専門セミナー（計量経済学の基礎）

(2017年度 春～夏学期 講義ノート)

平成 29 年 4 月 27 日 (木) 版

参考書『基本統計学 (第 3 版)』

(豊田・大谷・小川・長谷川・谷崎著，東洋経済新報社，2010 年)

谷崎　久志
大阪大学・経済学部

目 次

- この講義ノートは,

http://www2.econ.osaka-u.ac.jp/~tanizaki/class/2017

からダウンロード可。

- この講義ノートの文中のページは教科書『基本統計学 (第 3 版)』のページに対応。

序説 (P.1)

1. 統計的記述：

   資料の収集と整理 (平均値・分散・メディアン等の計算) $\Longrightarrow$ 第 1, 2 章

2. 統計的推測：

   標本から母集団の特徴をつかむこと

   (a) 標本： データを標本と考える

   (b) 母集団： 標本を含む全体

   (c) 母集団の特徴： 母集団の特性を表すパラメータ (母数という)

   (d) パラメータ (母数)： 平均，分散

   $\Longrightarrow$ 母数 (パラメータ) の推定と仮説検定が主な内容

1 度数分布 (P.3)

1.1 変数 (P.4)

変数の種類 (P.4)

1. 連続型変数： ある区間内の任意の実数値をとりうる変数 (身長，体重，温度，$\cdots$)

2. 離散型変数： 不連続な値しかとらない変数 (サイコロの出た目，家族数，$\cdots$)
   ただし，離散型変数を連続型変数とみなす場合も多い (例： 金額は離散型変数，2009 年の GDP は 470936.7 × 10 億円で，1 円に対して，GNP の値はあまりにも大きい)

データの種類 (P.9,10)

1. 時系列データ： 時間に依存するデータ (P.6 の表 1.1, 表 1.2, P.9 の表 1.4)

2. クロスセクション・データ (横断面データ)： 家計，企業等の一時点でのデータの系列 (P.10 の表 1.6)

1.2 度数分布 (P.4)

表 1.3 (P.7) のデータ (20 個の物体の重さ):

4.3	5.2	7.2	6.4	3.5	5.6	6.7	6.1	4.1	6.8
5.0	5.6	3.8	4.6	5.8	5.1	6.2	5.3	7.4	5.9

このデータを整理する。
$\Longrightarrow$ 表 1.4 (P.8)

階級値	階級境界値	度数
3.45	2.95〜3.95	2
4.45	3.95〜4.95	3
5.45	4.95〜5.95	8
6.45	5.95〜6.95	5
7.45	6.95〜7.95	2
	合計	20

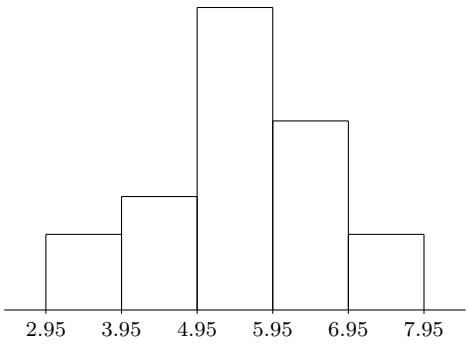
をもとにして，

表 1.4 20 個の物体の重さの度数分布表

階級値	階級境界値	度数	相対度数	累積度数	累積相対度数
3.45	2.95〜3.95	2	0.10	2	0.10
4.45	3.95〜4.95	3	0.15	5	0.25
5.45	4.95〜5.95	8	0.40	13	0.65
6.45	5.95〜6.95	5	0.25	18	0.90
7.45	6.95〜7.95	2	0.10	20	1.00
	合計	20	1.000		

を得る。小数第 2 位の 0.05 の単位で区間を分けている理由
$\longrightarrow$ 四捨五入の関係
小数第 1 位の 0.1 の単位で区間を分けた場合，境界値がどの階級に属するか区別できなくなる。(例えば，5.0 は 4.95 以上から 5.05 未満の間の数値)

図 1.1 20 個の物体の重さのグラフ (P.11)



グラフの形

- 右の裾野が広い $\Longrightarrow$ 右に歪んでいる
- 左の裾野が広い $\Longrightarrow$ 左に歪んでいる

グラフの作り方

1. 階級境界値：階級の境界を定める値
2. 階級値：階級境界値の中点
3. 度数：ある階級に属するデータの数
4. 度数分布表：各階級とその度数を表に表したもの
5. ヒストグラム：度数分布をグラフに表す
6. 相対度数 : 各階級の度数をデータの総数で割ったもの，すなわち，各階級に属するデータの割合
7. 累積度数：ある階級以下の度数を合計したもの
8. 累積相対度数 : ある階級以下の相対度数を合計したもの

2　代表値 (P.15)

度数分布表，ヒストグラム： 統計データを整理し，母集団に関する情報を得る一つの方法。
分布の状態を数値で表したい。
代表値： データを代表する値 $\Longrightarrow$ 平均値，分散，標準偏差，中央値 (メディアン)，最頻値 (モード)，$\cdots$

2.1　平均値 (P.16)

n 個のデータ： $x_1, x_2, \cdots, x_n$

算術平均 (P.16)：

$$\overline{x} = \frac{1}{n}(x_1 + x_2 + \cdots + x_n) = \frac{1}{n}\sum_{i=1}^{n} x_i$$

表 1.3 (P.7) のデータから

$$\overline{x} = \frac{1}{20}(4.3 + 5.2 + \cdots + 5.9) = 5.53$$

となる。

加重平均 (P.16)：

階級値	階級境界値 (以上)　(未満)	度数
m_1	$a_0 \sim a_1$	f_1
m_2	$a_1 \sim a_2$	f_2
$\vdots$	$\vdots$	$\vdots$
m_k	$a_{k-1} \sim a_k$	f_k
	合計	n

ただし，$m_1 = \dfrac{a_0 + a_1}{2}$, $m_2 = \dfrac{a_1 + a_2}{2}$, $\cdots$,
$m_k = \dfrac{a_{k-1} + a_k}{2}$ とする。

上のような度数分布表が利用可能なとき，

$$\overline{x} = \frac{1}{n}(f_1 m_1 + f_2 m_2 + \cdots + f_k m_k) = \frac{1}{n}\sum_{i=1}^{k} f_i m_i$$

として，平均値を計算することが出来る。$\Longrightarrow$ 加重平均 (各階級値を度数でウエイトづけして平均したもの)

$$\overline{x} = \sum_{i=1}^{k} \frac{f_i}{n} m_i$$

$\dfrac{f_i}{n}$ は相対度数である。
上の表のデータの平均を求めると，

$$\begin{aligned}\overline{x} &= \frac{1}{20}\Big(2 \times 3.45 + 3 \times 4.45 \\ &\quad +8 \times 5.45 + 5 \times 6.45 + 2 \times 7.45\Big) \\ &= 5.55\end{aligned}$$

階級の幅の選び方によって，多少，値は異なる。

2.2　分散，標準偏差 (P.20)

分散，標準偏差： データの散らばり具合を表す
分散，標準偏差が大きければ，データの存在する範囲が広い
標準偏差＝分散の平方根
分散 (s^2 で表す) の定義：

$$\begin{aligned}s^2 &= \frac{1}{n}\Big((x_1 - \overline{x})^2 + (x_2 - \overline{x})^2 + \cdots + (x_n - \overline{x})^2\Big) \\ &= \frac{1}{n}\sum_{i=1}^{n}(x_i - \overline{x})^2\end{aligned}$$

ただし，$\overline{x} = \dfrac{1}{n}\sum_{i=1}^{n} x_i$ とする。

標準偏差： s
分散の実際の計算には，

$$s^2 = \frac{1}{n}\sum_{i=1}^{n} x_i^2 - \overline{x}^2$$

を用いる。
なぜなら，

$$\begin{aligned} s^2 &= \frac{1}{n}\sum_{i=1}^{n}(x_i - \overline{x})^2 \\ &= \frac{1}{n}\sum_{i=1}^{n}(x_i^2 - 2\overline{x}x_i + \overline{x}^2) \\ &= \frac{1}{n}\Big(\sum_{i=1}^{n} x_i^2 - 2\overline{x}\sum_{i=1}^{n} x_i + \sum_{i=1}^{n}\overline{x}^2\Big) \\ &= \frac{1}{n}\Big(\sum_{i=1}^{n} x_i^2 - 2n\overline{x}^2 + n\overline{x}^2\Big) \\ &= \frac{1}{n}\Big(\sum_{i=1}^{n} x_i^2 - n\overline{x}^2\Big) \\ &= \frac{1}{n}\sum_{i=1}^{n} x_i^2 - \overline{x}^2 \end{aligned}$$

となる。
表 1.3 (P.7) のデータの分散を求めると，

$$\begin{aligned} s^2 &= \frac{1}{20}\Big((4.3 - 5.53)^2 + (5.2 - 5.53)^2 + \cdots \\ &\quad +(5.9 - 5.53)^2\Big) \\ &= 1.1591 \end{aligned}$$

または，

$$\begin{aligned} s^2 &= \frac{1}{20}(4.3^2 + 5.2^2 + \cdots + 5.9^2) - 5.53^2 \\ &= 1.1591 \end{aligned}$$

$s = 1.0766$ ===> 標準偏差
表 2.1 (P.17) の度数分布表からの計算では，

$$s^2 = \frac{1}{n}\sum_{i=1}^{k} f_i(m_i - \overline{x})^2$$

となる。ただし，$\overline{x} = \dfrac{1}{n}\displaystyle\sum_{i=1}^{k} f_i m_i$ とする。
実際の計算には，

$$s^2 = \frac{1}{n}\sum_{i=1}^{k} f_i m_i^2 - \overline{x}^2$$

を使う。
なぜなら，

$$\begin{aligned} s^2 &= \frac{1}{n}\sum_{i=1}^{k} f_i(m_i - \overline{x})^2 \\ &= \frac{1}{n}\sum_{i=1}^{k} f_i(m_i^2 - 2\overline{x}m_i + \overline{x}^2) \\ &= \frac{1}{n}\Big(\sum_{i=1}^{k} f_i m_i^2 - 2\overline{x}\sum_{i=1}^{k} f_i m_i + \overline{x}^2\sum_{i=1}^{k} f_i\Big) \\ &= \frac{1}{n}\Big(\sum_{i=1}^{k} f_i m_i^2 - 2n\overline{x}^2 + n\overline{x}^2\Big) \\ &= \frac{1}{n}\Big(\sum_{i=1}^{k} f_i m_i^2 - n\overline{x}^2\Big) \\ &= \frac{1}{n}\sum_{i=1}^{k} f_i m_i^2 - \overline{x}^2 \end{aligned}$$

となる。
上の表のデータの分散を求めると，

$$\begin{aligned} s^2 &= \frac{1}{20}\Big(2(3.45 - 5.55)^2 + 3(4.45 - 5.55)^2 \\ &\quad +8(5.45 - 5.55)^2 + 5(6.45 - 5.55)^2 \\ &\quad +2(7.45 - 5.55)^2\Big) \\ &= 1.19 \end{aligned}$$

または，

$$\begin{aligned} s^2 &= \frac{1}{20}(2 \times 3.45^2 + 3 \times 4.45^2 \\ &\quad +8 \times 5.45^2 + 5 \times 6.45^2 + 2 \times 7.45^2) - 5.55^2 \\ &= 1.19 \end{aligned}$$

すなわち，$s = 1.0909$，

2.3 範囲，四分位点，メディアン，モード (P.18)

- 範囲： 最大値－最小値
- 四分位点：
  25 %点 (第 1 四分位点)，50 %点 (第 2 四分位点)，75 %点 (第 3 四分位点) のこと
- 四分位範囲： 第 3 四分位点－第 1 四分位点

- メディアン（中央値）:
  大きい順に並べて，真ん中の値 (第 2 四分位点) $\longrightarrow$ 表 1.3 (P.7) のデータでは，大きい順に並べて 10 番目と 11 番目のデータの平均で，$(5.6 + 5.6)/2 = 5.6$

- モード（最頻値）:
  最も多い度数の階級値 $\longrightarrow$ 表 1.3 (P.7) のデータでは 5.45，階級の幅によって変わる

2.4 相関係数 (P.23)

2 変数データの組に関する代表値 $\Longrightarrow$ 共分散，相関係数
例： 100 人の家計からの消費と所得，身長と体重
n 組のデータ $(x_1, y_1), (x_2, y_2), \cdots, (x_n, y_n)$
共分散 s_{xy}

$$\begin{aligned} s_{xy} &= \frac{1}{n}\Big((x_1 - \overline{x})(y_1 - \overline{y}) + (x_2 - \overline{x})(y_2 - \overline{y}) \\ &\quad + \cdots + (x_n - \overline{x})(y_n - \overline{y})\Big) \\ &= \frac{1}{n}\sum_{i=1}^{n}(x_i - \overline{x})(y_i - \overline{y}) \\ &= \frac{1}{n}\sum_{i=1}^{n} x_i y_i - \overline{x}\,\overline{y} \end{aligned}$$

$s_{xy} > 0$： 正の相関 (x と y との関係はプラスの傾き)
$s_{xy} < 0$： 負の相関 (x と y との関係はマイナスの傾き)
$s_{xy} = 0$： 相関なし (x と y との関係は正負の傾きを決定できず)

相関 $\Longrightarrow$ 互いにかかわりを持つこと。相互に関係しあっていること。(『国語大辞典 (新装版)』小学館，1988)
相関の強弱を表す指標 $\Longrightarrow$ 相関係数 r

$$r = \frac{s_{xy}}{s_x s_y}$$

ただし，

$$s_x^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \overline{x})^2, \quad s_y^2 = \frac{1}{n}\sum_{i=1}^{n}(y_i - \overline{y})^2,$$

とし，s_x, s_y は x の標準偏差，y の標準偏差である。
$r > 0$： 正の相関 (x と y との関係はプラスの傾き)
$r < 0$： 負の相関 (x と y との関係はマイナスの傾き)
$r = 0$： 相関なし (x と y との関係は正負の傾きを決定できず)

r は，

$$-1 \le r \le 1$$

となる。
証明：
次のような t に関する式を考える。

$$f(t) = \frac{1}{n}\sum_{i=1}^{n}\Big((x_i - \overline{x})t - (y_i - \overline{y})\Big)^2,$$

平方和なので，必ずゼロ以上となる。よって，すべての t について，$f(t) \ge 0$ となるための条件を求めればよい。t に関する 2 次方程式の判別式がゼロ以下となる条件を求める。

$$\begin{aligned} f(t) &= t^2\frac{1}{n}\sum_{i=1}^{n}(x_i - \overline{x})^2 \\ &\quad + 2t\frac{1}{n}\sum_{i=1}^{n}(x_i - \overline{x})(y_i - \overline{y}) \\ &\quad + \frac{1}{n}\sum_{i=1}^{n}(y_i - \overline{y})^2 \\ &= s_x^2 t^2 + 2s_{xy}t + s_y^2 \ge 0 \end{aligned}$$

判別式

$$\frac{D}{4} = s_{xy}^2 - s_x^2 s_y^2 \le 0$$

$$\frac{s_{xy}^2}{s_x^2 s_y^2} \le 1,$$

$$-1 \le \frac{s_{xy}}{s_x s_y} \le 1,$$

を得る。
r が 1 に近いほど， 正の相関が強くなる (x と y のプロットが正の傾きで一直線上に近づく)。
r が -1 に近いほど， 負の相関が強くなる (x と y のプロットが負の傾きで一直線上に近づく)。
$r = -1, 1$ のとき，x と y は一直線上に並ぶ ($r = 1$ は正の傾き，$r = -1$ は負の傾き)。

3 計量経済学について

- 経済理論 (ミクロ，マクロ，財政，金融，国際経済，$\cdots$)

- データ (GNP，消費，投資，金利，為替レート，$\cdots$)

計量経済学 $\Longrightarrow$ 経済理論が現実に成り立つものかどうかを，データを用いて，統計的に検証する。

3.1 例 1： マクロの消費関数

$$C = f(Y)$$

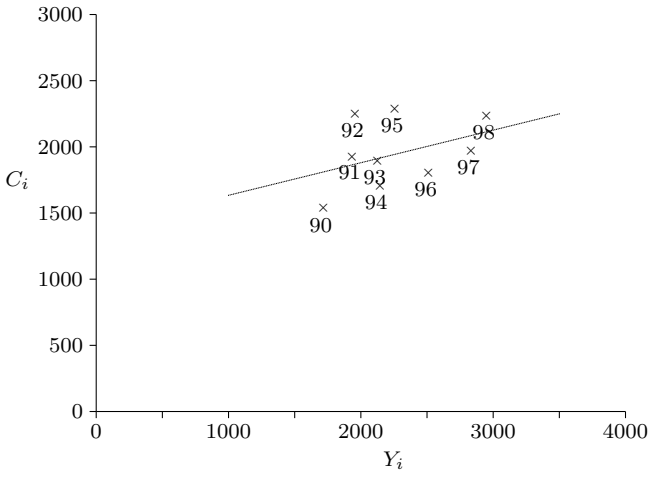
ただし，C は消費，Y は所得。

1. $Y \nearrow \Longrightarrow C \nearrow$
2. $\dfrac{\mathrm{d}C}{\mathrm{d}Y}$ = 限界消費性向 = 所得 1 円増加で消費が何円増加するか
3. すなわち，$\dfrac{\mathrm{d}C}{\mathrm{d}Y} > 0$

モデルの定式化

1. $C = a + bY$
2. $b = \dfrac{\mathrm{d}C}{\mathrm{d}Y}$ = 限界消費性向
3. a = 基礎消費 ($Y = 0$ のときに必要な消費)
4. 符号条件： $a > 0$，$b > 0$ (しかも，$1 > b$)

図 1： 消費 (C_i) と所得 (Y_i)



1. $\times \longrightarrow$ 実際のデータ
2. $(Y_i, C_i) \Longrightarrow t$ 期のデータ, i.e., $i = 1, 2, \cdots, 9$
3. $i = 1 \Longrightarrow$ 1990 年,

   $i = 2 \Longrightarrow$ 1991 年,

   $\cdots$,

   $i = 9 \Longrightarrow$ 1998 年,
1. 実際のデータを用いて，a, b を求める。
2. a, b を求める $\equiv$ 現実の経済構造を求める
3. その結果，もし $a > 0$，$1 > b > 0$ なら，経済理論は現実経済を説明していると言える。

3.2 例 2： 日本酒の需要関数

$$Q = f(Y, P_1, P_2)$$

ただし，Q は日本酒の需要量，Y は所得，P_1 は日本酒の価格，P_2 は洋酒の価格。

1. $Y \nearrow \Longrightarrow Q \nearrow$,

   $P_1 \nearrow \Longrightarrow Q \searrow$,

   $P_2 \nearrow \Longrightarrow Q \nearrow$
2. $\dfrac{\partial Q}{\partial Y} > 0$, $\dfrac{\partial Q}{\partial P_1} < 0$, $\dfrac{\partial Q}{\partial P_2} > 0$
3. 日本酒と洋酒は代替財
4. モデルの定式化 (A)

$$Q = a + b_1 Y + b_2 P_1 + b_3 P_2$$

5. Q, Y, P_1, P_2 を用いて，a, b_1, b_2, b_3 を求める (日本酒の需要構造を求める)。
6. 符号条件： $b_1 > 0$, $b_2 < 0$, $b_3 > 0$, a ?
7. t 期のデータ $(Q_i, Y_i, P_{1i}, P_{2i})$
8. n 組のデータ, i.e., $i = 1, 2, \cdots, n$
9. モデルの定式化 (B)

$$Q = a + b_1 Y + b_2 \frac{P_1}{P_2}$$

符号条件： $b_1 > 0$, $b_2 < 0$

10. モデルの定式化 (C)

$$\log(Q) = a + b_1 \log(Y) + b_2 \log(\frac{P_1}{P_2})$$

符号条件： $b_1 > 0, b_2 < 0$

11. モデル (A), (B), (C) のどれが最も現実的かを得られた結果から判断する。

4 行列について

A を 2×2 行列とすると，

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$$

と表される。
$a_{ij} = A$ の第 i 行，第 j 列の要素

a を 2×1 行列 (縦ベクトル) とすると，

$$a = \begin{pmatrix} a_1 \\ a_2 \end{pmatrix}$$

と表される。
$a_i = a$ の第 i 要素

a を 1×2 行列 (横ベクトル) とすると，

$$a = \begin{pmatrix} a_1 & a_2 \end{pmatrix}$$

と表される。
$a_i = a$ の第 i 要素

A を $n \times k$ 行列とすると，

$$A = \begin{pmatrix} a_{11} & \cdots & a_{1k} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nk} \end{pmatrix}$$

と表される。
$a_{ij} = A$ の第 i 行，第 j 列の要素 (ij 要素)

a を $n \times 1$ 行列 (縦ベクトル) とすると，

$$a = \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}$$

と表される。
$a_i = a$ の第 i 要素

a を $1 \times k$ 行列 (横ベクトル) とすると，

$$a = \begin{pmatrix} a_1 & \cdots & a_k \end{pmatrix}$$

と表される。
$a_i = a$ の第 i 要素

行列の等号： A，B を $n \times k$ 行列とする。$A = B$ は，すべての $i = 1, \cdots, n$, $j = 1, \cdots, k$ について，$a_{ij} = b_{ij}$ を意味する。ただし，a_{ij}, b_{ij} は，それぞれ，A, B の ij 要素とする。
$x = 3$, $y = 2$ の 2 つの等式を行列で表す。

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 3 \\ 2 \end{pmatrix} \qquad \text{または} \qquad \begin{pmatrix} x & y \end{pmatrix} = \begin{pmatrix} 3 & 2 \end{pmatrix}$$

行列の和と差： A, B を $n \times k$ 行列とする。

$$A + B = \begin{pmatrix} a_{11} & \cdots & a_{1k} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nk} \end{pmatrix} + \begin{pmatrix} b_{11} & \cdots & b_{1k} \\ \vdots & \ddots & \vdots \\ b_{n1} & \cdots & b_{nk} \end{pmatrix} = \begin{pmatrix} a_{11} + b_{11} & \cdots & a_{1k} + b_{1k} \\ \vdots & \ddots & \vdots \\ a_{n1} + b_{n1} & \cdots & a_{nk} + b_{nk} \end{pmatrix}$$

すなわち，$A + B$ の ij 要素は，$a_{ij} + b_{ij}$ となる。

$$A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \qquad B = \begin{pmatrix} 5 & 6 \\ 7 & 8 \end{pmatrix}$$

$$A + B = \begin{pmatrix} 1+5 & 2+6 \\ 3+7 & 4+8 \end{pmatrix} = \begin{pmatrix} 6 & 8 \\ 10 & 12 \end{pmatrix}$$

$$A - B = \begin{pmatrix} 1-5 & 2-6 \\ 3-7 & 4-8 \end{pmatrix} = \begin{pmatrix} -4 & -4 \\ -4 & -4 \end{pmatrix}$$

要素と行列の積： A を $n \times k$ 行列とする。c を スカラー (1×1 行列のこと) とする。

$$cA = c \begin{pmatrix} a_{11} & \cdots & a_{1k} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nk} \end{pmatrix} = \begin{pmatrix} ca_{11} & \cdots & ca_{1k} \\ \vdots & \ddots & \vdots \\ ca_{n1} & \cdots & ca_{nk} \end{pmatrix}$$

$$A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \qquad c = 5 \qquad \text{のとき}$$

$$cA = 5 \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} = \begin{pmatrix} 5 \times 1 & 5 \times 2 \\ 5 \times 3 & 5 \times 4 \end{pmatrix} = \begin{pmatrix} 5 & 10 \\ 15 & 20 \end{pmatrix}$$

行列と行列の積： A, B を $n \times k$，$k \times n$ 行列とする。

$$AB = \begin{pmatrix} a_{11} & \cdots & a_{1k} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nk} \end{pmatrix} \begin{pmatrix} b_{11} & \cdots & b_{1n} \\ \vdots & \ddots & \vdots \\ b_{k1} & \cdots & b_{kn} \end{pmatrix} = \begin{pmatrix} \sum_{m=1}^{k} a_{1m}b_{m1} & \cdots & \sum_{m=1}^{k} a_{1m}b_{mn} \\ \vdots & \ddots & \vdots \\ \sum_{m=1}^{k} a_{nm}b_{m1} & \cdots & \sum_{m=1}^{k} a_{1m}b_{mn} \end{pmatrix}$$

すなわち，AB は $n \times n$ 行列で，AB の ij 要素は，$a_{i1}b_{1j} + a_{i2}b_{2j} + \cdots + a_{ik}b_{kj} = \sum_{m=1}^{k} a_{ik}b_{kj}$ となる。

$$BA = \begin{pmatrix} b_{11} & \cdots & b_{1n} \\ \vdots & \ddots & \vdots \\ b_{k1} & \cdots & b_{kn} \end{pmatrix} \begin{pmatrix} a_{11} & \cdots & a_{1k} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nk} \end{pmatrix} = \begin{pmatrix} \sum_{m=1}^{n} b_{1m}a_{m1} & \cdots & \sum_{m=1}^{n} b_{1m}a_{mk} \\ \vdots & \ddots & \vdots \\ \sum_{m=1}^{n} b_{km}a_{m1} & \cdots & \sum_{m=1}^{n} b_{1m}a_{mk} \end{pmatrix}$$

すなわち，BA は $k \times k$ 行列で，BA の ij 要素は，$b_{i1}a_{1j} + b_{i2}a_{2j} + \cdots + b_{ik}a_{kj} = \sum_{m=1}^{k} a_{ik}b_{kj}$ となる。
このように，AB と BA の次元は異なる。

$$A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \qquad B = \begin{pmatrix} 5 & 6 \\ 7 & 8 \end{pmatrix}$$

$$AB = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \begin{pmatrix} 5 & 6 \\ 7 & 8 \end{pmatrix} = \begin{pmatrix} 1 \times 5 + 2 \times 7 & 1 \times 6 + 2 \times 8 \\ 3 \times 5 + 4 \times 7 & 3 \times 6 + 4 \times 8 \end{pmatrix} = \begin{pmatrix} 19 & 22 \\ 43 & 50 \end{pmatrix}$$

$$BA = \begin{pmatrix} 5 & 6 \\ 7 & 8 \end{pmatrix} \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} = \begin{pmatrix} 5 \times 1 + 6 \times 3 & 5 \times 2 + 6 \times 4 \\ 7 \times 1 + 8 \times 3 & 7 \times 2 + 8 \times 4 \end{pmatrix} = \begin{pmatrix} 23 & 34 \\ 31 & 46 \end{pmatrix}$$

一般的に，$AB \neq BA$ となる。

c をスカラーとする。

$$cAB = AcB = (Ac)B = A(cB) = ABc$$

c をどこで掛けても値は変わらない。

連立方程式：

$$\begin{cases} x + 2y = 3 \\ 4x + 5y = 6 \end{cases}$$

行列表示すると，

$$\begin{pmatrix} 1 & 2 \\ 4 & 5 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 3 \\ 6 \end{pmatrix}$$

となる。

また，

$$\begin{cases} x + 2y + 3z = 4 \\ 5x + 6y + 7z = 8 \\ 9x + 10y + 11z = 12 \end{cases}$$

行列表示すると，

$$\begin{pmatrix} 1 & 2 & 3 \\ 5 & 6 & 7 \\ 9 & 10 & 11 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 4 \\ 8 \\ 12 \end{pmatrix}$$

となる。

単位行列： 単位行列とは，対角要素 1，その他 0 となる行列であり，I で表す。

$$I = \begin{pmatrix} 1 & 0 & \cdots & & 0 \\ 0 & 1 & & & \\ \vdots & & \ddots & & \vdots \\ & & & 1 & 0 \\ 0 & & \cdots & 0 & 1 \end{pmatrix}$$

I が $n \times n$ 行列のとき，I_n と書くことも多い。

A を $n \times n$ 行列，x を $n \times 1$ 行列 (ベクトル) とする。

$$I_n A = AI_n = A \qquad I_n x = x$$

$$\begin{pmatrix} 1 & & 0 \\ & \ddots & \\ 0 & & 1 \end{pmatrix} \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix} = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix} \begin{pmatrix} 1 & & 0 \\ & \ddots & \\ 0 & & 1 \end{pmatrix} = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix}$$

$$\begin{pmatrix} 1 & & 0 \\ & \ddots & \\ 0 & & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

逆行列： A を $n \times n$ とする。A の逆行列とは，$AB = I_n$ または $BA = I_n$ となる B を指す。A も B も次元は同じ。B を A^{-1} と表す。
すなわち，A の逆行列は A^{-1} であり，A^{-1} の逆行列は A である。

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

のとき，

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

となる。

$$\begin{aligned} A^{-1}A &= \frac{1}{ad - bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix} \begin{pmatrix} a & b \\ c & d \end{pmatrix} \\ &= \frac{1}{ad - bc} \begin{pmatrix} da - bc & db - bd \\ -ca + ac & -bc + ad \end{pmatrix} \\ &= \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = I_2 \end{aligned}$$

$$\begin{aligned} AA^{-1} &= \begin{pmatrix} a & b \\ c & d \end{pmatrix} \times \frac{1}{ad - bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix} \\ &= \frac{1}{ad - bc} \begin{pmatrix} ad - bc & -ab + ba \\ cd - dc & -cb + da \end{pmatrix} \\ &= \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = I_2 \end{aligned}$$

連立方程式の解： A を $n \times n$ 行列，x と b を $n \times 1$ 行列 (ベクトル) とする。

$$Ax = b$$

両辺に A^{-1} を左から掛ける。

$$A^{-1}Ax = A^{-1}b$$

$A^{-1}A = I_n$ なので，

$$I_n x = A^{-1}b$$

となる。また，

$$I_n x = x$$

なので，x を A, b で表すと，

$$x = A^{-1}b$$

となる。

例

$$\begin{cases} x + 2y = 3 \\ 4x + 5y = 6 \end{cases}$$

の行列表示は，

$$\begin{pmatrix} 1 & 2 \\ 4 & 5 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 3 \\ 6 \end{pmatrix}$$

となる。
x, y の解は，

$$\begin{pmatrix} 1 & 2 \\ 4 & 5 \end{pmatrix}^{-1} \begin{pmatrix} 1 & 2 \\ 4 & 5 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 1 & 2 \\ 4 & 5 \end{pmatrix}^{-1} \begin{pmatrix} 3 \\ 6 \end{pmatrix}$$

なので，

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 1 & 2 \\ 4 & 5 \end{pmatrix}^{-1} \begin{pmatrix} 3 \\ 6 \end{pmatrix}$$

すなわち，

$$\begin{aligned} \begin{pmatrix} x \\ y \end{pmatrix} &= \begin{pmatrix} 1 & 2 \\ 4 & 5 \end{pmatrix}^{-1} \begin{pmatrix} 3 \\ 6 \end{pmatrix} \\ &= \frac{1}{1 \times 5 - 2 \times 4} \begin{pmatrix} 5 & -2 \\ -4 & 1 \end{pmatrix} \begin{pmatrix} 3 \\ 6 \end{pmatrix} \\ &= -\frac{1}{1 \times 3} \begin{pmatrix} 5 \times 3 - 2 \times 6 \\ -4 \times 3 + 1 \times 6 \end{pmatrix} = \begin{pmatrix} -1 \\ 2 \end{pmatrix} \end{aligned}$$

例

$$\begin{cases} x + 2y + 3z = 4 \\ 5x + 6y + 7z = 8 \\ 9x + 10y + 11z = 12 \end{cases}$$

の行列表示は，

$$\begin{pmatrix} 1 & 2 & 3 \\ 5 & 6 & 7 \\ 9 & 10 & 11 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 4 \\ 8 \\ 12 \end{pmatrix}$$

となる。x, y, z の解は，

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 \\ 5 & 6 & 7 \\ 9 & 10 & 11 \end{pmatrix}^{-1} \begin{pmatrix} 4 \\ 8 \\ 12 \end{pmatrix}$$

となる。

転置行列： A を $n \times k$ 行列とする。
A の ij 要素を a_{ij} とする。
A の転置行列 (A' または ${}^tA$) の ij 要素は，a_{ji} となる。

$$A = \begin{pmatrix} a_{11} & \cdots & a_{1k} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nk} \end{pmatrix}$$

$$A' = \begin{pmatrix} a_{11} & \cdots & a_{n1} \\ \vdots & \ddots & \vdots \\ a_{1k} & \cdots & a_{nk} \end{pmatrix}$$

A' は $k \times n$ となる。

$(A')' = A$

$$x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \qquad x' = \begin{pmatrix} x_1 & x_2 & \cdots & x_n \end{pmatrix}$$

5 回帰分析

5.1 重要な公式

1. $\sum_{i=1}^{n} X_i = n\overline{X}$
2. $\sum_{i=1}^{n} (X_i - \overline{X}) = 0$
3. $\sum_{i=1}^{n} (X_i - \overline{X})^2 = \sum_{i=1}^{n} X_i^2 - n\overline{X}^2$
4. $\sum_{i=1}^{n} (X_i - \overline{X})(Y_i - \overline{Y}) = \sum_{i=1}^{n} X_i Y_i - n\overline{X}\,\overline{Y} = \sum_{i=1}^{n} (X_i - \overline{X})Y_i = \sum_{i=1}^{n} (Y_i - \overline{Y})X_i$
5. 2 × 2 行列の逆行列の公式： $\begin{pmatrix} a & b \\ c & d \end{pmatrix}^{-1} = \frac{1}{ad-bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$

5.2 データについて

1. タイム・シリーズ (時系列)・データ： 添え字 i が時間を表す (第 i 期)。t を添え字に使う場合も多い。
2. クロス・セクション (横断面)・データ： 添え字 i が個人や企業を表す (第 i 番目の家計，第 i 番目の企業)。

6 最小二乗法について：単回帰モデル

最小二乗法とは，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である。

6.1 最小二乗法と回帰直線

$(X_1, Y_1), (X_2, Y_2), \cdots, (X_n, Y_n)$ 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，α, β 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。切片 α と傾き β をデータ $\{(X_i, Y_i),\ i = 1, 2, \cdots, n\}$ から推定することを考える。
ある基準の下で，α と β の推定値が求められたとしよう。それぞれ，$\hat{\alpha}$ と $\hat{\beta}$ とする。データ $\{(X_i, Y_i), i = 1, 2, \cdots, n\}$ と直線との関係は，

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

となる。すなわち，実際のデータ Y_i と直線上の値 $\hat{\alpha} + \hat{\beta} X_i$ との間には，誤差 $\hat{u}_i$（残差と呼ばれる）が生じる。

6.2 切片 α と傾き β の求め方

α, β のある推定値を $\hat{\alpha}, \hat{\beta}$ としよう。次のような関数 $S(\hat{\alpha}, \hat{\beta})$ を定義する。

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

これは残差平方和と呼ばれる。
このとき，

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

となるような $\hat{\alpha}, \hat{\beta}$ を求める（最小自乗法)。
最小化のためには，

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

を満たす $\hat{\alpha}, \hat{\beta}$ を求める。
すなわち，$\hat{\alpha}, \hat{\beta}$ は，

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

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

を満たす。
さらに，

$$\sum_{i=1}^{n} Y_i = n\hat{\alpha} + \hat{\beta}\sum_{i=1}^{n} X_i \tag{3}$$

$$\sum_{i=1}^{n} X_i Y_i = \hat{\alpha}\sum_{i=1}^{n} X_i + \hat{\beta}\sum_{i=1}^{n} X_i^2 \tag{4}$$

(3) 式の辺々を n で割って，

$$\frac{1}{n}\sum_{i=1}^{n} Y_i = \hat{\alpha} + \hat{\beta}\frac{1}{n}\sum_{i=1}^{n} X_i$$

すなわち，

$$\overline{Y} = \hat{\alpha} + \hat{\beta}\overline{X} \tag{5}$$

を得る。ただし，

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

とする。
さらに，$\sum_{i=1}^{n} X_i = n\overline{X}$ と (5) 式を利用して，$\hat{\alpha}$ を消去すると，

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

$\hat{\beta}$ で整理して，

$$\begin{aligned}\hat{\beta} &= \frac{\sum_{i=1}^{n} X_i Y_i - n\overline{XY}}{\sum_{i=1}^{n} X_i^2 - n\overline{X}^2} \\ &= \frac{\sum_{i=1}^{n}(X_i - \overline{X})(Y_i - \overline{Y})}{\sum_{i=1}^{n}(X_i - \overline{X})^2} = \frac{S_{XY}}{S_X^2}\end{aligned} \tag{6}$$

が得られ，$\hat{\alpha}$ は (5) 式から，

$$\hat{\alpha} = \overline{Y} - \hat{\beta}\overline{X} \tag{7}$$

となる。ただし，

$$S_{XY} = \frac{1}{n}\sum_{i=1}^{n}(X_i - \overline{X})(Y_i - \overline{Y})$$

$$S_X^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \overline{X})^2$$

とする。

または，行列を用いて解くこともできる。行列表示によって，

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

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

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

さらに，$\hat{\beta}$ について解くと，

$$\begin{aligned}\hat{\beta} &= \frac{n\sum_{i=1}^{n} X_i Y_i - (\sum_{i=1}^{n} X_i)(\sum_{i=1}^{n} Y_i)}{n\sum_{i=1}^{n} X_i^2 - (\sum_{i=1}^{n} X_i)^2} \\ &= \frac{\sum_{i=1}^{n} X_i Y_i - n\overline{X}\,\overline{Y}}{\sum_{i=1}^{n} X_i^2 - n\overline{X}^2} \\ &= \frac{\sum_{i=1}^{n}(X_i - \overline{X})(y_i - \overline{Y})}{\sum_{i=1}^{n}(X_i - \overline{X})^2}\end{aligned}$$

$\hat{\alpha}$ については，

$$\begin{aligned}\hat{\alpha} &= \frac{(\sum_{i=1}^{n} X_i^2)(\sum_{i=1}^{n} Y_i) - (\sum_{i=1}^{n} X_i)(\sum_{i=1}^{n} X_i Y_i)}{n\sum_{i=1}^{n} X_i^2 - (\sum_{i=1}^{n} X_i)^2} \\ &= \frac{\overline{Y}\sum_{i=1}^{n} X_i^2 - \overline{X}\sum_{i=1}^{n} X_i Y_i}{\sum_{i=1}^{n} X_i^2 - n\overline{X}^2} \\ &= \frac{\overline{Y}(\sum_{i=1}^{n} X_i^2 - n\overline{X}^2) - \overline{X}(\sum_{i=1}^{n} X_i Y_i - n\overline{Y}\,\overline{X})}{\sum_{i=1}^{n} X_i^2 - n\overline{X}^2} \\ &= \overline{Y} - \frac{\sum_{i=1}^{n} X_i Y_i - n\overline{Y}\,\overline{X}}{\sum_{i=1}^{n} X_i^2 - n\overline{X}^2}\overline{X} \\ &= \overline{Y} - \hat{\beta}\overline{X}\end{aligned}$$

となる。
回帰直線は，

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

として与えられる。$\hat{Y}_i$ は，X_i を与えたときの Y_i の予測値と解釈される。

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

i	X_i	Y_i
1	5	4
2	1	1
3	3	1
4	2	3
5	4	4

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

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

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

i	X_i	Y_i	X_i^2	X_iY_i
1	5	4	25	20
2	1	1	1	1
3	3	1	9	3
4	2	3	4	6
5	4	4	16	16
合計	$\sum X_i$	$\sum Y_i$	$\sum X_i^2$	$\sum X_iY_i$
	15	13	55	46
平均	$\overline{X}$	$\overline{Y}$		
	3	2.6		

表中では，$\displaystyle\sum_{i=1}^{n}$ を $\sum$ と省略して表記している。

よって，

$$\hat{\beta} = \frac{46 - 5 \times 3 \times 2.6}{55 - 5 \times 3^2} = \frac{7}{10} = 0.7$$
$$\hat{\alpha} = 2.6 - 0.7 \times 3 = 0.5,$$

となる。

注意事項：

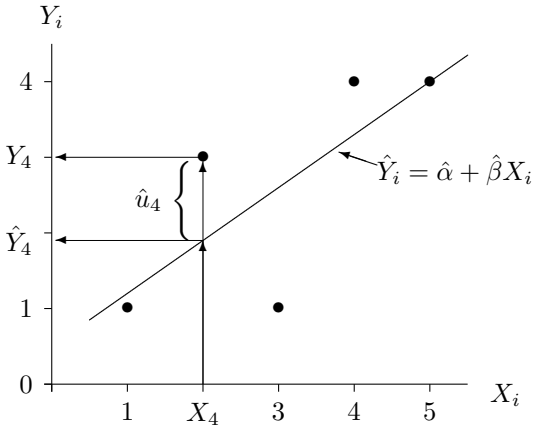
1. α, β は真の値で未知である。
2. $\hat{\alpha}$, $\hat{\beta}$ は α, β の推定値でデータから計算される。

回帰直線は，$\hat{Y}_i = \hat{\alpha} + \hat{\beta} X_i$ であり，上の数値例では，

$$\hat{Y}_i = 0.5 + 0.7 X_i,$$

となる。$\hat{Y}_1$, $\hat{Y}_2$, $\cdots$, $\hat{Y}_5$ として，次の表のように計算される。Y_i，X_i，$\hat{Y}_i$，$\hat{u}_i$ の関係が図 1 に描かれている。

図 1: Y_i，X_i，$\hat{Y}_i$，$\hat{u}_i$ の関係



i	X_i	Y_i	X_i^2	X_iY_i	$\hat{Y}_i$
1	5	4	25	20	4.0
2	1	1	1	1	1.2
3	3	1	9	3	2.6
4	2	3	4	6	1.9
5	4	4	16	16	3.3
合計	$\sum X_i$	$\sum Y_i$	$\sum X_i^2$	$\sum X_iY_i$	$\sum \hat{Y}_i$
	15	13	55	46	13
平均	$\overline{X}$	$\overline{Y}$			
	3	2.6			

$\hat{Y}_i$ を実績値 Y_i の予測値または理論値と呼ぶ。

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

$\hat{u}_i$ を残差と呼ぶ。Y_i, $\hat{Y}_i$, $\hat{u}_i$ の関係，$\hat{Y}_i$, X_i, $\hat{\alpha}$, $\hat{\beta}$ の関係は，

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

の式でまとめられる。

6.3 残差 $\hat{u}_i$ の性質について

$\hat{u}_i = Y_i - \hat{\alpha} - \hat{\beta} X_i$ に注意すると，(1) 式，(2) 式から，

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

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

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

が得られる。なぜなら，

$$\sum_{i=1}^{n}\hat{Y}_i\hat{u}_i=\sum_{i=1}^{n}(\hat{\alpha}+\hat{\beta}X_i)\hat{u}_i=\hat{\alpha}\sum_{i=1}^{n}\hat{u}_i+\hat{\beta}\sum_{i=1}^{n}X_i\hat{u}_i=0$$

となるからである。
数値例で確認してみよう。

i	X_i	Y_i	$\hat{Y}_i$	$\hat{u}_i$	$X_i\hat{u}_i$	$\hat{Y}_i\hat{u}_i$
1	5	4	4.0	0.0	0.0	0.00
2	1	1	1.2	-0.2	-0.2	-0.24
3	3	1	2.6	-1.6	-4.8	-4.16
4	2	3	1.9	1.1	2.2	2.09
5	4	4	3.3	0.7	2.8	2.31
合計	$\sum X_i$	$\sum Y_i$	$\sum \hat{Y}_i$	$\sum \hat{u}_i$	$\sum X_i\hat{u}_i$	$\sum \hat{Y}_i\hat{u}_i$
	15	13	13	0.0	0.0	0.0
平均	$\overline{X}$	$\overline{Y}$				
	3	2.6				

6.4 決定係数 R^2 について

Y_i, $\hat{Y}_i$, $\hat{u}_i$ の関係は，

$$Y_i=\hat{Y}_i+\hat{u}_i,$$

であった。$\overline{Y}$ を両辺から引くと，

$$(Y_i-\overline{Y})=(\hat{Y}_i-\overline{Y})+\hat{u}_i,$$

が得られる。さらに，両辺を二乗して，総和すると，

$$\begin{aligned}&\sum_{i=1}^{n}(Y_i-\overline{Y})^2\\&=\sum_{i=1}^{n}\big((\hat{Y}_i-\overline{Y})+\hat{u}_i\big)^2\\&=\sum_{i=1}^{n}(\hat{Y}_i-\overline{Y})^2+2\sum_{i=1}^{n}(\hat{Y}_i-\overline{Y})\hat{u}_i+\sum_{i=1}^{n}\hat{u}_i^2\\&=\sum_{i=1}^{n}(\hat{Y}_i-\overline{Y})^2+\sum_{i=1}^{n}\hat{u}_i^2\end{aligned}$$

となる。二つ目の等式の右辺第二項では，$\sum_{i=1}^{n}\hat{Y}_i\hat{u}_i=\overline{Y}\sum_{i=1}^{n}\hat{u}_i=0$ が使われている。まとめると，

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

を得る。さらに，両辺を左辺で割ると，

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

が得られる。それぞれの項は，

1. $\sum_{i=1}^{n}(Y_i-\overline{Y})^2\longrightarrow Y_i$ の全変動
2. $\sum_{i=1}^{n}(\hat{Y}_i-\overline{Y})^2\longrightarrow \hat{Y}_i$ (回帰直線) で説明される部分
3. $\sum_{i=1}^{n}\hat{u}_i^2\longrightarrow \hat{Y}_i$ (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 が，

$$R^2=\frac{\sum_{i=1}^{n}(\hat{Y}_i-\overline{Y})^2}{\sum_{i=1}^{n}(Y_i-\overline{Y})^2},\tag{8}$$

のように定義される。R^2 は Y_i のうち $\hat{Y}_i$（または，X_i）で説明できる比率を意味する。または，

$$R^2=1-\frac{\sum_{i=1}^{n}\hat{u}_i^2}{\sum_{i=1}^{n}(Y_i-\overline{Y})^2},\tag{9}$$

として書き換えることもできる。

R^2 の取り得る範囲: さらに，R^2 の取り得る範囲を求める。(8) 式の右辺の分子と分母は共に正なので，$R^2\geq 0$ となる。(9) 式の右辺では 1 から第二項の正の値（分子分母共に正）を差し引いているので，$R^2\leq 1$ となることが分かる。すなわち，R^2 の取り得る範囲は，

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

となる。
$R^2=1$ となる場合はすべての i について $\hat{u}_i=0$ となり，観測されたデータ (X_i,Y_i) は一直線上に並んでいる状態となる。
$R^2=0$ となる場合は二通りが考えられる。一つは，Y_i が X_i に影響されないときで，$\hat{\beta}=0$ の状態，すなわち，データが横軸に平行に一直線上に並んでいる状態となる。もう一つは，データが円状に散布していて，どこにも直線が引けない状態である（ちなみに，データが楕円上に散布している場合は，直線が引ける状態である）。

実際のデータを用いた場合は $R^2=0$ や $R^2=1$ という状況はあり得ない。R^2 が 1 に近づけば回帰式の当てはまりは良い，R^2 が 0 に近づけば回帰式の当てはまりは悪いと言える。しかし,「どの値よりも大きくなるべき」といった基準はない。慣習的には，メドとして 0.9 以上が当てはまりが良いと判断する。
データと R^2 との関係は，後述の 6.5 節で，数値例を挙げながら解説する。

R^2 の別の解釈: R^2 のもう一つの解釈をするために，R^2 の右辺の分子を，

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

と書き換える。最初の等式では，括弧二乗の一つに $\hat{Y}_i=Y_i-\hat{u}_i$ が用いられている。R^2 は,

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

と書き換えられる。この式では，R^2 が Y_i と $\hat{Y}_i$ の相関係数の二乗と解釈されることを意味する。なお，二つ目の等号の右式では，分子と分母に $\sum_{i=1}^n(\hat{Y}_i-\overline{Y})^2$ を掛けていることに注意せよ。
特に，単回帰の場合，$\hat{Y}_i=\hat{\alpha}+\hat{\beta}X_i$ と $\overline{Y}=\hat{\alpha}+\hat{\beta}\overline{X}$ を用いて，

$$\begin{aligned}\sum_{i=1}^n(\hat{Y}_i-\overline{Y})^2 &= \hat{\beta}^2\sum_{i=1}^n(X_i-\overline{X})\\ &= \hat{\beta}\sum_{i=1}^n(X_i-\overline{X})(Y_i-\overline{Y}),\end{aligned}$$

を利用すると，

$$\begin{aligned}R^2 &= \frac{\sum_{i=1}^n(\hat{Y}_i-\overline{Y})^2}{\sum_{i=1}^n(Y_i-\overline{Y})^2}\\ &= \frac{\hat{\beta}^2\sum_{i=1}^n(X_i-\overline{X})^2}{\sum_{i=1}^n(Y_i-\overline{Y})^2}\\ &= \left(\frac{\sum_{i=1}^n(X_i-\overline{X})(Y_i-\overline{Y})}{\sqrt{\sum_{i=1}^n(Y_i-\overline{Y})^2}\sqrt{\sum_{i=1}^n(X_i-\overline{X})^2}}\right)^2\\ &= \frac{S_{XY}^2}{S_X^2S_Y^2},\end{aligned}$$

としても書き換えられる。すなわち，単回帰の場合，決定係数は説明変数 X_i と被説明変数 Y_i との相関係数の二乗となる。

数値例： 決定係数の計算には以下の公式を用いる。

$$R^2=1-\frac{\sum_{i=1}^n\hat{u}_i^2}{\sum_{i=1}^nY_i^2-n\overline{Y}^2}$$

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

i	X_i	Y_i	$\hat{Y}_i$	$\hat{u}_i$	$\hat{u}_i^2$	Y_i^2
1	5	4	4.0	0.0	0.00	16
2	1	1	1.2	−0.2	0.04	1
3	3	1	2.6	−1.6	2.56	1
4	2	3	1.9	1.1	1.21	9
5	4	4	3.3	0.7	0.49	16
合計	$\sum X_i$	$\sum Y_i$	$\sum\hat{Y}_i$	$\sum\hat{u}_i$	$\sum\hat{u}_i^2$	$\sum Y_i^2$
	15	13	13	0.0	4.3	43
平均	$\overline{X}$	$\overline{Y}$				
	3	2.6				

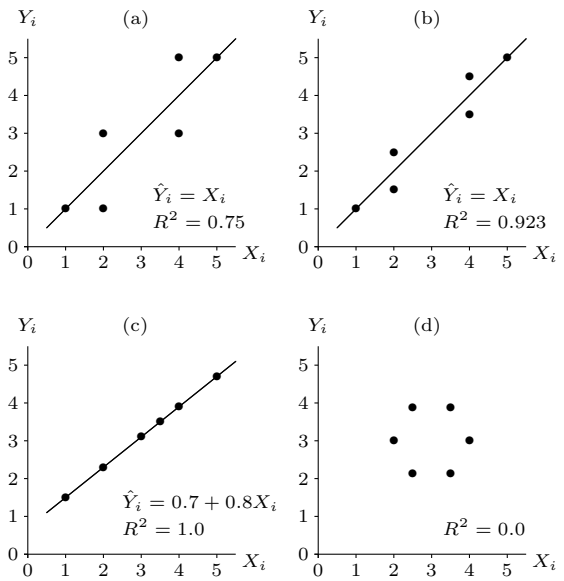
$\overline{Y}=2.6$，$\sum_{i=1}^n\hat{u}_i^2=4.3$，$\sum_{i=1}^nY_i^2=43$ なので，

$$R^2=1-\frac{4.3}{43-5\times2.6^2}=\frac{4.9}{9.2}=0.5326$$

6.5 決定係数の比較

次の数値例を用いて，決定係数の比較を行おう。X と Y のプロットしたものが図 2(a)∼(d) である。

図 2: 決定係数の比較



	(a)		(b)		(c)		(d)	
i	X_i	Y_i	X_i	Y_i	X_i	Y_i	X_i	Y_i
1	1	1	1	1	1	1.5	1	3
2	2	1	2	1.5	2	2.3	2.5	2.134
3	2	3	2	2.5	3	3.1	2.5	3.866
4	4	3	4	3.5	3.5	3.5	3.5	2.134
5	4	5	4	4.5	4	3.9	3.5	3.866
6	5	5	5	5	5	4.7	4	3

(a) と (b) のどちらの場合も，切片･傾きの値は $\hat{\alpha} = 0$，$\hat{\beta} = 1$ として計算されるが，決定係数について，(a) は 0.75，(b) は 0.923 となる（読者はチェックすること）。データのプロットと回帰直線は図 2 の (a) と (b) に描かれている。X_i はどちらも同じ数値とした。横軸 X が 2，4 のケースについて，(b) が (a) より直線に近くなるように，Y の値を変えてみた。(b) のデータの方が (a) より直線に近いために，決定係数が 0.923 と 1 に近い値となっているのが分かる。

(c) はデータが一直線上に並んでいる場合で，決定係数が 1 となる。決定係数がゼロとなるのは (d) の場合で，X と Y との関係を表す直線が描けない場合である。(d) の数値例では，X と Y との関係が円としているが，満遍なく散布している状態と考えてもらえれば良い。

6.6 まとめ

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

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

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。

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

ただし，$\hat{u}_i = Y_i - \hat{\alpha} - \hat{\beta}X_i$ である。計算に必要なものは，$\sum_{i=1}^{n} \hat{u}_i^2$，$\overline{Y}$，$\sum_{i=1}^{n} Y_i^2$ である。

7 最小二乗法について：重回帰モデル

k 変数の多重回帰モデルを考える。

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

X_{ji} は j 番目の説明変数の第 i 番目の観測値を表す。β_1, β_2, $\cdots$, β_k は推定されるべきパラメータである。すべての i について，$X_{1i} = 1$ とすれば，β_1 は定数項として表される。n 組のデータ $(Y_i,\, X_{1i},\, X_{2i},\, \cdots,\, X_{ki})$, $i = 1, 2, \cdots, n$ を用いて，β_1, β_2, $\cdots$, β_k を求める。
ある基準の下で，β_1, β_2, $\cdots$, β_k の解を $\hat{\beta}_1$, $\hat{\beta}_2$, $\cdots$, $\hat{\beta}_k$ としよう。データ $\{(X_i, Y_i),\, i = 1, 2, \cdots, n\}$ と直線との関係は，

$$Y_i = \hat{\beta}_1 X_{1i} + \hat{\beta}_2 X_{2i} + \cdots + \hat{\beta}_k X_{ki} + \hat{u}_i = \hat{Y}_i + \hat{u}_i,$$

となる。すなわち，すべての i について，実際のデータ Y_i と直線上の値 $\hat{Y}_i = \hat{\beta}_1 X_{1i} + \hat{\beta}_2 X_{2i} + \cdots + \hat{\beta}_k X_{ki}$ が一致することはあり得ないので，残差 $\hat{u}_i$ の二乗和を考える。
次のような関数 $S(\hat{\beta}_1, \hat{\beta}_2, \cdots, \hat{\beta}_k)$ を定義する。

$$S(\hat{\beta}_1, \hat{\beta}_2, \cdots, \hat{\beta}_k) = \sum_{i=1}^{n} u_i^2$$
$$= \sum_{i=1}^{n} (Y_i - \hat{\beta}_1 X_{1i} - \hat{\beta}_2 X_{2i} - \cdots - \hat{\beta}_k X_{ki})^2$$

このとき，

$$\min_{\hat{\beta}_1,\, \hat{\beta}_2,\, \cdots,\, \hat{\beta}_k} S(\hat{\beta}_1, \hat{\beta}_2, \cdots, \hat{\beta}_k)$$

となるような $\hat{\beta}_1, \hat{\beta}_2, \cdots, \hat{\beta}_k$ を求める。$\Longrightarrow$ 最小自乗法

最小化のためには，

$$\frac{\partial S(\hat{\beta}_1, \hat{\beta}_2, \cdots, \hat{\beta}_k)}{\partial \hat{\beta}_1} = 0$$
$$\frac{\partial S(\hat{\beta}_1, \hat{\beta}_2, \cdots, \hat{\beta}_k)}{\partial \hat{\beta}_2} = 0$$
$$\vdots$$
$$\frac{\partial S(\hat{\beta}_1, \hat{\beta}_2, \cdots, \hat{\beta}_k)}{\partial \hat{\beta}_k} = 0$$

を満たす $\hat{\beta}_1, \hat{\beta}_2, \cdots, \hat{\beta}_k$ となる。

すなわち，$\hat{\beta}_1, \hat{\beta}_2, \cdots, \hat{\beta}_k$ は，

$$\sum_{i=1}^{n}(Y_i - \hat{\beta}_1 X_{1i} - \hat{\beta}_2 X_{2i} - \cdots - \hat{\beta}_k X_{ki})X_{1i} = 0,$$
$$\sum_{i=1}^{n}(Y_i - \hat{\beta}_1 X_{1i} - \hat{\beta}_2 X_{2i} - \cdots - \hat{\beta}_k X_{ki})X_{2i} = 0,$$
$$\vdots$$
$$\sum_{i=1}^{n}(Y_i - \hat{\beta}_1 X_{1i} - \hat{\beta}_2 X_{2i} - \cdots - \hat{\beta}_k X_{ki})X_{ki} = 0,$$

を満たす。

さらに，

$$\sum_{i=1}^{n} X_{1i}Y_i = \hat{\beta}_1 \sum_{i=1}^{n} X_{1i}^2 + \hat{\beta}_2 \sum_{i=1}^{n} X_{1i}X_{2i} + \cdots + \hat{\beta}_k \sum_{i=1}^{n} X_{1i}X_{ki}$$
$$\sum_{i=1}^{n} X_{2i}Y_i = \hat{\beta}_1 \sum_{i=1}^{n} X_{1i}X_{2i} + \hat{\beta}_2 \sum_{i=1}^{n} X_{2i}^2 + \cdots + \hat{\beta}_k \sum_{i=1}^{n} X_{2i}X_{ki}$$
$$\vdots$$
$$\sum_{i=1}^{n} X_{ki}Y_i = \hat{\beta}_1 \sum_{i=1}^{n} X_{1i}X_{ki} + \hat{\beta}_2 \sum_{i=1}^{n} X_{2i}X_{ki} + \cdots + \hat{\beta}_k \sum_{i=1}^{n} X_{ki}^2$$

行列表示によって，

$$\begin{pmatrix} \sum X_{1i}Y_i \\ \sum X_{2i}Y_i \\ \vdots \\ \sum X_{ki}Y_i \end{pmatrix} = \begin{pmatrix} \sum X_{1i}^2 & \sum X_{1i}X_{2i} & \cdots & \sum X_{1i}X_{ki} \\ \sum X_{1i}X_{2i} & \sum X_{2i}^2 & \cdots & \sum X_{2i}X_{ki} \\ \vdots & \vdots & \ddots & \vdots \\ \sum X_{1i}X_{ki} & \sum X_{2i}X_{ki} & \cdots & \sum X_{ki}^2 \end{pmatrix} \begin{pmatrix} \hat{\beta}_1 \\ \hat{\beta}_2 \\ \vdots \\ \hat{\beta}_k \end{pmatrix}$$

が得られる。

$\hat{\beta}_1, \hat{\beta}_2, \cdots, \hat{\beta}_k$ についてまとめると，

$$\begin{pmatrix} \hat{\beta}_1 \\ \hat{\beta}_2 \\ \vdots \\ \hat{\beta}_k \end{pmatrix} = \begin{pmatrix} \sum X_{1i}^2 & \sum X_{1i}X_{2i} & \cdots & \sum X_{1i}X_{ki} \\ \sum X_{1i}X_{2i} & \sum X_{2i}^2 & \cdots & \sum X_{2i}X_{ki} \\ \vdots & \vdots & \ddots & \vdots \\ \sum X_{1i}X_{ki} & \sum X_{2i}X_{ki} & \cdots & \sum X_{ki}^2 \end{pmatrix}^{-1} \begin{pmatrix} \sum X_{1i}Y_i \\ \sum X_{2i}Y_i \\ \vdots \\ \sum X_{ki}Y_i \end{pmatrix}$$

を解くことになる。$\Longrightarrow$ コンピュータによって計算

$\sum_{i=1}^{n} X_{ji}X_{li}$，$\sum_{i=1}^{n} X_{ji}Y_i$ をそれぞれ $\sum X_{ji}X_{li}$，$\sum X_{ji}Y_i$ と表記する。

ただし，$j = 1, 2$，$l = 1, 2$ とする。

7.1 重回帰モデルにおける回帰係数の意味

結論： 他の変数の影響を取り除いての被説明変数への影響を表す。

$k = 2$ の単純なモデル：

$$Y_i = \beta_1 X_{1i} + \beta_2 X_{2i} + u_i, \qquad i = 1, 2, \cdots, n$$

β_1, β_2 の最小二乗推定量は，

$$\min_{\beta_1,\ \beta_2} \sum_{i=1}^{n}(Y_i - \beta_1 X_{1i} - \beta_2 X_{2i})^2$$

を解いて，$\hat{\beta}_1$，$\hat{\beta}_2$ が次のように得られる。

$$\begin{pmatrix} \hat{\beta}_1 \\ \hat{\beta}_2 \end{pmatrix} = \begin{pmatrix} \sum X_{1i}^2 & \sum X_{1i}X_{2i} \\ \sum X_{1i}X_{2i} & \sum X_{2i}^2 \end{pmatrix}^{-1} \begin{pmatrix} \sum X_{1i}Y_i \\ \sum X_{2i}Y_i \end{pmatrix}$$
$$= \frac{1}{(\sum X_{1i}^2)(\sum X_{2i}^2) - (\sum X_{1i}X_{2i})^2} \times \begin{pmatrix} \sum X_{2i}^2 & -\sum X_{1i}X_{2i} \\ -\sum X_{1i}X_{2i} & \sum X_{1i}^2 \end{pmatrix} \begin{pmatrix} \sum X_{1i}Y \\ \sum X_{2i}Y \end{pmatrix}$$
$$= \begin{pmatrix} \dfrac{(\sum X_{2i}^2)(\sum X_{1i}Y_i) - (\sum X_{1i}X_{2i})(\sum X_{2i}Y_i)}{(\sum X_{1i}^2)(\sum X_{2i}^2) - (\sum X_{1i}X_{2i})^2} \\ \dfrac{-(\sum X_{1i}X_{2i})(\sum X_{1i}Y_i) + (\sum X_{1i}^2)(\sum X_{2i}Y_i)}{(\sum X_{1i}^2)(\sum X_{2i}^2) - (\sum X_{1i}X_{2i})^2} \end{pmatrix}$$

一方，次の 2 つの回帰式を考える。

$$Y_i = \alpha_1 X_{2i} + v_i$$
$$X_{1i} = \alpha_2 X_{2i} + w_i$$

α_1，α_2 のそれぞれの最小二乗推定量を求めると，

$$\hat{\alpha}_1 = \frac{\sum X_{2i}Y_i}{\sum X_{2i}^2}, \qquad \hat{\alpha}_2 = \frac{\sum X_{2i}X_{1i}}{\sum X_{2i}^2}$$

となる。

$\hat{\alpha}_1$，$\hat{\alpha}_2$ を用いて，残差 $\hat{v}_i$，$\hat{w}_i$ を下記のようにそれぞれ求める。

$$\hat{v}_i = Y_i - \hat{\alpha}_1 X_{2i}, \qquad \hat{w}_i = X_{1i} - \hat{\alpha}_2 X_{2i}$$

$\hat{v}_i$，$\hat{w}_i$ は Y_i，X_{1i} から X_{2i} の影響を取り除いたものと解釈できる。

更に，次の回帰式を考える。

$$\hat{v}_i = \gamma \hat{w}_i + \epsilon_i$$

γ の最小二乗推定量 $\hat{\gamma}$ は $\hat{\beta}_1$ に一致することを示す。

$$\begin{aligned}\hat{\gamma} &= \frac{\sum \hat{w}_i \hat{v}_i}{\sum \hat{w}_i^2} = \frac{\sum (X_{1i} - \hat{\alpha}_2 X_{2i})(Y_i - \hat{\alpha}_1 X_{2i})}{\sum (X_{1i} - \hat{\alpha}_2 X_{2i})^2} \\ &= \frac{\sum X_{1i}Y_i - \hat{\alpha}_1 \sum X_{1i}X_{2i} - \hat{\alpha}_2 \sum X_{2i}Y_i + \hat{\alpha}_1\hat{\alpha}_2 \sum X_{2i}^2}{\sum X_{1i}^2 - 2\hat{\alpha}_2 \sum X_{1i}X_{2i} + \hat{\alpha}_2^2 \sum X_{2i}^2} \\ &= \frac{\sum X_{1i}Y_i - \dfrac{(\sum X_{2i}Y_i)(\sum X_{1i}X_{2i})}{\sum X_{2i}^2}}{\sum X_{1i}^2 - \dfrac{(\sum X_{1i}X_{2i})^2}{\sum X_{2i}^2}} \\ &= \frac{(\sum X_{2i}^2)(\sum X_{1i}Y_i) - (\sum X_{1i}X_{2i})(\sum X_{2i}Y_i)}{(\sum X_{1i}^2)(\sum X_{2i}^2) - (\sum X_{1i}X_{2i})^2} = \hat{\beta}_1,\end{aligned}$$

「Y_i から X_{2i} の影響を取り除いた変数」を被説明変数,「X_{1i} から X_{2i} の影響を取り除いた変数」を説明変数とした回帰係数が β_1 に等しい。

一般化： 次の回帰モデルを考える。

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

j 番目の回帰係数 β_j の意味は,「Y_i から $X_{1i}, \cdots, X_{j-1,i}, X_{j+1,i}, \cdots, X_{ki}$ (すなわち，X_{ji} 以外の説明変数) の影響を取り除いた変数」を被説明変数,「X_{ji} から $X_{1i}, \cdots, X_{j-1,i}, X_{j+1,i}, \cdots, X_{ki}$ (すなわち，X_{ji} 以外の説明変数) の影響を取り除いた変数」を説明変数とした回帰係数となる。

7.2 決定係数 R^2 と自由度修正済み決定係数 $\overline{R}^2$ について

また，決定係数 R^2 についても同様に表される。

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

ただし，$\hat{Y}_i = \hat{\beta}_1 X_{1i} + \hat{\beta}_2 X_{2i} + \cdots + \hat{\beta}_k X_{ki}$，$Y_i = \hat{Y}_i + \hat{u}_i$ である。

R^2 は，説明変数を増やすことによって，必ず大きくなる。なぜなら，説明変数が増えることによって，$\sum_{i=1}^n \hat{u}_i^2$ が必ず減少するからである。

R^2 を基準にすると，被説明変数にとって意味のない変数でも，説明変数が多いほど，よりよいモデルということになる。この点を改善するために，自由度修正済み決定係数 $\overline{R}^2$ を用いる。

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

$\sum_{i=1}^n \hat{u}_i^2/(n-k)$ は u_i の分散 σ^2 の不偏推定量であり，$\sum_{i=1}^n (Y_i - \overline{Y})^2/(n-1)$ は Y_i の分散の不偏推定量である。分散や不偏推定量の意味は，統計学の知識を必要とし，後述する。

R^2 と $\overline{R}^2$ との関係は，

$$\overline{R}^2 = 1 - (1 - R^2)\frac{n-1}{n-k},$$

となる。さらに，

$$\frac{1 - \overline{R}^2}{1 - R^2} = \frac{n-1}{n-k} \geq 1,$$

という関係から，$\overline{R}^2 \leq R^2$ という結果を得る。($k = 1$ のときのみに，等号が成り立つ。)

数値例： 今までと同じ数値例で，$\overline{R}^2$ を計算する。

i	X_i	Y_i	$\hat{Y}_i$	$\hat{u}_i$	$\hat{u}_i^2$	Y_i^2
1	5	4	4.0	0.0	0.00	16
2	1	1	1.2	−0.2	0.04	1
3	3	1	2.6	−1.6	2.56	1
4	2	3	1.9	1.1	1.21	9
5	4	4	3.3	0.7	0.49	16
合計	$\sum X_i$	$\sum Y_i$	$\sum \hat{Y}_i$	$\sum \hat{u}_i$	$\sum \hat{u}_i^2$	$\sum Y_i^2$
	15	13	13	0.0	4.3	43
平均	$\overline{X}$	$\overline{Y}$				
	3	2.6				

$\overline{Y} = 2.6$，$\sum_{i=1}^n \hat{u}_i^2 = 4.3$，$\sum_{i=1}^n Y_i^2 = 43$ なので，

$$\begin{aligned}R^2 &= 1 - \frac{\sum \hat{u}_i^2}{\sum Y_i^2 - n\overline{Y}^2} = 1 - \frac{4.3}{43 - 5 \times 2.6^2} \\ &= 1 - \frac{4.3}{9.2} = 0.5326\end{aligned}$$

となり，$\overline{R}^2$ は，

$$\begin{aligned}\overline{R}^2 &= 1 - \frac{\sum \hat{u}_i^2/(n-k)}{(\sum Y_i^2 - n\overline{Y}^2)/(n-1)} \\ &= 1 - \frac{4.3/(5-2)}{9.2/(5-1)} = 0.3768\end{aligned}$$

となる。

自由度について： 分子について，残差 $\hat{u}_i$ を求めるためには，$\hat{\beta}_1, \hat{\beta}_2, \cdots, \hat{\beta}_k$ の k 個の推定値を得なければならない。データ数 n から推定値の数 k を差し引いたものを**自由度** (degree of freedom) と呼ぶ。
一方，分母については，X_{1i} が定数項だとして，Y_i が定数項を除く $X_{2i}, X_{3i}, \cdots, X_{ki}$ に依存しない場合を考える。この場合，$\beta_2 = \beta_3 = \cdots = \beta_k = 0$ とするので，$\hat{u}_i = Y_i - \hat{\beta}_1$ となる。$\hat{u}_i$ を得るためには $\hat{\beta}_1$ だけを求めればよい。最小二乗法の考え方に沿って求めれば，$\hat{\beta}_1 = \overline{Y}$ となる（読者は確認すること）。すなわち，自由度は「データ数 $-$ 推定値の数 $= n - 1$」ということになる。
このように，決定係数の第二項目の分子・分母をそれぞれの自由度で割ることによって，自由度修正済み決定係数が得られる。

注意： R^2 や $\overline{R}^2$ を比較する場合，被説明変数が同じであることが重要である。被説明変数が対数かまたはそのままの値であれば，決定係数・自由度修正済み決定係数の大小比較は意味をなさない。ただし，被説明変数が異なる場合であっても，被説明変数を上昇率とするかそのままの値を用いるかの比較では，決定係数・自由度修正済み決定係数の大小比較はできないが，誤差項 u_i の標準誤差での比較は可能である (標準誤差の小さいモデルを採用する)。$\Longrightarrow$ 関数型の選択