

ポートフォリオ構築問題における非線形予測モデルの活用

猪瀬 悟史^{†a)} 鈴木 智也^{††b)}

Portfolio Selection Based on Nonlinear Time Series Prediction

Satoshi INOSE^{†a)} and Tomoya SUZUKI^{††b)}

あらまし マーコビッツの平均分散モデルでは、過去の收益率変動に基づいて確率分布を作成し、その平均値と標準偏差によって期待收益率とリスクを推定している。しかし実際の経済市場は複雑であるため、平均値を将来値とするような予測モデルが実用的であるとは考えにくい。そこで本研究は、より高次の非線形時系列予測モデルによって得られた予測値を期待收益率とみなすことで、推定精度の向上を実現した。しかし期待收益率の推定方法の変更により、リスクの定義に不一致が生じたため、本研究では経験的な予測誤差に基づいてリスクを算出した。この有用性を検証すべく、実データを用いた投資シミュレーションを行った結果、従来の平均分散モデルよりも高収益かつ安全に投資を実現できたので報告する。

キーワード 線形予測モデル、非線形予測モデル、ポートフォリオ理論、金融工学

1. まえがき

マーコビッツの平均分散モデル[1]によって分散投資の有用性が示され、組合せ比率を調整することで投資に伴うリスクを最小化することができる。この理論においては、将来に最も実現し得る收益率を期待收益率、その期待が外れる危険性をリスクとしており、これらを收益率の確率分布から推定する必要がある。しかし将来の確率分布を知ることは不可能であるため、過去の收益率変動による経験的な標本分布で代用することで、その平均値を期待收益率、標準偏差をリスクと定義している[1], [2]。

しかしこの操作は、移動平均という簡易な方法によって複雑な経済市場を予測することに相当し、現実的な予測モデルとして機能しない可能性がある。確かにFamaが述べるように金融市場は効率的でありランダムウォークするならば、予測モデルを高度化することは無意味である[3]。しかしその後の多くの実証研

究により、様々な金融市場の非効率性が報告されている。例えば、個別銘柄の收益率には非正規性、非独立性、非線形性が存在し、特に收益率を2乗した変動において高い自己相関性を確認できる。これはボラティリティ・クラスタリングと呼ばれ、金融市場の非効率性を裏付けている。なお、このような非効率性は特に日次收益率において確認され、週次や月次のように時間スケールを拡大すると諸特性は減弱し、その結果收益率変動はランダムウォークに近づくことが知られている[4]～[6]。その他、小形株効果や低PER効果など銘柄に関する規則性や、曜日効果や1月効果など季節性に関する規則性も確認され、これらは金融市場のアノマリーとして広く知られている[2], [7]。

このような規則性を取り込みながら、金融市場の予測モデルは発展してきた。特に收益率を2乗したボラティリティは收益率自体よりも強い規則性を有するため、GARCHモデルやその修正モデルが積極的に研究されている[8]。ここで、收益率変動にマルチングール性を仮定し期待收益率を0とすると、收益率の2乗であるボラティリティはリスクに相当する。ボラティリティであればGARCHモデルによる線形回帰式によってある程度予測できるため、マーコビッツの平均分散モデルにおけるリスク推定に役立つ。しかし收益率変動にマルチングール性を仮定することは、收益率の予

[†] 茨城大学大学院理工学研究科、日立市

Graduate School of Science and Engineering, Ibaraki University, 4-12-1 Nakanarusawa-cho, Hitachi-shi, 316-8511 Japan

^{††} 茨城大学工学部、日立市

College of Engineering, Ibaraki University, 4-12-1 Nakanarusawa-cho, Hitachi-shi, 316-8511 Japan

a) E-mail: 12nd202x@hcs.ibaraki.ac.jp

b) E-mail: tsuzuki@mx.ibaraki.ac.jp

測を諦めたことになる^(注1)。

一方、収益率自体の予測モデルには、古くから自己回帰（AR）モデルや自己回帰移動平均（ARMA）モデル等が用いられている[8]。しかし先述のとおり、ボラティリティのように明確な法則性をもたないため、予測精度はあまり高くない。そこで、マクロ経済指標や個別銘柄の財務指標などのファンダメンタルな要因も利用するマルチファクターモデルや、更に投資家の主観的な見通しをもとの自信度に応じて混合するブラック・リッターマンモデルが提案されている[2]。しかし本研究では、マーコビツツの平均分散モデルの改良を目的とするため、ヒストリカルな収益率変動のみによってポートフォリオを構築したい。そこでまず、平均分散モデルで用いられる移動平均予測をARモデルに拡張することで、将来の収益率を積極的に予測する^(注2)。しかし当然ながら、平均分散モデルにおける期待収益率の定義と異なるため、リスク評価においてつじつま合せが必要となる。そこでリスクを根本的に考えれば、期待値が実現しない可能性に相当するので、予測誤差に基づいてリスクを評価する。これにより期待収益率の予測モデルは選択自由になり、従来の移動平均予測に束縛されない。

更に本研究では、ARモデルに非線形的な近似能力をもたせた時系列予測モデルも検討する。本来のARモデルは線形自己回帰式であるため、説明変数と目的変数の関係を平面で近似する。しかし類似した変動パターンごとで細かく平面近似をすれば、大域的には非線形な曲面で近似できる。これはカオス変動を予測する局所線形近似法[9]～[12]に基づいているが、近似面として平面以外に曲面も許す柔軟さは、予測対象がカオスでなくても有意義である[12], [13]。なお本研究では、この局所線形近似法に基づくARモデルを非線形予測モデルと称する。この近似面に制約を与えることでARモデルや移動平均モデルを再現できるため、マーコビツツの平均分散モデルと親和性が高く、拡張しやすい予測モデルである。

まず次章でマーコビツツの平均分散モデルを紹介した後、それぞれの予測モデルでどの程度収益率を予測できるのか検証する(3.)。もし移動平均予測よりも高次の予測モデルによって予測精度を向上できれば、その予測値を平均分散モデルに用いればよい。しかし先述のようにリスクの定義を変更する必要があるため、過去の予測誤差によってリスクを評価できるように平均分散モデルを書換える(4.)。この妥当性を検証すべ

く、実際の株価収益率データを用いて投資シミュレーションを行い、運用成績の観点から従来手法との比較を行う(5.)。更に、ポートフォリオに利用する投資銘柄の選択法について検討しつつ(6.)、多銘柄のポートフォリオによるリスク分散効果を検証する(7.)。以上の結果について、週次や月次のような時間スケールの違いによる影響(8.)や、リーマン・ショック前後の異なる期間における有用性を検証(9.)した後、最後にまとめと今後の展望を述べる(10.)。

2. マーコビツツの平均分散モデルの概要

時刻 t における投資対象銘柄 i ($i = 1, 2, \dots, N$) の取引価格を $x_i(t)$ とすると、そのときの収益率 $r_i(t)$ は

$$r_i(t) = \frac{x_i(t) - x_i(t-1)}{x_i(t-1)} \quad (1)$$

として算出される。次に、未知である時刻 $t+1$ における収益率 $r_i(t+1)$ を考える。この収益率が $r_i^{(k)}(t+1)$ となる確率を $P_i^{(k)}$ とすると(ただし $\sum_{k=1}^K P_i^{(k)} = 1$), 期待収益率 $\tilde{r}_i(t+1)$ とその分散値 $\tilde{\sigma}_i^2(t+1)$ は次のように書ける。

$$\tilde{r}_i(t+1) = \sum_{k=1}^K P_i^{(k)} r_i^{(k)}(t+1) \quad (2)$$

$$\tilde{\sigma}_i^2(t+1) = \sum_{k=1}^K P_i^{(k)} \left[r_i^{(k)}(t+1) - \tilde{r}_i(t+1) \right]^2 \quad (3)$$

不確定性下での意思決定では、式(3)の標準偏差 $\tilde{\sigma}$ をリスクとみなす[1], [2]。

更に、組合せ比率 $\{d_i\}$ でポートフォリオを組むとき(ただし $\sum_{i=1}^N d_i = 1$), ポートフォリオの期待収益率 $\tilde{r}_p(t+1)$ とその分散値 $\tilde{\sigma}_p^2(t+1)$ は次のように書ける。

$$\tilde{r}_p(t+1) = \sum_{i=1}^N d_i \tilde{r}_i(t+1) \quad (4)$$

$$\tilde{\sigma}_p^2(t+1) = \sum_{i=1}^N \sum_{j=1}^N d_i d_j \tilde{\sigma}_{ij}(t+1) \quad (5)$$

ここで $\tilde{\sigma}_{ij}(t+1)$ は共分散であり、以下で得られる。

$$\tilde{\sigma}_{ij}(t+1) = \sum_{k=1}^K P_{ij}^{(k)} \left[r_i^{(k)}(t+1) - \tilde{r}_i(t+1) \right] \left[r_j^{(k)}(t+1) - \tilde{r}_j(t+1) \right] \quad (6)$$

(注1)：しかしオプション取引においてロングストラドルを組めば、収益率の予測は不要である。

(注2)：3. で述べるように、移動平均予測は AR モデルの全係数を統一した簡易モデルである。

$$\cdot [r_j^{(k)}(t+1) - \tilde{r}_j(t+1)] \quad (6)$$

$P_{ij}^{(k)}$ は、時刻 $t+1$ において、 $r_i^{(k)}(t+1)$ かつ $r_j^{(k)}(t+1)$ となる確率である。

ただし、これまでに登場した $P_i^{(k)}$, $P_{ij}^{(k)}$, $r_i^{(k)}(t+1)$ および $r_j^{(k)}(t+1)$ は未知なので、過去の収益率変動の標本分布を、将来の確率分布と見立てて推定される [2]。つまり、式 (2), (3), (6) を以下のように書き換える。

$$\begin{aligned} \tilde{r}_i(t+1) &= \bar{r}_i(t) \\ &= \frac{1}{T} \sum_{a=0}^{T-1} r_i(t-a) \end{aligned} \quad (7)$$

$$\begin{aligned} \tilde{\sigma}_i^2(t+1) &= \sigma_i^2(t) \\ &= \frac{1}{T} \sum_{a=0}^{T-1} [r_i(t-a) - \bar{r}_i(t)]^2 \end{aligned} \quad (8)$$

$$\begin{aligned} \tilde{\sigma}_{ij}(t+1) &= \sigma_{ij}(t) \\ &= \frac{1}{T} \sum_{a=0}^{T-1} [r_i(t-a) - \bar{r}_i(t)] \cdot [r_j(t-a) - \bar{r}_j(t)] \end{aligned} \quad (9)$$

ここで T は参照する過去のデータ長である。これらを式 (4) (5) に代入することで、

$$\tilde{r}_p(t+1) = \sum_{i=1}^N d_i \bar{r}_i(t) \quad (10)$$

$$\tilde{\sigma}_p^2(t+1) = \sum_{i=1}^N \sum_{j=1}^N d_i d_j \sigma_{ij}(t) \quad (11)$$

によって、将来時刻 $t+1$ におけるポートフォリオの期待収益率及びリスクを算出することができる。

更に、期待収益率 $\tilde{r}_p(t+1)$ を最大に、リスク $\tilde{\sigma}_p(t+1)$ を最小にする組合せ比率 $\{d_i\}$ を採用することで、合理的な投資を行うことができる。本研究では一般的な指標として、シャープレシオ S_r を導入し [14]、

$$S_r = \frac{\tilde{r}_p - r_f}{\tilde{\sigma}_p} \quad (12)$$

を最大にする組合せ比率 $\{d_i\}$ を採用する。ここで、 r_f はリスクフリーレートであり、無担保コールレート翌日物を使用した [15]。

なお、 \tilde{r}_p は買い取引における期待収益率を想定している。もし $\tilde{r}_p < 0$ の場合、そのまま買い取引を行えば損失となるので、全銘柄を空売りすることで \tilde{r}_p の符号を反転させる。その後、再度式 (12) の S_r を最大化することで最終的な組合せ比率 $\{d_i\}$ を得る。つまり

買い取引において $\tilde{r}_p < 0$ となる場合、 S_r を最小化する $\{d_i\}$ によって空売りすることになる。空売りは信用取引であるため現在の所持金を消費しないが、期待リターンの見積りを誤れば買戻し時に値上り分の損失を負うことになる。つまり空売りを行っても、期待リターンに対して背負うリスクは買い取引と同じなので、空売りする場合の取引額は、買い取引の場合と同額にした。更に、式 (12) の評価関数を最大化することは凸二次計画問題に相当するため、内点法を用いて最適解を算出した。

3. 時系列予測モデルの有用性

前章のポートフォリオ理論は式 (7) の期待収益率 $\tilde{r}_i(t+1)$ の推定において、移動平均予測を行っているとみなすことができる。確かに、価格変動のトレンドが安定している場合においては有効であろう。また、同一のポートフォリオを長期間保持することで短期的なノイズは相殺され、収益率は移動平均値に近づく場合もあるかもしれない。しかし、実際の市場の振舞いは複雑であり、同一のトレンドが長期間維持し続けることはまれである。したがって、トレンドの転換に柔軟に対応するには、短期的に期待収益率 $\tilde{r}_p(t+1)$ やリスク $\tilde{\sigma}_p(t+1)$ を再計算し、ポートフォリオを修正する必要がある。そこで本研究では、移動平均予測よりも緻密な時系列予測モデルを利用したい。

例えば、移動平均予測に用いたヒストリカルデータとその将来値の関係を、超平面によって近似すれば自己回帰 (AR) 予測モデルになる。この予測モデルを適用すれば、

$$\tilde{r}_i(t+1) = \sum_{a=0}^{T-1} \beta_a r_i(t-a) + \beta_T \quad (13)$$

によって式 (7) の期待収益率を計算できる。モデル係数 β_a ($a = 0, 1, \dots, T$) の推定には、過去 L 期分の時系列変動 $\{r_i(t), r_i(t-1), \dots, r_i(t-L+1)\}$ を学習データとして、最小二乗法を適用する。ここで、

$$\begin{aligned} &r_i(t-a) \\ &= [r_i(t-a), r_i(t-a-1), \dots, r_i(t-a-(T-1))] \end{aligned} \quad (14)$$

とし、

$$\mathbf{X} = \begin{bmatrix} r_i(t-1) & 1 \\ r_i(t-2) & 1 \\ \vdots & \vdots \\ r_i(t-(L-T)) & 1 \end{bmatrix}, \quad \mathbf{F} = [\beta_0, \beta_1, \dots, \beta_T]^t$$

$$\mathbf{Y} = [r_i(t), r_i(t-1), \dots, r_i(t-(L-T)+1)]^t$$

と書くと、 $\mathbf{Y} = \mathbf{X}\mathbf{F}$ の近似誤差を最小にする係数 \mathbf{F} は

$$\tilde{\mathbf{F}} = [\mathbf{X}^t \mathbf{X}]^{-1} \mathbf{X}^t \mathbf{Y} \quad (15)$$

によって推定できる。更に、この $\tilde{\mathbf{F}}$ を式 (13) で利用する。このように AR 予測モデルは一つの回帰超平面を有するので、線形予測モデルとして分類できる。一方、移動平均予測は、式 (13) 中の傾き β_a を全て $\frac{1}{T}$ に、切片 β_T を 0 に制約した AR 予測モデルとみなせる。

次に、より一般化した予測モデルを導入すべく、回帰超平面を曲面に可変できる非線形予測モデルを考える。その前処理として、学習データ $r_i(t)$ に対して埋込定理 [16], [17] を適用し、多次元状態ベクトル $\mathbf{v}_i(t)$ を構成する。

$$\mathbf{v}_i(t) = [r_i(t), r_i(t-\tau), \dots, r_i(t-(d-1)\tau)] \quad (16)$$

ここで、 τ は遅れ時間、 d は埋込次元を表す。なお $\tau = 1$ かつ $d = T$ とみなせば $\mathbf{v}_i(t) = \mathbf{r}_i(t)$ となるので、式 (13) は

$$\tilde{r}_i(t+1) = [\mathbf{v}_i(t) \ 1] \cdot \mathbf{F} \quad (17)$$

と書ける。しかし、非線形予測モデルでは、説明変数 $\mathbf{v}_i(t)$ の座標に応じて局所的に回帰超平面の傾き \mathbf{F} を可変させたい。これは過去の変動パターン $\mathbf{v}_i(t)$ に応じて、将来値 $r_i(t+1)$ との関係性 \mathbf{F} が変化することを想定している。つまり、 $\mathbf{v}_i(t)$ の座標が近い学習データほど予測にとって重要であるので、 \mathbf{F} の推定には重み付き最小二乗法を用いる。

まず、予測対象点 $\mathbf{v}_i(t)$ と各学習データ $\mathbf{v}_i(t-a)$ 間のユークリッド距離

$$l_i(t, a) = |\mathbf{v}_i(t) - \mathbf{v}_i(t-a)| \quad (18)$$

を算出し、重み係数を

$$w_i(t, a) = \exp(-bl_i(t, a)) \quad (19)$$

とする。次に、これらを対角成分に収めた対角行列を

$$\mathbf{W} = \begin{bmatrix} w_i(t,1) & 0 & \dots & 0 \\ 0 & w_i(t,2) & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \dots & 0 & w_i(t, L+(d-1)\tau+1) \end{bmatrix} \quad (20)$$

と書くと、 $\mathbf{WY} = \mathbf{WXF}$ の近似誤差を最小にする

係数 \mathbf{F} は

$$\tilde{\mathbf{F}} = [\mathbf{X}^t \mathbf{W}^t \mathbf{W} \mathbf{X}]^{-1} \mathbf{X}^t \mathbf{W}^t \mathbf{W} \mathbf{Y} \quad (21)$$

によって推定される。ここで

$$\mathbf{X} = \begin{bmatrix} \mathbf{v}_i(t-1) & 1 \\ \mathbf{v}_i(t-2) & 1 \\ \vdots & \vdots \\ \mathbf{v}_i(t-L+(d-1)\tau+1) & 1 \end{bmatrix},$$

$$\mathbf{Y} = [r_i(t), r_i(t-1), \dots, r_i(t-L+(d-1)\tau)]^t$$

である。

この $\tilde{\mathbf{F}}$ の推定は局所的には平面近似であるが、予測開始点 $\mathbf{v}_i(t)$ に応じて近似面が変わるために、大域的には曲面近似となる。このような局所線形近似法 [9]～[12] によって自己回帰予測モデルを非線形モデルに拡張したものを、本研究では非線形自己回帰予測モデルと呼ぶ。重み付けの方法については文献 [11] に倣い、 $b = 1$ とした。なお $\tau = 1$ かつ $b = 0$ として重み係数 $w_i(t, a)$ を全て 1 にすれば、非線形予測モデル（式 (21)）は線形予測モデル（式 (15)）と等価になる。更に係数 β_a を $\frac{1}{T}$ に、 β_T を 0 にすれば移動平均予測に帰着する。つまり非線形予測モデルは、線形予測と移動平均予測を内包するように一般化した予測モデルである。

収益率の予測精度の評価には、次式の相関係数 ξ を用いる。

$$\xi = \frac{\sum_i (r_i - \langle r \rangle)(\tilde{r}_i - \langle \tilde{r} \rangle)}{\sqrt{\sum_i (r_i - \langle r \rangle)^2} \sqrt{\sum_i (\tilde{r}_i - \langle \tilde{r} \rangle)^2}} \quad (22)$$

ここで r は真値データ、 \tilde{r} は予測値データである。また、 i はデータのインデックスを表し、 $\langle \cdot \rangle$ は i についての平均を表す。

各予測モデルの次数 T 及び非線形予測モデルの遅れ時間 τ の決定には、過学習による汎化能力の低下を防ぐために交差確認法 [18], [19] を適用する。まず学習データ全体をパート A とパート B に分割し、各パートの学習データとして他方のパートを予測する。そして式 (22) によってモデルのゆう度 ξ_A と ξ_B を得る。ここで予測されたパートは係数推定に用いられていないので、汎用性のあるゆう度を求めることができる。最終的には、予測モデルのゆう度を $\eta_i = (\xi_A + \xi_B)/2$ とし、これを最大化するモデルパラメータ T 及び τ を用いる。

表 1 各予測法による株価データ (200 銘柄) の予測精度 $\{\xi_i\}$ 。 $\bar{\xi}$ は平均値, σ_ξ は標準偏差である。太字の数値は、各項目の最良値を示す。

Table 1 Results of prediction accuracy $\{\xi_i\}$ by each prediction model. The index $\bar{\xi}$ means the mean value of $\{\xi_i\}$, and σ_ξ means the standard deviation of them. Each bold type means the best score of each category.

	移動平均予測	線形予測	非線形予測
$\bar{\xi}$	-0.033	0.051	0.053
σ_ξ	0.023	0.061	0.061
$\max \{\xi_i\}$	0.030	0.224	0.225
$\min \{\xi_i\}$	-0.089	-0.090	-0.093

さて、以上のプロセスを実際の株価データ [20] に適用してみる。予測モデルのゆう度 η は、2000 年 4 月 7 日以前の約 5 年間（土日祝日を除く 1250 日分の営業日）の学習データ ($L = 1250$) に対するものであり、予測精度 ξ の算出には、それ以降の約 5 年間の評価データ (1250 日分) を用いた。

表 1 に、全 200 銘柄に対する予測結果をまとめた。なお各予測モデルのモデルパラメータは、上記の交差確認法において全 200 銘柄の平均ゆう度を最大化するように設定した。具体的には、移動平均予測では $T = 211$ 、線形予測では $T = 5$ 、非線形予測では $T = 5$ 及び $\tau = 1$ と設定された。結果として、予測精度の平均値 $\bar{\xi}$ を比較すると、移動平均予測が最も悪い。

更に詳細に検証すべく、ウィルコクソンの符号順位和検定を行う。ここで「予測法 α の予測精度は、予測法 β の予測精度と差がない」という帰無仮説を立てる。次に銘柄 i の予測法 α の予測精度 $\xi_{i,\alpha}$ と予測法 β の予測精度 $\xi_{i,\beta}$ の差 g_i を算出し、 $|g_i|$ の小さい方から順位を付ける ($i = 1 \sim N$)。そして $\xi_{i,\alpha} > \xi_{i,\beta}$ の組の順位和と $\xi_{i,\alpha} < \xi_{i,\beta}$ の組の順位和のうち、小さい方を R として、標準正規分布に従う検定統計量 Z を次式によって算出する。

$$Z = \frac{\left| R - \frac{N(N+1)}{4} \right|}{\sqrt{\frac{N(N+1)(2N+1)}{24}}} \quad (23)$$

もし $Z \geq 1.645$ であれば有意水準 5% で、「予測法 α の予測精度は、予測法 β の予測精度より高い」ことになる。

まず線形、非線形予測法の予測精度が移動平均予測の予測精度に対し優位性があるか検定を行った。その結果、それぞれの検定統計量 Z は 9.504, 9.170 となつた。次に非線形予測が線形予測法より優位であるかを

検定したところ、 $Z = 0.215$ であった。これらの結果より、非線形予測は線形予測と比較して優位性を確認できなかったが、移動平均予測と比較すれば、線形予測や非線形予測の優位性は明らかである。

しかし線形予測及び非線形予測の予測精度 $\bar{\xi}$ はわずかに正であるため、実際の投資には役立たないかもしれない。だが一方では、経済データは予測不可能なランダム現象だとすら考えられているため [3], [21]、わずかでも正の予測精度を実現することは重要なことなのかもしれない。そこで実際の投資に対する有用性について、5. 以降の投資シミュレーションによって検証する。

4. 平均分散モデルの書換え

前章の時系列予測モデルをポートフォリオ理論に用いるには、従来の式 (7)～(9) を書き換える必要がある。そこで、これらの原形である式 (2), (3), (6) に戻って再考する。

式 (2) によれば、期待収益率 $\tilde{r}_i(t+1)$ はまさしく最も出現確率が高い将来値であるので、各予測モデルから得られた予測値を割り当てればよい。従来の式 (7) を移動平均予測と捉えれば、この考え方は従来と変わっていない。予測値を式 (4) に代入すれば、ポートフォリオの期待収益率 $\tilde{r}_p(t+1)$ を推定できる。

更に式 (3) では、予測誤差の期待値をリスク $\tilde{\sigma}_i^2(t+1)$ としている。そこで、式 (3) 中の予測誤差 $r_i^{(k)}(t+1) - \tilde{r}_i(t+1)$ を $\sigma_i^{(k)}(t+1)$ と書くと、

$$\tilde{\sigma}_i^2(t+1) = \sum_{k=1}^K P_i^{(k)} \left[\sigma_i^{(k)}(t+1) \right]^2 \quad (24)$$

となる。しかし、従来の平均分散モデルと同様に、 $P_i^{(k)}$ 及び $\sigma_i^{(k)}(t+1)$ は未知なので、過去の経験的な標本分布に基づいて推定することにする。つまり、過去の予測誤差 $r_i(t') - \tilde{r}_i(t')$ ($t' \leq t$) を記録しておき、

$$\tilde{\sigma}_i^2(t+1) = \frac{1}{T} \sum_{a=0}^{T-1} [r_i(t-a) - \tilde{r}_i(t-a)]^2 \quad (25)$$

に代入することでリスク $\tilde{\sigma}_i^2(t+1)$ を推定する。なお従来の平均分散モデルでは、式 (3) から式 (8) の変形において、 $\sigma_i^{(k)}(t+1)$ 中の $r_i^{(k)}(t+1)$ のみをヒストリカルデータで代用している。しかし本研究のように予測を陽に想定する場合、リスクの本質は予測誤差であるので、あえて $\sigma_i^{(k)}(t+1)$ 全体を過去の予測誤差に基づいて推定する。

なお、式(6)の共分散についても同様に、

$$\begin{aligned}\tilde{\sigma}_{ij}(t+1) &= \sum_{k=1}^K P_{ij}^{(k)} \sigma_i^{(k)}(t+1) \sigma_j^{(k)}(t+1) \\ &= \frac{1}{T} \sum_{a=0}^{T-1} [r_i(t-a) - \tilde{r}_i(t-a)] \cdot [r_j(t-a) - \tilde{r}_j(t-a)]\end{aligned}\quad (26)$$

とする。これを式(5)に代入すれば、ポートフォリオのリスク $\tilde{\sigma}_p^2(t+1)$ を推定できる。

しかし実際の使用では、式(26)の T について工夫が必要である。前章で述べた時系列予測モデルの観点では、この T はモデルの複雑さに相当し、 T が大きいほど過学習を引き起こす。よって交差確認法や赤池情報量基準に従うと、 T は小さい値になる場合がある。その結果、式(26)に用いる予測誤差のデータ長 T が短くなるため、ある 2 銘柄が偶然に逆相関になる可能性が高まる。そこで式(5)を変形すると、

$$\begin{aligned}\tilde{\sigma}_p^2 &= \sum_{i=1}^N \sum_{j=1}^N d_i d_j \tilde{\sigma}_{ij} \\ &= \sum_{i=1}^N d_i^2 \tilde{\sigma}_i + 2 \sum_{i=1}^N \sum_{j>i}^N d_i d_j \tilde{\sigma}_{ij}\end{aligned}\quad (27)$$

なる。ここで、ある 2 銘柄 ($i = 1, 2$) の相関係数 ρ は以下のように算出される。

$$\rho = \frac{\tilde{\sigma}_{12}}{\tilde{\sigma}_1 \tilde{\sigma}_2} \quad (28)$$

これらの銘柄が完全逆相関 $\rho = -1$ であると想定すると、 $\tilde{\sigma}_{12} = -\tilde{\sigma}_1 \tilde{\sigma}_2$ となる。このとき、 d_1 と d_2 以外の組合せ比率を 0 にすれば、

$$\begin{aligned}\tilde{\sigma}_p^2 &= d_1^2 \tilde{\sigma}_1^2 + d_2^2 \tilde{\sigma}_2^2 + 2d_1 d_2 \tilde{\sigma}_{12} \\ &= d_1^2 \tilde{\sigma}_1^2 + d_2^2 \tilde{\sigma}_2^2 - 2d_1 d_2 \tilde{\sigma}_1 \tilde{\sigma}_2 \\ &= (d_1 \tilde{\sigma}_1 - d_2 \tilde{\sigma}_2)^2\end{aligned}\quad (29)$$

となる。更に両辺の平方根をとれば、

$$\tilde{\sigma}_p = |d_1 \tilde{\sigma}_1 - d_2 \tilde{\sigma}_2| \quad (30)$$

が得られる。このとき、 $d_1 : d_2 = \tilde{\sigma}_2 : \tilde{\sigma}_1$ かつ $d_{i \notin \{1,2\}} = 0$ のポートフォリオを組めば $\tilde{\sigma}_p = 0$ となるため、式(12)のシャープレシオ S_r は発散する。つまり、2 銘柄の予測誤差が偶然に逆相関になれば、 S_r を最大化する組合せ比率はそれぞれの標準偏差 $\tilde{\sigma}_1$ と $\tilde{\sigma}_2$ のみによって決まる。しかしこの場合、期待リ

ターン $\{\tilde{r}_i\}$ は一切考慮されていないので、不適切なポートフォリオである。このようなポートフォリオが構築される危険性は、式(26)中の T が小さいほど拡大する。そこで以後の計算機実験では、 $\max\{S_r\}$ が有限値に収束するように T を過去 1 年分に調整した。

5. 投資シミュレーション

3. で述べた時系列予測モデルの活用、そして 4. で書き換えたポートフォリオ理論の効果を検証すべく、200 銘柄の実データ [20] を用いて投資シミュレーションを行う。本章ではまず $N = 2$ とし、2 銘柄のポートフォリオを組む。なお、投資期間中に投資銘柄を変更しない。したがって 200C₂ 通りの投資結果が得られるので、これらの統計量を算出する。なお、2000 年 4 月 8 日から投資行動を開始し、それ以降の約 5 年間（土日祝日を除く 1250 日分の営業日）において投資を続行する。以下に、1 日の投資行動の手順を述べる。

手順 1 学習データの読み込み

ポートフォリオに用いる N 種の銘柄について、本日から過去 L 日までの株価収益率データを計算機に読み込む。なお、式(1)の取引価格 x_i には始値を用いる。

手順 2 時系列予測モデルの活用

銘柄ごとに、翌営業日の期待収益率を予測し、その予測値を \tilde{r}_i とする。更に、式(9)または式(26)に基づいて、共分散 $\tilde{\sigma}_{ij}$ を算出する。

手順 3 ポートフォリオの算出

手順 2 で求めた統計量を式(4)と式(5)に代入することで、ポートフォリオの期待収益率 \tilde{r}_p とリスク $\tilde{\sigma}_p$ を得る。

手順 4 シャープレシオ S_r の最大化

最適な組合せ比率を決定すべく、式(12)のシャープレシオ S_r を最大化する。この際、 $\tilde{r}_p > 0$ であれば買取引を行い、 $\tilde{r}_p < 0$ であれば空売りすることで \tilde{r}_p の符号を反転させ、再度 S_r を最大化する。

手順 5 売買の注文（仕掛け）

手順 4 で最適化した組合せ比率で売買の注文を行う。当日の始値で指値注文を出し、現在所有する現金全てを投資資金とする。指値注文は、始値が決定した直後に行うため、シミュレーションでは毎回約定できると仮定した。

手順 6 利益の確定（手仕舞い）

翌営業日の取引市場が開始する際に、成行注文で手順 5 の反対売買を行い、翌営業日の始値で全ポジショ

表 2 全 200 銘柄に対する投資シミュレーション結果. それぞれの予測法における、資産の増幅率 λ に関する統計量を示す. なお、1 銘柄保持とは、技術介入が全くない状態であり、予測の効果を測るために基準となる. 元本割れ率とは、 $\lambda < 1$ となった割合である. 太字の数値は、各項目の最良値を示す.

Table 2 Results of investment simulation by all pairs of $200C_2$ portfolios. We show the asset growth rate λ of each prediction, and also show that of just holding each stock, which is no techniques and is used as the base score. The bottom percentages show the frequency that the final asset m_e was below the initial asset m_s , that is, $\lambda < 1$. Each bold type means the best score of each category.

	1 銘柄保持	移動平均予測	線形予測	非線形予測
$\bar{\lambda}$	1.203	0.815	5.939	6.220
$\hat{\lambda}$	1.083	0.572	3.064	3.186
σ_λ	0.690	0.841	9.253	9.769
$\max\{\lambda\}$	4.591	12.606	281.281	309.104
$\min\{\lambda\}$	0.176	0.025	0.069	0.067
VaR	0.303	0.172	0.522	0.547
元本割れ率	42.3 [%]	76.5 [%]	15.6 [%]	14.5 [%]

ンを手仕舞う（全資産を回収する）．もし期待収益率の予測が正しければ、 $|\tilde{r}_p|$ に応じた収益を得ることができる．更に投資を続ける場合は、手順 1 に戻る．

手順 1~6 を繰り返すことで 5 年間の投資活動を行った後、初期資産の増幅率 λ を以下のように計算する．

$$\lambda = \frac{m_e}{m_s} \quad (31)$$

ここで、 m_s は初期資産高、 m_e は最終資産高を表す． $\lambda > 1$ であるほど、投資戦略としての効果が高い．一方、 $\lambda < 1$ であれば損失となるので、そのような場合を元本割れと呼ぶ．

その結果を表 2 に示す．もし市場価格が上昇傾向であれば、単純に銘柄を保持するだけでも利益を得られる．このような技術を伴わない投資と比較するために、各 200 銘柄を単純に保持し続けたときに得られる λ も算出した．また、この λ は図 1 に示すように右裾が長い分布である．このような場合、パフォーマンスの指標として中央値が望ましいとされるため[22]、 λ に関する中央値 $\hat{\lambda}$ も算出した．

移動平均予測を行った場合は、単純に保持した場合の投資成績と比べて特段の差は見受けられない．しかし線形予測や非線形予測を適用すれば、資産の増幅率 λ の向上は明白である．特に $\max\{\lambda\}$ は飛躍的に上昇できている．一方、ポートフォリオを構築して得られた λ は右裾が長いため、標準偏差 σ_λ が大きい．その

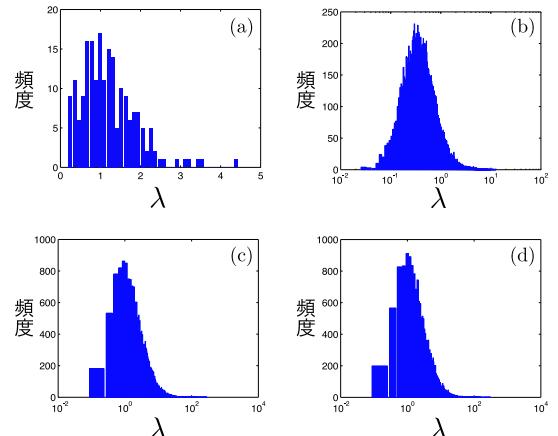


図 1 全 200 銘柄に対する資産の増幅率 λ の頻度分布：
(a) 1 銘柄保持、(b) 移動平均予測、(c) 線形予測、
(d) 非線形予測の場合

Fig. 1 Histograms of the asset growth rate λ given by (a) holding each stock, (b) the moving average prediction, (c) the AR prediction, or (d) the NAR prediction.

ため 1 銘柄保持が低リスクのように見える．しかし元本割れ率や信頼区間 95% で計算した VaR を比較すると、線形予測や非線形予測は 1 銘柄保持の結果を改善しており、損をする可能性が低い．つまり、本研究のように時系列予測モデルをポートフォリオ理論に組込むことで、高収益かつ安全に資産運用を行うことができる．

6. 投資銘柄の厳選

実際の投資では前章の投資シミュレーションと異なり、全 $200C_N$ 通りのポートフォリオを構築する必要はなく、使用する銘柄を選択することができる．その選択基準として、予測精度 ξ_i が高い銘柄を選ぶのが得策であるが、投資時において $\{\xi_i\}$ は未知である．しかし、もし予測モデルのゆう度 $\{\eta_i\}$ と予測精度 $\{\xi_i\}$ に相関関係が存在すれば、既知情報である η_i によってある程度 ξ_i を推測することができる．

この可能性を検証すべく、図 2 に全 200 銘柄の $\{\eta_i\}$ と $\{\xi_i\}$ の相関図を示す．投資開始日（2000 年 4 月 7 日）から過去 $L = 1250$ 点分の学習データに対するゆう度を η_i 、それ以降の 1250 営業日の評価データに対する予測精度を ξ_i としている．確かに、移動平均予測以外であれば、ある程度の相関関係を有することを確認できる．一方、移動平均予測ではそもそもゆう度 $\{\eta_i\}$ が低く、学習データの特徴を抽出できていない．

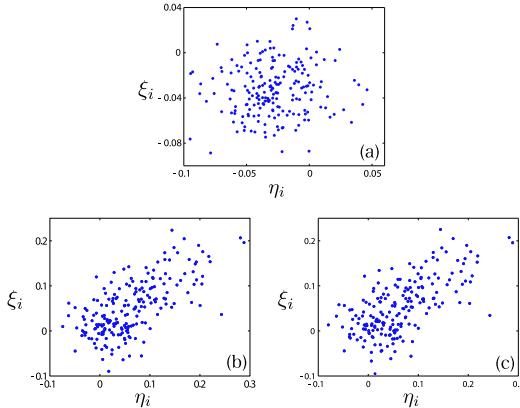


図 2 予測モデルのゆう度 $\{\eta_i\}$ と予測精度 $\{\xi_i\}$ の相関図：(a) 移動平均予測, (b) 線形予測, (c) 非線形予測の場合。なお、それぞれの相関係数は (a) 0.120, (b) 0.661, (c) 0.660 である。

Fig. 2 Correlation diagrams between the likelihood $\{\eta_i\}$ and the prediction accuracy $\{\xi_i\}$ given by (a) the moving average prediction, (b) the AR prediction, and (c) the NAR prediction. Each correlation is (a) 0.120, (b) 0.661 or (c) 0.660.

それゆえ予測精度 $\{\xi_i\}$ も低く、 $\{\eta_i\}$ と $\{\xi_i\}$ は無相関に近い。

次に、実際にゆう度 η_i が高い上位 n 銘柄に厳選し、予測精度の平均値 $\bar{\xi}$ と 2 銘柄ポートフォリオによる資産増幅率の平均値 $\bar{\lambda}$ を算出した。なお、 $\bar{\xi}$ は n 銘柄による平均値であり、 $\bar{\lambda}$ は nC_2 通りのポートフォリオによる平均値である。更に、各予測モデルにおいて $\{\eta_i\}$ が異なるので、選出される銘柄は各予測モデルで異なる。これらの結果を図 3 に示す。移動平均予測ではゆう度 $\{\eta_i\}$ と予測精度 $\{\xi_i\}$ の相関が低いことから、ゆう度に基づいて投資銘柄を厳選することは有効ではない。しかし移動平均予測以外では、ゆう度に基づいて投資銘柄を厳選することで予測精度 $\bar{\xi}$ の向上を確認できる。それに伴い、資産増幅率 $\bar{\lambda}$ も向上している。

図 3(a)において、線形予測と非線形予測の予測精度 $\bar{\xi}$ の差はわずかであるが、図 3(b)の資産増幅率 $\bar{\lambda}$ ではこの差が拡大している。つまり予測精度の向上が微小であっても、または表 1 のように予測精度 자체が微小であっても、投資成績に与える恩恵は大きいといえる。更に線形予測と非線形予測の $\bar{\lambda}$ において、上に凸のピークが存在する。これは、銘柄数 n が小さいほど予測精度が高い優秀な銘柄に厳選できるが、しかし各銘柄の予測誤差パターンが類似した場合、ポートフォリオによるリスク分散効果が低減するからである。

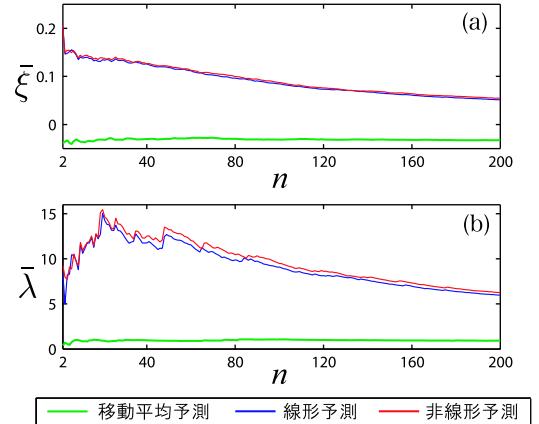


図 3 (a) 各予測モデルによる予測精度の平均値 $\bar{\xi}$ 、(b) 投資シミュレーションで得られた資産増幅率の平均値 $\bar{\lambda}$ 。ただし、予測モデルのゆう度 η_i に基づいて上位 n 銘柄に厳選した場合

Fig. 3 Results of the mean value of (a) prediction accuracy $\bar{\xi}$ and (b) the asset growth rate $\bar{\lambda}$, selecting the best n stocks based on the likelihood of each prediction model.

実際に n が小さいときは、特定の業界（主に、鉄道関連企業や製菓会社）が集中的に選ばれた。しかし n を増やすことで予測誤差パターンに多様性が生まれ、リスクを上手く分散できる 2 銘柄ペアが出現するため、 $n = 30$ 付近まで $\bar{\lambda}$ は上昇する。しかし n を増やし過ぎれば予測精度が低い銘柄までもが混入するため、 $\bar{\lambda}$ は低下する。

7. 多銘柄ポートフォリオの投資シミュレーション

前章の結果より、期待収益率の推定精度の向上が投資成績に大きく影響していることを確認できた。しかし 2 銘柄よりも多銘柄でポートフォリオを構築することで、より高いリスク分散効果を期待できる。この効果を検証すべく、投資シミュレーションを行った。なおポートフォリオを構築する銘柄は、6. でゆう度 η_i と予測精度 ξ_i に相関関係を確認できたため、ゆう度 η_i が高い銘柄から優先的に選択した。

投資シミュレーションの結果を図 4 に示す。ここで λ_p が、 n 銘柄ポートフォリオで得られた資産増幅率である。更にポートフォリオ効果を検証すべく、ポートフォリオを組まない場合と比較した。つまり、ポートフォリオに組み入れた n 銘柄をそれぞれ単独に投資し、資産増幅率の平均値 $\bar{\lambda}$ と中央値 $\hat{\lambda}$ も同時に示し

た。この単銘柄投資では、期待収益率 $\tilde{\lambda}_i(t+1)$ が正なら所持金分を買い建て、負なら所持金と同額分を売り建てた。更に時刻が $t+1$ になったときに、反対売買で決済した。

結果を見ると、それぞれの λ_p において上に凸のピークが存在している。これは図 3(b) と同様に、銘柄数 n の増加によりリスク分散効果の恩恵を得られるが、しかし n が多すぎると予測精度が低い銘柄までもが混入するため、ポートフォリオのパフォーマンスは低下する。したがって、そもそも予測精度が低い移動平均予測では、図 4(a) に示すように λ_p の値は低い。また先述のように、ポートフォリオを組む場合は式(12)中の $\tilde{\lambda}_p$ に基づいて全銘柄を同じポジションにするが、単銘柄投資の場合は個別に売り買いを決定できるため柔軟性がある。よって図 4(a) では、単銘柄投資の平均資産増幅率 $\bar{\lambda}$ の方がポートフォリオの資産増幅率 λ_p よりも高い傾向になったと考えられる。しかしポートフォリオにこのような制約があるとしても、図 4(b) (c) のように高次の予測モデルを用いることで、単銘柄投資と比較しても驚異的に資産を増幅させることができた。これはゆう度が高い多銘柄を同時にポートフォリオに組み込むことで高い予測精度を保ちつつ、かつリスク分散効果を十分に発揮できたことを示して

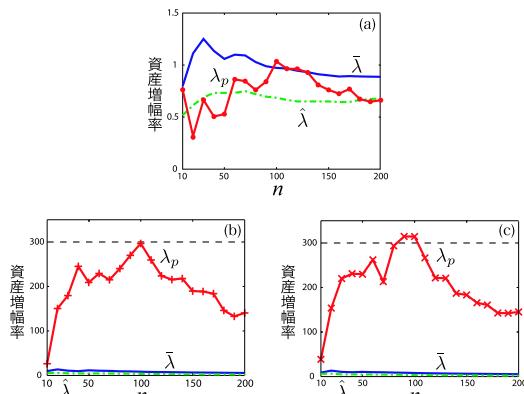


図 4 n 銘柄ポートフォリオから得られた資産の増幅率 λ_p と、単銘柄投資の資産増幅率の平均値 $\bar{\lambda}$ 及び中央値 $\hat{\lambda}$ 。それぞれ、(a) 移動平均予測、(b) 線形予測、(c) 非線形予測の場合

Fig. 4 Results of the asset growth rate given by making a portfolio with n stocks λ_p , the mean value of $\{\lambda_i\}$ given by each single stock investment $\bar{\lambda}$, and the median value of them $\hat{\lambda}$. Each figure is based on (a) the moving average prediction, (b) the AR prediction, and (c) the NAR prediction.

いる。特に非線形予測を用いると、初期資産を最大約 300 倍（年利約 213%）まで増やすことができた。しかし銘柄数 n を増やしすぎると、全体の予測精度は低下するため利益も低下する。したがって銘柄数 n には、リスク分散効果の向上と予測精度の低下におけるトレードオフが存在するが、この n を最適化する方法は今後の検討課題である。

8. 週次・月次データを用いた投資シミュレーション

前章までは日次データを対象としてきたが、従来よりポートフォリオ理論は主に月次などの時間スケールが大きいデータに対して用いられている。そこで本章では、2 週次データや月次データであっても本提案手法が有効であるか、予測精度と運用成績の観点から検証する。

まず、予測精度 $\{\xi_i\}$ の統計を表 3、表 4 に示す。表 1 と比較して、収益率の時間スケールが拡大するほど、どの予測法においても予測精度が低下している。特に予測精度の平均値 $\bar{\xi}$ において、移動平均予測が最も低い。これは 1. で述べたように、時間スケールの拡大に伴い収益率変動が有する諸特性が減弱し、ランダムウォークに近づいたためだと考えられる[4]～[6]。更に、特に金融データのように不等時間間隔で変動するジャンプ過程の場合、観測する時間スケールの

表 3 各予測法による 2 週次データ（200 銘柄）の予測精度 $\{\xi_i\}$ 。なお、 $\bar{\xi}$ は平均値、 σ_ξ は標準偏差である。

Table 3 Results of prediction accuracy $\{\xi_i\}$ using two weekly data for each prediction model. The index $\bar{\xi}$ means the mean value of $\{\xi_i\}$, and σ_ξ means the standard deviation of them.

	移動平均予測	線形予測	非線形予測
$\bar{\xi}$	-0.101	-0.011	-0.015
σ_ξ	0.070	0.098	0.100
$\max \{\xi_i\}$	0.108	0.218	0.205
$\min \{\xi_i\}$	-0.101	-0.252	-0.350

表 4 表 3 と同様。ただし、月次データ（200 銘柄）の予測精度 $\{\xi_i\}$

Table 4 The same as Table 3, but using monthly data.

	移動平均予測	線形予測	非線形予測
$\bar{\xi}$	-0.127	-0.014	-0.015
σ_ξ	0.105	0.133	0.132
$\max \{\xi_i\}$	0.161	0.328	0.348
$\min \{\xi_i\}$	-0.127	-0.370	-0.348

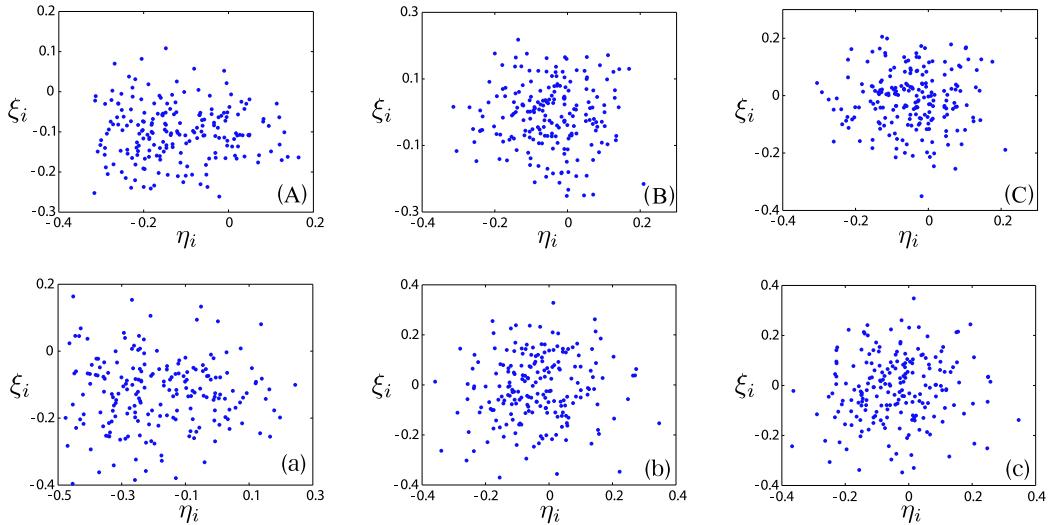


図 5 予測モデルのゆ度 $\{\eta_i\}$ と予測精度 $\{\xi_i\}$ の相関図：上図の (A)～(C) が 2 週次データ、下図の (a)～(c) が月次データの結果である。なお、(A) (a) は移動平均予測、(B) (b) は線形予測、(C) (c) は非線形予測の場合。それぞれの相関係数は、(A) -0.011 、(a) -0.006 、(B) -0.006 、(b) 0.086 、(C) 0.002 、(c) 0.100 である。

Fig. 5 Correlation diagrams between the likelihood $\{\eta_i\}$ and the prediction accuracy $\{\xi_i\}$ given by (A) (a) the moving average prediction, (B) (b) the AR prediction, and (C) (c) the NAR prediction. The upper figures (A)～(C) are based on two weekly data, and the bottom figures (a)～(c) are based on monthly data. Each correlation is (A) -0.011 , (a) -0.006 , (B) -0.006 , (b) 0.086 , (C) 0.002 , and (c) 0.100 .

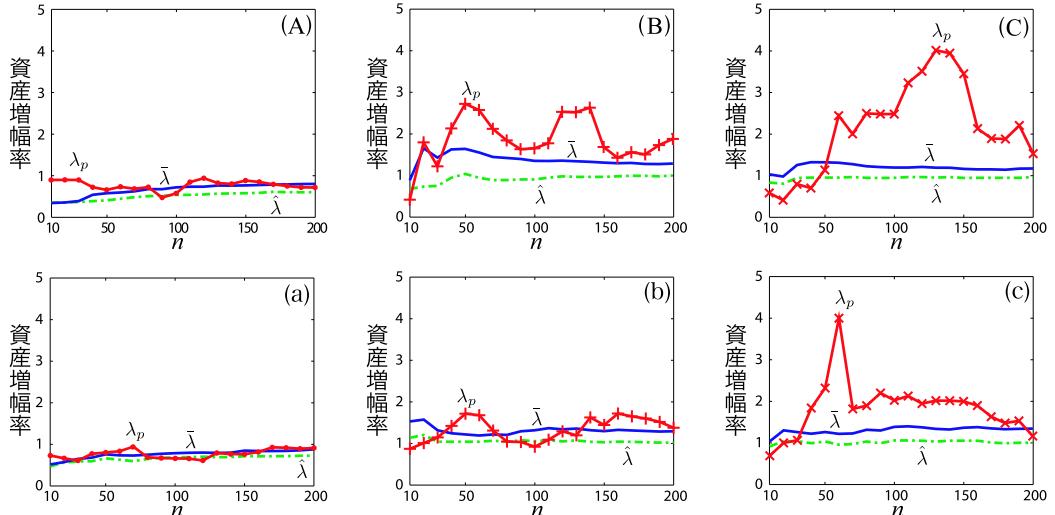


図 6 図 4 と同様。ただし、 n 銘柄ポートフォリオから得られた資産の増幅率 λ_p と、単銘柄投資の資産増幅率の平均値 $\bar{\lambda}$ 及び中央値 $\hat{\lambda}$ を示す。

Fig. 6 The same as Fig. 5, but showing the asset growth rate given by making a portfolio with n stocks λ_p , the mean value of $\{\lambda_i\}$ given by each single stock investment $\bar{\lambda}$, and the median value of them $\hat{\lambda}$.

拡大によって時系列予測が困難になることが示されている[23], [24]。これは、たとえシステムにダイナミクスが内在するとしても、サンプリング間隔の拡大によって元の力学的構造が破壊されるためである。

次に、2週次データ及び月次データ（全200銘柄）の $\{\eta_i\}$ と $\{\xi_i\}$ の相関図を図5に示す。先の結果のようにいずれの予測モデルでも予測困難であるため、学習データの特徴を抽出できずに $\{\eta_i\}$ と $\{\xi_i\}$ はほぼ無相関になっている。つまり、6.のようにゆう度に基づいた銘柄の選択は難しい。

最後に、多銘柄ポートフォリオを構築した場合の投資シミュレーション結果を図6に結果を示す。日次データを用いた図4と比較して、資産増幅率は圧倒的に低下している。これは予測精度の低下が根本的な原因である。ただし移動平均予測においては、日次データでも予測精度は低いので、大きな変化は見受けられない。

9. 他期間のデータを用いた有用性の検証

これまでの検証により、提案モデルは特に日次取引において有用性が高いことを確認した。しかしこれらの結果は1996年から2005年までのデータを用いた検証である。この期間はリーマン・ショック（2008年9月15日）以前であり、株価は比較的安定して変動していた。そこで本章では提案法の有用性をより厳しい状況下で評価すべく、リーマン・ショック後の日次収益率データ[25]に対して、これまでと同様の手法により検証する。

本章では2009年12月30日以前の約5年間（土日祝日を除く5年分の営業日）を学習データとし、それ以降の3年間において投資を続行する。まず、予測精度 $\{\xi_i\}$ の統計を表5に示す。なお各予測モデルのモデルパラメータは交差確認法より、移動平均予測では $T=96$ 、線形予測では $T=2$ 、非線形予測では $T=3$ 及び $\tau=4$ と設定された。表1と比較してどの予測法においても予測精度 $\bar{\xi}$ は若干低下している。これは推定されたモデルパラメータが変化したことからも、リーマン・ショック前後において収益率変動を生み出す力学的構造が変化したためと考えられる。しかし移動平均予測に対するウィルコクソンの符号付順位と検定の結果は、線形予測は $Z=6.606$ 、非線形予測は $Z=8.789$ であった。つまり信頼区間95%における棄却域は $Z \geq 1.645$ であることから、移動平均予測との有意差は明白である。更に、線形予測に対する非線形

表5 表1と同様。ただし、リーマン・ショック後の予測精度 $\{\xi_i\}$

Table 5 The same as Table 1, but using the data after the Lehman shock.

	移動平均予測	線形予測	非線形予測
$\bar{\xi}$	-0.052	0.016	0.038
σ_{ξ}	0.040	0.072	0.069
$\max\{\xi_i\}$	0.040	0.255	0.280
$\min\{\xi_i\}$	-0.208	-0.144	-0.103

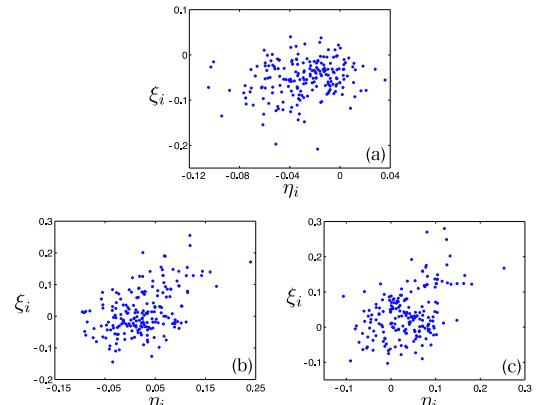


図7 図2と同様。ただし、リーマン・ショック後のデータに対する予測モデルのゆう度 $\{\eta_i\}$ と予測精度 $\{\xi_i\}$ 。なお、それぞれの相関係数は(a) 0.188, (b) 0.474, (c) 0.419である。

Fig. 7 The same as Fig. 2, but using the data after the Lehman shock. Each correlation is (a) 0.188, (b) 0.474, or (c) 0.419.

予測の検定結果は $Z=5.388$ であり、非線形予測の優位性も十分に確認することができた。

次に、リーマン・ショック後のデータの $\{\eta_i\}$ と $\{\xi_i\}$ の相関図を図7に示す。こちらも同様に相関は低下しているが、移動平均予測と比較すれば線形予測及び非線形予測は優位であり、ゆう度に基づいた投資銘柄の厳選は有効であると考えられる。

最後に、多銘柄ポートフォリオを構築した場合の投資シミュレーション結果を図8に示す。非線形予測を用いると、初期資産を最大約12倍まで増やすことができる。この場合の年利は約129%であり、図3の最大約300倍（年利約213%）と比較すれば、資産増幅率は低下している。これは表5に示すように、予測精度の低下が直接的な原因である。しかし移動平均予測と比較すれば、線形予測及び非線形予測の有用性は明白である。

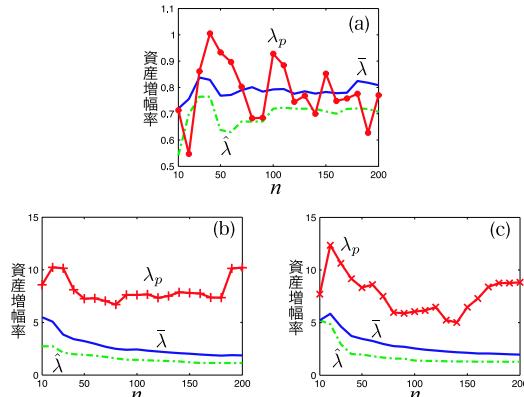


図 8 図 4 と同様、ただしリーマン・ショック後のデータを用いた場合

Fig. 8 The same as Fig. 4, but using the data after the Lehman shock.

10. む す び

マーコビツツの平均分散モデルでは、過去の収益率の移動平均値によって期待収益率の予測がなされていると解釈できる。そこで、より高次の時系列予測モデルを用いて収益率を予測したところ、予測精度を向上できたため、既存のポートフォリオ理論に時系列予測モデルを組む方法を検討した。移動平均予測以外の予測値を期待収益率とみなすには、リスクの評価式を変更する必要が生じたため、時系列予測が外れる可能性をリスクとした。この可能性は、過去の予測誤差に基づいて推定された。

このような変更に効果があるのかを検証すべく、実際の株価収益率データを用いて投資シミュレーションを行った。まずリーマン・ショック前の 1996 年から 2005 年までのデータを用いた。結果として、従来のポートフォリオ理論よりも高い投資成績が得られ、投資銘柄を厳選しない場合でも 5 年間で資産を約 6 倍（年利約 38%）に増やすことができた。更に初期資産を下回る確率を大幅に低減することができ、本手法によって低リスクかつ高収益な資産運用を行うことができる。

また、移動平均予測以外では、予測モデルのゆう度と予測精度に高い相関関係を確認できた。したがって、ゆう度を参照することで期待収益率を推定しやすい銘柄を選ぶことができ、銘柄の厳選によって更に投資成績を向上できた。更に、ゆう度が高い多銘柄を同時にポートフォリオに組むことで、高い予測精度かつ高

いリスク分散効果を発揮でき、利益を最大約 300 倍（年利約 213%）まで増やすことができた。しかし銘柄数を増やしすぎると、全体の予測精度は低下するため利益も低下した。この銘柄数の最適化が今後の課題である。

更に本手法の効果は、月次データなど時間スケールを拡大するとほとんど消滅した。これは時間スケールの拡大によって必然的に同一期間の学習データ数は減少するため、予測モデルの学習が困難になったと考えられる。更に、たとえ金融データに短時間スケールの特徴が存在したとしても必然的に全て破壊されるため、予測モデルの学習に利用できる情報が激減したと考えられる。一般的には長期的な展望でポートフォリオを組む場合が多いが、日次データの利点に着目することで時系列予測モデルを向上させる動機が生まれ、そして実際に投資成績を向上できることを示した点が本研究の特色である。

次にリーマン・ショック後の 2010 年から 2012 年までの期間では、予測精度の低下が見られた。しかしこのような場合においても、移動平均予測に対する線形予測及び非線形予測の優位性は高く、特に非線形予測は比較的優れた予測精度を維持できた。更に投資シミュレーションを行った結果、リーマン・ショック後であっても、利益を最大約 12 倍（年利約 129%）まで増やすことができた。

最後に今後の展望を述べる。証券会社に株式の委託売買を依頼した場合、取引額に応じて手数料が発生するが、本研究ではこの手数料を考慮していない。また、信用取引（空売り）の際に仲介業者に預ける証拠金についても考慮していない。よって本研究で得られた運用成績は、証券会社自身がディーラー業務として株式売買を行った場合を想定している。あるいは、1 回の取引額が 10 万円以下であれば委託手数料を必要としない証券会社が増えつつあるので、こういった証券会社を利用すれば、本研究で得られた運用成績は個人投資家にも当てはまる。しかし今後はこういった制約を取り除くために、手数料を陽に考慮したポートフォリオ選択法を検討する。特に本研究のように、短時間スケールのポートフォリオを組むほど取引回数が増加するので、手数料の影響が拡大する。上述のように利益を約 300 倍まで拡大できたが、手数料を考慮することでより現実的な結果を導けるだろう。それにはまずシャープレシオの改良やそれに基づいた投資戦略の再検討を行い、運用成績を再評価する必要がある。これ

らの課題については、誌面を改めて報告する予定である。

文 献

- [1] H.M. Markowitz, "Portfolio Selection," *J. Finance*, vol.7, no.1, pp.77–91, 1952.
- [2] 日本証券アナリスト協会編, 新・証券投資論 I&II, 日本経済新聞出版社, 2009.
- [3] E.F. Fama, "Efficient capital markets: A review of theory and empirical work," *J. Finance*, pp.383–417, 1970.
- [4] 刈屋武昭, 佃 良彦, 丸 淳子, 日本の株価変動, 東洋経済新報社, 1989.
- [5] G. Hawawini and D.B. Keim, "On the predictability of common stock returns: World-wide evidence review article," in *Handbooks in Operations Research and Management Science*, vol.9 ed. R.A. Jarrow, et al., pp.497–544, Elsevier B.V., 1995.
- [6] J.Y. Campbell, A.W. Lo, and A.C. Mackinlay, *The Econometrics of Financial Markets*, Princeton University Press, 1996.
- [7] 日本テクニカルアナリスト協会(編), 日本テクニカル分析大全, 日本経済新聞出版社, 2004.
- [8] 沖本竜義, 経済・ファイナンスデータの計量時系列分析, 朝倉書店, 2010.
- [9] E.N. Lorenz, "Atmospheric predictability as revealed by naturally occurring analogues," *J. Atmospheric Sciences*, vol.26, pp.636–646, 1969.
- [10] J.D. Farmer and J.J. Sidorowich, "Predicting chaotic time series," *Phys. Rev. Lett.*, vol.59, pp.845–848, 1987.
- [11] 池口 徹, 山田泰司, 小室元政, カオス時系列解析の基礎と応用, 合原一幸(編), 産業図書, 2000.
- [12] 松葉育雄, 非線形時系列解析, 朝倉書店, 2007.
- [13] B. LeBaron, *Nonlinear Forecasts for the S&P Stock Index, Nonlinear Modeling and Forecasting*, SFI Studies in the Sciences of Complexity, Addison-Wesley, 1992.
- [14] W.F. Sharpe, "Capital asset prices: A theory of market equilibrium under conditions of risk," *J. Finance*, vol.19, no.3, pp.425–442, 1964.
- [15] 無担保コールレート翌日物データは東京短資株式会社から入手. <http://www.tokyotanshi.co.jp/past/index2.shtml> (参照日 2012.1.15).
- [16] F. Takens, "Detecting strange attractors in turbulence," *Lecture Notes in Mathematics*, vol.898, pp.366–381, Springer-Verlag, 1981.
- [17] T. Sauer, J.A. Yorke, and M. Casdagli, "Embedology," *J. Statistical Physics*, vol.65, no.34, pp.579–616, 1991.
- [18] D.M. Allen, Mean square error of prediction as a criterion for selecting variables, *Technometrics*, vol.13, no.3, pp.469–475, 1971.
- [19] T. Suzuki, Y. Ueoka, and H. Sato, "Estimating structure of multivariate systems with genetic algorithms for nonlinear prediction," *Phys. Rev. E*, vol.80, 066208, 2009.
- [20] 株価データはパンローリング株式会社から購入. <http://www.panrolling.com/pansoft/data/> (参照日 2011.7.30).
- [21] B.G. Malkiel, *A Random Walk Down Wall Street: The Time-Tested Strategy for Successful Investing*, W.W. Norton & Company, 2007.
- [22] R.C. Merton and P.A. Samuelson, "Fallacy of the long-normal approximation to optimal portfolio decision-making over many periods," *J. Financial Economics*, pp.67–94, 1974.
- [23] T. Suzuki, "Appropriate time scales for nonlinear analyses of deterministic jump systems," *Phys. Rev. E*, vol.83, no.6, 066203, 2011.
- [24] 大塚陽介, 鈴木智也, "決定論的ジャンプ過程のシステム同定と長期予測に適したサンプリング手法の検討," 情処学論 数理モデル化と応用, vol.5, no.1, pp.30–39, 2012.
- [25] 株価データは Yahoo! ファイナンスから入手. <http://finance.yahoo.co.jp> (参照日 2013.1.10).

(平成 24 年 10 月 12 日受付, 25 年 2 月 4 日再受付)

猪瀬 悟史



平 22 茨城大・工・知能システム卒. 同年 4 月より同大大学院理工学研究科知能システム専攻博士前期課程に進学. 非線形時系列解析や数理ファイナンスに関する研究に従事.

鈴木 智也 (正員)



平 17 東京理科大学大学院理学研究科物理学専攻博士課程了. 理博. 同年東京電機大学工学部電子工学科助手, 平 18 より同志社大学工学部情報システムデザイン学科専任講師を経て, 平 21 より茨城大学工学部知能システム工学科准教授, 現在に至る. 非線形時系列解析, 複雑系, 金融工学に関する研究に従事. 電子情報通信学会, 情報処理学会, 日本物理学会, 日本テクニカルアナリスト協会各会員.