

Identification of a Hammerstein Model of the Stretch Reflex EMG using Cubic Splines

Erika J. Dempsey and David T. Westwick

Dept. Elec. & Comp. Eng., Univ. Calgary, 2500 University Dr. NW, Calgary, AB, T2N 1N4, Canada

Abstract—The use of cubic splines, instead of polynomials, in representing static nonlinearities in block structured models is considered. A system identification algorithm for the Hammerstein structure, a static nonlinearity followed by a linear filter, is developed in which the static nonlinearity is represented by a cubic spline. The identification algorithm, based on a separable least squares Levenberg-Marquardt optimization, is used to identify a Hammerstein model of the stretch reflex EMG recorded from a spinal cord injured patient. The resulting model provides more accurate predictions of the reflex EMG, even in novel data, than more conventional models which use polynomial representations of the nonlinearity. Furthermore, the spline based optimization appears to be less sensitive to its initialization than a similar polynomial-based approach.

Keywords— nonlinear system identification, physiological modeling, separable least squares, mean squared optimization, Levenberg-Marquardt iteration

I. INTRODUCTION

The Hammerstein cascade [10], a static nonlinearity followed by a linear filter, is often used to represent certain highly nonlinear systems. It is particularly useful for modeling systems, such as the stretch reflex, which contain “hard” nonlinearities, e.g. rectification and/or thresholding. Indeed, Hammerstein cascades can model some nonlinear systems, specifically those whose underlying structure is appropriate, using far fewer parameters than an equally accurate Volterra or Weiner series model.

Block structured models often use polynomials as static nonlinearities. Polynomials are linear in their parameters, and hence easy to estimate. Furthermore, fairly extreme nonlinearities can be represented using relatively few parameters. Finally, there is a direct mathematical relationship between the polynomial coefficients and corresponding Volterra kernels [10].

Several problems, however, have been identified with polynomial estimation, especially in regions where data is sparse [3]. Small disturbances in the data can produce significant differences in the interpolated values. High order polynomial solutions can also produce undesirable fluctuations, as evident in the polynomial shown in the upper panel of Fig. 3. Finally, polynomial extrapolation is notoriously difficult.

For these reasons, splines are often used instead of polynomials for function approximation [3]. Indeed, several researchers have suggested, but not demonstrated, the use of splines in nonlinear system identification [1], [2]. Unlike polynomials, fitting splines requires a nonlinear optimization. This added difficulty is most likely the reason for their limited application to date. Interest, however, appears to be growing in neural networks applications, where the replacement of sigmoidal nonlinearities with adaptive splines is being considered [4].

In this paper, we will develop an identification algorithm for

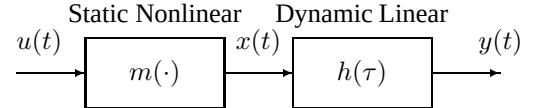


Fig. 1. Block diagram of a Hammerstein structure, consisting of a static nonlinearity, followed by a dynamic linear system.

Hammerstein systems in which the nonlinearity is represented by a cubic spline. The resulting technique will be used to identify models of the stretch reflex EMG from experimental data. Finally, identified models incorporating both polynomial and cubic spline nonlinearities will be compared.

II. THEORY

The Hammerstein model, shown in Fig. 1, consists of a static nonlinearity followed by a dynamic linear system. Here, $u(t)$ and $y(t)$ are its input and output, and $x(t)$ is the signal between the nonlinear and linear elements.

The linear filter will be represented by its impulse response (IRF), $h(\tau)$, which will have a memory length of T samples. Its output is computed using the convolution sum,

$$y(t) = \sum_{\tau=0}^{T-1} h(\tau)x(t-\tau) \quad (1)$$

The static nonlinearity, $m(\cdot)$, maps the input, $u(t)$, to the intermediate signal, $x(t)$. Traditionally, it is modeled with an order Q polynomial, in which case,

$$x(t) = m(u(t)) = \sum_{q=0}^Q c_q u^q(t) \quad (2)$$

where c_q is the q 'th order polynomial coefficient. Regardless of how the nonlinearity is represented, the output of a Hammerstein cascade can be written as,

$$y(t) = \sum_{\tau=0}^{T-1} h(\tau)m(u(t-\tau)) \quad (3)$$

In the special case of a *polynomial Hammerstein* model, where $m(\cdot)$ is parameterized with a polynomial, (3) becomes,

$$y(t) = \sum_{\tau=0}^{T-1} \sum_{q=0}^Q c_q h(\tau) u^q(t-\tau) \quad (4)$$

Report Documentation Page

Report Date 25OCT2001	Report Type N/A	Dates Covered (from... to) -
Title and Subtitle Identification of a Hammerstein Model of the Stretch Reflex EMG using Cubic Splines		Contract Number
		Grant Number
		Program Element Number
Author(s)	Project Number	
	Task Number	
	Work Unit Number	
Performing Organization Name(s) and Address(es) Dept. Elec. & Comp. Eng., Univ. Calgary, 2500 University Dr. NW, Calgary, AB, T2N 1N4, Canada		Performing Organization Report Number
Sponsoring/Monitoring Agency Name(s) and Address(es) US Army Research Development & Standardization Group (UK) PSC 803 Box 15 FPO AE 09499-1500		Sponsor/Monitor's Acronym(s)
		Sponsor/Monitor's Report Number(s)
Distribution/Availability Statement Approved for public release, distribution unlimited		
Supplementary Notes Papers from the 23rd Annual International conference of the IEEE Engineering in Medicine and Biology Society, October 25-28, 2001, held in Istanbul, Turkey. See also ADM001351 for entire conference on cd-rom.		
Abstract		
Subject Terms		
Report Classification unclassified	Classification of this page unclassified	
Classification of Abstract unclassified	Limitation of Abstract UU	
Number of Pages 4		

Techniques based on separable least squares (SLS) optimizations [9] have been developed for the identification of polynomial Hammerstein models [11]. Briefly, the Hammerstein cascade is represented by a parameter vector containing the filter weights and polynomial coefficients,

$$\begin{aligned}\theta &= [h(0) \ \dots \ h(T-1) \mid c_0 \ \dots \ c_Q]^T \\ &= [\theta_l^T \mid \theta_n^T]^T\end{aligned}\quad (5)$$

Its output is written $\hat{y}(t, \theta)$, explicitly showing the dependence on the parameter vector. Thus, $\hat{y}(t, \theta)$ is computed using (4), where the IRF and polynomial coefficients are taken from θ , defined in (5). Furthermore, identifying the model is equivalent to finding the parameter vector that minimizes the MSE,

$$V_N(\theta) = \frac{1}{N} \sum_{t=1}^N (y(t) - \hat{y}(t, \theta))^2$$

The SLS identification algorithm [11] hinges on the observation that for any given choice of polynomial coefficients, the output (4) is a linear function of the filter weights. This is reflected in (5), where θ is divided into linear and nonlinear parameters, θ_l and θ_n respectively. Since (4) is linear in θ_l , its optimal value, corresponding to any choice of polynomial coefficients, θ_n , can be found in closed form by solving a linear regression. Thus, θ_l is a function of θ_n , written $\theta_l(\theta_n)$. Similarly $\hat{y}(\theta_n)$, and $V_N(\theta_n)$ are also functions of the nonlinear parameters alone. Thus, an iterative, nonlinear, optimization is only needed to find θ_n . In this paper, we use the Levenberg-Marquardt algorithm. Thus, θ_n is updated using:

$$\theta_n^{(k+1)} = \theta_n^{(k)} + (J_s^T J_s + \delta_k I)^{-1} J_s^T \epsilon \quad (6)$$

where ϵ is a vector whose t 'th element is the error $\epsilon(t) = y(t) - \hat{y}(t)$, I is an identity matrix, and δ_k is a regularization parameter used to control the convergence rate and stability. J_s , the Jacobian, is matrix whose $[t, m]$ entry contains the partial derivative of the model output at time t with respect to the m 'th element of θ_n . Thus,

$$J_s(t, m) = \frac{\partial \hat{y}(t)}{\partial \theta_n(m)} \quad (7)$$

Note, however, that in computing J_s , the dependence of θ_l , the linear parameters, on θ_n must be taken into consideration. This can be done as follows. First compute J_l and J_n , the Jacobians with respect to the linear and nonlinear parameters, assuming that the parameters are independent of each other. Then, it can be shown [9] that

$$J_s = (I - P_l) J_n \quad (8)$$

where $P_l = J_l(J_l^T J_l)^{-1} J_l^T$ is an orthogonal projection onto the columns of the linear Jacobian, J_l .

Thus, to use the SLS algorithm, compute the Jacobian in the usual way, and then separate it into linear and nonlinear parts. Insert these into (8) to compute J_s , used in the L-M step (6).

A. Use of Cubic Splines

The SLS algorithm depends on the output, (4), being a linear function of the IRF weights, but does not depend on the representation of the static nonlinearity, $m(\cdot)$. All that is needed to use a different nonlinearity, (3), is to compute the nonlinear Jacobian, J_n , with respect to its parameters.

In this paper, the static nonlinearity, $x = m(u)$, will be modeled by a cubic spline instead of a polynomial. A cubic spline is defined by a series of M knots, (u_j, x_j) , for $j = 1, 2, \dots, M$, with $u_1 < u_2 < \dots < u_M$. Thus, the parameter vector describing the spline can be formed,

$$\theta_n = [u_1 \ \dots \ u_M \ x_1 \ \dots \ x_M]^T \quad (9)$$

In between each pair of knots, the spline is defined by a third degree polynomial. These are chosen so that the spline and its first two derivatives are all continuous functions. Hence, let u be a point in the input, between knots j and $j+1$, (i.e. $u_j \leq u \leq u_{j+1}$). The corresponding point in the output of the spline is computed from [8],

$$x = A(u)x_j + B(u)x_{j+1} + C(u)x_j'' + D(u)x_{j+1}'' \quad (10)$$

where $A(u)$, $B(u)$, $C(u)$ and $D(u)$ are defined by

$$\begin{aligned}A(u) &= \frac{u_{j+1} - u}{u_{j+1} - u_j} \\ B(u) &= 1 - A(u) \\ C(u) &= \frac{1}{6}(A^3(u) - A(u))(u_{j+1} - u_j)^2 \\ D(u) &= \frac{1}{6}(B^3(u) - B(u))(u_{j+1} - u_j)^2\end{aligned}$$

Note, however, that (10) depends on the second derivative of the the spline, $x'' = \partial^2 x / \partial u^2$, at the two adjacent knots. At first glance, it would appear that these second derivatives are required to define the spline completely. However, by forcing the first derivative of the spline to be continuous across the knots, one arrives at the equation [8],

$$\begin{aligned}\frac{u_j - u_{j-1}}{6} x_{j-1}'' + \frac{u_{j+1} - u_j}{3} x_j'' + \frac{u_{j+1} - u_j}{6} x_{j+1}'' \\ = \frac{x_{j+1} - x_j}{u_{j+1} - u_j} - \frac{x_j - x_{j-1}}{u_j - u_{j-1}}\end{aligned} \quad (11)$$

for $j = 2, \dots, M-1$. Thus, we have $M-2$ equations with which to define x'' at M locations, which is an under-determined problem. Typically, this is resolved by arbitrarily choosing the second derivatives at the two end points. In this paper, *natural* splines will be employed, where the second derivative is set to zero at the first and last knots. Thus (11) can be used to solve for x'' at the remaining $M-2$ knots. Once these have been computed, (10) can be used to compute the output of the cubic spline.

To use a cubic spline, instead of a polynomial, in the SLS Hammerstein identification algorithm, one need only compute the appropriate nonlinear Jacobian,

$$J_n = \left[\frac{\partial x(t)}{\partial u_1}, \ \dots, \ \frac{\partial x(t)}{\partial u_M}, \ \frac{\partial x(t)}{\partial x_1}, \ \dots, \ \frac{\partial x(t)}{\partial x_M} \right] \quad (12)$$

and then insert it into (8) to compute the separated Jacobian in the Levenberg-Marquardt step (6). The rest of the algorithm is used without modification.

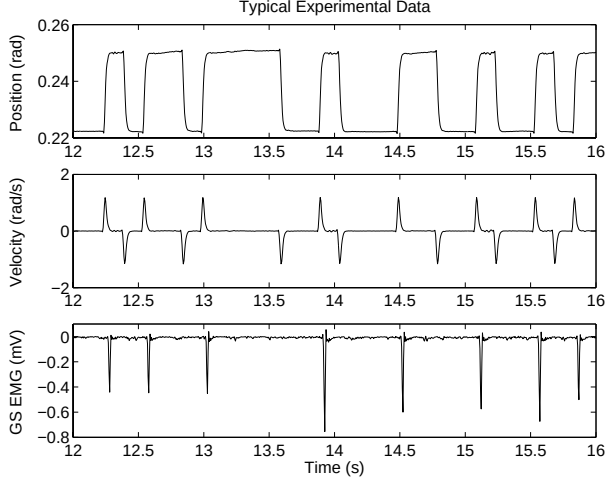


Fig. 2. Extract from a typical experimental trial showing 4 seconds of position, velocity and GS EMG.

III. EXPERIMENTAL RESULTS

To evaluate the utility of the cubic spline algorithm, we used it to estimate stretch reflex dynamics, the relationship between the ankle velocity and the resulting EMG, in spastic, spinal cord injured (SCI) patients. Previously, this system has been modeled as a Hammerstein cascade [5] in which the static non-linear element was found to resemble a half-wave rectifier and the IRF of the linear element resembled a delayed impulse. Since this model is being investigated for potential clinical use [7], methods that find efficient, unbiased estimates of its elements may prove to be very significant.

The experimental procedures have been described in detail elsewhere [6]. In summary, the subjects, due to a previous spinal cord injury, demonstrated an incomplete loss of motor function, spasticity and hyper-active stretch reflexes. They lay on their backs with their left foot attached to a rotary hydraulic actuator by a custom fitted fiber-glass boot. Ankle torque and position, and the EMG over the Gastrocnemius-Soleus(GS) muscles were recorded. A broad-band position perturbation, whose spectrum was shaped to preserve the stretch reflex [6], was applied while the subject maintained a constant background contraction. Figure 2 shows four of the 30 seconds, sampled at 200 Hz, of position, velocity and EMG data used in the analysis. Note that the velocity was calculated by numerically differentiating the measured position.

The technique developed in [11] was used to fit a polynomial Hammerstein model between the first 5000 points of the velocity and EMG records, hereafter called the training data. The starting point for this SLS optimization was the nonlinearity identified using a conventional correlation based scheme [5]. The resulting model, shown in Fig. 3 with dash-dotted lines, was then used to predict the remaining 1000 points of EMG data, the validation segment. The polynomial Hammerstein model accounted for 95.2% of the signal variance in the training sample and 94.5% in the validation set.

Next, a cubic spline Hammerstein cascade was identified

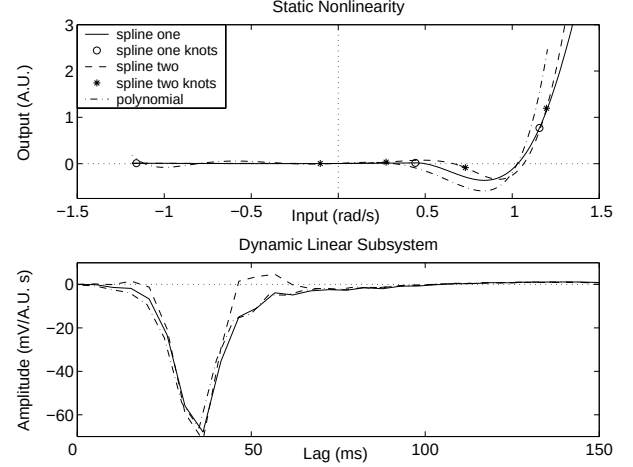


Fig. 3. Hammerstein models of the stretch reflex EMG using polynomial and cubic spline nonlinearities identified from experimental data.

from the training data. The initial spline had 5 knots along the line $m(u) = u$, equally spaced over the range of the experimental input. The SLS iteration described in this paper was then used to search for the optimal cubic spline nonlinearity and corresponding linear IRF. The results of this identification, referred to as “spline one”, are shown as solid lines in Fig. 3. The nonlinearity is shown in the upper panel: the knots are shown as circles, with the behavior between knots shown as a solid line. The prediction accuracy of this model was 95.3% VAF in the training data, and 95.6% VAF in the validation sample. Therefore, using the cubic spline algorithm decreased the residual variance by 2.1% and 25.0% in the training and validation sets, respectively.

If a general model of the system has been previously developed, points for the initial spline can be taken from its nonlinearity. For example, a cubic spline model, referred to as “spline two”, was generated from the same initial model as the polynomial cascade. The results are shown in Fig. 3 as dashed lines with the spline knots represented by asterisks. Spline two accounted for 95.5% of the variance in the training set and 96.0% in the validation set, giving a decrease of 6.7% and 37.5% of the residual variance in the two data segments with respect to the polynomial-based cascade.

Comparing the models produced by the two algorithms illustrates several potential advantages of the spline-based approach. First, its models produced more accurate predictions than did the polynomial-based scheme. Secondly, the “spline one” model resulted from an initialization, a simple linear fit, that contained no *a priori* information about the shape of the nonlinearity, suggesting that the spline-based algorithm may be less sensitive to its initialization than the polynomial approach. Furthermore, neither spline-based model included the oscillations at negative velocities present in the polynomial curve (see Fig. 3 upper panel). Finally, the dynamic linear subsystem (lower panel) of the “spline two” model shows a steeper response with a longer rest time at zero, which is more consistent with the propagation delay expected in the reflex.

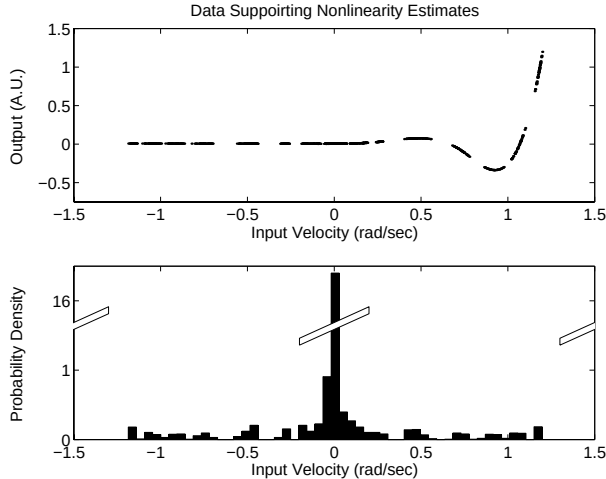


Fig. 4. Upper panel: input-output scatter plot of spline two. Lower panel: probability density histogram of input velocity.

In both the polynomial and spline nonlinearities, there is a large negative deflection between input velocities of 0.6 and 1.1 rad/sec. However, this deflection appears to be well supported by the experimental data. Indeed, the upper panel of Fig. 4 shows a scatter plot of the input and output of the nonlinearity from spline two. Since much of the deflection appears as a series of solid lines, it is evidently well supported by the training data. The lower panel shows a probability density histogram, also estimated from the training data, for the nonlinearity input. This suggests that, except near zero velocity, all points in the input are probed with almost equal frequency. Since the parameters describing a cubic spline (i.e. the knot positions) have primarily local effects on its shape, it is unlikely that the deflection is an artifact due to the analysis.

IV. DISCUSSION

In this paper, we have developed a cubic spline identification method for Hammerstein systems and used it to identify models of the stretch reflex EMG in SCI patients. Both spline-based models developed more accurate predictions than a polynomial based cascade identified from the same data. This applied to both the training and validation samples.

Iterative optimizations, such as the cubic-spline based technique developed in this study, are prone to finding sub-optimal local minima. For example, the results obtained by the polynomial method [11] are dependent on the choice of initial polynomial coefficients. In this study, see Fig.3, changing the initial spline knots produced slight variations in the final Hammerstein cascade. During the data analysis, several other initial splines were tested (results not shown). Starting with $m(u) = k$, a constant, led to a model in which the nonlinearity was offset and the IRF was highly oscillatory. This model was less accurate than even the polynomial cascade. On the other hand, when the initial knots were taken from $m(u) = u^2$ or $m(u) = u^3$, the optimization found models similar to splines one and two. Overall, the cubic spline method appeared to be

less dependent on its initial parameters than the polynomial-based algorithm. Future research will analyze the convergence behavior of the cubic spline Hammerstein algorithm.

The price of the cubic spline-based algorithm is an increased number of parameters needed to characterize the nonlinearity. A cubic spline Hammerstein cascade is still much more efficient than an equivalent Volterra series model, but uses more parameters than a polynomial Hammerstein model. In this study, a cubic spline with five knots, which has ten parameters, was compared to a sixth order polynomial, with seven coefficients. With the cubic spline initializations, including additional knots provided only a small increase in accuracy, but beginning with fewer than five resulted in significantly inferior models. Similarly, increasing the nonlinearity order had little effect on the accuracy of polynomial-based models. Finally, increasing the memory length of the IRFs led to slightly smaller residuals for both the cubic spline and the polynomial-based cascades.

These results suggest that cubic splines may be used in place of polynomials in a wide variety of block-structured models currently used to represent physiological systems. It appears that iterative identification algorithms based on splines may be less dependent on their initialization than similar polynomial-based schemes. Furthermore, models based on splines are likely to be relatively immune to the severe interpolation/extrapolation problems associated with polynomials.

ACKNOWLEDGEMENTS

Supported by the Natural Sciences and Engineering Research Council, and the U. Calgary Research Grants Office. The EMG data was provided by R.E. Kearney and M.M. Mirbagheri from the Dept. Biomed. Eng., McGill University.

REFERENCES

- [1] S.A. Billings and S.Y. Fakhouri. Identification of a class of nonlinear systems using correlation analysis. *Proc. IEE*, 125:691–697, 1978.
- [2] H.W. Chen, N. Ishii, and N. Suzumura. Structural classification of nonlinear systems by input and output measurements. *Int. J. Systems Science*, 17(5):741–774, 1986.
- [3] C. De Boor. *A practical guide to splines*. Springer-Verlag, New York, 1978.
- [4] S. Guarnieri, F. Piazza, and A. Uncini. Multilayer feedforward networks with adaptive spline activation function. *IEEE Trans. Neural Nets.*, 10(3):672–683, 1999.
- [5] R.E. Kearney and I.W. Hunter. Nonlinear identification of stretch reflex dynamics. *Ann. Biomed. Eng.*, 16:79–94, 1988.
- [6] R.E. Kearney, R.B. Stein, and L. Parameswaran. Identification of intrinsic and reflex contributions to human ankle stiffness dynamics. *IEEE Trans. Biomed. Eng.*, 44(6):493–504, 1997.
- [7] M.M. Mirbagheri, R.E. Kearney, and H. Barbeau. Stretch reflex behavior of spastic ankle under passive and active conditions. *Proc. IEEE EMBS*, 18:2325–2327, 1998.
- [8] W.H. Press, S.A. Teukolsky, W.T. Vetterling, and B.P. Flannery. *Numerical Recipes in C: The Art of Scientific Computing*. Cambridge University Press, Cambridge, 2nd ed., 1992.
- [9] J. Sjöberg and M. Viberg. Separable non-linear least squares minimization – possible improvements for neural net fitting. In *IEEE Wkshp. Neural Nets. Sig. Pro.*, 7:345–354, 1997.
- [10] D.T. Westwick and R.E. Kearney. Nonparametric identification of nonlinear biomedical systems, part I: Theory. *Crit. Rev. Biomed. Eng.*, 26:153–226, 1998.
- [11] D.T. Westwick and R.E. Kearney. Separable least squares identification of nonlinear Hammerstein models: Application to stretch reflex dynamics. *Ann. Biomed. Eng.*, To Appear 29(8), 2001.