

Предпочтения. Функция общественного выбора.

Предпочтения

- Агент принимает решения, сравнивая разные альтернативы. Если альтернатива x для него *не хуже*, чем альтернатива y , пишется

$$x \succeq y.$$

- *Безразличие*:

$$x \sim y, \quad \text{то есть } x \succeq y \text{ и } y \succeq x.$$

- *Строгое предпочтение*:

$$x \succ y, \quad \text{то есть } x \succeq y, \text{ но не } y \succeq x.$$

- Предпочтения *рациональны*, если они

- *полны*, то есть для любых x, y либо $x \succeq y$, либо $y \succeq x$
- *транзитивны*, то есть для любых x, y, z из $x \succeq y$ и $y \succeq z$ следует $x \succeq z$

Аксиоматика общественного выбора

- Есть N агентов и множество альтернатив A
- У каждого агента i есть рациональные предпочтения \succeq_i на A
- R - множество всех рациональных предпочтений на A
- P - множество всех рациональных предпочтений на A , строго различающих любые альтернативы
- *Профиль предпочтений* - $(\succeq_1, \dots, \succeq_N) \in R^N$
- *Функция общественных предпочтений* определена на каком-то подмножестве S множества R^N :

$$F : S \rightarrow R$$

$$(\succeq_1, \dots, \succeq_N) \mapsto F(\succeq_1, \dots, \succeq_N)$$

- Строгие предпочтения, соответствующие предпочтениям F обозначаются F_p
- **Функция общественных предпочтений удовлетворяет требованиям**
- *единогласия*, если для любой пары альтернатив $a, b \in A$ и любого профиля предпочтений $(\succeq_1, \dots, \succeq_N)$, из того, что $a \succ_i b$ для всех i следует $a F_p(\succeq_1, \dots, \succeq_N) b$.

- *независимости от посторонних альтернатив*, если общественные предпочтения относительно любой пары альтернатив a, b зависят только от индивидуальных предпочтений относительно этих альтернатив

- для любой пары альтернатив a, b и любой пары профилей (z_1, \dots, z_N) и (z'_1, \dots, z'_N) , для которых для любого i выполняется

$$a \succeq_i b \Leftrightarrow a \succeq'_i b \quad \text{и} \quad b \succeq_i a \Leftrightarrow b \succeq'_i a,$$

выполняется также

$$aF(z_1, \dots, z_N) b \Leftrightarrow aF(z'_1, \dots, z'_N) b,$$

$$bF(z_1, \dots, z_N) a \Leftrightarrow bF(z'_1, \dots, z'_N) a.$$

- *анонимности*, если имена агентов не играют никакой роли, то есть для любой перестановки индексов π , имеем $F(z_1, \dots, z_N) = F(z_{\pi(1)}, \dots, z_{\pi(N)})$

Обыкновенная диктатура

- Есть такой агент i^* , что для любой пары альтернатив $a, b \in A$ и любого профиля предпочтений (z_1, \dots, z_N) , из того, что $a \succ_{i^*} b$ следует $aF_p(z_1, \dots, z_N) b$.
 - Такой агент называется *диктатором*
- Диктатуры нередко встречаются в реальной жизни
 - Куба, КНДР, Туркменистан, Узбекистан, Сирия
- После 2-ой мировой войны они встречались особенно часто
 - СССР, Испания, Аргентина, Египет
 - Заметим, что это вовсе не маргинальные страны, как сейчас

Счёт Борда

- Множество альтернатив A конечно
- Агент i с предпочтениями \succeq_i присваивает каждой альтернативе a очки $c_i(a)$ следующим образом
 - $c_i(a) = n$, если n - порядковый номер альтернативы a в "рейтинге" агента i , заданном \succeq_i
 - если есть другие альтернативы b , которые столь же ценны как альтернатива a для агента i ($b \sim_i a$), то $c_i(a) = c_i(b) =$ среднему рейтингу этих альтернатив
- Для каждого профиля (z_1, \dots, z_N) общественные предпочтения задаются сложением рейтингов: $aF(z_1, \dots, z_N) b \Leftrightarrow \sum_i c_i(a) \leq \sum_i c_i(b)$
- Предпочтения Борда полны, транзитивны и удовлетворяют требованиям единогласия и анонимности (проверьте!), однако НЕ удовлетворяют требованию независимости от

посторонних альтернатив

Счёт Борда: зависимость от альтернатив

- Пусть есть два агента и три альтернативы - $\{a, b, c\}$
- Рассмотрим два профиля предпочтений

$$\begin{aligned} a \succ_1 c \succ_1 b, & \quad b \succ_2 a \succ_2 c \\ a \succ'_1 b \succ'_1 c, & \quad b \succ'_2 c \succ'_2 a \end{aligned}$$

- Для первого профиля общество предпочитает альтернативу a альтернативе b (так как $c_i(a) = 3, c_i(b) = 4$)
- Для второго профиля общество предпочитает альтернативу b альтернативе a (так как $c_i(a) = 4, c_i(b) = 3$)
- Однако относительный порядок a и b не изменился ни для одного агента!
- Требование независимости от посторонних альтернатив нарушается

Голосование большинством

- Общество предпочитает альтернативу a альтернативе b , если большинство предпочитают альтернативу a альтернативе b
- Три важных свойства голосования большинством
 - *анонимность*
 - *нейтральность к альтернативам* - если предпочтения всех участников обратить (то есть определить новые предпочтения $\succeq'_i : a \succeq'_i b \Leftrightarrow b \succeq_i a$), то и предпочтения общества обратятся
 - *отзывчивость* - если общество предпочитало альтернативу a альтернативе b (или безразлично к ним), а потом часть агентов пересмотрело отношение к альтернативе a в (строго) положительную сторону, то теперь общество строго предпочитает альтернативу a альтернативе b

Теорема Мэя

Будем записывать профили предпочтений как $\alpha = (\alpha_1, \dots, \alpha_N)$, где $\alpha_i \in \{-1, 0, 1\}$ для любого i ; соответственно, общественные предпочтения записываются как $F(\alpha_1, \dots, \alpha_N) \in \{-1, 0, 1\}$. Всего есть две альтернативы a и b .

- если $\alpha_i = -1$, то $\alpha \alpha_i b$ означает $a \prec_i b$
- если $\alpha_i = 0$, то $\alpha \alpha_i b$ означает $a \sim_i b$
- если $\alpha_i = 1$, то $\alpha \alpha_i b$ означает $a \succ_i b$

Теорема Мэя. *Общественные предпочтения $F(\alpha_1, \dots, \alpha_N)$ являются голосованием*

большинством тогда и только тогда, когда они удовлетворяют требованиям анонимности, отзывчивости и нейтральности к альтернативам.

- Осталось доказать только "тогда" (достаточность)

Доказательство теоремы Мэя

- Введем следующие обозначения:

- $n^+(\alpha_1, \dots, \alpha_N)$ - число агентов, предпочитающих a в профиле $(\alpha_1, \dots, \alpha_N)$
- $n^-(\alpha_1, \dots, \alpha_N)$ - число агентов, предпочитающих b в профиле $(\alpha_1, \dots, \alpha_N)$

Анонимность позволяет сделать вывод, что

$$F(\alpha_1, \dots, \alpha_N) = G(n^+(\alpha_1, \dots, \alpha_N), n^-(\alpha_1, \dots, \alpha_N))$$

- Если $n^+(\alpha_1, \dots, \alpha_N) = n^-(\alpha_1, \dots, \alpha_N)$, то $F(\alpha_1, \dots, \alpha_N) = 0$

$$\begin{aligned} F(\alpha_1, \dots, \alpha_N) &= G(n^+(\alpha_1, \dots, \alpha_N), n^-(\alpha_1, \dots, \alpha_N)) \\ &= G(n^+(-\alpha_1, \dots, -\alpha_N), n^-(-\alpha_1, \dots, -\alpha_N)) \\ &= F(-\alpha_1, \dots, -\alpha_N) \\ &= -F(\alpha_1, \dots, \alpha_N) \end{aligned}$$

(последнее равенство следует из нейтральности к альтернативам)

- Если $n^+(\alpha_1, \dots, \alpha_N) > n^-(\alpha_1, \dots, \alpha_N)$, то $F(\alpha_1, \dots, \alpha_N) = 1$

- Можно считать, что $\alpha_i = 1$ при $i \leq n^+$, $\alpha_i \leq 0$ при $i > n^+$

- Новый профиль $(\alpha'_1, \dots, \alpha'_N)$:

$$* \alpha'_i = \alpha_i = 1 \text{ при } i \leq n^- < n^+$$

$$\alpha'_i = 0 \quad \text{при } n^- < i \leq n^+$$

$$\alpha'_i = \alpha_i \leq 0 \text{ при } i > n^+$$

- Теперь

$$n^+(\alpha'_1, \dots, \alpha'_N) = n^-(\alpha_1, \dots, \alpha_N) = n^-(\alpha'_1, \dots, \alpha'_N),$$

и значит $F(\alpha'_1, \dots, \alpha'_N) = 0$. Можно воспользоваться отзывчивостью F

- Если $n^+(\alpha_1, \dots, \alpha_N) < n^-(\alpha_1, \dots, \alpha_N)$, то $F(\alpha_1, \dots, \alpha_N) = 0$

$$n^+(-\alpha_1, \dots, -\alpha_N) > n^-(-\alpha_1, \dots, -\alpha_N),$$

и, значит, $F(-\alpha_1, \dots, -\alpha_N) = 1$. Нейтральность относительно альтернатив дает

$$F(\alpha_1, \dots, \alpha_N) = -F(-\alpha_1, \dots, -\alpha_N) = -1. \blacksquare$$

Парадокс Кондорсе

- Голосование большинством - очень удобная вещь, когда есть всего две альтернативы.

Однако уже в случае трех альтернатив возникают проблемы

- **Парадокс Кондорсе:** *Голосование большинством может привести к нетранзитивному общественному выбору даже если все члены общества рациональны (и, следовательно, их предпочтения транзитивны)*

- Три агента и три альтернативы $\{a, b, c\}$
- Предпочтения:

$$a \succ_1 b \succ_1 c \quad c \succ_2 a \succ_2 b \quad b \succ_3 c \succ_3 a$$

- Голосование большинством дает $a \succ b \succ c \succ a$, то есть транзитивность нарушается

Теорема Эрроу

Конституцией называется отображение, сопоставляющее каждому возможному профилю рациональных индивидуальных предпочтений рациональное предпочтение (общественное предпочтение).

Предполагается, что число имеющихся альтернатив не меньше 3.

Если конституция (общественные предпочтения) удовлетворяет следующим требованиям:

- (i) *независимости от посторонних альтернатив,*
- (ii) *единогласия,*

то она является диктатурой.

Доказательство: общая схема

(Формулировка и первое доказательство: Кеннет Эрроу, 1951; это доказательство: Джон Геанокпос, 1996)

Возьмем какую-нибудь альтернативу b

Утверждение 1. *Если в каком-то профиле предпочтений каждый агент ставит b наверх или вниз своих предпочтений, то функция социального выбора тоже ставит b либо в самый верх, либо в самый низ социальных предпочтений*

Утверждение 2. *Существует такой агент $n^* = n^*(b)$, что при некотором профиле его голос переносит b с самого низа общественных предпочтений на самый верх*

Утверждение 3. *Агент $n^*(b)$ диктует обществу предпочтения относительно любой пары альтернатив a, c , не содержащей b*

Утверждение 4. *Агент $n^*(b)$ диктует обществу предпочтения относительно любой пары альтернатив a, b*

Доказательство теоремы Эрроу

\succeq будет обозначать предпочтения общества. Возьмем какую-нибудь альтернативу b

Утверждение 1. Если в каком-то профиле предпочтений каждый агент ставит b наверх или вниз своих предпочтений, то функция социального выбора тоже ставит b либо в самый верх, либо в самый низ социальных предпочтений

- От противного: пусть есть такие (различные) a и c , что $a \succeq b$ и $b \succeq c$
- Независимость от посторонних альтернатив позволяет переставить c выше a в предпочтениях каждого агента, при этом не изменяя порядок в парах a, b и b, c (потому что b стоит только в самом верху или самом низу)
- Единогласие дает $c \succeq a$
- Транзитивность дает $a \succeq c$, противоречие

Утверждение 2. Существует такой агент $n^* = n^*(b)$, что при некотором профиле его голос переносит b с самого низа общественных предпочтений на самый верх

- Рассмотрим какой-нибудь профиль предпочтений, в котором b стоит в самом низу предпочтений всех агентов
- Соответствующее общественное предпочтение тоже ставит альтернативу b в самый низ
- Пусть все агенты, с 1-го по N -ый, по очереди переставляют b на самый верх
- В какой-то момент b перескочит одним прыжком на самый верх (в худшем случае это произойдет на N -ом агенте, когда сработает единогласие)
- Назовем агента, при перемещении альтернативы b которым она перешла с самого низа на самый верх общественных предпочтений $n^* = n^*(b)$
- Назовем профилем α профиль предпочтений, сложившийся перед тем, как n^* переставил альтернативу b
- Назовем профилем β профиль предпочтений, сложившийся после того, как n^* переставил альтернативу b
- Общественное предпочтение, соответствующее профилю α , ставит b в самый низ, а соответствующее профилю β - на самый верх

Утверждение 3. Агент $n^*(b)$ диктует обществу предпочтения относительно любой пары альтернатив a, c , не содержащей b

- Мы хотим показать, что из $a \succ_{n^*} c$ следует $a \succ c$
- Создадим профиль γ из профиля β ,
 - переставив a выше b в предпочтениях агента n^*
 - позволив остальным агентам ($i \neq n^*$) расставить a и c произвольно относительно друг друга, не меняя при этом позиции b

- Из независимости от посторонних альтернатив следует, что
 - общественное предпочтение, соответствующее профилю γ , даст $a \succ b$ (так как все индивидуальные предпочтения относительно пары a, b такие же, как в профиле α)
 - общественное предпочтение, соответствующее профилю γ , даст $b \succ c$ (так как все индивидуальные предпочтения относительно пары b, c такие же, как в профиле β)
- Из транзитивности общественного предпочтения следует, что $a \succ c$
- Поскольку общественное предпочтение не зависит от посторонних альтернатив, мы показали, что из $a \succ_{n^*} c$ следует $a \succ c$

Утверждение 4. *Агент $n^*(b)$ диктует обществу предпочтения относительно любой пары альтернатив a, b*

- Возьмем какую-нибудь третью альтернативу c
- Повторяя рассуждения Утверждений 1 и 2 для этого c , получим, что существует какой-то агент $n^*(c)$, который, в частности, является диктатором для пары a, b (по Утверждению 3).
- Но наш первый диктатор $n^*(b)$ мог влиять на общественные предпочтения относительно пары a, b ! (Он же по определению переставляет b с низа общественного предпочтения на верх).
- Значит, $n^*(c)$ и есть $n^*(b)$, настоящий диктатор. ■

Монотонные предпочтения

- Функция (соответствие) группового выбора F *монотонна*, если в случае, когда в профиле предпочтений (z'_1, \dots, z'_N) отношение к альтернативе a не ухудшается по сравнению с профилем (z_1, \dots, z_N) у каждого агента i , из $a \in F(z_1, \dots, z_N)$ следует $a \in F(z'_1, \dots, z'_N)$
- **Пример 1** Зафиксируем i . Положим

$$F(z_1, \dots, z_N) = \{a \in A \mid a \succeq_i b \text{ для всех } b \in A\}$$

- **Пример 2**

$$F^{Par}(z_1, \dots, z_N) = \{a \in A \mid \text{нет такого } b, \text{ что } b \succ_i a \text{ для всех } i\}$$

- **Пример 3** Зафиксируем $a^* \in A$ ("статус-кво"). Положим

$$F^{a^*}(z_1, \dots, z_N) = \{a \in A \mid a \succeq_i a^* \text{ для всех } i\}$$

- **Пример 4** (соответствие Маскина) Снова зафиксируем $a^* \in A$. Положим

$$F(z_1, \dots, z_N) = F^{a^*}(z_1, \dots, z_N) \cap F^{Par}(z_1, \dots, z_N)$$

- Упражнение: *Пересечение или объединение монотонных соответствий монотонно.*

Теорема Мюллера - Сатеруайта

Теорема (Мюллер-Сатеруайт) Пусть функция общественного выбора $f: P \rightarrow A$ монотонна, каждая альтернатива $a \in A$ является значением функции f (для какого-то профиля предпочтений), и число альтернатив во множестве A не меньше трёх. Тогда f - диктаторская.

- Еще один результат о невозможности
- Доказательство более близко к исходному доказательству теоремы Эрроу
- Существенные ограничения
 - число альтернатив больше 2
 - область определения функции общественного выбора - все рациональные предпочтения

Доказательство

Утверждение 1 Функция f эффективна (то есть результат выбора не бывает Парето-доминируемым).

- Теперь возьмем альтернативу a и коалицию (подмножество множества агентов) K .
- Обозначим множество коалиций, которые форсируют a через $W(a)$. Это множество обладает следующими свойствами:
- $W(a)$ корректно определено
- I (множество всех агентов) принадлежит $W(a)$
- Если $K \in W(a)$ и $K \subset K'$, то $K' \in W(a)$
- Для двух различных альтернатив a, b , $K \in W(a)$ тогда и только тогда, когда $\neg K \in W(b)$
- Для всех a множество $W(a)$ на самом деле одно и то же (обозначим его W)

Утверждение 2 Если множество всех агентов I разбить на три непересекающиеся коалиции, то одна из них принадлежит W

- Рассмотрим какую-нибудь минимальную коалицию C в W и агента $i \in C$
- Получается, то ни одна из коалиций $C \setminus \{i\}$, $\{i\}$, $\neg C$ не могут принадлежать W , что противоречит Утверждению 2

Неманипулируемые функции

- **Утверждение** Функция общественного выбора неманипулируема тогда и только тогда, когда она монотонна
 - \Rightarrow : В профиле (z_i) альтернатива a стоит у каждого участника не выше, чем в профиле (z'_i) и $f(z_i) = a$. Можно считать, что (z_i) и (z'_i) различаются только у одного агента k

- $a \succ_k f(z'_i)$ и $f(z'_i) \succ'_k a$
- Из $a \succ_k f(z'_i)$ следует, что $a \succ'_k f(z'_i)$ (так как a стало выше)
- \Leftarrow : Надо проверить, что $f(z_k, \cdot) \succ_k f(z'_k, \cdot)$
- Допустим, $f(z_k, \cdot) \neq f(z'_k, \cdot)$ и $f(z_k, \cdot) \prec_k f(z'_k, \cdot)$
- Создадим $z_k^* = [b \succ a \succ \dots]$.
- Поскольку (z_k^*, \cdot) ставит a не ниже, чем (z_k, \cdot) , $f(z_k^*, \cdot) = a$
- Поскольку (z_k^*, \cdot) ставит b не ниже, чем (z'_k, \cdot) , $f(z_k^*, \cdot) = b$

Неманипулируемые механизмы

- **Теорема (Гиббард, 1973)** Пусть функция общественного выбора $f: \mathbf{P} \rightarrow A$ неманипулируема, каждая альтернатива $a \in A$ является значением функции f (для какого-то профиля предпочтений), и число альтернатив во множестве A не меньше трёх. Тогда f - диктаторская
- **Теорема (Гиббард - Сатеруайт)** Пусть множество предпочтений каждого агента на множестве A содержат все возможные предпочтения, каждая альтернатива $a \in A$ является значением функции f (для какого-то профиля предпочтений), а число альтернатив во множестве A не меньше трёх. Тогда функция общественного выбора f правдиво реализуется в доминирующих стратегиях тогда и только тогда, когда f - диктаторская.

Однопиковые предпочтения

- Бинарное отношение \geq называется *линейным порядком* на множестве A , если \geq
 - *полно*, то есть для любых различных a, b или $a \geq b$ и $b \geq a$
 - *рефлексивно*, то есть $a \geq a$ для любого a
 - *транзитивно*, то есть для любых a, b, c из $a \geq b$ и $b \geq c$ следует $a \geq c$
- Рациональные предпочтения \succ являются *однопиковыми* (по отношению к данному линейному порядку на множестве альтернатив A), если существует такая альтернатива $a^* \in A$, что
 - \succ возрастают на множестве $\{a | a \leq a^*\}$
 - \succ убывают на множестве $\{a | a \geq a^*\}$
- Пример A - отрезок чисел $[\alpha, \beta]$, \geq - обычный порядок
- Непрерывные предпочтения \succ однопиковы тогда и только тогда, когда они строго выпуклы (то есть для любого a и любых различных $b, c \succ a$ выполняется $\theta b + (1 - \theta)c \succ a$).
- Например, предпочтения, заданные строго квази-вогнутой функцией полезности,

являются однопиковыми

Медиана

- У всех агентов однопиковые предпочтения
- Для каждого профиля предпочтений (z_1, \dots, z_N) у каждого агента i есть точка счастья (наилучшими альтернативами) a_i
- Агент m , для которого число агентов с наилучшими альтернативами, превышающими (относительно линейного порядка на множестве A) a_m равно числу агентов с точками счастья меньше a_m , называется *медианой*
- Формально, агент $m \in I$ называется *медианой* для профиля (z_1, \dots, z_N) , если

$$\#\{i \in I | a_i \geq a_m\} \geq \frac{I}{2} \quad \text{и} \quad \#\{i \in I | a_i \leq a_m\} \geq \frac{I}{2}$$

- То есть медиан может быть несколько!

- Голосование большинством по каждой паре альтернатив: составим предпочтения общественного выбора \hat{F}

$$a \hat{F}(\zeta_1, \dots, \zeta_N) b \Leftrightarrow \#\{i \in I | a \succ_i b\} \geq \#\{i \in I | b \succ_i a\}$$

- Альтернатива, не уступающая никакой другой при попарном голосовании, называется победителем по Кондорсе.

Победитель по Кондорсе

- Утверждение Пусть \geq – линейный порядок на множестве альтернатив

$A, (z_1, \dots, z_N)$ – профиль однопиковых предпочтений и m – какая-то медиана. Тогда оптимальная альтернатива медианы, a_m не уступит никакой другой альтернативе при голосовании большинством, т.е. $a_m \hat{F}(\zeta_1, \dots, \zeta_N) b$ для любой альтернативы $b \in A$.

- Пусть $a_m > b$ (другой случай аналогичен).
- S – множество агентов с $a_i \geq a_m > b$. Для них $a_m \succ_i b$.
- $\#S \geq \frac{I}{2}$, так как m – медиана
- $\#\{i \in I | a_m \succ_i b\} \geq \#S \geq \frac{I}{2} \geq \#\bar{S} \geq \#\{i \in I | b \succ_i a_m\}$

- Из утверждения следует, что победитель по Кондорсе существует всегда, если предпочтения агентов однопиковы
- Утверждение Пусть \geq – линейный порядок на множестве альтернатив A , а число агентов I нечетно. Тогда голосование большинством по каждой паре альтернатив задает рациональные (полные и транзитивные) предпочтения общественного выбора

на множестве профилей однопиковых предпочтений, для которых не бывает безразличных альтернатив.

Другие неманипулируемые механизмы

- Что будет, если в качестве общественного выбора брать точку счастья не медианного агента, а, скажем, третьего снизу?
- **Пример: Левый диктатор** Для каждого профиля предпочтений (z_1, \dots, z_N) выберем самую левую (самую маленькую) точку счастья a_i
- Эта функция общественного выбора тоже неманипулируема!
- Заметим, что эта функция (как и выбор медианы и правый диктатор) анонимна
- Можно описать все анонимные неманипулируемые механизмы (Мулен, 1980)
- Зафиксируем в множестве A $n + 1$ вспомогательную альтернативу b_0, \dots, b_n и положим

$$f(z_1, \dots, z_N) = \text{медиана}(a_1, \dots, a_n, b_0, \dots, b_n)$$

- Если взять три неманипулируемые функции и определить новую как медиану этих трёх, то получится снова неманипулируемая функция
- В случае однопиковых предпочтений можно описать все неманипулируемые функции общественного выбора): любая такая функция получается применением операции медианы к диктаторским и постоянным функциям