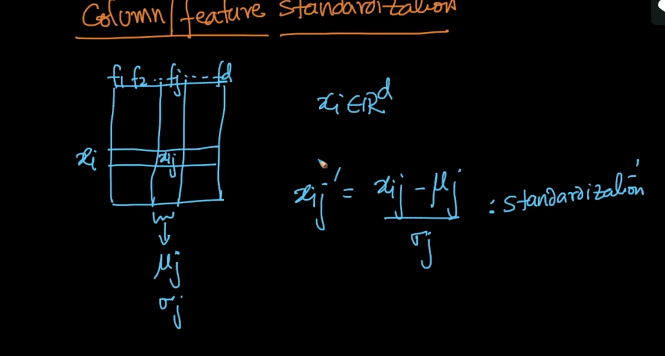
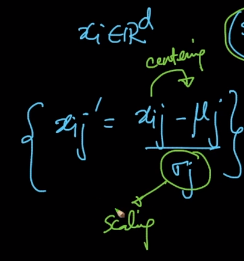
**Column Standardization:**

In standardization, let’s say we have a dataset and pick a point xij(xith point for Jth feature).  
in feature/Column standardization. Here we subtract the mean and divide it with standard deviation   
  
Even in LR, it’s mandatory to perform feature/Column standardization. The reason we perform standardization is because , if our features are in different scale, then it will impact our result. That’s the reason we perform standardization. We need to perform this before training data.  
Column stand. is also called as Mean centering and scaling.   
  
  
**Important Points**

why do we need to perform feature standardization since squashing takes care of the scale [0,1]?

 Reply   *Mar 05, 2019 06:47 AM*

Top of Form

Bottom of Form

**Applied AI Tech Admin**

Normalization is easily affected by the presence of outliers, whereas Standardization is less affected by Outliers. One of the hacks that could be used to still minimize the affect of outliers is to perform feature scaling with median in place of Mean and MAD in place of Standard deviation.  
In case, if we are using algorithms that make assumptions on the data, then it is recommended to go for Standardization.  
In case, if we are using algorithms that do not make assumptions on the data, then it is recommended to go for normalization.

 Reply *Mar 05, 2019 17:25 PM*

Top of Form

Bottom of Form

**Aaqib Naqati**

Sir, what i meant is applying sigmoid function on the equation already brings the signed distance to a consistent scale eradicating the problem of outliers partially and taking care of the scale issue...Why do we need to apply standardization after that?  
Question wasnt about the standardization and normalization!  
  
Thanks !

 Reply *Mar 06, 2019 07:46 AM*

Top of Form

Bottom of Form

**Applied\_AI**

Actually to be honest standardization is **NOT** required for LogisticRegression. Standardization is required for Linear Regression and KNN etc..Not LogisticRegression. But still, we do it. The reason is not the scale of the feature but faster convergence in SGD. All the algorithm that uses SGD or similar optimisation algorithms, for example, GLM models or Neural Network, standardization is performed for faster convergence of the model. For more reference read this https://stats.stackexchange.com/questions/48360/is-standardization-needed-before-fitting-logistic-regression

 Reply *Mar 06, 2019 10:18 AM*

Top of Form

Bottom of Form

**shaurabh**

Could you pls explain intuitively/geometrically "The reason is not the scale of the feature but faster convergence in SGD. All the algorithm that uses SGD or similar optimisation algorithms, for example, GLM models or Neural Network, standardization is performed for faster convergence of the model." as axplained above.I could understand the reference given.

 Reply *Apr 27, 2019 11:54 AM*

Top of Form

Bottom of Form

**Applied\_AI**

Most of the optimization algorithms relies on standadization of the data. Why beacause if you chalk out the math of differnt alogorithms that are optimized using SGD, you might find someting like mean of the distribution in different places of the equation which increases the complexity of the equation. After standardisation, the mean becomes zero which helps to ease out the math quite a lot and the optimization algorithm diminishes substantially. Empirically we can see the smoother convergence to local minima. For more insight check the PCA videos and you shall know it practially.

 Reply *Apr 28, 2019 00:10 AM*

Top of Form

Bottom of Form

**himshukla007**

so in LR we do standadization not for feature scaling but for faster convergence of model. right???

 Reply *Jul 05, 2019 14:54 PM*

Top of Form

Bottom of Form

**Applied\_AI**

Yes.