1 Linear Systems, Model Truncation, and Singular Perturbation

1.1 Linear State-Space Systems

We consider linear state-space systems

$$G : \begin{cases} \dot{x} = Ax + Bu, \quad x(0) = x_0 \\ y = Cx + Du \end{cases} \tag{1.1}$$

with *state* $x(t) \in \mathbb{R}^n$, *input* $u(t) \in \mathbb{R}^m$, and *output* $y(t) \in \mathbb{R}^p$. Knowledge of x_0 and $u(t)$ in the time interval $[0, T]$ determines $x(t)$ and $y(t)$ uniquely in the same interval by

$$x(t) = e^{At}x_0 + \int_0^t e^{A(t-\tau)}Bu(\tau)\,d\tau, \quad y(t) = Cx(t) + Du(t).$$

Often we assume A is a *Hurwitz matrix*, i.e., all eigenvalues of A are in the open left complex half plane and the system is (asymptotically) stable.

Large classes of models can be written in the form (1.1). Examples include

- Discretized partial differential equations (finite difference/finite element methods etc.), such as diffusion and wave equations; and
- Linearized nonlinear ordinary differential equations.

If the input-output mapping $u \mapsto y$ is the main interest, and not the state x, the system G can also be represented by its transfer function

$$G(s) = C(sI - A)^{-1}B + D \in \mathbb{C}^{p\times m}$$

for complex frequencies $s \in \mathbb{C}$. As a measure of system size, and to measure the distance between two different systems, we regularly use the H_∞-norm:

$$\begin{aligned} \|G\|_\infty &:= \sup_{s\in\mathbb{C}_+} \|G(s)\| \quad (\mathbb{C}_+ \text{ is the open complex right-half plane}) \\ &= \sup_{\omega} \|G(j\omega)\| \quad (G(s) \text{ has no poles in } \mathbb{C}_+) \end{aligned}$$

which is finite if, and only if, $G(s)$ is *stable* (has no poles in the closed right complex half plane). Here $\|G(s)\|$ denotes the largest singular value of the matrix $G(s)$ in the MIMO (Multi-Input–Multi-Output) case. In the SISO (Single-Input–Single-Output) case this is equal to the magnitude of the complex number $G(s)$.

1.2 Reduced Order Systems and Approximation Criteria

We identify the complexity of the system G with its order n. Some motivation for this definition are

- optimal controllers (LQG/H_2/H_∞) for G tend to have order of at least n, and
- the simulation time of the system (1.1) is strongly correlated to the number n of differential equations.

A reduced order system ("an approximation") of G is a state-space system G_r

$$G_r : \begin{cases} \dot{z} = A_r z + B_r u, \quad z(0) = z_0 \\ y_r = C_r z + D_r u \end{cases} \tag{1.2}$$

such that $z(t) \in \mathbb{R}^r$, where $r < n$, and $y_r(t) \in \mathbb{R}^p$. Note that the same input u is applied to both G and G_r.

Not only should G_r be of lower order than G, its trajectories should say something about the trajectories of G. Otherwise, we can hardly speak of an approximation. The main approximation criterium we will be interested in, in this course, is to make $\|G - G_r\|_\infty$ small. Motivation for this choice will be given throughout the course. One simple motivation is that it is a measure of the worst-case error. Other criteria, such as the relative criterion $\|G^{-1}(G - G_r)\|_\infty$ and the frequency-weighted criterion $\|W_1(G - G_r)W_2\|_\infty$ will also be discussed later in the course.

Since we mainly look at approximating input-output behavior $u \mapsto y$, choices of inputs and outputs are essential to get an approximation G_r that captures what you want.

1.3 Truncation and Singular Perturbation

Good approximations G_r can often be obtained by means of *truncation* or *singular perturbation (residualization)*. Both methods are done in two steps:

Step 1: Change the coordinates $x(t)$. That is, find a *suitable* invertible matrix $T \in \mathbb{R}^{n\times n}$ and transform the state-space model according to

$$\bar{A} = T^{-1}AT = \begin{pmatrix} \bar{A}_{11} & \bar{A}_{12} \\ \bar{A}_{21} & \bar{A}_{22} \end{pmatrix}, \ \bar{A}_{11} \in \mathbb{R}^{r\times r}, \qquad \bar{B} = T^{-1}B = \begin{pmatrix} \bar{B}_1 \\ \bar{B}_2 \end{pmatrix}, \ \bar{B}_1 \in \mathbb{R}^{r\times m},$$
$$\bar{C} = CT = \begin{pmatrix} \bar{C}_1 & \bar{C}_2 \end{pmatrix}, \ \bar{C}_1 \in \mathbb{R}^{p\times r}, \qquad \bar{D} = D.$$

Step 2: Depending on method, obtain G_r by:

- Truncation:

$$A_r = \bar{A}_{11}, \qquad B_r = \bar{B}_1$$
$$C_r = \bar{C}_1, \qquad D_r = \bar{D}.$$

- Singular perturbation:

$$A_r = \bar{A}_{11} - \bar{A}_{12}\bar{A}_{22}^{-1}\bar{A}_{21}, \qquad B_r = \bar{B}_1 - \bar{A}_{12}\bar{A}_{22}^{-1}\bar{B}_2$$
$$C_r = \bar{C}_1 - \bar{C}_2\bar{A}_{22}^{-1}\bar{A}_{21}, \qquad D_r = \bar{D} - \bar{C}_2\bar{A}_{22}^{-1}\bar{B}_2.$$

Note the following:

(P1) using truncation, we always have $G(\infty) = G_r(\infty)$; and

(P2) using singular perturbation, we always have $G(0) = G_r(0)$.

Hence, these methods achieve perfect approximation either at zero or at infinite frequency. Depending on how the reduced model will be used, one chooses one method or the other.

For successful application of these methods (usually it is not enough to have a good approximation only at a single frequency), we need to find good coordinate transformations T, and a suitable approximation order r. The following lectures will deal with this problem. Mostly, we will use the truncation method in the following, for simplicity.

1.4 Truncation $\sim$ Projection

The truncation method can also be seen as a projection in the original state space $\mathbb{R}^n$ to an r-dimensional subspace corresponding to the reduced state space $\mathbb{R}^r$. Consider the matrices

$$\begin{aligned} W^T &= \begin{pmatrix} I_r & 0_{r\times(n-r)} \end{pmatrix} T^{-1} \in \mathbb{R}^{r\times n} \quad (\text{"}x \to z\text{"}) \\ V &= T \begin{pmatrix} I_r \\ 0_{(n-r)\times r} \end{pmatrix} \in \mathbb{R}^{n\times r}, \qquad (\text{"}z \to x\text{"}) \end{aligned} \tag{1.3}$$

which combine coordinate transformation and truncation. Note that $P = VW^T \in \mathbb{R}^{n\times n}$ defines a (generally oblique) projection on $\mathbb{R}^n$ to the subspace spanned by the columns of V,

$$P = P^2,$$

because $W^TV = I_r$.

Given any projection $P = VW^T$ (i.e., $W^TV = I_r$), we can approximate G using a so-called a *Petrov-Galerkin projection*. (If $W^T = V^T$, this is called a *Galerkin projection*.) A Petrov-Galerkin projection of G, yielding G_r, is given by the state-space realization

$$\begin{aligned} A_r &= W^TAV, & B_r &= W^TB \\ C_r &= CV, & D_r &= D, \end{aligned}$$

which of course coincides with the truncated model in Section 1.3 when W^T and V are chosen as (1.3). To understand the Petrov-Galerkin projection and where this realization comes from, we can do the following analysis. Assume we want to try to express the solution $x(t) \in \mathbb{R}^n$ to the model G only in r variables. Such a solution can be written as $x(t) = Vz(t)$, where $z(t) \in \mathbb{R}^r$ and $V \in \mathbb{R}^{n\times r}$. V should be chosen such that its columns span a subspace where we believe the solution $x(t)$ will lie in. If the believed solution is inserted into the original state-space model G, we obtain

$$\dot{x} = V\dot{z} = AVz + Bu + E, \tag{1.4}$$

where E is the equation error. Now, $x(t) = Vz(t)$ is a solution to the original problem if, and only if, $E(t) = 0$ for all t. There are n equations in (1.4), but only r unknowns in z, and so the system is generally over determined. To find a unique solution, we can require that the projection of the residual E onto the subspace spanned by V is zero. This projection is given by W^T. Hence, we add the condition

$$W^TE(t) = 0, \quad \forall t$$

to (1.4). We then obtain the equation

$$\dot{z} = W^TAVz + W^TBu,$$

which exactly is the Petrov-Galerkin projection of (1.1). From this, $z(t)$ can be computed, and the projection of the resulting residual E onto V is zero.

1.5 Recommended Reading

Sections 3.2–3.2.3 in *Linear Robust Control* [1] discuss linear systems and the H_∞-norm. We will discuss signals, systems, and function spaces more during Exercise 1.

Sections 9.1–9.2.2 in *Linear Robust Control* [1] discuss model truncation and singular perturbation in more detail. The controllability and observability Gramians mentioned on page 315 will be introduced at a later stage in this course.

1.6 Exercises

EXERCISE 1.1 (Modal representation and truncation)

Assume that A has n distinct eigenvalues λ_i. Then there exists a coordinate transformation T such that

$$\bar{A} = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}, \quad \bar{B} = \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix}, \quad \bar{C} = \begin{pmatrix} c_1 & \dots & c_n \end{pmatrix}.$$

This is called a modal representation of G. The state corresponding to λ_i is often called the i-th mode of the system. Prove that using truncation on the modal representation we obtain

$$G(s) - G_r(s) = \sum_{i=r+1}^{n} \frac{c_i b_i}{s - \lambda_i}, \tag{1.5}$$

$$\|G - G_r\|_\infty \leq \sum_{i=r+1}^{n} \frac{|c_i b_i|}{|\text{Re}\, \lambda_i|}. \tag{1.6}$$

To show the error bound (1.6) you need that A is a Hurwitz matrix. Why? Discuss how an error bound like (1.6) can be used. What modes should be truncated?

EXERCISE 1.2 (Modal truncation)

What modes should be truncated in the following systems if $\|G - G_r\|_\infty$ should be small?

a)

$$\bar{A} = \begin{pmatrix} -1 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & -3 \end{pmatrix}, \quad \bar{B} = \begin{pmatrix} 1 \\ 0.1 \\ 0.1 \end{pmatrix}, \quad \bar{C} = \begin{pmatrix} 1 & 1 & 0.1 \end{pmatrix}.$$

b)

$$\bar{A} = \begin{pmatrix} -1 & 0 & 0 \\ 0 & -0.1 & 0 \\ 0 & 0 & -0.101 \end{pmatrix}, \quad \bar{B} = \begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix}, \quad \bar{C} = \begin{pmatrix} 1 & 1 & 1 \end{pmatrix}.$$

c)

$$\bar{A} = \begin{pmatrix} -1 & 0 & 0 \\ 0 & -0.1 & 0 \\ 0 & 0 & -0.101 \end{pmatrix}, \quad \bar{B} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \quad \bar{C} = \begin{pmatrix} 1 & 1 & 1 \end{pmatrix}.$$

d)

$$G(s) = \frac{(s+2)(s+4)(s+6)(s+8)}{(s+1)(s+3)(s+5)(s+7)}.$$

Relate your choices of truncated states to the error bound (1.6). Plot Bode diagrams of the systems G and the approximations G_r you construct.

EXERCISE 1.3

What are the necessary and sufficient conditions for controllability and observability of the states in a modal representation of G?

EXERCISE 1.4

Prove properties P1 and P2.