Email contact: joakim.nyberg@farmbio.uu.se

UPPSALA UNIVERSITET

Second order Taylor expansion of likelihood-based models for fast covariate and random effect model building

Yasunori Aoki[*], Joakim Nyberg[+] and Mats O. Karlsson[+]

[*]National Institute of Informatics, Japan

[+]Department of Pharmaceutical Biosciences, Uppsala University, Sweden.

Background and Objectives

Linearization of models with respect to the random effect variable has been proved to work well for continuous models and the FOCE(I) approximation [1]. The approach enables fast and sometimes more stable estimation of covariate and random effects models, once as a structural model has been identified. However, for likelihood-based models, first order mean and variance linearization is not feasible since the data are not implicitly assumed normally distributed. Instead, a second order Taylor expansion is needed.

The aim of this work is to implement a second order Taylor expansion to improve speed and stability for likelihood-based models.

Methods and Materials

The -2 log likelihood (-2LL) of the individual observation is approximated with second order Taylor series expansion, i.e.

$$l_{ij}(\hat{\boldsymbol{\eta}},\hat{\boldsymbol{\theta}}) \approx l_{ij}(\hat{\boldsymbol{\eta}}_{oj},\hat{\boldsymbol{\theta}}) + \nabla l_{ij}(\hat{\boldsymbol{\eta}}_{oj},\hat{\boldsymbol{\theta}})(\hat{\boldsymbol{\eta}}-\hat{\boldsymbol{\eta}}_{oj}) + \frac{1}{2}(\hat{\boldsymbol{\eta}}-\hat{\boldsymbol{\eta}}_{oj})^{\mathrm{T}} H_{ij}(\hat{\boldsymbol{\eta}}_{oj},\hat{\boldsymbol{\theta}})(\hat{\boldsymbol{\eta}}-\hat{\boldsymbol{\eta}}_{oj}),$$

where l_{ij} is the -2LL of the j^{th} observation of the i^{th} individual, ∇l_{ij} and H_{ij} are gradient vector and Hessian matrix with respect to random variable vector $\boldsymbol{\eta}$. $\boldsymbol{\theta}$ is a vector of fixed effects, $\hat{\boldsymbol{\eta}}_{oj}$ is the empirical Bayes estimate of $\hat{\boldsymbol{\eta}}$, see fig 1. The model is modified by substituting $\hat{\boldsymbol{\eta}} = \mathbf{g}(\boldsymbol{\eta},\boldsymbol{\theta})$. The function $\mathbf{g}$ can for example be used to introduce covariates or covariances. The key advantage of this approximation is that it does not require any additional computation of the structural model predictions, hence an extensive explorations of various functions $\mathbf{g}$ become feasible even for models with long computation time.

We now modify the model by letting $\hat{\boldsymbol{\eta}} = \mathbf{g}(\boldsymbol{\eta},\boldsymbol{\theta})$. This gives the following approximation

$$l_{ij}(\mathbf{g}(\boldsymbol{\eta},\boldsymbol{\theta}),\hat{\boldsymbol{\theta}}) \approx l_{ij}(\hat{\boldsymbol{\eta}}_{oj},\hat{\boldsymbol{\theta}}) + \nabla l_{ij}(\hat{\boldsymbol{\eta}}_{oj},\hat{\boldsymbol{\theta}})(\mathbf{g}(\boldsymbol{\eta},\boldsymbol{\theta})-\hat{\boldsymbol{\eta}}_{oj}) + \frac{1}{2}(\mathbf{g}(\boldsymbol{\eta},\boldsymbol{\theta})-\hat{\boldsymbol{\eta}}_{oj})^{\mathrm{T}} H_{ij}(\hat{\boldsymbol{\eta}}_{oj},\hat{\boldsymbol{\theta}})(\mathbf{g}(\boldsymbol{\eta},\boldsymbol{\theta})-\hat{\boldsymbol{\eta}}_{oj}),$$

Pseudo code for Base model

```
$ABBREVIATED COMRES=6

$ERROR
; Previous code that defines Y
TMP2=Y
 "LAST
 "  IF (EVID.EQ.0) THEN !Only obs records
 "   COM(1)=G(1,1)  !dY_dETA1
 "   COM(2)=G(2,1)  !dY_dETA2
 "   COM(3)=G(1,2)  !dY2_d2ETA1
 "   COM(4)=G(2,2)  !dY2_dETA2_dETA1
 "   COM(5)=G(2,3)  !dY2_d2ETA2
 "   COM(6)=TMP2 !The log likelihood
 "  ENDIF

; Save output with high precision
$TABLE FORMAT=s1PE30.15 ID DV ;Covariates;
EVID COM(6) ETA1 ETA2 COM(1) COM(2) COM(3)
COM(4) COM(5) NOPRINT ONEHEADER FILE=
2nd_order.dta
```

Figure 1. Pseudo code for additions needed in NONMEM to implement the 2nd order Taylor expansion of the model. This example assumes two random effects (ETAs).

Pseudo code for Taylor expanded model

```
$INPUT ID DV ;Covariates; EVID MYY EBE1 EBE2 DYDETA1
       DYDETA2 D2YDETA11 D2YDETA12 D2YDETA22

$DATA 2nd_order.dta IGNORE=@  IGNORE=(EVID.GT.0)

$PRED
  Y=0 ;No contribution to -2LL for other rows
  DELTA_ETA_1=(ETA(1)-EBE1)
  DELTA_ETA_2=(ETA(2)-EBE2)

  ; 0 and 1st order term
  TMP1=MYY+DYDETA1*DELTA_ETA_1+ DYDETA2*DELTA_ETA_2
  ; 2nd order terms
  TMP2=TMP1+1/2*(DELTA_ETA_1**2*D2YDETA11+
      DELTA_ETA_2**2*D2YDETA22)

  Y=TMP2+DELTA_ETA_1*DELTA_ETA_2*D2YDETA12

$THETA ;All structural THETAS removed expect e.g.
new covariates to test
$OMEGA VAR_ETA1 ; Estimate all OMEGAS
$OMEGA VAR_ETA2
```

This procedure can be performed automatically in PsN (ver. 4.8.0 and above), e.g.,

linearize run1.mod -second_order

Results

As can be seen from the above workflow, once the base model runs and all the necessary variables for the Taylor expansion are obtained, the covariate model building procedures can be done using Taylor expanded model (i.e., without evaluating the structural model). Once the final covariate model is built using the Taylor expanded model, as can be seen in this example, the covariate model structure and its parameters for the base model can be derived from the Taylor expanded model.

This workflow was tested using the following different models:

- minimal continuous-time Markov model for the Likert pain score (mCTMM) [2]
- first-order Markov model for Fatigue experienced by Sunitinib-treated patients (MARKOV) [3]
- bounded integer model for ADAS-cog score (BI) [4,5]

For mCTMM we have manually added covariate, and for MARKOV and BI, we have conducted extensive covariate search using the SCM method implemented in PsN [6,7].

Model	OFV of model w/o covariates	OFV of final model	Computation time including covariate model building*
mCTMM	60824.17	59994.69	554 sec
mCTMM using Taylor approximation	60826.96	60028.98	219 sec
MARKOV	6765.61	6765.61	9513 sec
MARKOV using Taylor approximation	6763.99	6763.99	697 sec + 24 sec**
BI	28122.99	27569.13	14100 sec
BI using Taylor approximation	28122.99	27569.10	168 sec + 357 sec**

Table I. The Objective function Value (OFV) and computation time.
*Computation time was measured with a MAXEVAL=0 run in NONMEM 7.4 with GCC version 6.3.0 on Intel(R) Xeon(R) E5645 @2.40GHz
**For the computation time for derivatives.

As can be seen from the above Table 1, some approximation error can be observed; however, a significant speed up in the computation was achieved.

For MARKOV, the SCM did not find any statistically significant covariates. For BI, the baseline Minimum Mental State Examination score was significant covariates for both baseline ADAS-cog score and disease progression. Both of these results were consistent between the original model and approximated model while it was significantly faster to run SCM using approximated model.

Final model from SCM of Taylor expanded model

```
$PROBLEM    Second order approximation
$INPUT      ID DV TIME BMMS AGE=DROP SEX=DROP EDU=DROP MYY EBE1 EBE2
EBE3 DYDETA1 DYDETA2 DYDETA3 D2YDETA11 D2YDETA12 D2YDETA22
D2YDETA13 D2YDETA23 D2YDETA33
$DATA      2nd_order.dta IGNORE=@
$PRED
;;; SLOPEBMMS-DEFINITION START
SLOPEBMMS = ( 1 + THETA(4)*(BMMS - 27))
;;; SLOPEBMMS-DEFINITION END
;;; SLOPE-RELATION START
SLOPECOV=SLOPEBMMS
;;; SLOPE-RELATION END

;;; BASEBMMS-DEFINITION START
BASEBMMS = ( 1 + THETA(3)*(BMMS - 27))
;;; BASEBMMS-DEFINITION END
;;; BASE-RELATION START
BASECOV=BASEBMMS
;;; BASE-RELATION END

TVBASE=THETA(1)
TVBASE = BASECOV*TVBASE
TVSLOPE=THETA(2)
TVSLOPE = SLOPECOV*TVSLOPE

DELTA_ETA_1 = (TVBASE-THETA(1)+ETA(1) - EBE1)
DELTA_ETA_2 = (TVSLOPE-THETA(2)+ETA(2) - EBE2)
TERM_O1_1 = DYDETA1 * DELTA_ETA_1
TERM_O1_2 = DYDETA2 * DELTA_ETA_2
TMP1 = MYY + TERM_O1_1 + TERM_O1_2
TERM_O2_1 = DELTA_ETA_1**2 * D2YDETA11
TERM_O2_2 = DELTA_ETA_2**2 * D2YDETA22
TMP2 = TMP1 + 1/2*(TERM_O2_1 + TERM_O2_2)
TERM_O3_12 = DELTA_ETA_1 * DELTA_ETA_2 * D2YDETA12
Y = TMP2 + TERM_O3_12 + TERM_O3_13

$THETA  1 FIX
$THETA  1 FIX
$THETA  (-0.333,-0.095057,0.143) ; BASEBMMS1
$THETA  (-0.333,-0.000616045,0.143) ; SLOPEBMMS1
$OMEGA  BLOCK(2)
 0.0561313  ;       BASE
 5.20099E-05 6.20319E-05  ;    2 SLOPE
$ESTIMATION MAXEVAL=9999 METHOD=1 LAPLACE -2LL
```

Conclusions

✓ A strategy to implement second-order Taylor expansion of the likelihood of the nonlinear mixed effect model was derived, which enables fast and stable assessments of covariates for likelihood-based models. The proposed method expands the use of the linearization technique [1,6,8] to also include likelihood-based models.

References:

[1] Khandelwal A, Harling K, Jonsson EN, Hooker AC, **Karlsson MO**. A Fast Method for Testing Covariates in Population PK/PD Models, AAPS J. 2011;13:464–472
[2] Schindler E and **Karlsson MO**. PAGE 26 (2017) Abstr 6077 [www.page-meeting.org/?abstract=6077]
[3] Hansson EK, Ma G, Amantea MA, French J, Milligan PA, Friberg LE, and **Karlsson MO**. PKPD Modeling of Predictors for Adverse Effects and Overall Survival in Sunitinib-Treated Patients With GIST, CPT Pharmacometrics Syst Pharmacol. 2013;2(12):e85
[4] **Karlsson MO** and Wellhagen G. A bounded integer model for rating and composite scale data. PAGANZ 2018
[5] Wellhagen G and **Karlsson MO**. A bounded integer model for rating and composite scale data. PAGE 27 (2018)
[6] Jonsson EN, **Karlsson MO**. Automated covariate model building within NONMEM., Pharm Res. 1998 Sep;15(9):1463-8.
[7] Keizer RJ, **Karlsson MO**, Hooker A. Modeling and Simulation Workbench for NONMEM: Tutorial on Pirana, PsN, and Xpose. CPT Pharmacometrics Syst Pharmacol. 2013 Jun 26;2:e50. doi: 10.1038/psp.2013.24.
[8] Svensson EM, **Karlsson MO**. Use of a linearization approximation facilitating stochastic model building, J Pharmacokinet Pharmacodyn. 2014;41:153-158