

不均一分散誤差を伴う線形回帰

モデル式

$$y_j = \mathbf{x}_j\boldsymbol{\beta} + \epsilon_j \quad (11.1)$$

において ϵ_j の分散が一定とは言えない場合の対処法としては次のようなアプローチが考えられます。

1. regress コマンド ([R] regress (*mwp-037*) 参照) において vce(robust) オプションを指定する方法。この場合には不均一分散に対する White の補正 (White correction for heteroskedasticity) を施した形で標準誤差の推定が行われます。
2. ϵ_j の分散値がデータとして与えられている場合には vwls コマンド ([R] vwls 参照) を使用することができます。vwls というのは variance-weighted least squares の略です。
3. arch コマンド ([TS] arch (*mwp-051*) 参照) の het() オプションを使うと、共変量を使って分散をモデル化することができます。ただし arch は時系列データ用のコマンドであるため、それを cross-sectional なデータに適用するためには、人為的に時間変数を生成した上で tsset を実行するという操作が必要となります。
4. OLS (ordinary least squares) に基づく regress コマンドの場合、その目標とするところは平均値の予測にあります。しかし平均値は外れ値 (outliers) による影響を受けやすいという特性を持っています。このため不均一分散の存在下においてはしばしば中央値 (median) 回帰が代りに用いられます。この中央値回帰の機能は qreg コマンド ([R] qreg 参照) によって提供されます。

■