will be presented here.

Assume that the underlying system is given by

$$\begin{aligned} x(k+1) &= Ax(k) + Bu(k) + \varepsilon_1(k) \\ y(k) &= Cx(k) + \varepsilon_2(k) \end{aligned} \tag{7.94}$$

where $\begin{bmatrix} \varepsilon_1(k) \\ \varepsilon_2(k) \end{bmatrix}$ is a zero-mean, i.i.d. vector sequence.

The system is assumed to be controllable (from $[u^T \ \varepsilon_1^T]^T$) and observable. In addition, the stochastic part of the system is assumed to be stationary. The objective is to identify from input-output data a state-space model

$$\begin{aligned} \tilde{x}(k+1) &= \tilde{A}\tilde{x}(k) + \tilde{B}u(k) + \tilde{\varepsilon}_1(k) \\ y(k) &= \tilde{C}\tilde{x}(k) + \tilde{\varepsilon}_2(k) \end{aligned} \tag{7.95}$$

that is equal to (7.94) in an *input-output sense*. We will assume for the sake of simplicity that the input sequence $u(k)$ used in the identification is a white noise sequence.

Consider the following optimal multi-step prediction equation (of finite memory):

$$\begin{aligned} \begin{bmatrix} y(k+1) \\ y(k+2) \\ \vdots \\ y(k+\bar{n}) \end{bmatrix} &= L_1 \begin{bmatrix} y(k-\bar{n}+1) \\ y(k-\bar{n}+2) \\ \vdots \\ y(k) \end{bmatrix} + L_2 \begin{bmatrix} u(k-\bar{n}+1) \\ u(k-\bar{n}+2) \\ \vdots \\ u(k) \end{bmatrix} \\ &\quad + L_3 \begin{bmatrix} u(k+1) \\ u(k+2) \\ \vdots \\ u(k+\bar{n}-1) \end{bmatrix} + \begin{bmatrix} e(k+1|k) \\ e(k+2|k) \\ \vdots \\ e(k+\bar{n}|k) \end{bmatrix} \end{aligned} \tag{7.96}$$

$$= \begin{bmatrix} y(k+1|k) \\ y(k+2|k) \\ \vdots \\ y(k+\bar{n}|k) \end{bmatrix} + \begin{bmatrix} e(k+1|k) \\ e(k+2|k) \\ \vdots \\ e(k+\bar{n}|k) \end{bmatrix} \tag{7.97}$$

$\bar{n} > n$ where n is the system order. $y(k+i|k)$ represents the optimal prediction of $y(k+i)$ on the basis of data $y(k-\bar{n}+1), \cdots, y(k)$ and $u(k-\bar{n}+1), \cdots, u(k+\bar{n}-1)$. $e(k+i|k)$ denotes the respective prediction error. $L_1 \in \mathcal{R}^{n_y \cdot \bar{n} \times n_y \cdot \bar{n}}$, $L_2 \in \mathcal{R}^{n_y \cdot \bar{n} \times n_u \cdot \bar{n}}$ and $L_3 \in \mathcal{R}^{n_y \cdot \bar{n} \times n_u \cdot (\bar{n}-1)}$ are functions of system matrices.

The optimal prediction error $e(k+i|k),\ i \leq 1 \leq \bar{n}$ is zero-mean and uncorrelated with $y(k-\bar{n}+1), \cdots, y(k)$ and $u(k-\bar{n}+1), \cdots, u(k+\bar{n}-1)$. Hence, unbiased, consistent esimates of L_1, L_2 and L_3 can be obtained by applying linear least squares identification. L_1, L_2 and L_3 are related to the system matrices and covariance matrices in a complex manner, and extracting the system matrices directly from L_1, L_2 and L_3 would involve a very difficult nonlinear optimization. It also requires a special parameterization of model matrices in order to prevent a loss of identifiability. Clearly, an alternative way to generate the system matrices is desirable.

We can rewrite the optimal predictions in (7.96) in terms of a Kalman filter estimate as follows:

$$\begin{bmatrix} y(k+1|k) \\ y(k+2|k) \\ \vdots \\ y(k+\bar{n}|k) \end{bmatrix} = \begin{bmatrix} C \\ CA \\ \vdots \\ CA^{\bar{n}-1} \end{bmatrix} x(k+1|k) + L_3 \begin{bmatrix} u(k+1) \\ u(k+2) \\ \vdots \\ u(k+\bar{n}-1) \end{bmatrix} \tag{7.98}$$

$x(k+1|k)$ represents an estimate of $x(k+1)$ that is obtained by running a nonsteady-state Kalman filter started with an initial estimate of

$x(k-\bar{n}+1|k-\bar{n})=0$ and initial covariance matrix corresponding to the open-loop, steady-state covariance of x.[1] Comparing (7.98) with (7.96), one can conclude that

$$\begin{bmatrix} C \\ CA \\ \vdots \\ CA^{\bar{n}-1} \end{bmatrix} x(k+1|k) = \begin{bmatrix} L_1 & L_2 \end{bmatrix} \begin{bmatrix} y(k-\bar{n}+1) \\ \vdots \\ y(k) \\ u(k-\bar{n}+1) \\ \vdots \\ u(k) \end{bmatrix} \tag{7.99}$$

Hence, the extended observability matrix and $\begin{bmatrix} L_1 & L_2 \end{bmatrix}$ have the same image space and examining the rank of the latter gives the system order.

In constructing a state-space model from input-output data, there clearly exists some degrees-of-freedom since the basis for the state vector can be chosen arbitrarily without affecting the input-output relation. This means that the extended observability matrix for the identified model (7.95) (denoted as Γ_o from this point on) can be *any* matrix (of dimension $(\bar{n}\cdot n_y)\times n$) that has the same image space as $\begin{bmatrix} L_1 & L_2 \end{bmatrix}$. Let the SVD of $\begin{bmatrix} L_1 & L_2 \end{bmatrix}$ be represented as follows:

$$\begin{bmatrix} L_1 & L_2 \end{bmatrix} = \begin{bmatrix} U_1 & U_2 \end{bmatrix} \begin{bmatrix} \Sigma_1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} V_1^T \\ V_2^T \end{bmatrix} \tag{7.100}$$

We choose $\Gamma_o = U_1\Sigma_1^{1/2}$. This defines the basis for the state vector. Let $\tilde{x}$ denote x written in terms of the above-defined basis. We then express the

[1] This interpretation does not hold in the case of time-correlated input sequence since future inputs can then contribute to the estimation of past outputs. However, a similar interpretation can be developed and the theory extends straightforwardly with some modifications.

system equation in terms of the new basis as follows:

$$\tilde{x}(k+1) = \tilde{A}\tilde{x}(k) + \tilde{B}u(k) + \tilde{\varepsilon}_1(k) \tag{7.101}$$

$$y(k) = \tilde{C}\tilde{x}(k) + \tilde{\varepsilon}_2(k) \tag{7.102}$$

The form of the state-space model that will actually be identified is the following Kalman filter equation for the above system:

$$\tilde{x}(k+2|k+1) = \tilde{A}\tilde{x}(k+1|k) + \tilde{B}u(k+1) + \underbrace{K(k+1)\zeta(k+1)}_{\tilde{\varepsilon}_1(k+1)} \tag{7.103}$$

$$y(k+1) = \tilde{C}\tilde{x}(k+1|k) + \underbrace{\zeta(k+1)}_{\tilde{\varepsilon}_2(k+1)} \tag{7.104}$$

$\tilde{x}(k+1|k)$ and $\tilde{x}(k+2|k+1)$ are two consecutive estimates generated from a nonsteady-state Kalman filter and $K(k+1)$ is the Kalman filter gain. ζ represents the innovation term (note $\zeta(k+1) = y(k+1) - \tilde{x}(k+1|k)$).

Now that the identification problem is well-defined, we discuss the construction of system matrices. In order to identify the system matrices using the relations in (7.103)–(7.104), we need data for the Kalman filter estimates $\tilde{x}(k+1|k)$ and $\tilde{x}(k+2|k+1)$. Let us define $\tilde{x}(k+2|k+1)$ and $\tilde{x}(k+1|k)$ as the estimates from the nonsteady-state Kalman filter for system (7.101), started with the initial estimate of $\tilde{x}(k-\bar{n}+1|k-\bar{n}) = 0$ and initial covariance given by the open-loop, steady-state covariance of $\tilde{x}$. Then, according to (7.99),

$$, \; {}_o\tilde{x}(k+1|k) = \begin{bmatrix} L_1 & L_2 \end{bmatrix} \begin{bmatrix} y(k-\bar{n}+1) \\ \vdots \\ y(k) \\ u(k-\bar{n}+1) \\ \vdots \\ u(k) \end{bmatrix} \tag{7.105}$$

Hence, the data for $\tilde{x}(k+1|k)$ can be found through the following formula:

$$\tilde{x}(k+1|k) = ,_o^{\dagger} \begin{bmatrix} L_1 & L_2 \end{bmatrix} \begin{bmatrix} y(k-\bar{n}+1) \\ \vdots \\ y(k) \\ u(k-\bar{n}+1) \\ \vdots \\ u(k-1) \end{bmatrix} \tag{7.106}$$

It is important to recognize that the data for $\tilde{x}(k+2|k+1)$ cannot be obtained by time-shifting the data for $\tilde{x}(k+1|k)$, since this will result in the Kalman filter estimate for $\tilde{x}(k+2)$ with a different starting estimate of $x(k-\bar{n}+2|k-\bar{n}+1) = 0$. Instead, one must start from the prediction equation below and follow the same procedure as before:

$$\begin{bmatrix} y(k+2|k+1) \\ y(k+3|k+1) \\ \vdots \\ y(k+\bar{n}|k+1) \end{bmatrix} = \hat{L}_1 \begin{bmatrix} y(k-\bar{n}+1) \\ y(k-\bar{n}+2) \\ \vdots \\ y(k+1) \end{bmatrix} + \hat{L}_2 \begin{bmatrix} u(k-\bar{n}+1) \\ u(k-\bar{n}+2) \\ \vdots \\ u(k+1) \end{bmatrix} + \hat{L}_3 \begin{bmatrix} u(k+2) \\ u(k+3) \\ \vdots \\ u(k+\bar{n}-1) \end{bmatrix} \tag{7.107}$$

Once the data for
$\begin{bmatrix} y^T(k+2|k+1) & y^T(k+3|k+1) & \cdots & y^T(k+\bar{n}|k+1) \end{bmatrix}^T$ are obtained by using the estimates, the data for $\tilde{x}(k+2|k+1)$ can be derived by multiplying them with the pseduo-inverse of $\hat{,}_o$ (which is $,_o$ with the last n_y rows eliminated).

Once the data for $\tilde{x}(k+1|k)$ and $\tilde{x}(k+2|k+1)$ are generated, one can find the system matrices by applying least squares identification to (7.103). Since $\zeta(k+1)$ is a zero-mean sequence that is independent of $\tilde{x}(k+1|k)$ and $u(k+1)$, the least squares method gives unbiased, consistent estimates of $\tilde{A}, \tilde{B}$ and $\tilde{C}$. The covariance matrix for $[\tilde{\varepsilon}_1^T \ \tilde{\varepsilon}_2^T]^T$ can also be computed from the residual sequence.

7.3.3.1 Properties and Issues

The subspace identification method we just described has the following properties (Comment: see Van Overschee and De Moor REF for proofs and discussions):

- The resulting model is asymptotically unbiased.
- The estimates for the covariance matrices are biased, however, due to the fact that (7.103) is a nonsteady-state Kalman filter. The approximation error diminishes as $\bar{n} \to \infty$.

Strengths of the method are that it requires only numerically stable, noniterative linear algebra operations only and that very little prior knowledge (an upper-bound on the system order) is needed to start up the algorithm. However, there are some drawbacks as well. Although the method yields an asymptotically unbiased model, very little can be said about the model quality obtained with finite data. In practice, one must always work with finite-length data sets. In addition, various nonideal factors like nonlinearity and nonstationarity make the residual sequence $e(k+i|k)$ in (7.96) become correlated with the regression data. Because of these reasons, L_1, L_2 obtained from the least squares identification (which are critical for determining the system order and generating data for the Kalman filter estimates) may have significant variance. Although expected errors in the estimates of these matrices can be quantified, it is difficult to say how these errors affect the final model quality (measured in terms of prediction error, frequency response error, etc.). One implication is that, in general, one needs a large amount of data in order to guarantee much success with these algorithms (which is only natural since these algorithms use very little prior knowledge). Another implication is that the above does not replace the traditional parametric identification, but complements it.

For instance, it has been suggested that the subspace methods be used to provide a starting estimate for the prediction error minimization.

Another related issue is that, becuase of the variance, the SVD of $\begin{bmatrix} L_1 & L_2 \end{bmatrix}$ is likely to show many more nonzero singular values than the intrinsic system order. In order not to overfit the data, one has to limit the system order by eliminating the *negligible* singular values in forming the , $_o$ matrix. In the context of model reduction, this is along the same line as the Hankel norm reduction. An alternative for deciding the system order and the basis for the states is to use the SVD of the matrix $\begin{bmatrix} L_1 & L_2 \end{bmatrix} Y$, where Y is the matrix whose columns contain the data for $\begin{bmatrix} y(k-\bar{n}+1)^T & \cdots & y(k)^T & u(k-\bar{n}+1)^T & \cdots & u(k-1)^T \end{bmatrix}^T$. In this case, the singluar values indicate how much of the output data are explained by different linear modes (in the 2-norm sense). In the context of model reduction, this corresponds to a frequency-weighting with the input spectrum (for the deterministic part). This step of determining the model order and basis is somewhat subjective, but is often critical.

Finally, the requirement that the stochastic part of the system be stationary should not be overlooked. If the system has integrating type disturbances, one can difference the input output data before applying the algorithm. Further low-pass filtering may be necessary not to over-emphasize the high frequency fitting (recall the discussion on the frequency-domain bias distribution).