時系列分析 5
多変量時系列モデルの性質、推定法、および予測

担当：長倉 大輔
（ながくらだいすけ）

多変量時系列モデルの性質、推定法、および予測

- 多変量時系列モデル

一変量時系列モデルがスカラー変数 y_t についてのモデルだったのに対して、多変量時系列モデルは $n\times 1$ ベクトル変数

$$\mathbf{y}_t = (y_{1t}, y_{2t}, \ldots, y_{nt})^{\mathrm{T}}$$

に対して、その時間を通じての確率的な挙動を記述するモデルである。ここで " T " は行列の転置を表す。

一変量モデルの拡張であり、概念的にはそこまで難しくないが、数学的には行列、ベクトルに関する知識が必要となってくる。

多変量時系列モデルの性質、推定法、および予測

■ 多変量時系列モデルの期待値

多変量時系列変数 $\mathbf{y}_t$ の**期待値**は

$$E(\mathbf{y}_t) = [\ E(y_{1t}), E(y_{2t}), \ldots., E(y_{nt})\ ]^{\mathrm{T}}$$

と定義される。つまりそれぞれの成分の期待値のベクトルである。これは一変量の時と同様、一般的には時点 t に依存するとする。

多変量時系列モデルの性質、推定法、および予測

- 多変量時系列モデルの自己共分散行列

多変量時系列変数 $\mathbf{y}_t$ の **k 次自己共分散行列**は

(i, j) 成分が $\mathrm{cov}(y_{i,t}, y_{j,t-k})$ である $n \times n$ 行列

(この場合対角成分は $y_{i,t}$ の k 次の自己共分散)。

である。つまりそれぞれベクトルの成分の共分散を並べたものである。

これは具体的には次のように表すことができる。

多変量時系列モデルの性質、推定法、および予測

- 多変量時系列モデルの自己共分散行列

多変量時系列変数 $\mathbf{y}_t$ の **k 次自己共分散行列**:

$$\operatorname{cov}(\mathbf{y}_t, \mathbf{y}_{t-k}) = \begin{bmatrix} \operatorname{cov}(y_{1,t}, y_{1,t-k}) & \operatorname{cov}(y_{1,t}, y_{2,t-k}) & \cdots & \operatorname{cov}(y_{1,t}, y_{n,t-k}) \\ \operatorname{cov}(y_{2,t}, y_{1,t-k}) & \operatorname{cov}(y_{2,t}, y_{2,t-k}) & \cdots & \operatorname{cov}(y_{2,t}, y_{n,t-k}) \\ \vdots & \vdots & \ddots & \vdots \\ \operatorname{cov}(y_{n,t}, y_{1,t-k}) & \operatorname{cov}(y_{n,t}, y_{2,t-k}) & \cdots & \operatorname{cov}(y_{n,t}, y_{n,t-k}) \end{bmatrix}$$

並び方に注意。これも一般的には時点 t に依存する。

$k = 0$ の時は $\mathbf{y}_t$ の**分散共分散行列**となり**対称行列**となる($k \neq 0$ の時には一般的には対称行列ではない)。

多変量時系列モデルの性質、推定法、および予測

- 多変量時系列モデルの自己共分散行列

また、多変量時系列変数 $\mathbf{y}_t$ の自己共分散行列は

$$\mathrm{cov}(\mathbf{y}_t, \mathbf{y}_{t-k}) = E[(\mathbf{y}_t - \boldsymbol{\mu}_t)(\mathbf{y}_{t-k} - \boldsymbol{\mu}_{t-k})^{\mathrm{T}}]$$

とも書ける(先ほどの表現はこれを実際に計算したものである)。ここで

$$\boldsymbol{\mu}_t = E(\mathbf{y}_t)$$

である。

多変量時系列モデルの性質、推定法、および予測

- 多変量(弱)定常過程

多変量時系列変数 $\mathbf{y}_t$ の平均、自己共分散は、一変量時系列変数の場合と同様、一般的には時点 t に依存する。

これらが時点 t に依存しない時、$\mathbf{y}_t$ は(弱)**定常過程**と言われる。

このような場合、今後はその期待値ベクトルと自己共分散行列を

$$\boldsymbol{\mu} = E(\mathbf{y}_t), \quad \boldsymbol{\Gamma}_k = \mathrm{cov}(\mathbf{y}_t, \mathbf{y}_{t-k})$$

という記号で表すとする。

多変量時系列モデルの性質、推定法、および予測

- 定常な多変量時系列の自己共分散行列の性質

定常な一変量時系列の自己共分散 γ_k は

$$\gamma_k = \gamma_{-k}$$

という性質をもっていた。定常な多変量時系列の**自己共分散行列**は

$$\mathbf{\Gamma}_k = \mathbf{\Gamma}_{-k}^{\mathrm{T}}$$

という性質を持っている。

多変量時系列モデルの性質、推定法、および予測

例題1 (自己共分散行列の性質)

多変量の定常自己共分散行列に対して

$$\boldsymbol{\Gamma}_k = \boldsymbol{\Gamma}_{-k}^{\mathrm{T}}$$

が成り立つことを確認せよ。

多変量時系列モデルの性質、推定法、および予測

- 定常な多変量時系列の自己相関行列

一変量の定常自己相関 ρ_k は $\rho_k = \gamma_k / \gamma_0$ と定義された。
多変量の**自己相関行列**は

$$\boldsymbol{\rho}_k = \mathrm{corr}(\mathbf{y}_t, \mathbf{y}_{t-k})$$

$$= \begin{bmatrix} \mathrm{corr}(y_{1,t}, y_{1,t-k}) & \mathrm{corr}(y_{1,t}, y_{2,t-k}) & \cdots & \mathrm{corr}(y_{1,t}, y_{n,t-k}) \\ \mathrm{corr}(y_{2,t}, y_{1,t-k}) & \mathrm{corr}(y_{2,t}, y_{2,t-k}) & \cdots & \mathrm{corr}(y_{2,t}, y_{n,t-k}) \\ \vdots & \vdots & \ddots & \vdots \\ \mathrm{corr}(y_{n,t}, y_{1,t-k}) & \mathrm{corr}(y_{n,t}, y_{2,t-k}) & \cdots & \mathrm{corr}(y_{n,t}, y_{n,t-k}) \end{bmatrix}$$

と定義される。ここで

$$\mathrm{corr}(y_{i,t}, y_{j,t-k}) = \frac{\mathrm{cov}(y_{i,t}, y_{j,t-k})}{\sqrt{\mathrm{var}(y_{i,t})}\sqrt{\mathrm{var}(y_{j,t-k})}}$$

である。

多変量時系列モデルの性質、推定法、および予測

■ 定常な多変量時系列の自己相関行列

対角行列 **D** を

$$\mathbf{D} = \begin{bmatrix} \mathrm{var}(y_{1t}) & 0 & \cdots & 0 \\ 0 & \mathrm{var}(y_{2t}) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \mathrm{var}(y_{nt}) \end{bmatrix}$$

と定義する。この時、自己相関行列は $\boldsymbol{\rho}_k = \mathbf{D}^{-1/2}\,\boldsymbol{\Gamma}_k\,\mathbf{D}^{-1/2}$ と書くことができる。これより

$$\boldsymbol{\rho}_k = \boldsymbol{\rho}_{-k}^{\mathrm{T}}$$

が成り立つ事がわかる。

多変量時系列モデルの性質、推定法、および予測

- ベクトルホワイトノイズ

ベクトル $\boldsymbol{\varepsilon}_t = (\varepsilon_{1t}, \varepsilon_{2t}, \ldots, \varepsilon_{nt})^{\mathrm{T}}$は

$$E(\boldsymbol{\varepsilon}_t) = \mathbf{0}, \quad \mathrm{cov}(\boldsymbol{\varepsilon}_t, \boldsymbol{\varepsilon}_{t-k}) = E(\boldsymbol{\varepsilon}_t \boldsymbol{\varepsilon}_{t-k}^{\mathrm{T}}) = \begin{cases} \boldsymbol{\Sigma}, & k = 0 \\ \mathbf{0} & k \neq 0 \end{cases}$$

を満たす時に**(ベクトル)ホワイトノイズ**と呼ばれる。ここで $\boldsymbol{\Sigma}$ は $n \times n$ **正定値行列**である。

これは一変量のホワイトノイズ $\varepsilon_t \sim \mathrm{W.N.}\ (\sigma^2)$を多変量に拡張したものである。

多変量時系列モデルの性質、推定法、および予測

- ベクトルホワイトノイズ

(ベクトル)ホワイトノイズは

(1) **弱定常過程**である。
(2) $\boldsymbol{\Sigma}$は一般的には(対称行列ではあるが) **対角行列ではない**。つまり同時点間の共分散は 0 ではない。

以後 $\boldsymbol{\varepsilon}_t$ が(ベクトル) ホワイトノイズである事を

$$\boldsymbol{\varepsilon}_t \sim \text{W.N.}\ (\boldsymbol{\Sigma})$$

と表記する事にする。

多変量時系列モデルの性質、推定法、および予測

- VAR モデル

ARモデルの多変量版として **Vector Autoregressive (VAR)モデル**がある。

p次のVARモデル(これは**VAR(p)**と書かれる)は次のように定義される。

$$\mathbf{y}_t = \mathbf{c} + \mathbf{\Phi}_1 \mathbf{y}_{t-1} + \cdots + \mathbf{\Phi}_p \mathbf{y}_{t-p} + \boldsymbol{\varepsilon}_t, \quad \boldsymbol{\varepsilon}_t \sim \mathrm{W.N.}(\mathbf{\Sigma})$$

ここで

$\mathbf{y}_t : n \times 1$ ベクトル、$\mathbf{c} : n \times 1$ベクトル、
$\mathbf{\Phi}_i : n \times n$ 行列、$\boldsymbol{\varepsilon}_t: n \times 1$ ベクトル

である。

多変量時系列モデルの性質、推定法、および予測

■ 2変量VAR モデル

2変量VAR(1)モデルは次のように書ける。

$$\begin{bmatrix} y_{1t} \\ y_{2t} \end{bmatrix} = \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} + \begin{bmatrix} \phi_{11} & \phi_{12} \\ \phi_{21} & \phi_{22} \end{bmatrix} \begin{bmatrix} y_{1t-1} \\ y_{2t-1} \end{bmatrix} + \begin{bmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{bmatrix},$$

$$\begin{bmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{bmatrix} \sim \text{W.N.}(\mathbf{\Sigma}),\ \mathbf{\Sigma} = \begin{bmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{bmatrix}.$$

これはまた次のように 2 本の回帰式として書く事もできる。

多変量時系列モデルの性質、推定法、および予測

- 2変量VAR モデル

$$\begin{cases} y_{1t} = c_1 + \phi_{11} y_{1t-1} + \phi_{12} y_{2t-1} + \varepsilon_{1t} \\ y_{2t} = c_2 + \phi_{21} y_{1t-1} + \phi_{22} y_{2t-1} + \varepsilon_{2t} \end{cases}$$

$$\mathrm{var}(\varepsilon_{1t}) = \sigma_1^2,\ \mathrm{var}(\varepsilon_{2t}) = \sigma_2^2,$$

$$\mathrm{cov}(\varepsilon_{1t}, \varepsilon_{2t}) = \rho\sigma_1\sigma_2.$$

どちらの書き方もよく使われる。

多変量時系列モデルの性質、推定法、および予測

- VAR モデルのラグオペレーターを用いた表現

VAR(p)モデルはラグオペレーターを用いて次のようにも表現される。

$$(\mathbf{I}_n - \mathbf{\Phi}_1 L - \cdots - \mathbf{\Phi}_p L^p)\mathbf{y}_t = \mathbf{c} + \boldsymbol{\varepsilon}_t,$$

ここで $\mathbf{I}_n$ は n 次単位行列である。また $\mathbf{y}_t$ が定常であればその期待値

$$\boldsymbol{\mu} = (\mathbf{I}_n - \mathbf{\Phi}_1 - \ldots - \mathbf{\Phi}_p)^{-1}\mathbf{c}$$

を用いて

$$\mathbf{y}_t - \boldsymbol{\mu} = \mathbf{\Phi}_1(\mathbf{y}_{t-1} - \boldsymbol{\mu}) + \cdots + \mathbf{\Phi}_p(\mathbf{y}_{t-p} - \boldsymbol{\mu}) + \boldsymbol{\varepsilon}_t,$$

と書くこともできる。

多変量時系列モデルの性質、推定法、および予測

例題2 (定常なVAR(p) モデルの期待値)

先ほどのVAR(p)モデルの期待値は

$$\boldsymbol{\mu} = (\mathbf{I}_n - \boldsymbol{\Phi}_1 - \ldots. - \boldsymbol{\Phi}_p)^{-1}\mathbf{c}$$

である事を示せ。

多変量時系列モデルの性質、推定法、および予測

- ユールウォーカー方程式

VAR(p)モデルの自己共分散行列は、次の**ユールウォーカー方程式(のような関係)の行列版**を満たす。

$$\boldsymbol{\Gamma}_k = \boldsymbol{\Phi}_1 \boldsymbol{\Gamma}_{k-1} + \dots + \boldsymbol{\Phi}_p \boldsymbol{\Gamma}_{k-p}$$

例えばVAR(1)の自己共分散行列は

$$\boldsymbol{\Gamma}_k = \boldsymbol{\Phi}_1 \boldsymbol{\Gamma}_{k-1} = \boldsymbol{\Phi}_1{}^2 \boldsymbol{\Gamma}_{k-2} = \dots = \boldsymbol{\Phi}_1{}^k \boldsymbol{\Gamma}_0$$

となる。

(注意: 1変量の場合と異なり、**自己相関行列**は上記のような関係を**満たさない**。教科書p.78の自己相関の関係は上記の自己共分散の関係の誤り)。

多変量時系列モデルの性質、推定法、および予測

- VAR モデルの定常性

VAR(p)はパラメータが次の条件を満たす時に(弱)定常過程となる。

(VAR(p)の定常性の条件)

$$|\mathbf{I}_n - \mathbf{\Phi}_1 z - \cdots - \mathbf{\Phi}_p z^p| = 0$$

の**z の全ての解の絶対値が 1 よりも大きい**。ここで $|\mathbf{A}|$ は行列 $\mathbf{A}$ の行列式を表す。

多変量時系列モデルの性質、推定法、および予測

- (例) 2変量VAR(1)の定常性の条件

例として2変量VAR(1)の定常性の条件がどうなるか見てみよう。次の2変量VAR(1)モデルを考える。

$$\mathbf{y}_t = \mathbf{c} + \mathbf{\Phi}_1 \mathbf{y}_{t-1} + \boldsymbol{\varepsilon}_t, \quad \boldsymbol{\varepsilon}_t \sim \mathrm{W.N.}(\mathbf{\Sigma})$$

ここで $\mathbf{y}_t = (y_{1t}, y_{2t})^{\mathrm{T}}$, $\quad \mathbf{\Phi}_1 = \begin{bmatrix} \phi_{11} & \phi_{12} \\ \phi_{21} & \phi_{22} \end{bmatrix}$ である。

定常性の条件は

$$|\mathbf{I}_2 - \mathbf{\Phi}_1 z| = 0$$

の z の全ての解の絶対値が 1 より大きいという事なのでこれを z について解いてみる

多変量時系列モデルの性質、推定法、および予測

- (例) 2変量VAR(1) の定常性の条件

$$
0 = |\mathbf{I}_2 - \mathbf{\Phi}_1 z| = \left| \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} - \begin{bmatrix} \phi_{11} z & \phi_{12} z \\ \phi_{21} z & \phi_{22} z \end{bmatrix} \right|
$$

$$
= \begin{vmatrix} 1 - \phi_{11} z & -\phi_{12} z \\ -\phi_{21} z & 1 - \phi_{22} z \end{vmatrix}
$$

$$
= (1 - \phi_{11} z)(1 - \phi_{22} z) - (-\phi_{12} z)(-\phi_{21} z)
$$

$$
= (\phi_{11}\phi_{22} - \phi_{12}\phi_{21}) z^2 - (\phi_{11} + \phi_{22}) z + 1
$$

$$
\Rightarrow \quad z = \frac{\phi_{11} + \phi_{22} \pm \sqrt{(\phi_{11} - \phi_{22})^2 + 4\phi_{12}\phi_{21}}}{2(\phi_{11}\phi_{22} - \phi_{12}\phi_{21})}
$$

となる。この z の2つの解の絶対値が1より大きければ定常となる。

多変量時系列モデルの性質、推定法、および予測

例題3

先ほどの2変量VAR(1)において$\phi_{21} = 0$ であるとすると定常条件はどのようになるか？

例題4

先ほどの2 変量 VAR(1)において$\phi_{21} = \phi_{12} = 0$ であるとすると定常条件はどのようになるか？

多変量時系列モデルの性質、推定法、および予測

- VARモデルの推定

VARモデルの推定には**最尤法**と**最小二乗法**の2 つがある。まず最尤法について見て行こう。

- VARモデルの尤度関数

一変量モデルの時と同様に尤度関数を求めるが、今回は$(\mathbf{y}_T^{\mathrm{T}}, \mathbf{y}_{T-1}^{\mathrm{T}}, \ldots, \mathbf{y}_1^{\mathrm{T}})^{\mathrm{T}}$ の結合密度関数を求めなければならない(転置を表す “T” と観測数を表す “T” の混同に注意。後者はイタリック)。

一変量の時と同様、結合密度関数を条件付き密度関数に分解する。

多変量時系列モデルの性質、推定法、および予測

- VARモデルの尤度関数

$(\mathbf{y}_T^{\mathrm{T}}, \mathbf{y}_{T-1}^{\mathrm{T}}, \ldots, \mathbf{y}_1^{\mathrm{T}})^{\mathrm{T}}$ の同時密度関数は、同時密度関数の条件付き密度関数への分解を用いて、

$$
\begin{aligned}
f(\mathbf{y}_T, \mathbf{y}_{T-1}, \ldots, \mathbf{y}_1) &= f(\mathbf{y}_T \mid \mathbf{y}_{T-1}, \ldots, \mathbf{y}_1) f(\mathbf{y}_{T-1} \mid \mathbf{y}_{T-2}, \ldots, \mathbf{y}_1) \times \\
&\quad \cdots \times f(\mathbf{y}_{p+1} \mid \mathbf{y}_p, \ldots, \mathbf{y}_1) f(\mathbf{y}_p, \ldots, \mathbf{y}_1) \\
&= f(\mathbf{y}_T \mid \mathbf{y}_{T-1}, \ldots, \mathbf{y}_{T-p}) f(\mathbf{y}_{T-1} \mid \mathbf{y}_{T-2}, \ldots, \mathbf{y}_{T-p-1}) \times \\
&\quad \cdots \times f(\mathbf{y}_{p+1} \mid \mathbf{y}_p, \ldots, \mathbf{y}_1) f(\mathbf{y}_p, \ldots, \mathbf{y}_1) \\
&= \left[\prod_{t=p+1}^{T} f(\mathbf{y}_t \mid \mathbf{y}_{t-1}, \ldots, \mathbf{y}_{t-p}) \right] f(\mathbf{y}_p, \ldots, \mathbf{y}_1)
\end{aligned}
$$

と表すことができる。

多変量時系列モデルの性質、推定法、および予測

■ VARモデルの尤度関数

よって対数尤度関数は

$$L(\boldsymbol{\Xi}) = \log f(\mathbf{y}_T, \mathbf{y}_{T-1}, \ldots, \mathbf{y}_1)$$
$$= \sum_{t=p+1}^{T} \log f(\mathbf{y}_t \mid \mathbf{y}_{t-1}, \ldots, \mathbf{y}_{t-p}) + \log f(\mathbf{y}_p, \ldots, \mathbf{y}_1)$$

となる。ここで $\boldsymbol{\Xi}$ は推定するパラメータを 1 つのベクトルにまとめたものである。

次に条件付き密度関数 $f(\mathbf{y}_t \mid \mathbf{y}_{t-1}, \ldots, \mathbf{y}_{t-p})$ を求める。

多変量時系列モデルの性質、推定法、および予測

- 多変量正規分布

$f(\mathbf{y}_t \mid \mathbf{y}_{t-1}, \ldots, \mathbf{y}_{t-p})$ を求めるためには $\boldsymbol{\varepsilon}_t$ $(n \times 1)$に何らかの分布を仮定しなくてはならない。

通常は $\boldsymbol{\varepsilon}_t$ に**多変量(n 変量)正規分布** を仮定する。

一般に、期待値ベクトル $\boldsymbol{\mu}$、分散共分散行列 $\boldsymbol{\Omega}$ の n 変量正規確率ベクトル $\mathbf{x} \sim N(\boldsymbol{\mu}, \boldsymbol{\Omega})$ の密度関数は

$$f(\mathbf{x}) = (2\pi)^{-\frac{n}{2}} \, |\boldsymbol{\Omega}|^{-\frac{1}{2}} \exp\left[-\frac{1}{2} (\mathbf{x}-\boldsymbol{\mu})^{\mathrm{T}} \boldsymbol{\Omega}^{-1} (\mathbf{x}-\boldsymbol{\mu}) \right]$$

によって与えられる。

多変量時系列モデルの性質、推定法、および予測

■ 多変量正規分布

多変量正規分布に関して(1変量正規分布の時と同様)以下の性質が成り立つ。

(1)定数(のベクトル)を足しても分布は正規分布のまま。
つまり $n \times 1$ ベクトル $\mathbf{x}$ が $\mathbf{x} \sim N(\boldsymbol{\mu}, \boldsymbol{\Omega})$ であれば
$\mathbf{y} = \mathbf{x} + \mathbf{c}$ ($\mathbf{c}$: $n \times 1$ の定数ベクトル) の分布は
$\mathbf{y} \sim N(\boldsymbol{\mu} + \mathbf{c}, \boldsymbol{\Omega})$ となる。

(2) 多変量正規確率変数の和はやはり多変量正規分布に従う。つまり ($\mathbf{x}$ と $\mathbf{y}$ は独立であり) $\mathbf{x} \sim N(\boldsymbol{\mu}_x, \boldsymbol{\Omega}_x)$,
$\mathbf{y} \sim N(\boldsymbol{\mu}_y, \boldsymbol{\Omega}_y)$ であれば $\mathbf{z} = \mathbf{x} + \mathbf{y}$ の分布は
$\mathbf{z} \sim N(\boldsymbol{\mu}_x + \boldsymbol{\mu}_y, \boldsymbol{\Omega}_x + \boldsymbol{\Omega}_y)$ となる。

多変量時系列モデルの性質、推定法、および予測

- VARモデルの最尤推定

VARモデルの最尤推定の場合は通常 $\boldsymbol{\varepsilon}_t \sim N(\mathbf{0}, \boldsymbol{\Sigma})$ を仮定する(平均ベクトル $\mathbf{0}$, 分散共分散行列 $\boldsymbol{\Sigma}$) 。この時、$\boldsymbol{\varepsilon}_t$ の密度関数は

$$f(\boldsymbol{\varepsilon}_t) = (2\pi)^{-\frac{n}{2}} |\boldsymbol{\Sigma}|^{-\frac{1}{2}} \exp\left(-\frac{1}{2} \boldsymbol{\varepsilon}_t^{\mathrm{T}} \boldsymbol{\Sigma}^{-1} \boldsymbol{\varepsilon}_t \right)$$

によって与えられる。

多変量時系列モデルの性質、推定法、および予測

- VARモデルの最尤推定

$\mathbf{y}_t$ の $\mathbf{y}_{t-1}, \mathbf{y}_{t-2}, \ldots, \mathbf{y}_{t-p}$ という条件付き密度関数を求めよう。今、

$$\mathbf{y}_t = \mathbf{c} + \mathbf{\Phi}_1 \mathbf{y}_{t-1} + \ldots + \mathbf{\Phi}_p \mathbf{y}_{t-p} + \boldsymbol{\varepsilon}_t$$

において $\boldsymbol{\varepsilon}_t \sim N(\mathbf{0}, \mathbf{\Sigma})$ であり、$\mathbf{y}_{t-1}, \mathbf{y}_{t-2}, \ldots, \mathbf{y}_{t-p}$ は定数扱いできるという事であるから $\mathbf{y}_t$ の $\mathbf{y}_{t-1}, \mathbf{y}_{t-2}, \ldots, \mathbf{y}_{t-p}$ という条件付き分布は多変量正規分布

$$N(\mathbf{c} + \mathbf{\Phi}_1 \mathbf{y}_{t-1} + \ldots + \mathbf{\Phi}_p \mathbf{y}_{t-p}, \mathbf{\Sigma})$$

となる。よってその密度関数は

多変量時系列モデルの性質、推定法、および予測

- VARモデルの最尤推定

$$f(\mathbf{y}_t \mid \mathbf{y}_{t-1},\dots,\mathbf{y}_{t-p}) = (2\pi)^{-\frac{n}{2}} \mid \boldsymbol{\Sigma} \mid^{-\frac{1}{2}} \exp\left(-\frac{1}{2}\mathbf{e}_t^{\mathrm{T}} \boldsymbol{\Sigma}^{-1}\mathbf{e}_t \right)$$

によって与えられる。ここで

$$\mathbf{e}_t = \mathbf{y}_t - \mathbf{c} - \boldsymbol{\Phi}_1\mathbf{y}_{t-1} - \cdots - \boldsymbol{\Phi}_p\mathbf{y}_{t-p}$$

である。

多変量時系列モデルの性質、推定法、および予測

- VARモデルの条件付き最尤推定

AR(p)モデルの時と同様に先ほどの対数尤度関数から $\log f(\mathbf{y}_p, \ldots, \mathbf{y}_1)$ を除いた条件付き対数尤度関数

$$L^C(\boldsymbol{\Xi}) = \sum_{t=1}^{T} \log f(\mathbf{y}_t \mid \mathbf{y}_{t-1}, \ldots, \mathbf{y}_{t-p})$$

をパラメーター $\mathbf{c}, \boldsymbol{\Phi}_1, \ldots, \boldsymbol{\Phi}_p, \boldsymbol{\Sigma}$ について最大化する事によって推定する事を**条件付き最尤推定法**と呼ぶ。

条件付き推定は前述の正確な最尤推定と漸近的な性質は同じになる(つまり標本数が多ければどちらを使っても推定効率あまり変わらない)。

多変量時系列モデルの性質、推定法、および予測

- VARモデルの最小二乗推定

係数 $\mathbf{\Phi}_k$ の(i, j)成分を $\phi_{ij}^{(k)}$ と表すと、VAR(p)モデルは n 本の回帰式

$$y_{1t} = c_1 + \sum_{i=1}^{n} \phi_{1i}^{(1)} y_{i,t-1} + \cdots + \sum_{i=1}^{n} \phi_{1i}^{(p)} y_{i,t-p} + \varepsilon_{1t}$$

$$y_{2t} = c_2 + \sum_{i=1}^{n} \phi_{2i}^{(1)} y_{i,t-1} + \cdots + \sum_{i=1}^{n} \phi_{2i}^{(p)} y_{i,t-p} + \varepsilon_{2t}$$

$$\vdots$$

$$y_{nt} = c_n + \sum_{i=1}^{n} \phi_{ni}^{(1)} y_{i,t-1} + \cdots + \sum_{i=1}^{n} \phi_{ni}^{(p)} y_{i,t-p} + \varepsilon_{nt}$$

より成り立つ。この 1 つ 1 つの式をOLSで推定する。

多変量時系列モデルの性質、推定法、および予測

- (条件付き)最尤法と最小二乗法の関係

VARモデルの係数 $\mathbf{\Phi}_k$ の条件付き最尤法による推定値と最少二乗法による推定値は**同じになる** (AR(p)モデルの時も同様の関係があった事を思いだそう)。

よってより簡便な最小二乗法による推定法がよく用いられる。

多変量時系列モデルの性質、推定法、および予測

- VARモデルによる1期先予測

以下では誤差項 ε_t は独立であると仮定する。VAR(p)モデルによる予測はAR(p)モデルによる予測と考え方は同じである。$\mathbf{y}_{t+1}$は

$$\mathbf{y}_{t+1} = \mathbf{c} + \mathbf{\Phi}_1\mathbf{y}_t + \mathbf{\Phi}_2\mathbf{y}_{t-1} + \ldots + \mathbf{\Phi}_p\mathbf{y}_{t-p+1} + \boldsymbol{\varepsilon}_{t+1}$$

であるので、情報集合 $\Omega_t = (\mathbf{y}_t, \mathbf{y}_{t-1}, \ldots, \mathbf{y}_1)$ が与えられた時の $\mathbf{y}_{t+1}$ の(最適)予測は

$$\hat{\mathbf{y}}_{t+1|t} = E(\mathbf{y}_{t+1} \mid \Omega_t) = \mathbf{c} + \mathbf{\Phi}_1\mathbf{y}_t + \mathbf{\Phi}_2\mathbf{y}_{t-1} + \cdots + \mathbf{\Phi}_p\mathbf{y}_{t-p+1}$$

となる。ここで $\hat{\mathbf{y}}_{t+1|t} = (\hat{y}_{1,t+1|t}, \hat{y}_{2,t+1|t} \ldots, \hat{y}_{n,t+1|t})^{\mathrm{T}}$ とする。

多変量時系列モデルの性質、推定法、および予測

- VARモデルによる1期先予測のMSE

第 i 番目の成分 y_{it} の1期先予測 $\hat{y}_{i,t+1|t}$ のMSEは

$$E[(\mathbf{y}_{t+1} - \hat{\mathbf{y}}_{t+1|t})(\mathbf{y}_{t+1} - \hat{\mathbf{y}}_{t+1|t})^{\mathrm{T}} \mid \Omega_t] = E(\boldsymbol{\varepsilon}_{t+1}\boldsymbol{\varepsilon}_{t+1}^{\mathrm{T}} \mid \Omega_t) = \boldsymbol{\Sigma}$$

であるので $\boldsymbol{\Sigma}$ の第 i 番目の対角成分となる。

- VARモデルによる h 期先予測

AR(p)モデルの時と同様、逐次予測で予測できる。
ただしMSEの計算はAR(p)モデルの時と同様に難しい。

宿題 (提出する必要はありません)

VAR(1)モデル:

$$\mathbf{y}_t = \mathbf{c} + \mathbf{\Phi}_1\mathbf{y}_{t-1} + \boldsymbol{\varepsilon}_t, \quad \boldsymbol{\varepsilon}_t \sim \text{i.i.d.}N(\mathbf{0}, \mathbf{\Sigma})$$

による、$\Omega_t = (\mathbf{y}_t^{\mathrm{T}}, \mathbf{y}_{t-1}^{\mathrm{T}}, \ldots, \mathbf{y}_1^{\mathrm{T}})^{\mathrm{T}}$ のもとでの h 期先最適予測は

$$\hat{\mathbf{y}}_{t+h|t} = (\mathbf{I}_n - \mathbf{\Phi}_1)^{-1}(\mathbf{I}_n - \mathbf{\Phi}_1^h)\mathbf{c} + \mathbf{\Phi}_1^h\mathbf{y}_t$$

で与えられ、またこの第 i 番目の成分のMSEは

$$\sum_{i=0}^{h-1}\mathbf{\Phi}_1^i\mathbf{\Sigma}(\mathbf{\Phi}_1^{\mathrm{T}})^i = \mathbf{\Sigma} + \mathbf{\Phi}_1\mathbf{\Sigma}\mathbf{\Phi}_1^{\mathrm{T}} + \cdots + \mathbf{\Phi}_1^{h-1}\mathbf{\Sigma}(\mathbf{\Phi}_1^{\mathrm{T}})^{h-1}$$

の第 i 番目の対角成分で与えられる事を示せ。

多変量時系列モデルの性質、推定法、および予測

- (例)VARモデルによる予測

以下はVAR(1)モデル:

$$\begin{bmatrix} y_{1t} \\ y_{2t} \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix} + \begin{bmatrix} 0.8 & 0.1 \\ 0.1 & 0.8 \end{bmatrix} \begin{bmatrix} y_{1t-1} \\ y_{2t-1} \end{bmatrix} + \begin{bmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{bmatrix},$$

$$\begin{bmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{bmatrix} \sim \mathrm{W.N.}(\mathbf{\Sigma}),\ \mathbf{\Sigma} = \begin{bmatrix} 1 & 0.2 \\ 0.2 & 1 \end{bmatrix}$$

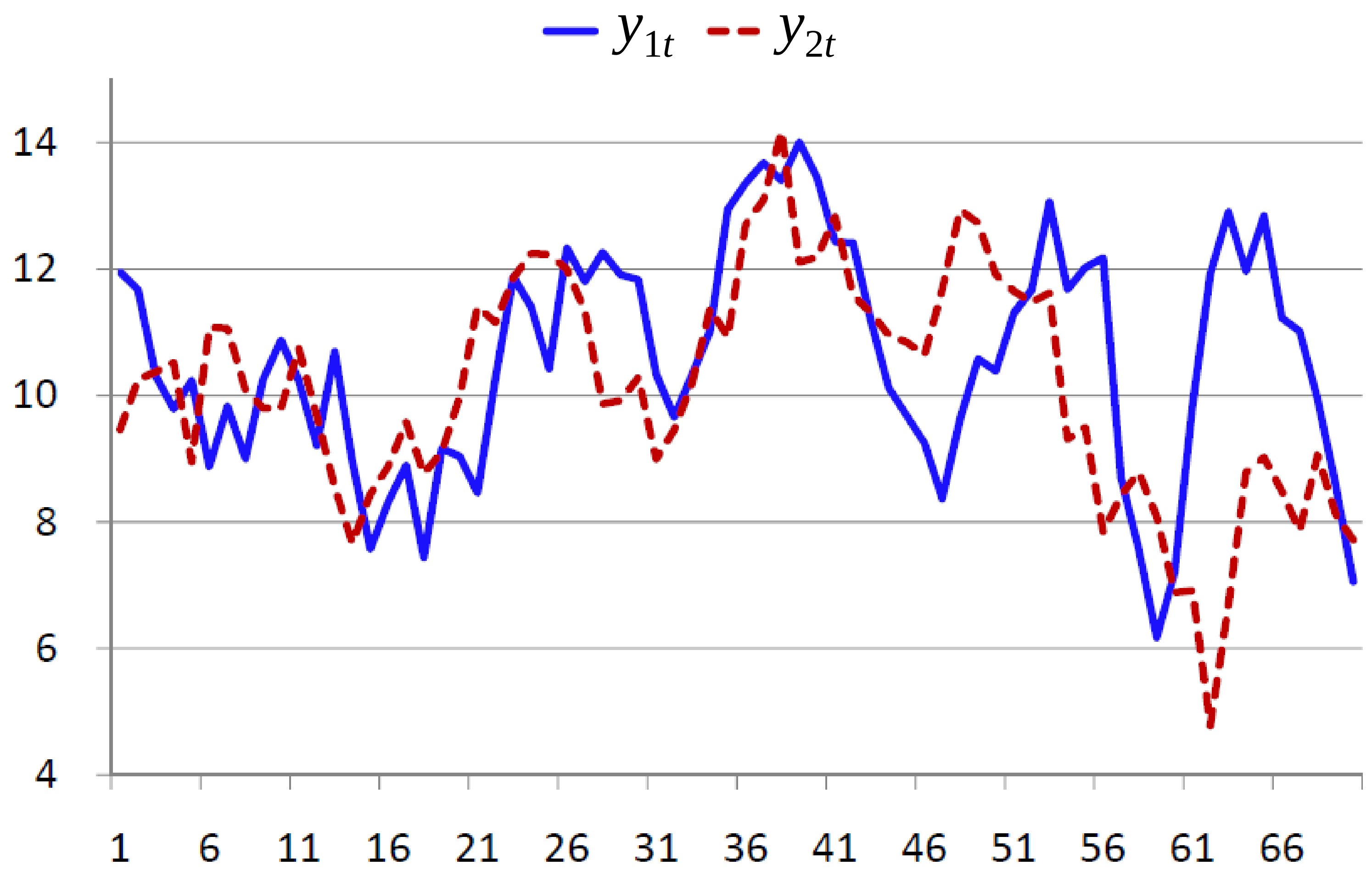
より T=70 のデータを発生させ、

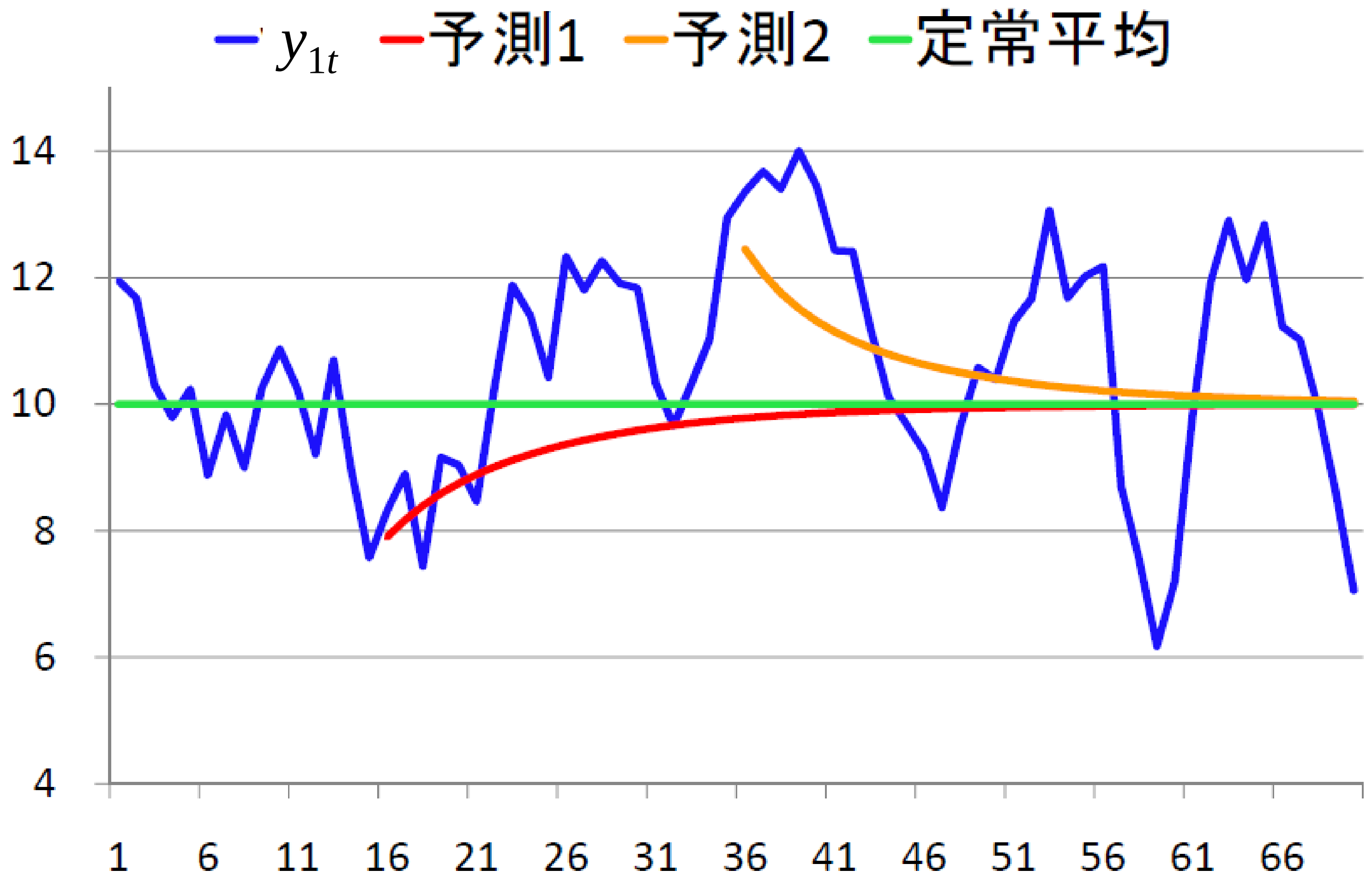
① y_{1t} と y_{2t} をプロットしたもの

② t =15 より y_{1t} の h =1~55期先最適予測 $\hat{y}_{1,15+h|15}$ をプロットしたもの(予測 1)

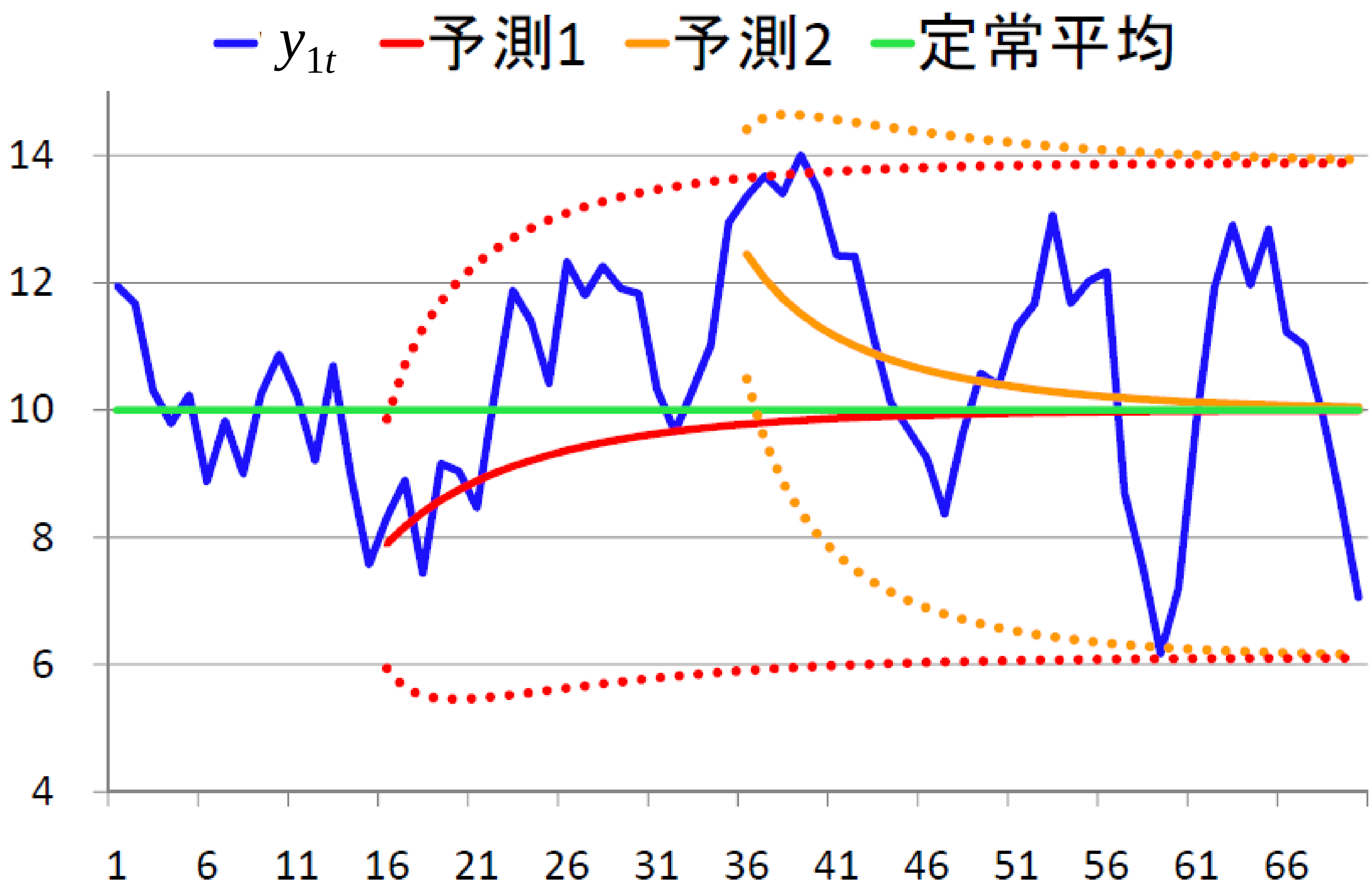
③ t =35 より y_{1t} の h =1~35期先最適予測 $\hat{y}_{1,35+h|35}$ をプロットしたもの(予測 2)

である。


y_{1t}
y_{2t}
14
12
10
8
6
4
1
6
11
16
21
26
31
36
41
46
51
56
61
66



y_{1t}
予測1
予測2
定常平均
14
12
10
8
6
4
1
6
11
16
21
26
31
36
41
46
51
56
61
66



y_{1t}
予測1
予測2
定常平均
14
12
10
8
6
4
1
6
11
16
21
26
31
36
41
46
51
56
61
66