

フラクタル時系列の性質を用いた 建材需要予測の一手法

竹林 渉, 植村 芳雄, 時永 祥三

1. まえがき

本研究では、フラクタル時系列の性質を用いた建材需要予測の一手法を示し、予測誤差の検討、フラクタル次元の推定、実際の需要予測の例を与える[1]。製品の需要予測についてはさまざまな方法が提案されているが、その多くが線形予測など時系列処理に基づく方法であり[2][3]、時系列が非定常であるケースでは予測誤差が大きくなる問題が存在する。

ところで、株価がフラクタルであることはいくつかの事例で示されており、需要データの1つである通信トラヒックについても同様の性質をもつケースが指摘されている[4][5][6][7]。本研究で示すように、需要データがある条件を満足して発生している場合には、その集計結果はフラクタルとして近似できることがわかる。われわれは、この現象を実際に住宅建材の需要にあてはめて分析を行うとともに、すでに開発しているフラクタル時系列の予測手法を適用して需要予測を行ってみる[8][9]。

フラクタル時系列の予測の原理は、フラクタル時系列をスケール関数の展開形式で表現されたインパルス応答関数と入力信号との畳み込みにより表現するモデルを仮定し、時間軸の伸長に対してインパルス応答の自己相似的な性質が保持されることを用いている[8][9]。研究では、実際に約60種類の建材需要への適用を行い、有効性を示す。

なお、本研究ではフラクタル性をもつ時系列に関して、自己相似性を用いれば経験的に予測精度のよい方法が得られることを示しており、この方法がフラクタル

理論と整合しているかの理論的検討などは別の機会に譲ることにする。

2. 需要の累計値のフラクタル性

2.1 ON/OFF ソースによるフラクタル時系列の生成

通信トラヒックのモデル化に関連して、最近、ON/OFF ソースを重ね合わせることにより、フラクタル時系列が生成できることが理論的に示され、LANなどのトラヒック解析に利用されている。ON/OFF ソースとは、ON 区間と OFF 区間が交互に現れるモデルであり、ON 区間では一定の間隔でパケット系列が発生している。パケット系列はLANなどの通信においてはネットワークを介して送信されるコンピュータデータであるが、製品需要の発生としてモデル化することも可能であろう。ON/OFF ソースの重ね合わせがフラクタルとなることの基本的な考え方は、統計的に独立である ON/OFF ソースを統計的に多重化すると Gaussian Noise となり、これを一定の時間区間にわたって積分すると fBm (fractional Brownianmotion) になることを利用している。fBm とはブラウン運動を非整数階積分して得られるものであり、フラクタル時系列の発生方法として用いられている。

本研究では同様な考え方により、日々の需要データが Gaussian Noise として近似でき、これを数日間にわたり集計することにより fBm となることを用いている。

以下では、文献[6][7]に示された、ON/OFF ソースを統計的に多重化した場合の時系列の性質についての理論式を引用しておく。この理論式は本研究では特に用いないが、需要カーブがフラクタルとなる根拠を与えている。

いま M 個の ON/OFF ソースを仮定する。関数 $W(t)$ を ON/OFF ソースにより発生される時系列と考え、 $W(t)=1$ であれば時刻 t でパケットがあると

たけばやし わたる 九州大学経済学部
〒812-8581 福岡市東区箱崎 6-19-1
うえむら よしお (株)住建産業
〒738-0022 廿日市市木材港南 1-1
ときなが しょうぞう 九州大学経済学部
〒812-8581 福岡市東区箱崎 6-19-1
受付 98.9.11 採択 99.9.10

し、 $W(t)=0$ であればパケットがないと定義する。ソースの区別を添字 m により行い、 M 個のソースを合計して得られる次の時系列を考える。

$$W_M(t) = \int_0^{Tt} \left(\sum_{m=1}^M W^{(m)}(u) \right) du \quad (1)$$

このとき、十分に大きな M と T をとった場合には、式(1)は次の関数に収束することが証明されている。

$$Y(t) = TM\mu_1 / (\mu_1 + \mu_2) + (T^{2H}L(T)M)^{1/2} \sigma_{lim} B_H(t) \quad (2)$$

ここで、 $B_H(t)$ は Hurst 指数が H である fBm である。 μ_1, μ_2 は ON, OFF 区間の長さの平均値であり、 $L(T)$ は ON, OFF 区間の長さの分布補関数が、変数が十分に大きな場所においてとる分布の近似形から得られる定数である。すなわち、ON, OFF 区間の長さの確率密度関数を $f_1(x), f_2(x)$ とした場合の分布関数を $F_1(x), F_2(x)$ としたときに、 $F_{1c}(x) = 1 - F_1(x)$, $F_{2c}(x) = 1 - F_2(x)$ を $x \rightarrow \infty$ において

$$F_{1c}(x) \sim l_1 x^{-\alpha_1} L_1(x), 1 < \alpha_1 < 2 \quad (3)$$

$$F_{2c}(x) \sim l_2 x^{-\alpha_2} L_2(x), 1 < \alpha_2 < 2 \quad (4)$$

として表現する。ここで、関数 $L_1(x), L_2(x)$ は x が大きい場合にゆっくり変動する関数 (slowly varying function) であり、定数や $\log x, (\log x)^{-1}$ などが相当する。また、 σ_{lim} は $L_1(x), L_2(x), \mu_1, \mu_2, \sigma_1, \sigma_2, \alpha_1, \alpha_2$ により表現される (σ_1, σ_2 は ON, OFF 区間の長さの分散、詳細は省略する)。

式(2)の証明は文献[6][7]に示されている。式(2)から次のような解釈が成り立つ。すなわち、ON/OFF ソースを統計的に多重化した時系列 $Y(t)$ は、観測区間全体に占める ON 区間の長さの割合に比例したトレンドをもち、このトレンドからの変化が fBm により記述できる。式(2)は Hurst パラメータ H を含んでおり、これは区間の長さの分布により求められる。

3. フラクタル時系列の予測手法

需要予測の方法は、現在までさまざまなものが提案されている。この詳細については省略するが、著者らが検討したなかで、方法の改善に役立つ経験としては次のようなものがある。まず、予測については線形回帰に基づく方法が簡単であり、理解しやすいと言える。しかし、この場合、時系列の変動が時間に依存する (時変システムである) ことにより、回帰係数を一定値にすることは予測精度を低下させる。また、大きな問題は需要そのものを観測しても正規乱数のような激しい変動をするので、そのまま予測の対象とす

ることは不適切であることがわかる。

以上のようなことを考慮すると、新しい需要予測の手法として、次のようなものが望ましいのではないかと結論に達した。まず、サンプルを集計した時系列を用い変動を緩和すること、線形予測を用いるが係数は時間的に変動するものであること、このようなことから、フラクタル時系列の自己相似性を用いる予測手法を考察の対象とすることとなった。

フラクタル時系列の特徴は、全体の図形を縮小したものが、図形の部分に類似していること (逆に部分図形を拡大すると全体に類似している) である。拡大縮小により完全に一致する図形は得られないが、概形は良く似たものになる。以下では、フラクタル理論を用いたフラクタル時系列の予測手法についてまとめておく [8][9]。

3.1 システムの同定

一般的な線形時変入出力システム

$$y(t) = \int_0^{t_0} h(t, t-\tau) x(\tau) d\tau, t > t_0 \quad (5)$$

を考察する。線形時変システムのインパルス応答関数 $h(t, \tau)$ がスケール関数 $\phi(t)$ を用いて次のように展開されると仮定する。

$$h(t, \tau) = \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} h_{ij} \phi_{N_i}(t) \phi_{N_j}(\tau) \quad (6)$$

ただし、

$$\phi_{N_i}(t) = \phi(2^{-N}t - i) \quad (7)$$

$$\phi(t) = \begin{cases} 1, & 0 < t \leq 1 \\ 0, & \text{otherwise} \end{cases} \quad (8)$$

特に、以下では入出力が同じ時系列である同定問題 ($y(t) = x(t)$ の場合) を考える。

いま、式(5)により同定された $y(t)$ と入力時系列 $x(t)$ との差の最小 2 乗近似を考え、これを最小化するようにインパルス応答関数の係数 h_{ij} を決定する。計算にあたり式(6)で $N=0$ とし、 j の範囲を $j=1 \sim 4$ に限定する。最適化の方法として最急降下法を用いる。

3.2 予測手法

次に、時間軸の伸長による予測について述べる。いま、 $T_s \leq t \leq T_e$ は時系列 $x(t)$ が観測される時間区間であり $T_1 = T_e - T_s$ とする。また、 $0 < t \leq T_2$ は $x(t)$ を予測する区間とする。なお、 T_1, T_2 の選び方には任意性があるが、期間 T_1 の時系列が、期間全体 T_2 の時系列と相似であるように選択する必要がある。

次の量を定義する。

$$b = a^D, a = T_2/T_1, T_2 > T_1. \quad (9)$$

D は時系列 $x(t)$ のフラクタル次元であり、 $1 < D < 2$ である。 $x(t)$ がフラクタル性を持つ場合には、その自己相似性により、 $0 < t \leq T_2$ において、次の式が近似的に成立する。

$$x(t) = b^{-1} \int_0^{bt_0} h\left(\frac{t}{b}, \frac{t-\tau}{b}\right) x(\tau) d\tau, t > bt_0 \quad (10)$$

スケール関数で近似された時系列のフラクタル性とは、インパルス応答関数により、以下のように説明される。インパルス応答関数 $h(t, \tau)$ の性質はスケール伸長されても時系列の自己相似性から、同じ線形予測の関係式を成立させる性質を保持するので、区間 T_1 で推定された $h(t, \tau)$ を b 倍したものをういれば、 a 倍に伸長した時間域においても、式(10)の予測モデルが成立する[2][7][8]。

このフラクタル性の説明は、コッホ曲線における相似形の出現を例としても説明できる。コッホ曲線を2次元グラフで描いた時にフラクタル図形の性質により図形は2次元的にひろがるので、横軸を a 倍に伸ばした空間には、もとの図形の b 個が入ることになる。

計算においては、 $a = \text{整数}$ となるように T_1, T_2 を選んでいる (例: $T_2 = 2T_1$)。サンプリングされた時系列を仮定し、サンプリング間隔を1とする。すなわち、式(5)では $t = t_0 + 1$ であり、式(10)では $t = bt_0 + b$ となるので、次の予測式となる。

$$y(bt_0 + b) = b^{-1} \int_0^{bt_0} h\left(\frac{t}{b}, \frac{t-\tau}{b}\right) x(\tau) d\tau, t > bt_0 \quad (11)$$

3.3 計算アルゴリズム

以下ではインパルス応答関数の計算や、予測の計算の手順を具体的に示す。式(6)を式(5)に代入する。

$$y(n+1) = \int_0^n \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} h_{ij} \phi(n-i) \phi(n-\tau-j) x(\tau) d\tau \quad (12)$$

ここで、 $t_0 = n$ とする。スケール関数は積分計算においてゼロとなる余分な項を除去するので、表現が簡単化でき

$$y(n+1) = \sum_{j=1}^B h_{nj} \theta(n-j) \quad (13)$$

となる。ここで $\theta(n-j)$ は $x(t)$ を積分することにより、次の式で与えられる。ここで B はインパルス応答の影響を制限するために、 $B=4$ にとっておく。

$$\theta(n-j) = \int_0^{n-j} \phi(n-\tau-j) x(\tau) d\tau \quad (14)$$

スケール関数の性質により

$$\theta(n-j) = \int_{n-j-1}^{n-j} x(\tau) d\tau \quad (15)$$

インパルス応答関数はモデル当てはめを行った結果である $y(t)$ と、もとのデータである $x(t)$ との差の2乗 $E_i = (y(t) - x(t))^2$ を最小にするように求めるので、式(13)より h_{ij} の修正値として次を得る。

$$\Delta h_{nj} = -2(x(n+1) - y(n+1)) \theta(n-j) \quad (16)$$

同様に式(11)に対応する表現として次を得る。

$$y(bn+b) = b^{-1} \sum_{j=1}^B h_{nj} \theta(n-j) \quad (17)$$

$$\theta(n-j) = \int_{nb-bj-b}^{nb-bj} x(\tau) d\tau \quad (18)$$

3.4 予測誤差

予測を議論する場合には、通常、現在までの観測データをもとにして、式(17)により1ステップ先のサンプル

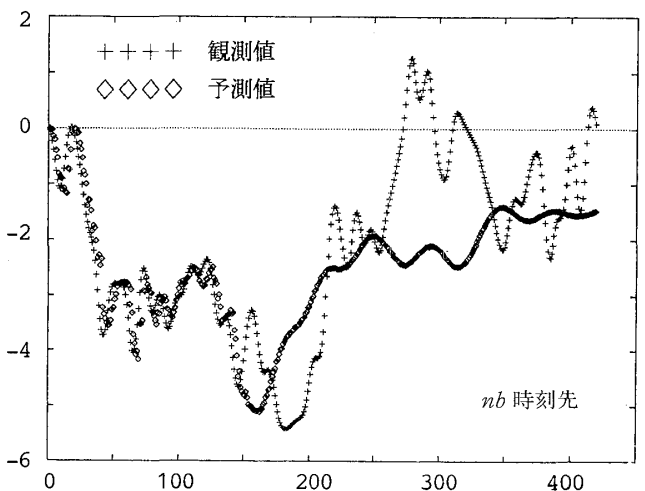
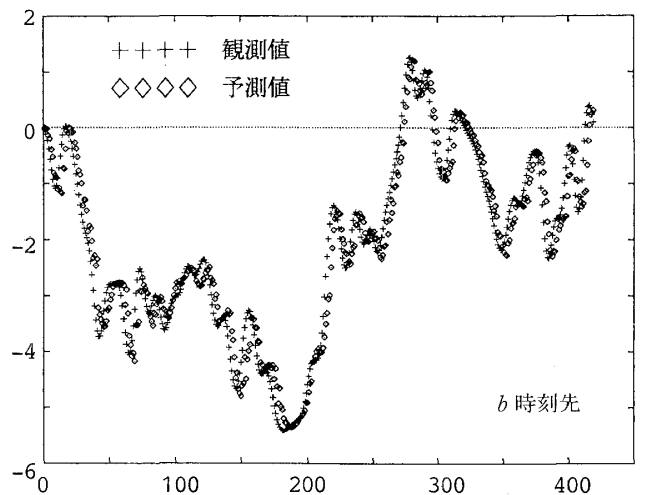


図1 fBm の予測 (b 時刻, nb 時刻先)

表1 fBmの**b**時刻先の予測誤差($a=2$)

| | 1000 サンプル | | |
|------|-----------|------|------|
| 次元 D | 1.80 | 1.50 | 1.25 |
| 予測誤差 | 0.25 | 0.19 | 0.17 |

表2 本手法の**nb**時刻先の予測誤差($D=1.5, a=2$)

| nb | 30 | 60 | 90 | 120 |
|-----------|-----|-----|-----|-----|
| 予測誤差 | 3.7 | 4.9 | 5.2 | 5.6 |

ルの予測値を計算することが行われる。これを、以下では**b**時刻先の予測とよぶ。

これに対して、式(17)に従って、逐次的に予測された値を観測値と見なして、予測を継続していく場合を考え、これにより n ステップ将来の値を予測する場合を nb 時刻先の予測とよんでおく。 nb 時刻先の予測においてもインパルス応答が計算されているので、線形予測をそのまま継続して用いる場合と比較して、予測誤差は極めて小さいものとなる。

図1にはフラクタル時系列 fBm の b 時刻先、 nb 時刻先の予測の様子を示している。

次に予測の平均誤差について述べる。まず、表1には $a=2$ の場合の b 時刻先の予測について、いくつかのフラクタル次元の時系列の平均予測誤差を示している(サンプル数は1000)。予測誤差の定義として、期間内の時系列の最大値と最小値の差(振幅)に対する予測誤差の割合(%)をとっている。これよりわかるように、0.3%程度の数値となっており、極めて小さな値である。

次に、表2に nb 時刻先の予測について予測誤差の平均値を示している。この結果が示すように nb 時刻先の予測についても誤差は6%程度におさまっている。

3.5 ウェーブレット係数によるフラクタル性の検証方法

本論文で示したフラクタル時系列の予測手法が適用可能であるためには、観測された時系列がフラクタルであることが必要である。通信トラヒックなどフラクタルであることを示す実証研究が存在する一方で、従来よりフラクタルの典型であるとされた株価時系列について、そのフラクタル性を疑問視する研究も存在する[12]。これらは、実証研究の難しさを示している。

以下では、時系列がフラクタルであるかどうかを検証する方法について述べるが、その中心は自己相似性とスペクトルが周波数のべき乗で低下する性質である。

本研究では、建材の日々の需要データについて、累計する処理を行うと、得られた時系列が近似的にフラクタルとなることを示す。

時系列がフラクタル(自己相似)であることは、通常、次の2つのことを確認することにより行われる[13]。

(m 個平均の分散)

もとの時系列を x_1, x_2, \dots としたとき、これらを m 個ずつ合計し平均をとった系列

$$x_i(m) = m^{-1} \sum_{j=(i-1)m+1}^{im} x_j \quad (19)$$

の分散である $var(x(m))$ が

$$var(x(m)) = \sigma m^{-\beta} \quad (20)$$

となること。ただし、 $0 < \beta < 1$ であり、 σ は $m=1$ の場合の分散である。

(1/f スペクトル)

時系列から計算したスペクトルが周波数のべき乗で低下する形状を示すこと。これは、いわゆる 1/f スペクトルと呼ばれる。

最初の分散の計算は容易に実行できる。後半のスペクトルの計算においては、フーリエ変換などの方法が可能であるが、ここでは解析精度がすぐれているウェーブレット変換による方法を用いる。

そのために、まず、時系列のウェーブレット変換の係数を計算しこれを用いて時系列がフラクタルであることを検証する方法について説明する[5][9][10]。ウェーブレット係数は相互に統計的にほぼ独立であり、高速フーリエ変換や R/S 解析(rescale range analysis)などの解析手法に比べて精度が良くなる。

時系列がフラクタル性をもつ場合には、そのスペクトルの時間平均が周波数のべき乗に反比例する性質を持つことが証明されている[10]。すなわち、時系列の分散、フラクタル次元を、それぞれ、 σ^2, D とするとスペクトルは式(21)のようになる。

$$S(\omega) = \sigma^2 / \omega^{-\gamma} \quad (21)$$

$$\gamma = 5 - 2D \quad (22)$$

与えられた時系列 $x(t)$ をウェーブレット変換する。

$$x(t) = \sum_n \sum_m x_n^m \phi_n^m(t). \quad (23)$$

$$x_n^m = \int_{-\infty}^{\infty} x(t) \phi_n^m(t) dt. \quad (24)$$

$\phi_n^m(t)$ はウェーブレット基本関数 $\phi(t)$ に対する次の

スケール, シフト変換とにより構成される.

$$\phi_n^m(t) = 2^{m/2} \phi(2^m t - n). \quad (25)$$

ここで, m, n は, スケール変換, シフト変換のインデックスである. $\phi(t)$ としては Daubechies のウェーブレットを用いる [11].

$x(t)$ 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.

$$\text{var}(x_n^m) = \sigma^2 2^{-m}. \quad (26)$$

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 [9][10].

式(26)の両変の対数をとると, m についての線形の直線となるので, 左辺により計算されるデータに対して回帰直線を当てはめ, この直線との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.

$$R_w = \left[\sum_m (\log(\text{var}(x_n^m)) - c_0 - c_1 m)^2 \right]^{1/2} / (MX_r) \quad (27)$$

ただし, c_0, c_1 は回帰直線を当てはめた場合の係数であり, M は添字 m の取り得る個数である. 具体的には, 時系列のサンプリング間隔を1とした場合に, データの補間を考えないケースでは, ウェーブレット関数 $\phi_n^m(t)$ の最小の台は1となるので, $m = -4, -3, -2, -1, 0$ などとなる. この場合は $M = 5$ である. また, X_r は $\log(\text{var}(x_n^m))$ の最大値と最小値との差であり, この値により2乗平方誤差を正規化している.

フラクタル時系列である条件であるスペクトル形状を検査する場合に, ウェーブレット係数の分散がウェーブレット係数のインデックスの回帰直線にどれだけ近いかを検証することにより実行できることを示した. しかし, これはあくまでも線形モデルでフラクタル性を検証したにすぎないので, 非線形モデルまで含めると, 更に別のモデルを当てはめることも可能である [14].

フラクタル時系列について, 式(27)の R_w を計算すると, 理論的にはゼロになる. しかし, 実際には計算誤差などが含まれるので, 有限の小さな値となる. コンピュータシミュレーションにより発生させた fBm について式(27)の R_w を計算すると, 平均して約 0.006 となる. 従って, この R_w の値がどの程度以下であれば

表3 ARMA モデルにおける R_w (時系列)

| | | | | | | |
|-------|-------|-------|-------|-------|-------|-------|
| p | 3 | 2 | 2 | 1 | 1 | 1 |
| q | 2 | 2 | 1 | 2 | 2 | 1 |
| R_w | 0.057 | 0.037 | 0.034 | 0.052 | 0.129 | 0.046 |

フラクタルと見なせるかを求める必要がある. ここでは, フラクタルとの対比で良く用いられ, 明らかに指数関数的に相関が減少する ARMA モデルについて R_w を計算し, このしきい値を推定する.

表3には, いくつかの次数の組合せ (AR 次数 p , MA 次数 q) をもつ ARMA モデルで発生させた時系列について平均的な R_w を計算した結果を示す. シミュレーションの詳細は省略するが, 表3に示されている平均値について分散は 0.002 程度であり極めて小さくなる.

以上のことから, 0.006 と 0.05 との中間の値を R_w のしきい値として採用し, この値より小さい値であればフラクタル性をもつと仮定する.

4. 建材需要の分析

4.1 需要の累計

以下では, 実際の建材の需要を集計したデータがフラクタルであるかどうかを検証してみる. すなわち, それぞれの日に発生する需要はランダムであるが, これを, 例えば, ある日を含めて8日前までを集計して時系列として表現すると, フラクタル性をもつことが示される. この理由としては, 日々の需要がランダムに (Gaussian Noise とみなす) 発生しており (ON/OFF ソースであると考え), これを数日間にわたり累計することにより多重化され, fBm となると解釈される.

ただし, 日曜祭日や, 取引が極めて少ない日については, データサンプルとしてそのまま残すと, 短期的な相関を発生しフラクタル性を無くす原因となるので, これらの日はデータから除外した.

建材の種類を 60 種類とり, それぞれについて次のようなデータ集計を行い, ウェーブレット係数から計算した R_w の平均値を求めた.

- (ケース1) 日々のデータ
- (ケース2) 4日単位の集計データ
- (ケース3) ある日を含めて8日前までを集計しその日の値とする
- (ケース4) ある日を含めて15日前までを集計しその日の値とする

表4 ケースごとの R_w の平均値

| | | | | |
|-------|-------|-------|-------|-------|
| ケース | 1 | 2 | 3 | 4 |
| R_w | 0.118 | 0.018 | 0.019 | 0.011 |

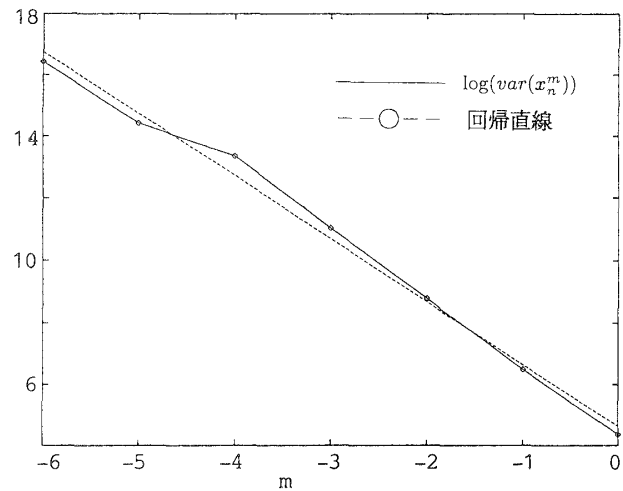
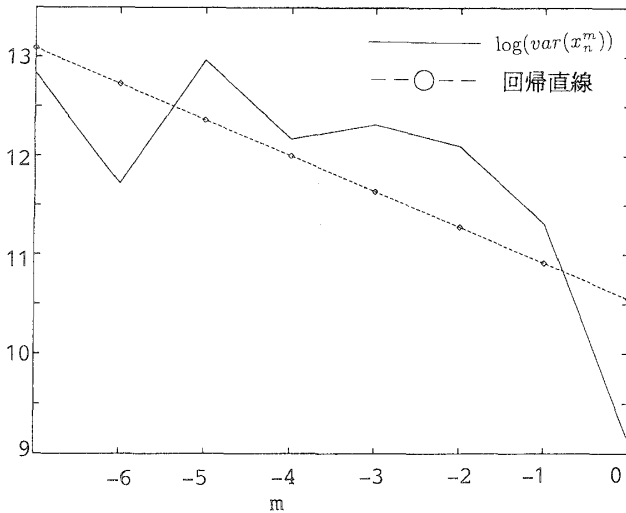
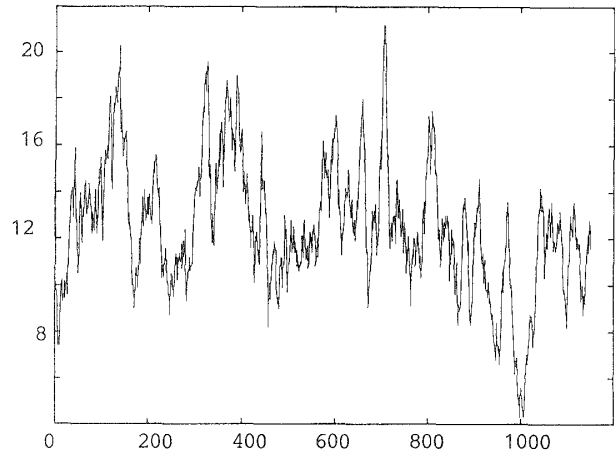
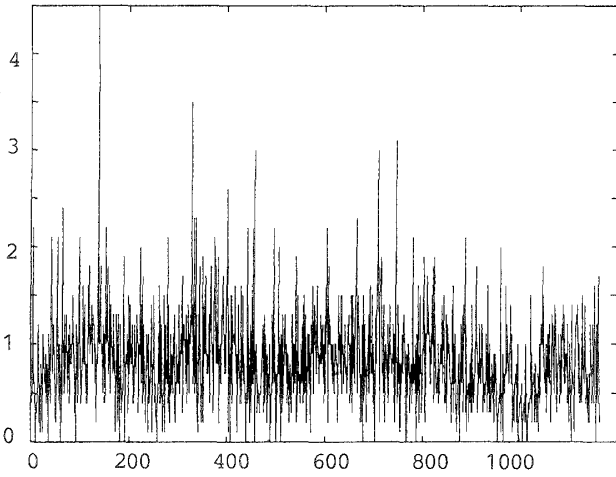


図2 ケース1の時系列の例(上)と $\log(\text{var}(x_n^m))$ の特性(下)

図3 ケース3の時系列の例(上)と $\log(\text{var}(x_n^m))$ の特性(下)

このようにして得られる R_w を求めたのが表4である。また図2, 3には, ケース1, 3の代表的な例について, 需要を示す時系列データと, m を横軸に $\log(\text{var}(x_n^m))$ を縦軸にとった場合の図と, これに当てはめた回帰直線を示している。

この表4よりわかるように, ケース1以外は時系列にフラクタル性が見られる。すなわち, 日々の需要値は Gaussian であるが, 4日単位の集計値や, 日々の需要を最低8, 15日間累計してその日の値とした場合には, フラクタルとみなすことができる。なお, 表4では平均値を示しているが, 多くの建材で, ケース2, 3, 4について顕著なフラクタル性を示していることがわかる (R_w の分散は0.013であり極めて小さい)。

なお, 表としては省略しているが, 式(20)で計算する β については, m にかかわらず一定値とならなければならないが, ケース2, 3, 4ではほぼこの条件が満足されている。ただし, 理想的なフラクタルである

fBm について式(20)の β を計算した場合と比較すると, m による変化が見られ, 現実に適用する場合には注意が必要である。フラクタル次元については説明を省略しているが, ケース2では約1.9, ケース3, 4では約1.7となっている。

4.2 日々の需要の性質

需要カーブを通信トラヒックにおける ON/OFF ソースの多重化とみなした場合には, 日々の需要が Gaussian で近似できることが前提となる。そこで, 以下では本研究でとりあげた例について, 日々の需要の性質を調べてみる。もちろん人工的なデータではないので, 近似的な視点からの分析となる。なお需要データによっては観測値そのものではなくその対数値をとった方が近似できる可能性が高くなるケースも存在する。

この検査には, QQプロットによる判断を用いている。表5にはQQプロットによる解析において傾向

表5 QQプロットにおける傾向線除去後の最大残差と平均

| | そのまま利用 | 対数値を利用 |
|-----|--------|--------|
| 最大値 | 0.81 | 0.80 |
| 平均値 | 0.19 | 0.18 |

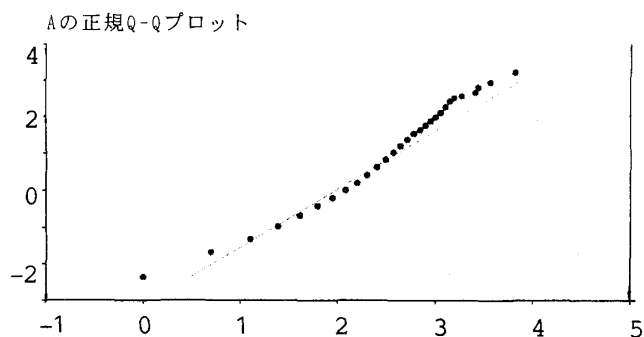


図4 日々の需要のQQプロットの例

線を除去したあとの残差の最大値と平均値を示している。需要データは60種類の建材であり、表5ではデータをそのまま解析した場合と、データの数値をとった場合の、それぞれの残差の最大値と平均値を示している。また、図4にはQQプロットの1つの例を示している。表5よりわかるように、QQプロットでは特に大きな偏移が見らず、日々の需要はほぼGaussianで近似できることがわかる。

なおデータの検査には χ^2 検定を利用することも考えられるが、一般にデータ数が大きくなると χ^2 値は大きくなり、検定の結果に影響を与えるので実施していない。

5. 需要予測への応用

5.1 需要予測の特性

次に、予測手法を実際の需要データに適用する場合について考察する。予測手法としては、節4.1で述べたケース3の場合の集計データに対する予測を用いる。需要データはフラクタル性をもつが、需要のデータ特有の問題点を持っている。以下にその概要を示す。

(1) トレンドの除去

需要カーブは一般にトレンドを含んでいる。多くの場合は単調な直線で近似される。本研究でとりあげている例では、ほとんどのケースで直線により傾向線を近似してさしつかえなかった。時系列の開始と終了の時点での値を考慮して傾向線を計算する方法をとった。

(2) 区間 T_1, T_2 の選択

理想的なフラクタル時系列と異なり、相互に縮小拡大の関係にある区間 (T_1, T_2) を見出すことが、やや複雑となる。予測をする場合には、 T_1 の区間でインパルス応答を計算し、これを T_2 における時系列に適用するがこの組合せが適切でない場合には予測誤差が大きくなる可能性がある。

この問題については、簡略化されたアルゴリズムにより、もとの時系列 $x(t)$ から最適な部分系列を抽出する方法を用いている。まず、 $x(t)$ を b 倍だけ軸にそって縮小しておき、更にこの振幅を1に正規化した時系列を $z(t)$ としておく。観測の開始時刻である T_s をゼロから一定の値だけ増加させながら、時系列 $x(t)$ の部分 ($T_s < t < T_e$) を求め、これを同様に振幅が1になるように正規化しておき、 $z(t)$ との差が最小となる T_s, T_e を求める。

(3) インパルス応答の計算時間

式(10)においてモデルを最適化するために、逐次近似計算によりインパルス応答を計算する必要がある。文献[8][9]においては、この計算を精度良く実行するために繰り返し計算の回数を20000回としている。この計算時間はかなりの負担となる。しかも、建材は多種類存在するので、それぞれの系列ごとに予測をする場合には、この計算時間が問題となる。

これについては、繰り返し計算の回数を減少させることにより解決した。具体的には、計算回数を500回程度まで減少させても、予測には大きな影響を与えないことが多くの数値例より確認できる。これにより大幅な計算時間の短縮がはかれる。

以上のことを考慮して、建材需要の予測について実施した結果を表6, 7にまとめている。図5には予測の例を示している。

これよりわかるように、建材需要はfBmのような理想的な挙動は見られないが、予測誤差も線形予測による結果よりはかなり良好なものとなっている。また、fBmの nb 時刻先の予測については現在の時刻からかなり離れた時点でも小さな予測誤差となっているが、建材需要の場合にはこれが約20時刻までは保証されることになっている。従って、現在から20日先までは予測が可能となる。現実には需要予測は15日をめどに見直しを行っているので、運用上では支障がないと考えられる。

なお、建材需要をサンプリングする場合、そのサンプル数はいくら必要であるかについては、次のようなことが経験的に言える。

表6 建材需要の b 時刻先の予測誤差 ($a=2$)

| | $T_1 = 200.$ | $T_1 = 300$ | $T_1 = 400$ |
|------|--------------|-------------|-------------|
| 予測誤差 | 1.21 | 1.39 | 1.17 |

表7 建材需要の nb 時刻先の予測誤差 ($a=2$)

| | $nb = 10$ | $nb = 20$ | $nb = 30$ | $nb = 40$ |
|------|-----------|-----------|-----------|-----------|
| 予測誤差 | 5.7 | 6.9 | 7.2 | 15.6 |

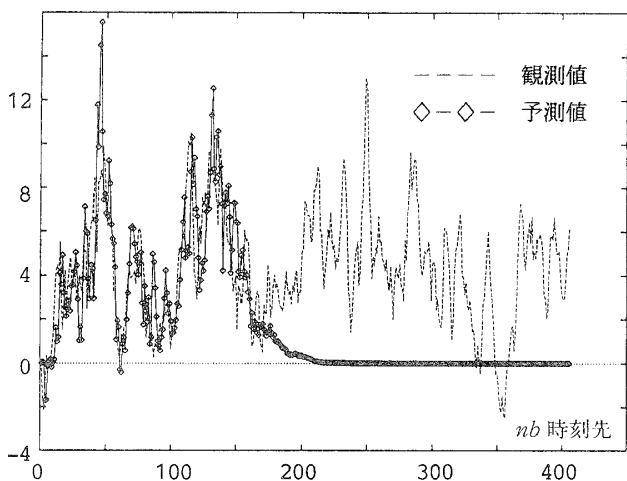
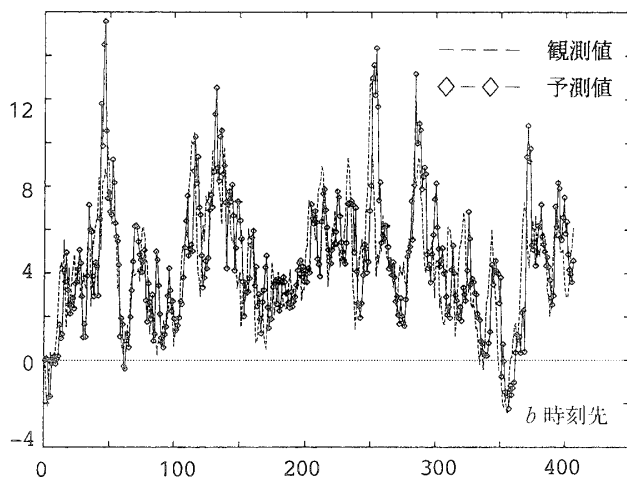


図5 建材需要時系列の b 時刻先, nb 時刻先予測の例

- (1) 日々データを集計する日数は最低でも4日であること。
- (2) 使用できるサンプル数が最低でも500個あること。
- (3) データより計算される時系列がフラクタルであることが必要であるが、サンプルの多くがゼロに近い値であったり、需要がない日が全体の1/8以上含まれている場合は方法を適用できない。

6. むすび

本研究では、フラクタル時系列の性質を用いた建材需要予測の一手法を展開するとともに、予測誤差、フラクタル次元の推定を示し、住宅建材の需要予測へと応用した。その結果建材の需要に固有の要因を考慮して適応的な修正を加えることにより、線形予測よりも良好な予測値を与えることを示した。

今後の課題としては、他の種類の商品の需要データへの適用と、時系列をデータベースとして構成し需要のパターンを検索する場合への利用などがあり、検討を行っていく予定である。

参考文献

- [1] 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OR学会春秋全国大会予稿集, pp. 1-2 (1997-10).
- [2] G. E. P. Box and G. M. Jenkins: *Time Series Analysis: Forecasting and Control*, rev. ed. San Francisco, Holden-Day (1976).
- [3] L. Ljung and T. Soederstroem: *Theory and Practice of Recursive Identification*, MIT Press (1983).
- [4] B. Mandelbrot: *The Fractal Geometry of Nature*, San Francisco, Freeman (1982).
- [5] 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. 42, No. 1, pp. 18-31 (1999).
- [6] M. S. Taqqu, W. Willinger, and R. Sherman: “Proof of a fundamental result in self-similar traffic modeling”, *Computer Communication Review (Proceeding of ACM SINCOMM '97)*, pp. 5-23 (1997).
- [7] W. Willinger, M. S. Taqqu, R. Sherman, and V. Wilson: “Self-similarity through high variability: Statistical analysis of Ethernet LAN traffic at the source level”, *IEEE/ACM Transaction on Networking*, Vol. 5, No. 1, pp. 71-86 (Feb. 1997).
- [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質を用いた予測手法とその応用”, 信学論(A), J79-A, 11, pp. 1793-1800 (1996-11).
- [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フラクタル時系列予測”, 信学論(A), J79-A, 12, pp. 1-9 (1996-12).
- [10] 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 (March 1992).
- [11] I. Daubechies: "Orthonormal bases of compactly-supported Wavelets", *Commun. Pure. Appl. Math.*, Vol. 41, pp. 909-996 (Nov. 1988).
- [12] 岸本一男: "フラクタルに関連した2つの実証研究", *数理科学*, No. 370, pp. 16-21 (1994).
- [13] 小沢利久: "いろいろな入力過程モデル (講座-待ち行列研究の新しい潮流(4))", *オペレーションズリサーチ*, Vol. 43, No. 12, pp. 680-686 (1998).
- [14] D. R. Cox, D. V. Hinkley, and O. E. Barndorff-Nielsen (eds): *Time Series Models: in econometrics, finance and other fields*, Chapman & Hall (1996).