

Ch 3.1: Linear Regression

Lecture 4 - CMSE 381

Prof. Elizabeth Munch

Michigan State University

::

Dept of Computational Mathematics, Science & Engineering

Fri Sep 6, 2024

Last time:

- 2.2 Assessing Model Accuracy

Announcements:

- Office Hours

Covered in this lecture

- Least squares coefficient estimates for linear regression
- Residual sum of squares (RSS)

Section 1

Simple Linear Regression

- Predict Y on a single predictor variable X

$$Y \approx \beta_0 + \beta_1 X$$

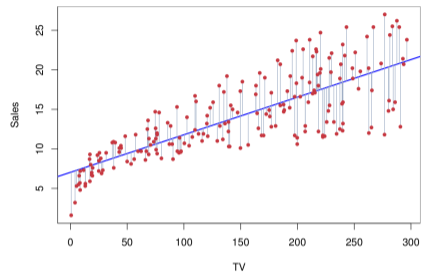
- " \approx " "is approximately modeled as"

Example

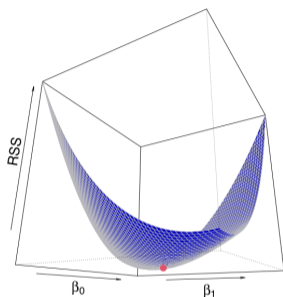
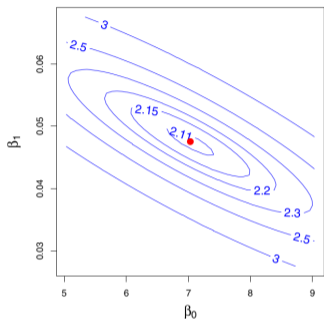
1		TV	Radio	Newspaper	Sales
2	1	230.1	37.8	69.2	22.1
3	2	44.5	39.3	45.1	10.4
4	3	17.2	45.9	69.3	9.3
5	4	151.5	41.3	58.5	18.5
6	5	180.8	10.8	58.4	12.9
7	6	8.7	48.9	75	7.2
8	7	57.5	32.8	23.5	11.8
9	8	120.2	19.6	11.6	13.2
10	9	8.6	2.1	1	4.8
11	10	199.8	2.6	21.2	10.6
12	11	66.1	5.8	24.2	8.6

Least squares criterion: Setup

How do we estimate the coefficients?



Least squares criterion: RSS



Residual sum of squares RSS is

$$\begin{aligned}RSS &= e_1^2 + \cdots + e_n^2 \\ &= \sum_i (y_i - \hat{\beta}_0 - \hat{\beta}_1 x_i)^2\end{aligned}$$

$$\text{sales} \approx \beta_0 + \beta_1 \text{TV}$$

Least squares criterion

Find β_0 and β_1 that minimize the RSS.

Least squares coefficient estimates

$$\min_{\beta_0, \beta_1} \sum_i (y_i - \hat{\beta}_0 - \hat{\beta}_1 x_i)^2$$

$$\frac{\partial RSS}{\partial \beta_0} = -2 \sum_i (y_i - \beta_0 - \beta_1 x_i) = 0$$

$$\frac{\partial RSS}{\partial \beta_1} = -2 \sum_i x_i (y_i - \beta_0 - \beta_1 x_i) = 0$$

$$\hat{\beta}_1 = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^n (x_i - \bar{x})^2}$$

$$\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x}$$

Coding group work

Next time

Lec #	Date			Reading	HW
1	Mon	8/26	Intro / First day stuff / Python Review Pt 1	1	
2	Wed	8/28	What is statistical learning?	2.1	
	Fri	8/30	Class Cancelled (Dr Munch out of town)		
	Mon	9/2	No class - Labor day		
3	Wed	9/4	Assessing Model Accuracy	2.2.1, 2.2.2	
4	Fri	9/6	Linear Regression	3.1	HW #1 Due Sun 9/8
5	Mon	9/9	More Linear Regression	3.1/3.2	
6	Wed	9/11	Even more linear regression	3.2.2	
7	Fri	9/13	Probably more linear regression	3.3	Hw #2 Due

Announcements

- Homework 1
 - ▶ Due Sun, Sep 8