2008－MPS－69 (9)
2008／5／16

株式クオンツモデルでの過適合

水田孝信[1*] 小林悟[1] 加藤徳史[1]

[1]スパークス・アセット・マネジメント株式会社 運用調査部

株価予測モデルにおける過剰適合について調べた。定量的分析を行うために、中間層が1層のニューラルネットワークを用いて中間層の数と汎化誤差の関係を調べた。その結果、中間層が多すぎると汎化誤差が上昇し、過剰適合が発生することが分かった。この現象は、株価予測モデルが“複雑すぎる”ために予測能力が低下することが起こりうることを示している。また、学習させるファクターが異なる2つのモデルの予測リターンを比べた結果、適切な学習を行ったときに最も予測が似てしまうことが分かった。

Over Fitting on the Quants Model of Stock Prices

Takanobu Mizuta, Satoru Kobayashi and Tokufumi Kato

[1]Department of Investment and Research, SPARX Asset Management Co.,Ltd.

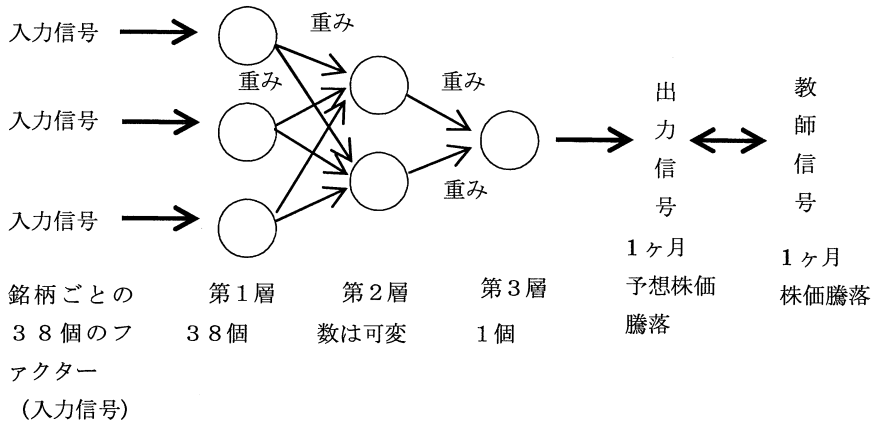
We discuss overfitting in stock price prediction models. In quantitative analyses, we investigate the relationships among the number of hidden layer units and generalization errors using artificial neural networks with one hidden layer. Our results show that over fitting occur as the number of hidden layer units increases. This phenomenon indicates that the prediction ability of a stock price model declines when the model is “too complex”.

1 はじめに

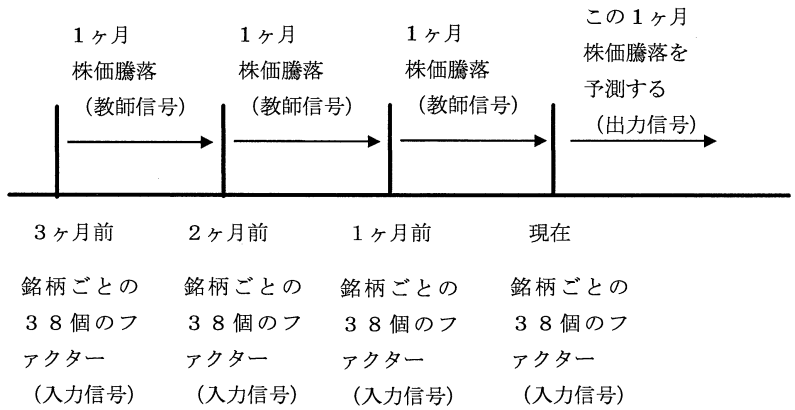
近年、資産運用業界では金融工学の発展に伴い、定量的な予測モデルを用いて資産運用を行うことが増えており、金融市場への影響も大きくなっている[1)]。このようなモデルを一般に、クオンモデルとよび、クオンツモデルを用いたファンドをクオンツファンドとよんでいる。通常、クオンツファンドを組成するときは、過去のデータを用いてバックテストによる検証を行う。バックテスト上での収益が高く、収益が悪化する時期が存在しないことが重要である。というのもバックテストの結果の良し悪しは、クオンツファンドを購入する顧客へ営業活動にとても重要であり、資産を安定的に増やしたいと願う顧客が多いからである。バックテストの結果が良いモデルを用いて実際に運用する場合が多いが、一方で、バックテストの結果はすばらしいのに、実際の運用となると結果がふるわないこともしばしば発生する。これはバックテストの結果を良くしようとするあまり、過去のデータに過剰にフィッティングを行った結果、未来の予測能力が落ちてしまう、過剰適合現象や過剰学習現象[2)~4)]であると推測できる。クオンツファンドによる過剰適合は、使用パラメータの組み合わせ方や演算方法を“複雑”にした結果、バックテストの結果はよくなるが、実際の運用では予測能力が落ちる現象である。過剰学習は、バックテストの結果をよくするためにモデルのパラメータを過剰に“精密”に設定し、予測能力が低下することである。このような現象は、例えば、有機化合物の有害性予測など他分野で多く報告[5)~6)]されているが、資産運用業界では、考慮されることはこれまであまりなかった。

株価予測モデルはさまざまな作成方法が存

* 連絡先：スパークス・アセット・
マネジメント株式会社
運用調査部
〒141-0032 東京都品川区大崎 1-11-2
ゲートシティ大崎イーストタワー16 階
E-mail: takanobu.mizuta@sparxgroup.com



(図１)　本稿で用いた３層パーセプトロン型ニューラルネットワークモデル。入力信号は各銘柄の３８個のファクターであり、第１層は３８個となる。教師信号はその後の１ヵ月の株価の騰落率で、各銘柄にひとつの値をとるため、第３層（出力信号）は１つとなる。第２層の数は可変であり、この数を変えることによりモデルの複雑さを変更する。

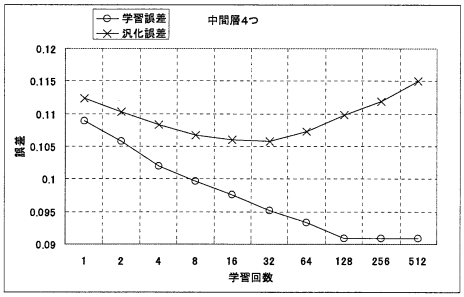


(図２)　学習方法の模式図。まず測定時点から１ヵ月前のファクター値を入力信号、その後の株価騰落率を教師信号とする。これを 1,000 銘柄に対して行う。同じことを、２ヶ月前、３ヶ月前も同様に行い学習は終了する。最後に現在のファクター値を入力し、出力信号を未来１ヵ月の株価騰落率の予想とする。

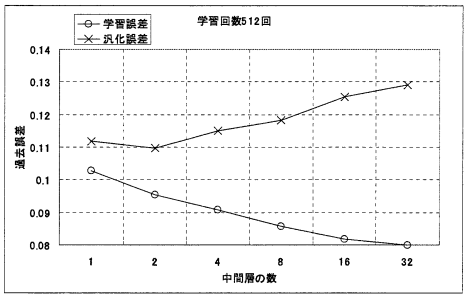
在し、一概にどちらが複雑か、どちらがより精密かを客観的に言うのは難しい。本研究では、このようなモデルの複雑さと精密さを定量的に表現するためにニューラルネットワーク [7]を用いて、株価予測モデルにおける複雑さや精密さと、予測能力の高さを比較した。

2　モデル

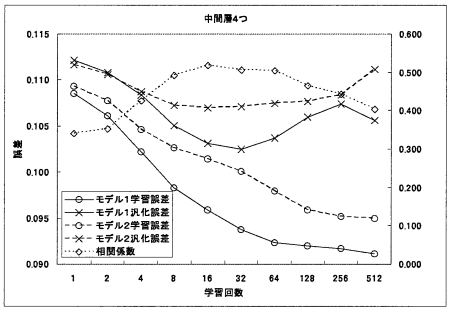
株価予測モデルはさまざまな作成方法が存在し、一概にどちらが複雑か、どちらがより精密かを客観的に言うのは難しい。そこで、それらを定量的にコントロールできるニューラルネットワークを用いて、モデルの複雑さや精密さと予測能力の関係を定量的に算出する。本稿では、３層パーセプトロン型 [8)~9)]のニューラルネットワークを用いる（図 1)。学習の回数を増やすことによりモデルは、より



(図３)　中間層を４つに固定し、学習回数を可変として学習誤差と汎化誤差を示した。学習回数が増えるにつれて学習誤差は減少しているが、汎化誤差は学習回数 32 回を境に上昇に転じている。



(図４)　学習回数を 512 回に固定し、中間層の数を可変として学習誤差と汎化誤差を示した。中間層の数が増えるにつれて学習誤差は減少しているが、汎化誤差は中間層が２つを超えると上昇に転じている。



(図５)　学習信号が異なる２つのモデルにおいて、中間層を４つに固定し学習回数を変化させたときの学習誤差、汎化誤差と２つモデルの出力信号の相関係数。汎化誤差が最小となる最適な学習回数付近で、２つのモデルの相関は高くなっている。その値は 0.5 を超えており、２つのモデルの予測リターンは、学習信号が違うにも関わらず、かなり似たものとなっている。

細かく重みを調整する。つまり、より精密なモデルになる。第２層の数を増やせば増やすほど、より多くの信号の組み合わせを行うことが可能になる。つまり、より複雑なモデルとなる。学習回数と第２層の数を変えることにより、モデルの精密さと複雑さを変更することが可能である。

この３層パーセプトロンを用いてクオンツモデルを作成する。学習信号は個別銘柄の３８個のファクター値である。学習手順を図２に示す。１ヶ月前のファクターを学習信号とし、その銘柄の１ヶ月前から現在までのリターンを教師信号とする。つまり、第１層は３８個で、これらのファクター値が入り、第３層は１つで予測リターンが出力されるモデルである。このとき各信号は 1～1,000 までの順位化し、これを 1,000 で割って約 0～1 までに規格化しておく。この学習を 1,000 銘柄すべてについて行う。つぎに２ヶ月前のファクターと２ヶ月前から１ヶ月までのリターンを教師信号とし全銘柄に対して学習を行う。これを３ヶ月前まで行い、このセットを１回の学習と数えることにする。あらかじめ決めた学習回数繰り返した後、現在のファクター値を入力すると、出力信号がこのモデルの予測リターンとなる。これを 1995 年 4 月から 2007 年 12 月の各月について行う。予測リターンと実際のリターンの差の２乗の平均を汎化誤差とよぶことにし、これが少ない方が優れたモデルであるといえる。また、各月の最終的な出力信号と教師信号の差の２乗の平均を学習誤差とよぶことにし、これが少ない方が過去のデータによく適合している、つまりバックテストが優れているモデルであるといえる。

3　分析結果

学習回数を 1,2,4,8,16,32,64,128,256,512、中間層の数を 1,2,4,8,16,32 とそれぞれ変化させて学習誤差と汎化誤差を計測した。学習回数、中間層の数の全ての組み合わせに対して最も低い汎化誤差となったのは、学習回数 32 回、中間層の数 4 つのときであり、最も低い学習誤差となったのは、学習回数 512 回、中間層 32 個の場合であった。

図３は中間層の数を 4 に固定し、学習回数を変化させ誤差を測定したものである。学習

回数が多いほうが学習誤差は減少しているが、汎化誤差は学習回数32回を境に上昇しており、過学習現象がはっきりと現れている。クオンツファンドに立ち返って考えると、パラメータ調整を精密にすればするほどバックテストは良くなっているが、予測能力は低下していることを示している。

次に学習回数を 512 回に固定して中間層の数を可変にした。図４は、中間層の数を変えながら、学習誤差、汎化誤差を示したものである。学習誤差は中間層の数を増やすほど低下している。つまり、モデルを複雑にすることによりバックテストの結果をよくすることが出来ている。しかし、汎化誤差を見ると中間層が２つのときに最も小さな値をとり、上昇に転じている。中間層２つでは学習誤差0.096 程度で飽和しているが、それぐらいシンプルなモデルが、最も予測能力が高くなっている。

図５は、学習信号が異なる２つのモデルにおいて、中間層を４つに固定し学習回数を変化させたときの誤差を示した。また、２つのモデルの出力信号（予測リターン）の相関も示している。どちらのモデルも過剰学習現象が起きており、また、汎化誤差が最小となる最適な学習回数付近で、２つのモデルの相関は高くなっている。その値は 0.5 を超えており、２つのモデルの予測リターンは、学習信号が大幅に違うにも関わらず、かなり似たものとなっている。クオンツファンドに当てはめて解釈すると、使用するファクターを工夫して差別化を行ったとしても、適切にチューニングすればするほど、差別化の効果は弱くなり他のファンドと保有銘柄が似てしまうことが起きてしまうことを示唆している。

4　まとめ

株価予測モデルにおいて、過剰適合現象や過剰学習現象が起こるか調べた。それらを定量的に調べるために、３層パーセプトロン型のニューラルネットワークモデルを用いた。中間層の数や学習回数と、汎化誤差の関係を調べた結果、これらの現象は確かに存在した。クオンツファンドにおけるモデルが精密すぎたり複雑すぎたりすることによって、予測能力を低下させることがありえることが示されたと考えられる。また、学習信号が異なる２つのモデルを比較した。使用するファクターが大きく異なるクオンツモデルでも、適切な精密さを持たせると、予測リターンが似たものになってしまうことを示した。

これまでクオンツファンドでは、このようなモデル選択理論的な考え方はあまりなかった。他分野の中には、これらの研究がかなり進んでいる分野もある。他分野での事例や手法が、ファンド運用でどれほど当てはまるか、今後も研究を深めていく必要がある。

参考文献

1) Khandani , A. E., and Lo, A. W.: What happened to the quants in August 2007?, Working paper series, Nov. 4, (2007)
2) Baum, E. B., and Haussler, D.: What size net gives valid generalization ?, Neural Compt. Vol. 1, pp. 151-160 (1989)
3) Weiss, S.M., and Kapouleas, I.: An empirical comparison of pattern recognition, neural nets, and machine learning classification methods, Proc. IJCAI 89, pp.781-787, (1989)
4) Tetko, I.V., Livingstone, D.J., and Luik, A.I.: Neural Network Studies. 1. Comparison of Overfitting and Overtraining, J. Chem. Inf. Comput. Sci., Vol. 35, pp.826-833 (1995)
5) Tanabe, K., Ohmori, N., Ono, S.;Suzuki, T., Matsumoto, T., Nagashima, U., and Uesaka, H.: Neural Network Prediction of Carcinogenicity of Diverse Organic Compound, J. Comput. Chem. Jpn., Vol.4, No.3, pp.89-100 (2005)
6) Peterson, K.L.: Artificial Neural Networks and Their Use in Chemistry, Rev. Comput. Chem., Vol. 16,pp.53-140 (2000)
7) McCulloch, W.S., and Pitts, W. H.: A logical calculus of the ideas immanent in nervous activity, Bulletin Math. Biophys., Vol. 5, pp.115-133 (1943)
8) Rosenblatt, F.: Principles of Neurodynamics (1962)
9) Rumelhart, D.E., Hinton, G.E., and Williams, R.J.: Learning representations by back-propagating errors, Nature, Vol. 323, pp.533-536 (1986)