

経済統計学講義ノート No.11

回帰分析 II

蛭川雅之

2022 年 9 月 12 日

目 次

1. 重回帰モデル
2. 最小 2 乗推定量の性質
3. 推定モデルのあてはまりの尺度
4. 古典的諸仮定
5. 最小 2 乗推定量の統計的特性
6. ガウス = マルコフ定理

1 重回帰モデル

- 前回、単回帰モデル

$$Y = \beta_0 + \beta_1 X + \epsilon$$

およびパラメータ (β_0, β_1) の最小 2 乗推定を学んだ。

- 説明変数を 1 つに限定するのは制約的である。
 - 多くの場合、 $k (\geq 2)$ 個の説明変数からなる**重回帰モデル**

$$Y = \beta_0 + \beta_1 X_1 + \cdots + \beta_k X_k + \epsilon$$

を利用する。

- 重回帰モデルについても、 n 個のデータ

$$\begin{aligned} & \{(Y_i, X_{1i}, \dots, X_{ki})\}_{i=1}^n \\ &= \{(Y_1, X_{11}, \dots, X_{k1}), \dots, (Y_n, X_{1n}, \dots, X_{kn})\} \end{aligned}$$

を使ってパラメータ $(\beta_0, \beta_1, \dots, \beta_k)$ を最小 2 乗推定する。

- 具体的には、残差 2 乗和

$$\sum_{i=1}^n e_i^2 = \sum_{i=1}^n (Y_i - b_0 - b_1 X_{1i} - \dots - b_k X_{ki})^2$$

を最小にする $(b_0, b_1, \dots, b_k)$ を求める。

- そのような $(b_0, b_1, \dots, b_k)$ を $(\beta_0, \beta_1, \dots, \beta_k)$ の最小 2 乗推定量 $(\hat{\beta}_0, \hat{\beta}_1, \dots, \hat{\beta}_k)$ と定義する。
- 単回帰モデルの場合と異なり、行列を用いなければ $(\hat{\beta}_0, \hat{\beta}_1, \dots, \hat{\beta}_k)$ を明示的に表現するのは困難である。

2 最小 2 乗推定量の性質

2.1 予測値・残差の性質

1. 残差

$$e_i = Y_i - \hat{Y}_i \quad (i = 1, \dots, n)$$

の和はゼロである。

$$\sum_{i=1}^n e_i = 0$$

2. 各説明変数と残差との積の和はゼロである。

$$\sum_{i=1}^n X_{1i}e_i = \dots = \sum_{i=1}^n X_{ki}e_i = 0$$

3. 回帰直線

$$Y = \hat{\beta}_0 + \hat{\beta}_1 X_1 + \cdots + \hat{\beta}_k X_k$$

は点 $(\bar{X}_1, \dots, \bar{X}_k, \bar{Y})$ を通る。

4. 予測値

$$\hat{Y}_i = \hat{\beta}_0 + \hat{\beta}_1 X_{1i} + \cdots + \hat{\beta}_k X_{ki} \quad (i = 1, \dots, n)$$

の平均 $\overline{\hat{Y}}$ は被説明変数の平均 $\bar{Y}$ に等しい。

2.2 分散分解

- 以下の表記を導入する。

1. 平均回りの被説明変数 Y の偏差 2 乗和を全変動 (Total Sum of Squares ; 略称 TSS) と呼ぶ。

$$TSS = \sum_{i=1}^n (Y_i - \bar{Y})^2$$

2. 平均回りの予測値 $\hat{Y}$ の偏差 2 乗和を回帰変動 (Explained Sum of Squares ; 略称 ESS) と呼ぶ。

$$ESS = \sum_{i=1}^n (\hat{Y}_i - \bar{\hat{Y}})^2 = \sum_{i=1}^n (\hat{Y}_i - \bar{Y})^2$$

3. 残差 e の 2 乗和を**残差変動** (Residual Sum of Squares ; 略称 **RSS**) と呼ぶ。

$$RSS = \sum_{i=1}^n e_i^2$$

- これら 3 つの 2 乗和には次のような関係がある。

$$\underbrace{\sum_{i=1}^n (Y_i - \bar{Y})^2}_{TSS} = \underbrace{\sum_{i=1}^n (\hat{Y}_i - \bar{Y})^2}_{ESS} + \underbrace{\sum_{i=1}^n e_i^2}_{RSS}$$
$$TSS = ESS + RSS$$

- この関係を**分散分解**という。

問題 1 “ln11.xlsx”のデータはノート No.9 と同時に配布したデータファイル“phillips.csv”と同一である。このとき、以下の点を確認せよ。

1. ノート No.9 結論 5 に基づいて計算した最小 2 乗推定値 $(\hat{\beta}_0, \hat{\beta}_1)$ が同ノートにある回帰分析の結果（直線）と一致する。
2. 予測値・残差の性質が成り立つ。
3. 分散分解が成り立つ。

3 推定モデルのあてはまりの尺度

3.1 重相関係数と決定係数

- 推定されたモデルのあてはまりの尺度として、**重相関係数 R** および**決定係数 R^2** が広く用いられる。
- R は被説明変数 Y とその予測値 $\hat{Y}$ に関するピアソン積率相関係数である。

$$R = r_{Y\hat{Y}} = \frac{\sum_{i=1}^n (Y_i - \bar{Y}) (\hat{Y}_i - \bar{\hat{Y}})}{\sqrt{\sum_{i=1}^n (Y_i - \bar{Y})^2} \sqrt{\sum_{i=1}^n (\hat{Y}_i - \bar{\hat{Y}})^2}}$$

- R^2 は回帰変動 (ESS) の全変動 (TSS) に対する比率である。

$$R^2 = \frac{ESS}{TSS} = 1 - \frac{RSS}{TSS}$$

■ R および R^2 の持つ性質

1. $0 \leq R, R^2 \leq 1$ である。
2. 決定係数 R^2 は重相関係数 R の 2 乗に等しい。
3. 単回帰モデルに限り、 (X, Y) のピアソン積率相関係数を r_{XY} とすると、

$$R^2 = r_{XY}^2$$

が成り立つ。

3.2 自由度修正済み決定係数

- 説明変数を追加すると、一般に R^2 は増加する（少なくとも減少はしない）。
 - － 無意味な説明変数を回帰モデルに追加しても R^2 は‘改善’する。
- 実際には、説明変数の追加 ($k \uparrow$) は自由度の減少 ($n - k - 1 \downarrow$) を伴う。
 - － R^2 を自由度調整した **自由度修正済み決定係数 $\bar{R}^2$**

$$\bar{R}^2 = 1 - \frac{RSS/(n - k - 1)}{TSS/(n - 1)} = 1 - (1 - R^2) \left(\frac{n - 1}{n - k - 1} \right)$$

が提案されている。

■ $\bar{R}^2$ の持つ性質

1. $\bar{R}^2$ の上限は 1 であるが、下限はゼロでなく負値も取りうる。
2. 同一の回帰モデルに対し、 $\bar{R}^2 \leq R^2$ である。
3. 回帰モデルに説明変数を追加すると、 $\bar{R}^2$ は増加する場合もあれば、不変もしくは減少する場合もある。

問題 2 データファイル“wage.csv”を使用して、 $Y = \ln(wage)$ を被説明変数とする回帰モデルを推定する。

1. まず $X = educ$ のみを説明変数とする単回帰モデルを最小 2 乗推定し、以下の点を確認せよ。

$$R^2 = \frac{ESS}{TSS} = 1 - \frac{RSS}{TSS}$$

$$R^2 = r_{XY}^2$$

$$\bar{R}^2 = 1 - \frac{RSS/(n - k - 1)}{TSS/(n - 1)} = 1 - (1 - R^2) \left(\frac{n - 1}{n - k - 1} \right)$$

$$\bar{R}^2 \leq R^2$$

2. 次に $(X_1, X_2) = (educ, exper)$ を説明変数とする重回帰モデルを最小 2 乗推定し、 R^2 および $\bar{R}^2$ が 1 で得られたものと比べて大きくなるか確認せよ。

4 古典的諸仮定

1. 回帰モデルは線形で、正しく定式化されており、加法的な誤差項を持つ。
2. 誤差項の平均はゼロである。
3. 説明変数は誤差項と無相関である。
4. 各観測値に付随する誤差項は相互に無相関である。
5. 誤差項の分散は均一である。
6. 説明変数はそれ以外の説明変数と完全な線形関数の関係を持たない
(= 説明変数間に完全な多重共線性は存在しない)。
7. 誤差項は正規分布に従う。

■ コメント

- 仮定 1-5 を満たす誤差項を古典的誤差項と呼ぶ。
 - － 仮定 1-5 に加えて仮定 7 を満たす誤差項を古典的正規誤差項と呼ぶ。
- 仮定 3 が満たされない場合、最小 2 乗推定量は一致性を持たない。
 - － 一致性を確保する対処法（操作変数法）は省略する。
- 仮定 4 または 5 が満たされない場合でも最小 2 乗推定量は一致性を持つ。
 - － ただし、最小 2 乗推定量は有効性を失う。
- 仮定 6 が満たされない場合、そもそも回帰モデルの識別ができない。
 - － 多重共線性の問題はダミー変数を説明する際に再考する。

5 最小 2 乗推定量の統計的特性

- 仮定 1-6 の下で、

$$E\left(\hat{\beta}_j\right) = \beta_j \quad (j = 0, 1, \dots, k) : [\text{不偏性}]$$

および

$$\hat{\beta}_j \xrightarrow{p} \beta_j \quad (j = 0, 1, \dots, k) : [\text{一貫性}]$$

の両方が成り立つ。

- 一貫性は $Var\left(\hat{\beta}_j\right) \rightarrow 0 \quad (j = 0, 1, \dots, k)$ に由来する。
- さらに仮定 7 (誤差項の正規性) が加わると、

$$\hat{\beta}_j \sim N\left(\beta_j, Var\left(\hat{\beta}_j\right)\right) \quad (j = 0, 1, \dots, k)$$

となる。

- ところで、 $Var(\hat{\beta}_j)$ はどのような形をとるのか？
 - 単回帰モデル

$$Y = \beta_0 + \beta_1 X + \epsilon$$

の場合、誤差分散 $\sigma^2 = Var(\epsilon)$ に対し、

$$Var(\hat{\beta}_0) = \sigma^2 \left\{ \frac{1}{n} + \frac{\bar{X}^2}{\sum_{i=1}^n (X_i - \bar{X})^2} \right\}$$

$$Var(\hat{\beta}_1) = \frac{\sigma^2}{\sum_{i=1}^n (X_i - \bar{X})^2}$$

である。

- これらは σ^2 に依存するため、観測できない。

- それでは、単回帰モデルについて $Var(\hat{\beta}_j)$ をどのように推定すればよいか？

– σ^2 を

$$s^2 = \frac{RSS}{n-2} = \frac{1}{n-2} \sum_{i=1}^n e_i^2$$

で推定する。

＊ 重回帰モデルの場合は

$$s^2 = \frac{RSS}{n-k-1} = \frac{1}{n-k-1} \sum_{i=1}^n e_i^2$$

となる。

- この推定量の正の平方根 $s = \sqrt{s^2}$ は、しばしば**回帰式の標準誤差 (Standard Error of Regression ; 略称 SER)** と呼ばれる。

– $Var(\hat{\beta}_0)$ および $Var(\hat{\beta}_1)$ は

$$\widehat{Var}(\hat{\beta}_0) = s^2 \left\{ \frac{1}{n} + \frac{\bar{X}^2}{\sum_{i=1}^n (X_i - \bar{X})^2} \right\}$$
$$\widehat{Var}(\hat{\beta}_1) = \frac{s^2}{\sum_{i=1}^n (X_i - \bar{X})^2}$$

と推定できる。

– これらの正の平方根

$$SE(\hat{\beta}_0) = \sqrt{\widehat{Var}(\hat{\beta}_0)}$$
$$SE(\hat{\beta}_1) = \sqrt{\widehat{Var}(\hat{\beta}_1)}$$

は $\hat{\beta}_0$ および $\hat{\beta}_1$ の標準誤差 (Standard Error ; 略称 SE) と呼ばれる。

■ コメント

- 回帰モデルの推定結果を報告する際には、係数推定値・標準誤差の両方を明示する。
 - － 係数推定値は点推定値、一方、標準誤差は推定の精度を示す。
- 例えば、問題 2-2 の推定結果は次のように報告すればよい。

$$\ln(\widehat{wage}) = \underset{(0.1120)}{5.5027} + \underset{(0.0066)}{0.0778educ} + \underset{(0.0033)}{0.0198exper}$$
$$n = 935, R^2 = 0.1309, \bar{R}^2 = 0.1290$$

6 ガウス＝マルコフ定理

定理 3 仮定 1-6 の下で、回帰モデル

$$Y = \beta_0 + \beta_1 X_1 + \cdots + \beta_k X_k + \epsilon$$

の最小 2 乗推定量 $(\hat{\beta}_0, \hat{\beta}_1, \dots, \hat{\beta}_k)$ は、 $(\beta_0, \beta_1, \dots, \beta_k)$ の全ての線形不偏推定量の中で最小の分散を持つ。

- 線形推定量とは、ある定数 c_i ($i = 1, \dots, n$) を用いて

$$\sum_{i=1}^n c_i Y_i$$

の形に表現できる推定量を指す。

- 最小 2 乗推定量は本当に線形推定量なのか？
 - － 話を簡単にするため、単回帰モデル

$$Y = \beta_0 + \beta_1 X + \epsilon$$

を考える。

- － まず $\hat{\beta}_1$ について、

$$\sum_{i=1}^n (X_i - \bar{X}) (Y_i - \bar{Y}) = \sum_{i=1}^n (X_i - \bar{X}) Y_i$$

から、

$$\hat{\beta}_1 = \frac{\sum_{i=1}^n (X_i - \bar{X}) (Y_i - \bar{Y})}{\sum_{i=1}^n (X_i - \bar{X})^2} = \sum_{i=1}^n \underbrace{\left\{ \frac{X_i - \bar{X}}{\sum_{i=1}^n (X_i - \bar{X})^2} \right\}}_{=c_i} Y_i$$

が成り立つ。

– 次に $\hat{\beta}_0$ について、

$$\begin{aligned}\hat{\beta}_0 &= \bar{Y} - \hat{\beta}_1 \bar{X} \\&= \frac{1}{n} \sum_{i=1}^n Y_i - \left[\sum_{i=1}^n \left\{ \frac{X_i - \bar{X}}{\sum_{i=1}^n (X_i - \bar{X})^2} \right\} Y_i \right] \bar{X} \\&= \sum_{i=1}^n \frac{1}{n} Y_i - \sum_{i=1}^n \left\{ \frac{\bar{X} (X_i - \bar{X})}{\sum_{i=1}^n (X_i - \bar{X})^2} \right\} Y_i \\&= \sum_{i=1}^n \underbrace{\left\{ \frac{1}{n} - \frac{\bar{X} (X_i - \bar{X})}{\sum_{i=1}^n (X_i - \bar{X})^2} \right\}}_{=c_i} Y_i\end{aligned}$$

も成り立つ。

- ガウス＝マルコフ定理の条件にさらに仮定 7（誤差項の正規性）が加わると、最小 2 乗推定量は最小分散不偏推定量となることが知られている。