2 Introduction of Discrete-Time Systems

This chapter concerns an important subclass of discrete-time systems, which are the linear and time-invariant systems excited by Gaussian distributed stochastic processes. In the following, it will be shown how these systems can be represented in state space or equivalently by ARMAV models. The modelling of linear time-invariant stochastic systems using univariate ARMA models has been covered extensively in many textbooks such as Box et al. [16], Harvey [36], Harvey [37] and Pandit et al. [88]. It is, however, only recently that modelling using multivariate ARMAV models has been covered systematically, see e.g. Aoki [11], Gevers et al. [29], Hannan et al. [33], Pandit [84], Pi et al. [92] and Piombo et al. [93]. Multivariate state space systems have been applied in e.g. control engineering for several decades, and the analytical manipulations and numerical processing of these systems have been covered extensively in many textbooks, see e.g. Goodwin et al. [31], Hannan et al. [33], Kailath [48], Ljung [71] and Middleton et al. [79].

Section 2.1 will introduce the state space representation of a linear and time-invariant discrete-time system excited by a Gaussian stochastic excitation. Besides the applied Gaussian stochastic excitation, a system may also be affected by noise or disturbance. In section 2.2, it is therefore investigated how to account for disturbance by means of a so-called innovation state space representation. Section 2.3, will then show how the state space system presented in section 2.1 can be converted to an equivalent ARMAV model. Similarly, in section 2.4 it is shown how the innovation state space representation presented in section 2.2 can be equivalently represented by an ARMAV model. The conversions from state space to the particular ARMAV models are unique. However, there are many ways to realise ARMAV models in state space, i.e. ways to select an internal structure of the state space representation. In section 2.5 one of these realizations will be presented.

2.1 Modelling of Discrete-Time Systems

In this section a state space representation of a stochastically excited discrete-time system is introduced. For simplicity, it will be assumed that the discrete-time system has as many inputs as there are outputs, i.e. that the dimension of the input process $\boldsymbol{u}(t_k)$ is the same as the dimension of the output process $\boldsymbol{y}(t_k)$.

2.1.1 The Stochastic State Space System

A linear and time-invariant discrete-time system can be represented in state space as

$$\begin{aligned} \boldsymbol{x}(t_{k+1}) &= \boldsymbol{A}\boldsymbol{x}(t_k) + \boldsymbol{B}\boldsymbol{u}(t_k) \\ \boldsymbol{y}(t_k) &= \boldsymbol{C}\boldsymbol{x}(t_k) + \boldsymbol{D}\boldsymbol{u}(t_k) \end{aligned} \tag{2.1}$$

The first equation is called the state equation and models the dynamic behaviour of the discrete-time system. The second equation is called the observation equation, since this equation controls which part of the output of the representation is observed. The input process $\boldsymbol{u}(t_k)$ is assumed to be a zero mean stationary Gaussian white noise, i.e.

$$
\begin{aligned}
E\left[\boldsymbol{u}(t_k)\right] &= \mathbf{0} \\
E\left[\boldsymbol{u}(t_k)\boldsymbol{u}^T(t_{k+n})\right] &= \boldsymbol{\Delta}\,\delta(n)
\end{aligned}
\tag{2.2}
$$

where $\delta(n)$ is the Kronecker Delta. The dimension of $\boldsymbol{u}(t_k)$ is assumed to be $p \times 1$, which implies that the dimension of the covariance matrix $\boldsymbol{\Delta}$ is $p \times p$. In the following the statistical properties in (2.2) will be abbreviated $NID(\mathbf{0},\boldsymbol{\Delta})$. Since (2.1) only involves linear operations the state vector process $\boldsymbol{x}(t_k)$ of the representation and the output process $\boldsymbol{y}(t_k)$ will also be zero-mean Gaussian distributed processes if the initial state vector is zero-mean, i.e.

$$
\begin{aligned}
E\left[\boldsymbol{y}(t_k)\right] &= \mathbf{0} \quad , \quad E\left[\boldsymbol{x}(t_k)\right] = \mathbf{0} \\
E\left[\boldsymbol{y}(t_k)\boldsymbol{y}^T(t_{k+n})\right] &= \boldsymbol{\Sigma}(n) \quad , \quad E\left[\boldsymbol{x}(t_k)\boldsymbol{x}^T(t_{k+n})\right] = \boldsymbol{\Pi}(n)
\end{aligned}
\tag{2.3}
$$

The statistical properties of these two processes are therefore fully described by the covariance functions $\boldsymbol{\Sigma}(n)$ and $\boldsymbol{\Pi}(n)$. Following the assumption made in the beginning of this section the dimension of $\boldsymbol{y}(t_k)$ is $p \times 1$. The dimension of the state vector $\boldsymbol{x}(t_k)$ is assumed to be $m \times 1$, with $m \geq p$. Since the state vector has the dimension m, the state space representation is said to be m-dimensional.

Based on the dimensions of the input, state, and output vectors the dimensions of the system matrices are:

- ☞ Transition matrix $\boldsymbol{A}$: $dim = m \times m$
- ☞ Input matrix $\boldsymbol{B}$: $dim = m \times p$
- ☞ Observation matrix $\boldsymbol{C}$: $dim = p \times m$
- ☞ Direct term matrix $\boldsymbol{D}$: $dim = p \times p$

The state space representation (2.1) of a discrete-time system is called a stochastic state space system, due to the stochastic input that results in a stochastic response. If a particular parameterization of the system matrices has been specified, (2.1) is referred to as a stochastic state space realization. In most cases the direct term, $\boldsymbol{D}\boldsymbol{u}(t_k)$, in the observation equation will be zero. However, there might be cases where this term is needed. As an example: If the accelerations of a sampled Gaussian white noise excited second-order system are extracted, it will be necessary to add a direct term, see Hoen [38]. Even though such a situation will not occur in this thesis, the presence of a direct term will be allowed for completeness in the following analysis.

A solution of (2.1) for $\boldsymbol{y}(t_{k+l})$, $l \geq 0$, can be obtained by recursive substitutions of itself to yield

$$
\begin{aligned}
\boldsymbol{y}(t_{k+l}) &= \boldsymbol{C}\boldsymbol{A}^{l}\boldsymbol{x}(t_k) + \sum_{j=1}^{l} \boldsymbol{C}\boldsymbol{A}^{l-j}\boldsymbol{B}\boldsymbol{u}(t_{k+j-1}) + \boldsymbol{D}\boldsymbol{u}(t_{k+l}) \\
\boldsymbol{u}(t_k) &\in NID(\boldsymbol{0},\boldsymbol{\Delta}) , \qquad \boldsymbol{x}(t_k) \equiv \boldsymbol{x}_0
\end{aligned}
\tag{2.4}
$$

The solution is split into a homogeneous part, which is represented by the first right-hand term, and a particular part represented by the second right-hand term.

2.1.2 Modelling of a Nonwhite Excitation

It is also possible to use (2.1) as a representation of a linear and time-invariant discrete-time system that is excited by nonwhite Gaussian distributed excitation. The only requirement is that the nonwhite Gaussian distributed stochastic input can be assumed to be generated from a linear and time-invariant shaping filter, see e.g. Melsa et al. [77]. Let the following system that is excited by a nonwhite Gaussian distributed input be represented by the following state space system

$$
\begin{aligned}
\boldsymbol{x}_1(t_{k+1}) &= \boldsymbol{A}_1\boldsymbol{x}_1(t_k) + \boldsymbol{B}_1\boldsymbol{w}(t_k) \\
\boldsymbol{y}(t_k) &= \boldsymbol{C}_1\boldsymbol{x}_1(t_k) + \boldsymbol{D}_1\boldsymbol{w}(t_k)
\end{aligned}
\tag{2.5}
$$

and assume that the nonwhite input $\boldsymbol{w}(t_k)$ can be obtained by filtering Gaussian white noise $\boldsymbol{u}(t_k)$ through a linear and time-invariant shaping filter of the form

$$
\begin{aligned}
\boldsymbol{x}_2(t_{k+1}) &= \boldsymbol{A}_2\boldsymbol{x}_2(t_k) + \boldsymbol{B}_2\boldsymbol{u}(t_k) \\
\boldsymbol{w}(t_k) &= \boldsymbol{C}_2\boldsymbol{x}_2(t_k) + \boldsymbol{D}_2\boldsymbol{u}(t_k)
\end{aligned}
\tag{2.6}
$$

These state space systems (2.5) and (2.6) are coupled, and they can be represented by a single state space system by defining the following augmented state vector

$$
\boldsymbol{x}(t_k) = \begin{bmatrix} \boldsymbol{x}_1(t_k) \\ \boldsymbol{x}_2(t_k) \end{bmatrix}
\tag{2.7}
$$

Insert the observation equation of (2.6) into the state equation of (2.5) and stack the obtained relation on top of the state equation of (2.6). Further, insert the observation equation of (2.6) into the observation equation of (2.5). In other words

$$
\begin{aligned}
\boldsymbol{x}_1(t_{k+1}) &= \boldsymbol{A}_1\boldsymbol{x}_1(t_k) + \boldsymbol{B}_1\big(\boldsymbol{C}_2\boldsymbol{x}_2(t_k) + \boldsymbol{D}_2\boldsymbol{u}(t_k)\big) \\
\boldsymbol{x}_2(t_{k+1}) &= \boldsymbol{A}_2\boldsymbol{x}_2(t_k) + \boldsymbol{B}_2\boldsymbol{u}(t_k) \\
\boldsymbol{y}(t_k) &= \boldsymbol{C}_1\boldsymbol{x}_1(t_k) + \boldsymbol{D}_1\big(\boldsymbol{C}_2\boldsymbol{x}_2(t_k) + \boldsymbol{D}_2\boldsymbol{u}(t_k))\big)
\end{aligned}
\tag{2.8}
$$

It is seen that the nonwhite input $\boldsymbol{w}(t_k)$ has been eliminated. Introducing the augmented state vector in (2.8), it is also seen that a state space system in the standard format is obtained, i.e. a system excited by Gaussian white noise. The system matrices of this augmented representation are defined as

$$\boldsymbol{A} = \begin{bmatrix} \boldsymbol{A}_1 & \boldsymbol{B}_1\boldsymbol{C}_2 \\ \mathbf{0} & \boldsymbol{A}_2 \end{bmatrix}, \quad \boldsymbol{B} = \begin{bmatrix} \boldsymbol{B}_1\boldsymbol{D}_2 \\ \boldsymbol{B}_2 \end{bmatrix}, \quad \boldsymbol{C} = \begin{bmatrix} \boldsymbol{C}_1 & \boldsymbol{D}_1\boldsymbol{C}_2 \end{bmatrix}, \quad \boldsymbol{D} = \boldsymbol{D}_1\boldsymbol{D}_2 \qquad (2.9)$$

So in conclusion:

☞ *If a linear and time-invariant discrete-time system is subjected to a stochastic input, that can be assumed to be generated by filtering Gaussian white noise through a linear and time-invariant shaping filter, then it will always be possible to represent the system by a Gaussian white noise excited state space system.*

Since it is a typical assumption that the excitation of a civil engineering structure can be modelled as filtered Gaussian white noise, see Hoen [38], Ibrahim [41], Kozin et al. [68], Melsa et al. [77] and Tajimi [106], the state space representation (2.1) is assumed to be fully capable of describing the linear and time-invariant dynamic behaviour of such a structure.

2.1.3 Properties of Stochastic State Space Systems

A linear and discrete-time system can be realised in state space in a number of ways. Due to this non-uniqueness care must be taken with regards to the parameterization of the system matrices and the dimension of the state space. If the dimension is too small there will most certainly be an information loss, compared to the system that is being modelled. On the other hand, if the dimension of the state space is too large, the state space realization will contain redundant information. Any of these cases should of course be avoided. Therefore, a realization should have a minimal state space dimension, where all modes in the system can be observed in the output, and excited by an appropriate input. Such a realization is called minimal. A realization is only minimal if it is observable and reachable, see Kailath [48].

Definition 2.1 - Observability and Reachability of a State Space Realization

An m-dimensional state space system is observable and reachable if the observability matrix $\boldsymbol{Q}_o(m)$, and the reachability matrix $\boldsymbol{Q}_r(m)$, defined as

$$\boldsymbol{Q}_o(m) = \begin{bmatrix} \boldsymbol{C} \\ \boldsymbol{CA} \\ \cdot \\ \cdot \\ \boldsymbol{CA}^{m-1} \end{bmatrix}, \qquad \boldsymbol{Q}_r(m) = \begin{bmatrix} \boldsymbol{B} & \boldsymbol{AB} & . & . & \boldsymbol{A}^{m-1}\boldsymbol{B} \end{bmatrix} \tag{2.10}$$

both have full rank m, see Kailath [48]. If a realization is observable, it is in principle possible to observe all dynamic modes of it, and if it is reachable, it is possible to transfer any initial state $\boldsymbol{x}(t_k)$ to an arbitrary state $\boldsymbol{x}(t_{k+m})$ in not more than m steps provided that the reachability matrix $\boldsymbol{Q}_r(m)$ is non-singular, i.e. has full rank m. □

The observability and reachability matrices are very important in many parts of system theory and system identification. One of the important properties of minimal state space systems is that one realization can be uniquely transformed to another by a similarity transformation.

Definition 2.2 - The Similarity Transformation of Minimal Realizations

Consider the m-dimensional state space system (2.1), define the following linear transformation of the state vector

$$\tilde{\boldsymbol{x}}(t_k) = \boldsymbol{T}\boldsymbol{x}(t_k) \tag{2.11}$$

where $\boldsymbol{T}$ is an $m \times m$ non-singular transformation matrix. The state space system (2.1) can then be expressed in terms of the new state vector $\tilde{\boldsymbol{x}}(t_k)$ as

$$\begin{aligned} \tilde{\boldsymbol{x}}(t_{k+1}) &= \tilde{\boldsymbol{A}}\tilde{\boldsymbol{x}}(t_k) + \tilde{\boldsymbol{B}}\boldsymbol{u}(t_k) \\ \boldsymbol{y}(t_k) &= \tilde{\boldsymbol{C}}\tilde{\boldsymbol{x}}(t_k) + \tilde{\boldsymbol{D}}\boldsymbol{u}(t_k) \end{aligned} \tag{2.12}$$

with the new system matrices defined as

$$\tilde{\boldsymbol{A}} = \boldsymbol{TAT}^{-1}, \quad \tilde{\boldsymbol{B}} = \boldsymbol{TB}, \quad \tilde{\boldsymbol{C}} = \boldsymbol{CT}^{-1}, \quad \tilde{\boldsymbol{D}} = \boldsymbol{D} \tag{2.13}$$

This linear transformation of one realization to another is called a similarity transformation. □

Similarity transformations are frequently used in system theory. One of the well-known similarity transformations is the eigenvalue / -vector decomposition, i.e. the modal decomposition, see e.g. Kailath [48].

2.2 Modelling of Discrete-Time Systems Affected by Noise

Besides the applied input to the state space system (2.1), there might be other inputs that in a more uncontrollable way contribute to the system response. This undesirable influence is characterized as disturbance or noise. In system identification using measured system response disturbance might be caused by different phenomena. The most obvious phenomena are noise generated by the sensors, and noise arising from roundoff errors during the A/D conversion. Disturbance might also be caused by an inadequacy of using a linear and time-invariant model. Such an inadequacy could be caused by a small non-linearity in the true system or by non-stationary excitation. In any case, noise will always be present in measured data and should therefore always be taken into account. It is therefore necessary to extend the stochastic state space system (2.1) with a noise model.

2.2.1 Modelling of Disturbance

In the context of representing linear and time-invariant discrete-time systems in state space, the different kinds of noise are usually divided into two categories:

- ☞ *Process noise.*
- ☞ *Measurement noise.*

The process noise should be regarded as one equivalent noise term that models all the noise sources causing the linear and time-invariant state space modelling to be inadequate. This could be the above-mentioned small non-linearities in the true system. The measurement noise should in a similar manner be regarded as an equivalent noise term that includes all the noise sources that disturb the measurements. This could be the above mentioned A/D roundoff errors and the sensor noise.

It will be assumed, that the process noise can be described by a zero-mean Gaussian white noise process $\boldsymbol{w}(t_k)$, and that the measurement noise in a similar way can be described by a zero-mean Gaussian white noise process $\boldsymbol{v}(t_k)$. The joint distribution of these two processes is given by

$$E\begin{bmatrix}\boldsymbol{w}(t_k)\\ \boldsymbol{v}(t_k)\end{bmatrix} = \boldsymbol{0} \ , \qquad E\left[\begin{bmatrix}\boldsymbol{w}(t_k)\\ \boldsymbol{v}(t_k)\end{bmatrix}\begin{bmatrix}\boldsymbol{w}^T(t_{k+n}) & \boldsymbol{v}^T(t_{k+n})\end{bmatrix}\right] = \begin{bmatrix}\boldsymbol{Q} & \boldsymbol{S}\\ \boldsymbol{S}^T & \boldsymbol{R}\end{bmatrix}\delta(n) \tag{2.14}$$

The process noise that disturbs the system should be added to the state equation in (2.1), whereas the measurement noise should be added to the observation equation. In other words

$$\begin{aligned}\boldsymbol{x}(t_{k+1}) &= \boldsymbol{A}\boldsymbol{x}(t_k) + \boldsymbol{B}\boldsymbol{u}(t_k) + \boldsymbol{w}(t_k)\\ \boldsymbol{y}(t_k) &= \boldsymbol{C}\boldsymbol{x}(t_k) + \boldsymbol{D}\boldsymbol{u}(t_k) + \boldsymbol{v}(t_k)\end{aligned} \tag{2.15}$$

This implies that the dimensions of $\boldsymbol{w}(t_k)$ and $\boldsymbol{v}(t_k)$ are $m \times 1$ and $p \times 1$, respectively. It also implies that the covariance matrices describing the disturbance have the following dimensions:

- ☞ $\boldsymbol{Q}$: $dim = m \times m$
- ☞ $\boldsymbol{R}$: $dim = p \times p$
- ☞ $\boldsymbol{S}$: $dim = p \times m$

The problem of noise contaminated systems is that it is only possible to predict the response. For state space systems, this prediction is accomplished by the construction of the associated Kalman filter.

2.2.2 The Steady-State Kalman Filter

Assume that measurements $\boldsymbol{y}(t_k)$ are available and that they are zero mean and Gaussian distributed. Also assume that the matrices $\{\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{C}, \boldsymbol{D}, \boldsymbol{\Delta}, \boldsymbol{Q}, \boldsymbol{R}, \boldsymbol{S}\}$ are known. It is realized that in system identification of civil engineering structures all these matrices are unknown, and some of them are impossible to determine. However, in order to carry out the statistical analysis, it is formally assumed that they are known. Now, since it is only possible to predict the system response optimal predictors of the state of the system and of the system response are needed.

Definition 2.3 - Optimal Predictor of the State Space System

The optimal in least-square sense one-step-ahead predictor $\hat{\boldsymbol{x}}(t_k|t_{k-1})$ of the state vector is defined as the conditional mean of $\boldsymbol{x}(t_k)$, given all previous measurements, collected in a vector Y^{k-1} defined as the following sequence $Y^{k-1} = \{\, \boldsymbol{y}(t_0), \boldsymbol{y}(t_1), \ldots, \boldsymbol{y}(t_{k-1})\,\}^T$. In other words

$$
\begin{aligned}
\hat{\boldsymbol{x}}(t_k|t_{k-1}) &= E\Big[\boldsymbol{x}(t_k)|\{\boldsymbol{y}(t_0),\ \boldsymbol{y}(t_1),\ \ldots\ ,\ \boldsymbol{y}(t_{k-1})\}^T\Big] \\
&= E\Big[\boldsymbol{x}(t_k)|Y^{k-1}\Big]
\end{aligned}
\tag{2.16}
$$

The optimal one-step-ahead predictor $\hat{\boldsymbol{y}}(t_k|t_{k-1})$ of the measured response $\boldsymbol{y}(t_k)$ is defined in a similar manner as the conditional mean of $\boldsymbol{y}(t_k)$, given Y^{k-1}

$$
\begin{aligned}
\hat{\boldsymbol{y}}(t_k|t_{k-1}) &= E\Big[\boldsymbol{y}(t_k)|Y^{k-1}\Big] \\
&= E\Big[\boldsymbol{C}\boldsymbol{x}(t_k) + \boldsymbol{D}\boldsymbol{u}(t_k) + \boldsymbol{v}(t_k)|Y^{k-1}\Big] \\
&= \boldsymbol{C}\hat{\boldsymbol{x}}(t_k|t_{k-1})
\end{aligned}
\tag{2.17}
$$

As seen $\hat{\boldsymbol{y}}(t_k|t_{k-1})$ is related to $\hat{\boldsymbol{x}}(t_k|t_{k-1})$ through the observation matrix $\boldsymbol{C}$. □

Given the system matrices of the state space system in (2.15), the measurements $\boldsymbol{y}(t_k)$, and the stochastic input $\boldsymbol{u}(t_k)$, the basic idea of Kalman filtering is to predict the state $\hat{\boldsymbol{x}}(t_k|t_{k-1})$ for $\boldsymbol{x}(t_k)$ in the sense that the state prediction error is as small as possible. This state prediction error is defined in the following.

Definition 2.4 - State Prediction Errors

Let $\varepsilon(t_k)$ represent the part of $\boldsymbol{x}(t_k)$ that cannot be predicted from Y^{k-1} by $\hat{\boldsymbol{x}}(t_k|t_{k-1})$ as

$$\boldsymbol{\varepsilon}(t_k) = \hat{\boldsymbol{x}}(t_k|t_{k-1}) - \boldsymbol{x}(t_k) \tag{2.18}$$

This part is termed the state prediction error. □

Since only observable states can be predicted, and since it must be possible to predict the complete state vector, the system has to be observable, see definition 2.1. Also, because it is only possible to predict the response of the system itself, there will always be a part of a measurement $\boldsymbol{y}(t_k)$ that cannot be predicted by $\hat{\boldsymbol{y}}(t_k|t_{k-1})$. This part is termed the innovation.

Definition 2.5 - Innovations

Let $\boldsymbol{e}(t_k)$ represent the part of $\boldsymbol{y}(t_k)$ that cannot be predicted from Y^{k-1} as

$$\begin{aligned}\boldsymbol{e}(t_k) &= \boldsymbol{y}(t_k) - \hat{\boldsymbol{y}}(t_k|t_{k-1}) \\ &= \boldsymbol{y}(t_k) - \boldsymbol{C}\hat{\boldsymbol{x}}(t_k|t_{k-1})\end{aligned} \tag{2.19}$$

Thus $\boldsymbol{e}(t_k)$ represents the new information in $\boldsymbol{y}(t_k)$ not contained in Y^{k-1}. For this reason $\boldsymbol{e}(t_k)$ is called the innovation. Since $\boldsymbol{y}(t_k)$ is assumed zero mean and Gaussian distributed $\boldsymbol{e}(t_k)$ is a zero-mean Gaussian white noise process that is fully described by the covariance matrix $\boldsymbol{\Lambda}$. In the following these statistical properties will be abbreviated $NID(\boldsymbol{0},\boldsymbol{\Lambda})$. □

In what follows, it is assumed that all transient behaviour has faded away and that a steady state exists. From Y^{k-1} the sequence Y^k can then be expressed as

$$Y^k = \{(Y^{k-1})^T,\ \boldsymbol{y}(t_k)\}^T \tag{2.20}$$

Combining this relation with definition 2.5 and lemma A.5 in appendix A, which concerns the properties of orthogonal projections of Gaussian stochastic variables, the following relations for the predictor $\hat{\boldsymbol{x}}(t_k|t_{k-1})$ are obtained

$$
\begin{aligned}
\hat{\boldsymbol{x}}(t_{k+1}|t_k) &= E[\boldsymbol{x}(t_{k+1})|Y^k] \\
&= E[\boldsymbol{x}(t_{k+1})|Y^{k-1},\boldsymbol{e}(t_k)] \\
&= E[\boldsymbol{x}(t_{k+1})|Y^{k-1}] + E[\boldsymbol{x}(t_{k+1})|\boldsymbol{e}(t_k)] \\
&= \hat{\boldsymbol{x}}(t_{k+1}|t_{k-1}) + E[\boldsymbol{x}(t_{k+1})\boldsymbol{e}^T(t_k)]\Big[E[\boldsymbol{e}(t_k)\boldsymbol{e}^T(t_k)]\Big]^{-1}\boldsymbol{e}(t_k) \\
&= \hat{\boldsymbol{x}}(t_{k+1}|t_{k-1}) + \boldsymbol{K}\boldsymbol{e}(t_k)
\end{aligned}
\tag{2.21}
$$

where the matrix $\boldsymbol{K}$ is called the steady-state Kalman gain. Based on the last relation in (2.21), it is possible to formulate the steady-state Kalman filter that is associated with the noise affected state space system in (2.15).

Theorem 2.1 - The Steady-State Kalman Filter of a Stochastically Excited System

The steady-state optimal state predictor, described in terms of the Kalman filter of the state space system with stochastic input, is given by

$$
\begin{aligned}
\hat{\boldsymbol{x}}(t_{k+1}|t_k) &= \boldsymbol{A}\hat{\boldsymbol{x}}(t_k|t_{k-1}) + \boldsymbol{K}\boldsymbol{e}(t_k) \\
\boldsymbol{e}(t_k) &= \boldsymbol{y}(t_k) - \boldsymbol{C}\hat{\boldsymbol{x}}(t_k|t_{k-1})
\end{aligned}
\tag{2.22}
$$

The first equation is called the update equation and the last equation the correction equation. $\boldsymbol{K}$ is the Kalman gain matrix, that includes the description of the disturbance on the system as well as the Gaussian white noise excitation. This matrix is defined as

$$
\boldsymbol{K} = \left(\boldsymbol{A}\boldsymbol{P}\boldsymbol{C}^T + \boldsymbol{B}\boldsymbol{\Delta}\boldsymbol{D}^T + \boldsymbol{S}\right)\boldsymbol{\Lambda}^{-1}
\tag{2.23}
$$

with Λ being the covariance matrix of the innovations $\boldsymbol{e}(t_k)$ given by

$$
\boldsymbol{\Lambda} = \boldsymbol{C}\boldsymbol{P}\boldsymbol{C}^T + \boldsymbol{D}\boldsymbol{\Delta}\boldsymbol{D}^T + \boldsymbol{R}
\tag{2.24}
$$

The covariance matrix $\boldsymbol{P}$ is obtained as the positive semi definite solution to the algebraic Riccati equation

$$
\boldsymbol{P} = \boldsymbol{A}\boldsymbol{P}\boldsymbol{A}^T + \boldsymbol{B}\boldsymbol{\Delta}\boldsymbol{B}^T + \boldsymbol{Q} - \boldsymbol{K}\boldsymbol{\Lambda}\boldsymbol{K}^T
\tag{2.25}
$$

Proof:

The following proof is a modified version of the traditional Kalman filter proofs found in e.g. Goodwin et al. [31] and Hannan et al. [33]. The modification lies in the fact that the excitation is usually assumed known and therefore deterministic. Assume that the Gaussian white noise excitation $\boldsymbol{u}(t_k)$ is uncorrelated with the disturbance, i.e. with $\boldsymbol{w}(t_k)$ and $\boldsymbol{v}(t_k)$. The steady-state prediction error covariance $\boldsymbol{P}$ is then defined as

$$\begin{aligned}\boldsymbol{P} &= E\Big[\{\hat{\boldsymbol{x}}(t_k|t_{k-1}) - \boldsymbol{x}(t_k)\}\{\hat{\boldsymbol{x}}(t_k|t_{k-1}) - \boldsymbol{x}(t_k)\}^T\Big] = \Pi - \Gamma \\ \Pi &= E\Big[\boldsymbol{x}(t_k)\boldsymbol{x}^T(t_k)\Big]\,, \quad \Gamma = E\Big[\hat{\boldsymbol{x}}(t_k|t_{k-1})\hat{\boldsymbol{x}}^T(t_k|t_{k-1})\Big]\end{aligned} \tag{2.26}$$

By taking the conditional mean of the state space equation of (2.15) given the measurements $Y^{k\text{-}1}$, the following relation is obtained

$$\begin{aligned}\hat{\boldsymbol{x}}(t_{k+1}|t_{k-1}) &= E\Big[\boldsymbol{x}(t_{k+1})|Y^{k-1}\Big] \\ &= E\Big[\boldsymbol{A}\boldsymbol{x}(t_k) + \boldsymbol{B}\boldsymbol{u}(t_k) + \boldsymbol{w}(t_k)|Y^{k-1}\Big] \\ &= \boldsymbol{A}\hat{\boldsymbol{x}}(t_k|t_{k-1})\end{aligned} \tag{2.27}$$

Substituting (2.27) into (2.21) gives the update equation in (2.22). From (2.22) the following expectation is obtained

$$\begin{aligned}\Gamma &= E\Big[\Big(\boldsymbol{A}\hat{\boldsymbol{x}}(t_k|t_{k-1}) + \boldsymbol{K}\boldsymbol{e}(t_k)\Big) \times \Big(\boldsymbol{A}\hat{\boldsymbol{x}}(t_k|t_{k-1}) + \boldsymbol{K}\boldsymbol{e}(t_k)\Big)^T\Big] \\ &= \boldsymbol{A}\Gamma\boldsymbol{A}^T + \boldsymbol{K}\Lambda\boldsymbol{K}^T\end{aligned} \tag{2.28}$$

and from (2.15) Π is given by

$$\begin{aligned}\Pi &= E\Big[\Big(\boldsymbol{A}\boldsymbol{x}(t_k) + \boldsymbol{B}\boldsymbol{u}(t_k) + \boldsymbol{w}(t_k)\Big) \times \Big(\boldsymbol{A}\hat{\boldsymbol{x}}(t_k) + \boldsymbol{B}\boldsymbol{u}(t_k) + \boldsymbol{w}(t_k)\Big)^T\Big] \\ &= \boldsymbol{A}\Pi\boldsymbol{A}^T + \boldsymbol{B}\Delta\boldsymbol{B}^T + \boldsymbol{Q}\end{aligned} \tag{2.29}$$

Subtracting (2.29) from (2.28) and using (2.26) gives (2.25). Finally, the covariance $E[\boldsymbol{x}(t_{k+1})\boldsymbol{e}^T(t_k)]$ in (2.21) is given by

$$\begin{aligned}E[\boldsymbol{x}(t_{k+1})\boldsymbol{e}^T(t_k)] &= E[\Big(\boldsymbol{A}\boldsymbol{x}(t_k) + \boldsymbol{B}\boldsymbol{u}(t_k) + \boldsymbol{w}(t_k)\Big) \times \\ &\qquad \Big(\boldsymbol{C}[\boldsymbol{x}(t_k) - \hat{\boldsymbol{x}}(t_k|t_{k-1})] + \boldsymbol{D}\boldsymbol{u}(t_k) + \boldsymbol{v}(t_k)\Big)^T] \\ &= \boldsymbol{A}E[\boldsymbol{x}(t_k)\Big(\boldsymbol{x}(t_k) - \hat{\boldsymbol{x}}(t_k|t_{k-1})\Big)^T]\boldsymbol{C}^T + \\ &\qquad \boldsymbol{B}E[\boldsymbol{u}(t_k)\boldsymbol{u}^T(t_k)]\boldsymbol{D}^T + E[\boldsymbol{w}(t_k)\boldsymbol{v}^T(t_k)] \\ &= \boldsymbol{A}\boldsymbol{P}\boldsymbol{C}^T + \boldsymbol{B}\Delta\boldsymbol{D}^T + \boldsymbol{S}\end{aligned} \tag{2.30}$$

since $\boldsymbol{x}(t_k) - \hat{\boldsymbol{x}}(t_k|t_{k-1})$ is orthogonal to $\hat{\boldsymbol{x}}(t_k|t_{k-1})$. The innovation covariance matrix is given by

$$
\begin{aligned}
\mathbf{\Lambda} &= E[\boldsymbol{e}(t_k)\boldsymbol{e}^T(t_k)] \\
&= E\Big[\Big(\boldsymbol{C}[\boldsymbol{x}(t_k) - \hat{\boldsymbol{x}}(t_k|t_{k-1})] + \boldsymbol{D}\boldsymbol{u}(t_k) + \boldsymbol{v}(t_k)\Big) \times \\
&\qquad \Big(\boldsymbol{C}[\boldsymbol{x}(t_k) - \hat{\boldsymbol{x}}(t_k|t_{k-1})] + \boldsymbol{D}\boldsymbol{u}(t_k) + \boldsymbol{v}(t_k)\Big)^T\Big] \\
&= \boldsymbol{CPC}^T + \boldsymbol{D}\mathbf{\Delta}\boldsymbol{D}^T + \boldsymbol{R}
\end{aligned}
\tag{2.31}
$$

By combining (2.30) with $\mathbf{\Lambda}$, the Kalman gain in (2.23) is obtained. Inserting (2.19) in definition 2.5 results in the correction equation of the Kalman filter recursion in (2.22). □

Thus, knowing the measured output $\boldsymbol{y}(t_k)$, the system matrices $\{\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{C}, \boldsymbol{D}\}$ and the covariance matrices $\{\mathbf{\Delta}, \boldsymbol{Q}, \boldsymbol{R}, \boldsymbol{S}\}$, it is possible to calculate the covariance matrix $\boldsymbol{P}$, the Kalman gain $\boldsymbol{K}$, and the innovation covariance matrix $\mathbf{\Lambda}$ before the filtering starts.

2.2.3 The Innovation State Space System

In system identification, it is the system matrices that are unknown. This means that the above Kalman filter is unknown. However, the filter reveals how the state space system that accounts for the presence of noise should look like. This special state space system is known as the innovation state space system, see e.g. Ljung [71].

$$
\begin{aligned}
\hat{\boldsymbol{x}}(t_{k+1}|t_k) &= \boldsymbol{A}\hat{\boldsymbol{x}}(t_k|t_{k-1}) + \boldsymbol{K}\boldsymbol{e}(t_k) \;, \quad \boldsymbol{e}(t_k) \in NID(\mathbf{0}, \mathbf{\Lambda}) \\
\boldsymbol{y}(t_k) &= \boldsymbol{C}\hat{\boldsymbol{x}}(t_k|t_{k-1}) + \boldsymbol{e}(t_k)
\end{aligned}
\tag{2.32}
$$

This state space system is obtained from the steady-state Kalman filter by rearranging the correction equation of (2.22). This system is assumed to be driven by a known innovation process, which is obtained from contributions of the stochastic input $\boldsymbol{u}(t_k)$, and the disturbance represented by $\boldsymbol{w}(t_k)$ and $\boldsymbol{v}(t_k)$. The dynamic behaviour of the system is still described by the system matrices $\boldsymbol{A}$ and $\boldsymbol{C}$, as in the noise-free case. This implies that e.g. the observability matrix $\boldsymbol{Q}_o(m)$ of an m-dimensional innovation state space system also is given by (2.10). On the other hand, to obtain the reachability matrix $\boldsymbol{Q}_r(m)$ of the innovation system the input matrix $\boldsymbol{B}$ in (2.10) must be substituted by the Kalman gain $\boldsymbol{K}$. Similarity transformations, can still be performed according to definition 2.2 by setting $\boldsymbol{B} = \boldsymbol{K}$ and $\boldsymbol{D} = \boldsymbol{I}$. A solution of (2.32) for $\boldsymbol{y}(t_{k+l})$, $l \geq 0$, can be obtained in a similar fashion as (2.4) by recursive substitutions of itself to yield

$$
\begin{aligned}
&\boldsymbol{y}(t_{k+l}) = \boldsymbol{CA}^l\hat{\boldsymbol{x}}(t_k|t_{k-1}) + \sum_{j=1}^{l}\boldsymbol{CA}^{l-j}\boldsymbol{K}\boldsymbol{e}(t_{k+j-1}) + \boldsymbol{e}(t_{k+l}) \\
&\boldsymbol{e}(t_k) \in NID(\mathbf{0}, \mathbf{\Lambda}) \;, \quad \hat{\boldsymbol{x}}(t_k|t_{k-1}) \equiv \hat{\boldsymbol{x}}_0 \;, \quad E[\hat{\boldsymbol{x}}_0] = \mathbf{0}
\end{aligned}
\tag{2.33}
$$

Again the solution is split into a homogeneous part, which is represented by the first right-hand term, and a particular part represented by the second right-hand term.

So in conclusion:

☞ *The innovation state space representation* (2.22) *models the dynamic behaviour of a linear and time-invariant discrete-time system, subjected to Gaussian white noise excitation and Gaussian white noise disturbance. Therefore, it is a more general description than the representation given in* (2.1).

2.3 ARMAV Modelling of Discrete-Time Systems

This section shows how the stochastic state space system is related to a stochastic difference equation system. If the applied excitation is Gaussian white noise, this stochastic difference equation system is equivalent to the ARMAV model defined in (1.2). This section will only focus on representing a discrete-time linear and time-invariant system that is not affected by disturbance. The inclusion of the disturbance will be shown in the next section.

Just as a continuous-time linear and time-variant system can be represented by a differential equation system, it is possible to represent a discrete-time linear and time-invariant system by a difference equation system. For a p-variate system the equation is

$$\begin{gathered} \boldsymbol{y}(t_k) + \boldsymbol{A}_1\boldsymbol{y}(t_{k-1}) + \ldots + \boldsymbol{A}_n\boldsymbol{y}(t_{k-n}) = f(\boldsymbol{u}(t_k)) \\ \boldsymbol{u}(t_k) \in NID(\boldsymbol{0},\boldsymbol{\Delta}) \end{gathered} \tag{2.34}$$

The system response $\boldsymbol{y}(t_k)$ is now described by a homogeneous part, which is an nth-order auto-regressive matrix polynomial, and a nonhomogeneous part being a function of the Gaussian white noise process $\boldsymbol{u}(t_k)$. It is assumed that $\boldsymbol{y}(t_k)$ and $\boldsymbol{u}(t_k)$ are vectors of dimension $p \times 1$, which implies that all the n auto-regressive coefficient matrices $\boldsymbol{A}_i$ will have the dimensions $p \times p$. It is claimed that a general solution of (2.34), for an arbitrary $k \geq 0$, can be represented on the following form

$$\boldsymbol{y}(t_k) = \boldsymbol{y}_h(t_k) + \boldsymbol{y}_p(t_k) \tag{2.35}$$

with $\boldsymbol{y}_p(t_k)$ being a fixed particular solution, i.e. one among many solutions, and $\boldsymbol{y}_h(t_k)$ a general solution of the homogeneous equation

$$\boldsymbol{y}(t_k) + \boldsymbol{A}_1\boldsymbol{y}(t_{k-1}) + \ldots + \boldsymbol{A}_n\boldsymbol{y}(t_{k-n}) = \boldsymbol{0} \tag{2.36}$$

Since $\boldsymbol{y}(t_k)$ in the homogeneous case can be sequentially computed by

$$\boldsymbol{y}(t_k) = -\boldsymbol{A}_1\boldsymbol{y}(t_{k-1}) - \ldots - \boldsymbol{A}_n\boldsymbol{y}(t_{k-n}) \tag{2.37}$$

the set of solutions of the homogeneous equations is a linear space of dimension np. The general solution is therefore given by (2.4) for an m-dimensional state vector, where $m = np$, see Gohberg et al. [30]. The first part of (2.4) is equal to $\boldsymbol{y}_h(t_k)$ and the last part to $\boldsymbol{y}_p(t_k)$ in (2.35). Since the general solution of a minimal system is not restricted to any specific realization, the following relations are in principle valid for any arbitrary m-dimensional minimal realizations. This is due to definition 2.2, stating that all minimal realizations are similar. From (2.4) the general solution to the homogeneous part of the solution is given by

$$\boldsymbol{y}_h(t_{k+l}) = \boldsymbol{C}\boldsymbol{A}^l\boldsymbol{x}(t_k) \;, \quad \begin{cases} l = 0,\ 1,\ \ldots \\ \boldsymbol{x}(t_k) \equiv \boldsymbol{x}_0 \end{cases} \tag{2.38}$$

Inserting (2.38) into (2.36) yields

$$\left(\boldsymbol{C}\boldsymbol{A}^n + \boldsymbol{A}_1\boldsymbol{C}\boldsymbol{A}^{n-1} + \ldots + \boldsymbol{A}_{n-1}\boldsymbol{C}\boldsymbol{A} + \boldsymbol{A}_n\boldsymbol{C}\right)\boldsymbol{x}(t_k) = \boldsymbol{0} \tag{2.39}$$

which can only be fulfilled for an arbitrary choice of $\boldsymbol{x}(t_k)$, if the contents inside the parentheses equals zero, i.e.

$$\boldsymbol{C}\boldsymbol{A}^n + \boldsymbol{A}_1\boldsymbol{C}\boldsymbol{A}^{n-1} + \ldots + \boldsymbol{A}_{n-1}\boldsymbol{C}\boldsymbol{A} + \boldsymbol{A}_n\boldsymbol{C} = \boldsymbol{0} \tag{2.40}$$

This relation reveals an important link between the auto-regressive coefficient matrices and the system matrices $\boldsymbol{A}$ and $\boldsymbol{C}$. The importance becomes clear in the following theorem.

Theorem 2.2 - The ARMAV(n,n) Model - Without Noise Modelling

A minimal realization of the state space system (2.1), described by $\{\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{C}, \boldsymbol{D}, \boldsymbol{\Delta}\}$, can be represented by an ARMAV(n,n) model, defined as

$$\begin{aligned} \boldsymbol{y}(t_k) + \boldsymbol{A}_1\boldsymbol{y}(t_{k-1}) + \boldsymbol{A}_2\boldsymbol{y}(t_{k-2}) + \ldots + \boldsymbol{A}_n\boldsymbol{y}(t_{k-n}) = \\ \boldsymbol{B}_0\boldsymbol{u}(t_k) + \boldsymbol{B}_1\boldsymbol{u}(t_{k-1}) + \ldots + \boldsymbol{B}_n\boldsymbol{u}(t_{k-n}) \\ \boldsymbol{u}(t_k) \in NID(\boldsymbol{0}, \boldsymbol{\Delta}) \end{aligned} \tag{2.41}$$

if the state space realization fulfils the dimensional requirement that the state dimension m, divided by the number of channels p, equals an integral value n. The n auto-regressive matrix coefficients, which all have the dimensions $p \times p$, are then given by

$$\begin{bmatrix} \boldsymbol{A}_n & \boldsymbol{A}_{n-1} & . & . & \boldsymbol{A}_2 & \boldsymbol{A}_1 \end{bmatrix} = -\boldsymbol{C}\boldsymbol{A}^n \boldsymbol{Q}_o^{-1}(n) \tag{2.42}$$

The $n+1$ moving average matrix coefficients, also of dimensions $p \times p$, are given by

$$\begin{bmatrix} \boldsymbol{B}_n & \boldsymbol{B}_{n-1} & . & . & \boldsymbol{B}_1 & \boldsymbol{B}_0 \end{bmatrix} = \begin{bmatrix} \boldsymbol{A}_n & \boldsymbol{A}_{n-1} & . & . & \boldsymbol{A}_1 & \boldsymbol{I} \end{bmatrix} \boldsymbol{T}(n+1) \tag{2.43}$$

with $\boldsymbol{T}(n+1)$ defined as

$$\boldsymbol{T}(n+1) = \begin{bmatrix} \boldsymbol{D} & \boldsymbol{0} & . & . & \boldsymbol{0} & \boldsymbol{0} \\ \boldsymbol{CB} & \boldsymbol{D} & . & . & \boldsymbol{0} & \boldsymbol{0} \\ . & . & . & . & . & . \\ . & . & . & . & . & . \\ \boldsymbol{C}\boldsymbol{A}^{n-2}\boldsymbol{B} & \boldsymbol{C}\boldsymbol{A}^{n-3}\boldsymbol{B} & . & . & \boldsymbol{D} & \boldsymbol{0} \\ \boldsymbol{C}\boldsymbol{A}^{n-1}\boldsymbol{B} & \boldsymbol{C}\boldsymbol{A}^{n-2}\boldsymbol{B} & . & . & \boldsymbol{CB} & \boldsymbol{D} \end{bmatrix} \tag{2.44}$$

where it should be noticed that $\boldsymbol{B}_0 = \boldsymbol{D}$. The Gaussian white noise input process $\boldsymbol{u}(t_k)$ is unaffected by this transformation.

Proof:

The auto-regressive coefficient matrices are obtained from (2.40) by the following rearrangement

$$\begin{bmatrix} \boldsymbol{A}_n & \boldsymbol{A}_{n-1} & . & . & \boldsymbol{A}_1 \end{bmatrix} \boldsymbol{Q}_o(n) = -\boldsymbol{C}\boldsymbol{A}^n \tag{2.45}$$

where the observability matrix, defined in (2.10) has been introduced. It is seen that the auto-regressive coefficient matrices can be determined from the system matrices $\boldsymbol{A}$ and $\boldsymbol{C}$ as

$$\begin{bmatrix} \boldsymbol{A}_n & \boldsymbol{A}_{n-1} & . & . & \boldsymbol{A}_1 \end{bmatrix} = -\boldsymbol{C}\boldsymbol{A}^n \boldsymbol{Q}_o^{-1}(n) \tag{2.46}$$

This equation is equal to (2.42). The auto-regressive coefficient matrices can only be determined for state space realizations where $\boldsymbol{Q}_o(n)$ is non-singular. If $\boldsymbol{Q}_o(n)$ is singular, methods for determination of the coefficient matrices may be found in Gawronski et al. [24] and Kailath [48]. The general solution (2.4) can be used to find an expression for $f(\boldsymbol{u}(t_k))$ in the following way. From the observation equation of (2.1) and (2.4), construct the following set of equations

$$\begin{bmatrix} \boldsymbol{y}(t_{k-n}) \\ \boldsymbol{y}(t_{k-n+1}) \\ . \\ . \\ \boldsymbol{y}(t_{k-1}) \\ \boldsymbol{y}(t_k) \end{bmatrix} = \begin{bmatrix} \boldsymbol{C} \\ \boldsymbol{CA} \\ . \\ . \\ \boldsymbol{CA}^{n-1} \\ \boldsymbol{CA}^{n} \end{bmatrix} \boldsymbol{x}(t_{k-n}) + \begin{bmatrix} \boldsymbol{D} & \boldsymbol{0} & . & . & \boldsymbol{0} & \boldsymbol{0} \\ \boldsymbol{CB} & \boldsymbol{D} & . & . & \boldsymbol{0} & \boldsymbol{0} \\ . & . & . & . & . & . \\ . & . & . & . & . & . \\ \boldsymbol{CA}^{n-2}\boldsymbol{B} & \boldsymbol{CA}^{n-3}\boldsymbol{B} & . & . & \boldsymbol{D} & \boldsymbol{0} \\ \boldsymbol{CA}^{n-1}\boldsymbol{B} & \boldsymbol{CA}^{n-2}\boldsymbol{B} & . & . & \boldsymbol{CB} & \boldsymbol{D} \end{bmatrix} \begin{bmatrix} \boldsymbol{u}(t_{k-n}) \\ \boldsymbol{u}(t_{k-n+1}) \\ . \\ . \\ \boldsymbol{u}(t_{k-1}) \\ \boldsymbol{u}(t_k) \end{bmatrix} \tag{2.47}$$

which can compactly be written as

$$\boldsymbol{Y}(t_{k-n},t_k) = \boldsymbol{Q}_o(n+1)\boldsymbol{x}(t_k) + \boldsymbol{T}(n+1)\boldsymbol{U}(t_{k-n},t_k) \tag{2.48}$$

Now introduce the matrix containing the auto-regressive coefficient matrices obtained from (2.42)

$$\boldsymbol{P} = \begin{bmatrix} \boldsymbol{A}_n & \boldsymbol{A}_{n-1} & . & . & \boldsymbol{A}_1 & \boldsymbol{I} \end{bmatrix} \tag{2.49}$$

and notice that (2.40) can be written as $\boldsymbol{PQ}_o(n+1) = \boldsymbol{0}$. Multiplying (2.48) from the left by (2.49) then yields

$$\begin{aligned} \boldsymbol{PY}(t_{k-n},t_k) &= \boldsymbol{PQ}_o(n+1)\boldsymbol{x}(t_k) + \boldsymbol{PT}(n+1)\boldsymbol{U}(t_{k-n},t_k) \\ &= \boldsymbol{PT}(n+1)\boldsymbol{U}(t_{k-n},t_k) \end{aligned} \tag{2.50}$$

The left hand-side of (2.50) is exactly the auto-regressive part of the difference equation system in (2.34), and the right-hand term is therefore equivalent to $f(\boldsymbol{u}(t_k))$, which is an nth order matrix polynomial in $\boldsymbol{u}(t_k)$. Since $\boldsymbol{u}(t_k)$ is a Gaussian white noise this matrix polynomial is a true moving average, which means that (2.50) really is an ARMAV(n,n) model. □

It is seen that it is possible to establish an equivalent ARMAV model to the minimal stochastic state space system. This ARMAV model is independent of the actual realization, i.e. no matter how a p-variate linear and time-invariant discrete-time system subjected to Gaussian white noise excitation is realized, the equivalent ARMAV model will always be the same.

So in conclusion

☞ *Any stochastic state space system, where the state space dimension m divided by the number of observed output p equals an integral value n can be represented by an ARMAV model if the observability matrix* $\mathbf{Q}_o(n)$ *is non-singular*

The ARMAV model can also be defined in terms of its transfer function $\boldsymbol{H}(q)$ introduced in (1.2) as

$$\begin{aligned} \boldsymbol{y}(t_k) &= \boldsymbol{A}^{-1}(q)\boldsymbol{B}(q)\boldsymbol{u}(t_k) = \boldsymbol{H}(q)\boldsymbol{u}(t_k) \\ \boldsymbol{A}(q) &= \boldsymbol{I} + \boldsymbol{A}_1 q^{-1} + \boldsymbol{A}_2 q^{-2} + \ldots + \boldsymbol{A}_n q^{-n} \\ \boldsymbol{B}(q) &= \boldsymbol{B}_0 + \boldsymbol{B}_1 q^{-1} + \boldsymbol{B}_2 q^{-2} + \ldots + \boldsymbol{B}_n q^{-n} \end{aligned} \tag{2.51}$$

where $\boldsymbol{A}(q)$ and $\boldsymbol{B}(q)$ are the auto-regressive and the moving average matrix polynomials. As was the case for the stochastic state space system (2.1), this ARMAV model is not very useful in practical system identification. An ARMAV model based on the innovation state space system will be more useful in this context, since such a model will take the presence of disturbance into account.

2.4 ARMAV Modelling of Discrete-Time Systems Affected by Noise

Establishing an ARMAV model on the basis of the innovation state space system is quite simple, since all the steps involved in such a conversion are the same as shown in the previous section. The proof of the following theorem will therefore resemble the proof of theorem 2.2, and will therefore not be given.

Theorem 2.3 - The ARMAV(*n*,*n*) Model - With Noise Modelling

A minimal realization of the innovation state space system (2.32), described by $\{\boldsymbol{A}, \boldsymbol{K}, \boldsymbol{C}, \Lambda\}$, can be equivalently represented by an ARMAV(n,n) model, defined as

$$\begin{aligned} \boldsymbol{y}(t_k) + \boldsymbol{A}_1\boldsymbol{y}(t_{k-1}) + \boldsymbol{A}_2\boldsymbol{y}(t_{k-2}) + \ldots + \boldsymbol{A}_n\boldsymbol{y}(t_{k-n}) &= \\ \boldsymbol{e}(t_k) + \boldsymbol{C}_1\boldsymbol{e}(t_{k-1}) + \ldots + \boldsymbol{C}_n\boldsymbol{e}(t_{k-n}) & \\ \boldsymbol{e}(t_k) \in NID(\mathbf{0}, \Lambda) & \end{aligned} \tag{2.52}$$

if the state space realization fulfils the dimensional requirement that the state dimension m, divided by the number of channels p, equals an integral value n. The n auto-regressive matrix coefficients, all of dimension $p \times p$, are given by

$$\begin{bmatrix} \boldsymbol{A}_n & \boldsymbol{A}_{n-1} & . & . & \boldsymbol{A}_2 & \boldsymbol{A}_1 \end{bmatrix} = -\boldsymbol{C}\boldsymbol{A}^n\boldsymbol{Q}_o^{-1}(n) \tag{2.53}$$

The n moving average matrix coefficients, also of dimensions $p \times p$, are given by

$$\begin{bmatrix} \boldsymbol{C}_n & \boldsymbol{C}_{n-1} & . & . & \boldsymbol{C}_1 & \boldsymbol{I} \end{bmatrix} = \begin{bmatrix} \boldsymbol{A}_n & \boldsymbol{A}_{n-1} & . & . & \boldsymbol{A}_1 & \boldsymbol{I} \end{bmatrix} \boldsymbol{T}(n+1) \tag{2.54}$$

with $\boldsymbol{T}(n+1)$ defined as

$$\boldsymbol{T}(n+1) = \begin{bmatrix} \boldsymbol{I} & \boldsymbol{0} & . & . & \boldsymbol{0} & \boldsymbol{0} \\ \boldsymbol{CK} & \boldsymbol{I} & . & . & \boldsymbol{0} & \boldsymbol{0} \\ . & . & . & . & . & . \\ . & . & . & . & . & . \\ \boldsymbol{CA}^{n-2}\boldsymbol{K} & \boldsymbol{CA}^{n-3}\boldsymbol{K} & . & . & \boldsymbol{I} & \boldsymbol{0} \\ \boldsymbol{CA}^{n-1}\boldsymbol{K} & \boldsymbol{CA}^{n-2}\boldsymbol{K} & . & . & \boldsymbol{CK} & \boldsymbol{I} \end{bmatrix} \tag{2.55}$$

The innovation process $\boldsymbol{e}(t_k)$ is unaffected by this transformation, and is still described the covariance matrix Λ. □

This ARMAV model can also be defined in terms of its transfer function $\boldsymbol{H}(q)$ as

$$\begin{aligned} \boldsymbol{y}(t_k) &= \boldsymbol{A}^{-1}(q)\boldsymbol{C}(q)\boldsymbol{e}(t_k) = \boldsymbol{H}(q)\boldsymbol{e}(t_k) \\ \boldsymbol{A}(q) &= \boldsymbol{I} + \boldsymbol{A}_1 q^{-1} + \boldsymbol{A}_2 q^{-2} + \ldots + \boldsymbol{A}_n q^{-n} \\ \boldsymbol{C}(q) &= \boldsymbol{I} + \boldsymbol{C}_1 q^{-1} + \boldsymbol{C}_2 q^{-2} + \ldots + \boldsymbol{C}_n q^{-n} \end{aligned} \tag{2.56}$$

where $\boldsymbol{A}(q)$ and $\boldsymbol{C}(q)$ are the auto-regressive and the moving average matrix polynomials. In chapter 5, system identification of multivariate linear and time-invariant discrete-time systems will be considered, and the models that will be used are the innovation state space system and the ARMAV model derived in this section. However, it is still important to be aware of the difference between the ARMAV model that accounts for the presence of noise and the model that does not. This is shown in chapter 4.

2.5 A State Space Realization of an ARMAV Model

Until now, the actual parameterization of the state space system has not been considered. This has not been necessary since the noise modelling and the conversion to ARMAV models are both procedures that are independent of the choice of realization. However, to convert an ARMAV model to state space makes it necessary to choose a realization. There is no unique way of doing this conversion. The choice of a state space realization may not be easy. If the realization is used as part of some algebraic manipulations, it is a good idea to choose a realization that is easy to construct from the auto-regressive and moving average coefficient matrices. On the other hand, if the realization is used in system identification software, the realization must be well-conditioned in order to be numerically efficient. In this section a state space realization that is easy to obtain from the coefficient matrices of the ARMAV model will be presented. This realization is used on several occasions in chapters 3-6 and is known as the observability canonical state space realization. This realization is constructed on the basis of the auto-regressive coefficient matrices and the impulse response function of the ARMAV model.

2.5.1 The Impulse Response Function of ARMAV Models

The ARMAV(n,n) model can equivalently be defined as a convolution of its impulse response function $\boldsymbol{h}(j)$ and the Gaussian white noise input.

$$\boldsymbol{y}(t_k) = \sum_{j=0}^{\infty} \boldsymbol{h}(j)\boldsymbol{u}(t_{k-j}) \ , \qquad \boldsymbol{u}(t_k) \in NID(\mathbf{0},\boldsymbol{\Delta}) \tag{2.57}$$

In this section, relations involving the impulse response function of this model will be given. Based on these relations, it is possible to construct the observability canonical state space realization. The $p \times p$ transfer function defined in (2.51) can also be expressed in terms of the impulse response function and the delay operator q as

$$\boldsymbol{H}(q) = \sum_{j=0}^{\infty} \boldsymbol{h}(j) q^{-j} \tag{2.58}$$

If $\boldsymbol{y}(t_k)$ is substituted in (2.51) with (2.57) and (2.58) the following relation is obtained

$$\begin{aligned} \left(\boldsymbol{I} + \boldsymbol{A}_1 q^{-1} + \ldots + \boldsymbol{A}_n q^{-n}\right)\left(\sum_{j=0}^{\infty} \boldsymbol{h}(j) q^{-j} \boldsymbol{u}(t_k)\right) = \\ \left(\boldsymbol{B}_0 + \boldsymbol{B}_1 q^{-1} + \ldots + \boldsymbol{B}_n q^{-n}\right)\boldsymbol{u}(t_k) \end{aligned} \tag{2.59}$$

By comparing the coefficients of equal powers of q, the following relationships between the coefficient matrices of the ARMAV model and its impulse response function $\boldsymbol{h}(k)$, for $k = 0$ to n, can be obtained

$$
\begin{aligned}
&\boldsymbol{h}(0) = \boldsymbol{B}_0 \\
&\boldsymbol{h}(1) + \boldsymbol{A}_1\boldsymbol{h}(0) = \boldsymbol{B}_1 \\
&\boldsymbol{h}(2) + \boldsymbol{A}_1\boldsymbol{h}(1) + \boldsymbol{A}_2\boldsymbol{h}(0) = \boldsymbol{B}_2 \\
&\quad . \\
&\quad . \\
&\boldsymbol{h}(n) + \boldsymbol{A}_1\boldsymbol{h}(n-1) + \ldots + \boldsymbol{A}_n\boldsymbol{h}(0) = \boldsymbol{B}_n
\end{aligned}
\tag{2.60}
$$

Based on (2.60) the following two recursive relations can be derived

$$
\boldsymbol{h}(k) = -\boldsymbol{A}_1\boldsymbol{h}(k-1) - \ldots - \boldsymbol{A}_k\boldsymbol{h}(0) + \boldsymbol{B}_k \,, \quad 0 \le k \le n \tag{2.61}
$$

$$
\boldsymbol{h}(k) = -\boldsymbol{A}_1\boldsymbol{h}(k-1) - \ldots - \boldsymbol{A}_n\boldsymbol{h}(k-n) \quad , \quad k > n \tag{2.62}
$$

The first relation should be used if $k \le n$, and the second if $k > n$. If the moving average polynomial is of lower order than the order of the auto-regressive polynomial, given by n, the remaining moving average coefficient matrices, $\boldsymbol{B}_i$ are zeroed in the above expressions. The last n equations of (2.60) can be rearranged as

$$
\begin{bmatrix} \boldsymbol{h}(1) \\ \boldsymbol{h}(2) \\ . \\ . \\ \boldsymbol{h}(n) \end{bmatrix}
=
\begin{bmatrix}
\boldsymbol{I} & \boldsymbol{0} & . \ . & \boldsymbol{0} & \boldsymbol{0} \\
\boldsymbol{A}_1 & \boldsymbol{I} & . \ . & \boldsymbol{0} & \boldsymbol{0} \\
. & . & . \ . & . & . \\
. & . & . \ . & . & . \\
\boldsymbol{A}_{n-1} & \boldsymbol{A}_{n-2} & . \ . & \boldsymbol{A}_1 & \boldsymbol{I}
\end{bmatrix}^{-1}
\begin{bmatrix}
\boldsymbol{B}_1 - \boldsymbol{A}_1\boldsymbol{B}_0 \\
\boldsymbol{B}_2 - \boldsymbol{A}_2\boldsymbol{B}_0 \\
. \\
. \\
\boldsymbol{B}_n - \boldsymbol{A}_n\boldsymbol{B}_0
\end{bmatrix}
\tag{2.63}
$$

which can be used for initialization of (2.61) and (2.62).

2.5.2 The Observability Canonical State Space Realization

The above relations can be used for prediction of the most probable value of the response $\boldsymbol{y}(t_{k+m})$ given the $\boldsymbol{u}(t_k)$ and $\boldsymbol{y}(t_s)$, for time steps $s \le k$. From (2.57) it can be verified that

$$
\boldsymbol{y}(t_{k+m}) = \sum_{j=0}^{\infty} \boldsymbol{h}(j)\boldsymbol{u}(t_{k+m-j}) = \sum_{j=0}^{m-1} \boldsymbol{h}(j)\boldsymbol{u}(t_{k+m-j}) + \sum_{j=m}^{\infty} \boldsymbol{h}(j)\boldsymbol{u}(t_{k+m-j}) \tag{2.64}
$$

The first sum has zero mean and it only involves inputs $\boldsymbol{u}(t_{k+1})$ to $\boldsymbol{u}(t_{k+m})$ and is as such independent of what has happened up to the time t_k. The second sum is known at the time t_k and the conditional mean $\hat{\boldsymbol{y}}(t_{k+m}|t_k)$ of $\boldsymbol{y}(t_{k+m})$, given response $\boldsymbol{y}(t_s)$, for time steps $s \leq k$, is as such equal to

$$\hat{\boldsymbol{y}}(t_{k+m}|t_k) = \sum_{j=m}^{\infty} \boldsymbol{h}(j)\boldsymbol{u}(t_{k+m-j}) \tag{2.65}$$

This equation can be formulated as

$$\begin{aligned} \hat{\boldsymbol{y}}(t_{k+m}|t_k) &= \sum_{j=m+1}^{\infty} \boldsymbol{h}(j)\boldsymbol{u}(t_{k+m-j}) + \boldsymbol{h}(m)\boldsymbol{u}(t_k) \\ &= \hat{\boldsymbol{y}}(t_{k+m}|t_{k-1}) + \boldsymbol{h}(m)\boldsymbol{u}(t_k) \end{aligned} \tag{2.66}$$

and for $m = 0$ this equation is equal to

$$\boldsymbol{y}(t_k) = \hat{\boldsymbol{y}}(t_k|t_{k-1}) + \boldsymbol{B}_0\boldsymbol{u}(t_k) \tag{2.67}$$

since the most probable value of $\hat{\boldsymbol{y}}(t_k|t_k)$ is $\boldsymbol{y}(t_k)$ itself. If (2.65) is combined with (2.61), and if $m \geq n$ the following result can be obtained

$$\begin{aligned} \hat{\boldsymbol{y}}(t_{k+m}|t_k) &= \sum_{j=m}^{\infty} \boldsymbol{h}(j)\boldsymbol{u}(t_{k+m-j}) \\ &= \sum_{j=m}^{\infty} \left(-\sum_{i=1}^{n} \boldsymbol{A}_i\boldsymbol{h}(j-i) + \boldsymbol{B}_m \right) \boldsymbol{u}(t_{k+m-j}) \\ &= -\sum_{i=1}^{n} \boldsymbol{A}_i \left(\sum_{j=m}^{\infty} \boldsymbol{h}(j-i)\boldsymbol{u}(t_{k+m-j}) \right) + \boldsymbol{B}_m\boldsymbol{u}(t_k) \\ &= -\sum_{i=1}^{n} \boldsymbol{A}_i\hat{\boldsymbol{y}}(t_{k+m-i}|t_k) + \boldsymbol{B}_m\boldsymbol{u}(t_k) \end{aligned} \tag{2.68}$$

Inserting this result into (2.66) yields

$$\begin{aligned} \hat{\boldsymbol{y}}(t_{k+m}|t_k) &= -\sum_{i=1}^{n} \boldsymbol{A}_i\hat{\boldsymbol{y}}(t_{k+m-i}|t_{k-1}) - \sum_{i=1}^{n} \boldsymbol{A}_i\boldsymbol{h}(m-i)\boldsymbol{u}(t_k) + \boldsymbol{B}_m\boldsymbol{u}(t_k) \\ &= -\sum_{i=1}^{n} \boldsymbol{A}_i\hat{\boldsymbol{y}}(t_{k+m-i}|t_{k-1}) + \boldsymbol{h}(m)\boldsymbol{u}(t_k) \end{aligned} \tag{2.69}$$

Based on these relations, it is possible to formulate the desired state space realization.

Theorem 2.4 - A State Space Realization of the ARMAV Model - Without Noise Modelling

A state space realization of the p-variate ARMAV(n,n) model (2.41) is given by

$$\begin{aligned} \boldsymbol{x}(t_{k+1}) &= \boldsymbol{A}\boldsymbol{x}(t_k) + \boldsymbol{B}\boldsymbol{u}(t_k) \ , \quad \boldsymbol{u}(t_k) \in NID(\boldsymbol{0},\boldsymbol{\Delta}) \\ \boldsymbol{y}(t_k) &= \boldsymbol{C}\boldsymbol{x}(t_k) + \boldsymbol{D}\boldsymbol{u}(t_k) \end{aligned} \tag{2.70}$$

with {$\boldsymbol{A}$, $\boldsymbol{B}$, $\boldsymbol{C}$, $\boldsymbol{D}$} defined as

$$\boldsymbol{A} = \begin{bmatrix} \boldsymbol{0} & \boldsymbol{I} & . & . & \boldsymbol{0} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} & . & . & \boldsymbol{0} & \boldsymbol{0} \\ . & . & . & . & . & . \\ . & . & . & . & . & . \\ \boldsymbol{0} & \boldsymbol{0} & . & . & \boldsymbol{0} & \boldsymbol{I} \\ -\boldsymbol{A}_n & -\boldsymbol{A}_{n-1} & . & . & -\boldsymbol{A}_2 & -\boldsymbol{A}_1 \end{bmatrix} \tag{2.71}$$

$$\boldsymbol{B} = \begin{bmatrix} \boldsymbol{I} & \boldsymbol{0} & . & . & \boldsymbol{0} & \boldsymbol{0} \\ \boldsymbol{A}_1 & \boldsymbol{I} & . & . & \boldsymbol{0} & \boldsymbol{0} \\ . & . & . & . & . & . \\ . & . & . & . & . & . \\ \boldsymbol{A}_{n-1} & \boldsymbol{A}_{n-2} & . & . & \boldsymbol{A}_1 & \boldsymbol{I} \end{bmatrix}^{-1} \begin{bmatrix} \boldsymbol{B}_1 - \boldsymbol{A}_1\boldsymbol{B}_0 \\ \boldsymbol{B}_2 - \boldsymbol{A}_2\boldsymbol{B}_0 \\ . \\ . \\ \boldsymbol{B}_n - \boldsymbol{A}_n\boldsymbol{B}_0 \end{bmatrix} \tag{2.72}$$

$$\boldsymbol{C} = \begin{bmatrix} \boldsymbol{I} & \boldsymbol{0} & . & . & \boldsymbol{0} & \boldsymbol{0} \end{bmatrix} \tag{2.73}$$

$$\boldsymbol{D} = \boldsymbol{B}_0 \tag{2.74}$$

and the state vector $\boldsymbol{x}(t_k)$ defined as

$$
\boldsymbol{x}(t_k) = \begin{bmatrix} \hat{\boldsymbol{y}}(t_k|t_{k-1}) \\ \hat{\boldsymbol{y}}(t_{k+1}|t_{k-1}) \\ . \\ . \\ \hat{\boldsymbol{y}}(t_{k+n-2}|t_{k-1}) \\ \hat{\boldsymbol{y}}(t_{k+n-1}|t_{k-1}) \end{bmatrix} \tag{2.75}
$$

This state space realization is called an observability canonical realization. Observability because the associated observability matrix $\boldsymbol{Q}_o(n)$ is a unity matrix. The $np \times np$ transition matrix $\boldsymbol{A}$ is in its present form also known as the companion matrix for the auto-regressive matrix polynomial. The $np \times p$ input matrix $\boldsymbol{B}$ consists of n impulse response functions. The $p \times np$ observation matrix $\boldsymbol{C}$ ensures that only the system response $\boldsymbol{y}(t_k)$ is returned.

Proof:

Combine (2.69) and (2.66), with increasing prediction horizon, to yield the following set of equations

$$
\begin{aligned}
\hat{\boldsymbol{y}}(t_{k+1}|t_k) &= \hat{\boldsymbol{y}}(t_{k+1}|t_{k-1}) + \boldsymbol{h}(1)\boldsymbol{u}(t_k) \\
\hat{\boldsymbol{y}}(t_{k+2}|t_k) &= \hat{\boldsymbol{y}}(t_{k+2}|t_{k-1}) + \boldsymbol{h}(2)\boldsymbol{u}(t_k) \\
&. \\
&. \\
\hat{\boldsymbol{y}}(t_{k+n-1}|t_k) &= \hat{\boldsymbol{y}}(t_{k+n-1}|t_{k-1}) + \boldsymbol{h}(n-1)\boldsymbol{u}(t_k) \\
\hat{\boldsymbol{y}}(t_{k+n}|t_k) &= -\boldsymbol{A}_n\hat{\boldsymbol{y}}(t_k|t_{k-1}) - ... - \boldsymbol{A}_1\hat{\boldsymbol{y}}(t_{k+n-1}|t_{k-1}) + \boldsymbol{h}(n)\boldsymbol{u}(t_k)
\end{aligned} \tag{2.76}
$$

This set of equations can compactly be written as

$$
\boldsymbol{x}(t_{k+1}) = \boldsymbol{A}\boldsymbol{x}(t_k) + \boldsymbol{B}\boldsymbol{u}(t_k) \tag{2.77}
$$

with $\boldsymbol{x}(t_k)$ defined by (2.75) and $\boldsymbol{A}$ by (2.71). The input matrix $\boldsymbol{B}$ is in its present form defined as

$$
\boldsymbol{B} = \begin{bmatrix} \boldsymbol{h}(1) \\ \boldsymbol{h}(2) \\ . \\ . \\ \boldsymbol{h}(n-1) \\ \boldsymbol{h}(n) \end{bmatrix} \tag{2.78}
$$

but by using (2.63), it can be expressed in terms of the coefficient matrices of the ARMAV model as in (2.72).

The response $\boldsymbol{y}(t_k)$ can be extracted by using (2.67) as

$$
\begin{aligned}
\boldsymbol{y}(t_k) &= \boldsymbol{y}(t_k|t_{k-1}) + \boldsymbol{B}_0\boldsymbol{u}(t_k) \\
&= \begin{bmatrix}\boldsymbol{I} & \boldsymbol{0} & . & . & \boldsymbol{0} & \boldsymbol{0}\end{bmatrix}\boldsymbol{x}(t_k) + \boldsymbol{B}_0\boldsymbol{u}(t_k) \\
&= \boldsymbol{C}\boldsymbol{x}(t_k) + \boldsymbol{D}\boldsymbol{u}(t_k)
\end{aligned}
\tag{2.79}
$$

which defines the observation matrix $\boldsymbol{C}$ and the direct term matrix $\boldsymbol{D}$. Equation (2.77) corresponds to the first equation in (2.70), and (2.79) to the second equation. □

In a similar fashion it is possible to convert the ARMAV model (2.52) that accounts for the presence of disturbance to an observability canonical state space realization. This is shown in the following theorem which will be given without proof.

Theorem 2.5 - A State Space Realization of the ARMAV Model - With Noise Modelling

A state space realization of the p-variate ARMAV(n,n) model (2.52) is given by

$$
\begin{aligned}
\hat{\boldsymbol{x}}(t_{k+1}|t_k) &= \boldsymbol{A}\hat{\boldsymbol{x}}(t_k|t_{k-1}) + \boldsymbol{K}\boldsymbol{e}(t_k) \ , \quad \boldsymbol{e}(t_k) \in NID(\boldsymbol{0},\boldsymbol{\Lambda}) \\
\boldsymbol{y}(t_k) &= \boldsymbol{C}\hat{\boldsymbol{x}}(t_k|t_{k-1}) + \boldsymbol{e}(t_k)
\end{aligned}
\tag{2.80}
$$

with {$\boldsymbol{A}$, $\boldsymbol{K}$, $\boldsymbol{C}$} defined as

$$
\boldsymbol{A} = \begin{bmatrix}
\boldsymbol{0} & \boldsymbol{I} & . & . & \boldsymbol{0} & \boldsymbol{0} \\
\boldsymbol{0} & \boldsymbol{0} & . & . & \boldsymbol{0} & \boldsymbol{0} \\
. & . & . & . & . & . \\
. & . & . & . & . & . \\
\boldsymbol{0} & \boldsymbol{0} & . & . & \boldsymbol{0} & \boldsymbol{I} \\
-\boldsymbol{A}_n & -\boldsymbol{A}_{n-1} & . & . & -\boldsymbol{A}_2 & -\boldsymbol{A}_1
\end{bmatrix}
\tag{2.81}
$$

$$
\boldsymbol{K} = \begin{bmatrix}
\boldsymbol{I} & \boldsymbol{0} & . & . & \boldsymbol{0} & \boldsymbol{0} \\
\boldsymbol{A}_1 & \boldsymbol{I} & . & . & \boldsymbol{0} & \boldsymbol{0} \\
. & . & . & . & . & . \\
. & . & . & . & . & . \\
\boldsymbol{A}_{n-1} & \boldsymbol{A}_{n-2} & . & . & \boldsymbol{A}_1 & \boldsymbol{I}
\end{bmatrix}^{-1}
\begin{bmatrix}
\boldsymbol{C}_1 - \boldsymbol{A}_1 \\
\boldsymbol{C}_2 - \boldsymbol{A}_2 \\
. \\
. \\
\boldsymbol{C}_n - \boldsymbol{A}_n
\end{bmatrix}
\tag{2.82}
$$

$$C = \begin{bmatrix} I & 0 & . & . & 0 & 0 \end{bmatrix} \tag{2.83}$$

and the state vector $\hat{x}(t_k|t_{k-1})$ defined as

$$\hat{x}(t_k|t_{k-1}) = \begin{bmatrix} \hat{y}(t_k|t_{k-1}) \\ \hat{y}(t_{k+1}|t_{k-1}) \\ . \\ . \\ \hat{y}(t_{k+n-2}|t_{k-1}) \\ \hat{y}(t_{k+n-1}|t_{k-1}) \end{bmatrix} \tag{2.84}$$

□

For univariate ARMA models several canonical state space realization exist, see e.g. Kailath [48]. However, for multivariate systems, the natural analogs of the univariate systems do not necessarily yield minimal realizations. This may be a severe drawback, unless one is able to reduce the realization to a minimal one. However, since reliable and applicable model reduction techniques exist, see e.g. Hoen [38], there is no cause for alarm in using the observability canonical state space realization. This particular representation of ARMAV models has been applied by many authors, see e.g. Andersen et al. [6], Aoki [11], Hannan et al. [33], Hoen [38], Ljung [71], Pandit et al. [86] and Prevosto et al. [94].

2.6 Summary

In this chapter two different representations of multivariate linear and time-invariant discrete-time systems, subjected to Gaussian white noise excitation, have been introduced. The two representations are the stochastic state space representation and representation by stochastic ARMAV models. These representations provide an equivalent description of the discrete-time system. It might seem meaningless to spend time on both representations. However, both representations have advantages. The manipulation of a state space representation of a system is very simple, since it only involves manipulations of a few system matrices. On the other hand, the recursive structure of the ARMAV model might reduce the computational effort in some applications. This chapter has concerned noise-free systems as well as systems affected by noise. In both cases, a state space representation and an equivalent ARMAV model have been derived. In the context of system identification based on measured data, it is very important to account for the presence of noise. So in this case, the innovation state space representation or the equivalent ARMAV model that account for the presence of noise should be used. On the other hand, in case of analytical studies of linear and time-invariant systems the noise-free representations would probably be preferred.