

リカレントニューラルネットワークによる為替レートの予測

伊藤 克輝^{†1} 飯間 等^{†2} 北村 能寛^{†3}京都工芸繊維大学 情報工学課程^{†1} 同 情報工学・人間科学系^{†2}早稲田大学 社会科学総合学術院^{†3}

1. はじめに

為替レートの決定に重要な役割を担うものとして、価格を指定せずに売買量を注文する成行注文が理論、実証的に注目されてきた。これに対して、価格を指定する指値注文の役割は未だ十分な研究蓄積がない。そこで、本論文では指値注文数が価格順に記録されている注文板の情報を活用した為替レートの予測を試みる。その際、従来よく用いられている線形モデルを用いた方法で高い予測精度を得ることは困難であると考えられる。一方、非線形モデルで為替レートのような時系列データを予測するリカレントニューラルネットワークが近年注目を集めている。そこで本論文では、リカレントニューラルネットワークである LSTM^[1]や GRU^[2]を用い、その入力を注文板の情報として5分毎の為替レートを予測する方法を提案する。また、これらの入力情報の中から為替レートの決定に特に重要な入力を特定する。

2. 提案する為替レート予測方法

注文板の情報を入力とするリカレントニューラルネットワークを用いて5分毎の為替レートの増減を予測する。単位時間を5分とした各離散時刻 t における為替レートの増減予測に用いる入力を、時刻 $t-1$ から時刻 t までの5分間に発生した種々の取引数、指値注文数、指値注文キャンセル数からなる11次元ベクトル $X_t = (x_{t1}, x_{t2}, \dots, x_{t11})$ とする。これらの詳細を表1に示す。ここで、気配値が良いとは、買い指値が高いことや売り指値が低いことを意味する。リカレントニューラルネットワークでは、直近 L 個の入力ベクトル X_t 、すなわち $X_{t-L+1}, X_{t-L+2}, \dots, X_t$ を順番に入力して時刻 t での予測を行う。

具体的なリカレントニューラルネットワークとして、LSTM と GRU を用いる。LSTM は、長期の依存性も短期の依存性も学習できることから注目されているモデルである。また GRU は、LSTM を少し単純化したモデルである。ネットワークの出力は為替レートの増減であるため、出力層の活性化関数はシグモイド関数、損失関数

は二値クロスエントロピーとする。出力が 0.5 より大きければ増加、小さければ減少と予測する。最適化手法には Adam^[3]を利用する。また、過学習の防止策として、10 エポック連続でテストデータに対する誤差が減少しなかった場合に学習を打ち切るという方法を採用する。

表1. 入力データ

x_{t1}	最良気配値の変化を伴う取引数
x_{t2}	最良気配値が変化しない取引数
x_{t3}	最良気配値を更新する指値注文数
x_{t4}	最良気配値での指値注文数
x_{t5}	最良気配値での指値注文のキャンセル数
x_{t6}	最良気配値を悪化させる指値注文のキャンセル数
x_{t7}	2番目に良い気配値での指値注文数
x_{t8}	2番目に良い気配値での指値注文のキャンセル数
x_{t9}	3番目以降に良い気配値での指値注文数
x_{t10}	3番目以降に良い気配値での指値注文のキャンセル数
x_{t11}	最良気配値外での取引数

3. 為替レート予測実験

実際の為替取引情報を用いた実験を行い、提案法の有効性を明らかにする。

使用するデータは2018年10月1日から2019年1月31日までの EUR/USD, USD/CNH, USD/JPY の取引情報(EBS Data Mine Level 2.0)であり、それぞれ前半80%をトレーニングデータ、残り20%をテストデータに分別した。

提案法では、隠れ層の数や各層の素子数、学習率、 L の値などを種々に変更して予測を行った。提案法による予測結果を、従来よく用いられている線形モデルと比較して提案法の有効性を検証する。

LSTM, GRU 両方のモデルで予測を行った結果、隠れ層の数は1層とした場合に速く最も高い正解率が得られた。また、その際の素子数は50とした場合に比較的良い正解率が得られたため、以降の実験では隠れ層は1層、素子数は50とした。表2に、EUR/USD のデータに対して、学習率と L の値を変化させたときの、テストデータとトレ

ーニングデータでの正解率と学習終了時のエポック数 N を示す。

表 2. 正解率と学習終了時のエポック数 N

	学習率	L	正解率		N
			テスト	トレーニング	
LSTM	0.001	5	0.730	0.762	18
		10	0.734	0.766	20
		25	0.734	0.768	22
	0.0001	5	0.735	0.755	103
		10	0.730	0.758	106
		25	0.738	0.757	101
GRU	0.001	5	0.731	0.762	25
		10	0.735	0.764	24
		25	0.734	0.768	25
	0.0001	5	0.733	0.755	122
		10	0.738	0.758	90
		25	0.738	0.758	104

表 2 より、いずれのパラメータを用いても正解率は 73% 以上と高い水準を獲得できていることがわかる。学習率が 0.001 の場合に比べて 0.0001 とした場合では学習終了時のエポック数 N が 5 倍程度増加するため、学習率は 0.001 が良いといえる。また、同様の理由でネットワークは LSTM を使い、 $L = 10$ とするのが良い。他の為替データに対しても同様の結果が得られた。

線形モデルを用いた場合の正解率は 0.509 となった。提案法での正解率が 0.73 程度であるため、提案法が有効であるといえる。

4. 重要な入力データの特定

表 1 に示した入力データの中から、為替レートの決定に特に重要な入力を特定する。この目的のために、11 個の入力データ中のいくつかの入力のみを用いて学習させる。入力の組合せは 2047 通りあり、これらすべての組合せで個別に学習させる。学習の結果、正解率の高い組合せに数多く出現する入力が必要であると考えられる。学習時のモデルとパラメータは、3 章で最も良い結果が得られたものを用いた。

図 1 は、2047 通りの組合せをテストデータの正解率順で並べ、その上位 20%、40%、60%、80%、100% で各入力は何回ずつ現れているかを表したものである。棒グラフは 11 本ずつ並んでおり、左から $x_{t1}, x_{t2}, \dots, x_{t11}$ の出現回数を表している。図 1 より、上位 20% までに x_{t3} と x_{t6} が数多く出現しており、これらの入力が重要であるといえる。特に x_{t3} は上位 40%、60% でも出現率が高い。最

も正解率が高かった組合せは $x_{t1}, x_{t2}, x_{t3}, x_{t5}, x_{t6}, x_{t8}, x_{t11}$ であり、その正解率は 0.750 であった。この値は表 2 の正解率よりも高くなっている。

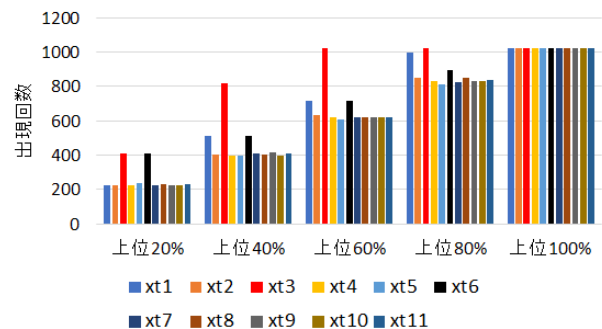


図 1. 各入力データの出現回数

5. おわりに

指値注文の情報に注目し、リカレントニューラルネットワークを用いて為替レートを予測する方法を提案した。提案法では、従来よく用いられている線形モデルと比較して高い正解率が得られた。また、注文板の情報の中で、最良気配値を更新する指値注文数と、最良気配値を悪化させる指値注文のキャンセル数が、為替レートの決定に重要であることを明らかにした。

謝辞

本研究は JSPS 科研費 JP19H01508、19K21704、公益財団法人全国銀行学術振興財団の助成を受けたものである。

参考文献

- [1] Sepp Hochreiter and Jürgen Schmidhuber (1997) "Long short-term memory", *Neural Computation*, Vol.9, No.8, pp.1735-1780.
- [2] Kyunghyun Cho, Bart van Merriënboer, Caglar Gulcehre, Dzmitry Bahdanau, Fethi Bougares, Holger Schwenk, and Yoshua Bengio (2014) "Learning Phrase Representations using RNN Encoder-Decoder for Statistical Machine Translation", arXiv:1406.1078.
- [3] Diederik P. Kingma and Jimmy Ba (2015) "Adam: A method for stochastic optimization", *Proceedings of International Conference on Learning Representations*.