

How do these methods help in reducing error in regularized linear regression

$$J(\vec{w}, b) = \frac{1}{2m} \sum_{i=1}^m (f_{\vec{w}, b}(x^{(i)}) - y^{(i)})^2 + \frac{\lambda}{2m} \sum_{j=1}^n w_j^2$$

Get more training examples	→	fixes	high	variance
Try smaller sets of features $x, x^2, \cancel{x^3}, \cancel{x^4}, \cancel{x^5} \dots$	→	fixes	high	variance
Try getting additional features	→	fixes	high	bias
Try adding polynomial features $(x_1^2, x_2^2, x_1 x_2, \text{etc})$	→	fixes	high	bias
Try decreasing $\lambda$	→	fixes	high	bias
Try increasing $\lambda$	→	fixes	high	variance

→ getting more training data makes the model get the best fitting for a lot of data, thus reducing overfitting.

→ reducing the number of features prevents giving the algorithm the flexibility of overfitting the data

→ adding more features is the opposite. By doing this, the algorithm gets more information to do better and thus reducing bias.

→ similar to adding more features. For eg. if a linear function is not working very well and resulting in a high bias then adding more polynomial features might make it fit better.

Decreasing  $\lambda$  means regularization not taking place properly, thus reducing high bias

Increasing  $\lambda$  means regularization will take place and will lead to reducing high variance.