

ウェーブレット分散と参照形式を利用した商品先物の暴落予想

上智大学大学院理工学研究科数学領域・院 本山 要

上智大学理工学部情報理工学科 加藤 剛

1 研究の目的

本研究は、ウェーブレット解析の中でも特にウェーブレット分散と呼ばれる概念を活用し、東京商品取引所 (TOCOM) における原油先物価格について、暴落の前兆現象をとらえることを目的としたものである。最終的には、暴落特有の特徴をもった参照図を作成し、新たなデータとの参照形式で暴落の注意喚起や警報に使用することを目標としている。

2 参照形式による暴落予想

参照形式とは、ある特徴を有する状況の情報を参照図として保存しておき、将来、参照図と似ている情報をもった状況が現れた場合、同じような状況が起こる可能性があるとして予測する方法である。具体的に、東京都の天気を予想する方法を例として挙げる。過去のデータから、東京都が晴れの日の前日の雲の位置、気温、湿度、気圧、降水量、風向き、風速、あるいは神奈川県や埼玉県についての天候に関わる情報を参照図として保存しておく。同じように東京都が曇、雨や雪の日の前日の情報を参照図として保存する。これらの参照図と現在の天候情報を比べ、「今日の天気は晴れの日の参照図に似ているから、明日東京は晴れるだろう」と予想する。

本研究は、このような参照形式を東京商品取引所 (TOCOM) が扱う原油先物価格の暴落予想に応用したものである。なお、値動きの状況を表す情報として様々なパラメータが存在するが、ウェーブレット分散の有効性を確かめるため、ウェーブレット分散の時間的推移のみを値動きの状況を表す情報としている。先物取引についての説明は付録 A、ウェーブレット分散については付録 B にそれぞれ記した。

3 データ解析

データ解析における計算の過程で誤った処理がされないように、参考文献 [3] と同様にして、TOCOM 提供のデータ 16 ヶ月分に対して前処理を行ったものが図 1 である。点線が入っている時刻は、後述する暴落時刻である。

3.1 参照図の作成

参照図を作成するにあたり、必要な用語を順次定義する。暴落の注意喚起や警報として利用するために、暴落を定義しなければならないが、暴落を理論的に定義することは困難である。本研究では、原油におけるサーキットブレーカー発動条件 2000 円の 75% である 1500 円を上回る下落がある場合を暴落とみなし、時間の間隔とあわせて次のように定義する。

定義 1 (暴落時刻)

次の不等式を満たす X_t が存在するとき、時刻 t を暴落時刻とする。

$$X_{t-(2^{16}+2^{14})} - X_t > 1500$$

定義 1 における時間 $2^{16} + 2^{14}$ は、TOCOM 原油先物取引立会時間内で数えると、4 時間 33 分 4 秒に相当する。

参考文献 [3] より、表 1 のように解析に関する設定を行った。はじめに幅を決定し、その後、時系列全体を隙間なく網羅できるように、幅の $1/4$ である 2^{14} でずらしを決定した。また、暴落の前兆をとらえやすい値を探索的に求め、着目スケールを決定した。前処理した 2014 年 9 月 1 日 0 時 0 分 0 秒から 2015 年 12 月 30 日 23 時 59 分 59

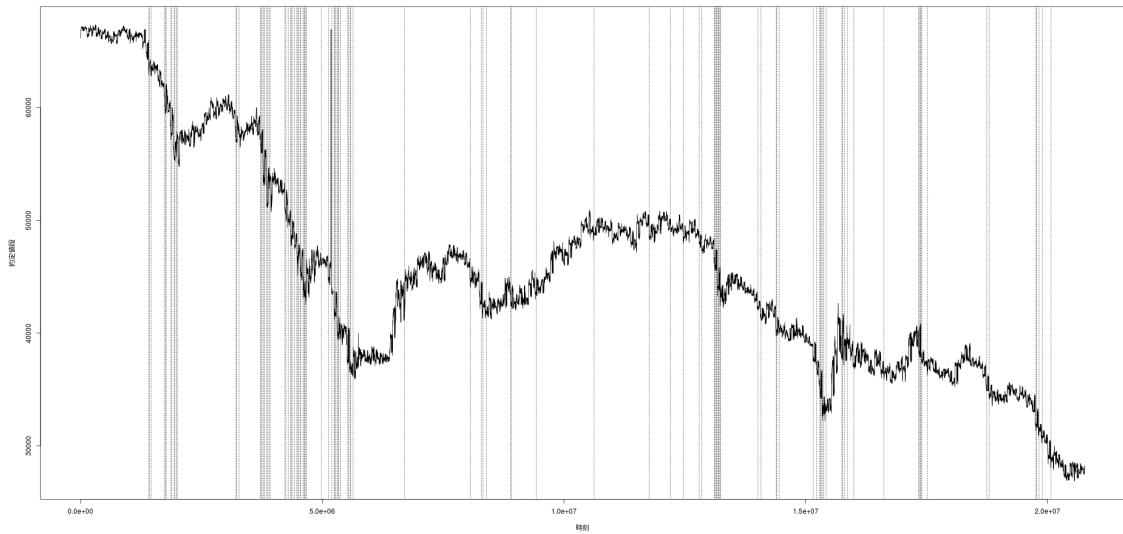


図1 2014年9月1日0時0分0秒から2015年12月30日23時59分59秒までの原油の約定値段

表1 解析に関する設定事項の整頓

幅 (MODWT における N のこと)	2 の 16 乗
ずらし	2 の 14 乗
着目スケール	$\tau_{11} = 2^{10}$

秒までの原油の約定値段のデータ数は 20, 768, 150 であったので、表 1 の設定事項から、1263 個のウェーブレット分散が得られた。

本研究の目的は、暴落の前兆現象をウェーブレット分散のパターンに照らし合わせてとらえることである。したがって、ウェーブレット分散の個数より多く暴落時刻を調べても、少ない回数だけ暴落時刻を調べても、情報に過不足が出てしまい、暴落時刻を特定する際に不都合である。よって、得られたウェーブレット分散の個数と同じ 1263 回暴落時刻を調べた。時刻におきかえると、時刻 $2^{16} + 2^{14}$ から 2^{14} おきに調べたことになる。その結果、1263 回中 111 回が暴落時刻に該当することがわかった。

ウェーブレット分散の 1 つの値だけで暴落の前兆をとらえることは困難である。そこで、2 つ以上のウェーブレット分散からなるベクトルをパターンとして定義する。

定義 2 (p 個の点をもつパターン)

$p > 1$ とし、隣り合う p 個のウェーブレット分散からなるベクトルを、 p 個の点をもつパターンと定める。

例 1 4 個の点をもつパターン A で 1263 個のウェーブレット分散を比較する場合、 $1263 - 4 + 1 = 1260$ 個の比較対象パターンが構成できる。

例 2 例 1 のウェーブレット分散からパターンを構成するイメージを図 2 に示す。図 2 で、垂線が下りている時刻を暴落が起きたか否かを調べた時刻 T とする。このとき、暴落時刻の定義にもとづく比較対象時刻は $T - (2^{16} + 2^{14}) \approx 2000000$ 付近の時刻である。また、時刻 T に対して、1~4 の範囲でそれぞれ計算された 4 つのウェーブレット分散の値のパターンが対応する。これを、ずらし 2^{14} でずらしていくことにより、1260 個のパターンが得られる。この場合、1 つのウェーブレット分散は、最大で 4 パターンに含まれる。

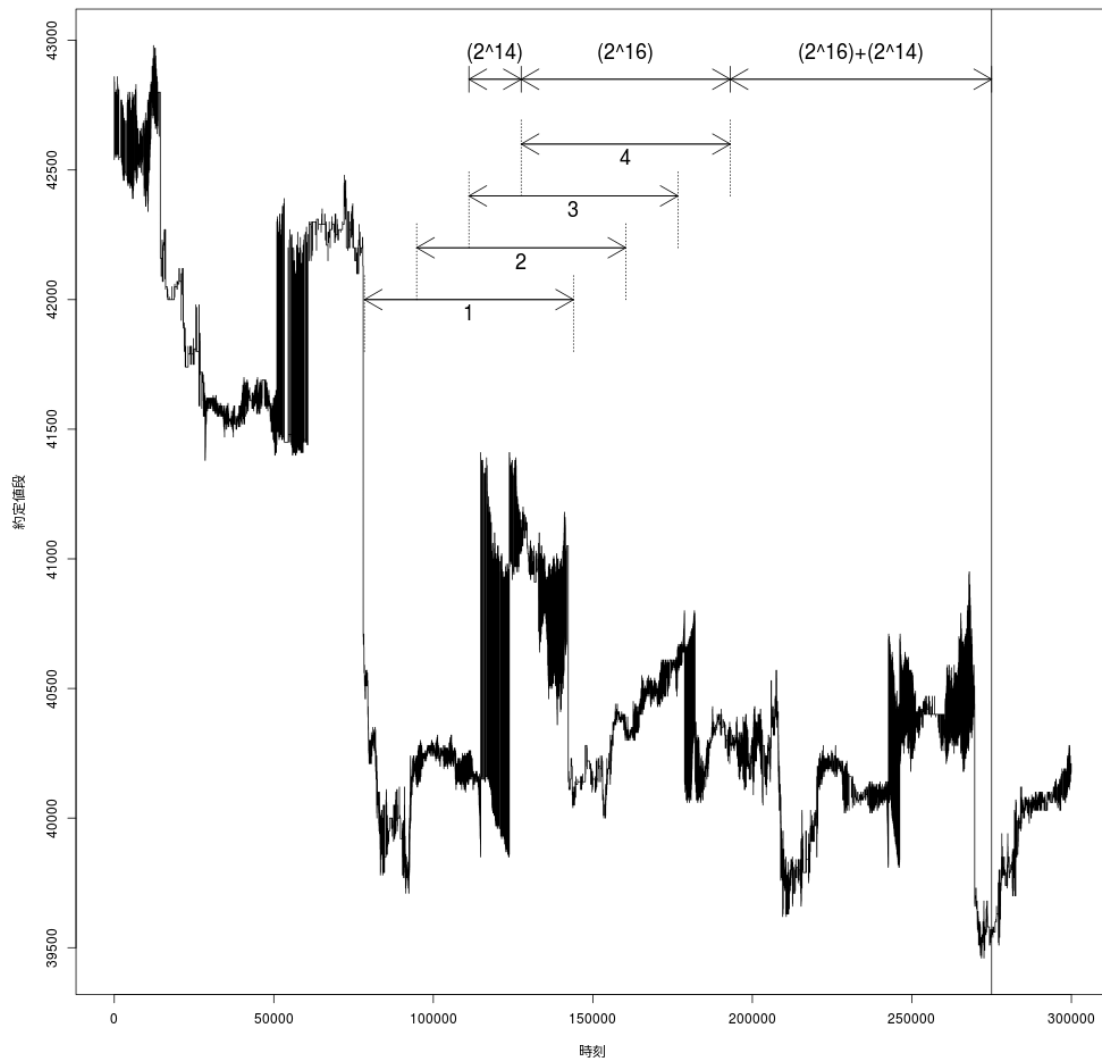


図2 暴落があるか調べた時刻とそれに対応するウェーブレット分散の対応図

1つのパターンで多くの暴落のパターンの特徴を備えることは困難と考えるのが自然である．そこで、複数のパターンで参照図を構成する．

定義 3 (m 個のパターンをもつ参照図)

m 個のパターンを要素とする集合を、 m 個のパターンをもつ参照図と定義する．

新たなデータから計算したウェーブレット分散にもとづくパターンと類似したものが参照図に存在するとき、暴落に対する注意喚起や警報を行う．パターンの類似度は、次の式で定義する距離 d で測る．

定義 4 p 個の点をもつパターン $A = \{a_1, \dots, a_p\}$, $B = \{b_1, \dots, b_p\}$ の距離 d を、

$$d = \sum_{i=1}^p (a_i - b_i)^2$$

で定める．

A と B の距離 d が短いほど類似度は高くなる。

類似度を測る際に実質的に同じ形だが、パターンを構成する値の平均と分散が大きく離れているせいで、類似度が低いと判断されてしまう。例えばパターン $\{1, 2\}$ と $\{100, 200\}$ は形は同じだが、構成する値の平均と分散が異なり、類似度が低く出てしまう。そこで、パターン同士の類似度を比較する前に、次のように各パターンに対し平均を 1、分散を 0 とする標準化の処理を施し、標準化後のパターンで類似度を測る。

定義 5 (パターンの標準化)

パターン $A = \{a_1, \dots, a_p\}$ の標準化とは、次の変換でパターン $B = \{b_1, \dots, b_p\}$ を得ることをいう。

$$b_i = (a_i - \bar{a})/\sigma_A$$

ここで、

$$\bar{a} = \frac{1}{p} \sum a_i, \quad \sigma_A = \sqrt{\sum (a_i - \bar{a})^2 / (p - 1)}$$

である。

3.2 参照図の作成

以上の議論を踏まえて、参照図の作成方法を説明する。

1. 使用するデータの期間

- (a) 2014 年 9 月 1 日 0 時 00 分 00 秒から 2014 年 12 月 30 日 23 時 59 分 59 秒の期間
- (b) 2014 年 9 月 1 日 0 時 00 分 00 秒から 2015 年 12 月 30 日 23 時 59 分 59 秒の期間

2. パターンを構成するウェーブレット分散の個数を決定

探索的に 4 個から 10 個までとした

3. 1 つの参照図を構成するパターン数を決定

4 個から 6 個までのいずれかとした

4. パターンを作成する方法を決定

(a) quantile based

暴落時刻に関するパターンをすべて選択し、選んだパターン数を N とする。互いのパターンの距離をすべて計算し、距離の値が低い順に並べ、1 から N の番号を付ける。そして、1 番, $1 + N/m$ 番, ... のように番号 1 から N/m ごとに番号を振られたパターン m 個を参照図とする。

(b) cluster based

暴落時刻に関するパターンをすべて選択し、選んだパターン数を N とする。 N 個のパターンにクラスタ分析を行い、 m 個のクラスタに分ける。各クラスタ内で対応する値の平均を計算して m 個のパターンを作り、参照図とする。

この手順に従えば、 $2 \times 7 \times 3 \times 2 = 84$ 種類の参照図が得られる。

3.3 結論

まず、得られた参照図の評価方法について説明する。評価は次の手順で行った。

1. パターンがもつ点の数 p を決定
2. p に応じて、全期間のウェーブレット分散 1263 個からパターンを構成
3. 参照図がもつパターンと、手順 2 で得られたパターンについてそれぞれ標準化を行う
4. 手順 3 で得られた参照図がもつパターンと全期間のパターンの距離を計算
5. p の値を変えて、手順 1~4 を実施

次に、評価項目について説明する。大まかに、次の2点で参照図を評価した。

- 信頼性
参照図がもつ、いずれかのパターンに近いパターンが見つかった時に、実際に暴落が起こる割合
- 性能
参照図がもつパターンが、全体の暴落の回数のうち、近いパターンとして検出できた割合

さらに、許容する類似度を探索的に変化させ、同様の評価を行った。

具体的な結果はシンポジウム当日に提示するが、結論として、信頼性と性能の両方の値が十分に高いといえる参照図はみつからなかった。しかし、信頼性は高く性能が低い参照図、逆に、信頼性は低い高性能参照図、さらに細かく見れば、信頼性が高いパターンを発見できた。したがって、これらを組み合わせることで、より実用的な参照図を作成できる可能性がある。

付録 A データ解析に必要な背景知識

参考文献 [4] および [5] から引用する形で、データの背景情報を記す。

A.1 商品先物の役割

本節における説明は、東京商品取引所のホームページに負う。商品先物取引とは、将来の一定期日に一定数の商品を契約時に取り決めた価格で買うことまたは売ることを約束する取引のことで、金融派生商品の1つである。将来における売買の値段を予め決めておくため、商品価格相場場の上下運動次第では利益を得ることもあれば損失を被ることもある。希望を表明すれば商品を授受することは可能であるが、契約時に決定した期日までに当初の取引と反対の取引を行うことで、差金決済だけで取引を終了することも可能である。

- リスクヘッジ
石油、貴金属やゴムなどの原材料品の価格は、受給バランスのみならず、政治・経済の情勢や気象条件などにより絶えず変動し、様々な企業の収益に大きな影響を与える。商品先物取引は、上記のリスク（物事の不確実性のこと）を軽減あるいは回避する手段を与える。たとえば、石油を継続的に買う立場の人間が石油の価格の高騰を予見できる場合、予見される高騰時の価格より低い価格での買いを約束する先物取引を結ぶことで、購入価格の高騰を抑えることができる。
- 公正な価格形成の場
商品先物市場には、現物を取り扱う企業の他に個人投資家や機関投資家など様々なプレーヤーが参加しているため、市場商品に関する多量の情報が集約される。故に、商品先物取引は公正で透明性の高い価格形成の場を提供する、市場の先物価格は公表され、信頼性の高い指標価格として応用される。
- 資産運用
商品先物取引価格は、総取引金額の5%~10%程度の証拠金を取引の担保として預け入れれば取引可能であり、差金決済によって取引を終了できる。ゆえに、商品先物取引は多くの投資家に資金運用の手段を与える。

実際の取引では、総取引額の5%~10%程度の現金を取引の担保として預けなければならない。これを、証拠金という。さらに、売買の過程で損失が一定額を超えた場合、取引を継続するためには追加の現金を預けなければならないというルールが設けられている。これを、追証拠金という。逆に言えば、希望する総取引金額と同等額の現金を取引開始時に用意する必要はないということである。ゆえに、オプション制度などと比較するとハイリスクかつハイリターンな金融派生商品といえる。

以上を踏まえ、商品先物取引の大きな特徴を以下に列挙しておく。

- 将来における売買を約束する取引である
- 商品価格の相場の変動によっては利益を得ることもあれば、損失を被ることもある
- 希望すれば商品の授受は可能であるが、反対売買による差金決済も可能である
- 総取引額 5%~10% 程度の証拠金を取引の担保として預ければ、取引に参加できる
- 売買の過程で損失が一定を超えた場合は、追証拠金を預けなければならない
- 現物を取り扱う当事者は商品先物取引を上手に使うことで価格変動リスクを抑えることができる

A.2 商品先物取引における専門用語の整理

商品先物取引における重要な専門用語をまとめておく、

- 約定価格
契約時に取り決められる価格のこと。約定値段ともいう。
- 限月
契約時に取り決められる期日のこと。正確には、契約時に取り決められる最終決済を行う月のことであり、各月における取引最終日と最終決済日は取引所によって決められている。最終取引日には、各月の 25 日や最終営業日などが採用されることが多い。なお、限月として選択することのできる月は商品ごとに取引所によって決められている。TOCOM では、金の場合は当月含めて 12 ヶ月以内の各偶数月、とうもろこしの場合は当月含めて 12 ヶ月以内の各奇数月、原油の場合は当月含めて 6 ヶ月以内の各月と決められている。
- 立会時間
取引所が決められている、取引を行うことのできる時間のこと。TOCOM では、前日の夜間立ち合いと当日の日中立ち合いの 2 つを合体させて 1 つの計算区域を定義している。故に、日中立ち合いには計算区域をもとにした日付と実日付は一致する。夜間立会中では、日付が変わるまでは計算区域をもとにした日付は実日付と比べて早まることになり、日付が変わった後も実日付が土日や祝日で取引所の非営業日である場合には、計算区域をもとにした日付と実日付にズレが生じる。

TOCOM における原油に関する種々の情報は、表 1 の形にまとめられる、

表 2 東京商品取引所の原油先物取引に関する情報

限月	新甫発会日の属する月の翌月から起算した 6 月以内の各月（6 限月制）
立会時間	日中：9:00～15:15，夜間：16:30～翌日 4:00
取引最終日	当限月の最終営業日（日中立会まで）
最終決済日	当限月の翌月第一営業日

付録 B ウェーブレット分散

参照図を作るもとになるウェーブレット分散について、参考文献 [3] の 22 章と [5] を引用する形で説明する。

B.1 ウェーブレット分散の特徴

ウェーブレット分散とは、時系列の分散を分解することで時系列に対する分散分析を可能にするものである。フーリエ変換にもとづくスペクトル密度も、これに含まれる。ウェーブレット分散による分散分析は、多くの点においてスペクトル密度による分散分析と類似している。ところが、実際に使用する者の立場から見ると両者には重要な違いがある。スペクトル密度は、連続的なフーリエ周波数の分散の分解である。分解によって得られる各々の成分は、特定の周波数を持つ正弦曲線に時系列がどの程度似ているかを反映するものである。その一方で、ウェーブレッ

ト分散は、スケールの離散集合の分解を与える。ここで、大まかな言い方をすれば、スケールとは時系列の平均値を取る時間の区間のことである。ウェーブレット分散による分解の各成分の長所は、ある特定のスケールにおける隣接した平均値の間にどれくらいの相違があるかを測れることにある。スケールという概念は、周期 (周波数の逆数) の概念とは別のものである。スケールも周期も同じ単位で測れるものであるが、後者は平均をとるという考えを伴わない。

B.2 最大重複離散ウェーブレット変換

ウェーブレット分散は、最大重複離散ウェーブレット変換 (maximal overlap discrete wavelet transform, MODWT) にもとづいている。そこで、参考文献 [2] を引用する形で、MODWT の定義と性質から議論を始める。MODWT はドゥベシウェーブレットフィルタ $\{\tilde{h}_{1,l}, l = 0, 1, \dots, L_1 - 1\}$ である。ここで、 $\tilde{h}_{1,0} \neq 0$, $\tilde{h}_{1,L_1-1} \neq 0$ である。したがって、ドゥベシフィルタは L_1 の幅を持つ。今後の議論のために $l < 0$ または $l \geq L_1$ に対して $\tilde{h}_{1,l} = 0$ と定義する。定義より、ドゥベシウェーブレットフィルタは次の3つの性質を満たさなければならない。

$$\sum_{l \in \mathbb{Z}} \tilde{h}_{1,l} = 0, \quad \sum_{l \in \mathbb{Z}} \tilde{h}_{1,l}^2 = 1, \quad \sum_{l \in \mathbb{Z}} \tilde{h}_{1,l} \tilde{h}_{1,l+2n} = 0, \quad n = \pm 1, \pm 2, \dots,$$

ここで \mathbb{Z} は整数全体を表す。もっとも単純なドゥベシフィルタは、 $L_1 = 2$ 、フィルタ係数を $\tilde{h}_{1,0} = 1/2$, $\tilde{h}_{1,1} = -1/2$ としたもので、ハールウェーブレットフィルタとよばれる。一般に、ウェーブレットフィルタは次で定義されるようなスケーリングフィルタとして知られる補完フィルタをもつ。

$$\tilde{g}_{1,l} = (-1)^{l+1} \tilde{h}_{1,L_1-l-1}, \quad l \in \mathbb{Z}$$

ハールの場合、 $\tilde{g}_{1,0} = 1/2$, $\tilde{g}_{1,1} = 1/2$, $\tilde{g}_{1,l} = 0$, ($l \neq 0, 1$) である。以下、ハールの場合を考えることとする。

$\{X_t, t = 0, 1, \dots, N-1\}$ を等間隔に観測された N 個の時系列データとする。すなわち、時刻 t_0 に観測されたデータが X_0 、サンプリング間隔を Δ とすれば、 X_t は $t_0 \Delta$ に観測されたデータとする。このとき、時系列データ $\{X_t\}$ とウェーブレットフィルタ、スケーリングフィルタからそれぞれ、単位 MODWT ウェーブレット係数 $\{\tilde{W}_{1,t}\}$ とそれに対応するスケーリング係数 $\{\tilde{V}_{1,t}\}$ を得る。

$$\tilde{W}_{1,0} = \frac{X_0 - X_{N-1}}{2}, \quad \tilde{W}_{1,t} = \frac{X_t - X_{t-1}}{2} \quad (t = 1, 2, \dots, N-1)$$

$$\tilde{V}_{1,0} = \frac{X_0 + X_{N-1}}{2}, \quad \tilde{V}_{1,t} = \frac{X_t + X_{t-1}}{2} \quad (t = 1, 2, \dots, N-1)$$

$t = 0$ の場合を除いて、スケーリング係数は時系列データの隣り合うふたつの観測値の平均値である。反対に、ウェーブレット係数は隣り合うふたつの観測値の差に比例する。

単位水準の場合、ウェーブレット係数はスケール $\tau_1 = 1$ での平均の差に相当し、スケーリング係数はスケール $\lambda_1 = 2$ での $\{X_t\}$ の隣り合うふたつの観測値の平均を計算することに相当する。より高水準 $\tau_j = 2^{j-1}$, $\lambda_j = 2^j$ の MODWT 係数も単位水準と同じように解釈することができる。高水準のウェーブレットフィルタ $\{\tilde{h}_{j,l}, l = 0, 1, \dots, L_j - 1\}$ とスケーリングフィルタ $\{\tilde{g}_{j,l}, l = 0, 1, \dots, L_j - 1\}$, $L_j \equiv (2^j - 1)(L_1 - 1) + 1$ を用いて、水準 $j > 1$ の係数を定義する。

$$\tilde{W}_{j,t} = \frac{1}{2^j} \left(\sum_{l=0}^{2^{j-1}-1} X_{t-l} - \sum_{l=0}^{2^{j-1}-1} X_{t-l-2^{j-1}} \right), \quad \tilde{V}_{j,t} = \frac{1}{2^j} \sum_{t=0}^{2^j-1} X_{t-l} \quad (t = 2^{j-1}, 2^{j-1} + 1, \dots, N-1)$$

スケーリング係数はスケール $\lambda_j = 2^j$ での平均、一方、ウェーブレット係数はスケール $\tau_j = 2^{j-1}$ での隣接した平均の差に比例する。なお、実際には MODWT ウェーブレット係数とスケーリング係数は上のように計算されるのではなく、ピラミッドアルゴリズムとして知られている方法で計算する。

B.3 ウェーブレット分散の推定方法

時系列, ウェーブレット係数, スケーリング係数をそれぞれ要素にもつ 3 つの N 次元列ベクトルを考える.

$$\mathbf{X} = {}^t(X_0, X_1, \dots, X_{N-1})$$

$$\widetilde{\mathbf{W}}_1 = {}^t(W_{1,0}, W_{1,1}, \dots, W_{1,N-1})$$

$$\widetilde{\mathbf{V}}_1 = {}^t(V_{1,0}, V_{1,1}, \dots, V_{1,N-1})$$

また, X の N 次元ユークリッドノルムの 2 乗を以下のように記す.

$$\|\mathbf{X}\|^2 \equiv \sum_{t=0}^{N-1} X_t^2$$

この量は, X がもつ”エネルギー”に相当する. MODWT の重要な点は

$$\|\mathbf{X}\|^2 = \|\widetilde{\mathbf{W}}_1\|^2 + \|\widetilde{\mathbf{V}}_1\|^2$$

が成り立つという意味で, エネルギーが保存することである. \mathbf{X} の標本平均を $\bar{X} = \sum_{t=0}^{N-1} X_t/N$ であらわすとき, その時系列の標本分散は以下のように表せる.

$$\begin{aligned} \hat{\sigma}_X^2 &\equiv \frac{1}{N} \sum_{t=0}^{N-1} (X_t - \bar{X})^2 = \frac{1}{N} \|\mathbf{X}\|^2 - \bar{X}^2 = \frac{1}{N} \|\widetilde{\mathbf{W}}_1\|^2 + \left(\frac{1}{N} \|\widetilde{\mathbf{V}}_1\|^2 - \bar{X}^2 \right) \\ &\equiv \hat{\sigma}_{\widetilde{\mathbf{W}}_1}^2 + \hat{\sigma}_{\widetilde{\mathbf{V}}_1}^2 \end{aligned}$$

ここで $\hat{\sigma}_{\widetilde{\mathbf{W}}_1}^2$ と $\hat{\sigma}_{\widetilde{\mathbf{V}}_1}^2$ は, それぞれ $\widetilde{\mathbf{W}}_1$ 及び $\widetilde{\mathbf{V}}_1$ の標本分散にとることができる (ウェーブレットフィルタの定義から, X_t が期待値 (母平均) を持つ場合, 緩い条件のもとで $\widetilde{W}_{1,t}$ の期待値は 0 とできる. 一方, ハールウェーブレットフィルタの場合に容易に直接確かめられるように, \widetilde{V}_1 の標本平均は常に \bar{X} に等しい). 以上から, 時系列 \mathbf{X} の標本平均は 2 つの構成要素に分解できることが分かった. 1 つの構成要素である $\hat{\sigma}_{\widetilde{\mathbf{W}}_1}^2$ は標準スケール $\tau_1 = 1$ における時系列の変動であって, もう 1 つの $\widetilde{\mathbf{V}}_1$ はスケール $\lambda_1 = 2$ における \mathbf{X} の平均を表している. あるいは $\widetilde{\mathbf{W}}_1$ と $\widetilde{\mathbf{V}}_1$ はそれぞれ $\hat{\sigma}_X^2$ の高周波数変動と $\hat{\sigma}_X^2$ の低周波数変動をとらえていると考えることもできる.

上述の考え方を最大水準 $J_0 \geq 1$ にまで適用することで, ウェーブレット分散による分散分析を定義できる. はじめに, 水準 $J_0 = 2$ の分散分析を考える. 具体的には, $\|\widetilde{\mathbf{V}}_1\|^2$ を 2 つの値の和, すなわち $\|\widetilde{\mathbf{W}}_2\|^2$ と $\|\widetilde{\mathbf{V}}_2\|^2$ の和に置き換える. $\|\widetilde{\mathbf{W}}_2\|^2$ はスケール $\tau_2 = 2$ の重み付き局所平均の隣り合うもの同士の変化に関連する量であり, $\|\widetilde{\mathbf{V}}_2\|^2$ はスケール $\lambda_2 = 4$ における重み付き平均に関する量である.

再帰的に $\|\widetilde{\mathbf{V}}_{j-1}\|^2$ を $\|\widetilde{\mathbf{W}}_j\|^2 + \|\widetilde{\mathbf{V}}_j\|^2$ に置き換えることで, J_0 での分解は以下のように求まる.

$$\|\mathbf{X}\|^2 = \sum_{j=1}^{J_0} \|\widetilde{\mathbf{W}}_j\|^2 + \|\widetilde{\mathbf{V}}_{J_0}\|^2 \quad \text{および} \quad \hat{\sigma}_X^2 = \sum_{j=1}^{J_0} \hat{\sigma}_{\widetilde{\mathbf{W}}_j}^2 + \hat{\sigma}_{\widetilde{\mathbf{V}}_{J_0}}^2$$

ただし,

$$\hat{\sigma}_{\widetilde{\mathbf{W}}_j}^2 \equiv \frac{1}{N} \|\widetilde{\mathbf{W}}_j\|^2, \quad \hat{\sigma}_{\widetilde{\mathbf{V}}_{J_0}}^2 \equiv \frac{1}{N} \|\widetilde{\mathbf{V}}_{J_0}\|^2 - \bar{X}^2$$

である.

$\hat{\sigma}_{\widetilde{\mathbf{W}}_j}^2$ を水準 j の経験ウェーブレット分散といい, $\hat{\sigma}_{\widetilde{\mathbf{W}}_{J_0}}^2$ を水準 J_0 の経験スケーリング分散とよぶことにする. $\hat{\sigma}_{\widetilde{\mathbf{W}}_j}^2$ はスケール $\tau_j = 2^{j-1}$ の重み付き局所平均の隣り合うもの同士の変化に関連する量であり, $\hat{\sigma}_{\widetilde{\mathbf{W}}_{J_0}}^2$ はスケール $\lambda_{J_0} = 2^{J_0}$ における重み付き平均に関連する量である.

B.4 ウェーブレット分散の推定方法

水準 J_0 での ANOVA は $j = 0$ から J_0 までの経験ウェーブレット分散と $j = J_0$ での経験スケーリング分散を求めることによって達成できる。しかし、経験ウェーブレット分散と経験スケール分散の値は、境界係数の値に大きな影響を受ける。特に、非定常性の強い信号を扱う場合は、時系列の最初の方と最後の値に大きな違いが生じていることが予想できる。この境界係数がもたらす影響は、対象の時系列のもつ特性とは全く関係のないものである。

ゆえに、ウェーブレット分散の推定を以下の形で行うことにする。

$$\hat{v}_{W_j}^2 \equiv \frac{1}{M_j} \sum_{t=2^j-1}^{N-1} \widetilde{W}_{j,t}^2$$

ただし、 $M_j = N - (2^j - 1)$ である。この改良は、参考文献 [3] にもとづいている。改良された推定量からは境界係数は削除されており、削除された境界係数の個数だけ全体を割る数が減らされている。

参考文献

- [1] Brockwell, P. J. and Davis, R. A., "Time Series: Theory and Methods", *Springer*, (1991).
- [2] Nason, G. P., "Wavelet Methods in Statistics with R", *Springer*, (2008).
- [3] T.S. ラオ 他 編 / 北川源四郎 他 監訳, "時系列分析ハンドブック", 朝倉書店 (2016)
- [4] 東京商品取引所ホームページ, <http://www.tocom.or.jp/jp/>
- [5] 若林光太, "ウェーブレット分散を応用した原油先物のリスクに関する考察", 上智大学大学院理工学研究科修士論文 (2015)