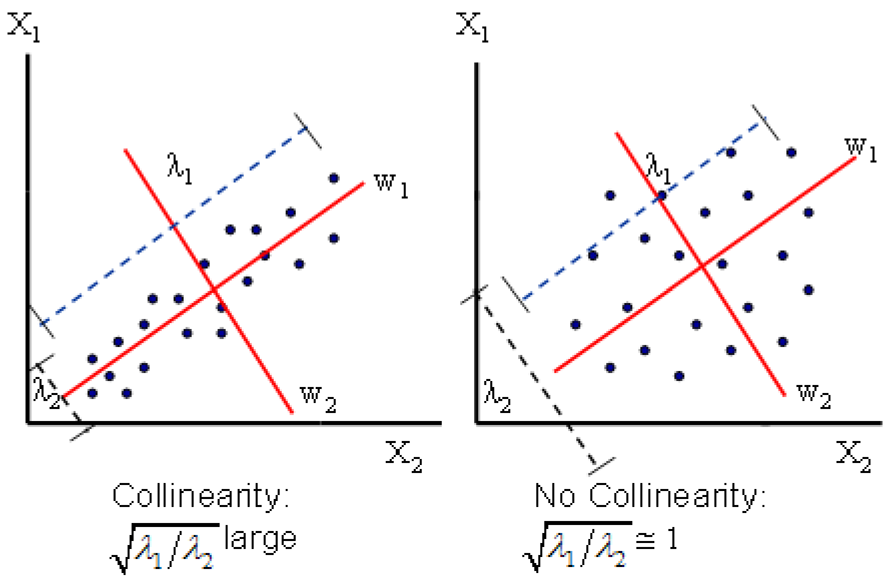
Print

Eigenanalysis and Multicollinearity

Collinearity occurs when linear combinations of some of the columns in the ***X*** matrix equal zero, or nearly zero. Geometrically, this occurs when at least one dimension of the ***X***-space has very little dispersion (shown in the left graph below).



When an independent variable has limited dispersion, its column in the ***X*** matrix is almost a multiple of a vector of ones, with the result that the variable is nearly collinear with the column for the intercept. The presence of collinearity is detected by *singular decomposition* of ***X*** or the *eigenanalysis* of ***X'X***.

A value λ is called the eigenvalue of the SSCP matrix ***X'X*** if there is a nonzero vector ***z*** such that (***X'X***)***z*** = λ***z***. The nonzero vector ***z*** is called the *eigenvector*. Eigenvalues and eigenvectors of the SSCP matrix ***X'X*** are closely related to the principal components of the matrix ***X***. Principal components (***W*** = ***XZ***) are linear combinations of the independent variables such that

- the principal components $\boldsymbol{w_i}$ are uncorrelated. In other words, they are all pairwise orthogonal.
- the first principal component has the largest variance of any linear function of the original variables (subject to a scale constant). The second principle component should explain most of the remaining variance that the previous component could not explain, and so on.

Principal components are obtained by computing the eigenvalues and eigenvectors of ***X'X***. The eigenvalues are the variances of the components, and the eigenvectors are the coefficients, $\boldsymbol{z_{ij}}$, of the linear equations that relate the principal components to the original variables, such that ***W*** = ***XZ***.

Principal components with very small variances (which correspond to large condition indices) are of interest in identifying sources of multicollinearity.

Close