6N-07

外国為替取引における通貨ペア選択を対象とした複利型強化学習と最適レバレッジの学習

塚本 智大 *　　松井 藤五郎 †‡

* 中部大学 大学院工学研究科 情報工学専攻　† 中部大学 生命健康科学部 臨床工学科

‡ 中部大学 工学部 情報工学科

1 はじめに

複利型強化学習 [1] は、報酬の替わりにリターン (利益率) を獲得するリターン型マルコフ決定過程において, 試行錯誤を通じてエージェントが将来に獲得する複利リターンを最大化する行動規則を学習する枠組みである。複利式のリターンを考える場合には，従来の強化学習のような期待割引収益の最大化は意味をなさないため、複利型強化学習では期待割引収益の替わりに二重指数関数を用いて割り引いた複利リターンの対数の期待値を最大化する。

FX と呼ばれる外国為替証拠金取引では、あらかじめ取引会社に証拠金を預け入れることで、証拠金の数倍から数百倍の金額で取引でき。これをレバレッジという。レバレッジとは、証拠金にかける「てこ」のようなもので、少ない資本金でも大きな取引ができる仕組みである。証拠金に対する取引可能金額の比率を、レバレッジの倍率という。

しかしながら、従来の複利型強化学習はレバレッジについて考慮いなかった。そこで我々は、複利型強化学習にレバレッジを導入し、レバレッジをかけた場合の複利リターンを最大化する方法 [2] について検討し、簡単な例題を用いてその有効性を確認した。本論文では、実際の外国為替市場のデータに基づいて通貨ペア選択問題を作成し、レバレッジを用いた複利型強化学習を適用してその有効性を検討する。

2 複利型強化学習とレバレッジの関係

複利型強化学習では、自分が保有する資産うちどれだけ投資するかを表す投資比率パラメーターが導入されている。

複利型強化学習は、将来にわたって得られる複利リターンを二重指数関数で割り引いたものの対数の期待値を最大化する行動規則を学習ものである。すなわち、複利型強化学習は以下の値を最大化する行動規則を学習する。

$$\mathrm{E}\left[\prod_{k=0}^{\infty}(1+l_{t+k+1}f_{t+k})^{\gamma^k}\right] \tag{1}$$

ここで、f_t は時刻 t における投資比率を表す。

従来の複利型強化学習は、投資比率が $0 \leq f_t < 1$ であり、レバレッジは考慮されていない。時刻 t における自己資本が p_t、レバレッジの最大倍率が l、投資比率が f_t のとき、最大で自己資本の l 倍の $p_t l$ まで投資することができ、投資比率を考慮すると実際の投資額は $p_t l f_t$ となる。

実際の投資額が自己資本を超える、すなわち $p_t l f_t - p_t = p_t(lf_t - 1) > 0$ のとき、自己資本を超える投資金は外国為替証拠金取引業者などから借りているものである。$lf_t <= 1$ のときは、投資金を借りる必要がないため、今後は $lf_t > 1$ のときについて考える。

時刻 t に実行した行動の結果として投資金に対するリターン R_{t+1} を獲得したとすると、投資金が $1+R_{t+1}$ 倍になり、投資金は $p_t l f_t(1+R_{t+1})$ になる。ここから、借りていた $p_t(lf_t - 1)$ を返済すると、手元には $p_t(1+R_{t+1}lf_t)$ が残る。自己資本は p_t から $p_t(1+R_{t+1}lf_t)$ に変化しており、自己資本に対する利益率は $R_{t+1}lf_t$ である。したがって、以下のように投資比率をレバレッジに置き換えることで複利型強化学習にレバレッジを導入することができる。

$$l_t = lf_t \tag{2}$$

従来の複利型強化学習では、オンライン勾配法を用いて投資比率 f_t を最適化する。レバレッジを導入した複利型強化学習においても、投資比率の最適化と同様にして、以下のようにしてレバレッジ l_t を最適化できる。

$$l_{t+1} = l_t + \eta_t \frac{R_{t+1}}{1+R_{t+1}l_t} \tag{3}$$

3 実験と考察

これまでの研究により、複利型強化学習にレバレッジを導入し、レバレッジを投資比率と同じようにオン

Compound reinforcement learning for currency selection in foreign exchange transactions and learning of optimum leverage

*Tomohiro Tsukamoto　†‡Tohgoroh Matsui

*Department of Computer Sceience, Graduate School of Engineering, Chubu University

†Department of Clinical Engineering, College of Life and Health Sciences, Chubu University

‡Department of Computer Science, College of Engineering, Chubu University

Copyright ©2017 Information Processing Society of Japan.
All Rights Reserved.

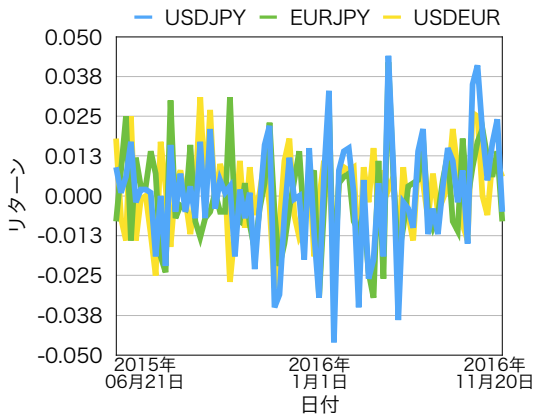

図 1: 為替取引データ

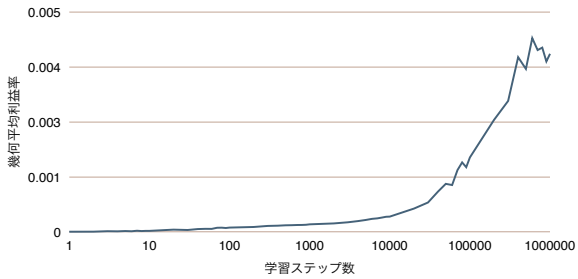

図 2: 幾何平均利益率

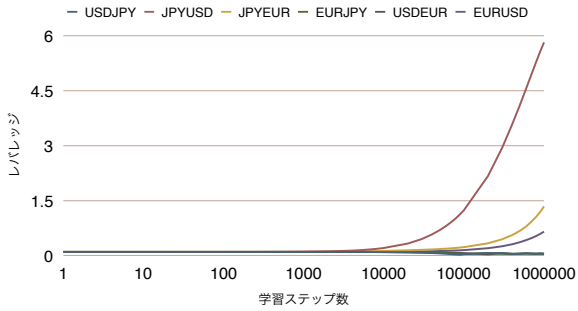

図 3: レバレッジ

ライン勾配法で最適化すると、最適なレバレッジを学習できるとわかった。しかし、これらの実験では、自分たちが定めた値のみの結果である。実際には、為替市場に応用できるものでなければならない。そこで実際の為替取引データを用いて幾何平均リターンを最大化するレバレッジの存在を確認することで、学習をしているか確かめる必要がある。

実験で使用した実際の外国為替市場データを図 1 に示す。2015 年 06 月 21 日から 2016 年 11 月 20 日までの毎週の取引データを示す。

対象期間における最初の 50 週分のリターンに対し、各リターンが 1/50 の確率で発生するものとしてバンディット問題を作成した。本論文では、ドル、円、ユーロの 3 つの通貨に対し、ドル円、円ドル、円ユーロ、ユーロ円、ユーロドル、ドルユーロの 6 組の為替取引を対象とした。ここで、ドル円はドル買い円売り取引に対応し、週の初めに円を売ってドルを買い、週の終わりに買ったドルを売って円を買い戻すことに相当する。6 組の通貨ペアのそれぞれに対応した 6 本腕バンディット問題とし、レバレッジを用いた複利型強化学習を用いてどの通貨ペアにどの程度投資するのが良いかを学習させる。

レバレッジ（投資比率）の初期値は 0.1 とし、レバレッジの学習率を $\eta = 0.01$ とした。強化学習における学習中の行動選択には $\epsilon = 0.2$ の ϵ-グリーディー選択を用い、学習率は $\alpha = 0.01$ とした。評価時はグリーディー選択を用いた。100 万ステップの学習を乱数のシードを変えて 100 回行い、幾何平均利益率と各行動のレバレッジの平均を求めた。

結果を図 2 と図 3 に示す。 学習が進むにしたがって幾何平均利益率が徐々に上昇し、最終的には約 0.004 となった。これは、長期的に見ると 1 回の取引ごとに 0.4%資産が増加することを意味している。

また、レバレッジを見ると、jpyusd, jpyeur, usdeur のレバレッジが大きくなっている。通貨ペアは対になっているため、一方で利益ができる場合はもう一方では利益は出ない。このため、それぞれの対の一方のレバレッジだけが大きくなるはずであり、実験結果と一致する。

4 まとめ

本論文では、実際の外国為替市場のデータに基づいて通貨ペア選択問題を作成し、レバレッジを用いた複利型強化学習を適用した。実験結果から、レバレッジを用いた複利型強化学習は実際の外国為替市場のデータに対しても最も利益が高い通貨ペアを学習し、レバレッジを学習できることが確認できた。今後は、レバッレジの収束性について検討したい。

参考文献

[1] 松井 藤五郎. 複利型強化学習—強化学習のファイナンスへの応用—. 計測と制御, 52(11):1022–1027, 2013.

[2] 塚本 智大, 松井 藤五郎. “レバレッジを用いた複利型強化学習.” 第 78 回情報処理学会全国大会, 3P-05, 2016.

Copyright ©2017 Information Processing Society of Japan.
All Rights Reserved.