

論文の内容の要旨

論文題目

時間順序データの可視化技法とこれによる連鎖的経済・金融事象の構造把握支援

氏 名 郷田慎一

経済・金融事象で、連鎖倒産、株価変動の連鎖等の連鎖的事象の存在が実務家の間で一般的かつ重要な事象と認識されている。しかし、分析やシミュレーションの際には独立の事象として扱われていたり、一部の関係を考慮に入れるに止まっている場合が多い。これは、①データの制約から連鎖構造を直接把握することができないこと、②既存の間接的構造推定手法では少量データでの推定や事象発生の前後関係を明示的に考慮に入れることが困難であることによる。

すなわち、連鎖的事象の分析ではデータの制約から間接的な構造推定を行う必要があるが、経済学・財務会計学等における関連研究には、企業の倒産確率の相関分析[1, 2]や株式市場間の相関分析[3]のように、先験的・理論的に構築された比較的少数の構造方程式を検証するアプローチのものが主流であった。また近時、仮説に基づき構築した人工市場でのシミュレーションを通じ市場での連鎖的事象の分析を行うアプローチも発展している[4]。しかし、構造についての仮説の導出方法は明確でなく、多数の事象を対象とした未知の関係抽出を主なテーマとした研究は限られる。

他方、工学から始まり、マーケティングや社会ネットワーク研究等では、多事象間の関係をデータ・マイニング的にモデリングし、本来のモデルを探索的に推論・可視化するグラフィカルモデリング手法が多く用いられる。ブーリアンネット、ベイジアンネット[5]、KeyGraph等が代表的手法である。KeyGraphは、統計的基準のみによる構造決定ではなく、人間によるチャンス発見のための手掛かりを提供することを目指すツールである[6]。このため低頻度の事象や多数のノード間の関係を比較的少量のデータからでも抽出・可視化する。ベイジアンネットは、事象の種類が多かったりサンプル数が少ない場合の構造探索への適用は困難である。他方、KeyGraphは、事象間の共起関係でリンクを作成し、無向グラフである。共起グラフの有向グラフ化ツールには、事象の依存関係の強弱で方向付けする”Word Colony[7]がある。KeyGraphでも時間順による有向グラフ化の試みがあったが[6]、時間発生順序が因果関係の方向と必ずしも一致しないこともあり、その後は、同様の試みはなされていない。時間経過とともに次々と発生する連鎖事象の構造推定のためには、事象発生の時間順序を明示的に織り込んだ構造推定・可視化の手法が必要である。本研究では、こうした課題に対応する新たな技法を開発し、連鎖的経済・金融事象の構造把握を支援することを目的としている。

このような本研究のねらいを実現するため、筆者はチャンス発見の技法に着目した。チャンス

発見技法では人間の持つ知見とコンピュータによるデータ分析・可視化の結果を相互に利用する二重螺旋プロセスにより、予兆や新たな機会の発見を促進する。時間やデータの制約や事象の独自性から、大量データによる統計分析が行いづらいビジネスの現場での意思決定支援に有効な手法であり、社会調査プロセスへの適用[8]をはじめ、ビジネス、医療等の実務での活用事例が蓄積されている。情報社会のプラットフォームのデザイン・検証プロセスへの螺旋的方法論の適用が提起されている[9]が、チャンス発見手法は社会・政策に係る幅広い分野での知見獲得支援にも適用可能であろう。

可視化は、二重螺旋プロセスでの人間とコンピュータのインタラクティブにおいて重要な手法である。可視化は、人間のメタ認知、他者との共感や知識の共有、知識の組合せ、等のプロセスを助け新たな価値の発見・掘起し・創造を支援する。視覚情報が認知獲得につながるプロセスがアンケート調査により検証されている[10]。可視化の際に、事象間の関係や変化の方向を示すことも重要である。編年型データを対象にデータの経年変化を視覚化することで新たな発見を支援するツールも提案されている[11]。特に、連鎖構造の推定においては事象発生時間順方向の可視化は決定的に重要である。本研究では、時間順序データを可視化する技法として、新たに、データの前処理のための時間順序法と、連鎖キーグラフ(図1)を提示した[12]。

(Sは起点業種、業種コード: C:土木・建築・工事業、M,N:製造業、W:卸売業、R:小売業、T:運送業、E:不動産業)

JaJa[20-20-15]

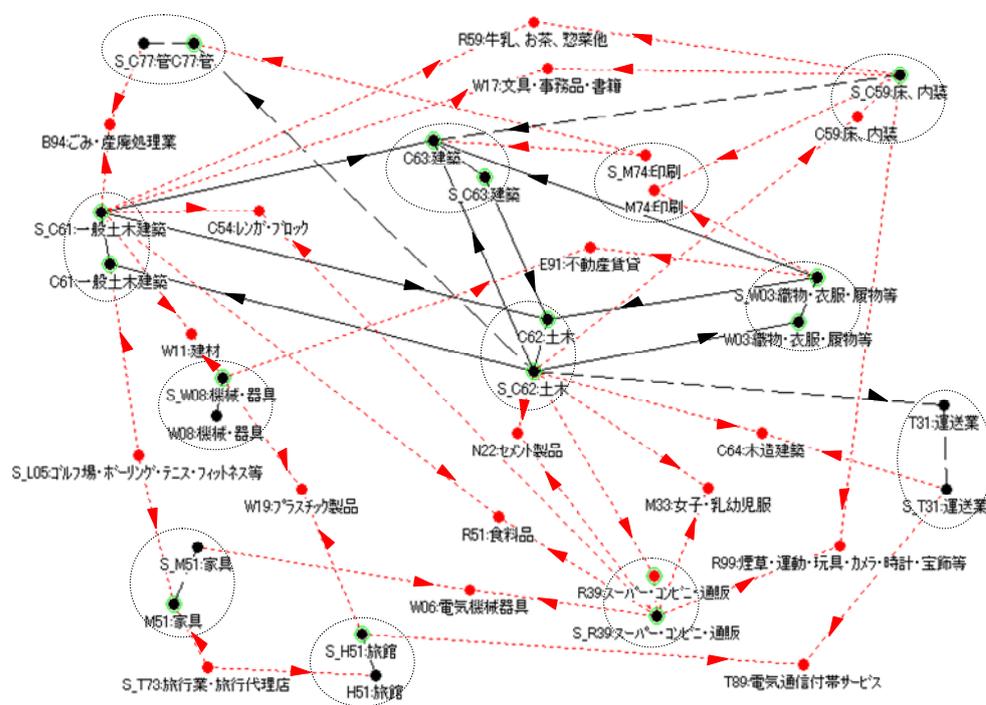


図1. 連鎖キーグラフ - 2005年の九州の連鎖倒産構造

時間順序法は、複数種類の事象の時系列データを用い複数事象間の共起関係を分析するためのデータ前処理の手法である。例えば、業種別の倒産発生の時系列データがある場合、
Step 1) ある倒産と、その後一定期間に発生した個々の倒産を1対としたデータを作成。
Step 2) 1対にした倒産のうち、先に生じた倒産に識別コード（“S”）を付す。
Step 3) 上記で作成したバスケットをキーグラフで解析し、グラフ化の際に、倒産が先に生じることの多い業種（“S”の付いたノード）から、後に起きることの多い業種（“S”の付いていないノード）に向けた矢印付きの線（エッジ）を描く。

というもので、図1のような図が描けることになる。この手法のメリットは

- ① 発現データのみで連鎖事象の構造推定を行うことができる
- ② 時間の前後を明示的に考慮に入れた構造把握ができる
- ③ 他の手法（ベイジアンネット、多変量解析等）に比べ比較的少数のデータでも構造推定を行え、稀な事象、短期間の構造変化に対応しやすい

という3点であり、冒頭に示した実務上の要請にこたえるものである。本論文では、連鎖キーグラフで連鎖倒産の構造を把握することについて、3つの視点で評価している。一つは倒産連鎖の影響予測力を検証したもの、二つ目は金融の専門家に対し新たな視点に気づく効果をアンケートによって評価したものであり、3つ目は、サブプライム問題について風説に囚われず説得力のある解釈をすべく試みたものである。

まず、倒産連鎖の影響予測力を以下の枠組みで評価した。連鎖キーグラフで把握したノードとエッジの関係をもとに、ある業種の倒産数を、その業種にエッジでつながっている起点業種で前月に発生した倒産数の関数としてモデル化し予測に使用したものである。

(例) グラフ: S_業種 10 → 業種 20 ← S_業種 30

(業種 10,30 で倒産が起きた後に業種 20 で倒産が起きる)

モデル: 業種 20 へ矢印が入る業種での倒産有無指数

= S_業種 10 での倒産有無指数 + S_業種 30 での倒産有無指数

結果として、05年の連鎖倒産構造を用いた05年実績の判別結果、同06年の予想結果とも、モデルにより倒産「あり」とし、実際に倒産「あり」となった確率は、全体(モデルを考慮しない)で倒産「あり」となった確率を上回った(各、1.74倍、1.27倍)。モデルの予測能力は相応にあると考えられる。

次に、実務家の連鎖倒産構造把握支援への適用実験を、実務的なチャンス発見の視点から行った。実務家が連鎖倒産構造を表した連鎖キーグラフを見ながら連鎖倒産が起きると思う業種ペアを選ぶアンケート調査を行い、連鎖キーグラフを見ることによる連鎖倒産構造把握の有効性の変化を見た。アンケート方法は、回答者として金融、土木業界等の実務者20人を集めた上で、回答者が連鎖倒産が起きると考える業種の組合せ(業種ペア)を、①1回目は何も見ないで、②2回目は連鎖キーグラフ(図1.2005年の九州の倒産データで作成)を見ながら、の2回、各々回答用紙に示した業種間に矢印を描く方法で選択させ結果を比較したものである。このうち1回目、2回目で選択した業種ペアが、2006年、2007年の九州の倒産データで作成した連鎖キーグラフでも、連鎖倒産が頻繁におきる業種として抽出されるか検証した。

適用実験の結果、1回目から2回目の間に、06年、07年のグラフと合致した選択ペアの件数は、06年、07年ともに増加。合致率(合致件数/選択したペア総件数)は、06年は横ばい、07年は大幅に上昇した。これは、回答者が2回目の回答時に連鎖キーグラフを見ることによって本質的な因果関係への気づきを掘り起こすことができ、未来の変化との合致率の低い業種ペアを選択から外し、逆に合致率の高い業種ペアを新規に選択したことによる。このように、連鎖キーグラフを見ながら選択を行うことで、連鎖倒産が頻繁に生じる業種ペア把握の有効性が向上することが確認できた。なお、回答者は、選択時には合致率を知らされておらず、意図的に高い合致率の業種ペアを選択したのではない。

さらに3つ目の評価として、リーマン・ショック前後の世界の主要市場の株価動向を、3つの時期に区分し、各期の株価下落の連鎖構造を時間順序技法と連鎖キーグラフを用い可視化し、シナリオを創出した[13]。リーマン・ショック以降、世界の株価・景気について、多くの場合「サブ・プライム問題→リーマン破綻→米国株下落→世界の株価下落→世界不況」という図式で語られた。しかし、この間の世界の主要株価下落の連動関係を連鎖キーグラフで把握し、主要な経済・政治イベントと合わせて解釈した場合、中国を中心としたアジア市場や原油に代表される資源市場が独自に先進国市場に先行して動きこれに日本市場が連動するシナリオにも注目する必要性があることを示した。

【参考文献】

- [1] リスク管理モデルに関する研究会「リスク管理モデルに関する研究会報告書」FISC,1999年
- [2] 橋本崇「与信ポートフォリオの信用リスク計量における資産相関について - 本邦のデフォルト実績データを用いた実証分析 -」日本銀行ワーキングペーパーシリーズ, No.08-J-10, 2008年6月
- [3] 小原沢則之「世界株式市場要因がポートフォリオを支配する」証券アナリストジャーナル, 2009年5月, pp41-50
- [4] 和泉潔, 松井宏樹, 松尾豊「人工市場とテキストマイニングの融合による市場分析」人工知能学会誌, Vol.22, No.4, pp397-404, 2007年
- [5] 繁樹算男, 植野真臣, 本村陽一『ベイジアンネットワーク概説』培風館, 2006年
- [6] 大澤幸生『チャンス発見のデータ分析』東京電機大学出版局, 2006年
- [7] 赤石美奈「文書群に対する物語構造の動的分解・再構成フレームワーク」人工知能学会論文誌, Vol.21, No.5, pp.428-438, 2006年
- [8] Y.Ohsawa, Y.Nara “Decision Process Modeling across Internet and Real World by Double Helical Model of Chance Discovery” New Generation Computing, 21(2003)109-121, Ohmsha and Springer
- [9] 國領二郎「情報社会のプラットフォーム: デザインと検証」情報社会学会誌, Vol.1 No.1 pp41-49, 2006年
- [10] Y.Ohsawa, Y.Maeda “Eyes Draw Auxiliary Lines in Interpreting Images” Journal of Computers, Vol. 4, No. 10, Oct. 2009
- [11] 赤石美奈, 伊藤直之, 箱石大, 石川徹也「編年型データ解析ツールの開発」第24回JSAI全国大会, 2010年
- [12] S.Goda, Y.Ohsawa “Estimation of Chain Reaction Bankruptcy Structure by Chance Discovery Method - with Time Order Method and Directed Keygraph” Journal of Systems Science and Systems Engineering, Vol.16, No.4, December 2007, Springer
- [13] 郷田慎一, 大澤幸生「近年の世界の株価の連鎖構造分析」社会・経済システム, 第31号, 2010年