

S A法の拡張と時系列予測を含んだ投資問題への応用

小林寛 木村充宏 松葉育雄

kobayasi@j90.tg.chiba-u.ac.jp

千葉大学大学院工学研究科

〒263 千葉県千葉市稲毛区弥生町 1-33

組合せ最適化問題の一解法であるS A法に多スタートと選択処理の要素を取り入れた拡張型S A法を提案する。また、最適化問題の投資問題における例として知られているポートフォリオ選択問題において必要となる予測収益率を、ニューラルネットワークによる時系列予測から得ることを試みる。本稿ではまずポートフォリオ問題に用いる収益率の時系列予測の方法と予測結果について述べ、次に拡張型S A法を紹介するとともにポートフォリオ問題に適用したときのシミュレーション結果を示す。

THE EXTENSION OF SIMULATED ANNEALING METHOD AND ITS APPLICATION TO INVESTMENT PROBLEM WITH PREDICTED VALUES OF TIME SERIES

Hiroshi Kobayashi ,Mituhiro Kimura ,and Ikuo Matsuba

Graduate School of Engineering ,Chiba University

1-33, Yayoi-tyou, Inage-ku, Chiba-shi, Chiba, 263 Japan

We propose the improved Simulated Annealing method using the multi-start and selection which are commonly used in GA for optimization problems. As an application of the present method, we try to solve the portfolio selection problem in which we estimate the return value by feed-back neural networks proposed previously.

1 はじめに

組合せ最適化問題の例として、経済における投資問題であるポートフォリオ選択問題[1]が知られている。その目的は投資の分散化によってリスクをなるべく小さくしようとすることである。最適なポートフォリオを組むための手法には、S A (Simulated Annealing) 法 [2][3] や G A (Genetic Algorithm)[4][5] などがあるが、どの手法においても数値データとして収益率データなどが必要となる。月間収益率で考えた場合、今月の最適ポートフォリオを組む段階では必要となる今月分の収益率はまだ未知である。そこで、先月分までの収益率データを用いて今月分の収益率をニューラルネットワークにより予測し、その予測収益率データを用いてポートフォリオを組むことにより、有効なポートフォリオが得られる可能性がある (図 1)。

以降ではポートフォリオ選択問題に沿って、ニューラルネットワークによる収益率予測と、最適化手法として S A 法を改良した手法を示し、ポートフォリオ問題に対するシミュレーション結果を示す。

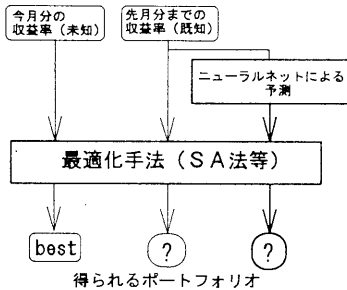


図 1 ポートフォリオ選択問題

2 ポートフォリオ選択問題

不確実な将来収益を持つ各種資産に対する投資家の最適投資行動に関するポートフォリオ選択理論は、1959 年にマーコビッツによって定式化されたのが始まりである [6]。まず、ニューラルネットワークによるポートフォリオ最適化のための定式化を行う。いま、 i 証券の月間収益率を R_i 、投資率を x_i とするとポートフォリオ全体の収益率は、

$$R = \sum_i R_i x_i \quad (1)$$

と表せる。 $\{x_i\}$ の状態はその銘柄が選択されたかそうでないかにしたがって 1, 0 を割り当てる。また、その定義から全証券に対して、 $\sum_i x_i = 1$ なる制約がある。投資に伴うリスクを分散と定義すると、

$$V = \sum_i \sum_j W_{ij} x_i x_j \quad (2)$$

を最小にする資産の組合せ $\{x_i\}$ を決める問題になる。ここで結合係数 W_{ij} は、通常過去の時系列データから求めた相関関数を使う場合が多い。式(1)を最大にし、式(2)を最小化する問題と考えると、投資比率が同じ場合には次のエネルギー関数を導入すればよい。

$$E = \sum_i \sum_j W_{ij} x_i x_j - A \sum_i R_i x_i + C \left(\sum_i x_i - N \right)^2 \quad (3)$$

ここで A, C は定数で、 N は選択されるべき銘柄の総数である。右辺第 3 項は選択銘柄数 N に対する制約条件である。この式(3)を最小化することで、最適解に近いポートフォリオ組合せ $\{x_i\}$ が得られる。

収益率 R_i には投資したい時期の値を用いるべきであるが、従来は過去の値を用いていたのでよい結果が得られなかった。そこで本稿では R_i にニューラルネットワークにより予測した収益率を用いる。 W_{ij} は各銘柄のある月の月間収益率を $R_i(h)$ とし、考える期間を H とすると、

$$W_{ij} = \frac{1}{H} \sum_{h=1}^H R_i(h) R_j(h) \quad (4)$$

と表せる。なお式(4)は R_i の相関の期間平均になっており、本稿では例として $H = 120$ (ヵ月) で計算しているので 1 ヶ月分の収益率の影響は小さいと考え、 W_{ij} の計算には予測値は入れないこととする。

3 ニューラルネットワークによる収益率予測

3.1 ニューラルネットワークの構造

次にニューラルネットワークによる収益率予測について述べる。最も簡素なニューラルネットワークモデルの構造は入力・中間・出力層からなり、各層は複数個のニューロンを構成し、各層のニューロン同士はシナプスによって結合している。通常ネットワークはバックプロパゲーション (以下 BP と略

す) 学習法により鍛え、その入出力に意味を持たせる。つまり、ある入力に対するするネットワークの出力を理想的な出力(教師)と比較しその誤差を少なくする方向へシナプス結合の重みを微小変動させる。このような学習を繰り返し行うことによりネットワークはいくつかのパターンを認識できるようになる。詳しい構造に関しては文献[7]にゆずる。

3.2 FBNNによる株式収益率データの予測

● 予測原理

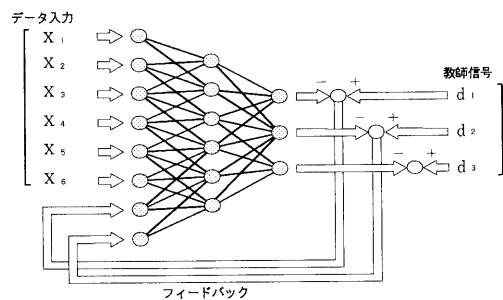
本稿では図2に示すようにフィードバック構造を持たせたニューラルネットワーク(Feed-Back Neural Network:以下FBNNと略す)を用いて収益率データの予想を行った。なお収益率データとしては1ヵ月単位の値(月間収益率)¹を用いている。その構造は入力層の6つの入力にそれぞれ過去6

ヵ月分のデータを与え、出力層の3つの出力に3ヵ月分の予想結果を出力させるものである。学習パターンについては1ヵ月ごとに時期をずらした入力・教師データ対を合成し、4から6パターン程度の学習パターンを用意した。また、出力の一部を教師信号との差分出力に加工し、フィードバックループを構成した。

収益率予測は、このような学習により得られたネットワークのフィードバックループを消去したものをを用いて、過去6ヵ月分の収益率データを入力し、3ヵ月分の予測収益率を得ている。

¹ データは日本証券経済研究所の「株式投資収益率」各号による。個別銘柄の月間収益率は、前月末に買い当月末に売った場合の収益率で、現金配当および株配・株主割当増資による収益を含む。

過去データの学習



学習パターンの合成

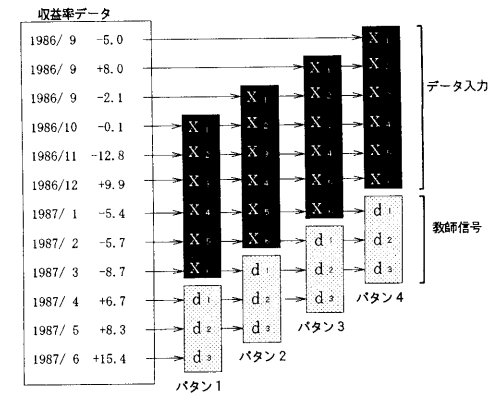


図2 時系列データ予想法

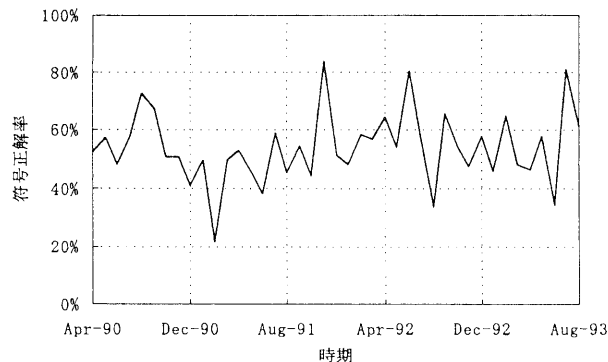


図3 符号正解率の時的推移

ここで、主なシミュレーション条件を示しておく。

ネットワーク構造：入力／中間／出力層＝6／3／3
 BP学習法：学習パターン数＝6、学習回数＝10000回、
 学習率＝0.4

● 予測結果

以上のFBNNを用いて東証一部上場銘柄のうち $M=100$ 銘柄の1985年5月～1993年8月までの100ヵ月について予想を行った。簡単のため予想の評価については株式収益率の符号正解率[%]を用いて表した。図3には1990年4月から1993年8月までの41ヵ月分の結果を示す。符号正解率とは予測値を x_i^* 、真値を x_i とした時

$$S(t) = \frac{\sum_{i=1}^M ps_i(t)}{\sum_{i=1}^M pb_i(t)} \quad ps_i(t) = \begin{cases} 1 \cdots \text{if}(x_i^* \cdot x_i > 0) \\ 0 \cdots \text{if}(x_i^* \cdot x_i \leq 0) \end{cases}$$

$$pb_i(t) = \begin{cases} 1 \cdots \text{if}(|x_i^*| > l[\%]) \\ 0 \cdots \text{otherwise} \end{cases}$$

のように予測値が ± 1 [%]以内の場合は評価対象としない符号正解率を算出した。以上の結果をまとめる。

1. 全範囲中91年半ばを中心とした2年あまりに良好な結果が得られている(図3)。
2. 時期別(100銘柄平均)で最高値83.7[%]を得た。

4 最適化手法

4.1 シミュレーテッドアニーリング(SA)法

組合せ最適化問題をニューラルネットワークを用いて解く場合、一般にエネルギー関数 $E(x)$ の最小化に帰着させ最適化する。その一手法であるSA法は、状態の確率的な変更で平衡状態を実現する。一般に多数の極小値を持つエネルギー関数に対して、SA法の本質は、状態変更確率に要素を導入することによりどの初期値から出発しても近似的に最小化できることである。

モンテカルロ法に従い元の最小化問題をボルツマン分布、

$$\exp\left(\frac{-E(x)}{T}\right) \quad (5)$$

の最大化に置き換える。 T は温度パラメータと呼ばれ、繰り返回数 t に依存して高温から低温に徐々に変化させる。本研究では、

$$T = \frac{T_0}{\log(1+t)} \quad (6)$$

とする。直感的には繰り返しの初期で高温にすることにより極小値から脱出する確率を大きくし、繰り返しの最後において低温化することで最小値の状態から逃げ出す確率を小さくすることである。しかし、SA法は一つの初期値からの探索であるので、状態空間の広い範囲の探索は難しいという欠点がある。その欠点を多数の初期値を取り入れ、解の選択を行うことで解決した手法について次で触れる。

4.2 拡張型SA法

本稿で提案する拡張型シミュレーテッドアニーリング法(以下拡張型SA法と略す)の特徴は、

- 多数の初期値を考える[8]。
- 解の選択処理を行う。

の二つである。なおこの概念は遺伝的アルゴリズム(以下GAと略す)的であるが、SA法における確率的な状態変更に対して、GAでは交叉という概念を用いている。蛇足ながら、GAにおいて交叉の代わりに乱数により個体の状態変更を行った場合にも、通常交叉を行った場合と同等のエネルギー最小化ができることが示せる。選択に関しては、探索の過程において出てくる多数の解に対して、エネルギーの大きい不良な解は切り捨て、その分エネルギーの小さい良好な解ほど多く残るような選択処理を行うことにより、効率的に最小値を探索できるようになっている。具体的な手順は、

1. 状態系列 $\mathbf{x}^i = \{x_1, \dots, x_M\}$ を L 個考える。
2. それぞれの系列に対してSA法と同様のモンテカルロ法に従う状態変更をおこなう。
3. $Q(\mathbf{x}^i) = \frac{\exp(-E(\mathbf{x}^i)/T)}{\sum_{j=1}^L \exp(-E(\mathbf{x}^j)/T)} \times L$ をそれぞれの状態系列に対して計算する。
4. それぞれの \mathbf{x}^i について $Q(\mathbf{x}^i)$ 個のコピーを作り、新たな状態系列を L 個作る。

図5にその概念図を示す。

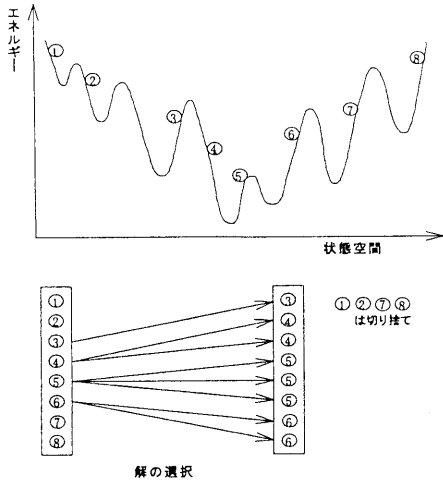


図5 拡張型SA法の概念図

例として次のようなエネルギー関数を持つランダムな系に対するの最小化を行い、拡張型SA法とSA法を比較する。

$$E = -\frac{1}{2} \sum_{ij} W_{ij} x_i x_j \quad (7)$$

ここで、 W_{ij} は平均0, 分散1のガウス分布、 x_i は-1から1までの値をとり、 $M=100$ とする。この問題は、各 x_i の結合のしかたがランダムなのでエネルギー関数の極小値が非常に多く、ポートフォリオ問題よりも最適解の探索が難しいと思われる。以下の図6にこのエネルギー関数の最小化を、SA法と拡張型SA法で行ったシミュレーション結果を示す。繰り返し数 $t=3000$ とした。

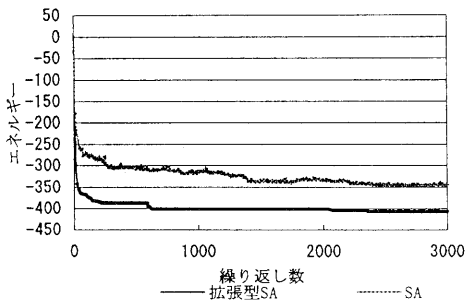


図6 エネルギー関数の最小化

拡張型SA法は単なるSA法の並列計算ではなく、解の探索過程で良好な解を選択しコピーしていくことで探索範囲を最適解付近に限定していくことができる。このシミュレーションからも拡張型SA法の方がエネルギーを小さくできることがわかる。しかし、多数の解を考えるために(図6の例では20個)、その分SA法よりも計算時間がかかるのが欠点である。次の章では、ポートフォリオに対するシミュレーション結果を示す。

5 ポートフォリオ選択問題への応用

計算機シミュレーションによるポートフォリオ選択の結果を以下に示す。使用した銘柄は100銘柄で、1984年1月から1993年12月までの月間収益率データを用いた。この100銘柄の内、同じ投資率で $N=10$ 銘柄選択する問題を取り上げた。この問題を先月、今月、FNNによる予測値のそれぞれの収益率データを用いて式(3)のもとに拡張型SA法により計算した。拡張型SA法の繰り返し数 $t=1000$ とし、状態系列の個数は $L=20$ とした。

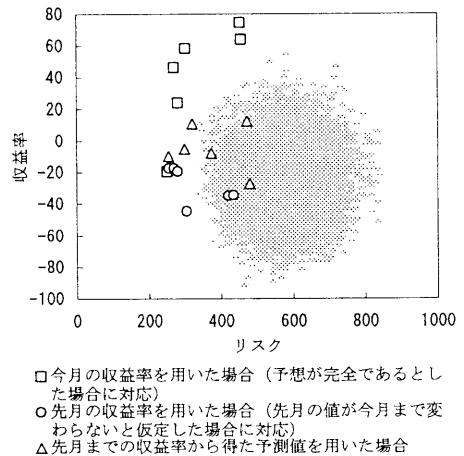


図7 収益率とリスク(1992年2月)

上の図7は選択された銘柄の良さを評価するため、1992年2月における収益率とリスクの指標を調べたものである。図に示した□は今月(2月)の取

益率を用いて求めたものであり、予測が完全であった場合を想定したものである。○は先月(1月)の収益率をそのまま用いて求めたものである。そして△が先月以前のデータからFBNNにより予測した収益率を用いて求めた値である。どの場合も式(3)において $C=100$ に固定し、 A の値を変化させてプロットしたものである。一方、小さな灰色の点は今月(2月)の収益率を用いて、対象銘柄からランダムに10銘柄を選択した16000個の組合わせに対する指標を示したものであり、同じく今月の収益率を用いて拡張型SA法で得た値よりも下方にある。つまり、固定したリスクに対してもっとも収益率が大きくなるような組合わせを選択できたことになる。しかし、実際にはポートフォリオを組む時点では今月の収益率は未知であるので、このような理想的なポートフォリオを組むことは一般に不可能である。そこで、FBNNによる予測収益率を用いて拡張型SA法を行った場合を見てみると、今月の収益率を用いた場合よりは当然劣るが先月の収益率を用いた場合よりはよい組合わせが選択できていることがわかる。

6 おわりに

本稿では、組合わせ最適化問題の解法としてよく知られているSA法を改良した拡張型SA法を紹介し、ニューラルネットワークによる時系列予測を含んだポートフォリオ選択問題に対して適用した。そして、拡張型SA法の有効性やポートフォリオ選択問題における収益率予測の重要性を示した。拡張型SA法に関しては、従来のSA法よりは確実に最適解探索能力は優れているといえるが初期値の数に比例して計算時間が長くなるという欠点がある。しかし、最近の計算機の処理能力の向上によりある程度許容できる範囲にあると考える。

収益率予測に関してはまだ実用的とはいえないが、実際の投資の場でポートフォリオ問題を考える場合には必ず予測した収益率のデータが必要になるという点でポートフォリオ選択問題とは切り離せない問題である。また、ニューラルネットワークによる予測で経済学の分野の人からよく指摘される点として、ネットワークがいわゆるブラックボックス化していて、時系列に含まれる経済的な要因がどのように使われて予測しているのかが解らないなどの指摘もある。

最後に、学習院大学経済学部の辰巳憲一教授に共に議論していただいたことに深く感謝する。なお、本研究の一部は文部省科学研究費補助金一般研究C(No.07650072)の援助によって行われた。

参考文献

- [1] 辰巳憲一「デリバティブと新金融商品の数学—基本と応用—」,東洋経済新報社(1995)
- [2] S. Kirkpatrick, C. D. Gellant Jr, and M. P. Vecchi: Optimization by Simulated Annealing, Science 220, 671 (1983)
- [3] 松葉育雄「ニューラルシステムによる情報処理」,昭見堂(1993)
- [4] 釜国男「遺伝的アルゴリズムによるポートフォリオ選択問題の一解法」,創価経済論集, XXIII, 3 (1993)
- [5] 玄光男「遺伝的アルゴリズムの国際会議とパッケージ1」, bit, 25, 12 (1993)
- [6] M. H. Markowitz: Portfolio Selection —Efficient Diversification Investments—, Yale University (1959)
- [7] 合原一幸編著 「ニューロ・ファジィ・カオス」より 松葉育雄 III.ニューラルネットワークの応用技術, オーム社(1993)
- [8] 茨木俊秀「組合せ最適化とスケジューリング問題: 新解法とその動向」, 計測と制御, 34-5, 340/346 (1995)