RELATIONS BETWEEN TIME DOMAIN AND FREQUENCY DOMAIN PREDICTION ERROR METHODS

Tomas McKelvey
Signal Processing, Dept. of Signals and Systems, Chalmers University of Technology, SE-412 96 Göteborg, Sweden

Keywords: estimation, system identification, time domain, frequency domain, maximum-likelihood estimation, prediction error method, discrete time systems, time series analysis, dynamical systems, spectral estimation, discrete Fourier transform

Contents

1. Introduction
1.1. Problem Formulation
1.2. Frequency Domain System Relations
2. Prediction error methods
2.1. Time Domain
2.2. Frequency Domain
2.2.1. Asymptotic Properties
2.3. A Comparison
2.4. Closed Loop
2.5. Frequency Domain ARX Case
2.5.1. ARX Example
3. Discussion
4. Numerical example
5. Conclusions
Acknowledgements
Glossary
Bibliography
Biographical Sketch

Summary

Prediction error techniques for joint identification of parametric transfer functions and noise models from measured data are considered. The time domain and frequency domain prediction error methods are presented and compared and their close connection with corresponding maximum-likelihood method is explored. Conditions are established under which the two techniques are asymptotically equivalent.

1. Introduction

Building mathematical models based on measured input and output signals of a dynamical system is known as system identification. Such models based on empirical information are important if the dynamical system is unknown or is only partially known and when it is in-feasible to derive a theoretical model from first principles. The availability of accurate models is important in order to derive high performance solutions, e.g., for model based control design or model based signal processing

Almost all measurements originating from real world devices intrinsically belong to the time domain, i.e. are samples of continuous time signals. Consequently a large part of the system identification methods and the theory developed around them deals with how to determine parametric models from such time domain measurements. Using samples of the Fourier transform of signals, here called *frequency domain data*, to build models is an alternative technique which has been frequently used for non-parametric estimates of the system transfer function.

A discussion of techniques to fit *parametric models* to noisy frequency domain data is the scope of this chapter. Such techniques can be traced back to the mid 1950s when Whittle combined classical inferential procedures, e.g. maximum-likelihood (ML) estimation, with spectral theory for time-series analysis.

A distinctive feature of frequency domain techniques is that modeling of continuous time systems from sampled data can be done in a straightforward fashion if a certain class of band-limited excitation signals is employed. This is a great advantage in contrast with the rather involved time domain techniques which, even in the noise free case, are only approximate if a finite set of sampled data is available. We will here however completely focus on the discrete time case and refer to *Estimation with known Noise Model* and *Continuous-time Identification* for more information on the continuous case.

The aim this chapter is to briefly introduce the time domain and frequency domain prediction error methods and highlight their close relation and present their own unique properties. More information regarding system identification in general can be found in *Frequency Domain System Identification* and *Prediction error methods*.

1.1. Problem Formulation

Let us assume that we are interested in obtaining a model of a system that can be described by the following linear time-invariant form

$$y(t) = G_0(q)u(t) + H_0(q)e(t) \qquad (1)$$

where $y(t), u(t)$ and $e(t)$ are the real valued output, input and noise signals, respectively. The operators $G_0(q)$ and $H_0(q)$ represent the discrete time linear transfer functions. We assume that H_0 is a stable an inversely stable monic filter. The noise signal $e(t)$ is assumed to be independent and identically distributed (i.i.d.) and zero mean with variance λ_0 and independent of the input signal $u(t)$. Later in Section 2.4 we will discuss the closed loop case where $u(t)$ and $e(t)$ are dependent through a feedback controller. The presentation is limited to the single input single output case where $u(t)$ and $y(t)$ are scalar valued quantities.

1.2. Frequency Domain System Relations

Let us define the discrete Fourier transform (DFT) of the signal $\{x(t)\}_{k=1}^{N}$ as

$$X_n(\omega) = \frac{1}{\sqrt{N}} \sum_{t=1}^{N} x(t) e^{-j\omega t} \; .$$

From the system relation (1) it is well known that

$$Y_N(\omega) = G_0(e^{j\omega}) U_N(\omega) + V_N(\omega) + O(\frac{1}{\sqrt{N}}) \qquad (2)$$

where V_N is asymptotically complex normally distributed with zero mean and variance $\Phi_v(\omega) \triangleq \lambda_0 \,|\, H_0(e^{j\omega}) \,|^2$. Furthermore, as $N \to \infty, V_N(\omega)$ and $V_N(\xi)$ are asymptotically independent whenever $\omega \neq \xi$ and $(\omega + \xi) \mod 2\pi \neq 0$.

Asymptotically as $N \to \infty$ we can consider the following frequency domain system equation to hold

$$Y(\omega) = G_0(e^{j\omega}) U(\omega) + H(e^{j\omega}) E_0(\omega) \; . \qquad (3)$$

where Y and U are the weak limits of Y_N and U_N and E_0 are the frequency domain innovations which are zero mean complex normally distributed with variance λ_0.

In our setup we assume that we can sample the relation (3) at a sequence of frequencies in the set $\Omega_N = \{\omega_k\}_{k=1}^{N}$ yielding the set $Z^N = \{Y_k, U_k \mid k = 1,..., N\}$ where $U_k = U(\omega_k)$ and $Y_k = Y(\omega_k)$. Hence, Y_k is a complex normally distributed random variable with mean $G_{0,\omega_k} U_k$ and variance $\Phi_v(\omega_k)$. Notice that the frequencies ω_k in the set Ω_N can have an arbitrary distribution. In practice these samples are obtained by the DFT and most often on the equidistant frequency grid $\omega_k = 2\pi k / N$, for $k = 0,..., N-1$.

Since the transformation of a signal from time to frequency domain using the DFT is nothing but a unitary transformation it might appear, at first sight, that nothing is gained by considering the estimation problem in the frequency domain. However, an important difference arises when the noise is colored. i.e. when $H(q) \neq 1$. In this case the samples of the system output $y(t)$ will be statistically dependent.

The unitary transformation, represented by the DFT, asymptotically decuples this statistical dependence. That is, in the frequency domain each sample is asymptotically (as the number of time domain data points tends to infinity) independent of the others. This provides an increased flexibility in the selection of which data to use for the estimation and the ability to combine data from different time domain experiments. If the noise transfer function $H_0(q)$ is known, the input and output time domain data can be pre-filtered with the filter $H_0(q)^{-1}$. Such an operation, which is also known as pre-whitening, will make the pre-filtered time domain samples white, i.e. statistically independent.

2. Prediction Error Methods

The aim is to find a model of (1) and to do so we construct a parameterized model

$$y(t) = G_\theta(q)u(t) + H_\theta(q)e(t) \tag{4}$$

where the transfer functions $G_\theta(q)$ and $H_\theta(q)$ are models of the system and noise transfer functions. The transfer functions are parameterized by a real valued vector θ. Let $D_{\mathcal{M}}$ denote the set of valid parameters. We assume $H_\theta(q)$ is a stable and inversely stable monic transfer function for all $\theta \in D_{\mathcal{M}}$ .we impose no particular structure on how the parameters enter into $G_\theta(q)$ or $H_\theta(q)$ and this enables the use of various parameterizations such as fraction of polynomials or state-space models. Hence, parameterized *gray-box models* which are partially unknown can also be used. To simplify notation in the sequel let $G_{0,\omega} \triangleq G_0(e^{j\omega})$ and $G_{\theta,\omega} \triangleq G_\theta(e^{j\omega})$ and similarly for the noise transfer function H.

2.1. Time Domain

Suppose input-output data in the time domain are given

$$z^N = \{y(t), u(t) \mid t = 1, \ldots, N\}$$

From the model (4) we can define the one-step ahead predictor

$$\hat{y}(t \mid \theta) = H_\theta(q)^{-1} G_\theta(q)u(t) + (I - H_\theta(q)^{-1})y(t).$$

The prediction error are defined as

$$\mathcal{E}(t,\theta) = y(t) - \hat{y}(t \mid \theta) = H_\theta(q)^{-1}(y(t) - G_\theta(q)u(t))$$

and would equal the white noise $e(t)$ if the output $y(t)$ was indeed generated by the model and for some sequences $u(t)$ and $e(t)$. The prediction error method (PEM) finds the model parameter θ by minimizing the sample variance of the prediction errors:

$$\begin{aligned}
V_N^{\mathrm{TD}}(\theta) &= \sum_{k=1}^{N} |\,\mathcal{E}(t,\theta)\,|^2 = \frac{1}{N}\sum_{k=1}^{N} |\,H_\theta(q)^{-1}(y(t) - G_\theta(q)u(t))\,|^2 \\
\hat{\theta}_N &= \arg\min_\theta V_N^{\mathrm{TD}}(\theta) \\
\hat{\lambda}_N &= V_N^{\mathrm{TD}}(\hat{\theta}_N)
\end{aligned} \tag{5}$$

where $\hat{\lambda}_N$ is the estimate of the variance of $e(t)$. Notice that the prediction errors are the deterministic system errors $y(t) - G_\theta(q)u(t)$ filtered through the inverse of the noise model. If $e(t)$ is normally distributed zero mean random variables with covariance λ, the prediction error method estimator is the maximum-likelihood estimator of θ if all initial conditions of the system and noise filters are zero. Hence it is often called the conditional maximum –likelihood estimator. Finally by applying Parseval's formula to the criterion function in (5) reveals that the PEM estimator minimizes the function

$$\int_{-\pi}^{\pi} \frac{|Y_N(\omega) - G_{\theta,\omega} U_N(\omega)|^2}{|H_{\theta,\omega}|^2} d\omega . \tag{6}$$

Please refer to *Prediction Error Method* for a more thorough treatment of the class of time domain prediction error methods.

-
-
-

Bibliography

Brillinger D.R. (1981). Time Series: Data Analysis and Theory. McGrawHill Inc., New York. [Provides an extensive collection of results of the statistical properties of the discrete Fourier transform of stochastic signals].

Chung K.L. (1974). *A Course in Probability Theory*. Academic Press, San Diego, CA. [This book contains many results on convergence of sums of random variables].

Levy E.C. (1959). Complex curve fitting. *IRE Trans. on Automatic Control* **AC4**, 37–44. [Describes a leastsquares technique to fit a continuoustransfer function to frequency domain data].

Ljung L. (1994). Building models from frequency domain data. In *IMA Workshop on Adaptive Control and Signal Processing*, Minneapolis, Minnesota. [Presents the frequency domain maximumlikelihood method for joint estimation of system and noise transfer functions].

Ljung L. (1999). *System Identification: Theory for the User*. Englewood Cliffs, New Jersey: PrenticeHall, 2nd edn. [This is an advanced textbook on the graduate level which covers most aspects of system identification both for the time domain and frequency domain problems].

Ljung L., Glover K. (1981). Frequency domain versus time domain methods in system identification. *Automatica* **17**(1), 71–86. [Gives a discussion and comparison of time domain parametric modeling methods and frequency domain nonparametric ones].

McKelvey T., Ljung L. (1997). Frequency domain maximum likelihood identification. In *Proc. of the 11th IFAC Symposium on System Identification*, vol. 4, pp. 1741–1746, Fukuoka, Japan. [Presents some asymptotic results for the multivariable frequency domain maximumlikelihood estimator].

Pintelon R., Guillaume P., Rolain Y., Schoukens J., Van Hamme H. (1994). Parametric identification of transfer functions in the frequency domain – A survey. *IEEE Trans. of Automatic Control* **94**(11), 2245–2260. [This overview paper provides a survey of many frequency domain parametric estimation techniques].

Pintelon R., Schoukens J. (2001). *System Identification A frequency domain approach*. IEEE Press. [This book provides extensive material on identification in general but with a focus on the practical use of frequency domain methods. The book also summarizes much of the important theoretical foundations.].

Whittle P. (1951). *Hypothesis testing in time series analysis*. Thesis. Uppsala University, Almqvist and Wiksell, Uppsala. Hafner, New York. [An early (first) account of using frequency domain techniques for parametric modeling of linear systems].

Biographical Sketch

Tomas McKelvey was born in 1966 in Lund, Sweden. He received his M.Sc. degree in Electrical Engineering from Lund Institute of Technology, Sweden in 1991 and his Ph.D. degree in Automatic Control from Linköping University, Sweden in 1995. From 1995 to 2000 he has held Assistant and Associate professor positions at Linköping University. Currently he is an Associate professor (docent) of Signal Processing in the department of Signals and Systems at Chalmers Technical University, Göteborg, Sweden. Between 1999 and 2000 he was a visiting researcher at University of Newcastle, Australia. Prof. McKelvey is currently the Chairman of International Federation of Automatic Control's (IFAC) Technical Committee on Modelling, Identification and Signal Processing and is also an associate editor for *Automatica*. His main scientific interests are system identification, spectral estimation, time series analysis and signal processing.