Ch 3.1: Linear Regression

Lecture 4 - CMSE 381

Michigan State University
::
Dept of Computational Mathematics, Science & Engineering

Wed Jan 21, 2026

Announcements

Last time:

- 2.2 Assessing Model Accuracy

Covered in this lecture

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

Section 1

Simple Linear Regression

Setup

- Predict Y on a single predictor variable X

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

Setup

- Predict Y on a single predictor variable X

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

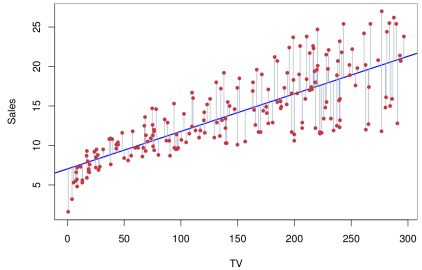
- "$\approx$" .... "is approximately modeled as"

Example

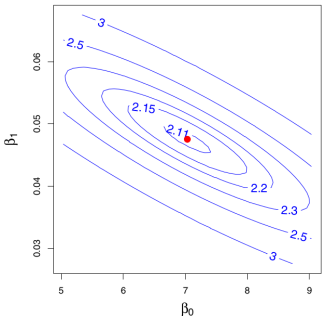
	TV	Radio	Newspaper	Sales
1	230.1	37.8	69.2	22.1
2	44.5	39.3	45.1	10.4
3	17.2	45.9	69.3	9.3
4	151.5	41.3	58.5	18.5
5	180.8	10.8	58.4	12.9
6	8.7	48.9	75	7.2
7	57.5	32.8	23.5	11.8
8	120.2	19.6	11.6	13.2
9	8.6	2.1	1	4.8
10	199.8	2.6	21.2	10.6
11	66.1	5.8	24.2	8.6

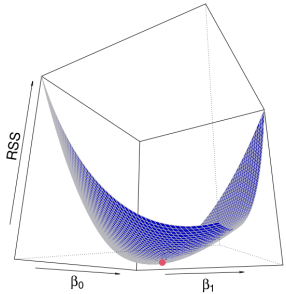
Least squares criterion: Setup

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

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

$$\frac{\partial RSS}{\partial \hat{\beta}_0} = -2 \sum_i (y_i - \hat{\beta}_0 - \hat{\beta}_1 x_i) = 0$$

$$\frac{\partial RSS}{\partial \hat{\beta}_1} = -2 \sum_i x_i (y_i - \hat{\beta}_0 - \hat{\beta}_1 x_i) = 0$$

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

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

Coding group work

Next time

Next time:

- More on simple linear regression!
- Evaluation of model etc.

CMSE381_S2026_Schedule : Sheet1

Lec #	Date		Topic	Reading	HW	Pop Quizzes	Notes
1	M	1/12	Intro / Python Review	1			
2	W	1/14	What is statistical learning	2.1		Q1	
3	F	1/16	Assessing Model Accuracy	2.2.1, 2.2.2			
	M	1/19	MLK - No Class				
4	W	1/21	Linear Regression	3.1		Q2	
5	F	1/23	More Linear Regression	3.1	HW #1 Due Sun 1/25		
6	M	1/26	Multi-linear Regression	3.2		Q3	
7	W	1/28	Probably More Linear Regression	3.3			
8	F	1/30	Last of the Linear Regression		HW #2 Due Sun 2/1		
9	M	2/2	Intro to classification, Bayes classifier, KNN classifier	2.2.3		Q4	
10	W	2/4	Logistic Regression	4.1, 4.2, 4.3.1-3			

Announcements

- Homework 1
  - Due Sun, Jan 25