A. S.

The paper is concerned with adaptive choice of options with binary losses. Convergence is studied and the rate of convergence of the well-known automaton algorithms of Lewis and Varshavskiy — Vorontsova that restructure the probability of choosing the options for minimizing the mean losses is estimated.