Формирование эффективной кооперации в повторяющихся играх

4 декабря 2018

Аннотация

В докладе, основанном на работах A.Abdulkadiroglu, K.Bagwell'a, W.Olszewski, M.Safronov'a и др., рассматриваются равновесия в повторяющихся симметричных играх с двумя участниками и неполной информацией, в которых оптимальная стратегия состоит в передаче выигрыша партнёру, способному лучше его инвестировать. Проблема в том, чтобы придумать систему стимулов, обеспечивающую доверие игроков и способную имплементировать их кооперативное поведение. Оказывается, что если взаимодействие участников достаточно длительно, причём они многократно оказываются в сходных ситуациях и придают большое значение будущим выигрышам (т.е. коэффициент дисконтирования близок к 1), то (в простых моделях) систематический обмен услугами, воплощающий принцип "ты мне, я тебе", является (в пределе) эффективной стратегией, отклоняться от которой никому не выгодно.

Доверительные инвестиции (M.Möbius, A.Abdulkadiroğlu, K.Bagwell)

В повторяющейся игре с двумя идентичными участниками на каждом этапе возможны три исхода: первый либо второй участник получает \$1 — оба с вероятностью $p \in (0, 1/2)$, либо ни один из участников не получает дохода — вероятность такого исхода $1 - 2p$. Каждому участнику персонально сообщают, получил ли он деньги, т.е. является ли он инвестором (таким образом, если он не получил, то он не знает, получил ли их его партнёр). Получив \$1, участник может 'проявить доверие', передав $x \in [0, 1]$ партнёру; процесс передачи и уровень инвестирования публично наблюдаемы. Исход инвестирования случаен. Возможны два исхода: либо инвестирование удачно и приносит доход kx (это происходит с вероятностью q, либо нет; в последнем случае инвестированные деньги пропадают. Предполагается, что $qk > 1$, т.е. в среднем инвестирование выгодно.

Получатель денег (trustee) знает о результате инвестирования, а инвестор нет. В случае успеха доверенное лицо может вернуть все деньги или их часть (включая доход) инвестору — немедленно или через некоторое время, однако если инвестору денег не возвращают, то он не знает, было ли инвестирование неудачным или доверенное лицо предпочло оставить деньги себе. Уровень дисконта δ одинаков для обоих участников. Эта модель позволяет исследовать уровень доверия и кооперации между игроками. Например, для некоторого класса стратегий оказывается, что если игрок 'задолжал' партнёру услугу, то чем длиннее 'нейтральный' интервал, в котором ни один из игроков не получает дохода, тем меньше инвестиция x.

Инвестиции в обмен на талоны

Игра с двумя идентичными участниками многократно повторяется. На каждом этапе возможны три исхода: первый либо второй участник получает \$1 — оба с вероятностью $p \in (0, 1/2)$, либо ни один из участников не получает дохода — вероятность такого исхода $1 - 2p$. Каждому участнику персонально сообщают, получил ли он деньги (таким образом, если он не получил, то он не знает, получил ли их его партнёр). Получив деньги, участник может передать их партнёру; при этом выигрыш партнёра составит $\gamma > 1$, т.е. больше, чем теряет участник. Уровень дисконта δ одинаков для обоих участников. Выигрыши нормализуются множителем $1 - \delta$. Хранение денег не предусмотрено: получив выигрыш, участник либо потребляет его, либо передаёт партнёру. Таким образом, эффективный общий выигрыш обоих участников составляет $v = 2p\gamma$ и достигается, когда каждый из участников передаёт свой выигрыш другому.

Талонная стратегия

Рассмотрим следующую *талонную стратегию*. Имеется $2n$ талонов, в начале каждого периода игрок i имеет $0 \leq k_i \leq 2n$ талонов, $k_1 + k_2 = 2n$. Если игрок i получает \$1 и $k_i < 2n$, то i-ый игрок передаёт свой выигрыш j-му ($j \neq i$), а в обмен получает один талон. Если же $k_i = 2n$, то i сам потребляет \$1. В начале игры $k_1 = k_2 = n$.

Предложение

$\forall \lambda > 0 \; \exists \underline{\delta} < 1 \; \exists n$ *такие, что* $\delta > \underline{\delta} \Rightarrow$ *талонная стратегия с* n *талонами даёт равновесие в повторяющейся игре с ожидаемым выигрышем* $> v/2 - \lambda$.

Обозначим через V_k ожидаемый выигрыш (continuation payoff) первого игрока, когда он имеет k талонов (эти выигрыши считаются перед тем, как игроки узнают о своих доходах в текущий период). Имеем

$$\begin{aligned}
V_k &= p[(1-\delta)\gamma + \delta V_{k-1}] + p\delta V_{k+1} + (1-2p)\delta V_k, \quad 1 \le k \le 2n-1;\\
V_0 &= p\delta V_0 + p\delta V_1 + (1-2p)\delta V_0;\\
V_{2n} &= p[(1-\delta)\gamma + \delta V_{2n-1}] + p[(1-\delta) + \delta V_{2n}] + (1-2p)\delta V_{2n},
\end{aligned}$$

где первое слагаемое соответствует возможности получения \$1 вторым игроком, второе — возможности получения \$1 первым игроком, а третье — случаю, когда ни один из игроков не получает дохода.

Лемма

$\forall n \ \forall \lambda > 0 \ \exists \underline{\delta}' > 0 \mid \forall \delta > \underline{\delta}'$ *ожидаемый выигрыш каждого из участников превосходит* $\dfrac{2n-1}{2n+1} p\gamma - \lambda$.

Доказательство

Талонная стратегия индуцирует марковскую цепь на состояниях $k = 0, \ldots, 2n$. Согласно эргодической теореме, существуют вероятности π_k, $k = 0, \ldots, 2n$ такие, что, независимо от начального состояния, через достаточно большое число шагов вероятность нахождения в состоянии k сколь угодно близка к π_k. Если

$$T = \begin{pmatrix} 1-p & p & 0 & . & 0 \\ p & 1-2p & p & . & . \\ 0 & . & . & . & . \\ . & . & . & . & . \\ . & . & . & . & . \\ . & . & . & . & 0 \\ 0 & . & p & 1-2p & p \\ 0 & . & 0 & p & 1-p \end{pmatrix} \tag{1}$$

— матрица перехода, в которой коэффициент в i-ой строке и j-ом столбце равен вероятности перехода из j-го состояния в i-ое, то вектор $\boldsymbol{\pi} = (\pi_0, \ldots, \pi_{2n})$ — собственный вектор матрицы T, отвечающий собственному значению 1. Нетрудно проверить, что $\pi_k = \frac{1}{2n+1}$, $k = 0, \ldots, 2n$. Поток платежей каждого из игроков в каждом состоянии $k \neq 0, 2n$ равен $p\gamma$, а предельная вероятность пребывания в состоянии, отличном от 0 и $2n$, равна $\frac{2n-1}{2n+1}$. Поэтому если δ близко к 1, то при $k \neq 0, 2n$ ожидаемый выигрыш ограничен снизу любым числом, меньшим $p\gamma(2n-1)/(2n+1)$. $\square$

Осталось доказать, что ни одному из игроков не выгодно отклоняться от талонной стратегии. Единственный способ отклониться в этой модели — это использовать полученный доход самому вместо того, чтобы передавать его партнёру. Выигрыш в этом случае составил бы $1 - \delta$, а проигрыш состоит в том, что в следующем периоде у этого игрока будет на один талон меньше.

Осталось доказать, что ни одному из игроков не выгодно отклоняться от талонной стратегии. Единственный способ отклониться в этой модели — это использовать полученный доход самому вместо того, чтобы передавать его партнёру. Выигрыш в этом случае составил бы $1-\delta$, а проигрыш состоит в том, что в следующем периоде у этого игрока будет на один талон меньше.

Лемма

Положим $\Delta_k = V_k - V_{k-1}$, $k = 1, \ldots, 2n$. $\forall n\ \exists \underline{\delta}'' < 1$, $\exists \kappa > 1$ *такое, что* $\Delta_k > \kappa(1-\delta)$, $k = 1, \ldots, 2n$.

Доказательство.

Из (1) следует, что

$$\begin{aligned}\Delta_k &= p\delta\Delta_{k-1} + p\delta\Delta_{k+1} + (1-2p)\delta\Delta_k, \quad k = 1,\ldots,2n-1;\\ \Delta_1 &= p(1-\delta)\gamma + p\delta\Delta_2 + (1-2p)\delta\Delta_1;\\ \Delta_{2n} &= p(1-\delta) + p\delta\Delta_{2n-1} + (1-2p)\delta\Delta_{2n}.\end{aligned}$$

При $\delta = 1$ решением этой системы уравнений является $\boldsymbol{\Delta} = 0$. В окрестности $\delta = 1$, $\boldsymbol{\Delta} = 0$, заменяя $\boldsymbol{\Delta}$ на $\dfrac{\partial\boldsymbol{\Delta}}{\partial\delta}$, получаем для $\dfrac{\partial\boldsymbol{\Delta}}{\partial\delta}$ систему уравнений, единственным решением которой является

$$\frac{\partial\Delta_k}{\partial\delta} = -\frac{(2n-k+1)\gamma + k}{2n+1} = -\gamma + \frac{(\gamma-1)k}{2n+1} < -1, \quad k = 1,\ldots,2n.$$

Поэтому при δ, близком к 1 и некотором $1 < \kappa$ требования Леммы выполнены. (γ *входит только в* Δ_1.) □

Повторяющиеся аукционы

В каждом периоде два игрока участвуют в аукционе первой цены; в случае равных ставок устраивается лотерея 50 на 50. Каждый из участников получает приватный сигнал об объекте, который может принимать два значения H или L. Эти сигналы симметричны, независимы и распределены со следующими вероятностями:

$$\text{сигнал 2-го игрока}\quad \begin{array}{c|cc} & L & H \\ \hline H & q & 2p \\ L & 1-2p-2q & q \end{array},$$

сигнал 1-го игрока

где $p, q \geq 0$, $p + q < 1/2$. Каждый из игроков оценивает объект на основании двух сигналов — своего и партнёра, причём оба имеют одинаковую функцию ценности v.

Оценки строго положительны и монотонны, причём свой сигнал имеет большее значение, чем сигнал оппонента, т.е. $v(H,H) > v(H,L) > v(L,H) > v(L,L) > 0$. Эффективный победитель даётся следующей таблицей:

$$\begin{array}{c} \text{сигнал 2-го игрока} \end{array} \quad \begin{pmatrix} H & \text{игрок 2} & \forall \\ L & \forall & \text{игрок 1} \\ & L & H \end{pmatrix}.$$
$$\text{сигнал 1-го игрока}$$

Описание эффективных талонных стратегий

В начале каждого периода i-ый игрок имеет $0 \leq k_i \leq 2n$, $k_1 + k_2 = 2n$ талонов, игра начинается с $k_1 = k_2 = n$. При $k_i \neq 0, 2n$ игроки правдиво сообщают друг другу свои сигналы (cheap talk), и если эти сигналы совпадают, то победитель данного периода определяется с помощью лотереи 50 на 50. Предложение ставок: a) ставка игрока ρ, если он получил сигнал H, а его соперник L или если они получили одинаковые сигналы, но он выиграл в лотерее; b) ставка нулевая, если он получил сигнал L, а его соперник H или если они получили одинаковые сигналы, но он проиграл в лотерее. Игрок с $2n$ талонами делает ставку ρ, а игрок с 0 талонами — ставку 0 независимо от сигналов. По окончании каждого периода количество талонов у победителя уменьшается на один, а у проигравшего увеличивается на один.

Предложение

$\forall \lambda > 0\ \exists \underline{\delta} < 1\ \ \exists n$ такие, что $\delta > \underline{\delta} \Rightarrow$ талонная стратегия с n талонами даёт равновесие в повторяющемся аукционе с общим выигрышем, отличающемся от эффективного общего выигрыша меньше, чем на λ.

Обозначим через V_k ожидаемый выигрыш первого игрока, когда он имеет k талонов, $0 \leq k \leq 2n$ (подсчёт выполнен до того, как игроки получат очередные сигналы). Имеем

$$\begin{aligned} V_k &= (1-\delta)C + \delta V_{k-1}/2 + \delta V_{k+1}/2, \qquad 1 \leq k \leq 2n-1, \\ V_0 &= \delta V_1, \\ V_{2n} &= (1-\delta)D + \delta V_{2n-1}, \end{aligned}$$

где $C = pv(H,H) + qV(H,L) + (1/2 - p - q)v(L,L)$ и $D = 2pv(H,H) + qv(H,L) + (1-2p-2q)v(L,L) + qv(L,H)$ — эффективные выигрыши первого участника за соответствующие периоды.

Лемма

$\forall n \ \forall \lambda > 0 \ \exists \underline{\delta}' > 0 \mid \forall \delta > \underline{\delta}'$ *имеем* $V_k > \dfrac{2n-1}{2n} C - \lambda$, $k = 0, \ldots, 2n$.

Доказательство.

Эргодическая теорема. В данном случае

$$\pi_k = \begin{cases} \frac{1}{2n}, & k = 1, \ldots, 2n-1, \\ \frac{1}{4n}, & k = 0, 2n \end{cases}.$$

$\square$

Лемма

$\forall n\ \exists \underline{\delta}'' < 1$ такое, что $\forall \delta > \underline{\delta}''$ талонная стратегия даёт равновесие в повторяющейся игре.

Доказательство

Проверим, что участникам выгодно правдиво сообщать о полученных сигналах. Если первый игрок солжёт, то в текущем периоде он потеряет

$$(1-\delta)\left[\frac{p}{2p+q}v(H,H)+\frac{q/2}{2p+q}v(H,L)\right], \qquad H \Rightarrow L$$

и выиграет

$$(1-\delta)\left[\frac{q/2}{1-2p-q}v(L,H)+\frac{1/2-p-q}{1-2p-q}v(L,L)\right], \qquad L \Rightarrow H.$$

Если первый игрок говорит, что получил L, а на самом деле получил H, то его шанс получить $k+1$ талон (вместо $k-1$) увеличивается на $\dfrac{p+q/2}{2p+q} = \dfrac{1}{2}$; если же он говорит, что получил H, а на самом деле получил L, то его шанс получить $k+1$ талон (вместо $k-1$) уменьшается на $\dfrac{1/2-p-q/2}{1-2p-q} = \dfrac{1}{2}$.
Поэтому нужно проверить, что

$$(1-\delta)\left[\frac{p}{2p+q}v(H,H) + \frac{q/2}{2p+q}v(H,L)\right] > \delta\frac{V_{k+1}-V_{k-1}}{2},$$

$$(1-\delta)\left[\frac{q/2}{1-2p-q}v(L,H) + \frac{1/2-p-q}{1-2p-q}v(L,L)\right] < \delta\frac{V_{k+1}-V_{k-1}}{2},$$

$k = 1, \ldots, 2n$.

Участникам также невыгодно отклоняться от талонной стратегии на этапе объявления ставок, поскольку если бы такое отклонение повлияло на исход аукциона, оно было бы обнаружено: проигравший стал бы победителем, а в этом случае участники переключились бы на плохую игру, описанную ниже и дающую выигрыш меньше половины эффективного. Этот вариант хуже талонной стратегии, при которой V_k сходится к половине общего эффективного выигрыша при всех $k = 0, \ldots, 2n$, коэффициенте дисконтирования δ, сходящемся к 1 и достаточно больших n. $\square$

Если один из игроков выиграет аукцион, когда "должен был" выиграть его партнёр (что невозможно при следовании талонным стратегиям), то игроки переключаются на плохую симметричную игру, в которой выигрыш каждого из игроков меньше половины эффективного. Опишем эту игру. Игрок, получивший сигнал L, делает ставку $v(L,L)$ с вероятностью 1. Игрок, получивший сигнал H, делает ставку в соответствии с распределением $G(b)$ с носителем в интервале

$$\left[v(L,L), E(v|H) - q\frac{v(H,L)-v(L,L)}{q+2p}\right],$$

где $E(v|H) = \frac{qv(H,L)}{q+2p} + 2p\frac{v(H,H)}{q+2p}$ — ожидаемая ценность объекта при условии получения сигнала H. В случае если игрок, получивший сигнал H, делает ставку b, его выигрыш составляет

$$q\frac{v(H,L)-b}{q+2p} + \frac{2p}{q+2p}G(b)[v(H,H)-b], \qquad (2)$$

поскольку он выигрывает с вероятностью 1 у противника, получившего сигнал L, и с вероятностью $G(b)$ у противника, получившего сигнал H.

Выражение (2) не зависит от b, так что, дифференцируя по b, получаем

$$\frac{2p}{q+2p}G'(b)[v(H,H)-b] = \frac{q}{q+2p} + \frac{2p}{q+2p}G(b),$$

так что $G(b) = \dfrac{c}{2p[v(H,H)-b]} - \dfrac{q}{2p}$, где c константа.
Поскольку $G(v(L,L)) = 0$, $c = q[v(H,H) - v(L,L)]$, так что

$$G(b) = \frac{q}{2p}\frac{v(H,H)-v(L,L)}{v(H,H)-b} - \frac{q}{2p}$$

и, в частности,

$$G\big(E(v|H) - q\frac{v(H,L)-v(L,L)}{q+2p}\big) = 1.$$

В этом равновесии выигрыш игрока, получившего сигнал L, равен нулю ($b = v(L,L)$), а выигрыш игрока, получившего сигнал H, равен $q\dfrac{v(H,L)-v(L,L)}{q+2p}$. Ожидаемый выигрыш равен $q[v(H,L) - v(L,L)]$, что строго меньше, чем половина эффективного выигрыша, равного
$pv(H,H) + qv(H,L) + (1/2 - p - q)v(L,L)$.

Дуополия Spulber'a

Имеются две фирмы, производящие один и тот же продукт. Затраты на производство единицы продукта у каждой из фирм могут составить либо $\underline{c}$ либо $\bar{c} > \underline{c}$ и формируются согласно независимым марковским случайным процессам. С вероятностью $p \in [1/2, 1)$ затраты фирм в следующем периоде совпадают с текущими затратами, а с вероятностью $1 - p$ отличаются от них. В каждом периоде фирмы одновременно назначают цены. Покупатель готов уплатить любую цену вплоть до $r > \bar{c} > \underline{c}$ за единицу продукта, но покупает у фирмы, которая предлагает меньшую цену. Фирмы максимизируют свой ожидаемый доход и дисконтируют будущий выигрыш с коэффициентом $\delta < 1$. В начальный период 0 затраты фирм принимают значения $\underline{c}$ и $\bar{c}$ с 50 %-ой вероятностью. Эффективный общий выигрыш равен

$$v = r - \frac{3}{4}\underline{c} - \frac{1}{4}\bar{c}.$$

Предложение

$\forall \lambda > 0\ \exists \underline{\delta} < 1\ \ \exists n$ такие, что $\delta > \underline{\delta} \Rightarrow$ талонная стратегия с n талонами даёт равновесие в повторяющемся аукционе с выигрышем каждой из фирм, превышающем $v/2 - \lambda$.

Талонные стратегии строятся следующим образом. Первоначально обе фирмы имеют по n талонов. Если в данном периоде одна из фирм устанавливает цену ниже, чем другая, то она продаёт товар покупателю, но даёт другой фирме (виртуальный) талон; если же у одной из фирм талонов больше нет, то она уступает другой фирме право обслужить покупателя в этом периоде, а взамен получает талон. Если одна из фирм уклоняется от талонного поведения, то фирмы начинают играть в “плохую динамическую игру” (см. ниже). При этом фирма с меньшими затратами устанавливает цену $r - \rho$, где ρ — малое число, а фирма с большими затратами — цену r. Если у второй фирмы не остаётся талонов, то первая назначает цену $r - \rho$, а вторая r.

Осталось описать “плохие” равновесия. Вначале, как только одна из фирм отклонится от предписанной цены, обе фирмы назначают цену $\gamma < r$, после чего с вероятностью μ обе назначают цену r, а с вероятностью $1 - \mu$ — цену γ. Как только в некотором периоде обе фирмы назначат цену r, они и в дальнейшем продолжают назначать максимальную цену с вероятностью 1. Когда коэффициент дисконтирования стремится к 1, выручка каждой фирмы ограничена $r/2 - \underline{c}/2 - \bar{c}/2$, что меньше эффективной выручки.

Посменные задания

Двое студентов живут в общей съёмной квартире и в каждом периоде должны произвести уборку. Затраты на уборку составляют либо $\underline{c}$, либо $\bar{c} > \underline{c}$, а выигрыш каждого от убранной квартиры 0, а от неубранной -1. Затраты участников распределены независимо и марковские по времени. В эффективной стратегии работу выполняет тот, у кого меньше затраты.

Талонная стратегия: если тебе выпало $\underline{c}$, а соседу $\bar{c}$, то убрать и дать талон, а если затраты одинаковы, то каждый выполняет уборку с вероятностью $1/2$. Если же у одного из участников не осталось талонов, то другой убирает независимо от своих затрат и выдаёт талон соседу. При отклонении плохое равновесие, а именно квартиру никто не убирает.

Недостатки и обобщения

При доказательстве того, что участники правдиво сообщают свой тип и им невыгодно отклоняться от предписанных стратегий существенно использовались состояния V_0 и V_{2n}, а также предписание играть в "плохую" игру. На мой взгляд, это слабые места.

При доказательстве возможности приблизить эффективные исходы простыми талонными стратегиями мы существенно использовали, что игра симметрична, в ней двое участников, а поступающие сигналы бывают двух типов. От симметричности можно отказаться, введя делимость талонов, что можно осуществить с помощью лотерей (так что участники получают талоны с некоторой вероятностью). Однако если сигналы бывают более, чем двух видов, то и такие рандомизированные талонные стратегии не обязательно аппроксимируют эффективные исходы (даже в симметричном случае). В случае, когда число участников больше двух, услуга может оказываться более, чем одним игроком, и выиграть от неё также могут несколько участников, так что неясно, кто должен давать и получать талоны.

Спасибо за внимание!