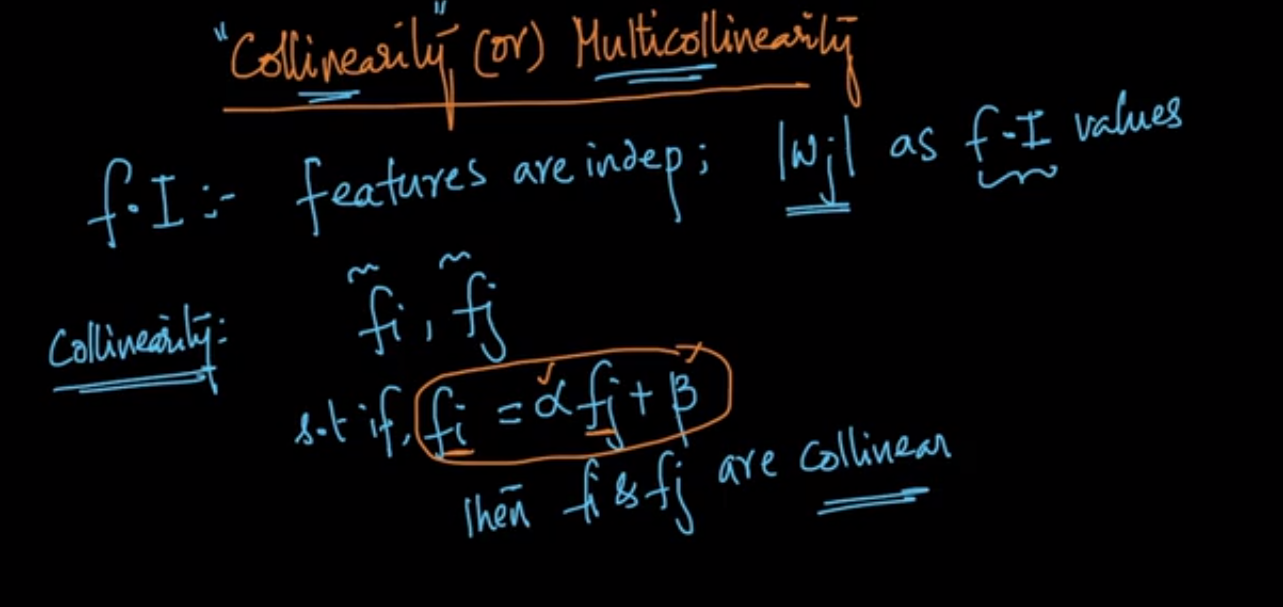
Since now we have studied about feature importance but it is only possible to check when features are independent.

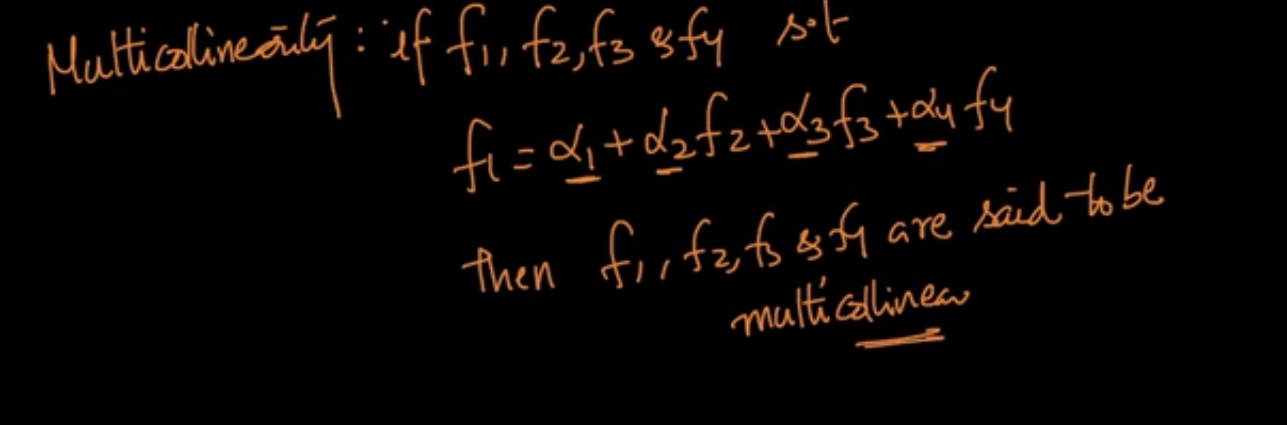
But when we check and find dependencies between features and one feature can be represented in terms of other one than it is called Collinear.



As shown in below image if we extend the phenomenon of collinearity to multi dimension it is said to be multicollinearity.

So suppose we have 4 features and one of them can be represented using other 3 features and some constant values.

Than the features are said to be multicollinear.



But the question is why we need to check if Wj is useful or not for F.I?

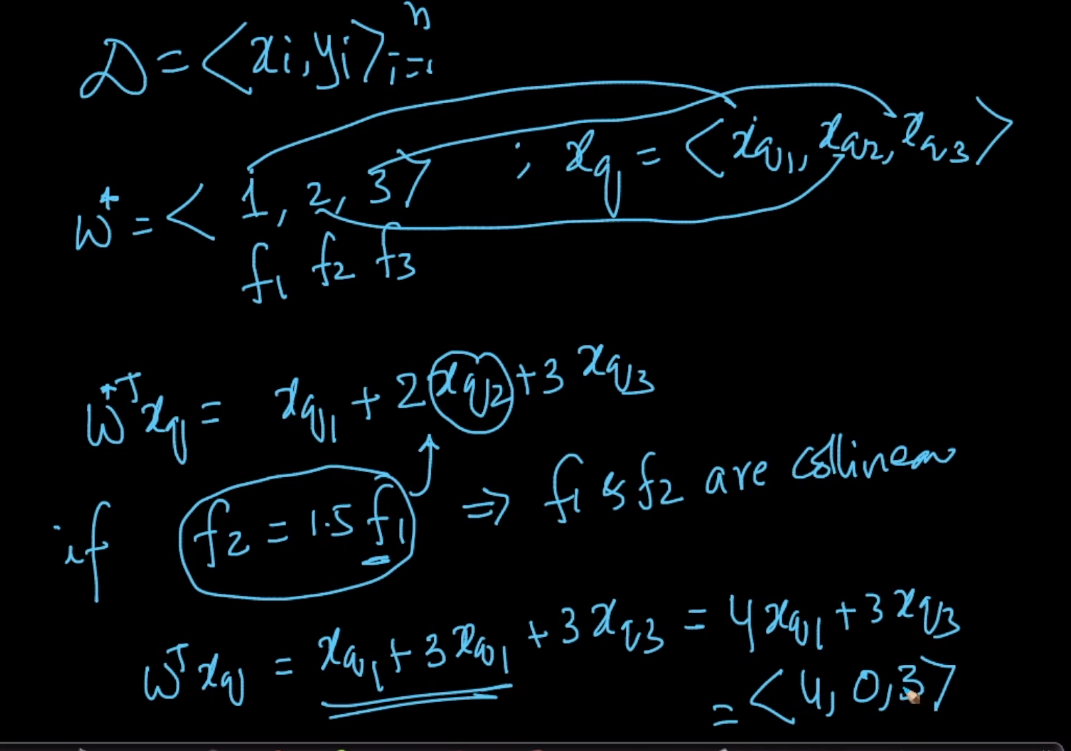
And how can we check if features are collinear?



So when we check and find that features are collinear it changes whole weight vector and hence changes feature importance.

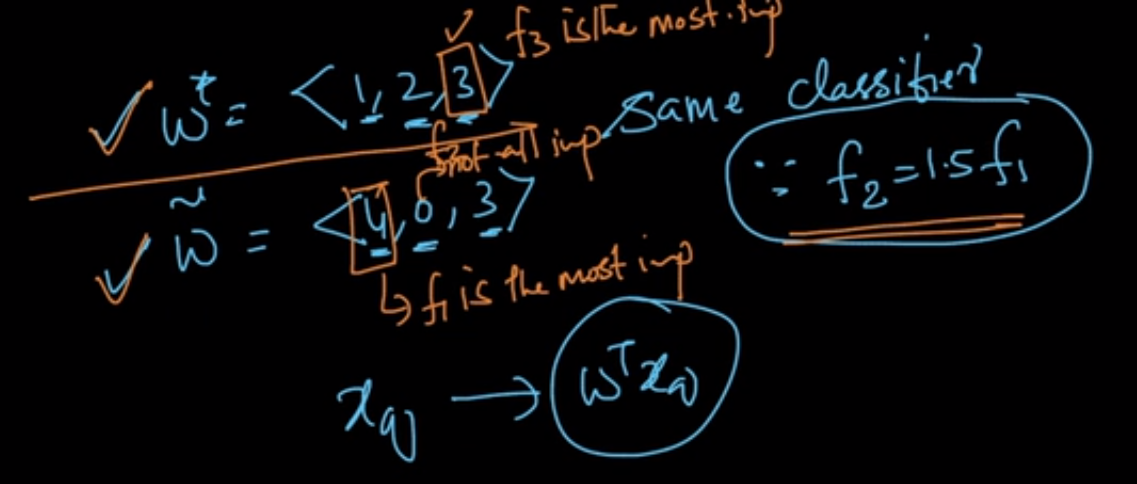
So below is the way to intuitively understand how the collinearity changes feature importance.

In below example we found that Xq2 can be represented in terms of Xq1 and so we replaces Xq2 with Xq1.

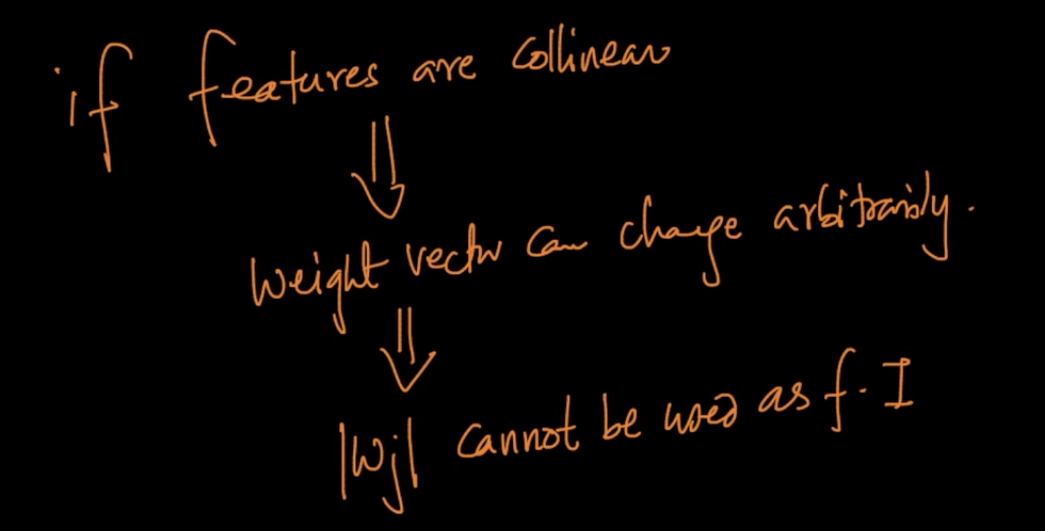


Now when we got our new vector it changes whole scenario for feature importance.

Before F3 was most important feature and after F1 is most important feature.



So below is representation how the Collinearity affects feature importance. And when Wj cannot be used to check feature importance



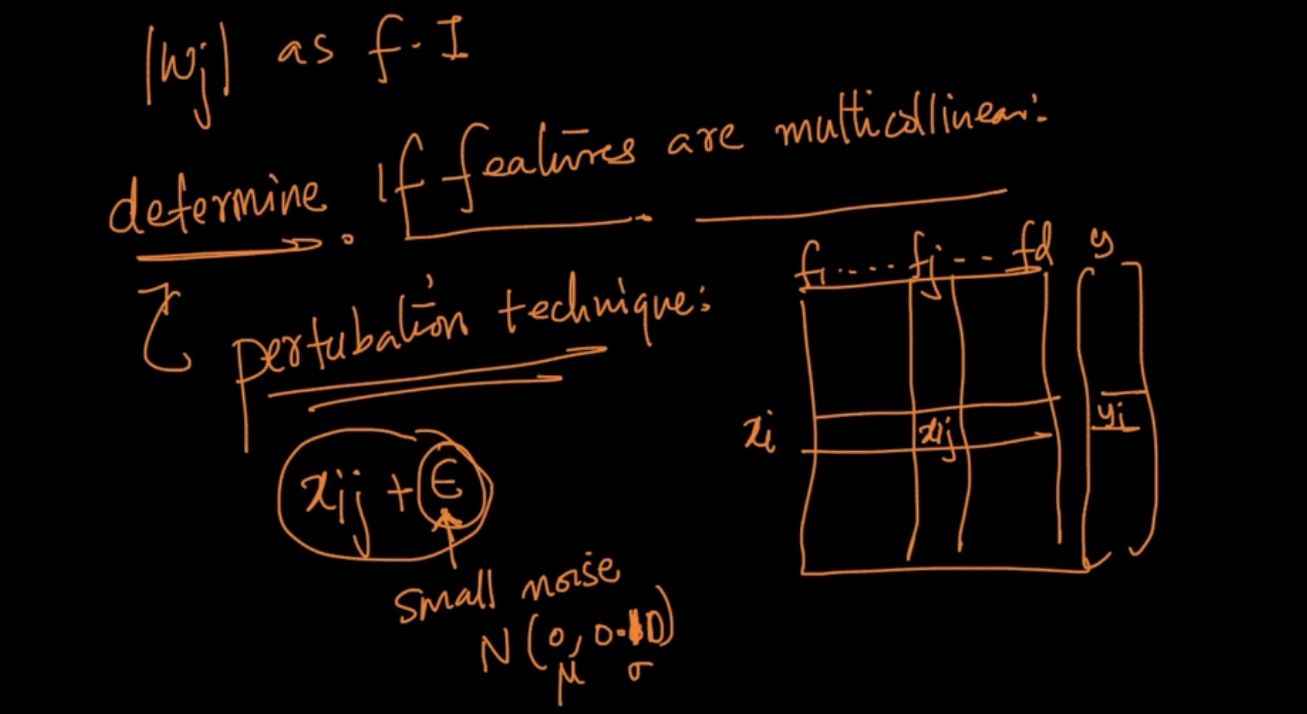
How can we check multicollinearity?

SO one of the technique is Perbutation technique.

In this technique we add some error/noise to all our Xij as shown in below image.

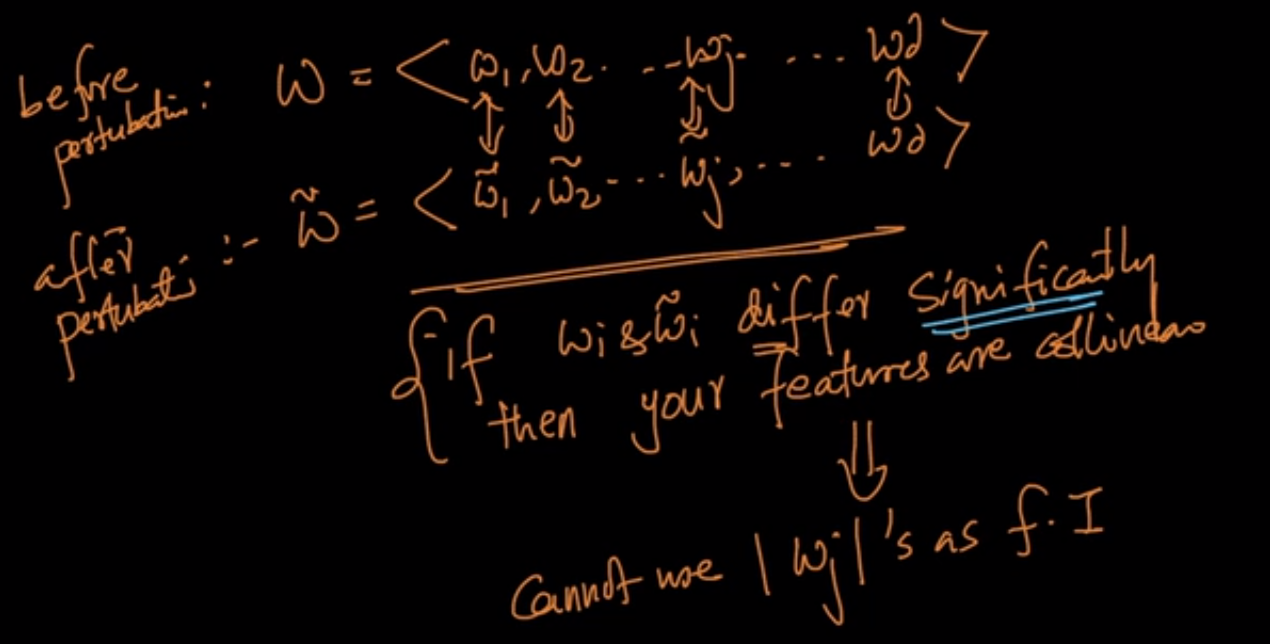
And hence will result in some changes in it std. deviation value.

Now we will get the new W.

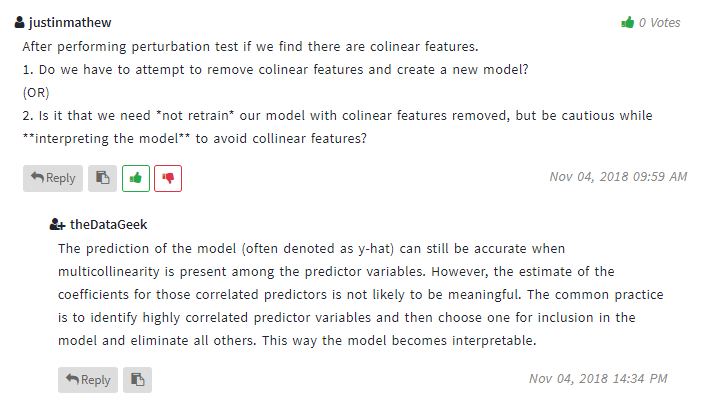


Now we will compare our new W that is W(Tilda) and previous W

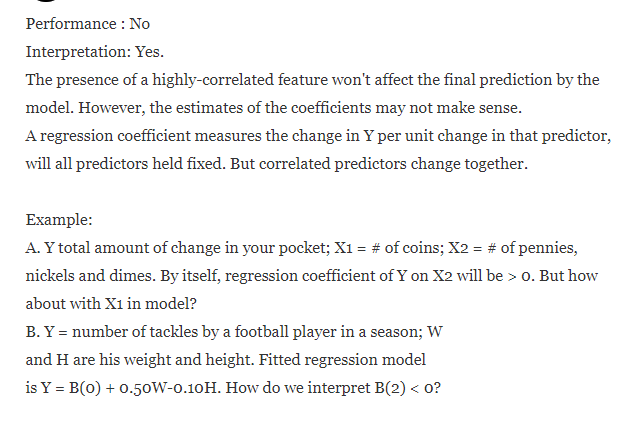
And if we find the both differ significantly then the features are collinear and hence Wj cannot be ussed as FI.

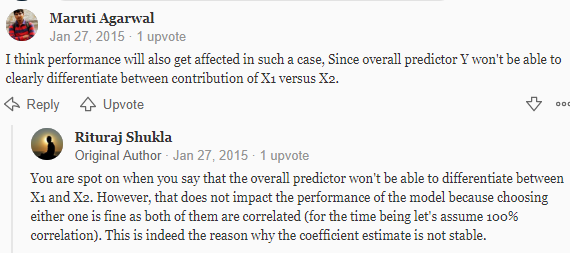


Comments:



**Does the performance of logistic regression get’s adversely affected by high correlated features.**





**What to do if multicollinearity is present**

