Теорема Робертса

Сергей Николенко

Теория экономических механизмов — ИТМО, весна 2008

Outline

1. Аффинные максимизаторы
   - Постановка задачи
   - Условия монотонности
2. Доказательство теоремы Робертса
   - Идея
   - Само доказательство

Суть задачи

- Мы всё время говорим о том, как бы нам ту или иную функцию социального выбора реализовать.
- То есть как построить механизм, который добивается нужного результата.
- Можно сказать, что в этом одна из вообще основных задач дизайна экономических механизмов.

Суть задачи

- Для функций ценности самого общего вида мы уже говорили о теореме Гиббарда-Саттертуэйта.
- Короче говоря, не получается там вообще ничего.
- Но вполне разумно было бы задать тот же вопрос для квазилинейных предпочтений. И на него, вообще говоря, ответа нет.
- Ответ есть только в частном (то есть, наоборот, общем) случае, который мы сейчас рассмотрим.

Определения и обозначения

- У механизма есть набор исходов $\mathcal{O}$ (их мы будем обозначать через $x, y, z, \ldots$).
- Есть N игроков, у каждого — свой тип $v_i \in V_i \subseteq \mathbb{R}^{|\mathcal{O}|}$, т.е. ценности, которые он может присвоить каждому исходу.
- $\mathbf{V} = V_1 \times V_2 \times \ldots \times V_N$ называется *неограниченным*, если $V_i = \mathbb{R}^{|\mathcal{O}|}$ для каждого i.
- Наш случай — неограниченное $\mathbf{V}$, т.е. ситуация, в которой каждый набор из $|\mathcal{O}|$ чисел представляет собой возможный тип игрока i.

Определения и обозначения

- Функция социального выбора $f : \mathbf{V} \to \mathcal{O}$. Можно без потери общности предположить, что f сюръективна.
- Механизм берёт с игроков платежи $p_i : \mathbf{V} \to \mathbb{R}$.
- Игроки квазилинейны, т.е. они хотят максимизировать себе

$$u_i = v_i(f(\mathbf{v})) - p_i(\mathbf{v}).$$

Определения и обозначения

- f правдиво реализуема, если существуют функции платежа p_i, для которых правдивость будет доминантной стратегией.
- Формально, для каждого i, каждого $\mathbf{v}_{-i} \in \mathbf{V}_{-i}$ и $v' \in V_i$

$$v_i(f(\mathbf{v})) - p_i(\mathbf{v}) \geq v_i(f(v_i', \mathbf{v}_{-i})) - p_i(v_i', \mathbf{v}_{-i}).$$

Определения и обозначения

- Пару лекций назад мы уже говорили, что все функции социального выбора, оптимизирующие всеобщее счастье, правдиво реализуемы VCG-платежами.
- Более того, так же реализуемы и взвешенные функции социального выбора.
- Задача в том, чтобы доказать обратное утверждение.

Теорема Робертса

Теорема

Пусть $|\mathcal{O}| \geq 3$, *и* $\mathbf{V}$ *неограничено. Тогда для каждой реализуемой функции социального выбора* f *существуют неотрицательные веса* $k_1, \ldots, k_N$, *не все равные нулю, и константы* C_x, $x \in \mathcal{O}$, *для которых для всех* $\mathbf{v} \in \mathbf{V}$

$$f(\mathbf{v}) \in \mathrm{argmax}_{x \in \mathcal{O}} \left\{ \sum_{i=1}^{n} k_i v_i(x) + C_x \right\}.$$

Условия монотонности

- Доказательства теоремы Робертса основаны на условиях монотонности, т.е. на том, что для правдивой реализуемости v_i должны удовлетворять этим условиям.
- Мы сейчас парочку сформулируем.

Условия монотонности

- W-MON — слабая монотонность. f удовлетворяет W-MON, если для всех $v_i, v_i', \mathbf{v}_{-i}$ при $f(\mathbf{v}) = x$ и $f(v_i', \mathbf{v}_{-i}) = y$
$$v_i'(y) - v_i(y) \geq v_i'(x) - v_i(x).$$
- Иначе говоря, если игрок i может изменить исход с x на y, изменив свою ставку, то *разность* его значений для y должна быть не меньше, чем разность для x.
- Разности можно (и нужно) использовать из-за квазилинейности.

Условия монотонности

- Лемма: Всякая доминантно реализуемая функция социального выбора f удовлетворяет W-MON.
- Во-первых, докажем (кстати, докажем!), что p_i не зависит от v_i.
- Пусть зависит, т.е. есть $\mathbf{v}_{-i}$, x, v_i, v_i' такие, что

$$f(v_i, \mathbf{v}_{-i}) = f(v_i', \mathbf{v}_{-i}) = x, \quad p_i(v_i, \mathbf{v}_{-i}) < p_i(v_i', \mathbf{v}_{-i}).$$

- Но тогда при типах $\mathbf{v}$ игроку i выгодно врать.

Условия монотонности

- Зафиксируем $\mathbf{v}_{-i}, v_i, v_i', x, y$ так, как в определении $W - MON$. Должно быть верно, что

$$v_i(x) - p_i(x, \mathbf{v}_{-i}) \geq v_i(y) - p_i(y, \mathbf{v}_{-i}),$$

  иначе при типе v_i игрок i может улучшить себе доход, соврав v_i'.
- Аналогично, $v_i'(y) - p_i(y, \mathbf{v}_{-i}) \geq v_i'(x) - p_i(x, \mathbf{v}_{-i})$.
- Из этих двух неравенств и получается искомое условие W-MON.

Условия монотонности

- Таким образом, W-MON необходимо.
- Как ни странно, оно для многих ситуаций достаточно.
- Но мы будем пользоваться не им, а другими похожими условиями.

Условия монотонности

- PAD — Positive Association of Differences. f удовлетворяет PAD, если для всех $v, v' \in V$ верно следующее: если $f(v) = x$, и $v_i'(x) - v_i(x) > v_i'(y) - v_i(y)$ для всех $y \in \mathcal{O} \setminus x$ и всех i, то $f(v')$ тоже равно x.
- PAD тоже рассматривает разности, и оно легко следует из W-MON.

Условия монотонности

- Лемма: Всякая доминантно реализуемая функция социального выбора f удовлетворяет PAD.
- Она удовлетворяет W-MON. Зафиксируем v, v' из определения PAD. Введём промежуточные векторы типов

$$v^i = (v'_1, \ldots, v'_i, v_{i+1}, \ldots, v_N).$$

- Тогда $f(v^0) = x$, $v^0 = v$, $v^N = v'$.

Условия монотонности

- Пусть $f(v^{i-1}) = x$, а $f(v^i) = y \neq x$.
- По W-MON, $v_i'(y) - v_i(y) \geq v_i'(x) - v_i(x)$, что противоречит предположению PAD.
- Значит, $f(v^i) = x$. Доказали по индукции.

Условия монотонности

- И в ещё одной форме мы будем пользоваться PAD.
- Обозначим векторы $\alpha, \beta \in \mathbb{R}^N$, и $\alpha > \beta$ значит строгое неравенство в каждой компоненте.
- Обозначим $\vec{0}$ нулевой вектор.

Условия монотонности

- Лемма: пусть f удовлетворяет PAD. Зафиксируем $\mathbf{v}, \mathbf{v}' \in \mathbf{V}$. Если $f(\mathbf{v}') = x$, и $\mathbf{v}'(y) - \mathbf{v}(y) > \mathbf{v}'(x) - \mathbf{v}(x)$, для некоторого $y \in \mathcal{O}$, то $f(\mathbf{v}) \neq y$.
- Для доказательства зафиксируем $\Delta = \mathbf{v}'(y) - \mathbf{v}(y) + \mathbf{v}(x) - \mathbf{v}'(x) \in \mathbb{R}^N$. Очевидно, $\Delta > \vec{0}$.
- Кроме того, $v_i(x) - v_i'(x) - \frac{\Delta_i}{2} = v_i(y) - v_i'(y) + \frac{\Delta_i}{2} > v_i(y) - v_i'(y)$.

Условия монотонности

- Определим новый тип $\mathbf{v}'' \in \mathbf{V}$:

$$v_i''(z) = \begin{cases} \min\{v_i(z), v_i'(z) + v_i(x) - v_i'(x)\} - \Delta_i, & z \neq x, y, \\ v_i(x) - \frac{\Delta_i}{2}, & z = x, \\ v_i(y), & z = y. \end{cases}$$

- Тогда $v_i''(y) - v_i(y) = 0 > v_i''(z) - v_i(z)$, и из PAD следует, что $f(\mathbf{v}'') = y$.

Условия монотонности

- С другой стороны, для $z \neq x, y$
  $v_i''(z) \leq v_i'(z) + v_i(x) - v_i'(x) - \Delta_i$, и
  $v_i''(x) - v_i'(x) = v_i(x) - v_i'(x) - \frac{\Delta_i}{2} > v_i''(z) - v_i'(z)$.
- Проверьте то же самое для $z = y$.
- Тогда из PAD получится, что $f(\mathbf{v}'') = x$. Противоречие.

Outline

1. Аффинные максимизаторы
   - Постановка задачи
   - Условия монотонности
2. Доказательство теоремы Робертса
   - Идея
   - Само доказательство

Идея доказательства

- Чтобы показать, что функция — это аффинный максимизатор, на самом деле надо изучать разности.
- Всё потому, что аффинная максимизация на самом деле эквивалентна системе неравенств:

$$\sum_{i=1}^{N} k_i(v_i(x) - v_i(y)) \geq C_y - C_x,$$

где $f(v) = x \neq y$.

- Мы будем изучать структуру этих самых разностей.

Идея доказательства

- Главное множество, которое мы будем изучать — это

$$P(x,y) = \{\alpha \in \mathbb{R}^N \mid \exists \mathbf{v} : \mathbf{v}(x) - \mathbf{v}(y) = \alpha, f(\mathbf{v}) = x\}.$$

- Т.е. если $f(\mathbf{v}) = x$, то $\mathbf{v}(x) - \mathbf{v}(y) \in P(x,y)$.

Идея доказательства

- В течение доказательства мы увидим, какова структура множеств $P(x, y)$.
- Мы покажем, что $P(x, y)$ — это полупространство. В частности, мы сделаем два важных замечания о структуре $P(x, y)$.
  1. $\alpha \in P(x, y)$ iff $-\alpha \notin P(y, x)$, и внутренности $P(x, y)$ и $P(y, x)$ не пересекаются.
  2. $P(x, y) + P(y, z) = P(x, z)$ (для внутренностей, по крайней мере).

Идея доказательства

- Как их использовать? Предположим, что $\vec{0} \in P(x, y)$ для всех $x, y \in \mathcal{O}$ (это не обязательно так, но предположим).
- Тогда по (2) все $P(x, y)$ равны (и равны, скажем, C).
- По (1), $C \cup -C = \mathbb{R}^N$: если $\alpha \notin C$, то $-\alpha \in C$.
- По (1), C — выпуклое множество: если $\frac{1}{2}(\alpha + \beta) \notin C$, то $-\frac{1}{2}(\alpha + \beta) \in C$, и по (2) $\alpha + \beta \in C$, и, значит, $\alpha + \beta - \frac{1}{2}(\alpha + \beta) \in C$. Противоречие.
- C и $-C$ покрывают всё пространство, выпуклы, и их внутренности не пересекаются. Значит, это подпространства.

Доказательство

- Итак, рассмотрим множество

$$P(x,y)=\{\alpha\in\mathbb{R}^N \mid \exists\mathbf{v}:\mathbf{v}(x)-\mathbf{v}(y)=\alpha, f(\mathbf{v})=x\}.$$

- Во-первых, поскольку f — сюръекция, $P(x,y)\neq\emptyset$ для любых x и y.
- Во-вторых, если $\alpha\in P(x,y)$, то для любого положительного $\delta\in\mathbb{R}^N$ $\alpha+\delta\in P(x,y)$.
- Потому что есть v: $f(v)=x$ и $\mathbf{v}(x)-\mathbf{v}(y)=\alpha$; давайте теперь увеличим $\mathbf{v}(x)$ на δ ($\mathbf{v}$ у нас неограниченные).

Доказательство

Лемма

Для всех $\alpha, \epsilon \in \mathbb{R}^N$, $\epsilon > \vec{0}$*:*

1. $\alpha - \epsilon \in P(x, y) \Rightarrow -\alpha \notin P(y, x)$.
2. $\alpha \notin P(x, y) \Rightarrow -\alpha \in P(y, x)$.

- Это будет свойство (1), которое мы упоминали выше.
- Докажем 1. Пусть $-\alpha \in P(y, x)$. Тогда существует $\mathbf{v}$ такой, что $\mathbf{v}(y) - \mathbf{v}(x) = -\alpha$, и $f(\mathbf{v}) = y$.
- $\alpha - \epsilon \in P(x, y)$, значит, существует $\mathbf{v}'$ такой, что $\mathbf{v}'(x) - \mathbf{v}'(y) = \alpha - \epsilon$, и $f(\mathbf{v}') = x$.

Доказательство

Лемма

Для всех $\alpha, \epsilon \in \mathbb{R}^N$, $\epsilon > \vec{0}$:

1. $\alpha - \epsilon \in P(x, y) \Rightarrow -\alpha \notin P(y, x)$.
2. $\alpha \notin P(x, y) \Rightarrow -\alpha \in P(y, x)$.

- Тогда получается, что $\mathbf{v}(x) - \mathbf{v}(y) = \alpha > \mathbf{v}'(x) - \mathbf{v}'(y)$, что противоречит лемме про PAD. Противоречие.

Упражнение. Докажите утверждение 2 этой леммы. Рассуждения будут примерно такими же.

Доказательство

- Итого мы пока что доказали, что внутренности $P(x, y)$ и $P(y, x)$ не пересекаются, и объединение $P(x, y)$ и $-P(y, x)$ составляет всё пространство.
- А свойство (2) показывало, что граница у $P(x, y)$ монотонно невозрастающая.
- Осталось только показать, что границы являются гиперплоскостями, и тогда мы докажем все нужные свойства $P(x, y)$.

Доказательство

Лемма

Для всех $\alpha, \beta, \epsilon^\alpha, \epsilon^\beta \in \mathbb{R}^N$, $\epsilon^\alpha, \epsilon^\beta > \vec{0}$

$$\alpha - \epsilon^\alpha \in P(x, y) \text{ и } \beta - \epsilon^\beta \in P(y, z) \quad \Rightarrow \quad \alpha + \beta - \frac{\epsilon^\alpha + \epsilon^\beta}{2} \in P(x, z).$$

- Выберем $w \neq x, y, z$, $\delta^w \in P(x, w)$, $\epsilon > \vec{0} \in \mathbb{R}^N$.
- Выберем такой $\mathbf{v}$, что

$$\mathbf{v}(x) - \mathbf{v}(y) = \alpha - \frac{\epsilon^\alpha}{2}, \quad \mathbf{v}(y) - \mathbf{v}(z) = \beta - \frac{\epsilon^\beta}{2},$$

$$\mathbf{v}(x) - \mathbf{v}(w) = \delta^w + \epsilon.$$

Доказательство

Лемма

Для всех $\alpha, \beta, \epsilon^\alpha, \epsilon^\beta \in \mathbb{R}^N$, $\epsilon^\alpha, \epsilon^\beta > \vec{0}$

$$\alpha - \epsilon^\alpha \in P(x, y) \text{ и } \beta - \epsilon^\beta \in P(y, z) \quad \Rightarrow \quad \alpha + \beta - \frac{\epsilon^\alpha + \epsilon^\beta}{2} \in P(x, z).$$

- По лемме о PAD, $f(v) = x$.
- Значит, $\alpha + \beta - \frac{\epsilon^\alpha + \epsilon^\beta}{2} = \mathbf{v}(x) - \mathbf{v}(z) \in P(x, z)$.

Доказательство

- Если бы было верно, что $\vec{0} \in P(x, y)$, мы бы уже всё доказали: из последней леммы следовало бы, что внутренности всех $P(x, y)$ равны.
- Но у нас это не обязательно так, поэтому мы множества сдвинем на

$$\gamma(x, y) = \inf\{p \in \mathbb{R} \mid p \cdot \vec{1} \in P(x, y)\}.$$

Упражнение. Докажите, что такое $\gamma(x, y)$ существует (т.е. что множество непусто и ограничено снизу).

Доказательство

Лемма

Для всех $x, y, z \in \mathcal{O}$*:*

1. $\gamma(x,y) = -\gamma(y,x)$;
2. $\gamma(x,z) = \gamma(x,y) + \gamma(y,z)$.

- Первый пункт: для всякого $\epsilon > 0$ $(\gamma(x,y) + \frac{\epsilon}{2}) \cdot \vec{1} \in P(x,y)$.
- Значит, по лемме, $(-\gamma(x,y) - \epsilon) \cdot \vec{1} \notin P(y,x)$.
- С другой стороны, $(\gamma(x,y) - \epsilon) \cdot \vec{1} \notin P(x,y)$, и $(-\gamma(x,y) + \frac{\epsilon}{)} \cdot \vec{1} \in P(y,x)$.

Доказательство

Лемма

Для всех $x, y, z \in \mathcal{O}$:

1. $\gamma(x, y) = -\gamma(y, x)$;
2. $\gamma(x, z) = \gamma(x, y) + \gamma(y, z)$.

- Второй пункт: $(\gamma(x, y) + \frac{\epsilon}{2}) \cdot \vec{1} \in P(x, y)$, $(\gamma(y, z) + \frac{\epsilon}{2}) \cdot \vec{1} \in P(y, z)$, значит, по лемме, $(\gamma(x, y) + \gamma(y, z) + \epsilon) \cdot \vec{1} \in P(x, z)$. Это в одну сторону.
- А в другую сторону из того, что $\gamma(z, x) \leq \gamma(z, y) + \gamma(y, x)$ и первого пункта этой леммы.

Доказательство

- Теперь можно сдвинуть множества $P(x, y)$:

$$C(x, y) = P(x, y) - \gamma(x, y) \cdot \vec{1}.$$

- Через $\dot{C}$ обозначим внутренность C.

Доказательство

Лемма

$\dot{C}(x,y) = \dot{C}(w,z)$ *для любых* $x, y, w, z \in \mathcal{O}$, $x \neq y$, $w \neq z$.

- По лемме, $\dot{P}(x,y) \subseteq \dot{P}(x,z) - \beta$ для любого $\beta \in \dot{P}(y,z)$, в частности, для $\beta = (\gamma(y,z) + \epsilon) \cdot \vec{1}$.
- Аналогично, $\dot{P}(x,z) \subseteq \dot{P}(w,z) - \alpha$ для $\alpha = (\gamma(w,x) + \epsilon) \cdot \vec{1}$.
- Т.е. $\dot{P}(x,y) \subseteq \dot{P}(w,z) - (\gamma(y,z) + \gamma(w,x)) \cdot \vec{1}$.

Доказательство

Лемма

$\dot{C}(x,y) = \dot{C}(w,z)$ для любых $x,y,w,z \in \mathcal{O}$, $x \neq y$, $w \neq z$.

- По предыдущей лемме, $\gamma(y,z) + \gamma(w,x) = \gamma(y,z) + \gamma(w,y) + \gamma(y,x) = \gamma(w,z) - \gamma(x,y)$.
- Значит, $\dot{P}(x,y) - \gamma(x,y)\cdot\vec{1} \subseteq \dot{P}(w,z) - \gamma(w,z)\cdot\vec{1}$.

Доказательство

Лемма

$\dot{C}(x, y) = \dot{C}(w, z)$ для любых $x, y, w, z \in \mathcal{O}$, $x \neq y$, $w \neq z$.

- Получилось, что $\dot{C}$ все равны. Обозначим $C = \dot{C}(x, y)$.

Упражнение. Докажите, что, например, $\dot{C}(x, y) = \dot{C}(y, x)$, т.е. все эти неравенства не обязательны.

Доказательство

Лемма

C выпукло.

- Пусть $\alpha, \beta \in C$. Зафиксируем разные $x, y, z \in \mathcal{O}$.
- $\gamma(x, y) \cdot \vec{1} + \alpha \in \dot{P}(x, y)$ и $\gamma(y, z) \cdot \vec{1} + \beta \in \dot{P}(y, z)$, значит, $\gamma(x, z) \cdot \vec{1} + \alpha + \beta \in \dot{P}(x, z)$, и $\alpha + \beta \in C$.

Доказательство

Лемма

C выпукло.

- Пусть $\alpha \in C$, но $\frac{\alpha}{2} \notin C$. Тогда $\frac{\alpha}{2} + \gamma(x,y) \cdot \vec{1} \notin P(x,y)$, значит, $-\frac{\alpha}{2} - \gamma(x,y) \cdot \vec{1} \in P(y,x)$, значит, $-\frac{\alpha}{2} \in C$, значит, $\frac{\alpha}{2} \in C$.
- Значит, $\frac{\alpha+\beta}{2} \in C$, т.е. C выпукло.

Доказательство

- Завершаем доказательство теоремы. Во-первых, $\vec{0} \notin \dot{C}$, т.к. он должен быть на границе, если он там.
- Значит, существует $\mathbf{k} \in \mathbb{R}^N$, для которого $\mathbf{k} \cdot \alpha \geq 0$ для любого $\alpha \in \bar{C}$ (в замыкании).
- Этот вектор $\mathbf{k}$ будет теми константами k_i, которые нас интересуют для аффинного максимизатора.

Доказательство

- Зафиксируем исход $x_0 \in \mathcal{O}$ и константы

$$C_x = \sum_{i=1}^{N} k_i \gamma(x_0, x).$$

- Докажем теперь все необходимые неравенства, т.е.

$$\sum_{i=1}^{N} k_i (v_i(x) - v_i(y)) \geq C_y - C_x,$$

где $f(v) = x \neq y$.

Доказательство

- Когда $f(v) = x \neq y$, $v(x) - v(y) \in P(x, y)$. Обозначим $\alpha = v(x) - v(y) - \gamma(x, y) \cdot \vec{1}$. Тогда $\alpha \in \bar{C}$.
- Значит, $\mathbf{k} \cdot \alpha \geq 0$. Так как $-\gamma(x, y) = \gamma(x_0, x) - \gamma(x_0, y)$,

$$\mathbf{k} \cdot \mathbf{v}(x) + C_x \geq \mathbf{k} \cdot \mathbf{v}(y) + C_y,$$

и тем самым получается утверждение теоремы.

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.