Bayesian Calibration of Stochastic Multistate Simulators

Oksana A. Chkrebtii
Department of Statistics, The Ohio State University

Joint work with Matthew T. Pratola (Statistics, The Ohio State University)

Forecasting from Complexity
IMA Workshop
April 24, 2018
Bayesian inference (calibration) problem

Spatio-temporal process $U_t = U(x_t, \theta)$, $x_t \in X$, $\theta \in \Theta$ is measured through an observation operator $G : U \rightarrow \mathcal{Y}$.

\[
Y_t \mid U_t, \theta, \xi \sim \mathcal{D}(GU_t), \ t = 1, \ldots, T,
\]

\[
U_t \mid \theta, \xi \sim \mathcal{M}(\theta, \xi), \ t = 1, \ldots, T
\]

Our goal is estimation and uncertainty quantification for parameters $\theta$
The dynamical (forward) model $\mathcal{M}$

$$U_t | \theta, \xi \sim \mathcal{M}(\theta, \xi), \ t = 1, \ldots, T$$

- Incorporates known physical laws
- Not available in closed form
- Forward simulation given $\theta, \xi$ is computationally expensive, making direct inference infeasible
- Computer model emulation framework for deterministic models $U_t = M_t(\theta, \xi)$ is well developed
- Lots of work in the data assimilation literature too!

We focus on the cases when the forward model is:

- deterministic with uncertainty
- stochastic (e.g. agent based, contains process noise)
Inference from deterministic simulators
Existing framework for deterministic models

Kennedy & O’Hagan (2001) and subsequent work set up the following data augmentation problem:

\[ Y_{1:T} \mid U_{1:T}, M_{1:T}(s_1), \ldots, M_{1:T}(s_M), \theta, \xi \sim \mathcal{D}(GU_{1:T}), \ s_i \in \Theta, \]

\[ M_{1:T}(s_1), \ldots, M_{1:T}(s_M) \mid U_{1:T}, \xi \sim \mathcal{N}(U_{1:T}, C), \ s_i \in \Theta, \]

\[ U_{1:T}(\cdot) \mid \xi \sim \mathcal{GP}(0, \xi), \]

\[ \theta, \xi \sim \mathcal{P}. \]

\( M_{1:T}(s_1), \ldots, M_{1:T}(s_M) \) are evaluations of the forward model at regimes (inputs) \( s_i \in \Theta, i = 1, \ldots, M \) in the parameter space. \( \xi \) is a covariance operator with a small number of non-zero eigenvalues.
Inference from stochastic simulators
The stochastic forward model

Kennedy & O’Hagan (2001) consider the case where, for each parameter setting $s \in \Theta$, the forward model output is fixed. However, this is often not the case.
A Hierarchical model representation

Ideally:

• For a given parameter regime $s_i \in \Theta$, we could generate ensembles “on demand” $\rightarrow$ exact simulation
• For a given parameter regime $s_i \in \Theta$, we could generate a very large number of ensembles $\rightarrow$ re-sampling from large ensemble provides close enough approximation

In reality:

• Small number $K$ of ensembles for each model run
• Additional dimension reduction required for the second layer of the hierarchical model
• We include the dimension reduction specifications within the hierarchical model, resulting in a fully probabilistic approach
• Multiple states require tensor valued models
Motivation: probability models for discretization uncertainty of ODEs/PDEs
Modeling uncertainty in the unknown solution

For fixed $\theta$, consider the ODE initial value problem,

$$\begin{cases} 
Du = f(t, u; \theta), & t \in (0, L], \theta \in \Theta \\
u = u_0 & t = 0.
\end{cases}$$

Unknown exact solution given $\theta$ introduces highly constrained model uncertainty. We provide a way to model this uncertainty: a probabilistic DE solver defined by the posterior* over $u$ as follows.

$$[u \mid \theta, N] = \int [u, Du, f \mid \theta, N] \, dDu \, df$$

Here $f$ are auxiliary variables that “interrogate” the model at each discretization grid point with some error (which we also model).

*Concentrates to exact solution as the maximum grid spacing approaches 0
Sequential Bayesian updating

Predictive probability distribution over the state and its derivative is:

\[
\begin{bmatrix}
Du(t_j) \\
u(t_k)
\end{bmatrix}
\mid f_n, \ldots, f_1
= \mathcal{GP}
\begin{bmatrix}
\begin{bmatrix}
m^n_t(t_j) \\
m^n(t_k)
\end{bmatrix}
, 
\begin{bmatrix}
C^n_t(t_j, t_j) \\
\int_0^{t_j} C^n_t(z, t_k)dz \\
\int_0^{t_k} C^n_t(t_j, z)dz \\
C^n(t_k, t_k)
\end{bmatrix}
\end{bmatrix}
\]

Marginal means and covariances with \( g_n = C^{-1}_n(s_n, s_n) + r_{n-1}(s_n) \) are:

\[
m^n_t(t_j) = m^{-1}_t(t_j) + C^{-1}_t(t_j, s_n) g_n^{-1} \left\{ f_n - m^{-1}_n(s_n) \right\},
\]

\[
m^n(t_k) = m^{-1}(t_k) + g^{-1} \int_0^{t_k} C^{-1}_t(z, s_n)dz \left\{ f_n - m^{-1}_t(s_n) \right\},
\]

\[
C^n_t(t_j, t_k) = C^{-1}_t(t_j, t_k) - C^{-1}_t(t_j, s_n) g_n^{-1} C^{-1}_t(s_n, t_k),
\]

\[
C^n(t_j, t_k) = C^{-1}(t_j, t_k) - g_n^{-1} \int_0^{t_j} C^{-1}_t(z, s_n)dz \left\{ \int_0^{t_k} C^{-1}_t(z, s_n)dz \right\}^\top,
\]
Example - simple ODE initial value problem

A simple example: the second order initial value ODE problem,

\[
\begin{align*}
D^2 u(t) &= \sin(2t) - u, \quad t \in [0, 10], \\
Du(0) &= 0, \quad u(0) = -1.
\end{align*}
\]

The exact solution, assumed unknown a priori, is

\[
u^*(t) = \left\{-4 \cos(t) + 2 \sin(t) - \sin(2t) + \cos(t)\right\}/(4 - 1).
\]

Five draws from the prior process for the state (left) and first derivative (right)
Example - simple ODE initial value problem

A simple example: the second order initial value ODE problem,

\[
\begin{align*}
D^2 u(t) &= \sin(2t) - u, \quad t \in [0, 10], \\
Du(0) &= 0, \quad u(0) = -1.
\end{align*}
\]

The exact solution, assumed unknown a priori, is

\[
u^*(t) = \{-4 \cos(t) + 2 \sin(t) - \sin(2t) + \cos(t)\}/(4 - 1).
\]

Five draws from the prior process for the state (left) and first derivative (right)
Example - simple ODE initial value problem

A simple example: the second order initial value ODE problem,

\[
\begin{align*}
D^2 u(t) &= \sin(2t) - u, \quad t \in [0, 10], \\
Du(0) &= 0, \quad u(0) = -1.
\end{align*}
\]

The exact solution, assumed unknown a priori, is

\[u^*(t) = \left\{-4 \cos(t) + 2 \sin(t) - \sin(2t) + \cos(t)\right\}/(4 - 1).\]
Example - simple ODE initial value problem

A simple example: the second order initial value ODE problem,

\[
\begin{cases}
    D^2 u(t) = \sin(2t) - u, & t \in [0, 10], \\
    Du(0) = 0, & u(0) = -1.
\end{cases}
\]

The exact solution, assumed unknown a priori, is
\[ u^*(t) = \left\{ -4 \cos(t) + 2 \sin(t) - \sin(2t) + \cos(t) \right\}/(4 - 1). \]
Example - simple ODE initial value problem

A simple example: the second order initial value ODE problem,

\[
\begin{cases}
D^2 u(t) = \sin(2t) - u, & t \in [0, 10], \\
Du(0) = 0, & u(0) = -1.
\end{cases}
\]

The exact solution, assumed unknown a priori, is
\[
u^*(t) = \left\{-4 \cos(t) + 2 \sin(t) - \sin(2t) + \cos(t)\right\}/(4 - 1).
\]

Five draws from the prior process for the state (left) and first derivative (right)
Example - Lorenz63 forward model

A probability statement over probable trajectories given fixed model parameters and initial conditions for the Lorenz63 model:

1000 draws for the probabilistic forward model for the Lorenz63 system given fixed initial states and model parameters in the chaotic regime.
Example - Lorenz63 forward model

1000 draws from forward model for Lorenz63 system at four fixed time points.
Example: Exact vs Dimension Reduced Inference in a Model of Protein Dynamics
Inference for a model of protein dynamics

JAK-STAT chemical signaling pathway model describes concentration of 4 STAT factors by a delay differential equation system on $t \in [0, 60]$,

$$
Du_t^{(1)}(\theta) = -\theta_1 u_t^{(1)}(\theta) EpoR_A(t) + 2 \theta_4 u_{t-\tau}^{(4)}(\theta)
$$

$$
Du_t^{(2)}(\theta) = \theta_1 u_t^{(1)}(\theta) EpoR_A(t) - \theta_2 \left( u_t^{(2)}(\theta) \right)^2
$$

$$
Du_t^{(3)}(\theta) = -\theta_3 u_t^{(3)}(\theta) + 0.5 \theta_2 \left( u_t^{(2)}(\theta) \right)^2
$$

$$
Du_t^{(4)}(\theta) = \theta_3 u_t^{(3)}(\theta) - \theta_4 u_{t-\tau}^{(4)}(\theta)
$$

$$
u_t^{(i)}(\theta) = \phi^{(i)}(t), \quad t \in [-\tau, 0], \quad i = 1, \ldots, 4
$$
Inference for a model of protein dynamics

States are observed indirectly through a nonlinear transformation:

\[ G^{(1)}(u, \theta) = \theta_5 (u_t^{(1)}(\theta) + 2u_t^{(3)}(\theta)) \]
\[ G^{(2)}(u, \theta) = \theta_6 (u_t^{(1)}(\theta) + u_t^{(2)}(\theta) + 2u_t^{(3)}(\theta)) \]
\[ G^{(3)}(u, \theta) = u_t^{(1)}(\theta) \]
\[ G^{(4)}(u, \theta) = \frac{u_t^{(3)}(\theta)}{u_t^{(2)}(\theta) + u_t^{(3)}(\theta)} \]

Observations are noisy measurements on the transformed states and forcing function at points \( t = \{ t_{ij} \}_{i=1,\ldots,4; j=1,\ldots,n_i} \)

\[ y(t) = G_{\theta_5, \theta_6} u_t(\theta_1, \ldots, \theta_4, \phi, \tau, EpoR_A) + \varepsilon_t \]
Exact inference

Draws from the marginal posterior distribution over the states (bottom row) and the states transformed via the observation process (top row); experimental measurements are shown in red.
Dimension-reduced model calibration

Draws from the marginal calibrated posterior with $m = 100$ model runs (top row), discrepancies $\delta_1$ and $\delta_2$ (middle row); experimental measurements are shown in red.
Dimension-reduced model calibration

Contours of the marginal stochastically calibrated posterior (gray) with $M = 100$ model runs each with $K = 10$ ensembles; exact posterior contours in black.
Thank you!

Some references:


Contact: oksana@stat.osu.edu