Spring 2015 Statistics 151 (Linear Models) : Lecture Sixteen

Aditya Guntuboyina

19 March 2015

1 Detecting Nonlinearity

We will continue our discussion of regression diagnostics today by looking at detecting nonlinearity.

The standard plots for this are (a) plot the response values y against each explanatory variable values, and (b) plot the residuals $\hat{e}$ against each explanatory variable values. The problem with these plots however is that they only look at the marginal effect of the ith explanatory variable on y and ignore the presence of the other explanatory variables. To correct this, one often looks at partial regression plots (also called added variable plots) and partial residual plots.

Let $Res(y, X^{-i})$ denote the vector of residuals obtained by regressing y on all the explanatory variables *except* the ith one. Also let $Res(x_i, X^{-i})$ denote the vector of residuals obtained by regressing x_i (x_i is the column of the X-matrix corresponding to the ith explanatory variable) on all the explanatory variables *except* the ith one. In the ith partial regression plot, one plots $Res(y, X^{-i})$ against $Res(x_i, X^{-i})$. This plot therefore looks at the relationship between y and x_i in the presence of all the other explanatory variables.

A remarkable feature of the ith partial regression plot is that if one performs a regression of $Res(y, X^{-i})$ against $Res(x_i, X^{-i})$, one gets the intercept estimate to be zero and the slope estimate will exactly equal $\hat{\beta}_i$. This fact can be proved for instance using the block matrix inverse formula (see wikipedia for this formula).

The partial residual plot is another plot for viewing the relationship between y and x_i in the presence of the other variables. This simply plots $y - \sum_{j \neq i} \hat{\beta}_j x_j$ against x_i. Because $y = \hat{y} + \hat{e} = \sum_j \hat{\beta}_j x_j + \hat{e}$, one can alternately describe the partial residual plot as plotting $\hat{e} + x_i \hat{\beta}_i$ against x_i. This also has the property that if one were to do simple regression, the fitted slope with be precisely $\hat{\beta}_i$.

See the R code for the rest of this lecture.