

## 科学研究費助成事業（学術研究助成基金助成金）研究成果報告書

平成25年6月7日現在

機関番号：12601

研究種目：挑戦的萌芽研究

研究期間：2012

課題番号：24651175

研究課題名（和文） 計算集約的経済分析の可能性：大変動・非定常・非線形を前提とした正確な実証分析

研究課題名（英文） Toward computer intensive methods in economic analysis, with special focus on non-Gaussian, non-stationary and non-linear factors

研究代表者

大西 立顕 (OHNISHI TAKAAKI)

東京大学・大学院経済学研究科・特任研究員

研究者番号：10376387

研究成果の概要（和文）：

金融危機やバブルにみられるように、経済現象において大変動・非定常・非線形な性質が及ぼす影響は非常に大きく、これらの性質が本質的に重要になる。様々な形式の経済データについて、これらの性質を前提とした正確な実証分析を行い、スーパーコンピュータを活用した計算集約的経済分析の手法を提案し、その有効性を検証した。

研究成果の概要（英文）：

As is often the case with financial crises or bubbles, non-Gaussian, non-stationary and non-linear factors have played a significant roles in economic phenomena, and in-depth understanding of such factors should be essential for future measures. Carrying out more detailed and accurate analyses on various types of economic data, we propose brand-new computer intensive methods in economic analysis by utilizing the unrivaled calculation ability of the supercomputer, and verify their effectiveness and usefulness.

交付決定額

(金額単位：円)

	直接経費	間接経費	合計
交付決定額	3,100,000	930,000	4,030,000

研究分野：複合新領域

科研費の分科・細目：社会・安全システム科学 社会システム工学・安全システム

キーワード：スーパーコンピュータ、ランダムシャッフル、非ガウス性、非定常性、非線形性、計算集約的手法、複雑ネットワーク、空間相関

### 1. 研究開始当初の背景

情報の電子化が進んだことにより、日々、様々な分野で膨大な量のデータが蓄積されるようになってきている。こうした大規模なデータを利活用することで、経済現象を実証科学的に研究し、社会やビジネスに役立てたりする試みが活発になってきている。また、コンピューターの計算能力の急速な向上に

伴い、今まではほとんど不可能に近かった、大量のデータを用いて大量の計算を行うような分析が可能になり、計算集約的手法による厳密な分析ができるようになってきている。

経済システムは、売り手と買い手の間のお金やモノの交換(相互作用)として捉えることができる。相互作用があっても個々がランダ

ムにゆらげば、全体のゆらぎは正規分布に従う。しかし、現実の経済システムでは、常に競争が働き相互作用が強いため、全体のゆらぎは極めて大きく、関連する変数はベキ分布に従う(大変動)。さらに、変動の性質や規則性は時々刻々変化するため、時期・環境の状態変化を考慮する必要もある(非定常性)。このような大変動・非定常・非線形な性質は経済現象の本質であり、既存のほとんどの統計的・数理的手法はこれらを前提としないため、必ずしも有効なものになっていない。

## 2. 研究の目的

経済現象において、大変動・非定常・非線形な性質が及ぼす影響は非常に大きく、これらの性質が現象の本質になっている。これらの性質を厳密に考慮した解析を行うには、適切なランダムシャッフルデータと比較したり、時期・環境を細分化するなどして、データドリブン(データ駆動)の思考で実データを基盤とした分析を行う必要がある。本研究では、様々な形式の経済データについてこれらの性質を前提とした正確な実証分析を行い、スーパーコンピュータを活用した計算集約的経済分析の手法を提案し、その有効性を検証した。

## 3. 研究の方法

計算集約的経済分析の例として、2変数X, Yの間の相関を求める問題を考える。実データから図1(a)のような散布図が得られたとする。通常、相関係数の検定では、Xの分布もYの分布も近似的に正規分布に従うと仮定するため、有意な相関であると判断されてしまう。しかし、実際には、このデータのXの分布とYの分布を保存したままで、並べ替えて対の作り方をランダムにしたデータを何個も作成し(一例を図1(b)に示す)、それらの相関係数の分布(図1(c))を用いて(分布を仮定せず)正確に検定を行うと、実際には無相関であることが判明する。あらゆる統計検定は、このような並べ替え検定により正確に行うことができる。つまり、莫大な量のランダムデータの計算に集約することで、仮定・近似を入れないで簡潔・明解に正確な分析を行うことが可能になる。値を順位に粗視化する順位相関係数などの通常のノンパラメトリック手法では、情報の削ぎ落としにより検出力が低下するが、並べ替え検定では値

をそのまま扱うため検出力の低下は生じない。このような計算集約的経済分析を用いることで、大変動・非定常・非線形な性質を厳密に扱った解析が可能になる。

計算集約的経済分析は莫大な量の計算を必要とするため、これまであまり行われてこなかった。本研究では、スーパーコンピュータによる超並列計算を用いてこれを行った。この手法では、全く同じ計算をパラメータや乱数の種だけを変えて何回も行う。そのため、基本的に各コアに独立に計算を行わせればよい。並列計算に適しており実装も容易である。

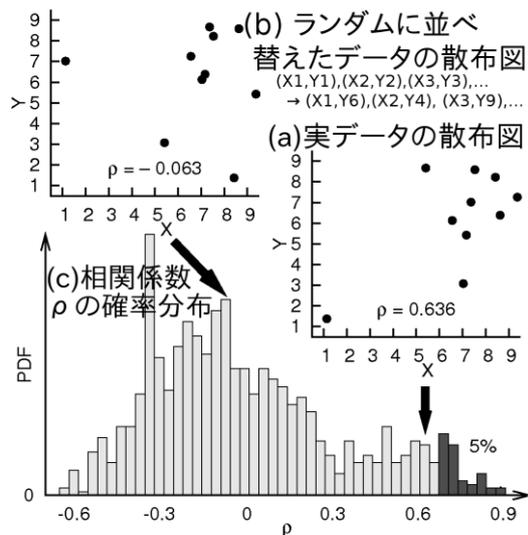


図1 計算集約的統計手法の例(並べ替え検定)

## 4. 研究成果

(1) 日本企業約100万社の取引関係・企業財務データを各ノード(企業)が多変量情報を持つ100万ノード・400万リンクの有向ネットワークとして分析した。企業間ネットワークは、同じ業種同士でつながりやすい傾向がある。そこで、業種の異質性を考慮した上で統計的有意なつながりを特定するために、各業種で次数(●←, ●→, ●⇔の数)を保存してランダムにつなぎ替えたランダムネットワークを作成した(図2, 3)。隣接する3企業間の取引関係を表現する構造(部分グラフ)に注目し、実ネットワークとランダムネットワークを比較した。その結果、3企業は図4のような構造を取りやすく、V字構造は出現しにくいことを明らかにした。さらに、各業種の経済学的機能を反映する特徴的なモチーフ構造を抽出した。

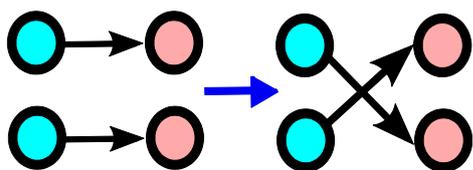


図 2 各業種で次数を保存したつなぎ替え

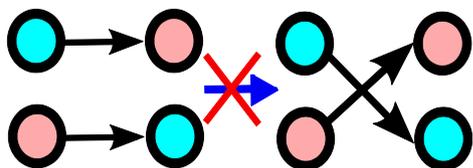


図 3 各業種で次数が保存されないつなぎ替えは行わない。

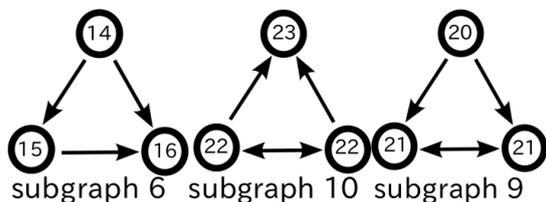


図 4 統計的に有意に出現する部分グラフ (ネットワークモチーフ)

(2) 首都圏の住宅地に関する地価公示価格のデータ(1974~2008年の年次, 地点数は年に約2000~6000地点)を年毎に分析した。対数価格は、バブル期を除き正規分布に従うことが判明した。しかし、バブル期は対数正規分布から統計的に有意に乖離しており、対数正規分布より裾の長いベキ分布(パレート分布)に従うことが分かった。つまり、価格は通常は対数正規分布に従うが、バブル期は物件間の価格裁定の及ぶ範囲(同一需給圏)が狭くなるため、ベキ分布に従う。この性質を用いて、価格分布から同一需給圏の大きさ(異質性の度合い)を推定した。ランダムに住宅地を一つ選び、その最近接住宅地を  $n-1$  個取り出し、これら  $n$  個の住宅地を一つの地域と定義するという操作を繰り返すことにより、全領域を複数の地域(一つの地域は  $n$  個の住宅地で構成される)に分割し、各地域内における価格分布を調べた。この操作による領域の分割は一意には定まらないため、乱数の種を変えて128通りの分割の仕方について分析した。その結果、近接住宅地数  $n$  を小さくして狭い領域に限定してみれば、バブル期においても価格は対数正規分布に従うことが分かった。つまり、バブル期には同一需給圏の大きさ(裁定が成立し一物一価になっている、価格が同質とみなせる地域の広

さ)  $n$  が小さく、地域の異質性が強く、各地域内では価格は同質だが異なる地域間では異質になっている。このように算出した同一需給圏の大きさ  $n$  はバブルの度合いを表す指標になると考えられる。そこで、 $n$ (全物件数)の場合からこの分析をはじめ、徐々に  $n$  の値を小さくしていくとき、はじめて価格分布が対数正規分布になる  $n$  の値(つまり、対数正規分布とみなせる最大の  $n$ )をバブルの度合いと定義した。分布が対数正規分布とみなせるかどうかは、コルモゴロフ・スミルノフ検定(有意水準5%)を用いて判定した。このようにしてバブルの度合いの年次推移を算出し、バブルの大きさを定量化することができた(図5)。

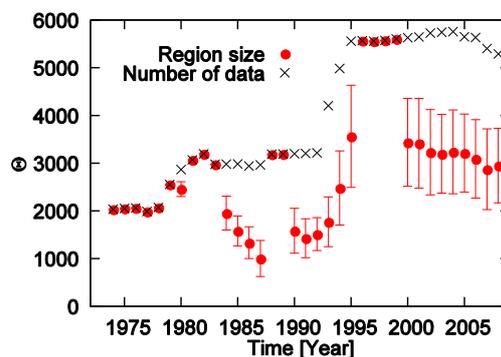


図 5 バブルの度合いの年次推移 (●). × は全地点数を示す。

(3) 首都圏の不動産の物件属性に関する多変量データについて、同じ属性間でデータをランダムに入れ替えたデータを作成し、クラスター分析などの多変量分析の手法で得られる結果を実データと比較し、統計的に有意な多変量相関を明らかにした。さらに、不動産の空間データについて、二次元空間上の物件分布の特徴を探るために、物件の緯度・経度をランダムに入れ替えたデータを作成し、さまざまな物件属性の統計的に有意な空間相関を抽出し、物件分布の空間構造を特定した。

(4) ビジネスニュースや新聞などの文字データについて、同じ品詞同士をランダムに入れ替えたデータを作成し、言語処理の手法で得られる結果の統計的吟味を行った。

(5) 市場価格変動の高頻度時系列にみられる非定常性は、ニュースのような外乱により引き起こされると考えられる。市場がニュースに反応して変動していることを調べるために、ロイター社のニュースアーカイブのデータ(2003~2011年の約2億件の記事)を言語処理の手法を用いて分析した。ニュースをその内容を表わすトピックスに分類した上で、

ニュースが株式市場の株価に及ぼす影響を解析した結果、株価のボラティリティの多くはニュースで説明つくことを明らかにした。

#### 5. 主な発表論文等

(研究代表者、研究分担者及び連携研究者には下線)

〔雑誌論文〕(計0件)

〔学会発表〕(計4件)

- ① 大西立顕, 水野貴之, 清水千弘, 渡辺努, “日本における住宅バブルの形成と崩壊の時空間構造”, 日本物理学会第68回年次大会, 2013年3月29日, 広島大学(広島県)
- ② Takaaki Ohnishi, “Spatial Heterogeneity and Real Estate Bubble”, Econophysics-Kolkata VII, 2012年11月11日, サハ核物理学研究所(インド)
- ③ 大西立顕, 高安秀樹, 高安美佐子, “ネットワークモチーフが明らかにする企業間の取引関係”, 平成24年度統数研共同研究集会「経済物理学とその周辺」第1回研究会, 2012年8月27日, キヤノングローバル戦略研究所(東京都)
- ④ Takaaki Ohnishi, Hideki Takayasu, Misako Takayasu, “Using Network Motif Analysis to Characterize Inter-firm Transactions”, NetSci2012, 2012年6月20日, ノースウェスタン大学(米国)

〔図書〕(計0件)

〔産業財産権〕

○出願状況(計0件)

名称：  
発明者：  
権利者：  
種類：  
番号：  
出願年月日：  
国内外の別：

○取得状況(計0件)

名称：  
発明者：  
権利者：  
種類：  
番号：  
取得年月日：  
国内外の別：

〔その他〕  
ホームページ等

#### 6. 研究組織

(1)研究代表者

大西 立顕 (OHNISHI TAKAAKI)

東京大学・大学院経済学研究科・特任研究員

研究者番号：10376387

研究者番号：

(2)研究分担者

( )

研究者番号：

(3)連携研究者

( )

研究者番号：