The Sensitivity Analysis of Mathematical Models Described by Differential Equations

Hossein Zivaripiran

Department of Computer Science

University of Toronto
Outline

- Modeling with Differential Equations (IVPs, DDEs)
- Modeling and Sensitivity Analysis
- Numerical Sensitivity Analysis of IVPs
- DDEs and Sensitivity Analysis
An Initial Value Problem (IVP) for Ordinary Differential Equations (ODEs)

\[ y'(t) = f(t, y(t)) \]

\[ y(t_0) = y_0 \]
An Initial Value Problem (IVP) for Ordinary Differential Equations (ODEs)

\[ y'(t) = f(t, y(t)) \]
\[ y(t_0) = y_0 \]

Retarded Delay Differential Equations (RDDEs)

\[ y'(t) = f(t, y(t), y(t - \sigma_1), \cdots, y(t - \sigma_\nu)) \text{ for } t_0 \leq t \leq t_F \]
\[ y(t) = \phi(t), \text{ for } t \leq t_0 \]

\[ \sigma_i = \sigma_i(t, y(t)) \geq 0 \] delay (constant / time dependent / state dependent)

\[ \phi(t) \] history function (constant / time dependent)
An Initial Value Problem (IVP) for Ordinary Differential Equations (ODEs)

\[ y'(t) = f(t, y(t)) \]
\[ y(t_0) = y_0 \]

Retarded Delay Differential Equations (RDDEs)

\[ y'(t) = f(t, y(t), y(t - \sigma_1), \ldots, y(t - \sigma_\nu)) \quad \text{for} \quad t_0 \leq t \leq t_F \]
\[ y(t) = \phi(t), \quad \text{for} \quad t \leq t_0 \]

\( \sigma_i = \sigma_i(t, y(t)) \geq 0 \) delay (constant / time dependent / state dependent)
\( \phi(t) \) history function (constant / time dependent)

Neutral Delay Differential Equations (NDDEs)

\[ y'(t) = f(t, y(t), y(t - \sigma_1), \ldots, y(t - \sigma_\nu), \]
\[ \quad y'(t - \sigma_{\nu+1}), \ldots, y'(t - \sigma_{\nu+\omega})) \quad \text{for} \quad t_0 \leq t \leq t_F \]
\[ y(t) = \phi(t), \quad y'(t) = \phi'(t), \quad \text{for} \quad t \leq t_0, \]
Modeling with DE - Some Areas of Application

<table>
<thead>
<tr>
<th>Area</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ecology</td>
<td>predator-prey</td>
</tr>
<tr>
<td>Epidemiology</td>
<td>spread of infections</td>
</tr>
<tr>
<td>Immunology</td>
<td>immune response models</td>
</tr>
<tr>
<td>HIV infection</td>
<td></td>
</tr>
<tr>
<td>Physiology</td>
<td>human respiration system</td>
</tr>
<tr>
<td>Neural Networks</td>
<td></td>
</tr>
<tr>
<td>Cell Kinetics</td>
<td></td>
</tr>
<tr>
<td>Chemical Kinetics</td>
<td>The Oregonator</td>
</tr>
<tr>
<td>Physics</td>
<td>Ring Cavity Lasers, two-body problem of electrodynamics</td>
</tr>
</tbody>
</table>
The Van der Pol Oscillator,

\[
\frac{d^2x}{dt^2} - \mu(1 - x^2)\frac{dx}{dt} + x = 0
\]

using \(y_1 = x\) and \(y_2 = \frac{dx}{dt}\),

\[
\begin{align*}
y_1'(t) &= y_2(t) \\
y_2'(t) &= \mu(1 - y_1^2)y_2 - y_1
\end{align*}
\]
The Van der Pol Oscillator with $\mu = 1, y_1(0) = 2, y_2(0) = 0$
An example of the famous Mackey-Glass DDEs proposed as a model for the production of white blood cells [Mackey and Glass 1977].

\[ y'(t) = \frac{2y(t - 2)}{1 + y(t - 2)^{9.65}} - y(t), \]

for \( t \) in \([0, 100]\). The history function is

\[ \phi(t) = 0.5 \quad \text{for} \quad t \leq 0. \]
The Mackey-Glass model
A neutral delay logistic Gause-type predator-prey system [Kuang 1991]

\[
y'_1(t) = y_1(t)(1 - y_1(t - \tau) - \rho y'_1(t - \tau)) - \frac{y_2(t)y_1(t)^2}{y_1(t)^2 + 1}
\]

\[
y'_2(t) = y_2(t) \left( \frac{y_1(t)^2}{y_1(t)^2 + 1} - \alpha \right)
\]

where \(\alpha = 1/10\), \(\rho = 29/10\) and \(\tau = 21/50\), for \(t\) in \([0, 30]\). The history functions are

\[
\phi_1(t) = \frac{33}{100} - \frac{1}{10}t
\]

\[
\phi_2(t) = \frac{111}{50} + \frac{1}{10}t
\]

for \(t \leq 0\).
The predator-prey model
- Classical Theory of Step by Step Integration for IVPs.
  - Runge-Kutta (RK).
  - Linear Multistep (LM).
Classical Theory of Step by Step Integration for IVPs.
- Runge-Kutta (RK).
- Linear Multistep (LM).

Continuous Solution using Polynomial Approximation.
- Continuous Runge-Kutta (CRK).
- Linear Multistep methods have natural approximating polynomials.
Classical Theory of Step by Step Integration for IVPs.
  ♦ Runge-Kutta (RK).
  ♦ Linear Multistep (LM).

Continuous Solution using Polynomial Approximation.
  ♦ Continuous Runge-Kutta (CRK).
  ♦ Linear Multistep methods have natural approximating polynomials.

DDEs : Combining an “interpolation” method (for evaluating delayed solution values) with an ODE integration method (for solving the resulting “ODE”).
Modeling with DE - Special Difficulties with DDEs

- Derivative Discontinuities
  In general
  \[
  \phi'(t_0) \neq f(t_0, \phi(t_0), \phi(t_0 - \sigma_1), \ldots, \phi(t_0 - \sigma_\nu))
  \]
  Due to the existence of delays, discontinuities propagate along the integration interval.
  Solution is smoothed for RDDEs but in general not for NDDEs.
  The RK and LM methods fail in presence of discontinuities.
  Treatment: Tracking Discontinuities and forcing them to be mesh points.

- Vanishing Delays
  \[
  \lim_{t \to t^*} \sigma_i(t, y(t)) = 0
  \]
  Causes the solver to choose a sequence of very small steps near \( t^* \). \( h \leq \sigma_i \)
  Treatment: Using Special Interpolants to pass \( t^* \)
■ Parameterized Models

♦ A parameterized IVP

\[
\begin{align*}
y'(t; p) &= f(t, y(t; p); p) \\
y(t_0) &= y_0(p)
\end{align*}
\]

For example \( p = [\mu] \) in the Van der Pol oscillator

\[
\begin{align*}
y'_1(t) &= y_2(t) \\
y'_2(t) &= \mu(1 - y_1^2)y_2 - y_1
\end{align*}
\]

♦ A simple parameterized DDE

\[
\begin{align*}
y'(t; p) &= f(t, y(t; p), y(t - \sigma(t; p)); p) \quad \text{for} \quad t_0(p) \leq t \\
y(t; p) &= \phi(t; p), \quad \text{for} \quad t \leq t_0(p)
\end{align*}
\]
Forward Sensitivity Analysis

The (first order) solution sensitivity with respect to the model parameter $p_i$ is defined as the vector

$$s_i(t; \mathbf{p}) = \left\{ \frac{\partial}{\partial p_i} \right\} y(t; \mathbf{p}), \quad (i = 1, \ldots, \mathcal{L})$$

The second order solution sensitivity with respect to the model parameters $p_i$ and $p_j$ is defined as the vector

$$r_{ij}(t; \mathbf{p}) = \left\{ \frac{\partial^2}{\partial p_j \partial p_i} \right\} s_i(t; \mathbf{p}) = \left\{ \frac{\partial^2}{\partial p_j \partial p_i} \right\} y(t; \mathbf{p}), \quad (i, j = 1, \ldots, \mathcal{L})$$
Adjoint Sensitivity Analysis

- We wish to evaluate the gradient \( \frac{\partial G}{\partial p_i} \) of

\[
G(p) = \int_{t_0}^{t_f} g(t, y, p) dt
\]

- or, alternatively, the gradient \( \frac{\partial g}{\partial p_i} \) of

\[
g(t, y, p)
\]

at time \( t_f \).

- More efficient if \( \text{dim}(g) < \text{dim}(y) \) and we need \( \frac{\partial}{\partial p} \) for many parameters.
Sensitivity information can be used to:

- Estimate which parameters are most influential in affecting the behavior of the simulation. Such information is crucial for
  - Experimental Design
  - Data Assimilation
  - Reduction of complex nonlinear models

- Study of Dynamical Systems: Periodic orbits, the Lyapunov exponents, chaos indicators, and bifurcation analysis are fundamental objects for the complete study of a dynamical system, and they require computation of the sensitivities with respect to the initial conditions of the problem.

- Evaluate optimization gradients and Jacobians in the setting of
  - Dynamic Optimization
  - Parameter Estimation
Finite Difference Approach

\[
\left\{ \frac{\partial}{\partial p_i} \right\} y(t; p) \approx \frac{y(t; p + e_i \Delta p_i) - y(t; p)}{\Delta p_i}
\]

Due to the rounding errors, the approximation is only \(O(\sqrt{Tol})\) with the best choice for \(\Delta p_i\).
 Finite Difference Approach

\[
\left\{ \frac{\partial}{\partial p_i} \right\} y(t; p) \approx \frac{y(t; p + e_i \Delta p_i) - y(t; p)}{\Delta p_i}
\]

Due to the rounding errors, the approximation is only \( O(\sqrt{Tol}) \) with the best choice for \( \Delta p_i \).

 Internal Differentiation

\[
y'(t; p) = f(t, y(t; p); p), \quad y(t_0) = y_0(p)
\]

\[
\downarrow
\]

Differentiation + Chain Rule + Clairaut’s Theorem

\[
\downarrow
\]

\[
s'_i = \frac{\partial f}{\partial y} s_i + \frac{\partial f}{\partial p_i}, \quad s_i(t_0) = \frac{\partial y_0(p)}{\partial p_i}, \quad (i = 1, \ldots, L)
\]
Finite Difference Approach

\[
\{ \frac{\partial}{\partial p_i} \} y(t; p) \approx \frac{y(t; p + e_i \Delta p_i) - y(t; p)}{\Delta p_i}
\]

Due to the rounding errors, the approximation is only \( O(\sqrt{Tol}) \) with the best choice for \( \Delta p_i \).

Internal Differentiation

\[
y'(t; p) = f(t, y(t; p); p), \quad y(t_0) = y_0(p)
\]

\[
\downarrow
\]

Differentiation + Chain Rule + Clairaut’s Theorem

\[
\downarrow
\]

\[
s'_i = \frac{\partial f}{\partial y} s_i + \frac{\partial f}{\partial p_i}, \quad s_i(t_0) = \frac{\partial y_0(p)}{\partial p_i}, \quad (i = 1, \ldots L)
\]

Taylor Series method using extended rules of AD [Barrio 2006].

\( \Rightarrow \) second(or higher)-order sensitivities.
The Van der Pol oscillator

\[
\begin{align*}
y'_1(t) &= y_2(t) \\
y'_2(t) &= \mu(1 - y_1^2)y_2 - y_1
\end{align*}
\]

\[
\begin{pmatrix}
    s'_{1(1)}(t) \\
    s'_{1(2)}(t)
\end{pmatrix}
= \begin{pmatrix}
    0 & 1 \\
    -2\mu y_1 y_2 - 1 & \mu(1 - y_1^2)
\end{pmatrix}
\begin{pmatrix}
    s_{1(1)} \\
    s_{1(2)}
\end{pmatrix}
+ \begin{pmatrix}
    0 \\
    (1 - y_1^2)y_2
\end{pmatrix}
\]

\[
\begin{align*}
s'_{1(1)}(t) &= s_{1(2)} \\
s'_{1(2)}(t) &= (-2\mu y_1 y_2 - 1)s_{1(1)} + \mu(1 - y_1^2)s_{1(2)} + (1 - y_1^2)y_2
\end{align*}
\]
The Van der Pol oscillator ($\Delta p = 0.2$). The decrease of $y_1$ at $t = 20$ can be explained as second order sensitivities being dominant ($\frac{\partial y_1}{\partial \mu} (t = 20) \gtrsim 0$, $\frac{\partial^2 y_1}{\partial \mu^2} (t = 20) < 0$).
\[ y'(t; p) = f(t, y(t; p), y(\alpha(t, y; p); p), y'(\alpha(t, y; p); p); p) \]
\[ y(t; p) = \phi(t; p), \text{ for } t \leq t_0(p) \]

\[ s'_i(t) = \frac{\partial f}{\partial y} s_i(t) + \frac{\partial f}{\partial y(\alpha_k)} \left( y'(\alpha_k) \left( \frac{\partial \alpha}{\partial y} s_i(t) + \frac{\partial \alpha}{\partial p} \right) + s_i(\alpha) \right) \]
\[ + \frac{\partial f}{\partial y'(\alpha)} \left( y''(\alpha) \left( \frac{\partial \alpha}{\partial y} s_i(t) + \frac{\partial \alpha}{\partial p} \right) + s'_i(\alpha) \right) \]
\[ + \frac{\partial f}{\partial p} \]
Hybrid ODE systems [Tolsma & Barton]

Continuous transition at $t = \lambda$,

$$y(\lambda^+) = y(\lambda^-)$$

$$\downarrow$$

$$\frac{\partial y}{\partial p_l}(\lambda^+) = \frac{\partial y}{\partial p_l}(\lambda^-) + (y'(\lambda^-) - y'(\lambda^+)) \frac{\partial \lambda}{\partial p_l}$$

Triggered by,

$$g(t, y, y'; p) = 0$$

$$\downarrow$$

$$\frac{\partial g}{\partial y'} \left( \frac{\partial}{\partial t} \left( \frac{\partial y}{\partial p_l} \right) + y'' \frac{\partial \lambda}{\partial p_l} \right) + \frac{\partial g}{\partial y} \left( \frac{\partial y}{\partial p_l} + y' \frac{\partial \lambda}{\partial p_l} \right) + \frac{\partial g}{\partial p_l} + \frac{\partial g}{\partial t} \frac{\partial \lambda}{\partial p_l} = 0$$
DDEs

Discontinuity points of $y'(t; p)$,

$$\Lambda(p) \equiv \{\lambda_1(p), \lambda_2(p), \ldots\}.$$  

Identified by,

$$\alpha(\lambda_{r+1}(p), y; p) = \lambda_r(p) \quad r = 1, 2, \ldots$$

Can be viewed as $\lambda_{r+1}(p)$ being a solution of,

$$\hat{g}(t, y; p) = \alpha(t, y; p) - \lambda_r(p) = 0.$$  

Using the result for hybrid ODEs,

$$\frac{\partial \lambda_{r+1}(p)}{\partial p_l} = -\frac{\partial \alpha}{\partial y} \frac{\partial y}{\partial p_l} + \frac{\partial \alpha}{\partial p_l} - \frac{\partial \lambda_r(p)}{\partial p_l}$$

for $r \geq 1$
Integration (first order sensitivities)

1. Initialize ($\lambda_1 = t_0(p)$).
2. $r \leftarrow 1$.
3. Integrate the equations up to a $C^1$-discontinuity point ($\lambda_{r+1}$).
4. Update the state variables (sensitivities) using
   \[
   \frac{\partial y}{\partial p_l}(\lambda^+_{r+1}) = \frac{\partial y}{\partial p_l}(\lambda^-_{r+1}) + (y'(\lambda^-_{r+1}) - y'(\lambda^+_{r+1})) \frac{\partial \lambda_{r+1}(p)}{\partial p_l}
   \]
5. $r \leftarrow r + 1$ and restart (3).
Mackey-Glass model

\[ y'(t) = \frac{2y(t-2)}{1 + y(t-2)^{9.65}} - y(t), \]

for \( t \) in \([0, 100]\). The history function is

\[ \phi(t) = 0.5 \quad \text{for} \quad t \leq 0. \]

Parameters are

\[ p = [\tau, n, A] \]

where

\[ y'(t) = \frac{2y(t-\tau)}{1 + y(t-\tau)^n} - y(t), \]

and

\[ \phi(t) = A \quad \text{for} \quad t \leq 0. \]
Mackey-Glass model
The predator-prey model

\[ \begin{align*}
    y'_1(t) &= y_1(t)(1 - y_1(t - \tau) - \rho y'_1(t - \tau)) - \frac{y_2(t)y_1(t)^2}{y_1(t)^2 + 1} \\
    y'_2(t) &= y_2(t) \left( \frac{y_1(t)^2}{y_1(t)^2 + 1} - \alpha \right)
\end{align*} \]

where \( \alpha = \frac{1}{10} \), \( \rho = \frac{29}{10} \) and \( \tau = \frac{21}{50} \), for \( t \) in \([0, 30]\). The history functions are

\[ \begin{align*}
    \phi_1(t) &= \frac{33}{100} - \frac{1}{10} t \\
    \phi_2(t) &= \frac{111}{50} + \frac{1}{10} t
\end{align*} \]

for \( t \leq 0 \). Parameters are

\[ p = \left[ \tau, \rho, \alpha, a, b, c, d \right] \]

where

\[ \begin{align*}
    \phi_1(t) &= a + b t \\
    \phi_2(t) &= c + d t
\end{align*} \]
The predator-prey model (structure-related parameters, $y_1$)
The predator-prey model (structure-related parameters, $y_2$)
The predator-prey model (history-related parameters, $y_1$)
■ The predator-prey model (history-related parameters, $y_2$)
Thank you for your attention!

Questions?