Title	合理化可能性問題のグラフ理論的・組合せ的構造
Sub Title	Combinatorial structure of the rationalizability test
Author	塩澤, 康平(Shiozawa, Kohei)
Publisher	慶應義塾経済学会
Publication year	2015
Jtitle	三田学会雑誌 (Mita journal of economics). Vol.108, No.3 (2015. 10) ,p.537(61)- 557(81)
JaLC DOI	10.14991/001.20151001-0061
Abstract	本稿では, 消費データの合理化可能性問題の背後に存在するネットワーク構造やグラフ構造を取り扱う。これにより問題の構造が明確になり, 合理化可能性を検証するための効率的なアルゴリズムを得ることができる。また, データがどの程度効用最大化問題に当てはまるかを測る指数の計算が, NP困難と呼ばれるクラスの最適化問題に属することが示される。そして, 新しい基準として上述のネットワーク構造に照らして自然で, 安易に計算可能なものが提案される。 This study examines a combinatorial or graph theoretic structure of the rationalizability test (revealed preference test) of a price-consumption dataset. This structure is common among linear and non-linear budget rationalizability test and clarifies the nature of the problem. The graph theoretic representation of the problem enables us to efficiently test the rationalizability of a given dataset and easily investigate the computational complexity of the violation sensitivity measures.
Notes	特集 : 経済の数理解析 : 数理経済学の新展開
Genre	Journal Article
URL	https://koara.lib.keio.ac.jp/xoonips/modules/xoonips/detail.php?koara_id=AN00234610-20151001-0061

「三田学会雑誌」108 巻 3 号（2015 年 10 月）

合理化可能性問題のグラフ理論的・組合せ的構造

塩澤康平*

（初稿受付 2015 年 8 月 20 日，査読を経て掲載決定 2015 年 9 月 28 日）

Combinatorial Structure of the Rationalizability Test

Kohei Shiozawa*

Abstract: This study examines a combinatorial or graph theoretic structure of the rationalizability test (revealed preference test) of a price-consumption dataset. This structure is common among linear and non-linear budget rationalizability test and clarifies the nature of the problem. The graph theoretic representation of the problem enables us to efficiently test the rationalizability of a given dataset and easily investigate the computational complexity of the violation sensitivity measures.

1　はじめに

ℓ 財の消費量 $x \in R^{\ell}_{+}$ と消費価格 $p \in R^{\ell}_{++}$ に関する n 組のデータ $\{(p_k, x_k)\}^{n}_{k=1}$ に対して，その各消費量データ $x_k (k = 1, \ldots, n)$ を最適化問題

$$\begin{aligned} &\max\ u(x) \\ &\text{s.t. } p_k \cdot x \leqq p_k \cdot x_k \end{aligned} \tag{1}$$

の最適解として支持するような効用関数 $u : R^{\ell}_{+} \to R$ は存在するか。Afriat（1967）に始まるこの問題を，顕示選好理論における合理化可能性問題と呼ぶ（本稿では，単に合理化可能性問題と呼ぶ）。この問題には常に自明な解が存在し，$u \equiv c$（ただし $c \in R$ はとある定数）とすればよいことがわか

近畿地区数理経済学ジョイント・セミナー（2015 春季）参加者の皆様，慶應義塾大学の尾崎裕之教授，および匿名の査読者の方から有益なコメントを頂きました。記して感謝申し上げます。

*　大阪大学大学院経済学研究科
Graduate School of Economics, Osaka University (nge013sk@student.econ.osaka-u.ac.jp)

る。それでは非自明な解，すなわちデータ $\{(p_k, x_k)\}_{k=1}^n$ を最適化問題（1）の解として支持するような非飽和的な効用関数 $u : R_+^\ell \to R$ が存在するかどうかが問題となる。この問題に対してよく知られた結果が以下の定理である。

定理 1（Afriat（1967）; Diewert（1973）; Varian（1982））．n 組のデータ $\{(p_k, x_k)\}_{k=1}^n$ に対して以下の 4 条件は同値である。

（i） ある非飽和な効用関数 $u : R_+^\ell \to R$ によってデータは合理化可能である。

（ii） データによって定義される以下の不等式系は解 $(U_k, \lambda_k)_{k=1}^n$（ただし $\lambda_k > 0$ $(k = 1\ldots, n)$）をもつ：$U_{k'} \leqq U_k + \lambda_k p_k \cdot (x_{k'} - x_k)$　すべての　$k, k' = 1, \ldots, n$　について。

（iii） データに対して以下の組合せ的な条件が成り立つ：$p_{k_0} \cdot (x_{k_1} - x_{k_0}) \leqq 0$, $p_{k_1} \cdot (x_{k_2} - x_{k_1}) \leqq 0$, ..., $p_{k_m} \cdot (x_{k_0} - x_{k_m}) \leqq 0$ ならば $p_{k_i} \cdot (x_{k_{(i+1)}} - x_{k_i}) = 0$ がすべての $i = 0, \ldots, m$ で成立する（ただし $m + 1 = 0$ とする）。

（iv） ある単調で凹な連続効用関数 $u : R_+^\ell \to R$ によってデータは合理化可能である。

この定理の条件（ii）の不等式系を Afriat の不等式系（Afriat's inequalities）と呼ぶ。また条件（iii）を巡回的整合性（cyclical consistency）条件と呼ぶ。この定理は Afriat（1967）によって $p_k \cdot (x_{k'} - x_k) \neq 0$（すべての $k \neq k'$ について）の場合について示された。Varian（1982）はこの仮定が満たされない（すなわち，とある $k \neq k'$ に対して $p_k \cdot (x_{k'} - x_k) = 0$ となりうる場合を含めた）場合に対して証明を与えた。また Varian（1983）では，効用関数の微分可能性を仮定した上で，この定理の特に（iv）⇒（ii）⇒（i）の部分が成立することについて簡単な証明を与えている。この定理の（iii）⇒（ii）の部分について Varian（1982）の与えたアルゴリズム的な証明よりも単純な議論である Fostel et al.（2004）なども知られている。（Fostel et al.（2004）はデータ数に関する数学的帰納法にもとづく簡単な証明である。）さらに，Geanakoplos（2013）により，この定理と von Neumann のミニマックス定理との関連も明らかになっている。

この定理にもとづき，データ $\{(p_k, x_k)\}_{k=1}^n$ が与えられたときに，その合理化可能性を検証する方法がいくつか考えられる。そのうちの 1 つは，Diewert（1973）の議論したように線形計画法により Afriat の不等式系を解く方法であり，この方法は単体法や内点法といった線形計画アルゴリズムにより，十分現実的な時間での検証が可能である。また，Varian（1982）では，巡回的整合性条件と同値な条件である GARP（generalized axiom of revealed preference; 顕示選好の一般化公理）を考え，合理化可能性を検証するための具体的な手順を提案しており，これは $O(n^3)$ の手間で計算できる十分に効率的な計算手法であるため，現在でも多くの研究でその方法が用いられている。

定理 1 は線形の予算制約をもつ最適化問題（1）に対する結果であるが，Forges and Minelli（2009）によって合理化可能性問題はより広範な，非線形予算制約を許す一般の場合に拡張されている。そ

のような一般的な定式化に対しても定理 1 とほぼ同様の結果が得られており，その結果は以下の定理にまとめられる。

定理 2（Forges and Minelli (2009)）. n 組のデータ $\{(x_k, B_k)\}_{k=1}^n$ を考える（ただし各 $k = 1, \ldots, n$ に対して $B_k := \{x \in R_+^\ell | g_k(x) \leqq 0\}$ であり $g_k : R_+^\ell \to R$ は単調な連続関数で $g_k(x_k) = 0$ を満たすものとする）。データ $\{(x_k, B_k)\}_{k=1}^n$ に対して以下の 3 条件は同値である。

（i） ある非飽和な連続効用関数 $u : R_+^\ell \to R$ が存在してデータを合理化できる：$g_k(x) \leqq g_k(x_k) \Rightarrow u(x) \leqq u(x_k)$。

（ii） データによって定義される以下の不等式系は解 $(U_k, \lambda_k)_{k=1}^n$（ただし $\lambda_k > 0\,(k = 1 \ldots, n)$）をもつ：$U_{k'} \leqq U_k + \lambda_k g_k(x_{k'})$ すべての $k, k' = 1, \ldots, n$ について。

（iii） データに対して以下の組合せ的な条件が成り立つ：$g_{k_0}(x_{k_1}) \leqq 0,\ g_{k_1}(x_{k_2}) \leqq 0,\ \ldots,\ g_{k_m}(x_{k_0}) \leqq 0$ ならば $g_{k_i}(x_{k_{(i+1)}}) = 0$ がすべての $i = 0, \ldots, m$ で成立する（ただし $m + 1 = 0$ とする）。

この一般的な形の定理に線形の場合も含まれていることは明らかである（実際，$p_k \in R_{++}^\ell$ かつ $x_k \in R_+^\ell$ なので，$g_k(x) := p_k \cdot (x - x_k)$ と定義すればこの $g_k : R_+^\ell \to R$ は定理 2 の仮定を満たす）。

以上のように，データの合理化可能性問題に関する不等式系による特徴づけが明らかになる一方で，Varian（1983），Piaw and Vhora（2003）らによって線形予算制約の合理化可能性問題と，古典的な組合せ最適化問題である最短経路問題との関係が指摘されている。また Kolesnikov et al.（2013）では Varian（1983）の指摘を，モンジュ・カントロビッチ問題と関連づけて議論している。さらに，Fujishige and Yang（2012）および Nobibon et al.（2014）では，グラフの強連結成分という構造を用いることによって，Varian（1982）の提案したものより計算効率のよいアルゴリズムによって合理化可能性の検証が可能であることが示されている。本稿の前半では，この合理化可能性問題と最短経路問題との関係，合理化可能性問題のもつ組合せ的な構造についての議論が，定理 2 に与えた非線形予算制約も許す一般の合理化可能性問題の場合でも同様に成り立つことを示す。この 2 つの場合は本質的に同様のネットワーク構造を有していることがわかる。またこの構造を利用することで Fujishige and Yang（2012）および Nobibon et al.（2014）の提案した合理化可能性の検証アルゴリズムが自然に現れる。

本稿の後半では，合理化可能性問題のもう 1 つの側面である「当てはまりの良さ基準（goodness of fit measure）」について議論する。上に見たことから明らかなように，合理化可能性のテストは「合理化可能か否か」という二者択一型のテストになっている。従って，例えばとあるデータが「合理化不可能」であるとき，その程度（どの程度「合理化不可能」なのか）を測る基準を考えることに意味が出てくる。このような基準を，逆の言い方であるが「当てはまりの良さ基準」と呼ぶ。これまでにも様々な基準が提案されており，その多くは組合せ最適化問題として定式化されるものであ

る。しかも，それらのうちいくつかが NP 困難と呼ばれる「難しい」クラスの最適化問題に属することが知られている（Smeulders et al., 2013, 2014）。本稿では特に Houtman and Maks（1985）によって提案された HM 指数と，Dean and Martin（2010, 2015）によって提案された最小費用指数（minimum cost index）に注目する。HM 指数はすでに Smeulders et al.（2014）によって NP 困難性が示されているが，最小費用指数についてはその計算の NP 困難性が Dean and Martin（2010, 2015）の議論によって示唆されるにとどまっている。本稿の前半で議論するような合理化可能性問題のグラフ理論的・ネットワーク的な構造を用いることで，HM 指数の計算問題が NP 困難と呼ばれる最適化問題のクラスに属することを最小フィードバック頂点集合問題（minimum feedback vertex set problem）と呼ばれる組合せ最適化問題への多項式時間還元により，Smeulders et al.（2014）よりも直接的に示すことができる(1)。また，Dean and Martin（2010, 2015）の定義するところの最小費用指数を計算する問題も NP 困難なクラスに属する最適化問題であることを，最大重み非閉路的部分グラフ問題（maximum acyclic subgraph problem）と呼ばれる NP 困難な組合せ最適化問題への多項式時間還元により示す。そして本稿の最後には，多項式時間アルゴリズムによって計算可能であり，前半に見た合理化可能性問題のグラフ理論的な構造に照らして自然なものとなるような，新しい当てはまりの良さ基準である強連結成分指数（strongly connected component index）を提案する。

2　合理化可能性問題と最短経路問題

合理化可能性問題と最短経路問題とを接続するのは Afriat の不等式系である（以下では非線形予算制約を許した一般の場合（定理 2）に沿って考える。先に注意したように，線形予算制約の場合は $g_k(x) := p_k \cdot (x - x_k)$ とすればデータ $\{(x_k, B_k)\}_{k=1}^n$ に関する定理 2 の仮定が満たされる。また，通常は Afriat の不等式系と呼ぶのは線形予算制約の場合であるが，本稿では特に区別することなく，一般の場合を特徴づける不等式系にもこの語を用いることとする）。すなわち，ある $2n$ 個の実数 $(U_k, \lambda_k)_{k=1}^n$（ただし $\lambda_k > 0$ $(k = 1 \ldots, n)$）が存在して

$$U_{k'} \leqq U_k + \lambda_k g_k(x_{k'}) \quad \text{すべての } k, k' = 1, \ldots, n \text{ について}$$

が成り立つ，という線形不等式系の可解条件である。一方で，古典的な組合せ最適化問題の 1 つに（単一始点全終点）最短経路問題がある。最短経路問題には，解（ある固定された地点から他の各地点への最短経路）が存在することの特徴づけとしてよく知られる不等式系があり，Afriat の不等式系はこの不等式系条件の一種であることがわかる。最短経路問題に関する不等式系条件とはすなわち，あ

(1)　Smeulders et al.（2014）でも指摘されている通り，Houtman and Maks（1985）ですでに最小フィードバック頂点集合問題と HM 指数との関係が議論されている。

る実数ベクトル $\pi \in R^{|V|}$ が存在して

$$\pi(u) - \pi(v) \leqq \ell((v, u)) \quad \text{すべての } (v, u) \in E \text{ について} \tag{2}$$

が成り立つ，という線形不等式系の可解条件である（ただし，下に定義する通り，集合 V と集合 E はそれぞれグラフ G の点の集合と辺の集合である）。以下に最短経路問題の定義と，この特徴づけをまとめる。

有限集合 V とその順序対の部分集合 $E \subset V \times V$ の組 $G = (V, E)$ を有向グラフと呼ぶ（並列辺やループは考えない）。V の要素を点，E の要素を辺と呼ぶ。点と辺の交互列 $v_0(v_0, v_1)v_1(v_1, v_2)v_2 \cdots v_{(m-1)}(v_{(m-1)}, v_m)v_m$ を有向路（v_0-v_m 有向路）と呼び，記号では $v_0 \to v_1 \to \cdots \to v_m$ と表す。有向路が $v_0 = v_m$ を満たすとき有向閉路と呼ぶ。各辺に（負の値を許した）実数値を対応させる関数 $c : E \to R$ を重み関数という。重み関数 $c : E \to R$ に関して，有向路（有向閉路）$v_0 \to v_1 \to \cdots \to v_m$ の重みを和 $\sum_{k=1}^{m} c((x_{k-1}, x_k))$ で定義する。点 $s \in V$ から点 $v \in V$ への最短経路とは，グラフ $G = (V, E)$ に存在する s-v 有向路のうち，$c : E \to R$ に関してその重みが最小になるものである。

最短経路問題は，グラフ $G = (V, E)$，重み関数 $c : E \to R$，およびスタート地点 $s \in V$ を与えられたとき，地点 s からほかの地点 $v \in V (v \neq s)$ への最短経路を求める問題として定式化される。最短経路問題には，その可解性（最短経路の存在）に対するよく知られた特徴づけがある。

定理 3（室田・塩浦（2013）参照）. $G = (V, E)$ を有向グラフとする。$c : E \to R$ を重み関数とする。 また，点 $s \in V$ からその他の点 $v \in V \, (v \neq s)$ へは少なくとも 1 つの有向路が存在すると仮定する。このとき，以下の 3 条件は同値である。

（i） 点 s から他の任意の点 v への最短経路が存在する。
（ii） ある実数ベクトル $\pi \in R^{|V|}$ が存在して以下が成り立つ： $\pi(u) - \pi(v) \leqq \ell((v, u)), \forall (v, u) \in E$.
（iii） 負の重みをもつ有向閉路は存在しない（任意の有向閉路の重みは非負である）。

負の重みをもつ閉路が存在するとき，その閉路に沿って有向路を延長すればいくらでも重みを減少させることができるため，その閉路を構成する任意の点への最短経路は存在しないことが直感的にわかる。この定理は，この直感が一般の場合にも成立して，その逆も成り立つことを示している。定理 3 の条件（ii）が先に述べた不等式系条件（2）である（この条件を満たす実数値ベクトル $\pi \in R^{|V|}$ はポテンシャルと呼ばれる）。そこで，合理化可能性問題に対するデータ $\{(x_k, B_k)\}_{k=1}^{n}$ に関連する有向グラフ $G = (V, E)$ とその重み $c(\lambda) : E \to R$ を以下のように構成する。(2)

（2） このグラフは Varian（1983），Piaw and Vhora（2003），Fujishige and Yang（2012），および Nobibon et al.（2014）において用いられたものと本質的に同様である。

まず，実数ベクトル $\lambda=(\lambda_1,\ldots,\lambda_n)\gg 0$ を任意に固定する。各消費データ x_k を 1 つの点に対応させ，各消費データ組 $(x_k,x_{k'})$ に辺を張り，その辺の重みを $\lambda_k g_k(x_{k'})\in R$ とする。形式的には，

$$V:=\{x_1,x_2,\ldots,x_n\}\ \text{ および }\ E:=\{(x_k,x_{k'})\,|\,k,k'=1,\ldots,n,\ \text{かつ}\ k\neq k'\} \tag{3}$$

によって V と E を定義し，

$$c((x_k,x_{k'}):\lambda):=\lambda_k g_k(x_{k'}) \tag{4}$$

によって $c(\lambda):E\to R$ を定義する。これによって合理化可能性問題の最短経路問題による言い換えが可能となる。

定理 4. n 組のデータ $\{(x_k,B_k)\}_{k=1}^n$ を考える（ただし各 $k=1,\ldots,n$ に対して $B_k:=\{x\in R_+^\ell|g_k(x)\leqq 0\}$ であり $g_k:R_+^\ell\to R$ は単調な連続関数で $g_k(x_k)=0$ を満たすものとする）。データ $\{(x_k,B_k)\}_{k=1}^n$ に対して以下の 5 条件は同値である。

（i） ある非飽和な連続効用関数 $u:R_+^\ell\to R$ が存在してデータを合理化できる： $g_k(x)\leqq g_k(x_k)\Rightarrow u(x)\leqq u(x_k)$。

（ii） データによって定義される以下の不等式系は解 $(U_k,\lambda_k)_{k=1}^n$（ただし $\lambda_k>0$ $(k=1\ldots,n)$）をもつ：$U_{k'}\leqq U_k+\lambda_k g_k(x_{k'})$　すべての　$k,k'=1,\ldots,n$　について。

（iii） データに対して以下の組合せ的な条件が成り立つ：$g_{k_0}(x_{k_1})\leqq 0,\ g_{k_1}(x_{k_2})\leqq 0,\ \ldots,\ g_{k_m}(x_{k_0})\leqq 0$ ならば $g_{k_i}(x_{k_{(i+1)}})=0$ がすべての $i=0,\ldots,m$ で成立する（ただし $m+1=0$ とする）。

（iv） ある実数ベクトル $\lambda=(\lambda_1,\ldots,\lambda_n)\gg 0$ が存在して，（3）と（4）によって定義する有向グラフ $G=(V,E)$ と重み関数 $c(\lambda):E\to R$ に関して，点 x_1 から他の各点への最短経路が存在する。

（v） ある実数ベクトル $\lambda=(\lambda_1,\ldots,\lambda_n)\gg 0$ が存在して，（3）と（4）によって定義する有向グラフ $G=(V,E)$ と重み関数 $c(\lambda):E\to R$ に関して，負の重みの閉路が存在しない。

ただし，データ $\{(x_k,B_k)\}_{k=1}^n$ が上記条件のいずれか（従ってすべて）を満たすとき，各条件に現れるベクトル $\lambda=(\lambda_1,\ldots,\lambda_n)\gg 0$ は共通のものとしてとれる。

証明. まず，定義したグラフ $G=(V,E)$ は定理 3 の前提を自明に満たしていることに注意する。というのも，定義（3）より x_1 から他の点 $x_k(k=2,\ldots,n)$ への自明な有向路 $(x_1,x_k)\in E$ が存在するからである。

（i）⇔（ii）⇔（iii）は定理 2 の結果である。また，（iv）⇔（v）は定理 3 の結果である。そして（ii）⇔（iv）はグラフの定義（3）と（4）に注意すると容易にわかる。実際，仮定 $g_k(x_k)=0$ より条件（ii）の不等式（Afriat の不等式）は $k=k'$ のとき自明に成立することに注意すると，条件（ii）は（3）と（4）により定義されるグラフ $G=(V,E)$ とそのコスト $c(\lambda):E\to R$ に関するポテン

シャル条件（定理 3 の条件（ii））と同値であることがわかる。従って，定理 3 の（ii）$\Leftrightarrow$（i）から本定理の（ii）$\Leftrightarrow$（iv）が導かれる。 □

この言い換えにより，データの合理化可能性問題と古典的な選好の表現定理（Debreu, 1954, Chapter 3）との類似性が明らかになる。古典的な選好の表現問題では，所与の順序構造 $\succsim \subset X \times X$ に対して，その順序構造と矛盾しないような基数表現 $u : X \to R$ の存在を問題としている。そして，データの合理化可能性問題においては，与えられた順序構造（データをもとに (3) と (4) で定義されるグラフ $G = (V, E)$ とその重み関数 $c(\lambda) : E \to R$）に対して，その順序構造と矛盾しないような数ベクトル（ポテンシャル $(U_k)_{k=1}^n$）が存在するかを問題としている。ただし，後者においては順序構造の符号を変えない範囲での調節 $(\lambda \gg 0)$ をすることが許されている。

さて，この最短経路問題による表現から，順序構造の符号を変えない範囲での調節 $\lambda \gg 0$ は，実は比率（すなわち $\sum_{k=1}^n \lambda_k = 1$ を満たすもの）のみを考えれば十分であることがわかる。

系 1. n 組のデータ $\{(x_k, B_k)\}_{k=1}^n$ を考える（ただし各 $k = 1, \ldots, n$ に対して $B_k := \{x \in R_+^\ell | g_k(x) \leqq 0\}$ であり $g_k : R_+^\ell \to R$ は単調な連続関数で $g_k(x_k) = 0$ を満たすものとする）。データ $\{(x_k, B_k)\}_{k=1}^n$ に対して以下の 2 条件は同値である。

（i） ある実数ベクトル $\lambda = (\lambda_1, \ldots, \lambda_n) \gg 0$ が存在して，(3) と (4) によって定義する有向グラフ $G = (V, E)$ と重み関数 $c(\lambda) : E \to R$ に関して，負の重みの閉路が存在しない。

（ii） ある実数ベクトル $\lambda = (\lambda_1, \ldots, \lambda_n) \gg 0$（ただし $\sum_{k=1}^n \lambda_k = 1$） が存在して，(3) と (4) によって定義する有向グラフ $G = (V, E)$ と重み関数 $c(\lambda) : E \to R$ に関して，負の重みの閉路が存在しない。

証明. （ii）$\Rightarrow$（i）は自明なので逆を示す。実数ベクトル $\lambda = (\lambda_1, \ldots, \lambda_n) \gg 0$ が存在して，(3) と (4) によって定義する有向グラフ $G = (V, E)$ と重み関数 $c(\lambda) : E \to R$ に関して，負の重みの閉路が存在しないとする。そこで新しい調整ベクトル $\bar{\lambda} \gg 0$ を $\bar{\lambda}_k := \lambda_k / (\sum_{h=1}^n \lambda_h)$ $(k = 1, \ldots n)$ によって定義する。このとき $\sum_{k=1}^n \bar{\lambda}_k = 1$ が成立することがわかる。また有向グラフ $G = (V, E)$ の閉路 $x_{k_0} \to \cdots \to x_{k_m} \to x_{k_0}$ を任意にとると

$$\bar{\lambda}_{k_0} g_{k_0}(x_{k_1}) + \cdots + \bar{\lambda}_{k_m} g_{k_m}(x_{k_0}) = \{\lambda_{k_0} g_{k_0}(x_{k_1}) + \cdots + \lambda_{k_{(m-1)}} g_{k_m}(x_{k_0})\} / (\sum_{h=1}^n \lambda_h)$$

が成り立つが，右辺の分子は非負（（i）の成立を仮定しているので）であり，分母は正なので全体として非負となる。すなわち，条件（ii）が成立することがわかる。 □

従って，合理化可能性を検証するには，どの座標も退化していない比率 $\lambda \gg 0$ により調節した順序構造（グラフ $G = (V, E)$ と重み関数 $c(\lambda) : E \to R$）に閉路が存在しないことがわかればよい。特

に，この系 1 と先の定理 4 により，合理化可能性問題のための Afriat の不等式系は以下のように精緻化されることがわかった：ある $(U_k, \lambda_k)_{k=1}^n$（ただし $\lambda_k > 0$ $(k = 1 \ldots, n)$ かつ $\sum_{k=1}^n \lambda_k = 1$ を満たす）が存在して $U_{k'} \leqq U_k + \lambda_k g_k(x_{k'})$（すべての $k, k' = 1, \ldots, n$ について）が成立する。

3　合理化可能性のグラフ表現とその検証方法

前節までは，合理化可能性条件の 1 つの表現である Afriat の不等式系をもとに，その最短経路問題による特徴づけと，その特徴づけを通じた Afriat 不等式系の精緻化が（線形予算制約の場合と同様に）成り立つことを見た。変数 $\lambda \gg 0$ の取りうる範囲は比率 $(\sum_{k=1}^n \lambda_k = 1)$ に制限してもよいことがわかり，λ の候補は格段に少なくなった。パラメータ λ を 1 つ固定すれば，最短経路問題に対するよく知られた効率的なアルゴリズムによって，グラフ $G = (V, E)$ とコスト関数 $c(\lambda) : E \to R$ に関して負の重みの閉路が存在するかどうかを検証することができる(3)。しかし，依然として無限個の候補 λ が存在するのでこの方法で検証することは実質的に不可能である。そこで本節では，先に定義したグラフの部分グラフ G_{np} を新たに定義することで合理化可能性条件を変数 λ に依存しない形に書き換え，現実的に検証可能な方法を考える。

系 1 で見たように，合理化可能性条件は「調整 λ によって負の閉路が存在しないようにできる」ことにほかならない。そこで，ここでは逆に「合理化不可能」な条件を考えてみると「どのように調整 λ を取っても負の閉路が存在してしまう」となる。そのような「合理化不可能」な状況は，例えば $e := (1, \ldots, 1)$ として有向グラフ $G = (V, E)$ と重み関数 $c(e) : E \to R$ に「非正の辺のみからなる負の閉路」$x_{k_0} \to \cdots \to x_{k_m} \to x_{k_0}$ が存在する場合に生起する。実際，そのような閉路では，どのように調節 $\lambda \gg 0$ を施した重み関数 $c(\lambda) : E \to R$（$c(\lambda)$ の定義は式（4）を参照）のもとでも

$$
\begin{aligned}
c((x_{k_0}, x_{k_1}); \lambda) + \cdots + c((x_{k_m}, x_{k_0}); \lambda) &= \lambda_{k_0} g_{k_0}(x_{k_1}) + \cdots + \lambda_{k_m} g_{k_m}(x_{k_0}) \\
&\leqq \min\{\lambda_{k_i} | i = 1, \ldots, m\}(g_{k_0}(x_{k_1}) + \cdots + g_{k_m}(x_{k_0})) \\
&= \min\{\lambda_{k_i} | i = 1, \ldots, m\}(c((x_{k_0}, x_{k_1}); e) + \cdots + c((x_{k_m}, x_{k_0}); e) < 0
\end{aligned}
$$

が成り立つので，負の閉路のままである。この観察をもとに，「非正の辺のみからなる負の閉路」に注目するために新たなグラフ G_{np} を定義する。

（3）　例えば Moore-Bellman-Ford アルゴリズム（Korte and Vygen, 2008, アルゴリズム 7.2）がよく知られたものであり，このアルゴリズムによると負の閉路が存在するか否かが $O(nm)$ の手間で検証可能である（Korte and Vygen, 2008, 系 7.8。ただし n をグラフの点の数，m をグラフの辺の数とする）。目下の状況では $m := n(n-1)$ なので，変数 $\lambda \gg 0$ を 1 つ固定することで，負の閉路が存在するか否かを $O(n^3)$ の手間で検証できるということになる。

グラフ $G_{np} := (V, E_{np})$ とコスト関数 $\ell_{np} : E_{np} \to R$ を

$$E_{np} := \{(x_k, x_{k'}) \,|\, k, k' = 1, \ldots, n,\ k \neq k',\ \text{かつ}\ c((x_k, x_{k'}); e) \leqq 0\}$$
$$c_{np}((x_k, x_{k'})) := c((x_k, x_{k'}) : e) \quad ((x_k, x_{k'}) \in E_{np}\ \text{に対して}) \tag{5}$$

と定義する。ただし $V = \{x_1, \ldots, x_n\}$ はこれまでと同じ消費量データの点である。すなわち，グラフ G_{np} はグラフ $G = (V, E)$ から重み関数 $c(e) : E \to R$ によって正の重みをもつような辺を除去し，重み c_{np} は $c(e)$ を残った非正の辺に制限したものである（すなわち，$c(e)|_{E_{np}} = c_{np}$ である)。先ほどの観察によると，合理化不可能性の十分条件は「G_{np} に負の重みの閉路が存在する」ことであった。次の定理はこれが合理化不可能性条件を（従って合理化可能性条件を）特徴づけることを示している。

定理 5. n 組のデータ $\{(x_k, B_k)\}_{k=1}^n$ を考える（ただし各 $k = 1, \ldots, n$ に対して $B_k := \{x \in R_+^\ell | g_k(x) \leqq 0\}$ であり $g_k : R_+^\ell \to R$ は単調な連続関数で $g_k(x_k) = 0$ を満たすものとする)。データ $\{(x_k, B_k)\}_{k=1}^n$ に対して以下の 4 条件は同値である。

(i) ある非飽和な連続効用関数 $u : R_+^\ell \to R$ が存在してデータを合理化できる： $g_k(x) \leqq g_k(x_k) \Rightarrow u(x) \leqq u(x_k)$。

(ii) データに対して以下の組合せ的な条件が成り立つ:$g_{k_0}(x_{k_1}) \leqq 0,\ g_{k_1}(x_{k_2}) \leqq 0,\ \ldots,\ g_{k_m}(x_{k_0}) \leqq 0$ ならば $g_{k_i}(x_{k_{(i+1)}}) = 0$ がすべての $i = 0, \ldots, m$ で成立する（ただし $m + 1 = 0$ とする)。

(iii) グラフ $G_{np} = (V, E_{np})$ の閉路は重み関数 $c_{np} : E_{np} \to R$ に関する負の辺を含まない。

(iv) グラフ $G_{np} = (V, E_{np})$ の強連結成分は重み関数 $c_{np} : E_{np} \to R$ に関する負の辺を含まない。(4)

証明. (i) $\Leftrightarrow$ (ii)： 定理 2 の結果である。

(ii) $\Leftrightarrow$ (iii)： 対偶を考えればグラフ G_{np} の定義より明らかである（ただし，定義 (5) において $c((x_k, x_{k'}); e) = g_k(x_{k'})$ であることに注意する)。実際，条件 (ii) の不成立は

$$g_{k_0}(x_{k_1}) \leqq 0,\ g_{k_1}(x_{k_2}) \leqq 0,\ \ldots,\ g_{k_{(m-1)}}(x_{k_m}) \leqq 0$$
$$\text{かつ}\ g_{k_i}(x_{k_{(i+1)}}) \neq 0 \quad (\text{とある}\ i \in \{1, \ldots, m\}\ \text{に対して。ただし}\ m + 1 = 0\ \text{とする})。$$

---

(4) グラフ $G = (V, E)$ の強連結成分とは点の部分集合 $V' \subset V$ で任意の点の 2 点 $v, u \in V'$ に対して，グラフ G に v-u 有向路と u-v 有向路が存在し，この条件について極大になるものである（すなわち「グラフ G 内で行き来できる」ような点の集まりのうち極大になるもののことである)。本稿では「強連結成分 V'」と「V' が誘導する G の部分グラフ」とをどちらも強連結成分と呼ぶことがある（ただし V' が誘導する G の部分グラフとは，グラフ $G' := (V', \{e \,|\, e = (v, u) \in E\ \text{かつ}\ v, u \in V'\})$，すなわち G のうち「V' で閉じている辺」のみに注目する部分グラフのことである)。

となるが，この条件の前半は「G_{np} の閉路が存在する」にほかならず，後半は「負の辺を含む」にほかならない。

（iii）$\Leftrightarrow$（iv）： 対偶を考える。定義より，任意の閉路はとある強連結成分に含まれるので「(iii) の否定 $\Rightarrow$（iv）の否定」が成り立つ。逆に，「(iv) の否定」が成り立つとする。負の辺 $(x_k, x_{k'})$ に対して $x_k, x_{k'}$ は同じ強連結成分に含まれるので，有向路 $x_{k'} \to x_{k_0} \to \cdots \to x_{k_m} \to x_k$ が存在する。このとき閉路 $x_k \to x_{k'} \to x_{k_0} \to \cdots \to x_{k_m} \to x_k$ は負の辺 $(x_k, x_{k'})$ を含む。よって「(iii) の否定」が成立する。 □

この定理によって，合理化可能性条件は変数 λ を含まない形で，グラフ理論的な条件に書き換えられた（証明からわかるように，これは巡回整合性条件のグラフ表現にほかならない）。そして，この結果の条件（iv）から，Fujishige and Yang（2012）および Nobibon et al.（2014）が線形予算制約のために提案した効率的な検証手順が，一般的な予算制約のもとでも実現できる。実際，SCCD アルゴリズム（strongly connected component decomposition algorithm; 強連結成分分解アルゴリズム）を用いることで直接的に条件（iv）の成立または不成立を検証することができる（Korte and Vygen, 2008, アルゴリズム 2.2）。すなわち，G_{np} に SCCD アルゴリズムを適用することで G_{np} の強連結成分分解を得る。そして，それらのうちで負の辺を含むものがあるかどうかを調べればよい。SCCD アルゴリズムの計算の手間は $O(n+m)$（ただし n はグラフの点の個数であり m は辺の個数）である（Korte and Vygen, 2008, 定理 2.19）。ここでは，G_{np} の辺は最大でも $n(n-1)$ 本なので SCCD アルゴリズムの手間は $O(n+(n(n-1))=O(n^2)$ で抑えられる。これは Varian（1982）の提案した「グラフの推移閉包をとる」ことによる検証手順がもつ計算の手間 $O(n^3)$ よりも効率的である。

4 「当てはまりの良さ基準」と NP 困難性

これまでに見た合理化可能性の検証は，データが与えられたときに以下の条件（上に見た様々な同値条件の 1 つ；定理 5 参照）

「グラフ $G_{np} = (V, E_{np})$ の閉路は $c_{np} : E_{np} \to R$ に関する負の辺を含まない」

を満たすか否かを問う形をしている。すなわち「合理化可能／不可能」のどちらであるかを二者択一的にテストするものになっている。そこで，例えば我々が合理化不可能なデータに直面したときに，そのデータが「どの程度合理化不可能か」を測るための基準を考えることに意味が出てくる。逆の言い方ではあるがこのような基準を「当てはまりの良さ基準（goodness of fit measure）」と呼ぶ。これまで様々な当てはまりの良さ基準が提案されている。代表的な基準として Afriat（1973），Houtman and Maks（1985），Varian（1990），Echenique et al.（2011），Smeulders et al.（2013）， Dean and Martin（2010, 2015）などの提案するものが挙げられる。これらは最適化問題として定式化されるものであるが，そのうちの多くがNP困難と呼ばれる「現状では多項式時間アルゴリズムが存在しそうもない」最適化問題のクラスに属するということがSmeulders et al.（2013, 2014）により示されている。

本節では，これらのうちの1つであるHoutman and Maks（1985）の提案する基準であるHM指数の計算がNP困難なクラスの最適化問題であることの証明を与える。これはSmeulders et al.（2014）によって示されている結果ではあるが，これまでに見たグラフ構造を道具として考えることにより，より直接的な証明を与えられることがわかる[(5)]。また，Dean and Martin（2010, 2015）らの提案するMC指数（minimum cost index; 最小費用指数）にも注目し，彼らが定義するところの指数を計算する問題もまたNP困難なクラスの最適化問題であることを見る[(6)]。なお，以下では線形予算制約の場合のみを考える。

HM指数

Houtman and Maks（1985）によるHM指数は「データのもつ情報の一部を無視することで合理化可能条件（の1つの形式）を満たすようになるような部分情報のうちで最大のもの」という概念として定義される。まず「データのもつ情報の一部を無視することで合理化可能条件（の1つの形式）を満たす」という概念を定義する。0-1ベクトル $v \in \{0,1\}^n$ を所与として以下の条件を考える（Houtman and Maks (1985)，Dean and Martin (2010, 2015)，および Heufer and Hjertstrand (2015) を参照）：

$$p_{k_0} \cdot (x_{k_1} - v_{k_0} x_{k_0}) \leqq 0,\ p_{k_1} \cdot (x_{k_2} - v_{k_1} x_{k_1}) \leqq 0,\ \ldots,\ p_{k_m} \cdot (x_{k_0} - v_{k_m} x_{k_m}) \leqq 0$$

ならば $p_{k_i} \cdot (x_{k_{(i+1)}} - v_{k_i} x_{k_i}) = 0$ がすべての $i = 0, \ldots, m$ で成立する（ただし $m+1=0$ とする）。この条件を v-巡回的整合性条件と呼ぶ[(7)]。そしてHM指数を

$$\text{HM 指数} := \max\{\sum_{k=1}^{n} v_k \,|\, v \in \{0,1\}^n \quad \text{かつ，} v\text{-巡回的整合性条件が成立}\ \} \qquad (6)$$

によって定義する[(8)]。これが先述の「データのもつ情報の一部を無視することで合理化可能性条件（の1つの形式）を満たすようになるような部分情報のうちで最大のもの」を表す数式にほかならない。

（5） Smeulders et al.（2014）ではこの結果を，HM指数の計算問題の決定問題（decision problem）版を考え，NP完全なクラスに属する別の古典的な決定問題の1つである安定集合問題（stable set problem）に多項式時間還元することで示している。本稿ではHM指数の計算問題そのものを，よく知られたNP困難な最適化問題であるMFVS問題（minimum feedback vertex set problem; 最小フィードバック頂点集合問題）に多項式時間還元する形で証明する。また，この方法による証明が可能であることはすでにHoutman and Maks（1985）の議論から示唆されている。

（6） こちらは，別のよく知られたNP困難な最適化問題であるMAS問題（maximum acyclic subgraph problem; 最大重み非閉路的部分グラフ問題）に多項式時間還元する形で示す。

ここで，HM 指数の計算問題を定式化する。

HM 指数の計算

インスタンス： 財の価格と消費量のデータ $\{(p_k, x_k)\}_{k=1}^n$。ただし $p_k \in R_{++}^{\ell}$, $x_k \in R_{+}^{\ell}$ $(k = 1, \ldots, n)$。

タスク： 式（6）によって定義される HM 指数を求める。

ここで，v-巡回的整合性条件の定義についての観察をする。 $p_k \in R_{++}^{\ell}, x_k \in R_{+}^{\ell}$ であるから，不等式 $p_k \cdot (x_{k'} - v_k x_k) \leqq 0$ は $v_k = 1$ のときにしか成立しない不等式である。すなわち，0-1 ベクトル $v \in \{0,1\}^n$ を所与とすると，任意の $k, k' = 1, \ldots, n$ に対して「$p_k \cdot (x_{k'} - v_k x_k) \leqq 0 \Leftrightarrow v_k = 1$ かつ $p_k \cdot (x_{k'} - x_k) \leqq 0$」にほかならない。 また，同様の同値関係が等式 $p_k \cdot (x_{k'} - v_k x_k) = 0$ についても成立する。この観察をもとに，0-1 ベクトル $v \in \{0,1\}^n$ を所与として，グラフ $G_{np}(v) := (V, E_{np}(v))$ を

$$E_{np}(v) := \{(x_k, x_{k'}) \mid k \neq k', v_k = 1, \text{ かつ } p_k \cdot (x_{k'} - x_k) \leqq 0\}$$

によって定義する。すなわち $G_{np}(v)$ は「G_{np} から $v_k = 0$ に対応する点 x_k から出るすべての辺 $(x_k, x_{k'})$ を取り除いたもの」にほかならない。この定義より，以下の命題が成立することがわかる（この命題は，対偶を考えれば上の注意と定義から明らかであるが，完全のため補遺にその証明を与える）。

命題 1. n 組のデータ $\{(p_k, x_k)\}_{k=1}^n$ と 0-1 ベクトル $v \in \{0,1\}^n$ に対して以下の 3 条件は同値である。

（i） v-巡回的整合性条件が成立する。

（ii） $G_{np}(v)$ の閉路は $p_k \cdot (x_{k'} - x_k) < 0$ なる組合せに対応する辺 $(x_k, x_{k'})$ を含まない。

（iii） G_{np} の $p_k \cdot (x_{k'} - x_k) < 0$ なる組合せに対応する辺 $(x_k, x_{k'})$ を含むような閉路は，$v_k = 0$ に対応する点 x_k を少なくとも 1 つ含む。

この命題により，HM 指数の定義（6）に対して以下の等式

（7） Houtman and Maks（1985）や Heufer and Hjertstrand（2015）では，Varian（1982）によって定義された GARP を用いて HM 指数を定義している。すなわち上の v-巡回的整合性条件と同値な v-GARP 条件を用いてこの「データのもつ情報の一部を無視することで合理化可能条件（の 1 つの形式）を満たす」という概念を定義している。本稿ではこれまでの議論との一貫性をもたせるために v-巡回的整合性条件を用いて議論する。

（8） Houtman and Maks（1985）および Heufer and Hjertstrand（2015）では，ここに定義した HM 指数をデータ数 n で割って規格化している。ここではその計算の NP 困難性を見ることが目的であるので，この規格化定数を無視して議論する。

$$
\begin{aligned}
\text{HM 指数} &= \max\{\sum_{k=1}^{n} v_k \,|\, v \in \{0,1\}^n \ \text{かつ，}v\text{-巡回的整合性条件が成立}\ \} \\
&= \max\{\sum_{k=1}^{n} v_k \,|\, v \in \{0,1\}^n \ \text{かつ，命題 1 の条件（iii）が成立}\ \} \\
&= n - \min\{\sum_{k=1}^{n} (1-v_k) \,|\, v \in \{0,1\}^n \ \text{かつ，命題 1 の条件（iii）が成立}\ \} \qquad (7)
\end{aligned}
$$

が成立する。ただし，最右辺に現れる最小化の目的 $\sum_{k=1}^{n}(1-v_k)$ は「$v_k = 0$ なる $k \in \{1,\ldots,n\}$ の個数」にほかならないことに注意する。

さて，等式（7）の最右辺第二項の最適化問題に注目すると「有向グラフ G_{np} が与えられたときにそのグラフに含まれる（ある条件を満たす）任意の閉路が少なくとも 1 点を含むような点の部分集合 $V' \subset V$ のうち要素数 $|V'|$ が最小になるものを求める」という形式の問題になっていることがわかる。Houtman and Maks（1985）および Dean and Martin（2010, 2015）にも指摘されているが，これは以下に示す MFVS 問題のもつ形式とよく似ていることがわかる。

MFVS

インスタンス： 有向グラフ $G = (U, E)$（自己ループや並列辺を含まない単純グラフとする）。

タスク： G の任意の有効閉路が少なくとも 1 点を含むような点の部分集合 U' のうち，要素数 $|U'|$ が最小のものを求める。

MFVS 問題は NP 困難な問題と知られている（Karp (1972)，Garey and Johnson（1979））。HM 指数の計算問題の NP 困難性は等式（7）を通して，上記の MFVS 問題からの多項式時間還元により示される。

定理 6. HM 指数の計算問題は NP 困難である。

証明. MFVS 問題のインスタンスである有向グラフ $G = (U, E)$ に対して，HM 指数の計算問題のインスタンス（価格と消費量のデータ組）$(p_k, x_k) \in R^n_{++} \times R^n_{+} (k = 1, \ldots, n)$ を構成する[(9)]（ただし $n := |V|$ である）。まず，消費量のデータ $x_k \in R^n_+$ $(k = 1, \ldots, n)$ を $x_k = (x_{k1}, \ldots, x_{k(k-1)}, x_{kk}, x_{k(k+1)}, \ldots, x_{kn}) := (0, \ldots, 0, 1, 0, \ldots, 0)$ によって定義する。次に，価格データ $p_k \in R^n_{++}$ を

$$
p_{kk'} = \begin{cases} 1 & (u_k, u_{k'}) \in E \text{ のとき} \\ 2 & k = k' \text{ のとき} \\ 3 & (u_k, u_{k'}) \notin E \text{ のとき} \end{cases}
$$

（9） このインスタンス構成のアイディアは Smeulders et al.（2014）の用いたものと本質的に同じである。

によって定義する。このとき，任意の $k, k' = 1, \ldots, n$ $(k \neq k')$ に対して

$$p_k \cdot (x_{k'} - x_k) = p_{kk'} - p_{kk} = \begin{cases} -1 & (u_k, u_{k'}) \in E \text{ のとき} \\ 1 & (u_k, u_{k'}) \notin E \text{ のとき} \end{cases}$$

となるので，(5) によって G_{np} を定義すると「$(u_k, u_{k'}) \in E \Leftrightarrow (x_k, x'_k) \in E_{np}$」が成立する。また $p_k \cdot (x_{k'} - x_k) = 0$ なる辺 $(x_k, x_{k'}) \in E_{np}$ は存在しないので，G_{np} の「$p_k \cdot (x_{k'} - x_k) < 0$ なる組合せに対応する辺 $(x_k, x_{k'})$ を含むような閉路」全体の成す集合は G_{np} の有向閉路全体の集合にほかならない。従って，0-1 ベクトル $v \in \{0,1\}^n$ に対して，同値関係「命題 1 の条件 (iii) が成立 $\Leftrightarrow$ G_{np} の任意の有向閉路は $v_k = 0$ に対応する点 x_k を少なくとも 1 つ含む $\Leftrightarrow$ 有向グラフ G の任意の有向閉路は $v_k = 0$ に対応する点 $u_k \in U$ を含む」が成立する。従って等式 (7) とあわせて，

$$\begin{aligned} \text{HM 指数} &= \max\{\sum_{k=1}^{n} v_k \,|\, v \in \{0,1\}^n \text{ かつ，命題 1 の条件 (iii) が成立}\} \\ &= n - \min\{\sum_{k=1}^{n} (1 - v_k) \,|\, v \in \{0,1\}^n \text{ かつ，命題 1 の条件 (iii) が成立}\} \\ &= n - \min\{\sum_{k=1}^{n} (1 - v_k) \,|\, v \in \{0,1\}^n \text{ かつ，} G \text{ の任意の有向閉路は } v_k = 0 \text{ に対応する点} \\ &\quad u_k \in U \text{ を含む}\} \end{aligned}$$

が成立する。従って，上に構成したデータ $\{(p_k, x_k)\}_{k=1}^{n}$ の HM 指数を用いれば，MFVS 問題の最適解を計算することができることがわかる。また，上記のインスタンス構成の手間はいずれも明らかに多項式時間で抑えられる。よって HM 指数を計算する定数時間アルゴリズムを仮定することで，MFVS 問題の多項式時間アルゴリズムが得られる。すなわち，HM 指数の計算問題は NP 困難であることがわかった。 □

Dean and Martin (2010, 2015) は HM 指数を最小集合被覆問題 (minimum set covering problem) に帰着することで計算できるという議論をしている[(10)]。また，Smeulders et al. (2014) では，定理 6 の事実を安定集合問題 (stable set problem) と呼ばれる NP 完全な決定問題に還元することで示している。彼らの証明は以下の構成をしている。まず，HM 指数の計算に関連した決定問題「HI-GARP」を定式化し，それを安定集合問題に多項式時間還元することで，HM 指数計算の決定問題版である「HI-GARP」の NP 困難性を示すというものである。本稿では HM 指数の計算問題そのものを NP 困難な最適化問題に多項式時間還元することで，その NP 困難性を示した。

(10) 上の結果から予想されることであるが，彼らは多項式時間帰着アルゴリズムを構成することには成功していない。実際，彼らの議論した手順には「有向グラフの単純閉路の列挙」をするアルゴリズムが含まれているが，これは入力するデータ数（グラフの点数および辺数）に対する多項式時間アルゴリズムではない（Johnson (1975)）。

MC 指数

Dean and Martin（2010, 2015）らの提案する「当てはまりの良さ基準」である MC 指数は以下のように定義される。[(11)]

$$\text{MC 指数} := \min\{ \sum_{(x_k, x_{k'}) \in E'} p_k \cdot (x_k - x_{k'}) \mid E' \subset E_{np} \ \text{かつ，} (V, E_{np} \setminus E') \text{ は有向閉路を含まない} \} \tag{8}$$

この定義は「データのもつ情報の一部を無視することで合理化可能性条件（の 1 つの形式）を満たすようになるような部分情報のうちで最大のもの」というアイディアにもとづいていることが見て取れる。実際，等式

$$\begin{aligned} \text{MC 指数} :&= \min\{ \sum_{(x_k, x_{k'}) \in E'} p_k \cdot (x_k - x_{k'}) \mid E' \subset E_{np} \ \text{かつ，} (V, E_{np} \setminus E') \text{ は有向閉路を} \\ &\quad \text{含まない} \} \\ &= \min\{ \sum_{(x_k, x_{k'}) \in E'} p_k \cdot (x_k - x_{k'}) \mid E' \subset E_{np} \ \text{かつ，} (V, E_{np} \setminus E') \text{ の有向閉路は} \\ &\quad \text{負の辺を含まない} \} \end{aligned}$$

が成り立つことがわかるので，MC 指数は「G_{np} のいくつかの辺を無視すれば合理化可能性条件の 1 つの形式（定理 5 の条件（iii）を参照）を満たすようにできるような辺の部分集合のなかで，その重みが $p_k \cdot (x_k - x_{k'}) \geqq 0$ による評価で最小になるもの」の値にほかならないことがわかり，「当てはまりの良さ基準」としての一定の合理性があることが認められる。

MC 指数と HM 指数との違いは「データのもつ情報をどのように無視するか」の決め方であり，HM 指数は「G_{np} において消費点 x_k から出るすべての枝 $(x_k, x_{k'})$」を 1 つの単位として無視していた（命題 1 の直前の議論を参照）のに対して，MC 指数では「G_{np} において消費点 x_k から出るある枝 $(x_k, x_{k'})$」を 1 つの単位として無視する問題になっていることがその定義からわかる。また，HM 指数では無視するデータの重みを 1 単位あたりすべて一定（すべて 1）としていたのに対して，MC 指数では無視するデータの重みを 1 単位あたり $p_k \cdot (x_k - x_{k'}) \geqq 0$ として評価している点も異なる。MC 指数の計算問題は以下のように定式化される。

（11） Dean and Martin（2010, 2015）では，以下に定義する値を定数 $\sum_{k=1}^{n} p_k \cdot x_k > 0$ で割ることで規格化しているが，ここではこの指数の計算問題が NP 困難であることを見るのが目的であるため，この規格化定数を無視している。また Dean and Martin（2010, 2015）では，Varian (1982) の方法（合理化可能性問題の GARP に基づく定式化）に従って二項関係 R_0 を $x_i R_0 x_j :\Leftrightarrow p_i \cdot (x_j - x_i) \leqq 0$ と定義し，MC 指数もこれを用いて定義しているが，G_{np} がこの二項関係の表現になっていることは定義から明らかであるので本稿ではグラフ表現 G_{np} を用いて MC 指数を定義する。

MC 指数の計算

インスタンス： 財の価格と消費量のデータ $\{(p_k, x_k)\}_{k=1}^n$。ただし $p_k \in R^\ell_{++}$, $x_k \in R^\ell_+$ $(k = 1, \dots, n)$。

タスク： 式（8）によって定義される MC 指数を求める。

以下で MC 指数の計算問題の NP 困難性を見る。まず，等式

$$
\begin{aligned}
\text{MC 指数} &= \min\{ \sum_{(x_k, x_{k'}) \in E'} p_k \cdot (x_k - x_{k'}) \mid E' \subset E_{np} \text{ かつ，} (V, E_{np} \setminus E') \text{ は有向閉路を含まない} \} \\
&= -\max\{ -(\sum_{(x_k, x_{k'}) \in E'} p_k \cdot (x_k - x_{k'})) \mid E' \subset E_{np} \text{ かつ，} (V, E_{np} \setminus E') \text{ は有向閉路を} \\
&\qquad \text{含まない} \}
\end{aligned}
$$

に注意すると，以下の等式

$$
\begin{aligned}
&\sum_{(x_k, x_{k'}) \in E_{np}} p_k \cdot (x_k - x_{k'}) - \text{MC 指数} \\
&= \max\{ \sum_{(x_k, x_{k'}) \in E_{np}} p_k \cdot (x_k - x_{k'}) - \sum_{(x_k, x_{k'}) \in E'} p_k \cdot (x_k - x_{k'}) \mid E' \subset E_{np} \text{ かつ，} \\
&\quad (V, E_{np} \setminus E') \text{ は有向閉路を含まない} \} \\
&= \max\{ \sum_{(x_k, x_{k'}) \in E_{np} \setminus E'} p_k \cdot (x_k - x_{k'}) \mid E' \subset E_{np} \text{ かつ，} (V, E_{np} \setminus E') \text{ は有向閉路を} \\
&\quad \text{含まない} \}
\end{aligned}
\tag{9}
$$

が成立することがわかる。この等式の最右辺は「データ $\{(p_k, x_k)\}_{k=1}^n$ に対応するグラフ G_{np} が与えられたとき，その辺の部分集合を消去することで，閉路を含まず残された重みを最大にする部分グラフを求める」という形式を有している。これは，以下のように定式化される MAS 問題と呼ばれる問題のもつ形式によく似ていることが観察できる。

MAS

インスタンス： 有向グラフ $G = (U, E)$ と正値の重み関数 $c : E \to R_+$（自己ループや並列辺を含まない単純グラフとする）。

タスク： $G' = (U, E')$ が非閉路的（有向閉路を含まない）となるような $E' \subset E$ のうちで，重み $\sum_{e \in E'} c(e)$ を最大にするものを求める。

MAS 問題は NP 困難な問題と知られている（Karp (1972)，Garey and Johnson (1979)）。そして，MC 指数の計算問題の NP 困難性は MAS 問題からの多項式時間還元により示される。

定理 7.　MC 指数の計算は NP 困難である。

証明.　MAS 問題のインスタンス $G = (U, E)$ と正値の重み関数 $c : E \to R_+$ に対して，MC 指数の計算問題のインスタンス（価格と消費量のデータ組）$(p_k, x_k) \in R^n_{++} \times R^n_+ (k = 1, \ldots, n)$ を構成する（ただし $n := |V|$ とする）。まず，消費量のデータ $x_k \in R^n_+$（$k = 1, \ldots, n$）を（HM 指数の場合と同じく）$x_k = (x_{k1}, \ldots, x_{k(k-1)}, x_{kk}, x_{k(k+1)}, \ldots, x_{kn}) := (0, \ldots, 0, 1, 0, \ldots, 0)$ によって定義する。次に，価格データ $p_k \in R^n_{++}$ を，定数 $a := \max\{c(e)|e \in E\} + 1$ を用いて，

$$p_{kk'} = \begin{cases} a - c((u_k, u_{'k})) & (u_k, u_{k'}) \in E \text{ のとき} \\ 2a & k = k' \text{ のとき} \\ 3a - c((u_k, u'_k)) & (u_k, u_{k'}) \notin E \text{ のとき} \end{cases}$$

によって定義する。このとき，任意の $k, k' = 1, \ldots, n$ $(k \neq k')$ に対して

$$p_k \cdot (x_{k'} - x_k) = p_{kk'} - p_{kk} = \begin{cases} -a - c((u_k, u_{k'})) & (u_k, u_{k'}) \in E \text{ のとき} \\ a - c((u_k, u_{k'})) & (u_k, u_{k'}) \notin E \text{ のとき} \end{cases} \tag{10}$$

となるので，(5) によって G_{np} を定義すると「 $(u_k, u_{k'}) \in E \Leftrightarrow (x_k, x'_k) \in E_{np}$ 」が成立する。この同値関係と等式 (9) より

$$\begin{aligned} &\sum_{(x_k, x_{k'}) \in E_{np}} p_k \cdot (x_k - x_{k'}) - \text{MC 指数} \\ &= \max\{ \sum_{(x_k, x_{k'}) \in E_{np} \setminus E'} p_k \cdot (x_k - x_{k'}) \mid E' \subset E_{np} \text{ かつ，} (V, E_{np} \setminus E') \text{ は有向閉路を含まない} \} \\ &= \max\{ \sum_{(u_k, u_{k'}) \in E \setminus E'} a + c((u_k, u_{k'})) \mid E' \subset E \text{ かつ，} (U, E \setminus E') \text{ は有向閉路を含まない} \} \\ &= a + \max\{ \sum_{(u_k, u_{k'}) \in E \setminus E'} c((u_k, u_{k'})) \mid E' \subset E \text{ かつ，} (U, E \setminus E') \text{ は有向閉路を含まない} \} \end{aligned} \tag{11}$$

という等式が成立する。また，等式 (10) と同値関係「$(u_k, u_{k'}) \in E \Leftrightarrow (x_k, x'_k) \in E_{np}$」より $\sum_{(x_k, x_{k'}) \in E_{np}} p_k \cdot (x_k - x_{k'}) = \sum_{e \in E} c(e) + a|E|$ であり，$a := \max\{c(e)|e \in E\} + 1$ であることを，等式 (11) と合わせると

$$\begin{aligned} &\sum_{e \in E} c(e) + (\max\{c(e)|e \in E\} + 1)(|E| - 1) - \text{MC 指数} \\ &= \max\{ \sum_{(u_k, u_{k'}) \in E_{np} \setminus E'} c((u_k, u_{k'})) \mid E' \subset E \text{ s.t } (V, E \setminus E') \text{ は有向閉路を含まない} \} \end{aligned}$$

が成り立つ。従って，上に構成したデータ $\{(p_k, x_k)\}_{k=1}^n$ の MC 指数を用いれば，MAS の最適解を計算することができる。以上の手続きの手間が多項式時間で抑えられることは明らかであるので，MC 指数の計算は NP 困難なクラスに属する問題であることがわかった。 □

5　多項式時間アルゴリズムをもつ「当てはまりの良さ基準」

前節で見たように，HM 指数と MC 指数の計算は NP 困難な最適化問題のクラスに属する。Smeulders et al.（2013, 2014）は，そのほかにも Varian（1990）や Echenique et al.（2011）の提案する指数の計算が NP 困難な問題であることを示している。従って，任意のデータ $\{(p_k, x_k)\}_{k=1}^n$ に対してこれらの指数を計算するような多項式時間の厳密アルゴリズムの存在は，（どの NP 完全問題にも多項式時間の厳密アルゴリズムが提案されていないという現状では）期待できそうもないということになる。

一方で Smeulders et al.（2014）では Afriat（1973）の指数が多項式時間アルゴリズムで計算できることを示し，さらに Smeulders et al.（2013）では Echenique et al.（2011）の提案する基準と似たアイディアによる新しい基準と，その指数の計算に対する多項式時間アルゴリズムを与えている。本節では，これらの指数と同様に多項式時間アルゴリズムで計算可能であり，しかもこれまでに見てきた合理化可能性問題のグラフ理論的な構造に照らして自然な解釈をもつような「当てはまりの良さ基準」を新たに提案する。

これまでの議論で見てきたように，データが合理化不可能であるとき，グラフ G_{np} は負の重みをもつ閉路をもつ。従って G_{np} は負の重みをもつ強連結成分をもつ。そして，任意の閉路はいずれかの強連結成分に含まれるので，データのもつ情報のうち，合理化不可能性に関わる部分はこれらの強連結成分にのみ含まれていることがわかる（定理 5 を参照）。従って，上記に挙げた様々な基準に加えて，「当てはまりの良さ基準」の候補として以下の SCC 指数（strongly connected component index; 強連結成分指数）が自然に考えられる：

$$\text{SCC 指数} := \frac{\displaystyle\sum_{h=1}^{m}\left(\sum_{(x_k, x_{k'})\in C_h} p_k\cdot(x_k - x_{k'})\right)}{\displaystyle\sum_{(x_k, x_{k'})\in E_{np}} p_k\cdot(x_k - x_{k'})}。\tag{12}$$

ただし $C_h (h = 1, \ldots, m)$ は G_{np} の各強連結成分で，「$(x_k, x_{k'}) \in C_h$」は $(x_k, x_{k'})$ が C_h の辺に含まれることを形式的に表すものとする。また，SCC 指数の定義式 (12) は分母が $\sum_{(x_k, x_{k'})\in E_{np}} p_k \cdot (x_k - x_{k'}) = 0$ のとき意味をもたないが，これはデータの合理化可能性を含意するため，そのようなデータに対しては SCC 指数 $:= 0$ と定義する。

この指数は「データがもしも合理化可能であるときある種の順序構造を保持することになる G_{np} に注目し，その総重みのうちで合理化不可能部分の重みが占める割合」を測るものであり，「当てはまりの良さ基準」として極めて自然な解釈をもつ。また，必ず $[0,1]$ 値をとるためデータ間での評価が安易であるという利点も期待できる。そして，明らかにこの指数には多項式時間アルゴリズムが存在する。

定理 8.　SCC 指数を計算するための $O(n^2)$ アルゴリズムが存在する。

これは，第 3 節でも見た強連結成分分解に対するよく知られたアルゴリズム（SCCD アルゴリズム）の存在より直ちに導かれる結果であり，この指数が現実的な量のデータに対して十分な早さで計算可能なことを示している。

補遺

命題 1 の証明。命題 1 の対偶版を証明する。対偶版は以下のような主張となる。

命題 1′.　n 組のデータ $\{(p_k, x_k)\}_{k=1}^{n}$ と 0-1 ベクトル $v \in \{0,1\}^n$ に対して以下の 3 条件は同値である。

（i）　v-巡回的整合性条件が成立しない。
（ii）　$G_{np}(v)$ のある閉路は $p_k \cdot (x_{k'} - x_k) < 0$ なる組合せに対応する辺 $(x_k, x_{k'})$ を含む。
（iii）　G_{np} のある閉路は $p_k \cdot (x_{k'} - x_k) < 0$ なる組合せに対応する辺 $(x_k, x_{k'})$ を含み，$v_k = 1$ に対応する点 x_k のみからなる。

証明.　（i）$\Longleftrightarrow$（ii）：条件（i）（v-巡回的整合性条件の不成立）は，

$$\begin{aligned}&p_{k_0} \cdot (x_{k_1} - v_{k_0} x_{k_0}) \leqq 0,\ p_{k_1} \cdot (x_{k_2} - v_{k_1} x_{k_1}) \leqq 0,\ \ldots,\ p_{k_m} \cdot (x_{k_0} - v_{k_m} x_{k_m}) \leqq 0\\ &\text{かつ}\ \ p_{k_i} \cdot (x_{k_{(i+1)}} - v_{k_i} x_{k_i}) < 0\ \ ({}^{\exists} i \in \{0, \ldots, m\},\ \text{ただし}\ m+1=0\ \text{とする}) \qquad (13)\end{aligned}$$

である。命題 1 の直前で注意したように，$p_k \in R^{\ell}_{++}$ かつ $x_k \in R^{\ell}_{+}$ なので，任意の $k, k' = 1, \ldots,$ n に対して

$$p_k \cdot (x_{k'} - v_k x_k) \leqq 0 \Leftrightarrow v_k = 1\ \text{かつ}\ p_k \cdot (x_{k'} - x_k) \leqq 0$$

$$p_k \cdot (x_{k'} - v_k x_k) < 0 \Leftrightarrow v_k = 1\ \text{かつ}\ p_k \cdot (x_{k'} - x_k) < 0$$

が成り立つ。よって，v-巡回的整合性条件の不成立（13）は，

$$p_{k_0} \cdot (x_{k_1} - x_{k_0}) \leqq 0,\ p_{k_1} \cdot (x_{k_2} - x_{k_1}) \leqq 0,\ \ldots,\ p_{k_m} \cdot (x_{k_0} - x_{k_m}) \leqq 0,$$

$$v_{k_i} = 1 \ \ ({}^{\forall} i \in \{0,\ldots,m\}), \text{ かつ } \ p_{k_i} \cdot (x_{k_{(i+1)}} - x_{k_i}) < 0$$
$$({}^{\exists} i \in \{0,\ldots,m\}, \text{ ただし } m+1=0 \text{ とする})$$

にほかならず，$G_{np}(v)$ の定義より，これは条件（ii）の成立にほかならない。

(ii) $\Longleftrightarrow$ (iii)：これは $G_{np}(v)$ の辺 $E_{np}(v)$ の定義 $E_{np}(v) := \{(x_k, x_{k'}) \mid k \neq k', v_k = 1, \text{ かつ } p_k \cdot (x_{k'} - x_k) \leqq 0\}$ と，G_{np} の辺 E_{np} の定義 $E_{np} := \{(x_k, x_{k'}) \mid k \neq k', \text{ かつ } p_k \cdot (x_{k'} - x_k) \leqq 0\}$ とを比較することにより明らかである。 □

参 考 文 献

[1] Afriat, S.N., 1967. The construction of utility functions from expenditure data. International Economic Review 8, 1, 67–77.

[2] Afriat, S.N., 1973. On a system of inequalities in demand analysis: An extension of the classical method. International Economic Review 14, 2, 460–472.

[3] Dean, M., Martin, D., 2010. How rational are your choice data? In Proceedings of the Conference on Revealed Preference and Partial Identification.

[4] Dean, M., Martin, D., 2015. Measuring rationality with the minimum cost of revealed preference violations. Review of Economics and Statistics (forthcoming).

[5] Debreu, G., 1954. Representation of a preference ordering by a numerical function. In Thrall, R.M., Coombs, C.H., Davis, R.L., eds. Decision Processes. Wiley, New York.

[6] Diewert, W.E., 1973. Afriat and revealed preference theory. The Review of Economic Studies 40, 3, 419–425.

[7] Echenique, F., Lee, S., Shum, M., 2011. The money pump as a measure of revealed preference violations. Journal of Political Economy 119, 6, 1201–1223.

[8] Forges, F., Minelli, E., 2009. Afriat's theorem for general budget sets. Journal of Economic Theory 144, 1, 135–145.

[9] Fostel, A., Scarf, H.E., Todd, M.J., 2004. Two new proofs of afriat's theorem. Economic Theory 24, 1, 211–219.

[10] Fujishige, S., Yang, Z., 2012. On revealed preference and indivisibilities. Modern Economy 3, 752–758.

[11] Garey, M.R., Johnson, D.S., 1979. Computers and intractability: A guide to the theory of NP-completeness. W. H. Freeman and Company, San Francisco.

[12] Geanakoplos, J.D., 2013. Afriat from maxmin. Economic Theory 54, 3, 443–448.

[13] Heufer, J., Hjertstrand, P., 2015. Consistent subsets: Computationally feasible methods to compute the Houtman-Maks-index. Economics Letters 128, 87–89.

[14] Houtman M., Maks, J., 1985. Determining all maximal data subsets consistent with revealed preference. Kwantitatieve Methoden 19, 89–104.

[15] Johnson, D.B., 1975. Finding all the elementary circuits of a directed graph. SIAM Journal on Computing 4, 1, 77–84.

[16] Karp, R.M., 1972. Reducibility among combinatorial problems. Complexity of Computer Computations 40, 4, 85–103.

[17] Kolesnikov, A.V., Kudryavtseva, O.V., Nagapetyan T., 2013. Remarks on Afriat's theorem and the Monge-Kantorovich problem. Journal of Mathematical Economics 49, 6, 501–505.
[18] Korte, B., Vygen, J., 2008. Combinatorial optimization: Theory and algorithms. 4th ed., Springer, Berlin.
[19] 室田一雄，塩浦昭義，2013. 離散凸解析と最適化アルゴリズム。朝倉書店。
[20] Nobibon, F.T., Smeulders, B., Spieksma, F.C.R., 2015. A note on testing axioms of revealed preference. Journal of Optimization Theory and Applications 166, 3, 1063–1070.
[21] Piaw T.C., Vhora, R.V., 2003. Afriat's theorem and negative cycles. mimeo.
[22] Smeulders, B., Cherchye, L., Spieksma, F.C.R., De Rock, B., 2013. The money pump as a measure of revealed preference violations: A comment. Journal of Political Economy 121, 6, 1248–1258.
[23] Smeulders, B., Spieksma, F.C.R., Cherchye, L., De Rock, B., 2014. Goodness-of-fit measures for revealed preference tests: Complexity results and algorithms. ACM Transactions on Economics and Computation Archive 2, 1, 3.
[24] Varian, H.R., 1982. The nonparametric approach to demand analysis. Econometrica 50, 4, 945–973.
[25] Varian, H.R., 1983. Non-parametric tests of consumer behavior. The Review of Economic Studies 50, 1, 99–110.
[26] Varian, H.R., 1990. Goodness-of-fit in optimizing models. Journal of Econometrics 46, 1, 125–140.

要旨: 本稿では，消費データの合理化可能性問題の背後に存在するネットワーク構造やグラフ構造を取り扱う。これにより問題の構造が明確になり，合理化可能性を検証するための効率的なアルゴリズムを得ることができる。また，データがどの程度効用最大化問題に当てはまるかを測る指数の計算が，NP 困難と呼ばれるクラスの最適化問題に属することが示される。そして，新しい基準として上述のネットワーク構造に照らして自然で，安易に計算可能なものが提案される。

キーワード: 顕示選好理論, 合理化可能性, 巡回的整合性（cyclical consistency）, GARP（generalized axiom of revealed preference; 顕示選好の一般化公理），最短経路問題