Fitting Solutions of Ordinary Differential Equations to Observed Data

S. Bart Childs and Michael R. Osborne

August 9, 1995

Abstract

The method of scoring is developed for the particular applications that arise in estimating systems of ordinary differential equations. Particular attention is paid to parametrising the solution set of the differential system in such a way that the estimation problem is well posed. It is shown that the method of scoring is very appropriate for this application. Numerical results obtained using the code `ps_quasi` are presented.

1 Introduction

Two basic problems considered in this paper are the estimation problem and the smoothing problem. These have in common that they require the fitting of the solutions of a system of ordinary differential equations to data corresponding to a realization of a particular solution trajectory observed in the presence of noise. Despite their potential for application the computational problems associated with these problems have received surprisingly little attention. For example, the definitive modern account of the numerical solution of boundary value problems [1] mentions them only in passing.

Let the differential equation be

$$\dot{\mathbf{x}} = \mathbf{f}(t, \mathbf{x}) \tag{1}$$

where $\mathbf{x} \in R^m$, and $\mathbf{f} \in R \times R^m \rightarrow R^m$. It will be assumed that $\mathbf{f}$ is at least three times continuously differentiable as a function of its arguments. In the estimation problem the system has additional structure expressed by

$$\mathbf{x} = \begin{bmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \end{bmatrix}, \mathbf{f} = \begin{bmatrix} \mathbf{f}_1 \\ 0 \end{bmatrix}, \tag{2}$$

where $\dim(\mathbf{x}_1) = \dim(\mathbf{f}_1) = m - p$. Thus, $\mathbf{x}_2$ can be interpreted as a constant vector corresponding to a set of p parameters to be estimated from the data. The general solution of (1) has the generic form

$$\mathbf{x} = \mathbf{x}(t, \boldsymbol{\alpha}) \tag{3}$$

where $\boldsymbol{\alpha} \in R^p$ parametrizes the solution set of (1). It will be assumed that $\mathbf{x}$ is at least twice continuously differentiable as a function of $\boldsymbol{\alpha} \in A$, the domain of allowable values. The data for the problem is assumed to have the form

$$y_i = \phi_i(\mathbf{x}(t_i, \boldsymbol{\alpha}^*)) + \epsilon_i, \quad i = 1, 2, \ldots, n \tag{4}$$

where the ϕ_i $i = 1, 2, \cdots, n$ are the observation functionals. Here these are specialized to have the form

$$\phi_i(\mathbf{x}) = \phi(t_i, \mathbf{x}(t_i)) \tag{5}$$

where ϕ is some suitably smooth function of its arguments. Frequently, the representer of ϕ is a constant vector $\mathbf{h}$ giving

$$\phi_i(\mathbf{x}) = h^T\mathbf{x}. \tag{6}$$

In addition it is assumed that $\boldsymbol{\alpha}^* \in A$ so there exists a *true* vector of parameters, and that the random variables $\epsilon_i, 1 = 1, 2, \ldots, n$, are independent, $\sim N(0, \sigma^2)$.

Given these assumptions, it is natural to estimate $\boldsymbol{\alpha}$, the vector of free parameters, by a solution of the minimization problem

$$\hat{\boldsymbol{\alpha}}_n = \arg\min_{\boldsymbol{\alpha}} K_n(\boldsymbol{\alpha}) \tag{7}$$

where

$$K_n(\boldsymbol{\alpha}) = \sum_{i=1}^{n} (y_i - \phi_i(\mathbf{x}(t_i, \boldsymbol{\alpha})))^2. \tag{8}$$

This corresponds to an application of the method of maximum likelihood; and this observation permits extension to other situations. For example, $\phi(\mathbf{x})$ could be the mean of a Poisson process and $y_i, i = 1, 2, \ldots, n$

corresponding counts so that $\mathcal{E}\{y_i\} = \phi(t_i, \mathbf{x})$. A discussion of techniques appropriate in the case is given in [4].

A distinguishing feature of this problem is that it is necessary to be careful in specifying $\boldsymbol{\alpha}$. Let us assume that the equation (1) satisfies some criterion of well-posedness such as dichotomy [1]. Then it is necessary that the method chosen for generating $\mathbf{x}(t_i, \boldsymbol{\alpha})$, which in turn depends on the method for imposing the parametrization of the solution set by $\boldsymbol{\alpha}$, be consistent with this criterion in the sense of providing a stable numerical integration procedure. For example, let boundary conditions

$$\begin{aligned} A_1\mathbf{x}(0) &= \boldsymbol{\alpha}_1 \\ A_2\mathbf{x}(1) &= \boldsymbol{\alpha}_2, \end{aligned} \tag{9}$$

be given where

$$\boldsymbol{\alpha} = \begin{pmatrix} \boldsymbol{\alpha}_1 \\ \boldsymbol{\alpha}_2 \end{pmatrix}. \tag{10}$$

Assuming that there exists a dichotomy, then it is sufficient in order that (1), (9) have a stably computable solution that

$$A_1 P(0) = A_1, A_2(I - P(1)) = A_2 \tag{11}$$

where $P(t)$ is the projection onto the decreasing solutions of

$$\dot{\mathbf{u}} = (\nabla_x \mathbf{f})\mathbf{u}. \tag{12}$$

Frequently A_1, A_2 are known *a priori*. For example, if (1) describes the approach to a stable equilibrium then $A_1 = I, A_2 = 0$ is appropriate and the method of simple shooting will do. The stability of partitions of the form of (9) is discussed in Osborne and Russell [5].

In the next section the asymptotic stability and asymptotic rates of convergence for large n are considered for the estimation problem.

Minimization of (7) using the Gauss-Newton algorithm, which corresponds to Fisher's method of scoring [4] in this case, is described in section (3). The excellent numerical properties of the scoring algorithm are stressed.

In section (4) the algorithm is applied to a numerical example derived from chemical reaction kinetics and the resulting data is used to illustrate the asymptotic results. The code used, `ps_quasi`, is described in Appendix A.

Remark 1.1 *The adjoining of the explicit boundary conditions as in (9) is not the only way to specify the comparison solutions* $\mathbf{x}(t, \boldsymbol{\alpha})$*. One other possibility is to consider the differential equation as a constraint on the minimization problem, in which case the additional degrees of freedom are taken up by the Lagrange multipliers. This approach has some interesting properties. For example, it is possible to implement the method in such a way as to exploit parallelism, and the parametrization by Lagrange multipliers appears to have interesting statistical advantages over the method described here (partial results are given in the report [3]. The implementation currently used to minimize (7) is based on Newton's method. It performs adequately, but it makes explicit use of second derivatives of the differential equation. Also, more work needs to be done on finding starting values for the iterations. This is a significant problem because initial values are required for* $\mathbf{x}_i, i = 1, 2, \ldots, n$ *and not just for the parameter vector* $\boldsymbol{\alpha}$ *of dimension* p*.*

2 The nonlinear least squares problem

In this section properties are derived for the minimization problem, (7), under the assumption that the parametrization chosen for the solution set of the differential equation is compatible with a stable numerical integration scheme in the sense of leading to moderately bounded values for $\mathbf{x}(t, \boldsymbol{\alpha}), \forall \boldsymbol{\alpha} \in B(\rho, \boldsymbol{\alpha}^*)$ where B is a ball of radius $\rho > 0$ centered on $\boldsymbol{\alpha}^*$. The necessary conditions for a minimum give

$$-\nabla_\alpha K_n = \sum_{i=1}^{n} (y_i - \phi_i(\mathbf{x})) \nabla_x \phi_i \nabla_\alpha \mathbf{x}(t_i, \boldsymbol{\alpha}) = 0. \tag{13}$$

Behavior for large values of n can be studied under the assumptions that the observations points $t_i^{(n)}, i = 1, 2, \ldots, n$ are chosen such that

$$\frac{1}{n} \sum_{i=1}^{n} g(t_i) \to \int_0^1 g(t) dw(t), n \to \infty, \tag{14}$$

where $w(t)$ is a density describing the limiting manner of selection of the observation points. This is not a serious assumption. Realisation of large data sets requires either a systematic or random mechanism for generating the sample points and the left hand side just becomes a numerical quadrature corresponding to the limiting integral. If this observation is combined with the law of large numbers in the form

$$\lim_{n \to \infty} \frac{1}{n} \left(\textstyle\sum_{i=1}^{n} \epsilon_i \nabla_\alpha \phi_i \;\middle|\; \boldsymbol{\alpha} \right) = 0, \; \forall \boldsymbol{\alpha} \in B(\rho, \boldsymbol{\alpha}^*), \; \rho > 0, \tag{15}$$

in the sense of almost sure convergence, then the limiting equation

$$\int_0^1 (\phi(t, \mathbf{x}(t, \boldsymbol{\alpha}^*)) - \phi(t, \mathbf{x}(t, \boldsymbol{\alpha})) \nabla_x \phi \nabla_\alpha \mathbf{x} dw(t) = 0 \tag{16}$$

is obtained. This equation is satisfied by $\boldsymbol{\alpha} = \boldsymbol{\alpha}^*$. It has an isolated solution provided

$$\frac{1}{n} \nabla_\alpha^2 K(\boldsymbol{\alpha}^*) = \frac{1}{n} \sum_{i=1}^{n} \left[\nabla_\alpha \phi_i^T \nabla_\alpha \phi_i - (y_i - \phi_i) \nabla_\alpha^2 \phi_i \right]$$

is nonsingular. This is well estimated for large n by its 'sample' expectation

$$\mathcal{I}_n = \frac{1}{n} \sum_{i=1}^{n} \nabla_\alpha \mathbf{x}(t_i, \hat{\boldsymbol{\alpha}}_n)^T \nabla_x \phi_i^T \nabla_x \phi_i \nabla_\alpha \mathbf{x}(t_i, \hat{\boldsymbol{\alpha}}_n). \tag{17}$$

This has the limiting form for large n

$$\mathcal{I} = \int_0^1 dw(t) \nabla_\alpha \mathbf{x}^T \nabla_x \phi^T \nabla_x \phi \nabla_\alpha \mathbf{x} \tag{18}$$

which is typically positive definite. Thus the result that $\boldsymbol{\alpha} = \boldsymbol{\alpha}^*$ is an isolated solution is generic in character. A criterion for the minimization problem to be well posed is that $\exists\, k_1, k_2,\ 0 < k_1 < k_2$, k_1 bounded away from zero, k_2 modest in size, such that

$$\begin{array}{rlcl} (i) & \sigma_1(\mathcal{I}) & \geq & k_1, \\ (ii) & \sigma_n(\mathcal{I}) & \leq & k_2. \end{array} \tag{19}$$

The condition on k_1 is a condition on the well posedness of the estimation problem. The condition on k_2 is a condition on the computational technique applied to solve (1) subject to (9) in the sense that a good algorithm and a good choice of parametrization of the solution set lead to small k_2 in the set of possible values. The context is exactly the one in which multiple shooting is favored over simple shooting when there is a nontrivial dichotomy.

Let $\hat{\boldsymbol{\alpha}}_n$ minimize (7). Consistency of this estimate is shown in [4], so that $\hat{\boldsymbol{\alpha}}_n \to \boldsymbol{\alpha}^*$ almost surely as $n \to \infty$. The rate of convergence of the estimate to the true value can be found from the limiting distribution. To calculate this, expand about $\boldsymbol{\alpha}^*$ to obtain

$$\begin{aligned} 0 &= \tfrac{1}{\sqrt{n}} \nabla_\alpha K_n(\hat{\boldsymbol{\alpha}}_n) \\ &= \tfrac{1}{\sqrt{n}} \nabla_\alpha K_n(\boldsymbol{\alpha}^*) + \tfrac{1}{n} \nabla_\alpha^2 K_n \sqrt{n}(\hat{\boldsymbol{\alpha}}_m - \boldsymbol{\alpha}) \\ &\quad + \text{terms smaller in probability} \end{aligned} \tag{20}$$

The central limit theorem applies to show that the first term on the rhs of (20) is asymptotically $\sim N(0, \mathcal{I})$. Inverting (20) to obtain $\sqrt{n}(\hat{\boldsymbol{\alpha}}_m - \boldsymbol{\alpha})$ then gives

$$\begin{aligned} \sqrt{n}(\hat{\boldsymbol{\alpha}}_m - \boldsymbol{\alpha}) &= \left(\tfrac{1}{n} \nabla_\alpha^2 K_n\right)^{-1} \tfrac{1}{\sqrt{n}} \nabla_\alpha K_n(\boldsymbol{\alpha}^*) + \text{ smaller terms} \\ &\sim N(0, \mathcal{I}^{-1}) \text{ asymptotically} \end{aligned} \tag{21}$$

One important consequence of this result is that there is a generic in probability convergence rate of $\mathcal{O}(\frac{1}{\sqrt{n}})$. It follows that there is little point in striving for high accuracy solutions of the differential equations as truncation errors $\mathcal{O}(n^{-2})$ can be expected to be well hidden by the stochastic error terms.

The result (21) can be used to generate confidence regions for the solution and confidence intervals for the individual solution components in the usual way. The sample covariance matrix $\mathcal{I}_n$ would be used, typically.

3 The scoring algorithm

Consider the linear least squares problem obtained by linearizing (7), (8). This linearization about $\boldsymbol{\alpha}_0$ gives

$$\hat{\boldsymbol{\delta}} = \arg\min_{\boldsymbol{\delta}} \sum_{i=1}^{n} (y_i - \phi_i(\mathbf{x}(t_i, \boldsymbol{\alpha}_0)) - \nabla_x \phi_i \nabla_\alpha \mathbf{x}\, \boldsymbol{\delta})^2 \tag{22}$$

It can be written in the form

$$\min_{\boldsymbol{\delta}} r^T r; \quad r = \mathbf{z} - A\boldsymbol{\delta}. \tag{23}$$

where $\mathbf{z}_i = y_i - \phi_i(\mathbf{x}(t_i, \boldsymbol{\alpha}_0))$ and $e_i^T A = \nabla_x \phi_i \nabla_\alpha \mathbf{x}$. Note that $\mathcal{I}_n = \frac{1}{n} A^T A$. Typically this linear subproblem to compute $\boldsymbol{\delta}$ is used in conjunction with a weak line search

$$\hat{\lambda} \approx \arg\min_{\lambda} K_n(\boldsymbol{\alpha} + \lambda \hat{\boldsymbol{\delta}}); \tag{24}$$

and $\boldsymbol{\alpha}$ is updated by

$$\boldsymbol{\alpha} \leftarrow \boldsymbol{\alpha} + \hat{\lambda}\hat{\boldsymbol{\delta}}. \tag{25}$$

In [4] it is shown that asymptotically, for large enough n, $\hat{\lambda} = 1$ will be accepted when $\boldsymbol{\alpha}$ is close to $\boldsymbol{\alpha}^*$. The rate of convergence of the scoring algorithm can be discussed by considering the fixed point iteration

$$\boldsymbol{\alpha} \leftarrow \boldsymbol{\alpha} - (A^T A)^{-1} A^T \mathbf{z} \tag{26}$$

which applies when $\lambda = 1$. Let

$$M = \nabla_\alpha \{\boldsymbol{\alpha} - (A^T A)^{-1} A^T \mathbf{z}\}|_{\boldsymbol{\alpha} = \hat{\boldsymbol{\alpha}}}. \tag{27}$$

Then (26) has $\hat{\boldsymbol{\alpha}}_n$ as a point of attraction provided

$$\varpi\{M\} < 1 \tag{28}$$

where ϖ is the spectral radius of the indicated matrix.

This matrix is given by

$$\begin{aligned} \nabla_\alpha \{\boldsymbol{\alpha} - (A^T A)^{-1} A^T \mathbf{z}\}|_{\boldsymbol{\alpha} = \hat{\boldsymbol{\alpha}}} & \\ &= I - (A^T A)^{-1} \nabla_\beta^2 K_n \\ &= -(A^T A)^{-1} \{\nabla_\beta^2 K_n - A^T A\} \\ &= -\left(\tfrac{1}{n} A^T A\right)^{-1} \left\{\tfrac{1}{n} \textstyle\sum_{i=1}^n \epsilon_i \nabla_\alpha^2 \phi(\boldsymbol{\alpha}^*) + \mathcal{O}(\|\boldsymbol{\alpha}^* - \hat{\boldsymbol{\alpha}}_n\|)\right\} \\ &= -\mathcal{I}_n^{-1} \left\{\tfrac{1}{n} \textstyle\sum_{i=1}^n \epsilon_i \nabla_\alpha^2 \phi_i(\boldsymbol{\alpha}^*) + \mathcal{O}(\|\boldsymbol{\alpha}^* - \hat{\boldsymbol{\alpha}}_n\|)\right\} \\ &\to 0, \quad n \to \infty, \quad \text{almost surely} \end{aligned} \tag{29}$$

We use a) the law of large numbers, b) consistency of the estimate $\hat{\boldsymbol{\alpha}}_n$, and c) that the limiting form of $\mathcal{I}_n$ for $n \to \infty$ is given by (18) and is assumed positive definite with bounded inverse. Then, it follows that

$$\varpi \to 0, \quad n \to \infty, \tag{30}$$

so that the scoring algorithm gives excellent ultimate convergence rates for n large enough.

In [4] it is argued that scoring possesses the virtues of the Newton method, and that it has the advantage of a natural merit function to use in this line search — this important advantage is not enjoyed by the Newton method in general. It has a potential difficulty in that its good features are enjoyed asymptotically. But even this can be used to advantage. What is required is that ϖ be small. If this is not so, then not only is the iteration slow, but the statistics are in doubt. Thus the results are worth having essentially when the algorithm works well. If ϖ corresponds to an isolated eigenvalue of (27), then an estimate of ϖ is given by

$$\varpi = \lim_{i \to \infty} \frac{\|\hat{\boldsymbol{\delta}}_i\|}{\|\hat{\boldsymbol{\delta}}_{i-1}\|}. \tag{31}$$

(This is just an application of the power method for finding the eigenvalue of largest magnitude M.) It will work best when convergence is slow, but returns a small value anyway when the iterates are converging rapidly.

There is an important geometrical characterisation of ϖ as the magnitude of the intrinsic curvature of the likelihood surface. Thus, ϖ is also transformation invariant.

4 An example

The the system of differential equations considered corresponds to the spring mass dashpot

$$\begin{aligned}
\dot{x}_1 &= x_2, && (32)\\
\dot{x}_2 &= -x_1 x_4 - x_2 x_3 + \sin(x_5 t), && (33)\\
\dot{x}_3 &= 0, && (34)\\
\dot{x}_4 &= 0, && (35)\\
\dot{x}_5 &= 0. && (36)
\end{aligned}$$

This system is stable for forward integration so a simple imbedding as an initial value problem is satisfactory. All results use the same estimates of the initial values. The true and "*estimated*" values are:

	$x_1(0)$	$x_2(0)$	$x_3(0)$	$x_4(0)$	$x_5(0)$
true	-3.00	-8.00	0.20	1.00	1.00
estimate	-4.00	-6.00	0.10	0.80	1.20
% error	33.3	-25	-50	-20	20

The cases run correspond to observed data with both normal and uniform error contamination, and with standard deviations of .1 and 1. for both distributions. Observations are made on the first component. The nuber of iterations required by the scoring algorithm as a function of n are as follows:

case	σ	25	50	100	200	400
N	.1	9	8	8	7	7
N	1.	*	14	12	9	9
U	.1	11	7	7	7	7
U	1.	25	9	8	8	9

The scoring iteration was run with a simple convergence test, and a tolerance of 10^{-6}. The results show the expected main features, including the predicted linear convergence with multiplier decreasing with n at a rate matching the doubling increase. This illustrates that the results are not dependent on the assumed form for the distribution of the errors. The iteration performs well, especially for the larger values of n.

5 References

References

[1] U.Ascher, R.Mattheij, and R.D.Russell. **The Numerical Solution of Boundary Value Problems**, Birkhauser, 1988.

[2] P.Erdi and J.Toth. **Mathematical Models of Chemical Reactions**, Princeton, 1989.

[3] M.R.Osborne. *Scoring with constraints*, Technical Report CS94-070, TAMU, 1994.

[4] M.R.Osborne. *Fisher's method of scoring*, Int. Stat. Rev., 60, p.99–117, 1992.

[5] M.R.Osborne and R.D.Russell. *Stability of BVP's, SINUM, 1987*