isvr Institute of Sound and Vibration Research

Multichannel Spectral Factorisation and System Identification for Active Control

I. Nagashiro

ISVR Technical Memorandum No 946

March 2005

University of Southampton

SCIENTIFIC PUBLICATIONS BY THE ISVR

Technical Reports are published to promote timely dissemination of research results by ISVR personnel. This medium permits more detailed presentation than is usually acceptable for scientific journals. Responsibility for both the content and any opinions expressed rests entirely with the author(s).

Technical Memoranda are produced to enable the early or preliminary release of information by ISVR personnel where such release is deemed to the appropriate. Information contained in these memoranda may be incomplete, or form part of a continuing programme; this should be borne in mind when using or quoting from these documents.

Contract Reports are produced to record the results of scientific work carried out for sponsors, under contract. The ISVR treats these reports as confidential to sponsors and does not make them available for general circulation. Individual sponsors may, however, authorize subsequent release of the material.

UNIVERSITY OF SOUTHAMPTON

INSTITUTE OF SOUND AND VIBRATION RESEARCH

SIGNAL PROCESSING & CONTROL GROUP

Multichannel Spectral Factorisation and System Identification for Active Control

by

Iwao Nagashiro

ISVR Technical Memorandum No.946

March 2005

Authorised for issue by
Prof S J Elliott
Group Chairman

ACKNOWLEDGMENT

I am grateful to Professor S. J. Elliott of ISVR, university of Southampton, for his suggestions, advices on this topic and many instructive discussions, also to all members of Signal Processing and Control Group of ISVR, for their help and kindness during my visiting of ISVR.

ABSTRACT

In recent years, a growing field of research in "adaptive control system" has resulted in a variety of adaptive automatons whose characteristics resemble certain characteristics of living systems and biological adaptive processes. Certainly adaptive control system is not limited to a single channel. A multichannel adaptive control system has been developed for many practical applications.

An adaptive algorithm called as multichannel filtered reference LMS adaptive algorithm is reported by S.J. Elliott. But when using this algorithm, the problems are the convergence speed may be limited by correlation between the reference signals and the control plant must be known very precisely. For this reason we consider to remove the correlation between input signals by using an inverse system of spectral factorisation or spectral estimation, also we try to identify control plant by using subspace method and remove cross talk between each channel.

In this report, based on a detailed investigation of past work, methods of multichannel spectral factorisation have been reviewed, and their relevance to adaptive feedforward control has been discussed. In order to achieve a good improvement for multichannel adaptive control problem, we first discussed some theoretical topics of control system and signal processing given as following.

1. multichannel spectral factorisation
2. multichannel spectral estimation
3. multichannel system identification

After this, we try to show some applications concern with these theoretical results.

4. spectral estimation of road noise in car and transfer function estimation in car.
5. discuss a multichannel active noise cancelling system in car.

These topics are discussed respectively in this report. Each chapter consists of basic theory and numerical example of computer simulation.

CONTENTS

FIGURES

Chapter 1 INTRODUCTION

1.1 General Background

Geophysicists gather sounding data from acoustics, radar and seismology to estimate and construct geophysical images. These images are used to visualise petroleum and mineral resource prospects, subsurface water, contaminant transport (environmental pollution), archaeology, lost treasure, even graves. In most practical situations they operate sounders along tracks on the earth surface (or track in the ocean, air) and record data from multiple sensors, i.e. they process multichannel input data to objectives deposited under earth surface. But can geophysicists feedback some effects to earth (ocean, air) based on the information they find, to control earthquake, tsunami or weather automatically? This is a problem of multichannel automatic control system.

Doctors gather time sequence data from electrocardiogram (ECG) or brain waves to describe work pattern of patient organ. These patterns are used to visualize movement, position or shape of patient organ, that will help make decisions during medical examination. In most practical measurements, they record data from multiple sensors, i.e., they process multichannel input data to visualise the situation in a patient's body. Then they will give some treatment by chemical medicine or surgical operation. But can doctors feedback some treatments based on the information they find, to control human organ or brain automatically? This is also a problem of multichannel automatic control system.

In recent years, a growing field of research in "adaptive control system" has resulted in a variety of adaptive automatons whose characteristics resemble certain characteristics of living systems and biological adaptive processes. These adaptive control systems have adjustable structure so that its behaviour or performance can be improved through contact with its environment, according to some desired criterion. Certainly adaptive control system is not limited to a single channel. A multichannel adaptive control system can be considered as following.

> For a given multichannel input signal, a multichannel adaptive control system can adjust these parameters inside its own structure to achieve an optimum state by means of a given cost function.

There many adaptive algorithm for signal channel adaptive control system. For example, least mean square (LMS) algorithm, recursive least square (RLS) algorithm et al. Many books said that it is easy to extend these algorithms from signal channel to multichannel, but things may not so easy as these authors expected. Although one can write a signal channel in vector or matrix form that corresponds a multichannel signal, but we must think about some condition for making these algorithm. The most important difference from multichannel system is that signals might be effected each other, for example, the input signals might have strong correlation between channels, a control signal might not only received by its own channel but also all of other channel, also the criterion might not describe the difference of each channel. Adaptive control systems can adaptive themselves depending on the processed information from signal, if it can not extract correct information, how can the system work as expected. Furthermore, because the correlation between individual channels, computing complexity might increase very quickly, generally not equal to channel number times one channel computation, it might be a power of one channel computation.

The problem we will discuss in this report is how to design a multichannel adaptive control system to ensure it works well and achieve a good performance.

1.2 General Block Diagram

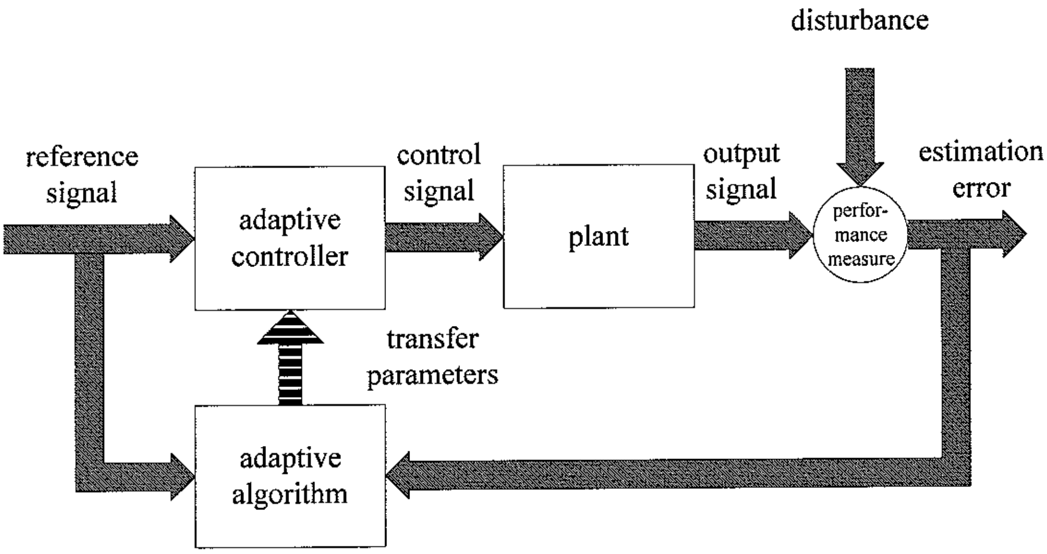
The system block diagram of our problem is shown in Figure 1. There are four parts in this diagram. Function of each individual part is described as following.

Adaptive controller: From reference signal vector, adaptive controller generates control signal vector to output to plant. This part has adjustable parameters in its own structure which will be updated by adaptive algorithm in real time.

Plant: Plant is a part to be controlled. When the control signal vector from adaptive controller is applied, plant will process or affect it to give output signal vector that should be a best estimation of disturbance.

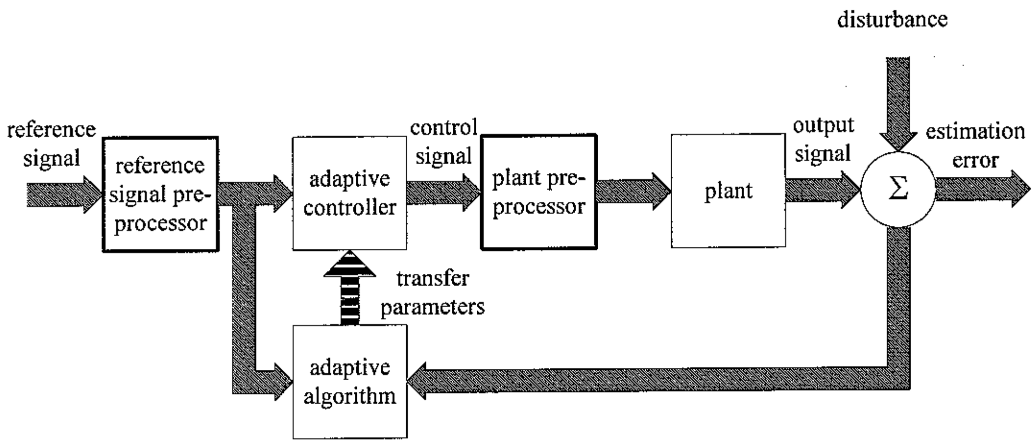
Performance measure: At the performance measure point, the plant output signal vector is combined with disturbance vector, this will result a residual error signal vector that will be used as a performance measurement of this control system.

Adaptive algorithm: This is the most important part in adaptive system. In this part, based on input signal vector, error signal vector and information of plant, adaptive algorithm will give update values of adjustable parameters in adaptive controller, according to a given cost function of performance.



• Figure 1 General block diagram of multichannel adaptive control system

It has been suggested by S. J. Elliott[1)] that the adaptive speed of the multichannel controller shown in Figure 1 could be improved by preconditioning the reference signals to extract their uncorrelated white noise components and preconditioning the plant with the inverse of its minimum phase component as shown in Figure 2.



• Figure 2 General block diagram of the preconditioned adaptive control system

The preconditioning required for the reference signals can be calculated using a spectral factorisation of their spectral density matrix and the plant preconditioning required the calculation of the minimum phase / all pass decomposition, or equivalently inner/outer factorisation. These topics are discussed respectively in later chapters of this report.

1.3 Topics of This Research

1.3.1 Digital Signal Processing Topics

For multichannel system, there are a number of input signals, generally with strong correlation between individuals signal. So the first problem is that if the correlation can be removed. The basic idea is to assume that the multichannel signal is generated from a multichannel uncorrelated white signal. If we can find the system for this and the inverse system exists, we can use the inverse system to get a multiple uncorrelated white signal from given multichannel input signal and use it to drive the adaptive system. There are several type models for the system structure, for example, moving average (MA) model and autoregress (AR) model. The MA model describes the spectral of input signals in a finite matrix polynomial. From the spectral expression, we can find a system transfer function matrix, this is a problem called multichannel spectral factorisation. Also the AR model describes the spectral matrix in an inverse of finite matrix polynomial. From this spectral expression, we can find a system transfer function matrix, this is a problem called multichannel spectral estimation. In our research, we tried to discuss both topics of multichannel spectral factorisation and multichannel spectral estimation.

1.3.2 Control Theory Topics

Multichannel adaptive control system treats a multi-input and multi-output (MIMO) plant as control objective. The first thing is to find a MIMO transfer function of the plant. This is a topic of system identification, especially subspace method developed in recent year. The second thing is remove cross-talk between multiple input signals and between multiple output signals. This problem is equivalent to inner-outer factorisation problem in general. If we can get an inner system and an outer system of given plant, we can try to remove inter-effect in plant by using inverse system of inner system and outer system. That will be much effective to improve the performance of adaptive system.

1.3.3 Application: Active Noise Control

As an application, these topic discussed above are applied to a multichannel active noise control (ANC) problem. For the noise cancelling problem in an enclosure, in order to achieve satisfactory performance in a relatively large dimension space, a number of cancelling loudspeakers and a number of error sensors are used, i.e., a multichannel ANC system is necessary. The feedforward ANC system consists of a number of reference microphones, a number of error microphones and a number of cancelling loudspeakers, and a multiple input multiple output adaptive controller which weight parameter matrix is able to be updated by an adaptive algorithm. In order to cancel the primary noise, an adaptive controller gives a secondary source by weighting each reference signal. The secondary source drives the plant, which is acoustic propagation space, then reach the error sensors, where these combined with the primary noise, thus resulting in the cancellation of both noises.

1.4 Outline of This Report

In this report, based on a detailed investigation of past work, we aim to find a better solution for multichannel adaptive control problem. For this purpose, we first discussed some theoretical topics of control system and signal processing given as following.

1. multichannel spectral factorisation

2. multichannel spectral estimation

3. multichannel system identification

After this, we try to show some applications concern with these theoretical results.

4. spectral estimation of road noise in car and transfer function estimation in car

5. discuss a multichannel active noise cancelling system in car

Following these topics, this report is organized in seven chapters. The outline each chapter is given as follows.

In chapter 1, a general background is given to describe the problem we will discuss in this report. Then we show a general system block diagram to refine the problem more precisely and also give the connection to adaptive filter, signal processing and control theory. At end of this chapter, an outline of this report is sketched.

In chapter 2, after giving a brief background on multichannel spectral factorization, the problem formulation in its basic form and canonical form are defined. Then some the necessary and sufficient conditions for these problem forms are described. In section 2, some important past works, including those of D.C. Youla, M.C. Davis, E.A. Robinson and T. Kailath, are described briefly. In section 3, an approach via a matrix binomial expression is provided. Basic principle is introduced with a detailed proof. Two flow charts are given to show the calculation procedure. Also a numerical example is solved step by step. In section 4, an approach via a multiple input multiple output state space equation is provided. Basic principle is introduced with a detailed proof. This approach requires the solution to a discrete algebraic Riccati equation (DARE). Like in section 3, the calculation procedure of this method is shown by a flow chart and a numerical example is given also.

In chapter 3, beginning from an overview of several models for spectral estimation, problem formulation is introduced to multichannel spectral estimation by using autoregressive (AR) model, then it is shown that this problem can result in a normal equation. In order to give a recursive solution of the normal equation, a Yule-Walker algorithm is derived, and all results is got together in a summary. After that a flow chart of the algorithm is shown for realizing the calculation procedure. Also a numerical example is shown.

In chapter 4, identification problem for multiple input multiple output system is discussed. From input data and output data, subspace method can provide coefficient matrices of state space expression of the system, of course then its multichannel transfer function can be obtained. At first, two basic theorems are described. Then based these theorems, the method to determine the coefficient matrices of state space equation are shown. Also in order the performance of this method, a numerical example is shown with comparing the output spectral of original system and estimated system.

In chapter 5, a spectral analysis of measured 6 channel data of car noise is given. The experiment condition and sensors position etc. are given in first section. Then spectral factorisation and spectral estimation discussed in previous chapter is applied to a measured 6 channel road noise data in car. It is confirmed here that the road noise can be whitening and uncorrelated by proposed method. Also from the simultaneous measured 2 channel input and 2 channel output data in car, the noise propagation space is identified by a state space system. These calculated results are shown in both numerical value and graphs.

In chapter 6, based on these results of whitening of multichannel input signals and system identification of control plant, an active noise cancelling system is constructed. For the assume of state space model of control plant, a adaptive control algorithm is derived based on steep-descent method, the performance improvement comparing with conventional method is discussed by a computer simulation.

In chapter 7, by evaluation of the contributions of our research, several conclusions are summarized and some comments on further research are given also.

Chapter 2 MULTICHANNEL SPECTRAL FACTORISATION

In this chapter, after given a brief background on multichannel spectral factorization, problem formulation for basic form and canonical form are defined. Then necessary and sufficient condition for general spectral density function, condition for canonical spectral factorisation and condition for canonical spectral factorisation of matrix polynomial are described. In section 2, some important past works, including those of D.C. Youla, M.C. Davis, E.A. Robinson and T. Kailath, are described briefly. He proposed a canonical spectral factorization by state space model. In section 3, an approach via a matrix binomial expression is provided, whose basic principle is introduced with a detailed proof. This method requires some complicated calculations, so two flow charts are given to show the calculation procedure. Also a numerical example is solved step by step. In section 4, an approach via a multiple input multiple output state space expression is provided. Basic principle is introduced with a detailed proof. This approach requires the solution to a discrete algebraic Riccati equation (DARE). The calculation procedure of this method is again shown by a flow chart and a numerical example is given also.

2.1 Background

Spectral factorization is a problem to factorise a given spectral function into a production of a factor function and its adjoint function. In continuous time domain, for a given spectral density function $\Phi(s)$, spectral factorization is a problem to find a stable minimum-phase function $G(s)$, such that $\Phi(s) = G(s)G(-s)$. When two or more channel such signals are to be operated on simultaneously, the spectral density function is given by a matrix function $\mathbf{\Phi}(s)$, whose *ij*th element is the cross-power density spectrum between the *i*th and *j*th input signal. In this case spectral factorization becomes a problem to find a stable minimum-phase matrix function such that $\mathbf{\Phi}(s) = \mathbf{G}(s)\mathbf{G}^T(-s)$. Also in discrete time domain, for scalar case, spectral factorization is problem to find a stable minimum-phase function $A(z)$ such that $S(z) = A(z)A(1/z)$. Similarly, with the multichannel problem, spectral factorization is a problem to find stable minimum-phase matrix function $\mathbf{A}(z)$ such that $\mathbf{S}(z) = \mathbf{A}(z)\mathbf{A}^T(1/z^*)$.

There are a variety of methods has been developed over the years for the computation of spectral factorisation. Because of significant progresses in digital computer and digital signal processing technology, spectral factorization problem in discrete time domain is used more commonly than in continuous time domain. For scalar-valued case, some methods named after Bauer, Levinson-Durbin, Schur, that is easy to calculate by digital computer have been reported. A survey of these methods is given by A. H. Sayed and T. Kailath in 2001[7)], which not only collects together the somewhat scattered result but also point out various relations and connections among them. Multichannel spectral factorisation is used to multichannel data in order to uncover the underlying components of measured signals, which is applied in several fields of engineering. For a general multichannel theory, the spectral factorisation problem is considerably more difficult. One reason is the condition $\mathbf{A}(z)$ of becomes more strict, particularly the condition on unit circle i.e., $z = e^{j\omega t}$. Another reason is due to computation complexity. There are several report on how to compute the multichannel spectral factorisation, including E. A. Robinson's method based on extracting zeros and T. Kailath's method based on discrete-time algebraic Riccati equation (DARE), each one has its own particularly good point.

2.2 Problem Formulation

2.2.1 Problem Formulation of Multichannel Spectral Factorization

For a p-channel stationary random signal vector

$$\mathbf{y}(n) = \begin{bmatrix} y_1(n) \\ y_2(n) \\ \vdots \\ y_p(n) \end{bmatrix} \tag{2.2.1}$$

We can give a $p \times p$ cross-correlation matrix

$$\mathbf{R}_i = E\left[\mathbf{y}(n+i)\mathbf{y}^{*T}(n)\right], (-\infty < i < \infty) \tag{2.2.2}$$

and of course,

$$\mathbf{R}_i = \mathbf{R}_{-i}^{*T}, (-\infty < i < \infty) \tag{2.2.3}$$

From these cross-correlation coefficient matrices, the spectral matrix is defined as

$$\mathbf{S}(z) = \sum_{i=-\infty}^{\infty} \mathbf{R}_i z^{-i} \tag{2.2.4}$$

For finite order of above equation, i.e. for the case of

$$\mathbf{R}_i = \mathbf{0}, (i > -m) \text{ or } (i < m) \tag{2.2.5}$$

(2.2.4) becomes following finite order equation, this is the form of power spectral matrix we will discuss here.

$$\mathbf{S}(z) = \sum_{i=-m}^{m} \mathbf{R}_i z^{-i} \tag{2.2.6}$$

Our problem is, for a given spectral matrix $\mathbf{S}(z)$, we wish to determine the p-channel minimum-phase operator $\mathbf{A}(z)$ such that

$$\mathbf{S}(z) = \mathbf{A}(z)\mathbf{A}_*(z) \tag{2.2.7}$$

where

$$\mathbf{A}(z) = \mathbf{A}_0 + \mathbf{A}_1 z^{-1} + \cdots \mathbf{A}_m z^{-m} \tag{2.2.8}$$

is a $p \times p$ matrix polynomial and $\mathbf{A}_0, \mathbf{A}_1, \ldots \mathbf{A}_m$ are each $p \times p$ constant matrices. And

$$\mathbf{A}_*(z) = \left[\mathbf{A}(1/z^*)\right]^{*T} = \mathbf{A}_0^{*T} + \mathbf{A}_1^{*T} z^1 + \cdots \mathbf{A}_m^{*T} z^m \tag{2.2.9}$$

is known as the para-Hermitian conjugate of $\mathbf{A}(z)$ and it is said as a adjoint system of $\mathbf{A}(z)$ also.

2.2.2 Problem Formulation of Multichannel Canonical Spectral Factorization

Notice that the solution of problem formulation in last section is not unique. In fact if a matrix polynomial $\mathbf{A}(z)$ is a solution of factorisation problem of (2.2.7), then for any arbitrary unitary matrix $\mathbf{D}$, which satisfy the condition of $\mathbf{DD}^{*T} = \mathbf{I}$, we can see that $\mathbf{A}(z)\mathbf{D}$ is a solution of (2.2.7) also. For this reason, we prefer to change the problem to a canonical form equivalently as following.

For a given spectral matrix expressed by a $p \times p$ matrix polynomial

$$\mathbf{S}(z) = \sum_{i=-m}^{m} \mathbf{R}_i z^{-i}$$

we wish to determine the p-channel minimum-delay operator $\mathbf{L}(z)$ such that

$$\mathbf{S}(z) = \mathbf{L}(z)\mathbf{\Sigma}\mathbf{L}_*(z) \tag{2.2.10}$$

where

$$\begin{aligned} \mathbf{L}(z) &= \mathbf{I}_p + \mathbf{L}_1 z^{-1} + \ldots\ldots + \mathbf{L}_m z^{-m} \\ &= \mathbf{I}_p + \sum_{i=1}^{m} \mathbf{L}_i z^{-i} \end{aligned} \tag{2.2.11}$$

is a $p \times p$ matrix polynomial, $\mathbf{I}_p$ is a $p \times p$ unity matrix and $\mathbf{L}_1, \ldots \mathbf{L}_m$ are each $p \times p$ constant matrices.

$$\begin{aligned} \mathbf{L}_*(z) &= [\mathbf{L}(1/z^*)]^{*T} \\ &= \mathbf{I}_p^{*T} + \mathbf{L}_1^{*T} z^1 + \ldots\ldots + \mathbf{L}_m^{*T} z^m \\ &= \mathbf{I}_p + \sum_{i=1}^{m} \mathbf{L}_i^* z^i \end{aligned} \tag{2.2.12}$$

is a para-Hermitian conjugate of $\mathbf{L}(z)$.

Every such problem admits a unique result with $\mathbf{\Sigma} > 0$ (where positive is meant in sense of matrices, i.e., that $\mathbf{\Sigma}$ is a positive-definite matrix) and $\mathbf{L}(z)$ a matrix polynomial of order m in z^{-1} that has all its roots strictly inside the unit circle, and such that $\mathbf{L}(\infty) = \mathbf{I}_p$. While we have obtained the solution $\{\mathbf{\Sigma}, \mathbf{L}(z)\}$, if $\mathbf{\Sigma}$ can be expressed in its decomposed form of*

$$\mathbf{\Sigma} = \mathbf{R}_0 \mathbf{R}_0^T \quad (2.2.13)$$

then we can connect $\mathbf{S}(z) = \mathbf{A}(z)\mathbf{A}_*(z)$ with $\mathbf{S}(z) = \mathbf{L}(z)\mathbf{\Sigma}\mathbf{L}_*(z)$ by

$$\begin{aligned} \mathbf{A}_0 &= \mathbf{R}_0 \\ \mathbf{A}_1 &= \mathbf{L}_1 \mathbf{R}_0 \\ &\cdots\cdots \\ \mathbf{A}_m &= \mathbf{L}_m \mathbf{R}_0 \end{aligned} \quad (2.2.14)$$

2.2.3 Condition for Spectral Factorization

Two basic properties

From the general definition of power spectral matrix given by (2.2.4), two important basic properties can be shown as following.

para-Hermitian symmetry:

$$\mathbf{S}(z) = [\mathbf{S}(1/z^*)]^{*T} = \mathbf{S}_*(z) \quad (2.2.15)$$

or Hermitian symmetry when $z = e^{j\omega}$:

$$\mathbf{S}(e^{j\omega}) = [\mathbf{S}(e^{j\omega})]^{*T} \quad (2.2.16)$$

Nonnegativity:

$$\mathbf{S}(e^{j\omega}) \geq 0 \quad (-\pi \leq \omega \leq \pi) \quad (2.2.17)$$

where nonegativity is meant in the sense of matrices, i.e., that $\mathbf{S}(e^{j\omega}) \geq 0$ is positive-semi-definite matrix.

* This can be done by Cholesky factorization easily. In MATLAB, the function of Cholesky factorization expresses a symmetric matrix as the product of a triangular matrix and its transpose $A = R^T R$, where R is an upper triangular matrix.

These two basic properties are necessary and sufficient condition for spectral density function, i.e. if power spectral matrix $\mathbf{S}(z)$ is given by (2.2.4), it will obey these two properties. Moreover, the converse is also true: a matrix function $\mathbf{S}(z)$ obeying these two properties must be a power spectral matrix given by (2.2.4), which means must be the z-transform of matrix-valued covariance sequence.

Condition for canonical spectral factorization

From the formulation of canonical spectral factorization problem in section 2.2.2, we can realize that $\mathbf{L}(z)$ must correspond to a minimum-phase system. This means $\mathbf{L}^{-1}(z)$ must analytic on the unit circle. This assumption rules out the possibility of having any zeros of $\mathbf{S}(z)$ on the unit circle. i.e., $\mathbf{S}(z)$ is assumed to be positive-definite on the unit circle as following.

$$\mathbf{S}(e^{j\omega}) > 0 \quad (-\pi \le \omega \le \pi) \tag{2.2.18}$$

which is equivalent to $\mathbf{S}(z)$ having full rank p everywhere on the unit circle, i.e.,

$$rank[\mathbf{S}(e^{j\omega})] = p \quad -\pi \le \omega \le \pi \tag{2.2.19}$$

When this holds, i.e., when $\mathbf{S}(z)$ is a rational matrix function and has maximal normal rank everywhere on the unit circle, one can show that it is always possible for perform the following canonical spectral factorization

$$\mathbf{S}(z) = \mathbf{L}(z)\mathbf{\Sigma}\mathbf{L}_*(z)$$

where

(i) $\mathbf{\Sigma}$ is a positive-definite matrix , $\mathbf{\Sigma} > 0$.

(ii) $\mathbf{L}(z)$ is a normalized to unity at infinity, $\mathbf{L}(\infty) = \mathbf{I}_p$

(iii) $\mathbf{L}(z)$ is analytic on and outside the unit circle $(|z| \ge 1)$.

(iv) $\mathbf{L}^{-1}(z)$ is analytic on and outside the unit circle $(|z| \ge 1)$.

We also can put (iii) and (iv) together as following.

(iii) $\mathbf{L}(z)$ is a rational minimum-phase function. That is, both $\mathbf{L}(z)$ and $\mathbf{L}^{-1}(z)$ are analytic on and outside the unit circle, or equivalently, $\mathbf{L}(z)$ has all its zeros and poles strictly inside the unit circle.

At last, we should mention that normalization $\mathbf{L}(\infty)=\mathbf{I}_p$ makes the factorization unique

Condition for canonical spectral factorization in matrix polynomial

Recall (2.2.6) that is the is the form of power spectral matrix we will discuss here.

$$\mathbf{S}(z)=\sum_{i=-m}^{m}\mathbf{R}_i z^{-i}$$

where

$$\mathbf{R}_i=\mathbf{R}_{-i}^{*}$$

such that it is nonnegative on the unit circle, $\mathbf{S}(e^{j\omega})\geq 0$ for $-\pi\leq\omega\leq\pi$. Then the following fact hold.

(i) If $z_0\neq 0$ is a zero of $\mathbf{S}(z)$ then z_0^{-*} is also a zero. It follows that if $\mathbf{S}(z)$ has m zeros $\{a_i\}$ on and inside the unit circle $(0<|z|\leq 1)$, then it also has m additional zeros $\{b_i=a_i^{-*}\}$ on and outside the unit circle $(1\leq|z|<\infty)$.

(ii) If $\mathbf{S}(z)$ is strictly positive on the unit circle then $\mathbf{L}(z)$ has all its zeros strictly inside the unit circle.

(iii) Assume the $\mathbf{R}_i$ are real-valued. Then the coefficient of $\mathbf{L}(z)$ are also real-valued. Moreover if z_0 is a complex root of $\mathbf{S}(z)$, then so are $\{z_0^{*},z_0^{-1},z_0^{-*}\}$.

2.3 Principal Approaches to Spectral Factorization

2.3.1 Approach by D.C. Youla[2)]

The fundamental theorem regarding the spectral factorization of rational matrices in continuous-time domain is contained in this paper. Using Youla's notation, the theorem is stated as following.

Let $\mathbf{G}(s) = \mathbf{G}_*(s)$ be a rational $n \times n$ para-conjugate Hermetian matrix of normal rank r which is non-negative on the real-frequency axis $s = j\omega$. Then, there exist an $r \times n$ rational matrix $\mathbf{H}(s)$ such that

(i) $\mathbf{G}(s) = \mathbf{H}(s)\mathbf{H}_*(s)$

(ii) $\mathbf{H}(s)$ and $\mathbf{H}^{-1}(s)$ are both analytic in $\mathrm{Re}\, s > 0$

(iii) $\mathbf{H}(s)$ is unique up to within a constant, unitary $r \times r$ matrix multiplier on the left; i.e., if $\mathbf{H}_1(s)$ also satisfies 1) and 2), $\mathbf{H}_1(s) = \mathbf{T}\mathbf{H}(s)$ where $\mathbf{T}$ is $r \times r$, constant and satisfies $\mathbf{T}^*\mathbf{T} = \mathbf{I}_r$.

(iv) Any factorization of the form $\mathbf{G}(s) = \mathbf{L}(s)\mathbf{L}_*(s)$ in which $\mathbf{L}(s)$ is $r \times n$, rational and analytic in $\mathrm{Re}\, s > 0$, is given by $\mathbf{L}(s) = \mathbf{V}(s)\mathbf{H}(s)$, $\mathbf{V}(s)$ being an arbitrary rational, regular $r \times r$ para-conjugate unitary matrix.

(v) If $\mathbf{G}(s)$ is analytic on the finite $s = j\omega$ axis, $\mathbf{H}(s)$ is analytic in the right semi-infinite strip $\mathrm{Re}\, s > -\tau$, $\tau > 0$.

(vi) If $\mathbf{G}(s)$ is analytic and $rank\mathbf{G}(s)$ is invariant on the finite $s = j\omega$ axis, $\mathbf{H}^{-1}(s)$ is analytic in the right semi-infinite strip $\mathrm{Re}\, s > -\tau_1$, $\tau_1 > 0$.

(vii) If $\mathbf{G}(s)$ is real, $\mathbf{H}(s)$ and $\mathbf{V}(s)$ are real and $\mathbf{T}$ is real orthogonal.

In this paper, several examples are given, but its computing step is not sorted out clearly and also have some difficulty to be carried out.

2.3.2 Approach by M. C. Davis[3)]

This paper presents a general factoring procedure for rational matrices in continuous time domain. By supposing $\mathbf{G}^{-1}(s)$ as a cascaded series as following,

$$\mathbf{G}^{-1}(s) = \mathbf{T}_n(s)\cdots\mathbf{T}_2(s)\mathbf{T}_1(s) \tag{2.3.1}$$

where $\mathbf{T}_i(s)$ is obtained by either of following two simple step.

(i) $\mathbf{T}_i(s)$ is equal to the identity matrix except for one or more *j*th diagonal elements $t_{jj}(s)$.

(ii) $\mathbf{T}_i(s)$ is equal to the identity matrix except for the off-diagonal elements of the *n*th row, $t_{nj}(s)$.

We can use the following update equation repeatedly.

$$\mathbf{\Phi}_{i+1}(s) = \mathbf{T}_i(-s)\mathbf{\Phi}_i(s)\mathbf{T}_i^T(s) \tag{2.3.2}$$

At last the spectrum matrix $\mathbf{\Phi}(s)$ of the input signals can be factored such that $\mathbf{\Phi}(s) = \mathbf{G}(-s)\mathbf{G}^T(s)$, where $\mathbf{G}(s)$ is a stable minimum-phase transfer function, then $\mathbf{G}(s)$ can be viewed as the system which would reproduce signals with the spectrum $\mathbf{\Phi}(s)$ when excited by n uncorrelated unit-density white noise sources.

2.3.3 Approach by E. A. Robinson[4)]

The section 5.4 of this book is titled in "Minimum-delay factorization of multichannel spectrum". Here a description of Robinson's approach is given using his notation. From a discrete time spectral density of matrix polynomial given as following,

$$\mathbf{\Psi}(z) = \sum_{i=-m}^{m} \mathbf{R}_i z^i \tag{2.3.2}$$

a method to determine the p-channel minimum-delay operator $\mathbf{A}(z)$ such that

$$\mathbf{\Psi}(z) = \mathbf{A}(z)\mathbf{A}_*(z) \tag{2.3.4}$$

is described, where

$$\mathbf{A}(z) = \mathbf{A}_0 + \mathbf{A}_1 z + \ldots\ldots + \mathbf{A}_m z^m \tag{2.3.5}$$

$$\mathbf{A}_*(z) = [\mathbf{A}(1/z^*)]^{*T} = \mathbf{A}_0^{*T} + \mathbf{A}_1^{*T} z^{-1} + \ldots\ldots + \mathbf{A}_m^{*T} z^{-m} \tag{2.3.6}$$

The method is based on via suppose that $\mathbf{A}(z)$ can be written into a binomial form

$$\mathbf{A}(z) = (\mathbf{I} - \mathbf{U}_1 z)(\mathbf{I} - \mathbf{U}_2 z)\cdots(\mathbf{I} - \mathbf{U}_m z)\mathbf{A}_0 \tag{2.3.7}$$

then determine matrix $\mathbf{U}_1, \mathbf{U}_2 \cdots \mathbf{U}_m$ and $\mathbf{A}_0$. At last determine matrix $\mathbf{A}_1, \mathbf{A}_2 \cdots \mathbf{A}_m$.

The method given in this section can be sort out as following.

(i) find all roots of det $\mathbf{\Psi}(z) = 0$.

(ii) choose all roots outside unity circle, and divide these roots into m group.

(iii) main loop: using each group roots to determine $\mathbf{U}_i$ matrix in binomial form

$(\mathbf{I} - \mathbf{U}_i z), (i = 1,2,...m)$.

(iv) determine the coefficient matrix $\mathbf{A}_0$ by Cholesky factorisation.

(v) transform binomial form into polynomial form to get

$$\mathbf{A}(z) = \mathbf{A}_0 + \mathbf{A}_1 z + \ldots\ldots + \mathbf{A}_m z^m$$

2.3.4 Approach by T. Kailath[5)]

The section 8.3 of this book is titled in "canonical spectral factorization". In this section, from a discrete time state space model given as following,

$$\begin{array}{l}\mathbf{x}_{i+1}=\mathbf{F}\mathbf{x}_i+\mathbf{G}\mathbf{u}_i, \\ \mathbf{y}_i=\mathbf{H}\mathbf{x}_i+\mathbf{v}_i\end{array}, \quad i\geq 0 \tag{2.3.8}$$

where $\mathbf{F}$, $\mathbf{G}$ and $\mathbf{H}$ are known constant matrices. $\mathbf{u}_i$ and $\mathbf{v}_i$ are zero-mean jointly stationary vector random variables that, along with the zero-mean random variable $\mathbf{x}_i$, satisfy the conditions

$$\left\langle \begin{bmatrix}\mathbf{x}_0\\ \mathbf{u}_i\\ \mathbf{v}_i\end{bmatrix}, \begin{bmatrix}\mathbf{x}_0\\ \mathbf{u}_j\\ \mathbf{v}_j\\ 1\end{bmatrix}\right\rangle = \begin{bmatrix}\mathbf{\Pi}_0 & 0 & 0\\ 0 & \begin{bmatrix}\mathbf{Q} & \mathbf{S}\\ \mathbf{S}^* & \mathbf{R}\end{bmatrix}\delta_{ij} & 0\end{bmatrix} \tag{2.3.9}$$

They express a discrete time spectral $\mathbf{S}_y(z)$ into following form.

$$\mathbf{S}_y(z)=\begin{bmatrix}\mathbf{H}(z\mathbf{I}-\mathbf{F})^{-1} & I\end{bmatrix}\begin{bmatrix}\mathbf{GQG}^* & \mathbf{GS}\\ \mathbf{SG}^* & \mathbf{R}\end{bmatrix}\begin{bmatrix}(z^{-1}\mathbf{I}-\mathbf{F}^*)^{-1}\mathbf{H}^*\\ \mathbf{I}\end{bmatrix} \tag{2.3.10}$$

Regarding the spectral matrix $\mathbf{S}_y(z)$ in above form, they show that its canonical spectral factorization can be obtained as

$$\mathbf{S}(z)=\mathbf{L}(z)\mathbf{R}_e\mathbf{L}^*(z^{-*}),\ \mathbf{L}(\infty)=\mathbf{I},\ \mathbf{R}_e>0 \tag{2.3.11}$$

where

$$\mathbf{L}(z)=\mathbf{I}+\mathbf{H}(z\mathbf{I}-\mathbf{F})^{-1}\mathbf{K}_p \tag{2.3.12}$$

$$\mathbf{L}^{-1}(z)=\mathbf{I}-\mathbf{H}(z\mathbf{I}-\mathbf{F}+\mathbf{K}_p\mathbf{H})^{-1}\mathbf{K}_p \tag{2.3.13}$$

$$\mathbf{K}_p=(\mathbf{FPH}+\mathbf{GS})\mathbf{R}_e^{-1} \tag{2.3.14}$$

$$\mathbf{R}_e=\mathbf{R}+\mathbf{HPH}^* \tag{2.3.15}$$

and $\mathbf{P}$ is the unique positive-semi-define solution to the discrete-time algebraic Riccati equation (DARE)

$$\mathbf{P} = \mathbf{F}\mathbf{P}\mathbf{F}^{*} + \mathbf{G}\mathbf{Q}\mathbf{G}^{*} - \mathbf{K}_{p}\mathbf{R}_{e}\mathbf{K}_{p} \tag{2.3.16}$$

Moreover, $\mathbf{F} - \mathbf{K}_{p}\mathbf{H}$ is stable, which, in addition to the stability of $\mathbf{F}$, will guarantee that $\mathbf{L}(z)$ is minimum-phase.

2.4 Solution via Binomial Approach

2.4.1 Principle (after Robinson's Approach)

For a given p-channel spectral density in matrix polynomial form as following.

$$\mathbf{S}(z)=\sum_{i=-m}^{m}\mathbf{R}_i z^{-i}$$

What we wish to do is to determine the p-channel minimum-delay operator $\mathbf{A}(z)$ such that

$$\mathbf{S}(z)=\mathbf{A}(z)\mathbf{A}_*(z) \tag{2.4.1}$$

where $\mathbf{A}(z)$ is a m-order $p\times p$ matrix polynomial

$$\mathbf{A}(z)=\mathbf{A}_0+\mathbf{A}_1 z^{-1}+......+\mathbf{A}_m z^{-m} \tag{2.4.2}$$

and $\mathbf{A}_*(z)$ is a adjoint polynomial of $\mathbf{A}(z)$, which is given by

$$\mathbf{A}_*(z)=[\mathbf{A}(1/z^*)]^{*T}=\mathbf{A}_0^{*T}+\mathbf{A}_1^{*T}z+......+\mathbf{A}_m^{*T}z^m \tag{2.4.3}$$

Suppose we can write $\mathbf{A}(z)$ into its binomial factors form as following.

$$\mathbf{A}(z)=(\mathbf{I}-\mathbf{U}_1 z^{-1})(\mathbf{I}-\mathbf{U}_2 z^{-1})\cdots(\mathbf{I}-\mathbf{U}_m z^{-1})\mathbf{A}_0 \tag{2.4.4}$$

where $\mathbf{U}_1,\mathbf{U}_2,\cdots\mathbf{U}_m$ are $p\times p$ constant matrices. Then also we can write $\mathbf{A}_*(z)$ as following.

$$\mathbf{A}_*(z)=\mathbf{A}_0^{*T}(\mathbf{I}-\mathbf{U}_m^{*T}z)\cdots(\mathbf{I}-\mathbf{U}_2^{*T}z)(\mathbf{I}-\mathbf{U}_1^{*T}z) \tag{2.4.5}$$

Hence the p-channel spectral matrix $\mathbf{S}(z)$ is given as

$$\mathbf{S}(z) = (\mathbf{I} - \mathbf{U}_1 z^{-1}) \cdots (\mathbf{I} - \mathbf{U}_m z^{-1}) \mathbf{A}_0 \mathbf{A}_0^{*T} (\mathbf{I} - \mathbf{U}_m^{*T} z) \cdots (\mathbf{I} - \mathbf{U}_1^{*T} z) \quad (2.4.6)$$

Next left-multiply (2.4.5) by $\det \mathbf{S}(z)\mathbf{S}^{-1}(z) = \mathrm{adj}\mathbf{S}(z)$, this yields

$$\begin{aligned} &\det \mathbf{S}(z)\mathbf{S}^{-1}(z)\mathbf{S}(z) = \det \mathbf{S}(z)\mathbf{I} \\ &= \mathrm{adj}\mathbf{S}(z)(\mathbf{I} - \mathbf{U}_1 z^{-1}) \cdots (\mathbf{I} - \mathbf{U}_m z^{-1}) \mathbf{A}_0 \mathbf{A}_0^{*T} (\mathbf{I} - \mathbf{U}_m^{*T} z) \cdots (\mathbf{I} - \mathbf{U}_1^{*T} z) \end{aligned} \quad (2.4.7)$$

Now the number of zeros of the polynomial is 2mp. (p is number of channels and m is the order of $\mathbf{A}(z)$). Moreover, because

$$\det \mathbf{S}(z) = \det \mathbf{A}(z) \det \mathbf{A}_*(z) \quad (2.4.8)$$

we see that if z_i is a zero of $\det \mathbf{S}(z)$, then $1/z_i^*$ is also a zero. One of this pair of zeros goes into the construction of $\mathbf{A}(z)$ and the other into the construction of $\mathbf{A}_*(z)$. and the choice is arbitrary for each of the mp pairs of zeros of $\det \mathbf{S}(z)$. As a result, there will 2^{mp} multichannel factorisations if we exclude the case of zeros on unit circle and other multiple zeros. In order to obtain the minimum-delay factorisation we choose these zeros which are smaller than unity in magnitude. Divide these mp zeros in m group, each group has p zeros. For the first group of zero $z_i, (i = 1,2,\ldots p)$ the adjoint factors into the product of a column $\mathbf{c}_i$ and a row $\mathbf{r}_i$, that is $\mathrm{adj}\mathbf{S}(z_i) = \mathbf{c}_i \mathbf{r}_i$ (cf. Section 2.4.4) which gives

$$\mathbf{0} = \mathbf{c}_i \mathbf{r}_i (\mathbf{I} - \mathbf{U}_1 z_i^{-1}) \cdots (\mathbf{I} - \mathbf{U}_m z_i^{-1}) \mathbf{A}_0 \mathbf{A}_0^{*T} (\mathbf{I} - \mathbf{U}_m^{*T} z_i) \cdots (\mathbf{I} - \mathbf{U}_1^{*T} z_i) \quad (2.4.9)$$

Thus we have $\mathbf{r}_i(\mathbf{I} - \mathbf{U}_1 z_i^{-1}) = \mathbf{0}$ i.e., $\mathbf{r}_i(z_i \mathbf{I} - \mathbf{U}_1) = \mathbf{0}$. Note that this characteristic equation of $\mathbf{U}_1$ and using this equation for all p zeros $z_i, (i = 1,2,\ldots p)$ in first group, we may obtain matrix $\mathbf{U}_1$ as a similarity transformation of its eigenvalue matrix $\mathbf{D}_1 = diag(z_1, z_2 \cdots z_p)$. (cf. Section 2.4.4).

Then for the second group, at the next zero z_i, $(i = p+1, p+2, \ldots 2p)$, because we also can write

$$\operatorname{adj}\mathbf{S}(z_i) = \mathbf{c}_i\mathbf{r}_i, \ (i = p+1, \ldots 2p) \tag{2.4.10}$$

where $\mathbf{c}_i$ is column vector and $\mathbf{r}_i$ is row vector, then

$$\mathbf{0} = \mathbf{c}_i\mathbf{r}_i(\mathbf{I} - \mathbf{U}_1 z_i^{-1})\cdots(\mathbf{I} - \mathbf{U}_m z_i^{-1})\mathbf{A}_0\mathbf{A}_0^{*T}(\mathbf{I} - \mathbf{U}_m^{*T} z_i)\cdots(\mathbf{I} - \mathbf{U}_1^{*T} z_i) \tag{2.4.11}$$

because $\mathbf{U}_1$ is known already, so the factor $\mathbf{c}_i\mathbf{r}_i(\mathbf{I} - \mathbf{U}_1 z_i)$ also becomes a product of a column vector $\mathbf{c}_i'$ and a row vector $\mathbf{r}_i'$, i.e., $\mathbf{c}_i\mathbf{r}_i(\mathbf{I} - \mathbf{U}_1 z_i) = \mathbf{c}_i'\mathbf{r}_i'$, thus we have $\mathbf{r}_i'(\mathbf{I} - \mathbf{U}_2 z_i^{-1}) = \mathbf{0}$ i.e., $\mathbf{r}_i(z_i\mathbf{I} - \mathbf{U}_2) = \mathbf{0}$. Note that this characteristic equation of $\mathbf{U}_2$ and using this equation for all p zeros $z_i, (i = p+1, p+2, \ldots 2p)$ in second group, hence we may obtain matrix $\mathbf{U}_2$ as a similarity transformation of its eigenvalue matrix $\mathbf{D}_2 = diag(z_{p+1}, z_{p+2} \cdots z_{2p})$. (cf. Section 2.4.4).

Likewise, we can then consecutively determine $\mathbf{U}_1$, $\mathbf{U}_2, \ldots \mathbf{U}_m$, thereby exhausting the mp zeros which have magnitude greater than one. Finally by let $z = 1$ * in following equation

$$\mathbf{S}(z) = (\mathbf{I} - \mathbf{U}_1 z^{-1})\cdots(\mathbf{I} - \mathbf{U}_m z^{-1})\mathbf{A}_0\mathbf{A}_0^{*T}(\mathbf{I} - \mathbf{U}_m^{*T} z)\cdots(\mathbf{I} - \mathbf{U}_1^{*T} z) \tag{2.4.12}$$

so we can get that

$$\mathbf{A}_0\mathbf{A}_0^{*T} = ((\mathbf{I} - \mathbf{U}_1)\cdots(\mathbf{I} - \mathbf{U}_m))^{-1}\mathbf{S}(1)((\mathbf{I} - \mathbf{U}_m^{*T} z)\cdots(\mathbf{I} - \mathbf{U}_1^{*T} z))^{-1} \tag{2.4.13}$$

* if $s(1) = 0$, we must select another z_0 for $s(z_0) \neq 0$

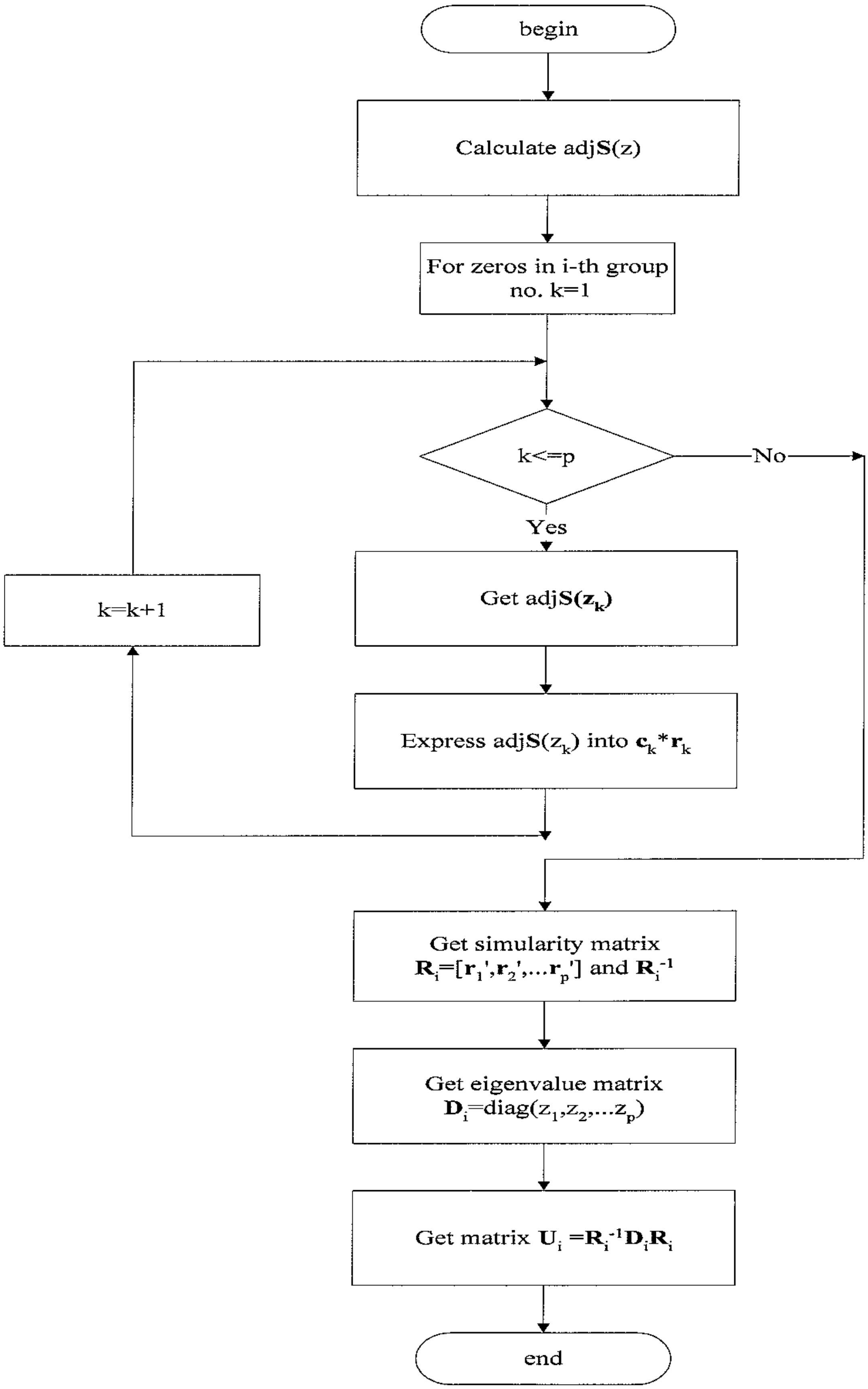
then by Cholesky factorisation we can get $\mathbf{A}_0$ and hence we obtain the required casual-chain minimum-delay transfer function in binomial form

$$\mathbf{A}(z) = (\mathbf{I} - \mathbf{U}_1 z^{-1})(\mathbf{I} - \mathbf{U}_2 z^{-1}) \cdots (\mathbf{I} - \mathbf{U}_m z^{-1})\mathbf{A}_0 \tag{2.4.14}$$

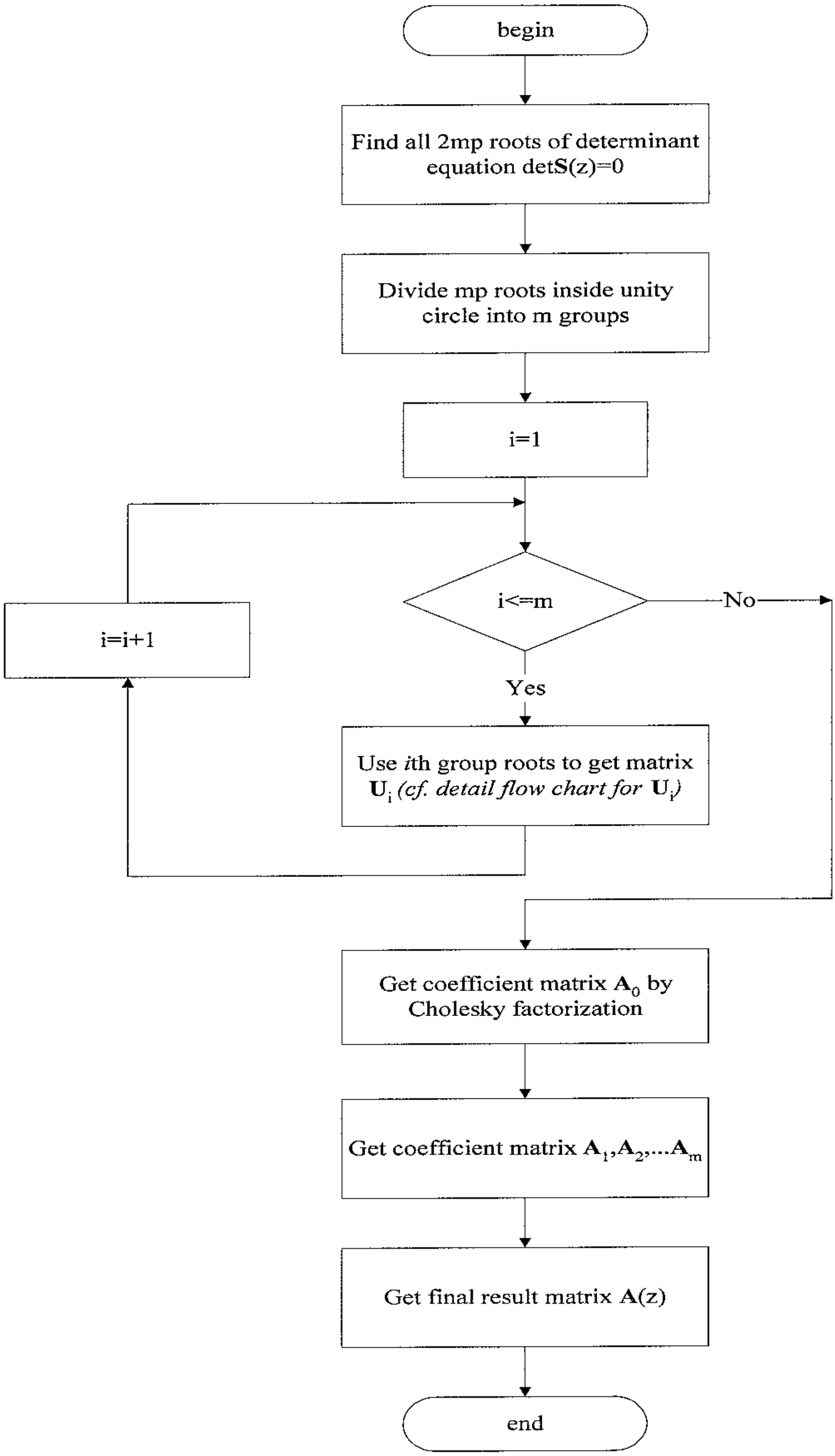
At last transform this expression into polynomial form, so that we can get the spectral factorization $\mathbf{A}(z)$ expected in (2.4.2).

2.4.2 Flow Chart

In order to describe the calculation procedures more precisely, two flow chart are shown in Figure 3 and Figure 4.



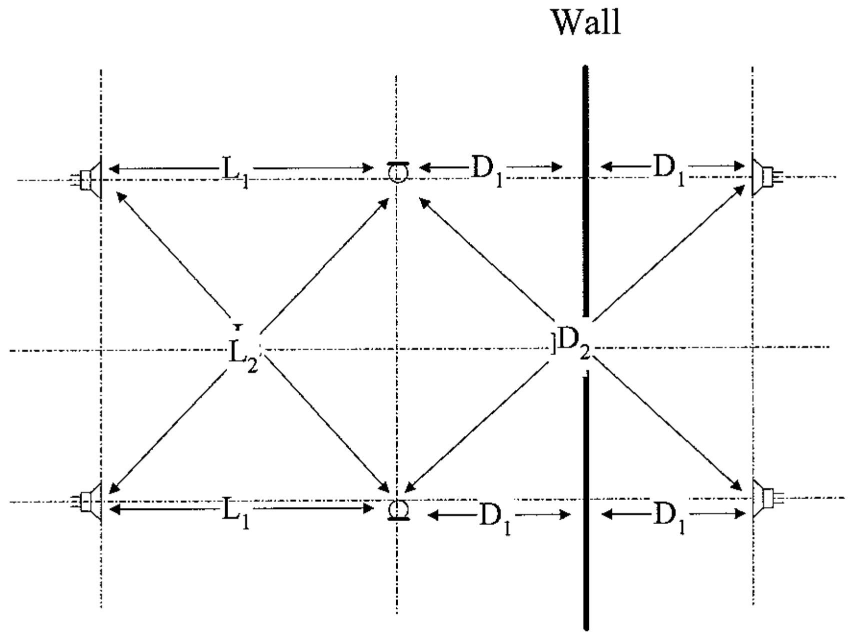
• Figure 3 Detail flow chart for calculating matrix $\mathbf{U}_i$



• Figure 4 Flow chart of presented spectral factorization method

2.4.3 Practical Example

Physical arrangement of acoustic system of two speakers and two microphones with wall reflection is shown in Figure 5.



• Figure 5 Physical arrangement of propagation between 2 loudspeakers and 2 microphones

From Figure 5, we can show the input of microphones as following.

$$\begin{aligned} x_1(k) &= c_1 n_1(k-N_1) + d_1 n_1(k-M_1) + c_2 n_2(k-N_2) + d_2 n_2(k-M_2) \\ x_2(k) &= c_2 n_1(k-N_2) + d_2 n_1(k-M_2) + c_1 n_2(k-N_1) + d_1 n_2(k-M_1) \end{aligned} \quad (2.4.15)$$

Applying z transform to above equations

$$\begin{aligned} X_1(z) &= \{c_1 z^{-N_1} + d_1 z^{-M_1}\} N_1(z) + \{c_2 z^{-N_2} + d_2 z^{-M_2}\} N_2(z) \\ X_2(z) &= \{c_2 z^{-N_2} + d_2 z^{-M_2}\} N_1(z) + \{c_1 z^{-N_1} + d_1 z^{-M_1}\} N_2(z) \end{aligned} \quad (2.4.16)$$

$$\begin{bmatrix} X_1(z) \\ X_2(z) \end{bmatrix} = \begin{bmatrix} c_1 z^{-N_1} + d_1 z^{-M_1} & c_2 z^{-N_2} + d_2 z^{-M_2} \\ c_2 z^{-N_2} + d_2 z^{-M_2} & c_1 z^{-N_1} + d_1 z^{-M_1} \end{bmatrix} \begin{bmatrix} N_1(z) \\ N_2(z) \end{bmatrix} \quad (2.4.17)$$

So that the transfer function matrix propagation between 2 loudspeakers and 2 microphones is given as following.

$$\mathbf{A}(z) = \begin{bmatrix} c_1 z^{-N_1} + d_1 z^{-M_1} & c_2 z^{-N_2} + d_2 z^{-M_2} \\ c_2 z^{-N_2} + d_2 z^{-M_2} & c_1 z^{-N_1} + d_1 z^{-M_1} \end{bmatrix}$$
$$= \begin{bmatrix} c_1 z^{-N_1} & c_2 z^{-N_2} \\ c_2 z^{-N_2} & c_1 z^{-N_1} \end{bmatrix} + \begin{bmatrix} d_1 z^{-M_1} & d_2 z^{-M_2} \\ d_2 z^{-M_2} & d_1 z^{-M_1} \end{bmatrix} \quad (2.4.18)$$

if $\Delta_1 = N_2 - N_1, \Delta_2 = M_1 - N_1, \Delta_3 = M_2 - N_1$, then

$$\mathbf{A}(z) = \mathbf{A}_{\min}(z)\mathbf{A}_{all}(z)$$
$$= \left(\begin{bmatrix} c_1 & c_2 z^{-\Delta_1} \\ c_2 z^{-\Delta_1} & c_1 \end{bmatrix} + \begin{bmatrix} d_1 z^{-\Delta_2} & d_2 z^{-\Delta_3} \\ d_2 z^{-\Delta_3} & d_1 z^{-\Delta_2} \end{bmatrix} \right) \begin{bmatrix} z^{-N_1} & 0 \\ 0 & z^{-N_1} \end{bmatrix} \quad (2.4.19)$$

where

$$\mathbf{A}_{\min}(z) = \begin{bmatrix} c_1 & c_2 z^{-\Delta_1} \\ c_2 z^{-\Delta_1} & c_1 \end{bmatrix} + \begin{bmatrix} d_1 z^{-\Delta_2} & d_2 z^{-\Delta_3} \\ d_2 z^{-\Delta_3} & d_1 z^{-\Delta_2} \end{bmatrix} \quad (2.4.20)$$

$$\mathbf{A}_{all}(z) = \begin{bmatrix} z^{-N_1} & 0 \\ 0 & z^{-N_1} \end{bmatrix} \quad (2.4.21)$$

$$\mathbf{A}_{\min}(z) = \begin{bmatrix} c_1 & 0 \\ 0 & c_1 \end{bmatrix} + \begin{bmatrix} 0 & c_2 \\ c_2 & 0 \end{bmatrix} z^{-\Delta_1} + \begin{bmatrix} d_1 & 0 \\ 0 & d_1 \end{bmatrix} z^{-\Delta_2} + \begin{bmatrix} 0 & d_2 \\ d_2 & 0 \end{bmatrix} z^{-\Delta_3} \quad (2.4.22)$$

Here we suppose that $\Delta_1 = 1, \Delta_2 = 1, \Delta_3 = 2$, so that the transfer function $\mathbf{A}(z)$ becomes

$$\mathbf{A}_{\min}(z) = \begin{bmatrix} c_1 & 0 \\ 0 & c_1 \end{bmatrix} + \begin{bmatrix} d_1 & c_2 \\ c_2 & d_1 \end{bmatrix} z^{-1} + \begin{bmatrix} 0 & d_2 \\ d_2 & 0 \end{bmatrix} z^{-2}$$
$$= \begin{bmatrix} c_1 + d_1 z^{-1} & c_2 z^{-1} + d_2 z^{-2} \\ c_2 z^{-1} + d_2 z^{-2} & c_1 + d_1 z^{-1} \end{bmatrix} = \begin{bmatrix} a_1(z) & a_2(z) \\ a_2(z) & a_1(z) \end{bmatrix} \quad (2.4.23)$$

where

$$a_1(z) = c_1 + d_1 z^{-1}$$
$$a_2(z) = c_2 z^{-1} + d_2 z^{-2} \quad (2.4.24)$$

And also the para-Hermitian conjugate of $\mathbf{A}(z)$ becomes

$$\mathbf{A}_{\min*}(z) = \left[\mathbf{A}_{\min*}(1/z^*)\right]^{*T} = \begin{bmatrix} c_1 & 0 \\ 0 & c_1 \end{bmatrix} + \begin{bmatrix} d_1 & c_2 \\ c_2 & d_1 \end{bmatrix} z^1 + \begin{bmatrix} 0 & d_2 \\ d_2 & 0 \end{bmatrix} z^2$$
$$= \begin{bmatrix} c_1 + d_1 z^1 & c_2 z^1 + d_2 z^2 \\ c_2 z^1 + d_2 z^2 & c_1 + d_1 z^1 \end{bmatrix} = \begin{bmatrix} a_1(z^{-1}) & a_2(z^{-1}) \\ a_2(z^{-1}) & a_1(z^{-1}) \end{bmatrix} \tag{2.4.25}$$

So that we can get a power spectrum function as following.

$$\mathbf{S}(z) = \mathbf{A}(z)\mathbf{A}_*(z) = \mathbf{A}_{\min}(z)\mathbf{A}_{all}(z)\mathbf{A}_{all*}(z)\mathbf{A}_{\min*}(z) \tag{2.4.26}$$

Notice that

$$\mathbf{A}_{all}(z)\mathbf{A}_{*all}(z) = \begin{bmatrix} z^{-N_1} & 0 \\ 0 & z^{-N_1} \end{bmatrix} \begin{bmatrix} z^{N_1} & 0 \\ 0 & z^{N_1} \end{bmatrix} = \mathbf{I} \tag{2.4.27}$$

then

$$\mathbf{S}(z) = \mathbf{A}_{\min}(z)\mathbf{A}_{\min*}(z)$$
$$= \begin{bmatrix} a_1(z) & a_2(z) \\ a_2(z) & a_1(z) \end{bmatrix} \begin{bmatrix} a_1(z^{-1}) & a_2(z^{-1}) \\ a_2(z^{-1}) & a_1(z^{-1}) \end{bmatrix} \tag{2.4.28}$$
$$= \begin{bmatrix} a_1(z)a_1(z^{-1}) + a_2(z)a_2(z^{-1}) & a_1(z)a_2(z^{-1}) + a_2(z)a_1(z^{-1}) \\ a_2(z)a_1(z^{-1}) + a_1(z)a_2(z^{-1}) & a_2(z)a_2(z^{-1}) + a_1(z)a_1(z^{-1}) \end{bmatrix}$$

then put $a_1(z)$ and $a_2(z)$ into this equation, we can get

$$\mathbf{S}(z) = \mathbf{A}_{\min}(z)\mathbf{A}_{\min*}(z)$$
$$= \begin{bmatrix} c_1 + d_1 z^{-1} & c_2 z^{-1} + d_2 z^{-2} \\ c_2 z^{-1} + d_2 z^{-2} & c_1 + d_1 z^{-1} \end{bmatrix} \begin{bmatrix} c_1 + d_1 z^1 & c_2 z^1 + d_2 z^2 \\ c_2 z^1 + d_2 z^2 & c_1 + d_1 z^1 \end{bmatrix}$$
$$= \begin{bmatrix} (c_1 d_1 + c_2 d_2) z^{-1} + (c_1 c_1 + d_1 d_1 & c_1 d_2 z^{-2} + (c_1 c_2 + d_1 d_2) z^{-1} + (c_2 d_1 \\ + c_2 c_2 + d_2 d_2) + (c_1 d_1 + c_2 d_2) z^1 & + c_2 d_1) + (c_1 c_2 + d_1 d_2) z^1 + c_1 d_2 z^2 \\ c_1 d_2 z^{-2} + (c_1 c_2 + d_1 d_2) z^{-1} + (c_2 d_1 & (c_1 d_1 + c_2 d_2) z^{-1} + (c_1 c_1 + d_1 d_1 \\ + c_2 d_1) + (c_1 c_2 + d_1 d_2) z^1 + c_1 d_2 z^2 & + c_2 c_2 + d_2 d_2) + (c_1 d_1 + c_2 d_2) z^1 \end{bmatrix}$$

(2.4.29)

Then we will try to factor this power spectrum function by the method last time.

step1 find all roots of $\det \mathbf{S}(z)=0$
When parameter c1=0.8,c2=0.6,d1=0.6,d2=0.4, then

$$\begin{aligned}&\det \mathbf{S}(z)=-0.1024z^{-4}-0.4608z^{-3}-0.4608z^{-2}+0.6912z^{-1}\\&+1.5872+0.6972z^{1}-0.4608z^{2}-0.4608z^{3}-0.1024z^{4}\\&=0\end{aligned}\tag{2.4.30}$$

By using a MATLAB function, we can find all roots of this determinate equation.

root =1.4142, -2.0000, 0.7071, -1.4142, -1.0000, -1.0000, -0.7071, -0.5000

There are 3roots outside the unit circle, 3 roots inside the unit circle and also two roots on the unit circle. We will choose all of these roots inside the unit circle and half of these roots on the unit circle. Then divide these roots into two groups and denote these roots as following.

Group 1: z1=-0.5000;

z2=-0.7071;

Group 2: z3=-1.0000;

z4=0.7071;

step2-1 use z1, z2 get r1, r2 vector

At first, calculate adjoint matrix of power spectral density $\mathbf{S}(z)$

$$\mathrm{adj}\mathbf{S}(z)=$$
$$\begin{bmatrix}(c_1d_1+c_2d_2)z^{-1}+(c_1c_1+d_1d_1 & -c_1d_2z^{-2}-(c_1c_2+d_1d_2)z^{-1}-(c_2d_1\\ +c_2c_2+d_2d_2)+(c_1d_1+c_2d_2)z^{1} & +c_2d_1)-(c_1c_2+d_1d_2)z^{1}-c_1d_2z^{2}\\ -c_1d_2z^{-2}-(c_1c_2+d_1d_2)z^{-1}-(c_2d_1 & (c_1d_1+c_2d_2)z^{-1}+(c_1c_1+d_1d_1\\ +c_2d_1)-(c_1c_2+d_1d_2)z^{1}-c_1d_2z^{2} & +c_2c_2+d_2d_2)+(c_1d_1+c_2d_2)z^{1}\end{bmatrix}\tag{2.4.31}$$

For zeros of z1=-0.5000, z2=-0.7071 in group 1, we can get

$$\mathrm{adj}\mathbf{S}(z_1)=\begin{bmatrix}-0.28 & -0.28\\ -0.28 & -0.28\end{bmatrix}=\begin{bmatrix}-0.28\\ -0.28\end{bmatrix}[1 \quad 1]\tag{2.4.32}$$

$$\text{adj}\mathbf{S}(z_2)=\begin{bmatrix}-0.007356 & 0.007356\\ 0.007356 & -0.007356\end{bmatrix}$$
$$=\begin{bmatrix}-0.007356\\ 0.007356\end{bmatrix}\begin{bmatrix}1 & -1\end{bmatrix} \quad (2.4.33)$$

$$\mathbf{r}_1=\begin{bmatrix}1 & 1\end{bmatrix} \quad (2.4.34)$$

$$\mathbf{r}_2=\begin{bmatrix}1 & -1\end{bmatrix} \quad (2.4.35)$$

step2-2 use r1,r2 vector to get U_1 matrix

We then assemble the matrix

$$\mathbf{R}=\begin{bmatrix}\mathbf{r}_1\\ \mathbf{r}_2\end{bmatrix}=\begin{bmatrix}1 & 1\\ 1 & -1\end{bmatrix} \quad (2.4.36)$$

$$\mathbf{R}^{-1}=\frac{1}{2}\begin{bmatrix}1 & 1\\ 1 & -1\end{bmatrix} \quad (2.4.37)$$

and the diagonal matrix

$$\mathbf{D}=\begin{bmatrix}z_1 & 0\\ 0 & z_2\end{bmatrix} \quad (2.4.38)$$

from which we can obtain the $\mathbf{U}_1$ matrix.

$$\mathbf{U}_1=\mathbf{R}^{-1}\mathbf{D}\mathbf{R}=\frac{1}{2}\begin{bmatrix}1 & 1\\ 1 & -1\end{bmatrix}\begin{bmatrix}z_1 & 0\\ 0 & z_2\end{bmatrix}\begin{bmatrix}1 & 1\\ 1 & -1\end{bmatrix}$$
$$=\begin{bmatrix}-0.60355 & 0.10355\\ 0.10355 & -0.60355\end{bmatrix} \quad (2.4.39)$$

step3-1 use z3, z4 get r3, r4 vector

For zeros of z3=-1.0000, z4=0.7071 in group 2, we can get

$$\text{adj}\mathbf{S}(z_3)=\begin{bmatrix}0.08 & 0.08\\ 0.08 & 0.08\end{bmatrix}=\begin{bmatrix}1\\ 1\end{bmatrix}\begin{bmatrix}0.08 & 0.08\end{bmatrix} \quad (2.4.40)$$

$$(\mathbf{I}-\mathbf{U}_1 z_3^{-1})\mathrm{adj}\mathbf{S}(z_3)$$

$$=\left(\begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix}-\begin{bmatrix}-0.60355 & 0.10355\\ 0.10355 & -0.60355\end{bmatrix}\left(\frac{1}{-1}\right)\right)\begin{bmatrix}1\\ 1\end{bmatrix}\begin{bmatrix}0.08 & 0.08\end{bmatrix} \tag{2.4.41}$$

$$=\begin{bmatrix}0.04 & 0.04\\ 0.04 & 0.04\end{bmatrix}=\begin{bmatrix}0.04\\ 0.04\end{bmatrix}\begin{bmatrix}1 & 1\end{bmatrix}$$

$$\mathrm{adj}\mathbf{S}(z_4)=\begin{bmatrix}3.047356 & -3.047356\\ -3.047356 & 3.047356\end{bmatrix}$$

$$=\begin{bmatrix}3.047356\\ -3.047356\end{bmatrix}\begin{bmatrix}1 & -1\end{bmatrix} \tag{2.4.42}$$

$$(\mathbf{I}-\mathbf{U}_1 z_4^{-1})\mathrm{adj}\mathbf{S}(z_4)$$

$$=\left(\begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix}-\begin{bmatrix}-0.6035 & 0.1035\\ 0.1035 & -0.6035\end{bmatrix}\left(\frac{1}{0.7071}\right)\right)\begin{bmatrix}1\\ -1\end{bmatrix}\begin{bmatrix}3.0473 & -3.0473\end{bmatrix} \tag{2.4.43}$$

$$=\begin{bmatrix}0.8926 & -0.8926\\ -0.8926 & 0.8926\end{bmatrix}=\begin{bmatrix}0.8926\\ -0.8926\end{bmatrix}\begin{bmatrix}1 & -1\end{bmatrix}$$

$$\mathbf{r}_3=\begin{bmatrix}1 & 1\end{bmatrix} \tag{2.4.44}$$

$$\mathbf{r}_4=\begin{bmatrix}1 & -1\end{bmatrix} \tag{2.4.45}$$

step3-2 use r3,r4 vector to get U_2 matrix

We then assemble the matrix

$$\mathbf{R}=\begin{bmatrix}\mathbf{r}_3\\ \mathbf{r}_4\end{bmatrix}=\begin{bmatrix}1 & 1\\ 1 & -1\end{bmatrix} \tag{2.4.46}$$

$$\mathbf{R}^{-1}=\frac{1}{2}\begin{bmatrix}1 & 1\\ 1 & -1\end{bmatrix} \tag{2.4.47}$$

and the diagonal matrix

$$\mathbf{D}=\begin{bmatrix}z_3 & 0\\ 0 & z_4\end{bmatrix} \tag{2.4.48}$$

from which we can obtain the $\mathbf{U}_2$ matrix.

$$\mathbf{U}_2 = \mathbf{R}^{-1}\mathbf{D}\mathbf{R} = \frac{1}{2}\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}\begin{bmatrix} z_3 & 0 \\ 0 & z_4 \end{bmatrix}\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} = \begin{bmatrix} -0.1465 & -0.8536 \\ -0.8536 & -0.1465 \end{bmatrix} \quad (2.4.49)$$

step4 get A_0 matrix by Cholesky factorization and then A_1 A_2 matrix

Consider $\mathbf{A}(z) = (\mathbf{I} - \mathbf{U}_1 z^{-1})(\mathbf{I} - \mathbf{U}_2 z^{-1})\mathbf{A}_0$, so that

$$\mathbf{S}(z) = (\mathbf{I} - \mathbf{U}_1 z^{-1})(\mathbf{I} - \mathbf{U}_2 z^{-1})(\mathbf{A}_0 \mathbf{A}_0^T)(\mathbf{I} - \mathbf{U}_2 z^1)(\mathbf{I} - \mathbf{U}_1 z^1) \quad (2.4.50)$$

$$\mathbf{S}(1) = (\mathbf{I} - \mathbf{U}_1)(\mathbf{I} - \mathbf{U}_2)(\mathbf{A}_0 \mathbf{A}_0^T)(\mathbf{I} - \mathbf{U}_2)(\mathbf{I} - \mathbf{U}_1) \quad (2.4.51)$$

$$\mathbf{A}_0 \mathbf{A}_0^T = (\mathbf{I} - \mathbf{U}_1)^{-1}(\mathbf{I} - \mathbf{U}_2)^{-1}\mathbf{S}(1)(\mathbf{I} - \mathbf{U}_2)^{-1}(\mathbf{I} - \mathbf{U}_1)^{-1} \quad (2.4.52)$$

Using a MATLAB function, we can get an upper triangular matrix by Cholesky factorisation. The run results are shown as following.

$$\mathbf{A}_0 = \begin{bmatrix} 0.8000 & 0 \\ 0 & 0.8000 \end{bmatrix} \quad (2.4.53)$$

$$\mathbf{A}_1 = -(\mathbf{U}_1 + \mathbf{U}_2)\mathbf{A}_0 = \begin{bmatrix} 0.6000 & 0.6000 \\ 0.6000 & 0.6000 \end{bmatrix} \quad (2.4.54)$$

$$\mathbf{A}_2 = \mathbf{U}_1 \mathbf{U}_2 \mathbf{A}_0 = \begin{bmatrix} 0 & 0.4000 \\ 0.4000 & 0 \end{bmatrix} \quad (2.4.55)$$

step5 get final result : transfer function

Finally, the required minimum-delay transfer function is given as following.

$$\mathbf{A}(z) = \mathbf{A}_0 + \mathbf{A}_1 z^{-1} + \mathbf{A}_2 z^{-2} = \begin{bmatrix} 0.8000 & 0 \\ 0 & 0.8000 \end{bmatrix} + \begin{bmatrix} 0.6000 & 0.6000 \\ 0.6000 & 0.6000 \end{bmatrix} z^{-1} + \begin{bmatrix} 0 & 0.4000 \\ 0.4000 & 0 \end{bmatrix} z^{-2} \quad (2.4.56)$$

step6 check result

From the factorisation result of step5, we can give $\mathbf{S}(z)$ as following.

$$\mathbf{S}(z)=\mathbf{A}(z)\mathbf{A}_*(z)$$
$$=\begin{bmatrix}0 & 0.3200\\0.3200 & 0\end{bmatrix}z^{-2}+\begin{bmatrix}0.7200 & 0.7200\\0.7200 & 0.7200\end{bmatrix}z^{-1}+\begin{bmatrix}1.5200 & 0.7200\\0.7200 & 1.5200\end{bmatrix} \quad (2.4.57)$$
$$+\begin{bmatrix}0.7200 & 0.7200\\0.7200 & 0.7200\end{bmatrix}z+\begin{bmatrix}0 & 0.3200\\0.3200 & 0\end{bmatrix}z^{2}$$

Otherwise from the theoretical result, when c1=0.8, c2=0.6, d1=0.6, d2=0.4

$$\mathbf{A}(z)=\begin{bmatrix}0.8 & 0\\0 & 0.8\end{bmatrix}+\begin{bmatrix}0.6 & 0.6\\0.6 & 0.6\end{bmatrix}z^{-1}+\begin{bmatrix}0 & 0.4\\0.4 & 0\end{bmatrix}z^{-2} \quad (2.4.58)$$

So that

$$\mathbf{S}(z)=\mathbf{A}(z)\mathbf{A}_*(z)$$
$$=\begin{bmatrix}0 & 0.3200\\0.3200 & 0\end{bmatrix}z^{-2}+\begin{bmatrix}0.7200 & 0.7200\\0.7200 & 0.7200\end{bmatrix}z^{-1}+\begin{bmatrix}1.5200 & 0.7200\\0.7200 & 1.5200\end{bmatrix} \quad (2.4.59)$$
$$+\begin{bmatrix}0.7200 & 0.7200\\0.7200 & 0.7200\end{bmatrix}z+\begin{bmatrix}0 & 0.3200\\0.3200 & 0\end{bmatrix}z^{2}$$

This is the same as previous factorisation result (2.4.57), so that we confirmed that (2.4.58) is a proper factorisation result of (2.4.29).

2.4.4 Appendix: Mathematical Proofs for Binomial Approach

Theorem of $\mathrm{adj}\mathbf{A}(z_i)=\mathbf{c}_i\mathbf{r}_i$

Theroem[6)]

Let

$$\mathbf{A}(z)=\mathbf{A}_0+\mathbf{A}_1z^{-1}+\ldots\ldots+\mathbf{A}_mz^{-m} \quad (2.4.60)$$

be a $p\times p$ polynomial matrix. Then

$$\det \mathbf{A}(z) = 0 \tag{2.4.61}$$

is called the determinantal equation, and its roots are denoted by $z_1, z_2, \ldots, z_m$, where m is the degree of polynomial $\det \mathbf{A}(z)$ in z. The (constant) matrix $\mathbf{A}(z_i)$ is obtained by substitution of any root z_i for z in the polynomial matrix $\mathbf{A}(z)$ is necessarily singular. When z_i is an unrepeated root, then the $p \times p$ matrix $\mathbf{A}(z_i)$ has rank $p-1$, and the adjugate evaluated at $z = z_i$, $\operatorname{adj}\mathbf{A}(z_i)$, is a $p \times p$ matrix with rank 1, and it is expressible as a product of the form

$$\operatorname{adj}\mathbf{A}(z_i) = \mathbf{c}_i \mathbf{r}_i \tag{2.4.62}$$

where $\mathbf{c}_i$ is a nonzero (constant) column vector, $\mathbf{r}_i$ is a nonzero (constant) row vector appropriate to given root z_i

Proof

Because $\det \mathbf{A}(z_i) = 0$, $p \times p$ matrix $\mathbf{A}(z_i)$ is singular. For if $\mathbf{A}(z_i)$ is simply degenerate, $\operatorname{adj}\mathbf{A}(z_i)$ by definition cannot be null. Hence, since the product $\mathbf{A}(z_i)\operatorname{adj}\mathbf{A}(z_i)$, which equals $\det \mathbf{A}(z_i)\mathbf{I}$ is null, $\operatorname{adj}\mathbf{A}(z_i)$ must be a matrix of unit rank. A square matrix $\mathbf{A}$ which has proportional columns and rows is rank of 1, and conversely. Such a matrix has effectively one linearly independent column and one linearly independent row, and is expressible as a product of the type $\mathbf{cr}$ where $\mathbf{c}$ is a nonzero column vector, $\mathbf{r}$ is a nonzero row vector.

Determine U matrix of binominal form

Case 1: 1-order matrix polynomial

In order to factorize a given spectral density matrix $\mathbf{S}(z)$ into $\mathbf{S}(z) = \mathbf{A}(z)\mathbf{A}_*(z)$, we have supposed that $\mathbf{A}(z)$ can be written into a binomial factors form as following.

$$\mathbf{A}(z) = (\mathbf{I} - \mathbf{U}_1 z^{-1})(\mathbf{I} - \mathbf{U}_2 z^{-1}) \cdots (\mathbf{I} - \mathbf{U}_m z^{-1})\mathbf{A}_0 \tag{2.4.63}$$

where $\mathbf{U}_1, \mathbf{U}_2, \cdots \mathbf{U}_m$ are $p \times p$ constant matrices. Then also $\mathbf{A}_*(z)$ can be written as following.

$$\mathbf{A}_*(z) = \mathbf{A}_0^{*T}(\mathbf{I} - \mathbf{U}_m^{*T} z) \cdots (\mathbf{I} - \mathbf{U}_2^{*T} z)(\mathbf{I} - \mathbf{U}_1^{*T} z) \tag{2.4.64}$$

For the most simple case of 1-rder matrix polynomial given as following.

$$\mathbf{A}(z) = \mathbf{A}_0 + \mathbf{A}_1 z^{-1} \tag{2.4.65}$$

we can express $\mathbf{A}(z)$ into binomial form

$$\mathbf{A}(z) = (\mathbf{I} - \mathbf{U}z^{-1})\mathbf{A}_0$$

where

$$\mathbf{A}_1 = -\mathbf{U}\mathbf{A}_0$$

Here we will show that how to calculate matrix $\mathbf{U}$. At beginning, for a given determinantal equation $\det \mathbf{S}(z) = 0$, suppose all of its roots z_i $(i = 1,2,...p)$ are distinct, finite and nonzero. From

$$\det \mathbf{S}(z_i) = \det \mathbf{A}(z_i) \det \mathbf{A}_*(z_i) = 0,\ (i = 1,2,...p), \tag{2.4.66}$$

we can get $\det \mathbf{A}(z_i) = 0$. Because $\mathbf{A}(z) = \mathbf{A}_0(\mathbf{I} - \mathbf{U}z^{-1})$, so that

$$\begin{aligned} &\det \mathbf{A}(z_i) = \det\{(\mathbf{I} - \mathbf{U}z_i^{-1})\mathbf{A}_0\} \\ &= \det(z_i \mathbf{I} - \mathbf{U}) \det\{\mathbf{A}_0 z_i^{-1}\} = 0 \end{aligned},\ (i = 1,2,...p) \tag{2.4.67}$$

Here because generally $\det\{\mathbf{A}_0 z_i^{-1}\} \neq 0$, then $\det(z_i \mathbf{I} - \mathbf{U}) = 0$. This is the same form of characteristic equation of matrix $\mathbf{U}$ and its eigenvalues equal to $\lambda = z_i$,

$(i=1,2,...p)$. Furthermore we can write $\text{adj}\mathbf{S}(z_i)=\mathbf{c}_i\mathbf{r}_i$, (Note that $\mathbf{c}_i$ and $\mathbf{r}_i$ are by no means unique), is where $\mathbf{c}_i$ is column vector and r_i is row vector, then

$$\begin{aligned}
&\det\mathbf{S}(z_i)=\text{adj}\,\mathbf{S}(z_i)\mathbf{S}(z_i)\\
&=\text{adj}\,\mathbf{S}(z_i)\{(\mathbf{I}-\mathbf{U}z_i^{-1})\mathbf{A}_0\mathbf{A}_0^{*T}(\mathbf{I}-\mathbf{U}^{*T}z_i)\}\\
&=\mathbf{c}_i\mathbf{r}_i(\mathbf{I}-\mathbf{U}z_i^{-1})\mathbf{A}_0\mathbf{A}_0^{*T}(\mathbf{I}-\mathbf{U}^{*T}z_i)
\end{aligned} \tag{2.4.68}$$

This means

$$\mathbf{c}_i\mathbf{r}_i(\mathbf{I}-\mathbf{U}z_i^{-1})=0 \tag{2.4.69}$$

$$\mathbf{r}_i(z_i\mathbf{I}-\mathbf{U})=0 \tag{2.4.70}$$

So we can see that $\mathbf{r}_i,(i=1,2,...p)$ is left eigenvector* of matrix $\mathbf{U}$. At last if we denote

---

* left eigenvector:

Any row vector that satisfy $\mathbf{xA}=\lambda\mathbf{x}$, $\mathbf{x}\neq 0$ is called as left eigenvector for eigenvalue λ. Because $\mathbf{A}^T\mathbf{x}^T=\lambda\mathbf{x}^T$, we can see that $\mathbf{x}^T$ is just the eigenvector of matrix $\mathbf{A}^T$. If x_i is left eigenvector of matrix $\mathbf{A}$, we can make a nonsingular matrix by using these left eigenvector as following.

$$\mathbf{T}=\begin{bmatrix}x_1\\ \vdots\\ x_m\end{bmatrix}.$$

Then from

$$\mathbf{TA}=[x_1\mathbf{A}\ \cdots\ x_m\mathbf{A}]=[\lambda_1x_1\ \cdots\ \lambda_mx_m]=\begin{bmatrix}\lambda_1 & & 0\\ & \ddots & \\ 0 & & \lambda_m\end{bmatrix}[x_1\ \cdots\ x_m]=\mathbf{\Lambda T}$$

So that

$$\mathbf{A}=\mathbf{T}^{-1}\mathbf{\Lambda T}$$

$$\mathbf{R} = \begin{bmatrix} \mathbf{r}_1 \\ \vdots \\ \mathbf{r}_m \end{bmatrix} \tag{2.4.71}$$

$$\mathbf{D} = diag(\lambda_1, \lambda_2 \cdots \lambda_p) = diag(z_1, z_2 \cdots z_p) \tag{2.4.72}$$

we can write matrix $\mathbf{U}$ as a similarity transformation $\mathbf{R}$ of its eigenvalue matrix $\mathbf{D}$ as following.

$$\mathbf{U} = \mathbf{R}^{-1}\mathbf{D}\mathbf{R} \tag{2.4.73}$$

Case 2: 2-order matrix polynomial

For a 2-order matrix polynomial

$$\mathbf{A}(z) = \mathbf{A}_0 + \mathbf{A}_1 z^{-1} + \mathbf{A}_2 z^{-2} \tag{2.4.74}$$

We can express $\mathbf{A}(z)$ into binomial form as following.

$$\mathbf{A}(z) = (\mathbf{I} - \mathbf{U}_1 z^{-1})(\mathbf{I} - \mathbf{U}_2 z^{-1})\mathbf{A}_0, \tag{2.4.75}$$

Here, we will show that how to calculate matrix $\mathbf{U}_1$ and $\mathbf{U}_2$.

At beginning, for a given determinantal equation $\det \mathbf{S}(z) = 0$, suppose all of its roots z_i $(i = 1,2,...p, p+1,...2p)$ are distinct, finite and nonzero. From

$$\det \mathbf{S}(z_i) = \det \mathbf{A}(z_i) \det \mathbf{A}_*(z_i) = 0, \ (i = 1,2,...p, p+1,...2p) \tag{2.4.76}$$

we can get $\det \mathbf{A}(z_i) = 0$. Because $\mathbf{A}(z) = (\mathbf{I} - \mathbf{U}_1 z^{-1})(\mathbf{I} - \mathbf{U}_2 z^{-1})\mathbf{A}_0$, so that

$$\begin{aligned} &\det \mathbf{A}(z_i) = \det \{(\mathbf{I} - \mathbf{U}_1 z_i^{-1})(\mathbf{I} - \mathbf{U}_2 z_i^{-1})\mathbf{A}_0\} \\ &= \det(z_i \mathbf{I} - \mathbf{U}_1) \det(z_i \mathbf{I} - \mathbf{U}_2) \det\{\mathbf{A}_0 z_i^{-2}\} \\ &= 0 \end{aligned} \tag{2.4.77}$$

$(i = 1,2,...p, p+1,...2p)$

Generally $\det\{\mathbf{A}_0 z_i^{-2}\} \neq 0$, so that we can get

$$\det(z_i\mathbf{I} - \mathbf{U}_1) = 0,\ (i = 1,2,...p) \tag{2.4.78}$$

$$\det(z_i\mathbf{I} - \mathbf{U}_2) = 0,\ (i = p+1,...2p) \tag{2.4.79}$$

where we divide all 2p zeros into 2 groups, the first group of p zeros $(z_i, i = 1,2,...p)$ are used for composing matrix $\mathbf{U}_1$ and the second group of p zeros $(z_i, i = p+1, p+2,...2p)$ are used for composing matrix $\mathbf{U}_2$.

(2.4.78) is characteristic equation of matrix $\mathbf{U}_1$. The eigenvalues for $\mathbf{U}_1$ equal to $\lambda = z_i$, $(i = 1,2,...p)$. Also (2.4.79) is characteristic equation of matrix $\mathbf{U}_2$.and the eigenvalues for $\mathbf{U}_2$ equal to $\lambda = z_i$, $(i = p+1, p+2,...2p)$.

$$\begin{aligned}\mathbf{S}(z_i) &= \mathbf{A}(z_i)\mathbf{A}_*(z_i)\\ &= (\mathbf{I} - \mathbf{U}_1 z_i^{-1})(\mathbf{I} - \mathbf{U}_2 z_i^{-1})\mathbf{A}_0\mathbf{A}_0^{*T}(\mathbf{I} - \mathbf{U}_2^{*T} z_i)(\mathbf{I} - \mathbf{U}_1^{*T} z_i)\end{aligned} \tag{2.4.80}$$

For the first group of p zeros, because we can write $\mathrm{adj}\mathbf{S}(z_i) = \mathbf{c}_i\mathbf{r}_i$, $(i = 1,2,...p)$, where $\mathbf{c}_i$ is column vector and $\mathbf{r}_i$ is row vector, so that

$$\begin{aligned}\det\mathbf{S}(z_i) &= \mathrm{adj}\mathbf{S}(z_i)\mathbf{S}(z_i)\\ &= \mathrm{adj}\mathbf{S}(z_i)\{(\mathbf{I} - \mathbf{U}_1 z_i^{-1})(\mathbf{I} - \mathbf{U}_2 z_i^{-1})\mathbf{A}_0\mathbf{A}_0^{*T}(\mathbf{I} - \mathbf{U}_2^{*T} z_i)(\mathbf{I} - \mathbf{U}_1^{*T} z_i)\}\\ &= \mathbf{c}_i\mathbf{r}_i(\mathbf{I} - \mathbf{U}_1 z_i^{-1})(\mathbf{I} - \mathbf{U}_2 z_i^{-1})\mathbf{A}_0\mathbf{A}_0^{*T}(\mathbf{I} - \mathbf{U}_2^{*T} z_i)(\mathbf{I} - \mathbf{U}_1^{*T} z_i)\end{aligned} \tag{2.4.81}$$

This means

$$\mathbf{c}_i\mathbf{r}_i(\mathbf{I} - \mathbf{U}_1 z_i^{-1}) = 0,\ (i = 1,2,...p) \tag{2.4.82}$$

$$\mathbf{r}_i(z_i\mathbf{I} - \mathbf{U}_2) = 0,\ (i = 1,2,...p) \tag{2.4.83}$$

So we can see that $\mathbf{r}_i, (i = 1,2,...p)$ is left eigenvector of matrix $\mathbf{U}_1$. This means if we denote

$$\mathbf{R}_1 = \begin{bmatrix} \mathbf{r}_1 \\ \vdots \\ \mathbf{r}_m \end{bmatrix} \tag{2.4.84}$$

$$\mathbf{D}_1 = diag(\lambda_1, \lambda_2 \cdots \lambda_p) = diag(z_1, z_2 \cdots z_p) \tag{2.4.85}$$

we can write matrix $\mathbf{U}_1$ as a similarity transformation $\mathbf{R}_1$ of its eigenvalue matrix $\mathbf{D}_1$ as following.

$$\mathbf{U}_1 = \mathbf{R}_1^{-1}\mathbf{D}_1\mathbf{R}_1 \tag{2.4.86}$$

Next for the second group of p zeros, because we can write

$$\mathrm{adj}\mathbf{S}(z_i) = \mathbf{c}_i\mathbf{r}_i, \ (i = p+1,...2p) \tag{2.4.87}$$

so that

$$\mathrm{adj}\mathbf{S}(z_i)(\mathbf{I} - \mathbf{U}_1 z_i) = \mathbf{c}_i\mathbf{r}_i(\mathbf{I} - \mathbf{U}_2 z_i) = \mathbf{c}'_i\mathbf{r}'_i, \ (i = p+1,...2p) \tag{2.4.88}$$

where $\mathbf{c}'_i$ is column vector and $\mathbf{r}'_i$ is row vector, then

$$\begin{aligned} \det\mathbf{S}(z_i) &= \mathrm{adj}\mathbf{S}(z_i)\mathbf{S}(z_i) \\ &= \mathrm{adj}\mathbf{S}(z_i)\{(\mathbf{I} - \mathbf{U}_1 z_i^{-1})(\mathbf{I} - \mathbf{U}_2 z_i^{-1})\mathbf{A}_0\mathbf{A}_0^{*T}(\mathbf{I} - \mathbf{U}_2^{*T} z_i)(\mathbf{I} - \mathbf{U}_1^{*T} z_i)\} \\ &= \mathbf{c}_i\mathbf{r}_i(\mathbf{I} - \mathbf{U}_1 z_i^{-1})(\mathbf{I} - \mathbf{U}_2 z_i^{-1})\mathbf{A}_0\mathbf{A}_0^{*T}(\mathbf{I} - \mathbf{U}_2^{*T} z_i)(\mathbf{I} - \mathbf{U}_1^{*T} z_i) \\ &= \mathbf{c}'_i\mathbf{r}'_i(\mathbf{I} - \mathbf{U}_2 z_i^{-1})\mathbf{A}_0\mathbf{A}_0^{*T}(\mathbf{I} - \mathbf{U}_2^{*T} z_i)(\mathbf{I} - \mathbf{U}_1^{*T} z_i) \end{aligned} \tag{2.4.89}$$

This means

$$\mathbf{c}'_i\mathbf{r}'_i(\mathbf{I} - \mathbf{U}_2 z_i^{-1}) = 0 \tag{2.4.90}$$

$$\mathbf{r}_i'(z_i\mathbf{I}-\mathbf{U}_2 z_i^{-1})=0 \tag{2.4.91}$$

So we can see that $\mathbf{r}_i'$ is eigenvector of matrix $\mathbf{U}_2$. This means if we denote

$$\mathbf{R}_2=\begin{bmatrix}\mathbf{r}_1'\\ \vdots \\ \mathbf{r}_m'\end{bmatrix} \tag{2.4.92}$$

$$\mathbf{D}_2=diag(\lambda_1,\lambda_2\cdots\lambda_p)=diag\left(z_{p+1},z_{p+2}\cdots z_{2p}\right) \tag{2.4.93}$$

we can write matrix $\mathbf{U}_2$ as a similarity transformation $\mathbf{R}_2$ of its eigenvalue matrix $\mathbf{D}_2$ as following.

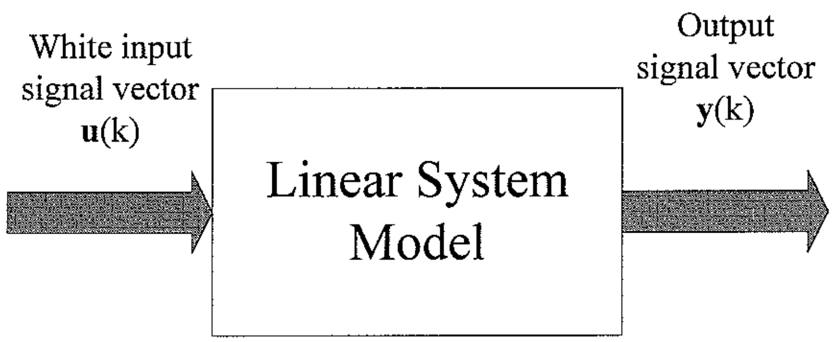
$$\mathbf{U}_2=\mathbf{R}_2^{-1}\mathbf{D}_2\mathbf{R}_2 \tag{2.4.94}$$

2.5 Solution via State Space Approach

2.5.1 Principle (after Kailath's Approach)

For a spectral factorisation method that not solve zeros of matrix polynomial is described by T. Kailath et al[7]. Let's consider the vector of random processes, we want estimate which cross-spectral density matrix, is a output of a multichannel linear system model with white input random signal vector as shown in Figure 6. Therefore by select a linear system model available to measured data vector, a modelling method of random processes can be obtained. Rewrite (2.2.11) as following, which means a moving average (MA) model is assumed for the linear system.

$$\begin{aligned}\mathbf{L}(z)&=\mathbf{I}_p+\mathbf{L}_1 z^{-1}+\ldots\ldots+\mathbf{L}_m z^{-m}\\ &=\mathbf{I}_p+\sum_{i=1}^{m}\mathbf{L}_i z^{-i}\end{aligned}$$



• Figure 6 Linear system model for multichannel spectral factorisation

When a white input signal vector $\mathbf{u}(z)$ is applied to this system, the output signal vector can be given as following.

$$\mathbf{y}(z) = \mathbf{L}(z)\mathbf{u}(z) \tag{2.5.1}$$

If the spectral matrix of white input signal vector is given by

$$\mathbf{S}_{uu}(z) = \boldsymbol{\Sigma} \tag{2.5.2}$$

then the spectral matrix of output signal vector can be expressed as following.

$$\mathbf{S}_{yy}(z) = \mathbf{L}(z)\mathbf{S}_{uu}(z)\mathbf{L}_{*}(z) = \mathbf{L}(z)\boldsymbol{\Sigma}\mathbf{L}_{*}(z) \tag{2.5.3}$$

State space description and transfer function matrix

From (2.5.1), the input-output relation for the linear system model $\mathbf{L}(z)$ can be written as

$$\mathbf{y}(n) = \mathbf{I}_p + \mathbf{L}_1\mathbf{u}(n-1) + \ldots\ldots + \mathbf{L}_m\mathbf{u}(n-m) \tag{2.5.4}$$

There are of course several ways for representing this relation in state space form. One option is the following realization. By define the coefficient matrices as following,

$$\mathbf{F} = \underbrace{\left.\begin{bmatrix} \mathbf{0}_p & & & \\ \mathbf{I}_p & \mathbf{0}_p & & \\ & \ddots & \ddots & \\ & & \mathbf{I}_p & \mathbf{0}_p \end{bmatrix}\right\} m}_{m} \tag{2.5.5}$$

$$\mathbf{G} = \left.\begin{bmatrix} \mathbf{L}_m \\ \mathbf{L}_{m-1} \\ \\ \mathbf{L}_1 \end{bmatrix}\right\} m \tag{2.5.6}$$

$$\mathbf{H} = \underbrace{\begin{bmatrix} \mathbf{0}_p & \mathbf{0}_p & \cdots & \mathbf{0}_p & \mathbf{I}_p \end{bmatrix}}_{m} \tag{2.5.7}$$

the moving average random signal vector $\mathbf{y}(n)$ can be regarded as the output of the following state space model.

$$\begin{aligned} \mathbf{x}(n+1) &= \mathbf{F}\mathbf{x}(n) + \mathbf{G}\mathbf{u}(n) \\ \mathbf{y}(n) &= \mathbf{H}\mathbf{x}(n) + \mathbf{u}(n) \end{aligned} \tag{2.5.8}$$

We can proof that the transfer function of this state space equation is $\mathbf{L}(z)$. (cf. section 2.5.4)

$$\begin{aligned} \mathbf{L}(z) &= \mathbf{I}_p + \mathbf{L}_1 z^{-1} + \ldots\ldots + \mathbf{L}_m z^{-m} \\ &= \mathbf{I}_p + \sum_{i=1}^{m} \mathbf{L}_i z^{-i} \end{aligned} \tag{2.5.9}$$

Note further that the transfer function $\mathbf{L}(z)$ must stable and have all its zeros strictly inside the unit circle. Using the matrix inversion lemma, we can find that

$$\mathbf{L}^{-1}(z) = \mathbf{I}_p - \mathbf{H}(z\mathbf{I}_{mp} - (\mathbf{F} - \mathbf{G}\mathbf{H}))^{-1}\mathbf{G} \tag{2.5.10}$$

So that the matrix $(\mathbf{F} - \mathbf{G}\mathbf{H})$ must also have all its eigenvalues strictly inside the unit circle.

In other words, we have shown so far that starting with a matrix polynomial $\mathbf{S}(z)$ that is strictly positive on the unit circle, there must exist a state space model of form (2.5.8) such that

(i) $\mathbf{F}$ is a stable matrix.

(ii) $(\mathbf{F} - \mathbf{G}\mathbf{H})$ is a stable.

(iii) The entries of $\mathbf{G}$ determine $\mathbf{L}(z)$ and the variance of $\mathbf{u}(n)$ determines $\mathbf{P}$.

(iv) Using these three facts, and the above state space model, we can now show how to determine the unknown matrix $\mathbf{G}$ and hence, the canonical factor $\mathbf{L}(z)$ and the variance $\boldsymbol{\Sigma}$.

Determine matrix $\mathbf{G}$ and canonical factor $\mathbf{L}(z)$ and variance $\boldsymbol{\Sigma}$

To begin with, the state space model is assumed to start in (2.5.8) and, therefore, the stability of $\mathbf{F}$ guarantees a stationary state vector process $\{\mathbf{x}(n)\}$. Let $\mathbf{C}$ denotes its covariance matrix,

$$\mathbf{C} = E[\mathbf{x}(n)\mathbf{x}^{*T}(n)] \tag{2.5.11}$$

It then from the state (2.5.8)

$$E[\mathbf{x}(n+1)\mathbf{x}^{*T}(n+1)] = \mathbf{F}E[\mathbf{x}(n)\mathbf{x}^{*T}(n)]\mathbf{F}^{*T} + \mathbf{G}E[\mathbf{u}(n)\mathbf{u}^{*T}(n)]\mathbf{G}^{*T} \tag{2.5.12}$$

Because

$$\mathbf{C} = E[\mathbf{x}(n)\mathbf{x}^{*T}(n)] = E[\mathbf{x}(n+1)\mathbf{x}^{*T}(n+1)] \tag{2.5.13}$$

for $\mathbf{x}$ is stationary process and

$$E[\mathbf{u}(n)\mathbf{u}^{*T}(n)] = \mathbf{S}_{uu}(z) = \boldsymbol{\Sigma} \tag{2.5.14}$$

for $\mathbf{u}$ is white stationary process, then covariance matrix $\mathbf{C}$ must satisfies the matrix equation

$$\mathbf{C} = \mathbf{F}\mathbf{C}\mathbf{F}^{*T} + \mathbf{G}\mathbf{R}_p\mathbf{G}^{*T} \tag{2.5.15}$$

It also follows from the output equation in (2.5.8) that

$$E[\mathbf{y}(n)\mathbf{y}^{*T}(n)] = \mathbf{H}E[\mathbf{x}(n)\mathbf{x}^{*T}(n)]\mathbf{H}^{*T} + E[\mathbf{u}(n)\mathbf{u}^{*T}(n)] \tag{2.5.16}$$

Because

$$\mathbf{R}_0 = E[\mathbf{y}(n)\mathbf{y}^{*T}(n)] \tag{2.5.17}$$

as shown in (2.5.2), then

$$\mathbf{R}_0 = \mathbf{HCH}^{*T} + \mathbf{\Sigma} \tag{2.5.18}$$

therefore

$$\mathbf{\Sigma} = \mathbf{R}_0 - \mathbf{HCH}^{*T} \tag{2.5.19}$$

Finally, we evaluate the inner product $E[\mathbf{x}(n+1)\mathbf{y}^{*T}(n)]$ in two different ways. The first way uses the state space model to conclude that

$$\begin{aligned}
&E[\mathbf{x}(n+1)\mathbf{y}^{*T}(n)] = E[(\mathbf{Fx}(n) + \mathbf{Gu}(n))(\mathbf{Hx}(n) + \mathbf{u}(n))^{*T}] \\
&= \mathbf{F}E[\mathbf{x}(n)\mathbf{x}^{*T}(n)]\mathbf{H}^{*T} + \mathbf{F}E[\mathbf{x}(n)\mathbf{u}^{*T}(n)] + \mathbf{G}E[\mathbf{u}(n)\mathbf{x}^{*T}(n)]\mathbf{H}^{*T} + \mathbf{G}E[\mathbf{u}(n)\mathbf{u}^{*T}(n)] \\
&= \mathbf{F}E[\mathbf{x}(n)\mathbf{x}^{*T}(n)]\mathbf{H}^{*T} + \mathbf{G}E[\mathbf{u}(n)\mathbf{u}^{*T}(n)] \\
&= \mathbf{FCH}^{*T} + \mathbf{G\Sigma}
\end{aligned} \tag{2.5.20}$$

where we used the fact that $\mathbf{u}(n)$ and $\mathbf{x}(n)$ are uncorrelated.

The second way to evaluate $E[\mathbf{x}(n+1)\mathbf{y}^{*T}(n)]$ is the following. As the first step, by write $\mathbf{x}(n)$ in block for each channel as following,

$$\mathbf{x}(n) = \begin{bmatrix} x_{p1}(n) \\ x_{p2}(n) \\ \vdots \\ x_{pm}(n) \end{bmatrix} \tag{2.5.21}$$

and solve $E[\mathbf{x}_{pk}(n+1)\mathbf{y}^{*T}(n)]$ individually, we can show that

$$E[\mathbf{x}_{pk}(n+1)\mathbf{y}^{*T}(n)] = \mathbf{R}_{m-k+1} \tag{2.5.22}$$

for $k = 1,2,...m$. Then put these results together, we easily verify that

$$
E[\mathbf{x}(n+1)\mathbf{y}^{*T}(n)] = E[\begin{bmatrix} \mathbf{x}_{p1}(n+1) \\ \mathbf{x}_{p2}(n+1) \\ \vdots \\ \mathbf{x}_{pm}(n+1) \end{bmatrix}\mathbf{y}^{*T}(n)]
$$

$$
= \begin{bmatrix} E[\mathbf{x}_{p1}(n+1)\mathbf{y}^{*T}(n)] \\ E[\mathbf{x}_{p2}(n+1)\mathbf{y}^{*T}(n)] \\ \vdots \\ E[\mathbf{x}_{pm}(n+1)\mathbf{y}^{*T}(n)] \end{bmatrix} = \begin{bmatrix} \mathbf{R}_m \\ \mathbf{R}_{m-1} \\ \vdots \\ \mathbf{R}_1 \end{bmatrix} \tag{2.5.23}
$$

We will denote this $mp \times p$ matrix by $\mathbf{N}$,

$$
\mathbf{N} = \begin{bmatrix} \mathbf{R}_m \\ \mathbf{R}_{m-1} \\ \vdots \\ \mathbf{R}_1 \end{bmatrix} \tag{2.5.24}
$$

So that we can have

$$
E[\mathbf{x}(n+1)\mathbf{y}^{*T}(n)] = \mathbf{N} \tag{2.5.25}
$$

This means we find that the desired inner product $E[\mathbf{x}(n+1)\mathbf{y}^{*T}(n)]$ is completely determined in terms of the given coefficients of spectral matrix $\mathbf{S}(z)$. Combining result of first step in (2.5.20) with result of second step in (2.5.25) we can obtain the equality

$$
\mathbf{FCH}^{*T} + \mathbf{G\Sigma} = \mathbf{N} \tag{2.5.26}
$$

This gives

$$
\mathbf{G} = (\mathbf{N} - \mathbf{FCH}^{*T})\mathbf{\Sigma}^{-1} \tag{2.5.27}
$$

which express the unknown $\mathbf{G}$ in term of the matrix $\mathbf{C}$. At last, substitute (2.5.27) into (2.5.15)

$$
\begin{aligned}
\mathbf{C} &= \mathbf{FCF}^{*T} + \mathbf{G\Sigma}G^{*T} \\
&= \mathbf{FCF}^{*T} + (\mathbf{N} - \mathbf{FCH}^{*T})\mathbf{\Sigma}^{-1}\mathbf{\Sigma}(\mathbf{\Sigma}^{-1})^{*T}(\mathbf{N} - \mathbf{FCH}^{*T})^{*T} \\
&= \mathbf{FCF}^{*T} + (\mathbf{N} - \mathbf{FCH}^{*T})(\mathbf{\Sigma}^{-1})^{*T}(\mathbf{N} - \mathbf{FCH}^{*T})^{*T}
\end{aligned} \tag{2.5.28}
$$

furthermore, substitute (2.5.19) into this equation, we can obtain

$$
\begin{aligned}
\mathbf{C} &= \mathbf{FCF}^{*T} + (\mathbf{N} - \mathbf{FCH}^{*T})(\mathbf{R}_0 - \mathbf{HCH}^{*T})^{-*T}(\mathbf{N} - \mathbf{FCH}^{*T})^{*T} \\
&= \mathbf{FCF}^{*T} + (\mathbf{N} - \mathbf{FCH}^{*T})(\mathbf{R}_0^{*T} - \mathbf{HC}^{*T}\mathbf{H}^{*T})^{-1}(\mathbf{N} - \mathbf{FCH}^{*T})^{*T} \\
&= \mathbf{FCF}^{*T} + (\mathbf{N} - \mathbf{FCH}^{*T})(\mathbf{R}_0 - \mathbf{HCH}^{*T})^{-1}(\mathbf{N} - \mathbf{FCH}^{*T})^{*T} \\
&= \mathbf{FCF}^{*T} - (\mathbf{FCH}^{*T} - \mathbf{N})(\mathbf{HCH}^{*T} - \mathbf{R}_0)^{-1}(\mathbf{FCH}^{*T} - \mathbf{N})^{*T}
\end{aligned}
\qquad (2.5.29)
$$

This is a discrete algebra Riccati equation regarding matrix $\mathbf{C}$.

Summary of solution

The above derivation therefore suggests the following procedure for finding $\mathbf{\Sigma}$ and the coefficients matrix of $\mathbf{L}(z)$ with $\mathbf{G}$, in the factorization $\mathbf{S}(z) = \mathbf{L}(z)\mathbf{\Sigma}\mathbf{L}_*(z)$.

(i) Define coefficient matrix of state space model in (2.5.8), that is

$$
\mathbf{F} = \underbrace{\begin{bmatrix} \mathbf{0}_p & & & \\ \mathbf{I}_p & \mathbf{0}_p & & \\ & \ddots & \ddots & \\ & & \mathbf{I}_p & \mathbf{0}_p \end{bmatrix}}_{m} \Bigg\} m
$$

$$
\mathbf{H} = \underbrace{\begin{bmatrix} \mathbf{0}_p & \mathbf{0}_p & \cdots & \mathbf{0}_p & \mathbf{I}_p \end{bmatrix}}_{m}
$$

$$
\mathbf{N} = \begin{bmatrix} \mathbf{R}_m \\ \mathbf{R}_{m-1} \\ \vdots \\ \mathbf{R}_1 \end{bmatrix}
$$

(ii) Determine the nonnegative solution $\mathbf{C}$ of the discrete algebra Riccati equation.

$$
\mathbf{C} = \mathbf{FCF}^{*T} - (\mathbf{FCH}^{*T} - \mathbf{N})(\mathbf{HCH}^{*T} - \mathbf{R}_0)^{-1}(\mathbf{FCH}^{*T} - \mathbf{N})^{*T}
$$

that results in a matrix

$$
\mathbf{F}_p = \mathbf{F} - (\mathbf{FCH}^{*T} - \mathbf{N})(\mathbf{HCH}^{*T} - \mathbf{R}_0)^{-1}\mathbf{H} \qquad (2.5.30)
$$

having all its eigenvalues strictly inside the unit circle. Such a nonnegative $\mathbf{C}$ that stabilizes $\mathbf{F}_p$ is guaranteed to exist by virtue of existence of model (2.5.8) itself. In fact, the nonnegative stabilizing $\mathbf{C}$ is unique.

(iii) Then set

$$\mathbf{\Sigma} = \mathbf{R}_0 - \mathbf{HCH}^{*T}$$

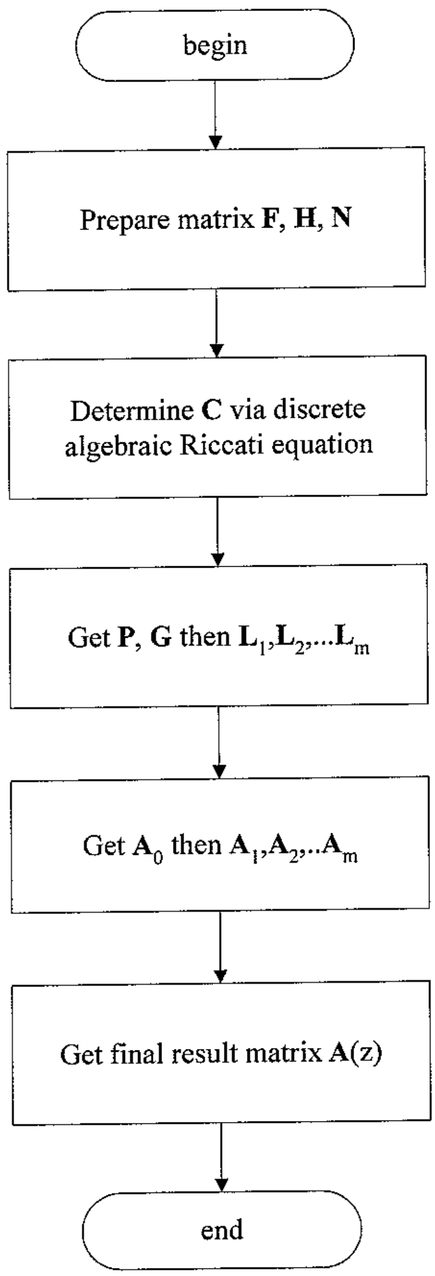
$$\mathbf{G} = (\mathbf{N} - \mathbf{FCH}^{*T})\mathbf{\Sigma}$$

where the component block of $\mathbf{G}$ matrix define the coefficients matrix of $\mathbf{L}(z)$ as following.

$$\mathbf{G} = \left.\begin{bmatrix} \mathbf{L}_m \\ \mathbf{L}_{m-1} \\ \\ \mathbf{L}_1 \end{bmatrix}\right\} m$$

2.5.2 Flow Chart

In order to describe the calculation procedures more precisely, a flow chart is shown in Figure 7.



• Figure 7 Flow chart of Riccati equation method for spectral factorisation

2.5.3 Practical Example

Let's use the problem in section 2.4.3 again. As we have shown in (2.4.29), we can write power spectrum function as following.

$$
\begin{aligned}
\mathbf{S}(z) &= \mathbf{A}(z)\mathbf{A}_*(z) \\
&= \begin{bmatrix} c_1 + d_1 z^{-1} & c_2 z^{-1} + d_2 z^{-2} \\ c_2 z^{-1} + d_2 z^{-2} & c_1 + d_1 z^{-1} \end{bmatrix} \begin{bmatrix} c_1 + d_1 z^{1} & c_2 z^{1} + d_2 z^{2} \\ c_2 z^{1} + d_2 z^{2} & c_1 + d_1 z^{1} \end{bmatrix} \\
&= \begin{bmatrix} (c_1 d_1 + c_2 d_2) z^{-1} + (c_1 c_1 + d_1 d_1 & c_1 d_2 z^{-2} + (c_1 c_2 + d_1 d_2) z^{-1} + (c_2 d_1 \\ + c_2 c_2 + d_2 d_2) + (c_1 d_1 + c_2 d_2) z^{1} & + c_2 d_1) + (c_1 c_2 + d_1 d_2) z^{1} + c_1 d_2 z^{2} \\ & \\ c_1 d_2 z^{-2} + (c_1 c_2 + d_1 d_2) z^{-1} + (c_2 d_1 & (c_1 d_1 + c_2 d_2) z^{-1} + (c_1 c_1 + d_1 d_1 \\ + c_2 d_1) + (c_1 c_2 + d_1 d_2) z^{1} + c_1 d_2 z^{2} & + c_2 c_2 + d_2 d_2) + (c_1 d_1 + c_2 d_2) z^{1} \end{bmatrix}
\end{aligned}
\qquad (2.5.31)
$$

Then we wish to find a minimum-delay operator

$$
\mathbf{A}(z) = \mathbf{A}_0 + \mathbf{A}_1 z^{-1} + \mathbf{A}_2 z^{-2} \qquad (2.5.32)
$$

such that

$$
\begin{aligned}
\mathbf{S}(z) &= \mathbf{A}(z)\mathbf{A}_*(z) \\
&= (\mathbf{A}_0 + \mathbf{A}_1 z^{-1} + \mathbf{A}_2 z^{-2})(\mathbf{A}_0^{*T} + \mathbf{A}_1^{*T} z^{1} + \mathbf{A}_2^{*T} z^{2})
\end{aligned}
\qquad (2.5.33)
$$

step1 Prepare F, H, N matrix

$$
\mathbf{O}_2 = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} \qquad (2.5.34)
$$

$$
\mathbf{I}_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \qquad (2.5.35)
$$

$$
\mathbf{F} = \begin{bmatrix} \mathbf{O}_2 & \mathbf{O}_2 \\ \mathbf{I}_2 & \mathbf{O}_2 \end{bmatrix} \qquad (2.5.36)
$$

$$
\mathbf{H} = \begin{bmatrix} \mathbf{O}_2 & \mathbf{I}_2 \end{bmatrix} \qquad (2.5.37)
$$

$$
\mathbf{R}_1 = \begin{bmatrix} c_1 d_1 + c_2 d_2 & c_1 c_2 + d_1 d_2 \\ c_1 c_2 + d_1 d_2 & c_1 d_1 + c_2 d_2 \end{bmatrix} \qquad (2.5.38)
$$

$$\mathbf{R}_2 = \begin{bmatrix} 0 & c_1 d_2 \\ c_1 d_2 & 0 \end{bmatrix} \quad (2.5.39)$$

$$\mathbf{N} = [\mathbf{R}_2 \quad \mathbf{R}_1] \quad (2.5.40)$$

step2 Determine C matrix via discrete algebra Riccati equation

$$\mathbf{C} = \begin{bmatrix} 0.16 & 0 & 0.24 & 0.24 \\ 0 & 0.16 & 0.24 & 0.24 \\ 0.24 & 0.24 & 0.88 & 0.72 \\ 0.24 & 0.24 & 0.72 & 0.88 \end{bmatrix} \quad (2.5.41)$$

step3 Get P, G matrices then L_1, L_2 matrix

$$\mathbf{P} = \begin{bmatrix} 0.64 & 0 \\ 0 & 0.64 \end{bmatrix} \quad (2.5.42)$$

$$\mathbf{G} = \begin{bmatrix} 0 & 0.5 \\ 0.5 & 0 \\ 0.75 & 0.75 \\ 0.75 & 0.75 \end{bmatrix} \quad (2.5.43)$$

$$\mathbf{L}_2 = \begin{bmatrix} 0 & 0.5 \\ 0.5 & 0 \end{bmatrix} \quad (2.5.44)$$

$$\mathbf{L}_1 = \begin{bmatrix} 0.75 & 0.75 \\ 0.75 & 0.75 \end{bmatrix} \quad (2.5.45)$$

step4 Get A_0 matrix then A_1, A_2 matrix

$$\mathbf{A}_0 = \begin{bmatrix} 0.8 & 0 \\ 0 & 0.8 \end{bmatrix} \quad (2.5.46)$$

$$\mathbf{A}_1 = \begin{bmatrix} 0.6 & 0.6 \\ 0.6 & 0.6 \end{bmatrix} \quad (2.5.47)$$

$$\mathbf{A}_2 = \begin{bmatrix} 0 & 0.4 \\ 0.4 & 0 \end{bmatrix} \tag{2.5.48}$$

step5 Get final result : transfer function

Finally, the required minimum-delay transfer function is given as following.

$$\begin{aligned} \mathbf{A}(z) &= \mathbf{A}_0 + \mathbf{A}_1 z^{-1} + \mathbf{A}_2 z^{-2} \\ &= \begin{bmatrix} 0.8000 & 0 \\ 0 & 0.8000 \end{bmatrix} + \begin{bmatrix} 0.6000 & 0.6000 \\ 0.6000 & 0.6000 \end{bmatrix} z^{-1} + \begin{bmatrix} 0 & 0.4000 \\ 0.4000 & 0 \end{bmatrix} z^{-2} \end{aligned} \tag{2.5.49}$$

step6 Check result

From result of step5, we can give spectral matrix $\mathbf{S}(z)$ as following.

$$\begin{aligned} \mathbf{S}(z) &= \mathbf{A}(z)\mathbf{A}_*(z) = (\mathbf{A}_0 + \mathbf{A}_1 z^{-1} + \mathbf{A}_2 z^{-2})(\mathbf{A}_0^{*T} + \mathbf{A}_1^{*T} z^1 + \mathbf{A}_2^{*T} z^2) \\ &= \mathbf{A}_2\mathbf{A}_0^{*T} z^{-2} + (\mathbf{A}_1\mathbf{A}_0^{*T} + \mathbf{A}_2\mathbf{A}_1^{*T})z^{-1} + (\mathbf{A}_0\mathbf{A}_0^{*T} + \mathbf{A}_1\mathbf{A}_1^{*T} + \mathbf{A}_2\mathbf{A}_2^{*T}) \\ &+ (\mathbf{A}_0\mathbf{A}_1^{*T} + \mathbf{A}_1\mathbf{A}_2^{*T})z + \mathbf{A}_2^{*T}\mathbf{A}_0 z^2 \end{aligned}$$

(2.5.50)

$$\begin{aligned} \mathbf{S}(z) &= \mathbf{A}(z)\mathbf{A}_*(z) \\ &= \begin{bmatrix} 0 & 0.3200 \\ 0.3200 & 0 \end{bmatrix} z^{-2} + \begin{bmatrix} 0.7200 & 0.7200 \\ 0.7200 & 0.7200 \end{bmatrix} z^{-1} + \begin{bmatrix} 1.5200 & 0.7200 \\ 0.7200 & 1.5200 \end{bmatrix} \\ &+ \begin{bmatrix} 0.7200 & 0.7200 \\ 0.7200 & 0.7200 \end{bmatrix} z + \begin{bmatrix} 0 & 0.3200 \\ 0.3200 & 0 \end{bmatrix} z^2 \end{aligned} \tag{2.5.51}$$

Otherwise from the theoretical result, when c1=0.8, c2=0.6, d1=0.6, d2=0.4

$$\mathbf{A}(z) = \begin{bmatrix} 0.8 & 0 \\ 0 & 0.8 \end{bmatrix} + \begin{bmatrix} 0.6 & 0.6 \\ 0.6 & 0.6 \end{bmatrix} z^{-1} + \begin{bmatrix} 0 & 0.4 \\ 0.4 & 0 \end{bmatrix} z^{-2} \tag{2.5.52}$$

$$\mathbf{A}_*(z) = \begin{bmatrix} 0.8 & 0 \\ 0 & 0.8 \end{bmatrix} + \begin{bmatrix} 0.6 & 0.6 \\ 0.6 & 0.6 \end{bmatrix} z^{1} + \begin{bmatrix} 0 & 0.4 \\ 0.4 & 0 \end{bmatrix} z^{2} \tag{2.5.53}$$

$$
\begin{aligned}
\mathbf{S}(z) &= \mathbf{A}(z)\mathbf{A}_*(z) \\
&= \begin{bmatrix} 0 & 0.3200 \\ 0.3200 & 0 \end{bmatrix} z^{-2} + \begin{bmatrix} 0.7200 & 0.7200 \\ 0.7200 & 0.7200 \end{bmatrix} z^{-1} + \begin{bmatrix} 1.5200 & 0.7200 \\ 0.7200 & 1.5200 \end{bmatrix} \\
&+ \begin{bmatrix} 0.7200 & 0.7200 \\ 0.7200 & 0.7200 \end{bmatrix} z + \begin{bmatrix} 0 & 0.3200 \\ 0.3200 & 0 \end{bmatrix} z^{2}
\end{aligned}
\qquad (2.5.54)
$$

This is the same as previous factorization result (2.5.46). So we confirmed that (2.5.49) is a proper factorization result of (2.5.30).

2.5.4 Appendix: Mathematical Proofs for State Space Approach

Transfer function of state space expression

Then the moving average random signal vector $\mathbf{y}(n)$ can be regarded as the output of the following state space model.

$$
\begin{aligned}
\mathbf{x}(n+1) &= \mathbf{F}\mathbf{x}(n) + \mathbf{G}\mathbf{u}(n) \\
\mathbf{y}(n) &= \mathbf{H}\mathbf{x}(n) + \mathbf{u}(n)
\end{aligned}
$$

where

$$
\mathbf{F} = \underbrace{\begin{bmatrix} \mathbf{0}_p & & & \\ \mathbf{I}_p & \mathbf{0}_p & & \\ & \ddots & \ddots & \\ & & \mathbf{I}_p & \mathbf{0}_p \end{bmatrix}}_{m} \Bigg\} m
$$

$$
\mathbf{G} = \begin{bmatrix} \mathbf{L}_m \\ \mathbf{L}_{m-1} \\ \vdots \\ \mathbf{L}_1 \end{bmatrix} \Bigg\} m
$$

$$
\mathbf{H} = \underbrace{\begin{bmatrix} \mathbf{0}_p & \mathbf{0}_p & \cdots & \mathbf{0}_p & \mathbf{I}_p \end{bmatrix}}_{m}
$$

Indeed, the transfer function of this state space model is given by

$$
\begin{aligned}
\mathbf{T}(z) &= \mathbf{I}_p + \mathbf{H}(z\mathbf{I}_{pm} - \mathbf{F})^{-1}\mathbf{G} \\
&= \mathbf{I}_p + \begin{bmatrix} \mathbf{0}_p & \cdots & \mathbf{0}_p & \mathbf{I}_p \end{bmatrix} \left(z \begin{bmatrix} \mathbf{I}_p & & & \\ & \mathbf{I}_p & & \\ & & \ddots & \\ & & & \mathbf{I}_p \end{bmatrix} - \begin{bmatrix} \mathbf{0}_p & & & \\ \mathbf{I}_p & \mathbf{0}_p & & \\ & \ddots & \ddots & \\ & & \mathbf{I}_p & \mathbf{0}_p \end{bmatrix} \right)^{-1} \begin{bmatrix} \mathbf{L}_m \\ \mathbf{L}_{m-1} \\ \vdots \\ \mathbf{L}_1 \end{bmatrix} \\
&= \mathbf{I}_p + z^{-1} \begin{bmatrix} \mathbf{0}_p & \cdots & \mathbf{0}_p & \mathbf{I}_p \end{bmatrix} \begin{bmatrix} \mathbf{I}_p & \mathbf{0}_p & & \mathbf{0}_p \\ -z^{-1}\mathbf{I}_p & \mathbf{I}_p & & \mathbf{0}_p \\ & \ddots & \ddots & \\ \mathbf{0}_p & & -z^{-1}\mathbf{I}_p & \mathbf{I}_p \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{L}_m \\ \mathbf{L}_{m-1} \\ \vdots \\ \mathbf{L}_1 \end{bmatrix}
\end{aligned}
\tag{2.5.55}
$$

Then by solving the following equation,

$$
\begin{bmatrix} \mathbf{I}_p & \mathbf{0}_p & \cdots & \mathbf{0}_p \\ -z^{-1}\mathbf{I}_p & \mathbf{I}_p & \cdots & \mathbf{0}_p \\ \vdots & \ddots & \ddots & \vdots \\ \mathbf{0}_p & \cdots & -z^{-1}\mathbf{I}_p & \mathbf{I}_p \end{bmatrix} \begin{bmatrix} \mathbf{B}_{p11} & \mathbf{B}_{p12} & \cdots & \mathbf{B}_{p1m} \\ \mathbf{B}_{p21} & \mathbf{B}_{p22} & \cdots & \mathbf{B}_{p2m} \\ \vdots & \ddots & \ddots & \vdots \\ \mathbf{B}_{pm1} & \mathbf{B}_{pm2} & \cdots & \mathbf{B}_{pmm} \end{bmatrix} = \begin{bmatrix} \mathbf{I}_p & \mathbf{0}_p & \cdots & \mathbf{0}_p \\ \mathbf{0}_p & \mathbf{I}_p & \cdots & \mathbf{0}_p \\ \vdots & \ddots & \ddots & \vdots \\ \mathbf{0}_p & \mathbf{0}_p & \cdots & \mathbf{I}_p \end{bmatrix}
\tag{2.5.56}
$$

We can get the inverse matrix

$$
\begin{aligned}
&\begin{bmatrix} \mathbf{I}_p & \mathbf{0}_p & \cdots & \mathbf{0}_p \\ -z^{-1}\mathbf{I}_p & \mathbf{I}_p & \cdots & \mathbf{0}_p \\ \vdots & \ddots & \ddots & \vdots \\ \mathbf{0}_p & \cdots & -z^{-1}\mathbf{I}_p & \mathbf{I}_p \end{bmatrix}^{-1} = \begin{bmatrix} \mathbf{B}_{p11} & \mathbf{B}_{p12} & \cdots & \mathbf{B}_{p1m} \\ \mathbf{B}_{p21} & \mathbf{B}_{p22} & \cdots & \mathbf{B}_{p2m} \\ \vdots & \ddots & \ddots & \vdots \\ \mathbf{B}_{pm1} & \mathbf{B}_{pm2} & \cdots & \mathbf{B}_{pmm} \end{bmatrix} \\
&= \begin{bmatrix} \mathbf{I}_p & \mathbf{0}_p & \cdots & \mathbf{0}_p \\ z^{-1}\mathbf{I}_p & \mathbf{I}_p & \cdots & \mathbf{0}_p \\ \vdots & \ddots & \ddots & \vdots \\ z^{-(m-1)}\mathbf{I}_p & z^{-(m-2)}\mathbf{I}_p & \cdots & \mathbf{I}_p \end{bmatrix}
\end{aligned}
\tag{2.5.57}
$$

So that we can get the transfer function matrix as following.

$$
\mathbf{T}(z)=\mathbf{I}_p+z^{-1}\begin{bmatrix}\mathbf{0}_p & \cdots & \mathbf{0}_p & \mathbf{I}_p\end{bmatrix}\begin{bmatrix}\mathbf{I}_p & \mathbf{0}_p & \cdots & \mathbf{0}_p\\ z^{-1}\mathbf{I}_p & \mathbf{I}_p & \cdots & \mathbf{0}_p\\ \vdots & \ddots & \ddots & \vdots\\ z^{-(m-1)}\mathbf{I}_p & z^{-(m-2)}\mathbf{I}_p & \cdots & \mathbf{I}_p\end{bmatrix}\begin{bmatrix}\mathbf{L}_m\\ \mathbf{L}_{m-1}\\ \vdots\\ \mathbf{L}_1\end{bmatrix}
$$

$$
=\mathbf{I}_p+z^{-1}\begin{bmatrix}z^{-(m-1)}\mathbf{I}_p & \cdots & z^{-1}\mathbf{I}_p & \mathbf{I}_p\end{bmatrix}\begin{bmatrix}\mathbf{L}_m\\ \mathbf{L}_{m-1}\\ \vdots\\ \mathbf{L}_1\end{bmatrix}
$$

$$
=\mathbf{I}_p+\begin{bmatrix}z^{-m}\mathbf{I}_p & \cdots & z^{-2}\mathbf{I}_p & z^{-1}\mathbf{I}_p\end{bmatrix}\begin{bmatrix}\mathbf{L}_m\\ \mathbf{L}_{m-1}\\ \vdots\\ \mathbf{L}_1\end{bmatrix}
$$

$$
=\mathbf{I}_p+\mathbf{L}_1z^{-1}+\cdots+\mathbf{L}_mz^{-m}
$$

$$
=\mathbf{I}_p+\sum_{i=1}^{m}\mathbf{L}_iz^{-i}=\mathbf{L}(z) \tag{2.5.58}
$$

Express $E[\mathbf{x}(n+1)\mathbf{y}^{*T}(n)]$ by spectral matrix

The second way to evaluate $E[\mathbf{x}(n+1)\mathbf{y}^{*T}(n)]$ is the following. As the first step, by write $\mathbf{x}(n)$ in block for each channel as following,

$$
\mathbf{x}(n)=\begin{bmatrix}x_{p1}(n)\\ x_{p2}(n)\\ \vdots\\ x_{pm}(n)\end{bmatrix} \tag{2.5.59}
$$

Put the $\mathbf{x}(n)$ into output equation of (3.1.8), we can get

$$
\begin{aligned}
\mathbf{y}(n)&=\mathbf{H}\mathbf{x}(n)+\mathbf{u}(n)\\
&=\begin{bmatrix}\mathbf{0}_p & \cdots & \mathbf{0}_p & \mathbf{I}_p\end{bmatrix}\begin{bmatrix}\mathbf{x}_{p1}(n)\\ \mathbf{x}_{p2}(n)\\ \vdots\\ \mathbf{x}_{pm}(n)\end{bmatrix}+\mathbf{u}(n)\\
&=\mathbf{x}_{pm}(n)+\mathbf{u}(n)
\end{aligned} \tag{2.5.60}
$$

so that

$$\mathbf{y}(n+1) = \mathbf{x}_{pm}(n+1) + \mathbf{u}(n+1) \tag{2.5.61}$$

therefore

$$\mathbf{x}_{pm}(n+1) = \mathbf{y}(n+1) - \mathbf{u}(n+1) \tag{2.5.62}$$

So we can conclude that

$$\begin{aligned} &E[\mathbf{x}_{pm}(n+1)\mathbf{y}^{*T}(n)] = E[(\mathbf{y}(n+1) - \mathbf{u}(n+1))\mathbf{y}^{*T}(n)] \\ &= E[\mathbf{y}(n+1)\mathbf{y}^{*T}(n)] = \mathbf{R}_1 \end{aligned} \tag{2.5.63}$$

where we used (2.2.2).

As the second step, let's consider $E[\mathbf{x}_{pm}(n+2)\mathbf{y}^{*T}(n)]$ now. From state space model in (2.5.8) we can get

$$\mathbf{x}(n+2) = \mathbf{F}\mathbf{x}(n+1) + \mathbf{G}\mathbf{u}(n+1) \tag{2.5.64}$$

Write it in block component, then we can get

$$\begin{aligned} &\begin{bmatrix} \mathbf{x}_{p1}(n+2) \\ \mathbf{x}_{p2}(n+2) \\ \vdots \\ \mathbf{x}_{pm}(n+2) \end{bmatrix} = \begin{bmatrix} \mathbf{0}_p & & & \\ \mathbf{I}_p & \mathbf{0}_p & & \\ & \ddots & \ddots & \\ & & \mathbf{I}_p & \mathbf{0}_p \end{bmatrix} \begin{bmatrix} \mathbf{x}_{p1}(n+1) \\ \mathbf{x}_{p2}(n+1) \\ \vdots \\ \mathbf{x}_{pm}(n+1) \end{bmatrix} + \begin{bmatrix} \mathbf{L}_m \\ \mathbf{L}_{m-1} \\ \vdots \\ \mathbf{L}_1 \end{bmatrix} \mathbf{u}(n+1) \\ &= \begin{bmatrix} \mathbf{0}_p \\ \mathbf{x}_{p1}(n+1) \\ \vdots \\ \mathbf{x}_{p(m-1)}(n+1) \end{bmatrix} + \begin{bmatrix} \mathbf{L}_m\mathbf{u}(n+1) \\ \mathbf{L}_{m-1}\mathbf{u}(n+1) \\ \vdots \\ \mathbf{L}_1\mathbf{u}(n+1) \end{bmatrix} \end{aligned} \tag{2.5.65}$$

then

$$\mathbf{x}_{pm}(n+2) = \mathbf{x}_{pm}(n+1) + \mathbf{L}_1\mathbf{u}(n+1) \tag{2.5.66}$$

so that

$$\begin{aligned} &\mathbf{y}(n+2) = \mathbf{x}_{pm}(n+2) + \mathbf{u}(n+2) \\ &= \mathbf{x}_{p(m-1)}(n+1) + \mathbf{L}_1\mathbf{u}(n+1) + \mathbf{u}(n+2) \end{aligned} \tag{2.5.67}$$

therefore

$$\mathbf{x}_{p(m-1)}(n+1) = \mathbf{y}(n+2) - \mathbf{L}_1\mathbf{u}(n+1) - \mathbf{u}(n+2) \tag{2.5.68}$$

So we can conclude that

$$E[\mathbf{x}_{p(m-1)}(n+1)\mathbf{y}^{*T}(n)] = E[(\mathbf{y}(n+2) - \mathbf{L}_1\mathbf{u}(n+1) - \mathbf{u}(n+2))\mathbf{y}^{*T}(n)]$$
$$= E[\mathbf{y}(n+2)\mathbf{y}^{*T}(n)] = \mathbf{R}_2 \tag{2.5.69}$$

By continuing this argument we easily verify that

$$E[\mathbf{x}(n+1)\mathbf{y}^{*T}(n)] = E[\begin{bmatrix} \mathbf{x}_{p1}(n+1) \\ \mathbf{x}_{p2}(n+1) \\ \vdots \\ \mathbf{x}_{pm}(n+1) \end{bmatrix}\mathbf{y}^{*T}(n)]$$
$$= \begin{bmatrix} E[\mathbf{x}_{p1}(n+1)\mathbf{y}^{*T}(n)] \\ E[\mathbf{x}_{p2}(n+1)\mathbf{y}^{*T}(n)] \\ \vdots \\ E[\mathbf{x}_{pm}(n+1)\mathbf{y}^{*T}(n)] \end{bmatrix} = \begin{bmatrix} \mathbf{R}_m \\ \mathbf{R}_{m-1} \\ \vdots \\ \mathbf{R}_1 \end{bmatrix} \tag{2.5.70}$$

We will denote this $mp \times p$ matrix by $\mathbf{N}$, that is

$$\mathbf{N} = \begin{bmatrix} \mathbf{R}_m \\ \mathbf{R}_{m-1} \\ \vdots \\ \mathbf{R}_1 \end{bmatrix} \tag{2.5.71}$$

Chapter 3 MULTICHANNEL SPECTRAL ESTIMATION

In this chapter, after an overview of several model for spectral estimation, the problem is introduced of multichannel spectral estimation using autoregressive (AR) model. This results in a set of normal equation, which can be solved recursively using a Yule-Walker algorithm. Flow chart of the algorithm is also shown for realizing the calculation procedure, together with a numerical example.

3.1 Spectral Estimation Method Overview

The various methods of spectral estimation can be categorized into nonparametric methods and parametric methods. Nonparametric methods are those in which the estimate of the power spectral density (PSD) is made directly from the signal itself. The simplest such method is the periodogram. An improved version of the periodogram is Welch's method. Parametric methods are those in those the signal whose PSD we want to estimate is assumed to be output of a linear system driven by white noise. Examples are Yule-Walker autoregressive (AR) method and Burg method. These method estimate the PSD by first estimating the parameters (coefficients) of linear system that hypothetically "generates" the signal. They tend to produce better results than classical nonparametric methods when the data length of the available signal is relative short.

The most commonly used linear system model is the all-pole model, a filter with all of zeros at the origin in the z-plane. The output of such a filter for white noise input is an autoregressive (AR) process. For this reason, these methods are sometimes referred to as AR methods of spectral estimation. The AR methods tend to adequately describe spectra of data that is "peak", that is, data whose PSD is large at certain frequencies. The data in many practical applications (such as speech) tends to have "peaky spectra" so that AR models are often useful. In addition, the AR models lead to a system of linear equations which is relatively simple to solve.

In many practical situations the data that are available are not limited to the output of a single channel but may well be the result of observations at the output of several channels. It is quite common in the fields of sonar, radar and seismic exploration to record data from multiple sensors. With this additional information it is then possible to estimate cross-spectra as well as auto-spectra. The cross-spectra are important in establishing linear filtering relationships between the time series. The multichannel spectral estimation problem is to estimate the auto-spectra of individual channels and the cross-spectra between all pairs of channels. For a given complex multichannel sequence

$$\mathbf{x}(n) = \begin{bmatrix} x_1(n) \\ x_2(n) \\ \vdots \\ x_p(n) \end{bmatrix} \tag{3.1.1}$$

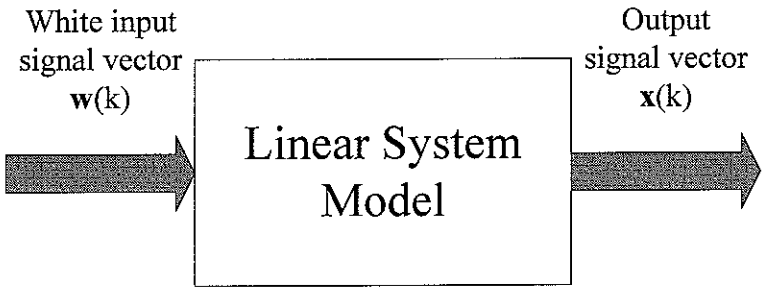
Typically, $x_i(n)$ represents the data observed at output of the ith channel. The power spectral density or cross-spectral matrix is defined as

$$\mathbf{S}(f) = \begin{bmatrix} S_{11}(f) & S_{12}(f) & \cdots & S_{1p}(f) \\ S_{21}(f) & S_{22}(f) & \cdots & S_{2p}(f) \\ \vdots & \vdots & \ddots & \vdots \\ S_{p1}(f) & S_{p2}(f) & \cdots & S_{pp}(f) \end{bmatrix} \tag{3.1.2}$$

The diagonal elements $S_{ii}(f)$ are the PSDs of the individual channels or auto-PSDs, while the off-diagonal elements $S_{ij}(f)$ for $i \neq j$ are the cross-PSDs between $x_i(n)$ and $x_j(n)$.

3.2 Spectral Estimation via Multichannel AR Model

As we illustrated in section 2.5, the vector of random processes, we want estimate which cross-spectral density matrix, can be assumed as a output of a multichannel linear system model with white input random signal vector as shown in Figure 8, which is same graph with Figure 6, but with different notation for input signal vector and output signal vector.



• Figure 8 Linear system model for multichannel spectral estimation

When a white input signal vector $\mathbf{u}(z)$ is applied to this system, the output signal vector can be given as following.

$$\mathbf{y}(z) = \mathbf{L}(z)\mathbf{u}(z)$$

Therefore by select a linear system model available to measured data vector, a modelling method of random processes can be obtained. Here we assume a autoregressive (AR) model as following.

$$\mathbf{D}(z) = \mathbf{I}_p + \sum_{i=1}^{m} \mathbf{A}(i) z^{-i}$$

$$\mathbf{L}(z) = \mathbf{D}^{-1}(z) = \left(\mathbf{I}_p + \sum_{i=1}^{m} \mathbf{A}(i) z^{-i} \right)^{-1} \qquad (3.2.1)$$

By adopting a vector autoregressive (AR) model (3.2.1), vector $\mathbf{x}(n)$ is assumed to evolve according to the mth-order autoregressive

$$\mathbf{x}(n) = -\sum_{k=1}^{m} \mathbf{A}(k)\mathbf{x}(n-k) + \mathbf{w}(n) \qquad (3.2.2)$$

where

$$\mathbf{A}(k) = \begin{bmatrix} a_{11}(k) & a_{12}(k) & \cdots & a_{1p}(k) \\ a_{21}(k) & a_{22}(k) & \cdots & a_{2p}(k) \\ \vdots & \vdots & \ddots & \vdots \\ a_{p1}(k) & a_{p2}(k) & \cdots & a_{pp}(k) \end{bmatrix} \quad k = 1, 2, \cdots, m \qquad (3.2.3)$$

are $p \times p$ matrices of multichannel AR model coefficients. The driving input signal

$$\mathbf{w}(n) = \begin{bmatrix} w_1(n) \\ w_2(n) \\ \vdots \\ w_p(n) \end{bmatrix} \tag{3.2.4}$$

is a set of uncorrelated white sequences with zero mean and so the covariance matrix

$$\boldsymbol{\Sigma} = E\left[\mathbf{w}(n)\mathbf{w}(n)^T\right] \tag{3.2.5}$$

is equal to the identity matrix. Our problem is described as following.

For a given measured data of vector random processes, how can we decide it's coefficients matrix $\mathbf{A}(k)$, $k = 1,2,\cdots,m$ of AR model, So that we can obtain an estimation of cross-spectral density matrix.

3.3 Multichannel Yule-Walker Algorithm

3.3.1 Multichannel Yule-Walker Equations

S. M. Key[8] has given a detail description of determining the optimal *m*th order forward and backward linear predictor for an AR(m) process.

The *m*th order forward predictor is

$$\hat{\mathbf{x}}^f(n) = -\sum_{i=1}^{m} \mathbf{A}^f(i)\mathbf{x}(n-i) \tag{3.3.1)}$$

The predictor error power is defined as the sum of prediction error powers for individual channels as following.

$$\rho^f = E\left[\left(\mathbf{x}(n) - \hat{\mathbf{x}}^f(n)\right)^T \left(\mathbf{x}(n) - \hat{\mathbf{x}}^f(n)\right)^*\right] \tag{3.3.2}$$

Alternatively, ρ^f may be view as the sum of diagonal elements or trace of the covariance matrix

$$\boldsymbol{\Sigma}^f = E\left[\left(\mathbf{x}(n) - \hat{\mathbf{x}}^f(n)\right)^* \left(\mathbf{x}(n) - \hat{\mathbf{x}}^f(n)\right)^T\right] \tag{3.3.3}$$

To minimize ρ^f, we can use the orthogonality principle for a vector space, then yields

$$E\left[\mathbf{x}^*(n-k)\left(\mathbf{x}(n)-\hat{\mathbf{x}}^f(n)\right)^T\right]=\mathbf{0} \quad k=1,2,\ldots,m \tag{3.3.4}$$

which results in the multichannel Wiener-Hopf equations

$$\mathbf{R}_{xx}(k)=-\sum_{i=1}^{m}\mathbf{R}_{xx}(k-i)\mathbf{A}^{fT}(i) \quad k=1,2,\ldots,m \tag{3.3.5}$$

The prediction error power matrix is

$$\begin{aligned}\mathbf{\Sigma}^f &= E\left[\left(\mathbf{x}(n)-\hat{\mathbf{x}}^f(n)\right)^*\left(\mathbf{x}(n)-\hat{\mathbf{x}}^f(n)\right)^T\right]\\ &=\mathbf{R}_{xx}(0)+\sum_{i=1}^{m}\mathbf{R}_{xx}(-i)\mathbf{A}^{fT}(i)\end{aligned} \tag{3.3.6}$$

As expected, $\mathbf{A}^f(i)=\mathbf{A}(i)$ and $\mathbf{\Sigma}^f=\mathbf{\Sigma}$ or the prediction coefficients are given by the AR filter parameters and the prediction error power matrix is given by the variance matrix of the white noise.

The mth order forward predictor is

$$\hat{\mathbf{x}}^b(n)=-\sum_{i=1}^{m}\mathbf{A}^b(i)\mathbf{x}(n+i) \tag{3.3.7}$$

The corresponding predictor error power is given as following.

$$\rho^b = E\left[\left(\mathbf{x}(n)-\hat{\mathbf{x}}^b(n)\right)^T\left(\mathbf{x}(n)-\hat{\mathbf{x}}^b(n)\right)^*\right] \tag{3.3.8}$$

and the prediction error power matrix is

$$\mathbf{\Sigma}^b = E\left[\left(\mathbf{x}(n)-\hat{\mathbf{x}}^b(n)\right)^*\left(\mathbf{x}(n)-\hat{\mathbf{x}}^b(n)\right)^T\right] \tag{3.3.9}$$

Using a similar development the Wiener-Hopf equation become

$$\mathbf{R}_{xx}(-k) = -\sum_{i=1}^{m} \mathbf{R}_{xx}(-k-i)\mathbf{A}^{bT}(i) \quad k = 1,2,\ldots,p \tag{3.3.10}$$

The prediction error power matrix is

$$\begin{aligned} \boldsymbol{\Sigma}^{b} &= E\left[\left(\mathbf{x}(n) - \hat{\mathbf{x}}^{b}(n)\right)^{*}\left(\mathbf{x}(n) - \hat{\mathbf{x}}^{b}(n)\right)^{T}\right] \\ &= \mathbf{R}_{xx}(0) + \sum_{i=1}^{m} \mathbf{R}_{xx}(i)\mathbf{A}^{bT}(i) \end{aligned} \tag{3.3.11}$$

From (3.3.5) (3.3.6) m-1 order forward predictors satisfy the equations

$$\begin{bmatrix} \mathbf{R}_{xx}(0) & \mathbf{R}_{xx}(-1) & \cdots & \mathbf{R}_{xx}(-(m-1)) \\ \mathbf{R}_{xx}(1) & \mathbf{R}_{xx}(0) & \cdots & \mathbf{R}_{xx}(-(m-2)) \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{R}_{xx}(m-1) & \mathbf{R}_{xx}(m-2) & \cdots & \mathbf{R}_{xx}(0) \end{bmatrix} \begin{bmatrix} \mathbf{I} \\ \mathbf{A}_{m-1}^{fT}(1) \\ \vdots \\ \mathbf{A}_{m-1}^{fT}(m-1) \end{bmatrix} = \begin{bmatrix} \Sigma_{m-1}^{f} \\ \mathbf{0} \\ \vdots \\ \mathbf{0} \end{bmatrix} \tag{3.3.12}$$

From equation (3.3.10) (3.3.11) m order backward predictors satisfy the equations

$$\begin{bmatrix} \mathbf{R}_{xx}(0) & \mathbf{R}_{xx}(-1) & \cdots & \mathbf{R}_{xx}(-(m-1)) \\ \mathbf{R}_{xx}(1) & \mathbf{R}_{xx}(0) & \cdots & \mathbf{R}_{xx}(-(m-2)) \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{R}_{xx}(m-1) & \mathbf{R}_{xx}(m-2) & \cdots & \mathbf{R}_{xx}(0) \end{bmatrix} \begin{bmatrix} \mathbf{A}_{m-1}^{bT}(m-1) \\ \mathbf{A}_{m-1}^{bT}(m-2) \\ \vdots \\ \mathbf{I} \end{bmatrix} = \begin{bmatrix} \mathbf{0} \\ \mathbf{0} \\ \vdots \\ \Sigma_{m-1}^{b} \end{bmatrix} \tag{3.3.13}$$

3.3.2 Derivation of Levinson Algorithm for Multichannel Yule-Walker Equations

It is assumed the solution of m-1 order forward predictors (3.3.12)

$$\mathbf{A}_{m-1}^{f}(i), i = 1,2,\ldots,m-1$$

are already available. By increasing

$$\Delta_{m}^{f} = \sum_{i=0}^{m-1} \mathbf{R}_{xx}(m-i)\mathbf{A}_{m-1}^{fT}(i) \text{ with } \mathbf{A}_{m-1}^{fT}(0) = \mathbf{I} \tag{3.3.14}$$

as last row to (3.3.12), we can get

$$\begin{bmatrix} \mathbf{R}_{xx}(0) & \mathbf{R}_{xx}(-1) & \cdots & \mathbf{R}_{xx}(-(m-1)) & \mathbf{R}_{xx}(-m) \\ \mathbf{R}_{xx}(1) & \mathbf{R}_{xx}(2) & \cdots & \mathbf{R}_{xx}(-(m-2)) & \mathbf{R}_{xx}(-(m-1)) \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \mathbf{R}_{xx}(m-1) & \mathbf{R}_{xx}(m-2) & \cdots & \mathbf{R}_{xx}(0) & \mathbf{R}_{xx}(-1) \\ \mathbf{R}_{xx}(m) & \mathbf{R}_{xx}(m-1) & \cdots & \mathbf{R}_{xx}(1) & \mathbf{R}_{xx}(0) \end{bmatrix} \begin{bmatrix} \mathbf{I} \\ \mathbf{A}_{m-1}^{fT}(1) \\ \vdots \\ \mathbf{A}_{m-1}^{fT}(m-1) \\ \mathbf{0} \end{bmatrix} = \begin{bmatrix} \Sigma_{m-1}^{f} \\ \mathbf{0} \\ \vdots \\ \mathbf{0} \\ \Delta_{m}^{f} \end{bmatrix} \tag{3.3.15}$$

Also it is assumed the solution of m-1 order forward predictors (3.3.13)

$$\mathbf{A}_{m-1}^{b}(i),\, i=1,2,...,m-1$$

are already available. By increasing

$$\Delta_m^b = \sum_{i=0}^{m-1} \mathbf{R}_{xx}(i-k)\mathbf{A}_{m-1}^{bT}(i) \text{ with } \mathbf{A}_{m-1}^{bT}(0) = \mathbf{I} \tag{3.3.16}$$

as first row to (3.3.13), we can get

$$\begin{bmatrix} \mathbf{R}_{xx}(0) & \mathbf{R}_{xx}(-1) & \cdots & \mathbf{R}_{xx}(-(m-1)) & \mathbf{R}_{xx}(-m) \\ \mathbf{R}_{xx}(1) & \mathbf{R}_{xx}(2) & \cdots & \mathbf{R}_{xx}(-(m-2)) & \mathbf{R}_{xx}(-(m-1)) \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \mathbf{R}_{xx}(m-1) & \mathbf{R}_{xx}(m-2) & \cdots & \mathbf{R}_{xx}(0) & \mathbf{R}_{xx}(-1) \\ \mathbf{R}_{xx}(m) & \mathbf{R}_{xx}(m-1) & \cdots & \mathbf{R}_{xx}(1) & \mathbf{R}_{xx}(0) \end{bmatrix} \begin{bmatrix} \mathbf{0} \\ \mathbf{A}_{m-1}^{bT}(m-1) \\ \vdots \\ \mathbf{A}_{m-1}^{bT}(1) \\ \mathbf{I} \end{bmatrix} = \begin{bmatrix} \Delta_{m}^{b} \\ \mathbf{0} \\ \vdots \\ \mathbf{0} \\ \Sigma_{m-1}^{b} \end{bmatrix} \tag{3.3.17}$$

Multiply (3.3.17) by the $L \times L$ matrix $\mathbf{K}_m^{fT}$, then add it to (3.3.15),

$$\begin{bmatrix} \mathbf{R}_{xx}(0) & \mathbf{R}_{xx}(-1) & \cdots & \mathbf{R}_{xx}(-(m-1)) & \mathbf{R}_{xx}(-m) \\ \mathbf{R}_{xx}(1) & \mathbf{R}_{xx}(2) & \cdots & \mathbf{R}_{xx}(-(m-2)) & \mathbf{R}_{xx}(-(m-1)) \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \mathbf{R}_{xx}(m-1) & \mathbf{R}_{xx}(m-2) & \cdots & \mathbf{R}_{xx}(0) & \mathbf{R}_{xx}(-1) \\ \mathbf{R}_{xx}(m) & \mathbf{R}_{xx}(m-1) & \cdots & \mathbf{R}_{xx}(1) & \mathbf{R}_{xx}(0) \end{bmatrix} \left\{ \begin{bmatrix} \mathbf{I} \\ \mathbf{A}_{m-1}^{fT}(1) \\ \vdots \\ \mathbf{A}_{m-1}^{fT}(m-1) \\ \mathbf{0} \end{bmatrix} + \begin{bmatrix} \mathbf{0} \\ \mathbf{A}_{m-1}^{bT}(m-1) \\ \vdots \\ \mathbf{A}_{m-1}^{bT}(1) \\ \mathbf{I} \end{bmatrix} \mathbf{K}_m^{fT} \right\} = \begin{bmatrix} \Sigma_{m-1}^{f} \\ \mathbf{0} \\ \vdots \\ \mathbf{0} \\ \Delta_m^{f} \end{bmatrix} + \begin{bmatrix} \Delta_m^{b} \\ \mathbf{0} \\ \vdots \\ \mathbf{0} \\ \Sigma_{m-1}^{b} \end{bmatrix} \mathbf{K}_m^{fT} \tag{3.3.18a}$$

$$\begin{bmatrix} \mathbf{R}_{xx}(0) & \mathbf{R}_{xx}(-1) & \cdots & \mathbf{R}_{xx}(-(m-1)) & \mathbf{R}_{xx}(-m) \\ \mathbf{R}_{xx}(1) & \mathbf{R}_{xx}(2) & \cdots & \mathbf{R}_{xx}(-(m-2)) & \mathbf{R}_{xx}(-(m-1)) \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \mathbf{R}_{xx}(m-1) & \mathbf{R}_{xx}(m-2) & \cdots & \mathbf{R}_{xx}(0) & \mathbf{R}_{xx}(-1) \\ \mathbf{R}_{xx}(m) & \mathbf{R}_{xx}(m-1) & \cdots & \mathbf{R}_{xx}(1) & \mathbf{R}_{xx}(0) \end{bmatrix} \begin{bmatrix} \mathbf{I} \\ \mathbf{A}_{m-1}^{fT}(1) + \mathbf{A}_{m-1}^{bT}(m-1)\mathbf{K}_m^{fT} \\ \vdots \\ \mathbf{A}_{m-1}^{fT}(m-1) + \mathbf{A}_{m-1}^{bT}(1)\mathbf{K}_m^{fT} \\ \mathbf{K}_m^{fT} \end{bmatrix} = \begin{bmatrix} \Sigma_{m-1}^{f} + \Delta_m^{b}\mathbf{K}_m^{fT} \\ \mathbf{0} \\ \vdots \\ \mathbf{0} \\ \Delta_m^{f} + \Sigma_{m-1}^{b}\mathbf{K}_m^{fT} \end{bmatrix} \tag{3.3.18b}$$

compare this equation with mth order forward predictor,

$$\begin{bmatrix} \mathbf{R}_{xx}(0) & \mathbf{R}_{xx}(-1) & \cdots & \mathbf{R}_{xx}(-(m-1)) & \mathbf{R}_{xx}(-m) \\ \mathbf{R}_{xx}(1) & \mathbf{R}_{xx}(2) & \cdots & \mathbf{R}_{xx}(-(m-2)) & \mathbf{R}_{xx}(-(m-1)) \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \mathbf{R}_{xx}(m-1) & \mathbf{R}_{xx}(m-2) & \cdots & \mathbf{R}_{xx}(0) & \mathbf{R}_{xx}(-1) \\ \mathbf{R}_{xx}(m) & \mathbf{R}_{xx}(m-1) & \cdots & \mathbf{R}_{xx}(1) & \mathbf{R}_{xx}(0) \end{bmatrix} \begin{bmatrix} \mathbf{I} \\ \mathbf{A}_m^{fT}(1) \\ \vdots \\ \mathbf{A}_m^{fT}(m-1) \\ \mathbf{A}_m^{fT}(m) \end{bmatrix} = \begin{bmatrix} \mathbf{\Sigma}_m^{f} \\ \mathbf{0} \\ \vdots \\ \mathbf{0} \\ \mathbf{0} \end{bmatrix} \tag{3.3.19}$$

it must be true that the last equation on right hand gives

$$\Delta_m^{f} = -\mathbf{\Sigma}_{m-1}^{f}\mathbf{K}_m^{fT} \tag{3.3.20a}$$

$$\mathbf{K}_m^{fT} = -\left(\mathbf{\Sigma}_{m-1}^{b}\right)^{-1}\mathbf{\Delta}_m^{f} \tag{3.3.21b}$$

Also the first equation results in

$$\mathbf{\Sigma}_m^{f} = \mathbf{\Sigma}_{m-1}^{f} + \mathbf{\Delta}_m^{b}\mathbf{K}_m^{fT} \tag{3.3.22}$$

and the solution must satisfy

$$\mathbf{A}_m^{f}(i) = \begin{cases} \mathbf{A}_{m-1}^{f}(i) + \mathbf{K}_m^{f}\mathbf{A}_{m-1}^{b}(m-i) & i = 1,2,\ldots,m-1 \\ \mathbf{K}_m^{f} \quad i = m \end{cases} \tag{3.3.23}$$

In a similar procedure we can generate the equations

$$\begin{bmatrix} \mathbf{R}_{xx}(0) & \mathbf{R}_{xx}(-1) & \cdots & \mathbf{R}_{xx}(-(m-1)) & \mathbf{R}_{xx}(-m) \\ \mathbf{R}_{xx}(1) & \mathbf{R}_{xx}(2) & \cdots & \mathbf{R}_{xx}(-(m-2)) & \mathbf{R}_{xx}(-(m-1)) \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \mathbf{R}_{xx}(m-1) & \mathbf{R}_{xx}(m-2) & \cdots & \mathbf{R}_{xx}(0) & \mathbf{R}_{xx}(-1) \\ \mathbf{R}_{xx}(m) & \mathbf{R}_{xx}(m-1) & \cdots & \mathbf{R}_{xx}(1) & \mathbf{R}_{xx}(0) \end{bmatrix} \left\{ \begin{bmatrix} \mathbf{I} \\ \mathbf{A}_{m-1}^{fT}(1) \\ \vdots \\ \mathbf{A}_{m-1}^{fT}(m-1) \\ \mathbf{0} \end{bmatrix} \mathbf{K}_m^{fT} + \begin{bmatrix} \mathbf{0} \\ \mathbf{A}_{m-1}^{bT}(m-1) \\ \vdots \\ \mathbf{A}_{m-1}^{bT}(1) \\ \mathbf{I} \end{bmatrix} \right\} = \begin{bmatrix} \Sigma_{m-1}^{f} \\ \mathbf{0} \\ \vdots \\ \mathbf{0} \\ \mathbf{\Delta}_m^{f} \end{bmatrix} \mathbf{K}_m^{fT} + \begin{bmatrix} \mathbf{\Delta}_m^{b} \\ \mathbf{0} \\ \vdots \\ \mathbf{0} \\ \Sigma_{m-1}^{b} \end{bmatrix} \tag{3.3.24a}$$

$$\begin{bmatrix} \mathbf{R}_{xx}(0) & \mathbf{R}_{xx}(-1) & \cdots & \mathbf{R}_{xx}(-(m-1)) & \mathbf{R}_{xx}(-m) \\ \mathbf{R}_{xx}(1) & \mathbf{R}_{xx}(2) & \cdots & \mathbf{R}_{xx}(-(m-2)) & \mathbf{R}_{xx}(-(m-1)) \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \mathbf{R}_{xx}(m-1) & \mathbf{R}_{xx}(m-2) & \cdots & \mathbf{R}_{xx}(0) & \mathbf{R}_{xx}(-1) \\ \mathbf{R}_{xx}(m) & \mathbf{R}_{xx}(m-1) & \cdots & \mathbf{R}_{xx}(1) & \mathbf{R}_{xx}(0) \end{bmatrix} \begin{bmatrix} \mathbf{K}_m^{fT} \\ \mathbf{A}_{m-1}^{fT}(1)\mathbf{K}_m^{fT} + \mathbf{A}_{m-1}^{bT}(m-1) \\ \vdots \\ \mathbf{A}_{m-1}^{fT}(m-1)\mathbf{K}_m^{fT} + \mathbf{A}_{m-1}^{bT}(1) \\ \mathbf{I} \end{bmatrix} = \begin{bmatrix} \Sigma_{m-1}^{f}\mathbf{K}_m^{fT} + \mathbf{\Delta}_m^{b} \\ \mathbf{0} \\ \vdots \\ \mathbf{0} \\ \mathbf{\Delta}_m^{f}\mathbf{K}_m^{fT} + \Sigma_{m-1}^{b} \end{bmatrix} \tag{3.3.24b}$$

which should represent the equations for the mth order backward predictor given as following.

$$\begin{bmatrix} \mathbf{R}_{xx}(0) & \mathbf{R}_{xx}(-1) & \cdots & \mathbf{R}_{xx}(-(m-1)) & \mathbf{R}_{xx}(-m) \\ \mathbf{R}_{xx}(1) & \mathbf{R}_{xx}(2) & \cdots & \mathbf{R}_{xx}(-(m-2)) & \mathbf{R}_{xx}(-(m-1)) \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \mathbf{R}_{xx}(m-1) & \mathbf{R}_{xx}(m-2) & \cdots & \mathbf{R}_{xx}(0) & \mathbf{R}_{xx}(-1) \\ \mathbf{R}_{xx}(m) & \mathbf{R}_{xx}(m-1) & \cdots & \mathbf{R}_{xx}(1) & \mathbf{R}_{xx}(0) \end{bmatrix} \begin{bmatrix} \mathbf{A}_m^{bT}(m) \\ \mathbf{A}_m^{bT}(m-1) \\ \vdots \\ \mathbf{A}_m^{bT}(1) \\ \mathbf{I} \end{bmatrix} = \begin{bmatrix} \mathbf{0} \\ \mathbf{0} \\ \vdots \\ \mathbf{0} \\ \mathbf{\Sigma}_m^b \end{bmatrix}$$

(3.3.25)

By the same argument as forward predictor, following results can be given.

$$\mathbf{\Delta}_m^b = -\mathbf{\Sigma}_{m-1}^f \mathbf{K}_m^{bT} \tag{3.3.26a}$$

$$\mathbf{K}_m^{bT} = -\left(\mathbf{\Sigma}_{m-1}^f\right)^{-1} \mathbf{\Delta}_m^b \tag{3.3.26b}$$

$$\mathbf{\Sigma}_m^b = \mathbf{\Sigma}_{m-1}^b + \mathbf{\Delta}_m^f \mathbf{K}_m^{bT} \tag{3.3.27}$$

$$\mathbf{A}_m^b(i) = \begin{cases} \mathbf{A}_{m-1}^b(i) + \mathbf{K}_m^b \mathbf{A}_{m-1}^f(m-i) & i = 1,2,\ldots,m-1 \\ \mathbf{K}_m^b & i = m \end{cases} \tag{3.3.28}$$

Because $\mathbf{\Delta}_m^b = \mathbf{\Delta}_m^{fH}$, so that (3.3.26b) becomes

$$\mathbf{K}_m^{bT} = -\left(\mathbf{\Sigma}_{m-1}^f\right)^{-1} \mathbf{\Delta}_m^{fH} \tag{3.3.29}$$

Also by put (3.3.26a) into (3.3.22), becomes

$$\mathbf{\Sigma}_m^f = \mathbf{\Sigma}_{m-1}^f - \mathbf{\Sigma}_{m-1}^f \mathbf{K}_m^{bT} \mathbf{K}_m^{fT} = \mathbf{\Sigma}_{m-1}^f \left(\mathbf{I} - \mathbf{K}_m^{bT} \mathbf{K}_m^{fT}\right) \tag{3.3.30}$$

and similarly by put (3.3.20a) into (3.3.27), becomes

$$\mathbf{\Sigma}_m^b = \mathbf{\Sigma}_{m-1}^b - \mathbf{\Sigma}_{m-1}^b \mathbf{K}_m^{fT} \mathbf{K}_m^{bT} = \mathbf{\Sigma}_{m-1}^b \left(\mathbf{I} - \mathbf{K}_m^{fT} \mathbf{K}_m^{bT}\right) \tag{3.3.31}$$

The algorithm is initialized by finding the solution for the first order linear predictors. From (3.3.12) and (3.3.23), the first order forward predictor coefficients are given by

$$\mathbf{A}_1^{fT}(1) = \mathbf{K}_1^{fT} = -(\mathbf{R}_{xx}(0))^{-1}\mathbf{R}_{xx}(1) \qquad (3.3.32)$$

and from (3.3.13) and (3.3.28), the first order backward predictor coefficients are given by

$$\mathbf{A}_1^{bT}(1) = \mathbf{K}_1^{bT} = -(\mathbf{R}_{xx}(0))^{-1}\mathbf{R}_{xx}(-1) \qquad (3.3.33)$$

Also from (3.3.12), the first order forward prediction error power matrix can be given as following.

$$\begin{aligned}\mathbf{\Sigma}_1^{f} &= \mathbf{R}_{xx}(0) + \mathbf{R}_{xx}(1)\mathbf{A}_1^{fT}(1) \\ &= \mathbf{R}_{xx}(0) - \mathbf{R}_{xx}(0)\mathbf{A}_1^{bT}(1)\mathbf{A}_1^{fT}(1) \\ &= \mathbf{R}_{xx}(0)\left(\mathbf{I} - \mathbf{K}_1^{bT}\mathbf{K}_1^{fT}\right)\end{aligned} \qquad (3.3.34)$$

Similarly, from (3.3.13), the first order backward prediction error power matrix can be given as following.,

$$\begin{aligned}\mathbf{\Sigma}_1^{b} &= \mathbf{R}_{xx}(0) + \mathbf{R}_{xx}(1)\mathbf{A}_1^{bT}(1) \\ &= \mathbf{R}_{xx}(0) - \mathbf{R}_{xx}(0)\mathbf{A}_1^{fT}(1)\mathbf{A}_1^{bT}(1) \\ &= \mathbf{R}_{xx}(0)\left(\mathbf{I} - \mathbf{K}_1^{fT}\mathbf{K}_1^{bT}\right)\end{aligned} \qquad (3.3.35)$$

3.3.3 Summary of Levinson Algorithm for Multichannel Yule-Walker Equations

Initialization:

$$\mathbf{A}_1^{fT}(1) = \mathbf{K}_1^{fT} = -(\mathbf{R}_{xx}(0))^{-1}\mathbf{R}_{xx}(1)$$

$$\mathbf{A}_1^{bT}(1) = \mathbf{K}_1^{bT} = -(\mathbf{R}_{xx}(0))^{-1}\mathbf{R}_{xx}(-1)$$

$$\mathbf{\Sigma}_1^{f} = \mathbf{R}_{xx}(0)\left(\mathbf{I} - \mathbf{K}_1^{bT}\mathbf{K}_1^{fT}\right)$$

$$\Sigma_1^b = \mathbf{R}_{xx}(0)\left(\mathbf{I} - \mathbf{K}_1^{fT}\mathbf{K}_1^{bT}\right)$$

Iteration for k=2,3,…,m

Reflection coefficients matrices:

$$\Delta_k^f = \sum_{i=0}^{m-1}\mathbf{R}_{xx}(k-i)\mathbf{A}_{k-1}^{fT}(i) \text{ with } \mathbf{A}_{k-1}^{fT}(0) = \mathbf{I}$$

$$\mathbf{K}_k^{fT} = -\left(\Sigma_{k-1}^b\right)^{-1}\Delta_k^f$$

$$\mathbf{K}_k^{bT} = -\left(\Sigma_{k-1}^f\right)^{-1}\Delta_k^{fH}$$

Predictor coefficient matrices:

$$\mathbf{A}_k^f(i) = \begin{cases}\mathbf{A}_{k-1}^f(i) + \mathbf{K}_k^f\mathbf{A}_{k-1}^b(k-i) & i = 1,2,\ldots,k-1 \\ \mathbf{K}_k^f & i = k\end{cases}$$

$$\mathbf{A}_k^b(i) = \begin{cases}\mathbf{A}_{k-1}^b(i) + \mathbf{K}_k^b\mathbf{A}_{k-1}^f(k-i) & i = 1,2,\ldots,k-1 \\ \mathbf{K}_k^b & i = k\end{cases}$$

Prediction error power matrices:

$$\Sigma_k^f = \Sigma_{k-1}^f\left(\mathbf{I} - \mathbf{K}_k^{bT}\mathbf{K}_k^{fT}\right)$$

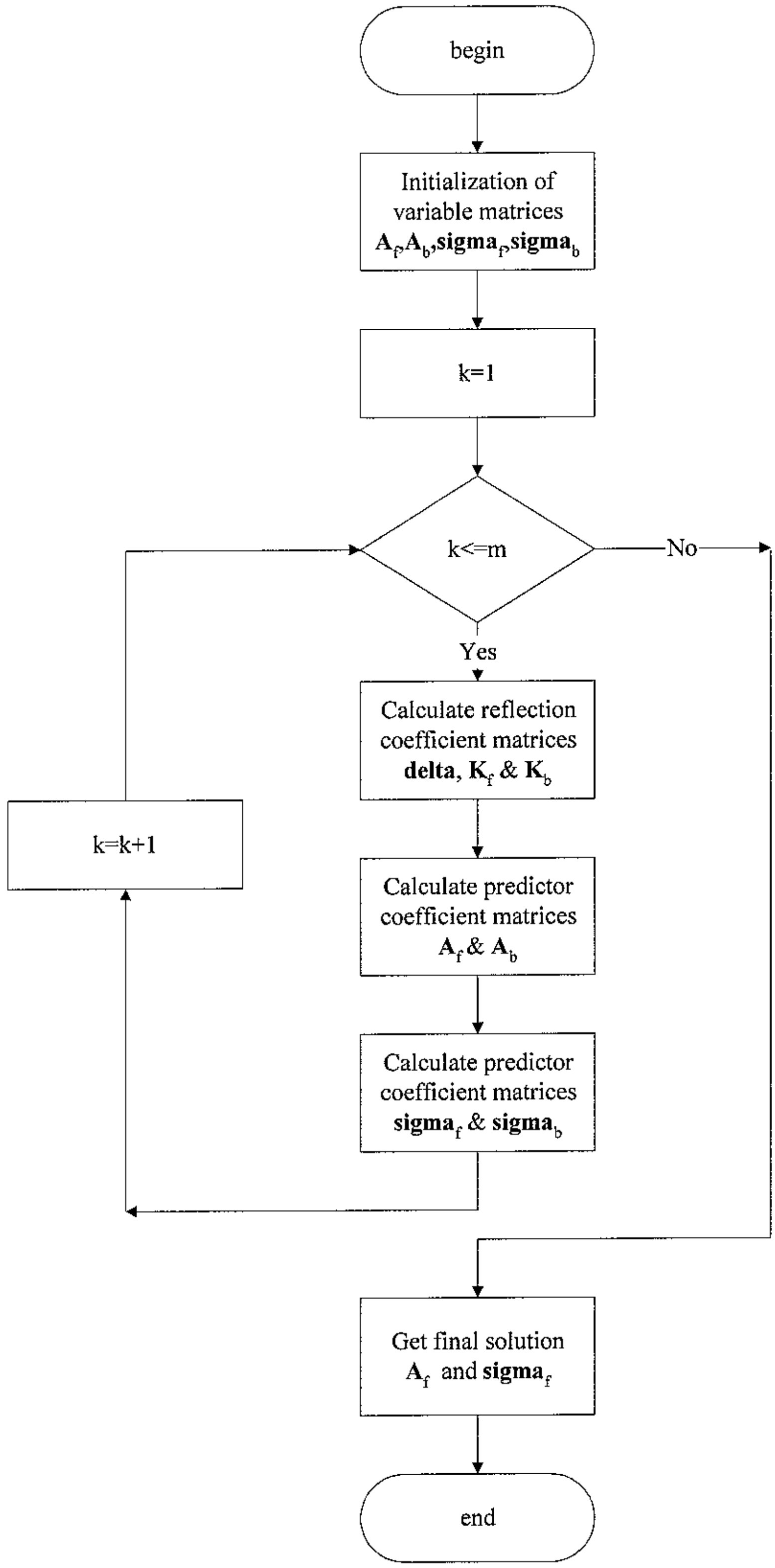
$$\Sigma_k^b = \Sigma_{k-1}^b\left(\mathbf{I} - \mathbf{K}_k^{fT}\mathbf{K}_k^{bT}\right)$$

The solution to the multichannel Yule-Walker equation is

$$\mathbf{A}(i) = \mathbf{A}_m^f(i) \quad i = 1,2,\ldots,m \tag{3.3.36}$$

$$\Sigma = \Sigma_m^f \tag{3.3.37}$$

3.3.4 Flow Chart



• Figure 9 Flow chart of multichannel Yule-Walk equation

3.3.5 Multichannel Spectral Estimation

For a given white noise vector with $p \times p$ covariance matrix $\mathbf{\Sigma}$, when apply it to a multichannel linear system with transfer function matrix $\mathbf{L}(z)$, The cross-spectral matrix of output vector is given as following.

$$\mathbf{S}(z) = \mathbf{L}(z)\mathbf{\Sigma}\mathbf{L}_{*}(z) \qquad (3.3.38)$$

For multichannel AR model,

$$\mathbf{L}(z) = \left(\mathbf{I}_p + \sum_{i=1}^{m} \mathbf{A}(i) z^{-i}\right)^{-1} \qquad (3.3.39)$$

where $\mathbf{A}(i), i = 1,2,\ldots,m$ are the results of multichannel Yule-Walker equation. By evaluating $\mathbf{S}(z)$ on the unit circle, i.e., $z = e^{j\omega T}$, the cross-spectral matrix is obtained as following

$$\mathbf{S}(\omega) = \mathbf{L}(\omega)\mathbf{\Sigma}\mathbf{L}_{*}(\omega) \qquad (3.3.40)$$

Furthermore, if we can get $\mathbf{\Sigma} = \mathbf{G}\mathbf{G}^T$, then transfer function can be given by $\mathbf{L}(\omega)\mathbf{G}$, we can get $\mathbf{S}(\omega)$ from p-channel white signals with covariance matrix $\mathbf{\Sigma} = \mathbf{I}_p$.

3.3.6 Computer Simulation Example

Example: The process that is to be considered is a real two channel AR(1) process with parameters of

$$\mathbf{A}(1) = \begin{bmatrix} -0.85 & 0.75 \\ -0.65 & -0.55 \end{bmatrix} \qquad (3.3.41)$$

$$\mathbf{\Sigma} = \mathbf{I} \qquad (3.3.42)$$

Results:

Using two set 10000 point white data as input signal, from measured output data of x1 and x2, by a MATLAB program, we can get results shown as following.

$$\mathbf{R}_{xx}(0)=\begin{pmatrix} r_{11}(0) & r_{12}(0) \\ r_{21}(0) & r_{22}(0) \end{pmatrix}=\begin{pmatrix} 23.1180 & 4.5267 \\ 4.5267 & 19.8562 \end{pmatrix} \tag{3.3.43}$$

$$\mathbf{R}_{xx}(1)=\begin{pmatrix} r_{11}(1) & r_{12}(1) \\ r_{21}(1) & r_{22}(1) \end{pmatrix}=\begin{pmatrix} 16.3072 & 17.4581 \\ -10.9915 & 13.8245 \end{pmatrix} \tag{3.3.44}$$

$$\mathbf{R}_{xx}(-1)=\begin{pmatrix} r_{11}(-1) & r_{12}(-1) \\ r_{21}(-1) & r_{22}(-1) \end{pmatrix}=\begin{pmatrix} 16.3072 & -10.9915 \\ 17.4581 & 13.8245 \end{pmatrix} \tag{3.3.45}$$

$$\mathbf{A}(1)=\left(-(\mathbf{R}_{xx}(0))^{-1}\mathbf{R}_{xx}(1)\right)^{T}=\begin{pmatrix} -0.8518 & 0.7477 \\ -0.6478 & -0.5486 \end{pmatrix} \tag{3.3.46}$$

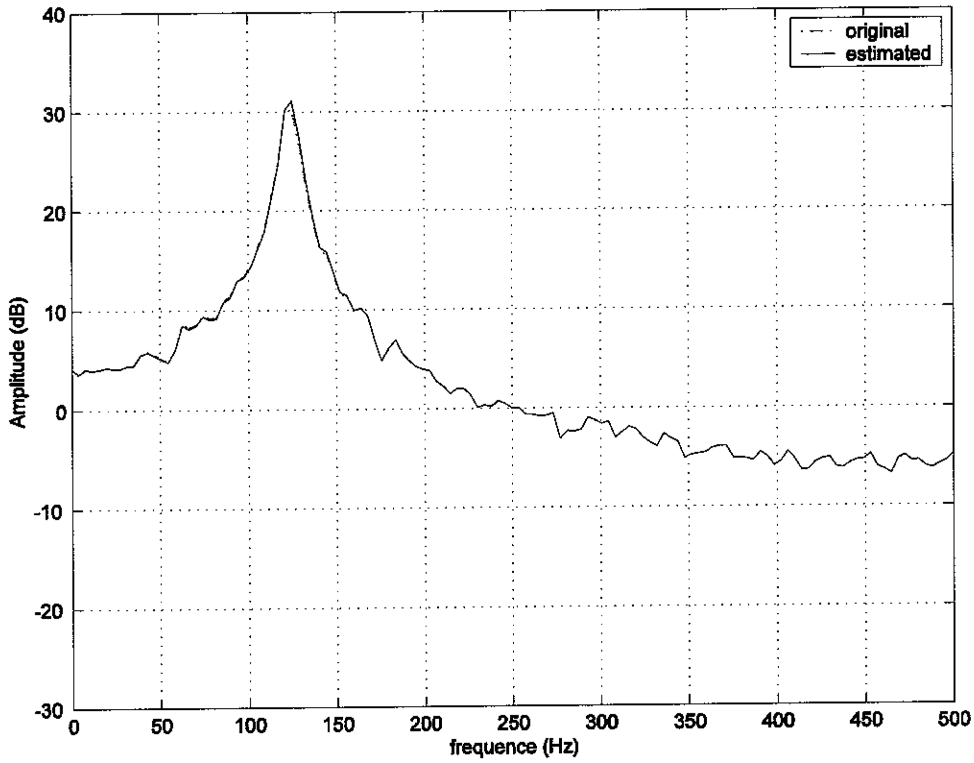
$$\mathbf{K}_1^{fT}=-(\mathbf{R}_{xx}(0))^{-1}\mathbf{R}_{xx}(1)=\begin{pmatrix} -0.8518 & -0.6478 \\ 0.7477 & -0.5486 \end{pmatrix} \tag{3.3.47}$$

$$\mathbf{K}_1^{bT}=-(\mathbf{R}_{xx}(0))^{-1}\mathbf{R}_{xx}(-1)=\begin{pmatrix} -0.5581 & -0.6404 \\ -0.7520 & -0.8422 \end{pmatrix} \tag{3.3.48}$$

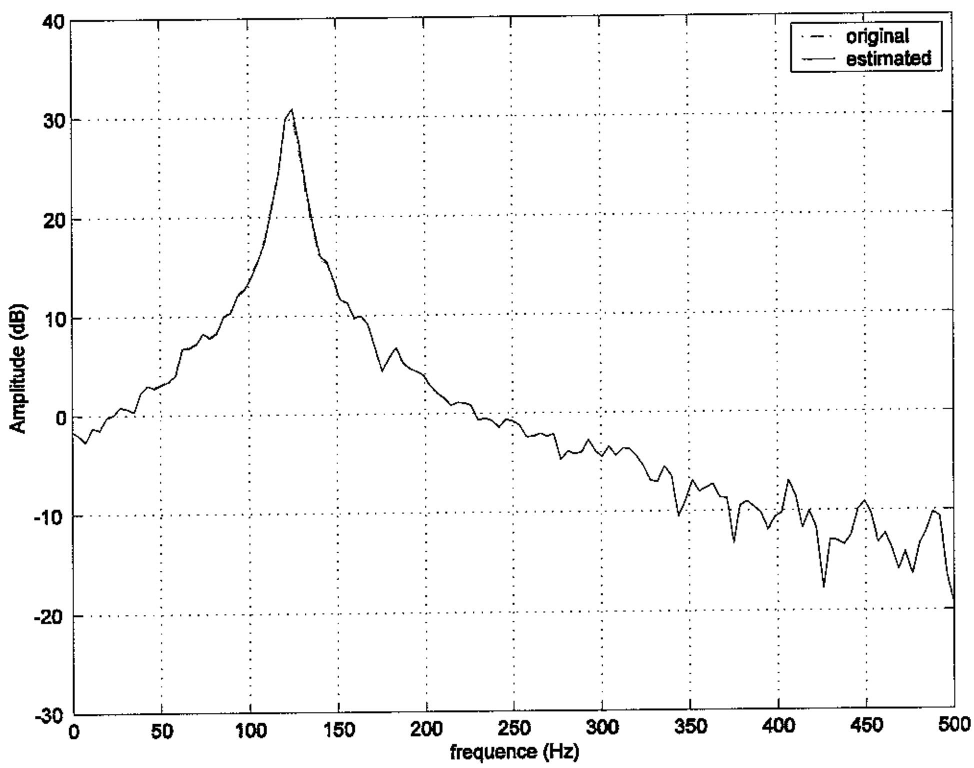
$$\mathbf{\Sigma}=\mathbf{R}_{xx}(0)\left(\mathbf{I}-\mathbf{K}_1^{bT}\mathbf{K}_1^{fT}\right)=\begin{pmatrix} -1.0085 & -0.0070 \\ -0.0070 & -0.9639 \end{pmatrix} \tag{3.3.49}$$

From (3.3.46), we can see a good estimation result of A(1) is achieved.

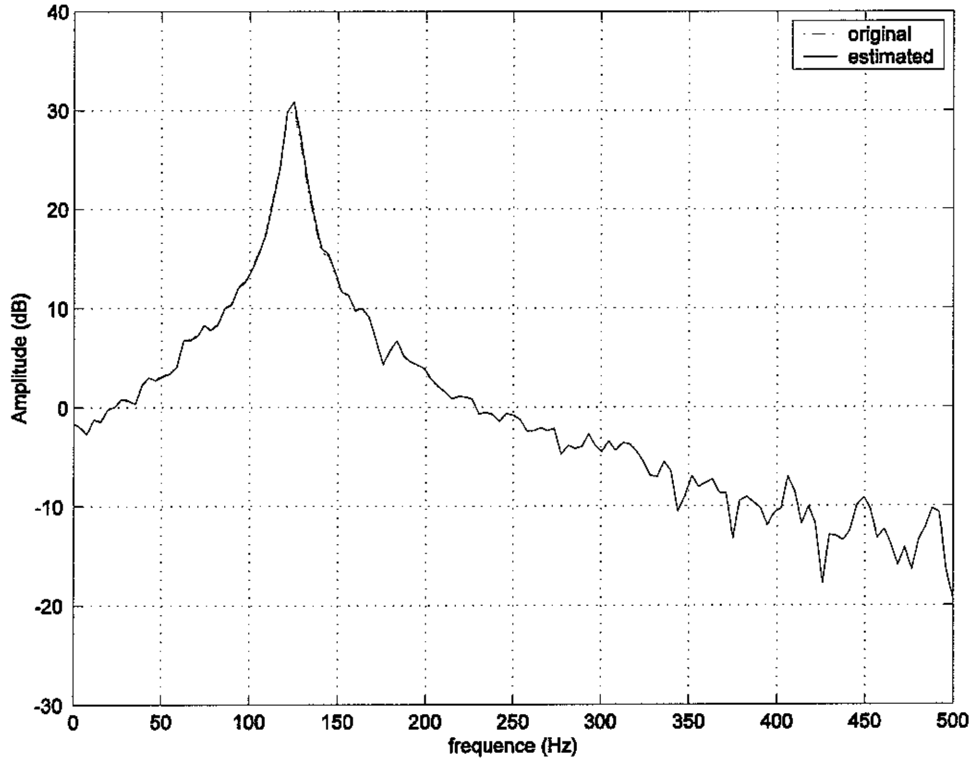
Graph of calculation results



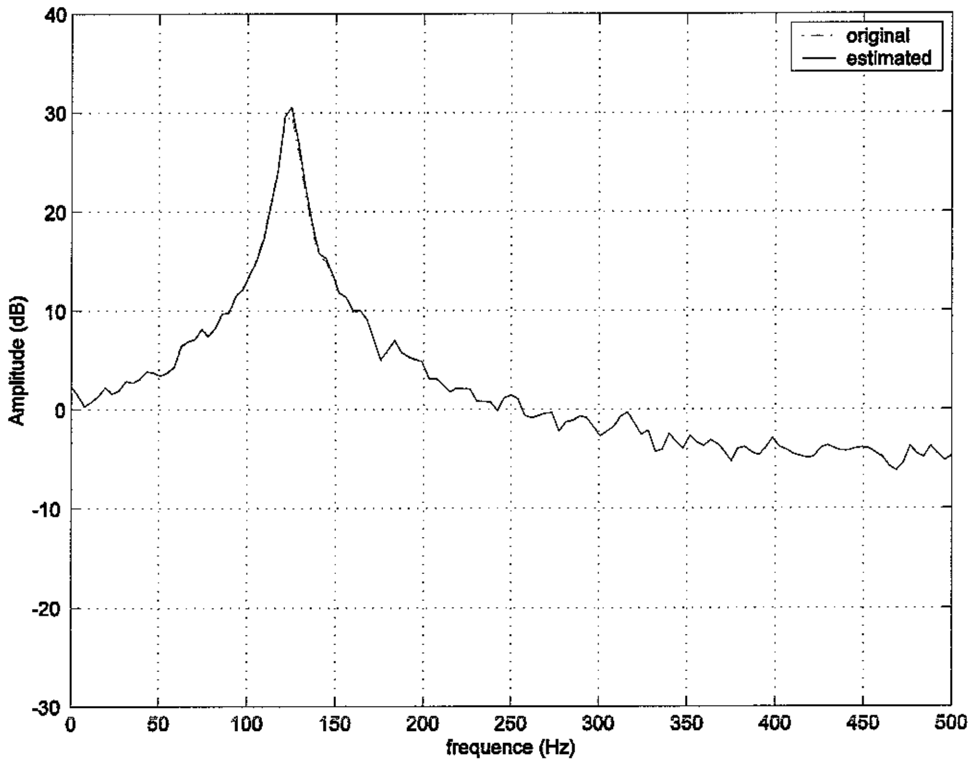
• Figure 10 Power spectral density of x1



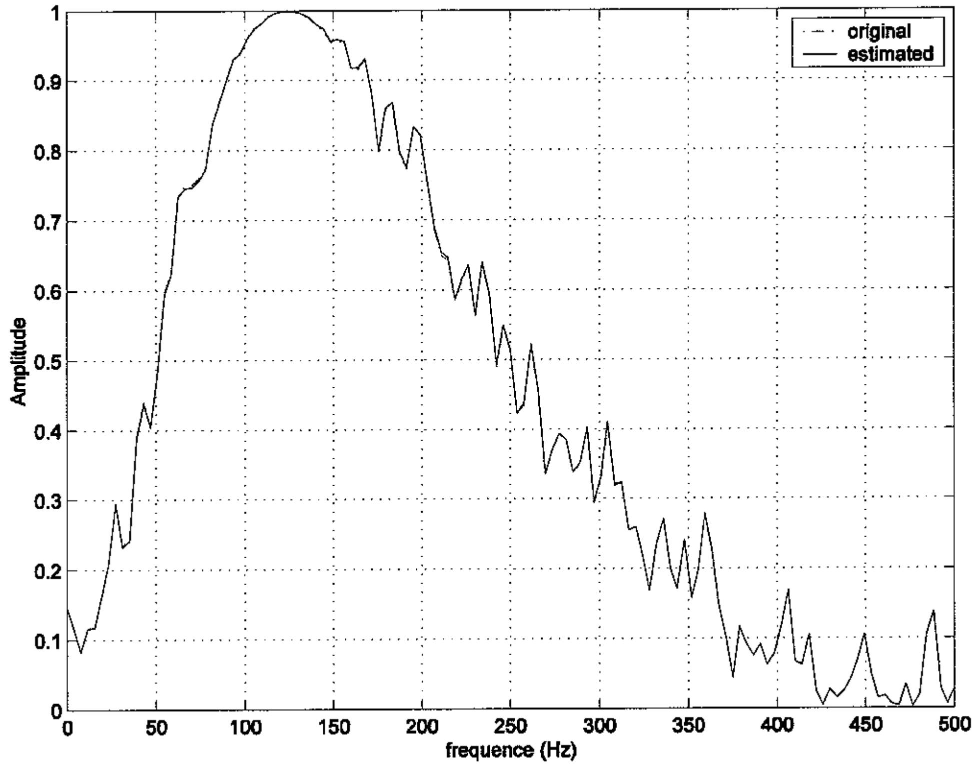
• Figure 11 Cross power spectral density of x1 and x2



• Figure 12 Cross power spectral density of x2 and x1



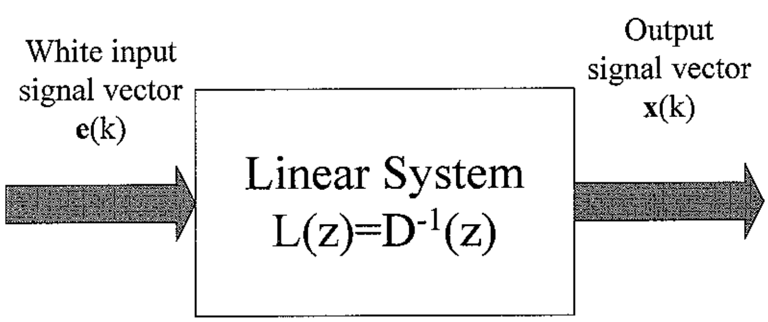
• Figure 13 Power spectral density of x2



• Figure 14 Coherence between x1 and x2

3.4 Whitening of Multichannel Random signal: Multichannel Prediction Error Filter

In previous section, it was seen that a multichannel random signal can be represented as a output of a linear system driven by multichannel white noise,



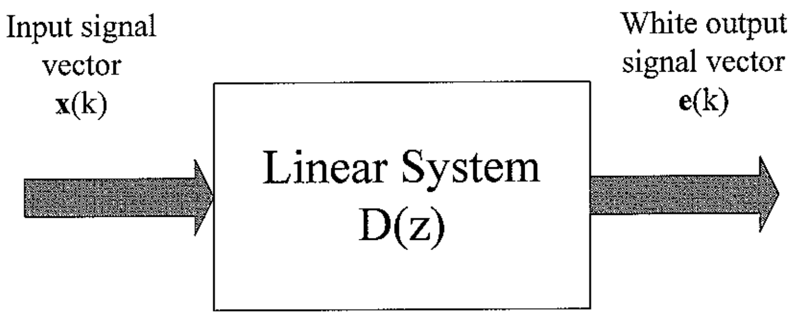
• Figure 15 AR model for random processes with transfer function $\mathbf{L}(z)$

where the linear system has a autoregressive (AR) model transfer function matrix given as following.

$$\mathbf{D}(z) = \mathbf{I}_p + \sum_{i=1}^{m} \mathbf{A}(i) z^{-i}$$

$$\mathbf{L}(z) = \mathbf{D}^{-1}(z) = \left(\mathbf{I}_p + \sum_{i=1}^{m} \mathbf{A}(i) z^{-i} \right)^{-1}$$

Here we think about the inverse problem. For a given multichannel random signal as input signal, can we find a linear system which produces corresponding output of multichannel white noise? The answer is very easy. If a multichannel random signal can be represent by AR model of Figure 14, then its inverse system is used to get multichannel white noise as shown in following.



• Figure 16 Whitening filter with transfer function $\mathbf{D}(z)$

Because

$$\mathbf{S}(z) = \mathbf{L}(z) \mathbf{\Sigma} \mathbf{L}_*(z)$$

The spectral matrix of output signal is given by

$$\mathbf{S}_{ee}(z) = \mathbf{L}^{-1}(z) \mathbf{S}(z) \mathbf{L}_*^{-1}(z) = \mathbf{\Sigma} \qquad (3.4.1)$$

As we have mentioned, because $\mathbf{\Sigma}$ is a symmetrical matrix, we can decompose it into $\mathbf{\Sigma} = \mathbf{G}\mathbf{G}^T$, then transfer function can be given by $\mathbf{L}(\omega)\mathbf{G}$, from p-channel random signal, we can get multichannel white signal with covariance matrix $\mathbf{\Sigma} = \mathbf{I}_p$.

$$
\begin{aligned}
\mathbf{S}_{ee}(z) &= (\mathbf{L}(z)\mathbf{G})^{-1}\mathbf{S}(z)(\mathbf{G}^T\mathbf{L}.(z))^{-1} \\
&= \mathbf{G}^{-1}\mathbf{L}^{-1}(z)\mathbf{S}(z)\mathbf{L}.^{-1}(z)\mathbf{G}^{T-1} \\
&= \mathbf{I}
\end{aligned}
\tag{3.3.2}
$$

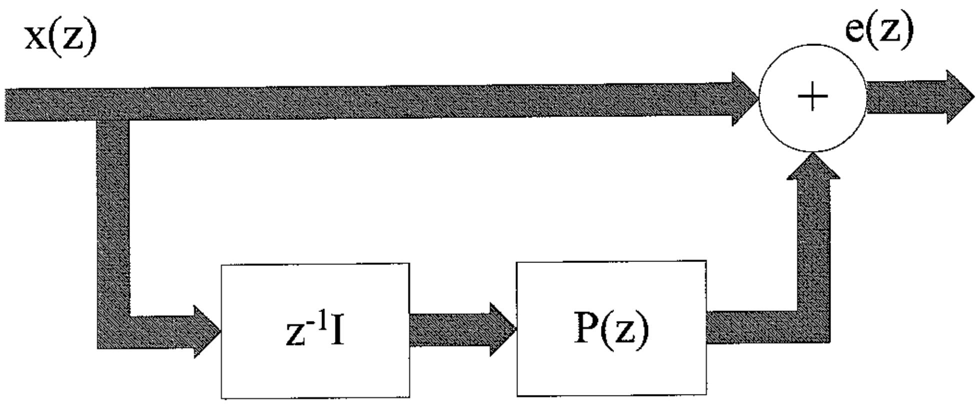
From transfer function matrix

$$
\begin{aligned}
\mathbf{D}(z) &= \mathbf{I}_p + \sum_{i=1}^{m} \mathbf{A}(i) z^{-i} \\
&= \mathbf{I}_p - z^{-1}\left(-\sum_{i=0}^{m-1} \mathbf{A}(i+1) z^{-i} \right) \\
&= \mathbf{I}_p - z^{-1}\mathbf{P}(z)
\end{aligned}
\tag{3.4.3}
$$

where

$$
\mathbf{P}(z) = -\sum_{i=0}^{m-1} \mathbf{A}(i+1) z^{-i} \tag{3.4.4}
$$

This system is called multichannel prediction error filter (MPEF)[9], its block diagram is shown in Figure 17, in which the delay operator is a diagonal matrix of delay.



• Figure 17 Block diagram of multichannel prediction error filter

The transfer function matrix $\mathbf{P}(z)$ is a multichannel FIR filter as shown in (3.3.4).

Let's have another look of this from frequency domain.

For the first step, from Figure 17, the transfer function of whole PEF is given as following.

$$\mathbf{PEF}(z) = \mathbf{I} - z^{-1}\mathbf{P}(z) \tag{3.4.5}$$

Suppose that spectral density matrix of input data can be factorise as following.

$$\mathbf{S}_{xx}(z) = \mathbf{F}(z)\mathbf{F}^T(z^{-1}) \tag{3.4.6}$$

The optimum least squares value for $\mathbf{P}(z)$ can be written as

$$\mathbf{P}_{opt}(z) = \{\mathbf{S}_{yx}(z)(\mathbf{F}^T(z^{-1}))^{-1}\}_+\mathbf{F}^{-1}(z) \tag{3.4.7}$$

where $\{\}_+$ is used to denote the causal part. Because

$$\mathbf{S}_{yx}(z) = z\mathbf{S}_{xx}(z) = z\mathbf{F}(z)\mathbf{F}^T(z^{-1}) \tag{3.4.8}$$

so we can get

$$\mathbf{P}_{opt}(z) = \{z\mathbf{F}(z))\}_+\mathbf{F}^{-1}(z) \tag{3.4.9}$$

If the matrix of spectral factors is written as

$$\mathbf{F}(z) = \mathbf{F}_0 + \mathbf{F}_1 z^{-1} + \mathbf{F}_2 z^{-2} + \cdots \tag{3.4.10}$$

so that

$$\{z\mathbf{F}(z)\}_+ = \mathbf{F}_1 + \mathbf{F}_2 z^{-1} + \cdots = z(\mathbf{F}(z) - \mathbf{F}_0) \tag{3.4.11}$$

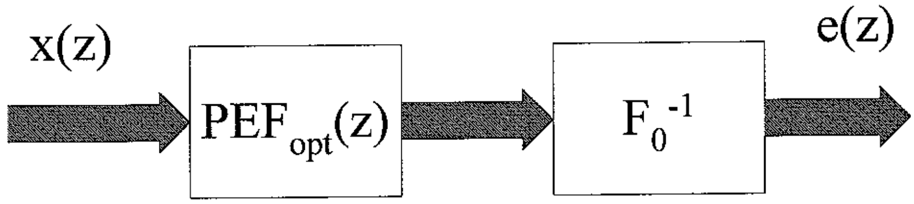
Thus we can get

$$\mathbf{P}_{opt}(z) = z(\mathbf{F}(z) - \mathbf{F}_0)\mathbf{F}^{-1}(z) = z(\mathbf{I} - \mathbf{F}_0\mathbf{F}^{-1}(z)) \tag{3.4.12}$$

$$\mathbf{PEF}_{opt}(z) = \mathbf{I} - z^{-1}\mathbf{P}_{opt}(z) = \mathbf{F}_0\mathbf{F}^{-1}(z) \tag{3.4.13}$$

$$\begin{aligned} \mathbf{S}_{ee}(z) &= \mathbf{PEF}_{opt}(z)\mathbf{S}_{xx}(z)\mathbf{PEF}_{opt}^T(z^{-1}) \\ &= \mathbf{F}_0\mathbf{F}^{-1}(z)\mathbf{F}(z)\mathbf{F}^T(z^{-1})\mathbf{F}^{-T}(z^{-1})\mathbf{F}_0^{\,T} \\ &= \mathbf{F}_0\mathbf{F}_0^{\,T} \end{aligned} \tag{3.4.14}$$

which has no concern with z, This means that error signals are white but correlated signal. The second step is use to decorrelate the output signal of PEF by using inverse matrix of $\mathbf{F}_0$, so finally we can get a uncorrelated, unit variance, white signal by a system shown in Figure 18.



• Figure 18 Decorrelation filter using prediction error filter

Chapter 4 SYSTEM IDENTIFICATION BY SUBSPACE THEORY

In this chapter, the identification problem for a multiple input multiple output system is discussed. From input data and output data, the subspace method can estimate the coefficient matrices of state space expression of the system, from which its multichannel transfer function can be obtained. At first, two basic theorems are described. Then based on these theorems, the method to determine the coefficient matrices of state space equation are shown. Also in order the performance of this method, a numerical example is shown, comparing the output spectral of the original system and the estimated system.

4.1 Basic Theorems

4.1.1 Definitions

The linear time invariant system is assumed to be represented by state space equation given as following.

$$\begin{aligned} \mathbf{x}_{k+1} &= \mathbf{A}\mathbf{x}_k + \mathbf{B}\mathbf{u}_k + \mathbf{w}_k \\ \mathbf{y}_k &= \mathbf{C}\mathbf{x}_k + \mathbf{D}\mathbf{u}_k + \mathbf{v}_k \end{aligned} \qquad (4.1.1)$$

where $\mathbf{x}_k \in \Re^n, \mathbf{u}_k \in \Re^m, \mathbf{y}_k, \mathbf{v}_k \in \Re^l$ and $\mathbf{w}_k \in \Re^m$. The unknown system matrices $\mathbf{A}, \mathbf{B}, \mathbf{C}, \mathbf{D}$ have appropriate dimensions. The process noise $\mathbf{w}_k$ and the measurement noise $\mathbf{v}_k$ are zero-mean white noise sequences, statistically independent of the input $\mathbf{u}_k$. They satisfy

$$E\left[\begin{pmatrix}\mathbf{w}_i\\ \mathbf{v}_i\end{pmatrix}\begin{pmatrix}\mathbf{w}_j^{\,T} & \mathbf{v}_j^{\,T}\end{pmatrix}\right]=\begin{bmatrix}\mathbf{Q} & \mathbf{S}\\ \mathbf{S} & \mathbf{R}\end{bmatrix}\cdot\delta_{ij} \tag{4.1.2}$$

Here, we would like to give several definitions of the system these will be useful in description of two import theorems.

1. extended observability matrix

$$\mathbf{\Gamma}_i=\begin{pmatrix}\mathbf{C}\\ \mathbf{CA}\\ \mathbf{CA}^2\\ \vdots\\ \mathbf{CA}^{i-1}\end{pmatrix} \tag{4.1.3}$$

2. the reversed extended controllability matrix

$$\mathbf{\Delta}_i=\begin{pmatrix}\mathbf{A}^{i-1}\mathbf{B} & \mathbf{A}^{i-2}\mathbf{B} & \cdots & \mathbf{AB} & \mathbf{B}\end{pmatrix} \tag{4.1.4}$$

3. the low block triangular Toeplitz matrix

$$\mathbf{H}_i=\begin{pmatrix}\mathbf{D} & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0}\\ \mathbf{CB} & \mathbf{D} & \mathbf{0} & \cdots & \mathbf{0}\\ \mathbf{CAB} & \mathbf{CB} & \mathbf{D} & \cdots & \mathbf{0}\\ \vdots & \vdots & \vdots & \ddots & \vdots\\ \mathbf{CA}^{i-2}\mathbf{B} & \mathbf{CA}^{i-3}\mathbf{B} & \mathbf{CA}^{i-4}\mathbf{B} & \cdots & \mathbf{D}\end{pmatrix} \tag{4.1.5}$$

4. block Hankel matrix of input signal

$$\mathbf{U}_{0|i-1}=\begin{pmatrix}\mathbf{u}_0 & \mathbf{u}_1 & \mathbf{u}_2 & \cdots & \mathbf{u}_{j-1}\\ \mathbf{u}_1 & \mathbf{u}_2 & \mathbf{u}_3 & \cdots & \mathbf{u}_j\\ \mathbf{u}_2 & \mathbf{u}_3 & \mathbf{u}_4 & \cdots & \mathbf{u}_{j+1}\\ \cdots & \cdots & \cdots & \cdots & \cdots\\ \mathbf{u}_{i-1} & \mathbf{u}_i & \mathbf{u}_{i+1} & \cdots & \mathbf{u}_{i+j-2}\end{pmatrix} \tag{4.1.6}$$

where 0|i-1 denotes 1st column component from u_0 to u_{i-1}.

5. block Hankel matrix of output signal

$$\mathbf{Y}_{0|i-1} = \begin{pmatrix} \mathbf{y}_0 & \mathbf{y}_1 & \mathbf{y}_2 & \cdots & \mathbf{y}_{j-1} \\ \mathbf{y}_1 & \mathbf{y}_2 & \mathbf{y}_3 & \cdots & \mathbf{y}_j \\ \mathbf{y}_2 & \mathbf{y}_3 & \mathbf{y}_4 & \cdots & \mathbf{y}_{j+1} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \mathbf{y}_{i-1} & \mathbf{y}_i & \mathbf{y}_{i+1} & \cdots & \mathbf{y}_{i+j-2} \end{pmatrix} \tag{4.1.7}$$

4.1.2 Theorem 1[10)]

A 2i row block Hankel matrix of input signal can be shown as following.

$$\mathbf{U}_{0|2i-1} = \begin{pmatrix} \mathbf{u}_0 & \mathbf{u}_1 & \mathbf{u}_2 & \cdots & \mathbf{u}_{j-1} \\ \mathbf{u}_1 & \mathbf{u}_2 & \mathbf{u}_3 & \cdots & \mathbf{u}_j \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \mathbf{u}_{i-1} & \mathbf{u}_i & \mathbf{u}_{i+1} & \cdots & \mathbf{u}_{i+j-2} \\ \mathbf{u}_i & \mathbf{u}_{i+1} & \mathbf{u}_{i+2} & \cdots & \mathbf{u}_{i+j-1} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \mathbf{u}_{2i-1} & \mathbf{u}_{2i} & \mathbf{u}_{2i+1} & \cdots & \mathbf{u}_{2i+j-2} \end{pmatrix} \tag{4.1.8}$$

This matrix can be partitioned as

$$\mathbf{U}_{0|2i-1} = \begin{bmatrix} \mathbf{U}_{0|i-1} \\ \mathbf{U}_{i|2i-1} \end{bmatrix} \tag{4.1.9}$$

Also a 2i row block Hankel matrix of output signal can be shown as following.

$$\mathbf{Y}_{0|2i-1} = \begin{pmatrix} \mathbf{y}_0 & \mathbf{y}_1 & \mathbf{y}_2 & \cdots & \mathbf{y}_{j-1} \\ \mathbf{y}_1 & \mathbf{y}_2 & \mathbf{y}_3 & \cdots & \mathbf{y}_j \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \mathbf{y}_{i-1} & \mathbf{y}_i & \mathbf{y}_{i+1} & \cdots & \mathbf{y}_{i+j-2} \\ \mathbf{y}_i & \mathbf{y}_{i+1} & \mathbf{y}_{i+2} & \cdots & \mathbf{y}_{i+j-1} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \mathbf{y}_{2i-1} & \mathbf{y}_{2i} & \mathbf{y}_{2i+1} & \cdots & \mathbf{y}_{2i+j-2} \end{pmatrix} \tag{4.1.10}$$

This matrix can be partitioned as

$$\mathbf{Y}_{0|2i-1} = \begin{bmatrix} \mathbf{Y}_{0|i-1} \\ \mathbf{Y}_{i|2i-1} \end{bmatrix} \tag{4.1.11}$$

The matrix input-output equations are defined in the following theorem,

$$\begin{aligned}
\mathbf{Y}_{0|i-1} &= \mathbf{\Gamma}_i \mathbf{X}_0 + \mathbf{H}_i \mathbf{U}_{0|i-1} + \mathbf{Y}^s_{0|i-1} \\
\mathbf{Y}_{i|2i-1} &= \mathbf{\Gamma}_i \mathbf{X}_i + \mathbf{H}_i \mathbf{U}_{i|2i-1} + \mathbf{Y}^s_{i|2i-1} \\
\mathbf{X}_i &= \mathbf{A}^i \mathbf{X}_0 + \mathbf{\Delta}_i \mathbf{U}_{0|i-1}
\end{aligned} \tag{4.1.12}$$

where

$$\mathbf{X}_i = \begin{pmatrix} \mathbf{x}_i & \mathbf{x}_{i+1} & \mathbf{x}_{i+2} & \cdots & \mathbf{x}_{i+j-1} \end{pmatrix} \tag{4.1.13}$$

4.1.3 Theorem 2[10)]

Here we introduce the projection of the future outputs onto the past and future inputs and the past outputs. The results can be described as a function of system matrices and the input-output block Hankel matrices.

We define the matrix $\mathbf{Z}_i$, $\mathbf{Z}_{i+1}$ as following.

$$\mathbf{Z}_i = \mathbf{Y}_{i|2i-1} \Big/ \begin{pmatrix} \mathbf{U}_{0|2i-1} \\ \mathbf{Y}_{0|i-1} \end{pmatrix} \tag{4.1.14a}$$

$$\mathbf{Z}_{i+1} = \mathbf{Y}_{i+1|2i-1} \Big/ \begin{pmatrix} \mathbf{U}_{0|2i-1} \\ \mathbf{Y}_{0|i} \end{pmatrix} \tag{4.1.14b}$$

where $\mathbf{A}/\mathbf{B} = \mathbf{A}\mathbf{B}^T(\mathbf{B}\mathbf{B}^T)^{-1}\mathbf{B}$. The row space of $\mathbf{A}/\mathbf{B}$ is equal to the projection of row space of A onto the row space of B. (4.1.14a) corresponds to the optimal prediction of $\mathbf{Y}_{i|2i-1}$ given $\mathbf{U}_{0|2i-1}$ and $\mathbf{Y}_{0|i-1}$ in a sense that $\left\|\mathbf{Y}_{i|2i-1} - \mathbf{Z}_i\right\|_F^2$ is minimized constrained to row space $\mathbf{Z}_i \subset$ row space $\begin{pmatrix} \mathbf{U}_{0|2i-1} \\ \mathbf{Y}_{0|i-1} \end{pmatrix}$.

By assume certain conditions, it can be shown that

$$\mathbf{Z}_i = \mathbf{\Gamma}_i \hat{\mathbf{X}}_i + \mathbf{H}_i \mathbf{U}_{i|2i-1} \tag{4.1.15a}$$

$$\mathbf{Z}_{i+1} = \mathbf{\Gamma}_{i-1}\hat{\mathbf{X}}_{i+1} + \mathbf{H}_{i-1}\mathbf{U}_{i-1|2i-1} \tag{4.1.15b}$$

where state vector is given by

$$\hat{\mathbf{X}}_i = \begin{pmatrix}\hat{\mathbf{x}}_i & \hat{\mathbf{x}}_{i+1} & \hat{\mathbf{x}}_{i+2} & \cdots & \hat{\mathbf{x}}_{i+j-1}\end{pmatrix} = \begin{bmatrix}\theta_1 & \theta_2 & \theta_3\end{bmatrix}\begin{bmatrix}\hat{\mathbf{X}}_0 \\ \mathbf{U}_{0|i-1} \\ \mathbf{Y}_{0|i-1}\end{bmatrix} \subset \Re^{n\times j} \tag{4.1.15c}$$

with $[\theta_1 \quad \theta_2 \quad \theta_3]$ is a function of system matrices and $\hat{\mathbf{X}}_0 = \mathbf{X}_0 / \mathbf{U}_{0|2i-1}$.

4.2 Determine the System Order

Let $\mathbf{T}$ be any rank deficient matrix of which the column space coincides with that of $\mathbf{\Gamma}_i$. Calculate the singular value decomposition

$$\mathbf{T} = \begin{pmatrix}\mathbf{U}_1 & \mathbf{U}_2\end{pmatrix}\begin{pmatrix}\mathbf{\Sigma}_1 & 0 \\ 0 & 0\end{pmatrix}\mathbf{V}^T \tag{4.2.1}$$

Since $\mathbf{T}$ is rank of n, the number of singular value different from zero will be equal to the order of the system.

4.3 Determine the System Matrices

From theorem 2

$$\hat{\mathbf{X}}_i = \mathbf{\Gamma}_i^{\,+}(\mathbf{Z}_i - \mathbf{H}_i\mathbf{U}_{i|2i-1}) = \mathbf{\Gamma}_i^{\,+}\mathbf{Z}_i - \mathbf{\Gamma}_i^{\,+}\mathbf{H}_i\mathbf{U}_{i|2i-1} \tag{4.3.1a}$$

$$\hat{\mathbf{X}}_{i+1} = \mathbf{\Gamma}_{i-1}^{+}(\mathbf{Z}_{i+1} - \mathrm{H}_{i-1}\mathbf{U}_{i+1|2i-1}) = \mathbf{\Gamma}_{i-1}^{+}\mathbf{Z}_{i+1} - \mathbf{\Gamma}_{i-1}^{+}\mathrm{H}_{i-1}\mathbf{U}_{i+1|2i-1} \tag{4.3.1b}$$

Here only H_i and H_{i-1} are unknown. From

$$\hat{\mathbf{X}}_{i+1} = \mathbf{A}\hat{\mathbf{X}}_i + \mathbf{B}\mathbf{U}_{i|i} + \begin{pmatrix}\mathbf{U}_{0|2i-1} \\ \mathbf{Y}_{0|i-1} \\ \hat{\mathbf{X}}_i\end{pmatrix}^{\perp} \tag{4.3.2a}$$

$$\mathbf{Y}_{i|i} = \mathbf{C}\hat{\mathbf{X}}_i + \mathbf{D}\mathbf{U}_{i|i} + \begin{pmatrix} \mathbf{U}_{0|2i-1} \\ \mathbf{Y}_{0|i-1} \\ \hat{\mathbf{X}}_i \end{pmatrix}^{\perp} \tag{4.3.2b}$$

where $(.)^{\perp}$ means orthogonal complement. (cf. section 4.5.3) we can get

$$\begin{pmatrix} \hat{\mathbf{X}}_{i+1} \\ \hline \mathbf{Y}_{i|i} \end{pmatrix} = \begin{pmatrix} \mathbf{A} \\ \hline \mathbf{C} \end{pmatrix} \hat{\mathbf{X}}_i + \begin{pmatrix} \mathbf{B} \\ \hline \mathbf{D} \end{pmatrix} \mathbf{U}_{i|i} + \begin{pmatrix} \mathbf{U}_{0|2i-1} \\ \mathbf{Y}_{0|i-1} \\ \hat{\mathbf{X}}_i \end{pmatrix}^{\perp} \tag{4.3.3}$$

Then from

$$\mathbf{\Gamma}_{i-1}^{+}\mathbf{Z}_{i+1} = \hat{\mathbf{X}}_{i+1} - \mathbf{\Gamma}_{i-1}^{+}\mathbf{H}_{i-1}\mathbf{U}_{i+1|2i-1} \tag{4.3.4}$$

By substitute (4.3.2a) into this equation, we can get

$$\mathbf{\Gamma}_{i-1}^{+}\mathbf{Z}_{i+1} = \mathbf{A}\hat{\mathbf{X}}_i + \mathbf{B}\mathbf{U}_{i|i} + \begin{pmatrix} \mathbf{U}_{0|2i-1} \\ \mathbf{Y}_{0|i-1} \\ \hat{\mathbf{X}}_i \end{pmatrix}^{\perp} - \mathbf{\Gamma}_{i-1}^{+}\mathbf{H}_{i-1}\mathbf{U}_{i+1|2i-1} \tag{4.3.5}$$

also using (4.3.1a) and (4.3.2b),

$$\begin{pmatrix} \mathbf{\Gamma}_{i-1}^{+}\mathbf{Z}_{i+1} \\ \hline \mathbf{Y}_{i|i} \end{pmatrix} = \begin{pmatrix} \mathbf{A} \\ \hline \mathbf{C} \end{pmatrix} (\mathbf{\Gamma}_i^{+}\mathbf{Z}_i - \mathbf{\Gamma}_i^{+}\mathbf{H}_i\mathbf{U}_{i|2i-1}) + \begin{pmatrix} \mathbf{B} \\ \hline \mathbf{D} \end{pmatrix} \mathbf{U}_{i|i} + \begin{pmatrix} \mathbf{U}_{0|2i-1} \\ \mathbf{Y}_{0|i-1} \\ \hat{\mathbf{X}}_i \end{pmatrix}^{\perp} - \mathbf{\Gamma}_{i-1}^{+}\mathbf{H}_{i-1}\mathbf{U}_{i+1|2i-1} \tag{4.3.6}$$

$$\begin{pmatrix} \mathbf{\Gamma}_{i-1}^{+}\mathbf{Z}_{i+1} \\ \hline \mathbf{Y}_{i|i} \end{pmatrix} = \begin{pmatrix} \mathbf{A} \\ \hline \mathbf{C} \end{pmatrix} \mathbf{\Gamma}_i^{+}\mathbf{Z}_i + \begin{pmatrix} \mathbf{K}_{12} \\ \hline \mathbf{K}_{22} \end{pmatrix} \mathbf{U}_{i|2i-1} + \begin{pmatrix} \mathbf{U}_{0|2i-1} \\ \mathbf{Z}_i \\ \hat{\mathbf{X}}_i \end{pmatrix}^{\perp} \tag{4.3.7}$$

In right side of this equation, from coefficient of 1st term, **A** matrix and **C** matrix can be obtained exactly. Then coefficient of 2nd term, which is a set of linear equation of **B** and **D**, matrix **B** and matrix **D** can be solved in the means of least mean square.

4.4 Example by N4SID of MATLAB

An example as following is considered for system identification by subspace method using the MATLAB function N4SID in system identification toolbox. This two inputs and two output system is represented in state space model

$$\begin{aligned} \mathbf{x}_{k+1} &= \mathbf{A}\mathbf{x}_k + \mathbf{B}\mathbf{u}_k + \mathbf{K}\mathbf{v}_k \\ \mathbf{y}_k &= \mathbf{C}\mathbf{x}_k + \mathbf{D}\mathbf{u}_k + \mathbf{v}_k \end{aligned} \tag{4.4.1}$$

where coefficient matrices are assumed to be given as following.

$$\mathbf{A} = \begin{bmatrix} 0.67 & 0.67 & 0 & 0 \\ -0.67 & 0.67 & 0 & 0 \\ 0 & 0 & -0.67 & -0.67 \\ 0 & 0 & 0.67 & -0.67 \end{bmatrix} \tag{4.4.2a}$$

$$\mathbf{B} = \begin{bmatrix} 0.6598 & 2.1256 \\ 1.9698 & 3.1201 \\ 4.3171 & 1.2050 \\ -2.6436 & -1.2356 \end{bmatrix} \tag{4.4.2b}$$

$$\mathbf{C} = \begin{bmatrix} -0.5749 & 1.0751 & -0.5225 & 0.1830 \\ 0.5687 & 2.3011 & 0.7520 & -0.1721 \end{bmatrix} \tag{4.4.2c}$$

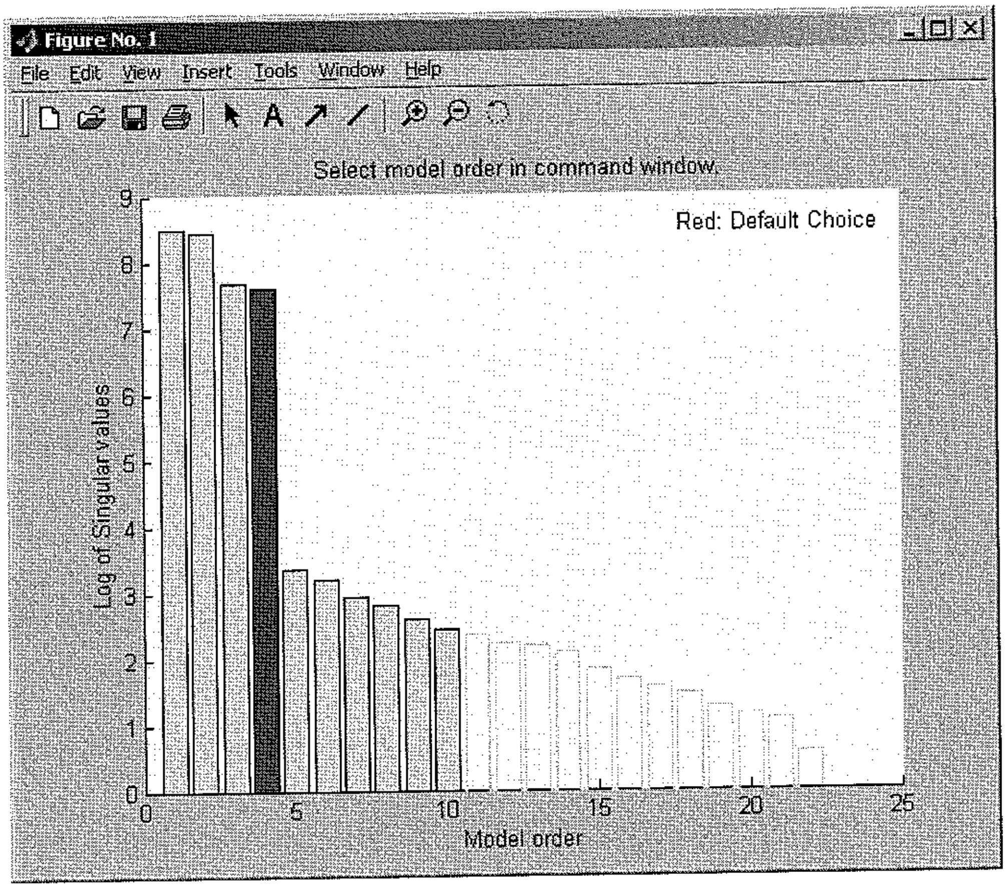
$$\mathbf{D} = \begin{bmatrix} -0.7139 & 1.2504 \\ 0.5478 & 0.9854 \end{bmatrix} \tag{4.4.2.d}$$

and with K as the Kalman gain matrix

$$\mathbf{K} = \begin{bmatrix} 0.2820 & -0.3041 \\ -0.7557 & 0.0296 \\ 0.1919 & 0.1317 \\ -0.3797 & 0.6538 \end{bmatrix} \tag{4.4.2e}$$

and u_k and v_k are independent white noise.

Create a input data and output data set of this system, the give these data to a MATLAB function to calculate coefficient matrices of example system . When program is run, a prompt of select model order is given as following. Generally we should choose an order such the singular values for higher order comparatively small in this picture and a default order is given in red.



• Figure 19 Singular values of the system

When we select default of order=4, calculation results can be given as following.

$$\mathbf{A}_e = \begin{bmatrix} 0.6076 & -0.7083 & -0.1127 & 0.0459 \\ 0.6216 & 0.7184 & -0.2592 & -0.0921 \\ -0.0755 & 0.0197 & -0.6162 & 0.7431 \\ -0.0728 & -0.0126 & -0.6015 & -0.7105 \end{bmatrix} \tag{4.4.3a}$$

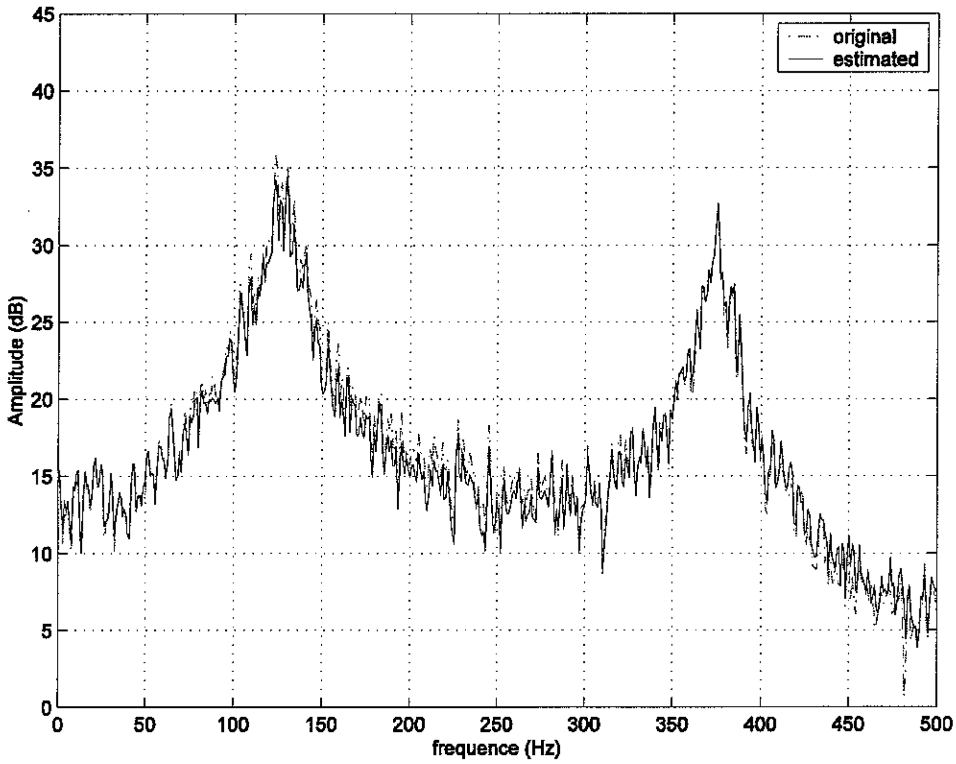
$$\mathbf{B}_e = \begin{bmatrix} -9.8771 & -15.9830 \\ -3.6107 & -6.5498 \\ -6.9798 & -0.6399 \\ -4.8123 & -1.1468 \end{bmatrix} \tag{4.4.3b}$$

$$\mathbf{C}_e = \begin{bmatrix} -0.1740 & 0.1443 & 0.2975 & 0.1241 \\ -0.4955 & -0.0946 & -0.3505 & -0.0789 \end{bmatrix} \quad (4.4.3c)$$

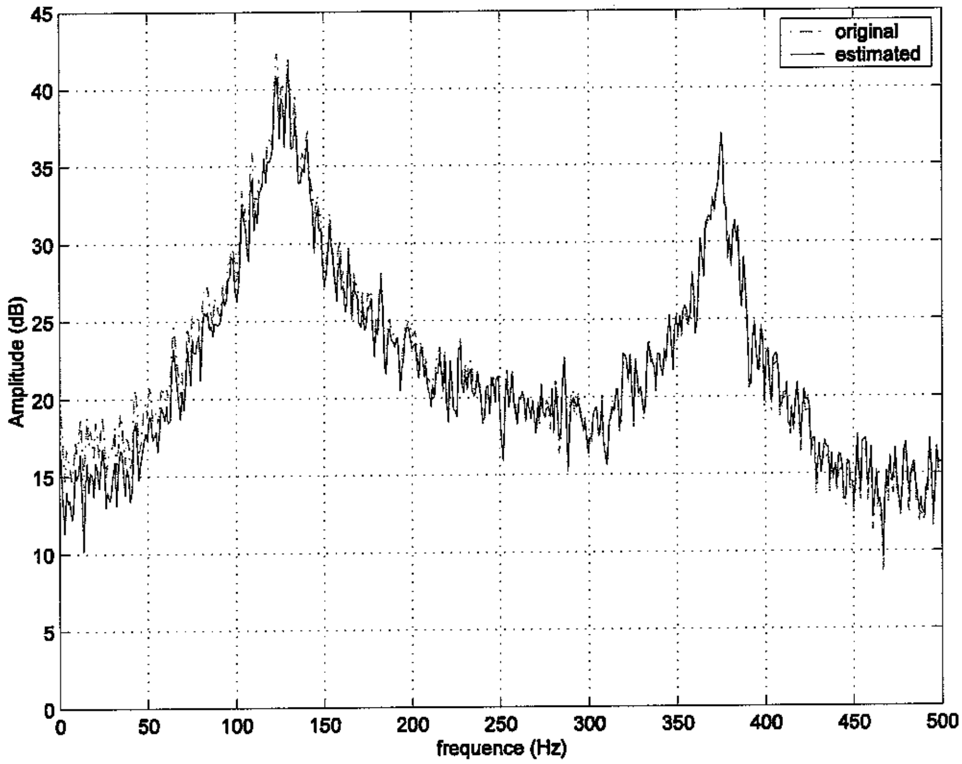
$$\mathbf{D}_e = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} \quad (4.4.3d)$$

$$\mathbf{K}_e = \begin{bmatrix} 0.6653 & -1.6501 \\ -1.1852 & -0.1693 \\ -0.0973 & -0.1051 \\ -0.7895 & 0.6058 \end{bmatrix} \quad (4.4.3e)$$

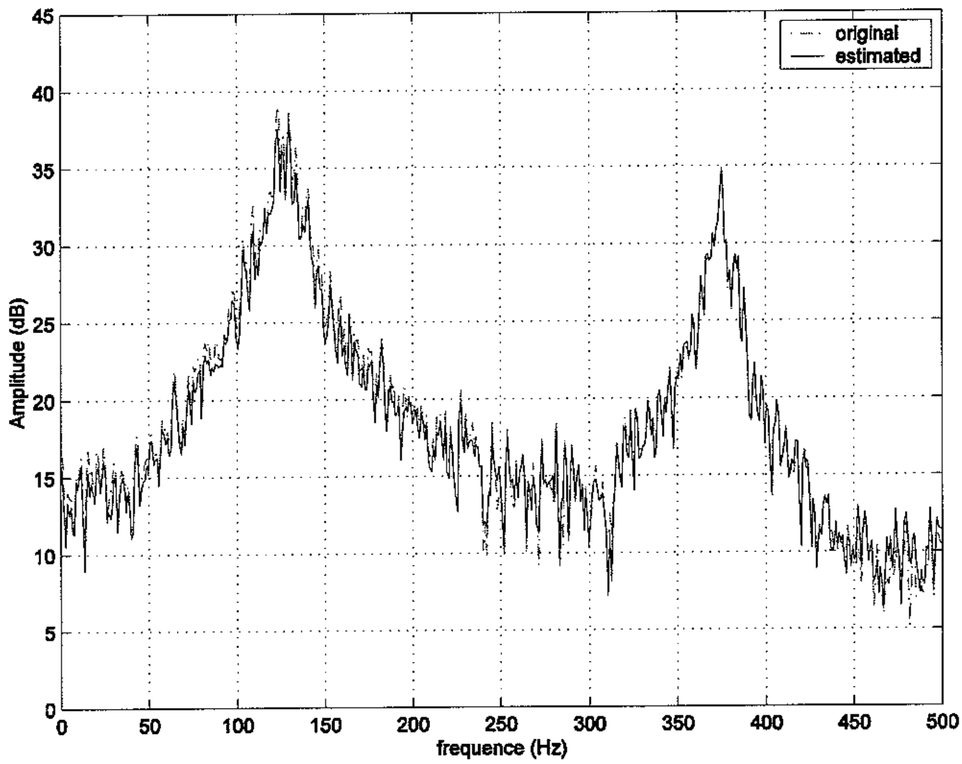
In order to check the performance of the results, we calculate the output signals for both original system and estimated system with same input signals. Then we get spectral data for both system and compare these results in graph shown in Figure 20 ~ Figure 23.



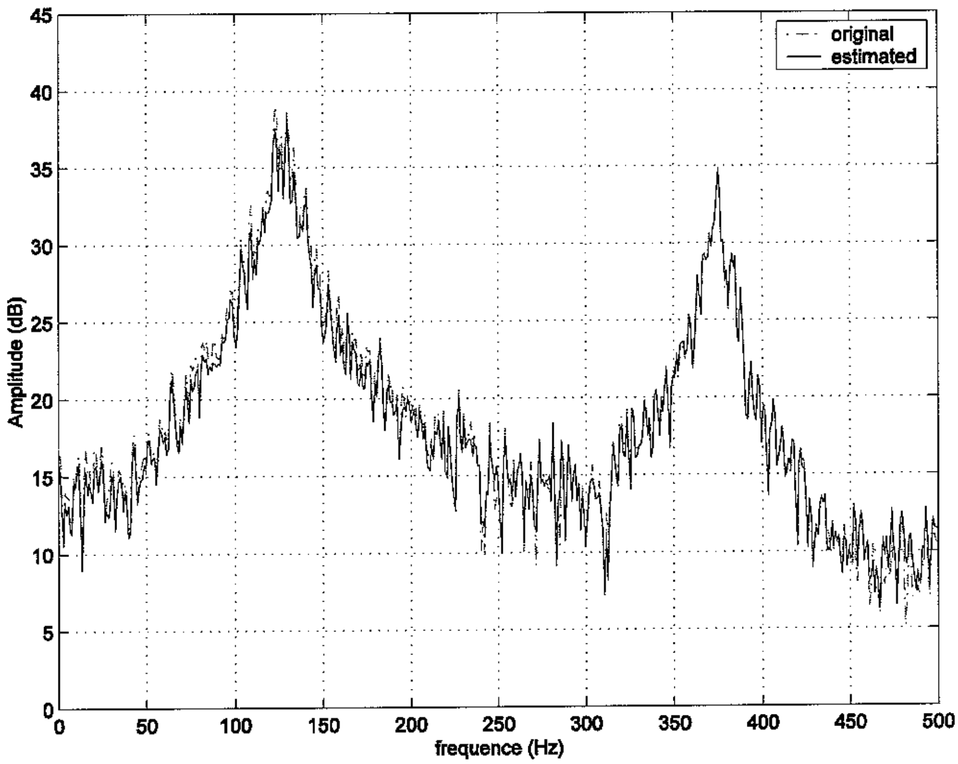
• Figure 20 Power spectral density of channel 1



• Figure 21 Cross-spectral density of channel 1 and channel 2



• Figure 22 Cross spectral density of channel 2 and channel 1



• Figure 23 Power spectral density of channel 2

4.5 Appendix: Matrix Subspace and Projections

4.5.1 Orthogonality

Definition 1: The norm of a vector $\mathbf{x} \in \Re^n$, denoted by $\|\mathbf{x}\|$, is given by the positive square root

$$\|\mathbf{x}\| = \sqrt{\mathbf{x}^T \mathbf{x}} \tag{4.5.1}$$

Suppose now that x and y are two vectors in $\Re^2$, and let $\mathbf{c} = \mathbf{x} - \mathbf{y}$. By the law of cosines for the triangle, we have

$$\|\mathbf{c}\|^2 = \|\mathbf{x} - \mathbf{y}\|^2 = \|\mathbf{x}\|^2 + \|\mathbf{y}\|^2 - 2\|\mathbf{x}\|\|\mathbf{y}\|\cos\theta \tag{4.5.2}$$

But from definition of norm,

$$\|\mathbf{x}-\mathbf{y}\|^2 = (\mathbf{x}-\mathbf{y})^T(\mathbf{x}-\mathbf{y}) = \|\mathbf{x}\|^2 - 2\mathbf{x}^T\mathbf{y} + \|\mathbf{y}\|^2 \qquad (4.5.3)$$

Substitut (4.5.2) into (4.5.3) and solving for $\cos\theta$ yields

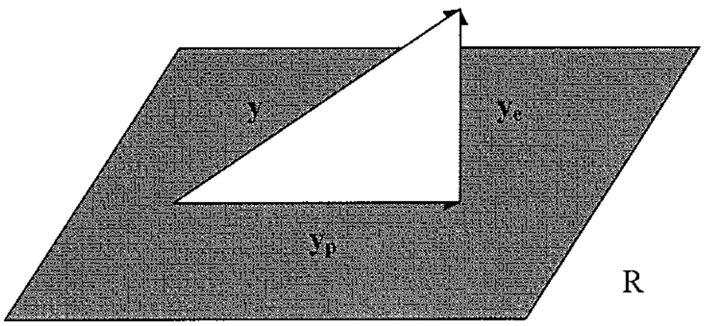
$$\cos\theta = \frac{\mathbf{x}^T\mathbf{y}}{\|\mathbf{x}\|\|\mathbf{y}\|} \qquad (4.5.4)$$

If the angle $\theta = 90^\circ$, then $\cos\theta = 0$, from (4.5.4) this means that $\mathbf{x}^T\mathbf{y} = 0$.

Definition 2: Two vectors $\mathbf{x},\mathbf{y} \in \Re^n$ are said to be orthogonal if their inner product $\mathbf{x}^T\mathbf{y} = 0$.

Definition 3: Two subspaces, S1 and S2, of $\mathbf{x} \in \Re^n$ are said to be orthogonal subspaces if for any $\mathbf{x} \in S_1$ and any $\mathbf{y} \in S_2$ the vectors x and y are orthogonal. If $\mathbf{x}_1,\ldots,\mathbf{x}_p$ is a basis for S1 and if $\mathbf{y}_1,\ldots,\mathbf{y}_q$ is a basis for S2 and, then S1 and S2 are orthogonal if $\mathbf{x}_i^T\mathbf{y}_j = 0$, $i = 1,\ldots,p, j = 1,\ldots,q$.

4.5.2 Orthogonal Projections



• Figure 24 $\mathbf{y}_p$ is the best approximation to $\mathbf{y}$ of any vector in R

The Importance of orthogonality comes from its use in deriving optimal approximations. The optimal solutions can be expressed in the language of linear algebra as follows: the shortest distance between a vector and a subspace occurs along a vector that is orthogonal to the subspace. Figure 24 shows a vector $\mathbf{y}$ and a subspace R. Of all vectors that connect the plane and vector $\mathbf{y}$, the vector $\mathbf{y}_e$ is the one with smallest norm. The vector $\mathbf{y}_e$ is orthogonal to the subspace R. The vector $\mathbf{y}_p$ is the vector in R that is the best approximation to $\mathbf{y}$. The vector $\mathbf{y}_p$ is called the projection of $\mathbf{y}$ onto R.

Assume that the subspace R is defined as the subspace spanned by a given set of basis $\mathbf{x}_1,\ldots,\mathbf{x}_r$. The question we would like to answer now is this: given a vector $\mathbf{y}$ and a basis for the subspace R, how can we compute $\mathbf{y}_p$, the projection of $\mathbf{y}$ onto R.

The answer is quit simple, and it comes from the fact "error vector"

$$\mathbf{y}_e = \mathbf{y} - \mathbf{y}_p \qquad (4.5.5)$$

is orthogonal to the subspace R, which means that error vector must be orthogonal to every vector in a basis for R.

$$\mathbf{y}_e^T \mathbf{x}_i = 0, \ i = 1,...,r \tag{4.5.6}$$

Because $\mathbf{y}_p$ is a vector in subspace, so that $\mathbf{y}_p$ can be expressed as a linear combination of the basis $\mathbf{x}_1,...,\mathbf{x}_r$, that is

$$\mathbf{y}_p = \alpha_1 \mathbf{x}_1 + \alpha_2 \mathbf{x}_2 + ... + \alpha_r \mathbf{x}_r \tag{4.5.7}$$

or in matrix form

$$\mathbf{y}_p = \mathbf{X}\boldsymbol{\alpha}, \text{ where } \mathbf{X} = [\mathbf{x}_1 \ \cdots \ \mathbf{x}_r] \text{ and } \boldsymbol{\alpha} = \begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_r \end{bmatrix} \tag{4.5.8}$$

So what we need to find is the expansion coefficient $\boldsymbol{\alpha}$ that satisfy the orthogonal conditions.

To be specific, substitute (4.5.5) into (4.5.6) we can get

$$(\mathbf{y} - \mathbf{y}_p)^T \mathbf{x}_i = 0, \ i = 1,...,r \tag{4.5.9}$$

Then put (4.5.8) into this equation, we can get

$$(\mathbf{y} - \mathbf{X}\boldsymbol{\alpha}^*)^T \mathbf{x}_i = 0, \ i = 1,...,r \tag{4.5.10}$$

or

$$(\mathbf{y} - \mathbf{X}\boldsymbol{\alpha}^*)^T \mathbf{X} = 0 \tag{4.5.11}$$

or

$$\mathbf{y}^T \mathbf{X} = \boldsymbol{\alpha}^{*T} \mathbf{X}^T \mathbf{X} \tag{4.5.12}$$

So we can solve above equation for the optimal coefficient vector as follows.

$$\boldsymbol{\alpha}^* = (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{y} \tag{4.5.13}$$

Finally we can get the projection of y onto R as follows.

$$\mathbf{y}_p = \mathbf{X}\boldsymbol{\alpha}^* = \mathbf{X}(\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T\mathbf{y} = \mathbf{P}_R\mathbf{y} \tag{4.5.14}$$

where the matrix $\mathbf{P}_R = \mathbf{X}(\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T$ is called the orthogonal projection matrix onto subspace R. Also we note that all orthogonal projection matrices have two properties as following.

A projection matrix is symmetric: $\mathbf{P}_R = \mathbf{P}_R{}^T$.

A projection matrix is idempotent: $\mathbf{P}_R\mathbf{P}_R = \mathbf{P}_R$.

4.5.3 Orthogonal Complement

Suppose we are given r linearly independent vectors in $\Re^n$, $\mathbf{x}_1,\ldots,\mathbf{x}_r$, where r<n. Because the set R of all linear combinations of r linearly independent vectors $\mathbf{x}_1,\ldots,\mathbf{x}_r$ is a r-dimensional subspace of $\Re^n$, we can have that

R={all linear combination of $\mathbf{x}_1,\ldots,\mathbf{x}_r$}

is a subspace of $\Re^n$. Given any subspace of $\Re^n$, we can always define its orthogonal complement $R^\perp$ as

$R^\perp$={ $\mathbf{y} \in \Re^n$ such that $\mathbf{y}^T\mathbf{x} = 0$ for all $\mathbf{x} \in R$ }

That is, elements of $R^\perp$ are orthogonal to every vector of R. From $\mathbf{x} = \sum_{i=1}^{r} \alpha_i \mathbf{x}_i$ then

$$\mathbf{y}^T\mathbf{x} = \mathbf{y}^T\sum_{i=1}^{r}\alpha_i\mathbf{x}_i = \sum_{i=1}^{r}\alpha_i\mathbf{y}^T\mathbf{x}_i = 0 \tag{4.5.15}$$

Chapter 5 ANALYSIS OF MULTICHANNEL SPECTRUM AND TRANSFER FUNCTION

In this chapter, as a practical application, a analysis of measured data set of car noise is reported. The experiment condition and sensors position etc. are given in first section. Then spectral factorisation and spectral estimation discussed in previous chapter is applied to a 2 channels road noise data, which are selected from a 6 channels measured data set in car. It is confirmed here that the road noise can be whitening and uncorrelated by proposed method. Also from 2 channels input and 2 channels output measured data inside a car, the transfer function of between 2 loudspeakers and 2 microphones is identified by subspace method. These calculated results are shown in both numerical values and graphs.

5.1 Measurement of Road Noise in Car

A measurement experiment of the road noise inside a car has been made by T. J. Sutton of ISVR in 1992[11]. As a practical problem of multichannel spectral factorisation and system identification, we will use these methods described in previous chapters to this measured experimental data set.

5.1.1 Data Set of Road Noise

General experiment condition:

Vehicles: Citroen AX, was driven in 5th gear at a steady speed of 60 kph over a rough surface (large chippings)

Measurement of road noise sources

Number of sensors: 6 accelerometers.

Position of sensors: In this particular front wheel drive vehicle, all six accelerometers were placed close to the front wheel.

Filename: CIT01.mat

Accelerometer 1: RH front floor close to rear wishbone connection. (z, vertical)

Filename: CIT02.mat

Accelerometer 2: RH wishbone (body chassis). (z, vertical)

Filename: CIT03.mat

Accelerometer 3: RH hub (y, lateral)

Filename: CIT04.mat

Accelerometer 4: LH front floor close to rear wishbone connection. (z, vertical)

Filename: CIT05.mat

Accelerometer 5: LH wishbone (body chassis). (z, vertical)

Filename: CIT06.mat

Accelerometer 6: RH hub (y, lateral)

Measurement of interior noise near ears of driver

Number of sensors: 2 microphones.

Filename: CIT07.mat

Microphone 1: RH microphone (out ear position)

Filename: CIT08.mat

Microphone 2: LH microphone (out ear position)

All of 6 accelerometers and 2 microphones are sampled simultaneously with sampling rate of 1kHz, The measurement duration is 5 minutes.

5.1.2 Data Set for Interior Response

General experiment condition:

Vehicles: Citroen AX

2 interior speakers: RH speaker, LH speaker

Number of sensors: 2 microphones.

Microphone 1: RH microphone (out ear position)

Microphone 2: LH microphone (out ear position)

Filename: tf2.mat

LH speaker is driven by white noise, LH speaker and 2 microphones are sampled simultaneously with sampling rate of 1kHz. The measurement duration is 77 seconds.

Column 1	time in second
Column 2	volts to LH speaker (white noise)
Column 3	RH microphone
Column 4	LH microphone

Filename: tf3.mat

RH speaker is driven by white noise, RH speaker and 2 microphones are sampled simultaneously with sampling rate of 1kHz. The measurement duration is 77 seconds.

Column 1	time in second
Column 2	volts to RH speaker (white noise)
Column 3	RH microphone
Column 4	LH microphone

5.2 Spectral Analysis of Road Noise

5.2.1 Spectral Analysis of Road Noise

From Sutton's data set, we selected accelerometer 1 and accelerometer 4 as a two channel random signal data. By using multichannel Levinson algorithm, a 15-order innovation system with AR model can be achieved.

AR model coefficient matrices of multichannel road noise

As we have shown in (3.2.1), the autoregressive (AR) model of a random data is described by multichannel transfer function shown as following.

$$\mathbf{L}(z) = \mathbf{D}^{-1}(z) = \left(\mathbf{I}_p + \sum_{i=1}^{m} \mathbf{A}(i) z^{-i} \right)^{-1}$$

By a MATLAB program, The coefficient matrices are obtained as following.

$$\mathbf{A}(1) = \begin{bmatrix} -1.6099 & -0.1517 \\ -0.2895 & -1.4837 \end{bmatrix} \qquad \mathbf{A}(2) = \begin{bmatrix} 2.6712 & 0.1047 \\ 0.2973 & 2.4969 \end{bmatrix}$$

$$\mathbf{A}(3) = \begin{bmatrix} -3.7685 & -0.1069 \\ -0.4674 & -3.3332 \end{bmatrix} \qquad \mathbf{A}(4) = \begin{bmatrix} 4.7637 & 0.0655 \\ 0.6265 & 4.1467 \end{bmatrix}$$

$$\mathbf{A}(5) = \begin{bmatrix} -5.3983 & 0.0336 \\ -0.6336 & -4.4727 \end{bmatrix} \qquad \mathbf{A}(6) = \begin{bmatrix} 5.6054 & -0.2568 \\ 0.5952 & 4.4061 \end{bmatrix}$$

$$\mathbf{A}(7) = \begin{bmatrix} -5.3749 & 0.3880 \\ -0.5717 & -3.9267 \end{bmatrix} \qquad \mathbf{A}(8) = \begin{bmatrix} 4.9354 & -0.5366 \\ 0.4432 & 3.3777 \end{bmatrix}$$

$$\mathbf{A}(9) = \begin{bmatrix} -0.40831 & 0.5480 \\ -0.3692 & -2.5456 \end{bmatrix} \qquad \mathbf{A}(10) = \begin{bmatrix} 3.1912 & -0.5711 \\ 0.2269 & 1.7629 \end{bmatrix}$$

$$\mathbf{A}(11) = \begin{bmatrix} -2.2475 & 0.4593 \\ -0.1773 & -1.1170 \end{bmatrix} \qquad \mathbf{A}(12) = \begin{bmatrix} 1.5430 & -0.3749 \\ 0.0746 & 0.6944 \end{bmatrix}$$

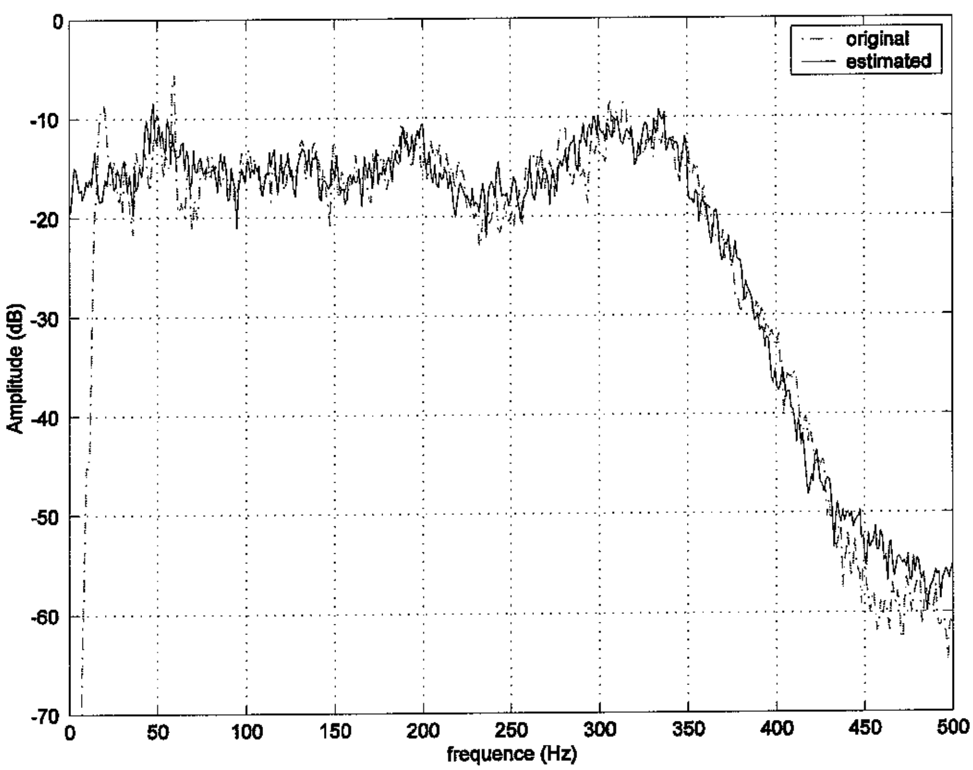
$$\mathbf{A}(13)=\begin{bmatrix}-0.8320 & 0.2184\\ -0.0797 & -0.2692\end{bmatrix} \qquad \mathbf{A}(14)=\begin{bmatrix}0.4071 & -0.1458\\ 0.0095 & 0.1368\end{bmatrix}$$

$$\mathbf{A}(15)=\begin{bmatrix}-0.1820 & 0.0900\\ -0.0121 & -0.0176\end{bmatrix}$$

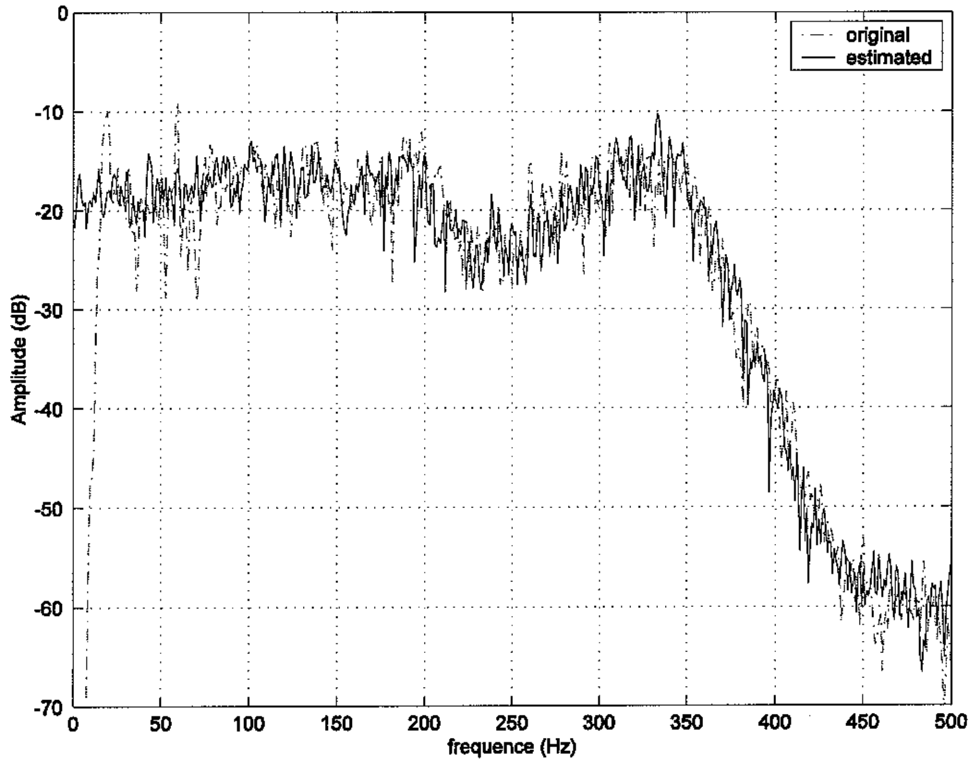
Also a covariance matrix $\mathbf{\Sigma}$ for input white noise vector is obtained as following.

$$\mathbf{\Sigma}=\begin{bmatrix}46.5340 & 8.5870\\ 8.5870 & 52.94\end{bmatrix}$$

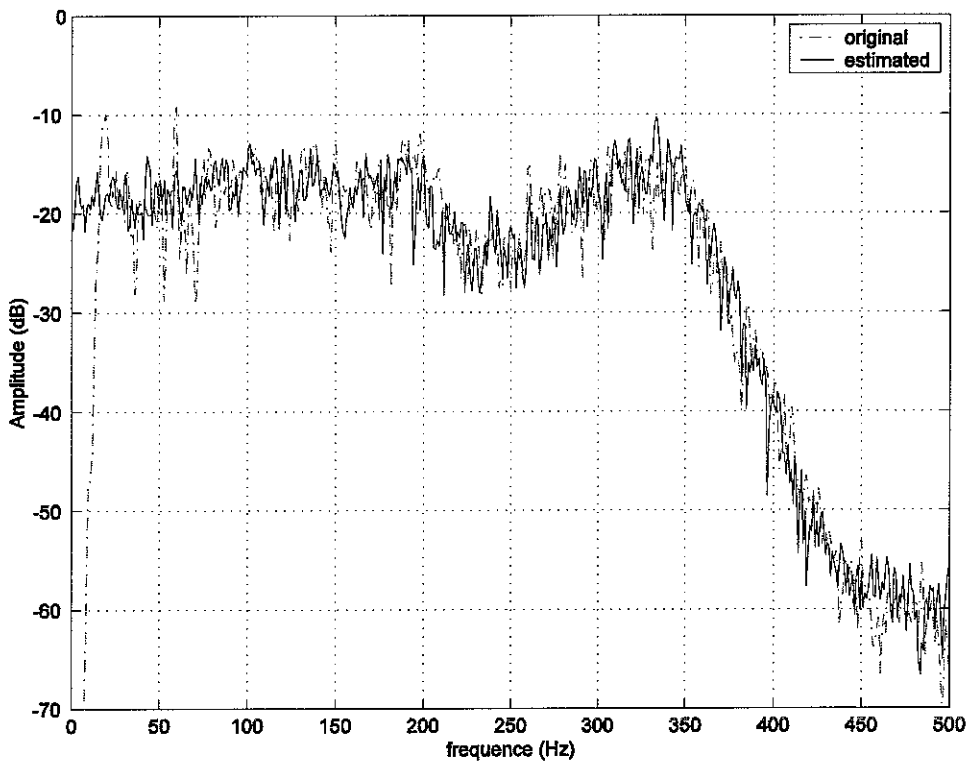
Comparison graph of spectral density between original signals and estimated signals.



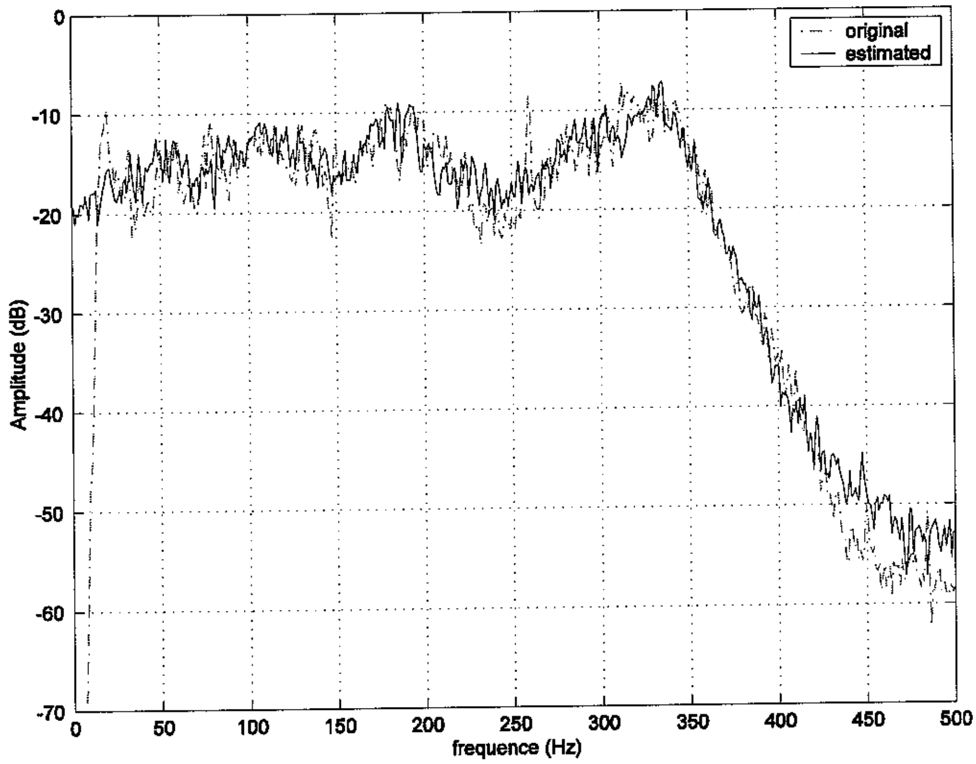
• Figure 25 Power spectral density of x1



• Figure 26 Cross power spectral density of x1 and x2

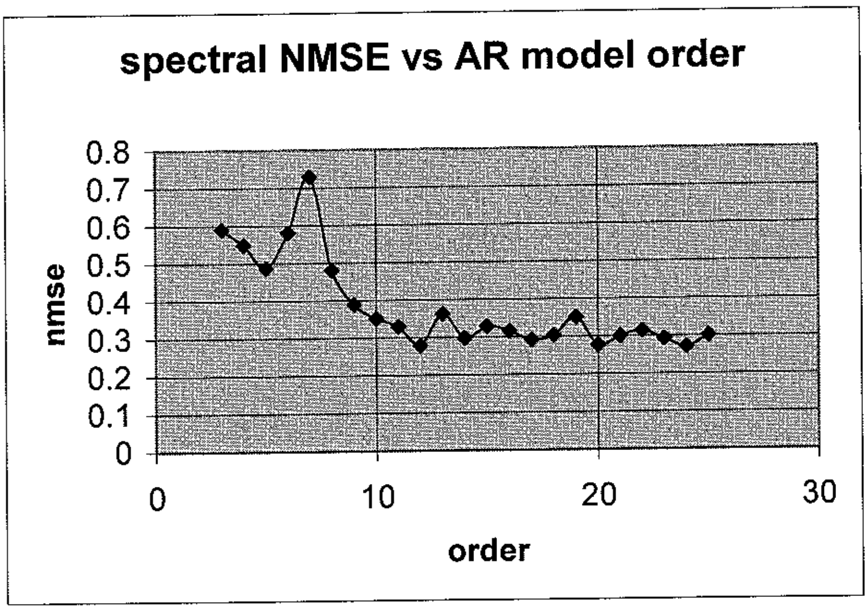


• Figure 27 Cross power spectral density of x2 and x1



• Figure 28 Power spectral density of x2

In order to give fine selection of system order, the graph of normalized mean square error (NMSE) vs system order is given as following.



• Figure 29 Normalized mean square error (NMSE) vs AR model order

5.2.2 Whitening of Road Noise

As we have discussed in section 3.4, by multichannel Levinson algorithm, for a given multichannel random signal, we can express its spectral as following.

$$\begin{aligned}\mathbf{S}_{xx}(z) &= \mathbf{L}(z)\mathbf{\Sigma}\mathbf{L}_{*}(z) \\ &= \mathbf{L}(z)\mathbf{G}\mathbf{S}_{uu}(z)\mathbf{G}^{T}\mathbf{L}_{*}(z) \\ &= \mathbf{A}^{-1}(z)\mathbf{G}\mathbf{S}_{uu}(z)\mathbf{G}^{T}\mathbf{A}_{*}^{-1}(z)\end{aligned}$$

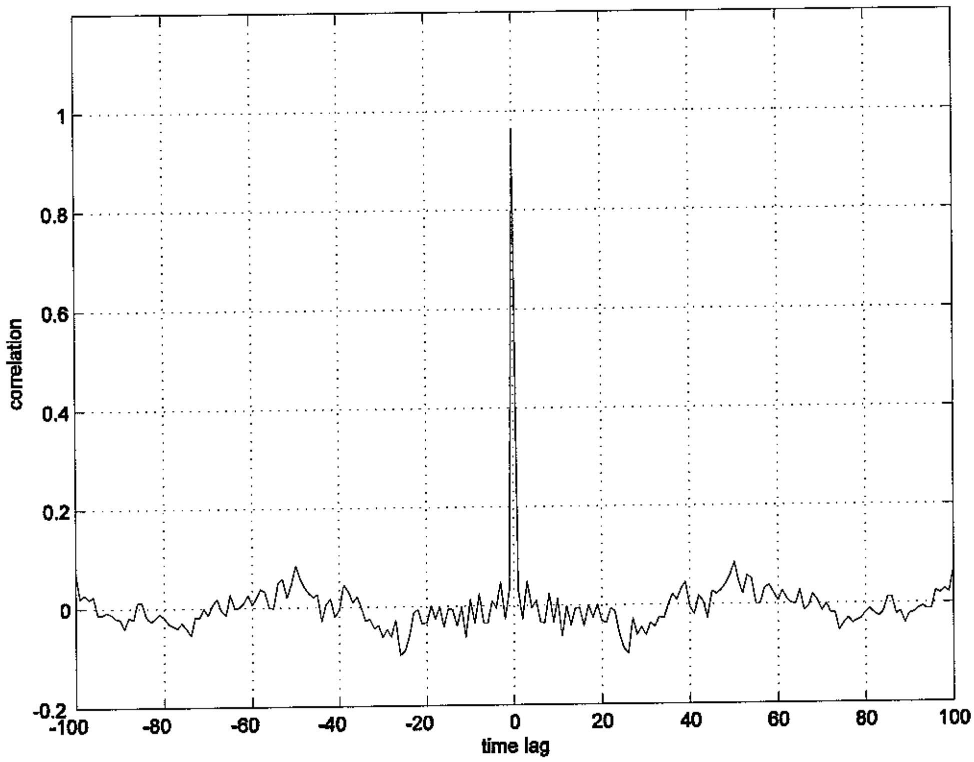
where

$$\mathbf{S}_{uu}(z) = \mathbf{I}$$

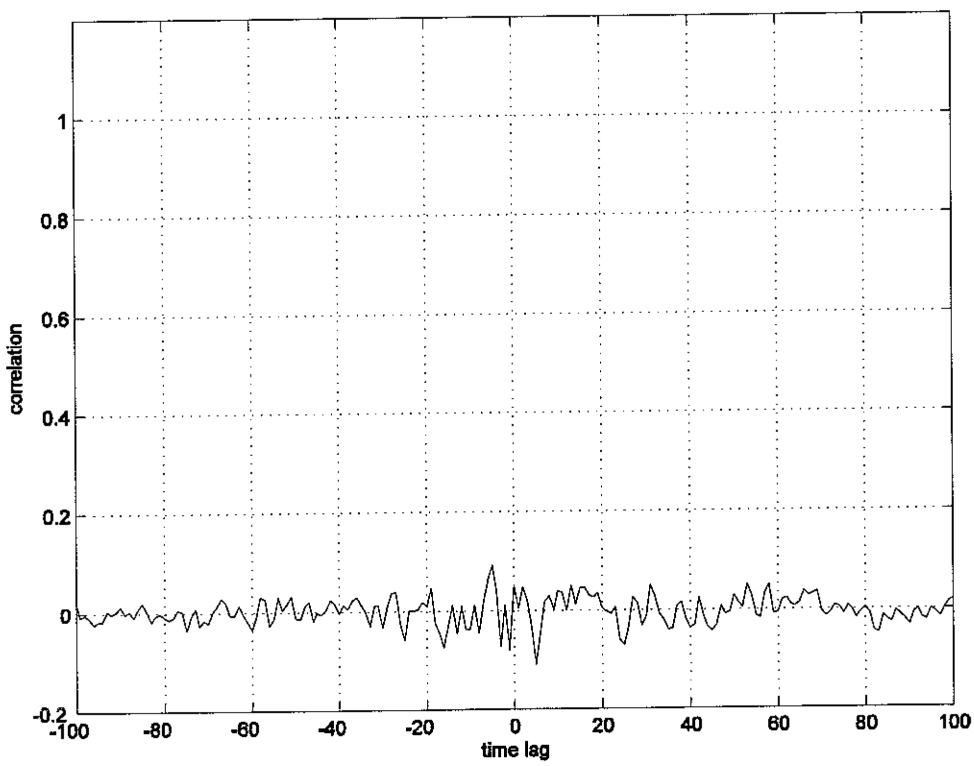
So by using an inverse system of calculated AR model, we can get a uncorrelated white noise with unity cross-spectral density matrix as following.

$$\begin{aligned}\mathbf{S}_{uu}(z) &= \mathbf{G}^{-1}\mathbf{A}(z)\mathbf{S}_{xx}(z)\mathbf{A}_{*}(z)\mathbf{G}^{T^{-1}} \\ &= \mathbf{I}\end{aligned}$$

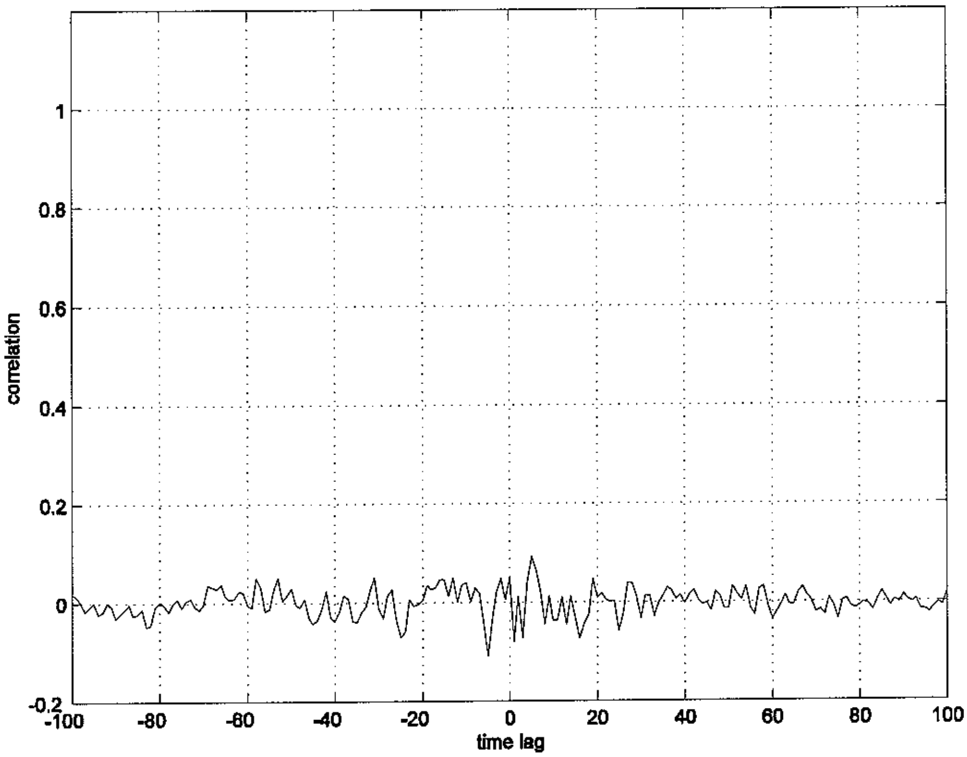
By a MATLAB program, for given 2 channel road noise, we can get a 2 channel whitening output signal, that have no correlation each other as shown in following Figure 30 ~ Figure 33.



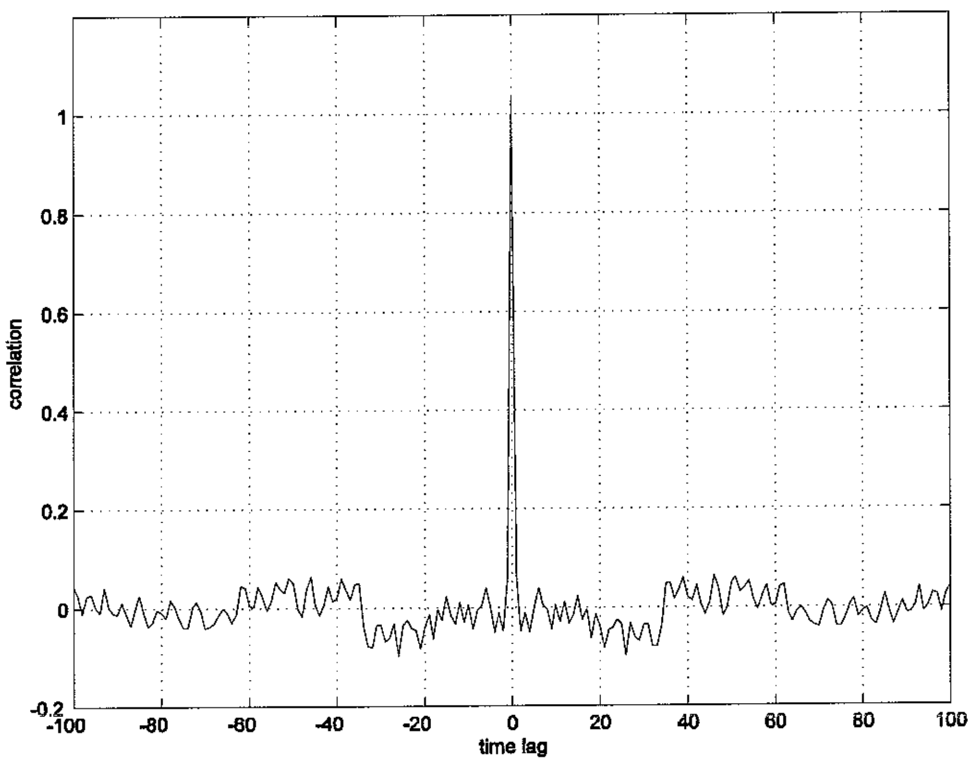
• Figure 30 Autocorrelation of u1



• Figure 31 Cross-correlation between u1 and u2



• Figure 32 Cross-correlation between u2 and u1



• Figure 33 Autocorrelation of u2

5.3 Interior Transfer Function in Car

5.3.1 System Identification by Subspace Method

From Sutton's data set for interior transfer function, we can calculate the interior response between loudspeakers and microphones inside the car. By a MATLAB program, we can get coefficient matrices of state space equation of the interior response, these results are shown as following.

System order=15

A matrix

Columns 1 through 8

0.6990	-0.7083	-0.0447	0.0186	0.0468	-0.0721	0.0444	0.0272
0.6600	0.5977	0.0842	-0.1435	-0.2862	0.2334	-0.1292	-0.0616
-0.0049	-0.0215	0.5606	-0.6745	0.3082	-0.2425	-0.1365	-0.1442
-0.0078	-0.0665	0.6099	0.5874	-0.0428	0.0060	-0.4649	0.1275
-0.0041	-0.0612	-0.2202	-0.0223	0.6104	0.6331	-0.4049	-0.1149
-0.0211	-0.0561	0.1082	-0.3469	-0.2720	0.3240	0.0341	0.5904
0.0158	0.0606	0.1470	0.0723	0.3929	0.1298	0.4036	0.5383
0.0061	0.0159	-0.0549	-0.0330	0.1415	-0.3391	-0.3526	0.3317
0.0026	0.0010	-0.0008	-0.0096	0.0822	-0.1284	0.0631	-0.2470
0.0001	0.0064	-0.0194	-0.0078	-0.0043	-0.0128	-0.1666	0.1062
0.0000	-0.0054	-0.0055	0.0185	-0.0107	0.0037	0.0943	0.1532
0.0007	0.0106	0.0073	-0.0087	-0.0077	0.0409	-0.1082	-0.0309
-0.0003	-0.0045	0.0036	0.0093	0.0065	-0.0058	0.0672	-0.0160
0.0001	0.0038	-0.0162	0.0072	-0.0334	0.0290	-0.0652	0.1569
0.0006	-0.0025	-0.0005	0.0147	0.0123	-0.0101	0.0651	0.0525

Columns 9 through 15

0.0044	0.0070	-0.0108	0.0077	0.0069	-0.0002	-0.0011
-0.0047	0.0704	0.0114	0.0264	-0.0132	0.1360	0.0099
0.0213	0.1035	0.0097	0.0668	0.0159	-0.0219	-0.0669
0.0976	-0.0548	-0.0920	0.0747	-0.0097	-0.0324	-0.0976
-0.0377	-0.1271	0.0125	0.0074	-0.0533	0.0087	-0.0394
0.0965	-0.3257	-0.0517	-0.2394	0.0049	-0.2315	-0.1240
0.1438	0.2705	-0.1751	0.1885	0.0413	0.3076	0.0481
-0.4546	-0.2997	0.1209	-0.1734	-0.0964	0.2662	0.2573
0.6090	-0.6391	-0.1587	0.0364	-0.0856	0.1539	0.1392

0.5440	0.3813	0.5760	-0.3819	0.0091	0.0289	0.1619
-0.0310	-0.3054	0.6841	0.5647	-0.0524	0.0528	-0.1644
0.0265	-0.0713	-0.0175	0.3344	0.7512	-0.3828	0.4994
-0.0593	-0.0958	0.1028	-0.3964	0.6372	0.5607	-0.5080
0.1152	-0.1227	-0.2462	0.0612	-0.0292	0.2416	0.0865
-0.0293	-0.0828	0.0387	-0.2941	0.0298	-0.2901	0.4095

B matrix

17.0955	-24.7914
401.8829	-409.2272
46.1383	-223.0384
56.3427	99.0702
-210.9566	16.2255
285.6818	-258.9477
-224.9644	333.8293
208.4473	283.3429
58.3675	80.3301
-273.3494	-27.5378
-44.4296	-97.9605
-116.9186	-91.0131
-34.7278	56.1166
-67.4704	-66.8519
-96.9916	86.0788

C matrix

Columns 1 through 8

-0.2183	-0.2666	-0.2483	-0.2691	-0.3931	0.0486	-0.4259	0.1147
0.1765	0.2678	-0.3989	0.0381	0.2742	-0.4703	-0.2387	0.2794

Columns 9 through 15

0.1476	0.2371	-0.1586	0.2576	0.0157	0.3043	-0.1014
0.2089	0.0349	-0.1556	0.0846	0.0670	-0.1935	-0.2400

D matrix

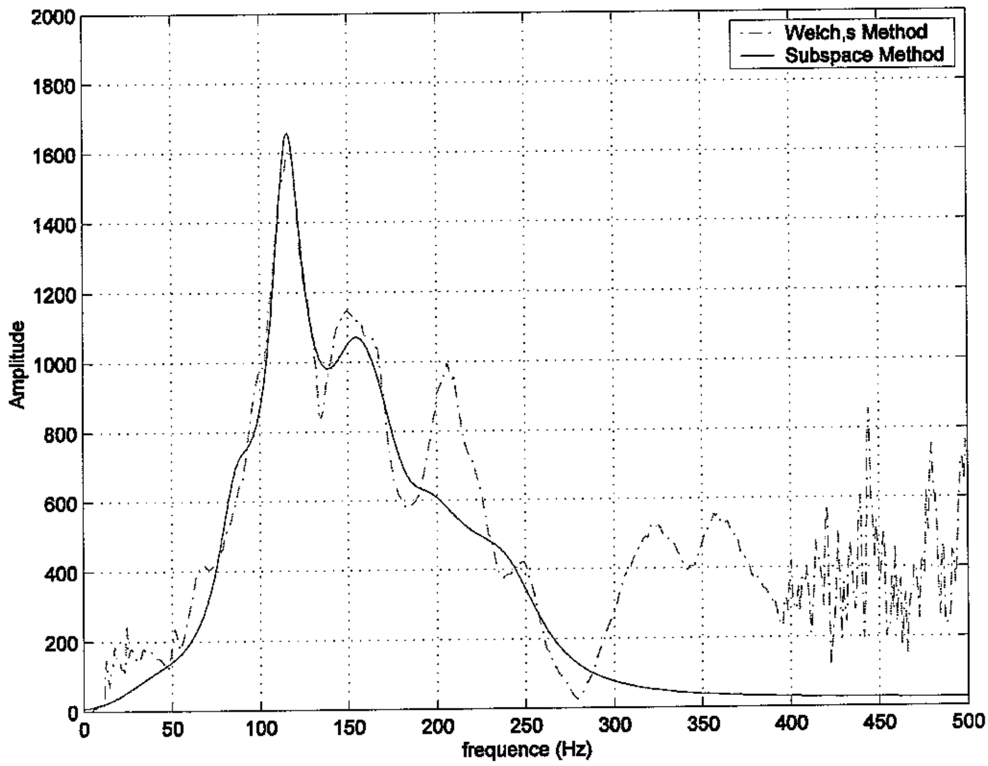
0	0
0	0

K matrix

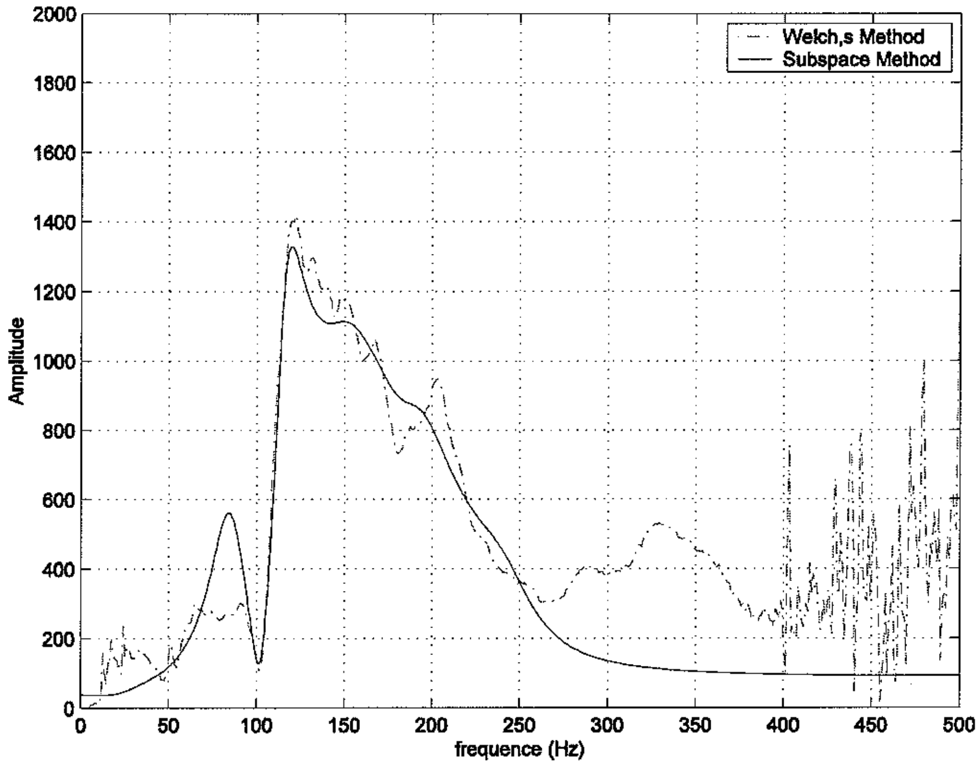
0.0379	-0.1698
-0.2599	-0.0865
0.0527	-0.4502
-0.0146	-0.1291
-0.1275	0.0179
-0.0467	-0.1833
-0.2151	-0.0564
0.0447	0.1332

0.0510	0.0015
-0.1051	0.2105
-0.0672	-0.1091
0.1062	-0.0983
0.0851	0.0383
0.0393	0.0214
-0.1084	0.0164

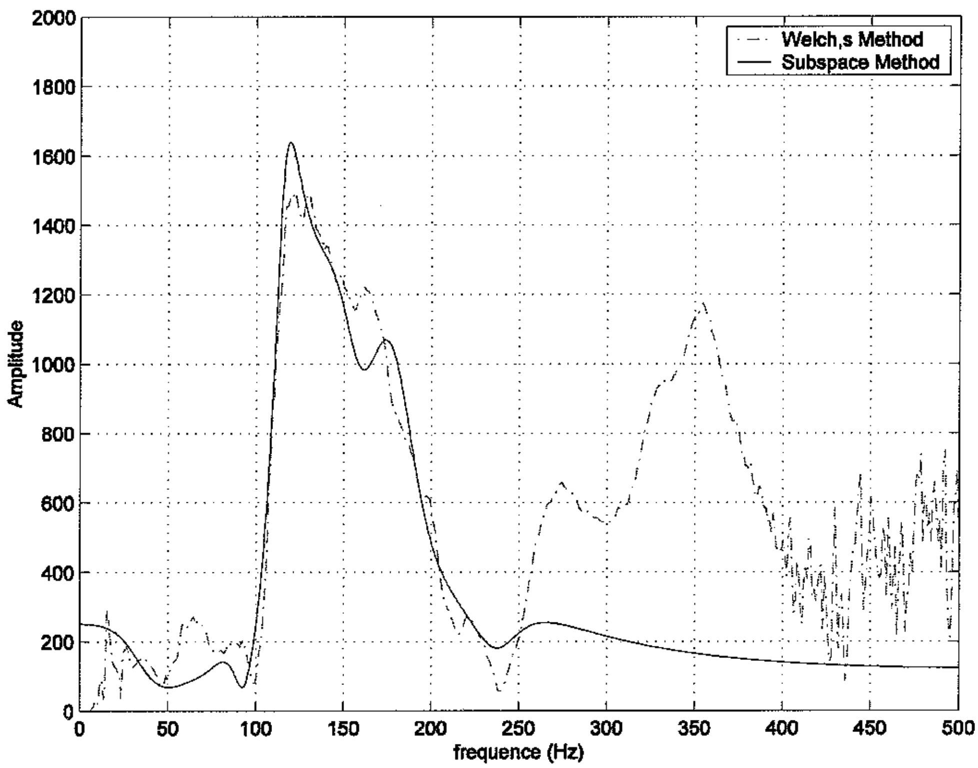
Comparison of transfer function



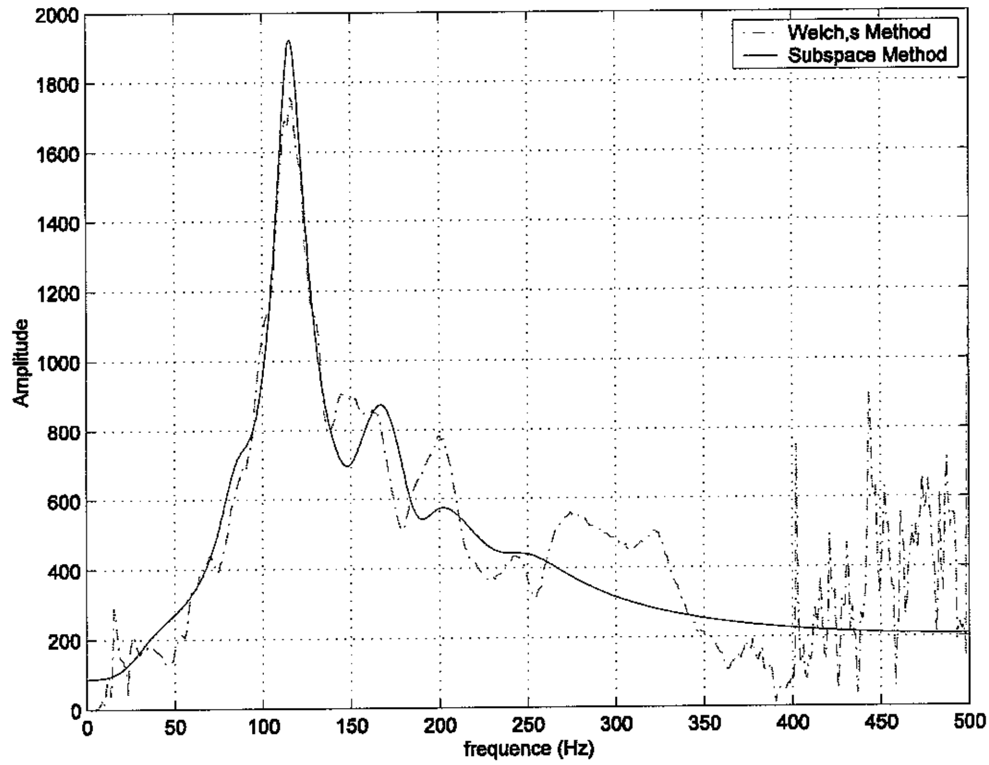
• Figure 34 transfer function from loudspeaker 1 to microphone 1



• Figure 35 transfer function from loudspeaker 1 to microphone 2



• Figure 36 transfer function from loudspeaker 2 to microphone 1



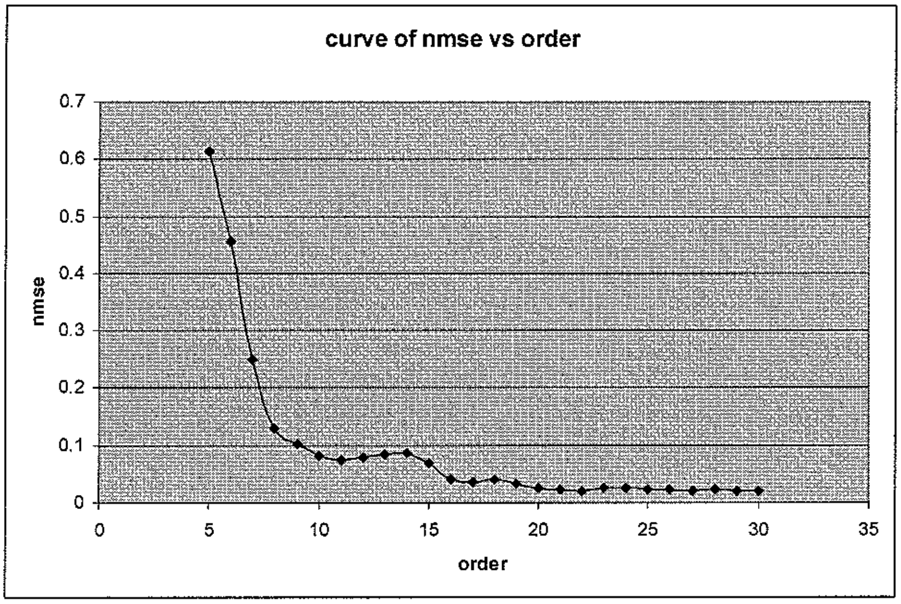
• Figure 37 transfer function from loudspeaker 2 to microphone 2

Fine system order selection

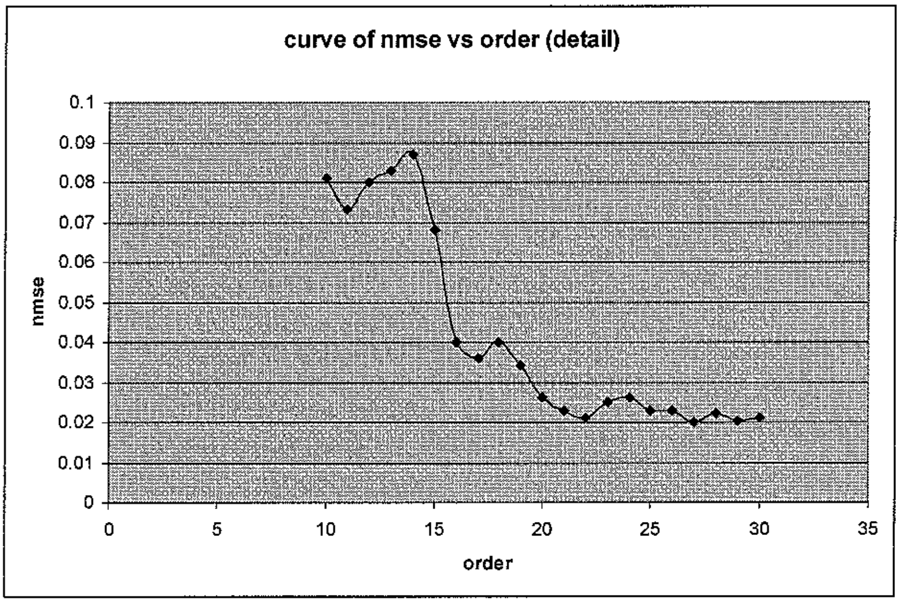
In order to give a fine method for system order selection, we calculated NMSE(normalized mean square error) , which is defined as following, for each system order.

$$nmse = \frac{\sum_{c=1}^{p}\sum_{i=1}^{m}(y_{ci} - \hat{y}_{ci})^2}{\sum_{c=1}^{p}\sum_{i=1}^{m} y_i^2}, \text{ with } p = 2, m = 10000$$

The results is shown in Figure 38 ~ Figure 39. These results show that it is better to select system order n=10 for a rough precision, and it is better to select system order n=20 for a fine precision,



• Figure 38 Curve of NMSE normalized mean square error) vs system order



• Figure 39 Detail curve of NMSE normalized mean square error) vs system order

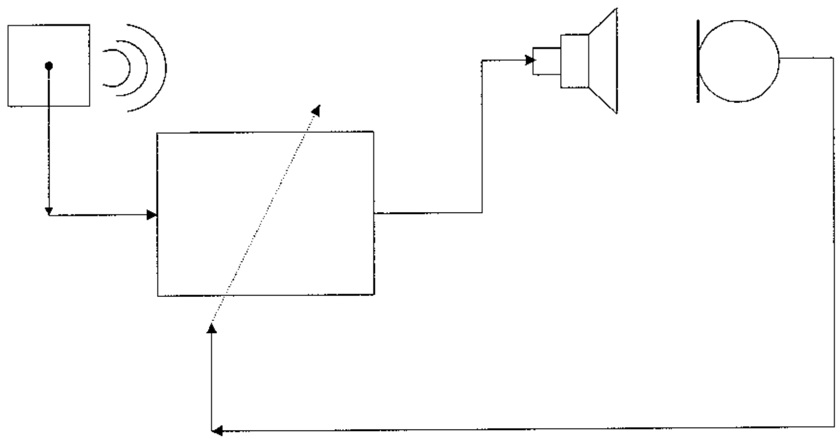
Chapter 6 APPLICATION TO MULTICHANNEL ACTIVE NOISE CONTROL SYSTEM

In this chapter, based on these results of whitening of multichannel input signals and system identification of control plant, an active noise control system is constructed. For the assume of state space model of control plant, a adaptive control algorithm is derived based on steep-descent method, the performance improvement comparing with conventional method is discussed by a computer simulation.

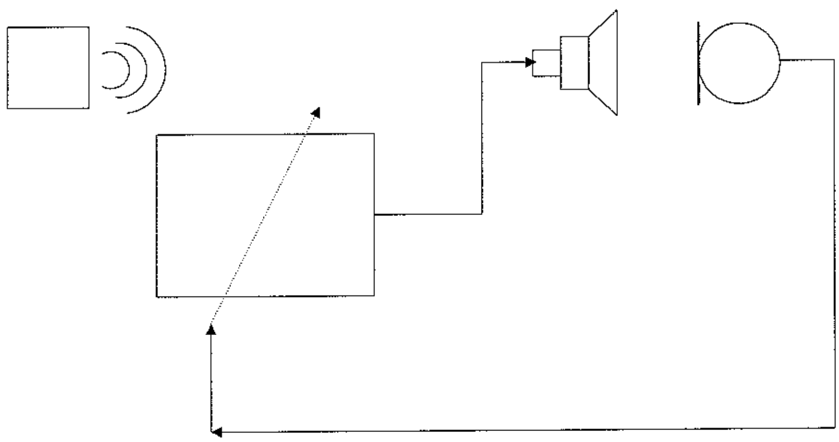
6.1 Active Noise Control: Principle and Background

Active noise control (ANC) involves an electroacoustic system that cancels the unwanted primary noise based on the principle of superposition; specifically, an antinoise of equal amplitude and opposite phase is generated and combined with the primary noise, thus resulting in the cancellation of both noises. The first work of ANC system is reported by Lueg in 1936[12]. After that by H. F. Olson and E. G. May, which uses a carefully designed amplifier marched to the response of the error sensor and secondary source in 1953[13]. Also a feedback system is reported by Conover as a means of reducing a transformer noise in 1956[14]. Since the characteristics of noise source and the environment are varying with time, therefore an ANC system must be adaptive in order to cope with variation. A modern type of ANC system utilizing adaptive filter, which adjusts their coefficients to minimise an error signal as a controller, has been developed for noise cancellation in air-conditioning duct system[15,16]. After that ANC based on adaptive filter theory is rapidly becoming the most effective method to reduce the noises that can otherwise be very difficult and expensive to control. ANC system is usually classified into two categories: feedforward system and feedback system, as shown in Figure 40(a) and 40(b).

The feedforward system requires a reference signal which is strongly related to the primary noise. The reference signal may be taken from the source itself or by measurement of the primary noise field. This signal is passed through an adaptive controller and on to the secondary loudspeaker. An error microphone measures residual noise field due to the primary and secondary sources combined, and the controller is adjusted to obtain the best noise reduction at the error microphone. Any adjustment of controller coefficients can happen slowly and so the stability problems with feedback systems are avoided. Thus it is not necessary to put secondary loudspeaker close to the error microphone. The drawback with feedforward approach is needed to obtain reference signal which are well related to the primary noise to be cancelled.



(a)



(b)

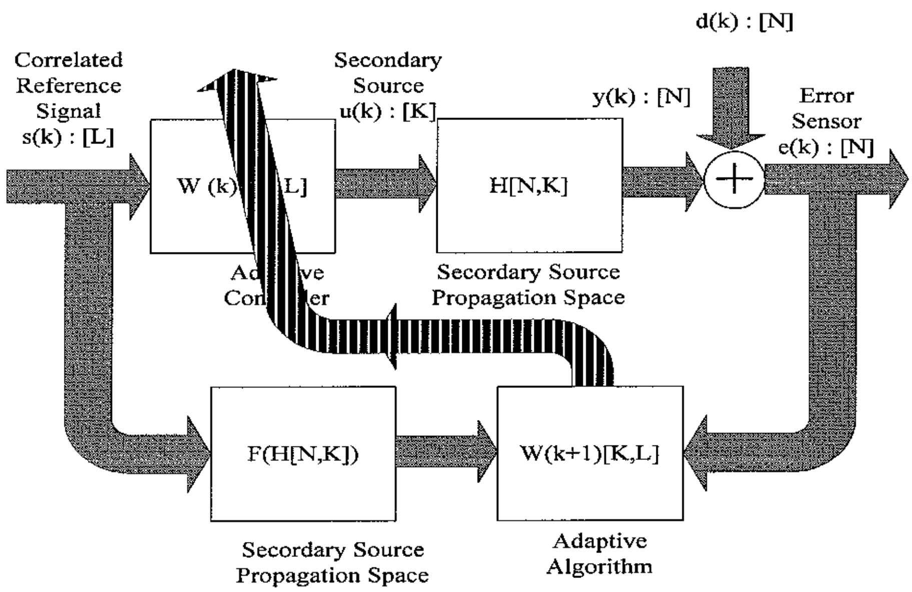
• Figure 40 Two categories of ANC system (a) Feedforward system (b) Feedback system

The feedback system not requires a reference signal measured from the primary noise. i.e., it only uses an error microphone to measure the residual noise. The controller uses adjustable negative gain feedback loop to drive the acoustic pressure at error microphone to zero through secondary loudspeaker. A significant advantage of this strategy is that no prior knowledge is required of the primary noise to be cancelled. The main disadvantage is that for loop stability the secondary source needs to be very close to the error microphone. Many applications are developed one after another, such as reduction of noise radiating from the outlet of an air-conditioning duct, active electronic muffler for automobile exhaust noise control, and active headset.

6.2 Multichannel Feedforward Active Noise Control System

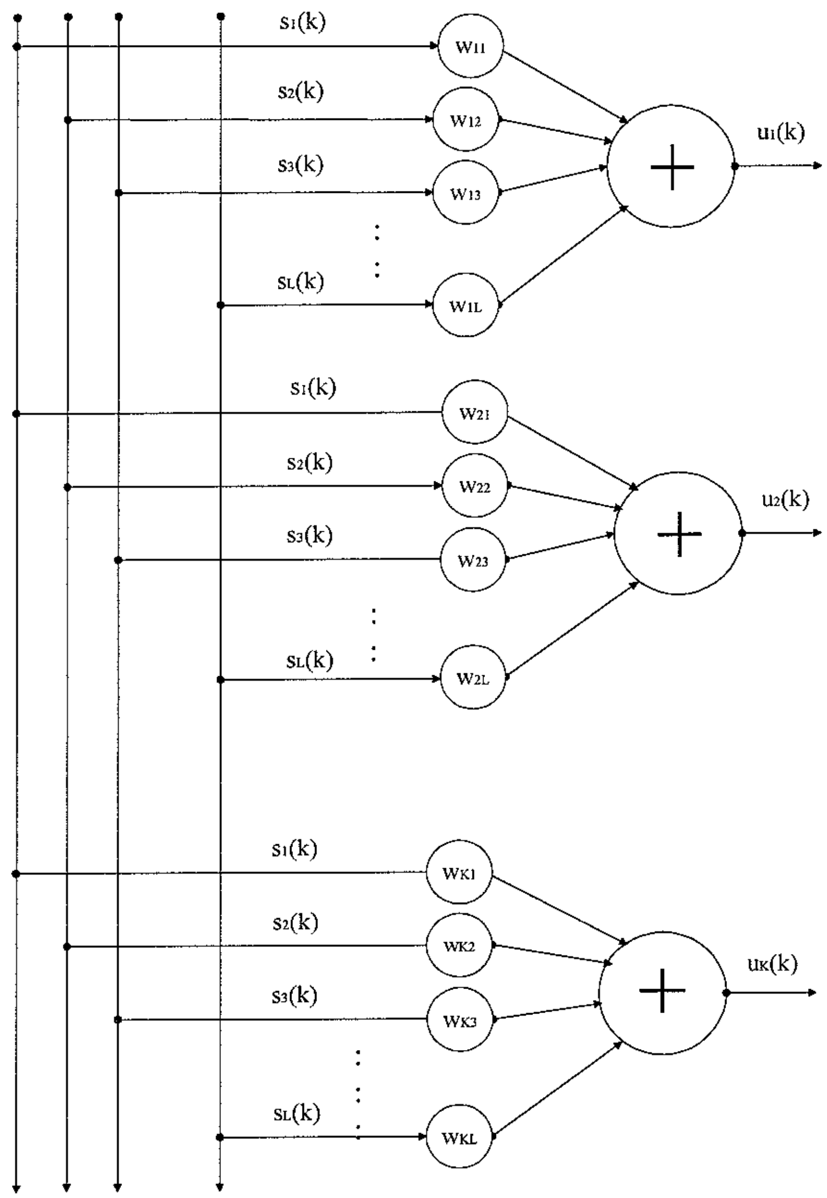
6.2.1 Basic Architecture

For the noise cancelling problem in an enclosure, in order to achieve satisfactory performance in a large dimension space, a number of cancelling loudspeakers and a number of error sensors are used, i.e., a multichannel ANC system is necessary. The feedforward ANC system consists of a number of reference microphones, a number of error microphones and a number of cancelling loudspeakers, and a multiple input multiple output adaptive controller which weight parameter matrix is able to be updated by an adaptive algorithm. The block diagram of basic feedforward ANC system is illustrated in Figure 41.



• Figure 41 Block diagram of feedforward ANC system

In order to cancel the primary noise d(k), adaptive controller gives a secondary source by weighting each reference signal. The secondary source is passed to plant, which is acoustic propagation space, then reach the error sensors, where its combined with the primary noise, thus resulting in the cancellation of both noises. The adaptive controller is implemented by a multiple input multiple output adaptive linear combiner with adjustable weight parameter as shown in Figure 42.



• Figure 42 Diagram of multiple input multiple output adaptive linear combiner

From L-input of reference signals, the $\mathbf{s}(k) = [s_1 \quad s_2 \quad \cdots \quad s_L]^T$ adaptive controller gives K-output $\mathbf{u}(k) = [u_1 \quad u_2 \quad \cdots \quad u_K]^T$ by combine each weight reference signal

$$
\begin{aligned}
u_1 &= \mathbf{w}_{11} * s_1 + \mathbf{w}_{12} * s_2 + \cdots + \mathbf{w}_{1L} * s_L \\
u_2 &= \mathbf{w}_{21} * s_1 + \mathbf{w}_{22} * s_2 + \cdots + \mathbf{w}_{2L} * s_L \\
&\cdots \\
u_K &= \mathbf{w}_{K1} * s_1 + \mathbf{w}_{K2} * s_2 + \cdots + \mathbf{w}_{KL} * s_L
\end{aligned}
\tag{6.2.1}
$$

By defining a $K \times L$ weight parameter matrix as following.

$$
\mathbf{W} = \begin{bmatrix} \mathbf{w}_{11} & \mathbf{w}_{12} & \cdots & \mathbf{w}_{1L} \\ \mathbf{w}_{21} & \mathbf{w}_{22} & \cdots & \mathbf{w}_{2L} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{w}_{K1} & \mathbf{w}_{K2} & \cdots & \mathbf{w}_{KL} \end{bmatrix}
\tag{6.2.2}
$$

where $\mathbf{w}_{ij} = \lfloor w_{ij1} \quad w_{ij2} \quad \cdots \quad w_{ijM} \rfloor (i = 1,2...K, j = 1,2...L)$. This equation can be write into a matrix from as following.

$$
\mathbf{u}(k) = \mathbf{W} * \mathbf{s}(k) \tag{6.2.3}
$$

The sound propagation space for secondary sources is described by multiple input multiple output state equation plant model given by following equation.

$$
\begin{aligned}
\mathbf{x}(k+1) &= \mathbf{Ax}(k) + \mathbf{Bu}(k) \\
\mathbf{y}(k) &= \mathbf{Cx}(k) + \mathbf{Du}(k) \\
\mathbf{x}(0) &= \mathbf{0}_M
\end{aligned}
\tag{6.2.4}
$$

where x(k) is a M degree state vector and y(k) is a N degree output vector and A, B, C, D are coefficient matrix of state equation. When a stochastic noise vector d(k) is added, the ANC system makes cancellation by output vector y(k), and the residual error noise vector e(k) will be detected by error sensors, it is given by

$$
\mathbf{e}(k) = \mathbf{d}(k) - \mathbf{y}(k) \tag{6.2.5}
$$

Suppose that the cost function is defined by mean-square-summation of error signals as following.

$$
J = E\left[\sum_{m=1}^{N} e_m^2(k)\right] = E\left[\mathbf{e}(k)^T \mathbf{e}(k)\right] \tag{6.2.6}
$$

Our problem is to minimize this cost function by adjusting weight parameters of the adaptive controller.

6.2.2 Equivalent System

Here we will give an equivalent system description for feedforward ANC system shown in Figure 41. If we rewrite the weight parameter matrix into a vector form as following,

$$\mathbf{w} = [\mathbf{w}_{11} \quad \mathbf{w}_{12} \quad \cdots \quad \mathbf{w}_{1L} \quad \mathbf{w}_{21} \quad \mathbf{w}_{22} \quad \cdots \quad \mathbf{w}_{2L} \quad \cdots \quad \mathbf{w}_{K1} \quad \mathbf{w}_{K2} \quad \cdots \quad \mathbf{w}_{KL}]^T \qquad (6.2.7)$$

and if we rewrite input signal vector into a matrix form as following,

$$\mathbf{S}(k) = \begin{bmatrix} s_1 & s_2 & \cdots & s_L & 0 & 0 & \cdots & 0 & \cdots & 0 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 & s_1 & s_2 & \cdots & s_L & \cdots & 0 & 0 & \cdots & 0 \\ & \vdots & & & & \vdots & & & \ddots & & \vdots & & \\ 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & \cdots & s_1 & s_2 & \cdots & s_L \end{bmatrix} \qquad (6.2.8)$$

we can get a alternative expression of (6.2.1).

$$\mathbf{u}(k) = \mathbf{S}(k)\mathbf{w} \qquad (6.2.9)$$

From the state space equation plant model, by denote that

$$\mathbf{H}_0 = \mathbf{D} \qquad (6.2.10)$$

$$\mathbf{H}_l = \mathbf{C}\mathbf{A}^{l-1}\mathbf{B}, \text{ where } l = 1,2,3...k \qquad (6.2.11)$$

we can give the output vector $\mathbf{y}(k)$ of state equation as following.

$$\begin{aligned} \mathbf{y}(k) &= [y_1 \quad y_2 \quad \cdots \quad y_N]^T = \sum_{l=0}^{k} \mathbf{H}_l \mathbf{u}(k-l) \\ &= \sum_{l=0}^{k} \mathbf{H}_l \mathbf{S}(k-l)\mathbf{w} = \left[\sum_{l=0}^{k} \mathbf{H}_l \mathbf{S}(k-l) \right] \mathbf{w} \end{aligned} \qquad (6.2.12)$$

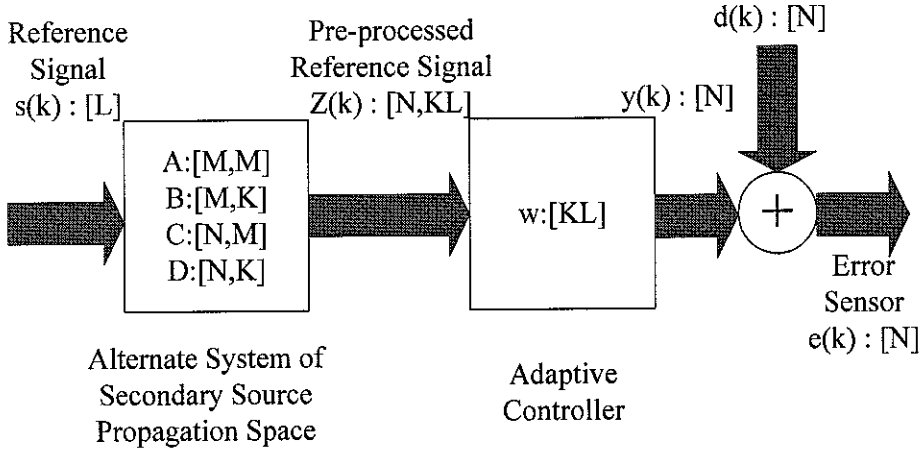
Here if we denote that

$$\mathbf{Z}(k) = \sum_{l=0}^{k} \mathbf{H}_l \mathbf{S}(k-l) \qquad (6.2.13)$$

then the output vector $\mathbf{y}(k)$ is described in matrix form as following.

$$\mathbf{y}(k) = \mathbf{Z}(k)\mathbf{w} \tag{6.2.14}$$

where $\mathbf{Z}(k)$ is a $N \times KL$ matrix. This equation shows that output vector $\mathbf{y}(k)$ can be given by a production of a pre-processed signal matrix of reference signal vector s(k) and the weight parameter vector of adaptive controller. So that it can be considered as a pre-process adaptive filter with KL weight parameters. The block diagram of this equivalent system is shown in Figure 43.



• Figure 43 Equivalent block diagram of Figure 41

The alternate system of secondary source propagation space can be expressed by following extended state equation.

$$\begin{aligned}\mathbf{V}(k+1) &= \mathbf{A}\mathbf{V}(k) + \mathbf{B}\mathbf{S}(k)\\ \mathbf{Z}(k) &= \mathbf{C}\mathbf{V}(k) + \mathbf{D}\mathbf{S}(k)\\ \mathbf{V}(0) &= \mathbf{0}_{M\times KL}\end{aligned} \tag{6.2.15}$$

where $\mathbf{S}(k)$ is a $K \times KL$ matrix and $\mathbf{V}(k)$ is a $M \times KL$ matrix. By using Kronecker product, this equation is able be expressed as following also.

$$\begin{aligned}\mathbf{V}(k+1) &= \mathbf{A}\mathbf{V}(k) + \mathbf{B} \otimes \mathbf{s}(k)\\ \mathbf{Z}(k) &= \mathbf{C}\mathbf{V}(k) + \mathbf{D} \otimes \mathbf{s}(k)\\ \mathbf{V}(0) &= \mathbf{0}_{M\times KL}\end{aligned} \tag{6.2.16}$$

6.2.3 Optimal Solution

From Figure 43, the multiple input multiple output ANC problem is replaced to a pre-process adaptive filter problem, so we can get Wiener solution easily to show optimal weight parameters and minimum value of cost function. By substitute (6.2.5) and (6.2.14) into (6.2.6), the cost function can be expressed as following.

$$
\begin{aligned}
\mathbf{J} &= E\left[\sum_{m=1}^{N} e_m^2(k)\right] = E\left[\mathbf{e}^T\mathbf{e}\right] \\
&= E\left[(\mathbf{d}-\mathbf{y})^T(\mathbf{d}-\mathbf{y})\right] \\
&= E\left[\mathbf{d}^T\mathbf{d} - \mathbf{d}^T\mathbf{y} - \mathbf{y}^T\mathbf{d} + \mathbf{y}^T\mathbf{y}\right] \\
&= E\left[\mathbf{d}^T\mathbf{d}\right] - E\left[\mathbf{d}^T\mathbf{y}\right] - E\left[\mathbf{y}^T\mathbf{d}\right] + E\left[\mathbf{y}^T\mathbf{y}\right] \\
&= E\left[\mathbf{d}^T\mathbf{d}\right] - 2E\left[\mathbf{d}^T\mathbf{Z}\right]\mathbf{w} + \mathbf{w}^T E\left[\mathbf{Z}^T\mathbf{Z}\right]\mathbf{w}
\end{aligned} \tag{6.2.17}
$$

Here, we denote that

$$\mathbf{P}_d = E\left[\mathbf{d}^T\mathbf{d}\right] \tag{6.2.18}$$

$$\mathbf{R}_{Zd} = E\left[\mathbf{d}^T\mathbf{Z}\right] \tag{6.2.19}$$

$$\mathbf{R}_{ZZ} = E\left[\mathbf{Z}^T\mathbf{Z}\right] \tag{6.2.20}$$

Then we can rewrite cost function as following,

$$\mathbf{J} = \mathbf{P}_d - 2\mathbf{R}_{Zd}\mathbf{w} + \mathbf{w}^T\mathbf{R}_{ZZ}\mathbf{w} \tag{6.2.21}$$

which is a quadratic function of weight parameter vector. The differentiation of cost function with respect to weight parameter vector can be derived as following.

$$\frac{\partial \mathbf{J}}{\partial \mathbf{w}} = -2\mathbf{R}_{Zd} + 2\mathbf{R}_{ZZ}\mathbf{w} \tag{6.2.22}$$

From $\frac{\partial \mathbf{J}}{\partial \mathbf{w}} = 0$, we can obtain an expression related to optimal weight parameter vector as following.

$$\mathbf{R}_{ZZ}\mathbf{w}^* = \mathbf{R}_{Zd} \tag{6.2.23}$$

Suppose that $\mathbf{R}_{ZZ}$ is a positive definite matrix, we can get the optimal weight parameter vector, which is expressed as following.

$$\mathbf{w}^* = \mathbf{R}_{ZZ}^{-1}\mathbf{R}_{Zd} \qquad (6.2.24)$$

when weight parameter vector take its optimal value, the cost function takes its minimum value as shown as following.

$$\mathbf{J}_{\min} = \mathbf{P}_d - \mathbf{R}_{Zd}{}^T\mathbf{R}_{ZZ}^{-1}\mathbf{R}_{Zd} \qquad (6.2.25)$$

And the cost function is also able to be shown as following.

$$\mathbf{J} = \mathbf{J}_{\min} + (\mathbf{w} - \mathbf{w}^*)^T \mathbf{R}_{ZZ}(\mathbf{w} - \mathbf{w}^*) \qquad (6.2.26)$$

6.2.4 Adaptive Algorithm

Based on the steepest-descent method, in order to make the cost function reach its minimum value, the updating algorithm of the weight parameter matrix is expressed by following formula.

$$\mathbf{w}(k+1) = \mathbf{w}(k) - \frac{1}{2}\mu\nabla\xi \qquad (6.2.27)$$

where ξ is transient estimation of cost function J in (6.2.6), which is written as following.

$$\xi = \mathbf{e}^T(k)\mathbf{e}(k) \qquad (6.2.28)$$

and $\nabla\xi$ denotes the gradient of ξ, which is given as following.

$$\nabla\xi = \frac{\partial\xi}{\partial\mathbf{w}} = -2\frac{\partial\mathbf{y}^T(k)}{\partial\mathbf{w}}\mathbf{e}(k) \qquad (6.2.29)$$

Because the $\mathbf{y}(k)$ is already given by (6.2.15), $\frac{\partial\mathbf{y}(k)}{\partial\mathbf{w}}$ is given as following,

$$\frac{\partial\mathbf{y}(k)}{\partial\mathbf{w}} = \mathbf{Z}(k) \qquad (6.2.30)$$

So that, $\nabla\xi$ can be given as following.

$$\nabla\xi = -2\mathbf{Z}^T(k)\mathbf{e}(k) \qquad (6.2.31)$$

By substitute this result to (6.2.26), the updating expression for weight parameters form can be obtained as following.

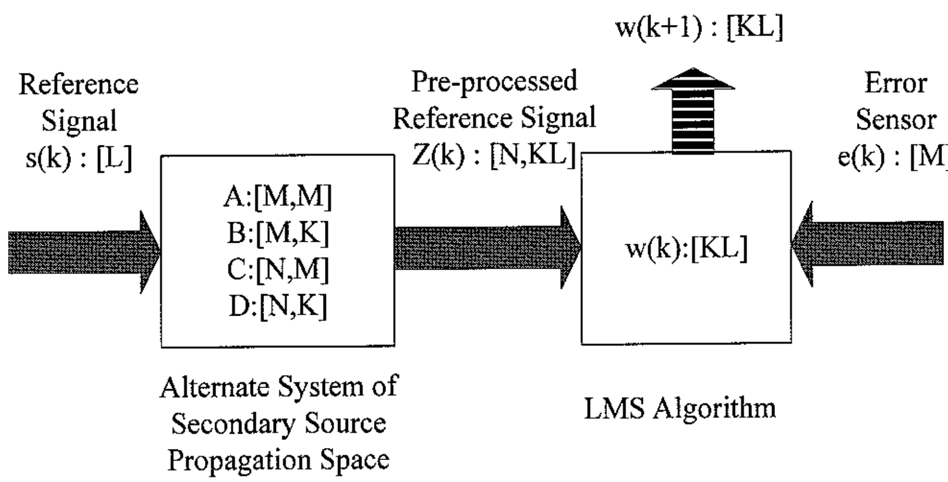
$$\mathbf{w}(k+1) = \mathbf{w}(k) + \mu \mathbf{Z}^T(k)\mathbf{e}(k) \tag{6.2.32}$$

Gathering (6.2.14) and (6.2.32) together, the adaptive algorithm can be written as following.

$$\mathbf{V}(k+1) = \mathbf{AV}(k) + \mathbf{BS}(k)$$
$$\mathbf{Z}(k) = \mathbf{CV}(k) + \mathbf{DS}(k)$$
$$\mathbf{V}(0) = \mathbf{0}_{M \times KL}$$

$$\mathbf{w}(k+1) = \mathbf{w}(k) + \mu \mathbf{Z}^T(k)\mathbf{e}(k)$$

This results extend filtered-X LMS algorithm to multiple input multiple output state space plant model, and its block diagram is shown as following.



• Figure 44 Block diagram of adaptive algorithm

Then we will have a look at convergence condition of proposed algorithm. From (6.2.26), we can obtain a different expression of $\frac{\partial J}{\partial \mathbf{w}}$ as following.

$$\nabla \xi = 2R_{ZZ}\left(\mathbf{w}(k) - \mathbf{w}^*\right) \tag{6.2.33}$$

From (6.2.24) and (6.2.27) we can get that

$$\begin{aligned} &\mathbf{w}(k+1)-\mathbf{w}^{*}=\mathbf{w}(k)-\mathbf{w}^{*}+\mu\nabla\xi \\ &=\left(\mathbf{w}(k)-\mathbf{w}^{*}\right)-\mu\mathbf{R}_{ZZ}\left(\mathbf{w}(k)-\mathbf{w}^{*}\right) \\ &=\left(1-\mu\mathbf{R}_{ZZ}\right)\left(\mathbf{w}(k)-\mathbf{w}^{*}\right) \end{aligned} \tag{6.2.34}$$

So that if we denote that $\mathbf{q}(k)=\mathbf{w}(k)-\mathbf{w}^{*}$ then (6.2.34) becomes to following equation.

$$\mathbf{q}(k+1)=\left(1-\mu\mathbf{R}_{ZZ}\right)\mathbf{q}(k) \tag{6.2.35}$$

Suppose that $\mathbf{R}_{ZZ}=\mathbf{Q}\Lambda Q^{-1}$, where $\Lambda = diag(\lambda_1,\lambda_2,\cdots,\lambda_M)$ is a eigenvalue matrix of matrix $\mathbf{R}_{ZZ}$. Also if we denote $\mathbf{v}(k)=\mathbf{Q}^{-1}\mathbf{q}(k)$, then (6.2.35) becomes to

$$\mathbf{v}(k+1)=\left(1-2\mu\Lambda\right)\mathbf{v}(k) \tag{6.2.36}$$

So it is the convergence condition is

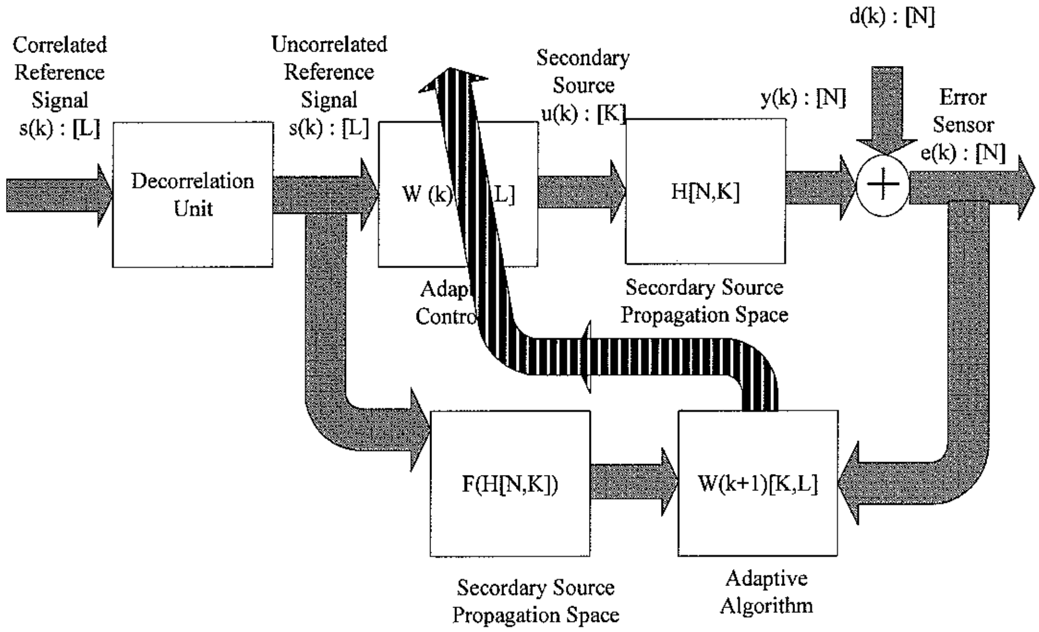
$$\max_i\left|1-2\mu\lambda_i\right|<1 \tag{6.2.37}$$

This means that convergence coefficient μ should satisfy following condition.

$$0<\mu<\frac{1}{\lambda_{\max}} \tag{6.2.38}$$

6.3 Improved Architecture of Multichannel Feedforward Active Noise Control System

It is clear that the adaptive algorithm has great influence on the performance of ANC system. An effective algorithm called as multichannel filtered reference LMS adaptive algorithm is reported by S.J. Elliott[17]. But when using this algorithm to update weight parameter matrix, the convergence speed may be limited by correlation between the reference signals. The best performance can be achieved if there is no correlation between individual input signal and each of input signals is white noise. By this reason we consider that if the correlation between input signals can be removed. An improved approach is given by Y. Tu and C. R. Fuller[18]. They proposed a adaptive decorrelation filter for preprocessing multichannel reference signal. But double stage adaptive unit may bring about unexpected slow down of convergence speed. As we have discussed in section 3.4, by using an inverse system of AR model, we can get uncorrelated white noise vector from a given random signal vector. If we can insert the inverse system in front of adaptive controller as a decorrelation unit, the we can reform the reference signal vector to uncorrelated white noise vector, this means the input of adaptive controller is uncorrelated signal, so that the multichannel filtered reference LMS adaptive algorithm will works effectively. From Figure 41, we can illustrate the block diagram of an improved ANC system shown in Figure 45.



• Figure 45 Block diagram of feedforward ANC system with decorrelatoin unit

6.4 Simulation of Noise Control inside Car

Based on above block diagram of improved architecture of multichannel feedforward active noise control system, we prepared a computer simulation as a numerical example. All of data used in this example is selected from Sutton's measured data set.

6.4.1 Simulation Conditions

Reference signals:

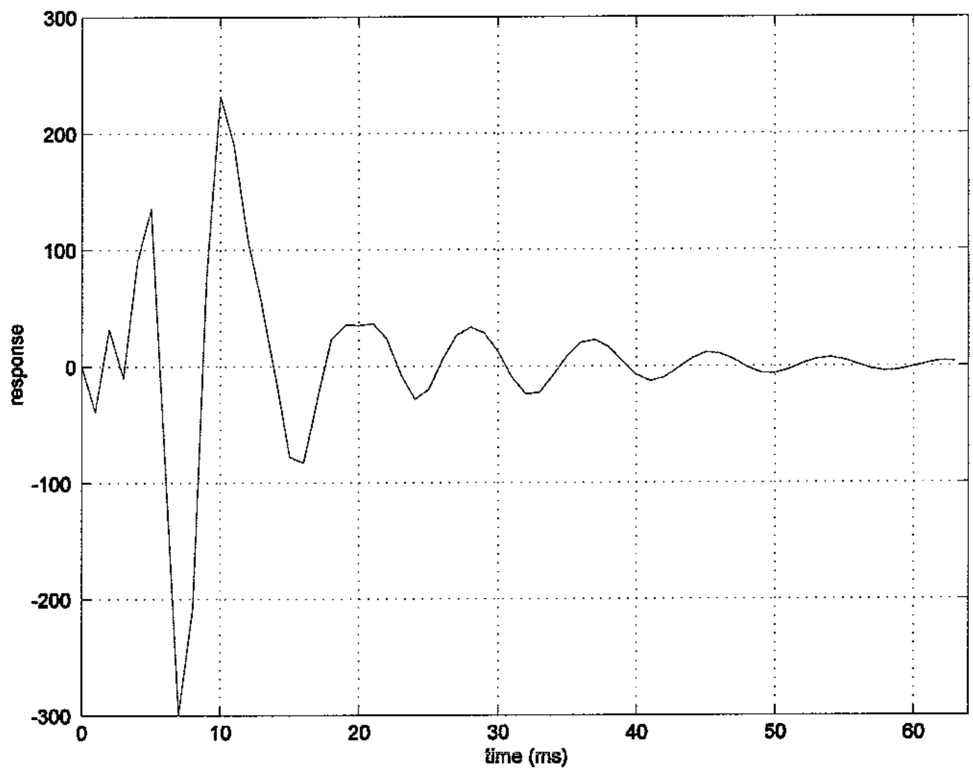
Channel 1: Accelerometer 1: RH front floor close to rear wishbone connection. (z, vertical)
Channel 2: Accelerometer 4: LH front floor close to rear wishbone connection. (z, vertical)

Road noises:

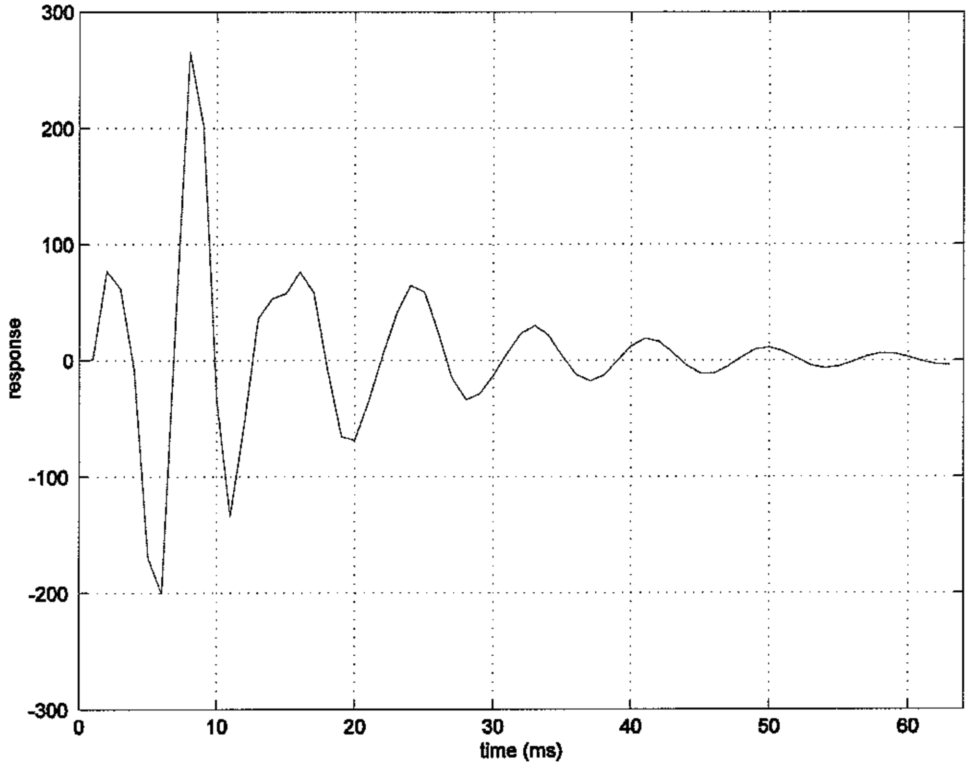
Channel 1: Microphone 1, RH microphone (out ear position)
Channel 2: Microphone 2, LH microphone (out ear position)

Plant:

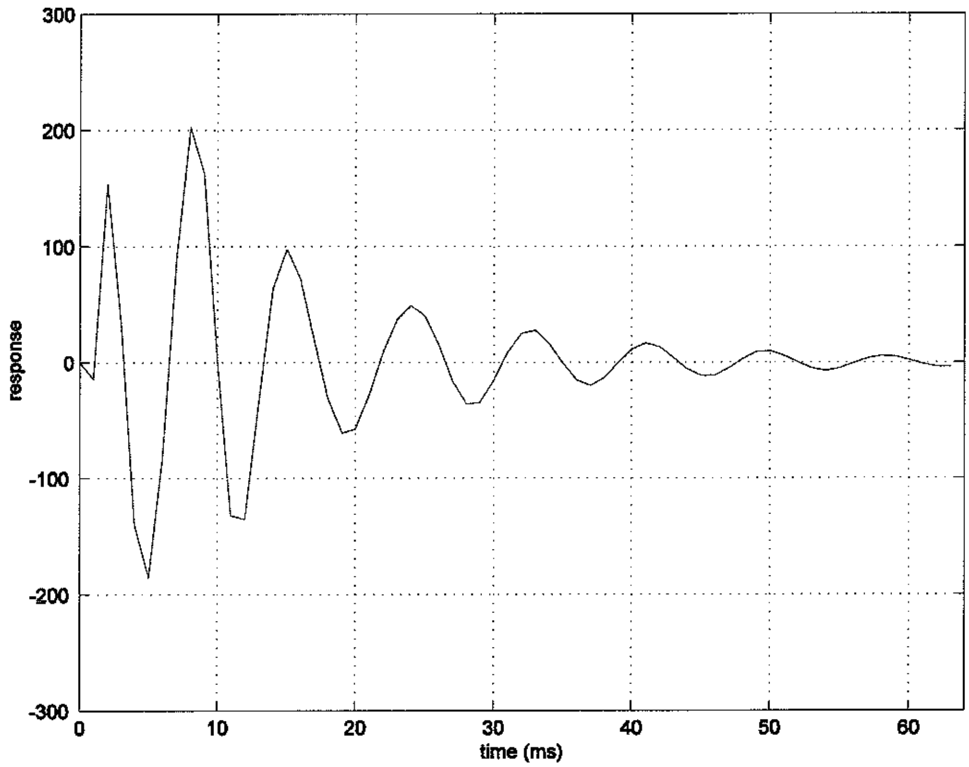
System identification results from data set for interior response by subspace method, which is transformed into a 2-input 2-output finite impulse response of 64 tap, given in Figure 46 ~ Figure 49.



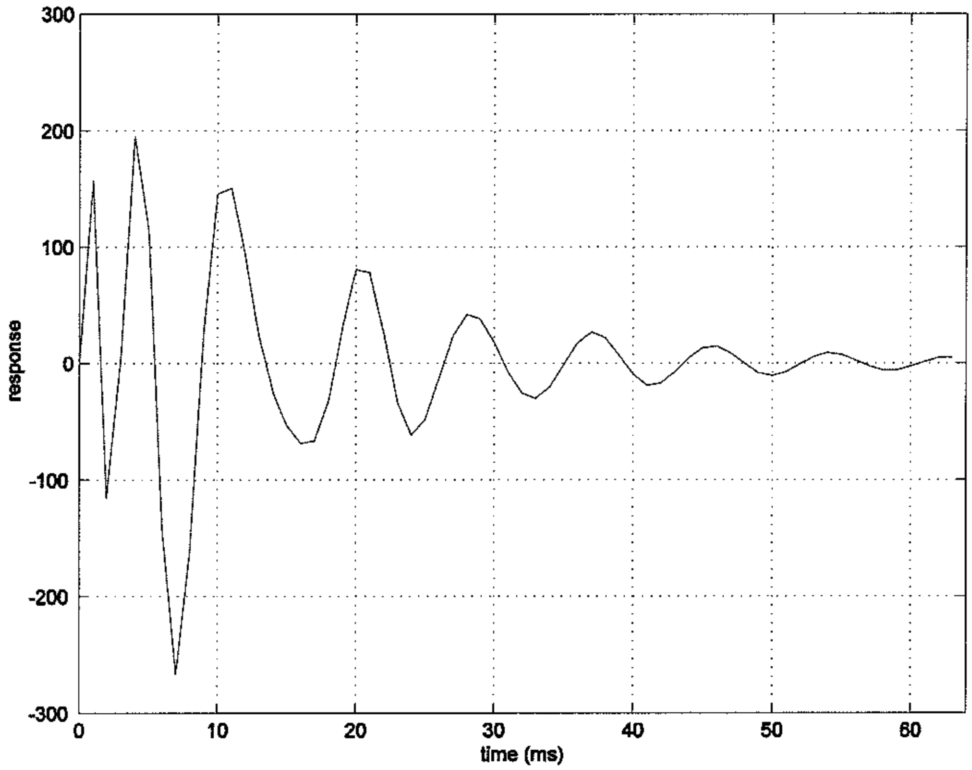
• Figure 46 impulse response from loudspeaker 1 microphone 1



• Figure 47 impulse response from loudspeaker 1 microphone 2



• Figure 48 impulse response from loudspeaker 2 microphone 2



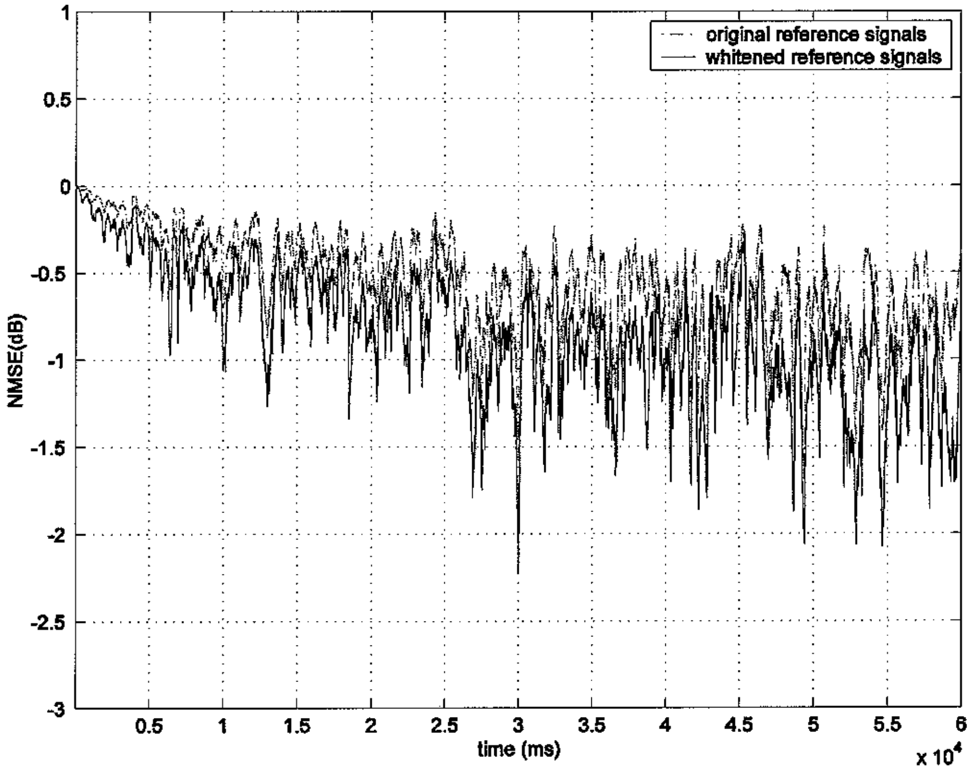
• Figure 49 impulse response from loudspeaker 2 microphone 2

6.4.2 Simulation Results

As a evaluating parameter of convergence characteristic, a normalized mean square error (NMSE) is used, which is defined as following.

$$NMSE = \frac{\sum_{i=1}^{m}(d_i - y_i)^2}{\sum_{i=1}^{m} d_i^2}$$

Figure 50 shows the learning curve of evaluating parameter of two simulation. The yellow line is for conventional method case. The black line is for improved method described in section 6.3. This result graph shows an obvious improvement of performance although that is not satisfied. The main reason is we used only 2 channels data as reference signals, whereas there are 6 channels data in Sutton's measured data set of car noise.



• Figure 50 Learning curve of 2 channel car ANC system

Chapter 7 CONCLUSIONS

In this report, methods of multichannel spectral factorisation has been reviewed, and their relevance to adaptive feedforward control has been discussed.

In chapter 2 and chapter 3, some works concerning with pre-processing multichannel reference signals have been described. In chapter 2, canonical spectral factorisation of matrix polynomial has been described. i.e. a multichannel moving average spectral model is assumed. For this problem, two solutions are given. The first one is important for basic concepts, which solves the matrix polynomial equation directly and get its minimum phase part. However, some complicated calculations are necessary for this method and it is difficult to reconstruct minimum phase part when zeros very near to unit circle. The second method tries to avoid solving the matrix polynomial equation directly via a state space expression that change the problem to a discrete algebraic Riccati equation (DARE). So by using DARE function in MATLAB, solutions have been achieved very easily. In chapter 3, multichannel spectral estimation problem is introduced by using autoregressive (AR) model, that leads a normal equation of coefficient matrices and its recursive solution of Yule-Walker algorithm is provided.

In chapter 4, some works concerning with multichannel control plant have been described. From input data and output data of multiple input multiple output system, subspace method for identification problem can provide coefficient matrices of state space expression of the system, of course then its multichannel transfer function of control plant can be obtained.

In chapter 5, based on these results of theoretical works have been applied to a measured data set of car noise. It has been confirmed here that the road noise can be whitening and uncorrelated by AR modelling with normal equation method, but it was found that the Robinson and state space MA method gave many zeros very close to the unit circle that gave a poorly conditioned solution. Also from the simultaneous measured 2 channel input and 2 channel output data in car, the noise propagation space can be identified very precisely.

In chapter 6, based on these results of whitening of multichannel reference signals and system identification of control plant, an active noise cancelling system has been constructed. Obvious improvement of performance in comparing with conventional method has been shown by a computer simulation.

The main conclusions of this report are:

- Performance of multichannel adaptive system is limited by correlation between individual channels occurs in both input (reference signals) and output (control plant) of adaptive controller.
- Spectral factorization and spectral estimation are very important, very effective method for whitening and uncorrelating multichannel signals.
- Subspace method has shown a precise result for multichannel system identification and very easy to be applied to identification of sound propagation space in car.
- An obvious performance improvement can be achieved by employing a whiten and uncorrelated referenced signals and precisely identified plant for multichannel adaptive system.

For future research, two points should be mentioned here.

- Although whitening and uncorrelating has been applied to reference signals, in order to achieve further improvement of performance, cross-talk of plant should be removed. A theoretical related matter is inner-outer factorisation and outer-inner factorisation, which is generalized concept of all-pass and minimum phase factorisation of signal channel system. So from identified system state space expression, further work to factor plant system into its inner part and outer part should be tried. By introducing a inverse system of outer part, it is considered as a virtual plant system with no cross-talk can be obtained and that will result good performance for multichannel adaptive control system.
- For systems with a large number of inputs and outputs, the centralized algorithms, as considered in this report, are not feasible anymore. If reference signals are white and uncorrelated, also if cross-talk of plant can be removed, a multichannel system can be considered as a set of signal channel system, this leads to decentralied control problem. In this case, algorithm for adaptive controller are just a parallel running algorithm for signal channel system, that could be realized very simple, whereas the multichannel filtered-X LMS algorithm need very complicated calculations. This will result many merits for practical problem including high response time and extension of frequency limitation.

REFERENCES

1) S. J. Elliott, "Optimal Controllers and Active Controllers for Multichannel Feedforward Control of Stochastic Disturbances," IEEE Trans. On Speech and Audio Processing, vol.48, no.4, pp1053-1060, April 2000.

2) D.C. Youla, "On the Factorization of Rational Matrices", IRE Trans. On Information Theory, vol.IT-7,pp.172-189,July 1961.

3) Davis, M.C., "Factoring the Spectral Matrix", IEEE Trans. On Automatic Control, AC-8, pp.296-305,1963.

4) Robinson, E.A. 'Multichannel time series analysis with digital computer programs' (Holden-Day, San Francisco, 1967)

5) Frazer, R. A., W.J. Duncan and A.R. Collar, 'Elementary Matrices' (Cambridge Univ. Press, Cambridge,1957)

6) T. Kailath, A. H. Sayed and B. Hassibi, "Linear Estimation" (Prentice Hall, 2000)

7) A. H. Sayed and T.Kailath, "A Survey of Spectral Factorization Methods", Numer. Linear Algebra Appl., vol.8,pp467-496,2001.

8) S. M. Kay, "Modern Spectral Estimation", Chapter 14.

9) J.G. Cook and S.J. Elliott, "Connection between multichannel prediction error filter and spectral factorisation", Electronics letters, vol.35, no.15, pp.1218-1220, July 1999.

10) P. Van. Overschee and B. De Moor, "N4SID: Subspace Algorithms for the Identification of Combined Deterministic Stochastic System", Automatica, vol.30, no.1,pp.75-93,1994

11) T. J. Sutton, "The Active Control of Random Noise in Automotive Interiors", Ph.D thesis, ISVR, university of Southampton, April 1992.

12) P.Lueg, "Process of silencing sound oscillation," U.S. Patent 2043416, June 9, 1936.

13) H. F. Olson and E.G. May, "Electronic sound absorber", J. Acoust. Soc. Am., 25, pp1130-1136, 1953.

14) W.B. Conover, "Fighting noise with noise", Noise Control 2, pp78-92, March 1956.

15) J.C. Burgess, "Active adaptice sound control in a duct: A computer simulation," J. Acoust. Soc Amer. Vol.70 pp.715-726, Sept. 1981.

16) G. E. Warnaka, J.Tichy and L.A. Poole, "Improvements in adaptive active attenuators," in Proc. Inter-noise, 1981, pp.307-310.

17) S. J. Elliott and P. A. Nelson, "Active noise control," IEEE signal processing magazine, vol.10, no.4, pp.12-35, Oct.1993.

18) Y. Tu and C. R. Fuller, "Multiple reference feedforward active noise control part 2: reference preprocessing and experimental results," Journal of Sound and Vibration, vol.233,no.5,pp.761-774, 2000.