7.2　最小二乗推定量の分散について

単回帰

$$Y_i = \alpha + \beta X_i + u_i, \qquad n = 1, 2, \cdots, n$$

仮定：$\mathrm{E}(u_i) = 0$

$\mathrm{V}(u_i) = \mathrm{E}(u_i^2) = \sigma^2$

$i \neq j$ について，$\mathrm{Cov}(u_i, u_j) = \mathrm{E}(u_i u_j) = \sigma_{ij}$ $\longleftarrow$ この仮定追加

系列相関を無視して，通常の最小二乗推定量は,

$$\widehat{\beta} = \sum_i \omega_i Y_i = \beta + \sum_i \omega_i u_i$$

である。ただし，$\omega_i = \dfrac{X_i - \overline{X}}{\sum_j (X_j - \overline{X})^2}$

$\mathrm{E}(\widehat{\beta})$ について,

$$\begin{aligned}
\mathrm{E}(\widehat{\beta}) &= \mathrm{E}(\beta + \sum_i \omega_i u_i) \\
&= \beta + \sum_i \omega_i \mathrm{E}(u_i) = \beta
\end{aligned}$$

$u_1, u_2, \cdots, u_n$ に系列相関があっても，$\widehat{\beta}$ は不偏推定量となる。

$\mathrm{V}(\widehat{\beta})$ について，

$$
\begin{aligned}
\mathrm{V}(\widehat{\beta}) &= \mathrm{V}(\beta + \sum_i \omega_i u_i) = \mathrm{V}(\sum_i \omega_i u_i) \\
&= \mathrm{E}((\sum_i \omega_i u_i)^2) \qquad \longleftarrow \text{分散の定義} \\
&= \mathrm{E}((\sum_i \omega_i u_i)(\sum_j \omega_j u_j)) \qquad \longleftarrow \text{添字一つ変更} \\
&= \mathrm{E}(\sum_i \sum_j \omega_i \omega_j u_i u_j) \\
&= \sum_i \sum_j \omega_i \omega_j \mathrm{E}(u_i u_j) \\
&= \sum_i \omega_i^2 \mathrm{E}(u_i^2) + \underset{i \neq j}{\sum_i \sum_j} \omega_i \omega_j \mathrm{E}(u_i u_j) \\
&= \sigma^2 \sum_i \omega_i^2 + \underset{i \neq j}{\sum_i \sum_j} \sigma_{ij} \omega_i \omega_j
\end{aligned}
$$

$$\neq \sigma^2 \sum_i \omega_i^2$$

したがって，$u_1, u_2, \cdots, u_n$ に系列相関があるとき，通常の最小二乗推定量 $\widehat{\beta}$ の分散の推定量は，$s^2 \sum_i \omega_i^2 + \sum_i \sum_{\substack{j \\ i \neq j}} s_{ij} \omega_i \omega_j$

とならなければならない。

s^2, s_{ij} は σ^2, σ_{ij} の推定量とする。

しかし，計量ソフトは $s^2 \sum_i \omega_i^2$ と計算する。

7.3　系列相関のもとで回帰式の推定

回帰式が

$$Y_i = \alpha + \beta X_i + u_i,$$
$$u_i = \rho u_{i-1} + \epsilon_i,$$

のときの推定を考える。ただし，$\epsilon_1, \epsilon_2, \cdots, \epsilon_n$ は互いに独立とする。

u_i を消去すると，

$$(Y_i - \rho Y_{i-1}) = \alpha(1-\rho) + \beta(X_i - \rho X_{i-1}) + \epsilon_i,$$

となり，

$Y_i^* = (Y_i - \rho Y_{i-1})$,
$X_i^* = (X_i - \rho X_{i-1})$
を新たな変数として，

$$Y_i^* = \alpha' + \beta X_i^* + \epsilon_i,$$

に最小二乗法を適用する。$\epsilon_1, \epsilon_2, \cdots, \epsilon_n$ は互いに独立とするなので，最小二乗法を適用が可能となる。ただし，$\alpha' = \alpha(1-\rho)$ の関係が成り立つことに注意。

より一般的に，回帰式が

$$Y_i = \beta_1 X_{1i} + \beta_2 X_{2i} + \cdots + \beta_k X_{ki} + u_i,$$
$$u_i = \rho u_{i-1} + \epsilon_i,$$

のときの推定を考える。ただし，$\epsilon_1, \epsilon_2, \cdots, \epsilon_n$ は互いに独立とする。

u_i を消去すると，

$$(Y_i - \rho Y_{i-1}) = \beta_1(X_{1i} - \rho X_{1,i-1})$$

$$
\begin{aligned}
&+ \beta_2(X_{1i} - \rho X_{2,i-1}) \\
&+ \cdots \\
&+ \beta_k(X_{1i} - \rho X_{k,i-1}) + \epsilon_i,
\end{aligned}
$$

となり，

$Y_i^* = (Y_i - \rho Y_{i-1})$,
$X_{1i}^* = (X_{1i} - \rho X_{1,i-1})$,
$X_{2i}^* = (X_{2i} - \rho X_{2,i-1})$,
$\cdots$,
$X_{ki}^* = (X_{ki} - \rho X_{k,i-1})$
を新たな変数として，

$$
Y_i^* = \beta_1 X_{1i}^* + \beta_2 X_{2i}^* + \cdots + \beta_k X_{ki}^* + \epsilon_i
$$

最小二乗法を適用する。$\epsilon_1, \epsilon_2, \cdots, \epsilon_n$ は互いに独立とするなので，最小二乗法を適用が可能となる。

ρ の求め方について： DW は近似的に $DW \approx 2(1 - \widehat{\rho})$ と表されるので，DW から ρ の推定値 $\widehat{\rho}$ を逆算して，

$Y_i^* = (Y_i - \widehat{\rho} Y_{i-1})$,

$X^*_{1i} = (X_{1i} - \widehat{\rho} X_{1,i-1})$,
$X^*_{2i} = (X_{2i} - \widehat{\rho} X_{2,i-1})$,
$\cdots$,
$X^*_{ki} = (X_{ki} - \widehat{\rho} X_{k,i-1})$
を新たな変数として，

$$Y^*_i = \beta_1 X^*_{1i} + \beta_2 X^*_{2i} + \cdots + \beta_k X^*_{ki} + \epsilon_i,$$

に最小二乗法を適用する。

8　不均一分散 (不等分散)

8.1　不均一分散 (不等分散) の意味と推定方法

回帰式が

$$Y_i = \alpha + \beta X_i + u_i$$

の場合を考える。X_i が外生変数，Y_i は内生変数，u_i は互いに独立な同一の分布を持つ攪乱項 (最小二乗法に必要な仮定) とする。「独立な同一の分布」の意味は「攪乱項 $u_1, u_2, \cdots, u_n$ はそれぞれ独立に平均ゼロ，分散 σ^2 の分布する」である。

分散が時点に依存する場合，代表的には，分散が他の変数 (例えば，z_i) に依存する場合，すなわち，u_i の平均はゼロ，分散は $\sigma_*^2 z_i^2$ の場合は，最小二乗法の仮定に反する。そのため，単純には，$Y_i = \alpha + \beta X_i + u_i$ に最小二乗法を適用できない。以下のような修正が必要となる。

$$
\begin{aligned}
\frac{Y_i}{z_i} &= \alpha \frac{1}{z_i} + \beta \frac{X_i}{z_i} + \frac{u_i}{z_i} \\
&= \alpha \frac{1}{z_i} + \beta \frac{X_i}{z_i} + u_i^*
\end{aligned}
$$

このとき，新たな攪乱項 u_i^* は平均ゼロ，分散 σ_*^2 の分布となる (すなわち,「同一の」分布)。

$$
\mathrm{E}(u_i^*) = \mathrm{E}\left(\frac{u_i}{z_i}\right) = \left(\frac{1}{z_i}\right)\mathrm{E}(u_i) = 0
$$

u_i の仮定 $\mathrm{E}(u_i) = 0$ が使われている。

$$
\mathrm{V}(u_i^*) = \mathrm{V}\left(\frac{u_i}{z_i}\right) = \left(\frac{1}{z_i}\right)^2 \mathrm{V}(u_i) = \sigma_*^2
$$

u_i の仮定 $\mathrm{V}(u_i) = \sigma_*^2 z_i^2$ が最後に使われている。

よって，$\dfrac{Y_i}{z_i}, \dfrac{1}{z_i}, \dfrac{X_i}{z_i}$ を新たな変数として，最小二乗法を適用することができる。

不均一分散の検定について

$$\widehat{u_i^2} = \gamma z_i^2 + \epsilon_i$$

を推定し，γ の推定値 $\widehat{\gamma}$ の有意性の検定を行う (通常の t 検定)。

z_i は回帰式に含まれる変数でもよい。例えば，u_i の平均はゼロ，分散は $\sigma_*^2 X_i^2$ の場合，各変数を X_i で割って，

$$\begin{aligned}\frac{Y_i}{X_i} &= \alpha\frac{1}{X_i} + \beta + \frac{u_i}{X_i} \\ &= \alpha\frac{1}{X_i} + \beta + u_i^*\end{aligned}$$

を推定すればよい。β は定数項として推定されるが，意味は限界係数 (すなわち，傾き) と同じなので注意すること。

8.2　最小二乗推定量の分散について

単回帰

$$Y_i = \alpha + \beta X_i + u_i, \qquad n = 1, 2, \cdots, n$$

仮定：$\mathrm{E}(u_i) = 0$

$\mathrm{V}(u_i) = \mathrm{E}(u_i^2) = \sigma_i^2 \longleftarrow$ この仮定追加

$i \neq j$ について，$\mathrm{Cov}(u_i, u_j) = \mathrm{E}(u_i u_j) = 0$

不均一分散を無視して，通常の最小二乗推定量は,

$$\widehat{\beta} = \sum_i \omega_i Y_i = \beta + \sum_i \omega_i u_i$$

である。ただし，$\omega_i = \dfrac{X_i - \overline{X}}{\sum_j (X_j - \overline{X})^2}$

$\mathrm{E}(\widehat{\beta})$ について,

$$\begin{aligned}\mathrm{E}(\widehat{\beta}) &= \mathrm{E}(\beta + \sum_i \omega_i u_i) \\ &= \beta + \sum_i \omega_i \mathrm{E}(u_i) = \beta\end{aligned}$$

$u_1, u_2, \cdots, u_n$ の分散が不均一であっても，$\widehat{\beta}$ は不偏推定量となる。

$\mathrm{V}(\widehat{\beta})$ について，

$$
\begin{aligned}
\mathrm{V}(\widehat{\beta}) &= \mathrm{V}(\beta + \sum_i \omega_i u_i) = \mathrm{V}(\sum_i \omega_i u_i) \\
&= \mathrm{E}((\sum_i \omega_i u_i)^2) \quad \longleftarrow \text{分散の定義} \\
&= \mathrm{E}((\sum_i \omega_i u_i)(\sum_j \omega_j u_j)) \quad \longleftarrow \text{添字一つ変更} \\
&= \mathrm{E}(\sum_i \sum_j \omega_i \omega_j u_i u_j) \\
&= \sum_i \sum_j \omega_i \omega_j \mathrm{E}(u_i u_j) \\
&= \sum_i \omega_i^2 \mathrm{E}(u_i^2) + \sum_i \sum_{\substack{j \\ i \neq j}} \omega_i \omega_j \mathrm{E}(u_i u_j) \\
&= \sum_i \sigma_i^2 \omega_i^2 \neq \sigma^2 \sum_i \omega_i^2
\end{aligned}
$$

したがって，$u_1, u_2, \cdots, u_n$ の分散が不均一であるとき，通常の最小二乗推定量 $\widehat{\beta}$ の分散の推定量は，$s_i^2 \sum_i \omega_i^2$ とならなければならない。

s_i^2 は σ_i^2 の推定量とする。

しかし，計量ソフトは $s^2 \sum_i \omega_i^2$ と計算する。

9　多重共線性について

回帰式が

$$Y_i = \alpha W_i + \beta X_i + u_i$$

の場合を考える。W_i, X_i が外生変数，Y_i は内生変数，u_i は互いに独立な攪乱項とする。$W_i = 1$ のとき，α は定数項となる。

W_i と X_i の相関が大きいことを多重共線性が強いと言う。

W_i と X_i の相関が大きい場合は，α, β の推定値は不安定になる。

極端な場合，W_i と X_i の相関が 1 の場合 (完全相関の場合) は，すべての i について，$W_i = \gamma X_i$ となる。この場合，回帰式は

$$\begin{aligned} Y_i &= \alpha W_i + \beta X_i + u_i \\ &= (\alpha\gamma + \beta) X_i + u_i \end{aligned}$$

となり，$\alpha\gamma+\beta$ を推定することは可能だが，α,β を別々に推定することはできなくなる。$Y_i = \alpha W_i + \beta X_i + u_i$ を推定した場合，$\alpha\gamma+\beta$ の推定値が一定値となる $\widehat{\alpha},\widehat{\beta}$ の組み合わせは無数に存在する。この意味で，$\widehat{\alpha},\widehat{\beta}$ は不安定であると言える。

厳密には，最小二乗法によると，

$$\sum_{i=1}^{n} u_i^2 = \sum_{i=1}^{n}(Y_i - \alpha W_i - \beta X_i)^2$$

を最小にする α,β をその推定値 $\widehat{\alpha},\widehat{\beta}$ とする。

すなわち，

$$\frac{\partial \sum u_i^2}{\partial \alpha} = -2\sum_{i=1}^{n}(Y_i - \alpha W_i - \beta X_i)W_i = 0$$

$$\frac{\partial \sum u_i^2}{\partial \beta} = -2\sum_{i=1}^{n}(Y_i - \alpha W_i - \beta X_i)X_i = 0$$

の連立方程式を解くことになる。

$$\sum_{i=1}^{n} Y_i W_i - \widehat{\alpha}\sum_{i=1}^{n} W_i^2 - \widehat{\beta}\sum_{i=1}^{n} X_i W_i = 0$$

$$\sum_{i=1}^{n} Y_i X_i - \widehat{\alpha} \sum_{i=1}^{n} W_i X_i - \widehat{\beta} \sum_{i=1}^{n} X_i^2 = 0$$

行列表示により，

$$\begin{pmatrix} \sum Y_i W_i \\ \sum Y_i X_i \end{pmatrix} = \begin{pmatrix} \sum W_i^2 & \sum X_i W_i \\ \sum W_i X_i & \sum X_i^2 \end{pmatrix} \begin{pmatrix} \widehat{\alpha} \\ \widehat{\beta} \end{pmatrix}$$

$\widehat{\alpha}, \widehat{\beta}$ について表すと，

$$\begin{pmatrix} \widehat{\alpha} \\ \widehat{\beta} \end{pmatrix} = \begin{pmatrix} \sum W_i^2 & \sum X_i W_i \\ \sum W_i X_i & \sum X_i^2 \end{pmatrix}^{-1} \begin{pmatrix} \sum Y_i W_i \\ \sum Y_i X_i \end{pmatrix}$$

逆行列を計算して，

$$\begin{pmatrix} \widehat{\alpha} \\ \widehat{\beta} \end{pmatrix} = \frac{1}{(\sum W_i^2)(\sum X_i^2) - (\sum W_i X_i)^2} \times \begin{pmatrix} \sum X_i^2 & -\sum X_i W_i \\ -\sum W_i X_i & \sum W_i^2 \end{pmatrix} \begin{pmatrix} \sum Y_i W_i \\ \sum Y_i X_i \end{pmatrix}$$

完全な多重共線性の場合 ($W_i = \gamma X_i$ の場合)，

$$(\sum W_i^2)(\sum X_i^2) - (\sum W_i X_i) = 0$$

となる。

また，

$$
\begin{aligned}
\mathrm{V}\begin{pmatrix}\widehat{\alpha}\\ \widehat{\beta}\end{pmatrix} &= \begin{pmatrix}\mathrm{V}(\widehat{\alpha}) & \mathrm{Cov}(\widehat{\alpha},\widehat{\beta})\\ \mathrm{Cov}(\widehat{\alpha},\widehat{\beta}) & \mathrm{V}(\widehat{\beta})\end{pmatrix}\\
&= \sigma^2\begin{pmatrix}\sum W_i^2 & \sum X_iW_i\\ \sum W_iX_i & \sum X_i^2\end{pmatrix}^{-1}\\
&= \frac{\sigma^2}{(\sum W_i^2)(\sum X_i^2)-(\sum W_iX_i)^2}\\
&\quad\times\begin{pmatrix}\sum X_i^2 & -\sum X_iW_i\\ -\sum W_iX_i & \sum W_i^2\end{pmatrix}
\end{aligned}
$$

となるので，完全な多重共線性の場合は，推定値の分散が無限大となる。推定値の分散が無限大という意味は，どこにパラメータがあるか分からないということを意味する。

簡単化のため，$\overline{W} = \dfrac{1}{n}\sum W_i = 0$，$\overline{X} = \dfrac{1}{n}\sum X_i = 0$ とする。W_i と X_i との相関係数を r とすると，

$$
r = \frac{\sum(W_i-\overline{w})(X_i-\overline{X})}{\sqrt{\sum(W_i-\overline{W})^2\sum(X_i-\overline{X})^2}}
$$

$$= \frac{\sum W_i X_i}{\sqrt{\sum W_i^2 \sum X_i^2}}$$

となる。さらに，r を用いて，$\mathrm{V}(\widehat{\alpha})$，$\mathrm{V}(\widehat{\beta})$ を求めると，

$$\begin{aligned}\mathrm{V}(\widehat{\alpha}) &= \frac{\sigma^2 \sum X_i^2}{(\sum W_i^2)(\sum X_i^2) - (\sum W_i X_i)^2} \\ &= \frac{\sigma^2}{(1 - r^2) \sum W_i^2}\end{aligned}$$

$$\begin{aligned}\mathrm{V}(\widehat{\beta}) &= \frac{\sigma^2 \sum W_i^2}{(\sum W_i^2)(\sum X_i^2) - (\sum W_i X_i)^2} \\ &= \frac{\sigma^2}{(1 - r^2) \sum X_i^2}\end{aligned}$$

が得られる。

これは，r が 1 または -1 に近づくにつれて (または，r^2 が 1 に近づくにつれて)，$\mathrm{V}(\widehat{\alpha})$，$\mathrm{V}(\widehat{\beta})$ は大きくなるということを意味する。

⟹ 係数の推定値の有意性が低くなる。

⟹ 本来は W_i や X_i が Y_i に影響を与えているにもかかわらず，統計的に有意な推定値は得られなくなるので，回帰分析によって理論モデルを立証しようという試みは成功しなくなる。

多重共線性の症状： 多重共線性が起こっていると考えられるケースは，

1. 推定値の符号が理論と合わない。
2. 決定係数 (R^2 や $\overline{R}^2$) は大きいのに，個々の t 値は小さい。
3. 観測値の数 (データ数) を少し増やすと，推定値が大きく変わる。
4. 説明変数を増減すると，推定値が大きく変動する。

等である。

10　F 検定について

複数の線形制約の検定を行う場合に F 検定が用いられる。

10.1　いくつかの例

例 1：コブ＝ダグラス型生産関数：　Q_i は生産量，K_i は資本，L_i は労働とする。生産関数を推定する。

$$\log(Q_i) = \beta_1' + \beta_2 \log(K_i) + \beta_3 \log(L_i) + u_i,$$

において，一次同時の制約 $\beta_2 + \beta_3 = 1$ を検定したい。すなわち，帰無仮説，対立仮説は以下のように表される。

帰無仮説 H_0：$\beta_2 + \beta_3 = 1$,
対立仮説 H_1：$\beta_2 + \beta_3 \neq 1$,

例 2：構造変化の検定：　n_0 期以前と $n_0 + 1$ 期以降とで経済構造が変化したと考えて推定を行う。しかも，定数項，傾き共に変化したと想定した場合，回帰式は以下のようになる。

$$Y_i = \alpha + \beta X_i + \gamma D_i + \delta D_i X_i + u_i,$$

ただし，

$$D_i = \begin{cases} 0, & i = 1, 2, \cdots, n_0 \text{ のとき}, \\ 1, & i = n_0 + 1, n_0 + 2, \cdots, n \text{ のとき}, \end{cases}$$

とする。構造変化が n_0+1 期で起こったかどうかを検定したい。すなわち，帰無仮説，対立仮説は以下のように表される。

帰無仮説 H_0： $\gamma=\delta=0$,
対立仮説 H_1： $\gamma\neq 0$，または，$\delta\neq 0$,

例 3：多重回帰モデルの係数の同時検定：　2 つの説明変数が含まれる場合を考える。

$$Y_i=\alpha+\beta X_i+\gamma Z_i+u_i,$$

のモデルにおいて，X_i と Z_i のどちらも，Y_i に影響を与えていないという仮説を検定したい。この場合，帰無仮説，対立仮説は以下のように表される。

帰無仮説 H_0： $\beta=\gamma=0$,
対立仮説 H_1： $\beta\neq 0$，または，$\gamma\neq 0$,

10.2　統計学の復習

$U\sim\chi^2(n)$，$V\sim\chi^2(m)$，U と V は独立とする。
このとき，

$$F=\frac{U/n}{V/m}\sim F(n,m)$$