Abstract

In order reduction of high order linear time invariant systems based on two-sided Krylov subspace methods, the Lanczos algorithm is commonly used to find the bases for input and output Krylov subspaces and to calculate the reduced order model by projection. However, this method can be numerically unstable even for systems with moderate number of states and can only find a limited number of biorthogonal vectors. In this paper, we present another method which we call Two-Sided Arnoldi and which is based on the Arnoldi algorithm. It finds two orthogonal bases for any pair of Krylov subspaces, with one or more starting vectors. This new method is numerically more robust and simpler to implement specially for nonsquare MIMO systems, and it finds a reduced order model with same transfer function as Lanczos. The Two-Sided Arnoldi algorithm can be used for order reduction of the most general case of a linear time invariant Multi-Input-Multi-Output (MIMO) systems. Furthermore, we present some suggestions to improve the method using a column selection procedure and to reduce the computational time using LU-factorization.

Key Words: Two-sided Arnoldi, Order reduction, Krylov subspace, Moment matching, Large scale systems.

1 Introduction

For the reduction of very high order systems (with application to circuit simulation and micro electro mechanical systems, for instance), Krylov subspace methods are probably one of the best choices today, e.g. [3, 5]. They define a projection from the high dimensional space of the original model to a lower dimensional space and vice versa and thereby find the reduced order model and match some of the characteristic parameters of the original and reduced order systems.

The system to be modelled is typically defined via a set of state equations

\[
\begin{align*}
E \dot{x}(t) &= Ax(t) + Bu(t), \\
y(t) &= Cx(t),
\end{align*}
\]  

(1)

where \( E, A \in \mathbb{R}^{n \times n} \), \( B \in \mathbb{R}^{n \times m} \) and \( C \in \mathbb{R}^{p \times n} \) are given matrices and the components of the vector valued functions \( u \in \mathbb{R}^{m}, y \in \mathbb{R}^{p} \) and \( x \in \mathbb{R}^{n} \) are the inputs, outputs and states of the system respectively. For single-input single-output (SISO) systems, \( p = m = 1 \) and matrices \( B \) and \( C \) reduce to vectors \( b \) and \( c^T \). Throughout this paper, it is assumed that \( A \) is nonsingular.

For system (1), moments (around zero) are defined as follows

\[ M_i = C(A^{-1}E)^iA^{-1}B, \quad i = 0, 1, \ldots. \]  

(2)

\( M_i \) is a \( p \times m \) matrix in multi input multi output (MIMO) case and is a scalar in single input single output (SISO) case. The \( M_i \)'s are related to the coefficients of a Taylor series expansion around zero for the transfer function of system (1) [5, 9].

The block Krylov subspace is defined

\[ K_q(A_1,B_1) = \text{colspan}\{B_1,A_1B_1,\ldots,A_1^{q-1}B_1\}, \]  

(3)
where $A_1 \in \mathbb{R}^{n \times n}$ and $B_1 \in \mathbb{R}^{n \times m}$, and starting vectors are located in the columns of the matrix $B_1$. The vectors which span the subspace are called the basic vectors.

Considering the state space representation (1), two Krylov subspaces $K_q(A^{-1}E, A^{-1}B)$ and $K_q(A^{-T}E^T, A^{-T}C^T)$ both with the same rank are used for model order reduction and are called input and output Krylov subspaces respectively.

In the next section, we will first recall order reduction by projection. In so-called one-sided methods, the most commonly used algorithm is Arnoldi, as introduced in section 3. Section 4 then presents the new two-sided Arnoldi approach, followed by some important implementation aspects in sections 5 to 7. Two application examples prove the ease of use and the numerical robustness of the method, and we conclude by a short comparison of one-sided Arnoldi, Lanczos, and two-sided Arnoldi.

2 Order reduction using Krylov subspace method

Consider a projection as follows:

$$\begin{align*}
    x &= Vx_r, \\
    V &\in \mathbb{R}^{n \times q}, x \in \mathbb{R}^n, x_r \in \mathbb{R}^q,
\end{align*}$$

where $q < n$. By applying this projection to system (1) and then multiplying the state equation by the transpose of some matrix $W \in \mathbb{R}^{n \times q}$, a model with reduced order $q$ can be found:

$$\begin{align*}
    \begin{cases}
        W^T E V x_r(t) &= W^T A V x_r(t) + W^T B u(t), \\
        y &= C V x_r(t).
    \end{cases}
\end{align*}$$

From this, the reduced order system in state space can be identified by the following matrices:

$$\begin{align*}
    E_r &= W^T E, & A_r &= W^T A, \\
    B_r &= W^T B, & C_r &= C V.
\end{align*}$$

In the reduced order model (5), an approximation of the full state vector $x$ can be found from $x_r$ by $\hat{x} = V x_r$. It can be proved that when $V$ is a basis of input Krylov subspace or $W$ is a basis of output Krylov subspace then the first $\frac{q}{m}$ or $\frac{q}{n}$ moments match respectively and the method is called one-sided Krylov method. When both $V$ and $W$ are bases of input and output Krylov subspaces then the first $\frac{q}{m} + \frac{q}{n}$ moments match and the method is called two-sided Krylov method [9].

**Theorem 1** In using two-sided Krylov subspace methods, the transfer function of the reduced order system (5) is independent of the particular choice of the bases $V$ and $W$ of the Krylov subspaces $K_q(A^{-1}E, A^{-1}B)$ and $K_q(A^{-T}E^T, A^{-T}C^T)$ [9].

According to theorem 1, any method that can find a pair of bases for input and output Krylov subspaces leads to a reduced order model with the same input-output behaviour.

3 Arnoldi algorithm

In many applications, it is desired to find an orthogonal basis for a given subspace to avoid numerical problems. The classical Arnoldi [1] procedure finds a set of orthonormal vectors that can be considered as a basis for a given Krylov subspace with one starting vector. To generalize this method for more than one starting vector, one way is using a block Arnoldi algorithm [8]. In the following, we briefly describe a version of the Arnoldi algorithm for more than one starting vector $[b_1 \cdots b_m] = B$, in which the resulting matrix of basis vectors is orthonormal.

Consider the Krylov subspace $K_q(A_1, B_1)$ with $m$ starting vectors. The algorithm finds a set of vectors with length one that are orthogonal to each other,

$$V^T V = I,$$

where the columns of the matrix $V$ are the basis for the given Krylov subspace. In the following algorithm, in each step one more vector orthogonal to all other previous vectors is constructed and then it is normalized to have length one. In a general case, when $q$ is not small enough, it can happen that not all of the basic vectors are linearly independent. Then, linearly dependent vectors must be deleted during the iterations (=deflation).
Algorithm 1 Arnoldi algorithm with deflation using modified Gram-Schmidt [6]:

0. Start: Delete all linearly dependent starting vectors to find \( m_1 \) independent starting vectors for the given Krylov subspaces then set 
\[
v_1 = \frac{b_1}{\sqrt{b_1^T b_1}},
\]
where \( b_1 \) is the first starting vector after deleting the dependent starting vectors.

1. For \( i = 2, 3, \cdots \), do,
   (a) Calculating the next vector: If \( i \leq m_1 \) the next vector is the \( i \)-th starting vector. Otherwise, the next vector is
   \[
r_i = A_1 v_{i-m_1}.
   \]
   (b) Orthogonalization: Set \( \hat{v}_i = r_i \) then for \( j=1 \) to \( i-1 \) do:
   \[
   h = \hat{v}_i^T v_j \tag{8}
   \]
   \[
   \hat{v}_i = \hat{v}_i - h v_j.
   \]
   (c) Normalization: If vector \( \hat{v}_i \) is zero, reduce \( m_1 \) to \( m_1 - 1 \) and if \( m_1 \) is nonzero go to step 1.a and if \( m_1 \) is zero break the loop. Otherwise, if \( \hat{v}_i \neq 0 \) the \( i \)-th column of matrix \( V \) is
   \[
v_i = \frac{\hat{v}_i}{\sqrt{\hat{v}_i^T \hat{v}_i}}.
   \]
   (d) Increase \( i \) and go to step 1.a.

By knowing that at each step of the Arnoldi algorithm, the \( i \)-th column of \( V \) is a linear combination of the first \( i \) basic vectors, it is not difficult to show that the vectors \( \{v_1, \cdots, v_i\} \) span the same space as the first \( i \) basic vectors of the given Krylov subspace.

One of the properties of a Krylov subspace is that, if the \( i \)-th basic vector related to starting vector \( b_j \) is a linear combination of all previous basic vectors, then all other basic vectors \( p_{i+m_1}, p_{i+2m_1}, \cdots \) related to starting vector \( b_j \) are linear combination of the first \( i-1 \) basic vectors. In step 1.b of algorithm 1, if \( \hat{v}_i \) is zero, it means that the \( i \)-th basic vector is a linear combination of the previous basic vectors and all basic vectors \( p_i, p_{i+m_1}, \cdots \) must be deleted in the next iterations. This can be done by going one step back and reducing the parameter \( m_1 \). If \( \hat{v}_i \) is nonzero, it means that the first \( i \) basic vectors are linearly independent because, as discussed above, \( \{v_1, \cdots, v_{i-1}, \hat{v}_i\} \) span the same space as the first \( i \) basic vectors and all of them are independent.

4 Two-sided Arnoldi algorithm

As mentioned in section 2, using any pair of bases of input and output Krylov subspaces in projection leads to match some of the first moments and finds a reduced system with the same transfer function. In most applications, the basic vectors used in the definition of Krylov subspaces tend to be almost linearly dependent even for moderate values of \( n \). So, they should not be used in numerical computations. Instead, there exist other suitable bases that can be applied in order reduction.

One of the ideas to avoid the numerical problems is orthogonalizing the vectors. By using biorthogonalization, the Lanczos algorithm [7] generates two sequences of basis vectors which span the input and output Krylov subspaces. Unfortunately, in practice this algorithm can only find a very limited number of biorthogonal basis vectors in many applications. In this section we introduce another method called Two-Sided Arnoldi to find the bases necessary for projection and finding the reduced order model. This new method in comparison to Lanczos is numerically more stable and easier to implement for the general representation of the systems specially for the nonsquare (\( m \neq p \)) MIMO systems. This method comprises the following steps:

Algorithm 2 Two-sided Arnoldi algorithm:

0. Choose the appropriate input and output Krylov subspaces.
1. Apply Arnoldi algorithm 1 to the input Krylov subspace to find the matrix \( V \).

2. Apply Arnoldi algorithm 1 to the output Krylov subspace to find the matrix \( W \).

3. Find the reduced order model by applying projection to the original system the same as equation 5.

Suppose that for a system with \( m \) inputs and \( p \) outputs, the reduced model with order \( q \) is desired.

In this case each moment has \( m \cdot p \) entries. Each column of matrices \( V \) and \( W \) leads to match one more row or column of the moment matrices. So, there is no need to increase \( q \) such that it is a multiple of both, \( m \) and \( p \). In two-sided methods \( m \cdot q + p \cdot q \) entries of the characteristic parameters of the original and reduced order models match.

5 Selection procedure

The moments of the MIMO system (1) are a \( p \times m \) matrices \( M_i \), where each column is related to an input and each row is related to an output of the system. As will be shown in this section, it is possible to match individual columns (and rows) up to a higher index \( i \) than other columns (and rows). In this way, for more important entries of the moment matrices the number of matched parameters is increased and for less important entries of the moment matrices the number of matched parameters is decreased, while the sum of the matched parameters is constant. This concept can be more useful when the number of inputs and outputs are high.

Consider only an input Krylov subspace. For a system with \( m \) inputs, there exist \( m \) starting vectors that must be orthogonalized and normalized to each other, and they produce the first \( m \) columns of matrix \( V \), \( v_1, \ldots, v_m \) see figure 1. For the next vector, there are \( m \) candidates, \( A^{-1}Ev_1, \ldots, A^{-1}Ev_m \), called \( r_1, \ldots, r_m \), that can be chosen to construct the vector \( v_{m+1} \). One of the best choices between \( r_1, \ldots, r_m \) is the vector which is ”most” linearly independent of all previous vectors. This vector can be identified in different ways. One geometrical dominance measure is the angle between the candidate vectors and the hyperspace constructed by all previous vectors. The sine of this angle can easily be calculated within Arnoldi algorithm as the division of the norm of the next vector before normalization (\( \hat{v} \) in Arnoldi algorithm 1) by the norm of the vector \( r_i, i = 1 \cdots m \). So in algorithm 1, \( d_i = \frac{\text{norm}(\hat{v}_i)}{\text{norm}(r_i)} \) can be used as a dominance measure for the \( m \) candidate vectors \( r_1, \ldots, r_m \).

Another dominance measure can be calculated using singular value decomposition (SVD) \([6]\), to find out ”how much” the vector \( r_i, i = 1 \cdots m \), is linearly dependent on the vectors \( v_1, \ldots, v_m \). To do that, SVD of the matrix \( [ \begin{array}{c} \mathbf{v}_1 \\ \vdots \\ \mathbf{v}_m \\ \mathbf{r} \end{array} ] \) must be calculated. The smallest singular value of this matrix shows the weight of the linear dependency of the columns. It must be noted that calculating this SVD is not too costly, since only the eigenvalues of a small symmetric \((m+1) \times (m+1)\) matrix must be calculated. For proper comparison, it is good to normalize all vectors \( r_i \) before calculating the SVD.

After calculating the dominance measures for \( m \) candidates, the vector with the largest measure can be chosen for becoming the next column of matrix \( V \). Suppose that it is \( r_j \), then by orthogonalization and normalization, vector \( v_{m+1} \) is calculated as described in Arnoldi algorithm 1. Then, the vector \( r_j \) must be substituted by the next candidate that is \( A^{-1}Ev_{m+1} \) and the dominance measures must be renewed for all candidates, using \( m + 1 \) columns of matrix \( V \), and so on, as illustrated in figure 1.

By this selection procedure, we may hope to choose those columns ”most” contributing to the controllability subspace, in each iteration. Applying the method to the output Krylov subspace can be done in the corresponding way, building up an approximate observability subspace.

6 Inverse of a large matrix

In order reduction of system (1) based on moment matching the matrix \( A^{-1}E \) and \( A^{-T}E^T \) are very important. So, the inverse of the large matrix \( A \) seems to be necessary. Calculating this inverse and then use it in the algorithm 1 is not recommended for numerical reasons. In order to find the vector \( x = A^{-1}Ev_{i-1} \), it is better to solve the linear equation \( Ax = Ev_{i-1} \) for \( x \) in each iteration to avoid numerical errors and to save time. (In this way, a total of \( q \) sets of linear equations are solved, whereas the calculation of matrix \( A^{-1} \) would require \( n \) sets of such equations.)

There exist many methods that can solve a linear equation and find an exact or approximate solution. One of the best method to find an exact solution is using LU-factorization \([6]\) and then solve two triangular linear equations by Gaussian elimination. Using this method in each iteration leads to a slow algorithm.
while the most time consuming part is finding the LU factorization of the large matrix. The remedy is finding the LU factorization of the matrix $A$ at the beginning and then solve only triangular linear equations in each iteration \(^1\). In this case, time is saved and the result is obtained very fast.

7 Generalization to rational Krylov method

The Two-Sided Arnoldi algorithm discussed in the previous sections was introduced for matching the moments around zero which leads to approximate the slow dynamics of the original system. However, sometimes it is advantageous to match the moments around points different from zero, \([10]\).

The results of this paper can be extended to matching the moments around $s_0 \neq 0$, called rational Krylov subspace method. The moments of the system (1) around $s_0$ are

$$M_i = C((A - s_0E)^{-1}E)^i(A - s_0E)^{-1}B,$$

(9)

where $i = 0, 1, \ldots$. By comparing equations (9) and (2), it can be seen that the matrix $A$ has been substituted by $A - s_0E$. If the subspaces $K_{q_1}((A - s_0E)^{-1}E, (A - s_0E)^{-1}B)$ and $K_{q_2}((A - s_0E)^{-1}E^T, (A - s_0E)^{-1}B^T)$ are used, the two methods are equivalent.

\(^1\)In MATLAB instead of the command $A\backslash v(:,i-1)$ the command $[L,U]=$LU$(A)$ must be added at the beginning of the algorithm and in each iteration the command $U\backslash(L\backslash v(:,i-1))$ can be used to solve triangular linear equations.
are considered, \( \frac{n}{m} + \frac{q}{p} \) moments around \( s_0 \) match. In two-sided Arnoldi, it is very simple to match the moments around two different points at the same time. By using subspaces \( K_{q_1}((A - s_0E)^{-1}E, (A - s_0E)^{-1}B) \) and \( K_{q_2}((A - s_1E)^{-T}E^T, (A - s_1E)^{-T}C^T) \), \( \frac{n}{m} \) moments around \( s_0 \) and \( \frac{q}{p} \) moments around \( s_1 \) match.

The next step to find the reduced order model is to find two bases for these Krylov subspaces. This can be done using the two-sided Arnoldi algorithm. The projection is then applied to the model (1), as described in equations (5) and (6) (i.e. \( A \) in equation (6) is not substituted by \( A - s_0E \)). With \( s_0 = 0 \), the reduced and original model have the same DC gain and steady state accuracy is achieved. Small values of \( s_0 \) will also find a reduced model with good approximation of slow dynamics.

8 Numerical examples

In order to illustrate the way that the algorithm works, it is applied to two different practical examples. The first example is the model of CD player\(^2\). The control task is to achieve track following, which basically amounts to pointing the laser spot to the track of pits on the CD that is rotating. The mechanism treated here, consists of a swing arm on which a lens is mounted by means of two horizontal leaf springs, figure 2. The rotation of the arm in the horizontal plane enables reading of the spiral-shaped disc tracks, and the suspended lens is used to focus the spot on the disc. Due to the fact that the disc is not perfectly flat, and due to irregularities in the spiral of pits on the disc, the challenge is to find a lowcost controller that can make the servosystem faster and less sensitive to external shocks [4]. So, finding a low order approximation of such a system seems to be necessary.

Figure 2: Schematic view of a rotating arm Compact Disc mechanism.

The model contains 120 states with 2 inputs and 2 outputs. The two-sided Arnoldi algorithm is applied to this system in three different ways: (1) By matching all moments around \( s = 0 \), (2) by matching all moments around \( s = 11000 \), and (3) by matching half of the moments around \( s_0 = 0 \) and half of them around \( s_1 = 11000 \) (this was done by considering the output Krylov subspace at \( s_0 = 0 \) and the input Krylov subspace at \( s_1 = 11000 \)). Figure 3 shows the largest singular value of the original model together with the order 12 approximation for matching moments only around zero, \( s_1 = 11000 \), and both of them using two-sided Arnoldi algorithm. In figure 4 the step responses of the original model and the reduced order model are illustrated, showing a good approximation for two responses and some acceptable errors for the two others. By matching only moments around zero, the infinity- and two-norms of the relative error are 0.0011 and 0.0081 respectively. To match moments around two points the infinity and two norm of the relative error are 0.0308 and 0.0227 respectively. As a result, we can say by matching the moments at two points, the Bode diagram fits very well. However, the norm of the error when only the moments around zero match is a little better.

The second example is a class of high energy Micro-Electro-Mechanical System (MEMS) actuators which integrates solid fuel with three silicon micromachined wafers [2]. It delivers either an impulse-bit

\(^2\)The model is available online at http://www.win.tue.nl/niconet/NIC2/benchmodred.html.
thrust or pressure waves within a submillimeter volume of silicon, by producing a high amount of energy from an ignitable substance contained within the microsystem. The microthruster fuel is ignited by passing an electric current through a polysilicon resistor embedded in a dielectric membrane, as shown in figure 5. After the ignition phase, sustained combustion takes place and forms a high-pressure, high-temperature ruptures and an impulse is imparted to the carrier frame as the gas escape from the tank. A two-dimensional axi-symmetric model is used, which after finite element based spatial discretization of the governing equations results in a linear system of 1071 ordinary differential equations.

The model contains 1071 states with 1 input and 3 outputs; the only relevant input signal is the unit step function. The two-sided Arnoldi algorithm is applied to this system in two different ways to find a reduced model with order 8: (1), in the normal way (reduced model 1), and (2), by choosing the next vector between some candidates (reduced model 2) as described in section 5. The errors for the step response of the system for these two reduced models are shown in table 1. In figure 6 the step responses of the original model and the reduced order models are shown. In this example, using selection of the next vectors reduces the error significantly.

9 Conclusion

In this paper, we presented a new algorithm called two-sided Arnoldi that can be used in order reduction in two-sided Krylov subspace methods. This method uses Arnoldi algorithm twice, first for the calculation
of a basis $V$ of the input Krylov subspace, then for the calculation of a basis $W$ for the output Krylov subspace (the reduced model then is (5)). Because it matches more moments, application of the two-sided Arnoldi algorithm to different technical systems has led to better results than one-sided Arnoldi algorithm, while classical Lanczos algorithm failed for numerical reasons. No problems in evaluating (6) were observed. Table 2 compares three algorithms in Krylov subspace methods.

The most important properties of this algorithm in comparison to Lanczos are:

- It is simple to implement specially for MIMO nonsquare systems.
- It is numerically stable.
- It leads to reduced model with the same transfer function as Lanczos algorithm.

We also propose some methods to reduce the approximation error using a selection procedure and reducing the computational effort using LU-factorization.

| Table 2: Comparison of three different algorithms in order reduction using Krylov subspaces ($q$ is the order of reduced model). |
|---|---|---|---|
|       | One-sided Arnoldi | Two-sided Arnoldi | Lanczos |
| No. of moments that match | $q/m$ | $q/m + q/p$ | $q/m + q/p$ |
| Numerical stability | stable | stable | unstable |

10 Acknowledgement

This work was supported by German research council (DFG) which is gratefully acknowledged. We also would like to thank the authors of the site http://www.win.tue.nl/niconet/NIC2 for providing us the model.

References


