Lecture 14

Controller Simplification

In the previous lecture we used convex optimization to solve optimal control design problems. The resulting controller could have very high order. In this short lecture we describe the use of balanced truncation for order reduction of the controller.

14.1 Model reduction by balanced truncation

Model reduction is a general problem in modeling and design of dynamical systems. Accurate physical modeling of processes can often lead to state-space equations with hundreds or thousands of states. Controller synthesis using the Youla parametrization and convex optimization can lead to very high order controllers. Hence there is a need for systematic way to reduce the model order of a given linear system. In general terms we would like to achieve

$$G_r(s) \approx G(s),$$

where the reduced system $G_r(s)$ has (much) lower order than the original system $G(s)$. The closeness between the systems can be measured in various ways. A common criterion is to measure the maximum deviation between the frequency responses of the systems, i.e.,

$$\|G_r - G\|_\infty.$$

Model reduction has been an active research area within control theory for the past few decades. The optimal model reduction problem is in general difficult (non-convex). Here we will briefly introduce the simple, popular technique known as *balanced truncation*. The basic idea is to remove states (or *modes*) that are both poorly controllable and poorly observable.

Hankel singular values

Starting from a high-order stable system,

$$\begin{aligned} \dot{x} &= Ax + Bu \\ y &= Cx + Du \end{aligned}$$

recall that the controllability and observability Gramians

$$W_c = \int_0^\infty e^{At} B B^T e^{A^T t} dt, \qquad W_o = \int_0^\infty e^{A^T t} C^T C e^{At} dt$$

can be found by solving the linear equations

$$\begin{aligned} AW_c + W_c A^T + BB^T &= 0, \\ A^T W_o + W_o A + C^T C &= 0. \end{aligned}$$

The *Hankel singular values* are defined as the square roots of the eigenvalues of $W_c W_o$:

$$\sigma_i = \sqrt{\lambda_i(W_c W_o)}$$

They measure the "energy" of each mode in the system and are usually ordered such that

$$\sigma_1 \geq \sigma_2 \geq \ldots \geq \sigma_n > 0.$$

A small Hankel singular value indicates a weakly controllable and observable state that is a candidate for elimination. We will see later that the Hankel singular values are independent of the realization.

EXAMPLE 14.1
Given the 6th order linear system

$$G(s) = \frac{1 - s}{s^6 + 3s^5 + 5s^4 + 7s^3 + 5s^2 + 3s + 1},$$

we introduce any state realization and calculate the Hankel singular values using Matlab. In order of size they are given by

$$\sigma = \begin{bmatrix} 1.984 & 1.918 & 0.751 & 0.329 & 0.148 & 0.004 \end{bmatrix}$$

We can see that the 6th Hankel singular value is very small compared to the others. □

Balanced realizations

Next we will performance a coordinate transformation that reveals the Hankel singular values. Given a stable system (A, B, C, D) with Gramians W_c and W_o, the variable transformation $\hat{x} = Tx$ gives the new state-space matrices

$$\hat{A} = TAT^{-1}, \quad \hat{B} = TB, \quad \hat{C} = CT^{-1}, \quad \hat{D} = D$$

and the new Gramians

$$\hat{W}_c = \int_0^\infty e^{\hat{A}t}\hat{B}\hat{B}^T e^{\hat{A}^T t} dt = \int_0^\infty Te^{At}BB^T e^{A^T t}T^T dt = TW_cT^T$$
$$\hat{W}_o = \int_0^\infty e^{\hat{A}^T t}\hat{C}^T\hat{C}e^{\hat{A}t} dt = \int_0^\infty T^{-T}e^{At}C^T Ce^{A^T t}T^{-1} dt = T^{-T}W_oT^{-1}$$

A particular choice of T gives

$$\hat{W}_c = \hat{W}_o = \Sigma = \begin{bmatrix} \sigma_1 & & 0 \\ & \ddots & \\ 0 & & \sigma_n \end{bmatrix},$$

i.e., the controllability and observability Gramians coincide and contain the Hankel singular values. The corresponding realization $(\hat{A}, \hat{B}, \hat{C}, \hat{D})$ is called a *balanced realization*.

The transformation matrix T can be found by the following calculations (not suitable for hand calculations). First compute the Cholesky decompositions

$$W_c = WW^T, \quad W_o = ZZ^T$$

and the singular value decomposition

$$W^T Z = U\Sigma V^T.$$

The balancing transformation is then given by

$$T = \Sigma^{-\frac{1}{2}}V^T Z^T, \quad T^{-1} = WU\Sigma^{-\frac{1}{2}}.$$

Notice that

$$\begin{bmatrix} \sigma_1^2 & & 0 \\ & \ddots & \\ 0 & & \sigma_n^2 \end{bmatrix} = \underbrace{(TW_cT^T)}_{\Sigma}\underbrace{(T^{-T}W_oT^{-1})}_{\Sigma} = TW_cW_oT^{-1}$$

which means that the Hankel singular values are independent of the coordinate system.

Balanced truncation

A small Hankel singular value σ_i corresponds to a state that is both weakly controllable and weakly observable. Hence, it can be truncated without much effect on the input-output behavior.

Consider a balanced realization

$$\begin{bmatrix} \dot{\hat{x}}_1 \\ \dot{\hat{x}}_2 \end{bmatrix} = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \begin{bmatrix} \hat{x}_1 \\ \hat{x}_2 \end{bmatrix} + \begin{bmatrix} B_1 \\ B_2 \end{bmatrix} u \qquad\qquad \Sigma = \begin{bmatrix} \Sigma_1 & 0 \\ 0 & \Sigma_2 \end{bmatrix}$$

$$y = \begin{bmatrix} C_1 & C_2 \end{bmatrix} \begin{bmatrix} \hat{x}_1 \\ \hat{x}_2 \end{bmatrix} + Du$$

with the lower part of the Gramian being $\Sigma_2 = \operatorname{diag}(\sigma_{r+1}, \ldots, \sigma_n)$. There are two standard ways to perform the reduction:

1. Simply remove $\hat{x}_2$ and keep (A_{11}, B_1, C_1, D). This method preserves the high-frequency $(\omega \to \infty)$ gain of the system.
2. (More commonly:) Set $\dot{\hat{x}}_2 = 0$, which in turn implies $0 = A_{21}\widehat{x}_1 + A_{22}\widehat{x}_2 + B_2 u$. This method preserves the DC gain of the system and yields the reduced system

$$\begin{cases} \dot{\hat{x}}_1 = (A_{11} - A_{12}A_{22}^{-1}A_{21})\hat{x}_1 + (B_1 - A_{12}A_{22}^{-1}B_2)u \\ y_r = (C_1 - C_2A_{22}^{-1}A_{21})\hat{x}_1 + (D - C_2A_{22}^{-1}B_2)u \end{cases}$$

Error bounds for balanced truncation

One way to measure the approximation error between the original system $G(s)$ and the reduced system $G_r(s)$ is

$$\|G - G_r\|_\infty = \max_\omega |G(i\omega) - G_r(i\omega)| = \sup_u \frac{\|y - y_r\|_2}{\|u\|_2}$$

For either of the truncation methods above, it holds that

$$\sigma_{r+1} \;\le\; \|G - G_r\|_\infty \;\le\; 2(\sigma_{r+1} + \cdots + \sigma_n)$$

Example 14.2

Again consider the system

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

Keeping $r = 3$ states and applying balanced truncation gives the reduced system (using the method of retaining the DC gain):

$$G_r(s) = \frac{0.3717s^3 - 0.9682s^2 + 1.14s - 0.5185}{s^3 + 1.136s^2 + 0.825s + 0.5185}$$

The error bounds for the approximation can be calculated as

$$0.329 \le \|G - G_r\|_\infty \le 0.963.$$

Checking the actual error, one obtains $\|G - G_r\|_\infty = 0.573$. The difference between the original and the reduced system can be studied in Figure 14.1. □

For unstable systems, the method above can be applied by first decomposing the system into its stable and non-stable parts:

$$G(s) = G_s(s) + G_{ns}(s)$$

Then the model reduction is performed only on $G_s(s)$; afterwards $G_{ns}(s)$ is added again.

(a)

(b)

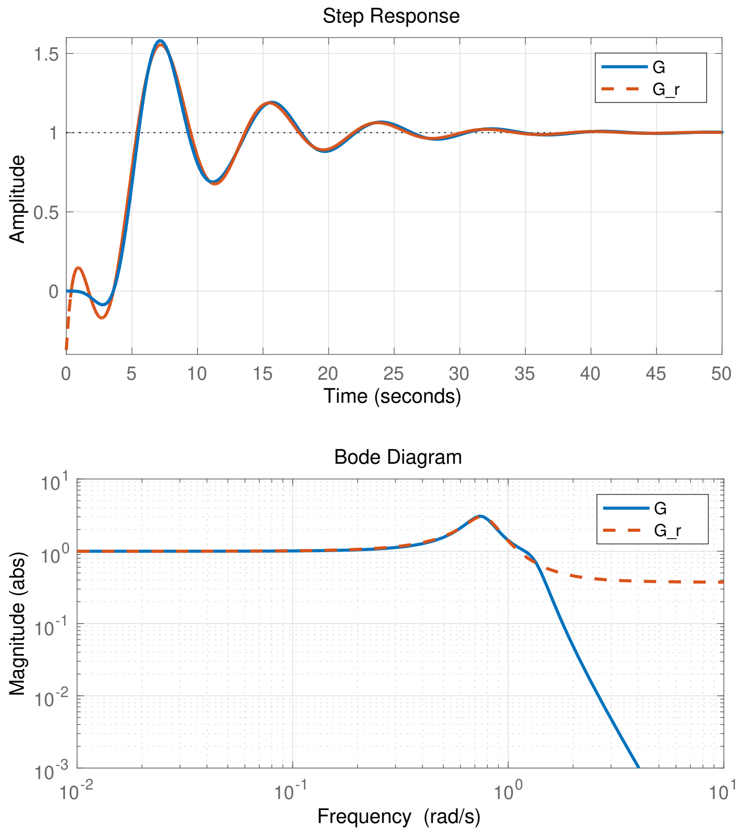

Figure 14.1 Comparison between 6th order original system G and 3rd order reduced system G_r: (a) step responses, (b) Bode magnitude diagrams.

14.2 Frequency-weighted balanced truncation

There are many extensions to balanced truncation. One common modification is to introduce frequency weighting, such that more emphasis is placed on making a good approximation of the system in a particular frequency range. For controllers, the cross-over frequency is often of particular interest.

As we saw above, the error bound

$$\max_{\omega} |G(i\omega) - G_r(i\omega)| = \sup_{u} \frac{\|y - y_r\|_2}{\|u\|_2} \leq 2\sigma_{r+1} + \cdots + 2\sigma_n$$

emphasizes all frequencies equally, but comparing a controller $C(s)$ with a reduced controller $C_r(s)$ in closed loop operation gives

$$|P(I + CP)^{-1}C - P(I + C_r P)^{-1}C_r| \approx |P(I + CP)^{-1}(C - C_r)|$$

Hence it is interesting to minimize the frequency weighted error

$$\max_{\omega} \Big| W(i\omega)[C(i\omega) - C_r(i\omega)] \Big|$$

with $W(i\omega) = P(i\omega)(I + C(i\omega)P(i\omega))^{-1}$.

For a general stable MIMO system we can aim to minimize

$$\max_{\omega} \Big\| W_o(i\omega)[G(i\omega) - G_r(i\omega)]W_i(i\omega) \Big\|$$

where

$$\begin{aligned} W_i(s) &= C_i(sI - A_i)^{-1}B_i + D_i \\ G(s) &= C(sI - A)^{-1}B + D \\ W_o(s) &= C_o(sI - A_o)^{-1}B_o + D_o \end{aligned}$$

We next find extended controllability and observabiliry Gramians S and O by solving

$$\begin{bmatrix} A & BC_i \\ 0 & A_i \end{bmatrix} \begin{bmatrix} S & S_{12} \\ S_{12}^T & S_{22} \end{bmatrix} + \begin{bmatrix} S & S_{12} \\ S_{12}^T & S_{22} \end{bmatrix} \begin{bmatrix} A & BC_i \\ 0 & A_i \end{bmatrix}^T + \begin{bmatrix} BD_i \\ B_i \end{bmatrix} \begin{bmatrix} BD_i \\ B_i \end{bmatrix}^T = 0$$

$$\begin{bmatrix} A & 0 \\ B_oC & A_o \end{bmatrix}^T \begin{bmatrix} O & O_{12} \\ O_{12}^T & O_{22} \end{bmatrix} + \begin{bmatrix} O & O_{12} \\ O_{12}^T & O_{22} \end{bmatrix} \begin{bmatrix} A & 0 \\ B_oC & A_o \end{bmatrix} + \begin{bmatrix} C^T D_o^T \\ D_o^T \end{bmatrix} \begin{bmatrix} D_oC & D_o \end{bmatrix} = 0$$

then change coordinates to make S and O equal and diagonal before truncating the realization of $G(s)$ to get $G_r(s)$ as before.