Envelopes: Methods for Improving Efficiency in Multivariate Statistics

Essentially a form of targeted dimension reduction, an envelopes is a construct for increasing efficiency in multivariate statistics without altering the traditional goals, sometimes producing gains equivalent to taking thousands of additional observations. This is made possible by recognizing that the data may contain unanticipated variation that is effectively immaterial to estimation or prediction. This notion leads to the central construct – an envelope – for enveloping the material variation and thereby reducing estimative variation and improving inference.

Envelopes link with some standard multivariate methodology. For instance, partial least squares regression depends fundamentally on an envelope at the population level, which opens the door to pursuing envelope estimators that can significantly improve upon partial least squares predictions.

We will begin with an intuitive introduction to envelopes and then describe some of their inner workings. This will be followed by a discussion of envelopes in multivariate linear regression. We will also describe briefly how to extend the scope of envelope methods well beyond linear models. The discussion will include several small examples for illustration. Emphasis will be placed on concepts and their potential impact on data analysis.