Diagnostics for Generalized Linear Model

Most of the same concepts we used for the general linear model have analogues in the generalized linear model. Hat values are based on the final iteration of the IWLS method for the matrix

$$\mathbf{H} = \mathbf{W}^{1/2}\mathbf{X}(\mathbf{X}'\mathbf{W}\mathbf{X})^{-1}\mathbf{X}'\mathbf{W}^{1/2}.$$

Several different types of residuals can be defined, such as the response residuals $Y_i - \widehat{\mu}_i = Y_i - g^{-1}(\widehat{\eta}_i)$, the standardized Pearson residuals,

$$R_{P_i} = \frac{Y_i - \widehat{\mu}_i}{\sqrt{\widehat{V}(Y_i|\eta_i)(1-h_i)}},$$

or the standardized deviance residuals,

$$R_{G_i} = sgn(Y_i - \widehat{\mu}_i)\sqrt{\frac{Y_i\left[g(Y_i) - g(\widehat{\mu}_i)\right] - b\left[g(Y_i)\right] + b\left[g(\widehat{\mu}_i)\right]}{a_i\widetilde{\phi}(1-h_i)}}$$

Approximations have also been developed for studentized residuals so that model refitting is not required. An approximation to Cook's distance is

$$D_i \equiv \frac{R_{P_i}^2}{(k+1)} \times \frac{h_i}{1-h_i}.$$

The text discusses approximate values for DFBETA_{ij} and DFBETAS_{ij} and an extension of added-variable plots for GLMs. Component-plus-residual plots are straightforward to extend to GLMs.