Dedication: Germund Dahlquist, 1925-2005.
http://www.siam.org/news.php?id=54

Numerical Questions in ODE Boundary Value Problems

M.R. Osborne

Mathematical Sciences Institute
Australian National University

CTAC'06 JCU Townsville

Outline

The boundary value problem

Consider the differential equation

$$\frac{d\mathbf{x}}{dt} = \mathbf{f}(t, \mathbf{x}),$$

where $\mathbf{x} \in R^m,\ \mathbf{f} \in R^m \times R \to R^m$, together with the boundary conditions

$$B_0\mathbf{x}(0) + B_1\mathbf{x}(1) = \mathbf{b}.$$

Special cases include the initial value problem where $B_0 = I,\ B_1 = 0$, while multi-point problems can be transformed to BVP form. Relatively weak conditions guarantee IVP solution locally. Non-trivial BVP involves a global statement. Can derive conditions using, for example, Newton-Kantorovich theorem.

Linear case

$$\mathbf{f}(t,\mathbf{x}) = A(t)\mathbf{x} + \mathbf{q}(t).$$

Let fundamental matrix $X(t,\xi)$ satisfy the IVP

$$\frac{dX}{dt} = A(t)X, \quad X(\xi,\xi) = I$$

then BVP has a solution provided $(B_0 + B_1X(1,0))$ has a bounded inverse. The Green's matrix is

$$\begin{aligned} G(t,s) &= X(t)\left[B_0X(0) + B_1X(1)\right]^{-1}B_0X(0)X^{-1}(s), t > s, \\ &= -X(t)\left[B_0X(0) + B_1X(1)\right]^{-1}B_1X(1)X^{-1}(s), t < s. \end{aligned}$$

Note G does not depend on the initial condition on X. The magnitude of G is an indicator of problem stability. Set *stability constant* $\alpha = \max_t \|G(t,s)\|_2$. go BVDI

Estimation provides target problem

Specialise $\mathbf{f} \leftarrow \mathbf{f}(t, \mathbf{x}, \beta)$ where $\beta \in R^p$. Given data

$$\mathbf{y}_i = H\mathbf{x}(t_i, \beta^*) + \varepsilon_i, \quad i = 1, 2, \cdots, n,$$

where $H : R^m \rightarrow R^k$, $nk > m + p$, and $\varepsilon_i \sim N\left(0, \sigma^2 I\right)$, estimate β.

Equivalent smoothing problem: $\mathbf{x} \leftarrow \begin{bmatrix} \mathbf{x}(t) \\ \beta \end{bmatrix}$, $\mathbf{f} \leftarrow \begin{bmatrix} \mathbf{f}(t, \mathbf{x}) \\ 0 \end{bmatrix}$.

Note problem is over-determined as stated and need not possess an exact solution. Thus seek a solution of best fit in an appropriate sense. Assume this problem has a well determined solution for n, the number of observations, large enough.

Data collection

1. Practical considerations can restrict interval on which observations can be made. An example is transient signals. Assumption is that sequences of observations $\{t_1, t_2, \cdots, t_n\} \subset [0,1]$ are possible for arbitrarily large n. The condition of a planned experiment is useful:

$$\frac{1}{n}\sum_{i=1}^{n} v(t_i) \to \int_0^1 v(t)d\rho(t).$$

   Here ρ implies an experimental mechanism.

2. Measurements for arbitrarily large t contain useful information on signal $\mathbf{x}(t)$. Stationary problems provide examples. Frequency estimation is problem of interest.

The problem setting

Mesh selection for integrating the ODE system is conditioned by two important considerations:

- The asymptotic analysis of the effects of noisy data on maximum likelihood estimates of the parameters shows that this gets small no faster than $O\left(n^{-1/2}\right)$ under planned experiment conditions. A higher rate ($O\left(n^{-3/2}\right)$) is theoretically possible in maximum likelihood estimates in the frequency estimation problem but direct maximization is not the way to obtain them.
- It is not difficult to obtain ODE discretizations that give errors at most $O\left(n^{-2}\right)$.

The problem setting

Mesh selection for integrating the ODE system is conditioned by two important considerations:

- The asymptotic analysis of the effects of noisy data on maximum likelihood estimates of the parameters shows that this gets small no faster than $O\left(n^{-1/2}\right)$ under planned experiment conditions. A higher rate ($O\left(n^{-3/2}\right)$) is theoretically possible in maximum likelihood estimates in the frequency estimation problem but direct maximization is not the way to obtain them.
- It is not difficult to obtain ODE discretizations that give errors at most $O\left(n^{-2}\right)$.

This suggests that the trapezoidal rule provides an adequate integration method. It is known to be endowed with attractive properties.

The objective

Estimation principles (least squares, maximum likelihood) consider the objective:

$$F(\mathbf{x}_c, \beta) = \sum_{i=1}^{n} \|\mathbf{y}_i - H\mathbf{x}(t_i, \beta)\|^2.$$

Methods differ in manner of generating comparison function values $\mathbf{x}(t_i, \beta),\ i = 1, 2, \cdots, n$.
Embedding: $\mathbf{x}(t_i, \beta, \mathbf{b})$ satisfies BVP

$$\frac{d\mathbf{x}}{dt} = \mathbf{f}(t, \mathbf{x}, \beta), \quad B_0\mathbf{x}(0) + B_1\mathbf{x}(1) = \mathbf{b}.$$

Introduces extra parameters $\mathbf{b}$. Needs method for choosing $B_0,\ B_1$. Must solve boundary value problem at each step. go GNM

The objective

Estimation principles (least squares, maximum likelihood) consider the objective:

$$F(\mathbf{x}_c, \beta) = \sum_{i=1}^{n} \|\mathbf{y}_i - H\mathbf{x}(t_i, \beta)\|^2.$$

Methods differ in manner of generating comparison function values $\mathbf{x}(t_i, \beta),\ i = 1, 2, \cdots, n$.

Simultaneous: ODE discretization information added as constraints

$$\mathbf{c}_i(\mathbf{x}_c) = \mathbf{x}_{i+1} - \mathbf{x}_i - \frac{h}{2}(\mathbf{f}_{i+1} + \mathbf{f}_i), \quad i = 1, 2, \cdots, n-1,$$

with $\mathbf{x}_i = \mathbf{x}(t_i, \beta)$. Methods typically correct solution and parameter estimates simultaneously. go SQP

Initial value stability (IVS)

Here the problem considered is:

$$\frac{d\mathbf{x}}{dt} = \mathbf{f}(t, \mathbf{x}), \quad \mathbf{x}(0) = \mathbf{b}.$$

The stability requirement is that solutions with close initial conditions $\mathbf{x}_1(0)$, $\mathbf{x}_2(0)$ remain close in an appropriate sense.

- $\|\mathbf{x}_1(t) - \mathbf{x}_2(t)\| \to 0, \ t \to \infty$. strong IVS.

Initial value stability (IVS)

Here the problem considered is:

$$\frac{d\mathbf{x}}{dt} = \mathbf{f}(t, \mathbf{x}), \quad \mathbf{x}(0) = \mathbf{b}.$$

The stability requirement is that solutions with close initial conditions $\mathbf{x}_1(0)$, $\mathbf{x}_2(0)$ remain close in an appropriate sense.

- $\|\mathbf{x}_1(t) - \mathbf{x}_2(t)\| \to 0,\ t \to \infty$. strong IVS.
- $\|\mathbf{x}_1(t) - \mathbf{x}_2(t)\|$ remains bounded as $t \to \infty$. weak IVS.

Initial value stability (IVS)

Here the problem considered is:

$$\frac{d\mathbf{x}}{dt} = \mathbf{f}(t, \mathbf{x}), \quad \mathbf{x}(0) = \mathbf{b}.$$

The stability requirement is that solutions with close initial conditions $\mathbf{x}_1(0)$, $\mathbf{x}_2(0)$ remain close in an appropriate sense.

- $\|\mathbf{x}_1(t) - \mathbf{x}_2(t)\| \to 0,\ t \to \infty$. strong IVS.
- $\|\mathbf{x}_1(t) - \mathbf{x}_2(t)\|$ remains bounded as $t \to \infty$. weak IVS.
- Computation introduces idea of stiff discretizations which preserve the stability characteristics of the original equation in the sense that decaying solutions are mapped onto decaying solutions. This is one area where there are genuine nonlinear results - for example, J.B's work on BN stability of Runge-Kutta methods.

Not all relevant IVPs are stable

The classical BVP solution method of multiple shooting provides an example. This requires computation of the multiple shooting matrix:

$$\begin{bmatrix} -X(t_2,t_1) & I & & & \\ & -X(t_3,t_2) & I & & \\ & & \cdots & \cdots & \\ B_0 & & & & B_1 \end{bmatrix}.$$

The IVP for computing $X(t_{i+1},t_i)$ could well be unstable in both forward and backward directions.

Dahlquist's famous "consistency + stability implies convergence as $h \to 0$" theorem does not require IVP stability, but it's setting implies exact arithmetic.

Multiple shooting in this form appears to require accurate computation of all solutions with the $\{t_i\}$ serving as controls. That is a weakness.

Constant coefficient case

Here

$$\mathbf{f}(t, \mathbf{x}) = A\mathbf{x} - \mathbf{q}$$

If A is non-defective then weak IVS requires the eigenvalues $\lambda_i(A)$ to satisfy $Re\lambda_i \leq 0$ while this inequality must be strict for strong IVS.

A one-step discretization of the ODE (ignoring $\mathbf{q}$ contribution) can be written

$$\mathbf{x}_{i+1} = T_h(A)\,\mathbf{x}_i.$$

where $T_h(A)$ is the amplification matrix. Here a stiff discretization requires the stability inequalities to map into the condition $|\lambda_i(T_h)| \leq 1$.

For the trapezoidal rule

$$\begin{aligned} |\lambda_i(T_h)| &= \left|\frac{1 + h\lambda_i(A)/2}{1 - h\lambda_i(A)/2}\right|, \\ &\leq 1 \text{ if } Re\{\lambda_i(A)\} \leq 0. \end{aligned}$$

Dichotomy: Key paper is de Hoog and Mattheij.

go DEBV This is the structural property that connects linear BVP stability with the detailed behaviour of the range of possible solutions.

Weak form: $\exists$ projection P depending on choice of X such that, given

$$S_1 \leftarrow \{XP\mathbf{w},\ \mathbf{w} \in R^m\}, \quad S_2 \leftarrow \{X(I-P)\mathbf{w},\ \mathbf{w} \in R^m\},$$

$$\phi \in S_1 \Rightarrow \frac{|\phi(t)|}{|\phi(s)|} \leq \kappa, \quad t \geq s,$$

$$\phi \in S_2 \Rightarrow \frac{|\phi(t)|}{|\phi(s)|} \leq \kappa, \quad t \leq s.$$

Computational context requires modest κ for $t, s \in [0,1]$.

If X satisfies $B_0X(0) + B_1X(1) = I$ then $P = B_0X(0)$ is a suitable projection in sense that for separated boundary conditions can take $\kappa = \alpha$. There is a basic equivalence between stability and dichotomy.

BVS restricts possible discretizations

- Dichotomy projection separates increasing and decreasing solutions. *Compatible* BC's pin down decreasing solutions at 0, growing solutions at 1.

BVS restricts possible discretizations

- Dichotomy projection separates increasing and decreasing solutions. *Compatible* BC's pin down decreasing solutions at 0, growing solutions at 1.
- Discretization needs similar property so given BC's exercise same control.

BVS restricts possible discretizations

- Dichotomy projection separates increasing and decreasing solutions. *Compatible* BC's pin down decreasing solutions at 0, growing solutions at 1.
- Discretization needs similar property so given BC's exercise same control.
- This requires solutions of ODE which are increasing (decreasing) in magnitude to be mapped into solutions of discretization which are increasing (decreasing) in magnitude.

BVS restricts possible discretizations

- Dichotomy projection separates increasing and decreasing solutions. *Compatible* BC's pin down decreasing solutions at 0, growing solutions at 1.
- Discretization needs similar property so given BC's exercise same control.
- This requires solutions of ODE which are increasing (decreasing) in magnitude to be mapped into solutions of discretization which are increasing (decreasing) in magnitude.

This property called di-stability by England and Mattheij who showed the TR is di-stable in constant coefficient case.

$$\lambda(A) > 0 \Rightarrow \left| \frac{1 + h\lambda(A)/2}{1 - h\lambda(A)/2} \right| > 1.$$

Bob Mattheij's example

Consider the differential system defined by

$$A(t) = \begin{bmatrix} 1-19\cos 2t & 0 & 1+19\sin 2t \\ 0 & 19 & 0 \\ -1+19\sin 2t & 0 & 1+19\cos 2t \end{bmatrix},$$

$$\mathbf{q}(t) = \begin{bmatrix} e^t\left(-1+19\left(\cos 2t-\sin 2t\right)\right) \\ -18e^t \\ e^t\left(1-19\left(\cos 2t+\sin 2t\right)\right) \end{bmatrix}.$$

Here the right hand side is chosen so that $\mathbf{z}(t) = e^t\mathbf{e}$ satisfies the differential equation. The fundamental matrix displays the fast and slow solutions:

$$X(t,0) = \begin{bmatrix} e^{-18t}\cos t & 0 & e^{20t}\sin t \\ 0 & e^{19t} & 0 \\ -e^{-18t}\sin t & 0 & e^{20t}\cos t \end{bmatrix}.$$

go OBC

Bob Mattheij's example

For boundary data with two terminal conditions and one initial condition :

$$B_0 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix}, \; B_1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \; \mathbf{b} = \begin{bmatrix} e \\ e \\ 1 \end{bmatrix},$$

the trapezoidal rule discretization scheme gives the following results.

	$\Delta t = .1$			$\Delta t = .01$		
$\mathbf{x}(0)$	1.0000	.9999	.9999	1.0000	1.0000	1.0000
$\mathbf{x}(1)$	2.7183	2.7183	2.7183	2.7183	2.7183	2.7183

Table: Boundary point values - stable computation

These computations are apparently satisfactory.

Bob Mattheij's example

For two initial and one terminal condition:

$$B_0 = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix}, \; B_1 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \; \mathbf{b} = \begin{bmatrix} 1 \\ e \\ 1 \end{bmatrix}.$$

The results are given in following Table.

	$\Delta t = .1$			$\Delta t = .01$		
$\mathbf{x}(0)$	1.0000	.9999	1.0000	1.0000	1.0000	1.0000
$\mathbf{x}(1)$	-7.9+11	2.7183	-4.7+11	2.03+2	2.7183	1.31+2

Table: Boundary point values - unstable computation

The effects of instability are seen clearly in the first and third solution components.

Nonlinear stability

The IVP/BVP stability requirements are restrictive in sense that solutions must not depart from classification as increasing/decreasing.

Nonlinear stability

The IVP/BVP stability requirements are restrictive in sense that solutions must not depart from classification as increasing/decreasing.
Important conflicting examples occur in dynamical systems. These

- can have a stable character - for example, limiting trajectories which attract neighboring orbits;
- clearly cannot satisfy the IVP/BVP stability requirements.

Nonlinear stability

The IVP/BVP stability requirements are restrictive in sense that solutions must not depart from classification as increasing/decreasing.
Important conflicting examples occur in dynamical systems. These

- can have a stable character - for example, limiting trajectories which attract neighboring orbits;
- clearly cannot satisfy the IVP/BVP stability requirements.

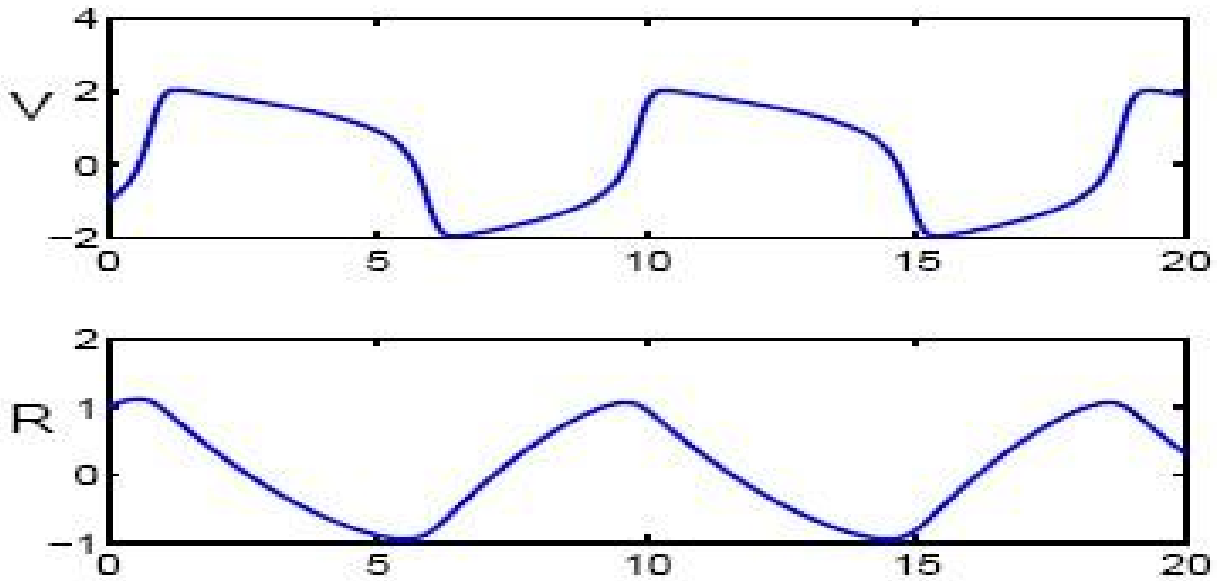
Limit cycle behavior provides a familiar example that is of this type. Also they can share some of the properties of stationary processes. Can algorithms for estimating frequency in such systems possess the $O(n^{-3/2})$ convergence rate?

Example 1 - preprint Hooker et al

FitzHugh-Nagumo equations $\alpha = .2, \beta = .2, \gamma = 1.$

$$\frac{dV}{dt} = \gamma\left(V - \frac{V^3}{3} + R\right),$$

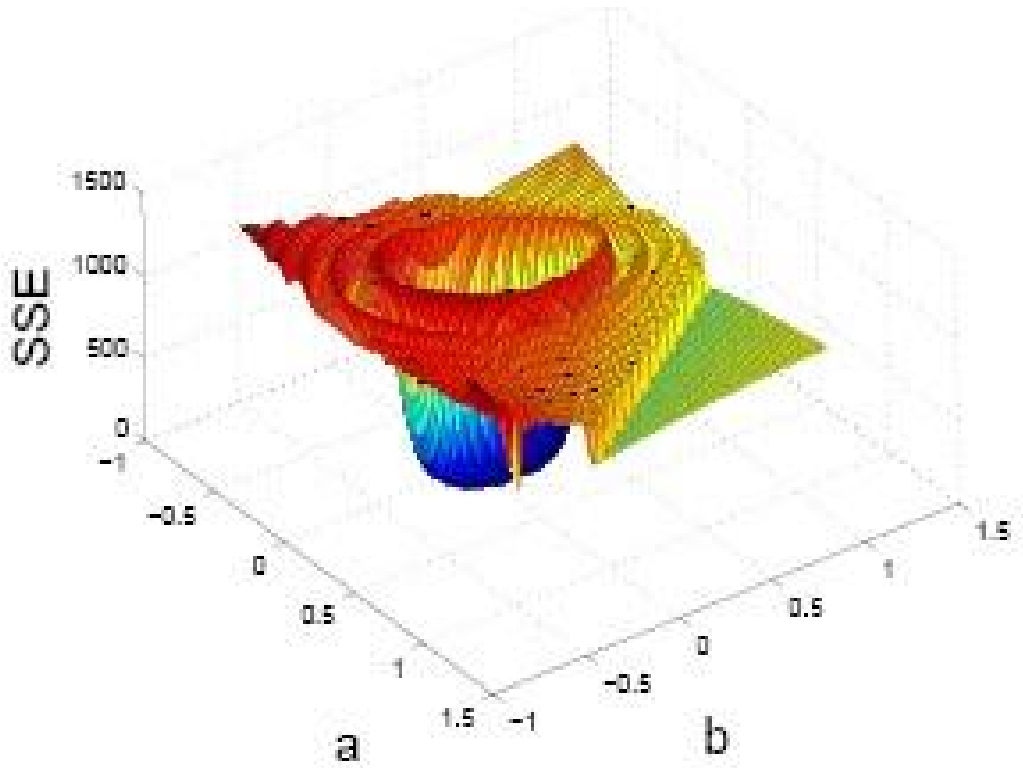
$$\frac{dR}{dt} = -\frac{1}{\gamma}\left(V - \alpha - \beta R\right).$$



Example 1 - preprint Hooker et al

FitzHugh-Nagumo equations $\alpha = .2, \beta = .2, \gamma = 1.$

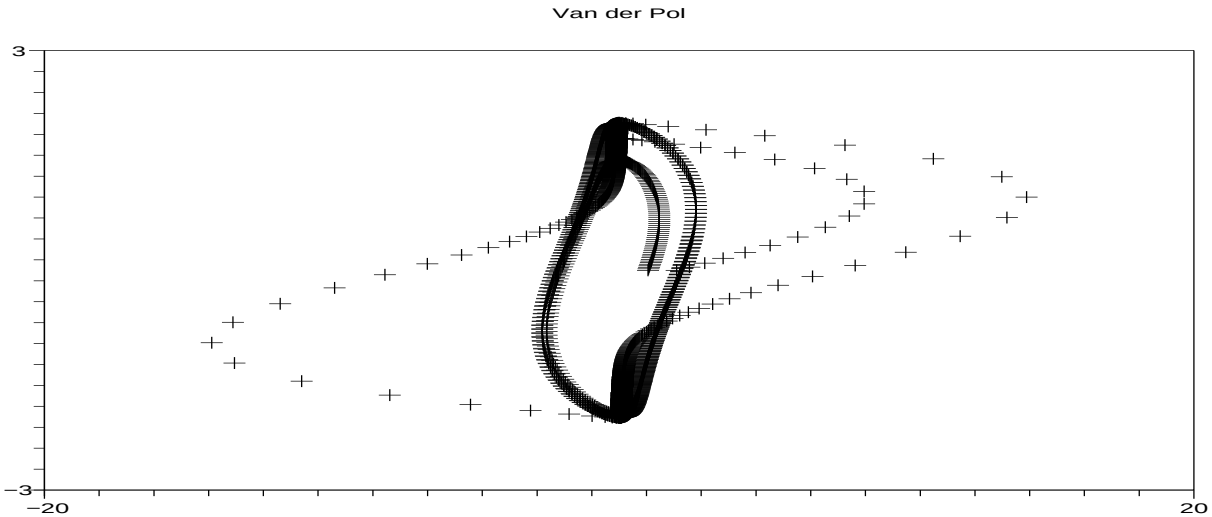
$$\frac{dV}{dt} = \gamma\left(V - \frac{V^3}{3} + R\right),$$
$$\frac{dR}{dt} = -\frac{1}{\gamma}\left(V - \alpha - \beta R\right).$$



Example 2 - Van der Pol equation 1

$$\frac{d^2x}{dt^2} - \lambda\left(1 - x^2\right)\frac{dx}{dt} + x = 0.$$

Reliable,"difficult" ODE example with difficulty increasing with λ.
scilab plot shows convergence to limit cycle for $\lambda = 1, 10$.

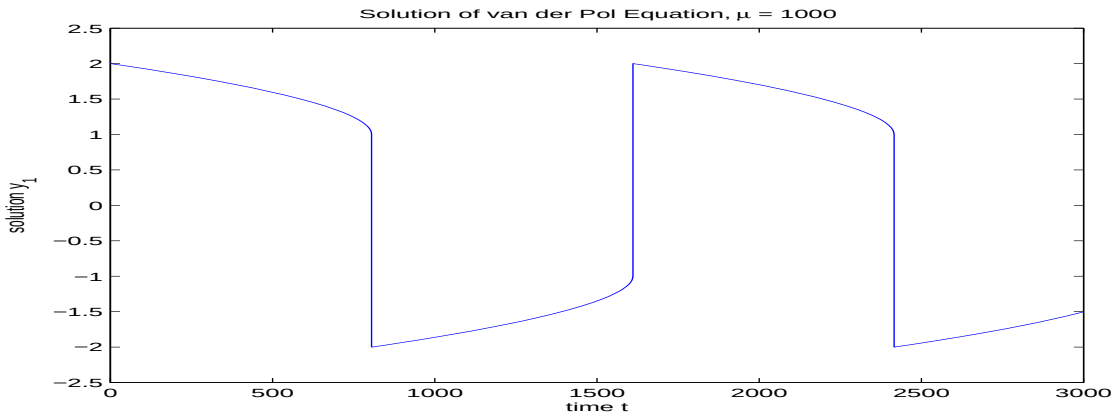


Example 2 - Van der Pol equation 2

Matlab also uses this example but result given is less useful as it gives state information but not derivative. $\lambda = 1000$. Also starting values are rather special as:

$$x(0) = 2 + \frac{1}{3}\alpha\lambda^{-4/3} - \frac{16}{27}\lambda^{-2}\ln(\lambda) + O(\lambda^{-2})$$

where $\alpha = 2.33811....$



Example 2 - BVP formulation 1

Transformation $s = 4t/T$ puts $1/2$ period onto $[0, 2]$. Set $x_3 = T/4$. The ODE becomes

$$\begin{aligned}\frac{dx_1}{ds} &= x_2,\\ \frac{dx_2}{ds} &= \lambda\left(1 - x_1^2\right)x_2x_3 - x_1x_3^2,\\ \frac{dx_3}{ds} &= 0.\end{aligned}$$

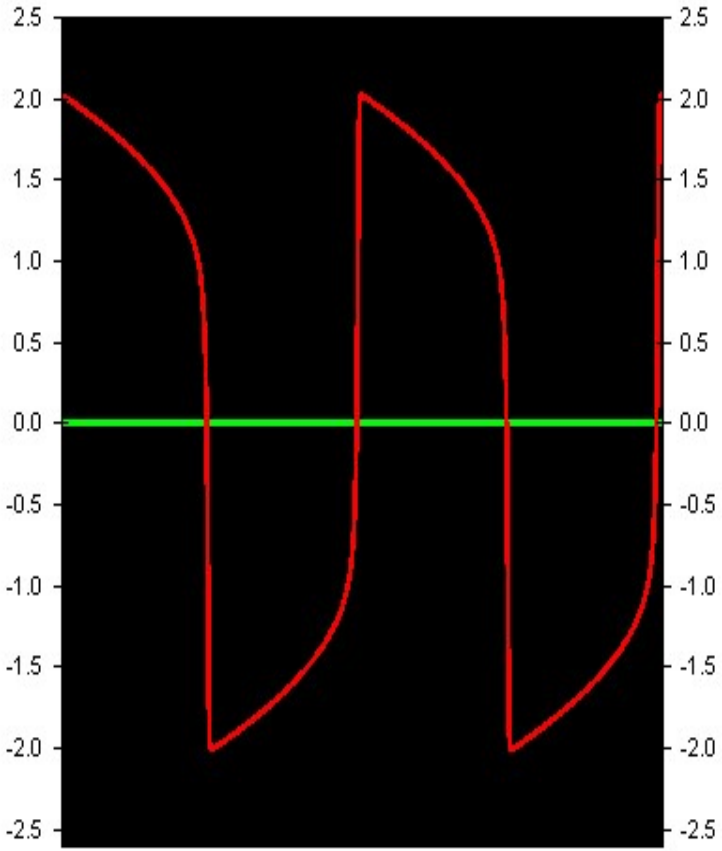
Boundary data is

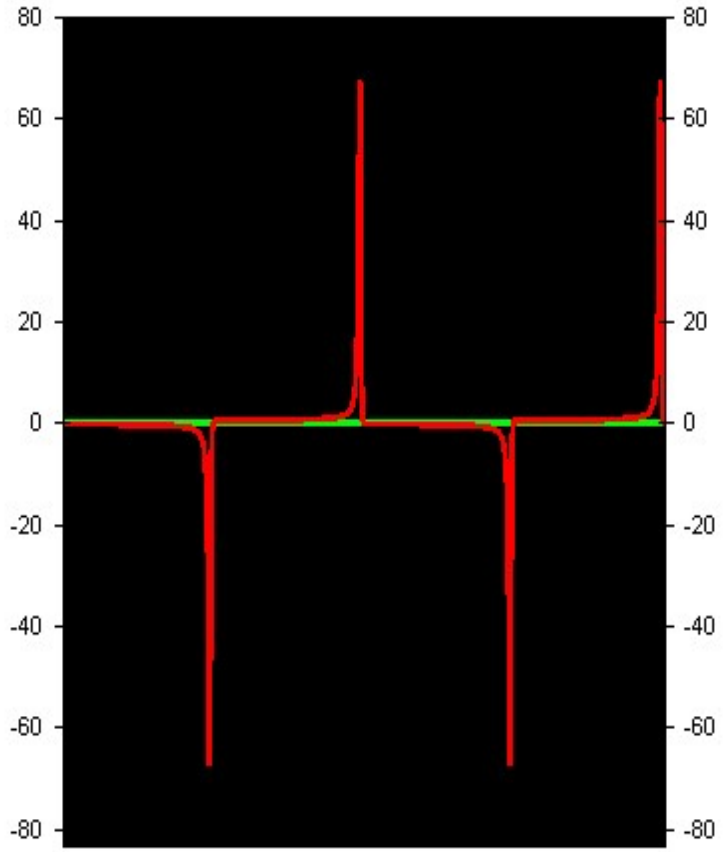
$$B_0 = \begin{bmatrix} 0 & 1 & 0\\ 0 & 0 & 0\\ 1 & 0 & 0\end{bmatrix}, B_1 = \begin{bmatrix} 0 & 0 & 0\\ 0 & 1 & 0\\ 1 & 0 & 0\end{bmatrix}, \mathbf{b} = 0.$$

Initial conditions good for $\lambda = 1$ and work for $\lambda \leq 5$.
Continuation with $\Delta\lambda = 1$ used for higher values.
$h = 1/100, 1/1000$.

Example 2 - BVP formulation 2

BVP results for $\lambda = 10$. Extra values by reflection.





Stability consequences

The ODE stability conditions provide sharp distinctions - in part because they are specifying global properties. Computational requirements force compromise.

In the IVP this is provided by various control devices: for example, automatic step length control. Two classes of computational stability problem:

- Difference approximation does not satisfy Dahlquist root condition $\rho(1) = 0; \rho(t) = 0, t \neq 1, \Rightarrow |t| < 1$. In this case errors grow with n and so are unbounded as $h \rightarrow 0$.
- In unstable problems a computed slow solution will be swamped eventually as a result of the growth of rounding error induced perturbations which grow like $\gamma \exp(Kt)$.

Stability consequences

The ODE stability conditions provide sharp distinctions - in part because they are specifying global properties. Computational requirements force compromise.

In BVP fudge dichotomy considerations to finite interval and ask for "moderate" κ. Can write down the inverse of the multiple shooting matrix as $h \to 0$ limit of corresponding inverses of discretization matrices. The limit can then be interpreted using the Green's matrix. Need to take advantage of di-stability. In practice a strictly unstable BVP is associated with a sensitive Newton iteration. Available tools include:

- adaptive mesh control;
- continuation.

System factorization

go OPT1 First problem is to set suitable boundary conditions. Expect good boundary conditions should lead to a relatively well conditioned linear system. Write the trapezoidal rule discretization as

$$\mathbf{c}_i\left(\mathbf{x}_i, \mathbf{x}_{i+1}\right) = \mathbf{c}_{ii}(\mathbf{x}_i) + \mathbf{c}_{i(i+1)}(\mathbf{x}_{i+1}).$$

Consider the factorization of the difference equation (gradient) matrix with first column permuted to end:

$$\left[\begin{array}{cccc|c} C_{12} & & & & C_{11} \\ C_{21} & C_{22} & & & \\ & & & & \\ \hline & & C_{(n-1)(n-1)} & C_{(n-1)n} & 0 \end{array}\right] \rightarrow Q \left[\begin{array}{ccc|c} & U & & V \\ \hline 0 & \cdots & H & G \end{array}\right]$$

This step is independent of the boundary conditions. go SVE

Optimal boundary conditions

The boundary conditions can be inserted at this point. This gives the system with matrix $\begin{bmatrix} H & G \\ B_1 & B_0 \end{bmatrix}$ to solve for $\mathbf{x}_1$, $\mathbf{x}_n$. Orthogonal factorization again provides a useful strategy.

$$\begin{bmatrix} H & G \end{bmatrix} = \begin{bmatrix} L & 0 \end{bmatrix} \begin{bmatrix} S_1^T \\ S_2^T \end{bmatrix}$$

It follows that the system determining $\mathbf{x}_1$, $\mathbf{x}_n$ is best conditioned by choosing

$$\begin{bmatrix} B_1 & B_0 \end{bmatrix} = S_2^T.$$

The conditions depend only on the ODE.

BC's for Mattheij example

go MatEx The "optimal" boundary matrices corresponding to $h = .1$ are given in the Table. These confirm the importance of weighting the boundary data to reflect the stability requirements of a mix of fast and slow solutions. The solution does not differ from that obtained when the split into fast and slow was correctly anticipated.

B_1			B_2		
.99955	0.0000	.02126	-.01819	0.0000	-.01102
0.0000	0.0000	0.0000	0.0000	1.0000	0.0000
.02126	0.0000	.00045	.85517	0.0000	.51791

Table: Optimal boundary matrices when $h = .1$

Gauss-Newton details

Let $\nabla_{(\beta,b)}\mathbf{x} = \left[\frac{\partial \mathbf{x}}{\partial \beta}, \frac{\partial \mathbf{x}}{\partial \mathbf{b}}\right]$, $\mathbf{r}_i = \mathbf{y}_i - H\mathbf{x}(t_i, \beta, \mathbf{b})$ then the gradient of F is

$$\nabla_{(\beta,b)}F = -2\sum_{i=1}^{n} \mathbf{r}_i^T H \nabla_{(\beta,b)}\mathbf{x}_i.$$

The gradient terms wrt β are found by solving the BVP's

$$B_0 \frac{\partial \mathbf{x}}{\partial \beta}(0) + B_1 \frac{\partial \mathbf{x}}{\partial \beta}(1) = 0,$$

$$\frac{d}{dt}\frac{\partial \mathbf{x}}{\partial \beta} = \nabla_x \mathbf{f} \frac{\partial \mathbf{x}}{\partial \beta} + \nabla_\beta \mathbf{f},$$

Gauss-Newton details

Let $\nabla_{(\beta,b)}\mathbf{x} = \left[\frac{\partial \mathbf{x}}{\partial \beta}, \frac{\partial \mathbf{x}}{\partial \mathbf{b}}\right]$, $\mathbf{r}_i = \mathbf{y}_i - H\mathbf{x}(t_i, \beta, \mathbf{b})$ then the gradient of F is

$$\nabla_{(\beta,b)}F = -2\sum_{i=1}^{n} \mathbf{r}_i^T H \nabla_{(\beta,b)}\mathbf{x}_i.$$

while the gradient terms wrt **b** satisfy the BVP's

$$B_0 \frac{\partial \mathbf{x}}{\partial \mathbf{b}}(0) + B_1 \frac{\partial \mathbf{x}}{\partial \mathbf{b}}(1) = I,$$
$$\frac{d}{dt}\frac{\partial \mathbf{x}}{\partial \mathbf{b}} = \nabla_x \mathbf{f} \frac{\partial \mathbf{x}}{\partial \mathbf{b}}.$$

Embedding: Again the Mattheij example

Consider the modification of the Mattheij problem with parameters $\beta_1^* = \gamma$, and $\beta_2^* = 2$ corresponding to the solution $\mathbf{x}(t, \beta^*) = e^t \mathbf{e}$:

$$A(t) = \begin{bmatrix} 1 - \beta_1 \cos \beta_2 t & 0 & 1 + \beta_1 \sin \beta_2 t \\ 0 & \beta_1 & 0 \\ -1 + \beta_1 \sin \beta_2 t & 0 & 1 + \beta_1 \cos \beta_2 t \end{bmatrix},$$

$$\mathbf{q}(t) = \begin{bmatrix} e^t \left(-1 + \gamma \left(\cos 2t - \sin 2t\right)\right) \\ -(\gamma - 1)e^t \\ e^t \left(1 - \gamma \left(\cos 2t + \sin 2t\right)\right) \end{bmatrix}.$$

In the numerical experiments optimal boundary conditions are set at the first iteration. The aim is to recover estimates of $\beta^*, \mathbf{b}^*$ from simulated data $e^{t_i} H\mathbf{e} + \varepsilon_i$, $\varepsilon_i \sim N(0, .01I)$ using Gauss-Newton, stopping when $\nabla F \mathbf{h} < 10^{-8}$.

Embedding: Again the Mattheij example

$$H = \begin{bmatrix} 1/3 & 1/3 & 1/3 \end{bmatrix}$$

$$H = \begin{bmatrix} .5 & 0 & .5 \\ 0 & 1 & 0 \end{bmatrix}$$

$n = 51,\ \gamma = 10,\ \sigma = .1$
14 iterations
$n = 51,\ \gamma = 20,\ \sigma = .1$
11 iterations
$n = 251,\ \gamma = 10,\ \sigma = .1$
9 iterations
$n = 251,\ \gamma = 20,\ \sigma = .1$
8 iterations

$n = 51,\ \gamma = 10,\ \sigma = .1$
5 iterations
$n = 51,\ \gamma = 20,\ \sigma = .1$
9 iterations
$n = 251,\ \gamma = 10,\ \sigma = .1$
4 iterations
$n = 251,\ \gamma = 20,\ \sigma = .1$
5 iterations

Here $\| \begin{bmatrix} B_1 & B_2 \end{bmatrix}_1 \begin{bmatrix} B_1 & B_2 \end{bmatrix}_k^T - I \|_F < 10^{-3}, k > 1.$

Lagrangian

go OPT2 Associated with the equality constrained problem is the Lagrangian

$$\mathcal{L} = F(\mathbf{x}_c) + \sum_{i=1}^{n-1} \boldsymbol{\lambda}_i^T \mathbf{c}_i .$$

The necessary conditions give:

$$\nabla_{\mathbf{x}_i} \mathcal{L} = 0,\ i = 1, 2, \cdots, n, \quad \mathbf{c}(\mathbf{x}_c) = 0.$$

The Newton equations determining corrections $\mathbf{dx}_c, \mathbf{d}\boldsymbol{\lambda}_c$ are:

$$\nabla^2_{\mathbf{xx}} \mathcal{L} \mathbf{dx}_c + \nabla^2_{\mathbf{x}\boldsymbol{\lambda}} \mathcal{L} \mathbf{d}\boldsymbol{\lambda}_c = -\nabla_{\mathbf{x}} \mathcal{L}^T,$$
$$\nabla_{\mathbf{x}} \mathbf{c}(\mathbf{x}_c)\, \mathbf{dx}_c = C \mathbf{dx}_c = -\mathbf{c}(\mathbf{x}_c),$$

Note sparsity! $\nabla^2_{\mathbf{xx}} \mathcal{L}$ is block diagonal, $\nabla^2_{\mathbf{x}\boldsymbol{\lambda}} \mathcal{L} = C^T$ is block bidiagonal.

SQP formulation

The Newton equations also correspond to necessary conditions for the QP:

$$\min_{\mathbf{dx}} \nabla_{\mathbf{x}} F \mathbf{dx}_c + \frac{1}{2} \mathbf{dx}_c^T M \mathbf{dx}_c; \quad \mathbf{c} + C\mathbf{dx}_c = 0,$$

in case $M = \nabla^2_{\mathbf{xx}} \mathcal{L}$, $\boldsymbol{\lambda}^u = \boldsymbol{\lambda}_c + \mathbf{d}\boldsymbol{\lambda}_c$. A standard approach is to use the constraint equations to eliminate variables. go GNM

$$\mathbf{dx}_i = \mathbf{v}_i + V_i \mathbf{dx}_1 + W_i \mathbf{dx}_n, \quad i = 2, 3, \cdots, n-1.$$

The reduced constraint equation is

$$G\mathbf{dx}_1 + H\mathbf{dx}_n = \mathbf{w}.$$

Is this variable elimination restricted by BVS considerations?

Null space method

Standard SQP approach. Let $C^T = \begin{bmatrix} Q_1 & Q_2 \end{bmatrix} \begin{bmatrix} U \\ 0 \end{bmatrix}$ then

Newton equations can be written

$$\begin{bmatrix} Q^T \nabla^2_{\mathbf{xx}} \mathcal{L} Q & \begin{bmatrix} U \\ 0 \end{bmatrix} \\ \begin{bmatrix} U^T & 0 \end{bmatrix} & 0 \end{bmatrix} \begin{bmatrix} Q^T \mathbf{dx}_c \\ \lambda^u \end{bmatrix} = - \begin{bmatrix} Q^T \nabla_{\mathbf{x}} F^T \\ \mathbf{c} \end{bmatrix}.$$

These can be solved in sequence

$$\begin{aligned} U^T Q_1^T \mathbf{dx}_c &= -\mathbf{c}, \\ Q_2^T \nabla^2_{\mathbf{xx}} \mathcal{L} Q_2 Q_2^T \mathbf{dx}_c &= -Q_2^T \nabla^2_{\mathbf{xx}} \mathcal{L} Q_1 Q_1^T \mathbf{dx}_c - Q_2^T \nabla_{\mathbf{x}} F^T, \\ U \lambda^u &= -Q_1^T \nabla^2_{\mathbf{xx}} \mathcal{L} \mathbf{dx}_c - Q_1^T \nabla_{\mathbf{x}} F^T. \end{aligned}$$

Stability test using Mattheij problem

$Q_1^T \mathbf{dx}_c$ estimates $Q_1^T \text{vec}\left\{e^{t_i}\right\}$ when $\mathbf{x}_c = 0$.

test results $n = 11$

.87665	-.97130	-1.0001
.74089	-1.0987	-1.3432
.47327	-1.2149	-1.6230
.11498	-1.3427	-1.8611
-.32987	-1.4839	-2.0366
-.85368	-1.6400	-2.1250
-1.4428	-1.8125	-2.1018
-2.0773	-2.0031	-1.9444
-2.7309	-2.2137	-1.6330
-3.3719	-2.4466	-1.1526

particular integral $Q_1^T x$

.87660	-.97134	-1.0001
.74083	-1.0988	-1.3432
.47321	-1.2150	-1.6231
.11491	-1.3428	-1.8612
-.32994	-1.4840	-2.0367
-.85376	-1.6401	-2.1250
-1.4429	-1.8125	-2.1019
-2.0774	-2.0032	-1.9444
-2.7310	-2.2138	-1.6331
-3.3720	-2.4467	-1.1527

Conclusion

- Problem stability considerations important but cannot be whole computational story.

Conclusion

- Problem stability considerations important but cannot be whole computational story.
- Embedding makes use of carefully constructed, explicit boundary conditions. Thus expect BVS restrictions must apply.

Conclusion

- Problem stability considerations important but cannot be whole computational story.
- Embedding makes use of carefully constructed, explicit boundary conditions. Thus expect BVS restrictions must apply.
- The variable eliminations form of the simultaneous method partitions variables into sets $\{\mathbf{x}_1,\ \mathbf{x}_n\}$, and $\{\mathbf{x}_2, \cdots, \mathbf{x}_{n-1}\}$ which are found sequentially. It relies implicitly on a form of BVS .

Conclusion

- Problem stability considerations important but cannot be whole computational story.
- Embedding makes use of carefully constructed, explicit boundary conditions. Thus expect BVS restrictions must apply.
- The variable eliminations form of the simultaneous method partitions variables into sets $\{\mathbf{x}_1,\ \mathbf{x}_n\}$, and $\{\mathbf{x}_2, \cdots, \mathbf{x}_{n-1}\}$ which are found sequentially. It relies implicitly on a form of BVS .
- The null space variant partitions the variables into the sets $\{Q_1^T \mathbf{x}_c\}$, $\{Q_2^T \mathbf{x}_c\}$. It appears at least as stable as the variable elimination procedure. Sparsity preserving implementation is straightforward.