Методы решения матричных игр

Сведение матричной игры к системе линейных неравенств

Для того, чтобы описать первый метод решения матричных игр, нам понадобится следующая теорема:
Пусть рассматривается игра с матрицей выигрыша A. Чтобы $x^0 = \left(x_1^0, x_2^0, \ldots, x_m^0\right)$ *была оптимальной смешанной стратегией первого игрока, необходимо и достаточно, чтобы выполнялись следующие неравенства:*

$$\sum_{i=1}^{m} a_{ij} x_i^0 \geq v \ , \quad j = 1, 2, \ldots, n \ ,$$

где v *— цена игры.*

Аналогично для второго игрока смешанная стратегия $y^0 = \left(y_1^0, y_2^0, \ldots, y_n^0\right)$ *будет оптимальной тогда и только тогда, когда*

$$\sum_{j=1}^{n} a_{ij} y_j^0 \leq v \ , \quad i = 1, 2, \ldots, m \ .$$

Кроме того, в силу определения смешанной стратегии,

$$\sum_{i=1}^{m} x_i = \sum_{j=1}^{n} y_j = 1 \ .$$

Таким образом, чтобы найти решение игры в смешанных стратегиях, необходимо решить приведенные выше системы неравенств. Рассмотрим на примере, как это можно сделать.

Игра на пальцах. Два игрока одновременно показывают один или два пальца. Одновременно с этим они говорят "один" или "два", пытаясь угадать, сколько пальцев покажет противник. После этого выигрыш подсчитывается следующим образом:

— если оба угадали, то объявляется ничья и выигрыш равен нулю;

— если оба не угадали, то объявляется ничья и выигрыш равен нулю;

— если угадал только один, то он объявляется победителем и его выигрыш равен сумме показанных пальцев.

Таким образом, у каждого игрока есть четыре стратегии:

S_1 : показать один палец и сказать "один";

S_2 : показать один палец и сказать "два";

S_3 : показать два пальца и сказать "один";

S_4 : показать два пальца и сказать "два".

Матрица выигрыша первого игрока выглядит следующим образом:

$$A = \begin{pmatrix} & S_1 & S_2 & S_3 & S_4 \\ S_1 & 0 & 2 & -3 & 0 \\ S_2 & -2 & 0 & 0 & 3 \\ S_3 & 3 & 0 & 0 & -4 \\ S_4 & 0 & -3 & 4 & 0 \end{pmatrix}$$

Матрица является кососимметричной, а потому цена игры v равна нулю. Применим сформулированную выше теорему.

Оптимальная смешанная стратегия первого игрока должна удовлетворять системе условий

$$\begin{aligned} -2x_2 + 3x_3 \quad &\geq 0 \\ 2x_1 - 3x_4 \quad &\geq 0 \\ -3x_1 + 4x_4 \quad &\geq 0 \\ 3x_2 - 4x_3 \quad &\geq 0 \\ x_1 + x_2 + x_3 + x_4 \quad &= 1. \end{aligned}$$

Попытаемся найти решение системы, предположив сначала, что все неравенства можно заменить равенствами, т.е. решение находится на границе допустимого множества. Легко можно найти, что единственно возможное решение первых четырех уравнений $x_1 = x_2 = x_3 = x_4 = 0$ противоречит последнему, пятому. Значит, сделанное предположение неверно. Предположим теперь, что только первое неравенство выполняется строго и снова попробуем решить систему линейных уравнений. Теперь нам везет больше и существует единственное решение:

$$x_1 = 0, \quad x_2 = \frac{4}{7}, \quad x_3 = \frac{3}{7}, \quad x_4 = 0 \ .$$

Проверим, удовлетворяет ли это решение сделанному предположению:

$$-2x_2 + 3x_3 = -2 \cdot \frac{4}{7} + 3 \cdot \frac{3}{7} = \frac{1}{7} > 0 \ .$$

Таким образом, все условия удовлетворены и мы нашли смешанную стратегию первого игрока, которая в силу сформулированной теоремы является оптимальной. Поскольку рассматриваемая игра симметрична, то она является оптимальной и для второго игрока. Если бы и в этот раз решения не существовало, мы продолжили бы перебирать варианты дальше. Очевидно, что с ростом размерности матрицы выигрыша, т.е. увеличением числа стратегий у игроков, число вариантов очень быстро растет.

Сведение матричной игры к задаче линейного программирования

Рассмотрим теперь другой метод, основанный на сведении матричной игры к задаче линейного программирования, для которой известны сравнительно эффективные методы решения. Еще раз обратимся к критерию оптимальности:

$$\begin{cases} \displaystyle\sum_{i=1}^{m} a_{ij} x_i \geq v \ , \quad j = 1, 2, \ldots, n \ , \\ \displaystyle\sum_{i=1}^{m} x_i = 1 \ , \quad x_i \geq 0, \quad i = 1, 2, \ldots, m \ . \end{cases}$$

Введем дополнительные переменные $x_{m+j} \geq 0, \quad j = 1, 2, \ldots, n$ так, чтобы выполнялись следующие равенства:

$$\begin{cases} \displaystyle\sum_{i=1}^{m} a_{ij} x_i - x_{m+j} = v \ , \quad j = 1, 2, \ldots, n \ , \\ \displaystyle\sum_{i=1}^{m} x_i = 1 \ , \quad x_i \geq 0 \ , \quad i = 1, 2, \ldots, m \ . \end{cases}$$

Обратите внимание, что дополнительные переменные не являются частью смешанной стратегии, т.е. им не соответствует никакой стратегии игрока.

Чтобы построить задачу линейного программирования, следует добавить к записанным выше условиям целевую функцию. Для этого выделим первое уравнение и поставим следующую задачу:

$$\begin{cases} \sum_{i=1}^{m} a_{i1}x_i - x_{m+1} = v \to \max, \\ \sum_{i=1}^{m} (a_{ij} - a_{i1})x_i - x_{m+j} + x_{m+1} = 0, \quad j = 1, 2, \ldots, n \ , \\ \sum_{i=1}^{m} x_i = 1, \quad x_i \geq 0 \ , \quad i = 1, 2, \ldots, m \ . \end{cases}$$

Ограничения здесь получены вычитанием первого уравнения из всех остальных.

Очевидно, что решение этой задачи является одновременно и решением игры. Найти его можно, к примеру, симплекс-методом. Подобную же процедуру можно проделать и для второго игрока. Легко показать, что это будет двойственная к первой задача. А значит, используя двойственность симплекс-метода, можно, решая одну задачу, получить сразу оптимальные смешанные стратегии обоих игроков. Рассмотрим в качестве примера игру, решение которой указанным способом можно найти графически.

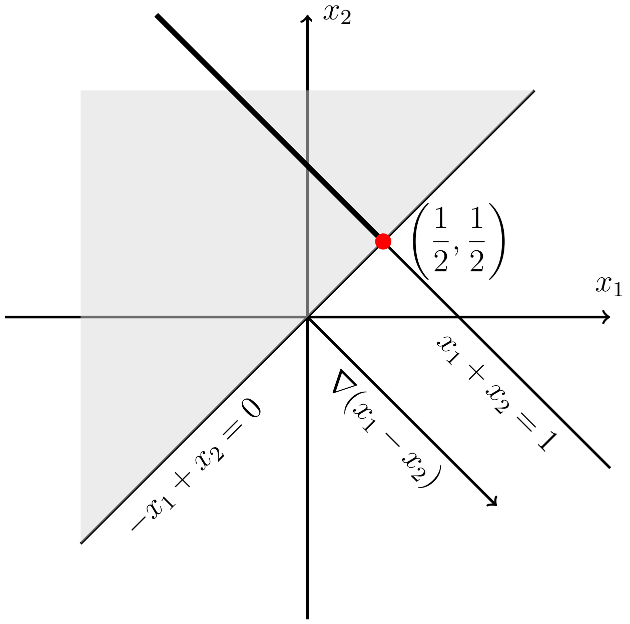
Игра в монеты. Два игрока, не глядя друг на друга, кладут на стол по одной монете вверх орлом или решкой по своему усмотрению. Если оба игрока выбрали одинаковые стороны, то первый игрок выигрывает и забирает обе монеты. В противном случае монеты забирает второй игрок. Матрица игры имеет вид:

$$A = \begin{pmatrix} & \text{Орел} & \text{Решка} \\ \text{Орел} & 1 & -1 \\ \text{Решка} & -1 & 1 \end{pmatrix}$$

Следуя изложенному методу, эту игру можно свести к задаче

$$\begin{cases} x_1 - x_2 \to \max \\ -x_1 + x_2 \geq 0 \\ x_1 + x_2 = 1 \ . \end{cases}$$

Нарисуем допустимую область :



Из рисунка видно, что решением задачи линейного программирования является точка $\left(\frac{1}{2}, \frac{1}{2}\right)$. Значит, решением игры будет смешанная стратегия, состоящая в том, чтобы в половине случаев выбирать орла, а в половине случаев решку.

Графоаналитический метод решения игр $2 \times n$ и $m \times 2$

Начнем с примера. Рассмотрим игру с матрицей выигрышей

$$A = \begin{pmatrix} 1 & 3 & 1 & 4 \\ 2 & 1 & 4 & 0 \end{pmatrix} .$$

Во-первых, проверим, что игра не имеет седловой точки. Далее, пусть игрок I, у которого всего две чистых стратегии, выбрал смешанную стратегию $x = (\xi, 1 - \xi)$, а его противник, игрок II, выбрал одну из своих чистых стратегий, например — первую. Тогда выигрыш первого игрока будет равен в среднем

$$V(x, 1) = 1 \cdot \xi + 2 \cdot (1 - \xi) = -\xi + 2 .$$

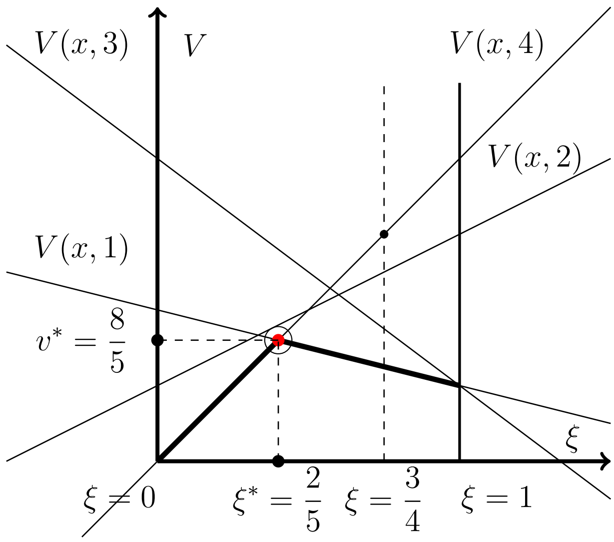
Аналогично. если игрок II выбирает свою вторую чистую стратегию, то выигрыш первого игрока равен

$$V(x, 2) = 3 \cdot \xi + 1 \cdot (1 - \xi) = 2\xi + 1 ,$$

и так далее,

$$V(x, 3) = 1 \cdot \xi + 4 \cdot (1 - \xi) = -3\xi + 4 ,$$
$$V(x, 4) = 4 \cdot \xi + 0 \cdot (1 - \xi) = 4\xi .$$

Изобразим эти выигрыши в пространстве (V, ξ).



Первый игрок желает получить гарантированный выигрыш. Значит, ему стоит рассчитывать на нижнюю оценку выигрыша, отмеченную на графике жирной линией. Поясним, почему.

Пусть, к примеру, первый игрок выбрал $\xi = \dfrac{3}{4}$. Тогда он может получить, вообще говоря, выигрыш $V(x,4) = 3$, как следует из графика. Однако цели игроков противоположны и игрок II, конечно, не будет выбирать 4-ую стратегию. Он выберет, напротив, 1-ую, худшую в данной ситуации для игрока I. Поэтому последний всегда должен исходить из самой пессимистической оценки, т.е. наименьшего гарантированного выигрыша.

Нижняя огибающая функций выигрыша первого игрока, которая отмечена жирной линией, является графиком функции $\min\limits_j V(x,j)$. Чтобы максимизировать свой выигрыш, игроку I надо найти $\max\limits_\xi \min\limits_j V(x,j)$, т.е. точку максимума нижней огибающей. Из графика видно, что точка является пересечением линий $V(x,1)$ и $V(x,4)$:

$$4\xi^* = -\xi^* + 2 \quad \Longrightarrow \quad \xi^* = \frac{2}{5}\ .$$

Выигрыш при этом равен $v^* = 4 \cdot \dfrac{2}{5} = \dfrac{8}{5}$. Это и будет цена игры, а оптимальная смешанная стратегия первого игрока имеет вид $x^* = \left(\dfrac{2}{5},\ \dfrac{3}{5}\right)$.

Найдем теперь оптимальную стратегию игрока II. Пусть она имеет вид

$$y^* = (\eta_1^*, \eta_2^*, \eta_3^*, \eta_4^*)\ , \quad \sum_i \eta_i^* = 1\ , \quad \eta_i^* \geq 0\ .$$

Тогда

$$v^* = \sum_{i=1}^{4} \eta_i^* \cdot V(x^*, i) = \eta_1^* \cdot \frac{8}{5} + \eta_2^* \cdot \frac{9}{5} + \eta_3^* \cdot \frac{14}{5} + \eta_4^* \cdot \frac{8}{5} = \frac{8}{5}\ .$$

Отсюда видно, что, так как $V(x^*,2) > \dfrac{8}{5}$ и $V(x^*,3) > \dfrac{8}{5}$, то $\eta_2^* = \eta_3^* = 0$. Другими словами, второму игроку вообще не имеет смысла использовать вторую и третью чистые стратегии и из матрицы выигрышей можно вычеркнуть второй и третий столбец. Тогда для нахождения η_1^* и η_4^* получаем систему

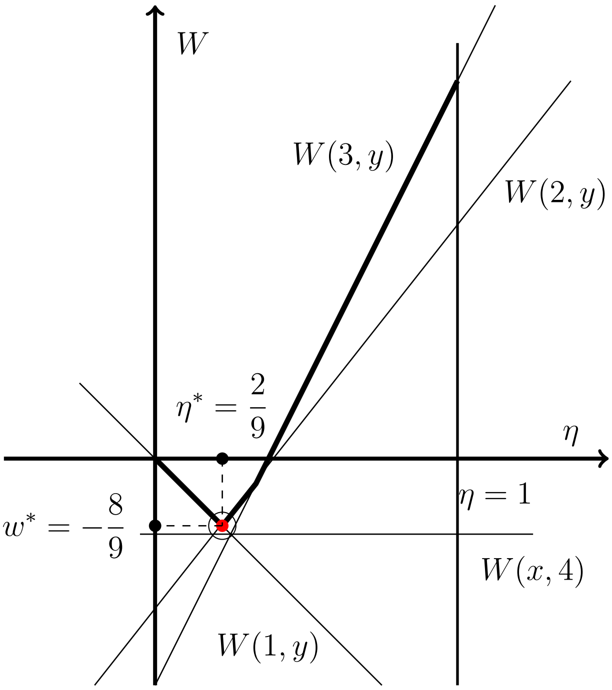
$$\begin{cases} \eta_1^* + 4\eta_4^* = \dfrac{8}{5} \\ \qquad\ \ 2\eta_1^* = \dfrac{8}{5} \end{cases},$$

решив которую, найдем $\eta_1^* = \dfrac{4}{5}$ и $\eta_4^* = \dfrac{1}{5}$. Значит, оптимальная смешанная стратегия игрока II имеет вид $y^* = \left(\dfrac{4}{5},\ 0,\ 0,\ \dfrac{1}{5}\right)$.

Упражнение. Решить матричную игру с матрицей выигрыша

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

Решение. Процедура решения аналогична изложенной выше. Кратко изложим его.



Судя по графику, оптимальная стратегия второго игрока имеет вид $y^* = \left(\dfrac{2}{9},\ \dfrac{7}{5}\right)$, а цена игры $w^* = -\dfrac{8}{9}$. Далее,

$$w^* = \sum_{i=1}^{4} \xi_i^* \cdot W(i, y^*) = -\xi_1^* \cdot \frac{8}{9} - \xi_2^* \cdot \frac{8}{9} - \xi_3^* \cdot \frac{2}{9} - \xi_4^* \cdot 1 = -\frac{8}{9} \ .$$

Снова приходим к выводу, что в смешанную стратегию только две чистые стратегии — первая и вторая — входят с ненулевой вероятностью. Для нахождения этих вероятностей решим систему

$$\begin{cases} -4\xi_1^* + 3\xi_2^* = -\dfrac{8}{9} \\ \qquad\quad -2\xi_2^* = -\dfrac{8}{9} \end{cases} ,$$

и найдем оптимальную стратегию первого игрока: $x^* = \left(\dfrac{5}{9},\ \dfrac{4}{9},\ 0,\ 0\right)$.

Метод Брауна-Робинсон (фиктивного разыгрывания)

Пусть игра задана матрицей A размерности $m \times n$. Каждое разыгрывание игры в чистых стратегиях будет далее называться партией. Метод Брауна-Робинсон — это итеративная процедура построения последовательности пар смешанных стратегий игроков, сходящейся к решению матричной игры.

В 1-ой партии оба игрока выбирают произвольную чистую стратегию. Пусть сыграно k партий, причем выбор стратегии в каждой партии запоминается. В $(k+1)$-ой партии каждый игрок выбирает ту чистую стратегию, которая максимизирует его ожидаемый выигрыш, если противник играет в соответствии с эмпирическим вероятностным распределением, сформировавшимся за k партий. Оценивается интервал для цены игры (об этом ниже) и, если он достаточно мал, процесс останавливается. Полученные при этом вероятностные распределения определяют смешанные стратегии игроков.

Пусть первые k партий первый игрок использовал i-ую стратегию ξ_i^k раз ($i = \overline{1, m}$), а второй игрок свою j-ую стратегию использовал η_j^k раз ($j = \overline{1, n}$). Тогда первый игрок полагает, что его противник будет следовать смешанной стратегии

$$y^k = \left(\frac{\eta_1^k}{k}, \ldots, \frac{\eta_n^k}{k}\right),$$

а второй игрок, в свою очередь, полагает, что первый последует согласно смешанной стратегии

$$x^k = \left(\frac{\xi_1^k}{k}, \ldots, \frac{\xi_m^k}{k}\right).$$

Значит, при следующем разыгрывании на $(k+1)$-ой партии они сыграют так:

$$\text{Первый игрок выбирает} \quad i_{k+1} : \sum_j a_{i_{k+1},j}\eta_j^k = \max_i \sum_j a_{ij}\eta_j^k = \bar{v}^k.$$

$$\text{Второй игрок выбирает} \quad j_{k+1} : \sum_i a_{i,j_{k+1}}\xi_i^k = \min_j \sum_i a_{ij}\xi_i^k = \underline{v}^k.$$

Через $\bar{v}^k$ и $\underline{v}^k$ обозначены k-ые приближения верхней и нижней цены игры соответственно. Легко видеть, что

$$\frac{\bar{v}^k}{k} = \max_i \sum_j a_{ij}\frac{\eta_i^k}{k} = \sum_j a_{i_{k+1},j}\frac{\eta_i^k}{k},$$

$$\frac{\underline{v}^k}{k} = \min_j \sum_i a_{ij}\frac{\xi_i^k}{k} = \sum_i a_{i,j_{k+1}}\frac{\xi_i^k}{k}.$$

Поэтому по определению цены игры получаем интервальную оценку для истинной цены игры v:

$$\max_k \frac{\underline{v}^k}{k} \leqslant v \leqslant \min_k \frac{\bar{v}^k}{k}.$$

Нетрудно показать, что итеративый процесс сходится к истинному значению игры, т.е.

$$\lim_{k\to\infty} \min_k \frac{\bar{v}^k}{k} = \lim_{k\to\infty} \max_k \frac{\underline{v}^k}{k} = v.$$

Заметим, что итерационный процесс метода Брауна-Робинсон не является, вообще говоря, монотонным. Кроме того, скорость сходимости метода быстро уменьшается с ростом размерности матрицы игры. Однако он обладает одним неоспоримым преимуществом, которое заключается в исключительной простоте программирования метода.