Comparison of Laplacian, Quasi-Random Parametric Expectation Maximization and Non-parametric Methods for Population Analysis of a Complex Dynamic System with Non-static BQL Data

Shuhua Hu, Bob Leary, Kevin Feng, Mike Dunlavey

Certara/Pharsight, Cary, NC, 27518

Introduction

Phoenix® NLME™ (Pharsight/Certara) provides several estimation methods that can be used to do population analysis for a system involved BQL data. These include the Laplacian method, the quasi-random parametric expectation maximization (QRPEM) method, and the non-parametric method. For all these methods, the cumulative distribution function evaluated at LOQ is involved in calculating the likelihood. This can be easily handled using the "bql" option in the observe statement.

- The Laplacian method involves explicit numerical optimization of an approximate marginal likelihood, which is obtained by approximating the marginal likelihood using a Laplacian approximation. Hence, it requires evaluation of second-order derivatives of the logarithm of the joint likelihood with respect to random effects.
- Unlike the Laplacian method, the QRPEM method is a member of expectation maximization methods. It is very similar to the importance sampling based Monte Carlo parametric expectation maximization (MCPEM) method, with the exception that samples are based on low discrepancy (quasi-random) sequence instead of the conventional pseudo-random sequence to compute the required integral. Note: The integral error resulting from the quasi-random sequence decays much faster (at rate $O(N^{-1})$ with N being the number of samples) than that using the pseudo-random sequence (at rate $O(N^{-1/2})$). Hence, the integrals in the QRPEM method converge much faster than those in the MCPEM method.

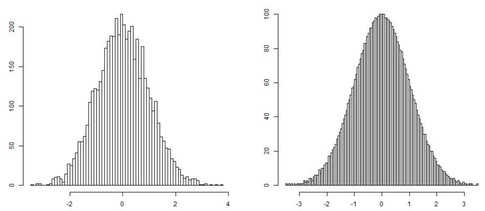


Figure 1: (left panel): histogram of 5000 normally distributed pseudo-random points; (right panel): histogram of 5000 normally distributed quasi-random points (right).

- Compared to parametric methods such as Laplacian and QRPEM methods, the non-parametric method makes no assumption on the distribution form of random effects. It involves using a discrete distribution to approximate the cumulative distribution function of random effects and hence the resulting marginal likelihood can be obtained analytically. Note: The non-parametric engine in Phoenix® NLME™ has the capability of optimizing both probabilities and associated support point positions.

Objectives

- To compare the capability of Laplacian and QRPEM engines in Phoenix® NLME™ for population analysis of a complex dynamic system with non-static BQL data;
- To use non-parametric engine as a post-processor for parametric runs to detect any serious violation of normality assumption such as bimodality.

Methods

The example used to test the capability of these three methods is a highly nonlinear, multi-scaled and long-term HIV dynamic model [1] with clinical data [1].

HIV Model [1]

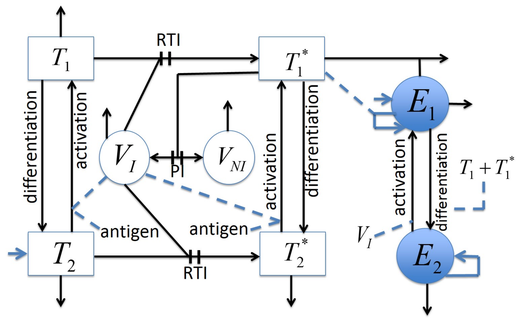


Figure 2: Diagram of HIV model proposed in [1]. Model states T_1 and T_1^* respectively denote the uninfected and infected activated CD4+ T cells, and T_2 and T_2^* respectively represent uninfected resting CD4+ T cells and latently infected CD4+ T cells. Model states V_I and V_{NI} respectively denote infectious and non-infectious virus, and E_1 and E_2 respectively represent activated and memory HIV-specific CD8+ T cells. PI and RTI denote protease inhibitor and reverse transcriptase inhibitor, respectively.

$$\dot{T}_1 = -d_1 T_1 - (1-\xi_1(t))k_1 V_I T_1 - \gamma_T T_1 + p_T\left(\tfrac{a_T V_I}{V_I+K_V} + a_A\right) T_2,$$

$$\dot{T}_1^* = (1-\xi_1(t))k_1 V_I T_1 - \delta T_1^* - mE_1 T_1^* - \gamma_T T_1^* + p_T\left(\tfrac{a_T V_I}{V_I+K_V} + a_A\right) T_2^*,$$

$$\dot{T}_2 = \lambda_T \tfrac{K_s}{V_I+K_s} + \gamma_T T_1 - d_2 T_2 - (1-f\xi_1(t))k_2 V_I T_2 - \left(\tfrac{a_T V_I}{V_I+K_V} + a_A\right) T_2,$$

$$\dot{T}_2^* = \gamma_T T_1^* + (1-f\xi_1(t))k_2 V_I T_2 - d_2 T_2^* - \left(\tfrac{a_T V_I}{V_I+K_V} + a_A\right) T_2^*,$$

$$\dot{V}_I = (1-\xi_2(t))10^3 N_T \delta T_1^* - cV_I - 10^3(1-\xi_1(t))\rho_1 k_1 T_1 V_I - 10^3(1-f\xi_1(t))\rho_2 k_2 T_2 V_I,$$

$$\dot{V}_{NI} = \xi_2(t)10^3 N_T \delta T_1^* - cV_{NI},$$

$$\dot{E}_1 = \lambda_E + \tfrac{b_{E1}T_1^*}{T_1^*+K_{b1}}E_1 - \tfrac{d_E T_1^*}{T_1^*+K_d}E_1 - \delta_{E1}E_1 - \gamma_E \tfrac{T_1+T_1^*}{T_1+T_1^*+K_\gamma}E_1 + \tfrac{p_E a_E V_I}{V_I+K_V}E_2,$$

$$\dot{E}_2 = \gamma_E \tfrac{T_1+T_1^*}{T_1+T_1^*+K_\gamma}E_1 + \tfrac{b_{E2}K_{b2}}{E_2+K_{b2}}E_2 - \delta_{E2}E_2 - \tfrac{a_E V_I}{V_I+K_V}E_2,$$

with initial condition $(T_1(0), T_1^*(0), T_2(0), T_2^*(0), V_I(0), V_{NI}(0), E_1(0), E_2(0))^T = (T_1^0, T_1^{*0}, T_2^0, T_2^{*0}, V_I^0, V_{NI}^0, E_1^0, E_2^0)^T$.

...

Clinical Data [1]

- Consisting of 14 patients followed for varying lengths of time between 2 and 6 years who all underwent antiretroviral therapy and had at least one treatment interruption;
- Including total CD4+ T cells and non-static BQL viral load (due to different assays used in the investigated period).

Setup

Both Laplacian and QRPEM engines started with same initial values and were implemented using parallel computations with 4 processors running in the same computer using the same ODE solver with the same setup. The QRPEM method uses 300 quasi-random sample points for evaluation of the required integral. For both engines, the sandwich method is used to calculate standard errors.

Results

The Laplacian method versus the QRPEM method

- The convergence status and runtime for the Lapacian method and the QRPEM estimation are summarized in the following table.

	QRPEM method	Laplacian method
return code	1	3
optimal -2LL	58.8452	2064.18
engine runtime (secs)	9126.656	9139.297
number of iterations	151	3
SE runtime (secs)	400.547	17258.172
SE status	successful	failed

Table 1: -2LL denotes twice the negative log likelihood, SE stands for standard error. Return code of 1 means sucessful convergence, and return code of 3 means that the line search step in the quasi-Newton direction failed to locate a sufficiently better objective value than the current value.

- For both QRPEM and Laplacian methods, we obtained reasonably good model fitting results for all the subjects (Figure 3 shows model fitting results for an example patient using the QRPEM estimation).
- For all the subjects except one, the residual mean square errors obtained using the QRPEM are smaller than the corresponding ones obtained by the Laplacian method.

This is expected for the following reasons.

- The QRPEM estimation does not involve numerical differentiation while the Laplacian method does, and hence the QRPEM method is more stable and reliable than the Laplacian method.
- The estimates obtained by the QRPEM can be made as accurate as desired to the true marginal likelihood estimates (through increasing the number of random sample points) while the Laplacian method cannot.

...

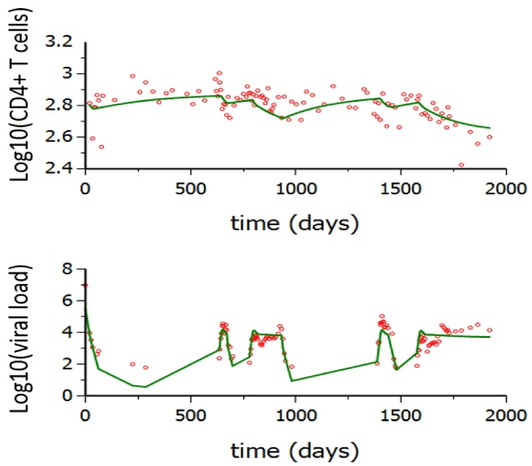
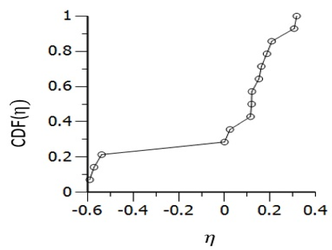


Figure 3: Model fitting results for an example patient, where red circles are the actual observations, and green solid line denotes the predicated model solution.

- The Q-Q plots for posthoc values obtained by both Laplacian and QRPEM methods suggest that there are discernible divergences from normal distributions for some random effects.
- Plots of estimated cumulative distribution functions (CDF) obtained by the non-parametric engine demonstrate that some of these random effects seem to have bimodal shape distributions (see figure shown below).



Conclusions

This example demonstrates that for population analysis of a complex dynamic system with complicated data, the QRPEM estimation is the method of choice and the nonparametric engine should be used as a post-processor whenever there is a doubt of the normality assumption.

References

[1] H.T. Banks, M. Davidian, S. Hu, G.M. Kepler and E.S. Rosenberg, Modeling HIV immune response and validation with clinical data, *Journal of Biological Dynamics*, 2 (2008), 357–385.