CHAPTER 14

INTERNAL BALANCING AND MODEL REDUCTION

Topics covered

- Internal Balancing
- Model Reduction via Internal Balancing
- Model Reduction via Schur Decomposition
- Hankel Norm Approximation

14.1 INTRODUCTION

Several practical situations such as the design of large space structures (LSS), control of power systems, and others, give rise to very large-scale control problems. Typically, these come from the discretization of distributed parameter problems and have thousands of states in practice. Enormous computational complexities hinder the computationally feasible solutions of these problems.

As a result, control theorists have always sought ways to construct *reduced-order models* of appropriate dimensions (depending upon the problem to be solved) which can then be used in practice in the design of control systems. This process is known as *model reduction.* The idea of model reduction is to construct a reduced-order model from the original full-order model such that the reduced-order model is close, in some sense, to the full-order model. The closeness is normally measured by the smallness of $\|G(s) - G_{\mathrm{R}}(s)\|$, where $G(s)$ and $G_{\mathrm{R}}(s)$ are, respectively, the transfer function matrices of the original and the reduced-order model. Two norms, $\|\cdot\|_\infty$ norm and the Hankel-norm are considered here. The problem of constructing a reduced-order model such that the Hankel-norm error is minimized is called an **Optimal Hankel-norm approximation problem**. A widely used practice of model reduction is to first find a balanced realization (i.e., a realization with controllability and observability Grammians equal to a diagonal matrix) and then to truncate the balanced realization in an appropriate manner to obtain a reduced-order model. The process is known as **balanced truncation method**. Balanced truncation does not minimize the H_∞ model reduction error, it only gives an upper bound. Balancing of a continuous-time system is discussed in **Section 14.2**, where two algorithms are described. **Algorithm 14.2.1** (Laub 1980; Glover 1984) constructs internal balancing of a stable, controllable and observable system, whereas **Algorithm 14.2.2** (Tombs and Postlethwaite 1987) is designed to extract a balanced realization, if the original system is not minimal. **Internal balancing** of a discrete-time system is described in **Section 14.3**.

In **Section 14.4**, it is shown (**Theorem 14.4.1**) that a reduced-order model constructed by **truncating a balanced realization (Algorithm 14.4.1) remains stable** and the H_∞-norm error is bounded.

A Schur method (**Algorithm 14.4.2**) for model reduction is then described. The Schur method due to Safonov and Chiang (1989) is designed to overcome the numerical difficulties in **Algorithm 14.4.1** due to the possible ill-conditioning of the balancing transforming matrices. In **Theorem 14.4.2**, it is shown that the transfer function matrix of the reduced-order model obtained by the Schur method is the same as that of the model reduction procedure via internal balancing using **Algorithm 14.2.1**. The Schur method, however, has its own computational problem. It requires computation of the product of the controllability and observability Grammians, which might be a source of round-off errors. The method, can be modified by using Cholesky factors of the Grammians which then leads to the **square-root algorithm (Algorithm 14.2.2)**.

The advantages of the Schur and the square-root methods can be combined into a **balancing-free square-root algorithm** (Varga 1991). This algorithm is briefly sketched in **Section 14.4.3**.

Section 14.5 deals with **Hankel-norm approximation.** A state-space characterization of all solutions to optimal Hankel-norm approximation due to Glover (1984) is stated (**Theorem 14.5.2**) and then an algorithm to compute an **optimal Hankel-norm approximation (Algorithm 14.5.1)** is described.

Section 14.6 shows how to obtain a **reduced-order model of an unstable system.**

The **frequency-weighted model reduction** problem due to Enns(1984) is considered in **Section 14.7**. The errors at the high frequencies can sometimes possibly be reduced by using suitable weights on the frequencies.

Finally, in **Section 14.8**, a numerical **comparison of different model reduction procedures** is given.

14.2 INTERNAL BALANCING OF CONTINUOUS-TIME SYSTEMS

Let (A, B, C, D) be an n-th order stable system that is both controllable and observable. Then it is known (Glover 1984) that there exists a transformation such that the transformed controllability and observability Grammians are equal to a diagonal matrix Σ. Such a realization is called a **balanced realization** (or **internally balanced realization**).

Internal balancing of a given realization is a preliminary step to a class of methods for model reduction, called **Balance Truncation Methods**. In this section, we describe two algorithms for internal balancing of a **continuous-time system**. The matrix D of the system (A, B, C, D) remains unchanged during the transformation of the system to a balanced system. **We, therefore, drop the matrix D from our discussions in this chapter.**

14.2.1 Internal Balancing of a Minimal Realization (MR)

Suppose that the n-th order system (A, B, C) is stable and minimal. Thus, it is both controllable and observable. Therefore, the controllability Grammian C_G and the observability Grammian O_G are symmetric and positive definite (**see Chapter 7**) and hence admit the Cholesky factorizations.

Let $C_G = L_c L_c^T$ and $O_G = L_o L_o^T$ be the respective Cholesky factorizations.

Let

$$L_o^T L_c = U \Sigma V^T \tag{14.2.1}$$

be the singular value decomposition (SVD) of $L_o^T L_c$.

Define now

$$T = L_c V \Sigma^{-1/2}, \tag{14.2.2}$$

where $\Sigma^{1/2}$ denotes the square root of Σ.

Then T is nonsingular, and furthermore using the expressions for C_G and Eq. (14.2.2), we see that the transformed controllability Grammian $\tilde{C}_G$ is

$$\tilde{C}_G = T^{-1} C_G T^{-T} = \Sigma^{1/2} V^T L_c^{-1} L_c L_c^T L_c^{-T} V \Sigma^{1/2} = \Sigma.$$

Similarly, using the expression for O_G and the Eqs. (14.2.1) and (14.2.2), we see that the transformed observability Grammian $\tilde{O}_G$ is

$$\begin{aligned}\tilde{O}_G &= T^T O_G T = \Sigma^{-1/2} V^T L_c^T L_o L_o^T L_c V \Sigma^{-1/2} \\ &= \Sigma^{-1/2} V^T V \Sigma U^T U \Sigma V^T V \Sigma^{-1/2} = \Sigma^{1/2} \cdot \Sigma^{1/2} = \Sigma.\end{aligned}$$

Thus, the particular choice of

$$T = L_c V \Sigma^{-1/2} \tag{14.2.3}$$

reduces both the controllability and observability Grammians to the same diagonal matrix Σ. The system $(\tilde{A}, \tilde{B}, \tilde{C})$, where the system matrices are defined by

$$\tilde{A} = T^{-1}AT, \qquad \tilde{B} = T^{-1}B, \qquad \tilde{C} = CT \tag{14.2.4}$$

is then a **balanced realization** of the system (A, B, C). The decreasing positive numbers $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_n$ in $\Sigma = \text{diag}(\sigma_1, \sigma_2, \ldots, \sigma_n)$, are the **Hankel singular values.**

The above discussion leads to the following algorithm for internal balancing.

Algorithm 14.2.1. *An Algorithm for Internal Balancing of a Continuous-Time MR*

Inputs.

A—The $n \times n$ state matrix.

B—The $n \times m$ input matrix.

C—The $r \times n$ output matrix.

Outputs.

T—An $n \times n$ nonsingular balancing transforming matrix.

$\tilde{A}$, $\tilde{B}$, $\tilde{C}$—The matrices of internally balanced realization:

$$\tilde{A} = T^{-1}AT, \qquad \tilde{B} = T^{-1}B, \qquad \tilde{C} = CT.$$

Assumptions. *(A, B) is controllable, (A, C) is observable, and A is stable.*

Result. *$T^{-1}C_G T^{-T} = T^T O_G T = \Sigma$, a diagonal matrix with positive diagonal entries.*

Step 1. Compute the controllability and observability Grammians, C_G and O_G, by solving, respectively, the Lyapunov equations:

$$AC_G + C_G A^T + BB^T = 0, \tag{14.2.5}$$

$$A^T O_G + O_G A + C^T C = 0. \tag{14.2.6}$$

(Note that since A is a stable matrix, the matrices C_G and O_G can be obtained by solving the respective Lyapunov equations above (see **Chapter 7**).)

Step 2. Find the Cholesky factors L_c and L_o of C_G and O_G:

$$C_G = L_c L_c^T \quad \text{and} \quad O_G = L_o L_o^T \tag{14.2.7}$$

Step 3. Find the SVD of the matrix $L_o^T L_c$: $L_o^T L_c = U \Sigma V^T$.

Step 4. Compute $\Sigma^{-1/2} = \text{diag}\left(\frac{1}{\sqrt{\sigma_1}}, \frac{1}{\sqrt{\sigma_2}}, \ldots, \frac{1}{\sqrt{\sigma_n}}\right)$, where $\Sigma = \text{diag}(\sigma_1, \sigma_2, \ldots, \sigma_n)$. (Note that σ_i, $i = 1, 2, \ldots, n$ are positive).

Step 5. Form $T = L_c V \Sigma^{-1/2}$

Step 6. Compute the matrices of the balanced realization:

$$\tilde{A} = T^{-1}AT, \quad \tilde{B} = T^{-1}B, \quad \text{and} \quad \tilde{C} = CT.$$

Remark

- The original method of Laub (1980) consisted in finding the transforming matrix T by diagonalizing the product $L_c^T O_G L_c$ or $L_o^T C_G L_o$, which is symmetric and positive definite. The method described here is mathematically equivalent to Laub's method and is numerically more effective.

Example 14.2.1. Consider finding the balanced realization using Algorithm 14.2.1 of the system (A, B, C) given by:

$$A = \begin{pmatrix} -1 & 2 & 3 \\ 0 & -2 & 1 \\ 0 & 0 & -3 \end{pmatrix}, \qquad B = (1, 1, 1)^T, \qquad C = (1, 1, 1).$$

Step 1. By solving the Lyapunov equation (14.2.5), we obtain

$$C_G = \begin{pmatrix} 3.9250 & 0.9750 & 0.4917 \\ 0.9750 & 0.3667 & 0.2333 \\ 0.4917 & 0.2333 & 0.1667 \end{pmatrix}.$$

Similarly, by solving the Lyapunov equation (14.2.6), we obtain

$$O_G = \begin{pmatrix} 0.5000 & 0.6667 & 0.7917 \\ 0.6667 & 0.9167 & 1.1000 \\ 0.7917 & 1.1000 & 1.3250 \end{pmatrix}.$$

Step 2. The Cholesky factors of C_G and O_G are:

$$L_c = \begin{pmatrix} 1.9812 & 0 & 0 \\ 0.4912 & 0.3528 & 0 \\ 0.2482 & 0.3152 & 0.0757 \end{pmatrix}, \qquad L_o = \begin{pmatrix} 0.7071 & 0 & 0 \\ 0.9428 & 0.1667 & 0 \\ 1.1196 & 0.2667 & 0.0204 \end{pmatrix}.$$

Step 3. From the SVD of $L_o^T L_c$ (using MATLAB function **svd**):

$$[U, \Sigma, V] = \mathbf{svd}(L_o^T L_c),$$

we have

$$\Sigma = \operatorname{diag}\begin{pmatrix} 2.2589 & 0.0917 & 0.0006 \end{pmatrix},$$

$$V = \begin{pmatrix} -0.9507 & 0.3099 & 0.0085 \\ -0.3076 & -0.9398 & -0.1488 \\ -0.0381 & -0.1441 & 0.9888 \end{pmatrix}.$$

Step 4.

$$\Sigma^{1/2} = \operatorname{diag}(1.5030, 0.3028, 0.0248).$$

Step 5. The transforming matrix T is:

$$T = L_c V \Sigma^{-1/2} = \begin{pmatrix} -1.2532 & 2.0277 & 0.6775 \\ -0.3835 & -0.5914 & -1.9487 \\ -0.2234 & -0.7604 & 1.2131 \end{pmatrix}.$$

Step 6. The balanced matrices are:

$$\tilde{A} = T^{-1}AT = \begin{pmatrix} -0.7659 & 0.5801 & -0.0478 \\ -0.5801 & -2.4919 & 0.4253 \\ 0.0478 & 0.4253 & -2.7422 \end{pmatrix}.$$

$$\tilde{B} = T^{-1}B = \begin{pmatrix} -1.8602 \\ -0.6759 \\ 0.0581 \end{pmatrix}, \qquad \tilde{C} = CT = \begin{pmatrix} -1.8602 & 0.6759 & -0.0581 \end{pmatrix}.$$

Verify:

$$T^{-1}C_{\mathrm{G}}T^{-\mathrm{T}} = T^{\mathrm{T}}O_{\mathrm{G}}T = \Sigma = \mathrm{diag}(2.2589, 0.0917, 0.0006).$$

Computational Remarks

- **The explicit computation of the product $L_{\mathrm{o}}^{\mathrm{T}}L_{\mathrm{c}}$ can be a source of round-off errors.** The small singular values might be almost wiped out by the rounding errors in forming the explicit product $L_{\mathrm{o}}^{\mathrm{T}}L_{\mathrm{c}}$. It is suggested that the algorithm of Heath *et al.* (1986), which computes the singular values of a product of two matrices without explicitly forming the product, be used in practical implementation of this algorithm.

MATLAB notes: The MATLAB function in the form:

SYSB = **balreal** (sys)

returns a balanced realization of the system (A, B, C). The use of the function **balreal** in the following format:

[SYSB, G, T, TI] = **balreal** (sys)

returns, in addition to $\tilde{A}$, $\tilde{B}$, $\tilde{C}$ of the balanced system, a vector G containing the diagonal of the Grammian of the balanced realization. The matrix T is the matrix of the similarity transformation that transforms (A, B, C) to $(\tilde{A}, \tilde{B}, \tilde{C})$ and TI is its inverse.

MATCONTROL notes: Algorithm 14.2.1 has been implemented in MATCONTROL function **balsvd**.

14.2.2 Internal Balancing of a Nonminimal Realization

In the previous section we showed how to compute the balanced realization of a stable minimal realization. Now we show how to obtain a balanced realization given a stable **nonminimal continuous-time realization.** The method is due to Tombs and Postlethwaite (1987) and is known as the **square-root method.** The algorithm is based on a partitioning of the SVD of the product $L_{\mathrm{o}}^{\mathrm{T}}L_{\mathrm{c}}$ and the balanced matrices $\tilde{A}$, $\tilde{B}$, $\tilde{C}$ are found by applying two transforming matrices L and Z to the matrices A, B, and C. The matrices L_o and L_c are, respectively, the Cholesky factors of the positive semidefinite observability and controllability matrices O_G and C_G.

The balanced realization in this case is of order $k(k < n)$ in contrast with the previous one where the balanced matrices are of the same orders as of the original model.

Let the SVD of $L_o^T L_c$ be represented as

$$L_o^T L_c = (U_1, U_2)\,\mathrm{diag}(\Sigma_1, \Sigma_2)(V_1, V_2)^T$$

where $\Sigma_1 = \mathrm{diag}(\sigma_1, \ldots, \sigma_k) > 0$, $\Sigma_2 = 0_{n-k \times n-k}$.

The matrices U_1, V_1^T, and Σ_1 are of order $n \times k$, $k \times n$, and $k \times k$, respectively.

Define now

$$L = L_o U_1 \Sigma_1^{-1/2}, \qquad Z = L_c V_1 \Sigma_1^{-1/2}$$

Then it has been shown in Tombs and Postlethwaite (1987) that the realization $(\tilde{A}, \tilde{B}, \tilde{C})$, where the matrices $\tilde{A}$, $\tilde{B}$, and $\tilde{C}$ are defined by $\tilde{A} = L^T A Z$, $\tilde{B} = L^T B$, and $\tilde{C} = CZ$ is balanced, truncated to k states, of the system (A, B, C).

Remark

- Note that no assumption on the controllability of (A, B) or the observability of (A, C) is made.

Algorithm 14.2.2. *The Square-Root Algorithm for Balanced Realization of a Continuous-Time Nonminimal Realization*

Inputs. *The system matrices A, B, and C of a nonminimal realization.*

Outputs. *The transforming matrices L, Z, and the balanced matrices $\tilde{A}$, $\tilde{B}$, and $\tilde{C}$.*

Assumption. *A is stable.*

Step 1. *Compute L_o and L_c, using the LDL^T decomposition of O_G and C_G, respectively. (Note that L_o and L_c may be symmetric positive semidefinite, rather than positive definite.)*

Step 2. *Compute the SVD of $L_o^T L_c$ and partition it in the form:*

$$L_o^T L_c = (U_1, U_2)\,\mathrm{diag}(\Sigma_1, \Sigma_2)(V_1, V_2)^T$$

where $\Sigma_1 = \mathrm{diag}(\sigma_1, \sigma_2, \ldots \sigma_k) > 0$.

Step 3. *Define*

$$L = L_o U_1 \Sigma_1^{-1/2} \quad \textit{and} \quad Z = L_c V_1 \Sigma_1^{-1/2}.$$

Step 4. *Compute the balanced matrices $\tilde{A}$, $\tilde{B}$, and $\tilde{C}$ as:*

$$\tilde{A} = L^T A Z, \quad \tilde{B} = L^T B, \quad \textit{and} \quad \tilde{C} = CZ.$$

Example 14.2.2. Let A, C be the same as in Example 14.2.1, and let $B = (1, 0, 0)^{\mathrm{T}}$. Thus, (A, B) is **not controllable.**

Step 1.

$$L_{\mathrm{o}} = \begin{pmatrix} 0.7071 & 0 & 0 \\ 0.9428 & 0.1667 & 0 \\ 1.1196 & 0.2667 & 0.0204 \end{pmatrix}, \qquad L_{\mathrm{c}} = \mathrm{diag}(0.7071, 0, 0).$$

Step 2. $U_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \qquad V_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \qquad \Sigma_1 = 0.5000, \qquad \kappa = 1.$

Step 3.

$$L = L_{\mathrm{o}} U_1 \Sigma_1^{-1/2} = \begin{pmatrix} 1 \\ 1.3333 \\ 1.5833 \end{pmatrix}, \qquad Z = L_{\mathrm{c}} V_1 \Sigma_1^{-1/2} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}.$$

Step 4. $\tilde{A} = -1$, $\tilde{B} = 1$, $\tilde{C} = 1$.

Thus, $(\tilde{A}, \tilde{B}, \tilde{C})$ is a **balanced realization** of order 1, since the realized system is both controllable and observable. Indeed both the controllability and observability Grammians for this realization are equal to 0.5.

MATCONTROL note: Algorithm 14.2.2 has been implemented in MATCONTROL function **balsqt**.

Numerical difficulties of Algorithm 14.2.1 and 14.2.2: Algorithm 14.2.1 of the last section can be numerically unstable in the case when the matrix T is ill-conditioned.

To see this, we borrow the following simple example from Safonov and Chiang (1989):

$$\text{Let} \left(\begin{array}{c|c} \mathrm{A} & \mathrm{B} \\ \hline \mathrm{C} & 0 \end{array} \right) = \left(\begin{array}{cc|c} -\frac{1}{2} & -\epsilon & \epsilon \\ 0 & -\frac{1}{2} & 1 \\ \hline 1 & \epsilon & 0 \end{array} \right).$$

The transforming matrix T of Algorithm 14.2.1 in this case is given by

$$T = \begin{pmatrix} \sqrt{\frac{1}{\epsilon}} & 0 \\ 0 & \sqrt{\epsilon} \end{pmatrix}.$$

Thus, as ϵ becomes smaller and smaller, T becomes more and more ill-conditioned. Indeed, when $\epsilon \to 0$, $\mathrm{Cond}(T)$ becomes infinite.

In such cases, the model reduction procedure via internal balancing becomes unstable.

Similarly, the **square-root algorithm (Algorithm 14.2.2) can be unstable if the matrices L and T are ill-conditioned.**

14.3 INTERNAL BALANCING OF DISCRETE-TIME SYSTEMS

In this section, we consider internal balancing of the stable discrete-time system:

$$\begin{aligned} x_{k+1} &= Ax_k + Bu_k, \\ y_k &= Cx_k. \end{aligned} \tag{14.3.1}$$

We assume that the system is controllable and observable, and give here a discrete analog of Algorithm 14.2.1.

The **discrete analog of Algorithm 14.2.2** can be similarly developed and is left as an (Exercise 14.11(b)).

The controllability Grammian $C_{\rm G}^{\rm D}$ and the observability Grammian $O_{\rm G}^{\rm D}$, defined by (Chapter 7):

$$C_{\rm G}^{\rm D} = \sum_{i=0}^{\infty} A^i BB^{\rm T}(A^{\rm T})^i$$

and

$$O_{\rm G}^{\rm D} = \sum_{i=0}^{\infty} (A^{\rm T})^i C^{\rm T} CA^i$$

satisfy, in this case, respectively, the discrete Lyapunov equations:

$$AC_{\rm G}^{\rm D}A^{\rm T} - C_{\rm G}^{\rm D} + BB^{\rm T} = 0 \tag{14.3.2}$$

and

$$A^{\rm T}O_{\rm G}^{\rm D}A - O_{\rm G}^{\rm D} + C^{\rm T}C = 0. \tag{14.3.3}$$

It can then be shown that the transforming matrix T defined by

$$T = L_{\rm c}V\Sigma^{-1/2}, \tag{14.3.4}$$

where $L_{\rm c}$, V, and Σ are defined in the same way as in the continuous-time case, will transform the system (14.3.1) to the internally balanced system:

$$\begin{aligned} \tilde{x}_{k+1} &= \tilde{A}\tilde{x}_k + \tilde{B}u_k, \\ \tilde{y}_k &= \tilde{C}\tilde{x}_k. \end{aligned} \tag{14.3.5}$$

The Grammians again are transformed to the same diagonal matrix Σ, the matrices $\tilde{A}$, $\tilde{B}$, and $\tilde{C}$ are defined in the same way as in the continuous-time case.

Example 14.3.1.

$$A = \begin{pmatrix} 0.0010 & 1 & 1 \\ 0 & 0.1200 & 1 \\ 0 & 0 & -0.1000 \end{pmatrix}, \qquad B = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \qquad C = (1, 1, 1).$$

(Note that the eigenvalues of A have moduli less than 1, so it is discrete-stable.)

Step 1. The discrete controllability and observability Grammians obtained, respectively, by solving (14.3.2) and (14.3.3) are:

$$C_{\mathrm{G}}^{\mathrm{D}} = \begin{pmatrix} 6.0507 & 3.2769 & 0.8101 \\ 3.2769 & 2.2558 & 0.8883 \\ 0.8101 & 0.8883 & 1.0101 \end{pmatrix},$$

and

$$O_{\mathrm{G}}^{\mathrm{D}} = \begin{pmatrix} 1 & 1.0011 & 1.0019 \\ 1.0011 & 2.2730 & 3.2548 \\ 1.0019 & 3.2548 & 5.4787 \end{pmatrix}.$$

Step 2. The Cholesky factors of C_{G}^{D} and $O_{\mathrm{G}}^{\mathrm{D}}$ are:

$$L_{\mathrm{c}}^{D} = \begin{pmatrix} 2.4598 & 0 & 0 \\ 1.3322 & 0.6936 & 0 \\ 0.3293 & 0.6482 & 0.6939 \end{pmatrix},$$

and

$$L_{\mathrm{o}}^{\mathrm{D}} = \begin{pmatrix} 1 & 0 & 0 \\ 1.0011 & 1.1273 & 0 \\ 1.0019 & 1.9975 & 0.6963 \end{pmatrix}$$

Step 3. The SVD of $(L_{\mathrm{o}}^{\mathrm{D}})^{\mathrm{T}} L_{\mathrm{c}}^{\mathrm{D}}$:

$$[U, \Sigma, V] = \mathbf{svd}(L_{\mathrm{o}}^{\mathrm{D}})^{\mathrm{T}} L_{\mathrm{c}}^{\mathrm{D}}$$

gives

$$\Sigma = \mathrm{diag}(\ 5.3574, \quad 1.4007, \quad 0.1238\ ),$$

$$V = \begin{pmatrix} 0.8598 & -0.5055 & 0.0725 \\ 0.4368 & 0.6545 & -0.6171 \\ 0.2645 & 0.5623 & 0.7835 \end{pmatrix}.$$

Step 4.

$$\Sigma^{1/2} = \mathrm{diag}(2.3146, 1.1835, 0.3519).$$

Step 5. The transforming matrix T is:

$$T = L_{\mathrm{c}}^{\mathrm{D}} V \,\Sigma^{-1/2},$$

$$= \begin{pmatrix} 0.9137 & -1.0506 & 0.5068 \\ 0.6257 & -0.1854 & -0.9419 \\ 0.3240 & 0.5475 & 0.4759 \end{pmatrix}.$$

Step 6. The balanced matrices are:

$$\tilde{A} = T^{-1} A T = \begin{pmatrix} 0.5549 & 0.4098 & 0.0257 \\ -0.4098 & -0.1140 & 0.2629 \\ 0.0257 & -0.2629 & -0.4199 \end{pmatrix},$$

$$\tilde{B} = T^{-1}B = \begin{pmatrix} 1.8634 \\ 0.6885 \\ 0.0408 \end{pmatrix},$$

$$\tilde{C} = CT = (1.8634, \ -0.6885, \ 0.0408).$$

Verify:

$$T^{-1}C_{\mathrm{G}}^{\mathrm{D}}T^{-\mathrm{T}} = T^{\mathrm{T}}O_{\mathrm{G}}^{\mathrm{D}}T = \Sigma = \mathrm{diag}(5.3574, \ 1.4007, \ 0.1238).$$

14.4 MODEL REDUCTION

Given an nth order realization (A, B, C) with the transfer function matrix $G(\lambda) = C(\lambda I - A)^{-1}B$, where "$\lambda$" is complex variable "$s$" for the continuous-time case and is the complex variable $z = (1 + s)/(1 - s)$ in the discrete time, the ideal **model reduction problem** aims at finding a state-space system of order $q < n$ such that the H_∞ error-norm

$$E = \|G(\lambda) - G_{\mathrm{R}}(\lambda)\|_\infty$$

is minimized over all state-space systems of order q, where $G_{\mathrm{R}}(\lambda)$ is the transfer function of the reduced-order model.

The exact minimization is a difficult computational task, and, in practice, a less strict requirement, such a guaranteed upper bound on E is sought to be achieved. We will discuss two such methods in this chapter:

- **Balanced Truncation Method**
- **The Schur Method**

We shall also describe briefly an **optimal Hankel-Norm Approximation** (HNA) method in **Section 14.5**. This optimal HNA method minimizes the error in Hankel norm (defined in **Section 14.5**). Furthermore, we will state another model reduction approach, called **Singular Perturbation** (SP) **Method** in **Exercise 14.23**. For properties of SP method, see Anderson and Liu (1989). Finally, an idea of **Frequency-Weighted Model Reduction** due to Enns (1984) will be discussed in **Section 14.7**.

14.4.1 Model Reduction via Balanced Truncation

As the title suggests, the idea behind model reduction via balanced truncation is to obtain a reduced-order model by deleting those states that are least controllable and observable (as measured by the size of **Hankel singular values**). Thus, if $\Sigma_{\mathrm{R}} = \mathrm{diag}(\sigma_1 I_{s_1}, \ldots, \sigma_N I_{s_N})$, is the matrix of Hankel singular values (which are arranged in decreasing order), obtained by a balanced realization, where s_i is the multiplicity of σ_i, and $\sigma_d \gg \sigma_{d+1}$, then the balanced realization implies that the states corresponding to the Hankel singular values $\sigma_{d+1} \ldots, \sigma_N$ are less controllable and less observable than those corresponding to $\sigma_1, \ldots, \sigma_d$. Thus, the reduced-order model obtained by eliminating these less controllable and less observable states are likely to retain some desirable information about the original system. Indeed, to this end, the following result (**Theorem 14.4.1**) holds. The idea of obtaining such a reduced-order model is due to Moore (1981). Part (a) of the theorem is due to Pernebo and Silverman (1982), and Part (b) was proved by Enns (1984) and independently by Glover (1984). We shall discuss here only the continuous-time case; the discrete-time case in analogous.

Theorem 14.4.1. *Stability and Error Bound of the Truncated Subsystem. Let*

$$G(s) = \left[\begin{array}{cc|c} A_\mathrm{R} & A_{12} & B_\mathrm{R} \\ A_{21} & A_{22} & B_2 \\ \hline C_\mathrm{R} & C_2 & 0 \end{array}\right] \tag{14.4.1}$$

be the transfer function matrix of an nth order internally balanced stable system with the Grammian $\Sigma = \mathrm{diag}(\Sigma_\mathrm{R}, \Sigma_2)$, *where*

$$\begin{aligned} \Sigma_\mathrm{R} &= \mathrm{diag}(\sigma_1 I_{s_1}, \ldots, \sigma_d I_{s_d}), \quad d < N \\ \Sigma_2 &= \mathrm{diag}(\sigma_{d+1} I_{s_{d+1}}, \ldots, \sigma_N I_{s_N}) \end{aligned} \tag{14.4.2}$$

and

$$\sigma_1 > \sigma_2 > \cdots > \sigma_d > \sigma_{d+1} > \sigma_{d+2} > \cdots > \sigma_N.$$

The multiplicity of σ_i is s_i, $i = 1, 2, \ldots, N$ and $s_1 + s_2 + \cdots + s_N = n$.

(a) Then the truncated system $(A_\mathrm{R}, B_\mathrm{R}, C_\mathrm{R})$ with the transfer function:

$$G_\mathrm{R}(s) = \left[\begin{array}{c|c} A_\mathrm{R} & B_\mathrm{R} \\ \hline C_\mathrm{R} & 0 \end{array}\right] \tag{14.4.3}$$

is balanced and stable.

(b) Furthermore, the error:

$$\|G(s) - G_\mathrm{R}(s)\|_\infty \le 2(\sigma_{d+1} + \cdots + \sigma_N). \tag{14.4.4}$$

In particular, if $d = N - 1$, then $\|G(s) - G_{N-1}(s)\|_\infty = 2\sigma_N$.

Proof. We leave the proof part (a) as an exercise **(Exercise 14.1)**. We assume that the **singular values σ_i are distinct** and prove part (b) only in the case $n = N$. The proof in the general case can be easily done using the proof of this special case and is also left as an exercise **(Exercise 14.1)**. The proof has been taken from Zhou *et al.* (1996, pp. 158–160).

Let

$$\phi(s) = (sI - A_{\mathrm{R}})^{-1}, \tag{14.4.5}$$

$$\psi(s) = sI - A_{22} - A_{21}\phi(s)A_{12}, \tag{14.4.6}$$

$$\bar{B}(s) = A_{21}\phi(s)B_{\mathrm{R}} + B_2, \tag{14.4.7}$$

$$\bar{C}(s) = C_{\mathrm{R}}\phi(s)A_{12} + C_2. \tag{14.4.8}$$

Then,

$$\begin{aligned} G(s) - G_{\mathrm{R}}(s) &= C(sI - A)^{-1}B - C_{\mathrm{R}}\phi(s)B_{\mathrm{R}}, \\ &= (C_{\mathrm{R}}, C_2)\begin{pmatrix} sI - A_{\mathrm{R}} & -A_{12} \\ -A_{21} & sI - A_{22} \end{pmatrix}^{-1}\begin{pmatrix} B_{\mathrm{R}} \\ B_2 \end{pmatrix} - C_{\mathrm{R}}\phi(s)B_{\mathrm{R}}, \\ &= \bar{C}(s)\psi^{-1}(s)\bar{B}(s). \end{aligned} \tag{14.4.9}$$

For $s = j\omega$, we have

$$\sigma_{\max}[G(j\omega) - G_{\mathrm{R}}(j\omega)] = \lambda_{\max}^{1/2}[\psi^{-1}(j\omega)\bar{B}(j\omega)\bar{B}^*(j\omega)\psi^{-*}(j\omega)\bar{C}^*(j\omega)\bar{C}(j\omega)], \tag{14.4.10}$$

where $\lambda_{\max}(M)$ denotes the largest eigenvalue of the matrix M.

Since the singular values are distinct, we have $\Sigma_2 = \mathrm{diag}(\sigma_{r+1}, \ldots, \sigma_n)$, and since Σ_2 satisfies

$$A_{22}\Sigma_2 + \Sigma_2 A_{22}^{\mathrm{T}} + B_2 B_2^{\mathrm{T}} = 0,$$

we obtain

$$\bar{B}(j\omega)\bar{B}^*(j\omega) = \psi(j\omega)\Sigma_2 + \Sigma_2\psi^*(j\omega).$$

Similarly, since Σ_2 also satisfies

$$\Sigma_2 A_{22} + A_{22}^{\mathrm{T}}\Sigma_2 + C_2^{\mathrm{T}}C_2 = 0,$$

we obtain

$$\bar{C}^*(j\omega)\bar{C}(j\omega) = \Sigma_2\psi(j\omega) + \psi^*(j\omega)\Sigma_2.$$

Substituting these expressions of $\bar{B}(j\omega)\bar{B}^*(j\omega)$ and $\bar{C}^*(j\omega)\bar{C}(j\omega)$ into (14.4.10), we obtain after some algebraic manipulations

$$\begin{aligned} \sigma_{\max}[G(j\omega) - G_{\mathrm{R}}(j\omega)] = \lambda_{\max}^{1/2}\{&[\Sigma_2 + \psi^{-1}(j\omega)\Sigma_2\psi^*(j\omega)] \\ &\times [\Sigma_2 + \psi^{-*}(j\omega)\Sigma_2\psi(j\omega)]\}. \end{aligned}$$

If $d = n - 1$, then $\Sigma_2 = \sigma_n$, and we immediately have

$$\sigma_{\max}[G(j\omega) - G_{\mathrm{R}}(j\omega)] = \sigma_n\lambda_{\max}^{1/2}\{[1 + \Theta^{-1}(j\omega)][1 + \Theta(j\omega)]\}$$

where $\Theta = \psi^{-*}(j\omega)\psi(j\omega)$.

Note that $\Theta^{-*} = \Theta$ is a scalar function. So, $|\Theta(j\omega)| = 1$.
Using the triangle inequality, we then have

$$\sigma_{\max}[G(j\omega) - G_{\mathrm{R}}(j\omega)] \leq \sigma_n[1 + |\Theta(j\omega)|] = 2\sigma_n$$

Thus, we have proved the result for one-step, that is, we have proved the result assuming that the order of the truncated model is just one less than the original model. Using this "one-step" result, Theorem 14.4.1 can be proved for any order of the truncated model (**Exercise 14.1**). ■

The above theorem shows that once a system is internally balanced, the balanced system can be truncated by eliminating the states corresponding to the less controllable and observable modes (as measured by the sizes of the Hankel singular values) to obtain a reduced-order model that still preserves certain desirable properties of the original model (see Zhou *et al.* 1996). However, the reduced-order model obtained this way does not minimize the H_∞ error.

Choosing the order of the Reduced Model

If the reduced-order model is obtained by truncating the states corresponding to the smaller Hankel singular values $\sigma_{d+1}, \ldots, \sigma_N$, then the order q of the reduced-order model is

$$q = \sum_{i=1}^{d} s_{\mathrm{i}},$$

where s_i is the multiplicity of σ_{i}.

Computationally, the decision on choosing which Heinkel singular values are to be ignored, has to be made in a judicious way so that the matrix $\sum$ which needs to be inverted to compute the balanced matrices does not become too ill-conditioned. Thus, the ratios of the largest Hankel singular value to the consecutive ones need to be monitored. See discussion in Tombs and Postlethwaite (1987).

Algorithm 14.4.1. *Model Reduction via Balanced Truncation*
Inputs. *The system matrices A, B, and C.*
Outputs. *The reduced-order model with the matrices A_{R}, B_{R}, and C_{R}.*
Assumption. *A is stable.*
Step 1. *Find a balanced realization.*
Step 2. *Choose q, the order of model reduction.*
Step 3. *Partition the balanced realization in the form*

$$A = \begin{pmatrix} A_{\mathrm{R}} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}, \qquad B = \begin{pmatrix} B_{\mathrm{R}} \\ B_2 \end{pmatrix}, \qquad C = (C_{\mathrm{R}}, C_2),$$

where A_{R} is of order q, and B_{R} and C_{R} are similarly defined.

The MATLAB function **modred** in the form

$$\mathrm{RSYS} = \mathrm{modred(SYS, ELIM)}$$

reduces the order of the model sys, by eliminating the states specified in the vector ELIM.

Example 14.4.1. Consider Example 14.2.1 once more. Choose $q = 2$. Then $A_\mathrm{R} =$ The 2×2 leading principal submatrix of $\tilde{A}$ is

$$\begin{pmatrix} -0.7659 & 0.5801 \\ -0.5801 & -2.4919 \end{pmatrix}.$$

The eigenvalues of A_R are: -0.9900 and -2.2678.

Therefore, A_R is stable.

The matrices B_R and C_R are:

$$B_\mathrm{R} = \begin{pmatrix} -1.8602 \\ -0.6759 \end{pmatrix}, \qquad C_\mathrm{R} = (-1.8602, 0.6759).$$

Let $G_\mathrm{R}(s) = C_\mathrm{R}(sI - A_\mathrm{R})^{-1}B_\mathrm{R}$.

Verification of the Error Bound: $\|G(s) - G_\mathrm{R}(s)\|_\infty = 0.0012$. Since $2\sigma_3 = 0.0012$, the error bound given by (14.4.4) is satisfied.

14.4.2 The Schur Method for Model Reduction

The numerical difficulties of model reduction via balanced truncation using Algorithm 14.2.1 or Algorithm 14.2.2 (because of possible ill-conditioning of the transforming matrices) can be overcome if orthogonal matrices are used to transform the system to another **equivalent system** from which the reduced-order model is extracted. Safonov and Chiang (1989) have proposed a Schur method for this purpose.

A key observation here is that in Algorithm 14.2.1, *the rows* $\{1, \ldots, d\}$ *and rows* $\{d+1, \ldots, n\}$ *of* T^{-1} *form bases for the left eigenspaces of the matrix* $C_\mathrm{G}O_\mathrm{G}$ *associated with the eigenvalues* $\{\sigma_1^2, \ldots, \sigma_d^2\}$ *and* $\{\sigma_{d+1}^2, \ldots, \sigma_n^2\}$, *respectively* (**Exercise 14.7**).

Thus the idea will be to replace the matrices T and T^{-1} (which can be very ill-conditioned) by the orthogonal matrices (which are perfectly conditioned) sharing the same properties.

The Schur method, described below, constructs such matrices, using the RSF of the matrix $C_\mathrm{G}O_\mathrm{G}$.

Specifically, the orthonormal bases for the right and left invariant subspaces corresponding to the "**large**" eigenvalues of the matrix $C_\mathrm{G}O_\mathrm{G}$ will be computed by finding the **ordered** real Schur form of $C_\mathrm{G}O_\mathrm{G}$.

Once the "large" eigenvalues are isolated from the "small" ones, the reduced order model preserving the desired properties, can be easily extracted.

We will now state the algorithm.

Algorithm 14.4.2. *The Schur Method for Model Reduction (Continuous-time System).*

Inputs.

A—The $n \times n$ state matrix.

B—The $n \times m$ control matrix.

C—The $r \times n$ output matrix.

q—The dimension of the desired reduced-order model.

Outputs.

$\hat{A}_{\mathrm{R}}$—The $q \times q$ reduced state matrix

$\hat{B}_{\mathrm{R}}$—The $q \times m$ reduced control matrix

$\hat{C}_{\mathrm{R}}$—The $r \times q$ reduced output matrix.

S_1 and S_2—Orthogonal transforming matrices such that $\hat{A}_{\mathrm{R}} = S_1^{\mathrm{T}} A S_2$, $\hat{B}_{\mathrm{R}} = S_1^{\mathrm{T}} B$, and $\hat{C}_{\mathrm{R}} = C S_2$

Assumption. *A is stable.*

Step 1. *Compute the controllability Grammian C_{G} and the observability Grammian O_{G} by solving, respectively, the Lyapunov equations (14.2.5) and (14.2.6).*

Step 2. *Transform the matrix $C_{\mathrm{G}} O_{\mathrm{G}}$ to the RSF Y, that is, find an orthogonal matrix X such that $X^{\mathrm{T}} C_{\mathrm{G}} O_{\mathrm{G}} X = Y$.*

Note: The matrix $C_{\mathrm{G}} O_{\mathrm{G}}$ does not have any complex eigenvalues. Thus, Y is actually an upper triangular matrix. Furthermore, the real eigenvalues are nonnegative.

Step 3. *Reorder the eigenvalues of Y in ascending and descending order, that is, find orthogonal matrices U and V such that*

$$U^{\mathrm{T}} Y U = U^{\mathrm{T}} X^{\mathrm{T}} C_{\mathrm{G}} O_{\mathrm{G}} X U = U_{\mathrm{S}}^{\mathrm{T}} C_{\mathrm{G}} O_{\mathrm{G}} U_{\mathrm{S}} = \begin{pmatrix} \lambda_1 & & * \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}, \quad (14.4.11)$$

$$V^{\mathrm{T}} Y V = V^{\mathrm{T}} X^{\mathrm{T}} C_{\mathrm{G}} O_{\mathrm{G}} X V = V_{\mathrm{S}}^{\mathrm{T}} C_{\mathrm{G}} O_{\mathrm{G}} V_{\mathrm{S}} = \begin{pmatrix} \lambda_n & & * \\ & \ddots & \\ 0 & & \lambda_1 \end{pmatrix}, \quad (14.4.12)$$

where $\lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_n$.

(Note that $\lambda_n = \sigma_1^2, \lambda_{n-1} = \sigma_2^2, \ldots$, and so on; where $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_n$ are the Hankel singular values.)

Step 4. *Partition the matrices U_S, V_S as follows:*

$$U_S = (U_{1S}, U_{2S}), \qquad V_S = (V_{1S}, V_{2S}).$$

Here U_{1S} contains the first $n-q$ columns of U_S and U_{2S} contains the remaining q columns. On the other hand, V_{1S} contains the first q columns of V_S and V_{2S} contains the remaining $n-q$ columns.

Note: Note that the columns of the V_{1S} and those of the matrix U_{1S} form, respectively, orthonormal bases for the right invariant subspace of $C_G O_G$ associated with the large eigenvalues $\{\sigma_1^2, \ldots, \sigma_q^2\}$ and the small eigenvalues $\{\sigma_{q+1}^2, \ldots, \sigma_n^2\}$. The columns of U_{2S} and V_{2S}, similarly, form orthonormal bases for the left invariant subspace of $C_G O_G$, with the large and the small eigenvalues, respectively.

Step 5. *Find the SVD of $U_{2S}^T V_{1S}$: $Q\Sigma R^T = U_{2S}^T V_{1S}$.*

Step 6. Compute *the transforming matrices:* $S_1 = U_{2S} Q \Sigma^{-1/2}$, $S_2 = V_{1S} R \Sigma^{-1/2}$

Step 7. *Form the reduced-order matrices:*

$$\hat{A}_R = S_1^T A S_2, \qquad \hat{B}_R = S_1^T B, \qquad \text{and } \hat{C}_R = C S_2. \tag{14.4.13}$$

Flop-count: Since the reduction to the Schur form using the QR iteration is an iterative process, an exact count cannot be given. The method just outlined requires approximately $100n^3$ flops.

Properties of the Reduced-Order Model by the Schur Method

The Schur method for model reduction does not give balanced realization. But the essential properties of the original model are preserved in the reduced-order model, as shown in the following theorem.

Theorem 14.4.2. *The transfer function matrix $\hat{G}_R(s) = \hat{C}_R(sI - \hat{A}_R)^{-1}\hat{B}_R$ obtained by the Schur method (Algorithm 14.4.2) is exactly the same as that of the one obtained via balanced truncation (Algorithm 14.4.1). Furthermore, the controllability and observability Grammians of the reduced-order model are, respectively, given by:*

$$\hat{C}_G^R = S_1^T C_G S_1, \qquad \hat{O}_G^R = S_2 O_G S_2.$$

Proof. We prove the first part and leave the second part as an exercise (**Exercise 14.9**).

Let the transforming matrix T of the internal balancing algorithm and its inverse be partitioned as:

$$T = (T_1, T_2), \tag{14.4.14}$$

and

$$T^{-1} = \begin{pmatrix} T_I \\ * \end{pmatrix}, \tag{14.4.15}$$

where T_1 is $n \times q$ and T_I is of order $q \times n$.

Then the transfer function $G_R(s)$ of the reduced-order model obtained by Algorithm 14.2.1 is given by

$$G_R(s) = C_R(sI - A_R)^{-1}B_R = CT_1(sI - T_I A T_1)^{-1}T_I B. \tag{14.4.16}$$

Again, the transfer function $\hat{G}_R(s)$ of the reduced-order model obtained by the Schur algorithm (Algorithm 14.4.2) is given by:

$$\hat{G}_R(s) = \hat{C}_R(sI - \hat{A}_R)^{-1}\hat{B}_R = CS_2(sI - S_1^T A S_2)^{-1}S_1^T B. \tag{14.4.17}$$

The proof now amounts to establishing a relationship between S_1, S_2 and T_1 and T_I.

Let's define

$$V_R = (V_{1S}\ U_{1S}), \qquad \text{and} \qquad V_L = \begin{pmatrix} U_{2S}^T \\ V_{2S}^T \end{pmatrix}. \tag{14.4.18}$$

Then, since the first q and the last $(n - q)$ columns of T^{-1}, V_R, and V_L^{-1} span, respectively, the right eigenspaces associated with $\sigma_1^2, \ldots, \sigma_q^2$ and $\sigma_{q+1}^2, \ldots, \sigma_n^2$, it follows that there exist nonsingular matrices X_1, and X_2, such that

$$V_R = T\begin{pmatrix} X_1 & 0 \\ 0 & X_2 \end{pmatrix} = V_L^{-1}\begin{pmatrix} E_1 & 0 \\ 0 & E_2 \end{pmatrix}. \tag{14.4.19}$$

From (14.4.18) and (14.4.19) we have

$$V_{1S} = T_1 X_1. \tag{14.4.20}$$

Thus,

$$S_2 = V_{1S} R \Sigma^{-1/2} = T_1 X_1 R \Sigma^{-1/2} \text{ (using (14.4.20))}.$$

Similarly,

$$\begin{aligned} S_1^T &= \Sigma^{-1/2} Q^T U_{2S}^T = \Sigma^{1/2} R^T (R\Sigma^{-1} Q^T) U_{2S}^T, \\ &= \Sigma^{1/2} R^T (U_{2S}^T V_{1S})^{-1} U_{2S}^T \\ &= \Sigma^{1/2} R^T (U_{2S}^T T_1 X_1)^{-1} U_{2S}^T \text{ (using (14.4.20))} = \Sigma^{1/2} R^T X_1^{-1} T_I. \end{aligned}$$

Thus,

$$\begin{aligned}\hat{G}_{\mathrm{R}}(s) &= CS_2(sI - S_1^{\mathrm{T}} A S_2)^{-1} S_1^{\mathrm{T}} B \\ &= CT_1 X_1 R \Sigma^{-1/2} \left(sI - \Sigma^{1/2} R^{\mathrm{T}} X_1^{-1} T_l A T_1 X_1 R \Sigma^{-1/2}\right)^{-1} \\ &\quad \times \left(\Sigma^{1/2} R^{\mathrm{T}} X_1^{-1} T_l B\right) \\ &= CT_1(sI - T_l A T_1)^{-1} T_l B = G_{\mathrm{R}}(s). \quad \blacksquare\end{aligned}$$

Note: Since $G_{\mathrm{R}}(s)$ of Theorem 14.4.1 and $\hat{G}_{\mathrm{R}}(s)$ of Theorem 14.4.2 are the same, from Theorem 14.4.1, we conclude that $\hat{A}_{\mathrm{R}}$ is stable and $\|G(s) - \hat{G}_{\mathrm{R}}(s)\|_\infty \leq 2\sum_{i=q+1}^{n} \sigma_{\mathrm{i}}$, where σ_{q+1} through σ_n are the $(q+1)$th through nth entries of the diagonal matrix Σ of the balancing algorithm, that is, σs are the Hankel singular values.

Relation to the square-root method: There could be large round-off errors in explicit computation of the matrix product $C_{\mathrm{G}} O_{\mathrm{G}}$. The formation of the explicit product, however, can be avoided by computing the Cholesky factors L_{c} and L_{o} of the matrices C_{G} and O_{G}, using **Algorithm 8.6.1** described in Chapter 8. This then leads to a **square-root** method for model reduction. We leave the derivation of the modified algorithm to the readers **(Exercise 14.10)**. For details, see Safonov and Chiang (1989).

Example 14.4.2.

$$A = \begin{pmatrix} -1 & 2 & 3 \\ 0 & -2 & 1 \\ 0 & 0 & -3 \end{pmatrix}, \qquad b = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \qquad C = (1, 1, 1), \qquad q = 2.$$

The system (A, B, C) is **stable, controllable**, and **observable.**

Step 1. Solving the Lyapunov equations (14.2.5) and (14.2.6) we obtain

$$C_{\mathrm{G}} = \begin{pmatrix} 3.9250 & 0.9750 & 0.4917 \\ 0.9750 & 0.3667 & 0.2333 \\ 0.4917 & 0.2333 & 0.1667 \end{pmatrix},$$

and

$$O_{\mathrm{G}} = \begin{pmatrix} 0.5000 & 0.6667 & 0.7917 \\ 0.6667 & 0.9167 & 1.1000 \\ 0.7917 & 1.1000 & 1.3250 \end{pmatrix}.$$

Step 2. The Schur decomposition Y and the transforming matrix X obtained using the MATLAB function **schur**:

$$[X, Y] = \mathrm{schur}(C_{\mathrm{G}} O_{\mathrm{G}})$$

are

$$X = \begin{pmatrix} -0.9426 & -0.3249 & 0.0768 \\ -0.2885 & 0.6768 & -0.6773 \\ -0.1680 & 0.6606 & 0.7317 \end{pmatrix},$$

$$Y = \begin{pmatrix} 5.1028 & -5.2629 & -1.0848 \\ 0 & 0.0084 & 0.0027 \\ 0 & 0 & 0 \end{pmatrix}.$$

Step 3. Since the eigenvalues of Y are in decreasing order of magnitude, we take

$$V_S = X.$$

Next, we compute U_S such that the eigenvalues of $U_S^T C_G O_G U_S$ appear in increasing order of magnitude:

$$U_S = \begin{pmatrix} 0.2831 & -0.8579 & 0.4289 \\ -0.8142 & 0.0214 & 0.5802 \\ 0.5069 & 0.5134 & 0.6924 \end{pmatrix},$$

$$U_S^T C_G O_G U_S = \begin{pmatrix} 0 & -0.0026 & 0.8663 \\ 0 & 0.0084 & -5.3035 \\ 0 & 0 & 5.1030 \end{pmatrix}.$$

Step 4. Partitioning U_S and $V_S = X$, we obtain

$$U_{1S} = \begin{pmatrix} 0.2831 \\ -0.8142 \\ 0.5069 \end{pmatrix}, \qquad U_{2S} = \begin{pmatrix} -0.8579 & 0.4289 \\ 0.0214 & 0.5802 \\ 0.5134 & 0.6924 \end{pmatrix}.$$

$$V_{1S} = \begin{pmatrix} -0.9426 & -0.3249 \\ -0.2885 & 0.6768 \\ -0.1680 & 0.6606 \end{pmatrix}, \qquad V_{2S} = \begin{pmatrix} 0.0768 \\ -0.6773 \\ 0.7317 \end{pmatrix}$$

Step 5. The SVD of the product $U_{2S}^T V_{1S}$ is given by:

$$[Q,\ \Sigma, R] = \mathrm{svd}(U_{2S}^T V_{1S})$$

$$Q = \begin{pmatrix} -0.4451 & 0.8955 \\ 0.8955 & 0.4451 \end{pmatrix}, \qquad R = \begin{pmatrix} -0.9348 & 0.3550 \\ 0.3550 & 0.9349 \end{pmatrix} \qquad \text{and}$$

$$\Sigma = \mathrm{diag}(1, 0.9441).$$

Step 6. The transforming matrices are:

$$S_1 = \begin{pmatrix} 0.7659 & -0.5942 \\ 0.5100 & 0.2855 \\ 0.3915 & 0.7903 \end{pmatrix} \quad \text{and} \quad S_2 = \begin{pmatrix} 0.7659 & -0.6570 \\ 0.5099 & 0.5458 \\ 0.3916 & 0.5742 \end{pmatrix}.$$

Step 7. The matrices of the reduced order model are:

$$\hat{A}_{\mathrm{R}} = S_1^{\mathrm{T}} A S_2 = \begin{pmatrix} 0.3139 & 1.7204 \\ -1.9567 & -3.5717 \end{pmatrix}, \quad \hat{B}_{\mathrm{R}} = S_1^{\mathrm{T}} B = \begin{pmatrix} 1.6674 \\ 0.4817 \end{pmatrix}, \quad \text{and}$$

$$\hat{C}_{\mathrm{R}} = C S_2 = (1.6674, 0.4631).$$

Verification of the properties of the reduced-order model: We next verify that the reduced-order model has desirable properties such as stability and the error bound (14.4.4) is satisfied.

1. The eigenvalues of $\hat{A}_{\mathrm{R}}$ are: $\{-0.9900, -2.2678\}$. Thus, $\hat{A}_{\mathrm{R}}$ is **stable. (Note that these eigenvalues are the same as those of A_{R} of order 2 obtained by Algorithm 14.4.1).**
2. The controllability Grammian $\hat{C}_{\mathrm{G}}^{\mathrm{R}}$ of the reduced order model is given by:

$$\hat{C}_{\mathrm{G}}^{\mathrm{R}} = S_1^{\mathrm{T}} C_{\mathrm{G}} S_1 = \begin{pmatrix} 3.5732 & -1.4601 \\ -1.4601 & 0.8324 \end{pmatrix}.$$

The eigenvalues of $\hat{C}_{\mathrm{G}}^{\mathrm{R}}$ are 4.2053, and 0.2003. Thus, $\hat{C}_{\mathrm{G}}^{\mathrm{R}}$ is positive definite.

It is easily verified by solving the Lyapunov equation $\hat{A}_{\mathrm{R}} \hat{C}_{\mathrm{G}}^{\mathrm{R}} + \hat{C}_{\mathrm{G}}^{\mathrm{R}} \hat{A}_{\mathrm{R}}^{\mathrm{T}} = -\hat{B}_{\mathrm{R}} \hat{B}_{\mathrm{R}}^{\mathrm{T}}$, that the $\hat{C}_{\mathrm{G}}^{\mathrm{R}}$ given above is indeed the controllability Grammian of the reduced order model.

Similar results hold for the observability Grammian $\hat{O}_{\mathrm{GR}}$.

Verification of the error bound: $\|G(s) - \hat{G}_{\mathrm{R}}(s)\|_{\infty} = 0.0012$.
Since $2\sigma_3 = 0.0012$, the error bound (14.4.4) is verified.

Comparison of the reduced order models obtained by balanced truncation and the schur method with the original Model: Figure 14.1 compares the errors of the reduced-order models with the theoretical error bound given by (14.4.4). Figure 14.2 compares the step response of the original model with the step responses of the reduced-order models obtained by balanced truncation and the Schur method.

MATCONTROL note: The MATCONTROL function **modreds** implements the Schur Algorithm **(Algorithm 14.4.2)** in the following format:

$$[A_{\mathrm{R}}, B_{\mathrm{R}}, C_{\mathrm{R}}, S, T] = \mathrm{modreds}(A, B, C, d).$$

The matrices A_{R}, B_{R}, C_{R} are the matrices of the reduced-order model of dimension d. The matrices S and T are the transforming matrices.

14.4.3 A Balancing-Free Square-Root Method for Model Reduction

By computing the matrices L and T a little differently than in the square-root method (**Algorithm 14.2.2**), the main advantages of the Schur method and the square-root method can be combined. The idea is due to Varga (1991).

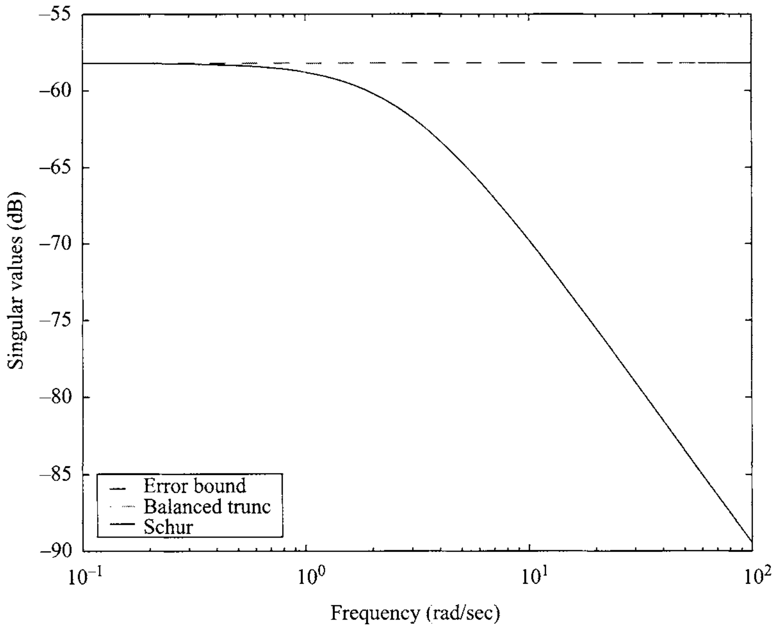

FIGURE 14.1: Theoretical error bound and errors of the reduced-order models.

Consider the **economy** QR factorizations of the matrices $L_c V_1$ and $L_o^T U_1$:

$$L_c V_1 = XW, \qquad L_o^T U_1 = YZ,$$

where W and Z are nonsingular upper triangular and X and Y are orthonormal matrices.

Then L and T defined by

$$L = (Y^T X)^{-1} Y^T, \qquad Z = X$$

are such that the system $(\tilde{A}, \tilde{B}, \tilde{C})$ with $\tilde{A} = LAZ$, $\tilde{B} = LB$, and $\tilde{C} = CZ$ form a **minimal realization** and therefore can be used to obtain a reduced-order model.

Example 14.4.3. Let's consider Example 14.2.2 once more.

Then,

$$X = \begin{pmatrix} -1 \\ 0 \\ 0 \end{pmatrix}, \qquad Y = \begin{pmatrix} -1 \\ 0 \\ 0 \end{pmatrix}, \qquad W = -0.7071, \qquad Z = -0.7071.$$

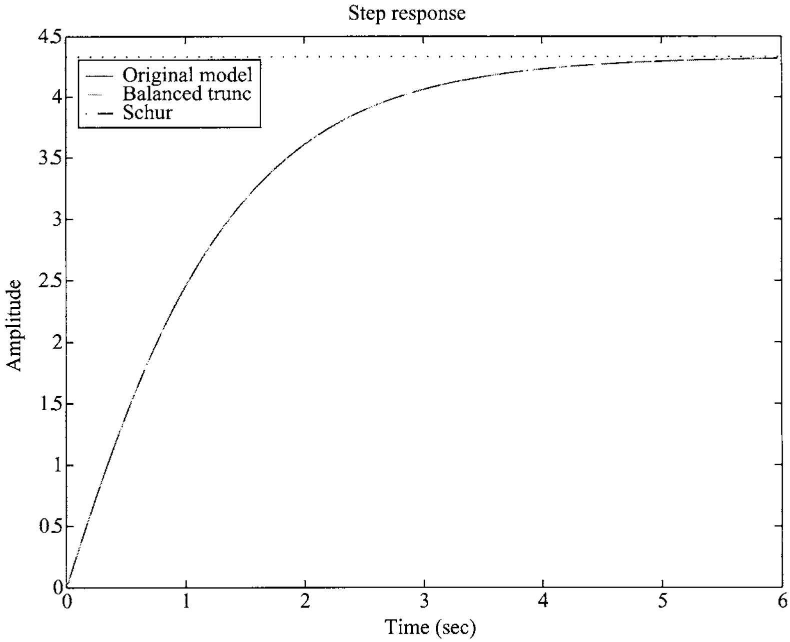

FIGURE 14.2: Step responses of the original and the reduced-order models.

The matrices L and Z in this case are

$$L = (-1 \;\; 0 \;\; 0), \qquad Z_2 = \begin{pmatrix} -1 \\ 0 \\ 0 \end{pmatrix}.$$

The matrices $\tilde{A}$, $\tilde{B}$ and $\tilde{C}$ are: $\tilde{A} = -1$, $\tilde{B} = -1$, $\tilde{C} = -1$.

14.5 HANKEL-NORM APPROXIMATIONS

Let (A, B, C) be a stable realization of $G(s) = C(sI - A)^{-1}B$. Then the **Hankel-norm** of $G(s)$ is defined as

$$\|G(s)\|_H = \lambda_{\max}^{1/2}(C_G O_G), \tag{14.5.1}$$

where C_G and O_G are the controllability and observability Grammians, respectively, and $\lambda_{\max}(M)$ stands for the largest eigenvalue of M.

The **optimal Hankel-norm approximation problem** is the problem of finding an approximation $\hat{G}(s)$ of McMillan degree $k < n$ such that the norm of the error $\|G(s) - \hat{G}(s)\|_H$ is minimized.

The following theorem gives an achievable lower bound of the error of an approximation in Hankel-norm. Proof can be found in Glover (1984) or in Zhou *et al.* (1996, pp. 189–190).

Theorem 14.5.1. *Let $G(s)$ be a stable rational transfer function with Hankel singular values $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_k \geq \sigma_{k+1} \cdots \geq \sigma_n > 0$. Then for all stable $\hat{G}(s)$ of McMillan degree $\leq k$*

$$\| G(s) - \hat{G}(s) \|_{\tilde{H}} \geq \sigma_{k+1}.$$

We now give a result on characterization of all solutions to optimal Hankel-norm approximations and then state an algorithm to find an optimal Hankel-norm approximation.

The presentation here is based on Glover (1984). For proofs and other details, the readers are referred to the paper of Glover or the book by Zhou *et al.* (1996).

14.5.1 A Characterization of All Solutions to the optional Hankel-Norm Approximation

The following theorem gives necessary and sufficient conditions for $\hat{G}(s)$ to be an optimal Hankel-norm approximation to $G(s)$. (For proof, see Glover (1984, lemma 8.1).)

Theorem 14.5.2. *Let $G(s) = C(sI - A)^{-1}B$ be a stable, rational, $m \times m$ transfer function with singular values*

$$\sigma_1 \geq \sigma_2 \geq \sigma_3 \geq \cdots \geq \sigma_k > \sigma_{k+1} = \sigma_{k+2} = \cdots = \sigma_{k+p} > \sigma_{k+p+1} \geq \cdots \geq \sigma_n > 0.$$

Let $\hat{G}(s)$ be of McMillan degree $k \leq n - p$. Then $\hat{G}(s)$ is an optimal Hankel-norm approximation to $G(s)$ if and only if there exists $(\hat{A}, \hat{B}, \hat{C})$, P_e, Q_e such that

(a) *$\hat{G}(s)$ is the stable part of*

$$\hat{C}(sI - \hat{A})^{-1}\hat{B}. \tag{14.5.2}$$

(b) *The matrices P_e and Q_e satisfy*

(*i*) $$A_eP_e + P_eA_e^T + B_eB_e^T = 0, \tag{14.5.3}$$

(*ii*) $$A_e^TQ_e + Q_eA_e + C_e^TC_e = 0 \tag{14.5.4}$$

(*iii*) $$P_eQ_e = \sigma_{k+1}^2 I, \tag{14.5.5}$$

where A_e, B_e *and* C_e *are defined by*

$$A_e = \begin{pmatrix} A & 0 \\ 0 & \hat{A} \end{pmatrix}, \qquad B_e = \begin{pmatrix} B \\ \hat{B} \end{pmatrix}, \qquad C_e = (C, \ -\hat{C}). \quad (14.5.6)$$

and

(c) *If* P_e *and* Q_e *are partitioned conformally with* A_e *in (14.5.6) as:*

$$P_e = \begin{pmatrix} P_{11} & P_{12} \\ P_{12}^T & P_{22} \end{pmatrix}, \qquad Q_e = \begin{pmatrix} Q_{11} & Q_{12} \\ Q_{12}^T & Q_{22} \end{pmatrix},$$

then

$$\mathrm{In}(P_{22}) = \mathrm{In}(Q_{22}) = (k, l, 0). \quad (14.5.7)$$

Further, $\dim(\hat{A}) = k + l$ *can be chosen* $\leq n + 2k - 1$.

We now give a construction of $\hat{A}$, $\hat{B}$, $\hat{C}$ that satisfy the Eqs. (14.5.3)–(14.5.7) for a **balanced realization** (A, B, C) of $G(s)$, which will be the basis of an algorithm for Hankel-norm approximation. The construction, however, remains valid for a more general class of realization, and the details can be found in Glover (1984).

Theorem 14.5.3. *Let* (A, B, C) *be a balanced realization of* $G(s)$ *with the matrix of the singular values*

$$\Sigma = \mathrm{diag}(\sigma_1, \sigma_2, \ldots, \sigma_n),$$

$$\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_k > \sigma_{k+1} = \sigma_{k+2} = \cdots = \sigma_{k+p} > \sigma_{k+p+1} \geq \cdots \geq \sigma_n > 0.$$

Partition $\Sigma = (\Sigma_1, \sigma_{k+1} I_p), \sigma_{k+1} \neq 0$, *and then partition* A, B, C *conformally:*

$$A = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}, \qquad B = \begin{pmatrix} B_1 \\ B_2 \end{pmatrix}, \qquad C = (C_1, C_2). \quad (14.5.8)$$

Define now

$$\hat{A} = \Gamma^{-1}(\sigma_{k+1}^2 A_{11}^T + \Sigma_1 A_{11} \Sigma_1 - \sigma_{k+1} C_1^T U B_1^T), \quad (14.5.9)$$

$$\hat{B} = \Gamma^{-1}(\Sigma_1 B_1 + \sigma_{k+1} C_1^T U), \quad (14.5.10)$$

$$\hat{C} = C_1 \Sigma_1 + \sigma_{k+1} U B_1^T, \quad (14.5.11)$$

where

$$\Gamma = \Sigma_1^2 - \sigma_{k+1}^2 I \quad (14.5.12)$$

and U *is such that*

$$\mathrm{U} = -C_2 \left(B_2^T\right)^{\dagger}, \quad (14.5.13)$$

where '†' denotes generalized inverse.

Then A_e, B_e, and C_e defined by (14.5.6) satisfy (14.5.3)–(14.5.5), with

$$P_e = \begin{pmatrix} \Sigma_1 & 0 & I \\ 0 & \sigma_{k+1} I & 0 \\ I & 0 & \Sigma_1 \Gamma^{-1} \end{pmatrix}, \tag{14.5.14}$$

$$Q_e = \begin{pmatrix} \Sigma_1 & 0 & -\Gamma \\ 0 & \sigma_{k+1} I & 0 \\ -\Gamma & 0 & \Sigma_1 \Gamma \end{pmatrix}. \tag{14.5.15}$$

Based on Theorems 14.5.2 and 14.5.3, the following algorithm can be written down for finding a Hankel-norm approximation of a balanced realization (A, B, C) of $G(s)$.

Algorithm 14.5.1. *An Algorithm for optimal Hankel-Norm Approximation of a Continuous-Time System*

Inputs.

1. *The matrices A, B, and C of a stable realization $G(s)$.*
2. *k—McMillan degree of Hankel-norm approximation*

Outputs. *The matrices $\hat{A}_{11}$, $\hat{B}_1$, and $\hat{C}_1$ of a Hankel-norm approximation $\hat{G}(s)$ of McMillan degree k.*

Assumptions.

1. *A is stable.*
2. *The Hankel singular values σ_i are such that $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_k > \sigma_{k+1} > \sigma_{k+2} \geq \cdots \geq \sigma_n > 0$.*

Step 1. *Find a balanced realization $(\tilde{A}, \tilde{B}, \tilde{C})$ of $G(s)$ using* **Algorithm 14.2.1** *or* **Algorithm 14.2.2,** *whichever is appropriate.*

Step 2. *Partition $\Sigma = \operatorname{diag}(\Sigma_1, \sigma_{k+1})$ and then order the balanced realization $(\tilde{A}, \tilde{B}, \tilde{C})$ conformally so that*

$$\tilde{A} = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}, \qquad \tilde{B} = \begin{pmatrix} B_1 \\ B_2 \end{pmatrix}, \qquad \tilde{C} = (C_1, C_2).$$

(Note that A_{11} is $(n-1) \times (n-1)$).

Step 3. *Compute the matrix U satisfying* (14.5.13) *and form the matrices Γ, $\hat{A}$, $\hat{B}$, and $\hat{C}$ using Eqs.* (14.5.9)–(14.5.12).

Step 4. Block diagonalize *$\hat{A}$ to obtain $\hat{A}_{11}$:*

(a) *Transform $\hat{A}$ to an upper real Schur form and then order the real Schur form so that the eigenvalues with negative real parts appear first; that is, find an orthogonal matrix V_1 such that $V_1^T \hat{A} V_1$ is in upper real Schur* *form and then find another orthogonal matrix V_2 such that*

$$V_2^{\mathrm{T}} V_1^{\mathrm{T}} \hat{A} V_1 V_2 = \begin{pmatrix} \hat{A}_{11} & \hat{A}_{12} \\ 0 & \hat{A}_{22} \end{pmatrix}, \tag{14.5.16}$$

where the eigenvalues of $\hat{A}_{11}$ have negative real parts and those of $\hat{A}_{22}$ have positive real parts. (Note that $\hat{A}_{11}$ is $k \times k$).

(b) *Solve the Sylvester equation for $X \in \mathbb{R}^{k\times(n-k-1)}$ (using* **Algorithm 8.5.1***):*

$$\hat{A}_{11}X - X\hat{A}_{22} + \hat{A}_{12} = 0 \tag{14.5.17}$$

(c) *Let*

$$T = V_1 V_2 \begin{pmatrix} I & X \\ 0 & I \end{pmatrix} = (T_1, T_2), \tag{14.5.18}$$

$$S = \begin{pmatrix} I & -X \\ 0 & I \end{pmatrix} V_2^{\mathrm{T}} V_1^{\mathrm{T}} = \begin{pmatrix} S_1 \\ S_2 \end{pmatrix}. \tag{14.5.19}$$

Step 5. *Form*

$$\hat{B}_1 = S_1 \hat{B}, \tag{14.5.20}$$

$$\hat{C}_1 = \hat{C} T_1. \tag{14.5.21}$$

Example 14.5.1. Consider Example 14.2.1 once more. Then

Step 1. The balanced matrices $\tilde{A}$, $\tilde{B}$, and $\tilde{C}$ are the same as of Example 14.2.1.

Step 2.

$$\Sigma = \mathrm{diag}(2.2589, 0.0917, 0.0006), \qquad \Sigma_1 = \mathrm{diag}(2.2589, 0.0917).$$

$k = 2$ and $\sigma_3 = 0.0006$.

$$A_{11} = \begin{pmatrix} -0.7659 & 0.5801 \\ -0.5801 & -2.4919 \end{pmatrix}, \qquad B_1 = \begin{pmatrix} -1.8602 \\ -0.6759 \end{pmatrix}, \qquad B_2 = (0.0581)$$

$$C_1 = (-1.8602, 0.6759), \qquad C_2 = (-0.0581).$$

Step 3.

$$U = 1, \qquad \Gamma = \mathrm{diag}(5.1026, 0.0084),$$

$$\hat{A} = \begin{pmatrix} -0.7659 & 0.0235 \\ -14.2961 & -2.4919 \end{pmatrix}, \quad \hat{B} = \begin{pmatrix} -0.8235 \\ -7.3735 \end{pmatrix}, \quad \hat{C} = (-4.2020, -0.0620).$$

Step 4.

(a) $\hat{A}$ is already in upper Schur form. Thus, $V_1 = I$.

The eigenvalues of $\hat{A}$ are -0.9900, -2.2678. Since both have negative real parts, no reordering is needed.

Thus, $V_2 = I$, $\hat{A}_{11} = \hat{A}$, $\hat{A}_{12} = 0$, $\hat{A}_{22} = 0$.

(b) $X = 0$.
(c)

$$T = \begin{pmatrix} I & 0 \\ 0 & I \end{pmatrix}, \qquad T_1 = I, \qquad T_2 = I.$$
$$S = \begin{pmatrix} I & 0 \\ 0 & I \end{pmatrix}, \qquad S_1 = I, \qquad S_2 = I.$$

Step 5.

$$\hat{B}_1 = \hat{B}, \qquad \hat{C}_1 = \hat{C}.$$

Obtaining an error bound: Next, we show how to construct a matrix $\hat{D}_1$ such that with

$$\hat{G}(s) = \hat{D}_1 + \hat{C}_1(sI - \hat{A}_{11})^{-1}\hat{B}_1,$$

an error bound for the approximation $\|G(s) - \hat{G}(s)\|_\infty$ can be obtained.

Define

$$\hat{B}_2 = S_2\hat{B}, \qquad \hat{C}_2 = \hat{C}T_2, \qquad \hat{D}_1 = -\sigma_{k+1}U. \tag{14.5.22}$$

Step 6. Update now $\hat{D}_1$ as follows:

6.1. Find a balanced realization of the system $(-\hat{A}_{22}, \hat{B}_2, \hat{C}_2, \hat{D}_1)$, say (A_3, B_3, C_3, D_3). Compute the Hankel singular values of this balanced system and call them $\mu_1, \mu_2, \ldots, \mu_{n-k-1}$.

6.2. Let q be an integer greater than or equal to $r + m$, where r and m are the number of outputs and inputs, respectively. Define $Z, Y \in \mathbb{R}^{q\times(n-k-1)}$ by

$$Z = \begin{pmatrix} B_3^{\mathrm{T}} \\ 0 \end{pmatrix}, \qquad Y = \begin{pmatrix} C_3 \\ 0 \end{pmatrix}.$$

Denote the ith columns of Z and Y by z_i and y_i, respectively.

6.3. For $i = 1, 2, \ldots, n-k-1$ do

(i) Find Householder matrices H_1 and H_2 such that

$$H_1 y_i = -\begin{pmatrix}\alpha & 0 & \cdots & 0\end{pmatrix}^{\mathrm{T}}$$

and

$$H_2 z_i = -\begin{pmatrix}\beta & 0 & \cdots & 0\end{pmatrix}^{\mathrm{T}}.$$

(ii) Define

$$U_i = H_1 \begin{pmatrix} -\alpha/\beta & 0 & 0 & 0 \\ 0 & 0 & I_{r-1} & 0 \\ 0 & I_{m-1} & 0 & 0 \\ 0 & 0 & 0 & I_{q-r-m+1} \end{pmatrix} H_2.$$

(iii) If $i < n-k+1$, then for $j = i+1$ to $(n-k+1)$ do

$$y \equiv -(y_j\mu_j + Uz_j\mu_{\rm i})(\mu_{\rm i}^2 - \mu_j^2)^{-1/2},$$
$$z_j \equiv (z_j\mu_j + U^{\rm T}y_j\mu_{\rm i})(\mu_{\rm i}^2 - \mu_j^2)^{-1/2},$$
$$y_j = y$$

(iv) Compute $\hat{D}_1 = \hat{D}_1 + (-1)^i\mu_{\rm i}\begin{pmatrix} I_r & 0\end{pmatrix} U \begin{pmatrix} I_m \\ 0\end{pmatrix}$

Theorem 14.5.4. *(An Error Bound).* $\|G(s) - \hat{G}(s)\|_\infty \leq \sigma_{k+1} + \mu_1 + \mu_2 + \cdots + \mu_{n-k-1}$.

Example 14.5.2. Consider $k = 2$ and

$$A = \begin{pmatrix} -1 & 2 & -1 & 3 \\ 0 & -2 & 2 & 0 \\ 0 & 0 & -3 & -2 \\ 0 & 0 & 0 & -4 \end{pmatrix}, \quad B = \begin{pmatrix} 1 & -2 \\ 2 & 0 \\ -1 & 5 \\ 2 & 3 \end{pmatrix}, \quad C = \begin{pmatrix} -1 & 0 & 2 & -3 \\ 1 & 1 & -2 & 1 \end{pmatrix}$$

Step 1. The Hankel singular values:

$$\{4.7619 \quad 1.3650 \quad 0.3614 \quad 0.0575\}$$

Step 2.

$$\tilde{A} = \begin{pmatrix} -1.1663 & 1.7891 & -0.2132 & 0.3266 \\ -0.0919 & -2.5711 & 0.7863 & -1.0326 \\ 0.3114 & 2.6349 & -3.7984 & 1.5960 \\ -0.3641 & 0.5281 & -0.2582 & -2.4642 \end{pmatrix},$$

$$\tilde{B} = \begin{pmatrix} 3.3091 & -0.3963 \\ -0.2903 & 2.6334 \\ -0.6094 & -1.5409 \\ 0.4947 & -0.1967 \end{pmatrix}$$

$$\tilde{C} = \begin{pmatrix} -2.4630 & 1.3077 & 0.9011 & 0.1600 \\ 2.2452 & -2.3041 & 1.3905 & -0.5078 \end{pmatrix}$$

The permutation matrix that does the reordering is:

$$P = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}.$$

The reordered balanced realization gives

$$\Sigma_1 = \begin{pmatrix} 4.7619 & 0 & 0 \\ 0 & 1.3650 & 0 \\ 0 & 0 & 0.0575 \end{pmatrix}, \qquad \Sigma_2 = 0.3614 = \sigma_3.$$

$$A_{11} = \begin{pmatrix} -1.1663 & 1.7891 & 0.3266 \\ -0.0919 & -2.5711 & -1.0326 \\ -0.3641 & 0.5281 & -2.4642 \end{pmatrix}, \qquad A_{12} = \begin{pmatrix} -0.2132 \\ 0.7863 \\ -0.2582 \end{pmatrix},$$

$$A_{21} = \begin{pmatrix} 0.3114 & 2.6349 & 1.5960 \end{pmatrix}, \qquad A_{22} = -3.7984.$$

$$B_1 = \begin{pmatrix} 3.3091 & -0.3963 \\ -0.2903 & 2.6334 \\ 0.4947 & -0.1967 \end{pmatrix}, \qquad B_2 = \begin{pmatrix} -0.6094 & -1.5409 \end{pmatrix},$$

$$C_1 = \begin{pmatrix} -2.4630 & 1.3077 & 0.1600 \\ 2.2452 & -2.3041 & -0.5078 \end{pmatrix}, \qquad C_2 = \begin{pmatrix} 0.9011 \\ 1.3905 \end{pmatrix}.$$

Step 3.

$$U = \begin{pmatrix} 0.2000 & 0.5057 \\ 0.3086 & 0.7804 \end{pmatrix}.$$

$$\Gamma = \begin{pmatrix} 22.5447 & 0.0000 & 0.0000 \\ 0.0000 & 1.7326 & 0.0000 \\ 0.0000 & 0.0000 & -0.1273 \end{pmatrix}, \quad \hat{A} = \begin{pmatrix} -1.1872 & 0.4948 & -0.0019 \\ 0.0063 & -2.3615 & 0.0072 \\ 0.3687 & 1.5211 & 2.5932 \end{pmatrix}$$

$$\hat{B} = \begin{pmatrix} -0.7022 & 0.0756 \\ 0.3225 & -1.8376 \\ 0.1306 & 0.9841 \end{pmatrix}, \qquad \hat{C} = \begin{pmatrix} 11.5617 & -2.2454 & 0.0090 \\ -10.9487 & 2.4347 & -0.0295 \end{pmatrix}.$$

Step 4.

$$V_1 = \begin{pmatrix} -0.3754 & -0.9222 & -0.0930 \\ 0.8936 & -0.3335 & -0.3006 \\ -0.2462 & 0.1960 & -0.9492 \end{pmatrix}, \qquad V_2 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

$$\hat{A}_{11} = \begin{pmatrix} -2.3661 & 0.4343 \\ 0 & -1.1847 \end{pmatrix}, \qquad \hat{A}_{12} = \begin{pmatrix} 1.3682 \\ -0.7792 \end{pmatrix}, \qquad \hat{A}_{22} = 2.5953.$$

Solution of (14.5.17) gives

$$X = \begin{pmatrix} 0.2577 \\ -0.2061 \end{pmatrix}$$

and then

$$T_1 = \begin{pmatrix} -0.3754 & -0.9222 \\ 0.8936 & -0.3335 \\ -0.2462 & 0.1960 \end{pmatrix}, \qquad T_2 = \begin{pmatrix} 0.0003 \\ -0.0015 \\ -1.0530 \end{pmatrix},$$

$$S_1 = \begin{pmatrix} -0.3515 & 0.9710 & -0.0015 \\ -0.9413 & -0.3954 & 0.0003 \end{pmatrix},$$

$$S_2 = \begin{pmatrix} -0.0930 & -0.3006 & -0.9492 \end{pmatrix}.$$

Step 5. Using (14.5.20)–(14.5.22), we obtain

$$\hat{B}_1 = \begin{pmatrix} 0.5597 & -1.8124 \\ 0.5335 & 0.6558 \end{pmatrix}, \qquad \hat{C}_1 = \begin{pmatrix} -6.3494 & -9.9113 \\ 6.2934 & 9.2788 \end{pmatrix},$$

$$\hat{B}_2 = \begin{pmatrix} -0.1556 & -0.3889 \end{pmatrix}, \qquad \hat{C}_2 = 10^{-1}\begin{pmatrix} -0.0237 \\ 0.2383 \end{pmatrix},$$

$$\hat{D}_1 = \begin{pmatrix} -0.0723 & -0.1828 \\ -0.1115 & -0.2820 \end{pmatrix}.$$

Step 6.

6.1. The matrices of the balanced realization of the system $(-\hat{A}_{22}, \hat{B}_2, \hat{C}_2, \hat{D}_1)$ are:

$$A_3 = \begin{pmatrix} -2.5953 \end{pmatrix}, \qquad B_3 = 10^{-1}\begin{pmatrix} -0.3720 & -0.9298 \end{pmatrix},$$

$$C_3 = 10^{-1}\begin{pmatrix} -0.0991 \\ 0.9966 \end{pmatrix}, \qquad D_3 = \begin{pmatrix} -0.0723 & -0.1828 \\ -0.1115 & -0.2820 \end{pmatrix}.$$

The system (A_3, B_3, C_3, D_3) has only one Hankel Singular value $\mu_1 = 0.0019$.

6.2. Taking $q = r + m = 4$, we obtain

$$Z = 10^{-1}\begin{pmatrix} -0.3720 \\ -0.9298 \\ 0.0000 \\ 0.0000 \end{pmatrix}, \qquad Y = 10^{-1}\begin{pmatrix} -0.0991 \\ 0.9966 \\ 0.0000 \\ 0.0000 \end{pmatrix}$$

6.3. $i = 1, \alpha = \beta = 0.1001$

$$H_1 = \begin{pmatrix} -0.0989 & 0.9951 & 0.0000 & 0.0000 \\ 0.9951 & 0.0989 & 0.0000 & 0.0000 \\ 0.0000 & 0.0000 & 1. & 0.0000 \\ 0.0000 & 0.0000 & 0.0000 & 1.0000 \end{pmatrix},$$

$$H_2 = \begin{pmatrix} 0.3714 & 0.9285 & 0.0000 & 0.0000 \\ 0.9285 & -0.3714 & 0.0000 & 0.0000 \\ 0.0000 & 0.0000 & -1.0000 & 0.0000 \\ 0.0000 & 0.0000 & 0.0000 & -1.0000 \end{pmatrix}$$

$$U_1 = \begin{pmatrix} 0.0368 & 0.0919 & -0.9951 & 0.0000 \\ -0.3696 & -0.9239 & -0.0989 & 0.0000 \\ 0.9285 & -0.3714 & 0.0000 & 0.0000 \\ 0.0000 & 0.0000 & 0.0000 & -1.0000 \end{pmatrix},$$

$$\hat{D}_1 = \begin{pmatrix} -0.0723 & -0.1829 \\ -0.1108 & -0.2803 \end{pmatrix}.$$

Verification: Let $\hat{G}(s) = \hat{C}_1(sI - \hat{A}_{11})^{-1}\hat{B}_1 + \hat{D}_1$. Then,

$$0.3627 = \|G(s) - \hat{G}(s)\|_\infty \leq \sigma_3 + \mu_1 = 0.3633.$$

Also, $\|G(s) - \hat{G}(s)\|_H = 0.3614 = \sigma_3$.

Figure 14.3 compares the step response of the original model with that of the Hankel-norm approximation model.

MATCONTROL note: Algorithm 14.5.1 has been implemented in Matcontrol function **hnaprx**.

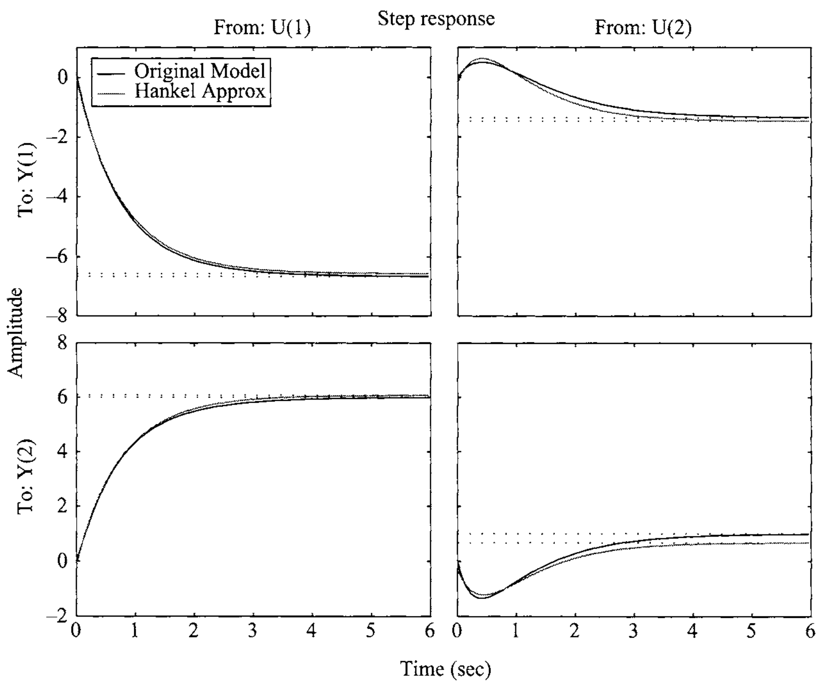

FIGURE 14.3: Step responses of the original and Hankel-norm approximation models.

14.6 MODEL REDUCTION OF AN UNSTABLE SYSTEM

We have so far considered model reduction of a stable system.

However, model reduction of an unstable system can also be performed. Varga (2001) has proposed two approaches. The first approach consists of finding only the reduced-order model of the stable part and then including the unstable part in the resulting reduced model. The second approach is based on computing a stable rational coprime factorization of the transfer function matrix and then reducing the stable system. We describe just the first approach here. For details of the second approach, see Varga (2001).

Step 1. Decompose the transfer function matrix $G(\lambda)$ additively as:

$$G(\lambda) = G_S(\lambda) + G_U(\lambda)$$

such that $G_S(\lambda)$ is the stable part and $G_U(\lambda)$ is the unstable part.

Step 2. Find a reduced-order model $G_{RS}(\lambda)$ of the stable part $G_S(\lambda)$.

Step 3. The reduced-order model $G_R(\lambda)$ of $G(\lambda)$ is then given by

$$G_R(\lambda) = G_{RS}(\lambda) + G_U(\lambda).$$

Computational remarks. The decomposition in Step 1 can be performed by block-diagonalizing the matrix A using the procedure of Step 4 of Algorithm 14.5.1.

14.7 FREQUENCY-WEIGHTED MODEL REDUCTION

In this section, we consider the frequency-weighted model reduction, proposed by Enns (1984). Specifically, the following problem is considered.

Given a stable transfer function matrix $G(s) = C(sI - A)^{-1}B$ and the two input and output weighting transfer function matrices $W_i = C_i(sI - A_i)^{-1}B_i$, and $W_o = C_o(sI - A_o)^{-1}B_o$, find a reduced-order model (A_R, B_R, C_R) with

$$G_R(s) = C_R(sI - A_R)^{-1}B_R$$

such that $\| W_o(G - G_R)W_i \|_\infty$ is minimized and $G(s)$ and $G_R(s)$ have the same number of unstable poles.

The effect of weighting on the model reduction is the possible reduction of the errors at the high frequencies.

The weighting model reduction problem can be solved in a similar way as the model reduction procedure by balanced truncation described in Section 14.4.

First, we note that the state space realization for the weighted transfer matrix is given by

$$W_{\mathrm{o}} G W_{\mathrm{i}} = \bar{C}(sI - \bar{A})^{-1}\bar{B},$$

where

$$\bar{A} = \begin{pmatrix} A & 0 & BC_{\mathrm{i}} \\ B_{\mathrm{o}}C & A_{\mathrm{o}} & 0 \\ 0 & 0 & A_{\mathrm{i}} \end{pmatrix}, \qquad \bar{B} = \begin{pmatrix} 0 \\ 0 \\ B_{\mathrm{i}} \end{pmatrix}, \qquad \bar{C} = (0, C_{\mathrm{o}}, 0). \tag{14.7.1}$$

Let $\hat{C}_{\mathrm{G}}$ and $\hat{O}_{\mathrm{G}}$ be the solutions to the Lyapunov equations:

$$\bar{A}\hat{C}_{\mathrm{G}} + \hat{C}_{\mathrm{G}}(\bar{A})^{\mathrm{T}} + \bar{B}\bar{B}^{\mathrm{T}} = 0,$$

$$\hat{O}_{\mathrm{G}}\bar{A} + (\bar{A})^{\mathrm{T}}\hat{O}_{\mathrm{G}} + (\bar{C})^{\mathrm{T}}\bar{C} = 0.$$

Then the input weighted Grammian $\bar{C}_{\mathrm{G}}$ and the output weighted Grammian $\bar{O}_{\mathrm{G}}$ are defined by

$$\bar{C}_{\mathrm{G}} \equiv (I_n, 0)\hat{C}_{\mathrm{G}}\begin{pmatrix} I_n \\ 0 \end{pmatrix} \quad \text{and} \quad \bar{O}_{\mathrm{G}} \equiv (I_n, 0)\hat{O}_{\mathrm{G}}\begin{pmatrix} I_n \\ 0 \end{pmatrix}.$$

It can be shown (**Exercise 14.21**) that $\bar{C}_{\mathrm{G}}$ and $\bar{O}_{\mathrm{G}}$ satisfy:

$$\begin{aligned} &\begin{pmatrix} A & BC_{\mathrm{i}} \\ 0 & A_{\mathrm{i}} \end{pmatrix}\begin{pmatrix} \bar{C}_{\mathrm{G}} & \bar{C}_{\mathrm{G}_{12}} \\ \bar{C}_{\mathrm{G}_{12}}^{\mathrm{T}} & \bar{C}_{\mathrm{G}_{22}} \end{pmatrix} + \begin{pmatrix} \bar{C}_{\mathrm{G}} & \bar{C}_{\mathrm{G}_{12}} \\ \bar{C}_{\mathrm{G}_{12}}^{\mathrm{T}} & \bar{C}_{\mathrm{G}_{22}} \end{pmatrix}\begin{pmatrix} A^{\mathrm{T}} & 0 \\ C_{\mathrm{i}}^{\mathrm{T}}B^{\mathrm{T}} & A_{\mathrm{i}}^{\mathrm{T}} \end{pmatrix} \\ &\quad + \begin{pmatrix} 0 \\ B_{\mathrm{i}} \end{pmatrix}\begin{pmatrix} 0 & B_{\mathrm{i}}^{\mathrm{T}} \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \end{aligned} \tag{14.7.2}$$

$$\begin{aligned} &\begin{pmatrix} \bar{O}_{\mathrm{G}} & \bar{O}_{\mathrm{G}_{12}} \\ \bar{O}_{\mathrm{G}_{12}}^{\mathrm{T}} & \bar{O}_{\mathrm{G}_{22}} \end{pmatrix}\begin{pmatrix} A & 0 \\ B_{\mathrm{o}}C & A_{\mathrm{o}} \end{pmatrix} + \begin{pmatrix} A^{\mathrm{T}} & C^{\mathrm{T}}B_{\mathrm{o}}^{\mathrm{T}} \\ 0 & A_{\mathrm{o}}^{\mathrm{T}} \end{pmatrix}\begin{pmatrix} \bar{O}_{\mathrm{G}} & \bar{O}_{\mathrm{G}_{12}} \\ \bar{O}_{\mathrm{G}_{12}}^{\mathrm{T}} & \bar{O}_{\mathrm{G}_{22}} \end{pmatrix} \\ &\quad + \begin{pmatrix} 0 \\ C_{\mathrm{o}}^{\mathrm{T}} \end{pmatrix}(0 \;\; C_{\mathrm{o}}) = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}. \end{aligned} \tag{14.7.3}$$

Consider now two special cases.

Case 1. $W_{\mathrm{i}} = I$. Then $\bar{C}_{\mathrm{G}}$ can be obtained from

$$\bar{C}_{\mathrm{G}}A^{\mathrm{T}} + A\bar{C}_{\mathrm{G}} + BB^{\mathrm{T}} = 0.$$

Case 2. $W_{\mathrm{o}} = I$. Then $\bar{O}_{\mathrm{G}}$ can be obtained from

$$\bar{O}_{\mathrm{G}}A + A^{\mathrm{T}}\bar{O}_{\mathrm{G}} + C^{\mathrm{T}}C = 0.$$

Now, let T be a nonsingular matrix such that

$$T\bar{C}_{\mathrm{G}}T^{\mathrm{T}} = (T^{-1})^{\mathrm{T}}\bar{O}_{\mathrm{G}}T^{-1} = \mathrm{diag}(\bar{\Sigma}_1, \bar{\Sigma}_2); \tag{14.7.4}$$

that is, the matrix T makes the realization balanced.

Let $\bar{\Sigma}_1 = \mathrm{diag}(\sigma_1 I_{s_1}, \ldots, \sigma_r I_{s_r})$ and $\bar{\Sigma}_2 = \mathrm{diag}(\sigma_{r+1} I_{s_r+1} \cdots \sigma_n I_{s_n})$.

Partition the system (TAT^{-1}, TB, CT^{-1}) accordingly; that is

$$TAT^{-1} = \begin{pmatrix} \bar{A}_{\mathrm{R}} & \bar{A}_{12} \\ \bar{A}_{21} & \bar{A}_{22} \end{pmatrix}, \qquad TB = \begin{pmatrix} \bar{B}_{\mathrm{R}} \\ \bar{B}_2 \end{pmatrix},$$

and

$$CT^{-1} = (\bar{C}_{\mathrm{R}}, \bar{C}_2).$$

Then $(\bar{A}_{\mathrm{R}}, \bar{B}_{\mathrm{R}}, \bar{C}_{\mathrm{R}})$ is a weighted reduced-order model.

If the full-order original model is minimal, then $\bar{\Sigma}_1 > 0$.

Unfortunately, the stability of the reduced-order model here cannot, in general, be guaranteed.

However, there are some special cases of weightings for which the reduced-order models are stable (**Exercise 14.22**). Also, no a priori error found for the approximate is known.

14.8 SUMMARY AND COMPARISONS OF MODEL REDUCTION PROCEDURES

We have described the following techniques for model reduction of a stable system:

(i) The balanced truncation procedure (**Algorithm 14.4.1**)
(ii) The Schur method (Algorithm **14.4.2**)
(iii) The Hankel-norm approximation algorithm (**Algorithm 14.5.1**)
(iv) Frequency-weighted model reduction (**Section 14.7**).

For the first two methods (i)-(ii), the error satisfies

$$\|G(s) - G_{\mathrm{R}}(s)\|_{\infty} \leq 2(\sigma_{d+1} + \sigma_{d+2} + \cdots + \sigma_N),$$

In the method (i), $G_{\mathrm{R}}(s)$ is obtained by truncating the balanced realization of $G(s)$ to the first $(s_1 + s_2 \ldots + s_d)$ states, where s_i is the multiplicity of σ_i. For the method (ii), $G_R(s)$ is obtained by Algorithm 14.4.2. For a similar error bound for the method (iii), see Theorem 14.5.4. Furthermore, for this method, the reduced-order model $G_{\mathrm{R}}(s)$ has the property: $\inf \| G - G_{\mathrm{R}}(s) \|_H = \sigma_{k+1}$, where $G_{\mathrm{R}}(s)$ is of McMillan degree k.

The weighted model reduction procedure in Section 14.7 does not enjoy any of the above properties. Even the stability in general cannot be guaranteed. Stability, however, in some special cases can be proved. See Enns (1984) for details. Discussion of this section has been taken from Zhou *et al.* (1996).

If the system is not stable, model reduction is still possible using the three simple steps of Section 14.6.

The balanced truncation procedure for model reduction (Algorithm 14.4.1) and Algorithm 14.5.1 need computation of a balanced realization. Two algorithms (**Algorithms 14.2.1** and **14.2.2**) have been described for this purpose. Both these algorithms suffer from the danger of possible ill-conditioning of the transforming matrices. **However, the methods usually work well in practice for well-equilibrated systems.**

The Schur method has been designed to avoid such possible ill-conditioning.

Unfortunately, because of the requirement of explicitly computing the product of the controllability and observability Grammians, *the Schur method is usually less accurate for moderately ill-conditioned systems than the square-root method* (see Varga 2001). The main advantages of the balanced truncation procedure and the Schur method have been combined in the **balanced-free square-root method** by Varga (1991). Numerical experiments performed by Varga (2001) show that the accuracy of this method is usually better than either of the Schur methods or the balanced truncation method using the square-root algorithm for balancing.

Finally, we remark that it is **very important that the system be scaled properly for the application of the balanced-truncation or the Hankel-norm approximation method.** One way to do this is to attempt to reduce the 1-norm of the scaled system matrix

$$S = \begin{pmatrix} Z^{-1}AZ & Z^{-1}B \\ CZ & 0 \end{pmatrix}, \text{ where } Z \text{ is a positive definite matrix.}$$

Note that the Hankel singular values are not affected by such a coordinate transformation; in particular, by coordinate scaling of diagonal matrices.

For a comparative study of different model reduction algorithms and detailed description of available software, see Varga (2001). See also Varga (1994).

14.9 SOME SELECTED SOFTWARE

14.9.1 MATLAB Control System Toolbox

State-space models

balreal	Grammian-based Balancing of state-space realization.
modred	Model state reduction.
ssbal	Balancing of state-space model using diagonal similarity.

14.9.2 MATCONTROL

BALSVD	Internal balancing using the SVD
BALSQT	Internal balancing using the square-root algorithm
MODREDS	Model reduction using the Schur method
HNAPRX	Hankel-norm approximation.

14.9.3 CSP-ANM

Model reduction

- The Schur method for model reduction is implemented as `DominantSubsystem` [*system*, `Method`→`SchurDecomposition`].
- The square-root method for model reduction is implemented as `DominantSubsystem` [*system*, `Method`→`SquareRoot`].

14.9.4 SLICOT

Model reduction

AB09AD	Balance and truncate model reduction
AB09BD	Singular perturbation approximation based model reduction
AB09CD	Hankel-norm approximation based model reduction
AB09DD	Singular perturbation approximation formulas
AB09ED	Hankel-norm approximation based model reduction of unstable systems
AB09FD	Balance and truncate model reduction of coprime factors
AB09GD	Singular perturbation approximation of coprime factors
AB09ID	Frequency-weighted model reduction based on balanced truncations
AB09KD	Frequency-weighted Hankel-norm approximation
AB09MD	Balance and truncate model reduction for the stable part
AB09ND	Singular perturbation approximation based model reduction for the stable part.

State-space transformations

TB01ID	Balancing a system matrix for a given triplet.

14.9.5 MATRIX$_X$

Purpose: Convert a discrete dynamic system into an internally balanced dynamic form.

Syntax: [SB, SIGMASQ, T] = DBALANCE (SD, NS)

Purpose: Compute a reduced order form of a discrete-time system.
Syntax: [SR, NSR] = DMREDUCE (SD, NS, KEEP)

Purpose: Compute a reduced-order form of a continuous system.
Syntax: [SR, NSR] = MREDUCE (S, NS, KEEP)

Purpose: Perform model structure determination.
Syntax: [THETA, COR, COV] = MSD (X, Y)
The other software packages dealing with model reduction include:

- **MATRIX$_X$ Model Reduction Module** (1998) by B.D.O. Anderson and B. James.
- **μ-Analysis and Synthesis Toolbox 1.0** by G. Balas, J. Doyle, K. Glover, A. Packard and R. Smith (1998).
- **Robust Control Toolbox** 2.0 by R.Y. Chiang and M.G. Safonov.

14.10 SUMMARY AND REVIEW

The chapter covers the topics:

- Internal balancing
- Model reduction
- Hankel-norm approximation.

Internal Balancing

Given an $n \times n$ stable minimal realization (A, B, C), there always exists a transformation T that simultaneously diagonalizes both the controllability and observability Grammians to the same diagonal matrix $\Sigma = \text{diag}(\sigma_1, \ldots, \sigma_n)$, where $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r > \sigma_{r+1} > \sigma_{r+2} > \cdots > \sigma_n$. The numbers $\sigma_i, i = 1, \ldots, n$ are the **Hankel singular values.**

In this case, the transformed system $(\tilde{A}, \tilde{B}, \tilde{C})$, is called **internally balanced. Algorithms 14.2.1** and **14.2.2** compute balanced realization of a **continuous-time system.** The internal balancing of a **discrete-time system is discussed in Section 14.3.**

Model Reduction

The problem of model reduction is the problem of constructing a qth order model from a given nth order model $(n > q)$ in such a way that the reduced qth order model is close to the original system in some sense. The precise mathematical definition of model reduction appears in **Section 14.4**.

Model reduction via internal balancing: Once a system is internally balanced, a desired reduced-order model can be obtained by eliminating the states corresponding to the less controllable and observable modes (**Algorithm 14.4.1**).

Theorem 14.4.1 shows that a truncated model is also balanced and stable, and furthermore, if $G(s)$ and $G_R(s)$ are the respective transfer functions of the original and the truncated model, then

$$\|G(s) - G_R(s)\|_\infty \leq 2(\sigma_{d+1}, \ldots, \sigma_N),$$

where the states corresponding to $\sigma_{d+1}, \ldots, \sigma_N$ are eliminated.

The Schur method for model reduction: There are some numerical difficulties associated with the procedure of finding a reduced order model via internal balancing using Algorithms 14.2.1 and 14.2.2. The transforming matrix T in Algorithm 14.2.1 and the matrices L and Z in Algorithm 14.2.2 can be, in some cases, highly ill-conditioned. An alternative method (**Algorithm 14.4.2**) for model reduction based on the real Schur decomposition of the product of the controllability and observability Grammians, is described in **Section 14.4. The transforming matrix T in this case is orthogonal, and, therefore, well-conditioned.** The Schur method does not give an internally balanced system; however, the essential properties of the original system are preserved. In fact, **Theorem 14.4.2** shows that the transfer function matrix obtained by the Schur method is exactly the same as that of the one obtained via **Algorithm 14.4.1**.

A possible numerically difficulty with Algorithm 14.4.2 is the explicit computation of the product of the controllability and observability Grammians. In this case, instead of explicitly computing the controllability and observability Grammians, their Cholesky factors can be computed using the Hammarling algorithm (**Algorithm 8.6.1**) in Chapter 8.

Combining the advantages of the Schur method and the square-root algorithm, a **balancing-free square root method** has been developed. This is described in **Section 14.4.3**.

Hankel-norm approximation: Given a stable $G(s)$, the problem of finding a $\hat{G}(s)$ of McMillan degree k such that $\|G(s) - \hat{G}(s)\|_H$ is minimized is called an optimal Hankel-norm approximation.

A characterization of all solutions to Hankel-norm approximation is given in Section 14.5.1 (**Theorem 14.5.2**). An algorithm (**Algorithm 14.5.1**) for computing an optimal Hankel-norm approximation is then presented.

Model Reduction of an Unstable System

The model reduction of an unstable system can be achieved by decomposing the model into its stable and unstable part, followed by finding a model reduction of the stable part and finally adding the reduced-order model of the stable part with the unstable part. This is descried in **Section 14.6**. For this and another approach, based on stable rational coprime factorization, see Varga (2001).

Weighted Model Reduction

Sometimes the errors at high frequencies in a reduced-order model can be reduced using weights on the model. This is discussed in **Section 14.7**.

Comparison of the Model Reduction Procedures

The model reduction procedures are summarized and a brief comparative discussion of different procedures is presented in **Section 14.8.**

14.11 CHAPTER NOTES AND FURTHER READING

The internal balancing algorithms, **Algorithms 14.2.1** and **14.2.2** are due to Laub (1980) and Tombs and Postlethwaite (1987), respectively. The idea of model reduction via balanced truncation was first introduced by Moore (1981).

The stability property of the truncated subsystem (part (a) of Theorem 14.4.1) was obtained by Pernebo and Silverman (1982) and the error bound (part (b) of Theorem 14.4.1) is due to Glover (1984) and Enns (1984).

The Schur algorithm for model reduction and Theorem 14.4.2 is due to Safonov and Chiang (1989). The balancing-free square-root method for model reduction is due to Varga (1991).

The Hankel-norm approximation problem was introduced and solved by Glover in a celebrated paper (Glover 1984). Besides the topic of Hankel-norm approximation of a transfer function, the paper contains many other beautiful results on systems theory and linear algebra. A good discussion of this topic can also be found in the book by Zhou *et al.* (1996). See also Glover (1989).

For results on discrete-time balanced model reduction, see Al-Saggaf and Franklin (1987), and Hinrichsen and Pritchard (1990).

The idea of frequency weighted model reduction is due to Enns (1984). Other subsequent results on this and related topics can be found in Al-Saggaf and Franklin (1988), Glover (1986, 1989), Glover *et al.* (1992), Hung and Glover (1986), Liu and Anderson (1990), Zhou (1993), etc.

For a discussion on Balanced Stochastic Truncation (BST) method, see Zhou *et al.* (1996).

The idea of singular perturbation approximation is due to Liu and Anderson (1989).

For an optimal Hankel norm approximation procedure with stable weighting functions, see Hung and Glover (1986).

The other papers on Hankel norm approximation include Kung and Lin (1981) and Latham and Anderson (1986).

A recent book by Obinata and Anderson (2000) deals exclusively with the topic of model reduction.

The paper by Green (1988) deals with stochastic balanced realization. For more on this topic, see Zhou *et al.* (1996).

Exercises

14.1 Prove part (a) of Theorem 14.4.1 and fill in the missing details of part (b), whenever indicated in the book.

14.2 Let

$$G(s) = \left[\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right].$$

Suppose that there exists a symmetric matrix $P = \mathrm{diag}(P_1, 0)$, with P_1 nonsingular, such that

$$AP + PA^{\mathrm{T}} + BB^{\mathrm{T}} = 0.$$

Partition $G(s)$ conformably with P as

$$\begin{bmatrix} A & B \\ C & D \end{bmatrix} = \left[\begin{array}{ccc} \hat{A}_{\mathrm{R}} & A_{12} & \hat{B}_{\mathrm{R}} \\ A_{21} & A_{22} & B_2 \\ \hline \hat{C}_{\mathrm{R}} & C_2 & D \end{array}\right].$$

Then prove that $\left[\begin{array}{cc} \hat{A}_{\mathrm{R}} & \hat{B}_{\mathrm{R}} \\ \hline \hat{C}_{\mathrm{R}} & D \end{array}\right]$ is also a realization of $G(s)$. Moreover, if $\hat{A}_{\mathrm{R}}$ is stable, then $(\hat{A}_{\mathrm{R}}, \hat{B}_{\mathrm{R}})$ is controllable.

14.3 Based on the result of Exercise 14.2 develop a method for extracting a controllable subsystem from a stable noncontrollable system.

14.4 (Zhou *et al.* (1996)) Let $G(s)$ be the same as in Exercise 14.2. Suppose that there exists a symmetric matrix $Q = \mathrm{diag}(Q_1, 0)$, with Q_1 is nonsingular, such that $QA + A^{\mathrm{T}}Q + C^{\mathrm{T}}C = 0$. Partition the realization (A, B, C, D) conformably with Q as in Exercise 14.2. Then prove that

$$\left[\begin{array}{cc} \hat{A}_{\mathrm{R}} & \hat{B}_{\mathrm{R}} \\ \hline \hat{C}_{\mathrm{R}} & D \end{array}\right]$$

is also a realization of $G(s)$. Prove further that $(\hat{A}_{\mathrm{R}}, \hat{C}_{\mathrm{R}})$ is observable if $\hat{A}_{\mathrm{R}}$ is stable.

14.5 Based on Exercise 14.4, develop a method for extracting an observable subsystem from a stable nonobservable system.

14.6 Construct your own example to illustrate the numerical difficulties of Algorithm 14.2.1.

14.7 Prove that the rows $\{1, \ldots, d\}$ and the rows $\{d+1, \ldots, n\}$ of T^{-1} in Algorithm 14.2.1, form bases for the left eigenspaces of the matrix $C_G O_G$ associated with the eigenvalues $\{\sigma_1^2, \ldots, \sigma_d^2\}$, and $\{\sigma_{d+1}^2, \ldots, \sigma_n^2\}$, respectively.

14.8 Prove that the columns of the matrix V_{1S} and those of the matrix U_{2S} in the Schur algorithm (Algorithm 14.4.2) for model reduction, form orthonormal bases for the right and left invariant subspace of $C_G O_G$ associated with the large eigenvalues $\sigma_1^2, \ldots, \sigma_d^2$.

14.9 Prove that the controllability and observability Grammians of the reduced-order model obtained by the Schur algorithm (Algorithm 14.4.2) are, respectively, given by $\hat{C}_G^R = S_1^T C_G S_1$ and $\hat{O}_G^R = S_2 O_G S_2$, where C_G and O_G are the controllability and observability Grammians of the original model.

14.10 (a) Modify the Schur algorithm for model reduction by making use of Hammarling's algorithm (**Algorithm 8.6.1**) so that the explicit formation of the product $C_G O_G$ is avoided, and only the Cholesky factors L_c and L_o are computed. (Consult Safonov and Chiang (1989)).

(b) Work out an example to demonstrate the superiority of this modified Schur algorithm over the Schur algorithm.

14.11 (a) Prove that the matrix T defined by (14.3.4) transforms the discrete-time system (14.3.1) to the balanced system (14.3.5).

(b) Work out a discrete analog of Algorithm 14.2.2.

14.12 (Zhou *et al.* (1996)). Let

$$G(s) = \left[\begin{array}{c|c} A & B \\ \hline C & O \end{array}\right]$$

be the transfer function of a balanced realization. Then prove that

$$\sigma_1 \leq \|G\|_\infty \leq \int_0^\infty \|Ce^{At}B\|dt \leq 2\sum_{i=1}^{N} \sigma_i.$$

14.13 Construct an example to show that if the diagonal entries of the matrix Σ of the balanced Grammian are all distinct, then every subsystem of the balanced system is asymptotically stable. Construct another example to show that this condition is only sufficient.

14.14 Construct an example to show that the bound of Theorem 14.4.1 can be loose if the quantities σ_i, $i = 1, \ldots, n$ are close to each other.

(**Hint:** Construct a stable realization $G(s)$ such that $G^T(-s)G(s) = I$ and then construct a balanced realization of $G(s)$. Now make a small perturbation to this balanced realization and work with this perturbed system.)

14.15 (a) Develop a Schur method for model reduction of the discrete-time system.

(b) Give a simple example to illustrate the method.

(c) Give a flop-count of the method.

14.16 *Minimal realization using block diagonalization* (*Varga 1991*). Consider the following algorithm:

Step 1. Reduce A to block diagonal form and update B and C, that is, find a nonsingular matrix T such that $T^{-1}AT = \text{diag}(\bar{A}_1, \bar{A}_2, \ldots, \bar{A}_r)$, $T^{-1}B = (\bar{B}_1, \bar{B}_2, \ldots, \bar{B}_r)$, $CT = (\bar{C}_1, \bar{C}_2, \ldots, \bar{C}_r)$. (see Exercise 8.10).

Assume that the diagonal blocks in A have disjoint spectra.

Step 2. Find an MR of each of the system $(\bar{A}_i, \bar{B}_i, \bar{C}_i), i = 1, \ldots, r$ using Algorithm 14.2.1 or 14.2.2, as appropriate.

Let $(\hat{A}_i, \hat{B}_i, \hat{C}_i), i = 1, \ldots, r$ be the computed MR of $(\bar{A}_i, \bar{B}_i, \bar{C}_i)$ in Step 2.

Then show that the system $(\hat{A}, \hat{B}, \hat{C})$ defined by:

$$\hat{A} = \operatorname{diag}(\hat{A}_1, \hat{A}_2, \ldots, \hat{A}_r), \qquad \hat{B} = \begin{pmatrix} \hat{B}_1 \\ \hat{B}_2 \\ \vdots \\ \hat{B}_r \end{pmatrix}, \qquad \hat{C} = (\hat{C}_1, \hat{C}_2, \ldots, \hat{C}_r),$$

is an MR of (A, B, C).

14.17 Using the matrix version of bilinear transformation:

$$s = \frac{z-1}{z+1},$$

prove that the Hankel singular values of the discrete and continuous systems are identical.

(**Hint:** Obtain the system matrices (A, B, C) for the continuous-time system from the system matrices $(\bar{A}, \bar{B}, \bar{C})$ of the discrete-time system and then show that the controllability Grammian (observability Grammian) of (A, B, C) is the same as the controllability Grammian (observability Grammian) of $(\bar{A}, \bar{B}, \bar{C})$.)

14.18 Using the bilinear transformation of Exercise 14.17 and Algorithm 14.5.1, find an optimal Hankel-norm approximation for the discrete-time system defined by the matrices in Example 14.3.1.

14.19 Write down a discrete analog of Algorithm 14.2.2 and apply the algorithm to the system (A, B, C) defined by

$$A = \begin{pmatrix} 0.0001 & 1 & 0 \\ 0 & 0.1200 & 1 \\ 0 & 0 & 0 \end{pmatrix}, \qquad B = (1, 1, 0)^{\mathrm{T}}, \qquad C = (1, 1, 1).$$

14.20 (Safonov and Chiang 1989). Consider the system (A, B, C) given by

$$A = \begin{pmatrix} -6 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & -8 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -10 & 3 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & -8 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -13 & -3 & 9 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & -8 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & -8 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & -14 & -9 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & -8 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -2 \end{pmatrix},$$

$$B = \begin{pmatrix} 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 10^{-3} \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 10^{-3} \end{pmatrix}^{\mathrm{T}},$$

$$C = \begin{pmatrix} 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 5 \times 10^5 \\ 0 & 0 & 0 & 0 & 0 & 0 & -6 & 1 & -2 & 5 \times 10^5 \end{pmatrix}.$$

Find a reduced-order model of order 4 using

(a) Balanced truncation via Algorithms 14.2.1 and 14.2.2.

(b) The Schur method (Algorithm 14.4.2).

Compare the results with respect to the condition numbers of the transforming matrices and the $\|\cdot\|_\infty$ norm errors.

14.21 Prove that the weighting Grammians $\bar{C}_{\mathrm{G}}$ and $\bar{O}_{\mathrm{G}}$ are given by the equations (14.7.2) and (14.7.3).

14.22 Consider the two special cases of the frequency-weighted model reduction:

Case 1. $W_{\mathrm{i}}(s) = I$ and $W_{\mathrm{o}}(s) \neq I$,

Case 2. $W_{\mathrm{i}}(s) \neq I$ and $W_{\mathrm{o}}(s) = I$.

Prove that the reduced-order model $(\bar{A}_{\mathrm{R}}, \bar{B}_{\mathrm{R}}, \bar{C}_{\mathrm{R}})$ is stable provided that it is controllable in Case 1 and is observable in Case 2.

(**Hint:** Write the balanced Grammian $\Sigma = \mathrm{diag}(\Sigma_1, \Sigma_2)$. Then show that

$$\bar{A}_{\mathrm{R}}\Sigma_1 + \Sigma_1\bar{A}_{\mathrm{R}}^{\mathrm{T}} + \bar{B}_{\mathrm{R}}\bar{B}_{\mathrm{R}}^{\mathrm{T}} = 0, \quad \text{and} \quad \bar{A}_{\mathrm{R}}^{\mathrm{T}}\Sigma_1 + \Sigma_1\bar{A}_{\mathrm{R}} + \bar{C}_{\mathrm{R}}^{\mathrm{T}}\bar{C}_{\mathrm{R}} = 0).$$

Work out an example to illustrate the result.

14.23 *Singular perturbation approximations.* Let $(\bar{A}, \bar{B}, \bar{C})$ be a balanced realization of (A, B, C). Partition the matrices $\bar{A}, \bar{B}, \bar{C}$ as:

$$\bar{A} = \begin{pmatrix} \bar{A}_{11} & \bar{A}_{12} \\ \bar{A}_{21} & \bar{A}_{22} \end{pmatrix}, \qquad \bar{B} = \begin{pmatrix} \bar{B}_1 \\ \bar{B}_2 \end{pmatrix}, \qquad \bar{C} = (\bar{C}_1, \bar{C}_2).$$

Then the system $(\hat{A}, \hat{B}, \hat{C})$ defined by

$$\hat{A} = \bar{A}_{11} + \bar{A}_{12}(\gamma I - \bar{A}_{22})^{-1}\bar{A}_{21}, \qquad \hat{B} = \bar{B}_1 + \bar{A}_{12}(\gamma I - \bar{A}_{22})^{-1}\bar{B}_2,$$
$$\hat{C} = \bar{C}_1 + \bar{C}_2(\gamma I - \bar{A}_{22})^{-1}\bar{A}_{21}$$

is called **the balanced singular perturbation approximation** of $(\bar{A}, \bar{B}, \bar{C})$ (Liu and Anderson 1989). ($\gamma = 0$ for a continuous-time system and $\gamma = 1$ for a discrete-time system).

(a) Compute singular perturbation approximations of the system in Example 14.2.1

using Algorithms 14.2.1 and 14.2.2.

(b) Show how the balancing-free square root method in Section 14.4.3 can be modified to compute singular perturbation approximation (**Hint.** Find the SVD of $Y^{\mathrm{T}}X$ and then compute L and Z from the matrices of the SVD). See Varga (1991).

(c) Apply the modified balancing-free square-root method in (b) to the system in Example 14.2.1 and compare the results.

References

Al-Saggaf U.M. and Franklin G.F., "An error bound for a discrete reduced order model of a linear multivariable sytem", *IEEE Trans. Autom. Control*, Vol. AC-32, pp. 815–819, 1987.

Al-Saggaf U.M. and Franklin G.F. "Model reduction via balanced realizations: an extension and frequency weighting techniques," *IEEE Trans. Autom. Control*, Vol. AC-33(7), pp. 687–692, 1988.

Balas G., Doyle J., Glover K., Packard A. and Smith R. *μ-Analysis and Synthesis Toolbox 3.0.4*, The MathWorks Inc., Natick, MA, 1998.

Chiang R.Y. and Safonov M.G. *Robust Control Toolbox 2.0.6.*, The MathWorks Inc., Natick, MA, 1997.

Enns D.F. "Model reduction with balanced realizations: An error bound and a frequency weighted generalization", *Proc. 23rd IEEE Conf. Dec. Control*, pp. 127–132, 1984.

Glover K. "Robust stabilization of linear multivariable systems: relations to approximation", *Int. J. Control*, vol. 43(3), pp. 741–766, 1986.

Glover K. *A Tutorial on Hankel-norm Approximation, Data to Model*, Springer-Verlag, New York, 1989. (Willems J.C. ed.)

Glover K., Limebeer D.J.N., and Hung Y.S. "A structured approximation problem with applications to frequency weighted model reduction," *IEEE Trans. Autom. Control*, Vol. AC-37(4), pp. 447–465, 1992.

Glover K. "All optimal Hankel-norm approximations of linear multivariable systems and their L^{∞}-error bounds," *Int. J. Cont.*, 39, pp. 1115–1193, 1984.

Green M. "Balanced stochastic realizations," *Lin. Alg. Appl.*, Vol. 98, pp. 211–247, 1988.

Heath M.T., Laub A.J., Paige C.C., and Ward R.C. "Computing the singular value decomposition of the product of two matrices," *SIAM J. Sci. Stat. Comput.*, Vol. 7, pp. 1147–1159, 1986.

Hinrichsen D. and Pritchard A.J. "An improved error estimate for reduced-order models of discrete time system," *IEEE Trans. Autom. Control*, Vol. AC-35, pp. 317–320, 1990.

Hung Y.S. and Glover K. "Optimal Hankel-norm approximation of stable systems with first order stable weighting functions," *Syst. Control Lett.*, Vol. 7, pp. 165–172, 1986.

Kung S.K. and Lin D.W. "Optimal Hankel norm model reduction: multivariable systems," *IEEE Trans. Autom. Control*, Vol. AC-26, pp. 832–852, 1981.

Latham G.A. and Anderson B.D.O. "Frequency weighted optimal Hankel-norm approximation of stable transfer function," *Syst. Control Lett.*, Vol. 5, pp. 229–236, 1986.

Laub A.J. "Computation of 'balancing' transformations," *Proc. 1980 JACC*, Session FA8-E, 1980.

Liu Y. and Anderson B.D.O. "Singular perturbation approximation of balanced systems," *Int. J. Control*, Vol. 50, pp. 1379–1405, 1989.

Liu Y. and Anderson B.D.O. "Frequency weighted controller reduction methods and loop transfer recovery," *Automatica*, Vol. 26(3), pp. 487–497, 1990.

MATLAB, *Control System Toolbox* 4.2, The MathWorks Inc., Natick, MA, 1998.

MATRIX$_X$, *Xmath Model Reduction module*, ISI, Santa Clara, CA, January 1998.

Moore B.C. "Principal component analysis in linear systems: controllability, observability and model reduction," *IEEE Trans. Autom. Control*, Vol. AC-26, pp. 17–31, 1981.

Obinata G. and Anderson B.D.O. *Model Reduction for Control System Design*, Springer-Verlag, New York, 2000.

Pernebo L. and Silverman L.M. "Model reduction via balanced state space representation," *IEEE Trans. Autom. Control*, Vol. AC-27(2), pp. 382–387, 1982.

Safonov M.G. and Chiang R.Y. "A Schur method for balanced-truncation model reduction," *IEEE Trans. Autom. Control*, Vol. 34, pp. 729–733, 1989.

Tombs M.S. and Postlethwaite I. "Truncated balanced realization of a stable non-minimal state-space system," *Int. J. Control*, Vol. 46, pp. 1319–1330, 1987.

Varga A. "Balancing-free square-root algorithm for computing singular perturbation approximations," *Proc. 30th IEEE Conf. Dec. Control*, pp. 1062–1065, 1991.

Varga A. "Numerical methods and software tools for model reduction," *Proc. 1st MATHMOD Conf., Vienna*, Troch I. and Breitenecker F., eds. Vol. 2, pp. 226–230, 1994.

Varga A. "Model reduction software in the SLICOT library," *Applied and Computational Control, Signals and Circuits: Recent Developments*, Datta B.N. *et al.*, eds. pp. 239–282, Kluwer Academic Publisher, Boston, 2001.

Zhou K (with Doyle J.C. and Glover K.), *Robust and Optimal Control,* Prentice Hall, Upper Saddle River, NJ, 1996.

Zhou K. "Frequency weighted model reduction with $\mathcal{L}_\infty$ error bounds," *Syst. Control Lett.*, Vol. 21, pp. 115–125, 1993.