

スペクトル解析，ウェーブレット解析

時系列分析は，自己相関などの時間的な情報にもとづいて時間領域で行われるのが普通であるが，本章で説明するウェーブレット解析とスペクトル解析は，そのような時間領域における分析を補完，あるいは代替するものである．スペクトル解析は，時系列データをフーリエ変換により周波数領域に変換し，周波数成分の強さなどの情報を抽出する．他方，ウェーブレット解析は，時間と周波数という2つの領域における情報を同時に取り込もうとする方法である．

以下では，手法が開発された歴史的順序に従い，まず，スペクトル解析について述べ，次に，スペクトル解析と比較しつつ，ウェーブレット解析について説明する．

21.1 スペクトル解析

1次元の離散的確率過程 $\{y_t\}$ が定常であるとする．すなわち，期待値は一定で，自己共分散は時間差だけに依存するものと仮定する．このとき，平均と自己共分散を次のように表す．

$$E(y_t) = \mu, \quad \text{Cov}(y_t, y_{t+h}) = \gamma(h) = \gamma(-h)$$

平均 μ や時差 h の自己共分散 $\gamma(h)$ は，母集団のパラメータであり，一般に未知である．時間領域における時系列分析では，とくに，自己共分散，あるいは自己相関に関する情報が重要であり，その推定が分析の出発点となる．

他方，本節のメイン・テーマであるスペクトル解析では，自己共分散をフーリエ変換することにより，舞台を周波数領域に移して情報の抽出に努める．こうすることにより，時間領域ではみえにくかった事実を浮かび上がらせることが可能となる．

21.1.1 スペクトル密度関数

スペクトル解析においては，自己共分散を係数とするフーリエ級数を考える．

$$f(\lambda) = \frac{1}{2\pi} \sum_{h=-\infty}^{\infty} \gamma(h) e^{-ih\lambda}, \quad (-\pi \leq \lambda \leq \pi) \quad (21.1)$$

ここで，右辺の無限級数は収束するものと仮定する．そのための十分条件は，

$$S = \sum_{h=-\infty}^{\infty} |\gamma(h)| < \infty \quad (21.2)$$

となることである。(18.1) 式の $f(\lambda)$ を $\{y_t\}$ のスペクトル密度関数 (spectral density function), あるいは, スペクトラム (spectrum) とよぶ。スペクトラムは, 原点对称, 周期 2π の周期関数である。したがって, $f(\lambda)$ の挙動は $[0, \pi]$ で考えれば十分である。このとき, λ は周波数の意味合いをもち, $2\pi/\lambda$ は周期となる。周期は, 周波数が π のときに最小値 2 をとり, 周波数 0 で無限大となる。また, スペクトラムは, つねに非負である。このことは, 定義からは自明でないが,

$$f_T(\lambda) = \frac{1}{2\pi T} E \left(\left| \sum_{t=1}^T (y_t - \mu) e^{-i\lambda t} \right|^2 \right) = \frac{1}{2\pi T} \sum_{h=-(T-1)}^{T-1} (T - |h|) \gamma(h) e^{-ih\lambda}$$

において, $f_T(\lambda)$ は非負で, $T \rightarrow \infty$ のとき, $f(\lambda)$ に収束することによる。

逆に, スペクトラムが与えられたときに, (21.1) 式の両辺に $e^{ih\lambda}$ をかけて, $[-\pi, \pi]$ の範囲で λ に関して定積分することにより, 自己共分散をスペクトラムのフーリエ変換

$$\gamma(h) = \int_{-\pi}^{\pi} e^{ih\lambda} f(\lambda) d\lambda \quad (21.3)$$

から求めることができる。特に, 分散はスペクトラムを区間 $[-\pi, \pi]$ で積分したものに等しくなる。この意味で, スペクトラム $f(\lambda)$ は, 周波数 λ における分散 (の定数倍) であるという直感的な解釈が成り立つ。そして, $f(\lambda)$ の値が大きければ大きいほど, 時系列に含まれる周波数 λ , あるいは周期 $2\pi/\lambda$ に対応する変動の程度が強いことを意味する。

例題 21.1 分散一定の無相関過程, すなわち, $\gamma(0) = \sigma^2$, $\gamma(h) = 0$ ($h \neq 0$) の場合, スペクトラムは, (21.1) 式の定義から, $f(\lambda) = \sigma^2 / (2\pi)$ と定数になる。このことは, 各周波数が同一の変動をもたらすことを意味しており, それは白色光の波長としての性質と同じである。このことから, 無相関過程は白色雑音 (white noise) とよばれる。

例題 21.2 AR (1) 過程 $y_t = \phi y_{t-1} + \varepsilon_t$, $|\phi| < 1$, $\{\varepsilon_t\} \sim \text{i.i.d. } (0, \sigma^2)$ の場合には, $\gamma(h) = \sigma^2 \phi^{|h|} / (1 - \phi^2)$ であるから,

$$\begin{aligned} f(\lambda) &= \frac{\sigma^2}{2\pi(1-\phi^2)} \left(1 + \sum_{h=1}^{\infty} \phi^h (e^{-ih\lambda} + e^{ih\lambda}) \right) \\ &= \frac{\sigma^2}{2\pi(1-\phi^2)} \left(1 + \frac{\phi e^{-i\lambda}}{1 - \phi e^{-i\lambda}} + \frac{\phi e^{i\lambda}}{1 - \phi e^{i\lambda}} \right) \\ &= \frac{\sigma^2}{2\pi} \frac{1}{|1 - \phi e^{i\lambda}|^2} = \frac{\sigma^2}{2\pi} \frac{1}{|\phi(e^{i\lambda})|^2} \end{aligned}$$

となる。ただし, $\phi(x) = 1 - \phi x$ である。図 21.1 には, $\sigma^2 = 1$ の場合の 2 つの AR (1) 過程のスペクトラムが描かれている。実線は $\phi = 0.6$, 点線は $\phi = -0.6$ の場合のスペクトラムである。前者は低周波 (= 長周期) 成分の影響が強く, 逆に, 後者は高周波 (= 短周期) 成分の影響が強いことがわかる。両者の周波数成分の相違は, 前者の自己相関が正の値のまま減衰する (= 長周期が支配的) のに対して, 後者の自己相関は正負交互に振動しながら減衰する (= 短周期が支配的) ことによる。

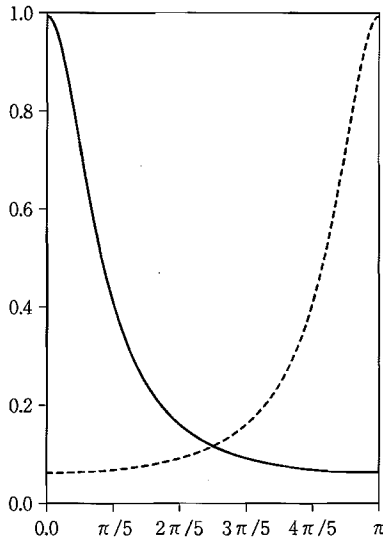


図 21.1 AR (1) のスペクトラム

21.1.2 スペクトル分布関数

定常過程に従う時系列でも、必ずしも、スペクトラムが存在するとは限らない。たとえば、次の時系列を考えてみよう。

$$x_t = A \cos \theta t + B \sin \theta t \quad (21.4)$$

ここで、 A と B は、ともに平均0、分散 σ^2 で、たがいに無相関な確率変数、 θ は $(0, \pi)$ に属する正の定数である。このとき、自己共分散は、

$$\gamma(h) = \text{Cov}(x_t, x_{t+h}) = \sigma^2 \cos \theta h$$

となり、 $h \rightarrow \pm\infty$ としても減少することなく振動的である。このことから、スペクトラムは定義されないことがわかる。

しかし、このような場合でも、スペクトル分布関数 (spectral distribution function) はつねに存在する。スペクトル分布関数 $F(\lambda)$ とは、 $[-\pi, \pi]$ 上で定義され、単調非減少、右連続、原点に関して対称な増分をもつ非負の関数で、 $F(-\pi) = 0$ 、 $F(\pi) = \gamma(0)$ となるものである。

スペクトラムと自己共分散が1対1対応すると同様に、スペクトル分布関数も自己共分散と一意的に対応する。具体的には、次の定理が成り立つ (Anderson, 1971)。

【定理 21.1】 定常過程の自己共分散 $\gamma(h)$ は、スペクトル分布関数 $F(\lambda)$ を使って、次のように表すことができる。

$$\gamma(h) = \int_{-\pi}^{\pi} e^{ih\lambda} dF(\lambda) \quad (21.5)$$

さらに、 $F(\lambda)$ が微分可能な場合はスペクトラム $f(\lambda)$ が存在し、導関数 $F'(\lambda)$ が

スペクトラム $f(\lambda)$ となる。このとき, (21.5) 式の関係は, (21.3) 式で表現される。□

上の定理で使われる積分はスティルチェス積分であり, いまの場合は, 単調非減少な有界変動関数 $F(\lambda)$ に関する積分である。

例題 21.3 (21.4) 式で定義された時系列 $\{x_t\}$ のスペクトル分布関数は,

$$F(\lambda) = \begin{cases} 0 & (\lambda < -\theta) \\ \sigma^2/2 & (-\theta \leq \lambda < \theta) \\ \sigma^2 & (\theta \leq \lambda) \end{cases}$$

となる。すなわち, スペクトル分布関数は階段関数であり, ジャンプが起きるのは, 時系列 $\{x_t\}$ に含まれる周波数成分 θ , および $-\theta$ においてである。このとき, 実際, 次のことが成り立つ。

$$\int_{-\pi}^{\pi} e^{ih\lambda} dF(\lambda) = \frac{\sigma^2}{2} (e^{-ih\theta} + e^{ih\theta}) = \sigma^2 \cos h\theta = \gamma(h)$$

21.1.3 線形過程のスペクトラム

(21.1) 式に従ってスペクトラムを計算することは, 一般に面倒である。ここでは, 具体的なモデルが与えられた場合には, 自己共分散を使うことなくスペクトラムが求められることを示そう。そのために, 次の線形過程 (linear process) を考えよう。

$$y_t = \sum_{j=0}^{\infty} \alpha_j \varepsilon_{t-j}, \quad \alpha_0 = 1, \quad \sum_{j=0}^{\infty} \alpha_j^2 < \infty, \quad \{\varepsilon_t\} \sim \text{i.i.d.}(0, \sigma^2) \quad (21.6)$$

まず, $\{\varepsilon_t\}$ のスペクトラムは, $f_\varepsilon(\lambda) = \sigma^2/(2\pi)$ である。このとき, $\{y_t\}$ の自己共分散を $\gamma_y(h)$ とすれば, 定理 21.1 を使って, 次の関係が得られる。

$$\begin{aligned} \gamma_y(h) &= \text{Cov}(y_t, y_{t+h}) = E\left(\sum_{j=0}^{\infty} \alpha_j \varepsilon_{t-j} \sum_{k=0}^{\infty} \alpha_k \varepsilon_{t+h-k}\right) \\ &= \sum_{j=0}^{\infty} \sum_{k=0}^{\infty} \alpha_j \alpha_k E(\varepsilon_{t-j} \varepsilon_{t+h-k}) = \sum_{j=0}^{\infty} \sum_{k=0}^{\infty} \alpha_j \alpha_k \int_{-\pi}^{\pi} e^{i(h+j-k)\lambda} f_\varepsilon(\lambda) d\lambda \\ &= \int_{-\pi}^{\pi} e^{ih\lambda} \left| \sum_{j=0}^{\infty} \alpha_j e^{ij\lambda} \right|^2 f_\varepsilon(\lambda) d\lambda \end{aligned}$$

定理 21.1 を再度使うことにより, $\{y_t\}$ のスペクトラムが, 次の定理のように得られる。

【定理 21.2】 (21.6) 式の線形過程 $\{y_t\}$ のスペクトラム $f_y(\lambda)$ は,

$$f_y(\lambda) = \left| \sum_{j=0}^{\infty} \alpha_j e^{ij\lambda} \right|^2 f_\varepsilon(\lambda) = \frac{\sigma^2}{2\pi} \left| \sum_{j=0}^{\infty} \alpha_j e^{ij\lambda} \right|^2 = \frac{\sigma^2}{2\pi} |\alpha(e^{i\lambda})|^2 \quad (21.7)$$

で与えられる。ここで, $\alpha(x) = \sum_{j=0}^{\infty} \alpha_j x^j$ である。

定理 21.2 が示唆することは, L をラグ・オペレータ, $\beta(L)$ をラグ多項式として, 2つの定常過程 $\{x_t\}$ と $\{y_t\}$ が, $y_t = \beta(L)x_t$ の関係にあるならば, $\{y_t\}$ のスペクトラムは, $\{x_t\}$ のスペクトラムに $|\beta(e^{i\lambda})|^2$ を乗じたものになる, ということである。工学では, $\beta(e^{i\lambda})$ を周波数応答関数 (frequency response function), $|\beta(e^{i\lambda})|^2$

をパワー伝達関数 (power transfer function) という。したがって、2つのスペクトラムは、パワー伝達関数で結び付けられる。

以上のことを使えば、ARMA (p, q) 過程のスペクトラムを求めることは簡単である。実際、次の定理に従えばよい。

[定理 21.3] 定常な ARMA (p, q) 過程 $\phi(L)y_t = \theta(L)\varepsilon_t$ のスペクトラム $f_y(\lambda)$ は、

$$f_y(\lambda) = \frac{\sigma^2}{2\pi} \frac{|\theta(e^{i\lambda})|^2}{|\phi(e^{i\lambda})|^2} \quad (21.8)$$

で与えられる。ただし、

$$\phi(L) = 1 - \phi_1 L - \dots - \phi_p L^p, \quad \theta(L) = 1 - \theta_1 L - \dots - \theta_q L^q \quad \square$$

例題 21.4 次の2つの AR (2) モデルを考えよう。

$$x_t = 1.5x_{t-1} - 0.75x_{t-2} + \varepsilon_t, \quad y_t = 0.8y_{t-1} - 0.64y_{t-2} + \xi_t$$

ここで、 $\{\varepsilon_t\} \sim \text{i.i.d.}(0, \sigma_x^2)$, $\{\xi_t\} \sim \text{i.i.d.}(0, \sigma_y^2)$ である。定理 18.3 から、それぞれのスペクトラムは

$$f_x(\lambda) = \frac{\sigma_x^2}{2\pi} \frac{1}{|1 - 1.5e^{i\lambda} + 0.75e^{2i\lambda}|^2}, \quad f_y(\lambda) = \frac{\sigma_y^2}{2\pi} \frac{1}{|1 - 0.8e^{i\lambda} + 0.64e^{2i\lambda}|^2}$$

となる。図 21.2 には、 $\sigma_x^2 = 1$, $\sigma_y^2 = 6$ の場合に、これら2つのスペクトラムが図示されている。実線は $f_x(\lambda)$ 、点線は $f_y(\lambda)$ である。 $f_x(\lambda)$ のピークは $\pi/6$ 、 $f_y(\lambda)$ のピークは $\pi/3$ であり、それぞれ、周期 12, 6 に対応する。これらのピークは、2つのモデルの特性方程式の根を極形式 $re^{i\omega}$ で表したときの周波数 ω となる。

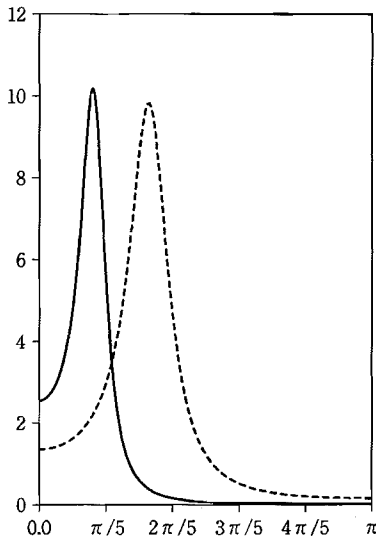


図 21.2 AR (2) のスペクトラム

時系列データを分析する場合は, 階差変換などをしてから分析することが多い. このような変換の意味をスペクトラムの観点から解釈することができる. たとえば, 1期前との階差 $x_t - x_{t-1} = (1-L)x_t$ では, フィルター $1-L$ が使われるが, そのパワー伝達関数は,

$$\psi_1(\lambda) = |1 - e^{i\lambda}|^2 = 2(1 - \cos \lambda)$$

となり, 低周波数 (長周期) の値が小さく, 高周波数 (短周期) の値が大きくなるような関数である. したがって, 階差変換後のデータは, 高周波数成分は保持しつつ, 低周波数成分の影響が除去されるものと考えられる. このような役割をするフィルターをハイパス・フィルター (high-pass filter) とよぶ. 逆に, 集計演算, たとえば, $x_t + x_{t-1}$ のフィルター $1+L$ のパワー伝達関数は,

$$\psi_2(\lambda) = |1 + e^{i\lambda}|^2 = 2(1 + \cos \lambda)$$

となり, 前とは逆に, 低周波数成分は保持しつつ, 高周波数の影響を除去するものである. このようなフィルターは, ローパス・フィルター (low-pass filter) とよばれる.

なお, 季節性を含むようなデータでは, 1年前との階差 (四半期データでは4期前, 月次データでは12期前) をとる場合が多い. たとえば, 月次データの階差変換 $x_t - x_{t-12}$ では, フィルター $1-L^{12}$ のパワー伝達関数は,

$$\psi_3(\lambda) = |1 - e^{i2\pi\lambda}|^2 = 2(1 - \cos 12\lambda)$$

であり, 季節周波数 $2\pi j/12$ ($j=1, 2, \dots, 6$) における値が0となる. したがって, 季節階差は季節変動を除去する働きをもつことがわかる.

21.1.4 スペクトラムの推定

スペクトラムは, 一般に未知であるので, データから推定する必要がある. もっともナイーブな推定量は, 次のように構成できる. まず, 時系列データ y_1, \dots, y_T を与えられて, 標本自己共分散

$$\hat{\gamma}(h) = \frac{1}{T} \sum_{t=1}^{T-h} (y_t - \bar{y})(y_{t+h} - \bar{y}), \quad (h=0, 1, \dots, T-1)$$

を計算する. そして, スペクトラムの定義 (21.1) の中の $\gamma(h)$ を $\hat{\gamma}(h)$ でおきかえることにより, ピリオドグラム (periodogram)

$$I(\lambda) = \frac{1}{2\pi} \sum_{h=-(T-1)}^{T-1} \hat{\gamma}(h) e^{-ih\lambda} = \frac{1}{2\pi} |w(\lambda)|^2 \quad (21.9)$$

を得る. ここで, $w(\lambda)$ は, データ y_1, \dots, y_T の有限フーリエ変換

$$w(\lambda) = \frac{1}{\sqrt{T}} \sum_{t=1}^T (y_t - \bar{y}) e^{it\lambda}$$

である. ピリオドグラムは, 標本自己共分散の有限フーリエ変換であるから, 逆に, 標本自己共分散は,

$$\hat{\gamma}(h) = \int_{-\pi}^{\pi} e^{ih\lambda} I(\lambda) d\lambda$$

と表される。この関係は、スペクトラムと母集団の自己共分散との関係を示した (21.3) 式の標本バージョンである。

ピリオドグラムは、スペクトラムの推定量として、次の性質をもつ。なお、 $I(0) \equiv 0$ であるので、原点は除外して考える。

【定理 21.4】 1次元の時系列 $\{y_t\}$ が定常、反転可能な ARMA (p, q) モデル

$$\phi(L)(y_t - \mu) = \theta(L)\varepsilon_t, \quad \{\varepsilon_t\} \sim \text{i.i.d.}(0, \sigma^2)$$

に従い、誤差項 $\{\varepsilon_t\}$ が4次のモーメントをもつとする。このとき、(21.9) 式で定義されたピリオドグラムは、次の性質をもつ。

$$\begin{aligned} \lim_{T \rightarrow \infty} T[E(I(\lambda)) - f(\lambda)] &= -\frac{1}{2\pi} \sum_{h=-\infty}^{\infty} |h| \gamma(h) e^{-ih\lambda}, \quad (\lambda \neq 0) \\ \lim_{T \rightarrow \infty} V(I(\lambda)) &= \begin{cases} f^2(\lambda) & (\lambda \neq 0, \pm\pi \text{ のとき}) \\ 2f^2(\lambda) & (\lambda = \pm\pi \text{ のとき}) \end{cases} \\ \lim_{T \rightarrow \infty} \text{Cov}(I(\omega), I(\lambda)) &= 0, \quad (\omega \neq \pm\lambda) \end{aligned}$$

上の定理が伝えるメッセージは、ピリオドグラムが、① スペクトラムの漸近的不偏推定量である。② しかし、一致性はない。③ 異なる周波数間では、漸近的に無相関である、ということである。

ピリオドグラムは、スペクトラムの推定量としては一致性をもたないので、それ自体では有用でない。しかし、漸近的な不偏性と、異なる周波数間での漸近的な無相関性があるので、当該の周波数まわりでの平均をとることにより、一致性をもたらしうな推定量を得ることができる。すなわち、推定量は、 λ のまわりの $2m+1$ 個の基本周波数 $\lambda_j = 2\pi j/T$ におけるピリオドグラムの平均として、

$$\bar{I}(\lambda) = \frac{1}{2m+1} \sum_{j \in A_{2m+1}(\lambda)} I(\lambda_j) \quad (21.10)$$

を考えることになる。ここで、 $A_{2m+1}(\lambda)$ は、周波数 λ に近接する $2m+1$ 個の基本周波数からなる集合である。項数 $2m+1$ が標本サイズ T とともに大きくなるならば、ただし、 $m/T \rightarrow 0$ となるような発散のスピードならば、次のことが成り立つ (Fuller, 1996)。

$$\lim_{T \rightarrow \infty} (2m+1) V(\bar{I}(\lambda)) = \begin{cases} f^2(\lambda) & (\lambda \neq 0, \pm\pi) \\ 2f^2(\lambda) & (\lambda = 0, \pm\pi) \end{cases}$$

もっと一般に、標本自己共分散にもとづくスペクトラムの一致推定量のクラスは、

$$\hat{f}(\lambda) = \sum_{h=-K_T}^{K_T} w_h \hat{\gamma}(h) e^{-ih\lambda}$$

と表される。ここで、 $\{w_h\}$ は、 $w_h = w_{-h}$ となるようなウェイトである。また、 K_T は、 $T \rightarrow \infty$ のとき、 $K_T \rightarrow \infty$ 、 $K_T/T \rightarrow 0$ となるような自然数である。ウェイトを $w_h = 1 - (|h|/K_T)$ としたものは Bartlett 推定量、 $w_h = 1 - (h^2/K_T^2)$ としたものは Parzen 推定量として知られている (Anderson, 1971)。

21.1.5 周波数領域における分析の利点

周波数領域で分析することの1つの利点について述べるために, 再度, データ y_1, \dots, y_T の有限フーリエ変換を取り上げる. ここでは, 基本周波数における変換

$$w(\lambda_j) = \frac{1}{\sqrt{T}} \sum_{t=1}^T (y_t - \bar{y}) e^{it\lambda_j}, \quad \lambda_j = \frac{2\pi}{T}([-T/2] + j), \quad (j=1, \dots, T)$$

を考える. なお, この場合のフーリエ変換は, $\lambda_j \neq 0$ ならば平均修正をしなくても同一の結果をもたらす. なぜなら, $e^{iT\lambda_j} = 1$ であるから, 次のことが成り立つことによる.

$$\sum_{t=1}^T e^{it\lambda_j} = \frac{e^{i\lambda_j}(1 - e^{iT\lambda_j})}{1 - e^{i\lambda_j}} = \frac{e^{i\lambda_j}(1 - e^{2\pi i j})}{1 - e^{i\lambda_j}} = 0, \quad (\lambda_j \neq 0)$$

上のフーリエ変換を簡潔に表現するために, 次の T 次元ベクトルと $T \times T$ 行列を定義しよう.

$$\mathbf{w} = \begin{pmatrix} w(\lambda_1) \\ \vdots \\ w(\lambda_T) \end{pmatrix}, \quad \mathbf{y} = \begin{pmatrix} y_1 \\ \vdots \\ y_T \end{pmatrix}, \quad \mathbf{e} = \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}, \quad \mathbf{P} = \begin{pmatrix} P_1 \\ P_2 \\ \vdots \\ P_T \end{pmatrix}$$

ここで, $P_j = (e^{-i\lambda_j}, \dots, e^{-iT\lambda_j})/\sqrt{T}$ である. このとき, 上のフーリエ変換は, 次のように簡潔に表現することができる.

$$\mathbf{w} = \mathbf{P}(\mathbf{y} - \bar{y}\mathbf{e})$$

行列 \mathbf{P} の共役転置行列を \mathbf{P}^* と表せば, $\mathbf{P}^*\mathbf{P} = \mathbf{I}_T$ (\mathbf{I}_T は, T 次の単位行列) が成り立つから, \mathbf{P} はユニタリー行列である. そして, \mathbf{y} の共分散行列を Σ とすれば, Σ は \mathbf{P} により, ほぼ対角化されて, 次の重要な近似式が成り立つ (Fuller, 1996).

$$\mathbf{P}\Sigma\mathbf{P}^* \cong 2\pi\mathbf{D}_T, \quad \mathbf{D}_T = \text{diag}\{f(\lambda_1), \dots, f(\lambda_T)\} \quad (21.11)$$

ここで, $\text{diag}\{a_1, \dots, a_n\}$ は, 対角要素が a_1, \dots, a_n であるような対角行列を表す.

上の事実を使えば, 定常な ARMA モデルの推定を周波数領域で行えば, 時間領域の場合よりも一般に容易となる. 実際, 観測値ベクトル \mathbf{y} に正規性を仮定して, $\mathbf{y} \sim N(\mu\mathbf{e}, \Sigma)$ とすれば, 時間領域における対数尤度関数は, 次のように変形することができる.

$$\begin{aligned} L(\boldsymbol{\theta}) &= -\frac{T}{2} \log(2\pi) - \frac{1}{2} \log |\Sigma| - \frac{1}{2} (\mathbf{y} - \mu\mathbf{e})' \Sigma^{-1} (\mathbf{y} - \mu\mathbf{e}) \\ &\approx -\frac{T}{2} \log(2\pi) - \frac{1}{2} \log |\mathbf{P}\Sigma\mathbf{P}^*| - \frac{1}{2} (\mathbf{P}(\mathbf{y} - \bar{y}\mathbf{e}))^* (\mathbf{P}\Sigma\mathbf{P}^*)^{-1} (\mathbf{P}(\mathbf{y} - \bar{y}\mathbf{e})) \\ &\approx -\frac{T}{2} \log(2\pi) - \frac{1}{2} \log |2\pi\mathbf{D}_T| - \frac{1}{4\pi} \mathbf{w}^* \mathbf{D}_T^{-1} \mathbf{w} \\ &= -\frac{T}{2} \log(2\pi) - \frac{1}{2} \sum_{j=1}^T \log(2\pi f(\lambda_j)) - \frac{1}{2} \sum_{j=1}^T \frac{I(\lambda_j)}{f(\lambda_j)} \end{aligned}$$

ここで, $\boldsymbol{\theta}$ は, ARMA モデルに含まれる未知パラメータのベクトルである. この最

後の表現では、行列式や逆行列の計算が回避されている。これは、近似にもとづく尤度（擬似尤度）であるが、時間領域における厳密な尤度よりも計算しやすいことが了解されよう。この点が有限フーリエ変換の効用であり、周波数領域で考えることの利点となっている。

21.1.6 多変量時系列のスペクトラム

いままでの議論を多変量の場合に拡張しよう。時系列 $\{y_t\}$ を n 次元の離散的定常確率過程として、平均ベクトルと自己共分散行列を、

$$E(y_t) = \mu, \quad \text{Cov}(y_t, y_{t+h}) = \Gamma(h) = \Gamma'(-h) \quad (21.12)$$

により定義する。ラグ h の自己共分散行列 $\Gamma(h)$ は、 $h=0$ のとき分散行列となり対称であるが、 $h \neq 0$ のときは、一般には対称でない。

このような多変量時系列 $\{y_t\}$ に対して、スペクトラムを

$$f(\lambda) = \frac{1}{2\pi} \sum_{h=-\infty}^{\infty} \Gamma(h) e^{-ih\lambda} \quad (21.13)$$

で定義する。ただし、右辺の無限和が収束するものと仮定する。この場合のスペクトラムは、 $n \times n$ の行列関数であり、エルミート行列となる。すなわち、 $f(\lambda) = f^*(\lambda)$ が成り立つ。このことから、 $f(\lambda)$ を実部と虚部に分けると、

$$f(\lambda) = c(\lambda) + iq(\lambda) = f^*(\lambda) = c'(\lambda) - iq'(\lambda)$$

となるので、 $c(\lambda)$ は対称、 $q(\lambda)$ は歪対称となる。前者をコスpectrum (cospectrum)、後者をカドラチャー・スペクトラム (quadrature spectrum) とよぶ。また、 $f(\lambda)$ の (j, k) 要素 $f_{jk}(\lambda)$ は、 j 番目と k 番目の時系列の相互スペクトラム (cross spectrum) とよばれ、一般に複素数値となる。

相互スペクトラム $f_{jk}(\lambda)$ を、

$$f_{jk}(\lambda) = c_{jk}(\lambda) + iq_{jk}(\lambda) = \alpha_{jk}(\lambda) e^{i\phi_{jk}(\lambda)}$$

と表したとき、

$$\alpha_{jk}(\lambda) = \sqrt{c_{jk}^2(\lambda) + q_{jk}^2(\lambda)} = |f_{jk}(\lambda)|$$

は、相互振幅スペクトラム (cross-amplitude spectrum) とよばれ、

$$\phi_{jk}(\lambda) = \tan^{-1}(q_{jk}(\lambda)/c_{jk}(\lambda))$$

は、位相スペクトラム (phase spectrum) とよばれる。また、

$$\rho_{jk}(\lambda) = \frac{c_{jk}^2(\lambda) + q_{jk}^2(\lambda)}{f_{jj}(\lambda) f_{kk}(\lambda)}$$

は、コヒーレンス (coherency) とよばれ、周波数 λ における 2 つの時系列の相関の強さを表す。

一般の n 次元線形過程

$$y_t = \sum_{j=0}^{\infty} A_j \varepsilon_{t-j}, \quad \sum_{j=0}^{\infty} \|A_j\|^2 < \infty, \quad \{\varepsilon_t\} \sim \text{i.i.d.}(\mathbf{0}, \Sigma) \quad (21.14)$$

のスペクトラムを考えよう。ここで、 $\|A\|$ は、行列 A のノルムであり、 $A'A$ の最大

固有値の正の平方根で定義される。スカラーの場合と同様に, $\{y_t\}$ のスペクトラム $f_y(\lambda)$ が与えられれば, 自己共分散は,

$$\Gamma y(h) = \int_{-\pi}^{\pi} e^{ih\lambda} f_y(\lambda) d\lambda \quad (21.15)$$

から求めることができる。他方,

$$\begin{aligned} \Gamma y(h) &= \sum_{j=0}^{\infty} \sum_{k=0}^{\infty} A_j E(\varepsilon_{t-j} \varepsilon'_{t+h-k}) A'_k = \sum_{j=0}^{\infty} \sum_{k=0}^{\infty} A_j \int_{-\pi}^{\pi} f_{\varepsilon}(\lambda) e^{i(j+h-k)\lambda} d\lambda A'_k \\ &= \int_{-\pi}^{\pi} e^{ih\lambda} A(e^{i\lambda}) f_{\varepsilon}(\lambda) A'(e^{-i\lambda}) d\lambda \end{aligned} \quad (21.16)$$

と表すことができる。ただし,

$$A(e^{i\lambda}) = \sum_{j=0}^{\infty} A_j e^{ij\lambda}$$

である。したがって, (21.15) と (21.16) 式から,

$$f_y(\lambda) = A(e^{i\lambda}) f_{\varepsilon}(\lambda) A'(e^{-i\lambda}) = \frac{1}{2\pi} A(e^{i\lambda}) \Sigma A'(e^{-i\lambda}) \quad (21.17)$$

を得る。同様にして, 定常な n 次元 ARMA (p, q) モデル

$$\Phi(L)(y_t - \mu) = \Theta(L)\varepsilon_t$$

のスペクトラムは, (21.17) 式の表現で, $A(z)$ を $\Phi^{-1}(z)\Theta(z)$ に代えればよい。

21.1.7 多変量スペクトラムの応用

線形過程 (21.14) の誤差項 $\{\varepsilon_t\}$ の分散行列 Σ が, 対角行列 $\Sigma = \text{diag}(\sigma_{11}, \dots, \sigma_{nn})$ であるとする。このとき, 第 j 番目の時系列のスペクトラムは,

$$f_{jj}(\lambda) = \frac{1}{2\pi} \sum_{k=1}^n |a_{jk}(e^{i\lambda})|^2 \sigma_{kk} \quad (21.18)$$

で与えられる。ここで, $a_{jk}(e^{i\lambda})$ は, $A(e^{i\lambda})$ の (j, k) 要素である。スペクトラム $f_{jj}(\lambda)$ は, 第 j 系列の周波数 λ における分散の意味合いをもっている。その分散が, (21.18) 式のように, n 個の系列の分散に分解されるから, 次の量

$$RPC_{k \rightarrow j}(\lambda) = \frac{|a_{jk}(e^{i\lambda})|^2 \sigma_{kk}}{\sum_{k=1}^n |a_{jk}(e^{i\lambda})|^2 \sigma_{kk}}$$

は, j 番目の系列の変動に占める k 番目の系列の影響の割合を表すと考えられる。RPC は, 相対的パワー寄与率 (relative power contribution) とよばれ, マクロ経済変数間の因果関係の分析に使われている (山本, 1988)。

21.1.8 長期記憶過程のスペクトラム

定常な ARMA 過程のスペクトラムは, $[-\pi, \pi]$ で連続であるが, 不連続なスペクトラムをもつような確率過程として, 次のモデルを考えよう。

$$(1-L)^d y_t = u_t = \frac{\theta(L)}{\phi(L)} \varepsilon_t, \quad \{\varepsilon_t\} \sim \text{i.i.d.}(0, \sigma^2) \quad (21.19)$$

ここで、 $u_t = \phi^{-1}(L)\theta(L)\varepsilon_t$ は定常な ARMA (p, q) に従う。また、 d は、 $-1/2 < d < 1/2$ となる実数である。以上の条件のもとで、 $\{y_t\}$ は反転可能な定常過程となり、 $y_t = (1-L)^{-d}\phi^{-1}(L)\theta(L)\varepsilon_t$ は、(18.6) のような線形過程表現をもつ (Hosking, 1981)。モデル (21.19) を ARFIMA (p, d, q) モデルとよぶ。

定理 21.2 を使うと、ARFIMA (p, d, q) モデルに従う時系列 $\{y_t\}$ のスペクトラムは、

$$f_y(\lambda; d) = \frac{\sigma^2}{2\pi} \frac{1}{|1 - e^{i\lambda}|^{2d}} \frac{|\theta(e^{i\lambda})|^2}{|\phi(e^{i\lambda})|^2} = \frac{\sigma^2}{2\pi} \frac{1}{\left(4 \sin^2 \frac{\lambda}{2}\right)^d} \frac{|\theta(e^{i\lambda})|^2}{|\phi(e^{i\lambda})|^2} \quad (21.20)$$

で与えられる。このスペクトラムは、 d が正ならば、原点では発散して不連続である。実際、 $\lambda \rightarrow 0$ のとき、 $f_y(\lambda; d) = O(\lambda^{-2d})$ となり、長期的な周期成分が支配的となる。図 21.3 には、そのような場合のスペクトラムが例示されている。これは、ARFIMA $(2, d, 0)$ で、 $d=0.4$ 、 $\phi_1=0.8$ 、 $\phi_2=-0.64$ 、 $\sigma^2=1$ の場合のスペクトラムである。一般に、ARFIMA モデルの自己共分散 $\gamma(h)$ は、 $h \rightarrow \infty$ のとき、 $\gamma(h) = O(h^{2d-1})$ となり、 d が正ならば絶対総和不可能で、非常にゆっくりと 0 に減少する。このような性質をもつ時系列を長期記憶時系列 (long-memory time series) という。

ARFIMA (p, d, q) モデルでは、差分パラメータ d の推定が重要となる。そのための方法としては、時間領域における最尤法のほかに、周波数領域における最尤法や、あとで説明するウェーブレット領域における最尤法が使われる。

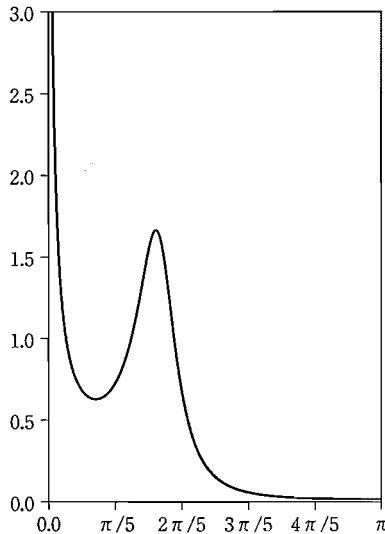


図 21.3 ARFIMA $(2, 0.4, 0)$ モデルのスペクトラム

21.2 ウェーブレット解析

ウェーブレット解析は、きわめて実用的な理由から工学の分野で生まれた手法である。そのルーツは、1980年代初頭にフランスの石油探査技師J. モルレーが考案した解析手法にあるとされている(榊原, 1995)。モルレーは、油床が存在する地層を特定化するために地中に振動を与え、人工的に作り出された地震波を解析する研究に従事していたが、実際に観測される地震波は、油床以外の地層や異質物の影響により、多くの不規則で局所的なノイズを含んでおり、伝統的なフーリエ変換によるスペクトル解析の方法では満足いく結果が得られなかった。そこで、短い波(ウェーブレット)を局所的に拡大縮小したり、平行移動したりすることにより、地震波を時間と周波数の両面から解析する手法を考案したのである。

図 21.4 には、ある離散時間確率過程 $\{x_i\}$ から得られたデータ(標本サイズ=512)のウェーブレット変換が示されている。1番上の時系列は原系列である。2番目以降の系列は、原系列を高周波数から低周波数の順にレベルごとにウェーブレット変換した系列である。高周波数の変換のほうが、短い周期に対応した変動をみる必要があるため、より多くの時点で計算されている。これらがウェーブレット領域を形成している。各レベルの時間的な変動をみることにより、原系列は、高周波よりも低周波

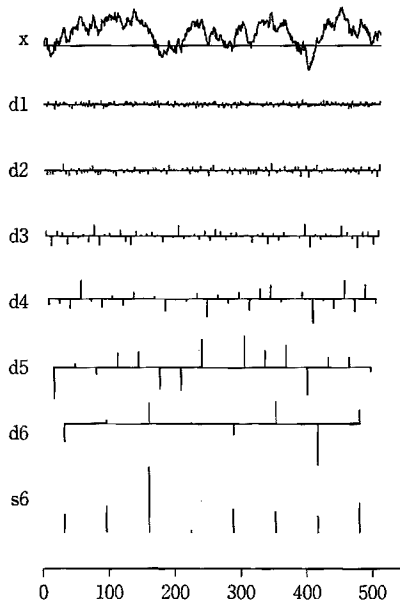


図 21.4 原系列とウェーブレット変換

の変動のほうが大きいことがわかる。ウェーブレット解析は、これらのウェーブレット変換系列からさまざまな統計量を計算して、原系列に対して想定されるモデルのパラメータに関する統計的推測を行うものである。

21.2.1 ウェーブレット変換 (1) : 連続時間確率過程の場合

連続時間 $t \in R$ で定義される時系列 $\{x(t)\}$ を考えよう。ここで、 $x(t)$ は R 上で 2 乗可積分な関数空間 $L^2(R)$ に属する。すなわち、

$$\|x\|^2 = \int_{-\infty}^{\infty} x^2(t) dt < \infty$$

であると仮定する。このような $x(t)$ のウェーブレット解析を行うためには、ウェーブレット関数 $\psi(t)$ を導入する。 $\psi(t)$ は、 $x(t)$ と同様に、 $L^2(R)$ に属する関数であり、次の 2 つの条件を満たすものと仮定される。

$$\int_{-\infty}^{\infty} \psi(t) dt = 0, \quad \int_{-\infty}^{\infty} \psi^2(t) dt = 1 \quad (21.21)$$

第 1 の条件は、 $\psi(t)$ が 0 のまわりを上下運動する様子を想起させる。第 2 の条件は、その上下運動の全体が有限であることを示唆する。

このとき、 $x(t)$ の連続ウェーブレット変換 (continuous wavelet transform: CWT) は、次のように定義される。

$$C_{a,b}(\psi, x) = \frac{1}{\sqrt{a}} \int_{-\infty}^{\infty} x(t) \psi\left(\frac{t-b}{a}\right) dt = \int_{-\infty}^{\infty} x(t) \tilde{\psi}_{a,b}(t) dt. \quad (21.22)$$

ここで、 $\tilde{\psi}_{a,b}(t) = \psi((t-b)/a)/\sqrt{a}$ ($0 < a < \infty$, $-\infty < b < \infty$) である。CWT は、2 つのパラメータ a と b をもっている。 a は、波長の拡大・縮小を制御するスケール・パラメータであり、 a が小さいほど波長を縮小し、大きくなるにつれて拡大する。この意味で、 $1/a$ が周波数の役割を果たすことになる。他方、 b はシフト・パラメータであり、波長の拡大・縮小の基準時点を調整する時間的な役割をもっている。

ウェーブレット関数のもっとも簡単な例として、次に定義される Haar 関数 ψ_H (x) を取り上げよう。

$$\psi_H(x) = \begin{cases} 1 & (0 \leq x < 1/2) \\ -1 & (1/2 \leq x < 1) \\ 0 & (\text{それ以外のとき}) \end{cases}$$

この場合の CWT は、

$$C_{a,b}(\psi_H, x) = \frac{1}{\sqrt{a}} \left[\int_b^{b+a/2} x(t) dt - \int_{b+a/2}^{b+a} x(t) dt \right]$$

となる。これは、時点 b からはじまって、幅 $a/2$ の隣りあう 2 つの区間上での $x(t)$ の平均 (の定数倍) の差を計算したものである。パラメータ a を大きくすれば、 $x(t)$ の大局的なふるまいをみることができ、小さくすれば、局所的なふるまいをみることが可能になる。もっと一般のウェーブレット関数を使った CWT では、このような単純な解釈は困難であるが、基本的な役割は同じである。

21.2.2 ウェーブレット変換 (2) : 離散時間確率過程の場合

離散時間確率過程 $\{x_t\}$ からの観測値を列ベクトル $\mathbf{x} = (x_1, x_2, \dots, x_T)'$ で表す。ここで、標本サイズ T は $T=2^J$ (J は自然数) であると仮定する。このとき、 \mathbf{x} の離散ウェーブレット変換 (discrete wavelet transform: DWT) とは、次の変換

$$\mathbf{w} = \mathbf{W}\mathbf{x}, \quad \mathbf{w} = \begin{pmatrix} w_1 \\ \vdots \\ w_j \\ \vdots \\ v_j \end{pmatrix}, \quad \mathbf{W} = \begin{pmatrix} W_1 \\ \vdots \\ W_j \\ \vdots \\ V_j \end{pmatrix} \quad (21.23)$$

で、以下で述べる条件を満たすものをいう。

\mathbf{W} はウェーブレット変換行列とよばれる直交行列であり、連続時間のウェーブレット関数 $\psi(t)$ に対応するものである。その構成部分 W_j は、レベル j の変換行列であり、 $T/2^j$ 個の行からなる行列である。最後の V_j は W_j と同様に行ベクトルであり、すべての要素は $1/\sqrt{T}$ からなっている。ウェーブレット変換行列は、データにフィルター演算をするフィルターの役割を果たしており、Daubechies のフィルターなど、幅の異なるさまざまなフィルターがある。

他方、 \mathbf{w} はウェーブレット係数ベクトルであり、通常は、単にウェーブレットとよばれるものである。その構成部分 w_j はレベル j のウェーブレットとよばれ、 $T/2^j$ 個の要素からなるベクトルである。ウェーブレットは、レベル 1 においてもっとも解像度の高い、高周波の変換をもたらす。そのために、局所的な時点での計算が必要であり、その結果、もっとも多くの要素 $T/2$ 個を含むことになる。すなわち、レベル 1 では、ハイパス・フィルターによる演算が行われる。他方、レベル j が大きくなるにつれて、しだいに解像度の低い、低周波の変換に移行するので、大局的な時点での計算となり、要素も少なくなる。すなわち、ローパス・フィルターによる演算が行われる。実際、 w_j はスカラーである。また、 \mathbf{w} の最後の要素である v_j もスカラーであり、レベル J のスケーリング係数とよばれる。 V_j の定義から、 $v_j = V_j \mathbf{x} = \sqrt{T} \bar{x}$ となる。

ウェーブレット変換の計算を、(21.23) 式の行列演算で実行すれば、計算量は T^2 のオーダーとなる。これに対して、 \mathbf{W} が疎な行列であることを使って、 T のオーダーで実行される効率的なアルゴリズムを考えることができる。それは、ウェーブレットをレベルごとに逐次計算する方法であり、計算量を $O(T)$ に減らすもので、ピラミッド・アルゴリズムとよばれる (Percival and Walden, 2000)。

ウェーブレット変換は、上で説明した DWT に限られるわけではない。DWT においては、レベル j のウェーブレット w_j は $T/2^j$ 個の要素からなっている。これに対して、各レベルに同数の T 個の要素をもたらし、しかも、レベル j においては周波数を 2^j 個の周波数帯に等分割したうえで、各周波数帯に $T/2^j$ 個の要素を振り分ける変換がある。このようにして得られる変換を DWPT (離散ウェーブレット・パケット変換) という (Bruce and Gao, 1996; Percival and Walden, 2000)。DWPT は、

各レベルにおいて、よりきめの細かい周波数成分を抽出する働きがあるので、季節性を含むようなデータの分析に適している (Gençay-Selçuk-Whitcher, 2002)。

DWPTは、DWTと同様に直交ウェーブレット変換である。他方、非直交変換も提案されている。その1つが、Percival and Walden (2000)においてMODWT (maximal overlap DWT)と命名され、Bruce and Gao (1996)においてNon-Decimated DWTと名付けられた変換である。通常のDWTが、レベル j に $T/2^j$ 個のウェーブレットをもたらすのに対して、MODWTは、各レベルに T 個のウェーブレットをもたらす。この点ではDWPTと同様にみえるが、DWPTは各レベルで周波数帯という新たなパラメータを導入した直交変換であるという点で異なる。なお、DWPTに対しても、各レベルの各周波数帯に T 個のウェーブレット係数をもたらす非直交変換が考えられる。これをMODWPT (maximal overlap DWPT)という。

MODWTあるいはMODWPTは、変換の際に重複やムダを生じる。計算量も $T \log_2 T$ に増えることが知られている。しかし、① 標本サイズが2のべき乗でなくともよい、② 結果が時系列の初期時点のとり方に依存しない、③ 原系列との位相のずれを生じない、などのDWTやDWPTにない利点がある。また、あとで述べるウェーブレット分散の推定量は、これらのウェーブレットにもとづくほうが、より精度の高い結果をもたらすことなどが知られている (Percival, 1995)。他方、直交性がないので、最尤推定には短所となるなど、それぞれのウェーブレット変換の長所、短所については、さらに検討する必要があるだろう。

21.2.3 ウェーブレット変換の利点

データをウェーブレット変換することの長所は何であろうか。本節では、このことについて、2つの観点から説明する。

a. フーリエ変換との比較

ウェーブレット変換の第1の利点は、時間と周波数の局在性 (=時間領域と周波数領域における原系列の局所的な挙動に関する情報の抽出能力)である。その具体的な意味合いを考えるために、次の人工的な4種類のデータ (標本サイズ $T=128$)を取り上げよう。

$$\begin{aligned} x_{1t} &= \cos\left(\frac{\pi t}{16}\right), & x_{2t} &= \begin{cases} x_{1t} & (t=56, \dots, 72), \\ 0 & (\text{そのほか}), \end{cases} \\ x_{3t} &= \frac{x_{1t}}{20} + x_{2t}, & x_{4t} &= \begin{cases} 0.7t & (t=56, \dots, 72), \\ 10+0.7t & (\text{そのほか}). \end{cases} \end{aligned}$$

データの特徴としては、 x_1 は周期的データ、 x_2 はダミー的データ、 x_3 はこれらの混合、 x_4 は切片の一時的シフトを含むトレンド・データである。これらのデータに対して、FFT (有限フーリエ変換)とDWTを適用して、説明力を比較しよう。比較の基準として、標準化された累積2乗和

表 21.1 FFT と DWT のふるまいの比較

データ	x_1	x_2	x_3	x_4
原系列	105	15	38	112
FFT	2	25	23	25
DWT	27	11	16	12

$$C_t = \frac{\sum_{i=1}^t |\hat{x}_{(i)}|^2}{\sum_{i=1}^T |x_i|^2} \quad (t=1, \dots, T) \quad (21.24)$$

を使うことにする。ここで、 $\{\hat{x}_{(i)}\}$ は、原系列 $\{x_i\}$ を FFT あるいは DWT によって変換した系列 $\{\hat{x}_i\}$ を絶対値の大きい順に並べ替えた順序統計量である。 $\{C_t\}$ は 0 から 1 へ単調に増加するが、より早く増加するほうが説明力があると考えられる。

表 21.1 には、 C_t の値が 0.99 をはじめてこえる t の値を、FFT, DWT および原系列に対して示してある。もちろん、この値が小さいほうが望ましい。予想されることであるが、FFT は周期的データ x_1 , 原系列はダミー的データ x_2 に対して比較的ふるまいがよい。しかし、FFT と原系列は、それ以外のデータに対しては相性が悪い。他方、DWT は x_1 に対しては次善のふるまいをしているが、ダミー的データ x_2 , 混合的データ x_3 , およびシフトをとまなうトレンド・データ x_4 に対しては最善のふるまいをしていることがわかる。

b. ARFIMA 過程のウェーブレット変換

離散時間時系列 $\{x_t\}$ が、(21.19) 式で定義された ARFIMA (p, d, q) モデルに従う場合を考えよう。そこで述べたように、 $u_t = \phi^{-1}(L)\theta(L)\varepsilon_t$ が定常で、しかも、差分パラメータ d が $1/2$ よりも小さいならば、 $\{x_t\}$ は定常となる。ここでは非定常の場合も考慮して、 d は任意の正数とする。とくに、 $d=1$ の場合は、単位根系列となる。なお、非定常な場合には、厳密な意味でのスペクトラムは存在しないが、それに相当するもの考えることができる。実際、(18.20) 式で定義された定常な場合のスペクトラムを d が任意の正数の場合にも拡張して考え、形のうえでは定常な場合と同様に議論することが正当化される (Solo, 1992)。この場合のスペクトラムは、原点の近傍で周波数の逆数 (のべき乗) に比例することから、そのようなスペクトラムをもつような確率過程を $1/f$ 過程とよぶ。

非定常な ARFIMA (p, d, q) モデルに従う時系列に対しては、ウェーブレット変換に使うフィルターの幅 m が $m \geq 2d$ ならば、ウェーブレット系列は、ほぼ無相関の定常過程になり、しかも、レベルごとに一定の分散をもつことが知られている (Percival and Walden, 2000)。すなわち、レベル j のウェーブレットを W_{jt} ($j=1, 2, \dots, J; t=1, \dots, T/2^j$) とするとき、

$$\text{Cov}(W_{js}, W_{kt}) \approx 0 \quad (s \neq t), \quad V(W_{jt}) \approx \sigma_j^2$$

が成り立つ。分散 σ_j^2 は、レベル j のウェーブレット分散とよばれる。

上記の性質を利用して、ARFIMA モデルの差分パラメータ d の推定をウェーブレット領域で行うことができる。まず、レベル j のウェーブレットのスペクトラムは、 $|H_{jm}(\lambda)|^2 f_x(\lambda)$ で与えられる。ここで、 $|H_{jm}(\lambda)|^2$ は、フィルターの幅 m 、レベル j のウェーブレット変換のパワー伝達関数、 $f_x(\lambda)$ は原系列 $\{x_t\}$ のスペクトラムである。このとき、レベル j のウェーブレット分散は時点に依存せずに、

$$\sigma_j^2 = \int_{-\pi}^{\pi} |H_{jm}(\lambda)|^2 f_x(\lambda) d\lambda = \frac{\sigma^2}{2\pi 4^d} \int_{-\pi}^{\pi} \frac{|H_{jm}(\lambda)|^2 |\theta(e^{i\lambda})|^2}{\left(\sin^2 \frac{\lambda}{2}\right)^d |\phi(e^{i\lambda})|^2} d\lambda \quad (21.25)$$

で与えられる。

(21.25) 式で与えられたウェーブレット分散の表現をさらに簡単にしよう。そのために、ウェーブレット変換のフィルターがオクターブ周波数帯 $[-2\pi/2^j, -2\pi/2^{j+1}]$ と $[2\pi/2^{j+1}, 2\pi/2^j]$ 上のバンドパス・フィルターで近似できる場合を考えよう。すなわち、パワー伝達関数が、

$$|H_{jm}(\lambda)|^2 \approx \begin{cases} 2^j & (2\pi/2^{j+1} \leq |\omega| \leq 2\pi/2^j \text{ のとき}) \\ 0 & (\text{そのほか}) \end{cases}$$

となる場合である。バンドパス・フィルターは理想的なフィルターであり、有限幅のフィルターでは実現不可能であるが、現実にはフィルターの幅 m を大きくすることで上の近似式が正当化される。さらに、各オクターブ周波数帯上で $|\theta(e^{i\lambda})|^2 / |\phi(e^{i\lambda})|^2$ を定数とみなすことができれば、次の近似式が成り立つ。

$$\begin{aligned} \sigma_j^2 &\approx \frac{\sigma^2}{2\pi} \frac{2^{j+1}}{4^d} \int_{2\pi/2^{j+1}}^{2\pi/2^j} \frac{|\theta(e^{i\lambda})|^2}{\left(\sin^2 \frac{\lambda}{2}\right)^d |\phi(e^{i\lambda})|^2} d\lambda \approx K_1 2^j \int_{2\pi/2^{j+1}}^{2\pi/2^j} \lambda^{-2d} d\lambda \\ &= K_2 4^{jd} \end{aligned} \quad (21.26)$$

この近似式において、未知のウェーブレット分散 σ_j^2 を、与えられたデータの DWT から得られる各レベルの標本分散 $\hat{\sigma}_j^2$ で置きかえることにより、レベルごとの対数線形回帰式

$$\log \hat{\sigma}_j^2 = a + d \log 4^j + e_j \quad (j = j_0, j_0+1, \dots, j_1) \quad (21.27)$$

が得られる。これから、差分パラメータ d を最小 2 乗推定により求めることができる。ただし、ウェーブレット回帰では、近似 $\sin^{-2d} \omega \approx \omega^{-2d}$ を使っており、これは低周波においてのみ有効な近似である。したがって、高周波に対応するレベル 1 では無効なので、 j_0 は少なくとも 2 とする。

ウェーブレット分散の推定は、 $j=1$ の場合に最大の自由度をもたらすが、この場合を排除する点は、ウェーブレット回帰の短所である。他方、 j が大きいと標本分散の自由度が小さくなり、推定量の精度が悪くなるので、あまり大きくてはいけなく、という制約も課せられる。実際、 $T=2^j$ の場合、レベルは J まで考えられるが、利用可能なデータは、レベル J で 1 個、レベル $J-1$ で 2 個、レベル $J-2$ では 4 個で

ある。したがって, j_1 は, 多くとも $J-3$ 程度とするのが合理的である。

例として, 図 21.4 に示したデータを取り上げよう。実は, このデータは, (21.19) 式の ARFIMA モデルにおいて, $d=0.9$, $\phi(L)=\theta(L)=1$ として, $T=2^J=512$ ($J=9$) 個のデータを生成したものである。図 21.5 は, このデータに対して, (21.27) 式の対数線形回帰をあてはめた様子を図示したものである。実線は 5 個の点 ($j_0=2, j_1=6$) にもとづく回帰であり, d の推定値は 0.974 であった。他方, 点線は 4 個の点 ($j_0=2, j_1=5$) にもとづく回帰であり, d の推定値は 0.937 であった。このような実験を 1000 回くり返した結果, 5 個と 4 個の点それぞれから得られた d の推定値の平均は, 0.856 と 0.859, 標準偏差は, 0.093 と 0.088 であり, 4 個の点にもとづく回帰のほうが, わずかではあるがよい結果をもたらした。

ウェーブレット回帰は, 不均一分散があるにもかかわらず最小 2 乗推定を行うものであるが, その欠点を補い, さらに精度のよい推定を行うために, ウェーブレットにもとづく MLE を考えることができる。誤差項に正規性を仮定すれば, 対数尤度関数は,

$$l(\alpha) = -\frac{T}{2} \log(2\pi\sigma^2) - \frac{1}{2} \log|\Phi| - \frac{1}{2\sigma^2} \mathbf{x}'\Phi^{-1}\mathbf{x} \quad (21.28)$$

で与えられる。ここで, α はモデルの未知パラメータすべてからなるベクトルである。また, $\sigma^2\Phi$ は観測値ベクトル $\mathbf{x}=(x_1, \dots, x_T)'$ の共分散行列であるが, 一般の ARFIMA (p, d, q) の場合には, 明示的な表現は不可能であり, 共分散行列の各要素の計算そのものも面倒となる。そして, 対数尤度の最大化のためには, くり返し計

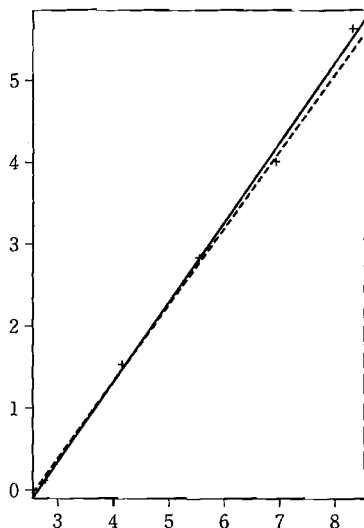


図 21.5 ウェーブレット回帰

算のアルゴリズムが使われるが、その際、共分散行列の逆行列および行列式の計算がくり返し行われ、その計算量は T が大きくなるにつれて膨大なものとなる。

このような困難を回避するための1つの方法は、周波数領域におけるMLEを考えることである。しかし、その場合のMLEは、 d が1をこえる場合は一致性が保証されない (Velasco and Robinson, 2000)。それに対して、ウェーブレット領域におけるMLEは、原系列が非定常であっても、ウェーブレット変換系列は定常的となり、MLEは一致性が保証される。ウェーブレット領域におけるMLEは、次のように求めることができる。まず、観測値ベクトル \mathbf{x} のDWTを $\mathbf{w} = \mathbf{W}\mathbf{x}$ とすると、DWTがほぼ無相関になるという性質を利用して、(21.28) 式の尤度関数は次のように表現される。

$$\begin{aligned} l(\mathbf{a}) &= -\frac{T}{2} \log(2\pi\sigma^2) - \frac{1}{2} \log |\mathbf{W}\Phi\mathbf{W}'| - \frac{1}{2\sigma^2} (\mathbf{W}\mathbf{x})' (\mathbf{W}\Phi\mathbf{W}')^{-1} (\mathbf{W}\mathbf{x}) \\ &\approx -\frac{T}{2} \log(2\pi\sigma^2) - \frac{1}{2} \sum_{j=1}^J T_j \log h_j - \frac{1}{2\sigma^2} \sum_{j=1}^J \frac{\mathbf{w}'_j \mathbf{w}_j}{h_j}. \end{aligned} \quad (21.29)$$

ここで、 $T_j = T/2^j$ であり、 \mathbf{w}_j はレベル j のウェーブレットである。また、

$$h_j = \frac{2^{j+1}}{2\pi 4^d} \int_{2\pi/2^{j+1}}^{2\pi/2^j} \frac{|\theta(e^{i\lambda})|^2}{\left(\sin^2 \frac{\lambda}{2}\right)^d |\phi(e^{i\lambda})|^2} d\lambda$$

であり、定積分はパラメータの値が与えられれば数値積分で計算することができる。

ウェーブレット回帰およびウェーブレットにもとづく最尤推定、さらに周波数領域における各種の推定量については、Jensen (2000)、Percival and Walden (2000)、Tanaka (2004) などにおいて、シミュレーション実験により比較検討がなされている。これらの結果から、次のような事実が見だされた。

- (1) 周波数領域におけるMLEは、差分パラメータ d が1以下の場合には非常に良好である。しかし、 d が1をこえるとふるまいが悪くなる。
- (2) ウェーブレット回帰にもとづく推定量は、 $d \leq 1$ の場合にはよくないが、 $d > 1$ ならば、周波数領域におけるMLEよりも優れている。しかし、 d が大きくなるにつれて、下方へのバイアスがみられる。
- (3) ウェーブレット領域におけるMLEは、 $d \geq 1$ の場合でも優れている。また、 $d < 1$ の場合でも、良好なふるまいをしている。 [田中勝人]

■文 献

- 猪狩 惺 (1996). 実解析入門, 岩波書店.
 榭原 進 (1995). ウェーブレットビギナーズガイド, 東京電機大学出版局.
 謝 東深・鈴木 武 (2002). ウェーブレットと確率過程入門, 内田老鶴圃.
 田中勝人 (2005). ウェーブレット解析の統計学への応用について, 数学, 57(1), 50-69.
 山本 拓 (1988). 経済の時系列分析, 創文社.
 Anderson, T. W. (1971). *The Statistical Analysis of Time Series*, John Wiley & Sons.
 Bruce, A. and Gao, H. Y. (1996). *Applied Wavelet Analysis with S-PLUS*, Springer.

- Daubechies, I. (1992). *Ten Lectures on Wavelets*, SIAM.
- Flandrin, P. (1989). On the spectrum of fractional Brownian motions, *IEEE Transactions on Information Theory*, **35**, 197-199.
- Fuller, W. A. (1996). *Introduction to Statistical Time Series*, 2nd ed., John Wiley & Sons.
- Gençay, R., Selçuk, F. and Whitcher, B. (2002). *An Introduction to Wavelets and Other Filtering Methods in Finance and Economics*, Academic Press.
- Hosking, J. R. M. (1981). Fractional differencing, *Biometrika*, **68**, 165-176.
- Jensen, M. J. (2000). An alternative maximum likelihood estimator of long-memory processes using compactly supported wavelets, *Journal of Economic Dynamics and Control*, **24**, 361-387.
- McCoy, E. J. and Walden, A. T. (1996). Wavelet analysis and synthesis of stationary long-memory processes, *Journal of Computational and Graphical Statistics*, **5**, 26-56.
- Percival, D. B. (1995). On Estimation of the Wavelet Variance, *Biometrika*, **82**, 619-631.
- Percival, D. B. and Walden, A. T. (2000). *Wavelet Methods for Time Series Analysis*, Cambridge University Press.
- Solo, V. (1992). Intrinsic random functions and the paradox of $1/f$ noise, *SIAM Journal of Applied Mathematics*, **52**, 270-291.
- Tanaka, K. (1996). *Time Series Analysis: Nonstationary and Noninvertible Distribution Theory*, John Wiley & Sons.
- Tanaka, K. (2004). Frequency domain and wavelet-based estimation for long-memory signal plus noise models, In *State Space ad Unobserved Component Models, Festschrift in honor of Professor J. Durbin*, Cambridge University Press. pp. 75-91.
- Velasco, C. and Robinson, P. (2000). Whittle pseudo-maximum likelihood estimation for non-stationary time series, *Journal of the American Statistical Association*, **95**, 1229-1243.