第3章　最小二乗法

はじめに

　計量経済学で最もひんぱんに使うのが最小二乗法である。２つ以上の変数の関係を統計的に推計できるものである。

　最も簡単な消費関数は、Ｃを消費、Ｙを所得として次の式で表される。

Ｃ=a+bY

　理論的に考える場合は、係数 a,b をある定数として考えることができるが、実証的に考える場合は、係数が現実的な数字でなければならない。

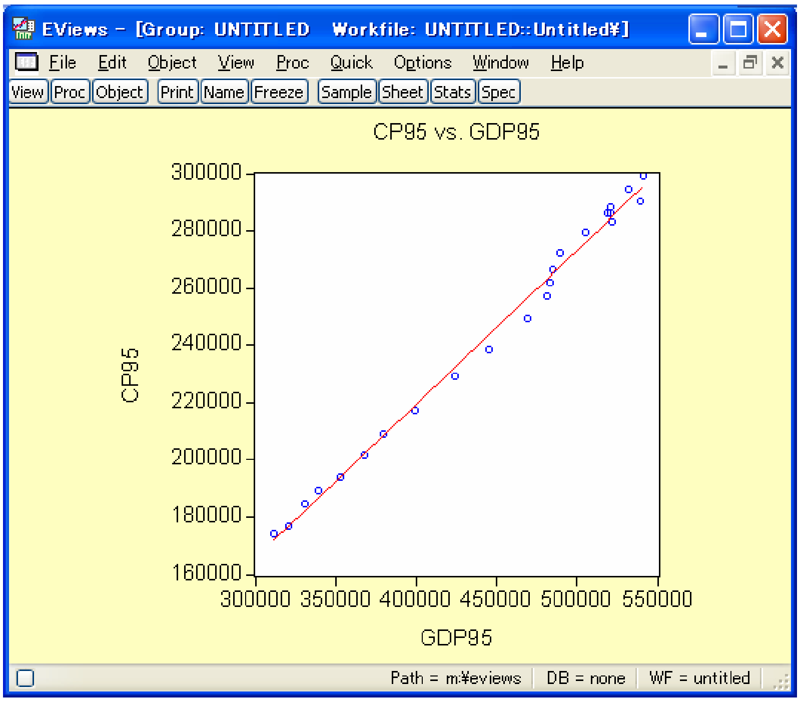
　この定数は、Ｃをｙ軸、Ｙをｘ軸として散布図を書いてみると意味するところがわかる。a はさまざまな消費と消費の組み合わせを示す点が示す直接の切片、ｂは傾きを表すことがわかる。目分量でも、何らかの直線が引けるが、そのなかで、最も望ましい性質を持っているのが最小二乗法による a と b の推計である。

散布図による確認

　実質ＧＤＰ（GDP95）と実質民間最終消費（CP95）で説明すると、まず２変数をグループ（第１章参照）として開く。グループオブジェクトのメニューバーで、

View/Graph/Scatter/Scatter with regression...

 で、次のオプションは何もせずに Ok を押すと縦軸に GDP95、横軸に CP95 をとったグラフが出てくる。この 2 変数の関係を最もよく表す直線も表示される。



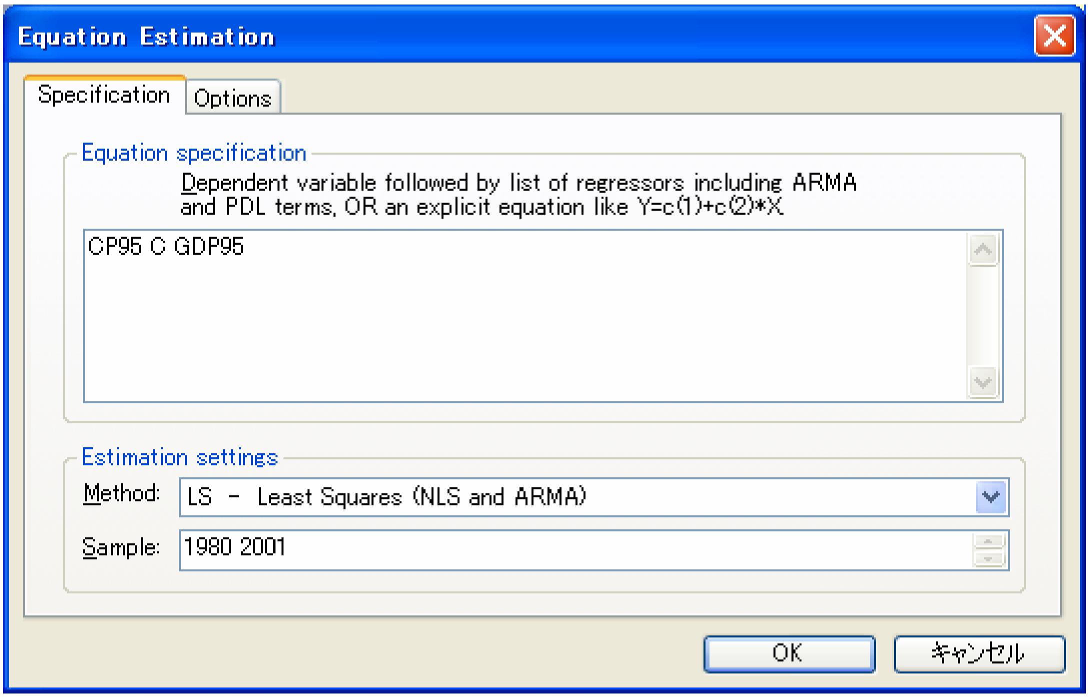
最小二乗法の推計法

Eviews で最小二乗法を推計するには、メイン画面のメニューバーで、
Quick/Estemate equation...
を選ぶと変数を入力する画面になる。
CP95=a+bGDP95+e という式を推計したい場合は、次のように入力する。

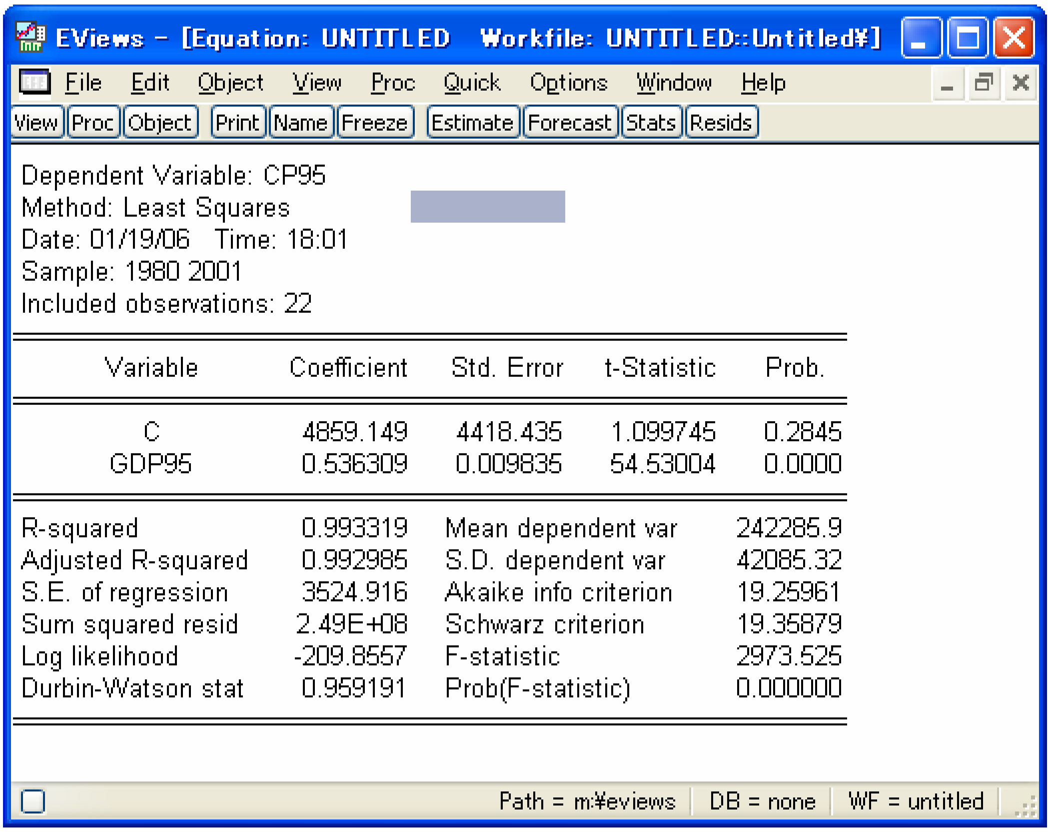
CP95＝c(1)＋c(2)*GDP95

c(1)と c(2)は推計すべき係数を表す。c は係数（Coefficient）の頭文字をとっている。
あるいは簡単に変数を並べても同じことを意味する。

CP95 C GDP95



OK を押すと推計結果が出力される。定数項はCで表される。出力されるさまざまな統計量についてまとめたのが、表である。詳しい説明は山澤（2002）など参照。


EViews - [Equation: UNTITLED Workfile: UNTITLED::Untitled¥]

File Edit Object View Proc Quick Options Window Help

View Proc Object Print Name Freeze Estimate Forecast Stats Resids

Dependent Variable: CP95
Method: Least Squares
Date: 01/19/06 Time: 18:01
Sample: 1980 2001
Included observations: 22

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	4859.149	4418.435	1.099745	0.2845
GDP95	0.536309	0.009835	54.53004	0.0000

R-squared	0.993319	Mean dependent var	242285.9
Adjusted R-squared	0.992985	S.D. dependent var	42085.32
S.E. of regression	3524.916	Akaike info criterion	19.25961
Sum squared resid	2.49E+08	Schwarz criterion	19.35879
Log likelihood	-209.8557	F-statistic	2973.525
Durbin-Watson stat	0.959191	Prob(F-statistic)	0.000000

Path = m:¥eviews DB = none WF = untitled

項目	説明	説明
Dependent Variable: CP95	被説明変数が CP95	
Method: Least Squares	最小二乗法	
Sample: 1980 2002	推計期間が 1980 年から 2002 年	
Included observations: 23	サンプル数が 23	
Variable Coefficient	係数	
Std. Error	係数の標準誤差	
t-Statistic	ｔ値	およそ 2 以上が望ましい。
Prob.	ｔ値の現れる確率	0.05（５%）以下が望ましい。
C	定数項	
GDP95	説明変数が GDP95	
R-squared	決定係数	1 に近いほど当てはまりがよい。
Adjusted R-squared	自由度修正済み決定係数	同上

S.E. of regression	回帰の標準誤差	
Sum squared resid	残差の二乗和	小さいほど当てはまりがよい。
Log likelihood	残差の対数尤度	大きいほど当てはまりがよい。
Durbin-Watson stat	ダービンワトソン比	２に近いほど残差に自己相関がない。
Mean dependent var	被説明変数の平均	
S.D. dependent var	被説明変数の分散	
Akaike info criterion	赤池情報規準	小さいほど当てはまりがよい。
Schwarz criterion	シュワルツ情報規準	小さいほど当てはまりがよい。
F-statistic	Ｆ値	大きいほどよい。
Prob(F-statistic)	Ｆ値の現れる確率	0.05%以下が望ましい。

t 値は、ある係数が意味のある係数かどうかを示している。大きいほど、その係数がゼロから離れていると考えられる。「Prob.」は「係数がゼロ」と仮定したときに、t 値が現れる確率を示している。確率が小さいと、「係数がゼロ」という仮定が棄却されることを示す。

当てはまりを示す指標の代表的なものは決定係数で、ゼロから１の間の値をとる。サンプルが少ないと決定係数は大きめに出るので、自由度修正済み決定係数を使うほうがよい。

Ｆ値は、「すべての係数がゼロである」という仮説を検定する場合に使う。Ｆ値の現れる確率が小さいほど、同仮説が棄却される。

推計値と残差のグラフ

最小二乗法は次の式の a やｂを推計することであるが、説明変数と係数の組み合わせである a+bY が被説明変数のＣをすべて説明する場合はほとんどなく、誤差 e が生じる。

Ｃ=a+bY+e

説明変数と係数によって説明される部分を推計値と呼ぶ。Ｃ’で表すと、推計値は次式である。

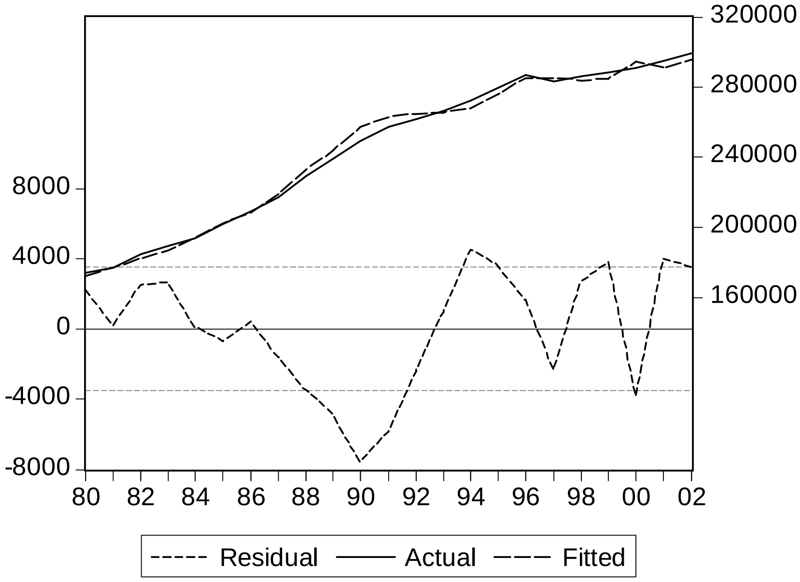
Ｃ’=a+bY

推計値や残差の動きを確認するのは、回帰式がうまく推計できたかどうかを検討するう

えで非常に重要である。EVｉｅｗｓでは、次の手順で、推計値と残差のグラフを見ることができる。

Equation オブジェクト：

[View]→[Actual,Fitted,Residual]→[Actual,Fitted,Residual Graph]



残差を一つの系列として保存する場合は、

推計値を一つの系列として保存する場合は、

予測

式を推計するこができると、被説明変数の予測を行うことができる。予測とは、実績値のない部分を推計することである。

Ｃ=a+bY

Ｙの将来の値を何らかの形で予測すれば、上記の関係を使ってＣの予測値が作られる。将来に関しては誤差の情報はないので、通常ゼロと置く。誤差はゼロではあるが、誤差の分散はある。

$C=a+bY+C$