

ハンケル動的モード分解を用いた車業界群の株価予測

日大生産工(学部) ○小森 麗菜
日大生産工学部 佐々木 真

1. まえがき

株価の動向の予測は金融取引戦略において重要な役割を持つ。市場の動向をいかに正確に把握し高い投資収益率を生み出すかが焦点となる。しかし、株価の変動は多数の要因により動きが複雑であるため予測が難しい。これらの解析・予測の手法は複数存在しており、長期短期記憶(Long Short-Term Memory,LSTM)¹⁾というリカレントニューラルネットワーク(Recurrent Neural Network,RNN)の一種や、1980年以降価格予測で頻繁に使われるARCH(Autoregressive Conditional Heteroskedasticity)モデル⁵⁾、テキストデータを用いた株価予測⁴⁾やそれらを複合した手法などがある。LSTMでの株価予測は従来のRNNの問題点であった長期依存の学習を可能にし、より精度の高い予測ができる¹⁾。

このように株価時系列データに対する予測研究が進んでいる。ただ、多くの場合、ある特定の業界に属している企業の株価はある程度、相関して変動すると期待される。業界群内の変化を系統的にモデル化出来れば、より効果的な予測となる可能性がある。

本研究では動的モード分解(Dynamic Mode Decomposition,DMD)という実例が少ない解析手法を使い予測精度を高めることを定量化する。

2. 動的モード分解の概略

動的モード分解(DMD)は流体などの時間の進みによって変化する複雑なデータを解析する手法の一つである²⁾。これを用いて複数の株価の変動を解析して次にどのような動きをするのか予測する。

動的モード分解は時空間観測データから直接データが従う発展方程式を推定する手法である。支配方程式は形式的に以下のようになる。

$$\frac{dX}{dt} = F(X, t) \cong AX \quad (1)$$

ここで、 X が観測値、 $F(X,t)$ は X の発展を決める非線形方程式である。ここで非線形演算子 F を線形演算子 A で印字するという過程をおく。すると、式(1)より A は次のように求めることができる。

$$A = \frac{dX}{dt} X^{-1} \quad (2)$$

このように観測値のみを使って時間発展演算子を推定することが、いったん A が求まれば X が次のように決まる。

$$X = X(0)e^{At} \quad (3)$$

以上のように観測データのみから時間発展の予測が可能となる。この一連の手法をDMDと呼ぶ。

このような標準的な動的モード分解は線形近似値を用いるため、長時間の予測が困難であった。そのため、ハンケル行列とスパース化を併用することでより少ない誤差で長時間の予測を可能とする手法が提案された⁶⁾。

3. 車業界群の株価推移

本研究で自動車関連の会社を中心に、85個の株価の2021年1月1日から2023年12月31日までの二年間データを解析に使用した。対数化した株価をグラフ化したものをFig.1に示す。



Fig. 1 株価のグラフ

株価の平均値をそれぞれ出した時の最小値は120.61円、最大値は67823.55円となっている。そのままの株価では価格変動が大きく、予測精度が低くなる可能性があるため、株価を対数化している。

これらの獲得した株価データから二つのハンケル行列を作った。ハンケル行列は入力する時系列データをスライド幅 s を Δt ずらしたデータを h 層積み上げたものである。通常の対角成分と垂直方向の対角線と並行となる行列成分がすべて等しくなる正方行列である。この行列に動的モード分解を適用することでモード数が増える。モード数の増加はデータの空間情報を見かけ上増やすことができ、時間・空間情報量の差異を抑制し、長時間の時空間的な特徴を捉えることが可能になる。

この時stackとslideというハイパラメータの数値を操作することでハンケル行列の大きさを変えることができる。

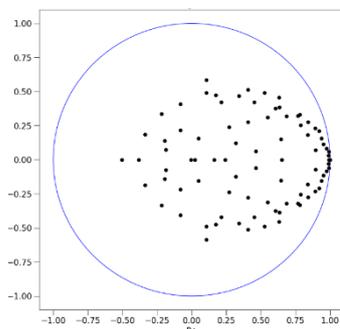


Fig. 2 固有値分解

1に近いほど影響の大きいモードである。

次に時間発展DVとモードFDを示すグラフを実部と虚部に分けて抽出した。時間発展DVでは各モードがどのように発展するかを示しており、モードFDでは複数のモードが時間に沿ってどのように進展するかを示す。

その後オリジナルデータとDMDによる再現を出力して比較し、誤差をプロットした。

最後に評価関数に $l1$ の正規化を加える。正規化とは、目的変数に寄与しない関数の係数を小さくして実質的な特徴量の数を減らし、過学習を抑える手法である。 $l1$ 正規化を用いることでスパース化を行う。再度時系列データの予測と予測評価を行った。

4. 実験結果

実際のデータとDMDでによる予測はFig.3のようになった。また、読み込んだ株価群の推移と未来予測をFig.4に示す。

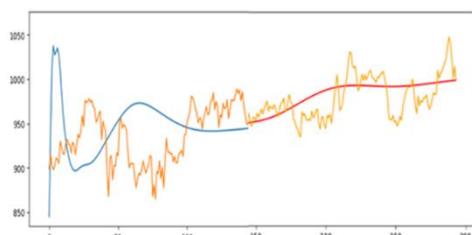


Fig. 3 実際のデータと DMD 再現と予測

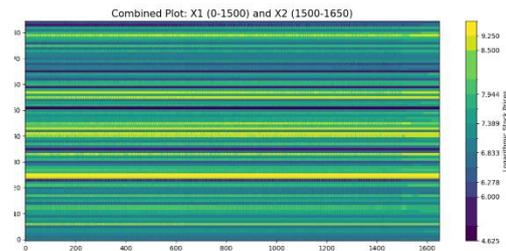


Fig. 4 株価全体の推移

Fig3上図のオレンジの線が実際の株価の推移で青い線が読み込んだ株価データの再現、赤線が未来予測した部分である。

Fig4は横軸の0から1500までが実際に予測に利用した範囲で、1500から1650が未来予測の範囲である。

青線の再現は序盤に変動が激しく再現が上手くできていないが、予測の再現になると再現のものと同様の傾向のものになった。

5. まとめ

本研究では、拡張型モード分解を用いることで株価データの時系列予測を行えると考え研究を行った。投資では高い利益率を出すことが必要となる。株価予測を活用することで金融業界のさらなる発展の助力となる。今回の研究から株価の変動の大きな傾向を再現した。今後さらに予測精度を高めることで繊細な株価の変動の再現を行う。

参考文献

- 1) 高橋健太,藤戸四恩,古川隆証,法政大学経営システム工学令和3年度卒論概要集,(2021)
- 2) 佐々木真,河原吉伸,草場彰, J. Plasma Fusion Res. Vol.97, No.2 (2021),pp.79-85
- 3) 塚越友希, 日本大学生産工学部電気電子工学科令和4年度卒業論文,(2023)
- 4) 井口勝太,湯本高行, DEIM2022 B23-1,(2022),pp.7
- 5) 北田宙士,柘植綾香,菅井友駿,北栄輔, 日本機械学会計算力学講演会講演論文集,第28巻,(2015),No.183
- 6) M.R.Jovanović,P.J.Schmid,and J.W.Nichols,Phys.Fluids26,024103 (2014)
- 7) 小山直紀,山田朋人,山田正, 河川技術論文集,第30巻,(2024),pp.459-464