Proceedings of the
44th IEEE Conference on Decision and Control, and
the European Control Conference 2005
Seville, Spain, December 12-15, 2005

Control Theoretic Splines with Deterministic and Random Data

Y. Zhou M. Egerstedt C. Martin

Abstract— **In this paper we give a basic derivation of smoothing splines and through this derivation we show that the basic smoothing spline construction can be separated into a filtering problem on the raw data and an interpolating spline construction. Both the filtering algorithm and and the interpolating spline construction can be effectively implemented. We allow hard constraints (such as boundary values) on the dynamics and we allow data that is subject to error. We are thus constructing smoothing splines with hard constraints.**

I. Introduction

Estimation and smoothing for data sets that contain deterministic and random data present difficulties not present in purely random data sets. Yet such data sets are very common in practice and if the nature of the data is not respected conclusions may be drawn that have little relation to reality. In this paper we will present a unified treatment of such problems. We will extend the theory of smoothing splines to cover such situations. Some of the techniques that we will use have been developed in papers by Egerstedt and Martin, [11], [6], [3] and their colleagues. The main technical contribution of this paper will be to show that many of these problems can be cast as minimum norm problems in suitable Hilbert spaces. This approach unifies a series of problems that have been solved by Egerstedt, Zhou, Sun and Martin, [14], [15], [16]. We will see that this method rests on the Hilbert space methods developed by Luenberger in [5].

A very simple problem is to determine the volume of water contained in a playa lake in West Texas, [9]. These are transient water supplies that because of their formation are almost perfectly circular. If a transect is made across the center of the lake it is possible to obtain a fairly good estimate of the volume. At the boundary of the lake the depth of the water is 0cm. However the depth is measured by a graduate student wading through the lake and measuring the depth at a series of points. These measurement are quite random. The bottom of the lake is silted and so it is not clear where the bottom of the probe rests and the measurement is made by reading the depth of a marked probe. The data set then consists of two deterministic values at the boundary and a series of random numbers representing the depth at a series of predetermined points.

In population studies the census is taken every ten years and whether correct or not the values of the census are considered to be absolute for many purposes. Estimates are made of populations within a given city at irregular intervals between the censuses. Thus if it is necessary to study the growth or decline of a city over a long period of time deterministic data is available at ten year intervals and estimated data is available at shorter and often irregular time periods. Thus the data set consists of deterministic census data and random estimated data. Estimates such as the report by the State of California, [8], are a necessary and critical part of planning for governments.

For most individuals in the United States their home is the principle component of their financial portfolio. The question of the value of the portfolio is of interest in a variety of economic indicators, [7]. When the home is purchased there is a firm monetary value that can be measured and when the home is sold there is a firm value. In between the value is less certain. Almost every individual can give you an estimate of the value but unless a formal appraisal is done there may be a very large error in the estimate. This results in a data set with a few deterministic values, the purchase price, the selling price and formal appraisals and many random values that are estimates by the owner.

These problems all have in common some data that can be assumed to be exact and some data that is subject to error. The goal of this paper is to find a common frame work to treat all such linear problems.

In this paper we will consider the problem of approximating discrete or continuous data using the dynamics of a linear controlled system. The system may have hard constraints such as boundary values and/or hard constraints at internal values. The data will be assume to be noisy with known statistics. A contribution of this paper is to formulate these problems as a general class of minimal norm problems in Hilbert space. Egerstedt, Sun and Martin, [6], [11], have formulated interpolation problems as minimal norm problems but the general problems of smoothing spines have not to this point been so formulated. The advantage is more than conceptual in that the smoothed data is immediately available as is the smooth functional approximation. Thus we will be able to split the problems into an estimation problem and a problem of finding the interpolating splines. Both of these problems have fast numerical implementations.

II. Statement of the basic problem

In this section we state the basic problem of smoothing splines and construct the solution. Here we show that the construction splits into two parts in a very natural way. Ultimately this will allow the implementation of fast

Y. Zhou is with the Division of Mathematics, Stockholm University, SE-106 91 Stockholm, Sweden. Email: yishao@math.su.se

M. Egerstedt is with the School of Electrical and Computer Engineering, Georgia Institute of Technology, Atlanta, GA 30332, USA. Email: magnus@ece.gatech.edu

C. Martin is with the Department of Mathematics and Statistics, Texas Tech University, Lubbock, TX 79409, USA. Email: martin@.math.ttu.edu

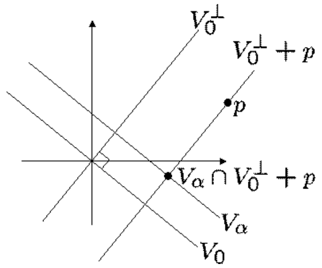

Fig. 1. The general process for finding the point on an affine variety (V_α) in a Hilbert space closest to a given data point (p).

algorithms for smoothing spline constructions. The basic idea of the construction is to define a linear variety, in a Hilbert space, that contains all of the constraints. The data is then defined as a point in the Hilbert space and the optimization reduces to finding the point on the affine variety that is closest (in the sense of the norm in the Hilbert space) to the data point. We know that we can construct this point by finding the orthogonal complement of the linear variety that defines the affine variety and constructing the intersection of the affine variety with the orthogonal complement, as seen in Figure 1. In this process we follow Luenberger, [5].

A. The definitions

Let

$$\dot{x} = Ax + bu, \quad y = cx$$

be a controllable and observable linear system with initial data $x(0) = x_0$. As will be seen later it makes for the smoothest approximation if we have

$$cb = cAb = cA^2b = \cdots = cA^{n-2}b = 0 \qquad \text{(II.1)}$$

where n is the dimension of the system. The initial data will be chosen as part of the algorithm. We will later see that the use of an initial value problem is unnecessary.

Let a data set be given as

$$D = \{(t_i, \alpha_i) : \ i = 1, \cdots, N\}$$

and assume that $t_i > 0$ and let $T_f = t_N$. Our goal is to find a control $u(t)$ that minimizes

$$J(u, x_0) = \int_0^{T_f} u^2(t)dt + (\hat{y} - \hat{\alpha})^T Q(\hat{y} - \hat{\alpha}) + x_0^T R x_0$$

where Q and R are positive definite matrices. The vector $\hat{y}$ has components

$$y_i = y(t_i) = ce^{At_i}x_0 + \int_0^{t_i} ce^{A(t_i-s)}bu(s)ds$$

and the vector $\hat{\alpha}$ has components α_i.

It is convenient to define the functions

$$\ell_i(s) = \begin{cases} ce^{A(t_i-s)}b & t_i > s \\ 0 & t_i < s \end{cases}$$

Note that if the assumption on zeros, equation II.1, holds then $\ell_i(s)$ is $n-1$ times differentiable at t_i. We can now write

$$y_i = ce^{At_i}x_0 + \int_0^{T_f} \ell_i(s)u(s)ds = ce^{At_i}x_0 + \langle \ell_i, u\rangle_L.$$

Now let $\beta_i = R^{-1}e^{A^Tt_i}c^T$ and we have

$$\begin{aligned} y_i &= ce^{At_i}x_0 + \int_0^{T_f} \ell_i(s)u(s)ds \\ &= \langle \beta_i, x_0\rangle_R + \langle \ell_i, u\rangle_L \end{aligned}$$

where we define the inner product

$$\langle x, w\rangle_R = x^T R w.$$

B. The Hilbert space and the affine variety

Let

$$\mathcal{H} = L_2[0, T_f] \oplus R^n \oplus R^N$$

with norm

$$\|(u; x; d)\|^2 = \int_0^{T_f} u^2(t)dt + d^TQd + x^TRx.$$

We define the affine variety of constraints, V, in $\mathcal{H}$ as

$$V = \{(u; x; d) : \quad d_i = \langle \beta_i, x\rangle_R + \langle \ell_i, u\rangle_L\}$$

Note that V is of infinite dimension and finite co-dimension. We will construct the orthogonal complement of V in $\mathcal{H}$.

$$\begin{aligned} &V^\perp = \{(u'; x'; d') : \\ &\quad \forall (u; x; d) \in V \ \langle u', u\rangle_L + \langle d', d\rangle_Q + \langle x', x\rangle_R = 0\}. \end{aligned}$$

Now we have

$$\begin{aligned} \langle d', d\rangle_Q &= \sum_{i=1}^N \langle d', e_i\rangle_Q d_i \\ &= \sum_{i=1}^N \langle d', e_i\rangle_Q [\langle \beta_i, x\rangle_R + \langle \ell_i, u\rangle_L] \\ &= \sum_{i=1}^N [\langle d', e_i\rangle_Q \langle \beta_i, x\rangle_R + \langle d', e_i\rangle_Q \langle \ell_i, u\rangle_L] \\ &= \sum_{i=1}^N [\langle \langle d', e_i\rangle_Q \beta_i, x\rangle_R + \langle \langle d', e_i\rangle_Q \ell_i, u\rangle_L] \\ &= \langle \sum_{i=1}^N \langle d', e_i\rangle_Q \beta_i, x\rangle_R + \langle \sum_{i=1}^N \langle d', e_i\rangle_Q \ell_i, u\rangle_L \end{aligned}$$

Therefore we have

$$\begin{aligned} 0 &= \langle u', u\rangle_L + \langle x', x\rangle_R + \langle d', d\rangle_Q = \\ &= \langle u', u\rangle_L + \langle x', x\rangle_R + \langle \sum_{i=1}^N \langle d', e_i\rangle_Q \beta_i, x\rangle_R + \\ &+ \langle \sum_{i=1}^N \langle d', e_i\rangle_Q \ell_i, u\rangle_L = \\ &= \langle x' + \sum_{i=1}^N \langle d', e_i\rangle_Q \beta_i, x\rangle_R + \langle u' + \sum_{i=1}^N \langle d', e_i\rangle_Q \ell_i, u\rangle_L \end{aligned}$$

From this we see that

$$V^\perp = \{(u';x';d'): \quad x' + \sum_{i=1}^N \langle d', e_i\rangle_Q \beta_i = 0,$$
$$u' + \sum_{i=1}^N \langle d', e_i\rangle_Q \ell_i = 0\}$$

C. The intersection

Before constructing the intersection two things must be verified. The first is that V is nonempty and the second is that V is closed. Both follow from the fact that V is the graph of a continuous mapping from $L_2[0,T] \times R^n$ to R^N. Equating quantities from V and $V^\perp + p$ ($p = (0;0;\hat{\alpha})$ is the data point) we have from the definition of V and some rearrangement of terms

$$\begin{aligned} d_i &= \langle \beta_i, x\rangle_R + \langle \ell_i, u\rangle_L \\ &= -\sum_{j=1}^N \langle d', e_j\rangle_Q \langle \beta_i, \beta_j\rangle_R - \sum_{j=1}^N \langle d', e_j\rangle_Q \langle \ell_i, \ell_j\rangle_L \end{aligned}$$

Now, equating d with $\hat{y}$ and d' with $\hat{y} - \hat{\alpha}$, we get

$$\begin{aligned} y_i &= -\sum_{j=1}^N \langle \hat{y} - \hat{\alpha}, e_j\rangle_Q \langle \beta_i, \beta_j\rangle_R - \\ &- \sum_{j=1}^N \langle \hat{y} - \hat{\alpha}, e_j\rangle_Q \langle \ell_i, \ell_j\rangle_L \\ &= -e_i^T GQ(\hat{y} - \hat{\alpha}) - e_i^T FQ(\hat{y} - \hat{\alpha}) \end{aligned}$$

where G is the Grammian of the β_is and F is the Grammian of the l_is. Note that since the ℓ_is are linearly independent F is invertible. In more compact form we have

$$\hat{y} = -(GQ + FQ)(\hat{y} - \hat{\alpha})$$

or finally we have that

$$(I + GQ + FQ)\hat{y} = (GQ + FQ)\hat{\alpha}. \qquad \text{(II.2)}$$

By rewriting $I + GQ + FQ = (Q^{-1} + F + G)Q$ and since F and Q are positive definite and G is positive semidefinite the matrix $(I + GQ + FQ)$ is invertible and we find $\hat{y}$ as linear function of the data $\hat{\alpha}$. This $\hat{y}$ is the smoothed estimate of the data $\hat{\alpha}$. Using $\hat{y}$ we can then calculate both the optimal control and the optimal initial condition. We summarize with the following theorem.

Theorem 2.1: Let

$$\dot{x} = Ax + bu, \quad y = cx$$

be a controllable and observable linear system with initial data $x(0) = x_0$ and let a data set be given as

$$D = \{(t_i, \alpha_i): \ i = 1, \cdots, N\}$$

and assume that $t_i > 0$ and let $T = t_N$. Let the cost function be given as

$$J(u, x_0) = \int_0^{T_f} u^2(t)dt + (\hat{y} - \hat{\alpha})^T Q(\hat{y} - \hat{\alpha}) + x_0^T R x_0$$

where Q and R are positive definite matrices. The vector $\hat{y}$ has components

$$y_i = y(t_i) = ce^{At_i}x_0 + \int_0^{t_i} ce^{A(t_i - s)}bu(s)ds$$

and the vector $\hat{\alpha}$ has components α_i. Minimizing J over $u \in L_2[0,t]$ and $x_0 \in R^n$ we have that the optimal smoothed data is given by

$$\hat{y} = (I + GQ + FQ)^{-1}(GQ + FQ)\hat{\alpha}, \qquad \text{(II.3)}$$

the optimal control is given by

$$u = \sum_{i=1}^N \langle [I + (I + GQ + FQ)^{-1}(GQ + FQ)]\hat{\alpha}, e_i\rangle_Q \ell_i \qquad \text{(II.4)}$$

and the optimal initial condition is given by

$$x_0 = \sum_{i=1}^N \langle [I + (I + GQ + FQ)^{-1}(GQ + FQ)]\hat{\alpha}, e_i\rangle_Q \beta_i \qquad \text{(II.5)}$$

III. Interpolating splines with initial data

For interpolating splines we are required to find a control that drives the output y through the points in the data set D. This can be expressed in terms of additional constraints of the form

$$\alpha_i = \langle \beta_i, x_0\rangle_R + \langle \ell_i, u\rangle_L.$$

The goal is thus to find a control and an initial condition that minimizes

$$J(u, x_0) = \int_0^{T_f} u^2(t)dt + x_0^T R x_0$$

subject to the constraints. Just as for smoothing splines we define the Hilbert space to be

$$\mathcal{H} = L_2[0, T_f] \times R^n.$$

Now the affine variety of constraints is given by

$$V_{\hat{\alpha}} = \{(u;x_0): \ \alpha_i = \langle \beta_i, x_0\rangle_R + \langle \ell_i, u\rangle_L\}$$

Here the goal is to find the point in $V_{\hat{\alpha}}$ of minimum norm. The procedure is much the same as for smoothing splines. We first must verify that V_α is nonempty. This follows from the hypothesized controllability of the linear system. We construct the $V_0^\perp$ and construct the intersection

$$V_0^\perp \cap V_{\hat{\alpha}},$$

which contains a single point.

After some calculation we have

$$V_0^\perp = \{(u';x'): u' = \sum_{i=1}^N \tau_i \ell_i, x' = \sum_{i=1}^N \tau_i \beta_i\}.$$

After some more calculation we have that the optimal u is in fact given by

$$u = \sum_{i=1}^N e_i^T (F + G)^{-1} \hat{\alpha} \ell_i$$

and the optimal initial condition is given by

$$x_0 = \sum_{i=1}^{N} e_i^T (F+G)^{-1}\hat{\alpha}\beta_i.$$

This is just a slight generalization of the construction given in [11] and hence the details are left out.

For cubic splines the classical construction reduces to solving a system of equations of the form $Ax = b$ where A is tridiagonal and of course this a much faster procedure. In [13] the construction of interpolating splines is reduced to solving banded matrices. However, in both cases additional constraints are required to make the problem have a unique solution. With the procedure developed here the additional constraints are unnecessary because of the optimization. Neither the classical cubic splines nor the procedure developed in [13] can easily handle the optimal initial data.

IV. Smoothing and Estimation for Problems with Additional Hard Constraints

In a series of papers Willsky and coauthors [1], [2] and Krener [4] considered an estimation problem based on a stochastic boundary value problem. In this section we consider a similar problem in which the smoothing spline is generated by linear system for which there are hard constraints. The constraints may occurs as boundary values but they may also occur as fixed internal values or even as linear operator constraints on the solution. We will show that many of these problems can be formulated and solved with the machinery that we have established. The basic idea is that we have a data set in which each data point is of the form $\alpha_i = f(t_i) + \epsilon_i$ where $f(t_i)$ is deterministic and the ϵ_i is the value of random variable. The goal is to produce a curve (the spline) that better approximates $f(t)$. This is, of course, a standard statistical assumption, [12].

We begin by considering a general boundary value problem. Let the boundary condition be given by

$$\Phi x(0) + \Psi x(T) = h \qquad \text{(IV.1)}$$

where we let $h \in R^k$. This, of course, includes the classical two point boundary value formulations and other problems of interest. We note that since

$$x(T_f) = e^{AT_f}x(0) + \int_0^{T_f} e^{A(T_f-s)}bu(s)ds$$

the specific dependence on $x(T_f)$ can be removed and the boundary constraint simply becomes

$$Px(0) + \Psi\int_0^{T_f} e^{A(T_f-s)}bu(s)ds = h, \qquad \text{(IV.2)}$$

where

$$P = \Phi + \Psi e^{AT_f}.$$

Note that if there is any solution to IV.1 then by the controllability hypothesis there is a solution to IV.2. We hypothesize that there is at least one solution of IV.1.

We now define the Hilbert space to be

$$\mathcal{H} = L_2[0,T_f] \times R^n \times R^N$$

with norm

$$\|(u; x_0; y)\|^2 = \int_0^{T_f} u^2(t)dt + x_0^T R x_0 + y^T R^N y.$$

We define the constraint variety to be

$$V_h = \{(u; x; d) : \ d_i = \langle \beta_i, x\rangle_R + \langle \ell_i, u\rangle_L, \\ Px + \Psi \textstyle\int_0^{T_f} e^{A(T_f-s)}bu(s)ds = h\}.$$

We first prove the following lemma.

Lemma 4.1: V_0 is a closed subspace of $\mathcal{H}$.

Proof: The mapping

$$(u; x) \to \Psi\int_0^{T_f} e^{A(T_f-s)}bu(s)ds + Px$$

, with domain $L_2[0,T_f] \times R^n$ is continuous and hence the subspace

$$\begin{aligned} W &= \{(u,x) \in L_2[0,T_f]\times R^n : \ Px \\ &+ \Psi\int_0^{T_f} e^{A(T_f-s)}bu(s)ds = 0\} \end{aligned}$$

is closed. Now the mapping from W to R^N defined by

$$d_i = \langle \beta_i, x\rangle_R + \langle \ell_i, u\rangle_L$$

is continuous and hence we appeal to the closed graph theorem to finish the proof. □

We now construct $V_0^\perp$.

Lemma 4.2:

$$\begin{aligned} V_0^\perp &= \{(u'; x'; d') : \\ & x' = -\sum_{i=1}^{N}\langle d', e_i\rangle_Q \beta_i + R^{-1}P^T\lambda, \\ & u' = -\sum_{i=1}^{N}\langle d', e_i\rangle_Q \ell_i + (\Psi e^{A(T_f-t)})^T\lambda\} \end{aligned}$$

for some $\lambda \in R^k$.

Proof: The first part of the construction is exactly the same as in subsection II-B and from there we have

$$\begin{aligned} V_0^\perp &= \{(u'; x'; d') : \ \langle x' + \sum_{i=1}^{N}\langle d', e_i\rangle_Q\beta_i, x\rangle + \\ & +\langle u' + \sum_{i=1}^{N}\langle d', e_i\rangle_Q \ell_i, u\rangle = 0\}. \end{aligned}$$

Now the relationship does not hold for all x and u but only for those x and u for which equation IV.2 holds. Multiplying by λ^T, $\lambda \in R^k$, we can rewrite equation IV.2 as

$$\langle R^{-1}P^T\lambda, x\rangle_R + \langle (\Psi e^{A(T_f-t)})^T\lambda, u\rangle_L = 0. \qquad \text{(IV.3)}$$

From this we conclude that

$$x' + \sum_{i=1}^{N}\langle d', e_i\rangle_Q \beta_i = R^{-1}P^T\lambda$$

and

$$u' + \sum_{i=1}^{N} \langle d', e_i \rangle_Q \ell_i = (\Psi e^{A(T_f - t)})^T \lambda,$$

and the lemma follows. □

It remains to construct the intersection $V_h \cap (V_0^{\perp} + p)$ to find the optimal point. This construction is technically more complicated than the simple smoothing spline but the technique is identical.

The unique point in the intersection is defined as the solution of the following system of four equations in the unknowns u, x_0, y and λ, obtained by identifying x and x' with x_0, d with $\hat{y}$, and d' with $\hat{y} - \hat{\alpha}$.

$$u = -\sum_{i=1}^{N} \langle \hat{y} - \hat{\alpha}, e_i \rangle_Q \ell_i + b^T e^{A^T (T_f - t)} \Psi^T \lambda \tag{IV.4}$$

$$x_0 = -\sum_{i=1}^{N} \langle \hat{y} - \hat{\alpha}, e_i \rangle_Q \beta_i + R^{-1} (\Phi + \Psi e^{A T_f})^T \lambda \tag{IV.5}$$

$$h = P x_0 + \int_0^{T_f} \Psi e^{A(T_f - s)} b u(s) ds \tag{IV.6}$$

$$y_i = \langle \beta_i, x_0 \rangle_R + \langle \ell_i, u \rangle_L \tag{IV.7}$$

We begin by eliminating x_0 and u from equation IV.7 by substituting equations IV.4 and IV.5. After some manipulation we have

$$\begin{aligned} y_i &= e_i^T G(\hat{y} - \hat{\alpha}) - e_i^T F(\hat{y} - \hat{\alpha}) + \beta_i^T P^T \lambda \\ &+ \int_0^{T_f} \ell_i(s) b^T e^{A^T (T_f - s)} \Psi^T ds \lambda \end{aligned}$$

Since $\beta_i = R^{-1} e^{A^T t_i} c^T$ let

$$\beta = R^{-1} (e^{A^T t_1} c^T, \cdots, e^{A^T t_N} c^T) = R^{-1} E$$

to obtain

$$\hat{y} = -G(\hat{y} - \hat{\alpha}) - F(\hat{y} - \hat{\alpha}) + E^T R^{-1} P^T \lambda + \Lambda \lambda,$$

where

$$\Lambda = \int_0^{T_f} l(s) b^T e^{A^T (T_f - s)} \Psi^T ds.$$

We will now use equation IV.6 to obtain a second equation in λ and $\hat{y}$.

$$\begin{aligned} h \;=\; & P\Big[-\sum_{i=1}^{N} \langle \hat{y} - \hat{\alpha}, e_i \rangle_Q \beta_i + \\ & \qquad + R^{-1} P^T \lambda\Big] + \\ & + \int_0^{T_f} \Psi e^{A(T_f - s)} b \Big[-\sum_{i=1}^{N} \langle \hat{y} - \hat{\alpha}, e_i \rangle_Q \ell_i + \\ & \qquad + b^T e^{A^T (T_f - s)} \Psi^T \lambda\Big] ds \end{aligned}$$

We make the following observation:

$$\begin{aligned} \sum_{i=1}^{N} \langle \hat{y} - \hat{\alpha}, e_i \rangle_Q \beta_i &= \sum_{i=1}^{N} \beta_i e_i^T Q (\hat{y} - \hat{\alpha}) \\ &= R^{-1} E Q (\hat{y} - \hat{\alpha}) \end{aligned}$$

We now define

$$M = \sum_{i=1}^{N} \int_0^{T_f} \Psi e^{A(T_f - s)} b \ell_i(s) e_i^T ds Q$$

and hence

$$\sum_{i=1}^{N} \int_0^{T_f} \Psi e^{A(T_f - s)} b \langle \hat{y} - \hat{\alpha}, e_i \rangle_Q \ell_i(s) ds = M(\hat{y} - \hat{\alpha}).$$

Using these two constructions we then have

$$\begin{aligned} h = P(-R^{-1} E Q (\hat{y} - \hat{\alpha})) + P R^{-1} P^T \lambda - \\ -M(\hat{y} - \hat{\alpha}) + \Psi \Gamma \Psi^T \lambda, \end{aligned} \tag{IV.8}$$

where Γ is the controllability Grammian

$$\Gamma = \int_0^{T_f} e^{A(T_f - s)} b b^T e^{A^T (T_f - s)} ds.$$

By combining these two expressions linking $\hat{y}$ and λ gives the following linear equation system

$$\begin{pmatrix} I + (G+F)Q & -E^T R^{-1} P^T - \Lambda \\ P R^{-1} E Q - M & P R^{-1} P^T + \Psi \Gamma \Psi^T \end{pmatrix} \begin{pmatrix} \hat{y} \\ \alpha \end{pmatrix} = \\ = \begin{pmatrix} (G+F) Q \hat{\alpha} \\ h + (P R^{-1} E Q + M) \hat{\alpha} \end{pmatrix} \tag{IV.9}$$

where Γ is the controllability Grammian

$$\Gamma = \int_0^{T_f} e^{A(T_f - s)} b b^T e^{A^T (T_f - s)} ds.$$

Using equation IV.9 we can solve for $\hat{y}$ and for λ. These values can be used in equations IV.4 and IV.5 to uniquely determine the optimal control and the the optimal initial condition. As before we see that the optimal estimate of the data is obtained independently of the control.

Remark:*The matrix E is a Grammian like matrix that determine if the intimal data can be recovered from sampled observational data. There are no known necessary and sufficient conditions for E to have full rank. This problem was studied originally by Smith and Martin and was reported in [10]. It is also interesting that the controllability Grammian arises in the formulation of the equation IV.8. The reason for the controllability Grammian to appear is more obvious when one considers the simpler problem of optimally moving between affine subspaces. This problem is studied in [17].*

A. Multiple Point boundary value problems

In this case we have a hard constraint of the form

$$\Phi_1 x(r_1) + \cdots + \Phi_k x(r_k) = h$$

and the data set

$$D = \{(t_i, \alpha_i) : \; i = 1, \cdots, N\}$$

and we assume without loss of generality that

$$\{r_i : \; i = 1, \cdots, k\} \cup \{t_i : \; i = 1, \cdots, N\} = \emptyset.$$

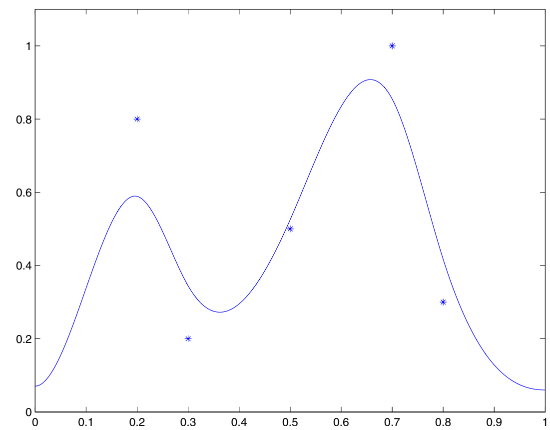

Fig. 2. Periodic splines: Here the boundary value is given by $x(0) = x(T_f)$. Depicted are $y(t)$ (solid) and $\alpha_i,\ i = 1, \dots, 4$ (stars).

We again make the assumption that there exist at least one set of vectors a_i such that

$$\Phi_1 a_1 + \cdots + \Phi_k a_k = h.$$

We construct the variety of constraints and note that we can replace $x(r_i)$ with

$$e^{Ar_i} x(0) + \int_0^{r_i} e^{A(r_i - s)} bu(s) ds.$$

Thus the constraint depends only on u and x_0. We use the Hilbert space

$$\mathcal{H} = L_2[0, T_f] \times \mathbb{R}^n \times \mathbb{R}^N.$$

The constraint variety is

$$V_h = \{(u; x_0; \hat{y}) : \ y_i = \langle \beta_i, x_0 \rangle + \langle \ell_i, u \rangle_L, \\ \sum_{i=1}^k \Phi_i e^{Ar_i} x_0 + \sum_{i=1}^k \int_0^{T_f} \Phi_i \ell_{r_i}(s) u(s) ds = h\}.$$

As before we construct the orthogonal complement to V_0 and then determine the intersection

$$V_h \cap (V_0^{\perp} + (0; 0; \hat{\alpha})).$$

We leave this construction to the reader.

V. Example

In this section we will present some examples of problems that fit this generalized boundary value formulation. We let

$$A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix},\ b = \begin{pmatrix} 0 \\ 1 \end{pmatrix},\ c = \begin{pmatrix} 1 & 0 \end{pmatrix},\ T_f = 1$$
$$t_1 = 0.2,\ t_2 = 0.3,\ t_3 = 0.5,\ t_4 = 0.7,\ t_5 = 0.8$$
$$\hat{\alpha} = \begin{pmatrix} 0.8 & 0.2 & 0.5 & 1 & 0.3 \end{pmatrix}^T$$
$$Q = 10^4 I_5,\ R = 10^4 I_2,\ (I_p = p \times p \text{ identity matrix}).$$

Example 1: Periodic splines

We first study the situation when we insist that $x(0) = x(T_f)$. In this case we have that $\Phi = -\Psi = I_2$, while $h = 0$. The solution is depicted in Figure 2.

Example 2: Two point boundary value problems

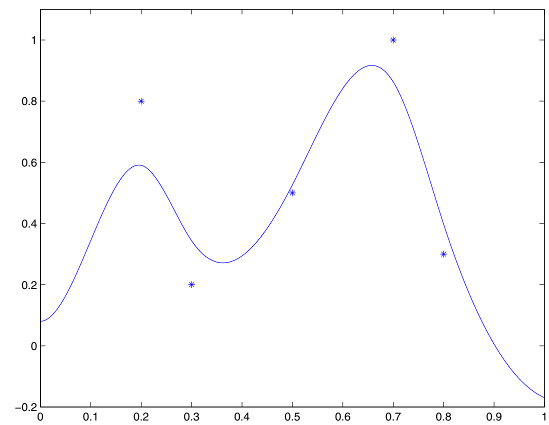

Fig. 3. Boundary value problem: $(1, 1)x(0) - (1, 1)x(T_f) = 1$.

We now let the boundary constraint be encoded by $\Phi = \begin{pmatrix} 1 & 1 \end{pmatrix}$, $\Psi = -\Phi$, $h = 1$, which implies that the boundary values are given by the set

$$\{(x_0, x_{T_f}) \mid (1, 1)x_0 - (1, 1)x_{T_f} = 1\}.$$

The solution is given in Figure 3.

References

[1] Adams, M.B.; Willsky, A.S.; Levy, B.C. Linear estimation of boundary value stochastic processes. I. The role and construction of complementary models. IEEE Trans. Automat. Control 29 (1984), no. 9, 803–811.

[2] Adams, M.B.; Willsky, A.S.; Levy, B.C. Linear estimation of boundary value stochastic processes. II. 1-D smoothing problems. IEEE Trans. Automat. Control 29 (1984), no. 9, 811–821.

[3] M. Egerstedt and C.F. Martin. Optimal control and monotone smoothing splines. Symposium on New Trends in Nonlinear Dynamics, Springer Verlag. To appear 2003.

[4] Krener, A.J. Boundary value linear systems. Systems analysis (Conf., Bordeaux, 1978), pp. 149–165, Astrisque, 75-76, Soc. Math. France, Paris, 1980.

[5] Luenberger, D.G. Optimization by vector space methods. John Wiley & Sons, Inc., New York-London-Sydney 1969

[6] Martin, C. F., Sun, S., Egerstedt, M. Optimal control, statistics and path planning. Computation and Control, VI (Bozeman, MT, 1998). Math. Comput. Modelling 33 (2001), no. 1-3, 237–253.

[7] Martin, R.F., Consumption, durable goods, and transaction costs, Dissertation, University of Chicago, 2002

[8] D. Sheya, M. Martindale, L. Gage, State of California, Department of Finance, Historical City/County Population Estimates, 1991-1998, with 1990 Census Counts. Sacramento, California, May 1998

[9] Smith, L. M. 2003. Playas of the Great Plains. University of Texas Press, Austin.

[10] Smith, J., Martin, C.

[11] Sun, S., Egerstedt, M. B., Martin, C. F. Control theoretic smoothing splines. IEEE Trans. Automat. Control 45 (2000), no. 12, 2271–2279.

[12] Wahba, G., Spline models for observational data. CBMS-NSF Regional Conference Series in Applied Mathematics, 59. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1990.

[13] Zhang, Z., Tomlinson, J., Martin, C., Splines and linear control theory, Acta Applican- dae Mathematicae, 49, pp.1-34, 1997.

[14] Zhou, Y., Dayawansa, W., Martin, C., Control theoretic smoothing splines are approximate linear filters, submitted

[15] Zhou, Y., Martin, C., A regularized solution to the Birkhoff interpolation problem, Commun. Inf. Syst. (2004)

[16] Zhou, Y., Egerstedt, M., Martin, C. Optimal approximation of functions. Commun. Inf. Syst. 1 (2001), no. 1, 101–112.

[17] Zhou, Y., Egersted, M. Martin, C. Optimal trajectories between affine subspaces. submitted.