Statistics and Data Analysis

Unit 04 – Lecture 03: Multiple Linear Regression

Tofik Ali

School of Computer Science, UPES Dehradun

February 14, 2026

https://github.com/tali7c/Statistics-and-Data-Analysis

Quick Links

Overview Model Interpretation Exercises Demo Summary

Agenda

Learning Outcomes

- Write the multiple linear regression model

Learning Outcomes

- Write the multiple linear regression model
- Interpret a coefficient as a partial effect

Learning Outcomes

- Write the multiple linear regression model
- Interpret a coefficient as a partial effect
- Explain dummy variables for categories (basic)

Learning Outcomes

- Write the multiple linear regression model
- Interpret a coefficient as a partial effect
- Explain dummy variables for categories (basic)
- Explain adjusted R-squared (intuition)

Model: Key Points

- y = b0 + b1 x1 + b2 x2 + ...

Model: Key Points

- $y = b0 + b1\ x1 + b2\ x2 + \ldots$
- Each coefficient is a partial effect (others fixed)

Model: Key Points

- $y = b0 + b1\ x1 + b2\ x2 + \dots$
- Each coefficient is a partial effect (others fixed)
- Scaling helps when using regularization

Model: Key Formula

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \epsilon$$

Interpretation: Key Points

- Dummy variables encode categories

Interpretation: Key Points

- Dummy variables encode categories
- Adjusted R^2 penalizes unnecessary predictors

Interpretation: Key Points

- Dummy variables encode categories
- Adjusted R^2 penalizes unnecessary predictors
- Multicollinearity can harm interpretability

Exercise 1: Partial effect

Model: yhat=5 + 0.8x1 + 2.0x2. Interpret coefficient 2.0.

Solution 1

- Holding x1 fixed, +1 in x2 increases yhat by 2.0 units.

Exercise 2: Dummy variable

Urban=1, Rural=0. If coef(Urban)=10, interpret.

Solution 2

- Urban has predicted y about 10 units higher than Rural (all else equal).

Exercise 3: Adjusted R^2

Why use adjusted R^2 when comparing models with different number of predictors?

Solution 3

- Because R^2 never decreases, adjusted R^2 penalizes extra predictors.

Mini Demo (Python)

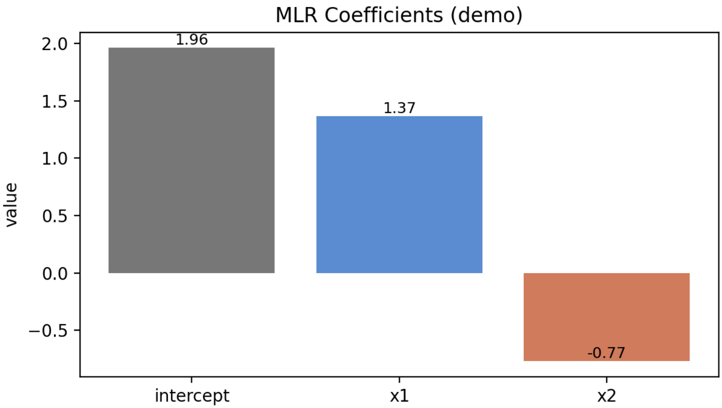
Run from the lecture folder:

```
python demo/demo.py
```

Outputs:

- `images/demo.png`
- `data/results.txt`

Demo Output (Example)



Summary

- Key definitions and the main formula.

Summary

- Key definitions and the main formula.
- How to interpret results in context.

Summary

- Key definitions and the main formula.
- How to interpret results in context.
- How the demo connects to the theory.

Exit Question

Why does adding a useless feature still increase (or keep) R^2?