
OLSQ (HCTYPE=頑健SEタイプ, HI, ROBUSTSE, SILENT, TERSE,
UNNORM, WEIGHT=ウェイトする変数名, WTYPE=ウェイトのタイプ)
従属変数 独立変数リスト ;

機能 :

OLSQ は TSP における基本的な回帰プロシジャです。この方法は、従属変数を 1 組の独立変数に回帰して回帰係数の最小 2 乗推定値を計算します。オプションを使えば、分散不均一性を修正する加重最小 2 乗推定値や、誤差項の分散不均一性に対して頑健な標準誤差を得ることができます (自己相関に対して頑健な推定値は GMM コマンドを参照して下さい)。

使用法 :

基本的な OLSQ コマンドでは、最初に従属変数を指定し、次に独立変数を与えます。(通常行うように) 定数項が回帰式に必要なならば、独立変数のリスト中に特別な変数である C か CONSTANT を含めて下さい。独立変数の数は自由ですが、1 つのステートメントあたりの引数の数の制限と使用可能な作業領域によって制約を受けます。もちろんその数は、利用するデータ観測値の数によっても制限を受けます。

回帰が計算される標本期間は、計算時に指定されている SMPL によって決められます。もし現在の標本期間で欠測値があるなら、その観測値はサンプルから落とされ、欠測値のある各系列に対して警告メッセージがプリントされます。計算に使われた残りの観測値数は回帰計算の結果としてプリントされます。この場合 @RES と @FIT は欠測値を含み、ダービン・ワトソン統計量は標本ギャップに対して調整されます。

OLSQ コマンドに含まれる独立変数には、PDL 変数 (多項式分布ラグ) と同様にラグ変数やリード変数を含むことができます。この PDL 変数は、回帰式に多くのラグ付き変数が含まれる場合に、係数間に滑らかな多項式を想定して自由に動ける係数の数を減らす方法です。PDL 変数の与え方については、PDL の項を参照して下さい。

オプション :

HCTYPE= 計算する分散不均一致標準誤差のタイプ (0 と 3 の間の値で、既定値は 2 です)。このオプションは ROBUSTSE を意味します。一般に分散・共分散行列の推定値は次の形を取ります。

$$V = (X'X)^{-1} \left(\sum e_i^2 x_i' x_i / d_i \right) (X'X)^{-1}$$

オプション HCTYPE は d_i に対して用いられる式を決めます。

HCTYPE	d_i	
0	1	(通常の Eicker-White 漸近式)
1	$(T - k)/T$	(有限標本の自由度を用いる)
2	$1 - h_i$	(e が真に分散均一なら不偏)
3	$(1 - h_i)^2$	(ジャックナイフ近似)

ここで h_i は「ハット」行列 (以下の HI を参照) の対角要素である。もしある i に対して $1 - h_i$ がゼロなら、 $(T - k)/T$ が代入される。HCTYPE=2 と HCTYPE=3 の両方とも、よい有限標本の性質を持っている。詳細は Davidson and MacKinnon, pp.552-556 を参照して下さい。

HI/NOHI この指定によりハット行列 $X(X'X)^{-1}X$ の対角要素を @HI に保存するかどうか決めます。この対角要素は、標準的な行列計算では計算が面倒な(データの誤り、はずれ値などの)「影響力の強い」観測値の検出に役立ちます。次のようなステートメント

```
SELECT @HI > 2*@NCOEF/@NOB ;
```

によって影響力の強い観測値と判定します。詳しくは Belsley, Kuh, and Welsch, あるいは Krasker, Kuh, and Welsch を参照して下さい。

NORM/UNNORM は観測値数を合計してウェイトを正規化するかどうか指示します。このオプションは係数推定値やほとんどの統計量には影響を与えませんが、このオプションを使えば加重された変数も、加重されない変数も変数値の大きさが平均で同じになり、推定結果の解釈が容易になります。もし WEIGHT オプションが指定されていなければ、NORM は何の効果も与えません。

ROBUSTSE/NOROBUST は White などによって提案された方法で、係数推定値の分散、標準誤差と対応する t -統計量を計算します。これらの分散の推定値は、攪乱項の分散が均一でない場合も、その分散がモデルの独立変数と相関を持っていたとしても一致性を持ちます。しかし攪乱項が独立でない場合にも、一致性を持ちません。厳密な公式は、参考文献を参照して下さい。

SILENT/NOSILENT は、すべての回帰の結果を省略します。このコマンドは数多くの回帰式を計算して、特定の推定結果だけが欲しいときに役立ちます。これらは@変数で得ることができ、これらは SILENT オプションがオンでも保存されます。TERSE も参照して下さい。

TERSE/NOTERSE は、観測値数、尤度関数の対数と係数、標準誤差、 t 統計量という最小のアウトプットを出力します。SILENT も参照して下さい。

WEIGHT= 観測値をウェイト付するのに用いる変数名です。回帰式を計算する前に、観測値には正規化したウェイト系列の平方根が掛けられます(上の NORM オプションを参照)。この系列は変数の分散の逆数に比例しています。もし特定の観測値についてウェイトの値がゼロならば、その観測値は計算から除去され、自由度の計算からも排除されます。

WTYPE= HET あるいは REPEAT がウェイトのタイプで、特に指定しないときは REPEAT になり、このときウェイトは繰り返し回数(直接に尤度関数に掛けます)になります。これはグループ化されたデータに対して用いられ、UNNORM オプションと一致します。WTYPE=HET はウェイトが分散不均一であることを意味しています(尤度関数には σ_T^2 を通じてのみ入ってきます)。回帰アウトプットにおけるこれらの2つのオプションの唯一の違いは、対数尤度関数の値です(全ての係数、標準誤差等は同じです)。WTYPE=HET では、対数尤度は対数ウェイトを含んでおり、既定値の WTYPE=REPEAT はこれを含んでいません。

例：

この例のモデルの消費関数を推定します。

```
OLSQ CONS C GNP ;
```

次の例では、人口変数を重みとして、各州における一人住いの若年層の比率を、社会地理的な変数の上に回帰します。回帰式は一人当りの数値を使っていますから、誤差項の分散は人口の逆数に比例しています。

```
OLSQ (WEIGHT=POP) YOUNG C RSALE URBAN CATHOLIC ;
```

他の例として

```
OLSQ (ROBUSTSE) LOGP C LOGP(-1) LOGR ;
OLSQ TBILL C RATE(4,12,FAR) ;
```

を挙げておきます。

アウトプット：

OLSQ のアウトプットは方程式の名前と従属変数の名前で始まります。次に推定結果の良さを示す様々な統計量が与えられます。残差平方和、回帰の標準誤差、 R^2 、残差の自己相関に対するダービン・ワトソン比、定数項以外の回帰の全係数がゼロであるという仮説に対する F 統計量などです。もし回帰式が定数項を含まず、従属変数から平均が差し引かれていなければ、 F 統計量は意味がありません。そのほかの（ほとんどが自己相関と分散不均一性の検定である）様々な回帰診断に対しては REGOPT コマンドを参照して下さい。続いて右辺の変数名、係数の推定値、標準誤差と対応する t -統計量が表の形式で与えられます。回帰の前に REGOPT(PVPRINT) T; とすると、 P -値と t 統計量の有意性に対する星印がプリントできます。

もし REGOPT コマンドで選択されていれば、分散共分散行列と相関行列がこの表の後にプリントされます。独立変数にラグ付きの従属変数があれば通常のダービン・ワトソン統計量は偏りがあり、系列相関に対しては別の検定が計算されます。この統計量は、残差を従属変数とし、残差のラグ付きの値を右辺の変数の一つとする補助回帰を計算し、ラグ付き残差変数の有意性を検定します。詳しくは、ダービンの論文を参照して下さい。この検定法は、同じラグ付き従属変数の場合における AR1 回帰での標準誤差を修正するのに用いる方法と非常に良く似ています。この検定法は、ダービンの h -検定よりもより一般的です。なぜならば従属変数のラグが、複数個の右辺の変数として回帰式に含まれる場合も、この方法は有効だからです。もし WEIGHT コマンドや SMPL にギャップがあったり、OLSQ Y C X Y(-1); や PDL における様に、明示的にラグ付き従属変数を指定する代わりに) GENR コマンドによって計算されたために、ラグ付き従属変数を右辺の変数の中から検出できなければ、この検定統計量は計算されません。

PLOTS オプションが指定されていれば、従属変数の実際値と推定値がプロットされ、また残差がプリントされます。OLSQ は計算結果のほとんどを、後の計算で使うためにデータ領域に保存します。

次ページの表は、OLSQ コマンドの後で使える変数のリストです。

方法：

OLSQ は行列方程式

$$b = (X'X)^{-1}X'y$$

を計算します。ここで X 独立変数の行列、 y は従属変数のベクトルです。この回帰式の計算（そして TSP のその他すべての回帰タイプの推定値）は非常に正確なもので、内積の計算の前に X 行列を直交正規行列に変換してから、逆行列を求め、そして結果の逆変換を行います（本マニュアルの ORTHON を参照）。

OLSQ は Longley のデータについて、64 ビット倍精度の IBM370 で同じ計算をして精度を確かめてあります。6 桁の精度を持つことが分かりました。Wampler の論文で指摘されている Lauchli の問題についても、5 桁まで OLSQ は正確な答えを出します。指数関数の値が小さくなり過ぎると、回帰式は計算できなくなります。この状態が起きる前に、OLSQ は異常を感知して、係数をゼロにすることにより変数を一個落とし、警告をプリントします。これらの特別な回帰に対する結果は、OPTION DOUBLE; が有効である場合はもっと正確になります。これは、データが有効数字が通常より大きくなるためです。

変数	タイプ	長さ	変数の説明
@LHV	リスト	1	従属変数名
@SSR	スカラー	1	残差平方和
@S2	スカラー	1	残差の分散
@S	スカラー	1	回帰式の標準誤差
@YMEAN	スカラー	1	従属変数の平均値
@SDEV	スカラー	1	従属変数の標準偏差
@NOB	スカラー	1	観測値数
@DW	スカラー	1	ダービン・ワトソン値
@DH	スカラー	1	ダービンの h (ラグ付の従属変数)
@DHALT	スカラー	1	ダービンの h 代替統計 (ラグ付の従属変数)
@LMHET	スカラー	1	LM 分散不均一性検定
@JB	スカラー	1	残差の正規性の Jarque-Bera(LM) 検定
@RESET2	スカラー	1	2 次の X の欠測の Ramsey の RESET 検定
@RSQ	スカラー	1	R^2
@ARSQ	スカラー	1	調整済み R^2
@FST	スカラー	1	F 統計量
@NCOEF	スカラー	1	係数の数
@NCID	スカラー	1	識別された係数の数 (非ゼロ SE)
@SSRO	スカラー	1	元データの SSR(加重回帰)
@...O	スカラー	1	元データの @S20, @S0, ... @ARSQ0
@RNMS	リスト	変数の数	右辺の変数名
@COEF	ベクトル	変数の数	係数推定値
@SES	ベクトル	変数の数	標準偏差
@T	ベクトル	変数の数	t 統計量
%T	ベクトル	変数の数	t 統計量の p 値
@VCOV	行列	変数の数×変数の数	係数推定値の共分散行列
@RES	系列	観測値数	残差系列
@FIT	系列	観測値数	従属変数の推定値
@LOGL	スカラー	1	対数尤度
@HI	系列	観測値数	HI オプションの場合のハット行列の対角要素

回帰式が PDL 変数や SDL 変数を含む場合は、次の変数も保存されます。

@SLAG	スカラー	1	ラグ係数の和
@MLAG	スカラー	1	ラグ係数の平均
@LAGF	ベクトル	観測値数	『分解』後の推定されたラグ係数

注: REGOPT(NOPRINT) LAGF ; は PDL 変数のラグプロットをオフにします。

参考文献:

Belsley, David A., Kuh, Edwin, and Welsch, Roy E., **Regression Diagnostics: Identifying Influential Data and Sources of Collinearity**, John Wiley & Sons, New York, 1980, pp. 11-18.

Davidson, Russell, and James G. MacKinnon, **Estimation and Inference in Econometrics**, Oxford University Press, New York, 1993, pp. 552-556.

Durbin, J., "Testing for Serial Correlation in Least Squares Regression where Some of the Regressors are Lagged Dependent Variables," **Econometrica**, 38, p. 420.

Judge et al, **The Theory and Practice of Econometrics**, John Wiley & Sons, New York, 1981, pp.11-18,126-144.

Krasker, William S., Kuh, Edwin, and Welsch, Roy E., "Estimation for Dirty Data and Flaed Models," **Handbook of Econometrics, Volume I**, Griliches and Intrilligator eds., North-Holland Publishing Company, New York, 1983, pp. 660-664.

Longley, James W., "An Appraisal of Least Squares Programs for the Electronic Computer from the Point of View of the User," **JASA**, 1967, pp.818-841.

Maddala, G. S., **Econometrics**, McGraw Hill Book Company, New York, 1977 pp. 104-127, 257-268.

Pindyck, Robert S., and Daniel L. Rubinfeld, **Econometric Models and Economic Forecasts**, McGraw Hill Book Company, New York, 1976, Chapter 2,3,4.

Wampler, Roy H., "Test Procedures and Test Problems for Least Squares Algorithms," **Journal of Econometrics**, 12, pp. 3-21.

White, Halbert, "A Heteroskedasticity-Consistent Covariance Matrix and a Direct Test for Heteroskedasticity," **Econometrica**, 48, 1980, pp. 721-746.