

数理モデルを用いたフィッシャーの投資原則 に関する研究

中村大樹[†] 施建明^{††}

投資家フィッシャーの投資原則をサポートベクトルマシン等の数理モデルとAHPや飽和集合を用いて分析を行い、彼の投資原則の有用性を検証する手法を提案していく。

1. はじめに

日本経済は1990年のバブル崩壊後から失われた20年と言われるように経済成長はほとんどしていない。2011年度の4-6月期のGDPの実質成長率が年率換算で1.3%減となり、三四半期連続でマイナスとなった。2010年度日本のGDP実質成長率が3.94%であり、世界各国182カ国で93位であった。2011年度は日本のGDPの実質成長率は当初政府見通し1.5%だったが、震災の影響から0.5%に下方修正された。昨年の各国のGDPの数値から推察するに、先進諸国では最下位となる可能性が高い。

国内市場の拡大と共に国内企業も成長していった。しかし今後日本のGDPが伸びる可能性は薄く、国内市場は縮小傾向となる。この縮小される日本の市場で日本企業は成長しなくてはならない。

企業の成長は多くの要因が複雑に絡まって起こる。それらの主要要因を見つけ、定量的に分析することは難しい。そこで本研究では企業の成長するための主要因を明らかにするために、フィッシャー(Philip Arthur Fisher:1907-2004)の投資理論、サポートベクターマシン、飽和集合列挙集合、AHPを用いて、何が企業の成長要因になるのかを検証する手法を提案する。

2. サポートベクトルマシン

サポートベクターマシン(Support Vector Machine:SVM)は1960年代に[Vapnic]ら[4]によって提案された超平面による特徴空間の分割法から始まり、ここから線形サポートベクターマシン(Linear Support Vector Machine:LSVM)が誕生した。線形サポートベクターマシンは線形分離可能である場合は良い認識率を達成することが出来るが、全ての問題を線形分離可能であることはない[2]。

一般的には線形分離が不可能である場合が多い。線形分離不可能であるときの対応として、非線形サポートベクターマシンがある。非線形サポートベクターマシンの主な手法にソフトマージン法と写像を利用するものがある。ソフトマージン法は識別面よりも他群への侵入を許可するものである。写像を利用するものは写像 $\phi:R^n \rightarrow H$ を用いて、サンプルデータをより高次元な H に写して、 H 上で線形分離を行うものである。しかし写像を用いた場合、空間 H の次元が大きいと計算時間がかかる。そこで $k(x_i, x_j) = \phi(x_i) \cdot \phi(x_j)$ となるような、カーネル関数 K が存在するならば、カーネルトリックと呼ばれる手法により計算量を少なくすることが出来る。

提案する手法では、検証する成長要因の個数で次元が変化するが、フィッシャー成長要因が15個であることから、彼の投資理論から考えると次元の最大数は15次元となる。飽和集合列挙を始める段階で使用する場合は10個以下の次元となる可能性もある。

数値として初めから判断することが出来るフィッシャーの成長要因はある程度きれいに分離できることが予想される。成長要因の判断基準が数字ではなく、判断する人の主観に任される部分はAHPを用いて数値化する。

全体として数値で判断することが出来る部分は多いため、[成長できる企業]と[成長できない企業]との振り分けることは成功すると予想される。項目ごとの振り分けも行うため、どの項目が全体の振り分け結果とかなりしているかも判断できる。

3. 飽和集合

飽和集合列挙とはデータベースから特徴的なパターン・ルールを列挙する方法である。頻出集合の情報を失わずに、不要な頻出集合を排除し解を絞り込むことが出来る。すべての頻出集合が「その頻出集合がどのトランザクションに含まれているか」に注目してグループ分けをする。すると各グループの極大元は必ず唯一に定まる。唯一に定まった極大元を飽和集合と呼ぶ。

飽和集合を全て列挙すれば、各グループから1つずつ代表を列挙したことを意味し、頻出度合い、および含まれるトランザクションの組に関しては、全ての可能性を網羅することができる。以下に例を示す[3]。

トランザクションデータベース

(このデータベースは1つの項目が1つの行に対応し、各行に書かれた数値がその項目に含まれるアイテムに対応する)

[†] 室蘭工業大学 情報電子工学系 1923074@gmail.com, ^{††} shi@mmm.muroran-it.ac.jp

T1 [1 2 3 4 5 7]
T2 [4 5 6 7]
T3 [2 3 4 5]
T4 [2 4 5 7]
T5 [1 2 3]

T1, T2, T4 に含まれるアイテムの集合で最も大きなものは {4 5 7} であり、これが飽和集合である。この飽和集合がこのグループにおいての最大の特徴的なパターンとなる。

与えられたデータベース T とサポート値 $\theta(\lfloor T(P) \geq \theta \rfloor)$ を満たす) に対して頻出飽和集合を全く漏らさず、かつ重複なく列挙する問題を頻出飽和集合列挙と呼ぶ。この問題を解くアルゴリズムに極大クリーク列挙アルゴリズムから派生した出力数多項式時間アルゴリズム (LCM) や頻出集合に列挙を基礎とするものがある [3]。

飽和集合列挙を利用することで、フィッシャーの成長要因と実際成長している企業の要因との関係を見つけることが出来る。またフィッシャーの成長要因において特徴的に似ているものを抽出して絞り込むことも出来ると考えている。要因を絞り込むことで、サポートベクターマシンの計算量を減らすことが出来る。飽和集合を使う箇所を変える事で成長要因の検証に関して、様々な視点を持つことが出来るとも考えられる。

4. 研究

フィッシャーの成長要因は全部で 15 項目存在する [1]。フィッシャーの企業の成長要因が実際に成長している企業の要因とどれだけ共通性を持つか、または違うものになるかの検証手法を提案する。

- Step 1. 企業のデータベースを作成する。
- Step 2. 企業のデータベースだけでは指標が足りない場合、AHP を使い、教師データを作成する。
- Step 3. step 1 と step 2 から得られたデータを用いて SVM で、成長する企業と成長しない企業に振り分ける。
- Step 4. 分けられたグループ内の特徴を飽和集合列挙で抽出する。
- Step 5. 抽出された特徴とフィッシャーの 15 項目との共通性とかいり性を探る。

step5. の説明

フィッシャーの投資原則を判断基準に企業のデータベースを用いて、SVM で 成長する企業と成長しない企業に振り分ける。そこで振り分けられた結果と企業のデータベースからのみ得られた結果が似たようなグループ分けであれば、フィッシャーの投資原則が成長企業の要因となることが検証できる。また得られた結果が異なるものであれば、どの要因が実際の結果とかいりしているかを判断することが出来る。そのためフィッシャーの成長要因 15 項目を個別に検証する必要がある。

5. 今後の展開

成長している企業の要因とフィッシャーの成長するための要因 15 個がどのぐらい共通性を持つか、大量の企業データから今回の提案手法の有用性を検証する。

この問題は線形分離が不可能である可能性が高いため、その対応として非線形サポートベクターマシンで考える必要がある。フィッシャーの成長要因だけに関して飽和集合列挙を用いれば、彼の投資原則において本質的な成長要因を見つけることも出来る。

またサポートベクターマシンによって成長する企業、成長しにくい企業など振り分けられた場合、属している位置から適切な戦略の指針となるものを提案する必要性がある。

フィッシャーの投資原則において、今回の提案手法の有用性を検証できれば成長に関してすでに提案されている別の要因も検証することもできる。これにより最も適切な成長要因の組み合わせを発見することができる可能性もあり、企業の成長要因の複雑さを解消する新たな手法となる可能性がある。

参考文献

- [1] フィリップ・フィッシャー：フィッシャーの「超」成長株投資、フォレスト出版、2000。
- [2] ジョンソンテイラー：サポートベクターマシン入門、共立出版、2005。
- [3] 宇野毅明、有村博紀：飽和集合列挙アルゴリズムを用いた大規模データベースからのルール発見手法、統計数理 53 (2), 317-329, 2005。
- [4] 山口貴大：サポートベクターマシンによる倒産予測、2001。