第5卷 第4期 控制理论与应用 Vol.5, No.4
1988年12月 CONTROL THEORY AND APPLICATIONS Dec., 1988

Some Complementary Relations Between Time Domain Method and Frequency Domain Method for System Identification

Yuan Zhendong
(Department of Mathematics, East China Normal University Shanghai)

Abstract

This paper gives some complementary pictures of the time domain approach and the frequency domain approach in system identification. Two main time domain methods are compared by a frequency domain criterion. Some interesting properties of the time domain methods for estimating transfer functions are given. An analytic solution of optimal input design for system identification is discussed.

1. Introduction

By frequency domain methodsfor system identification, we mean sine wave testing (frequency analysis), conventional spectral analysis, the maximum entropy method or simply using frequency domain expression to calculate things related to the identification procedure. These methods are used to estimate frequency responses, pulse responses, or transfer functions. By time domain methods for system identification, we mean least square method, maximum likelihood method, prediction error method or output error methods. Using these methods, we can estimate the coefficients of the differential equation, the difference equation or the state equation.

Historically, frequency domain methods seem to dominate the theory and practice of system identification in control engineering applications up to sixties. Since the end of the sixties, the interest in time domain methods has increased. The control literature on identification is now apparently dominated by time domain methods.

However, frequency domain methods are still very useful in

Manuscript received Apr. 6, 1987.

practical applications. Moreover, the transfer function plays a very important role for understanding the properties of a linear system. Even if we work with a parametric, time domain model, the accuracy of the transfer function of that model will tell us how successful certain applications involving the model will be. The interplay between parametric model estimate and their frequency domain properties is crucial in adaptive control area. Some complementary pictures of these two approaches have been given in Ljung and Glover (1981).

In this paper, we shall make some further discussion on the complementary relationships of these two approaches. We compare two main time domain methods, the prediction error method and the output error method, by frequency domain criteria in Section 2. We review some recent results on time domain methods for estimating transfer functions by Ljung and Yuan (1985), in Section 3. Section 4 gives an analytic solution of the optimal input design problems. Finally we give some conclusions in Section 5.

2. Comparision Between Output Error Method and Prediction Error Method

Consider the following discrete time single-input-single-output system

$$y(t)=G(\theta,q^{-1})u(t)+H(\theta,q^{-1})e(t) \qquad (2.1)$$

where $u(t)$, $y(t)$ are the input and the output at time t, respectively. $\{e(t)\}$ is a sequence of white noise. θ is unknown parameter vector. $G(\theta,z^{-1})$ is the transfer function of $u(t)$ and $y(t)$ and $H(\theta,z^{-1})$ the transfer function of the stochastic disturbance, $H(o)=1$.

According to the model (2. 1), the prediction of $y(t)$ is given by (2.2)

$$\hat{y}(t|\theta)=[1-H^{-1}(\theta,q^{-1})]y(t)+H^{-1}(\theta,q^{-1})G(\theta,q^{-1})u(t) \qquad (2.2)$$

The prediction error

$$\varepsilon(t,\theta)=y(t)-\hat{y}(t|\theta) \qquad (2.3)$$

For a system of type (2. 1), we consider a special case that

$$\theta=\begin{pmatrix}\alpha\\ \delta\end{pmatrix},\ \alpha\in A\subset R^p,\ \delta\in D\subset R^r$$

and $p+r=s$.

The true system is

S: $y(t)=G(\alpha_0,q^{-1})u(t)+H(\delta_0,q^{-1})e(t)$ (2.4)

The model set, in prediction Error Method, is discrbed by

M: $y(t)=G(\alpha,q^{-1})u(t)+H(\delta,q^{-1})e(t)$ (2.5)

The prediction error (2.3) turns out to be

$$\varepsilon(t,\theta)=H^{-1}(\delta,q^{-1})(y(t)-G(\alpha,q^{-1})u(t))$$
$$=H^{-1}(\delta,q^{-1})(G(\alpha_0,q^{-1})u(t)+H(\delta_0,q^{-1})e(t)-G(\alpha,q^{-1})u(t))$$

Hence

$$\varepsilon(t,\theta)=H^{-1}(\delta,q^{-1})(G_0(q^{-1})-G(\alpha,q^{-1}))u(t)+\frac{H_0(q^{-1})}{H(\delta,q^{-1})}e(t) \quad (2.6)$$

where

$$\theta=\begin{pmatrix}\alpha\\ \delta\end{pmatrix},\ G_0(q^{-1})=G(\alpha_0,q^{-1}),\ H_0(q^{-1})=H(\delta_0,q^{-1})$$

Introduce a scalar valued criterion of fit

$$\varepsilon^2(t,\theta)$$

After N data have been collected, we can form the criterion function

$$V_N(\theta)=\frac{1}{N}\sum_{t=1}^{N}\varepsilon^2(t,\theta) \quad (2.7)$$

The prediction error estimate of θ is obtained by minimizing $V_N(\theta)$, over $\theta\in D_m$. The minimizing element, denoted by $\hat{\theta}_N$, will thus give us the resulting estimate of the system: $M(\theta)$.

It can be shown under quite general condition that the prediction error estimate $\hat{\theta}_N$ tends with probability one to the value $\hat{\theta}$ that minimzes

$$V(\theta)\triangleq\lim_{N\to\infty}EV_N(\theta)=E\varepsilon^2(t,\theta) \quad (2.8)$$

(See, Ljung (1978)) where $V(\theta)$ is called the limit criterion function, and

$$\hat{\theta}=\arg\min_{\theta\in D_M}V(\theta) \quad (2.9)$$

Similarly, in the output error method, the model set can be represented by

M_1: $y(t)=G(\alpha,q^{-1})u(t)+v(t)$ (2.10)

where $v(t)$ represents some stochastic disturbance of unspecified character. The output error is that

$$\varepsilon_1(t,\alpha)=y(t)-y_m(t) \quad (2.11)$$

where

$$y_m(t)=G(\alpha,q^{-1})u(t) \quad (2.12)$$

The output error estimate of α is obtained by minimizing the following criterion function

$$V'_N(\alpha) = \frac{1}{N}\sum_{t=1}^{N} \varepsilon_1^2(t,\alpha) \qquad (2.13)$$

The output error estimate, denoted by $\tilde{\alpha}_N$, tends with probability one to $\tilde{\alpha}$, where

$$\tilde{\alpha} = \arg\min_{\alpha} V'(\alpha) \qquad (2.14)$$

and

$$V'(\alpha) = E\varepsilon_1(t,\alpha) \qquad (2.15)$$

For comparision, we consider a frequency domain criterion

$$\int_{-\pi}^{\pi} |G_0(e^{i\omega}) - G(\theta,e^{i\omega})|^2 Q(\omega)d\omega \qquad (2.16)$$

where $G_0(e^{i\omega})$ is the true system transfer function, $G(\theta,e^{i\omega})$ represents the model set, $Q(\omega)$ is a weighting function $Q(\omega)>0$, $\forall\omega$. The best approximation of the true system in the frequency domain, is obtained by minimizing the criterion function (2.16). How do we select the weighting function $Q(\omega)$? IF the true system belongs to the model set M; $S\in M$, then the optimal $G(\theta,e^{i\omega}) = G_0(e^{i\omega})$ regardless of $Q(\omega)$ as $Q(\omega)>0, \forall\omega$ and hence the integral (2.16) is greater than or equal to zero. $G(\theta,e^{i\omega}) = G_0(e^{i\omega})$ the integral equals zero. If $S\notin M$, the weighting function will determine the bias distribution, where in the frequency band we require to have the best fit. A high value of $Q(\omega)$ at certain frequencies indecates that we require that the bias in $G(\theta,e^{i\omega})$ to be small at these frequencies.

For comparision between the prediction error method and the output error method, we give the following definition.

Definition 2.1 If $F_0(z^{-1})$ is a known transfer function of a low-pass filter, $Q(\omega)>0$, and $S\notin M$.

$$\hat{\eta} = \arg\min_{\eta} \int_{-\pi}^{\pi} |F_0^{-1}(e^{i\omega})||G(\eta,e^{i\omega}) - G_0(e^{i\omega})|^2 Q(\omega)d\omega$$

and

$$\tilde{\eta} = \arg\min_{\eta} \int_{-\pi}^{\pi} |G(\eta,e^{i\omega}) - G_0(e^{i\omega})|^2 Q(\omega)d\omega$$

then the estimate $G(\hat{\eta},e^{i\omega})$ is said to be more emphasized at the high frequency band than $G(\tilde{\eta},e^{i\omega})$ or simply say, $\hat{\eta}$ is more emphasized at high frequency band than $\tilde{\eta}$.

Theorem 2.1 If $S\notin M\cup M_1$, $H(\hat{\delta},e^{i\omega})$ is a low-pass filter transfer function, where $\hat{\theta}^T=(\hat{\alpha}^T,\hat{\delta}^T)$ in (2.9), then the limit prediction error estimate $\hat{\alpha}$ is more emphasized at the high frequency band than the output error estimate $\tilde{\alpha}$.

proof. In view of (2.6), (2.8) and (2.9)

$$\begin{pmatrix}\hat{\alpha}\\ \hat{\delta}\end{pmatrix}=\arg\min_{\alpha,\delta}\left\{\int_{-\pi}^{\pi}|H^{-1}(\delta,e^{i\omega})|^2|G(\alpha,e^{i\omega})-G_0(e^{i\omega})|^2\phi_u(\omega)d\omega+\lambda\int_{-\pi}^{\pi}\left|\frac{H_0(e^{i\omega})}{H(\delta,e^{i\omega})}\right|d\omega\right\}$$

Hence

$$\hat{\alpha}=\arg\min_{\alpha}\int_{-\pi}^{\pi}|H^{-1}(\hat{\delta},e^{i\omega})|^2|G(\alpha,e^{i\omega})-G_0(e^{i\omega})|^2\phi_u(\omega)d\omega \qquad (2.17)$$

From (2.11),

$$\begin{aligned}\varepsilon_1(t,\theta)&=y(t)-G(\alpha,q^{-1})u(t)\\&=(G_0(q^{-1})-G(\alpha,q^{-1}))u(t)+H_0(q^{-1})e(t)\end{aligned} \qquad (2.18)$$

According to (2.14), (2.15) and (2.18)

$$\tilde{\alpha}=\arg\min_{\alpha}\left\{\int_{-\pi}^{\pi}|G(\alpha,e^{i\omega})-G_0(e^{i\omega})|^2\phi_u(\omega)d\omega\right. \qquad (2.19)$$

Compare (2.19) with (2.17) and use the definition 2.1, thus completes the proof of the theorem.

Remark This theorem means that if the true system does not belong to the model set, we use the prediction error estimate or the output error estimate in the recursive version, then for large enough N (the number of data), by using the prediction error method, the resulting transfer function has smaller bias at the high frequency band than the other.

Wahlberg (1983) discussed the similar problem, but he considered the limit case only.

3. Properties of Time Domain Identification for Estimating Transfer Function

To estimate a transfer function is basically a non-parametric problem. One common approach is the frequency domain one: the cross-spectrum of the input and the output is estimated, as well as the input spectrum and a transfer function estimate is formed as their ratio.

In this section, we consider a discrete time stochastic linear system, generally, it can be expressed as

$$y(t) = G_0(q^{-1})u(t) + v(t) \tag{3.1}$$

where $u(t)$ and $y(t)$ are respectively its input and output, $v(t)$ is the stochastic disturbance. $v(t)$ is a zero mean stationary process. Expand $G_0(q^{-1})$

$$G_0(q^{-1}) = \sum_{k=1}^{\infty} g_k^0 q^{-k} \tag{3.2}$$

The complex-valued function

$$G(e^{i\omega}) = \sum_{k=1}^{\infty} g_k^0 e^{-ik\omega}; \quad -\pi \leqslant \omega \leqslant \pi \tag{3.3}$$

is the transfer function for model (3.1) and (3.2).

In conventional spectral analysis, the transfer function estimate is given by the following expressions

$$\overline{G}(e^{i\omega}) = \overline{\phi}_{yu}(i\omega)\overline{\phi}_u^{-1}(\omega), \tag{3.4}$$

here $\overline{\phi}_{yu}(i\omega)$ and $\overline{\phi}_u(\omega)$ are their smoothed spectral estimates,

$$\overline{\phi}_{yu}(i\omega) = \frac{1}{2\pi} \sum_{k=-M}^{M} w(k)\hat{r}_{yu}(k)e^{-ik\omega}$$

$$\overline{\phi}_u(\omega) = \frac{1}{2\pi} \sum_{k=-M}^{M} w(k)\hat{r}_u(k)e^{-ik\omega}$$

where

$$\hat{r}_{yu}(k) = \frac{1}{N} \sum_{t=1}^{N-k} y(t+k)u(t)$$

$$\hat{r}_u(k) = \frac{1}{N} \sum_{t=1}^{N-k} u(t+k)u(t)$$

A comprehensive discussion can be found in Jenkins and Watts (1968). It is well known that the variance of the spectral estimate does not decrease as $N \to \infty$, one may use the window function $w(k)$ to decrease it. The useful window functions include Bartlett window, Parzen window etc. The accuracy of this method depends on the lag of window and the character of the input signals.

In Ljung and Yuan (1985), the following method is given for estimating the transfer functions. Consider the model set

$$M:\ y(t)=G_d(q^{-1})u(t)+v(t) \tag{3.5}$$

where

$$G_d(q^{-1})=\sum_{k=1}^{d} g_k q^{-k} \tag{3.6}$$

Using the output error method to estimate the parameter θ in the model (3.5) and (3.6), we obtain the time domain estimate of the transfer function

$$G_d(\hat{\theta}_N,e^{i\omega})=\sum_{k=1}^{d}\hat{g}_k e^{-ik\omega};\quad -\pi\leqslant\omega\leqslant\pi \tag{3.7}$$

where

$$\theta^T=(g_1,\cdots,g_d) \tag{3.8}$$

Denote that

$$\hat{G}_{d(N)}(e^{i\omega})=G_d(\hat{\theta}_N,e^{i\omega}) \tag{3.9}$$

here $d(N)$ denotes the model order used when N data have been collected. The transfer function is parametrized as a black box and no given order is chosen a priori. This means that the model orders may increase to infinity when the number of observed data tends to infinity, i. e. $d(N)\to\infty$, as $N\to\infty$. The parameters are only vehicles for arriving at a transfer function estimate. Under certain regularity conditions, essentially independent of model structure and noise model, the basic results are given as the following:

(1) $\hat{G}_{d(N)}(e^{i\omega})\to G_0(e^{i\omega})$ as $N\to\infty$, with probability one. (3.10)

(2) $\dfrac{N}{d(N)}\operatorname{cov}\hat{G}_{d(N)}(e^{i\omega})\to\phi_v(\omega)/\phi_u(\omega)$ as $N\to\infty$ (3.11)

(3) $\sqrt{\dfrac{N}{d(N)}}\,(\hat{G}_{d(N)}(e^{i\omega})-E\hat{G}_{d(N)}(e^{i\omega}))\in$ As $N(0,\phi_v(\omega)/\phi_u(\omega))$, (2.12)

where $\phi_v(\omega)$ and $\phi_u(\omega)$ are the disturbance and input spectral densites, respectively. These results are useful for optimal input design and other design problems.

Remark The problem of estimating the transfer function of a multivariable, linear, stochastic system has been considered in Yuan and Ljung (1984). Similar results have been obtained where $G_0(q^{-1})$ is the pxm polynomial matrix and $(N/d(N))\operatorname{cov}\hat{G}_{d(N)}(e^{i\omega})\longrightarrow\Phi_u^{-1}(\omega)$

$$E\int_{-\pi}^{\pi} | G(\hat{\theta}_N \cdot e^{i\omega}) - G_0(e^{i\omega}) |^2 Q(\omega) d\omega$$

$$\approx E\int_{-\pi}^{\pi} \left(\frac{\partial G}{\partial \theta}\bigg|_{\theta=\theta_0} \right) (\hat{\theta}_N - \theta_0)(\hat{\theta}_N - \theta_0)^T \left(\overline{\frac{\partial G}{\partial \theta}}\bigg|_{\theta=\theta_0} \right)^T Q(\omega) d\omega$$

$$= \text{trace}\left\{ E(\hat{\theta}_N - \theta_0)(\hat{\theta}_N - \theta_0)^T \int_{-\pi}^{\pi} \left(\overline{\frac{\partial G}{\partial \theta}}\bigg|_{\theta=\theta_0} \right)^T \cdot \left(\frac{\partial G}{\partial \theta}\bigg|_{\theta=\theta_0} \right) Q(\omega) d\omega \right\}$$

where the bar "–" denotes conjugate and "T" the transponse.
Therefore

$$V(\phi_u(\omega)) \approx \text{trace } P_N \cdot W \tag{4.3}$$

Here P_N is the covariance matrix of $\hat{\theta}_N$, and W is a weighting matrix derived from $Q(\omega)$, G_0 and the model parametrization, i. e.

$$W = \int_{-\pi}^{\pi} \left(\overline{\frac{\partial G}{\partial \theta}}\bigg|_{\theta=\theta_0} \right)^T \left(\frac{\partial G}{\partial \theta}\bigg|_{\theta=\theta_0} \right) Q(\omega) d\omega \tag{4.4}$$

Minimization of (4.3) under the constraint (4.2c) has been studied, e. g. in the aforementioned references. Some structural results are known, such as how many different sinusoids are required to reach the minimum, etc. The actual calculation of the optimal spectrum is however numerically complicated.

In this section, we consider the following criterion

$$J = \lim_{N\to\infty} \frac{N}{d(N)} \int_{-\pi}^{\pi} E| \hat{G}_N(e^{i\omega}) - G_0(e^{i\omega}) |^2 Q(\omega) d\omega \tag{4.5}$$

Under certain regularity conditions, from (3.10) and (3.11), we obtain

$$J = \int_{-\pi}^{\pi} \frac{\phi_v(\omega)}{\phi_u(\omega)} Q(\omega) d\omega$$

If $\phi_e(\omega) = 1$, then

$$J = \int_{-\pi}^{\pi} \frac{|H_0(e^{i\omega})|^2 Q(\omega)}{\phi_u(\omega)} d\omega \tag{4.6}$$

The optimal input design problem will be to obtain

$$\min_{\phi_u(\omega)} J(\phi_u(\omega)) \tag{4.7a}$$

with the constraint

$$\int_{-\pi}^{\pi} \phi_u(\omega)d\omega = 1 \qquad (4.7b)$$

This is a typical constrained variation problem. In view of Gelfand and Fomin (1963), the optimal input spectral density must satisfy the following equation

$$\frac{-|H_0(e^{i\omega})|^2 Q(\omega)}{(\phi_u^*(\omega))} + \lambda = 0 \qquad (4.8)$$

where λ is a constant.
From (4.8), we obtain

$$\phi_u^*(\omega) = \text{const} \cdot |(H_0(e^{i\omega})|\sqrt{Q(\omega)} \qquad (4.9)$$

where the "const" is determined by (4.7b), i.e.

$$\int_{-\pi}^{\pi} \phi_u^*(\omega)d\omega = 1. \qquad (4.10)$$

An appealing aspect of such a result is that the optimal input is easy to determine It depends on the chosen weighting function $Q(\omega)$ in a very natural way. For an output error method we have $H(\theta, e^{i\omega}) \equiv 1$. Hence the optimal input is independent of the system in that case. A systematic discussion on this topic can be found in Yuan and Ljung (1985).

5. Conclusion

Some prospective pictures on the complementary relations have been shown in this paper. We believe that the gap between the time domain approach and the frequency domain approach will be filled in the future. A good system identification package should include examples of both techniques. This package will facilitate both theoretitical researchs and practical applications.

References

〔1〕 Goodwin, G. C. and R. L. Payne, Dynamic System Identification: Experiment Design and Data Analysis, Acadmic Press, New York. (1977).

〔2〕 Gelfand, I. M. and S. V. Fomin, Calculus of Variations. Prentice-Hall Inc., (1963), 43.

〔3〕 Ljung, L. Convergence Analysis of Parametric Identification Methods. IEEE Trans., AC-23, (1978), 770-783.

〔4〕 Ljung, L. and P. E. Caines, Asymptotic Normality of Prediction Error Estimation for Approximate System Models. Stochastics, 3, (1979), 29-46.

〔5〕 Ljung, L. and K. Glover, Frequency Domain Versus Time Domain Methods in System Identification. Automatica, 17, (1981), 71-86.

〔6〕 Ljung, L. and T.Soderstrom Theory and Practice of Recursive Identification, MIT Press, (1983).

〔7〕 Ljung L. and Z. D. Yuan, Asymptotic Properties of Black Box Identification of Transfer Functions, IEEE Trans., AC 30:6, (1985), 514-530.

〔8〕 Mehra, R. K., Choice of Input Signals. In P. Eykhoff (ed) Trends and Progress in System Identification, Pergamon Press, Oxford, (1981).

〔9〕 Wahlberg, B., Connections Between Frequency Domain Validity Criteria and Time Domain Identification Methods. Technical Report LiTH-ISY-I-0608 Linkoping University, Sweden, (1983).

〔10〕 Yuan, Z. D. and L. Ljung, Optimal Input Design by Frequency domain Criteria The 21th IEEE CDC Conference, Orlando, Florida, USA, (1982), 1005-1006.

〔11〕 Yuan, Z. D. and L. Ljung, Black-box Identification of Multivariable Transfer Functions-Asymptotic Properties and Optimal Input Design. Int. J. of Control, 40:2, (1984), 233-256.

〔12〕 Yuan, Z. D. and L. Ljung, Unprejudiced Optimal Input Design for Identification of Transfer Functions. Automatica, 21:6, (1985), 697-708.

系统辨识中时域方法与频域方法的相辅相成关系

袁 震 东

（华东师范大学数学系，上海）

摘 要

本文给出时域辨识与频域辨识相互补充的一些事实。用频域准则对时域辨识作比较，可弄清不同时域辨识的频域意义。用时域辨识可获得传递函数估计的有意义的性质。本文还包括这些性质在最优输入设计上的应用。