

# 経験知識にもとづく時系列データの予測システム

高野 義士†

田中 靖士†

杉村 博‡

松本 一教†‡

† 神奈川工科大学情報学部情報工学科

‡ 神奈川工科大学大学院情報工学専攻

## 1 はじめに

本論文では、時系列データを対象として、経験知識を活用して、時系列データの動向を予測するシステムを提案する。時系列データとは、株価の変動、プラントのセンサデータ、毎日の気温や視聴率、商品販売数など、時間軸で連続しているデータのことであり、産業界のあらゆる場面や日常に多数ある。これらデータを解析することは、未来予測や市場調査を行う上で重要である。時系列データの予測を行うシステムは以前から提案されている。数学的なデータ解析手法を用いる手法 [1] や AI の技術を用いる手法もある [2]。

一方、各々の時系列データに対して、様々な経験的な知識が蓄積されている場合が多く、それらが活用されている。たとえば株価データの場合には、チャートの分析により多くの経験則が知られている。このような各分野での経験知識をシステムに組み込んで、自動的に活用することが期待される。

本論文では、典型的な時系列データである株価データを用いて、その経験知識試作システムにより、活用のための課題を検討し評価する。

## 2 知識表現とマッチング

時系列データとは、実数値の列  $x = (v_1, v_2, \dots, v_x)$  のことである。時系列データ上の知識とは、時系列データのパターン  $p = (u_1, u_2, \dots, u_p)$  のことをいう。ただし、一般的にパターンは比較的短いものであることが多い。時系列データの処理では、各時刻での値そのものが重要となる場合もあるが、値の変化の割合や率が重要となることも多い。

本論文では、率を重要視する立場とする。従って各データは、適当な方法によりある範囲に正規化されて

### A system for analysis of time-series data based on empirical knowledge

†Yoshiaki TAKANO, †Yasushi TANAKA,

‡Hiroshi SUGIMURA and †‡Kazunori MATSUMOTO

†Department of Information and Computer Sciences, Kanagawa Institute of Technology, ‡Course of Information and Computer Sciences, Kanagawa Institute of Technology

1030 Shimo-ogino, Atsugi-shi, Kanagawa 243-0292, JAPAN

{s055090, s055101}@cce.kanagawa-it.ac.jp,

hiroshi.sugimura@gmail.com, matumoto@ic.kanagawa-it.ac.jp

いるとして扱う。このような正規化されたパターンはグラフ形式で可視化することができ、たとえば図1ようになる。2つのパターン  $p_1, p_2$  が与えられたとき、その等価性を判定することは重要である。この目的のために、DPW マッチングの手法を用いることができる。

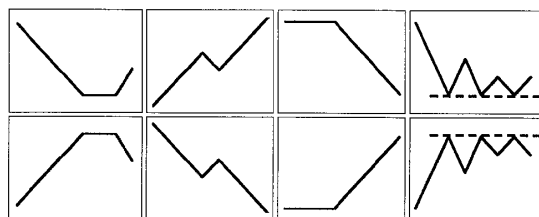


図 1: 8 種の参照パターン

DTW(Dynamic Time Warping) マッチングとは、パターンの要素間に定義された類似度にもとづいて、パターンの伸縮まで考慮に入れたマッチング方式である [3]。この方法ではいくつかのパラメータを事前に定めて、パターンの類似度  $s(p_1, p_2)$  を計算できる。ここで、 $p_1$  と  $p_2$  の長さが異っていても構わない。与えられた実数値  $\delta$  に対して  $s(p_1, p_2) < \delta$  であれば、 $p_1$  と  $p_2$  が一致すると判定する。 $\delta$  のことを許容相異度と呼ぶ。このようにして、時系列上のパターンを比較する方法を定めて実装できる。

## 3 予測システム開発と評価

### 3.1 システムの概要

時系列データに対する経験知識は様々なものがあり、知識の使われ方も様々である。

本論文では、実験として株価チャートに関する多数ある知識の中から最も良く知られている売買タイミングに関する経験知識を用いることにした [4]。この知識は、図 1 に示す 8 種のパターンからなる。上段の 4 種が買いパターンであり、下段の 4 種が売りパターンと呼ばれる。買いパターンチャートがこのような動けば、今後の株価が上昇すると予想して買いのタイミングとみなす。売りパターンならばその逆である。

このパターンの利用方法の概念を図2に示す。

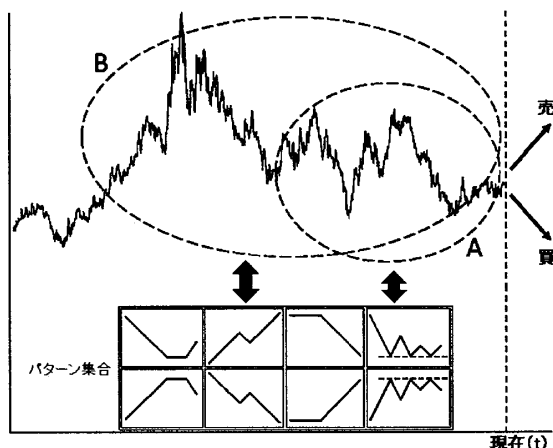


図2: パターンの利用方法

各々のパターン  $p_i$  に対して、現在 ( $t$ ) より過去の振舞  $X$  と比較する。このとき、どの程度の長さの過去の振舞に注目するかが問題となる。すなわち、図中の  $A$  に示すように  $X$  を短く取るか、 $B$  のように長く取るか決める必要がある。この研究では後に述べる様に実験的に検討した。

#### 4 実験方法

1990年から2003年までの13年間の東京証券取引場でのデータを用いて実験を行う。株式の売買に関する実験なので、売買シミュレーションが必要となる。今回は図3に示す最も基本的で単純な売買方法のみを実施した。

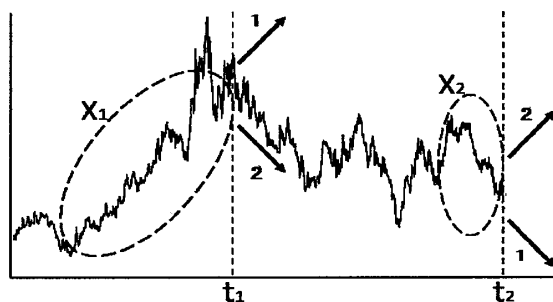


図3: 売買方法

時点  $t_1$  において、その直前の振舞  $X_1$  が、経験知識の買い(上昇)パターンとマッチした場合には(図中の1の状況)、このタイミングで  $t_i$  の時点での価格で1株購入する。その後、時点  $t_2$  において、直前の振舞  $X_2$  が売り(下降)パターンとマッチした場合には(図中の

1の状況)、この時点の価格で1株のみを売却する。なお、いずれの場合においても、売買の手数料は考慮しない。信用売りの利用やポートフォリオ構築などは今回は行っていない。

#### 4.1 実験結果

同業種からランダムに選んだ6社を対象に、短期振舞(100日)と長期振舞(200日)での売買シミュレーションによる利益の結果を表1に示す。期間の選択には、一般的に株価チャート分析で用いられる移動平均の13週(91日)と26週(182日)を参考にした。利益とは、1990年に株価取引を開始した時点での資本金が2003年には何倍になっているかを示す。

表1: 売買による利益の結果

会社名	利益(100日)	利益(200日)
A社	5.206	4.461
B社	1.339	0.295
C社	-3.054	-2.415
D社	1.529	1.425
E社	1.021	1.440
F社	0.167	0.950
平均	1.035	1.026

#### 5 おわりに

本論文では、株価データを用いて、経験知識の試作システムにより、時系列データの予測を試みた。

今後の課題として、システムに使用する経験知識の選択、マッチング精度の改善、振舞期間の調整などが挙げられる。

#### 参考文献

- [1] James Douglas Hamilton, Time Series Analysis, Princeton University Press, 1994.
- [2] Nils J. Nilsson, Artificial Intelligence : A New Synthesis, Morgan Kaufmann Publishers, 1998.
- [3] Donald J. Berndt, James Clifford, Finding Patterns in Time Series : A Dynamic Programming Approach, Advances in Knowledge Discovery and Data Mining, pp.229-248, 1996.
- [4] 星和貴, 買い時・売り時がピタリとわかる株価チャートの読み方, 明日香出版社, 2002.