Department of Statistics

The University of Wisconsin
Madison, Wisconsin

UWIS-DS-69-217

TECHNICAL REPORT NO. 217

October 1969

ON THE NUMERICAL SOLUTION OF FREDHOLM INTEGRAL EQUATIONS OF THE FIRST KIND

by

Grace Wahba

University of Wisconsin

Typist: Wanda R. Gray

This research was partially supported by the Mathematics Research Center under Contract DA-31-124-ARO-D-462 and by the Air Force Office of Scientific Research Office of Aerospace Research, United States Air Force under AFOSR Grant No. AF-AFOSR-69-1803.

ON THE NUMERICAL SOLUTION OF FREDHOLM INTEGRAL EQUATIONS OF THE FIRST KIND

by

Grace Wahba

1. Introduction

1.1 Statement and Some History of the Problem

We are interested in the numerical solution of a Fredholm integral equation of the first kind, namely

$$u(t) = \int_S K(t,s)\, z(s)\,ds \quad t \in T \tag{1.1}$$

where S, T are intervals, $K(t,s)$ is a given kernel on $T \times S$ with appropriate properties and $u(t)$ is given for $t = t_1, t_2, \ldots, t_n$, $t_i \in T$.

It has been noted by a number of authors (see for example [7], [11], [14]) that replacing the integral in (1.1) by a quadrature approximation and inverting the resulting matrix does not give satisfactory results. No matter how large n is, there are many functions z, including highly oscillatory ones, such that

$$u(t_i) = \int_S K(t_i,s) z(s)\,ds, \quad i = 1,2,\ldots,n \tag{1.2}$$

A discrete approximation of one of these functions is obtained by the matrix inversion

This research was partially supported by the Mathematics Research Center under Contract DA-31-124-AFO-D-462 and by the Air Force Office of Scientific Research Office of Aerospace Research, United States Air Force under AFOSR Grant No. AF-AFOSR-69-1803.

technique. Which one is actually obtained, apparently depends more on the quadrature procedure than anything else. Two methods in the literature seem to have resulted in satisfactory numerical examples. The first method, called the method of regularization of Tihonov [12] [13] and studied experimentally by Tihonov and Glasko [14], goes as follows:

Let

$$M_m^{\alpha}(z,u) = \int_T (\hat{u}(t)-u(t))^2 dt + \alpha \int_S (L_m z(s))^2 ds \qquad (1.3a)$$

where u is a given function in $L_2(T)$, $\hat{u}$, depending on z is given by

$$\hat{u}(t) = \int_S K(t,s)z(s)ds, \qquad (1.3b)$$

$\alpha > 0$, and L_m is an m^{th} order linear differential operator with continuous positive coefficients. Assume $K(t,s)$ is continuous, and that

$$0 = \int_S K(t,s)z(s)ds, \quad t\varepsilon T \qquad (1.4)$$

implies that $z(s) \equiv 0$, $s\varepsilon S$.

Then, it is shown in [13], that, for every fixed $u\varepsilon L_2[T]$, and every $\alpha > 0$, there exists a unique 2m times differentiable function $\hat{z} = \hat{z}_{\alpha}$ which minimizes $M_m^{\alpha}(z,u)$.

Tihonov and Glasko [14] provide an argument that L_m defined by

$$L_m z = z' \tag{1.5}$$

fits into the general theory if $K > 0$ on $T \times S$. They then experiment with a numerical algorithm based on finding $\hat{z}$ to minimize $M_m^\alpha(z,u)$ as follows. Quadrature points $\{s_i\}_{i=1}^k$, $s_i \varepsilon S$, and $\{t_j\}_{j=1}^n$, $t_j \varepsilon T$ are chosen. Let t_o be the left boundary of T, define

$$\bar{z} = (\bar{z}_1, \bar{z}_2, \ldots, \bar{z}_k)'$$

$$\bar{u} = (u(t_1), u(t_2), \ldots, u(t_n))'$$

and $\bar{z}_o = 0 \cdot \bar{u}$ is a given vector. Define $\bar{M}_1^\alpha(\bar{z},\bar{u})$ as

$$\bar{M}_1^\alpha(\bar{z},\bar{u}) = \sum_{j=1}^{n} d_j \left(\sum_{i=1}^{k} K(t_j,s_i)\sigma_i \bar{z}_i - u(t_j) \right)^2$$

$$+ \alpha \sum_{i=0}^{k-1} \frac{(\bar{z}_{i+1} - \bar{z}_i)^2}{(s_{i+1} - s_i)^2} \tag{1.6}$$

where $\{d_j\}_{j=1}^n$, $d_j > 0$, and $\{\sigma_i\}_{i=1}^k$ are appropriately chosen quadrature coefficients. $\bar{M}_1^\alpha(\bar{z},\bar{u})$ is to be thought of as a discretized version of $M_1^\alpha(z,u)$ of (1.3) with L_m given by

(1.5). Let the $n \times k$ matrix $\bar{K}$, the $n \times n$ (diagonal) matrix D and the $k \times k$ symmetric matrix Q be defined by setting

$$\bar{M}_1^\alpha(\bar{z},\bar{u}) = (\bar{K}\bar{z}-\bar{u})'D(\bar{K}\bar{z}-u) + \alpha\bar{z}'Q\bar{z} \tag{1.7}$$

and matching the coefficients of $\bar{z}$, $\bar{u}$ and $\alpha\bar{z}$ in (1.6) and (1.7). For a given $\bar{u}$, the vector $\hat{\bar{z}}$ which minimizes $\bar{M}_1^\alpha(\bar{z},\bar{u})$ is well known to be given by

$$\hat{\bar{z}} = (\bar{K}'D\bar{K}+\alpha Q)^{-1}\bar{K}'D\bar{u} \tag{1.8}$$

It can be verified by direct manipulations upon the matrices involved that another formula for $\hat{\bar{z}}$ is

$$\hat{\bar{z}} = Q^{-1}\bar{K}^1(\bar{K}Q^{-1}\bar{K}'+\alpha D^{-1})^{-1}\bar{u} \tag{1.9}$$

Tihonov and Glasko compute $\hat{\bar{z}}$ from (1.8) as a discrete approximation to the function $\hat{z}_\alpha(s)$ which minimizes $M_1^\alpha(z,u)$. They do this experimentally, for several values of α, by beginning with a $\bar{u}$ obtained from a known function u(t) satisfying

$$u(t) = \int_S K(t,s)\, z^*(s)ds$$

where z^* is a given (smooth) experimental function. For certain values of α, the results are "good", i.e. the components $\hat{\bar{z}}_i$ of $\hat{\bar{z}}$ satisfy

$$\hat{\bar{z}}_i \approx z^*(s_i).$$

More recently, Ribiere [8] has also studied the method of regularization. In both [8] and [14] the experimental results suggest that there is an optimum choice of α. However, a theory making precise the optimal choice is apparently not available.

The second method, discussed by Strand and Westwater [11] and called "statistical estimation" of the solution, is as follows: Let the $n \times k$ matrix $\bar{K}$ be defined as before. Let $Z = (Z_1, Z_2, \ldots, Z_k)'$ and $\epsilon = (\epsilon_1, \epsilon_2, \ldots, \epsilon_n)'$ be normally distributed zero mean random vectors with (prior) covariance

$$EZZ' = Q^{-1}$$

$$E\epsilon\epsilon' = \alpha D^{-1}$$

$$EZ\epsilon' = 0$$

Let $U = (U_1, U_2, \ldots, U_n)$ be a normally distributed random vector defined by

$$U = \bar{K}Z + \epsilon \tag{1.10}$$

It is well known that

$$E(Z|U=\bar{u}) = Q^{-1}\bar{K}'(\bar{K}Q^{-1}\bar{K}'+\alpha D^{-1})^{-1}\bar{u} \qquad (1.11)$$

Strand and Westwater replace the equation (1.1) with the model (1.10), where the unknown vector $\bar{z}$ is replaced by the random vector Z with prior distribution $N(0,Q^{-1})$, and ϵ is to be thought of as a "noise" vector in dependent of Z and having the distribution $N(0,\alpha D^{-1})$. The estimate for the "random" vector Z given the "data" $\bar{u}$ is taken as the right hand side of (1.9). Hence the method of regularization and that of statistical estimation are in practice the same for appropriate choice of D and Q. Strand and Westwater performed numerical experiments (again beginning with a smooth known solution z^*) analogous to those of Tihonov and Glasko, and again obtained "good" results.

1.2 Purpose of this note and basic assumptions

Both of these methods can be embedded in the general theory of the approximation of continuous linear functionals in a reproducing kernel Hilbert space. The overall purpose of this note is to demonstrate this statement in some considerable practical detail. As a byproduct of this theory we obtain a rationale for choosing α, to minimize a certain error bound. We also obtain pointwise error bounds on the solution involving a Hilbert space norm of the (unknown) solution z. These results provide a criteria for the selection of a good Hilbert space in which to operate. Several practical examples are given.

We assume the function z to be an element of a reproducing kernel Hilbert space $\mathcal{H}$ of real-valued functions defined on S. This means that the linear functionals N_s, $s \in S$ defined by

$$N_s z = z(s) \quad , \quad z \in \mathcal{H} \tag{1.12}$$

are continuous for every $s \in S$. In this case, for $\forall s \in S$ there exists an element $\delta_s \in \mathcal{H}$ for which

$$\langle \delta_s, z \rangle = z(s) \quad , \quad \forall z \in \mathcal{H}. \tag{1.13}$$

If we define the kernel R(s,s') on S × S' as

$$\langle \delta_s, \delta_{s'} \rangle = R(s,s') \tag{1.14}$$

then R(s,s) is positive definite. Let R_{s_o} be that function defined on S whose value at s is given by

$$R_{s_o}(s) = R(s_o, s) \tag{1.15}$$

By the Moore-Aronszajn Theorem (see [2]), to every positive definite kernel R(s,s') defined on S × S there corresponds a unique Hilbert space $\mathcal{H} = \mathcal{H}_R$ with the following properties:

1) $R_{s_o} \in \mathcal{H}$ $\qquad \forall s_o \in S$

2) $\langle z, R_{s_o} \rangle = z(s_o)$, $\quad z \in \mathcal{H}$, $s_o \in S$

Hence δ_s of (1.13) is given by

$$\delta_s = R_s \tag{1.16}$$

and

$$\langle R_s, R_{s'} \rangle = R(s,s') \tag{1.17}$$

Equation (1.17) is the source of the terminology "reproducing kernel". The elements $\{R_s, s \in S\}$ clearly span $\mathcal{H}$.

We make sufficient assumptions on K to ensure that the linear functionals Λ_t defined on $\mathcal{H}_R$ by

$$\Lambda_t z = \int_S K(t,s) z(s) ds \quad , \quad z \in \mathcal{H}_R \tag{1.18}$$

are continuous for every fixed $t \in T$. The problem of estimating $z(s)$ for a particular $s \in S$ given that $z \in \mathcal{H}_R$ and the information of (1.1), namely, $\Lambda_{t_i} z = u(t_i)$, $i = 1,2,\ldots,n$ may then be viewed as that of approximating the continuous linear functional N_s defined by (1.12) by the continuous linear functionals $\{\Lambda_{t_i}\}_{i=1}^n$ defined by (1.18). There is a growing body of literature on the approximation of one continuous linear functional by several others (see [3] [4], [9]) familiar to these working with splines. We actually consider the cases where $u(t_i)$ are known only up to some error (experimental error) or the values $\eta_{t_i}(s)$ of the representers η_{t_i} of the continuous linear functionals Λ_{t_i} are known only imperfectly (quadrature error). We show that these cases both lead in a natural way to the same family of algorithms, which contains the method (s) of regularization and statistical estimation.

To discuss the problem practically we must know when Λ_t defined by (1.18) is continuous, and how to finds its representer η_t. The answer is given by Theorem (1.1) and its corollary which will conclude section 1.

Section 2 provides statements of all the Hilbert space lemmas that are used in the sequel. They are generally well known and may be proved by elementary methods, no proofs are provided. Section 3 provides examples of some reproducing kernel Hilbert spaces. Concrete examples are fairly hard to come by. Section 4 presents the main theorems concerning approximations to the solution z and their properties. For concreteness sake, but without loss of generality, the theorems are stated with respect to a particular reproducing kernel Hilbert space which is related to the method of regularization as discussed in [12], [13], [14]. Section 5 discusses the introduction of quadrature error and appropriate ways of dealing with it, and a rationale for choosing α is given. In Section 6, the precise relationship between statistical estimation, and the approximation of continuous linear functionals is described, and it is shown how in general one obtains the same numerical result from the two approaches.

We conclude this Section with Theorem 1.1 and its corollary.

1.3 A Preliminary Theorem

Theorem 1.1. Let Λ be a continuous linear functional on $\mathcal{H}_R$, and let $\psi \varepsilon \mathcal{H}_R$ be defined by

$$\Lambda z = \langle \psi, z \rangle \quad , \quad z \varepsilon \mathcal{H}_R \tag{1.19}$$

Then $\psi(s)$, $s \varepsilon S$ is given by the formula:

$$\psi(s) = \Lambda R_s \tag{1.20}$$

Conversely, let $\tilde{\Lambda}$ be a linear functional defined on $\mathcal{H}_R$ and suppose that the function $\tilde{\psi}$ defined in S by

$$\tilde{\psi}(s) = \tilde{\Lambda} R_s \tag{1.21}$$

satisfies $\tilde{\psi} \varepsilon \mathcal{H}_R$, then $\tilde{\Lambda}$ is continuous.

Proof. Since $\psi(s) \varepsilon \mathcal{H}_R$, $\psi(s) = \langle \psi, R_s \rangle = \Lambda R_s$. Conversely, if $\tilde{\psi} \varepsilon \mathcal{H}_R$, then the continuous linear functional $\tilde{\tilde{\Lambda}}$ defined by

$$\tilde{\tilde{\Lambda}} z = \langle\tilde{\psi}, z\rangle \tag{1.22}$$

coincides with $\tilde{\Lambda}$ on the span of $\{R_s, s\varepsilon S\}$, since $\tilde{\tilde{\Lambda}} R_s = \tilde{\psi}(s)$. But the span of $\{R_s, s\varepsilon S\}$ is $\mathcal{H}_R$.

Corollary. Let K be such that

$$\int_S \int_S K(t,s)R(s,s')K(t,s')dsds' = \theta^2(t) \tag{1.23}$$

is well defined and finite as a Riemann integral for each fixed $t\varepsilon T$. Then the linear functional Λ_t defined by

$$\Lambda_t z = \int_S K(t,u)z(u)du \ , \quad z\varepsilon \mathcal{H}_R \tag{1.24}$$

is continuous on $\mathcal{H}_R$ and

$$\Lambda_t z = \langle\eta_t, z\rangle, \quad t\varepsilon T, \quad z\varepsilon \mathcal{H}_R$$

where $\eta_t \ \varepsilon \mathcal{H}_R$ is defined by

$$\eta_t(s) = \Lambda_t R_s = \int_S K(t,u)R_s(u)du \ , \tag{1.25}$$

and $||\Lambda_t||^2 = \theta^2(t)$.

Proof. The hypotheses on K imply that there exists (for each fixed t), a triangular array

$$
\begin{array}{l}
s_{11} \\
s_{12}, s_{22} \\
\vdots \\
s_{1k}, s_{2k}, \ldots, s_{kk} \\
\vdots
\end{array}
$$

such that the Riemann sums

$$\sum_{i=1}^{k-1} \sum_{j=1}^{\ell-1} K(t,s_{ik})R(s_{ik},s_{j\ell})K(t,s_{j\ell})(s_{i+1,k}-s_{ik})(s_{j+1,\ell}-s_{j\ell})$$

converge to $\theta^2(t)$, as $k, \ell \to \infty$. But then, using (1.17), it follows that the sequence $\eta_t^{(k)}$, $k = 1,2,\ldots$ defined by

$$\eta_t^{(k)} = \sum_{i=1}^{k-1} K(t,s_{ik})(s_{i+1,k}-s_{ik})R_{s_{ik}} \qquad k = 1,2,\ldots$$

is a Cauchy sequence in $\mathcal{H}_R$, $||\eta_t^{(k)}||^2 \to \theta^2(t)$, and $\eta_t^{(k)}$ converges pointwise (in s) to η_t given by (1.25), which is therefore in $\mathcal{H}_R$.

2. Hilbert Space Lemmas

In the lemmas of this section we have the following: $m \leq n$, $\mathcal{H}$ is a Hilbert space,

$$\mathcal{H} = \mathcal{H}_0 \oplus \mathcal{H}_1$$

where $\mathcal{H}_0$ is an m-dimensional subspace, $\mathcal{H}_1 = \mathcal{H}_0^{\perp}$ and P_0 and P_1 are the projection operators onto $\mathcal{H}_0$ and $\mathcal{H}_1$ respectively. $\{\eta_i\}_{i=1}^n$ are, in each lemma, n elements in satisfying the following two conditions

(i) $\{P_0\eta_i\}_{i=1}^n$ span $\mathcal{H}_0$

(ii) $\{P_1\eta_i\}_{i=1}^n$ are linearly independent.

Let

$$P_1\eta_i = \xi_i \ , \quad i = 1,2,\ldots,n$$

let $\{\phi_\nu\}_{\nu=1}^m$ be a (fixed) orthonormal basis for $\mathcal{H}_0$, and let X and Σ be the m × n and n × n matrices with entries defined by

$$[X]_{\mu i} = \langle\phi_\mu,\eta_i\rangle \ , \quad \mu = 1,2\ldots m \qquad (2.1)$$

$$x = 1,2\ldots n \ .$$

$$[\Sigma]_{ij} = \langle\xi_i,\xi_j\rangle$$

Conditions (i) and (ii) guarantee that X and Σ are of full rank. We let ϕ, ξ and η be vectors of m, n and n elements of $\mathcal{H}$ respectively, given by

$$\phi = (\phi_1,\phi_2,\ldots\phi_m)$$

$$\xi = (\xi_1,\xi_2,\ldots\xi_n)$$

$$\eta = (\eta_1,\eta_2,\ldots\eta_n)$$

and $\bar{u}$, be a given vector of real numbers,

$$\bar{u} = (u_1, u_2, \ldots u_n).$$

Lemma 2.1. The solution to the problem: Find $z \varepsilon \mathcal{H}$ to minimize $\langle P_1 z, P_1 z \rangle$ subject to $\langle z, \eta_i \rangle = u_i$, $i = 1,2,\ldots n$ is unique, and is given by $\hat{z}$,

$$\hat{z} = \phi(X\Sigma^{-1}X')^{-1}X\Sigma^{-1}\bar{u}' + \xi(\Sigma^{-1} - \Sigma^{-1}X'(X\Sigma^{-1}X')^{-1}X\Sigma^{-1})\bar{u}' \ . \quad (2.2)$$

Lemma 2.2. Let $\delta_o \varepsilon \mathcal{H}$ be given. The solution to the problem: Find $y \varepsilon \mathcal{H}$ of the form

$$y = \sum_{i=1}^{n} c_i \eta_i \quad (2.3)$$

where $\bar{c} = (c_1, c_2, \ldots c_n)$ is a vector of real numbers, to minimize

$$||P_1(\delta_o - y)||^2 \quad (2.4)$$

subject to

$$||P_o(\delta_o - y)||^2 = 0 \quad (2.5)$$

is unique and is given by $\hat{\delta}_o$,

$$\hat{\delta}_o = \bar{\phi}(X\Sigma^{-1}X')^{-1}X\Sigma^{-1}\eta' + \bar{\xi}(\Sigma^{-1} - \Sigma^{-1}X'(X\Sigma^{-1}X')^{-1}X\Sigma^{-1})\eta' \tag{2.6}$$

where $\bar{\phi}$ and $\bar{\xi}$ are the vectors of real numbers (depending on δ_o), given by

$$\bar{\phi} = (\langle\delta_o,\phi_1\rangle, \langle\delta_o,\phi_2\rangle, \ldots \langle\delta_o,\phi_m\rangle)$$

$$\bar{\xi} = (\langle\delta_o,\xi_1\rangle, \langle\delta_o,\xi_2\rangle, \ldots \quad \ldots \langle\delta_o,\xi_n\rangle). \tag{2.7}$$

Let the bounded linear operators A and A^* be defined by (2.2) and (2.6) respectively as

$$Az = \hat{z} \quad , \quad \text{if} \quad u_i = \langle z,\eta_i\rangle, \ i = 1,2,\ldots n, z\varepsilon\mathcal{H}$$

$$A^*\delta_o = \hat{\delta}_o \ , \quad \delta_o \varepsilon \mathcal{H}$$

Inspection of (2.2) and (2.6) reveals that A and A^* are each idempotent, and are adjoint to each other, that is:

Lemma 2.3 If z is of the form of the right hand side of (2.2) for some $\bar{u}$, then $z = \hat{z}$.

Lemma 2.4

$$\langle\hat{z},\delta_o\rangle = \langle z,\hat{\delta}_o\rangle = \langle\hat{z},\hat{\delta}_o\rangle, \quad z,\delta_o \varepsilon \mathcal{H} \tag{2.8}$$

Lemma 2.5

$$\langle z-\hat{z},\delta_o\rangle^2 = \langle z-\hat{z},\delta_o-\hat{\delta}_o\rangle^2 \leq ||P_1(z-\hat{z})||\,||P_1(\delta_o-\hat{\delta}_o)||^2$$

$$\leq ||P_1 z||^2\,||P_1(\delta_o-\hat{\delta}_o)||^2 \quad (2.9)$$

Furthermore, a calculation gives

Lemma 2.6

$$||P_1(\delta_o-\hat{\delta}_o)||^2 = \langle P_1\delta_o,P_1\delta_o\rangle - \bar{\xi}\Sigma^{-1}\bar{\xi}'$$

$$+ (\bar{\phi}-\bar{\xi}\Sigma^{-1}X')(X\Sigma^{-1}X')^{-1}(\bar{\phi}-\bar{\xi}\Sigma^{-1}X')^{1}$$

(2.10)

where $\bar{\phi}$ and $\bar{\xi}$ are given by (2.7).

Lemma 2.7 Let $V = \{v_{ij}\}$ be a non-negative definite $n \times n$ matrix with ij^{th} entry of V^{-1} given by v^{ij}. Then the solution to the problem: Find $z \in \mathcal{H}$ to minimize

$$M(z) = \sum_{i,j=1}^{n} (\langle z,\eta_i\rangle-u_i)v^{ij}(\langle z,\eta_j\rangle-u_j) + \langle P_1 z,P_1 z\rangle$$

(2.11)

is unique, and is given by $\tilde{z}$,

$$\tilde{z} = \phi(XS^{-1}X')^{-1}XS^{-1}\bar{u}' + \xi(S^{-1}-S^{-1}X'(XS^{-1}X')^{-1}XS^{-1})\bar{u}'$$

(2.12)

where

$$S = \sum + V. \tag{2.13}$$

Lemma 2.8 Let $\{\varepsilon_i\}_{i=1}^n$ be n elements in $\mathcal{H}$ satisfying $\langle\varepsilon_i,\varepsilon_j\rangle = v_{ij}$, and let $\delta_o \varepsilon \mathcal{H}$ be given. The solution to the problem: Find $y\varepsilon \mathcal{H}$ of the form

$$y = \sum_{i=1}^{n} d_i(\eta_i+\varepsilon_i) \tag{2.14}$$

where $\bar{d} = (d_1,d_2,\ldots d_n)$ is a vector of real numbers to be found, to minimize

$$||\delta_o - \sum_{i=1}^{n} d_i\eta_i||^2 + ||\sum_{i=1}^{n} d_i\varepsilon_i||^2 \tag{2.15}$$

subject to

$$||P_o(\delta_o - \sum_{i=1}^{n} d_i\eta_i)||^2 = 0 \tag{2.16}$$

is unique, and is given by $\tilde{\delta}_o$,

$$\tilde{\delta}_o = \bar{\phi}(XS^{-1}X')^{-1}XS^{-1}(\eta+\varepsilon)' + \bar{\xi}(S^{-1}-S^{-1}S'(XS^{-1}X')^{-1}XS^{-1})(\eta+\varepsilon)' \tag{2.17}$$

where $\bar{\phi}$, $\bar{\xi}$, depending on δ_o are given by (2.7), S is given by (2.13), and the vector of functions $(\eta+\varepsilon)$ is given by

$$\eta+\varepsilon = (\eta_1+\varepsilon_1, \eta_2+\varepsilon_2, \ldots \eta_n+\varepsilon_n) \tag{2.18}$$

Lemma 2.9 Let z satisfy $\langle z, \eta_i+\varepsilon_i\rangle = u_i$, $i = 1,2,\ldots n$, then if $\tilde{z}$ is given by (2.12) and $\tilde{\delta}_o$ by (2.17), then

$$\langle z, \tilde{\delta}_o\rangle = \langle \tilde{z}, \delta_o\rangle \tag{2.19}$$

Lemma 2.10 For $\bar{d} = (d_1, d_2, \ldots d_n)$ any vector of real numbers, and z and $\tilde{z}$ as in Lemma 2.9, $\tilde{\delta}_o$ as in (2.17)

$$\langle z-\tilde{z}, \delta_o\rangle^2 = \langle z, \delta_o-\tilde{\delta}_o\rangle^2 \leq 2[||z||^2(||\delta_o - \sum_{i=1}^{n} d_i\eta_i||^2 + ||\sum_{i=1}^{n} d_i\varepsilon_i||^2)] \tag{2.20}$$

If $\varepsilon_i \varepsilon \mathcal{H}_1$, $i = 1,2,\ldots n$, then $\delta_o - \tilde{\delta}_o \,\varepsilon\, \mathcal{H}_1$, and

$$\langle z-\tilde{z}, \delta_o\rangle^2 \leq 2[||P_1 z||^2(||\delta_o - \sum_{i=1}^{n} d_i\eta_i||^2 + ||\sum_{i=1}^{n} d_i\varepsilon_i||^2)] \tag{2.21}$$

3. Examples of Reproducing Kernel Hilbert Spaces

3.1 Generalities

Let S, Ω each be the real line or a closed subset, let G(s,u) be defined on S × Ω with the property that $G(s,u) \in L_2(\Omega)$ for every fixed s, and suppose further that

$$0 = \int_{\Omega} G(s,u)p(u)du, \quad p\varepsilon L_2(\Omega), \quad s\varepsilon S \tag{3.1}$$

implies that $p = 0$. Then the range of the operator G defined by

$$(Gp)(s) = \int_{\Omega} G(s,u)p(u)du \tag{3.2}$$

is a reproducing kernel Hilbert space of functions defined on S. If we call this space $\mathcal{H}_1$, then it has reproducing kernel R_1 given by

$$R_1(s,s') = \int_{\Omega} G(s,u)G(s',u)du \tag{3.3}$$

with inner product

$$\langle f_1,f_2\rangle = \int_{\Omega} p_1(u)p_2(u)du \tag{3.4a}$$

where

$$f_i(s) = (Gp_i)(s) = \int_{\Omega} G(s,u)p_i(u)du, \quad i = 1,2 \tag{3.5}$$

Suppose $\mathcal{H}_1$ is contained in some larger Hilbert space of functions defined on S, and $\{\phi_\nu\}_{\nu=1}^m$ are m orthonormal functions on S all in $\mathcal{H}_1^{\perp}$. Then $\mathcal{H}_o$, the m dimensional space spanned by $\{\phi_\nu\}_{\nu=1}^m$ is a reproducing kernel space with reproducing kernel R_o given by

$$R_o(s,s') = \sum_{\nu}^{m} \phi_\nu(s)\phi_\nu(s') \qquad (3.6)$$

and $\mathcal{H} = \mathcal{H}_o \oplus \mathcal{H}_1$ is a reproducing kernel space with reproducing kernel R given by

$$R(s,s') = R_o(s,s') + R_1(s,s') \qquad (3.7)$$

3.2 Reproducing Kernel Spaces Associated with Green's Functions

Let L_m be an m^{th} order linear differential operator on [0,1] with m-dimensional null space. Let the linear functionals M_μ be defined by

$$M_\mu f = \sum_{\nu=1}^{m} \gamma_{\mu\nu} f^{(\nu-1)}(0), \quad \mu = 1,2,...m \qquad (3.8)$$

where the matrix $\{\gamma_{\mu\nu}\}$ is non-singular. Let G(s,u) be the Green's function for the equation $L_m f = g$, $M_\mu f = 0$, $\mu = 1,2,...m$, and let $\{\phi_\nu\}_{\nu=1}^{m}$ span the null space of L_m with

$$M_\mu \phi_\nu = \delta_{\mu\nu} \qquad (3.9)$$

where $\delta_{\mu\nu}$ is the Kronecker δ. We may take $\mathcal{H}_1$ as the collection $\{f: L_m f \varepsilon L_2[0,1], M_\mu f = 0, \mu = 1,2...m,\}$ and $\mathcal{H}_0$

the space spanned by $\{\phi_\nu\}_{\nu=1}^m$. $\mathcal{H} = \mathcal{H}_0 \oplus \mathcal{H}_1$ then has inner product

$$\langle f_1, f_2 \rangle = \sum_{\nu=1}^{m} (M_\nu f_1)(M_\nu f_2) + \int_0^1 (L_m f_1)_s (L_m f_2)_s ds \qquad (3.10a)$$

and

$$||P_1 f||^2 = \int_0^1 (L_m f)_s^2 ds \; . \qquad (3.10b)$$

Two simple examples are:

a) $$L_m = \frac{d^m}{ds^m}$$

$$G(s,u) = \frac{(s-u)_+^{m-1}}{(m-1)!} \qquad (3.11)$$

$$\phi_\nu(s) = \frac{s^{\nu-1}}{(\nu-1)!}, \quad \nu = 1,2 \ldots m$$

$$\gamma_{\mu\nu} = 1, \; \mu = \nu, \; = 0 \text{ otherwise}$$

b) $$L_m = \prod_{\nu=1}^{m} (D+\alpha_\nu), \quad \{\alpha_\nu\}_{\nu=1}^m \text{ distinct positive real numbers}$$

$$G(s,u) = \sum_{\nu=1}^{m} c_\nu e^{-\alpha_\nu (s-u)} \quad s \geq u, \quad c_\nu^{-1} = \prod_{j \neq \nu} (\alpha_j - \alpha_\nu)$$

$$= 0 \qquad \text{otherwise} \qquad (3.12)$$

$$\phi_\nu(s) = c_\nu e^{-\alpha_\nu s}$$

$$M_\nu f = \prod_{\ell \neq \nu} (D+\alpha_\ell)$$

where $(D+\alpha_\nu)f = f' + \alpha_\nu f$. This type of reproducing kernel space may be used to construct spline functions. (See [3])

3.3 Spaces of Functions with Rapidly Decreasing Fourier Transform

Let $W(\lambda)$ be a complex-valued Hermitian function of λ on $(-\infty,\infty)$ with $W(\lambda)\varepsilon L_2(-\infty,\infty)$ and $|W(\lambda)| > 0$, all real λ. Let $G(\tau)$ be the inverse Fourier-Plancherel Transform of $W(\lambda)$ which we write as

$$G(\tau) = \int_{\infty}^{\infty} e^{-i\tau\lambda}\, W(\lambda)d\lambda \quad . \tag{3.13}$$

Let

$$R(s,s') = \int_{-\infty}^{\infty} G(s-u)G(s'-u)du = \int_{-\infty}^{\infty} e^{-i(s-t)\lambda} W(\lambda)W^*(\lambda)d\lambda \tag{3.14}$$

$R(s,s')$ is positive definite, and, since $|W(\lambda)| > 0$,

$$0 = \int_{-\infty}^{\infty} G(s-u)p(u)du, \quad -\infty < s < \infty,\ p\varepsilon L_2 \tag{3.15}$$

implies $p(u) = 0$. $\mathcal{H}$, the reproducing kernel space for $R(s,s')$ is the collection of all functions f of the form

$$f(s) = \int_{-\infty}^{\infty} G(s-u)p(u)du \quad p\varepsilon L_2(-\infty,\infty) \tag{3.16}$$

with inner product (by Parseval's theorem), given by

$$\langle f_1, f_2 \rangle = \int_{-\infty}^{\infty} \frac{F_1(\lambda) F_2^*(\lambda)}{W(\lambda) W^*(\lambda)} d\lambda \tag{3.17}$$

where F_i is the Fourier-Plancherel Transform of f_i. $\mathcal{H}$ contains only functions whose Fourier transforms $F(\lambda)$ decrease sufficiently rapidly that $F(\lambda)/W(\lambda) \varepsilon L_2(-\infty,\infty)$.

For a concrete example, let q, p be non-negative integers with $q < p$, let $\{\alpha_\mu\}_{\mu=1}^{p}$, $\{\beta_\nu\}_{\nu=1}^{q}$ be p + q distinct positive real numbers, let

$$\begin{aligned} P(\lambda) &= \prod_{\mu=1}^{p} (i\lambda - \alpha_\mu) \\ Q(\lambda) &= \prod_{\nu=1}^{q} (i\lambda - \beta_\nu) \\ W(\lambda) &= Q(\lambda)/P(\lambda). \end{aligned} \tag{3.18}$$

We have

$$\begin{aligned} G(\tau) &= \sum_{\mu=1}^{p} c_\mu e^{-\alpha_\mu \tau}, \quad \tau \geq 0 \\ &= 0 \end{aligned} \tag{3.19}$$

where

$$c_\mu = \prod_{k=1}^{q} (\beta_k - \alpha_\mu) / \prod_{j \neq \mu} (\alpha_j - \alpha_\mu)$$

and

$$R(s,s') = \sum_{\nu=1}^{p} \theta_\nu e^{-\alpha_\nu |s-s'|}$$

$$\theta_\nu = \sum_{\mu=1}^{p} \frac{c_\mu c_\nu}{(\alpha_\mu + \alpha_\nu)} \tag{3.20}$$

(for $q = 0$, $p = 1$ we have $R(s,s') = e^{-\alpha|s-s'|}$).

Letting $m = p-q$, $\mathcal{H}$ is here the collection of all functions with absolutely continuous m-1st derivative and square integrable m^{th} derivative. Let

$$G_\Omega(\tau) = \sum_{\nu=1}^{q} d_\nu e^{-\beta_\nu \tau}, \quad \tau \geq 0$$

$$= 0 \tag{3.21}$$

where

$$d_\nu^{-1} = \prod_{j \neq \nu} (\beta_j - \beta_\nu) \ .$$

Let

$$(G_\Omega f)_t = \int_{-\infty}^{\infty} G_\Omega(t-s) f(s) ds \tag{3.22}$$

and let

$$L_p f = \prod_{\mu=1}^{p} (D+\alpha_\mu) f \,. \qquad (3.23)$$

For $f \varepsilon \mathcal{H}$, $(G_Q f)$ has p-1 absolutely continuous and p square integrable derivatives, and the inner product in is given by

$$\langle f_1, f_2 \rangle = \int_{-\infty}^{\infty} (L_p G_Q f_1)_t (L_p G_Q f_2)_t dt. \qquad (3.24)$$

3.4 Spaces of Band Limited Functions

Let $\mathcal{F}_\Lambda$ be the subset of $L_2(-\infty,\infty)$ of functions whose Fourier transforms vanish outside $[-\Lambda,\Lambda]$.

Let

$$G(\tau) = \int_{-\infty}^{\infty} s(\tau-u) w(u) du \qquad (3.25)$$

where

$$s(\tau) = \frac{\sin \Lambda\tau}{\pi\tau}$$

$$(3.26)$$

$$w(u) = \int_{-\Lambda}^{\Lambda} e^{iu\lambda} W(\lambda) d\lambda$$

and $W(\lambda)$ is assumed to be a Hermitian function of λ with no real zeroes in $[-\Lambda,\Lambda]$. The equation

$$0 = \int_{-\infty}^{\infty} G(t-u) p(u) du \qquad -\infty < t < \infty \qquad (3.27)$$

has solutions in $L_2(-\infty,\infty)$ but not in $\mathcal{F}_\Lambda$, since (3.25)-(3.27) imply $W(\lambda)P(\lambda) = 0$, $-\Lambda \leq \lambda \leq \Lambda$, where $P(\lambda)$ is the Fourier-Plancherel transform of $p(u)$. (Note that $S(\lambda)$, the Fourier-Plancherel transform of $\sin \Lambda\tau/\pi\tau$ is $S(\lambda) = 1$, $|\lambda| \leq \Lambda$, $= 0$ otherwise.)

$\mathcal{F}_\Lambda$ is a reproducing kernel Hilbert space with inner product

$$<f_1,f_2> = \int_{-\Lambda}^{\Lambda} \frac{F_1(\lambda)F_2(\lambda)}{W(\lambda)W^*(\lambda)} d\lambda \tag{3.28}$$

where F_i is the Fourier-Plancherel transform of f_i, $i = 1,2$. The reproducing kernel is

$$R(s,s') = \int_{-\Lambda}^{\Lambda} e^{-i\lambda(s-s')} W(\lambda)W^*(\lambda)d\lambda \; . \tag{3.29}$$

If we set

$$W(\lambda)^2 = 1, \quad |\lambda| \leq \Lambda$$
$$= 0 \quad \text{otherwise} \tag{3.30}$$

then

$$R(s,s') = \frac{\sin \Lambda(s-s')}{\pi(s-s')} \tag{3.31}$$

and $\mathcal{F}_\Lambda$ is a Hilbert subspace of $L_2(-\infty,\infty)$. Slepian and Pollack [10] consider the space of bandlimited functions with reproducing kernel (3.31). A restatement of the sampling theorem, tells us that the functions η_j defined by

$$\eta_j(s) = \frac{1}{\sqrt{2\Lambda}} \int_{-\Lambda}^{\Lambda} e^{i(s-2\pi j/\Lambda)} W(\lambda)d\lambda, \quad j = \ldots -1,0,1,\ldots \tag{3.32}$$

are a complete orthonormal basis for $\mathcal{F}_\Lambda$

4. Typical Theorems

A variety of theorems now fall out by applying Theorem 1.1 and the lemmas of Section 2 to the reproducing kernel Hilbert spaces of Section 3. For concreteness only we state them with reference to the example of Section 3.2, since this example provides a direct comparison with the method of regularization as discussed by Tihonov and Glasko.

Thus, let L_m be an m^{th} order linear differential operator on [0,1] with m-dimensional null space spanned by $\{\phi_\nu\}_{\nu=1}^m$, and $\mathcal{H} = \{z: L_m z \varepsilon L_2[0,1], z^{(m-1)} \text{ absolutely continuous}\}$. Let G(s,u) be the Greens function for the equation

$$L_m f = g, \quad f^{(\nu)}(0) = 0, \quad \nu = 0,1,2,\ldots m-1$$ [1]

---

[1] The reader may verify that the choice of the matrix $\{\lambda_{\mu\nu}\}$ in (3.8) is irrelevant to the solution of our problems, we take it as I.

and

$$R_1(s,s') = \int_0^1 G(s,u)G(s',u)\,du \qquad (4.1)$$

$$R(s,s') = \sum_{\nu=1}^{m} \phi_\nu(s)\phi_\nu(s') + \int_0^1 G(s,u)G(s',u)\,du$$

(4.2)

We suppose that $K(t,s)$ satisfies that

$$\int_0^1\int_0^1 K(t_i,s)R(s,s')K(t_i,s')\,ds\,ds' \qquad i = 1,2,\ldots n$$

is well defined and finite as a Riemann integral.

Let η_i ($\eta_i \varepsilon \mathcal{H}$) be defined by

$$\eta_i(s) = \int_0^1 K(t_i,u)R(s,u)\,du \qquad (4.3)$$

Define the elements $\xi_i \ \varepsilon \mathcal{H}$, and the matrices X and Σ by

$$P_i\eta_i = \xi_i, \quad \xi_i(s) = \int_0^1 K(t_i,u)R_1(s,u)\,du, \ i = 1,2,\ldots n,$$

(4.4)

$$X = \{\chi_{\mu i}\}, \ \mu = 1,2,\ldots m, \ i = 1,2,\ldots n$$

$$\Sigma = \{\sigma_{ij}\}, \ i,j = 1,2,\ldots n$$

$$\chi_{\mu i} = \langle \eta_i, \phi_\mu \rangle \quad \int_0^1 K(t_i,u)\phi_\mu(u)du \tag{4.5}$$

$$\sigma_{ij} = \langle \xi_i, \xi_j \rangle = \int_0^1\int_0^1 K(t_i,s)R_1(s,s')K(t_j,s')dsds' \tag{4.6}$$

<u>Theorem 4.1</u> Let $\bar{u} = (u_1, u_2, \ldots u_n)$ be a vector of n real numbers, let $V = \{v_{ij}\}$ be an $n \times n$ symmetric positive definite matrix with $V^{-1} = \{v^{ij}\}$, and suppose that matrices X and Σ defined by (4.5) and (4.6) be of full rank. For $z \varepsilon \mathcal{H}$, define the numbers $\hat{u}_i$ by

$$\hat{u}_i = \hat{u}_i(z) = \langle \eta_i, z \rangle = \int_0^1 K(t_i,s)z(s)ds \tag{4.7}$$

$$i = 1,2,\ldots n$$

Then

(i) There is a unique solution $\hat{z}_\alpha$ to the problem:

Find $z \varepsilon \mathcal{H}$ to minimize

$$M(z) = \sum_{i,j=1}^{n} (\hat{u}_i - u_i)v^{ij}(\hat{u}_j - u_j) + \alpha \int_0^1 (L_m z)_s^2 ds \tag{4.8}$$

given by

$$\hat{z}_\alpha = \phi(X\Gamma_\alpha^{-1}X')^{-1}X\Gamma_\alpha^{-1}\bar{u}' + \xi(\Gamma_\alpha^{-1}-\Gamma_\alpha^{-1}X'(X\Gamma_\alpha^{-1}X')^{-1}X\Gamma_\alpha^{-1})\bar{u}' \tag{4.9}$$

where ϕ and ξ are the vectors of functions

$$\phi = (\phi_1,\phi_2,\ldots\phi_m)$$

$$\xi = (\xi_1,\xi_2,\ldots\xi_n)$$

and

$$\Gamma_\alpha = \sum + \alpha V \tag{4.10}$$

(ii) $\lim_{\alpha\to 0} \hat{z}_\alpha = \hat{z}_o$, given by

$$\hat{z}_o = \phi(X\sum^{-1}X')^{-1}X\sum^{-1}\bar{u} + \xi(\sum^{-1}-\sum^{-1}X'(X\sum^{-1}X')^{-1}X\sum^{-1})\bar{u}' \tag{4.11}$$

is the (unique) solution to the problem: Find $z \varepsilon \mathcal{H}$ to minimize

$$\int_0^1 (L_m z)_s^2 ds$$

subject to $\hat{u}_i(z) = u_i$, $i = 1,2,\ldots n$.

(iii) For any $z \in \mathcal{H}$ satisfying

$$u_i = \int_0^1 K(t_i,s)z(s)ds \qquad i = 1,2,\ldots n \tag{4.12}$$

and $\hat{z}_o$ given by (4.11) we have the pointwise error bounds

$$|\hat{z}_o(s)-z(s)|^2 \leq \sigma^2(s)\{\int_0^1 (L_m z)_s^2\, ds\} \tag{4.13}$$

where

$$\sigma^2(s) = R_1(s,s) - \xi(s)\textstyle\sum^{-1}\xi(s)' \tag{4.14a}$$

$$+ (\bar{\phi}(s)-\bar{\xi}(s)\textstyle\sum^{-1}X')(X\sum^{-1}X')^{-1}(\bar{\phi}(s)-\bar{\xi}(s)\sum^{-1}X')'$$

and

$$\bar{\phi}(s) = (\phi_1(s),\phi_2(s),\ldots\phi_m(s)$$

$$\bar{\xi}(s) = (\xi_1(s),\xi_2(s),\ldots\xi_n(s)) \tag{4.14b}$$

(iv) If $z \in \mathcal{H}$ has the form

$$z(s) = \sum_{\nu=1}^{m} c_\nu\phi_\nu(s) + \sum_{i=1}^{n} d_i\xi_i(s) \tag{4.15}$$

where $\bar{d} = (d_1,d_2,\ldots d_n)$ is any vector satisfying $X\bar{d}' = 0$, and we set u_i in (ii) as

$$u_i = \int_0^1 K(t_i,s)z(s)ds \tag{4.16}$$

then the solution $\hat{z}_o$ in (ii) satisfies

$$\hat{z}_o = z \tag{4.17}$$

Proof: Assertion (i) is Lemma 2.7 and (ii) is Lemma 2.1. The remarkable error bound of (iii) is obtained from Lemmas 2.5 and 2.6 upon setting δ_o as the representer of the continuous linear functional N_s defined by $N_s z = z(s)$, $z \varepsilon \mathcal{H}$, that is, $\delta_o = R_s$. In this case we have

$$|\hat{z}_o(s)-z(s)| = |<\hat{z}_o-z,R_s>|$$

$$<R_s,\phi_\mu> = \phi_\mu(s) \tag{4.18}$$

$$<R_s,\S_i> = \xi_i(s)$$

(iv) is another way of writing Lemma 2.3.

Other minimization problems may be handled within the context of the geometry of reproducing kernel Hilbert space. For example, the solution to the problem: Find $z \varepsilon \mathcal{H}$ to minimize $M(z)$ of (4.8) subject to the linear inequalities

$$a_\ell \leq z(s_\ell) \leq b_\ell \ , \quad \ell = 1,2,...k \ , \quad s_\ell \varepsilon S \tag{4.19}$$

maybe reduced to a standard problem of minimizing a (finite) quadratic form subject to linear inequalities, as follows:

Inequalities (4.19) maybe written

$$a_\ell \le \langle R_{s_\ell}, z\rangle \le b_\ell, \quad \ell = 1,2,\ldots k \qquad (4.20)$$

Then, any solution z to this minimization problem must be of the form

$$z = \sum_{\nu=1}^{m} c_\nu \phi_\nu + \sum_{i=1}^{n} d_i \xi_i + \sum_{\ell=1}^{k} e_\ell R_{s_\ell} \qquad (4.21)$$

for some coefficients $\{c_\nu\}$, $\{d_i\}$ and $\{e_\ell\}$. By substituting (4.21) into (4.20) and (4.8), the minimization problem reduces to the standard problem of minimizing a quatratic form in the coefficients, subject to a set of linear inequalities. This minimization problem may be recognized as a problem in control theory. See, for example [5].

Side conditions which are precisely enough to specify $P_0 z$ give especially simple looking answers and error bounds. For example:

Theorem 4.2. Under the assumptions of Theorem 4.1:

(i) The solution to the problem: find $z \varepsilon \mathcal{H}$ to minimize (4.8) subject to the boundary conditions $\mathcal{B}$

$$\mathcal{B}_\theta: M_\nu z = \langle \phi_\nu, z\rangle = \theta_\nu, \quad \nu = 1,2\ldots m \qquad (4.22)$$

is unique and is given by $\hat{z}_\alpha$,

$$\hat{z}_\alpha = \phi\bar{\theta}' + \xi\Gamma_\alpha^{-1}(\bar{u}-X'\bar{\theta}') \qquad (4.23)$$

where $\bar{\theta}=(\theta_1,\theta_2,\ldots\theta_m)$

ii) $\lim_{\alpha\to 0} \hat{z}_\alpha = \hat{z}_o$, where

$$\hat{z}_o = \phi\bar{\theta}' + \xi\textstyle\sum^{-1}(\bar{u}-X'\bar{\theta}') \qquad (4.24)$$

is the solution to the problem: Find $z\varepsilon\mathcal{H}$ to minimize

$$\int_0^1 (L_m z)_s ds \qquad (4.25)$$

subject to

$$u_i = \int_0^1 K(t_i,s)z(s)ds, \quad i = 1,2,\ldots n \qquad (4.26)$$

$$M_\nu z = \theta_\nu, \quad \nu = 1,2,\ldots m. \qquad (4.22)$$

(iii) If $z\varepsilon\mathcal{H}_0$ satisfies

$$u_i = \int_0^1 K(t_i,s)z(s)ds \qquad (4.27a)$$

and

$$M_\nu z = \theta_\nu, \quad \nu = 1,2,\ldots m \qquad (4.27b)$$

then $\hat{z}_o$ of (4.24) satisfies

$$|z(s)-\hat{z}_o(s)|^2 \leq \underset{\sim}{\sigma}^2(s)\{\int_0^1 (L_m z)_u^2 du\} \tag{4.28a}$$

where

$$\underset{\sim}{\sigma}^2(s) = R_1 \ (s,s) - \bar{\xi}(s)\Sigma^{-1}\bar{\xi}(s)' \tag{4.28b}$$

and $\xi(s)$ is given by (4.14b).

Proof: $z\varepsilon\mathcal{H}$, $M_\nu z = \theta_\nu$, $\nu = 1,2,\ldots m$ imply that

$$P_o z = \sum_{\nu=1}^{m} \theta_\nu \phi_\nu$$

and

$$\int_0^1 K(t_i,s)z(s)ds = \langle\eta_i,z\rangle = \langle\xi_i,P_1 z\rangle + \sum_{\nu=1}^{m} \theta_\nu \chi_{\nu i} \tag{4.29}$$

In this case (4.8) may be written

$$M(z) = \sum_{i,j} (\langle\xi_i,P_1 z\rangle-(u_i-\sum_{\nu=1}^{m}\theta_\nu\chi_{\nu i}))v^{ij}(\langle\xi_j,P_1 z\rangle-(u_j-\sum_{\nu=1}^{m}\theta_\nu\chi_{\nu j}))$$

$$+ \alpha||P_1 z|| \tag{4.30}$$

Since $\{\xi_i\}_{i=1}^n$ and $P_1 z$ are in $\mathcal{H}_1$, we may find $(P_1 z)$ to minimize (4.30) or $(P_1 z)$ to minimize (4.25) subject to

$$u_i - \sum_{\nu=1}^{m} \theta_\nu \chi_{\nu i} = \langle \xi_i, P_1 z \rangle, \quad i = 1,2,\ldots n$$

via Lemmas 2.7 and 2.1 respectively, by setting m = 0, $\mathcal{H} = \mathcal{H}_1$ in these lemmas. To prove (iii) let P_ξ be the projection operator onto the subspace spanned by $\{\xi_i\}_{i=1}^n$. If $z \in \mathcal{H}$ satisfies (4.27a) and (4.27b), then

$$P_\xi z = P_1 \hat{z}_o \tag{4.31a}$$

$$P_o z = P_o \hat{z}_o \tag{4.31b}$$

and hence

$$z - \hat{z}_o = (P_1 - P_\xi) z. \tag{4.32}$$

But

$$\begin{aligned} z(s) - \hat{z}_o(s) &= \langle z - \hat{z}_o, R_s \rangle \\ &= \langle (P_1 - P_\xi) z, R_s \rangle \\ &= \langle (P_1 - P_\xi) z, (P_1 - P_\xi) R_s \rangle \end{aligned} \tag{4.33}$$

Hence

$$|z(s) - \hat{z}_o(s)| \le \|(P_1 - P_\xi) z\| \|(P_1 - P_\xi) R_s\| \le \|P_1 z\| \|(P_1 - P_\xi) R_s\| \tag{4.34}$$

But

$$\underset{\sim}{\sigma}^2(s) = ||(P_1 - P_\xi)R_s||^2 \tag{4.35}$$

A direct comparison with the solution of the methods of regularization and statistical estimation obtains, as follows:

Set $V^{-1} = D$ in (4.8) where D is the diagonal matrix of (1.7), and set $\theta = 0$ in (4.23). Choose $\{s_i\}_{i=1}^k$, $s_i \varepsilon [0,1]$, as in Section 1 and let

$$\bar{u} = (u_1, u_2, \ldots u_n)'$$

be a given vector of real numbers and

$$\bar{z} = (\bar{z}_1, \bar{z}_2, \ldots \bar{z}_k)'$$

where

$$\bar{z}_i = \hat{z}_\alpha(s_i), \quad \hat{z}_\alpha = \xi\Gamma_\alpha^{-1}\bar{u}$$

Then

$$\bar{z} = \Xi(\textstyle\sum + \alpha D^{-1})^{-1}\bar{u} \tag{4.36}$$

where Ξ is the $k \times n$ matrix with i, j^{th} entry

$$\xi_j(s_i) = \int_0^1 K(t_j, u)R_1(s_i, u)du \tag{4.37}$$

and $\sum$ is the $n \times n$ matrix with i, j^{th} entry

$$\langle \xi_i, \xi_j \rangle = \int_0^1 \int_0^1 K(t_i,u) R_1(u,u') K(t_j,u') du du' \qquad (4.38)$$

If $\{\sigma_\ell\}_{\ell=1}^k$ are a set of suitably chosen quadrature coefficients, then we have, for purposes of comparison, approximately

$$\xi_j(s_i) \approx \sum_{\ell=1}^k K(t_j,u_\ell) R_1(s_i,u_\ell)\sigma_\ell \qquad (4.39)$$

$$\langle \xi_i, \xi_j \rangle \approx \sum_{\ell=1}^k \sum_{\ell'=1}^k K(t_i,u_\ell) R_1(u_\ell,u_{\ell'}) K(t_j,u'_\ell)\sigma_\ell \sigma_{\ell'} \qquad (4.40)$$

Let $\bar{K}$ be the $n \times k$ matrix (of (1.7)), with i, j^{th} entry $K(t_j,s_i)\sigma_i$, and $\bar{R}$ be the $k \times k$ matrix with ij^{th} entry $R_1(s_i,s_j)$. Then we may write (4.36) as

$$\bar{z} \approx \bar{R}\bar{K}(\bar{K}\bar{R}\bar{K}'+\alpha D^{-1})^{-1}\bar{u} \qquad (4.41)$$

We may identify $\bar{R}$ with Q^{-1} of section 1 as follows: Suppose $f \in \mathcal{H}_1$, then if $L_m f = g$, we have

$$f(s) = \int_0^1 G(s,u) g(u) du \qquad (4.42)$$

and

$$\int_0^1 (L_m f)^2 ds = \int_0^1 g^2(s)ds \tag{4.43}$$

Let $\bar{f} = (f(s_1), f(s_2), \ldots f(s_k))'$, $\bar{g} = (g(s_1), g(s_2), \ldots g(s_k))'$, $\bar{G}$ be the k × k matrix with ij^{th} entry $G(s_i, s_j)\sigma_j$ and D_σ be the k × k diagonal matrix with iith entry σ_i.

Since

$$R(s,s') = \int_0^1 G(s,u)G(s',u)du, \tag{4.44}$$

we have

$$\bar{R} \approx \bar{G}D_\sigma^{-1}\bar{G}' \tag{4.45}$$

also

$$\bar{f} \approx \bar{G}g \tag{4.46}$$

$$f'Qf \approx \int_0^1 (L_m f)_s^2 ds = \int_0^1 g^2(s)ds \approx g'D_\sigma g \approx f'G^{-1'}D_\sigma G^{-1}f = f'\bar{R}^{-1}f \tag{4.47}$$

Some questions of convergence may be answered as follows: For simplicity, we consider that the boundary values β_θ of the solution z to (1.1) are known. We may then consider, without loss of generality that z and $\hat{z}_o$ are in $\mathcal{H}_1$.

Let the family of functions $\xi_t \varepsilon \mathcal{H}_1$, $0 \leq t \leq 1$ be defined by

$$\xi_t(s) = \int_0^1 K(t,u)R_1(s,u)du \quad 0 \leq t \leq 1 \tag{4.48}$$

If the family of functions $\{\xi_i,\ 0\leq t\leq 1\}$ span the (separable) space $\mathcal{H}_1$, then $\underset{\sim}{\sigma}^2(s)$ of (4.28), will tend to 0 for each s as the set $\{t_i\}_{i=1}^n$ becomes dense in [0,1], by (4.35). A necessary and sufficient condition that $\{\xi_t,\ 0\leq t\leq 1\}$ span $\mathcal{H}_1$, is that

$$<\xi_t, z> = 0 \quad 0 \leq t \leq 1,\ z\varepsilon \mathcal{H}_1, \tag{4.49}$$

implies that z = 0. But (4.49) may be rewritten

$$\int_0^1 K(t,s)z(s)ds = 0, \quad 0 \leq t \leq 1, \quad z\varepsilon \mathcal{H}_1 \tag{4.50}$$

imples z = 0. There does not seem to be a straight forward general way of establishing a rate at which

$$\underset{\sim}{\sigma}^2(s) = ||(P_1 - P_\xi)R_s||^2 \tag{4.35}$$

tends to zero, as, say $\sup|t_{i+1}-t_i| \to 0$, if indeed such a rate exists. However, results have been obtained regarding the convergence of $||(P_1-P_\xi)\delta||$ when $\delta\varepsilon\mathcal{H}_1$ is "very smooth".

Thus error rates for the pointwise approximation of very smooth z, or for the approximation of continuous linear functionals with very smooth representers, are available. These results will appear separately.

5. The Introduction of Quadrature Formulae and the Choice of α

If the integral

$$\xi_i(s) = \int_0^1 K(t_i,u)R_1(s,u)du \tag{4.4}$$

can be evaluated analytically at values of s for which it is desired to estimate z(s), and $\chi_{\mu i}$ and σ_{ij} of (4.5) and (4.6) are known exactly, and computational and experimental errors are negligible, then it is natural to estimate z(s) by $\hat{z}_o(s)$ of Theorem 4.1 or Theorem 4.2. The purpose of this section is to study the situation where $\{\chi_{\mu i}\}$, $\{\sigma_{ij}\}$ and $\{\xi_i(s),\}_{i=1}^{n}$ must be evaluated by quadratures, where quadrature error is the primary source of error. Let $\{s_k\}_{k=1}^{N}$, $s_k \varepsilon [0,1]$, and $\{\omega_k\}_{k=1}^{N}$ be suitably chosen quadrature points and quadrature coefficients, respectively. We show that this situation leads in a natural way to estimating z by the solution to the problem: Find $z\varepsilon \mathcal{H}$ to minimize

$$\bar{M}(z) = \sum_{i=1}^{n} (u_i - \sum_{k=1}^{N} K(t_i,s_k)\omega_k z(s_k))^2 + \lambda \int_0^1 (L_m z)_s^2 ds \tag{5.1}$$

where λ is chosen to approximate the mean square quadrature error. Define $\hat{\eta}_i \in \mathcal{H}$ by

$$\hat{\eta}_i = \sum_{i=1}^{N} K(t_i,s_k)\omega_k R_{s_k} \qquad i = 1,2,\ldots n \tag{5.2}$$

Letting R_{1,s_k} be that element of $\mathcal{H}_1$ whose value at s is given by $R_{1,s_k}(s) = R_1(s,s_k)$, $k = 1,2,\ldots N$, define $\hat{\xi}_i \in \mathcal{H}_1$ by

$$\hat{\xi}_i = \sum_{k=1}^{N} K(t_i,s_k)\omega_k R_{1,s_k} , \qquad i = 1,2,\ldots n \tag{5.3}$$

Then $P_1\hat{\eta}_i = \hat{\xi}_i$, $i = 1,2,\ldots n$. $\hat{\chi}_{\mu i}$ given by

$$\hat{\chi}_{\mu i} = \langle\hat{\eta}_i,\phi_\mu\rangle = \sum_{k=1}^{N} K(t_i,s_k)\omega_k\phi_\mu(s_k) \tag{5.4}$$

is a quadrature approximation to $\chi_{\mu i}$, $\hat{\sigma}_{ij}$ given by

$$\hat{\sigma}_{ij} = \langle\hat{\xi}_i,\hat{\xi}_j\rangle = \sum_{k=1}^{N}\sum_{\ell=1}^{N} K(t_i,s_k)\omega_k R_1(s_k,s_\ell)K(t_j,s_\ell)\omega_\ell \tag{5.5}$$

is a quadrature approximation to σ_{ij} and $\hat{\xi}_i(s)$ given by

$$\hat{\xi}_i(s) = \sum_{k=1}^{N} K(t_i,s_k)\omega_k R_1(s,s_k) \tag{5.6}$$

is a quadrature approximation to $\xi_i(s)$ of (4.4), for each s.

Define $\varepsilon_i \in \mathcal{H}$ by

$$\varepsilon_i = \eta_i - \hat{\eta}_i \tag{5.7}$$

The problem may now be viewed as that of approximating $z(s) = \langle R_s, z\rangle$ from the information $u_i = \langle \hat{\eta}_i + \varepsilon_i, z\rangle$, $i = 1,2,\ldots n$, or, alternatively approximating R_s by $\{\hat{\eta}_i + \varepsilon_i\}_{i=1}^n$; where the $\{\varepsilon_i\}_{i=1}^n$ are unknown.

Let

$$y = \sum_{i=1}^{n} d_i(\hat{\eta}_i + \varepsilon_i)$$

where $\bar{d} = (d_1, d_2, \ldots d_n)$ are to be found so that y is a good approximation to R_s. If we try to choose $\bar{d}$ to minimize $||R_s - y||$ in the error bound

$$|\langle z, R_s - y\rangle| \le ||z||\,||R_s - y|| \tag{5.8}$$

it is necessary to know $\langle R_s, \varepsilon_i\rangle = \varepsilon_i(s)$, $i = 1,2,\ldots n$, which is assumed unknown.

We will choose $\bar{d}$ subject to the constraint

$$||P_o(R_s - \sum_{i=1}^{n} d_i \hat{\eta}_i)||^2 = 0$$

Then

$$||R_s - y||^2 \le 2\{||P_1(R_s - \sum_{i=1}^{n} d_i \hat{\eta}_i)||^2 + ||\sum_{i=1}^{n} d_i \varepsilon_i||^2\}$$

(5.9)

Let V be the matrix with i,jth entry $v_{ij} = \langle \varepsilon_i, \varepsilon_j \rangle$ given by

$$\langle \varepsilon_i, \varepsilon_j \rangle = \int_0^1 K(t_i,u)du\{\int_0^1 K(t_j,v)R(u,v)dv - \sum_{k=1}^{N} K(t_j,s_k)R(u,s_k)\omega_k\}$$

$$- \sum_{\ell=1}^{N} K(t_i,s_\ell)\{\int_0^1 K(t_j,v)R(s_\ell,v)dv - \sum_{k=1}^{N} K(t_j,s_k)R(s_\ell,s_k)\omega_k\}$$

(5.10)

and let $\hat{X}$ and $\hat{\Sigma}$ be the n × m and n × n matrices with entries $\{\hat{\chi}_{\mu i}\}$ and $\{\hat{\sigma}_{ij}\}$ respectively, given by (5.4) and (5.5).

By Lemma 2.8 with $\delta_0 = R_s$, the solution $\tilde{R}_s$ to the problem: Find y of the form

$$y = \sum_{i=1}^{n} d_i(\hat{\eta}_i + \varepsilon_i) \qquad (5.11)$$

to minimize

$$||P_1(R_s - \sum_{i=1}^{n} d_i \hat{\eta}_i)||^2 + ||\sum_{i=1}^{n} d_i \varepsilon_i||^2 \qquad (5.12)$$

subject to

$$||P_o(R_s - \sum_{i=1}^{n} d_i \hat{\eta}_i)||^2 = 0 \tag{5.13}$$

is given by

$$\tilde{R}_s = \bar{\phi}(s)(\hat{X}\hat{S}^{-1}X')^{-1}\hat{X}\hat{S}^{-1}(\hat{\eta+\varepsilon})$$

$$+ \bar{\hat{\xi}}(s)(\hat{S}^{-1} - \hat{S}^{-1}\hat{X}(\hat{X}\hat{S}^{-1}\hat{X}')^{-1}\hat{S}^{-1})(\hat{\eta+\varepsilon}) \tag{5.14a}$$

where

$$\bar{\phi}(s) = (\phi_1(s), \phi_2(s), \ldots \phi_m(s))$$

$$\bar{\hat{\xi}}(s) = (\hat{\xi}_1(s), \hat{\xi}_2(s), \ldots \hat{\xi}_n(s)) \tag{5.14b}$$

$$\hat{S} = \sum + V$$

and

$$\hat{\eta+\varepsilon} = (\hat{\eta}_1+\varepsilon_1, \hat{\eta}_2+\varepsilon_2, \ldots \hat{\eta}_n+\varepsilon_n)'.$$

Any "optimal" approximation to R_s will depend on the unknown V. Thus it is desirable here to approximate V using whatever information is available. A plausible approximation is

$$V \approx \lambda I \tag{5.15}$$

where λ is a "guestimate" of the mean diagonal element of V,

$$\lambda \approx \frac{1}{n} \sum_{i=1}^{n} \langle \varepsilon_i, \varepsilon_i \rangle, \tag{5.16}$$

this "questimate" being based on (5.10) with i = j and the properties of the quadrature formula being used.

Let $\tilde{R}_{s,\lambda} \in \mathcal{H}$ be given by (5.14) with V replaced by λI, that is,

$$\tilde{R}_{s,\lambda} = \bar{\phi}(s)(\hat{X}\hat{S}_\lambda^{-1}\hat{X}')^{-1}\hat{X}\hat{S}_\lambda^{-1}(\hat{\eta+\varepsilon})$$

$$+ \hat{\bar{\xi}}(s)(\hat{S}_\lambda^{-1} - \hat{S}_\lambda^{-1}\hat{X}(\hat{X}\hat{S}_\lambda^{-1}\hat{X}')^{-1}\hat{S}_\lambda^{-1})(\hat{\eta+\varepsilon}) \tag{5.17}$$

where

$$\hat{S}_\lambda = \hat{\sum} + \lambda I.$$

Then an estimate $\hat{z}_\lambda$ for z(s) is defined by

$$\hat{z}_\lambda(s) = \langle z, \tilde{R}_{s,\lambda} \rangle \tag{5.18}$$

with

$$|z(s) - \hat{z}_\lambda(s)| \leq ||z|| \, ||R_s - \tilde{R}_{s,\lambda}||$$

The function $\hat{z}_\lambda$ defined by (5.18) is in $\mathcal{H}$ and may be written

$$\hat{z}_\lambda = \phi(\hat{X}\hat{S}_\lambda^{-1}\hat{X}')\hat{X}\hat{S}_\lambda^{-1}\bar{u} \tag{5.19}$$

$$+ \hat{\xi}(\hat{S}_\lambda^{-1} - \hat{S}_\lambda^{-1}\hat{X}(\hat{X}\hat{S}_\lambda^{-1}\hat{X}')^{-1}\hat{S}_\lambda^{-1})\bar{u}$$

where ϕ and $\hat{\xi}$ are vectors of elements of $\mathcal{H}$ with ϕ as in (4.9) and

$$\hat{\xi} = (\hat{\xi}_1, \hat{\xi}_2, \ldots \hat{\xi}_n)$$

with $\hat{\xi}_i$ given by (5.3). By Lemma 2.7, $\hat{z}_\lambda$ is the solution to the problem: Find $z \varepsilon \mathcal{H}$ to minimize

$$\sum_{i=1}^{n} (u_i - \langle\hat{\eta}_i, z\rangle)^2 + \lambda \int_0^1 (L_m z)_s^2 ds \qquad (5.20)$$

where

$$\langle\hat{\eta}_i, z\rangle = \sum_{k=1}^{N} K(t_i, s_k)\omega_k z(s_k), \qquad i = 1,2,\ldots n$$

and λ is of the order of magnitude of the mean square quadrature error. Thus, if the primary source of error in forming a computational estimate of z(s) is quadrature error, then this shows that an appropriate choice for the regularizing parameter α of (1.6) is as λ, an estimate of the mean square quadrature error, as defined by (5.10) and (5.16).

We have not mentioned the choice of quadrature formula. Once the quadrature points $\{s_k\}_{k=1}^N$ are chosen, the choice of best formula in the sense of Sard is equivalent to approximating the element $\eta_i \varepsilon \mathcal{H}$ by a linear estimation of the elements $\{R_{s_k}\}_{k=1}^N$. The optimum coefficients in this case are readily seen to depend on the unknown integrals. Hence a convenient quadrature formula which allows a "guestimate" of λ should be used.

As is widely known, as soon as there are experimental or computational errors, there is a point of diminishing returns in choosing n too large. If, e.g. $K(t,s)$ is continuous then $||\Sigma^{-1}||$ and $||\hat{\Sigma}^{-1}|| \to \infty$ as n becomes large, where $||\cdot||$ is the spectral norm. We will quantify this statement and indicate a mitigating technique. To simplify the equations, we let $M_\nu z = 0$, $\nu = 1,2,\ldots m$. Then we may let

$$u_i = \langle \hat{\xi}_i + \varepsilon_i , z \rangle = \int_0^1 K(t_i,s)z(s)ds, \quad i = 1,2,\ldots n \qquad (5.21)$$

where $\{\hat{\xi}_i\}_{i=1}^n$ are given by (5.3) and $\varepsilon_i = \xi_i - \hat{\xi}_i$. Then, by Theorem 4.2 the solution to the problem: Find $z \varepsilon \mathcal{H}_1$ to minimize

$$\sum_{i=1}^n (u_i - \langle \hat{\xi}_i , z \rangle)^2 + \lambda \int_0^1 (L_m z)_s^2 ds \qquad (5.22)$$

is given by $\hat{z}_\lambda$,

$$\hat{z}_\lambda = \hat{\xi}(\hat{\Sigma} + \lambda I)^{-1}\bar{u} \qquad (5.23)$$

and

$$|z(s)-\hat{z}_\lambda(s)|^2 = |\langle z-\hat{z}_\lambda, R_{1,s}\rangle| =$$

$$|\langle z, R_{1,s}-\tilde{R}_{1,s,\lambda}\rangle| \leq ||z|| \, ||R_{1,s}-\tilde{R}_{1,s,\lambda}|| \quad (5.24)$$

where

$$\tilde{R}_{1,s,\lambda} = \sum_{i=1}^{n} d_i(\hat{\xi}_i+\varepsilon_i)$$

with $\bar{d} = (d_1,d_2,\ldots d_n)$ given by

$$\bar{d} = \hat{\xi}(s)(\hat{\Sigma}+\lambda I)^{-1}$$

with $\hat{\bar{\xi}}(s)$ given by (5.14b).

Now

$$||R_{1,s}-\tilde{R}_{1,s,\lambda}||^2 \leq 2\{||R_{1,s}-\sum_{i=1}^{n} d_i\hat{\xi}_i||^2+||d_i\varepsilon_i||^2\}$$

(5.25)

If V is the matrix with ij^{th} entry $v_{ij}, = \langle\varepsilon_i,\varepsilon_j\rangle$, then the term in brackets in (5.25) may be expanded as

$$R_1(s,s) - 2\bar{\hat{\xi}}(s)(\hat{\Sigma}+\lambda I)^{-1}\bar{\hat{\xi}}(s)' + \bar{\hat{\xi}}(s)(\hat{\Sigma}+\lambda I)^{-1}\hat{\Sigma}(\hat{\Sigma}+\lambda I)^{-1}\bar{\hat{\xi}}(s)'$$

$$+ \bar{\hat{\xi}}(s)(\hat{\Sigma}+\lambda I)^{-1}V(\hat{\Sigma}+\lambda I)^{-1}\bar{\hat{\xi}}(s)' \qquad (5.26)$$

and some algebraic **manipulation** gives that (5.26) is equal to

$$R_1(s,s) - \bar{\hat{\xi}}(s)\hat{\Sigma}^{-1}\bar{\hat{\xi}}(s)' + \bar{\hat{\xi}}(s)\hat{\Sigma}^{-1/2}A\hat{\Sigma}^{-1/2}\bar{\hat{\xi}}(s)' \qquad (5.27)$$

where $\hat{\Sigma}^{-1/2}$ is the symmetric square root of $\hat{\Sigma}^{-1}$ and

$$A = (\hat{\Sigma}+\lambda I)^{-1}(\hat{\Sigma}^{1/2}V\hat{\Sigma}^{1/2}+\lambda^2 I)(\hat{\Sigma}+\lambda I)^{-1} \qquad (5.28)$$

Letting $P_{\hat{\xi}}$ denote the projection operator onto the subspace spanned by $\{\hat{\xi}_i\}_{i=1}^n$, we have

$$||P_{\hat{\xi}}R_{1,s}||^2 = \bar{\hat{\xi}}(s)\hat{\Sigma}^{-1}\bar{\hat{\xi}}(s)' \qquad (5.29)$$

Lines (5.25), (5.27), (5.28) and (5.29) yield the bound

$$||R_{1,s}-\tilde{R}_{1,s,\lambda}||^2 \le 2\Big[||R_{1,s}-P_{\hat{\xi}}R_{1,s}||^2 + ||P_{\hat{\xi}}R_{1,s}||^2 \cdot$$

$$\{||(\hat{\Sigma}+\lambda I)^{-1}||\cdot||V+\lambda I||\}\Big] \qquad (5.30)$$

If $V = \lambda I$, the term in **curly brackets in** (5.30) can be replaced by $\lambda||(\hat{\Sigma}+\lambda I)^{-1}||$. In practice as N, the number of quadrature points becomes large, one expects $||V||$ to decrease to a finite limit imposed by round off errors. As $n(n<<N)$, the number of data points, becomes large, $||R_{1,s}-P_{\hat{\xi}}R_{1,s}||$ decreases, however, $\lambda||(\hat{\sum}+\lambda I)^{-1}||^2$ will increase for λ bounded below since $||(\hat{\sum})^{-1}|| \to \infty$ as $n \to \infty$.

Let $H = \{h_{\ell i}\}$ be a $p \times n$ matrix of real numbers of rank p. Let

$$\hat{\hat{\xi}}_\ell = \sum_{i=1}^{n} h_{\ell i}\hat{\xi}_i$$

$$\hat{\varepsilon}_\ell = \sum_{i=1}^{n} h_{\ell i}\varepsilon_i \tag{5.31}$$

$$\underline{u}_\ell = \sum_{i=1}^{n} h_{\ell i}u_i = \sum_{i=1}^{n} h_{\ell i}\int_0^1 K(t_i,s)z(s)ds, \quad \ell = 1,2,\ldots p.$$

If we estimate z(s) assuming the information $z\varepsilon\mathcal{H}_1$,

$$\langle z,\hat{\hat{\xi}}_\ell+\hat{\varepsilon}_\ell\rangle = \underline{u}_\ell, \quad \ell = 1,2,\ldots p$$

as before, only replacing the set $\{\hat{\xi}_i\}_{i=1}^n$ by $\{\hat{\hat{\xi}}_\ell\}_{\ell=1}^p$ and the numbers $\{u_i\}_{i=1}^n$ by $\{\underline{u}_\ell\}_\ell^p$, an error bound analogous to the right hand side of (5.30) is obtained of the form

$$2(||R_{1,s}-P_{\hat{\hat{\xi}}}R_{1,s}||^2 + ||P_{\hat{\hat{\xi}}}R_{1,s}||^2\{||(H\hat{\sum}H'+\lambda I)^{-1}||\cdot||HVH'+\lambda I||\}) \tag{5.32}$$

where $P_{\hat{\xi}}$ is the projection operator onto the subspace spanned by $\{\hat{\xi}_\ell\}_{\ell=1}^p$. If the rows of H are orthonormal vectors in Euclidean n-space, then the term in curly brackets in (5.32) can only be decreased, as compared to the term in brackets in (5.20), while the first term in (5.32) may, for large n, and n-p, may not increase much, as compared to the first term in (5.30). If $V \approx \lambda I$, then the "optimum" choice of H to minimize the term in brackets in (5.32) is to choose the rows of H as the eigen vectors corresponding to the p largest eigenvalues of $\hat{\Sigma}$.

6. 'Statistical Estimation' of Solutions to Integral Equations

It is far from coincidental that the method of regularization and the method of statistical estimation lead to the same numerical solution. Let $R(s,s')$ be a continuous positive definite kernel on $S \times S$, and $\mathcal{H}_R$ the reproducing kernel Hilbert space associated with R. Let $Z(s)$, $s \epsilon S$ be a stochastic process (i.e. a family of random variables indexed by s), with $EZ(s) = 0$ [2] and

$$EZ(s)Z(s') = R(s,s') \qquad (6.1)$$

[2] without real loss of generality

Let K(t,s) satisfy the assumptions of the corollary to Theorem 1.1 and consider the stochastic model

$$U(t) = \int_S K(t,s)Z(s)ds \qquad (6.2)$$

where observations will be taken on the random variables $U(t_i)$, i = 1,2,...n. U(t) is a well defined random variable for every $t\varepsilon T$. We have been studying the (deterministic) model

$$u(t) = \int_S K(t,s)z(s)ds \qquad (6.3)$$

where $z\varepsilon \mathcal{H}_R$, and the numbers $u(t_i)$, i = 1,2,...n are available. The purpose of this section is to demonstrate rigorously that the same numerical solution to the integral equation is obtained wether the true 'solution' is considered to be an element $z\varepsilon\mathcal{H}_R$ or a realization of a stochastic process Z(s) with covariance R(s,s').

Let $\mathcal{H}_Z$ be the Hilbert space spanned by the stochastic process $\{Z(s), s\varepsilon S\}$. (See [6]). $\mathcal{H}_Z$ is defined as follows: All random variables Y which are finite linear combinations of the form

$$Y = \sum_{\ell} a_{\ell} Z(s_{\ell}) \qquad s_{\ell}\varepsilon S \qquad (6.4)$$

are in $\mathcal{H}_Z$. An inner product on the linear manifold of all such finite linear combinations is

$$\langle Y_1, Y_2 \rangle = EY_1Y_2, \qquad (6.5)$$

and $\mathcal{H}_Z$ is the closure of this linear manifold with the given inner product. The precise source of the duality between 'deterministic' and statistical models is the following (well known) fact:

$\mathcal{H}_Z$ is isometrically isomorphic to $\mathcal{H}_R$ under the isomorphism induced by the correspondance "~",

$$Z(s) \sim R_s, \qquad \forall s \varepsilon S \qquad (6.6)$$

Furthermore, the random variable $Y \varepsilon \mathcal{H}_Z$ corresponds to the element $\eta \varepsilon \mathcal{H}_R$ if and only if

$$EYZ(s) = \langle \eta, R_s \rangle = \eta(s), \qquad s \varepsilon S \qquad (6.7)$$

The family of random variables $U(t)$, $t \varepsilon T$ are all in $\mathcal{H}_Z$ by our assumptions on $K(t,s)$ and

$$U(t) \sim \eta_t \qquad (6.8)$$

where η_t is defined, for each t, by (1.25). Let $\{\phi_\nu\}_{\nu=1}^m$ be a specified set of m orthonormal elements in $\mathcal{H}_R$, and let

$\{\rho_\nu\}_{\nu=1}^m$ be the m random variables in $\mathcal{H}_Z$ which correspond to $\{\phi_\nu\}_{\nu=1}^m$ under the isomorphism induced by (6.6). It can then be shown that $Z(s)$, $s\varepsilon S$ has a representation of the form

$$Z(s) = \sum_{\nu=1}^{m} \rho_\nu \phi_\nu(s) + Z_1(s) \qquad (6.9)$$

where $Z_1(s)$ is a zero mean stochastic process with

$$EZ_1(s)Z_1(s') = R_1(s,s') = R(s,s') - \sum_{\nu=1} \phi_\nu(s)\phi_\nu(s') \qquad (6.10)$$

and $E\rho_\nu\rho_\mu = \delta_{\mu,\nu}$, $\mu,\nu = 1,2...m$. Let $\hat{Z}(s)$ be, for each s, that random variable in the subspace of $\mathcal{H}_Z$ spanned by $\{U(t_i)\}_{i=1}^n$ which minimizes

$$E(\hat{Z}(s) - Z(s))^2 \qquad (6.11)$$

subject to

$$E(\hat{Z}(s) - Z(s)|\rho_\nu, \nu = 1,2...n) = 0. \qquad (6.12)$$

It follows from Lemma 2.2 and (6.6) with the identifications $Z(s) \sim \delta_o$, $U(t_i) \sim \eta_i$, $\rho_\nu \sim \phi_\nu$, and $\int K(t_i,s)Z_1(s)ds \sim \xi_i$, that

$$\hat{Z}(s) = \phi(s)(X\Sigma^{-1}X')^{-1}X\Sigma^{-1}U'$$

$$+ \xi(s)(\Sigma^{-1} - \Sigma^{-1}X(X\Sigma^{-1}X')^{-1}X\Sigma^{-1})U' \qquad (6.13)$$

where

$$\phi(s) = (\phi_1(s), \phi_2(s), \ldots \phi_m(s))$$

$$\xi(s) = (\xi_1(s), \xi_2(s), \ldots \xi_n(s))$$

X and Σ are as in (2.1), and U is the vector of random variables given by

$$U = (U(t_1), U(t_2), \ldots U(t_n))$$

Thus, the numerical value $\hat{Z}(s)$ based on the model (6.2) and a "realization" $U(t_i) = u_i$, $i = 1,2,\ldots n$, is exactly the same as the numerical value of $\hat{z}_o(s)$ of (4.11), based on the model (6.3). An identical statement may be made about $\hat{z}_\lambda(s)$ of (5.19) if we replace (6.2) by the stochastic model

$$U(t_i) - \sum_{k=1}^{N} K(t_i,s_k)\omega_k Z(s_k) + \epsilon(t_i) \quad i = 1,2,\ldots n \qquad (6.14)$$

where

$$E\,\epsilon(t_i) = 0$$

$$E\,\epsilon(t_i)\epsilon(t_j) = 0, \qquad i \neq j$$

$$= \lambda, \qquad i = j$$

$$E\,\epsilon(t_i)Z(s_k) = 0, \qquad i = 1,2,\ldots n$$
$$k = 1,2,\ldots N$$

If we let $\mathcal{H}_R^*$ be the Hilbert space of all continuous linear functionals on $\mathcal{H}_R$, then $\mathcal{H}_R^*$ is consequently isometrically isomorphic to $\mathcal{H}_Z$ under the correspondance

$$Z(s) \sim N_s$$

where N_s is the continuous linear functional defined by

$$N_s z = z(s), \qquad z \in \mathcal{H}_R$$

Then

$$U(t_i) \sim \Lambda_{t_i}$$

where Λ_{t_i} is the continuous linear functional defined by

$$\Lambda_{t_i} z = \int_S K(t_i, s) z(s) ds, \qquad z \varepsilon \mathcal{H}_R$$

It is seen that the geometry for approximating Z(s) by $\{U(t_i)\}_{i=1}^n$ is exactly the same as the geometry for approximatin N_s by $\{\Lambda_{t_i}\}_{i=1}^n$.

An experimenter approaching the problem with the model (6.2) chooses the prior covariance R(s,s') of (6.1) according to his belief or past experience concerning Z(s). The numerical analyst, beginning with (6.3) should choose an R such that the norm of the solution z in $\mathcal{H}_R$ is known or believed to be small.

It is clear that algorithms for the numerical solution of a broad variety of (linear) equations can, in fact, be identified with prediction problems on stochastic processes in this manner.

REFERENCES

1. Ahlberg, T.H., Nilson, E.W., and Walsh, J.L., The Theory of Splines and Their Applications, (1967), Academic Press, New York.

2. Aronszajn, N., Theory of reproducing kernels, Trans. Am. Math. Soc., 68 (1950), pp. 337-404.

3. de Boor, C. and Lynch, R.E., On splines and their minimum properties, J. Math. Mech., 15 (1967), pp. 953-69.

4. Golomb, M. and Weinberger, H.F., Optimal approximation and error bounds, pp. 117-90 of On Numerical Approximation (R.E. Langer, Ed.), University of Wisconsin Press, Madison, Wisconsin (1959).

5. Mangasarian, O.L., and Schumaker, L.L., Splines via optimal control, to appear in the proceedings of the Conference on Approximations, with Special Emphasis on Splines, held at the University of Wisconsin, May 4-6, 1969.

6. Parzen, E., An approach to time series analysis, Ann. Math. Statis., 32, (1961), pp. 951-989.

7. Phillips, David L., A technique for the numerical solution of certain integral equations of the first kind, J. Assn. Comput. Mach. 9 (1962), pp. 84-97.

8. Ribiere, G., Regularisation d'Operateurs, R.I.R.O. 1, 5, (1967), pp. 57-79.

9. Sard, A., Linear Approximation, (Mathematical Surveys, No. 9), American Mathematical Society, Providence, R. I. (1963).

10. Slepian, D. and Pollak, H.O., Prolate Spheroidal Wave Functions, Fourier Analysis, and Uncertainty - I, Bell System Technical Journal, 40, (1961), pp. 43-64.

11. Strand, O.N. and Westwater, E.R., Statistical estimation of the numerical solution of a Fredholm integral equation of the first kind, J.A.C.M. 15, 1, (1968), pp. 100-114.

12. Tihonov, A.N., Solution of incorrectly formulated problems and the regularization method, Soviet Mathematics 4, #4, (1963), pp. 1035-38.

13. Tihonov, A.N., Regularization of incorrectly posed problems, Soviet Mathematics 4, #6, (1963), pp. 1624-27.

14. Tihonov, A.N. and Glasko, V.B., An approximate solution of Fredholm integral equations of the first kind, Ž. Vyčisl. Mat. i Mat. Fiz. 4 (1964), pp. 564-571.