FIND NORMALIZING TRANSFORMATIONS

The idea (very sketchy) behind the "Find Normalizing Transformations" command (on the "Transformations" menu on scatterplot matrix):

(See pp. 322 – 324 and 329 – 330 for a little more detail.)

This command will look for appropriate "scaled power transformations" – that is, functions

$$v^{(\lambda)} = \begin{cases} \dfrac{v^{\lambda} - 1}{\lambda} & if\ \lambda \neq 0 \\ \log(v) & if\ \lambda = 0 \end{cases}.$$

(Recall that these are the functions that we saw earlier on the transformation slidebars.)

One possible idea: e.g., if v = y, look for λ to minimize RSS(λ) = the RSS from regressing $y^{(\lambda)}$ on the terms.

This has a problem: The units of RSS(λ) will be different for different λ's; that is, the different RSS(λ)'s are not in the same scale.

[Note: This points out a general problem in using RSS for comparing models: It is not meaningful for comparing models when data has been transformed, since scales are different.]

A possible remedy here: Instead consider "modified scaled power transformations":

$$z^{(\lambda)} = y^{(\lambda)}[GM(y)]^{1-\lambda},$$

where

$GM(y)$ = geometric mean of $y_1, y_2, \ldots, y_n$
$= (y_1\, y_2 \ldots y_n)^{1/n}$.

Note that GM(y) has the same units as y, so $z^{(\lambda)}$ also has the same units as y.

To handle several variables simultaneously: Minimize an analogous function of the matrix of sums of squares and cross products (analogues of SXX, SXY, etc.)

Note: This is a tool to try; it is not guaranteed to work in all cases.
    e.g., it is impossible to transform an indicator variable for a categorical variable to normality.
    However, using the tool gives a better chance that regression techniques will apply.

Example: Big Mac