Uncertainty Propagation

Setting:

- We assume that we have determined distributions for parameters
  - e.g., Bayesian inference, prior experiments, expert opinion

\[
\begin{align*}
\dot{T}_1 &= \lambda_1 - d_1 T_1 - (1 - \epsilon) k_1 VT_1 \\
\dot{T}_2 &= \lambda_2 - d_2 T_2 - (1 - f\epsilon) k_2 VT_2 \\
\dot{T}_1^* &= (1 - \epsilon) k_1 VT_1 - \delta T_1^* - m_1 ET_1^* \\
\dot{T}_2^* &= (1 - f\epsilon) k_2 VT_2 - \delta T_2^* - m_2 ET_2^* \\
\dot{V} &= N_T \delta(T_1^* + T_2^*) - cV - [(1 - \epsilon) \rho_1 k_1 T_1 + (1 - f\epsilon) \rho_2 k_2 T_2] V \\
\dot{E} &= \lambda_E + \frac{b_E(T_1^* + T_2^*)}{T_1^* + T_2^* + K_b} E - \frac{d_E(T_1^* + T_2^*)}{T_1^* + T_2^* + K_d} E - \delta_E E
\end{align*}
\]

Goal: Construct statistics for quantities of interest

- e.g., Expected viral load in HIV patient with appropriate uncertainty intervals

- Note: Often involves moderate to high-dimensional integration

\[
\mathbb{E}[V(t)] = \int_{\mathbb{R}^6} V(t, q) \rho(q) dq
\]
Uncertainty Propagation

Issues:

• Uncertainty propagation and computation of statistical quantities of interest much more difficult for PDE models.

• e.g. Sinko-Streifer model: $u(t, x)$: Number of fish of size $x$ at time $t$

$$\frac{\partial u}{\partial t} + \frac{\partial (gu)}{\partial x} = -\mu u$$

$$g(t, x)u(t, x)\big|_{x=x_0} = \int_{x_0}^{x_1} k(t, \xi)u(t, \xi)d\xi$$

$$u(0, x) = \Phi(x)$$

Random Field Representation:

$$g(x) = \sum_{j=1}^{p} q_j \phi_j(x)$$

Quantity of Interest:

$$\mathbb{E}[u(t, x)] = \int_{\mathbb{R}^p} u(t, x, q)\rho(q)dq$$

Issues:

• How do we efficiently propagate input uncertainties through models? Surrogate models.

• How do we approximately integrate in moderate to high dimensions; e.g., $p = 50$-100? Monte Carlo sampling, sparse grid quadrature
Uncertainty Propagation

Sampling-Based Approaches:
• Quadrature: Monte Carlo, Latin hypercube, Sobol
• Interval definitions and construction
• Prediction intervals for HIV model via DRAM algorithm

Numerical Analysis-Based Approaches:
• Next presentation by C. Webster
Numerical Quadrature

Motivation: Computation of expected values requires approximation of integrals

\[ \mathbb{E}[u(t, x)] = \int_{\mathbb{R}^p} u(t, x, q) \rho(q) dq \]

Numerical Quadrature:

\[ \int_{\mathbb{R}^p} f(q) \rho(q) dq \approx \sum_{r=1}^{R} f(q^r) w^r \]

Questions:

• How do we choose the quadrature points and weights?
  – E.g., Newton-Cotes; e.g., trapezoid rule

\[ \int_{a}^{b} f(q) dq \approx \frac{h}{2} \left[ f(a) + f(b) + 2 \sum_{r=1}^{R-2} f(q^r) \right] \]

\[ q^r = a + hr, \quad h = \frac{b - a}{R - 1} \]
Numerical Quadrature

Motivation: Computation of expected values requires approximation of integrals

\[ \mathbb{E}[u(t, x)] = \int_{\mathbb{R}^p} u(t, x, q) \rho(q) dq \]

Numerical Quadrature:

\[ \int_{\mathbb{R}^p} f(q) \rho(q) dq \approx \sum_{r=1}^{R} f(q^r) w^r \]

Questions:

• How do we choose the quadrature points and weights?
  – E.g., Newton-Cotes, Gaussian algorithms
Numerical Quadrature

Numerical Quadrature:

\[ \int_{\mathbb{R}^p} f(q) \rho(q) \, dq \approx \sum_{r=1}^{R} f(q^r) w^r \]

Questions:

- Can we construct nested algorithms to improve efficiency?
  - E.g., employ Clenshaw-Curtis points
Numerical Quadrature

Questions:

- How do we reduce required number of points while maintaining accuracy?

**Tensored Grids:** Exponential growth

**Sparse Grids:** Same accuracy

<table>
<thead>
<tr>
<th>$p$</th>
<th>$R_\ell$</th>
<th>Sparse Grid $R$</th>
<th>Tensored Grid $R = (R_\ell)^p$</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>9</td>
<td>29</td>
<td>81</td>
</tr>
<tr>
<td>5</td>
<td>9</td>
<td>241</td>
<td>59,049</td>
</tr>
<tr>
<td>10</td>
<td>9</td>
<td>1581</td>
<td>$&gt; 3 \times 10^9$</td>
</tr>
<tr>
<td>50</td>
<td>9</td>
<td>171,901</td>
<td>$&gt; 5 \times 10^{47}$</td>
</tr>
<tr>
<td>100</td>
<td>9</td>
<td>1,353,801</td>
<td>$&gt; 2 \times 10^{95}$</td>
</tr>
</tbody>
</table>
Numerical Quadrature

Problem:
• Accuracy of methods diminishes as parameter dimension $p$ increases
• Suppose $f \in C^\alpha([0, 1]^p)$
• Tensor products: Take $R_\ell$ points in each dimension so $R = (R_\ell)^p$ total points
• Quadrature errors:
  
  Newton-Cotes: $E \sim \mathcal{O}(R_\ell^{-\alpha}) = \mathcal{O}(R^{-\alpha/p})$
  
  Gaussian: $E \sim \mathcal{O}(e^{-\beta R_\ell}) = \mathcal{O}
  \left(e^{-\beta \sqrt[p]{R}}\right)$
  
  Sparse Grid: $E \sim \mathcal{O}
  \left(R^{-\alpha \log (R) \frac{(p-1)(\alpha+1)}{(p-1)(\alpha+1)}}\right)$
Numerical Quadrature

Problem:

• Accuracy of methods diminishes as parameter dimension \( p \) increases
• Suppose \( f \in C^\alpha([0, 1]^p) \)
• Tensor products: Take \( R_\ell \) points in each dimension so \( R = (R_\ell)^p \) total points
• Quadrature errors:
  
  Newton-Cotes: \( E \sim O(R_\ell^{-\alpha}) = O(R^{-\alpha/p}) \)
  
  Gaussian: \( E \sim O(e^{-\beta R_\ell}) = O(e^{-\beta \sqrt{R}}) \)
  
  Sparse Grid: \( E \sim O(R^{-\alpha} \log(R) \frac{(p-1)(\alpha+1)}{p}) \)

• Alternative: Monte Carlo quadrature

\[
\int_{\mathbb{R}^p} f(q) \rho(q) dq \approx \frac{1}{R} \sum_{r=1}^{R} f(q^r), \quad E \sim \left( \frac{1}{\sqrt{R}} \right)
\]

• Advantage: Errors independent of dimension \( p \)
• Disadvantage: Convergence is very slow!
**Numerical Quadrature**

**Problem:**
- Accuracy of methods diminishes as parameter dimension $p$ increases
- Suppose $f \in C^\alpha([0, 1]^p)$
- Tensor products: Take $R_\ell$ points in each dimension so $R = (R_\ell)^p$ total points
- Quadrature errors:
  - Newton-Cotes: $E \sim \mathcal{O}(R_\ell^{-\alpha}) = \mathcal{O}(R^{-\alpha/p})$
  - Gaussian: $E \sim \mathcal{O}(e^{-\beta R_\ell}) = \mathcal{O}\left(e^{-\beta \sqrt{R}}\right)$
  - Sparse Grid: $E \sim \mathcal{O}\left(R^{-\alpha \log(R)} (p-1)(\alpha+1)\right)$

- Alternative: Monte Carlo quadrature
  $$\int_{\mathbb{R}^p} f(q) \rho(q) dq \approx \frac{1}{R} \sum_{r=1}^{R} f(q^r) , \quad E \sim \left(\frac{1}{\sqrt{R}}\right)$$

**Conclusion:** For high enough dimension $p$, monkeys throwing darts will beat Gaussian and sparse grid techniques!

- Advantage: Errors independent of dimension $p$
- Disadvantage: Convergence is very slow!
Monte Carlo Sampling Techniques

Issues:

- Very low accuracy and slow convergence
- Random sampling may not “randomly” cover space …
Monte Carlo Sampling Techniques

Issues:

• Very low accuracy and slow convergence
• Random sampling may not “randomly” cover space …
Bayesian Model Calibration – HIV Example

**Model:**
\[
\begin{align*}
\dot{T}_1 &= \lambda_1 - d_1 T_1 - (1 - \varepsilon) k_1 VT_1 \\
\dot{T}_2 &= \lambda_2 - d_2 T_2 - (1 - f \varepsilon) k_2 VT_2 \\
\dot{T}_1^* &= (1 - \varepsilon) k_1 VT_1 - \delta T_1^* - m_1 ET_1^* \\
\dot{T}_2^* &= (1 - f \varepsilon) k_2 VT_2 - \delta T_2^* - m_2 ET_2^* \\
\dot{V} &= N T_1 \delta (T_1^* + T_2^*) - cV - [(1 - \varepsilon) \rho_1 k_1 T_1 + (1 - f \varepsilon) \rho_2 k_2 T_2] V \\
\dot{E} &= \lambda_E + \frac{b_E (T_1^* + T_2^*)}{T_1^* + T_2^* + K_b} E - \frac{d_E (T_1^* + T_2^*)}{T_1^* + T_2^* + K_d} E - \delta_E E
\end{align*}
\]

**Verification:** Why do we trust results??
- Compare results from different algorithms; e.g., DRAM and Gibbs

**Parameter Chains and Densities:**
\[
q = [b_E, \delta, d_1, k_2, \lambda_1, K_b]
\]
Propagation of Uncertainty in Models – HIV Example

Parameter Densities:

- $b_L$
- $\delta$
- $d_1$
- $k_2$
- $\lambda_1$
- $K_b$

Techniques:

- Sample from parameter densities to construct prediction intervals for QoI.
- Slow convergence rate $O(1/\sqrt{M})$
- 100-fold more evaluations required to gain additional place of accuracy.
- Significant numerical analysis used to efficiently propagate densities. C. Webster
Forward Uncertainty Propagation: Linear Models

Linear Models: Analytic mean and variance relations

Example: Linear stress-strain relation

\[ \gamma_i = E e_i + E_2 e_i^3 + \varepsilon_i , \ i = 1, \ldots, n \]

Model Statistics:

Let \( \overline{E} \), \( \overline{E_2} \) and \( \text{var}(E) \), \( \text{var}(E_2) \) denote parameter means and variance. Then

\[
\mathbb{E}[E e_i + E_2 e_i^3] = \overline{E} e_i + \overline{E_2} e_i^3
\]

\[
\text{var}[E e_i + E_2 e_i^3] = e_i^2 \text{var}(E) + e_i^6 \text{var}(E_2) + 2e_i^4 \text{cov}(E, E_2)
\]

Response Statistics: Assume measurement errors uncorrelated from model response.

\[
\mathbb{E}[\gamma_i] = \overline{E} e_i + \overline{E_2} e_i^3
\]

\[
\text{var}[\gamma_i] = e_i^2 \text{var}(E) + e_i^6 \text{var}(E_2) + 2e_i^4 \text{cov}(E, E_2) + \text{var}(\varepsilon_i)
\]
Forward Uncertainty Propagation: Sampling Methods

Strategy: Randomly sample from parameter and measurement error distributions and propagate through model to quantify response uncertainty.

Advantages:
• Applicable to nonlinear models.
• Parameters can be correlated and non-Gaussian.
• Straight-forward to apply and convergence rate is independent of number of parameters.
• Can directly incorporate both parameter and measurement uncertainties.
• No additional cost for DRAM if interpolating.

Disadvantages:
• Very slow convergence rate: $\mathcal{O}(1/\sqrt{M})$ where $M$ is the number of samples.
• 100-fold more evaluations required to gain additional place of accuracy.
Confidence, Credible and Prediction Intervals

Data: $\mathcal{Y} = [\mathcal{Y}_1, \cdots, \mathcal{Y}_n]$ of iid random observations

Confidence Interval (Frequentist): A $100 \times (1 - \alpha)\%$ confidence interval for a fixed, unknown parameter $q_0$ is a random interval $[L_c(\mathcal{Y}), U_c(\mathcal{Y})]$, having probability at least $1 - \alpha$ of covering $q_0$ under the joint distribution of $\mathcal{Y}$.

Credible Interval (Bayesian): A $100 \times (1 - \alpha)\%$ credible interval is that having probability at least $1 - \alpha$ of containing $q$.

Strategy: Sample out of parameter density $\rho_Q(q)$
Confidence, Credible and Prediction Intervals

Data: $\gamma = [\gamma_1, \cdots, \gamma_n]$ of iid random observations

Prediction Interval: A $100 \times (1 - \alpha)$% prediction interval for a future observable $\gamma_{n+1}$ is a random interval $[L_c(\gamma), U_c(\gamma)]$ having probability at least $1 - \alpha$ of containing $\gamma_{n+1}$ under the joint distribution of $(\gamma, \gamma_{n+1})$.

Example: Consider linear model

$$\gamma_i = q_0 + q_1 x_i + \varepsilon_i, \ i = 1, \cdots, n$$