

ベイジアンネットワークを用いた株価指数の動向予測

Prediction of Stock Index by Using Bayesian Network

上山 薫, 上島 康孝, 左 毅, 北 栄輔
名古屋大学大学院情報科学研究科

本研究ではベイジアンネットワーク(BN)を用いた株価指数の変動予測について述べる。最初に2007年1月から10月の予測を行ったところ、テクニカル分析よりも高い的中率を示した。続いて、2007年と日本のバブル崩壊期的的中率の変化を詳しく調査したところ、的中率の低下をもとに、問題が顕在化した時期でなく、問題の原因が生じた時期をある程度推測できることが分かった。

This paper describes the application of the prediction of stock index by using Bayesian network. In the prediction of FTSE100 in 2007 January - October, the prediction accuracy of Bayesian network was better than that of technical analysis. Next, one observed the history of the prediction accuracy of index in 2007 and during the years of the asset-inflated economy in Japan. The results indicated that the reduction of prediction accuracy is effective for finding the reasons of the problems.

1 緒論

ベイジアンネットワークは確率的な因果関係をモデル化するグラフィカルモデリングの一つである[1]。あらかじめ設定しておいた個々のノードについての因果関係の有無を有向グラフで表し、また、因果関係の強さを条件付確率で表す。この因果関係の強さを求める方法にはいくつかの方法があるが、本研究ではK-2アルゴリズムを使用する。K-2アルゴリズムはよくばり法を元に作られたアルゴリズムであり、全木探索よりも少ない計算量でネットワークを構築する事が可能である。

本研究では、ベイジアンネットワークを用いて、前日までの3大証券取引所の株価指数を事前データとして学習させ、当日の指標の動き(上がり、下がり)を予測するシミュレーションを行う。予測する株価指数を持つ市場だけでなく、他市場の株価指数も学習する理由は、各市場の挙動について、「グローバルなマーケットは相互に関連し合っているという考え方は今から10年前には疑いの目で見られていたが、現在では最も重要なテクニカル分析の原則のひとつになっている。」[2]という指摘があるからである。予測された株価変動によって仮想的に売買を行う。これによって得られる収支を、いくつかの代表的な取引手法によって得られた収支と比較する。最後に、2007年と日本のバブル崩壊期的的中率の変化を調査し、的中率の低下とサブプライムローンやバブルの現象との関係を議論する。

2 研究背景

2-1 3大証券取引所の株価指数

現在世界で株式における3大市場と呼ばれている市場は、アメリカのニューヨーク証券取引所、イギリスのロンドン証券取引所、そして日本の東京証券取引所である。この3市場は、取引高が世界の株式市場の中で上位なだけでなく、お互いがある程度離れた場所にあり、その地域で代表的な都市の市場であるため、ニューヨーク証券取引所はアメリカ圏、ロンドン証券取引所はヨーロッパ圏、東京証券取引所はアジア圏の経済の状況を表す市場として、よく話題に上る。

ニューヨーク証券取引所は世界一上場審査が厳しいとされる。株価指標は「ダウ工業株30種平均」(Dow Jones Industrial Average)と呼ばれ、アメリカのさまざまな業種の代表的な銘柄を選出して、平均株価をリアルタイムで公表する株価平均型株価指数である。

ロンドン証券取引所は株式取引高が世界で2位の取引所である。代表的な株価指数はFinancial Times Stock Exchange 100(FTSE 100)で、これは1983年末の株価を基準値1000とした、時価総額加重平均型株価指数である。

東京証券取引所は世界第3位の証券取引所であり、市場第一部、市場第二部並びに新興企業向けのマザーズ、公社債市場がある。代表的な株価指数としては、東証第一部上場株の時価総額の合計を終値ベースで評価し、基準日である1968年1月4

日の時価総額を100として指数化する東証株価指数、東京証券取引所第一部に上場する約1700銘柄の株式のうち225銘柄を対象として算出する日経平均などがある。

2-2 株価指数の予測

株価指数の予測手法は大きく、テクニカル分析とファンダメンタル分析に分けられる。

ファンダメンタル分析では、経済の基礎的要因、つまり、経済力、政治力、その他要因を分析し、需要関係などを基にして株価指数の変動を予測する。ファンダメンタル分析は市場の現在の様子を幅広い原因から分析できるので、これからの市場の成長性を予測するのに優れている。しかし個々の要因同士が反対の結論を導くことも多くあり、さらにその判断はアナリストが行うため、人によって結果にムラが出る。さらにデータの収集や分析に時間がかかるため、短期的な投資には向かないという欠点がある。

これに対して、テクニカル分析では容易に入手できる過去と現在の価格、取引高、出来高の3つのみにデータを絞って将来の予測を行う。テクニカル分析は、価格の動きを研究していけば最終的に市場動向が予測できるという考え方である[2]。この方法も人によって結果にムラが出来てしまう。また、過去の経験のみで動向を求めると、大企業の倒産や合併など、価格には反映されるが、過去に数字として出ていないものには対応できないという欠点をもつ。

市場参加者はファンダメンタル分析とテクニカル分析のいずれかだけを用いて予測しているわけではない。お互いの欠点を補うために、どちらかを重視しつつも、もう一方の分析も考慮に入れて予測していくのが一般的である。しかし、2つの方法を組み合わせて利用しても、どちらの方法にしてもどの情報を使うか、それをどのように解釈するのかが個人によって違うために、結果がまちまちになってしまう。本研究では、ベイジアンネットワークを利用することで、市場参加者のムラが入らない予測を行うことができる。

3 株価予測

3-1 利用するデータ

東京、ロンドン、ニューヨークの各市場は時差のために開場している時間帯が異なる。そこで、本研究では、予測したい日の前市場から3日分の3大市場の株価指数を用いてベイジアンネットワークの学習を行う。つまり、市場は、1日の中で東京、ロンドン、ニューヨークの順に開いており、例えば、ロンドン市場のその日の株価指数に対してのネットワークを組む場合、使われるパラメータは3日前までの東京市場からその日の東京

市場、その間に開場されたロンドン、ニューヨーク市場のデータとなる。

これらの確率変数の状態としては、確率変数が示す市場と、確率変数が示す市場の前日の同市場との価格を比較して、上昇するか、下降するか2値とする。すなわち、前日よりその市場の株価指数が上昇する場合は1、下降する場合は-1とする。

3-2 ベイジアンネットワーク

ベイジアンネットワークは、確率変数、確率変数間の条件付依存関係、その条件付確率の3つによって定義されるネットワーク状の確率的モデルである。確率変数をノード、確率変数間の条件付依存関係をノード間に張られたリンクとして表す。リンクは有向であり、リンクの先にあるノードを子ノード、リンクの元にあるノードを親ノードとして表現する。条件付確率は、親ノードがある値を取った時に、子ノードがある値をとる条件付確率のことである。

3-3 グラフ構造探索のアルゴリズム

ベイジアンネットワークの構造学習アルゴリズムとして、本研究では現実的な時間でグラフの構造を探索、学習することのできるK-2アルゴリズム[3]を用いる。K-2アルゴリズムは確率変数間の半順序関係が決まっていなければ正確な結果が出ないという欠点がある。本研究で用いるパラメータは、各市場の株価指数が上がったか否かを表すものだけである。各市場は開いている時間に差があるため、各指標間の依存関係は一方に定義できる。過去の情報は未来に影響を与えるが、未来の情報は過去の情報に影響を与えない。従って、K-2アルゴリズムを用いることは妥当と考えられる。

4 解析結果

4-1 実験方法

1998年、1999年、2007年の3種類を予測期間とする。学習期間は予測期間の前2年間分とする。予測期間中において株価指数が上昇、もしくは下降する確率が55%以上であった場合に売買を行う。市場の開場と同時に、株価指数が上がると予測されるならば買い、下がると予測されるならば売りの取引を行う。55%以上の確率で結果が得られない場合取引を行わないものとする。そして、市場が開場するとき、開始とは反対の取引を行うことによって、その日の利益を確定する。これを1年間分繰り返して、的中率と総利益額を求める。

シミュレーション結果の比較対象として、テクニカル分析のなかからサイコロイド分析と一目均衡表を用いたトレンド分析

をとる。

サイコロイド分析は、株価指数が上がり続けた後には大きな下落があるという経験から考えられたものであって、株価指数の傾向は同じ状態で何日も続かないという考え方に基づいている。本研究では、過去10日間の株価指数の上がり下がりを用いる。過去10日中7日以上割合で株価指数が上がっているならば市場の開始時に売り、逆の状態なら買いの取引を行うこととする。利益を確定するための反対取引は、ベイジアンネットワークを用いた検証に準じて市場が閉まる時に行うものとする。

一目均衡表を用いたトレンド分析は、今の株価指数の価格傾向が以前と比べてずれているならば、それを修正する方向に価格が動くという考え方に基づく分析方法である。これは、株価指数は時間と共に動くものであり、過去の変動と現在の変動を比べることで相場が分かるというテクニカル分析の考え方に基づいている。基準線と転換線という2つの指標を使いトレンドを分析して売り買いを行う。このときの、基準線は(過去26日間の高値+安値)/2で求められ、転換線は(過去9日間の高値+安値)/2で求められるものである。転換線が基準線より高い場合は、以前よりも現在の方が買いの気が強く、価格が上昇すると予測し市場開始時に買いを、逆ならば売りの取引を行う。反対取引は、他と同じ方法で行う。

4-2 2007年の予測

2007年の株価指数を、ベイジアンネットワーク(BN)を利用したモデル(BN)、サイコロイド分析(TA1)、一目均衡表を用いたトレンド分析(TA2)の3種類の分析方法を用いて予測する。ロンドン市場を予測対象市場とする。ベイジアンネットワークの学習期間は2005年の1月1日から2006年の12月31日まで、予測時期は2007年の1月1日から10月19日までとする。

次に各分析方法で得られた結果と実際のロンドン市場の指標であるFTSE100の上がり下がりと比較する。サイコロイド分析による結果とFTSE100の実際の変動の比較を図1に、トレンド分析による結果とFTSE100の実際の変動の比較を図2に、ベイジアンネットワークによる結果とFTSE100の実際の変動の比較を図3に示す。縦軸が1をとる場合、FTSEの変動と予測結果が同じ、すなわち予測が当たったことを示し、-1をとる場合、予測が外れたことを示す。0の場合は、予測において取引を行えなかった日である。また、各分析方法での取引回数、的中回数、的中率、総利益を表1にまとめる。

図1に着目すると、サイコロイド分析(TA1)では取引日がかなり少ないことがわかる。表1から分かるように10ヶ月以上の期間にも関わらず取引回数は59回である。予測が当たった回数は29回であり、的中率は50%弱である。

図2に着目すると、トレンド分析(TA2)では常に買いか売りを出しているため、他の方法よりも取引回数が多いのが特徴である。しかし、予測が外れている日も多く、成果は上げられていないとも取れる。表1にあるように実際取引した回数204回に対して、予測が当たった回数は102回という結果になっており、的中率は50%である。

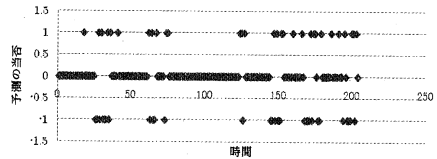


図1 サイコロイド分析(TA1)の結果

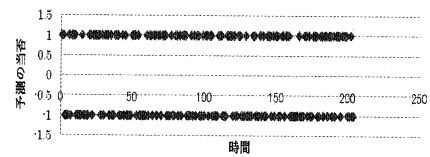


図2 テレンド分析(TA2)の結果

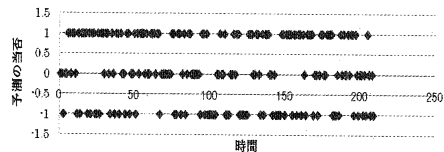


図3 ベイジアンネットワーク(BN)の結果

表1 各分析方法の比較(解析例1)

	取引回数	的中回数	的中率	総利益
TA1	59	29	0.49	696
TA2	204	103	0.50	-104.1
BN	146	86	0.59	1640.4

図3と表1を見れば、ベイジアンネットワークによる結果(BN)での取引回数はトレンド分析(TA2)よりも少ないが、サイコロイド分析(TA1)よりも多いことがわかる。実取引回数142回のうち予想が当たった回数は84回であり、的中率は約60%である。

以上をまとめると、取引回数では、ベイジアンネットワークはトレンド分析よりも少ないことがわかる。しかし、予測が当たった的中率は、ベイジアンネットワークはテクニカル分析よりもかなり高い。テクニカル分析的な中率は、どちらの手法でも約50%なのに対して、ベイジアンネットワークの的中率は約60%である。これは、ベイジアンネットワークが株価指数の予測に対して、旧来の手法よりも優れている可能性を示唆してい

る。

以上よりベイジアンネットワークがある程度の有用性を示すことが分かった。しかし、図3についてさらに詳しく見ていくと、横軸の値が70を示すあたりからの中率が急に低下していることが分かる。この時期は、図1のサイコロイド分析で取引が行われなくなった時期と重なっている。

ところで、同時期のFTSE100の変動を見ると横軸が70の辺りでは目立った株価変動は見られないが、150付近では価格が暴落している。実は150付近で生じた暴落は米国のサブプライムローンの影響による。この事に注目すると、興味深いことが見える。実は、日経新聞紙上でサブプライムローンの回収想定が甘いと指摘されたのが、2007年3月19日であり、その日付は、先ほどまでの図で言えば横軸で60付近であり、ベイジアンネットワークが予想的中を少なくした時期とおよそ一致している。

このことにより、一つの仮説が導き出せる。すなわち、ベイジアンネットワークを用いた予測的中率は、サブプライムローン問題によって起こる市場の変化によって悪化した。つまり、ベイジアンネットワークによる予測的中率の変化が、市場の状態変化、少なくとも、株価指数の変動がそれまでと比べて異常になったことを示しているのではないかと考えられる。

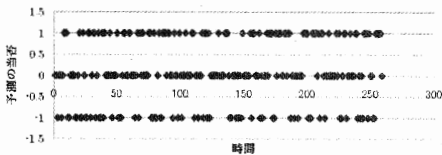


図4 1990/7/1-1991/6/30の予測結果

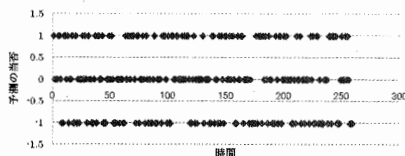


図5 1989/7/1-1990/6/30の予測結果

4-3 バブル期の予測

日本のバブル崩壊期を検討するために、2種類の予測期間を用いて行う。第1は、バブル崩壊が起きたとされる1991年2月を含む、1990年の7月1日から1991年の6月31日までを予測期間としたものである。第2は、バブル崩壊の原因とされる日銀の低金利政策からの転換があった1989年6月を学習期間と予測期間の境目として検証を行うものである。

(1) 1990年の7月1日から1991年の6月31日まで

ベイジアンネットワークを用いた予測結果を図4に示す。これよりバブル崩壊期(1991年2月)の前後で予測的中精度に大きな変化がないことが分かる。予測期の取引回数150回に対して的中回数88回であって、約59%的中率をもっており、2007年と比較して特に悪化したわけではない。この検証では、価格が上昇傾向にある期間を学習期間とし、価格が下降傾向にある期間を予測期間とした。つまり、株価指数の価格傾向が、それまでと比べて変わったとしても、ベイジアンネットワークを用いた予測の精度はそれほど落ちないといえる。

(2) 1987年7月1日から1989年6月31日まで

ベイジアンネットワークによる予測結果を図5に示す。図4と図5を比較すると、図5は的中率が低下していることが分かる。図5では取引回数160回に対して的中回数81回、的中率は約50%である。的中率が低下した理由として、予測期間で学習期間にはなかった(最終的にバブル崩壊を予想させたと思われる)情報が市場に広がり、市場状態が変わったためであると考えられる。

これらの結果より、ベイジアンネットワークによる予測的中率の変化によって、市場の状態変化を予測できると考えられる。

5 結論

本研究では、ベイジアンネットワークを用いた株価指数の変動予測について述べた。3大市場の株価指数を用いてベイジアンネットワークを学習し、そのネットワークを予測に適用した。

最初に2007年度の株価変化を予測したところ、約60%的中率を示した。この結果は、サイコロイド分析、一目均衡表を用いたトレンド分析よりも高いものであった。続いて、バブル期の株価指数動向の予測にベイジアンネットワークを適用した。その結果、バブル崩壊時期の前後では予測精度に大きな変化はないが、バブル崩壊の原因となる時期の前後では的中率が低下することが分かった。

以上のことより、BNによる予測精度の低下が市場の挙動変化予測にある程度役立つことが分かった。

参考文献

- [1] 志賀元紀, ベイジアンネットワーク, 2005
- [2] ジョン・J・マーフィー (長尾慎太郎, 関本博英 訳), 市場間分析入門 パンローリング, 2005
- [3] G. Cooper and E. Herskovits: A Bayesian method for the induction of probabilistic networks from Data, *Machine Learning*, vol.9 (1992), 309.