Copyright © 1999 IFAC
14th Triennial World Congress, Beijing, P.R. China
315
D-2b-10-1

TRAJECTORY PLANNING WITH SMOOTHING SPLINES

Shan Sun *Magnus Egerstedt ** Clyde F. Martin ***

* *Department of Mathematics and Statistics*
Texas Tech University
Lubbock, TX 79409, USA
** *Department of Optimization and Systems*
KTH
Stockholm, Sweden
*** *Department of Mathematics and Statistics*
Texas Tech University
Lubbock, TX 79409, USA

1. INTRODUCTION

In this paper we continue the development begun in [Zhang, et al] of splines related to linear control systems. We refer to [Zhang, et al] for the rationale for the development. Classical polynomial splines and the splines developed in [Zhang, et al] are interpolating splines; that is, they are required to pass through specific points at specific times. In most applications, including trajectory planning, this is overly restrictive. We are usually content if the trajectory passes close to an assigned point at an assigned time. In this paper we will develop such generalized splines which are reflective of the dynamics of the underlying system.

Grace Wahaba, [Wahaba] and her school have developed a theory of splines for statistics and they long ago recognized that interpolating splines are overly restrictive. In statistics a theory of smoothing splines, see [Silverman], has developed. These are almost exactly what are needed in many applications. These approximations are indeed splines but the nodal points are determined by the algorithm rather than being predetermined. All that is guaranteed is that the curves pass close to the desired points. We continue that development using the basic tool of the control of linear systems as was developed in [Zhang, et al].

The outline of the paper is as follows. In Section 2 we derive the basic spline functions. In Section 3 we show that by a change of basis we can develop numerically stable algorithms for the calculation of these splines. In Section 4 we prove a convergence result for smoothing splines using results form the the theory of numerical quadrature. These results seem to be new even in the polynomial smoothing spline case. Detailed results about the rates of convergence are beyond the scope of this paper.

2. DERIVATION

We consider the linear system

$$\begin{aligned} \dot{x} &= Ax + bu \\ y &= cx, \quad x \in R^n. \end{aligned} \tag{1}$$

We further assume that

$$cb = cAb = \cdots = cA^{k-2}b = 0$$

for some $k \leq n$ and we assume that the system is both controllable and observable. Our goal is to produce a control law $u(t)$ which drives a trajectory close to a sequence of set points at fixed times while maintaining control of the growth of the control function $u(t)$. We will label the points and times as

$$\{(t_i, \alpha_i): \quad i = 1, \cdots, m\} \tag{2}$$

where we assume that

$$0 < t_1 < t_2 < \cdots < t_m \leq T.$$

ISBN: 0 08 043248 4

Let

$$g_{t_i}(t) = \begin{cases} ce^{A(t_i-t)}b & t < t_i \\ 0 & \text{otherwise} \end{cases}$$

and let

$$\beta_i = ce^{At_i}x_0.$$

We also define a set of linear functionals as

$$L_{t_i}(u) = \int_0^T g_{t_i}(t)u(t)dt.$$

We now formulate a cost function of the form

$$J(u) = \sum_{i=1}^m w_i(y(t_i) - \alpha_i)^2 + \rho \int_0^T u(t)^2 dt \quad (3)$$

where the weights w_i and ρ are chosen to be strictly positive. Our goal is to minimize the quadratic functional J subject to the affine constraint

$$y(t) = ce^{At}x_0 + \int_0^t ce^{A(t-s)}bu(s)ds.$$

Now using the notation that we have developed we have

$$y(t_i) = \beta_i + \int_0^T g_{t_i}(t)u(t)dt = \beta_i + L_{t_i}(u).$$

We substitute this into equation (3) to obtain

$$J(u) = \sum_{i=1}^m w_i(L_{t_i}(u) + \beta_i - \alpha_i)^2 + \rho \int_0^T u(t)^2 dt$$

and letting $\beta_i - \alpha_i = \gamma_i$ we have

$$J(u) = \sum_{i=1}^m w_i(L_{t_i}(u) + \gamma_i)^2 + \rho \int_0^T u^2(t)dt \quad (4)$$

Our goal is to minimize this functional over the Hilbert space of square integrable functions on the interval $[0, T]$. We calculate the Frechet derivative in the form

$$\begin{aligned} &\lim_{\alpha \to 0} \frac{1}{\alpha}(J(u + \alpha h) - J(u)) \\ &= \sum_{i=1}^m 2w_i L_{t_i}(h)(L_{t_i}(u) + \gamma_i) + 2\rho \int_0^T h(t)u(t)dt \\ &= 2\int_0^T [\sum_{i=1}^m w_i g_{t_i}(t)(L_{t_i}(u) + \gamma_i) + \rho u(t)]h(t)dt. \end{aligned}$$

Now to ensure that u is a minimum we must have that the Frechet derivative vanishes but this can only happen if

$$\sum_{i=1}^m w_i g_{t_i}(t)(L_{t_i}(u) + \gamma_i) + \rho u(t) = 0. \quad (5)$$

We now consider the operator

$$T(u) = \sum_{i=1}^m w_i g_{t_i}(t) L_{t_i}(u) + \rho u(t).$$

We can rewrite this operator in the following form

$$T(u) = \int_0^T (\sum_{i=1}^m w_i g_{t_i}(t) g_{t_i}(s))u(s)ds + \rho u(t). \quad (6)$$

Our goal is to show that the operator T is one to one and onto.

Lemma 2.1. The set of functions $\{g_{t_i}(t) : \ i = 1, \cdots, m\}$ are linearly independent.

Proof: The proof is obvious.

We now establish that the operator T is one to one.

Lemma 2.2. The operator T is one to one for all choices of $w_i > 0$ and $\rho > 0$.

Proof: Suppose $T(u_0) = 0$ then this implies that

$$\sum_{i=1}^m w_i g_{t_i}(t) L_{t_i}(u_0) + \rho u_0(t) = 0$$

and hence that

$$\sum_{i=1}^m w_i g_{t_i}(t) a_i + \rho u_0(t) = 0,$$

where a_i is the constant $L_{t_i}(u_0)$. This implies that any solution u_0 of $T(u_0) = 0$ is in the span of the set $\{g_{t_i}(t) : \ i = 1, \cdots, m\}$. Now consider a solution of the form

$$u_0(t) = \sum_{i=1}^m \tau_i g_{t_i}(t)$$

and evaluate $T(u_0)$ to obtain

$$\sum_{i=1}^m w_i g_{t_i}(t) L_{t_i}(\sum_{j=1}^m \tau_j g_j(t)) + \rho \sum_{i=1}^m \tau_i g_{t_i}(t) = 0.$$

Thus for each i

$$w_i \sum_{j=1}^m L_{t_i}(g_j)\tau_j + \rho\tau_i = 0.$$

The coefficient τ is then the solution of a set of linear equations of the form

$$(DG + \rho I)\tau = 0$$

where D is the diagonal matrix of the weights w_i and G is the Grammian with $g_{ij} = L_{t_i}(g_j)$. Now consider the matrix $DG + \rho I$ and multiply on the left by D^{-1} and consider the scalar

$$x^T(G + \rho D^{-1})x = x^T Gx + \rho x^T D^{-1}x > 0$$

since both terms are positive. Thus for positive weights and positive ρ the only solution is $\tau = 0$. It remains to show that the operator T is onto.

ISBN: 0 08 043248 4

Lemma 2.3. For $\rho > 0$ and $w_i > 0$ the operator T is onto.

Proof: Suppose T is not onto. Then there exists a nonzero function f such that $\int_0^T f(t)T(u)(t)dt = 0$ for all u. We have after some manipulation

$$\int_0^T f(t)T(u)(t)dt =$$

$$\int_0^T \left[\int_0^T \sum_{i=0}^{m} w_i g_{t_i}(t) g_{t_i}(s) f(t)dt + \rho f(s) \right] u(s)ds = 0$$

and hence that

$$\int_0^T \sum_{i=0}^{m} w_i g_{t_i}(t) g_{t_i}(s) f(t)dt + \rho f(s) = 0.$$

By the previous lemma the only solution of this equation is $f = 0$ and hence T is onto. We have proved the following proposition.

Proposition 2.1. The functional

$$J(u) = \sum_{i=1}^{m} w_i (L_{t_i}(u) + \gamma_i)^2 + \rho \int_0^T u^2(t)dt$$

has a unique minimum.

We now use equation (5) to find the optimal solution. As in the proof that T is one to one we look for a solution of the form

$$u(t) = \sum_{i=1}^{m} \tau_i g_{t_i}(t).$$

Substituting this into equation (5) we have upon equating coefficients of the $g_{t_i}(t)$ the system of linear equations

$$(DG + \rho I)\tau = D\gamma \tag{7}$$

where γ is the vector of γ_is from equation 4. As in the proof of the lemma the coefficient matrix is invertible and hence the solution exists and is unique. The resulting curve $y(t)$ is a spline. The major difference is that the nodal points are determined by the optimization instead of being predetermined. It is interesting to calculate the difference between the nodal points γ_k and the values of the spline function. We have that

$$L_t(u_N) = \sum_{i=1}^{N} \tau_i L_t(g_{t_i} = \sum_{i=1}^{N} \int_0^T g_t(s) g_{t_i}(s)ds.$$

Now evaluating at $t = t_k$ we have

$$L_{t_k}(u_N) = \sum_{i=1}^{N} \tau_i \int_0^T g_{t_k}(s) g_{t_i}(s)ds =$$

$$e_k^T G\tau = e_k^T[\gamma - \rho D^{-1}\tau]$$

and thus

$$L_{t_k}(u_N) - \gamma_k = -\rho w_k^{-1}\tau_k.$$

Now τ is a linear function of the data γ and hence the spline depends linearly on the data.

Inverting the matrix $DG + \rho I$ is not trivial. Since it is a Grammian we can expect it to be badly conditioned. In the next section we will show that by a change of basis we can produce a much better conditioned system.

3. CONVERGENCE

In this section we will prove that the smoothing splines defined above converge in a statistical sense. We make a few assumptions which we believe are necessary.

Assumption 3.1. For the system of equation 1 we assume $cb = \cdots = cA^{n-2}b = 0$

Assumption 3.2. The matrix A has only real eigenvalues.

Assumption 3.3. The set $\{t_i : i = 1, 2, \cdots\}$ are contained in the interval $[0, T]$.

Assumption 3.4. Let $f(t)$ be a C^∞ function on an interval that contains $[0, T]$.

Let u_N be the control that optimizes the functional

$$J_N(u) = \frac{1}{2N} \sum_{i=1}^{N} w_{iN}(L_{t_{iN}}(u) - f(t_{iN}))^2 \tag{8}$$

$$+\frac{\rho}{2} \int_0^T u(t)^2 dt \tag{9}$$

and let u^* be the control that optimizes

$$J(u) = \frac{1}{2} \int_0^T (L_t(u) - f(t))^2 dt + \frac{\rho}{2} \int_0^T u^2(t)dt \tag{10}$$

We make the following important assumption:

Assumption 3.5. The sequence of quadratures defined by the numbers w_{iN}, t_{iN} converge for all continuous functions defined on $[0, T]$, i.e.

$$\lim_{N\to\infty} \frac{1}{2N} \sum_{i=1}^{N} w_{iN} h(t_{iN}) = \frac{1}{2} \int_0^T h(t)dt$$

We will prove the following theorem.

ISBN: 0 08 043248 4

Theorem 3.1. Under the assumptions 4.1-4.5 the sequence of controls $\{u_N(t)\}_{m=1}^{\infty}$ converges to the function $u^*(t)$ in the L_2 norm and the sequence of smoothing splines $\{L_t(u_N)\}_{N=1}^{\infty}$ likewise converges to $L_t(u^*)$ in L_2 norm.

Proof of theorem: We begin the proof by showing that u^* exists and is unique. We show this by explicit construction. We have shown previously that the functions u_N exist and are unique. Then we will argue that the minimizers of the functionals J_N converge to the minimizer of J. We divide the proof into a series of lemmas whose proffs are omitted for the sake of space.

Lemma 3.1. The function u^* exists and is unique.

Lemma 3.2. The optimal spline $L_t(u^*)$ is given by

$$L_t(u^*) = \begin{pmatrix} e_1^T & 0 \end{pmatrix} \exp\left(\begin{pmatrix} A & -e_n e_n^T \\ -e_1 e_1^T & -A^T \end{pmatrix} t\right) \begin{pmatrix} \hat{w}(0) \\ \lambda(0) \end{pmatrix}$$

$$-\int_0^t \begin{pmatrix} e_1^T & 0 \end{pmatrix} \exp\left(\begin{pmatrix} A & -e_n e_n^T \\ -e_1 e_1^T & -A^T \end{pmatrix} (t-s)\right) \begin{pmatrix} e_n \\ 0 \end{pmatrix} F(s)ds$$

and the optimal control is given by

$$u^*(t) = e_n^T \lambda(t)$$

where

$$\frac{d}{dt}\begin{pmatrix} \hat{w}(t) \\ \lambda(t) \end{pmatrix} = \begin{pmatrix} A & -e_n e_n^T \\ -e_1 e_1^T & -A^T \end{pmatrix} \begin{pmatrix} \hat{w}(t) \\ \lambda(t) \end{pmatrix} - \begin{pmatrix} e_n \\ 0 \end{pmatrix} F(t)$$

with data

$$\hat{w}(0) = w_0 \quad \lambda(T) = 0.$$

We now finish the proof of the theorem by proving convergence. From our assumption that f is C^∞, the control u^* is at least C^∞. We know that the minimizer of $J(u)$ is unique and we have show that the minimizer of $J_N(u)$ is unique. We also know from the general theory of optimization [Luenberger] that the minimizer of a quadratic functional is given by the unique zero of the Frechet derivative of the functional. Calculating the Frechet derivatives we have the following two linear functionals, $DJ_N(u)$ and $DJ(u)$. It is clear that for each u and w that $DJ_N(u)(w)$ converges to $DJ(u)(w)$ provided the quadrature scheme converges for a sufficiently general class of functions. We now rewrite the Frechet derivatives in terms of inner products by the simply expediency of interchanging the order of integration. We have that the convergence is independent of w since

$$\sum_{i=1}^{N} g_{t_{iN}}(s) w_{iN} (L_{t_{iN}}(u) - f(t_{iN})) + u(s) \quad (12)$$

converges to

$$\int_0^T g_t(s)(L_t(u) - f(t))dt + u(s) \quad (13)$$

for every $s \in [0, T]$. We are now concerned with the convergence of linear operators rather than linear functionals.

Let

$$B(s)(u) = \int_0^T g_t(s) L_t(u) dt + u(s) \quad (14)$$

and let

$$B_N(s)(u) = \sum_{i=1}^{N} g_{t_{iN}}(s) w_{iN} L_{t_{iN}}(u) + u(s) \quad (15)$$

Let

$$b(s) = \int_0^T g_t(s) f(t) dt$$

and let

$$b_N(s) = \sum_{i=1}^{N} g_{t_{iN}} w_{iN} f(t_{iN}).$$

Now it is clear that $b_N(s)$ converges to $b(s)$ pointwise and hence in L_2 norm. Thus given ϵ for N sufficiently large we have

$$|B_N(s)(u_N - u^*)| < \epsilon$$

Now we know that $B_N(s)x = b_N(s)$ has a unique solution and hence that $B_N(s)$ is non singular. We can thus conclude that

$$u_N(s) - u^*(s)$$

converges to 0 pointwise and hence in L_2 norm. Thus the theorem is proven.

Comments: Lemma 4.2 is important for it shows exactly how the continuous spline is dependent on the data. We see from the theorem that the spline is the convolution of the function F with a kernel that is the semigroup of a Hamiltonian system. We also see that the that since the control is optimal with respect to the cost function that the resulting feedback controlled system is stable and hence that perturbations in F are not blown up but die quite quickly. this however is really just straight forward results from the theory of linear quadratic optimal control.

We obtain as corollaries three important results.

Corollary 3.1. Let $w_{iN} = \frac{1}{N}$ and let the sequence of t_i's be the observed values of a random variable uniformly distributed in the interval $[0, T]$ then the sequence of smoothing splines $\{L_t(u_N)\}_{N=1}^{\infty}$ converges to $L_t(u^*)$ in L^2 norm.

ISBN: 0 08 043248 4

Corollary 3.2. Let $w_{iN} = \frac{1}{N}$ and let $t_{iN} = \frac{iT}{N}$ (Riemann Sum) then the sequence of smoothing splines $\{L_t(u_N)\}_{N=1}^{\infty}$ converges to $L_t(u^*)$ in L^2 norm.

Corollary 3.3. Let w_{iN}, t_{iN} be defined by a Gaussian quadrature scheme then the sequence of smoothing splines $\{L_t(u_N)\}_{N=1}^{\infty}$ converges to $L_t(u^*)$ in L^2 norm.

Proof of Corollaries: The corollaries are presented in the order of the rate of convergence. We begin with Corollary 3.1. This result is based on the "law of large numbers" and while the rate of convergence is painfully slow there are minimal assumptions about the location of points making it an extremely useful result. We state this standard result of ease of reference.

Theorem 3.2. (Law of Large Numbers). Let $x_1, x_2, \cdots, x_n$ be independent and identically distributed random variables whose probability density function is denoted by $\mu(x)$. Let

$$I = \int_{-\infty}^{\infty} f(x)\mu(x)dx$$

exist, then

$$\frac{1}{n}\sum_{i=1}^{n} f(x_i)$$

converges to I in probability with n.

For Corollary 3.1 we have taken

$$\mu(x) = \begin{cases} \frac{1}{T} & 0 \le x \le T \\ 0 & \text{otherwise} \end{cases}$$

in the theorem. It is worth noting that we could have modified the cost function $J(u)$ to obtain a more general result. If we take

$$J(u) = \int_0^{\infty} (L_t(u) - f(t))^2 \mu(t)dt + \int_0^{\infty} u^2(t)dt \qquad (16)$$

most of the previous result still hold. However there are problems that need to be resolved such as what is the nature of the optimal solution and how does one interpret splines on an infinite interval. There are certain aspects of this that will be investigated in a future paper. For example if $\mu(t) = \lambda e^{-\lambda t}$ this would produce a spline like smoothing function which would have some predictive value for future times which of course is one of the more serious drawbacks of spline approximation.

For Corollary 3.2 we appeal to the general theory of Riemann sums and use the following theorem which can be found in Davis's classical book, [Davis].

Theorem 3.3. (Convergence of Riemann Sums). Let f(x) be continuous in $[a, b]$. then

$$\left| \int_a^b f(x)dx - h\sum_{i=1}^{n} f(a+kh) \right| \le (b-a)w(\frac{b-a}{n})$$

where

$$w(\delta) = \max_{x_1 - x_2 \le \delta} |f(x_1) - f(x_2)|, \quad a \le x_1, x_2 \le b.$$

There are various refinements of this theorem in the literature and we refer the reader to [Davis and Rabinowitz] for a survey and interpretation of the literature. For this result the rate of convergence is of the order of $\frac{1}{n}$ which is an improvement over the rate of convergence given by the law of large numbers which is only $\frac{1}{\sqrt{n}}$. There are various improvement which can be made along this line. For example we can use multiple point trapezoidal rules and multiple point Simpson's rules to obtain polynomial convergence of various orders. Again we refer the reader to [Davis and Rabinowitz] for many examples.

The results for Gaussian quadrature are quite diverse. Technically we have used the following thereon related to Legendre quadrature.

Theorem 3.4. (Legendre Quadrature). Let

$$E_n(f) = \int_0^T f(x)dx - \sum_{k=1}^{n} w_k f(x_k)$$

where the x_k are the zeros of the Legendre polynomial of degree n and the weights w_k are the weights of the associated quadrature scheme, then

$$E_n(f) = \frac{2^{2n+1}(n!)^4}{(2n+1)[(2n)!]^3} f^{(2n)}(\eta).$$

From this result we see that the order of convergence is non polynomial. We also see that it becomes harder and harder to give precise estimates because of the difficulty of estimating the higher derivatives of f. In [Davis and Rabinowitz] there are numerous results on the rates of convergence of some of the classical quadrature schemes but there does not seem to be a general procedure for finding rates for arbitrary weight functions.

ISBN: 0 08 043248 4

4. CONCLUSION

Polynomial smoothing splines have played an important role in nonparametric statistics. When data is corrupted by noise interpolating splines fail to represent the underlying process because of their inherent variation. Smoothing splines overcome this drawback at the expense of the introduction of more complexity and at the expense of obscuring the underlying process by the addition of the smoothing factor. In this paper we have shown that the smoothing splines converge to a very natural object that can be represented as the output of a forced Hamiltonian system associated with the continuous version of the cost function of the optimal control system.

This paper makes two contributions. The first is to show that the smoothing splines have a natural object to which they converge. This is parallel to the result for interpolating deterministic splines. The second contribution is to show that polynomial smoothing splines can naturally be considered as special case of a more general class of splines associated with linear control theory. This is important in cases in which the underlying process is dynamic.

In this paper we have shown that the discrete spline converges to the continuous analog and we have shown that the rate of convergence of the cost function is dependent on the rate of convergence of the corresponding quadrature method. The quadrature method influences the spline through the choice of points and through the choice of weights. We have not studied how the choice of smoothing parameter enters into this calculation nor have we studied the rate of convergence of the spline itself. These topics are beyond the scope of this paper.

References

Zhang, Z., Tomlinson, J., and Martin, C. Splines and linear control theory, *Acta Math. Appl.* **49** 1-34, 1997

Davis, P. and Rabinowitz, P. **Methods of Numerical Integration**, Academic Press, New York, 1984

Davis, P. **Interpolation and Approximation**, Ginn(Blaisdell), Boston, 1963

Wahaba, G. **Spline models for observational data**, CBMS-NSF Regional Conference Series in Applied Mathematics, SIAM, Philadelphia, 1990

Silverman, B. Some aspects of the spline smoothing approach to non-parametric regression curve fitting, *J.R.Statist.Soc. B.* 1985, 1-52

Doolin, B., and Martin, C. **Introduction to Differential Geometry of Control Engineers**, Marcel Dekker, New York, 1990

Martin, C. Finite escape time for riccati differential equations, *Systems and Control Letters*, **1**, (1981), 121–131.

Luenberger, D. **Optimization by Vector Space Methods**, Wiley, New York, 1969

ISBN: 0 08 043248 4