Using shooting approaches to generate initial guesses for ODE parameter estimation

Jonathan Calver (supervised by Wayne Enright)

Department of Computer Science
University of Toronto

AMMCS
Waterloo, Ontario, Canada
August 19, 2019

Outline

Parameter Estimation for ODEs

Shooting Approaches

Generating Initial Guesses

Proposed Methods

Numerical Example

ODEs

We consider the initial value problem (IVP),

$$\dot{\boldsymbol{y}}(t) = \boldsymbol{f}(t, \boldsymbol{y}(t), \boldsymbol{p})$$
$$\boldsymbol{y}(0) = \boldsymbol{y}_0$$
$$t \in [0, T]$$

- Numerous numerical methods exist for approximating $\boldsymbol{y}(t, p)$ that satisfy this IVP

ODEs

We consider the initial value problem (IVP),

$$\dot{\boldsymbol{y}}(t) = \boldsymbol{f}(t, \boldsymbol{y}(t), \boldsymbol{p})$$
$$\boldsymbol{y}(0) = \boldsymbol{y}_0$$
$$t \in [0, T]$$

- Numerous numerical methods exist for approximating $\boldsymbol{y}(t, p)$ that satisfy this IVP
- Most methods simulate the **trajectory**, $\boldsymbol{y}(t, p)$, starting from $\boldsymbol{y}(0)$. (e.g. continuous Runge-Kutta with defect control Enright and Yan (2010))

Parameter Estimation

- The goal of parameter estimation is to estimate the parameters, $\hat{\boldsymbol{p}}$, that make the model best fit the observed data $(\hat{\boldsymbol{y}}(t_i) = \boldsymbol{y}(t_i) + N(0, \sigma^2),\ i = 1, \ldots, n_o\ )$.

Parameter Estimation

- The goal of parameter estimation is to estimate the parameters, $\hat{\boldsymbol{p}}$, that make the model best fit the observed data ($\hat{\boldsymbol{y}}(t_i) = \boldsymbol{y}(t_i) + N(0, \sigma^2),\ i = 1, \ldots, n_o$).
- This is a non-linear least squares problem.

$$
\begin{aligned}
\hat{\boldsymbol{p}} &= \arg\min_{\boldsymbol{p}} \sum_{i=1}^{n_o} \frac{\|\hat{\boldsymbol{y}}(t_i) - \boldsymbol{y}(t_i, \boldsymbol{p})\|^2}{2\sigma^2} \\
&= \arg\min_{\boldsymbol{p}} L(\boldsymbol{p})
\end{aligned}
$$

Parameter Estimation

- The goal of parameter estimation is to estimate the parameters, $\hat{\boldsymbol{p}}$, that make the model best fit the observed data $(\hat{\boldsymbol{y}}(t_i) = \boldsymbol{y}(t_i) + N(0, \sigma^2),\ i = 1, \ldots, n_o\ )$.
- This is a non-linear least squares problem.

$$\begin{aligned} \hat{\boldsymbol{p}} &= \arg\min_{\boldsymbol{p}} \sum_{i=1}^{n_o} \frac{\|\hat{\boldsymbol{y}}(t_i) - \boldsymbol{y}(t_i, \boldsymbol{p})\|^2}{2\sigma^2} \\ &= \arg\min_{\boldsymbol{p}} L(\boldsymbol{p}) \end{aligned}$$

We will consider using **shooting approaches** - Given $\boldsymbol{p}$, simulate the ODE IVP to approximate $y(t, \boldsymbol{p})$ whenever evaluating $L(\boldsymbol{p})$.

Shooting Approaches

Simple (or single) shooting (e.g. Bard (1974))

- Choose an initial guess, $\boldsymbol{p}_o$

Common criticisms:

Shooting Approaches

Simple (or single) shooting (e.g. Bard (1974))

- Choose an initial guess, $\boldsymbol{p}_o$
- Apply a suitable local optimizer to find $\hat{\boldsymbol{p}}$.

Common criticisms:

Shooting Approaches

Simple (or single) shooting (e.g. Bard (1974))

- Choose an initial guess, $\boldsymbol{p}_o$
- Apply a suitable local optimizer to find $\hat{\boldsymbol{p}}$.

Common criticisms:

- might converge to a local minimum instead of $\hat{\boldsymbol{p}}$

Shooting Approaches

Simple (or single) shooting (e.g. Bard (1974))

- Choose an initial guess, $\boldsymbol{p}_o$
- Apply a suitable local optimizer to find $\hat{\boldsymbol{p}}$.

Common criticisms:

- might converge to a local minimum instead of $\hat{\boldsymbol{p}}$
- simulation may fail for some values of $\boldsymbol{p}$ before time T

Shooting Approaches

Simple (or single) shooting (e.g. Bard (1974))

- Choose an initial guess, $\boldsymbol{p}_o$
- Apply a suitable local optimizer to find $\hat{\boldsymbol{p}}$.

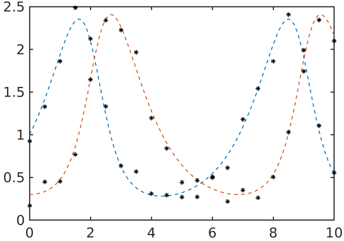
Common criticisms:

- might converge to a local minimum instead of $\hat{\boldsymbol{p}}$
- simulation may fail for some values of $\boldsymbol{p}$ before time T

If $\boldsymbol{p}_o$ is close enough to $\hat{\boldsymbol{p}}$, single shooting typically works quite well.

Example: Barnes Problem

Lotka-Volterra Predator-Prey Model

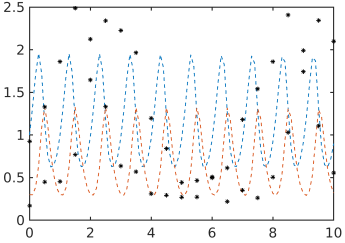


- oscillatory trajectory
- global minimum sometimes found with single shooting.

$$y_1'(t) = p_1 y_1(t) - p_2 y_1(t) y_t(2)$$
$$y_2'(t) = p_2 y_1(t) y_t(2) - p_3 y_2(t)$$
$$p = [1, 1, 1]$$
$$y(0) = [1, 0.3]$$

Example: Barnes Problem

Lotka-Volterra Predator-Prey Model

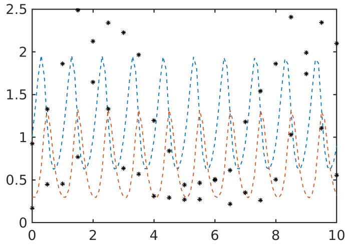


- oscillatory trajectory
- global minimum sometimes found with single shooting.
- local minimum sometimes found with single shooting.

$$
\begin{aligned}
y_1'(t) &= p_1 y_1(t) - p_2 y_1(t) y_t(2) \\
y_2'(t) &= p_2 y_1(t) y_t(2) - p_3 y_2(t) \\
p &= [1, 1, 1] \\
y(0) &= [1, 0.3]
\end{aligned}
$$

Example: Barnes Problem

Lotka-Volterra Predator-Prey Model



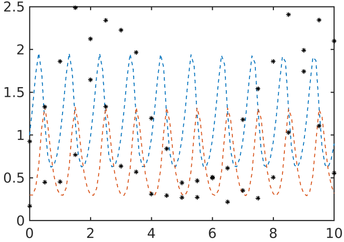
- oscillatory trajectory
- global minimum sometimes found with single shooting.
- local minimum sometimes found with single shooting.

$$
\begin{aligned}
y_1'(t) &= p_1 y_1(t) - p_2 y_1(t) y_t(2) \\
y_2'(t) &= p_2 y_1(t) y_t(2) - p_3 y_2(t) \\
p &= [1, 1, 1] \\
y(0) &= [1, 0.3]
\end{aligned}
$$

Insight: Once the trajectory begins to diverge, how well it fits the rest of the data isn't really useful.

Example: Barnes Problem

Lotka-Volterra Predator-Prey Model



- oscillatory trajectory
- global minimum sometimes found with single shooting.
- local minimum sometimes found with single shooting.

$$
\begin{aligned}
y_1'(t) &= p_1 y_1(t) - p_2 y_1(t) y_t(2) \\
y_2'(t) &= p_2 y_1(t) y_t(2) - p_3 y_2(t) \\
p &= [1, 1, 1] \\
y(0) &= [1, 0.3]
\end{aligned}
$$

Idea: Only fit to the first part of the data initially.

More Robust Shooting Approaches

Incremental (or progressive) shooting (Michalik et al., 2009; Krogh et al., 1985)

- simulate the ODE IVP to approximate $y(t, \boldsymbol{p})$ from $t = 0$ up to s ($s \leq T$).

More Robust Shooting Approaches

Incremental (or progressive) shooting (Michalik et al., 2009; Krogh et al., 1985)

- simulate the ODE IVP to approximate $y(t, \boldsymbol{p})$ from $t = 0$ up to s ($s \leq T$).
- Apply a suitable local optimizer to find $\hat{\boldsymbol{p}}(s)$.

More Robust Shooting Approaches

Incremental (or progressive) shooting (Michalik et al., 2009; Krogh et al., 1985)

- simulate the ODE IVP to approximate $y(t, \boldsymbol{p})$ from $t = 0$ up to s ($s \leq T$).
- Apply a suitable local optimizer to find $\hat{\boldsymbol{p}}(s)$.
- Set $\boldsymbol{p}_o = \hat{\boldsymbol{p}}(s)$.

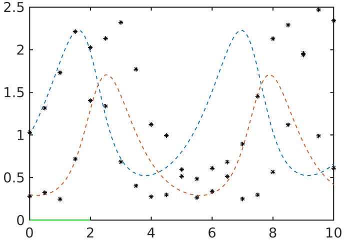
More Robust Shooting Approaches

Incremental (or progressive) shooting (Michalik et al., 2009; Krogh et al., 1985)

- simulate the ODE IVP to approximate $y(t, \boldsymbol{p})$ from $t = 0$ up to s ($s \leq T$).
- Apply a suitable local optimizer to find $\hat{\boldsymbol{p}}(s)$.
- Set $\boldsymbol{p}_o = \hat{\boldsymbol{p}}(s)$.
- Increment s and repeat until $s = T$. ($\hat{\boldsymbol{p}}(T) = \hat{\boldsymbol{p}}$)

Example: Barnes Problem

Lotka-Volterra Predator-Prey Model

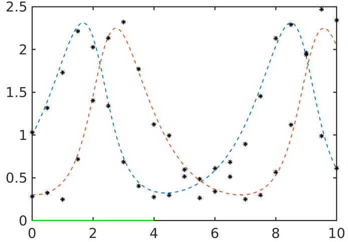


Initially $s = 2$, we increment by 2

- Incremental shooting can succeed here - less sensitive to $\boldsymbol{p}_o$.

Example: Barnes Problem

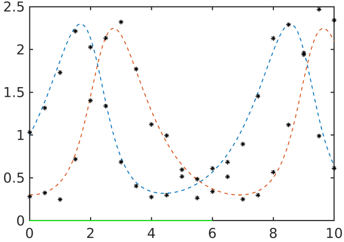
Lotka-Volterra Predator-Prey Model



- Incremental shooting can succeed here - less sensitive to $\boldsymbol{p}_o$.

Example: Barnes Problem

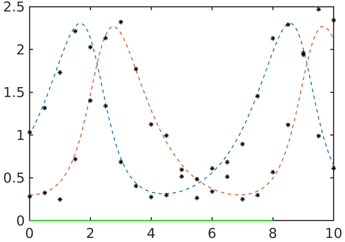
Lotka-Volterra Predator-Prey Model



- Incremental shooting can succeed here - less sensitive to $\boldsymbol{p}_o$.

Example: Barnes Problem

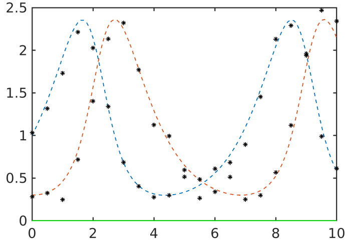
Lotka-Volterra Predator-Prey Model



- Incremental shooting can succeed here - less sensitive to $\boldsymbol{p}_o$.

Example: Barnes Problem

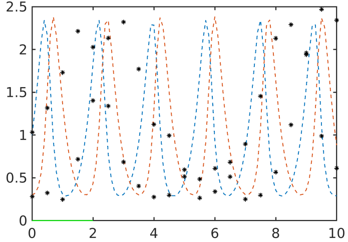
Lotka-Volterra Predator-Prey Model



- Incremental shooting can succeed here - less sensitive to $\boldsymbol{p}_o$.

Example: Barnes Problem

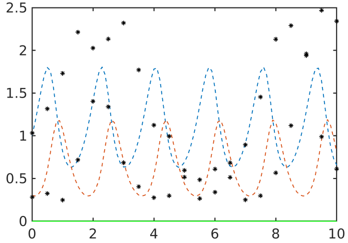
Lotka-Volterra Predator-Prey Model



- Incremental shooting can succeed here - less sensitive to $\boldsymbol{p}_o$.
- Poor local minimum still sometimes found with incremental shooting.

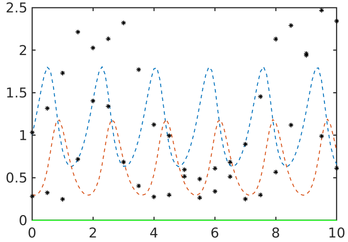
Example: Barnes Problem

Lotka-Volterra Predator-Prey Model



- Incremental shooting can succeed here - less sensitive to $\boldsymbol{p}_o$.
- Poor local minimum still sometimes found with incremental shooting.

Example: Barnes Problem
Lotka-Volterra Predator-Prey Model

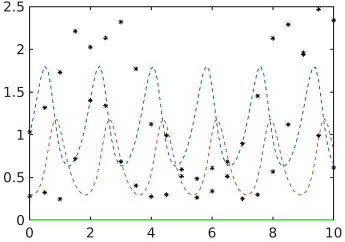


- Incremental shooting can succeed here - less sensitive to $\boldsymbol{p}_o$.
- Poor local minimum still sometimes found with incremental shooting.

Insight: It might help to use all of the observations

Example: Barnes Problem

Lotka-Volterra Predator-Prey Model



- Incremental shooting can succeed here - less sensitive to $\boldsymbol{p}_o$.
- Poor local minimum still sometimes found with incremental shooting.

Idea: reset the trajectory and keep simulating

More Robust Shooting Approaches

Multiple shooting (Bock and Plitt, 1984; Peifer and Timmer, 2007)

- Split the interval $[0, T]$ into n_{MS} sub-intervals (i.e. generate a mesh, $\{0 = \tau_0, \tau_1, \ldots, \tau_{n_{MS}} = T\}$)

More Robust Shooting Approaches

Multiple shooting (Bock and Plitt, 1984; Peifer and Timmer, 2007)

- Split the interval $[0, T]$ into n_{MS} sub-intervals (i.e. generate a mesh, $\{0 = \tau_0, \tau_1, \ldots, \tau_{n_{MS}} = T\}$)
- simulate the ODE IVP to approximate $y(t, \boldsymbol{p})$ on each of the n_{MS} sub-intervals.

More Robust Shooting Approaches

Multiple shooting (Bock and Plitt, 1984; Peifer and Timmer, 2007)

- Split the interval $[0, T]$ into n_{MS} sub-intervals (i.e. generate a mesh, $\{0 = \tau_0, \tau_1, \ldots, \tau_{n_{MS}} = T\}$)
- simulate the ODE IVP to approximate $y(\boldsymbol{t}, \boldsymbol{p})$ on each of the n_{MS} sub-intervals.
- Add an equality constraint at each sub-interval boundary $y(\tau_i)^- = y(\tau_i)^+$

More Robust Shooting Approaches

Multiple shooting (Bock and Plitt, 1984; Peifer and Timmer, 2007)

- Split the interval $[0, T]$ into n_{MS} sub-intervals (i.e. generate a mesh, $\{0 = \tau_0, \tau_1, \ldots, \tau_{n_{MS}} = T\}$)
- simulate the ODE IVP to approximate $y(t, \boldsymbol{p})$ on each of the n_{MS} sub-intervals.
- Add an equality constraint at each sub-interval boundary $y(\tau_i)^- = y(\tau_i)^+$
- Apply a suitable local optimizer to find $\hat{\boldsymbol{p}}$.

More Robust Shooting Approaches

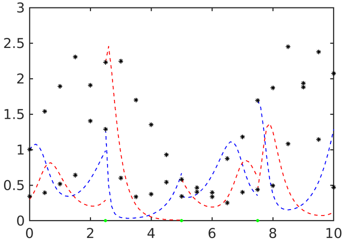
Multiple shooting (Bock and Plitt, 1984; Peifer and Timmer, 2007)

- Split the interval $[0, T]$ into n_{MS} sub-intervals (i.e. generate a mesh, $\{0 = \tau_0, \tau_1, \ldots, \tau_{n_{MS}} = T\}$)
- simulate the ODE IVP to approximate $y(t, \boldsymbol{p})$ on each of the n_{MS} sub-intervals.
- Add an equality constraint at each sub-interval boundary $y(\tau_i)^- = y(\tau_i)^+$
- Apply a suitable local optimizer to find $\hat{\boldsymbol{p}}$.

Can initialize each $y(\tau_i)^+$ using nearby observed data points

Example: Barnes Problem

Lotka-Volterra Predator-Prey Model

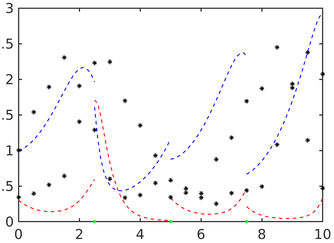


We used $n_{MS} = 4$, uniform spacing

- multiple shooting can succeed here - less sensitive to $\boldsymbol{p}_o$.

Example: Barnes Problem

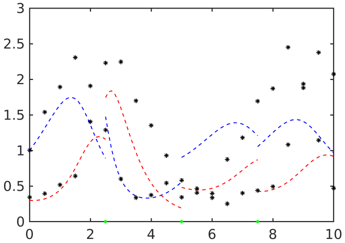
Lotka-Volterra Predator-Prey Model



- multiple shooting can succeed here - less sensitive to $\boldsymbol{p}_o$.

Example: Barnes Problem

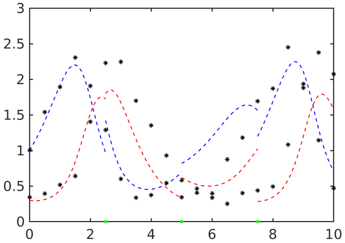
Lotka-Volterra Predator-Prey Model



- multiple shooting can succeed here - less sensitive to $\boldsymbol{p}_o$.

Example: Barnes Problem

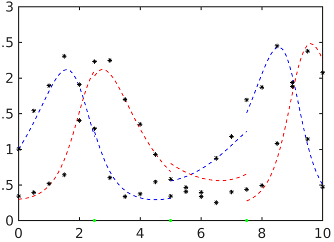
Lotka-Volterra Predator-Prey Model



- multiple shooting can succeed here - less sensitive to $\boldsymbol{p}_o$.

Example: Barnes Problem

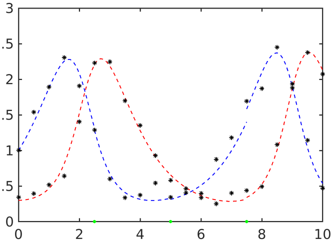
Lotka-Volterra Predator-Prey Model



- multiple shooting can succeed here - less sensitive to $\boldsymbol{p}_o$.

Example: Barnes Problem

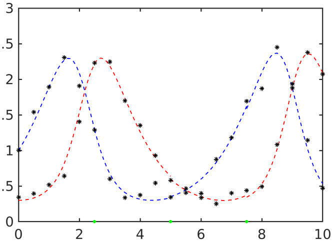
Lotka-Volterra Predator-Prey Model



- multiple shooting can succeed here - less sensitive to $\boldsymbol{p}_o$.

Example: Barnes Problem

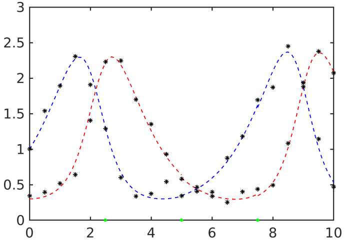
Lotka-Volterra Predator-Prey Model



- multiple shooting can succeed here - less sensitive to $\boldsymbol{p}_o$.

Example: Barnes Problem

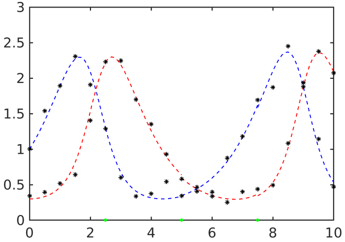
Lotka-Volterra Predator-Prey Model



- multiple shooting can succeed here - less sensitive to $\boldsymbol{p}_o$.
- local minimum still sometimes found with multiple shooting.

Example: Barnes Problem

Lotka-Volterra Predator-Prey Model

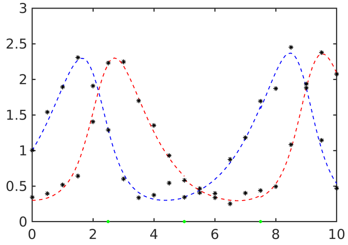


- multiple shooting can succeed here - less sensitive to $\boldsymbol{p}_o$.
- local minimum still sometimes found with multiple shooting.

Try 100 random samples for $\boldsymbol{p}_o$ from $[0, 10]^3$

Example: Barnes Problem

Lotka-Volterra Predator-Prey Model



- multiple shooting can succeed here - less sensitive to $\boldsymbol{p}_o$.
- local minimum still sometimes found with multiple shooting.

Try 100 random samples for $\boldsymbol{p}_o$ from $[0, 10]^3$

	simple	incremental	multiple
convergence	5%	39 %	48 %
median time	0.1s	0.36s	0.15s

Methods for obtaining a suitable $\boldsymbol{p}_o$

- Smooth and match estimator (SME) (Gugushvili et al., 2012)

Methods for obtaining a suitable $\boldsymbol{p}_o$

- Smooth and match estimator (SME) (Gugushvili et al., 2012)
  - Numerical Discretization based estimation (DBE) is similar (Wu et al., 2012)

Methods for obtaining a suitable $\boldsymbol{p}_o$

- Smooth and match estimator (SME) (Gugushvili et al., 2012)
  - Numerical Discretization based estimation (DBE) is similar (Wu et al., 2012)
  - Gradient Matching (Macdonald and Husmeier, 2015) in the ML literature

$$\boldsymbol{p}_{\text{sme}} = \arg\min_{\boldsymbol{p}} \int_0^T \left\|(\tilde{\boldsymbol{y}}'(t) - \boldsymbol{f}(t, \tilde{\boldsymbol{y}}(t), \boldsymbol{p}))\right\|^2 dt.$$

Methods for obtaining a suitable $\boldsymbol{p}_o$

- Smooth and match estimator (SME) (Gugushvili et al., 2012)
  - Numerical Discretization based estimation (DBE) is similar (Wu et al., 2012)
  - Gradient Matching (Macdonald and Husmeier, 2015) in the ML literature
  - Varah (1982); Bellman and Roth (1971) and others recognized that if one uses the observed values of $\boldsymbol{y}(t)$ to approximate $\boldsymbol{y}'(t)$, then one can formulate a related least squares problem,

$$\boldsymbol{p}_{\text{sme}} = \arg\min_{\boldsymbol{p}} \int_0^T \left\| (\tilde{\boldsymbol{y}}'(t) - \boldsymbol{f}(t, \tilde{\boldsymbol{y}}(t), \boldsymbol{p})) \right\|^2 dt.$$

Methods for obtaining a suitable $\boldsymbol{p}_o$

- Smooth and match estimator (SME) (Gugushvili et al., 2012)
  - Numerical Discretization based estimation (DBE) is similar (Wu et al., 2012)
  - Gradient Matching (Macdonald and Husmeier, 2015) in the ML literature
  - Varah (1982); Bellman and Roth (1971) and others recognized that if one uses the observed values of $\boldsymbol{y}(t)$ to approximate $\boldsymbol{y}'(t)$, then one can formulate a related least squares problem,

$$\boldsymbol{p}_{\text{sme}} = \arg\min_{\boldsymbol{p}} \int_0^T \left\| (\tilde{\boldsymbol{y}}'(t) - \boldsymbol{f}(t, \tilde{\boldsymbol{y}}(t), \boldsymbol{p})) \right\|^2 dt.$$

- $\tilde{\boldsymbol{y}}$ is obtained by smoothing the observations

Integral smooth and match (INT-SME)

- Gugushvili et al. (2012) proposed using the integral form of the ODE IVP instead.

$$\min_{\boldsymbol{y}_0,\boldsymbol{p}} \int_0^T \left\| \tilde{y}(t) - \left(\boldsymbol{y}_0 + \int_0^t f(s, \tilde{y}(s), \boldsymbol{p})\, ds\right) \right\|^2 dt,$$

where $\boldsymbol{y}_0$ are the initial conditions to be estimated.

Integral smooth and match (INT-SME)

- Gugushvili et al. (2012) proposed using the integral form of the ODE IVP instead.

$$\min_{\boldsymbol{y}_0, \boldsymbol{p}} \int_0^T \left\| \tilde{y}(t) - \left(\boldsymbol{y}_0 + \int_0^t f(s, \tilde{y}(s), \boldsymbol{p})\, ds\right) \right\|^2 dt,$$

where $\boldsymbol{y}_0$ are the initial conditions to be estimated.

This is the method we will consider for the rest of this talk

Integral smooth and match (INT-SME)

- Gugushvili et al. (2012) proposed using the integral form of the ODE IVP instead.
- Similar approach to DBE that uses the integral form also recently appeared in the UQ literature (Green and Rindler, 2019)

$$\min_{\boldsymbol{y}_0,\boldsymbol{p}} \int_0^T \left\| \tilde{y}(t) - \left(\boldsymbol{y}_0 + \int_0^t f(s, \tilde{y}(s), \boldsymbol{p})\, ds\right) \right\|^2 dt,$$

where $\boldsymbol{y}_0$ are the initial conditions to be estimated.

This is the method we will consider for the rest of this talk

Case of Linear Parameters

- Sometimes the parameters appear linearly in $f(t, \boldsymbol{y}(t), \boldsymbol{p})$.
- $\boldsymbol{p} = \boldsymbol{r}$

$$f(t, \boldsymbol{y}(t), \boldsymbol{p}) = G(t, \boldsymbol{y}(t))\boldsymbol{r}$$

- INT-SME becomes:

$$\arg\min_{\boldsymbol{r}, \boldsymbol{y}_0} \int_0^T \left\| \tilde{\boldsymbol{y}}(t) - (\boldsymbol{y}_0 + [\int_0^t G(s, \tilde{\boldsymbol{y}}(s))\, ds]\, \boldsymbol{r}) \right\|^2 dt.$$

Case of Linear Parameters

- Sometimes the parameters appear linearly in $f(t, \boldsymbol{y}(t), \boldsymbol{p})$.
- $\boldsymbol{p} = \boldsymbol{r}$

$$f(t, \boldsymbol{y}(t), \boldsymbol{p}) = G(t, \boldsymbol{y}(t))\boldsymbol{r}$$

- INT-SME becomes:

$$\arg\min_{\boldsymbol{r}, \boldsymbol{y}_0} \int_0^T \left\| \tilde{\boldsymbol{y}}(t) - (\boldsymbol{y}_0 + [\int_0^t G(s, \tilde{\boldsymbol{y}}(s))\, ds]\, \boldsymbol{r}) \right\|^2 dt.$$

INT-SME is a linear least squares problem

Case of Linear Parameters

- Sometimes the parameters appear linearly in $f(t, \boldsymbol{y}(t), \boldsymbol{p})$.
- $\boldsymbol{p} = \boldsymbol{r}$

$$f(t, \boldsymbol{y}(t), \boldsymbol{p}) = G(t, \boldsymbol{y}(t))\boldsymbol{r}$$

- INT-SME becomes:

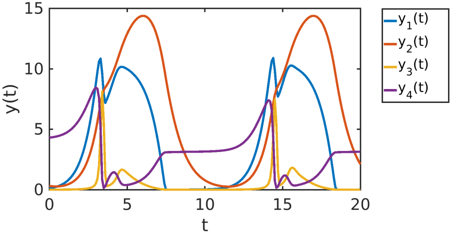
$$\arg\min_{\boldsymbol{r}, \boldsymbol{y}_0} \int_0^T \left\| \tilde{\boldsymbol{y}}(t) - \left(\boldsymbol{y}_0 + \left[\, \int_0^t G(s, \tilde{\boldsymbol{y}}(s))\, ds \right] \boldsymbol{r}\right) \right\|^2 dt.$$

For a given $\tilde{\boldsymbol{y}}(t)$, this gives only one candidate $\boldsymbol{p}_o$

Calcium Ion Test Problem [Kummer et al. (2000) Peifer and Timmer (2007)]

$$G^*{}_\alpha{}' = k_1 + k_2 G^*{}_\alpha - k_3 PLC^* \frac{G^*{}_\alpha}{G^*{}_\alpha + Km_1} - k_4 Ca_{cyt} \frac{G^*{}_\alpha}{G^*{}_\alpha + Km_2},$$
$$PLC^{*\prime} = k_5 G^*{}_\alpha - k_6 \frac{PLC^*}{PLC^* + Km_3},$$
$$Ca_{cyt}{}' = k_7 PLC^* Ca_{cyt} \frac{Ca_{er}}{Ca_{er} + Km_4} + k_8 PLC^* + k_9 G^*{}_\alpha - k_{10} \frac{Ca_{cyt}}{Ca_{cyt} + Km_5} - k_{11} \frac{Ca_{cyt}}{Ca_{cyt} + Km_6},$$
$$Ca_{er}{}' = -k_7 PLC^* Ca_{cyt} \frac{Ca_{er}}{Ca_{er} + Km_4} + k_{11} \frac{Ca_{cyt}}{Ca_{cyt} + Km_6},$$

- 11 linear parameters (k) to be estimated
- 6 non-linear parameters (Km) [held fixed]
- 6.5% relative noise added to true trajectory sampled at 200 uniformly spaced times
- initial conditions assumed known



Calcium Ion Test Problem Results

method	converged	avg time of converged (s)
SS	4 %	44 ± 16
MS	49 %	48 ± 58

Table: Results reported in Peifer and Timmer (2007). The trajectory simulations were performed using ODESSA and a Gauss-Newton optimizer. They generated one set of noisy data and ran each of simple shooting (SS) and multiple shooting (MS) with $n_{MS} = 17$ from 250 random initial guesses for $\boldsymbol{p}_o$, which were drawn uniformly from $[0, 1]^{11}$.

- We generated 100 sets of noisy data

Calcium Ion Test Problem Results

method	converged	avg time of converged (s)
SS	4 %	44 ± 16
MS	49 %	48 ± 58

Table: Results reported in Peifer and Timmer (2007). The trajectory simulations were performed using ODESSA and a Gauss-Newton optimizer. They generated one set of noisy data and ran each of simple shooting (SS) and multiple shooting (MS) with $n_{MS} = 17$ from 250 random initial guesses for $\boldsymbol{p}_o$, which were drawn uniformly from $[0, 1]^{11}$.

- We generated 100 sets of noisy data
- 63% success rate if we use INT-SME to generate $\boldsymbol{p}_o$ and then use simple shooting.

Calcium Ion Test Problem Results

method	converged	avg time of converged (s)
SS	4 %	44 ± 16
MS	49 %	48 ± 58

Table: Results reported in Peifer and Timmer (2007). The trajectory simulations were performed using ODESSA and a Gauss-Newton optimizer. They generated one set of noisy data and ran each of simple shooting (SS) and multiple shooting (MS) with $n_{MS} = 17$ from 250 random initial guesses for $\boldsymbol{p}_o$, which were drawn uniformly from $[0, 1]^{11}$.

- We generated 100 sets of noisy data
- 63% success rate if we use INT-SME to generate $\boldsymbol{p}_o$ and then use simple shooting.
- time: $1.1s \pm 2.2s$

Proposed Methods
Motivation

Why does INT-SME sometimes fail?
The integrals,

$$\int_0^t G(s, \tilde{\boldsymbol{y}}(s))\, ds,$$

may accumulate errors.

- only use a subset of the data? (like progressive shooting)

Proposed Methods
Motivation

Why does INT-SME sometimes fail?
The integrals,

$$\int_0^t G(s, \tilde{\boldsymbol{y}}(s))\, ds,$$

may accumulate errors.

- only use a subset of the data? (like progressive shooting)
- restart the integrals periodically? (like multiple shooting)

Why does INT-SME sometimes fail?
The integrals,

$$\int_0^t G(s, \tilde{\boldsymbol{y}}(s))\, ds,$$

may accumulate errors.

- only use a subset of the data? (like progressive shooting)
- restart the integrals periodically? (like multiple shooting)
- try a different smoother? (i.e. change $\tilde{\boldsymbol{y}}(t)$)

Proposed Methods

Motivation

Why does INT-SME sometimes fail?
The integrals,

$$\int_0^t G(s, \tilde{\boldsymbol{y}}(s))\, ds,$$

may accumulate errors.

- only use a subset of the data? (like progressive shooting)
- restart the integrals periodically? (like multiple shooting)
- try a different smoother? (i.e. change $\tilde{\boldsymbol{y}}(t)$)
- try a different quadrature rule?

Proposed Methods

Motivation

Why does INT-SME sometimes fail?
The integrals,

$$\int_0^t G(s, \tilde{\boldsymbol{y}}(s))\, ds,$$

may accumulate errors.

- only use a subset of the data? (like progressive shooting)
- restart the integrals periodically? (like multiple shooting)
- try a different smoother? (i.e. change $\tilde{\boldsymbol{y}}(t)$)
- try a different quadrature rule?
- we used simple shooting for the final optimization - could try incremental or multiple shooting instead

Proposed Methods
Motivation

Why does INT-SME sometimes fail?
The integrals,

$$\int_0^t G(s, \tilde{\boldsymbol{y}}(s))\, ds,$$

may accumulate errors.

- only use a subset of the data? (like progressive shooting)
- restart the integrals periodically? (like multiple shooting)
- try a different smoother? (i.e. change $\tilde{\boldsymbol{y}}(t)$)
- try a different quadrature rule?
- we used simple shooting for the final optimization - could try incremental or multiple shooting instead

We consider the first two ideas here.

Proposed Methods

First, we define INT-SME(s):

$$\min_{\boldsymbol{r}} \int_0^{s} \left\| \tilde{\boldsymbol{y}}(t) - (\boldsymbol{y}_0 + [\, \int_0^t G(\tau, \tilde{\boldsymbol{y}}(\tau))\, d\tau]\, \boldsymbol{r}) \right\|^2 dt,$$

where $0 < s \leq T$.

Incremental Shooting - only consider data up to time s.

Proposed Methods

INT-SME(m,s)

$$\sum_{i=1}^{m} \left[\int_{t_i}^{t_{i+1}} \left\| \tilde{y}(t) - \left(y(t_i)^{+} + \left[\int_{t_i}^{t} G(\tau, \tilde{\boldsymbol{y}}(\tau))\, d\tau \right] r \right) \right\|^2 dt \right],$$

- m is the number of intervals used

Proposed Methods

INT-SME(m,s)

$$\sum_{i=1}^{m}\left[\int_{t_i}^{t_{i+1}}\left\|\tilde{y}(t)-\Big(y(t_i)^{+}+\big[\int_{t_i}^{t}G(\tau,\tilde{\boldsymbol{y}}(\tau))\,d\tau\big]r\Big)\right\|^2 dt\right],$$

- m is the number of intervals used
- the set of t_i's partition the interval from 0 to s.

Proposed Methods

INT-SME(m,s)

$$\sum_{i=1}^{m} \left[\int_{t_i}^{t_{i+1}} \left\| \tilde{y}(t) - \left(y(t_i)^+ + [\, \int_{t_i}^{t} G(\tau, \tilde{\boldsymbol{y}}(\tau))\, d\tau] \, r \right) \right\|^2 dt \right],$$

- m is the number of intervals used
- the set of t_i's partition the interval from 0 to s.
- In our numerical experiments, we use uniform partitions.

Proposed Methods

INT-SME(m,s)

$$\sum_{i=1}^{m}\left[\int_{t_i}^{t_{i+1}}\left\|\tilde{y}(t)-\left(y(t_i)^{+}+\left[\int_{t_i}^{t}G(\tau,\tilde{\boldsymbol{y}}(\tau))\,d\tau\right]r\right)\right\|^2 dt\right],$$

- m is the number of intervals used
- the set of t_i's partition the interval from 0 to s.
- In our numerical experiments, we use uniform partitions.
- Unlike true multiple shooting, we do not enforce equality constraints at the end of each shooting interval.

Proposed Methods

Let $S = \{s_1, s_2, \ldots s_{N_{PS}}\}$.

INT-SME(S):

$$\min_{\boldsymbol{p} \in P_o} L(\boldsymbol{p}),$$

where the i'th element in P_o, is $\boldsymbol{p}_i = \text{INT-SME}(s_i)$

Proposed Methods

Let $S = \{s_1, s_2, \ldots s_{N_{PS}}\}$.

INT-SME(S):

$$\min_{\boldsymbol{p} \in P_o} L(\boldsymbol{p}),$$

where the i'th element in P_o, is $\boldsymbol{p}_i = \text{INT-SME}(s_i)$

Note that unlike INT-SME(m,s), this procedure *does* evaluate $L(p)$, so it may be more expensive.

Proposed Methods

Let $S = \{s_1, s_2, \ldots s_{N_{PS}}\}$.

INT-SME(S):

$$\min_{\boldsymbol{p} \in P_o} L(\boldsymbol{p}),$$

where the i'th element in P_o, is $\boldsymbol{p}_i = \text{INT-SME}(s_i)$

Note that unlike INT-SME(m,s), this procedure *does* evaluate $L(p)$, so it may be more expensive.

Calcium Ion Results

m \ s	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
1	96	98	98	92	93	85	95	92	94	95	92	84	72	66	65	63
2	93	100	92	90	98	98	98	100	95	95	98	91	80	66	55	64
4	98	100	98	92	97	100	100	99	98	97	98	96	94	93	86	91
8	98	100	99	100	100	100	99	99	100	100	100	99	98	99	98	100
16	99	100	100	98	100	100	100	100	100	100	100	100	100	100	99	100
PS	96	97	100	98	100	100	100	99	100	100	100	100	100	100	100	100

Table: What percentage of times the final optimization succeeded when INT-SME(m,s) was used to generate $\boldsymbol{p}_o$ using different values of m and the data up to time s.

- top row: simple shooting is poor if we use all the data (63%)

Calcium Ion Results

m \ s	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
1	96	98	98	92	93	85	95	92	94	95	92	84	72	66	65	63
2	93	100	92	90	98	98	98	100	95	95	98	91	80	66	55	64
4	98	100	98	92	97	100	100	99	98	97	98	96	94	93	86	91
8	98	100	99	100	100	100	99	99	100	100	100	99	98	99	98	100
16	99	100	100	98	100	100	100	100	100	100	100	100	100	100	99	100
PS	96	97	100	98	100	100	100	99	100	100	100	100	100	100	100	100

Table: What percentage of times the final optimization succeeded when INT-SME(m,s) was used to generate $\boldsymbol{p}_o$ using different values of m and the data up to time s.

- top row: simple shooting is poor if we use all the data (63%)
- multiple shooting improves as we add more intervals

Calcium Ion Results

m \ s	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
1	96	98	98	92	93	85	95	92	94	95	92	84	72	66	65	63
2	93	100	92	90	98	98	98	100	95	95	98	91	80	66	55	64
4	98	100	98	92	97	100	100	99	98	97	98	96	94	93	86	91
8	98	100	99	100	100	100	99	99	100	100	100	99	98	99	98	100
16	99	100	100	98	100	100	100	100	100	100	100	100	100	100	99	100
PS	96	97	100	98	100	100	100	99	100	100	100	100	100	100	100	100

Table: What percentage of times the final optimization succeeded when INT-SME(m,s) was used to generate $\boldsymbol{p}_o$ using different values of m and the data up to time s. The last line, PS, corresponds to using INT-SME(S), with $S = \{5, 6, \ldots, s\}$.

- top row: simple shooting is poor if we use all the data (63%)
- multiple shooting improves as we add more intervals
- bottom row: INT-SME(S) generates more candidates for $\boldsymbol{p}_o$ as we move right in the table

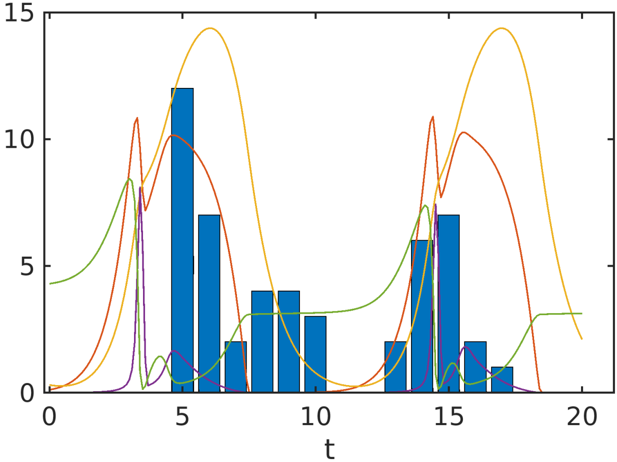
Additional Remarks

m \ s	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
1	0.87	0.84	0.92	0.89	1.00	0.85	0.83	0.98	0.91	0.85	0.86	1.01	1.45	1.48	1.31	1.65
2	0.93	0.84	0.85	0.93	0.91	0.86	0.89	0.93	1.05	0.80	0.93	1.18	1.17	1.55	1.32	1.57
4	0.91	0.77	0.84	0.97	0.80	0.83	0.83	0.89	0.86	0.85	0.93	0.92	1.04	1.08	1.31	1.19
8	0.87	0.81	0.86	0.86	0.81	0.81	0.83	0.81	0.82	0.85	0.85	0.86	0.94	0.98	0.90	0.95
16	0.98	0.92	0.93	0.89	0.90	0.92	0.92	0.95	0.92	0.95	0.91	0.96	0.98	1.04	1.01	1.03
PS	0.94	0.82	0.82	0.79	0.80	0.79	0.83	0.79	0.82	0.76	0.81	0.79	0.79	0.78	0.77	0.78

Table: Time taken for the full procedures.

- in all cases, run time is 0.8s - 1.6s.
- most of this time is for the final optimization.
- PS-INT-SME generates a better $\boldsymbol{p}_o$ as more candidates are considered.

Which INT-SME(s) is PS-INT-SME([5:20]) using for $\boldsymbol{p}_o$?



Histogram of which INT-SME(s) is being used as $\boldsymbol{p}_o$, overlaid with the true trajectory

- The best $\boldsymbol{p}_o$'s don't use all of the data.

Summary

- Discussed shooting approaches for ODE IVP parameter estimation

Summary

- Discussed shooting approaches for ODE IVP parameter estimation
- Demonstrated how the same ideas can be applied when determining $\boldsymbol{p}_o$.

Summary

- Discussed shooting approaches for ODE IVP parameter estimation
- Demonstrated how the same ideas can be applied when determining $\boldsymbol{p}_o$.
  - Proposed progressive and multiple shooting versions of INT-SME.

Summary

- Discussed shooting approaches for ODE IVP parameter estimation
- Demonstrated how the same ideas can be applied when determining $\boldsymbol{p}_o$.
  - Proposed progressive and multiple shooting versions of INT-SME.
  - Demonstrated the performance of the procedures on the Calcium Ion problem from the literature.

Summary

- Discussed shooting approaches for ODE IVP parameter estimation
- Demonstrated how the same ideas can be applied when determining $\boldsymbol{p}_o$.
  - Proposed progressive and multiple shooting versions of INT-SME.
  - Demonstrated the performance of the procedures on the Calcium Ion problem from the literature.

Joint work with Wayne Enright and Jienan Yao

Thanks for Listening

Questions?

Y. Bard. *Nonlinear parameter estimation*. Academic press, 1974.

R. Bellman and R.S. Roth. The use of splines with unknown end points in the identification of systems. *Journal of Mathematical Analysis and Applications*, 34(1):26–33, 1971.

H.G. Bock and K.J Plitt. A multiple shooting algorithm for direct solution of optimal control problems. *Proceedings 9th IFAC World Congress Budapest*, pages 243–247, 1984.

W.H. Enright and L Yan. The reliability/cost trade-off for a class of ode solvers. *Numerical Algorithms*, 53(2):239–260, 2010.

D. K. E. Green and F. Rindler. Model inference for ordinary differential equations by parametric polynomial kernel regression. 2019. URL `https://arxiv.org/abs/1908.02105`.

S. Gugushvili, C.A.J. Klaassen, et al. $\sqrt{n}$-consistent parameter estimation for systems of ordinary differential equations: bypassing numerical integration via smoothing. *Bernoulli*, 18(3): 1061–1098, 2012.

F.T. Krogh, J.P. Keener, and W.H. Enright. Reducing the number of variational equations in the implementation of multiple shooting. *Numerical Boundary Value ODEs*, pages 121–135, 1985.

U. Kummer, L. F. Olsen, C. J. Dixon, A. K. Green, E. Bornber-Bauer, and G. Baier. Switching from simple to complex oscillations in calcium signaling. *Biophys. J.*, 79(3): 1188–1195, 2000.

B. Macdonald and D. Husmeier. Gradient matching methods for computational inference in mechanistic models for systems biology: a review and comparative analysis. *Frontiers in bioengineering and biotechnology*, 3:180, 2015.

C. Michalik, R. Hannemann, and W. Marquardt. Incremental single shooting - a robust method for the estimation of parameters in dynamical systems. *Computers & Chemical Engineering*, 33(7):1298 – 1305, 2009. ISSN 0098-1354.

M. Peifer and J. Timmer. Parameter estimation in ordinary differential equations for biochemical processes using the method of multiple shooting. *IET System Biology*, 1(2):78–88, 2007.

J.M. Varah. A spline least squares method for numerical parameter estimation in differential equations. *SIAM J. of Sci. and Stat. Comput.*, 3(1):28–46, 1982.

H. Wu, H. Xue, and A. Kumar. Numerical discretization-based estimation methods for ordinary differential equation models via penalized spline smoothing with applications in biomedical research. *Biometrics*, 68(2):344–352, 2012.