

On ODE estimation algorithms

M.R.Osborne *

Contents

1	Introduction	2
2	Methods	3
2.1	Estimation via Embedding	3
2.2	Simultaneous Estimation	5
3	Equivalence	6
4	Consistency	7
5	Convergence rate results	9
6	In Conclusion	9

Abstract

This paper addresses aspects of the problem of estimating the parameters of a system of ordinary differential equations (ODE) given data derived from noisy observations on the state variables. There are two main classes of method for attacking this problem, and their equivalence and effectiveness (consistency) are discussed. Recent rate of convergence results for the major implementation techniques are summarized, and some matters requiring further consideration indicated.

*Mathematical Sciences Institute, Australian National University, ACT 0200, AUSTRALIA. <mailto:Mike.Osborne@maths.anu.edu.au>

1 Introduction

The problem under consideration is that of estimating the vector of parameters $\boldsymbol{\beta} \in R^p$ in the system of ordinary differential equations

$$\frac{d\mathbf{x}}{dt} = \mathbf{f}(t, \mathbf{x}, \boldsymbol{\beta}), \quad (1)$$

where $\mathbf{x}, \mathbf{f} \in R^m$, given observed data:

$$\mathbf{y}_i = \mathcal{O}\mathbf{x}^*(t_i, \boldsymbol{\beta}^*) + \boldsymbol{\varepsilon}_i, \quad i = 1, 2, \dots, n, \quad (2)$$

$$\mathcal{O} \in R^m \rightarrow R^k, \quad \mathbf{y}_i \in R^k, \quad k \leq m,$$

$$\boldsymbol{\varepsilon}_i \in R^m, \quad \sim N(0, \sigma^2 I_k), \quad \text{independent.} \quad (3)$$

It is assumed that $t_i \in [0, 1]$, $i = 1, 2, \dots, n$. The assumption of normal errors is standard enough, but it does have some implications (Section 5). It means that the maximum likelihood estimators of the “true” parameter vector $\boldsymbol{\beta}^*$ and corresponding state variable values $\mathbf{x}(t, \boldsymbol{\beta}^*)$ are found by solving the nonlinear least squares problem

$$\hat{\boldsymbol{\beta}}_n = \arg \min_{\boldsymbol{\beta} \in R^p} \sum_{i=1}^n \|\mathbf{y}_i - \mathcal{O}\mathbf{x}_i\|^2 \quad (4)$$

where the $\mathbf{x}_i = \mathbf{x}(t_i, \boldsymbol{\beta})$, and the values of $\mathbf{x} = \mathbf{x}(t, \boldsymbol{\beta})$ are constrained by the differential equation. Typically a discretized version of the ODE is solved to produce trial values for the optimization problem, and for our purposes it is assumed that the trapezoidal rule produces solution values that are sufficiently accurate. It has the form:

$$\mathbf{c}(\mathbf{x}_c)_i = \mathbf{x}_{i+1} - \mathbf{x}_i - \frac{\Delta t}{2} (\mathbf{f}(t_{i+1}, \mathbf{x}_{i+1}, \boldsymbol{\beta}) + \mathbf{f}(t_i, \mathbf{x}_i, \boldsymbol{\beta})) = 0, \quad (5)$$

where $(\mathbf{x}_c)_i = \mathbf{x}_i$, $i = 1, 2, \dots, n$, $\mathbf{x}_c \in R^{n \times m}$. This recurrence has the important sparsity structure

$$\mathbf{c}(\mathbf{x}_c)_i = \mathbf{c}_{ii}(\mathbf{x}_i) + \mathbf{c}_{i(i+1)}(\mathbf{x}_{i+1}), \quad i = 1, 2, \dots, n-1. \quad (6)$$

Methods for solving the optimization problem (4) fall into two general classes called here *embedding* [6] and *simultaneous* [1], [7]. The embedding method provides a formal link between the problem and the closely related regression problem, but this connection involves some arbitrary choices which can affect performance. However, it does allow the method to connect with an important body of theory which can be used to validate the procedure. The simultaneous class avoids this specification uncertainty and should be

capable of being implemented more efficiently in many cases. However, justification is technically a more difficult problem. These points are illustrated here by considering both the problem of the equivalence of the methods and the asymptotic question of consistency – do the methods produce the true parameter vector in the limit as the number of observations grows without bound? It is pointed out that completely satisfactory answers are not yet available to all questions of interest.

2 Methods

2.1 Estimation via Embedding

The embedding approach leads to an unconstrained optimization problem which can be solved by such standard methods as the Gauss-Newton algorithm. It removes the differential equation constraint on the state variable $\mathbf{x}(t, \boldsymbol{\beta})$ by embedding the differential equation into a parametrised family of boundary value problems which is solved explicitly at each step in order to generate trial solution values. It imposes boundary conditions:

$$B_1 \mathbf{x}(0) + B_2 \mathbf{x}(1) = \mathbf{b},$$

where $B_1, B_2 \in R^m \rightarrow R^m$ are assumed known while \mathbf{b} is a vector of additional parameters which must be determined as part of the estimation process. The key requirement is that the resulting system has a numerically well determined solution $\mathbf{x}(t, \boldsymbol{\beta}, \mathbf{b})$ for all $\boldsymbol{\beta}, \mathbf{b}$ in a large enough neighborhood of $\boldsymbol{\beta}^*, \mathbf{b}^*$ where \mathbf{b}^* is determined by the true state variable values

$$\mathbf{b}^* = B_1 \mathbf{x}^*(0, \boldsymbol{\beta}^*) + B_2 \mathbf{x}^*(1, \boldsymbol{\beta}^*).$$

So far this leaves open the selection of appropriate B_1, B_2 . A hint is provided by the need to calculate $\mathbf{x}(t, \boldsymbol{\beta}, \mathbf{b})$ and $\nabla_{\boldsymbol{\beta}} \mathbf{x}, \nabla_{\mathbf{b}} \mathbf{x}$ at each Gauss-Newton step. Common to these computations is the solution of a sequence of linear problems obtained by linearising the ODE about the current solution estimates. For example, $\nabla_{\mathbf{b}} \mathbf{x}$ satisfies the linear system

$$\begin{aligned} \frac{d}{dt} \nabla_{\mathbf{b}} \mathbf{x} - \nabla_x \mathbf{f} \nabla_{\mathbf{b}} \mathbf{x} &= 0, \\ B_1 \nabla_{\mathbf{b}} \mathbf{x}(0) + B_2 \nabla_{\mathbf{b}} \mathbf{x}(1) &= I. \end{aligned}$$

Computation of the trapezoidal rule approximations (5) to these quantities requires the inversion of the matrix

$$F = \begin{bmatrix} C_{11} & C_{12} & & & & \\ & C_{22} & C_{23} & & & \\ & & & \ddots & & \\ & & & & C_{(n-1)(n-1)} & C_{(n-1)n} \\ B_1 & & & & & B_2 \end{bmatrix}$$

where $C_{ij} = \nabla_x \mathbf{c}_{ij}$.

Idea: Choose B_1, B_2 so this matrix is well conditioned at $\mathbf{x}^*(t, \boldsymbol{\beta}^*)$.

Computation: Begin by permuting the first block column of F to the last position. A transformation of the first $n - 1$ block rows of the permuted matrix to block upper triangular form by orthogonal S yields

$$S^T F P = \begin{bmatrix} R_{11} & R_{12} & 0 & \cdots & 0 & R_{1n} \\ & R_{22} & R_{23} & \cdots & 0 & R_{2n} \\ & & & \cdots & \vdots & \vdots \\ & & & & R_{(n-1)(n-1)} & R_{(n-1)n} \\ & & & & B_2 & B_1 \end{bmatrix}. \quad (7)$$

The factorization affects quantities which depend on the ODE only. Thus the last two block rows determine the conditioning of the transformed matrix. Make a second orthogonal factorization

$$\begin{bmatrix} R_{(n-1)(n-1)} & R_{(n-1)n} \end{bmatrix} = \begin{bmatrix} U^T & 0 \end{bmatrix} \begin{bmatrix} Q_1^T \\ Q_2^T \end{bmatrix}.$$

Then $\begin{bmatrix} B_2 & B_1 \end{bmatrix} = Q_2^T$ provides the desired conditions. The embedding method has the advantages:

- An estimate of the boundary conditions can be computed by the above procedure given suitable initial \mathbf{x}_c^0 .
- It is readily adapted to make use of standard Gauss-Newton nonlinear least squares solvers and ODE boundary value software.
- The availability of good boundary value software is important if the ODE is difficult.

It has the disadvantages:

- Good initial conditions are important. What happens if $1 - \left\| Q_2^T(\mathbf{x}_c^0)^T Q_2(\mathbf{x}_c^*) \right\|$ is close to 1?

- The economics of solving a nonlinear boundary value problem for every function evaluation needs watching.
- The extra parameters \mathbf{b} are not directly relevant to the problem formulation.

2.2 Simultaneous Estimation

Let

$$\mathbf{r}_i = \mathbf{y}_i - \mathcal{O}\mathbf{x}_i,$$

$$\Phi(\mathbf{x}_c) = \frac{1}{2n} \sum_{i=1}^n \|\mathbf{r}_i\|^2.$$

Then the simultaneous method formulates the estimation problem as a constrained nonlinear least squares problem for the $nm + p$ unknowns $\mathbf{x}_c, \boldsymbol{\beta}$:

$$\begin{bmatrix} \widehat{\mathbf{x}}_c \\ \widehat{\boldsymbol{\beta}}_n \end{bmatrix} = \arg \min_{\mathbf{x}_c, \boldsymbol{\beta}} \Phi(\mathbf{x}_c); \quad \mathbf{c}(\mathbf{x}_c, \boldsymbol{\beta}) = 0. \quad (8)$$

Solution of (8) falls within the scope of standard methods of sequential quadratic programming [2]. However, it is important to note that the number of constraints increases as the discretization of the ODE is refined. This provides a context in which it is necessary to exploit the sparsity structure of the problem formulation.

To summarise the solution process, introduce the problem Lagrangian

$$\mathcal{L}(\mathbf{x}_c, \boldsymbol{\beta}, \boldsymbol{\lambda}_c) = \Phi(\mathbf{x}_c) + \sum_{i=1}^{n-1} \boldsymbol{\lambda}_i^T \mathbf{c}_i(\mathbf{x}_c, \boldsymbol{\beta}). \quad (9)$$

The necessary conditions for a stationary point give

$$\nabla_{((x,\beta),\lambda)} \mathcal{L}(\mathbf{x}_c, \boldsymbol{\beta}, \boldsymbol{\lambda}_c) = 0. \quad (10)$$

The corresponding Newton iteration is

$$\begin{bmatrix} \nabla_{(x,\beta)}^2 \mathcal{L} & C^T \\ C & 0 \end{bmatrix} \begin{bmatrix} \Delta \mathbf{x} \\ \Delta \boldsymbol{\beta} \\ \Delta \boldsymbol{\lambda} \end{bmatrix} = - \begin{bmatrix} \nabla_{(x,\beta)} \mathcal{L}^T \\ \mathbf{c}_c \end{bmatrix},$$

where $C = \nabla_{(x,\beta)} \mathbf{c}_c \in R^{nm+p} \rightarrow R^{(n-1)m+p}$. The simultaneous method has the advantages:

- It is completely specified given initial estimates of $\mathbf{x}_c, \boldsymbol{\beta}, \boldsymbol{\lambda}_c$. It can be shown that the choice $\boldsymbol{\lambda}_c = 0$ is typically suitable in large samples.
- Economy - the simultaneous method avoids the inner iterations needed to implement the Newton steps needed to solve nonlinear boundary value problems required in the embedding method.

It has the disadvantages:

- The number of constraints grows without bound as the discretization is refined.
- This means that so does the number of constraint second derivatives that must be computed.
- It is more difficult to formulate solution strategies such as mesh refinement that prove necessary for difficult problems as the state variables are known exactly only at the solution.

3 Equivalence

Superficially the embedding and simultaneous methods look rather different. This is not misleading. The relatively arbitrary component in the embedding method has been noted, while the simultaneous method has a surprising depth of structure. Perhaps the most obvious feature in common is that they address the same problem! However, some progress can be made on the question of equivalence. Let $S_E(\mathbf{x}_E)$ and $S_S(\mathbf{x}_S)$ be the minimum sums of squares of residuals achieved for given n at single isolated points in a large enough ball $B(\mathbf{x}^*, \rho(n))$ by the embedding and simultaneous methods respectively. Then \mathbf{x}_E satisfies the constraints on the simultaneous method so that

$$S_S(\mathbf{x}_S) \leq S_S(\mathbf{x}_E) = S_E(\mathbf{x}_E),$$

but direct substitution gives

$$B_1 \mathbf{x}_S(0) + B_2 \mathbf{x}_S(1) = \mathbf{b}_S.$$

Thus \mathbf{x}_S can generate comparison quantities for the embedding problem. It follows by a similar argument that

$$S_E(\mathbf{x}_E) \leq S_E(\mathbf{x}_S) = S_S(\mathbf{x}_S).$$

Thus $S_E(\mathbf{x}_E) = S_S(\mathbf{x}_S)$. It follows that $\mathbf{x}_E = \mathbf{x}_S$ provided the solutions are isolated and $B(\mathbf{x}^*, \rho(n))$ is large enough.

The problem with this derivation of equivalence is the condition on the size of $\rho(n)$. More satisfactory would be results of the kind:

- The identity of stationary points for the two procedures would follow if satisfaction of necessary conditions for either the embedding or simultaneous methods could be deduced from satisfaction of the other.
- If consistency results of the kind

$$\mathbf{x}_E \xrightarrow[n \rightarrow \infty]{a.s.} \mathbf{x}^*, \quad \mathbf{x}_S \xrightarrow[n \rightarrow \infty]{a.s.} \mathbf{x}^*,$$

obtain then there is the chance of choosing $\rho(n)$ tending to 0 with $n^{-\alpha}$, $\alpha < \frac{1}{2}$. This is not always possible [5]. It would then be possible to deduce identity of the methods for large enough n by the preceding argument. A direct proof of consistency for the simultaneous method appears to be lacking.

4 Consistency

ODE estimation by the embedding method becomes a conventional maximum likelihood estimation problem if the boundary value problems are solved exactly. Here methods for showing consistency of maximum likelihood can be applied to the embedding method also. The argument sketched here follows [3]. It avoids the usual assumption of knowledge of the global maximum of the likelihood function and extends without difficulty to the case in which the likelihood is only evaluated approximately. The maximum likelihood problem has the form

$$\hat{\boldsymbol{\beta}}_n = \arg \max_{\boldsymbol{\beta}} \mathbf{L}_n(\mathbf{y}, \boldsymbol{\beta}) = \arg \max_{\boldsymbol{\beta}} \sum_{i=1}^n L(\mathbf{y}_i, t_i, \boldsymbol{\beta}).$$

where L corresponds to the log of the relevant probability density. Assume the t_i equispaced, then

$$\frac{1}{n} \nabla_{\boldsymbol{\beta}} \mathbf{L}(\mathbf{y}, \boldsymbol{\beta}) \xrightarrow[n \rightarrow \infty]{a.s.} \int_0^1 \mathcal{E}^* \{ \nabla_{\boldsymbol{\beta}} L(\mathbf{y}, t, \boldsymbol{\beta}) \} dt, \quad (11)$$

where the expectation is evaluated using the true density. This gives a limiting form of the necessary conditions and it follows from a standard identity that $\boldsymbol{\beta} = \boldsymbol{\beta}^*$ is a solution. The Kantorovich form of Newton's method can be used to show that $\hat{\boldsymbol{\beta}}_n \xrightarrow[n \rightarrow \infty]{a.s.} \boldsymbol{\beta}^*$. The idea is to apply this to solve the necessary conditions

$$\nabla_{\boldsymbol{\beta}} \mathbf{L}_n(\mathbf{y}, \boldsymbol{\beta}) = 0 \quad (12)$$

starting from β^* . The result (11) can be used to show that the exact limiting solution leads to small residuals in (12). The Kantorovitch result can then be used to show that $\widehat{\beta}_n$ is close to β^* almost surely.

The Kantorovich Theorem required has the following statement. Let $\mathcal{J}_n = \nabla_{\beta\beta}^2 \mathbf{L}$, and $S_\varrho = \{\beta; \|\beta - \beta_0\| < \varrho\}$. If the following conditions are satisfied :

(i) $\|\mathcal{J}_n(\mathbf{u}) - \mathcal{J}_n(\mathbf{v})\| \leq K_1 \|\mathbf{u} - \mathbf{v}\|, \forall \mathbf{u}, \mathbf{v} \in S_\varrho,$

(ii) $\|\mathcal{J}_n(\beta_0)^{-1}\| = K_2,$

(iii) $\left\| \mathcal{J}_n(\beta_0)^{-1} \frac{1}{n} \nabla_x \mathbf{L}_n(\mathbf{y}; \beta_0)^T \right\| = K_3,$ and

(iv) $\xi = K_1 K_2 K_3 < \frac{1}{2},$

then the Newton iteration converges to a point $\widehat{\beta} \in S_\varrho$ satisfying the estimating equation (12), and $\widehat{\beta}$ is the only root in S_ϱ . The step to the solution $\widehat{\beta}$ is bounded by

$$\left\| \widehat{\beta} - \beta_0 \right\|_2 < 2K_3 < \rho.$$

The consistency result for the embedding method which assumes exact integration requires modification to take account of discretization error which causes the objective function to differ from the true likelihood for all finite n . The embedding consistency result can be extended to two important cases:

1. when each differential equation discretization grid \mathbf{K}_n corresponds to the observation grid \mathbf{T}_n ; and
2. when the discretization is made on a fixed grid $t_j \in \mathbf{K}$ independent of $\mathbf{T}_n, n \rightarrow \infty$.

The maximum mesh spacing $\Delta t \rightarrow 0, n \rightarrow \infty$ in the first case so the solution of the discretized problem tends to that of the differential equation at a satisfactory rate. For the trapezoidal rule this is $O(\Delta t^2)$ which is significantly faster than any relevant stochastic rate (typically $O(\Delta t^{1/2})$). In the second case Δt is fixed and finite. This means that truncation error effects persist in the solution of the discretized problem as the size of the observation set $|T_n| \rightarrow \infty$.

Typical results are the following:

1. If $\Delta t \rightarrow 0$ then consistency follows using an argument similar to that in the exact integration case. The idea is to start the iteration for each n at the exact integration solution $\widehat{\beta}_n, \widehat{\mathbf{b}}_n$ and use knowledge of the

discretization error to show that $K_3 = O(\Delta t)^2$ so this start is close to the finite grid solution $\beta_\Delta^n, \mathbf{b}_\Delta^n$. Consistency now follows from the consistency for exact integration

2. If Δt fixed, small enough, then the best result possible is

$$\begin{bmatrix} \beta_\Delta^n \\ \mathbf{b}_\Delta^n \end{bmatrix} \subset S \left(\begin{bmatrix} \beta^* \\ \mathbf{b}^* \end{bmatrix}, O(\Delta t^2) \right), n \rightarrow \infty.$$

It uses $K_3 = O(\Delta t)^2, \forall n = |T_n|$ large enough.

5 Convergence rate results

The Gauss-Newton method for nonlinear least squares minimization is typically the method of choice in the embedding method. It has the key feature that the evaluation of second derivatives is avoided in the approximate Hessian. This has the consequential advantages of strong positive definiteness properties and excellent scale invariance. A key result in this case is that the convergence rate approaches second order asymptotically if the discretization error tends to zero as $|T_n| \rightarrow \infty$ [3]. However, if Δt fixed, small enough, so that discretization error effects persist, then the convergence rate is reduced to a fast first order rate if needed function values are found by linear interpolation.

The Bock iteration is the method of choice in the simultaneous method. Here the Newton iteration is modified by setting the constraint second derivatives to 0 in estimating the augmented matrix. The condition needed for ignoring the constraint curvatures is that the associated Lagrange multipliers be small as $\Delta t \rightarrow 0$. It can be shown that the multipliers are $O(\Delta t)^{1/2}$ and that, as a consequence, the Bock iteration has a similar convergence rate behaviour to Gauss-Newton provided the error terms are normally distributed [4]. This is a stronger condition on the error structure than that required for the convergence rate estimates for the embedding methods. The assumptions needed for the Gauss-Newton results that have been derived for the embedding method require the weaker conditions that the errors are independent and have bounded variance [3].

6 In Conclusion

The two main approaches to the ODE estimation problem have been considered from the point of view of their equivalence and consistency. Consistency

of the embedding method based on trapezoidal rule discretization strategies follows relatively easily from standard results in regression analysis. Thus consistency of the simultaneous method follows from results establishing the equivalence of the two approaches. While a relatively simple argument serves to establish a form of equivalence, it is pointed out that deeper results would have significance. One important step would be an independent proof of consistency for the simultaneous method. Both the important Gauss-Newton and Bock iteration algorithms make use of a strategy of ignoring certain second order partial derivatives. It is of interest that rather different assumptions on measurement error distributions contrast the convergence rate results obtained for the two methods.

References

- [1] H.G. Bock. Recent advances in parameter identification techniques in ODE. In P. Deuffhard and E. Hairer, editors, *Numerical Treatment of Inverse Problems in Differential and Integral Equations*, pages 95–121. Birkhäuser, 1983.
- [2] J. Nocedal and S.J. Wright. *Numerical Optimization*. Springer Verlag, 1999.
- [3] M.R. Osborne. Fisher’s method of scoring. *Int. Stat. Rev.*, 86:271–286, 1992.
- [4] M.R. Osborne. The bock iteration for the ode estimation problem. 2008. in preparation.
- [5] B.J. Quinn and E.J. Hannan. *The Estimation and Tracking of Frequency*. Cambridge University Press, Cambridge, United Kingdom, 2001.
- [6] Childs S.B. and M.R. Osborne. Fitting solutions of ordinary differential equations to observed data. In R.L. May and A.K. Easton, editors, *Computational Techniques and Applications: CTAC95*, pages 193–198. World Scientific Publishing Co., 1996.
- [7] I. Tjoa and L.T. Biegler. Simultaneous solution and optimization strategies for parameter estimation of differential-algebraic systems. *Ind. Eng. Chem. Res.*, 30:376–385, 1991.