

УДК 519.626:519.832.2

О МЕТОДЕ ФИКТИВНЫХ НЕИЗВЕСТНЫХ ДЛЯ ЧИСЛЕННОГО РЕШЕНИЯ МАТРИЧНЫХ ИГР

Е. В. Чижонков¹

Предложен новый подход к решению симметричных матричных игр, использующий введение фиктивных неизвестных. Показано, что на этой основе специализированными алгоритмами можно определять как частные оптимальные стратегии, так и решения минимальной длины. Проведенные расчеты демонстрируют вычислительную эффективность подхода для игр умеренной размерности. Работа выполнена при частичной финансовой поддержке РФФИ (код проекта 09-01-00625а).

Ключевые слова: симметричные матричные игры, фиктивные неизвестные, задача наименьших квадратов, итерационные методы, вариационные неравенства, решение минимальной длины.

Введение. С точки зрения практического решения матричных игр, в настоящее время доминирующим подходом является их сведение к задачам линейного программирования с последующим решением либо методом внутренней точки, либо симплекс-методом [1, 2]. Однако развитие теоретических представлений о приближенном решении игровых задач не ограничивается только этим направлением. Например, продолжает развиваться классический алгоритм фиктивного розыгрыша, известный как метод Брауна–Робинсон [3, 4]. Достаточно регулярно появляются работы, посвященные его модификациям, учитывающим специфику различных классов матриц [5, 6]. При этом, конечно, следует отметить, что происходит не только развитие известных подходов, но и периодическое обновление взглядов на традиционные постановки. Например, в работе [7] был предложен новый подход к приближенному решению матричных игр, основанный на сведении исходной задачи к вариационному неравенству специального вида. В частности, это позволило сконструировать и обосновать предобусловленные итерационные методы, аналогичные используемым при численном решении больших и плохо обусловленных систем линейных алгебраических уравнений.

В настоящей статье рассмотрено развитие указанного подхода, а именно: продемонстрировано, что одновременное приближенное определение как неизвестных из исходной постановки, так и их дополнения в виде специального набора фиктивных неизвестных, может быть вполне эффективно с вычислительной точки зрения реализовано для матричных игр умеренной размерности (порядка 10^2 – 10^3).

Статья организована следующим образом. Сначала симметричная матричная игра сводится к эквивалентной постановке определения некоторого вектора, компоненты которого состоят одновременно из компонент исходного и фиктивного векторов неизвестных.

Затем изучается решение сформулированной задачи путем сведения ее к конструктивно вычисляемому решению задачи наименьших квадратов с условием неотрицательности (NNLS — Non-Negative Least Squares). Этот процесс иллюстрируется численными экспериментами с матричной игрой, имеющей единственное решение.

Далее описаны примеры алгоритмов решения матричных игр, которые используют фиктивные неизвестные и основаны на методах решения сеточных вариационных неравенств. Там же приведены результаты расчетов, демонстрирующие не только обоснованность, но и практическую работоспособность такого рода алгоритмов.

Наконец, обсуждается применение фиктивных неизвестных для выделения единственного решения минимальной длины сначала эквивалентной задачи с расширенным набором неизвестных, а затем — только исходной матричной игры. С этой целью строится специальная регуляризация эквивалентной задачи, для решения которой используется известный алгоритм NNLS [8].

Теоретические обсуждения сопровождаются расчетами для специально сконструированной модельной игровой задачи, имеющей однопараметрическое семейство решений. В заключение кратко систематизируются полученные результаты и определяются возможные направления дальнейших исследований.

¹ Московский государственный университет им. М. В. Ломоносова, механико-математический факультет, Ленинские горы, 119899, Москва; профессор, e-mail: chizhonk@mech.math.msu.su

1. Постановка задачи. Рассмотрим симметричную матричную игру G в смешанных стратегиях:

$$\sum_{j=1}^r d_{ij}y_j \leq 0, \quad i = 1, 2, \dots, r, \tag{1}$$

$$y_j \geq 0, \quad j = 1, 2, \dots, r, \tag{2}$$

$$\sum_{j=1}^r y_j = 1, \tag{3}$$

где d_{ij} — элементы заданной кососимметрической матрицы $D = -D^T$, называемой, как правило, матрицей выигрышей. Хорошо известно, что решение $\mathbf{y} = (y_1, \dots, y_r)^T$ игры G всегда существует, но не обязано быть единственным; часто это утверждение для произвольной матричной игры называют теоремой Неймана, или основной теоремой матричных игр (см., например, [9, 10]). Будем предполагать также, что G не имеет решений в чистых стратегиях, что делало бы постановку элементарной; в рассматриваемом случае отсюда следует, что в матрице D отсутствуют нулевые строки (столбцы).

Отметим, что рассмотрение процесса решения только симметричных игр не ограничивает общности подхода, так как для произвольной игры с $(m \times l)$ -матрицей существуют ее различные симметричные расширения [10]. Более того, если одновременно существуют решения прямой и двойственной задач линейного программирования, то они могут быть решены сведением именно к симметричной матричной игре [11]. Из вышесказанного следует, что приоритетным направлением для разработки приближенных методов решения матричных игр представляются симметричные игры.

Обозначим через $K_s = \{\mathbf{x} \in \mathbb{R}^s, x_i \geq 0, i = 1, 2, \dots, s\}$ неотрицательный ортант в пространстве \mathbb{R}^s и будем называть его далее множеством ограничений. В частности, неравенства (2) означают, что $\mathbf{y} \in K_r$. Введем полезное

Определение. Две задачи будем называть эквивалентными, если множества их решений связаны взаимно однозначным отображением.

Весьма содержательным примером эквивалентности задач является симметризация матричной игры, более элементарным — аффинное преобразование матрицы выигрышей. Воспользуемся последним для нормировки матрицы D в (1): будем считать, что ее элементы умножены на подходящую положительную константу (реально, поделены на максимальный модуль) так, что теперь $d_{ij} \in [-1, 1]$. Такое преобразование не влияет на теоретические рассуждения и не меняет решения задачи, однако оно весьма полезно для вычислительной устойчивости матрично-векторных умножений. В дальнейшем постановку (1)–(3) с $|d_{ij}| \leq 1$ будем называть **задачей I**.

2. Фиктивные неизвестные. Введем фиктивные неизвестные $z_i \geq 0, i = 1, 2, \dots, r$, в неравенства (1) таким образом, чтобы превратить их в равенства $\sum_{j=1}^r d_{ij}y_j + z_i = 0, i = 1, 2, \dots, r$, матричная форма которых имеет вид

$$D\mathbf{y} + E\mathbf{z} = 0. \tag{4}$$

Здесь и далее будем обозначать через E единичную матрицу соответствующего размера, т.е. матрицу с элементами $e_{ii} = 1$ при $i = j$ и $e_{ij} = 0$ при $i \neq j$ (в данном случае ее размер равен $r \times r$); через \mathbf{z} обозначен вектор фиктивных неизвестных.

Будем называть **задачей II** следующую постановку: найти пару векторов $(\mathbf{y}; \mathbf{z})$, для которых справедливо

$$D\mathbf{y} + E\mathbf{z} = 0, \quad \mathbf{y} \in K_r, \quad \mathbf{z} \in K_r, \tag{5}$$

и, кроме того, для вектора \mathbf{y} выполнено условие нормировки (3).

Легко заметить, что задачи I и II эквивалентны. Действительно, для произвольного решения задачи I конструктивно вычисляются значения фиктивных переменных; с другой стороны, из произвольного решения задачи II решение задачи I получается их простым отбрасыванием.

Отметим, что использование неотрицательных фиктивных переменных встречалось и ранее. Например, в [12] описан способ сведения матричной игры с произвольной ценой к задаче линейного программирования. Настоящая же работа в первую очередь посвящена тому, что одновременное нахождение и решения, и соответствующих ему фиктивных переменных представляется более удобным и содержательным с вычислительной точки зрения.

3. Наименьшие квадраты с условием неотрицательности. Пусть C, \mathbf{x} , и \mathbf{f} — соответственно $(m \times n)$ -матрица, n -вектор и m -вектор. Рассмотрим следующую задачу: найти минимум $\|C\mathbf{x} - \mathbf{f}\|$ при

условии $\mathbf{x} \geq 0$. Следуя [8], далее будем называть ее задачей NNLS, а для обычной задачи наименьших квадратов (без ограничений по знаку) примем обозначение $C\mathbf{x} \cong \mathbf{f}$. Для отыскания численного решения поставленной задачи имеется хорошо известный алгоритм с одноименным названием, для которого доказана сходимость за конечное число шагов [8]. Приведем его для полноты изложения.

Все входные данные (вектор \mathbf{f} и матрица C) являются на входе определенными. Дополнительно требуются рабочие векторы \mathbf{w} и \mathbf{v} длины n . В ходе выполнения алгоритма также определяются индексные множества \mathcal{L} и \mathcal{P} . Переменные с индексами, принадлежащими множеству \mathcal{L} , имеют только нулевые значения. С другой стороны, переменные с индексами, принадлежащими множеству \mathcal{P} , могут принимать ненулевые значения так, как это определяется решением задачи. Если значение такой переменной неположительно, то алгоритм либо изменит его на положительное, либо, установив нулевое значение, переместит соответствующий индекс из множества \mathcal{P} в множество \mathcal{L} . На выходе алгоритма вектор \mathbf{x} будет содержать решение, а вектор \mathbf{w} — двойственный к нему.

Алгоритм NNLS ($C, m, n, \mathbf{f}, \mathbf{x}, \mathbf{w}, \mathbf{v}, \mathcal{P}, \mathcal{L}$):

1. Положить $\mathcal{P} := \emptyset$, $\mathcal{L} := \{1, 2, \dots, n\}$, $\mathbf{x} := 0$.
2. Вычислить n -вектор $\mathbf{w} := C^T(\mathbf{f} - C\mathbf{x})$.
3. Если множество \mathcal{L} пусто или $w_j \leq 0$ для всех $j \in \mathcal{L}$, то перейти к шагу 12.
4. Найти индекс $t \in \mathcal{L}$, такой, что $w_t = \max\{w_j : j \in \mathcal{L}\}$.
5. Переместить индекс t из множества \mathcal{L} в множество \mathcal{P} .
6. Определить матрицу $C_{\mathcal{P}}$ размерности $m \times n$ следующим образом:

$$\text{столбец } j \text{ матрицы } C_{\mathcal{P}} := \begin{cases} \text{столбец } j \text{ матрицы } C, & \text{если } j \in \mathcal{P}, \\ 0, & \text{если } j \in \mathcal{L}. \end{cases}$$

Вычислить n -вектор \mathbf{v} как решение задачи наименьших квадратов $C_{\mathcal{P}}\mathbf{v} \cong \mathbf{f}$. Заметим, что эта задача определяет только компоненты v_j , где $j \in \mathcal{P}$, поэтому следует положить $v_j := 0$ для $j \in \mathcal{L}$.

7. Если $v_j > 0$ для всех $j \in \mathcal{P}$, то положить $\mathbf{x} := \mathbf{v}$ и перейти к шагу 2.
8. Найти индекс $q \in \mathcal{P}$, такой, что $x_q/(x_q - v_q) = \min\{x_j/(x_j - v_j) : v_j \leq 0, j \in \mathcal{P}\}$.
9. Положить $\alpha := x_q/(x_q - v_q)$.
10. Положить $\mathbf{x} := \mathbf{x} + \alpha(\mathbf{v} - \mathbf{x})$.
11. Переместить из множества \mathcal{P} в множество \mathcal{L} все индексы $j \in \mathcal{P}$, для которых $x_j = 0$. Перейти к шагу 6.
12. Комментарий: вычисления закончены.

Полученный после применения алгоритма NNLS вектор \mathbf{x} удовлетворяет соотношениям

$$x_j > 0, \quad j \in \mathcal{P}, \quad (6)$$

$$x_j = 0, \quad j \in \mathcal{L}, \quad (7)$$

и является решением задачи наименьших квадратов

$$C_{\mathcal{P}}\mathbf{x} \cong \mathbf{f}. \quad (8)$$

Двойственный вектор \mathbf{w} удовлетворяет соотношениям

$$w_j = 0, \quad j \in \mathcal{P}, \quad (9)$$

$$w_j \leq 0, \quad j \in \mathcal{L}, \quad (10)$$

$$\mathbf{w} := C^T(\mathbf{f} - C\mathbf{x}). \quad (11)$$

Соотношения (6), (7), (9)–(11) — это условия Куна–Таккера, характеризующие решение \mathbf{x} задачи NNLS. Соотношение (8) является следствием (7), (9) и (11).

Построим задачу NNLS, эквивалентную задаче II. Обозначим через \mathbf{x} вектор, объединяющий решение матричной игры \mathbf{y} и фиктивные неизвестные \mathbf{z} так, что $\mathbf{x} = (\mathbf{y}; \mathbf{z}) \in \mathbb{R}^n$, $n = 2r$, а также определим вектор \mathbf{f} , играющий в данном случае нормирующую роль: $\mathbf{f} = (0, 0, \dots, 0, 1)^T \in \mathbb{R}^m$, $m = r + 1$. Теперь матрица C для искомой задачи будет выглядеть следующим образом:

$$C = \left(\begin{array}{c|c} D & E \\ \hline 1 \ 1 \ \dots \ 1 & 0 \ 0 \ \dots \ 0 \end{array} \right), \quad (12)$$

где вертикальные и горизонтальные черты для наглядности отделяют соответствующие блоки.

Будем называть **задачей III** задачу NNLS с указанными выше матрицей C и векторами \mathbf{x} , \mathbf{f} .

Легко видеть, что задачи II и III эквивалентны. Действительно, для произвольного решения задачи II заключаем, что оно удовлетворяет задаче NNLS с величиной $\rho(\mathbf{x}) = \|\mathbf{C}\mathbf{x} - \mathbf{f}\|$, равной нулю. Отсюда, в частности, следует существование такого решения задачи III. С другой стороны, произвольное решение задачи III с минимально возможным $\rho(\mathbf{x}) = 0$ (а такое имеется, см. выше) в точности является решением задачи II, так как $(r+1)$ -е соотношение задачи NNLS, связанное с последней строкой матрицы \mathbf{C} и последней компонентой вектора \mathbf{f} , определяет условие нормировки (3) исходной матричной игры, а предыдущие r соотношений равносильны равенству (5).

Приведем описание вычислительного эксперимента, основанного на приближенном решении задачи III, с помощью пакета подпрограмм lawson-hanson [13], разработанного в соответствии с монографией [8]. Здесь удобно воспользоваться задачей с известным решением, причем вполне удовлетворительной для этого является формулировка матричных игр Мендельсона [14], решение которых не зависит от размерности матрицы выигрышей. Напомним, что в этом случае кососимметрическая матрица \mathbf{D} с элементами d_{ij} , $1 \leq i, j \leq r$, из (1) определяется при $i > j$ так: если $i = j + 1$, то $d_{ij} = 1/2$; если $i \geq j + 2$, то $d_{ij} = -1$. В отличие от оригинальной постановки в приведенных формулах учтена нормировка $|d_{ij}| \leq 1$. Легко проверить, что единственным решением такой матричной игры будет вектор $\mathbf{y} = (1/4, 1/2, 1/4, 0, 0, \dots, 0)^T \in \mathbb{R}^r$; при этом вектор фиктивных переменных за исключением первых трех компонент будет отличным от нуля. Может также оказаться полезной игра с матрицей $\hat{\mathbf{D}} = -\mathbf{D}$, в этом случае ее решением будет инверсия вектора \mathbf{y} , т.е. вектор $\hat{\mathbf{y}} = (0, 0, \dots, 0, 1/4, 1/2, 1/4)^T \in \mathbb{R}^r$ с соответствующим изменением вектора \mathbf{z} . Заметим, что в наших расчетах результаты для обеих игр оказались примерно одинаковыми.

В табл. 1 приведены в зависимости от размерности r времена счета T в секундах для получения приближенного решения задачи III на настольном компьютере весьма ограниченной производительности (Pentium IV с тактовой частотой 3.0 ГГц).

Таблица 1

r	250	500	1000	2000
T	0.418	3.34	26.3	207.0

Эти расчеты иллюстрируют два примечательных эффекта:

- 1) доступность за разумное время решения матричной игры умеренной размерности ($r \leq 2000$),
- 2) предельно низкий показатель в асимптотике вычислительных затрат (порядка $O(r^3)$).

Значимость первого из них обусловлена простотой идеи и вычислительной устойчивостью алгоритма NNLS, причем по форме метод итерационный, а по сути — прямой (конечно, в предположении отсутствия ошибок округлений). Второй эффект, несмотря на большую привлекательность, скорее всего объясняется спецификой модельной задачи. Можно заметить, что алгоритм NNLS в качестве оценки снизу требует конечного числа итераций с вычислительными затратами на каждой, эквивалентными решению обычной задачи наименьших квадратов, что суммарно приводит к асимптотике не ниже $O(r^3)$. При этом обсуждение сходимости в [8] совершенно не гарантирует даже полиномиальных затрат метода с произвольным показателем. Тем не менее, отмеченная выше привлекательность подхода для решения матричных игр умеренной размерности вполне позволяет конкурировать с симплекс-методом [2], используемым для эквивалентных задач линейного программирования.

4. Вариационные неравенства и включения. Запишем матрицу из (5) в виде $(D | E)$ (здесь вертикальная черта разделяет блоки — матрицы D и E) и проведем ее симметризацию:

$$A = (D | E)^T(D | E) \equiv \begin{pmatrix} D^T D & D^T E \\ D & E \end{pmatrix}. \tag{13}$$

Определим задачу IV как задачу отыскания вектора $\mathbf{x} = (\mathbf{y}; \mathbf{z}) \in \mathbb{R}^{2r}$, удовлетворяющего ограничениям

$$A\mathbf{x} = 0, \quad \mathbf{x} \in K_{2r}, \tag{14}$$

и дополнительному условию нормировки (3) для его первой составляющей \mathbf{y} .

Легко видеть, что задачи I и IV действительно эквивалентны (см. формальное доказательство в [7]), так как:

1) произвольное решение \mathbf{y} исходной матричной игры конструктивно (и поэтому однозначно) порождает соответствующее решение \mathbf{x} ;

2) в обратную сторону — ядра матриц A и $(D | E)$ совпадают и, кроме того, соотношения (14) не допускают нетривиальных решений $\mathbf{x} = (\mathbf{y}; \mathbf{z})$ с тождественно нулевой первой компонентой \mathbf{y} (так как из $\|\mathbf{y}\| = 0$ необходимо следует $\|\mathbf{z}\| = 0$).

Из сказанного следует важное замечание, что решение задачи IV удобно находить в два этапа: сначала определить какое-либо нетривиальное решение $(\tilde{\mathbf{y}}; \tilde{\mathbf{z}})$ соотношений (14), а затем отнормировать его по

правилу $\mu = \sum_{i=1}^r \tilde{y}_i \neq 0$, $\mathbf{y} = \tilde{\mathbf{y}}/\mu$, $\mathbf{z} = \tilde{\mathbf{z}}/\mu$.

Перейдем к вариационной формулировке. Пусть $n = 2r$. Обозначим $J(\mathbf{x}) = \frac{1}{2}(A\mathbf{x}, \mathbf{x}) + I_K(\mathbf{x})$, $\mathbf{x} \in \mathbb{R}^n$, $I_K(\mathbf{x})$ — индикаторная функция выпуклого замкнутого множества K : $I_K(\mathbf{x}) = \begin{cases} 0 & \text{при } \mathbf{x} \in K, \\ +\infty & \text{при } \mathbf{x} \notin K. \end{cases}$

Имеет место следующее полезное утверждение (частный случай леммы 1.6 из [15]): пусть $A = A^T \geq 0$, тогда эквивалентны задачи

$$\mathbf{x}^* = \arg \min_{\mathbf{x} \in \mathbb{R}^n} J(\mathbf{x}), \quad (15)$$

$$\mathbf{x}^* \in \mathbb{R}^n : (A\mathbf{x}^*, \mathbf{x} - \mathbf{x}^*) + I_K(\mathbf{x}) - I_K(\mathbf{x}^*) \geq 0 \quad \forall \mathbf{x} \in \mathbb{R}^n, \quad (16)$$

$$A\mathbf{x}^* + P(\mathbf{x}^*) \ni 0, \quad (17)$$

где $P(\mathbf{x}) = \{\mathbf{z} \in \mathbb{R}^n : (\mathbf{z}, \mathbf{y} - \mathbf{x}) \leq 0 \quad \forall \mathbf{y} \in K\}$ — субдифференциал $\partial I_K(\mathbf{x})$.

В нашем случае множество ограничений — неотрицательный ортант: $K = \prod_{i=1}^n [0, +\infty)$.

Отсюда следует, что $I_K(\mathbf{x})$ — сепарабельная функция и, соответственно, $\partial I_K(\mathbf{x})$ — диагональный и монотонный оператор, причем обратный к нему оператор также имеет диагональную структуру и определяется как $(\partial I_K)^{-1} = \text{diag}(H_{[0,+\infty)}, \dots, H_{[0,+\infty)})$, где $H_{[0,+\infty)}(x)$ — функция Хевисайда:

$$H_{[0,+\infty)}(x) = \begin{cases} 0 & \text{при } x < 0, \\ [0, +\infty) & \text{при } x = 0. \end{cases}$$

Таким образом, кроме задачи IV дополнительно имеем три эквивалентные формулировки (15)–(17), причем выбирать для решения можно любую, руководствуясь исключительно соображениями вкуса. В нашем случае (итерационного решения) представляется наиболее удобным иметь дело с включением (17): $A\mathbf{x} + P(\mathbf{x}) \ni 0$, где $A = A^T \geq 0$, $P = \partial I_K$, K — выпуклое замкнутое множество ограничений K_n . Этот выбор связан как с обоснованием соответствующих алгоритмов, так и со способами их реализации, детали которых можно уточнить в [7, 15, 16]. Отметим, что наиболее полезной в данном случае является монография [15], в которой указанные вопросы освещены с достаточной строгостью и необходимой полнотой.

5. Итерационные методы. Для решения (17) сначала рассмотрим простейшие алгоритмы релаксационного типа. Определим матрицу $D_A = \text{diag}(a_{11}, \dots, a_{nn})$, тогда точечный экстраполированный (иногда говорят — релаксированный) метод Якоби имеет вид

$$D_A \frac{\mathbf{x}^{k+1} - \mathbf{x}^k}{\tau} + A\mathbf{x}^k + P(\mathbf{x}^{k+1}) \ni 0. \quad (18)$$

Для нашей постановки алгоритм (18) сходится с произвольного начального приближения $\mathbf{x}^0 \in \mathbb{R}^n$ к некоторому решению неравенства (16) при любом $\tau \in (0, 2/\Delta)$, где постоянная Δ определяется из матричного неравенства $A \leq \Delta D_A$. В данном случае допускается прямое решение включения, так как вспомогательная задача

$$D_A \mathbf{v} + \tau P(\mathbf{v}) \ni \mathbf{g} \quad (19)$$

с известным вектором \mathbf{g} , $\tau > 0$ и монотонным диагональным оператором $P(\mathbf{v}) = (p_1(v_1), \dots, p_n(v_n))^T$ распадается на n скалярных соотношений

$$a_{ii}v_i + \tau p_i(v_i) \ni g_i, \quad i = 1, 2, \dots, n, \quad (20)$$

которые можно реализовать элементарными вычислениями.

Действительно, из (20) следует $v_i = \begin{cases} g_i/a_{ii} & \text{при } g_i > 0, \\ 0 & \text{при } g_i \leq 0. \end{cases}$

Отметим, что все $a_{ii} > 0$, $i = 1, 2, \dots, n$, что следует из отсутствия нулевых строк (столбцов) в матрице D исходной симметричной игры G .

Рассмотрим метод последовательной верхней релаксации (SOR — Successive Over-Relaxation). Представим матрицу A в следующем виде: $A = D_A + L_A + L_A^T$, где матрица D_A определена выше, а L_A —

строго нижняя треугольная подматрица A . Выбирая теперь предобусловливатель в виде $B = D_A + \tau L_A$, получаем метод SOR

$$B \frac{\mathbf{x}^{k+1} - \mathbf{x}^k}{\tau} + A\mathbf{x}^k + P(\mathbf{x}^{k+1}) \ni 0. \tag{21}$$

Аналогично классической версии для систем линейных алгебраических уравнений, в случае нашей постановки алгоритм (21) сходится с произвольного начального приближения $\mathbf{x}^0 \in \mathbb{R}^n$ к некоторому решению неравенства (16) при любом $\tau \in (0, 2)$.

Здесь также допускается прямое решение включения, так как вспомогательная задача $B\mathbf{v} + \tau P(\mathbf{v}) \ni \mathbf{g}$ с треугольной матрицей B и диагональным оператором P решается рекуррентно. Так как $b_{ij} = 0$ для $j > i$, то последовательно решаются одномерные задачи

$$b_{ii}v_i + \tau p_i(v_i) \ni g_i - \sum_{j < i} b_{ij}v_j, \quad i = 1, 2, \dots, n.$$

Аналогично (20) имеем $b_{ii} = a_{ii} > 0, i = 1, 2, \dots, n$, поэтому приведенная формула не содержит сингулярностей.

Рассмотрим более сложный алгоритм типа Дугласа–Рэкфорда:

$$B \frac{\mathbf{x}^{k+1} - \mathbf{x}^k}{\tau} + A\mathbf{x}^k + P\left(B(\mathbf{x}^{k+1} - \mathbf{x}^k) + \mathbf{x}^k\right) \ni 0, \tag{22}$$

где $B = E + \tau A$; он сходится с произвольного начального приближения $\mathbf{x}^0 \in \mathbb{R}^n$ к некоторому решению неравенства (16) при любом $\tau > 0$. Численная реализация метода (22) может быть осуществлена с помощью процедуры расщепления

$$\frac{\mathbf{x}^{k+1/2} - \mathbf{x}^k}{\tau} + A\mathbf{x}^k + P(\mathbf{x}^{k+1/2}) \ni 0, \tag{23}$$

$$B(\mathbf{x}^{k+1} - \mathbf{x}^k) = \mathbf{x}^{k+1/2} - \mathbf{x}^k. \tag{24}$$

Первое соотношение метода, т.е. (23), есть простая итерация, аналогичная (19), в то время как второе соотношение (24) состоит в решении вспомогательной линейной системы и может рассматриваться как уточнение итерационного приближения $\mathbf{x}^{k+1/2}$. В расчетах, описываемых ниже, решение системы (24) было реализовано следующим образом. Сначала подпрограммами из пакета EISPACK [17], реализующими QL-алгоритм, решалась полная проблема собственных значений для матрицы A из (13). В результате получалось конструктивное разложение $A = QSQ^T$, где матрица Q состоит по столбцам из векторов собственного ортонормированного базиса, а элементы $s_{ii} \geq 0, i = 1, 2, \dots, n$, составляют диагональную матрицу S . Далее, учитывая, что $B = E + \tau A$, обратная матрица B^{-1} представлялась в виде $B^{-1} = Q\hat{S}Q^T$, где элементы диагональной матрицы \hat{S} имели вид $\hat{s}_{ii} = \frac{1}{1 + \tau s_{ii}}, i = 1, 2, \dots, n$. В результате решение системы (24) сводилось к двум умножениям на ортогональные матрицы и покомпонентному делению вспомогательного вектора на фиксированные положительные числа. В качестве реальной альтернативы такому способу можно привести предварительное разложение Холецкого матрицы B с последующим обращением двух треугольных матриц на итерации [18]. Однако по порядку вычислительные затраты обоих подходов одинаковы; кроме того, векторы, составляющие матрицу Q , использовались при построении начального приближения.

В методах (18), (21), (22) выбор начального приближения очень важен. Дело в том, что исходная матричная игра может иметь неединственное решение и, кроме того, тривиальный вектор $\mathbf{x} \equiv 0$ формально удовлетворяет включению (17). Поэтому при практическом использовании алгоритмов всегда следует контролировать нетривиальность получаемого приближенного решения и при необходимости повторять вычисления, стартуя с других начальных векторов. В описываемых ниже расчетах с целью избежать ортогональности вектора \mathbf{x}^0 искомому решению \mathbf{x} применялся следующий прием. При вычислении разложения матрицы A выделялись собственные векторы $\mathbf{q}_1, \mathbf{q}_2, \dots, \mathbf{q}_l, l \geq 1$, образующие базис ядра, т.е. соответствующие нулевым значениям $s_{ii}, i = 1, 2, \dots, l$, и из них строилась “средневзвешенная” линейная комбинация $\mathbf{x}^0 = \frac{1}{\sqrt{l}} \left(\sum_{i=1}^l \mathbf{q}_i \right)$. Следует отметить, что количество l базисных векторов ядра в общем случае не мало ($l \geq r$) и по порядку величины совпадает с размерностью r исходной симметричной матричной игры.

После того как зафиксирован выбор начального приближения, следует уточнить критерий останова итераций. Так как модельная постановка с играми Мендельсона имеет единственное (и известное аналитически!) решение, то удобно контролировать ошибку на итерациях не по “полному” вектору \mathbf{x}^k , а по его первой составляющей \mathbf{y}^k — приближению к решению игры \mathbf{y} , не используя для этого фиктивные переменные. С этой целью на каждой итерации вычислялась евклидова норма ошибки, и итерации останавливались, когда начальная ошибка уменьшалась в 1000 раз.

В табл. 2–4 приведены результаты численных экспериментов с описанными выше начальным приближением и критерием останова для алгоритмов (18), (21), (22).

Таблица 2

r	10	20	50	100
IT_{opt}	385	541	7575	37363
τ_{opt}	0.45	0.25	0.125	0.08

Рассмотрим сначала расчеты по методу Якоби из табл. 2. В первой строке указана размерность исходной игры Мендельсона, во второй строке приведено количество итераций до останова при оптимальном значении итерационного параметра τ , а в третьей, соответственно, — сами оптимальные значения τ . Легко проследить две тенденции по числам таблицы: сильную зависимость количества итераций от размерности и монотонное убывание оптимальных значений итерационного параметра. Отсюда следует вывод, что без веских дополнительных оснований метод Якоби применять не следует.

Таблица 3 содержит результаты вычислений по методу верхней релаксации (SOR) и имеет по сравнению с табл. 2 дополнительную строку. В ней приведены расчеты при $\tau = 1$, т.е. для частного случая, когда метод SOR представляет собой метод Гаусса–Зейделя. Отметим, что для взятого метода оптимизация по параметру не играет принципиальной роли: с ростом размерности игры оптимальное значение τ стремится к единице, а при фиксированной размерности — количества итераций по порядку совпадают.

Таблица 3

r	10	20	50	100
$IT_{\tau=1}$	117	319	838	1678
IT_{opt}	104	231	710	1471
τ_{opt}	1.2	1.3	1.1	1.1

Формальное сравнение с методом Якоби явно указывает, что при сопоставимых затратах на реализацию метод SOR сходится значительно быстрее. Поэтому практическая рекомендация однозначна: для решения небольших серий задач рекомендуется метод SOR, который даже без оптимизации (при $\tau = 1$) приводит к вполне удовлетворительным результатам.

В табл. 4 приведены расчеты по методу типа Дугласа–Рэкфорда, ее структура полностью идентична табл. 3. В данном случае вторая строка предназначена для иллюстрации сильной чувствительности метода к изменению параметра τ . Несложно заметить, что здесь оптимизация по параметру уменьшает количество итераций примерно в три раза. Сравнение метода расщепления с методом SOR показывает более высокую скорость сходимости первого, что вполне согласуется с более высокими вычислительными затратами на итерацию. Однако высокая чувствительность скорости сходимости метода типа Дугласа–Рэкфорда к изменению итерационного параметра позволяет рекомендовать в качестве основного для решения задачи IV метод Гаусса–Зейделя (метод SOR при $\tau = 1$) как достаточно простой в реализации и не сильно проигрывающий в асимптотике вычислительной работы.

Таблица 4

r	10	20	50	100	200
$IT_{\tau=1}$	118	255	871	1862	4831
IT_{opt}	44	87	296	537	1522
τ_{opt}	0.40	0.40	0.35	0.35	0.30

6. Выделение единственного решения с минимальной нормой. Так как множество оптимальных стратегий произвольной матричной игры (в том числе симметричной) непусто, выпукло, замкнуто и ограничено [10], то представляется удобным зафиксировать единственное решение задачи I как решение, имеющее минимальную евклидову норму. Факт единственности в данном случае следует из указанных свойств множества решений. Действительно, предположив, что имеются два решения \mathbf{y}_1 и \mathbf{y}_2 с одинаковой минимальной нормой, немедленно получаем противоречие, так как существующее, в силу выпуклости, решение $\tilde{\mathbf{y}} = \frac{1}{2}(\mathbf{y}_1 + \mathbf{y}_2)$ должно обладать еще меньшей нормой. Дальнейшее содержание раздела посвящено описанию способа получения этого единственного решения, которое удобно обозначить через \mathbf{y}_{\min} .

Рассмотрим простейшую симметричную матричную игру, имеющую неединственное решение. Например, ее несложно построить на основе уже применявшихся для расчетов игр Мендельсона. Пусть D_l — игра Мендельсона с матрицей размерности $l \times l$, $l \geq 3$, которая имеет единственное решение $\mathbf{y}_l = (1/4, 1/2, 1/4, 0, \dots, 0)^T \in \mathbb{R}^l$. Рассмотрим блочно диагональную матрицу D размерности $r \times r$, $r = 2l$, на диагонали которой расположены последовательно две матрицы D_l . Легко понять, что решением игры с матрицей выигрышей D будет однопараметрическое семейство $\mathbf{y} = \gamma(\mathbf{y}_l; 0) + (1 - \gamma)(0; \mathbf{y}_l)$, $\gamma \in [0, 1]$,

где нулем обозначен подвектор размерности l . В данном случае решению \mathbf{y}_{\min} соответствует значение $\gamma = 1/2$.

При попытке решить построенную матричную игру с помощью эквивалентной задачи NNLS соответствующий алгоритм упорно выдает решение, соответствующее $\gamma = 1$, в то время как решение для $\gamma = 0$ ничем его не хуже. Отмеченное равноправие связано с тем, что оба указанных решения являются крайними точками выпуклого замкнутого множества оптимальных стратегий игры с матрицей D [10]. С позиции формализма такой результат (отыскание только \mathbf{y} с $\gamma = 1$) вполне корректен, так как получаемые при этом векторы \mathbf{x} и \mathbf{w} удовлетворяют соотношениям (6)–(11). А объяснение этого явления достаточно простое: указанным выше решениям с различными γ соответствуют непересекающиеся наборы индексов, формирующие индексные множества \mathcal{L} и \mathcal{P} . Другими словами, если исходная игра с матрицей D имеет неединственное решение, то алгоритм NNLS с матрицей C , определенной по формуле (12), сформирует вектор \mathbf{x} , соответствующий какому-то из них.

С другой стороны, известно, что задача NNLS с матрицей C размерности $m \times n$, $m > n$, имеющей полный столбцовый ранг, имеет единственное решение [19]. Это приводит к идее регуляризации использовавшейся ранее в разделе 3 задачи NNLS.

Действительно, вместо задачи III с матрицей C , имеющей столбцовый ранг не выше r , рассмотрим задачу NNLS следующего вида: найти минимум функции $\|C_1\mathbf{x} - \mathbf{f}_1\|^2 + \alpha\|\mathbf{x}\|^2$ при условии $\mathbf{x} \geq 0$, где α — положительный параметр. Это означает, что матрицу C размерности $(r + 1) \times 2r$ (см. формулу (12)) следует заменить на матрицу C_1 размерности $(2r + 1) \times 2r$, имеющую вид

$$C_1 = \left(\begin{array}{c|ccc} D + \alpha E & & & E \\ \hline 0 & & & \alpha E \\ \hline 1 & 1 & \dots & 1 & | & 0 & 0 & \dots & 0 \end{array} \right),$$

и, соответственно, переопределить нормирующий вектор $\mathbf{f} = (0, 0, \dots, 0, 1)^T \in \mathbb{R}^{2r+1}$. Смысл вектора $\mathbf{x} = (\mathbf{y}; \mathbf{z})$ при этом остается прежним, однако значение функции $\rho_\alpha^2(\mathbf{x}) = \|C_1\mathbf{x} - \mathbf{f}_1\|^2 + \alpha\|\mathbf{x}\|^2$ уже будет ненулевым.

Обратим внимание, что для матрицы C_1 выполнены все условия единственности решения при произвольном положительном значении параметра α . Действительно, наличие полного столбцового ранга зависит только от невырожденности верхнего левого блока — матрицы $D + \alpha E$, что всегда имеет место, так как матрица D является по условию кососимметрической.

Такая постановка задачи NNLS позволяет выделять решение задачи II (вектор $\mathbf{x} = (\mathbf{y}; \mathbf{z})$), имеющее минимальную длину, с помощью варьирования (в данном случае — уменьшения) параметра α . Приведем результаты численных экспериментов, иллюстрирующих сказанное. Зафиксируем построенную на основе игр Мендельсона игру с однопараметрическим семейством решений при $r = 100$ (т.е. с $l = 50$). Несложно проверить, что решение \mathbf{x}_{\min} задачи II минимальной длины соответствует решению \mathbf{y} при $\gamma = 1/2$. В табл. 5 по строкам приведены значения параметра регуляризации α , евклидовой нормы уклонения приближенного расширенного решения \mathbf{x}_α от точного \mathbf{x}_{\min} , евклидовой нормы уклонения приближенного решения исходной матричной игры \mathbf{y}_α от точного \mathbf{y}_{\min} и целевой функции $\rho_\alpha(\mathbf{x}_\alpha)$. Результаты расчетов демонстрируют сходимость по параметру α решения регуляризованной задачи к искомому решению минимальной длины, причем асимптотически невязка $\rho_\alpha(\mathbf{x}_\alpha)$ убывает линейно, ошибка расширенного решения $\|\mathbf{x}_\alpha - \mathbf{x}_{\min}\|$ — квадратично, а ошибка для решения исходной игры $\|\mathbf{y}_\alpha - \mathbf{y}_{\min}\|$ — как $O(\alpha^4)$.

Таблица 5

α	0.1	0.05	0.01	0.001	0.0001
$\ \mathbf{x}_\alpha - \mathbf{x}_{\min}\ $	2.21	0.779	1.18×10^{-2}	1.18×10^{-4}	1.18×10^{-6}
$\ \mathbf{y}_\alpha - \mathbf{y}_{\min}\ $	3.27×10^{-2}	3.98×10^{-3}	1.04×10^{-6}	1.05×10^{-10}	1.05×10^{-14}
$\rho_\alpha(\mathbf{x}_\alpha)$	0.367	0.225	4.83×10^{-2}	4.84×10^{-3}	4.84×10^{-4}

Приведенная регуляризация не гарантирует выделения решения минимальной длины исходной задачи I, так как квадрат длины вектора фиктивных переменных в общем случае заранее не известен и определяется только после решения задачи. Однако в нашем распоряжении имеется весьма эффективный инструмент для управления этой длиной. Введем в матрицу C_1 положительный параметр σ , что соответствует нормировке (покомпонентному делению на σ) фиктивных неизвестных, в результате полу-

ЧИМ

$$C_\sigma = \left(\begin{array}{c|ccc} D + \alpha E & & & \sigma E \\ \hline 0 & & & \alpha E \\ \hline 1 & 1 & \dots & 1 & | & 0 & 0 & \dots & 0 \end{array} \right). \quad (25)$$

Несложно заметить, что увеличение параметра σ не меняет свойств матрицы, обеспечивающих единственность решения задачи NNLS, а лишь уменьшает относительную долю длины фиктивных неизвестных \mathbf{z} по отношению к общей длине вектора \mathbf{x} . Это означает, что двухпараметрическая постановка задачи NNLS с матрицей C_σ , определенной по формуле (25), позволяет конструктивно приблизиться к решению исходной задачи I минимальной длины: для этого нужно сначала взять достаточно большое значение σ , а затем провести серию расчетов, уменьшая значение параметра α . В табл. 6 приведены результаты численных экспериментов для значения $\sigma = 100$, остальные величины размерностей и параметров точно такие же, как при построении табл. 5. В данном случае иллюстрируется более быстрое приближение к точному решению за счет “понижения веса” фиктивных переменных; при этом все указанные выше асимптотические характеристики полностью сохраняются.

Таблица 6

α	0.1	0.05	0.01	0.001	0.0001
$\ \mathbf{x}_\alpha - \mathbf{x}_{\min}\ $	1.55×10^{-3}	3.91×10^{-4}	1.57×10^{-5}	1.57×10^{-7}	1.57×10^{-9}
$\ \mathbf{y}_\alpha - \mathbf{y}_{\min}\ $	2.41×10^{-6}	1.52×10^{-7}	2.44×10^{-10}	2.44×10^{-14}	2.44×10^{-18}
$\rho_\alpha(\mathbf{x}_\alpha)$	4.35×10^{-2}	2.18×10^{-2}	4.36×10^{-3}	4.36×10^{-4}	4.36×10^{-5}

Заключение. Основным содержанием работы является расширение понятия решения симметричной матричной игры с помощью введения фиктивных неизвестных и демонстрация целесообразности этого подхода с вычислительной точки зрения. При этом расширенное решение может быть использовано в различных целях: как для традиционного определения какой-либо оптимальной стратегии, так и для выделения единственного решения, имеющего минимальную длину. Последнее представляется достаточно важным, так как в случае неединственности решения исходной задачи всегда хочется иметь фиксированный в каком-то смысле (т.е. отличающийся от других) предел, к которому будут стремиться приближения (итерации) численного метода.

С этой целью в работе рассмотрены алгоритмы двух типов: решения задачи наименьших квадратов с неотрицательными ограничениями и решения вариационных неравенств специального вида. При этом приведено только обсуждение теории и численных экспериментов, так как обоснование сходимости и корректности алгоритмов изложено в других работах. Главный вывод таков: введение фиктивных неизвестных делает вполне доступным и прозрачным решение матричных игр умеренной размерности (порядка 10^2 – 10^3). Кроме того, изученные методы могут служить в качестве “строительных” блоков при конструировании различных версий многоуровневого метода для приближенного решения больших матричных игр [20], что делает содержание статьи еще более актуальным.

Наконец, определяя целевую функцию приближенного метода как поиск решения симметричной игры минимальной длины, представляется достаточно интересным сначала получить теоретический ответ на вопрос, к какому именно решению приводит тот или иной метод (например, прямо-двойственный алгоритм решения эквивалентной задачи линейного программирования [1]), а затем попытаться изменить его конструкцию так, чтобы он порождал решение минимальной длины. Скорее всего, последнее не всегда возможно. Однако смысл замечания таков: введение фиктивных переменных расширяет не только практические возможности, но и границы теоретических представлений в приближенном решении матричных игр.

СПИСОК ЛИТЕРАТУРЫ

1. *Nemirovski A.S., Todd M.J.* Interior-point methods for optimization // Acta Numerica. 2008. **17**. 191–234.
2. *Васильев Ф.П., Иванюцкий А.Ю.* Линейное программирование. М.: Факториал Пресс, 2008.
3. *Brown G.W.* Iterative solution of games by fictitious play // Activity Analysis of Production and Allocation / Ed. by T. C. Koopmans. New York: Wiley, 1951. 374–376.
4. *Robinson J.* An iterative method of solving a game // Ann. Math. 1951. **54**. 296–301.
5. *Gaas S.I., Zafra P.M., Qui Z.* Modified fictitious play for solving matrix games and linear programming problems // Comp. Oper. Res. 1995. **22**. 893–903.

6. *Washburn A.* A new kind of fictitious play // *Naval Research Logistics*. 2001. **48**. 269–280.
7. *Чижевков Е.В.* Итерационное решение матричных игр методами сеточных вариационных неравенств // *Журн. вычисл. матем. и матем. физики*. 2010. **50**, № 8. 1367–1380.
8. *Лоусон Ч., Хенсон Р.* Численное решение задач метода наименьших квадратов. М.: Наука, 1986.
9. *Нейман Дж., Моргенштерн О.* Теория игр и экономическое поведение. М.: Наука, 1970.
10. *Воробьев Н.Н.* Теория игр для экономистов-кибернетиков. М.: Наука, 1985.
11. *Оуэн Г.* Теория игр. М.: Изд-во ЛКИ, 2008.
12. *Крушевский А.В.* Теория игр. Киев: Вища школа, 1977.
13. www.netlib.org
14. *Mendelsohn N.S.* A psychological game // *American Mathematical Monthly*. 1946. **53**. 86–88.
15. *Лапин А.В.* Итерационные методы решения сеточных вариационных неравенств. Казань: Изд-во Казанского государственного университета, 2008.
16. *Гловински Р., Лионс Ж.-Л., Трёмольер Р.* Численное исследование вариационных неравенств. М.: Мир, 1979.
17. Зарубежные библиотеки и пакеты программ по вычислительной математике / Ред. У. Кауэлл. М.: Наука, 1993.
18. *Голуб Дж., Ван Лоун Ч.* Матричные вычисления. М.: Мир, 1999.
19. *Vjörck A.* Numerical methods for least squares problems. Philadelphia: SIAM, 1996.
20. *Чижевков Е.В.* Многоуровневый метод решения больших матричных игр // *Вычислительные методы и программирование*. 2009. **10**, № 2. 141–153.

Поступила в редакцию
29.08.2011
