# System Identification Coursework

**Yi Zhou 9971603**

## Q1

The ARX model is defined as

$$A(z)y(t) = B(z)u(t-nk) + e(t)$$

In system identification toolbox, the array (na nb nk) is selected as the number of poles, the number of zeros minus one and time delay respectively.

Testing with the orders of (1:4 1:3 1:10), the best-fit model selected using least square method is shown as below.

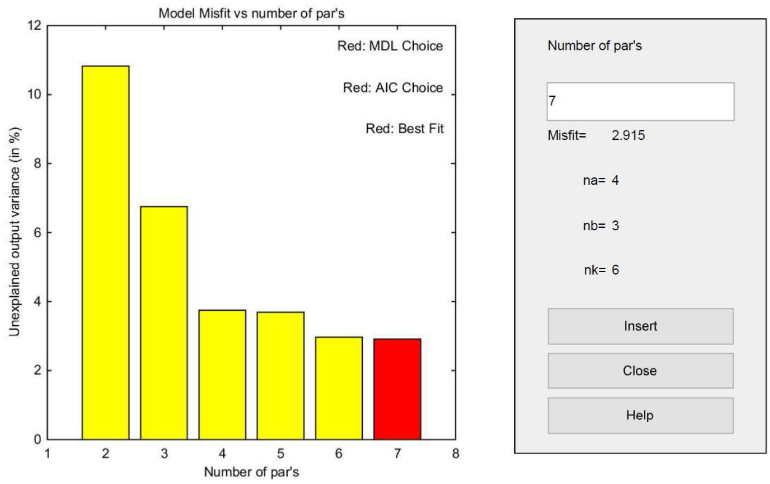


Figure 1 Best fit parameter

## Q2

The selected ARX model is

$$A(z)y(t) = B(z)u(t) + e(t)$$

$$A(z) = 1 - 0.6277 \cdot z^{-1} - 0.3087 \cdot z^{-2} + 0.1548 \cdot z^{-3} + 0.07447 \cdot z^{-4}$$

$$B(z) = 0.07737 \cdot z^{-6} + 0.1206 \cdot z^{-7} + 0.07305 \cdot z^{-8}$$

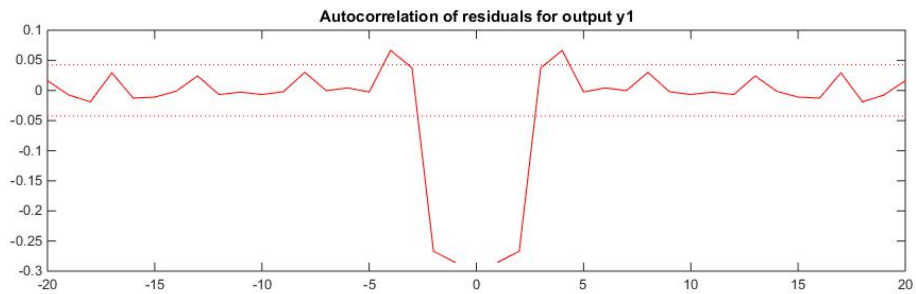


Figure 2 Autocorrelation of residuals of ARX model

According to the auto correlation plot, the sharp peak indicates that the dynamics of the error is not suitable for the data sets. Thus the assumption of the dynamics of error in ARX model is not correct. Alternatively, the OE model, whose error is directly added to the output, is tested in Q3.

## Q3

The OE model, i.e.output error model, is defined as

# System Identification Coursework
**Yi Zhou 9971603**

$$y(t) = \frac{B(q)}{F(q)} \cdot u(t - nk) + e(t)$$

In system identification toolbox, the array (nb nf nk) is selected as the order of numerator, denominator and time delay respectively.

To find the coefficient, ten combinations are tested. Under the condition that the system has at most 4 poles, the nf should range from 1 to 4. The nb should be less than the nf. The delay term nk can be estimated as near 6, according to the ARX model. The simulated results are shown in table 1.

| Selected Order OE : (nb nf nk) | Fitting Rate | Maximum error of Auto-correlation* | Maximum error of cross-correlation* |
|---|---|---|---|
| (3 4 4) | 86.87 | 0.03032 | 0.05857 |
| (3 4 5) | 86.93 | 0.03279 | 0.01955 |
| (3 4 6) | 86.93 | 0.03285 | 0.01937 |
| (3 4 7) | 76.08 | 0.01966 | 0.8386 |
| (1 3 6) | 86.29 | 0.03265 | 0.2246 |
| (1 4 6) | 86.85 | 0.02953 | 0.075 |
| (2 3 5) | 86.45 | 0.02804 | 0.1668 |
| (2 3 6) | 86.92 | 0.03282 | 0.0197 |
| (2 4 5) | 86.87 | 0.03042 | 0.05862 |
| (2 4 6) | 86.92 | 0.03249 | 0.02736 |

Table 1. The properties of of different orders

* Absolute value of each term. The 99% confidence interval for auto-correlation is [-0.03001,0.03001]. The 99% confidence interval for cross-correlation is [-0.03,0.03]

**Q4**

According to the fitting rate and two errors, three cases ,which are marked as red, are selected at first. It can be seen from the table that the time delay should equals to 6. Thus case (3 4 5) is dropped. Comparing the poles and zeros map, the case (3 4 6) is equivalent to to the case (2 3 6), due to the zero-pole cancellation. Therefore, the case (2 3 6) is chosen as best fit.

$$y(t) = \frac{0.007658 \cdot z^{-6} + 0.04157 \cdot z^{-7}}{1 - 1.659 \cdot z^{-1} + 1.005 \cdot z^{-2} - 0.2283 \cdot z^{-3}} u(t) + e(t)$$

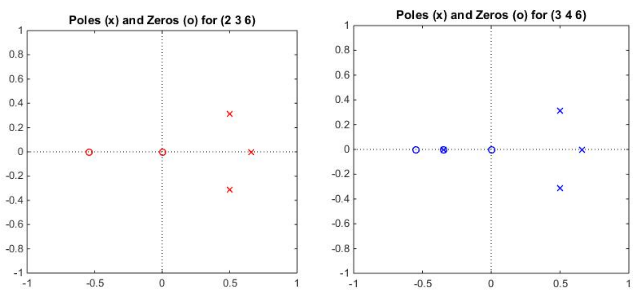


Figure 3 poles and zeros for (2 3 6) and (3 4 6)

According to residual analysis plot, the autocorrelation of residuals for output is almost inside the 99% confidence interval, indicating the error model of oe is preferable. The cross correlation lays inside the 99% confidence interval completely, indicating the selected time delay is suitable for the system.

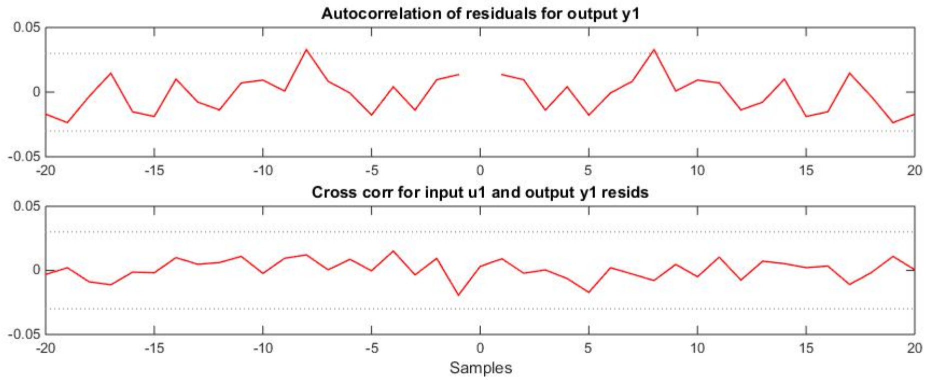


Figure 4 Autocorrelation and cross correlation of selected oe model

**Q5**

Nonparametric method directly estimate the system from frequency domain. It does not assume any structure of the system, other than it is linear[1].

It can be realized by the function spa, in short of spectral analysis, from the system identification toolbox. However, in order to better understand it, the matlab code is programmed by myself. The key idea is simply taking the FFT,i.e.fast Fourier transform, of the output and input signals. Then the transfer function can be computed as

$$G(\omega) = \frac{FFT(y)}{FFT(u)}$$

For better illustration of bode plot, a tricky method, referred as hamming window in textbook, is used to plot the smooth curve. It should be noticed that the FFT of PRBS is sparse in frequency domain, which means there is a series of zeros between two peaks. Furthermore, the model is linear so that the frequencies of output should be same as those of input. Thus, what is expected from the measured frequency should be the frequency of the those peaks. Only these data points in the transfer function , namely from index of 30 to the end with interval of 68, are used for Bode plot.

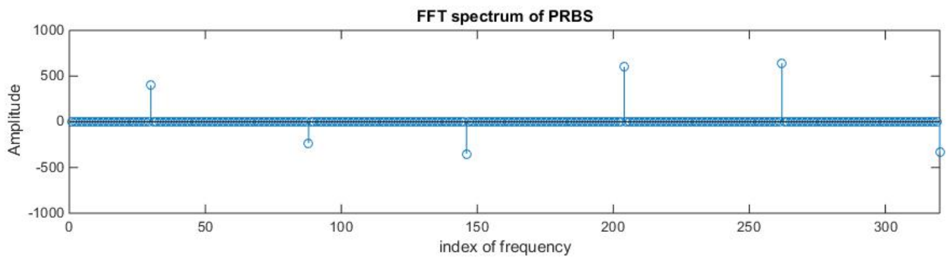


Figure 5 Spectrum of PRBS signal

The bode plots of non-parametric model and ARX model are compared in figure 6. Both the amplitude and phase curves fits well in low frequency. However, in high frequency, there are discrepancies between these two methods, that is the ARX model predicts a bounce. In other words, it predicts the amplifying of high-frequency noise. The assumed dynamics of the error should contribute to this abnormal phenomenon. Therefore, it can be inferred that the ARX model is not suitable for this system. In contrast, the bode plot of oe model is consistent with the non parametric method. Thus, the oe model should better fit this system.

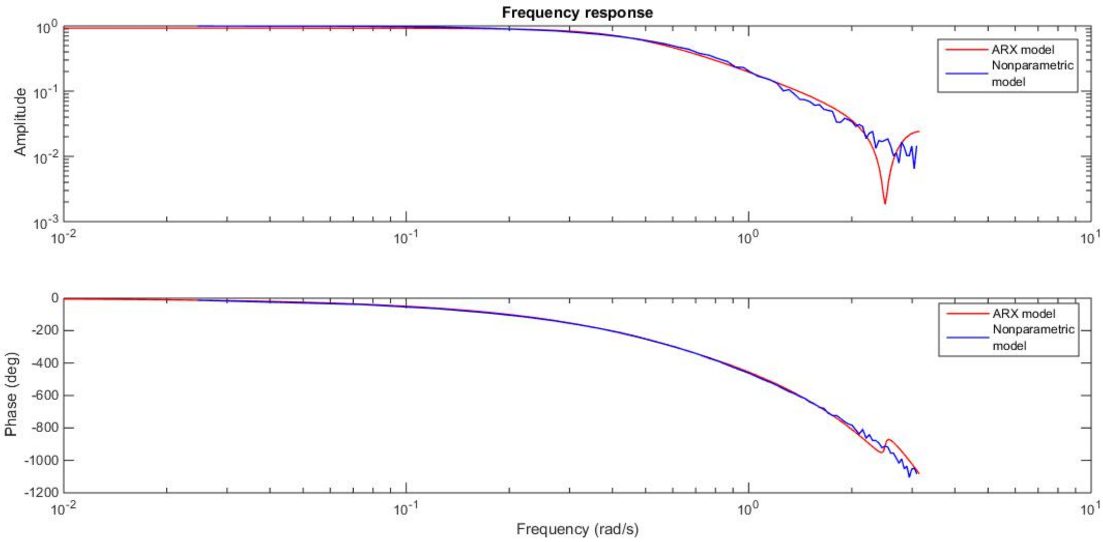


Figure 6 Bode plot for ARX model and nonparametric model

## Q6

To test the nonlinear effect of the system, the homogeneity principle method[2] is used. Ten times the input and output data. Again computing the new frequency domain model. Suppose the system is linear, the harmonics should be merely generated by the noise. However, it can be seen that the peak of second harmonic and higher do change when the input and output change. This phenomenon indicates there exists some nonlinear terms in the system.

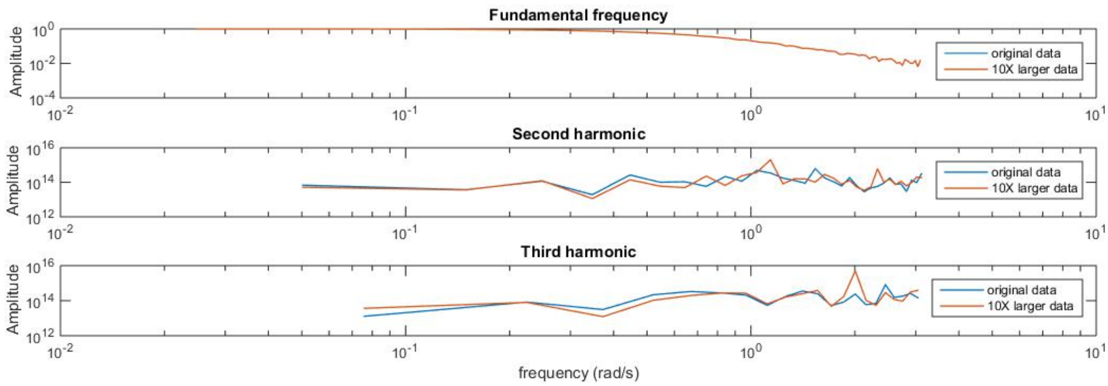


Figure 7 The Bode plot for fundamental frequency and harmonics

Reference

[1] Ljung, L. (2002). System identification toolbox user''s guide. Journal of Aircraft.

[2] Vanhoenacker, K., Schoukens, J., Swevers, J., & Vaes, D. (2002). Summary and comparing overview of techniques for the detection of non-linear distortions. *Die Pharmazie, 57*(7), 1241-1256.