

第 24 回 回帰分析 (3.4, 13.1–13.2.1)

村澤 康友

2026 年 1 月 6 日

今日のポイント

1. $E(Y|X)$ を与える式を, Y の X 上への回帰モデルという. 線形な回帰モデルを線形回帰モデルといふ. 線形回帰モデルの説明変数の係数を回帰係数といふ.
2. Y の X 上への線形回帰モデルにおける回帰係数は X から Y への限界効果を表す. $\ln Y$ の $\ln X$ 上への線形回帰モデルにおける回帰係数は X に対する Y の弾力性を表す.
3. 残差 2 乗和を最小にするように回帰係数を定める方法を通常の最小 2 乗法 (OLS) といふ. OLS 問題の 1 階の条件を整理した式を正規方程式といふ. OLS 問題の解を回帰係数の OLS 推定量 (値) といふ.

4 今日のキーワード 4

5 次回までの準備 4

1 回帰モデル

1.1 回帰 (p. 258)

(X, Y) を確率ベクトルとする. X から Y を予測したい (身長→体重, 所得→消費など).

定義 1. $E(Y|X)$ を求ることを, Y を X に回帰するといふ.

注 1. X がカテゴリー変数ならカテゴリーごとの平均を求めるだけ.

注 2. $F_{Y|X}(\cdot)$ が求まれば理想的.

1.2 回帰モデル (p. 258)

定義 2. $E(Y|X)$ を与える式を, Y の X 上への回帰モデル (回帰式, 回帰関数) といふ.

注 3. すなわち

$$E(Y|X) = r(X)$$

定義 3. 説明する方の変数を説明変数といふ.

定義 4. 説明される方の変数を被説明変数といふ.

定義 5. $U := Y - E(Y|X)$ を回帰の誤差項といふ.

注 4. 誤差項を用いて回帰モデルを表すと

$$Y = r(X) + U$$

$$E(U|X) = 0$$

1	回帰モデル	1
1.1	回帰 (p. 258)	1
1.2	回帰モデル (p. 258)	1
1.3	線形回帰モデル (p. 258)	1
1.4	単回帰と重回帰 (p. 270)	2
2	限界効果と弾力性	2
2.1	限界効果	2
2.2	弾力性 (p. 259, p. 277)	2
3	最小 2 乗法	3
3.1	定数項なしの単回帰モデル	3
3.2	定数項ありの単回帰モデル (p. 260)	3

注 5. すなわち

$$E(Y|X) = \alpha + \beta X$$

注 6. $X, Y > 0$ なら $\ln Y$ の $\ln X$ 上への線形回帰モデルを考えることも多い。すなわち

$$E(\ln Y|\ln X) = \alpha + \beta \ln X$$

定義 7. 線形回帰モデルの説明変数の係数を**回帰係数**という。

定理 1. $(X, Y)' \sim N(\mu, \Sigma)$ とする。ただし

$$\mu := \begin{pmatrix} \mu_X \\ \mu_Y \end{pmatrix}, \quad \Sigma := \begin{bmatrix} \sigma_X^2 & \sigma_{XY} \\ \sigma_{YX} & \sigma_Y^2 \end{bmatrix}$$

このとき

$$Y|X \sim N\left(\mu_{Y|X}, \sigma_{Y|X}^2\right)$$

ただし

$$\mu_{Y|X} := \mu_Y + \frac{\sigma_{XY}}{\sigma_X^2}(X - \mu_X)$$

$$\sigma_{Y|X}^2 := \sigma_Y^2 - \frac{\sigma_{XY}^2}{\sigma_X^2}$$

証明。同時 pdf を周辺 pdf で割って条件付き pdf を求める。詳細は略。 \square

注 7. $E(Y|X)$ は X の 1 次式。 $\text{var}(Y|X)$ は X に依存しない。

1.4 単回帰と重回帰 (p. 270)

定義 8. 定数項以外に説明変数が 1 つしかない線形回帰モデルを**単回帰モデル**という。

定義 9. 定数項以外に説明変数が複数ある線形回帰モデルを**重回帰モデル**という。

注 8. すなわち

$$E(Y|X_1, \dots, X_k) = \alpha + \beta_1 X_1 + \dots + \beta_k X_k$$

定義 10. 重回帰モデルの回帰係数を**偏回帰係数**といふ。

2 限界効果と弾力性

2.1 限界効果

定義 11. X の 1 単位の増加に対する Y の変化を X から Y への**限界効果**といふ。

定理 2. Y の X 上への線形回帰モデルにおける回帰係数は X から Y への限界効果を表す。

証明。 Y の X 上への線形回帰モデルは

$$Y = \alpha + \beta X + U$$

$$E(U|X) = 0$$

X が 1 単位増えると Y は β 単位増える (X が連続なら微分する)。 \square

2.2 弾力性 (p. 259, p. 277)

定義 12. X の 1 %の増加に対する Y の変化率を X に対する Y の**弾力性**といふ。

注 9. 式で表すと

$$\epsilon := \frac{dY/Y}{dX/X} \approx \frac{\Delta Y/Y}{\Delta X/X}$$

定理 3. $\ln Y$ の $\ln X$ 上への線形回帰モデルにおける回帰係数は X に対する Y の弾力性を表す。

証明。 $\ln Y$ の $\ln X$ 上への線形回帰モデルは

$$\ln Y = \alpha + \beta \ln X + U$$

$$E(U|\ln X) = 0$$

したがって

$$\begin{aligned} \beta &= \frac{d \ln Y}{d \ln X} \\ &= \frac{dX}{d \ln X} \frac{dY}{dX} \frac{d \ln Y}{dY} \\ &= \left(\frac{d \ln X}{dX} \right)^{-1} \frac{dY}{dX} \frac{d \ln Y}{dY} \\ &= \left(\frac{1}{X} \right)^{-1} \frac{dY}{dX} \frac{1}{Y} \\ &= \frac{dY/Y}{dX/X} \end{aligned}$$

\square

3 最小 2 乗法

3.1 定数項なしの单回帰モデル

2 変量データを $((y_1, x_1), \dots, (y_n, x_n))$ とする.
 y_i の x_i 上への定数項なしの单回帰モデルは

$$E(y_i|x_i) = \beta x_i$$

β を求めたい (図 1 左).

定義 13. 実際の y_i と適当な回帰係数を用いた y_i の予測値との差を y_i の **残差**という.

注 10. $\beta = b$ としたら y_i の残差は

$$e_i := y_i - bx_i$$

誤差 $u_i := y_i - \beta x_i$ とは異なる.

定義 14. 残差 2 乗和を最小にするように回帰係数を定める方法を**通常の最小 2 乗法** (*Ordinary Least Squares, OLS*) という.

注 11. すなわち OLS 問題は

$$\begin{aligned} \min_b & \sum_{i=1}^n (y_i - bx_i)^2 \\ \text{and } & b \in \mathbb{R} \end{aligned}$$

ただし \mathbb{R} は実数の集合.

定義 15. OLS 問題の 1 階の条件を整理した式を**正規方程式**といふ.

注 12. 残差 2 乗和は b に関する凸関数なので, 1 階の条件は最小化の必要十分条件.

注 13. OLS 問題の 1 階の条件は

$$\sum_{i=1}^n (-x_i)2(y_i - b^*x_i) = 0$$

すなわち

$$\sum_{i=1}^n x_i(y_i - b^*x_i) = 0$$

正規方程式は

$$\sum_{i=1}^n x_i y_i - b^* \sum_{i=1}^n x_i^2 = 0$$

定義 16. OLS 問題の解を回帰係数の **OLS 推定量 (値)** という.

注 14. 正規方程式より $\sum_{i=1}^n x_i^2 \neq 0$ なら β の OLS 推定量は

$$b^* = \frac{\sum_{i=1}^n x_i y_i}{\sum_{i=1}^n x_i^2}$$

$\sum_{i=1}^n x_i^2 = 0$ だと解は存在しない.

定義 17. 回帰係数の OLS 推定量 (値) を用いた予測を**回帰予測**といふ.

注 15. y_i の回帰予測は $\hat{y}_i := b^*x_i$.

定義 18. 実際の値と回帰予測との差を**回帰残差**といふ.

注 16. y_i の回帰残差は $e_i^* := y_i - \hat{y}_i$.

3.2 定数項ありの单回帰モデル (p. 260)

y_i の x_i 上への定数項ありの单回帰モデルは

$$E(y_i|x_i) = \alpha + \beta x_i$$

(α, β) を求めたい (図 1 右). $(\alpha, \beta) = (a, b)$ とし
たら y_i の残差は

$$e_i := y_i - a - bx_i$$

OLS 問題は

$$\begin{aligned} \min_{a,b} & \sum_{i=1}^n (y_i - a - bx_i)^2 \\ \text{and } & a, b \in \mathbb{R} \end{aligned}$$

1 階の条件は

$$\begin{aligned} \sum_{i=1}^n (-1)2(y_i - a^* - b^*x_i) &= 0 \\ \sum_{i=1}^n (-x_i)2(y_i - a^* - b^*x_i) &= 0 \end{aligned}$$

すなわち

$$\begin{aligned} \sum_{i=1}^n (y_i - a^* - b^*x_i) &= 0 \\ \sum_{i=1}^n x_i(y_i - a^* - b^*x_i) &= 0 \end{aligned}$$

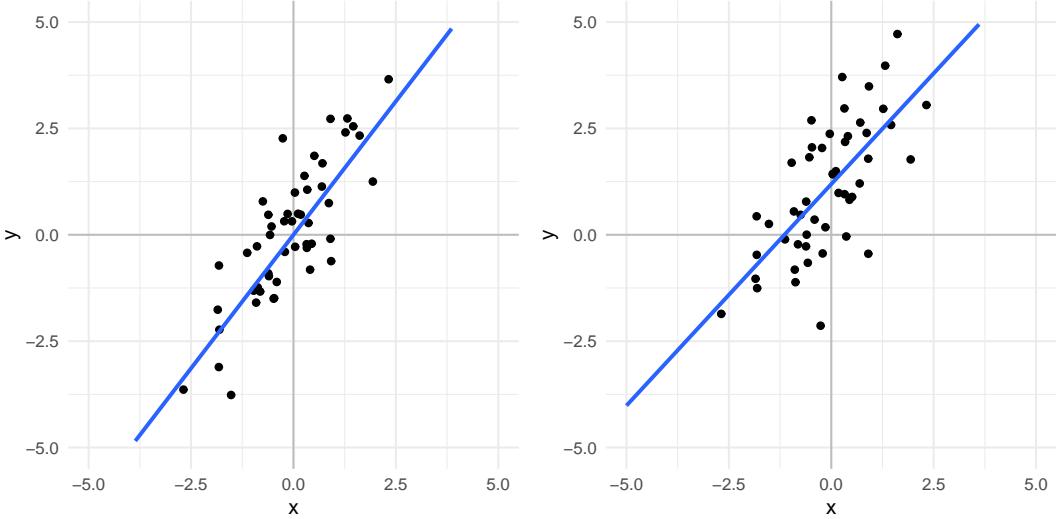


図 1 定数項なしの単回帰モデル（左）と定数項ありの単回帰モデル（右）

正規方程式は

$$\begin{aligned} \sum_{i=1}^n y_i - na^* - b^* \sum_{i=1}^n x_i &= 0 \\ \sum_{i=1}^n x_i y_i - a^* \sum_{i=1}^n x_i - b^* \sum_{i=1}^n x_i^2 &= 0 \end{aligned}$$

この連立方程式の解が (α, β) の OLS 推定値.

注 17. 正規方程式の第 1 式を n で割ると

$$\bar{y} - a^* - b^* \bar{x} = 0$$

したがって $a^* = \bar{y} - b^* \bar{x}$. 1 階の条件より

$$\begin{aligned} \sum_{i=1}^n (x_i - \bar{x})(y_i - a^* - b^* x_i) \\ = \sum_{i=1}^n x_i(y_i - a^* - b^* x_i) - \bar{x} \sum_{i=1}^n (y_i - a^* - b^* x_i) \\ = 0 \end{aligned}$$

$a^* = \bar{y} - b^* \bar{x}$ を左辺に代入すると

$$\sum_{i=1}^n (x_i - \bar{x})[(y_i - \bar{y}) - b^*(x_i - \bar{x})] = 0$$

これを解くと

$$b^* = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^n (x_i - \bar{x})^2}$$

これは σ_{XY}/σ_X^2 の推定量（値）と理解できる.

例 1. 某大学 1 年生の大学での GPA (colgpa) の高校での GPA (hsgpa) への単回帰（図 2）.

$$\widehat{\text{colgpa}} = 0.920577 + 0.524173 \cdot \text{hsgpa}$$

4 今日のキーワード

回帰, 回帰モデル, 説明変数, 被説明変数, 誤差項, 線形回帰モデル, 回帰係数, 単回帰モデル, 重回帰モデル, 偏回帰係数, 限界効果, 弾力性, 残差, 通常の最小 2 乗法 (OLS), 正規方程式, OLS 推定量 (値), 回帰予測, 回帰 (OLS) 残差

5 次回までの準備

復習 教科書第 3 章 4 節, 第 13 章 1-2.1 節, 復習テスト 24

予習 教科書第 13 章 2.2-3 節

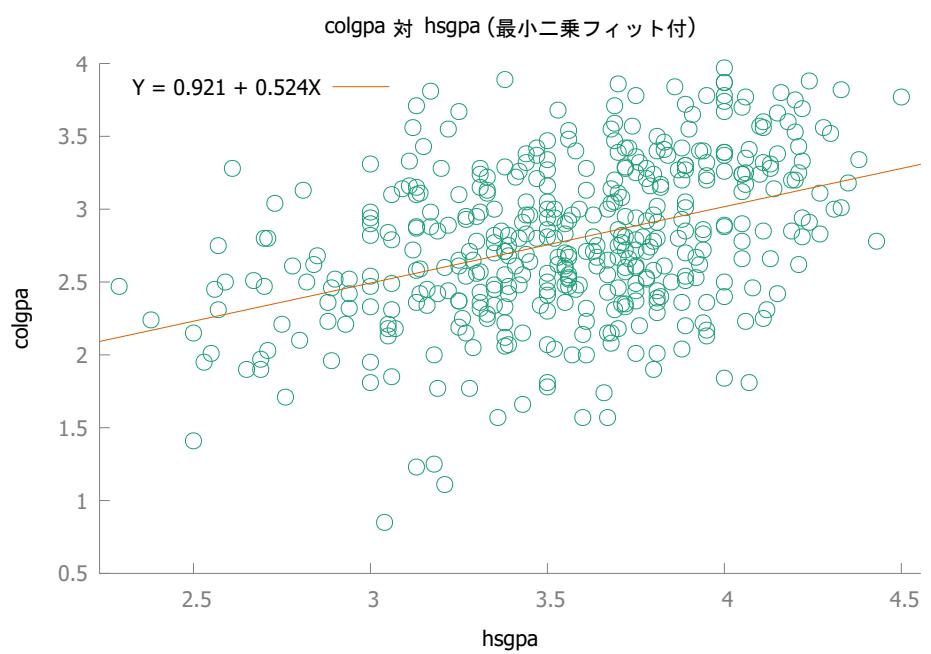


図 2 某大学 1 年生の大学での GPA (colgpa) の高校での GPA (hsgpa) への単回帰