

Mock Interview: Machine Learning

What is Linear Regression?

I have a data matrix $\mathbf{X} \in n \times d$ where n is the size of the training data and d is the dimension of the data. Let's call the model parameter \mathbf{w} , and the target \mathbf{y} . What are the dimensions of \mathbf{w} and \mathbf{y} ?

Assume that $\mathbf{d} = 1$. Solve for \mathbf{w} in terms of \mathbf{y} and \mathbf{X} .

HINTS:

- What's the dimension of \mathbf{X} ? What is the dimension of \mathbf{w} ?
- Let's call them \mathbf{x} (since it is an $n \times 1$ vector) and w (since it is a scalar).
- What's the loss function you want to optimize for? (Squared loss). Write out the loss function.
- What are you optimizing it with respect to?
- $\text{minimize}_{\mathbf{w}} \|\mathbf{x}w - \mathbf{y}\|_2^2$ is the optimization problem. Now take the derivative, set it to zero and solve for w . You will arrive at $w = \frac{\mathbf{x}^T \mathbf{y}}{\mathbf{x}^T \mathbf{x}}$.

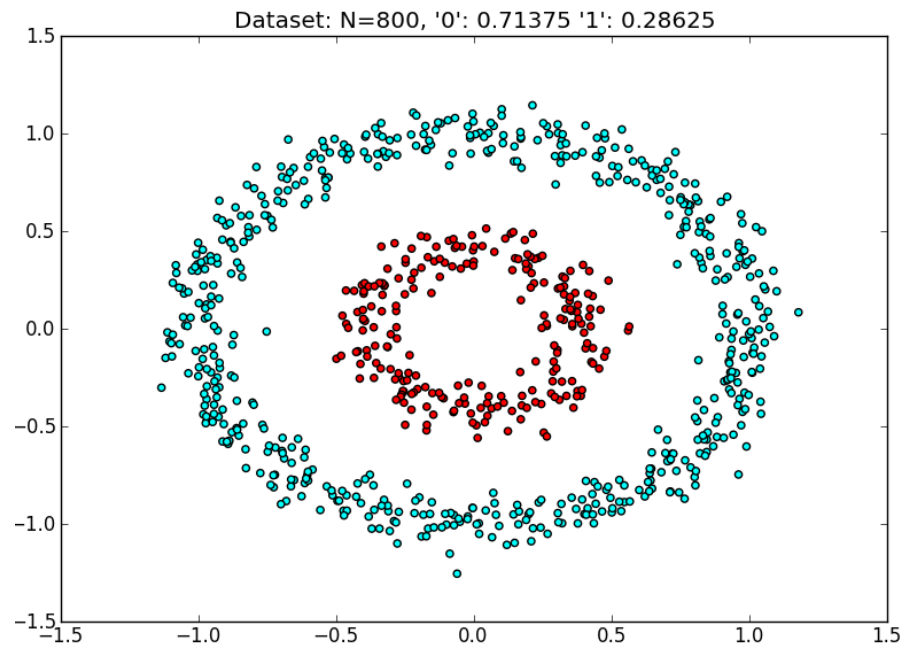
Can you generalize your solution $w = \frac{\mathbf{x}^T \mathbf{y}}{\mathbf{x}^T \mathbf{x}}$ to case where $\mathbf{d} > 1$?

$$\mathbf{w} = (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{y}.$$

Briefly ask about the matrix inversion.

Let's go back to your loss function. Is your loss function missing anything?

HINTS: Regularization. Trade of between model complexity and minimizing the loss.



Consider the plot above. Will linear regression be a good model for the data below? Explain your answer.