Доклады Академии наук СССР
1971. Том 199, № 5

УДК 519.95 *МАТЕМАТИКА*

Ю. В. ЧАЙКОВСКИЙ

О НЕПРЕРЫВНОМ ПРОЦЕССЕ РЕШЕНИЯ МАТРИЧНОЙ ИГРЫ

(Представлено академиком П. С. Новиковым 12 II 1971)

1. Известный метод решения игр Брауна — Неймана, несколько модифицированный, можно интерпретировать ([1]) как процесс игры двух однородных коллективов. Пусть, как и ранее, члены двух бесконечных однородных коллективов играют попарно в матричную игру $A = \|a_{ij}\|$, $i = 1, \ldots, m$, $j = 1, \ldots, n$, $|a_{ij}| \leqslant 1$, и пусть $x = (x_1, \ldots, x_m)$, $y = (y_1, \ldots, y_n)$ — состояния коллективов (зависящие от непрерывного времени t). Введем обозначения: $M(x, y) = xAy$ — средний по первому коллективу выигрыш; $M_i(y) = \sum\limits_j a_{ij} y_j$ — средний выигрыш по тем членам первого коллектива, которые делают i-е действие; $-M_j(x)$ — средний выигрыш по тем членам второго коллектива, которые делают j-е действие; $\varphi_i = \max\{0, M_i(y) - M(x, y)\}$ — выгодность i-го действия; $\psi_j = \max\{0, M(x, y) - M_j(x)\}$ — выгодность j-го действия; x^0, y^0 — оптимальные стратегии; v — цена игры A. Основной результат ([1]) состоял в следующем. Если член первого коллектива переходит за единицу времени от k-го действия к i-му $(i \neq k)$ с вероятностью $P_{ki} = \varphi_i$ (аналогично $P_{lj} = \psi_j$), то

$$c_1 / (t + c_2) \leqslant |M(x(t), y(t)) - v| \leqslant \sum \varphi_i + \sum \psi_j \leqslant c_3 / (t + c_4), \quad c_i > 0, \tag{1}$$

а если $P_{ki} = \varepsilon \bar{\varphi}_i$, где $\bar{\varphi}_i = \operatorname{sgn} \varphi_i$, $0 < \varepsilon \leqslant 1 / \nu$, $\nu = \max\{m - 1, n - 1\}$, то

$$|M(x(t), y(t)) - v| \leqslant \sum \varphi_i + \sum \psi_j \leqslant (m + n - 2) e^{-\varepsilon t}. \tag{2}$$

Таким образом, для оптимального* поведения однородных коллективов в матричной игре достаточно, чтобы каждому члену каждого коллектива на каждом интервале $(t, t + \Delta t)$ сообщался с вероятностью $\varepsilon \Delta t$ один из номеров действий, выгодности которых положительны. Мы покажем, что это условие в определенном смысле необходимо: если члену коллектива сообщается лишь знак выгодности выполняемого им действия, то поведение коллективов неоптимально (п. 2).

Далее, сравнение оценок (1) и (2) показывает, что использование самих выгодностей, а не их знаков, лишь замедляет сходимость. Тем не менее, полезное увеличение используемой информации возможно. Мы опишем эвристический метод, дающий, по-видимому, лучшую сходимость, чем (2), для почти всех игр (п. 3). Это существенно при численном решении игр, поскольку в этом случае ε ограничивается сверху допустимой ошибкой результата.

2. Пусть $\rho_k = \max\{0, M(x, y) - M_k(y)\}$ — невыгодность k-го действия; $\sigma_l = \max\{0, M_l(x) - M(x, y)\}$ — невыгодность l-го действия и пусть вероятности переходов таковы:

$$P_{ki} = \varepsilon f(x_i) \bar{\rho}_k, \qquad \bar{\rho}_k = \operatorname{sgn} \rho_k, \qquad f(\cdot) > 0, \quad i \neq k,$$

$$P_{lj} = \varepsilon f(y_j) \bar{\sigma}_l, \qquad \bar{\sigma}_l = \operatorname{sgn} \sigma_l, \qquad j \neq l,$$

где $f(x_i)$ — предпочтение нового действия в зависимости от его распрост-

* Оптимальность понимается в смысле наличия сходимости к (x^0, y^0).

раненности x_i в коллективе. Теперь так же, как в ([1]), для x, y получим систему

$$\dot{x}_i = \varepsilon\left[f(x_i)\sum_{k=1}^{m} x_k\bar{\rho}_k - x_i\bar{\rho}_i\sum_{k=1}^{m} f(x_k)\right],$$

$$\dot{y}_j = \varepsilon\left[f(y_j)\sum_{l=1}^{n} y_l\bar{\sigma}_l - y_j\bar{\sigma}_j\sum_{l=1}^{n} f(y_l)\right].$$

Система не имеет иных устойчивых точек кроме (x^0, y^0). Эти точки действительно устойчивы при некоторых f, если A имеет седловую точку. Убедиться в этом можно, оценив производную непрерывной функции $\sum x_k\rho_k + \sum y_l\sigma_l$. Однако приведем пример, когда (x^0, y^0) неустойчива по Ляпунову при любой f, дифференцируемой в этой точке. Пусть A кососимметрична, игра вполне смешанна (все $x_i^0 > 0$, $y_j^0 > 0$) и $\sum_i a_{ij} = \sum_j a_{ij} = 0$.

Тогда $x_i^0 = 1 / m$, $y_j^0 = 1 / m$. При этом существует область G такая, что (x^0, y^0) принадлежит замыканию G, а функция $M(x, y)$ удовлетворяет в G условиям теоремы Четаева о неустойчивости ([2]). Область G ограничена условиями: 1) $M(x, \tilde{y}) > 0$; 2) y_j заключены между $1 / m$ и $\tilde{y}_j$; 3) x_i заключены между $1 / m$ и y_i. Смысл условий 2), 3) в том, что при их выполнении производная (вдоль траектории) $M(x, y) > 0$. Тот факт, что G непуста, легко увидеть на примере матрицы 3×3, у которой $a_{ij} = \pm 1$ при $i \neq j$.

3. Пусть теперь

$$P_{ki} = \varepsilon f(x_i)\bar{\varphi}_i, \qquad i \neq k; \qquad P_{lj} = \varepsilon f(y_j)\bar{\psi}_j, \qquad j \neq l.$$

Тогда x, y подчиняются системе

$$\dot{x}_i = \varepsilon(f(x_i)\bar{\varphi}_i - x_i\bar{\beta}), \qquad \dot{y}_j = \varepsilon(f(y_j)\bar{\psi}_j - y_j\bar{\gamma}),$$

где $\bar{\beta} = \sum f(x_i)\bar{\varphi}_i$, $\bar{\gamma} = \sum f(y_j)\bar{\psi}_j$. Для непрерывной функции $\beta + \gamma$, где $\beta = \sum f(x_i)\varphi_i$, $\gamma = \sum f(y_j)\psi_j^*$, равной нулю в точке (x^0, y^0), можно найти

$$\dot{\beta} + \dot{\gamma} = -\varepsilon[\beta\bar{\beta} + \gamma\bar{\gamma} - \sum(f(x_i) - \bar{\beta}x_i)\varphi_i f'(x_i) - \sum(f(y_j) - \\ - \bar{\gamma}y_j)\psi_j f'(y_j)], \tag{3}$$

где f' — производная. Возьмем

$$f(b) = \begin{cases} 1 - b & \text{при } 0 \leqslant b \leqslant {}^1/_2, \\ {}^1/_2 & \text{при } {}^1/_2 \leqslant b \leqslant 1. \end{cases}$$

В этом случае аналогично (2) получим

$${}^1/_2|M(x(t),\ y(t)) - v| \leqslant \beta + \gamma \leqslant (m + n - 2)e^{-\varepsilon t/2},$$

так что экспоненциальная сходимость обеспечена. Если же известно, что все $x_i^0 < {}^1/_2$, все $y_j^0 < {}^1/_2$, то, начав поиск решения с точки $x_i = 1 / m$, $y_i = 1 / n$, можно гарантировать лучшую сходимость, чем сходимость согласно ([2]), поскольку теперь в ([3]) все слагаемые отрицательны. Подобный метод поиска должен, по-видимому, сокращать среднее по всевозможным играм время поиска. Дело в том, что почти все игры имеют единственное решение, причем это решение задается формулами, симметричными относительно элементов a_{ij} ([3]). Поэтому представляется естественным допущение, что с ростом $\mu = \min\{m, n\}$ растет вероятность того, что компоненты x_i^0, y_j^0 (произвольной игры размерности μ) меньше ${}^1/_2$.

В предположении, что $x_i^0 \to 0$, $y_j^0 \to 0$ при $\mu \to \infty$, получаем, что средняя (по играм размерности μ) верхняя интегральная оценка стремится при $\mu \to \infty$ к

$$|M(x(t), y(t)) - v| \leqslant \sum \varphi_i + \sum \psi_j \leqslant (m + n - 2)\, e^{-2\varepsilon t}.$$

Автор весьма признателен А. Л. Лунцу за уточнение формулировки задачи и внимательное обсуждение результатов.

Институт электронных управляющих машин
Москва

Поступило
1 I 1971

ЦИТИРОВАННАЯ ЛИТЕРАТУРА

[1] Ю. В. Чайковский, Экономика и матем. методы, 1971 (в печати). [2] Л. С. Гноенский и др., Математические основы теории управляемых систем, М., 1969. [3] Х. Ф. Боннебласт и др., В сборн. Матричные игры, М., 1961.