

ЧИСЛЕННОЕ РЕШЕНИЕ БИМАТРИЧНЫХ ИГР

Предложен эффективный игровой алгоритм поиска равновесия по Нэшу в биматричных играх, основанный на методах линейного программирования и теории двойственности.

Поиск ситуации равновесия по Нэшу биматричной игры в смешанных стратегиях игроков обычно проводится с помощью метода полного перебора [1, 2]. Для этого строятся системы линейных алгебраических уравнений (СЛАУ) вида

$$\left\{ \begin{array}{l} A'q' = ue; \\ p'^T B' = ve'^T; \\ \sum_{i \in I} p'_i = 1; \\ \sum_{j \in J} q'_j = 1 \end{array} \right.$$

относительно неизвестных p' , q' , u , v , в которых векторы e , e' таковы, что имеют все координаты, равные единице, множества значений индексов $I \subset \{1, 2, \dots, m\}$, $J \subset \{1, 2, \dots, n\}$, а A' , B' являются подматрицами платежных матриц игроков A , B , имеющих размер $m \times n$. Здесь m (n) — число чистых стратегий первого (второго) игрока. Пара (p', q') есть искомая ситуация равновесия в смешанных стратегиях лишь при условии того, что указанная СЛАУ имеет неотрицательное решение $p' \geq 0$, $q' \geq 0$. Таким образом, для поиска равновесия игры требуется решить 2^{m+n} систем. Экспоненциальная сложность метода перебора делает его практически неприменимым. В статье предложен алгоритм предположительно полиномиальной сложности, апробированный в вычислительных экспериментах.

Постановка задачи. В биматричной игре [1, 2] каждый игрок располагает конечным множеством чистых стратегий. Игра полностью определяется $m \times n$ матрицами A , B , которые являются табличным заданием функций выигрыша игроков 1 и 2. Игрок 1 выбирает строку i , одновременно с ним игрок 2 выбирает столбец j . После этого игрок 1 получает выигрыш a_{ij} , а игрок 2 — выигрыш b_{ij} .

Смешанная стратегия игрока — это случайная величина, принимающая значения в множестве его чистых стратегий. Смешанную стратегию будем отождествлять с выбираемым игроком 1 (игроком 2) вероятностным распределением $p(q)$: игрок 1 (игрок 2) с вероятностью

$p_i(q_j)$ выбирает свою i -ю (j -ю) чистую стратегию ($i = 1, 2, \dots, m$; $j = 1, 2, \dots, n$).

При использовании смешанных стратегий игроки по возможности стремятся увеличить свои *средние выигрыши*, которые, будучи записанными в матричной форме, имеют вид $p^T Aq$, $p^T Bq$. При этом множества смешанных стратегий игроков образуют симплексы размерности $m - 1$ и $n - 1$ соответственно. Сказанное кратко выразим в форме включений $p \in \Sigma_p$, $q \in \Sigma_q$, где, напомним, симплекс — это $\Sigma_p = \left\{ p : p \geq 0, \sum_{i=1}^m p_i = 1 \right\}$.

Ситуацией *равновесия по Нэшу* в смешанных стратегиях называется пара (p^*, q^*) , удовлетворяющая условию

$$(\forall p, q) p^T Aq^* \leq p^{*T} Aq^*, \quad p^{*T} Bq \leq p^{*T} Bq^*.$$

В игре всегда существует хотя бы одна ситуация равновесия в смешанных стратегиях [1, 2]. Цель настоящей работы — разработка *игрового* алгоритма поиска одного из равновесий. Под этим понимается многошаговая схема, моделирующая процесс игры как поочередный выбор своих стратегий ее участниками, приводящий к равновесию по Нэшу.

Метод поиска равновесия. Поиск равновесия игры сведем к решению следующей многоэкстремальной задачи математического программирования. Требуется найти глобальный максимум целевой функции

$$F(p, u, q, v) = p^T(A + B)q - u - v \rightarrow \max_{p, u, q, v} \quad (1)$$

при ограничениях

$$p^T B \leq ue^T, \quad Aq \leq ve', \quad p \in \Sigma_p, \quad q \in \Sigma_q. \quad (2)$$

Очевидна следующая

Теорема 1. *Глобальный максимум целевой функции (1) на множестве (2), имеющий значение, равное нулю, достигается лишь в точках (p^*, u^*, q^*, v^*) , отвечающих ситуациям равновесия игры (p^*, q^*) , в которых величины v^*, u^* — средние выигрыши игроков 1 и 2 соответственно.*

Сформулированный результат дает основание для рассмотрения следующей игры двух лиц, в которой участники при выборе своих стратегий (p, u) и (q, v) руководствуются целевыми функциями

$$p^T(A + B)q - u \rightarrow \max_{p, u: p^T B \leq ue^T, p \in \Sigma_p} \quad (3)$$

и

$$p^T(A + B)q - v \rightarrow \max_{q, v: Aq \leq ve, q \in \Sigma_q} \quad (4)$$

соответственно.

При любой фиксированной стратегии партнера каждая из функций выигрыша (3), (4) является линейной по переменной (p, u) или (q, v) , представляющей выбираемую этим игроком стратегию. Таким образом, (3), (4) – это взаимосвязанные задачи линейного программирования (ЛП).

Замечание 1. В случае $A + B = 0$ задачи ЛП (3), (4) взаимно двойственны [3]. Как известно [1, 2], их решение дает ситуацию равновесия исходной, в указанном случае, *антагонистической* игры.

Замечание 2. Всякая ситуация равновесия (p', u', q', v') в игре (3), (4) есть точка максимума функции (1) по каждой из переменных $x = (p, u)$ или $y = (q, v)$ в отдельности при фиксированном значении $(y'$ или $x')$ другой переменной.

Теорема 2. *Если в ситуации равновесия (p', u', q', v') обе задачи ЛП (3) ($q = q'$) и (4) ($p = p'$) имеют единственное решение, то точка (p', u', q', v') является локальным максимумом функции (1) на множестве (2) по совокупности переменных.*

Доказательство. Рассмотрим приращение целевой функции (1) в точке (p', u', q', v') . По условию теоремы имеющееся билинейное слагаемое имеет более высокий порядок малости по сравнению с линейными, являющимися приращениями целевых функций (3), (4) в их точках максимума. Поэтому приращение отрицательно при малых вариациях переменных (p, u, q, v) . Следовательно, по любому допустимому направлению функция (1) локально убывает.

Равновесия по Нэшу в игре (3), (4) будем называть *стационарными* точками целевой функции F (ввиду теоремы 2 и замечания 2).

Игровая задача (3), (4) намного проще исходной, так как в ней игроки приходят к равновесию, поочередно делая свои “ходы”, начиная игру из произвольной начальной ситуации.

Теорема 3. *Пусть вектор p^{t+1} – решение задачи (3), $q = q^t$, q^{t+1} – решение задачи (4), $p = p^t$, $t = 0, 1, \dots$, причем элемент $q^0 \in \Sigma_q$ выбран произвольно. Тогда $p^t \rightarrow p'$, $q^t \rightarrow q'$, $t \rightarrow \infty$, где векторы (p', q') определяют некоторую ситуацию равновесия в игре (3), (4).*

Доказательство. Значения целевой функции (1) $F(p^{t+1}, u^{t+1}, q^t, v^t)$, $t = 0, 1, \dots$, по построению образуют монотонно возрастающую, а следовательно, сходящуюся последовательность. Ввиду компактности множеств смешанных стратегий игроков можно считать, что последовательность $\{p^t, q^t\}$ также сходится [4]. Предельный переход в условиях задач ЛП (3), (4) показывает, что пара (p', q') служит решением задач ЛП (3), (4), т.е. является равновесием.

Через λ, μ обозначим двойственные переменные в экстремальной задаче (1), (2), с помощью которых снимаются ограничения [5] в форме

неравенств в условиях (2). Введем функцию Лагранжа, которая по определению имеет вид [5]

$$L = F(p, q, u, v) + (ue^T - p^T B)\lambda + \mu(v e' - Aq).$$

Множители Лагранжа λ , μ являются также двойственными переменными в задачах ЛП (3), (4). Исходя из двойственных к ним задач ЛП [3], несложно убедиться, что двойственные переменные удовлетворяют ограничениям $\mu \in \Sigma_p$, $\lambda \in \Sigma_q$.

Для упрощения записи далее будем опускать переменные u, v и упоминать лишь о паре переменных (p, q) , говоря об оптимизации функции (1) на множестве (2). Пусть (p^*, q^*) — стационарная точка функции (1), а λ^* , μ^* — соответствующие двойственные переменные в задачах ЛП (3) ($p = p^*$) и (4) ($q = q^*$).

Согласно теореме 2 и замечанию 2, итерационный процесс из теоремы 3, сходясь к стационарной точке функции F , вообще говоря, не дает решения исходной игры. Для того чтобы выходить из областей притяжения стационарных точек функции (1) с целью попадания в искомый глобальный максимум, будем использовать следующее свойство экстремальной задачи (1), (2).

Теорема 4. Ситуация (p^*, q^*) является равновесием по Нэшу в исходной игровой задаче тогда и только тогда, когда в стационарной точке целевой функции (1) $p^* = \mu^*$, $q^* = \lambda^*$.

Доказательство. Ситуация равновесия (p^*, q^*) является точкой глобального экстремума F (см. теорему 1). В ней значения функций Лагранжа L и F совпадают [5], поэтому с учетом включений $\mu \in \Sigma_p$, $\lambda \in \Sigma_q$ получим

$$0 = F = L = (p^* - \mu^*)^T Aq^* + p^{*T} B(q^* - \lambda^*).$$

Отсюда следует, что нулевой уровень целевой функции достигается при выборе множителей Лагранжа $\mu^* = p^*$, $\lambda^* = q^*$.

Обратное утверждение доказывается аналогичными рассуждениями, применяемыми к точкам экстремума целевых функций в задачах ЛП (3), (4). В случае задачи (3), $q^* = \lambda^*$, имеем

$$p^{*T}(A + B)q^* - u^* = p^{*T}(A + B)q^* - u^* + (u^* e^T - p^{*T} B)\lambda^* = p^{*T} Aq^*.$$

Для задачи (4) получим $p^{*T} Bq^*$. Сложив найденные значения и вычтя величину, которая дважды вошла в эту сумму, вычислим значение целевой функции (1) в рассматриваемой точке:

$$F = p^{*T} Aq^* + p^{*T} Bq^* - p^{*T}(A + B)q^* = 0.$$

Применение теоремы 1 завершает доказательство.

Замечание 3. Игра (3), (4) является матричной и поэтому имеет конечное множество равновесий по Нэшу. В достаточном условии теоремы 4 можно потребовать выполнения лишь одного из равенств $p^* = \mu^*$ или $q^* = \lambda^*$, так как другое в этом случае обязано выполняться.

Докажем это. Пусть для определенности $q^* = \lambda^*$. Тогда по условию дополняющей нежесткости

$$(\forall j)(p^{*T}B - u^*)_j q_j^* = 0$$

закключаем, что

$$\max_{q \in \Sigma_q} p^* B q = u^*. \quad (5)$$

Оптимизация (3) эквивалентна поиску максимума функции Лагранжа задачи (3), т.е. [5]

$$p^T(A+B)q^* - u + (ue^T - p^T B)q^* \equiv p^T A q^* \rightarrow \max_{p \in \Sigma_p}.$$

Так как решение этой задачи есть точка p^* , то вместе с (5) это доказывает равновесность пары стратегий (p^*, q^*) в исходной игре. По теореме 1 и необходимому условию из теоремы 4 заключаем, что $p^* = \mu^*$.

Алгоритм поиска равновесия. В соответствии с теоремами 3, 4 и замечанием 3 предлагается использовать следующую многошаговую схему поиска равновесия.

Этап 0. Один из игроков, пусть это игрок 2, первым выбирает одну из своих смешанных стратегий $q^0 \in \Sigma_q$ и сообщает ее игроку 1.

Этап 1. Оба игрока, начиная с игрока 1, поочередно выбирают свои оптимальные (в текущих ситуациях) стратегии в игре (3), (4) и сообщают их партнеру. По достижении некоторой ситуации равновесия (p', q') перейти к выполнению этапа 2.

Этап 2. Если $q' \neq \lambda'$, то один из игроков, пусть это игрок 2, изменяет свое решение — вместо стратегии q' применяет стратегию λ' и сообщает ее игроку 1. Перейти к действиям этапа 1.

На этапе 1 разворачивается итерационный процесс из теоремы 3. Если проверяемое на этапе 2 условие не выполнено, то ситуация (p', q') есть искомое равновесие по Нэшу исходной игры (замечание 3).

Вычислительный пример. Пусть $m = n = 4$, а платежи игроков представлены матрицей $((a_{ij}, b_{ij}))$ вида

$$\begin{pmatrix} (1, 3) & (1, 2) & (0, 1) & (3, 1) \\ (3, 2) & (2, 3) & (2, 0) & (0, 1) \\ (2, 3) & (1, 0) & (1, 2) & (3, 1) \\ (1, 0) & (3, 2) & (0, 2) & (2, 3) \end{pmatrix}.$$

Этап 0. Возьмем $q^0 = (0, 0,5, 0,5, 0)$.

Этап 1. Вычисления дают ситуацию (p', q') : $p' = (0, 0,33, 0, 0,67)$, $q' = (0, 0,67, 0,33, 0)$, в которой $F = -0,33$. Двойственные переменные $\lambda' = (0, 0,67, 0, 0,33)$, $\mu' = (0, 0, 0, 1)$.

Этап 2. Заменяем стратегию q' на λ' (при этом изменилось значение целевой функции на $F = -0,45$) и выполним вычисления этапа 1.

В результате найдем искомое равновесие

$$p^* = (0,33, 0, 0, 0,67), \quad q^* = (0, 0,33, 0, 0,67), \quad v^* = 2,33, \quad u^* = 2.$$

Всего потребовалось решить четыре задачи ЛП.

Заключение. Задачи ЛП имеют полиномиальную сложность решения [6]. Это дает основание предположить, что предложенный в работе алгоритм имеет такую же вычислительную сложность.

СПИСОК ЛИТЕРАТУРЫ

1. Петросян Л. А., Зенкевич Н. А., Семина Е. А. Теория игр. – М.: Высш. шк., 1998.
2. Волков И. К., Загоруйко Е. А. Исследование операций. – М.: Изд-во МГТУ им. Н.Э. Баумана, 2004.
3. Ашманов С. А. Линейное программирование. – М.: Наука, 1981.
4. Васильев Ф. П. Численные методы решения экстремальных задач. – М.: Наука, 1980.
5. Иоффе А. Д., Тихомиров В. М. Теория экстремальных задач. – М.: Наука, 1974.
6. Хачян Л. Г. Полиномиальный алгоритм в линейном программировании // ДАН СССР. – 1979. – Т. 244. – С. 1093–1096.

Статья поступила в редакцию 24.12.2007

Николай Семенович Васильев окончил МГУ им. М.В. Ломоносова в 1974 г. Д-р физ.-мат. наук, профессор кафедры “Высшая математика” МГТУ им. Н.Э. Баумана. Автор более 50 научных работ в области теории оптимального управления и моделирования распределенных телекоммуникационных систем.

N.S. Vasiliev graduated from the Lomonosov Moscow State University in 1974. D. Sc. (Phys.-Math.), professor of “Higher Mathematics” department of the Bauman Moscow State Technical University. Author of more than 50 publications in the field of theory of optimal control and simulation of distributed telecommunication systems.