1B-05

リカレントニューラルネットワークによる ボラティリティ・クラスタリングのモデル化

五島　圭一†　　高橋　大志‡　　寺野　隆雄§

† 東京工業大学 大学院総合理工学研究科　　‡ 慶應義塾大学 大学院経営管理研究科

§ 東京工業大学 情報理工学院

1　はじめに

金融資産のリスク指標であるボラティリティには，自己相関があることが広く知られており，ボラティリティが上昇（下降）するとしばらくボラティリティの高い（低い）期間が続く．この現象はボラティリティ・クラスタリングと呼ばれ，どのようにモデル化をして，予測に用いるかは金融経済学や金融工学における重要な研究課題の一つである．なぜならば，金融資産のボラティリティを予測することは，リターンの区間予測することに繋がり，また，保有している金融資産がどれだけリスクにさらされていて，どのように変化していくかを把握することを可能にする [3]．そのため，これまで様々なボラティリティ変動モデルが開発されてきた．

ボラティリティ変動モデルの一つに ARCH 型モデルがある．ARCH 型モデルは，将来の条件付きボラティリティについて，過去の価格に関する情報を用いて予測が行われる．そして，予測を行うために有用な変数（特徴量）の選択は，人手で設計されることが多い．例えば，GARCH モデルでは，過去の条件付きボラティリティと過去の予測誤差が用いられる．また，レバレッジ効果を加味した GJR-GARCH・EGARCH や長期記憶性を加味した FIEGARCH などがある [4]．

このように，研究者や実務家が現実のマーケットを観察し，当てはまりのより良いモデルを開発してきた．この課題に対して，近年注目を集めているディープニューラルネットワーク [2] を用いれば，これまで人手によって行っていた特徴量の選択や特徴量同士の組み合わせを自動でモデリングを出来る可能性がある．そこで本研究では，ディープニューラルネットワークの中でも，時系列データへの適用に関して，大きな成果を挙げているリカレントニューラルネットワークを用いて，ボラティリティ・クラスタリングのモデリングを試みる．具体的には，RNN (Recurrent Neural Network) 及び派生モデルである LSTM (Long Short-Term Memory) による日経平均株価の条件付きボラティリティの予測タスクを通じて探求する．

Modeling the Volatility Clustering with Recurrent Neural Networks
†Keiichi GOSHIMA　‡Hiroshi TAKAHASHI　§Takao TERANO
†Interdisciplinary Graduate School of Science and Engineering, Tokyo Institute of Technology
‡Graduate School of Business Administration, Keio University
§School of Computing, Tokyo Institute of Technology

2　実験設定

2.1　データ

日経 NEEDS から 2000 年 1 月から 2016 年 4 月までの日経平均株価の日次データを用いた．一般的に，t 期におけるリターン y_t は，(1) 式のように定義することが出来る．

$$y_t = \mu_t + u_t = \mu_t + \sigma_t \epsilon_t \tag{1}$$

ここで，μ_t は期待収益率，u_t は予測誤差，σ_t はボラティリティ，ϵ_t は確率変数である．渡部・佐々木 (2006)[5] では，日経平均株価のリターンは自己相関がないことや平均値がゼロであることから，ARCH 型モデルによる推定の際には，1 期前において予測可能な変動 μ_t をゼロとしている．本研究でもこれに倣い，μ_t はゼロとする．

2.2　モデルの概要

本研究のモデルは，RNN 層及び LSTM 層において，k 期から $t-1$ 期までにおける各期の予測誤差 $(u_k, u_{k+1}, \cdots, u_{t-1})$ の二乗を入力として受け取り，全結合層に渡した後，恒等関数によって，t 期のボラティリティσ_t の二乗を出力する関数である．このとき，全結合層にて，Dropout 及び ElasticNet による正則化を行う．RNN 層及び LSTM 層内部の活性化関数は，tanh 関数を用いる．加えて，Bi-directional RNN 及び Bi-directional LSTM を用いる．モデルパラメータの最適化手法には Adam を使用し，ミニバッチのサイズは 50，epoch 数は 50 とした．ただし，Early-stopping を採用したため，3 回続けて二乗誤差が良くならなかった場合，そこで学習を打ち切るようになっている．モデルの実装には，Keras* を用いた．RNN によるボラティリティ・クラスタリングモデルの関数を $f_{RNN-ARCH}$，LSTM によるボラティリティ・クラスタリングモデルの関数を $f_{LSTM-ARCH}$ とす

*https://keras.io/

ると，(2) 式と (3) 式のように表すことが出来る．

$$\sigma_t^2 = f_{RNN-ARCH}(u_k^2, u_{k+1}^2, \cdots, u_{t-1}^2) \quad (2)$$

$$\sigma_t^2 = f_{LSTM-ARCH}(u_k^2, u_{k+1}^2, \cdots, u_{t-1}^2) \quad (3)$$

本研究では，モデルのハイパーパラメータについては，ElasticNet の l1l2 は [(1.0, 1.0)，(10.0, 10.0)，(100.0, 100.0)] から探索し，Dropout レートは 0.3 とする．隠れ層の数は 100，200，500 をそれぞれ分析案 d する．

2.3 検証方法

前節のモデルを用いて，2015 年 5 月から 2016 年 4 月までの条件付きボラティリティの予測タスクに取り組む．本実験では，k から $t-1$ までを 100 期とする．例えば，2016 年 4 月 1 日の条件付きボラティリティを予測する場合は，2015 年 11 月 4 日から 2016 年 3 月 31 日までの 100 営業日間の予測誤差を用いることになる．そして，モデルのパラメータは，2016 年 3 月 31 日以前のデータを用いて推定する．

真なるボラティリティは，通常観察されるものではないので，実現ボラティリティやインプライド・ボラティリティ，ヒストリカル・ボラティリティなどを真なる値として ARCH 型モデルは評価されることが多い．本研究では，t 期のボラティリティσ_t については，予測誤差 u_t を用いるものとする．すなわち，$t-1$ 期以前のデータについて，教師スコア Y は u_{t-k}^2，トレーニングデータ X は $(u_{t-k-100}^2, u_{t-k-99}^2, \cdots, u_{t-k-1}^2)$ となるようなデータセットを作成し，$f_{LSTM-ARCH}$ 及び $f_{RNN-ARCH}$ を推定し，そして，推定した関数に $(u_{t-100}^2, u_{t-99}^2, \cdots, u_{t-1}^2)$ を入力することで，t 期のボラティリティσ_t を予測する．

2015 年 5 月から 2016 年 4 月まで 244 営業日間について，毎期ごとにモデルのパラメータを再推定し，翌期のボラティリティを予測する．トレーニングデータ数は，t 期において用いることの出来るすべてのデータを使用するため，2000 年 1 月以降のすべてを用いることになり，毎期モデルを推定する際には，1 ずつ増えることになる．このとき，トレーニングデータのうち，ランダムに抽出した 10%をデベロップメントデータとする．

推定したボラティリティの予測精度について，GARCH(1,1) モデルとの比較を通じて，実験結果を考察する．GARCH(1,1) モデルは，Hansen and Lunde (2005)[1] において 330 の ARCH 型モデルを比較したところ，最良のモデルであると報告されているためである．GARCH(1,1) モデルについても，毎期ごとにモデルのパラメータを再推定し，ボラティリティの予測をする．モデルの予測精度の評価は，(4) 式のように推定したボラティリティσ と真なるボラティリティh との平均二乗誤差 (MSE) によって行う．ここでも，真なるボラティリティh は，予測誤差 u とする．

$$MSE \equiv \frac{1}{n}\sum_{t=1}^{n}(\sigma_t^2 - h_t^2)^2 \quad (4)$$

3 実験結果

表 1 は，モデル別の MSE をまとめたものである．表 1 を見ると，RNN と LSTM について，GARCH(1,1) と同等の予測精度であることを示している．中には，GARCH(1,1) よりも MSE が小さいモデルもあることが分かる．

4 おわりに

本研究では，リカレントニューラルネットワークを用いて，ボラティリティ・クラスタリングのモデリングを試みた．実験の結果，RNN や LSTM を用いたモデルは，GARCH(1,1) モデルと同等の予測精度が得られる可能性を示した．他の期間での実験や日経平均株価以外の金融資産の実験，より精緻なパラメータの探索について，今後の課題である．

参考文献

[1] P. R. Hansen and A. Lunde. A forecast comparison of volatility models: does anything beat a GARCH(1,1)? *Journal of Applied Econometrics*, Vol. 20, No. 7, pp. 873–889, 2005.

[2] 岡谷貴之. 深層学習. 講談社, 2015.

[3] 沖本竜義. 経済・ファイナンスデータの計量時系列分析. 朝倉書店, 2010.

[4] 渡部敏明. ボラティリティ変動モデル. 朝倉書店, 2000.

[5] 渡部敏明. Arch 型モデルと realized volatility によるボラティリティ予測とバリュー・アット・リスク. Vol. 25, No. 2, pp. 39–74, 2006.

表 1: モデル別の MSE

モデル	GARCH(1,1)	RNN-ARCH			LSTM-ARCH		
隠れ層数		100	200	500	100	200	500
MSE	39.70	40.36	38.65	39.23	39.92	40.85	40.15