Interfacing Symbolic Computation and Numerical Calculation

Alessandra R. Brazzale
Department of Mathematics
Swiss Federal Institute of Technology
1015 Lausanne, Switzerland
alessandra.brazzale@epfl.ch

Abstract

The fate of a new theory is to a large extent determined by the time intervening its development and the availability of a set of tools that supports practical applications. This last step is crucial if the methods developed are to become widely accepted and used. During the last two decades, considerable progress has been made in small-sample asymptotics leading to a well-established theory on higher-order approximations. Nevertheless, these methods are still under-used in practice, a main reason being the forbidding mathematics involved in their derivation. But the undoubtedly largest obstacle to a wider use is the lack of software.

What makes the implementation of small-sample asymptotics so challenging is the trade-off between algebraic complexity and numerical efficiency, which affects crucially the choice of programming environment in which to implement the methods. The first task is best accomplished through symbolic computation. Computer algebra systems such as Maple, Mathematica and Reduce allow all quantities required to be computed in a rather straightforward way, and only need a limited amount of input. However, real data analysis involves extensive numerical computation. The implementation of the methods of interest in a numerical computing environment enlarges the potential user base by a considerable factor. It entails only one drawback: for each class of models considered, one has to supply all necessary algebraic input.

At the time of writing, there exists no environment able to accomplish algebraic and numerical tasks in an efficient way. One alternative is to use the interfaces available between S-Plus and computer algebra systems such as Mathematica or Maple. Ideally, the analysis is performed in S-Plus, but all necessary algebraic output is produced by symbolic computation and passed back to S-Plus. Until now, this possibility has only been used in a pseudo-automatic form, where the user is responsible for the interaction between the two environments. Our aim is to investigate if and to what extent this step might be fully automated.

RÉSUMÉ

Différentes méthodes asymptotiques d'ordre élevé ont été développées durant ces quinze dernières années furnissant des approximations très précises dans le cadre de l'inférence paramétrique. Malgré une théorie bien développée, ces méthodes sont peu utilisées dans la pratique statistique. Elles sont souvent considérées comme trop difficile à calculer, mais l'obstacle majeur provient principalement du manque d'automatisation.

Les implémentations déjà existantes utilisent des logiciels de calcul formel tels que Maple, Mathematica ou Reduce pour la dérivation d'expressions algébriques, mais font recours à des environnements statistiques pour tous les calculs numériques. Jusqu'à maintenant l'interaction entre ces deux familles de logiciels a été gérée de façon pseudo-automatique. Nous aimerions examiner dans quelle mesure et à quel point ce dernier pas peut être automatisé.