IDENTIFICATION OF TIME-VARYING HAMMERSTEIN SYSTEMS

J. C. Ralston and A. M. Zoubir

Signal Processing Research Centre, QUT,
GPO Box 2434, Brisbane Q. 4001 Australia.
email: j.ralston@qut.edu.au

ABSTRACT

We consider the identification of systems which are both time-varying and nonlinear. This class of systems is more likely to be encountered in practice, but is often avoided due to the difficulties that arise in modelling and estimation. We attempt to address this problem by considering a new time-varying nonlinear model, the time-varying Hammerstein model, which effectively characterises time-variation and nonlinearity in a simple manner. Using this model we formulate a procedure to find least-squares estimates of the coefficients. The model is general and can be used when little is known about the time-variation of the system. In addition, we do not require that the input is stationary or Gaussian. Finally, an application to automobile knock modelling is presented, where a time-varying nonlinear model is seen to more accurately characterise the system than a time-varying linear one.

1. INTRODUCTION

Nonlinear system identification is concerned with determining an appropriate model in order to describe the relationship between the input and output signals of an unknown system. Once a candidate model has been chosen, the identification task is one of parameter estimation. System identification is important for modelling, prediction, detection, control and equilisation [1]–[4].

We are particularly interested in identifying nonlinear systems which exhibit time-varying characteristics, since they arise so frequently in control, physiology, communications and engineering. There are currently very few general identification procedures available to solve this problem. We demonstrate the use of a new time-varying nonlinear model, the *time-varying* Hammerstein model, which is able to characterise both the time-variation and the nonlinearity of the system in a simple manner.

In Section 2, we present the model and formulate the identification procedure. We discuss the approach in Section 3, and then give an example in Section 4 where we model the transmission characteristics of a combustion engine under knocking conditions.

2. FORMULATION OF THE IDENTIFICATION PROCEDURE

2.1. The Model

The Volterra series is an important non-parametric model for nonlinear system identification [3], which assumes that an analytic relationship between a random input $X(t)$ and output $Y(t)$, $t = 0, 1, \dots$ of a time-varying nonlinear system may be expressed as

$$Y(t) = \sum_{\tau_1=0}^{m} g_1(t,\tau_1)X(t-\tau_1) + \sum_{\tau_1=0}^{m}\sum_{\tau_2=0}^{m} g_2(t,\tau_1,\tau_2)X(t-\tau_1)X(t-\tau_2) + \cdots \quad (1)$$

where the set of functions $\{g_n(t,\tau_1,\dots,\tau_n)\}$, $n = 1, 2, \dots$ are the time-varying Volterra kernels, and m denotes the extent of the system's memory. The Volterra kernels have the interpretation of characterising the linear, quadratic and higher order interactions of the system, and thus are physically meaningful representations.

As the time-varying kernels consist of many coefficients, there are often difficulties in visualisation and interpretation. Additionally, the estimation of the Volterra kernels using a cross-correlation based approach requires the computation of time-varying higher order spectra. This presents severe analytic and estimation problems, particularly when the input is non-Gaussian [3].

We propose an alternative to the model in (1). We extend a model that has been widely used for time-invariant nonlinear system identification, known in the literature as the Hammerstein model (e.g. see [5]). We define a new time-varying Hammerstein model as consisting of a static nonlinear function, followed by a time-varying linear filter,

$$Y(t) = \frac{B_t(z^{-1})}{1 + A_t(z^{-1})}\psi(X(t)) + N(t), \quad (2)$$

where $X(t)$ and $Y(t)$ are the observed input and output signals respectively and $N(t)$ is stationary, white, zero-mean noise, independent to $X(t)$, and z^{-1} is the unit delay operator. ψ is a zero memory nonlinear function. We model memory with a rational filter having time-varying coefficients in order to reduce the number of parameters required

0-7803-2431-5/95 $4.00 © 1995 IEEE

to characterise the system which is expressible as,

$$\frac{B_t(z^{-1})}{1+A_t(z^{-1})} = \frac{b(t,0)+b(t,1)z^{-1}+\cdots+b(t,m)z^{-m}}{1+a(t,1)z^{-1}+\cdots+a(t,n)z^{-n}}.$$

The advantage of this model is that it is simple and capable of characterising time-varying and nonlinear behaviour. A block diagram of this model is shown in Figure 1.

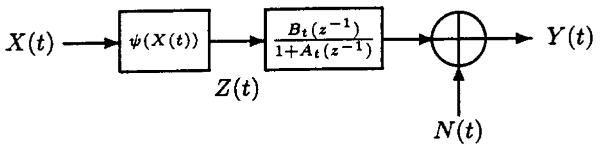

Figure 1: The time-varying Hammerstein model.

We assume that the nonlinear function can be approximated with a pth order polynomial[1] with coefficients $\{\alpha_i\}$, $i = 1, \ldots, p$. This step enables us to transform the single-input single-output nonlinear model into a multi-input single-output time-varying *linear* model. This is achieved by treating each term of the polynomial as a separate input to the time-varying filter. We can assume that α_1 is unity without loss of generality. This configuration is shown in Figure 2, where the notation $(\cdot)^i$ means $X(t)^i$.

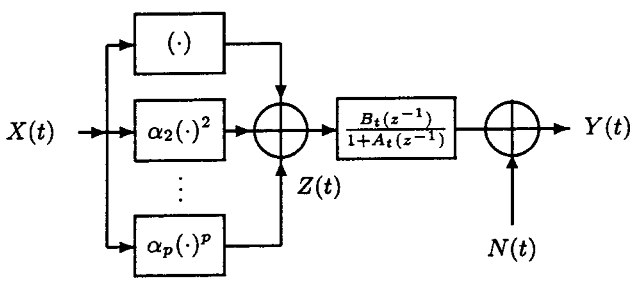

Figure 2: The multi-input Hammerstein model.

The output of the nonlinear function, $Z(t)$, can be expressed as

$$Z(t) = \sum_{i=1}^{p} \alpha_i X(t)^i.$$

Thus the total system output is

$$\begin{aligned} Y(t) &= \frac{B_t(z^{-1})}{1+A_t(z^{-1})}\psi(X(t)) + N(t) \\ &= \frac{B_t(z^{-1})}{1+A_t(z^{-1})}Z(t) + N(t). \end{aligned}$$

Re-arrangement of the above equation leads to

$$\left(1+A_t(z^{-1})\right)Y(t) = B_t(z^{-1})Z(t) + \left(1+A_t(z^{-1})\right)N(t)$$

which becomes

$$\begin{aligned} Y(t) &= B_t(z^{-1})Z(t) - A_t(z^{-1})Y(t) + E(t) \\ &= \sum_{\tau=0}^{m} b(t,\tau)Z(t-\tau) - \sum_{v=1}^{n} a(t,v)Y(t-v) + E(t) \end{aligned} \quad (3)$$

[1]Note that if ψ were known, the estimation task is straightforward given that $X(t)$ is observable.

with $E(t) = \left(1+A_t(z^{-1})\right)N(t)$. The model in (3) is nonlinear in the parameters, however we can perform a linearisation by setting $b_i(t,\tau) = \alpha_i b(t,\tau)$. We now substitute this and $Z(t)$ into (3) to yield

$$Y(t) = \sum_{i=1}^{p}\sum_{\tau=0}^{m} b_i(t,\tau)X(t-\tau)^i - \sum_{v=1}^{n} a(t,v)Y(t-v) + E(t). \quad (4)$$

The model has thus been linearised at the expense of increased parameterisation. The problem now is to estimate $b_i(t,\tau)$, $i = 1, \ldots, p$, $\tau = 0, \ldots, m$ and $a(t,v)$, $v = 1, \ldots, n$ for a nominated p, m and n.

2.2. Estimation Procedure

Our model assumes that the same basic time-varying nonlinear behaviour operates on each realisation. Given multiple realisations of the input and output processes $X_r(t)$ and $Y_r(t)$ for $r = 1, \ldots, R$ and $t = 0, \ldots, T-1$ we can write (4) for the rth realisation as

$$Y_r(t) = \sum_{i=1}^{p}\sum_{\tau=0}^{m} b_i(t,\tau)X_r(t-\tau)^i - \sum_{v=1}^{n} a(t,v)Y_r(t-v) + E_r(t)$$

We can now re-write the above as a set of linear equations

$$\begin{aligned} \mathbf{y}_t &= \begin{bmatrix} \mathbf{X}_t & \mathbf{X}_t^2 & \ldots & \mathbf{X}_t^p & \mathbf{Y}_t \end{bmatrix} \begin{bmatrix} b_1(t,\tau) \\ b_2(t,\tau) \\ \vdots \\ b_p(t,\tau) \\ -a(t,v) \end{bmatrix} + \mathbf{e}_t \\ &\equiv \mathbf{D}_t\mathbf{g}_t + \mathbf{e}_t, \end{aligned} \quad (5)$$

where $\mathbf{y}_t$ is an $[R \times 1]$ vector, $\mathbf{X}_t$ is a $[R \times (m+1)]$ matrix with the (j,k)th element $X_j(t-k)$, $\mathbf{Y}_t$ is a $[R \times m]$ matrix with (j,k)th element $Y_j(t-k)$, $\mathbf{D}_t$ is an $[R \times M]$ matrix, $\mathbf{g}_t$ is an $[M \times 1]$ vector and $M = (m+1)p + n$ and $\mathbf{e}_t$ is an $[R \times 1]$ vector for $t = 0, \ldots, T-1$. The notation $\mathbf{X}_t^p$ here means that each element of $\mathbf{X}_t$ is taken to the pth power. We now attempt to find least-squares estimates for $\mathbf{g}_t$.

We generally require that $R > M$, representing the over-determined case and thus $\mathbf{D}_t$ is non-square. The singular value decomposition allows us to find the best low rank approximation for this problem [6]. The Moore-Penrose pseudo inverse of $\mathbf{D}_t$, $\mathbf{D}_t^{\#}$, leads to least-squares estimates for $\mathbf{g}_t$ at each time instant t, $t = 0, \ldots, T-1$:

$$\hat{\mathbf{g}}_t = \mathbf{D}_t^{\#}\mathbf{y}_t.$$

Once an estimate for $\mathbf{g}_t$ has been found, the polynomial coefficients can be estimated via $\hat{\alpha}_i = \hat{b}_i(t,\tau)/\hat{b}_1(t,\tau)$ for $i = 2, \ldots, p$ over $\tau = 0, \ldots, m$ and $t = 0, \ldots, T-1$. As a result there exists an amount of redundancy in the polynomial coefficient estimates. We choose to take the median to determine the estimates for $\{\alpha_i\}$, since it is relatively robust against outliers.

3. DISCUSSION

3.1. Volterra Series Representation

Given that the time-varying linear filter can be represented by the time-varying impulse response $g(t,\tau)$, we can relate the time-varying nonlinear model to an equivalent time-varying Volterra series representation. This is done by equating like orders of the two models. The relationship between the nth order Volterra kernel, $h_n(\tau_1,\ldots,\tau_n)$, in (1) and the model in (2) is given by

$$\begin{aligned} h_n(\tau_1,\ldots,\tau_n) &= g(t,\tau_1)\delta(\tau_1-\tau_2)\cdots\delta(\tau_1-\tau_n) \\ &\equiv \alpha_n g(t,\tau_1)\prod_{q=1}^{n-1}\delta(\tau_1-\tau_{q+1}) \end{aligned}$$

Although the model in (2) is not as general as the time-varying Volterra series, it has the advantage that the model can be easily visualised and interpreted.

3.2. Reductions for Simpler Structures

We now discuss how the configuration of the model simplfies when additional structural assumptions can be made about the system.

3.2.1. Transversal and recursive filter

The conditions when $n = 0$ and $m = 0$ in (4) respectively correspond to special forms of transversal or recursive filter configurations. For the traversal filter, we have a similar form to (5), but with $\mathbf{D}_t$ a $[R \times mp]$ matrix and $\mathbf{g}_t$ a $[(m+1)p \times 1]$ vector for $t = 0,\ldots,T-1$. For the recursive filter, we have $\mathbf{D}_t$ a $[R \times n]$ matrix and $\mathbf{g}_t$ a $[n \times 1]$ vector.

3.2.2. Time-invariant nonlinear filter

For a time-invariant nonlinear model ($a(t,\tau) \equiv a(\tau)$ and $b(t,\tau) \equiv b(\tau)$) we only require one input-output observation ($R = 1$). We have $\mathbf{y} = \mathbf{D}_t\mathbf{g} + \mathbf{e}$ with $\mathbf{y}$ a $[T \times 1]$ vector, $\mathbf{D}_t$ a $[T \times l]$ vector, $\mathbf{g}$ a $[l \times 1]$ vector and $\mathbf{e}_t$ a $[T \times 1]$ vector for $t = 0,\ldots,T-1$.

3.2.3. Time-varying linear filter

If the model is linear ($p = 0$), then the problem reduces to the estimation of a single-input single-output time-varying linear filter. Thus we have $\mathbf{D}_t$ a $[R \times m]$ matrix and $\mathbf{g}_t$ a $[m \times 1]$ vector for $t = 0,\ldots,T-1$.

3.3. Attaining Input-Output Records

The identification procedure requires that the same time-varying behaviour operates on each realisation. We cannot generally segment a single input-output observation into individual records since the output is generally non-stationary. As a result obtaining records may prove more difficult in practice. However, input-output records for the identification procedure can be collected for certain classes of signals, such as cyclostationary signals [7]. Thus the method is particularly well suited to applications such as rotating machinery or periodic phenomena.

4. APPLICATION: KNOCK TRANSMISSION MODELLING

The system identification method was first verified using simulated input-output data, where we obtained good results. We now present an example to indicate the use and potential of the approach to a practical system identification problem arising from the automotive industry.

4.1. Background

It is known from spark ignition engine theory that increasing the compression ratio results in increased engine efficiency [8], but this also increases the occurrence of an abnormal combustion phenomenon called *knock*. Knock needs to be avoided as it results in an excessively noisy, over-heated and inefficient engine, and leads to premature mechanical failure. System identification techniques can be used for determining the optimal positioning of sensors for detecting the knocking condition [1].

A number of physical factors motivate the use of a time-varying nonlinear model: the motion of the piston, the rapid pressure variation during combustion, non-uniform acoustic losses during the knocking condition and nonlinear resonances varying over vapour temperature [4].

4.2. The Identification Process

We compared the time-varying Hammerstein model to a time-varying linear model with real knock data[2]. Cylinder pressure and engine vibration signals constituted the system input and output records respectively, measured under strong knocking conditions for 150 cycles over $t = 0,1,\ldots,375$. Typical input and output signals are shown in Figures 3 and 4. Note that the multiple realisations required for the identification procedure were easily obtained since the knock signals were assumed to be cyclostationary.

A second order polynomial was used as the nonlinear function, with system memory $m = 15, n = 0$. Figure 5 shows the estimated time-varying impulse response. The estimated polynomial coefficients were $\{1, .21\}$. Figure 6 compares the mean-square error between the observed and predicted outputs of the two models over the set of realisations; the dashed line corresponds to the time-varying model and the solid line to the time-varying nonlinear case. The poorer performance of the time-varying linear model suggests that a better choice of model is a time-varying nonlinear one. The results obtained here demonstrate the potential of this method for identifying time-varying nonlinear systems.

[2]**Acknowledgements**: We would like to thank Prof. J. F. Böhme from the Signal Theory Division of Ruhr University Bochum and Volkswagen AG, Wolfsburg, Germany, for kindly providing the knock data used in this paper.

5. CONCLUSIONS

We have proposed a new approach for time-varying nonlinear system identification. This procedure can be used when little is known about the time-varying dynamics of the system. The approach is simple in concept and in implementation. The method does not require that the input is stationary or Gaussian, which is an assumption that is often necessary for other system identification strategies.

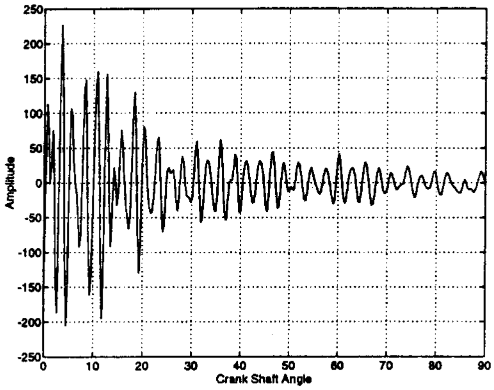

Figure 3: Cylinder pressure (Input).

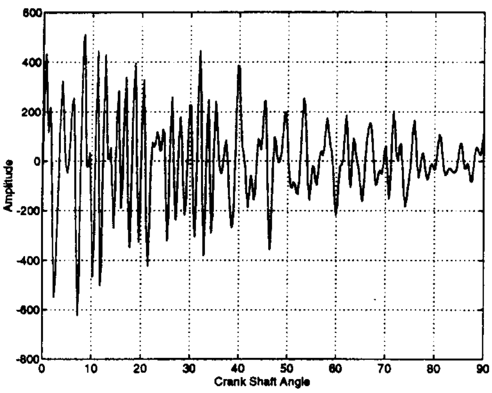

Figure 4: Engine vibration (Output).

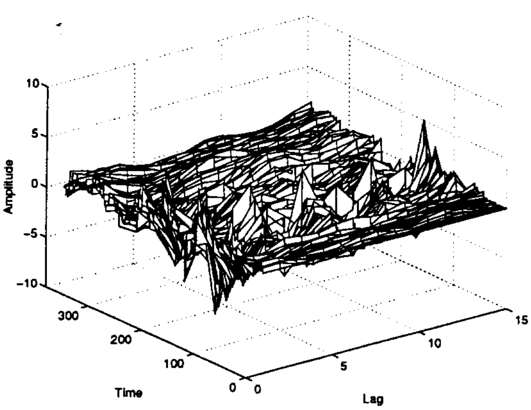

Figure 5: Estimated time-varying linear impulse response.

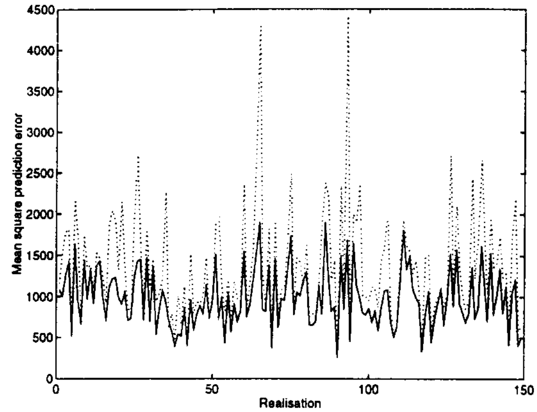

Figure 6: Comparison of mean-square prediction error for the two models – solid line: time-varying nonlinear model, dashed line: time-varying linear.

6. REFERENCES

[1] A. M. Zoubir, "Identification of second-order Volterra systems driven by non-Gaussian stationary processes," in *Proceedings of the SPIE Conference on Advanced Signal Processing Algorithms, Architectures and Implementations* (T. Luk, ed.), (San Diego, USA), pp. 327–338, Proc. SPIE, July 1992.

[2] J. S. Bendat, *Nonlinear System Analysis and Identification from Random Data.* New York: John-Wiley & Sons, 1990.

[3] M. B. Priestley, *Non-Linear and Non-Stationary Time Series Analysis.* London: Academic Press, 1988.

[4] A. M. Zoubir and J. F. Böhme, "Application of higher order spectra to the analysis and detection of knock in combustion engines," in *Higher Order Statistical Signal Processing* (B. Boashash, E. J. Powers, and A. M. Zoubir, eds.), ch. 9, Melbourne: Longman Cheshire, 1995.

[5] W. Greblicki, "Non-parametric orthogonal series identification of Hammerstein systems," *Int. J. Systems Sci.*, vol. 20, no. 12, pp. 2355–2367, 1989.

[6] L. L. Scharf, *Statistical Signal Processing.* New York: Addison-Wesley, 1991.

[7] W. A. Gardner, "Identification of systems with cyclostationary input and correlation input/output measurement noise," *IEEE Transactions on Automatic Control*, vol. 35, pp. 449–452, April 1990.

[8] T. D. Eastrop and A. McConkey, *Applied Thermodynamics for Engineering Technologists.* NY: Longman, 1970.