Matrix Equations and Model Reduction

Peter Benner
Max Planck Institute for Dynamics of Complex Technical Systems
and
Otto-von-Guericke Universität Magdeburg, Faculty of Mathematics

benner@mpi-magdeburg.mpg.de

July 28, 2013

Contents

Chapter 1

Introduction

Chapter 2

Mathematical Basics

2.1 Numerical Linear Algebra

The Singular Value Decomposition

One essential tool from (numerical) linear algebra for data compression and dimension reduction is the singular value decomposition (SVD) of a matrix. It is related to the *four fundamental subspaces* of a matrix. These are the *column space* (or *range*), the *row space*, the *nullspace* (or *kernel*), and the *cokernel*. The SVD exists for any matrix as the following theorem shows.

Theorem 2.1 *Let $A \in \mathbb{R}^{m\times n}$, then there exist orthogonal $U \in \mathbb{R}^{m\times m}$ and $V \in \mathbb{R}^{n\times n}$, such that*

$$A = U\Sigma V^T,\ \Sigma = \begin{cases} \begin{bmatrix} \Sigma_1 \\ 0 \end{bmatrix}, & m \geq n \\ \begin{bmatrix} \Sigma_1 & 0 \end{bmatrix}, & m \leq n \end{cases} \quad \textit{and} \quad \Sigma_1 = \begin{bmatrix} \sigma_1 & & \\ & \ddots & \\ & & \sigma_{\min(m,n)} \end{bmatrix} \tag{2.1}$$

with

$$\sigma_1 \geq \ldots \sigma_s > \sigma_{s+1} = \ldots = \sigma_{\min(m,n)} = 0 \quad \textit{for} \quad s = \operatorname{rank}(A).$$

Proof: Let $x \in \mathbb{R}^n, y \in \mathbb{R}^m$ satisfy $Ax = \sigma_1 y$, $\sigma_1 = \|A\|_2$ and $\|x\|_2 = 1 = \|y\|_2$ ($\exists$, since $\{\|x\|_2 = 1\}$ compact). Complete x and y to orthonormal bases (ONB) of $\mathbb{R}^n$ and $\mathbb{R}^m$, respectively, to obtain orthogonal $U \in \mathbb{R}^{m\times m}$ with $U(:,1) = y$ and $V \in \mathbb{R}^{n\times n}$ with $V(:,1) = x$. It follows

$$U^T A V = \begin{bmatrix} \sigma_1 & w^T \\ 0 & B \end{bmatrix} =: A_1, \quad (V = [x.v_2, \ldots, v_n],\ U = [y, u_2, \ldots, u_m]).$$

Now it holds that

$$\|A_1 \begin{bmatrix} \sigma_1 \\ w \end{bmatrix}\|_2^2 = \|\begin{bmatrix} \sigma_1^2 + w^T w \\ Bw \end{bmatrix}\|_2^2 \geq (\sigma_1^2 + w^T w)^2.$$

As $\| \begin{bmatrix} \sigma_1 \\ w \end{bmatrix} \|_2^2 = \sigma_1^2 + w^T w$ we get

$$\|A_1\|_2^2 \geq \sigma_1^2 + w^T w = \|A\|_2^2 + w^T w = \|A_1\|_2^2 + w^T w$$

and hence $w = 0$. The proof follows inductively by applying the same arguments to B.

(For the rank argument observe that "$\sigma_1 = 0 \Leftrightarrow A = 0$" in the above argument and $\|B\|_2 \leq \|A\|_2$ if B is a submatrix of $A \Rightarrow \sigma_j \geq \sigma_{j+1}$ and $\sigma_{s+1} = 0$.) □

Definition 2.2 *Let $A \in \mathbb{R}^{m\times n}$. The factorization (2.1) is the* singular value decomposition (SVD) of A, *the nonnegative numbers σ_j are its* singular values, *and the columns of U and V are the corresponding* left *and* right singular vectors, *respectively.*

We have the following important observation relating the four fundamental subspaces to the SVD:

Corollary 2.3 *For $A \in \mathbb{R}^{m\times n}$ we have*

a) $\text{colspace}(A) \equiv \text{range}\,(A) = \text{span}\{u_1, \ldots, u_s\}$,

b) $\text{rowspace}(A) \equiv \text{range}\left(A^T\right) = \text{span}\{v_1, \ldots, v_s\}$,

c) $\text{null}(A) \equiv \ker(A) = \text{span}\{v_{s+1}, \ldots, v_n\}$,

d) $\text{cokernel}(A) \equiv \text{null}(A^T) = \text{span}\{u_{s+1}, \ldots, u_m\}$.

Note that once this result is established, it is trivial to prove the dimension/rank-nullity theorems:

$$\text{null}(A) \oplus \text{range}\left(A^T\right) = \mathbb{R}^n, \tag{2.2a}$$
$$\text{range}\,(A) \oplus \text{null}(A^T) = \mathbb{R}^m. \tag{2.2b}$$

Before coming to the key observation which makes the SVD so important in dimension reduction, we also recall that the SVD is related to matrix norms.

Corollary 2.4 *Let $A \in \mathbb{R}^{m\times n}$ have an SVD as in* (2.1). *Then*

a) $\|A\|_F = \sqrt{\sigma_1^2 + \ldots + \sigma_s^2}$,

b) $\|A\|_2 = \sigma_1$ $(= \sigma_{\max} := \max\{\sigma \geq 0, \sigma \text{ singular value of } A\})$.

c) If $s = \min\{m, n\}$, then $\text{cond}_2\,(A) = \frac{\sigma_1}{\sigma_s}$.

d) If $m = n$ and A is invertible, then $\|A^{-1}\|_2 = \sigma_n$.

Also note that $\sigma_{\min} := \min\{\sigma > 0, \sigma \text{ singular value of } A\}$ is the size (measured in the spectral norm) of the smallest perturbation to make a non-singular (or, more generally, full-rank) A a singular (or rank-deficient) matrix. A proof of this fact can be deduced from the following, more general statement, that will be the basis or at least the motivation for a number of model reduction algorithms in the following.

Theorem 2.5 (Schmidt-Mirsky/Eckart-Young Theorem[1]) *Let $A \in \mathbb{R}^{m\times n}$ have an SVD as in (2.1). For $k \leq \operatorname{rank}(A)$, the best rank-k approximation to A in the spectral norm is given by*

$$A_k := \sum_{j=1}^{k} \sigma_j u_j v_j^T, \tag{2.3}$$

and the approximation error is

$$\|A - A_k\|_2 = \min_{\operatorname{rank}(B)=k} \|A - B\|_2 = \sigma_{k+1} \quad \text{for} \quad k \leq \operatorname{rank}(A). \tag{2.4}$$

Proof: See Exercise 1.4. □

A particular instance of the above theorem is that A can be exactly represented as

$$A = A_s = \sum_{j=1}^{s} \sigma_j u_j v_j^T, \qquad s = \operatorname{rank}(A),$$

In a "thin" SVD this can employed, by, e.g., computing only $U_1 = [\, u_1, \ldots, u_s \,]$ and/or $V_1 = [\, v_1, \ldots, v_s \,]$, depending on which of the four subspaces is required, or whether only the "minimal" representation of A through its singular vectors is required.

Note that the above theorem also holds if the spectral norm is replaced by the Frobenius or any Schatten-p norm.

Remark 2.6 *Also note that the* Moore-Penrose pseudoinverse *of A, i.e., the unique matrix $A^+ = X$ satisfying*

(i) $AXA = A$,

(ii) $XAX = X$,

(iii) $(AX)^T = AX$,

(iv) $(XA)^T = XA$,

is given by

$$A^+ = V \begin{bmatrix} \hat{\Sigma}^{-1} & 0 \\ 0 & 0 \end{bmatrix} U^T,$$

where $\hat{\Sigma} = \operatorname{diag}\{\sigma_1, \ldots, \sigma_s\}$. This implies that the (minimum 2-norm) solution of the least squares problem $\min_{x\in\mathbb{R}^n} \|b - Ax\|$ is $x_* = A^+ b$.

[1] E. Schmidt 1907, Eckart/Young 1936, Mirsky 1960

As a motivation for the use of the SVD in dimension reduction/data compression, we cosnsider the following example:

Example 2.7 *MATLAB: Clown, Gatlinburg*

Krylov Subspaces

Another important ingredient from numerical linear algebra is the Krylov subspace.

Definition 2.8 *Let $A \in \mathbb{R}^{n\times n}$, $b \in \mathbb{R}^n$, and $q \in \mathbb{N}$. The subspace*

$$\mathcal{K}_q(A,b) = \operatorname{span}\{b, Ab, A^2b, \ldots, A^{q-1}b\} \subset \mathbb{R}^n \tag{2.5}$$

is called the Krylov subspace (of order q) *corresponding to A, b, while*

$$K(A,b,q) = \left[b, Ab, A^2b, \ldots, A^{q-1}b\right] \in \mathbb{R}^{n\times q}$$

is the corresponding Krylov matrix.

The following result summarizes important properties of Krylov subspaces and corresponding matrices.

Theorem 2.9 *Let $A \in \mathbb{R}^{n\times n}$ be nonsingular, $b \in \mathbb{R}^n$ and $A^{-1}b =: x^* \in \mathbb{R}^n$ be the solution of the linear system $Ax = b$. The following statements are equivalent:*

a) The vectors $b, Ab, A^2b, \ldots, A^kb$ are linearly dependent.

b) $\mathcal{K}_k(A,b) = \mathcal{K}_{k+1}(A,b) = \mathcal{K}_j(A,b)\ \forall j > k$.

c) $\mathcal{K}_k(A,b)$ is an A-invariant subspace, that is $A\mathcal{K}_k(A,b) \subseteq \mathcal{K}_k(A,b)$.

d) $x^* \in \mathcal{K}_k(A,b)$.

There is an important relation of Krylov subspaces and matrix decompositions.

Theorem 2.10 *Let $A \in \mathbb{R}^{n\times n}$, $Q \in \mathbb{R}^{n\times n}$ orthogonal satisfying*

$$Q^TAQ = T =: H = \begin{bmatrix} h_{11} & h_{12} & \ldots & \ldots & h_{1,n} \\ h_{21} & h_{22} & \ldots & \ldots & h_{2,n} \\ 0 & h_{32} & h_{33} & \ldots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \ldots & 0 & h_{n-1,n} & h_{n,n} \end{bmatrix}, \tag{2.6}$$

then with $q_1 := Qe_1$ it holds:

a) $Q^TK(A,q_1,n) = R = \left[\ \diagdown\ \right]$, i.e, $Q \cdot R$ is a QR decomposition of $K(A,q_1,n)$.

b) If R is nonsingular, then H is unreduced (i.e., $h_{k+1,k} \neq 0$ for all $k = 1, \ldots, n-1$).

c) Let $K = \underset{k=1,\ldots,n}{\operatorname{argmin}}\{r_{kk} = 0\}$, then $K - 1 = \underset{k=1,\ldots,n-1}{\operatorname{argmin}}\{h_{k+1.k} = 0\}$.

Proof:

$$\begin{aligned} Q^T K(A,\ q_1,\ n) &= [Q^T q_1,\ Q^T A Q Q^T q_1,\ (Q^T A Q)^2 Q^T q_1,\ \ldots\ (Q^T A Q)^{n-1} Q q_1] \\ &= [e_1,\ He_1,\ H^2 e_1,\ \ldots,\ H^{n-1} e_1] = R \end{aligned}$$

with $r_{11} = 1$. Furthermore, $r_{22} = h_{21},\ r_{33} = h_{32}h_{21},\ \ldots, r_{kk} = h_{k+1,k} \cdots h_{21} \Rightarrow R$ nonsingular, hence from $r_{kk} \neq 0\ \forall k$ it follows $h_{k+1,k} \neq 0 \forall k$, i.e. T is unreduced.

$$r_{k-1,k-1} \neq 0, r_{kk} = 0 \Rightarrow h_{k+1,k} = 0 \quad \Rightarrow \quad c).$$

□

Remark 2.11 *The decomposition (2.6) is called* Hessenberg decomposition *of A. Note that if $A = A^T$, i.e., A is symmetric, the matrix H in the Hessenebrg form will necessarily become symmetric tridiagonal.*

The above theorem devises a way to compute orthogonal bases of Krylov subsapces: evaluating (2.6) in the form

$$AQ = QH$$

columnwise leads to algorithms which compute orthonormal bases of $\mathcal{K}_q(A, b)$ column by column. For general nonsymmetric matrices, this leads to the *Arnoldi algorithm*, while for symmetric A, one obtains the *Lanczos iteration* (see Exercise 1.2).

One obtains the *Arnoldi recursion*

$$AQ_k = Q_k H_k + h_{k+1,k} q_{k+1} e_{k+1}^T \tag{2.7}$$

(and in a similar fashion, the Lanczos recursion in which the only difference is that with symmetric A, H_k is symmetric, too, and thus tridiagonal). It follows

$$H_k = Q_k^T A Q_k,$$

which is the *matrix Rayleigh quotient* of A with respect to Q_k.

2.2 Systems and Control Theory

Consider the linear, time-invariant (LTI) system

$$\begin{aligned} \dot{x}(t) &= Ax(t) + Bu(t), \quad t > 0, \quad x(0) = x^0, && \text{(2.8a)} \\ y(t) &= Cx(t) + Du(t), \quad t \geq 0, && \text{(2.8b)} \end{aligned}$$

where $A \in \mathbb{R}^{n\times n}$ is the state matrix, $B \in \mathbb{R}^{n\times m}$, $C \in \mathbb{R}^{p\times n}$, $D \in \mathbb{R}^{p\times m}$, and $x^0 \in \mathbb{R}^n$ is the initial state of the system. Here, n is the order (or state-space dimension) of the system. The particular model imposed by (2.8), given by a differential equation describing the behavior of the states x and an algebraic equation describing the outputs y is called a *state-space representation*. Alternatively, the relation between inputs and outputs can also be described in the *frequency domain* by an algebraic expression. Applying the *Laplace transform*

$$\mathcal{L} : x(t) \mapsto x(s) = \int_0^\infty e^{-st} x(t)\, dt \quad (\Rightarrow\ \dot{x}(t) \mapsto sx(s))$$

to the two equations in (2.8), and denoting the transformed arguments as $x(s)$, $y(s)$, $u(s)$ where s is the Laplace variable, we obtain

$$\begin{aligned} sx(s) - x(0) &= Ax(s) + Bu(s), \\ y(s) &= Cx(s) + Du(s). \end{aligned}$$

By solving for $x(s)$ in the first equation and inserting this into the second equation, we obtain

$$y(s) = \big(C(sI_n - A)^{-1}B + D\big)\, u(s) + C(sI_n - A)^{-1}x^0.$$

For a zero initial state, the relation between inputs and outputs is therefore completely described by the *transfer function*

$$G(s) := C(sI_n - A)^{-1}B + D. \tag{2.9}$$

Many interesting characteristics of an LTI system are obtained by evaluating $G(s)$ on the positive imaginary axis, that is, setting $s = \jmath\omega$. In this context, ω can be interpreted as the operating frequency of the LTI system.

In this section, we introduce some basic notation and properties of LTI systems used throughout this paper. More detailed introductions to LTI systems can be found in many textbooks [GL95, HP05, Son98, ZDG96] or handbooks [Lev96, Mut99]. We essentially follow these references here without further citations, but many other sources can be used for a good overview on the subjects covered in this section.

2.2.1 Analysis of Control Systems

Definition 2.12 *An LTI system is stable if all its poles are in the left half plane and it asymptotically (or Lyapunov or exponentially) stable if all poles are in the open left half plane* $\mathbb{C}^- := \{z \in \mathbb{C} \,|\, \Re(z) < 0\}$.

Sufficient for asymptotic stability is that A is asymptotically stable (or *Hurwitz*), i.e., the spectrum of A, denoted by $\Lambda(A)$, satisfies $\Lambda(A) \subset \mathbb{C}^-$. Note that by abuse of notation, often the terminology *stable system* is used for asymptotically stable systems.

First, we will ask the question whether for a given initial state x^0 the target x^1 can be reached for some admissible input function u, where we will denote the set of admissible functions by $\mathcal{U}_{ad}$. Controllability is only related to the state equation (2.8a). So we will ignore the output equation for the moment.

Definition 2.13 *(Controllability) Consider the target (the state to be reached)* $x^1 \in \mathbb{R}^n$.

a) The control system (2.8) *with initial value* $x(t_0) = x^0$ *is* controllable to x^1 in time $t_1 > t_0$ *if there exists* $u \in \mathcal{U}_{ad}$ *such that* $x(t_1; u) = x^1$. *(Equivalently,* (t_1, x^1) *is* reachable from (t_0, x^0).*)*

b) x^0 *is* controllable to x^1 *if there exists a* $t_1 > t_0$ *such that* (t_1, x^1) *can be reached from* (t_0, x^0).

c) If the system is controllable to x^1 *for all* (t_0, x^0) *with* $x^0 \in \mathbb{R}^n$, *it is* (completely) controllable.

The controllability set w.r.t. x^1 *is defined as* $\mathcal{C} := \bigcup_{t_1 > t_0} \mathcal{C}(t_1)$ *where*

$$\mathcal{C}(t_1) := \{x^0 \in \mathcal{X}; \exists u \in \mathcal{U}_{ad} : x(t_1; u) = x^1\}.$$

The last definition characterizes (complete) controllability as $\mathcal{C} = \mathbb{R}^n$ if we have an LTI system. Now we want to characterize controllability for linear systems. First, we need the solution of linear systems.

Proposition 2.14 *For the LTI system* (2.8), *the fundamental solution of the uncontrolled part of the LTI system, i.e., the homogeneous linear ODE* $\dot{x} = Ax$ *is* $\Phi(t, s) = e^{A(t-s)}$ *and therefore*

$$x(t) = e^{At}x^0 + \int_0^t e^{A(t-s)}Bu(s)ds = e^{At}(x^0 + \int_0^t e^{-As}Bu(s)ds), \tag{2.10}$$

$$y(t) = Ce^{At}\left(x^0 + \int_0^t e^{-As}Bu(s)ds\right) + Du(t). \tag{2.11}$$

Proof: See theory of linear differential equations. □

An important property of $\Phi(t, s) = e^{A(t-s)}$ is the *semigroup property*:

$$\Phi(t, s) = \Phi(t, \hat{t})\Phi(\hat{t}, s) = e^{A(t-\hat{t})}e^{A(\hat{t}-s)}$$

This follows from the uniqueness of the solution of the initial value problem $\dot{x} = Ax,\ x(0) = x^0$.

When does such an $u \in \mathcal{U}_{ad}$ exist such that x^1 can be reached from x^0? Assume $u(t) = B^T e^{-A^T t}c$, where $c \in \mathbb{R}$ is a constant. Then the solvability of

$$x^1 = x(t_1) = e^{At_1}x^0 + \int_0^{t_1} e^{A(t_1-t)}Bu(t)dt$$

is equivalent to

$$e^{-At_1}x^1 - x^0 = \underbrace{\int_0^{t_1} e^{-At}BB^Te^{-A^Tt}dt}_{=:P(0,t_1)}\, c.$$

We can uniquely solve for c (and, hence obtain an admissible control function $u(t)$) if the *finite time Gramian* $P(0,t_1)$ is nonsingular, which, as it is positive semidefinite in any case, implies

$$P(0,t_1) := \int_0^{t_1} e^{-At}BB^Te^{-A^Tt}dt > 0. \tag{2.12}$$

It should be noted that due to continuity of the integrand, positive semidefiniteness of $P(0,t)$ for one $t > 0$ implies positive semidefiniteness for all t, which in case of stability of A is equivalent to positive definiteness of the *(infinite) controllability Gramian*

$$P := \int_0^{\infty} e^{As}BB^Te^{A^Ts}ds. \tag{2.13}$$

Definition 2.15 *The* controllability matrix *of an LTI system as in* (2.8) *is*

$$K(A,B) := [B, AB, A^2B, \ldots, A^{n-1}B] \in \mathbb{R}^{n\times n\cdot m}.$$

The controllability matrix provides a nice characterization of *controllability*.

Theorem 2.16 *For LTI systems:* $\mathcal{C}(t_1) = \text{range}(K(A,B))\ \forall t_1 > 0$.

Theorem 2.16 shows that for a controllable system,

a) $\mathcal{C}(t_1) = \mathcal{C}(t_2)\ \forall t_1, t_2 > 0$. Hence $\mathcal{C} \equiv \mathcal{C}(t)$ and all controllability concepts are the same for LTI systems! Moreover, as $K(A,B)$ is independent of the target x^1, we can also conclude that $\mathcal{C}(t)$ is independent of the target, and thus for controllable systems, $\mathcal{C} = \mathbb{R}^n$ for all targets $x \in \mathbb{R}^n$.

b) $P(0,t_1) > 0 \Leftrightarrow P(0,t_2) > 0$ for $t_1, t_2 > 0$ arbitrary. In particular, for a controllable system, the infinite controllability Gramian satisfies $P > 0$.

Another useful characterization is the Hautus–(Popov)–Lemma.

Theorem 2.17 *The following are equivalent.*

a) The LTI system is controllable.

b) $\text{rank}\,(K(A,B)) = n$

c) If z is a left eigenvector of A, then $z^*B \neq 0$.

d) $\operatorname{rank}([\lambda I - A, B]) = n \ \forall \lambda \in \mathbb{C}$ *(Hautus test).*

e) The controllability Gramian P is positive definite.

Example 2.18

$$A = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix},\ B = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \quad \Longrightarrow \quad K(A,B) = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$
$$\Longrightarrow \quad \operatorname{rank}(K(A,B)) = 2$$
$$\Longrightarrow \quad (A,B) \text{ controllable.}$$

Alternatively using the Hautus–test we get:
As for $\lambda \notin \Lambda(A) : \operatorname{rank}(\lambda I - A) = n$, we only need to check $\lambda \in \Lambda(A) = \{\pm 1\}$.

$$\lambda = 1 : \lambda I - A = \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix},\ \text{left eigenvector is } \begin{bmatrix} 1 \\ 1 \end{bmatrix},\quad [1\ 1]B = 1 \neq 0$$
$$\lambda = -1 : \lambda I - A = \begin{bmatrix} -1 & -1 \\ -1 & -1 \end{bmatrix},\ \text{left eigenvector is } \begin{bmatrix} 1 \\ -1 \end{bmatrix},\quad [1\ -1]B = -1 \neq 0.$$

Hence the system is controllable. □

In many situations, full state information is not available. This rises the question whether we can get enough information about the system from the output equation (2.8b). The question is: suppose we have

$$y(t) = \tilde{y}(t)$$

corresponding to two trajectories $x, \tilde{x}$ obtained by to the same input function $u(t)$. Can we conclude that $x(t_0) = \tilde{x}(t_0)$, or even stronger that $x(t) = \tilde{x}(t)$ for $t \leq t_0, t \geq 0$ (past/future)? (Note that $(x(t_0) = \tilde{x}(t_0)$ is sufficient as trajectory uniquely determined. In other words, is the mapping $x^0 \to y(t)$ injective?)

Definition 2.19 *An LTI system is* reconstructable (observable) *if for solution trajectories $x(t), \tilde{x}(t)$ of* (2.8a) *obtained with the same input function u, we have*

$$\begin{aligned} y(t) &= \tilde{y}(t) \quad \forall t \leq t_0 \quad (\forall t \geq t_0) \\ \Longrightarrow \quad x(t) &= \tilde{x}(t) \quad \forall t \leq t_0 \quad (\forall t \geq t_0). \end{aligned}$$

Note: For LTI systems it is easy to see that reconstructability and observability are equivalent.

Theorem 2.20 *(duality) An LTI system is reconstructable if and only if the* dual system $\dot{x}(t) = -A^T x(t) + -C^T u(t)$ *is controllable.*

Theorem 2.21 *An LTI system is observable if and only if the* observability Gramian

$$Q := \int_0^\infty e^{A^T t} C^T C e^A t dt \tag{2.14}$$

is positive definite.

Duality also yields the Hautus–Popov–observability Lemma.

Corollary 2.22 *The following conditions are equivalent:*

a) The LTI system (2.8) *is reconstructable.*

b) The LTI system (2.8) *is observable.*

c) The observability matrix

$$\mathcal{O}(A,C) = \begin{bmatrix} C \\ CA \\ CA^2 \\ \vdots \\ CA^{n-1} \end{bmatrix} \in \mathbb{R}^{np\times n} \text{ has rank } n.$$

d) If $Ax = \lambda x$, *then* $C^T x \neq 0$.

e) $\operatorname{rank}\begin{bmatrix} \lambda I - A \\ C \end{bmatrix} = n.$

Note: a system that is both controllable and observable is called *minimal* for reasons which become obvious later when considering realizations.

For systems that are not controllable, there is the weaker concept of *stabilizability*, for which we have an analogous criterion.

Theorem 2.23 *The following conditions are equivalent:*

a) The LTI system (2.8) is stabilizable, i.e., $\exists F \in \mathbb{R}^{m\times n}$ *with* $\Lambda(A+BF) \subset \mathbb{C}^-$.

b) If $p^*A = \tilde{\lambda}p^*$ *and* $\operatorname{Re}(\lambda) \geq 0$, *then* $p^*B \neq 0$.

c) $\operatorname{rank}([A-\lambda I, B]) = n \quad \forall \lambda \in \mathbb{C}$ *with* $\operatorname{Re}(\lambda) \geq 0$.

d) In the Kalman decomposition of $(A,B), \Lambda(A_3) \subset \mathbb{C}^-$.

Definition 2.24 *(dual concept of stabilizability) An LTI system is* detectable *if for any solution* $x(t)$ *of* $\dot{x} = Ax$ *with* $Cx(t) \equiv 0$ *we have* $\lim_{t\to\infty} x(t) = 0$.
(We can not observe all of x, *but the unobservable part is stable.)*

Corollary 2.25 *The following conditions are equivalent:*

a) An LTI system is detectable.

b) (A^T, C^T) *is stabilizable.*

c) $Ax = \lambda x, \operatorname{Re}(\lambda) \geq 0 \Rightarrow C^T x \neq 0$.

d) $\operatorname{rank}\begin{bmatrix} \lambda I - A \\ C \end{bmatrix} = n$ *for all* $\lambda, \operatorname{Re}(\lambda) \geq 0$.

e) In the observability Kalman decomposition *of* (A^T, C^T),

$$W^T A W = \begin{bmatrix} A_1 & 0 \\ A_2 & A_3 \end{bmatrix}, CW = [C_1\ 0], \Lambda(A_3) \subset \mathbb{C}^-.$$

Note: In the Kalman decomposition, $\tilde{x} := W^T x$,

$$\begin{aligned} \dot{\tilde{x}}_1 &= A_1 \tilde{x}_1 \\ \dot{\tilde{x}}_2 &= A_2 \tilde{x}_1 + A_3 \tilde{x}_2 \\ y &= C_1 \tilde{x}_1 \end{aligned}$$

$\tilde{x}_2$ has no influence on the output. The states $\tilde{x}_2$ are called *unobservable*, $\lambda \in \Lambda(A_3)$ *unobservable modes*.

2.2.2 Realizations

A realization of an LTI system is the set of the four matrices

$$(A, B, C, D) \in \mathbb{R}^{n\times n} \times \mathbb{R}^{n\times m} \times \mathbb{R}^{p\times n} \times \mathbb{R}^{p\times m}$$

corresponding to (2.8). In general, an LTI system has infinitely many realizations as its transfer function is invariant under state-space transformations,

$$\mathcal{T} : \begin{cases} x & \to & Tx, \\ (A, B, C, D) & \to & (TAT^{-1}, TB, CT^{-1}, D), \end{cases} \tag{2.15}$$

as the simple calculation

$$D + (CT^{-1})(sI - TAT^{-1})^{-1}(TB) = C(sI_n - A)^{-1}B + D = G(s)$$

demonstrates. But this is not the only non-uniqueness associated to LTI system representations. Any addition of states that does not influence the input-output relation, meaning that for the same input u the same output y is achieved, leads to a realization of the same LTI system. Two simple examples are

$$\begin{aligned} \frac{d}{dt}\begin{bmatrix} x \\ x_1 \end{bmatrix} &= \begin{bmatrix} A & 0 \\ 0 & A_1 \end{bmatrix}\begin{bmatrix} x \\ x_1 \end{bmatrix} + \begin{bmatrix} B \\ B_1 \end{bmatrix} u(t), \quad y(t) = \begin{bmatrix} C & 0 \end{bmatrix}\begin{bmatrix} x \\ x_1 \end{bmatrix} + Du(t), \\ \frac{d}{dt}\begin{bmatrix} x \\ x_2 \end{bmatrix} &= \begin{bmatrix} A & 0 \\ 0 & A_2 \end{bmatrix}\begin{bmatrix} x \\ x_2 \end{bmatrix} + \begin{bmatrix} B \\ 0 \end{bmatrix} u(t), \quad y(t) = \begin{bmatrix} C & C_2 \end{bmatrix}\begin{bmatrix} x \\ x_2 \end{bmatrix} + Du(t), \end{aligned}$$

for arbitrary matrices $A_j \in \mathbb{R}^{n_j \times n_j}$, $j = 1, 2$, $B_1 \in \mathbb{R}^{n_1 \times m}$, $C_2 \in \mathbb{R}^{p \times n_2}$ and any $n_1, n_2 \in \mathbb{N}$. An easy calculation shows that both of these systems have the same transfer function $G(s)$ as (2.8) so that

$$(A, B, C, D),\ \left(\begin{bmatrix} A & 0 \\ 0 & A_1 \end{bmatrix}, \begin{bmatrix} B \\ B_1 \end{bmatrix}, \begin{bmatrix} C & 0 \end{bmatrix}, D\right),\ \left(\begin{bmatrix} A & 0 \\ 0 & A_2 \end{bmatrix}, \begin{bmatrix} B \\ 0 \end{bmatrix}, \begin{bmatrix} C & C_2 \end{bmatrix}, D\right)$$

are both realizations of the same LTI system described by the transfer function $G(s)$ in (2.9). Therefore, the order n of a system can be arbitrarily enlarged without changing the input-output mapping. On the other hand, for each system there exists a unique minimal number of states which is necessary to describe the input-output behavior completely. This number $\hat{n}$ is called the *McMillan degree* of the system. A *minimal realization* is a realization $(\hat{A}, \hat{B}, \hat{C}, \hat{D})$ of the system with order $\hat{n}$. Note that only the McMillan degree is unique; any state-space transformation (2.15) leads to another minimal realization of the same system. Finding a minimal realization for a given system can be considered as a first step of model reduction as redundant (non-minimal) states are removed from the system. Sometimes this is part of a model reduction procedure, e.g. optimal Hankel norm approximation, and can be achieved via balanced truncation.

An important fact is the following:

Theorem 2.26 *A realization (A, B, C, D) of the LTI system (2.8) is minimal if and only if (A, B) is controllable and (A, C) is observable.*

Although realizations are highly non-unique, stable LTI systems have a set of invariants with respect to state-space transformations that provide a good motivation for finding reduced-order models. To see this, we note that the controllability Gramian P from (2.13) and the observability Gramian Q from (2.14) of a stable LTI system (2.8) also satisfy the pair of *Lyapunov equations*

$$AP + PA^T + BB^T = 0, \tag{2.16a}$$
$$A^TQ + QA + C^TC = 0. \tag{2.16b}$$

From Lyapunov stability theory (see, e.g., [LT85, Chapter 13]) it is clear that for stable A, the Lyapunov equations in (2.16) have unique positive semidefinite solutions P and Q. We have already seen that $P > 0$ is equivalent to controllability and $Q > 0$ is equivalent to observability. By Theorem 2.26, controllability plus observability is equivalent to minimality of the system so that for minimal systems, all eigenvalues of the product PQ are strictly positive real numbers. The square roots of these eigenvalues, denoted in decreasing order by

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

are known as the *Hankel singular values (HSVs)* of the LTI system and are invariants of the system: let

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

be the transformed realization with associated controllability Lyapunov equation

$$0 = \hat{A}\hat{P} + \hat{P}\hat{A}^T + \hat{B}\hat{B}^T = TAT^{-1}\hat{P} + \hat{P}T^{-T}A^TT^T + TBB^TT^T.$$

This is equivalent to

$$0 = A(T^{-1}\hat{P}T^{-T}) + (T^{-1}\hat{P}T^{-T})A^T + BB^T.$$

The uniqueness of the solution of the Lyapunov equation (see, e.g., [LT85]) implies that $\hat{P} = TPT^T$ and, analogously, $\hat{Q} = T^{-T}QT^{-1}$. Therefore,

$$\hat{P}\hat{Q} = TPQT^{-1},$$

showing that $\Lambda(\hat{P}\hat{Q}) = \Lambda(PQ) = \{\sigma_1^2, \ldots, \sigma_n^2\}$. Note that extending the state-space by non-minimal states only adds HSVs of magnitude equal to zero, while the non-zero HSVs remain unchanged.

An important (and name-inducing) type of realizations are *balanced realizations*. A realization (A, B, C, D) is called *balanced* iff

$$P = Q = \begin{bmatrix} \sigma_1 & & \\ & \ddots & \\ & & \sigma_n \end{bmatrix};$$

that is, the controllability and observability Gramians are diagonal and equal with the decreasing HSVs on their respective diagonal entries. For a minimal realization there always exists a balancing state-space transformation of the form (2.15) with nonsingular matrix $T_b \in \mathbb{R}^{n\times n}$:

Theorem 2.27 *Given a minimal LTI system* (2.8) *with realization* (A, B, C, D)*, a balanced realization is given by the state-space transformation with*

$$T_b := \Sigma^{-\frac{1}{2}}V^TR,$$

where $P = S^TS$ *and* $Q = R^TR$ *(e.g., Cholesky decompositions) and*

$$SR^T = U\Sigma V^T$$

is the SVD of SR^T*.*

Proof: First note that $T^{-1} = S^TU\Sigma^{-\frac{1}{2}}$, then the result follows by simple algebraic manipulations. □

Remark 2.28 *For non-minimal systems the Gramians can also be transformed into diagonal matrices with the leading* $\hat{n} \times \hat{n}$ *submatrices equal to* $\operatorname{diag}(\sigma_1, \ldots, \sigma_{\hat{n}})$*, and*

$$\hat{P}\hat{Q} = \operatorname{diag}(\sigma_1^2, \ldots, \sigma_{\hat{n}}^2, 0, \ldots, 0);$$

see, e.g., [TP87].

2.2.3 Frequency Domain and Norms

A stable transfer function defines a mapping

$$G : \mathcal{L}_2 \to \mathcal{L}_2 : u \to y = Gu \tag{2.17}$$

where the two function spaces denoted by $\mathcal{L}_2$ are actually different spaces and should more appropriately be denoted by $\mathcal{L}_2(\mathbb{C}^m)$ and $\mathcal{L}_2(\mathbb{C}^p)$, respectively. As the dimension of the underlying spaces will always be clear from the context, i.e., the dimension of the transfer function matrix $G(s)$ or the dimension of input and output spaces, we allow ourselves the more sloppy notation used in (2.17). The function space $\mathcal{L}_2$ contains the square integrable functions in the frequency domain, obtained via the Laplace transform of the square integrable functions in the time domain, usually denoted as $\mathcal{L}_2(-\infty, \infty)$. The $\mathcal{L}_2$-functions that are analytic in the open right half plane $\mathbb{C}^+$ form the *Hardy space* $\mathcal{H}_2$. Note that $\mathcal{H}_2$ is a closed subspace of $\mathcal{L}_2$. Under the Laplace transform $\mathcal{L}_2$ and $\mathcal{H}_2$ are isometric isomorphic to $\mathcal{L}_2(-\infty, \infty)$ and $\mathcal{L}_2[0, \infty)$, respectively. (This is essentially the *Paley-Wiener Theorem* which is the Laplace transform analog of Parseval's identity for the Fourier transform.) Therefore it is clear that the frequency domain spaces $\mathcal{H}_2$ and $\mathcal{L}_2$ can be endowed with the corresponding norms from their time domain counterparts. Due to this isometry, our notation will not distinguish between norms for the different spaces so that we will denote by $\|f\|_2$ the induced 2-norm on any of the spaces $\mathcal{L}_2(-\infty, \infty)$, $\mathcal{L}_2$, $\mathcal{L}_2[0, \infty)$, and $\mathcal{H}_2$. Using the definition (2.17), it is therefore possible to define an operator norm for G by

$$\|G\| := \sup_{\|u\|_2 \leq 1} \|Gu\|_2.$$

It turns out that this operator norm equals the $\mathcal{L}_\infty$-norm of the transfer function G, which for rational transfer functions can be defined as

$$\|G\|_\infty := \sup_{\omega \in \mathbb{R}} \sigma_{\max}(G(\jmath\omega)). \tag{2.18}$$

The $p \times m$-matrix-valued functions G for which $\|G\|_\infty$ is bounded, i.e., those essentially bounded on the imaginary axis, form the function space $\mathcal{L}_\infty$. The subset of $\mathcal{L}_\infty$ containing all $p \times m$-matrix-valued functions that are analytical and bounded in $\mathbb{C}^+$ form the Hardy space $\mathcal{H}_\infty$. As a consequence of the maximum modulus theorem, $\mathcal{H}_\infty$ functions must be bounded on the imaginary axis so that the essential supremum in (2.18) simplifies to a supremum for rational functions G. Thus, the $\mathcal{H}_\infty$-norm of the rational transfer function $G \in \mathcal{H}_\infty$ can be defined as

$$\|G\|_\infty := \sup_{\omega \in \mathbb{R}} \sigma_{\max}(G(\jmath\omega)). \tag{2.19}$$

A fact that will be of major importance throughout this paper is that the transfer function of a stable LTI system is rational with no poles in the closed right-half plane. Thus, $G \in \mathcal{H}_\infty$ for all stable LTI systems.

Although the notation is somewhat misleading, the $\mathcal{H}_\infty$-norm is the 2-induced operator norm. Hence the sub-multiplicativity condition

$$\|y\|_2 \leq \|G\|_\infty \|u\|_2 \tag{2.20}$$

holds. This inequality implies an important way to tackle the model reduction problem: suppose the original system and the reduced-order model (**??**) are driven by the same input function $u \in \mathcal{H}_2$, so that

$$y(s) = G(s)u(s), \qquad \hat{y}(s) = \hat{G}(s)u(s),$$

where $\hat{G}$ is the transfer function corresponding to (**??**); then we obtain the error bound

$$\|y - \hat{y}\|_2 \leq \|G - \hat{G}\|_\infty \|u\|_2. \tag{2.21}$$

Due to the aforementioned Paley-Wiener theorem, this bound holds in the frequency domain and the time domain. Therefore a goal of model reduction is to compute the reduced-order model so that $\|G - \hat{G}\|_\infty$ is smaller than a given tolerance threshold.

Bibliography

[ANS] ANSYS, Inc., http://www.ansys.com. *ANSYS*.

[BDD+00] Z. Bai, J. Demmel, J. Dongarra, A. Ruhe, and H. van der Vorst, editors. *Templates for the Solution of Algebraic Eigenvalue Problems: A Practical Guide*, volume 11 of *Software - Environments - Tools*. SIAM Publications, Philadelphia, PA, 2000.

[BHK11] P. Benner, M.E. Hochstenbach, and P. Kürschner. Model order reduction of large-scale dynamical systems with Jacobi-Davidson style eigensolvers. In *2011 International Conference on Communications, Computing and Control Applications (CCCA)*, March 2011. 6 pages, available at `http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=6031208`.

[BL04] J.K. Bennighof and R.B. Lehoucq. An automated multilevel substructuring method for eigenspace computation in linear elastodynamics. *SIAM J. Sci. Comput.*, 25(6):2084–2106, 2004.

[CB68] R.R. Craig and M.C.C. Bampton. Coupling of substructures for dynamic analysis. *AIAA J.*, 6:1313–1319, 1968.

[Dav66] E.J. Davison. A method for simplifying linear dynamic systems. *IEEE Trans. Automat. Control*, AC–11:93–101, 1966.

[GL95] M. Green and D.J.N Limebeer. *Linear Robust Control*. Prentice-Hall, Englewood Cliffs, NJ, 1995.

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

[Guy68] R.J. Guyan. Reduction of stiffness and mass matrices. *AIAA J.*, 3:380, 1968.

[GV96] G.H. Golub and C.F. Van Loan. *Matrix Computations*. Johns Hopkins University Press, Baltimore, third edition, 1996.

[HP05] D. Hinrichsen and A.J. Pritchard. *Mathematical Systems Theory I*. Springer-Verlag, Berlin, 2005.

[LA86] Y. Liu and B.D.O. Anderson. Controller reduction via stable factorization and balancing. *Internat. J. Control*, 44:507–531, 1986.

[Lev96] W.S. Levine, editor. *The Control Handbook*. CRC Press, 1996.

[LSY98] R.B. Lehoucq, D.C. Sorensen, and C. Yang. *ARPACK User's Guide*, volume 6 of *Software - Environments - Tools*. SIAM Publications, Philadelphia, PA, 1998.

[LT85] P. Lancaster and M. Tismenetsky. *The Theory of Matrices*. Academic Press, Orlando, 2nd edition, 1985.

[Mar66] S.A. Marschall. An approximate method for reducing the order of a linear system. *Contr. Eng.*, 10:642–648, 1966.

[MLP96] N. Martins, L.T.G. Lima, and H.J.C.P. Pinto. Computing dominant poles of power system transfer functions. *IEEE Transactions on Power Systems*, 11(1):162–170, feb 1996.

[MQ03] N. Martins and P.E.M. Quintao. Computing dominant poles of power system multivariable transfer functions. *IEEE Transactions on Power Systems*, 18(1):152–159, feb 2003.

[MSC] MSC.Software Corporation, http://www.mscsoftware.com. *MSC.Nastran*.

[Mut99] A.G.O. Mutambara. *Design and Analysis of Control Systems*. CRC Press, Boca Raton, FL, 1999.

[Rix04] D.J Rixen. A dual Craig–Bampton method for dynamic substructuring. *J. Comput. Appl. Math.*, 168(1–2):383–391, 2004.

[RM06a] J. Rommes and N. Martins. Efficient computation of multivariable transfer function dominant poles using subspace acceleration. *IEEE Transactions on Power Systems*, 21(4):1471–1483, nov. 2006.

[RM06b] J. Rommes and N. Martins. Efficient computation of transfer function dominant poles using subspace acceleration. *IEEE Transactions on Power Systems*, 21(3):1218–1226, aug. 2006.

[RS08] J. Rommes and G. L. G. Sleijpen. Convergence of the dominant pole algorithm and Rayleigh quotient iteration. *SIAM Journal on Matrix Analysis and Applications*, 30(1):346–363, 2008.

[Son98] E.D. Sontag. *Mathematical Control Theory*. Springer-Verlag, New York, NY, 2nd edition, 1998.

[TP87] M.S. Tombs and I. Postlethwaite. Truncated balanced realization of a stable non-minimal state-space system. *Internat. J. Control*, 46(4):1319–1330, 1987.

[Van00] P. Van Dooren. Gramian based model reduction of large-scale dynamical systems. In D.F. Griffiths and G.A. Watson, editors, *Numerical Analysis 1999. Proc. 18th Dundee Biennial Conference on Numerical Analysis*, pages 231–247, London, UK, 2000. Chapman & Hall/CRC.

[van02] H. van der Vorst. Computational methods for large eigenvalue problems. In P.G. Ciarlet and J.L. Lions, editors, *Handbook of Numerical Analysis*, volume 8, pages 3–179. North-Holland/ Elsevier, Amsterdam, 2002.

[Var91] A. Varga. Efficient minimal realization procedure based on balancing. In *Prepr. of the IMACS Symp. on Modelling and Control of Technological Systems*, volume 2, pages 42–47, 1991.

[Var95] A. Varga. Enhanced modal approach for model reduction. *Math. Model. Syst.*, 1(2):91–105, 1995.

[ZDG96] K. Zhou, J.C. Doyle, and K. Glover. *Robust and Optimal Control*. Prentice-Hall, Upper Saddle River, NJ, 1996.