時系列解析入門 - 基礎と準備

1. 時系列の分類
   - 連続時間時系列と離散時間時系列
     さらに離散時間時系列には等間隔と不等間隔があるが、解析を行うのは等間隔の離散時間時系列がほとんど
   - 一変量時系列と多変量時系列
   - 定常時系列と非定常時系列
     変動が時間的に変化しないものを定常、変動や平均が時間とともに変化するものを非定常とよぶ
   - ガウス型時系列と非ガウス型時系列
     時系列の分布が正規分布に従うものをガウス型という
   - 線形時系列と非線形時系列
   - 欠損値と異常値

2. 時系列解析の目的
   - 記述
     グラフ化、標本自己共分散関数・標本自己相関関数・ピリオドグラムなどの基本統計量による特徴の表現など
   - モデリング
     与えられた時系列に対し、その変動の仕方を表現する時系列モデルを構成し、時系列の確率的構造を解析すること
   - 予測
   - 信号抽出

3. 時系列データの前処理
   時系列データに非定常性が見られた場合にデータを処理し定常化する場合がある。
   - 変数変換
     対数変換： $z_n = \log(y_n)$
     変動の分散が一様になったり、誤差分布が正規分布になったりする場合がある。
     対数変換を特殊な場合として含む方法に Box-Cox 変換がある。

     ロジット変換： $z_n = \log\left(\frac{y_n}{1-y_n}\right)$

     $(0,1)$上の値を取る時系列の場合にこの変換を施すことによって$(-\infty,\infty)$上の値をとる時系列に変換できる。

- 差分（階差）

時系列が顕著なトレンドを含む場合は、差分系列を求めて解析する場合がある。

$$z_n = \Delta y_n = y_n - y_{n-1}$$

これは、$y_n = a + bn$ と表される場合には、$z_n = \Delta y_n = b$ となって直線の傾きを除去できるということから直感的にも理解しやすい。

さらに $y_n = a + bn + cn^2$ の場合にはさらに z_n の差分を求めると、

$$\begin{aligned}\Delta z_n &= z_n - z_{n-1} \\ &= \Delta y_n - \Delta y_{n-1} \\ &= (b + 2cn) - (b + 2c(n-1)) \\ &= 2c\end{aligned}$$

となって、2 次成分および 1 次成分を除去できる。

周期性のあるデータでは、一周期前の値との差分を利用することがある。

$$\Delta_p y_n = y_n - y_{n-p}$$

- 前期比・前年同期比

経済時系列では前期比や前年同期比を分析することがある。

$$z_n = \frac{y_n}{y_{n-1}}$$

$$x_n = \frac{y_n}{y_{n-p}}$$

時系列 y_n が $y_n = T_n \omega_n$ とトレンドとノイズの積で表され、T_n が $(1-\alpha)T_{n-1}$ によって成長率 α で変化する場合、

$$z_n = \frac{y_n}{y_{n-1}} = \frac{T_n \omega_n}{T_{n-1}\omega_{n-1}} = (1-\alpha)\frac{\omega_n}{\omega_{n-1}}$$

となって成長率 α を検出できる。

時系列 y_n が周期 p の周期関数 s_n のノイズの積として、$y_n = s_n \bullet \omega_n, s_n = s_{n-1}$ として表せる場合は

$$x_n = \frac{y_n}{y_{n-p}} = \frac{s_n \omega_n}{s_{n-p}\omega_{n-p}} = \frac{\omega_n}{\omega_{n-p}}$$

となって周期成分を除去できる。

- 移動平均

変動の激しい時系列を滑らかにする簡便な方法として移動平均がある。

時系列 y_n が与えられる時、$2k+1$ 項の移動平均は、

$$T_n = \frac{1}{2k-1}\sum_{j=-k}^{k} y_{n+j}$$

で定義される。

もとの時系列が、$y_n = t_n + \omega_n, t_n = a + bn$ と直線とノイズの和で表される場合、T_n の平均は t_n と同じで分散は ω_n の分散の $1/(2k+1)$ となる。

移動平均を一般化したものとして、$\sum_{j=-k}^{k}\omega_j = 1, \omega_j \geq 0$ を満たす重み係数を用いて定義される重みつき移動平均、

$$T_n = \sum_{j=-k}^{k}\omega_j y_{n-j}$$

がある。

また移動平均の定義において、平均をメディアンに置き換えると、移動メディアン $T_n = \mathrm{median}\{T_{n-k},...,T_n,...T_{n+k}\}$ が得られる。

誤植等はご指摘ください。

（この項目・了）