

2012 年度 大問 2

hari64boli64 (hari64boli64@gmail.com)

2025 年 4 月 23 日

1 問題

線形回帰モデル

2 解答

(1)

$$\hat{\beta} = (X^{\top} X)^{-1} X^{\top} y$$

(2)

$$\mathbb{E}[\hat{\beta}] = \beta$$

(3)

$$\Sigma = \sigma^2 (X^{\top} X)^{-1}$$

(4)

$$\mathbb{E}[(y_{n+1} - x_{n+1}^{\top} \beta)^2] = \sigma^2 (1 + x_{n+1}^{\top} (X^{\top} X)^{-1} x_{n+1})$$

3 おまけ

Listing 1 problem4

```

1 import numpy as np
2 import matplotlib.pyplot as plt
3 from tqdm import tqdm
4
5 squared_errors = []
6 answers = []
7 for _ in tqdm(range(1000)):
8     n = np.random.randint(1, 100)
9     p = np.random.randint(1, 10)
10    X = np.random.random((n, p))
11    beta = np.random.random(p)
12    sigma = np.random.random()
13    epsilon = np.random.normal(0, sigma, n)
14    y = X.dot(beta) + epsilon
15    if np.linalg.matrix_rank(X) != p:
16        continue
17    beta_hat = np.linalg.inv(X.T.dot(X)).dot(X.T).dot(y)
18    mse = 0
19    for _ in range(100):
20        x_np1 = np.random.random(p)
21        y_np1 = x_np1.T.dot(beta) + np.random.normal(0, sigma)
22        squared_error = (y_np1 - x_np1.T.dot(beta_hat)) ** 2
23        mse += squared_error
24    mse /= 100
25    squared_errors.append(mse)
26    answers.append(
27        (sigma**2) * (1 + x_np1.T.dot(np.linalg.inv(X.T.dot(X)).
28            dot(x_np1)))
29
30 plt.figure(figsize=(6, 6))
31 plt.scatter(answers, squared_errors)
32 plt.title("Simulated Squared Error vs. Theoretical Squared Error")
33
34 plt.xlabel("Theoretical Squared Error")
35 plt.ylabel("Simulated Squared Error")
36 plt.xscale("log")
37 plt.yscale("log")
38 plt.xlim(min(answers + squared_errors), max(answers +
39     squared_errors))
40 plt.ylim(min(answers + squared_errors), max(answers +
41     squared_errors))
42 plt.minorticks_on()
43 plt.grid(which="major", linestyle="-", linewidth="0.5", color="
44     black")

```

```

41 plt.grid(which="minor", linestyle="--", linewidth="0.5", color="
    black")
42 plt.savefig("2.png")

```

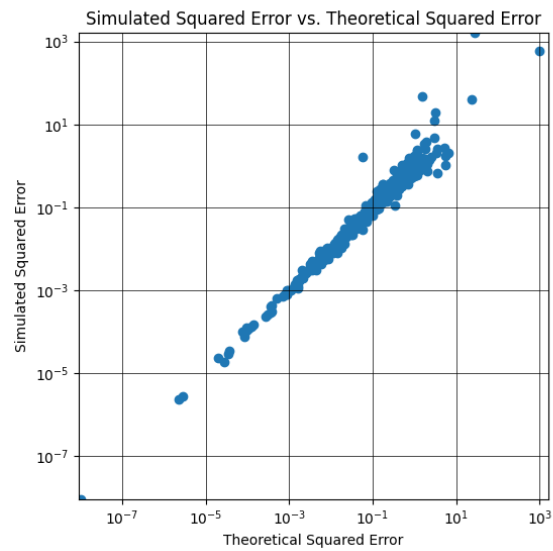


図 1 問題 4 のグラフ (シミュレーション値と理論値)