

# 数理工学実験レポート

## 第 3 章（最小二乗法）

学籍番号 1029366161 中塚一瑛

2025 年 10 月 30 日

### 概要

## 目次

1	はじめに	2
2	最小二乗法とその評価方法	2
2.1	原理・方法	2
2.2	課題 1（重回帰）	3
2.3	課題 2（多項式回帰）	5
2.4	課題 3（分散の観測誤差：Cauchy）	7
2.5	課題 4（入力域の制約と設計）	9
3	重み付き最小二乗法	11
3.1	原理と方法	11
3.2	課題 5（2 次元出力：重み付き最小二乗法）	12
3.3	課題 6（2 次元出力・異分散 2 群に対する重み付き最小二乗）	15
4	逐次最小二乗法	18
5	アルゴリズム比較	18
6	結論	18
付録 A	使用コード一覧	18
付録 B	補足導出	19

## 1 はじめに

本実験第2回では、最小二乗法の基礎とその実装手法を学ぶことを目的とする。具体的には、観測データからのパラメータ推定、重み付き・逐次最小二乗法、データ分割・推定値の合成を通して、推定精度と計算効率の違いを実験的に比較する。

## 2 最小二乗法とその評価方法

### 2.1 原理・方法

観測モデルを

$$y_i = f(\theta, x_i) + w_i \quad (1)$$

とする。加法雑音  $w_i$  は平均 0、観測ごとに独立、同一分布であり、共分散は有限と仮定する。ここでは  $f(\theta, x) = \phi(x)\theta$  としてパラメータに対して線形とし、最小二乗問題

$$\min_{\theta} \sum_{i=1}^N \|y_i - \phi(x_i)\theta\|^2 \quad (2)$$

を解く。解は

$$\hat{\theta}_N = \left( \sum_{i=1}^N \phi_i^\top \phi_i \right)^{-1} \sum_{i=1}^N \phi_i^\top y_i, \quad \phi_i := \phi(x_i). \quad (3)$$

また雑音分散が未知の場合の推定量および推定誤差共分散は

$$\hat{\sigma}^2 = \frac{1}{N-n} \sum_{i=1}^N \|y_i - \phi_i \hat{\theta}_N\|^2, \quad (4)$$

$$\widehat{\text{Cov}}(\hat{\theta}_N) = \hat{\sigma}^2 \left( \sum_{i=1}^N \phi_i^\top \phi_i \right)^{-1} \quad (5)$$

で与えられる。

当てはまりの評価には決定係数

$$C = \frac{\sum_{i=1}^N \|\phi_i \hat{\theta}_N - \bar{y}\|^2}{\sum_{i=1}^N \|y_i - \bar{y}\|^2}, \quad \bar{y} = \frac{1}{N} \sum_{i=1}^N y_i \quad (6)$$

を用いる。以上は資料 3.2 の線形最小二乗および評価に対応する。

本実験では  $\phi(x)$  として設計した特徴量から行列

$$X = \begin{bmatrix} \phi(x_1)^\top \\ \vdots \\ \phi(x_N)^\top \end{bmatrix} \quad (7)$$

を構成し、観測データを

$$y = [y_1; \dots; y_N] \quad (8)$$

として用いる．パラメータ推定量は

$$\hat{\theta} = (X^\top X)^{-1} X^\top y \quad (9)$$

で与えられる．

残差は

$$r = y - X\hat{\theta} \quad (10)$$

とし，雑音分散推定量は

$$\hat{\sigma}^2 = \frac{\|r\|^2}{N-p} \quad (11)$$

( $p$  は列数) とする．

パラメータ推定の誤差共分散は

$$\hat{\sigma}^2 (X^\top X)^{-1} \quad (12)$$

で与えられる．

決定係数は

$$C = \frac{\|X\hat{\theta} - \bar{y}\mathbf{1}\|^2}{\|y - \bar{y}\mathbf{1}\|^2} \quad (13)$$

と定義する．

以下の R 関数が上記推定を実装している．

```
1 regression_simple <- function(x, y){
2   theta_hat <- solve(t(x) %*% x) %*% t(x) %*% y
3   sigma2_hat <- as.numeric(t(y - x %*% theta_hat) %*%
4     (y - x %*% theta_hat) /
5     (nrow(x) - ncol(x)))
6   err_cov_mat <- sigma2_hat * solve(t(x) %*% x)
7   det_coef <- sum((x %*% theta_hat - mean(y))^2) /
8     sum((y - mean(y))^2)
9   list(theta_hat=theta_hat, err_cov_mat=err_cov_mat,
10     sigma2_hat=sigma2_hat, det_coef=det_coef)
11 }
```

## 2.2 課題 1 (重回帰)

■モデル  $y_i = x_i^\top \theta + w_i$  ( $i = 1, \dots, N$ ).  $x_i \in \mathbb{R}^2$ .  $w_i$  は独立，同分散  $\sigma^2$ ，平均 0.  $X = [x_1^\top; \dots; x_N^\top] \in \mathbb{R}^{N \times 2}$ ,  $y = [y_1; \dots, y_N] \in \mathbb{R}^N$ .

■推定量 最小二乗推定量

$$\hat{\theta}_N = (X^\top X)^{-1} X^\top y.$$

残差  $r = y - X\hat{\theta}_N$  により

$$\hat{\sigma}^2 = \frac{\|r\|_2^2}{N-p}, \quad p=2, \quad \widehat{\text{Cov}}(\hat{\theta}_N) = \hat{\sigma}^2 (X^\top X)^{-1}.$$

決定係数

$$R^2 = \frac{\|X\hat{\theta}_N - \bar{y}\mathbf{1}\|_2^2}{\|y - \bar{y}\mathbf{1}\|_2^2}, \quad \bar{y} = \frac{1}{N} \sum_{i=1}^N y_i.$$

■収束確認  $N \in \{2, 4, 8, \dots, 2^{13} = 8192\}$  で  $\hat{\theta}_N$  を計算し,  $N$  を横軸とする片対数図で各成分を同一図に描く.

■結果 (全データ  $N = 10000$ )

$$\hat{\theta}_N = \begin{bmatrix} 1.506551 \\ 1.997696 \end{bmatrix}, \quad \widehat{\text{Cov}}(\hat{\theta}_N) = \begin{bmatrix} 9.866491 \times 10^{-5} & -4.081657 \times 10^{-7} \\ -4.081657 \times 10^{-7} & 1.005248 \times 10^{-4} \end{bmatrix}.$$

決定係数

$$R^2 = 0.8629734.$$

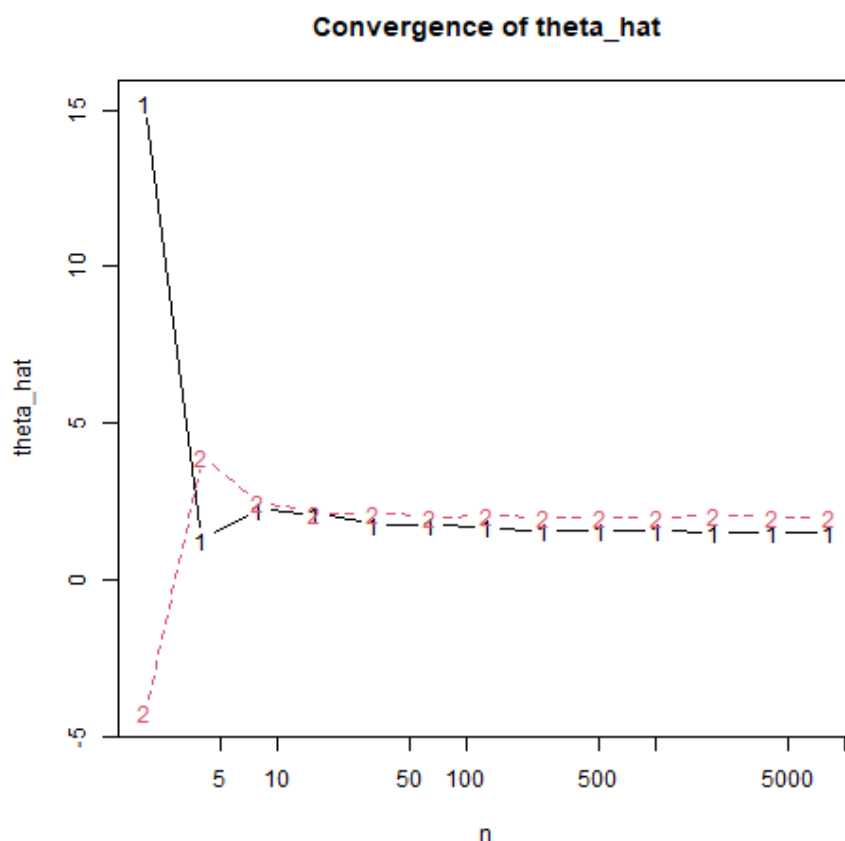


図 1:  $\hat{\theta}_N$  の収束 (横軸  $N = 2, 4, \dots, 8192$  の片対数)

#### ■Rコード

```
1 # データ読み込み
2 data <- read.csv("datas/mmse_kadai1.csv", header=FALSE,
3                 col.names = c("x1", "x2", "y"))
4 x <- as.matrix(data[, c("x1", "x2")])
5 y <- as.matrix(data[, "y"])
6
7 # 最小二乗 (原理・方法で用いる関数)
8 regression_simple <- function(x, y){
9   theta_hat <- solve(t(x) %*% x) %*% t(x) %*% y
10  sigma2_hat <- as.numeric(t(y - x %*% theta_hat) %*%
11                          (y - x %*% theta_hat) / (nrow(x) - ncol(x)))
12  err_cov_mat <- sigma2_hat * solve(t(x) %*% x)
```

```

13  det_coef <- sum((x %*% theta_hat - mean(y))^2) /
14      sum((y - mean(y))^2)
15  list(theta_hat=theta_hat, err_cov_mat=err_cov_mat,
16      sigma2_hat=sigma2_hat, det_coef=det_coef)
17 }
18
19 # 部分データでの実験
20 expl <- function (x, y, n){
21   x <- x[1:n, , drop=FALSE]
22   y <- y[1:n, , drop=FALSE]
23   result <- regression_simple(x, y)
24   list(theta_hat = result$theta_hat,
25       err_cov_mat = result$err_cov_mat,
26       det_coef = result$det_coef)
27 }
28
29 # 収束図の作成
30 plot_expl <- function(x, y, out="graphs/task1.png"){
31   ns <- 2^(1:13) # 2,...,8192
32   theta_hats <- matrix(NA_real_, nrow=length(ns), ncol=ncol(x))
33   for(i in seq_along(ns)){
34     theta_hats[i, ] <- as.vector(expl(x, y, ns[i])$theta_hat)
35   }
36   dir.create(dirname(out), showWarnings=FALSE, recursive=TRUE)
37   png(out, width=960, height=600, res=120)
38   matplot(ns, theta_hats, type="b", log="x",
39       xlab="N", ylab=expression(hat(theta)),
40       main="Convergence of OLS estimates")
41   legend("bottomright",
42       legend=paste0("theta[", 1:ncol(x), "]"),
43       lty=1:ncol(x), pch=1:ncol(x))
44   dev.off()
45 }
46
47 # 実行例 (全データ)
48 N_max <- nrow(x)
49 th <- expl(x, y, N_max)$theta_hat
50 Vhat <- expl(x, y, N_max)$err_cov_mat
51 R2 <- expl(x, y, N_max)$det_coef
52 print(th); print(Vhat); print(R2)
53
54 # 収束図
55 plot_expl(x, y)

```

## ■考察

### 2.3 課題2 (多項式回帰)

■モデル  $y_i = \varphi(x_i)\theta + w_i$ ,  $\varphi(x) = [1 \ x \ x^2 \ x^3]$ ,  $i = 1, \dots, N$ . 雑音  $w_i$  は独立, 同分散  $\sigma^2$ , 平均 0.  $X = [\varphi(x_1); \dots; \varphi(x_N)] \in \mathbb{R}^{N \times 4}$ .

■推定量 推定量・共分散・決定係数は課題1と同じ形式 ( $p = 4$ ) で計算する.

■収束確認  $N \in \{4, 8, 16, \dots, 8192\}$  で  $\hat{\theta}_N$  を計算し,  $N$  を横軸とする片対数図で各成分を同一図に描画する.

■結果 (全データ  $N = 10000$ )

$$\hat{\theta}_N = \begin{bmatrix} -0.50902942 \\ 1.97586067 \\ 0.19774405 \\ -0.09866691 \end{bmatrix}, \quad \widehat{\text{Cov}}(\hat{\theta}_N) = \begin{bmatrix} 2.023442 \times 10^{-3} & -1.186497 \times 10^{-5} & -1.348312 \times 10^{-4} & 3.971165 \times 10^{-7} \\ -1.186497 \times 10^{-5} & 6.753015 \times 10^{-4} & -1.898545 \times 10^{-7} & -3.794719 \times 10^{-5} \\ -1.348312 \times 10^{-4} & -1.898545 \times 10^{-7} & 1.603910 \times 10^{-5} & 6.351820 \times 10^{-8} \\ 3.971165 \times 10^{-7} & -3.794719 \times 10^{-5} & 6.351820 \times 10^{-8} & 2.528719 \times 10^{-6} \end{bmatrix}$$

標準誤差 (対角の平方根):  $\text{SE}(\hat{\theta}_0) \approx 0.0450$ ,  $\text{SE}(\hat{\theta}_1) \approx 0.0260$ ,  $\text{SE}(\hat{\theta}_2) \approx 0.00400$ ,  $\text{SE}(\hat{\theta}_3) \approx 0.00159$ .

決定係数:

$$R^2 = 0.461855.$$

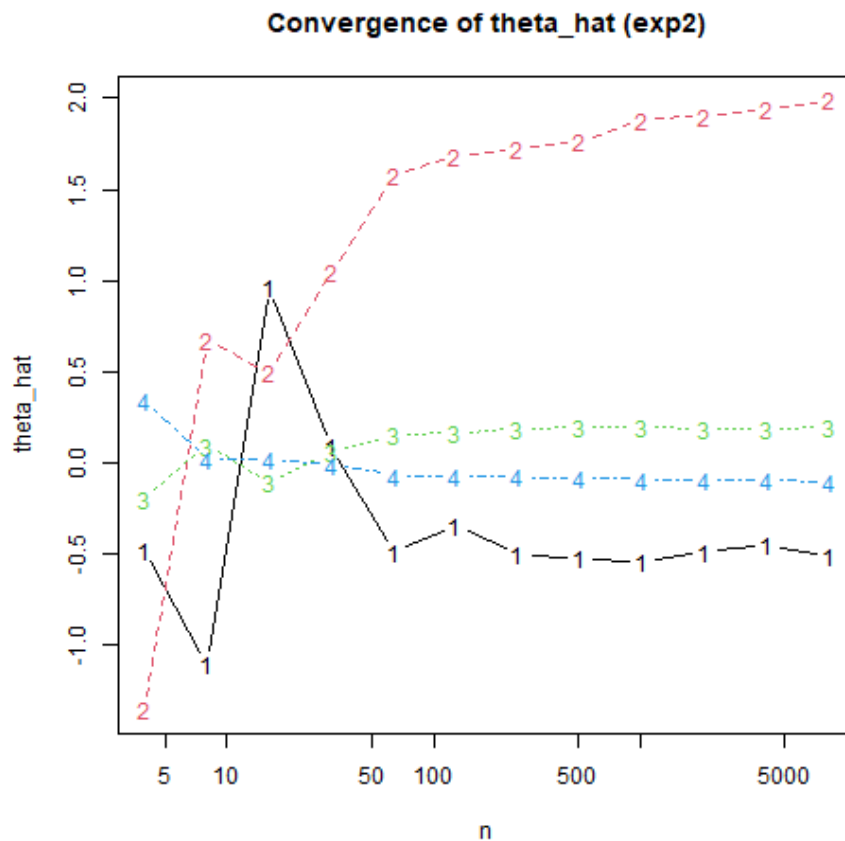


図 2:  $\hat{\theta}_N$  の収束 (横軸  $N = 4, 8, \dots, 8192$  の片対数)

■Rコード

```
1 # データ
2 data <- read.csv("datas/mmse_kadai2.csv", header=FALSE,
3                 col.names=c("x1", "y"))
4 x0 <- rep(1, nrow(data))
5 x1 <- as.matrix(data[, "x1"])
6 x2 <- x1^2; x3 <- x1^3
7 x <- cbind(x0, x1, x2, x3)
```

```

8 y <- as.matrix(data[, "y"])
9
10 # OLS 基本関数 (課題1と同じ)
11 regression_simple <- function(x, y){
12   theta_hat <- solve(t(x) %*% x) %*% t(x) %*% y
13   sigma2_hat <- as.numeric(t(y - x %*% theta_hat) %*%
14     (y - x %*% theta_hat) / (nrow(x) - ncol(x)))
15   err_cov_mat <- sigma2_hat * solve(t(x) %*% x)
16   det_coef <- sum((x %*% theta_hat - mean(y))^2) /
17     sum((y - mean(y))^2)
18   list(theta_hat=theta_hat, err_cov_mat=err_cov_mat,
19     sigma2_hat=sigma2_hat, det_coef=det_coef)
20 }
21
22 # 部分データ実験
23 exp2 <- function(x, y, n){
24   x <- x[1:n, , drop=FALSE]
25   y <- y[1:n, , drop=FALSE]
26   result <- regression_simple(x, y)
27   list(theta_hat=result$theta_hat,
28     err_cov_mat=result$err_cov_mat,
29     det_coef=result$det_coef)
30 }
31
32 # 収束図
33 plot_exp2 <- function(x, y, out="graphs/task2.png"){
34   ns <- 2^(2:13) # 4,...,8192
35   theta_hats <- matrix(NA_real_, nrow=length(ns), ncol=ncol(x))
36   for(i in seq_along(ns)){
37     theta_hats[i, ] <- as.vector(exp2(x, y, ns[i])$theta_hat)
38   }
39   dir.create(dirname(out), showWarnings=FALSE, recursive=TRUE)
40   png(out, width=960, height=600, res=120)
41   matplot(ns, theta_hats, type="b", log="x",
42     xlab="N", ylab=expression(hat(theta)),
43     main="Convergence of OLS estimates (poly degree 3)")
44   legend("bottomright",
45     legend=paste0("theta[", 0:(ncol(x)-1), "]"),
46     lty=1:ncol(x), pch=1:ncol(x))
47   dev.off()
48 }
49
50 # 実行例
51 N_max <- nrow(x)
52 print(exp2(x, y, N_max)$theta_hat)
53 print(exp2(x, y, N_max)$err_cov_mat)
54 print(exp2(x, y, N_max)$det_coef)
55 plot_exp2(x, y)

```

## ■考察

### 2.4 課題3 (分散の観測誤差: Cauchy)

■モデルと注意  $y_i = x_i^\top \theta + w_i$  ( $i = 1, \dots, N$ ),  $x_i \in \mathbb{R}^2$ . 観測誤差  $w_i \sim \text{Cauchy}(0, 1)$  独立同分布. Cauchy は二乗可積分でないため  $\mathbb{E}[w_i^2]$  が存在せず, 最小二乗法の通常仮定 (有限分散と大数の

法則)は満たされない。よって  $\sigma^2$  や  $\text{Cov}(\hat{\theta})$  は理論上定義できない。以下の分散・標準誤差・ $R^2$  は便宜的な値であり、統計的保証はない。

■推定 (形式的な OLS) 推定量・共分散・決定係数は課題 1 と同じ形式 ( $p = 2$ ) で計算する。ただし観測誤差が Cauchy 分布であり理論保証がないことに注意。

■結果 (全データ  $N = 10000$ )

$$\hat{\theta}_N = \begin{bmatrix} 2.373074 \\ 1.537311 \end{bmatrix}, \quad \widehat{\text{Cov}}(\hat{\theta}_N) = \begin{bmatrix} 9.234588215 & -0.003642121 \\ -0.003642121 & 9.264075136 \end{bmatrix}.$$

参考：対角の平方根は  $\text{SE}(\hat{\theta}_1) \approx 3.0388$ ,  $\text{SE}(\hat{\theta}_2) \approx 3.0437$ . 決定係数 (参考値) は

$$R^2 = 0.0002841468.$$

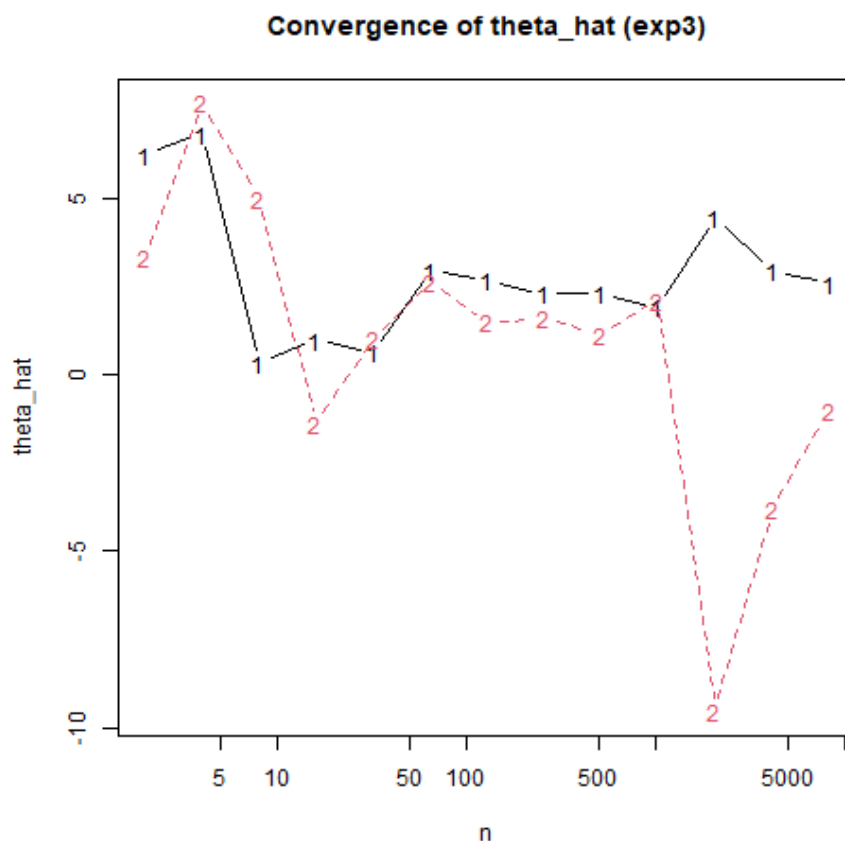


図 3:  $\hat{\theta}_N$  の非収束例 (横軸  $N = 2, 4, \dots, 8192$  の片対数)

#### ■Rコード

```
1 # データ読み込み
2 data3 <- read.csv("datas/mmse_kadai3.csv", header=FALSE,
3                   col.names=c("x1", "x2", "y"))
4 x <- as.matrix(data3[, c("x1", "x2")])
5 y <- as.matrix(data3[, "y"])
6
7 # 形式的なOLS (課題1と同様)
8 regression_simple <- function(x, y){
```



```

9   theta_hat <- solve(t(x) %*% x) %*% t(x) %*% y
10  sigma2_hat <- as.numeric(t(y - x %*% theta_hat) %*%
11                        (y - x %*% theta_hat) / (nrow(x) - ncol(x)))
12  err_cov_mat <- sigma2_hat * solve(t(x) %*% x)
13  det_coef <- sum((x %*% theta_hat - mean(y))^2) /
14            sum((y - mean(y))^2)
15  list(theta_hat=theta_hat, err_cov_mat=err_cov_mat,
16       sigma2_hat=sigma2_hat, det_coef=det_coef)
17 }
18
19 # 部分データでの実験
20 exp3 <- function(x, y, n){
21   x <- x[1:n, , drop=FALSE]
22   y <- y[1:n, , drop=FALSE]
23   res <- regression_simple(x, y)
24   list(theta_hat=res$theta_hat,
25        err_cov_mat=res$err_cov_mat,
26        det_coef=res$det_coef)
27 }
28
29 # 収束図
30 plot_exp3 <- function(x, y, out="graphs/task3.png"){
31   ns <- 2^(1:13) # 2,...,8192
32   ns <- ns[ns <= nrow(x)]
33   theta_hats <- matrix(NA_real_, nrow=length(ns), ncol=ncol(x))
34   for(i in seq_along(ns)){
35     theta_hats[i, ] <- as.vector(exp3(x, y, ns[i])$theta_hat)
36   }
37   dir.create(dirname(out), showWarnings=FALSE, recursive=TRUE)
38   png(out, width=960, height=600, res=120)
39   matplot(ns, theta_hats, type="b", log="x",
40           xlab="N", ylab=expression(hat(theta)),
41           main="Convergence of OLS under Cauchy noise")
42   legend("topright", legend=paste0("theta[",1:ncol(x),"]"),
43         lty=1:ncol(x), pch=1:ncol(x))
44   dev.off()
45 }
46
47 # 実行例
48 N_max <- nrow(x)
49 print(exp3(x, y, N_max)$theta_hat)
50 print(exp3(x, y, N_max)$err_cov_mat)
51 print(exp3(x, y, N_max)$det_coef)
52 plot_exp3(x, y)

```

■考察 資料にあるように、Cauchy 誤差では外れ値の影響が支配的で、 $\hat{\theta}_N$  は  $N$  を増やしても安定しにくいことが確認できた。

## 2.5 課題 4 (入力域の制約と設計)

■設定 課題 2 と同じ  $\varphi(x) = [1 \ x \ x^2 \ x^3]$ , 真の  $\theta$ , 観測誤差  $w_i \sim \mathcal{N}(0, 9)$  とする. ただし入力  $x_i \in [0, 1]$ ,  $i = 1, \dots, 10000$ .  $X = [\varphi(x_1); \dots; \varphi(x_N)] \in \mathbb{R}^{N \times 4}$ ,  $y = [y_1; \dots; y_N]$ .

■推定量 推定量・共分散・決定係数は課題 1 と同じ形式 ( $p = 4$ ) で計算する.

## ■Rコード

```

1 # 課題4:  $x \in [0,1]$ ,  $\phi(x)=[1 \ x \ x^2 \ x^3]$ 
2 data <- read.csv("datas/mmse_kadai4.csv", header=FALSE,
3                 col.names=c("x1","y"))
4 x0 <- rep(1, nrow(data))
5 x1 <- as.matrix(data[, "x1"])
6 x2 <- x1^2; x3 <- x1^3
7 x <- cbind(x0, x1, x2, x3)
8 y <- as.matrix(data[, "y"])
9
10 regression_simple <- function(x, y){
11   theta_hat <- solve(t(x) %*% x) %*% t(x) %*% y
12   sigma2_hat <- as.numeric(t(y - x %*% theta_hat) %*%
13                             (y - x %*% theta_hat) / (nrow(x) - ncol(x)))
14   err_cov_mat <- sigma2_hat * solve(t(x) %*% x)
15   det_coef <- sum((x %*% theta_hat - mean(y))^2) /
16               sum((y - mean(y))^2)
17   list(theta_hat=theta_hat, err_cov_mat=err_cov_mat,
18        sigma2_hat=sigma2_hat, det_coef=det_coef)
19 }
20
21 exp4 <- function(x, y, n){
22   x <- x[1:n, , drop=FALSE]
23   y <- y[1:n, , drop=FALSE]
24   regression_simple(x, y)
25 }
26
27 N_max <- nrow(x)
28 res4 <- exp4(x, y, N_max)
29 print(res4$theta_hat)
30 print(res4$err_cov_mat)
31 print(res4$det_coef) # R^2

```

## ■結果 ( $N = 10000$ )

$$\hat{\theta}_{(4)} = \begin{bmatrix} \text{ここに数値} \\ \text{ここに数値} \\ \text{ここに数値} \\ \text{ここに数値} \end{bmatrix}, \quad \widehat{\text{Cov}}(\hat{\theta}_{(4)}) = \begin{bmatrix} \text{ここに } 4 \times 4 \text{ 行列} \end{bmatrix}, \quad R^2 = \begin{bmatrix} \text{ここに数値} \end{bmatrix}.$$

(上の枠に、直上の R 出力を転記)

## ■課題 2.1 との比較 ( $N = 10000$ ) 課題 2.1 の推定結果：

$$\hat{\theta}_{(2.1)} = \begin{bmatrix} -0.50902942 \\ 1.97586067 \\ 0.19774405 \\ -0.09866691 \end{bmatrix}, \quad R^2 = 0.461855.$$

評価観点：

- 分散は  $\text{Var}(\hat{\theta}) = \sigma^2(X^\top X)^{-1}$  に比例.  $x \in [0,1]$  では列  $\{1, x, x^2, x^3\}$  が強く相関しやすく,  $X^\top X$  の条件が悪化し  $\hat{\theta}$  の不確かさが増えやすい.
- $R^2$  は母集団の  $y$  の散らばり ( $SST$ ) に依存. 入力域が狭いと  $SST$  が小さく, 雑音分散が同じなら  $R^2$  は低下しやすい.

### ■考察（どうデータを取るべきか）

- 目的は  $\text{Var}(\hat{\theta})$  の縮小（ $= X^\top X$  を「大きく」「良条件」に）。
- 入力設計： $x$  を区間全体で広く配置（例：一様），中心化・標準化して  $[-1, 1]$  に写像，または直交基底（Legendre/Chebyshev）で回帰。
- 最適化観点：A/D 最適設計を用いて  $\text{tr}((X^\top X)^{-1})$  や  $\det((X^\top X)^{-1})$  を最小化する点集合を選ぶ。

## 3 重み付き最小二乗法

### 3.1 原理と方法

■原理 各観測  $i = 1, \dots, N$  で  $y_i \in \mathbb{R}^m$ ,  $X_i \in \mathbb{R}^{m \times p}$  とし

$$y_i = X_i \theta + w_i, \quad \mathbb{E}[w_i] = 0, \quad \text{Cov}(w_i) = V \text{ (既知)},$$

を仮定する。 $Q := V^{-1}$  とおく。WLS は重み付き残差平方和

$$J(\theta) = \sum_{i=1}^N (y_i - X_i \theta)^\top Q (y_i - X_i \theta)$$

を最小化する推定で，正規方程式は

$$S \hat{\theta} = b, \quad S := \sum_{i=1}^N X_i^\top Q X_i, \quad b := \sum_{i=1}^N X_i^\top Q y_i.$$

したがって

$$\hat{\theta} = \left( \sum_{i=1}^N X_i^\top Q X_i \right)^{-1} \left( \sum_{i=1}^N X_i^\top Q y_i \right).$$

$V$  が既知のとき

$$\text{Cov}(\hat{\theta}) = S^{-1}.$$

$V = \sigma^2 \Sigma$  のようにスケール未知なら

$$\hat{\sigma}^2 = \frac{\sum_{i=1}^N r_i^\top \Sigma^{-1} r_i}{Nm - p}, \quad r_i := y_i - X_i \hat{\theta}, \quad \widehat{\text{Cov}}(\hat{\theta}) = \hat{\sigma}^2 \left( \sum_{i=1}^N X_i^\top \Sigma^{-1} X_i \right)^{-1}.$$

■方法 (1) 各  $i$  の設計行列  $X_i$  と観測  $y_i$  を用意（ $X_i$  は  $m \times p$ ）。(2)  $Q = V^{-1}$  を決めて  $S = \sum X_i^\top Q X_i$ ,  $b = \sum X_i^\top Q y_i$  を計算。(3)  $\hat{\theta} = S^{-1}b$ 。

■実装 R による実装例を以下に示す。

```
1 regression_multiple <- function(x, y, V = NULL, Q = NULL){
2   # y: n×1行列（またはベクトル）
3   if (is.null(ncol(y))) {
4     y <- as.matrix(y, ncol = 1)
5   }
6   if (length(dim(x)) != 3) {
```

```

7      stop("x must be a 3-dimensional array")
8  }
9  n <- dim(x)[3]
10 m <- ncol(y)
11 p <- dim(x)[1] # 特徴量数
12 S <- matrix(0, p, p)
13 b <- matrix(0, p, 1)
14 T <- matrix(0, p, p)
15 if (is.null(V)) {
16     V <- diag(m)
17 }
18 if (is.null(Q)) {
19     Q <- solve(V)
20 }
21 # m=1 (スカラー出力) の場合はQ,Vをスカラー化
22 if (m == 1) {
23     Q <- as.numeric(Q)
24     V <- as.numeric(V)
25     for (i in 1:n) {
26         xi <- as.matrix(x[, , i]) # (p,1)
27         S <- S + xi %*% t(xi) * Q
28         b <- b + xi * Q * y[i, 1]
29         T <- T + xi %*% t(xi) * V
30     }
31 } else {
32     for (i in 1:n) {
33         xi <- as.matrix(x[, , i])
34         yi <- matrix(y[i, ], nrow = m, ncol = 1)
35         S <- S + t(xi) %*% Q %*% xi
36         b <- b + t(xi) %*% Q %*% yi
37         T <- T + t(xi) %*% V %*% xi
38     }
39 }
40 theta_hat <- solve(S) %*% b
41 err_cov_mat <- solve(S) %*% T %*% solve(S)
42 return(list(theta_hat = theta_hat, err_cov_mat = err_cov_mat, S = S, b = b, T = T
43 ))

```

### 3.2 課題 5 (2 次元出力：重み付き最小二乗法)

■課題の内容 入力  $x_i \in \mathbb{R}$  ( $i = 1, \dots, 1000$ ) に対し,  $m = 2$  次元出力の線形モデル

$$y_i = \phi(x_i)\theta + w_i, \quad \phi(x) = \begin{bmatrix} 1 & x \\ 1 & x^2 \end{bmatrix}, \quad w_i \sim \mathcal{N}(0, V), \quad V = \begin{bmatrix} 100 & 0 \\ 0 & 1 \end{bmatrix},$$

を考える. ここで  $\theta \in \mathbb{R}^2$  を推定対象とする.

■推定法 最小二乗 (OLS) の推定値は

$$\hat{\theta}_N = \left( \sum_{i=1}^N \phi_i^\top \phi_i \right)^{-1} \sum_{i=1}^N \phi_i^\top y_i,$$

で与えられる ( $\phi_i = \phi(x_i)$ ). 観測雑音の共分散  $V$  が既知で, その影響を反映した重み付き最小二乗 (WLS) を用いると,  $V = W^{-2}$  を満たす  $W$  に対し

$$\hat{\theta}_N = \left( \sum_{i=1}^N \phi_i^\top W^{-2} \phi_i \right)^{-1} \sum_{i=1}^N \phi_i^\top W^{-2} y_i = \left( \sum_{i=1}^N \phi_i^\top Q \phi_i \right)^{-1} \sum_{i=1}^N \phi_i^\top Q y_i,$$

となる ( $Q = W^{-2} = V^{-1}$ ).

■実装 `datas/mmse_kadai5.csv` を読み込み,  $x$  から  $\phi(x) = \begin{bmatrix} 1 & x \\ 1 & x^2 \end{bmatrix}$  を構成して  $x \in \mathbb{R}^{2 \times 2 \times N}$  の 3 次元配列とし,  $y \in \mathbb{R}^{N \times 2}$  を観測行列とした. 推定は関数 `regression_multiple` を用い, **OLS** は無重み ( $Q = I$ ), **WLS** は  $V = \text{diag}(100, 1)$  を渡し内部で  $Q = V^{-1}$  を用いる設定とした. また,  $N \in \{4, 8, 16, 32, 64, 128, 256, 512, 1000\}$  に対する  $\hat{\theta}_N$  の収束の様子を片対数 ( $x$  軸のみ対数) で可視化した.

■結果 全データ ( $N = 1000$ ) を用いた推定結果は次のとおりである.

**OLS**

$$\hat{\theta}_{\text{OLS}} = \begin{bmatrix} 2.994567 \\ -2.068971 \end{bmatrix}, \quad \widehat{\text{Cov}}(\hat{\theta}_{\text{OLS}}) = \begin{bmatrix} 5.7621 \times 10^{-4} & -1.5041 \times 10^{-4} \\ -1.5041 \times 10^{-4} & 2.9684 \times 10^{-4} \end{bmatrix}.$$

**WLS**

$$\hat{\theta}_{\text{WLS}} = \begin{bmatrix} 2.939084 \\ -1.986465 \end{bmatrix}, \quad \widehat{\text{Cov}}(\hat{\theta}_{\text{WLS}}) = \begin{bmatrix} 2.39993 \times 10^{-1} & -9.66541 \times 10^{-2} \\ -9.66541 \times 10^{-2} & 4.99126 \times 10^{-2} \end{bmatrix}.$$

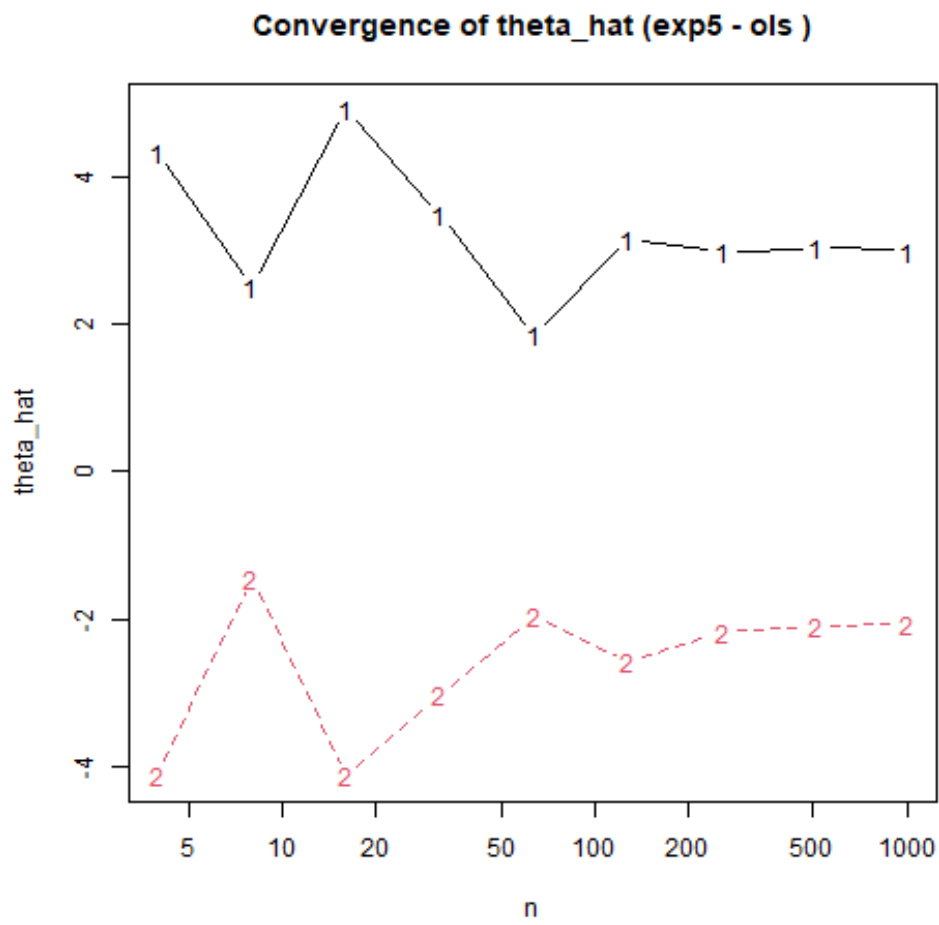


図 4:  $\hat{\theta}_N$  の収束 (OLS)

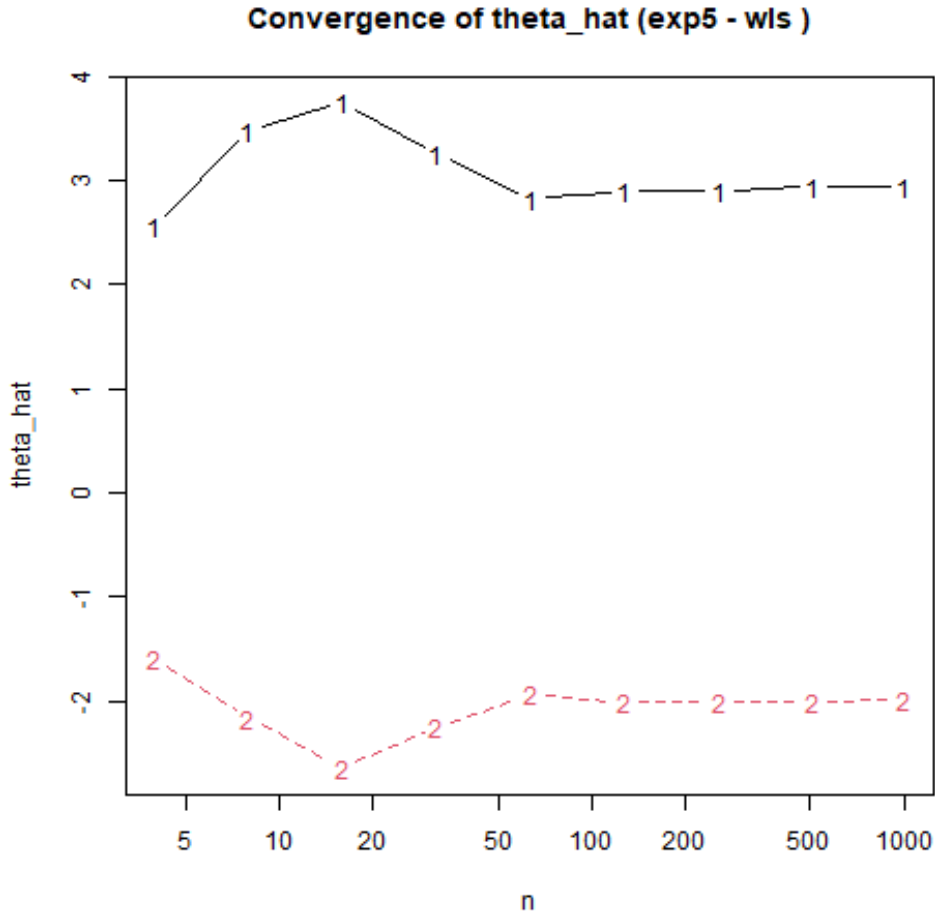


図 5:  $\hat{\theta}_N$  の収束 (WLS,  $V = \text{diag}(100, 1)$ )

## ■考察

### 3.3 課題 6 (2 次元出力・異分散 2 群に対する重み付き最小二乗)

■課題の内容 入力  $x_i \in \mathbb{R}$  に対し,

$$y_i = \phi(x_i)\theta + w_i, \quad \phi(x) = \begin{bmatrix} 1 & x \\ 1 & x^2 \end{bmatrix}, \quad \theta \in \mathbb{R}^2,$$

という  $m = 2$  次元出力の線形モデルを考える. 雑音は  $N$  サンプルのうち前半  $i = 1, \dots, \text{num}$  が  $w_i \sim \mathcal{N}(0, V_1)$ , 後半  $i = \text{num} + 1, \dots, N$  が  $w_i \sim \mathcal{N}(0, V_2)$  に従うとする. 本課題では  $V_1 = \text{diag}(100, 1)$ ,  $V_2 = \text{diag}(2, 1)$ ,  $\text{num} = 500$  を既知として推定を行う.

■推定法 (GLS/WLS)  $Q_k = V_k^{-1}$  ( $k = 1, 2$ ) とおくと, 一般化最小二乗 (WLS) の正規方程式は

$$S\hat{\theta}_N = b, \quad S = \sum_{i=1}^{\text{num}} \phi_i^\top Q_1 \phi_i + \sum_{i=\text{num}+1}^N \phi_i^\top Q_2 \phi_i, \quad b = \sum_{i=1}^{\text{num}} \phi_i^\top Q_1 y_i + \sum_{i=\text{num}+1}^N \phi_i^\top Q_2 y_i,$$

より

$$\hat{\theta}_N = S^{-1}b.$$

標本ベースの誤差共分散推定は（実装に合わせて）

$$\widehat{\text{Cov}}(\hat{\theta}_N) = S^{-1} T S^{-1}, \quad T = \sum_{i=1}^{\text{num}} \phi_i^\top V_1 \phi_i + \sum_{i=\text{num}+1}^N \phi_i^\top V_2 \phi_i.$$

なお OLS は  $Q_1 = Q_2 = I$  とした特別な場合である.

■実装 `datas/mmse_kadai6.csv` から  $x$  を読み,  $\phi(x)$  を用いて  $x \in \mathbb{R}^{2 \times 2 \times N}$ ,  $y \in \mathbb{R}^{N \times 2}$  を構成した. 推定は関数 `regression_multiple_different_distributions` により, 前半  $V_1$ , 後半  $V_2$  の重みで  $S, b, T$  を分割加算して  $\hat{\theta}_N$  と  $\widehat{\text{Cov}}(\hat{\theta}_N)$  を得た. また  $N$  を `ns = seq(4, 1000, length.out=20)` で走らせ,  $\hat{\theta}_N$  の収束の様子を折れ線で可視化した ( $x$  軸は線形目盛).

■結果 全データ ( $N = 1000$ , `num` = 500) の推定結果は以下のとおりである.

**WLS** ( $V_1 = \text{diag}(100, 1)$ ,  $V_2 = \text{diag}(2, 1)$ )

$$\hat{\theta}_{\text{WLS}} = \begin{bmatrix} 2.994202 \\ -2.014699 \end{bmatrix}, \quad \widehat{\text{Cov}}(\hat{\theta}_{\text{WLS}}) = \begin{bmatrix} 6.019710 \times 10^{-2} & -2.227335 \times 10^{-2} \\ -2.227335 \times 10^{-2} & 1.276759 \times 10^{-2} \end{bmatrix}.$$

**OLS (比較)**

$$\hat{\theta}_{\text{OLS}} = \begin{bmatrix} 3.188356 \\ -2.092183 \end{bmatrix}, \quad \widehat{\text{Cov}}(\hat{\theta}_{\text{OLS}}) = \begin{bmatrix} 5.692084 \times 10^{-4} & -1.402629 \times 10^{-4} \\ -1.402629 \times 10^{-4} & 2.842674 \times 10^{-4} \end{bmatrix}.$$



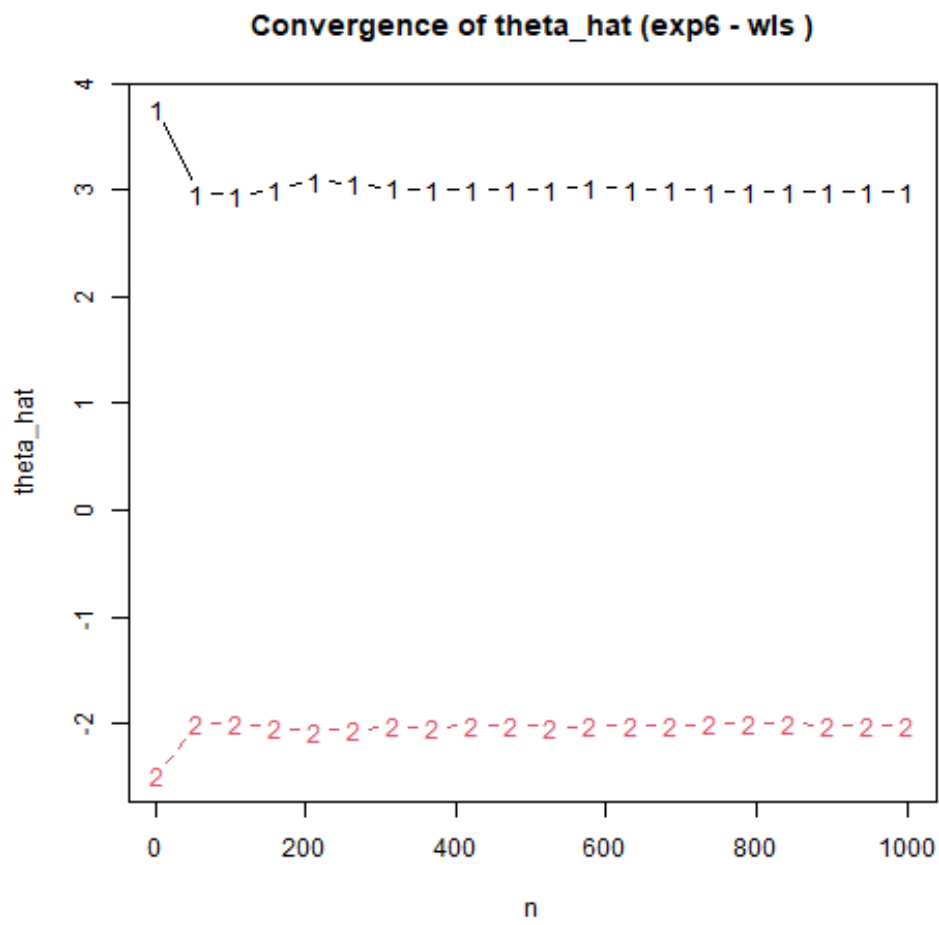


図 6:  $\hat{\theta}_N$  の収束 (WLS,  $V_1/V_2$  異分散)

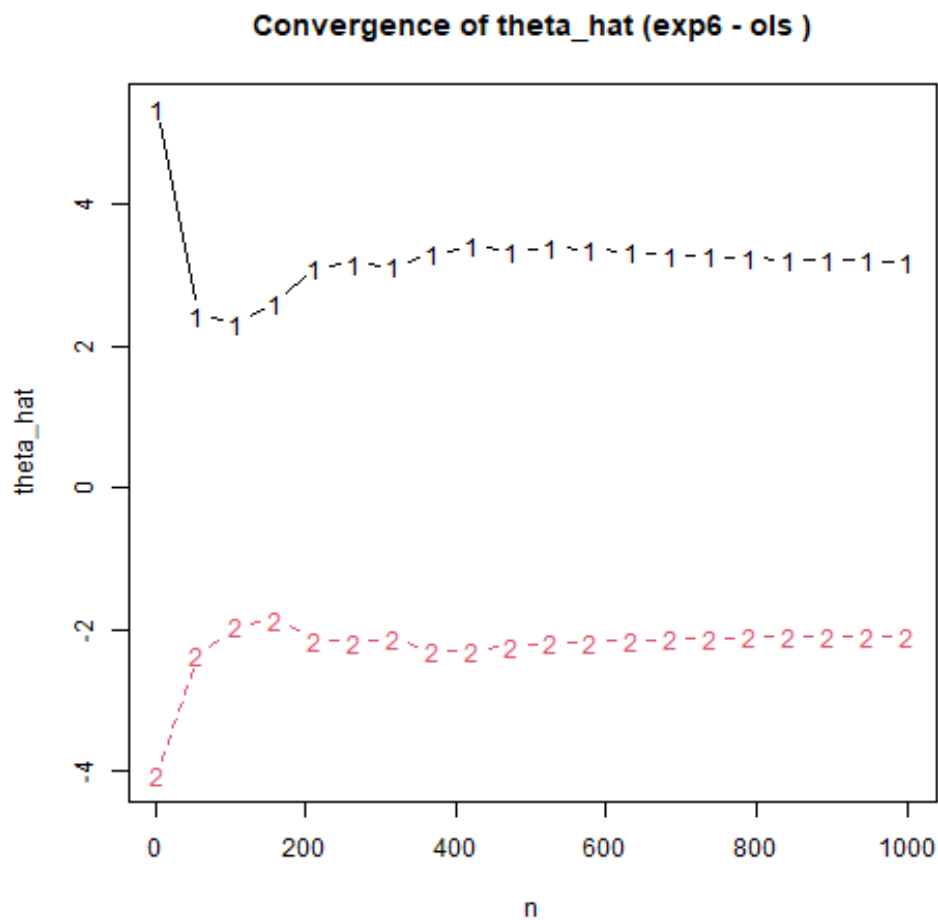


図 7:  $\hat{\theta}_N$  の収束 (OLS)

## ■考察

### 4 逐次最小二乗法

### 5 アルゴリズム比較

表や図を用いて複数手法の比較を行う。計算量・精度・安定性・実行時間の観点でまとめる。

### 6 結論

本章での結論と今後の課題を箇条書きでまとめる。

## 付録 A 使用コード一覧

主要スクリプトと入手先を列挙する。

- 実験コード (R) : <https://github.com/<your-repo>>

- データ生成スクリプト：`scripts/generate_data.jl`

## 付録 B 補足導出

必要な導出や補助的な数学的議論を載せる。

## 参考文献

## 参考文献

[1] 数理工学実験（2025 年度配布資料）。