人工知能学会研究会資料
SIG-FIN-012-02

収益率分布予測に時系列モデルを用いたポートフォリオ最適化

Portfolio Optimization Using Time Series Models for Predicting the Distribution of Returns

吉住遼　水野眞治　高野祐一　西山昇
Ryo Yoshizumi　Shinji Mizuno　Yuichi Takano　Noboru Nishiyama

東京工業大学大学院社会理工学研究科
Graduate School of Decision Science and Technology, Tokyo Institute of Technology

Abstract: In recent years, portfolio optimization with the use of a vector autoregressive (VAR) model to estimate future stock returns has been subject of research. In this paper, we propose an optimization model that uses time series models to predict not only the future returns but also the conditional variance-covariance matrix of returns. More specifically, the future returns are predicted by using the VAR model, and the conditional variance-covariance matrix is estimated by using a dynamic conditional correlation (DCC) multivariate GARCH model. We evaluate the out-of-sample investment performance of our model using historical data of U.S. stock market.

1 はじめに

株式や債券などの資産価格は時間の推移とともに変動するため，将来の価格は不確実性を持つ．金融では一般にこのような不確実性をリスクととらえる．複数の資産の組み合わせのことをポートフォリオといい，資産運用を行う際，不確実性の低い資産に投資したり，相関の小さい複数の資産に分散投資することによってポートフォリオ全体としてリスクを抑えることができる．複数の投資対象への資産配分を決定する問題をポートフォリオ選択問題という．

Markowitz [5] は，ポートフォリオの収益率の分散をリスクとし，ポートフォリオの収益率の期待値を一定以上とする制約のもと，収益率の分散を最小化するように投資比率を決定する平均・分散モデルを提案した．ポートフォリオの期待収益率と分散を求める方法としては，過去に観測された収益率の時系列データの標本平均，標本分散が用いられる場合が多い．しかし経済環境は絶えず変化するため，過去の実現値による標本平均，標本分散を将来の投資決定のために利用することは適切ではないと考えられる．よって，本研究では，将来の収益率の分布を予測し，その予測に基づいて資産配分を決定するポートフォリオ最適化に着目する．

近年の研究として，DeMiguel et al. [2] は，一期間ベクトル自己回帰 (VAR(1)) モデルによる収益率予測を用いたポートフォリオ最適化モデルを提案した．このときポートフォリオのリスクは条件付き分散とし，VAR(1) モデルから予測した条件付き期待収益率と実測値との残差の標本分散を用いている．

しかし，一般に収益率の分散は時間の推移とともに変動し自己相関を持つことが知られている (ボラティリティ・クラスタリング)．よって，ポートフォリオのリスク指標とする条件付き分散についても，時系列モデルを用いて予測を行うことによって，ポートフォリオの収益率を高めることができる可能性がある．

分散が変動すると仮定して，分散を予測する時系列モデルは，収益率予測とは別に様々な研究がおこなわれている．Engle は過去の 2 乗誤差で分散を予測する ARCH モデル [3] を提案し，Bollerslev は ARCH モデルを拡張し，過去の 2 乗誤差と過去の分散によって分散を予測する GARCH モデル [1] を提案した．これらのモデルは一つの資産の分散のみを考えるモデルであるが，ポートフォリオの分散を考える際は資産間の共分散の予測も必要である．共分散を予測するモデルとして Engle は GARCH モデルで各資産の分散を予測し，資産間の相関についても GARCH モデルと同様の考え方でモデル化することで，分散と相関から分散共分散行列を予測する DCC-GARCH モデル [4] を提案した．

本研究では，VAR モデルで収益率を予測し，条件付き分散共分散行列を DCC-GARCH モデルで予測するポートフォリオ最適化モデルを提案する．また米国株式市場の株価データを用いて，運用のシミュレーションを行い，運用結果から提案モデルの有効性を検証する．

2　ポートフォリオ最適化

複数の投資対象資産の中から投資家にとって最も好ましいように，どの投資対象にどれだけ投資をしたらよいかという資産配分を決定する問題をポートフォリオ選択問題という．投資対象資産 $i \in \mathcal{A}$ $(i = 1, \cdots, N)$ の収益率を確率変数とみなし，$\tilde{r}_i$ と書く．資産 i への投資比率を x_i とする．よって，ポートフォリオの収益率は $\sum_{i \in \mathcal{A}} \tilde{r}_i x_i$ となり，$\tilde{r}_i$ の期待値を $\mu_i = \mathrm{E}[\tilde{r}_i]$ とすると，ポートフォリオの期待収益率は

$$\mathrm{E}[\sum_{i \in \mathcal{A}} \tilde{r}_i x_i] = \sum_{i \in \mathcal{A}} \mathrm{E}[\tilde{r}_i] x_i = \sum_{i \in \mathcal{A}} \mu_i x_i \tag{1}$$

である．一方，資産 i と j の収益率の共分散を $\sigma_{i,j} = \mathrm{E}[(\tilde{r}_i - \mu_i)(\tilde{r}_j - \mu_j)]$ とすると，ポートフォリオの収益率の分散は

$$\begin{aligned}
&\mathrm{E}[(\sum_{i \in \mathcal{A}} \tilde{r}_i x_i - \sum_{i \in \mathcal{A}} \mu_i x_i)^2] \\
&= \sum_{i \in \mathcal{A}} \sum_{j \in \mathcal{A}} \mathrm{E}[(\tilde{r}_i - \mu_i)(\tilde{r}_j - \mu_j)] x_i x_j \\
&= \sum_{i \in \mathcal{A}} \sum_{j \in \mathcal{A}} \sigma_{i,j} x_i x_j
\end{aligned} \tag{2}$$

と表せる．Markowitz [5] は，ポートフォリオの収益率の分散をリスクととらえ，ポートフォリオの収益率の期待値を一定以上とする制約のもと，収益率の分散を最小化する平均・分散モデルを提案した．本研究では，リスク回避度 γ $(0 \leq \gamma \leq 1)$ を用いて分散と期待値の加重和を最小化する問題を解く．すなわち，目的関数を

$$\text{minimize } \gamma \sum_{i \in \mathcal{A}} \sum_{j \in \mathcal{A}} \sigma_{i,j} x_i x_j - (1 - \gamma) \sum_{i \in \mathcal{A}} \mu_i x_i \tag{3}$$

とし，制約条件として，投資比率の和が 1 になる制約と空売り禁止制約を設ける．

$$\text{subject to} \qquad \sum_{i \in \mathcal{A}} x_i = 1 \tag{4}$$

$$x_i \geq 0 \ (\forall i \in \mathcal{A}) \tag{5}$$

ポートフォリオの収益率の期待値と分散には，過去に観測された収益率の時系列データの標本平均と標本分散が用いられる場合が多い．過去の $t = 1$ 時点から $t = T - 1$ 時点までの $T - 1$ 期間に観測された資産 i の収益率の実測値 $r_{i,t}$ とすると，標本平均を用いた資産 i の期待収益率 μ_i は

$$\mu_i = \frac{1}{T-1} \sum_{t=1}^{T-1} r_{i,t} \tag{6}$$

であり，標本共分散 $\sigma_{i,j}$ は以下のように表せる．

$$\sigma_{i,j} = \frac{1}{T-1} \sum_{t=1}^{T-1} (r_{i,t} - \mu_i)(r_{j,t} - \mu_j) \tag{7}$$

3　VAR モデルによる収益率予測

DeMiguel et al. [2] はポートフォリオの運用成績を改善するために，1 期間ベクトル自己回帰モデル (VAR(1) モデル) を用いて各資産の収益率を予測した．

時点 1 から時点 $T - 1$ までの各時点 t の収益率 $\tilde{r}_{i,t}$ の実現値 $r_{i,t}$ が与えられているとする．VAR(1) モデルでは資産 i の時点 t の収益率 $\tilde{r}_{i,t}$ は 1 期前のすべての資産の収益率の影響を受けると考える．時点 t の資産 i の収益率の撹乱項を $\tilde{u}_{i,t}$ とし，平均 0，ある分散共分散行列を持ち，任意の i, j, t, s (ただし，$t \neq s$) に対して $\tilde{u}_{i,t}$ と $\tilde{u}_{j,s}$ は互いに独立と仮定する．パラメータを $c_i, \phi_{i,k}$ とすると VAR(1) モデルは以下のように表される．

$$\tilde{r}_{i,t} = c_i + \sum_{k=1}^{N} \phi_{i,k} \tilde{r}_{k,t-1} + \tilde{u}_{i,t} \tag{8}$$

パラメータは最小 2 乗法によって求めることができる．

$t - 1$ 時点までの確率変数の実現値が与えられたときの $\tilde{r}_{i,t}$ の期待値を条件付き期待値という．条件付き期待値は求めたパラメータ $c_i, \phi_{i,k}$ を用いて以下のように表せる．

$$\mathrm{E}_{t-1}[\tilde{r}_{i,t}] = c_i + \sum_{k=1}^{N} \phi_{i,k} r_{k,t-1} \tag{9}$$

条件付き期待値を用いて，時点 T の予想収益率を $\hat{r}_{i,T} = \mathrm{E}_{T-1}[\tilde{r}_{i,T}]$ とする．一方，リスクはモデルから推定した予測値と実際の値との残差を用いた条件付き分散を用いる．

$$\begin{aligned}
&\sum_{i \in \mathcal{A}} \sum_{j \in \mathcal{A}} \mathrm{E}_{T-1}[(\tilde{r}_{i,T} - \hat{r}_{i,T})(\tilde{r}_{j,T} - \hat{r}_{j,T})] x_i x_j \\
&= \sum_{i \in \mathcal{A}} \sum_{j \in \mathcal{A}} \mathrm{E}_{T-1}[\tilde{u}_{i,T} \tilde{u}_{j,T}] x_i x_j
\end{aligned} \tag{10}$$

ここからは，$h_{i,j,t} = \mathrm{E}_{t-1}[\tilde{u}_{i,t} \tilde{u}_{j,t}]$ とおく．DeMiguel et al. [2] では時点 T の条件付き共分散 $h_{i,j,T}$ は予測残差 $u_{i,t}$ の標本共分散を用いて推定している．すなわち，

$$h_{i,j,T} = \frac{1}{T-1} \sum_{t=1}^{T-1} (r_{i,t} - \hat{r}_{i,t})(r_{j,t} - \hat{r}_{j,t}) \tag{11}$$

となる．よって，VAR(1) モデルによる収益率予測に基づいて資産配分を決定するポートフォリオ最適化モデルの目的関数は

$$\text{minimize } \gamma \sum_{i \in \mathcal{A}} h_{i,j,T} x_i x_j - (1 - \gamma) \sum_{i \in \mathcal{A}} \hat{r}_{i,T} x_i \tag{12}$$

となる．DeMiguel et al. [2] は VAR モデルで収益率を予測したポートフォリオ最適化モデルが，過去の収益率の標本平均と標本分散を用いたモデルよりも運用成果が改善することを示した．

4 DCC-GARCHモデルによるリスク予測

先行研究では，ポートフォリオのリスクは条件付き分散とし，モデルから予測した条件付き期待収益率と実測値との残差の標本分散を用いている．しかし，一般に収益率の分散は時間の推移とともに変動し自己相関を持つことが知られている．よって，条件付き分散についても，時系列モデルで予測を行うことによって，ポートフォリオの収益率を高めることができる可能性がある．そこで本研究では，条件付き分散の予測に DCC-GARCH モデル [4] を用いるポートフォリオ最適化モデルを提案する．

条件付き共分散 $h_{i,j,t}$ のモデル化を考えるために，まず，時点 t の資産 i の条件付き分散 $h_{i,i,t}$ のみを考える．資産価格の変動において，収益率の分散は一旦上がると高い水準がしばらく続き，低下すると低い水準が続く傾向（ボラティリティ・クラスタリング）が知られており，過去の変動の大きさの影響を受ける．また特に直近の時点の変動ほど与える影響が大きいと考えられる．そこで，1 期前の 2 乗誤差の実現値を $u_{i,t-1}^2$，予測値を $h_{i,i,t-1}$，パラメータを切片 ω_i (≥ 0)，係数 α_i，β_i，ただし，$0 < \alpha_i, \beta_i < 1$ とし，資産 i の条件付き分散を

$$h_{i,i,t} = \omega_i + \alpha_i u_{i,t-1}^2 + \beta_i h_{i,i,t-1} \tag{13}$$

とモデル化する．上式は過去 m 期間のデータを用いて次のように変形できる．

$$h_{i,i,t} = \omega_i \sum_{k=1}^{m} \beta_i^{k-1} + \alpha_i \sum_{k=1}^{m} \beta_i^{k-1} u_{i,t-k}^2 + \beta_i^m h_{i,i,t-m} \tag{14}$$

このモデルのパラメータは各資産間の相関がないと仮定し最尤法で推定する．

ポートフォリオの分散を考えるためには資産間の条件付き共分散も推定する必要がある．よって，ここでは条件付き相関をモデル化し，先に推定した条件付き分散と条件付き相関から共分散を計算する DCC モデルを利用する．

収益率の条件付き相関係数 $\rho_{i,j,t}$ は以下のように表せる．

$$\begin{aligned}\rho_{i,j,t} &= \frac{\mathrm{E}_{t-1}[\tilde{u}_{i,t}\tilde{u}_{j,t}]}{\sqrt{\mathrm{E}_{t-1}[\tilde{u}_{i,t}^2]\mathrm{E}_{t-1}[\tilde{u}_{j,t}^2]}} &(15)\\ &= \frac{h_{i,j,t}}{\sqrt{h_{i,i,t}h_{j,j,t}}} &(16)\end{aligned}$$

すなわち条件付き共分散は以下のようになる．

$$h_{i,j,t} = \rho_{i,j,t}\sqrt{h_{i,i,t}h_{j,j,t}} \tag{17}$$

ここで，$\sqrt{h_{i,i,t}}$ の対角行列を $D_t = \mathrm{diag}(\sqrt{h_{i,i,t}})$ とし，時点 t の相関係数 $\rho_{i,j,t}$ を要素とする相関行列を R_t とすると，時点 t の分散共分散行列 H_t は以下のように表せる．

$$H_t = D_t R_t D_t \tag{18}$$

条件付き相関係数も GARCH モデルと同様の考え方でモデル化する．ただし，すべての資産間の相関について異なるパラメータを考えると推定するパラメータ数が非常に多くなるので，相関係数の変動の構造は任意の 2 資産に対してすべてパラメータが同じと仮定し，標準化して考える．

標準化した残差を $\epsilon_{i,t} = u_{i,t}/\sqrt{h_{i,i,t}}$ とし，$q_{i,j,t} = \mathrm{E}_{t-1}[\tilde{\epsilon}_{i,t}\tilde{\epsilon}_{j,t}]$ とする．そして，パラメータ a, b を用いて，GARCH モデルと同様に以下のように表す．

$$q_{i,j,t} = (1-a-b)\bar{q}_{i,j} + a\epsilon_{i,t-1}\epsilon_{j,t-1} + bq_{i,j,t-1} \tag{19}$$

ただし，$a, b > 0,\ a + b < 1$ とする．ここで $\bar{q}_{i,j}$ は標本相関を用いている．パラメータ a と b は任意の i と j に対して同じとする．このようにするとパラメータ数は GARCH モデルで $3 \times N$ 個，DCC モデルで 2 個となる．このようにしたとき条件付き相関係数 $\rho_{i,j,t}$ は

$$\begin{aligned}\rho_{i,j,t} &= \frac{\mathrm{E}_{t-1}[\tilde{u}_{i,t}\tilde{u}_{j,t}]}{\sqrt{\mathrm{E}_{t-1}[\tilde{u}_{i,t}^2]\mathrm{E}_{t-1}[\tilde{u}_{j,t}^2]}} &(20)\\ &= \frac{\sqrt{h_{i,i,t}h_{j,j,t}}\mathrm{E}_{t-1}[\tilde{\epsilon}_{i,t}\tilde{\epsilon}_{j,t}]}{\sqrt{h_{i,i,t}\mathrm{E}_{t-1}[\tilde{\epsilon}_{i,t}^2]\cdot h_{j,j,t}\mathrm{E}_{t-1}[\tilde{\epsilon}_{j,t}^2]}} &(21)\\ &= \frac{q_{i,j,t}}{\sqrt{q_{i,i,t}q_{j,j,t}}} &(22)\end{aligned}$$

で計算される．(14) 式で求めた分散と (22) 式の相関係数から (18) 式によって共分散を求める．このようにして推定した $h_{i,j,T}$ を (12) 式の最適化モデルに用いる．

5 数値実験

本研究の提案モデルの有効性を検証するために運用のシミュレーションを行った．既存モデルとの比較をもとにどのような場合において提案モデルが有効か考察する．データは Kenneth R. French — Data Library [1] からダウンロードした 6 Portfolios Formed on Size and Book-to-Market [Daily] を用いた．これは図 1 に示すように，米国の上場企業株式を時価総額で 2 分割し (大きいほうを Big，小さいほうを Small とする)，PBR (純資産株価比率) の逆数で 3 分割して (高いほうから High，Middle，Low とする)，6 つのクラス (S-H，S-M，S-L，B-H，B-M，B-L) に分割し，各クラスに属する株式の日次収益率を時価総額で加重平均した指標である．

[1] http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/index.html

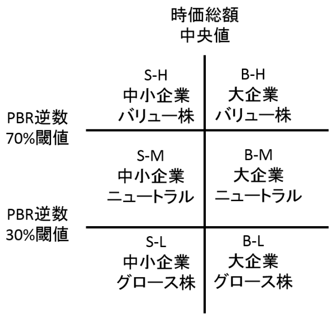

図 1: 米国株式インデックス

それぞれのモデルから求めた投資比率をもとに 250 日 (2012 年 11 月 5 日から 2013 年 10 月 31 日までの約 1 年) の運用を行って収益率の平均と分散を求める．その際，予測モデルのパラメータ，最適投資比率は毎日計算し直し，リバランスを行う．リスク回避度のパラメータ γ は $0,0.001,\cdots,0.009,0.01,0.02,\cdots,0.09,0.1,0.2,\cdots,0.9,1$ の 29 通りの値を設定する．横軸に 250 日間の収益率の分散，縦軸に収益率の平均をとり，γ ごとにプロットする．左上にプロットがあるほどよい運用成果といえる．

図 2〜4 では，古典的な平均分散モデル（mean-cov），VAR モデルによって収益率を予測する先行研究 [2] のモデル（VAR-cov），収益率と分散共分散行列の両方を予測する提案モデル（VAR-DCC）の運用成績を比較している．なお，図 2，3，4 はそれぞれ各時点で直前の 250 日，750 日，1000 日を学習期間とし，その期間の収益率データを用いて投資比率を決定している．

VAR-cov モデルと VAR-DCC モデルを比較すると，学習期間が，250 日，750 日，1000 日のいずれの場合でもリスク回避度が中ほどのとき提案モデルの方が運用成績が上回っている．両モデルとも収益率の予測方法は同じであるため，リスク回避度 γ が 0 に近いときは似たような投資になる．逆にリスク回避度が大きいときは VAR-DCC モデルは VAR-cov モデルに比べ劣っている．また，VAR-DCC モデルと mean-cov モデルとの比較では，学習期間 250 日ではリスク選好的なときに mean-cov モデルを上回るが，リスク回避的なときは mean-cov モデルの方がよい結果となる．1000 日という長期の学習期間を取ったときは mean-cov モデルの方が，VAR-DCC モデルよりも良い結果となった．リスク回避的なとき，mean-cov モデルと VAR-cov モデルは似たような投資になる．さらに，学習期間数の比較では，リスク選好的なときは学習期間 250 日の VAR-DCC モデルが最も運用成績がよく，リスク回避的なときは学習期間 1000 日の mean-cov モデルがよい結果となった．

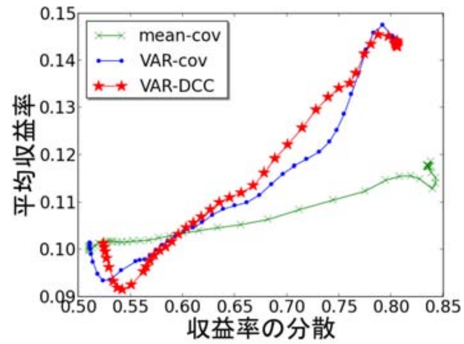

図 2: 運用成績の比較 (学習期間 250 日)

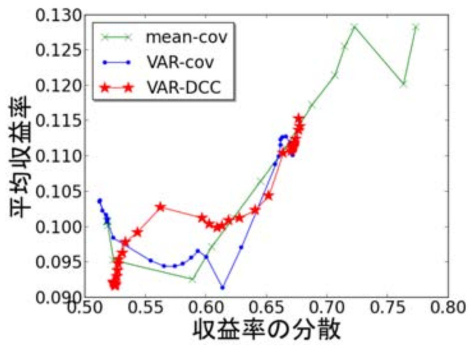

図 3: 運用成績の比較 (学習期間 750 日)

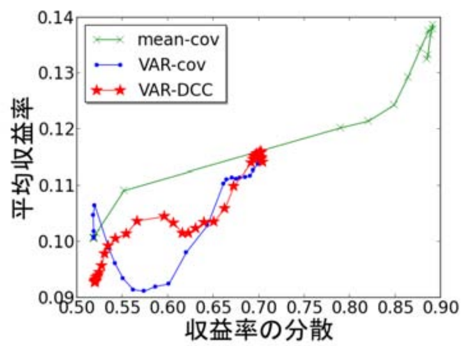

図 4: 運用成績の比較 (学習期間 1000 日)

6 まとめと今後

本研究では，VAR モデルで収益率を予測し，条件付き分散共分散行列を DCC-GARCH モデルで予測する最適化モデルを構築した．また，米国の株価インデックスを用いた運用シミュレーションで平均分散モデル，収益率を VAR モデルで予測する先行研究 [2] のモデル，収益率と分散共分散行列を共に予測する提案モデルを比較した結果，リスク回避度が中ほどのとき提案モデルが先行研究 [2] のモデルよりも良い運用結果を得られることがわかった．

今後はモデルのパラメータ推定の正確さを高める手法を考えるとともに，リスク予測をポートフォリオ最適化に取り入れることが運用成績の向上につながるか，さらなる実験を通して検証していく．

参考文献

[1] T. Bollerslev: Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, Vol. 31, pp. 307-327 (1986).

[2] V. DeMiguel, F.J. Nogales, R. Uppal: Stock return serial dependence and out-of-sample portfolio performance. to appear in *The Review of Financial Studies*.

[3] R. Engle: Autoregressive conditional heteroskedasticity with estimates of the variance of U.K. inflation. *Econometrica*, Vol. 50, pp. 987-1008 (1982).

[4] R. Engle: Dynamic conditional correlation: a simple class of multivariate generalized autoregressive conditional heteroskedasticity models. *Journal of Business and Economic Statistics*, Vol. 20, pp. 339-350 (2002).

[5] H.M. Markowitz: Portfollio selection. *Journal of Finance*, Vol. 7, pp. 77-91 (1952).