

ベイジアンネットワークの株価予測への応用

左 毅

目次

第 1 章	緒論	1
1.1	研究背景と研究目的	1
1.2	本論文の構成	6
第 2 章	研究背景	9
2.1	株価指数	9
2.2	線形モデル	10
2.2.1	時系列分析法	10
2.2.2	SV モデル	13
2.2.3	D1 モデル	14
2.3	非線形モデル	14
2.3.1	マルコフ連鎖モンテカルロ法 (MCMC)	14
2.3.2	フラクタル	16
2.3.3	ニューラルネットワーク	16
2.3.4	隠れマルコフモデル	17
2.4	提案手法	18
2.4.1	ベイジアンネットワークの概要	20
2.4.2	提案手法の概要	20
第 3 章	ベイジアンネットワーク	25
3.1	概要	25
3.2	関連する概念	27
3.2.1	条件付き確率	27
3.2.2	ベイズ推定	28
3.3	グラフ構造の決定	29
3.3.1	構造の探索アルゴリズム	29

3.3.2	グラフ構造の評価関数	31
3.3.3	K2 アルゴリズム	33
3.4	確率推論	33
第 4 章	提案手法 1	37
4.1	概要	37
4.2	アルゴリズム	37
4.2.1	株価リターンの算出	37
4.2.2	株価リターンの離散化	37
4.2.3	グラフ構造の決定	38
4.2.4	株価リターンの予測	39
4.3	日経平均株価への適用	39
4.3.1	離散化法によるネットワークの比較	41
4.3.2	離散化法の予測精度への影響	41
4.3.3	離散値数の予測精度への影響	42
4.4	トヨタ自動車の株価への適用	47
4.4.1	離散化法によるネットワークの比較	49
4.4.2	離散化法の予測精度への影響	49
4.4.3	離散値数の予測精度への影響	50
第 5 章	提案手法 2	55
5.1	概要	55
5.2	アルゴリズム	56
5.2.1	株価リターンの算出	56
5.2.2	株価リターンの離散化	56
5.2.3	グラフ構造 B の決定	56
5.2.4	株価リターンの予測	56
5.2.5	予測誤差の算出	56
5.2.6	予測誤差の離散化	56
5.2.7	グラフ構造 B' の決定	57
5.2.8	株価リターンの再予測	57
5.3	日経平均株価への適用	57

5.3.1	予測精度の比較 (1)	58
5.3.2	予測精度の比較 (2)	63
5.3.3	計算時間	63
5.4	トヨタ自動車株価への適用	66
5.4.1	予測精度の比較 (1)	66
5.4.2	予測精度の比較 (2)	70
第 6 章	提案手法 3	73
6.1	概要	73
6.2	アルゴリズム	73
6.3	数値実験	74
6.4	考察	76
第 7 章	結論	79
7.1	本論文のまとめ	79
7.2	今後の課題	80
	謝辞	83
	参考文献	85

目 次

2.1	ニューラルネットワークの3層模式図	17
2.2	隠れマルコフモデル	18
2.3	日経平均株価の株価リターン	19
2.4	日経平均株価の株価リターンの頻度分布図	19
2.5	ベイジアンネットワーク	20
2.6	時系列分析法と提案手法の比較	21
3.1	ベイジアンネットワークの応用例	26
3.2	確率変数の条件付き依存関係	27
3.3	3ノードの場合の構造候補数	29
3.4	変数ノードの全順序関係	30
3.5	K2アルゴリズムによる3ノードの場合の構造候補数	30
3.6	グラフ構造決定アルゴリズム	34
4.1	日経平均株価によって決定されたベイジアンネットワーク(等分割クラスタリング)	40
4.2	日経平均株価によって決定されたベイジアンネットワーク(ワード法)	41
4.3	離散値数とAICの関係(ワード法)	43
4.4	トヨタ自動車株価の株価リターンの頻度分布図	47
4.5	トヨタ自動車株価によって決定されたベイジアンネットワーク(等分割クラスタリング)	48
4.6	トヨタ自動車株価によって決定されたベイジアンネットワーク(ワード法)	48
4.7	離散値数の精度への影響(ワード法)	51
5.1	株価リターンと予測誤差の全順序関係	57
5.2	株価リターンと誤差で決定されたベイジアンネットワーク(日経平均株価)	62
5.3	提案手法2と時系列分析法及び提案手法1の予測精度の比較(日経平均株価)	63

5.4	予測期間が異なる場合の予測値と実測値の比較 (日経平均株価)	64
5.5	提案手法 1 と提案手法 2 の計算コストの比較	65
5.6	トヨタ自動車株価の株価リターンと予測誤差で決定されたベイジアンネットワーク	66
5.7	予測値と実測値の比較 (トヨタ自動車株価)	71
5.8	予測期間が異なる場合の予測値と実測値の比較 (トヨタ自動車株価)	72
6.1	提案手法 3 で決定されたベイジアンネットワーク (日経平均株価)	74
6.2	提案手法 3 と提案手法 1 及び提案手法 2 の計算コストの比較	76

表 目 次

3.1	例の事前確率と条件付き確率	27
4.1	日経平均株価リターンの分類区間 (等分割クラスタリング)	40
4.2	日経平均株価リターンの分類区間 (ウォード法)	40
4.3	提案手法 1 と時系列分析の予測精度の比較 (日経平均株価)	42
4.4	日経平均株価の株価リターンにおける離散値数と AIC の関係	43
4.5	日経平均株価リターンの分類区間 (ウォード法, $L=2$)	43
4.6	日経平均株価リターンの分類区間 (ウォード法, $L=3$)	44
4.7	日経平均株価リターンの分類区間 (ウォード法, $L=4$)	44
4.8	日経平均株価リターンの分類区間 (ウォード法, $L=5$)	44
4.9	日経平均株価リターンの分類区間 (ウォード法, $L=7$)	45
4.10	日経平均株価リターンの分類区間 (ウォード法, $L=8$)	45
4.11	日経平均株価リターンの分類区間 (ウォード法, $L=9$)	46
4.12	日経平均株価リターンの分類区間 (ウォード法, $L=10$)	46
4.13	トヨタ自動車株価リターンの分類区間 (等分割クラスタリング)	47
4.14	トヨタ自動車株価リターンの分類区間 (ウォード法)	48
4.15	提案手法 1 と時系列分析の予測精度の比較 (トヨタ自動車株価)	50
4.16	トヨタ自動車株価リターンにおける離散値数と AIC の関係	50
4.17	トヨタ自動車株価リターンの分類区間 (ウォード法, $L=2$)	51
4.18	トヨタ自動車株価リターンの分類区間 (ウォード法, $L=3$)	51
4.19	トヨタ自動車株価リターンの分類区間 (ウォード法, $L=4$)	51
4.20	トヨタ自動車株価リターンの分類区間 (ウォード法, $L=6$)	52
4.21	トヨタ自動車株価リターンの分類区間 (ウォード法, $L=7$)	52
4.22	トヨタ自動車株価リターンの分類区間 (ウォード法, $L=8$)	52
4.23	トヨタ自動車株価リターンの分類区間 (ウォード法, $L=9$)	53
4.24	トヨタ自動車株価リターンの分類区間 (ウォード法, $L=10$)	53

5.1	日経平均株価予測誤差における離散値数と AIC の関係	58
5.2	日経平均株価リターンの分類区間 (ウォード法, $L=2$)	59
5.3	日経平均株価の予測誤差の分類区間 (ウォード法, $L=3$)	59
5.4	日経平均株価の予測誤差の分類区間 (ウォード法, $L=4$)	59
5.5	日経平均株価の予測誤差の分類区間 (ウォード法, $L=5$)	59
5.6	日経平均株価の予測誤差の分類区間 (ウォード法, $L=6$)	60
5.7	日経平均株価の予測誤差の分類区間 (ウォード法, $L=7$)	60
5.8	日経平均株価の予測誤差の分類区間 (ウォード法, $L=8$)	60
5.9	日経平均株価の予測誤差の分類区間 (ウォード法, $L=9$)	61
5.10	日経平均株価の予測誤差の分類区間 (ウォード法, $L=10$)	61
5.11	提案手法 2 と時系列分析法及び提案手法 1 の予測精度の比較 (日経平均株価)	62
5.12	予測期間が異なる場合の予測値と実測値の比較 (日経平均株価)	64
5.13	提案手法 1 と提案手法 2 の計算時間の比較	65
5.14	トヨタ自動車株価予測誤差における離散値数と AIC の関係	67
5.15	トヨタ自動車株価の予測誤差の分類区間 (ウォード法, $L=2$)	67
5.16	トヨタ自動車株価の予測誤差の分類区間 (ウォード法, $L=3$)	67
5.17	トヨタ自動車株価の予測誤差の分類区間 (ウォード法, $L=4$)	67
5.18	トヨタ自動車株価の予測誤差の分類区間 (ウォード法, $L=5$)	68
5.19	トヨタ自動車株価の予測誤差の分類区間 (ウォード法, $L=6$)	68
5.20	トヨタ自動車株価の予測誤差の分類区間 (ウォード法, $L=7$)	68
5.21	トヨタ自動車株価の予測誤差の分類区間 (ウォード法, $L=8$)	69
5.22	トヨタ自動車株価の予測誤差の分類区間 (ウォード法, $L=9$)	69
5.23	トヨタ自動車株価の予測誤差の分類区間 (ウォード法, $L=10$)	70
5.24	提案手法 2 と時系列分析法及び提案手法 1 の予測精度の比較 (トヨタ自動車 株価)	71
5.25	予測期間が異なる場合の予測値と実測値の比較 (トヨタ自動車株価)	71
6.1	日経平均株価予測誤差における離散値数と AIC の関係	75
6.2	提案手法 3 と提案手法 2 の予測精度の比較 (日経平均株価)	75
6.3	提案手法 3 と提案手法 1 及び提案手法 2 の計算時間の比較	77

第1章 緒論

1.1 研究背景と研究目的

株価の精度良い予測は、投資家個人による投資行動の決定だけでなく、企業の経営方針や国家運営などさまざまな問題に関わっている。毎日の生活では、株価に関するニュースを新聞やテレビで頻繁に耳にしたり目にしたりすることが多い。最近では、株価の低迷が長期化していることから、円高傾向がこれ以上続くと自動車、機械、電機など輸出産業の業績に対するさらなる悪影響が懸念される。そこで、株価予測に関する研究が広く行われている。

株価とは、株式の価格であり、特定企業の株式の売買取引の際に付けられる価格と株式市場全体の平均株価の2種類がある。平均株価として、代表的なものには東京証券取引所の日経平均株価（Nikkei Stock Average-N225）[11]、ニューヨーク証券取引所のダウ平均株価（Dow Jones Industrial Average-Dow）[12]とロンドン証券取引所のFTSE100種総合株価指数（Financial Times Stock Exchange 100 Index-FTSE100）[13]の3つがあり、时时刻刻各自の市場で変動していて、国家の景気動向に対する1つの指標としてとらえられることも多い。

株式市場における株価の暴騰や暴落が、いつ頃、どのような原因によって起こるかについて、多くの人々が強い関心を持っている。弱気相場とは、数ヶ月から数年にわたる株式市場の低迷を意味する。弱気相場は、必ずしも暴落を伴うわけではないが、暴落と結び付けられることが多い。

史上最も有名な暴落は1929年のウォール街大暴落であり[7][14][19]、これは通常ブラック・サズデーと呼ばれる。1920年代にアメリカを中心として世界の経済は急成長し、家電製品、自動車、航空機などの発明が普及の黄金期であった。この時代は、一般的に「黄金の20年代」または「狂騒の20年代」と呼ばれる。この時代の先陣を切ったラジオ産業大手のRCAや自動車産業大手のGMの株価は急上昇した。金融会社株についてもウォール街の銀行家がゴールドマンサックスなど投資信託会社株を買い好調だった。投資家は信用取引にレバレッジを効かせて得た株式市場からのリターンに有頂天だったが、その年の夏

には経済に限界が見え 10 月前半には株価も低下していくことが誰の目にも明らかだった。この株価不安に投資家は浮き足立ち事態は一気に現実化した。10 月 24 日（ブラック・サーズデー）には株価急落の第 1 波が襲来した。さらに追い立てるように 10 月 28 日のブラック・マンデー、10 月 29 日のブラック・チューズデーが続いた。大恐慌が底をつく 1932 年 7 月時点でダウ式平均株価は 89% 下落した。

また、1987 年 10 月 19 日、ニューヨーク証券取引所の株価が大暴落し、NY ダウは 508 ドル、率にして 22.61% という大幅な下げを記録した。この暴落は一取引日の損失としてはウォール街の歴史で最大である。アメリカの新聞や雑誌はこの日をブラック・マンデーと呼び、1929 年 10 月 24 日のブラック・サーズデーに対比して、大恐慌の再来ではないかと書きたてた [7] [14] [19]。ニューヨークの株価大暴落はたちまち世界中に伝播して、ロンドン、パリ、フランクフルト、東京などの取引所では 19 日から 20 日にかけていずれも株価が暴落し、中には香港のように混乱を恐れて取引所を閉鎖するところまで出てきた。1987 年の暴落の結果、ニューヨーク証券市場にサーキット・ブレーカーという取引制限が導入された。サーキット・ブレーカーとは、強制的市場停止のことであり、取引期間中に市場株価が大きく下落すれば、いつでも発動される。冷却期間をおくことで投資家のパニックを防ぐために設けられた制度である。

日本証券経済研究所によれば、1987 年の日本の株式投資はブラック・マンデーという大幅な株価の下落があったにもかかわらず、14.6% の収益率を示した。更に、1988 年の投資も株価の上昇が続いたため、22.4% の収益率に達していた。また、1989 年末の時価総額を比較すると、日本 4.26 兆ドル、アメリカ 2.90 兆ドルとなっていた。これは、日本の株式市場が当時世界一の規模だったということを示している。1989 年 12 月に日経平均は 38915.87 円まで上昇して、そして、1990 年 10 月に 20221.86 円まで下落した [7] [8] [14]。1990 年の株価の急落とその後の低迷の要因として、よく指摘されるのは、「金融引締め」、「株式需給の悪化（海外投資家の売却、裁定取引など）」、「その他一時的要因（イラクのクウェート侵攻と石油供給不安）」である。

2008 年の暴落は、2007 年までに NY ダウが史上最高値を更新し続けたことに端を発する。その景気拡大をささえていたサブプライムローンは 2006 年頃から安定的な運用を疑問視され始めていたが、なお、それは拡大を続けていた。世界において、これはサブプライムローン危機と呼ばれる。2007 年 6 月 22 日、米大手証券ベアスターン傘下のヘッジファンドがサブプライムローンに関連した運用に失敗したことが明らかになると、7 月 10 日には米格付け機関のムーディーズがサブプライムローンを組み込んだ住宅ローン担保証券 RMBS

の大幅格下げを発表し、金融収縮が始まった。その後一時平静を取り戻し、NYダウは10月には再度史上最高値を更新する。しかし、その裏で住宅価格の下落は止まらず、各金融会社から大きな損失が公表されたことにより再度金融不安が拡大した。2008年9月15日、ニューヨーク証券取引所4位（当時）のリーマンブラザーズの破綻と同3位のメリルリンチのバンク・オブ・アメリカによる救済合併、翌16日には、米最大の保険会社AIGの経営危機による国営化が明らかになると金融収縮の懸念から、アメリカは緊急経済安定化法案を議会で提出することになる。しかし、9月29日下院でそれが否決されるとNYダウは史上最悪となる-777.68ドルをつけるなど、一気に金融信用収縮が加速した。10月1日には下院で修正案が可決されたものの、時既に遅く、その毒は欧州に回り、6日にはロシアのRTS指数が約19%程度下落し、アイスランドでは対ユーロでクローナが約30%程度下落し、同国では全ての銀行が国有化されるなど、未曾有の世界同時金融危機が本格化した。翌9日、NYダウが-678.91ドルをつけると、日経平均は翌10日881.06円安（-9.62%、過去5番目）と暴落した。今なお、世界中の株式市場では依然として低迷傾向にある。

したがって、世界の株式市場はこれから数ヶ月、数年、また数十年後にどのような動きをするのかは、私達の経済の未来にかかわる大きな疑問である。1999年以来、インターネットの普及や株式手数料の完全自由化によって、個人投資家による株式投資がますます盛んになっている。その結果、ブラック・マンデーやバブル崩壊のような大きな出来事が発生した時、直接被害を受ける投資家が以前より多くなっている。最近では、サブプライムローン問題に起因した2008年9月の世界同時株安で多くの人が損を被ったことが記憶に新しい。これらのことから、効果的な株価予測を必要とする投資家が、以前より一層多くなっている。

株価の予測方法には、大きく分けてテクニカル分析（チャート分析）とファンダメンタルズ分析がある [21] [22]。テクニカル分析とは過去に発生した株価の変化から将来の株価を予測する手法である。多くの経済学者や金融工学者はこれを「根拠が無く、科学的理論とはいえない」と批判している。一方、市場予測の根拠を過去の株価の変動ではなく、現実の要因（株価リターンなど）に求める手法をファンダメンタルズ分析という。現実の要因としては、景気や国際環境などの外部環境、企業の財務体質や業績予想が考慮される。このため、ファンダメンタルズ分析のほうが信頼性が高いといえる。しかし、ファンダメンタルズ分析では「信頼性の高い情報」という特徴がある反面、「複雑で難解な分析」、「情報の分析に時間がかかる」という欠点をもっている。特に分析時間の点からは、短期投資では利用しにくいと言える。

また、過去3、40年の経済発展とともに金融工学 [35] [36] が発展し、その過程でさまざま

なテクニカル分析法が開発された。金融工学において用いられる株価分析法は大きく線形分析法と非線形分析法 [39] [42] [65] に分けることができる [21]。線形分析法 [15] [23] [24] [28] [29] [30] 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、その重み係数を相関解析などにより決定する。これらのなかで、代表的な手法は時系列分析法である。AutoRegressive(AR) モデル [23] [29] [30] は自己回帰モデルであり、今期株価を過去の株価の線形和で近似する。Moving Average(MA) モデル [23] [29] [30] は移動平均モデルであり、今期の株価を過去の攪乱項の移動加重和で近似する。AutoRegressive Moving Average(ARMA) モデル [23] [29] [30] は AR と MA を組み合わせたモデルである。さらに、AutoRegressive Conditional Heteroskedasticity(ARCH) モデル [23] [24] [28] [29] [30] は、ボラティリティをリスクの指標として用い、その変化を捉えられる時系列モデルである。ARCH モデルは、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ている。その後、ARCH の拡張として、多数の改良型 Generalized ARCH(GARCH) [23] [29] [30] [48] や Exponential GARCH(EGARCH) [23] [29] [30] [34] が提案されている。このようなモデルは確率的ボラティリティ変動 (Stochastic Volatility-SV) モデル [23] [90] [93] [96] [97] [98] [99] [100] [101] と呼ばれ、ボラティリティのショック持続性を捉えることができる。また、天候デリバティブにおいて、気温を予測するために、Dischel-D1 モデル [102] [103] がよく用いられる。これは、平均回帰過程モデルを拡張したモデルとして、今期値を過去値の移動加重和と 1 期先の値の線形結合で近似する。これらの線形モデルは、指数関数型崩壊・増殖あるいは周期的の変動しか表現できないので、予期できないランダム事象を表現するために、正規乱数を用いる。

しかし、近年のいくつかの研究 [5] [104] [105] [106] によれば、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ので、正規乱数を用いたモデルの場合には予測精度が低下することが予想される。また、カオス理論 [41] によれば、ランダムな入力から不規則の出力を導く過程では、得られた予測値と実測値とのずれが大きく、十分な予測精度を得られない可能性があることがわかる。そこで、株価市場の複雑性、不安定性、非線形性を考慮し、幾つかの非線形モデルが株価予測に適用されている。

非線形分析法で、近年最も注目されている方法は、マルコフ連鎖モンテカルロ法 (Markov Chain Monte Carlo methods-MCMC) [9] [10] [16] [91] [92] である。代表的なものとしては、M-H アルゴリズム (Metropolis-Hastings Algorithm) [16] [107] やギブスサンプリング (Gibbs Sampling) [16] [108] がある。十分に多くの回数の試行を行った後のマルコフ連鎖の状態を求める標本は目標分布に近似する。試行の回数を増やすとともにサンプルの

品質も向上する。しかし、十分な精度を要求される場合、計算コストがかなりかかるのは、MCMC の不可避の問題である。

また、池田ら [6] はフラクタル (Fractal) 性 [109] [110] をもつ時系列の予測手法を提案している。フラクタルとはフランスの数学者ブノワ・マンデルブロ (Benoit Mandelbrot) が導入した幾何学の概念で、図形の部分と全体が自己相似になっているものなどをいう。物理学・自然現象の解明、信号処理、図形符号化など工学分野をはじめ、さまざまに応用がなされている。フラクタルの研究者によると、マンデルブロは株価チャートを見ていてフラクタルの着想を得たという。株価のような時系列の一部分を拡大、縮小した場合に同じような形状をもつ相似形が出現する性質を説明するフラクタル理論が展開されている。なお、株価予測に適用する時、時系列性をもつ株価はフラクタル性を仮定する必要がある。

もう1つ注目される方法は、グラフィカルモデルである [46] [111]。グラフィカルモデルでは、さまざまな事象が相互に関連しあった問題において、いくつかの情報が手に入ったとき、観測されなかった情報をしばしば推測しなくてはならない。グラフィカルモデルのうち、最も有名なものがニューラルネットワーク (Neural Network-NN) [72] [73] [74] である。ニューラルネットワークは、脳機能に見られるいくつかの特性を計算機上のシミュレーションによって表現することを目指した数学モデルであり、パターン認識やデータマイニングをはじめ、さまざまな分野において応用されている [52] [53] [54] [75] [76] [77] [78]。近年では、ニューラルネットワークを用いて株価予測の研究が活発に行われている [49] [79] [80] [81]。計算機の性能向上は大規模データを扱うことを容易にし、これにいつそう拍車をかけている。先行の研究のいくつかは、このような工学的手法が株価予測にある程度は役に立つことを示している。ただ、株価予測に対して、入出力間の関係が見えなく、株価どうしの依存関係の解釈が困難である指摘が示されている。もう1つのグラフィカルモデルは、隠れマルコフモデル (Hidden Markov Model-HMM) [56] [69] [71] がある。隠れマルコフモデルは確率モデルの一つで、「システムがパラメータ未知のマルコフ過程である」と仮定し、観測可能な情報からその未知のパラメータを推定する。これも、パターン認識への応用 [68] [70] のほか、株価予測への研究 [47] [57] [58] が示されている。なお、パラメータ間のマルコフ性を仮定する。

本研究では、ベイジアンネットワーク [1] [4] [25] [31] [32] [40] を用いて株価変化を推定することを提案する [26] [27]。ベイジアンネットワークは、確率的なネットワークモデルである。確率変数をノードで表し、変数間の因果関係を非循環有向リンクで、変数間の定量的な依存関係を条件付き確率によって表す。最初の提案手法 1 では、このベイジアンネット

ワークを用いて、過去の株価リターンから将来の株価リターンを予測する。このとき、株価リターンの分布を正規分布などによって仮定する必要がない。なお、ベイジアンネットワークでは株価のような連続値をとる変数を扱うことができない。そこで、本研究では、クラスタリング手法 [50] [51] [59] [63] [64] を用いて株価を離散値に変換する。各ノードは、離散化された株価を確率変数として有し、各ノードの離散値間の因果関係をベイジアンネットワークによって推定する。

その後、提案手法2について述べる。提案手法2では、提案手法1で過去の株価リターンを予測した後、その予測値の誤差を評価する。誤差分布をクラスタリング手法によって離散化して誤差分布の離散値を求め、誤差ノードを定義する。誤差ノードを株価リターンノードと共に用いてベイジアンネットワークを構築し、これを株価リターン予測に用いる。また、提案方法1では、求められるネットワークはNaiveベイズ構造である。そこで、提案手法2では、計算コストを下げるためにNaiveベイズ構造を仮定してネットワーク構造を探索する。

最後に、提案手法3について述べる。提案手法3は提案手法2と同じアルゴリズムであるが、提案手法2と異なり、ベイジアンネットワークを構築する時、Naiveベイズ構造を仮定しないでネットワーク構造を探索する。

解析例として日経平均株価とトヨタ自動車の株価を用いる。日経平均株価は、そのリターンが正規分布に比較的近い場合の例として扱うのに対して、トヨタ自動車株価はリターンが正規分布から離れている例として用いる。数値実験を行い、提案手法1, 2, 3による予測結果の精度をAR, MA, ARMA, ARCHと比較し、提案手法の有効性を検討する。

1.2 本論文の構成

次章以降の本論文の構成は以下のようになっている。

第2章において、研究背景として、各予測手法について説明する。時系列分析法の定義とパラメータ次数の決定方法まで説明した後、非線形モデルMCMC、池田らが提案したフラクタル時系列予測法について説明する。次に、グラフィカルモデルとして、ニューラルネットワークと隠れマルコフモデルについて説明した後、提案手法について紹介し、他のモデルとの比較を述べる。

第3章において、ベイジアンネットワークのアルゴリズムについて説明する。最初に、ベイジアンネットワークについて紹介する。次に、ベイジアンネットワークにおける条件付

き確率，グラフ決定アルゴリズム，確率推論を説明する．グラフ決定アルゴリズムでは，探索アルゴリズムとしての K2 アルゴリズムとネットワークの評価指標としての K2Metric について詳細に説明する．

第 4, 5, 6 章において，提案手法 1, 2, 3 のそれぞれのアルゴリズムと数値実験について説明する．数値実験の対象として日経平均株価とトヨタ自動車株価を用い，本研究で提案する手法を時系列分析手法と比較する．日経平均株価は，そのリターンが正規分布に比較的近い場合の例として扱うのに対して，トヨタ自動車株価はリターンが正規分布から離れている例として用いる．

第 7 章において，本論文の結論を述べるとともに，本研究に関連する今後の展望を述べる．

第2章 研究背景

2.1 株価指数

株価指数 [8] [11] [12] [13] [22] とは、株式市場の状況を示すために、各銘柄の株価を一定の計算方法によって統合し、数値化したものである。つまり、株価全体の動きを端的に表しているものだと言える。さらに言うならこの指数はその時期の経済の様子自体を表すものである。株式市場全体の様子を見るには、なんらかの「平均」があると便利であるが、各銘柄はそれぞれ発行株式数が異なり、株価のスケールも異なるので、単純に銘柄の株価を平均しても、株式市場全体を表した数字にはならない。そこで、「平均」を出すために、一定のルールに従って計算に用いる銘柄を決め、それら銘柄の株価を元に計算を行う。そうして算出されたものが株価指数になる。銘柄の選定ルールと計算方法によって、さまざまな株価指数がある。代表的なものとして日経平均株価 (Nikkei Average Stock-N225) [11] , ダウ・ジョーンズ工業平均株価 (Dow Jones Industrial Average-Dow) [12] , FTSE100 種総合株価指数 (Financial Times Stock Exchange 100 Index-FTSE100) [13] などがある。

日経平均株価は日本の株式市場を代表する株価指数である [11] 。60年に及ぶ歴史があり、内外の投資家や株式市場関係者に最もよく知られている。日経平均株価は「ダウ式平均」によって算出する指数である。基本的には225銘柄の株価の平均値であるが、分母(除数)の修正などで株式分割や銘柄入れ替えなど市況変動以外の要因を除去して指数値の連続性を保っている。日経平均株価の計算式 [11] は式(2.1)と式(2.2)で与えられる。

$$\text{日経平均株価} = \frac{\text{225 銘柄の指数採用株価合計}}{\text{除数}} \quad (2.1)$$

$$\text{指数採用株価} = \text{株価} \times \frac{50}{\text{みなし額面}} \quad (2.2)$$

東証第一部上場銘柄のうち取引が活発で流動性の高い225銘柄を選定し、ダウ平均株価の株価平均型方式を基にした計算方法で修正平均を算出する。この銘柄の修正方法が「ダウ式」と呼ばれ、日経平均は「ダウ式修正平均」と言われる。

2.2 線形モデル

2.2.1 時系列分析法

時系列分析法 [15] [23] [24] [28] [29] [30] とは、時間経過ごとに記録された数値列からモデルを作成して、将来の予測を行う分析手法のことである。時系列の例としては、気温や、株価、為替レートなどがある。株価予測の代表的な手法は時系列分析である AutoRegressive (AR) モデル [23] [29] [30] は自己回帰モデルであり、今期株価を過去の株価の線形結合で近似する。Moving Average (MA) モデル [23] [29] [30] は移動平均モデルであり、今期株価を過去の攪乱項の移動加重和で近似する。AutoRegressive Moving Average (ARMA) モデル [23] [29] [30] [67] は AR と MA を組み合わせたモデルである。さらに、AutoRegressive Conditional Heteroskedasticity (ARCH) モデル [23] [24] [28] [29] [30] [66] は、ボラティリティをリスクの指標として用いて、その変化を捉えられる時系列モデルである。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 (GARCH, EGARCH) [23] [34] [29] [30] [48] が提案されている。時系列分析の手法は、計量経済学をはじめ、各分野で活用されている。

1. AR モデル

AR モデル $AR(p)$ では、 t 期の株価リターン r_t を p 期前までのリターン r_{t-i} と、擾乱項 u_t によって次のように表す。

$$r_t = \alpha_0 + \sum_{i=1}^p \alpha_i r_{t-i} + u_t \quad (2.3)$$

ここで、 α_i はモデルのパラメータである。 u_t は平均 0、分散 σ^2 のホワイトノイズに従う。

本研究では、文献 [23] にしたがって、予測誤差を最小とするように AR モデルのパラメータ次数 p を決定する。文献 [23] では、予測誤差が独立同一の正規分布にしたがうと仮定すれば、予測誤差の平均 2 乗誤差は予測誤差の標準偏差 σ と同値なので、予測誤差を最小にする基準として式 (2.4) で定義される AIC を用い、これを最小とするようにモデルパラメータを決定している。

$$AIC = \ln \sigma^2 + \frac{2p}{T} \quad (2.4)$$

ここで、 σ^2 は AR モデルの残差標本の分散であり、 T は株価リターン標本の総数である。

$\bar{\sigma}^2$ は次式で求められる .

$$\bar{\sigma}^2 = \frac{1}{T-p} \sum_{t=p+1}^T (r_t - \bar{r}_t)^2 \quad (2.5)$$

ここで r_t は式 (2.3) より求められる t 期のリターンの期待値を , \bar{r}_t は実測値を示す .

解析例では $p = 1, 2, \dots, 10$ について解析を行い , 式 (2.4) が最小値となる p を採用する .

2. MA モデル

MA モデル MA(q) では , t 期の株価リターン r_t を過去の攪乱項 u_{t-j} によって以下のように表す .

$$r_t = \beta_0 + \sum_{j=1}^q \beta_j u_{t-j} + u_t \quad (2.6)$$

ここで , β_j はモデルのパラメータである .

本研究では , 文献 [23] にしたがって , 予測誤差を最小とるように MA モデルのパラメータ次数 q を決定する . 文献 [23] では , 予測誤差が独立同一の正規分布にしたがうと仮定すれば , 予測誤差の平均 2 乗誤差は予測誤差の標準偏差 σ と同値なので , 予測誤差を最小にする基準として式 (2.7) で定義される AIC を用い , これを最小とるようにモデルパラメータを決定している .

$$\text{AIC} = \ln \bar{\sigma}^2 + \frac{2q}{T} \quad (2.7)$$

ここで , $\bar{\sigma}^2$ は次式で求められる .

$$\bar{\sigma}^2 = \frac{1}{T-q} \sum_{t=q+1}^T (r_t - \bar{r}_t)^2 \quad (2.8)$$

解析例では , MA モデルのパラメータ次数 q が $q = 1, 2, \dots, 10$ について解析を行い , 式 (2.7) が最小値を取るときの q を採用する .

3. ARMA モデル

ARMA モデル ARMA(p, q) では , t 期の株価リターン r_t を株価リターン r_{t-i} と攪乱項 u_{t-j} によって次のように表す .

$$r_t = \sum_{i=1}^p \alpha_i r_{t-i} + \sum_{j=1}^q \beta_j u_{t-j} + u_t \quad (2.9)$$

本研究では , 文献 [23] にしたがって , 予測誤差を最小とるように ARMA モデルのパラメータ次数 p, q を決定する . 文献 [23] では , 予測誤差が独立同一の正規分布にしたがう

と仮定すれば、予測誤差の平均 2 乗誤差は予測誤差の標準偏差 $\bar{\sigma}$ と同値なので、予測誤差を最小にする基準として式 (2.10) で定義される AIC を用い、これを最小とするようにモデルパラメータを決定している。

$$\text{AIC} = \ln \bar{\sigma}^2 + \frac{2(p+q)}{T} \quad (2.10)$$

ここで、 $\bar{\sigma}^2$ は次式で求められる。

$$\bar{\sigma}^2 = \frac{1}{T - (p+q)} \sum_{t=(p+q)+1}^T (r_t - \bar{r}_t)^2 \quad (2.11)$$

解析例では、ARMA モデルのパラメータ次数 p, q が $p = 1, 2, \dots, 10$ と $q = 1, 2, \dots, 10$ について解析を行い、式 (2.10) が最小値を取るときの p, q を採用する。

4. ARCH モデル

ARCH モデル ARCH(p, q) では、 t 期の株価リターン r_t を次式で表す。

$$r_t = \alpha_0 + \sum_{i=1}^p \alpha_i r_{t-i} + u_t \quad (2.12)$$

また、擾乱項 u_t は次式で与えられる。

$$u_t = \sigma_t z_t \quad (2.13)$$

ここで、 $\sigma_t > 0$ であり、 z_t は平均 0、分散 1 の正規乱数である。

Engle の ARCH モデル [23] [24] [28] [29] [30] によれば、 σ_t^2 は次式で近似される。

$$\sigma_t^2 = \beta_0 + \sum_{j=1}^q \beta_j u_{t-j}^2 \quad (2.14)$$

本研究では、文献 [23] にしたがって、予測誤差を最小とするように ARCH モデルのパラメータ次数 p, q を決定する。予測誤差の平均 2 乗誤差は予測誤差の標準偏差 $\bar{\sigma}$ と同値なので、予測誤差を最小にする基準として次式で定義される AIC を用いる。

$$\text{AIC} = \ln \bar{\sigma}^2 + \frac{2(p+q)}{T} \quad (2.15)$$

ここで、 $\bar{\sigma}^2$ は ARCH モデルの残差標本 $\sigma_{p+q+1}^2 - \bar{u}_{p+q+1}^2, \sigma_{p+q+2}^2 - \bar{u}_{p+q+2}^2, \dots$ の分散であり、 T は株価リターン標本の総数である。 $\bar{\sigma}^2$ は次式で求められる。

$$\bar{\sigma}^2 = \frac{1}{T - (p+q)} \sum_{t=(p+q)+1}^T (\sigma_t^2 - \bar{u}_t^2) \quad (2.16)$$

ここで、 σ_t^2 は t 期の株価リターン r_t の条件付き分散であり、 \bar{u}_t^2 は t 期株価リターンの予測誤差の 2 乗である。

解析例では、ARCH モデルのパラメータ次数 p, q が $p = 1, 2, \dots, 10$ と $q = 1, 2, \dots, 10$ について解析を行い、式 (2.15) が最小値を取るときの p, q を採用する。

5. GARCH モデルと EGARCH モデル

Engle の ARCH モデルを拡張したのが Bollerslev の提案した GARCH (Generalized-eralized AutoRegressive Conditional Heteroscedastic) モデル [23] [29] [30] [48] である。ARCH モデルの条件付き分散が過去の分散だけでなく過去の残差の 2 乗にも依存するという関係を組み込んでいる。ARCH から GARCH への拡張は、ちょうど AR を ARMA に拡張したものと対応している。 σ_t^2 は次式で近似される。

$$\sigma_t^2 = \beta_0 + \sum_{j=1}^q \beta_j u_{t-j}^2 + \sum_{k=1}^s \omega_k \sigma_{t-k}^2 \quad (2.17)$$

ここで、パラメータに非負制約を課すのは、ボラティリティ σ_t^2 の非負性を保証するためである。

Nelson [34] は、ボラティリティを被説明変数とするのではなく、その対数値を被説明変数とすることによって、こうしたパラメータの非負制約を取り除いている。さらに、提案した EGARCH モデルによれば、ARCH モデルや GARCH モデルでは捉えることのできないボラティリティ変動の非対称性を捉えることができる。

2.2.2 SV モデル

確率的ボラティリティモデル (Stochastic Volatility-SV) [23] [90] [93] [94] [95] [96] [97] [98] [99] [101] は、数理ファイナンスにおいてオプションなどのデリバティブや有価証券を評価するのに使われるモデルである。原資産となる有価証券のボラティリティを確率過程として取り扱うことからこのような名称となっている。

ブラック・ショールズ方程式ではボラティリティは定数として取り扱われ時間や原資産価格の変動に影響されないと仮定している。しかしこのモデルでは、インプライド・ボラティリティがオプションの権利行使価格や権利行使期日によって異なることを示唆するボラティリティ・スマイル (Volatility Smile) やボラティリティ・スキュー (Volatility Skew) を説明できない。そこで確率的ボラティリティモデルでは、原資産価格のボラティリティは

時間あるいは原資産価格などの状態変数の変化により影響を受けると仮定することで、より正確なモデル化を可能としている。

SV モデルでは、ボラティリティの対数値の変動を線形の ARMA モデルによって定式化する。通常、次のような SV モデルが用いられる。ここで、 η_t は過去と独立かつ同一な平均 0、分散 σ_η^2 の正規分布に従うことを表している。

$$\ln(\sigma_t^2) = \omega + \phi \ln(\sigma_{t-1}^2) + \eta_t \quad (2.18)$$

また、 σ_t^2 ではなくその対数値を考えるのは、 σ_t^2 が負にならないようにするためである。さらに、この式では、こうしたボラティリティのショックの持続性を ϕ によって測ることができる。

2.2.3 D1 モデル

気温予測に用いられるモデルに Dischel の D1 モデル [102] [103] がある。Dischel の気温モデルでは温暖化傾向が考慮されておらず、自己回帰性に関しては前日の気温のみの影響を反映させている。日付 t における気温 T_t は以下のように表される。

$$T_t = \alpha \theta_t + (1 - \alpha) T_{t-1} + u_t \quad (2.19)$$

T_t は t 日の気温、 θ_t は t 日目にあたる日付の平均気温、 α は重み係数、 u_t は平均 μ 、分散 σ^2 の正規分布に従う独立な確率変数とする。気温の予測と同じように、近年、株価、金利、為替レートへの研究もいくつか示されている。

2.3 非線形モデル

2.3.1 マルコフ連鎖モンテカルロ法 (MCMC)

マルコフ連鎖モンテカルロ法 (Markov Chain Monte Carlo-MCMC) [9] [10] [16] [91] [92] は、多変量の確率分布からサンプルを得るための一連の手法である。代表的なものとしていわゆるギブス・サンプラーがある。ここで、「サンプルを得る」というのは、実世界で調査や実験をしてサンプルを得るという意味ではなく、数式などで明示的に与えられた確率分布からの具体的なサンプルを得るという意味である。

パラメータを $\theta = \theta_1, \theta_2, \dots, \theta_n$, 観測データを y , パラメータの事後分布を $P(\theta|y)$ とする. 次のように反復的に θ を発生させることにより, 事後分布からの標本を得るのがギブス・サンプラーである.

Step1. まず初期値として $\theta = \theta_1^{(0)}, \theta_2^{(0)}, \dots, \theta_n^{(0)}$ を適当な分布から発生させる.

Step2. $\theta = \theta_1^{(i)}, \theta_2^{(i)}, \dots, \theta_n^{(i)}$ ($i > 0$) が得られたら

Step2a. $\theta_1^{(i+1)}$ を $P(\theta_1^{(i+1)}|\theta_1^{(i)}, \theta_2^{(i)}, \dots, \theta_n^{(i)})$ から発生させる.

Step2b. $\theta_2^{(i+1)}$ を $P(\theta_2^{(i+1)}|\theta_1^{(i+1)}, \theta_2^{(i)}, \dots, \theta_n^{(i)})$ から発生させる.

Step2c. $\theta_3^{(i+1)}$ を $P(\theta_3^{(i+1)}|\theta_1^{(i+1)}, \theta_2^{(i+1)}, \dots, \theta_n^{(i)})$ から発生させる.

同様に $\theta_4^{(i)}, \dots, \theta_n^{(i)}$ を順次発生させていく. $\theta^{(i+1)} = \theta_1^{(i+1)}, \theta_2^{(i+1)}, \dots, \theta_n^{(i+1)}$ を得たら, i を $i+1$ として Step2 に戻り繰り返す. 最後に, θ の系列 $\theta^{(0)}, \theta^{(1)}, \dots, \theta^{(m)}$ の標本を得る.

例えば, (X_1, X_2, X_3) の 3 つの変数があり, 各変数は状態 -1 または 1 をとるとする. ギブス・サンプラーで書き換える順番は $X_1 \quad X_2 \quad X_3$ となる.

1. 初期状態を $(1, 1, 1)$ とする.
2. X_1 を -1 と置き換え, 他の 2 つをそのままにした状態を次の候補 $(-1, 1, 1)$ とする.
3. 候補 $(-1, 1, 1)$ の確率 $r = P(X_1 = -1|X_2 = 1, X_3 = 1)$ を求める.
4. $0 < R < 1$ の一様乱数を生成する.
5. $R < r$ なら候補 $(-1, 1, 1)$ を採用する (受理, accept). そうでなければ, 現在の $(1, 1, 1)$ をそのまま次の状態とする (棄却, reject).

このようにして, 採用された $(-1, 1, 1)$ において X_2 を -1 と置き換え, 次の状態の候補 $(-1, -1, 1)$ とする. ここで, 2 から 5 まで評価して $R < r$ と仮定し, 候補の $(-1, -1, 1)$ を棄却して, 現在の $(-1, 1, 1)$ を採用することとする. 続いて, 採用された $(-1, 1, 1)$ によって X_3 を -1 と置き換え, 次の状態の候補 $(-1, 1, -1)$ とする. ここで, 2 から 5 まで評価して $R < r$ と仮定し, 候補の $(-1, 1, -1)$ を採用することとする. このようにして, 新しい標本 $(-1, 1, -1)$ が生成され, 標本空間 $\{(1, 1, 1), (-1, 1, -1)\}$ が得られる.

ギブス・サンプラーはさまざまな確率分布からの乱数を発生する方法である. 一様乱数や正規乱数など標準的な確率分布に従う乱数の発生は, 多くのソフトウェアに組み込み関数として用意されている.

2.3.2 フラクタル

フラクタル研究の始まりは、イギリスの気象学者ルイス・フライ・リチャードソン (Lewis Fry Richardson) [109] の国境線に関する検討である。国境を接するスペインとポルトガルは、国境線の長さとしてそれぞれ 987 km と 1214 km と別の値を主張していた。リチャードソンは、国境線の長さは用いる地図の縮尺によって変化し、縮尺と国境線の長さがそれぞれ対数を取ると直線状に相関することを発見した。このような特徴をフラクタルと名付けて一般化したのがマンデルブロである。フラクタルの研究者高安秀樹 [110] によると、マンデルブロは株価チャートを見ていてフラクタルの着想を得たという。

フラクタルの特徴は直感的に理解できることと、数学的に厳密に定義することが非常に難しい。マンデルブロはフラクタルを「ハウスドルフ次元が位相次元を厳密に上回るような集合」と定義した。完全に自己相似なフラクタルにおいては、ハウスドルフ次元はミンコフスキー次元と等しくなる。

フラクタルを定義する際の問題には次のようなものがある。

1. 「不規則すぎる」と正確な意味が存在しない。
2. 「次元」の定義が唯一でない。
3. 物体が自己相似である方法がいくつも存在する。
4. 全てのフラクタルが再帰的に定義されるとは限らない。

2.3.3 ニューラルネットワーク

ニューラルネットワーク (Neural Network) [72] [73] [74] は、脳機能に見られるいくつかの特性を計算機上のシミュレーションによって表現することを目指した数学モデルである。代表的なニューラルネットワークは、三層パーセプトロンモデルの例を図 2.1 に示す。研究の源流は生体の脳のモデル化であるが、神経科学の知見の改定などにより次第に脳モデルとは乖離が著しくなり、生物学や神経科学との区別のため、人工ニューラルネットワーク (Artificial Neural Network-ANN) とも呼ばれる。

ニューラルネットワークの学習は、教師信号 (正解) の入力によって問題に最適化されていく教師あり学習と、教師信号を必要としない教師なし学習に分けられる。明確な解答が用意される場合には教師あり学習が、データ・クラスタリングには教師なし学習が用いられる。結果としていずれも次元削減されるため、画像や統計など多次元量のデータでか

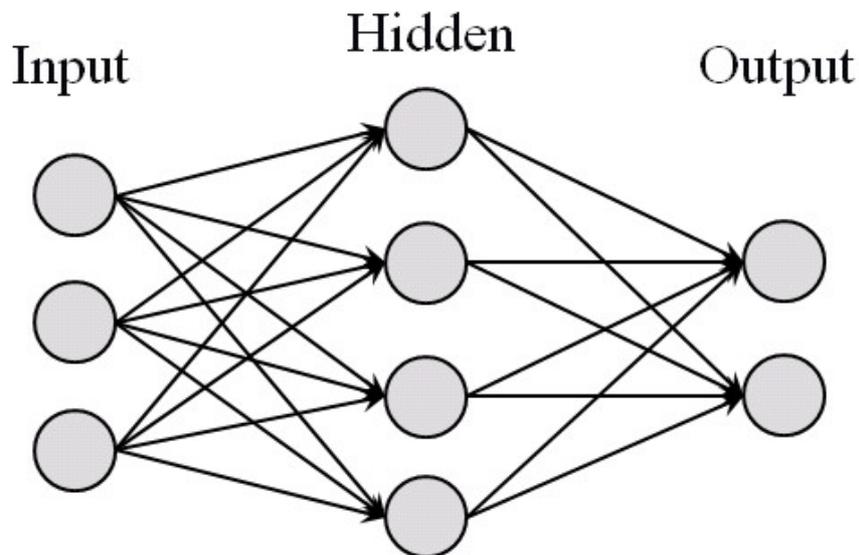


Fig. 2.1: ニューラルネットワークの3層模式図

つ線形分離不可能な問題に対して、比較的小さい計算量で良好な解を得られることが多い。このことから、パターン認識やデータマイニングをはじめ、さまざまな分野において応用されている。

1986年にRumelhart, Hinton, Williams [82] によってパーセプトロンの問題を解決したバックプロパゲーション（誤差逆伝播法）が提案された。このモデルも入力層、中間層（隠れ層）、出力層の3層から構成されるフィードフォワード型のモデルで、基礎はパーセプトロンである。

2.3.4 隠れマルコフモデル

隠れマルコフモデル（Hidden Markov Model-HMM）[56] [69] [71] は確率モデルの一つである。「システムがパラメータ未知のマルコフ過程である」と仮定し、観測可能な情報からその未知のパラメータを推定する。音声認識、ゲノミクス、形態素解析（自然言語処理）などに応用されている [68] [70]。連続的かつ伸縮しうる信号列のパターン抽出には適しているが、反面、長い距離をはさんで呼応しているような信号列からのパターン認識には、間の距離の長さに応じて状態数を増やす必要があり、計算量の観点から実用的ではない。また、局所最適に陥りやすいため、対象に応じて適切なパラメータの初期値を設定する（適切なモデルトポロジーを導入する）必要がある。隠れマルコフモデルの例をは図 2.2 に示す。

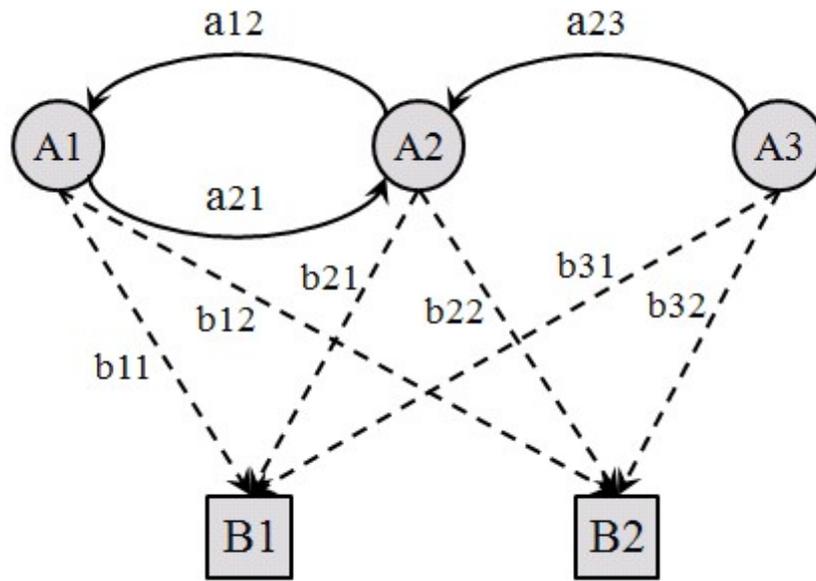


Fig. 2.2: 隠れマルコフモデル

A は内部の状態，a は内部状態の遷移確率，B は出力状態，b は出力確率とする．

観測系列のデータから生成する過程は次のようになる．

1. 初期状態と初期状態の確率を決定する．
2. 遷移状態 $S(t+1) = A_j$ を現在の状態 $S(t) = A_i$ と状態遷移確率 a_{ij} によって決定する．
3. 状態遷移する際に出力状態 B_k を出力確率 b_{jk} によって決定する．
4. step 2 に戻る．

2.4 提案手法

AR, MA, ARMA, ARCH など線形モデルでは株価データの分布が正規分布に従うことを仮定している．しかし，実際の株価変動の解析から，株価の頻度分布は完全には正規分布に従わないことが指摘されている．例として，日経平均株価の株価リターンを図 2.3 に示す．この頻度分布に，それから求めた正規分布を重ねた図が図 2.4 である．これより，分布は正規分布に近いけれども，少し異なっていることがわかる．後述するように，このずれは一般企業の株価ではもっと大きい場合がある．原らは TOPIX の頻度分布を解析し，ボ

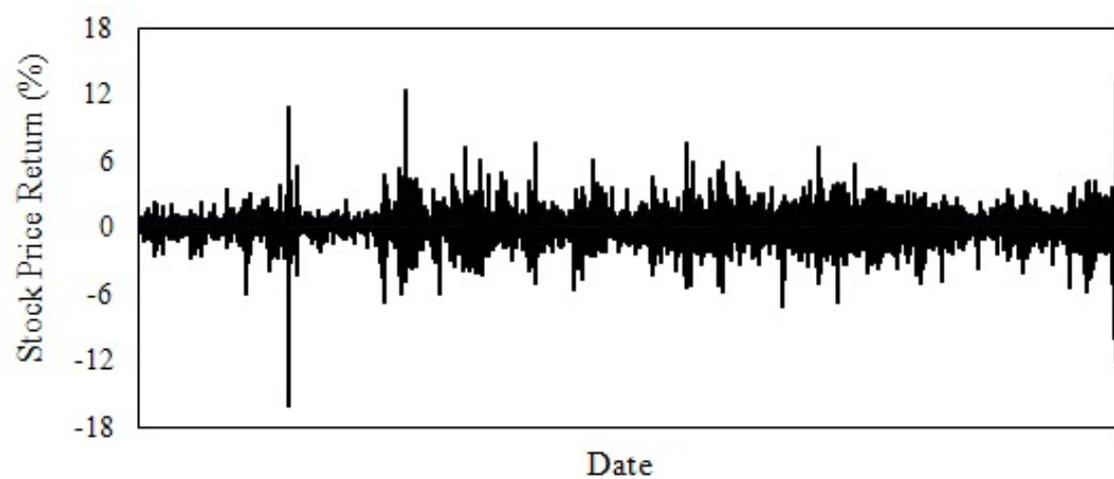


Fig. 2.3: 日経平均株価の株価リターン

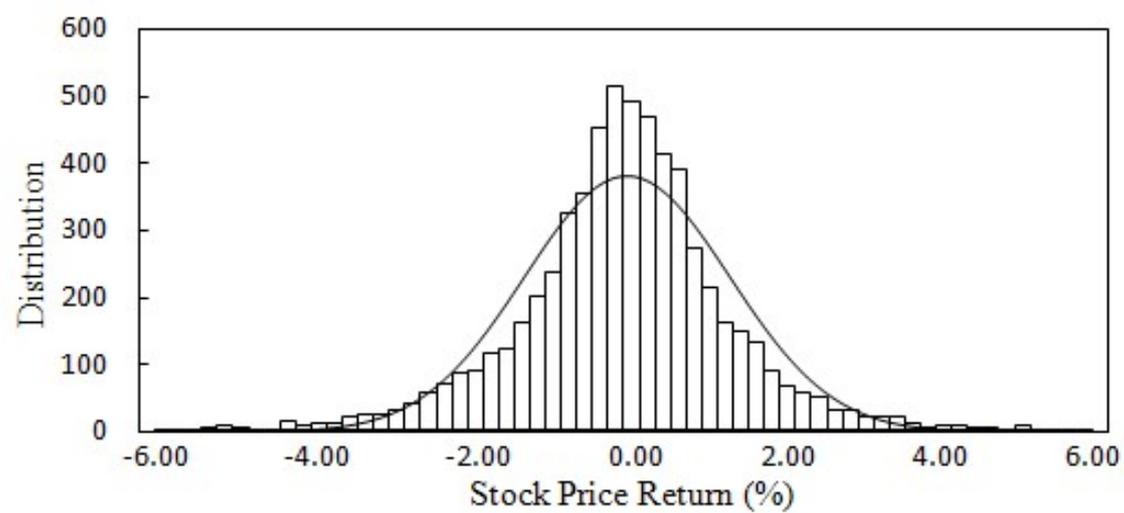


Fig. 2.4: 日経平均株価の株価リターンの頻度分布図

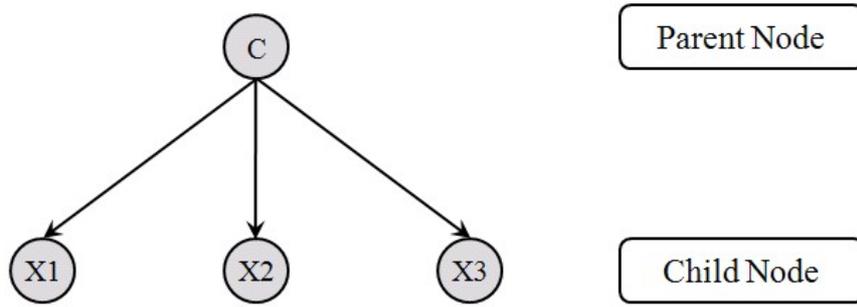


Fig. 2.5: ベイジアンネットワーク

ラティリティを σ とすると、実際の頻度分布ではイベントが $\pm\sigma$ 内にある頻度は正規分布よりも大きく、また、 $\pm 3\sigma$ の外にある頻度は正規分布の3倍ほどであることを指摘している [5]。このことを改善するために、頻度分布の確率密度関数の尖度や突度を修正する研究が示されている。これに対して、本研究ではベイジアンネットワークを用いて株価の時系列分析を行う。

2.4.1 ベイジアンネットワークの概要

ベイズ推定によれば、過去に起きた事象の発生頻度から未来に起こる事象の発生頻度を求めることができる。そして、ベイズ推定に基礎をおくベイジアンネットワーク [1] [4] [25] [31] [32] [40] を用いれば、確率変数をノードで、変数間の因果関係をリンクで表すことで因果関係を表す非循環有向グラフを得ることができる。確率変数間の条件付き依存関係をリンクとして表し、リンクの先にあるノードを子ノード、リンクの元にあるノードを親ノードと呼ぶ (図 2.5)。

本研究では、このベイジアンネットワークを用いて、過去の株価リターンから今期株価リターンを予測する。

2.4.2 提案手法の概要

時系列分析法では、株価リターンや誤差の分布が正規分布に従うと仮定し、今期の株価リターンを過去の株価リターンと擾乱項の線形結合で近似している。しかし、株価の頻度分布は完全には正規分布に従わないことが指摘されている。

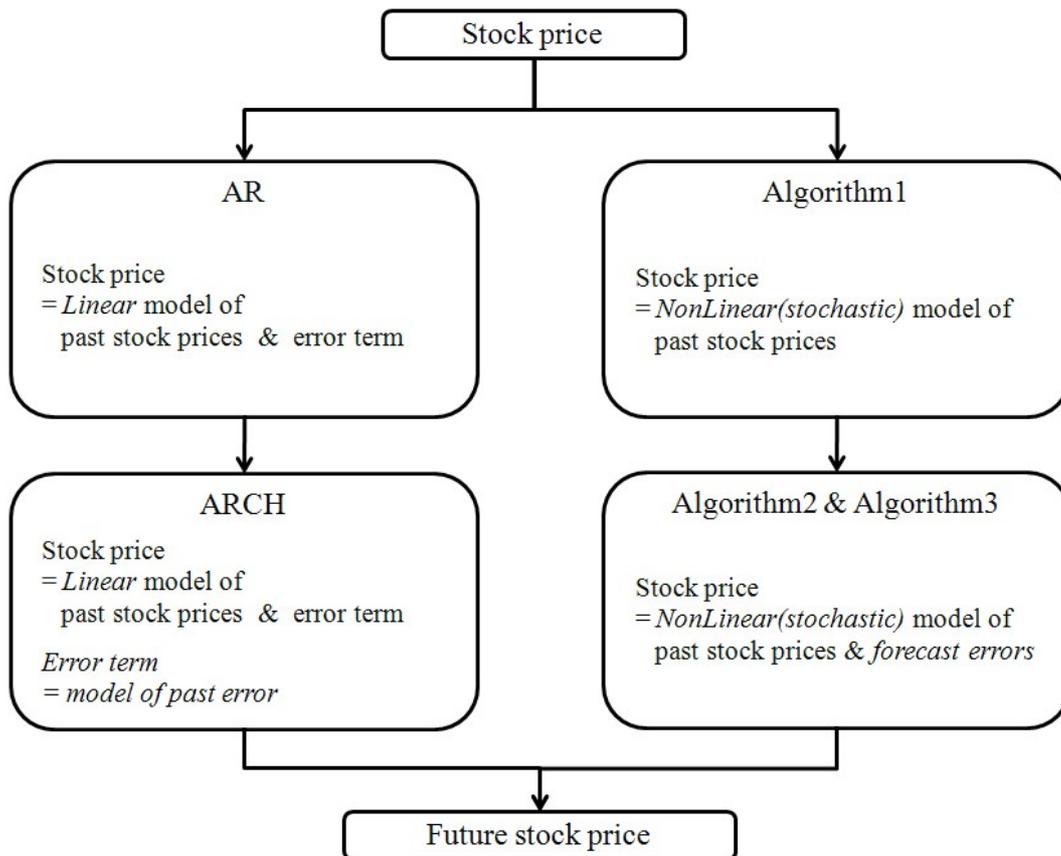


Fig. 2.6: 時系列分析法と提案手法の比較

ベイジアンネットワークでは、確率変数をノード、確率変数間の定性的な依存関係をリンクとする非循環有向グラフで可視化し、変数間の定量的な依存関係を条件付き確率によってモデル化する。本研究の提案手法では、ベイジアンネットワークを用いて、過去の株価リターンどうしの因果関係をモデル化し、これを予測に用いている。このとき分布には正規乱数などによる仮定を用いないので、時系列分析法の問題点を改善できる可能性がある。

本研究の提案手法1では、株価リターンだけを確率変数ノードとしてベイジアンネットワークを構築し、これを用いて将来の株価リターンを予測する。この方法では、株価リターンの分布を正規分布などによって仮定する必要がない。ただし、ベイジアンネットワークではノードにおかれる確率変数は連続値をとることができないので、クラスタリング手法などを用いて株価リターンを離散化する操作が必要である。従来の時系列分析手法では、今期の株価リターンを過去の説明変数と攪乱項の線形結合で近似する。例えば、ARモデルでは株価リターンを過去の株価リターンと攪乱項で表す。これに対して、提案手法1では今期株価リターンと過去の株価リターンとの間の非線形関係式をベイジアンネットワークによって定義して予測するという見方もできる(図2.6)。

本研究の提案手法2では、提案手法1の予測精度向上を目的として、株価リターンだけでなくその予測誤差をも用いてベイジアンネットワークを構築する。従来手法のベイジアンネットワークを用いて、すでに測定されている過去の株価リターンを予測すれば、その予測値と実際値から誤差を評価できる。提案手法2では、この誤差値のノードを定義し、株価リターンのノードとともに用いて、株価リターン予測用のベイジアンネットワークを構築する。ただし、提案手法1で得られた結果を考慮して計算コストを下げるために Naive ベイズ構造 [3] を仮定してネットワーク構造を探索する。

提案手法2の着想は、時系列分析法におけるARモデルとARCHモデルの関係にある。図2.6に示すように、ARモデルでは予測株価リターン r_t を過去リターン r_{t-i} と攪乱項 u_t で近似する。ARCHモデルではARモデルにおける攪乱項 u_t のボラティリティを過去の攪乱項 u_{t-i} で近似する。つまり、ARCHモデルでは予測株価リターン r_t を過去リターン r_{t-i} と過去の攪乱項 u_{t-i} で近似している。一方、ベイジアンネットワークを用いた提案手法1では、予測株価リターン r_t を過去の株価リターン r_{t-i} のみの非線形関係式として定義するためにベイジアンネットワークを用いるのに対して、提案手法2では予測株価リターン r_t を過去の株価リターン r_{t-i} と過去の予測誤差 u_{t-i} の両者の非線形関係式として定義するためにベイジアンネットワークを用いている。ただし、ベイジアンネットワークではノードにおかれる確率変数は連続値をとることができないので、クラスタリング手法などを用い

て予測誤差を離散化する操作が必要である。

しかし，ARCH モデルと提案手法 2 の予測精度向上に対するアプローチは少し異なる．ARCH モデルでは，AR モデルにおける擾乱項 u_t のボラティリティを過去の擾乱項 u_{t-i} で近似することにより，ボラティリティを時間に対して可変として精度を向上している．これに対して，提案手法 2 では，提案手法 1 で既知の株価リターンを予測して予測誤差を計測し，その誤差を元に提案手法 1 で用いたベイジアンネットワークモデル自体を補正することで精度を改善しようとしている．結果として，ARCH モデルでは予測株価リターン r_t を過去リターン r_{t-i} と擾乱項 u_t の線形結合で近似するのに対して，提案手法 2 では予測株価リターン r_t を過去リターン r_{t-i} と擾乱項 u_t の確率的依存関係に基づく非線形関係によってモデル化しているといえる．

さらに，提案手法 2 の改善手法として提案手法 3 を示す．提案手法 3 は提案手法 2 と同様に，誤差ノードを株価リターンノードと共に用いてベイジアンネットワークを構築するが，Naive ベイズ構造を仮定しないでネットワーク構造を探索する．提案手法 3 によって，構築できたネットワークを Tree Augmented Naive Bayes (TAN) [3] と呼ぶことにする．提案手法 3 によるネットワークは提案手法 2 よりも複雑になる．したがって，計算コストは提案手法 2 より増加するが，予測精度は提案手法 2 よりも向上すると期待される．

第3章 ベイジアンネットワーク

3.1 概要

ベイジアンネットワーク [1] [4] [25] [31] [32] [40] とは、過去の事象が現在の事象の確率に影響を与えるというベイズ推定に基づく確率的なグラフィカルモデルである。ベイジアンネットワークは、確率変数、確率変数間の条件付き依存関係とその条件付き確率の3つによって定義される。それを可視化したものが図 3.1 である。ここで、確率変数をノードで表し、確率変数間の条件付き依存関係をノード間に張られたリンクとして表している。リンクは有向であり、リンクの先にあるノードを子ノード、リンクの元にあるノードを親ノードとして表現する。条件付き確率は、親ノードがある値をとった時に、子ノードがある値をとる条件付き確率である。こうして各子ノードとその親ノード間にリンクを張って構築したベイジアンネットワークによって、これらの変数間の確率的な依存関係がモデル化できる。

ベイジアンネットワークの応用においては、基本的に4つのステップがある。それは、確率変数の決定、ネットワーク構造の選択、条件付き確率の獲得、確率推論である。

次に、簡単な例でベイジアンネットワークの予測アプローチを述べる。

朝、出かける時芝生が濡れていることに気づいた。その原因を考えると、「スプリンクラーが動いた」と「雨が降った」の2つが考えられる。また、雨が降った場合、スプリンクラーが動かなかったこととなる。

ここで次の問題を考える。芝生が濡れている時、雨が降った可能性はどの程度だろうか？

1. 確率変数の決定

例の状況によって、事象の雨降り (R)、スプリンクラー (S)、芝生濡れ (G) を抽出し、変数ノードとする。

- $R = r$ は雨が降る。
- $S = s$ はスプリンクラーが動く。

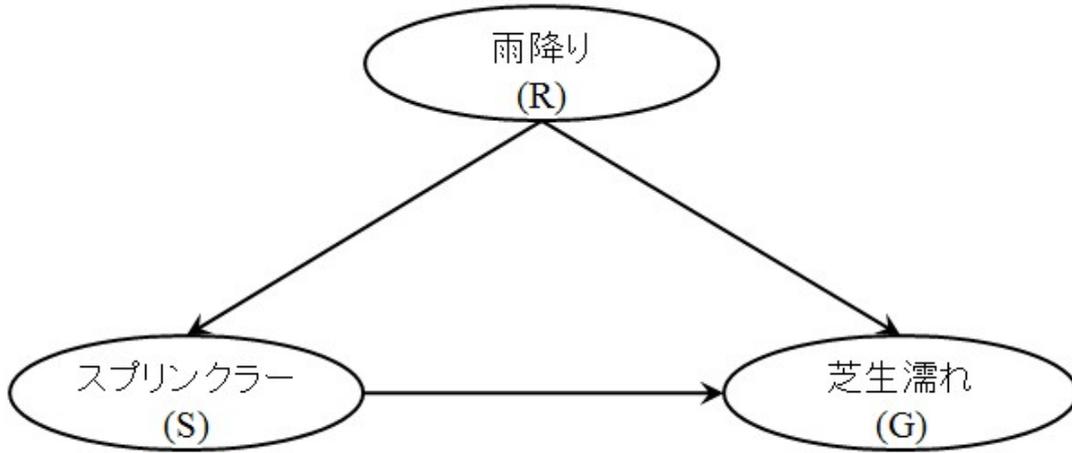


Fig. 3.1: ベイジアンネットワークの応用例

- $G = g$ は芝生が濡れる .

と表記する . それぞれの否定はバーをつけて ,

- $R = \bar{r}$ は雨が降らない .
- $S = \bar{s}$ はスプリンクラーが動かない .
- $G = \bar{g}$ は芝生が濡れない .

と表記する . これらのノードそれぞれは , その事象が生起したか生起しなかったかの 2 値である .

2. ネットワーク構造の選択

ここで , 雨降り (R) , スプリンクラー (S) と芝生濡れ (G) でこの 3 つの変数間の因果関係を図 3.1 のように表す .

3. 条件付き確率の獲得

図 3.1 において他のノードに影響されないノード R の事前確率を評価する . これを表 3.1(a) に示す . そして , ノード R とノード S , 及び (ノード R , ノード S) とノード G の条件付き確率を表 3.1(b) と表 3.1(c) に示す .

4. 確率推論

Table. 3.1: 例の事前確率と条件付き確率

Rain		Sprinkler		Sprinkler	Rain	Grass Wet	
T	F	T	F	F	F	T	F
0.2	0.8	0.4	0.6	F	F	0.0	1.0
		0.01	0.99	F	T	0.8	0.2
				T	F	0.9	0.1
				T	T	0.99	0.01

(a) 雨降り

(b) スプリンクラー

(c) 芝生濡れ

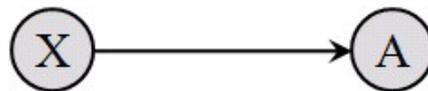


Fig. 3.2: 確率変数の条件付き依存関係

この時，芝生が濡れている場合，雨が降った確率は $P(R = r|G = g)$ となる．ベイズ推定を用いて，式 (3.1) のように与えられる．

$$P(R = r|G = g) = \frac{\sum_{S \in (s, \bar{s})} P(G = g, R = r)}{\sum_{R \in (r, \bar{r}), S \in (s, \bar{s})} P(G = g, R, S)} \quad (3.1)$$

そして，事象 R の全ての想定される原因事象で式 (3.2) のように展開される．式の中に，表 3.1 の条件付き確率によって，解答は 35.77% となる．

$$P(R = r|G = g) = \frac{\sum_{S \in (s, \bar{s})} P(R = r)P(S|R = r)P(G = g|R = r, S)}{\sum_{R \in (r, \bar{r}), S \in (s, \bar{s})} P(R)P(S|R)P(G = g|R, S)} \quad (3.2)$$

3.2 関連する概念

3.2.1 条件付き確率

条件付き確率 (Conditional Probability) [37] [38] とは，ある事象 X が起こるとい条件の下で事象 A が起こる確率のことである．これを $P(A|X)$ と書く．

確率変数 A が X に依存していることを図 3.2 に $X \rightarrow A$ と表現し， X を親ノード， A を子ノードと呼ぶ．子ノードの親ノードに対する依存関係は条件付き確率 $P(A|X)$ で表され

る．これは次式で与えられる．

$$P(A|X) = \frac{P(A \cap X)}{P(X)} \quad (3.3)$$

$P(A \cap X)$ は，同時確率または結合確率（Joint Probability）と呼ばれ，2つの事象がどちらも起こる確率である．

ここで，親ノード X が取り得る状態が I 個あり，そのひとつを記号 X^i とする．一方，子ノード A の取り得る状態が J 個あり，そのひとつを A^j とする．このとき

$$\begin{array}{ccc} P(A^1|X^1), P(A^2|X^1), & \dots, & P(A^J|X^1) \\ \vdots & & \vdots \\ P(A^1|X^I), P(A^2|X^I), & \dots, & P(A^J|X^I) \end{array}$$

を表にしたものを条件付き確率表（Conditional Probability Table-CPT）と呼ぶ．

3.2.2 ベイズ推定

ベイズ推定 [20] [43] [44] [45] とは，ある証拠に基づいて，その原因となった事象を推定するための確率論的方法である．ここで， X を原因， A をそれに対する証拠（つまり原因によって起きたと想定される事象）とすると，条件付き確率によって，ある事象 A が起こるとい条件の下で事象 X が起こる確率を条件付き確率と呼ぶ．これを $P(X|A)$ と書き，次式で与えられる．

$$P(X|A) = \frac{P(X \cap A)}{P(A)} \quad (3.4)$$

また，3.2.1 節の条件付き確率によって，次式が与えられる．

$$P(X \cap A) = P(A|X)P(X) \quad (3.5)$$

つまり， $P(X|A)$ は

$$P(X|A) = \frac{P(A|X)P(X)}{P(A)} \quad (3.6)$$

と表される．

$P(X|A)$ が事後確率， $P(X)$ が事前確率， $P(A|X)$ が尤度関数と呼ばれる．そして，分母の $P(A)$ は，全ての想定される原因事象 $X = X^1, X^2, \dots, X^I$ から

$$P(A) = \sum_{X^i}^{X^I} P(X^i)P(A|X^i) \quad (3.7)$$

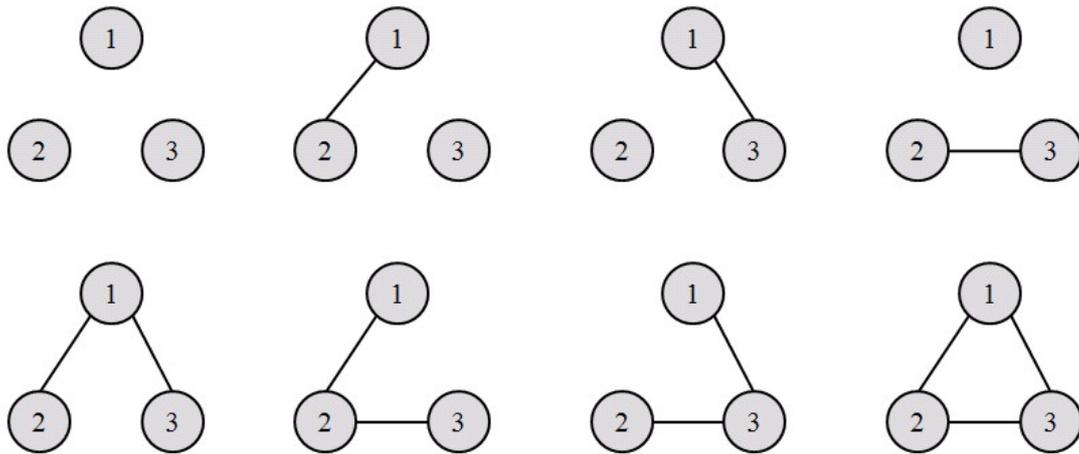


Fig. 3.3: 3 ノードの場合の構造候補数

と求められる． $P(A|X^i)$ は 3.2.1 節の条件付き確率表から得られる．

A の親ノードが

$$X = \{X_1, X_2, \dots, X_N\} \quad (3.8)$$

のように複数ある場合，その親ノード集合を $Pa(A)$ と表現することにする．子ノードの親ノードに対する依存関係は条件付き確率 $P(X|Pa(A))$ で表される．このとき，分母の $P(A)$ は，

$$P(A) = P(A|Pa(A)) \prod_{X_n}^{X_N} P(X_n|Pa(X_n)) \quad (3.9)$$

と求められる．

3.3 グラフ構造の決定

3.3.1 構造の探索アルゴリズム

探索アルゴリズムとしてよく用いられるのは，全探索アルゴリズム，K2 アルゴリズム（欲張り法），遺伝的アルゴリズムによる探索法である．

1. 全探索アルゴリズム

図 3.3 では，3 つの確率変数ノードで無向なグラフを 8 種類に示している．ここで，リンクの方向を考えると，構築できるネットワークの組み合わせ総数は 25 種類となる．確率変



Fig. 3.4: 変数ノードの全順序関係

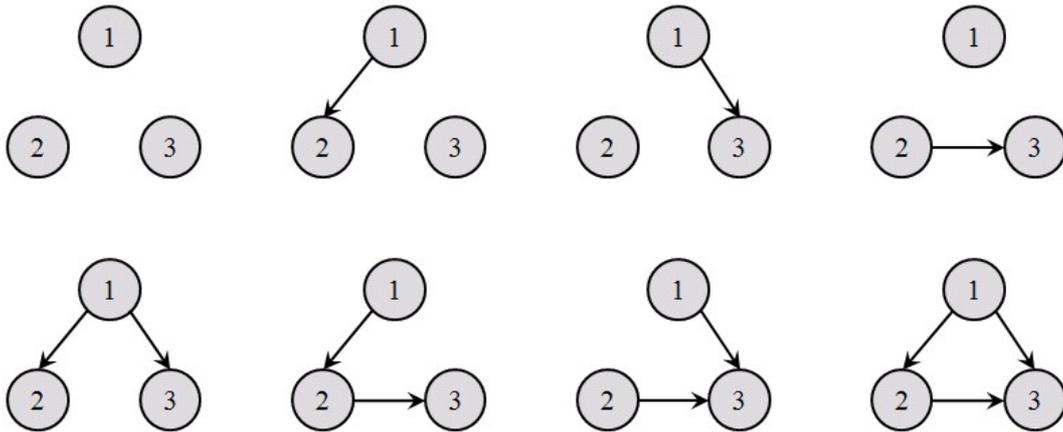


Fig. 3.5: K2 アルゴリズムによる 3 ノードの場合の構造候補数

数ノード n が既知の場合，有向なネットワークの構築数は次式で与えられる [112] .

$$F(n) = \sum_{i=1}^n (-1)^{i+1} (nC_i) 2^{i(n-i)} F(n-i) \quad (3.10)$$

このようにして， $n = 2$ の時構造数は 3， $n = 3$ の時構造数は 25， $n = 5$ の時構造数は 29281， $n = 10$ の時構造数はおよそ 4.2×10^{18} となる．この探索問題は NP (Non-deterministic Polynomial) 問題であり，何らかの工夫により探索計算量を減じなければならぬ [1] .

2. K2 アルゴリズム

前項のような問題に対して，ヒューリスティックを用いた手法が有効であると考えられる．その中でも最も有名な手法は Cooper の K2 アルゴリズムである [25] . このアルゴリズムは，あらかじめ変数間の全順序関係の制約が必要である．つまり，図 3.4 のような全順序関係が与えられる時，探索すべき構造数は図 3.5 のように 8 となる．

K2 アルゴリズムは、変数間の全順序関係に基づいて欲張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 [1] .

3. 遺伝的アルゴリズムによる探索

K2 アルゴリズムは変数ノード間の全順序関係を制約としてグラフ構造を探索する．しかし、このような関係がまったくない場合は、K2 アルゴリズムの適用も難しい．このような場合のために、植野ら、Larranaga ら、Wong らは、遺伝的アルゴリズムを用いた構造探索手法を提案している [113] [114] [115] .

これらの手法では、構造を示す行列 $M = m_{ij}$ と定義する．

$$M = \left\{ \begin{array}{cccc} m_{11}, m_{12}, & \cdots, & m_{1n} \\ \vdots & & m_{ij} & \vdots \\ m_{n1}, m_{n2}, & \cdots, & m_{nn} \end{array} \right\} \quad (3.11)$$

m_{ij} はノード i からノード j にリンクが張られれば 1 として、 m_{ij} はノード i からノード j にリンクが張られなければ 0 とする．このグラフ構造行列の各要素を遺伝的アルゴリズムの遺伝子とみなして適用する手法である．真の構造を見つけるのが困難な大規模なグラフ構造の同定では、非常に有効であることが報告されている [1] .

本研究では、K2 アルゴリズムを用いてグラフ構造の探索を行う．K2 アルゴリズムを用いるためには、確率変数間の全順序関係があらかじめわかっている必要があるが、株価データには時系列性に基づく全順序関係があるため、K2 アルゴリズムを効果的に適用できる．このようにすれば、全順序関係をうまく扱うことができ、実際には親ノード集合の数に制約をつけることにより、より探索空間を小さくすることができる．

3.3.2 グラフ構造の評価関数

一般的に良好なネットワークを選択するため、2つの方法がある．1つは、情報量規準 [17] [18] と呼ばれ、統計モデルの良さを評価するための指標である．統計学の世界において、非常に有名な指標は、赤池が 1971 年に考案し 1973 年に発表した赤池情報量規準 (Akaike's Information Criterion-AIC) である．AIC は、「モデルの複雑さと、データとの適合度とのバランスを取る」ために使用される．例えば、ある測定データを統計的に説明するモデルを作成することを考える．この場合、パラメータの数や次数を増やせば増やすほど、その測

定データとの適合度を高めることができる．その反面，ノイズなどの偶発的な（測定対象の構造と無関係な）変動にも無理にあわせてしまうため，同種のデータには合わなくなる．この問題を避けるには，モデル化のパラメータ数を抑える必要があるが，実際にどの数に抑えるかは難しい問題である．AIC は，この問題に1つの解を与える．具体的にはAIC最小のモデルを選択すれば，多くの場合，良いモデルが選択できる [1] ．

もう1つは，モデルの予測分布を求めることである．あるモデルが特定されれば，そのモデルの予測分布を求めることにより，予測精度の高いモデルを選択できる基準を作成することができると思われ．

ディレクレ分布を事前分布とした一般的なベイジアンネットワークにおける予測分布を求めると

$$BDMetric = \prod_{i=1}^N \prod_{j=1}^M \frac{\Gamma(N'_{ij})}{\Gamma(N'_{ij} + N_{ij})} \prod_{k=1}^L \frac{\Gamma(N'_{ijk} + N_{ijk})}{\Gamma(N'_{ijk})} \quad (3.12)$$

となる．これは，BD (Bayesian Dirichlet) Metric と呼ばれる．

式 (3.12) では， $\Gamma(i)$ は， i が1以上の自然数のとき， $\Gamma(i) = (i-1)!$ として計算される．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． N'_{ij} と N'_{ijk} はハイパーパラメータであり，次式の関係で与えられる．

$$N'_{ij} = \sum_{k=1}^L N'_{ijk} \quad (3.13)$$

Cooper によって，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，次式と導かれている．式 (3.14) は，K2Metric [1] [2] [25] と呼ばれる．

$$K2Metric = \prod_{i=1}^N \prod_{j=1}^M \frac{(L-1)!}{(N_{ij} + L - 1)!} \prod_{k=1}^L N_{ijk}! \quad (3.14)$$

ここで， N_{ij} と N_{ijk} には次の関係がある．

$$N_{ij} = \sum_{k=1}^L N_{ijk} \quad (3.15)$$

K2Metric は，式 (3.12) より，ハイパーパラメータの $N'_{ijk} = 1$ をあてはめると，一致することが確認される．K2Metric は，BDMetric を特殊化したものであり，ネットワークのリンク数とノード数が大きい場合，BDMetric よりも計算量を大幅に減らすことができる [1] ．

本研究では, K2Metric を用いて, 予測分布を最大化するモデルを選択することにより, 最適なモデル選択を行うことを狙う.

3.3.3 K2 アルゴリズム

本研究では, K2Metric [1] [2] [25] をネットワークの評価指標として採用し, K2 アルゴリズム [1] [25] を用いてネットワークのグラフ構造を決定する.

探索アルゴリズムをに, アルゴリズムを以下に示す (図 3.6).

1. $i = 1$
2. 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.
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. $j = i + 1, \dots, N$ 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 2 へ戻る.

3.4 確率推論

ベイジアンネットワークの確率推論 [33] 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. 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 の条件付き確率 $P(x_i|e)$ を求める.

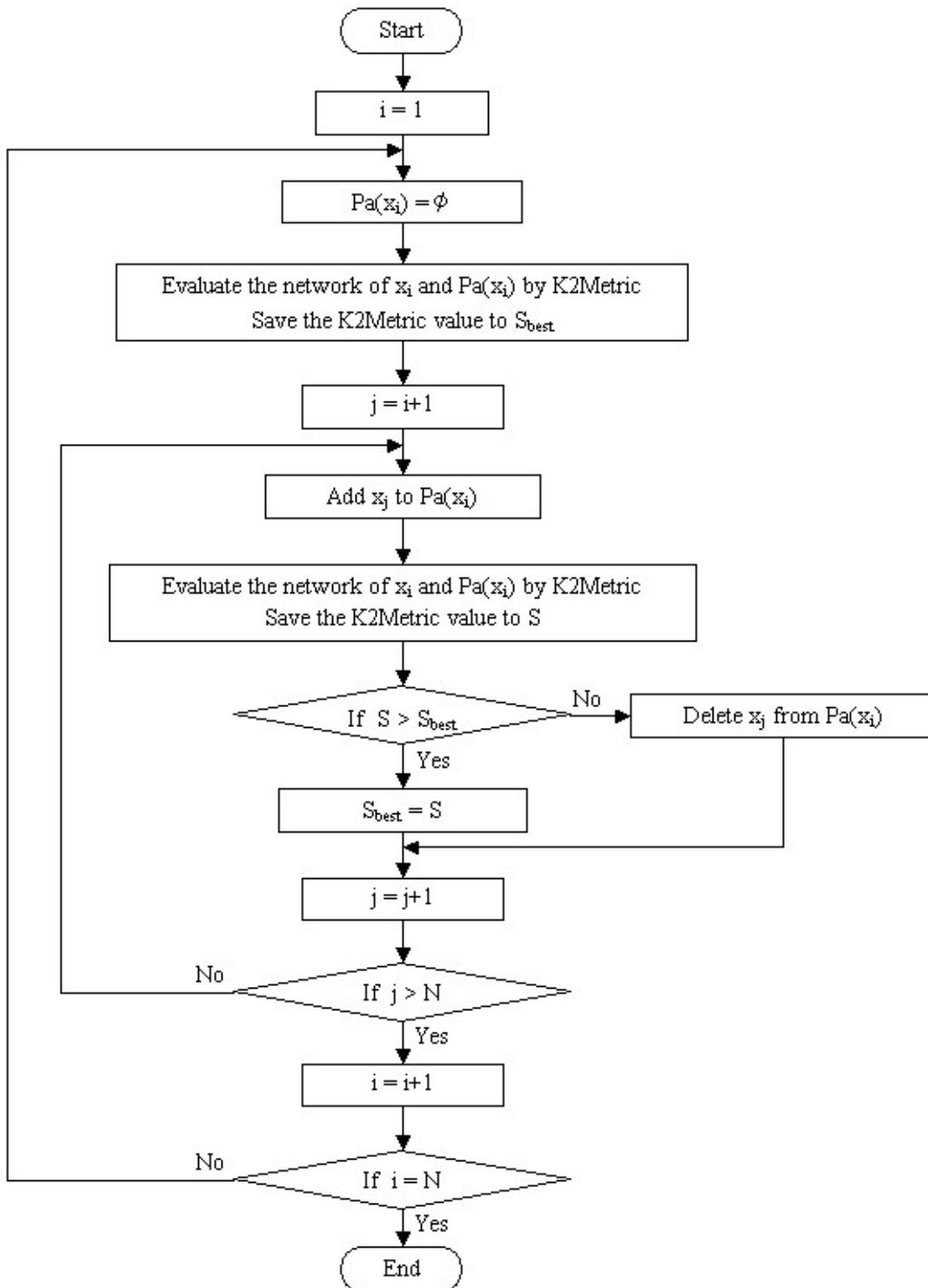


Fig. 3.6: グラフ構造決定アルゴリズム

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、例えば確率変数 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。

$$P(x_i = X^l|e) = \frac{\sum_{j=1, j \neq i}^N \sum_{x_j=X^1}^{X^L} P(x_1, \dots, x_i = X^l, \dots, x_N, e)}{\sum_{j=1}^N \sum_{x_j=X^1}^{X^L} P(x_1, \dots, x_N, e)} \quad (3.16)$$

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, \dots, X^L について総和を取ることを意味する。

第4章 提案手法1

4.1 概要

提案手法1では、株価リターンノードのみでベイジアンネットワークを構築し、これを用いて将来の株価リターンを予測する。ただし、ベイジアンネットワークではノードにおかれる確率変数は連続値をとることができないので、クラスタリング手法などを用いて株価リターンを離散化する操作が必要である。そこで、本章では、等分割クラスタリングとウォード法を用いて比較する。

提案手法1のアルゴリズムを整理し直すと以下ようになる。

1. 株価リターンを算出する。
2. 株価リターンを離散化する。
3. 離散化された過去の株価リターンからベイジアンネットワーク B を決定する。
4. B を用いて株価リターンを予測する。

4.2 アルゴリズム

4.2.1 株価リターンの算出

本研究では、株価指数の終値の日次データから求めた株価リターンを扱う。定義式を以下に示す。

$$r_t = (\ln P_t - \ln P_{t-1}) \times 100 \quad (4.1)$$

ここで、 P_t は t 期の株価終値を、 r_t は t 期の株価リターンを示す。

4.2.2 株価リターンの離散化

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できない。しかし、株価リターンは連続値なので、それを何らかの方法で離散化する必要がある。各ノ-

ドは確率変数として離散化された株価リターンを持つので、3.3.2節で述べた各ノードの状態総数は株価の離散値総数と等しい。そこで、ノードの状態総数に当たる離散化した株価リターンの離散値総数を L 、離散値を r^l とする。

離散値の集合は以下のように表される。

$$\{r^1, r^2, \dots, r^L\} \quad (4.2)$$

離散化のために、本研究では等分割クラスタリングを用いた方法とワード法を用いた方法を比較する。

クラスタリング法 [50] [51] [59] により離散化されたクラスタを C_l 、その重心を c_l とする。離散値 r^l には、各クラスタの重心 c_l をとることとする。つまり、

$$\{r^1, r^2, \dots, r^L\} = \{c_1, c_2, \dots, c_L\} \quad (4.3)$$

1. 等分割クラスタリングによる離散化

等分割クラスタリングでは、過去の株価リターンの頻度分布図において、それぞれのクラスタに含まれるイベント数がほぼ均等になるように複数のクラスタに等分割する。

2. ワード法による離散化

クラスタを C_i 、 C_i の重心を c_i 、データを z とする。ワード法 [83] [84] [85] [86] [87] では、各対象からその対象を含むクラスタの重心までの距離の二乗の総和を最小化する。ワード法の評価式は次式で表される。

$$D(C_i, C_j) = E(C_i \cup C_j) - E(C_i) - E(C_j) \quad (4.4)$$

ただし

$$E(C_i) = \sum_{z \in C_i} d(z, c_i)^2 \quad (4.5)$$

ここで、 $d(z, c_i)$ は z と c_i のユークリッド距離を示す。

4.2.3 グラフ構造の決定

本研究では、3.3節のアルゴリズムより、3.3.2節の K2Metric をネットワークの評価指標として採用し、3.3.3節の K2 アルゴリズムを用いてネットワークのグラフ構造を決定する。

4.2.4 株価リターンの予測

クラスタリングにより離散化された過去の株価リターンを用いて予測用のベイジアンネットワークを決定する．本研究では K2 アルゴリズムを用いているので，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．株価では時系列に基づく全順序関係があるので，これに従ってネットワークを決定する．

離散化された過去の株価リターンからネットワークを決定し，これを B とする．これを用いて株価リターン $r_t = r^l$ となる確率を推定した結果を $P(r^l|B)$ と定義する． t 期の株価リターンとして， $P(r^l|B)$ が最大となる r^l を選択する．つまり，

$$r_t = \arg \max_{r^l} (P(r^l|B)) \quad (4.6)$$

4.3 日経平均株価への適用

最初の例題として日経平均株価をとる．式 (4.1) で変換した日次リターンを図 2.3 に，リターンの頻度分布を図 2.4 に示す．

ベイジアンネットワークを決定するために 1985 年 2 月 22 日から 2008 年 12 月 30 日までの 6000 日間の日次株価リターンを使用する．このベイジアンネットワークを用いて 2008 年 12 月 1 日から 2008 年 12 月 30 日の株価を予測し，相関係数 (Correlation Coefficient-CC)，平均二乗誤差 (Root Mean Square Error-Ave.Error)，最大誤差 (Maximum Error-Max.Error)，最小誤差 (Minimum Error-Min.Error) を用いて，実際の株価に対する精度を比較する．実際の株価リターンを r_t ，予測の株価リターンを r'_t とすると，相関係数，平均二乗誤差，最大誤差，最小誤差は以下のように定義される．

$$\text{Correlation Coefficient(CC)} = \frac{\sum_{t=1}^n (r_t - \bar{r})(r'_t - \bar{r}')}{\sqrt{\sum_{t=1}^n (r_t - \bar{r})^2} \cdot \sqrt{\sum_{t=1}^n (r'_t - \bar{r}')^2}} \quad (4.7)$$

$$\text{Average Error(Ave.Error)} = \sqrt{\frac{1}{n} \sum_{t=1}^n (r_t - r'_t)^2} \quad (4.8)$$

$$\text{Maximum Error(Max.Error)} = \arg \max_t |r_t - r'_t| \quad (4.9)$$

$$\text{Minimum Error(Min.Error)} = \arg \min_t |r_t - r'_t| \quad (4.10)$$

ここで， n はデータの組の総数を， \bar{r}, \bar{r}' はそれぞれ r_t, r'_t の平均値を示す．

Table. 4.1: 日経平均株価リターンの分類区間 (等分割クラスタリング)

Cluster	$(C_l)_{\min}, (C_l)_{\max}$	Size	$c_l(r^l)$
C_1	[-16.14%, -1.12%)	980	-2.18%
C_2	[-1.12%, -0.42%)	980	-0.73%
C_3	[-0.42%, -0.00%]	980	-0.20%
C_4	(+0.00%, 0.44%]	1020	0.22%
C_5	(0.44%, 1.07%]	1020	0.72%
C_6	(1.07%, 13.23%]	1020	2.04%

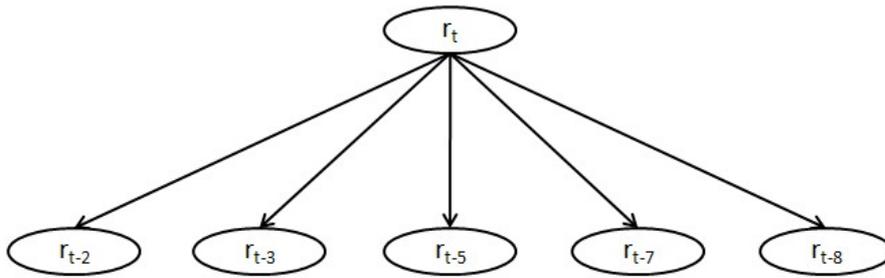


Fig. 4.1: 日経平均株価によって決定されたベイジアンネットワーク (等分割クラスタリング)

Table. 4.2: 日経平均株価リターンの分類区間 (ワード法)

Cluster	$(C_l)_{\min}, (C_l)_{\max}$	Size	$c_l(r^l)$
C_1	[-16.14%, -3.03%)	150	-4.32%
C_2	[-3.03%, -0.85%)	1123	-1.59%
C_3	[-0.85%, -0.00%]	1667	-0.37%
C_4	(+0.00%, 1.23%]	2207	0.52%
C_5	(1.23%, 3.80%]	796	1.98%
C_6	(3.80%, 13.23%]	57	5.47%

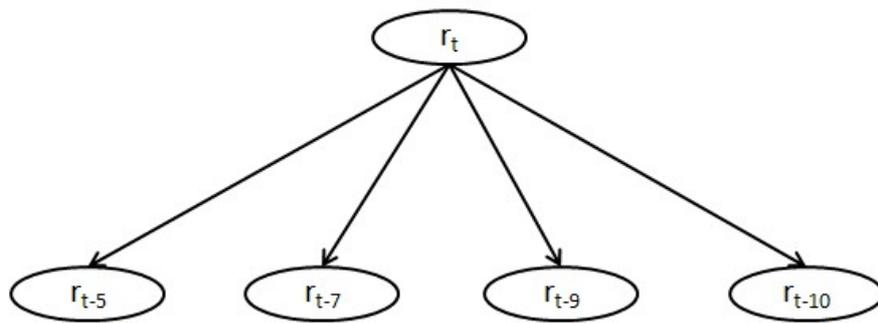


Fig. 4.2: 日経平均株価によって決定されたベイジアンネットワーク (ワード法)

4.3.1 離散化法によるネットワークの比較

クラスタ数を6個として、4.2節で説明した等分割クラスタリングで離散化した離散値を確率変数としてネットワークを構成する。クラスタ数を6個として等分割クラスタリングによって得られた離散値を表4.1に示す。表中で、 $(C_l)_{\min}$ と $(C_l)_{\max}$ はそのクラスタに含まれるデータ点の株価リターンの最小値と最大値を示す。Size と $c_l(r^l)$ は、そのクラスタに含まれるデータ点総数とそれらの値の平均値を示す。データから決定されたベイジアンネットワークを図4.1に示す。この場合、株価 r_t は2日前の株価 r_{t-2} 、3日前の株価 r_{t-3} 、5日前の株価 r_{t-5} 、7日前の株価 r_{t-7} 、8日前の株価 r_{t-8} と関連づけられていることがわかる。

次に、クラスタ数を6個として、4.2節で説明したワード法で離散化してネットワークを構成する。ワード法で株価リターンを離散化した結果を表4.2に示す。この離散化を用いて決定したネットワークを図4.2に示す。この場合、株価 r_t は5日前の株価 r_{t-5} 、7日前の株価 r_{t-7} 、9日前の株価 r_{t-9} 、10日前の株価 r_{t-10} と関連づけられており、等分割クラスタリングの場合と異なっていることがわかる。

4.3.2 離散化法の予測精度への影響

予測結果を実際の株価と比較した結果を表4.3に示す。ここで、Algorithm1(Uniform) と Algorithm1(Ward) は、それぞれ等分割クラスタリングによる離散化を用いた結果とワード法によるクラスタリングを用いた結果を示す。また、AR, MA, ARMA, ARCH は、それぞれ AR モデル, MA モデル, ARMA モデル, ARCH モデルによる結果である。それぞれにおいて、1000回シミュレーションを行い、その予測平均値を示す。比較に用いたモデルのパラメータは、2.2.1節で述べた方法により決定している。

Table. 4.3: 提案手法1と時系列分析の予測精度の比較(日経平均株価)

	Max.Error	Min.Error	Ave.Error	CC
AR(2)	6.4808	0.0056	2.7452	0.6928
MA(2)	6.4808	0.0399	2.7345	0.6942
ARMA(2,2)	6.6313	0.2826	2.7751	0.6840
ARCH(2,9)	6.5119	0.1069	2.7331	0.6974
Algorithm1(Uniform)	5.2753	0.2227	2.5362	0.7479
Algorithm1(Ward)	4.2383	0.2442	2.0727	0.8371

表 4.3 から, Algorithm1(Uniform) は AR, MA, ARMA, ARCH モデルより相関係数は高く, 平均二乗誤差はほとんど同じであることがわかる. また, Algorithm1(Ward) の予測結果の相関係数は Algorithm1 (Uniform) や AR モデルなどよりも高くなっており, Algorithm1(Ward) の平均二乗誤差は AR モデルなどよりも小さくなっていることがわかる. このことから, 離散化法を変更することで, 提案手法の精度が改善できることがわかる.

4.3.3 離散値数の予測精度への影響

まず, 株価リターンだけをノードとしてネットワークを決定する. このとき, 株価リターンの離散値数によって結果が異なるので, 離散値総数 $L = 2, 3, \dots, 10$ についてベイジアンネットワークを決定し, AIC(式 (4.11)) が最も小さな値を示すものを用いる. ここで, $\bar{\sigma}^2$ は残差標本の分散であり, T は残差の総数である. $\bar{\sigma}^2$ は式 (4.12) で求められる. 異なる離散値総数 L に対して決定されたネットワークの AIC を表 4.4 と図 4.3 に示す. これより, 離散値総数 $L = 6$ のとき AIC が最も小さいことがわかる. このときの具体的な離散値やクラスタのパラメータを表 4.2 に示す. また, 離散値総数 $L = 2 \sim 5, 7 \sim 10$ のときの具体的な離散値やクラスタのパラメータを表 4.5 から表 4.12 に示す.

$$AIC = \ln \bar{\sigma}^2 + \frac{2L}{T} \quad (4.11)$$

$$\bar{\sigma}^2 = \frac{1}{T - N} \sum_{t=N+1}^T (r_t - \bar{r}_t)^2 \quad (4.12)$$

Table. 4.4: 日経平均株価の株価リターンにおける離散値数と AIC の関係

Discrete number L	AIC
2	2.0442
3	1.7830
4	1.6066
5	1.5193
6	1.4597
7	1.8168
8	1.7478
9	1.8126
10	1.7729

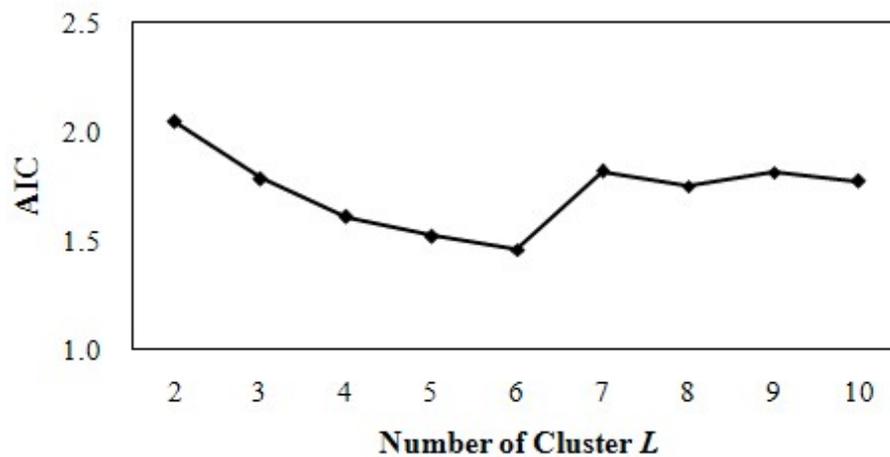


Fig. 4.3: 離散値数と AIC の関係 (ワード法)

Table. 4.5: 日経平均株価リターンの分類区間 (ワード法, $L=2$)

Cluster	$(C_l)_{\min}, (C_l)_{\max}$	Size	$c_l(r^l)$
C_1	[-16.138%, -0.730%]	1440	-1.77%
C_2	(-0.730%, 13.235%]	4560	0.55%

Table. 4.6: 日経平均株価リターンの分類区間 (ウォード法, L=3)

Cluster	$(C_l)_{\min}, (C_l)_{\max}$	Size	$c_l(r^l)$
C_1	[-16.138%, -0.730%]	1440	-1.77%
C_2	(-0.730%, 0.947%]	3415	0.10%
C_3	(0.947%, 13.235%]	1145	1.92%

Table. 4.7: 日経平均株価リターンの分類区間 (ウォード法, L=4)

Cluster	$(C_l)_{\min}, (C_l)_{\max}$	Size	$c_l(r^l)$
C_1	[-16.138%, -3.000%)	149	-4.30%
C_2	[-3.000%, -0.730%)	1291	-1.48%
C_3	(-0.730%, 0.947%]	3415	0.10%
C_4	(0.947%, 13.235%]	1145	1.92%

Table. 4.8: 日経平均株価リターンの分類区間 (ウォード法, L=5)

Cluster	$(C_l)_{\min}, (C_l)_{\max}$	Size	$c_l(r^l)$
C_1	[-16.138%, -3.000%)	149	-4.30%
C_2	[-3.000%, -0.730%)	1291	-1.48%
C_3	[-0.730%, 0.947%]	3415	0.10%
C_4	(0.947%, 3.800%]	1089	1.73%
C_5	(3.800%, 13.235%]	56	5.45%

Table. 4.9: 日経平均株価リターンの分類区間 (ウォード法, L=7)

Cluster	$(C_l)_{\min}, (C_l)_{\max}$	Size	$c_l(r^l)$
C_1	[-16.138%, -3.000%)	149	-4.30%
C_2	[-3.000%, -0.730%)	1291	-1.48%
C_3	[-0.730%, -0.065%]	1329	-0.37%
C_4	(-0.065%, 0.947%]	2086	0.39%
C_5	(0.947%, 1.871%]	740	1.36%
C_6	(1.871%, 3.800%]	349	2.54%
C_7	(3.800%, 13.235%]	56	5.45%

Table. 4.10: 日経平均株価リターンの分類区間 (ウォード法, L=8)

Cluster	$(C_l)_{\min}, (C_l)_{\max}$	Size	$c_l(r^l)$
C_1	[-16.138%, -3.000%)	149	-4.30%
C_2	[-3.000%, -1.810%)	357	-2.28%
C_3	[-1.810%, -0.730%)	934	-1.17%
C_4	[-0.730%, -0.065%]	1329	-0.37%
C_5	(-0.065%, 0.947%]	2086	0.39%
C_6	(0.947%, 1.871%]	740	1.36%
C_7	(1.871%, 3.800%]	349	2.54%
C_8	(3.800%, 13.235%]	56	5.45%

Table. 4.11: 日経平均株価リターンの分類区間 (ウォード法, L=9)

Cluster	$(C_l)_{\min}, (C_l)_{\max}$	Size	$c_l(r^l)$
C_1	[-16.138%, -9.849%)	5	-11.66%
C_2	[-9.849%, -3.000%)	145	-4.05%
C_3	[-3.000%, -1.810%)	357	-2.28%
C_4	[-1.810%, -0.730%)	934	-1.17%
C_5	[-0.730%, -0.065%]	1329	-0.37%
C_6	(-0.065%, 0.947%]	2086	0.39%
C_7	(0.947%, 1.871%]	740	1.36%
C_8	(1.871%, 3.800%]	349	2.54%
C_9	(3.800%, 13.235%]	56	5.45%

Table. 4.12: 日経平均株価リターンの分類区間 (ウォード法, L=10)

Cluster	$(C_l)_{\min}, (C_l)_{\max}$	Size	$c_l(r^l)$
C_1	[-16.138%, -9.849%)	5	-11.66%
C_2	[-9.849%, -3.000%)	144	-4.05%
C_3	[-3.000%, -1.810%)	357	-2.28%
C_4	[-1.810%, -0.730%)	934	-1.17%
C_5	[-0.730%, -0.065%]	1329	-0.37%
C_6	(-0.065%, 0.947%]	2086	0.39%
C_7	(0.947%, 1.871%]	740	1.36%
C_8	(1.871%, 3.800%]	349	2.54%
C_9	(3.800%, 6.212%]	47	4.77%
C_{10}	(6.212%, 13.235%]	9	9.02%

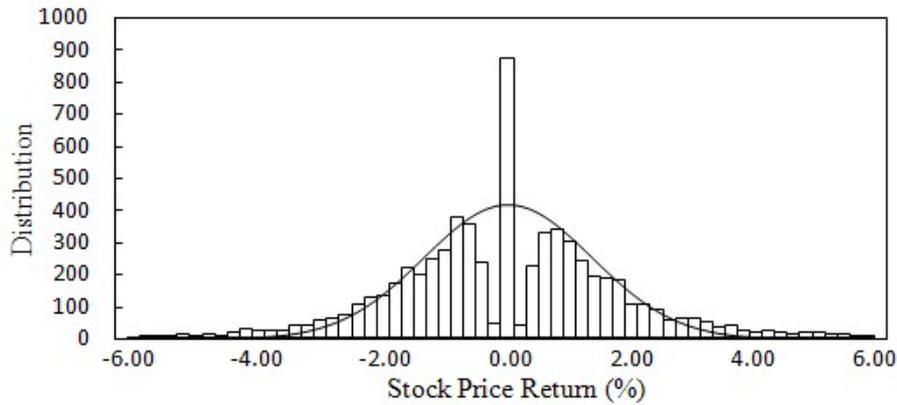


Fig. 4.4: トヨタ自動車株価の株価リターンの頻度分布図

Table. 4.13: トヨタ自動車株価リターンの分類区間 (等分割クラスタリング)

Cluster	$(C_l)_{\min}, (C_l)_{\max}$	Size	$c_l(r^l)$
C_1	$[-21.146\%, -1.702\%]$	1343	-2.25%
C_2	$[-1.702\%, -0.000\%]$	1343	-0.61%
C_3	$[-0.000\%, -0.000\%]$	720	0.00%
C_4	$[+0.00\%, 1.153\%]$	1297	0.66%
C_5	$(1.153\%, 16.264\%]$	1297	2.51%

4.4 トヨタ自動車の株価への適用

次の例題として、トヨタ自動車の株価データを取り上げる。トヨタ自動車株価の日次リターンの頻度分布を図 4.4 に示す。この図より、分布が正規分布からかなり外れており、多峰性をもつことを示している。特に、図 4.4 からわかるように株価リターン 0% 付近にかなり多くのデータが分布していることがわかる。

ベイジアンネットワークを決定するために 1985 年 2 月 22 日から 2008 年 12 月 30 日までの 6000 日間の日次株価リターンを使用する。このベイジアンネットワークを用いて 2008 年 12 月 1 日から 2008 年 12 月 30 日の株価を予測し、相関係数、平均二乗誤差、最大誤差、最小誤差を用いて、実測の株価と比較する。

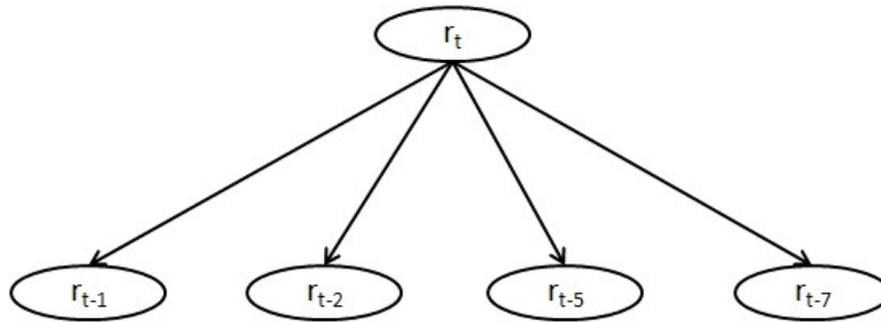


Fig. 4.5: トヨタ自動車株価によって決定されたベイジアンネットワーク (等分割クラスタリング)

Table. 4.14: トヨタ自動車株価リターンの分類区間 (ワード法)

Cluster	$(C_l)_{\min}, (C_l)_{\max}$	Size	$c_l(r^l)$
C_1	$[-21.146\%, -3.900\%)$	133	-5.55%
C_2	$[-3.900\%, -1.285\%)$	1053	-2.10%
C_3	$[-1.285\%, 0.000\%]$	2221	-0.47%
C_4	$(0.000\%, 2.530\%]$	2152	1.08%
C_5	$(2.530\%, 16.264\%]$	441	4.12%

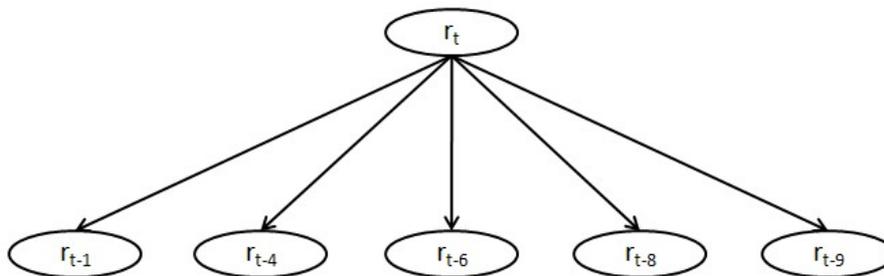


Fig. 4.6: トヨタ自動車株価によって決定されたベイジアンネットワーク (ワード法)

4.4.1 離散化法によるネットワークの比較

先に述べたように、トヨタ自動車の株価リターンでは、株価リターン 0% 付近にかなり多くのデータが分布している。そこで、クラスタ数を 5 個として、4.2 節で説明した等分割クラスタリングで離散化してネットワークを構成する。

等分割クラスタリングによる離散値を表 4.13 に示す。表中で、 $(C_l)_{\min}$ と $(C_l)_{\max}$ はそのクラスタに含まれるデータ点の株価リターンの最小値と最大値を示す。Size と $c_l(r^l)$ は、そのクラスタに含まれるデータ点総数とそれらの値の平均値を示す。データから決定されたベイジアンネットワークを図 4.5 に示す。この場合、株価 r_t は 1 日前の株価 r_{t-1} 、2 日前の株価 r_{t-2} 、5 日前の株価 r_{t-5} 、7 日前の株価 r_{t-7} と関連づけられていることがわかる。

次に、クラスタ数を 5 個として、4.2 節で説明したウォード法で離散化してネットワークを構成する。

ウォード法で株価リターンを離散化した結果を表 4.14 に示す。この離散化を用いて決定したネットワークを図 4.6 に示す。この場合、株価 r_t は 1 日前の株価 r_{t-1} 、4 日前の株価 r_{t-4} 、6 日前の株価 r_{t-6} 、8 日前の株価 r_{t-8} 、9 日前の株価 r_{t-9} と関連づけられている。このように、クラスタリング手法を変更することで、得られるベイジアンネットワークはかなり異なることがわかる。

4.4.2 離散化法の予測精度への影響

予測結果を実際の株価と比較した結果を表 4.15 に示す。ここで、Algorithm1(Uniform) と Algorithm1(Ward) は、それぞれ等分割クラスタリングによる離散化を用いた結果とウォード法によるクラスタリングを用いた結果を示す。また、AR, MA, ARMA, ARCH は、それぞれ AR モデル, MA モデル, ARMA モデル, ARCH モデルによる結果である。それぞれにおいて、1000 回シミュレーションを行い、その予測平均値を示す。比較に用いたモデルのパラメータは、2.2.1 節で述べた方法により決定する。

表 4.15 から、Algorithm1(Uniform), AR, MA, ARMA, ARCH モデルの予測結果では、相関係数が 0.5 より小さく、平均二乗誤差は 4 を超えており、日経平均株価での結果に比べるとかなり精度が低くなっている。これに対して、Algorithm1(Ward) では、相関係数は 0.6469 とほかの方法よりかなり大きい。また、平均二乗誤差は 3.1942 と他の方法よりかなり小さい値を示している。Algorithm1(Ward) の結果は日経平均株価での結果に比べると精度が低くなっているが、他の手法よりは良い結果を示している。

Table. 4.15: 提案手法1と時系列分析の予測精度の比較(トヨタ自動車株価)

	Max.Error	Min.Error	Ave.Error	CC
AR(9)	10.3189	0.2119	4.0866	0.4689
MA(6)	10.2584	0.0242	4.1016	0.4679
ARMA(9,6)	10.6026	0.0539	4.1434	0.4457
ARCH(9,9)	10.2630	0.1632	4.0669	0.4692
Algorithm1(Uniform)	13.1577	0.2871	4.5762	0.4788
Algorithm1(Ward)	6.8142	0.0491	3.1942	0.6469

Table. 4.16: トヨタ自動車株価リターンにおける離散値数とAICの関係

Discrete number L	AIC
2	2.8452
3	2.5892
4	2.3913
5	2.3243
6	2.5309
7	2.7240
8	2.4935
9	2.5941
10	2.5994

4.4.3 離散値数の予測精度への影響

日経平均の場合と同様に, 時系列分析手法のモデルパラメータ L は, これらを 2 から 10 までとってシミュレーションを行い, AIC(式 (4.11)) が最も小さな値を示す場合のモデルパラメータを用いる. 表 4.16 と図 4.7 により, 離散値総数 $L = 5$ のとき AIC が最も小さいことがわかる. このときの具体的な離散値やクラスタのパラメータは表 4.14 の通りである. また, 離散値総数 $L = 2 \sim 4, 6 \sim 10$ のときの具体的な離散値やクラスタのパラメータを表 4.17 から表 4.24 に示す.

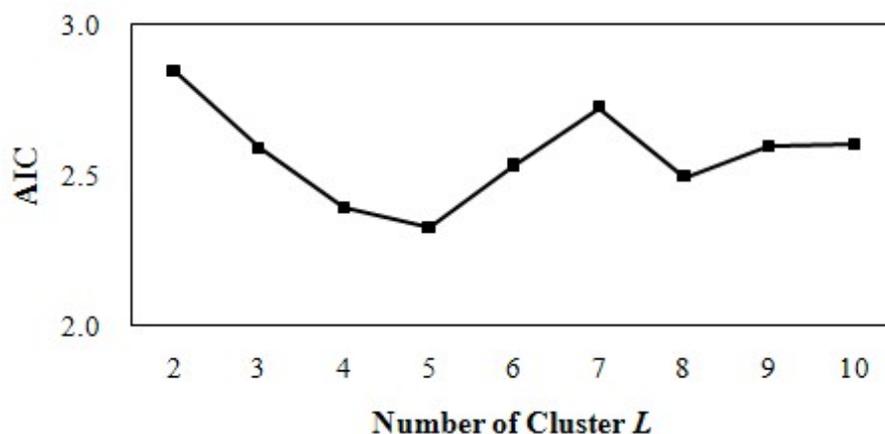


Fig. 4.7: 離散値数の精度への影響 (ウォード法)

Table. 4.17: トヨタ自動車株価リターンの分類区間 (ウォード法, L=2)

Cluster	$(C_l)_{\min}, (C_l)_{\max}$	Size	$c_l(r^l)$
C_1	[-21.146%, -1.285%]	1186	-2.48%
C_2	(-1.285%, 16.264%]	4814	0.64%

Table. 4.18: トヨタ自動車株価リターンの分類区間 (ウォード法, L=3)

Cluster	$(C_l)_{\min}, (C_l)_{\max}$	Size	$c_l(r^l)$
C_1	[-21.146%, -1.285%]	1186	-2.48%
C_2	(-1.285%, 2.530%]	4373	0.28%
C_3	(2.530%, 13.235%]	441	4.12%

Table. 4.19: トヨタ自動車株価リターンの分類区間 (ウォード法, L=4)

Cluster	$(C_l)_{\min}, (C_l)_{\max}$	Size	$c_l(r^l)$
C_1	[-21.146%, -1.285%]	1186	-2.48%
C_2	(-1.285%, 0.000%]	2221	0.28%
C_3	(0.000%, 0.947%]	2152	1.07%
C_4	(2.530%, 13.235%]	441	4.12%

Table. 4.20: トヨタ自動車株価リターンの分類区間 (ウォード法, L=6)

Cluster	$(C_l)_{\min}, (C_l)_{\max}$	Size	$c_l(r^l)$
C_1	[-21.146%, -3.900%)	133	-5.55%
C_2	[-3.900%, -1.285%)	1053	-2.10%
C_3	[-1.285%, 0.000%)	2221	-0.47%
C_4	(0.000%, 2.530%]	2152	1.07%
C_5	(2.530%, 4.580%]	318	3.23%
C_6	(4.580%, 16.264%]	123	6.42%

Table. 4.21: トヨタ自動車株価リターンの分類区間 (ウォード法, L=7)

Cluster	$(C_l)_{\min}, (C_l)_{\max}$	Size	$c_l(r^l)$
C_1	[-21.146%, -3.900%)	133	-5.55%
C_2	[-3.900%, -1.285%)	1053	-2.10%
C_3	[-1.285%, 0.000%)	2221	-0.47%
C_4	(0.000%, 0.986%]	2086	0.59%
C_5	(0.986%, 2.530%]	740	1.57%
C_6	(2.530%, 4.580%]	318	3.23%
C_7	(4.580%, 16.264%]	123	6.42%

Table. 4.22: トヨタ自動車株価リターンの分類区間 (ウォード法, L=8)

Cluster	$(C_l)_{\min}, (C_l)_{\max}$	Size	$c_l(r^l)$
C_1	[-21.146%, -9.541%)	11	-11.44%
C_2	[-9.541%, -3.900%)	133	-5.01%
C_3	[-3.900%, -1.285%)	1053	-2.10%
C_4	[-1.285%, 0.000%)	2221	-0.47%
C_5	(0.000%, 0.986%]	2086	0.59%
C_6	(0.986%, 2.530%]	740	1.57%
C_7	(2.530%, 4.580%]	318	3.23%
C_8	(4.580%, 16.264%]	123	6.42%

Table. 4.23: トヨタ自動車株価リターンの分類区間 (ウォード法, L=9)

Cluster	$(C_l)_{\min}, (C_l)_{\max}$	Size	$c_l(r^l)$
C_1	[-21.146%, -9.541%)	11	-11.44%
C_2	[-9.541%, -3.900%)	133	-5.01%
C_3	[-3.900%, -1.285%)	1053	-2.10%
C_4	[-1.285%, 0.000%]	2221	-0.47%
C_5	(0.000%, 0.986%]	2086	0.59%
C_6	(0.986%, 2.530%]	740	1.57%
C_7	(2.530%, 4.580%]	318	3.23%
C_8	(4.580%, 8.110%]	107	5.76%
C_9	(8.110%, 16.264%]	16	10.80%

Table. 4.24: トヨタ自動車株価リターンの分類区間 (ウォード法, L=10)

Cluster	$(C_l)_{\min}, (C_l)_{\max}$	Size	$c_l(r^l)$
C_1	[-21.146%, -9.541%)	11	-11.44%
C_2	[-9.541%, -3.900%)	133	-5.01%
C_3	[-3.900%, -1.285%)	288	-2.97%
C_4	[-3.900%, -1.285%)	765	-1.77%
C_5	[-1.285%, 0.000%]	2221	-0.47%
C_6	(0.000%, 0.986%]	2086	0.59%
C_7	(0.986%, 2.530%]	740	1.57%
C_8	(2.530%, 4.580%]	318	3.23%
C_9	(4.580%, 8.110%]	107	5.76%
C_{10}	(8.110%, 16.264%]	16	10.80%

第5章 提案手法2

5.1 概要

提案手法2では、提案手法1の予測精度向上を目的として、株価リターンだけでなくその予測誤差も用いてベイジアンネットワークを構築する。提案手法1で構成したベイジアンネットワークを用いて、すでに測定されている過去の株価リターンを予測すれば、その予測値と実際値から誤差を評価できる。提案手法2では、この誤差値のノードを定義し、株価リターンのノードとともに用いて、株価リターン予測用のベイジアンネットワークを構築する。ところで、提案手法1ではネットワーク構造を仮定しなかったにもかかわらず、結果としてNaiveベイズ構造のネットワークが得られた。そこで、提案手法2では、計算コストを下げるためにNaiveベイズ構造を仮定してネットワーク構造を探索する。

提案手法2のアルゴリズムを整理し直すと以下のようになる。

1. 株価リターンを算出する。
2. 株価リターンを離散化する。
3. 離散化された過去の株価リターンからベイジアンネットワーク B を決定する。
4. B を用いて株価リターンを予測する。
5. 予測値と実測値の誤差を求める。
6. 予測誤差を離散化する。
7. 離散化された予測誤差を加えてベイジアンネットワーク B' を決定する。
8. B' を用いて株価リターンを再予測する。

5.2 アルゴリズム

5.2.1 株価リターンの算出

式 (4.1) より株価指数の終値の日次データから求めた株価リターンを求める。

5.2.2 株価リターンの離散化

4.2.2 節に示したように、株価リターンの分布を離散化する。そのとき、クラスタリング法により離散化されたクラスタを C_l 、その重心を c_l とする。離散値 r^l には、式 (4.3) のように各クラスタの重心 c_l をとる。

5.2.3 グラフ構造 B の決定

4.2.3 節と同じように、3.3 節のアルゴリズムより、3.3.2 節の K2Metric をネットワークの評価指標として採用し、3.3.3 節の K2 アルゴリズムを用いてネットワークのグラフ構造を決定する。

5.2.4 株価リターンの予測

ネットワーク B を与える時、式 (4.6) を用いて予測値を求める。

上記のステップ 5.2.1 節 ~ 5.2.4 節は、提案手法 1 とまったく同じとなる。

5.2.5 予測誤差の算出

すでに測定された過去の t 期において、式 (4.6) より求めた株価リターンの予測値を r_t 、実測値を \bar{r}_t とする。このとき、予測誤差 u_t は次式で与えられる。

$$u_t = r_t - \bar{r}_t \quad (5.1)$$

5.2.6 予測誤差の離散化

予測誤差は連続値なので、ベイジアンネットワークで適用できるように 4.2 節で示したワード法によって離散化する。



Fig. 5.1: 株価リターンと予測誤差の全順序関係

予測誤差離散値を u^l , 離散値総数を L とすると , 予測誤差離散値の集合は以下のように表される .

$$\{u^1, u^2, \dots, u^L\} \quad (5.2)$$

離散値総数 L は次のようにして定める . 異なる L について AIC を計算し , AIC が最小値を示す L を採用する . AIC は式 (4.11) と式 (4.12) で定義されたものを用いる .

予測誤差の離散値を求める方法を箇条書きにすると以下ようになる .

1. 営業日ごとの予測誤差を式 (5.1) より求める .
2. 横軸を予測誤差 , 縦軸を頻度 (その予測誤差を取った日数) として頻度分布図を描く .
3. 頻度分布図を異なる離散値総数 (クラスタ数) でクラスタリングする .
4. 式 (4.11) で示される AIC が最小となる離散値総数 L とそのときの離散値を採用する .

5.2.7 グラフ構造 B' の決定

過去の株価リターン離散値とその予測誤差の離散値を用いて , 図 5.1 に示す全順序関係に従って , 3.3 節のアルゴリズムによりネットワークを決定する .

5.2.8 株価リターンの再予測

決定されたネットワークを B' , これを用いて株価リターン $r_t = r^l$ となる確率を推定した結果を $P(r^l|B')$ とする . $P(r^l|B')$ が最大となる r^l を t 期の株価リターンとする . つまり ,

$$r_t = \arg \max_{r^l} (P(r^l|B')) \quad (5.3)$$

5.3 日経平均株価への適用

最初の例題として日経平均株価を取り上げる . 1985 年 2 月 22 日から 2008 年 12 月 30 日までの日次株価リターンを用いてベイジアンネットワークを決定し , それを用いて 2008 年

Table. 5.1: 日経平均株価予測誤差における離散値数と AIC の関係

Discrete number L	AIC
2	1.6659
3	1.4587
4	1.5764
5	1.5132
6	1.3203
7	1.4596
8	1.4599
9	1.4602
10	1.4606

12月1日から2008年12月30日の株価リターンを予測する。そして、提案手法2を時系列分析手法や提案手法1と比較する。既に述べたように、時系列分析手法のモデルパラメータ p, q は、これらを1から10までとってシミュレーションを行い、AICが最も小さな値を示す場合のモデルパラメータを用いる。

ベイジアンネットワークを用いた手法のうち提案手法1と提案手法2についても、1期前から10期前までの株価リターンや予測誤差のノードを用いてベイジアンネットワークを決定する。

5.3.1 予測精度の比較(1)

まず、株価リターンだけをノードとしてネットワークを決定する時、提案手法1の数値実験を参照する。

次に、株価リターンとその誤差の両方をノードとしてネットワークを決定する。株価リターンの離散値総数を $L = 6$ として固定した上で、誤差の離散値総数を $L = 2, 3, \dots, 10$ としてベイジアンネットワークを決定し、AIC(式(4.11))を計算した結果を表5.1に示す。これより、離散値総数 $L = 6$ のとき AIC が最も小さいことがわかる。このときの具体的な離散値やクラスタのパラメータを表5.6に示す。また、離散値総数 $L = 2 \sim 5, 7 \sim 10$ のときの具体的な離散値やクラスタのパラメータを表5.2から表5.5、表5.7から表5.10に示す。

次に、3.3節のアルゴリズムにより、株価リターンと予測誤差で決定されたネットワーク

Table. 5.2: 日経平均株価リターンの分類区間 (ウォード法, L=2)

Cluster	$(C_l)_{\min}, (C_l)_{\max}$	Size	$c_l(r^l)$
C_1	[-16.527%, -0.238%]	3111	-1.39%
C_2	(-0.238%, 17.538%]	2879	0.99%

Table. 5.3: 日経平均株価の予測誤差の分類区間 (ウォード法, L=3)

Cluster	$(C_l)_{\min}, (C_l)_{\max}$	Size	$c_l(r^l)$
C_1	[-16.527%, -0.238%]	3111	-1.39%
C_2	(-0.238%, 2.300%]	2533	0.62%
C_3	(2.300%, 17.538%]	346	3.70%

Table. 5.4: 日経平均株価の予測誤差の分類区間 (ウォード法, L=4)

Cluster	$(C_l)_{\min}, (C_l)_{\max}$	Size	$c_l(r^l)$
C_1	[-16.527%, -2.138%)	565	-3.37%
C_2	[-2.138%, -0.238%]	2546	-0.95%
C_3	(-0.238%, 2.300%]	2533	0.62%
C_4	(2.300%, 17.538%]	346	3.70%

Table. 5.5: 日経平均株価の予測誤差の分類区間 (ウォード法, L=5)

Cluster	$(C_l)_{\min}, (C_l)_{\max}$	Size	$c_l(r^l)$
C_1	[-16.527%, -4.910%)	59	-6.91%
C_2	[-4.910%, -2.138%)	506	-2.96%
C_3	[-2.138%, -0.238%]	2546	-0.95%
C_4	(-0.238%, 2.300%]	2533	0.62%
C_5	(2.300%, 17.538%]	346	3.70%

Table. 5.6: 日経平均株価の予測誤差の分類区間 (ワード法, L=6)

Cluster	$(C_l)_{\min}, (C_l)_{\max}$	Size	$c_l(r^l)$
C_1	[-16.527%, -4.910%]	59	-6.91%
C_2	[-4.910%, -2.138%]	506	-2.96%
C_3	[-2.138%, -0.238%]	2546	-0.95%
C_4	(-0.238%, 1.182%]	1951	0.32%
C_5	(1.182%, 2.300%]	582	1.63%
C_6	(2.300%, 17.538%]	346	3.70%

Table. 5.7: 日経平均株価の予測誤差の分類区間 (ワード法, L=7)

Cluster	$(C_l)_{\min}, (C_l)_{\max}$	Size	$c_l(r^l)$
C_1	[-16.527%, -4.910%]	59	-6.91%
C_2	[-4.910%, -2.138%]	506	-2.96%
C_3	[-2.138%, -1.040%]	956	-1.51%
C_4	[-1.040%, -0.238%]	1590	-0.61%
C_5	(-0.238%, 1.182%]	1951	0.32%
C_6	(1.182%, 2.300%]	582	1.63%
C_7	(2.300%, 17.538%]	346	3.70%

Table. 5.8: 日経平均株価の予測誤差の分類区間 (ワード法, L=8)

Cluster	$(C_l)_{\min}, (C_l)_{\max}$	Size	$c_l(r^l)$
C_1	[-16.527%, -4.910%]	59	-6.91%
C_2	[-4.910%, -2.138%]	506	-2.96%
C_3	[-2.138%, -1.040%]	956	-1.51%
C_4	[-1.040%, -0.238%]	1590	-0.61%
C_5	(-0.238%, 1.182%]	1951	0.32%
C_6	(1.182%, 2.300%]	582	1.63%
C_7	(2.300%, 5.188%]	309	3.29%
C_8	(5.188%, 17.538%]	37	7.11%

Table. 5.9: 日経平均株価の予測誤差の分類区間 (ウォード法, L=9)

Cluster	$(C_l)_{\min}, (C_l)_{\max}$	Size	$c_l(r^l)$
C_1	[-16.527%, -4.910%)	59	-6.91%
C_2	[-4.910%, -2.138%)	506	-2.96%
C_3	[-2.138%, -1.040%)	956	-1.51%
C_4	[-1.040%, -0.238%]	1590	-0.61%
C_5	(-0.238%, 0.327%]	1101	0.03%
C_6	(0.327%, 1.182%]	850	0.69%
C_7	(1.182%, 2.300%]	582	1.63%
C_8	(2.300%, 5.188%]	309	3.29%
C_9	(5.188%, 17.538%]	37	7.11%

Table. 5.10: 日経平均株価の予測誤差の分類区間 (ウォード法, L=10)

Cluster	$(C_l)_{\min}, (C_l)_{\max}$	Size	$c_l(r^l)$
C_1	[-16.527%, -9.589%)	7	-11.59%
C_2	[-9.589%, -4.910%)	52	-6.28%
C_3	[-4.910%, -2.138%)	506	-2.96%
C_4	[-2.138%, -1.040%)	956	-1.51%
C_5	[-1.040%, -0.238%]	1590	-0.61%
C_6	(-0.238%, 0.327%]	1101	0.03%
C_7	(0.327%, 1.182%]	850	0.69%
C_8	(1.182%, 2.300%]	582	1.63%
C_9	(2.300%, 5.188%]	309	3.29%
C_{10}	(5.188%, 17.538%]	37	7.11%

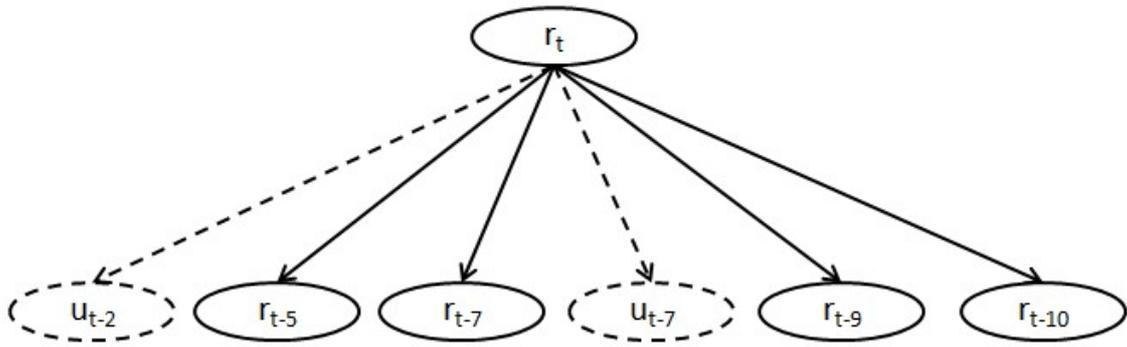


Fig. 5.2: 株価リターンと誤差で決定されたベイジアンネットワーク (日経平均株価)

Table. 5.11: 提案手法 2 と時系列分析法及び提案手法 1 の予測精度の比較 (日経平均株価)

	Max.Error	Min.Error	Ave.Error	CC
AR(2)	6.4665	0.0214	2.7477	0.6929
MA(2)	6.5148	0.0174	2.7538	0.6912
ARMA(2,2)	6.6085	0.2443	2.7696	0.6855
ARCH(2,9)	6.4525	0.0976	2.7440	0.6932
Algorithm1	4.2383	0.2442	2.0727	0.8371
Algorithm2	4.2383	0.1192	1.9331	0.8629

を図 5.2 に示す．この場合，株価リターン r_t は 5 日前の株価リターン r_{t-5} ，7 日前の株価リターン r_{t-7} ，9 日前の株価リターン r_{t-9} ，10 日前の株価リターン r_{t-10} ，2 日前の誤差 u_{t-2} ，7 日前の誤差 u_{t-7} と関連づけられていることがわかる．予測誤差ノードを取り入れると，異なるネットワークが導かれることがわかる．

求めたネットワークで予測を行い，時系列分析法や提案手法 1 と比較した結果を表 5.11 と図 5.3 に示す．ここで，AR,MA,ARMA,ARCH は，それぞれの時系列分析法による結果を，Algorithm1 は図 4.2 のネットワークによる予測結果を，Algorithm2 は図 5.2 のネットワークによる予測結果を示す．それぞれの数値は相関係数 (CC)，平均二乗誤差 (Ave.Error)，最大誤差 (Max.Error)，最小誤差 (Min.Error) を示す．

表 5.11 と図 5.3 から提案手法 2 は最も大きな相関係数を示しており，平均二乗誤差と最大誤差も最小となっていることがわかる．最小誤差はどの方法でも同程度であるが，平均二乗誤差では提案手法 2 は時系列分析法よりも 30% 程度，提案手法 1 よりも 6% 程度改善

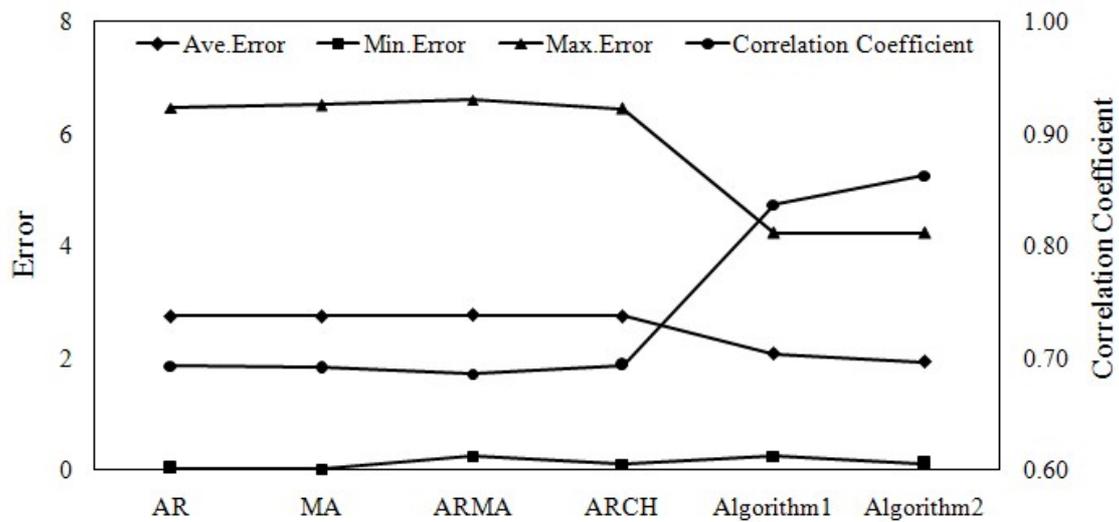


Fig. 5.3: 提案手法 2 と時系列分析法及び提案手法 1 の予測精度の比較 (日経平均株価)

されていることがわかる。最大誤差については、提案手法 2 と提案手法 1 の精度は同程度となっている。これは提案手法 2 と提案手法 1 のいずれにおいても、予測精度の改善とは予測確率の改善であって、予測値として示される株価リターンは表 5.6 における $c_l(r^l)$ の 6 種類でしかないことによると考えられる。

5.3.2 予測精度の比較 (2)

前節の実験結果では、予測期間は学習期間に含まれていた。そこで、学習期間に含まれない期間の株価リターンを予測することを考える。上記で示した手法を用いて、2009 年 1 月 1 日から 3 月 31 日における取引日での株価リターン予測を行う。

提案手法 2 を時系列分析法や提案手法 1 と比較した結果を表 5.12 と図 5.4 に示す。これらの中で提案手法 2 は最も大きな相関係数を示している。最小誤差はどの方法でも大きな違いはないけれども、平均二乗誤差と最大誤差では提案手法 2 が最も小さい値を示しており、提案手法 2 の平均二乗誤差は、時系列分析法よりも 6% 程度、最大誤差は 15% 程度改善していることがわかる。

5.3.3 計算時間

誤差ノードを用いる提案手法 2 は、誤差ノードを用いない提案手法 1 に比べて計算時間が増えると予想される。そこで、日経平均株価リターン予測について、提案手法 2 と提案手

Table. 5.12: 予測期間が異なる場合の予測値と実測値の比較 (日経平均株価)

	Max.Error	Min.Error	Ave.Error	CC
AR(2)	5.7648	0.0604	2.4923	0.9278
MA(2)	5.7746	0.0758	2.4953	0.9276
ARMA(2,2)	5.9389	0.0076	2.5150	0.9250
ARCH(2,9)	5.8342	0.0178	2.5116	0.9268
Algorithm1	6.2948	0.0057	2.9905	0.9099
Algorithm2	4.9016	0.0105	2.3467	0.9375

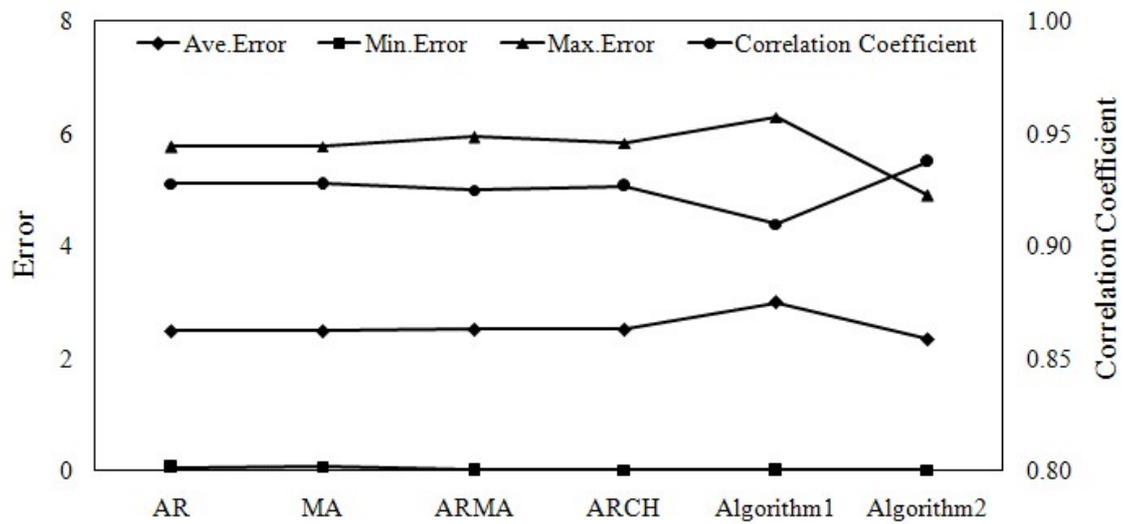


Fig. 5.4: 予測期間が異なる場合の予測値と実測値の比較 (日経平均株価)

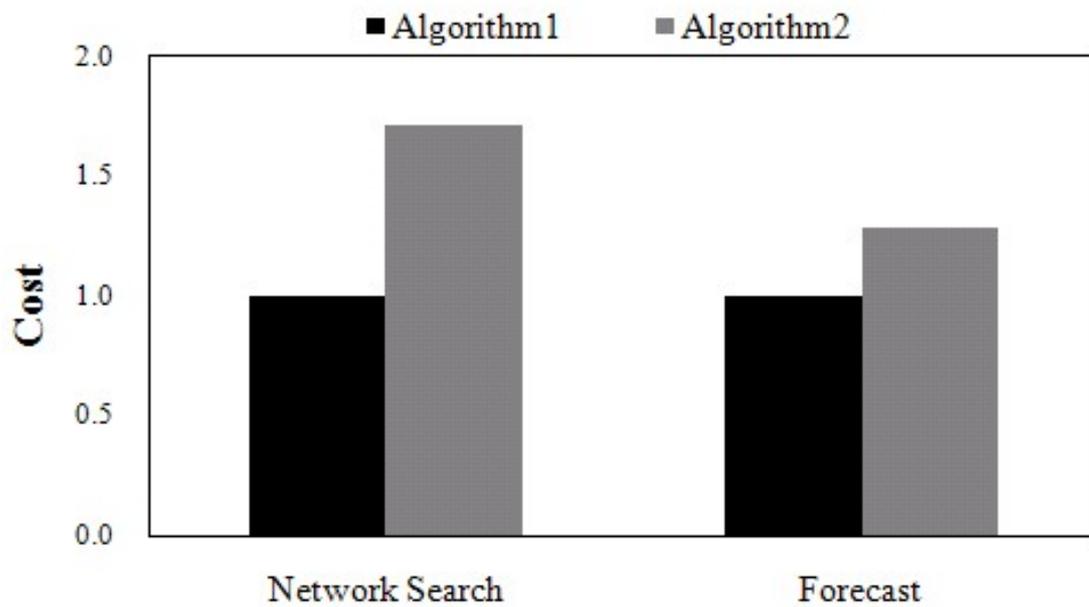


Fig. 5.5: 提案手法 1 と提案手法 2 の計算コストの比較

Table. 5.13: 提案手法 1 と提案手法 2 の計算時間の比較

	Network Search	Forecast
Algorithm1	529sec	0.691sec
Algorithm2	908sec	0.887sec

法 1 の計算時間を比較する．計算時間は，ベイジアンネットワークの学習にかかる時間と，決定されたベイジアンネットワークを用いて株価リターンを予測するためにかかる時間に分けて示す．

比較結果を表 5.13 に示す．これより，提案手法 1 ではネットワーク決定に 529 秒かかり，株価リターン予測に 0.691 秒かかることがわかる．一方，提案手法 2 では，それぞれ 908 秒と 0.887 秒である．このように，株価リターン予測時間はネットワーク決定時間よりかなり短いことがわかる．

そして，図 5.5 をみると，ネットワーク決定にかかる時間において，提案手法 2 は提案手法 1 のおよそ 1.7 倍となっている．計算時間は増大しているけれども，最初にネットワーク決定のために 1 回だけ計算すればよい．また，予測にかかる時間について検討すると提案手法 2 は提案手法 1 の 1.3 倍程度であり，ネットワーク決定にかかる時間に比べれば計算時間の増加は比較的小さいといえる．

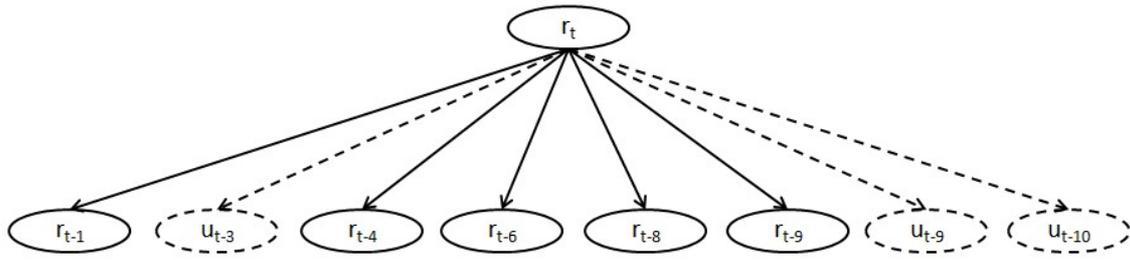


Fig. 5.6: トヨタ自動車株価の株価リターンと予測誤差で決定されたベイジアンネットワーク

5.4 トヨタ自動車株価への適用

次の例題として、トヨタ自動車の株価データを取り上げる。ベイジアンネットワークを決定するために1985年2月22日から2008年12月30日までの日次株価リターンを使用する。このベイジアンネットワークを用いて2008年12月1日から2008年12月30日の株価リターンを予測する。そして、提案手法2を時系列分析手法や提案手法1と比較する。日経平均株価の場合と同様に、時系列分析手法のモデルパラメータ p, q は、これらを1から10までとってシミュレーションを行い、AICが最も小さな値を示す場合のモデルパラメータを用いる。

また、ベイジアンネットワークを用いた手法のうち提案手法1と提案手法2についても、1期前から10期前までの株価リターンや予測誤差のノードを用いてベイジアンネットワークを決定する。

5.4.1 予測精度の比較(1)

株価リターンとその誤差の両方をノードとしてネットワークを決定する。株価リターンの離散値総数を $L = 5$ として固定した上で、誤差の離散値総数を $L = 2, 3, \dots, 10$ としてベイジアンネットワークを決定し、AIC(式(4.11))を計算した結果を表5.14に示す。これより、離散値総数 $L = 3$ のときAICが最も小さいことがわかる。このときの具体的な離散値やクラスタのパラメータを表5.16に示す。また、離散値総数 $L = 2, 4 \sim 10$ のときの具体的な離散値やクラスタのパラメータを表5.15および表5.17から表5.23に示す。

次に、3.3節のアルゴリズムにより、株価リターンと予測誤差で決定されたネットワークを図5.6に示す。この場合、株価リターン r_t は1日前の株価リターン r_{t-1} 、4日前の株価リターン r_{t-4} 、6日前の株価リターン r_{t-6} 、8日前の株価リターン r_{t-8} 、9日前の株価リター

Table. 5.14: トヨタ自動車株価予測誤差における離散値数と AIC の関係

Discrete number L	AIC
2	2.3225
3	2.0729
4	2.1618
5	2.4284
6	2.1506
7	2.3820
8	2.3823
9	2.6758
10	2.7555

Table. 5.15: トヨタ自動車株価の予測誤差の分類区間 (ワード法, $L=2$)

Cluster	$(C_l)_{\min}, (C_l)_{\max}$	Size	$c_l(r^l)$
C_1	[-25.262%, 0.139%]	3359	-1.55%
C_2	(0.139%, 19.981%]	2631	1.66%

Table. 5.16: トヨタ自動車株価の予測誤差の分類区間 (ワード法, $L=3$)

Cluster	$(C_l)_{\min}, (C_l)_{\max}$	Size	$c_l(r^l)$
C_1	[-25.262%, -2.770%]	566	-4.36%
C_2	[-2.770%, 0.139%]	2793	-0.98%
C_3	(0.139%, 19.980%]	2631	1.66%

Table. 5.17: トヨタ自動車株価の予測誤差の分類区間 (ワード法, $L=4$)

Cluster	$(C_l)_{\min}, (C_l)_{\max}$	Size	$c_l(r^l)$
C_1	[-25.262%, -2.770%]	566	-4.36%
C_2	[-2.770%, 0.139%]	2793	-0.98%
C_3	(0.139%, 3.243%]	2333	1.20%
C_4	(3.243%, 17.538%]	298	5.22%

Table. 5.18: トヨタ自動車株価の予測誤差の分類区間 (ウォード法, L=5)

Cluster	$(C_l)_{\min}, (C_l)_{\max}$	Size	$c_l(r^l)$
C_1	[-25.262%, -2.770%)	566	-4.36%
C_2	[-2.770%, -1.205%)	978	-1.88%
C_3	(-1.205%, 0.139%]	1815	-0.50%
C_4	(0.139%, 3.243%]	2333	1.20%
C_5	(3.243%, 17.538%]	298	5.22%

Table. 5.19: トヨタ自動車株価の予測誤差の分類区間 (ウォード法, L=6)

Cluster	$(C_l)_{\min}, (C_l)_{\max}$	Size	$c_l(r^l)$
C_1	[-25.262%, -5.582%)	95	-7.32%
C_2	[-5.582%, -2.770%)	471	-3.76%
C_3	[-2.770%, -1.205%)	978	-1.88%
C_4	(-1.205%, 0.139%]	1815	-0.50%
C_5	(0.139%, 3.243%]	2333	1.20%
C_6	(3.243%, 17.538%]	298	5.22%

Table. 5.20: トヨタ自動車株価の予測誤差の分類区間 (ウォード法, L=7)

Cluster	$(C_l)_{\min}, (C_l)_{\max}$	Size	$c_l(r^l)$
C_1	[-25.262%, -5.582%)	95	-7.32%
C_2	[-5.582%, -2.770%)	471	-3.76%
C_3	[-2.770%, -1.205%)	978	-1.88%
C_4	(-1.205%, 0.139%]	1815	-0.50%
C_5	(0.139%, 1.266%]	1356	0.66%
C_6	(1.266%, 3.243%]	977	1.95%
C_7	(3.243%, 17.538%]	298	5.22%

Table. 5.21: トヨタ自動車株価の予測誤差の分類区間 (ウォード法, L=8)

Cluster	$(C_l)_{\min}, (C_l)_{\max}$	Size	$c_l(r^l)$
C_1	[-25.262%, -5.582%)	95	-7.32%
C_2	[-5.582%, -2.770%)	471	-3.76%
C_3	[-2.770%, -1.205%]	978	-1.88%
C_4	(-1.205%, 0.139%]	1815	-0.50%
C_5	(0.139%, 1.266%]	1356	0.66%
C_6	(1.266%, 3.243%]	977	1.95%
C_7	(3.243%, 6.242%]	240	4.36%
C_8	(6.242%, 17.538%]	58	8.76%

Table. 5.22: トヨタ自動車株価の予測誤差の分類区間 (ウォード法, L=9)

Cluster	$(C_l)_{\min}, (C_l)_{\max}$	Size	$c_l(r^l)$
C_1	[-25.262%, -16.433%)	2	-20.85%
C_2	[-16.433%, -5.582%)	93	-7.03%
C_3	[-5.582%, -2.770%)	471	-3.76%
C_4	[-2.770%, -1.205%]	978	-1.88%
C_5	(-1.205%, 0.139%]	1815	-0.50%
C_6	(0.139%, 1.266%]	1356	0.66%
C_7	(1.266%, 3.243%]	977	1.95%
C_8	(3.243%, 6.242%]	240	4.36%
C_9	(6.242%, 17.538%]	58	8.76%

Table. 5.23: トヨタ自動車株価の予測誤差の分類区間 (ウォード法, L=10)

Cluster	$(C_l)_{\min}, (C_l)_{\max}$	Size	$c_l(r^l)$
C_1	[-25.262%, -16.433%]	2	-20.85%
C_2	[-16.433%, -5.582%]	93	-7.03%
C_3	[-5.582%, -2.770%]	471	-3.76%
C_4	[-2.770%, -1.205%]	978	-1.88%
C_5	[-1.205%, 0.139%]	1815	-0.50%
C_6	[0.139%, 1.266%]	1356	0.66%
C_7	[1.266%, 3.243%]	977	1.95%
C_8	[3.243%, 6.242%]	240	4.36%
C_9	[6.242%, 9.582%]	45	7.69%
C_{10}	[9.582%, 17.538%]	13	12.47%

ン r_{t-9} , 3 日前の誤差 u_{t-3} , 9 日前の誤差 u_{t-9} , 10 日前の誤差 u_{t-10} と関連づけられていることがわかる。予測誤差ノードを取り入れると, 異なるネットワークが導かれることがわかる。

提案手法2による予測結果を時系列分析法や提案手法1と比較した結果を表5.24と図5.7に示す。ここで, AR, MA, ARMA, ARCH は, それぞれの時系列分析法による結果を, Algorithm1 は図4.6のネットワークによる予測結果を, Algorithm2 は図5.6のネットワークによる予測結果を示す。それぞれの数値は相関係数(CC), 平均二乗誤差(Ave.Error), 最大誤差(Max.Error), 最小誤差(Min.Error)を示す。

表5.24と図5.7から, 提案手法2は最も大きな相関係数を示しており, 平均二乗誤差と最大誤差も最小となっていることがわかる。また, 最小誤差はどの方法でも同程度である。平均二乗誤差をみると, 提案手法2は時系列分析法よりも30%程度向上している。提案手法1と比べると, 平均誤差は12%程度, 相関係数は9%程度改善している。

5.4.2 予測精度の比較(2)

前節の実験結果では, 予測期間は学習期間に含まれていた。そこで, 学習期間に含まれない期間の株価リターンを予測することを考える。上記で示した手法を用いて, 2009年1月1日から3月31日の取引日における株価リターン予測を行う。

Table. 5.24: 提案手法 2 と時系列分析法及び提案手法 1 の予測精度の比較 (トヨタ自動車株価)

	Max.Error	Min.Error	Ave.Error	CC
AR(9)	10.4027	0.0937	4.0789	0.4708
MA(6)	10.3414	0.2012	4.0853	0.4699
ARMA(9,6)	10.5354	0.0711	4.1409	0.4454
ARCH(9,9)	10.2655	0.0302	4.0709	0.4706
Algorithm1	6.8142	0.0491	3.1942	0.6469
Algorithm2	5.2172	0.1327	2.8178	0.7056

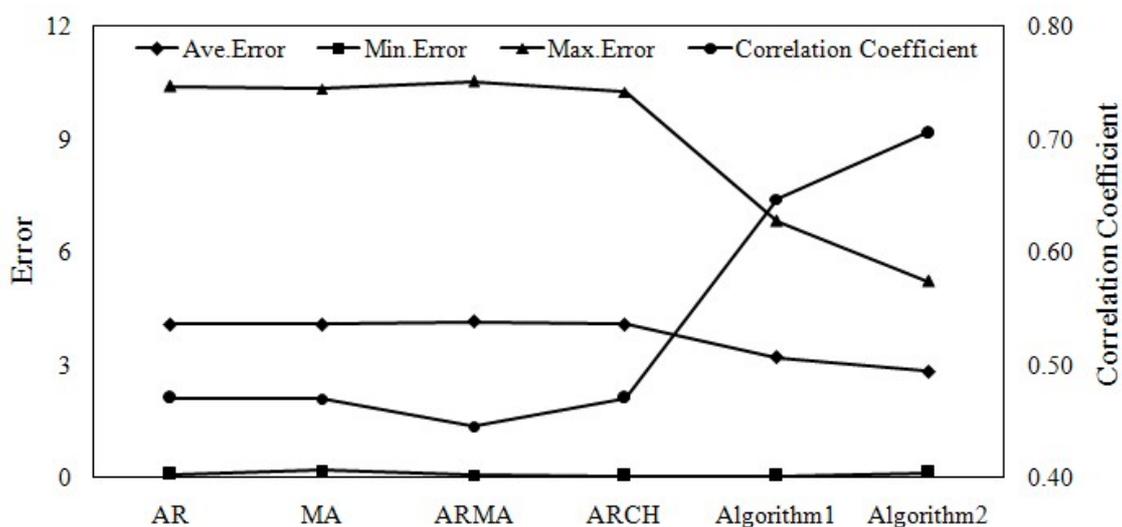


Fig. 5.7: 予測値と実測値の比較 (トヨタ自動車株価)

Table. 5.25: 予測期間が異なる場合の予測値と実測値の比較 (トヨタ自動車株価)

	Max.Error	Min.Error	Ave.Error	CC
AR(9)	7.5091	0.0615	2.6657	0.7448
MA(6)	7.6259	0.0319	2.6859	0.7417
ARMA(9,6)	7.1204	0.0401	2.6739	0.7427
ARCH(9,9)	8.0839	0.0597	2.6992	0.7527
Algorithm1	7.8579	0.0415	3.1494	0.7122
Algorithm2	5.7047	0.0818	2.6313	0.7984

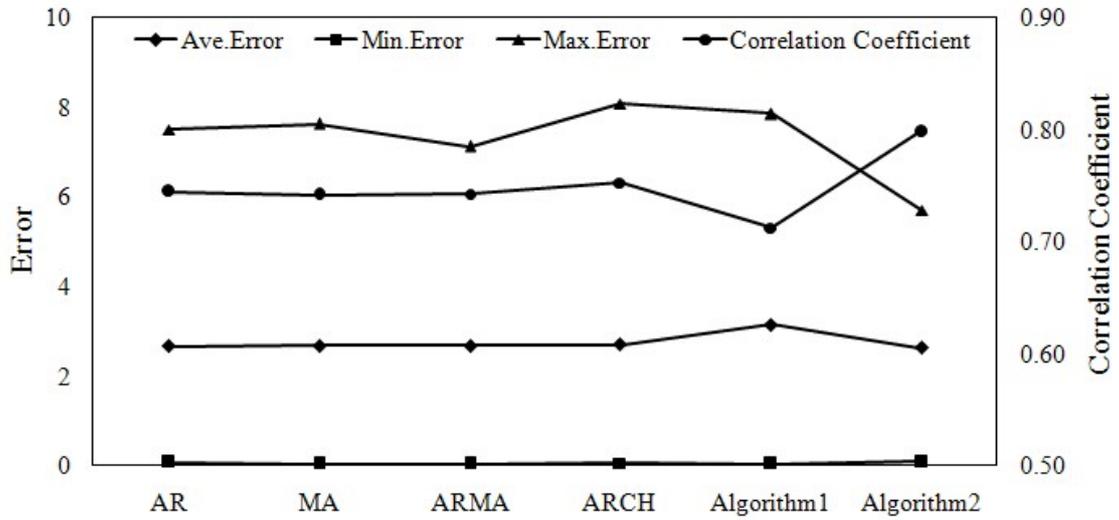


Fig. 5.8: 予測期間が異なる場合の予測値と実測値の比較 (トヨタ自動車株価)

提案手法2を時系列分析法や提案手法1と比較した結果を表5.25と図5.8に示す。これらの中で提案手法2は最も大きな相関係数を示している。また、最小誤差はどの方法でも0に近い値を示し、平均二乗誤差と最大誤差は提案手法が最も小さい値を示している。提案手法2の平均二乗誤差は時系列分析法よりも2%程度、最大誤差は20%程度改善していることがわかる。

第6章 提案手法3

6.1 概要

提案手法3は、提案手法2と同じアルゴリズムに基づいているが、予測用のベイジアンネットワークを構築する時、Naive ベイズ構造を仮定しないでグラフ構造の探索を行う。結果として、提案手法3では、構築されたネットワークがNaive ベイズ構造より複雑になっており、各子ノードが複数の親ノードを持っている。提案手法3は、複雑なネットワークを用いることで精度を向上しようとしている。

その一方で、提案手法3は提案手法2に比べて、ネットワーク構築にかかる計算時間がかなり増加すると予想される。そこで、本章では、提案手法3について予測精度の改善と計算コストについて検討する。

6.2 アルゴリズム

提案手法3のアルゴリズムは提案手法2と同じであり、下記のようになる。ただ、ネットワークを構築する時、Naive ベイズ構造を仮定しない。

1. 4.2.1 節の式 (4.1) により株価リターンを算出する。
2. 4.2.2 節のウォード法で株価リターンを離散化する。
3. 3.3 節のアルゴリズムにより離散化された過去の株価リターンからベイジアンネットワーク B を決定する。
4. B を用いて株価リターンを予測する。
5. 5.2.5 節の式 (5.1) により予測値と実測値の誤差を求める。
6. 5.2.6 節のウォード法で予測誤差を離散化する。
7. 3.3 節のアルゴリズムにより離散化された予測誤差を加えてベイジアンネットワーク B' を決定する。

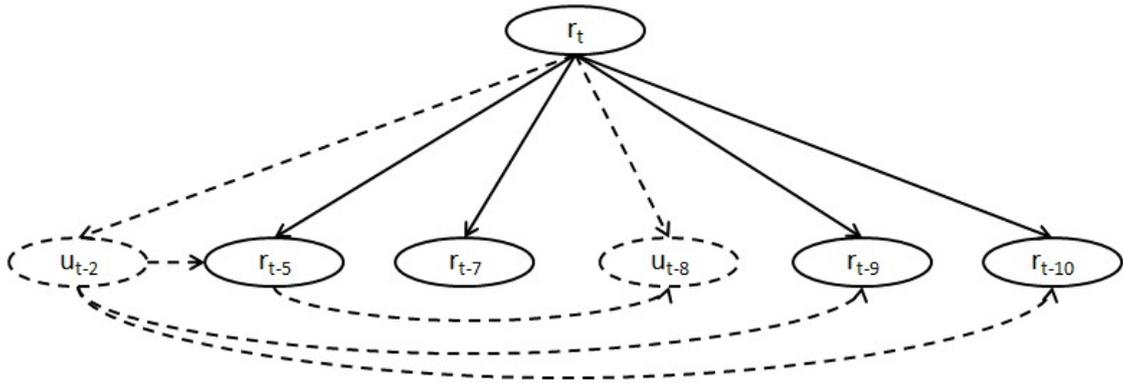


Fig. 6.1: 提案手法3で決定されたベイジアンネットワーク (日経平均株価)

8. B' を用いて株価リターンを再予測する .

6.3 数値実験

数値実験の例題として日経平均株価をとる . 1985年2月22日から2008年12月30日までの日次株価リターンを用いてベイジアンネットワークを決定し , それを用いて2009年1月1日から3月31日における取引日での株価リターン予測を行う . そして , 提案手法3を提案手法2と比較する .

ベイジアンネットワークを用いた手法のうち提案手法2と提案手法3についても , 1期前から10期前までの株価リターンや予測誤差のノードを用いてベイジアンネットワークを決定する .

まず , 株価リターンだけをノードとしてネットワークを決定する時 , 提案手法2の数値実験を参照する . そして , 株価リターンとその誤差の両方をノードとしてネットワークを決定する . 株価リターンの離散値総数を $L = 6$ として固定した上で , 誤差の離散値総数を $L = 2, 3, \dots, 10$ として離散化する . 具体的な離散値やクラスタのパラメータは提案手法2と同じになっており , 表5.2から表5.10に示す .

次に , ベイジアンネットワークを決定し , AIC(式(4.11))を計算した結果を表6.1に示す . これより , 離散値総数 $L = 4$ のとき AIC が最も小さいことがわかる . このときの具体的な離散値やクラスタのパラメータは表5.4に示す .

3.3節のアルゴリズムにより , 株価リターンと予測誤差で決定されたネットワークを図6.1に示す . この場合 , 株価リターン r_t は5日前の株価リターン r_{t-5} , 7日前の株価リターン

Table. 6.1: 日経平均株価予測誤差における離散値数と AIC の関係

Discrete number L	AIC
2	1.9960
3	1.9142
4	1.7684
5	1.7869
6	1.8828
7	1.9734
8	1.9097
9	2.1211
10	2.1466

Table. 6.2: 提案手法 3 と提案手法 2 の予測精度の比較 (日経平均株価)

	Max.Error	Min.Error	Ave.Error	CC
Algorithm2	4.9016	0.0105	2.3467	0.9375
Algorithm3	5.4338	0.0272	2.3361	0.9391

r_{t-7} , 9 日目の株価リターン r_{t-9} , 10 日目の株価リターン r_{t-10} , 2 日目の誤差 u_{t-2} , 8 日目の誤差 u_{t-8} と関連づけられていることがわかる。予測誤差ノードを取り入れると, 提案手法 2 の図 5.2 に比べてかなり複雑なグラフ構造のネットワークが導かれることがわかる。

求めたネットワークで予測を行い, 提案手法 2 と比較した結果を表 6.2 に示す。Algorithm2 は図 5.2 のネットワークによる予測結果を, Algorithm3 は図 6.1 のネットワークによる予測結果を示す。それぞれの数値は相関係数 (CC), 平均二乗誤差 (Ave.Error), 最大誤差 (Max.Error), 最小誤差 (Min.Error) を示す。

表 6.2 から, 最小誤差は同程度で, 最大誤差は少し悪化していることがわかる。しかし, 平均二乗誤差は提案手法 3 のほうがよく, 0.5% 程度改善されており, 相関係数でも提案手法 3 のほうが大きくなっている。

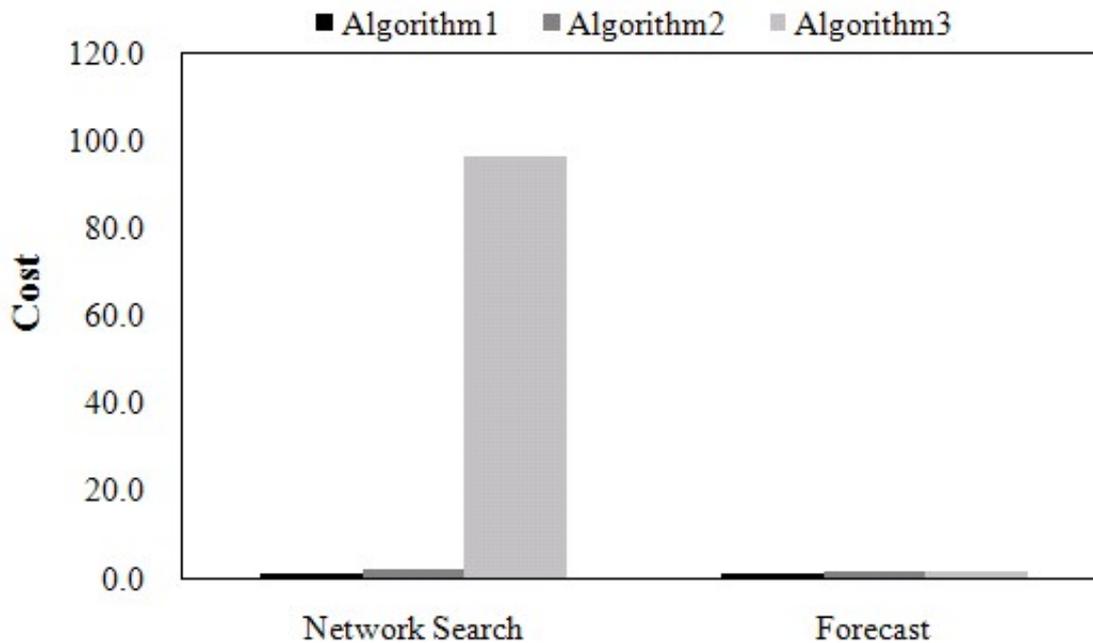


Fig. 6.2: 提案手法3と提案手法1及び提案手法2の計算コストの比較

6.4 考察

提案手法3は、提案手法2と同じアルゴリズムに基づいて、予測用のベイジアンネットワークを構築する時、Naive ベイズ構造を仮定しないでグラフ構造の探索を行う方法である。

そこで、日経平均株価リターン予測について、提案手法3と提案手法2及び提案手法1の計算時間を比較する。計算時間は、ベイジアンネットワークの学習にかかる時間と、決定されたベイジアンネットワークを用いて株価リターンを予測するためにかかる時間に分けて示す。

比較結果を表6.3に示す。これより、提案手法1と提案手法2ではネットワーク決定に529秒と908秒かかり、株価リターン予測に0.691秒と0.881秒かかることがわかる。一方、提案手法3では、それぞれ50902秒と0.889秒である。このことから、株価リターン予測にかかる時間はネットワーク決定にかかる時間よりかなり短いことがわかる。

次に、図6.2をみると、ネットワーク決定にかかる時間において、提案手法3は提案手法2と提案手法1のおよそ56倍と96倍になっている。ネットワーク決定にかかる時間はかなり増大しているが、これは最初にネットワーク決定のために1回だけ計算すればよい。また、予測にかかる時間について検討すると、提案手法3は提案手法2及び提案手法1とほぼ同じ程度であり、提案手法2と提案手法1のおよそ1倍と1.3倍程度となっている。ネッ

Table. 6.3: 提案手法 3 と提案手法 1 及び提案手法 2 の計算時間の比較

	Network Search	Forecast
Algorithm1	529sec	0.691sec
Algorithm2	908sec	0.881sec
Algorithm3	50902sec	0.889sec

トワーク決定にかかる時間に比べれば計算時間の増加はかなり小さいことがわかる。

第7章 結論

7.1 本論文のまとめ

本研究では、ベイジアンネットワークを用いた株価リターンの予測手法について述べた。

提案手法1では、株価リターンの分布をクラスタリング法により離散化し、その離散値を用いてベイジアンネットワークを決定し、今期の株価リターンを予測する。数値実験の対象として日経平均株価とトヨタ自動車株価をとり、提案手法1による予測結果をARモデル、MAモデル、ARMAモデル、ARCHモデルによる結果と比較した。

最初に日経平均株価を用いたシミュレーションにおいて、株価リターンの離散化法として等分割クラスタリングとワード法を比較した。等分割クラスタリングを用いる場合、提案手法1はARモデル、MAモデル、ARMAモデル、ARCHモデルに比べて予測誤差と相関係数の両方でより良い結果を示した。これに対して、ワード法を用いる場合、等分割クラスタリングを用いる方法より誤差は15%ほど小さくなり、相関係数も改善された。次の実験では、株価リターンが正規分布からずれているトヨタ自動車株価を用いて、ARモデル、MAモデル、ARMAモデル、ARCHモデルと提案手法1を比較した。ワード法を用いた提案手法1を時系列分析法と比較したところ、ARモデル、MAモデル、ARMAモデル、ARCHモデルより高い予測精度を示した。以上のことから、ベイジアンネットワークを用いる予測法に一定の有効性があることを確認した。

次に、ARモデルとARCHモデルの関係から着想した提案手法2について述べた。提案手法1では、過去の株価リターンが特定の値を取るときに、次の時点の株価リターンが特定の値を取る確率を推測するためのベイジアンネットワークを構築し、これを将来株価の予測に用いる。これに対して、提案手法2では、最初に提案手法1で過去の既知の株価リターンを予測し、その予測精度を評価する予測誤差を株価リターンとともに用いてノードを定義してベイジアンネットワークを構築する。

数値実験例として、提案手法1と同じように日経平均とトヨタ自動車の株価リターン予測を扱った。構成したネットワークを比較すると、いずれの場合においても、提案手法2で構築したネットワークには予測誤差ノードが含まれており、過去の予測精度が将来の予

測に影響していることがわかった。日経平均の株価リターン予測に適用した結果、平均誤差では、提案手法2は時系列分析法よりも30%程度、ベイジアンネットワークを用いた提案手法1よりも6%程度改善された。また、トヨタ自動車の株価リターンに適用した場合、平均誤差をみると、提案手法2は時系列分析法よりも30%程度向上している。ベイジアンネットワークを用いた提案手法1と比べても平均誤差は12%程度、相関係数は9%程度改善された。これらのことより、提案手法2にも一定の有効性があることが示された。また、計算時間について提案手法1と提案手法2を比較したところ、株価リターン予測にかかる計算時間に比べてネットワーク決定にかかる計算時間の比重が大きいことがわかった。

その後、提案手法2と同じアルゴリズムに基づきながら、ネットワークを構築する時、Naive ベイズ構造を仮定しないでネットワーク構造を探索する提案手法3を示した。例題として、日経平均株価を扱った。日経平均の株価リターンに適用した結果、平均誤差と相関係数において、提案手法3は、提案手法2より改善できたことがわかった。しかし、提案手法3の計算時間はかなり増加した。

7.2 今後の課題

最後に、提案手法が解決すべき今後の課題についても述べておきたい。

第1は、株価リターンの離散化法である。株価リターンを離散化して近似する場合、本研究では等分割クラスタリング法とワード法によるクラスタリング法を比較し、ワード法のほうが精度が良いことを示した。しかし、予測精度を上げるためにはクラスタ数を増やして、表現できる数値を増やすことが望ましいにもかかわらず、あまり細かく離散化するとベイジアンネットワークでの予測精度が低下する。そこで、精度を下げずに細かい離散化を可能とするような他のクラスタリング手法(K-Means [60] [61] [62], C-Means [55] [88] [89] など)についても検討する必要がある。

第2は、ベイジアンネットワークで予測を行うためには、ある程度多くのデータが必要となることである。そこで、より少ないデータ数で良い精度を得られるようにアルゴリズムの改良が望まれる。また、提案手法では、ベイジアンネットワークを学習するためにある程度多くのデータが必要となるのに対して、従来の時系列分析手法では、もっと少ないデータ数で予測することができる。そこで、組み合わせて使用するなど、今後両者の特長を生かした活用法についても検討を進めたい。

第3は、提案手法2と提案手法3の計算コストである。予測精度の点からみると提案手法

3が最も良い精度を示している。しかし、提案手法3におけるネットワーク決定にかかる計算コストはかなり大きい。そこで、今後は提案手法3の計算精度を犠牲にしないで、計算時間を縮小する方法について検討を進めていく予定である。

提案手法は、従来ある時系列分析法に比べて予測精度において一定の改善を示すことができたが、まだ十分とはいえない。特に、本研究を始めるきっかけとなった株式市場の暴騰や暴落といった現象は、残念ながら予測できていない。今後は、予測に用いるデータの種類を増やすこと、市場トレンドの転換点ごとに自動的に予測手法を変えていくレジームスイッチングなどの手法を組み合わせることで、市場の暴騰や暴落の予測につなげていきたいと考えている。

謝辞

本研究の遂行ならびに本論文の作成にあたり，名古屋大学大学院情報科学研究科北栄輔教授には，博士課程（前期）を含め5年間終始激励して頂き，懇切丁寧な御指導，御助力を賜りました．ここに深く感謝の意を表します．

更に，名古屋大学大学院情報科学研究科有田隆也教授，大岡昌博教授には，ご多忙中にも関わらず，大変貴重な御指導，御助言を賜りました．先生方の懇切丁寧な御助力のおかげで本論文を完成させることができました．心より感謝致します．

また，名古屋大学情報科学研究科複雑系科学専攻創発システム論講座の皆様には，本研究の遂行にあたり多大なご協力を頂きました．ここに厚くお礼申し上げます．

左 毅

参考文献

- [1] 繁榊算男, 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] D.Heckerman, D.Gerger and D.M.Chickering : Learning Bayesian Networks : The Combination of Knowledge and Statistical Data, *Machine Learning*, Vol.20, pp.197-243, 1995
- [3] J.Cheng and R.Greiner : Learning Bayesian Belief Network Classifiers : Algorithms and System, *Advances in artificial intelligence : 14th Biennial Conference of the Canadian Society for Computational Studies of Intelligence*, Vol.2056, pp.141-151, Springer, 2001
- [4] 本村陽一, 岩崎弘利 : ベイジアンネットワーク技術 ユーザ・顧客のモデル化と不確実性推論, 東京電機大学出版局, 2006
- [5] 原章, 長尾智晴 : 自動グループ構成手法 ADG による人工株式市場の構築と解析, *情報処理学会論文誌*, Vol.43, No.7, pp.2292-2299, 2002
- [6] 池田欽一, 時永祥三 : フラクタル時系列の予測手法を用いた株価予測とその応用, *日本オペレーションズ・リサーチ学会論文誌*, Vol.42, No.1, pp.18-31, 1999
- [7] E.Chancellor : *Devil Take the Hindmost:A History of Financial Speculation*, Plume, 2000
- [8] 大村敬一, 川北英隆 : 日本の株式市場, 東洋経済新報社, 1992
- [9] 小西貞則, 越智義道, 大森裕浩 : 計算統計学の方法:ブートストラップ・EM アルゴリズム・MCMC, 朝倉書店, 2008
- [10] 伊庭幸人, 種村正美, 大森裕浩, 和合肇, 佐藤整尚, 高橋明彦 : マルコフ連鎖モンテカルロ法とその周辺, 岩波書店, 2005
- [11] 日経平均プロフィール, <http://www.nikkei.co.jp/nkave>

- [12] Dow Jones Index, <http://www.djindexes.com/>
- [13] Financial Times Stock Exchange, <http://www.ftse.com/>
- [14] D.Sornette : *Why Stock Markets Crash:Critical Events in Complex Financial Systems*, Princeton University Press, 2002
- [15] 北川源四郎 : 時系列解析における数値的方法:計算統計学的方法, 計算機統計学会, Vol.15, No.2, pp.159-170, 2003
- [16] 中妻照雄 : ファイナンスのための MCMC 法によるベイズ分析, 三菱経済研究所, 2003
- [17] 下平英寿, 伊藤秀一, 久保川達也, 竹内啓 : モデル選択:予測・検定・推定の交差点, 岩波書店, 2004
- [18] 小西貞則, 北川源四郎 : 情報量規準, 朝倉書店, 2004
- [19] 奥村宏 : 日本の株式市場:投機時代の株価はこう決まる, ダイヤモンド社, 1988
- [20] 伊庭幸人 : ベイズ統計と統計物理, 岩波書店, 2003
- [21] R.Bookstaber : *The Complete Investment Book:Trading Stocks, Bonds, and Options With Computer Applications*, Scott Foresman Trade, 1985
- [22] 釜江廣志 : 証券分析の基礎, 有斐閣, 1995
- [23] 渡部敏明 : 日本の株式市場におけるボラティリティーの変動:ARCH 型モデルによる分析, 三菱経済研究所, 1997
- [24] R.F.Engle and V.K.Hg : Measuring and Testing the Impact of News on Volatility, *Journal of Finance*, Vol.48, No.5, pp. 1749-1778, 1993
- [25] G.F.Cooper and E.Herskovits : A Bayesian Method for the Induction of Probabilistic Networks from Data, *Machine Learning*, Vol.9, No.4, pp.309-347, 1992
- [26] 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
- [27] 左毅, 原田昌朗, 北栄輔 : ベイジアンネットワークを用いた株価予測法の精度改善, 情報処理学会論文誌トランザクション 数理モデル化と応用, Vol.4, No.4, pp.92-103, 2011

- [28] R.F.Engle : Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation, *Econometrica*, Vol.50, No.4, pp.987-1008, 1982
- [29] G.Box, G.M.Jenkins and G.Reinsel : *Time Series Analysis:Forecasting and Control*, Prentice-Hall, 1994
- [30] P.J.Brockwell and R.A.Davis, *Time Series:Theory and Methods*, Springer, 2009
- [31] J.Pearl and S.Russell, *Bayesian Networks*, MIT Press, 2002
- [32] I.Ben-Gal : *Bayesian Networks*, John Wiley & Sons, 2007
- [33] J.Pearl : *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference*, Morgan Kaufmann, 1988
- [34] D.B.Nelson : Conditional Heteroskedasticity in Asset Returns:A New Approach, *Econometrica*, Vol.59, No.2, pp.347-370, 1991
- [35] H.Takayasu : *Practical Fruits of Econophysics*, Springer, 2006
- [36] 森棟公夫, 中窪文男, 富安弘毅, 中園美香 : ファイナンス計量分析入門, 東洋経済新報社, 2008
- [37] R.Meester : *A Natural Introduction to Probability Theory*, Birkhauser, 2008
- [38] A.Klenke : *Probability Theory:A Comprehensive Course*, Springer, 2008
- [39] H.Kantz and T.Schreiber : *Nonlinear Time Series Analysis*, Cambridge University Press, 2004
- [40] T.Koski and J.Noble : *Bayesian Networks:An Introduction*, Wiley, 2009
- [41] W.D.Dechert : *Chaos Theory in Economics: Methods, Models and Evidence*, Edward Elgar Pub, 1996
- [42] T.Terasvirta, D.Tjostheim and C.W.J.Granger : *Modelling Nonlinear Economic Time Series*, Oxford Univ Pr, 2011
- [43] 中妻照雄 : 入門ベイズ統計学, 朝倉書店, 2007

- [44] 繁榊算男 : ベイズ統計入門, 東京大学出版会, 1985
- [45] 照井 伸彦 : ベイズモデリングによるマーケティング分析, 東京電機大学出版局, 2008
- [46] D.Koller and N.Friedman : *Probabilistic Graphical Models:Principles and Techniques*, The MIT Press, 2009
- [47] R.S.Mamon and R.J.Elliott : *Hidden Markov Models in Finance*, Springer-Verlag, 2007
- [48] T.Bollerslev : Generalized Autoregressive Conditional Heteroskedasticity, *Journal of Econometrics*, Vol.31, pp.307-327, 1986
- [49] A.S.Weigend and S.Shi : Predicting Daily Probability Distributions of S&P500 Returns, *Journal of Forecasting*, Vol.19, No.4, pp.375-392, 2000
- [50] J.M.Mendel : Fuzzy Logic Systems for Engineering:A Tutorial, *Proceedings of the IEEE*, Vol.83, No.3, pp.345-377, 1995
- [51] T.Munakata and Y.Jani : Fuzzy Systems; An Overview, *Communications of the ACM*, Vol.37, No.3, pp.68-76, 1994
- [52] B.E.Rosen : Neural Network Moving Averages for Time Series Prediction, *SPIE*, Vol.1966, No.2, pp.448-456, 1993
- [53] M.Nelson, T.Hiil, W.Remus and M.O'Connor : Time Series Forecasting using Neural Networks: should the data be deseasonalized first?, *Journal of Forecasting*, Vol.18, No.5, pp.359-367, 1999
- [54] J.Faraway and C.Chatfield : Time Series Forecasting with Neural Networks: a comparative study using the air line data, *Journal of the Royal Statistical Society*, Vol.47, No.2, pp.231-250, 1998
- [55] J.C.Bezdek, R.Ehrlich, W.Full : FCM:The Fuzzy C-Means Clustering Algorithm, *Computers&Geosciences*, Vol.10, No.2-3, pp.191-203, 1984
- [56] A.B.Poritz : Hidden Markov Models:a guided tour, *ICASSP*, Vol.1, No.11-14, pp.7-13, 1988

- [57] M.R.Hassan, B.Nath and M.Kirley : A Fusion Model of HMM, ANN and GA for Stock Market Forecasting, *Expert Systems with Applications*, Vol.33, No.1, pp.171-180, 2007
- [58] M.R.Hassan : A combination of HMM and Fuzzy model for Stock Market Forecasting, *Neurocomputing*, Vol.72, No.16-18, pp.3439-3446, 2009
- [59] A.K.Jain, M.N.Murty and P.J.Flynn : Data Clustering:a review, *ACM Computing Surveys*, Vol.31, No.3, pp.264-323, 1999
- [60] J.A.Hartigan and M.A.Wong : A K-Means Clustering Algorithm, *Applied Statistics*, Vol.28, No.1, pp.100-108, 1979
- [61] T.Kanungo, D.M.Mount, N.S.Netanyahu, C.D.Piatko, R.Silverman and A.Y.Wu : An Efficient k-Means Clustering Algorithm:Analysis and Implementation, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol.24, No.7, pp.100-108, 2002
- [62] S.P.Lloyd : Least Squares Quantization in PCM, *IEEE Trans. Information Theory*, Vol.28, No.2, pp.129-137, 1982
- [63] M.Sugeno and T.Yasukawa : A Fuzzy-Logic-Based Approach to Qualitative Modeling, *IEEE Trans. Fuzzy Systems*, Vol.1, No.1, pp.1-7, 1993
- [64] M.Sugeno and K.Tanaka : Successive Identification of a Fuzzy Model and its Applications to Prediction of a Complex System, *Fuzzy Sets and Systems*, Vol.42, No.3, pp.315-334, 1991
- [65] E.K.Berndt, B.H.Hall and R.E.Hall : Estimation and Inference in Nonlinear Structural Models, *Annals of Economic and Social Measurement*, Vol.3, No.4, pp.103-116, 1974
- [66] R.F.Engle : Estimates of the Variance of U.S. Inflation Based upon the ARCH Model, *Journal of Money, Credit and Banking*, Vol.19, No.3, pp.286-301, 1983
- [67] A.A.Weiss : ARMA models with ARCH errors, *Journal of Time Series Analysis*, Vol.5, No.3, pp.129-143, 1984

- [68] A.Krogh, M.Brown, I.S.Mian, K.Sjolander and D.Haussler : Hidden Markov Models in Computational Biology:Applications to protein modeling, *J Mol Biol*, Vol.235, No.5, pp.1501-1531, 1994
- [69] L.E.Baum and J.A.Eagon : An inequality with applications to statistical estimation for probabilistic functions of Markov processes and to a model for ecology, *Bulletin of the American Mathematical Society*, Vol.73, No.3, pp.360-363, 1967
- [70] L.R.Rabiner : A Tutorial on Hidden Markov Models and Selected Applications in Speech Recognition, *Proceedings of the IEEE*, Vol.77, No.2, pp.257-286, 1989
- [71] W.H.Press, S.A.Teukolsky, W.T.Vetterling and B.P.Flannery : *Numerical Recipes 3rd Edition:The Art of Scientific Computing*, Cambridge University Press, 2007
- [72] J.A.Anderson : *An Introduction to Neural Networks*, The MIT Press, 1995
- [73] S.Haykin : *Neural Networks:A Comprehensive Foundation*, Prentice Hall, 1998
- [74] J.Lawrence : *Introduction to Neural Networks*, California Scientific Software Press, 1994
- [75] S.Seneff : A computational model for the peripheral auditory system: Application of speech recognition research, Acoustics, Speech, and Signal Processing, *IEEE International Conference on ICASSP '86*, Vol.11, pp.1983-1986, 1986
- [76] R.Lippmann : Review of neural networks for speech recognition, *Neural Computation*, Vol.1, No.1, pp.1-38, 1989
- [77] E.Botha, E.Barnard and D.Casasent : Optical neural networks for image analysis:imaging spectroscopy and production systems, *IEEE International Conference on Neural Networks*, Vol.1, pp.541-546, 1988
- [78] M.Egmont-Petersen, D.Ridder and H.Handels : Image processing with neural networks-a review, *Pattern Recognition*, Vol.35, No.10, pp.2279-2301, 2002
- [79] S.Margarita, P.Terna and A.Beltratti : *Neural Networks for Economic and Financial Modeling*, International Thomson Computer Press, 1995

- [80] R.Lippmann : An introduction to computing with neural nets, *IEEE ASSP*, Vol.4, No.2, pp.4-22, 1987
- [81] P.D.McNelis : *Neural Networks in Finance: Gaining Predictive Edge in the Market*, Academic Press, 2005
- [82] D.E.Rumelhart, G.E.Hinton and R.J.Willimas : Learning Internal Representations by Error Propagation, *Parallel distributed processing*, Vol.1, pp.318-362, 1986
- [83] J.H.Ward : Hierarchical Grouping to Optimize an Objective Function, *Journal of the American Statistical Association*, Vol.58, No.301, pp.236-244, 1963
- [84] L.Kaufman and P.JRousseeuw : *Finding Groups in Data: An Introduction to Cluster Analysis*, New York:Wiley, 1990
- [85] B.S.Everitt, S.Landau and M.Leese : *Clustering Analysis*, Wiley, 2001
- [86] J.A.Hartigan : *Algorithms for Clustering Data*, New York;London:Wiley, 1975
- [87] A.K.Jain and R.C.Dubes : *Algorithms for Clustering Data*, Prentice Hall College Div, 1988
- [88] J.C.Dunn : A fuzzy relative of the ISODATA process and its use in detecting compact well-separated clusters, *Journal of Cybernetics*, Vol.3, No.3, pp.32-57, 1973
- [89] J.C.Bezdek : A Convergence Theorem for the Fuzzy ISODATA Clustering Algorithms, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol.PAMI-2, No.1, pp.1-8, 1980
- [90] S.Chib, F.Nardari and N.Shephard : Markov chain Monte Carlo methods for stochastic volatility models, *Journal of Econometrics*, Vol.108, No.2, pp.281-316, 2002
- [91] S.Chib and E.Greenberg : Markov Chain Monte Carlo Simulation Methods in Econometrics, *Econometric Theory*, Vol.12, No.3, pp.409-431, 1996
- [92] B.Eraker : MCMC Analysis of Diffusion Models with Application to Finance, *Journal of Business and Economic Statistics*, Vol.19, No.2, pp.177-191, 1996

- [93] S.Kim, N.Shephard and S.Chib : Stochastic Volatility: Likelihood Inference and Comparison with ARCH Models, *the Review of Economic Studies*, Vol.65, No.3, pp.361-393, 1998
- [94] P.K.Clark : A Subordinated Stochastic Process Model with Finite Variance for Speculative Prices, *Econometrica*, Vol.41, No.1, pp.135-155, 1973
- [95] E.Jacquier, N.G.Polson and P.E.Rossi : Bayesian Analysis of Stochastic Volatility Models, *Journal of Business and Economic Statistics*, Vol.12, No.4, pp.371-389, 1994
- [96] J.Hull and A.White : The Pricing of Options on Assets with Stochastic Volatilities, *Journal of Finance*, Vol.42, No.2, pp.281-300, 1987
- [97] J.B.Wiggins : Option values under stochastic volatility: Theory and empirical estimates, *Journal of Financial Economics*, Vol.19, No.2, pp.351-372, 1987
- [98] A.Melino and S.M.Turnbull : Pricing foreign currency options with stochastic volatility, *Journal of Econometrics*, Vol.45, No.1-2, pp.239-265, 1990
- [99] G.E.Tauchen and M.Pitts : The Price Variability-Volume Relationship on Speculative Markets, *Econometrica*, Vol.51, No.2, pp.485-505, 1983
- [100] A.R.Gallant, D.Hsieh and G.Tauchen : On Fitting A Recalcitrant Series: The Pound/Dollar Exchange Rate, 1974-83, in Nonparametric and Semiparametric Methods in Econometrics and Statistics, *Proceedings of the Fifth International Symposium in Economic Theory and Econometrics*, Cambridge University Press, 1991
- [101] T.G.Andersen : Return Volatility and Trading Volume: An Information Flow Interpretation of Stochastic Volatility, *Journal of Finance*, Vol.51, No.1, pp.169-204, 1996
- [102] B.Dischel : The D1 stochastic temperature model for valuing weather futures and options, *Applied Derivatives Trading*, 1999
- [103] B.Dischel : Shaping History for the Weather Risk Market, *Weather Risk Special Report from Risk Magazine and Energy and Power Risk Management*, 1999

- [104] X. Gavaix : *Stock Trade Patterns could Predict Dinancial Earthquakes*, MIT News, 2003
- [105] B.B. Mandelbrot : The Variation of Certain Speculative Prices, *Journal of Business*, Vol.36, No.4, pp.394-419, 1963
- [106] R.N. Mantegna, H.E. Stanley : Scaling Behaviour in the Dynamics of an Economic Index, *Nature*, Vol.376, No.6535, pp.46-49, 1995
- [107] X.Gavaix : Understanding the Metropolis-Hastings Algorithm, *American Statistician*, Vol.49, No.4, pp.327-335, 1995
- [108] 三井秀俊 : Gibbs Sampler を用いたベイズ推定法によるオプション評価の方法, *産業経営研究*, Vol.25, pp.69-88, 2003
- [109] B.B.Mandelbrot : *The Fractal Geometry of Nature*, W.H.Freeman, 1982
- [110] 高安秀樹 : フラクタル, 朝倉書店, 1986
- [111] C.M.Bishop : *Pattern Recognition and Machine Learning*, Springer, 2006
- [112] R.W.Robinson : Counting Unlabeled Acyclic Digraphs, *Lecture Notes in Mathematics*, Vol.622, pp.28-43, 1977
- [113] P.Larranaga, C.M.H.Kuijpers, R.H.Murga, Y.Yurramendi : Learning Bayesian Network Structures by Searching For the Best Ordering With Genetic Algorithms, *IEEE Transactions on Systems, Man and Cybernetics*, Vol.26, No.4, pp.487-493, 1996
- [114] 植野真臣, 河竹千春, 大西仁, 繁榎算男 : ネットワーク型テスト理論における構造最適化法, *電子情報通信学会技術研究報告*, Vol.93, No.541, pp.124-145, 1994
- [115] M.L.Wong, W.Lam,K.S.Leung : Using Evolutionary Programming and Minimum Description Length Principle for Data Mining of Bayesian Networks, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol.21, No.2, pp.174-178, 1999