ORTHOGONAL MATRIX DECOMPOSITIONS IN SYSTEMS AND CONTROL

P. M. VAN DOOREN
Dept. Math. Eng., Université catholique de Louvain, Belgium

Abstract. In this paper we present several types of orthogonal matrix decompositions used in systems and control. We focus on those related to eigenvalue and singular value problems and include generalizations to several matrices.

1. Introduction

In systems and control theory, one often uses state space models to represent a dynamical system. In such models the relation between inputs $u(t) \in \mathcal{R}^m$ and outputs $y(t) \in \mathcal{R}^p$ is described via the use of a state $x(t) \in \mathcal{R}^n$ and a system of first order differential equations :

$$\left\{ \begin{array}{rcl} E\dot{x}(t) & = & Ax(t) + Bu(t) \\ y(t) & = & Cx(t) + Du(t), \end{array} \right. \tag{1}$$

where E and A are real $n \times n$ matrices and B, C and D are real $n \times m$, $p \times n$ and $p \times m$ matrices, respectively. In the above model (1) the input, output and state vectors are continuous time functions. An analogous model is used for discrete time vectors functions u_k, y_k and x_k, now involving a system of first order difference equations :

$$\left\{ \begin{array}{rcl} Ex_{k+1} & = & Ax_k + Bu_k \\ y_k & = & Cx_k + Du_k. \end{array} \right. \tag{2}$$

If one takes the Laplace transform of (1) and the z-transform of (2), then both can be represented by the system of algebraic equations :

$$\left\{ \begin{array}{rcl} \lambda Ex(.) & = & Ax(.) + Bu(.) \\ y(.) & = & Cx(.) + Du(.), \end{array} \right. \tag{3}$$

where λ stands for the differential operator and the difference operator in the two respective cases. The transfer function of this model is then obtained by eliminating the state $x(.)$:

$$T(\lambda) = C(\lambda E - A)^{-1}B + D, \tag{4}$$

and describes directly the relation between inputs and outputs :

$$y(.) = T(\lambda)u(.).$$

Notice finally, that the transfer function (and hence the input/output behavior) is not affected by the system equivalence transformation

$$\{E, A, B, C, D\} \Longrightarrow \{\hat{E}, \hat{A}, \hat{B}, \hat{C}, \hat{D}\} \doteq \{SET, SAT, SB, CT, D\}. \tag{5}$$

An important subclass of these models consists of the so-called standard state space models where $E = I$, in which case the state vector $x(.)$ is given explicitly by the first equation in (3). The equivalence transformation (5) now becomes a similarity transformation since $\hat{E} = ST$ must also be the identity :

$$\{A, B, C, D\} \Longrightarrow \{\hat{A}, \hat{B}, \hat{C}, \hat{D}\} \doteq \{T^{-1}AT, T^{-1}B, CT, D\}. \tag{6}$$

Although these state space models are not the only ones used for systems and control purposes, they are the ones that have been most heavily studied as far as numerical algorithms are concerned (see e.g. [15]). We will assume for the sequel that the system under consideration is given in such a form and that the model parameters are actually known (i.e. the system has already been identified). Once the system is given, one typically has to analyze its properties (frequency response, poles/zeros, stability, robustness, ...) and then design a particular controller in order to improve some characteristics or to satisfy certain design criteria (tracking, robustness, optimality criteria, ...).

2. Condensed versus canonical forms

Many analysis and design problems are well understood these days and their theoretical solution is often described in terms of so called canonical forms, which have been defined for state space models of multivariable linear systems. These forms are typically very sparse since they are described with a minimum number of parameters. Therefore they often allow to efficiently characterize all solutions to a particular problem, which is of course very appealing. Unfortunately, it has been shown that these forms are also very sensitive to compute and amount to a coordinate transformation that can be very poorly conditioned. For most analysis and design problems encountered in linear system theory, one can as well make use of so-called *condensed forms*, which can be obtained under orthogonal transformations [25]. Such transformations have become a major tool in the development of reliable numerical linear algebra algorithms. A first reason for this is the numerical sensitivity of the problem at hand. The sensitivity (or *conditioning*) of several problems in linear algebra can be expressed in terms of norms, singular values or angles between spaces and each of these are invariant under orthogonal transformations. These transformations therefore allow to reformulate the problem in a new coordinate system which is more appropriate for solving the problem, and this without affecting its sensitivity. A second reason in the numerical stability of the algorithm used for solving the problem. Most decomposition involving orthogonal transformations can be obtained by a sequence of Givens or Householder transformations which can be performed in a numerically stable manner. The concatenation of such transformations can also be performed in a backward stable manner because numerical errors resulting from previous steps are indeed maintained in norm throughout subsequent steps since these transformations (and their inverse) have 2-norm equal to 1. Condensed forms obtained under orthogonal transformations are therefore more appropriate for solving several systems and control problems involving (generalized) state space models.

We explain this below with the special example of analyzing the poles of a single-input/single-output system given in standard state space model (i.e. $E = I$, $S = T^{-1}$ and $m = p = 1$ in the above models). For such models the poles are the eigenvalues of the matrix A and the classical form describing the fine structure of these eigenvalues is the Jordan canonical form. So we choose the similarity transformation (6) where $\hat{A}_J = T^{-1}AT$ is in Jordan canonical form. For convenience, we give the transformed system $\{\hat{A}_J, \hat{B}_J, \hat{C}_J, \hat{D}_J\}$ in the form of a compound matrix (the reason of this will become apparent later) :

$$\left[\begin{array}{c|c} \hat{B}_J & \hat{A}_J \\ \hline \hat{D}_J & \hat{C}_J \end{array}\right] \doteq \left[\begin{array}{c|ccccccc} \times & \lambda_1 & 1 & 0 & 0 & 0 & 0 & 0 \\ \times & 0 & \lambda_1 & 1 & 0 & 0 & 0 & 0 \\ \times & 0 & 0 & \lambda_1 & 0 & 0 & 0 & 0 \\ \times & 0 & 0 & 0 & \lambda_2 & 1 & 0 & 0 \\ \times & 0 & 0 & 0 & 0 & \lambda_2 & 0 & 0 \\ \times & 0 & 0 & 0 & 0 & 0 & \lambda_3 & 0 \\ \times & 0 & 0 & 0 & 0 & 0 & 0 & \lambda_4 \\ \hline \times & \times & \times & \times & \times & \times & \times & \times \end{array}\right], \qquad (7)$$

where we assume for simplicity that the eigenvalues λ_i are real. Notice that we chose our example such that there is only one Jordan block associated with each individual eigenvalue, but in general this does not have to be the case. This form not only describes the poles of the system, but contains more information, like e.g. the partial fraction description of the transfer function. But a large disadvantage of the form is that it requires a state space transformation T to put A in its Jordan form $\hat{A}_J$, and that the norms T and T^{-1} can not be bounded in general. On the other hand, when one restricts T to be orthogonal, then so is T^{-1} and both are bounded in norm. Under such transformations, one can always reduce A to triangular form, called the *Schur form*, which also has the eigenvalues on diagonal :

$$\left[\begin{array}{c|c} \hat{B}_S & \hat{A}_S \\ \hline \hat{D}_S & \hat{C}_S \end{array}\right] \doteq \left[\begin{array}{c|ccccccc} \times & \lambda_1 & \times & \times & \times & \times & \times & \times \\ \times & 0 & \lambda_1 & \times & \times & \times & \times & \times \\ \times & 0 & 0 & \lambda_1 & \times & \times & \times & \times \\ \times & 0 & 0 & 0 & \lambda_2 & \times & \times & \times \\ \times & 0 & 0 & 0 & 0 & \lambda_2 & \times & \times \\ \times & 0 & 0 & 0 & 0 & 0 & \lambda_3 & \times \\ \times & 0 & 0 & 0 & 0 & 0 & 0 & \lambda_4 \\ \hline \times & \times & \times & \times & \times & \times & \times & \times \end{array}\right]. \tag{8}$$

If one is only interested in computing the poles of the system, it is well known that the latter form is numerically much more reliable and actually requires less computations that the Jordan form [7].

The theorem for general $n \times n$ pencils $\lambda E - A$ is known as the generalized Schur form, and applies to regular pencils (i.e. $\det(\lambda E - A) \not\equiv 0$). Its so-called *generalized eigenvalues* are the roots of $\det(\lambda E - A) = 0$.

Theorem 1 [12]
There always exist orthogonal transformations Q and Z that transform a regular pencil $\lambda E - A$ to

$$Q^T(\lambda E - A)Z = \lambda E_S - A_S, \tag{9}$$

where E_S is upper triangular and A_S is block upper triangular with a 1×1 diagonal block corresponding to each real generalized eigenvalue and a 2×2 diagonal block corresponding to each pair of complex conjugate generalized eigenvalues (such matrices are called *quasi triangular*). This decomposition exists for every ordering of eigenvalues in the quasi triangular form. ■

If $E = I$ one retrieves the standard (quasi triangular) Schur decomposition $A_S = U^T A U$ based on an orthogonal similarity transformation by taking $U = Z = Q$. Notice that if E is invertible one also has

$$Q^T A E^{-1} Q = A_S E_S^{-1}, \quad Z^T E^{-1} A Z = E_S^{-1} A_S,$$

which are both quasi triangular matrices. Then Q and Z of the generalized Schur form can be obtained from the standard Schur forms of AE^{-1} and

$E^{-1}A$ but this "detour" should be avoided since E may be badly conditioned in general. One of the most important uses of this form is the computation of orthogonal bases for eigenspaces. Consider any (block) triangular decomposition where we partitioned the invertible matrix X conformably :

$$X^{-1}AX = \begin{bmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{bmatrix}, \quad X = \begin{bmatrix} X_1 & X_2 \end{bmatrix}. \tag{10}$$

Then $AX_1 = X_1A_{11}$ which implies that $\mathcal{X} = ImX_1$ satisfies the condition for an *invariant subspace*

$$A\mathcal{X} \subset \mathcal{X}.$$

When X is orthogonal (as in the Schur decomposition) the basis X_1 is of course orthogonal as well. The corresponding concept for the generalized eigenvalue problem $\lambda E - A$ is that of *deflating subspace* defined by the condition

$$\dim(A\mathcal{X} + E\mathcal{X}) = \dim\mathcal{X}.$$

For E invertible this is easily shown to be equivalent to $E^{-1}A\mathcal{X} \subset \mathcal{X}$ and hence each deflating subspace of $\lambda E - A$ is and invariant subspace of $E^{-1}A$. The first k columns of the right transformation Z [12] are therefore an orthogonal basis for a deflating subspace of the pencil $\lambda E - A$. We refer to [12], [7], [21] for a more rigorous discussion. The use of these eigenspaces in control shows up in the solution of several matrix equations.

We illustrate this again with a standard eigenvalue problem (i.e. $E = I$). Suppose one wants to solve the quadratic matrix equation (of dimension $q \times p$) :

$$M_{21} - XM_{11} + M_{22}X - XM_{12}X = 0 \tag{11}$$

in the $q \times p$ matrix X, then it is easily verified that this is equivalent to

$$\begin{bmatrix} I_p & 0 \\ -X & I_q \end{bmatrix} \begin{bmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{bmatrix} \begin{bmatrix} I_p & 0 \\ X & I_q \end{bmatrix} = \begin{bmatrix} \hat{M}_{11} & \hat{M}_{22} \\ 0 & \hat{M}_{22} \end{bmatrix}, \tag{12}$$

where $\hat{M}_{11} \doteq M_{11} + M_{12}X$, $\hat{M}_{12} = M_{12}$, $\hat{M}_{22} \doteq M_{22} - XM_{12}$, and $\hat{M}_{21} = 0$ since it is precisely equation (11). But (12) is a similarity transformation on the $(p+q) \times (p+q)$ matrix M partitioned in the 4 blocks M_{ij} $i = 1,2,$ $j = 1,2$. The block triangular decomposition says that the eigenvalues of M are the union of those of $\hat{M}_{11}$ and of $\hat{M}_{22}$ and that the columns of $\begin{bmatrix} I_p \\ X \end{bmatrix}$ span an invariant subspace of the matrix M corresponding to the p eigenvalues of $\hat{M}_{11}$ [10]. Let us suppose for simplicity that M is simple, i.e. that it has distinct eigenvalues. Then every invariant subspace of a particular dimension p is spanned by p eigenvectors. Therefore, let $\begin{bmatrix} X_{11} \\ X_{21} \end{bmatrix}$ be a matrix whose columns are p arbitrary eigenvectors, then it is a basis for the corresponding invariant subspace. If moreover X_{11} is invertible then the columns of $\begin{bmatrix} I_p \\ X \end{bmatrix}$ with $X = X_{21}X_{11}^{-1}$ span the same subspace and hence X is a solution of the quadratic matrix equation (11). One shows that the eigenvalues corresponding to the selected eigenvectors will be the eigenvalues of $\hat{M}_{11}$ after applying the transformation (12). This approach actually yields *all solutions* X provided M is simple and the matrices X_{11} defined above are invertible. But it requires the computation of all the eigenvectors, which is obtained from a diagonalizing similarity transformation. One shows that when M has repeated eigenvalues, one should compute its Jordan canonical to find all solutions of the quadratic matrix equation (11) [15]. The disadvantage of this approach is that it involves the construction of a transformation T that may be badly conditioned.

But any invariant subspace has also an orthogonal basis, and in general these basis vectors will not be eigenvectors since eigenvectors need not be orthogonal to each other. It turns out that such a basis is exactly obtained by the Schur decomposition (8). One can always compute an orthogonal similarity transformation that quasi triangularizes the matrix M. If we then partition the triangular matrix with a $p \times p$ leading block :

$$\begin{bmatrix} U_{11} & U_{12} \\ U_{21} & U_{22} \end{bmatrix}^T \begin{bmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{bmatrix} \begin{bmatrix} U_{11} & U_{12} \\ U_{21} & U_{22} \end{bmatrix} = \begin{bmatrix} \tilde{M}_{11} & \tilde{M}_{22} \\ 0 & \tilde{M}_{22} \end{bmatrix} \tag{13}$$

then it follows than the columns of $\begin{bmatrix} U_{11} \\ U_{21} \end{bmatrix}$ also span an invariant subspace of M and that the columns of $\begin{bmatrix} I_p \\ X \end{bmatrix}$ with $X = U_{21}U_{11}^{-1}$ span the same subspace, provided U_{11} is invertible [10]. This approach has the advantage that it uses numerically reliable coordinate transformations in (13) but the disadvantage that only one invariant subspace is directly obtained that way. It turns out that in several applications one only needs one invariant subspace. Typical examples arise in applications involving continuous time systems :

– the algebraic Riccati equation $XBR^{-1}B^TX - XA - A^TX - Q = 0$ from optimal control. Here the relevant matrix is

$$M = \begin{bmatrix} A & -BR^{-1}B^T \\ -Q & -A^T \end{bmatrix}$$

and the matrix $\hat{M}_{11}$ must contain all eigenvalues of M in the left half plane

– the Lyapunov equation $AX + XA^T + Q = 0$ occurring in stability analysis. Here

$$M = \begin{bmatrix} -A^T & 0 \\ Q & A \end{bmatrix}$$

and $\hat{M}_{11} = -A^T$

– the Sylvester equation $AX - XB + C = 0$ where

$$M = \begin{bmatrix} B & 0 \\ C & A \end{bmatrix}$$

and $\hat{M}_{11} = B$.

In each of these cases one has a well defined spectrum in mind for the matrix $\tilde{M}_{11}$ after transformation, and so only one invariant subspace has to be computed. We point out that Lyapunov and Sylvester equations can be viewed as special linear cases of the quadratic Riccati equation, which is extensively discussed in [10] [15]. The efficient calculation of the linear equations is discussed in [1] and is based on the Schur forms of A and B. This so-called *Schur* approach was proposed in a number of papers [15] and is now the recommended basic technique to solve these problems, although improvements are still being found in this area.

The generalized eigenvalue counterpart involves deflating subspaces and arises in applications involving discrete time systems. We refer to [15] for more details. Besides the computations of eigenspaces it is also shown there that the Schur form plays an important role in simulation of dynamical systems and in the computation of system and frequency responses. In fact, any matrix function of A can be evaluated efficiently using the Schur form of A, and the same holds for matrix functions of $E^{-1}A$ using the generalized Schur form of $\lambda E - A$.

Another class of orthogonal coordinate transformations (which we do not want to elaborate on here) are the so-called Hessenberg and staircase forms. A staircase form has a typical echelon-like form for a couple of matrices as shown below for the (A, B) pair of a system with 3 inputs, 6 states and 2 outputs :

$$\left[\begin{array}{c|c} \hat{B}_c & \hat{A}_c \\ \hline \hat{D}_c & \hat{C}_c \end{array}\right] \doteq \left[\begin{array}{ccc|ccccccc} X & \times & \times & \times & \times & \times & \times & \times & \times & \times \\ 0 & X & \times & \times & \times & \times & \times & \times & \times & \times \\ 0 & 0 & X & \times & \times & \times & \times & \times & \times & \times \\ 0 & 0 & 0 & 0 & X & \times & \times & \times & \times & \times \\ 0 & 0 & 0 & 0 & 0 & X & \times & \times & \times & \times \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & X & \times & \times \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & X & \times \\ \hline \times & \times & \times & \times & \times & \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times & \times & \times & \times & \times & \times \end{array}\right]. \qquad (14)$$

The X elements are nonzero and define exactly the controllability indices of the system [20]. This form is the orthogonal counterpart of the Brunovsky canonical form, which is more sparse but can only be obtained under non-orthogonal state space transformation. These have been developed in several papers and have numerous applications, including : controllability, observability, minimality of state space and generalized state space models, pole placement and robust feedback, observer design and Kalman filtering. They also exist for generalized state space systems in which case the additional matrix $\hat{E}_c$ is upper triangular. We refer to [25] [20] for more details on this form and its applications in systems and control.

A third important condensed form is the singular value decomposition of a matrix. It is given by the following theorem.

Theorem 2 [9]
Let A be a $m \times n$ real matrix of rank r. Then there exist orthogonal matrices U and V of dimension $m \times m$ and $n \times n$ respectively, such that

$$U^T A V = \Sigma = \begin{bmatrix} \Sigma_r & 0 \\ 0 & 0 \end{bmatrix} \tag{15}$$

where $\Sigma_r = \mathrm{diag}\{\sigma_1, \cdots, \sigma_r\}$ with $\sigma_1 \geq \cdots \geq \sigma_r > 0$. ■

Its main uses are detecting the rank of the matrix A in a reliable fashion, finding lower rank approximations and finding orthogonal bases for the kernel and image of A. In systems and control these problems show up in identification, balancing and finding bases for various geometric concepts related to state space models. We refer to [23] [14] [13] [15] for an extensive discussion of these applications.

3. Matrix sequences

In this section we look at orthogonal decompositions of a sequence of matrices. These typically occur in the context of discrete linear time varying systems :

$$\left\{ \begin{array}{rcl} E_k x_{k+1} & = & A_k x_k + B_k u_k \\ y_k & = & C_k x_k + D_k u_k , \end{array} \right. \tag{16}$$

arising e.g. from a discretization of a continuous time system.

Let us first consider the input to state map over a period of time $[1, N]$ (notice that in the control literature one prefers to start from $k = 0$, but starting from $k = 1$ turns out to be more convenient here). If the matrices E_k are all invertible then clearly each state x_k is well defined by these equations. One is often interested in the zero input behavior (i.e. the homogeneous system) which yields an explicit expression for the final state x_{N+1} in terms of the initial state x_1 :

$$x_{N+1} = \Phi_{N,1} x_1 \tag{17}$$

where

$$\Phi_{N,1} = E_N^{-1} A_N \cdots E_2^{-1} A_2 E_1^{-1} A_1 \tag{18}$$

is the state transition matrix over the interval $[1, N]$. If the matrices E_k are *not* invertible, these this expression does not make sense, but still one may be able to solve the input to state map when imposing boundary conditions

$$F_N x_{N+1} - F_1 x_1 + f = 0. \tag{19}$$

We can then rewrite these equations is the following matrix form :

$$\begin{bmatrix} -A_1 & E_1 & & & \\ & -A_2 & E_2 & & \\ & & \ddots & \ddots & \\ & & & -A_N & E_N \\ F_1 & & & & -F_N \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_N \\ x_{N+1} \end{bmatrix} = \begin{bmatrix} B_1 u_1 \\ B_2 u_2 \\ \vdots \\ B_N u_N \\ f \end{bmatrix}, \tag{20}$$

which will have a solution provided the (square) matrix in the left hand side is invertible.

3.1. PERIODIC BOUNDARY VALUE PROBLEMS

A periodic system is a set of difference equations (16) where now the coefficient matrices vary periodically with time, i.e. $M_k = M_{k+K}$, $\forall k$ and for $M = E, A, B, C$ and D. The period is the smallest value of K for which this holds. It was shown in [18] that a periodic system of period K is *solvable and conditionable* (i.e. has a well defined solution for suitably chosen boundary conditions F_1, F_N) for all N, provided the pencil

$$\lambda \mathcal{E} - \mathcal{A} \doteq \begin{bmatrix} -A_1 & \lambda E_1 & & \\ & \ddots & \ddots & \\ & & -A_{K-1} & \lambda E_{K-1} \\ \lambda E_K & & & -A_K \end{bmatrix} \tag{21}$$

is regular (i.e. $\det(\lambda \mathcal{E} - \mathcal{A}) \not\equiv 0$). Such periodic systems are said to be *regular*. Two point boundary value problems for regular periodic systems have thus unique solutions for *any* time interval N, provided the boundary conditions are suitably chosen.

Let us now choose a time interval equal to one period ($N = K$) and introduce boundary conditions that allow us to define eigenvectors and eigenvalues of periodic boundary value problems. Since these concepts are typically linked to homogeneous systems, we impose :

$$u_k = 0, k = 1, \ldots, K, \quad f = 0, \tag{22}$$

and for the invariance of the boundary vectors we impose :

$$F_1 = sI_n, \; F_K = cI_n, \implies sx_1 = cx_{K+1} \quad \text{with} \quad c^2 + s^2 = 1, \tag{23}$$

which yields :

$$\begin{bmatrix} -A_1 & E_1 & & & \\ & -A_2 & E_2 & & \\ & & \ddots & \ddots & \\ & & & -A_K & E_K \\ sI_n & & & & -cI_n \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_K \\ x_{K+1} \end{bmatrix} = 0. \tag{24}$$

After some algebraic manipulations one shows that this is equivalent to :

$$(z\mathcal{E}_b - \mathcal{A}_b)\mathbf{x}(1,K) \doteq \begin{bmatrix} -A_1 & E_1 & & \\ & \ddots & \ddots & \\ & & -A_{K-1} & E_{K-1} \\ zE_K & & & -A_K \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_K \end{bmatrix} = 0 \tag{25}$$

where $z = s/c$. This condition says that $\mathbf{x}(1,K)$ is an eigenvector of the pencil in the left hand side, with eigenvalue z. The pencils $z\mathcal{E}_b - \mathcal{A}_b$ and $\lambda\mathcal{E} - \mathcal{A}$ are closely related [22]. In case all matrices E_k are invertible, it follows e.g. that

$$\det(z\mathcal{E}_b - \mathcal{A}_b) = c.\det(zI_n - \Phi_{K,1}), \quad \text{and} \quad \det(\lambda\mathcal{E} - \mathcal{A}) = c.\det(\lambda^K I_n - \Phi_{K,1})$$

where $\Phi_{K,1} = E_K^{-1}A_K \cdots E_2^{-1}A_2E_1^{-1}A_1$ is the so-called *monodromy matrix* of the periodic system. For more details on the relations between generalized eigenvectors and eigenvalues of these pencils we refer to [22]. A key decomposition for computing these generalized eigenvalues and eigenvectors is described in the next section.

3.2. PERIODIC SCHUR FORM

The role played by the generalized Schur form for time invariant systems is now replaced by a very similar orthogonal decomposition, called the *periodic Schur form.*

Theorem 3 [3]
Let the $n \times n$ matrices E_k and A_k, $k = 1, \ldots, K$ be such that the pencil $\lambda\mathcal{E} - \mathcal{A}$ is regular. Then there always exist orthogonal transformations Q_k and Z_k, $k = 1, \ldots, K$ such that

$$\begin{bmatrix} Q_1^T & & & \\ & Q_2^T & & \\ & & \ddots & \\ & & & Q_K^T \end{bmatrix} \begin{bmatrix} -A_1 & \lambda E_1 & & \\ & \ddots & \ddots & \\ & & -A_{K-1} & \lambda E_{K-1} \\ \lambda E_K & & & -A_K \end{bmatrix} \begin{bmatrix} Z_1 & & & \\ & Z_2 & & \\ & & \ddots & \\ & & & Z_K \end{bmatrix}$$

$$= \begin{bmatrix} -\hat{A}_1 & \lambda \hat{E}_1 & & \\ & \ddots & \ddots & \\ & & -\hat{A}_{K-1} & \lambda \hat{E}_{K-1} \\ \lambda \hat{E}_K & & & -\hat{A}_K \end{bmatrix}, \tag{26}$$

where the transformed matrices $\hat{A}_k$ and $\hat{E}_k$ are all upper triangular, except for one matrix – say, $\hat{A}_1$ – which is quasi triangular. ■

The relation with the standard Schur form is that if the E_k matrices are invertible, then the monodromy matrix $\Phi_{K,1}$ is transformed by the orthogonal similarity Z_1 to its Schur form :

$$\hat{\Phi}_{K,1} \doteq \hat{E}_K^{-1}\hat{A}_K \cdots \hat{E}_1^{-1}\hat{A}_1 = Z_1^T(E_K^{-1}A_K \cdots E_1^{-1}A_1)Z_1 = Z_1^T\Phi_{K,1}Z_1.$$

Since all matrices are triangular it follows that all transformed monodromy matrices $\hat{\Phi}_{K+k-1,k}$ are quasi triangular as well, and with the same ordering of eigenvalues. From Theorem 1 it follows that the ordering of the eigenvalues can be chosen arbitrarily and hence that there exists a periodic Schur form associated with any eigenvalue ordering.

We point out here that the transformations Z_k and Q_k can also be applied directly to the system (16). Define indeed a new state $\hat{x}_k \doteq Z_k^T x_k$ and multiply the top equation of (16) by Q_k^T then we obtain an equivalent system

$$\left\{ \begin{array}{rcl} \hat{E}_k\hat{x}_{k+1} & = & \hat{A}_k\hat{x}_k + \hat{B}_k u_k \\ y_k & = & \hat{C}_k\hat{x}_k + D_k u_k, \end{array} \right. \tag{27}$$

where $\hat{B}_k \doteq Q_k^T B_k$, $\hat{C}_k \doteq C_k Z_k$, and $\hat{E}_k$ and $\hat{A}_k$ are upper triangular, except for $\hat{A}_1$ which is quasi triangular. This is a very special coordinate system : the "bottom" equation in (27) (or the bottom 2 equations if $\hat{A}_1$ has a bottom 2×2 block) is now "decoupled" from the rest of the system. Since the ordering of the eigenvalues in the Schur form can always be chosen arbitrarily, one can choose this decoupled system to be the one with the smallest eigenvalue in absolute value and hence the "easiest" to integrate numerically [3]. Once this "bottom" component of the state has been computed, one substitutes this in the next component(s), which is then also decoupled from the rest of the system, and so on. This coordinate system is therefore very appealing for simulation purposes [11].

3.3. PERIODIC CONTROL SYSTEMS

The periodic Schur form has several other applications in control problems involving periodic discrete time systems. In optimal control of such a periodic system one considers e.g. the problem :

$$\begin{array}{c} \text{Minimize } J = \sum_{k=1}^{\infty} z_k^T Q_k z_k + u_k^T R_k u_k \\ \text{subject to } H_k z_{k+1} = F_k z_k + G_k u_k \end{array} \tag{28}$$

where the matrices Q_k, R_k, F_k, G_k, H_k are periodic with period K. In order to solve this variational problem, one needs to solve the Hamiltonian equations which is a periodic homogeneous system of difference equations (16) in the state z_k and co-state λ_k of the system [16]. The correspondences with (16) are :

$$x_k \doteq \begin{bmatrix} \lambda_k \\ z_k \end{bmatrix}, E_k \doteq \begin{bmatrix} -G_k R_k^{-1} G_k^T & H_k \\ F_k^T & 0 \end{bmatrix}, A_k \doteq \begin{bmatrix} 0 & F_k \\ H_k^T & Q_k \end{bmatrix}. \tag{29}$$

For finding the periodic solutions to the underlying periodic Riccati equation one has to find the stable invariant subspaces of the monodromy matrices $\Phi_{K+k-1,k}$ [2]. Clearly the periodic Schur form is useful here as well as the reordering of eigenvalues [3].

In pole placement of periodic systems, again the periodic Schur form and reordering is useful when one wants to extend Varga's pole placement algorithm [27] to periodic systems. Consider the system

$$\begin{array}{c} E_k z_{k+1} = A_k z_k + D_k u_k \\ \text{with state feedback } u_k = F_k z_k + v_k \end{array} \tag{30}$$

where the matrices A_k, E_k, D_k, F_k are periodic with period K. This results in the closed loop system

$$E_k z_{k+1} = (A_k + D_k F_k) z_k + D_k v_k \tag{31}$$

of which the underlying time invariant eigenvalues are those of the matrix :

$$\Phi_{K,1}^{(F)} \doteq E_K^{-1}(A_K + D_K F_K) \cdots E_2^{-1}(A_2 + D_2 F_2) E_1^{-1}(A_1 + D_1 F_1). \tag{32}$$

In the above equation it is not apparent at all how to choose the matrices F_k to assign particular eigenvalues of $\Phi_{K,1}^{(F)}$. Yet when the matrices A_k, E_k are in the triangular form (26), one can choose the F_k matrices to have only nonzero elements in the last column. This will preserve the triangular form of the matrices $A_k + D_k F_k$ and it is then trivial to choose e.g. one such column vector to assign one eigenvalue. In order to assign the other eigenvalues one needs to *reorder* the diagonal elements in the periodic Schur form and each time assign another eigenvalue with the same technique. For more details, we refer to [17].

Other applications of the periodic Schur form are the solution of periodic Lyapunov and Sylvester equations. Since these are special cases of periodic Riccati equations, they can also be solved via the periodic Schur form. These equations show up in problems of stability analysis, decoupling, balancing and so on [17], [28].

3.4. GENERALIZED QR DECOMPOSITION

The basic equation in the matrix sequences occurring in the homogeneous two point boundary value problem defined earlier is $E_k x_{k+1} = A_k x_k$. We now analyze the system of equations when both E_k and A_k are singular, and more specifically, what is needed to be able to define singular values of sequences of such equations. For this we first analyze a single equation

$$Ey = Ax. \tag{33}$$

The singular value decomposition is originally defined for a single matrix M and is closely related to its URV decomposition. If an $m \times n$ matrix M has rank $r = \min\{m, n\}$ then there exists orthogonal matrices U and V such that

$$M = U \begin{bmatrix} \tilde{M} & 0 \\ 0 & 0 \end{bmatrix} V^T$$

where $\tilde{M}$ is an $r \times r$ invertible matrix. This so-called URV decomposition can be viewed as a two sided QR decomposition and can be obtained in a finite number of operations [7]. The first r columns of U in fact span the image of M and the first r columns of V are the orthogonal complement of the kernel of M. So if in the equation

$$y = Mx$$

we constrain the vector y to $\mathrm{Im}M$ and x to $\mathrm{Ker}M^\perp$, then M is invertible since it essentially is represented by $\tilde{M}$. This can be applied to the implicit system (33). If A and E are invertible then

$$Ey = Ax \Leftrightarrow y = E^{-1}Ax \Leftrightarrow A^{-1}Ey = x.$$

Do these equations still make sense when the matrices are singular ? Let us consider spaces $\mathcal{X}$ and $\mathcal{Y}$ such that $\mathcal{Z} = E\mathcal{Y}$, $\mathcal{Z} = A\mathcal{X}$ and $\dim \mathcal{X} = \dim \mathcal{Y} = \dim \mathcal{Z}$. The largest such subspace Z will be $\mathcal{Z} \doteq ImA \cap ImE$. Provided we choose $x \in \mathcal{X}$ then there exists a solution $y \in \mathcal{Y}$ satisfying $Ey = Ax$, and then the mapping $y = E^{-1}Ax$ makes sense. But the solutions are not unique, unless we constrain y to be e.g. in $\mathrm{Ker}E^{\perp}$. Conversely, provided we choose $y \in \mathcal{Y}$ then there exists a solution $x \in \mathcal{X}$ satisfying $Ey = Ax$. Again $A^{-1}Ey = x$ then make sense and x is unique e.g. if $x \in \mathrm{Ker}A^{\perp}$.

This discussion does not change when transforming to a coordinate system

$$z_q = QEV^T y_v = QAU^T x_u$$

where U, V and Q are orthogonal. One can always find matrices Q, U, V such that in the new coordinate system this equation has the form :

$$\begin{bmatrix} z_1 \\ z_2 \\ z_3 \\ z_4 \end{bmatrix} = \begin{bmatrix} 0 & \tilde{E} & \times \\ 0 & 0 & \times \\ 0 & 0 & R_s \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = \begin{bmatrix} 0 & \tilde{A} & \times \\ 0 & 0 & R_t \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \tag{34}$$

where R_s, R_t are invertible matrices of dimension s and t, respectively, and $\tilde{E}$ and $\tilde{A}$ are square invertible of dimension $r = \dim ImA \cap ImE$. In the new coordinate system it is clear that we need to take $y_3 = 0$, $x_3 = 0$ in order to make the system compatible and $y_1 = 0$, $x_1 = 0$ in order to make the solution unique. Defining $\tilde{x} = x_2$ and $\tilde{y} = y_2$ we obtain the equation

$$\tilde{E}\tilde{y} = \tilde{A}\tilde{x},$$

which has a well defined solution $\tilde{y} = \tilde{E}^{-1}\tilde{A}\tilde{x}$. The singular values we are interested in are of course those of the reduced order map $\tilde{E}^{-1}\tilde{A}$. The decomposition (34) is called the generalized QR decomposition and can be extended to sequences of equations $E_k x_{k+1} = A_k x_k$. As in the single equation case, one can again extract from a possibly singular system of equations a lower dimensional one that has all matrices $\tilde{E}_k$ and $\tilde{A}_k$ nonsingular. The relevant singular values are then those of

$$\tilde{E}_K^{-1}\tilde{A}_K \cdots \tilde{E}_1^{-1}\tilde{A}_1.$$

For the details, we refer to [4]. In the next section we show how to extract from such a nonsingular sequence, the singular values by only applying orthogonal transformations to the sequence of matrices $\tilde{E}_k$, $\tilde{A}_k$.

3.5. BIDIAGONAL AND SINGULAR VALUE DECOMPOSITION

Here we consider a sequence of $n \times n$ matrices E_k, A_k which are invertible, and we want to compute the singular values of the state transition matrix over a time interval $[1, N]$.

$$\Phi_{N,1} \doteq E_N^{-1} A_N \cdots E_2^{-1} A_2 E_1^{-1} A_1.$$

It is clear that one has to perform left and right transformations on $\Phi_{N,1}$ to diagonalize $U^T \Phi_{N,1} V = \Sigma$, but these will only affect E_N and A_1. But one can insert pairs of orthogonal transformations in between the factors of this expression without altering the result. The following theorem shows how to use these degrees of freedom to find the singular values of $\Phi_{N,1}$.

Theorem 4 [8]
Let the $n \times n$ matrices E_k and A_k, $k = 1, \ldots, N$ be invertible. Then there always exist orthogonal transformations Q_k, $k = 1, \ldots, N$ and Z_k, $k = 1, \ldots, N+1$ such that

$$\begin{bmatrix} Q_1^T & & \\ & \ddots & \\ & & Q_N^T \end{bmatrix} \begin{bmatrix} -A_1 & E_1 & & \\ & \ddots & \ddots & \\ & & -A_N & E_N \end{bmatrix} \begin{bmatrix} Z_1 & & & \\ & \ddots & & \\ & & Z_N & \\ & & & Z_{N+1} \end{bmatrix}$$

$$= \begin{bmatrix} -\hat{A}_1 & \hat{E}_1 & & \\ & \ddots & \ddots & \\ & & -\hat{A}_N & \hat{E}_N \end{bmatrix}, \tag{35}$$

where all matrices $\hat{E}_k$ and $\hat{A}_k$ are upper triangular and the product $\hat{\Phi}_{N,1} \doteq \hat{E}_N^{-1} \hat{A}_N \cdots \hat{E}_1^{-1} \hat{A}_1$ is diagonal (alternatively, there is an algorithm of complexity $O(Nn^3)$ which *bidiagonalizes* $\hat{\Phi}_{N,1}$). ■

The proof of this result is very simple. Choose $Z_1 = V$ and $Z_{N+1} = U$. Then alternately choose the matrices Q_k to triangularize $\hat{A}_k = Q_k^T (A_k Z_k)$ (for $k = 1, \ldots, N$) and Z_k to triangularize $\hat{E}_{k-1} = (Q_{k-1}^T E_{k-1}) Z_k$ (only for $k = 2, \ldots, N$ since Z_1 and Z_{N+1} are already fixed). All matrices but $\hat{E}_N$ in the expression for $\hat{\Phi}_{N,1}$ are now upper triangular by construction, but since all factors are invertible and the product is diagonal (or bidiagonal), $\hat{E}_N$ must be upper triangular as well. The finite algorithm for the bidiagonalization could also be derived this way, but we refer to [8] for a constructive and numerically stable algorithm.

The bidiagonalization has been shown to yield very accurate results despite the fact that the singular values of such product can become very large and very small as N tends to grow [8]. The singular values of the computed bidiagonal are then computed to high relative accuracy using an appropriate technique [6]. This decomposition can e.g. be used to find the "dominant directions" of the state transition map $\Phi_{N,1}$ over a finite time interval $[1, N]$. In the case that the discrete time system (16) comes from a discretization of a nonlinear continuous time system it is known that the singular values of $\Phi_{N,1}$ are closely related to the Lyapunov exponents of the corresponding continuous time system (provided the discretization is sufficiently "fine") [5]. Singular values also show up in robustness aspects of dynamical systems [24].

4. Concluding remarks

In this paper we surveyed a number of orthogonal matrix decompositions arising in systems and control. We pointed out that they lead to numerically reliable algorithms for solving quite a large range of problems in this area. This is mainly due to the fact that orthogonal transformations are backward stable when implemented correctly, and that they do not affect the conditioning of the problem at hand. Although we did not mention all uses of orthogonal transformations in this area, we gave references for further reading on this important area of research.

Acknowledgments

This work was supported by the National Science Foundation under grant CCR-9796315. Parts of this paper presents research results of the Belgian Programme on Inter-university Poles of Attraction, initiated by the Belgian State, Prime Minister's Office for Science, Technology and Culture. The scientific responsibility rests with its authors.

References

1. R.H. Bartels and G.W. Stewart (1972), Solution of the equation $AX + XB = C$, *Comm. ACM* **15**, pp. 820–826.
2. S. Bittanti, P. Colaneri and G. de Nicolao (1988), The difference periodic Riccati equation for the periodic prediction problem, *IEEE Trans. Aut. Contr.* **33**, pp. 706–712.
3. A. Bojanczyk, G. Golub and P. Van Dooren (1992), The periodic Schur decomposition. Algorithms and applications, in *Advanced Signal Processing Algorithms, Architectures, and Implementations III*, Proceedings of SPIE, the International Society for Optical Engineering **1770**, pp. 31–42, Ed. F.T. Luk, Bellingham, Wash., USA.
4. B. De Moor and P. Van Dooren (1992), Generalizations of the singular value and QR decomposition, *SIAM Matr. Anal. & Applic.* **13**, pp. 993–1014.
5. L. Dieci, R. Russell and E. Van Vleck (1997), On the computation of Lyapunov exponents for continuous dynamical systems, *SIAM Numer. Anal.* **34**, pp. 402–423.

6. K.V. Fernando and B.N. Parlett (1994), Accurate singular values and differential qd algorithms, *Numerische Mathematik* **67**, pp. 191–229.
7. G. Golub and C. Van Loan (1989), *Matrix Computations*, 2nd edition, The Johns Hopkins University Press, Baltimore, Maryland.
8. G. Golub, K. Solna and P. Van Dooren (1995), Computing the singular values of products and quotients of matrices, in *SVD in Signal Processing III, Algorithms, Architectures and Applications*, Eds. M. Moonen, B. De Moor, Elsevier, Amsterdam, 1995, also *SIAM Matrix Anal. & Appl.*, to appear.
9. G. Golub and V. Kahan (1965), Calculating the singular values and pseudo-inverse of a matrix, *SIAM Numer. Anal.* **2**, pp. 205–224.
10. A. Laub (1990), Invariant subspace methods for the numerical solution of Riccati equations, in *The Riccati equation*, Eds. Bittanti, Laub, Willems, Springer Verlag.
11. R.M. Mattheij and S.J. Wright (1993), Parallel stabilized compactification for ODEs with parameters and multipoint conditions, *Applied Numerical Mathematics* **13**, pp. 305–333.
12. C. Moler and G. Stewart (1973), An algorithm for the generalized matrix eigenvalue problem, *SIAM J. Numer. Anal.* **10**, pp. 241–256.
13. M. Moonen, B. De Moor, L. Vandenberghe, and J. Vandewalle (1989), On- and off-line identification of linear state-space models, *Int. J. Contr.* **49**, pp. 219–232.
14. B.C. Moore (1981), Principal component analysis in linear systems: Controllability, observability, and model reduction, *IEEE Trans. Aut. Contr.* **26**, pp. 17–31.
15. R. Patel, A. Laub and P. Van Dooren (1993), *Numerical Linear Algebra Techniques for Systems and Control*, (Eds.) IEEE Press, Piscataway NJ.
16. A. Sage and C. White (1977), *Optimum Systems Control*, 2nd Ed., Prentice-Hall, New Jersey.
17. J. Sreedhar and P. Van Dooren (1994), A Schur approach for solving some periodic matrix equations, in *Systems and Networks : Mathematical Theory and Applications*, pp. 339–362, Eds. U. Helmke, R. Mennicken, J. Saurer, Mathematical Research **77**, Akademie Verlag, Berlin.
18. J. Sreedhar and P. Van Dooren (1999), Periodic descriptor systems: Solvability and conditionability, *IEEE Trans. Aut. Contr.*, to appear.
19. G.W. Stewart and J.-G. Sun (1990), *Matrix Perturbation Theory*, Academic Press, New York.
20. P. Van Dooren (1981), The generalized eigenstructure problem in linear system theory, *IEEE Trans. Aut. Contr.* **26**, pp. 111–129.
21. P. Van Dooren (1981), A generalized eigenvalue approach for solving Riccati equations, *SIAM Sci. & Stat. Comp.* **2**, pp. 121–135.
22. P. Van Dooren (1999), Two point boundary value and periodic eigenvalue problems, in *Proceedings 1999 IEEE CACSD Symposium*, Hawaii, to appear.
23. P. Van Dooren (1999), The singular value decomposition and its applications in systems and control, in *Wiley Encyclopedia of Electrical and Electronics Engineering*, Wiley and Sons, New York, to appear.
24. P. Van Dooren and V. Vermaut (1997), On stability radii of generalized eigenvalue problems, in *Proceedings European Control Conf.*, paper FR-M-H6.
25. P. Van Dooren and M. Verhaegen (1985), On the use of unitary state-space transformations, *Linear Algebra and its Role in Linear System Theory*, Ed. B.N. Datta, pp. 447–463, Providence, RI: AMS Contemporary Mathematics Series, **47**.
26. P. Van Dooren and M. Verhaegen (1988), Condensed forms for efficient time invariant Kalman filtering, *SIAM Sci. & Stat. Comp.* **9**, pp. 516–530.
27. A. Varga (1981), A Schur method for pole assignment, *IEEETrans. Aut. Contr.* **26**, pp. 517–519.
28. A. Varga (1997), Periodic Lyapunov equations : some applications and new algorithms, *Int. J. Contr.* **67**, pp. 69–87.