10/3(一) 出題
12/12(一) 出題(一)
10/14(五) 3,4 補課
1/6(五) 3,4 確認

期中考 11/7(一) (0補考) 期末考 1/9(一) (0補考) 自主學習週 11/21-25.

回歸分析 · Montgomery Peck & Vining

Introduction to linear regression analysis's 5th ed

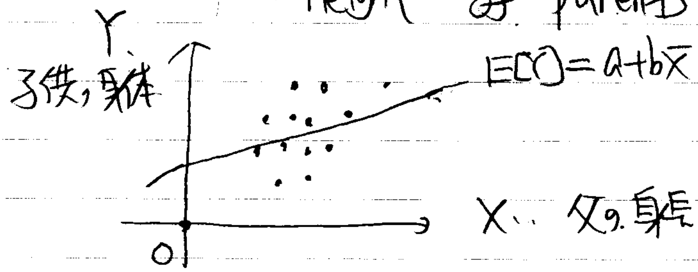
· 作業、期中考、期末考 { 、線形代數
  30%, 30%, 40% 、微積分
(資料分析 etc) 、確率統計
(、R
、SAS

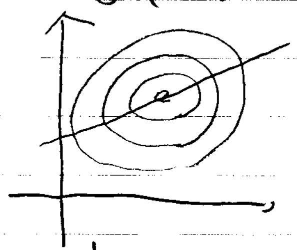
Regression. finding the relationship between variables discovering how variables affect other variables. (response variables or dependent, or dependent variables.)

(covariates, independent variables, predictor variables or regressors

· Francis Galton. フランシス、ゴルトン.

… height of parents / height of offsprings.





regression line. regression towards mediocrity

Chuyu culture

Model (モデル) $Y = f(X) + \varepsilon$

- X… independent variable
- Y… dependent variable
- ε… random error, $E[\varepsilon | X] = 0$

$$E[Y|X] = E[f(X) + \varepsilon | X] = f(X)$$

(regression function)

Y = response variable (反応変数)

- quantitative … 量的 … { discrete (離散), continuous (連続) }
- qualitative … 質的 … { nominal … (名義), oridinal (順序) }

model…

1. f … unknown
   - smooth or piecewise smooth
   - non-parametric model
2. $f = g(x|\theta)$ (g: known, θ: unknown)
3. $f(X) = \underbrace{\beta_0}_{\text{切片}} + \underline{\beta_1} X$ simple linear regression (単回帰 model)

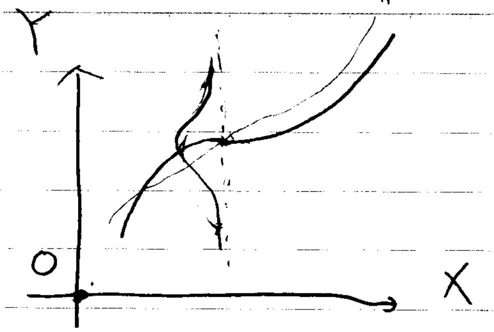
data: (X_i, Y_i) 数据

$(X_1, Y_1), (X_2, Y_2), \ldots (X_n, Y_n)$

9/12
[8]

No.
Date 9 / 12

$Y = f(X) + \varepsilon$. true



Data. $(X_1, Y_1) - (X_n, Y_n)$

- interpolation
- extrapolation.

Yの条件付分布 $(Y|X)$

- Model specification
- Model fitting
- Model checking
- Model validation.

• Data Collection (データの収集)

1. retrospective study (回顧的) → 過去のデータに基づくということだろうか?
2. observational study
3. designed experiment (デザインされた実験)

- multiple covariates

$Y = f(X_1, X_2 \dots X_k) + \varepsilon$

$\alpha + f_1(X_1) + \dots f_k(X_k)$

$\beta_0 + \beta_1 X_1 + \dots \beta_k X_k$

linear regression

パラメータの線形

linear in parameters

X_1, X_2

(- $y = \beta_0 + \beta_1 x_1^2 + \varepsilon$
- $y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_1^2 + \dots \varepsilon$

これらも線形回帰モデル

No.
Date : :

Simple linear regression.（線形単回帰）

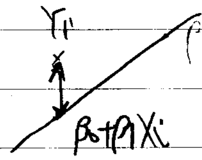
Data... $(X_1,Y_1)\ldots,(X_n,Y_n)$ are coming from $Y=\beta_0+\beta_1X+\varepsilon$
$E[\varepsilon]=0$

つまり $Y_i=\beta_0+\beta_1X_i+\underbrace{\varepsilon_i}_{\text{i.i.d}}$ $(i=1,2,3\ldots)$ という事である

$E[\varepsilon_i]=0$

Least Squares Fitting...（最小二乗法）

$$\arg\min_{\beta_0,\beta_1}\sum_{i=1}^{n}(Y_i-\beta_0-\beta_1X_i)^2 \to F(\beta_0,\beta_1)$$



（これの解は $\hat{\beta}_0,\hat{\beta}_1$ と書く）

$$\underbrace{\frac{\partial F(\beta_0,\beta_1)}{\partial\beta_0}=0,\ \frac{\partial F(\beta_0,\beta_1)}{\partial\beta_1}=0}$$

これを計算すると
$$\begin{cases}-2\sum_{i=1}^{n}(Y_i-\beta_0-\beta_1X_i)=0\\ -2\sum_{i=1}^{n}(Y_i-\beta_0-\beta_1X_i)(-X_i)=0\end{cases}$$

となり
$$\begin{cases}\beta_0=\bar{Y}-\beta_1\bar{X}\\ \sum X_iY_i-\beta_0\sum_{i=1}^{n}X_i-\beta_1\sum_{i=1}^{n}X_i^2=0\end{cases}$$

注 $S_{xx}=\sum(X_i-\bar{X})^2$
$S_{xy}=\sum(X_i-\bar{X})(Y_i-\bar{Y})$

$\hat{\beta}_1=\frac{S_{xy}}{S_{xx}}$, $\hat{\beta}_0=\bar{Y}-\frac{S_{xy}}{S_{xx}}\bar{X}$

が得られる。

$E[\varepsilon_i]=0$

$V[\varepsilon_i]=\sigma^2$

Gauss-Markov Conditions

$E[\varepsilon_i\varepsilon_j]=0\quad(i\neq j)$

No.

Date 9 12 ⑩

このとき $\hat\beta_0, \hat\beta_1$ を最小二乗推定量 (Least Square Estimator) と言う

② $E[\hat\beta_0]=E[\bar Y-\hat\beta_1\bar X]=\beta_0+\beta_1\bar X-\beta_1\bar x=\beta_0$

① $E[\hat\beta_1]=E\left[\frac{S_{xy}}{S_{xx}}\right]=\frac{\sum_{i=1}^n x_i(\beta_0+\beta_1x_i)-\bar X\sum_{i=1}^n(\beta_0+\beta_1x_i)}{S_{xx}}=\frac{\beta_1\left\{\sum_{i=1}^n x_i^2-\frac{(\sum x_i)^2}{n}\right\}}{S_{xx}}$

$=\beta_1$

次に分散に関して、$V[\hat\beta_1]=V\left[\frac{S_{xy}}{S_{xx}}\right]=V\left[\sum_{i=1}^n\frac{(x_i-\bar x)Y_i}{S_{xx}}\right]$

いま $Y_1\sim Y_n$ は互いに独立であるので $\sum_{i=1}^n V\left[\frac{(x_i-\bar x)Y_i}{S_{xx}}\right]$

$(\because \varepsilon_i : \text{iid})$

$$\sum_{i=1}^n\frac{(x_i-\bar x)^2}{(S_{xx})^2}\underbrace{V[Y_i]}_{\sigma^2}=\frac{1}{(S_{xx})^2}\sum_{i=1}^n\sigma^2(x_i-\bar x)^2$$

$=\frac{\sigma^2}{S_{xx}}$ となる

$$V[\hat\beta_0]=Var[\bar Y-\hat\beta_1\bar x]=\underbrace{Var[\bar Y]}_{\frac{\sigma^2}{n}}+\underbrace{\bar x^2\,Var[\hat\beta_1]}_{\frac{\sigma^2}{S_{xx}}\bar x^2}-\underbrace{2\bar x\cdot cov[\bar Y,\hat\beta_1]}_{0}$$

$$\left(\frac{1}{n}+\frac{\bar x^2}{S_{xx}}\right)\sigma^2$$

以下、行列を用いても

$$\underbrace{\begin{pmatrix}Y_1\\ \vdots\\ Y_n\end{pmatrix}}_{Y}=\underbrace{\begin{pmatrix}1 & x_1\\ \vdots & \vdots\\ 1 & x_n\end{pmatrix}}_{X}\underbrace{\begin{pmatrix}\beta_0\\ \beta_1\end{pmatrix}}_{\beta}+\begin{pmatrix}\varepsilon_1\\ \vdots\\ \varepsilon_n\end{pmatrix}$$

$$L=(Y-X\beta)^t(Y-X\beta)=(Y^t-\beta^tX^t)(Y-X\beta)=Y^tY-\beta^tX^tY-Y^tX\beta+\beta^tX^tX\beta$$

No.
Date 9 : 12 (10)

residuals … $e_i = Y_i - \hat{Y}_i$ (i=1~n)
残差

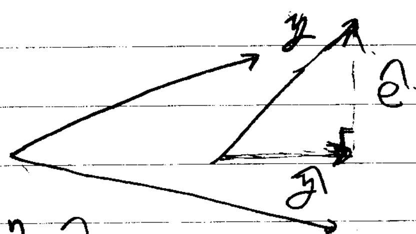
・真のmodel …… $Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i$
・仮定のmodel … $\hat{Y}_i = \hat{\beta}_0 + \hat{\beta}_1 X_i$

Properties (残差の性質)

1. $\sum_{i=1}^{n} e_i = \sum_{i=1}^{n} (\underbrace{Y_i - \bar{Y}}_{\to 0} - \hat{\beta}_1 \underbrace{(X_i - \bar{X})}_{\to 0})$
2. $\sum_{i=1}^{n} \hat{Y}_i = \sum_{i=1}^{n} Y_i$ (ここで Y_i は観測値データと同じ)
3. LS regression line passes $(\bar{X}, \bar{Y})$
4. $\sum_{i=1}^{n} X_i e_i = 0$

ベクトル空間で考えると、正射影として

$$X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \quad Y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \quad e = \begin{pmatrix} e_1 \\ \vdots \\ e_n \end{pmatrix}$$



5. $\sum_{i=1}^{n} \hat{y}_i e_i = 0$

この辺りを復習しておく。

9/19.~

- Simple Linear Regression.

$$Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i \quad (i = 1 \sim n)$$

Gauss Markov Conditions
$$\begin{cases} E[\varepsilon_i] = 0 & (i = 1 \sim n) \\ V[\varepsilon_i] = \sigma^2 & \\ E[\varepsilon_i \varepsilon_j] = 0 & (i \neq j) \end{cases}$$

- Least Square Estimators (LSE)

$$\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x}, \quad \hat{\beta}_1 = \frac{S_{xy}}{S_{xx}}$$

$$E[\hat{\beta}_0] = \beta_0, \quad E[\hat{\beta}_1] = \beta_1$$

$$Var[\hat{\beta}_0] = \frac{\sigma^2}{S_{xx}}, \quad Var[\hat{\beta}_1] = \sigma^2\left(\frac{1}{n} + \frac{\bar{x}^2}{S_{xx}}\right)$$

- Residuals（残差）

$$e_i = Y_i - \hat{Y}_i = Y_i - \hat{\beta}_0 - \hat{\beta}_1 x_i \quad (i = 1 \sim n)$$

いまいち失因，内容

- Estimation of σ^2.（σ^2の推定）

、$\sigma^2 = E[\varepsilon_i^2]$ であるが ε_iは直接に観測されない。

、$SS_{Res} = \sum_{i=1}^{n} e_i^2 \left(= \sum_{i=1}^{n} (y_i - \hat{y}_i)^2\right)$
（またはSS_{Error}）

No.
Date 9/19

$E[SS_{Res}] = (n-2)\sigma^2$ となる。

$$\hat{\sigma}^2 = \frac{1}{n-2} SS_{Res} = MS_{Res} \quad \text{(Residual Mean Square)}$$

$$SS_{Res} = \sum_{i=1}^{n} (y_i - \bar{y} + \hat{\beta}_1 \bar{x} - \hat{\beta}_1 x_i)^2$$
$$= \sum_{i=1}^{n} \{ (y_i - \bar{y}) - \hat{\beta}_1 (x_i - \bar{x}) \}^2$$
$$= \sum_{i=1}^{n} \{ (y_i - \bar{y})^2 - 2\hat{\beta}_1 (x_i - \bar{x})(y_i - \bar{y}) + \hat{\beta}_1^2 (x_i - \bar{x})^2 \}$$
$$= S_{yy} - \frac{2 S_{xy}^2}{S_{xx}} + \left(\frac{S_{xy}}{S_{xx}} \right)^2 \cdot S_{xx}$$
$$= \underbrace{S_{yy}}_{SST} - \frac{S_{xy}^2}{S_{xx}}$$ となる。

• Analysis of Variance …… (分散分析)

$$SST = \sum_{i=1}^{n} (y_i - \bar{y})^2 = \sum_{i=1}^{n} ((y_i - \hat{y}_i) + (\hat{y}_i - \bar{y}))^2$$
$$= \sum_{i=1}^{n} \{ \underbrace{(y_i - \hat{y}_i)^2}_{SS_{Res}} + \underbrace{2(y_i - \hat{y}_i)(\hat{y}_i - \bar{y})}_{e_i \cdot (\hat{y}_i - \bar{y}) = 0} + \underbrace{(\hat{y}_i - \bar{y})^2}_{SSR} \}$$

ここは先週に登場した数式で説明がつく。

$$= \underbrace{SS_{Res}}_{\text{モデルにおいて説明できないところ}} + \underbrace{SS_R}_{\text{モデルが説明できるところ}}$$

No.
Date 9/19.

全分散の公式

- $V[Y] = E[V[Y|X]] + V[E[X|Y]]$

先ほどの式に戻って $SS_{RES} = SS_T - \hat{\beta}_1 \cdot S_{xy} = SS_T - \underbrace{\frac{S_{xy}^2}{S_{xx}}}_{SS_R}$

これが回帰モデルによる変動に一致することはすぐに分かる

- Degrees of Freedom (自由度) → 先ほど $E[SS_{RES}] = \sigma^2(n-2)$ から、

$$\underbrace{\sum_{i=1}^{n}(Y_i-\bar{Y})^2}_{\text{自由度}(n-1)} = \underbrace{\sum_{i=1}^{n}(Y_i-\hat{Y}_i)^2}_{\text{自由度}(n-2)} + \underbrace{\sum_{i=1}^{n}(\hat{Y}_i-\bar{Y})^2}_{\text{自由度}1}$$

- Mean Squares.

$$\begin{cases} MS_R = \dfrac{SS_R}{1} \\ MS_{RES} = \dfrac{SS_{RES}}{n-2} \end{cases}$$

- ANOVA Table....

	SS	df	MS	検定統計量 F
Source of Variation	Sum of Square	Degree of Freedom	Mean Square	
Regression (回帰)	SS_R	1	MS_R	$\frac{MS_R}{MS_{RES}}$
Residual (残差)	SS_{RES}	$n-2$	MS_{RES}	
計	SS_T	$n-1$		

Chryv culture

No.

Date 9/19

• H_0: (帰無仮説) vs H_1 (対立仮説) に対して

: $\beta_1 = 0$ vs $\beta_1 \neq 0$

$$E[MS_{Res}] = \sigma^2, \quad E[MS_R] = \begin{cases} \sigma^2 & \text{if } (\beta_1 = 0) \\ \sigma^2 + \beta_1^2 S_{xx} & \text{if } (\beta_1 \neq 0) \end{cases}$$

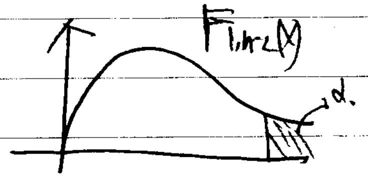
$$F_0 = \frac{MS_R}{MS_{Res}}$$ (もし $\beta_1 = 0$ であれば F_0 は1になるが $\beta_1 \neq 0$ であれば F_0 はそれより大きくなる。と考えられる。)

$$= t_0^2$$

$$t_0 = \frac{\hat{\beta}_1 - 0}{\sqrt{\left(\frac{\hat{\sigma}^2}{S_{xx}}\right)}} \qquad \left(\hat{\sigma}^2 = \frac{SS_{Res}}{n-2}\right)$$

• Normal Assumption + (Gauss-Markov Condition)

$\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n \sim N(0, \sigma^2)$



$F_0 \sim F_{1,n-2}$ under H_0

Reject $H_0: \beta_0 = 0$ when $F_0 > F_{1,n-2}(1-\alpha)$

• Confidence Interval (信頼区間)

$\hat{\beta}_1$ …… $\frac{\hat{\beta}_1 - \beta_1}{\sqrt{\hat{\sigma}^2 / S_{xx}}} \sim t_{n-2}$ に従うので

No.
Date 9 19

~~β_1の100(1-α)%~~

$\hat{\beta}_1 \pm t_{n-2}\left(\frac{\alpha}{2}\right) \cdot \sqrt{\frac{\hat{\sigma}^2}{S_{xx}}}$ を信頼区間と呼ぶ。

次に σ^2の信頼区間　$\hat{\sigma}^2 = MS_{RES} \sim \sigma^2 \cdot \chi_{n-2} / (n-2)$

$$1-\alpha = Pr\left(\chi^2_{n-2,\frac{\alpha}{2}} < \frac{\hat{\sigma}^2(n-2)}{\sigma^2} < \chi_{n-2,1-\frac{\alpha}{2}}\right)$$

$$= Pr\left(\frac{\hat{\sigma}^2(n-2)}{\chi^2_{n-2,1-\frac{\alpha}{2}}} < \sigma^2 < \frac{\hat{\sigma}^2(n-2)}{\chi^2_{n-2,\frac{\alpha}{2}}}\right)$$

これが信頼区間となる。

- Coefficient of Determination : R^2 (決定係数)

$$R^2 = \frac{SSR}{SST} = \frac{SSR}{SSR + SS_{RES}} \quad (0\sim1\text{の値をとる。})$$

$$= 1 - \frac{SS_{RES}}{SSR + SS_{RES}}$$

- 次に $\beta_0 + \beta_1 x_0$ の信頼区間について

$$\frac{(\hat{\beta}_0 + \hat{\beta}_1 x_0) - (\beta_0 + \beta_1 x_0)}{\sqrt{\hat{\sigma}^2 \cdot \left(\frac{1}{n} + \frac{(x_0 - \bar{x})^2}{S_{xx}}\right)}} \sim t_{n-2}$$

$$Var(\hat{\beta}_0 + \hat{\beta}_1 x_0) = Var[\bar{Y} + \hat{\beta}_1(x_0 - \bar{x})]$$

$$= Var[\bar{Y}] + (x_0 - \bar{x})^2 Var(\hat{\beta}_1) + 0$$

No.

Date :

予測区間…

- Prediction Interval for a new observation

$Y_0 = \beta_0 + \beta_1 x_0 + \varepsilon_0 \quad \varepsilon_0 \sim N(0, \sigma^2)$

$\dfrac{\hat\beta_0 + \hat\beta_1 x_0 - Y_0}{\sqrt{\hat\sigma^2\left(1 + \frac{1}{n} + \frac{(x_0 - \bar x)^2}{S_{xx}}\right)}} \sim t_{n-2}$ となる。

$\mathrm{Var}[\hat\beta_0 + \hat\beta_1 x_0 - \beta_0 - \beta_1 x_0 - \varepsilon_0]$

$= \mathrm{Var}[\bar Y] + (x_0 - \bar x)^2 \mathrm{Var}(\hat\beta_1) + \mathrm{Var}(\varepsilon_0)$

この部分だけ増加する

- Regression through the origin (原点を通る回帰モデル)

(つまり $\beta_0 = 0$ のモデル)

「 $Y = \beta_1 X + \varepsilon$ 」

LSE for β_1 … $\displaystyle\min_{\beta_1} \sum_{i=1}^{n} (Y_i - \beta_1 x_i)^2$

(最小二乗推定量)

$\hat\beta_1 = \dfrac{\sum_{i=1}^{n} X_i Y_i}{\sum_{i=1}^{n} X_i^2} \qquad E[\hat\beta_1] = \beta_1 \qquad \mathrm{Var}(\hat\beta_1) = \dfrac{\sigma^2}{\sum_{i=1}^{n} x_i^2} \leq \dfrac{\sigma^2}{S_{xx}}$

$\hat\sigma^2 = MS_{Res} = \dfrac{\sum_{i=1}^{n} (Y_i - \hat\beta_1 x_i)^2}{n-1}$ → 10/14, X-1 [illegible]

$\hat Y_i = \hat\beta_1 x_i$ Intercept β_0 がない model

Chryu culture

No.
Date

$$R^2 = \frac{\sum \hat{Y_i}^2}{\sum Y_i^2}$$

- Random Design … (X is random)

「$Y = \beta_0 + \beta_1 X + \varepsilon$」

$E[Y|X] = \beta_0 + \beta_1 x$, $E[\varepsilon|X] = 0$

$(X, Y)^t \sim$ 2次元正規分布… $N\left(\begin{pmatrix}\mu_1\\ \mu_2\end{pmatrix}, \begin{pmatrix}\sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2\end{pmatrix}\right)$

$Y|X=x \sim N\left(\mu_2 + \rho\frac{\sigma_2}{\sigma_1}(x-\mu_1), \sigma_2^2(1-\rho^2)\right)$

(公式)

$E[Y|X] = E(Y) + \mathrm{cov}[Y, X]\, \Sigma_X^{-1}(x - E(X)) =$

$V(Y|X) = V(Y) - \mathrm{cov}(Y,X) \cdot \Sigma_X^{-1} \mathrm{cov}(Y,X)$

MLE = LSE.

Sample correlation coefficient.

$$r = \frac{S_{xy}}{(S_{xx} \cdot S_{yy})^{\frac{1}{2}}} = \frac{S_{xy}}{(S_{xx} SS_T)^{\frac{1}{2}}}$$

$H_0: \rho = 0$. $H_1: \rho \neq 0$

$$\hat{\beta_1} = \left(\frac{SS_T}{S_{xx}}\right)^{\frac{1}{2}} \cdot r$$

No. 9
Date : :

(A)
Chapter 2. Homework

…… ~~4~~, ~~7~~, 10. 12, 17, 25, 26, 27, 32, 33.

10/14までに提出。

Chapter 3. Multiple Linear Regression (9/26~

$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots \beta_k X_k + \varepsilon$

$E[\varepsilon] = 0$

- $X_1, X_2, \ldots X_k$: Covariates
- Y : Response Variable
- $\beta_0, \beta_1 \ldots \beta_k$: Unknown Parameters

β_1… $X_2 \sim X_k$を固定した状況下で X_1の値が1増加したときに変動する量を表している。

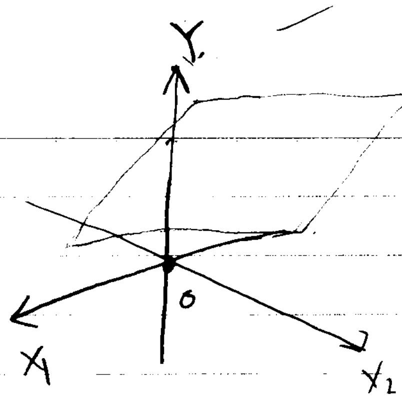
例、多項式回帰（これも重回帰と言える。）

X…

$Y = \beta_0 + \beta_1 X + \beta_2 X^2 + \cdots \beta_k X^k + \varepsilon$

No
Date 9/26

Interactions.



(平面)

- 交互作用のない2変量

$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \varepsilon$

- 交互作用を考慮したmodel

$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_1 X_2 + \varepsilon$

$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_1 X_2 + \beta_4 X_1^2 + \beta_5 X_2^2 + \varepsilon$

- データ (Data)

$(Y_i, X_{i1}, X_{i2}, \ldots X_{ik}) \quad i = 1 \sim n$

model … $Y_i = \beta_0 + \beta_1 X_{i1} + \ldots \beta_k X_{ik} + \varepsilon_i \quad (i = 1 \sim n)$

Gauss-Markov-Conditions $E[\varepsilon_i] = 0, \quad V[\varepsilon_i] = \sigma^2 \quad (i = 1 \sim n)$

$Cov[\varepsilon_i, \varepsilon_j] = 0 \quad (i \neq j)$

- 最小二乗法による推定

$S(\beta_0, \beta_1, \ldots \beta_k) = \sum_{i=1}^{n} (Y_i - \beta_0 - \beta_1 X_{i1} - \cdots - \beta_k X_{ik})^2$

$\underset{\begin{pmatrix}\beta_0 \\ \vdots \\ \beta_k\end{pmatrix}}{\arg\min}\ S(\beta_0, \beta_1, \ldots \beta_k)$ を求める

No.
Date 9/26

$$\begin{pmatrix} \frac{\partial}{\partial \beta_0} S(\beta_0, \cdots, \beta_k) \\ \frac{\partial}{\partial \beta_1} S(\beta_0, \cdots, \beta_k) \\ \vdots \\ \frac{\partial}{\partial \beta_k} S(\beta_0, \cdots, \beta_k) \end{pmatrix} = \frac{\partial S}{\partial \boldsymbol{\beta}} = 0$$ と計算して内積良い

$$\frac{\partial S}{\partial \beta_j} = \sum_{i=1}^{n} 2(Y_i - \beta_0 - \beta_1 X_{i1} \cdots - \beta_k X_{ik}) \cdot (-X_{ij}) = 0 \quad (j = 1 \sim k) \quad (j \geq 1)$$

しかしこのまま計算すると少しめんどうなので細かく表す

モデルの行列表示を与える

$$\begin{pmatrix} Y_1 \\ \vdots \\ Y_n \end{pmatrix} = \underbrace{\begin{pmatrix} 1 & X_{1,1} & \cdots & X_{1,k} \\ 1 & X_{2,1} & \cdots & X_{2,k} \\ & \vdots & & \\ 1 & X_{n,1} & \cdots & X_{n,k} \end{pmatrix}}_{X} \underbrace{\begin{pmatrix} \beta_0 \\ \vdots \\ \beta_k \end{pmatrix}}_{\boldsymbol{\beta}} + \varepsilon$$

これを利用して $(Y - X\beta)^t (Y - X\beta) = S(\beta)$

となるので $\frac{\partial S}{\partial \boldsymbol{\beta}} = \frac{\partial}{\partial \boldsymbol{\beta}} (Y^t Y - \beta^t X^t Y - Y^t X\beta + \beta^t X^t X\beta)$

Chey culture

$$\frac{\partial}{\partial \beta}\left({}^t\beta\, {}^tX Y\right) = \begin{pmatrix} \frac{\partial}{\partial \beta_1}\, {}^t\beta\, {}^tX Y \\ \vdots \\ \frac{\partial}{\partial \beta_n}\, {}^t\beta\, {}^tX Y \end{pmatrix} = \begin{pmatrix} {}^tXY\text{の1列目} \\ \vdots \\ {}^tXY\text{のn列目} \end{pmatrix} = {}^tXY$$

$$\frac{\partial}{\partial \beta}\left({}^t\beta A \beta\right) = \frac{\partial}{\partial \beta} \sum_{i=1}^{n} \sum_{j=1}^{n} \beta_i \beta_j \cdot a_{ij}$$ であるから

$=$

$$\frac{\partial S}{\partial \beta} = -2\,{}^tXY + 2\,{}^tX X\beta = 0$$

$$\therefore \beta = ({}^tXX)^{-1}\, {}^tXY \Rightarrow \text{これは} \hat{\beta} \text{となる}$$

また残差に関して $e = Y - X\hat{\beta} = Y - X({}^tXX)^{-1}\,{}^tXY = \begin{pmatrix} e_1 \\ \vdots \\ e_n \end{pmatrix}$

回帰に関して $\hat{y} = X\hat{\beta} = \underbrace{X({}^tXX)^{-1}\,{}^tX}_{\text{これはHと書かれる。}}Y$

No. 9/26
Date

$$X = (1\ x_1 \cdots x_k)$$

$$1_n = \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} \updownarrow n \qquad x_j = \begin{pmatrix} x_{1,j} \\ \vdots \\ x_{n,j} \end{pmatrix}$$ $\boxed{\text{各データ1〜n の } x_j \text{ の値}}$ と見なせる

$$X\beta = (1_n\ x_1 \cdots x_k)\begin{pmatrix} \beta_0 \\ \vdots \\ \beta_k \end{pmatrix}$$

$= \beta_0 1 + \beta_1 x_1 + \cdots \beta_k x_k$ である。

$\hat{y} = X\hat{\beta} = H\underline{y}$ 観測値

$$H = X(X^tX)^{-1}X^t$$

$$HH = X(X^tX)^{-1}\underbrace{X^tX(X^tX)^{-1}}_{I}X^t = X(X^tX)^{-1}X^t = H$$

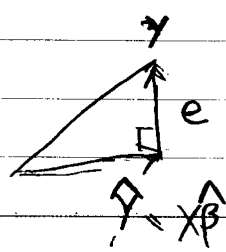
$\therefore H^2 = H$ である。

$$e = y - \hat{y} = (I-H)y$$

・$X^te = X^t(I - X(X^tX)^{-1}X^t)y$

$= (X^t - X^t)y = Oy = 0$ となる

図



正射影と思う

$$X^te = 0$$

No.
Date 9/26

$$\hat{\beta} = (X^t X)^{-1} X^t Y$$

$$E[\hat{\beta}] = \begin{pmatrix} E[\hat{\beta}_0] \\ \vdots \\ E[\hat{\beta}_k] \end{pmatrix}$$ の要素 $= E[(X^t X)^{-1} X^t Y]$

$$= (X^t X)^{-1} X^t E[Y] = (X^t X)^{-1} X^t E[X\beta + \varepsilon] = (X^t X)^{-1} X^t X \beta = \beta$$ とする

$$V[\hat{\beta}] = E[(\hat{\beta} - \underset{E[\hat{\beta}]}{\underset{\parallel}{\beta}})(\hat{\beta} - \beta)^t] =$$

$(\mathrm{COV}[\hat{\beta}])$

分散、共分散行列

$\underline{k+1} \times \underline{k+1}$ の行列

$$\hat{\beta} - \beta = (X^t X)^{-1} X^t Y - \beta = (X^t X)^{-1} X^t (Y - X\beta)$$

$$(\hat{\beta} - \beta)(\hat{\beta} - \beta)^t = (X^t X)^{-1} X^t (Y - X\beta)(Y - X\beta)^t ((X^t X)^{-1} X^t)^t$$

$$\therefore E[(\hat{\beta} - \beta)(\hat{\beta} - \beta)^t] = (X^t X)^{-1} X^t \underbrace{E[(Y - X\beta)(Y - X\beta)^t]}_{\substack{\parallel \\ \sigma^2 I}} ((X^t X)^{-1} X^t)^t$$

$$= \sigma^2 \cdot (X^t X)^{-1} X^t ((X^t X)^{-1} X^t)^t$$

$$= \sigma^2 \cdot (X^t X)^{-1} X^t X ((X^t X)^{-1})^t$$

$$= \sigma^2 \cdot ((X^t X)^{-1})^t = \sigma^2 \cdot (X^t X)^{-1}$$

No.
Date 9.26.

~~SS_{Error}~~ (3節)

$$\shortparallel$$

• $SS_{Residual}$ …… $\sum_{i=1}^{n} e_i^2 = e^t e$ である。

$$\left((I-H)y\right)^t (I-H)y = y^t (I-H)^t (I-H)y$$

$$= y^t\left(I - \cancel{H^t I} - IH + H^t H\right) \qquad H^t = H,\ H^2 = H$$

$$= y^t (I-H) y$$

$$= (X\beta+\varepsilon)^t (I-H)(X\beta+\varepsilon)$$

（Xβの部分に適用することで消去）

$$= \varepsilon^t (I-H)\varepsilon$$

$$E\left[\varepsilon^t \underbrace{(I-H)}_{M}\varepsilon\right] = E\left[\sum_{i=1}^{n}\sum_{j=1}^{n} M_{ij}\,\varepsilon_i \varepsilon_j\right] = \sigma^2 \cdot tr(M)$$

$$= \sigma^2 tr(I-H)$$

$$= \sigma^2(\underbrace{tr I}_{n} - tr H)$$

$n \times (k+1)$ $\quad (k+1)\times n$

$$tr H = tr\left[\underbrace{X}_{A}\underbrace{(X^t X)^{-1} X^t}_{B}\right] = tr(AB) = tr(BA)$$

$$= tr(I_{k+1}) = k+1.$$

$$= \sigma^2 (n - \underbrace{p}_{k+1})$$

• SS_{Total} ‥ $\sum_{i=1}^{n}(y_i - \bar{y})^2 = \left(y - \begin{pmatrix}1\\ \vdots\\ 1\end{pmatrix}\bar{y}\right)^t \left(y - \begin{pmatrix}1\\ \vdots\\ 1\end{pmatrix}\bar{y}\right)$

スカラー値 ←

$$= (y - \bar{y}\mathbf{1})^t (y - \mathbf{1}\bar{y})$$

$$= (\hat{y} + e - \bar{y}\mathbf{1})^t (\hat{y} + e - \bar{y}\mathbf{1})$$

ベクトルで考えると直交 ことから

$$= \underbrace{(\hat{y} - \bar{y}\mathbf{1})^t (\hat{y} - \bar{y}\mathbf{1})}_{SS_{Reg}} + \underbrace{(\hat{y} - \bar{y}\mathbf{1})^t e + e^t (\hat{y} - \bar{y}\mathbf{1})}_{0} + \underbrace{e^t e}_{SS_E}$$

Chuyu culture

・R^2 … $\frac{SS_{R}}{SS_T} = 1 - \frac{SS_E}{SS_T}$ $(0 \leqq R^2 \leqq 1)$」

これ $\begin{pmatrix} Y \\ X_1 \\ \vdots \\ X_k \end{pmatrix}$ の分布（ここでは $X_1 \dots X_k$ は確率変数と見なす）

$$\sim N\left(\underbrace{\begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}}_{\mu}, \underbrace{\begin{pmatrix} \sigma_1^2 & \Sigma_{1,2} \\ \Sigma_{2,1} & \Sigma_{2,2} \end{pmatrix}}_{\Sigma}\right)$$

同時確率密度関数

$$\frac{1}{\sqrt{2\pi}^{k+1} \cdot |\Sigma|^{\frac{1}{2}}} \exp\left(-\frac{1}{2}\left((y\ x_1 \cdots x_k) - \underbrace{(\mu_1\ \mu_2^t)}_{\mu^t}\right) \Sigma^{-1} \left\{\begin{pmatrix} y \\ x_1 \\ \vdots \\ x_k \end{pmatrix} - \underbrace{\begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}}_{\mu}\right\}\right)$$

これに $Y \mid X_1, X_2 \dots X_k$ の条件付分布について考える

、平均：$\mu_1 + \Sigma_{12}\Sigma_{22}^{-1}(X - \mu_2)$

、分散：$\sigma_1^2 - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}$

の正規分布 $N(\underbrace{\mu_1 + \Sigma_{12}\Sigma_{22}^{-1}(X - \mu_2)}_{= X\beta}, \underbrace{\sigma_1^2 - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}}_{= \sigma^2})$

、MLEを求めると＝LSEに一致

（要証明）

No.
Date 10/14(金) (前までの復習を兼ねて)

・重回帰分析 $\underset{n\times 1}{Y} = \underset{n\times p}{X}\underset{p\times 1}{\beta}+\varepsilon.$

・LSE : $\hat\beta = (X^tX)^{-1}X^tY$

$E[\hat\beta]=\beta \qquad V[\hat\beta]=(X^tX)^{-1}\sigma^2$

・$\hat{Y}\beta = \underbrace{X(X^tX)^{-1}X^t}_{H}Y$

$E[\hat Y] = X\beta = E[Y]$

$V[\hat Y] = X\,\mathrm{cov}[\hat\beta]\,X^t = \sigma^2 X(X^tX)^{-1}X^t = \sigma^2 H$

・$e = Y-\hat Y = (I-H)Y$

$V[e]=\sigma^2(I-H)$

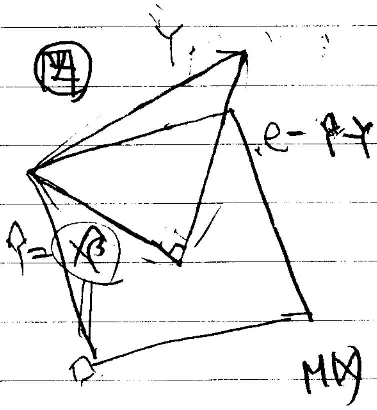
$\left(V[e_i]=\sigma^2(1-H_{ii})\right)$

$SS_{res}=e^te$ (残差平方和) $\left(\hat\sigma^2=\frac{1}{n-p}e^te\right)$

$\underbrace{\sum_{i=1}^n Y_i^2}_{\text{Total}} = \underbrace{\sum_{i=1}^n \hat Y_i^2}_{\text{predicted}} + \underbrace{\sum_{i=1}^n e_i^2}_{\text{residual}}$

$Y^tY = \hat Y^t\hat Y + e^te$

図



No. Date 10/14

Xの計画行列に $\begin{pmatrix}1\\ \vdots\\ 1\end{pmatrix}$ のvectorが入っている

- Intercept β_0 in the model. (β_0がモデルに入っているとき)

$\sum_{i=1}^{n} e_i = 0$ つまり $\begin{pmatrix}1\\ \vdots\\ 1\end{pmatrix}^t \begin{pmatrix}e_1\\ \vdots\\ e_n\end{pmatrix} = e_1 + \cdots + e_n = 0$

$$\underbrace{\sum_{i=1}^{n} e_i^2}_{SS_E} = \sum_{i=1}^{n} Y_i^2 - \sum_{i=1}^{n} \hat{Y}_i^2 = \underbrace{\sum_{i=1}^{n}(Y_i - \bar{Y})^2}_{SS_T} - \underbrace{\sum_{i=1}^{n}(\hat{Y}_i - \bar{Y})^2}_{SS_{reg}}$$

$$R^2(\text{決定係数}) = \frac{SS_{Residual}}{SS_T} = 1 - \frac{SS_{reg}}{SS_T}$$

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

$= \underbrace{\sum_{i=1}^{n} e_i \hat{Y}_i}_{\substack{= \\ 0 \\ (e^t \cdot \hat{Y})}} - \bar{Y}\underbrace{\sum_{i=1}^{n} e_i}_{\substack{= \\ 0 \\ (1_n^t e = 0)}} + \sum_{i=1}^{n}(\hat{Y}_i - \bar{Y})^2 = \underbrace{\sum_{i=1}^{n}(\hat{Y}_i - \bar{Y})^2}_{= SS_{reg}}$ である。

またこのことから $R^2 = \dfrac{\left(\sum_{i=1}^{n}(Y_i - \bar{Y})(\hat{Y}_i - \bar{Y})\right)^2}{\sum_{i=1}^{n}(Y_i - \bar{Y})^2 \sum_{i=1}^{n}(\hat{Y}_i - \bar{Y})^2}$ $= SS_{reg}$ と書けるので

∴ R = yと$\hat{Y}$の標本相関係数と言える。

(つまり Rが1に近ければ、$\hat{Y}$がYに近いので良いモデルと考えられる)

No.
Date 10/14

ANOVA Table :

Source	自由度	SS
回帰	$p-1$ (k)	$\hat{\beta}^t X^t Y - n\bar{Y}^2$
残差	$n-p$	$Y^tY - \hat{\beta}^t X^t Y$
Total	$n-1$	$Y^tY - n\bar{Y}^2$

$$= SS_{reg} = Y^t H Y - n\bar{Y}^2$$

$$Y^t\left(H - \frac{J}{n}\right)Y$$

- Intercept β_0 not in the model.

(β_0 がモデルに入っていないとき)

ANOVA Table :

why

Source	自由度	SS
回帰	k	$\hat{Y}^t\hat{Y} = \hat{\beta}^t X^t Y$
残差	$n-k$	$e^t e$

SS_T = Y^tY とかに
$\bar{Y}$ を 0 におきかえたもの

- Best Linear Unbiased Estimator (BLUE)

$\underset{q\times n}{A}\,\underset{n\times 1}{Y}$ が $\underset{q\times p}{L}\,\underset{p\times 1}{\beta}$ の BLUE (最良線形不偏推定量) で

あるとは $E[AY] = L\beta$ (for all β) が成り立ち

分散共分散行列の差 $V[CY] - V[AY]$ が

任意の $L\beta$ の不偏推定量 CY に対し

positive semi-definite になることをいう
(半正定値行列)

No.
Date 10/14

ガウス・マルコフ・モデル

$Y = X\beta + \varepsilon \quad (E[\varepsilon]=0)$

定義

$l^t\beta$ が推定可能 (estimable) であるとは $l \in \mathrm{span}[x_1, \cdots, x_n]$ と同値である

$(X = [x_1 \cdots x_n])$

※

Gauss-Markov Theorem

ガウスマルコフの条件下で $l^t\hat{\beta}$ は推定可能な関数 $l^t\beta$ の BLUE になれる。

証明

$l^t\beta$ 推定可能 $\Leftrightarrow \exists \underset{n\times 1}{c},\ l^t\beta = E[c^tY] = c^tX\beta$ (for all β)

($l^t \to$ Lk にても同じ) $\Leftrightarrow \exists c,\ l^t = c^tX$

ここで c^tY を任意の線形不偏推定量 ($l^t\beta$ の) とする。

$c^tX = l^t$ より $V[c^tY] - V[l^t\hat{\beta}] = \sigma^2 c^t I c - \sigma^2 l^t (X^tX)^{-1} l$

$= \sigma^2 c^t (I - X(X^tX)^{-1}X^t) c$

$= \sigma^2 c^t (I-H) c$

$= \mathrm{Var}[c^t e] \geq 0$

No.

Date 10/14

• Centered Model: $\underline{k+1=p}$

$$\begin{pmatrix} Y_1 \\ \vdots \\ Y_n \end{pmatrix} = \begin{pmatrix} 1 & X_{1,1} & \cdots & X_{1,k} \\ \vdots & \vdots & & \vdots \\ 1 & X_{n,1} & \cdots & X_{n,k} \end{pmatrix} \begin{pmatrix} \beta_0 \\ \vdots \\ \beta_k \end{pmatrix} + \varepsilon$$

$$= \beta_0 1_{n\times1} + \begin{pmatrix} x_{1,1} & \cdots & x_{1,k} \\ \vdots & & \vdots \\ x_{1,n} & & x_{n,k} \end{pmatrix} \beta_{(0)} + \varepsilon \qquad \beta_{(0)} = \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_k \end{pmatrix}$$

$$= \underbrace{\beta_0 1_{n\times1} + \underbrace{\begin{pmatrix} \sum_{j=1}^{k} \beta_j \overline{x_j} \\ \vdots \\ \sum_{j=1}^{k} \beta_j \overline{x_j} \end{pmatrix}}_{\sum_{j=1}^{k} \beta_j \overline{x_j} \cdot 1_{n\times1}}}_{\gamma_0 1_{n\times1}} + \underbrace{\begin{pmatrix} x_{1,1}-\overline{x_1} & \cdots & x_{1,k}-\overline{x_k} \\ \vdots & & \vdots \\ x_{n,1}-\overline{x_1} & & x_{n,k}-\overline{x_k} \end{pmatrix}}_{= Z} \beta_{(0)} + \varepsilon$$

$$= \gamma_0 1_{n\times1} + Z\beta_{(0)} + \varepsilon$$ と書ける

$$= \underbrace{(1 \ Z)}_{\text{新しい計画行列}} \begin{pmatrix} \gamma_0 \\ \beta_{(0)} \end{pmatrix} + \varepsilon$$ と書ける

最小二乗推定量 $$\begin{pmatrix} \hat{\gamma}_0 \\ \hat{\beta}_{(0)} \end{pmatrix} = \left\{ (1 \ Z)^t (1 \ Z) \right\}^{-1} (1 \ Z)^t Y$$

$$= \left\{ \begin{pmatrix} 1^t \\ Z^t \end{pmatrix} (1 \ Z) \right\}^{-1} (1 \ Z)^t Y$$

Chuyu culture

No.
Date 10/14

$$= \begin{pmatrix} 1^t 1 & 1^t Z \\ Z^t 1 & Z^t Z \end{pmatrix}^{-1} (1\ Z)^t Y \qquad \text{注}\ \ 1^t Z = 0 \ \ (Z^t 1)$$

$$= \begin{pmatrix} n & 0 \\ 0 & Z^t Z \end{pmatrix}^{-1} \begin{pmatrix} 1^t \\ Z^t \end{pmatrix} Y = \begin{pmatrix} n & 0 \\ 0 & Z^t Z \end{pmatrix}^{-1} \begin{pmatrix} n\bar{Y} \\ Z^t Y \end{pmatrix}$$

$$= \begin{pmatrix} n^{-1} & 0 \\ 0 & (Z^t Z)^{-1} \end{pmatrix} \begin{pmatrix} n\bar{Y} \\ Z^t Y \end{pmatrix} = \begin{pmatrix} \bar{Y} \\ (Z^t Z)^{-1} Z^t Y \end{pmatrix}$$ と書ける

$$\hat{Y} = (1\ Z) \begin{pmatrix} \hat{\beta}_0 \\ \hat{\beta}_{(1)} \end{pmatrix} = X\hat{\beta}$$ になるはずである
（元々は同じもの）

$$e = Y - \hat{Y}$$

$$e^t e = Y^t Y - n\bar{Y}^2 - Y^t Z (Z^t Z)^{-1} Z^t Y$$ となる

$$R^2 = \frac{Y^t Z (Z^t Z)^{-1} Z^t Y}{Y^t Y - n\bar{Y}^2} = \frac{(Y - \bar{Y}1)^t Z (Z^t Z)^{-1} Z^t Y}{(Y - \bar{Y}1)^t (Y - \bar{Y}1)}$$

= sample multiple correlation between Y and $X_1 \sim X_k$

$$V\begin{bmatrix} Y \\ X_1 \\ \vdots \\ X_k \end{bmatrix} = \begin{pmatrix} \sigma_Y^2 & \sigma_{YX} \\ \sigma_{YX} & \Sigma_X \end{pmatrix}$$

multiple-correlation between Y and $\begin{pmatrix} X_1 \\ \vdots \\ X_k \end{pmatrix}$ is

$$\frac{(\sigma_{YX}^t \Sigma_{XX}^{-1} \sigma_{YX})^{\frac{1}{2}}}{\sigma_Y}$$

10/14

（最後は高速に進んだので要確認）