

Міністерство освіти і науки України
Харківський національний університет радіоелектроніки
Кафедра програмної інженерії

ЗВІТ

Практичної роботи № 1
з дисципліни «Інтелектуальний аналіз даних»
на тему «Метод найменших квадратів (OLS)»

Виконав

студент групи ІПЗм-24-2

Голодніков Дмитро

Перевірив

ст. викл. Онищенко К.Г.

Харків 2024

1 МЕТА РОБОТИ

Вивчити метод найменших квадратів (Ordinary Least Squares, OLS) для побудови регресійних моделей. Провести аналіз реальних економічних даних підприємства. Оцінити якість моделі та перевірити виконання припущень OLS.

2 ХІД ВИКОНАННЯ РОБОТИ

2.1 Опис даних

Використано дані виробничо-економічної діяльності машинобудівного підприємства. Залежна змінна Y_1 - продуктивність праці. Незалежні змінні X_1 - X_{10} характеризують різні аспекти діяльності підприємства.

2.2 Кореляційний аналіз

Побудовано матрицю кореляцій для визначення зв'язків між змінними та виявлення потенційної мультиколінеарності.

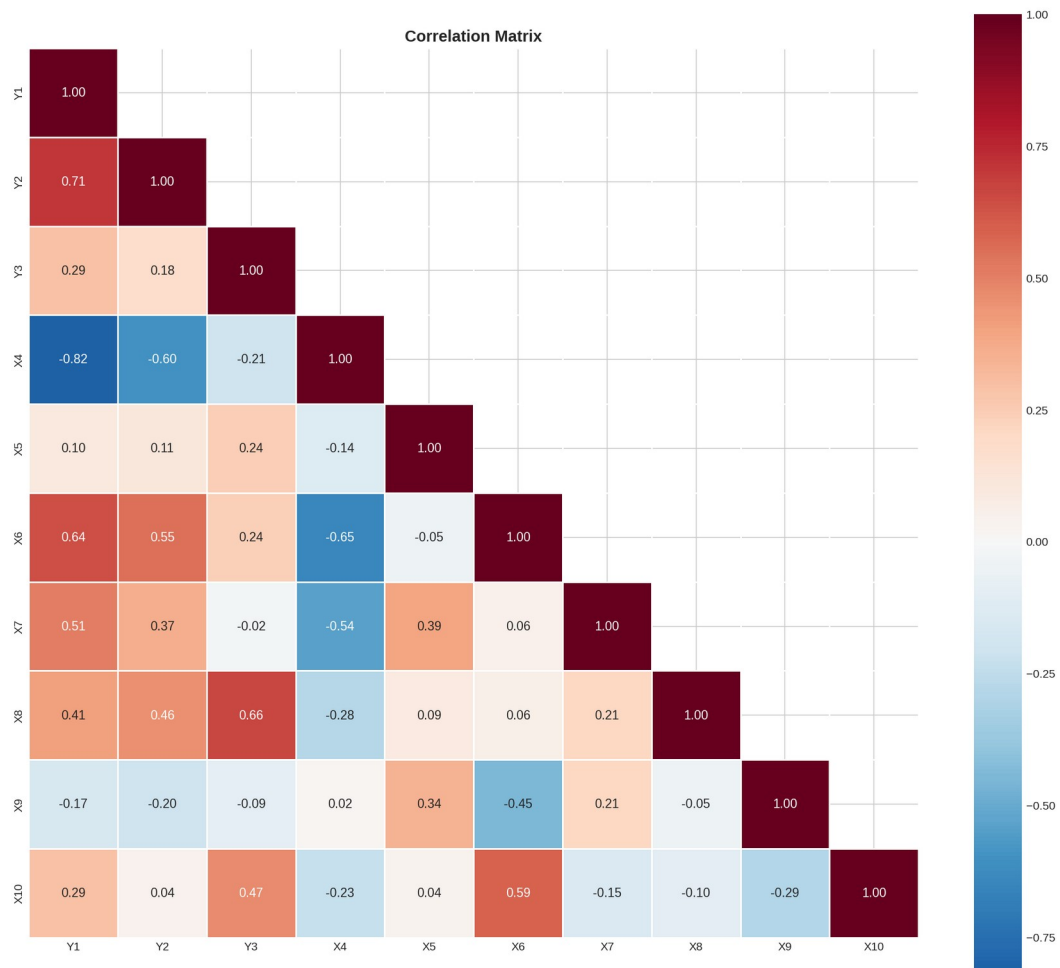


Рис. 2.1 - Матриця кореляцій змінних

2.3 Побудова OLS моделі

Побудовано регресійну модель методом найменших квадратів. Реалізація OLS регресії з використанням statsmodels:

```
import statsmodels.api as sm

# Підготовка даних
X = data[['X1', 'X2', 'X5', 'X6', 'X8']] # Відібрані предиктори
y = data['Y1'] # Продуктивність праці

# Додавання константи (intercept)
X_const = sm.add_constant(X)

# Побудова OLS моделі
model = sm.OLS(y, X_const).fit()

# Виведення результатів
print(model.summary())
```

```
# Ключові метрики
r_squared = model.rsquared
adj_r_squared = model.rsquared_adj
f_statistic = model.fvalue
p_values = model.pvalues
```

Проведено оцінку значущості коефіцієнтів та загальної якості моделі.

2.4 Діагностика моделі

Проведено діагностику моделі: аналіз залишків, перевірка нормальності розподілу, виявлення впливових спостережень.

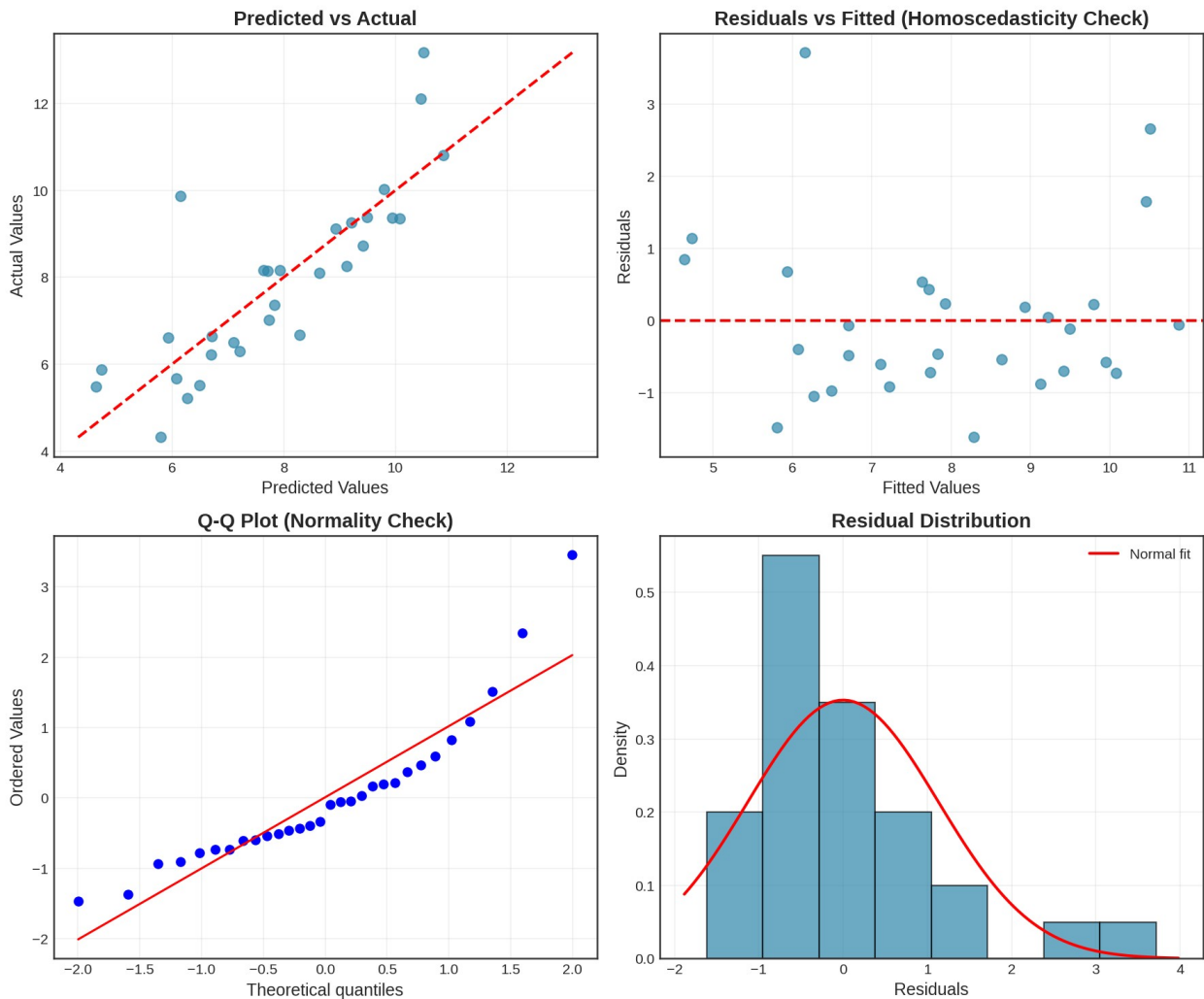


Рис. 2.2 - Діагностичні графіки OLS регресії

3 ВИСНОВКИ

У ході виконання практичної роботи було застосовано метод найменших квадратів для аналізу економічних даних підприємства. Проведено кореляційний

аналіз та відібрано значущі предиктори. Побудовано та оцінено OLS регресійну модель. Виконано діагностику моделі та перевірено виконання припущень методу. Результати можуть бути використані для прогнозування продуктивності праці.

ПОСИЛАННЯ

Код проєкту доступний у репозиторії GitHub: <https://github.com/na-naina/data-analysis-khnure>