Model order reduction by balance truncation

Martin Ambrožič * Manuel Baumann * Carlos Echeverría Serur *

November 26, 2012

Abstract

The purpose of this report is to present some clarifying remarks together with proofs of some relevant theorems that were not dealt with in the presentation but that were studied for a deeper understanding of the balanced truncation method. In this scheme, the states which require a large amount of energy and/or yield small amounts of observation energy are eliminated. This process requires a phase of balancing, after which the observability and controllability Gramians are diagonal and identical. Thereby, states that are *hard to reach* as well as states that are *hard to observe* are easily identified. Following is a phase of truncation, in which states satisfying both properties are eliminated. An error bound for the truncated system can be derived. It is found that the reduction preserves the stability of the system.

In Section 1, we identify the suitable states for the reduction of the model in terms of a relation of the controllability and observability Gramians. In Section 2, we present a reduction of the system which eliminates the uncontrollable and unobservable states without the introduction of an error. In Section 3 we give a detailed proof of the preservation of stability of the reduced system. The section consists of a detailed work-out of the proof of Theorem 7.7 found in [Ant05].

Keywords: Model order reduction, Balanced truncation, controllability and observability Gramians, Hankel singular values.

1 Characterization of suitable states for elimination

In this section, we show two relations that are used in the balanced truncation algorithm in order to identify those states that are *hard to reach* and those which are *hard to observe*. In our presentation we point out that because both relations only depend respectively on the controllability and observability Gramians, they can be used to eliminate those states that are at the same time hard to reach and hard to observe.

Theorem: Let $A \in \mathbb{R}^{n,n}$, $t_f > 0$, and $x_f \in \mathbb{R}^n$ be a reachable state of the system $\dot{x} = Ax(t) + Bu(t)$ with controllability Gramian

$$P(t_f) := \int_0^{t_f} e^{At} BB^T e^{A^T t} dt. \tag{1}$$

*Delft University of Technology. {mambrozic,mbaumann,cecheverriaser}@tudelft.nl

Let $u(\cdot) \in L_2([0,t_f],\mathbb{R}^m)$ with

$$\hat{u}(t) = B^T e^{A^T(t_f-t)} P^{-1}(t_f) x_f. \tag{2}$$

Then the following holds true

(a) $\|\hat{u}\|^2 = x_f^T P(t_f)^{-1} x_f$.

(b) For all $u(\cdot) \in L_2([0,t_f],\mathbb{R}^m)$ which control the system from $x(0) = x_0$ to $x(t_f) = x_f$, there holds $\|\hat{u}\|_{L_2} \leq \|u\|_{L_2}$.

Proof. (a) Let $\mathcal{B}_{t_f} : L_2([0,t_f],\mathbb{R}^m) \to \mathbb{R}^n$ be the reachability map of $\dot{x} = Ax(t) + Bu(t)$, given by:

$$\mathcal{B}_{t_f} : u(\cdot) \mapsto x(t_f) = \int_0^{t_f} e^{A(t_f-\tau)} Bu(\tau) d\tau. \tag{3}$$

Then,

$$\hat{u}(t) = \mathcal{B}_{t_f}^* P(t_f)^{-1} x_f,$$

with the adjoint $\mathcal{B}_{t_f}^*$ derived in Appendix A.

It follows:

$$\begin{aligned}
\|\hat{u}\|^2 &= \left\langle \mathcal{B}_{t_f}^* P(t_f)^{-1} x_f, \mathcal{B}_{t_f}^* P(t_f)^{-1} x_f \right\rangle_{L_2} \\
&= \left\langle P(t_f)^{-1} x_f, \mathcal{B}_{tf} \mathcal{B}_{tf}^* P(t_f)^{-1} x_f \right\rangle_{\mathbb{R}^n} \\
&= \left\langle P(t_f)^{-1} x_f, P(t_f) P(t_f)^{-1} x_f \right\rangle_{\mathbb{R}^n} \\
&= \left\langle P(t_f)^{-1} x_f, x_f \right\rangle_{\mathbb{R}^n} = x_f^T P(t_f)^{-1} x_f
\end{aligned}$$

(b) In order to show that $\hat{u}$ defined in (2) is the optimal input that drives the system from $x_0 = 0$ to $x(t_f) = x_{t_f}$, we consider

$$u(\cdot) := \hat{u}(\cdot) + v(\cdot) \in \mathcal{B}_{t_f}, \quad \text{since we require } v(\cdot) \in ker\mathcal{B}_{t_f}.$$

Then

$$\begin{aligned}
\|u\|_{L_2}^2 &= \|\hat{u} + v\|_{L_2}^2 \\
&= \|\hat{u}\|_{L_2}^2 + 2\langle \hat{u}, v\rangle_{L_2} + \|v\|_{L_2}^2 \\
&= \|\hat{u}\|_{L_2}^2 + 2\left\langle \mathcal{B}_{t_f}^* P(t_f)^{-1} x_f, v \right\rangle_{L_2} + \|v\|_{L_2}^2 \\
&= \|\hat{u}\|_{L_2}^2 + 2\left\langle P(t_f)^{-1} x_f, \underbrace{\mathcal{B}_{t_f} v}_{=0} \right\rangle_{L_2} + \|v\|_{L_2}^2 \\
&= \|\hat{u}\|_{L_2}^2 + \|\hat{v}\|_{L_2}^2 \geq \|\hat{u}\|_{L_2}^2
\end{aligned}$$

□

Theorem: Consider the zero-input system $\dot{x}(t) = Ax(t), x(0) = x_0, y(t) = Cx(t)$ for $t_f > 0, x_0 \in \mathbb{R}^n$ with the observability Gramian

$$Q(t_f) = \int_0^{t_f} e^{A^T t} C^T C e^{At} \, dt. \tag{4}$$

Then it holds

$$x_0^T Q(t_f) x_0 = \|y(\cdot)\|_{L_2}^2$$

Proof.

$$x_0^T Q(t_f) x_0 = \int_0^{t_f} x_0^T e^{A^T t} C^T C e^{At} x_0 \; dt = \int_0^{t_f} \|C e^{At} x_0\|^2 \; dt = \|y(\cdot)\|_{L_2}^2$$

□

2 Errorless model order reduction

Now, we present that uncontrollable and unobservable states can be eliminated without any error. In our presentation, we have shown that there exists a transformation S, such that (A, B, C, D) is equivalent to $(S^{-1}AS, S^{-1}B, CS, D)$, where

$$S^{-1}AS = \begin{bmatrix} A_{11} & 0 & A_{13} & 0 \\ A_{21} & A_{22} & A_{23} & A_{24} \\ 0 & 0 & A_{33} & 0 \\ 0 & 0 & A_{43} & A_{44} \end{bmatrix}, \quad S^{-1}B = \begin{bmatrix} B_1 \\ B_2 \\ 0 \\ 0 \end{bmatrix}, \quad CS = \begin{bmatrix} C_1 & 0 & C_3 & 0 \end{bmatrix}$$

In particular, $[A_{11}, B_1, C_1, D]$ is both controllable and observable. Determining the transfer function we obtain:

$$\begin{aligned}
G(s) &= C(sI - A)^{-1}B + D \\
&= \begin{bmatrix} C_1 & 0 & C_3 & 0 \end{bmatrix} \begin{bmatrix} sI - A_{11} & 0 & -A_{13} & 0 \\ -A_{21} & sI - A_{22} & -A_{23} & -A_{24} \\ 0 & 0 & sI - A_{33} & 0 \\ 0 & 0 & -A_{43} & sI - A_{44} \end{bmatrix}^{-1} \begin{bmatrix} B_1 \\ B_2 \\ 0 \\ 0 \end{bmatrix} + D \\
&= \begin{bmatrix} C_1 & 0 & C_3 & 0 \end{bmatrix} \begin{bmatrix} \begin{bmatrix} sI - A_{11} & 0 \\ -A_{13} & sI - A_{22} \end{bmatrix}^{-1} & \star \\ 0 & \star \end{bmatrix} \begin{bmatrix} B_1 \\ B_2 \\ 0 \\ 0 \end{bmatrix} + D \\
&= \begin{bmatrix} C_1 & 0 \end{bmatrix} \begin{bmatrix} sI - A_{11} & 0 \\ -A_{13} & sI - A_{22} \end{bmatrix}^{-1} \begin{bmatrix} B_1 \\ B_2 \end{bmatrix} + D \\
&= \begin{bmatrix} C_1 & 0 \end{bmatrix} \begin{bmatrix} (sI - A_{11})^{-1} & 0 \\ \star & \star \end{bmatrix} \begin{bmatrix} B_1 \\ B_2 \end{bmatrix} + D \\
&= C_1(sI - A_{11})^{-1}B_1 + D.
\end{aligned}$$

Hence, the elimination of uncontrollable and unobservable states preserves the transfer function. It can therefore be considered as errorless model order reduction.

3 Preservation of stability

In our presentation, the following theorem which guarantees that stability is preserved by balanced truncation has been presented.

Theorem: Given a stable full order system (A, B, C, D), then the reduced system $(\tilde{A}, \tilde{B}, \tilde{C}, \tilde{D})$ obtained by balanced truncation has the following properties:

(a) $\tilde{A}$ has no eigenvalues in the open right half plane, i.e. the reduced system is stable.

(b) If, additionally, $\sigma_i \neq \sigma_j$, for $i = 1, ..., r, j = r + 1, ..., n$, then $\tilde{A}$ has no eigenvalues on $i\mathbb{R}$, i.e. the reduced system is asymptotically stable.

Proof. (a) In [Ant05], **Lemma 6.15** it is proved that for $AP + PA^* = Q$, with $Q \geq 0$ the following holds true

$$\text{If P has no eigenvalues on } i \cdot \mathbb{R}, \text{ then } \pi(A) \leq \pi(P),$$

where $\pi(\cdot)$ denotes the number of eigenvalues in the open right half plane of the considered matrix.

By the balanced truncation algorithm, the following Lyapunov equation holds

$$\begin{aligned} \tilde{A}\tilde{\Sigma} + \tilde{\Sigma}\tilde{A}^* &= -\tilde{B}\tilde{B}^* \\ \Leftrightarrow \tilde{A}(-\tilde{\Sigma}) + (-\tilde{\Sigma})\tilde{A}^* &= \tilde{B}\tilde{B}^*, \end{aligned}$$

with $\tilde{B}\tilde{B}^* \geq 0$. Since $\tilde{\Sigma}$ is positive definite, we conclude $-\tilde{\Sigma}$ is negative definite. Therefore, no eigenvalue of $-\tilde{\Sigma}$ lies in the open right half plane and by **Lemma 6.15**, the same holds for the eigenvalues of $\tilde{A}$.

(b) Let $P = Q = \begin{bmatrix} \tilde{\Sigma} & \\ & \Sigma_{22} \end{bmatrix}$ be the Gramians of the full order system after balancing, i.e. $\tilde{\Sigma}$ and Σ_{22} are diagonal and $\tilde{\Sigma}$ is the Gramian of the reduced system. We have to show that if $\tilde{\Sigma}$ and Σ_{22} have no diagonal elements in common, then $\tilde{A}$ has no eigenvalues on the imaginary axis.

Let us assume the contrary, i.e. let $\tilde{A}$ have eigenvalues on the imaginary axis. Let us further assume that $\lambda = i\omega$ is the only eigenvalue on the imaginary axis and $\sigma_1 = 1$ is the largest Hankel singular value with multiplicity equal to 1. Let v be the corresponding eigenvector,

$$\tilde{A}v = \lambda v.$$

Then, it follows

$$v^*\tilde{A}^* = \lambda^* v^* = -\lambda v^*.$$

Consider the Lyapunov equation

$$\tilde{A}^*\tilde{\Sigma} + \tilde{\Sigma}\tilde{A} + \tilde{C}^*\tilde{C} = 0 \tag{5}$$

and multiply (5) from the left with v^* and from the right with v:

$$\begin{aligned} v^*\tilde{A}^*\tilde{\Sigma}v + v^*\tilde{\Sigma}\tilde{A}v + v^*\tilde{C}^*\tilde{C}v &= 0 \\ -\lambda v^*\tilde{\Sigma}v + \lambda v^*\tilde{\Sigma}v + \|\tilde{C}v\| &= 0 \\ \Rightarrow \tilde{C}v &= 0 \end{aligned}$$

Multiplying (5) only from the right with v and using $\tilde{C}v = 0$,we derive

$$\begin{aligned} \tilde{A}^*\tilde{\Sigma}v + \tilde{\Sigma}\tilde{A}v + \tilde{C}^*\tilde{C}v &= 0 \\ \tilde{A}^*\tilde{\Sigma}v + \tilde{\Sigma}\tilde{A}v &= 0 \\ (\tilde{A}^* + \lambda I)\tilde{\Sigma}v &= 0 \end{aligned}$$

Consider the Lyapunov equation

$$\tilde{A}\tilde{\Sigma} + \tilde{\Sigma}\tilde{A}^* + \tilde{B}\tilde{B}^* = 0, \tag{6}$$

together with $(\tilde{A}^* + \lambda I)\tilde{\Sigma}v = 0$. Then, a left multiplication of (6) with $v^*\tilde{\Sigma}$ and a right multiplication with $\tilde{\Sigma}v$ leads to:

$$\underbrace{v^*\tilde{\Sigma}\tilde{A}}_{-\lambda v^*\tilde{\Sigma}}\tilde{\Sigma}\tilde{\Sigma}v + v^*\tilde{\Sigma}\tilde{\Sigma}\underbrace{\tilde{A}^*\tilde{\Sigma}v}_{\lambda\tilde{\Sigma}v} + v^*\tilde{\Sigma}\tilde{B}\tilde{B}^*\tilde{\Sigma}v = 0 \;\Rightarrow\; \tilde{B}^*\tilde{\Sigma}v = 0.$$

Furthermore, right multiplication of (6) with $\tilde{\Sigma}v$ leads to:

$$(\tilde{A} - \lambda I)\tilde{\Sigma}^2 v = 0. \tag{7}$$

Since (7) is an eigenequation of the eigenvalue λ, we conclude that $\Sigma^2 v$ is a multiple of v, i.e.

$$\Sigma^2 v = \alpha v,$$

for some real α. Since this is another eigenequation, now for the pair (v, α), we can conclude due to the special structure of the matrix $\tilde{\Sigma}^2$ that v can be taken as $v = [1, 0, ..., 0]^T$.

Next, we consider the full-order Lyapunov equation

$$\begin{bmatrix} \tilde{A} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \begin{bmatrix} \tilde{\Sigma} & \\ & \Sigma_{22} \end{bmatrix} + \begin{bmatrix} \tilde{\Sigma} & \\ & \Sigma_{22} \end{bmatrix} \begin{bmatrix} \tilde{A}^* & A_{21}^* \\ A_{12}^* & A_{22}^* \end{bmatrix} + \begin{bmatrix} B_1 \\ B_2 \end{bmatrix} \begin{bmatrix} B_1^* & B_2^* \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$

and obtain the next equation from the $(1, 2)$ position:

$$A_{21}\tilde{\Sigma} + \Sigma_{22}A_{12}^* + B_2B_1^* = 0. \tag{8}$$

Similary, we derive

$$\Sigma_{22}A_{21} + A_{12}^*\tilde{\Sigma} + C_2^*C_1 = 0. \tag{9}$$

Let us denote the first column of A_{21}, A_{12}^* by a, b, respectively. Multiplying (8)-(9) from the right by $v = [1, 0, ..., 0]^T$, we obtain

$$a + \Sigma_{22}b = 0, \quad \Sigma_{22}a + b = 0. \tag{10}$$

Since, by assumption, $\tilde{\Sigma}$ and Σ_{22} don't have common diagonal entries, the following can be deduced from (10)

$$\left\{ \begin{array}{l} \Sigma_{22}^{-1}a + b = 0 \\ \Sigma_{22}a + b = 0 \end{array} \right\} \;\Rightarrow\; (\Sigma_{22}^{-1} - \Sigma_{22})a = 0 \;\Rightarrow\; a = 0,$$

where the last implication holds true because there are no ones on the diagonal of Σ_{22} and, hence, also Σ_{22}^{-1}. Therefore, the matrix $(\Sigma_{22}^{-1} - \Sigma_{22})$ has full rank.

As a consequence, we have derived that $[v^*, 0]^* \in \mathbb{R}^n$ is an eigenvector of the whole matrix A corresponding to the eigenvalue $\lambda = i\omega$. This is, however, a contradiction to the reachability of (A, B). Therefore, $\tilde{A}$ can not have eigenvalues on the imaginary axis. □

4 Concluding remarks on the project work

This report contains complementary information to the presentation on *Model order reduction by balanced truncation* for the MASTERMATH course in Systems and Control. It consists of a set of theorems, proofs and concepts, which, for pedagogic reasons, were not presented in detail in the presentation but had much importance in the preparatory work. The idea of approximating the transfer functions in the $\mathcal{H}_\infty$-norm comes from the breakthrough paper by Glover, [Glo84]. The study of the characterization of suitable states for model order reduction as well as the errorless reduction of uncontrollable and unobservable states was adapted from [Rei12]. The proof presented in the Section 4 on the preserving of stability is a detailed expansion and explanation of that found in [Ant05]. Many asseverations were used without a proof and were used as background knowledge found on [PW98]. The balanced truncation algorithm was tested with a numerical example, inspired by [Rei12].

References

[Ant05] Athanasios C. Antoulas. *Approximation of Large-Scale Dynamical Systems (Advances in Design and Control).* Society for Industrial and Applied Mathematics, Philadelphia, PA, USA, 2005.

[Glo84] Keith Glover. All optimal Hankel-norm approximations of linear multivariable systems and their L^∞-error bounds. *Int. J. Control*, 39:1115–1193, 1984.

[PW98] J.W. Polderman and J.C. Willems. *Introduction to Mathematical Systems Theory: A Behavioral Approach.* Texts in Applied Mathematics Series. Springer-Verlag, 1998.

[Rei12] Timo Reis. Model order reduction. *Lecture notes, University Hamburg*, 2012.

A The adjoint of $\mathcal{B}_{t_f}$

Consider the reachability map

$$\begin{aligned}\mathcal{B}_{t_f} : L_2([0,t_f],\mathbb{R}^m) &\to \mathbb{R}^n \\ u(\cdot) \mapsto x(t_f) &= \int_0^{t_f} e^{A(t_f-\tau)}Bu(\tau)\ d\tau)\end{aligned}$$

Then, the corresponding adjoint operator maps from $\mathbb{R}^n$ to $L_2([0,t_f],\mathbb{R}^m)$ and is given by

$$\begin{aligned}\mathcal{B}^*_{t_f} : \mathbb{R}^n &\to L_2([0,t_f],\mathbb{R}^m) \\ x &\mapsto B^T e^{A^T(t_f-\cdot)}x(\cdot)\end{aligned}$$

Proof.

$$\begin{aligned}\left\langle x, \mathcal{B}_{t_f}u(\cdot)\right\rangle_{\mathbb{R}^n} &= x^T\int_0^{t_f} e^{A(t_f-\tau)}Bu(\tau)d\tau \\ &= \int_0^{t_f} x^T e^{A(t_f-\tau)}Bu(\tau)d\tau \\ &= \int_0^{t_f} (B^T e^{A^T(t_f-\tau)}x)^T u(\tau)d\tau \\ &= \left\langle B^T e^{A^T(t_f-\cdot)}x, u(\cdot)\right\rangle_{L_2}\end{aligned}$$

□

Note, that the following relation holds true:

$$\begin{aligned}\mathcal{B}_{t_f}\mathcal{B}^*_{t_f}x &= \int_0^{t_f} e^{A(t_f-\tau)}B(\mathcal{B}^*_{t_f}x)(\tau)\ d\tau \\ &= \int_0^{t_f} e^{A(t_f-\tau)}B(B^T e^{A^T(t_f-\tau)}x)\ d\tau \\ &= \int_0^{t_f} e^{A(t_f-\tau)}BB^T e^{A^T(t_f-\tau)}\ d\tau\ x \\ &= \int_0^{t_f} e^{A\tau}BB^T e^{A^T\tau}\ d\tau\ x = P(t_f)x,\end{aligned}$$

where $P(t_f)$ is the controllability Gramian defined by (1).