A Posteriori Error Estimates and Adaptive Error Control for Probabilistic Sensitivity Analysis

Don Estep

Department of Statistics
Department of Mathematics
Colorado State University

Collaborators: T. Butler, A. Malqvist, S. Tavener

SIAM CS&E 2011
Introduction
Stochastic Sensitivity Analysis for Differential Equations

We consider a differential equation with stochastic parameters

\[
\begin{align*}
L(U; \lambda) &= 0, \\
U &= \text{data, \quad on boundaries}
\end{align*}
\]

\(\lambda\) is a random vector of parameters, known by distribution or a collection of samples

The target value is a quantity of interest \(Q(U)\), for a functional \(Q\)

\(Q(U)\) is an implicit function of \(\lambda\)

The goal is to describe the stochastic properties of \(Q(U)\)

For example, we may wish to compute the cumulative distribution function
Approximations Used for Sensitivity Analysis

There are two approximations used in practice

- **(Deterministic)** We compute numerical solutions $\tilde{U} \approx U$
- **(Stochastic)** We compute only a finite number of sample values $Q(U(\lambda_i))$

These sources interact: As parameter values vary, solution behavior varies, and so does numerical accuracy

As an example, consider a sensitivity analysis for the Lorenz equations

\[
\begin{align*}
\dot{U}_1 &= -10U_1 + 10U_2, \\
\dot{U}_2 &= \lambda U_1 - U_2 - U_1U_3, \\
\dot{U}_3 &= U_1U_2 - \frac{8}{3}U_3, \\
U(0) &= (-.69742, -7.008, 25.1377), \\
\lambda &\sim \text{Unif}(25, 31)
\end{align*}
\]
The quantity of interest is $U_2(10)$

We use 1000 samples and compute numerical solutions with two accuracies:

Solutions with fixed time steps

Solutions with error less than .00001

Numerical error ruins the results
We desire relatively accurate error estimates for the approximate statistics that account for both sources of error.

These are a posteriori in the sense that the estimates are computed after the approximate statistics are computed.

If we interpret the random variation in the parameters as statistical error, then combining sensitivity analysis with estimates of numerical and sampling error is an uncertainty quantification.
Analysis for a Cumulative Distribution Function
Approximating a cdf

We estimate the error in an approximate cumulative distribution function

\( U = U(\lambda) \) is the solution of a differential equation depending implicitly on a random parameter \( \lambda \)

\( Q(U) \) for some functional \( Q \) is the quantity of interest

We approximate the cdf

\[
F(t) = P(\{\lambda : Q(U(\lambda)) \leq t\}) = P(Q \leq t)
\]

using a finite number of approximate sample values \( \{\tilde{Q}^n\} \):

\[
\tilde{F}_N(t) = \frac{1}{N} \sum_{n=1}^{N} I(\tilde{Q}^n \leq t),
\]

\( I \) is the indicator function

\( \tilde{Q}^n = Q(\tilde{U}^n) \) is computed using the numerical solution \( \tilde{U}^n \approx U^n \)
A Decomposition of the Error

We use the sample distribution function

\[ F_N(t) = \frac{1}{N} \sum_{n=1}^{N} I(Q^n \leq t) \]

The error is decomposed into the stochastic and deterministic components

\[ |F(t) - \tilde{F}_N(t)| \leq |F(t) - F_N(t)| + |F_N(t) - \tilde{F}_N(t)| = I + II \]

We assume that there is an error estimate

\[ \tilde{Q}^n - Q^n \approx \mathcal{E}^n \]
Properties of a Sample cdf

\( F_N \) has very desirable properties, e.g.

- As a function of \( t \), \( F_N(t) \) is a distribution function
- For each fixed \( t \), \( F_N(t) \) is a random variable corresponding to the sample
- It is an unbiased estimator, i.e., \( E(F_N) \equiv E(F) \)
- \( N F_N(t) \) has exact binomial distribution for \( N \) trials and probability of success \( F(t) \)
- \( \text{Var}(F_N(t)) = F(t)(1 - F(t))/N \to 0 \) as \( N \to \infty \), and \( F_N \) converges in mean square to \( F \) as \( N \to \infty \)
Bounds on an Sample cdf

There are several ways to bound the error with high probability for large $N$

A standard measure of accuracy is the Kolmogorov–Smirnov distance

$$\sup_{t \in \mathbb{R}} |F_N(t) - F(t)|$$

A result that is uniform in $t$: For any $\epsilon > 0$,

$$P \left( \sup_{t \in \mathbb{R}} |F_N(t) - F(t)| \leq \left( \frac{\log(\epsilon^{-1})}{2N} \right)^{1/2} \right) \geq 1 - \epsilon$$

For example, using $\epsilon = .05$ gives a bound that holds with probability .95
Bounds on an Sample cdf

Since \( \{I(Q^n \leq t)\} \) are iid Bernoulli, the Chebyshev inequality implies that for \( \epsilon > 0 \),

\[
P \left( |F(t) - F_N(t)| \leq \left( \frac{F(t)(1 - F(t))}{N\epsilon} \right)^{1/2} \right) > 1 - \epsilon
\]

Since

\[
F(t)(1-F(t)) = F_N(t)(1-F_N(t)) + (F(t)-F_N(t))(1-F(t)-F_N(t))
\]

For \( \epsilon > 0 \),

\[
P \left( |F(t) - F_N(t)| \leq \left( \frac{F_N(t)(1 - F_N(t))}{N\epsilon} \right)^{1/2} + \frac{1}{2N\epsilon} \right) > 1 - \epsilon
\]
We next estimate the effect of using approximate values on the sample distribution function

$$\left|F_N(t) - \tilde{F}_N(t)\right| = \left|\frac{1}{N} \sum_{n=1}^{N} \left(I(\tilde{Q}^n \leq t) - I(Q \leq t)\right)\right|$$
Effect of Approximate Values

We next estimate the effect of using approximate values on the sample distribution function

\[ |F_N(t) - \tilde{F}_N(t)| = \frac{1}{N} \sum_{n=1}^{N} \left( I(\tilde{Q}^n \leq t) - I(Q \leq t) \right) \]

\[ = \frac{1}{N} \sum_{n=1}^{N} \left( I(Q^n + \mathcal{E}^n \leq t) - I(Q \leq t) \right) \]

\[ = \frac{1}{N} \sum_{n=1}^{N} \left( I(Q^n \leq t \leq Q^n + |\mathcal{E}^n|) \right) \]

\[ + \frac{1}{N} \sum_{n=1}^{N} \left( I(Q^n - |\mathcal{E}^n| \leq t \leq Q^n) \right) \]
Effect of Approximate Values

We obtain

$$\left| F_N(t) - \tilde{F}_N(t) \right| \leq \left| \frac{1}{N} \sum_{n=1}^{N} (I(Q^n - |E^n| \leq t \leq Q^n + |E^n|)) \right|$$

Expanding with $Q^n = \tilde{Q}^n - \mathcal{E}^n$ yields the computable estimate

$$\left| F_N(t) - \tilde{F}_N(t) \right| \leq \left| \frac{1}{N} \sum_{n=1}^{N} \left( I(\tilde{Q}^n - |\mathcal{E}^n| \leq t \leq \tilde{Q}^n + |\mathcal{E}^n|) \right) \right|$$

To obtain a computable estimate on I, we use

$$F_N(t)(1 - F_N(t)) = \tilde{F}_N(t)(1 - \tilde{F}_N(t))$$
$$+ (F_N(t) - \tilde{F}_N(t))(1 - F_N(t) - \tilde{F}_N(t))$$

along with bounds from the analysis of II
A Posteriori Error Estimate

Theorem
For any $\epsilon > 0$,

\[ |F(t) - \tilde{F}_N(t)| \leq \left( \frac{\tilde{F}_N(t)(1 - \tilde{F}_N(t))}{N \epsilon} \right)^{1/2} \]

\[ + 2 \left| \frac{1}{N} \sum_{n=1}^{N} \left( I(\tilde{Q}^n - |\mathcal{E}^n| \leq t \leq \tilde{Q}^n + |\mathcal{E}^n|) \right) \right| + \frac{1}{2N \epsilon} \]

with probability greater than $1 - \epsilon$
A Practical Point

Note that

\[
\left| \frac{1}{N} \sum_{n=1}^{N} \left( I(\tilde{Q}^n - |E^n| \leq t \leq \tilde{Q}^n + |E^n|) \right) \right|
\]

is itself an expected value

If \( M < N \) and \( N' = \{n_1 < \cdots < n_M\} \) is a set chosen randomly from \( \{1, \ldots, N\} \), we can use the unbiased estimator

\[
\left| \frac{1}{M} \sum_{n \in N'} \left( I(\tilde{Q}^n - |E^n| \leq t \leq \tilde{Q}^n + |E^n|) \right) \right|
\]

The error decreases as \( O(1/\sqrt{M}) \), which gives reasonable accuracy when \( N \) is large.
Numerical Example
Numerical Example

The estimate trades accuracy of the bound for increasing the probability that the bound is larger than the error.

We sample a given distribution and add random errors to each value.

<table>
<thead>
<tr>
<th>First computation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sample distribution</td>
</tr>
<tr>
<td>Error distribution</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Second computation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sample distribution</td>
</tr>
<tr>
<td>Error distribution</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Third computation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sample distribution</td>
</tr>
<tr>
<td>Error distribution</td>
</tr>
</tbody>
</table>
Approximate (solid line) and true (dashed line) cdf.

Left: first computation with $N = 5000$, $\delta = .001$.

Middle: second computation with $N = 2000$, $\delta = .0001$.

Right: third computation with $N = 500$, $\delta = .05$. 
Plots of the 95% confidence level bound.

Left: first computation with \( N = 5000, \delta = .001 \).

Middle: second computation with \( N = 2000, \delta = .0001 \).

Right: third computation with \( N = 500, \delta = .05 \).
Numerical Example

Left: Difference between the estimate and bound versus the error.
Right: Relative difference versus the error.
Application to a Stochastic Elliptic Problem
A Stochastic Elliptic Problem

Compute a quantity of interest $Q(U)$ where $U$ (a.s.) solves

$$
\begin{aligned}
-\nabla \cdot A \nabla U &= f, \quad x \in \Omega, \\
U &= 0, \quad x \text{ in } \partial \Omega,
\end{aligned}
$$

- $f \in L^2(\Omega)$
- $\Omega$ is a convex polygonal domain with boundary $\partial \Omega$
- $A$ is a stochastic function, uniformly coercive, uniformly bounded with piecewise smooth dependence on its inputs, and has continuous and bounded covariance functions

This implies $Q(U)$ is a random variable
The Numerical Solution Method

We use an unusual numerical solution method

We assume

\[ \mathcal{A} = a + A, \]

- \( a \) is a uniformly coercive, bounded, piecewise smooth function
- \( A \) is the random perturbation and \( |A(x)| < a(x) \)
- \( A(x) = \sum_{\kappa \in \mathcal{K}} A^\kappa \chi_\kappa(x), \quad x \in \Omega, \)
- \( \mathcal{K} \) is a finite nonoverlapping polygonal partition of \( \Omega \)
- \( \chi_\kappa \) is the characteristic function for the set \( \mathcal{K} \)
- \( (A^\kappa) \) is a random vector
Random perturbations on a $9 \times 9$ grid
The Numerical Solution Method

To compute numerical solutions, we use

- A finite element method to compute $\tilde{U}$
- Lion’s nonoverlapping domain decomposition to localize the elliptic solves to regions in which the random perturbation is constant
- A truncated Neumann series to approximate the solves of the local perturbed elliptic problems

This allows the statistics and computational “loops” to be “swapped”

The number of linear systems that have to be solved is independent of the number $N$ of random samples

Discretization parameters include mesh size $h$, number of domains $D$, number of iterations in D.D. $I$, and number of terms in the Neumann series $P$
The required a posteriori error estimate using an adjoint problem and variational analysis is

**Theorem**

\[ |Q(U) - Q(\tilde{U})| \lesssim |(f, \Phi) - (A\nabla \tilde{U}, \nabla \Phi)| + \text{h.o.t.} \]

where \( \Phi \) is an approximation to the solution of the adjoint problem

These estimates are robustly accurate
Generalized Adaptive Computation

We can use the a posteriori error bound to derive a generalized adaptive algorithm.

The algorithm balances work effort to control both stochastic and deterministic errors.

Since the bound consists of a sum of positive quantities, optimal effort is obtained by the Principle of Equidistribution.

We begin with a small number of samples, values for $I$ and $P$, and large mesh size.

We iteratively solve with increasing accuracy by adjusting the discretization parameters on the each iteration to balance sources of error.
Oil Reservoir Simulation

The model uses a realistic permeability field

We use a $27 \times 27$ uniform partition of $[0, 1] \times [0, 1]$ for $A$

$A$ is uniformly distributed in $\pm 20\%$ of the value of $a$

There is an “injection” site in the lower left corner of a square domain, where $f = 1$, while $f = -1$ in the opposite corner
A typical solution

The band of low permeability at $x \approx 0.2$ creates a large pressure drop.
Oil Reservoir Simulation

We set

- $\epsilon = .05$
- Error tolerance = .15
- Initially:
  - 100 domain decomposition iterations
  - 1 term in the Neumann series
  - 30 samples
- We fix a fine space mesh

The error tolerance is reached in 4 iterations

The final values are

- 800 domain decomposition iterations
- 4 terms in the Neumann series
- 240 samples
Oil Reservoir Simulation

D. Estep: A Posteriori Error Estimates and Adaptive Error Control for Probabilistic Sensitivity Analysis 33/35
Final Comments
Final Comments

We have derived analogous results for

- A stochastic moment $q$ with an unbiased estimator $Q$
- $p$-quantiles

The results apply to general differential equation, requiring only the availability of an accurate a posteriori error estimate for quantities of interest.

We have applied the results to the inverse problem of determining parameter distributions given observations on the output.