Vol.2012-MPS-88 No.21
2012/5/17

ベイジアンネットワークを用いた株式指数予測の改良

原田 昌朗[†1] 左 毅[†1] 北 栄輔[†1]

本研究では，ベイジアンネットワークを用いて将来株価を予測する方法について述べる．従来の方法では予測誤差を測定するために予測値と実測値の差を用いていたが，本研究では離散化された予測値と実測値の差を用いることで，アルゴリズムを簡単化する．また，以前の研究では過去 10 期までのデータを用いて株価を予測するベイジアンネットワークを構築しているのに対して，本研究では 5，10，15 期の 3 種類を比較する．解析例において．過去の株価収益率だけを用いる場合と，その予測誤差をも用いる方法を比較する．決定されたネットワークを比較すると，前者がナイーブベイズ構造となったのに対して，後者はより複雑なネットワーク構造となっている．そして，実データに対する相関係数と最小二乗誤差で比較したところ，後者は前者に対して精度向上が見られる．

Improvement of Stock Price Prediction by Using Bayesian Network

Masaaki Harada,[†1] Yi Zuo[†1] and Eisuke Kita[†1]

Application of Bayesian network to stock price prediction is described in this study. While, in the previous study, the prediction error is estimated by the difference between the actual and the predicted stock prices which are represented in the continuum values. In this study, the stock prices described in the discrete values are adopted for estimating the prediction error, which can simplify the algorithm. Besides, the effect of the number of the past stock prices for predicting the future stock price is also discussed. The prediction results by using the past stock price alone are compared with them by the past stock price and their prediction error in the numerical example. Bayesian network in the latter case is more complicated than that in the former case. The prediction accuracy in the latter case is better than that in the former case.

1. 緒　　論

株価指数の予測には事例列分析法が広く用いられている．代表的な時系列分析法には，Auto Regressive (AR) モデル，Moving Average (MA) モデル，AutoRegressive Moving Average (ARMA) モデル，AutoRegressive Conditional Heteroskedasticity(ARCH) モデル等がある．特に，ARCH モデルは 1980 年代初期に提案されて以来，株や金利などによる価格変化の分析・予測に大きな威力を発揮し，その後多数の改良型が提案されている．

これらのモデルでは，予測しようとする株価リターンを説明変数の線形結合で近似し，重み係数を相関解析などにより決定する．このとき，株価リターンの頻度分布は正規乱数に従うと仮定されている．しかし，近年のいくつかの研究により，株価変化は正規乱数に完全には従わないことが指摘されている．そこで，著者らの研究グループでは，ベイジアンネットワーク[1)] を用いた株価リターン推定法を提案している[2)]．ベイジアンネットワークを用いれば確率変数の条件付き因果関係をグラフネットワークとして表現できる．そこで，ベイジアンネットワークを用いて，過去の株価の条件付き因果関係を確率的にモデル化し，それを用いて将来株価を予測する．従来の方法では予測誤差を測定するために予測値と実測値の差を用いていたが，本研究では予測値と実測値の離散値の差を用いることで，アルゴリズムを簡単化する．また，以前の研究では過去 10 期までのデータを用いて株価を予測するベイジアンネットワークを構築しているのに対して，本研究では 5，10，15 期の 3 種類を比較する．

本論文の構成は以下のようになっている．2 節でベイジアンネットワークの理論を述べる．3 節で今回提案する手法を述べ，4 節で解析結果を示す．5 節は本論文の結論である．

2. ベイジアンネットワーク

2.1 ベイジアンネットワークと条件付き確率表

ベイジアンネットワークでは，確率変数間の定性的な依存関係を確率変数をノードとし，依存関係をリンクとする非循環有向グラフで可視化する．さらに，変数間の定量的な依存関係を変数間の条件付き確率によって表現する．

確率変数 x_i が x_j に依存していることを $x_j \rightarrow x_i$ と表現し，x_j を親ノード，x_i を子ノードと呼ぶ．x_i の親ノードが複数ある場合，その親ノード集合を $Pa(x_i)$ と表現することにす

†1 名古屋大学大学院情報科学研究科
Graduate School of Information Science, Nagoya University

る．子ノードの親ノードに対する依存関係は条件付き確率 $P(x_i|Pa(x_i))$ で表される．

ここで，親ノード集合 $Pa(x_i)$ が取り得る状態が M 個あり，そのひとつを記号 Y^m とする．一方，子ノード x_i の取り得る状態が L 個あり，そのひとつを X^l とする．このとき

$$\begin{array}{c} P(X^1|Y^1), P(X^2|Y^1), \cdots, P(X^L|Y^1) \\ \vdots \qquad\qquad \vdots \\ P(X^1|Y^M), P(X^2|Y^M), \cdots, P(X^L|Y^M) \end{array}$$

を表にしたものを条件付き確率表 (CPT) とよぶ．

2.2 グラフ構造の決定

本研究では，K2Metric[1),3)] をネットワークの評価値として採用し，K2 アルゴリズム[1)] を用いてネットワークのグラフ構造を決定する．

全ノード数を N，子ノード x_i が取りうる状態の総数を L，親ノード集合 $Pa(x_i)$ が取りうる状態の総数を M と表す．また，ノード x_i について，その親ノード集合 $Pa(x_i)$ が状態 Y^j を，ノード x_i が状態 X^k を取る場合の個数を N_{ijk} とする．事前分布が一様分布であるとすると，与えられたノード集合から構築されたネットワークについて，K2Metric[1),3),4)] は次式で与えられる．

$$K2 = \prod_{i=1}^{N}\prod_{j=1}^{M}\frac{(L-1)!}{(N_{ij}+L-1)!}\prod_{k=1}^{L}N_{ijk}! \tag{1}$$

ここで，N_{ij} と N_{ijk} には次の関係がある．

$$N_{ij} = \sum_{k=1}^{L}N_{ijk} \tag{2}$$

K2 アルゴリズム[1)] はよくばり法を元に作られており，全木探索より少ない計算量でネットワークを構築できる．このアルゴリズムでは，あらかじめ変数間の全順序関係がわかっている必要がある．本研究で扱う株価では，時系列に基づく全順序関係がある．K2 アルゴリズムは，その全順序制約をうまく用いて，ある変数の親変数の組合せを探索する．

探索アルゴリズムを以下に示す．

(1) ノード x_i に対する親ノード集合 $Pa(x_i)$ を空集合 ϕ として定義する．

(2) $i=1$

(3) x_i と $Pa(x_i)$ から構成されるネットワークについて K2Metric を評価し，これを S_{best} とする．

(4) $j=i+1,\cdots,N$ について以下の操作を行う．

  (a) x_j を $Pa(x_i)$ に加える．

  (b) x_i と $Pa(x_i)$ から構成されるネットワークについて K2Metric を評価し，これを S とする．

  (c) $S > S_{best}$ でないならば，x_j を $Pa(x_i)$ から除外する．

(5) $i=i+1$ とし，$i \leq N$ ならばステップ (3) へ戻る．

2.3 確率推論

ベイジアンネットワークの確率推論では，確率変数の確定値から，知りたい確率変数の事後確率を求め，これを用いて期待値などを計算する．

条件付き依存関係が成立しているネットワークにおいて，確率変数の確定値 e に対する知りたい確率変数 x_i の事後確率を $P(x_i|e)$ とすると，これは以下の手順で行われる[1)]．

(1) 観測された変数の値 e をノードにセットする．

(2) 知りたい確率変数 x_i の条件付確率 $P(x_i|e)$ を求める．

知りたい確率変数 x_i の各値についての事後確率 $P(x_i|e)$ は，あり得るすべての状態で平均化する周辺化によって求める．そのために，条件付き確率表を用いる．

周辺化によれば，たとえば確率変数 x_i が状態 X^l をとる確率 $P(x_i=X^l|e)$ は次式で与えられる．

$$P(x_i=X^l|e) = \frac{\displaystyle\sum_{j=1,j\neq i}^{N}\sum_{x_j=X^1}^{X^L}P(x_1,\cdots,x_i=X^l,\cdots,x_N,e)}{\displaystyle\sum_{j=1}^{N}\sum_{x_j=X^1}^{X^L}P(x_1,\cdots,x_N,e)} \tag{3}$$

ここで，$\sum_{x_j=X^1}^{X^L}$ は確率変数 x_j の取り得るすべての場合 $X^1, X^2, \cdots, X^L$ について総和を取ることを意味する．

3. 提案手法

3.1 提案手法のプロセス

本研究では，あらかじめ測定されている株価収益率と，それをベイジアンネットワークで予測した誤差の両方を確率変数ノードとしてとって株価収益率を予測する．

提案手法の株価収益率予測プロセスを以下に示す．

(1) 株価収益率をウォード法で離散化する．

(2) 離散化された過去の株価収益率によりベイジアンネットワーク B を決定する．

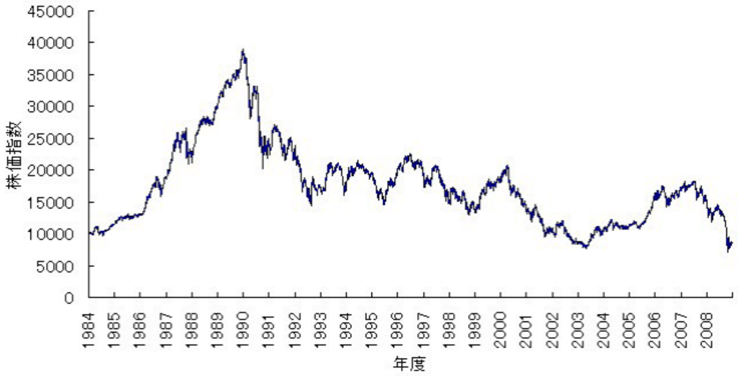

図 1　1984 年-2008 年の日経平均株価

(3)　B を用いて株価収益率を予測する．
(4)　得られた予測株価収益率と実際の株価収益率より予測誤差を求める．
(5)　予測誤差を離散化する．
(6)　離散化された過去の株価収益率, 及び予測誤差によりベイジアンネットワーク B' を決定する．
(7)　B' を用いて株価収益率を予測する．

個々の処理については，以下で説明する．

3.2　株価収益率の離散化

本研究では，1985 年から 2008 年までの 6000 日の日経平均株価のデータを用いる (図 1). 次式を用いて株価から株価収益率を求める[5)].

$$r_t = (\ln P_t - \ln P_{t-1}) \times 100 \tag{4}$$

ここで P_t は t 期の株価終値を，r_t は t 期の株価収益率を示す. ベイジアンネットワークの各ノードが持つ確率変数では連続値を扱うことが出来ないため，連続値である株価収益率をウォード法[6)] により離散化する．

離散化した株価の離散値を r^l, 離散値総数を L とおくと, 離散値の集合は以下のように表される．

$$r^l = \{r^1, r^2, \cdots, r^L\} \tag{5}$$

3.3　ベイジアンネットワーク B の決定と株価収益率予測

日経平均株価から求めた株価収益率から予測用のベイジアンネットワークを決定する．K2 アルゴリズムを適用するための確率変数の全順序関係としては，株価指数の時系列における順序を用いる．ウォード法を用いて離散化された過去の株価収益率だけを用いて株価収益率予測用のベイジアンネットワーク B を決定する．

このとき，t 期の株価収益率 $r_t = r^l$ となる確率を $P(r^l|B)$ と定義する．そして，t 期の株価収益率として $P(r^l|B)$ が最大となる r^l を選択する．つまり,

$$\bar{r}_t = \arg\max_{r^l}(P(r^l|B)) \tag{6}$$

3.4　予測誤差の測定と離散化

t 期での r^l の実際の値を r^l_t，ベイジアンネットワーク B による予測値を $\bar{r}^l_t$ とする．誤差はステップ関数である r^l と $\bar{r}^l$ との差として次式で定義される．

$$Er_t = r^l_t - \bar{r}^l_t \tag{7}$$

Er^l をウォード法で離散化する．離散化された株価収益率を Er^l，離散値総数を L とおくと離散値の集合は以下のように表される．

$$Er^l = \{Er^1, \cdots, Er^L\} \tag{8}$$

3.5　ベイジアンネットワーク B' の決定と株価収益率予測

日経平均株価から求めた株価収益率と，そのネットワーク B による予測値から求めた誤差から予測用のベイジアンネットワーク B' を決定する．

K2 アルゴリズムを適用するための確率変数の全順序関係として，株価指数と予測誤差の時系列における順序を用いる．K2 アルゴリズムによって決定したベイジアンネットワークを B' とする．これを用いて株価収益率 $r_t = r^l$ となる確率を推定した結果を $P'(r^l|B')$ と定義する．t 期の株価収益率として $P'(r^l|B')$ が最大となる r^l を選択する．つまり,

$$\bar{r}_t = \arg\max_{r^l}(P'(r^l|B')) \tag{9}$$

4.　数 値 実 験

解析対象として日経平均株価を用いる．学習期間として 1985 年から 2008 年までの 6000 営業日分をとり，予測期間として 2008 年 12 月の 1ヶ月間をとる．

4.1　ベイジアンネットワーク B による予測

株価収益率のクラスタ数を 6 個としてウォード法を用いて離散化する．この離散値を用いて，予測する日の過去 5,10,15 日間のデータを用いて予測を行う．決定したネットワークを

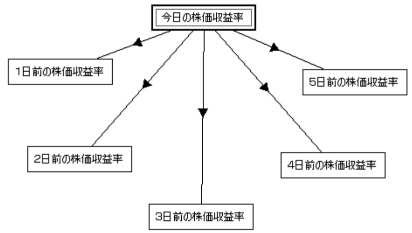

図 2 ベイジアンネットワーク図 (誤差ノード無し,5 日前までの株価利用)

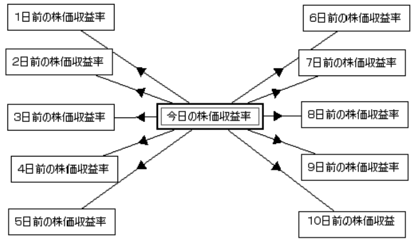

図 3 ベイジアンネットワーク図 (誤差ノード無し,10 日前までの株価利用)

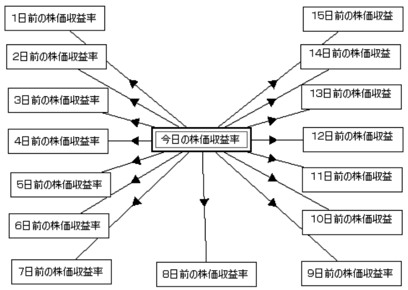

図 4 ベイジアンネットワーク図 (誤差ノード無し,15 日前までの株価利用)

図 2，3，4 に示す．これらより得られたネットワークがナイーブベイズ構造をしていることがわかる．また，予測精度の比較を表 1 に示す．表には実データに対する相関係数と最小二乗誤差を示している．結果より，相関性があまり強くないことがわかる．

表 1 実験 1 結果表

学習ノード	相関係数	最小二乗誤差
5 日前	0.056474	3.339943
10 日前	0.189224	4.074093
15 日前	0.380121	3.786995

4.2 ベイジアンネットワーク B' による予測

次に，株価収益率と予測誤差の両方を用いて予測した結果を示す．データをウォード法を用いて 6 クラスタに離散化して解析を行う．

決定したネットワークを図 5，6，7 に示す．先の実験のネットワークがナイーブベイズ構造であったのに対して，今回のネットワークは複雑な構造となっていることがわかる．次に，予測精度の比較を表 2 に示す．結果より，相関係数と平均二乗誤差の両方とも改善していることがわかる．

5. 結　　論

ベイジアンネットワークを用いれば確率変数の条件付き因果関係をグラフネットワークとして表現できる．そこで，本研究では，ベイジアンネットワークを用いて，過去の株価の条件付き因果関係を確率的にモデル化し，それを用いて将来株価を予測する方法について述べた．従来の方法では予測誤差を測定するために予測値と実測値の差を用いていたが，本研究では離散化された予測値と実測値の差を用いることで，アルゴリズムを簡単化した．また，以前の研究では過去 10 期までのデータを用いて株価を予測するベイジアンネットワークを構築しているのに対して，本研究では 5，10，15 期の 3 種類を比較した．

解析例において．過去の株価収益率だけを用いる場合と，その予測誤差をも用いる方法を比較した．ネットワークを比較すると，前者がナイーブベイズ構造となったのに対して，後者はより複雑なネットワーク構造となった．そして，実データに対する相関係数と最小二乗誤差で比較したところ，後者は前者に対して精度向上が見られた．

今後は，株価変化率の離散化数の調節やレジームスイッチングモデルの利用で，予測精度を改善したいと考えている．

参 考 文 献

1) 繁桝算男, 本村陽一, 植野真臣. ベイジアンネットワーク概説. 培風館, 2006.
2) 左毅, 北栄輔. ベイジアンネットワークを用いた株価予測について. 情報処理学会論文

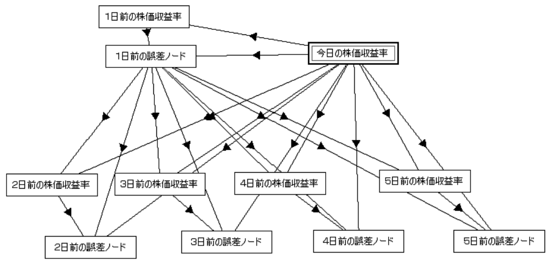

図 5 ベイジアンネットワーク図 (誤差ノード有り,5 日前までの株価利用)

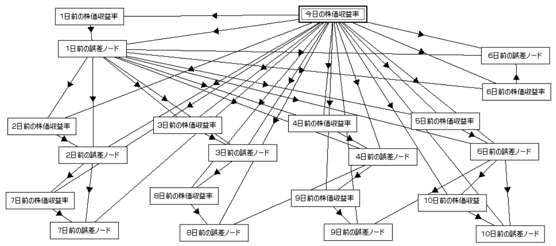

図 6 ベイジアンネットワーク図 (誤差ノード有り,10 日前までの株価利用)

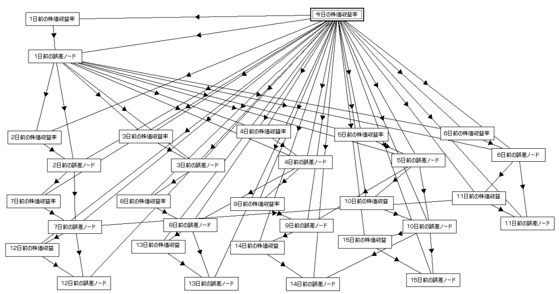

図 7 ベイジアンネットワーク図 (誤差ノード有り,15 日前までの株価利用)

誌トランザクション 数理モデル化と応用, Vol.3, No.3, pp. 80–90, 2010.

3) D.Heckerman, D.Geiger, and D.Chickering. Learning Bayesian networks: The combination of knowledge and statistical data. *Machine Learning*, Vol.20, No.3, pp. 197–243, 1995.

表 2 実験 2 結果表

学習ノード	相関係数	最小二乗誤差
5 日前	0.73905	2.018397
10 日前より	0.941114	1.015445
15 日前より	0.970635	0.790162

4) G.F. Cooper and Herskovits E. A Bayesian method for the induction of probabilistic networks from data. *Machine Learning*, Vol.9, No.4, pp. 309–347, 1992.

5) 渡部敏明. 日本の株式市場におけるボラティリティの変動–ARCH 型モデルによる分析–. 三菱経済研究所, 1997.

6) J.H. Ward. Hierarchical grouping to optimize an objective function. *Journal of the American Statistical Association*, Vol.58, pp. 236–244, 1963.