2017/06/02

藪友良

ガウス・マルコフの定理

既に確認した通り、OLS 推定量は、普遍性を満たした推定量です。実は、普遍性を満たすパラメータの推定法は多数存在します。ガウス・マルコフの定理は、OLS 推定量が普遍性を満たす線形推定量の中で最も分散が小さいこと、を示しています。つまり、OLS 推定量は、普遍性を満たす線形推定量の中で最も優れた推定量なのです。

> **ガウス・マルコフの定理**では、仮定 1、3、4、5 のもとで、OLS 推定量が**線形不偏推定量**(Lienar Unbiased Estimator)の中で最も分散が小さいことを示している。このため、OLS 推定量は**最良線形不偏推定量**（Best Linear Unbiased Estimator: **BLUE**）と呼ばれる。

ガウス・マルコフの定理は、OLS 推定量を用いることの根拠の 1 つとなっています。しかし、仮定 1、3、4、5 が成立しなければ、この定理も成立しないことに注意してください。

[証明]　ここで、Y_1、Y_2、…、Y_n の一次結合で表される不偏推定量を考えます。まず、線形推定量とは何でしょうか。計量経済学では、線形推定量とは、被説明変数 Y_1、Y_2、…、Y_n の一次結合、つまり

$$\beta^* = \sum_{i=1}^{n} d_i Y_i \tag{1}$$

と定義されます。ただし、d_i は加重（ウェイト）となります（加重は説明変数 X の関数であってもかまいません)。

ここで(1)式に $Y_i=\alpha+\beta X_i+u_i$ を代入すると、線形推定量は

$$\beta^* = \sum_{i=1}^{n} d_i(\alpha+\beta X_i+u_i) = \alpha\sum_{i=1}^{n} d_i + \beta\sum_{i=1}^{n} d_i X_i + \sum_{i=1}^{n} d_i u_i \tag{2}$$

とも表せます。不偏推定量とは、$E[\beta^*]=\beta$ となる推定量です。(2)式の期待値をとると、

$$E[\beta^*]=\alpha\sum_{i=1}^{n}d_i+\beta\sum_{i=1}^{n}d_iX_i+E\left[\sum_{i=1}^{n}d_iu_i\right]$$

$$=\alpha\sum_{i=1}^{n}d_i+\beta\sum_{i=1}^{n}d_iX_i$$

となります。式展開では、仮定 1（X_i は非確率変数）、仮定 3（$E[u_\mathrm{i}]=0$）から $E[\Sigma d_iu_\mathrm{i}]=\Sigma d_iE[u_\mathrm{i}]=0$ であること、を用いました。したがって、

① $\Sigma d_i=0$、② $\Sigma d_iX_i=1$

が成立すれば、普遍性が満たされることになります。制約①②を課すことで、線形推定量(2)は、線形不偏推定量となり、

$$\beta^*=\beta+\sum_{i=1}^{n}d_iu_i \tag{3}$$

と表せます[1]。

OLS 推定量は、線形不偏推定量です。ここで、

$$c_i=\frac{(X_i-\overline{X})}{\sum_{j=1}^{n}(X_j-\overline{X})^2} \tag{4}$$

と定義しましょう。このとき、OLS 推定量の定義から、

$$\hat{\beta}=\frac{\sum_{i=1}^{n}(X_i-\overline{X})(Y_i-\overline{Y})}{\sum_{j=1}^{n}(X_j-\overline{X})^2}=\frac{\sum_{i=1}^{n}(X_i-\overline{X})Y_i}{\sum_{j=1}^{n}(X_j-\overline{X})^2}$$

$$=\sum_{i=1}^{n}\left[\frac{(X_i-\overline{X})}{\sum_{j=1}^{n}(X_j-\overline{X})^2}\right]Y_i=\sum_{i=1}^{n}c_iY_i$$

となります（OLS 推定量は他の推定量と区別するため、加重として d_i ではなく c_i という記号を用いました）。読者は、$\Sigma c_i=0$、$\Sigma c_iX_i=1$ が成立することを確認してください。

線形不偏推定量の分散は、(3)式から、

$$Var(\beta^*)=E\left[(\beta^*-\beta)^2\right]=E\left[\left(\sum_{i=1}^{n}d_iu_i\right)^2\right]$$

1 たとえば、β の推定量$(Y_2-Y_1)/(X_2-X_1)$は線形不偏推定量です（練習問題 11.5 参照）。これは $d_1=-1/(X_2-X_1)$、$d_2=1/(X_2-X_1)$、それ以外の $d_i=0$ とします。このとき、$\Sigma d_i=0$、$\Sigma d_iX_i=1$ が成立することは簡単に確認できます。

となります。ここで仮定 4、5 から

$$E\left[\left(\sum_{i=1}^{n} d_i u_i\right)^2\right] = \sigma^2 \sum_{i=1}^{n} d_i^2$$

となります。さらに、$d_i = d_i - c_i + c_i$ となることから、

$$\sigma^2 \sum_{i=1}^{n} d_i^2 = \sigma^2 \sum_{i=1}^{n} \{(d_i - c_i) + c_i\}^2 = \sigma^2 \sum_{i=1}^{n} \{(d_i - c_i)^2 + c_i^2 + 2(d_i - c_i)c_i\}$$

とも表せます。まず、上式の第 3 項は 0 になることを確認しましょう。制約①②から

$$\sum_{i=1}^{n} c_i d_i = \frac{\sum_{i=1}^{n}(X_i - \overline{X})d_i}{\sum_{j=1}^{n}(X_j - \overline{X})^2} = \frac{\sum_{i=1}^{n} d_i X_i - \overline{X}\sum_{i=1}^{n} d_i}{\sum_{j=1}^{n}(X_j - \overline{X})^2}$$

$$= \frac{1}{\sum_{j=1}^{n}(X_j - \overline{X})^2} = \sum_{i=1}^{n} c_i^2$$

となります[2]。したがって、第 3 項は

$$\sum_{i=1}^{n} \{(d_i - c_i)c_i\} = \sum_{i=1}^{n} c_i d_i - \sum_{i=1}^{n} c_i^2 = 0$$

以上から、線形不偏推定量の分散は、

$$\sigma^2 \sum_{i=1}^{n} d_i^2 = \sigma^2 \sum_{i=1}^{n} \{(d_i - c_i)^2 + c_i^2\}$$

$$= \sigma^2 \sum_{i=1}^{n} c_i^2 + \sigma^2 \sum_{i=1}^{n} (d_i - c_i)^2$$

と表せます。第 1 項、第 2 項ともに 2 乗和ですから 0 以上となります。ここで、第 1 項は d_i に依存しません。これに対し、第 2 項は、線形不偏推定量が OLS 推定量であるとき（$d_i = c_i$）最小となります。したがって、線形不偏推定量は、OLS 推定量であるとき分散は最小となるのです。 [終]

---

2 最後の等式は、詳しくは

$$\sum_{i=1}^{n} c_i^2 = \sum_{i=1}^{n}\left[\frac{(X_i - \overline{X})}{\sum_{j=1}^{n}(X_j - \overline{X})^2}\right]^2 = \frac{\sum_{i=1}^{n}(X_i - \overline{X})^2}{\left[\sum_{j=1}^{n}(X_j - \overline{X})^2\right]^2} = \frac{1}{\sum_{j=1}^{n}(X_j - \overline{X})^2}$$

から確認できます。