SAMPLING DYNAMICAL SYSTEMS

Melvin J. Hinich
Applied Research Laboratories
The University of Texas at Austin
Austin, TX 78713-8029, USA
(512) 835-3278 (Voice) 835-3259 (Fax)
hinich@mail.la.utexas.edu

ABSTRACT

Linear dynamical systems are widely used in many different fields from engineering to economics. One simple but important class of such systems is called the *single input transfer function model*. Suppose that all the variables of the system are sampled for a period using a fixed sample rate. The central issue of this paper is the determination of the smallest sampling rate that will yield a sample that will allow the investigator to identify the discrete-time representation of the system. A critical sampling rate exists that will identify the model. This rate, called the Nyquist rate, is twice the highest frequency component of the system. Sampling at a lower rate will result in an identification problem that is serious. The standard assumptions made about the model and the unobserved innovation errors in the model protect the investigators from the identification problem and resulting biases of undersampling. The critical assumption that is needed to identify an undersampled system is that at least one of the exogenous time series is white noise.

1. INTRODUCTION

Linear dynamical systems are widely used in many different fields from engineering to economics. One simple but important class of such systems is called the *single input transfer function model*. Suppose that all the variables of the system are sampled for a period using a fixed sample rate. The central issue of this paper is the determination of the smallest sampling rate that will yield a sample that will allow the investigator to identify the discrete-time representation of the system. The determination of the minimal sufficient sampling rate is a mathematical problem that was solved years ago using Fourier transforms. A critical sampling rate exists that will identify the model. This rate, called the Nyquist rate (p.388, Anderson, 1971) is twice the highest frequency component of the system. The importance of the Nyquist rate for system identification is known in the science and engineering spectral analysis literature yet it has been largely ignored in the literature that applies the time domain methodology popularized by Box and Jenkins (1970).

The standard approach is to start with a discrete-time linear model. An alternative approach, strongly advocated by Wymer (1972,1997) and Bergstrom (1990), is to start with a linear stochastic differential equation model for the system. Telser (1967) discusses the identification problem inherent in estimating the parameters of a difference equation using a data series that is a moving sum of discrete-time observations. Telser also recognized the connection between the parameter identification problem for discrete-time data and the aliasing of the period of a sinusoid. Phillips (1973) addresses the identification of parameters of a continuous time differential equation using discrete-time data. Phillips shows that it is possible to identify a finite parameter dynamical system by assuming linear constraints on the structural matrix even when the stochastic disturbance is aliased. His paper should have led to an important set of advances in time series model identification but it did not catch on, perhaps because as with Telser's paper it was overshadowed by the popularity of the Box and Jenkins point and click methodology.

The sampling rate issue is also confused or ignored in the econometrics literature extending the Box and Jenkins methodology to economics (see Granger and Newbold, 1976, and Harvey, 1981). The standard reason usually given by time series econometricians for ignoring the sampling issue is that it is irrelevant for the identification and estimation problem for a discrete-time linear system model. The parameters of the model are estimated by sample autocorrelations and the sample autocorrelations are unbiased. A more subtle reason for ignoring the sampling rate is that the sampling rate used to collect the data was fixed when the data was collected and it thus can not be changed. This point is well made by Telser in the paper previously cited.

These arguments are true but they deflect attention away from the fact that the models used and the assumptions made about the unobserved innovation errors in the model protect the investigators from the identification problem and resulting biases of undersampling. The critical assumption needed to identify an undersampled system is that at least one of the exogenous time series is white noise.

The standard form for a *causal* linear transfer function model in continuous time is as follows where $x(t)$ denotes the input time series and $y(t)$ denotes the output:

$$(1.1) \qquad y(t) = \int_0^\infty h(s)x(t-s)ds$$

The function $h(t)$ is the *impulse response* of the model. In engineering and science applications, the time series are called *signals* and (1.1) is called a filtering operation where the input signal $x(t)$ is filtering by the impulse response to yield the output signal $y(t)$. Assume that $h(t)=0$ for t > T. The impulse response has *finite support*.

The input and output signals are sampled to produce a set of data. Since the problem is mathematical and not statistical there is no reason to add a noise signal in (1.1) and the signals are functions of time and not continuous time stochastic processes. The sampling issues discussed in this paper apply to any statistical time series model.

2. *BANDLIMITED SAMPLING*

If $x(t)$ and $h(t)$ are absolutely integrable, the Fourier transforms

(2.1) $$X(f) = \int_{-\infty}^{\infty} x(t)\exp(-i2\pi t)dt \quad \text{and} \quad H(f) = \int_{-\infty}^{\infty} h(t)\exp(-i2\pi t)dt$$

exist and $Y(f) = H(f)X(f)$. Since $x(t)$ is real $X(-f) = X^*(f)$, the complex conjugate of $X(f)$ and similarly for the complex transfer function $H(f)$.

Assume that the set of positive support for $|X(f)|$ is $(-f_o, f_o)$ for some frequency f_o. This frequency is the bandlimit of $x(t)$. The transfer function $H(f)$ has infinite support since $h(t)$ is finite.

Suppose that the signal is sampled at the Nyquist rate $2f_o$, or equivalently at a fixed sampling interval $\tau = 1/(2f_o)$. Then the discrete time version of the model (1.1) is

(2.1) $$y(t_k) = \sum_{n=0}^{N} h_o(t_n)x(t_{k-n}) \qquad \text{where } h_o(t) = \int_{-\infty}^{\infty} h(s)\frac{\sin(2\pi f_o(t-s))}{\pi(t-s)}ds,$$

$t_k = k\tau$ and $N = T/\tau$ (Chapter 10, Bracewell, 1986). If τ is much smaller than T then the impulse response parameters $h_o(k\tau) \approx h(k\tau)$ with an error of order $O(\tau^{-1})$.

The discrete-time convolution of the finite $h_o(k\tau)$ sequence with the $x(k\tau)$ sequence to yield the $y(k\tau)$ sequence is a set of linear equations which can be solved to obtain the impulse response parameters $h_o(k\tau)$ for k=*1, ... , N.*

Suppose that the sampling rate used is f_o rather than the Nyquist rate $2f_o$. Then every other $h_0((k-n)\tau)$ and $x(k\tau)$ are missing in the system of equations (2.1). It thus is impossible to solve for the N values of $h_o(k\tau)$. For example, suppose that $y(k\tau) = x(k\tau) + a\,x((k-1)\tau)$. Then the two equations for k=*1* and *2* for times $t_1 = \tau$ and $t_2 = 3\tau$ are

(2.3) $$y(\tau) = x(\tau) + a\,x(0) \quad \text{and} \quad y(3\tau) = x(3\tau) + a\,x(2\tau).$$

These equations can not be solved to find a since $x(0)$ and $x(2\tau)$ are not observed. Using more equations is fruitless since the $x(k\tau)$ for even values of k are not observed. The parameter a is **not estimable** since the system is **not identified**. The investigator must interpolate the missing values in order to estimate a. Interpolation requires some prior knowledge about the functional form of the input.

So far the sampling issue has been separated from the stochastic linear model problem which is the motivation for this exposition. Let us turn to the stochastic model.

3. STOCHASTIC TRANSFER FUNCTION MODEL

Suppose that the $\{x(k\tau)\}$ in expression (2.1) is a sequence of observations of a zero mean random process with a given joint distribution. The covariance function of $\{x(k\tau)\}$ is $c_{xx}(k\tau) = Ex(n\tau)x((n+k)\tau)$. Then the cross covariance function for $\{x(k\tau)\}$ and $\{y(k\tau)\}$ is

(3.1) $$c_{xy}(k\tau) = \sum_{n=0}^{N} h_o(n\tau)c_{xx}((k-n)\tau) .$$

This system of linear equations is used to solve for the parameters $h_o(k\tau)$. If the system in (2.1) were used to obtain an ordinary least squares fit of the parameters, then the solution would be the solution using (3.1) with sample estimates of the covariances $c_{xx}(k\tau)$ and the cross covariances $c_{xy}(k\tau)$ ignoring end effects. Consider the covariances and cross covariances to be known values to simplify exposition.

Once again if the processes are sampled at a slower rate than $2f_o$ then the solution of the linear system (3.1) will produce a distorted estimate of the filter parameters unless the investigator can produce a valid interpolation for the missing covariances. The impulse response is not identified.

An example is helpful here. Suppose that the impulse response is $h_o(n\tau) = 10\cos(2\pi n\tau / N)$ and the input is a first order autoregressive process

AR(ρ) whose innovations variance is one. Thus the covariance of the input is $c_{xx}(k\tau) = \rho^{k\tau}(1-\rho^2)^{-1}$. Assume that the processes are sampled at the rate $f_o/2$ and thus every fourth value of the processes is observed. Figure 1 compares the impulse response recovered from a least squares solution of the under identified system for ρ =0.9 with the skip sampled true impulse response. Figure 2 displays the results for a sampling rate of $f_o/6$. The undersampling produces a distorted picture of the response of the system.

4. *IDENTIFICATION BY WHITE NOISE*

There is a special case for which a subset of the impulse response parameters will be identified. Suppose that $\{x(k\tau)\}$ is white noise, that is $c_{xx}(k\tau)=0$ for all $k \neq 0$ Then $c_{xy}(k\tau) = h_o(k\tau)\sigma_x^2$ from (3.1). In this case, if the processes are sampled at a slower rate than Nyquist, the estimated impulse response parameters will be an under-sampled version of the filter parameters. For example if $h_o(k\tau) = \exp(-ck\tau)$ and the process is sampled at a rate of $f_o/10$ the recovered filter parameters will be $\exp(-ck10\tau)$ for $k \leq N/10$. The recoverd impulse response will provide good short-term prediction for the $y(k10\tau)$.

If one could control the input then it is obvious that one would use white noise input. It is the time series equivalent of an orthogonal design in the statistical design of experiments literature. In the more general dynamical systems the state space representation is of the form

$$\mathbf{y}(k\tau) = \mathbf{A}\mathbf{y}((k-1)\tau) + \mathbf{e}(k\tau) \quad (4.1)$$

where $\mathbf{y}(k\tau)$ is an n dimensional vector of observed exogenous and endogenous time series, $\mathbf{A}$ is a nonsingular system matrix and $\mathbf{e}(k\tau)$ is an n dimensional vector of unobserved exogenous random inputs which are called *innovations* in economics. The innovations sequence models the real input to the linear system and are not a mathematical representation. If on the other hand expression (4.1) is seen as a statistical model to represent the correlations in the data, then one should question the validity of using this statistical model to make statements about causal relationships in the true system.

The system matrix is identified if the innovations are jointly white. This is the generalization of the white input in expression (3.1). If the system is under-sampled then the eigenfunctions of the system are similarly under-sampled but their pattern is not distorted.

The identification assumption on the unobserved innovations time series is the mathematically equivalent to the identification of the impulse response of a linear transfer function with an important distinction. The innovations are not observed while the input of the transfer function is. If nature is obliging and makes the innovations white to help the investigator, then all is well. If not then the covariance structure of the input must be modeled.

Another approach is to reject the Markov model (4.1) and use a properly sampled multivariate transfer function model for forecasting. A multivariate transfer function is an intellectually and technically valid approach to modeling and forecasting a linear system where the input can be measured. Transfer functions are widely applied in engineering and science. But it is not in favor among most time series econometricians.

5 CONCLUSIONS

The results presented in this paper pose a real problem for macroeconomists who use time series models to model economics systems. It is impossible to obtain high frequency data for standard macroeconomic series such as interest rates, output, and prices. The highest frequency macroeconomic data available is monthly. Thus the analyst must use the dynamical model (4.1) and hope that the innovations are white. If the model is a good approximation to reality then the analyst can get some sense of the dynamical response of the system. Otherwise only the trends can be analyzed.

Acknowledgments

I wish to thank Chris Brooks, Bart Bronnenberg, Houston Stokes and Phil Wild for reading the manuscript and giving me constructive criticism

REFERENCES

Anderson, T.W. (1971), *The Statistical 'Analysis of Time Series*. New York: John Wiley & Sons.

Bergstrom, A.R. (1990), *Continuous Time Econometric Modeling*. Oxford: Oxford University Press.

Box, G.E.P. and G.M. Jenkins (1970), *Time Series Analysis: Forecasting and Control*, San Francisco: Holden-Day.

Bracewell, R.N. (1986), *The Fourier Transform and Its Applications*, Second Edition, Revised. New York: McGraw-Hill.

Granger, C.W.J. and P. Newbold (1976), *Forecasting Economic Time Series.* New York: Academic Press.

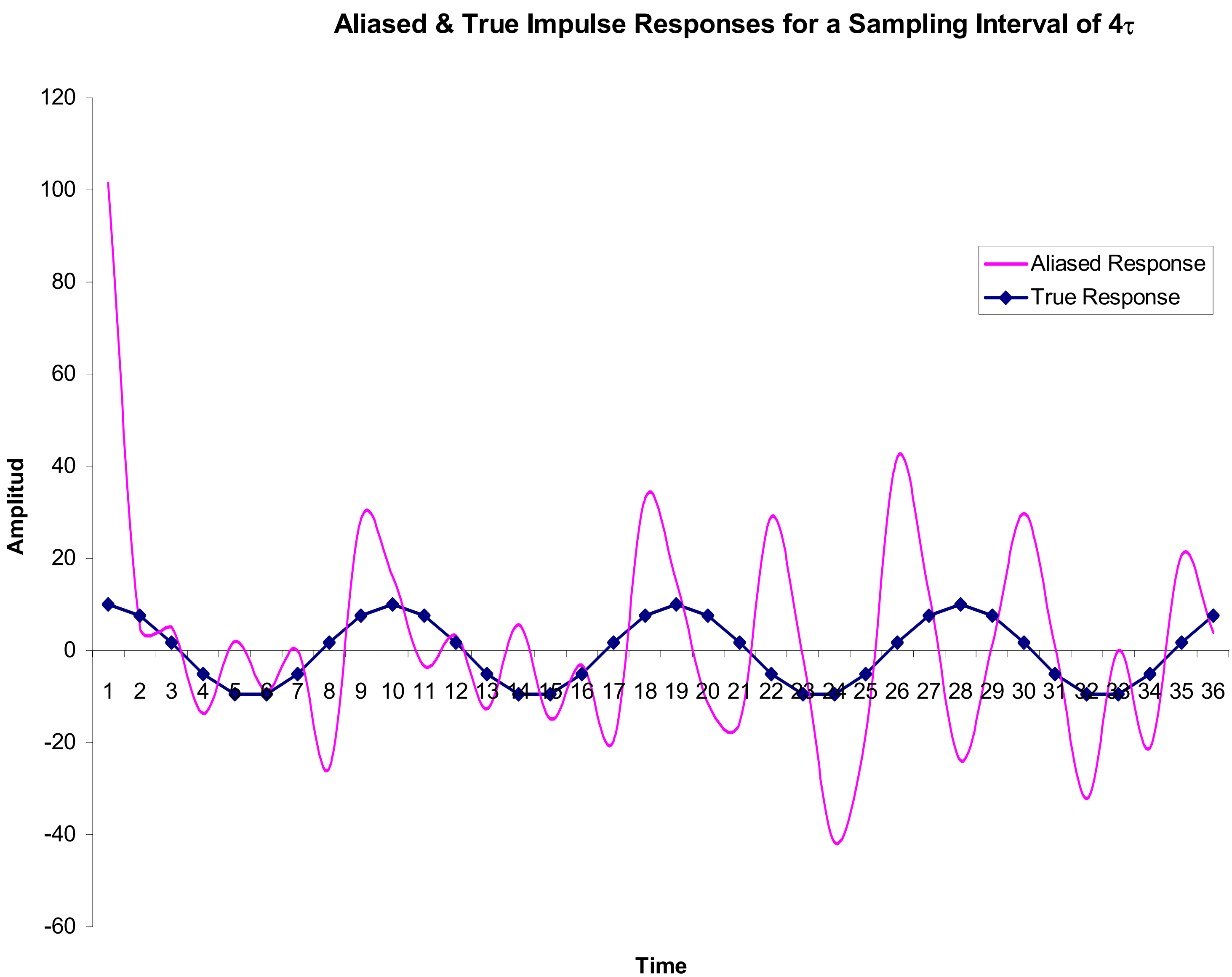
Harvey, A.C (1981), T*ime Series Models*. Oxford: Philip Allan.

Phillips, P.C.B (1973), The problem of identification in finite parameter continuous time models, *Journal of Econometrics 1*, 351-362.

Telser, L.G. (1967), Discrete samples and moving sums in stationary stochastic processes, *Journal of the American Statistical Association 62*, 484-499.

Wymer, C.R. (1972), Econometric estimation of stochastic differential equation systems, *Econometrica 40*, 565-577.

Wymer, C.R. (1997), Structural nonlinear continuous-time models in econometrics, *Macroeconomic Dynamics 1 (2)*, 518-548.



Aliased & True Impulse Responses for a Sampling Interval of 4τ
120
100
80
60
40
20
0
-20
-40
-60
Amplitud
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36
Time
Aliased Response
True Response

Figure 1

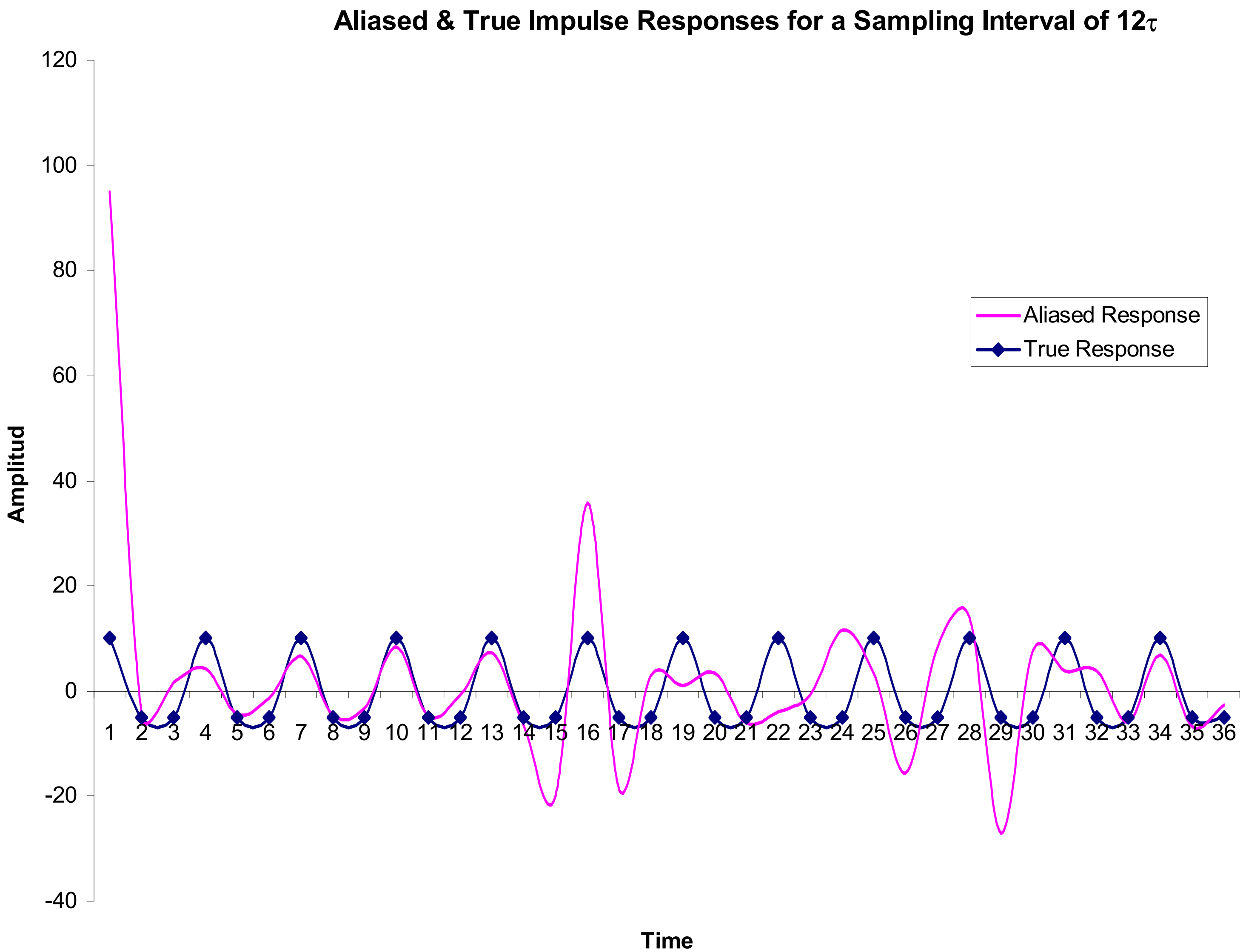

Aliased & True Impulse Responses for a Sampling Interval of 12τ
Amplitud
Time
Aliased Response
True Response
120
100
80
60
40
20
0
-20
-40
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36

Figure 2