Доклады Академии наук СССР
1970. Том 192, № 1

УДК 518.9

МАТЕМАТИКА

А. Ю. ЛЕВИН

ПОВТОРЕНИЕ ИГР ДВУХ ЛИЦ НА БОЛЬШИХ ИНТЕРВАЛАХ ВРЕМЕНИ

(Представлено академиком А. Н. Колмогоровым 21 X 1969)

Теория антагонистических игр двух лиц после Бореля и фон Неймана развивалась и обобщалась в различных направлениях; по этому поводу см., например, ([1]). Предлагаемый ниже подход, по-видимому, не обсуждался, хотя представляется весьма естественным. Суть его, кратко говоря, состоит в учете фактора времени.

Рассмотрим встречу двух игроков, на протяжении которой они проводят между собой ряд партий в одну и ту же матричную игру с нулевой суммой. Каждая партия продолжается положительное время; наряду с платежной матрицей $A = \| a_{ij} \|$ задана также матрица $T = \| t_{ij} \|$, где t_{ij} — средняя продолжительность партии при выборе I и II игроками соответственно i-й и j-й чистых стратегий. Будем предполагать (хотя это, возможно, и не очень существенно), что конечны соответствующие дисперсии σ_{ij}^2 продолжительностей партий. Общее время встречи либо фиксировано заранее, либо, более общо, есть случайная величина, независимая по отношению к ходу игры. (Таким образом, число партий во встрече также есть случайная величина, которая, однако, зависит от хода игры). Нас будет интересовать случай, когда математическое ожидание t времени встречи велико по сравнению со всеми $t_{ij}\sigma_{ij}$. Это позволит игнорировать «краевой эффект» последней партии, которая — если она не окончилась в момент окончания встречи — может в зависимости от регламента либо аннулироваться, либо доигрываться, либо приводить к некоторому частичному платежу. Краевой эффект сильно осложняет нахождение точных оптимальных стратегий; однако, если считать t большим и подходить к делу с асимптотической точки зрения, вкладом единичной партии можно пренебречь.

Предполагается, что оба игрока, как и обычно, стремятся максимизировать гарантированное значение математического ожидания своего выигрыша за встречу. Какие стратегии они должны применять?

Ясно, что стандартные оптимальные по Нейману стратегии для игры с матрицей A здесь, вообще говоря, непригодны, поскольку они не учитывают фактора времени: небольшие частые выигрыши могут оказаться выгоднее крупных, но редких. Имеются два основных случая, когда стратегии, оптимальные в игре с матрицей A, остаются оптимальными (точнее, асимптотически оптимальными, см. ниже) и здесь: а) когда игра с матрицей A безобидна; б) когда все t_{ij} равны. Это важные, но все же частные случаи; во многих реальных игровых ситуациях продолжительность партий сильно колеблется в зависимости от избираемых игроками стратегий.

Может показаться, что задача сводится к решению игры с матрицей $H = \| h_{ij} \|$, где $h_{ij} = a_{ij} / t_{ij}$ суть платежи «на единицу времени». Это верно лишь отчасти. Именно, легко видеть, что если у матрицы H есть седловая точка h_{kl}, то k-я и l-я чистые стратегии игроков I и II соответственно асимптотически оптимальны; при этом I выигрывает за одну встречу в среднем $\sim h_{kl}t$. Но в наиболее интересном случае, когда H не обладает седловой точкой, оптимальные смешанные стратегии игры с матрицей H не являются, вообще говоря, асимптотически оптимальными для рассматриваемой игры.

Способ определения асимптотически оптимальных стратегий в общем случае основан на решении следующего скалярного уравнения:

$$(\text{цена игры с матрицей } A - \lambda T) = 0. \tag{1}$$

Так как все $t_{ij} > 0$, то левая часть (1) строго убывает по λ и уравнение (1), очевидно, имеет единственный корень λ_0. *Стратегии, оптимальные по Нейману для игры с матрицей $A - \lambda_0 T$, являются асимптотически оптимальными для игры, описанной выше; при этом средний выигрыш I игрока за встречу равен $\lambda_0 t + O(1)$, где $O(1)$ мажорируется величиной, не зависящей от t.*

Термин асимптотически оптимальные означает, что применение этих стратегий приводит к среднему выигрышу за встречу, не более чем на $O(1)$ отклоняющемуся от среднего выигрыша при строго оптимальных стратегиях; здесь $O(1)$ опять-таки не зависит лишь от t. Также и в отдельно взятой встрече вероятность выполнения неравенства

$$\left| \frac{\text{выигрыш I}}{t} - \lambda_0 \right| > \varepsilon \quad (\varepsilon > 0 \text{ произвольно})$$

при оптимальном поведении игроков стремится к нулю с ростом t (при фиксированных распределениях платежей и продолжительностей партий). Отметим, что строго оптимальные стратегии не являются, вообще говоря, однородными во времени; для их отыскания (которое, по-видимому, представляет собой трудоемкую задачу) необходимо знать функции распределения времени встречи и отдельных партий, а также регламент последней партии.

Наметим вкратце схему доказательства. Так как игра с матрицей $A - \lambda_0 T$ безобидна, то, согласно сделанному выше замечанию, оптимальные стратегии этой игры асимптотически оптимальны во встрече с платежной матрицей $A - \lambda_0 T$ и матрицей продолжительностей T. При этом средний выигрыш I в такой встрече есть, очевидно, $O(1)$ для сколь угодно больших t. Сравним теперь встречу с матрицами A, T и встречу с матрицами $A - \lambda_0 T, T$. При любом конкретном выборе игроками своих смешанных стратегий в первом случае I получит за встречу в среднем на $\lambda_0 t$ (с точностью до $O(1)$) больше, чем во втором. Действительно, после каждой отдельной партии, кроме незакончившейся, I получает в среднем дополнительно $\lambda_0 t_{ij}$, где i, j — номера реализованных чистых стратегий; но математическое ожидание суммы t_{ij} по всем сыгранным за встречу партиям с точностью до $O(1)$ совпадает с t. Сопоставляя эти соображения (которые допускают аккуратное обоснование), приходим к требуемому.

Уравнение (1) ввиду монотонности левой части можно приближенно решить с помощью «половинения». Начальный отрезок определяется по матрице H согласно легко проверяемому неравенству

$$\max_i \min_j h_{ij} \leqslant \lambda_0 \leqslant \min_j \max_i h_{ij}.$$

Если оптимальные стратегии в игре с матрицей $A - \lambda_0 T$ единственны, то после некоторого числа шагов удастся «нащупать» активные чистые стратегии игроков. Обозначая соответствующие подматрицы A и T через A_0, T_0 (в случае единственности они квадратны), можно после этого найти λ_0 как корень алгебраического уравнения $\det \| A - \lambda T \| = 0$, лежащий в полученном интервале.

За исключением последнего замечания сказанное распространяется и на бесконечные игры. При этом в общем случае следует, естественно, говорить не об (асимптотически) оптимальных, а об (асимптотически) ε-оптимальных стратегиях.

В заключение приведем численный пример. Пусть

$$A = \begin{pmatrix} 1 & -2 & 4 \\ 0 & 3 & -2 \end{pmatrix}, \quad T = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix}.$$

Если I применит во встрече стратегию $(^1/_2;\ ^1/_2)$, оптимальную для него в обычном смысле, т. е. без учета времени, то при лучшем ответе II (2-я стратегия) I выиграет за одну встречу в среднем $\sim {}^1/_7 t$ $(t \geqslant 1)$. В игре с матрицей $H = \left\| \frac{a_{ij}}{t_{ij}} \right\| = {}^1/_{15} \begin{pmatrix} 15 & -15 & 20 \\ 0 & 9 & -5 \end{pmatrix}$ оптимальной по Нейману для I является стратегия $(^2/_7;\ ^5/_7)$; однако при ее применении I будет даже проигрывать — в среднем $\sim {}^1/_{18} t$ за встречу (если II изберет 3-ю стратегию). Решая уравнение (1), находим $\lambda_0 = {}^1/_6 (45 - \sqrt{1929}) \approx 0{,}18$. Таким образом, при асимптотически оптимальных стратегиях I и II, которые приближенно равны (0,47; 0,53) и (0; 0,59; 0,41), математическое ожидание выигрыша I за встречу $\sim 0{,}18t$.

Автор признателен В. М. Гранину и Е. Н. Садовскому за обсуждение.

Воронежский
государственный университет

Поступило
20 X 1969

ЦИТИРОВАННАЯ ЛИТЕРАТУРА

[1] Р. Д. Льюс, Х. Райфа, Игры и решения, ИЛ, 1961.