DEPARTMENT OF STATISTICS

University of Wisconsin
Madison, Wisconsin 53706

TECHNICAL REPORT NO. 530

June 1978

DETERMINATION OF AN OPTIMAL MESH
FOR A COLLOCATION-PROJECTION METHOD FOR
SOLVING TWO POINT BOUNDARY VALUE PROBLEMS

Manohar L. Athavale
IBM Corporation
Hopewell Junction, New York 12533

and

Grace Wahba*
Department of Statistics
University of Wisconsin
Madison, Wisconsin 53706

TYPIST: Mary E. Arthur

*This research was partially sponsored by the Office of Naval Research
under Grant No. N00014-77-C-0675.
Abstract

We consider mesh-point optimization for certain collocation-projection methods for solving boundary value problems (BVP) for ordinary differential equations. Consider the BVP

\[(L_m x)(t) = g(t), \quad t \in [0,1]\]

\[x^{(v)}(0) = x^{(v)}(1) = 0, \quad v = 0, 1, \ldots, m/2-1\]

where \(L_m\) is an \(m^{\text{th}}\) order linear differential operator. We assume conditions guaranteeing that there is a unique solution in a given (reproducing kernel) Hilbert space \(H_R\), and that \(|(L_m x)(t)| \leq M |x|_R\), \(0 \leq t \leq 1\), for all \(x \in H_R\), and some \(M\), where \(|\cdot|_R\) is the norm in \(H_R\). For a given mesh \(T_N = \{t_{iN}\}_{i=1}^N\), we let \(x_N\), our approximate solution to the BVP be that element in \(H_R\) of minimal norm which satisfies the boundary conditions and matches the data on \(T_N\), that is \((L_m x_N)(t_{iN}) = g(t_{iN}), i = 1, 2, \ldots, N\). \(x_N\) is both a collocation and an orthogonal projection approximation to \(x\), and for certain \(H_R\) equivalent to the Sobolev space \(W^2_2(r)\), \(x_N\) is a spline approximant to \(x\).

We are interested in choosing the mesh \(T_N\) so that \(||x-x_N||_R|\) is as small as possible. The optimal mesh we are after depends on the unknown \(x\). Under certain circumstances a mesh behaving essentially like an optimal mesh can be characterized by a cumulative distribution function \(F^*\) on \([0,1]\), which depends on \(x\). A (nearly) optimal mesh \(T_N^* = \{t_{iN}^*\}\) is determined by solving \(F^*(t_{iN}^*) = i/(N+1), i = 1, 2, \ldots, N\). \(F^*\) has been given by Sacks and Ylvisaker [7] and
Wahba [8,11,12], under various conditions. In this paper we show how an estimate $F_n^*$ of $F^*$ can be computed from data, starting with an arbitrary (nice) mesh with $n$ points. Once $F_n^*$ is obtained, then a new mesh, say $\hat{t}_N = (\hat{t}_{iN})^N_{i=1}$ can be obtained as $F_n^*(\hat{t}_{iN}) = i/(N+1)$, $i = 1,2,...,N$, and the final approximate solution $x_N$ computed using the estimated approximately optimal mesh $\hat{T}_N$. From a different point of view, if data points $g(t_{iN})$ are obtained from an experiment and are expensive to measure, this approach is that of the sequential design of an experiment. Data at a preliminary uniform mesh (design) are obtained, and this data is used to obtain an improved mesh (design). These results apply to more general linear operator equations.
1. Introduction and Preliminaries

Consider the boundary value problem (BVP)

\[(L_m x)(t) = g(t), \quad t \in [0, 1]\]

\[x^{(v)}(0) = x^{(v)}(1) = 0, \quad v = 0, 1, \ldots, m/2 - 1 \quad (1.1)\]

where \(L_m\) is an \(m\)th order differential operator with an \(m\)-dimensional null space. We suppose that the domain of \(L_m\) is a reproducing kernel Hilbert space (RKHS) which contains the null space of \(L_m\).

Letting \(|\cdot|_R\) be the norm in \(H_R\), we further suppose

\[a) \quad |(L_m x)(t)| \leq M||x||_R, \quad x \in H_R, \quad t \in [0, 1]\]

\[b) \quad |x^{(v)}(i)| \leq M||x||_R, \quad x \in H_R, \quad v = 0, 1, \ldots, m/2 - 1, \]

\[i = 0, 1\]

Conditions a) and b) are always satisfied, for example if \(H_R\) is the Sobolev space \(W^2_2(r)\),

\[W^2_2(r) = \{x: x, x', \ldots, x^{(r-1)} \text{ abs. cont.}, x^{(r)} \in L^2_2[0, 1]\}\]

with \(r > m\). Then by the Riesz representation theorem there exist \(\{n_t, t \in [0, 1]\}\), and \(\{R_{iv}, i = 0, 1, v = 0, 1, \ldots, m/2 - 1\}\) in \(H_R\) such that

\[(L_m x)(t) = \langle n_t, x \rangle_R, \quad t \in [0, 1]\]

\[x^{(v)}(i) = \langle R_{iv}, x \rangle_R, \quad i = 0, 1, v = 0, 1, \ldots, m/2 - 1.\]

Given a mesh \(T_N = \{t_i\}_{i=1}^N \equiv \{t_i\}_{i=1}^N\), and data \(g(t_i), i = 1, 2, \ldots, N\), we take as an approximate solution, that element \(x_N\) in \(H_R\) of minimal \(H_R\) norm satisfying the boundary conditions:
\[ \langle R_{iv}, x_N \rangle_R = 0, \quad i = 0,1, \quad v = 0,1, \ldots, m/2-1 \]

and matching the data on the given mesh

\[ \langle n_{i_1}, x \rangle_R = g(t_i), \quad i = 1,2,\ldots,N. \]

We will assume

c) \{ n_t, t \in [0,1], R_{iv}, \quad i = 0,1, \quad v = 0,1, \ldots, m/2-1 \} are linearly independent and span \( H_R \).\]

Condition c) can be shown to be satisfied, if, e.g.

\[ L_m = \sum_{j=0}^{m} a_j(t) D^m \] (1.2)

with \( \omega > a_m(t) > 0 \), and \( H_R = W^r_2, \quad r > m \).

It is not hard to show (see [2]) that conditions a), b) and c) guarantee a unique solution in \( H_R \) to the BVP for any \( g \in L_m(H_R) \), and that \( x_N \) is uniquely determined for any mesh and data vector. \( x_N \) is both a collocation approximation and an orthogonal projection approximation to \( x \). Letting \( V_N = \text{span}\{ n_{t_1}, \ldots, n_{t_N} \}, S_m = \text{span}\{ R_{iv}, \quad i = 0,1, \quad v = 0,1, \ldots, m/2-1 \}, x_N \) is the orthogonal projection of \( x \) onto \( V_N \oplus S_m \). This type of approximate solution was suggested in [5]. If the reproducing kernel for \( H_R \) is taken, e.g. as

\[ R(s,t) = \sum_{i=0}^{r-1} \frac{s^i t^j}{(j!)^2} + \int_0^1 \frac{(s-u)^{r-1}(t-u)^{r-1}}{[(r-1)!]^2} \, du, \quad (1.3) \]
then $H_R$ is topologically equivalent to $W_2^{(r)}$. In this case the $R_{i,v}$ are polynomials, $x_{t_i}(\cdot)$ is a polynomial spline of degree $2r-1$, possessing $2r-m-2$ continuous derivatives and a single knot at $t_i$, and $x_N$ is a polynomial spline of degree $2r-1$ and continuity class $2r-m-2$ with knots at the collocation points. To verify this last assertion, let $R_s(\cdot) = R(s,\cdot)$ be the representer of the evaluation functional at $s$ in $H_R$, then

$$R_{i,v}(s) = \langle R_{i,v}, R_s \rangle_R = \left. \frac{a(v)}{a(t_i)} R(s,t) \right|_{t=t_i}$$ (1.4)

$$n_{t_i}(s) = \langle n_{t_i}, R_s \rangle_R = \sum_{v=0}^{m-1} a_v(t_i) \left. \frac{a(v)}{a(t_i)} R(s,t) \right|_{t=t_i}$$ (1.5)

and it can be checked that the functions of $s$ on the right have the claimed properties. (Details may be found in [2]).

We would like to choose $T_N$ to minimize $\|x-x_N\|_R$. We do not know how to do this, but will do something very close, as follows:

Let $V = \text{span}(n_t, t \in [0,1])$. By c), $H_R = V \oplus S_m$. Now, $S_m$ is of dimension $m$. The codimension of $V$ is also $m$, since $V^\perp$ is of dimension $m$. This follows since $V^\perp$ is the null space of $L_m$, since $\langle n_t, x \rangle_R = 0, t \in [0,1]$ if and only if $L_mx = 0$. Thus each $x \in H_R$ has a unique decomposition

$$x = y + z$$ (1.6)

with $y \in V$ and $z \in S_m$.

Letting $P_S$ and $P_{V_N}$ be the orthogonal projectors in $H_R$ onto $S_m \oplus V_N$ and $V_N$ respectively, we have
\[ ||x-x_N||_R = ||x-P_Sx||_R = ||(y+z)-P_S(y+z)||_R \]
\[ \leq ||y-P_Sy||_R + ||z-P_SZ||_R \]
\[ = ||y-P_Sy||_R \]
\[ \leq ||y-P_{V_N}y||_R. \]

We will be able to choose \( T_N \) with the goal of minimizing \( ||y-P_{V_N}y||_R \).

Let
\[ F(t) = \int_0^t f(s)ds \]
where \( f > 0 \) and \( F(1) = 1 \), and let \( T_N = T_N(f) \) be determined by

\[ T_N = \{t_{i_N}\}_{i=1}^N, \text{ as} \]

\[ F(t_{i_N}) = i/(N+1), \quad i = 1, 2, \ldots, N. \quad (1.7) \]

A uniform mesh corresponds to \( f(s) = 1 \), \( s \in [0,1] \). It is known from [7,8,12], that, under some further conditions on \( y \), the large \( N \) behavior of \( ||y-P_{V_N}y||_R \) can be described fairly precisely as a function of the mesh cumulative distribution function (c.d.f.) \( F \) appearing in (1.7). Furthermore the \( F \), call it \( F^* \) which (loosely) minimizes \( ||y-P_{V_N}y||_R \) is known from [7,8,12], and, it depends on the unknown \( y \). In this paper we construct, starting from an arbitrary (nice) mesh with \( n \) points, and the values of \( g \) on this mesh, an approximation \( F_n^* \) to \( F^* \). Once \( F_n^* \) is obtained, a new mesh \( \hat{T}_N = \{t_{\hat{i_N}}\} \), say, can be determined from

\[ F_n^*(t_{\hat{i_N}}) = i/(N+1), \quad i = 1, 2, \ldots, N. \]
Then this new mesh can be used to compute the final approximant $x_N$.

This technique has a greater generality than BVP's, we indicate its usefulness for more general linear operator equations at the end.

We now proceed to describe the results from [12] that we need.

Let $Q(s, t) = <\eta_s, \eta_t>_R = (L_m \eta_t)(s) = L_m(s) L_m(t) R(s, t)$, where $L_m(s)$ means $L_m$ applied to a function of $s$. Let $H_Q$ be the RKHS with RK $Q$. There exists an isometric isomorphism between $V$ and $H_Q$ generated by the correspondence

$\eta_t \in V \sim Q_t \in H_Q$

where $Q_t(\cdot) = Q(t, \cdot)$ is the representer of evaluation at $t$ in $H_Q$.

To see this note that

$<\eta_s, \eta_t>_R = Q(s, t) = <Q_s, Q_t>_Q$.

Furthermore $y \in V \sim u \in H_Q$ iff $L_m y = u$. Since $H_R = V \oplus V^\perp$, where $V^\perp$ is the null space of $L_m$, $L_m(H_R) = L_m(V) = H_Q$. Let $P_{TN}$ be the orthogonal projector in $H_Q$ onto $span(Q_t, \ldots, Q_{TN})$. Then if $y \in V \sim u \in H_Q$, by the aforementioned isometric isomorphism,

$||y - P_{V_N} y||_R = ||u - P_{TN} u||_Q$. \hspace{1cm} (1.8)

Consider $u \in H_Q$ with a representation

$u(t) = \int_0^1 Q(t, s) \rho(s) ds \hspace{1cm} (1.9)$

for some $\rho \in L_2$. It is for $u$ of the form (1.9) that $||u - P_{TN} u||_Q$
can be described in terms of the mesh c.d.f. We have \( y \in V \cap u \) of 
\begin{equation}
(1.8) \text{ if}
\end{equation}
\begin{equation}
y(t) = \frac{1}{n} \int n_t(s) \rho(s) ds,
\end{equation}
equivalently
\begin{equation}
(L_m y)(t) = \int_0^1 Q(t,s) \rho(s) ds = u(t).
\end{equation}
(We note that \( u \parallel a \) unless \( S_m = V^1 \).)

We will henceforth assume that \( y \) has a representation (1.10). If 
\( H_R = W_2^{(r)} \), then (1.10) entails that \( x \in W_2^{(m+2q)} = W_2^{(2r-m)} \).

It is supposed in [12] (loosely), that \( Q(s,t) \) has the continuity properties of a Green's function of a self adjoint linear differential operator of order 2q, with
\begin{equation}
\lim_{s \to t} \frac{s^{2q-1}}{3s^{2q-1}} Q(s,t) - \lim_{s \to t} \frac{s^{2q-1}}{3s^{2q-1}} Q(s,t) = (-1)^m \alpha(t)
\end{equation}
with \( \alpha \) and \( \alpha' \) continuous. (More generally the hypotheses of the theorem in [12, Section 2] are being assumed.) If \( L_m \) and \( H_R \) are given by (1.2) and (1.3) then \( q = r-m \) and \( \alpha(t) = 1/a_m(t) \). With these assumptions, the linear functionals \( N_{j,k} u^{(k)}(t_j) \) are continuous in \( H_Q \) for \( k = 0,1,\ldots,q-1 \). Let \( Q_{t_j,k} \) be the representer of \( N_{j,k} \) in \( H_Q \), and let \( P_{q,T} \) be the projection operator in \( H_Q \) onto \( \text{span}\{Q_{t_j,k}^{(k)}\}_{j=1}^{n} \), \( k=0,\ldots,q-1 \), and let \( P_{T} \) be the projection operator in \( H_Q \) onto \( \text{span}\{Q_{t_j,k}^{(k)}\}_{j=1}^{n} \). Then for \( u \in H_Q \),
\begin{equation}
\inf_{T} ||u-P_{q,T} u||_Q \leq \inf_{T} ||u-P_{T} u||_Q \leq \inf_{T} ||u-P_{q,T} u||_Q,
\end{equation}
where \( \inf \) means the infimum is taken over all \( \xi \)-point designs. It is known that if \( q = 2 \), the right hand inequality becomes an equality. See [8] for details. The reason for presenting these inequalities, is, that an exact asymptotic expression is available for
\[ ||u-P_{q,T} u||_Q \] and is given by
Theorem [11,12]

Let \( f \) be a strictly positive density, \( F(t) = \int_0^t f(u) du \), let \( \rho \) and \( \alpha \) be as in (1.10) and (1.12) and suppose \( \rho \) and \( \alpha \) are strictly positive, continuous and possess bounded first derivatives. Let

\[ T_N = T_N(f) = \{ t_{in} \}_{i=1}^N, N = 1, 2, \ldots \] be determined by

\[ F(t_{in}) = \frac{i}{(N+1)}, \quad i = 1, 2, \ldots \]

Then

\[
||u-P_{q,T_n}u||_Q^2 = \frac{1}{N^q} \frac{(q!)^2}{(2q)!(2q+1)!} \int_0^1 \frac{\rho^2(s)\alpha(s)}{f^{2q}(s)} ds \left( 1+o(1) \right). \tag{1.14}
\]

where \( o(1) \to 0 \) as \( N \to \infty \). Furthermore, by using a Holder inequality and the fact that \( \int_0^1 f(s)ds = 1 \), we have

\[
\int_0^1 \frac{\rho^2(s)\alpha(s)}{f^{2q}(s)} ds \geq \left[ \int_0^1 \left[ \frac{\rho^2(s)\alpha(s)}{f^{2q}(s)} \right]^{1/2q} ds \right]^{2q+1}. \tag{1.15}
\]

and the lower bound on the right of (1.15) is achieved if and only if \( f = f^* \) given by

\[
f^*(s) = \frac{\left[ \frac{\rho^2(s)\alpha(s)}{f^{2q}(s)} \right]^{1/(2q+1)}}{\int_0^1 \left[ \frac{\rho^2(u)\alpha(u)}{f^{2q+1}(u)} \right]^{1/(2q+1)} du}, \tag{1.16}
\]

that is

\[
F^*(t) = \int_0^1 \left[ \frac{\rho^2(s)\alpha(s)}{f^{2q+1}(s)} \right]^{1/(2q+1)} ds / \left[ \int_0^1 \left[ \frac{\rho^2(s)\alpha(s)}{f^{2q}(s)} \right]^{1/(2q+1)} ds. \tag{1.17}
\]

We remark that all the regularity conditions on \( \rho \) and \( \alpha \) are probably not necessary but are artifacts of earlier proofs.

The result of this paper is as follows. Given a uniform mesh \( \{ t_{in} \}_{i=1}^n \), we show how an approximation \( F^*_n \) to \( F^* \) may be obtained from
the data vector \( g(t_{1n}), \ldots, g(t_{nn}) \), using coefficients which are an intermediate step in the calculation of \( x_n \).

In Section 2 we define \( F_n^* \) and show that \( \lim_{n \to \infty} F_n^*(t) = F^*(t) \), \( 0 \leq t \leq 1 \). In Section 3 we briefly mention some numerical results. In Section 4 we note how the results apply to more general linear operator equations. In Section 5 we note how the problem is formally equivalent to an optimal quadrature problem and compare it to other work.

2. The estimate \( F_n^* \) of \( F^* \)

Given an arbitrary mesh \( T_n \) of distinct points \( \{ t_{in} \}_{i=1}^n = \{ t_i \}_{i=1}^n \), then \( x_n^* \), that element of minimal \( H^* \) norm satisfying the boundary conditions and \( (L^* x_n)(t_i) = g(t_i) \), has a representation

\[
x_n(t) = \sum_{\nu=0}^{m/2-1} d_{\nu} R_{\nu}(t) + \sum_{i=1}^{n} c_i \eta_{t_i}(t) + \sum_{\nu=0}^{m/2-1} d_{1\nu} R_{1\nu}(t)
\]

where the \( R_{\nu} \) and \( \eta_{t_i} \) have been given in (1.4) and (1.5).

Forcing \( x_n \) to satisfy the boundary conditions and the differential equation at the mesh points leads to the following system of \( n+m \) equations in the \( c \)'s and \( d \)'s:

\[
\sum_{\nu=0}^{m/2-1} d_{\nu} R_{\nu}(0) + \sum_{i=1}^{n} c_i \eta_{t_i}(0) + \sum_{\nu=0}^{m/2-1} d_{1\nu} R_{1\nu}(0) = 0, \quad \mu = 0, 1, \ldots, m/2-1
\]

\[
\sum_{\nu=0}^{m/2-1} d_{\nu} (L^* R_{\nu})(t_j) + \sum_{i=1}^{n} c_i (L^* \eta_{t_i})(t_j) + \sum_{\nu=0}^{m/2-1} d_{1\nu} (L^* R_{1\nu})(t_j) = g(t_j), \quad j = 1, 2, \ldots, n
\]

\[
\sum_{\nu=0}^{m/2-1} d_{\nu} R_{\nu}(1) + \sum_{i=1}^{n} c_i \eta_{t_i}(1) + \sum_{\nu=0}^{m/2-1} d_{1\nu} R_{1\nu}(1) = 0, \quad \mu = 0, 1, \ldots, m/2-1
\]

(2.2)
These equations are equivalent to (2.1) and

\[ \langle x_n, R_{0v} \rangle_R = 0, \quad v = 0, 1, \ldots, m/2 - 1 \]

\[ \langle x_n, n_i \rangle_R = g(t_i), \quad i = 1, 2, \ldots, n \]

\[ \langle x_n, R_{1v} \rangle_R = 0, \quad v = 0, 1, \ldots, m/2 - 1. \] (2.3)

The system resulting from (2.2) is not particularly well suited for computation. When R is given by e.g. (1.3), the solution can be expressed in terms of a B-spline basis, resulting in a linear system involving band matrices. See [2] for details of the calculation.

As an estimate \( F^*_n \) of \( F^* \) of (1.17) we take

\[ F^*_n(t) = \frac{1}{t} \int_0^t \left[ \rho_n(s) \alpha(s) \right]^{1/(2q+1)} ds \int_0^1 \left[ \rho_n(s) \alpha(s) \right]^{1/(2q+1)} ds \] (2.4)

where \( \rho_n \) is an estimate of the \( \rho \) appearing in (1.10). The function \( \rho_n \) is formed from the vector \( (c_1, c_2, \ldots, c_n) \) appearing in (2.2) as the piecewise linear function on \([0,1]\) joining the points \((0,0), (t_1, c_1/h), (t_2, c_2/h), \ldots, (t_n, c_n/h), (0,1)\), where \( h = 1/(n+1) \). Our main result is:

Theorem. Let \( \alpha \) and \( \rho \) be strictly positive and continuous, and (without loss of generality), let \( t_i = i/(n+1) \), \( i = 1, 2, \ldots, n \). Then

\[ \lim_{n \to \infty} F^*_n(t) = F^*(t), \quad t \in [0,1] \]
Proof:

We first show

\[
\lim_{n \to \infty} \int_0^1 \rho_n(s)Q(t,s)ds = \int_0^1 \rho(s)Q(t,s)ds \equiv u(t), \quad t \in [0,1].
\] (2.5)

Using the trapezoidal rule in each interval on the integral on the left hand side we have

\[
\int_0^1 \rho_n(s)Q(t,s)ds = \sum_{i=1}^{n-1} \left[ (\rho_n(t_i)Q(t_i)+\rho_n(t_{i+1})Q(t_{i+1}))\frac{h}{2} + O(h^2) \right]
\]

\[
= \sum_{i=1}^{n} c_i Q_{t_i}(t) + \sum_{i=1}^{n-1} O(h^2)
\] (2.6)

We will next show that \( \| \sum_{i=1}^{n} c_i Q_{t_i} - P_y y \|_R \to 0 \), where \( y \) is the element in the decomposition \( x = y + z, \quad y \in V, \quad z \in S_m \).

This will guarantee that \( \| \sum_{i=1}^{n} c_i Q_{t_i} - P_y u \|_0 \to 0 \) by the isometric isomorphism, and, since \( \| u - P_y u \|_0 \to 0 \), it will follow that

\( \| \sum_{i=1}^{n} c_i Q_{t_i} - u \|_0 \to 0 \), and hence \( \sum_{i=1}^{n} c_i Q_{t_i}(t) \to u(t) \), and then (2.5) holds. Let \( \gamma_1, \ldots, \gamma_m \) be an orthonormal basis for \( S_m \). We may write

\[
x = \sum_{i=1}^{n} \tilde{c}_i \tilde{n}_{t_i} + \delta + \sum_{i=1}^{m} \tilde{\theta}_i \gamma_i
\] (2.7)

where

\[
\sum_{i=1}^{n} \tilde{c}_i \tilde{n}_{t_i} = P_y y
\] (2.8)

\[
\delta = (I-P_y) y
\]

Also, we may write

\[
x_n = \sum_{i=1}^{n} c_i \tilde{n}_{t_i} + \sum_{j=1}^{m} \tilde{\theta}_j \gamma_j
\] (2.9)

where \( c = (c_1, \ldots, c_n) \) is as in (2.1) and (2.2).
Letting $Q_n$ be the $n \times n$ matrix with $ij$th entry $\langle n_{t_i}, n_{t_j} \rangle_R = (L_{p_{t_i}})(t_j)$, and $\Sigma$ be the $n \times m$ matrix with $ij$th entry $\langle n_{t_i}, \gamma_j \rangle_R$, we have that (2.1) and (2.2) are equivalent to

$$
\begin{align*}
    x_n &= (n_{t_1}, \ldots, n_{t_n}, \gamma_1, \ldots, \gamma_m) \\
    \Sigma' &= I
\end{align*}
$$

\[
\begin{pmatrix}
    n_{t_1} \\ \\
    \vdots \\ \\
    n_{t_n}
\end{pmatrix}
\]

(2.10)

Now, using (2.7), gives

$$
\begin{align*}
    \begin{pmatrix}
        \langle x, n_{t_1} \rangle_Q \\
        \vdots \\
        \langle x, n_{t_n} \rangle_Q
    \end{pmatrix}
    &= Q_n \tilde{c} + \Sigma \tilde{\theta} \\
    \begin{pmatrix}
        \langle x, \gamma_1 \rangle_Q \\
        \vdots \\
        \langle x, \gamma_m \rangle_Q
    \end{pmatrix}
    &= \Sigma' \tilde{c} + I \tilde{\theta} + \left( \begin{pmatrix}
        \langle x, \gamma_1 \rangle_Q \\
        \vdots \\
        \langle x, \gamma_m \rangle_Q
    \end{pmatrix} \right)
\end{align*}
$$

(2.11)

where $\tilde{c} = (\tilde{c}_1, \ldots, \tilde{c}_n)$, $\tilde{\theta} = (\tilde{\theta}_1, \ldots, \tilde{\theta}_m)$, and $I$ is the $m \times m$ identity.

Substituting (2.11) into (2.10) gives

$$
    x_n = \sum_{i=1}^{n} \tilde{c}_i n_{t_i} + \sum_{j=1}^{m} \tilde{\theta}_j \gamma_j + P_{S^\delta}
$$

where $P_{S^\delta}$ is the projection of $\delta$ onto span $(V_n \oplus S_m)$. Since $S_m$ may not be orthogonal to $V_n$, $P_{S^\delta}$ will have a decomposition $P_{S^\delta} = \delta_0^\delta + \delta_1^\delta$. 

where $\delta_0 \in S_m$ and $\delta_1 \in V_n$. Thus

$$\sum_{i=1}^{n} c_i \eta_{i} \cdot \sum_{i=1}^{n} c_i \eta_{i} = \sum_{i=1}^{n} c_i \eta_{i} - P_V y = \delta_1,$$

and we want to show that $||\delta_1||_R \to 0$. But the angle between $V_n$ and $S_m$ is bounded from below by the angle between $V$ and $S_m$, and this entails the existence of a constant good for all $n$ such that $||\delta_1||_R \leq \text{const} ||P_S \delta||_R \leq \text{const} ||y-P_V y||_R \to 0$. Thus we have completed the proof of (2.5).

Equation (2.5) says $(\rho_n, Q_t)_{L_2} \to (\rho, Q_t)_{L_2}$, all $t \in [0,1]$. Since $Q$ is of full rank and $H_0$ is dense in $L_2[0,1]$, this entails that $||\rho_n - \rho||_{L_2} \to 0$. It remains to show that $\rho_n \to \rho$ in $L_2$ entails that

$$t \int_0^t [\rho_n^2(u)\alpha(u)]^{1/(2q+1)} du \to t \int_0^t [\rho^2(u)\alpha(u)]^{1/(2q+1)} du.$$

Letting $\rho = \rho_n + \varepsilon_n$ and using the fact that

$$|\rho|^{2/(2q+1)} - |\varepsilon_n|^{2/(2q+1)} \leq |\rho_n|^{2/(2q+1)} \leq |\rho|^{2/(2q+1)} + |\varepsilon_n|^{2/(2q+1)}$$

and a Minkowski inequality gives the theorem.

We remark that we may compare the efficiency of a uniform mesh ($f(t) = 1, F(t) = t$) with the optimum mesh determined by $F^*$, $F^*(t) = \int_0^t f^*(s)ds$, by looking at the ratio

$$\frac{1}{\int_0^1 [\rho^2(s)\alpha(s)]^{1/2q} ds} \frac{1}{\int_0^1 [\rho^2(s)\alpha(s)]^{2q+1} ds}.$$

Clearly, if $\rho^2(s)\alpha(s) = \text{const}$, this ratio is one, and the uniform mesh is optimum. The greater the difference between the "geometric" mean
and the arithmetic mean of $\rho^2(s)a(s)$, the greater the benefit of obtaining an optimum mesh.

The iterative determination of the mesh may be repeated any number of times, but the tradeoff between the cost of iteration and increased accuracy will depend on the problem. If data are determined experimentally and are costly to obtain then a multistage procedure becomes more attractive.

3. Numerical Results

Numerical results appear in [2]. For solving the problem the second time with the new mesh, an equivalent B-spline basis is used for S. See [2,3]. The details of the B-spline formulation, considerations in the selection of $H_R$ as well as other computational parameters and certain numerical comparisons may be found in [2]. A summary of numerical results for the problem

$$x''(t) + 10t x(t) = (-\pi^2+10t) \sin \pi t$$

$$x(0) = x(1) = 0$$

is given in Table 1. Since the actual solution $x(t) = \sin \pi t$ is known the maximum error using the optimized as well as a uniform mesh can be computed and are tabulated in Table 1.
<table>
<thead>
<tr>
<th>$r$</th>
<th>$n$</th>
<th>$N$</th>
<th>APPROXIMATELY OPTIMUM MESH-ERROR</th>
<th>UNIFORM MESH-ERROR</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>10</td>
<td>15</td>
<td>.35-5</td>
<td>.31-4</td>
</tr>
<tr>
<td></td>
<td></td>
<td>25</td>
<td>.23-6</td>
<td>.28-5</td>
</tr>
<tr>
<td></td>
<td></td>
<td>35</td>
<td>.50-7</td>
<td>.57-6</td>
</tr>
<tr>
<td></td>
<td></td>
<td>45</td>
<td>.13-7</td>
<td>.16-6</td>
</tr>
<tr>
<td></td>
<td>20</td>
<td>15</td>
<td>.63-6</td>
<td>.31-4</td>
</tr>
<tr>
<td></td>
<td></td>
<td>25</td>
<td>.78-7</td>
<td>.28-5</td>
</tr>
<tr>
<td></td>
<td></td>
<td>35</td>
<td>.12-7</td>
<td>.57-6</td>
</tr>
<tr>
<td></td>
<td></td>
<td>45</td>
<td>.87-7</td>
<td>.16-6</td>
</tr>
<tr>
<td>6</td>
<td>10</td>
<td>15</td>
<td>.23-6</td>
<td>.22-5</td>
</tr>
<tr>
<td></td>
<td></td>
<td>25</td>
<td>.67-8</td>
<td>.91-7</td>
</tr>
<tr>
<td></td>
<td>20</td>
<td>15</td>
<td>.80-7</td>
<td>.22-5</td>
</tr>
<tr>
<td></td>
<td></td>
<td>25</td>
<td>.17-7</td>
<td>.91-7</td>
</tr>
</tbody>
</table>
4. Application to Design of Experiments for More General Linear Operator Equations

The iterative, or sequential procedure for choosing a mesh clearly has more general application than to the solution of BVP's. Consider for example the Fredholm integral equation of the first kind

\[ g(t) = \int_0^1 K(t,s)x(s)ds \]  \hspace{1cm} (4.1)

arising in many experimental situations. If \( x \in H_R \), then \( y \in H_Q \)

where \( Q(s,t) = \int \int \int_{0}^{1} K(t,u)R(u,v)K(s,v)du dv \) (See [9]). If \( x_N \) is that element of minimal \( H_R \) norm satisfying \( (Kx_N)(t) = g(t) \) for \( t \in T_N \), and \( K \) is of full rank, then \( \|x-x_N\|_R = \|g-P_{T_N}g\|_Q \) and the procedure for choosing \( T_N \) proceeds as before. However first kind equations are better solved by e.g. regularization than by collocation for reasons noted by many people (see [13] for details). Provided that \( K \) is not "too" compact (as an operator from \( L_2 \) to \( L_2 \)) and data points are expensive to obtain (as frequently happens in this context) an iterative procedure for choosing additional points may well turn out to be important. For a discussion of the experimental design problem when \( g \) is observed with noise, and regularization is used to solve (4.1), see [14].

5. Remarks

We note that the approach here contrasts with that of deBoor and Swartz [4]. They suppose \( x \) has \( 2q+m \) continuous derivatives (their \( k \) is our \( q \)), they use local piecewise polynomial approximating functions of degree \( m+q-1 \) to approximate \( x \) by collocation and obtain pointwise \( O(N^{-2q}) \) convergence rates at certain special points when
certain collocation points are zeroes of the $q$th Legendre polynomial in subintervals. The approximations to $x$ here are piecewise polynomials of degree $2r-1 = 2q+2m-1$. Here pointwise convergence rates of $O(N^{2q})$ hold uniformly over $[0,1]$ for any mesh determined by a nice $F$ and $y$ satisfies (1.11), which is equivalent to $x \in W^{(2q+m)}$ and $x$ satisfies some further boundary conditions. The convergence proof has been given in [10, Theorem 3]. The proof in [10] essentially uses $|x(t)-x_N(t)| = |<x-x_N,R_t-R_{tN}>| \leq ||x-x_N||_R ||R_t-R_{tN}||_R$ where $R_{tN}$ is the projection of $R_t$ onto $S_m \oplus V_N$, and $||x-x_N||_R$ and $||R_t-R_{tN}||_R$ are each $O(N^{-q})$.

We note that this optimal mesh problem is formally an optimal quadrature problem, as follows. If

$$g(t) = \int_0^1 Q(t,s)\rho(s)ds$$

then $g$ is a representor of weighted integration in $H_Q$,

$$<g,u> = \int_0^1 \rho(s)u(s)ds, \quad u \in H_Q$$

Letting $P_N g = \sum_{t=1}^N w_i Q_t$, then a quadrature formula is obtained since

$$<P_N g,u> = \sum_{t=1}^N w_i Q_{t,i} u(t_i) = \sum_{t=1}^N w_i u(t_i) .$$

An error bound is

$$\frac{1}{\int_0^1 \rho(s)u(s)ds} - \sum_{t=1}^N w_i u(t_i) = |<g-P_N g,u>| \leq ||g-P_N g||_Q ||u||_Q .$$
References


