

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	ネットワーク系	犬塚研究室
	No. 21115087	武田 祐太

1 はじめに

投資において、単一の金融商品への投資はリスクが高く、分散投資がリスク分散に有効であるとされている。そこで、資産運用をプロに委託する投資信託という金融商品がある。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。

本稿では、ファンドと銘柄のネットワーク構造とクラスタリングの手法を用いてファンドの区分分けを行い、分散投資のためのポートフォリオ構成を試みる。

2 投資信託の構造

ファンド情報を得るため、Yahoo!ファイナンスより国内株式型投資信託 691 個を抽出した。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、銘柄業種、投資割合などの情報が含まれている。ファンドが投資先としている銘柄についてファンドと銘柄の間にリンクを張ったネットワークは、ファンド間や銘柄間にはリンクの無い二部グラフ構造となっている。

3 分散投資の実験と評価

抽出したデータについて次の 2 タイプ合計 4 つの手法を用いてファンドのクラスタリングを試みた。

3.1 属性ベクトルを用いた手法

第 1 のタイプは、ネットワーク構造を用いず、銘柄 769 についてファンドが投資割合上位 10 銘柄として投資しているか否かを 0/1 で表現した長さ 769 のベクトルと、それらの間のユークリッド距離に基づく方法である。このデータを対象に (1)*k*-means 法および (2) 階層型クラスタリングであるウォード法 [1] によるクラスタリングを行った。

3.2 ネットワーク構造を用いた手法

第 2 のタイプは、ネットワーク構造を利用したものである。5 つ以上の銘柄に共通投資しているファンド間にリンクをひいたネットワークについて (3) コミュニティ分割の手法である GN アルゴリズム [2] を適用した。また、(4)HITS アルゴリズム [3] と *k*-means 法を組み合わせた手法も試みた。HITS アルゴリズムでは、銘柄に投資しているファンドの数を各銘柄の次数とし、各ファンドの投資している銘柄の次数の合計を各ファンドに与える。この値を hub 値としたとき、hub 値が低いファンドは他のファンドが投資しているような人気の高い銘柄への投資を行っておらず、他のファンドと類似性の低いポートフォリオ構成となっていると言える。そこで、hub 値が閾値以下のファンドを「独自性の高いファンド」として選別した。

4 評価

4.1 銘柄・業種のカバー率および分散

分散投資を目的に、各手法に基づき *k* 個のファンド (*k*=5,10,20,30,50) によりポートフォリオを作成する。

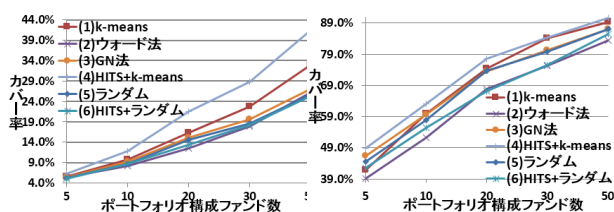


図 1: 銘柄カバー率

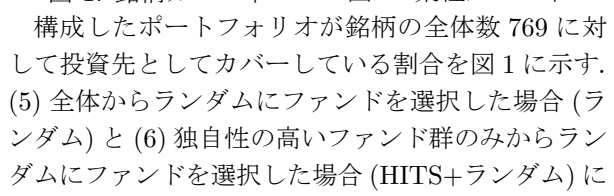


図 2: 業種カバー率

構成したポートフォリオが銘柄の全体数 769 に対して投資先としてカバーしている割合を図 1 に示す。(5) 全体からランダムにファンドを選択した場合 (ランダム) と (6) 独自性の高いファンド群のみからランダムにファンドを選択した場合 (HITS+ランダム) についても計算した。また、業種の全体数 34 に対してのカバー割合について同様に計算した結果を図 2 に示す。さらに、投資の偏り具合を検証するため、ポートフォリオの構成ファンド全体が各銘柄に対して投資している数を重みとし、投資先となり得る最大数である $k \times 10$ (業種の場合は全業種数である 34) 次元で構成されたベクトルについて標準分散を計算した。この値が低ければ低いほど、均等な投資が出来ていることを意味する。その結果、分散についても銘柄・業種ともに HITS+k-means により最も良い結果が得られた。

4.2 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、前節のものに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。

5 おわりに

本稿の評価では、HITS アルゴリズムで独自性の高いファンドを選別し、それ以外のファンドを *k*-means 法でクラスタリングする手法によって他の手法よりも優れた結果が得られた。

本稿の実験では、HITS アルゴリズムを用いた際に hub 値のヒストグラムの目視によって独自性が高いとする閾値を決定したが、この閾値の決定方法についての検討が今後の課題として挙げられる。

参考文献

- [1] 神島 敏弘, データマイニング分野のクラスタリング手法 (1), 人工知能学会誌, pp.59-65, (2003)
- [2] M.E.J.Newman and M.Girvan, "Finding and evaluating community structure in networks". Phys. Rev. E (2004)
- [3] Jon Kleinberg, "Authoritative sources in a hyperlinked environment", ACM-SIAM Symposium on Discrete Algorithms, pp.668-677 (1998)

発表

- [1] 武田 祐太, 松井 藤五郎, 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投資信託の分類」, 情報処理学会全国大会, (2013)