Pre-Filtering In Robust Model Estimation-A Brief Tour

Dalvinder kaur mangaldalvinder@yahoo.com
Assistant Professor/ Electronics & Instrumentation Department
MD University Rohtak
TITS, Bhiwani, 127021
Haryana, India

Lillie Dewan l_dewanin@yahoo.com
Professor/ Electrical Engineering department
NIT, Kurukshetra, 132119
Haryana, India

Abstract

Presence of noise has significant effect on the system identification and parameter estimation. To have accurate system models cleaner data is required which can be obtained if noise is reduced by prefiltering In this paper an attempt has been made to survey the literature on the prefiltering methods in system identification.

Keywords: Pre-filtering 1, Nonarithmatic filters 2, Gaussian noise 3, Moving average 4, Robust H∞ filtering 5

1. INTRODUCTION

The adaptive ARMA system identification is usually realized by the adaptive equation error and output error algorithm [1], thus, the system identification will be influenced by additive noise caused by various reasons such as disturbance and measurement noise. Measured noise means the measured data invariably contains noise attributed to a no. of causes; sensors and measurement devices, modeled and unmodelled disturbances or unaccounted and unidentified sources. Measurement noise particularly of non – Gaussian type is known to be problematic in identification and parameter estimation and relatively small amount may wreak havoc on linear estimation schemes. Opportunities for improved system identification exist using data prefiltering and noise reducing techniques. In this paper an attempt has been made to study various prefiltering techniques in robust model estimation given in literature. Brief introduction to mathematical model used for identification is given in Section2. Data filtering and three methods form the literature has been discussed in section3 followed by conclusions in section4.

2. MATHEMATICAL FOUNDATION AND PARAMETER ESTIMATION

The autoregressive-moving average (ARMA) model uses present and past inputs with past outputs to determine the present output. The ARMA process is represented by the difference equation

$$Y(n) + a_1\, y(n-1) + ... + a_M\, y(n-M) = b_0\, u(n) + b_1\, u(n-1) + ... + b_K\, u(n-k)$$
(1)

Where $a_1, ..., a_M$ and $b_0, ..., b_K$ are called the ARMA parameters. Further information regarding the development of the ARMA model can be found, e.g., in [2, 3, 4]. The ARMA system described by the difference equation (2) can be re-written in the form:

$$y(n) = -\sum_{i=1}^{M} a_i \; y(n-i) + \sum_{i=0}^{k} b_i \; u(n-i)$$

(2)

Compact (2) by writing it in regression model form where the regressor vector is defined by

$$\varphi(n) = (y(n-1), ..., y(n-M), u(n), ..., u(n-K))$$

(3)

and the vector parameter by

$$\theta = (-a_1, ..., -a_M, b_0, b_1, ..., b_K)$$

(4)

Given (3), (4) and (5) can be expressed as an inner product

$$y(n) = \theta\varphi^T (n)$$

(5)

Note that the parameter vector in (6) is not dependent on n; therefore, no index is specified for θ in (4) or (5), nor will one be specified in the remainder of this text.

The output-error equivalence method is used to calculate the estimated output and update the parameter estimates. It is achieved by replacing the parameter vector with its time dependent estimate, $\hat{\theta}$ as in

$$\hat{y}(n) = \hat{\theta}\varphi^T (n) \tag{6}$$

Here $\hat{\theta}$ is defined as the most recent parameter vector estimates

$$\hat{\theta} = (-\hat{a}_1, ..., -\hat{a}_M, \hat{b}_0, \hat{b}_1, ..., \hat{b}_K)$$

(7)

Subsequently the output estimate (7) can be expressed as

$$\hat{y}(n) = -\sum_{1=1}^{M} \hat{a}_i \; y(n-i) + \sum_{i=0}^{K} \hat{b}_i \; u(n-i)$$

(8)

The parameter estimates are adjusted incrementally to reduce the error between (5) and (6). The idea is to update the previous parameter values for $\hat{\theta}$ so that the residual error ultimately approaches to zero. This is accomplished using a gradient search technique with a simple quadratic to calculate the error between the models and system outputs.

3. DATA PRE FILTERING

Data pre-filtering is systematically employed in linear identification, mostly for anti aliasing reasons and to reduce the effect of high frequency disturbances in order to increase the signal to noise ratio. It is well known from the linear identification literature [2] that pre filtering may be used to improve model accuracy in a specified frequency band e.g. as a result of control specification which narrow the frequency region over which accurate models are actually needed. In [3] it is pointed out that data prefiltering will introduce bias into the estimates even in the noise free case and a detailed analyses of mean level induced bias is performed. The addition of filters in the identification strategies reduces the impact of noise on parameter estimates.

3.1 Prefiltering using moving average process

IIR prefilter is used in the application of signal frequency estimation [5]. The process of IIR filtering will add extraneous poles to the original signal model and the computational burden is much increased since the estimator order has to be much higher than the order of the estimator plus the IIR prefilter. FIR prefiltering methods [6] were proposed to increase the signal to noise ratio of original measurement before the singular value decomposition (SVD) based Prony's method was applied to estimate the location of single pose. By FIR prefiltering not only the AR but also the MA coefficient functions are estimated [7]. The optimal moving average coefficients are extracted from the estimated AR coefficients by solving the normal equation SVD and spectral factorization in the frequency domain. Liu and Doraiswami in[8] have emphasized moving average FIR prefiltering. Algorithm given by them is as follow:

3.1.1 Robust estimation of the denominator coefficient

The discrete time domain signal model is assumed to be

$$s(n) = \sum_{i=1}^{M_0} a_i \; s(n-i) + \sum_{i=0}^{K} b_i \; \delta(n-i), \qquad \text{I} \le M_0$$

(9)

Where $s(n)$ is the signal, $\{a_i\}$ is the AR coefficient set and $\{b_i\}$ is the MA coefficient set, respectively. $\delta(n)$ is Kronecker delta function. The measurement model is

$$y(n) = s(n) + v(n), \qquad 0 \le n \le N-1$$

(10)

By using LPCA, the AR coefficients $\{a_i\}$ can be obtained from $y(n)$ in the sense of minimum least squared error. The derivation of standard form LPCA can be found in [7], [8], [9]. LPCA provides high resolution estimates when the signal SNR is high. However its performance degrades severely when the SNR is low.

Let us contact a new signal $\bar{y}(n)$ which is the summing average of the measurement $y(n)$ with data length L, that is

$$\bar{y}(n) = \frac{1}{L} \sum_{i=n}^{n+L-1} y(i)$$

(11)

Hence,

$$\bar{y}(z) = F(z)s(z) + F(z)v(z)$$

(12)

Where

$$F(z) = \frac{(1 + z + \ldots z^{L-1})}{L}, \quad \text{or}$$

$$F(z) = f_1 + f_2\, z + \ldots + f_L\, z^{L-1}$$

(13)

In the general case clearly, the $\bar{y}(z)$ and $s(z)$ have the same signal poles location excepting $\bar{y}(z)$ has additional poles due to the measurement noise. Hence, the LPCA can be applied to $\bar{y}(n)$ instead of $y(n)$. The advantage of translating $y(n)$ into $\bar{y}(n)$ is that the strong perturvation of the noise with high level of variance can be greatly reduced if the noise mean η is small (considering the prefiltering be a moving average process) and it is more natural using a linear ARMA model to describe $\bar{y}(n)$ other than the stochastic process $y(n)$ with stronger noise pattern.

3.1.2 Estimation of numerator Coefficients

The estimation procedure of numerator coefficients is also divided into two steps. First, the numerator coefficients of the smoothed measurement $\bar{y}(n)$ are obtained through the similar procedure describe in section II by minimizing the error between $\bar{y}(n)$ and the impulse response $\bar{h}(n)$ generated by the estimated denominator coefficients. Second, the numerator coefficients of the original signal $s(n)$ are calculated by using a squared error minimization and spectral factorization process in the frequency domain.

The ARMA model predictor output of $\bar{y}(n)$ is given by $\vec{y}_{arma}(n) = Z^{-1}\{\bar{b}(z)/\bar{a}(z)\}$, where

$$\bar{a}(z) = 1 - \sum_{i=1}^{M} \bar{a}_i \ z^{-i}, \qquad \bar{b}(z) = \sum_{i=0}^{M} \bar{b}_i \ z^{-i} \tag{14}$$

$\vec{y}_{arma}(n)$ can be expressed by

$$\vec{y}_{arma}(n) = \sum_{i=0}^{M} \bar{b}_i \ \bar{h}(n-i) \tag{15}$$

Where $\bar{h}(n) = Z^{-1}\{1/\bar{a}(z)\}$, that is

$$\bar{h}(n) = \sum_{i=1}^{M} \bar{a}_i \ \bar{h}(n-i) + \delta(n), \qquad \bar{h}(0) = 1$$

$\delta(n)$ is the Kronecker delta.

Eq. 10 is expressed in the matrix form $\vec{y}_{arma} = \overline{H}\bar{b}$, where $\bar{b} = [\bar{b}_0 \bar{b}_1 \ldots \bar{b}_M]^T$, and

$$\hat{Y}_{arma}(n) = \begin{bmatrix} \hat{y}_{arma}(n) \\ \hat{y}_{arma}(n+1) \\ . \\ . \\ \hat{y}_{arma}(n+N-1) \end{bmatrix}$$

$$H = \begin{bmatrix} h(n) & h(n-1) & \ldots & h(n-M) \\ h(n+1) & h(n) & \ldots & h(n-M+1) \\ . & . & \ldots & . \\ . & . & \ldots & . \\ h(n+N-1) & h(n+N-2) & \ldots & h(n+N-M-1) \end{bmatrix} \tag{16}$$

By minimizing $j(\bar{b}) = (\bar{Y} - \overline{H}\bar{b})^T \ (\bar{Y} - \overline{H}\bar{b})$, where $\bar{Y} = [\bar{y}(n)\bar{y}(n+1)\ldots\bar{y}(n+N-1)]^T$. The optimal numerator estimates are obtained from the solution of normal Eq. $\overline{H}^T \ \overline{H}\bar{b} = \overline{H}^T \ \bar{Y}$

The above procedure forms an ARMA model estimate of the smoothed measurement $\bar{y}(z)$, or $\bar{s}(z)$ when the mean of $v(n)$ is zero. However, our objective is to obtain the ARMA model parameters of the original signal $s(n)$. The estimate of $s(z)$, namely $\hat{s}(z)$ is calculated from

$$\hat{s}(z) = \frac{\bar{s}(z)}{F(z)}$$

This operation is numerically unstable since in general $F(z)$ will have unstable roots. Hence the estimate of $s(z)$ is obtained from

$$\min \left\| \hat{s}z - \frac{\bar{s}(z)}{F(z)} \right\|$$

Solution of above formula is obtained by spectral factorization which yields

$$\hat{s}(z) = \left(\frac{\bar{s}(z)}{F(z)} \right)$$

Where $(G(z))$, denotes the stable part of $G(z)$.

3.2 Non arithmetic filtering

Applying filters to input and output signals can enhance parameter estimation but may concentrate only on specific frequency bands. Low-, high-, and band-pass filters isolate portions of signals providing opportunities for closer approximations. This technique provides parameter estimates for specific frequency ranges, sacrificing model optimality over the entire operating spectrum. Filters used in this fashion reduce the effects of noise and outliers by obliterating data. Filtering techniques are applied to all of the pertinent signals before parameter estimation is attempted. However, traditional filters used in this manner have the potential to alter the data while attempting to reduce noise and important data can be lost or compromised and introduce new problems.

A class of non arithmetic filters has been developed by the authors [4] and applied to signal smoothing for improved parameter estimation and system identification. This class of filters has been proven successful in reducing induced computational errors such as coefficient quantification and round off. Additionally, this class of filters characteristically eliminates impulsive and Gaussian distributed noise.

The non arithmetic filtering theory employed [4] requires some basic assumptions. It is assumed that data sequences come from a finite totally ordered set (TOS) of values *S,* and that any subset of data sequence values may be ordered. The sampling rate used in measuring the signals is unspecified and is not restricted to a uniform rate. There exists a distance function (metric) on *S* which can be user defined. Also**,** a median-type operator is defined so as to always produce an element residing in *S,* unlike the usual median operator which may use averaging [10].

The non arithmetic filtering technique demonstrated in [10] evolved as a natural extension of the weighted majority with minimum range (WMMR) filter [4]. The WMMR utilizes a technique of dividing filtering windows into overlapping sub windows in combination with a weighting scheme. The WMMR prefacers' very well, but its highly computational methodology introduces noise in the form of round-off and averaging errors. This consequence led to the development of a nonarithmetic filter. A straight-forward demonstration of the application of nonarithmetic filtering can be obtained using MATLAB System Identification Toolbox [I1].

3.3 Robust H_2 and H_∞ Filtering

By using H_2 and H_∞ filtering problem the controller is designed such that worst case induced L_2 gain from process noise to estimation error is minimized. Here an upper bound is tried first and then bound is minimized using techniques based on Riccati equations or LMIs. For a class of uncertain continuous – time systems defined by

$$\begin{bmatrix} \dot{x}(t) \\ y(t) \end{bmatrix} = \quad M(\Delta(t)) \begin{bmatrix} x(t) \\ d(t) \end{bmatrix}$$
(17)

Where $x(0) = x_0$, and $M(\Delta(t))$ is given by

$$M(\Delta(t)) = \begin{bmatrix} A & B \\ C & D \end{bmatrix} + \begin{bmatrix} L_1 \\ L_2 \end{bmatrix} \Delta(t)(I - H\Delta(t))^{-1} \begin{bmatrix} R_1 & 0 \end{bmatrix}$$
(18)

Where $x(t) \in \Re^n$

Are the states, $d(t) \in \Re^{nd}$ is the process and measurement noise, and $y(t) \in \Re^m$ are the measurements. $A, C, B, D, L_1, L_2, R_1,$ and H are known constant matrices with appropriate dimensions. The uncertainty matrix $\Delta(.)$ is norm bounded, time - varying and with problem – specific block – diagonal structure. The set of uncertain matrix values are denoted by

$\Delta := \{$ diag$\left(\delta_1 I_{q1} \ldots\ldots \delta_l I_{ql} . \Delta_{l+1} \ldots\ldots \Delta_{l+\int}\right) : \|\Delta\| \le 1 . \delta_1 \in \Re . \Delta_i \in \Re^{q_i \times q_i} \} \subset \Re^{n_p \times n_p}$.

Associated with Δ, define subspaces S and G as

S$:= \left\{ diag\left(S_1 \ldots S_l . \lambda_{1Iql} + 1 \ldots\ldots \lambda_s I_{ql+\int}\right) : S_i = S_i^T \in \Re^{q_i \times q_i} . \lambda_j \in \Re \right\}$

$G := \left\{ diag\left(G_1 \ldots\ldots G_l . 0_{ql+1} \ldots 0_{ql+\int}\right) : G_i = -G_i^T \in \Re^{q_i \times q_i} \right\}$.

Note that if $l = 0, f = 1,$ then $\Delta = \left\{ \Delta \in \Re^{n_p \times n_q} : \|\Delta\| \le 1 \right\}$, referred to as *unstructured* uncertainty, denoted Δ_u. In this case, S$= \left\{ \lambda I : \lambda \in \Re \right\}$, and every element of G is 0.

This LFT representation of uncertainty is widely used in robust control theory; for instance, in [12] and [13]. In this note, we assume the representation (17) is well – posed over Δ, meaning that $\det(I - H\Delta) \ne 0$ for all $\Delta \in \Delta$. Under this assumption, the uncertain part can be isolated from known part and the system written as

$$\dot{x}(t) = Ax(t) + Bd(t) + L_1 p(t)$$

$$y(t) = Cx(t) + Dd(t) + L_2 p(t)$$

$$q(t) = R_1 x(t) + Hp(t)$$

$$p(t) = \Delta(t) q(t).$$

$$\Delta(t) \in \Delta$$

Given $L \in \Re^{r \times n}$, the objective is to design a linear, full order filter to estimate $z(t) := Lx(t).$ The filter structure is constrained to: $\dot{\hat{x}}(t) = A_\int \hat{x}(t) + B_\int y(t) . \hat{z}(t) = L_\int \hat{x}(t)$, where $A_\int \in \Re^{n \times n}$, $B_\int \in \Re^{n \times m}$ and $L_\int \in \Re^{r \times n}$ are constant matrices. Define estimation error $e(t) := z(t) - \hat{x}(t)$. Let $\eta(t) := \left[x(t)^T \; \hat{x}(t)^T \right]^T$ denote the states of the augmented system

$$\dot{\eta}(t) = \left[\overline{A} + \overline{L} \Delta(t)(I - H\Delta(t))^{-1} \overline{E} \right] \eta(t) + \overline{B} d(t)$$
(19)

$$e(t) = \overline{C} \eta(t)$$
(20)

Where

$$\overline{A} = \begin{bmatrix} A & 0 \\ B_{\int} C & A_{\int} \end{bmatrix}, \quad \overline{L} = \begin{bmatrix} L_1 \\ B_{\int} L_2 \end{bmatrix}, \quad \overline{E} = \begin{bmatrix} R_1 & 0 \end{bmatrix}$$

$$\overline{B} = \begin{bmatrix} B \\ B_{\int} D \end{bmatrix} \qquad \overline{C} = \begin{bmatrix} L & -L_{\int} \end{bmatrix}.$$

In [14], two problems are considered. The H_2 problem (similar to the Kalman filter) has a stochastic interpretation. In (17), assume d is zero mean white noise, with $\varepsilon\left[d(t)d(l)^T\right] = \delta(t-l)I_{nd}$, where $\delta(t)$ is the Dirac Delta function. The H_2 performance objective is $\sigma := \lim_{T\to\infty} \sigma_T$, where $\sigma_T = \sup_{\Delta(.)\in\Delta} \varepsilon\left\{(1/T)\int_0^T e^T(t)e(t)dt\right\}$. The notation $\sup_{\Delta(.)\in\Delta}$ denotes the supremum over all piecewise continuous functions $\Delta : \Re \rightarrow$. The design objective is to minimize σ (by choice of $A_{\int}$, $B_{\int}, L_{\int}$) subject to (19) and (20). The H_∞ problem defines the performance measure as $\rho := \sup_{\Delta(.)\in\Delta} \sup\|d\|_2 \neq 0\left(\|e\|_2 / \|d\|_2\right)$. (a worst case induced L_2 operator norm). The design objective is to minimize ρ (by choice of $A_{\int}, B_{\int}, L_{\int}$) subject to (19) and (20).

In [14] it is assumed that (17) is quadratically stable, namely the existence of a positive-definite matrix P such that $A_\Delta^T P + PA_\Delta \prec 0$ for all $\Delta \in \Delta$. Here, $A_\Delta := A + L_1\Delta(I - H\Delta)^{-1}R_1$. This is a typical assumption for all work in this area.

4. CONCLUSION

Three main prefiltering techniques to reduce the effects of process noise on parameter estimation have been presented from the literature. A known – arithmetic filtering technique used only input and output data with no a – priori or posteriori system noise information. Where as robust H_2 and H_∞ filters can be used not only for the reduction of the noise but also considers unstructured, non bounded uncertainty case, the upper bounce are directly minimized yielding less conservative results finally one can pose robust filtering as a more general robust control problem, simply used the adhoc iteration methods obtaining adequate results.

5. REFERENCES

1. J.J.Shynk. “Adaptive IIR Filtering”. IEEE ASP Magazine, April, 1989.
2. Ljung, L. “Svstem Identification: Theow for the User”. T. Kailith ed., New Jersey: Prentice-Hall, 1987.
3. Haykin, S.S. “Adaptive Filter Theory”. 2nd ed. Englewood Cliffs, New Jersey: Prentice-Hall, 1991.
4. Longbotham, H.G. and D. Eberly. “Statistical Properties, Fixed Points, and Decomposition with WMMR Filter”. JOW of Mathematical Imaging and Vkion: 99- 116,1992.
5. S. Kay. “Accurate frequency estimation at low signal to noise ratio". IEEE Trans. ASSP: 540-547, June 1984.
6. R.Kumaresan and Y. Feng. “FIR prefiltering improves Prony's method". IEEE Tran. Signal Processing, 39 (2): 736-741, March, 1991.
7. R.Kumaresan and D.W.Tufts. “Estimating the parameters of exponentially damped sinusoids and pole zero modelling in noise”. IEEE Tran. ASSP, 30: 833- 840, Dec. 1982.
8. R. Doraiswami and W. Liu. “Real-time estimation of the parameters of power system small signal oscillations". IEEE Power Engineering Society IEEEIPES 1992, Winter Meeting, New York, USA, and also to appear in IEEE Trans. on Power Systems.
9. B. Friedlander and B. Porat. “The modified Yula- Walker method of ARMA spectral estimation". IEEE Tran. Aerospace and Electronic Syst., 20:158-173, Mar. 1984.

10. Jack J., Harris and Cheryl B. Schrader. "Application of nonarithmatic filtering techniques for improved system identification". 37th IEEE, Florida USA, Dec. 1998.
11. Ljung, L. User's Guide: MATLAB Svstem Identification Toolbox, The Mathworks, Inc., 1995.
12. L. EI Ghaoui and G. Calafiore. "Robust filtering for descrite-time systems with bounded noise and parametric uncertainty". IEEE Trans. Autom. Control, 46 (7): 1084-1089, Jul, 2001.
13. K. Zhou, J.C. Doyle, and K. Glover. "Robust and Optimal control". Upper Saddle River, NJ: Prentice-Hall, 1996.
14. J. C. Geromel, M.C. De oliveira, and J. Bernussou. "Robust filtering of descrete -time linear systems with parameter dependent Lyapunov functions". SIAM J. Control optim.41 (3): 700-711, 2002.