LIETUVOS MATEMATIKOS RINKINYS, VII Nr. 3, 1968
ЛИТОВСКИЙ МАТЕМАТИЧЕСКИЙ СБОРНИК, VII № 3, 1968

АНТАГОНИСТИЧЕСКАЯ ДИНАМИЧЕСКАЯ ИГРА НА ОСНОВЕ ПОВТОРЕНИЯ БИМАТРИЧНОЙ ИГРЫ

С. П. ВАКРИНЕНЕ

Динамическую игру двух лиц можно формулировать так. Первый игрок первоначально имеет капитал p_0, второй игрок – r_0. Проводится конечная игра, согласно результатам которой капиталы игроков изменяются. Если $A=\|a_{ij}\|$, $B=\|b_{ij}\|$ – матрицы выигрышей первого и второго игроков в этой игре, то капиталы изменяются согласно формулам

$$p_1=p_0+a_{i_1j_1},$$
$$r_1=r_0+b_{i_1j_1},$$

где i_1 и j_1 – в первом шагу выбранные строка и столбец. Вообще

$$p_{k+1}=p_k+a_{i_{k+1}j_{k+1}},$$
$$r_{k+1}=r_k+b_{i_{k+1}j_{k+1}}, \qquad k=0,\ 1,\ \ldots,$$

причем на $k+1$-ом шагу игроки знают p_k и r_k. Игра останавливается, когда $(p_k,\ r_k)$ выходит из некоторой области L. Окончательный выигрыш первого игрока – $M(p_n,\ r_n)$, а второго – $N(p_n,\ r_n)$ (n соответствует моменту окончания игры).

Случай, когда $A=-B$, $M=-N$ был рассмотрен Л. С. Шэпли [1] и Дж. Милнором [2], а случай многих повторяемых матричных игр – Г. Эвереттом [3] и др.

Оказалось, что похоже решается и игра, когда повторяется биматричная игра, но правило окончания игры имеет антагонистический характер:

$$L=\{(p,\ r): p<C_1,\ r<C_2\},$$
$$M(p,\ r)=f(p,\ r),\quad N(p,\ r)=-f(p,\ r),$$

где $f(p,\ r)$ – ограниченная в конечной области R функция, где

$$R=\left\{(p,\ r):\begin{array}{c} p_0\leqslant p\leqslant \bar{C}=C_1+\max\limits_{i,j} a_{ij} \\ r_0\leqslant r\leqslant \mathbf{C}=C_2+\max\limits_{i,j} b_{ij}\end{array}\right\}.$$

Всюду в дальнейшем мы будем предполагать, что $a_{ij}>0$ и $b_{ij}\geqslant 0$ для всех i и j. Эту игру будем обозначать через Γ.

Вводим следующие обозначения. Пусть $X(C)$ является смешанной максиминной стратегией в игре с матрицей C, $Y(C)$ – соответственно минимаксной стратегией. Пусть $U(p,\ r)$ – ограниченная функция. Локальной U – стратегией первого игрока будем называть стратегию поведения, следуя которой для каждого k первый игрок на $k+1$-ом шагу (когда имеются накопленные капиталы p_k и r_k) применяет стратегию $X\left(\|U(p_k+a_{ij},\ r_k+b_{ij})\|\right)$.

Аналогично, локальная U – стратегия второго игрока – это стратегия $Y\left(\|U(p_k+a_{ij},\ r_k+b_{ij})\|\right)$ для каждого k.

Теорема. *Значение игры* Γ *существует и равно* $V(p_0,\ r_0)$, *где* $V(p,\ r)$ *есть ограниченное в конечной области единственное решение функционального уравнения*

$$\Phi(p,\ r)=\begin{cases}\operatorname{val}\|\Phi(p+a_{ij},\ r+b_{ij})\|, & \text{если}\quad (p,\ r)\in L,\\ f(p,\ r), & \text{если}\quad (p,\ r)\notin L.\end{cases}\tag{1}$$

Локальные V – *стратегии будут оптимальными.*

Доказательство. Прежде всего докажем, что функциональное уравнение (1) имеет ограниченное в конечной области решение $V(p,\ r)$. Решение будем конструировать методом тераций. Пусть

$$\Phi_0(p,\ r)=\begin{cases}\inf\limits_{(p,\,r)\in R} f(p,\ r), & \text{если}\quad (p,\ r)\in L,\\ f(p,\ r), & \text{если}\quad (p,\ r)\notin L.\end{cases}$$

Далее, пусть $\Phi_n(p,\ r)=T\Phi_{n-1}(p,\ r)$, где

$$T\Phi(p,\ r)=\begin{cases}\operatorname{val}\|\Phi(p+a_{ij},\ r+b_{ij})\|, & (p,\ r)\in L,\\ \Phi(p,\ r), & (p,\ r)\notin L.\end{cases}$$

Оператор T сконструирован так, что его неподвижная точка является решением функционального уравнения (1). Будем доказывать существование неподвижной точки.

Когда $(p,\ r)\in L$

$$\Phi_1(p,\ r)=\operatorname{val}\|\Phi_0(p+a_{ij},\ r+b_{ij})\|\geqslant$$

$$\geqslant\operatorname{val}\|\inf_{(p,\,r)\in R} f(p,\ r)\|=\inf_{(p,\,r)\in R} f(p,\ r)=\Phi_0(p,\ r).$$

Когда $(p,\ r)\in L$, $\Phi_1(p,\ r)=\Phi_0(p,\ r)$. Так как из $\Phi_n\geqslant\Phi_{n-1}$ следует $\Phi_{n+1}=T\Phi_n\geqslant T\Phi_{n-1}=\Phi_n$, получаем, что последовательность $\{\Phi_n(p,\ r)\}$ монотонно возрастает по n. Из ограниченности членов нашей последовательности следует, что $\lim\limits_n\Phi_n(p,\ r)=V(p,\ r)$ существует, а тем самым в силу непрерывности оператора T и ограниченное в конечной области решение функционального уравнения (1) существует.

Докажем, что $V(p_0,\ r_0)$ есть значение нашей игры.

Если стратегии обоих игроков на $k+1$-ом шаге смешанные, то накопленные капиталы p_{k+1} и r_{k+1} суть случайные величины. Чтобы подчеркнуть это, мы обозначим их через $\dot p_{k+1}$ и $\dot r_{k+1}$.

Теперь, пусть первый игрок применяет локальную V – стратегию, а второй игрок – любую стратегию. Тогда, по определению максимина, случайная величина $V(\dot p_{k+1},\ \dot r_{k+1})$ должна иметь условное математическое ожидание, удовлетворяющее неравенству

$$E\{V(p_{k+1},\ \dot r_{k+1})\mid V(p_k,\ r_k),\ \ldots,\ V(p_0,\ r_0)\}\geqslant\operatorname{val}\|V(p_k+a_{ij},\ r_k+b_{ij})\|.$$

Так как $V(p,\ r)$ является решением функционального уравнения (1), получаем

$$E\{V(\dot p_{k+1},\ \dot r_{k+1})\mid V(p_k,\ r_k),\ \ldots,\ V(p_0,\ r_0)\}\geqslant V(p_k,\ r_k).\tag{2}$$

Последнее неравенство действительно для всех k. Случайный процесс, удовлетворяющий неравенству (2) называется полумартингалом (см. [4], стр. 286). Для ограниченных полумартингалов $\lim V(p_k, r_k)$ существует с вероятностью 1, и условное математическое ожидание предельного значения удовлетворяет неравенству:

$$E\{\lim_k V(p_k, r_k) | V(p_0, r_0)\} \geqslant V(p_0, r_0). \quad (3)$$

Ограниченность $V(p, r)$ была доказана ранее. Кроме того, $(p_l, r_l) = (p_{l+1}, r_{l+1}) = \ldots = (p^*, r^*)$, как только $(p_l, r_l) \notin L$ и, следовательно, $V(p_l, r_l) = V(p_{l+1}, r_{l+1}) = \ldots = f(p^*, r^*)$ и левую сторону неравенства (3) можно переписать следующим образом:

$$E\{f(p^*, r^*)\} \geqslant V(p_0, r_0),$$

откуда получаем, что, применяя локальную V – стратегию, первый игрок может гарантировать себе средний выигрыш $V(p_0, r_0)$. Аналогично показывается, что, применяя свою локальную V – стратегию, второй игрок может не дать первому получить в среднем больше $V(p_0, r_0)$.

Таким образом, указанные стратегии поведения суть оптимальные стратегии в описанной игре, а $V(p_0, r_0)$ есть значение игры.

Институт физики и математики
Академии наук Литовской ССР

Поступило в редакцию
20.I.1967

ЛИТЕРАТУРА

1. L. S. Shapley, Stochastic games, Proceedings of the National Academy of Sciences, U.S.A., 39, 1095—1100 (1953).
2. J. Milnor, L. S. Shapley, On games of survival Contributions to the theory of games, val. III Princenton, 1957.
3. H. Everett, Recursive games, Contributions to the theory of games, val. III, Princenton, 1957.
4. Дж. Л. Дуб, Вероятностные процессы, ИЛ, М., 1956.

ANTAGONISTINIS DINAMINIS LOŠIMAS, KAI KARTOJAMAS BIMATRICINIS LOŠIMAS

S. VAKRINIENĖ

(Reziumė)

Nagrinėjamas antagonistinis dinaminis lošimas, kuriame kartojamas bimatricinis lošimas. Surasta lošimo reikšmė ir abiejų lošėjų optimalios strategijos.

THE ANTAGONISTIC DYNAMIC GAME BASED ON REPEATING OF THE BIMATRIX GAME

S. VAKRINIENĖ

(Summary)

The zero-sum game, in which bimatrix game is repeated, is investigated. The existence of the value of this game is proved and optimal strategies are given for both players.