

Understanding Stochastic Subspace Identification

Rune Brincker

Department of Structural and Environmental Engineering
University of Aalborg, Sohngaardsholmsvej 57
9000 Aalborg, Denmark

Palle Andersen

Structural Vibration Solutions A/S
Niels Jernes Vej 10, 9220 Aalborg East, Denmark

Nomenclature

$\mathbf{y}(t)$	System response in continuous time
$\mathbf{M}, \mathbf{D}, \mathbf{K}$	Mass, damping and stiffness matrices
$\mathbf{f}(t)$	Force vector
$\mathbf{x}(t)$	State vector
$\mathbf{A}, \mathbf{B}, \mathbf{C}$	State space matrices
\mathbf{A}_c	System matrix in continuous time
\mathbf{y}_k	System response in discrete time
\mathbf{A}_d	System matrix in discrete time
\mathbf{Y}	System response matrix
N	Number of data points
M	Number of measurement channels
\mathbf{R}_k	Covariance matrix at time lag k
\mathbf{Y}_h	Block Hankel matrix
$\mathbf{Y}_{hp}, \mathbf{Y}_{hf}$	Past and future half part of the Block Hankel matrix
\mathbf{O}	Projection matrix
$\mathbf{\Gamma}_s$	Observability matrix
\mathbf{X}_0	Kalman State matrix
$[\mu_i], \mathbf{\Psi}$	Poles and eigenvectors of discrete system matrix
$[\lambda_i], \mathbf{\Phi}$	Poles and eigenvectors of 2^{nd} order differential equation

Abstract

The data driven Stochastic Subspace Identification techniques is considered to be the most powerful class of the known identification techniques for natural input modal analysis in the time domain. However, the techniques involves several steps of "mysterious mathematics" that is difficult to follow and to understand for people with a classical background in structural dynamics. Also the connection to the classical correlation driven time domain techniques is not well established. The purpose of this paper is to explain the different steps in the SSI techniques of importance for modal identification and to show that most of the elements in the identification techniques have simple counterparts in the classical time domain techniques.

Introduction

Stochastic Subspace Identification (SSI) modal estimation algorithms have been around for more than a decade by now. The real break-through of the SSI algorithms happened in 1996 with the publishing of the book by van Overschee and De Moor [1]. A set of MATLAB files were distributed along with this book and the readers could easily convince themselves that the SSI algorithms really were a strong and efficient tool for natural input modal analysis. Because of the immediate acceptance of the effectiveness of the algorithms the mathematical framework described in the book where accepted as a de facto standard for SSI algorithms.

However, the mathematical framework is not going well together with normal engineering understanding. The reason is that the framework is covering both deterministic as well as stochastic estimation algorithms. To establish this kind of general framework more general mathematical concepts has to be introduced. Many mechanical engineers have not been trained to address problems with unknown loads enabling them to get used to concepts of stochastic theory, while many civil engineers have been trained to do so to be able to deal with natural loads like wind, waves and traffic, but on the other hand, civil engineers are not used to deterministic thinking. The book of van Overschee and De Moor [1] embraces both engineering worlds and as a result the general formulation presents a mathematics that is difficult to digest for both engineering traditions.

It is the view point of the present authors, that going back to a more traditional basis of understanding for addressing the response of structural systems due to natural input (ambient loading) makes things more easy to understand. In this paper, we will look at the SSI technique from a civil engineering (stochastic) point of view. We will present the most fundamental steps of the SSI algorithms based on the use of stochastic theory for Gaussian distributed stochastic processes, where everything is completely described by the correlation functions in time domain or by the spectral densities in frequency domain.

Most modal people still like to think about vibrations in continuous time, and thus the discrete time formulations used in SSI are not generally accepted. Therefore a short introduction is given to discrete time models and it is shown how simple it is to introduce the description of free responses in discrete time. In the SSI technique it seem mysterious to many people why the response data is gathered together in a Block Hankel matrix, orders of magnitude larger than the original amount of data. Therefore the structure of the Block Hankel matrix is related to traditional covariance estimation, and it is shown how the subsequent so-called Projection of this Hankel matrix onto itself can be explained in terms of covariances and thus results in a set of free responses for the system. Then finally it is explained how the physics can be estimated by performing a singular value decomposition of the projection matrix.

It is avoided to get into discussions about how to estimate the statistical part of the general model. Normally when introduced to the SSI technique, people will start looking at the innovation state space formulation involving mysterious Kalman states and a Kalman gain matrix that has nothing to do with the physics. This makes most engineers with a normal background in dynamics fall of the train. In this formulation, the general model is bypassed, however the mysterious Kalman states are introduced and explained as the states for the free responses estimated by the projection. Thus, this is an invitation to the people that were disappointed in the first place to get back on the track, take a another ride with the SSI train to discover that most of what you will see you can recognize as generalized procedures well established in classical modal analysis.

The discrete time formulation

We consider the stochastic response from a system as a function of time

$$(1) \quad \mathbf{y}(t) = \begin{Bmatrix} y_1(t) \\ y_2(t) \\ \vdots \\ y_M(t) \end{Bmatrix}$$

The system can be considered in classical formulation as a multi degree of freedom structural system

$$(2) \quad \mathbf{M}\ddot{\mathbf{y}}(t) + \mathbf{D}\dot{\mathbf{y}}(t) + \mathbf{K}\mathbf{y}(t) = \mathbf{f}(t)$$

Where $\mathbf{M}, \mathbf{D}, \mathbf{K}$ is the mass, damping and stiffness matrix, and where $\mathbf{f}(t)$ is the loading vector. In order to take this classical continuous time formulation to the discrete time domain the easiest way is to introduce the State Space formulation

$$(3) \quad \mathbf{x}(t) = \begin{Bmatrix} \mathbf{y}(t) \\ \dot{\mathbf{y}}(t) \end{Bmatrix}$$

Here we are using the rather confusing terminology from systems engineering where the states are denoted $\mathbf{x}(t)$ (so please don't confuse this with the system input, the system input is still $\mathbf{f}(t)$). Introducing the State Space formulation, the original 2nd order system equation given by eq. (2) simplifies to a first order equation

$$(4) \quad \begin{aligned} \dot{\mathbf{x}}(t) &= \mathbf{A}_c \mathbf{x}(t) + \mathbf{B} \mathbf{f}(t) \\ \mathbf{y}(t) &= \mathbf{C} \mathbf{x}(t) \end{aligned}$$

Where the system matrix \mathbf{A}_c in continuous time and the load matrix \mathbf{B} is given by

$$(5) \quad \begin{aligned} \mathbf{A}_c &= \begin{bmatrix} 0 & \mathbf{I} \\ -\mathbf{M}^{-1}\mathbf{K} & -\mathbf{M}^{-1}\mathbf{D} \end{bmatrix} \\ \mathbf{B} &= \begin{bmatrix} 0 \\ \mathbf{M}^{-1} \end{bmatrix} \end{aligned}$$

The advantage of this formulation is that the general solution is directly available, se for instance Kailath [2]

$$(6) \quad \mathbf{x}(t) = \exp(\mathbf{A}_c t) \mathbf{x}(0) + \int_0^t \exp(\mathbf{A}_c (t - \tau)) \mathbf{B} \mathbf{f}(\tau) d\tau$$

Where the first term is the solution to the homogenous equation and the last term is the particular solution. To take this solution to discrete time, we sample all variables like $y_k = y(k\Delta t)$ and thus the solution to the homogenous equation becomes

$$(7) \quad \begin{aligned} \mathbf{x}_k &= \exp(\mathbf{A}_c k\Delta t) \mathbf{x}_0 = \mathbf{A}_d^k \mathbf{x}_0 \\ \mathbf{A}_d &= \exp(\mathbf{A}_c \Delta t) \\ \mathbf{y}_k &= \mathbf{C} \mathbf{A}_d^k \mathbf{x}_0 \end{aligned}$$

Here one should not be confused by the fact that we calculate the exponential function of a matrix, this construction is simply defined by its power series, and in practice is calculated by performing a eigen-value decomposition of the involved matrix and then taking the exponential function of the eigen values. Note that the system matrix in continuous time and in discrete is time is not the same.

The Block Hankel Matrix

In discrete time, the system response is normally represented by the data matrix

$$(8) \quad \mathbf{Y} = [\mathbf{y}_1 \ \mathbf{y}_2 \ \cdots \ \mathbf{y}_N]$$

Where N is the number of data points. To understand the meaning of the Block Hankel matrix, it is useful to consider a more simple case where we perform the product between two matrices that are modifications of the

data matrix given by eq. (7). Let $\mathbf{Y}_{(1:N-k)}$ be the data matrix where we have removed the last k data points, and similarly, let $\mathbf{Y}_{(k:N)}$ be the data matrix where we have removed the first k data points, then

$$(9) \quad \hat{\mathbf{R}}_k = \frac{1}{N-k} \mathbf{Y}_{(1:N-k)} \mathbf{Y}_{(k:N)}^T$$

Is an unbiased estimate of the correlation matrix at time lag k . This follows directly from the definition of the correlation estimate, se for instance Bendat and Piersol [4]. The Block Hankel matrix \mathbf{Y}_h defined in SSI is simply a gathering of a family of matrices that are created by shifting the data matrix

$$(10) \quad \mathbf{Y}_h = \begin{bmatrix} \mathbf{Y}_{(1:N-2s)} \\ \mathbf{Y}_{(2:N-2s+1)} \\ \vdots \\ \mathbf{Y}_{(2s:N)} \end{bmatrix} = \begin{bmatrix} \mathbf{Y}_{hp} \\ \mathbf{Y}_{hf} \end{bmatrix}$$

The upper half part of this matrix is called “the past” and denoted \mathbf{Y}_{hp} and the lower half part of the matrix is called “the future” and is denoted \mathbf{Y}_{hf} . The total data shift is $2s$ and is denoted “the number of block rows” (of the upper or lower part of the Block Hankel matrix). The number of rows in the Block Hankel matrix is $2sM$, the number of columns is $N - 2s$.

The Projection

Here comes what in many peoples opinion is one of the most mysterious operations in SSI. In van Overschee and De Moor [1] the projection is introduced as a geometrical tool and is explained mainly in this context. However, dealing with stochastic responses, projection is defined as a conditional mean. Specifically, in SSI the projection of the future unto the past defines the matrix

$$(11) \quad \mathbf{O} = \mathbf{E}(\mathbf{Y}_{hf} | \mathbf{Y}_{hp})$$

A conditional mean like this can for Gaussian processes be totally described by its covariances, se for instance Melsa & Sage [3]. Since the shifted data matrices also defines covariances, it is not so strange that the projection can be calculated directly as also defined by van Overschee and De Moor [1]

$$(12) \quad \mathbf{O} = \mathbf{Y}_{hf} \mathbf{Y}_{hp}^T (\mathbf{Y}_{hp} \mathbf{Y}_{hp}^T)^{-1} \mathbf{Y}_{hp}$$

The last matrix in this product defines the conditions, the first four matrices in the product introduces the covariances between channels at different time lags. A conditional mean like given by eq. (10) simply consist of free decays of the system given by different initial conditions specified by \mathbf{Y}_{hp} . The matrix is $sM \times sM$ and any column in the matrix \mathbf{O} is a stacked free decay of the system to a (so far unknown) set of initial conditions. Using eq. (7) any column in \mathbf{O} can be expressed by

$$(13) \quad \mathbf{o}_{col} = \mathbf{\Gamma}_s \mathbf{x}_0$$

$$\mathbf{\Gamma}_s = \begin{bmatrix} \mathbf{C} \\ \mathbf{C}\mathbf{A}_d \\ \mathbf{C}\mathbf{A}_d^2 \\ \vdots \\ \mathbf{C}\mathbf{A}_d^{s-1} \end{bmatrix}$$

Now, if we knew the so-called observability matrix Γ_s , then we could simply find the initial conditions directly from eq. (13) (it is a useful exercise to simulate a system response from the known system matrices, use the SSI standard procedure to find the matrix \mathbf{O} and then try to estimate the initial conditions directly from eq. (13)).

The Kalman States

The so-called Kalman states are simply the initial conditions for all the columns in the matrix \mathbf{O} , thus

$$(14) \quad \mathbf{O} = \Gamma_s \mathbf{X}_0$$

Where the matrix \mathbf{X}_0 contains the so defined Kalman states at time lag zero. Again, if we knew the matrix Γ_s , then we could simply find all the Kalman states directly from eq.(14), however, since we don't know the matrix Γ_s , we cannot do so, and thus we have to estimate the states in a different way. The trick is to use the SVD on the \mathbf{O} matrix

$$(15) \quad \mathbf{O} = \mathbf{U}\mathbf{S}\mathbf{V}^T$$

And then define the estimate of the matrix Γ_s and the Kalman state matrix states \mathbf{X}_0 by

$$(16) \quad \begin{aligned} \hat{\Gamma} &= \mathbf{U}\mathbf{S}^{1/2} \\ \hat{\mathbf{X}}_0 &= \mathbf{S}^{1/2}\mathbf{V}^T \end{aligned}$$

The so defined procedure for estimating the matrices $\hat{\Gamma}$ and $\hat{\mathbf{X}}_0$ is not unique. A certain arbitrary similarity transformation can be shown to influence the individual matrices, but can also be shown not to influence the estimation of the system matrices.

A note on the name "Kalman states". The Kalman state matrix $\hat{\mathbf{X}}_0$ is the Kalman state matrix for time lag zero. If we remove one block row of \mathbf{O} from the top, and then one block row of Γ_s from the bottom, then similarly we can estimate the Kalman state matrix $\hat{\mathbf{X}}_1$ at time lag one. Thus by subsequently removing block rows from \mathbf{O} all the Kalman states can be defined. Using the Kalman states a more general formulation for estimating also the noise part of the stochastic response modeling can be established. However, in this paper we focus on explaining how the system matrices can be found, and in this context, there is no further need for Kalman states.

Estimating the system matrices

The system matrix \mathbf{A}_d can be found from the estimate of the matrix Γ by removing one block from the top and one block from the bottom yielding

$$(17) \quad \hat{\Gamma}_{(2:s)} \hat{\mathbf{A}}_d = \hat{\Gamma}_{(1:s-1)}$$

And thus, the system matrix $\hat{\mathbf{A}}_d$ can be found by regression. The observation matrix \mathbf{C} can be found simply by taking the first block of the observability matrix

$$(18) \quad \hat{\mathbf{C}} = \hat{\Gamma}_{(1)}$$

Modal Analysis and practical issues

Now we are finally back to something like what we normally do in the field of structural vibrations. First step of finding the modal parameters is to perform an eigenvalue decomposition of the system matrix $\hat{\mathbf{A}}_d$

$$(19) \quad \hat{\mathbf{A}}_d = \mathbf{\Psi}[\mu_i]\mathbf{\Psi}^{-1}$$

The continuous time poles λ_i are found from the discrete time poles μ_i by

$$(20) \quad \mu_i = \exp(\lambda_i)$$

Leading to the well known formulas

$$(21) \quad \begin{aligned} \lambda_i &= \frac{\ln(\mu_i)}{\Delta T} \\ \omega_i &= |\lambda_i| \\ f_i &= \frac{\omega_i}{2\pi} \\ \zeta_i &= \frac{\text{Re}(\lambda_i)}{|\lambda_i|} \end{aligned}$$

The mode shape matrix is found from

$$(22) \quad \mathbf{\Phi} = \mathbf{C}\mathbf{\Psi}$$

and the job is done – from a modal point of view – if we are able to make up our mind about the size of the Block Hankel matrix. As we have seen earlier, the number s defines the size of the Block Hankel matrix, and thus also the size of the projection matrix \mathbf{O} . However, the number sM defines the number of eigenvalues in our model, thus sM defines the model order. Normally we would like to vary the model order to establish a stabilization diagram. This can of course be done by establishing a series of Block Hankel matrices of different size, but it is somewhat easier, instead of varying the size of the Block Hankel matrix, to vary the number of singular values used in eq. (16). Thus in practice the size of the Block Hankel matrix defines the maximum model order, and the actual model order is varied by varying the number of singular values taken into account when performing the singular value decomposition of the projection matrix. The maximum number of eigen values sM must be adjusted to a reasonable level to incorporate the needed range of models.

References

- [1] Peter van Overschee and Bart De Moor: Subspace Identification for Linear Systems. Kluwer Academic Publishers, 1996
- [2] T. Kailath: Linear Systems. Prentice Hall Inc., 1980.
- [3] James L. Melsa and Andrew P. Sage: An Introduction to Probability and Stochastic Processes. Prentice-Hall Inc., 1973.
- [4] Julius S. Bendat and Allan G. Piersol: Random Data - Analysis and Measurement Procedures. John Wiley & Sons, 1986.