F-019

高頻度価格時系列の特徴と予測可能性

田中美栄子，元山智弘(*)，小牧信也(**)
鳥取大学工学部知能情報工学科，鳥取大学工学研究科知能情報工学専攻(*)
東芝プロセスソフトウエア株式会社(**)
mieko@ike.tottori-u.ac.jp

Abstract: 価格変動は通常ランダムウオークと考えられているが，一方ではこれを支配するなんらかの法則があると考えられる証拠もいくつか存在する．本稿では高頻度価格変動が普遍的に持つと思われる性質のうち，１万点程度以上の範囲にわたる平均値としてはティックごとに反発傾向に向かうにもかかわらず，数百点程度以下のスケールでは統計分布においてもまた複数為替の同時裁定機会がバースト的に出現することからも牽引傾向が存在するという，一見矛盾する二つの現象を示す高頻度時系列の予測可能性を議論する．

１．価格変動の科学

価格変動の従う力学がどのようなものであるかはまだわかっていない．競争市場で決まる価格変動にはランダム性が高いので，数学的に一番簡単なモデルとしてランダムウオークが使われるが，必要な自由度や，一つの法則が及ぶ範囲，それにそもそも何らかの一般的な規則が存在するのかどうかについても，明らかでない．

ランダム性が非常に高い場合，株式や為替への投資が万人に対して同等の機会を与えるゲームとなり，投資市場が成立する．一方，何らかの規則に従う場合は，その規則に従って行動する人が勝ち組となり，そうでない人は負けることになる．一般には長期間にわたって有効な規則はないので，長期投資には素人でも安心して参加できるが短期投資において素人が玄人に勝つことは困難であるとされている．

一方，価格変動がフラクタル性をもつことは多くの例から知られているが，このことに基づいて暴落の予兆を読み取る可能性について多くの研究が行われてきた．暴落や暴騰などの急激な変化は，人間の生活を脅かす危険性が高く，安全な生活を保障するためにはできれば回避することが望ましい．経済問題は環境問題と同じく，万人の安全に関わる重要な問題でありながら，グローバルな視野を必要とし，すぐには解にたどり着けない非効率性のため，これまで後回しにされてきたが，近年の計算機の高速化と大容量化によって急激に手に届く問題としてクローズアップされるようになった．すなわち科学としての定量的な解析を待つ状況ができあがったのである[1].

我々は高頻度価格変動データの統計的性質を計算機解析によって調べてきた．最初に確かめたかったことは，価格変動の統計分布がある種の規則に従うという通説であった．数秒単位の価格変化の分布は正規分布ではなくて，中心が狭く裾野の広い，ファットテールの分布をし，ある時間間隔の範囲では Levy の安定分布で近似できることもある[2]，という

結果は様々の実データで検証できることがわかった.
また，このように短い時間間隔での外国為替価格を同時比較すると，複数の為替を数十秒以内に売り買いすれば「濡れ手に粟」の差益を得る可能性，すなわち裁定機会があることもわかった．このような状態はいったん生じるとしばらく続く傾向が見られ，バースト誤りのようなものと仮定した扱いで理解できる．つまりある種のマルコフ型モデルで記述できるような性質が見られ，一旦その状態に遷移するとしばらくその状態に留まり易くなる，という遷移確率を持つであろう事が推測されるのである．通常は裁定機会のない状態にいるが，裁定機会のある状態に移行すると高い確率でその状態に留まるという傾向が見られるのである[3].

さらに一つの為替変動の時系列をある大きさ（L とする）の束に区切ったものから条件付確率を計算すると，明らかな規則性が見えることがわかった．自由度が多すぎると見えなくなるので，上下運動だけに絞って条件付確率の時系列を自動計算すると，データの及ぶ限り一定値からほとんど外れないことがわかった．つまり，上昇と下降を 1 と 0 で表せば

$$P(1|0)=P(0|1)=0.7 \tag{1}$$

となり，一つ前の動きとは反対方向に動く．条件なしの上昇と下降の動きはほぼ等確率

$$P(1)=P(0)=0.5 \tag{2}$$

であるから，明らかに一つ前の動きと強い相関があるのみならず，この規則が時期に依らず一定であり，データの種類にもほぼ無関係であることから，一般性を持った何らかの規則がその裏に存在することが予想されるのである．また，どこまで記憶が及ぶのかを見るために条件を長くしてゆくと大まかに言って,

$$P(1|00)=0.8 =P(1|0)+0.1 \tag{3}$$

$$P(1|01)=0.6=P(1|0)-0.1 \tag{4}$$

等と枝分かれしてゆくが，深さ 3 ステップくらいで打ち止めとなり飽和してしまう.

$$P(1|0000)=P(1|0001) \tag{5}$$

というように，ずっと前まで考慮しても何も新しい知見は得られない．これは為替のみでなく，取引量の多い人気株でも同じような振る舞いをする．これは予測が可能であることを意味する．100%確実な予測はできないが，確率の多い方に賭ければ 70%ほどの確率で予測が当たることになるわけで，回数を多く取れば誤差はいくらでも小さくできる[4]．このようなゲームで賭を行えば 7 割勝って 3 割負けであるから，相当な勝ちをおさめることが出来るわけである．また，二つ続けて同じ方向へ動いた後は，8 割を超える確率で逆方向に移動する．しかしこれは条件を満たす状況が生起する確率が 1 より小さいため，的中率を上げる為の方便として直接利用することはできない.

「次の取引」がいつ行われるかを知らなければこの知見を実際に使うことは難しいが，制度として公平な投資市場を作りたければ，取引間隔を十分広く取らざるを得ないような制度に移行する，というような対処もできるわけである．

2．環境の微細構造を拾う試み

前章で議論した条件付確率の定値性はＬが十分大きい場合を前提にしている．このＬが1,000以上ならば条件付確率は漸近値の２倍以下の範囲に収まり，5,000以上で漸近値から誤差20%以内の範囲に入り，10,000以上ならば，漸近値から10%以下の範囲に入る．このような理由から，前章での条件付確率をデータから求める際には，データを区切る大きさＬを２万ティックまたはそれ以上として求めた比率として算出した．その結果が式(1)－(5)の規則であった．

しかし，先に指摘した統計分布がファットテールになることは過去の動きに対してポジティブフィードバック，すなわち同じ方向に動く傾向がある事に関連しており[5]，反発性の高い規則とは相容れない．この謎を解くには、まずＬより細かい構造について知る必要がある．微細構造がある種のパターンを持つけれども単純な構造を持たない場合には，しばしば進化計算による自動学習が用いられる[6]．

そこでティックごとの価格変化の方向を｛下降，不動，上昇｝＝{0,1,2}の３方向とし，これを環境と見立てることにより，条件付確率の最適値を自動学習する遺伝アルゴリズムで正答率を評価関数とした進化計算を試みた．用いたデータはCQGによる米ドル対日本円のbid値で1995-2001にわたり１千万点弱の長さをもつ[7]．遺伝子数を100とし，長さＮ点の学習期間を終了すると世代交代を行って100個の遺伝子を良い方から数個残し，残りを選ばれた遺伝子の１点をランダムに変化させて作った突然変異種で置き換えて，学習結果が収束，すなわち最適遺伝子が変化しなくなるまで繰返した．このアルゴリズムで，条件付確率の条件部分の長さｎを１から７まで変化させ，また学習期間Ｎをいろいろに変化させて正答率がどこまで上がるかを調べた．この方法では正答率は68%程度に留まり，ｎは１－4くらいまでが一番良い結果を与えるが，これ以上大きくしても正答率が上昇することはなかった[8]．

以上の方法では，価格変化の方向のみに限定しているため，例えば上昇－下降－上昇という変化の後で量的に価格が上昇している場合も下落している場合も同じ条件101となってしまう．この点を改良するため，過去ｎティックの平均値に比べて現在値が上昇しているか下落しているかという条件に置き換えて，上と同じ条件付確率の遺伝子を進化させるアルゴリズムを適用して見ると上記の方法より2%近く正答率が上がり，約70%が得られた[9]．

3．考察

微細構造に決まったパターンが見えるならば，遺伝アルゴリズムによる自動学習で得ら

れた遺伝子は正答率 70%を越えるであろうことが予測された．結果としてそのあたりを上回らなかったことは，学習アルゴリズムに改良の余地があるか，あるいは決まったパターンが見えにくいかのどちらかである．

二つの方法を比較すると，前者の，変化の符号のみを問題にした場合よりも，後者の，変化値まで考慮した方が僅かに良い正答率を得たということは，変化の符号と過去の変化値の間に相関が存在することを意味する．いずれも最適遺伝子の正答率を評価値として採用したが，これを確率分布に置き換えることで正答率の向上が得られるかどうかを調べる余地はあると考えられる．

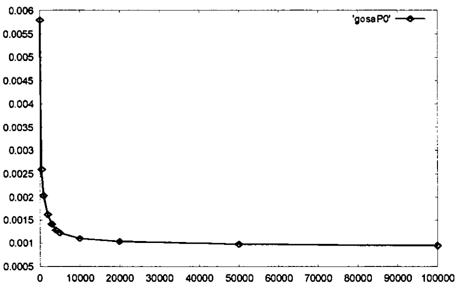

Fig.1 条件付確率の分散とデータ長 L との関係．L>10,000 で漸近値から 10%以内に安定する．

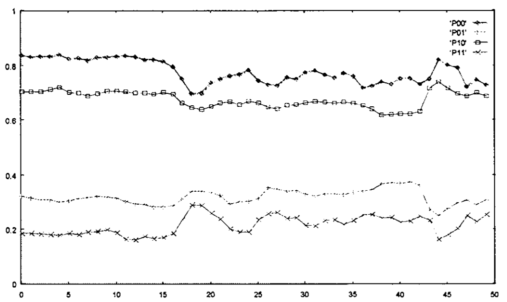

Fig.2 条件付確率 P(1|00), P(1|10), P(1|01),P(1|11)（上から下へ）は 1995-2001 年にわたり定値をとる．

参考文献

1. e.g. 入門、経済物理学、暴落はなぜ起こるのか？:森谷博之監訳:PHP 出版 (2004); 原著:Sornette, D., Why Stock Market Crash: Critical Events in Complex Financial Systems:Princeton University Press (2003)

2. Weron, R.: Levy Stable Distribution Revisited: Tail Index > 2 Does Not Exclude Levy Stable Regime: International Journal of Modern Physics C 12 (2) (2001) 209-223

3. Tanaka-Yamawaki, M., and Komaki, S.: A Numerical Analysis on the Triangular Arbitrage Chances:nolta(2002)

4. Tanaka-Yamawaki, M.: Stability of Markovian Structure Observed in High Frequency Foreign Exchange Data: Ann. Inst. Statist. Math. (AISM) 55 (2003) 437-446

5. Tanaka-Yamawaki, M.: Two-phase Oscillatory Patterns in an Positive Feedback Agent Model: Physica A 324 (2003) 380-387

6. Holland, J.H.: Adaptation in Natural and Artificial Systems: An Introductory Analysis with Applications to Biology, Control, and Artificial Intelligence, MIT Press (1992)

7. We thank Mr. Moriya and the members of SOKENDAI econophysics meeting for sharing the data.

8. Tanaka-Yamawaki, M. and Motoyama, T.: Evolution of Investment Strategy in an Environment Offered by the Real Tick Data: Proceedings of the 9-th International Symposium on Artificial Life and Robotics: Beppu, Oita, January 28-30 (2004) 313-316

9. Tanaka-Yamawaki, M. and Motoyama, T.: Predicting the Tick-wise Price Fluctuations by Means of Evolutional Computation: to appear in Proceedings of CEC2004.