

論文

SINGLE 法による勤労者世帯支出構造の変化点の検出*

—「家計調査」月次データを利用して—

橋 本 紀 子
荒 木 孝 治

要 旨

個別消費関数の推定では、費目間の相互関係を把握するため体系的なモデルが利用される。同時方程式体系における構造変化の検出は困難であるため、従来、時系列的なデータを用いた需要分析では、期間内の支出構造変化の有無やその内容についての検討はあまり行われてこなかった。

本稿は、この問題を克服するため、画像処理の分野で開発されてきたデータ解析手法 SINGLE 法を援用することにより、費目間の偏相関関係を洗い出し、それをもとに観察期間中の構造変化の有無や変化点を検出し、また各時点における費目間構造について検討を試みた。

『家計調査』の月次公表データを用いた検証の結果、2000 年以降の 15 年間に、二人以上世帯の勤労者世帯の支出行動には構造変化と見なされる変化が見られ、5つのフェーズに区分して考えられるとの結果を得た。また、それぞれの期間における費目間構造や中心となる費目を検出し、さらには階層的クラスタリングにより、観察期間中、どの時期の費目間構造が類似性を持つかについても検討を行った。

キーワード：家計調査；費目間構造；構造変化；SINGLE 法
経済学文献季報分類番号：15-71；15-72；16-10

1. はじめに

家計の支出行動の分析では、費目間の相互関係を把握するため体系的なアプローチに基づく需要モデルがよく用いられる。しかし、体系的モデルでは、パラメータ数を抑えるため対象とする費目数を増やすことが難しい。また、時系列データを用いて支出構造の変化を検出しようとする際、たとえ多くの費目がある一時点で変化したと仮定しても、その時点を特定するのは困難である。この結果、時系列データを用いた多品目需要関数の分析において、構造変化の問題はこれまでほとんど取り上げられてこなかった。

*) 本稿は科学研究費補助金基盤研究 (B)「周期統計調査のマイクロデータによるコーホート分析－文化需要の実証分析－」(研究代表者：勝浦正樹)(課題番号：23330073)の研究成果の一部である。

これに対して fMRI (functional Magnetic Resonance Imaging: 機能的磁気共鳴画像) データの分析において、複数の時系列データの統計的な依存関係の強さをはかるための手法である SINGLE (Smooth Incremental Graphical Lasso Estimation) 法 [4, 5] が開発されている。これは脳の関心のある複数の領域の血流量 (p 本) を計測したデータからそれらの関係の構造をネットワーク (グラフ) として抽出するものである。時系列に非定常性を仮定し、データの観測時点におけるグラフを描くことにより、そのパターンの変化を知ることができる。また、時系列の本数 p が大きい場合、ペアの関係数は $p(p-1)/2$ あるため、組み合わせ数が大きくなる。SINGLE 法では、関係のあり方にスパース (疎ら) 性を仮定することにより、観測期間が短い (データ数が少ない) 場合でも効率的にグラフを推定することができる。本稿ではこのような考え方に立ち、費目間の相関構造にスパース性と非定常性を仮定し、費目間のネットワーク構造の時間による変化を調べる。この手法を用いることにより、家計の費目の支出額の時系列データについて、期間中に構造変化が生じていないかどうか、生じているとするとそれはどのような変化であるかを把握することが期待できる。利用するデータは、総務省統計局『家計調査』[7] の費目別支出額月次データである。

本稿の構成は以下の通りである。第2節では今回利用するデータについて予備分析を行う。本稿では「家計調査」の二人以上世帯のうち勤労者世帯を対象とした月次データを利用するが、季節性が見られるか、見られた場合それを除去するかどうかを検討する。第3節で構造変化時点の検出とグラフ構造の推定手順について、偏相関係数の推定と SINGLE 法の考え方についてそれぞれ説明する。第4節で得られた結果を示す。まず、期間中の偏相関係数の動きについて見た後、それらの推移から変化点が存在するか、すなわち構造変化が生じているかどうかを検討する。もし変化が生じていた場合には、構造変化前後における費目間ネットワーク図 (グラフ) を描き、費目間構造の変化を検討する。さらに費目間構造の特徴について、グラフの密度により費目間構造の稠密さを、次数中心性により関係の中心となる費目を、クラスター分析により期間中の費目間構造の類似性について調べる。本稿の分析により、今回対象とした期間 (2000 年 1 月から 2014 年 11 月までの 179 ヶ月) にはいくつかの構造変化時点が検出される。最後に第5節で、構造変化の生じた背景や要因について、期間中の家計を取り巻く変数の動きを合わせ見ることで推察を行っていく。

2. 利用するデータと予備分析

本稿の分析では、総務省統計局『家計調査』の全国の二人以上世帯のうち勤労者世帯の支出額月次データを用いる。観察期間は、家計調査の調査対象が大きく変更された 2000 年 1 月以降、2014 年 11 月までの 179 ヶ月である。また、支出の対象となる費目は 1. 食料、2. 住居、

3. 光熱・水道、4. 家具・家事用品、5. 被服及び履物、6. 保健医療、7. 交通・通信、8. 教育、9. 教養娯楽、10. その他の消費支出の10費目である。なお、全ての支出額データは実質化している。

実質化した支出額データ（元データと表記している）の時系列プロットを図1に示す。図1より、データの変動に季節性があることがわかる。これをより詳細に見るために月別の箱ひげ図を図2（ここでは特に動きの大きかった6つの費目のみ表示）に示す。これより多くの変数に年次での季節性が見られることがわかる。そこで、データ解析環境 R[6] の関数 `decompose()` を用いて季節変動を抽出し（データに季節要素、トレンド要素、ランダム要素の加法性を仮定している）、それを除去することにより季節調整を行った¹⁾。季節調整後のデータの動き（図3）より、季節性が弱まっていることが確認できる。このデータの散布図行列を作成すると図4となる。図4より、若干の外れ値が見られるものの、2つの費目の同時分布はほぼ2次元正規分布をしていると判断できる。

以下、この季節調整後のデータを分析の対象としていく。

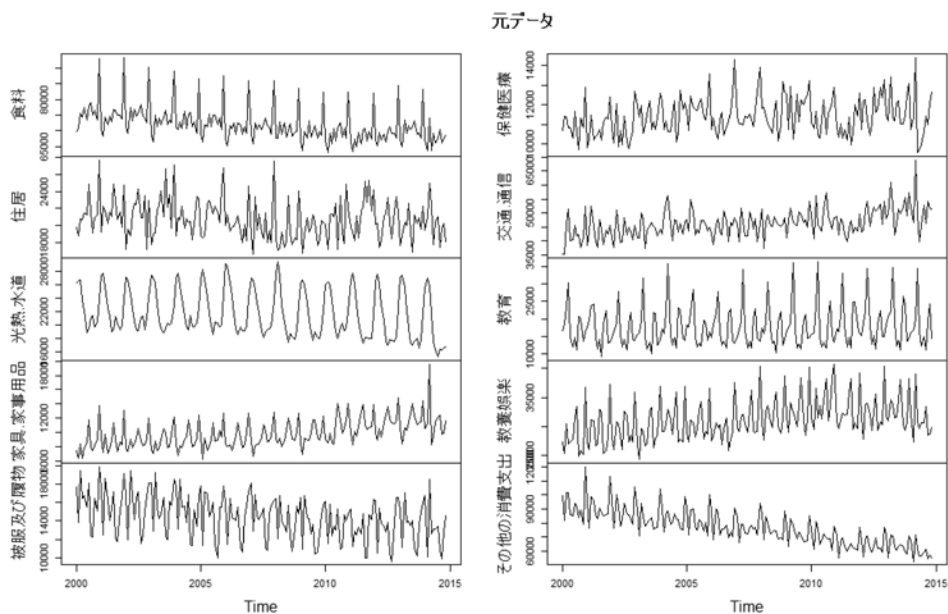


図1 元データの時系列プロット

1) 以下、データ分析は全て R を用いて行っている。

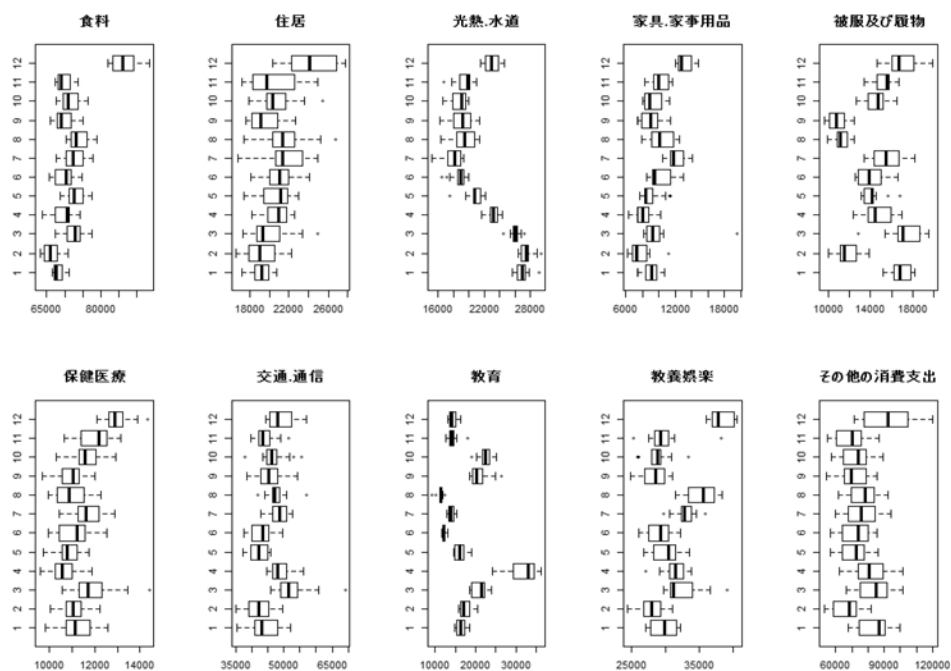


図2 月別の箱ひげ図

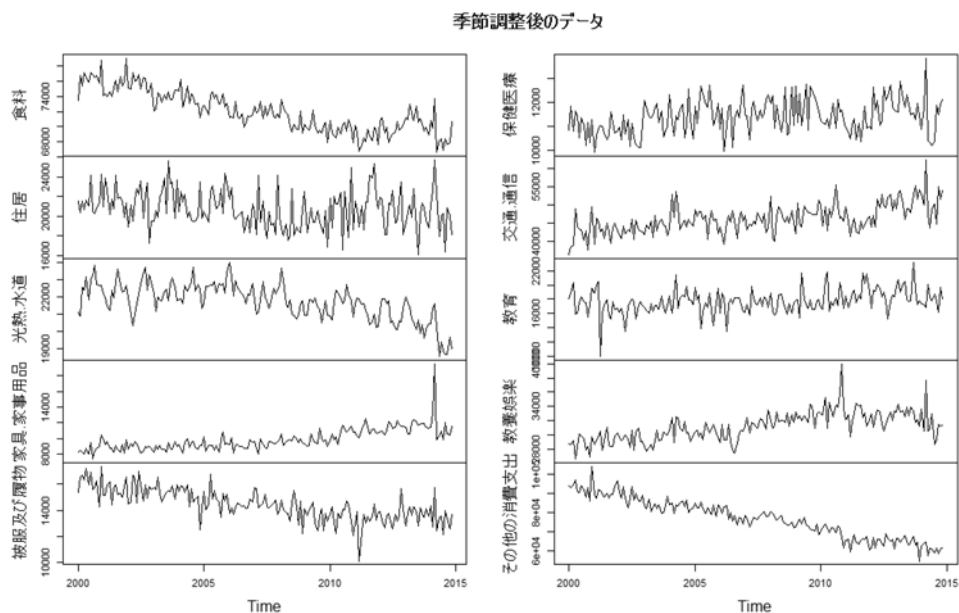


図3 季節調整後のデータの時系列プロット

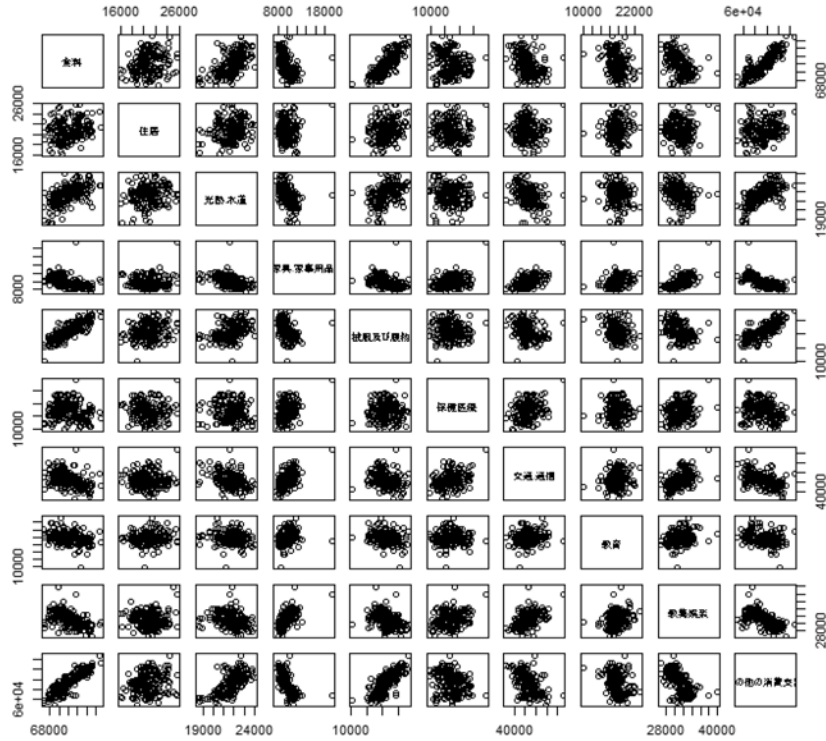


図 4 散布図行列

3. グラフ構造の推定

3.1 偏相関係数

p 次元確率変数の時系列 $X_i \in R^{1 \times p}$ ($i = 1, 2, \dots, T$) を考える。 $X_i = (X_i^1, X_i^2, \dots, X_i^p)^T$ は、第 i 時点における p 個の費目の支出額であり、 $X_i \sim N(\mu_i, \Sigma_i)$ とする (本稿の場合、 $p = 10$, $T = 179$)。ここで、 Σ_i は p 行 p 列の共分散行列であり、この逆行列 Σ_i^{-1} を θ_i とし、その要素を $\theta_{i,kl}$ ($k = 1, 2, \dots, p; l = 1, 2, \dots, p$) とする。 θ_i を精度行列または集中行列という。

第 i 時点での費目 k と l の間の偏相関係数を $\rho_{i,kl}$ とするとき、これは式(1)に示すように、 θ_i の要素で表すことができる。

$$\rho_{i,kl} = -\frac{\theta_{i,kl}}{\sqrt{\theta_{i,kk}\theta_{i,ll}}} \quad (1)$$

式(1)より、 θ_i の要素 $\theta_{i,kl}$ が 0 のとき、偏相関係数 $\rho_{i,kl}$ は 0 となるため、これに対応する費目 k と l は、これらを除く $p-2$ 個の他の費目に関して条件づけたとき、第 i 時点において独立であると判断することができる [3, 5]。さらにこれらの情報をもとに、費目間のネットワーク構造を推定することができる。しかし、 θ_i は未知のためこれを推定する必要がある (推

定値を $\hat{\theta}_i$ とする)。 $\hat{\theta}_i$ を求める目的は、これを用いて要素（今の場合、費目）間の関係の強さ、さらには独立性を判断することができるからである。

費目をノード（点）とし、その集合を V とする。 $\hat{\theta}_i$ にもとづいて偏相関が 0 でないと判断できるとき、つまり費目間に関係があるときにエッジ（線）を描くことにより、第 i 時点での費目間のネットワーク図（グラフ）を作成することができる。すなわち、第 i 時点におけるエッジの集合を E_i とするとき、 $\hat{\theta}_i$ を用いてノード間の関係をグラフ $G_i = (V, E_i)$ として推定することができるのである。

以上、偏相関係数を推定することにより費目間の関係やそのネットワーク構造について明らかにできることを見たが、問題はいかにして偏相関係数を推定するかである。そのために本稿で用いる SINGLE 法について、次項で説明する。

3.2 SINGLE 法について

Monti ら [4, 5] は、fMRI（機能的磁気共鳴画像）データの解析法として SINGLE (Smooth Incremental Graphical Lasso Estimation) 法を提案した。SINGLE 法では、(a) グラフの構造に関するスパース（疎ら）性、(b) 時間等質性を仮定する。(a) は、グラフのエッジはそれほど多くない、つまりエッジが疎らであるという仮定である。(b) は、グラフ構造は近接する時間ではそれほど変化しない、つまり、変化は滑らかに生じるという性質の仮定である。

SINGLE 法では、損失関数 $l\{\theta_i\}$ を式(2)で定義し、精度行列 θ_i ($i = 1, \dots, T$) を推定する。

$$l\{\theta_i\} = \underbrace{\sum_{i=1}^T \{-\log \det \Theta_i + \text{trace}(S_i \Theta_i)\}}_{\text{-対数尤度}} + \underbrace{\lambda_1 \sum_{i=1}^T \|\Theta_i\|_1}_{\text{lasso 罰則}} + \underbrace{\lambda_2 \sum_{i=2}^T \|\Theta_i - \Theta_{i-1}\|_1}_{\text{fused lasso 罰則}} \quad (2)$$

ここで、 $\|\cdot\|_1$ は行列の各要素の絶対値の和 ($\|A\|_1 = \sum_{i=1}^p |A_{ij}|$) を示す。

式(2)の右辺の第 1 項は、多変量正規分布を仮定したときの対数尤度のマイナスの項である。最尤法では、対数尤度を最大とするパラメータを推定するが、ここではそれをマイナスに設定しているため、最小とするパラメータを推定することになる。第 2 項は、グラフのスパース性を確保するための項であり、lasso 罰則 [8] という。 λ_1 が正のとき、 θ_i の要素の多くが 0 でないと、この項が大きくなり、損失関数は小さくならない。そのため、 θ_i の要素をできるだけ 0 に縮小し、小さなものは 0 にする効果を持つ。第 3 項は、時間等質性を確保するための項であり、fused lasso 罰則 [9] という。 θ_i の要素の隣接した時点間の差の小さなものを 0 に縮小し、真に差が大きなもののみを検出する効果を持つ。

λ_1 は lasso 罰則の強さ、 λ_2 は fused lasso 罰則の強さを調整するためのパラメータであり、チューニングパラメータともいう。 $\lambda_1 = \lambda_2 = 0$ のとき、式(2)は最尤法に対応する。チューニングパラメータの値が大きくなると罰則が強くなり、 θ_i の要素や $\theta_i - \theta_{i-1}$ の要素の小さなものは 0 に縮小される。

式(2)は既に説明したように、最尤法に対して、2つの罰則項 (lasso と fused lasso) を付加した形になっている。第1項と第2項の計を最小にする形で θ_i を推定する手法がグラフィカル lasso [2] という。なお、こうした罰則項を用いたパラメータの推定法を一般に正則化法という。

式(1)による θ_i の推定には、標本共分散行列 S_1, S_2, \dots, S_T が一般に用いられる。時間の推移により構造が変化していないことを仮定できる場合、通常、全期間の標本共分散行列が等しいとし ($S_1 = S_2 = \dots = S_T = S$)、全データをプールした

$$S = \frac{1}{T-1} \sum_{i=1}^T (X_i - \bar{x})^T (X_i - \bar{x}) \quad (3)$$

を推定値として用いる。SINGLE 法では、共分散行列の推定に時間依存性を導入するために、ガウスカーネル (Gaussian kernel)

$$K_h(i, j) = \exp \left\{ \frac{-(i-j)^2}{h} \right\} \quad (4)$$

を利用して (式中の h をカーネル幅という)、次の形で推定する。

$$\bar{x}_i = \frac{\sum_{j=1}^T K_h(i, j) X_j}{\sum_{j=1}^T K_h(i, j)} \quad S_i = \frac{\sum_{j=1}^T K_h(i, j) (X_j - \bar{x}_j)^T (X_j - \bar{x}_j)}{\sum_{j=1}^T K_h(i, j)} \quad (5)$$

SINGLE 法では、損失関数 $l\{\theta_i\}$ を最小にする θ_i の計算には、交互方向乗数法 (ADMM: Alternating Directions Method of Multipliers) [1] を用いている。

4. 得られた結果

4.1 偏相関の推定

正則化法では、正則化パラメータの最適な設定が必要となる。SINGLE 法では、AIC (赤池情報量規準) を用いて λ_1 および λ_2 の推定を行う。実際には式(4)のカーネル幅 h の推定も必要である。SINGLE 法で求めたパラメータ λ_1 , λ_2 , h の最適値の組に対して、これらを少しずつ変化させたところ、より小さな AIC を持つパラメータの組 $\lambda_1 = 0.11$, $\lambda_2 = 0.4$, $h = 34$ が存在したので、以下、このパラメータ値を用いた結果を示す。

推定された偏相関係数 (PC: Partial Correlation) の観測期間中の推移を図5に示す²⁾。費目数は $p = 10$ なので、これらの組合せ数は $p \times (p - 1) / 2 = 45$ となる。図5では、この組合せ

2) 以下、「推定された」は、文脈の中で分かるため省略する。

数の折れ線グラフが示されている。一つのグラフが、ある2費目間の偏相関係数の、2000年1月以降の179ヶ月間にわたる推移を示している。グラフの見出し、たとえば「PC: 8:10」は、費目8と10の偏相関係数の推移図であることを示す。

この図より、相手の費目によらず、10. その他の消費支出に関わる偏相関係数は0.0からほとんど変化していないため、10. その他の消費支出は、データ期間中（ほぼ）他の費目との関係がなかったことがわかる。また、{1. 食料, 2. 住居} {3. 光熱・水道, 5. 被服及び履物} {3, 9. 教養娯楽} の組合せも期間を通じてほぼ関係が見られなかった。

費目間の関係は、基本的に正あるいは負の領域だけで推移しているが、少数ながらいくつか符号が変化している組み合わせがある。{2. 住居, 3. 光熱・水道} {2, 6. 保健医療} {2, 7. 交通・通信} {3, 6. 保健医療} {3, 7} {4. 家具・家事用品, 8. 教育} {7, 8} である。主に {2, 3, 6, 7, 8} が関わる組合せにおいて、時間経過のなかで関係の正負が変化している。

期間中費目間の変化が見られた場合、多くの費目間の関係は正の領域で動いている。負の偏相関のみで推移したのは、{1. 食料, 8. 教育} {5. 被服及び履物, 8} の組合せのみであった。

変化の動きを見ると、組合せ {1, 4} {1, 9} {2, 3} {3, 4} {4, 5} {4, 6} {4, 7} {4, 9} {4, 10} {6, 7} のように、初期や中期には関係がない、あるいは弱かったものが近年になって関係が強くなっているものが見られる。また、組合せ {1, 5} {1, 6} {2, 4} {2, 5} {2, 7} {2, 8} {2, 9} {5, 9} {7, 9} のように、初期にあった関係が中期において関係がなくなったり、弱くなったりしているものもある。この場合も近年には関係が強くなる傾向が見られる。こうした偏相関の関係を時点ごとにグラフに示すと、179個のグラフを得る。

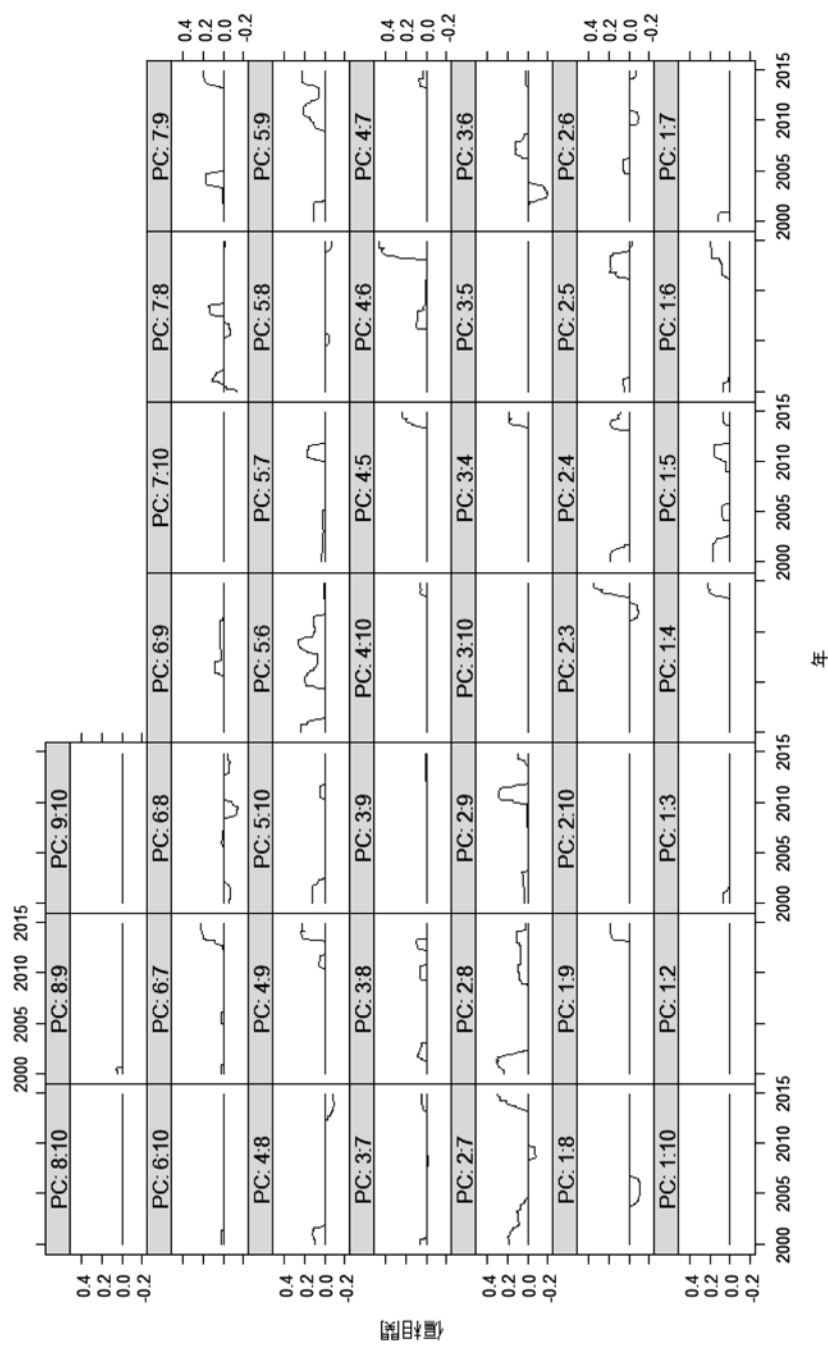


図 5 推定された偏相関係数の推移

4. 2 得られたグラフの分析

得られたグラフは179個あり、これらの変化の全体としてのパターンを知るために、推定した45組の偏相関の変化点を調べる。推移を一つのグラフにまとめると図6のようになる。図より、2001年から2002年、2013年から2014年にかけて多くの偏相関の値が変化していることがわかる。

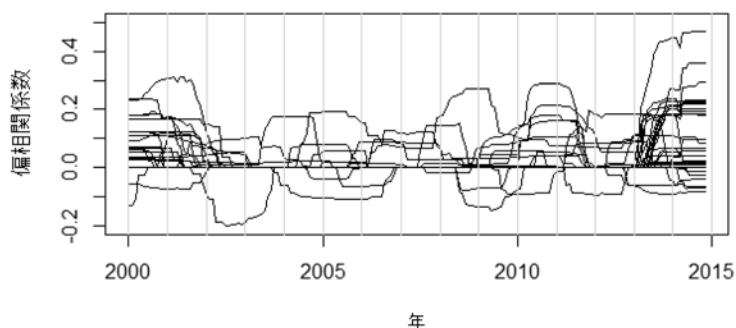


図6 偏相関係数の変化

変化点を推定するために、図5に示した各組合せの偏相関係数の推移データから個別に変化点を検出し、それを集約して全体としての変化点を推定する³⁾。集約した変化点のデータの密度関数を推定すると、図7のようになる。

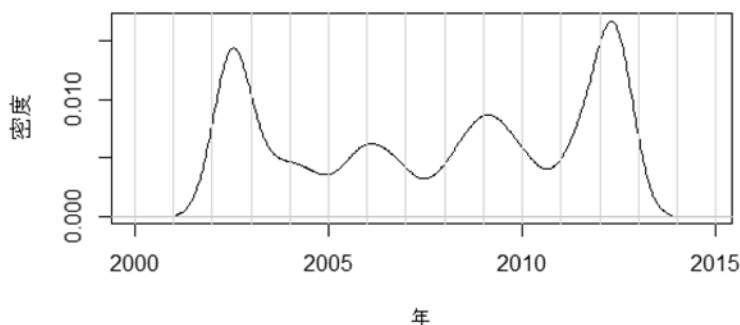


図7 変化点の推定

図7より、期間中に4つの変化点が全体としてあることがわかる。それは、時点32（2002年8月）、57（2004年9月）、75（2006年3月）、150（2012年6月）の付近である。そこで、変化点のほぼ中間点にあたる15（2001年3月）、45（2003年9月）、70（2005年10月）、110（2009

3) 変化点の推定には、Rのパッケージ ecp を利用した。

年2月)、160(2013年4月)におけるネットワーク図を作成すると図8になる。

図8の(a)と(b)はそれぞれ同一のネットワーク図であるが、(a)では費目間の関係の有無や符号、強さを時点間に対応させて見るために費目を円周上に固定して配置している。(b)では費目間の相互の関係が見やすくなるよう配置している。

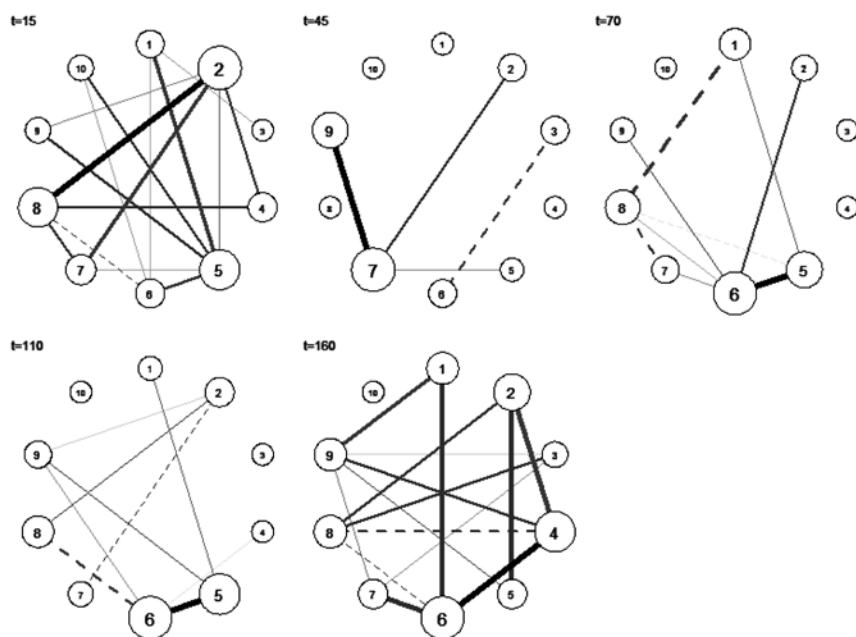
まず(a)より、5つの期間それぞれにおける費目間の関係の有無や符号、強さを見る。ここでエッジ(線)の実線は正の相関を、破線は負の相関を示す。エッジの太さは、相関係数の絶対値の大きさに応じている。また、ノードの大きさは、後に検討するグラフにおける次数中心性の大きさに応じている。本図より、偏相関係数の推移(図5)でも見られたように、当初と最近の費目間の関係は密であるが、中期では関係が疎らになっていることがわかる。

次に(b)より、費目間の関係、ネットワークの特徴を見ていく。部分グラフの中で3以上のノードで構成される完全グラフ(すべてのノードが連結されたグラフ)をクリークという。このような密接に関連する部分グラフがネットワーク内に存在しているかどうかを、図8(b)の5つのグラフについて見てみる。

順にクリークを抽出すると(エッジの太さは考えていない)、15期におけるグラフでは、 $\{1, 5, 6\}$ $\{2, 5, 9\}$ $\{2, 7, 8\}$ $\{2, 5, 7\}$ $\{2, 4, 8\}$ $\{5, 6, 10\}$ の6つのクリークがある。45期におけるグラフではクリークはなく、70期のグラフでは、 $\{5, 6, 8\}$ $\{6, 7, 8\}$ $\{1, 5, 8\}$ の3つ、110期では、 $\{5, 6, 9\}$ の1つ、160期では $\{3, 7, 9\}$ $\{4, 6, 8\}$ $\{2, 4, 8\}$ の3つがある。またクリークに費目10が含まれるのは15期におけるグラフのみであり、費目3が含まれるのは160期におけるグラフのみである。また、どのグラフにおいても、共通したクリークは存在しない。また、ノード数が4以上のクリークも存在しない。

以下、ネットワーク分析の手法[8]を用いて、得られたグラフをさらに分析していく。

(a)



(b)

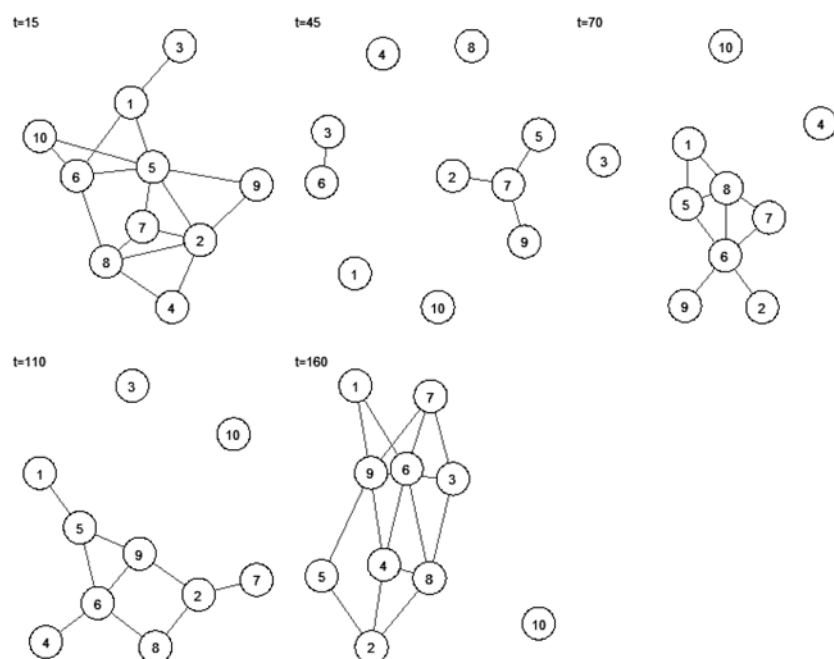


図8 変化点前後におけるネットワーク図

(1) グラフの密度

費目間の関係の密接さを調べるために、ネットワーク分析で用いられる指標である「密度」(density)を求める。密度は、グラフが持つことができるエッジの全ての数に対する実際に存在するものの割合と定義される。密度は0から1の値を取り、密度が0のとき、エッジが全く存在しない空グラフであり、1のとき、すべてのエッジが存在する完全グラフである。1に近いほど密度が高いと判断する。先に求めたクリークは、密度が1のサブグラフである。

密度の月次での推移グラフを図9に示す。図9より、期首の2000年から密度が減少し始め、2003年から2004年にかけて最もスパースになり、その後2006年にかけて少し上昇した後下降し、2011年半ばまでまた少し上昇した後に下降し、2012年初頭から大きく上昇していることが見てとれる。

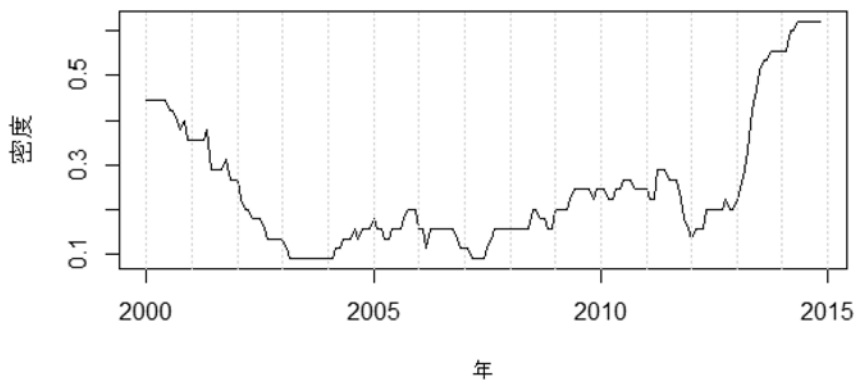


図9 グラフの密度の推移

(2) グラフの中心性

次に、グラフにおける各ノード(費目)の役割を分析するために、次数中心性(degree)を求める。次数中心性はノードが持つエッジの数であり、これが大きいノードはそのグラフで中心的な役割を果たしていると判断することができる。各費目の次数中心性の推移を図10に示す。

全体を通じて、10. その他の消費支出の次数中心性は低く、ほぼ一定で推移している。3. 光熱・水道および4. 家具・家事用品は、当初一定であったが、終期に増えるパターンを示している。他の費目に関しては、密度の高い初期と末期で中心性が高くなっている。密度の低い中期において、6. 保健医療の次数中心性が高い。中期では、保健医療を中心に費目間の関係が成り立っていることがわかる。密度のグラフ(図9)において、中期に2つのピークが見られたが、これには、5. 被服及び履物、8. 教育が関与していると考えられる。

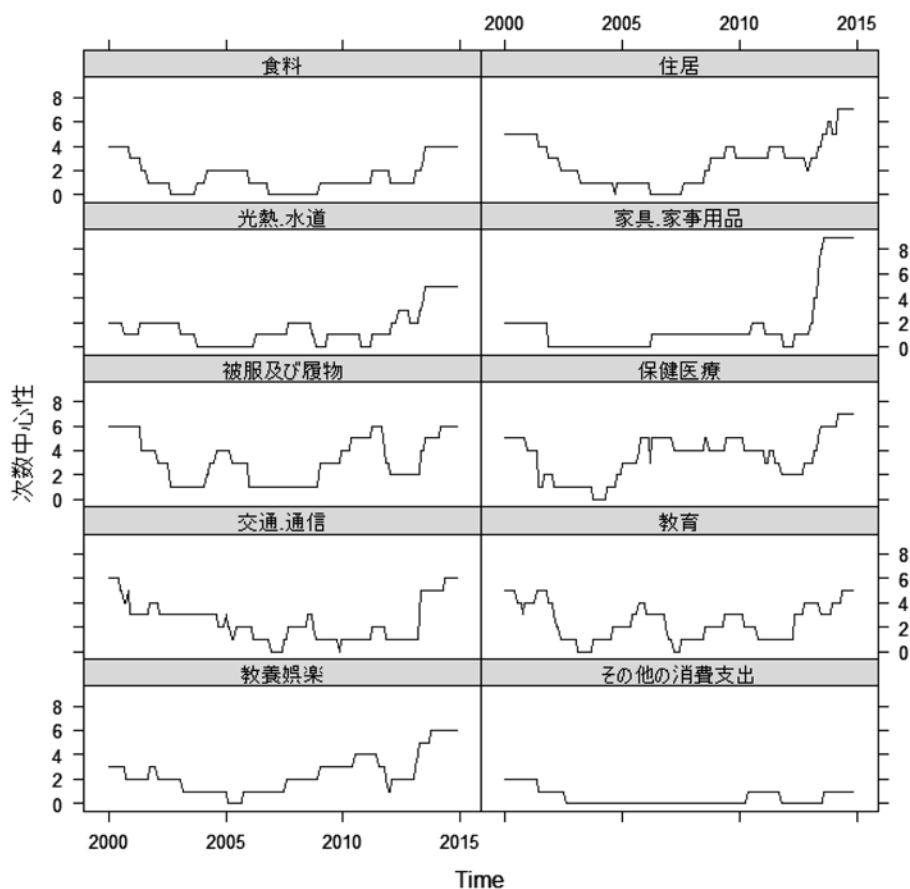


図 10 次数中心性の推移

(3) グラフのクラスタリング

ここではクラスタリング手法を用いて、179 個あるグラフの中で、どのグラフが同じような構造となっているかを見る。階層的クラスタリング（マンハッタン距離、完全連結法を適用）により樹状図（デンドログラム）を作成すると、図 11 のようになる。

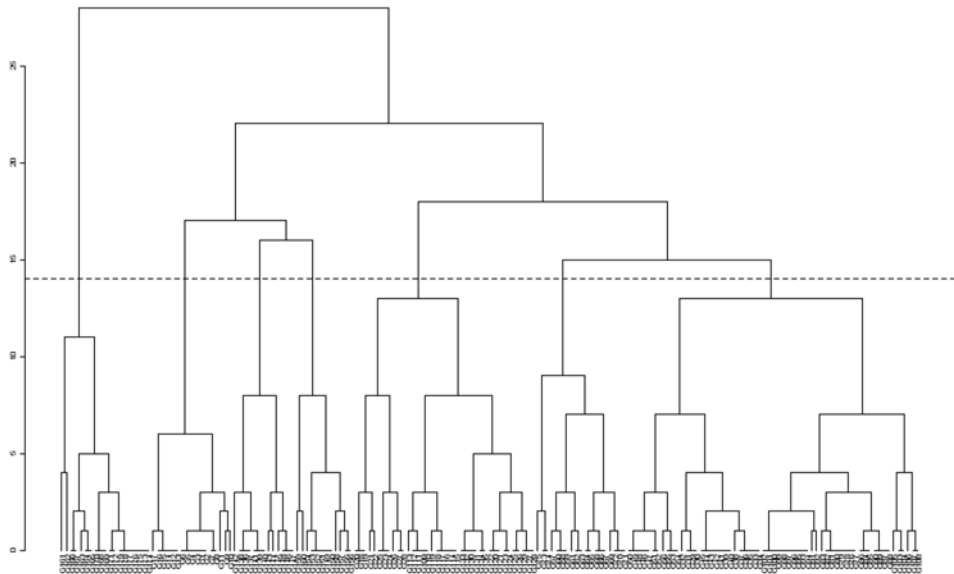


図 11 階層的クラスタリングの結果―全体

図 11 の樹状図を縦軸の高さ 14 でカットすると（図中の水平の破線）、図 12 に示す 7 つの枝（Branch）を得る。Branch 2, 3, 4 が一つのグループ、Branch 5, 6, 7 がもう一つのグループを構成し、Branch 1 が他から一番離れている。

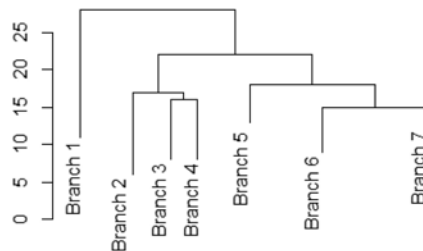


図 12 階層的クラスタリングの結果―グルーピング

次に各枝の詳細を図 13-1、図 13-2 に示す。図中の「G 番号」、たとえば G161 は第 161 期のグラフを意味する。Branch 1 は、G161 ～ G179（2013 年 5 月～2014 年 11 月）までの末期のグラフのグループ、Branch 2 は、G1 ～ G17（2000 年 1 月～2001 年 5 月）までの初期のグラフのグループ、Branch 3 は、G136 ～ G148（2011 年 4 月～2012 年 4 月）までのグループ、Branch 4 は、G149 ～ G160（2012 年 5 月～2013 年 4 月）のグループ、Branch 6 は、

G57～G75（2004年9月～2006年3月）のグループである。Branch 5はG18～G28（2001年6月～2002年4月）とG109～G135（2009年1月～2011年3月）のグループであり、Branch 7は、G29～G56（2002年5月～2004年8月）とG76～G108（2006年4月～2008年12月）のグループである。Branch 5と7は、異なる期間のグラフのグループから構成されている。

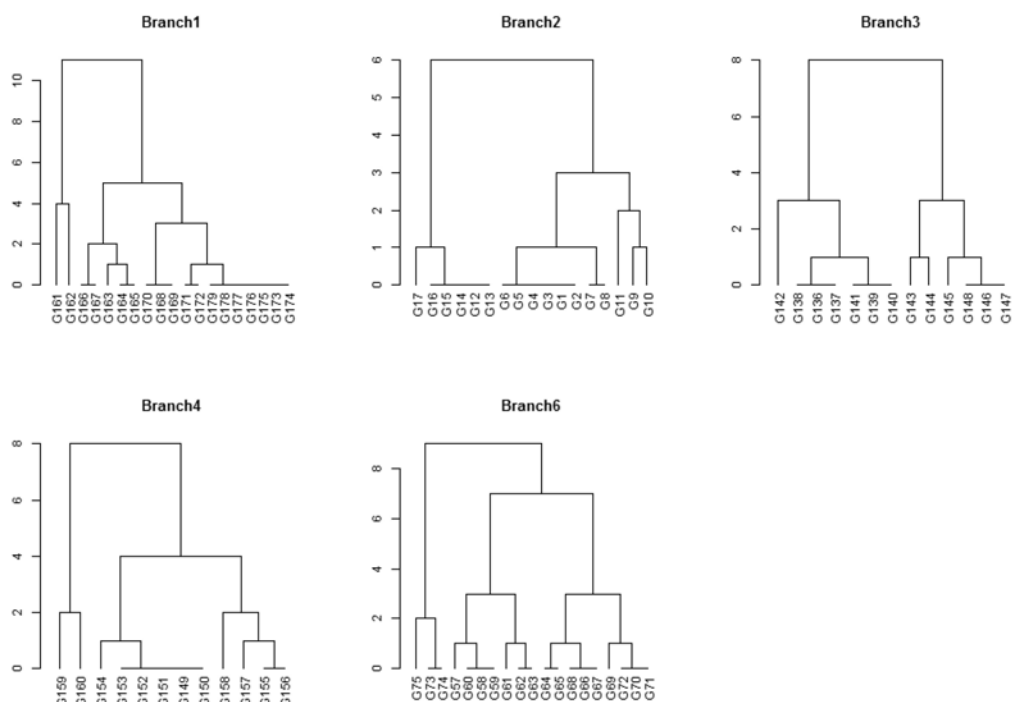


図 13-1 枝を構成する要素の詳細 1 — Branch 1, 2, 3, 4, 6

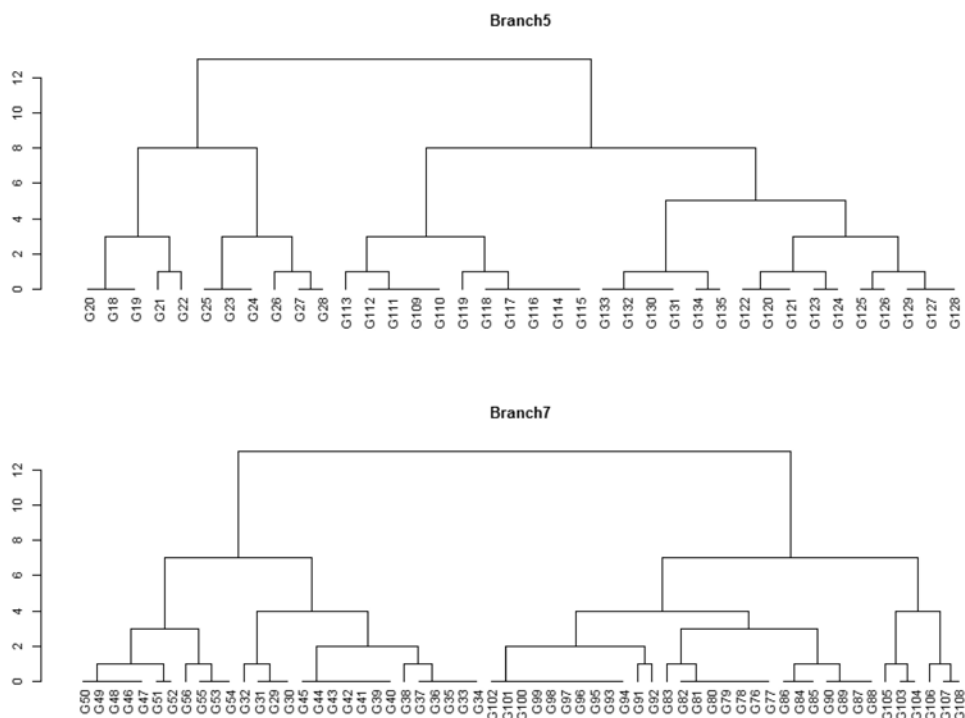


図 13-2 枝を構成する要素の詳細 2 - Branch 5, 7

年別にグループを並べると、Branch 2、Branch 5（前半）、Branch 7（前半）、Branch 6、Branch 7（後半）、Branch 5（後半）、Branch 3、Branch 4、Branch 1となる。これを図示すると図 14 となる。ネットワークの形としては、Branch 5に含まれる 2001 年 6 月～2002 年 4 月と 2009 年 1 月～2011 年 3 月が似ており、また、Branch 7に含まれる 2002 年 5 月～2004 年 2 月と 2006 年 4 月～2008 年 12 月が似ていると判断することができる。また、Branch 2と Branch 3、Branch 4も似ていると判断することができる。

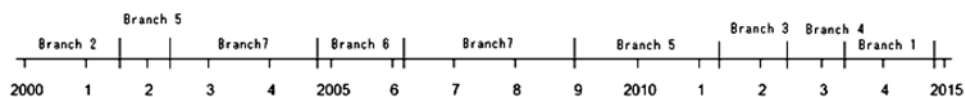


図 14 階層的クラスタリングによるグループの推移

5. 得られた結果の意味と残された課題

以上、家計調査の月次データを用いて、SINGLE 法を利用し、二人以上の勤労者世帯の支出行動において、2000 年以降の 15 年間に 4 度の構造変化が生じていること、また、それぞ

れの期間における費目間構造、その中心となる費目、どの時期の費目間構造が類似性を持つかについての検討を行った。

さて、本分析の結果、2000年1月から2014年11月までの179ヶ月間の勤労者世帯の支出行動は、5つのフェーズ、つまり、

- ① 2000年1月から2002年8月（32期）
- ② 2002年9月から2006年3月（75期）
- ③ 2006年4月から2009年2月（110期）
- ④ 2009年3月から2012年6月（150期）
- ⑤ 2012年7月から2014年11月（179期）

に区分して考えられるという結果が得られた。なお、括弧内の期間数は、その時点が今回用いたデータの何期目（単位は月）に当たるかを示している（以下、同様）。

図14で見た階層クラスタリングの結果とつきあわせると、①はBranch 2とBranch 5の前半、②がBranch 7の前半とBranch 6、③がBranch 7の後半、④がBranch 5の後半とBranch 3、⑤がBranch 4とBranch 1に対応している。図12のグルーピングの結果を見ると、グループが近いBranch 5、Branch 6、Branch 7は観察期間の中頃に集まっており、図9よりその時期のグラフの密度が低いこと、図8よりその時期の費目の関係は疎らであることがわかる。逆に密度が高く費目の関係性が強い時期は期間の初期と末期、①と⑤のフェーズに対応するが、クラスタリングのグルーピングでは類似性が見られたBranch 2とBranch 3およびBranch 4が初期と末期に分かれて存在している。また密度が急激に高まった時期はBranch 4とBranch 1に対応するが、図12より、Branch 1は他のグループとは異なる費目間構造をしていると判断されるため、この点についてはより詳細に見ていく必要がある。

そこで本稿で得られた変化点やクラスタリングの結果について、その背景や意味を探るため、観察期間中の経済動向を概観し、合わせて考察していく。

2000年春にITバブルが崩壊した後、日本経済はなかなか回復基調に戻れなかったが、その後、2002年2月（26期）から2008年2月（98期）頃にかけて「いざなぎ景気」と呼ばれる73か月にわたる戦後最も長い景気拡大が生じた。しかし、サブプライム問題（世界金融危機）から派生した2008年9月（105期）のリーマンショックは若干のタイムラグを持って日本経済にも大きな影響を与える。四半期ベースで見たとき、2009年第1四半期の対前年比実質経済成長率は-9.4%と大きく落ち込んだ。その後2010年には5%前後の成長が見られたものの、2011年第2、第3四半期や2012年第4四半期にはゼロあるいはマイナス成長の状態に陥っている。なお2012年の暮れに安倍政権が登場し、アベノミクスと呼ばれる経済政策に着手した。2013年4月（160期）には日銀が「異次元」と称される金融緩和を行い、

2014年4月(172期)には消費税の税率が5%から8%に17年ぶりに上げられている。より詳細な検討が必要ではあるが、階層的クラスタリングによるグルーピングでBranch 1(2013年5月、161期以降)が他の期間と大きく異なると推定されたのは、アベノミクスの影響によるとも考え得る。

実質経済成長率のグラフを図15に示す。データは四半期ベースのため、3ヶ月間に同じ値を与えることにより月次データとして表示している。また、今回の分析で得られた勤労者世帯支出行動の変化点を図中に縦の実線で示している(以下同じ)。成長率の動きと合わせてみると、①の期間はいざなぎ景気が始まるまでの景気低迷期、②はいざなぎ景気の前半、③はいざなぎ景気後半からリーマンショックによる景気の谷まで、④や⑤はその後の復調期に当たる。①や②など経済の動向により期間の意味合いが説明できるところもあるが、好景気と急激に景気が悪化する時期が一つの期間と見なされていたり、同じいざなぎ景気あるいはリーマンショックからの復調期が別々の期間とみなされたりしているのか、その説明もうまく行うことができない。

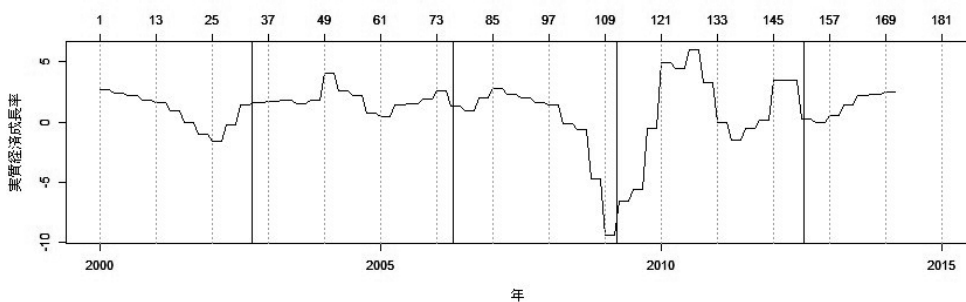


図15 実質経済成長率の推移

成長率は経済の動向を示す代表的な指標であるが、マクロ経済指標であるため、個々の世帯が感じる景況感とずれが生じる可能性がある。企業の収益が賃金にどのように反映されるか、ラグの長さはどれだけか、といった問題もある。そこで、次に、今回分析対象とした二人以上勤労者世帯の収入の動きについて見る(図16)。これも、それぞれの期間の実収入はある程度似通った動きや傾向を見せるものの、推定されたフェーズをうまく説明することはできない。

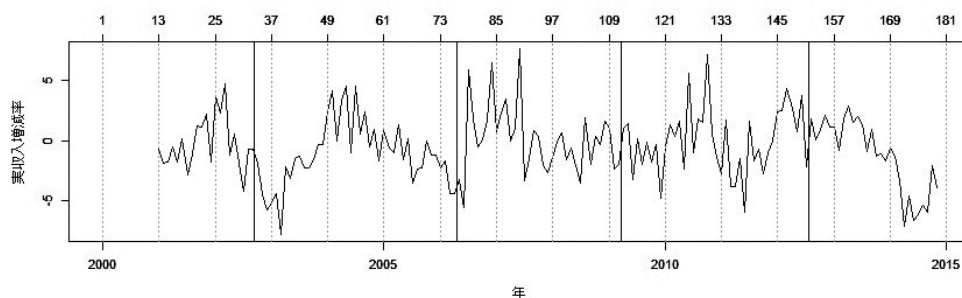


図 16 家計の実収入の増減率（月次、対前年同月比）

家計は、得られた収入に、必要な場合は資産の売却、貯蓄の取り崩しなどを行い、その結果消費支出を行う。支出行動の構造変化を考える際には、実収入の動き以外にそれらの動向についても考慮する必要があるであろう。また、実際に獲得できた実収入の動きに左右されるのは当然であるが、今後の経済動向への期待が支出行動に影響する、いわゆるマインドの問題も少なからずあると思われる。そこで、家計が頻繁に耳にする可能性のある指標のうち、分かりやすく、自らを含めた雇用状況を肌身で感じる指標として、有効求人倍率⁴⁾（厚生労働省、「一般職業紹介状況（職業安定業務統計）」）の期間中の動きを見るため、それを図 17 に示す。

図 17 より、32 期までの①の期間、有効求人倍率は 0.6 前後で停滞、75 期までの②の期間は上昇し 1.0 を超え、その後 110 期頃まで大きく下落（③）、それ以降、有効求人倍率は④（150 期まで）に上昇、⑤（150 期以降）はさらに上昇を続けていることがわかる。このように、④と⑤の期間がなぜ分かれるかについての説明は未解明なものの、本稿で得られた支出行動の動きと、有効求人倍率の動きにはかなりの相関がみられることがわかる。

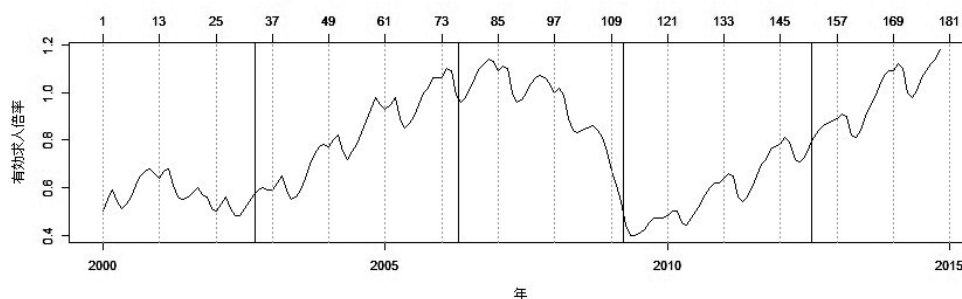


図 17 有効求人倍率の推移

4) 求人求職数には新規学卒者関連は含まれないが、パートタイム労働は含まれている。

本稿では SINGLE 法により、支出行動における構造変化の有無を見いだせる可能性を示したが、その経済学的な解釈を成立させるにはまだ数多くの検討、考察が必要である。上で有効求人倍率の推移が本稿で得られた変化点と呼応していることを示したが、このような対応がどのような意味を持つのか、家計の支出行動と有効求人倍率の動きをつなぐ第三の要因が存するのか、また、実収入以外の収入要因、他の支出行動に影響を及ぼす要因についての検討は、記して今後の課題としておきたい。

参考文献

- [1] Boyd, S. and L. Vandenberghe (2004). *Convex Optimization*, Cambridge University Press.
- [2] Friedman, T., T. Hastie, and R. Tibshirani (2008). Sparse inverse covariance estimation via the graphical lasso, *Biostatistics*, 9, 432-441.
- [3] Hastie, T., R. Tibshirani, and J. Friedman (2009). *The Elements of Statistical Learning*, 2nd ed., Springer. (杉山将他訳 (2014) 『統計的学習の基礎—データマイニング・推論・予測』, 共立出版).
- [4] Monti, R. P., C. Anagnostopoulos and G. Montana (2014). Package 'SINGLE', <http://cran.ism.ac.jp/web/packages/SINGLE/index.html>. R package version 1.3.
- [5] Monti, R. P., P. Hellyer, D. Sharp, R. Leech, C. Anagnostopoulos, and G. Montana (2014). Estimating time-varying brain connectivity networks from functional MRI time series, *NeuroImage*, 103, 427-443.
- [6] R Core Team (2014). R: A Language and Environment for Statistical Computing, R Foundation for Statistical Computing, <http://www.R-project.org/>.
- [7] 総務省統計局, <http://www.stat.go.jp/data/kakei/>.
- [8] 鈴木努 (2009) 『R で学ぶデータサイエンス 8 ネットワーク分析』 共立出版.
- [9] Tibshirani, R. (1996). Regression shrinkage and selection via the lasso, *Journal of the Royal Statistical Society B*, 58, 267-288.
- [10] Tibshirani, R., M. Saunders, S. Rosset, J. Zhu, and K. Knight (2005). Sparsity and smoothness via the fused lasso, *Journal of the Royal Statistical Society B*, 67, 91-108.