

系列データマイニングによる
株価変動パターンの抽出

東京工芸大学工学部コンピュータ応用学科

宇田川 佳久

1

目次

1. 研究の背景と概要
2. これまでの株価変動分析
3. 類似株価の探索手法の提案
4. 実験結果
5. まとめと今後の研究方針

2

1. 研究の背景と概要

- 株価の安定は、企業活動の安定、ひいては一国の経済活動の安定に繋がる。
- 一方、株価の変動は、財政・金融政策、政治的安定など多くの要因が絡み合って決まる複雑なもの。
- 株価変動のメカニズムを解明し、経済社会の持続的な発展のためのモデルを確立することが本研究の目的である。

3

研究の方針

- 営業日数(シーケンスの長さ)と変動率をパラメータとする類似株価変動パターンの探索手法の提案
- 探索実験結果
 - 類似パターンが検索できるか？
 - 数日後の株価を予測できるか？
- 株価と景気変動を関連つける研究であり、
- 景気変動を実感するのは人間であることから、分析対象は、日経平均株価の終値とする。

4

目次

1. 研究の背景と概要
2. これまでの株価変動分析
3. 類似株価の探索手法の提案
4. 実験結果
5. まとめと今後の研究方針

5

2. 従来の株価変動分析

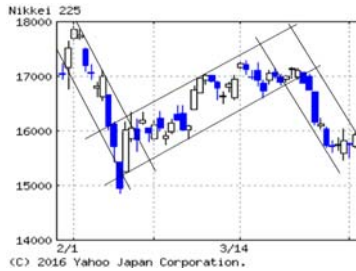
- ファンダメンタル分析
 - 企業の財務分析⇒株価の割高性
- テクニカル分析
 - (前提1) 投資家の行動を反映した結果として株価が決まる。
⇒ 投資家を動かした要因を分析しなくても、株価を分析すれば、将来を含めた株式市場の状況が分かる
 - (前提2) 株価は一定のトレンドを形成する。
⇒ 投資パターンがある
 - (前提3) 株価は繰り返す。

テクニカル分析方法が異なると、将来の予測が逆であることも珍しくない ⇒ 改善の余地がある。

6

トレンドチャネル

- 株価は、一定期間上昇あるいは下降する。
- 要因の例
 - 政府による経済活動への刺激策
 - 米国・欧州における景気動向
 - 地域紛争



7

移動平均線

- 一定期間(5日, 25日, 75日など)の移動平均線を見ると一定の傾向を見出すことができる。



8

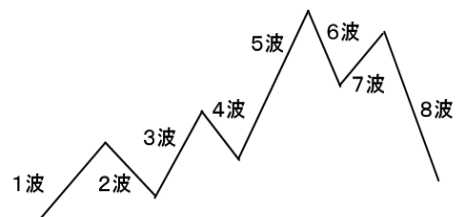
移動平均線

- グランビルの法則
 - 株価上昇判断: 下落していた移動平均線が横ばいから上昇に転じ、株価が移動平均線を超えたとき、以降、株価が上昇する。
 - 株価下降判断: 移動平均線が上昇後、横ばいになるか下降に転じ、株価が移動平均線を下回ったとき、以降、株価が下降する。
- ゴールデンクロス: 上昇基調で、短期線と中期線が交差すること(傾きが正)
- デッドクロス: 下降基調で、短期線と中期線が交差すること(傾きが負)

9

エリオット波動法則

- 以下に示したパターンを繰り返すとする法則(1934年頃)
- 上昇または下落した幅のフィボナッチ数列から導かれる値(0.382, 0.618, 1.618など)に収まることが多い。



10

株価変動の例



11

目次

1. 研究の背景と概要
2. これまでの株価変動分析
3. 類似株価の探索手法の提案
4. 実験結果
5. まとめと今後の研究方針

12

数式による定義

- 前日比 ($i \geq 1$) (高値低値圏の影響を除く)

$$R_i \equiv (p_i - p_{i+1}) / p_{i+1}$$

- 株価シーケンス変動率(N, S) $\equiv \sum_{i=1}^N |R_{S+i} - R_i|$

株価シーケンス変動率とは、基準となる株価シーケンス R_i ($i=1, \dots, N$) と S ($S \geq 1$) 営業日離れた株価シーケンスとの変動率の差の合計値と定義する。

13

類似株価の探索手法の提案

株価シーケンス変動率(N, S) とは



14

数式による定義

- 類似株価シーケンス集合SSSP(Set of Similar Stock Prices)とは、**株価シーケンス変動率が許容値 ϵ 以下である株価シーケンスの最初の営業日の日付の集合**

$$SSSP(N, \epsilon) \equiv \{ \text{Date}(S) \mid$$

$$S \in \{S \geq 1\} \wedge \exists S \left(\sum_{i=1}^N |R_{S+i} - R_i| \leq \epsilon \right) \}$$

15

数式による定義

- K 営業日前の株価を基準とした、類似株価シーケンス集合SSSP(Set of Similar Stock Prices) の定義

$$SSSP(N, \epsilon, \text{Date}(k)) \equiv \{ \text{Date}(S) \mid S \in \{S \neq 0 \text{ である整数} \}$$

$$\wedge \exists S \left(\sum_{i=k}^{N+k-1} |R_{S+i} - R_i| \leq \epsilon \right) \}$$

16

目次

1. 研究の背景と概要
2. これまでの株価変動分析
3. 類似株価の探索手法の提案
4. 実験結果
5. まとめと今後の研究方針

17

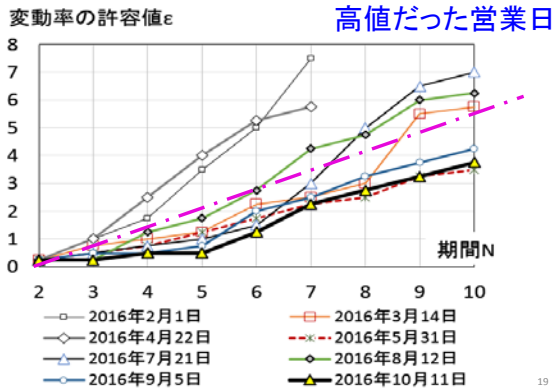
4. 実験結果

- 目的: 期間と変動率に対する**類似株価パターン**の発生数の関係を定量的に明らかにする。
- 基準日: 高値だった8日間
- 方法: 期間 N を 2 から 13日, 変動率 ϵ を 0.25% から 7.5% まで 0.25% 刻みで変化させて, 類似株価パターンを探索した。
- 株価データは, 2010年8月2日から2016年10月26日までの1,529件

$N \setminus \epsilon$	0.25	0.5	0.75	1	1.25	1.5	1.75	2	2.25	2.5	2.75	3	3.25	3.5	3.75	4	4.25	4.5	4.75	5	5.25	5.5	5.75	6	6.25	6.5	6.75	7	7.25	7.5								
2	24	62	144	232	367	491	601	721	840	952	1053	1122	1167	1197	1219	1233	1242	1248	1251	1253	1254	1255	1256	1257	1258	1259	1260	1261	1262	1263	1264	1265	1266	1267	1268	1269	1270	
3	11	10	25	55	100	165	253	349	433	533	648	738	823	903	978	1048	1113	1173	1228	1278	1323	1363	1403	1443	1483	1523	1563	1603	1643	1683	1723	1763	1803	1843	1883	1923	1963	2003
4	0	2	4	12	32	58	93	130	184	242	309	390	473	558	643	728	813	898	983	1068	1153	1238	1323	1408	1493	1578	1663	1748	1833	1918	2003	2088	2173	2258	2343	2428	2513	
5	0	1	2	2	4	13	30	37	69	96	124	167	219	271	323	375	427	479	531	583	635	687	739	791	843	895	947	999	1051	1103	1155	1207	1259	1311	1363	1415	1467	1519
6	0	0	0	0	1	2	6	11	19	34	49	71	90	109	128	147	166	185	204	223	242	261	280	299	318	337	356	375	394	413	432	451	470	489	508	527	546	565
7	0	0	0	0	0	0	0	0	2	4	6	10	12	16	20	24	28	32	36	40	44	48	52	56	60	64	68	72	76	80	84	88	92	96	100	104	108	112
8	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
9	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

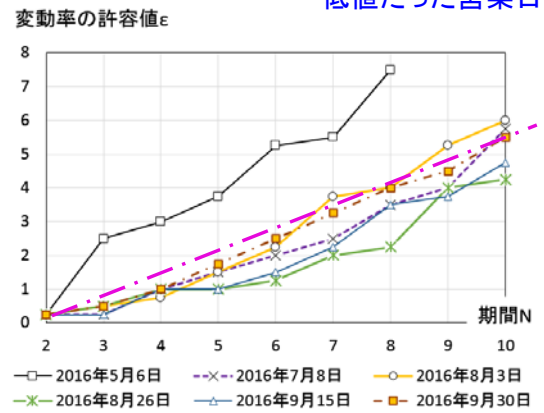
18

最初に類似パターンが見つかったときの期間N(X軸)と変動率 ϵ (Y軸)についてのグラフ



19

低値だった営業日

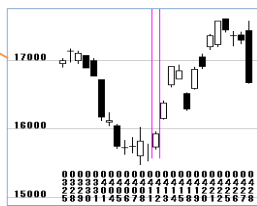
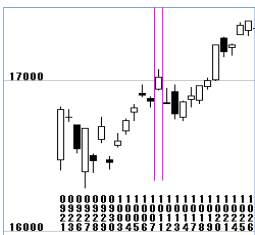


2016/12/3

20

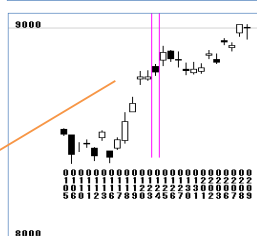
類似パターンが検索できるか？

N=4 で検索された例



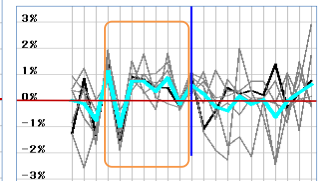
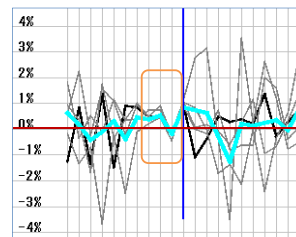
N=7 で検索された例

2016/12/3



数日後の株価を予測できるか？

N=7の方がN=4に比べてわずかに良いが、差はほとんどない。



2016/12/3

22

5. まとめ

- 従来のテクニカル分析手法を概観した。
- 新たな探索手法として類似株価シーケンスの集合 SSSP(Set of Similar Stock Prices)を提案した。
- 最初に類似パターンが見つかったときの期間 N と変動率の許容値 ϵ は、5日から10日間の範囲で直線傾向を示す。
- 高値圏、低値圏で同じ傾向がある。
- 直近の株価変動を予測することに関しては、N=7の方がN=3に比べてわずかに良い(明確ではない)。

23

今後の研究方針

- 株価を特徴付けるパラメータを増やした検索方法を開発する。パラメータとしては、始値と終値の差、高値、安値、前日終値と始値の差(いわゆる窓明け)などを対象とする。
- 日経平均株価に関し、株価の変動傾向(上げ局面、下げ局面、迷いなど)分類に沿った追加実験を行う。
- 投資家の行動を動かす要因は、テクニカル分析が対象としていた株価変動だけではない。政治・経済状況が株価変動に与える影響について研究する。

謝辞

- 匿名の査読者の皆様に感謝いたします
- 本研究は、JSPS科研費 基盤研究(C)一般 JP16K00161 の助成を受けたものです。

24