

7章：非線形時系列モデル

ここでは7章6節と9節の内容を再現しよう。具体的には、非線形性の検定、TARモデル、LSTARモデル、ESTARモデルの推定である。

1. 非線形性の検定

教科書の7章6節の例（失業率）を用いて、非線形性に関する検定をしよう。まず、HPからUNRATE.XLSダウンロードしWorkfileに読み込もう。このデータは米国の月次失業率（1960年1月～2013年6月）である。また、失業率unrateの階差をdurateとして定義する。

```
genr durate =d(unrate)
```

そして、「QUICK」→「Estimate Equation」を選択し、

```
durate c durate(-1) durate(-2) durate(-3) durate(-4) durate(-12)
```

と入力する(ラグは1から4まで連続しているので、durate c durate(-1 to -4) durate(-12)としてもよい)。OLSで推定するため、推定方法はLSを選択しよう。そしてOKをクリックすると、以下の結果が得られる。

Dependent Variable: DURATE				
Method: Least Squares				
Date: 11/27/17 Time: 11:20				
Sample (adjusted): 1961M02 2013M06				
Included observations: 629 after adjustments				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.000549	0.006166	0.088961	0.9291
DURATE(-1)	0.058121	0.039374	1.476101	0.1404
DURATE(-2)	0.227569	0.038673	5.884396	0.0000
DURATE(-3)	0.188357	0.038647	4.873758	0.0000
DURATE(-4)	0.140142	0.039080	3.586017	0.0004
DURATE(-12)	-0.127695	0.035694	-3.577481	0.0004
R-squared	0.180601	Mean dependent var	0.001472	
Adjusted R-squared	0.174025	S.D. dependent var	0.170068	
S.E. of regression	0.154563	Akaike info criterion	-0.886933	
Sum squared resid	14.88336	Schwarz criterion	-0.844540	
Log likelihood	284.9403	Hannan-Quinn criter.	-0.870465	
F-statistic	27.46263	Durbin-Watson stat	2.012268	
Prob(F-statistic)	0.000000			

最後に、残差をe、理論値（予測値）をpredとして記録しておこう。これは、

```
genr e = resid
```

```
genr pred=durate-e
```

と入力すればよい（residはEquationの残差である）。次に、残差と理論値をもとに、線形性の仮定が妥当であったかを検証しよう。

RESET

非線形性の検定として RESET を行う。このため、残差系列を説明変数と回帰予測値のべき乗に回帰する。「Quick」→「Estimate Equation」を選択し、

`e c pred^2 pred^3 pred^4 durate(-1 to -4) durate(-12)`

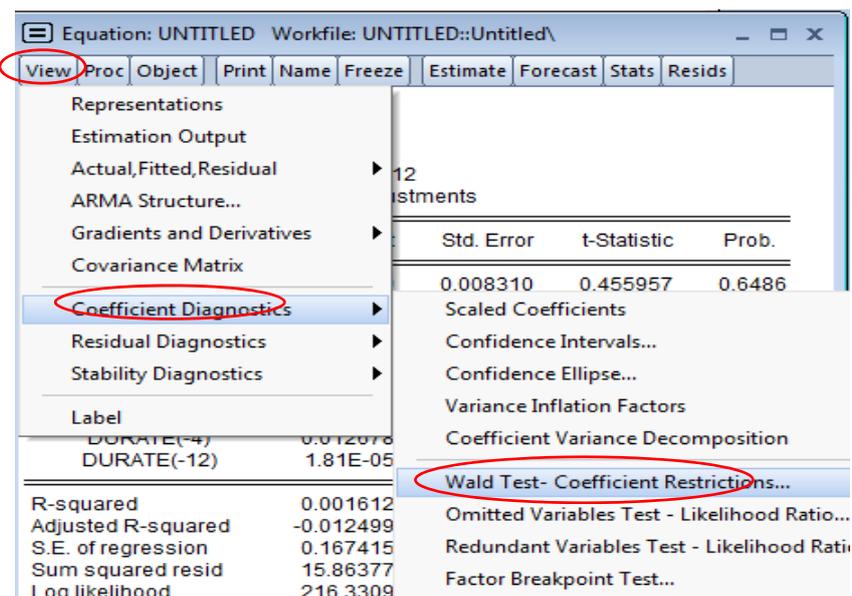
と入力すると以下の結果が得られる (pred^2 とは pred の 2 乗を表す)。

Dependent Variable: E
Method: Least Squares
Date: 11/27/17 Time: 11:24
Sample (adjusted): 1961M02 2013M06
Included observations: 629 after adjustments

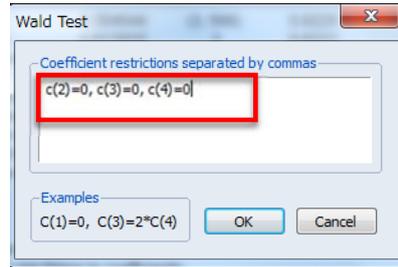
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-0.006158	0.007309	-0.842512	0.3998
PRED^2	1.586283	1.007938	1.573790	0.1160
PRED^3	10.35561	7.003735	1.478584	0.1398
PRED^4	-33.93546	18.91378	-1.794219	0.0733
DURATE(-1)	-0.028927	0.041925	-0.689963	0.4905
DURATE(-2)	-0.044212	0.046985	-0.940965	0.3471
DURATE(-3)	-0.039290	0.044781	-0.877389	0.3806
DURATE(-4)	-0.035810	0.044075	-0.812482	0.4168
DURATE(-12)	0.014168	0.040231	0.352165	0.7248

R-squared	0.006856	Mean dependent var	-3.27E-18
Adjusted R-squared	-0.005959	S.D. dependent var	0.153947
S.E. of regression	0.154405	Akaike info criterion	-0.884273
Sum squared resid	14.78132	Schwarz criterion	-0.820684
Log likelihood	287.1039	Hannan-Quinn criter.	-0.859572
F-statistic	0.534982	Durbin-Watson stat	2.002193
Prob(F-statistic)	0.830466		

pred^2、pred^3、pred^4 の係数の t 値はすべて小さいため、有意な結果となっていない。しかし、個々の係数が 0 から有意に離れていないとしても、すべてのべき乗項の係数が同時に 0 とした帰無仮説が棄却されるとは限らない。帰無仮説 (pred^2、pred^3、pred^4 の係数が同時に 0) とした F 統計量を調べてみよう。Equation Window から「View」→「Coefficient Diagnostics」→「Wald Test」と選択する。



そうすると以下の画面がでてくるので、係数制約を明示的に書こう。



ここでは、 pred^2 、 pred^3 、 pred^4 は 2 番目、3 番目、4 番目の係数なので、

$$c(2)=0, c(3)=0, c(4)=0$$

と入力し、OK をクリックする ($c(2)=c(3)=c(4)=0$ としても同じ結果が得られる)。

Wald Test:
Equation: Untitled

Test Statistic	Value	df	Probability
F-statistic	1.426619	(3, 620)	0.2339
Chi-square	4.279857	3	0.2328

Null Hypothesis: C(2)=0,C(3)=0,C(4)=0
Null Hypothesis Summary:

Normalized Restriction (= 0)	Value	Std. Err.
C(2)	1.586283	1.007938
C(3)	10.35561	7.003735
C(4)	-33.93546	18.91378

Restrictions are linear in coefficients.

これをみると、F 値は 1.426 であり、その p 値は 0.234 である。したがって、帰無仮説を棄却できない。以上から、非線形性を否定する結果が得られたといえる。

マクラウド=リー検定

マクラウド=リー検定のために、補助回帰式を推定する。「Quick」→「Estimate Equation」とし、

$$e^2 c e(-1)^2 e(-2)^2$$

と入力すると、以下の画面が表示される。

Equation: UNTITLED Workfile: UNTITLED::Untitled¥

View Proc Object Print Name Freeze Estimate Forecast Stats Resids

Dependent Variable: E^2
Method: Least Squares
Date: 11/27/17 Time: 11:27
Sample (adjusted): 1961M04 2013M06
Included observations: 627 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.018077	0.002083	8.680223	0.0000
$E(-1)^2$	0.143016	0.039848	3.589039	0.0004
$E(-2)^2$	0.095534	0.039840	2.397956	0.0168
R-squared	0.033919	Mean dependent var		0.023735
Adjusted R-squared	0.030822	S.D. dependent var		0.042817
S.E. of regression	0.042152	Akaike info criterion		-3.490307
Sum squared resid	1.108706	Schwarz criterion		-3.469058
Log likelihood	-1097.211	Hannan-Quinn criter.		-3.482051
F-statistic	10.95421	Durbin-Watson stat		1.995485
Prob(F-statistic)	0.000021			

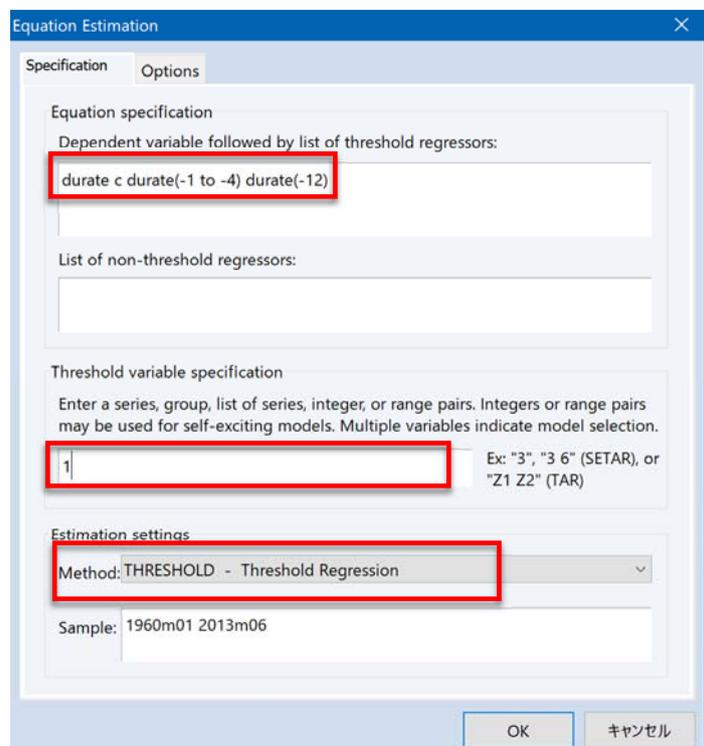
この結果から、 $e(-1)^2$ は 1%有意であり、 $e(-2)^2$ も 5%で有意となる。また、 $e(-1)^2$ と $e(-2)^2$ の係数が同時に 0 という帰無仮説を検定する F 値は 10.95 となっており、帰無仮説は棄却される（ここで表示される F 値は定数項以外の係数がすべて 0 という仮説検定に対応している）¹。

以上から、RESET では非線形性が検出されなかった一方、マクラウド=リー検定では、非線形が検出された。

2. 閾値自己回帰モデル(TAR)

ここでは、失業率のデータを用いて、閾値自己回帰（TAR）モデルを推定してみよう。

「QUICK」→「Estimate Equations」と選択し、Methodを THRESHOLD とする²。



ここで推定式は、以前と同様、`durate c durate(-1 to -4) durate(-12)`である。また、Threshold variable specification は 1 としよう(遅れのパラメータが $d=1$ ということ)³。ここで OK をクリックすると以下の画面が表示される。

¹ この F 値が正しいことを確認したいなら、Equation Window から「View」→「Coefficient Diagnostics」→「Wald Test」と選択し、 $c(2)=c(3)=0$ と入力すればよい。そうすると、同じ結果が得られることが確認できる。

² EViews の古い version だと Method の中に THRESHOLD がないので注意してほしい。

³ Threshold variable specification において、整数を入力すると、EViews は TAR モデルの遅れのパラメータと判断する。たとえば、3 と入力すると、`dunrate(-3)`の値に応じてモデルが変化すると考える。もし threshold variable が被説明変数のラグでなく、他の変数、たとえば失業率のラグ `unrate(-1)`であれば、Threshold variable specification に整数でなく、`unrate(-1)`と入力すればよい。

Dependent Variable: DURATE
 Method: Threshold Regression
 Date: 11/29/17 Time: 09:55
 Sample (adjusted): 1961M02 2013M06
 Included observations: 629 after adjustments
 Threshold type: Bai-Perron tests of L+1 vs. L sequentially determined thresholds
 Threshold variable: DURATE(-1)
 Threshold selection: Trimming 0.15, Max. thresholds 5, Sig. level 0.05
 Threshold value used: 0.07799999

Variable	Coefficient	Std. Error	t-Statistic	Prob.
DURATE(-1) < 0.07799999 -- 453 obs				
C	-0.005089	0.008856	-0.574625	0.5658
DURATE(-1)	-0.045480	0.067511	-0.673663	0.5008
DURATE(-2)	0.117178	0.048500	2.416043	0.0160
DURATE(-3)	0.181127	0.047674	3.799244	0.0002
DURATE(-4)	0.161966	0.046740	3.465262	0.0006
DURATE(-12)	-0.125236	0.040310	-3.106827	0.0020
0.07799999 <= DURATE(-1) -- 176 obs				
C	-0.071263	0.022761	-3.130850	0.0018
DURATE(-1)	0.383925	0.102634	3.740725	0.0002
DURATE(-2)	0.365583	0.067780	5.393665	0.0000
DURATE(-3)	0.113141	0.066857	1.692288	0.0911
DURATE(-4)	0.079416	0.068516	1.159087	0.2469
DURATE(-12)	-0.155242	0.074064	-2.096045	0.0365
R-squared	0.215156	Mean dependent var	0.001472	
Adjusted R-squared	0.201164	S.D. dependent var	0.170068	
S.E. of regression	0.152003	Akaike info criterion	-0.910942	
Sum squared resid	14.25570	Schwarz criterion	-0.826157	
Log likelihood	298.4912	Hannan-Quinn criter.	-0.878007	
F-statistic	15.37672	Durbin-Watson stat	2.006114	

この結果から、閾値は 0.078、SSR は 14.256 であることが確認できる。また、それぞれのシステムにおける係数の値も表示されている。

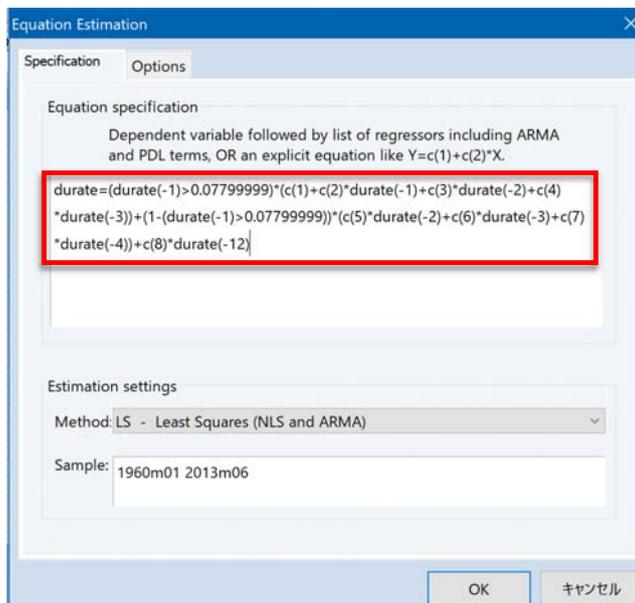
TAR モデルの妥当性を考えるうえで F 検定を行いたいが、教科書の 7 章で言及されているように閾値を推定したため、通常の F 分布表を用いることができない。そこで、Equation Window において、「view」→「Threshold Specification」を選択する。以下の画面は、推定結果の一部だけを切り取ったものである。0 vs 1 とは、閾値がないケースを帰無仮説とし、対立仮説を閾値が 1 つとした F 検定の値である。Scaled F 値は 28.29 であり、有意水準 5% とした臨界値 20.08 を上回っている。したがって、閾値を考えないモデルより、閾値モデルの方が優れていることが分かる。また、1 vs 2(帰無仮説は閾値が 1 つ、対立仮説は閾値が 2 つ)をみると、Scaled F 値は 14.42 であり、臨界値 22.11 を下回る。したがって、閾値は 1 つだけであるといえる。

Sequential F-statistic determined thresholds:			
			1
Threshold Test	F-statistic	Scaled F-statistic	Critical Value**
0 vs. 1 *	4.716394	28.29836	20.08
1 vs. 2	2.403164	14.41898	22.11

* Significant at the 0.05 level.

** Bai-Perron (Econometric Journal, 2003) critical values.

ここで、 $I_t\Delta u_{t-4}$ 、 $(1-I_t)\Delta u_{t-1}$ 、定数項 ($\Delta u_{t-1} \leq 0.078$ の場合) は有意ではなかったため、これらの変数を除いた方がよいかもしれない。また $I_t\Delta u_{t-12}$ と $(1-I_t)\Delta u_{t-12}$ の係数はほぼ同じのため、単純に Δu_{t-12} のみをモデルに含めるべきであるかもしれない。以下では、これらを考慮したモデルを推定してみよう。これは「QUICK」→「Estimate Equations」と選択し、Equation specification に式を入力すればよい。



入力が面倒な読者は、以下を画面にはりつけてもらいたい。

$$\text{durate}=(\text{durate}(-1)>0.07799999)*(c(1)+c(2)*\text{durate}(-1)+c(3)*\text{durate}(-2)+c(4)*\text{durate}(-3))+(1-(\text{durate}(-1)>0.07799999))*(c(5)*\text{durate}(-2)+c(6)*\text{durate}(-3)+c(7)*\text{durate}(-4))+c(8)*\text{durate}(-12)$$

そして OK をクリックすると、以下の結果が得られる。

```

Dependent Variable: DURATE
Method: Least Squares (Gauss-Newton / Marquardt steps)
Date: 11/29/17   Time: 10:51
Sample (adjusted): 1961M02 2013M06
Included observations: 629 after adjustments
DURATE=(DURATE(-1)>0.07799999)*(C(1)+C(2)*DURATE(-1)+C(3)
*DURATE(-2)+C(4)*DURATE(-3))+(1-(DURATE(-1)>0.07799999))
*(C(5)*DURATE(-2)+C(6)*DURATE(-3)+C(7)*DURATE(-4))+C(8)
*DURATE(-12)

```

	Coefficient	Std. Error	t-Statistic	Prob.
C(1)	-0.070106	0.022578	-3.105080	0.0020
C(2)	0.391663	0.102065	3.837394	0.0001
C(3)	0.392758	0.062153	6.319238	0.0000
C(4)	0.118301	0.065650	1.801994	0.0720
C(5)	0.118897	0.048218	2.465846	0.0139
C(6)	0.179811	0.046840	3.838861	0.0001
C(7)	0.159258	0.046059	3.457731	0.0006
C(8)	-0.131361	0.035253	-3.726238	0.0002
R-squared	0.212728	Mean dependent var		0.001472
Adjusted R-squared	0.203854	S.D. dependent var		0.170068
S.E. of regression	0.151747	Akaike info criterion		-0.920571
Sum squared resid	14.29981	Schwarz criterion		-0.864048
Log likelihood	297.5195	Hannan-Quinn criter.		-0.898614
Durbin-Watson stat	2.031547			

SSR(Sum squared resid)は、他のモデルよりも低下しているが、これは単に説明変数が減少したためである。AIC と SBC をみても、他のモデルよりも低下していることが確認できる。

3. LSTAR モデル

TAR モデルでは閾値を境にしてモデルが急激に移行するような状況をとらえることができるが、緩やかに構造が変化することも考えられる。このような場合、平滑推移モデル (smooth transition model) が有用となってくる。教科書の 7 章 9 節の例 (仮想データの LSTAR モデル推定) を再現してみよう。

LSTAR.XLS を Workfile に読み込もう。このデータは、次の LSTAR モデル

$$y_t = 1 + 0.9y_{t-1} + (3 - 1.7y_{t-1})/[1 + \exp(-10(y_{t-1} - 5))] + \varepsilon_t$$

から 250 個の実現値を発生させたものである。まずは、系列 y に関して AR(1) モデルを推定しよう。「QUICK」 → 「Estimate Equation」とし、 y c $y(-1)$ と入力し OK をクリックすると、次の推定結果が得られる。

$$y_t = 0.277996 + 0.552494y_{t-1} + e_t$$

(1.503229) (10.41585)

そして、モデルの残差と被説明変数の理論値を記録しておこう。これは

genr e = resid

genr pred = y-e

と入力すればよい。

ここで、教科書で行われている非線形性の検定を試みよう。まず、RESET を行うため、「Quick」 → 「Estimate Equation」を選択する。そうすると、Equation Estimation Window が表示されるので、Equation Specification に

e c y(-1) pred^2 pred^3 pred^4

と入力し、OK をクリックする。そうすると、以下の推定結果が得られる。これは教科書の結果とほぼ同じである。

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.932197	0.219872	4.239725	0.0000
Y(-1)	0.710069	0.078553	9.039396	0.0000
PRED^2	0.058359	0.089885	0.649260	0.5168
PRED^3	-0.156936	0.016719	-9.386596	0.0000
PRED^4	-0.034198	0.007061	-4.843551	0.0000

R-squared	0.540318	Mean dependent var	-4.01E-17
Adjusted R-squared	0.532782	S.D. dependent var	2.865955
S.E. of regression	1.958976	Akaike info criterion	4.202597
Sum squared resid	936.3714	Schwarz criterion	4.273229
Log likelihood	-518.2233	Hannan-Quinn criter.	4.231027
F-statistic	71.70040	Durbin-Watson stat	2.142555
Prob(F-statistic)	0.000000		

非線形項 (pred^2, pred^3, pred^4) の有無に関する F 統計量を計算してみよう。これは Equation Window から「View」 → 「Coefficient Diagnostics」 → 「Wald Test」と選択し、

c(3)=c(4)=c(5)=0 と入力することで、F 値=95.60 という結果が得られる。以上から、何らかの非線形性があることが確認できる。

次に、テラスバータの検定を試みよう。これは $d=1$ であれば Equation Specification に $e\ c\ y(-1)\ y(-1)^2\ y(-1)^3\ y(-1)^4$ と入力すればよい。非線形項 ($y(-1)^2$ 、 $y(-1)^3$ 、 $y(-1)^4$) の有無に関する F 統計量を求めて、係数が同時に 0 という帰無仮説が棄却されることを確認してほしい。また、 $y(-1)^4$ の係数の t 値が十分に大きいことも確認してほしい。したがって、ESTAR モデルではなく LSTAR モデルが選ばれる。

同様に、 $d=2$ であれば

$e\ c\ y(-1)\ y(-1)*y(-2)\ y(-1)*y(-2)^2\ y(-1)*y(-2)^3$

と入力すればよい。自分で推定してみると、遅れのパラメータが $d=1$ の方が、 $d=2$ よりも当てはまりが良いことが分かるだろう。

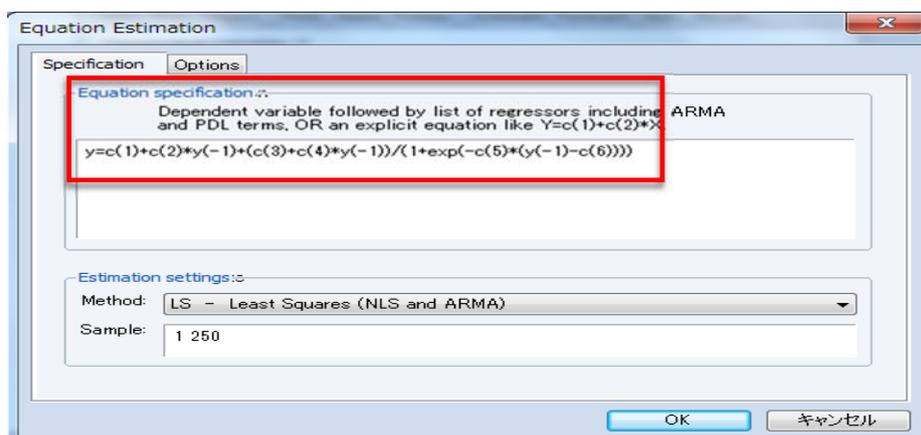
以下では、適切な推移変数として 1 期前の被説明変数の値を採用する。また、以下の LSTAR モデルを推定する。

$$y_t = \alpha_0 + \alpha_1 y_{t-1} + (\beta_0 + \beta_1 y_{t-1}) / (1 + \exp(-\gamma(y_{t-1} - c))) + \varepsilon_t$$

これは「QUICK」→「Estimate Equation」を選択し、Equation Specification に推定式を

$y=c(1)+c(2)*y(-1)+(c(3)+c(4)*y(-1))/(1+\exp(-c(5)*(y(-1)-c(6))))$

と入力すればよい。



そして OK をクリックすると、以下の推定結果が得られる。

View	Proc	Object	Print	Name	Freeze	Estimate	Forecast	Stats	Resids
Equation: UNTITLED Workfile: UNTITLED::Untitled\									
Dependent Variable: Y Method: Least Squares Date: 09/15/15 Time: 13:11 Sample (adjusted): 2 250 Included observations: 249 after adjustments Convergence achieved after 18 iterations Y=C(1)+C(2)*Y(-1)+(C(3)+C(4)*Y(-1))/(1+EXP(-C(5)*(Y(-1)-C(6))))									
		Coefficient	Std. Error	t-Statistic	Prob.				
	C(1)	0.941528	0.065259	14.42765	0.0000				
	C(2)	0.922875	0.020439	45.15201	0.0000				
	C(3)	-5.862800	2.820713	-2.078482	0.0387				
	C(4)	-1.179275	0.479755	-2.458080	0.0147				
	C(5)	11.20696	1.654241	6.774683	0.0000				
	C(6)	5.002021	0.016015	312.3354	0.0000				
	R-squared	0.921818	Mean dependent var		0.620246				
	Adjusted R-squared	0.920209	S.D. dependent var		3.438228				
	S.E. of regression	0.971204	Akaike info criterion		2.803242				
	Sum squared resid	229.2068	Schwarz criterion		2.888000				
	Log likelihood	-343.0036	Hannan-Quinn criter.		2.837358				
	F-statistic	573.0268	Durbin-Watson stat		2.118327				
	Prob(F-statistic)	0.000000							

この結果をまとめると、

$$y_t = 0.941 + 0.923y_{t-1} + (-5.86 - 1.18y_{t-1})/[1 + \exp(-11.207(y_{t-1} - 5.00))] + \varepsilon_t$$

(14.43) (45.15) (-2.03) (-2.45) (6.77) (312.33)

もし LSTAR ではなく、ESTAR モデルとして

$$y_t = \alpha_0 + \alpha_1 y_{t-1} + (\beta_0 + \beta_1 y_{t-1})(1 - \exp(-\lambda(y_{t-1} - c)^2)) + \varepsilon_t$$

を推定したいのであれば、Equation Specification に推定式を

$$y=c(1)+c(2)*y(-1)+(c(3)+c(4)*y(-1))*(1-\exp(-c(5)*(y(-1)-c(6))^2))$$

と入力すればよい。

4. 単位根と非線形過程

7章 11 節で紹介したエンダース=グレンジャー(Enders and Granger,1998)の結果を再現してみよう。GRANGER.XLS を読み込んで、金利スプレッドを定義する。

$$\text{genr } s=r_10-r_short$$

ここでモデルは、

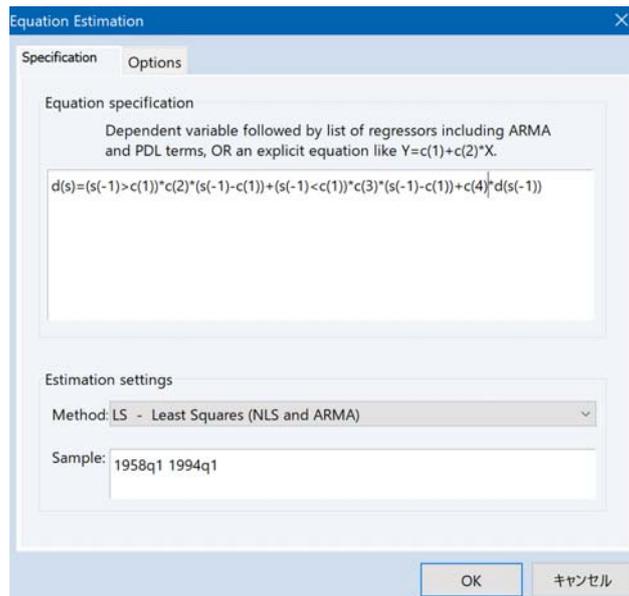
$$\Delta s_t = I_t \rho_1 (s_{t-1} - \tau) + (1 - I_t) \rho_2 (s_{t-1} - \tau) + \alpha_1 \Delta s_{t-1} + \varepsilon_t$$

$$I_t = \begin{cases} 1 & s_{t-1} > \tau \text{ のとき} \\ 0 & s_{t-1} \leq \tau \text{ のとき} \end{cases}$$

としよう。

このモデルを推定するには、「Quick」→「Estimate Equation」を選択し、Equation Estimation Window の Equation Specification に以下を入力すればよい。

$$d(s)=(s(-1)>c(1))*c(2)*(s(-1)-c(1))+s(-1)<c(1))*c(3)*(s(-1)-c(1))+c(4)*d(s(-1))$$



そして OK をおすと、以下の画面が出力される。

	Coefficient	Std. Error	t-Statistic	Prob.
C(1)	-0.270001	0.900360	-0.299881	0.7647
C(2)	-0.065547	0.049055	-1.336209	0.1837
C(3)	-0.285971	0.128005	-2.234067	0.0271
C(4)	0.171728	0.083228	2.063341	0.0409

R-squared	0.106185	Mean dependent var	0.006154
Adjusted R-squared	0.086894	S.D. dependent var	0.895665
S.E. of regression	0.855866	Akaike info criterion	2.554169
Sum squared resid	101.8185	Schwarz criterion	2.637045
Log likelihood	-178.6231	Hannan-Quinn criter.	2.587846
Durbin-Watson stat	2.032683		

ここで τ は-0.27 であり、 ρ_1 は-0.065、 ρ_2 は-0.285 として推定されている。

次に、調整が階差に依存するモデルを推定してみよう。

$$\Delta s_t = I_t \rho_1 (s_{t-1} - \tau) + (1 - I_t) \rho_2 (s_{t-1} - \tau) + \alpha_1 \Delta s_{t-1} + \varepsilon_t$$

$$I_t = \begin{cases} 1 & \Delta s_{t-1} > 0 \text{ のとき} \\ 0 & \Delta s_{t-1} \leq 0 \text{ のとき} \end{cases}$$

これは Equation Specification において以下を入力すれば教科書と同じ結果が得られる。

$$d(s) = (d(s(-1)) > 0) * c(1) * (s(-1) - c(2)) + (d(s(-1)) < 0) * c(3) * (s(-1) - c(2)) + c(4) * d(s(-1))$$