

経済因果チェーン検索のシステム紹介と応用

Economic Causal Chain Search System and its Application

和泉 潔^{*1}
Kiyoshi Izumi

坂地 泰紀^{*1}
Hiroki Sakaji

^{*1} 東京大学大学院工学系研究科 システム創成学専攻
Department of Systems Innovation, School of Engineering, the University of Tokyo

In this research, we introduce a system for searching causal relationships of events related to economics and finance in a chain-like manner from a database extracted from economic text data. This system also introduces service application ideas such as inferring economic ripple effects from events entered by users and presenting related stocks tracing the causal chain from news articles.

1. はじめに

経済ニュース記事を読んでいると、株価の動きや商品の売り上げ、雇用や貿易など様々な経済事象に関する原因と結果の関係の記述がよく出てくる。例えば「将来の少子高齢化が〇〇を引き起こし、それに関連した〇〇に関する需要が増えることが予想される」とか「現在の株価下落は〇〇による市場参加者のリスク警戒心理の高まりを反映している」など、ある事象の経済的な波及効果やその事象の原因について述べている。

しかし経済現象のように人間の行動が関係する事象の因果関係を、数値データだけから統計的に解析することは難しい。なぜなら、人間が原因事象をどのように認識して、それに対してどのような行動を取るのかという行動ルールが因果関係の鍵となっているからである。時と場合によって因果関係が変化する。自然科学的現象のように数値データの統計分析で、客観的かつ普遍的な因果系列を取り出すことはほぼ不可能である。

そこで、本研究では人間が認識した因果関係が含まれていると思われる経済テキストデータを解析し、経済分野に関わる因果関係のデータベースを構築する。さらに、特定の事象を表すフレーズから派生する因果系列を検索する手法を開発する。本手法を用いて、ユーザが入力した語句に対して因果系列を表示し、ユーザが適切な系列に選択したり適切でない系列を削除したりできるシステムを実装した。本システムで作成された因果チェーン、経済的な波及効果と要因列挙が可能となる。

本稿では、因果チェーン検索システムと、本システムやアルゴリズムを用いた応用例を紹介する。

2. 経済的因果関係の抽出とデータベース構築

人間が認知した因果関係について記述されたテキストデータを解析し、そこから因果関係を抽出する技術が必要となる。本システムでは、決算短信という上場企業が業績や財務状況を開示するために定期的に発行しているテキストから、手がかり表現を用いた手法[坂地 2015]により因果関係を抽出した。

- 使用したテキストデータ: 2012 年 10 月から 2018 年 5 月に約 2,300 社が発行した約 2 万個の決算短信テキスト
- 抽出した因果関係: 1,078,542 個

抽出された因果関係は、決算短信の発行された日付、決算短信を発行した銘柄などと一緒にデータベースに保存される。

3. 因果チェーンの構築アルゴリズム

上述の経済因果データベースから、特定の語句から関連する因果系列を構築する手法には、[西村 2018][Nishimura 2018]を基に改良した手法を用いた。具体的には以下の 4 ステップからなる(図 1)。

1. ユーザがフレーズを入力する。最初は終端ノードを入力テキストとする。
2. 終端ノードの表現と一定の類似性がある因果関係を経済因果データベースから選択し、ネットワーク追加候補とする。
3. Step 2 で抽出したネットワーク追加候補と終端ノードの組み合わせについて因果関係ノード間の類似度を計算する。因果関係ノード間の類似度が閾値 α 以上あるときにノードをネットワークに追加してネットワークを拡張する。
4. Step 3 でネットワークに追加したノードを終端ノードとして Step 2, Step 3 を繰り返す。

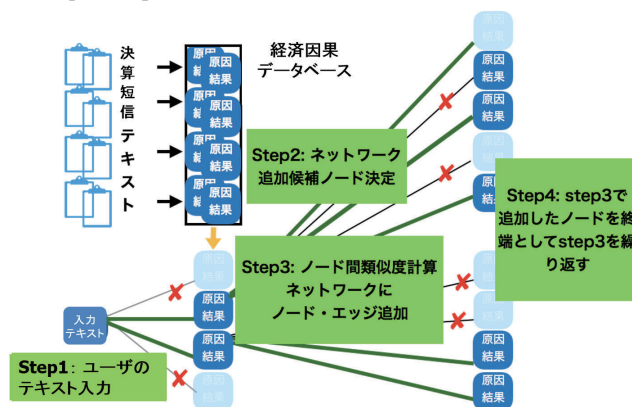


図 1. 因果チェーン構築のアルゴリズム
([西村 2018] [Nishimura 2018]を基に作成)

4. 経済因果チェーン検索システム

上述の因果チェーン構築アルゴリズムを基に、ユーザが提示された因果関係を編集できる機能を加えた経済因果チェーン検索システムを実装した[坂地 2018]。以下に本システムの動作の様子を説明する。

最初に開始テキストをユーザが入力する(図 2). その際に, 原因から結果への波及効果の検索か結果から原因への要因検索かを指定する. さらに, 因果チェーンを構築する対象となる因果関係が発行された期間を限定することも可能である.

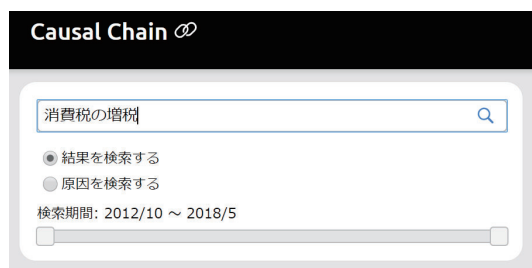


図 2. 開始テキストを入力するテキストボックス

テキストボックスの右にある検索ボタンをクリックすると, 開始テキストに連鎖する因果関係を表示する(図 3). デフォルトでは関連性の高い順に 3 個の因果関係を表示する. より多くの因果関係を見たいければ, 「もっと見る」ボタンをクリックし因果関係ノードの表示を増やすことが出来る. 因果関係の各ノードに対して, ユーザが適切な因果でないかと判断した場合は, 各ノードの右上にある削除ボタンを押して, ノードを削除できる.

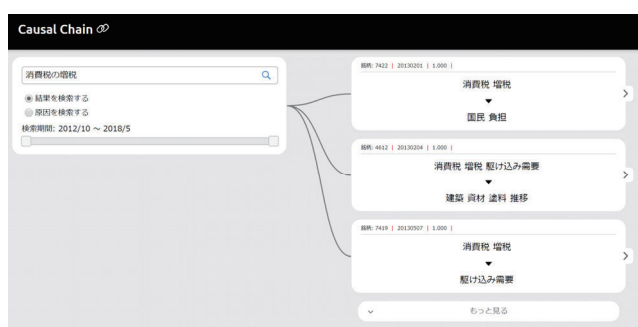


図 3. 連鎖する因果関係の表示

各ノードからさらに因果チェーンを伸ばしたい場合は, 各ノードの右側にある「>」ボタンをクリックすると, クリックされたノードを終端ノードとして関連する因果関係が追加される(図 4).

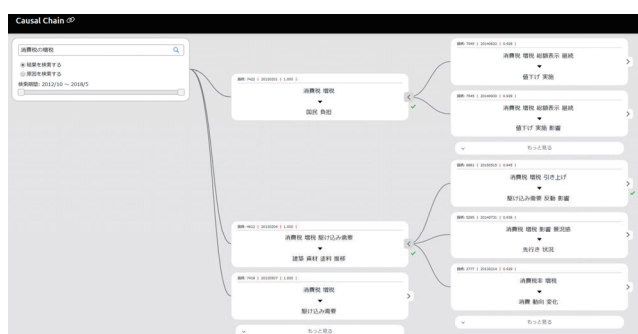


図 4. 因果チェーンの追加

上記の因果チェーンの構築作業を繰り返し, ユーザが必要とする因果チェーンを構築できたら, 構築した因果チェーンをファイルに保存することが出来る.

5. 経済因果チェーン検索システムの応用例

経済因果チェーン検索システムや因果チェーン検索アルゴリズムを応用したサービスの例を以下に示す.

5.1 個人投資家向けサービス

一般の個人投資家, 特に資産運用を始めたばかりの個人投資家には, 資産運用を判断するための情報が難解であることがある. この難解さの原因の一つとして, 自分たちが日常生活で得てきた知識と金融の専門的な知識との間に大きなギャップがあることである. 日常的な事象から金融市場動向へは, いくつかの経済事象の因果関係の連鎖が存在する. 提案手法により, このギャップを埋めるような知識を提供するサービスを実装できる.

(1) ニュース内容の波及効果と背景情報の提示

ユーザが日常的に見るニュース記事の中に気になる語句があれば, そこを指定しその語句から派生して関連する経済事象を本手法により提示することができる(図 5). 波及効果の中には, ある企業の業績や経営状況に直接関わる事象も含まれるので, ユーザが様々なニュースから関連銘柄を検索するサービスを構築することが可能である. また, 逆に指定した語句の要因となり得る事象を検索してユーザに表示することもできる.

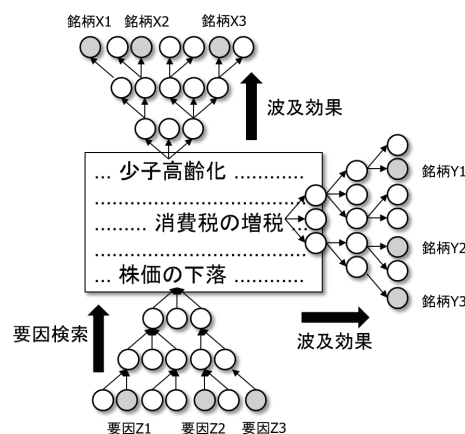


図 5. ニュース内容の波及効果と背景情報の提示サービス

(2) 質問応答システム

金融機関が個人投資家向けに提供する対話エージェントサービスに提案手法を応用可能である. 個人投資家があまり金融専門知識を有していない場合, 対話システムのエージェントに対して, 質問をしてくる場面が多くなると考えられる. 自分の経験や興味に関連する一般的な話題を切り口にして, 関連する経済事象や銘柄を質問してきた場合に, 提案手法により関連する事象や銘柄と, なぜ関連すると思われるかを答えることができる(図 6). また, 事象を引き起こした要因も返答できる.

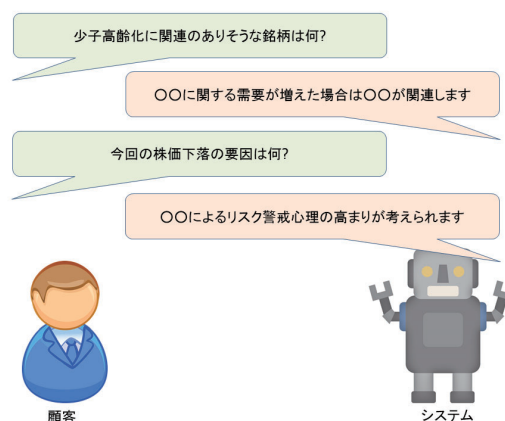


図 6. 個人投資家向けの質問応答システム

5.2 金融実務家向けの業務支援

提案手法により、アナリストや営業担当者など金融機関の様々な部門の実務者の業務支援システムを構築することができる。

(1) レポート作成支援

金融機関のアナリストは、顧客向けに特定の銘柄や市場全体の動向に関するレポートを定期的に発行する。提案手法により、アナリストがレポートを作成する際のコンテンツを決めることを支援することができる。例えば、ある事象から解説対象の市場に対して何か波及効果がなかったかを検索し、レポートにこの事象を書くべきかどうか決める。また、ある市場の動きに対して、潜在的な要因を検索し、レポートに書くべき要因に漏れがないかどうかをチェックできる。他にも、ある事象 X が別の事象 Y を引き起こすとレポートに書いた場合に、事象 X から事象 Y への因果系列を検索して、因果の飛躍がないかどうかチェックすることができる。途中にある因果系列のステップ数が長い場合には、途中の因果の一部をレポートに追記して、読者に理解しやすいようにする。

(2) セールス支援

上述の個人投資家向けの質問応答システムと同様の機能で、金融機関の営業担当者が顧客にセールスを行う際に、顧客の興味や関心と関連する銘柄を検索することができる。顧客の属性や興味に関連してあらかじめ関連銘柄を検索しておけば、営業活動の支援となり得る。また、顧客からの質問に対しても、上述の質問応答システムで、返答すべき内容の候補を得ることができる。このような対面営業以外でも、オンライン上での資産運用アドバイスで、基本的な相談であれば本手法により自動化を支援することができる。

5.3 多種の因果データベースとの連鎖

決算短信以外のテキストデータから因果関係データベースを構築し、多様な分野の因果データベースを組み合わせることによって、以下のような解析や応用をすることができる。

(1) 業務報告書・社内文書

各企業には、通常業務において担当者が作成した社内文章が存在する。例えば、金融機関では営業部門や融資部門が対応した顧客に関して、業務報告書や稟議報告書などを作成している。また、アナリストやトレーダーなどの市場の専門家が、作成した相場見通しに関する定期的な分析レポートもある。このような今までの業務で蓄積された社内文章には、有用な業務知識や専門知識が多く含まれおり、大量に集めて解析することによりデータを保有する独自の知識として高い価値を有することがある。例えば、地域の金融機関の業務報告書を機械学習手法で解析し、その地域の詳細な経済活動指標を作成した例がある[坂地 2019]。同様に、本提案手法を用いて、その企業しか持っていない社内文章から因果データベースを構築し、決算短信などの一般的な文章からの因果データベースと組み合わせれば、その企業の独自知識を活用した新たな因果系列を探索できる。

(2) 特許文書

金融以外の分野でも、有用な因果データベースを構築できるテキストデータは多数ある。例えば特許文書の中には、特定の技術がどのような効果をもたらすのかという因果関係の情報が多く含まれている。特許文書から構築した因果データベースを用いて、因果チェーンを探索すれば、複数の技術の組み合わせによりどのような効果が期待できるかということを見つけること

ができる。また、特許文書の因果データベースと経済分野の因果データベースを合わせれば、技術の経済的波及効果や技術に対するニーズを調べることができる。

(3) ミッシングリンクの発見

上述の経済や技術分野だけでなく、より一般的な分野での因果関係をニューステキストなどから構築し、これらの因果データベースを組み合わせれば、現在は存在していない重要な因果関係を探索することができる。現状 X から開始した波及効果 $\{x_1, x_2, \dots, x_n\}$ を検索する。次に、達成したい状況 Y から因果関係を遡って要因 $\{y_1, y_2, \dots, y_m\}$ を検索する。最後に、 x_i と y_j の組み合わせについて、 x_i を原因として y_j が結果となるような現象や技術がないかを検討してみる。こうすることによって、目標状況 Y を達成するために、現在は存在しないが重要な技術や現象を発見することを支援できる。

6. まとめ

本研究では、経済・金融に関わる事象の因果関係を、経済テキストデータから抽出したデータベースから連鎖的に検索するシステムの紹介をした。本システムにより、ユーザが入力したイベントからの経済的な波及効果を推論したり、ニュース記事からの因果チェーンをたどった関連銘柄を提示したりするなどのサービス応用のアイデアを紹介した。

謝辞

本システムを実装するにあたって、日名 洋平 氏と松田 健児 氏に多大な助力をいただいた。ここに深謝の意を表する。

参考文献

- [坂地 2015] 坂地 泰紀, 酒井 浩之, 増山 茂: 決算短信 pdf から原因・結果表現の抽出, 電子情報通信学会論文誌 D, vol. 98, no.5, pp.811-822, 2015.
- [西村 2018] 西村 弘平, 坂地 泰紀, 和泉 潔: ベクトル表現を用いた因果関係連鎖の抽出, 人工知能学会 第 20 回金融情報学研究会, pp.50-53, 2018 年 3 月.
- [Nishimura 2018] K. Nishimura, H. Sakaji, K. Izumi: Creation of Causal Relation Network using Semantic Similarity, 第 32 回人工知能学会全国大会, p. 1P104, 2018 年 6 月.
- [坂地 2018] 坂地 泰紀, 和泉 潔, 松島 裕康: 金融テキストマイニングの基づいた投資家支援プラットフォームの開発, 人工知能学会 第 21 回金融情報学研究会, 2018 年 10 月.
- [坂地 2019] 坂地 泰紀, 和泉 潔, 松島 裕康, 川瀬 和哉, 林 寛, 接触履歴を用いた地域景況インデックスの自動生成, 日本知能情報ファジィ学会誌 (to be appeared)