

# 計量経済分析 I

**Thu., 8:50-10:20**

**Room # 4 (法経講義棟)**

- The prerequisite of this class is **Basic Statistics** (統計基礎) (by Prof. Fukushige, Tue., 16:20-17:50, this semester) and **Econometrics** (エコノメトリックス) (by Prof. Benjamin Poignard, undergraduate level, next semester, 『計量経済学』山本拓著, 新世社).
- The class of **Introductory Econometrics** (計量経済基礎) (by Prof. Benjamin Poignard, 10:30-12:00 on Tue., this semester) should be registered.

# Contents

<b>1</b>	<b>最小二乗法について (復習)</b>	<b>11</b>
1.1	最小二乗法と回帰直線 . . . . .	11
1.2	切片 $\alpha$ と傾き $\beta$ の推定 . . . . .	12
1.3	残差 $\widehat{u}_i$ の性質について . . . . .	21
1.4	決定係数 $R^2$ について . . . . .	23
1.5	まとめ . . . . .	29
<b>2</b>	<b>Regression Analysis (回帰分析)</b>	<b>31</b>
2.1	Setup of the Model . . . . .	31
2.2	Ordinary Least Squares Estimation . . . . .	36
2.3	Properties of Least Squares Estimator . . . . .	40

<b>3</b>	<b>多重回帰</b>	<b>53</b>
3.1	推定量の性質 . . . . .	57
<b>4</b>	<b>系列相関：<i>DW</i> について</b>	<b>64</b>
4.1	<i>DW</i> について . . . . .	64
4.2	系列相関のもとで回帰式の推定 . . . . .	72
<b>5</b>	<b>不均一分散 (不等分散)</b>	<b>77</b>
<b>6</b>	<b>推定量の求め方</b>	<b>80</b>
6.1	最小二乗法 . . . . .	80
6.2	最尤法 . . . . .	82
6.2.1	変数変換 . . . . .	105
6.2.2	回帰分析への応用 . . . . .	107

6.2.3	誤差項に系列相関がある場合 . . . . .	116
6.3	尤度比検定 . . . . .	120
<b>7</b>	<b>Time Series Analysis (時系列分析)</b>	<b>135</b>
7.1	Introduction . . . . .	135
7.2	Time Series Models (時系列モデル) . . . . .	139
7.3	Autoregressive Model (自己回帰モデル or AR モデル) . . . . .	142
7.4	MA Model . . . . .	172
7.5	ARMA Model . . . . .	184
7.6	ARIMA Model . . . . .	189
7.7	SARIMA Model . . . . .	190
7.8	Optimal Prediction . . . . .	191
7.9	Identification (識別, または, 同定) . . . . .	194
7.10	Example of SARIMA using Consumption Data . . . . .	197

7.11	ARCH and GARCH Models . . . . .	203
<b>8</b>	<b>Vector Autoregressive (VAR) Model – Causality, Impulse Response Function and etc</b>	<b>208</b>
8.1	Autocovariance Matrix and Autocorrelation Matrix . . . . .	211
8.2	Granger Causality Test (グレンジャー因果性テスト) . . . . .	212
8.3	Impulse Response Function (インパルス応答関数): . . . . .	219
<b>9</b>	<b>Unit Root (単位根) and Cointegration (共和分)</b>	<b>224</b>
9.1	Unit Root (単位根) Test (Dickey-Fuller (DF) Test) . . . . .	224
9.2	Serially Correlated Errors . . . . .	238
9.2.1	Augmented Dickey-Fuller (ADF) Test . . . . .	238
9.3	Cointegration (共和分) . . . . .	243
9.4	Testing Cointegration . . . . .	253

9.4.1	Engle-Granger Test . . . . .	253
9.4.2	Error Correction Representation . . . . .	255
9.5	Testing the Number of Cointegrating Vectors . . . . .	261
<b>10</b>	<b>GMM (Generalized Method of Moments, 一般化積率法)</b>	<b>263</b>
<b>11</b>	<b>Bayesian Estimation (ベイズ推定)</b>	<b>270</b>
11.1	Introduction . . . . .	271
11.2	Inference . . . . .	275
11.2.1	Point Estimate . . . . .	275
11.2.2	Interval Estimate . . . . .	276
11.2.3	Marginal Likelihood (周辺尤度) . . . . .	276
11.3	Example: Linear Regression . . . . .	277
11.4	On Prior Distribution . . . . .	280

11.4.1	Non-informative Prior . . . . .	280
11.4.2	Jeffreys' Prior . . . . .	281
11.5	Evaluation of Expectation . . . . .	282
11.5.1	Evaluation of Expectation: Numerical Integration . . . . .	283
11.5.2	Evaluation of Expectation: Monte Carlo Integration . . . . .	287
11.5.3	Evaluation of Expectation: Random Number Generation . . . . .	290
11.6	Sampling Method I: Random Number Generation . . . . .	291
11.6.1	Uniform Distribution: $U(0, 1)$ . . . . .	291
11.6.2	Transforming $U(0, 1)$ : Continuous Type . . . . .	300
11.6.3	Inverse Transform Method . . . . .	323
11.6.4	Using $U(0, 1)$ : Discrete Type . . . . .	326

**References** **327**

11.7	Sampling Method II: Random Number Generation . . . . .	332
------	--	-----

11.7.1	Rejection Sampling (棄却法) . . . . .	332
11.7.2	Importance Resampling (重点的リサンプリング) . . . . .	353
11.7.3	Metropolis-Hastings Algorithm (メトロポリスーハスティ ングス・アルゴリズム) . . . . .	361
11.7.4	Ratio-of-Uniforms Method . . . . .	384
11.7.5	Gibbs Sampling . . . . .	398
	<b>References</b>	<b>410</b>
	<b>12 Bayesian Estimation — Examples</b>	<b>417</b>
12.1	Heteroscedasticity Model . . . . .	417
12.1.1	Introduction . . . . .	418
12.1.2	Multiplicative Heteroscedasticity Regression Model . . . . .	420
12.1.3	Bayesian Estimation . . . . .	426

12.1.4	Monte Carlo Study . . . . .	432
12.1.5	Summary . . . . .	450
12.1.6	Appendix: Are $M = 5000$ and $N = 10^4$ Sufficient? . . . . .	454
12.2	Autocorrelation Model . . . . .	456
12.2.1	Introduction . . . . .	458
12.2.2	Setup of the Model . . . . .	459
12.2.3	Monte Carlo Experiments . . . . .	469
12.2.4	Summary . . . . .	484
<b>References</b>		<b>485</b>
12.3	Marginal Likelihood, Convergence Diagnostic and so on . . . . .	490
12.3.1	Marginal Likelihood (周辺尤度) . . . . .	490
12.3.2	Convergence Diagnostic (収束判定) . . . . .	494

代表的テキスト：

・ J.D. Hamilton (1994) *Time Series Analysis*

沖本・井上訳 (2006) 『時系列解析(上・下)』

・ A.C. Harvey (1981) *Time Series Models*

国友・山本訳 (1985) 『時系列モデル入門』

・ 沖本竜義 (2010) 『経済・ファイナンスデータの計量時系列分析』

# 1 最小二乗法について (復習)

経済理論に基づいた線型モデルの係数の値をデータから求める時に用いられる手法  $\Rightarrow$  最小二乗法

## 1.1 最小二乗法と回帰直線

$(X_1, Y_1), (X_2, Y_2), \dots, (X_n, Y_n)$  のように  $n$  組のデータがあり,  $X_i$  と  $Y_i$  との間に以下の線型関係を想定する。

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

$X_i$  は説明変数,  $Y_i$  は被説明変数,  $\alpha, \beta$  はパラメータとそれぞれ呼ばれる。

上の式は回帰モデル (または, 回帰式) と呼ばれる。目的は, 切片  $\alpha$  と傾き  $\beta$  をデータ  $\{(X_i, Y_i), i = 1, 2, \dots, n\}$  から推定すること,

データについて：

1. タイム・シリーズ (時系列) ・データ：  $i$  が時間を表す (第  $i$  期)。
2. クロス・セクション (横断面) ・データ：  $i$  が個人や企業を表す (第  $i$  番目の家計，第  $i$  番目の企業)。

## 1.2 切片 $\alpha$ と傾き $\beta$ の推定

次のような関数  $S(\alpha, \beta)$  を定義する。

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

このとき，

$$\min_{\alpha, \beta} S(\alpha, \beta)$$

となるような  $\alpha, \beta$  を求める (最小自乗法)。このときの解を  $\hat{\alpha}, \hat{\beta}$  とする。

最小化のためには,

$$\frac{\partial S(\alpha, \beta)}{\partial \alpha} = 0$$

$$\frac{\partial S(\alpha, \beta)}{\partial \beta} = 0$$

を満たす  $\alpha, \beta$  が  $\widehat{\alpha}, \widehat{\beta}$  となる。すなわち,  $\widehat{\alpha}, \widehat{\beta}$  は,

$$\sum_{i=1}^n (Y_i - \widehat{\alpha} - \widehat{\beta}X_i) = 0, \quad (1)$$

$$\sum_{i=1}^n X_i(Y_i - \widehat{\alpha} - \widehat{\beta}X_i) = 0, \quad (2)$$

を満たす。さらに,

$$\sum_{i=1}^n Y_i = n\widehat{\alpha} + \widehat{\beta} \sum_{i=1}^n X_i, \quad (3)$$

$$\sum_{i=1}^n X_i Y_i = \widehat{\alpha} \sum_{i=1}^n X_i + \widehat{\beta} \sum_{i=1}^n X_i^2,$$

行列表示によって,

$$\begin{pmatrix} \sum_{i=1}^n Y_i \\ \sum_{i=1}^n X_i Y_i \end{pmatrix} = \begin{pmatrix} n & \sum_{i=1}^n X_i \\ \sum_{i=1}^n X_i & \sum_{i=1}^n X_i^2 \end{pmatrix} \begin{pmatrix} \widehat{\alpha} \\ \widehat{\beta} \end{pmatrix},$$

逆行列の公式:

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}^{-1} = \frac{1}{ad - bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$$

$\widehat{\alpha}, \widehat{\beta}$  について, まとめて,

$$\begin{aligned} \begin{pmatrix} \widehat{\alpha} \\ \widehat{\beta} \end{pmatrix} &= \begin{pmatrix} n & \sum_{i=1}^n X_i \\ \sum_{i=1}^n X_i & \sum_{i=1}^n X_i^2 \end{pmatrix}^{-1} \begin{pmatrix} \sum_{i=1}^n Y_i \\ \sum_{i=1}^n X_i Y_i \end{pmatrix} \\ &= \frac{1}{n \sum_{i=1}^n X_i^2 - (\sum_{i=1}^n X_i)^2} \begin{pmatrix} \sum_{i=1}^n X_i^2 & -\sum_{i=1}^n X_i \\ -\sum_{i=1}^n X_i & n \end{pmatrix} \begin{pmatrix} \sum_{i=1}^n Y_i \\ \sum_{i=1}^n X_i Y_i \end{pmatrix} \end{aligned}$$

さらに,  $\widehat{\beta}$  について解くと,

$$\widehat{\beta} = \frac{n \sum_{i=1}^n X_i Y_i - (\sum_{i=1}^n X_i)(\sum_{i=1}^n Y_i)}{n \sum_{i=1}^n X_i^2 - (\sum_{i=1}^n X_i)^2}$$

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

連立方程式の (3) 式から,

$$\widehat{\alpha} = \bar{Y} - \widehat{\beta}\bar{X}$$

となる。ただし,

$$\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i, \quad \bar{Y} = \frac{1}{n} \sum_{i=1}^n Y_i,$$

とする。

数値例： 以下の数値例を使って，回帰式  $Y_i = \alpha + \beta X_i$  の  $\alpha$ ,  $\beta$  の推定値  $\widehat{\alpha}$ ,  $\widehat{\beta}$  を求める。

$i$	$Y_i$	$X_i$
1	6	10
2	9	12
3	10	14
4	10	16

$\widehat{\alpha}$ ,  $\widehat{\beta}$ を求めるための公式は

$$\widehat{\beta} = \frac{\sum_{i=1}^n X_i Y_i - n \overline{X} \overline{Y}}{\sum_{i=1}^n X_i^2 - n \overline{X}^2}$$
$$\widehat{\alpha} = \overline{Y} - \widehat{\beta} \overline{X}$$

なので、必要なものは  $\overline{X}$ ,  $\overline{Y}$ ,  $\sum_{i=1}^n X_i^2$ ,  $\sum_{i=1}^n X_i Y_i$  である。

$i$	$Y_i$	$X_i$	$X_i Y_i$	$X_i^2$
1	6	10	60	100
2	9	12	108	144
3	10	14	140	196
4	10	16	160	256
合計	$\sum Y_i$ 35	$\sum X_i$ 52	$\sum X_i Y_i$ 468	$\sum X_i^2$ 696
平均	$\bar{Y}$ 8.75	$\bar{X}$ 13		

よって,

$$\hat{\beta} = \frac{468 - 4 \times 13 \times 8.75}{696 - 4 \times 13^2} = \frac{13}{20} = 0.65$$

$$\hat{\alpha} = 8.75 - 0.65 \times 13 = 0.3$$

となる。

注意事項：

1.  $\alpha, \beta$  は真の値で未知
2.  $\widehat{\alpha}, \widehat{\beta}$  は  $\alpha, \beta$  の推定値でデータから計算される

回帰直線は

$$\widehat{Y}_i = \widehat{\alpha} + \widehat{\beta}X_i,$$

として与えられる。

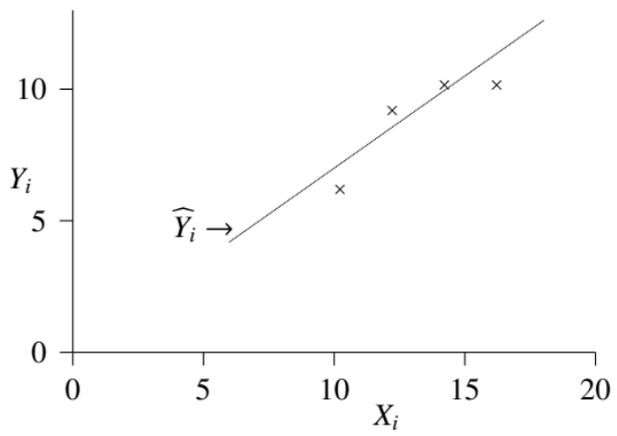
上の数値例では、

$$\widehat{Y}_i = 0.3 + 0.65X_i$$

となる。

$i$	$Y_i$	$X_i$	$X_i Y_i$	$X_i^2$	$\widehat{Y}_i$
1	6	10	60	100	6.8
2	9	12	108	144	8.1
3	10	14	140	196	9.4
4	10	16	160	256	10.7
合計	$\Sigma Y_i$	$\Sigma X_i$	$\Sigma X_i Y_i$	$\Sigma X_i^2$	$\Sigma \widehat{Y}_i$
	35	52	468	696	35.0
平均	$\bar{Y}$	$\bar{X}$			
	8.75	13			

図 2 :  $Y_i, X_i, \widehat{Y}_i$



$\widehat{Y}_i$  を実績値  $Y_i$  の予測値または理論値と呼ぶ。

$$\widehat{u}_i = Y_i - \widehat{Y}_i,$$

$\widehat{u}_i$  を残差と呼ぶ。

$$Y_i = \widehat{Y}_i + \widehat{u}_i = \widehat{\alpha} + \widehat{\beta}X_i + \widehat{u}_i,$$

さらに、 $\bar{Y}$  を両辺から引いて、

$$(Y_i - \bar{Y}) = (\widehat{Y}_i - \bar{Y}) + \widehat{u}_i,$$

### 1.3 残差 $\widehat{u}_i$ の性質について

$\widehat{u}_i = Y_i - \widehat{\alpha} - \widehat{\beta}X_i$  に注意して、(1) 式から、

$$\sum_{i=1}^n \widehat{u}_i = 0,$$

を得る。(2) 式から、

$$\sum_{i=1}^n X_i \widehat{u}_i = 0,$$

を得る。  $\widehat{Y}_i = \widehat{\alpha} + \widehat{\beta}X_i$  から,

$$\sum_{i=1}^n \widehat{Y}_i \widehat{u}_i = 0,$$

を得る。なぜなら,

$$\begin{aligned} \sum_{i=1}^n \widehat{Y}_i \widehat{u}_i &= \sum_{i=1}^n (\widehat{\alpha} + \widehat{\beta}X_i) \widehat{u}_i \\ &= \widehat{\alpha} \sum_{i=1}^n \widehat{u}_i + \widehat{\beta} \sum_{i=1}^n X_i \widehat{u}_i \\ &= 0 \end{aligned}$$

である。

$i$	$Y_i$	$X_i$	$\widehat{Y}_i$	$\widehat{u}_i$	$X_i\widehat{u}_i$	$\widehat{Y}_i\widehat{u}_i$
1	6	10	6.8	-0.8	-8.0	-5.44
2	9	12	8.1	0.9	10.8	7.29
3	10	14	9.4	0.6	8.4	5.64
4	10	16	10.7	-0.7	-11.2	-7.49
合計	$\sum Y_i$	$\sum X_i$	$\sum \widehat{Y}_i$	$\sum \widehat{u}_i$	$\sum X_i\widehat{u}_i$	$\sum \widehat{Y}_i\widehat{u}_i$
	35	52	35.0	0.0	0.0	0.00

## 1.4 決定係数 $R^2$ について

次の式

$$(Y_i - \bar{Y}) = (\widehat{Y}_i - \bar{Y}) + \widehat{u}_i,$$

の両辺を二乗して、総和すると、

$$\begin{aligned}\sum_{i=1}^n (Y_i - \bar{Y})^2 &= \sum_{i=1}^n \left( (\widehat{Y}_i - \bar{Y}) + \widehat{u}_i \right)^2 \\ &= \sum_{i=1}^n (\widehat{Y}_i - \bar{Y})^2 + 2 \sum_{i=1}^n (\widehat{Y}_i - \bar{Y}) \widehat{u}_i + \sum_{i=1}^n \widehat{u}_i^2 \\ &= \sum_{i=1}^n (\widehat{Y}_i - \bar{Y})^2 + \sum_{i=1}^n \widehat{u}_i^2\end{aligned}$$

となる。まとめると、

$$\sum_{i=1}^n (Y_i - \bar{Y})^2 = \sum_{i=1}^n (\widehat{Y}_i - \bar{Y})^2 + \sum_{i=1}^n \widehat{u}_i^2$$

を得る。さらに、

$$1 = \frac{\sum_{i=1}^n (\widehat{Y}_i - \bar{Y})^2}{\sum_{i=1}^n (Y_i - \bar{Y})^2} + \frac{\sum_{i=1}^n \widehat{u}_i^2}{\sum_{i=1}^n (Y_i - \bar{Y})^2}$$

それぞれの項は、

1.  $\sum_{i=1}^n (Y_i - \bar{Y})^2 \Rightarrow y$  の全変動

2.  $\sum_{i=1}^n (\widehat{Y}_i - \bar{Y})^2 \Rightarrow \widehat{Y}_i$  (回帰直線) で説明される部分

3.  $\sum_{i=1}^n \widehat{u}_i^2 \Rightarrow \widehat{Y}_i$  (回帰直線) で説明されない部分

となる。

回帰式の当てはまりの良さを示す指標として、決定係数  $R^2$  を以下の通りに定義する。

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

または,

$$R^2 = 1 - \frac{\sum_{i=1}^n \widehat{u}_i^2}{\sum_{i=1}^n (Y_i - \bar{Y})^2},$$

として書き換えられる。

または、 $Y_i = \widehat{Y}_i + \widehat{u}_i$  と

$$\begin{aligned}\sum_{i=1}^n (\widehat{Y}_i - \bar{Y})^2 &= \sum_{i=1}^n (\widehat{Y}_i - \bar{Y})(Y_i - \bar{Y} - \widehat{u}_i) \\ &= \sum_{i=1}^n (\widehat{Y}_i - \bar{Y})(Y_i - \bar{Y}) - \sum_{i=1}^n (\widehat{Y}_i - \bar{Y})\widehat{u}_i \\ &= \sum_{i=1}^n (\widehat{Y}_i - \bar{Y})(Y_i - \bar{Y})\end{aligned}$$

を用いて、

$$\begin{aligned}R^2 &= \frac{\sum_{i=1}^n (\widehat{Y}_i - \bar{Y})^2}{\sum_{i=1}^n (Y_i - \bar{Y})^2} \\ &= \frac{(\sum_{i=1}^n (\widehat{Y}_i - \bar{Y})^2)^2}{\sum_{i=1}^n (Y_i - \bar{Y})^2 \sum_{i=1}^n (\widehat{Y}_i - \bar{Y})^2} \\ &= \left( \frac{\sum_{i=1}^n (\widehat{Y}_i - \bar{Y})(Y_i - \bar{Y})}{\sqrt{\sum_{i=1}^n (Y_i - \bar{Y})^2 \sum_{i=1}^n (\widehat{Y}_i - \bar{Y})^2}} \right)^2\end{aligned}$$

と書き換えられる。すなわち、 $R^2$  は  $Y_i$  と  $\widehat{Y}_i$  の相関係数の二乗と解釈される。

$$\sum_{i=1}^n (Y_i - \bar{Y})^2 = \sum_{i=1}^n (\widehat{Y}_i - \bar{Y})^2 + \sum_{i=1}^n \widehat{u}_i^2 \text{ から, 明らかに,}$$

$$0 \leq R^2 \leq 1,$$

となる。 $R^2$  が 1 に近づけば回帰式の当てはまりは良いと言える。しかし、 $t$  分布のような数表は存在しない。したがって、「どの値よりも大きくなるべき」というような基準はない。

慣習的には、メドとして 0.9 以上を判断基準にする。

数値例： 決定係数の計算には以下の公式を用いる。

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

計算に必要なものは、 $\widehat{u}_i = Y_i - (\widehat{\alpha} + \widehat{\beta}X_i)$ ,  $\bar{Y}$ ,  $\sum_{i=1}^n Y_i^2$  である。

$i$	$Y_i$	$X_i$	$\widehat{Y}_i$	$\widehat{u}_i$	$\widehat{u}_i^2$	$Y_i^2$
1	6	10	6.8	-0.8	0.64	36
2	9	12	8.1	0.9	0.81	81
3	10	14	9.4	0.6	0.36	100
4	10	16	10.7	-0.7	0.49	100
合計	$\sum Y_i$	$\sum X_i$	$\sum \widehat{Y}_i$	$\sum \widehat{u}_i$	$\sum \widehat{u}_i^2$	$\sum Y_i^2$
	35	52	35.0	0.0	2.30	317

$\sum \widehat{u}_i^2 = 2.30$ ,  $\bar{X} = 13$ ,  $\bar{Y} = 8.75$ ,  $\sum_{i=1}^n Y_i^2 = 317$  なので,

$$R^2 = 1 - \frac{2.30}{317 - 4 \times 8.75^2} = 1 - \frac{2.30}{10.75} = 0.786$$

## 1.5 まとめ

$\widehat{\alpha}$ ,  $\widehat{\beta}$ を求めるための公式は

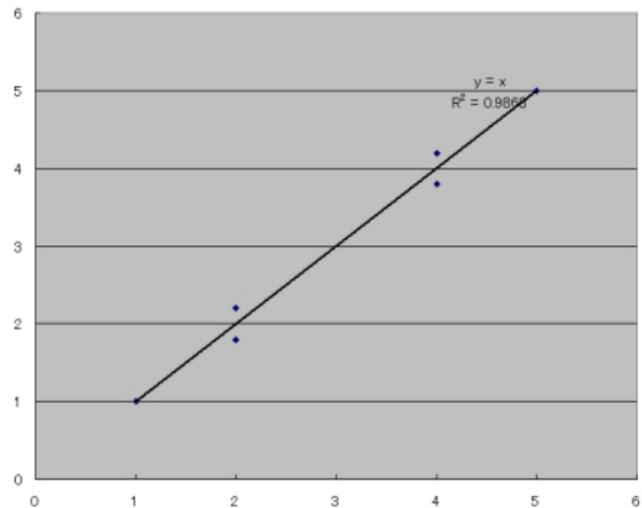
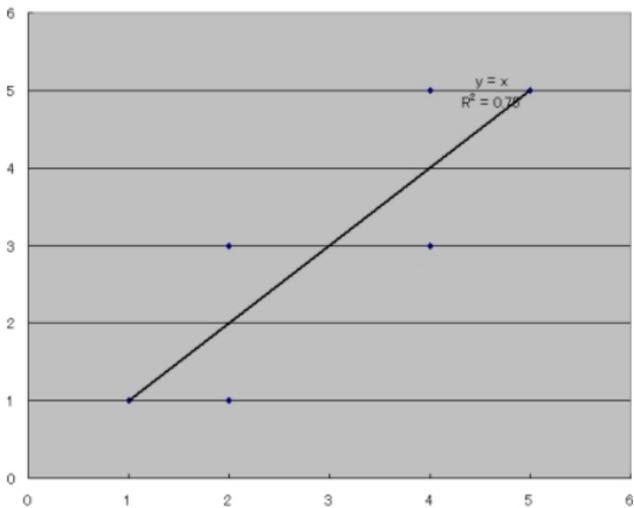
$$\widehat{\beta} = \frac{\sum_{i=1}^n X_i Y_i - n\overline{X}\overline{Y}}{\sum_{i=1}^n X_i^2 - n\overline{X}^2}$$
$$\widehat{\alpha} = \overline{Y} - \widehat{\beta}\overline{X}$$

なので、必要なものは  $\overline{X}$ ,  $\overline{Y}$ ,  $\sum_{i=1}^n X_i^2$ ,  $\sum_{i=1}^n X_i Y_i$  である。

決定係数の計算には以下の公式を用いる。

$$R^2 = 1 - \frac{\sum_{i=1}^n \widehat{u}_i^2}{\sum_{i=1}^n (Y_i - \overline{Y})^2} = 1 - \frac{\sum_{i=1}^n \widehat{u}_i^2}{\sum_{i=1}^n Y_i^2 - n\overline{Y}^2}$$

計算に必要なものは,  $\sum \widehat{u}_i^2$ ,  $\overline{Y}$ ,  $\sum_{i=1}^n Y_i^2$  である。



## 2 Regression Analysis (回帰分析)

### 2.1 Setup of the Model

When  $(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)$  are available, suppose that there is a linear relationship between  $y$  and  $x$ , i.e.,

$$y_i = \beta_1 + \beta_2 x_i + u_i, \quad (4)$$

for  $i = 1, 2, \dots, n$ .  $x_i$  and  $y_i$  denote the  $i$ th observations.

→ **Single (or simple) regression model** (単回帰モデル)

$y_i$  is called the **dependent variable** (従属変数) or the **explained variable** (被説明変数), while  $x_i$  is known as the **independent variable** (独立変数) or the **explanatory (or explaining) variable** (説明変数).

$\beta_1 = \mathbf{Intercept}$  (切片),  $\beta_2 = \mathbf{Slope}$  (傾き)

$\beta_1$  and  $\beta_2$  are unknown **parameters** (パラメータ, 母数) to be estimated.

$\beta_1$  and  $\beta_2$  are called the **regression coefficients** (回帰係数).

$u_i$  is the unobserved **error term** (誤差項) assumed to be a random variable with mean zero and variance  $\sigma^2$ .

$\sigma^2$  is also a parameter to be estimated.

$x_i$  is assumed to be **nonstochastic** (非確率的), but  $y_i$  is **stochastic** (確率的) because  $y_i$  depends on the error  $u_i$ .

The error terms  $u_1, u_2, \dots, u_n$  are assumed to be mutually independently and identically distributed, which is called ***iid***.

It is assumed that  $u_i$  has a distribution with mean zero, i.e.,  $E(u_i) = 0$  is assumed.

Taking the expectation on both sides of (4), the expectation of  $y_i$  is represented as:

$$\begin{aligned} E(y_i) &= E(\beta_1 + \beta_2 x_i + u_i) = \beta_1 + \beta_2 x_i + E(u_i) \\ &= \beta_1 + \beta_2 x_i, \end{aligned} \tag{5}$$

for  $i = 1, 2, \dots, n$ .

Using  $E(y_i)$  we can rewrite (4) as  $y_i = E(y_i) + u_i$ .

(5) represents the true regression line.

Let  $\hat{\beta}_1$  and  $\hat{\beta}_2$  be estimates of  $\beta_1$  and  $\beta_2$ .

Replacing  $\beta_1$  and  $\beta_2$  by  $\hat{\beta}_1$  and  $\hat{\beta}_2$ , (4) turns out to be:

$$y_i = \hat{\beta}_1 + \hat{\beta}_2 x_i + e_i, \tag{6}$$

for  $i = 1, 2, \dots, n$ , where  $e_i$  is called the **residual** (残差).

The residual  $e_i$  is taken as the experimental value (or realization) of  $u_i$ .

We define  $\hat{y}_i$  as follows:

$$\hat{y}_i = \hat{\beta}_1 + \hat{\beta}_2 x_i, \quad (7)$$

for  $i = 1, 2, \dots, n$ , which is interpreted as the **predicted value** (予測値) of  $y_i$ .

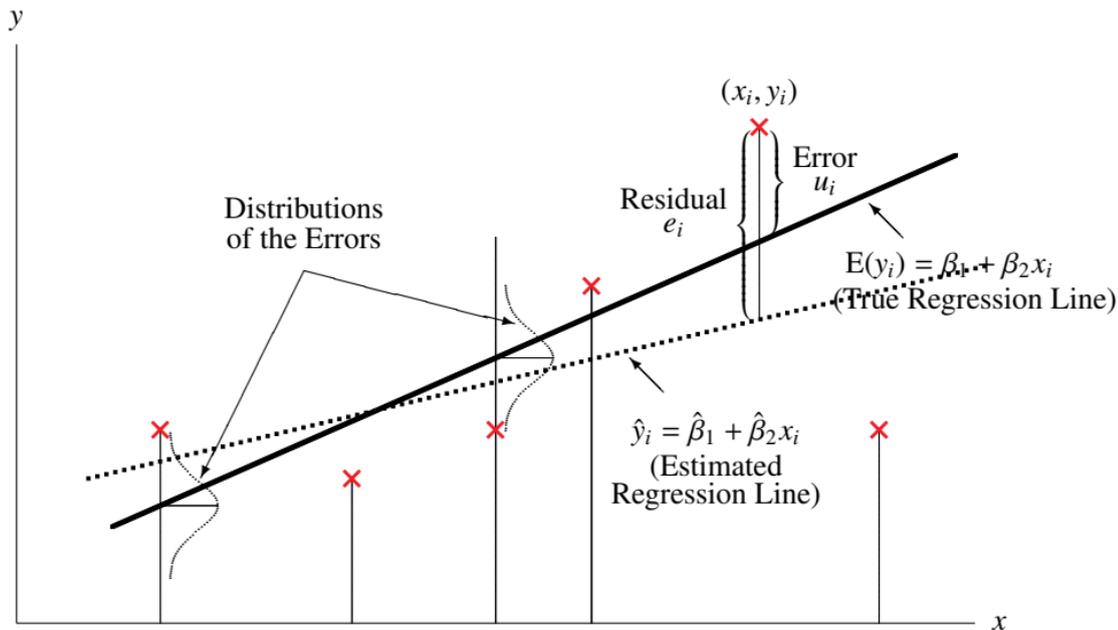
(7) indicates the estimated regression line, which is different from (5).

Moreover, using  $\hat{y}_i$  we can rewrite (6) as  $y_i = \hat{y}_i + e_i$ .

(5) and (7) are displayed in Figure 1.

Consider the case of  $n = 6$  for simplicity.  $\times$  indicates the observed data series.

**Figure 1. True and Estimated Regression Lines (回帰直線)**



The true regression line (5) is represented by the solid line, while the estimated regression line (7) is drawn with the dotted line.

Based on the observed data,  $\beta_1$  and  $\beta_2$  are estimated as:  $\hat{\beta}_1$  and  $\hat{\beta}_2$ .

In the next section, we consider how to obtain the estimates of  $\beta_1$  and  $\beta_2$ , i.e.,  $\hat{\beta}_1$  and  $\hat{\beta}_2$ .

## 2.2 Ordinary Least Squares Estimation

Suppose that  $(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)$  are available.

For the regression model (4), we consider estimating  $\beta_1$  and  $\beta_2$ .

Replacing  $\beta_1$  and  $\beta_2$  by their estimates  $\hat{\beta}_1$  and  $\hat{\beta}_2$ , remember that the residual  $e_i$  is given by:

$$e_i = y_i - \hat{y}_i = y_i - \hat{\beta}_1 - \hat{\beta}_2 x_i.$$

The sum of squared residuals is defined as follows:

$$S(\hat{\beta}_1, \hat{\beta}_2) = \sum_{i=1}^n e_i^2 = \sum_{i=1}^n (y_i - \hat{\beta}_1 - \hat{\beta}_2 x_i)^2.$$

It might be plausible to choose the  $\hat{\beta}_1$  and  $\hat{\beta}_2$  which minimize the sum of squared residuals, i.e.,  $S(\hat{\beta}_1, \hat{\beta}_2)$ .

This method is called the **ordinary least squares estimation** (最小二乘法, **OLS**).

To minimize  $S(\hat{\beta}_1, \hat{\beta}_2)$  with respect to  $\hat{\beta}_1$  and  $\hat{\beta}_2$ , we set the partial derivatives equal to zero:

$$\begin{aligned}\frac{\partial S(\hat{\beta}_1, \hat{\beta}_2)}{\partial \hat{\beta}_1} &= -2 \sum_{i=1}^n (y_i - \hat{\beta}_1 - \hat{\beta}_2 x_i) = 0, \\ \frac{\partial S(\hat{\beta}_1, \hat{\beta}_2)}{\partial \hat{\beta}_2} &= -2 \sum_{i=1}^n x_i (y_i - \hat{\beta}_1 - \hat{\beta}_2 x_i) = 0.\end{aligned}$$

The second order condition for minimization is:

$$\begin{pmatrix} \frac{\partial^2 S(\hat{\beta}_1, \hat{\beta}_2)}{\partial \hat{\beta}_1^2} & \frac{\partial^2 S(\hat{\beta}_1, \hat{\beta}_2)}{\partial \hat{\beta}_1 \partial \hat{\beta}_2} \\ \frac{\partial^2 S(\hat{\beta}_1, \hat{\beta}_2)}{\partial \hat{\beta}_2 \partial \hat{\beta}_1} & \frac{\partial^2 S(\hat{\beta}_1, \hat{\beta}_2)}{\partial \hat{\beta}_2^2} \end{pmatrix} = \begin{pmatrix} 2n & 2 \sum_{i=1}^n x_i \\ 2 \sum_{i=1}^n x_i & 2 \sum_{i=1}^n x_i^2 \end{pmatrix}$$

should be a positive definite matrix.

The diagonal elements  $2n$  and  $2 \sum_{i=1}^n x_i^2$  are positive.

The determinant:

$$\begin{vmatrix} 2n & 2 \sum_{i=1}^n x_i \\ 2 \sum_{i=1}^n x_i & 2 \sum_{i=1}^n x_i^2 \end{vmatrix} = 4n \sum_{i=1}^n x_i^2 - 4 \left( \sum_{i=1}^n x_i \right)^2 = 4n \sum_{i=1}^n (x_i - \bar{x})^2$$

is positive.  $\implies$  The second-order condition is satisfied.

The first two equations yield the following two equations:

$$\bar{y} = \hat{\beta}_1 + \hat{\beta}_2 \bar{x}, \tag{8}$$

$$\sum_{i=1}^n x_i y_i = n \bar{x} \hat{\beta}_1 + \hat{\beta}_2 \sum_{i=1}^n x_i^2, \tag{9}$$

where  $\bar{y} = \frac{1}{n} \sum_{i=1}^n y_i$  and  $\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$ .

Multiplying (8) by  $n\bar{x}$  and subtracting (9), we can derive  $\hat{\beta}_2$  as follows:

$$\hat{\beta}_2 = \frac{\sum_{i=1}^n x_i y_i - n\bar{x}\bar{y}}{\sum_{i=1}^n x_i^2 - n\bar{x}^2} = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^n (x_i - \bar{x})^2}. \quad (10)$$

From (8),  $\hat{\beta}_1$  is directly obtained as follows:

$$\hat{\beta}_1 = \bar{y} - \hat{\beta}_2 \bar{x}. \quad (11)$$

When the observed values are taken for  $y_i$  and  $x_i$  for  $i = 1, 2, \dots, n$ , we say that  $\hat{\beta}_1$  and  $\hat{\beta}_2$  are called the **ordinary least squares estimates** (or simply the **least squares estimates**, 最小二乘推定値) of  $\beta_1$  and  $\beta_2$ .

When  $y_i$  for  $i = 1, 2, \dots, n$  are regarded as the random sample, we say that  $\hat{\beta}_1$  and  $\hat{\beta}_2$  are called the **ordinary least squares estimators** (or the **least squares estimators**, 最小二乘推定量) of  $\beta_1$  and  $\beta_2$ .

## 2.3 Properties of Least Squares Estimator

Equation (10) is rewritten as:

$$\begin{aligned}\hat{\beta}_2 &= \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^n (x_i - \bar{x})^2} = \frac{\sum_{i=1}^n (x_i - \bar{x})y_i}{\sum_{i=1}^n (x_i - \bar{x})^2} - \frac{\bar{y} \sum_{i=1}^n (x_i - \bar{x})}{\sum_{i=1}^n (x_i - \bar{x})^2} \\ &= \sum_{i=1}^n \frac{x_i - \bar{x}}{\sum_{i=1}^n (x_i - \bar{x})^2} y_i = \sum_{i=1}^n \omega_i y_i.\end{aligned}\tag{12}$$

In the third equality,  $\sum_{i=1}^n (x_i - \bar{x}) = 0$  is utilized because of  $\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$ .

In the fourth equality,  $\omega_i$  is defined as:  $\omega_i = \frac{x_i - \bar{x}}{\sum_{i=1}^n (x_i - \bar{x})^2}$ .

$\omega_i$  is nonstochastic because  $x_i$  is assumed to be nonstochastic.

$\omega_i$  has the following properties:

$$\sum_{i=1}^n \omega_i = \sum_{i=1}^n \frac{x_i - \bar{x}}{\sum_{i=1}^n (x_i - \bar{x})^2} = \frac{\sum_{i=1}^n (x_i - \bar{x})}{\sum_{i=1}^n (x_i - \bar{x})^2} = 0,\tag{13}$$

$$\sum_{i=1}^n \omega_i x_i = \sum_{i=1}^n \omega_i (x_i - \bar{x}) = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{\sum_{i=1}^n (x_i - \bar{x})^2} = 1, \quad (14)$$

$$\sum_{i=1}^n \omega_i^2 = \sum_{i=1}^n \left( \frac{x_i - \bar{x}}{\sum_{i=1}^n (x_i - \bar{x})^2} \right)^2 = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{\left( \sum_{i=1}^n (x_i - \bar{x})^2 \right)^2} = \frac{1}{\sum_{i=1}^n (x_i - \bar{x})^2}. \quad (15)$$

The first equality of (14) comes from (13).

From now on, we focus only on  $\hat{\beta}_2$ , because usually  $\beta_2$  is more important than  $\beta_1$  in the regression model (4).

In order to obtain the properties of the least squares estimator  $\hat{\beta}_2$ , we rewrite (12) as:

$$\begin{aligned} \hat{\beta}_2 &= \sum_{i=1}^n \omega_i y_i = \sum_{i=1}^n \omega_i (\beta_1 + \beta_2 x_i + u_i) \\ &= \beta_1 \sum_{i=1}^n \omega_i + \beta_2 \sum_{i=1}^n \omega_i x_i + \sum_{i=1}^n \omega_i u_i = \beta_2 + \sum_{i=1}^n \omega_i u_i. \end{aligned} \quad (16)$$

In the fourth equality of (16), (13) and (14) are utilized.

## [Review] Random Variables:

Let  $X_1, X_2, \dots, X_n$  be  $n$  random variables, which are mutually independently and identically distributed.

**mutually independent**  $\implies f(x_i, x_j) = f_i(x_i)f_j(x_j)$  for  $i \neq j$ .

$f(x_i, x_j)$  denotes a joint distribution of  $X_i$  and  $X_j$ .

$f_i(x)$  indicates a marginal distribution of  $X_i$ .

**identical**  $\implies f_i(x) = f_j(x)$  for  $i \neq j$ .

[End of Review]

### **[Review] Mean and Variance:**

Let  $X$  and  $Y$  be random variables (continuous type), which are independently distributed.

### **Definition and Formulas:**

- $E(g(X)) = \int g(x)f(x)dx$  for a function  $g(\cdot)$  and a density function  $f(\cdot)$ .
- $V(X) = E((X - \mu)^2) = \int (x - \mu)^2 f(x)dx$  for  $\mu = E(X)$ .
- $E(aX + b) = aE(X) + b$  and  $V(aX + b) = a^2V(X)$ .
- $E(X \pm Y) = E(X) \pm E(Y)$  and  $V(X \pm Y) = V(X) + V(Y)$ .

**[End of Review]**