Sensitivity analysis, parameter inversion, and design of experiments under uncertainty

Alen Alexanderian

Department of Mathematics
North Carolina State University

October 9, 2020
Global Sensitivity analysis (GSA)

**model:** \( y = f(\theta) \)
- Evaluating \( f(\theta) \) requires evaluating a mathematical model (ODE, PDE, DAE, Stochastic model, ...)
- \( \theta \in \mathbb{R}^p \) vector of uncertain model parameters

**global sensitivity analysis (GSA):**
- *find entries of \( \theta \) that are “most important to model output”*
- Various ways of defining parameter importance: variance based, derivative based, moment independent, etc.

**why do GSA?**
- Understand the model better
- Reduce dimensionality of input parameter (accelerate forward/inverse UQ problems)
- Simplify the model ...
Some standard GSA measures

variance based:

\[ S_i = \frac{\text{Var} \left\{ \mathbb{E} \{ f(\theta) | \theta_i \} \right\}}{\text{Var} \left\{ f(\theta) \right\}}, \quad i = 1, \ldots, p \]

estimation requires Monte-Carlo sampling

derivative based:

\[ \nu_i = \mathbb{E} \left\{ \left( \frac{\partial f}{\partial \theta_i} \right)^2 \right\}, \quad i = 1, \ldots, p \]

and others: moment independent; Morris screening; activity scores; ...

model:  \( y = f(\theta) \)

- high dimensional \( \theta \)
- expensive to evaluate \( f \): models governed by complex physics systems
- entries of \( \theta \) might be correlated
- results depend on how parameter uncertainty is modeled; robustness issues
- \( f \) might be time/space dependent \( \Rightarrow \) high-dimensional output
- stochasticity in the model, \( y = f(\theta, \omega) \); e.g., stochastic compartment models, stochastically forced dynamical systems;
- problems that are defined at multiple scales

in this talk:

we will discuss GSA across physical scales
GSA across scales: application to stochastic chemical systems
Stochastic chemical kinetics: basic setup

- species $S_1, S_2, \ldots, S_N$
- state vector $\mathbf{X}(t) = (X_1(t), \ldots, X_N(t))^T$
  
  $X_i(t) =$ number of molecules of $i$th species at $t$

- $M$ reactions; each reaction has
  - a stoichiometric (state change) vector $\nu_j$
  - a reaction rate $k_j$ (we model as uncertain)
  - a propensity function $a_j(\mathbf{X})$
  
  \[
  a_j(\mathbf{X}(t))dt = \text{Prob}\{\text{reaction } j \text{ occurs in time interval } [t, t + dt]\}
  \]

**example:** two species model

\[
\begin{align*}
S_1 & \xrightarrow{k_1} S_2 \\
S_2 & \xrightarrow{k_2} S_1
\end{align*}
\]

\[
\begin{align*}
a_1(\mathbf{X}) &= k_1 X_1 \\
\nu_1 &= \begin{bmatrix} -1 \\ 1 \end{bmatrix} \\
a_2(\mathbf{X}) &= k_2 X_2 \\
\nu_2 &= \begin{bmatrix} 1 \\ -1 \end{bmatrix}
\end{align*}
\]
Multiple scales

- **microscale**: track number of molecules of each species over time
  \[ X(t, \omega) \in \mathbb{Z}^N_{\geq 0} \] (use Stochastic Simulation Algorithm)

- **mesoscale**: track species evolution via real valued stochastic process
  (Chemical Langevin Equation)
  \[ Y(t, \omega) \in \mathbb{R}^N_{\geq 0} \] (use methods for SDEs)

- **macroscale**: track species concentration using a system of ODEs
  (reaction rate equations — RREs)
  \[ y(t) \in \mathbb{R}^N_{\geq 0} \] (use methods for ODEs)
Michaelis-Menten system:

\[ S + E \xrightarrow{k_1} C \]
\[ C \xrightarrow{k_2} S + E \]
\[ C \xrightarrow{k_3} P + E \]

**species:**  
- \( S \) (substrate),  
- \( E \) (enzyme),  
- \( C \) (complex);  
- \( P \) (product).

**uncertain parameters:**  
\( \theta = (k_1, k_2, k_3)^T \)

**quantity of interest (QoI):**  
average product concentration:

\[ f(\theta, \omega) = \frac{1}{T} \int_0^T [P](t, \theta, \omega) \, dt \]

**GSA:** with respect to reaction rates

**question:** can we use the macroscale model for efficient GSA of the microscale model?
Michaelis–Menten simulation
Michaelis–Menten simulation

Black lines: solution of RREs
GSA across scales in a nutshell

stochastic model: \( y = f(\theta, \omega) \)

for example: \( f(\theta, \omega) = \frac{1}{T} \int_0^T [P](t, \omega, \theta) \, dt \)

\[
\begin{align*}
  f(\theta, \omega) & \xrightarrow{\text{GSA}} \{S_j(\omega)\}_{j=1}^M \\
  \tilde{f}(\theta) & \xrightarrow{\text{GSA}} \{\tilde{S}_j\}_{j=1}^M
\end{align*}
\]

questions: does this diagram commute? in what sense? what assumptions are needed?
Relation between micro and macro scale models

**system size:** \( V = \text{Avogadro number} \times \text{system volume} \)

**evolution of stochastic species concentrations:**

\[
Z^V(t) = Z^V(0) + \sum_{j=1}^{M} \frac{1}{V} \nu_j Y_j \left( V \int_0^t \tilde{a}_j(Z^V(s)) \, ds \right)
\]

**thermodynamic limit:**

\[
\lim_{V \to \infty} \sup_{s \leq t} |Z^V(s) - y(s)| = 0 \quad \text{a.s. for all } t > 0
\]

where \( y(t) \) satisfies the RREs

\[
\frac{dy}{dt} = F(y(t)), \quad y(0) = y_0, \quad \text{with } F(y) = \sum_{j=1}^{M} \nu_j \tilde{a}_j(y)
\]

**example of stochastic QoI:**

\[
f(\theta, \omega) = \int_0^T Z^V_1(t, \theta, \omega) \, dt
\]

**the corresponding “deterministic” QoI:**

\[
\tilde{f}(\theta) = \int_0^T y_1(t, \theta) \, dt
\]
Numerical results: back to the Michaelis-Menten system

\[ S + E \xrightleftharpoons[k_1]{k_2} C \xrightarrow{k_3} P + E \]

\[ f(\theta, \omega) = \int_0^T [P](\theta, \omega, t) \, dt \quad \theta = (k_1, k_2, k_3)^T \]

convergence of \( T_1(\omega) \) in distribution for increasing \( V \)

(\( T_1 \): total Sobol’ index corresponding to \( k_1 \))
Numerical results: back to the Michaelis-Menten system

\[
S + E \xrightleftharpoons[k_1]{k_2} C \xrightarrow{k_3} P + E
\]

\[
f(\theta, \omega) = \int_0^T [P](\theta, \omega, t) \, dt \quad \theta = (k_1, k_2, k_3)^T
\]

convergence of \( T_2(\omega) \) in distribution for increasing \( V \)
Numerical results: back to the Michaelis-Menten system

\[ S + E \xrightleftharpoons[k_2]{k_1} C \xrightarrow{k_3} P + E \]

\[ f(\theta, \omega) = \int_0^T [P](\theta, \omega, t) \, dt \quad \theta = (k_1, k_2, k_3)^T \]

convergence of \( T_3(\omega) \) in distribution for increasing \( V \)
GSA across scales: summary

- can use the (deterministic) RREs for efficient GSA of stochastic chemical system
- numerical illustrations on biochemical reaction networks
- analysis of convergence
- many interesting remaining questions ...

Inverse problems and optimal design of experiments
Example: Diffusive transport of a contaminant with uncertain initial condition

- **Governing PDE (forward model):**
  advection-diffusion equation

- **Unknown/uncertain parameter:** initial concentration field

- **Inverse problem:** Use a vector $d$ of space/time sensor measurements of concentration to reconstruct the initial state
2D Model problem

Forward problem: time dependent advection-diffusion

\[ u_t - \kappa \Delta u + \mathbf{v} \cdot \nabla u = 0 \quad \text{in } \mathcal{D} \times [0, T] \]
\[ u(0, \mathbf{x}) = m \quad \text{in } \mathcal{D} \]
\[ \kappa \nabla u \cdot \mathbf{n} = 0 \quad \text{on } \partial \mathcal{D} \times [0, T] \]

- \( m \): unknown initial condition
- \( \mathbf{v} \): velocity field
Solution of the forward problem

$t = 0$

$t = 1$

$t = 2$

$t = 3$
The inverse problem: reconstruct initial condition

The inverse problem of finding the unknown initial state based on sensor data

\[
\min_m \frac{1}{2} \|Bu(m) - d\|^2 + \frac{\alpha}{2} \langle Am, m \rangle
\]

where

\[
Bu(m) = Bu_t - \kappa \Delta u + v \cdot \nabla u = 0 \quad \text{in } D \times [0, T] \\
u(0, x) = m \quad \text{in } D \\
k \nabla u \cdot n = 0 \quad \text{on } \partial D \times [0, T]
\]

- \(B\): observation operator
- \(d = [d_1^T \ d_2^T \ \cdots \ d_{nt}^T]^T, \ d_i \in \mathbb{R}^{ns}, ns = \text{number of sensors} \)
- \(u\) linear in \(m\), \(u = Sm \implies \text{linear parameter-to-observable map: } \mathcal{F} = BS \)
- Can rewrite the optimization problem as

\[
\min_m J(m) := \frac{1}{2} \|\mathcal{F}m - d\|^2 + \frac{\alpha}{2} \langle Am, m \rangle
\]
Solving the inverse problem

- Derivative of $\mathcal{J}$

$$D\mathcal{J}(m)(\tilde{m}) = \frac{d}{d\varepsilon} \mathcal{J}(m + \varepsilon\tilde{m}) \big|_{\varepsilon=0} = \langle \mathcal{F}^*(\mathcal{F}m - d) + \alpha Am, \tilde{m} \rangle$$

- Action of $\mathcal{F}^*$

$$\mathcal{F}^* y = p(\cdot, 0), \text{ where } p \text{ is solution of the adjoint equation}$$

$$-p_t - \nabla \cdot (pv) - \kappa \Delta p = -\mathcal{B}^* y$$

$$p(T) = 0$$

$$(vp + \kappa \nabla p) \cdot n = 0$$

- Optimality condition

$$(\mathcal{F}^* \mathcal{F} + \alpha A)m = \mathcal{F}^* d \quad \text{discretize} \quad (F^* F + \alpha A)m = F^* d$$

Solve the linear system using an iterative method, e.g. conjugate gradient
Solving the inverse problem: numerical results

Truth

Sensor sites

Reconstruction
Solving the inverse problem: numerical results

Truth

Sensor sites

Reconstruction
How to place sensors in an “optimal” way?

- Can formulate the optimal sensor placement problem as an optimal experimental design (OED) problem
- Can consider a statistical formulation of the inverse problem
- In addition to a reconstruction, we can also compute a statistical distribution of the parameters, conditioned on experimental data
- Find sensor locations so as to optimize the statistical quality of the reconstructed/inferred parameter
- In context of inverse problems a Bayesian formulation is natural
Bayesian inference: Bayes’ formula

$$\pi_{\text{post}}(m|d) \propto \pi_{\text{like}}(d|m) \pi_{\text{prior}}(m)$$

- $\pi_{\text{post}}(m|d)$: posterior pdf of $m$
- $\pi_{\text{like}}(d|m)$: pdf of $d$ given $m$ (data likelihood)
- $\pi_{\text{prior}}(m)$: prior pdf of $m$

pdf = probability density function

Bayes, T., An Essay towards Solving a Problem in the Doctrine of Chances. By the Late Rev. Mr. Bayes, FRS Communicated by Mr. Price, in a Letter to John Canton, AMFRS. Philosophical Transactions, 1763.

Laplace, P.S., Théorie analytique des probabilités. 1820.
Bayesian linear inverse problems

Assume linear parameter-to-observable map:

\[ d = Fm + \eta \]

and assume prior is Gaussian

\[ \pi_0(m) \propto \exp(-\frac{1}{2}m^T\Gamma_{\text{prior}}^{-1}m) \]

Then, the posterior pdf is

\[ \pi_{\text{post}}(m|d) \propto \exp\left\{ -\frac{1}{2}(m - m_{\text{MAP}})^T(F^T\Gamma_{\text{noise}}^{-1}F + \Gamma_{\text{prior}}^{-1})(m - m_{\text{MAP}}) \right\} \]

\[ \Rightarrow \mu_{\text{post}} = \mathcal{N}(m_{\text{MAP}}, \Gamma_{\text{post}}) \]

\[ \Gamma_{\text{post}}^{-1} = F^T\Gamma_{\text{noise}}^{-1}F + \Gamma_{\text{prior}}^{-1} \quad (= D_m^2(-\log \pi_{\text{post}})) \]

\[ m_{\text{MAP}} = \arg \min_m \frac{1}{2}\|Fm - d\|^2_{\Gamma_{\text{noise}}^{-1}} + \frac{1}{2}\langle \Gamma_{\text{prior}}^{-1}m, m \rangle \]

An important problem structure:

\[ H_{\text{misfit}} \text{ is low rank, and } \Gamma_{\text{prior}}^{1/2}H_{\text{misfit}}\Gamma_{\text{prior}}^{1/2} \text{ is even more so ...} \]
Bayesian inversion of the initial condition for 2D advection-diffusion equation

- Posterior mean, and posterior variance

\[
\nu = m_{\text{MAP}} + \Gamma_{\text{post}}^{1/2} M^{-1/2} n, \quad n \sim \mathcal{N}(0, I)
\]

- Posterior samples:
The optimal experimental design problem

A grid of candidate locations for observation points

- **Experimental design**:
  locations of observation points / sensors

  \[
  \text{design} := \left\{ x_1, \ldots, x_{N_s}, w_1, \ldots, w_{N_s} \right\}
  \]

- **Bayesian inversion**:
  data + likelihood, prior \(\xrightarrow{}\) posterior distribution of inversion parameter

- **Optimal experimental design (OED)**:
  Find sensor locations that result in minimized posterior uncertainty
Commonly used OED criteria

Bayesian A-optimal experimental design:

\[
\min_{\mathbf{w} \in S} \text{tr}\left[\Gamma_{\text{post}}(\mathbf{w})\right] + \gamma P(\mathbf{w})
\]  

Bayesian D-optimal experimental design:

\[
\min_{\mathbf{w} \in S} -\frac{1}{2} \log \det(I + \Gamma_{\text{prior}}^{1/2} H_{\text{misfit}}(\mathbf{w}) \Gamma_{\text{prior}}^{1/2}) + \gamma P(\mathbf{w})
\]

- Need trace/log-determinant of high-dimensional operators
- Need many applications of the forward operator \( \Rightarrow \) many PDE solves
- OED much harder for nonlinear inverse problems
- Ingredients of efficient OED methods: randomized matrix methods, use of problem structure, low-rank approximations, iterative solvers, gradient based optimization algorithms, adjoint based gradient/Hessian computation, sparsifying penalty methods, ...

A-optimal design: the variance field

Optimal

Sub-optimal
A-optimal design: the variance field

Optimal

Sub-optimal
A-optimal design: the variance field

Optimal

Sub-optimal
OED for 3D model (parameter dim $\sim 10^4$)
Global sensitivity analysis: apportion uncertainty in the model output to different sources of uncertainty in the model input parameters

- Interesting directions: theory and computational methods for global sensitivity analysis with high-dimensional inputs/outputs; GSA for stochastic models; GSA across scales; sensitivity analysis of inverse problems

- Applications: flow through porous media, contaminant transport, radioactive waste storage, biotransport in cancerous tumors, chemical kinetics, biochemistry, epidemiology, pharmacokinetics ...

- NCSU faculty collaborators: Pierre Gremaud, Ralph Smith

- Students: Helen Cleaves, Mike Merritt, Isaac Sunseri
Optimal placement for Bayesian inverse problems: find optimal placements of measurement point to minimize the uncertainty in reconstructed parameters

- Interesting research directions: scalable algorithms for sensor placement for linear inverse problems governed by PDEs (randomized methods in numerical linear algebra, low-rank approximations, optimization with exact penalty method, ...); Optimal sensor placement for nonlinear inverse problems, (bi-level PDE-constrained optimization, adjoint based derivative computation, ... ), optimal experimental design under model uncertainty (marginalization, stochastic optimization, ...); sequential design of experiments

- Applications: porous medium flow, contaminant source identification, radiation detection in urban environments

- NCSU faculty collaborators: Arvind Saibaba, Ralph Smith

- Students: Isaac Sunseri, Bekah White
Relevant courses

<table>
<thead>
<tr>
<th>Course Code</th>
<th>Course Name</th>
</tr>
</thead>
<tbody>
<tr>
<td>MA 580 - 780</td>
<td>Numerical Analysis</td>
</tr>
<tr>
<td>MA 515 - 715</td>
<td>Analysis</td>
</tr>
<tr>
<td>MA 573 - 574</td>
<td>Mathematical Modeling</td>
</tr>
<tr>
<td>MA 534 - 734</td>
<td>Partial Differential Equations</td>
</tr>
<tr>
<td>MA 546 - 747</td>
<td>Applied Probability and Stochastic Processes</td>
</tr>
<tr>
<td>MA 540</td>
<td>Uncertainty Quantification</td>
</tr>
<tr>
<td>MA 587</td>
<td>Numerical Solution of PDEs – Finite Element Method</td>
</tr>
<tr>
<td>MA 798</td>
<td>Inverse Problems (offered this spring)</td>
</tr>
</tbody>
</table>
Some references: sensitivity analysis


Some references: OED and inverse problems


