

# 不等間隔時系列のフラクタル解析

熊谷善彰\*

\* 慶應義塾大学産業研究所

Fractal Analysis for Non-equidistant Time Series

Yoshiaki Kumagai \*

\* Keio Economic Observatory

## Abstract

We propose a new method to describe fractal structure of data, which are non-equidistant in physical time, and apply this method to fractional Brownian motions. Using extended extreme values, we define functions independent of time scale. Moreover, a kind of fractal dimension is measured. In high frequency financial data, observations can occur at varying time intervals. Using these functions, we can analyze non-equidistant data without interpolation or evenly resampling. Moreover, the problem of choosing the appropriate time scale is avoided. Lastly, these functions are related to a viewpoint of investor with constant transaction costs.

## 1 はじめに

近年、金融市場におけるデータベースの整備により、個々の取引のデータ（ティックデータ）などの High Frequency Data を対象とする研究が盛んに行われている。個々の取引のように異なる時間間隔で発生するデータを等間隔の離散時間系列に変換する場合、幾つかの方法が考えられる。まず、物理的時間で測ったときに等間隔になるように、一定間隔（5分毎、30分毎など）でサンプリングし直す方法がある。この方法では、スキップされるデータの情報は失われる。つぎに、データの発生順序を時間とみなす方法がある。データの発生順に番号を振り、この通し番号を時間とみなす。これは、出来高の累積を価格過程の先導過程とする方法 [1] と同じ考え方に基づいている。また、一般に、金融時系列において用いる時間尺度の選択については多くの議論がある。主な候補としては、物理的時間・市場の開いている時間・取引回数あるいは出来高の累積の3つが挙げられる。これらにはいずれも長所と短所があり、決定的なものはない [2]。連続時間システムを不等間隔サンプリングする場合とは異なり、金融時系列における不等間隔データは、取引そのものが時間的に不等間隔なので、サンプリング間隔をより細かくすることで解決することもできない。

そこで本稿では、時間尺度の選択に依存しない方法で時系列データのフラクタル解析を行う。通常の時系列グラフのフラクタル解析においては、時間軸が粗視化される。ここでは変動を粗視化することによって、時系列グラフのフラクタル性を測定する。極値の定

義を拡張して、変動の粗視化の程度に応じた極値を定義する。拡張された極値の間の変動は、粗視化の程度の関数となっている。この極値間の変動から、一種のフラクタル次元である折畳次元を定義する。これらの計算結果は、時間尺度として何を採用するかに関係なく、データの時間的順序のみに依存している。言い換えると、価格の変動そのものを時間尺度としている。本稿では、フラクタルブラウン運動についての計算例から、この折畳次元の有用性を示す。円相場のティックデータについての計算例も併せて紹介する。

## 2 変動の粗視化

関数  $g(s)$  に対して、粗視化の程度に対応した極値を以下のように定義する [3]。

$g(s)$  を離散時間  $\{s_0 = 0, s_1, s_2, \dots, s_n = S\}$  上の実数値関数とする。このとき、 $\{s_0 = 0, s_1, s_2, \dots, s_n = S\}$  は等間隔である必要はない。 $g(t - d_1) < g(t) - C, g(t + d_2) < g(t) - C$  なる  $d_1, d_2 > 0$  が存在して、区間  $[t - d_1, t + d_2]$  内では、 $g$  が  $t$  において最大をとるとき  $g(t)$  を「 $C$  で粗視化した極大値」と呼ぶ。極小についても同様に定義する。 $g(t - d_1) > g(t) + C, g(t + d_2) > g(t) + C$  なる  $d_1, d_2 > 0$  が存在して、区間  $[t - d_1, t + d_2]$  内では、 $g$  が  $t$  において最小をとるとき  $g(t)$  を「 $C$  で粗視化した極小値」と呼ぶ。

例えば、関数  $g(s)$  を Fig. 1 において (o) で示すような、離散時間の不等間隔データとする。破線のようにデータの点と点を結んで補間した場合、 $t_{01}, t_1$  などの点では極大、 $t_{02}, t_2$  などの点では極小をとる。しかし、 $C$  で粗視化した場合の関数の極値は、極大  $g(t_1)$  と、極小  $g(t_2)$  のみとなる。図で考える場合は、横軸に平行な幅  $C$  のスリットを使って時系列グラフ  $g(s)$  を観察するとよい。横軸に平行を保ちながらスリットを移動し、時系列グラフを辿る。グラフのスリット内に見えている部分ができる限り連結になるようにスリットを移動する。スリットの中に入っている部分が連結でなくなるところに、 $C$  で粗視化したときの極値が存在する。 $C = 0$  とした場合が、 $g(s)$  の各点を Fig. 1 の破線のように結んだ連続関数における通常の極値に相当する。

時系列  $g(s)$  について、期間  $[0, S]$  を固定して、 $C$  で粗視化した極値を考える。極値をとる時点に時間的順序に従って  $t_1^C, t_2^C, \dots$  のように添字を付けると、極大と極小が交互に現れる。時系列  $g(s)$  を固定し、期間  $[0, S]$  内における極値の個数を  $C$  の関数として  $m(C)$  と表す。 $g$  の極値から極値までの変化の絶対値を累積した総変動  $R(C)$  は、

$$(2.1) \quad R(C) \equiv \sum_{i=1}^{m(C)-1} |g(t_{i+1}^C) - g(t_i^C)|$$

と表すことができる。 $m$  は階段状の単調減少関数で右方連続であるが、左方連続ではない。総変動  $R$  も  $m$  と同じく、右方連続だが、左方連続ではない単調減少関数である。範囲 (レンジ) 以上、つまり、 $C_L \geq \max g(t) - \min g(t)$  のときは、 $m(C_L) = 0, R(C_L) = 0$  である。また、 $m(0)$  は、Fig. 1 の破線のように  $g(s)$  の各点を結んだ連続関数の、期間  $[0, S]$

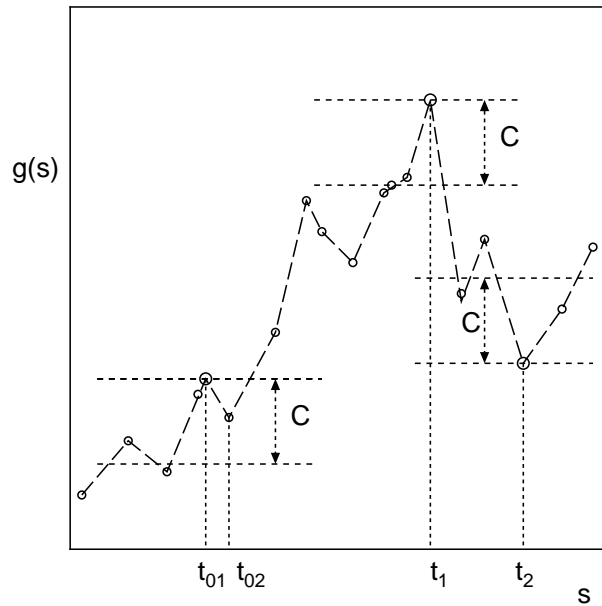


Fig.1: Extreme values determined by a scale  $C$ . The circles ( $\circ$ ) represent a function  $g(s)$ . The function  $g$  attains its local maximum determined by a scale  $C$  at  $t_1$ , but does not attain its maximum at  $t_{01}$ . Besides,  $g$  attains its local minimum with  $C$  at  $t_2$ , not at  $t_{02}$ . Using a slit: a pair of parallels whose distance is  $C$  (denoted by *dashed lines*), we can detect these extreme values.

内における通常の極値の個数である。そして、 $R(0)$  は個々の変動の絶対値の総和  $R(0) = \sum_{k=1}^n |g(s_k) - g(s_{k-1})|$  である。時間的に等間隔な離散時間系列  $g(s), s = 1, 2, \dots, S$  に適用すると、 $R(0) = \sum_{k=1}^S |g(k) - g(k-1)|$  であり、これをデータ数で割る  $R(0)/S$  と絶対偏差が得られる。つまり、 $R(C)$  は粗視化の尺度  $C$  に対応した絶対偏差ともいえる。

さらに、 $R$  を用いて、変動の粗視化による一種のフラクタル次元  $D_f$  を求めることができる。

$$(2.2) \quad R(C)/C \sim C^{-D_f}$$

の関係が成立しているとき、 $D_f$  を折畳次元と呼ぶ [5]。この定義について、フラクタル次元として一般的な容量次元との類推によって、簡単に説明する。容量次元の定義は以下のようなものである [9]。同じ大きさの立方体 (球でも良い) を用いて、対象の図形を無駄のないように最少の個数で覆う。立方体の一辺の長さ  $\varepsilon$  の関数として、図形を覆うために必要な最少の個数  $N(\varepsilon)$  を求める。図形にフラクタル構造があるとき、必要な立方体の個数  $N(\varepsilon)$  と立方体の大きさ  $\varepsilon$  の対数比は一定になる。そこで、

$$(2.3) \quad N(\varepsilon) \sim \varepsilon^{-D_c}$$

の関係が成立しているとき、 $D_c$  を容量次元と呼ぶ（あるいはボックスカウンティング次元、箱個数次元とも呼ばれる.）。折畳次元の定義と比較すると、 $C$  が長方形の一辺の長さ、 $R(C)/C$  が長方形の個数に対応している。折畳次元の場合は容量次元と異なり、時系列グラフを覆う長方形の横の辺、つまり、時間軸方向の長さは一定ではない。

次に、この2つの関数から

$$(2.4) \quad \Pi(C) \equiv R(C) - Cm(C)$$

を定義する。この関数の価格時系列における意味は、次のようである。値幅  $C$  を取引費用と考える。ある取引費用  $C$  で価格時系列を粗視化し、その極値で取引を行う。そのとき取引費用を差し引いた利益が  $\Pi(C)$  に相当する。極大と極小の差が取引費用  $C$  よりも小さい変動で売買を行うと損失が生じる。しかし、ここでは取引費用で粗視化した極値  $t_1^C, t_2^C, t_3^C, \dots$  のみで取引している。つまり、 $\Pi(C)$  は事後的に最適な取引を行った場合の一単位あたりの利益である [4]。  $\Pi$  は  $m$  や  $R$  の不連続点をなくした非負の単調減少連続関数である [5]。  $\Pi$  は  $m, R$  の不連続点では微分可能ではない。

ここで、以上の関数  $m, R, \Pi$  はデータの発生順序のみに依存していて、時間尺度の選択には独立である。

容易に分かるように、折畳次元は単調増加あるいは単調減少関数の場合は、1 となる。時系列について  $C$  に対する  $R(C)/C$  を両対数グラフにプロットすると、 $C$  の小さい部分では傾きが緩やかで1に近くなる。一方、 $C$  が大きい部分では傾きは急になる。その中間の部分に存在する直線状の部分の傾きから折畳次元を測定する、

### 3 計算例

まず、人工的に発生させたデータに対する計算例を示す。フラクタルブラウン運動は、ブラウン運動  $B(t)$  を一般化したもので、パラメータ  $0 < H < 1$  に対して、

$$(3.1) \quad B_H(x) \equiv \frac{1}{\Gamma(H + \frac{1}{2})} \int_{-\infty}^x (x - x')^{H-1/2} dB(x')$$

で定義される [7] ( $\Gamma(\cdot)$  はガンマ関数である.)。フラクタルブラウン運動の時系列グラフは自己アフィンという、広い意味での自己相似性をもっている [9]。つまり、任意の  $h > 0$  に対して、

$$(3.2) \quad B_H(x + hX) - B_H(x) \stackrel{d}{=} h^H \{B_H(x + X) - B_H(x)\}$$

の関係が成立する。フラクタルブラウン運動の時系列を発生させるアルゴリズムは幾つか知られている。ここでは、スペクトル密度を近似するフーリエ・フィルタリング法 [8] を用

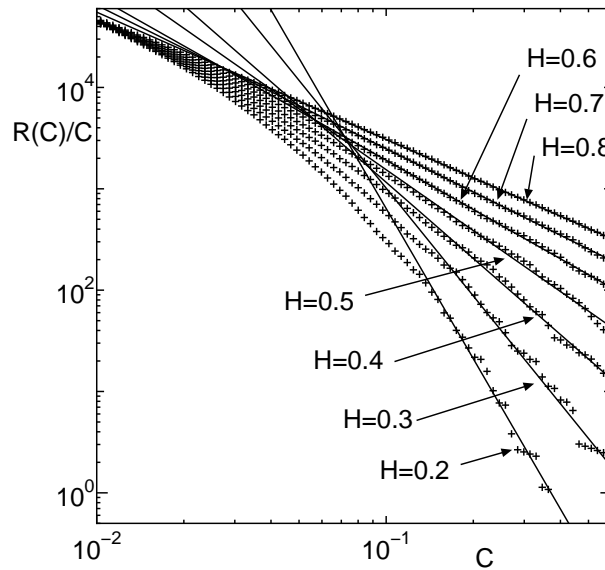


Fig.2: Double logarithmic plot of  $R(C)/C$  with  $C$ . The pluses (+) represent the result from fractional Brownian motions (fBm) with parameter  $H = 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8$

いる。他には、中点変位法、逐次ランダム加算法 [10] などの解像度を上げていくことにより自己アフィンな時系列グラフを構成する方法がある。

時系列  $X(t)$  のフーリエ変換を  $F(f, T)$  としよう。スペクトル密度

$$(3.3) \quad S(f) \equiv \lim_{T \rightarrow \infty} \frac{1}{T} |F(f, T)|^2$$

$$\text{where} \quad F(f, T) \equiv \int_0^T X(t) e^{-2\pi i f t} dt$$

が  $1/f^2$  に比例する場合は、 $H = 0.5$  の通常のブラウン運動である。一般に、スペクトル密度  $S(f)$  が  $1/f^\beta$  に比例する過程は、 $H = (\beta - 1)/2$  のフラクタルブラウン運動に対応する。そこで、パラメータ  $H$  に対して、 $\beta = 2H + 1$  を求め、スペクトル密度の形状が  $1/f^\beta$  となる時系列を、逆フーリエ変換によって作成する。ここでは、パラメータ  $H$  を、0.2 から 0.8 まで 0.1 ずつ変化させて、それぞれ 100000 個のデータを作成した。Figure 2 には  $C$  に対する  $R(C)/C$  の両対数グラフを示した。Table 1 は、グラフの右側部分に直線を当てはめたときの傾きから求めた折畳次元である。ここで求められた折畳次元  $D_f$  は、 $H$  の逆数である潜在次元  $D_l = 1/H$  [6] にほぼ等しくなる。

次に、価格時系列データについて、計算例を示す。

Figure 3 には、2000 年 5 月 8 日月曜日から 12 日金曜日の円ドル相場のティックデータについて、一日毎に計算した結果を示す。Figure 4 には、2000 年 8 月 21 日月曜日から 25

Table 1: Folding dimension of fractional Brownian motion with parameter  $H$

$H$	$D_f$	$D_l$
0.2	4.9253	5.0000
0.3	3.5336	3.3333
0.4	2.4952	2.5000
0.5	1.9962	2.0000
0.6	1.6148	1.6667
0.7	1.4163	1.4286
0.8	1.2562	1.2500

日金曜日の、同じく円ドル相場のティックデータについて、一日毎に計算した結果を示す。両図とも比較のため、Fig. 2のフラクタルブラウン運動の結果の一部を併せて表示した。8月24日は日本銀行のゼロ金利解除によりボラティリティの増大した日である。

## 4 結び

本稿では、時間尺度に独立な関数を用いて時系列データを解析する手法を紹介し、フラクタルブラウン運動・円相場ティックデータについての計算例を示した。金融時系列では、市場において価格過程を進展させる時間として複数の候補がある。また、近年盛んに研究されるようになったティックデータなどのHigh Frequency Dataは、データ自体の時間間隔が一定ではない。本稿の方法により、データの情報を失うあるいは歪めることなく、不等間隔データのフラクタル性を分析することが可能である。今後、この方法を用いて様々な金融時系列データの分析を試みたい。

## 謝辞

本稿の作成にあたっては、岩田暁一氏（東京工科大）・森平爽一郎氏（慶應義塾大）・辻幸民氏（慶應義塾大）から、多くの有益な御助言を頂いた。本稿で用いた円ドルレートのティックデータは慶應義塾大学サイバー・トレーディング/リスクラボより提供を受けた。本稿の成果の一部は科研費「非線形動学の方法による経済変動の研究」にかかわる研究会（富山大）において報告されたが、その際、稲葉敏夫氏（早稲田大）・田中辰雄氏（慶應義塾大）をはじめ、出席者から示唆に富むコメントを頂いた。ここに厚く御礼申し上げたい。

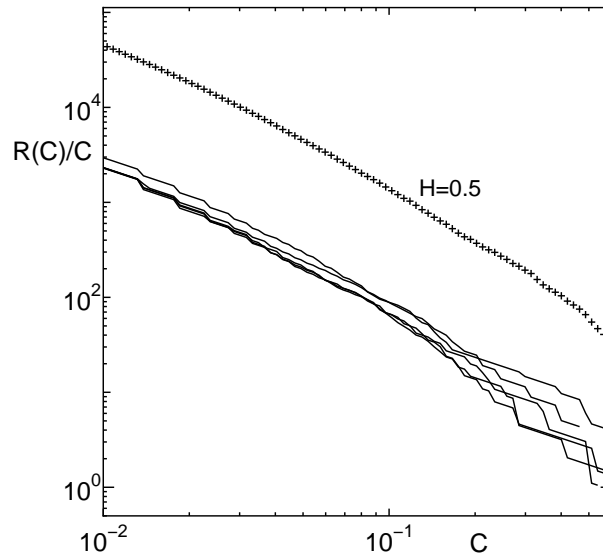


Fig.3: Double logarithmic plot of  $R(C)/C$  with  $C$  for USD/JPY exchange rate. The lines represent the result from the data of each day from May 8, 2000 to May 12, 2000. The pluses (+) represent the result from a fBm with parameter  $H = 0.5$

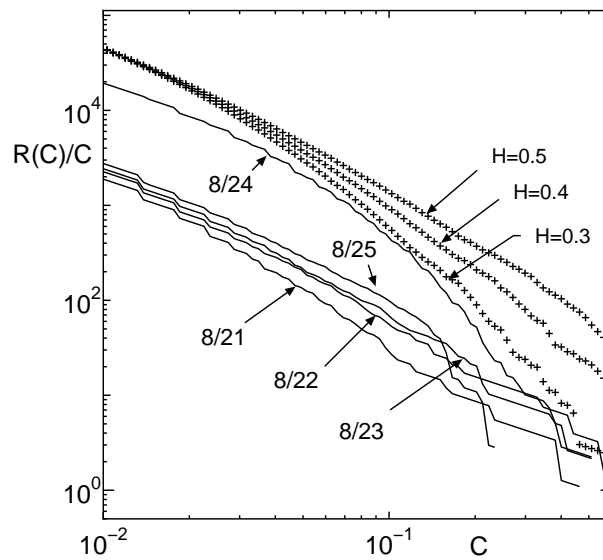


Fig.4: Double logarithmic plot of  $R(C)/C$  with  $C$  for USD/JPY exchange rate. The lines represent the result from the data of each day from Aug. 21, 2000 to Aug. 25, 2000. The pluses (+) represent the result from fBm with parameter  $H = 0.3, 0.4, 0.5$

## References

- [1] Clark, P. K., Subordinated stochastic process model with finite variance for speculative prices, *Econometrica* 41: (1973) 135-155.
- [2] Goodhart, C. A. E. and O'Hara, M., High frequency data in financial markets: Issues and applications, *J. of Empirical Finance*, 4: (1997) 73-114.
- [3] 熊谷善彰, 取引費用と目標時点を考慮した投機的市場モデル, *三田商学研究* 42, 4: (1999) 71-92.
- [4] 熊谷善彰, 取引費用と目標時点を考慮した出来高・流動性の分析, *三田商学研究* 42, 6: (2000) 95-117.
- [5] Kumagai, Y., Time-space scaling of financial time series, in *Empirical Science of Financial Fluctuations* (ed. Takayasu, H.), Springer-Verlag, Tokyo, (in press).
- [6] Mandelbrot, B. B., Fractals in physics: Squig clusters, diffusions, fractal measures, and the unicity of fractal dimensionality, *J. Stat. Phys.* 34, (1984), 895-930.
- [7] Mandelbrot, B. B. and van Ness, J. W., Fractional Brownian motion, fractional noises and applications. *SIAM Review* 10,4 (1968) 422-437.
- [8] Saupe, D., Algorithms for random fractals in *The Science of Fractal Images* (eds. Peitgen, H. and Saupe, D.), Springer-Verlag, New York, 1988.
- [9] 高安秀樹, フラクタル, 朝倉書店, 1986.
- [10] Voss, R. F., Random fractal forgeries. in *Fundamental Algorithms in Computer Graphics* (ed. Earnshaw, R. A.), Springer-Verlag, Berlin, 1985.



熊谷善彰 (正会員) 〒 108-8345 東京都港区三田 2-15-45

1998 年慶應義塾大学大学院商学研究科博士課程単位取得退学. 理学修士・商学修士. 現在, 同大学産業研究所特別研究員. 金融時系列データの研究に従事. 日本統計学会, 日本経済学会, 日本ファイナンス学会, 日本金融・証券計量・工学学会, 日本 FP 学会会員.