

## 1 はじめに

投資信託（ファンド）とは、投資家から集めた資金を運用し、その成果を分配する金融商品のことである。投資家は運用方針の決定やリスク管理のためにファンドの性格を把握しておく必要があるが、類似したファンドを把握することは難しい。そこで武田ら [1] はファンドが投資している銘柄について着目しクラスタリングとネットワーク構造を用いてファンドの分析を行ったが他の多くのファンドとは違った銘柄に投資しているファンドに対してうまくクラスタリングを行うことができなかった。そこで本論文ではファンドの性格を分析するために従来研究ではできなかった独自性の高いファンドもクラスタリングできる手法について提案する。

## 2 投資信託の構造

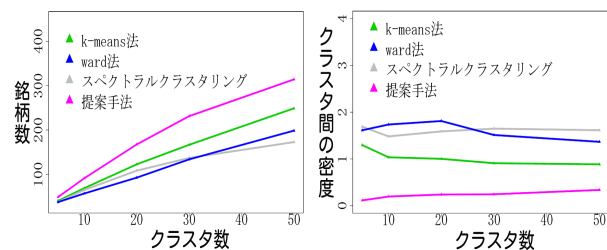
Yahoo!ファイナンスの提供する MORNINGSTAR の投資信託情報の中から、投資割合上位 10 銘柄の公開されている国内株式型投資信託を抽出した。このデータには各ファンドの投資割合上位 10 銘柄の銘柄名、投資割合などの情報が含まれている。抽出したファンド総数は 496 であり投資先銘柄総数は 688 個であった。各銘柄ごとに投資されているファンドの数を調べると、ほとんどの銘柄は 5 程度のファンドからしか投資されていない一方、100 や 200 を超えるファンドから投資されている銘柄の存在が確認できる。これは似通った投資をしているファンドが多数存在していることを意味する。本稿では他の多くのファンドとは違った銘柄に投資しているファンドを独自性の高いファンドと呼び、多くのファンドと似通った銘柄に投資しているファンドを独自性の低いファンドと呼ぶ。

## 3 投資信託クラスタリングの問題点

ファンドの投資割合が高い上位銘柄を属性とした属性ベクトルに対し、 $k$ -means 法、階層的クラスタリング手法の一つであるワード法、スペクトラルクラスタリングを行う。そして「ファンド間で共通する投資銘柄の数を重みとしたネットワーク」の密度について調べる。各手法のクラスタごとの密度を調べると独自性の高いファンド群が一まとめにされその密度は小さかった。武田はこの問題を解決するために HITS アルゴリズムで独自性の高いファンドを取り除いた後で  $k$ -means 法を適用する手法を提案した。しかし HITS アルゴリズムでは独自性の高いファンドを選別できない可能性があり、また独自性の高いファンドもまたファンドの性格を知る上で重要であるという考えから取り除かずにクラスタリングできる手法を提案する。

## 4 提案手法と評価

スペクトラルクラスタリングはグラフ分割の問題としてクラスタリングを行うがこの手法は対象を類似度行列に変換する必要がある、変換する手法に結果が大きく影響される。そこでファンド間の類似度は共通する投資銘柄の数に比例すると考えから、第 3 節で用いた共通銘柄数を重みとしたネットワークを直接類似度行列として与える手法を提案する。また、分散投資を目的に  $k$ -means 法、ワード法、スペクトラルクラスタリング、提案手法でファンドを  $k$  個にクラスタリングし ( $k=5,10,20,30,50$ )、各クラスタからランダムに一つ取り出した  $k$  個のファンドで構成されたポートフォリオについて、重複しない銘柄数とポートフォリオ内の密度について評価すると提案手法がすべてのクラスタ数  $k$  において優れていた。銘柄数は多いほど、密度は小さいほどよい。



左：銘柄数 右：密度

## 5 一般性についての検討

前節の実験では投資割合上位 10 銘柄を用いてクラスタリングを行ったが、次に上位 5 銘柄の情報のみを用いてクラスタリングを行いポートフォリオの構成を試みた。評価に上位 10 銘柄を用いて比較したところ、前節のものと近い結果が得られた。このことから、投資割合上位 11 銘柄以降の銘柄情報を用いた場合についても同様の手法が有効であると考えられる。

## 6 おわりに

本稿では独自性の高いファンドもクラスタリングできる手法を提案し、それが幅広い銘柄への投資に役立つことを確認した。またクラスタ内のファンドのパフォーマンスを調べることで同じ投資先でのファンドの優劣を調べられると考えられる。

## 参考文献

[1] 武田 祐太, 犬塚 信博, 「投資信託と銘柄のネットワーク構造に基づいた分散投資への試み」, 名古屋工業大学平成 24 年度卒業論文

## 発表論文

[1] 榎原 拓将, 松井藤五郎, 犬塚信博, 武藤敦子, 「スペクトラルクラスタリングを用いた投資信託分析」, 人工知能学会第 28 回全国大会, 2014 (発表予定) .