CS 613 - Machine Learning

Assignment 2 - Linear Regression
Alex Lapinski
Fall 2016

10/20/2016

1 Theory

1.1 Standardize Data

$\mu = \frac{-2+-5+-3+0+-8+-2+1+5+-1+6}{10} = \frac{-9}{10} = -0.9$

$\sigma = \sqrt{\frac{(-2+0.9)^2+(-5+0.9)^2+(-3+0.9)^2+(0+0.9)^2+(-8+0.9)^2+(-2+0.9)^2+(1+0.9)^2+5+0.9)^2+(-1+0.9)^2+(6+0.9)^2}{10-1}}$

$\sigma = \sqrt{\frac{(-1.1)^2+(-4.1)^2+(-2.1)^2+(0.9)^2+(-7.1)^2+(-1.1)^2+(1.9)^2+(5.9)^2+(-0.1)^2+(6.9)^2}{9}}$

$\sigma = \sqrt{\frac{1.21+16.81+4.41+0.81+50.41+1.21+3.61+34.81+0.01+47.61}{9}}$

$\sigma = \sqrt{\frac{160.9}{9}} = \sqrt{17.87} = 4.23$

$$\begin{bmatrix} -2 & 1 \\ -5 & -4 \\ -3 & 1 \\ 0 & 3 \\ -8 & 11 \\ -2 & 5 \\ 1 & 0 \\ 5 & -1 \\ -1 & -3 \\ 6 & 1 \end{bmatrix} = \begin{bmatrix} (-2+0.9)/4.23 & 1 \\ (-5+0.9)/4.23 & -4 \\ (-3+0.9)/4.23 & 1 \\ (0+0.9)/4.23 & 3 \\ (-8+0.9)/4.23 & 11 \\ (-2+0.9)/4.23 & 5 \\ (1+0.9)/4.23 & 0 \\ (5+0.9)/4.23 & -1 \\ (-1+0.9)/4.23 & -3 \\ (6+0.9)/4.23 & 1 \end{bmatrix} = \begin{bmatrix} -0.26 & 1 \\ -0.97 & -4 \\ -0.50 & 1 \\ 0.21 & 3 \\ -1.68 & 11 \\ -0.26 & 5 \\ 0.45 & 0 \\ 1.40 & -1 \\ -0.02 & -3 \\ 1.63 & 1 \end{bmatrix}$$

1.2 Add Bias

$$\begin{bmatrix} 1 & -0.26 & 1 \\ 1 & -0.97 & -4 \\ 1 & -0.50 & 1 \\ 1 & 0.21 & 3 \\ 1 & -1.68 & 11 \\ 1 & -0.26 & 5 \\ 1 & 0.45 & 0 \\ 1 & 1.40 & -1 \\ 1 & -0.02 & -3 \\ 1 & 1.63 & 1 \end{bmatrix}$$

1.3 Find Theta

$\theta = (X^T X)^- 1 X^T Y$

$$X = \begin{bmatrix} 1 & -0.26 \\ 1 & -0.97 \\ 1 & -0.50 \\ 1 & 0.21 \\ 1 & -1.68 \\ 1 & -0.26 \\ 1 & 0.45 \\ 1 & 1.40 \\ 1 & -0.02 \\ 1 & 1.63 \end{bmatrix} Y = \begin{bmatrix} 1 \\ -4 \\ 1 \\ 3 \\ 11 \\ 5 \\ 0 \\ -1 \\ -3 \\ 1 \end{bmatrix}$$

$$X^T X = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ -0.26 & -0.97 & -0.50 & 0.21 & -1.68 & -0.26 & 0.45 & 1.40 & -0.02 & 1.63 \end{bmatrix} \begin{bmatrix} 1 & -0.26 \\ 1 & -0.97 \\ 1 & -0.50 \\ 1 & 0.21 \\ 1 & -1.68 \\ 1 & -0.26 \\ 1 & 0.45 \\ 1 & 1.40 \\ 1 & -0.02 \\ 1 & 1.63 \end{bmatrix}$$

$$X^T X = \begin{bmatrix} 10 & 0.00 \\ 0.00 & 9.01 \end{bmatrix}$$

$$(X^T X)^- 1 = \frac{1}{|X^T X|} \begin{bmatrix} 9.01 & 0 \\ 0 & 10 \end{bmatrix} = \frac{1}{10 * 9.01 - 0} \begin{bmatrix} 9.01 & 0 \\ 0 & 10 \end{bmatrix} = \begin{bmatrix} 9.01/90.1 & 0 \\ 0 & 10/90.1 \end{bmatrix} = \begin{bmatrix} 0.01 & 0 \\ 0 & 0.11 \end{bmatrix}$$

$$(X^T X)^- 1 X = \begin{bmatrix} 0.01 & 0.00 \\ 0.00 & 0.11 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ -0.26 & -0.97 & -0.50 & 0.21 & -1.68 & -0.26 & 0.45 & 1.40 & -0.02 & 1.63 \end{bmatrix}$$

$$(X^T X)^- 1 X^T = \begin{bmatrix} 0.1 & 0.1 & 0.1 & 0.1 & 0.1 & 0.1 & 0.1 & 0.1 & 0.1 & 0.1 \\ -0.03 & -0.11 & -0.06 & 0.02 & -0.19 & -0.03 & 0.05 & 0.16 & 0 & 0.18 \end{bmatrix}$$

$$(X^TX)^-1X^TY = \begin{bmatrix} 0.1 & 0.1 & 0.1 & 0.1 & 0.1 & 0.1 & 0.1 & 0.1 & 0.1 & 0.1 \\ -0.03 & -0.11 & -0.06 & 0.02 & -0.19 & -0.03 & 0.05 & 0.16 & 0 & 0.18 \end{bmatrix} \begin{bmatrix} 1 \\ -4 \\ 1 \\ 3 \\ 11 \\ 5 \\ 0 \\ -1 \\ -3 \\ 1 \end{bmatrix} = \begin{bmatrix} 1.4 & -1.75 \end{bmatrix}$$

$\theta = \begin{bmatrix} 1.4 & -1.75 \end{bmatrix}$

2 Closed Form Linear Regression

1. Final Model: $y = 3343.27586207 + 1036.63016523x_{:,1} - 295.66859639x_{:,2}$
2. Root Mean Squared Error: 653.76012597

3 S-Folds Cross-Validation

1. With S = 5 (as required), RMSE = 636.315054765
2. With S = 6 (to test other values), RMSE = 621.823391267
3. With S = 7 (to test other values), RMSE = 640.86061133

4 Locally-Weighted Linear Regression

1. With k = 1 (as required), RMSE = 242.531418672
2. With k = 2 (to test other values), RMSE = 435.493965086
3. With k = 0.5 (to test other values), RMSE = 216.575016394
4. With k = 0.75 (to test other values), RMSE = 209.853739868

5 Gradient Descent

Results when $\alpha = 0.01$

- Final model: $y = 3343.27574946 + 1036.63010327x_{:,1} + -295.66858438x_{:,2}$
- After 1712 iterations
- RMSE = 653.760066367

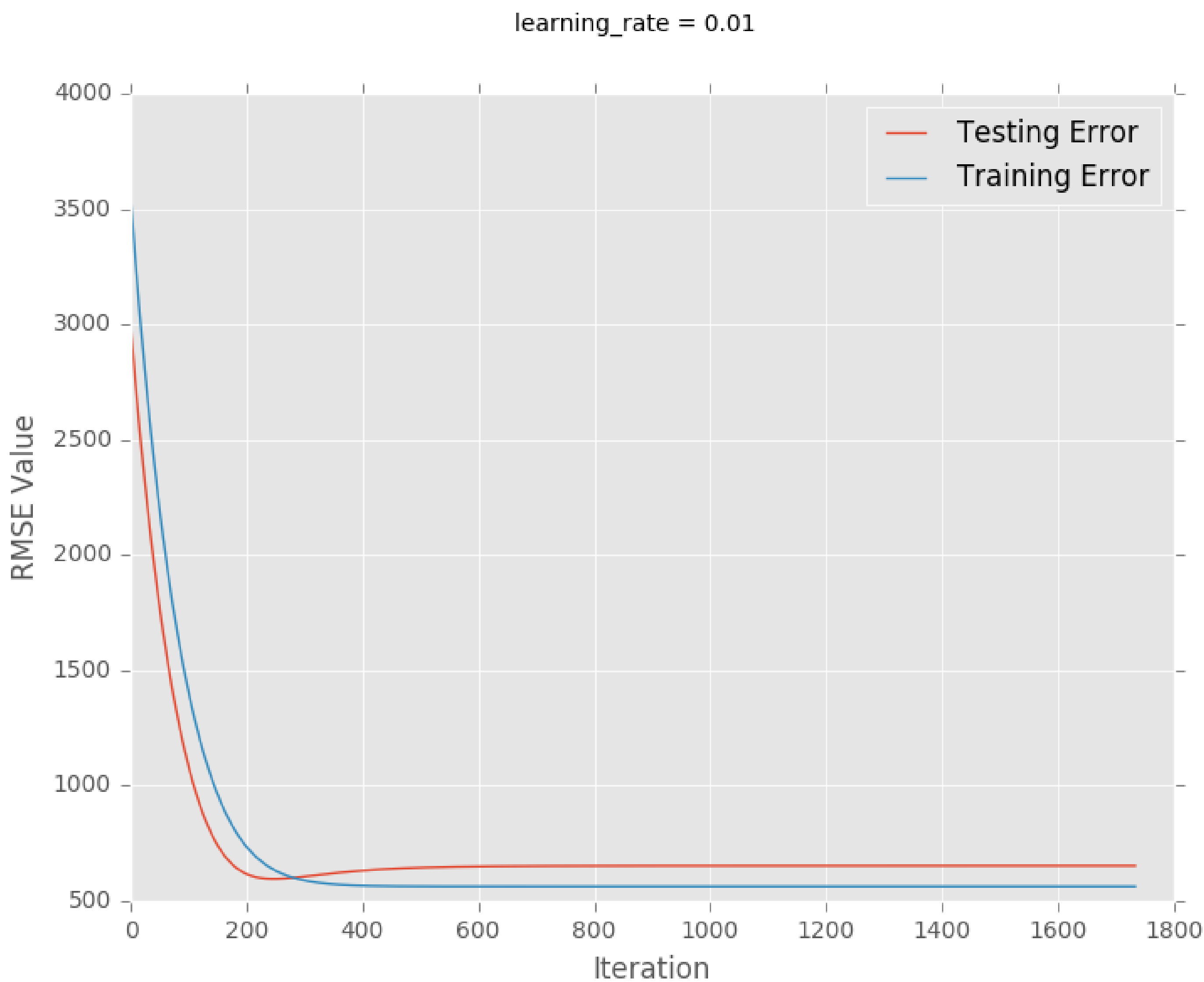

Figure 1: Gradient Descent Progress

Results when $\alpha = 0.1$

- Final model: $y = 3343.27584461 + 1036.63015469x_{:,1} + -295.66859449_{:,2}$
- After 181 iterations
- RMSE = 653.760097303