Efficient Numerical Methods for Nonlinear Parameter Estimation

Johannes P. Schlöder

Interdisciplinary Center for Scientific Computing (IWR)
Heidelberg University

based on joint work with
A. Altmann-Dieses, H. G. Bock, K. Hatz, S. Lenz and E. Kostina

BISON Seminar
Automatic Control Lab, ETH Zurich, March 15, 2013
Overview

• Parameter Estimation Problems
  Example: Photosynthesis
  Differential Equation and Optimal Control Models & Data
  Optimization Boundary Value Problems

• Structure Exploiting Numerical Methods for
  Optimization Boundary Value Problems

  • The Direct Multiple Shooting Method for Parameter Estimation
  • The Generalized Gauss Newton Method
  • Assessment of the Statistical Error of the Parameter Estimates

• "Proof of Concept" and "Real World" Applications
Parameter Estimation: Match Model to Data

$M(p)$ model response
- data
Example:
The Light Reaction in Photosynthesis
The Light Reaction in Photosynthesis

3 experiments:
- fluorescence measured at living tobacco leave
- with 3 different laser intensities
The Light Reaction in Photosynthesis

electron transport chain in photosynthesis

mathematical model:
nonlinear ODE with 6 states, and 4+2 system parameters

\[
\begin{align*}
\dot{y}_1 &= (k_a + k_3 (p_{tot} - y_6)) y_1 + k_3 y_3 y_6 \\
\dot{y}_2 &= k_a y_1 - (k_1 + k_3 (p_{tot} - y_6)) y_2 + k_{-1} y_3 + k_3 y_6 (1 - \sum_{i=1}^{5} y_i) \\
\dot{y}_3 &= k_1 y_2 - (k_a + k_{-1}) y_3 \\
\dot{y}_4 &= k_u y_3 - k_2 y_4 + k_{-2} y_5 \\
\dot{y}_5 &= k_3 y_1 (p_{tot} - y_6) + k_2 y_4 - (k_a + k_{-2} + k_3 y_6) y_5 \\
\dot{y}_6 &= -k_3 (1 - \sum_{i=1}^{5} y_i) y_6 + k_3 (y_1 + y_2) (p_{tot} - y_6) + (p_{tot} - y_6) k_{\text{lim}} \\
\end{align*}
\]

with

\[
k_a = \frac{I_2 (1 - p_{2T})}{1 - p_{22} - p_{2T} + p_{22} p_{2T} (y_1 + y_3 + y_5)}
\]
The Light Reaction in Photosynthesis

fluorescence: nonlinear function of the states

\[ F(t) = \left\{ \frac{1 - p_{2T} - p_{22}}{p_{2T}} + \frac{1 - (y_1 + y_3 + y_5)}{1 + D} \right\} \cdot S \cdot I_2 \]

\[ D = p_{22} p_{2T} (y_1 + y_3 + y_5) / (1 - p_{2T} - p_{22}) \]

must estimate

- 4 system par’s & 1 observation par. jointly for all experiments
- 3x2 experiment dependent parameters

data: from 3 experiments with 96 measurements each
The Problem Formulation
Model: Differential Algebraic Equations (DAE) 
Optimal Control Problems (OCP)

\[ \dot{y} = f(t, y, z, p, q, u) \]
\[ 0 = g(t, y, z, p, q, u) \]

\( y \) „differential“ states 
\( z \) „algebraic“ states

\( p \): (unknown) system parameters (PE)  
\( q \): control parameters, \( u \): control functions (given)

- ordinary differential equations .... → PDE reactive flow 
  + further constraints possible

- or: optimal control problems

\[
\min_{y,u} \sum_j \gamma_j \Phi_j^M (y(T), p) \\
\text{s.t. } \dot{y}(t) = f(y(t), u(t), p) \\
0 \leq c(y(t), u(t), p) \\
0 = r(y(t_0), y(T))
\]

\( y \) „differential“ states  
\( u \) control functions  
\( \gamma \) weighting parameters  
\( p \) system parameters
The Experimental Data

- measurements

\[ \eta_{ij} = b_{ij}(t_i, y(t_i), z(t_i), p) + \varepsilon_{ij} \quad j \in \text{Ind}(i) \]

- measurement functions \( b_{ij} \), with add'l calibration parameters
- measurement errors \( \varepsilon_{ij} \)
- from multiple experiments, under varying conditions
  - instationary states
  - stationary
  - bifurcations
  - oscillations

Each has specific model!
The Parameter Estimation Problem (DAE)

DAE process model

\[
\begin{align*}
\dot{y} &= f(t, y, z, p, q, u) \\
0 &= g(t, y, z, p, q, u) \\
x &:= (y, z)
\end{align*}
\]

Data

\[
\begin{align*}
\eta_{ij} &= b_{ij}(t_i, x(t_i), p, q) + \varepsilon_{ij} \\
\varepsilon_{ij} &\in N(0, \sigma_{ij}^2)
\end{align*}
\]

determine \(p\) and \(x\)

\[
\min_{x, p} \sum_{i, j} \frac{(\eta_{ij} - b_{ij}(t_i, x(t_i), p, q))^2}{\sigma_{ij}^2}
\]

\[
\begin{align*}
\dot{y} &= f(t, y, z, p, q, u) \\
0 &= g(t, y, z, p, q, u) \\
d(x(t_0), \ldots, x(t_f), p, q) &= 0, \quad \text{or} \geq 0
\end{align*}
\]

Boundary value problem
The Multiple Experiment Case (DAE)

DAE process model

\[
\dot{y}_k = f_k(t, y_k, z_k, p, q_k, u_k) \\
0 = g(t, y_k, z_k, p, q_k, u_k)
\]

data

\[
\eta_{ijk} = b_{ijk}(t_{ik}, x(t_{ik}), p, q) + \varepsilon_{ijk} \\
\varepsilon_{ijk} \in N(0, \sigma_{ijk}^2)
\]

determine \(p\) and \(x_k\) \((k=1,\ldots,\#\ exp)\)

\[
\min_{x_k,p} \sum_{k}^{\#\ exp} \sum_{i,j} (\eta_{ijk} - b_{ijk}(t_{ik}, x_k(t_{ik}), p, q_k))^2
\]

\[
\dot{y}_k = f_k(t, y_k, z_k, p, q_k, u_k) \\
0 = g_k(t, y_k, z_k, p, q_k, u_k) \\
d_k(x_k(t_{0k}),\ldots, x_k(t_{fk}), p, q_k) = 0, \quad \text{or} \geq 0
\]

family of boundary value problems
The Choice of Norms

- least squares norm: Legendre 1805, Gauss 1809
  (normally distributed measurement error)

but much can be said in favour of
- $\ell_1$-norm: Boscovic 1758, Laplace 1812
  robust against outliers
  (Laplace-distributed measurement error)

Kostina '02, '04
Direct Methods for Constrained Parameter Estimation
Direct "All-at-Once" Boundary Value Problem (BVP) Methods

- the IVP approach: "single shooting"
  - integrate DAE over whole interval to yield $x(t;x_0,p)$ resp., solve OCP for given $\gamma$, $p$
  - eliminate all - infinite - variables in favour of unknown parameters $\gamma$, $p$,
  - plug into suitable optimizer
Direct "All-at-Once" Boundary Value Problem (BVP) Methods

- the IVP approach: "single shooting"
  - integrate DAE over whole interval to yield $x(t;x_0,p)$ resp., solve OCP for given $\gamma$, $p$
  - eliminate all - infinite - variables in favour of unknown parameters $\gamma$, $p$
  - plug into suitable optimizer

- the BVP approach: discretize the DAE/OCP, and solve simultaneously
  - optimization problem,
  - *discretized BVP as equality constraint or necessary conditions for discretized OCP* plus further constraints
    in one loop
The Direct Multiple Shooting Method for Parameter Estimation in DAE

parameterize/discretize DAE by the *multiple shooting* method, i. e.,

- choose suitable mesh
  
  \[ t_0 < t_1 < ... < t_m = t_f \]

- introduce state variables at nodes \( t_i \)
  
  \[ s^x_i \triangleq x(t_i), \quad s^z_i \triangleq z(t_i) \]

as additional optimization variables

alternatives: collocation on finite elements, finite differences, ...

Biegler

PARFIT

Bock, Bär, Schl. '78, '81, '83, '87 ff
Bock, Eich, Schl. '88, Kostina '01, '04
The Direct Multiple Shooting Method for Parameter Estimation in DAE

- Integrate relaxed DAE on multiple shooting subintervals \([t_i, t_{i+1}]\)

\[
\begin{align*}
\dot{y} &= f(t, y, z, p) \\
0 &= g(t, y, z, p) - \alpha_i(t) g(t_i, s_i^y, s_i^z, p)
\end{align*}
\]

\(\begin{cases}
\alpha_i(t_i) = 1, \\
\alpha_i(t) \to 0 \quad (t \to \infty)
\end{cases}\)

- Jumps and relaxation terms must vanish at the solution – additional continuity and consistency conditions replace DAE

\[
\begin{align*}
y(t; s_i^y, s_i^z, p) &= 0 \\
0 &= y(t_{i+1}; s_i^y, s_i^z, p) - s_{i+1}^y \quad (i=0, \ldots, m-1)
\end{align*}
\]
Result:
Constrained Nonlinear Least Squares Problem

[X] \begin{align*}
\min_X & \left\| F_1(X) \right\|_2^2 \\
\text{s.t.} & \quad F_2(X) = 0, \text{ or } \geq 0
\end{align*}

solution by Newton-type methods

\[ X^{k+1} = X^k + t^k \Delta X^k \]

where \( \Delta X^k \) solves a constrained linear least squares problem (CLLS)

[CLLS] \begin{align*}
\min_{\Delta X} & \left\| F_1(X^k) + J_1(X^k)\Delta X \right\|_2^2 \\
\text{s.t.} & \quad F_2(X^k) + J_2(X^k)\Delta X = 0, \text{ or } \geq 0
\end{align*}

solution by generalized inverse

\[ \Delta X^k := -J^+(X^k)F(X^k) \]

with \[ J^+ = J^+JJ^+ \]
Block-Sparse Structures of Jacobian

- super-structures from multiple experiments
- structure from multiple shooting
- sub-structures from spatial discretization of PDE
- sub-sub-structures from sparse state equations

# experiments 1 - 100

# multiple shooting points typically 10 - 40

must exploit special block structures → condensing
FAQ: Why Multiple Shooting?

• key property: discretized states as add'l optimization variables
  • allows *better initial guesses*, using information about process, which helps avoid "far away" local minima
  • *reduces nonlinearity*, even down to one-step convergence
  • is *numerically stable*, suitable even for highly unstable, e.g. chaotic dynamics

• efficient parallel implementation

• adaptive accuracy discretization strategies

• add'l advantage of multiple shooting
  • state-of-the-art solvers for DAE applicable
  • treatment of discontinuities (hybrid systems) e.g. phase changes, hysteresis, ...

"adaptive accuracy" realized through integrator
An Unstable Test Problem
Unstable Test Problem

- state equations:

\[ \begin{align*}
\dot{x}_1 &= x_2, \\
\dot{x}_2 &= \mu^2 x_1 - (\mu^2 + p^2) \sin pt, \\
\end{align*} \]

\[ x_1(0) = 0, \quad x_2(0) = \pi \quad t \in [0,1] \]

- special solution for "true" parameter value \( p = \pi \):

\[ \begin{align*}
x_1(t) &= \sin \pi t, \\
x_2(t) &= \pi \cos \pi t
\end{align*} \]

- pseudo random measurement noise, \( \sigma = 0.05 \)
Unstable Test Problem - Single Shooting

initial trajectory with $p=1$, and with $p=\text{float}(\pi)$ in 64 bit

initial value problem is extremely ill-conditioned!
Unstable Test Problem - General Solution

\[ x_1 = \sin pt + \varepsilon_1 \sinh \mu t + \varepsilon_2 \cosh \mu t; \quad \varepsilon_1 = \frac{x_2(0) - p}{\mu} \]

\[ x_2 = p \cos pt + \varepsilon_1 \mu \cosh \mu t + \varepsilon_2 \mu \sinh \mu t; \quad \varepsilon_2 = x_1(0) \]

Eigenvalues of \( f_x(t, x(t), p) \) are \( \lambda_{1,2} = \pm \mu \)

Error propagation: \( \exp(\pm \mu t) \)!

\( \mu = 60 \), i.e. error propagation over \([0,1]\) is \( 10^{27} \) - highly unstable!
Unstable Test Problem - Multiple Shooting

- initial trajectory for $p=1$ - convergence after 4 iterations
- parameter estimation problem is well-conditioned!

*works!*
Efficiency of Boundary Value Problem Methods

Theorem

Assumptions

Dense exact data for all states available
Model equations linear in parameters
Initial guesses for states: given data
Length of multiple shooting (resp. collocation) intervals $h \to 0$

Then

One step convergence to true parameter value $p^*$

$$p^1 = p^0 + \Delta p^0 = p^* + O(h)$$

Reduction of Nonlinearity by Decoupling

Advantages of BVP approach also in case of non-dense noisy data
Lotka-Volterra Problem
Lotka-Volterra: Model and Data

\[
\begin{align*}
\dot{x_1} &= -p_1 x_1 + p_2 x_1 x_2 \\
\dot{x_2} &= +p_3 x_2 - p_4 x_1 x_2
\end{align*}
\]

- $x_1$: predators
- $x_2$: preys

DE linear in par's

Data: $\sigma = 5\%$
Comparison: Single vs. Multiple Shooting

Single Shooting
Convergence after 20 iterations

Multiple Shooting
Convergence after 4 iterations

Initial guesses
\[ p_1 = 0.5 \]
\[ p_2 = 0.5 \]
\[ p_3 = -0.5 \]
\[ p_4 = -0.2 \]
Lotka-Volterra: Solution with Multiple Shooting

Initial Trajectory

Solution Trajectory

Initial guesses:
\[ p_1 = 0.5 \quad p_2 = 0.5 \]
\[ p_3 = -0.5 \quad p_4 = -0.2 \]

Solution:
\[ p_1 = 1.01 \pm 0.02 \quad p_2 = 1.01 \pm 0.03 \]
\[ p_3 = 0.99 \pm 0.02 \quad p_4 = 1.01 \pm 0.03 \]
Example Photosynthesis
Photosynthesis: 3-Experiment Initial Trajectories for Multiple Shooting

Multiple shooting with 20 gridpoints

Initial guesses

<table>
<thead>
<tr>
<th></th>
<th>$k_3$</th>
<th>$p_{22}$</th>
<th>$p_{2T}$</th>
<th>$P_{\text{tot}}$</th>
<th>$S$</th>
<th>$I_2$</th>
<th>$k_{\text{lim}}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Exp 1</td>
<td>20</td>
<td>0.5</td>
<td>0.4</td>
<td>10</td>
<td>10</td>
<td>300</td>
<td>1</td>
</tr>
<tr>
<td>Exp 2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>210</td>
</tr>
<tr>
<td>Exp 3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>150</td>
</tr>
</tbody>
</table>
Photosynthesis: 3-Experiment Solution

12 iterations, 3 damped

<table>
<thead>
<tr>
<th></th>
<th>$k_3$</th>
<th>$p_{22}$</th>
<th>$p_{2T}$</th>
<th>$p_{\text{tot}}$</th>
<th>$S$</th>
</tr>
</thead>
<tbody>
<tr>
<td>solution</td>
<td>17.3</td>
<td>0.0710</td>
<td>0.841</td>
<td>11.5</td>
<td>18.1</td>
</tr>
<tr>
<td>standard error$^1$</td>
<td>±0.76</td>
<td>±0.015</td>
<td>±0.015</td>
<td>±0.84</td>
<td>±1.0</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>100%</th>
<th>70%</th>
<th>50%</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$I_2$</td>
<td>$k_{\text{lim}}$</td>
<td>$I_2$</td>
</tr>
<tr>
<td>solution</td>
<td>195.</td>
<td>1.07</td>
<td>143.</td>
</tr>
<tr>
<td>standard error$^1$</td>
<td>±9.0</td>
<td>±0.36</td>
<td>±6.9</td>
</tr>
</tbody>
</table>

estimated; multiplication by 4.5 yields 95% confidence intervals

Baake, Schl. `92
Some Algorithmic Features
The crucial requirement for practical use: numerics must be "derivative-free" for the user!

- adaptive integrators for ODE and relaxed DAE
- fast and error controlled computation of 1st and higher order derivatives
  - combining "automatic differentiation" of model equations and
  - "internal differentiation" of adaptive discretization scheme
- treatment of implicitly given discontinuities and jumps in dynamics
- in forward or reverse (adjoint) mode

Evaluation of CLLS
Computation of 1st and Higher Order Derivatives

e.g. DAESOL, RKFSWT
Bauer et al. '98
Albersmeyer '05
Kirches '06
Parallel Evaluation and Decomposition

1. Evaluation of functions and gradients: parallel on interval level
2. Parallel condensing

Variants: Orthogonal transformations for unstable systems
Block Gauss elimination for stable systems
Fast Sequential Solution
Reduced Generalized Gauss Newton

Idea:
Use initial, multipoint, DAE consistency conditions, ...
to reduce number of directional derivatives of IVP solutions to minimum

1. DAE consistency
   \[ A^y \Delta S^y + A^z \Delta S^z + \alpha = 0 \]
   Computation of solution mf \[ N\delta + \Delta \hat{S}, \delta \text{ free} \]

2. Continuity
   \[ G^y \Delta S^y + G^z \Delta S^z - \Delta S^y_+ + m = 0 \]
   Insertion of solution mf
   \[ \begin{bmatrix} G^y & G^z \end{bmatrix} N\delta + \begin{bmatrix} G^y & G^z \end{bmatrix} \Delta \hat{S} - \Delta S^y_+ + m = 0 \]
   \[ (# \text{Gradients} : # \Delta S^y + # \Delta S^z) \]
   \[ (# \text{Gradients} : # \delta + 1) \]

# Needed directional derivatives = # Degrees of freedom + 1

\[ \rightarrow \text{PDE} \]
Treatment of Condensed System

Large-scale linear constrained system is reduced to condensed system in \( n \) variables

\[
\begin{align*}
\min & \quad ||A_1v + b_1||_\alpha \\
\text{s.t.} & \quad A_{2E}v + b_{2E} = 0 \\
& \quad A_{2I}v + b_{2I} \geq 0 \\
& \quad v \in R^n \\
& \quad \alpha = 1, 2, \infty
\end{align*}
\]

\( n \leq \) number of states + parameters \( n \) small!

Solution: \( l_1, l_\infty \) Modifications of „Adaptive Method“ (Gabasov, Kirillova, Kostina, ...)

\( l_2 \) Orthogonal Transformations & Active Set Strategies + Elimination

Rank determination possible \( \Rightarrow \) regularization
Convergence of Constrained Gauss-Newton

- **local linear convergence** of full step method with asymptotic rate $\kappa$

\[ \kappa := \sup_{\Delta X} \left\| (J^+ (X^* + \Delta X) - J^+ (X^*)) F(X^*) \right\| / \|\Delta X\| < 1, \text{ at } X = X^* \]

- **advantage**: method not attracted by large residual local minima $X^*$, so called "statistically unstable" estimates - cannot be interpreted as continuous deformation of "true parameter"!

- **global convergence**: by efficient new strategies based on "affine invariance" principles - guarantee full step in local domain of convergence

Bock, Kostina, Schl. '00
Statistical Sensitivity Analysis for Constrained Case

- need to know uncertainty of parameter estimate $X^*(\epsilon)$ depending on measurements errors $\epsilon \in N(0, \beta^2 I)$

\[ X^*(\epsilon) \in N(X^*, C) \]

- first order expansion:

\[ X^*(\epsilon) - X^* = -J(X^*)^+ \begin{pmatrix} \epsilon \\ 0 \end{pmatrix}, X^* := X^*(0) \]

- covariance matrix approximation:

\[ C := \mathbb{E} \left( J(X^*)^+ \begin{pmatrix} \epsilon \\ 0 \end{pmatrix} \begin{pmatrix} \epsilon \\ 0 \end{pmatrix}^T \right) = J(X^*)^+ \begin{pmatrix} \beta^2 I & 0 \\ 0 & 0 \end{pmatrix} J(X^*)^+^T \]
Statistical Sensitivity Analysis for Constrained Case

- confidence ellipsoid $G$, includes “true value” with error probability $\approx \alpha$

$$G := \{ X \mid X - X^* = -J^+ (X^*) \varepsilon, \| \varepsilon \|^2 \leq \gamma(\alpha) \}, \gamma(\alpha) := v_1 F^{1-\alpha} (v_1, v_2)$$

- Lemma: $G$ can be enclosed by confidence box

$$= 2\delta_i$$

- need to compute only $C_{ii}^{1/2}$, "standard deviations" of parameters

$\rightarrow$ basis for optimum experimental design
Applications
Initial Satellite Orbit Determination

in cooperation with
Application: Orbit Determination of Satellites

aim:
• fast and reliable determination of satellite orbit after faulty launch
• to perform correction maneuvers

*Spaceflight Now*, Feb 21, 2002: "...the Ariane 5 launcher had propelled the ARTEMIS satellite into a transfer orbit that was lower then expected, with the apogee at only 17000 km rather than the nominal 36000 km..."

• ill-conditioned problem, 6 orbit parameters, only range and angle measurements from 2 different ground stations
• very poor initial guesses
Satellite Orbit Determination: Model

Kepler equations augmented by perturbations

\[ \dot{r}(t) = v(t), \quad \dot{v}(t) = -\frac{GM}{\|r(t)\|^3} r(t) + \text{pert}(r(t), v(t), t) \]

due to external forces

- gravitational forces of sun and moon
- inhomogeneities of earth’s gravitational field
- air drag
- solar radiation pressure

Results in small but complex nonlinear differential equation system in six states with discontinuities in right-hand side

Montenbruck, Gill '00
Measurements from Different Stations

Perth ground station, Australia
-31.803°, +115.885°

Artemis measurements

+ 76 range
+ 27 azimuth and elevation angle pairs

Malindi ground station, Kenya
-2.996°, 40.196°
Artemis Test Scenarios

Parameters to be estimated: Six orbit elements at a given time (epoch)

Initial guesses for parameters

\[
s_0^\beta (\beta) = (1 - \beta) \begin{pmatrix} s_{0,\text{sol}} + \beta \end{pmatrix} s_{0,\text{expected}}
\]

= \begin{pmatrix} 5046.07 \\ -7018.14 \\ -263.990 \\ 7.87936 \\ 2.02064 \\ -0.401513 \end{pmatrix}

ESA reference (\beta=0) solution

Initial guesses for Multiple Shooting variables by projection

range measurements

angle measurements

corresponding orbits
**Artemis Results**

**Comparison of Single Shooting and Multiple Shooting**

<table>
<thead>
<tr>
<th>ESA reference test case</th>
<th>BAHN (single shooting)</th>
<th>PARFITBAHN (single shooting)</th>
<th>PARFITBAHN (multiple shooting)</th>
</tr>
</thead>
<tbody>
<tr>
<td>β</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.01</td>
<td>6</td>
<td>5</td>
<td>4</td>
</tr>
<tr>
<td>0.1</td>
<td>No convergence</td>
<td>19</td>
<td>5</td>
</tr>
<tr>
<td>0.2</td>
<td>No convergence</td>
<td>No convergence</td>
<td>10</td>
</tr>
<tr>
<td>0.4</td>
<td>No convergence</td>
<td>No convergence</td>
<td>12</td>
</tr>
<tr>
<td>0.6</td>
<td>No convergence</td>
<td>No convergence</td>
<td>19</td>
</tr>
<tr>
<td>0.8</td>
<td>No convergence</td>
<td>No convergence</td>
<td>13</td>
</tr>
<tr>
<td>1.0</td>
<td>No convergence</td>
<td>No convergence</td>
<td>23</td>
</tr>
<tr>
<td>1.5</td>
<td>No convergence</td>
<td>No convergence</td>
<td>32</td>
</tr>
<tr>
<td>2.0</td>
<td>No convergence</td>
<td>No convergence</td>
<td>24</td>
</tr>
<tr>
<td>3.0°</td>
<td>Hyperbolic</td>
<td>No convergence</td>
<td>30</td>
</tr>
<tr>
<td>5.0</td>
<td>Hyperbolic</td>
<td>No convergence</td>
<td>21</td>
</tr>
</tbody>
</table>

*Multiple Shooting works for very poor initial parameter guesses*

Lenz, Bock, Schl., Kostina, Gienger, Ziegler ‘10
Application: Orbit Determination of Satellites

red: nominal, green: actual orbit of Artemis
Transport and Degradation of Xenobiotics in Soil
Transport and Degradation of Xenobiotics in Soil

- Investigation of fate of xenobiotics
- Expensive lysimeter experiments for registration
- To be replaced by computer experiments
- Here: parameter estimation
- Later: optimal mini-lysimeter experiments
  - optimal irrigation
  - optimal solute application
  - Optimal sampling

Altmann-Dieses, Bock, Schl., Richter '02
Field experiment: Water Transport (K. Aden)

- loamy sand without vegetation
- time-domain reflectrometry (TDR): hourly
- measurements of water content in 7, 15 and 20 cm
PDE-Model: Richards Equation

\[ \frac{\partial \theta}{\partial t} = \frac{\partial}{\partial z} \left( D(\theta) \frac{\partial \theta}{\partial z} - K(\theta) \right) \]

\[ K(\theta) = K_s \Theta^{1/2} \left[ 1 - \left( 1 - \frac{\Theta^n}{(n-1)} \right)^{1-1/n} \right]^2, \quad \Theta = \frac{\theta - \theta_r}{\theta_s - \theta_r} \]

\[ D(\theta) = K(\theta) \tilde{C}(\theta) \]

\[ \tilde{C}(\theta) = \frac{1}{\alpha n m} \left( \Theta^{-1/m} - 1 \right)^{-m} \Theta^{-1/m} \frac{1}{\theta - \theta_r}, \quad m = 1 - \frac{1}{n} \]

- Initial condition: Linear interpolation of 7cm, 15cm, 20cm at begin of experiments (Oct 28, 1997)
- Upper boundary: Dirichlet condition (TDR data in 7 cm)
- Lower boundary: Dirichlet condition (TDR data in 20 cm)
Transport and Degradation of Xenobiotics in Soil

Result: Estimates for $n$, and $K_s$

<table>
<thead>
<tr>
<th></th>
<th>guess</th>
<th>estimate</th>
</tr>
</thead>
<tbody>
<tr>
<td>$n$</td>
<td>1.5</td>
<td>$1.262 \pm 0.0024$</td>
</tr>
<tr>
<td>$\alpha$</td>
<td>0.05</td>
<td>$0.0324 \pm 0.0024$</td>
</tr>
<tr>
<td>$K_s$</td>
<td>35.0</td>
<td>$20.92 \pm 1.68$</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>$\alpha$</th>
<th>$K_s$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$n$</td>
<td>0.14</td>
<td>-0.61</td>
</tr>
<tr>
<td>$\alpha$</td>
<td>-</td>
<td>-0.94</td>
</tr>
</tbody>
</table>
Identification of Cerebral Palsy Gaits
Cerebral Palsy Gaits

before surgery

after surgery

Assumption: Movement is optimal

Task: Find suitable optimal control problem!

... K. Hatz in coop with S. Wolf (Orthopedics HD)
Identification of Cerebral Palsy Gaits

Model for patient’s motion:

- 48 states $\mathbf{q}$: 3 global coordinates, 3 global angles, 18 local joint angles (generalized coordinates), and the corresponding velocities $\dot{\mathbf{q}}$
- 18 controls $\mathbf{u}$
- constrained multibody system
- formulated with HuMAoS Toolbox (INRIA, France)
Inverse Optimal Control for Cerebral Palsy Gaits

\[
\min_{y,u,p,\gamma} \sum_{i,j} \frac{(\eta_{ij} - b_{ij}(t_i, y(t_i)))^2}{\sigma_{ij}^2}
\]

s.t.

\[
\min_{y:=(q, \dot{q}), u} \sum_k \gamma_k C_k [q, \dot{q}, u]
\]

s.t.

\[
\dot{y}(t) = f(t, y(t), u(t), p) \quad 0 \leq c(t, y(t), u(t), p) \quad 0 = r(y(t_0), y(T))
\]

\[
\sum_k \gamma_k = 1, \quad \gamma_k \geq 0 \quad \forall k
\]

Possible criteria \(C_k\): stability, energy, duration, ...

Challenges: discontinuities in states, large-scale, ...
Efficient Direct All-at-once Approach

after discretization of optimal control problem with direct multiple shooting

\[
\min_{x, p, \gamma, \lambda, \mu} \sum_{i,j} \frac{(\eta_{ij} - b_{ij}(t_i, y(t_i)))^2}{\sigma_{ij}^2}
\]

s.t. \(0 = y(t_{i+1}; t_i, s_i, w_i, p) - s_{i+1}\)
\(0 \leq \tilde{c}(s_0, \ldots, s_{nt}, w, p)\)
\(0 = r(s_0, s_{nt})\)
\(0 = \nabla_x L(x, \gamma, \lambda, \mu)\)
\(0 \leq \mu\)
\(0 \geq \mu \tilde{c}(s_0, \ldots, s_{nt}, w, p)\)
\(1 = \sum_k \gamma_k\)
\(0 \leq \gamma_k \quad \forall k\)

Multiple Shooting discretization
x:=(s,w)
s multiple shooting variables
w control variables

KKT conditions
L Lagrangian
\(\lambda, \mu\) adjoints

objective regularization

large scale constrained ls-problem with complementarity condition
Measurements for Cerebral Palsy Gait

Vicon data
Identified Cerebral Palsy Gait

measured gait

identified gait

very good agreement
Measured and Estimated OCP Gait

... K. Hatz

green: measured gait blue: estimated optimal OCP gait
Summary

- Parameter Estimation in Differential Equations
- Optimization Boundary Value Problems
  - Hierarchical Optimization Problems
- Direct Multiple Shooting
- Some Applications

Ongoing work

Real-Time Moving Horizon Estimation
Optimum Experimental Design
Thank you very much for your attention!