Re: transformation of regressors to remove collinearity

Source: http://sci.tech−archive.net/Archive/sci.stat.math/2008−03/msg00088.html

* *From*: sangdonlee@xxxxxxxxx
* *Date*: Tue, 11 Mar 2008 07:42:43 −0700 (PDT)

On Mar 11, 10:08 am, papu <prac...@xxxxxxxxx> wrote:

> Actually because of model governance rules we are
> required to use a certain minimum number of variables. The VIFs need
> to be less than 3 and correlation coeffcient should be less than 0.7.
> I have some predictors that have correlation coefficient 0.75 or
> so..they are not highly correlated but not desirable either..

If I were you, I would do the following (to make the long story short)

1. The first way (possibly the best in my opinion) is to combine/
delete the correlated
variables based on previous results or theory. For example, the track
width
and wheelbase of a vehicle can be combined into planview.
(planview=trackwidth*wheelbase). This approach may result in a model
with very few predictors.

> Hence I
> thought why not separate the correlated and uncorrelated parts between
> the predictors using geometric decomposition or projections and use
> those transformation as predictors. This way the collinearity is zero.
> But the question remains how to decompose these predictors and will
> the predictive power of the transformed predictors remain the same as
> a group.− Hide quoted text −

2. Apply PCA (principal component analysis) and then PCR (principal
component
regression) or PLS (partial least square). All of them are geometric
projections/decomposition of the predictors. They are "mere"
mathmatical transformation of predictors and they provide
correlational structure among predictors (hopely).

By the way, if your data are from observational studies (not from
"controlled" experiments where the predictors are orthogonal, samples

are balanced, confounding variables are controlled), the
interpretation should be very careful (almost futile
in my experience). As you already know, prediction and interpretation
are completely seperate issues.

Hope this helps.

Sangdon Lee, Ph.D.,
GM Tech. Center

.