Московский государственный университет имени М.В. Ломоносова
Факультет вычислительной математики и кибернетики
Кафедра исследования операций

Байрамкулов Аслан Магомедович

Исследование некоторых моделей покера двух лиц

Магистерская диссертация

Научный руководитель:
к.ф-м.н., старший научный сотрудник
Белянкина Т.В.

Москва, 2019

Содержание

1 Введение

В представленной работе рассматриваются некоторые простые модели покера двух лиц. Формализация моделей изложена в [2] и в [4]. Частично эти модели рассматриваются в [5]. При этом в некоторых расчетах был использован математический аппарат, описанный в [6].

Все рассмотренные модели покера проистекают одна из другой. Каждая следующая по сути является усложнением предыдущей. Самый начальный и примитивный вариант игры - игра двух лиц с одним раундом и фиксированной ставкой. Далее мы можем предоставить опцию первому игроку выбрать ставку из фиксированного множества. Далее увеличить количество раундов. В перспективе можно рассматривать модели покер нескольких лиц.

В разделе 2 формализуется модель покер, вводятся обозначения, определяется вид оптимальных стратегий каждого из игроков. Оптимальные стратегии формируются на основе средней функции выигрыша игрока. Представлено доказательство оптимальности стратегий. Также в конце этого раздела исследуется вопрос единственности оптимальных стратегий. Показано, что существует несколько оптимальных стратегий у первого игрока. Такие дополнительные стратегии соответствуют блеф-стратегии игрока.

В разделе 3 модель чуть усложняется. А именно первый игрок получает возможность выбрать ставку из множества(ранее ставка была фиксированной). Вначале рассматривается простейший вариант с двумя ставками, но затем исследуется общий случай с n ставками. Также, как и в предыдущем разделе, происходит подробный вывод оптимальных стратегий. В конце раздела исследуются асимптотические свойства полученных стратегий.

В разделе 4 исследуется модель покера с несколькими раундами ставок(до этого был лишь 1 раунд). Предполагается, что колода карт состоит из старших и младших карт. Старшая карта достается первому игроку с вероятностью P. Оба игрока знают значение P, но лишь первый игрок знает, какая карта ему досталась. В модели допускается, что на старшей карте первый игрок всегда ставит, а на младшей он либо блефует, либо пасует. Также, как и для всех предыдущих моделей, здесь приведен вывод оптимальных стратегий.

Хочется отметить, что строгий математический вывод оптимальных решений в таких моделях положительно сказывается на покерную индустрию. В настоящее время покер развивается. Создаются все новые и новые программы, предлагающие расчет оптимальных стратегий в тех или иных ситуациях. Причем это актуально не только для безлимитного техасского холдема, так и для других(чуть менее популярных видов покера), например: пот-лимитная омаха, стад, разз и другие. Наукоемкие исследования позволяют создать более качественный покерный софт и помочь игрокам улучшить свои покерные навыки.

2 Покер. Теоретико-игровая модель

В качестве математической модели этой игры рассмотрим игру двух лиц. В начале игры игроки делают взнос, равный единице. После этого они получают две карты достоинством x и y, не имея информации о карте противника. Первый игрок имеет выбор, либо спасовать, и тогда он теряет свой взнос, либо сделать ставку c, где $c > 1$. В последнем случае ход переходит ко второму игроку, и он имеет те же возможности. Если он пасует, то теряет свой взнос, в противном случае игроки открывают карты и тогда выигрывает тот, чья карта больше. Заметим, что значения карт игроков случайны, и в этом случае мы должны определить вероятностный характер всевозможных исходов.

Предположим, что значения карт лежат в интервале от 0 до 1 и все эти значения равновероятны, т. е. случайные величины x и y имеют равномерное распределение на интервале [0, 1]. Давайте определим стратегии в этой игре. Каждый из игроков знает лишь свою карту, и значит, его решение должно основываться лишь на этом. Поэтому под стратегией первого игрока в этой игре мы будем понимать функцию $\alpha(x)$ — вероятность сделать ставку, если на руках у него карта x. Поскольку α — вероятность, то ее значения $0 \leq \alpha \leq 1$ и $\overline{\alpha} = 1$ - α — это вероятность спасовать. Аналогично, если первый игрок сделал ставку, то стратегией второго игрока будет функция $\beta(y)$ - вероятность поддержать ставку, если у него на руках карта y. Понятно, что $0 \leq \beta(y) \leq 1$. В процессе игры будут выпадать различные комбинации карт, и следовательно, выигрыш каждого из игроков будет случайным. В качестве критерия будем использовать среднее значение выигрыша. В данной игре если стратегии игроков α, β выбраны, то в силу условий игры средний выигрыш первого игрока определяется соотношениями:

- 1, с вероятностью $\overline{\alpha}(x)$
- 1, с вероятностью $\alpha(x)\overline{\beta}(y)$
- $(c+1)sgn(x-y)$, с вероятностью $\alpha(x)\beta(y)$

где функция $sgn(x-y)$ определяется следующим образом:

$$sgn(x-y) = \begin{cases} 1, & x > y \\ -1, & x < y \\ 0, & x = y \end{cases}$$

В силу этих соотношений средний выигрыш первого игрока будет равен

$$H(\alpha,\beta) = \int_0^1 \int_0^1 \Big[-\overline{\alpha}(x) + \alpha(x)\overline{\beta}(y) + (c+1)sgn(x-y)\alpha(x)\beta(y) \Big] dxdy \tag{1}$$

Теперь игра полностью определена. Мы определили стратегии в игре и выигрыш. Первый игрок желает максимизировать, а второй игрок — минимизировать средний выигрыш (1).

2.1 Оптимальные стратегии

Нетрудно понять, какой вид должны иметь оптимальные стратегии. Для этого представим выигрыш (1) в следующем виде, выделив члены, содержащие $\alpha(x)$:

$$H(\alpha,\beta) = \int_0^1 \alpha(x) \Big[1 + \int_0^1 \Big(\overline{\beta}(y) + (c+1)sgn(x-y)\beta(y) \Big) dy \Big] dx - 1 \tag{2}$$

Обозначим выражение в квадратных скобках в (2) за $Q(x)$. Тогда из (2) следует, что оптимальная стратегия первого игрока, максимизирующая его выигрыш, должна иметь следующий вид.

$$\alpha^*(x) = \begin{cases} 1, & Q(x) > 0 \\ 0, & Q(x) < 0 \\ \forall, & Q(x) = 0 \end{cases}$$

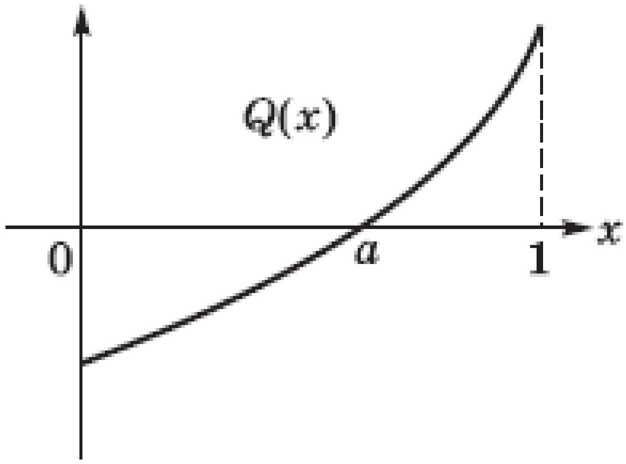

Рис. 1: Функция Q(x)

Поскольку функция $sgn(x-y)$, а вместе с ней и функция $Q(x)$ неубывающая, то, как видно из рис. 1 (допустим, что функция $Q(x)$ пересекает ось Ox в точке $0 \le a \le 1$), оптимальная стратегия $\alpha^*(x)$ должна определяться некоторым порогом a, и если полученная карта x имеет значение, меньшее, чем a, то следует пасовать, в противном случае — делать ставку.

Аналогично, представляя выигрыш $H(\alpha, \beta)$ в виде

$$H(\alpha,\beta) = \int_0^1 \beta(y)\Big[\alpha(x)(-(c+1)sgn(y-x)-1)dx\Big]dy+ \\ +\int_0^1 (2\alpha(x)-1)dx \quad (3)$$

мы получаем, что оптимальная стратегия второго игрока $\beta^*(y)$ также определяется некоторым пороговым значением b, больше которого он уравнивает ставку, и пасует в противном случае. Найдем эти оптимальные пороги a^*, b^*.

Предположим, что первый игрок использует стратегию α с порогом a. Тогда проигрыш второго игрока, согласно (3), имеет вид

$$H(\alpha, \beta) = \int_0^1 \beta(y)G(y)dy + 2(1-a) - 1 \qquad (4)$$

$$G(y) = \int_a^1 \Big[-(c+1)sgn(y-x) - 1 \Big] dx$$

Проведя вычисления, получаем

$$G(y) = \begin{cases} c(1-a), & y < a \\ -2(c+1)y + a(c+2) + c, & y \geq a \end{cases}$$

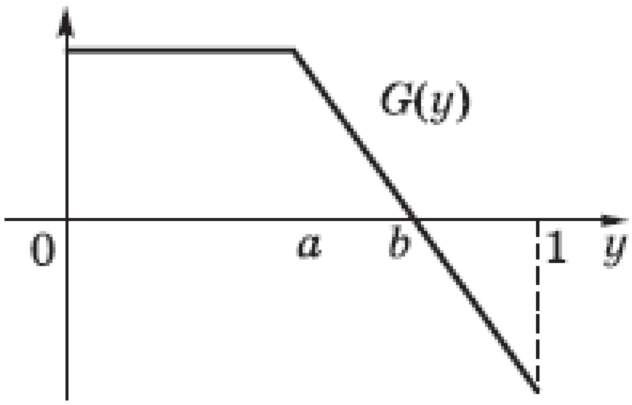

Рис. 2: Функция G(y)

На рис. 2 показан вид функции $G(y)$, из которого видно, что оптимальный порог b определяется соотношением $-2(c+1)b + a(c+2) + c = 0$, откуда находим

$$b = \frac{a(c+2)+c}{2(c+1)}. \qquad (5)$$

Таким образом, оптимальный порог второго игрока однозначно определяется значением порога первого игрока. При этом минимальное значение проигрыша второго игрока равно

$$H(\alpha,\beta) = \int_b^1 \beta(y)G(y)dy + 2(1-a) - 1 =$$
$$\int_b^1 \Big[-2(c+1)y + a(c+2) + c \Big] dy + 2(1-a) - 1. \quad (6)$$

После интегрирования получим

$$H(\alpha,\beta) = -(c+1)(1-b^2) + [a(c+2)+c](1-b) - 2a + 1 =$$
$$(c+1)b^2 - b[a(c+2)+c] + ac. \quad (7)$$

Подставляя оптимальное значение порога b из (5) в (7), получим представление минимального проигрыша второго игрока как функцию аргумента a:

$$H(a) = \frac{(a(c+2)+c)^2}{4(c+1)} - \frac{(a(c+2)+c)^2}{2(c+1)} + ac =$$
$$\frac{(c+2)^2}{4(c+1)}\Big[-a^2 + 2a\frac{c^2}{(c+2)^2} - \frac{c^2}{(c+2)^2}\Big]. \quad (8)$$

Поскольку a является стратегией первого игрока, естественно, что он будет стараться максимизировать минимальный проигрыш второго игрока (8). Таким образом, окончательно мы приходим к задаче, в которой нужно найти максимум параболы:

$$h(a) = -a^2 + 2a\frac{c^2}{(c+2)^2} - \frac{c^2}{(c+2)^2}.$$

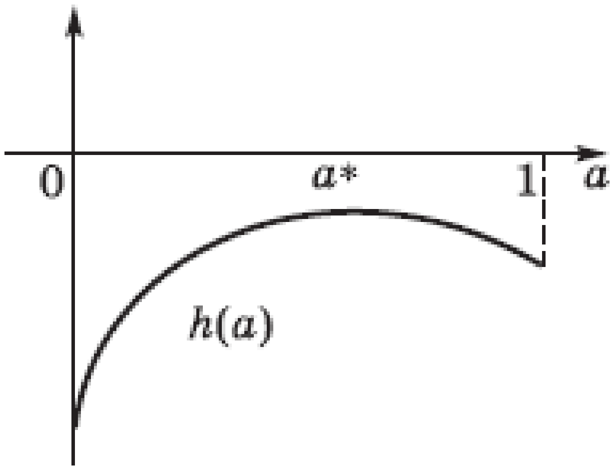

Рис. 3: Функция h(a)

Вид этой параболы изображен на рис. 3. Ее максимум достигается в точке

$$a^* = (\frac{c}{c+2})^2$$

$$0 \le a^* \le 1.$$

Подставляя это значение в (5), найдем оптимальный порог второго игрока:

$$b^* = \frac{c}{c+2}.$$

При этом значение выигрыша первого игрока (оно будет наилучшим и для первого, и для второго игрока) найдем, подставив оптимальный порог a^* в (8):

$$H^* = H(a^*, b^*) = \frac{(c+2)^2}{4(c+1)}\Big[(\frac{c}{c+2})^4 - (\frac{c}{c+2})^2\Big] = -(\frac{c}{c+2})^2.$$

Мы видим, что значение игры отрицательно, то есть первый игрок находится в невыгодном положении.

2.2 Особенности оптимального поведения в покере

Итак, оптимальные стратегии игроков и значение игры найдены. Мы получили, что оптимальный порог первого игрока меньше, чем у второго, т.е. он должен быть более осторожен. Значение игры отрицательно, это связано с тем, что ход первого игрока дает некоторую информацию о его карте второму игроку.

Теперь обсудим вопрос о единственности оптимальных стратегий. Как мы видели из рис. 2, в случае, если первый игрок использует оптимальную стратегию $\alpha^*(x)$ с порогом $a^* = (\frac{c}{c+2})^2$, то наилучшим ответом второго игрока будет также использование пороговой стратегии $\beta^*(y)$ с порогом $b^* = \frac{c}{c+2}$, т.е. оптимальная стратегия второго игрока определяется однозначно.

Теперь зафиксируем стратегию второго игрока с порогом b^* и найдем наилучший ответ первого игрока. Для этого опять обратимся к выражению 2 и вычислим функцию $Q(x)$.

Если $x < b^*$, то:

$$Q(x) = 1 + \int_0^1 \Big(\overline{\beta}(y) + (c+1)sgn(x-y)\beta(y)\Big)dy =$$
$$= 1 + \int_0^{b^*} dy - \int_{b^*}^1 (c+1)dy = 1 + b^* - (c+1)(1-b^*) = 0.$$

Если $x \geq b^*$, то:

$$Q(x) = 1 + b^* + \int_{b^*}^x (c+1)dy - \int_x^1 (c+1)dy =$$
$$= 1 + b^* + (c+1)(x-b^*) - (c+1)(1-x) = 2(c+1)x - c(b^*+1).$$

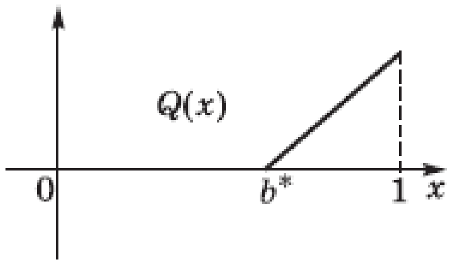

Рис. 4: Функция Q(x)

Как видно из рис. 4, функция $Q(x)$ на интервале $(b^*, 1]$ положительна, и, следовательно, если у первого игрока карта $x > b^*$, его наилучшим решением является сделать ставку. Однако если значение x лежит в интервале $[0, b^*]$, то $Q(x) = 0$ и $\alpha^*(x)$ может принимать любые значения, так как это не влияет на значение выигрыша (2). Конечно, найденная нами стратегия с порогом a^* удовлетворяет этому условию. Однако интересно, нет ли какой-либо другой стратегии первого игрока $\alpha(x)$, для которой оптимальная стратегия второго игрока совпадет с $\beta^*(y)$.

Оказывается, такие стратегии существуют. Рассмотрим, например, стратегию $\alpha(x)$ такого вида. Если $x \geq b^*$, то по-прежнему, первый игрок делает ставку, если же $x < b^*$, то с вероятностью $p = \frac{2}{c+2}$ он делает ставку и с вероятностью $\overline{p} = 1 - p = \frac{c}{c+2}$ пасует. Найдем наилучший ответ второго игрока для стратегии такого вида. Как и раньше, представим выигрыш в виде (4). Функция $G(y)$ теперь имеет вид:

$$G(y) = \int_0^{b^*} p(-(c+1)sgn(y-x)-1)dx + \int_{b^*}^1 (-(c+1)sgn(y-x)-1)dx.$$

Если $y < b^*$, то:

$$G(y) = p\int_0^y (-(c+1)-1)dx + p\int_y^{b^*} (c+1-1)dx + \int_{b^*}^1 (c+1-1)dx =$$
$$= -2p(c+1)y + pcb^* + c(1-b^*).$$

Если $y \geq b^*$, то:

$$G(y) = p\int_0^{b^*} (-c-2)dx + \int_{b^*}^{y} (-c-2)dx + \int_y^1 cdx =$$
$$= -2(c+1)y + (c+2)b^*(1-p) + c.$$

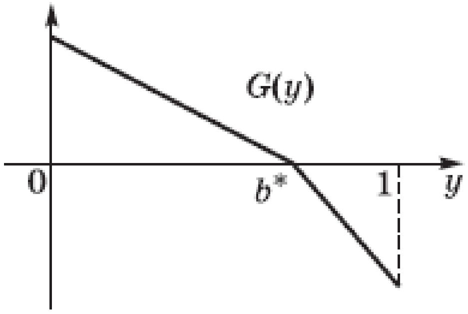

Рис. 5: Функция G(y)

Вид функции $G(y)$ изображен на рис.5. Заметим, что в силу выбора p $G(b^*) = 0$ и, следовательно, наилучшей стратегией второго игрока по-прежнему остается стратегия $\beta^*(y)$.

Таким образом, мы нашли еще одно решение данной игры. Это решение принципиально отличается от предыдущего для первого игрока. Теперь, даже имея малое значение карты на руках, первый игрок может сделать ставку. Этот эффект в карточных играх называется блефом. Игрок изображает, что у него на руках большая карта, понуждая противника сказать спасовать. Отметим, однако, что вероятность блефа тем меньше, чем больше значение ставки c. Например, для $c = 100$ вероятность блефа должна быть меньше 0.02.

3 Модель покера с переменной ставкой

В рассмотренной выше модели покера ставка c была фиксирована, однако, в реальной игре она может меняться. Рассмотрим модификацию данной модели с переменной ставкой. Итак, так же как и раньше, первый и второй игроки делают в начале игры взнос, равный единице. После этого они получают две карты достоинством x и y, не имея информации о карте противника. На первом шаге первый игрок делает ставку $c(x)$, которая зависит от значения его карты x. Ход переходит ко второму игроку, и он может либо спасовать, и потерять свой взнос в банке, либо поддержать вызов противника. Тогда он добавляет в банк $c(x)$, игроки открывают карты, и выигрывает тот из них, чья карта больше. Итак, в данной модели первый игрок выигрывает либо единицу, либо $(1+c(x))sgn(x-y)$. Задачей здесь является найти оптимальную функцию $c(x)$ и оптимальный ответ второго игрока. Эту проблему сформулировал в конце 1950-х годов Беллман.

Рассмотрим вначале дискретную модель данной игры. Предположим, что ставка первого игрока может принимать одно значение из следующего фиксированного набора $0 < c_1 < ... < c_n$. Тогда стратегией первого игрока является смешанная стратегия $\alpha(x) = (\alpha_1(x), ..., \alpha_n(x))$, где $\alpha_i(x)$ - вероятность поставить в игре c_i, $i = 1, ..., n$, если значение его карты равно x. Тогда $\sum_{i=1}^n \alpha_i = 1$. Стратегией второго игрока будет стратегия поведения $\beta(y) = (\beta_1(y), ..., \beta_n(y))$, где $\beta_i(y)$ - вероятность уравнять ставку c_i, $0 \le \beta_i \le 1$ при выбранной карте y. Соответственно $\overline{\beta_i}(y) = 1 - \beta_i(y)$ - вероятность паса при ставке c_i и карте y.

Ожидаемый выигрыш первого игрока равен:

$$H(\alpha, \beta) = \int_0^1 \int_0^1 \sum_{i=1}^n \Big[\alpha_i(x)\overline{\beta_i}(y) + (1+c_i)sgn(x-y)\alpha_i(x)\beta_i(y)\Big] dxdy. \tag{9}$$

Вначале разберем случай $n = 2$.

3.1 Модель покера с двумя ставками

Итак, предположим, что первый игрок в зависимости от выбранной карты x может поставить одну из двух ставок c_1 или c_2, $c_1 < c_2$. Таким образом, его стратегию можно определить с помощью функции $\alpha(x)$. $\alpha(x)$ - это вероятность ставки c_1. Соответственно $\overline{\alpha}(x) = 1 - \alpha(x)$ - вероятность ставки c_2. У второго игрока стратегия определяется с помощью двух функций $\beta_1(y), \beta_2(y)$, которые обозначают вероятности поддержать соответственно ставки c_1 и c_2. Функция выигрыша (9) примет вид

$$H(\alpha, \beta) = \int_0^1 \int_0^1 \Big[\alpha(x)\overline{\beta_1}(y) + (1 + c_1)sgn(x - y)\alpha(x)\beta_1(y) + \\ (1 - \alpha(x))\overline{\beta_2}(y) + (1 + c_2)sgn(x - y)(1 - \alpha(x))\beta_2(y)\Big]dxdy. \quad (10)$$

Найдем вначале вид оптимальной стратегии второго игрока. Выделим в (10) слагаемые с β_1 и β_2. Они имеют вид

$$\int_0^1 \beta_1(y)dy\Big[\int_0^1 \alpha(x)(-1 + (1 + c_1)sgn(x - y))dx\Big], \quad (11)$$

$$\int_0^1 \beta_2(y)dy\Big[\int_0^1 (1 - \alpha(x))(-1 + (1 + c_2)sgn(x - y))dx\Big]. \quad (12)$$

Функция $sgn(x-y)$ не возрастает по y, следовательно, и выражения в квадратных скобках в (11)-(12), обозначим их $G_i(y)$, $i = 1, 2$, также представляют собой невозрастающие функции по y (см. рис. 6). Предположим, что функции $G_i(y)$ пересекают ось Oy на интервале $[0, 1]$ в точках b_i, $i = 1, 2$.

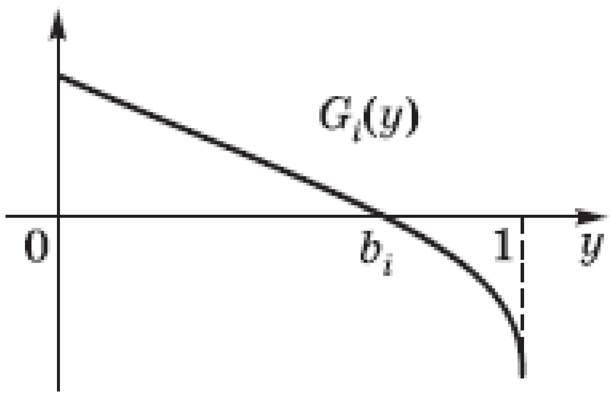

Рис. 6: Функция $G_i(y)$

Второй игрок заинтересован минимизировать выражения (11)-(12). Чтобы интегралы в этих выражениях принимали минимальные значения, необходимо, чтобы $\beta_i(y)$ было равно нулю для $G_i(y) > 0$ и равно 1 для $G_i(y) < 0$, $i = 1, 2$. Отсюда его оптимальная стратегия имеет вид $\beta_i(y) = I(y \geq b_i)$, $i = 1, 2$, где $I(A)$ - индикатор множества A. То есть второй игрок поддерживает ставку первого при достаточно больших картах (превышающих порог b_i, $i = 1, 2$). Поскольку $c_1 < c_2$, то естественно предположить, что порог для поддержания большей ставки также должен быть больше, т. е. $b_1 < b_2$. Пороги b_1, b_2 определяются уравнениями $G_i(b_i) = 0$, $i = 1, 2$ или согласно (11)-(12),

$$\int_0^{b_1} (-2 - c_1)\alpha(x)dx + \int_{b_1}^1 c_1\alpha(x)dx = 0,$$

$$\int_0^{b_2} (-2 - c_2)\overline{\alpha}(x)dx + \int_{b_2}^1 c_1\overline{\alpha}(x)dx = 0.$$

Теперь перейдем к построению оптимальной стратегии первого игрока $\alpha(x)$. Выделим в выигрыше (10) выражение с $\alpha(x)$:

$$\int_0^1 \alpha(x)dx\Big[\int_0^1(\beta_2(y)-\beta_1(y)+sgn(x-y)((1+c_1)\beta_1(y)-(1+c_2)\beta_2(y)))dy\Big].$$

Выражение в квадратных скобках обозначим $Q(x)$. Тогда для тех x, где $Q(x) < 0$, оптимальная стратегия $\alpha(x)$ будет равна нулю, для x, где $Q(x) > 0$, $\alpha(x) = 1$. Там же, где $Q(x) = 0$, $\alpha(x)$ может принимать произвольные значения. После преобразований

$$Q(x) = \int_0^x (c_1\beta_1(y) - c_2\beta_2(y))dy + \int_x^1 ((2 + c_2)\beta_2(y) - (2 + c_1)\beta_1(y))dy.$$

Производная функции $Q(x)$ равна:

$$Q^{'}(x) = (2 + 2c_1)\beta_1(x) - (2 + 2c_2)\beta_2(x).$$

С учетом вида стратегий $\beta_i(x)$, $i = 1, 2$, производная принимает вид:

$$Q^{'}(x) = \begin{cases} 0, & x \in [0, b_1] \\ 2 + 2c_1, & x \in (b_1, b_2) \\ -2(c_2 - c_1) & x \in [b_2, 1] \end{cases}.$$

Таким образом функция $Q(x)$ на отрезке $[0, b_1]$ постоянна, на интервале (b_1, b_2) возрастает и на отрезке $[b_2, 1]$ убывает.

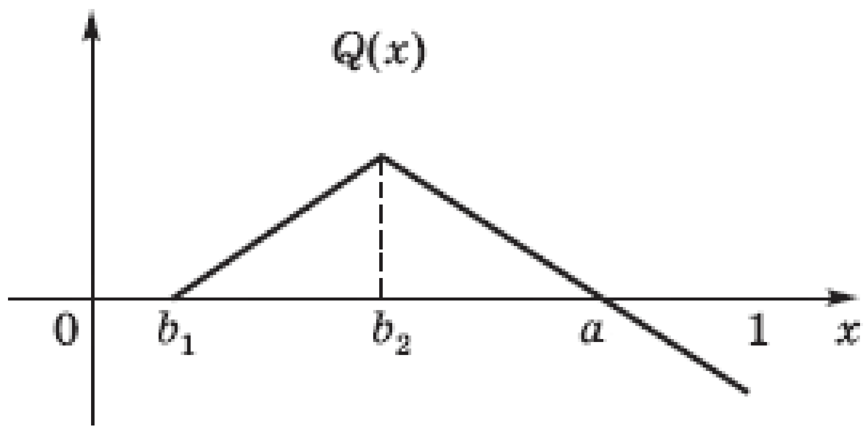

Рис. 7: Функция Q(x)

Потребуем, чтобы на интервале $[0, b_1]$ функция $Q(x)$ была равна нулю и пересекала ось Ox в некоторой точке a на интервале $[b_2, 1]$ (см. рис. 7). Тогда будут иметь место соотношения $b_1 < b_2 < a$. Для этого необходимо выполнение условий

$$Q(0) = \int_{b_2}^{1} (2 + c_2) dy - \int_{b_1}^{1} (2 + c_1) dy = 0$$

и

$$Q(a) = \int_{b_1}^{a} c_1 dy - \int_{b_2}^{a} c_2 dy + \int_{a}^{1} (c_2 - c_1) dy = 0.$$

Упрощая эти условия, получим

$$(1 - b_1)(2 + c_1) = (1 - b_2)(2 + c_2),$$

$$(c_2 - c_1)(2a - 1) = c_2 b_2 - c_1 b_1.$$

При выполнении этих условий оптимальная стратегия первого игрока будет иметь вид

$$\alpha(x)=\begin{cases} 1, & x\in(b_1,a) \\ 0, & x\in[a,1] \\ \forall, & x\in[0,b_1] \end{cases}.$$

Таким образом условия (11)-(12) можно переписать в виде

$$\int_0^{b_1}\alpha(x)dx=\frac{c_1(a-b_1)}{2+c_1}=b_1-\frac{c_2(1-a)}{2+c_2}. \tag{13}$$

Таким образом, параметры для определения оптимальных стратегий игроков определяются системой уравнений

$$(1-b_1)(2+c_1)=(1-b_2)(2+c_2) \tag{14}$$

$$(c_2-c_1)(2a-1)=c_2b_2-c_1b_1 \tag{15}$$

$$\frac{c_1(a-b_1)}{2+c_1}=b_1-\frac{c_2(1-a)}{2+c_2}. \tag{16}$$

Можно показать, что система уравнений (14)-(16) имеет решение $0\le b_1<b_2\le a\le 1$. Ниже мы покажем это для общего случая.

Особенностью оптимальной стратегии первого игрока является то, что на интервале $[0,b_1]$ его стратегия может быть произвольной, лишь бы выполнялось условие (13). Это соответствует стратегии блефа в игре, поскольку для маленьких карт первый игрок может объявить высокую ставку в игре. У второго игрока оптимальная стратегия предписывает при маленьких картах выходить из игры, а при достаточно больших - поддерживать ту или иную ставку противника. Например, при $c_1=2$, $c_2=4$ оптимальные параметры такие: $b_1=0.345$, $b_2=0.563$, $a=0.891$. При картах со значениями меньше 0.345 первый игрок блефует, если же карты больше этого значения, но меньше 0.891 он ставит ставку 2, а при картах больших 0.891 он ставит ставку 4. Второй игрок поддерживает соответственно ставку 2, если его карты в интервале $[0.345,0.563]$, и 4, если они больше 0.563. В других случаях он пасует.

3.2 Модель покера с n ставками

Теперь предположим, что первый игрок после получения карты x может объявить ставку из следующего набора значений $0 < c_1 < \ldots < c_n$. Тогда его стратегией будет смешанная стратегия $\alpha(x) = (\alpha_1(x), ..., \alpha_n(x))$, где $\alpha_i(x)$ представляет вероятность объявить ставку c_i. После этого ход переходит ко второму игроку. В зависимости от выбранной карты y, он может спасовать и потерять свой взнос в банке, либо продолжить игру. В последнем случае он должен уравнять ставку первого игрока. Игроки вскрывают карты, и выигрывает тот, чье значение больше. Стратегией второго игрока является стратегия поведения $\beta(y) = (\beta_1(y), ..., \beta_n(y))$, где $\beta_i(y)$ - вероятность уравнять ставку первого игрока, которая равна c_i, $i = 1, ..., n$. Функция выигрыша имеет вид

$$H(\alpha,\beta) = \int_0^1\int_0^1\sum_{i=1}^n\Big[\alpha_i(x)\overline{\beta_i}(y) + (1+c_i)sgn(x-y)\alpha_i(x)\beta_i(y)\Big]dxdy. \quad (17)$$

Найдем вначале оптимальную стратегию второго игрока. Для этого перепишем функцию (17) в виде

$$H(\alpha,\beta) = \sum_{i=1}^n\int_0^1\beta_i(y)dy\Big[\int_0^1\alpha_i(x)(-1+(1+c_i)sgn(x-y))dx\Big]+1. \quad (18)$$

Обозначим выражение в квадратных скобках как $G_i(y)$. Для каждой фиксированной стратегии $\alpha(x)$ и ставки c_i второй игрок заинтересован минимизировать (18). Поэтому его оптимальная стратегия для любого $i = 1, ..., n$ должна иметь вид

$$\beta_i(y) = \begin{cases} 1, & G_i(y) < 0 \\ 0, & G_i(y) > 0 \end{cases},$$

$$G_i(y) = -(2+c_i)\int_0^y \alpha_i(x)dx + c_i\int_y^1 \alpha_i(x)dx.$$

Легко видеть, что функция $G_i(y)$ не возрастает по y. Кроме того:

$$G_i(0) = c_i\int_0^1 \alpha_i(x)dx \geq 0,$$

$$G_i(1) = -(2+c_i)\int_0^1 \alpha_i(x)dx \leq 0.$$

Следовательно, всегда существует корень b_i уравнения $G_i(y) = 0$ (см. рис.6).

b_i удовлетворяет уравнению:

$$\int_0^{b_i} \alpha_i(x)dx = \frac{c_i}{2+c_i}\int_{b_i}^1 \alpha_i(x)dx. \tag{19}$$

Таким образом, оптимальная стратегия второго игрока принимает вид

$$\beta_i(y) = \begin{cases} 0, & 0 \leq y < b_i \\ 1, & b_i \leq y \leq 1 \end{cases},$$

$i = 1, ..., n$.

Заметим, что набор b_i, $i = 1, ..., n$, удовлетворяющий уравнению (19) существует для любой стратегии $\alpha(x)$.

Теперь перейдем к построению оптимальной стратегии первого игрока. Представим выигрыш (17) в виде

$$H(\alpha, \beta) = \sum_{i=1}^{n}\int_0^1 \alpha_i(x)Q_i(x)dx, \tag{20}$$

$$Q_i(x) = \int_0^1 (\overline{\beta_i}(y) + (1+c_i)sgn(x-y)\beta_i(y))dy =$$
$$= b_i + (1+c_i)\Big(\int_0^x \beta_i(y)dy - \int_x^1 \beta_i(y)dy\Big). \quad (21)$$

Для каждого x первый игрок ищет стратегию $\alpha(x)$, которая максимизирует выигрыш (20). Заметим, что это другая оптимизационная задача, отличная от той, которую мы исследовали в рассуждениях для второго игрока.

Здесь $\alpha(x)$ - смешанная стратегия, $\sum_1^n \alpha_i(x) = 1$. Максимальное значение выигрыша (20) достигается при таком $\alpha(x)$, что $\alpha_i(x) = 1$, если для данного x $Q_i(x)$ принимает значения большие, чем другие $Q_j(x)$, $j \neq i$, и $\alpha_i(x) = 0$ в противном случае. Если же для данного значения x все значения $Q_i(x)$ совпадают, то $\alpha(x)$ может принимать произвольные значения.

Будем искать оптимальную стратегию $\alpha(x)$ в специальном виде. Пусть на интервале $[0, b_1)$ все функции $Q_i(x)$ совпадают, т.е. $Q_1(x) = ... = Q_n(x)$. Это будет соответствовать стратегии блефа для первого игрока. Положим $a_1 = b_1$. Далее пусть на интервале $[a_1, a_2)$ функция $Q_1(x) > max\{Q_j(x), j \neq 1\}$, на интервале $[a_2, a_3)$ функция $Q_2(x) > max\{Q_j(x), j \neq 2\}$ и т.д. На интервале $[a_n, 1]$ предположим, что максимальное значение принимает функция $Q_n(x)$. Тогда оптимальная стратегия первого игрока примет вид

$$\alpha_i(x) = \begin{cases} \forall, & x \in [0, b_1] \\ 1, & x \in [a_i, a_{i+1}) \\ 0, & x \notin [a_i, a_{i+1}), x \notin [0, b_1] \end{cases}.$$

Уточним вид функции $Q_i(x)$. Упростив (21), получим:

$$Q_i(x) = \begin{cases} b_i - (1+c_i)(1-b_i), & 0 \le x < b_i \\ (1+c_i)(2x-1) - c_i b_i, & b_i \le x \le 1 \end{cases}.$$

На интервале $[0, b_i]$ функция $Q_i(x)$ принимает постоянные значения. Потребуем, чтобы эти значения для всех функций $Q_i(x)$, $i =$

$1, ..., n$ были одинаковы, т.е.

$$b_i - (1 + c_i)(1 - b_i) = k, i = 1, ..., n$$

Тогда все b_i, $i = 1, ..., n$ можно определить по формуле

$$b_i = \frac{1 + k + c_i}{2 + c_i} = 1 - \frac{1 - k}{2 + c_i}, i = 1, ..., n. \tag{22}$$

Из (22) сразу следует, что $b_1 < ... < b_n$. Это соответствует интуитивным соображениям, что второй игрок должен поддерживать более высокую ставку при больших картах.

На интервале $[b_i, 1]$ функция $Q_i(x)$ является линейной. Обозначим точки пересечения функций $Q_{i-1}(x)$ и $Q_i(x)$ через a_i, $i = 2, ..., n$. При этом $a_1 = b_1$. Для того, чтобы оптимальная стратегия $\alpha(x)$ имела необходимый вид, потребуем, чтобы $a_1 < a_2 < ... < a_n$. Тогда на интервале $[a_i, a_{i+1})$ функция $Q_i(x)$ будет максимальной, $i = 1, ..., n$. График расположения функций $Q_i(x)$, $i = 1, ..., n$ представлен на рис. 8.

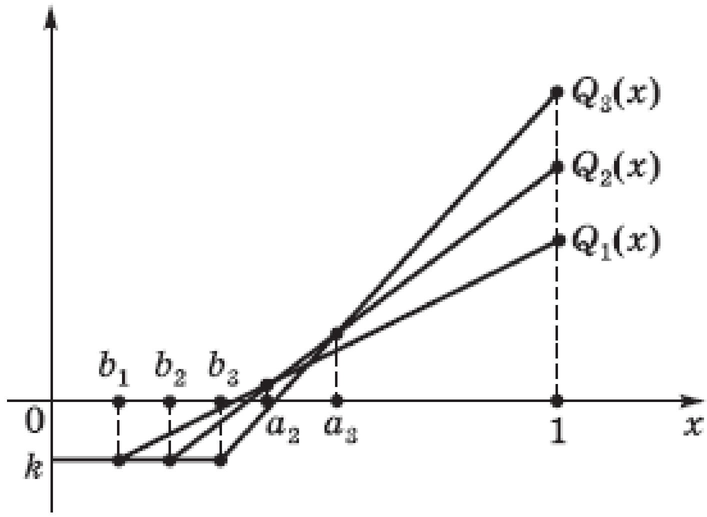

Рис. 8: Функции $Q_i(x)$

Точки пересечения a_i находятся из уравнений

$$(1+c_{i-1})(2a_i-1)-c_ib_i=(1+c_i)(2a_i-1)-c_ib_i, i=2,...,n.$$

Преобразуем и получим:

$$a_i=1-\frac{\overline{k}}{(2+c_{i-1})(2+c_i)}, i=1,...,n,$$

где $\overline{k}=1-k$.

Остается определить k. Напомним, что оптимальные пороги стратегии второго игрока b_i удовлетворяли уравнению (19), которое с учетом вида стратегии α примет вид

$$\int_0^{b_1}\alpha_i(x)dx=\frac{c_i}{2+c_i}(a_{i+1}-a_i), i=1,...,n. \tag{23}$$

Складывая все уравнения (23) и учитывая условие $\sum_{i=1}^n \alpha_i(x)=1$, получим

$$b_1=\sum_{i=1}^n\frac{c_i}{2+c_i}(a_{i+1}-a_i).$$

Откуда

$$\frac{1+k+c_1}{2+c_1}=\overline{k}A,$$

где

$$A=\sum_{i=1}^n\frac{c_i(c_{i+1}-c_{i-1})}{(2+c_{i-1})(2+c_i)^2(2+c_{i+1})}.$$

Здесь в сумме мы считаем $c_0=-1$, $c_{n+1}=\infty$. Отсюда

$$k = 1 - \frac{2 + c_1}{A(2 + c_1) + 1}.$$

Очевидно, что A и вместе с ним $\overline{k}$ положительны. Следовательно, последовательность a_i является монотонной, $a_1 < a_2 < ... < a_n$. Кроме того, все пороги a_i лежат в промежутке $[0, 1]$.

Подытожим проделанные рассуждения. Оптимальная стратегия первого игрока имеет вид

$$\alpha_i^*(x) = \begin{cases} \forall f : (23), & x \in [0, b_1] \\ 1, & x \in [a_i, a_{i+1}) \\ 0, & x \notin [0, b_1], x \notin [a_i, a_{i+1}) \end{cases},$$

где

$$a_i = 1 - \frac{\overline{k}}{(2 + c_{i-1})(2 + c_i)},$$

$$i = 2, ..., n.$$

Заметим, что на интервале $[0, b_1)$ первый игрок блефует. При малых значениях карт он может поставить любую ставку. Для определенности можно разбить отрезок $[0, b_1)$, последовательно начиная от нуля, на интервалы длиной $c_i(a_{i+1} - a_i)/(2 + c_i)$, $i = 1, ..., n$ (их сумма по построению равна b_1) и положить $\alpha_i^*(x) = 1$ на соответствующем интервале. Для $x > b_1$ первый игрок должен ставить ставку c_i на интервале $[a_i, a_{i+1})$.

Оптимальная стратегия второго игрока определяется как

$$\beta_i^*(y) = \begin{cases} 0, & 0 \le y < b_i \\ 1, & b_i \le y \le 1 \end{cases},$$

где $b_i = (1 + k + c_i)/(2 + c_i)$, $i = 1, ..., n$.

Значение игры найдем из (20):

$$H(\alpha^*, \beta^*) = \sum_{i=1}^{n} \int_0^1 \alpha_i^*(x) Q_i(x) dx =$$

$$\int_0^{b_1} k \sum_{i=1}^{n} \alpha_i^*(x) dx + \sum_{i=1}^{n} \int_{a_i}^{a_{i+1}} Q_i(x) dx =$$

$$kb_1 + \sum_{i=1}^{n} (a_{i+1} - a_i) \Big[(1 + c_i)(a_i + a_{i+1}) - (1 + c_i + c_i b_i) \Big]. \quad (24)$$

Например, для игры, в которой размеры ставок равны $c_1 = 1$, $c_2 = 3$, $c_3 = 6$, последовательно находим параметры

$$A = \frac{c_1(c_2+1)}{(2+c_1)^2(2+c_2)} + \frac{c_2(c_3-c_2)}{(2+c_1)(2+c_2)^2(2+c_3)} + \frac{c_3}{(2+c_2)(2+c_3)^2} = 0.122.$$

$$k = 1 - \frac{2+c_1}{A(2+c_1)+1} = -1.193.$$

И затем оптимальные стратегии

$$b_1 = \frac{1+k+c_1}{2+c_1} = 0.269,$$

$$b_2 = \frac{1+k+c_2}{2+c_2} = 0.561,$$

$$b_3 = \frac{1+k+c_3}{2+c_3} = 0.725,$$

$$a_1 = b_1 = 0.269,$$

$$a_2 = 1 - \frac{1-k}{(2+c_1)(2+c_2)} = 0.854,$$

$$a_3 = 1 - \frac{1-k}{(2+c_2)(2+c_3)} = 0.945.$$

Наконец, значение игры равно

$$H(\alpha^*, \beta^*) = -0.117.$$

Значение игры отрицательно, поэтому игра невыгодна для первого игрока.

3.3 Асимптотические свойства стратегий в модели покера с переменной ставкой

Вернемся к проблеме, сформулированной в начале главы. Предположим, что получив карту x, первый игрок может объявить ставку $c(x)$, которая может принимать произвольное значение из R. Воспользуемся результатами, полученными в предыдущем разделе. Выберем некоторое положительное число B и нанесем на отрезке $[0, B]$ равномерную сеть $\{\frac{B}{n}, \frac{2B}{n}, ..., B\}$, где n - некоторое целое число. Будем считать узлы этой сетки значениями ставок в игре, т. е. $c_i = \frac{iB}{n}$. При этом мы будем неограниченно увеличивать n и i, но таким образом, чтобы сохранялось равенство $\frac{iB}{n} = c$ для некоторого c. После этого мы будем неограниченно увеличивать B. Найдем предельные значения параметров, определяющих оптимальные стратегии игроков в игре с такими ставками. Найдем вначале предел A:

$$A = \sum_{i=1}^{n} \frac{c_i(c_{i+1} - c_{i-1})}{(2 + c_{i-1})(2 + c_i)^2(2 + c_{i+1})} = \sum_{i=1}^{n} \frac{B\frac{i}{n}2\frac{B}{n}}{(2 + B\frac{i-1}{n})(2 + B\frac{i}{n})^2(B\frac{i+1}{n})}.$$

При $n \to \infty$ данная интегральная сумма сходится к интегралу

$$A \to \int_0^B \frac{2c}{(2+c)^4}dc = \frac{1}{12} - \frac{3B}{3(2+B)^3} - \frac{1}{3(2+B)^2},$$

который при $B \to \infty$ имеет предел $A = \frac{1}{12}$. Отсюда получаем предельное значение для k, $k = 1 - \frac{2}{2A+1} = -\frac{5}{7}$.

Теперь мы можем определить пороговое значение для блефа первого игрока. $b_1 = a_1 = 1 - \frac{1-k}{2+B/n} - > 1 - \frac{1+\frac{5}{7}}{2} = \frac{1}{7}$. Итак, если у первого игрока значение карт меньше $\frac{1}{7}$, он блефует. Теперь перейдем к определению размера ставки, которую первый игрок должен ставить в зависимости от значения его карты x. Согласно определению найденной выше оптимальной стратегии первого игрока $\alpha^*(x)$, он ставит ставку c_i в интервале $\big[a_i, a_{i+1}\big)$, где

$$a_i = 1 - \frac{1-k}{(2+c_{i-1})(2+c_i)}.$$

Таким образом, ставка $c = c(x)$, которая соответствует карте x, удовлетворяет уравнению

$$x = 1 - \frac{1-k}{(2+c)^2},$$

откуда находим

$$c(x) = \sqrt{\frac{12}{7(1-x)}} - 2. \tag{25}$$

Выражение (25) неотрицательно при $x \geq \frac{4}{7}$, следовательно, при $\frac{1}{7} \leq x < \frac{4}{7}$ первый игрок не ставит ничего. При значении карт $x \geq \frac{4}{7}$ размер его ставки определяется формулой (25).

Определим асимптотическое поведение второго игрока. При $y < \frac{1}{7}$ он пасует. Если же $y \geq \frac{1}{7}$, второй игрок согласно (22) поддерживает заявленную ставку первого игрока c при значении карт, больших

$$y \geq 1 - \frac{1-k}{2+c} = 1 - \frac{12}{7(2+c)}.$$

Поведение оптимальных стратегий обоих игроков представлено на рис. 9.

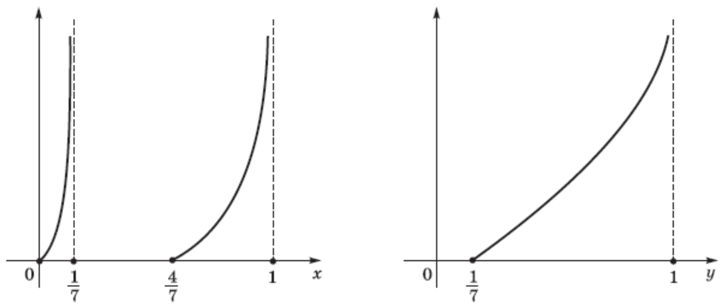

Рис. 9: Оптимальные стратегии в покере

Остается вычислить предельное значение игры. Воспользуемся выражением (24). Переходя к пределу, получим

$$kb_1+\int_0^\infty \frac{2(1-k)}{(2+c)^3}(1+c)\Big[2(1-\frac{1-k}{(2+c)^2})-(1+\frac{c(1+k+c)}{(1+c)(2+c)})\Big]dc=$$
$$-\frac{5}{49}+\frac{24}{7}\int_0^\infty \frac{1+c}{(2+c)^3}\Big(1-\frac{24}{7(2+c)^2}-\frac{c(\frac{2}{7}+c)}{(1+c)(2+c)}\Big)dc=-\frac{5}{49}+\frac{18}{49}=\frac{13}{49}. \tag{26}$$

Заметим, однако, что к предельному значению (26) нужно добавить выигрыш, который получается при значении карт первого игрока в интервале $\left[\frac{1}{7},\frac{4}{7}\right]$, т.е.

$$\int_{\frac{1}{7}}^{\frac{4}{7}}\Big(\frac{1}{7}+\int_{\frac{1}{7}}^{1} sgn(x-y)dy\Big)dx=\frac{3}{49}\int_{\frac{1}{7}}^{\frac{4}{7}}(2x-\frac{8}{7})dx=-\frac{6}{49}. \tag{27}$$

Складывая (26) и (27), получим

$$H(\alpha^*,\beta^*)=\frac{13}{49}-\frac{6}{49}=\frac{1}{7}.$$

Итак, предельное значение игры равно $\frac{1}{7}$, она выгодна первому игроку. Оптимальные стратегии полностью определены. Интересно, что выигрыш и стратегии определяются через число 7, которое неожиданно появилось в результате аналитических вычислений.

4 Модель покера с несколькими раундами ставок

Рассмотрим модель, в рамках которой игра длится более одного раунда. Пусть n - количество раундов. Перед началом игры оба игрока делают обязательный взнос a, $a > 0$. Затем первый игрок получает лучшую карту с вероятностью $P \geq 0$ и худшую, соответственно, с вероятностью $1 - P \geq 0$. В каждом раунде $0 < k < n$ первый игрок может сделать ставку $b_k > 0$, либо спасовать. Если он пасует, то происходит вскрытие карт и выигрывает игрок с лучшей картой. Если же первый игрок делает ставку, то ход переходит ко второму игроку. У него есть 2 варианта: спасовать или уравнять ставку первого игрока. Если он пасует, то игра заканчивается и деньги забирает первый игрок. Иначе, начинается новый раунд $k+1$. Если игра дошла до раунда $k = n$, то в случае ставки от первого игрока и уравнивания этой ставки вторым игроком происходит вскрытие и выигрывает игрок с лучшей картой.

Предполагается, что на лучшей карте первый игрок всегда делает ставку(повышает). На худшей карте он либо сразу пасует, либо повышает до раунда k и пасует в следующем раунде. Это соответствует стратегии блефа(стратегия k). Всего он располагает $n + 1$ стратегией. Аналогично, стратегия $j = 1, ..., n$ второго игрока означает, что он уравнивает ставки до раунда j, затем пасует. Если $j = 0$, то второй игрок пасует сразу. Матрица выигрышей при этом определяется следующим образом: $a_{ij} = Ps_j - (1 - P)s_i$, если $i \leq j$, $a_{ij} = s_j$, в противном случае.

Введем следующие обозначения:

$$r_k = \frac{b_k}{2(a + b_1 + ... + b_{k-1}) + b_k},$$

$$Q_k = (\prod_{j=1}^{k}(1 + r_j))^{-1}, Q_0 = 1,$$

$$P_k = \frac{s_k + a}{2s_k}, s_k = a + b_1 + ... + b_k, s_0 = a, k = 1, ..., n.$$

$$p_k = \frac{\prod_{i=k}^{n}(1 + r_i) - 1}{\prod_{j=k-1}^{b}(1 + r_i) - 1}$$

Лемма. *Для всех* $k = 1, ..., n$ *выполняется неравенство* $Q_k < P_k$.

Доказательство:

$$Q_k = \frac{(2s_0 + b_1)...(2s_{k-1} + b_k)}{2^k s_1 ... s_k} = Q_{k-1}\frac{2s_{k-1} + b_k}{2s_k} = Q_{k-1}\frac{2s_k - b_k}{2s_k}.$$

Отсюда следует, что $Q_k < Q_{k-1}$.

Покажем, что $P_k < P_{k-1}$.

$$\frac{P_k}{P_{k-1}} = \frac{s_k + a}{2s_k}\frac{2s_{k-1}}{s_{k-1} + a} = \frac{s_{k-1}(s_k + a)}{s_k(s_{k-1} + a)} = \frac{s_{k-1}s_k + as_k - ab_k}{s_{k-1}s_k + as_k} =$$
$$= 1 - \frac{ab_k}{s_{k-1}s_k + as_k} < 1.$$

Следовательно:

$$Q_k = Q_{k-1}\frac{2s_k - b_k}{2s_k},$$

$$P_k = P_{k-1}\Big(1 - \frac{ab_k}{s_{k-1}s_k + as_k}\Big).$$

Исходная лемма верна, если:

$$Q_{k-1}\frac{2s_k - b_k}{2s_k} < P_{k-1}\Big(1 - \frac{ab_k}{s_{k-1}s_k + as_k}\Big),$$

$$Q_{k-1} < P_{k-1}\Big(1 - \frac{ab_k}{s_{k-1}s_k + as_k}\Big)\frac{2s_k}{2s_k - b_k}, \tag{28}$$

Проведем доказательство по индукции. При $k = 2$ покажем, что $Q_2 < P_2$:

$$Q_2 - P_2 = \frac{(2a+b_1)(2a+2b_1+b_2)}{4(a+b_1)(a+b_1+b_2)} - \frac{2a+b_1+b_2}{2(a+b_1+b_2)} =$$
$$= \frac{(2a+b_1)(2a+2b_1+b_2) - 2(2a+b_1+b_2)(a+b_1)}{4(a+b_1)(a+b_1+b_2)} =$$
$$= -\frac{b_1b_2}{4(a+b_1)(a+b_1+b_2)} < 0$$

Пусть верно $Q_{k-1} < P_{k-1}$. Покажем, что и $Q_k < P_k$. Из (28) следует, что нам необходимо проверить условие:

$$\Big(1 - \frac{ab_k}{s_{k-1}s_k + as_k}\Big)\frac{2s_k}{2s_k - b_k} > 1.$$

Если оно выполняется, то лемма доказана.

$$\Big(1 - \frac{ab_k}{s_{k-1}s_k + as_k}\Big)\frac{2s_k}{2s_k - b_k} = \frac{2(s_k^2 - s_kb_k + as_k - ab_k)}{(s_k - b_k + a)(2s_k - b_k)} =$$
$$= \frac{2s_k^2 - 2s_kb_k + 2as_k - 2ab_k}{2s_k^2 - 2s_kb_k + 2as_k - 2ab_k + b_k(b_k + a - s_k)} = \frac{A}{A + b_k(b_k + a - s_k)}.$$

Так как $b_k + a - s_k < 0$, то $\frac{A}{A+b_k(b_k+a-s_k)} > 1$. Отсюда немедленно следует, что $Q_k < P_k$. Что и требовалось доказать.

Утверждение. *При $Q_k < P \leq Q_{k-1}$, $P < P_n$ значение игры равно a, второй игрок пасует в первом раунде(использует оптимальную стратегию 0), а первый игрок имеет оптимальную стратегию вида*

$$x_i^* = \begin{cases} 0, & i = 0, k+1, ..., n \\ r_i Q_i/(1-P), & i = 1, ..., k-1 \\ (Q_{k-1} - P)/(1-P), & i = k \end{cases}.$$

Если $P_k \leq P < P_{k-1}$, то матрица игры имеет седловую точку: значение игры равно a, первый игрок использует любую оптимальную стратегию $i \in \{k, ..., n\}$, а второй игрок - оптимальную стратегию 0.

Доказательство:

Покажем, что стратегия x_i^* действительно оптимальна. Для начала рассмотрим столбец матрицы a_{ij}, $j = 1$. Умножим скалярно этот столбец на соответствующую стратегию.

$$\frac{r_1 Q_1}{1-P}(2P-1)s_1 + s_1 \frac{\sum_{i=2}^{k-1} r_i Q_i + Q_{k-1} - P}{1-P}) =$$
$$= s_1 \frac{r_1 Q_1}{1-P}(2P-1) + s_1 \frac{1 - r_1 Q_1 - P}{1-P} =$$
$$= \frac{s_1}{1-P}(r_1 Q_1 (2P-1) + 1 - P - r_1 Q_1) =$$
$$= \frac{s_1}{1-P}(2(P-1) r_1 Q_1 + 1 - P) =$$
$$= s_1(1 - 2r_1 Q_1) = s_1(1 - \frac{b_1}{s_1})) = a.$$

Покажем, что для каждого столбца $j = 1, ..., k$ соответствующие скалярные произведения равны между собой и их значения при этом равны a. Если это, действительно так, то утверждение доказано.

Фиксируем столбец j. Тогда соответствующее скалярное произведение равно:

$$\sum_{i=1}^{j}\frac{r_iQ_i}{1-P}(Ps_j-(1-P)s_i)+s_j\frac{\sum_{i=j+1}^{k-1}r_iQ_i+Q_{k-1}-P}{1-P}=$$

$$=\sum_{i=1}^{j}\frac{r_iQ_iPs_j}{1-P}-\sum_{i=1}^{j}\frac{r_iQ_is_i(1-P)}{1-P}+s_j(\frac{1-\sum_{i=1}^{j}r_iQ_i}{1-P})=$$

$$=\frac{Ps_j(1-Q_j)}{1-P}+s_j\frac{Q_j-P}{1-P}-(s_jQ_j-a)=$$

$$=\frac{Ps_j-Ps_jQ_j-Ps_j+s_jQ_j}{1-P}-s_jQ_j+a=a.$$

То есть для каждого столбца $j=1,...,k$ значения одинаковы и равны a. Отсюда следует, что стратегий x_i^* оптимальная. Что и требовалось доказать.

В процессе доказательства мы использовали следующую лемму.

Лемма. *Для всех j выполняется равенство $s_jQ_j-a=\sum_{i=1}^{j}r_iQ_is_i$.*

Доказательство:

Доказательство проведем по индукции. При $j=1$:

$$s_1Q_1-a=r_1Q_1s_1,$$

$$(1-r_1)s_1Q_1=a,$$

$$(1-r_1)s_1Q_1=a,$$

$$(\frac{1-r_1}{1+r_1})s_1=a.$$

Последнее верно, так как:

$$s_{j-1}=(\frac{1-r_j}{1+r_j})s_j.$$

Пусть лемма верна при j. Докажем, что она верна при $j+1$.

$$s_{j+1}Q_{j+1} - a = \sum_{i=1}^{j+1} r_i Q_i s_i.$$

Вычтем из этого равенства равенство при j, получаем:

$$s_{j+1}Q_{j+1} - s_j Q_j = r_{j+1}Q_{j+1}s_{j+1},$$

$$s_{j+1}\left(\frac{1 - r_{j+1}}{1 + r_{j-1}}\right)Q_j = s_j Q_j,$$

$$s_{j+1}\left(\frac{1 - r_{j+1}}{1 + r_{j-1}}\right) = s_j.$$

Последнее верно. Следовательно равенство выполняется при $j+1$, т.е. лемма верна.

Утверждение. *Если $P < Q_n$, то оптимальным решением для второго игрока будет пас на раунде k с вероятностью $p = r_k$, оптимальным решением для первого игрока будет ставка с лучшей картой. С худшей картой оптимальной будет следующая стратегия: ставка в первом раунде с вероятностью $p_1 = \frac{P}{1-P}\frac{1-Q_n}{Q_n}$, ставка в раунде $k > 1$ с вероятностью p_k. Значение игры $V = \frac{a(2P-Q_n)}{Q_n}$.*

Доказательство:

Пусть A_{ij} - ожидаемый выигрыш первого игрока при использовании им стратегии i и при использовании вторым игроком стратегии j.

$$A_{ij} = \begin{cases} Ps_j - (1-P)s_i, & 0 \le i \le j \le n \\ s_j, & 0 \le j < i \le n \end{cases}.$$

Пусть $(\sigma_0, ..., \sigma_n)$ - смешанная стратегия второго игрока, где σ_j является вероятностью того, что второй игрок уравняет ставку первого игрока ровно j раз. Если второй игрок использует эту стратегию, а первый использует стратегию i, то средний выигрыш равен:

$$V_i = \sum_{j=0}^{n} A_{ij}\sigma_j = \sum_{j=0}^{i-1} s_j\sigma_j + P\sum_{j=i}^{n} s_j\sigma_j - (1-P)s_i\sum_{j=i}^{n}\sigma_j =$$

$$= P\sum_{j=0}^{n} s_j\sigma_j + (1-P)\sum_{j=0}^{i-1} s_j\sigma_j - (1-P)s_i\sum_{j=i}^{n}\sigma_j.$$

Найдем σ_i из условия независимости V_i от i.

$$V_k - V_{k-1} = (1-P)\Big[(s_k + s_{k-1})\sigma_{k-1} - (s_k - s_{k-1})\sum_{j-k-1}^{n}\sigma_j\Big] = 0.$$

Отсюда следует для $k = 1, ..., n$:

$$\frac{\sigma_{k-1}}{\sum_{j=k-1}^{n}\sigma_j} = \frac{s_k - s_{k-1}}{s_k + s_{k-1}} = r_k. \tag{29}$$

Уравнение (29) описывает σ_k. Фактически, его левая часть представляет собой вероятность паса вторым игроком в раунде k.

$$V_0 = P\sum_{j=0}^{n} s_j\sigma_j - (1-P)s_0,$$

$$V_n = \sum_{j=0}^{n-1} s_j\sigma_j + Ps_n\sigma_n - (1-P)s_n\sigma_n = \sum_{j=0}^{n} s_j\sigma_j - 2(1-P)s_n\sigma_n.$$

Из условия $V_0 = V_n$ получаем, что $\sum_{j=0}^{n} s_j\sigma_j = 2s_n\sigma_n - s_0$. Следовательно, $V_0 = 2Ps_n\sigma_n - s_0$. Используя (29) имеем $V_0 = 2Ps_n\prod_{i=1}^{n}(1 - r_i) - s_0$.

Так как $1-r_k=\frac{2s_{k-1}}{s_k+s_{k-1}}$, а $1+r_k=\frac{2s_k}{s_k+s_{k-1}}$, то

$$s_n\prod_{k=1}^{n}(1-r_k)=s_0\prod_{k=1}^{n}(1+r_k).$$

$$V_0=a\Big(2P\prod_{k=1}^{n}(1+r_k)-1\Big)=a\Big(\frac{2P}{Q_n}-1\Big).$$

Теперь перейдем к первому игроку. Пусть $\pi=(\pi_0,...,\pi_n)$ - смешанная стратегия первого игрока, где π_i - вероятность сделать ровно i ставок. Если первый игрок использует такую смешанную стратегию, а второй игрок использует стратегию j, то средний выигрыш при $0\le j\le n$ равен:

$$W_j=\sum_{i=0}^{n}\pi_iA_{ij}=\sum_{i=0}^{j}\pi_i(Ps_j-(1-P)s_i)+\sum_{i=j+1}^{n}\pi_is_j=$$
$$=Ps_j+(1-P)s_j\sum_{i=j+1}^{n}\pi_i-(1-P)\sum_{i=0}^{j}s_i\pi_i.$$

Приравнивая W_j к W_{j-1} для $0<j\le n$, получим:

$$P(s_j-s_{j-1})+(1-P)(s_j-s_{j-1})\sum_{i=j+1}^{n}\pi_i-(1-P)(s_j+s_{j-1})\pi_j=0,$$

$$\pi_j=\frac{P}{1-P}r_j+r_j\sum_{i=j+1}^{n}\pi_i.$$

Так как $\pi_n=r_n\frac{P}{1-P}$ и $\sum_{i=j+1}^{n}\pi_i=\frac{P}{1-P}\Big[\prod_{i=j+1}^{n}(1+r_i)-1\Big]$, то поведение первого игрока в раунды $1<j\le n$ описывается следующим образом:

$$p_j = \frac{\sum_{i=j}^{n} \pi_i}{\sum_{i=j-1}^{n} \pi_i} = \frac{\prod_{i=j}^{n}(1+r_i)-1}{\prod_{i=j-1}^{n}(1+r_i)-1}.$$

Стратегия в первом раунде:

$$p_1 = 1 - \pi_0 = \sum_{i=1}^{n} \pi_i = \frac{P}{1-P}\Big[\prod_{i=1}^{n}(1+r_i)-1\Big].$$

Покажем, что $W_0 = V_0$.

$$W_0 = Ps_0 + (1-P)s_0\sum_{i=1}^{n} \pi_i - (1-P)s_0\pi_0 = 2Ps_0\prod_{i=1}^{n}(1+r_i) - s_0 = V_0.$$

То есть $W_0 = V_0$ является значением игры. А значит утверждение верно. Что и требовалось доказать.

5 Заключение

В данной работе представлены следующие результаты:

1. Расчет оптимальных стратегий в модели покера с фиксированной ставкой.

2. Исследование единственности оптимальных стратегий.

3. Расчет оптимальных стратегий в модели с двумя ставками.

4. Расчет оптимальных стратегий в модели с конечным числом ставок.

5. Исследование асимптотических свойств стратегий.

6. Расчет оптимальных стратегий в модели с n раундами ставок.

При исследовании моделей были сделаны предположения о равномерном распределении значения карты x, а также независимость случайных величин x, y. В рамках этих предположений мы рассчитали в численном виде стратегии для каждой рассмотренной модели покера.

Список литературы

[1] Морозов В.В., Байрамкулов А.М. О решении простейшей модели покера с двумя участниками. Тезисы докладов конференции "Тихоновские чтения". М.: МАКС-Пресс, 2018.

[2] Мазалов В.В. Математическая теория игр и приложения. СПб.: Издательство «Лань», 2016.

[3] Мазалов В.В., Менчер А.Э., Токарева Ю.С. Переговоры. Математическая теория. СПб.: Издательство «Лань», 2012.

[4] Ferguson T., Ferguson C. The endgame in poker. - https://www.math.ucla.edu/ tom/papers/poker3.pdf.

[5] Ferguson T., Ferguson C. On the Borel and von Neumann Poker Models. - https://www.math.ucla.edu/ tom/papers/poker1.pdf.

[6] Васин А.А., Морозов В.В. Теория игр и модели математической экономики. М.: МАКС-Пресс, 2005.