

arima - 自己回帰移動平均モデル 【 評価版 】

arima は擾乱が自己回帰移動平均 (ARMA: autoregressive moving-average) 過程として表現される単変量モデルのフィットを行います。arima は従属変数の自己回帰項のみによって記述された ARMA モデルのみならず、独立変数を含む形の ARMAX モデルにも対応しています。

1. ARMA 過程	
2. ARIMA モデル	Example 1
	Example 2
3. 乗法的季節変動モデル	Example 3
4. ARMAX モデル	Example 4
5. 動的予測	
補足 1	
補足 2	

1. ARMA 過程

次に示すのは 1 次の ARMA 過程 ARMA(1,1) のモデル式です。

$$y_t = \mathbf{x}_t \boldsymbol{\beta} + \mu_t \quad (1a)$$

$$\mu_t = \rho \mu_{t-1} + \theta \epsilon_{t-1} + \epsilon_t \quad (1b)$$

(1a) 式は構造方程式 (structural equation) と呼ばれます。これに対し (1b) 式は擾乱 (disturbance) の特性を規定する数式です。ただし $\epsilon_t \sim i.i.d. N(0, \sigma^2)$ 、すなわち ϵ_t は白色ノイズ (white-noise) と呼べる擾乱であるとします。

このモデル式の本質は擾乱項 μ_t の中に自己相関 (autoregressive) 項 $\rho \mu_{t-1}$ と移動平均 (moving-average) 項 $\theta \epsilon_{t-1}$ を含む点です。すなわち周期 t における擾乱は単なる白色ノイズ ϵ_t 以外に 1 期前の μ_t, ϵ_t の影響を引きずった構造となっています。その影響の度合いはパラメータ ρ と θ によって規定されるわけです。

ARMA(1,1) のモデル式を ARMA(p, q) に一般化したときの擾乱は

$$\mu_t = \rho_1 \mu_{t-1} + \cdots + \rho_p \mu_{t-p} + \theta_1 \epsilon_{t-1} + \cdots + \theta_q \epsilon_{t-q} + \epsilon_t$$

のように記述できるわけですが、これに $\mu_t = y_t - \mathbf{x}_t \boldsymbol{\beta}$ という式を代入すると

$$y_t = \mathbf{x}_t\boldsymbol{\beta} + \rho_1(y_{t-1} - \mathbf{x}_{t-1}\boldsymbol{\beta}) + \rho_2(y_{t-2} - \mathbf{x}_{t-2}\boldsymbol{\beta}) + \cdots + \rho_p(y_{t-p} - \mathbf{x}_{t-p}\boldsymbol{\beta}) + \theta_1\epsilon_{t-1} + \cdots + \theta_q\epsilon_{t-q} + \epsilon_t \quad (2)$$

という ARMA(p, q) のモデル式が得られます。これに対しラグ演算子 L を用いた数式を

$$\rho(L^p) = 1 - \rho_1L - \rho_2L^2 - \cdots - \rho_pL^p \quad (3a)$$

$$\theta(L^q) = 1 + \theta_1L + \theta_2L^2 + \cdots + \theta_qL^q \quad (3b)$$

のように定義すると (ただし $L^j y_t = y_{t-j}$ を意味する) ARMA(p, q) は

$$\rho(L^p)(y_t - \mathbf{x}_t\boldsymbol{\beta}) = \theta(L^q)\epsilon_t \quad (4)$$

という簡単な式で表現できることになります。

ARMA 過程と言った場合、従属変数のみからなる、すなわち独立変数 \mathbf{x}_t を含まないモデルを対象にすることが多いわけですが、ここでは ARMA(1, 1) 過程と ARMA(2, 2) 過程について具体的なモデル式を誘導しておきます。

(1) ARMA(1, 1) モデル

ARMA(p, q) 過程の一般式 (4) より

$$(1 - \rho_1L)(y_t - \beta_0) = (1 + \theta_1L)\epsilon_t$$

ラグ演算子を適用すると

$$y_t - \beta_0 - \rho_1(y_{t-1} - \beta_0) = \epsilon_t + \theta_1\epsilon_{t-1}$$

従って

$$y_t - \beta_0 = \rho_1(y_{t-1} - \beta_0) + \theta_1\epsilon_{t-1} + \epsilon_t \quad (5)$$

というのが ARMA(1, 1) 過程のモデル式となります。

(2) ARMA(2, 2) モデル

同様に ARMA(2, 2) 過程のモデル式は

$$y_t - \beta_0 = \rho_1(y_{t-1} - \beta_0) + \rho_2(y_{t-2} - \beta_0) + \theta_1\epsilon_{t-1} + \theta_2\epsilon_{t-2} + \epsilon_t \quad (6)$$

となります。



用例の中には ARIMA(p, d, q) 過程という表現が出てきます (I は integrated の略) が、それは d 階差分の時系列が ARMA(p, q) 過程に従うことを意味します。



ARIMA モデルのモデル式に関する補足については補足 2 を参照ください。

2. ARIMA モデル

▷ Example 1: ARIMA(1,1,1) モデル

[TS] `arima` の Example 1 には Example データセット `wpi1.dta` を使用した ARIMA(1, 1, 1) モデルの用例が紹介されています。

```
. use http://www.stata-press.com/data/r17/wpi1.dta *1
```

このデータセット中には 1960q1 から 1990q4 に至る期間中のデータが四半期ごとに記録されています。

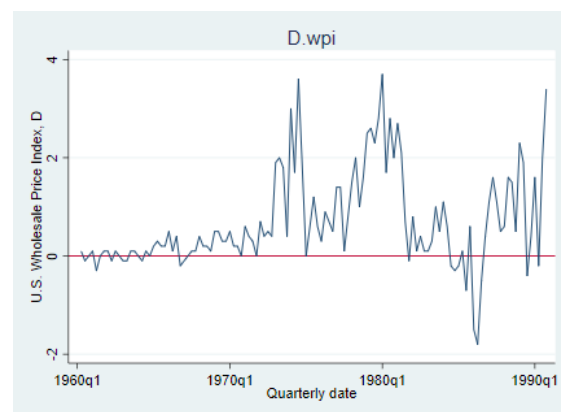
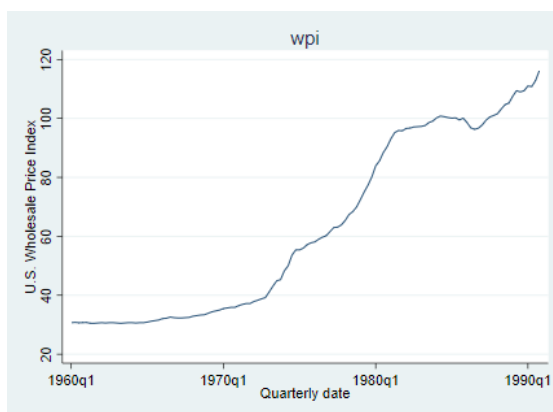
```
. list wpi t if _n <= 4 | _n >= (_N - 3), separator(4) *2
```

	wpi	t
1.	30.7	1960q1
2.	30.8	1960q2
3.	30.7	1960q3
4.	30.7	1960q4
121.	111	1990q1
122.	110.8	1990q2
123.	112.8	1990q3
124.	116.2	1990q4

次に示すのは `wpi` と `D.wpi` (`wpi` の 1 階差分) についてのプロットです。

```
. twoway (line wpi t), title(wpi) *3
```

```
. twoway (line D.wpi t), yline(0) title(D.wpi)
```



*1 メニュー操作 : File ▷ Example Datasets ▷ Stata 17 manual datasets と操作、Time-Series Reference Manual [TS] の `arima` の項よりダウンロードする。

*2 メニュー操作 : Data ▷ Describe data ▷ List data

*3 メニュー操作 : Graphics ▷ Twoway graph (scatter, line, etc.) 詳細については補足 1 を参照。

このグラフから明らかなように原系列 `wpi` は定常とは言えないので、ここでは階差系列 `D.wpi` を解析対象とします。

ARMA モデルを `arima` に対し指定する場合、`arima(p,d,q)` という略式指定を用いる方法と `ar()`, `ma()` 指定による方法の 2 種類が選択できます。モデルに含めるべきラグ次数が AR 項については 1 から p 、MA 項については 1 から q と連続している場合には略式指定を用いることができます。これに対し指定すべきラグ次数が 1 と 4 といった形で不連続な場合には、`numlist` が指定できる `ar()`, `ma()` インタフェースを用いることになります。

なお、`arima(p,d,q)` インタフェースを用いた場合には差分の次数 d を明示することができます。その場合、例えば `arima(1,1,1)` という形でモデルを規定したとするなら、従属変数として指定するのは `D.wpi` ではなく `wpi` となる点に注意してください。

最初に `arima(p,d,q)` インタフェースを用いて `arima` の実行を行ってみます。この場合、従属変数としては `wpi` を指定します。

- Statistics ▷ Time series ▷ ARIMA and ARMAX models と操作
- Model タブ: Dependent variable: `wpi`
ARIMA(p,d,q) specification: $p = d = q = 1$

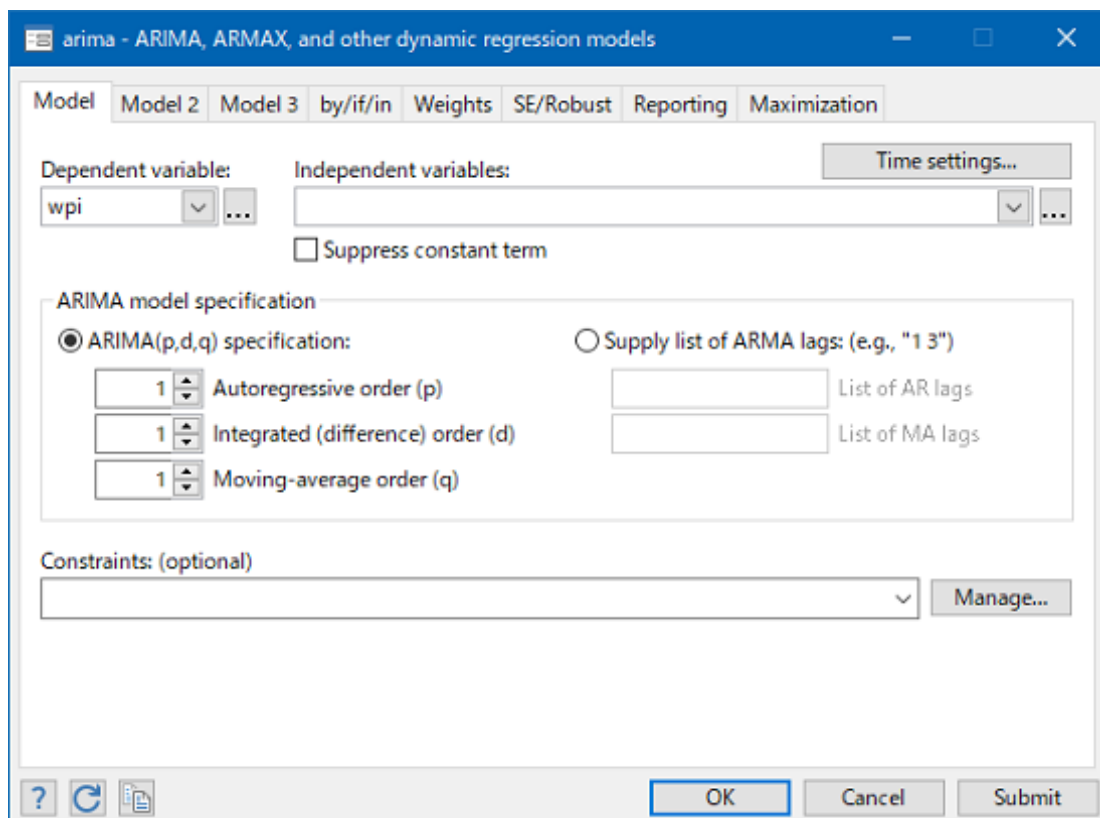


図 1 arima ダイアログ – Model タブ

```

. arima wpi, arima(1,1,1)

(setting optimization to BHHH)
Iteration 0:  log likelihood = -139.80133
Iteration 1:  log likelihood = -135.6278
Iteration 2:  log likelihood = -135.41838
Iteration 3:  log likelihood = -135.36691
Iteration 4:  log likelihood = -135.35892
(swapping optimization to BFGS)
Iteration 5:  log likelihood = -135.35471
Iteration 6:  log likelihood = -135.35135
Iteration 7:  log likelihood = -135.35132
Iteration 8:  log likelihood = -135.35131

ARIMA regression

Sample: 1960q2 thru 1990q4                Number of obs   =       123
                                           Wald chi2(2)    =       310.64
Log likelihood = -135.3513                Prob > chi2     =       0.0000

```

D.wpi	OPG				
	Coefficient	std. err.	z	P> z	[95% conf. interval]
wpi					
_cons	.7498197	.3340968	2.24	0.025	.0950019 1.404637
ARMA					
ar					
L1.	.8742288	.0545435	16.03	0.000	.7673256 .981132
ma					
L1.	-.4120458	.1000284	-4.12	0.000	-.6080979 -.2159938
/sigma	.7250436	.0368065	19.70	0.000	.6529042 .7971829

Note: The test of the variance against zero is one sided, and the two-sided confidence interval is truncated at zero.

結果が意味するところは (5) 式より次のようになります。

$$\Delta wpi_t - 0.750 = 0.874 \cdot (\Delta wpi_{t-1} - 0.750) - 0.412 \cdot \epsilon_{t-1} + \epsilon_t$$

なお sigma として 0.725 とレポートされていますがこれはホワイトノイズ擾乱 ϵ の標準偏差推定値を意味します。

評価版では割愛しています。

▷ Example 2: 加法的季節変動モデル

評価版では割愛しています。

3. 乗法的季節変動モデル

評価版では割愛しています。

4. ARMAX モデル

評価版では割愛しています。

5. 動的予測

評価版では割愛しています。

補足 1 – グラフ作成コマンド操作

評価版では割愛しています。

補足 2 – ARIMA のモデル式

評価版では割愛しています。

