Residuals of a linear model for correlated data with measurement errors

Ants Kaasik

University of Tartu, Estonia, email: ants.kaasik@ut.ee

Keywords: correlated observations, errors-in-variables model, phylogenetic analysis.

We consider the case, where a linear model has been set up for variables Y and $X_1, \ldots, X_n$. All the variables in the model are observed with error. When calculating the residuals for such a model, errors in the X variables will have important consequences that cannot be ignored when performing further analyses with these residuals. I show that this can be thought of as measurement errors carrying over to the residuals and the process is analyzed in detail.

Such a model is quite typical in phylogenetic analyses (analysis where traits measured for a species correspond to a sample element and the sample is treated as correlated data because of the shared evolution of the species) in a situation where the relation between two traits is sought and one (or both of them) need to be corrected for the value(s) of some other trait(s). See e.g. [1] and [2] for examples.

References

[1] Revell, L. J. (2009). Size-correction and principal components for interspecific comparative studies. *Evolution* **63**(12), 3258–3268.

[2] Ives, A. R., Midford, P. E., Garland, T. Jr. (2007). Within-species variation and measurement error in phylogenetic comparative methods. *Systematic Biology* **56**(2), 252–270.