deeplearning.ai

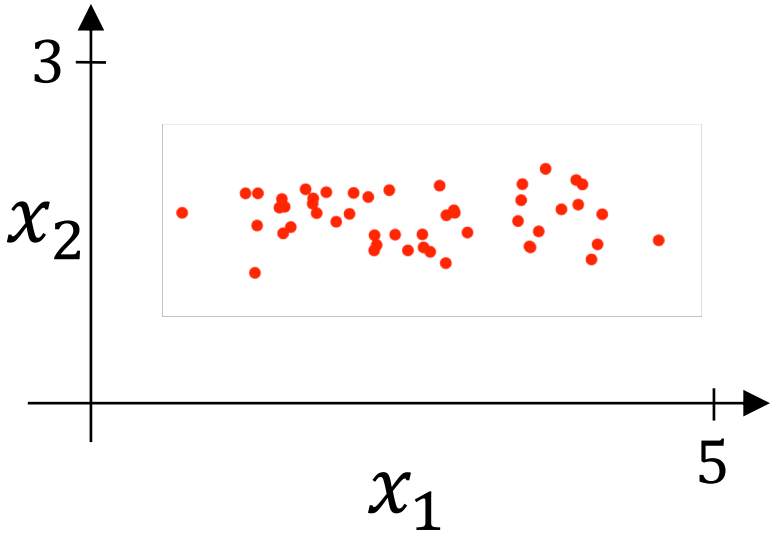
Setting up your optimization problem

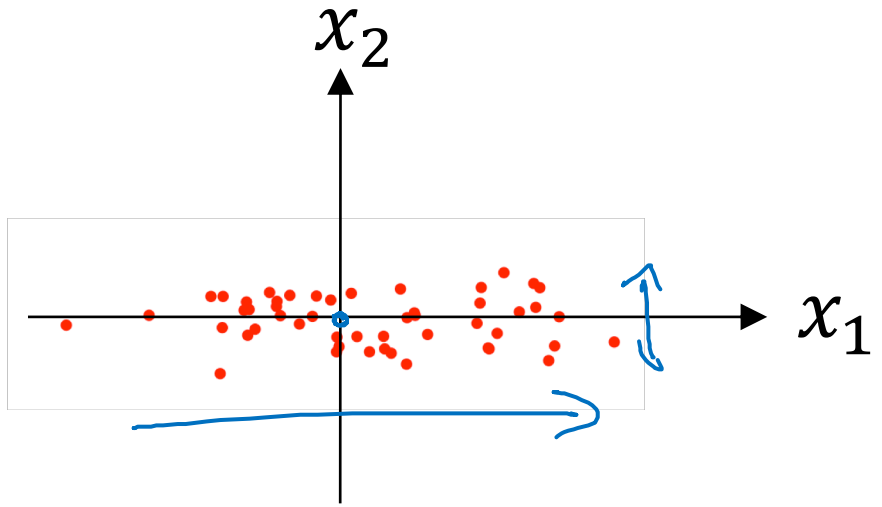
Normalizing inputs

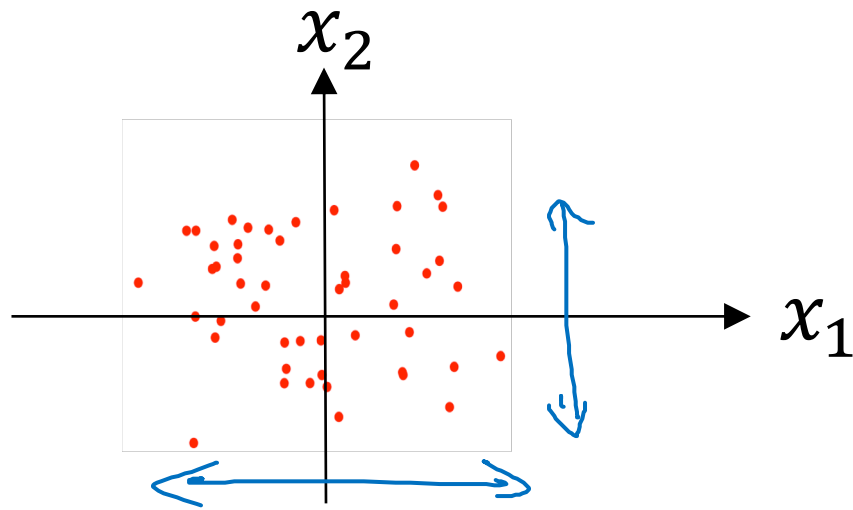
Normalizing training sets

$x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$

just guarantees that all your features on a similar scale and will usually help your learning algorithm run faster.







Subtract mean:

$\mu = \frac{1}{m} \sum_{i=1}^{m} x^{(i)}$

$x := x - \mu$

Normalize variance

$\sigma^2 = \frac{1}{m} \sum_{i=1}^{m} x^{(i)} **2$

↑ element-wise

$x /= \sigma^2$

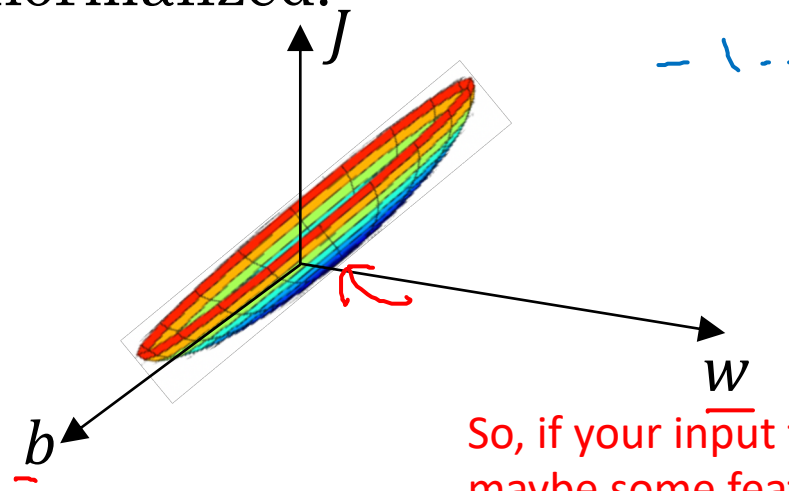
Use same μ, σ^2 to normalize test set.

Andrew Ng

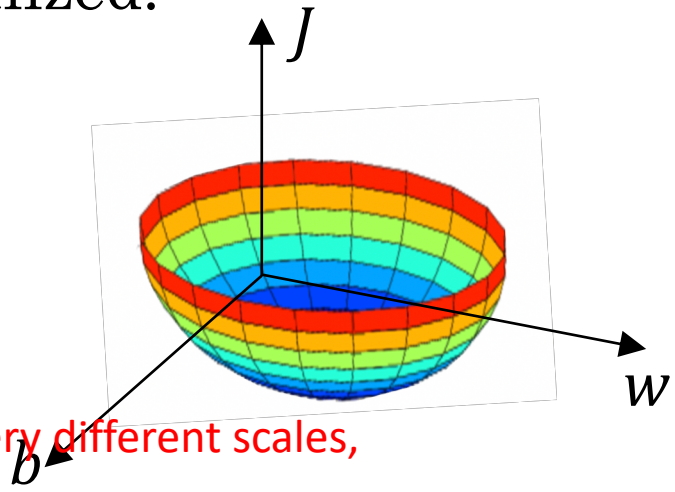
Why normalize inputs?

$$J(w,b) = \frac{1}{m}\sum_{i=1}^{m} \mathcal{L}\left(\hat{y}^{(i)}, y^{(i)}\right)$$

Unnormalized:



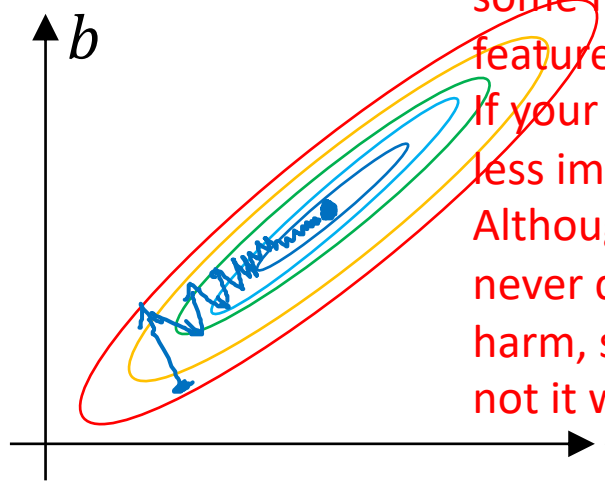
Normalized:

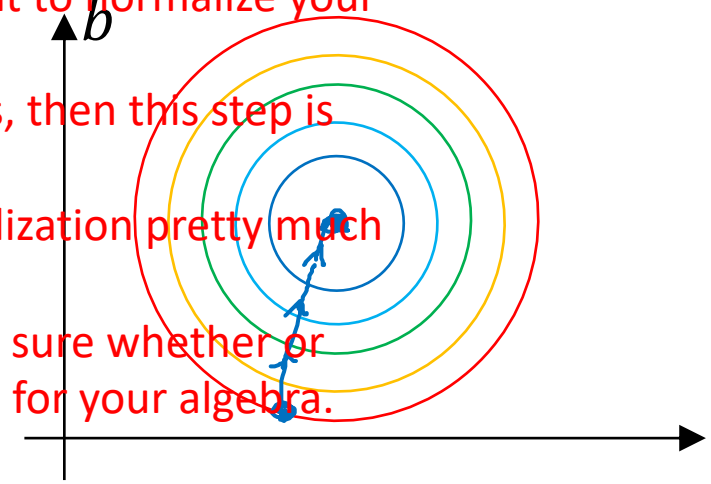


So, if your input features came from very different scales, maybe some features are from 0 to 1, some from 1 to 1,000, then it's important to normalize your features.

If your features came in on similar scales, then this step is less important.

Although performing this type of normalization pretty much never does any harm, so I'll often do it anyway if I'm not sure whether or not it will help with speeding up training for your algebra.





Andrew Ng