

制約付きニューラルネットワークの混合モデルによる株価予測

1X07C033-4 小笠原匠
指導教員 後藤正幸

1 研究背景・目的

近年、証券市場の発達により、株式投資を行う多くの投資家にとって、株価予測のニーズは高くなっているといえる。

ここで、株式投資における代表的な手法として、不測の損失と利益の機会を予測するテクニカル分析手法が挙げられ、これにニューラルネットワーク (NN) を用いた手法が提案されている [1]。この手法は、7つの NN を用いて株価予測を行うことで、一定の結果を得ている。しかし、すべての NN の出力を同等に評価していることから、さらに予測精度を改善する余地があると考えられる。NN の予測精度の向上には、パラメータ数を適切に設定することが必要である。これは、一般に NN のパラメータ数が多いと、過学習がおきるため、未学習データには適合しないといわれるためである。一方、橋川ら [2] の研究では、複数の NN を用いて予測を行う場合、すべての NN の出力を同等に評価して予測を行うのではなく、事後予測分布をもとに予測を行うことがベイズ最適であることが示されている。

そこで、本研究では、合理的に NN の構造に制約を加えることで過学習を抑える方法を提案するとともに事後予測分布に基づいた混合モデルを用いて株価予測を行うことで、株価の予測精度を向上させることを目的とする。また、実際の TOPIX 数値データを用いて、提案手法の有効性を示す。

2 準備

2.1 テクニカル分析手法

テクニカル分析手法とは、株価に関する短期移動平均線 (6週平均) と長期移動平均線 (13週平均) の交わった時点で株式の売買を行う手法をいう。なお、短期移動平均線が長期移動平均線を下から上に突き抜けた点をゴールデンクロス (GC)、短期移動平均線が長期移動平均線を上から下に切り下がった点をデッドクロス (DC) という。一般にテクニカル分析手法は有用といわれるが、株価が急激に上昇または下落している場合には GC、DC が発生した時点ですでに意思決定のタイミングを逸している場合が多く、GC、DC の発生を事前に予測が必要である [1]。

2.2 ニューラルネットワーク

NN は、脳機能のメカニズムを計算機上でシミュレートして表現することを目指した数学モデルである。入力層の第 i ユニットの入力、出力を x_i^1, y_i^1 とし、中間層の第 j ユニットの入力、出力を x_j^2, y_j^2 とし、出力層の第 k ユニットの入力、出力を x_k^3, y_k^3 とする。入力層と中間層、中間層と出力層の結合荷重を $\omega_{ij}^{12}, \omega_{jk}^{23}$ とする。また、中間層と出力層の閾値を θ_j^2, θ_k^3 とする。入力層、中間層、出力層の入出力関数を f_{in}, f_{hid}, f_{out} とする。 f_{in} を線形関数、 f_{hid}, f_{out} をシグモイド関数とする。このとき、NN の出力層第 k ユニットの出力は、

$$y_k^3 = f_{out} \left[\sum_{j=1}^H \omega_{jk}^{23} f_{hid} \left(\sum_{i=1}^I \omega_{ij}^{12} f_{in}(x_i^1) + \theta_j^2 \right) + \theta_k^3 \right] \quad (1)$$

で与えられる。

なお、結合荷重、閾値の学習には、出力層からの出力と教師信号 (正解) との誤差を中間層と出力層の結合荷重及び出力層の閾値、入力層と中間層の結合荷重及び中間層の閾値の順に修正していく Back Propagation 学習を用いる。入力ベ

クトル $x_i = (x_{i1}, \dots, x_{iI})$ に対する出力を y_i とし、学習データが与えられたもとの、未学習データの入力ベクトル x_{n+1} が与えられたときに出力 y_{n+1} を予測する問題と捉えることができる。

3 従来手法

馬場ら [1] の研究では、TOPIX 数値の変化率、米ドル変化率、ダウ変化率、ナスダック変化率等 21 個の入力変数を用いて GC、DC の予測を行っている。入力変数について、特定の入力変数を 1% 変化させたときの出力の変化の大きい順に順位づけを行い、上位 21 個、18 個、16 個、14 個、12 個、10 個、8 個の入力変数を用いて、7つの NN を構築する。なお、中間層の数は、入力変数の数と同じ数に設定する。

教師信号に関しては、GC を 0 とし、GC が発生する 4 週間前の教師信号を 1.0、3、5 週間前を 0.8、2、6 週間前を 0.6、1、7 週間前を 0.4、そして、GC が発生する 8 週間前から DC までの教師信号を 0.15 とした。一方、DC を 0 とし、DC が発生する 4 週間前の教師信号を -1.0、3、5 週間前を -0.8、2、6 週間前を -0.6、1、7 週間前を -0.4、そして、GC が発生する 8 週間前から DC までの教師信号を -0.15 とした。図 1 に株価の動きと教師信号の与え方の例を示す。

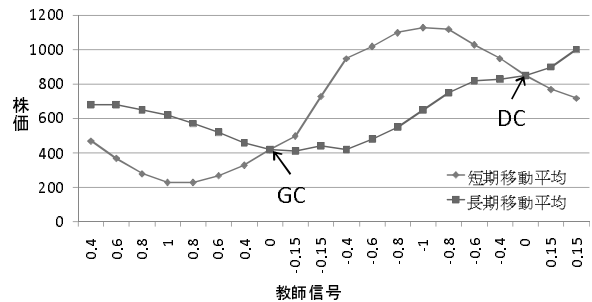


図 1. 教師信号の設定 (例)

予測に関しては、7つの NN の出力を用いて行う。具体的には、以下の A、B のどちらかの条件を満たした場合に GC (DC) が生じると予測する。

条件 A: 全出力が正 (負) かつ出力の絶対値の平均が 0.5 以上
条件 B: 過半数の出力が正 (負) かつ正 (負) の出力のうち、0.7 以上 (-0.7 以下) の出力が過半数を占め、さらに全体の出力の平均値が (0.5 以上 -0.5 以下)

4 提案手法

各 NN の予測精度を改善するために制約を加え、パラメータ数を減少させた NN を提案し、すべての NN の出力を同等に評価している点を改善するために混合モデル [2] を適用する。

4.1 制約を加えたニューラルネットワーク

7つの NN の予測精度が低い理由としては、NN の表現能力の豊かさのために過学習をおこしている可能性があり、この問題を解消するためには、対象問題に適した構造をもつ NN を構築すべきである。従来手法では、TOPIX 数値の変化率、米ドル変化率、ユーロ変化率、ダウ変化率、ナスダック変化率等 21 個の入力変数を用いて、GC、DC の予測を行っているが、これらの入力変数は、直近の TOPIX 数値

の推移、為替相場、海外市場等の要因別のカテゴリに分類することができ、これらのカテゴリの影響が GC, DC の発生に起因していると考えられる。そのため、NN の構造として、中間層への入力が入力層からの出力を反映したものである必要はない。

そこで、従来手法で用いた入力変数 21 個を要因別にカテゴリ 6 つに分けて、入力層と中間層の結合荷重についてカテゴリ外のを 0 と制約を課すことでネットワークに構造を加えた NN を提案する。図 2 にその構造を示す。

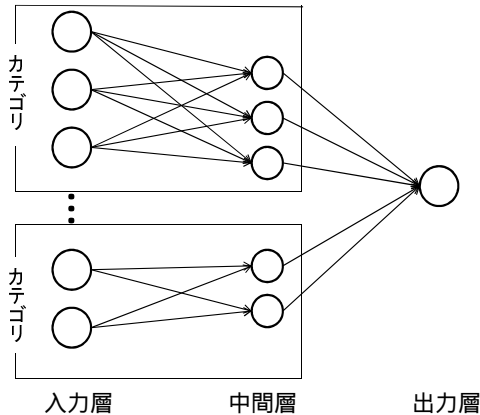


図 2. 制約を加えたニューラルネットワーク

4.2 混合モデル

本来、すべての NN の出力を用いて予測を行う場合、予測精度が相対的に高い NN の出力を高く評価し、低い NN の出力を低く評価すべきである。

橋川ら [2] は複数の NN を用いて予測を行う場合、事後予測分布をもとに予測を行うことがベイズ最適であることを示している。このことから、本研究では、ベイズ最適な予測を行うため、事後予測分布の出力結果が 0.5 以上の場合に GC が発生、-0.5 以下の場合に DC が発生すると予測するモデルを提案する。なお、事後予測分布を厳密に計算すると、パラメータ空間上の積分操作が必要になり計算が困難となるため、橋川らの研究では、ラプラスの漸近展開を用いてこの積分を排除し漸近近似的に事後予測分布を得る方法を提案している。本研究はこの手法を株価予測に援用したものである。ここで、各 NN の重み H_k は、以下のように示される。

$$H_k = \left(\frac{1}{2\pi\sigma^2} \right)^{\frac{n}{2}} \exp \left\{ -\frac{1}{2\sigma^2} \sum_{i=1}^n (y_i - f(x_i, \hat{\theta}_k, M_k))^2 \right\} \left(\frac{n}{2\pi} \right)^{-\frac{s_k}{2}} / Z \quad (2)$$

M_k は NN モデル ($1 \leq k \leq 7$)、 $\hat{\theta}_k$ は学習後得られたモデル M_k の結合荷重および閾値、 σ は教師信号 y と出力 $f(x_i, \hat{\theta}_k, M_k)$ の誤差分散、 n は学習データ数、 s_k は、モデル M_k のパラメータ数を指す。

5 実験および考察

5.1 実験設定及び概要

実験は、制約を加えた NN の予測精度を向上させること、ならびに GC, DC の予測精度を向上させることを示すために 2 段階で行った。

まず、テストデータにおける NN の出力と教師信号の二乗誤差について、従来手法と提案手法について各 NN の比較を行う。なお、提案手法における入力変数の数、中間層の数は従来手法と等しく、21 個、18 個、16 個、14 個、12 個、10 個、8 個と設定する。

次に、従来手法と 4.2 節の提案手法で、GC, DC の予測を行う。評価尺度として、精度、予測成功率を以下に定義する。

精度：GC(DC) が発生する 1~7 週間前に GC(DC) の予測を行った比率

予測成功率：GC(DC) の予測があった 1~7 週間後に実際に GC(DC) が発生する比率

なお、学習期間 3 年、予測期間 1 年とし、実験データは、2004~2006 年の TOPIX 数値データを用いて学習を行い、2007~2009 年をテストデータとして予測を行う。

5.2 実験結果及び考察

表 2 に、テストデータにおける NN の出力と教師信号の二乗誤差損失の実験結果を示す。

表 2. NN の二乗誤差損失の比較

No	入力変数	従来手法	制約付き NN
1	21 個	58.707	37.063
2	18 個	67.685	47.097
3	16 個	50.959	30.860
4	14 個	60.839	27.673
5	12 個	53.146	21.765
6	10 個	44.423	22.741
7	8 個	25.250	22.751

表 2 より、入力変数にかかわらず、従来手法よりも提案した制約付き NN の方が二乗誤差損失が小さくなっている。入力層と中間層の結合荷重に制約を加えることによって、NN の過学習が抑えられ、予測精度の向上につながったと考えられる。

また、表 3 に GC, DC の予測を行った実験結果を示す。

表 3. 従来手法と提案手法の比較

	従来手法		提案手法	
	精度	予測成功率	精度	予測成功率
GC	0.80	0.50	1.00	0.60
DC	0.60	0.78	1.00	0.83

精度、予測成功率共に従来手法よりも提案手法の方が高くなっている。これは、予測精度が向上した個々の NN モデルを事後予測分布に基づき混合することにより、株価予測の予測精度の向上につながったと考えられる。

また、提案手法の GC が DC よりも予測成功率が低かった点に関して、GC, DC の予測精度に差があるというのは、両者の対称性を考慮すると矛盾するとも考えられる。しかし、これは 2008 年の 8 月~11 月に多く GC の予測をしているが予測が外れてしまったことが原因であるといえる。リーマンショックが 2008 年 9 月 15 日に起きており、この影響を少なからず受けて、予測が外れてしまった可能性がある。現に、この時期の誤った予測を除外すれば、GC の予測成功率は、0.8 となり、DC とほぼ等しいため、GC と DC の予測精度にあまり変わりはなく、直感的にも自然な結果になる。

6 まとめと今後の課題

本研究では、NN の入力層と中間層の結合荷重に制約を加え、複数の NN を混合することによって、株価予測の予測精度を高められることを示した。しかし、予測の対象である GC, DC 自体は頻繁に起こるものではないため、サンプル数が比較的少なくなっている。そのため、データ数が増加した際にも適用できるかどうか検証する必要がある。

参考文献

- [1] 馬場剛夫, 西田将人, 甲斐良隆, “ニューラルネットワークを活用した従来型テクニカル分析手法改善の試み,” 電気学会論文誌 C, 126 巻, 11 号, pp.1324-1331, 2006
- [2] 橋川弘紀, 後藤正幸, 依信彦, “階層型ニューラルネットワークの混合モデルによるベイズ最適な予測について,” 電子情報通信学会論文誌 Vol.J80-D II, No.7, pp1919-1928, 2002