

Multiple Linear Regression

- Here our model is

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_p X_p + \epsilon,$$

- We interpret β_j as the *average* effect on Y of a one unit increase in X_j , *holding all other predictors fixed*. In the advertising example, the model becomes

$$\text{sales} = \beta_0 + \beta_1 \times \text{TV} + \beta_2 \times \text{radio} + \beta_3 \times \text{newspaper} + \epsilon.$$

Interpreting regression coefficients

- The ideal scenario is when the predictors are uncorrelated — a *balanced design*:
 - Each coefficient can be estimated and tested separately.
 - Interpretations such as “*a unit change in X_j is associated with a β_j change in Y , while all the other variables stay fixed*”, are possible.
- Correlations amongst predictors cause problems:
 - The variance of all coefficients tends to increase, sometimes dramatically
 - Interpretations become hazardous — when X_j changes, everything else changes.
- *Claims of causality* should be avoided for observational data.

The woes of (interpreting) regression coefficients

“Data Analysis and Regression” Mosteller and Tukey 1977

- a regression coefficient β_j estimates the expected change in Y per unit change in X_j , *with all other predictors held fixed*. But predictors usually change together!

The woes of (interpreting) regression coefficients

“Data Analysis and Regression” Mosteller and Tukey 1977

- a regression coefficient β_j estimates the expected change in Y per unit change in X_j , *with all other predictors held fixed*. But predictors usually change together!
- Example: Y total amount of change in your pocket; $X_1 = \#$ of coins; $X_2 = \#$ of pennies, nickels and dimes. By itself, regression coefficient of Y on X_2 will be > 0 . But how about with X_1 in model?

The woes of (interpreting) regression coefficients

“Data Analysis and Regression” Mosteller and Tukey 1977

- a regression coefficient β_j estimates the expected change in Y per unit change in X_j , *with all other predictors held fixed*. But predictors usually change together!
- Example: Y total amount of change in your pocket; $X_1 = \#$ of coins; $X_2 = \#$ of pennies, nickels and dimes. By itself, regression coefficient of Y on X_2 will be > 0 . But how about with X_1 in model?
- $Y =$ number of tackles by a football player in a season; W and H are his weight and height. Fitted regression model is $\hat{Y} = b_0 + .50W - .10H$. How do we interpret $\hat{\beta}_2 < 0$?

Two quotes by famous Statisticians

“Essentially, all models are wrong, but some are useful”

George Box

Two quotes by famous Statisticians

“Essentially, all models are wrong, but some are useful”

George Box

“The only way to find out what will happen when a complex system is disturbed is to disturb the system, not merely to observe it passively”

Fred Mosteller and John Tukey, paraphrasing George Box

Estimation and Prediction for Multiple Regression

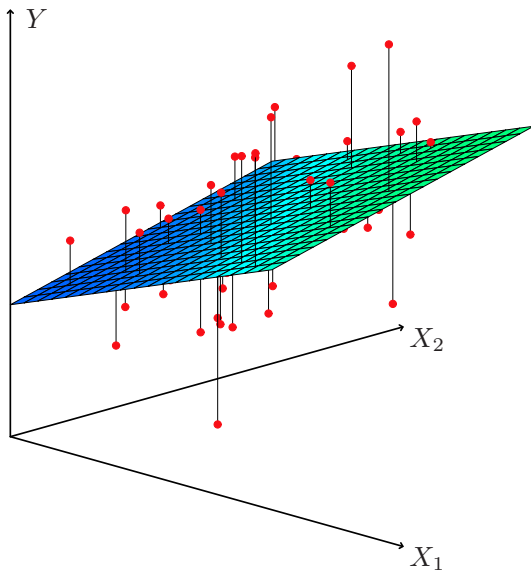
- Given estimates $\hat{\beta}_0, \hat{\beta}_1, \dots, \hat{\beta}_p$, we can make predictions using the formula

$$\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x_1 + \hat{\beta}_2 x_2 + \dots + \hat{\beta}_p x_p.$$

- We estimate $\beta_0, \beta_1, \dots, \beta_p$ as the values that minimize the sum of squared residuals

$$\begin{aligned} \text{RSS} &= \sum_{i=1}^n (y_i - \hat{y}_i)^2 \\ &= \sum_{i=1}^n (y_i - \hat{\beta}_0 - \hat{\beta}_1 x_{i1} - \hat{\beta}_2 x_{i2} - \dots - \hat{\beta}_p x_{ip})^2. \end{aligned}$$

This is done using standard statistical software. The values $\hat{\beta}_0, \hat{\beta}_1, \dots, \hat{\beta}_p$ that minimize RSS are the multiple least squares regression coefficient estimates.



Results for advertising data

| | Coefficient | Std. Error | t-statistic | p-value |
|-----------|-------------|------------|-------------|----------|
| Intercept | 2.939 | 0.3119 | 9.42 | < 0.0001 |
| TV | 0.046 | 0.0014 | 32.81 | < 0.0001 |
| radio | 0.189 | 0.0086 | 21.89 | < 0.0001 |
| newspaper | -0.001 | 0.0059 | -0.18 | 0.8599 |

Correlations:

| | TV | radio | newspaper | sales |
|-----------|--------|--------|-----------|--------|
| TV | 1.0000 | 0.0548 | 0.0567 | 0.7822 |
| radio | | 1.0000 | 0.3541 | 0.5762 |
| newspaper | | | 1.0000 | 0.2283 |
| sales | | | | 1.0000 |