Symbolic modelling of drug trial data

Rikard Nordgren

September 18, 2019

1 Background

Pharmacometrics is a field within the pharmaceutical sciences. The pharmacometrics research group at Uppsala University develop and use mathematical models to understand drug and disease mechanisms, and to optimise drug development and therapy. (For more information see our webpage at `http://farmbio.uu.se/forskning/researchgroups/farmakometri/`). The research group has a small team of developers that develop the computational methods that are used by the researchers. All tools are made available as open source and is widely used by academia and pharmaceutical industry across the globe.

2 Introduction

We mainly work with nonlinear mixed effects models to describe for example how the blood concetration of a drug varies over time or the effect of a drug on a disease. The models we typically work with are represented as source code or "almost" source code that is used to fit the model parameters, predict and to simulate new data. Here is an example of a model for pharmacokinetics of the drug phenobarbital, which is used to treat epilepsy of newborns.

```
;; 1. Based on: 5
$PROBLEM    PHENOBARB SIMPLE MODEL
$DATA      pheno.dta IGNORE=@
$INPUT      ID TIME AMT WGT APGR DV FA1 FA2
$SUBROUTINE ADVAN1 TRANS2
$PK
      TVCL=THETA(1)*WGT
      TVV=THETA(2)*WGT
IF(APGR.LT.5) TVV=TVV*(1+THETA(3))
      CL=TVCL*EXP(ETA(1))
      V=TVV*EXP(ETA(2))
```

```
      S1=V

$ERROR
      W=F
      Y=F+W*EPS(1)
      IPRED=F         ;  individual-specific prediction
      IRES=DV-IPRED   ;  individual-specific residual
      IWRES=IRES/W    ;  individual-specific weighted residual

$THETA (0,0.00469307) ; CL
$THETA  (0,1.00916) ; V
$THETA  (-.99,.1)
$OMEGA  DIAGONAL(2)
 0.0309626  ;       IVCL
 0.031128  ;        IVV
$SIGMA  0.0130865
$ESTIMATION METHOD=1 INTERACTION
$COVARIANCE UNCONDITIONAL
$TABLE      ID TIME DV AMT WGT APGR IPRED PRED RES CWRES NPDE NOAPPEND
            NOPRINT ONEHEADER FILE=sdtab1
```

All these operations are purely numerical. In some cases it is of great benefit to be able to use the underlying symbolic (or mathematical) relationships of the model code to simplify certain types of operations. Using symbolics in statistical modelling is of increasing interest, see for example `http://people.cs.uchicago.edu/~mrocklin/tempspace/sympystats.pdf`. Currently we see use cases for this in methods that search for significant covariates in models. For example the body weight might affect the clearance of a drug from the blood stream.

3 The project

This project would be about investigating the use of symbolic statistics for pharmaceutical modelling using the python sympy package. You will explore the sympy package and by hand translate a model from a fortran like notation into a sympy set of equations. The next step would be to use this symbolic form of the model to test different operations. Depending on the time available one or more of the following could be tested:

- Calculate values of derived covariates
- Symbolic calculation of derivatives
- Evaluation of model predictions

- Symbolic simulation
- Determine relationships between different parameters

4 Goal

The goal is to re-implement a real patient data pharmacometric model in python using sympy and to do one or more useful things using the symbolic model. The students will learn a bit about pharmacometric models, numeric vs symbolic computing and the python sympy package.