第 14 回　分布ラグと MIDAS モデル

村澤 康友

2023 年 1 月 17 日

今日のポイント

1. 経済時系列は様々な頻度で観測される．観測頻度が異なる時系列を含む多変量時系列を混合頻度時系列という．
2. $\{x_t\}$ を所与とした $\{y_t\}$ の p 次の分布ラグモデルは，任意の t について $y_t = \alpha + \beta(\mathrm{L})x_t + u_t$，ただし $\beta(\mathrm{L}) := \beta_0 + \beta_1\mathrm{L} + \cdots + \beta_p\mathrm{L}^p$ で $\{u_t\}$ は平均 0 の I(0)．p が大きい場合は分布ラグ $\beta(\mathrm{L})$ の形状を制約する．
3. 混合頻度時系列を直接的に定式化するモデルを混合データ抽出（MIDAS）モデルという．$\{y_t\}$ を低頻度系列，$\{x_t\}$ を $\{y_t\}$ の S 倍の頻度で観測される高頻度系列とすると，p 次の MIDAS 回帰モデルは，任意の t について $\mathrm{E}(y_t|\boldsymbol{X}_t) = \alpha + \beta\left(\mathrm{L}^{1/S}\right)x_t$，ただし $\beta\left(\mathrm{L}^{1/S}\right) := \beta_0 + \beta_1\mathrm{L}^{1/S} + \cdots + \beta_p\mathrm{L}^{p/S}$．

目次

1 混合頻度時系列

1.1 経済時系列の頻度

経済時系列は様々な頻度で観測される．

年次　国民経済計算（SNA）

四半期　四半期別 GDP 速報（QE）

月次　各種経済指標（失業率・生産指数・物価指数など）

日次　各種金融指標（金利・株価・為替レートなど）

複数の時系列を扱う場合，通常は最も低頻度の時系列に頻度を揃える（低頻度の時系列を補間して高頻度に揃える場合もある）．

例 1. 実質 GDP と失業率の関係（オークンの法則）の分析なら四半期に揃える．

1.2 超短期予測（ナウキャスト）

実質 GDP は四半期系列であり，かつ公表が遅い．直近の実質 GDP の予測には，足元の月次系列の情報が役立つ．

定義 1. 直近の予測を**超短期予測（ナウキャスト）**という．

注 1. 「ナウキャスト」は気象学（天気予報）の用語だが，最近は経済予測でも使われる．

定義 2. 高頻度系列を低頻度に揃えて低頻度系列を予測する式を**ブリッジ方程式**という．

注 2. まず高頻度系列の予測値を作成して低頻度に変換し，それを用いてブリッジ方程式で低頻度系列を予測する．

1.3 混合頻度時系列

定義 3. 観測頻度が異なる時系列を含む多変量時系列を**混合頻度時系列**という．

注 3. 混合頻度時系列の分析には 2 つのアプローチがある．

MIDAS モデル 低頻度系列と高頻度系列の関係を分布ラグモデルで定式化し，（非線形）最小 2 乗法で係数を推定する．

潜在変数モデル 低頻度系列を欠損値をもつ高頻度系列とみなして状態空間モデルを定式化し，カルマン・フィルターで係数と欠損値を推定する．

2 分布ラグ

2.1 分布ラグモデル

$\{x_t\}, \{y_t\}$ を I(0) とする．$\{x_t\}$ から $\{y_t\}$ を予測したい．

定義 4. $\{x_t\}$ を所与とした $\{y_t\}$ の p **次の分布ラグモデル**は，任意の t について

$$y_t = \alpha + \beta(\mathrm{L})x_t + u_t$$

ただし $\beta(\mathrm{L}) := \beta_0 + \beta_1\mathrm{L} + \cdots + \beta_p\mathrm{L}^p$ で $\{u_t\}$ は平均 0 の I(0).

注 4. p が無限なら $|\beta(1)| < \infty$ が必要.

注 5. ラグ多項式を使わずに書くと，任意の t について

$$y_t = \alpha + \sum_{s=0}^{p} \beta_s x_{t-s} + u_t$$

$(\alpha, \beta_0, \ldots, \beta_p)$ は OLS で推定できる．ただし p が大きいと推定値が不安定になる．

注 6. $w(\mathrm{L}) := \beta(\mathrm{L})/\beta(1)$ と正規化すると分布ラグの形状が分かりやすい．すなわち任意の t について

$$y_t = \alpha + \beta w(\mathrm{L})x_t + u_t$$

ただし $\beta := \beta(1)$.

2.2 コイック・ラグ

$\{y_t\}$ に AR(1) の動学モデルを仮定する．すなわち任意の t について

$$y_t = \alpha + \phi y_{t-1} + \beta x_t + u_t$$

ただし $|\phi| < 1$ で $\{u_t\}$ は平均 0 の I(0)．(α, ϕ, β) は OLS で推定できる．

定理 1. 任意の t について

$$y_t = \frac{\alpha}{1-\phi} + \beta \sum_{s=0}^{\infty} \phi^s x_{t-s} + \sum_{s=0}^{\infty} \phi^s u_{t-s}$$

証明. 逐次代入より任意の t について

$$\begin{aligned}
y_t &= \alpha + \phi(\alpha + \phi y_{t-2} + \beta x_{t-1} + u_{t-1}) \\
&\quad + \beta x_t + u_t \\
&= \ldots \\
&= \alpha(1 + \phi + \cdots) + \beta(x_t + \phi x_{t-1} + \cdots) \\
&\quad + u_t + \phi u_{t-1} + \cdots \\
&= \frac{\alpha}{1-\phi} + \beta \sum_{s=0}^{\infty} \phi^s x_{t-s} + \sum_{s=0}^{\infty} \phi^s u_{t-s}
\end{aligned}$$

□

注 7. すなわち幾何級数の分布ラグが得られる．

定義 5. 幾何級数の分布ラグを**コイック・ラグ**という．

2.3 アーモン・ラグ

p を有限とする．p が大きい場合，$\{\beta_s\}$ を低次の多項式で表現すれば，推定する係数を減らせる．

定義 6. m **次のアーモン・ラグ**は，$s = 0, \ldots, p$ について

$$\beta_s := \gamma_0 + \gamma_1 s + \cdots + \gamma_m s^m$$

ただし $m < p$.

定理 2. m 次のアーモン・ラグをもつ p 次の分布ラグモデルは，任意の t について

$$\begin{aligned}y_t = \alpha + \gamma_0 \sum_{s=0}^{p} x_{t-s} + \gamma_1 \sum_{s=1}^{p} s x_{t-s} + \cdots \\ + \gamma_m \sum_{s=1}^{p} s^m x_{t-s} + u_t\end{aligned}$$

証明. 代入して式変形すると，任意の t について

$$\begin{aligned}y_t &= \alpha + \sum_{s=0}^{p} \left(\gamma_0 + \gamma_1 s + \cdots + \gamma_m s^m\right) x_{t-s} + u_t \\ &= \alpha + \gamma_0 \sum_{s=0}^{p} x_{t-s} + \gamma_1 \sum_{s=0}^{p} s x_{t-s} + \cdots \\ &\quad + \gamma_m \sum_{s=0}^{p} s^m x_{t-s} + u_t \\ &= \alpha + \gamma_0 \sum_{s=0}^{p} x_{t-s} + \gamma_1 \sum_{s=1}^{p} s x_{t-s} + \cdots \\ &\quad + \gamma_m \sum_{s=1}^{p} s^m x_{t-s} + u_t\end{aligned}$$

□

注 8. $(\alpha, \gamma_0, \ldots, \gamma_m)$ は OLS で推定できる．

2.4 正規化指数アーモン・ラグ

$\{\beta_s\}$ が非負なら多項式を指数変換する．すなわち $s = 0, \ldots, p$ について

$$\begin{aligned}\beta_s &:= \exp(\gamma_0 + \gamma_1 s + \cdots + \gamma_m s^m) \\ &= \exp(\gamma_0) \exp(\gamma_1 s + \cdots + \gamma_m s^m)\end{aligned}$$

正規化すると，$s = 0, \ldots, p$ について

$$\begin{aligned}w_s &:= \frac{\beta_s}{\sum_{r=0}^{p} \beta_r} \\ &= \frac{\exp(\gamma_1 s + \cdots + \gamma_m s^m)}{\sum_{r=0}^{p} \exp(\gamma_1 r + \cdots + \gamma_m r^m)}\end{aligned}$$

定義 7. m **次の正規化指数アーモン・ラグ**は，$s = 0, \ldots, p$ について

$$w_s := \frac{\exp(\gamma_1 s + \cdots + \gamma_m s^m)}{\sum_{r=0}^{p} \exp(\gamma_1 r + \cdots + \gamma_m r^m)}$$

ただし $m < p$.

注 9. したがって任意の t について

$$\begin{aligned}y_t = \alpha + \beta \sum_{s=0}^{p} \frac{\exp(\gamma_1 s + \cdots + \gamma_m s^m)}{\sum_{r=0}^{p} \exp(\gamma_1 r + \cdots + \gamma_m r^m)} x_{t-s} \\ + u_t\end{aligned}$$

$(\alpha, \beta, \gamma_1, \ldots, \gamma_m)$ は非線形最小 2 乗法で推定できる．

3 MIDAS

3.1 MIDAS モデル

$\{y_t\}$ を低頻度系列，$\{x_t\}$ を $\{y_t\}$ の S 倍の頻度で観測される高頻度系列とする．すなわち任意の t について $x_t, x_{t-1/S}, \ldots, x_{t-(S-1)/S}$ を観測する．（任意の t について $x_t, x_{t+1/S}, \ldots, x_{t+(S-1)/S}$ を観測するとみなす場合もある．）

例 2. $\{y_t\}$ が四半期，$\{x_t\}$ が月次なら，例えば第 1 四半期は 1 月に $x_{1/3}$，2 月に $x_{2/3}$，3 月に (x_1, y_1) を観測する．

定義 8. 混合頻度時系列を直接的に定式化するモデルを**混合データ抽出**（*mixed-data sampling, MIDAS*）**モデル**という．

注 10. 潜在変数モデルは間接的に定式化する．

3.2 MIDAS 回帰モデル

高頻度系列 $\{x_t\}$ から低頻度系列 $\{y_t\}$ を予測したい．時点 t までの $\{x_t\}$ の観測値を行列で表すと，任意の t について

$$\boldsymbol{X}_t := \begin{bmatrix} x_1 & \ldots & x_{1-(S-1)/S} \\ & \vdots & \\ x_t & \ldots & x_{t-(S-1)/S} \end{bmatrix}$$

定義 9. y_t **の** $\boldsymbol{X}_t$ **上への** p **次の** *MIDAS* **回帰モデル**は，任意の t について

$$\mathrm{E}(y_t | \boldsymbol{X}_t) = \alpha + \beta \left(\mathrm{L}^{1/S}\right) x_t$$

ただし $\beta\left(\mathrm{L}^{1/S}\right) := \beta_0 + \beta_1 \mathrm{L}^{1/S} + \cdots + \beta_p \mathrm{L}^{p/S}$.

注 11. $w(\mathrm{L}) := \beta(\mathrm{L})/\beta(1)$ と正規化すると，任意の t について

$$\mathrm{E}(y_t | \boldsymbol{X}_t) = \alpha + \beta w \left(\mathrm{L}^{1/S}\right) x_t$$

ただし $\beta := \beta(1)$．ラグ多項式を使わずに書くと，任意の t について

$$\mathrm{E}(y_t|\boldsymbol{X}_t) = \alpha + \beta \sum_{s=0}^{p} w_s x_{t-s/S}$$

注 12．係数が多い場合は分布ラグの形状を制約する．右辺に y_{t-1} を追加する場合もある（コイック・ラグ）．

3.3 U-MIDAS モデル

定義 10. 分布ラグの形状を制約しない MIDAS モデルを**制約なしの** *MIDAS*（*unrestricted MIDAS, U-MIDAS*）**モデル**という．

注 13．係数が少なければ U-MIDAS モデルでよい．

例 3. アメリカの実質 GDP（四半期）と鉱工業生産指数（月次）の対数階差系列に関する分布ラグの MIDAS 回帰による推定結果：U-MIDAS（図 1）と 2 次の正規化指数アーモン・ラグ（図 2）．[*1]

4 今日のキーワード

超短期予測（ナウキャスト），ブリッジ方程式，混合頻度時系列，分布ラグモデル，コイック・ラグ，アーモン・ラグ，正規化指数アーモン・ラグ，混合データ抽出（MIDAS）モデル，MIDAS 回帰モデル，制約なしの MIDAS（U-MIDAS）モデル

5 次回までの準備

復習 復習テスト 14

提出 宿題 14，復習テスト 9〜14

試験 (1) 教科書を読む (2) 用語の定義を覚える (3) 復習テストを自力で解く (4) 過去問に挑戦

[*1] gretl は期首に低頻度系列を観測すると想定している．例えば第 1 四半期は 1 月に (x_1, y_1)，2 月に $x_{1+1/3}$，3 月に $x_{1+2/3}$ を観測する．そのため期末に低頻度系列を観測する場合（例えばフロー変数），分析の際に時点をずらす必要がある．

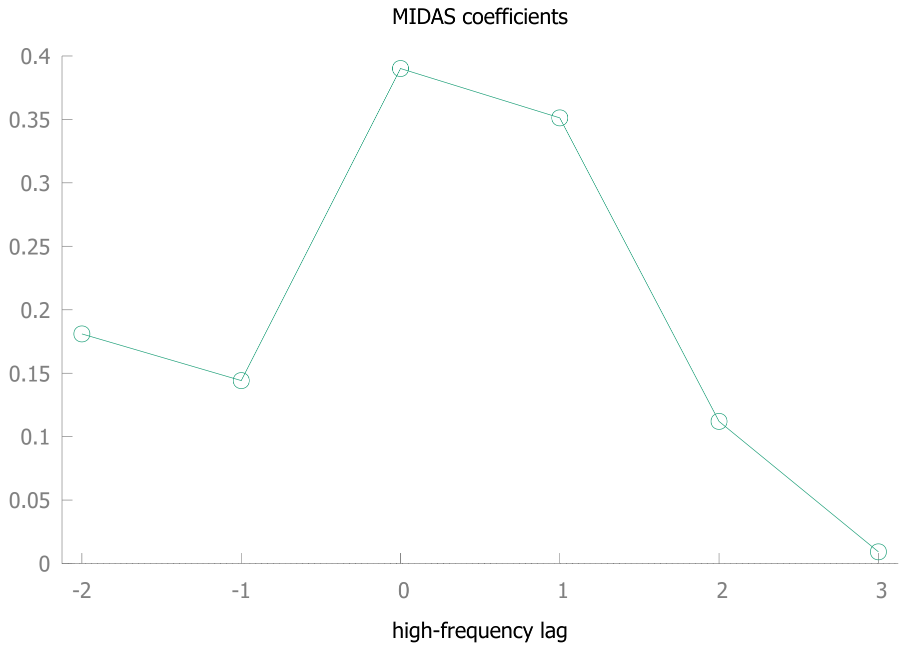

図 1　実質 GDP（四半期）と鉱工業生産指数（月次）の対数階差系列に関する分布ラグ（U-MIDAS）

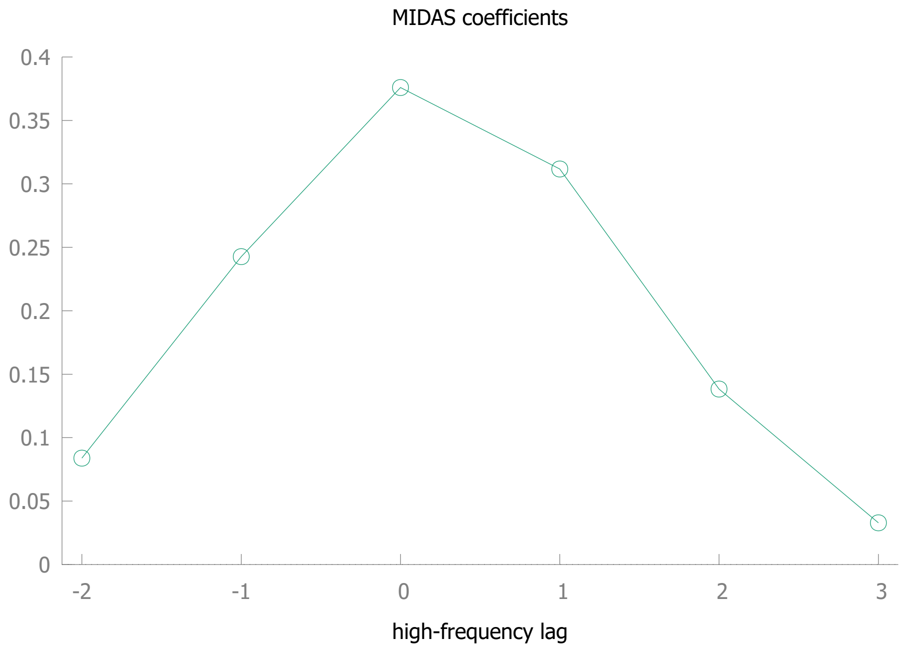

図 2　実質 GDP（四半期）と鉱工業生産指数（月次）の対数階差系列に関する分布ラグ（2 次の正規化指数アーモン・ラグ）