Multicollinearity Diagnostics Guidelines

This chart presents guidelines (or rules of thumb) for using the three diagnostic measures to determine whether there is multicollinearity. These guidelines will be helpful when you look at your output. A correlation coefficient that is near +1 or -1 indicates a high degree of linear relationship between two regressors and might suggest multicollinearity. Similarly, a VIF in excess of 10 indicates strong multicollinearity. Conventionally, condition index values between 10 and 30 suggest weak dependencies and between 30 and 100 indicate moderate dependencies among predictors. However, condition index values greater than 100 indicate strong multicollinearity.

The proportion of variation explained by the principle components is another measure that you can use, in combination with the condition index, to diagnose multicollinearity. PROC REG calculates variance proportions for each term in the model. Along with the condition index, variance proportions can be used to identify the sets of Xs (predictors) that are multicollinear. Variance proportions greater than 0.5 indicate which terms are correlated. The proportion of variation tells how much (what percentage) of the variance of the parameter estimate (the coefficient) is associated with each eigenvalue. A high proportion of variance of an independent variable coefficient reveals a strong association with the eigenvalue. Thus, if an eigenvalue is small enough (which means a large condition index) and some independent variables show a high proportion of variation with respect to the eigenvalue--that is, in excess of 0.5--we might conclude that these independent variables have significant linear dependency (or correlation).

Copyright © 2017 SAS Institute Inc., Cary, NC, USA. All rights reserved.

Close