

2

2010 年度 学部 エコノメトリックス 第 5 回講義メモ

2010 年 10 月 19 日

2.4 最小 2 乗推定量の性質

ここでは、推定量の「良さ」の基準に照らして、最小 2 乗推定量が持つ性質を見てゆくことにしよう。

推定量の「良さ」の基準

いま、パラメータ（母数） θ の大きさ（sample size） n の標本データに基づく推定量を $\hat{\theta}_n$ で表すことにする。統計学では、こういった推定量の「良さ」を表すものとして、次の 3 つの基準を考える。

- (1) 不偏性 (unbiasness)
- (2) 一致性 (consistency)

3

(3) 効率性 (efficiency)

実は最小 2 乗推定量は、上記の 3 つをすべて満たしている推定量である。以下では、3 つの基準を説明しながら、最小 2 乗推定量の性質を見てゆくことにしよう。

(1) 不偏性

不偏性とは、推定量の期待値がパラメータの真値に等しい、すなわち、

$$E(\hat{\theta}_n) = \theta \quad \forall n$$

であることをいう。

第 3 回の授業で、単回帰式 $Y_i = \alpha + \beta X_i + \epsilon_i$ における β の最小 2 乗

4

推定量 b の期待値が、

$$E(b) = E\left(\sum_{i=1}^n w_i Y_i\right) = \beta + E\left(\sum_{i=1}^n w_i \epsilon_i\right) = \beta$$

となることを示した。したがって、最小 2 乗推定量は不偏性を満たしていることになる。

(2) 一致性

一致性とは、 $n \rightarrow \infty$ のとき、推定量がパラメータの真値に収束することをいう。 $\lim_{n \rightarrow \infty} \hat{\theta}_n = \theta$ と表現したいところであるが、これは**正しくない**。実は、推定量は確率変数であるので、数列の極限と違った概念、「確率変数列の収束」(stochastic convergence) 概念を導入する必要がある。

5

定義 2.4.1 確率収束 (convergence in probability)

$Z_1, Z_2, \dots$ を確率変数列（確率変数の数列）であるものとする。このとき、確率変数 S について、

$$\lim_{n \rightarrow \infty} P(|Z_n - S| > \epsilon) = 0 \quad \forall \epsilon > 0$$

が成立するとき、 Z_n ($n = 1, 2, \dots$) は S に確率収束するといい、

$$\text{plim}_{n \rightarrow \infty} Z_n = S$$

または

$$Z_n \xrightarrow{p} S \quad \text{as } n \rightarrow \infty$$

と表す。また S を Z_n の確率極限 (stochastic limit) ともいう。

6

定数は分散ゼロの確率変数として考えることもできるので、この定義は S が定数の場合を含んでいる。したがって、推定量が一致性を満たすとは、

$$\text{plim}_{n \rightarrow \infty} \hat{\theta}_n = \theta$$

で表される。

最小 2 乗推定量の一致性を示すためには、次の結果を利用する必要がある。

7

定理 2.4.2

Z_n, S_n ($n = 1, 2, \dots$) を確率変数列、 γ, δ を定数とする。このとき、 $\text{plim}_{n \rightarrow \infty} Z_n = \gamma, \text{plim}_{n \rightarrow \infty} S_n = \delta$ であるなら、

- (a) $\text{plim}_{n \rightarrow \infty} (Z_n \pm S_n) = \gamma \pm \delta$
- (b) $\text{plim}_{n \rightarrow \infty} (Z_n S_n) = (\text{plim}_{n \rightarrow \infty} Z_n) (\text{plim}_{n \rightarrow \infty} S_n) = \gamma \delta$
- (c) もし $\delta \neq 0$ であるならば、

$$\text{plim}_{n \rightarrow \infty} \left(\frac{Z_n}{S_n} \right) = \frac{\text{plim}_{n \rightarrow \infty} Z_n}{\text{plim}_{n \rightarrow \infty} S_n} = \frac{\gamma}{\delta}$$

が成立する。

さて、 β の最小 2 乗推定量 b_n の確率極限を求めることにしよう。仮定 1～仮定 4 が成立し、また X_i について

$$\frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^2 \longrightarrow Q < +\infty \quad \text{as } n \rightarrow \infty$$

8

が成り立てば、大数の (弱) 法則より

$$\frac{1}{n} \sum_{i=1}^n (X_i - \bar{X}) \epsilon_i \xrightarrow{p} E((X_i - \bar{X}) \epsilon_i) = 0 \quad \text{as } n \rightarrow \infty$$

が成立するので、

$$\begin{aligned} \text{plim}_{n \rightarrow \infty} b_n &= \beta + \text{plim}_{n \rightarrow \infty} \sum_{i=1}^n w_i \epsilon_i \\ &= \beta + \text{plim}_{n \rightarrow \infty} \left(\frac{\frac{1}{n} \sum_{i=1}^n (X_i - \bar{X}) \epsilon_i}{\frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^2} \right) \\ &= \beta + \frac{\text{plim}_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X}) \epsilon_i}{\text{plim}_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^2} \\ &= \beta \end{aligned}$$

が導かれ、最小 2 乗推定量は一致性を満たしていることが確認できる。

9

(3) 効率性

効率性とは、推定量の精度に関する概念の一つである。「精度が高い」ということを、「データの情報を無駄にしていない」とこと読み変え、「効率」をその意味で捉えている。

統計学では、推定量の精度を平均 2 乗誤差 (mean squared error) で考える。

$$\text{MSE}(\hat{\theta}_n) \equiv E[(\hat{\theta}_n - \theta)^2]$$

平均 2 乗誤差は、

$$\begin{aligned} \text{MSE}(\hat{\theta}_n) &= E \left[\left((\hat{\theta}_n - E(\hat{\theta})) + (E(\hat{\theta}) - \theta) \right)^2 \right] \\ &= \text{Var}(\hat{\theta}_n) + (E(\hat{\theta}) - \theta)^2 \end{aligned}$$

10

のように、「推定量の分散」と「推定量の偏りの 2 乗」の和で表すことができる。

したがって、不偏推定量においては、平均 2 乗誤差が最小となる (すなわち最も精度の高い) 推定量は、分散が最小となる推定量に一致する。これが、効率性という概念である。

最小 2 乗推定量が、線形推定量で不偏であるもののうち、最も分散が小さいもの (最良線形不偏推定量 (BLUE: best linear unbiased estimator)) であることを示したのが Gauss=Markov 定理である。

定理 2.4.3 Gauss=Markov 定理

β の最小 2 乗推定量 b_n は最良線形不偏推定量である。

[証明]

いま、 $\tilde{b}_n$ を任意の線形推定量であるものとする。 $\tilde{b}_n$ は線形推定量である

11

から $c_i \equiv w_i + d_i$ を用いて、 $\tilde{b}_n = \sum_{i=1}^n c_i Y_i$ で表現できる。よって、

$$\begin{aligned} \tilde{b}_n &= \sum_{i=1}^n c_i (\alpha + \beta X_i + \epsilon_i) \\ &= \alpha \sum_{i=1}^n c_i + \beta \sum_{i=1}^n c_i X_i + \sum_{i=1}^n c_i \epsilon_i \end{aligned}$$

になる。いま、 $\tilde{b}_n$ が不偏推定量であるならば、

$$\sum_{i=1}^n c_i = 0 \quad \text{かつ} \quad \sum_{i=1}^n c_i X_i = 1 \quad (2.4.1)$$

が成立していなければならない。ところで、最小 2 乗推定量の性質より、 $\sum_{i=1}^n w_i = 0$ かつ $\sum_{i=1}^n w_i X_i = 1$ という関係が成立しているから、

12

(2.4.1) 式は

$$\sum_{i=1}^n d_i = 0 \quad \text{かつ} \quad \sum_{i=1}^n d_i X_i = 0 \quad (2.4.2)$$

であることを意味する。このとき、 $\tilde{b}_n$ は

$$\begin{aligned} \text{Var}(\tilde{b}_n) &= \sigma^2 \left(\sum_{i=1}^n c_i^2 \right) = \sigma^2 \left(\sum_{i=1}^n (w_i + d_i)^2 \right) \\ &= \sigma^2 \left(\sum_{i=1}^n w_i^2 + 2 \sum_{i=1}^n w_i d_i + \sum_{i=1}^n d_i^2 \right) \\ &= \sigma^2 \left(\sum_{i=1}^n w_i^2 + \sum_{i=1}^n d_i^2 \right) \geq \sigma^2 \sum_{i=1}^n w_i^2 = \text{Var}(b_n) \end{aligned}$$

となり、 b_n の分散が線形不偏推定量の下限であることが示された。(証了)