

富士通株と北陸電力株との株価変動パターンの比較 —ニューラルネットを用いた経路依存性の測定—

高見茂雄

要約：株価はランダムに動くかそれとも経路依存性があるのか、程度の差はあれど銘柄でも両者の特色がある。しかし、銘柄によってどちらの傾向がより強くみられるか、経路依存性を計測する尺度は何か、これがわれわれの問題意識の出発点である。われわれは値動きの激しい電気機器業界のなかから富士通株、値動きが安定的な電力業界から北陸電力株を選び、両者の株価変動パターンをニューラルネットを用いて分析した。われわれは以下の結論を得た。富士通株は過去の経路依存性を示す野線情報を咀嚼して翌週の終値相場が形成される傾向にあるのに対し、北陸電力株ではほとんど当週の終値だけで翌週の終値相場が形成されるランダム性の強い線形に近い関係が見られる。

キーワード：ニューラルネット、野線、相場変動、経路依存性の測定

1. はじめに

株価変動は蕘谷（2001，p.279）が主張するように「期種や観測期間に関わりなく、株価がランダム・ウォークするという主張もランダム・ウォークしないという主張もともに誤りである。ランダム・ウォークするときもしないときもある、というあいまいな結論が実は正しいであろう。」という主張は妥当である。また、銘柄によってもランダムな動きを示しがちな株とテクニカルな動きすなわち経路依存性を示した動きが見られがちな株とに分かれるであろう。このうち、蕘谷（2001）が指摘した同一銘柄での期間による経路依存性（あるいは裏返してランダム性）の違いは例えば regime switching モデルで説明す

る試みが見られる¹。しかし、相場転換点を表す背後のパラメータを設定する必要がある点で難点があり、モデル設定に恣意性が入り込む恐れがある。そこで、高見・上坂（2004）が試みたように、そもそも大局的に説明するモデル設定は断念し、せいぜい過去1ヶ月くらいの週足で局所的に経路依存性を観測する方向が考えられる。一方、銘柄間による経路依存性の強弱の測定についてはどのようなアプローチが考えられるであろうか。

ひとつには ARMA (p, q) モデルのように将来の株価が過去の株価に依存させる要素と正規分布に従うホワイトノイズ要素とから構成されるモデルを想定し、前者 p と後者 q のパラメータの数と特性を評価するアプローチが考えられる²。しかし、前者 p のパラメータ設定では過去の終値がどう関連付けられているかを示すに過ぎず、終値から次の終値にどのような相場挙動を示したかがモデルに反映されていない。また、2つの銘柄間を比較するにあたり、それぞれ (p, q) の設定の当てはまりのよさ悪さが生じ同一モデルでの比較が適当でない場合がありうるなどの問題点が残る。この点ですべてのデータを用いて最適かつ一意的モデルを追及するアプローチは適当とはいえない。そこで、われわれはデータを学習用と検証用にわけ、前者データから得られたモデルを後者データで検証し、モデル構造の特徴を観察するニューラルネット手法を用いる³。もうひとつのアプローチには、株価収益率 $\ln(P_{t+1}/P_t)$ の実際の分布を観察し正規分布にどれだけ近いかに判定することも考えられる。しかし、正規分布に適合する確率が高ければランダム性が強いとは言えるかもしれないが、逆に正規分布に従わないとしても必ずしも経路依存性が高いとは直接言えない点、つまり経路依存性の強い分布を具体的に想定することはできない点では難点がある。しかし、このアプローチはデータの特徴を概観するには活用でき

1 Dacco & Satchell (1999)

2 ブロックウェル・デービス(2000)

3 豊田 (2001, p.32) はニューラルネットなどのデータマイニング手法の特徴は交差妥当性 (Cross Validation) の標準的使用と最適性・一意性を重視しない思想であると指摘している。

るので、3節で用いることにする。

ところで、これらのアプローチでは2銘柄の大局的相場変動を説明しようとの暗黙の前提がおかれているが、期間の取り方によって2銘柄挙動の個性の違いが強く出る時期とそうでない時期とが現れるおそれがある。それゆえ、銘柄間比較であっても局所的な期間に限定してどれだけ経路依存性要素が含まれるかを分析することが望ましい。そこで、本論文においても高見・上坂(2004)のように、せいぜい過去1ヶ月くらいの週足データを対象にニューラルネットを用いて2銘柄間の比較を行う。

ニューラルネットとはデータマイニング手法のひとつであり、図1のように人間の脳細胞ニューロンに見立てる。つまり、入力信号がどのようにシナプスを活性化させ、いかなる経路で電気信号を送り、最終的に出力させるかを独自の探索的計算方法によって解を求める手法である。目的関数は理論値と実際値との最小2乗誤差で、その極小化を計算目的とし、計算結果はシナプス間を結ぶ軸索の重みの強弱で表現される。ブラックボックスとみなされていた入力データと出力データとの関係にある程度可視化がはかられるという効果が期待できる。また、その特徴としては逆伝播学習法という自動的微分イタレーションアルゴリズムであり、微分演算の簡素化のため、シグモイド関数変換を行って

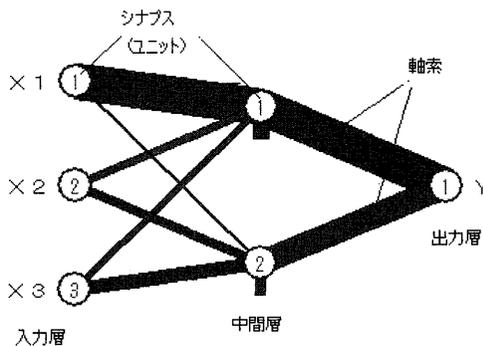


図1 ニューラルネット説明図

ることである。モデル構造や因果性を特定することなく試行錯誤で何らかの法則性を見つけようとするときに向いており、複雑な非線形関係も中間層を設けることで表現できる点がメリットである。一方、得られた解は局所的最適解であって大局的最適解ではないこと、微分演算の速度を調整するパラメータなどの設定に職人芸が要求されること、特定の入力ベクトル要素が出力データにインパクトを与えていることは観察できてもその背後の意味づけが難しいことなどはデメリットと言える。

ニューラルネットを用いた株式相場等に関する先行研究はあまり多くはない。Kimoto et al. (1990)、馬場 (1998) は TOPIX を対象に外生的マクロ経済変数と株価移動平均などのテクニカル指標を入力変数に、買い・売りシグナルを出力変数にとり、トレード益のパフォーマンスを評価している。中田 (1992) は債券先物の予測システムを説明しているが、同様に外生的マクロ経済変数とテクニカル指標を入力変数に用いている。Yoon and Swales (1991) は企業属性や戦略などのカテゴリー変数を入力変数に加え、40社の個別株を対象に分析を行なっている。これらの研究は資産価格の変動を過去の挙動だけではなく、外生変数を導入して分析しているが、われわれの関心は経路依存性の観測であるので、あくまで過去の株価データのみを取り上げる。また、Kamijo and Tanigawa (1990) はトライアングルを形成する16社個別株の描く罫線情報（ローソク足）をもとに、どのようなときにトライアングルをブレイクするかをニューラルネットで分析している。この研究は従来職人芸に頼っていたトライアングルチャートの判定をニューラルネットで探索できることを示すことを目的としており、直接われわれの研究には応用できないものの示唆に富む。というのは罫線から得られる情報は過去の終値と終値との経路をある程度反映しているからである。すなわち図2のように罫線は上昇した相場は白い実体で、下落した相場は黒い実体で描かれ、たとえば、上昇した例で示せば曲線のような正確な株価挙動は分からないにしても、高値と実体との距離である上影、実体と安値との距離である下影の長さを示すことでその日の相場挙動の

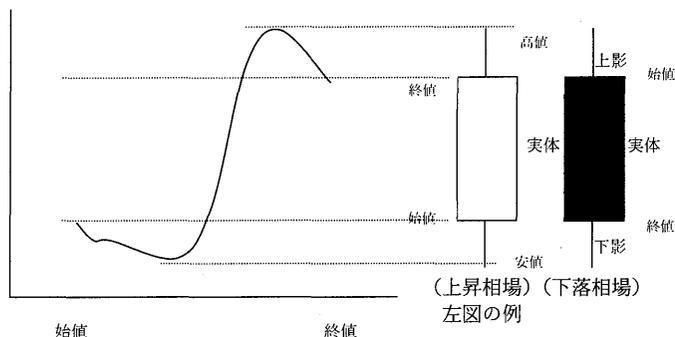


図2 罫線と相場挙動との関係

重要な一面を伝える情報をもっている。そこでわれわれは取り扱う株式銘柄は異なるものの高見・上坂（2004）と同様入力データに過去の終値以外に上影，下影，実体の罫線情報をも採用することにする。

本論文の構成は以下の通りである。2節では富士通株と北陸電力株を選択した経緯について述べる。3節では両者株価データについて原データを説明するとともに、いかに加工し前処理を行ったかを述べる。4節ではソフトウェア Neurosim/ L V4をいかに処理して分析結果を導いたかを述べる。5節では Neurosim/ L V4による分析結果を解説し、6節では結論としてインプリケーションと限界、今後の課題について述べる。

2. 個別株選択の経緯

われわれの研究の目的は個別株の経路依存性を測定する尺度を考察することである。そこでできるだけコントラストのある株式2銘柄を選択し分析を行うが、ランダム性の強い株と経路依存性の強い個別株を選択することが求められる。しかもできる限り恣意性を排除して選択しなければならない。しかし、3,500社を超える銘柄のチャートを見て判断することは困難であり、その判定も職人芸を要求される。そこで、第1次近似として、株価変動の激しい株と安定的な株とから2銘柄選択することにした。株価変動の激しい株は流動性に富み、

売り手と買い手との思惑が交錯し相場形成がなされており、相場参加者は将来の相場予想に罫線情報を参考にすることが多い。そのため、くせのある動きすなわち経路依存性が強いと考えられる。一方、安定的な株は長期保有者が多く、日々の浮動株売買取引は揺らぎのような振る舞いをするかもしれない。これは仮説に過ぎないが、コントラストのある株式2銘柄を抽出することは最低限期待できる。

取引の激しい株の尺度はベータ値が用いられる。東京証券取引所（2004）の「規模別・業種別ベータ値と決定係数」から業種別ベータ値を検討すると、1999年10月から2004年9月までの期間でベータ値の高い業種は証券業1.79、卸売業1.62、情報通信業1.51、電気機器業1.39である。ベータ値の低い業種は電気・ガス業-0.07である。ベータ値の高い業種では日本株式市場を代表する銘柄が比較的多くあると思われる電気機器業を選択した。ベータ値の低い業種は電気・ガス業が妥当だが、電気業界から選択することにした。そのうち候補として、電気通信業から7社、電気業界から3社選択した。表1はそれぞれの銘柄のベータ値とボラティリティー（株価収益率の標準偏差）を掲げている。この表を検討するに、高いベータ値を示す銘柄は高いボラティリティーを示す傾向にあるが、値動きの激しさの尺度となるのはボラティリティーである。その基準から値動きの激しい銘柄は富士通、ソニーが候補になり、安定的な銘柄は北陸電力とするのが妥当といえる。

表1 候補銘柄の東証ベータ値とボラティリティー

	TOPIX	ソニー	シャープ	三洋電機	松下	東芝	日立	富士通	東電	関西電力	北陸電力
ベータ値	1.00	1.51	1.52	1.11	0.72	1.49	1.56	1.84	-0.06	-0.17	0.01
ボラティリティー	4.63%	13.23%	12.03%	12.21%	8.10%	11.17%	9.99%	12.63%	6.13%	4.59%	4.35%

(対象期間 1999.10- 2004.9)

しかし、これら数値の計算において東京証券取引所は配当金を考慮した月次株価収益率を基準にしている。一方われわれは週次データで分析を行いたい。というのは、日足データでは前日終値と当日始値との間に窓が多くみられることが予想され、変動性が極端に高くなる恐れがあるからである。その反面、月

次ではデータ数が少なく、また局所的といっても何ヶ月か前ものデータを扱うには時間間隔が空き過ぎる感がある。そこで、対象データの期間も異なることもあり、東洋経済新報社（2003）より対象期間1992年1月第1週から2002年12月第4週までの574週終値データを用いて、週次ボラティリティーを計算しなおした⁴。

表2 ボラティリティー計算結果

	TOPIX	ソニー	シャープ	三洋電機	松下	東芝	日立	富士通	東電	関西電力	北陸電力
週次	19.58%	34.34%	32.92%	37.76%	28.58%	33.88%	32.84%	37.82%	25.04%	20.17%	17.15%
参考											
日次	20.06%	34.57%	34.63%	39.62%	29.41%	34.83%	33.06%	36.01%	24.86%	23.02%	18.38%
月次	18.43%	36.46%	37.26%	34.89%	27.60%	33.35%	31.93%	38.26%	21.90%	16.23%	15.04%

(対象期間 1992年1月から2002年12月まで)

表2を観察すると単位時間の長さによって数値は変わってくるが、おおむねボラティリティーがいちばん高い銘柄は富士通、低いのは北陸電力であるので、この2銘柄を対象として分析を進める。

3. 分析対象株価データ

3.1. 原データ

われわれの分析対象の原データは東洋経済新報社（2003）から富士通株と北陸電力株であり、対象期間は1992年1月第1週から2002年12月第4週までの573週足4本値（始値、高値、安値、終値）である。単に終値だけではないのは始値、高値、安値は経路情報を与えるからである。

3.2. 原データの性質

まず、11年間の大局的相場変動を終値のグラフで見ると、図2の通りである。比較のために TOPIX も補っている。

4 ソニーと電力3社株は対象期間内で株式分割を行っているが、分割後の株式価値にあわせ分割以前の株価を調整して求めた。配当金は収益率計算において考慮しなかった。

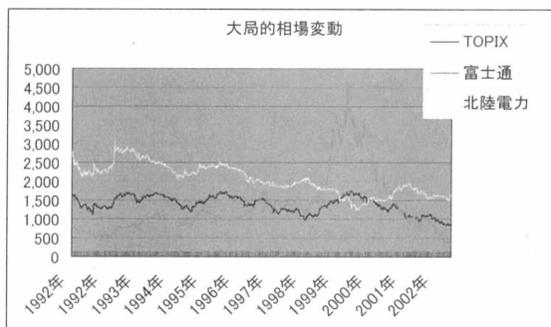


図3 富士通、北陸電力株の値動き（週足）

図3で富士通株は1999年くらいから2,000円を超え急上昇し4,660円のピークをつけた後で2000年に急落するといった明らかなトレンドを示しているのに対し、北陸電力株は安定的である。しかし、富士通株でも他の期間はどうなのか、局面が変わることにかかわるパラメータはどのようなものがあるかなど大局的株価変動評価には困難な問題点を残す。そこで、われわれはせいぜい過去4期前、1ヶ月前までの局所的な期間に限定し、研究を進める。

次に、週足終値の株価収益率 $\ln(P_{t+1}/P_t)$ をとることで、トレンドを除去した標本分布データが得られる。この分布が正規分布に近ければ、ある程度ランダムな挙動を示していることが示唆される。そこで、2節のリストした企業の株も含め、それぞれ算出した平均値、標準偏差で基準化を行い、歪度、尖度を求めた数値が表3のパネルA、 χ^2 検定、コルモゴロフ・スミルノフ検定で正規分布の適合度の検定を行った結果⁵をパネルBに記載した。

パネルAから伺えるのは歪度、尖度も正規分布に近い分布を示す銘柄はなさそうである。また電力業界株はどれも高い尖度を示しているのが特筆され

5 Excel アドオンソフト Crystal Ball® を使用して求めた。Crystal Ball® は Decisioneering 社の製品であり、日本における発売元は（株）構造計画研究所である。

表 3 各銘柄の株価収益率分布の特徴

パネルA											
	富士通	北陸電力	TOPIX	ソニー	シャープ	三洋電機	松下	東芝	日立	東電	関西電力
歪度	0.3264	1.0445	0.1896	-0.0306	-0.0762	0.3643	0.1925	0.1048	0.1546	2.2363	1.7188
尖度	1.4915	10.5233	0.7465	2.7162	0.7073	0.9858	0.2361	0.5483	0.7096	20.4605	16.9450

パネルB											
	富士通	北陸電力	TOPIX	ソニー	シャープ	三洋電機	松下	東芝	日立	東電	関西電力
χ^2 値	47.0262	142.925	20.7609	35.9442	29.5742	41.8778	35.1588	28.4398	27.3927	83.1518	81.8429
(p値)	0.0015	0.0000	0.5356	0.0308	0.1292	0.0065	0.0037	0.1616	0.1967	0.0000	0.0000
K-S値	0.0573	0.0824	0.0266	0.0518	0.0375	0.054	0.038	0.0377	0.047	0.0865	0.0691

る。パネル B の適合度の検定では χ^2 検定における p 値は正規分布とみなせる確率を意味するが 0.5 を超えているのは TOPIX のみであり、コルモゴロフ・スミルノフ検定でも 0.03 未満であることが判断基準⁶であるが、これも松下、東芝が比較的近いものの TOPIX のみが充足する。われわれの研究対象の個別株は候補も含め、その株価収益率が正規分布を示していないと結論づけることができる。しかし、だからといってすべてトレンド性があるとは言えないので、以下ニューラルネット手法を用いて検討する。

3.3. 加工手順

われわれは高見・上坂 (2004) 同様、富士通製のニューラルネットソフトウェア Neurosim/ L V4 を使用するが、そのためには Excel データベースの加工を行う必要がある。その手順を富士通株の例をとって以下説明する。北陸電力株の場合も全く同様である。

表 4 原データの加工手順その 1 (富士通株の例)

	始値	高値	安値	終値	上影	下影	実体
19920101	844	850	802	814	6	12	-30
19920102	804	819	791	800	15	9	-4
19920103	801	825	785	787	24	2	-14
19920104	790	804	781	804	0	9	14
19920201	804	809	790	790	5	0	-14
.....							
20021204	339	358	336	351	7	3	12
20021205	346	346	339	339	0	0	-7

6 エバンズ・オルソン (1999, p.103) を参照。

原データは1992年1月第1週から2002年12月第5週までの574週の4本値で表4の太線部分の各行の4種類のデータが各週の値を示している。この4種類の値から罫線情報を表す上影，下影，実体を週ごと計算する。計算手順は以下の通りである。

上影 = 高値 - max(始値, 終値), 下影 = min(始値, 終値) - 安値,
 実体 = 終値 - 始値

表5 原データの加工手順その2 (富士通株の例)

rand()	日付	IN	IN	IN	IN	IN	IN	IN	IN	IN	IN	IN	OUT
		上影t-4	下影t-4	実体t-4	Ct-4	上影t-3	下影t-3	実体t-3	Ct-3			IN	OUT
0.245645	19920101	6	12	-30	814	15	9	-4	800	*		790	755
0.119709	19920102	15	9	-4	800	24	2	-14	787	*		755	745
0.962256	19920103	24	2	-14	787	0	9	14	804	*		745	712
0.999407	19920104	0	9	14	804	5	0	-14	790	*		712	690
0.750911	19920201	5	0	-14	790	3	0	-31	755	*		690	658
.....													
0.528578	20021103	16	14	18	389	14	16	47	438	*		336	351
0.969179	20021104	14	16	47	438	11	17	-27	416	*		351	339
この行 以下カッ トされる	20021201	11	17	-27	416	5	2	-37	374	*			
	20021202	5	2	-37	374	2	3	-34	336	*			
	20021203	2	3	-34	336	7	3	12	351	*			
	20021204	7	3	12	351	0	0	-7	339	*			
	20021205	0	0	-7	339								

同じ点線ブロックをコピーする

次に、表4から上影，下影，実体，終値 (Ct) の列の組み合わせを作成し、表5の点線枠のように、1期ごとずらして重ね合わせ合計20列の入力データを作成、また、終値だけでt+1期に該当する出力データ列も作成する。すなわち表5で1992年1月第1週の行ではt-4期のデータは表4の1992年1月第1週の間から作成されているのに対し、t-3期のデータは表4の1992年1月第2週の間から作成され、t-2期のデータは表4の1992年1月第3週の間から作成されるなど同じ要領で1期ごとずらしている。その結果データ行数は569行に減少する。表5で見出し行のINとあるのはニューラルネット入力層のデータであることを示し、OUTとあるのは出力層のデータであることを示す。1列目のrand()とあるのは次のステップで並べ替えるために用いる便宜である。

Neurosim/ L V4を作動させる最後の準備として、表5の各行のデータを表

6のようにアットランダムに並べ替える（そのため第2列の日付はアットランダムに並べ替えられている）。

表6 原データの加工手順その3（富士通株の例）

rand()	日付	IN 上影t-4	IN 下影t-4	IN 実体t-4	IN Ct-4	IN 上影t-3	IN 下影t-3	IN 実体t-3	IN Ct-3	.	IN Ct	OUT Ct+1
0.001541	20000101	120	80	-1090	3820	110	50	-200	4000	.	3760	3730
0.003382	19951205	8	11	13	1505	22	24	8	1488	.	1523	1465
0.004644	19971104	10	20	50	1430	20	10	30	1450	.	1400	1400
0.005048	20021004	3	17	-51	437	10	12	-44	393	.	389	438
0.005771	20010103	20	0	261	2080	55	12	-150	1910	.	1742	1685
.....
0.99704	19951102	10	20	-40	1190	10	20	-10	1170	.	1240	1180
0.999383	19980302	10	20	-50	1340	10	20	-20	1320	.	1450	1475

表6のどの行のデータも同一の内容のデータがもれなくダブリなく表5の行のいずれかにある。表5の第1行に一様分布関数からランダムに発生された数値が表示してあるが、この数値が小さいものから並べ替えることで、各行のデータ全体を並べ替えることができる仕組みになっている。アットランダムに並べ替える目的は、われわれは局所的な相場変動を調べることにあり、たとえば1999年から2000年まで乱高下を示した富士通株のような過去からの大局的相場変動の影響を除去するためである。一方、4節で述べるようにニューラルネットではすべてのデータを学習用データと検証用データに分割するが、もし、学習用データにだけにそれらの範囲のデータが含まれるだけなら、検証力が非常に弱くなる。そこで、恣意性を避けつつ乱高下期のデータができるだけ学習用データと検証用データに共通して含まれるようにランダムに並べ替えた。われわれはこの並べ替え処理を複数回行うことで原データのもつ構造を分析して行く⁷。

7 Kamijo & Tanigawa (1990) は分析の対象こと異なるが、16種のパターンのうち15種を学習用データにとり、残りの1種を検証用データにとって相互に適合度を検証している点でわれわれのアプローチと共通する。

4. ソフトウェア Neurosim/ L V4の処理

4.1. 処理のプロセス

ニューラルネット分析にあたり、われわれは複数のソフトウェアを調べたが、操作が容易な富士通（株）製作の Neurosim/L V4 for Windows を用いることにした⁸。また中田（1992）は Neurosim を用いた先行研究⁹であり、広く実務で使われていることも理由のひとつである。Neurosim/L V4の操作は図4の手順を踏む。各プロセスの概略は以下の通りである。一見して、「2.学習パラメータの設定」いかんで満足できる結果が得られるかどうかが決まる。この点では確かにある程度の職人芸が要求されるが、試行錯誤を続けて行くしかない。つまり、「3.学習実行」でソフトが起動しているとき、収束率とよばれる設定された許容誤差にモデルが納まる確率値の挙動、ネットワーク図などを参考に計算している範囲、スピードなどをパラメータで微調整して行くのである。その収束率は70%以上が目処とされる（田辺（2003），p.31）。「4.学習結果出力」の段階で満足が行っても、「6.相関係数出力」で学習用データと検証用データとの相関係数に大きなブレがある場合は「2.学習パラメータの設定」に戻りプロセスを続ける。

8 田辺（2003）は Neurosim の学習書である。

9 Kimoto et al. (1990) も富士通（株）のサポートがあり、Neurosim を使用しているものと推察される。

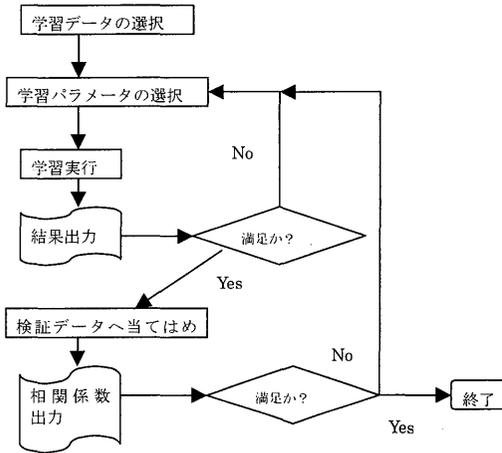


図4 Neurosim 処理計算手順

各プロセスの概要

1. 学習データの選択：表6データの中で最初から、398番目の行までを学習データとし、その部分のデータのみを読み込む。残り399番目から569番目の行までを検証用データとする。学習用データ7割、検証用データ3割にデータを分割するのは交差妥当性判定のためよく用いられる割合である¹⁰。
2. 学習パラメータの選択：中間層設定、逆伝播学習計算方法の設定（4.2節で詳しく述べる）
3. 学習実行：この過程はソフトの自動計算。収束の早さなどモニターできる（50,000回の演算は約3～5分かかる）
4. 学習結果出力：収束率、ネットワーク図などで確認できる

10 ベリー&ゴードン（2002）ではトレーニングセット、テストセット、評価セットと3分割し、6:3:1の比率を提唱している（p.257）。ただし、よく知られたツールは4:3:3とも述べている。豊田（2001）は実例として、2分割、7:3の比率を用いている（p.82）。われわれはデータを並べ替えることでトレンド除去しているので、3分割の必要はないと考え、ポピュラーな7:3の比率で2分割を採用した。

5. 検証用データへ当てはめ：4.で得られた重みの集合で理論値を計算する
6. 相関係数出力：学習用，検証用ともそれぞれ，理論値と実際値との相関係数が計算される

4.2. 学習パラメータの設定

われわれは試行錯誤で学習パラメータを調整した結果，表7の設定で両者の株式データを扱うことにした。

表7 学習パラメータ設定

学習パラメータ	意味	default値	設定値
変換後Max	シグモイド関数の最大値	1	0.9
変換後min	シグモイド関数の最小値	0	0.1
中間層の数	中間層のニューロン数	選択	2
学習回数	学習データを一巡計算する回数	10000	50000
許容誤差	学習を終了させる誤差の値	0.1	0.03
学習定数 s	重みの修正式に成長を抑制させる効果がある	0	0.00005
学習定数 ϵ	誤差曲線にける定数。重みの修正量の感度	5	3
学習定数 α	重みの修正式の慣性項	0.4	0.1
乱数	初期値ランダムシード	1	1

順を追って設定値の根拠を説明する。変換後の Max, Min は田辺 (2003, p.82) のように高い精度が要求される場合の設定である。もし入力データで最大値近辺の数値が集中していた場合アルゴリズムがスムーズに進まないおそれがあるからである。中間層の数2は最低限の数を選択した。ネットワーク図から判断して3以上に増やしても収束率はさほど改善せず，非線形の興味ある軸索結合はみられなかったからである。学習回数は途中で計算が終わることがみられなかったので，デフォルト設定の5倍の回数に変更した。許容誤差は高見・上坂 (2004) で触れた根拠を踏襲したが，TOPIX ヒストリカルボラティリティ 20.08%を前提に5営業日後理論的に着地する 1 倍シグマの範囲を目処としてい

る。デフォルトより厳格な設定である。学習定数 s の 0.0005 はある程度軸索の枝刈りが促進される意図で設けた。われわれの入力データは 20 種類もあるので、重要な入力データを目立たせる意図である。学習定数 ε と学習定数 α の組み合わせは計算がスムーズに行きつつもあまり計算対象の空間が動かないように配慮して設定した。乱数の設定はデフォルトのままである。

4.3. シミュレーション内容

われわれは 4.2 の学習パラメータ設定にて、富士通株、北陸電力株の表 6 フォーマットのデータをそれぞれ 10 通りの並べ替えを行い、合計 20 回のシミュレーションを行った。それぞれのシミュレーションで出力したアウトプットは収束率数値、学習用データと検証用データの相関係数数値、ネットワーク図、および重み集合数値である。5 節ではそのシミュレーション結果を紹介するとともに、考察を加える。

5. 分析結果

5.1. シミュレーション数値結果

われわれの行ったシミュレーション結果は表 8 の通りである。

表 8 シミュレーション結果

富士通											単位%	
シミュレーション(何番目)	1	2	3	4	5	6	7	8	9	10	平均値	中央値
収束率	89	95	88	88	93	90	87	93	93	92	90.8	91
相関係数												
学習用データ	99.4	99.4	98.7	99.3	99.2	99.2	99.3	99.2	99.3	99.2	99.22	99.25
検証用データ	98.4	98.4	99.1	98.6	99.2	98.9	98.6	99.1	98.4	98.8	98.75	98.7

北陸電力											単位%	
シミュレーション(何番目)	1	2	3	4	5	6	7	8	9	10	平均値	中央値
収束率	84	85	78	79	85	86	80	83	81	85	82.6	83.5
相関係数												
学習用データ	98.9	99.0	98.9	99.1	99.0	99.0	99.1	99.0	99.0	99.0	99.0	99.0
検証用データ	99.1	99.0	99.2	98.6	99.1	99.1	98.3	99.0	99.0	98.8	98.92	99.0

富士通株は北陸電力株よりどのシミュレーションでも収益率は高いが、両者とも目処である 70% をはるかに上回っている点で満足できる結果といえる。学

習パラメータ設定を変えることにより北陸電力株の収束率が向上することは考えられるが、両者同じ基準で比較することに意味があるので、行わなかった。次に、相関係数であるが、両者ともどのシミュレーションでもほとんど100%に近い数値であり、これも満足できる結果といえる。総じて、このシミュレーション結果は信頼に足るものである。

5.2. ネットワーク図

図5は富士通株の第1回目シミュレーションのネットワーク図で10回のシミュレーションのなかで9回見られた典型的なパターンである。

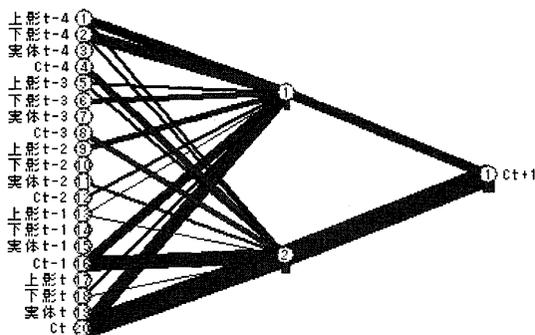


図5 ネットワーク図（富士通株シミュレーション1回目）

まず、10回通じて中間層ユニットから出力層ユニットに同一符号（濃い色のマイナス符号）の2本の太い軸索が描かれている。次に入力ユニットから中間層ユニットへの軸索であるが、まず枝刈りされずにかかなり残っている¹¹（40個のうち20個）。そのうち、太い軸索は Ct, Ct-1から2つの中間層ユニットへ出ている同一符号のものだが、上影 t-4, 下影 t-4, Ct-4, 上影 t-3, Ct-3, 実体

11 ただし、Neurosim では見やすくするために小さな数値の重みの軸索はカットするように表示されており、数値として必ずしも0をとってはいない。

t-2のように過去の終値あるいは罫線データから逆符号（薄い色のプラス符号）の軸索が伸びている。他のシミュレーションでもこのようなパターンが見られた（逆符号の発生している入力層ユニットが異なっている場合があるが、過去のデータから数多く軸索が伸びている点で共通点がある）。

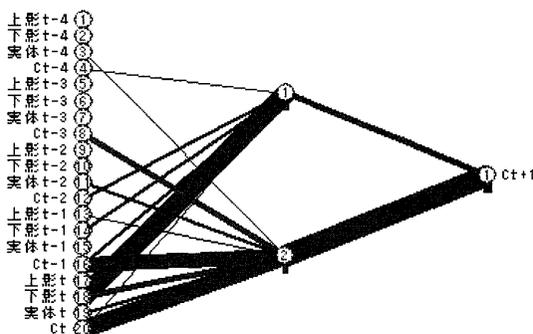


図6 ネットワーク図（富士通株シミュレーション5回目）

ところが、シミュレーションの中には例外的に図6のように入力層ユニットから伸びている軸索数が少なく（14個）、逆符号のものが少ないものも見られた。しかしこの場合も中間層ユニットから出力層ユニットに延びる軸索は2本とも残っている。

一方、北陸電力株からは対称的なネットワーク図が得られた。図7の第1回目シミュレーションのネットワーク図は10回すべてに見られる典型的な図である。すなわち、中間層ユニットのうち上部から出力層ユニットに向かう軸索は枝刈りされている。したがって、入力層ユニットから上部の中間層ユニットに向かう軸索があっても最終的に出力層ユニットに向かう影響度はほとんどない。その結果、入力層ユニットから下部の中間層ユニットに向かう軸索だけでほとんど説明されることになる。しかも、このうち Ct から発する軸索の太さが他を凌駕しており、結局 Ct と Ct+1 との対応、すなわち線形に近い関係を示唆している。

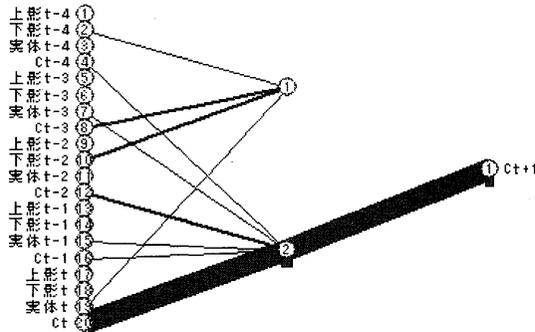


図7 ネットワーク図（北陸電力株シミュレーション5回目）

5.3. 重み集合数値結果

5.2節のネットワーク図の比較では軸索の太さを見ることで、両者株の特徴は対称的であることを概観したが、5.3節ではシミュレーション毎の重み数値データを用いて両者の違いを測定することを試みる。

まず、入力層 i ユニットから中間層 j ユニットへの軸索の太さを w_{ij} ($i=1,2,\dots,20; j=1,2$)

中間層 j ユニットから出力層ユニットへの軸索の太さを v_1 または v_2 とすると入力層 i ユニットが最終的に出力ユニットに向かう力の強さは符号も考慮して、 $w_{i1}v_1 + w_{i2}v_2$ の値が候補になる。しかし、ニューラルネットではシナプスの入力と出力との間に非線形なシグモイド関数で変換しているため不適切である¹²。

そこで、Neurosim のメニューに組み込まれた「影響度」を用いて分析することにした¹³。表9は両者の出力結果のサンプルとしてシミュレーションの2回目までを示している。

12 すなわち重回帰モデルのような線形モデルであれば係数がそのまま影響度を示すが、ニューラルネットのような非線形モデルでは係数の他にローカルな点の位置情報も加味しなければならない。

表9 影響度出力サンプル

シミュレーション(何番)	富士通		北陸電力	
	1	2	1	2
上影t-4	0.0859	0.0114	-0.0137	0.0000
下影t-4	-0.0335	0.0043	0.0191	0.0206
実体t-4	0.0231	-0.0405	-0.0340	-0.0028
Ct-4	-0.1577	0.0002	-0.0477	0.0007
上影t-3	-0.0192	-0.0415	0.0188	0.0213
下影t-3	-0.0482	-0.0221	0.0022	0.0021
実体t-3	0.0003	-0.0747	0.0438	0.0805
Ct-3	-0.1422	-0.2342	0.0000	0.0001
上影t-2	0.0766	-0.0126	-0.0132	-0.0126
下影t-2	-0.0044	-0.0440	-0.0207	-0.0602
実体t-2	-0.0942	-0.0847	-0.0178	-0.0084
Ct-2	0.0307	0.0710	0.0977	0.0632
上影t-1	0.0140	-0.0002	-0.0318	-0.0254
下影t-1	-0.0012	-0.0123	0.0261	0.0169
実体t-1	-0.0178	0.0064	0.0525	0.0031
Ct-1	0.5151	0.3881	0.0531	0.2684
上影t	-0.0164	-0.0046	0.0116	0.0054
下影t	-0.0109	0.0804	0.0177	0.0074
実体t	-0.0272	0.0216	0.0133	0.0290
Ct	0.7788	0.6794	0.7831	0.5392
合計	0.9516	0.6914	0.9603	0.9484

見出しの列で上影t-4から Ct までの20種類の入力ユニットの名称は図5から図7までのいちばん左のユニットの名称に対応する。しかし、この表では中間層を経由してどう伝わるかの情報が埋没されており、最終的に主力層ユニットにどのくらいのインパクトで伝わるか、符号かどうかを示すにとどまっている。合計の値は1にはならない。このように観察するとシミュレーションごとに合計の値が異なっていることから、加工して比較する必要がある。

そこで、符号情報は捨象して、まず各入力層ユニットは20個のユニット全体に対しどのくらいのインパクトを持っているか、絶対値を取り構成比を計算したのが表10である。

13 学習データに限り適用される尺度である。原理は入力層の値が出力層の値にどれだけインパクトがあるか、表6の入力行ごとに微係数をもとめそれらの平均をとった値である。入力値の値によって行ごと微係数値が異なるのでやや粗い方法とされる（田辺（2003, p.37））。

表10 影響度構成比（絶対値ベース％）

シミュレーション(何番目)	富士通		北陸電力	
	中央値	平均値	中央値	平均値
上影t-4	1.2%	1.7%	0.7%	0.6%
下影t-4	0.8%	1.1%	1.4%	1.2%
実体t-4	1.4%	1.9%	0.3%	0.7%
Ct-4	0.5%	1.4%	4.2%	4.3%
上影t-3	1.7%	1.7%	1.1%	1.1%
下影t-3	1.1%	1.3%	0.4%	0.7%
実体t-3	0.5%	0.8%	1.9%	2.4%
Ct-3	7.4%	8.4%	0.0%	0.1%
上影t-2	0.9%	1.1%	1.2%	1.3%
下影t-2	1.2%	1.3%	2.4%	2.7%
実体t-2	3.7%	3.8%	1.5%	1.5%
Ct-2	1.2%	2.1%	7.7%	7.7%
上影t-1	0.9%	1.3%	1.4%	1.7%
下影t-1	0.1%	0.5%	0.8%	1.0%
実体t-1	0.8%	1.8%	1.6%	1.7%
Ct-1	26.7%	26.1%	8.0%	9.8%
上影t	0.8%	1.6%	0.7%	0.8%
下影t	1.1%	1.8%	1.4%	1.3%
実体t	2.7%	3.5%	2.6%	2.6%
Ct	37.1%	36.8%	55.7%	56.9%
合計	100.0%	100.0%	100.0%	100.0%

5.2節と符合する点で、富士通株では過去の終値、罫線情報での影響度が北陸電力株より高い傾向にある（たとえば、以下中央値の比較でCt-3の7.4%、実体t-2の3.7%）。また、北陸電力株ではCtが全体の55.7%を占めるのに対し、富士通株では37.1%に過ぎず、同じ終値情報でも1期前のCt-1も26.7%と高い割合を加味している（北陸電力は8.0%）。反面、過去の情報でも北陸電力株の方がウェイトの高いユニットもあり、いちがいに言えない。

次に、5.2節では符号も考慮に入れてネットワーク図を観察したが、絶対値を取らずに構成比を計算した結果が表11である。

表11 影響度構成比（符号考慮ベース％）

シミュレーション(何番 目)	富士通		北陸電力	
	中央値	平均値	中央値	平均値
上影t-4	2.3%	3.2%	0.0%	-0.3%
下影t-4	0.8%	-0.2%	2.0%	1.7%
実体t-4	2.2%	1.8%	-0.3%	-1.0%
Ct-4	0.0%	-1.0%	-6.0%	-6.4%
上影t-3	-2.9%	-3.1%	1.7%	1.5%
下影t-3	-2.5%	-2.2%	0.4%	0.7%
実体t-3	0.0%	-0.9%	2.5%	2.8%
Ct-3	-12.5%	-18.3%	0.0%	0.0%
上影t-2	0.3%	0.7%	-1.7%	-1.8%
下影t-2	-2.0%	-2.8%	-3.5%	-3.7%
実体t-2	-6.3%	-7.4%	-2.2%	-2.2%
Ct-2	1.9%	2.9%	10.3%	11.2%
上影t-1	2.0%	2.4%	-2.1%	-2.4%
下影t-1	-0.1%	-0.2%	1.2%	1.3%
実体t-1	0.7%	1.7%	0.1%	-0.7%
Ct-1	45.4%	49.5%	12.7%	13.8%
上影t	-1.2%	-2.4%	1.0%	1.2%
下影t	1.5%	1.2%	1.7%	1.3%
実体t	4.6%	4.5%	3.2%	2.8%
Ct	64.9%	70.7%	82.1%	80.2%
合計	100.0%	100.0%	100.0%	100.0%

表11は表10に比べてやや分かりにくいですが、過去の入力層ユニットのなかで違いが顕著なのはCt-3である。富士通株は12.5%のインパクトを持って全体の力の逆方向の影響度があるのに対し、北陸電力株では全くインパクトがない。実体t-2も同様である。また表10と同様にCtの集中度は北陸電力が格段に高い(64.9%に対し82.1%)。ただし、北陸電力株でもマイナス符号の過去の入力層ユニットは多く5.2.節で感覚的に考察したほどの同一符号集中度は確認できない。

6. 結論

われわれは株式個別銘柄によって株価変動の経路依存性はどの程度あるか、どう観察・測定するかを問題と捉え、ニューラルネット手法を用い、局所的に個別銘柄を比較することを試みた。われわれの研究は限定的である。選択した個別銘柄は富士通株と北陸電力株の2銘柄であり、Neurosimのシミュレーショ

ンも合計20回にしか過ぎない。ただし、銘柄選択はできるだけ恣意性を排除することを試み候補を絞って行ったうえでの選択であるし、シミュレーションもネットワーク図で確認しながらの試行錯誤であり、実際はパラメータ設定を変えることで100回以上は行っている。

われわれが得たインプリケーションは富士通株では北陸電力株に比べて当週の終値だけでなく過去の終値、罫線情報も加味して翌週末終値が決まる傾向にあり、より経路依存性が見られるというものである。また過去のデータのうち Ct-3と実体 t-2が顕著であった。両者とも逆符号であることも特徴である。たとえば3週間前の終値水準が相対的に安ければ安いほど、また t-2期が下落相場であればあるほど、翌週終値は高くなるという傾向を示している。これらインプリケーションは罫線情報で天井や大底を予想する経験知がある程度銘柄によって有効であることを示唆するものである。

しかし、解決すべき課題はまだ多く残されている。本質的な問題ではニューラルネットは最適解を一意的に求める手法ではないので、大まかな傾向があることだけしか言えない。現にわれわれは2銘柄10回ずつのシミュレーションを行ったが、各軸策の太さ・符号はシミュレーションによって、すなわち学習データの集合によって異なっていた。これはニューラルネットの宿命ではあるものの、大まかな傾向がより精緻な傾向になるようにデータ採択の基準を見直すべきである。また、20種類の入力データは独立して同時に作用することが前提とされている。ところが、実務上の罫線に基づく相場予想判断はたとえば「3期続けて連騰したときは・・・」というように因果律を暗黙の前提として、条件つきモデルを考慮しなければならないだろう。その解決方法は職人芸から判断したパターンを学習データにとりモデルを形成する方向が考えられる。

また、技術的問題としてわれわれは Neurosim ソフト自体のロジックを十分検討していない。入力層ユニットの影響度は平均値数値のみでは満足できない。ソフトを改造してローカルな位置情報も分析するのも課題である。

最後に、問題を拡げるようではあるが、やはり2銘柄比較は研究の入り口に

過ぎない。経路依存性の測定技術の精緻化とともに、業種特性、業績データなども加味しつつ分析を拡張して行きたい。

参考文献

Dacco, Robert and S. Satchell (1999) "Why Do Regime-Switching Models Forecast So Badly?" *Journal of Forecasting*, Vol.18, p.1 - 16.

Kamijo, Ken-ichi and T. Tanigawa (1990) "Stock Price Pattern Recognition: A Recurrent Neural Network Approach," *Proceedings of the IEEE International Joint Conference on Neural Networks*, San Diego, CA, p.1215 - 1211.

Kimoto, Takashi, K. Asakawa, M. Yoda and M. Takeika (1990) "Stock Market Prediction System with Modular Neural Networks," *Proceedings of the IEEE International Joint Conference on Neural Networks*, San Diego, CA, p.11 -16.

Law, Averill M. and W. D. Kelton (2000) "Simulation Modeling and Analysis, third edition.," MacGraw Hill.

Yoon, Youngohc and G. Swales (1991) "Predicting Stock Price Performance: A Neural Network Approach," *Proceedings of the IEEE 24th Annual International Conference of Systems Sciences*, p.156 - 162.

ブロックウェル, P.J., R.A.デービス (2000) 「入門 時系列解析と予測」(逸見等訳) CPA 出版

馬場則夫 (1998) 「ニューラルネットを活用した株式売買意思決定支援システムの構築」 *MTEC Journal*, Vol.11, p.3 - 41.

ベリー, マイケル J. A., ゴードン・リノフ (2002) 「マスタリング・データマイニング 理論編」(江原淳他6名共訳, 海文堂, 原題: "Mastering Data Mining," John Wiley and Sons, Inc.

エバンス, ジェームス, デビット・オルソン (1999) 「リスク分析・シミュレーション入門」(服部正太監訳, 共立出版, 原題: "Introduction to Simulation and Risk Analysis," Prentice Hall, Inc.)

藁谷千風彦 (2001) 「金融データの統計分析」 東洋経済新報社

中田匡 (1992) 「様々な相場展開に対応できるニューロ応用相場予測システム」 日経インターネットシステム 別冊1992夏号, p.64 - 71.

日本テクニカル・アナリスト協会編 (1992) 「日本の株価分析 第3版」 日本経済新聞社

高見茂雄・上坂博亨 (2004) 「TOPIX 相場変動のニューラルネットによる可視化」 日本経営財務研究学会第28回全国大会, 平成16年10月16日

田辺和俊 (2003) 「Neurosim/L によるニューラルネットワーク入門」 日刊工業新聞社

豊田秀樹 (2001) 「金鉱を掘り当てる統計学」 講談社

東京証券取引所 (2004) 「TOPIX β Value」 Vol.27 (2004年9月)

東洋経済新報社 (2003) 「株価 CD-ROM 2003」 東洋経済新報社

吉富康成 (2002) 「ニューラルネットワーク」 朝倉書店

原稿提出日：2004年11月30日