Introduction to Structural & Practical Identifiability

Marisa Eisenberg
University of Michigan, Ann Arbor
Identifiability

• Identifiability—Is it possible to uniquely determine the parameters from the data?

• Important problem in parameter estimation

• Many different approaches - statistics, applied math, engineering/systems theory

On Structural Identifiability

R. Bellman
Department of Mathematics, Electrical Engineering and Medicine
University of Southern California, Los Angeles, California

AND

K. J. Åström
Division of Automatic Control
Lund Institute of Technology
Lund, Sweden
Identifiability

- Practical vs. Structural
- Broad, sometimes overlapping categories
- Noisy vs. perfect data
- Example: $y = (m_1 + m_2)x + b$
- Unidentifiability - can cause serious problems when estimating parameters
- Identifiable combinations
Structural Identifiability

• Assumes best case scenario - data is known perfectly at all times

• Unrealistic!

• But, necessary condition for practical identifiability with real, noisy data
Structural Identifiability

- Reveals identifiable combinations and how to restructure the model so that it is identifiable
- Can give a priori information, help direct experiment design
Categories to consider

- Structural vs. practical identifiability
- Analytical vs. numerical methods
- Global vs. local results (in parameter space)
Key Concepts

- Identifiability vs. unidentifiability
- Practical vs. structural, local vs. global
- Can be in between, e.g. quasi-identifiable
- Identifiable Combinations
- Reparameterization
- Related questions: observability, distinguishability & model selection
Reparameterization

- Identifiable combinations - parameter combinations that can be estimated

- Once you know those, why reparameterize?

- Estimation issues - reparameterization provides a model that is input-output equivalent to the original but identifiable

- Often the reparameterized model has ‘sensible’ biological meaning (e.g. nondimensionalized, in terms of $R_0$, etc.)
Methods we’ll talk about today

• Differential Algebra Approach - structural identifiability, global, analytical method

• Fisher information matrix - structural or practical, local, analytical or numerical method

• Profile likelihood - structural or practical, local, numerical method
Simple Methods

- Simulated data approach

- If you have a small system, you can even plot the likelihood surface (typically can’t though—more on this with profile likelihoods)
Analytical Methods for Structural Identifiability
Analytical Methods for Structural Identifiability

- **Laplace transform** - linear models only

- **Taylor series approach** - more broad application, but only local info & may not terminate

- **Similarity transform approach** - difficult to make algorithmic, can be difficult to assess conditions for applying theorem

- **Differential algebra approach** - rational function ODE models, global info

Analytical Methods for Structural Identifiability

- **Laplace transform** - linear models only

- **Taylor series approach** - more broad application, but only local info & may not terminate

- **Similarity transform approach** - difficult to make algorithmic, can be difficult to assess conditions for applying theorem

- **Differential algebra approach** - rational function ODE models, global info

Differential Algebra Approach

- Basic idea: use substitution & differentiation to eliminate all variables except for observed output (y)
- Clear (divide by) the coefficient for highest derivative term(s)
- This is called the input-output equation(s)
- Contains all structural identifiability info for the model
Differential Algebra Approach

- Use the coefficients to solve for identifiability of the model
- If unidentifiable, determine identifiable combinations
- Find identifiable reparameterization of the model?
- Easier to see with an example—
2-Compartment Example

- Linear 2-Comp Model
  \[
  \begin{align*}
  \dot{x}_1 &= u + k_{12} x_2 - (k_{01} + k_{21}) x_1 \\
  \dot{x}_2 &= k_{21} x_1 - (k_{02} + k_{12}) x_2 \\
  y &= x_1 / V
  \end{align*}
  \]
  - state variables (x)
  - measurements (y)
  - known input (u) (e.g. IV injection)
2-Compartment Example

\[
\begin{align*}
\dot{x}_1 &= u + k_{12} x_2 - (k_{01} + k_{21}) x_1 \\
\dot{x}_2 &= k_{21} x_1 - (k_{02} + k_{12}) x_2 \\
y &= x_1 / V
\end{align*}
\]
2-Compartment Example

\[
\begin{align*}
\dot{x}_1 &= x_1 + V k_{12} x_2 - (k_{01} + k_{21}) x_1 \\
\dot{x}_2 &= k_{21} x_1 - (k_{02} + k_{12}) x_2
\end{align*}
\]
2-Compartment Example

\[ \dot{yV} = u + k_{12}x_2 - (k_{01} + k_{21})yV \]

\[ \dot{x}_2 = k_{21}x_1 - (k_{02} + k_{12})x_2 \]
2-Compartment Example

\[ \dot{y}_2 = k_{21} x_1 k_{12} (x_{202} - (k_{012} + k_{021}) x_{221}) y V \]
2-Compartment Example

\[
\dot{y} + (k_{01} + k_{21} + k_{12} + k_{02}) \dot{y} - \\
(k_{12} k_{21} - (k_{02} + k_{12})(k_{01} + k_{21}))y - u(k_{12} + k_{02})/V - \dot{u}/V = 0
\]
2-Compartment Example

\[
\begin{align*}
\dot{y} + (k_{01} + k_{21} + k_{12} + k_{02})\dot{y} & - (k_{12}k_{21} - (k_{02} + k_{12})(k_{01} + k_{21}))y - u\left(k_{12} + k_{02}\right)/V - \dot{u}/V = 0
\end{align*}
\]
2-Compartment Example

\[ \ddot{y} + \left( k_{01} + k_{21} + k_{12} + k_{02} \right) \dot{y} - \left( k_{12} k_{21} - (k_{02} + k_{12})(k_{01} + k_{21}) \right) y - u \left( \frac{k_{12} + k_{02}}{V} \right) = 0 \]

\[ \left( k_{01} + k_{21} + k_{12} + k_{02} \right) \]

\[ \left( k_{12} + k_{02} \right) / V \]

\[ \left( k_{12} k_{21} - (k_{02} + k_{12})(k_{01} + k_{21}) \right) \]

\[ 1 / V \]
2-Compartment Example

\[
\begin{align*}
&\left(k_{01} + k_{21} + k_{12} + k_{02}\right) \\
&\left(k_{12}k_{21} - (k_{02} + k_{12})(k_{01} + k_{21})\right) \\
&\left(k_{12} + k_{02}\right) / V \\
&1 / V
\end{align*}
\]
2-Compartment Example

\[
\frac{1}{V} \\
\left( k_{12} + k_{02} \right) / V \\
\left( k_{01} + k_{21} + k_{12} + k_{02} \right) \\
\left( k_{12}k_{21} - (k_{02} + k_{12})(k_{01} + k_{21}) \right)
\]
2-Compartment Example

\[
\frac{1}{V} = a_1
\]

\[
\left( k_{12} + k_{02} \right) / V = a_2
\]

\[
\left( k_{01} + k_{21} + k_{12} + k_{02} \right) = a_3
\]

\[
\left( k_{12} k_{21} - \left( k_{02} + k_{12} \right) \left( k_{01} + k_{21} \right) \right) = a_4
\]
2-Compartment Example

\[
\frac{1}{V} = a_1 \Rightarrow V = \frac{1}{a_1}
\]

\[
\left( k_{12} + k_{02} \right) / V = a_2
\]

\[
\left( k_{01} + k_{21} + k_{12} + k_{02} \right) = a_3
\]

\[
\left( k_{12} k_{21} - (k_{02} + k_{12})(k_{01} + k_{21}) \right) = a_4
\]
2-Compartment Example

\[
1/V = a_1 \implies V = 1/a_1
\]

\[
(k_{12} + k_{02})/V = a_2
\]

\[
(k_{01} + k_{21} + k_{12} + k_{02}) = a_3
\]

\[
(k_{12}k_{21} - (k_{02} + k_{12})(k_{01} + k_{21})) = a_4
\]
2-Compartment Example

\[
1 / V = a_1 \Rightarrow V = 1 / a_1
\]

\[
\left( k_{12} + k_{02} \right) / V = a_2
\]

\[
\left( k_{01} + k_{21} + k_{12} + k_{02} \right) = a_3
\]

\[
\left( k_{12} k_{21} - (k_{02} + k_{12})(k_{01} + k_{21}) \right) = a_4
\]
2-Compartment Example

\[
1 / V = a_1 \Rightarrow V = 1 / a_1
\]

\[
\frac{(k_{12} + k_{02})}{V} = a_2
\]

\[
\left( k_{01} + k_{21} - k_{12} + k_{02} \right) = a_3
\]

\[
\left( k_{12} k_{21} - (k_{02} + k_{12})(k_{01} + k_{21}) \right) = a_4
\]
2-Compartment Example

\[
\frac{1}{V} = a_1 \Rightarrow V = \frac{1}{a_1}
\]

\[
\frac{(k_{12} + k_{02})}{V} = a_2
\]

\[
\left( k_{01} + k_{21} - k_{12} + k_{02} \right) = a_3
\]

\[
\left( k_{12} k_{21} \right) - \left( k_{02} + k_{12} \right) \left( k_{01} + k_{21} \right) = a_4
\]

Unidentifiable
2-Compartment Example

\[ \dot{x}_1 = u + k_{12}x_2 - (k_{01} + k_{21})x_1 \]
\[ \dot{x}_2 = k_{21}x_1 - (k_{02} + k_{12})x_2 \]
\[ y = \frac{x_1}{V} \]
2-Compartment Example

\[
\begin{align*}
\dot{x}_1 &= u + k_{12}x_2 - (k_{01} + k_{21})x_1 \\
\dot{x}_2 &= k_{21}x_1 - (k_{02} + k_{12})x_2 \\
y &= x_1 / V
\end{align*}
\]

Let \( x_2 = k_{12}x_2 \)
2-Compartment Example

\[
\begin{align*}
\dot{x}_1 &= u + k_{12}x_2 - (k_{01} + k_{21})x_1 \\
\dot{x}_2 &= k_{21}x_1 - (k_{02} + k_{12})x_2 \\
y &= x_1 / V
\end{align*}
\]

Let \( x_2 = k_{12}x_2 \)

\[
\begin{align*}
\dot{x}_1 &= u + x_2 - (k_{01} + k_{21})x_1 \\
\dot{x}_2 &= k_{12}k_{21}x_1 - (k_{02} + k_{12})x_2 \\
y &= x_1 / V
\end{align*}
\]

Or add information about one of the parameters
Differential Algebra Approach

- View model & measurement equations as differential polynomials
- Reduce the equations using Gröbner bases, characteristic sets, etc. to eliminate unmeasured variables (x)
- Yields input-output equation(s) only in terms of known variables (y, u)
- Use coefficients to test model identifiability

Differential Algebra Approach

• From the coefficients, can often determine:
  • Simpler forms for identifiable combinations
  • Identifiable reparameterizations for model
  • Not always easy by eye—use Gröbner bases & other methods to simplify
  • Note about scaling as a useful first step (cf. nondimensionalization)
Differential Algebra Approach

• Convenient as a way to prove identifiability results for relatively broad classes of models
Numerical Methods for Identifiability Analysis
Numerical Approaches to Identifiability

• Analytical approaches can be slow, sometimes have limited applicability

• Wide range of numerical approaches

  • Sensitivities/Fisher Information Matrix
  • Profile Likelihood

  • Many others (e.g. Bayesian approaches, etc.)
Numerical Approaches to Identifiability

- Most can do both structural & practical identifiability
- Wide range of applicable models, often (relatively) fast
- Typically only local
Simple Simulation Approach

• Simulate data using a single set of ‘true’ parameter values

• Without noise for structural identifiability

• With noise for practical identifiability (in this case generate multiple realizations of the data)
Simple Simulation Approach

• Fit your simulated data from multiple starting points and see where your estimates land

• If they all return to the ‘true’ parameters, likely identifiable, if they do not—may be problems

• Note—unidentifiability when estimating with ‘perfect’, noise-free simulated data is most likely structural
transmission pathway. To assess the efficacy of different intervention strategies, this was particularly relevant for waterborne disease models because of the public health importance of distinguishing multiple transmission pathways. To help guide public health practice, modeling efforts have increasingly been used to help understand the transmission dynamics of waterborne pathogens. This is particularly relevant for waterborne disease models because of the public health importance of distinguishing multiple transmission pathways.

The water data significantly decreases the variability on estimates of the pathogen shedding rate, which was not available using case measurements in the water. The inclusion of a second series of data measurements alone. The pathogen shedding rate, which was not available using case measurements in the water also gives additional information on the variabilities of the parameters, particularly those involved in the waterborne transmission. The variability of parameters involved in the waterborne transmission has been of interest and commonly encountered in public health applications. This has been particularly important in the cholera epidemic in Haiti (Garcia et al., 2011), making the issue of parameter identifiability an important and commonly encountered problem in public health applications.

To determine the pathogen shedding rate, which was not available using case measurements in the water also gives additional information on the variabilities of the parameters, particularly those involved in the waterborne transmission. The variability of parameters involved in the waterborne transmission has been of interest and commonly encountered in public health applications. This has been particularly important in the cholera epidemic in Haiti (Garcia et al., 2011), making the issue of parameter identifiability an important and commonly encountered problem in public health applications.

To determine the pathogen shedding rate, which was not available using case measurements in the water also gives additional information on the variabilities of the parameters, particularly those involved in the waterborne transmission. The variability of parameters involved in the waterborne transmission has been of interest and commonly encountered in public health applications. This has been particularly important in the cholera epidemic in Haiti (Garcia et al., 2011), making the issue of parameter identifiability an important and commonly encountered problem in public health applications.

Table 1

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\mu$</td>
<td>0.1</td>
</tr>
<tr>
<td>$\beta_I$</td>
<td>0.2</td>
</tr>
<tr>
<td>$\gamma$</td>
<td>0.3</td>
</tr>
<tr>
<td>$\xi$</td>
<td>0.4</td>
</tr>
<tr>
<td>$\beta_W$</td>
<td>0.5</td>
</tr>
</tbody>
</table>

Fig. 8. Examples of simulated data sets using least squares estimation with four noise distributions (left to right): Poisson, Gaussian, negative binomial with variance equal to 5 times the mean, and negative binomial with variance equal to 50 times the mean.

Fig. 7. Scatterplots showing parameter estimates for 100 simulated data sets using least squares estimation for Poisson noise. True parameters (indicated as given in Table 1).

Fig. 10. This yields a wider range of parameter estimates, as given in Table 1.
Parameter Sensitivities

- Output sensitivity matrix (design matrix)
- Closely related to identifiability
- Insensitive parameters
- Dependencies between columns

\[ X = \begin{pmatrix}
\frac{\partial y(t_1)}{\partial p_1} & \cdots & \frac{\partial y(t_1)}{\partial p_n} \\
\vdots & \ddots & \vdots \\
\frac{\partial y(t_m)}{\partial p_1} & \cdots & \frac{\partial y(t_m)}{\partial p_n}
\end{pmatrix} \]
Fisher Information Matrix

• FIM - $N_p \times N_p$ matrix

\[ [\mathcal{I}(\theta)]_{i,j} = \mathbb{E} \left[ \left( \frac{\partial}{\partial \theta_i} \log f(X; \theta) \right) \left( \frac{\partial}{\partial \theta_j} \log f(X; \theta) \right) \left| \theta \right. \right] \]

• Useful in testing practical & structural ID - represents amount of information that the output $y$ contains about parameters $p$

• Cramer-Rao Bound: $\text{FIM}^{-1} \leq \text{Cov}(p)$

• $\text{Rank(FIM)} = \text{number of identifiable parameters/combinations}$
Fisher Information Matrix

- For identifiability analysis, often more useful to consider (sometimes denoted the sensitivity FIM):

\[ F = X^T X \]

\[ X = \begin{pmatrix} \frac{\partial y(t_1)}{\partial p_1} & \cdots & \frac{\partial y(t_1)}{\partial p_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial y(t_m)}{\partial p_1} & \cdots & \frac{\partial y(t_m)}{\partial p_n} \end{pmatrix} \]

- Can also derive as usual FIM with assumption of normally distributed measurement error with fixed variance (e.g. 1)
Identifiability & the FIM

- Covariance matrix/confidence interval estimates from Cramér-Rao bound: \( \text{Cov} \geq \text{FIM}^{-1} \)
  - e.g. large confidence interval \( \Rightarrow \) probably at least practically unID
  - Often can detect structural unID as ‘near-infinite’ (gigantic) variances in \( \text{Cov} \sim \text{FIM}^{-1} \)
Identifiability & the FIM

- **Rank of the FIM** is number of identifiable combinations/parameters - can do a lot by testing sub-FIMs and versions of the FIM

- Use FIM to find blocks of related parameters & how many to fix (not estimate)

- Identifiable combinations - can often see what parameters are related, but don’t know form

  - Interaction of combinations
Connections with sloppiness, active subspaces

- Use eigenvalues & eigenvectors to find sensitive/identifiable/stiff/active directions vs. insensitive/unidentifiable/sloppy/inactive

- E.g. in active subspaces, from Constantine (2015):

\[ C = \int (\nabla f)(\nabla f)^T \rho(\theta) \, d\theta \]

\[ \nabla f = \begin{bmatrix} \frac{\partial f}{\partial \theta_1} \\ \vdots \\ \frac{\partial f}{\partial \theta_n} \end{bmatrix} \]

- Can write this as the weighted average sFIM:

\[ C = \int F(f; \theta) \rho(\theta) \, d\theta. \]

- In FIM form, QOI could be univariate or multivariate
Identifiability & the FIM

- But, be careful—FIM is local & asymptotic
- Local approximation of the curvature of the likelihood

Brouwer, Meza, Eisenberg 2017

Raue et al. 2010
Maximum Likelihood
Parameter Estimation

- Basic idea: parameters that give model behavior that more closely matches data are ‘best’ or ‘most likely’

- Frame this from a statistical perspective (inference, regression)

- Can determine ‘most likely’ parameters or distribution, confidence intervals, etc.
How to frame this statistically?

• Maximum Likelihood Approach

• Idea: rewrite the ODE model as a statistical model, where we suppose we know the general form of the density function but not the parameter values.

• Then if we knew the parameters we could calculate probability of a particular observation/data:

\[ P(z \mid p) \]

data  parameters
Maximum Likelihood

- **Likelihood Function**

\[ P(z \mid p) = f(z, p) = L(p \mid z) \]

- Re-think the distribution as a function of the data instead of the parameters

- E.g. \[ f(z \mid \mu, \sigma^2) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(z - \mu)^2}{2\sigma^2}\right) = L(\mu, \sigma^2 \mid z) \]

- Find the value of p that maximizes \( L(p \mid z) \) - this is the maximum likelihood estimate (MLE) (most likely given the data)
Likelihood Function

PDF given a parameter value
Likelihood Function

Move the parameter and the distribution shifts.
Likelihood Function
Likelihood Function
Likelihood Function

Parameter value vs. Data value

PDF given a parameter value
Likelihood Function

Likelihood function given data
Example - ODE Model with Gaussian Error

• Model:
  
  \[ \dot{x} = f(x, t, p) \]
  
  \[ y = g(x, t, p) \]

• Suppose data is taken at times \( t_1, t_2, \ldots, t_n \)

• Data at \( t_i = z_i = y(t_i) + e_i \)

• Suppose error is gaussian and unbiased, with known variance \( \sigma^2 \) (can also be considered an unknown parameter)
Example - ODE Model with Gaussian Error

- The measured data $z_i$ at time $i$ can be viewed as a sample from a Gaussian distribution with mean $y(x, t, p)$ and variance $\sigma^2$

- Suppose all measurements are independent (is this realistic?)
Example - ODE Model with Gaussian Error

- Then the likelihood function can be calculated as:

  Gaussian PDF: \( f(z_i \mid \mu, \sigma^2) = \frac{1}{\sqrt{2\pi}\sigma} \exp \left( -\frac{(z_i - \mu)^2}{2\sigma^2} \right) \)
Example - ODE Model with Gaussian Error

Then the likelihood function can be calculated as:

Gaussian PDF:

\[
f(z_i | \mu, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma}} \exp\left(-\frac{(z_i - \mu)^2}{2\sigma^2}\right)
\]

Formatted for model:

\[
f(z_i | y(x, t, p), \sigma^2) = \frac{1}{\sqrt{2\pi\sigma}} \exp\left(-\frac{(z_i - y(t, p))^2}{2\sigma^2}\right)
\]
Example - ODE Model with Gaussian Error

• Then the likelihood function can be calculated as:

  Gaussian PDF: \[ f(z_i \mid \mu, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma}} \exp\left( -\frac{(z_i - \mu)^2}{2\sigma^2} \right) \]

  Formatted for model:
  \[
  f(z_i \mid y(x, t_i, p), \sigma^2) = \frac{1}{\sqrt{2\pi\sigma}} \exp\left( -\frac{(z_i - y(t_i, p))^2}{2\sigma^2} \right)
  \]

  Likelihood function assuming independent observations:
  \[
  L(y(t_i, p), \sigma^2 \mid z_1, \ldots, z_n) = f(z_1, \ldots, z_n \mid y(t_i, p), \sigma^2) = \prod_{i=1}^{n} f(z_i \mid y(t_i, p), \sigma^2)
  \]
Example - ODE Model with Gaussian Error

- It is often more convenient to minimize the Negative Log Likelihood (-LL) instead of maximizing the Likelihood

- Log is well behaved, minimization algorithms common

\[-LL = -\ln\left(\left(\frac{1}{2\pi\sigma^2}\right)^{n/2} \exp\left(- \frac{\sum_{i=1}^{n} (z_i - y(t_i, p))^2}{2\sigma^2}\right)\right)\]
Example - ODE Model with Gaussian Error

\[-LL = \frac{n}{2} \ln(2\pi) + n \ln(\sigma) + \frac{\sum_{i=1}^{n}(z_i - y(t_i, p))^2}{2\sigma^2}\]

If \(\sigma\) is known, then first two terms are constants & will not be changed as \(p\) is varied—so we can minimize only the 3rd term and get the same answer

\[
\min_p (-LL) = \min_p \left( \frac{\sum_{i=1}^{n}(z_i - y(t_i, p))^2}{2\sigma^2} \right)
\]
Example - ODE Model with Gaussian Error

• Similarly for denominator:

\[
\min_p (-LL) = \min_p \left\{ \sum_{i=1}^{n} \frac{(z_i - y(t_i, p))^2}{2\sigma^2} \right\} = \min_p \left\{ \sum_{i=1}^{n} (z_i - y(t_i, p))^2 \right\}
\]

• This is just least squares!

• So, least squares is equivalent to the ML estimator when we assume a constant known variance
Maximum Likelihood Summary for ODEs

- Can calculate other ML estimators for different distributions
- Not always least squares-ish! (mostly not)
- Although surprisingly, least squares does fairly decently a lot of the time
Example - Poisson ML

• Model:
  \[ \dot{x} = f(x, t, p) \]
  \[ y = g(x, t, p) \]

• Data \( z_i \) is assumed to be Poisson with mean \( y(t_i) \)

• Assume all data points are independent

• Poisson PMF:
  \[ f(z_i \mid y(t_i)) = \frac{y(t_i)^{z_i} e^{-y(t_i)}}{z_i!} \]
Poisson ML

- Negative log likelihood:

$$-LL = - \ln \left( \prod_{i=1}^{n} \frac{y(t_i)^{z_i} e^{-y(t_i)}}{z_i!} \right)$$

$$= - \sum_{i=1}^{n} \ln \left( \frac{y(t_i)^{z_i} e^{-y(t_i)}}{z_i!} \right)$$

$$= - \sum_{i=1}^{n} z_i \ln(y(t_i)) + \sum_{i=1}^{n} y(t_i) + \sum_{i=1}^{n} \ln(z_i)$$

- Last term is constant
Example - Poisson ML

• Poisson ML Estimator:

\[
\min_p (-LL) = \min_p \left( -\sum_{i=1}^{n} z_i \ln(y(t_i)) + \sum_{i=1}^{n} y(t_i) \right)
\]

• Other common distributions - negative binomial (overdispersion), zero-inflated poisson or negative binomial, etc.
Maximum Likelihood Summary for ODEs

- Basic approach - suppose only measurement error
- Data is given by distribution where model output is the mean
- Suppose each time point of data is independent
- Use PDF/PMF to calculate the likelihood
- Take the negative log likelihood, minimize this over the parameter space
Profile Likelihoods
Profile Likelihood

- Want to examine likelihood surface, but often high-dimensional

- Basic Idea: ‘profile’ one parameter at a time, by fixing it to a range of values & fitting the rest of the parameters

- Gives best fit at each point

- Evaluate curvature of likelihood to determine confidence bounds on parameter (and to evaluate parameter uncertainty)
Profile Likelihood

• Choose a range of values for parameter $p_i$

• For each value, fix $p_i$ to that value, and fit the rest of the parameters

• Report the best likelihood/RSS/cost function value for that $p_i$ value

• Plot the best likelihood values for each value of $p_i$—this is the profile likelihood
Profile Likelihoods

- Identifiable
- Structurally unidentifiable
- Practically unidentifiable
Profile Likelihood & ID

- Can generate confidence bounds based on the curvature of the profile likelihood
- Flat or nearly flat regions indicate identifiability issues
- Can generate simulated ‘perfect’ data to test structural identifiability
Profile Likelihood

- Can also help reveal the form of identifiable combinations
  - Look at relationships between parameters when profiling
  - However, can be problematic when too many degrees of freedom
- Similar to pairwise plots with sampling-based methods (e.g. MCMC)
Dengue Model Example

\[\frac{dS_h}{dt} = \mu(1 - S_h) - \beta^*_m S_h I_m\]
\[\frac{dE_h}{dt} = \beta^*_m S_h I_m - \alpha E_h - \mu E_h\]
\[\frac{dI_h}{dt} = \alpha E_h - \eta I_h - \mu I_h\]
\[\frac{dR_h}{dt} = \eta I_h - \mu R_h\]
\[\frac{dA}{dt} = \xi^*(S_m + E_m + I_m)(1 - A) - \mu_a^* A\]
\[\frac{dS_m}{dt} = A - \beta_{hm} S_m I_h - \mu_m S_m\]
\[\frac{E_m}{dt} = \beta_{hm} S_m I_h - \gamma E_m - \mu_m E_m\]
\[\frac{I_m}{dt} = \gamma E_m - \mu_m I_m\]
Measurement Model & Structural Identifiability

- Measure human incidence data, \( y = \kappa_h \alpha E_h \), integrated to weekly incidence
- Differential algebra approach and FIM-based approaches show structural identifiability
A

Simulated epidemic curve
2010 dengue incidence in Kaohsiung

\[ \beta_{mh} = 14.15 \]
\[ \xi = 2.03 \]
\[ \beta_{hm} = 0.03 \]
\[ \mu_a = 4.18 \]
\[ \mu_m = 0.32 \]
\[ \kappa_h = 1546.74 \]
What about practical identifiability?
What about practical identifiability?
How does this affect $R_0$?

$$R_0 = \sqrt{\frac{S_m \alpha \beta_{hm} \beta_{mh} \gamma}{(\alpha + \mu)(\eta + \mu)(\gamma + \mu_m)\mu_m}}.$$
Practically Identifiable Combinations

\[ R_0 = \sqrt{\frac{S_m \alpha \beta_{hm} \beta_{mh} \gamma}{(\alpha + \mu)(\eta + \mu)(\gamma + \mu_m)\mu_m}}. \]
Intervention predictions

Fit1:
\[ \begin{align*}
\beta_{mh} &= 14.15 \\
\xi &= 2.03 \\
\beta_{hm} &= 0.03 \\
\mu_a &= 4.18 \\
\mu_m &= 0.32 \\
\kappa_h &= 1546.74
\end{align*} \]

Fit2:
\[ \begin{align*}
\beta_{mh} &= 38.10 \\
\xi &= 0.13 \\
\beta_{hm} &= 0.02 \\
\mu_a &= 0.15 \\
\mu_m &= 0.42 \\
\kappa_h &= 1625.42
\end{align*} \]
Some potential issues

\[
\begin{align*}
\dot{x}_1 &= k_1 x_2 - (k_2 + k_3 + k_4) x_1 \\
\dot{x}_2 &= k_4 x_1 - (k_5 + k_1) x_2 \\
y &= x_1 / V
\end{align*}
\]
Example Model

\[
\begin{align*}
\dot{x}_1 &= k_1 x_2 - (k_2 + k_3 + k_4)x_1 \\
\dot{x}_2 &= k_4 x_1 - (k_5 + k_1)x_2 \\
y &= x_1 / V
\end{align*}
\]
Example Model

\[
\dot{x}_1 = k_1 x_2 - (k_2 + k_3 + k_4) x_1
\]
\[
\dot{x}_2 = k_4 x_1 - (k_5 + k_1) x_2
\]
\[
y = x_1 / V
\]

Eisenberg & Hayashi 2014, in review
Conclusions

• Many related questions and potential issues when connecting models to data: observability, distinguishability & model selection, reparameterization & model/parameter reduction, and more

• Many other methods! (eigenvalues of FIM, sloppy models, active subspaces, Bayesian methods, & more)

• Depending on amount of data, model complexity, model type, and more, different approaches may work in different circumstances
Conclusions

• Identifiability—an important question to address when estimating model parameters

• Common problem in math bio (identifiability-robustness tradeoff)

• Many approaches, both numerical and analytical
Cholera epidemics
Cholera

- Waterborne disease caused by bacterium *V. cholerae*
- Profuse, watery diarrhea, vomiting, dehydration
- Up to 50% fatal if untreated
- Infection-derived immunity
- Treatment: oral or IV rehydration
- Direct & environmental transmission

Yemen, 2017 (Credit: Yahya Arhab/European Pressphoto Agency)
SIWR Model

SIWR Model Equations

\[
\frac{ds}{dt} = \mu - \beta_w ws - \beta_i si - \mu s
\]

\[
\frac{di}{dt} = \beta_w ws + \beta_i si - \gamma i - \mu i
\]

\[
\frac{dw}{dt} = \xi (i - w)
\]

\[
\frac{dr}{dt} = \gamma i - \mu r
\]

\[y = ki\]

\[R_0 = \frac{\beta_I + \beta_W}{\mu + \gamma}\]
Structural Identifiability

- SIWR model is \textit{globally structurally identifiable}
  - but, identifiability can be lost if $\xi \to \infty$
Practical Identifiability
SIWR Identifiability

- SIWR model structurally identifiable
  - Identifiability can be lost if $\xi \to \infty$
- Practical identifiability - dependence between $\beta_W$ and $\xi$
  - Water measurements (pathogens or adding environmental forcing) can improve practical & structural identifiability

Eisenberg, Robertson, Tien 2013, JTB
Haiti Cholera Outbreak

The severity of the cholera outbreak in Haiti emphasizes the fundamental importance of clean water and sanitation for health and well-being. One of the goals of this proposal is to highlight this fact by explicitly incorporating data on water quality and sanitation into dynamic models of cholera epidemics. Given the extensive body of research in mathematical epidemiology, surprisingly little work has been done on modeling waterborne diseases—and even less on incorporating empirical measures of water quality and sanitation into the types of mathematical models for which the theory has been developed. This is of direct practical importance. By identifying “hot spots” of cholera risk and understanding their role in disease transmission through mathematical modeling, public health officials can pinpoint priority areas for intervention. Insights from our modeling efforts in Haiti will also be relevant for future outbreaks in other regions of the world. One of the primary concerns in epidemic situations is forecasting where the outbreak is likely to spread. This involves calibration of mathematical models of the outbreak before the disease has reached a given area, which is a challenging task.

The requested funds in this proposal will be used to lay the groundwork for evaluating intervention efforts for cholera in Haiti. A key component of the short-term preventative interventions to date involves education campaigns through a variety of media, including posters, radio announcements and songs, and cell phone text messages. The effectiveness of these education campaigns has not been evaluated. By establishing contacts with the UN Health Cluster overseeing the coordination of these campaigns, we will be able to work towards obtaining data on both the coverage level of these education messages, as well as on cholera case data at a corresponding level of spatial resolution. Mathematical and statistical models will then be used to quantify the impact of these intervention efforts. Of particular interest are the use of text messages for cholera education. In our previous work on cholera in Haiti, we have established a collaboration with Digicel, the primary cell phone carrier in Haiti. Linus Bengtsson and colleagues at the Karolinska Institute (Sweden) initially forged an agreement with Digicel in order to...

Figure "One panel from an educational poster about cholera made by the Haitian MSPP."

The image shows various scenes related to the Haiti Cholera Outbreak, including water sources, sanitation facilities, and health workers. The poster in the image is written in Creole: "Jete poupou ak vomisman nan latrin."
Cholera & the environment

The recent natural disasters are an important facet of predicting epidemic dynamics in Haiti—much of the country's already poor water and sanitation infrastructure was destroyed in the recent January earthquake. The damage from the earthquake displaced approximately 950,000 people. Over one million of whom remain in tent camps, without electricity, running water, or sewage disposal. This disarray exacerbated the risk of infectious disease, particularly waterborne diseases such as cholera, contributing to the spread of the outbreak. Additionally, after the earthquake many residents of the major cities such as Port au Prince, the capital of Haiti, fled the cities to return to the outlying departments. The resulting higher population densities in rural areas is likely to have affected disease spread. To compound these issues, flooding due to the subsequent Hurricane Tomas is believed to have caused a resurgence in the epidemic, highlighting the direct link between the status of available water and the course of the epidemic.

Aspects of social and human behavior also affect cholera dynamics and case counts. The social stigma associated with cholera is severe, with at least 2,000 lynchings reported within the Grande Anse department. Cholera victims and their families may be reluctant to reveal that a sickness or death is due to cholera, in some cases hiding the body of the deceased. This makes it difficult to evaluate mortality and case counts within the community at large outside of hospitals, particularly in more remote villages.

The ongoing cholera outbreak in Haiti thus provides an example of the type of public health crisis where insights from cholera modeling are needed rapidly from incomplete data, and where forecasting the spatial dynamics of the cholera outbreak is an important but difficult task. Currently, daily cholera cases and hospitalizations by Department are available through the Haitian Ministere de la Sante Publique et de la Population (MSPP), with the outbreak beginning in the St. Marc region of the Artibonite Department in October 2017. Additionally, due to the ensuing relief efforts following the earthquake, a number of unusual data sets are available from Haiti. These include highly detailed information on the spatial location and population sizes.
Rainfall Data

- NASA TRMM Data - satellite precipitation data (resolution 0.25° \times 0.25°) averaged over each area

- USGS Rain Gauges in the Morne Gentilehomme and Foret de Pins regions
SIWR Model & Rainfall

\[ \frac{ds}{dt} = \mu - \beta_w f_{\text{rain}}(t) ws - \beta_1 si - \mu s \]
\[ \frac{di}{dt} = \beta_w f_{\text{rain}}(t) ws + \beta_1 si - \gamma i - \mu i \]
\[ \frac{dw}{dt} = \xi (i - w) \]
\[ \frac{dr}{dt} = \gamma i - \mu r \]
\[ y = ki \]

Eisenberg, Kujbida, Tuite, Fisman Tien, *Epidemics* 2013
Figure 8: Model fits (black line) to case data (grey circles) with the last two data points dropped. Subsequent model predictions compared to data points not used for fitting shown in red. Data sources are, clockwise: HAS data, IDP Camps data, MSPP Port au Prince (PaP) data, and MSPP national case counts.

11

Figure 8: Model fits (black line) to shortened case data (grey circles) with the last two (leftmost column), three (center column), or four (rightmost column) data points dropped. Subsequent model predictions compared to data points not used for fitting shown in red. Data sources are, from the top row to bottom row: HAS data, IDP Camps data, MSPP Port au Prince (PaP) data, and MSPP national case counts.

9
Cholera in Maela refugee camp
Cholera in Maela, Thailand

- Established in 1984
- ~50,000 refugees, mostly Karen

MAP: Mae La Temporary Shelter, the largest Burmese refugee camp in Thailand, established in 1984. Between 50,000 and 100,000 refugees live in the camp, with the population fluctuating based on various factors (e.g., season, political climate).
Cholera in Maela, Thailand

- Thai MoPH, US-CDC, & NGO’s provide medical care & water, sanitation, & hygiene (WaSH)

- Solidarités provides clean water & toilets

- River through Maela, springs, private wells & taps also serve as water sources

- Cholera outbreaks every 1-2 years
Modeling Cholera in Maela
Modeling Cholera in Maela

- Model includes demographic dynamics (based on camp data from 2010)
  - Population sizes, migration & dynamics by age group (<15 and adults)
  - Births & deaths
- Model also simulates observed cases vs. unreported (e.g. due to being asymptomatic)
Parameter Estimation

• However, similar issues with parameters—more transmission pathways $\rightarrow$ more unidentifiability issues

• As with basic SIWR, with perfect data the model is ID, however with noise/uncertainty it becomes unID

• Simplified model—assumed all human-human transmission parameters equal (i.e. adult-child = child-child = adult-adult, etc.)
Modeling Cholera in Maela

Adult Cholera Cases – Maela 2010

Child Cholera Cases – Maela 2010
Cholera Conclusions

• Identifiability is a key first step to modeling cholera dynamics

• Transmission pathways often unidentifiable, but water/rainfall data can improve estimates

• Useful for forecasting

• Potential for combining data sets—e.g. incidence + climate data
Questions?

comic by Olivia Walch (UM):
http://imogenquest.net