

実市場データをもとにポートフォリオ運用を行う エージェントを用いた人工市場による株価予測

Stock Price Prediction by the Artificial Market
using Agents Managing Portfolio on Real Market Data.

水田孝信¹ 和泉潔²

Takanobu MIZUTA¹, Kiyoshi IZUMI²

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

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

² 産業技術総合研究所 デジタルヒューマン研究センター

² Digital Human Research Center, Advanced Industrial Science and Technology, Japan

Abstract: We simulate stock returns using the artificial market using agents managing portfolio on real market data. In this model to couple real markets and an artificial market, cash inflows (outflows) to (from) agents are estimated by stock returns of the real markets. Our model can predict stock returns on some level.

1 はじめに

近年、人工市場を用いた研究がいくつかの成果をあげてきた。Artur et al. [1]は人工市場を用いて、Soros [2,3]が提唱した再帰性(reflexivity)、つまり、市場価格への正の自己参照的な期待価格の形成が存在し示ることを示した。他にも、バブル発生メカニズムの研究[4,5]、価格変動のカオス的特性の研究[6]、ファイナンスの重要理論である効率的期待仮説の検証[7]、外国為替市場の介入タイミングの検証[8,9]など多くの研究がある。しかし、これらの人工市場は単一の市場のみを取り扱っていたり、非常に少ない銘柄だけを売買したりすることが前提のモデルであった。

これに対して、実際の株式市場での多くの機関投資家の取引は、リスク分散を考えて数百の銘柄を組み合わせたポートフォリオ運用を行っている[10]。このようなポートフォリオ最適化の問題に、ニューラルネットワークなどの機械学習をもちいた先行研究があった[11]。これらの機械学習においては、人工市場のようにトレーダー同士の相互作用を考慮していない。

そこで本研究では、人工市場を複数銘柄の取引に対応できるように拡張を行い、複数銘柄取引の市場シミュレーションを行った。この際、ファクターにベットすることによりポートフォリオ運用を行うエ

ージェントを使用した。ファクターの値は実際の日本株式の実市場のデータを用いた。さらに、各エージェントの資金出入量を実市場のデータと人工市場でのシミュレーション結果が整合的になるように推定し、実際の投資家分布に近くなるように、人工市場のエージェント分布を決定した。よって、本研究のモデルは、実市場と人工市場の結合モデルである。

これにより、人工市場が予測する株価と実際の株価を比較することができ、その予測能力を評価した。さらに、機械学習エージェントの予測能力に関して、人工市場の予測株価も追加学習させると、予測能力が向上するかどうか調べた。

2 結合モデル

まず、人工市場に入れたエージェントについて述べる。以下の6種類のエージェントを5体ずつ、計30体入れた。

1. ノイズ
ランダムに売買する
2. ベータ
ベータが高い銘柄を買い、低いものを売る
3. 時価総額
時価総額が高い銘柄を買い、低いものを売る
4. 資本/時価総額
資本/時価総額が高い銘柄を買い、低いものを売る

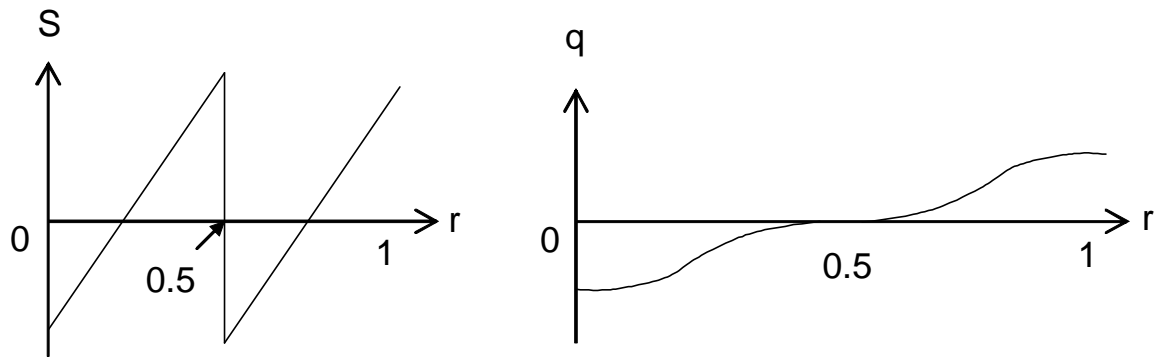


図1：(左)各エージェントの発注価格 S と各銘柄の魅力度 r の関係。(右)各エージェントの発注量 q と r の関係。

5. 12ヶ月リターン

過去12ヶ月の株価リターンが高いものを買、低いものを売

6. NN(ニューラルネットワーク)

ベータ、時価総額、資本/時価総額、過去12ヶ月リターンを学習信号、その後の1ヵ月リターンを教師信号として、過去12ヶ月分学習させ、未来の株価を予測して売買をする

ベータ、時価総額、資本/時価総額、12ヶ月リターンの4つのファクターは、Fama and French [12]が提唱した3つのファクターに過去12ヶ月リターンを加えた4ファクターモデル[13]で用いられるファクターであり、株価を説明するファクターとして広く知られている。

各エージェントは1ヵ月に一度、売買の判断を行う。本研究では、実際に日本で上場している銘柄のデータを用いた。まず、各時点での時価総額上位500銘柄に対してエージェント j は買いたい順位をつける。銘柄 i の魅力度 $r_j^i = (500 - \text{順位}) / 500$ を計算した後、多少の乱数も混ぜる。最も買いたい銘柄の r_j^i は1、最も売りたい銘柄の r_j^i は0となる。また、0.5以上を買いたい銘柄、0.5以下を売りたい銘柄と考える。次に発注量と発注値段を以下の式で決定する。(図1参照)

$$S_j^i = \frac{1}{10} S_{c,j} \left(|2r_j^i - 1| - \frac{1}{2} \right) \text{sgn}(2r_j^i - 1) \quad (1)$$

$$q_j^i = k_j q_{0,j}^i = |2r_j^i - 1|^{1/q_{c,j}^2} (1/q_{c,j}^2 + 1) \text{sgn}(2r_j^i - 1) \quad (2)$$

ここで、 S_j^i はエージェント j の銘柄 i への発注価格、 $S_{c,j}$ はエージェントごとの定数、 q_j^i はエージェント j の銘柄 i への発注量、 $q_{c,j}$ はエージェントごとの定数、 k_j は実市場から決定される係数(後で詳しく述べる)、 sgn は符号関数で、 $\text{sgn}(x) = 1 (x > 0)$ のとき、 $\text{sgn} = -1 (x < 0)$

のとき)、 $\text{sgn} = 0 (x = 0)$ のときである。

発注価格 S_j^i は、買い($r_j^i > 0.5$)のとき r_j^i が高いものは積極的に買うため、高くなる。 r_j^i がそんなに高くないものは安ければ買ってもよいという意味で、 S_j^i を低く設定する。また売り($r_j^i < 0.5$)のほうはその逆で、 r_j^i が非常に低いものは積極的に売りたいので S_j^i を低くし、 r_j^i がそんなに低くないものは高い値段で売れるのなら売りたいという意味で、 S_j^i を高く設定する。また、係数 $S_{c,j}$ によってどれくらい無理して売買するかが表される。

発注量 q_j^i はある程度 r_j^i が高いものは満遍なくもち、 r_j^i が低いものは満遍なく空売りするポートフォリオを組むことを想定して設定した。これは買持ちのポジションと売りもちのポジションを同額にするマーケットニュートラルを意識したものである。 $q_{c,j}$ が大きいほど r_j^i が極端に高い(低い)銘柄を集中投資する。本稿では $q_{c,j} \propto S_{c,j}$ とした。というのも、一般に、集中投資する投資家は値段を無理してでも買いにくい傾向があると考えられているからである。つまり、集中投資する運用者のほうが売買コストをかけることを仮定した。

$S_{c,j}, q_{c,j}$ は同一種類内(例えば、同じノイズエージェント同士)では異なる値とした。同じ種類のエージェントでも銘柄集中度が全てことなるエージェントになるようにした。

$S_{c,j}, q_{c,j}$ が決まった後、板寄せ方式で株価を決定する[14]。ここでは1ヶ月に一回、板寄せを一度だけ行いザラバは存在しない単純なモデルを採用している。

さて、 k_j はエージェント j への資金流入を示す。エージェントに資金が流入した場合は株式の保有量を大きくするために発注量が大きくなる。そのため、 k_j は大きくなるようにする。また、逆に資金がエー

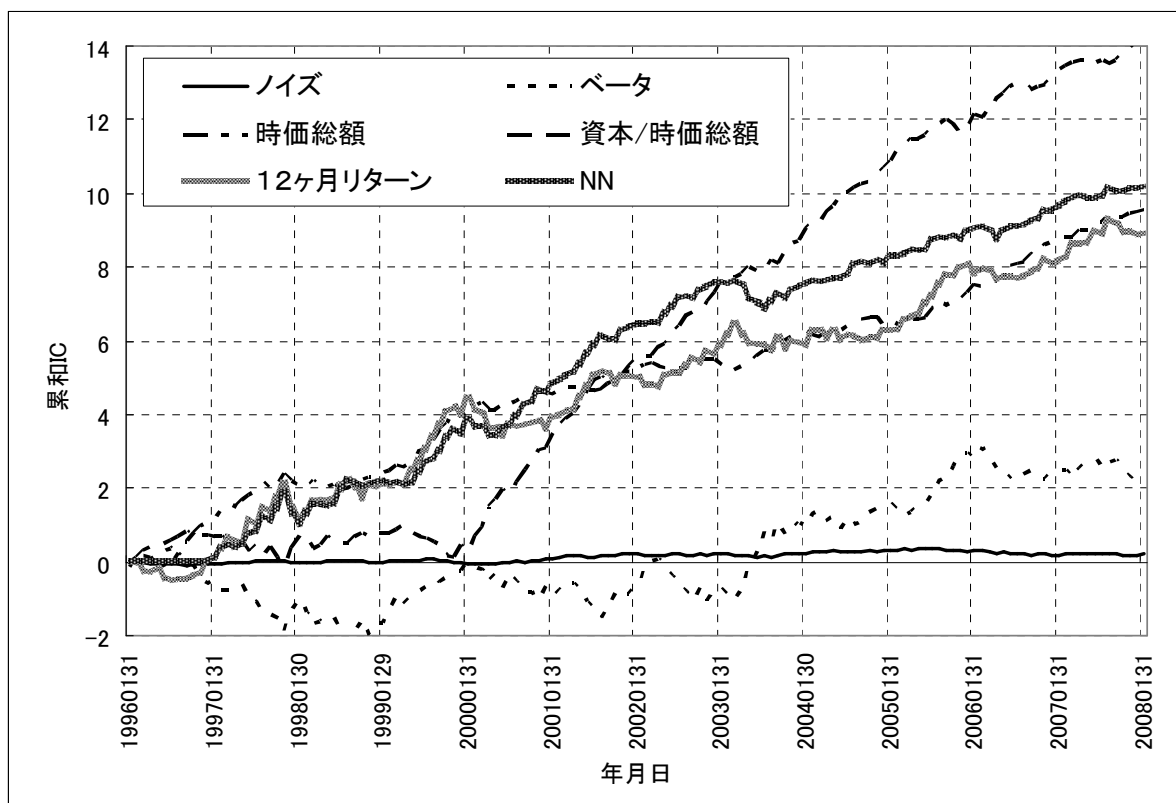


図2：各種類のエージェントの平均 r_j^i とその後1ヶ月の各銘柄のリターンとの順位相関 (IC) を月次で累和したものを時系列で示した。2000年以降、資本/時価総額とリターンに相関があることは良く知られている。また、NN (ニューラルネットワークによるモデル) も IC はプラスであるが資本/時価総額ほどのパフォーマンスではない。

エージェントから流出する場合は、保有している株式を手放さなければならない。これを表現するために k_j を負の値とする。なぜならば、魅力的 (r_j^i が高い) である銘柄をもともと保有していたと仮定し、魅力的である保有銘柄を処分売りしなければならなくなると考える。資金の流出が大きければ大きいほど k_j が負の方向に大きくなるとする。

k_j の値は、実市場のデータを用いて決定する。これにより、実際の市場参加者の分布と同じようなエージェントの構成を得ることができる。これらのエージェントを人工市場に入れることにより、今後の実市場の市場参加者の行動をシミュレーションする。 k_j は1ヵ月前から現在までの株価リターンを既知とし、人工市場上で最もその過去1ヶ月リターンを再現するように決める。つまり、過去1ヶ月のリターンを説明するエージェント分布を求める。より具体的には、

$$G(k_j) = \sum_i (Ra^i - Rr^i)^2 \quad (3)$$

で定義される $G(k_j)$ を最小にすることにより求める。

ここで Ra^i はある k_j のときの銘柄 i のリターン、 Rr^i は実際の過去リターンである。 $G(k_j)$ は複雑な関数形となるため、最小値を求めるために、PSO (Particle Swarm Optimization) [15] を使用した。

k_j , Sc_j , qc_j の決定を行い、その時点でのデータを用いて人工市場による価格形成を行い株価シミュレーションを行う。この一連の作業を1996年1月から2008年1月まで各月で行った。

3 結果

実市場と人工市場との接続モデルの結果を見る前に、まず、各エージェントが実市場でどれくらいのパフォーマンスなのかを見ておく。図2は各種類のエージェントの平均 r_j^i とその後1ヶ月の各銘柄のリターンとの順位相関 (IC: information coefficient) を月次で累和したものを時系列で示した。グラフが右上に伸びていけばパフォーマンスがよかったことを示す。1999年ごろを除いて、資本/時価総額とリターンに高い相関があることは良く知られている[16]。

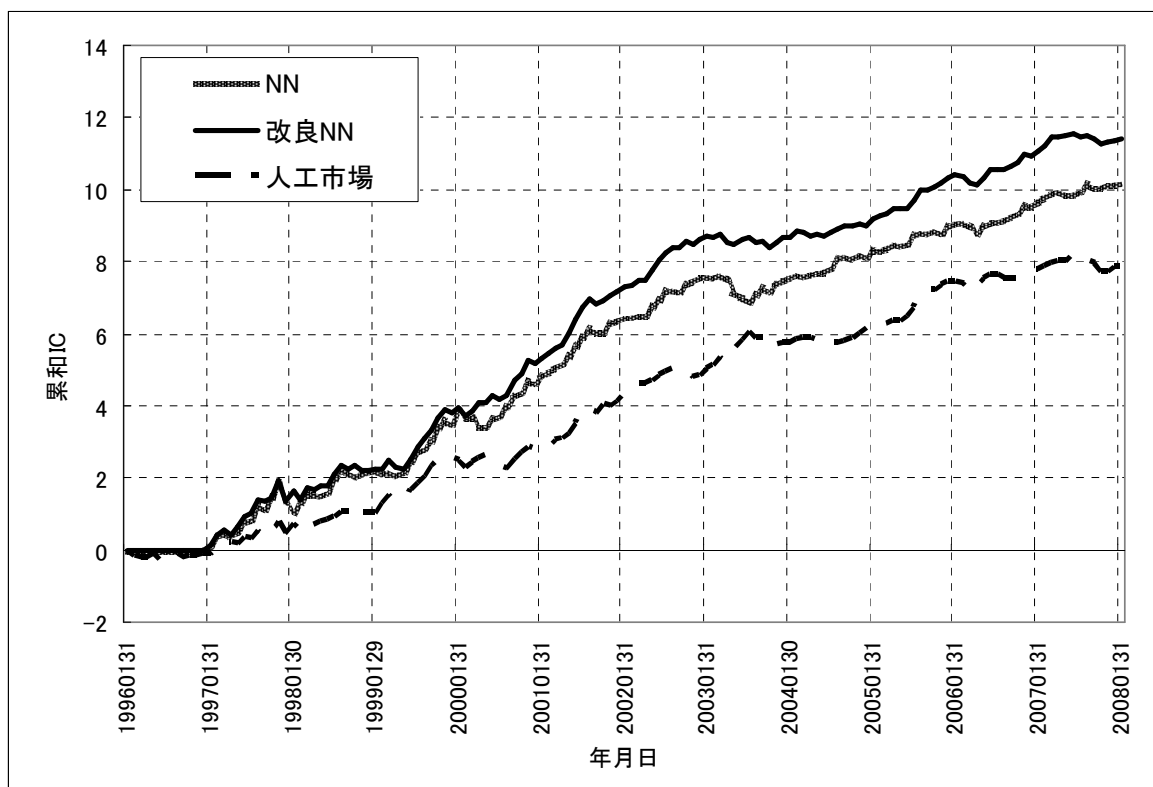


図3：NN、人工市場がシミュレートした株価（人工市場）、NNにさらに人工市場の株価も学習させたモデル（改良NN）について、平均 r_j とその後 1 ヶ月の各銘柄のリターンとの順位相関（IC）を月次で累和したものを時系列で示した。人工市場は高い IC を示しているが、NN にはおよばない。改良NNは NN と比べ IC が改善されている。

また、NN（ニューラルネットワークによるモデル）も IC はプラスであるが資本/時価総額ほどのパフォーマンスではない。ノイズとベータはほとんどリターンを取れていない。

次に、実市場と人工市場の接続モデルの予測パフォーマンスを見る。図3に接続モデルがシミュレーションした株価リターンと実際の株価リターンの順位相関の累和を示している。NNエージェントほどではないが、株価リターンをある程度予測できている。さらに、NNエージェントに人工市場がシミュレーションした株価もあわせて学習させた改良NNエージェントについても調べた。改良NNのパフォーマンスは、改良前のNNに比べ、改善されている。

図4は、各種類のエージェントの資金の出入量 k_j の平均の累和を時系列で示した。IC とはほとんど関係なく推移しており、単純に IC が高いエージェントに資金を流入させて $G(k_j)$ を最小化しているわけではないことが分かる。つまり、単純に有効であったファクターにベットして作成されたファクターとは全く異なる作られ方をしていると考えられる。

4 まとめ

実市場データをもとにポートフォリオ運用を行うエージェントを用い、実市場と人工市場を接続したモデルによる株価予測を行った。その結果、予測株価リターンは実際のリターンと正の相関を持っており、ある程度有効な株価シミュレーションができていることがわかった。

さらに NN を用いて過去データを学習するエージェントに、人工市場の株価シミュレーション結果も合わせて学習させた結果、パフォーマンスが改善することが分かった。

しかしながら今回のモデルは単純化しすぎており、実務上必要な株価予測精度は得られていないと考えている。

今後の課題として、さらに現実に近い複雑なエージェントを入れ、さらにシミュレーション精度を上げることが考えられる。今回は、ファクターにベットするエージェントが中心であったが、テクニカル重視のエージェントや外国株市場や債券市場などの他の市場との関係で取引しているエージェントを加

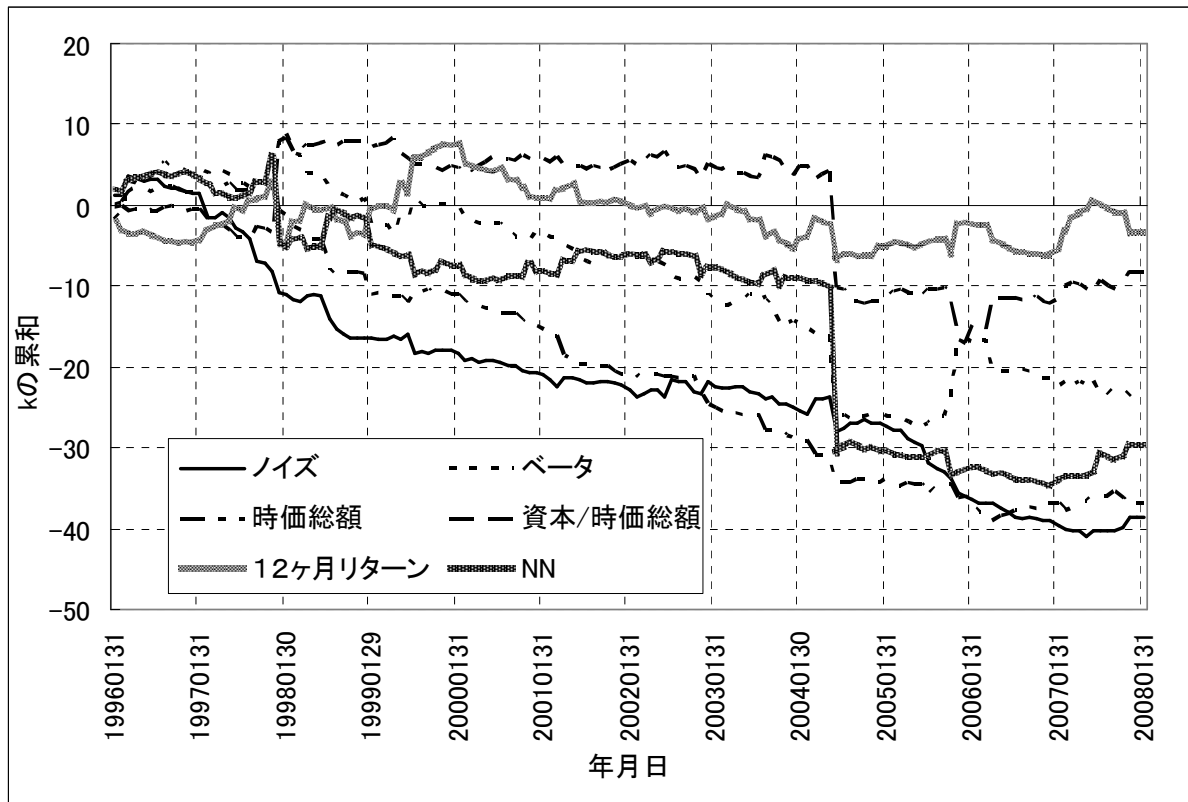


図4：各種類のエージェントの資金の出入量 k_i の平均の累和を時系列で示した。IC とはほとんど関係なく推移しており、単純に IC が高いエージェントに資金を流入させて $G(k_i)$ を最小化しているわけではないことが分かる。

えることが考えられる。また、実市場との接続方法として、今回は株価リターンを説明する資金流量を考えたが、この決定方法や決定精度にはまだ改善の余地があると考えている。決定方法に関しては、エージェントごとの残高情報を保持しそれを考慮して資金流量を扱うことが考えられる。また精度としては、各投資家の保有残高やその増減の実測値を用いることも考えられる。

このようにして精度を高め、実市場の再現性が高い人工市場を構築し、実務的に使えるレベルの株価シミュレーターを開発することが今後の研究課題である。

参考文献

- [1] Arthur, W., Holland, J., LeBaron, B., Palmer, R., and Tayler, P.: Asset Pricing Under Endogenous Expectations in an Artificial Stock Market, Technical report, Santa Fe Insitute (1996)
- [2] Soros, G: The Theory of Reflexivity. New York: Soros Fund Management, (1994)
- [3] Soros G: ソロスの錬金術 (訳書), 総合法令出版, (1996)
- [4] Steglitz, K. and Shapiro, D.: Simulating the Madness of Crowds: Price Bubbles in an Auction-Mediated Robot Market, Computational Economics, Vol. 12, (1998)
- [5] Mizuta, H., Steglitz, K., and Lirov, E.: Effects of Price Signal Choices on Market Stability, in WEHIA 99 (1999)
- [6] Lux, T.: The Socio-Economic Dynamics of Speculative Markets: Interacting Agents, Chaos, and the Fat tails of Return Distributions, Journal of Economic Behavior & Organization, Vol. 33, pp. 143-165 (1998)
- [7] Chen, S.-H. and Yeh, C.-H.: Genetic Programming and the Efficient Market Hypothesis, in Koza, J., Goldberg, D., and Fogel, D. eds., Genetic Programming: Proceedings of the 1st Annual Conference, pp. 45-53, the MIT Press (1996)
- [8] 松井宏樹, 東条敏: 人工市場アプローチによる仲介エージェントを用いた為替介入効果の分析, 人工知能学会論文誌, 20 卷, pp36-45, (2005)
- [9] 和泉潔, 松井宏樹, 松尾豊: 人工市場とテキストマイニングの融合による市場分析, 人工知能学会論文誌, 22 卷, pp.397-404, (2007)
- [10] 吉野貴晶: 定量分析入門, 日本経済新聞社,

(2003)

- [1 1] Shadbolt J., Taylor J.: Neural Networks and the Financial Markets; Predicting, Combining, and Portfolio Optimisation Springer, (2002)
- [1 2] Fama.E.F and French.K.R.: The Cross-Section of Expected Stock Return, Journal of Finance 47, pp.427-465, (1992)
- [1 3] Carhart, M.: On persistence of mutual fund performance, Journal of Finance 52, pp.57-82, (1997)
- [1 4] 和泉潔 : 人工市場 市場分析の複雑系アプローチ, 森北出版, (2003)
- [1 5] Kennedy, J. and Eberhart, R.: Particle Swarm Optimization, Proceedings of IEEE the International Conference on Neural Networks, (1995)
- [1 6] 大庭昭彦: 最新金融工学に学ぶ資産運用戦略, 野村証券金融工学研究センター, (2008)