Projecting uncertainty through nonlinear ODEs

Youdong Lin¹, Mark Stadherr¹, George Corliss², Scott Ferson³

¹University of Notre Dame
²Marquette University
³Applied Biomathematics
Uncertainty

- **Artifactual uncertainty**
  - Too few polynomial terms
  - Numerical instability
  - Can be reduced by a better analysis

- **Authentic uncertainty**
  - Genuine unpredictability due to input uncertainty
  - Cannot be reduced by a better analysis
Uncertainty propagation

• We want the prediction to ‘break down’ if that’s what should happen

• But we don’t want artifactual uncertainty
  – Wrapping effect
  – Dependence problem
  – Repeated parameters
Problem

• Nonlinear ordinary differential equation (ODE)

\[ \frac{dx}{dt} = f(x, \theta) \]

with uncertain \( \theta \) and initial state \( x_0 \)

• Information about \( \theta \) and \( x_0 \) comes as
  – Interval ranges
  – Probability distribution
  – Something in between
Model
Initial states (range)
Parameters (range)

Notre Dame

List of constants plus remainder
Inside VSPODE

- Interval Taylor series (à la VNODE)
  - Dependence on time

- Taylor model
  - Dependence of parameters

(Comparable to COSY)
Representing uncertainty

• Cumulative distribution function (CDF)
  – Gives the probability that a random variable is smaller than or equal to any specified value

\[ F \text{ is the CDF of } \theta, \text{ if } F(z) = \text{Prob}(\theta \leq z) \]

We write: \( \theta \sim F \)
Example: uniform

\[ \text{Prob}(\theta \leq 2.5) = 0.75 \]
Another example: normal

\[ \text{Prob}(\theta \leq 2.5) = 0.90 \]
P-box (probability box)

Interval bounds on an CDF
Marriage of two approaches

Point value \rightarrow\ \text{Interval}

\downarrow

\text{Distribution} \rightarrow\ \text{P-box}
Probability bounds analysis

• All standard mathematical operations
  – Arithmetic (+, −, ×, ÷, ^, min, max)
  – Transformations (exp, ln, sin, tan, abs, sqrt, etc.)
  – Other operations (and, or, ≤, envelope, etc.)

• Quicker than Monte Carlo

• Guaranteed (automatically verified)
What are the bounds on the distribution of the sum of $A+B$?
# Cartesian product

<table>
<thead>
<tr>
<th>$A+B$ independence</th>
<th>$A \in [1, 3]$</th>
<th>$A \in [2, 4]$</th>
<th>$A \in [3, 5]$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$p_1 = 1/3$</td>
<td>$p_2 = 1/3$</td>
<td>$p_3 = 1/3$</td>
</tr>
<tr>
<td>$B \in [2, 8]$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$q_1 = 1/3$</td>
<td>$A+B \in [3, 11]$</td>
<td>$A+B \in [4, 12]$</td>
<td>$A+B \in [5, 13]$</td>
</tr>
<tr>
<td></td>
<td>prob=1/9</td>
<td>prob=1/9</td>
<td>prob=1/9</td>
</tr>
<tr>
<td>$B \in [6, 10]$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$q_2 = 1/3$</td>
<td>$A+B \in [7, 13]$</td>
<td>$A+B \in [8, 14]$</td>
<td>$A+B \in [9, 15]$</td>
</tr>
<tr>
<td></td>
<td>prob=1/9</td>
<td>prob=1/9</td>
<td>prob=1/9</td>
</tr>
<tr>
<td>$B \in [8, 12]$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>prob=1/9</td>
<td>prob=1/9</td>
<td>prob=1/9</td>
</tr>
</tbody>
</table>
$A + B$ under independence
When independence is untenable

Suppose $X \sim F$ and $Y \sim G$. The distribution of $X+Y$ is bounded by

$$\begin{bmatrix}
\sup_{z=x+y} \max(F(x) + G(y) - 1, 0), \\
\inf_{z=x+y} \min(F(x) + G(y), 1)
\end{bmatrix}$$

whatever the dependence between $X$ and $Y$

Similar formulas for operations besides addition
Example ODE

\[
dx_1/dt = \theta_1 x_1(1 - x_2) \\
dx_2/dt = \theta_2 x_2(x_1 - 1)
\]

What are the states at \( t = 10 \)?

\[
x_0 = (1.2, 1.1)^T \\
\theta_1 \in [2.99, 3.01] \\
\theta_2 \in [0.99, 1.01]
\]

VSPODE

- Constant step size \( h = 0.1 \), Order of Taylor model \( q = 5 \),
- Order of interval Taylor series \( k = 17 \), QR factorization
Calculation of $X_1$

\[
1.916037656181642 \times \theta_1^0 \times \theta_2^1 + 0.689979149231081 \times \theta_1^1 \times \theta_2^0 + \\
-4.690741189299572 \times \theta_1^0 \times \theta_2^2 + -2.275734193378134 \times \theta_1^1 \times \theta_2^1 + \\
-0.450416914564394 \times \theta_1^2 \times \theta_2^0 + -29.788252573360062 \times \theta_1^0 \times \theta_2^3 + \\
-35.200757076497972 \times \theta_1^1 \times \theta_2^2 + -12.401600707197074 \times \theta_1^2 \times \theta_2^1 + \\
-1.349694561113611 \times \theta_1^3 \times \theta_2^0 + 6.062509834147210 \times \theta_1^0 \times \theta_2^4 + \\
-29.50312865048253 \times \theta_1^1 \times \theta_2^3 + -25.744336555602068 \times \theta_1^2 \times \theta_2^2 + \\
-5.563350070358247 \times \theta_1^3 \times \theta_2^1 + -0.222000132892585 \times \theta_1^4 \times \theta_2^0 + \\
218.607042326120308 \times \theta_1^0 \times \theta_2^5 + 390.260443722081675 \times \theta_1^1 \times \theta_2^4 + \\
256.315067368131281 \times \theta_1^2 \times \theta_2^3 + 86.029720297509172 \times \theta_1^3 \times \theta_2^2 + \\
15.322357274648443 \times \theta_1^4 \times \theta_2^1 + 1.094676837431721 \times \theta_1^5 \times \theta_2^0 + \\
[ 1.1477537620811058, 1.1477539164945061 ]
\]

where $\theta$'s are centered forms of the parameters; $\theta_1 = \theta_1 - 3$, $\theta_2 = \theta_2 - 1$
uniform

normal
min, max, mean, var

precise
Calculation of $X_1$

$$1.916037656181642 \times \theta_1^0 \times \theta_2^1 + 0.689979149231081 \times \theta_1^1 \times \theta_2^0 +$$
$$-4.690741189299572 \times \theta_1^0 \times \theta_2^2 + -2.275734193378134 \times \theta_1^1 \times \theta_2^1 +$$
$$-0.450416914564394 \times \theta_1^2 \times \theta_2^0 + -29.788252573360062 \times \theta_1^0 \times \theta_2^3 +$$
$$-35.200757076497972 \times \theta_1^1 \times \theta_2^2 + -12.401600707197074 \times \theta_1^2 \times \theta_2^1 +$$
$$-1.349694561113611 \times \theta_1^3 \times \theta_2^0 + 6.062509834147210 \times \theta_1^0 \times \theta_2^4 +$$
$$-29.503128650484253 \times \theta_1^1 \times \theta_2^3 + -25.744336555602068 \times \theta_1^2 \times \theta_2^2 +$$
$$-5.563350070358247 \times \theta_1^3 \times \theta_2^1 + -0.222000132892585 \times \theta_1^4 \times \theta_2^0 +$$
$$218.607042326120308 \times \theta_1^0 \times \theta_2^5 + 390.260443722081675 \times \theta_1^1 \times \theta_2^4 +$$
$$256.315067368131281 \times \theta_1^2 \times \theta_2^3 + 86.029720297509172 \times \theta_1^3 \times \theta_2^2 +$$
$$15.322357274648443 \times \theta_1^4 \times \theta_2^1 + 1.094676837431721 \times \theta_1^5 \times \theta_2^0 +$$

$$[ 1.1477537620811058, 1.1477539164945061 ]$$

where $\theta$’s are centered forms of the parameters; $\theta_1 = \theta_1 - 3, \theta_2 = \theta_2 - 1$
Results for uniform p-boxes

![Graphs of X₁ and X₂]
Probability

normals

min, max, mean, var
Still repetitions of uncertainties

\[ 1.916037656181642 \times \theta_1^0 \times \theta_2^1 + 0.689979149231081 \times \theta_1^1 \times \theta_2^0 + \\ -4.690741189299572 \times \theta_1^0 \times \theta_2^2 + -2.275734193378134 \times \theta_1^1 \times \theta_2^1 + \\ -0.450416914564394 \times \theta_1^2 \times \theta_2^0 + -29.788252573360062 \times \theta_1^0 \times \theta_2^3 + \\ -35.200757076497972 \times \theta_1^1 \times \theta_2^2 + -12.401600707197074 \times \theta_1^2 \times \theta_2^1 + \\ -1.349694561113611 \times \theta_1^3 \times \theta_2^0 + 6.062509834147210 \times \theta_1^0 \times \theta_2^4 + \\ -29.503128650484253 \times \theta_1^1 \times \theta_2^3 + -25.744336555602068 \times \theta_1^2 \times \theta_2^2 + \\ -5.563350070358247 \times \theta_1^3 \times \theta_2^1 + -0.222000132892585 \times \theta_1^4 \times \theta_2^0 + \\ 218.607042326120308 \times \theta_1^0 \times \theta_2^5 + 390.260443722081675 \times \theta_1^1 \times \theta_2^4 + \\ 256.315067368131281 \times \theta_1^2 \times \theta_2^3 + 86.029720297509172 \times \theta_1^3 \times \theta_2^2 + \\ 15.322357274648443 \times \theta_1^4 \times \theta_2^1 + 1.09467683741721 \times \theta_1^5 \times \theta_2^0 + \\ [1.1477537620811058, 1.1477539164945061] \]
Subinterval reconstitution

• Subinterval reconstitution (SIR)
  – Partition the inputs into subintervals
  – Apply the function to each subinterval
  – Form the union of the results

• Still rigorous, but often tighter
  – The finer the partition, the tighter the union
  – Many strategies for partitioning

• Apply to each cell in the Cartesian product
Discretizations
Contraction from SIR

Best possible bounds reveal the authentic uncertainty
Precise distributions

- Uniform distributions (iid)
- Can be estimated with Monte Carlo simulation
  - 5000 replications
- Result is a p-box even though inputs are precise
Results are (narrow) p-boxes
Not automatically verified

• *Monte Carlo cannot yield validated results*
  – Though can be checked by repeating simulation

• Validated results can be achieved by modeling inputs with (narrow) p-boxes and applying probability bounds analysis

• Converges to narrow p-boxes obtained from infinitely many Monte Carlo replications
What are these distributions?

“bouquet”

“tangle”
Conclusions

- VSPODE is useful for bounding solutions of parametric nonlinear ODEs
- P-boxes and Risk Calc software are useful when distributions are known imprecisely
- Together, they rigorously propagate uncertainty through a nonlinear ODE

{Intervals, Distributions, P-boxes} -> {Initial states, Parameters}
To do

• Subinterval reconstitution accounts for the remaining repeated quantities

• Integrate it more intimately into VSPODE
  – Customize Taylor models for each cell

• Generalize to stochastic case ("tangle") when inputs are given as intervals or p-boxes
Acknowledgments

- U.S. Department of Energy (YL, MS)
- NASA and Sandia National Labs (SF)
More information

Mark Stadtherr (markst@nd.edu)

Scott Ferson (scott@ramas.com)
end