Теоремы о невозможности

Сергей Николенко

Теория экономических механизмов
Computer Science Club, осень 2008

Outline

Голосования

- Мы начинаем с вопроса, который, на первый взгляд, даже не лежит в области экономики.
- Мы будем рассматривать *голосования*, возможные схемы голосований и то, к каким результатам они могут привести.
- Голосование — простой и естественный частный случай экономического механизма.
- Множество возможных исходов: кандидаты A, B и C.

Голосования

- У каждого участника голосования (агента) есть определённый порядок на этих исходах (ограничимся случаем, когда этот порядок линейный).
- Например, кандидат A мне нравится больше, чем B, а B — больше, чем C:
$$A \succ B \succ C.$$
- Этот порядок можно рассматривать как скрытую функцию предпочтений агента.
- Функция социального выбора определяет, какой кандидат должен победить при том или ином соотношении голосов.

Голосования

- Этот частный случай вместе с тем оказывается и наиболее общим.
- Мы не предполагаем вообще никаких ограничений, никакой структуры на множестве предпочтений каждого из агентов; любой исход может оказаться на любом месте в его внутренней функции предпочтения.
- Поэтому неудивительно, что не всё получится реализовать. Наши результаты о невозможности окажутся весьма полезными в доказательстве результатов о невозможности в теории экономических механизмов.

Наши цели

- Что бы мы хотели получить от системы голосования? Каковы цели, которых мы будем (безуспешно) пытаться достигнуть?
- Рассмотрим простой и понятный случай голосования: случай, когда в нём участвует ровно один агент.
- Какими самыми базовыми, самыми естественными свойствами будет обладать множество предпочтений одного агента?

Наши цели

1. *Транзитивность*: если $A \succ B$ и $B \succ C$, то $A \succ C$.
2. *Попарная независимость предпочтений*: выбор между A и B зависит только от того, как соотносятся друг с другом A и B в моём «персональном рейтинге», и никак не зависит от положения там других альтернатив $C, D, \ldots$ Т.е. если вам предлагают выбор между персиком и апельсином, ваши предпочтения насчёт яблок не должны на этот выбор повлиять.

Наши цели

3 *Положительная ассоциированность*: если мои предпочтения изменились к лучшему для какой-либо альтернативы, то в результате голосования шансы этой альтернативы на победу могут только возрасти.

- Например, если раньше был профиль предпочтений $A \succ B \succ C$, а сейчас $B \succ A \succ C$, то шансы B в любом голосовании, даже против C, не должны от этого ухудшиться.

Наши цели

4. *Единогласие*. Если *все* участники голосования предпочитают возможный исход A другому возможному исходу B, то в результате голосования не может быть выбран B.

- Четвёртое свойство является по сути свойством функции социального выбора, а не свойством одного-единственного агента, как первые три (это просто оптимальность по Парето).

Когда всё получается

- Предположим, что возможных исходов всего два, то есть голосование превратилось в референдум.
- Тогда можно предложить систему голосования: выбирать нужную альтернативу большинством голосов.
- Выбор простым большинством из двух исходов удовлетворяет всем четырём интересующим нас свойствам.
- Однако для трёх и более возможных исходов голосования дело гораздо хуже...

Кондорсе

- 1785 год, Мари Жан Антуан Николя, маркиз де Кондорсе.
- Фактически первым начал применять математику к общественным наукам.
- Во время революции критиковал «конституцию монтаньяров», был против казни Людовика XVI — в общем, наработал на арест и приговор.

Метод Кондорсе

- Кондорсе разработал хороший метод голосования для нескольких кандидатов.
    1. Каждый участник ранжирует кандидатов.
    2. Считают «микроматчи» между кандидатами: в профиле каждого участника сравниваются все пары кандидатов и одному из пары записывается победа.
    3. Выигрывает тот, у кого суммарно больше побед.
- Но такие методы не всегда работают...

Парадокс Кондорсе

- Рассмотрим три возможных исхода A, B и C и трёх участников x, y и z. Предположим, что их предпочтения распределены так:

$$\begin{aligned} A &\succ_x B &\succ_x C, \\ B &\succ_y C &\succ_y A, \\ C &\succ_z A &\succ_z B. \end{aligned}$$

  Иначе говоря, предпочтения трёх участников получаются циклическим сдвигом одного линейного порядка.

Парадокс Кондорсе

- Если на выбор предложат A и B, то x и z проголосуют за A, и будет избран A: $A \succ B$.
- В борьбе между B и C выяснится, что $B \succ C$.
- Но если предложат выбор между A и C, то y и z проголосуют за C, и окажется, что $C \succ A$!

Парадокс Кондорсе

- В парадоксе Кондорсе нарушается *транзитивность* «мнения большинства».
- В частности, это значит, что если пытаться решать вопрос референдумами, то придётся вечно ходить по кругу.
- А метод Кондорсе даёт ничью.

Следствие парадокса Кондорсе: семь вариантов

- Можно придумать и ещё более любопытные следствия.
- Пусть у нас три агента — 1, 2, 3 — и семь возможных вариантов: A, B, C, D, E, F, G.

Следствие парадокса Кондорсе: семь вариантов

- И предпочтения у них вот какие:

1	2	3
A	C	D
B	D	A
C	A	G
D	F	B
E	G	C
F	B	E
G	E	F

Следствие парадокса Кондорсе: семь вариантов

- Они решают вопрос голосованием в следующем порядке:
  1. A против D: 1 в меньшинстве, D идёт дальше.
  2. D против C: 1 и 2 проводят дальше C.
  3. C против B: B побеждает и проходит в следующий бой.
  4. B против G: несмотря на то, что G ну совсем не мил агенту 1, он побеждает.
  5. G против F: выигрывает F.
  6. F против E: E побеждает, 2 в меньшинстве.
- В итоге победило *совсем* не оптимальное по Парето решение: вариант E никому особенно не интересен.
- Более того, есть несколько вариантов, которые *у каждого агента* стоят выше.

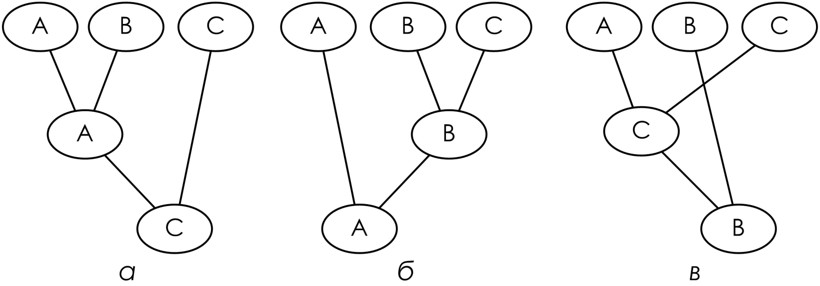
Следствие парадокса Кондорсе: поправка Сазерленда

- Другое следствие парадокса Кондорсе: то, что результат голосования может зависеть от порядка, в котором проводятся референдумы или другие «подголосования».
- Вспомним парадокс Кондорсе:

$$\begin{aligned} A &\succ_x B \succ_x C, \\ B &\succ_y C \succ_y A, \\ C &\succ_z A \succ_z B. \end{aligned}$$

Следствие парадокса Кондорсе: поправка Сазерленда

- Вот как результат зависит от порядка:



Следствие парадокса Кондорсе: поправка Сазерленда

- Получается, что результат при одних и тех же предпочтениях кардинально зависит от формата голосования!
- А значит, тот, кто контролирует формат голосования, имеет существенное преимущество и может победить, даже оказавшись в меньшинстве.
- Это приводит к тому, что попарная независимость предпочтений тоже не выполняется.

Следствие парадокса Кондорсе: поправка Сазерленда

- Рассмотрим ситуацию, в которой есть ровно две альтернативы: A и B, причём большинство хочет выбрать A.
- Тогда простым большинством A без проблем выберут.
- Но если у меньшинства получится построить такую третью возможность C, что при выборах $C \succ A$ и $B \succ C$, то это меньшинство сможет, установив правильный порядок выборов (сначала A против C, затем B против победителя), провести B, а не A.

Следствие парадокса Кондорсе: поправка Сазерленда

- В политике такие ситуации редко, но действительно возникают на практике.
- Они называются «поправки-убийцы» (killer amendments).
- Пример: в США сенаторов поначалу выбирали не прямым всенародным голосованием, а законодательными органами соответствующего штата. В том, чтобы ввести голосования на пост сенатора, заключалась 17-я поправка к Конституции США, которая в конце концов всё же была принята в 1913.
- Но приняли её не сразу.

Следствие парадокса Кондорсе: поправка Сазерленда

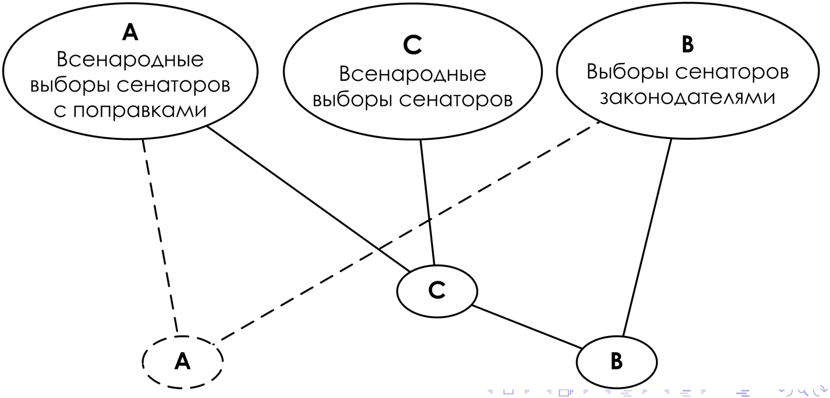
- Южные сенаторы беспокоились, что если федеральное государство возьмёт выборы сенаторов под свой контроль, то северяне-республиканцы сделают что-нибудь ужасное, например допустят на выборы чернокожих.
- Был достигнут компромисс: билль, который вводил прямые выборы сенаторов, но содержал поправки, ограничивающие контроль федерального правительства над выборами в южных штатах.
- Его поддерживало большинство (это была возможность A), и на прямом голосовании между этим биллем и тем, чтобы вообще не вводить прямые выборы (возможность B), билль бы прошёл.

Следствие парадокса Кондорсе: поправка Сазерленда

- Однако сенатор Сазерленд, лидер меньшинства, которое было против выборов сенаторов вообще, внёс поправку-убийцу C: предложение о прямых выборах сенаторов без каких-либо поправок про южные штаты.
- Сазерленд сначала запустил голосование между A и C.
- Меньшинство Сазерленда проголосовало за C, северяне-республиканцы тоже проголосовали за C, и C победило A.
- Затем встал выбор между C и B, Сазерленд внезапно «изменил свою точку зрения» и стал голосовать не за C, а за B, то есть против выборов совсем.

Следствие парадокса Кондорсе: поправка Сазерленда

- В результате билль C сначала выполнил свою функцию и выбил поддерживаемый большинством билль A, а затем не прошёл на следующих выборах.



Следствие парадокса Кондорсе: поправка Сазерленда

- Этот же пример демонстрирует, что правдивости при таких выборах тоже лучше не ждать: меньшинство, стоявшее против выборов вообще, здесь было вынуждено сначала голосовать за них, чтобы затем иметь возможность провалить этот исход на следующих выборах.
- И он же показывает, что попарная независимость тоже недоступна: ведь по этому свойству выбор между A и B не должен зависеть от наличия или отсутствия третьей альтернативы C.

Выборы президента РФ

- Давайте рассмотрим теперь совершенно реальную систему, по которой выбирают президента РФ.
- В первом туре участвуют все кандидаты, и если никто не набирает больше 50%, то двое лидеров выходят во второй тур.
- Предположим, что у нас есть три кандидата на высокий пост и 27 избирателей.

Выборы президента РФ

- Распределение предпочтений:

К-во избирателей	**6**	**6**	**6**	**4**	**2**	**3**
Барсуков	1	2	3	2	3	1
Гризлев	2	3	1	1	2	3
Углеводский	3	1	2	3	1	2

- В первом туре Барсуков наберёт 9 голосов, Углеводский — 8, а Гризлев — 10.
- Однако во втором туре ситуация изменится, и победит Барсуков, набрав 15 голосов против 12 у Гризлева. Пока всё нормально.

Выборы президента РФ

- Предположим, что Барсуков, пытаясь победить Гризлева в первом туре, сумел воздействовать на сердца некоторых избирателей:
  - трое из четырёх с распределением $2 \succ 1 \succ 3$ переместили Барсукова на первое место;
  - двое с распределением $3 \succ 2 \succ 1$ изменили его на $2 \succ 3 \succ 1$.
- Итого получается:

К-во избирателей	**9**	**8**	**6**	**1**	**3**
Барсуков	1	2	3	2	1
Гризлев	2	3	1	1	3
Углеводский	3	1	2	3	2

Выборы президента РФ

- В первом туре Барсуков действительно выигрывает с бóльшим отрывом, получив 12 голосов.
- Однако во второй тур теперь выходит не Гризлев, а Углеводский, который в итоге побеждает Барсукова с счётом $14:13$.
- Барсуков сделал распределение строго лучше для себя, но в итоге сменил победу на поражение. И всё это во вполне естественной системе голосования, по которой действительно выбирают президента РФ...
- Что же делать? Можно ли что-то сделать?

Outline

Парадокс Кондорсе

- Вспомним парадокс Кондорсе: пусть у нас три участника, у них есть свои предпочтения на трёх исходах, и мы хотим решить дело голосованием.
- Предпочтения таковы:

$$x \succ_1 y \succ_1 z,$$
$$z \succ_2 x \succ_2 y,$$
$$y \succ_3 z \succ_3 x.$$

- Получится, что нарушилась транзитивность.

Рациональность

Определение

Профиль предпочтений $\succeq$ называется рациональным, если он является линейным порядком, то есть любые два исхода сравнимы и выполняется условие транзитивности: для всяких $x, y, z \in \mathcal{O}$, если $x \succeq y$ и $y \succeq z$, то $x \succeq z$.

Профили предпочтений

- Мы будем предполагать, что профили предпочтений бывают всякие.
- Например, всякие рациональные — их множество мы обозначим через $\mathcal{R}$.
- Или вообще всякие профили, лишь бы любые два исхода были различимы: множество таких исходов мы обозначим через $\mathcal{P}$.
- Если агентов N, то, значит, множество всевозможных предпочтений будет в этих обозначениях $\mathcal{R}^N$ или $\mathcal{P}^N$.

Функции социального выбора

- Функция социального выбора — некоторая функция f с областью определения $\mathcal{R}^N$ или $\mathcal{P}^N$ и областью значений $\mathcal{O}$, которая по данным предпочтениям агентов выбирает исход.
- Чуть обобщим это определение и будем считать, что функция социального выбора выдаёт не один исход, а слабый линейный порядок на имеющихся исходах $\succeq_{f(\succeq_1,\ldots,\succeq_N)}$.

Эффективность по Парето

Определение

Пусть пара исходов $x, y \in \mathcal{O}$ *такова, что для каждого агента* i *исход* x *не хуже, и при этом для какого-нибудь агента он строго лучше: для всех* i $x \succeq_i y$, *и существует такое* j, *что* $x \succ_j y$. *Тогда функция социального выбора* f *называется эффективной по Парето, если для каждой такой пары исходов результат функции социального выбора* $f(\succeq_1, \ldots, \succeq_N)$ *ставит* x *перед* y: $x \succeq_f y$.

Диктаторские функции

Определение

Функция социального выбора f называется диктаторской, если существует такой агент h, что для любых $x, y \in \mathcal{O}$ и любого профиля $(\succeq_1, \ldots, \succeq_N)$ $x \succeq_f y$ тогда и только тогда, когда $x \succ_h y$.

- Диктаторская функция социального выбора делает точно такой же выбор, как один из представленных агентов.
- У диктаторской функции получится соответствовать нужным свойствам, точно так же, как у предпочтений одного агента это получается.
- А беда в том, что ничего другого-то и не получится.

Формулировка

Теорема (Эрроу)

Пусть множество возможных исходов $\mathcal{O}$ состоит из не менее чем трёх элементов, и возможны все рациональные профили ($\mathcal{R}$) или все профили, в которых любые две альтернативы различимы ($\mathcal{P}$). Тогда всякая функция социального выбора f, которая оптимальна по Парето и удовлетворяет условию попарной независимости, является диктаторской.

Определяющие наборы агентов

- Для данного F будем говорить, что набор агентов $S \subset [N]$:
  - *определяющий для x перед y*, если когда каждый агент в S предпочитает $x \succ y$ *и* каждый агент в $[N] \setminus S$ предпочитает $y \succ x$, F выбирает x;
  - *определяющим*, если он определяющий для любой пары $\{x, y\}$;
  - *полностью определяющим*, если когда каждый агент из S предпочитает $x \succ y$, F тоже выбирает x.

Доказательство

1 Если для некоторых x и y набор $S \subset [N]$ является определяющим для x перед y, то $\forall z \neq x$ набор S является определяющим для x перед z и $\forall z \neq y$ набор S является определяющим для z перед y.

- Если $z = y$, доказывать нечего.
- Если $z \neq y$, рассмотрим такой профиль $(\succeq_1, \ldots, \succeq_N)$, что

$$
\begin{aligned}
x \succ_i y \succ_i z \quad & \forall i \in S, \\
y \succ_i z \succ_i x \quad & \forall i \in [N] \setminus S.
\end{aligned}
$$

Доказательство

1. Если для некоторых x и y набор $S \subset [N]$ является определяющим для x перед y, то $\forall z \neq x$ набор S является определяющим для x перед z и $\forall z \neq y$ набор S является определяющим для z перед y.

- Тогда, значит, по свойству определяющего набора f должна предпочесть x перед y.
- А по оптимальности по Парето, F предпочитает y перед z.
- Значит, F предпочитает x перед z.
- Осталось сослаться на попарную независимость.

Доказательство

2 Если для некоторых x и y набор $S \subset [N]$ является определяющим для x перед y, и z — третья альтернатива, то набор S является определяющим для z перед w и для w перед z для всех $w \neq z \in \mathcal{O}$.

- По шагу 1, S определяющий для z перед y и для x перед z.
- Применим снова шаг 1 для пары $\{x, z\}$ и альтернативы w; значит, S определяющий для w перед z.
- Аналогично для пары $\{z, y\}$.

Доказательство

3. Если для некоторых $\{x, y\} \subset \mathcal{O}$ S определяющий для x перед y, то S определяющий.

- Доказательство сразу следует из шага 2 и из того, что третья альтернатива существует.

Доказательство

4. Если S определяющий и T определяющий, то $S \cap T$ тоже определяющий.

- Рассмотрим тройку альтернатив $\{x, y, z\} \subset \mathcal{O}$ и профиль $(\succeq_1, \ldots, \succeq_N)$ такой, что

$$\begin{array}{ll} z \succ_i y \succ_i x & \forall i \in S \setminus (S \cap T), \\ x \succ_i z \succ_i y & \forall i \in S \cap T, \\ y \succ_i x \succ_i z & \forall i \in T \setminus (S \cap T), \\ y \succ_i z \succ_i x & \forall i \in [N] \setminus (S \cup T). \end{array}$$

Доказательство

4 Если S определяющий и T определяющий, то $S \cap T$ тоже определяющий.

- Тогда $z \succ_f y$, потому что $S = (S \cap T) \cup (S \setminus (S \cap T))$ — определяющий.
- И $x \succ_f z$, потому что T — определяющий.
- Значит, $x \succ_f y$, и по попарной независимости $S \cap T$ определяющий для x перед y. Значит, он вообще определяющий.

Доказательство

5 Для любого $S \subset [N]$ либо S определяющий, либо его дополнение $[N] \setminus S$ определяющий

- Рассмотрим тройку альтернатив $\{x, y, z\} \subset \mathcal{O}$ и профиль $(\succeq_1, \ldots, \succeq_N)$ такой, что

$$x \succ_i z \succ_i y \quad \forall i \in S,$$
$$y \succ_i x \succ_i z \quad \forall i \in [N] \setminus S.$$

- Тогда либо $x \succ_f y$, и S определяющий для x перед y, либо $y \succ_f x$.
- Если $y \succ_f x$, то по Парето $x \succ_f z$, и, значит, $y \succ_f z$; значит, $y \succ_f z$ определяющий для y перед z.

Доказательство

6. Если S определяющий и $S \subset T$, то T определяющий.

- По Парето пустой набор не может быть определяющим.
- Значит, $[N] \setminus T$ не может быть определяющим, потому что тогда и $\emptyset = S \cap ([N] \setminus T)$ будет определяющим.

Доказательство

7 Если $S \subset [N]$ определяющий, и $|S| > 1$, то есть строгое подмножество $S' \subsetneq S$, тоже являющееся определяющим набором.

- Рассмотрим $h \in S$. Если $S \setminus \{h\}$ определяющий, то всё.
- Если нет, то $[N] \setminus (S \setminus \{h\})$ определяющий, и $\{h\} = S \cap ([N] \setminus (S \setminus \{h\}))$ определяющий.

Доказательство

8. Для некоторого $h \in [N]\ \{h\}$ определяющий.

- Нужно просто итерировать шаг 7.

Доказательство

9 Если $S \subset [N]$ определяющий, то для всех x и y S полностью определяющий для x перед y.

- Нужно получить, что для всех $T \subset [N] \setminus S$ $x \succ_f y$, если все агенты из S предпочитают $x \succ y$, из T — $x \succeq y$, остальные — $y \succ x$.
- Рассмотрим третью альтернативу и профиль $(\succeq_1, \ldots, \succeq_N)$ такой, что

$$\begin{aligned} &x \succ_i z \succ_i y && \forall i \in S, \\ &x \succ_i y \succ_i z && \forall i \in T, \\ &y \succ_i z \succ_i x && \forall i \in [N] \setminus (S \cup T). \end{aligned}$$

Доказательство

9 Если $S \subset [N]$ определяющий, то для всех x и y S полностью определяющий для x перед y.

- Рассмотрим третью альтернативу и профиль $(\succeq_1, \ldots, \succeq_N)$ такой, что

$$\begin{aligned} &x \succ_i z \succ_i y && \forall i \in S, \\ &x \succ_i y \succ_i z && \forall i \in T, \\ &y \succ_i z \succ_i x && \forall i \in [N] \setminus (S \cup T). \end{aligned}$$

- Тогда $x \succ_f z$, потому что $S \cup T$ определяющий, и $z \succ_f y$, потому что S определяющий. Значит, $x \succ_f y$, что и требовалось.

Доказательство

10 Если $\{h\}$ определяющий, то h — диктатор.

- Это в точности следует из определения полностью определяющего набора.

Если в $\mathcal{O}$ два элемента

- Где мы пользовались тем, что $|\mathcal{O}| \geq 3$? Что будет, если $|\mathcal{O}| = 2$?

Если в $\mathcal{O}$ два элемента

- Где мы пользовались тем, что $|\mathcal{O}| \geq 3$? Что будет, если $|\mathcal{O}| = 2$?
- На самом деле, если $|\mathcal{O}| = 2$, то теорема неверна.
- Например, функция социального выбора «большинство голосов» в данном случае и недиктаторская, и правдиво реализуемая в доминантных стратегиях.

Outline

О чём всё это

- Мы уже столько замечательных примеров рассмотрели: и правдивые механизмы получаются, и реализующие социальную функцию в доминантных стратегиях, и вообще всё прекрасно.
- Ну, не могло же оно всё так и продолжаться, где-то должен быть подвох.
- Сейчас мы рассмотрим один из самых больших подвохов этой теории.

Суть теоремы

- Оказывается, что всё-таки не всякие механизмы существуют.
- Сейчас мы сформулируем определение довольно узкого и «нечестного» класса социальных функций — *диктаторских*, т.е. таких, которые выгодны одному конкретному участнику.
- А потом докажем, что никаких других реализовать в доминантных стратегиях нельзя...

Диктаторские функции социального выбора

Определение

Функция социального выбора f называется диктаторской, если существует такой агент i, что для всех $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_N) \in \boldsymbol{\Theta}$

$$f(\boldsymbol{\theta}) = \{x \in X \mid u_i(x, \theta_i) \geq u_i(y, \theta_i) \text{ для всех } y \in X\}.$$

- Проще говоря, функция социального выбора всегда выбирает один из вариантов, оптимальных для i-го агента.

Монотонные функции социального выбора

- Вспомним определение: множество нижнего контура возможного исхода o при агенте i типа θ_i — это

$$L_i(o,\theta_i) = \{o' \in \mathcal{O} : u_i(o,\theta_i) \geq u_i(o',\theta_i)\}.$$

Определение

Функция социального выбора f называется монотонной, если для каждого θ, если θ' таково, что $L_i(f(\theta),\theta_i) \subseteq L_i(f(\theta),\theta_i')$ для всех i, то $f(\theta) = f(\theta')$.

Монотонные функции социального выбора

Определение

Функция социального выбора f называется монотонной, если для каждого θ, если θ' таково, что $L_i(f(\theta), \theta_i) \subseteq L_i(f(\theta), \theta_i')$ для всех i, то $f(\theta) = f(\theta')$.

- То есть если $f(\theta) = x$, и при переходе к θ' ни у одного агента ни один исход, который раньше был хуже x, не стал строго лучше x, то x должен остаться его социальным выбором.

Порядки предпочтений

- Важным для нас понятием будут порядки на возможных исходах $\mathcal{O}$, которые для каждого агента задают, что ему больше нравится.
- Нам не так важно, сколько именно агент получит (u_i), сколько то, что он исход o_1 ценит выше, чем o_2, но ниже, чем o_3.

Порядки предпочтений

- Обозначим через $\mathcal{P}$ множество всех линейных порядков на $\mathcal{O}$.
- Через $\mathcal{R}_i$ — множество порядков, которые может реализовывать агент i.

Теорема Гиббарда–Саттертуэйта

Теорема (Гиббарда-Саттертуэйта)

Предположим, что:

- *множество возможных исходов* $\mathcal{O}$ *конечно и состоит не менее чем из трёх элементов:* $|O| \geq 3$;
- *все исходы реализуются:* $f(\theta) = \mathcal{O}$;
- *каждый агент может реализовывать любое рациональное множество предпочтений:* $\mathcal{R}_i = \mathcal{P}$.

Тогда функция социального выбора f *правдиво реализуема в доминантных стратегиях тогда и только тогда, когда она диктаторская.*

Справа налево

- Очевидно, что диктаторская f правдиво реализуема в доминантных стратегиях (проверьте!).
- Дальше будем доказывать слева направо.

Структура доказательства

- Доказывать будем так:
  1. Если $\mathcal{R}_i = \mathcal{P}$ для всех i, и f правдиво реализуема в доминантных стратегиях, то f монотонна.
  2. Если $\mathcal{R}_i = \mathcal{P}$ для всех i, f монотонна, и $f(\Theta) = \mathcal{O}$, то f эффективна ex post.
  3. Если f монотонна и эффективна ex post, то она диктаторская.
- Это будут наши три леммы.

Доказательство леммы 1

1. Если $\mathcal{R}_i = \mathcal{P}$ для всех i, и f правдиво реализуема в доминантных стратегиях, то f монотонна.

- Рассмотрим два профиля типов θ и θ', для которых $L_i(f(\theta), \theta_i) \subseteq L_i(f(\theta), \theta_1')$. Хотим показать, что $f(\theta) = f(\theta')$.
- Т.к. f правдиво реализуема, то $f(\theta_1', \theta_2, \ldots, \theta_N) \in L_1(f(\theta), \theta_1) \subseteq L_1(f(\theta), \theta_i')$ и $f(\theta) \in L_1(f(\theta_1', \theta_2, \ldots, \theta_N), \theta_1')$.

Доказательство леммы 1

1 Если $\mathcal{R}_i = \mathcal{P}$ для всех i, и f правдиво реализуема в доминантных стратегиях, то f монотонна.

- Т.к. порядки линейные (всё сравнимо), из этого следует, что $f(\theta_1', \theta_2, \ldots, \theta_N) = f(\theta)$.
- Далее, $f(\theta_1', \theta_2', \theta_3, \ldots, \theta_N) = f(\theta_1', \theta_2, \ldots, \theta_N) = f(\theta)$. И т.д., в общем, $f(\theta) = f(\theta')$.

Доказательство леммы 2

2 Если $\mathcal{R}_i = \mathcal{P}$ для всех i, f монотонна, и $f(\theta) = \mathcal{O}$, то f эффективна ex post.

- Напомним, что «эффективна ex post» означает, что уже после того, как агенты сыграют по своим стратегиям, для каждого возможного значения θ нельзя сместить равновесие туда, где всем будет лучше.
- Предположим противное. Пусть существуют такие $\theta \in \mathbf{\Theta}$, $y \in X$ и i, что
$$u_i(y, \theta_i) > u_i(f(\theta), \theta_i)$$
(не равно, т.к. нет несравнимых исходов).

Доказательство леммы 2

2 Если $\mathcal{R}_i = \mathcal{P}$ для всех i, f монотонна, и $f(\theta) = \mathcal{O}$, то f эффективна ex post.

$$u_i(y, \theta_i) > u_i(f(\theta), \theta_i)$$

- Воспользуемся теперь тем, что $f(\theta) = \mathcal{O}$. Это значит, что есть такой $\theta' \in \boldsymbol{\Theta}$, что $f(\theta') = y$.
- А теперь воспользуемся тем, что все предпочтения в $\mathcal{P}$ возможны. Выберем такой вектор $\theta'' \in \boldsymbol{\Theta}$, что

$$\forall i \ \forall x \neq f(\theta), y \quad u_i(y, \theta_i'') > u_i(f(\theta), \theta_i'') > u_i(z, \theta_i'').$$

Доказательство леммы 2

2 Если $\mathcal{R}_i = \mathcal{P}$ для всех i, f монотонна, и $f(\theta) = \mathcal{O}$, то f эффективна ex post.

$$\forall i \ \forall x \neq f(\theta), y \quad u_i(y, \theta_i'') > u_i(f(\theta), \theta_i'') > u_i(z, \theta_i'').$$

- Поскольку $L_i(y, \theta_i') \subset L_i(y, \theta_i'')$ для всех i, по монотонности $f(\theta'') = f(\theta)$. Противоречие, т.к. $y \neq f(\theta)$.

Доказательство леммы 3

3 Если f монотонна и эффективна ex post, то она диктаторская.

- Эта лемма следует из теоремы Эрроу о невозможности.

Упражнение. Доказать, что лемма 3 из неё действительно вытекает.

Спасибо за внимание!

- Lecture notes и слайды будут появляться на моей homepage:
  http://logic.pdmi.ras.ru/~sergey/index.php?page=teaching
- Присылайте любые замечания, решения упражнений, новые численные примеры и прочее по адресам: sergey@logic.pdmi.ras.ru, snikolenko@gmail.com
- Заходите в ЖЖ smartnik.