Subject: regression optimization
Posted by Klemen on Thu, 17 Jun 2010 07:35:43 GMT
View Forum Message <> Reply to Message

Hello all,

I have a variable to be explained with 12 attributes (at the moment all of them being arrays of 300 by 100, but in the future it can be even larger than 1000 by 1000 array). I would like to preform a moving window analysis. For each window I want to estimate the optimal multiple linear regression parameters.

Optimal means that it is not necessary that all of 12 variables are used (some of them are correlated). So I want to get out an equation that has 2,3,4... parameters and provide the best statistics.

The problem is also that not all the values within a moving window can be used - some pixels contain no data. Defining NaN for these values and using the code similar to the one written by Bringfried Stecklum ( http://groups.google.com/group/comp.lang.idl-pvwave/browse_t hread/ thread/17613c70b78f1ac4/6891d260db6c7c93? lnk=gst&q=regression#6891d260db6c7c93), I can test which of attributes might be significant for the regression.

The question is, how to proceed. From those attributes that I know that they are correlated among each other, I want to use just the one that explains the most variability. I can somehow imagine to select the final 2-4 attributes that should be used for multiple regression without using any FOR loops. But how do I do the final step - estimation of multiple regression parameters without using any loops? Any idea?

Thank you!
Klemen