Подсчёт нетранзитивных троек в методе парных сравнений

П. П. Рябов

§ 1. Введение

В социально-экономических приложениях существенная часть информации может быть получена только от экспертов и требует специальных методов анализа. Одним из самых старых методов экспертных оценок являются парные сравнения. В этом методе мы считаем, что эксперт из любых двух объектов может выбрать более предпочтительный или объявить, что эти объекты эквивалентны. Метод парных сравнений используются также при проведении соревнований (круговые турниры, далее в работе «турниры») и при анализе предпочтений индивида. В статье мы будем пользоваться турнирной терминологией.

Наличие полной информации о встречах позволяет делать выводы независимо от порядка игр участников, однако не решает проблему итогового упорядочивания. Например, для турнира без ничьих может найтись нетранзитивная тройка игроков a, b и c: a выиграл у b, b — у c, а c — у a. Тогда, как бы мы ни упорядочили игроков a, b и c в итоговой таблице, результат будет вызывать сомнения. Среди турниров без ничьих турнир, в котором отсутствуют нетранзитивные тройки, определён однозначно. Мы будем называть такой турнир **упорядоченным**.

Существуют разные метрики для сравнения турнира с упорядоченным. В данной статье для сравнения используется число нетранзитивных троек. Проблема такого сравнения заключается в сложности подсчёта этих троек при очень большом числе участников. Г. Дэвид [1, с. 20] вывел формулу для турнира без ничьих, позволяющую подсчитать число нетранзитивных троек через число побед каждого участника. Если в турнире допускаются ничьи, упорядоченный турнир можно определить по-разному (мы рассмотрим транзитивные и полутранзитивные турниры). А. Заславский [3, 4] обобщил формулу

Г. Дэвида для сравнения турнира с транзитивным турниром, но коэффициенты в формуле оказались плохо интерпретируемы. В этой статье мы покажем, почему нет другой линейной формулы для сравнения турнира с транзитивным и почему не существует линейной формулы для сравнения турнира с полутранзитивным.

§ 2. Определения

Сопоставим игрокам вершины графа. Если игрок a выиграл у b, то направим ребро из b в a. Если игроки сыграли вничью, то ребро будет направлено к обеим вершинам.

Обозначим через a_j число троек A_j в турнире. Пусть в турнире участвовало n игроков. Число побед, поражений и ничьих игрока i $(1 \leqslant i \leqslant n)$ обозначим через v_i, l_i и d_i соответственно.

На рис. 1 указаны всевозможные результаты встреч трёх игроков. Символом A_j, $1 \leqslant j \leqslant 7$, будем обозначать вид тройки.

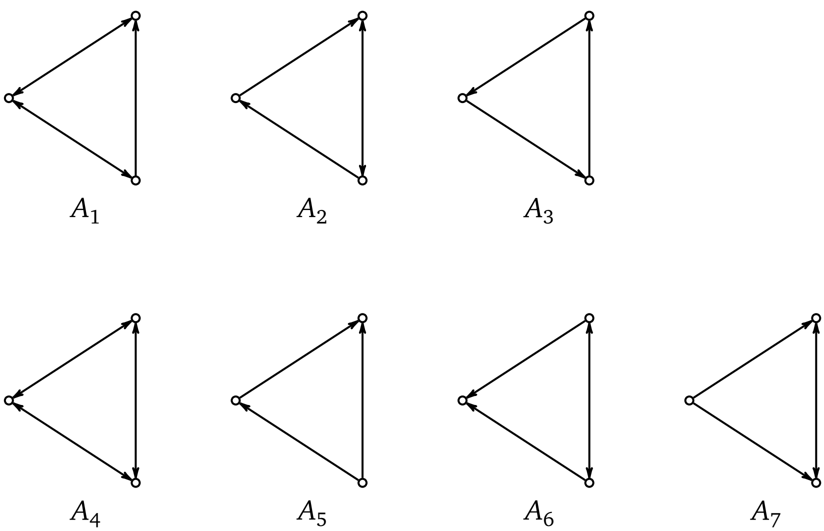

Рис. 1

Определение 1. **Полутранзитивным турниром** называется турнир, в котором не существует игроков $v_1, \ldots, v_k$ таких, что игрок v_1 выиграл или сыграл вничью с v_2, игрок v_2 выиграл или сыграл вничью с v_3, и т. д., игрок v_k выиграл или сыграл вничью с v_1 (при этом среди указанных игр не более одной ничьей).

Определение 2. **Транзитивным турниром** будем называть турнир, в котором отсутствуют тройки игроков A_1, A_2 и A_3.

Замечание. Упорядоченные турниры можно определять и с помощью циклов. **Транзитивным турниром** называется турнир, в котором не существует игроков $v_1, \ldots, v_k$ таких, что игрок v_1 выиграл или сыграл вничью с v_2, игрок v_2 выиграл или сыграл вничью с v_3, и т. д., игрок v_k выиграл или сыграл вничью с v_1. При этом хотя бы одна из данных игр отлична от ничейной. **Полутранзитивным турниром** называется турнир, в котором не существует игроков $v_1, \ldots, v_k$ таких, что игрок v_1 выиграл у v_2, игрок v_2 выиграл у v_3, и т. д., игрок v_k выиграл у v_1. Эквивалентность определений доказана в [3]. Однако для сравнения турнира с упорядоченными гораздо легче считать число троек A_1, A_2, A_3, чем число циклов неопределённой длины.

§ 3. Основной результат

Для турнира без ничьих Г. Дэвид вывел следующую формулу [1, с. 20]:

$$\frac{n(n-1)(2n-1)}{12} - \frac{1}{2}\sum_{i=1}^{n} v_i^2 = a_3.$$

А. Заславский обобщил формулу Г. Дэвида для турниров с ничьими [3, 4].

$$6a_3 + 3a_2 + a_1 = \frac{(n^3-n) - \sum_{i=1}^{n} d_i(d_i+2) - 3\sum_{i=1}^{n}(v_i - l_i)^2}{4}.$$

Покажем, что в любой формуле, связывающей линейную комбинацию числа троек A_1, A_2 и A_3 с числом побед/ничьих/поражений каждого игрока, коэффициенты при a_1, a_2 и a_3 определены однозначно с точностью до умножения на константу.

Теорема 1. *Для турнира с n участниками верны следующие равенства*:

$$\sum_{i=1}^{n}(v_i l_i) = a_2 + a_5 + 3a_3,$$

$$\sum_{i=1}^{n}(v_i d_i) = a_1 + a_2 + 2a_7,$$

$$\sum_{i=1}^{n}(d_i l_i) - \frac{n(n-1)(n-2)}{6} = a_6 - a_3 - a_4 - a_5 - a_7,$$

$$\sum_{i=1}^{n}(2v_i^2 + d_i) - (n-1)n = 4a_5 + 4a_6,$$

$$\sum_{i=1}^{n}(2l_i^2 + d_i) - (n-1)n = 4a_5 + 4a_7.$$

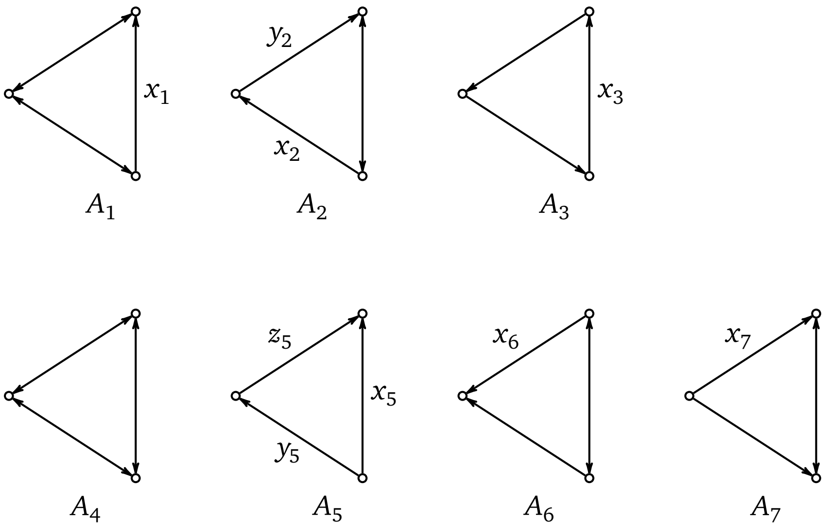

Рис. 2

Доказательство. Рассмотрим произвольный турнир, в котором не все партии закончились вничью. В этом турнире возьмём двух игроков a и b таких, что игрок a выиграл у b, и заменим результат их встречи на ничейный. Подсчитаем, как изменится число различных троек в турнире и число побед, поражений и ничьих каждого игрока. На рис. 2 числа x_i, y_j, $i = 1, 2, 3, 5, 6, 7$, $j = 2, 5$ и z_5 обозначают количество троек данного вида с ребром ab (в тройках A_2 и A_5 рёбра неравнозначны, поэтому важно учитывать, где расположено ребро ab; в тройку A_4 не могут одновременно входить игроки a и b). Через $\triangle$ обозначим изменение величины при замене результата игры на ничейный (из нового значения мы вычитаем первоначальное). Далее приведены формулы изменения числа троек каждого вида в турнире при замене результата игры между a и b на ничейный.

$$
\begin{aligned}
\triangle a_1 &= -x_1 + x_2 + y_2 + x_6 + x_7,\\
\triangle a_2 &= -x_2 - y_2 + x_3 + x_5,\\
\triangle a_3 &= -x_3,\\
\triangle a_4 &= x_1,\\
\triangle a_5 &= -x_5 - y_5 - z_5,\\
\triangle a_6 &= y_5 - x_6,\\
\triangle a_7 &= z_5 - x_7.
\end{aligned}
\tag{1}
$$

Через число троек с ребром ab также можно выразить число побед v_a, v_b, поражений l_a, l_b и ничьих d_a, d_b игроков a и b до замены результата их встречи:

$$
\begin{aligned}
v_a &= x_5 + z_5 + x_6 + 1, & v_b &= y_2 + x_3 + z_5,\\
l_a &= x_2 + x_3 + y_5, & l_b &= x_5 + y_5 + x_7 + 1,\\
d_a &= x_1 + y_2 + x_7, & d_b &= x_1 + x_2 + x_6.
\end{aligned} \tag{2}
$$

После замены результата встречи на ничью число побед у a уменьшилось на единицу, число ничьих увеличилось на единицу, число поражений не изменилось; у игрока b число поражений уменьшилось на единицу, число ничьих увеличилось на единицу, количество побед не изменилось. Теперь, используя формулы (2), запишем равенства:

$$
\begin{aligned}
&\triangle(v_a^2) = -1 - 2x_5 - 2z_5 - 2x_6, && \triangle(v_b^2) = 0,\\
&\triangle(l_a^2) = 0, && \triangle(l_b^2) = -1 - 2x_5 - 2y_5 - 2x_7,\\
&\triangle(d_a^2) = 1 + 2x_1 + 2y_2 + 2x_7, && \triangle(d_b^2) = 1 + 2x_1 + 2x_2 + 2x_6,\\
&\triangle(d_a l_a) = x_2 + x_3 + y_5, && \triangle(d_b l_b) = x_5 + y_5 + x_7 - x_1 - x_2 - x_6,\\
&\triangle(v_a d_a) = -x_1 - y_2 - x_7 + x_5 + z_5 + x_6, && \triangle(v_b d_b) = y_2 + x_3 + z_5,\\
&\triangle(v_a l_a) = -x_2 - x_3 - y_5, && \triangle(v_b l_b) = -y_2 - x_3 - z_5,\\
&\triangle(d_a) = 1, && \triangle(d_b) = 1.
\end{aligned}
$$

Поскольку результаты остальных участников турнира не поменялись, верны следующие формулы:

$$
\begin{aligned}
&\triangle\left(\sum_{i=1}^{n} v_i^2\right) = -1 - 2x_5 - 2z_5 - 2x_6,\\
&\triangle\left(\sum_{i=1}^{n} l_i^2\right) = -1 - 2x_5 - 2y_5 - 2x_7.\\
&\triangle\left(\sum_{i=1}^{n} d_i^2\right) = 2 + 4x_1 + 2x_2 + 2y_2 + 2x_6 + 2x_7,\\
&\triangle\left(\sum_{i=1}^{n} d_i l_i\right) = x_3 + x_5 - x_1 - x_6 + x_7 + 2y_5,\\
&\triangle\left(\sum_{i=1}^{n} v_i d_i\right) = -x_1 + x_3 - x_7 + x_5 + 2z_5 + x_6,\\
&\triangle\left(\sum_{i=1}^{n} v_i l_i\right) = -x_2 - y_2 - 2x_3 - y_5 - z_5,\\
&\triangle\left(\sum_{i=1}^{n} d_i\right) = 2.
\end{aligned} \tag{3}
$$

Приравнивая (1) и (3), получаем равенства:

$$\triangle\Big(\sum_{i=1}^{n} v_i l_i\Big) = -x_2 - y_2 - 2x_3 - y_5 - z_5 = \triangle a_2 + \triangle a_5 + 3\triangle a_3,$$

$$\triangle\Big(\sum_{i=1}^{n} (v_i d_i)\Big) = -x_1 + x_3 - x_7 + x_5 + 2z_5 + x_6 = \triangle a_1 + \triangle a_2 + 2\triangle a_7,$$

$$\triangle\Big(\sum_{i=1}^{n} (d_i l_i)\Big) = x_3 + x_5 - x_1 - x_6 + x_7 + 2y_5 = \triangle a_6 - \triangle a_3 - \triangle a_4 - \triangle a_5 - \triangle a_7,$$

$$\triangle\Big(\sum_{i=1}^{n} (2v_i^2 + d_i)\Big) = -4x_5 - 4z_5 - 4x_6 = 4\triangle a_5 + 4\triangle a_6,$$

$$\triangle\Big(\sum_{i=1}^{n} (2l_i^2 + d_i)\Big) = -4x_5 - 4y_5 - 4x_7 = 4\triangle a_5 + 4\triangle a_7.$$

Заменяя победы и поражения на ничьи, получаем турнир целиком из ничейных игр. В нём выполнено равенство

$$\sum_{i=1}^{n} d_i = n(n-1),$$

а число троек A_4 равно $n(n-1)(n-2)/6$. Отсюда нетрудно видеть, что правая и левая части формул из условия теоремы меняются одинаково при замене результатов игр на ничейные, а в турнире из ничьих имеют одинаковые значения. □

Замечание. Выражения величины Δ через v_i, l_i и d_i не единственные. Для любого i выполняется равенство $v_i + d_i + l_i = n - 1$.

Рассмотрим $a_1, \ldots, a_7$ как векторы. Тогда векторы

$$a_2 + a_5 + 3a_3, \quad a_1 + a_2 + 2a_7, \quad a_6 - a_3 - a_4 - a_5 - a_7, \quad 4a_5 + 4a_6 \quad \text{и} \quad 4a_5 + 4a_7$$

образуют пятимерное векторное подпространство V в семимерном векторном пространстве W, порождённом линейными комбинациями из $\{a_j\}$, $1 \leqslant j \leqslant 7$.

Теорема 2. *Пусть число участников турнира не меньше* 6. *Тогда все линейные комбинации элементов* a_j, $1 \leqslant j \leqslant 7$, *значения которых можно определить при любых возможных количествах побед/поражений/ничьих каждого игрока, лежат в* V.

Доказательство. Покажем, что при числе участников турнира $n \geqslant 6$ невозможно получить информацию о количествах троек вида A_2 и A_3

и любой их линейной комбинации. Для этого рассмотрим два турнира: в одном все игроки сыграли вничью, кроме четвёрки игроков, которые сыграли как показано на рис. 3 слева, а во втором все участники сыграли вничью, кроме четырёх игроков, которые сыграли, как показано на рис. 3 справа. Для обоих турниров множества, образованные тройками из чисел побед, поражений, ничьих каждого игрока, совпадают. Однако в первом турнире нет троек A_2 и 2 тройки A_3, а во втором 2 тройки A_2 и 1 тройка A_3. Следовательно, единственная линейная комбинация, о которой можно получить информацию, это $a_2 + 2a_3$.

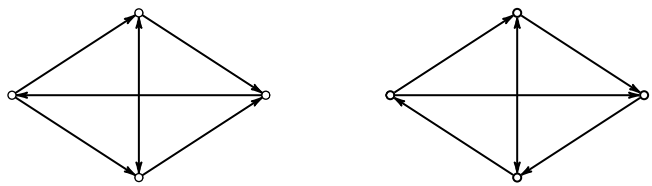

Рис. 3

Рассмотрим другую пару турниров, в каждом из которых все игроки сыграли вничью, кроме шестёрки игроков, игры которых, отличные от ничейных, показаны на рис. 4 (слева — в первом турнире, справа — во втором). В первом турнире нет троек A_3 и есть 6 троек A_2, во втором нет троек A_2 и есть 2 тройки A_3. Множества образованные тройками из чисел побед, ничьих, поражений каждого игрока, для обоих турниров совпадают. Поэтому о комбинации $a_2 + 2a_3$ также не всегда возможно получить информацию.

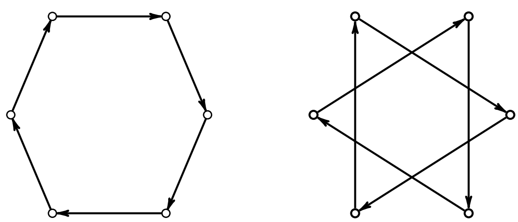

Рис. 4

Линейные комбинации, о которых можно получить информацию, образуют векторное подпространство. Размерность пространства $\{a_2, a_3\}$ равна двум; $\dim W = 7$. Следовательно, векторное пространство, о котором можно получить информацию, имеет размерность не больше 5. Однако $\dim V = 5$. Следовательно, ни для одной линейной комбинации, не лежащей в V, мы не сможем однозначно определить значения при любых возможных количествах побед/поражений/ничьих каждого игрока. □

Замечание. Если число участников турнира меньше шести, то по числу побед/ничьих/поражений каждого игрока можно всегда узнать, например, число троек A_4.

Следствие 1. *Не существует формулы, связывающей линейную комбинацию числа троек A_2 и A_3 с числом побед/ничьих/поражений каждого игрока.*

Следствие 2. *В любой формуле, связывающей линейную комбинацию числа троек A_1, A_2 и A_3 с числом побед/ничьих/поражений каждого игрока, коэффициенты при a_1, a_2 и a_3 относятся как* $1:3:6$.

Благодарности

Автор благодарен своему научному руководителю А. А. Заславскому за ценные наставления и замечания.

Список литературы

[1] *Дэвид Г.* Метод парных сравнений. М.: Статистика, 1978.

[2] *Заславский А. А.* Геометрия парных сравнений // Автоматика и телемеханика. 2007. № 3. С. 181–186.

[3] *Заславский А. А., Френкин Б. Р.* Математика турниров. М.: МЦНМО, 2009.

[4] *Заславский А. А., Шевлякова А. Н.* Геометрический метод анализа парных сравнений. Результаты вычислительного эксперимента // Вестник МЭИ. 2010. № 6. С. 5–12.

[5] *Заславский А. А.* О логичных и нелогичных турнирах // Квант. 1997. № 5. С. 11–13.

Павел Павлович Рябов, ВШЭ, Москва
pavel_ryabov_98@mail.ru