

SEARCH



RESOURCES

CONCEPTS

- ✓ 13. Minimizing Error Functions
- ✓ 14. Mean vs Total Error
- ✓ 15. Mini-batch Gradient Descent
- ✓ 16. Quiz: Mini-Batch Gradient De...
- ✓ 17. Absolute Error vs Squared Err...
- ✓ 18. Linear Regression in scikit-lea...
- ✓ 19. Higher Dimensions
- ✓ 20. Multiple Linear Regression
- ✓ 21. Closed Form Solution
- ✓ 22. (Optional) Closed form Solut...
- 23. Linear Regression Warnings
- 24. Polynomial Regression
- 25. Quiz: Polynomial Regression
- 26. Regularization
- 27. Quiz: Regularization
- 28. Feature Scaling
- 29. Outro

## (Optional) Closed form solution math

In this optional section, we'll develop the math of the closed form solution, which we introduced in the last video. First, we'll do it for the 2-dimensional case, and then for the general case.

### 2-Dimensional solution

Our data will be the values  $x_1, x_2, \dots, x_m$ , and our labels will be the values  $y_1, y_2, \dots, y_m$ . Let's call our weights  $w_1$ , and  $w_2$ . Therefore, our predictions are  $\hat{y}_i = w_1 x_i + w_2$ . The mean squared error is

$$E(w_1, w_2) = \frac{1}{2m} \sum_{i=1}^m (\hat{y}_i - y_i)^2$$

We need to minimize this error function. Therefore, the factor of  $\frac{1}{m}$  can be ignored. Now, replacing the value of  $\hat{y}$ , we get

$$\begin{aligned} E(w_1, w_2) &= \sum_{i=1}^m (\hat{y}_i - y_i)^2 \\ &= \frac{1}{2} \sum_{i=1}^m (w_1 x_i + w_2 - y_i)^2 \end{aligned}$$

Now, in order to minimize this error function, we need to take the derivatives with respect to  $w_1$  and  $w_2$  and set them equal to 0.

Using the chain rule, we get

$$\begin{aligned} \frac{\partial E}{\partial w_1} &= \sum_{i=1}^m (w_1 x_i + w_2 - y_i) x_i \\ &= w_1 \sum_{i=1}^m x_i^2 + w_2 \sum_{i=1}^m x_i - \sum_{i=1}^m x_i y_i \end{aligned}$$

and

$$\begin{aligned} \frac{\partial E}{\partial w_2} &= \sum_{i=1}^m (w_1 x_i + w_2 - y_i) \\ &= w_1 \sum_{i=1}^m x_i + w_2 \sum_{i=1}^m 1 - \sum_{i=1}^m y_i \end{aligned}$$

Setting the two equations to zero gives us the following system of two equations and two variables (where the variables are  $w_1$  and  $w_2$ ).

$$\begin{aligned} w_1 \left( \sum_{i=1}^m x_i^2 \right) + w_2 \left( \sum_{i=1}^m x_i \right) &= \sum_{i=1}^m x_i y_i \\ w_1 \left( \sum_{i=1}^m x_i \right) + w_2 (m) &= \sum_{i=1}^m y_i \end{aligned}$$

We can use any method to solve 2 equations and 2 variables. For example, if we multiply the first equation by  $\sum_{i=1}^m x_i$ , the second one by  $m$ , subtract them to obtain a value for  $w_1$ , and then replace this value in the first equation, we get the following:



Mentor Help

Ask a mentor on our Q&amp;A platform