

令和 6 年 4 月 19 日現在

機関番号：13901

研究種目：挑戦的研究（萌芽）

研究期間：2020～2023

課題番号：20K20755

研究課題名（和文）パターン認識技術を用いて財務データから企業の定性的属性を読み取る技法の開発

研究課題名（英文）Developing a System for Reading Qualitative Attributes of Firms from Financial Data with Pattern Recognition Method

研究代表者

小沢 浩（OZAWA, HIROSHI）

名古屋大学・経済学研究科・教授

研究者番号：40303581

交付決定額（研究期間全体）：（直接経費） 4,000,000円

研究成果の概要（和文）：財務諸表は企業の健康診断書とも言われ、その数字から企業の特徴を知ることができるとされている。そこで、パターン認識の技術を使って、財務諸表の数字から企業の定性的属性を判別に挑戦しようというのが本研究の目的であった。パターン認識の方法としては、品質工学におけるタグチ・メソッドのうちRT法を用いる。識別する定性的属性としては、まずは業種の特定を試みた。業種の特定に成功した後は、債券格付け、戦略などの識別にも挑戦する予定であった。

研究成果の学術的意義や社会的意義

1966年から2004年までの財務データを用いて、様々な業種の識別に挑戦した。その結果、商社、百貨店、電力、大手私鉄、造船、カメラの6業種については、各業種の特徴が明確で、概ね識別に成功した。ただし、期間について、1966年頃には明確に識別できているものが、年を追うごとに業種の区別が曖昧になってしまい、識別の精度が低下してしまう。理由としては、企業の体質自体が変化して業種間の違いが縮小していること、会計制度の変更により財務諸表から企業の属性の違いが読み取りにくくなっていることが考えられる。より精度を高めることに挑戦してきたが、決定的な分析手法を確立するには至っていない。

研究成果の概要（英文）：Financial statements are said to be a health certificate of a company, and it is believed that the characteristics of a company can be learned from its figures. Therefore, the purpose of this study was to try to identify qualitative attributes of a company from its financial statement figures by using pattern recognition techniques. As a method of pattern recognition, the RT method of the Taguchi method in quality engineering is used. As for the qualitative attributes to be identified, we first attempted to identify the industry. After successfully identifying the industry, we planned to try to identify bond ratings, strategies, and so on.

研究分野：会計学

キーワード：経営分析 財務諸表 パターン認識

科研費による研究は、研究者の自覚と責任において実施するものです。そのため、研究の実施や研究成果の公表等については、国の要請等に基づくものではなく、その研究成果に関する見解や責任は、研究者個人に帰属します。

1. 研究開始当初の背景

財務諸表（貸借対照表・損益計算書）には、企業の性質が反映されていると言われている。例えば、「製鉄業は固定資産の比率が高い」「差別化戦略をとる企業は売上総利益率が高い」「薄利多売型の企業は総資本回転率が高い」などである。そうであれば、逆に、経営分析指標の特徴から、財務諸表の持ち主企業を特定することもできるはずである。つまり、製鉄、自動車、商社などの業種ごとに教師データ（標本データ）を用意しておいて、企業名や業種を伏せた「未知の財務諸表」をそれと比較することで、伏せられた業種を特定したり、さらには、業種に限らず、企業の戦略、市場におけるポジション、財務的健全性などの様々な属性を読み取ったりすることができるかもしれない。このことを、品質工学で用いられるパターン認識の技法（MT システム）を用いて試してみようというのが本研究の背景にある考えである。

2. 研究の目的

本研究の目的は、品質工学で用いられるパターン認識の技法（MT システム）を用いた、新しい財務データ分析の方法を確立し、これによって、従来はテキスト情報などによる定性的情報に基づいて判断していた企業の諸属性を、財務データから読み取ることが可能にしようとするのであった。近年の会計分野では、非財務情報を定量化する研究があるが、本研究はその逆である。

本研究で仕様する MT システム（マハラノビス・タグチ・システム）とは、田口玄一氏がマハラノビス距離（点と点の距離を単純に計算するユークリッド距離ではなく、点と点の相関関係を考慮した特殊な距離）の計算方法を発展させたものである。MT システムには様々な技法が包含されているが、本研究では、その中の T 法⁽³⁾（=RT 法）を用いる。元来は、工業製品の品質不良を発見するための方法であったが、手書き文字の識別や、特殊健康診断などの分野にも広く応用されている。

その概要は、次の通りである。まず、教師データとなる複数のサンプルを用意して、その特徴量を取り出す。このデータの集合を、「標本データ」「単位空間」という。一方で、属性が未知の（評価したい）サンプルの特徴量を取り出す。これを「信号データ」という。特徴量のデータは計算によって、 β 値、SN 比(η)という 2 つの数値に集約される。これらから、標本データと信号データの距離（マハラノビス距離）を計算し、それによって両者の類似性を判定する。MT システムでは、文字認識と同じ要領で、波形の特徴をデータ化し、その類似性を評価することもできる。そのため、財務数値の時系列の変動についても、その類似性を、人間の視覚に頼るのではなく、数値的に評価することが可能になる。このことは、例えば、財務データから、企業の成長や衰退のパターンを読み取ることや、キャッシュフローや売掛金残高の四半期毎の変化から財務的健全性を評価することなどに応用できる。

3. 研究の方法

本研究課題の核心は、「正常なサンプル（企業群）から標本データを取り出すこと」である。これはさらに 2 つに分けられる。1 つは、「正常」と定義される企業群を選抜することである。例えば、業種を識別するためには、誰が見ても製鉄業、自動車製造業、商社と認識できる企業を選抜し、戦略を識別するためには、誰が見ても同じ戦略をとっていると認識できる企業を選抜する必要がある。企業を分析対象とする場合、この選抜は非常に困難であることが予想される。分析対象が製品や人間であれば、正常な製品や健康な人間の状態についての共通認識が形成されているが、企業の場合にはサンプルの個性のばらつきが大きいいため、何が正常であるかについての共通認識がないからである。しかし、幸いなことに、我が国には非常に多くの企業があり、それらのデータへのアクセスも容易になっている。また、分類の切り口となる属性も、無数に考えることができる。本研究では、その中から、同質で正常とみなすことのできる企業群や属性を探し出そうとした。

もう 1 つは、財務諸表の様々な項目、例えば、売上高、売上総利益、営業利益、経常利益、総資産、流動資産などの中から、どの項目を特徴量として取り出すかを定めることである。これについては、財務諸表分析に関する豊富な先行研究を参考にして、ある程度の目星を付けることができていた。ただし、先行研究とは分析の方法が異なるため、多様な組み合わせ、様々な前処理の方法などについて、膨大な試行錯誤を繰り返さなければならなかった。また、先行して行ったパイロット的な分析から、例えば、1965～1999 年までは、財務データから商社・自動車・鉄鋼の 3 業種を識別することができたが、2000 年以降は識別できなくなっているなど、財務データが変質していることも分かっていた。こうした問題を回避する方法も検討しなくてはならなかった。

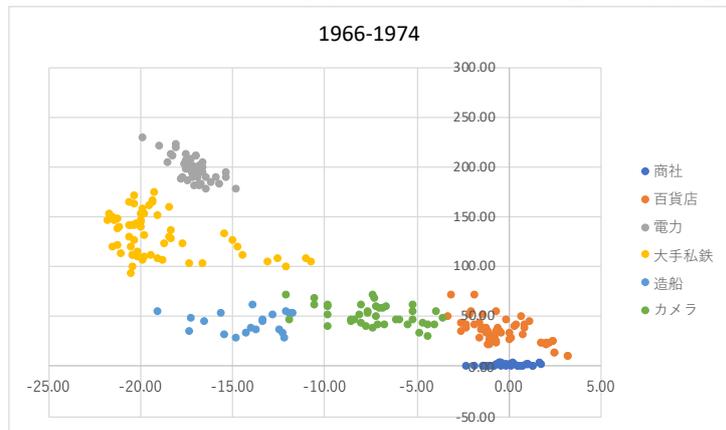
4. 研究成果

本研究では、まず、財務諸表の数値から企業が属する業種を識別することを試みた。日経 NEEDS Financial Quest の財務データベースを使って、東証業種分類の各業種に属する企業のうち、1966 年から 2004 年までのデータが揃っている企業を分析対象とした。また、試行錯誤の結果、財務諸表の項目のうち、棚卸資産、有形固定資産、無形固定資産、売上高、営業利益を特性値として取り上げた。これは、棚卸資産が取り扱い品目の影響を受けているであろうこと、有形固定資産が製造設備などの投資の大きさに影響を受けているであろうこと、無形固定資産が特許やブランドなどの開発投資に影響を受けているであろうこと、売上高が企業活動の規模を表していること、営業利益が企業の収益性を最もよく反映しているであろうことを想定して選択した結果である。なお、これらの数値は、貨幣価値変動の影響を受け、期間比較に影響がでるため、試算表の貸方項目である「負債+自己資本+収益」を分母とした比率に変換して用いた。

計算の結果、業種によって、業種区分内の企業の類似性の高低、他の業種との識別性の高低において違いがみられた。本研究は、挑戦的萌芽研究という、いわばパイロット的な研究であることから、全ての業種を同列に扱うことはせず、まずは、業種区分内の企業の類似性が高く、また、他の業種との識別性が高い業種を選定し、商社、百貨店、電力、大手私鉄、造船、カメラの各業種を取り上げ、これらの企業の識別精度を高めようとした。

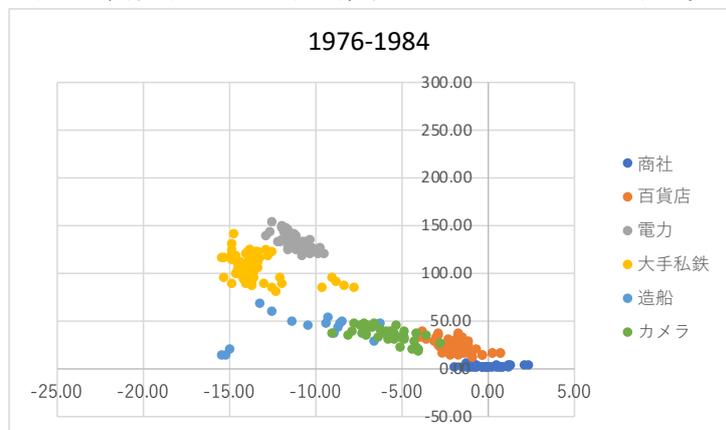
上述の業種の中では、商社と電力が最も顕著に識別されるため、商社のデータうち、1966 年から 1974 年までの偶数年のデータを標準データとした。また、同じ期間の各業種に属する企業のデータを信号データとした。このときの識別状態を表したものが図表 1 である。

図表 1 商社(1966-1974)を標準データとしたときの識別状態



続いて、商社のデータのうち、1976 年から 1984 年までの偶数年のデータを標準データ、同期間の各企業のデータを信号データとして分析を試みた。このときの識別状態を表したものが図表 2 である。図表 1 と比べると、百貨店とカメラ、カメラと造船の区分が曖昧になり、識別の精度が落ちていることが分かる。また、業種ごとに点が凝集していることから、各業種内の企業の類似性が高くなっていることが分かる。

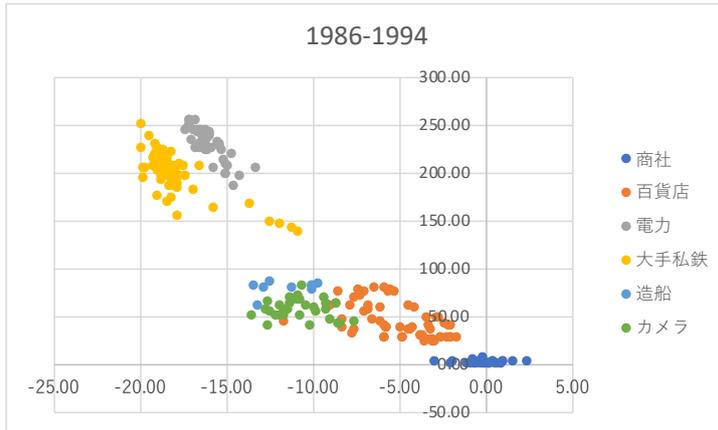
図表 2 商社(1976-1984)を標準データとしたときの識別状態



同様に、1986 年から 1994 年の偶数年の商社のデータを標準データ、同期間の各企業のデータを信号データとして分析した結果が図表 3 である。カメラと増減の区分が一層曖昧になって識別精度が落ちていることが分かる。また、特に百貨店において点の凝集性が小さくなり、企業

ごとに業績に違いが生じていることがわかる。

図表3 商社(1986-1994)を標準データとしたときの識別状態



この他に、標準データとする業種や期間を変えたり、特性値とする財務諸表項目を変えたり、分析の前処理方法を変えるなど、様々な工夫をしたが、上述の分析より良い結果を得ることはできなかった。ただし、条件設定には無数の組み合わせがあることから、本研究の成果がこの分析方法の限界というわけではなく、可能性はまだ残されている。引き続き、いろいろな条件設定を工夫することが今後の課題である。

なお、ここまでの研究の経験から、一般に言われるように、「財務諸表の数値には企業の性質が反映されている」というのは、間違いではないが、極めて限られた範囲、あるいは、特徴が顕著な業種間でのみ当てはまる命題であるといえる。また、原因は明らかでないが、古い財務諸表ほどこの命題が当てはまり、最近の財務諸表になるほど、この命題が当てはまらなくなる傾向がある。

5. 主な発表論文等

〔雑誌論文〕 計1件（うち査読付論文 0件 / うち国際共著 0件 / うちオープンアクセス 0件）

1. 著者名 小沢 浩	4. 巻 73(8)
2. 論文標題 会計時評：多元的経営分析指標 - 財務諸表の3項目以上の比率を計算する	5. 発行年 2021年
3. 雑誌名 企業会計	6. 最初と最後の頁 128-129
掲載論文のDOI（デジタルオブジェクト識別子） なし	査読の有無 無
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

〔学会発表〕 計0件

〔図書〕 計0件

〔産業財産権〕

〔その他〕

-

6. 研究組織

氏名 (ローマ字氏名) (研究者番号)	所属研究機関・部局・職 (機関番号)	備考
---------------------------	-----------------------	----

7. 科研費を使用して開催した国際研究集会

〔国際研究集会〕 計0件

8. 本研究に関連して実施した国際共同研究の実施状況

共同研究相手国	相手方研究機関
---------	---------