

ウェーブレット変換を用いたマルチフラクタル次元 の定義とその性質を用いた時系列特徴抽出

時永, 祥三

<https://doi.org/10.15017/1097>

出版情報：経済學研究. 69 (5/6), pp.51-64, 2003-03-31. 九州大学経済学会
バージョン：published
権利関係：



ウェーブレット変換を用いたマルチフラクタル次元の定義 とその性質を用いた時系列特徴抽出

時永 祥三

1 まえがき

マルチフラクタル過程は、時系列データや空間的パターンの相互依存性を分析したり、時間空間的な微細な構造を解析することに用いられる [1]-[5]。特に、フラクタル時系列は、少ないパラメータでランダムな変動を記述できるため、株価の変動のモデルなどに用いられている [3]-[5]。

本論文では、マルチフラクタル過程の定義を用いて時系列を特徴づけるとともに、これを用いた特徴抽出の方法について述べる。特に、従来のマルチフラクタル次元の定義が、時間域でなされていたのに対して、ウェーブレット変換を用いた定義を導入する [6]。一般に、フラクタル時系列はウェーブレット変換を用いることにより簡単な表現となるので、解析に有利である [8]-[12]。具体的には、ウェーブレット係数 x_n^m の高次モーメントが、マルチフラクタルと1つの関係式で結ばれることを用いる。

このようにウェーブレット係数の高次モーメントが典型的な時系列パターンごとに推定されると、時系列の特徴成分の抽出が可能となる。すなわち、フラクタル波形、周期波形、ノイズなどの基本波形ごとのマルチフラクタルの特徴をテンプレートとして準備しておき、分類が未知である時系列のマルチフラクタルとの照合を行い、認識する手法を用いる。この場合、ウェーブレット係数の高次モーメント領域での一致をとるために、遺伝的アルゴリズムを用いる [16][17]。同様に、基本波形が重畳した波形を、その基本波形に分解する操作について、本論文の手法を拡張する。更に、応用例として、実際の株価データについて本論文の手法を適用して、急激な変動などの特徴的な部分パターンを抽出する例を示す。

以下、2. では、マルチフラクタル時系列のパラメータ表現について述べる。3. では、時系列変動の表現について、ウェーブレット変換との対応関係を示す。4. ではこれらに基づく時系列特徴抽出を説明し、5. では、人工的に生成した基本波形を、本論文の手法で認識する方法について述べ、6. では、株価時系列の特徴抽出について述べる。

2 フラクタル時系列のパラメータ表現

2.1 マルチフラクタル次元とその性質

フラクタル時系列は特徴的な確率過程であり、基本的にはブラウン運動の非整数階積分(微分も含む)により定義される。代表的な表現として、次のものがある。通常の積分とは異なり非整数回微積

分する。このような時系列を fBm(fractional Brownian motion) とよんでいる。初等的な積分の理論を用いれば、正弦波の複素数表現である $\exp(i\omega t)$ を n 回微分した関係式は、次のようになる [3]。

$$d^n/dt^n \exp(i\omega t) = (i\omega)^n \exp(i\omega t) \quad (1)$$

ここで、 ω は角周波数とよばれ、周期的な波形の周期を T としたときに

$$\omega = 2\pi/T \quad (2)$$

により与えられる。また、複素数の指数関数は次のようになる。

$$\exp(i\omega t) = \cos(\omega t) + i \sin(\omega t) \quad (3)$$

n が負である場合には積分と見なされる。これを複数の周波数を含む一般のケースに拡張する。もとの時系列を $f(t)$ とし、そのフーリエ変換を $F(\omega)$ とすると、微積分された結果について、フーリエ変換により相互に次の関係式が成立する。

$$d^n f(t)/dt^n = (2\pi)^{-1} \int_{-\infty}^{\infty} (-i\omega)^n F(\omega) \exp(i\omega t) d\omega \quad (4)$$

ここで、もとの時系列 $f(t)$ をブラウン運動とすると、そのスペクトルは次のようになる。ただし、 $\alpha = 5 - 2D$, $D = D_B = 1.5$ 。なお、この式は一般のフラクタルについても成立する (ただし、 $D \neq 1.5$)。

$$S(\omega) \propto \omega^{-\alpha} \quad (5)$$

一方、これを微積分した関数のスペクトルは、次のようになる。

$$S(\omega) \propto \omega^{-\alpha+2n} \quad (6)$$

従って、 $1 < \alpha < 3$ であるとき、関数のグラフのフラクタル次元を D とすると、

$$D = D_B + n \quad (7)$$

という関係が成立する。

ここで、時系列の性質を決めるものとして、フラクタル次元 D および分散 σ^2 がある。簡単にいえば、フラクタル次元の大きな時系列の波形は滑らかになり、次元が小さい場合には細かな変動を含む波形となる。また、分散は波形の全体の振幅の大きさを制御するパラメータとなる。

フラクタル次元は、もとの時系列がフラクタル時系列でない場合にはあまり意味のある指標ではないが、一般の時系列の性質を大まかにとらえる場合に、拡張して用いられている。また、時系列だけでなく、画像解析にまで拡張して用いられている。この場合には、2次元のウェーブレット変換を用いる。

フラクタル次元は、1つのフラクタル時系列に対して1つだけ定まるものであるので、これを計算するには、ウェーブレット変換やスペクトル、あるいはハースト指数 H を計算するための時系列の算術計算を実行する。 $H = 2 - D$ で定義される。特に、ウェーブレット係数の対数を取り、dilation index を横軸としてプロットすると、線形となるので、この傾きからフラクタル次元を算出することができる (分散は切片から計算する)。

このように、フラクタル次元はフラクタル時系列の特徴を示すには便利な指標であるが、一般の時系列には適用できない。あるいは、画像解析に関しても、フラクタル画像あるいはフラクタル符合化には有効であるが、一般の画像には適用できない。

このような制限を除去するため、次元の定義を拡張することが行われている。これは、一般化次元とも呼ばれているが、次のような定義を用いる。

$$D_q = \lim_{r \rightarrow 0} \frac{1}{q-1} \frac{\log \sum_i p_i^q}{\log r} \quad (8)$$

ここで、 p_i は、対象が幾何学的に r のサイズで分割された場合に i 番目の特徴パターンが存在する確率であり、 D_q は一般化次元である。

例えば、画像に関して、この計算を実施するには、白黒の画像を仮定し、画像の中の部分に含まれる画素の数をカウントし、これから分布の密度関数を推定する。この操作を、画像全体について実施すると、画素の密度が計算できる。

しかし、このままでは単独の画像の粗さしか判定できないが、この確率について指数をとる(べき乗する)ことにより、強調する度合を変化させることができる。すなわち、指数が正である場合には密度の高い部分の存在が強調され、指数が更に大きくなるとこの傾向が強まる。これに対して、指数をマイナスにしていくと、密度の小さい部分が強調されることになる。

すなわち、フラクタル的である画像は、画像の部分を細分化した場合に、その密度が相対的に変化しないので、次元は一定値に近くなる。すなわち、細かな変動を含む画像のマルチフラクタルは大きくは変化しない。これに対して、濃淡が明確である画像についてはマルチフラクタルの変化が大きい。これを模式的に示したものが図1である。

このように、マルチフラクタルを導入することにより、時系列や画像の微細な構造や性質を判別することが可能である。経済時系列解析の分野でいえば、株価の変動を、これまでとは異なる基準で判別することができる。

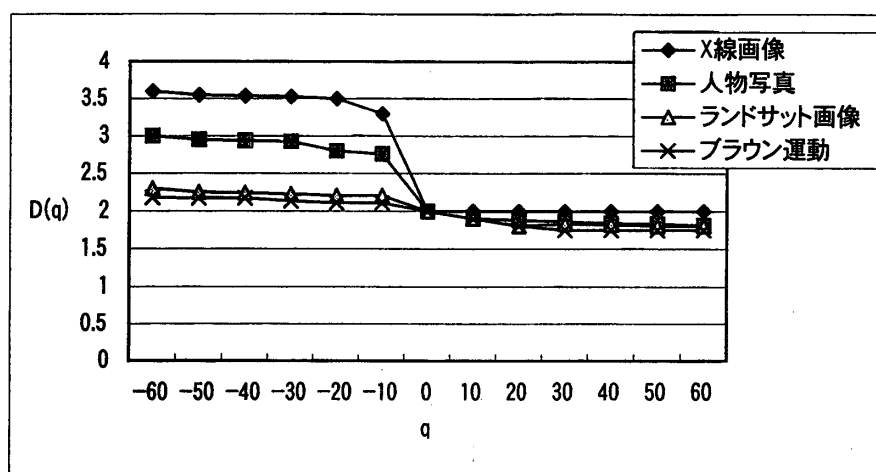


図1. マルチフラクタルによる観測データの詳細分析

2.2 フラクタル時系列とウェーブレット変換

時系列を認識したり正弦波などの信号要素に分解する場合には、時間域での表現をそのまま用いることが可能であるが、フラクタル時系列の場合には、時系列のウェーブレット変換を用いる方が便利である。

最初に、フラクタル時系列のウェーブレット変換の性質について述べておく。フラクタル時系列は非定常時系列であるので、正確な意味ではスペクトルは存在しないが、平均的な形として、次の式により与えられる [13]。

$$S(\omega) = \sigma^2 \omega^{-\gamma}, \gamma = 5 - 2D \quad (9)$$

ここで D, σ^2 は時系列のフラクタル次元と分散である。

フラクタル時系列のウェーブレット変換と分散、フラクタル次元との関係を求める。与えられた時系列 $x(t)$ をウェーブレット変換する。

$$x(t) = \sum_n \sum_m x_n^m \psi_n^m(t) \quad (10)$$

$$x_n^m = \int_{-\infty}^{\infty} x(t) \psi_n^m(t) dt \quad (11)$$

$\psi_n^m(t)$ はウェーブレット基本関数 $\psi(t)$ に対する次のスケール、シフト変換とにより構成される。

$$\psi_n^m(t) = 2^{m/2} \psi(2^m t - n) \quad (12)$$

ここで、 m, n は、スケール変換、シフト変換のインデックス (dilation index, translation index) である。インデックスのとおり得る範囲を $m \in M, n \in N$ と表現する。 $\psi(t)$ としては Daubechies のウェーブレットを用いる [13]。

$x(t)$ がフラクタル性をもつことから、ウェーブレット係数 x_n^m の満たすべき条件として、次の関係式が得られる [14]。

$$\text{var}(x_n^m) = \sigma^2 2^{-\gamma m} \quad (13)$$

この関係式は、ウェーブレット基本関数の性質などを用いると証明できる。いま、式 (10) の両辺のスペクトルをとると、次のようになる。

$$S(\omega) = \sum_m P_m(\omega) |\Psi(2^{-m}\omega)|^2 \quad (14)$$

ここで、 $\Psi(\omega)$ はウェーブレット基本関数 $\psi(t)$ のフーリエ変換、また、 $P_m(\omega)$ はウェーブレット係数のスペクトルである。 x_n^m が局所的に変化しているので、時間を表す添字 n について計算した値である。ウェーブレット基本関数のフーリエ変換は帯域フィルタの特性をもっているので、式 (14) が、前の式 (9) と一致するには

$$P_m(\omega) = \sigma^2 2^{-\gamma m} \quad (15)$$

となることが必要であることが分かる。すなわち、 $P_m(\omega)$ は ω に依存しない関数となる。 x_n^m から求めた分散とスペクトルとは等価であるので、式 (13) が成り立つ。

この両辺の対数をとると、 m について線形の関係式が得られる。このデータに回帰直線を当てはめ、直線との2乗平方誤差 R_w の大きさによりフラクタル性を判定できる。シミュレーションにより R_w を求めると、0.002 程度の極めて小さな値となる。

2.3 マルチフラクタルへの拡張

容量フラクタル次元を拡張したものとしてマルチフラクタル次元が導入され、一般に、次のような式で定義されている [3] (前に上げた式 (8) の再掲)。

$$D_q = \lim_{r \rightarrow 0} \frac{1}{q-1} \frac{\log \sum_i p_i^q}{\log r} \quad (16)$$

ここで、 p_i は、対象が幾何学的に r のサイズで分割された場合に i 番目の特徴パターンが存在する確率であり、 D_q は一般化次元である。

この次元 D_q は、 q の値に応じて確率の小さいサンプルを過剰に評価するか、あるいは、逆に、確率の高いものを過剰に評価するかの違いとして解釈される。時系列や画像などの実際のデータ解析の結果から言えるように、変動が明確な時系列や、パターンが明確になっている画像では、次元の変化が大きくなりランダム時系列やランダム画像では、次元の変化が小さい。

フラクタル時系列に対して、このマルチフラクタルをウェーブレット領域で表現すると、 q 次のフラクタル次元 (q は偶数とする) の大きさ D_q と、ウェーブレット変換係数の q 次モーメント ${}^{(q)}x_n^m$ との間に次のような関係式を得る。

$${}^{(q)}x_n^m = \sigma^2 2^{-\gamma m q / 2} \quad (17)$$

すなわち、式 (13) のようにウェーブレット係数の 2 次のモーメントを計算した式が、スペクトルと直接関係しており、これからフラクタル次元が計算できたように、 q 次のウェーブレット係数のモーメントから一般化次元を計算することができる。これは、次のように示すことができる。

いま $q = 2$ の場合を考察する。時系列 $y(t)$ を a 倍に伸長した時系列を $y_a(t)$ とすると、この時系列の自己相関関数 $R_{y_a}(\tau)$ と、もとの時系列の自己相関関数 $R_y(\tau)$ の間には、次の関係が成立する。

$$R_{y_a}(\tau) = E[y_a(t)y_a(t-\tau)] = a^{-2H} R_y(a\tau) \quad (18)$$

$$H = 2 - D, \gamma = 2H + 1 \quad (19)$$

この両辺をフーリエ変換すると、変数変換の関係式を用いて、次の式が得られる。

$$S_{y_a}(a\omega) = a^\gamma S_y(\omega) \quad (20)$$

スペクトルの形状は、 $S_{y_a}(\omega)$ と $S_y(\omega)$ とでは同じなので、次が得られる。

$$S_y(a\omega) = a^\gamma S_y(\omega) \quad (21)$$

次に、式 (21) の両辺を ω で偏微分して、 $a \rightarrow 1$ とすると、次の式を得る。

$$\omega S_y'(\omega) = -\gamma S_y(\omega) \quad (22)$$

この微分方程式が成立すべき関係を求めることにより、式 (18) が得られる。

この計算を q 次のモーメント計算に拡張する。すなわち、 $q = 2K$ とすると、 q 次の自己相関関数は

$$R_{y_a}(\tau_1, \tau_2, \dots, \tau_q) = E[x_a(t)y_a(t-\tau_1)y_a(t-\tau_2)\dots]$$

$$y_a(t - \tau_q)] = a^{-2HK} R_y(a\tau_1, a\tau_2, \dots, a\tau_q) \quad (23)$$

この両辺をフーリエ変換すると、式 (18) および変数変換の関係式を用いて

$${}^{(q)}S_y(a\omega) = a^{-\gamma K} S_y(\omega) \quad (24)$$

が求められる。

次に、式 (24) の両辺を ω で偏微分して、成立すべき関係を求めることにより、次に示す式が得られる。

$${}^{(q)}S(\omega) = \sigma_y \omega^{-\gamma K} \quad (25)$$

この表現まで得られると、式 (13) と同様にして、ウェーブレット係数による関係式を求めることができる。

3 マルチフラクタルの特徴表現

3.1 時系列変動モーメントとマルチフラクタル

マルチフラクタルには、時間不変のいくつかの性質が知られている。以下では、前に示したウェーブレット係数の高次モーメントと、時系列変動モーメントとの関係を考察する [5]。

いま、次で定義される時系列 $x(t)$ の変動の q 次モーメントを求めてみる。

$$M(q, l) = E(|\delta_l(x(t))|^q) = E(|x(t+l) - x(t)|^q) \quad (26)$$

この式の指数表現として、次が用いられる。

$$M(q, l) = K_q l^{\zeta_q} \quad (27)$$

ここで、指数部分 ζ_q を ζ_q スペクトルと呼んだりする [5]。

ブラウン運動、あるいは fBm (fractional Brownian motion) などの典型的な自己相似時系列においては、時間軸を伸長した時系列ともとの時系列の変動に関して、次のような関係がある。

$$\delta_l x(t) = (l/L)^H \delta_L x(t/l) \quad (28)$$

これを用いて、次のような関係が得られる。

$$\zeta^q = qH \quad (29)$$

一般の時系列では、このような線形の関係は見られないで、モーメント次数を表す変数 q に関して、非線形の convex 関数となる。

すでに述べたように、時系列の高次スペクトルとウェーブレット係数のモーメントとは等価な関係にあり、また、フラクタル時系列の係数の高次モーメントの対数は、フラクタル次元に関して線形の関係にある。

従って、時系列のマルチフラクタルとウェーブレット係数のモーメントとが直接結ばれる関係式が得られると、時系列を特徴づける場合に便利である。

図2に示しているのは、文献[5]において解析されている円先物価格と、これを理論的にモデル化してマルチフラクタルによるモデルで生成した時系列の例である。図3には文献[5]において定義されている $M(q, l)$ の例を示している。この図3の例では時系列に対する指数を求めている。次元は、ほとんど指数について線形であり、係数部分の影響は小さくなっている。

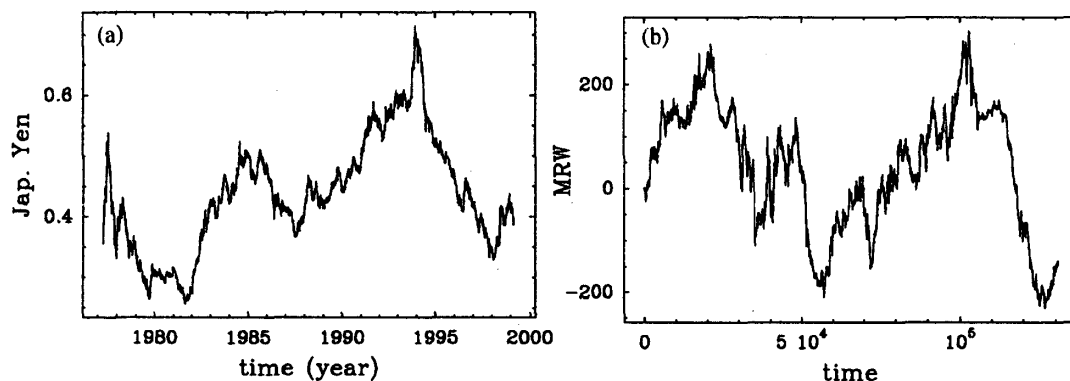


図2. マルチフラクタル時系列の例 (左が円先物価格, 右がマルチフラクタルによるモデルで生成した時系列)

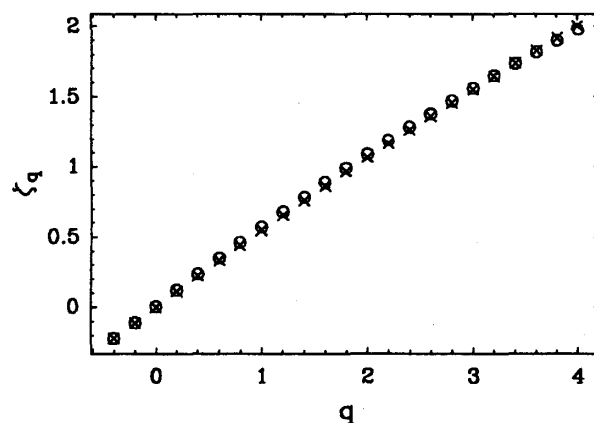


図3. ζ_q スペクトルの例 (○が円先物価格, ×がマルチフラクタルによるモデルで生成した時系列の場合)

3.2 一般化次元と変動のモーメント

これまで、マルチフラクタルの定義として、時系列の変動分のモーメントの表現を用いてきた。この定義と、式(16)に示す一般化次元との相互関係は、次のように説明できる。

いま、式(16)における確率 p_i を、時系列が時刻 t から時刻 $t+l$ までに変化する大きさである $\delta_{ix}(t)$ の分布であると解釈する。例えば、変動の単位を Δx とした場合に、 i 番目の変動の大きさは $i\Delta x$ であり、この大きさの変動が発生する確率が p_i である。変動は正と負とがあるので、 i は正と負にわたる数値である。 $|i\Delta x|$ が大きい場合には、急激な変動が観測される。

これに対して、変動分のモーメントを計算するように利用する方法を考える。確率の q 次のモーメントではなく、変動分を用いることにより、現実の観測データとの対応がとりやすい。

これらを考慮して、次のような一般化次元の定義式を提案する。この定義はマルチフラクタルによる画像分類に用いられており、効果を上げている。

$$D_q = -\lim_{l \rightarrow 0} \frac{1}{q} \frac{\log \sum_i (i \Delta x)^q p_i}{\log l} \quad (30)$$

この式では、式 (16) で定義される一般化次元と同様に、時系列変動分のモーメントは、 q によって変動分の大きな部分を強調したり、小さな部分を強調したりすることに対応している。出現確率 p_i により期待値を計算している。

また、式 (27) の対数をとると、 $\log M(q, l) = \log K_q + \zeta_q \log l$ となるので、 ζ_q/q という値は、一般化フラクタル次元に対応していると考えられる。

3.3 ウェーブレット係数による表現

時系列の変動分 $\delta_l(x(t))$ に関する表現を、ウェーブレット変換を用いて示してみる。式 (26) における時間域における差異を $l = \Delta n$ と表す。時系列の変動分 $\delta_l x(t)$ の q 次モーメント期待値の値は、次のようになる。

$$M(q, l) = E\left[\left(\sum_m \sum_n x_n^m 2^{m/2} \psi(2^m t - n + \Delta n) - \sum_m \sum_n x_n^m 2^{m/2} \psi(2^m t - n)\right)^q\right] \quad (31)$$

この式を展開して計算を行うが、最初に q が偶数の場合 ($q = 2K$) を考える。 $q = 2$ の場合、ウェーブレット関数の直交性、およびウェーブレット関数の 2 乗の期待値が一定の値となることを用いると、次の表現を得る。

$$M(q, l) = cE\left[\sum_n \sum_m (x_n^m)^2\right] \quad (32)$$

従って、一般の偶数次数の場合の期待値を求めると、次のようになる。

$$M(q, l) = cE\left[\sum_m \sum_n (x_n^m)^K\right] \quad (33)$$

このように、次数が偶数の場合には、フラクタル時系列と、これ以外の時系列との間の、時系列の変動分 $\delta_l x(t)$ のモーメントにおける表現には差異がない。次数が奇数の場合には、モーメントは負になるケースも存在するが、同様の表現を求めることができる。

3.4 ウェーブレット係数のモーメントによる特徴づけ

次に、 ζ_q スペクトルとウェーブレット係数との間の関係を議論する。これまで述べたように、時系列の変動分 $\delta_l x(t)$ のモーメントを、ウェーブレット係数により求めることができる。式 (27) に示すように、 ζ_q スペクトルと一般化フラクタル次元との間に関係が成立するとが予想され、特に、時系列がフラクタルの場合には、一般化フラクタル次元が一定となり、表現が簡単化できる。しかし、この場合でも、最初にかかっている係数 K_q は、フラクタル時系列の場合のように解析的に求まるケースでも

簡単ではない。

係数 K_q の非線形性の影響は無視できないので、式 (31) の両辺の対数を取り、指数 q で割った値である $\log M(q, l)/q$ はフラクタル時系列の場合でも一定値ではない。このようなことを考慮して、以下では、ウェーブレット係数の高次モーメントの対数のパターンを直接用いることにする。

$$\log^{(q)} x_n^m \quad (34)$$

時系列ごとの特徴的なパターンについて、モーメント次数ごとに簡潔なグラフとしてまとめることも可能であるが、本論文では時系列の認識に重点を置いているので、次数 q ごとにテンプレートを用意して、これを認識に用いることにする。更に、ウェーブレット係数のモーメントの特徴について、ウェーブレット係数の dilation index m を横軸としての特徴を描く方法を用いる。これを以下では簡単のためウェーブレット係数モーメント特徴パターンとよんでおく。

多くのケースでは、時系列は正弦波などの確定的な変化をする成分と、フラクタル信号、およびノイズで構成されていると仮定できる。図 4 にはいくつかの典型的な時系列に対するウェーブレット係数モーメントを示している。この特徴は、次のようにまとめられる。

(1) フラクタル時系列

ウェーブレット係数モーメントの対数は dilation index m についてすべて線形となる。

(3) 正弦波

1 次モーメントでは特定の dilation index の付近に大きな値が発生する。この両側にむかっての値は線形的に減少していく。

(3) ノイズ

全体に平坦なパターンとなるが、中央付近で最大となってから両側にかけて 4 次関数の形状で緩やかに減少していく。

これより分かるように、それぞれの基本波形のパターンの違いは大きいので、時系列の種類の特定特徴部分の抽出に有効である。

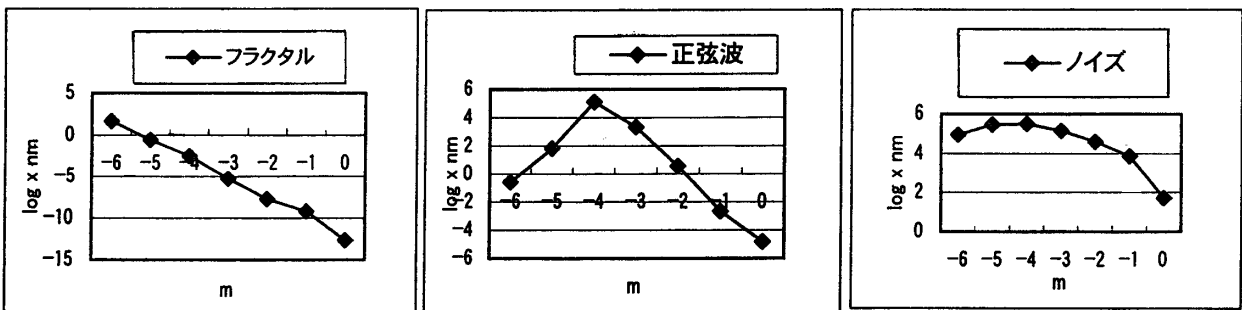


図 4. 基本波形のウェーブレット係数 2 次モーメントの対数

4 マルチフラクタル観測による特徴部分の抽出

4.1 時系列分類とマルチフラクタル

以下では、マルチフラクタルの応用として、ウェーブレット係数から計算した一般化次元の変化により時系列の特徴と思われる部分を抽出する方法を提案する。

本論文では、フラクタル時系列を基本的な時系列であると仮定し、これにさまざまな種類の時系列が重畳しているとする。これまでの議論で明らかのように、ウェーブレット変換の高次モーメントを計算することにより、一般化次元の大きさを知ることができるので、これを用いた、やや詳細な次元の変化を検出することができる。

図 4 に示すように、ウェーブレット係数 2 次モーメントの対数の領域でも、おおよその特異的な時系列の発生を検出することができる。しかし、得られる情報が制限されることと、株価変動など、一般に時間域での変化を知ることが必須となるケースが多い。従って、本論文では、高次モーメントまで用いたパターンの検出を行う。

$$d_i = \sum_q \sum_{m \in M} (\log^{(q)} x_n^m - \log^{(q)} y_n^m)^2 \quad (35)$$

ここで、 d_i は i 番目の近似計算における誤差であり、 $\log^{(q)} x_n^m, \log^{(q)} y_n^m$ は、それぞれ、基本波形、認識対象である時系列のウェーブレット係数のモーメント (n についての期待値) である。

このアルゴリズムを次に示す。

(1) テンプレートの準備

時系列がいくつかの基本波形に分解可能と仮定し、これらの基本波形についてウェーブレット変換の高次モーメントを計算し、パターン照合のためのテンプレートとして準備する。

(2) ウェーブレット係数の高次モーメント計算

式 (9) に従って、認識の対象とする時系列のウェーブレット係数の高次モーメントを計算する。この場合、基本波形の発生時刻が特定されている場合には、直接的に GA によりパターン照合を行う。そうでない場合には、過渡的な波が発生する部分を特定するために、ウェーブレット係数を計算する時間域を重複させながら、移動させて計算を行う (スライディングウィンドウ)。

(3) 過渡的な変化部分の特定

ウェーブレット係数の高次モーメントのパターンを観測して、これらが直線から乖離しているかどうかを検出する。ウェーブレット係数の高次モーメントの変化パターンをテンプレートとして用いて、おおよその時系列成分を推定しておく。具体的には、それぞれの成分を仮定し、これらの高次モーメントの合計が、観測された高次モーメントのパターンに近くなるように、成分にかける重みを推定していく。この操作に GA を適用する。

(4) 基本波のパラメータ推定

フラクタル時系列などの成分に分解したあとに、フラクタル時系列の成分のパラメータ、すなわち、フラクタル次元 (相似性次元) および分散を推定する。正弦波の場合には中心周波数とレベル、ノイズの場合にはノイズレベルを推定する。

5 応用例

5.1 人工的に発生させたフラクタル時系列

以下では、本報告で述べた時系列の特徴抽出について、あらかじめ出現するパターンが分かっている場合について、シミュレーションをもとにして、認識の能力を確認する。これまでの議論で明らかのように、マルチフラクタルの視点から見た場合に、特徴的な時系列パターンは、フラクタル時系列、周期波形、およびランダムノイズに分類できるであろう。フーリエ展開可能な波形は、周期波形の集合であり、周波数が近接していない場合には、その周波数まで分離できる可能性がある。

最初に、時間的に区分された領域（以下ではセグメントとよぶ）において、上に述べた3つの波形が、一定の時間だけ持続して出現するケースを考察する。それぞれのセグメントの長さは、マルチフラクタルの検出能力を考慮して、1000 サンプルとしておく。また、それぞれのセグメントにおける振幅の値も、同程度となるように調整しておく。

パターン認識の方法としてテンプレートを用いるが、後で述べる複数の基本波形の重畳として時系列が表現された場合への拡張を考慮して、遺伝的アルゴリズム (GA: Genetic Algorithm) による方法を用いる。GA において、個体を構成するビットとしては、次のようなものを仮定する。

(1) フラクタル時系列の場合

フラクタル次元および分散

(2) 周期波形の場合

周波数および振幅値

(3) ノイズの場合

ノイズレベルおよび4次の関数の係数

このような未知のパラメータを推定するシステムを、それぞれ別々に構成して、最も推定の誤差が小さくなるケースを最終的なグループとして確定する。すなわち、ストリングがフラクタル次元と分散であるシステム、周波数と振幅値であるシステム、ノイズレベルの3つのシステムを独立的に作動させこのシステムの適合度の最大値が求まった時点でどのシステムの予測誤差は最小であるかにより認識結果を決定する。

個体 i の適合度としては、次で定義されるような基本波形についてのウェーブレット係数で表現された変動分のモーメント q ごとの基本パターンと、同様にウェーブレット係数で表現された認識対象の時系列の変動分のモーメント q ごとの値の2乗誤差の逆数である。実際にはモーメント q ごとのウェーブレット係数の個数は M_r 個であることに注意する。式 (35) の近似誤差の逆数を用いて、 i 番目の個体の適合度を $S_i = 1/d_i$ により計算する。

このようなシミュレーションを、50 個の時系列について実施した結果を表に示している。これらのデータについての認識率に大きく影響するのが、時系列に付加されるノイズレベルである。従って、シミュレーションの結果を整理する場合に、ノイズレベルに応じた認識率を示している。

具体的には、平均ゼロ分散1の正規乱数 $\varepsilon(t)$ を発生しておいて、生成された時系列 $x(t)$ に対して、 $R\delta\varepsilon(t)$ を加算した時系列に対して、上に示す認識方法を適用する。ただし、 R は時系列の振幅である。 δ が大きいほど、他のカテゴリに分類される可能性が大きくなることが予測される。

表1では生成された時系列のカテゴリと、認識されたカテゴリの一致度を認識率として示している。これより明らかのように、本報告で述べる時系列認識の手法では、ノイズレベルが比較的大きな場合

でも、正確な認識が可能となっている。

表 1. 基本波形の時系列の認識結果 (判別誤り)

	category 1	category 2	category 3
$\delta = 0.3$	0.12	0.15	0.12
$\delta = 0.2$	0.04	0.04	0.04
$\delta = 0.1$	0.0	0.0	0.0
$\delta = 0.02$	0.0	0.0	0.0

5.2 基本波形の重畳

ここで示すような、マルチフラクタルを用いた時系列認識の有効性を評価するために、以下では、一般的に認識で用いられる方法との比較分析を行う。具体的には、ウェーブレット係数の 2 次モーメントだけを用いた時系列分解の方法との認識精度の比較である。

前の例題としてとりあつかっているように、3 つの基本波形が重畳されたケースを仮定し、その基本波形を分離し推定する方法である。重畳された基本波形のパラメータの推定の方法として GA を用いる [16][17]。GA における個体のストリング構成は 6 つのパラメータを接続したものである。

表 2 に認識結果を示す。結果にはそれぞれの基本波形が含まれていると判定される確率を示しており欠損した場合には識別誤りとしている。これより分かるように、マルチフラクタルに相当するウェーブレット係数の高次のモーメントを用いた認識により分離が可能なケースの限界における付加されたノイズレベルが推定できる。いる。

表 2. 基本波形の抽出結果 (判定誤差)

	category 1	category 2	category 3
$\delta = 0.3$	0.33	0.34	0.34
$\delta = 0.2$	0.11	0.12	0.15
$\delta = 0.1$	0.02	0.04	0.04
$\delta = 0.02$	0.01	0.02	0.02

6 株価変動の検出

6.1 株価時系列のフラクタル性

株価のフラクタル性については多くの研究事例があるので、以下では、簡単に特徴だけをまとめる。株価のフラクタル次元は、ほぼ、 $D = 1.5$ であり、フラクタル性を検証する R_w の値も 0.02 以下であり、小さい。このように、ブラウン運動の仮定は一般には成立している。この値は、業種や観測期間に依存しない性質である [7][11]。

しかし、日々の相対リターンの対数値をとり、分布を調べると、正規分布からずれていることが検証できる例が多い。更に、相対リターンの自己相関について、ほとんどゼロであるが、1 次のラグだけは絶対値は小さいが、負になっている。このことは、市場の効率性を否定する根拠ともなる。

株価の分散である volatility についても、一定であることを前提に、オプション価格の推定などが行われているが、ARCH テストなどを実施すると、多くのケースで一定の仮説が棄却される。

このように株価変動のベースになる性質は変わらないが短時間の範囲でこの仮定を乱すような変化が発生していると考えられる。

6.2 株価時系列の特徴抽出

以下ではマルチフラクタルの状況だけを観測して株価時系列にフラクタル性以外の変動要因が加わっているかを検出するシミュレーションを行う。なお、本論文の方法を用いてどのような変動成分が含まれているかを特定することもできるが、スペースの関係で省略する。

フラクタルからの乖離を測定すればよいので、ウェーブレット係数の高次のモーメントの対数に、 m を変数とする回帰直線を当てはめ、この両者の 2 乗誤差を求める。

$$r^2 = \sum_{q=2,4,6} \left(\left[\sum_{m \in M} \log^{(q)} x_n^m - a_0 - a_1 m \right]^2 \right) / M_T \quad (36)$$

表 3 には、1983-1993 年において日本の証券市場で観測された時系列に対する検証の結果を示している。式 (36) の値が、表に示すしきい値 T_i を越える確率 p を求めている。横の欄は観測期間の区分である。

特に統計的な検定をする手段はないので、典型的な時系列について計算されたマルチフラクタルの値から推定すると、しきい値を 0.3 に設定し、これを越えるケースを特定した場合、株価にはフラクタル以外に短期的な変動が含まれていると言える。

表 3. 株価におけるマルチフラクタルの観測

	1983-1987	1988-1993
$T_1 = 0.5$	0.01	0.02
$T_2 = 0.3$	0.09	0.12
$T_3 = 0.1$	0.19	0.18

7 むすび

本論文では、マルチフラクタル過程の定義を用いて時系列を特徴づけ、これを用いた特徴抽出について述べた。すなわち、ウェーブレット係数 x_n^m の高次モーメントが、マルチフラクタル次元と 1 つの関係式で結ばれることを用いた。応用例として、フラクタル波形、周期波形、ノイズなどの単独の波形が連続する場合の認識株価の特徴抽出についてシミュレーション結果を示した。

今後の課題として、さまざまな現実の時系列認識への応用、フラクタル表面などの 3 次元データへの拡張があり、今後検討していく予定である。

参考文献

- [1] B.Mandelbrot: The Fractal Geometry of Nature, Freeman, San Francisco, 1982.

- [2] B.Mandelbrot and N.Van: “The fractional Brownian motions, fractional noises and applications”, SIAM Rev.,10, pp.422-436 ,1968.
- [3] B.Mandelbrot: “Multifractal measure, especially for the geophysicist”,Pure and Applied Geophysics, 131, pp.5-42, 1989.
- [4] B.Mandelbrot:Fractals and Scaling in Finance, Risk New York, Springer, 1997.
- [5] E.Bacry,J.Delour and J.F.Muzy: “Modeling financial time series using multifractal random walks”,Physica A, vol.229, pp.84-92,2001.
- [6] 高木 昇,時永 祥三: “マルチフラクタル時系列の性質を用いた特徴抽出とその応用”,電子情報通信学会技術研究報告,NLP2002-61,pp.75-80,2002.
- [7] 時永祥三: 「複雑系による経済モデル分析」,九州大学出版会,2000.
- [8] 時永祥三,森保 洋,宮崎明雄,島津宣之: “時系列のフラクタル性を用いた予測手法とその応用”,信学論 (A), vol.J79-A,no.11,pp.1793-1800,1996.
- [9] 時永祥三,森保 洋,宮崎明雄,島津宣之: “スケール伸長変換およびウェーブレット変換によるパラメータ推定を用いた時系列予測”,信学論 (A),vol.J79-A,no.12,pp.2054-2062, 1996.
- [10] 池田欽一,時永祥三: “スケール伸長変換を用いたフラクタル表面の推定とその応用”,信学論 (A), vol.J81-A,no.8, pp.1127-1133,1998.
- [11] 池田欽一,時永祥三: “フラクタル時系列の予測手法を用いた株価予測とその応用”,日本オペレーションズリサーチ学会論文誌, vol.40,no.1,pp.18-31, 1998.
- [12] 竹林 渉,植村芳雄,時永祥三: “フラクタル時系列の性質を用いた需要予測の一手法”,日本オペレーションズリサーチ学会誌,2000年2月号,no.2,pp.51-59, 2000.
- [13] I.Daubechies: “Orthonormal bases of compactly supported wavelets”, Commun.Pure Appl.Math, vol.41,no.7,pp.909-996 (1988).
- [14] G.W.Wornell and A.V.Oppenheim: “Estimation of fractal signals from noisy measurement using Wavelets”, IEEE Trans. Signal Processing, vol.40,no.3, pp.611-623 , 1992.
- [15] G.W.Wornell: “Wavelet-based representation for the 1/f family of fractal process”, Proc. IEEE, vol.81,no.10, pp.1428-1450 ,1993.
- [16] J.H.Holland: *Adaption in Natural and Artificial Systems*, Univ.of Michigan Press, 1975, MIT Press, 1992.
- [17] D.E.Goldberg: *Genetic Algorithm in Search, Optimization and Machine Learning*, Los Altos, California, Morgan Kaufmann, 1989.

[九州大学大学院経済学研究院 教授]