

MDP Approach in Econophysics

東京工業大学大学院理工学研究科 家田雅志¹

本稿では、市場価格変動モデルの評価に MDP(Markov Decision Process) を用いる方法を提案する。まずはじめに、MDP を紹介し、次に具体的な評価関数を導入する。最後に、簡単な価格変動モデルに対して MDP による評価を行い、モデル改良について考察する。

1 緒言

近年、詳細な市場データの入手が可能になり、これに基づいて様々な市場価格変動モデルが提案されている。これらのモデルは価格変動の確率分布がファットテールを持つことなど、市場の様々な特徴を再現することも報告されており、今後さらなる発展が期待されている。

しかし、これらのモデルに対して、実際に投資を行う観点から評価するアプローチはあまり行われていないように見受けられる。そこで、本研究では与えられた市場価格変動モデルに基づいて最適投資戦略を求め、そのパフォーマンスからモデルを改良する手法の確立を目的としている。

2 MDP

MDP(Markov Decision Process) とはエージェントの意思決定を含んだ確率過程である。MDP は、状態空間 $\mathcal{S}$ (系が取りうる状態からなる集合)・行動空間 $\mathcal{A}$ (エージェントの行動の選択肢)・状態遷移確率・利得 η (通常は $\mathcal{S}$ と $\mathcal{A}$ の元を引数に持つ関数)・政策 $\pi = \{\pi_t\}$ (エージェントの意思決定) の 5 つの要素で構成される。そして、状態遷移確率が次の条件を満たす場合を MDP と呼ぶ。

$$Prob(s_{t+1} | s_0, a_0, \dots, s_{t-1}, a_{t-1}, s_t, a_t) = Prob(s_{t+1} | s_t, a_t) \quad (1)$$

ただし、 $s_t \in \mathcal{S}$ は時刻 t での系の状態、 a_t は時刻 t でのエージェントの行動である。 a_t は $a_t = \pi_t(\cdot)$ のように政策から決定される。

MDP が使用される問題の代表例は、(1) を満たす状態遷移確率が与えられた場合の最適化問題である。また、強化学習理論の数学的な基礎² を与えるものとしての側面も持っている。

¹E-mail: mieda@mikan.ap.titech.ac.jp

²学習理論を考える際に問題となるのは、学習する対象に対して正解が確実に存在するのか、その正解に対して学習結果が確実に結果が収束するのか、の二点である。MDP は前者の証明を与える。

3 評価関数

本研究では、最適投資戦略の算出方法として MDP による最適化のフレームワークを採用した。このため、本稿では以降、 s_t と a_t をそれぞれ時刻 t での投資対象の市場価格と投資行動、 π を投資戦略と考える。

さて、一般的な MDP を用いた最適化問題において、最適化度合いを判定する評価基準として用いられているものは、予測割引利得と呼ばれる関数である。(詳細は [1] などを参照されたい。)

しかし、予測割引利得はいくつかの理由から投資戦略の計算には向いていない。最も主要な理由は、予測割引利得を用いる場合に、状態遷移確率が時間変動しないことが条件として含まれているためである。つまり、 p_t を具体的な確率(密度)関数としたときに、その関数が時間変化しないこと

$$Prob(s_{t+1} | s_t, \pi_t) \neq p_t(s_{t+1} | s_t, \pi_t)$$

が条件として課せられているのである。しかし、この条件を満たす市場価格変動モデルで現実的な市場を表現することは非常に難しい。

そのため、本研究では投資手法を売買差益を得るもの限定し、評価関数を次のように定義する。

$$V_T^\pi(s) = E \left[\sum_{t=0}^T \pi_t (s_T - s_t) \middle| s_0 = s \right] = \sum_{t=0}^T \sum_{s_T, s_t} \pi_t (s_T - s_t) Prob(s_T, s_t | s_0 = s) \quad (2)$$

ここで、 s_t は 1 ユニットあたりの投資対象の市場価格、 a_t は $+1 \cdot 0 \cdot -1$ のいずれかの値をとり、それぞれ 1 ユニット購入・何もしない・1 ユニット売却を意味する。また、投資行動の選択肢 $\mathcal{A}_t$ をそれまでの投資履歴 $h_t \equiv a_0 + a_1 + \dots + a_{t-1}$ により次のように制限する。

$$\mathcal{A}_t(-1) = \{0, +1\}, \mathcal{A}_t(0) = \{-1, 0, +1\}, \mathcal{A}_t(+1) = \{-1, 0\}$$

これにより、同一のポジションを 2 つ以上とることを禁止し、別のポジションへ移るためには一度保持しているポジションを解消することが義務付けられる。

この評価関数 (2) に対する最適投資戦略の存在は、状態空間(市場価格)が離散有限であること、それにより政策も離散となることから自明である。

4 価格変動モデルとその評価

本稿で評価を行う価格変動モデルを次のように設定する。

$$s_{t+1} = s_t + z_t \quad (3)$$

ただし、 z_t は実際の市場データから生成した価格変動確率分布より発生させる乱数である。このモデルは価格変動をランダムウォークをベースに、実際の市場データを用いて価格変動確率分布がファットテールを持つように拡張した、ごく単純なモデルである。

続いて、評価のためのシミュレーション条件を設定する。投資対象は外国為替市場(ドル円)とする。価格変動確率分布を生成するための市場データは、2005年の10分足データを使用する³。このデータから生成される価格変動分布は図1左側のグラフのようになる。

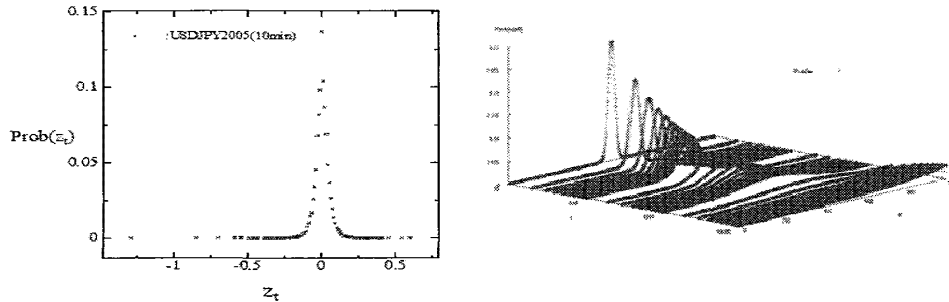


図 1: 価格変動分布

図1右側のグラフは、このモデルを用いた時に、各ステップごとに初期値からどの程度の価格変動がどの程度の確率で生じるかを示したものである。ここで、1000ステップ以上になると明らかにグラフの形が歪んでいる⁴。そのため、投資期間 T は実時間との関連も考慮し、700ステップ(実時間1週間相当)とする。そして、2005年のデータを調査した結果、700ステップの間に生じるレートの変動幅は ± 5 円よりも十分に小さいため、想定レート変動レンジ(想定レンジ)は ± 5 円とし、自然境界条件とみなす。

以上の条件の下、(2)を用いて最適投資戦略 π を計算し、2006年に投資を行った結果を図2に示す。

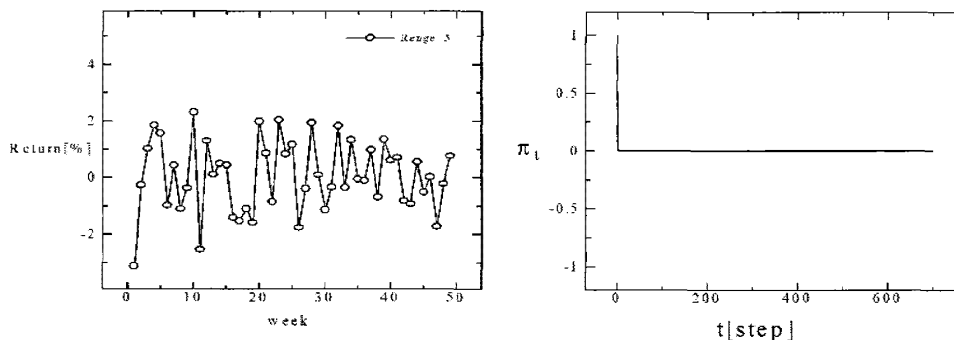


図 2: 最適投資戦略による投資結果

図2左側は利益を初期投資額比のパーセンテージで表示したもの、右側は算出された最適投資戦略を示している。この図から最適投資戦略は最初にロングポジションをとるのみであると算出されたことがわかる。これは先にも表れた2005年の上昇基調を反映しただけの結果である。

³従って、シミュレーション上の1ステップは現実時間の10分に相当する。

⁴これは、2005年のドル円レートが上昇基調にあったことに起因している。2006年も通年で同様の傾向があると判断することは非常に危険であるから、投資期間 T は1000ステップよりも小さくする必要がある。

この結果から、このモデルには調整局面(ある程度のレート上昇の後には下降圧力がかかる)に相当する要素が不足していると考えることができる。この調整局面をモデルに組み込むために、簡易的な方法ではあるが境界条件を変更する。境界条件を反射壁に変更し、想定レンジを狭めて、強力な抵抗線的作用を組み込む。境界条件を変更した結果を図3に示す。

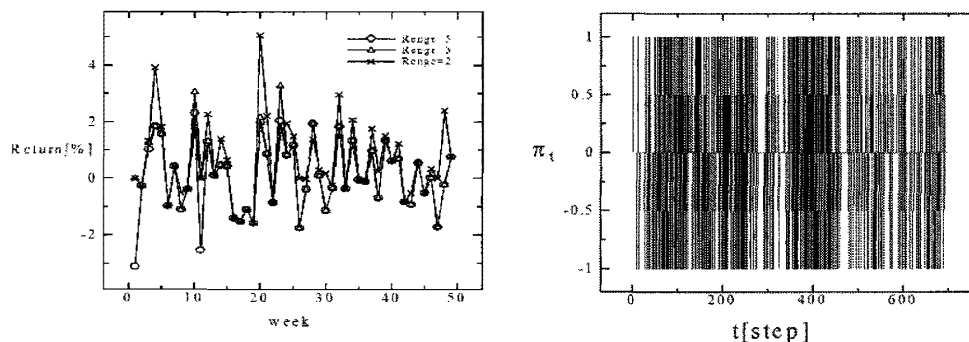


図 3: 最適投資戦略による投資結果 (モデル改良後)

図3左側は投資のパフォーマンスを、右側は想定レンジを ± 2 円とした場合に算出された最適投資戦略を示している。図からわかるように、トレード頻度が十分に上がり、パフォーマンスも改善されていることがわかる。

5 おわりに

本稿では、与えられた価格変動モデルに対して MDP に基づいた最適投資戦略を求め、そのパフォーマンスからモデルを改良する一連の流れを示した。これにより、MDP によるモデル評価が有効に作用する可能性を示すことができた。

今回は非常に単純なモデルに対して評価を行ったが、すでに提唱されている現実的なモデルに対しても同様の評価を行うことで、更なる発展が期待できると考えている。

また、今回用いたモデル (3) を

$$s_{t+1} = f(s_t) + z_t \quad (4)$$

のように変更する。そして f の関数形として回復力を与えるポテンシャルを導入するなどして、調整局面に対応したモデルへと拡張する。これにより、Markov 性を保ったままの市場モデルがどれだけのパフォーマンスを発揮できるかを検討することも、価値があるのではないかと考えている。

参考文献

- [1] M. L. Puterman, Markov Decision Processes, John Wiley & Sons, (1994).
- [2] R.N. Mantegna and H.E. Stanley, An Introduction to Econophysics, Cambridge University Press, (2000).