

# 強化学習を用いた金融市場取引の取引ルール獲得に関する研究

温井 慧, 徐 春暉、安藤 雅和

千葉工業大学大学院社会システム研究科

**Abstract:** AI 画像認識や AlphaGo に代表させるゲーム AI などの成功により機械学習が注目を浴びている。またその高い精度から様々な分野への応用が進んでいる。特に自動運転技術やゲーム AI の成功により強化学習の応用が活発になってきている。現在、機械学習を用いて金融市場取引における戦略の獲得に関する研究も活発化している。本研究では、外国為替証拠金取引（以下 FX）における強化学習を用いて取引ルールの獲得を目指す。

## 1. はじめに

1996 年の外国為替取引の完全自由化により FX 取引が誕生した。現在では、50 万人を超えるトレーダーがいるとされている。FX は、24 時間取引可能なことや手数料が低いことにより日本では、人気の投資先となっている。FX は、証拠金（保証金）を業者に預託し、為替変動によって売買益をだす。差金決済が可能なことから多額の利益を得ることができる反面、過大な損失を被ることもある。そのためトレーダーは、取引ルールを設定しそれに従い取引することでこのリスクを回避している。しかしこの取引ルールは経験など主観によるものが多くトレーダーの実力に左右される。そこでコンピュータを駆使し自動的にルールに従い取引をするシステムトレードを導入し機械的に取引をする投資家も増加している。また昨今の AI ブームから人工知能を導入し価格の予測、戦略を獲得する研究も行われている。

本研究では、予測値、市場の状態、投資家の状態をもとに強化学習を使用し取引を学習させる。取引内容を分析し取引ルールを構築していく。本稿では、先行研究の紹介、手法の提案を行う。

## 2. 研究目的

本研究では、FX における取引ルールの獲得を目指す。以下のルールを決定することが本研究の目的である。

- 買い注文のルール  
 $x$  分後の予測値が  $y\%$  と上昇と予測が出れば買い注文。
- 売り注文のルール  
買い注文を行った  $x$  分後に売り注文を決定するかどうか。  
予測値が外れた場合、売り注文を行うかどうか。

## 3. 先行研究

現在、取引戦略の獲得に強化学習を応用する研究が活発になってきている。松井らは深層強化学習を拡張した複利型深層強化学習の提案を行い日本国債における取引戦略の獲得を目指し利益率を向上させた。[1]。また和田らは、深層強化学習を用いて株価、テクニカル指標などの時系列情報から戦略の獲得を目指し高い利益率をだした[2]。上記のほかにも先行研究が存在する。また高い利益率を出している研究も多い。

本研究は、これらの先行研究の内容を参考に

学習を進め取引ルールの獲得を行っていく。

## 4. 提案手法

### 4.1 先行研究との相違

3章で提案した手法は、強化学習を使い戦略の獲得を目指している。しかし戦略および売買ルールに関して明示的なパラメータが存在していない。本研究では、実験後取引の分析しパラメータを明らかにしていく。また多くの先行研究では、状態にテクニカル指標などの分析指標を与えるものが多いが予測値を与えるものは少ない。

### 4.2 取引対象

本研究では、円ドル相場を使用し学習を行う。また使用する値は、みずほ銀行が公開している日次データを使用する。

使用する期間は、2017年1月4日から2019年12月30日のデータを使用する

### 4.3 予測手法

予測モデルは、時系列モデルであるARIMA過程を使用する。

本研究では、過去1年分のデータを使用し1日後の価格を

$$y_t - y_{t-d} = c + \varepsilon_t + \sum_{i=1}^p \varphi_i y_{t-i} + \sum_{i=1}^q \theta_i \varepsilon_{t-i}$$

を解くことによって求められる。モデルの選定は、AIC（赤池情報量基準）を使用する。検定の結果使用するモデルは、ARIMA(2, 1, 7)となった。予測手法は、1日後の価格を前日までの価格を使用する。予測結果は、図1のようになった。

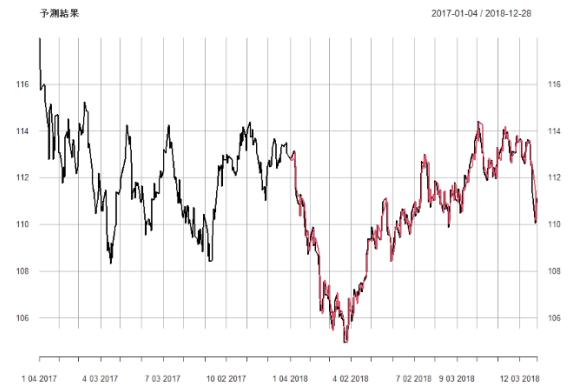


図 1 予測結果

使用した期間は、2017年1月4日から2018年12月28日までである。予測期間は、2018年1月4日から12月28日である。また予測精度は以下のようになった。

表 1 予測精度

ME	RMSE	MAE
-0.0094	0.4748	0.3750

MPE	MAPE
-0.0094	0.3401

### 4.4 深層強化学習

強化学習とは、機械学習の手法の一つである。強化学習では、与えられた環境の中で価値（利益）を最大化させるようにエージェントを学習させていく。この手法は、様々な技術に導入されておりゲームのプレイや株式取引の戦略などに導入されている。代表的な学習手法としてQ学習がある。

しかしこの手法では、状態が多い場合学習を進めることが困難になる。そこでQ学習に用いるパラメータの一つであるQ値をニューラルネットワークで近似する深層強化学習がある。本研究では、この深層強化学習のアプローチの一つである深層Qネットワーク（DQN）を使用する。

#### 4.5 状態

以下の三項目を環境からエージェントが受け取る状態  $s_t$  と定義する。

- ① 為替データ (USDJPY)
- ② 一日後の予測値
- ③ 所持金 (USD, JPY)

#### 4.6 行動

以下の三項目をエージェントが行う行動  $a$  と定義する。

- ① Hold
- ② Buy
- ③ Sell

#### 4.7 報酬

学習の目的は、取引初日と取引最終日の日本円を比較したときの利益率が高いことである。しかし最終日に報酬を与えると取引をしない可能性が発生する。したがって売り注文を実行した時点で利益に応じ報酬  $r_t$  を与える。

$$r_t = \begin{cases} \frac{P_{Sell} - P_{Buy}}{P_{Buy}} \times N_{Sell} & a = Sell \\ 0 & a = Buy \\ 0 & a = Hold \end{cases}$$

ここで  $P_{Sell}$  は、売り注文時の日本円、 $P_{Buy}$  は、買い注文時の日本円、 $N_{Sell}$  は、売り注文を行った米ドルの価格である。

#### 4.8 学習設定

訓練期間は、2018年1月4日から12月28日の一年間、テスト期間は、2019年1月4日から12月30日の一年間になる。訓練期間で取引をすべて終えるまでを1エピソードとする。

#### 4.9 分析

学習後の取引内容を可視化し分析し第2章で述べた取引ルールを決定する。

#### 5. 研究結果と考察

発表時に詳細な結果と考察を述べる。

#### 参考文献

- [1] 松井藤五郎, 片桐雅治, 金融取引戦略獲得のための複利型深層強化学習, 人工知能学会
- [2] 和田裕貴, 長尾智晴, “深層強化学習による株式売買戦略の構築”, 情報処理学会第79回全国大会 2017
- [3] 小林弘幸, 和泉潔, 松島裕康, 坂地泰紀, 島田尚, “強化学習による高頻度取引戦略の構築”, 第24回 人工知能学会 金融情報学研究会 2020
- [4] 松井 藤五郎, 後藤 卓, ”強化分析を用いた金融市场取引戦略の獲得と分析” 人工知能学会誌 24巻3号 2009
- [5] 常井祥太, 穴田一, 深層強化学習を用いた投資戦略の構築, 情報処理学会第80回全国大会, 2018
- [6] 岡谷 貴之, ”深層学習” 講談社 第18刷発行 2018
- [7] 斎藤 康毅, ”ゼロから作る Deep Learning Python で学ぶディープラーニングの理論と実装” オライリージャパン 初版第9刷発行 2017
- [8] Nikhil Buduma 著, 太田 満久, 藤原 秀平 監訳, 牧野 聰訳, ”実装 Deep Learning Python と TensorFlow で学ぶ次世代の機械学習アルゴリズム” オライリージャパン 初版第1刷発行 2018
- [9] 小川雄太郎, つくりながら学ぶ! 深層強化学習 PyTorch による実践プログラミング” マイナビ出版 初版第3刷発行 2018
- [10] 定常時系列の解析に使われる ARMA モデル・SARIMA モデルとは?, AVILEN AI trend, [https://ai-trend.jp/basic-study/time-series-analysis/sarima\\_model/](https://ai-trend.jp/basic-study/time-series-analysis/sarima_model/)