ORNL/TM--11688

DE91 007872

Engineering Physics and Mathematics Division

BAYESIAN APPROXIMATION OF SOLUTIONS TO LINEAR ORDINARY DIFFERENTIAL EQUATIONS

Karole Herzog[1]
Max D. Morris
Toby J. Mitchell
Mathematical Sciences Section
Engineering Physics and Mathematics Division
Oak Ridge National Laboratory
P.O. Box 2009, Bldg. 9207A
Oak Ridge, Tennessee 37831-8083

[1]K. Herzog's current address is: Dept. of Statistics, VPI & State Univ., Blacksburg, Virginia 24061-0439.

Date Published: November 1990

Research sponsored by the Applied Mathematical Sciences Research Program, Office of Energy Research, U. S. Department of Energy under contract DE-AC05-84OR21400 with the Martin Marietta Energy Systems, Inc., and by the U. S. Department of Energy through the Oak Ridge Science and Engineering Research Semester Program through Oak Ridge Associated Universities.

Prepared by the
OAK RIDGE NATIONAL LABORATORY
Oak Ridge, Tennessee 37831
Managed By
MARTIN MARIETTA ENERGY SYSTEMS, INC.
for the
U.S. DEPARTMENT OF ENERGY
under Contract No. DE-AC05-84OR21400

MASTER

DISTRIBUTION OF THIS DOCUMENT IS UNLIMITED

TABLE OF CONTENTS

1. INTRODUCTION 1

2. APPROXIMATING THE SOLUTION OF A LINEAR DIFFERENTIAL EQUATION . 3

2.1 The Approximation 3
2.2 Example 1 7
2.3 Selecting Enforcement Sites 8
2.4 Example 1 Revisited 9
2.5 Choice of Correlation Function; Estimating θ 9

3. EXAMPLES 12

3.1 Example 2 13
3.2 Example 3 16
3.3 Example 4 20

4. COMPUTATION 20

5. DISCUSSION 21

5.1 Design 22
5.2 Choice of prior process 22
5.3 Computation 22
5.4 Extensions 23

REFERENCES 24

LIST OF FIGURES

Figure 1. Plot of $\tau_m(t) = \hat{h}_m(t) - h(t)$ and $\varepsilon_m(t) = y(t) - \hat{y}_m(t)$, $m = 5, 6, 10$, for Example 1 (revisited), where the approximation $\hat{y}$ is based on the nonnegative cubic correlation with $\theta = 20$. Points that correspond to the m enforcement sites are marked on the curves by dots.

Figure 2. Plot of the maximum absolute error in $\hat{y}$, as a function of the number of enforcement sites m, for Example 1 (revisited), where $\hat{y}$ is based on the nonnegative cubic correlation with $\theta = 20$.

Figure 3. Plots of true solution y and approximate solutions $\hat{y}_m = 10, 11, 12, 13$, for Example 2, where the approximation is based on the Gaussian correlation with $\theta = 0.42$ for $m = 10, 11$, and $\theta = 0.25$ for $m = 12, 13$. Points that correspond to the m enforcement sites are marked on $\hat{y}$ by dots.

Figure 4. Plot of $\tau_m(t) = \hat{h}_m(t) - h(t)$ and $\varepsilon_m(t) = y(t) - \hat{y}_m(t)$, $m = 10, 11, 13$ for Example 2, where the approximation $\hat{y}$ is based on the Gaussian correlation with $\theta = 0.42$. Points that correspond to the m enforcement sites are marked on the curves by dots.

Figure 5. Plot of $\tau_m(t) = \hat{h}_m(t) - h(t)$ and $\varepsilon_m(t) = y(t) - \hat{y}_m(t)$, $m = 10, 11, 17$ for Example 2, where the approximation $\hat{y}$ is based on the nonnegative cubic correlation with $\theta = 13$. Points that correspond to the m enforcement sites are marked on the curves by dots.

Figure 6. Plot of the maximum absolute error in $\hat{y}$, as a function of the number of enforcement sites m, for Example 2, where $\hat{y}$ is based on the nonnegative cubic correlation with $\theta = 13$.

Figure 7. Plot of the solution function y for the second order differential equation of Example 3 (initial value problem) and Example 4 (two point boundary value problem). Points that correspond to the 5 enforcement sites are marked by dots. The approximation $\hat{y}$, based on the Gaussian correlation with $\theta = 9$ is virtually the same as y in both examples.

BAYESIAN APPROXIMATION OF SOLUTIONS TO LINEAR ORDINARY DIFFERENTIAL EQUATIONS

Karole Herzog, Max Morris, and Toby Mitchell

ABSTRACT

An approach to numerically solving linear ordinary differential equations, based on statistical Bayesian prediction, is described. Preliminary results on the details of choice of correlation parameters and experimental design are given, using first- and second-order example problems.

Key Words: Bayesian prediction, collocation methods, correlation function, experimental design, interpolation, spline, stochastic process.

1. INTRODUCTION

We describe here a general method for approximating the solution of a linear ordinary differential equation of form

$$\sum_{j=0}^{n} \alpha_j(t) y^{(j)}(t) = f(t), \tag{1}$$

for all t in a specified interval T, where necessary side conditions are supplied. Here we shall consider only first and second-order equations ($n \leq 2$), but the approach is easily generalized to higher orders. There exists a large collection of methods that are applicable to problems of the form (1); see Isaacson and Keller (1966, Chapter 8) and Gear (1971) for a review of many of these. In contrast to previous approaches, the method we present here is based on a Bayesian statistical formulation, under which uncertainty about the solution function y is expressed by means of the random function (equivalently, stochastic process) Y. This is a direct generalization of the Bayesian use of a random variable to represent uncertainty about a scalar quantity. The *data* on which we shall base the estimate of y consists of the side conditions and a finite number of *enforcements* of the differential equation at selected values of t. The conditional stochastic process given these data, or Bayesian *posterior* process, provides a predictive distribution for y at each value of t in T, and we take the mean of this process as the approximation $\hat{y}$ of the function y.

It should be noted that this is a report of preliminary research results. If this methodology is to eventually become part of the practical body of numerical tools for solving differential equations, it will need to be made computationally competitive with methods presently used. However, our work to this point has concentrated simply on the feasibility of the approach, rather than on its competitiveness. Thus, for example, we have not yet compared the computational effort required with the requirements of methods currently in use.

This kind of approach was used in a different context by Currin, Mitchell, Morris, and Ylvisaker (1988) and Sacks, Welch, Mitchell, and Wynn (1989) for predicting an unknown function y on domain T, given y(t) at a set of *design sites* $D = \{t_i, i=1, ..., n\}$. (We shall often refer to particular values of t as sites, following a common practice in geological applications of similar methods.) There, prior knowledge of y was represented by a Gaussian stochastic process Y, usually taken to be stationary to express prior impartiality. Thus $E[Y(t)] = \mu$ and $Var[Y(t)] = \sigma^2$, where μ and σ^2

are the same for all t in T, and, for $s, t \in T$, $\mathrm{Corr}[Y(t), Y(s)] = R(s-t)$, a function only of the distance between s and t. Following observation of the values of y at each site in D, the posterior process is also Gaussian and Y(t) has posterior mean and variance:

$$\hat{y}(t) = E[Y(t) \mid \vec{y}_D] = \mu + \vec{\sigma}_{Dt}^{T} \sigma_{DD}^{-1} (\vec{y}_D - \vec{\mu}_D) \tag{2}$$

$$\sigma^2(t) = \mathrm{Var}[Y(t) \mid \vec{y}_D] = \sigma^2 - \vec{\sigma}_{Dt}^{T} \sigma_{DD}^{-1} \vec{\sigma}_{Dt} \,. \tag{3}$$

Here, $\vec{\sigma}_{Dt}$ is the vector of prior covariances between Y at the site t and the design sites D, and σ_{DD} is the prior covariance matrix for Y at the design sites. The elements of $\vec{y}_D$ are the observed values of y on D, and the elements of $\vec{\mu}_D$ are the prior expected values of Y(t) (all equal to μ) for each $t \in D$. It can be shown that $\hat{y}$ satisfies the given data exactly; it is therefore an interpolating function.

In order to implement equations (2) and (3), one needs to specify μ, σ, and the correlation function R. Two examples of R used by Currin et al. are the Gaussian correlation function (not to be confused with a Gaussian stochastic process),

$$R(d) = e^{-d^2/\theta} \tag{4}$$

and the nonnegative cubic correlation function,

$$\begin{aligned} R(d) &= 1 - 6\left(\frac{d}{\theta}\right)^2 + 6\left(\frac{|d|}{\theta}\right)^3 && |d| < \frac{\theta}{2} \\ R(d) &= 2\left(1 - \frac{|d|}{\theta}\right)^3 && \frac{\theta}{2} \le |d| < \theta \\ R(d) &= 0 && |d| \ge \theta. \end{aligned} \tag{5}$$

where $d = s-t$, $s, t \in T$, and $0 < \theta < \infty$.

Since

$$\vec{\sigma}_{Dt}^{T} = \sigma^2 \times [\, R(t - t_1), R(t - t_2), \cdots, R(t - t_n) \,] \,,$$

the posterior mean function $\hat{y}$ is seen to be a constant (μ) plus a linear combination of the n basis functions $R_i = R(t - t_i)$. Thus, if R is given by (4), $\hat{y}$ is analytic (because R is), and if R is given by (5), $\hat{y}$ is a cubic spline (because R is). We shall use these correlation functions in our examples below, where the application is to the approximation of the solution of linear differential equations.

2. APPROXIMATING THE SOLUTION OF A LINEAR DIFFERENTIAL EQUATION

From a statistical point of view, the primary difference between the interpolation problem described above and the problem of approximating the solution of a differential equation is the nature of what may be called data. We shall follow essentially the approach outlined above for interpolation, modified to use the data available in a differential equation.

2.1 The Approximation

In order to admit a unique solution, an nth-order linear ordinary differential equation (1) is generally accompanied by n side conditions that must be satisfied in addition to the differential equation. In this discussion, we will use conditions of the form:

$$y^{(k(i))}(t_i) = c_i\,, \quad i = 1, 2, \cdots n\,, \tag{6}$$

where the ith condition specifies the value c_i of the k(i)th derivative at site t_i, $k(i) \le n$. We denote by $B = \{t_i, i=1, ..., n\}$ the list of sites corresponding to the set of side conditions. These are not necessarily distinct; for example, an *initial value problem* specifies all side conditions at the same site.

Other information about y is available in the differential equation (1) itself. Of course, if we could find a function that satisfied the side conditions as well as the differential equation at all sites in T, it would be the exact solution, y. Unfortunately, this is generally not feasible. Instead, we shall use as data the relationship specified by (1) at only a finite set of m *enforcement sites* $D = \{t_i, i = n+1, \cdots, n+m\}$:

$$\sum_{j=0}^{n} \alpha_j(t_i) y^{(j)}(t_i) = f(t_i)\,, \quad t_i \in D\,. \tag{7}$$

We regard enforcement sites as design sites, because we are free to select them. Although some sites in B and D may coincide, it is convenient for notation and bookkeeping to treat B and D as distinct lists. We shall call their concatenation C.

We represent uncertainty about the solution function y by the nth-order differentiable Gaussian stochastic process Y. Let

$$\vec{Z} = \begin{bmatrix} Y^{(k(1))}(t_1) \\ Y^{(k(2))}(t_2) \\ \cdots \\ Y^{(k(n))}(t_n) \\ F(t_{n+1}) \\ F(t_{n+2}) \\ \cdots \\ F(t_{n+m}) \end{bmatrix} \tag{8}$$

where

$$\sum_{j=0}^{n} \alpha_j(t) Y^{(j)}(t) = F(t) \ . \tag{9}$$

The elements of $\vec{Z}$ are all linear functions of Y and its first n derivatives at the sites of side conditions and enforcement, i.e.,

$$\vec{Z} = A\vec{Y}^*, \tag{10}$$

where $\vec{Y}^*$ is the [(n+m)(n+1)]-vector:

$$\vec{Y}^* = \begin{bmatrix} Y^{(0)}(t_1) \\ Y^{(1)}(t_1) \\ \cdots \\ Y^{(n)}(t_1) \\ \text{-----} \\ \cdots \\ \cdots \\ \text{-----} \\ Y^{(0)}(t_{n+m}) \\ Y^{(1)}(t_{n+m}) \\ \cdots \\ Y^{(n)}(t_{n+m}) \end{bmatrix} \tag{11}$$

and A is an (n+m)×(n+m)(n+1) matrix of known constants. (In practice, any columns of A which contain only zeros could be deleted from the matrix, but for notational simplicity we will not do so here. We demonstrate the formation of A in examples to be given later.)

Then the knowledge provided by the data is that

$$\vec{Z} = \vec{z} \tag{12}$$

where

$$\vec{z} = \begin{bmatrix} c_1 \\ c_2 \\ \cdots \\ c_n \\ \\ f(t_{n+1}) \\ \\ f(t_{n+2}) \\ \\ \cdots \\ f(t_{n+m}) \end{bmatrix} . \tag{13}$$

Since the process Y is Gaussian and nth-order differentiable, the joint prior distribution of $\vec{Z}$ and Y(t) at any site is multinormal. The posterior distribution of Y(t), given (12), is therefore also multinormal (see, for example, Morrison (1967)), and its mean and variance are given by:

$$\hat{y}(t) = E[Y(t) \mid \vec{z}] = \mu + \vec{\sigma}_{Ct}^{T}\, \sigma_{CC}^{-1}(\vec{z} - \vec{\mu}_z) \tag{14}$$

$$\sigma^2(t) = Var[Y(t) \mid \vec{z}] = \sigma^2 - \vec{\sigma}_{Ct}^{T}\sigma_{CC}^{-1}\vec{\sigma}_{Ct} . \tag{15}$$

Here

$$\vec{\sigma}_{Ct} = Cov[\vec{Z}, Y(t)] = Cov[A\vec{Y}^*, Y(t)] = ACov[\vec{Y}^*, Y(t)], \tag{16}$$

where $Cov[\vec{Y}^*, Y(t)]$ is the vector of covariances between $\vec{Y}^*$ and Y(t);

$$\sigma_{CC} = Cov[\vec{Z}, \vec{Z}] = Cov[A\vec{Y}^*, A\vec{Y}^*] = ACov[\vec{Y}^*, \vec{Y}^*]A^T, \tag{17}$$

where $Cov[\vec{Y}^*, \vec{Y}^*]$ is the matrix whose (i, j)th element is the covariance between the ith and jth elements of $\vec{Y}^*$;

$$\vec{\mu}_z = E[\vec{Z}] = E[A\vec{Y}^*] = AE[\vec{Y}^*]. \tag{18}$$

The posterior process depends, therefore, on certain covariances and expectations involving $\vec{Y}^*$, which are immediate once covariances and expectations involving the derivatives of Y are available. It can be shown that, for stationary Gaussian processes with R suitably differentiable,

$$E[Y^{(j)}(t)] = 0\,, \quad j > 0\,, \tag{19}$$

and

$$\mathrm{Cov}[Y^{(j)}(t)\,,\, Y^{(k)}(s)] = \sigma^2(-1)^j R^{(j+k)}(s-t). \tag{20}$$

From (11), (18), and (19),

$$\vec{\mu}_z = \mu A\vec{e}, \tag{21}$$

where $\vec{e}$ is the [(n+m)(n+1)]-vector with "1" in positions j(n+1)+1, j=0, 1, 2, $\cdots$, n+m−1, and "0" everywhere else, i.e., $\vec{e}$ has "1" in the same positions that Y has the superscript "(0)" in (11).

We shall use $\hat{y}$ in (14) as our approximation to the solution of the differential equation (1) subject to the specified side conditions (6). It can be shown that $\hat{y}$ necessarily satisfies (6) and (7), i.e., our approximation satisfies the side conditions and also satisfies the differential equation at the enforcement sites. Our method is therefore a type of *collocation* method. (See, e.g., Prenter 1975, Chapter 8.) We shall not be concerned here with the posterior variances given by (15); effective use of these calculations for design and inference requires a more intensive investigation.

At present, we automatically choose μ by the method of maximum likelihood. The probability density function of $\vec{Z}$ is

$$\phi(\vec{z}) = \frac{1}{(2\pi)^{(n+m)/2}\det(\sigma_{CC})^{1/2}}\exp[-\tfrac{1}{2}(\vec{z}-\vec{\mu}_z)^T\sigma_{CC}^{-1}(\vec{z}-\vec{\mu}_z)].$$

The likelihood function is the prior probability density function for $\vec{Z}$, evaluated at the value $\vec{z}$ given by (13), and regarded as a function of the parameters of the stochastic process. It is convenient to work with the natural log of the likelihood:

$$L = -\tfrac{1}{2}[(n+m)\ln(2\pi) + \ln\det(\sigma_{CC}) + (\vec{z}-\vec{\mu}_z)^T\sigma_{CC}^{-1}(\vec{z}-\vec{\mu}_z)]. \tag{22}$$

For fixed θ and σ the maximum with respect to μ occurs at the unique stationary point, which can be easily found by differentiating L with respect to μ and equating the resulting expression to 0. One obtains:

$$\hat{\mu} = \frac{\vec{w}^T \sigma_{CC}^{-1} \vec{z}}{\vec{w}^T \sigma_{CC}^{-1} \vec{w}}, \tag{23}$$

where $\vec{w} = A\vec{e}$.

With (23) substituted for μ in (14), $\hat{y}$ depends only on the choice of correlation function R, including its parameter(s) θ. (The prior variance σ^2 does not affect $\hat{y}$, since it appears only as a constant factor in $\vec{\sigma}_{Ct}$ and σ_{CC}, and effectively "cancels out".) The choice of correlation function is still very much an open issue. In our examples so far, we have considered only the Gaussian (4) and nonnegative cubic (5) correlations, and have experimented with various choices of θ in each case.

2.2 Example 1

Consider the simple differential equation:

$$y(t) - y'(t) = 0 \qquad 0 \le t \le 1 \qquad y(0) = 1 \ . \tag{24}$$

The solution is $y(t) = e^t$, which we shall use to compare with our approximation.

There is one side condition site $B = \{t_1\} = \{0\}$. Suppose we choose the m=6 enforcement sites $D = \{t_2, \cdots, t_7\} = \{0, 0.12, 0.35, 0.61, 0.86, 1.00\}$. Then

$$A = \begin{bmatrix} 1 & 0 & 0 & 0 & \cdot\ \cdot & 0 & 0 \\ 0 & 0 & \alpha_0(t_2) & \alpha_1(t_2) & \cdot\ \cdot & 0 & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot\ \cdot & \cdot & \cdot \\ 0 & 0 & 0 & 0 & \cdot\ \cdot & \alpha_0(t_7) & \alpha_1(t_7) \end{bmatrix} \tag{25}$$

where $\alpha_0(t) = 1$ and $\alpha_1(t) = -1$, and

$$\vec{z} = \begin{bmatrix} c_1 \\ f(t_2) \\ f(t_3) \\ \cdot \\ \cdot \\ f(t_7) \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 0 \\ \cdot \\ \cdot \\ 0 \end{bmatrix}. \tag{26}$$

To obtain the covariances needed to compute $\vec{\sigma}_{Ct}$ and σ_{CC} through (16), (17), and (20), we use the Gaussian correlation function (4) with $\theta = 1.0$. The approximation (14) works very well in this case: the maximum absolute error, $|\varepsilon(t)| = |y(t) - \hat{y}(t)|$, evaluated at 101 equally spaced sites covering T, is 0.00021. We shall visit this example again later, using the nonnegative cubic correlation.

2.3 Selecting Enforcement Sites

In an attempt to develop a formal method for choosing the enforcement sites, we first adopted as a design criterion the maximization of det (σ_{CC}). This was contrived by analogy with the approach of Currin, et al. (1988) to the construction of designs for prediction of y on T given knowledge of y at the design sites. There the design criterion for prediction based on (2) and (3) was maximization of det (σ_{DD}). Although the analogy is weak (because the given knowledge here is of a different type), this criterion is useful for numerical reasons. Since σ_{CC} is the matrix of the linear system that must be solved in (14) and (23), the choice of design heavily influences the extent of numerical difficulty encountered in the prediction. We have found that designs which maximize the determinant of this matrix often lead to relatively fewer numerical difficulties. In the few cases we have considered, the "optimal" designs found by our computer algorithm place the design sites at roughly equal intervals. Until more rigorous approaches to design construction are developed, this kind of design seems a reasonable choice.

Alternatively, we tried an adaptive approach, which starts with an equispaced design and then adds new enforcement sites one at a time, each one at the value of t where the absolute value of the function

$$\tau(t) = \hat{f}(t) - f(t) = \sum_{j=0}^{n} \alpha_j(t)\hat{y}^{(i)}(t) - f(t). \tag{27}$$

is currently largest. (We typically restrict this search to the finite set $\dot{T}_{101}$, which consists of 101 equally spaced sites covering T.) This strategy is based on the hope that enforcement of the differential equation at the value of t where it currently fails most badly will improve the approximation to y. After the addition of each new site, $\hat{y}$ and τ are updated and the process continues; as already noted, τ becomes 0 at each new enforcement site. Ideally, we want τ to be 0 everywhere in T, for then $\hat{y}$ would satisfy (1). However, we have not developed any convergence theory. Our "stopping rules" are quite vague: we stop when $|\tau|$ seems satisfactorily small. We are sometimes forced to stop sooner (or to pause and alter θ), because σ_{CC} becomes increasingly ill-conditioned for fixed θ as we add enforcement sites. Nevertheless, we have had good results with this adaptive method so far, and we shall use it in the remaining examples of this report.

2.4 Example 1 Revisited

We consider again the first-order differential equation (24), this time with the five equally spaced enforcement sites $D = \{0, 0.25, 0.50, 0.75, 1.00\}$, and the nonnegative cubic correlation (5) with $\theta = 20$. The maximum $|\tau(t)|$ on $\dot{T}_{101}$ is 0.02675; this is attained at t=0.12, so the sixth enforcement site is placed there. After updating $\hat{y}$ and τ, we find the maximum $|\tau(t)|$ to be 0.01865 at $t = 0.37$, and the seventh enforcement site is added there. This process is continued, and three further enforcement sites are added at 0.62, 0.06, and 0.18, in that order. The functions $\tau_m(t)$ and $\varepsilon_m(t) = y(t) - \hat{y}_m(t)$ are shown in Figure 1 for m=5, 6, and 10, where the subscript m refers to the number of enforcement sites used in the approximation. Figure 2 shows the maximum value of the absolute error $|\varepsilon_m(t)|$ on $\dot{T}_{101}$ as a function of m. The final maximum absolute error of approximation is 0.00385, which we think is respectable, although it is still not as low as that obtained above using the Gaussian correlation function with six enforcement sites.

2.5 Choice of Correlation Function; Estimating θ

As in the prediction context (Section 1), the nonnegative cubic correlation function leads automatically to approximations that are cubic splines and the Gaussian correlation function leads to approximations that are analytic. This can be seen by noting that $\hat{y}$ in (14) depends on t through the elements of $\vec{\sigma}_{Ct}$; examination of (20) shows that these elements have the same form as the correlation function R and its derivatives.

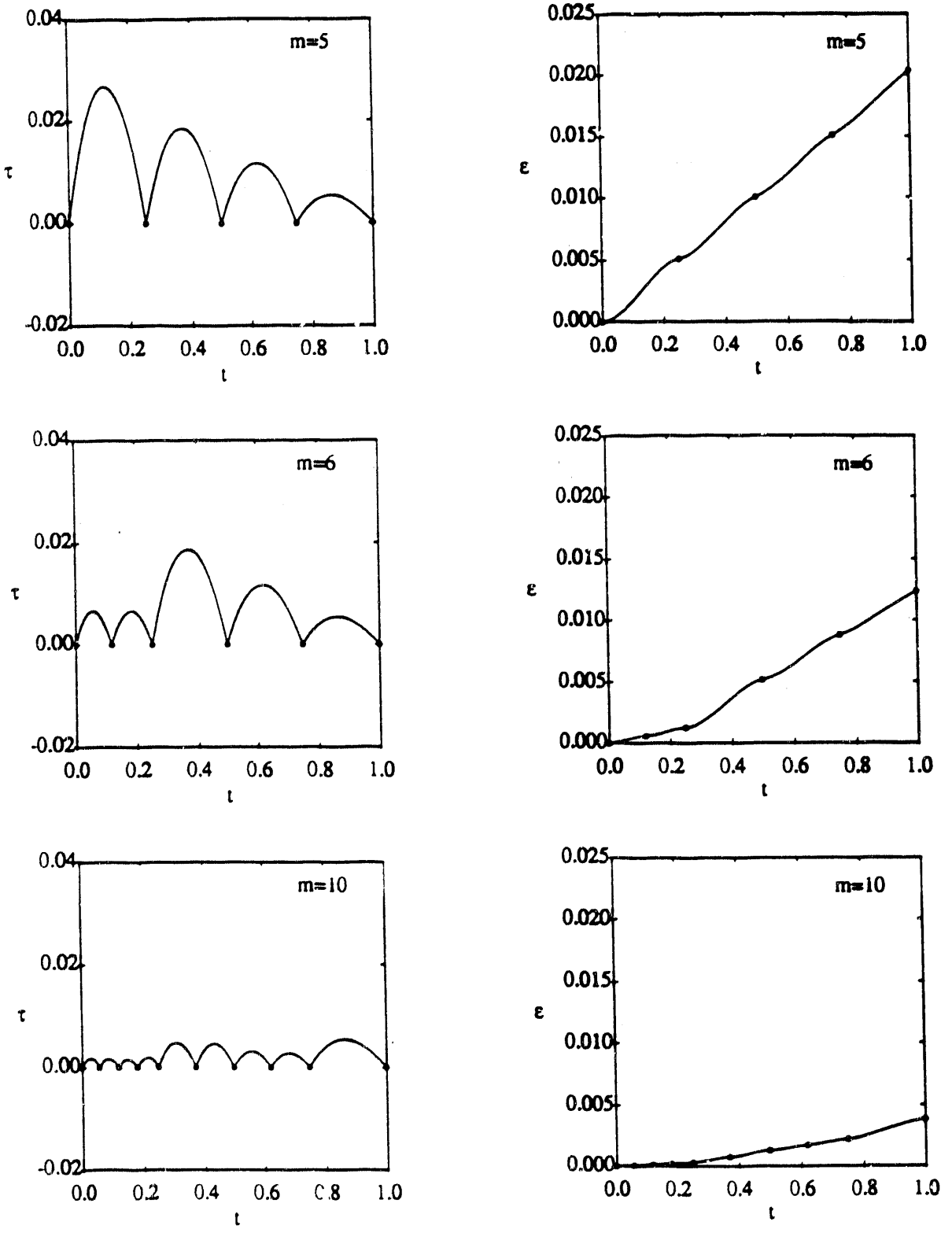

Figure 1. Plot of $\tau_m(t) = \hat{h}_m(t) - h(t)$ and $\varepsilon_m(t) = y(t) - \hat{y}_m(t)$, $m = 5, 6, 10$, for Example 1 (revisited), where the approximation $\hat{y}$ is based on the nonnegative cubic correlation with $\theta = 20$. Points that correspond to the m enforcement sites are marked on the curves by dots.

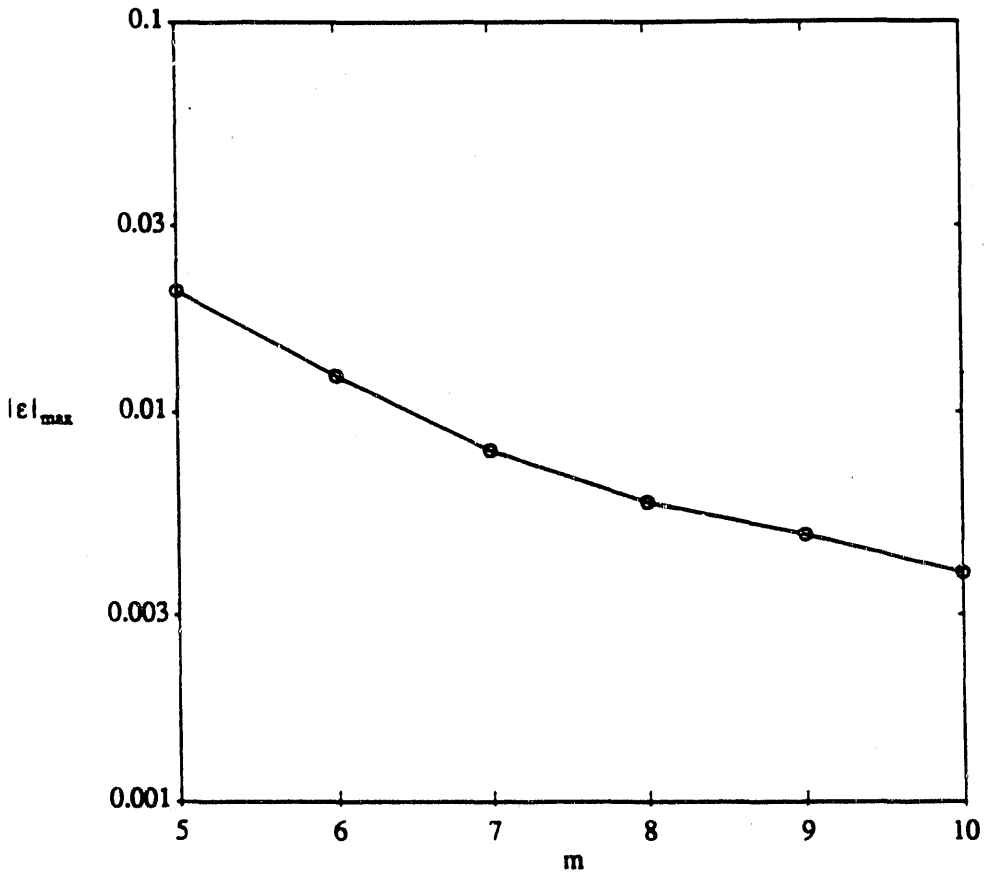

Figure 2. Plot of the maximum absolute error in $\hat{y}$, as a function of the number of enforcement sites m, for Example 1 (revisited), where $\hat{y}$ is based on the nonnegative cubic correlation with $\theta = 20$.

Our limited experience suggests that the Gaussian correlation can work very well indeed, as in Example 1, but its performance seems quite sensitive to the design D and to θ. The nonnegative cubic correlation seems more forgiving, and is better suited to the sequential addition of enforcement sites for fixed θ. A disadvantage is that it can only be used for first-order equations; this can be seen by noting that R must be differentiable through order 2n in order to define all the covariances required (see (20)).

We have tried two approaches to the choice of θ for a given correlation function, one based on maximizing the likelihood (22) and the other based on minimizing the size of the function τ (27). For fixed θ it is easy to obtain maximum likelihood estimates for μ and σ, but after substituting these values into (22), maximization with respect to θ is not so direct. Although the likelihood function seemed well behaved in all the cases we considered, the search needs to be handled delicately. As θ increases, the determinant of σ_{CC} tends to become very small, so the computation of (22) becomes unreliable. Unfortunately, maximization of L frequently forces the search in this direction. After making some superficial efforts to overcome this difficulty, we abandoned the attempt to choose θ via maximum likelihood in favor of an alternative criterion based on τ. We observed that the absolute value of this function tended to be relatively small when the approximation was good and relatively large when the approximation was bad. As a result, we tentatively adopted a criterion which chooses θ to minimize the sum of squared values of $\tau(t)$ on $\dot{T}_{101}$. Although we have not investigated the theoretical properties of this or similar criteria, it has not so far presented the numerical difficulties we encountered with maximum likelihood, largely because it does not seem to favor such high values of θ. We used this criterion for choosing the initial values of θ in the Examples 3 and 4 below, where it worked well, and we think it might prove helpful in automating parameter estimation after further development. Computation of τ for many values of θ should be avoided if possible, however, because of computational expense (Section 4).

3. EXAMPLES

We applied the method described above to three additional examples where the solution is known. The problems are of the textbook variety, but serve as generic examples of several kinds of differential equations that are of much interest. The first is a first-order equation that is considered *stiff* because its solution contains a rapidly decaying component. The second and

third examples are based on a second-order differential equation. They differ from each other only by the nature of their side conditions, the first of them being an initial value problem and the second being a two-point boundary value problem.

3.1 Example 2

Problem:

$$50y(t) + y'(t) = 50t^2 + 2t \qquad 0 \le t \le 1 \qquad y(0) = \frac{1}{3}. \tag{28}$$

The solution is

$$y(t) = \frac{e^{-50t}}{3} + t^2. \tag{29}$$

Our initial design consists of the 10 enforcement sites D = {0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1.0}. The form of A and of $\vec{z}$ are similar to (25) and (26), but A has 11 rows, $\alpha_0(t) = 50$, $\alpha_1(t) = 1$, $c_1 = 1/3$, and $f(t) = 50t^2 + 2t$. Our initial approximation, which is shown in Figure 3a, is based on the Gaussian correlation function with $\theta = 0.42$. (This and other values of θ were chosen by trial and error so as to give reasonably low values of $|\tau|$ without encountering numerical difficulties.) We find the maximum $|\tau(t)|$ to be 8.96, at $t = 0$, and add a new enforcement site there. The resulting approximation is shown in Figure 3b. We then find the maximum $|\tau(t)|$ to be 3.17, at $t = 0.03$. After adding a new enforcement site there, we encounter numerical problems, which are alleviated by decreasing θ to 0.25. The maximum $|\tau(t)|$ surprisingly shoots up to 20.6, at $t = 0.97$. The approximation, although good for $t < 0.75$, becomes much worse near the right end of the interval. (See Figure 3c.) The addition of a thirteenth enforcement site, at $t = 0.97$, improves matters greatly. Figure 3d shows $\hat{y}$ for the resulting 13-point design; the maximum absolute error in $\hat{y}$ on $\dot{T}_{101}$ is 0.010, and the root mean squared error is 0.0045. The maximum $|\tau(t)|$ is 0.66 at $t = 0.06$, but when we try to add an enforcement site there, we encounter more numerical problems. We go no further, although we presumably could relax θ again and continue. The functions τ_m and ε_m are shown in Figure 4, for m=10, 11, and 13.

We now repeat this example using the nonnegative cubic correlation function with $\theta = 13$. We find that this correlation function allows us to use more enforcement sites in our design without encountering numerical problems. Also the τ and ε functions become smaller (as measured by

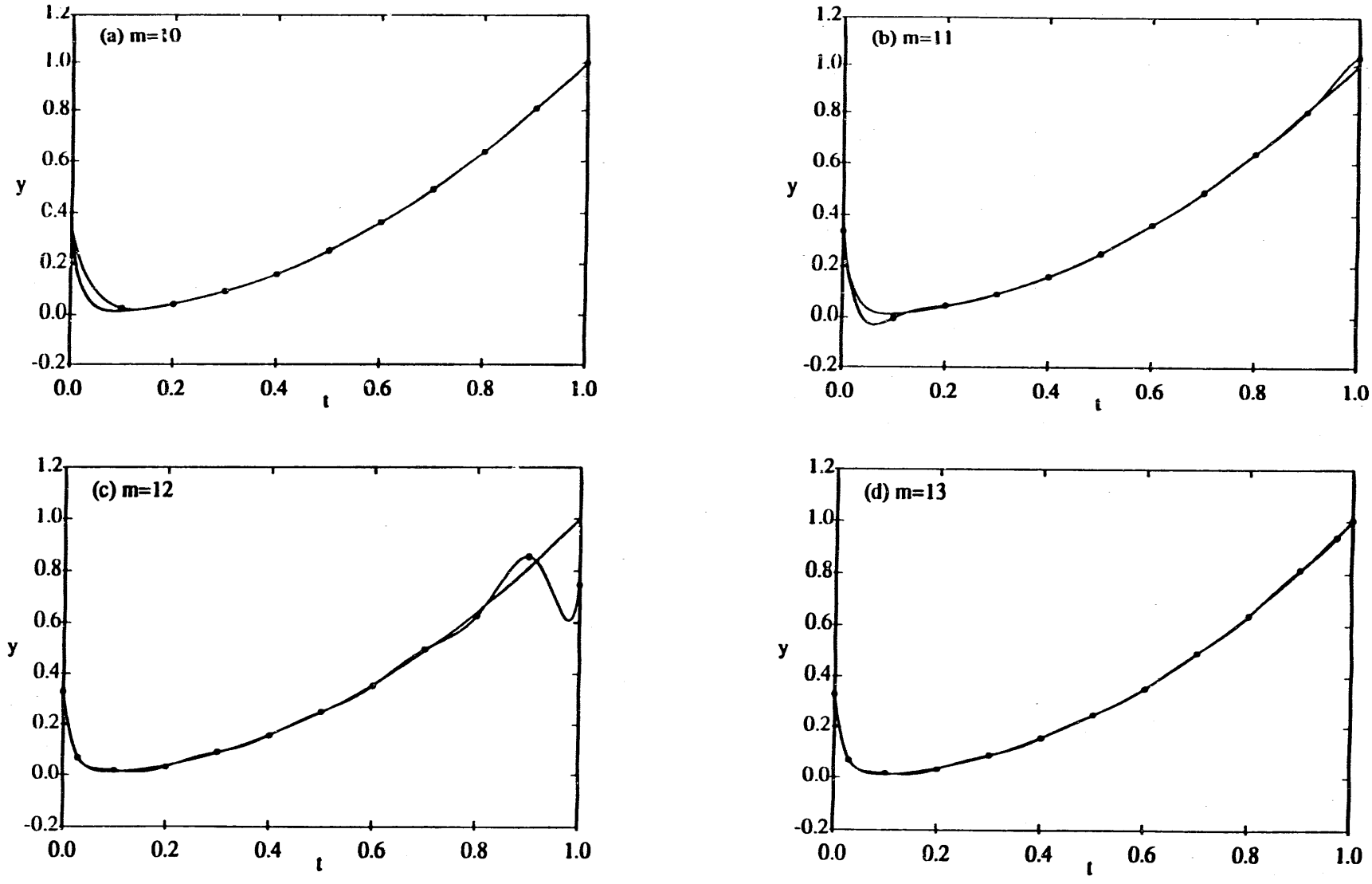

Figure 3. Plots of true solution y and approximate solutions $\hat{y}_m$, m = 10, 11, 12, 13, for Example 2, where the approximation is based on the Gaussian correlation with $\theta = 0.42$ for M = 10, 11 and $\theta = 0.25$ for m = 12, 13. Points that correspond to the m enforcement sites are marked on $\hat{y}$ by dots.

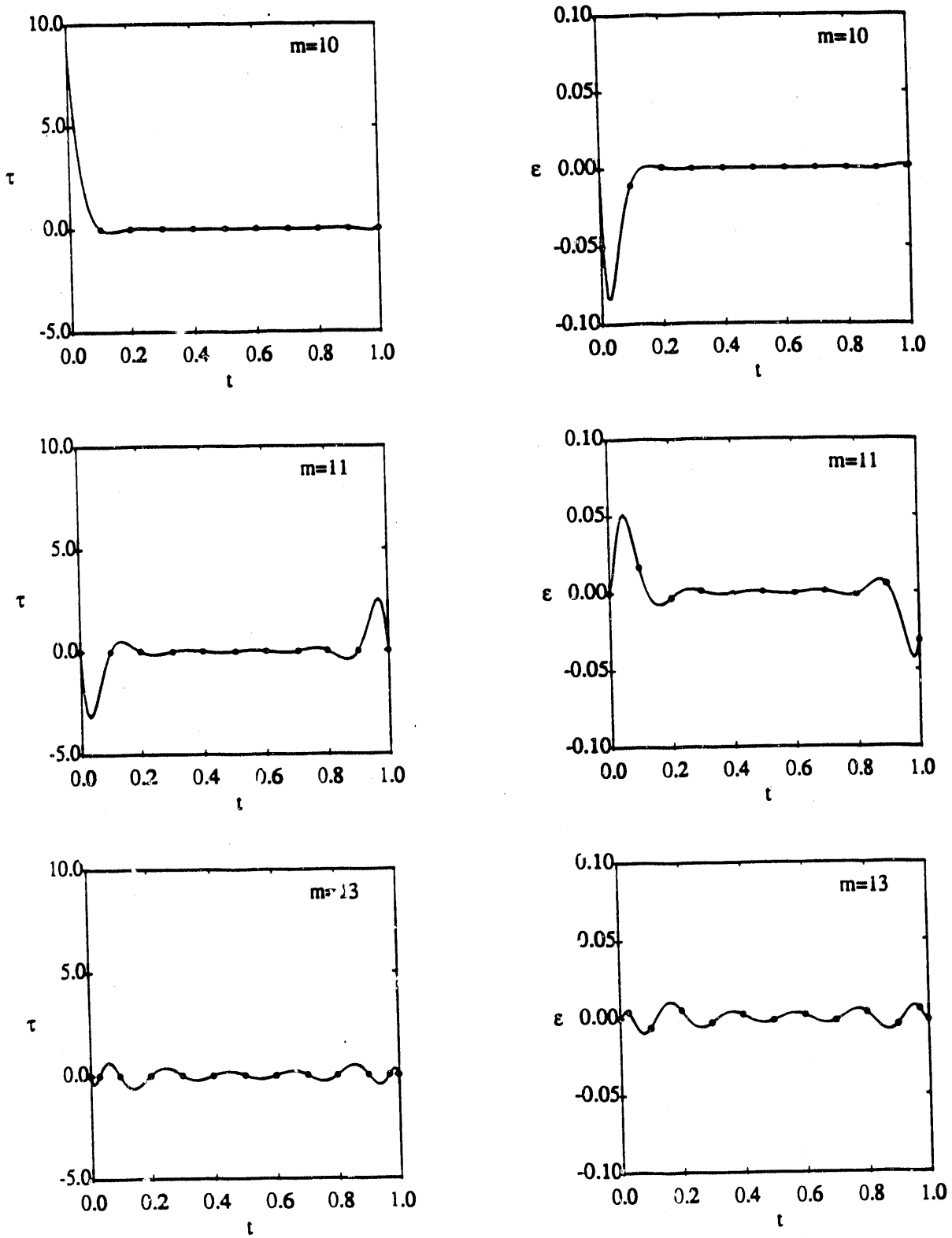

Figure 4. Plot of $\tau_m(t) = \hat{h}_m(t) - h(t)$ and $\varepsilon_m(t) = y(t) - \hat{y}_m(t)$, $m = 10, 11, 13$, for Example 2, where the approximation $\hat{y}$ is based on the Gaussian correlation with $\theta = 0.42$ for $m = 10, 11$ and $\theta = 0.25$ for $m = 12, 13$. Points that correspond to the m enforcement sites are marked on the curves by dots.

the maximum and sum of squared absolute values on $\dot{T}_{101}$) with the addition of each new site. After starting with the same initial 10-point design (from 0.1 to 1.0 in steps of 0.1), we add the 7 new points: 0, 0.04, 0.02, 0.01, 0.07, 0.03, 0.05, in that order, again using the strategy of adding the new point where $|\tau|$ is currently largest. For this 17-point design, the maximum absolute error in $\hat{y}$ on $\dot{T}_{101}$ is 0.00131 and the root mean squared error is 0.00032. The plot of $\hat{y}$ (not shown) is visually indistinguishable from the true solution y. The functions τ_m and ε_m are shown in Figure 5, for m=10, 11, and 17. Figure 6 shows the maximum absolute error in $\hat{y}$ on $\dot{T}_{101}$ as a function of the number of enforcement sites. We could continue to add enforcement sites without numerical difficulties, but this does not seem warranted given the accuracy of the approximation obtained at this stage.

3.2 Example 3

In order to approximate the solution of second-order linear differential equations, we require a correlation function that is at least four times differentiable. The nonnegative cubic correlation function does not have this property; therefore, we use only the Gaussian correlation function, which can be differentiated as many times as needed, for the second-order examples.

Problem:

$$2y(t) - 2ty'(t) + t^2y''(t) = t^3\ln(t) \qquad 1 \le t \le 2\,, \quad y(1) = 1, \quad y'(1) = 0. \tag{30}$$

The solution to this problem is

$$y(t) = \frac{7}{4}t + \frac{t^3}{2}\ln(t) - \frac{3}{4}t^3, \tag{31}$$

which is plotted in Figure 7. Our design consists of the 5 equally spaced sites from $t = 1.2$ through $t = 2.0$. The 7×21 matrix A has the form:

$$\begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \cdot & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & \cdot & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \alpha_0(t_3) & \alpha_1(t_3) & \alpha_2(t_3) & \cdot & 0 & 0 & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \cdot & \alpha_0(t_7) & \alpha_1(t_7) & \alpha_2(t_7) \end{bmatrix}$$

where $\alpha_0(t) = 2$, $\alpha_1(t) = -2t$, and $\alpha_2(t) = t^2$. The vector $\vec{z}$ is given by (13), with $c_1 = 1$, $c_2 = 0$, and $f(t) = t^3\ln(t)$.

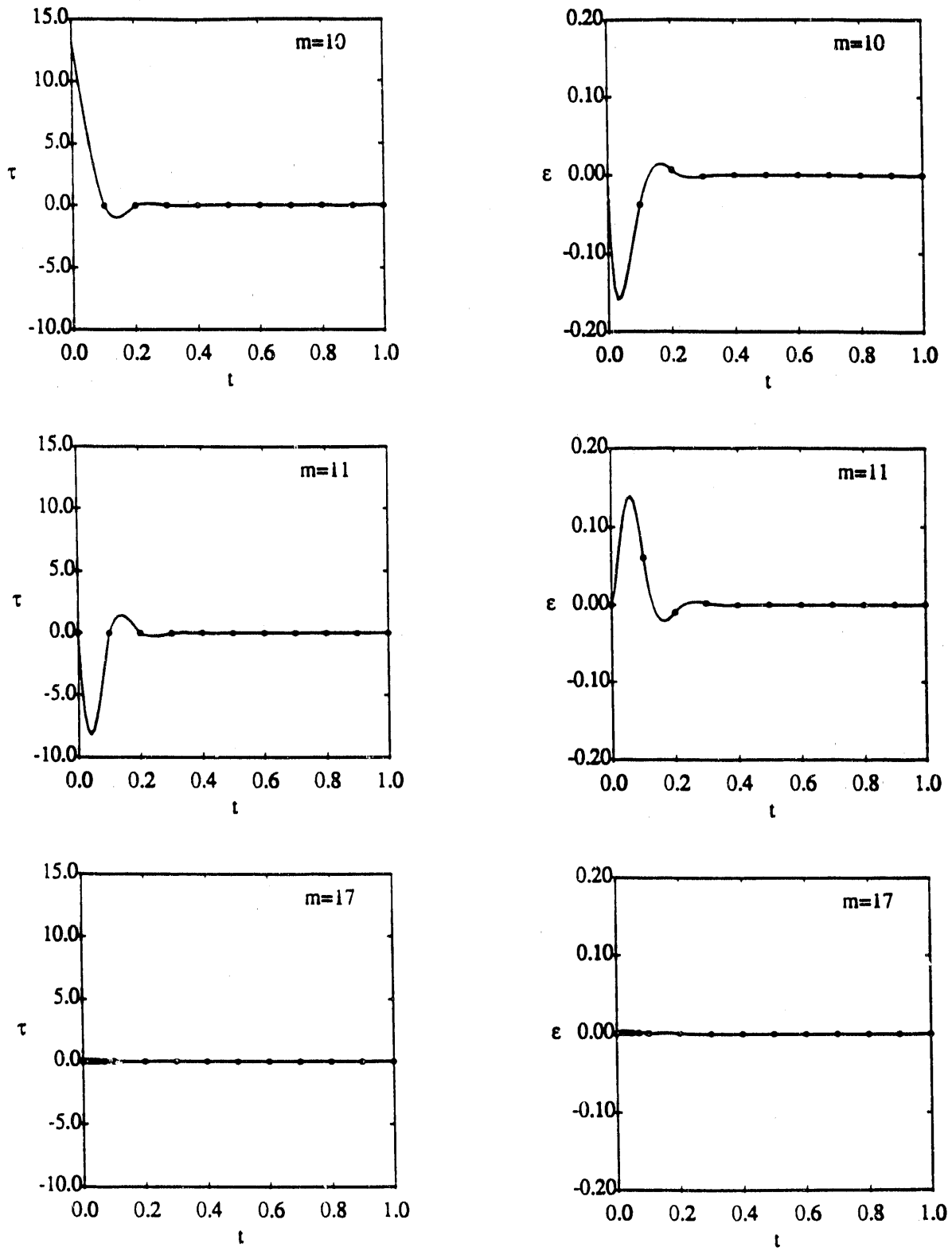

Figure 5. Plot of $\tau_m(t) = \hat{h}_m(t) - h(t)$ and $\varepsilon_m(t) = y(t) - \hat{y}_m(t)$, $m = 10, 11, 17$, for Example 2, where the approximation $\hat{y}$ is based on the nonnegative cubic correlation with $\theta = 13$. Points that correspond to the m enforcement sites are marked on the curves by dots.

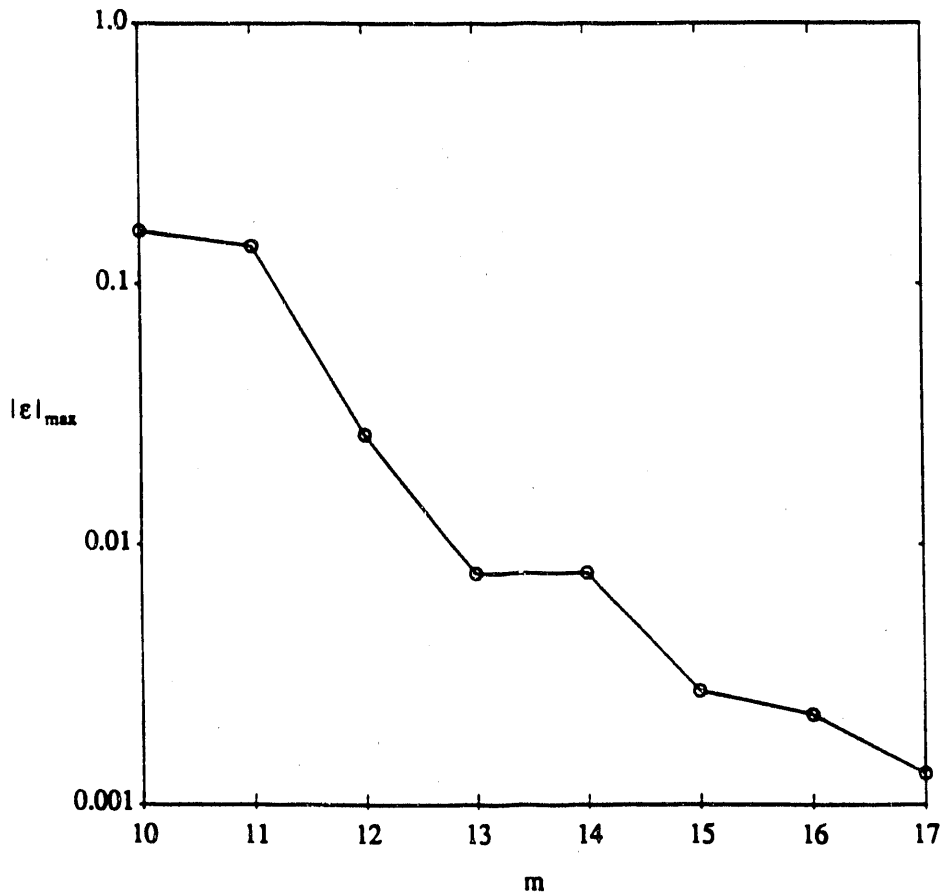

Figure 6. Plot of the maximum absolute error in $\hat{y}$, as a function of the number of enforcement sites m, for Example 2, where $\hat{y}$ is based on the nonnegative cubic correlation with $\theta = 13$.

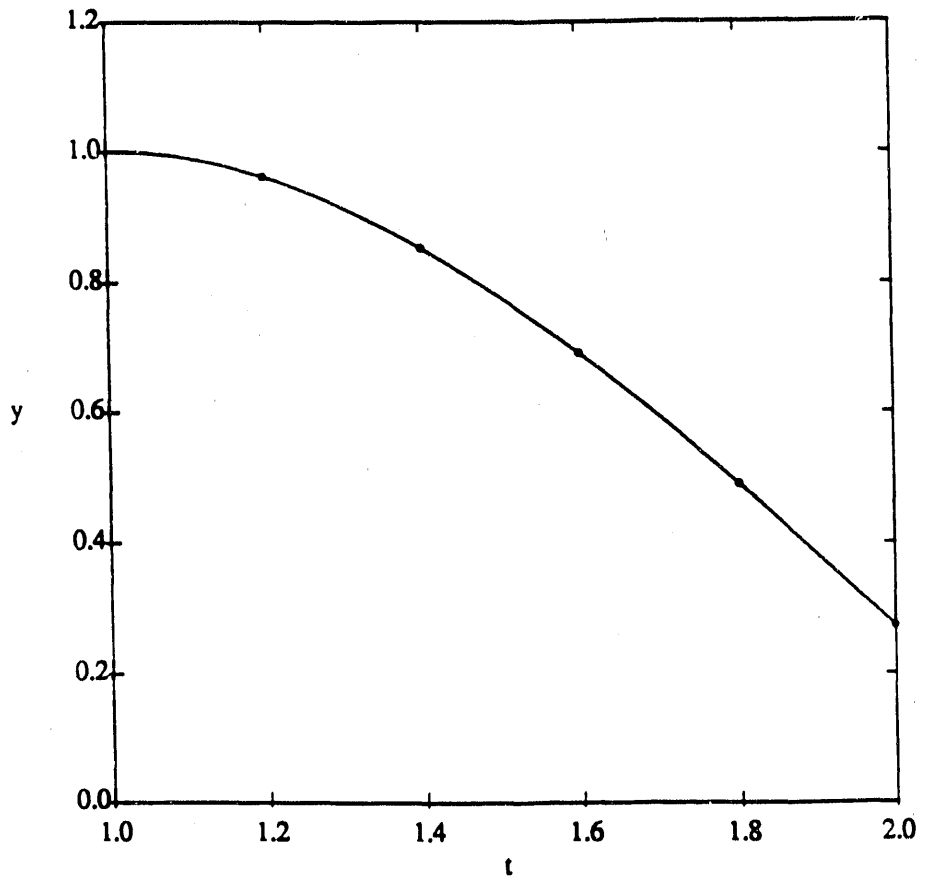

Figure 7. Plot of the solution function y for the second order differential equation of Example 3 (initial value problem) and Example 4 (two point boundary value problem). Points that correspond to the 5 enforcement sites are marked by dots. The approximation $\hat{y}$, based on the Gaussian correlation with $\theta = 9$ is virtually the same as y in both examples.

Our approximation, based on the Gaussian correlation function with $\theta = 9$, turns out to be exceptionally good in this example; the maximum absolute error in $\hat{y}$ on $\dot{T}_{101}$ is 7.84×10^{-6}.

3.3 Example 4

We conclude the examples using the differential equation of Example 3 (30), but stated as a two-point boundary value problem rather than as an initial value problem. The boundary conditions are:

$$y(1) = 1, \quad y(2) = 4\ln(2) - 5/2 = 0.27258872 \tag{32}$$

and the solution is the same as given in (31).

Using the same 5-point design and value of θ as in Example 3, the largest absolute difference between our approximation and the true solution on $\dot{T}_{101}$ is 2.17×10^{-7}. Although such accuracy cannot be expected in general, this example raises our hopes that the method will prove competitive.

4. COMPUTATION

We have written two programs for producing the approximation $\hat{y}$. They are similar in most respects, but one was written specifically for first-order equations and the other for second-order equations. By means of an input file, the user provides the number of side condition sites n, and for each of these, the location t_i, the order of derivative k(i), and the value c_i of $y^{(k(i))}(t_i)$, $i = 1, \cdots, n$. The number of enforcement sites m and their locations $t_{n+1}, \cdots, t_{n+m}$ are provided in the same input file. The differential equation (1) is specified through the user-defined functions $\alpha_0, \cdots, \alpha_n$, and f. For first-order equations, the user is given a choice of the Gaussian or nonnegative cubic correlation functions; this information is supplied interactively by the user in response to a prompt. For second-order equations, the Gaussian correlation is assumed. In both cases, the value of θ is also supplied interactively.

The method of computing the posterior expectation and variance of Y follows the matrix formulation described in Section 2. The program constructs the matrix A and the vector $\vec{e}$ and uses (17) and (20) to construct σ_{CC}. It computes $\hat{y}$ in (14) as

$$\hat{y}(t) = \hat{\mu} + \vec{\sigma}_{Ct}^{T}\vec{\beta} \tag{33}$$

where $\hat{\mu}$ in (23) is computed as

$$\hat{\mu} = \frac{\vec{\alpha}^{T}\vec{z}}{\vec{\alpha}^{T}A\vec{e}} \tag{34}$$

and the vectors $\vec{\alpha}$ and $\vec{\beta}$ are computed as solutions to the linear systems:

$$\sigma_{CC}\vec{\alpha} = A\vec{e} \tag{35}$$

$$\sigma_{CC}\vec{\beta} = \vec{z} - \hat{\mu}A\vec{e}, \tag{36}$$

respectively. The program also computes $\tau(t)$ in (27) for all t in $\dot{T}_{101}$. The derivatives of $\hat{y}$ needed for this are computed from

$$\hat{y}^{(j)}(t) = \frac{\partial^{j}\vec{\sigma}_{Ct}^{T}}{\partial t^{j}}\vec{\beta} \ . \tag{37}$$

The main computational work is in the solution of the two (n+m)-dimensional linear systems (35) and (36), both of which have the same coefficient matrix σ_{CC}. In the examples we have considered here, m+n ranged from 6 to 17. When new enforcement sites are added to an existing design, for fixed θ, computational savings can be achieved by updating, for example, the Cholesky factorization of σ_{CC} used in the solution of the linear systems. We have not yet incorporated this into the program.

5. DISCUSSION

This report has demonstrated the application of Bayesian statistical inference to the solution of linear differential equations. By viewing the differential equation and its side conditions as sources of data about the unknown solution function y, we have used simple Bayesian techniques to produce the posterior mean function $\hat{y}$, which serves as our approximation to y.

Our main concern here has been to show that this approach can work well, in the sense of producing good approximations easily. We have shown that for either the Gaussian or nonnegative cubic correlation function, with selected values of θ, it is possible to obtain a very good approximation of the solution to both first-order and second-order linear differential equations by this method. However, we have treated lightly many of the following important issues, which will need to be resolved before the ultimate utility of the method can be determined.

5.1 Design

One of the advantages of a Bayesian approach is that there exist accepted ways to quantify such concepts as "amount of information". However, we have not really exploited this in formulating design criteria. That is, we have not formally addressed the question of where to place additional sites in order to maximize the gain in information about y. Our current design method is to start with an equispaced design and then add new sites where the value of $|\tau|$ is largest. This is intuitively appealing and seems to work well, but lacks a rigorous foundation.

5.2 Choice of prior process

Although we restricted our attention to Gaussian prior processes, stochastic processes other than the Gaussian could have been considered. However, such processes would, in general, greatly complicate computation of $\hat{y}$. Furthermore, given freedom to choose the form of an appropriate correlation function, use of Gaussian processes may not pose a serious limitation in practice.

Within the framework of Gaussian processes, there is much room for the investigation of correlation functions other than the two we have used here. A nonnegative quintic correlation would be useful for second-order problems, since the cubic can be used only for first-order equations. The Gaussian correlation function, which is infinitely differentiable, can be used for differential equations of any order, but requires considerable care to avoid numerical problems.

Both the correlation functions we considered here depend on a single parameter θ. In practice, we have found it useful to select a value of θ which approximately minimizes the sum of squares of τ over a representative finite subset of T, but this minimization can be computationally expensive.

5.3 Computation

The main computational disadvantage of our method is that it requires the solution of an $(n+m)\times(n+m)$ system of linear algebraic equations, and the coefficient matrix of the system may not be well-conditioned. This is offset somewhat by the fact that $(n+m)$ does not need to be very large to obtain a good approximation. One possible way of speeding up the computation is to choose the correlation function and the design so that σ_{CC} is sparse. For example, if one uses an equispaced design in which the distance between two neighboring sites is δ, and a nonnegative cubic correlation function with $\delta < \theta < 2\delta$, σ_{CC} will be tridiagonal.

5.4 Extensions.

An appealing feature of the Bayesian approach is that it extends naturally to systems of linear ordinary differential equations (o.d.e.'s) and to linear partial differential equations (p.d.e.'s). For the former, prior uncertainty about the unknown functions is represented by independent stochastic processes, each of which can be of the same form considered in this report. Each o.d.e. is then considered a source of data for the Bayesian method. The mechanics of the approximation are the same as for single o.d.e.'s but the system of linear algebraic equations to be solved is likely to be much larger. The design issues are likely to be more interesting, also. It is not clear whether one should enforce the whole system at all design sites, or take a more selective approach.

The Bayesian approach to linear p.d.e.'s is the same as it is to linear o.d.e.'s. The only obvious complication is that T is multidimensional, and suitable correlation functions need to be defined. This can be done using the *product correlation rule*, as is done for the prediction problem (Currin et al. 1988, Sacks et al. 1989). Once the intersite correlations are defined, the mechanics are again the same as for linear o.d.e.'s. Our experience with the prediction problem leads us to believe that we may be able to achieve good approximations to the solution of linear p.d.e.'s with fairly simple and economical designs, i.e., without having to impose a dense, well-structured mesh on T. Further extension to systems of linear p.d.e.'s leads to problems that are larger in scale, but require no fundamental change of approach.

Ultimately, of course, we would like to extend our approach to nonlinear differential equations, but the way is not so clear. Our approach for linear differential equations benefits greatly from the fact that, for Gaussian prior processes, the data provided by the differential equation(s) are normally distributed, a priori, with easily derivable means and covariances. For nonlinear equations, this is not the case. Our inclination is to attempt to develop an iterative method, based on the solution of appropriately linearized equations at each stage.

These extensions will require much more work than we have so far undertaken. Although we can only speculate about their eventual success, we are encouraged by the results we have obtained in the few simple examples described in this report.

REFERENCES

[1] Currin, C., Mitchell, T., Morris, M., and Ylvisaker, D. (1988), "A Bayesian Approach to the Design and Analysis of Computer Experiments", ORNL-6498, available from National Technical Information Service, 5285 Port Royal Road, Springfield, VA 22161.

[2] Gear, C. W. (1971), *Numerical Initial Value Problems in Ordinary Differential Equations*, Prentice-Hall, Englewood Cliffs, New Jersey.

[3] Isaacson, E. and Keller, H.B. (1966), *Analysis of Numerical Methods*, Wiley, New York.

[4] Morrison, D.F. (1967), *Multivariate Statistical Methods*, McGraw-Hill, New York.

[5] Prenter, P.M. (1975), *Splines and Variational Methods*, Wiley, New York.

[6] Sacks, J., Welch, W.J., Mitchell, T.J., and Wynn, H.P. (1989), "Design and Analysis of Computer Experiments" (with discussion), *Statist. Sci.* **4**, 409-435.

END

DATE FILMED

03 / 05 / 91