Utilitarian Theorems and Equivalence of Utility Theories

細矢祐誉

中央大学

September 16, 2023

功利主義定理

X は社会の状態を表す集合とし、X 上に社会の倫理的順序 $\succsim_0$ と、個人の選好順序の列 $(\succsim_1, ..., \succsim_n)$ が存在するとする。また、これらに対する効用関数の導出についてのルールがあり、v は $\succsim_0$ を表現していて、u_i は $\succsim_i$ を表現しているとする。このとき、いくつかの追加的な公理の下で

$$v(x) = \sum_{i=1}^{n} a_i u_i(x) + b$$

が成り立つことを証明するのが、功利主義定理と呼ばれるタイプの定理である。

ハーサニーの功利主義定理（1）

功利主義定理は古くは1950年代初頭からあるようだが、最も有名なのはHarsanyi (1955) による期待効用関数の定理である。

定理（Harsanyi）

$\mathscr{P}$ は X 上の有限くじの集合とし、$(\succsim_0, ..., \succsim_n)$ はこの上の弱順序の族であり、$v : X \to \mathbb{R}$ は $\succsim_0$ のNM効用、$u_i : X \to \mathbb{R}$ は $\succsim_i$ のNM効用であるとする。このとき、以下の二つの条件は同値である。

(i) すべての個人 $i \in \{1, ..., n\}$ に対して $P \sim_i Q$ ならば、$P \sim_0 Q$ である。

(ii) ある $a_1, ..., a_n \in \mathbb{R}$ と $b \in \mathbb{R}$ に対して以下が成り立つ。

$$v(x) = \sum_{i=1}^{n} a_i u_i(x) + b$$

ハーサニーの功利主義定理（2）

二つほど補足しておく。第一に、公理の中に「パレート改善を社会の選好はよいと判断する」という公理がない。この結果、$a_i > 0$ は証明できない。ハーサニーの定理の条件だと、その個人を悪くすればするほど社会がよいと判断するような個人が存在する可能性は否定できない。一応、パレート原理を追加で仮定すると、a_i をすべて正である「ように取れる」ことは証明できる。ただし、a_i の一意性が証明できないため、a_i が必ず正になることは証明できない。
第二に、ハーサニーはこの定理の証明を少し間違えていて、仮定していない性質を追加で使っている。この問題は、その後に Hammond (1992) がより洗練された証明を提出したことで解決したが、発表された当初は問題を含む定理であったことに注意しなければならない。

ハーサニーの功利主義定理（3）

通常、功利主義の文脈においては、効用関数の数値の差は改善の程度を表現することが要請される。しかし、ハーサニーが使ったのは「NM効用」という、まったく異なる文脈のものだった。この定理の後に、ハーサニーはNM効用が改善の程度を表現するという趣旨で長い論考をしている。

しかし、NM効用は確率的な事象に関連する効用概念であり、したがって改善の程度に対応する意味を持たないと考える人間も多かった。例として、Luce and Raiffa (1957) は、５割の確率で 10 ドル失うくじと確実に 9 ドル失うことが等価な個人を挙げて、「9 ドル失ってからさらに追加で 1 ドル失うことは、元の 9 ドルを失うことの二倍嫌われている、という解釈は取りたくない」と述べている。

ハーサニーの功利主義定理（4）

より徹底的にハーサニーと戦ったのは Diamond (1967) や Rawls (1971)、そして Sen (1970,1973,1976,1981) などである。特にセンはハーサニーと Theory and Decision 誌上で何度も論戦を交わしている。センらの主張としては、ハーサニーのセットアップ自体が問題であって、たとえば社会がNM効用を持つことなどが批判の対象となっていた。
実はこれらの批判は二つに分かれていて、数学的なものとそうでないものである。ルース＝レイファは数学的な批判ではなかったし、ロールズも同様である。一方でダイアモンドやセンは公理を攻撃していた。ハーサニーは数学的な批判以外にはほとんど反応しておらず、おそらくは自身の結果は数学的に証明できるため、相手の好みと関係なく受け入れざるを得ない結果だと考えていたと思われる。
この、ハーサニーの考えが正しいか否かというのを深く考察していくのが今回の研究である。

準可分社会（1）

主結果のための準備として、Harvey(1999) の功利主義定理を紹介しよう。これは Alt (1936) による、改善の強度を表す関数についての結果である。しかしその結果を説明する前に、準可分社会という用語を定義しておく。社会の状態を表す集合を X で表したとき、X 上の弱順序の族 $(\succsim_0, ..., \succsim_n)$ を X 上の**社会**と呼ぶことにする。この社会が**準可分** (semi-separable) であるという用語を以下のように定義する：$i \in \{1, ..., n\}$ に対してそれぞれ $x_i \in X$ を取ってきたとき、$x \sim_i x_i$ がすべての i に対して成り立つ $x \in X$ が必ず存在するとき、社会は準可分であると言う。
理解しやすくするために、**可分** (separable) という用語も定義しておこう。これは、$X = \prod_{i=1}^n X_i$ であり、$\succsim_i$ は X_i 以外の要素に影響を受けないような社会を指す用語である。

準可分社会（2）

外部性のない純粋交換経済では、Ω_i を i の消費集合としたとき、$X = \prod_{i=1}^n \Omega_i$ が社会の状態の全体である。そして個人の選好 $\succsim_i$ は Ω_i 上で定義されるため、これは可分な社会である。
一方で生産が入ってきた場合、$X = \prod_{i=1}^n \Omega_i \times \prod_{j=1}^m Y_j$ が社会の状態の全体であるため、これは可分ではない。しかし、
$z^i = (x_1^i, ..., x_n^i, y_1^i, ..., y_m^i)$ としたときに、$z = (x_1^1, ..., x_n^n, y_1^1, ..., y_m^1)$ とすれば $z \sim_i z^i$ なので、準可分である。
注意して欲しいのは、z^i がすべて実現可能だったとしても z は実現可能であるとは限らないということである。このように、社会が準可分かどうかを判定するためには実現可能性は関係ない。

準可分社会（3）

サイドペイメントを許す外部性つきの純粋交換経済は、t_i を i が受け取る金額（負なら支払う金額）として、$X = \mathbb{R}^n \times \prod_{i=1}^n \Omega_i$ が社会の状態を表す。そして、$\prod_{i=1}^n \Omega_i$ 上で定義された実数値関数 $u_i(x)$ を用いて、

$$U_i(t, x) = u_i(x) + t_i$$

が表現する順序を i が持っているとすると、これは準可分社会である。なぜなら、どの個人も t_i を適切に変えることによって、任意の効用水準に調整できるからである。ここでもやはり、t を運用するための資金の出所について考える必要はない。
外部性だけではなく公共財のモデルなども同様で、準可分社会は思った以上に多くのモデルをカバーしている。

アルトの効用関数（1）

さらに準備として、アルトの効用関数について説明しておく。$\geq$ は X^2 上で定義された弱順序とする。$(x, y, z, w) \in \geq$ は $[x, y] \geq [z, w]$ と書き、$[x, y] > [z, w]$ および $[x, y] = [z, w]$ はそれぞれ対応する強い順序関係と無差別関係を表すとする。実数値関数 u が $\geq$ を表現するアルトの効用関数であるとは、

$$[x, y] \geq [z, w] \Leftrightarrow u(x) - u(y) \geq u(z) - u(w)$$

であることを意味する。
解釈としては、$[x, y] \geq [z, w]$ は、y から x への移動がもたらす改善が、w から z への移動がもたらす改善よりも強いことを意味している。

アルトの効用関数（2）

X^2 上の弱順序 $\geq$ が X 上の弱順序 $\succsim$ に適合しているという用語を、

$$x \succsim y \Leftrightarrow [x, y] \geq [y, y]$$

が成り立っていることで定義する。適合している弱順序が存在するとき、$\geq$ に対するアルトの効用関数 u は、

$$x \succsim y \Leftrightarrow [x, y] \geq [y, y] \Leftrightarrow u(x) - u(y) \geq u(y) - u(y) \Leftrightarrow u(x) \geq u(y)$$

となるため、通常の意味で $\succsim$ を表現する効用関数となっている。同様に、X^2 上の社会 $(\geq_0, ..., \geq_n)$ が X 上の社会 $(\succsim_0, ..., \succsim_n)$ に適合しているという用語を、すべての i について $\geq_i$ が $\succsim_i$ と適合していることで定義する。

パレート原理

あとひとつ用語を与えておく。X 上の社会 $(\succsim_0, ..., \succsim_n)$ が与えられているとき、x が y を**パレートの意味で改善する**とは、すべての $i \in \{1, ..., n\}$ に対して $x \succsim_i y$ であり、かつ少なくとも一つの i については $x \succ_i y$ であることを指す。社会 $(\succsim_0, ..., \succsim_n)$ が**パレート原理**に従うとは、x が y をパレートの意味で改善してるときには必ず $x \succ_0 y$ であることを言う。

ハーヴェイの功利主義定理

ハーヴェイの功利主義定理は以下の通りである。

定理 (Harvey)

X は連結なハウスドルフ位相空間であり、X 上の社会 $(\succsim_0, ..., \succsim_n)$ とそれと適合した X^2 上の社会 $(\geq_0, ..., \geq_n)$ が存在して、$\geq_0$ は連続なアルトの効用 v を、各 $\geq_i$ も連続なアルトの効用 u_i を持っているとする。また社会 $(\succsim_0, ..., \succsim_n)$ は準可分かつパレート原理に従うとする。このとき、以下の二条件は同値である。

(I) すべての個人 $i \in \{1, ..., n\}$ に対して $[x, y] =_i [z, w]$ ならば、$[x, y] =_0 [z, w]$ である。

(II) ある $a_1, ..., a_n > 0$ と $b \in \mathbb{R}$ に対して以下が成り立つ。

$$v(x) = \sum_{i=1}^{n} a_i u_i(x) + b$$

社会の確率的拡張

もうひとつだけ概念を考えておく。ハーヴェイの定理では我々は X を通常の位相空間と仮定していたが、ハーサニーの定理では X は単純くじの空間であった。これらを区別する表記が欲しいので、我々は X 上の単純くじ、つまり台が有限な確率測度の空間を $\mathscr{P}$ で表す。元の空間 X が位相空間である場合には、$\mathscr{P}$ には弱∗位相を入れる。$\mathscr{P}$ 上の順序 $\succeq$ が X 上の順序 $\succsim$ の確率的拡張であるとは、ディラック測度 δ_x を用いて

$$x \succsim y \Leftrightarrow \delta_x \succeq \delta_y$$

が成り立つことを言う。同様に、$\mathscr{P}$ 上の社会 $(\succeq_0, ..., \succeq_n)$ が X 上の社会 $(\succsim_0, ..., \succsim_n)$ の確率的拡張であるという用語も、各 i について $\succeq_i$ が $\succsim_i$ の確率的拡張であるとして定義する。

主結果

以下の結果が主たる成果である。

定理

X は連結なハウスドルフ位相空間であり、$(\succsim_0, ..., \succsim_n)$ は X 上のパレート原理に従う準可分な社会で、ハーヴェイの定理の仮定をすべて満たす適合した X^2 上の社会 $(\geq_0, ..., \geq_n)$ と、ハーサニーの定理の仮定をすべて満たす確率的拡張である $\mathscr{P}$ 上の社会 $(\succeq_0, ..., \succeq_n)$ を持っているとする。さらに、各 $\succeq_i$ を表現するNM効用は連続であるとする。そしてハーサニーの定理の公理 (i) と、ハーヴェイの定理の公理 (I) が成り立っているとし、$\succsim_i \neq X^2$ とならない個人が少なくとも二人はいるとする。このとき、任意の個人 $i \in \{1, ..., n\}$ について、NM効用とアルトの効用は（正アフィン変換を除いて）一致する。

例（1）

主結果を理解しやすいように例を作る。いま、簡単化のために $X = \mathbb{R}^2_+$ とし、$n = 2$ で、i は x_i にのみ関心を持つと仮定する。$u_i(x_i)$ は i のアルトの効用を、$u_i^*(x_i)$ はNM効用を表すとし、また社会についても $v(x_1, x_2)$ がアルトの効用、$v^*(x_1, x_2)$ がNM効用とする。ハーサニーの定理とハーヴェイの定理が両方成り立つことを仮定すると

$$v(x_1, x_2) = a_1 u_1(x_1) + a_2 u_2(x_2) + b,$$

$$v^*(x_1, x_2) = a_1^* u_1^*(x_1) + a_2^* u_2^*(x_2) + b^*$$

と書けることになるが、ここでは簡単化のために $a_1 = a_2 = a_1^* = a_2^* = 1,\ b = b^* = 0$ として話を進める。したがって、

$$v(x_1, x_2) = u_1(x_1) + u_2(x_2),\ v^*(x_1, x_2) = u_1^*(x_1) + u_2^*(x_2)$$

が以後仮定される。

例（2）

u_i と u_i^* は同じ順序を表現し、v と v^* も同じ順序を表現しているとする。仮にここで、$u_1(x_1) = x_1$ かつ $u_1^*(x_1) = \sqrt{x_1}$ であったとしよう。この場合、なにか矛盾が生ずるだろうか？
たとえば $i = 2$ がすべてについて無差別と判断する個人である場合を考えよう。この場合、$u_2(x_2)$ と $u_2^*(x_2)$ は共に定数関数なので、定数 c_1, c_1^* を用いて

$$v(x_1, x_2) = u_1(x_1) + c_1,\ v^*(x_1, x_2) = u_1^*(x_1) + c_1^*$$

となり、この場合には v と v^* は同じ順序を表現するため、なにも矛盾は生じない。したがって、$i = 2$ がすべてについて無差別と判断する個人の場合には、アルトの効用とNM効用は正アフィン変換ではない関係を持ちうることがこの例からわかる。

例（3）

今度は $i=2$ が無差別と判定しない x_2^1, x_2^2 が存在するとしよう。一般性を失うことなく $u_2^*(x_2^2) - u_2^*(x_2^1) = \varepsilon > 0$ と仮定する。ここで、$x_1^k = (k\varepsilon)^2$ と定義すると、$u_1^*(x_1) = \sqrt{x_1}$ であるから、

$$u_1^*(x_1^{k+1}) + u_2^*(x_2^1) = u_1^*(x_1^k) + u_2^*(x_2^2)$$

であるため、

$$v^*(x_1^{k+1}, x_2^1) = v^*(x_1^k, x_2^2)$$

であり、よって $(x_1^{k+1}, x_2^1) \sim_0 (x_1^k, x_2^2)$ でなければならない。

例（4）

したがって
$$v(x_1^{k+1}, x_2^1) = v(x_1^k, x_2^2)$$
を得るため、
$$u_1(x_1^{k+1}) + u_2(x_2^1) = u_1(x_1^k) + u_2(x_2^2)$$
であり、整理して
$$u_1(x_1^{k+1}) - u_1(x_1^k) = u_2(x_2^2) - u_2(x_2^1)$$
を得る。ここで $u_1(x_1) = x_1$ なので、これは $x_1^{k+1} - x_1^k$ がすべての k について同じという意味になる。ところが明らかに
$$x_1^{k+1} - x_1^k = (2k+1)\varepsilon^2$$
で、右辺は k によって異なるため、矛盾が生ずる。
このように、ひとつでも $i = 2$ が無差別と感じないペアがある場合には、u_1 と u_1^* が正アフィン変換で一致しないことは許されないのである。これが、主結果の具体的な意味である。

主結果の補足（１）

ルース＝レイファの批判は、NM効用関数をアルトの効用関数であるかのように議論するべきではないというものだった。しかしハーサニーとハーヴェイの二つの功利主義定理が同時に成り立つ環境では、実は二つの効用概念は一致してしまうというのが我々の結果である。この意味でハーサニーの言う通り結果は数学的であって、ルース＝レイファの批判は的外れである。
しかし一方で、なぜこんなことが起こったのかについて深く考えてみる必要がある。実際のところ、ハーサニーの定理とハーヴェイの定理のどちらにも、NM効用とアルトの効用を結びつけるような公理は存在しない。なので、一見してわからないが実は極めて強い公理が、どこかに潜んでいるのではないかという疑いを抱くことになる。

主結果の補足（2）

NM効用やアルトの効用が存在するという仮定は、ないと当然それらの一致も示せないという意味で必要であるが、それが強いかと言われるとピンと来ない。一方で u_i^* が連続だという仮定は追加的であるが、アルトの効用は普通連続なので、NM効用が連続でないと一致は証明できず、だから必要だという意味で自然である。そうなると (i) と (I) の二つの公理が怪しくなってくる。(I) については、強いところを見いだすことはできない。しかし (i) はけっこうな強さである。以下、最も簡単な可分社会、つまり $X = \mathbb{R}^n$ であり、$\succsim_i$ が $u_i(x_i)$ という関数で表現できている状況を考えてみよう。このとき、(i) が意味するのは、x_i にのみ確率的に作用する任意のくじについて、社会のリスクに対する態度（確実性等価やリスク回避度）は個人 i のそれと一致しなければならない、という意味になる。

主結果の補足（３）

Harsanyi (1975) は、ダイアモンドやセンの批判に答えて、「他人のためや社会のために動くときには、自身のために動くときよりも高度に合理的な態度を要求されると考えるのが自然である」と述べた。これはハーサニー自身は社会の選好に独立性公準を課すことの擁護のつもりだったようであるが、我々からすると、この文章は「社会は個人より慎重に動くべきである」というような別の意味合いにも取れる。実際、センは「高度に合理的な態度」が独立性を意味するという考えに疑問を呈している。そして、「高度に合理的な態度」が「より慎重な態度＝よりリスク回避的な態度」を意味するのであれば、ハーサニーの公理 (i) はむしろそれを禁止しているのである。この意味では、ハーサニーの公理 (i) は彼が考えているよりずっと強い仮定であり、これがすべての根源ではないかと考えられる。

主結果の補足（4）

したがって、ルース＝レイファの批判を現代化しようとすれば、この公理 (i) の是非ということになるだろう。この公理を認めざるを得ないと考えるのならば、アルトの効用とNM効用は一致せざるを得ない。対偶を取れば、アルトの効用とNM効用の一致を否定するためには、この公理を否定せざるを得ないのである。この公理が妥当か否かを問う、という形で、我々はルース＝レイファの批判の現代化に成功したと考えている。

主結果の補足（5）

ただし、これらは「社会」という言葉を文字通りに捉えた場合である。実際のところ、社会と呼んでいるのは個人かも知れない。我々は意思決定に際して複数の基準を総合して決定する必要がある場合がある。たとえば、おいしい食事を取りたいけれど、医者から高血圧について注意されており、塩分を控えめにしなければならない等である。この場合、$\succsim_1$ が通常の味の好み、$\succsim_2$ が塩分の量についての選好とすると、$(\succsim_1, \succsim_2)$ を総合して決定した食事の好み $\succsim_0$ が実際の我々の好みだということになる。このようにあくまで個人の意思決定の範囲内として問題を捉える場合には、公理 (i) は自然な公理と見なしうるかもしれない。ただし、その場合には準可分などの他の仮定について、より詳細に吟味する必要があると思われる。

補足：NM効用の連続性（1）

主結果でNM効用の連続性を仮定したが、この仮定のための十分条件として知られているのは $\mathscr{P}$ における弱∗位相での $\succsim$ の閉性である。しかし、この仮定はNM効用の有界性も同時に意味してしまう。そこで、NM効用の連続性の必要十分条件を示してみよう。まず、$Y \subset X$ とする。このとき、台が Y に含まれる $\mathscr{P}$ の元全体を $\mathscr{P}_Y$ と書くことにする。$\mathscr{P}$ 上の弱順序 $\succsim$ の Y への制限という言葉を、$\succsim \cap (\mathscr{P}_Y)^2$ を意味する言葉として定義しておく。$\mathscr{P}$ 上の弱順序 $\succsim$ が**列連続性**を満たすとは、$X_n \subset \mathrm{int}.X_{n+1}$ と $\cup_n X_n = X$ を満たす集合列 (X_n) が存在して、$\succsim$ の X_n への制限が常に $(\mathscr{P}_{X_n})^2$ の相対位相で閉であることを言う。

補足：NM効用の連続性（2）

定理

X はチコノフ空間、$\mathscr{P}$ はその上の単純くじの集合、$\succsim$ は $\mathscr{P}$ 上の弱順序とする。このとき、以下の二条件は同値である。

1) $\succsim$ を表現する連続なNM効用関数 $u : X \to \mathbb{R}$ が存在する。
2) $\succsim$ は独立かつ列連続である。

補足：この議論を通じて $\mathscr{P}$ には弱*位相を入れているが、これはつまりネット (P_ν) が P に収束するということが、任意の有界連続関数 v についてネット $(E_{P_\nu}[v])$ が $E_P[v]$ に収束するという意味になるような位相という意味である。ただ、一般のハウスドルフ位相空間でそのような位相が存在するかどうかについては精査していない。X がチコノフ空間であれば大丈夫であるが、これを緩めると上の位相がハウスドルフにならないという問題がある。

Thank you for your attention.