More Coping With Collinearity

Biased estimation

A different approach for collinear data is to use a biased estimation method in order to reduce the variance of the regression estimator. Some examples of biased estimation are listed below.

Ridge Regression

For standardized regression coefficients the ridge-regression estimator is

$$\mathbf{b}_d^* = (\mathbf{R}_{XX} + d\mathbf{I}_k)^{-1}\mathbf{r}_{XY},$$

for some constant $d \geq 0$. The estimator for $d = 0$ is the usual least-squares estimator for standardized coefficients. By adding a constant to the diagonal of the $\mathbf{R}_{XX} = [1/(n-1)]\mathbf{Z}_X'\mathbf{Z}_X$ matrix, we remove some of the variance-inflating problems of collinearity, at the cost of adding some bias. It can be shown that the ridge regression estimator 'shrinks' the coefficient estimate closer to zero as d increases. By rewriting the ridge regression estimator as $\mathbf{b}_d^* = \mathbf{U}\mathbf{b}^*$, it is shown in the text that the bias and variance of $\mathbf{b}_d^*$ are

$$\begin{aligned} \text{bias}(\mathbf{b}_d^*) &= (\mathbf{U} - \mathbf{I})\beta^* \text{ and} \\ V(\mathbf{b}_d^*) &= \frac{\sigma_\varepsilon^{*2}}{n-1}(\mathbf{R}_{XX} + d\mathbf{I}_k)^{-1}\mathbf{R}_{XX}(\mathbf{R}_{XX} + d\mathbf{I}_k)^{-1}. \end{aligned}$$

It has been shown by Hoerl and Kennard (1970) that it is always possible to choose a positive value for d so that the MSE of $\mathbf{b}_d^*$ is less than the MSE of $\mathbf{b}^*$. However, the challenge in practice is in finding an optimal choice of d. There is an extensive literature on methods for the choice of d, some of which are covered in Myers (1990).

Regression on Principal Components

Another biased estimation approach is to predict $\mathbf{y}$ by principal components from the set of covariates instead of the original variables. Since principal components are orthogonal, there is no variance inflation problem with these new variables (the principal components). However, since we usually discard some of the principal components we introduce bias in the estimator of $\boldsymbol{\beta}$ as we also reduce the variance of the estimator. More details can be found in Myers (1990).

Use of Prior Information About the Regression Coefficients

As discussed in the text, use of prior information may help with collinearity, from simple respecification of the model to a more sophisticated Bayesian analysis.