

Linear Regression II

Forrest Davis

September 16, 2025

Fall 2025 COSC410A: Applied Machine Learning
Colgate University

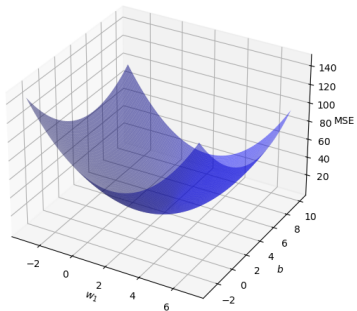


Figure 1: Mean-squared error with different parameters.

$$\partial_{\mathbf{w}} \frac{1}{m} \|\mathbf{X}\mathbf{w} - \mathbf{y}\|^2 = 0$$

Gradient Descent

$$\mathbf{X} = \begin{bmatrix} x_1^{(1)} & x_2^{(1)} & x_3^{(1)} \\ x_1^{(2)} & x_2^{(2)} & x_3^{(2)} \\ x_1^{(3)} & x_2^{(3)} & x_3^{(3)} \\ x_1^{(4)} & x_2^{(4)} & x_3^{(4)} \end{bmatrix} \text{ and } \mathbf{y} = \begin{bmatrix} y^{(1)} \\ y^{(2)} \\ y^{(3)} \\ y^{(4)} \end{bmatrix}$$

$$\mathbf{w} = \begin{bmatrix} w_1 \\ w_2 \\ w_3 \end{bmatrix} \text{ and } b$$

$$P(\mathbf{y}|\mathbf{X}) = \prod_{i=1}^m P(y^{(i)}|\mathbf{x}^{(i)})$$