

平成21年度 情報工学コース卒業研究報告要旨

石井（健）研究室	氏 名	田 島 裕 之
卒業研究題目	機械学習を用いたポートフォリオの設計	
<p>株式等に代表されるリスク資産には、常に価格変動というリスクが付きまとう。リスクを低減するための方法の1つに、分散投資がある。分散投資とは、異なる性質を持つ複数の資産に対して投資することである。分散投資した資産の組み合わせを、ポートフォリオと呼ぶ。多くの異なる資産を組み入れるほど、ポートフォリオ全体のリスクは分散される。資金が多ければ、同時に保有できる資産の数が多いため、リスク分散は容易である。しかし、機関投資家とは異なり、個人投資家は多くの資産に分散投資を行うのに十分な資金を有していない場合が多い。また、多数の資産を個人で管理するのは難しい。したがって、少数の資産でもリスクを効率よく分散させることができるような、安定した運用成績を上げるポートフォリオを構築する手法が求められる。</p> <p>本研究では、安定した運用成績を上げるポートフォリオを構築することを目的とする。ここで言う安定した成績とは、市場全体の平均価格に連動することを指す。株価指数は、その市場における平均的な価格変動を表している場合が多い。よって、株価指数に連動するポートフォリオは、その市場において最大限にリスクを分散させた結果であると言える。そこで本研究では、ポートフォリオの安定性の基準に東証株価指数 (TOPIX) を用いることとする。ポートフォリオと TOPIX の乖離度を、両者の変動率の平均二乗誤差によって求める。乖離度の低いポートフォリオを、安定した運用成績を上げるポートフォリオと見なす。本研究では、TOPIX との乖離度の低いポートフォリオを、機械学習によって選別する手法を提案する。</p> <p>提案手法は、多数のポートフォリオの集合を入力とし、その中から TOPIX との乖離度の低いポートフォリオを選別し、出力とする。まず、ポートフォリオを構成する銘柄について、株価に関する指標を、異なる集計日数を用いて求める。次に、求めた指標毎に、ポートフォリオから度数分布を作成する。そして、度数分布の各データ区間における頻度を成分とした特徴ベクトルを、ポートフォリオから生成する。その後、生成した特徴ベクトルを、Nearest Neighbor 法によって識別する。Nearest Neighbor 法を用いた識別によって、入力となる多数のポートフォリオの中から、低乖離度クラスと識別されるポートフォリオを選別する。</p> <p>提案手法の有効性を検証するため、2種類の検証を行った。学習期間とテスト期間を同一期間とした検証では、適合率は 0.296 という結果が得られた。この適合率は、ランダムにポートフォリオを選択した場合に低乖離度となる確率 0.201 を上回った、しかし、テスト期間を学習期間の次の期間とした検証では、適合率は 0.232 に留まり、ランダムに選択した場合の適合率 0.217 との差はほとんど見られなかった。また、どちらの検証の適合率も十分高い数値ではなかったため、提案手法には、用いた特徴の見直しや、識別手法の再検討等、改善の余地が多いと思われる。</p>		