Applied Mathematical Sciences, Vol. 3, 2009, no. 38, 1849 - 1859

Model Reduction and Moment Matching

Ismail bin Mohd and Mustafa bin Mamat

Jabatan Matematika Fakulti Sains dan Teknologi
Universiti Malaysia Terengganu

Yosza bin Dasril

Jabatan Industrial Electrics Faculty of Electrics & Comp Engineering
Universiti Teknikal Malaysia, Melaka

Fudziah Ismail

Jabatan Matematika Fakulti Sains
Universiti Putra Malaysia

Farikhin

Jabatan Matematika Fakulti Sains dan Teknologi
Universiti Malaysia Terengganu
and
Jurusan Matematika FMIPA Undip Semarang Indonesia
farikhin2@yahoo.com

Abstract

In this paper, we proposed a simple way to find model reduction of dynamical system $\begin{cases} \frac{d}{dt}x(t) = \mathbf{A}\mathbf{x}(t) + \mathbf{B}u(t) \\ y(t) = \mathbf{C}^{\mathbf{T}}x(t) \end{cases}$ where x: $\mathbf{R} \rightarrow \mathbf{R}^n$ is a state vector, u: $\mathbf{R} \rightarrow \mathbf{R}^p$ is input function, y: $\mathbf{R} \rightarrow \mathbf{R}^q$ is a output function, $\mathbf{A} \in \mathbf{R}^{n \times n}$, $\mathbf{B} \in \mathbf{R}^{n \times p}$, and $\mathbf{C} \in \mathbf{R}^{n \times q}$ are the system matrices. Furthermore, we show that error output of single input single output system can be estimated over a certain class of input functions.

Mathematics Subject Classification: 93A30, 34C20

Keywords: model reduction, Krylov subspace, Arnoldi algorithm and moment matching

1. Introduction

Consider dynamical system

$$\begin{cases} \frac{d}{dt}x(t) = \mathbf{A}\mathbf{x}(t) + \mathbf{B}u(t) \\ y(t) = \mathbf{C}^{\mathbf{T}}x(t) \end{cases} \quad (1)$$

x: $\mathbf{R} \to \mathbf{R}^n$ is a state vector, u: $\mathbf{R} \to \mathbf{R}^p$ is input function, y: $\mathbf{R} \to \mathbf{R}^q$ is a output function, $\mathbf{A} \in \mathbf{R}^{n \times n}$, $\mathbf{B} \in \mathbf{R}^{n \times p}$, and $\mathbf{C} \in \mathbf{R}^{n \times q}$ are the system matrices. Matrix $\mathbf{A}$ is allowed to be a singular matrix and we assumed that the matrix $s\mathbf{I} - \mathbf{A}$ is not singular. In most practical case,we have $n > p > q$ [1-6, 9-11 and 13].

Model reduction is a procedure to find reduced-order model, say of r-th order

$$\begin{cases} \frac{d}{dt}x_r(t) = \mathbf{A}_r\mathbf{x}_r(t) + \mathbf{B}_r u_r(t) \\ y_r(t) = \mathbf{C}_r^{\mathbf{T}}x_r(t) \end{cases} \quad (2)$$

x_r : $\mathbf{R} \to \mathbf{R}^r$ is a state vector, u_r : $\mathbf{R} \to \mathbf{R}^p$ is input function, y_r : $\mathbf{R} \to \mathbf{R}^q$ is a output function, $\mathbf{A} \in \mathbf{R}^{r \times r}$, $\mathbf{B} \in \mathbf{R}^{r \times p}$, and $\mathbf{C} \in \mathbf{R}^{r \times q}$ and $r < n$, in such a way that the transfer functions are close in some sense. If $G(s) = \mathbf{C}(s\mathbf{I} - \mathbf{A})^{-1}\mathbf{B}$ and $G_r(s) = \mathbf{C}_r(s\mathbf{I} - \mathbf{A}_r)^{-1}\mathbf{B}_r$ are transfer function of original model and transfer function of reduced-order model, respectively, then $\|G(s) - G_r(s)\|$ is less than a predefined tolerance [1, 2 and 4]. In other words, reduced system $G_r(s)$ approximates original system $G(s)$ well. Model reduction is to apply in circuit electrics (VSLI design), wave surge forecast, vibrations system , and biological system [1, 2, 10 and 11].

Model reduction can be approximated by Krylov subspace method, singular value decomposition (SVD) method, or combining between Krylov subspace and SVD method. Model reduction is to apply in circuit electrics (VSLI design), wave surg forecast, vibrations system , and biological system.

On his textbook, Antoulas (2005) discussed a survey of model reduction methods [1]. Futhermore, we refer a survey paper of model reducion in [7-10] . Model reduction using Krylov subpace is constructed by moment matching. This method have devoleped with others approach. Gallivan et. al. (2003) proposed model reduction using Krylov subspace based on interpolation theory, and more precisely on Pade approximation [3]. Others researchers have developed this method based on modified Arnoldi algorithm. Model reduction using Krylov subspace can de devoleped by matrices arise in these classes can be viewed as multiple copies of certain subspace of the state space of the original system. See for detail in [6].

In this paper, we propose a simple way to find reduced-order system based on sparseness and structure matrix from original system.

The paper is organized as follow. In section 2 we give some fundamental properties of Krylov subspace. Also we give how to construct orthogonal basis for Krylov subspace. Next section , we discussed moment matching and to related by model reduction. Section 4, we present a simple way to find reduced-order system. Finally, conclusion is written in section 5.

In this paper, we use the following notations. Symbol $\mathbf{R}^{m\times n}$ devote set of real matrices have m rows and n columns. The Kronecker product (or tensor product for matrices) is devoted by $\otimes$. Let $\mathbf{A} = \lfloor a_{i,j} \rfloor \in \mathbf{R}^{m\times n}$ and $\mathbf{B} = \lfloor b_{i,j} \rfloor \in \mathbf{R}^{a\times b}$. Kronecker product of **A** and **B** is defined

$$\mathbf{A}\otimes\mathbf{B} = \begin{bmatrix} a_{11}\mathbf{B} & a_{12}\mathbf{B} & \cdots & a_{1n}\mathbf{B} \\ a_{21}\mathbf{B} & a_{22}\mathbf{B} & \cdots & a_{2n}\mathbf{B} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1}\mathbf{B} & a_{m2}\mathbf{B} & \cdots & a_{mn}\mathbf{B} \end{bmatrix} \in \mathbf{R}^{am\times bn}$$

Futhermore, we define norm on vectors or matrices. For all $v \in \mathbf{R}^n$, we define norm of the vector as $\|v\|_2 = \sqrt{\langle v,v\rangle}$ where $\langle -,- \rangle$ is inner product on $\mathbf{R}^n$. For all matrix $\mathbf{A},\mathbf{B} \in \mathbf{R}^n$ we define the inner product $\langle \mathbf{A},\mathbf{B} \rangle = \text{trace}\left(\mathbf{A}^T\mathbf{B}\right)$ and the norm $\|\mathbf{A}\|_{\text{trace}} = \sqrt{\text{trace}\left(\mathbf{A}^T\mathbf{A}\right)}$. The properties of Kronecker product and norm of matrix can be found in [7 and 8].

2. Krylov Subspace

A given matrix $\mathbf{A} \in \mathbf{R}^{n\times n}$ and vector $\mathbf{b} \in \mathbf{R}^n$. The Krylov subspace is defined

$$K_m(\mathbf{A},\mathbf{b}) = \text{span}\left\{\mathbf{b},\mathbf{A}\mathbf{b},\mathbf{A}^2\mathbf{b},\cdots,\mathbf{A}^{m-1}\mathbf{b}\right\} \quad (3)$$

The i-th basis vector in Krylov subspace (3) is linear combination of the previous (i-1) vectors. In other words, the p$^{\text{th}}$ basic vector can be written as linear combination of the first (p-1) vectors. Clearly, $\dim K_1(\mathbf{A},\mathbf{b}) = 1$ and $\dim K_m(\mathbf{A},\mathbf{b}) \le m$. Scaling and shift by identity matrix (= **I**) are not important in Krylov subspace, since $K_m(\mathbf{A},\mathbf{b}) = K_m(\alpha\mathbf{A}+\mathbf{I},\mathbf{b})$ for any nonzero scalar α.

Theorem 1. *Consider Krylov subspace* $K_m(\mathbf{A},\mathbf{b})$ *for* $m = 1, 2, \cdots$.
If $K_m(\mathbf{A},\mathbf{b}) = K_{m+1}(\mathbf{A},\mathbf{b})$ *for some integer m >0, then* $K_m(\mathbf{A},\mathbf{b}) = K_p(\mathbf{A},\mathbf{b})$ *for each integer* $p \ge m > 0$.

Krylov subspace $K_m(\mathbf{A},\mathbf{b})$ = $\text{span}\{\mathbf{b},\mathbf{Ab},\mathbf{A}^2\mathbf{b},\cdots,\mathbf{A}^{m-1}\mathbf{b}\}$ is also called *reachability space* or *controlability space* in control systems comunnity.
Two methods to construct a basis for Krylov subspace $K_m(\mathbf{A},\mathbf{b})$, which are Arnoldi algorithm and lanzcos algorithm. Based on the two methods, Krylov subspace has three main uses: iterative solution of linear equations $\mathbf{Ax}=\mathbf{y}$, iterative approximation of eigenvalue of **A**, and approximation of dynamical systems by moment matching.

The Arnoldi algorithm is an orthogonal projection method for calculating the orthonormal basis for Krylov subspace $K_m(\mathbf{A},\mathbf{b})$. The algorithm can be written the following

Arnoldi Algorithm

[1] Data : **A** and **b**
[2] $\mathbf{v}_1 = \mathbf{b}/\|\mathbf{b}\|$
[3] for j = 1, 2, …, m-1
[4] $h_{i,j} = \mathbf{v_i}^T\mathbf{Av_i}$ for i = 1, 2, …, j
[5] $\mathbf{u}_j = \mathbf{Av}_j - \sum_{i=1}^{j} h_{i,j}\mathbf{v}_i$
[6] $h_{i,j} = \|\mathbf{u_j}\|$
[7] stop if $h_{i,j}$ = 0
[8] $\mathbf{v}_{j+1} = \mathbf{u}_j / h_{j+1,j}$
[9] End for

Outputs of Arnoldi algorithm are matrices $\mathbf{V_m} = [\mathbf{v}_1,\mathbf{v}_2,\cdots,\mathbf{v_m}]$ and

$$\mathbf{H}_m = \begin{bmatrix} h_{1,1} & h_{1,2} & \cdots & h_{1,m-1} & h_{1,m} \\ 0 & h_{2,2} & \cdots & h_{2,m-1} & h_{2,m} \\ 0 & 0 & \ddots & \vdots & \vdots \\ \vdots & \vdots & & h_{m,m-1} & h_{m,m} \\ 0 & 0 & \cdots & 0 & h_{m+1,m} \end{bmatrix}$$

The matrix $\mathbf{H}_m$ is called $(m+1)\times m$ upper Hessenberg matrix .

Theorem 2. *A given matrix* $\mathbf{A}\in\mathbf{R}^{n\times n}$ *and vector* $\mathbf{b}\in\mathbf{R}^n$. *If the Arnoldi procedure does not stop before the* m^{th} *step, then the vectors*

$$\mathbf{v}_1,\mathbf{v}_2,\cdots,\mathbf{v_m}$$

form an orthonormal basis of the Krylov subspace $K_m(\mathbf{A},\mathbf{b})$.

We devoted $\overline{\mathbf{H}}_m = \begin{bmatrix} h_{1,1} & h_{1,2} & \cdots & h_{1,m-1} & h_{1,m} \\ 0 & h_{2,2} & \cdots & h_{2,m-1} & h_{2,m} \\ 0 & 0 & \ddots & \vdots & \vdots \\ \vdots & \vdots & & h_{m-1,m-1} & h_{m-1,m} \\ 0 & 0 & \cdots & h_{m,m-1} & h_{m,m} \end{bmatrix}$

The square matrix $\overline{\mathbf{H}}_m$ is obtained from $\mathbf{H}_m$ by deleting its last row.

Theorem 3

Let $\mathbf{A}$ *(n × n) matrix and* $\mathbf{v}$ *vector. Devote* $\mathbf{V}_m$, $\mathbf{H}_m$, *and* $\overline{\mathbf{H}}_m$ *matrices as described previous. The following relations hold:*

(a). $\mathbf{A}\mathbf{V}_m = \mathbf{V}_m \overline{\mathbf{H}}_m + h_{m+1,m}\mathbf{v}_{\mathbf{m+1}}\mathbf{e_m}^T$ *where is* $\mathbf{e}_m = (0,0,\cdots,1) \in \mathbf{R}^m$

(b). $\mathbf{A}\mathbf{V}_m = \mathbf{V}_{m+1}\mathbf{H}_m$

(c). $\mathbf{V}_m^T \mathbf{A}\mathbf{V}_m = \mathbf{H}_m$

3. Moment Matching and Model Reduction

Consider D is a subset on C complex plane and $z_0 \in D$. The function analytic of f on D can be writen by the Laurent expansion

$$f(z) = \sum_{k=-\infty}^{\infty} c_k (z - z_0)^k \qquad (4)$$

for all $z_0 \in D$. The coeficients c_k is called **k^th^-moment** of $f(z)$ at z_0. If $f(z)$ is a rational function, then $f(z)$ are analytic on the complement of the set of their poles. For complex analysis, we refer [14].

The problem of moment matching can be described as follow: a given a sequence of complex numbers $\{s_1, s_2, \cdots, s_{k_1}\}$ and j = 1, 2, …, j_1 and a function $f(z)$ which is analytical in neighbourhood of points of $\{s_1, s_2, \cdots, s_{k_1}\}$, find a strictly proper real rational function $F(z)$ of degree n with no poles at $\{s_1, s_2, \cdots, s_{k_1}\}$ such that

$$\frac{d^j f(s_k)}{ds^j} = \frac{d^j F(s_k)}{ds^j} \qquad (5)$$

for all j = 1, 2, …, j_1 and k = 1, 2, …, k_1.

Dynamical system (1) can also be represented in frequency domain using Laplace transform. Recall that Laplace transform is defined

$$\hat{f}(s) = \int_0^{\infty} f(t)e^{-st}dt, \text{ for } s \in \mathrm{C} \qquad (6)$$

Let $\hat{x}(s)$, $\hat{y}(s)$, and $\hat{u}(s)$ denoted Laplace transform of *x(t), y(t),* and *u(t)* respectively. Then taking the Laplace transform of (1), we have

$$\begin{cases} s\hat{x}(s) = \mathbf{A}\hat{x}(s) + \mathbf{B}\hat{u}(\mathbf{s}) \\ \hat{y}(s) = \mathbf{C}\hat{x}(s) \end{cases} \tag{7}$$

If we eliminate $\hat{x}(s)$ in equation of (7), we have most important concept on linear system

$$G(s) = \mathbf{C}(s\mathbf{I} - \mathbf{A})^{-1}\mathbf{B} \tag{8}$$

The function of $G(s)$ is called *transfer function*. Equation of (8) can also be written $G(s) = \mathbf{CX}$, where $\mathbf{X}$ is a solution of linear system $(s\mathbf{I} - \mathbf{A})\mathbf{X} = \mathbf{B}$.
By expanding $(s\mathbf{I} - \mathbf{A})^{-1}$ around $x = \frac{1}{s} = 0$, we have

$$(s\mathbf{I} - \mathbf{A})^{-1} = \left(\tfrac{1}{s}\right)\left(\mathbf{I} - \tfrac{\mathbf{1}}{\mathbf{s}}\mathbf{A}\right)^{-1} = (x)(\mathbf{I} - x\mathbf{A})^{-1}$$
$$= x\mathbf{I} + x\mathbf{A} + x^2\mathbf{A}^{\mathbf{2}} + \cdots$$

Hence, transfer function of (8) become

$$G(s) = \mathbf{C}(s\mathbf{I} - \mathbf{A})^{-1}\mathbf{B}$$
$$= \sum_{i=0}^{\infty} \left(\mathbf{CA}^i\mathbf{B}\right)s^{-i-1} \tag{9}$$

The transfer function is also a rational function. The coeficients matrix $\mathbf{CA}^i\mathbf{B}$ are known as i^{th} moment of $G(s)$ at $x = \frac{1}{s} = 0$ (around ∞). It is called *markov parameters*. In this case, we have $\mathbf{CA}^i\mathbf{B} = (i!)^{-1} \left.\frac{d^iG(s)}{ds^i}\right|_{s=\infty}$ for all i.

Model reduction can be approximated by moment matching. The goal is to find transfer function of reduced-order system $G_r(s)$ that interpolate transfer function of original system $G(s)$ and a certain number of its derivatives at the selected points s_k in complex plane so that

$$\frac{d^jG(s)}{ds^j} = \frac{d^jG_r(s)}{ds^j} \tag{10}$$

for j = 1, 2, …, j_1 and k = 1, 2, …, k_1, where k_1 is the number of interpolation points and j_1 is the numbers of moments at each s_k.
The moments are extremely ill-conditioned to compute [1 & 4]. Many investigators proposed model reduction based Krylov subspace that satisfied equation of (10) without computing the moments explicitly [1, 4 and 9].

4. Moment Matching and Krylov Subspace

Cayley-Hamilton theorem state that every square matrix satisfies its own polynomial characteristic, see for detail in [7]. One important use of Cayley-Hamilton theorem is to write $\mathbf{A}^{\mathbf{k}}$, for all $k > n$, as linear combinations of $\mathbf{I}$, $\mathbf{A}$, $\mathbf{A}^{\mathbf{2}}$, …, $\mathbf{A}^{\mathbf{n-1}}$. Therefore expanding $(s\mathbf{I} - \mathbf{A})^{-1}\mathbf{B}$ can be done as the following

$$(s\mathbf{I}-\mathbf{A})^{-1}\mathbf{B} = \left(p_0(s)\mathbf{I} + p_1(s)\mathbf{A} + p_2(s)\mathbf{A}^2 + \cdots + p_{n-1}(s)\mathbf{A}^{n-1}\right)\mathbf{B}$$

or

$$(s\mathbf{I}-\mathbf{A})^{-1}\mathbf{B} = \begin{bmatrix}\mathbf{B} & \mathbf{AB} & \cdots & \mathbf{A}^{n-1}\mathbf{B}\end{bmatrix}\begin{bmatrix} p_0(s) \\ p_1(s) \\ \vdots \\ p_{n-1}(s)\end{bmatrix}$$

If $\det(s\mathbf{I}-\mathbf{A})$ can be evaluated then polynomials $p_0(s)$, $p_1(s)$, …, and $p_{n-1}(s)$ can also be found.

In order to Krylov subspace can be applied to model reduction, Arnoldi algorithm have been modified. By taking norm of vector $\|v\|_2$ replaced with norm of matrix $\|\mathbf{A}\|_{\text{trace}}$, we get general Arnoldi algorithm that it can be used any pair matrix (**A, B**) as data input.

Commonly, the main goal of model reduction based on Arnoldi algorithm is to find projecting matrices **H**, $\mathbf{V} \in \mathbf{R}^{n\times r}$ such that $\mathbf{H}^{\mathbf{T}}\mathbf{V} = \mathbf{I} \in \mathbf{R}^{r\times r}$. Then we set $\mathbf{A_r} = \mathbf{H^T AV}$, $\mathbf{B_r} = \mathbf{H^T B}$, and $\mathbf{C_r} = \mathbf{CV}$. See detail in [1 and 3].

We propose a simple way to find projecting matrices **H**, $\mathbf{V} \in \mathbf{R}^{n\times r}$ as the following. Firstly, we apply r step Arnoldi algorithm to the pair matrix (**A, B**) to obtain the matrices $\mathbf{V} = \mathbf{V}_r$ and $\mathbf{H} = \mathbf{H}_r$. We set $\mathbf{A_r} = \mathbf{H}_r \otimes \mathbf{I}$, $\mathbf{B_r} = \|\mathbf{B}\|_{\text{trace}}(e_1 \otimes \mathbf{I})$, and $\mathbf{C_r} = \mathbf{CV_r}$.

From (9), $G(s)$ is transfer function of original system with $m_i = \mathbf{CA}^i\mathbf{B}$ is i^{th}-moment $G(s)$. If $G_r(s) = \mathbf{C}_r(s\mathbf{I}-\mathbf{A}_r)^{-1}\mathbf{B}_r$ is transfer function of reduced order system, we have

$$\mathrm{M_i} = \mathbf{C}_r\mathbf{A}_r{}^i\mathbf{B}_r = \mathbf{CV_r}\,(\mathbf{H}_r \otimes \mathbf{I})^{\mathbf{i}}\,\|\mathbf{B}\|_{\text{trace}}(e_1 \otimes \mathbf{I}) \qquad (11)$$

as i^{th}-moment $G_r(s)$.

Theorem 4. *Let* $\mathbf{V}_r$ *and* $\mathbf{H}_r$ *can be yielded by r steps Arnoldi algorithm to pair (***A,B***), where* **A** *and* **B** *are matrices from* (1). *Let* $\mathbf{A_r} = \mathbf{H}_r \otimes \mathbf{I}$, $\mathbf{B_r} = \|\mathbf{B}\|_{\text{trace}}(e_1 \otimes \mathbf{I})$, and $\mathbf{C_r} = \mathbf{CV_r}$. *The original system* (1) *and reduced order system* (2) *are first r moments same, i.e.* $\mathrm{M_i} = \mathbf{C}_r\mathbf{A}_r{}^i\mathbf{B}_r$ *for* i = , 1, 2, , …., r-1.

Proof.

Let $\mathbf{V}_r = [\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_r]$ and $\mathbf{H}_r$ is $(r+1)\times r$ upper Hessenberg matrix are obtained by r step Arnoldi algorithm for a pair matrix (**A, B**). The square matrix $\overline{\mathbf{H}}_r$ obtained from $\mathbf{H}_r$ by deleting its last row.

Using fact $(\mathbf{P}\otimes\mathbf{Q})(\mathbf{R}\otimes\mathbf{S}) = (\mathbf{PR})\otimes(\mathbf{QS})$, we have $(\mathbf{A_r})^{\mathbf{i}} = \mathbf{H}_r \otimes \mathbf{I}$ for i = , 1, 2, , …., r-1. Furthermore, we have $\overline{\mathbf{H}}_r\, e_1 \otimes \mathbf{I} = (\overline{\mathbf{H}}_r \otimes \mathbf{I})(e_1 \otimes \mathbf{I})$

Since $\mathbf{B} = \|\mathbf{B}\|_{\text{trace}}\,\mathbf{v_1} = \|\mathbf{B}\|_{\text{trace}}\,\mathbf{V}_r\,(e_1 \otimes \mathbf{I})$, we obtained

$$\begin{aligned} M_i &= \mathbf{C}_r \mathbf{A}_r{}^i \mathbf{B}_r \\ &= \mathbf{CV_r}\,(\mathbf{H}_r \otimes \mathbf{I})^{\mathbf{i}}\ \|\mathbf{B}\|_{\text{trace}}\,(e_1 \otimes \mathbf{I}) \\ &= \|\mathbf{B}\|_{\text{trace}}\ \mathbf{CV_r}\,(\mathbf{H}_r \otimes \mathbf{I})^{\mathbf{i}}\,(e_1 \otimes \mathbf{I}) \\ &= \|\mathbf{B}\|_{\text{trace}}\ \mathbf{CV_r}\,(\mathbf{H}_r \otimes \mathbf{I})^{\mathbf{i-1}}\,(\overline{\mathbf{H}}_r\ e_1 \otimes \mathbf{I}) \end{aligned}$$

The process is continued, hence

$$\begin{aligned} M_i &= \|\mathbf{B}\|_{\text{trace}}\ \mathbf{CV_r}\,((\overline{\mathbf{H}}_r)^i e_1 \otimes \mathbf{I}) \\ &= \|\mathbf{B}\|_{\text{trace}}\ \mathbf{CA^iV_r}\,(e_1 \otimes \mathbf{I}) = \mathbf{CA^i}\ \|\mathbf{B}\|_{\text{trace}}\ \mathbf{v_1} = \mathbf{CA^iB} \end{aligned}$$

for $i = 0, 1, 2, \cdots, r-1$. In other words, The original system (1) and reduced order system (2) are first r moments same.

The proof is finished. ♦

Generally, we want to approximate output $y(t)$ by $y_r(t)$ over large class of input $u(t)$. Different measure of approximation and different choice of class of input function will lead to different model reduction goals.

A given dynamical system like (1)

$$\begin{cases} \dfrac{d}{dt}x(t) = \mathbf{A}\mathbf{x}(t) + \mathbf{b}u(t) \\ y(t) = \mathbf{c}^{\mathbf{T}} x(t) \end{cases} \tag{12}$$

where x: $\mathbf{R} \to \mathbf{R}^n$ is a state vector, u: $\mathbf{R} \to \mathbf{R}$ is input function, y: $\mathbf{R} \to \mathbf{R}$ is a output function, $\mathbf{A} \in \mathbf{R}^{n\times n}$, $\mathbf{b} \in \mathbf{R}^n$, and $\mathbf{c} \in \mathbf{R}^n$ are the system matrices.

The dynamical system is called single input single output (SISO) system. After this, we assume that class of input function u: $\mathbf{R} \to \mathbf{R}$ and its Laplace transform contained in $L_2(\mathbf{R})$ and $\int_0^\infty |u(t)|^2 dt \le 1$. The class is devoted by $D_2(\mathbf{R})$.

In $D_2(\mathbf{R})$, we use the norm of the usual Hardy space is given by

$$\|G\|_{\mathrm{H}} = \sup_{x>0}\ \frac{1}{2\pi}\int_R |G(x+iy)|^2 dy$$

for all $G \in D_2(\mathbf{R})$. We refer [12] for Hardy space.

Model reduction process of (12) will yield reduced-order system

$$\begin{cases} \dfrac{d}{dt}x(t) = \mathbf{A}_r\mathbf{x}(t) + \mathbf{b}_r u(t) \\ y(t) = \mathbf{c}_r{}^{\mathbf{T}} x(t) \end{cases} \tag{13}$$

where x: $\mathbf{R} \to \mathbf{R}^n$ is a state vector, u: $\mathbf{R} \to \mathbf{R}$ is input function, y: $\mathbf{R} \to \mathbf{R}$ is a output function, $r < n$, $\mathbf{A} \in \mathbf{R}^{r\times r}$, $\mathbf{b} \in \mathbf{R}^r$, and $\mathbf{c} \in \mathbf{R}^r$ are the system matrices.

Let $g(s)$ and $g_r(s)$ are transfer function (12) and (13), respectively. We see that $\hat{y}(s) = g(s)\hat{u}(s)$ and $\hat{y}_r(s) = g_r(s)\hat{u}(s)$. Next theorem, we show that error estimate to approximate $y(t)$ by $y_r(t)$ over class of input $u(t) \in D_2(\mathbf{R})$.

Theorem 5. *A given system (12) and its transfer functions g(s). The system (13) is reduced-order system 12 with* $\mathbf{A_r} = \overline{\mathbf{H}}_r$, $\mathbf{b_r} = \|\mathbf{b}\|_2\,\mathbf{e}_1$, and $\mathbf{c_r} = \mathbf{cV_r}$, *where* $\overline{\mathbf{H}}_r$ *and* $\mathbf{V}_r = [\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_r]$ *are obtained by applying r steps Arnoldi algorithm to a pair* ($\mathbf{A}$, $\mathbf{b}$). *The square matrix* $\overline{\mathbf{H}}_r$ *obtained from* $\mathbf{H}_r$ *by deleting its last row.*
If $u \in D_2(\mathbf{R})$ *then*

$$\sup_{t>0} |y(t) - y_r(t)| \leq h_{r+1,r}\, \|\mathbf{b}\|_2 \left\|\mathbf{c}(s\mathbf{I} - \mathbf{A})^{-1}\mathbf{V}_{r+1}\right\|_H \left\|\mathbf{e}_r^T (s\mathbf{I} - \overline{\mathbf{H}}_r)^{-1}\mathbf{e}_1\right\|_H$$

Where $\mathbf{e}_r^T = (0,0,\cdots,1) \in \mathbf{R}^r$ *and* $\mathbf{e}_1 = (1,0,0,\cdots)$.

Proof.

Let *g(s)* and $g_r(s)$ are devoted transfer functions of original system and reduced-order system, respectively. We see that $\hat{y}(s) = g(s)\hat{u}(s)$ and $\hat{y}_r(s) = g_r(s)\hat{u}(s)$.

We claim that $\sup_{t>0} |y(t) - y_r(t)| \leq \|g(s) - g_r(s)\|_H$

Using fact invers Laplace transform,

$$\sup_{t>0} |y(t) - y_r(t)| = \sup_{t>0} \left|\frac{1}{2\pi}\int_R (\hat{y}(it) - \hat{y}_r(it))e^{ixt}dt\right|$$

$$\leq \frac{1}{2\pi}\int_R |(\hat{y}(it) - \hat{y}_r(it))|dt$$

Since $\hat{y}(s) - \hat{y}_r(s) = (g(s) - g_r(s))\hat{u}(s)$, we have

$$\sup_{t>0} |y(t) - y_r(t)| \leq \frac{1}{2\pi}\int_R |(g(it) - g_r(it))||\hat{u}(it)|dt$$

if we use Cauchy-Schwarz inequality for integral, then

$$\frac{1}{2\pi}\int_R |(g(it) - g_r(it))||\hat{u}(it)|dt$$

$$\leq \left(\frac{1}{2\pi}\int_R |(g(it) - g_r(it))|^2 dt\right)^{1/2} \left(\int_R |u(t)|^2 dt\right)^{1/2}$$

$$\leq \left(\frac{1}{2\pi}\int_R |(g(it) - g_r(it))|^2 dt\right)^{1/2} = \|g - g_r\|_H$$

whenever $u \in D_2(\mathbf{R})$.

Hence

$$\sup_{t>0} |y(t) - y_r(t)| \leq \|g - g_r\|_H \tag{14}$$

whenever $u \in D_2(\mathbf{R})$.

Using Theorem 3.a, $\mathbf{A}\mathbf{V}_r = \mathbf{V}_r\,\overline{\mathbf{H}}_r + h_{r+1,r}\mathbf{v}_{r+1}\mathbf{e}_r^T$

$\Leftrightarrow$ $s\mathbf{V}_r - \mathbf{A}\mathbf{V}_r = s\mathbf{V}_r - \mathbf{V}_r\,\overline{\mathbf{H}}_r - h_{r+1,r}\mathbf{v}_{r+1}\mathbf{e}_r^T$

$\Leftrightarrow$ $(s\mathbf{I} - \mathbf{A})\,\mathbf{V}_r = \mathbf{V}_r\,(s\mathbf{I} - \overline{\mathbf{H}}_r) - h_{r+1,r}\mathbf{v}_{r+1}\mathbf{e}_r^T$

Since $g(s) = \mathbf{c}(s\mathbf{I} - \mathbf{A})^{-1}\mathbf{b}$ and $g_r(s) = \mathbf{c}_r(s\mathbf{I} - \mathbf{A}_r)^{-1}\mathbf{b}_r$, it implies

$g(s) - g_r(s)$

$= \mathbf{c}(s\mathbf{I} - \mathbf{A})^{-1}\mathbf{b} - \mathbf{c}_r(s\mathbf{I} - \mathbf{A}_r)^{-1}\mathbf{b}_r$

$$= \mathbf{c}(s\mathbf{I}-\mathbf{A})^{-1}\mathbf{b} - (\mathbf{c}\mathbf{V}_r)\left(s\mathbf{I}-\overline{\mathbf{H}}_r\right)^{-1}\left(\|\mathbf{b}\|_2\mathbf{e}_1\right)$$
$$= \mathbf{c}(s\mathbf{I}-\mathbf{A})^{-1}(\mathbf{b} - (s\mathbf{I}-\mathbf{A}))\mathbf{V}_r\left(s\mathbf{I}-\overline{\mathbf{H}}_r\right)^{-1}\left(\|\mathbf{b}\|_2\mathbf{e}_1\right)$$
$$= \mathbf{c}(s\mathbf{I}-\mathbf{A})^{-1}(\mathbf{b} - \mathbf{V}_r(s\mathbf{I} - \overline{\mathbf{H}}_r) - h_{r+1,r}\mathbf{v}_{r+1}\mathbf{e}_r{}^T)\left(s\mathbf{I}-\overline{\mathbf{H}}_r\right)^{-1}\left(\|\mathbf{b}\|_2\mathbf{e}_1\right)$$
$$= \mathbf{c}(s\mathbf{I}-\mathbf{A})^{-1}\left(b-\left(\|\mathbf{b}\|_2\mathbf{V}_r\mathbf{e}_1\right)+h_{r+1,r}\|\mathbf{b}\|_2\mathbf{V}_{r+1}\mathbf{e}_r{}^T\left(s\mathbf{I}-\overline{\mathbf{H}}_r\right)^{-1}\mathbf{e}_1\right)$$

By Arnoldi process, $\mathbf{b} = \|\mathbf{b}\|\mathbf{v}_1 = \|\mathbf{b}\|\mathbf{V}_r\mathbf{e}_1$ and implies

$$g(s) - g_r(s) = \mathbf{c}(s\mathbf{I}-\mathbf{A})^{-1}\left(h_{r+1,r}\|\mathbf{b}\|_2\mathbf{V}_{r+1}\mathbf{e}_r{}^T\left(s\mathbf{I}-\overline{\mathbf{H}}_r\right)^{-1}\mathbf{e}_1\right)$$

hence

$$\|g-g_r\|_H \le h_{r+1,r}\|\mathbf{b}\|_2\left\|\mathbf{c}(s\mathbf{I}-\mathbf{A})^{-1}\mathbf{V}_{r+1}\right\|_H\left\|\mathbf{e}_r^T\left(s\mathbf{I}-\overline{\mathbf{H}}_r\right)^{-1}\mathbf{e}_1\right\|_H \qquad (15)$$

From (14) and (15), we get

$$\sup_{t>0}|y(t)-y_r(t)| \le h_{r+1,r}\|\mathbf{b}\|_2\left\|\mathbf{c}(s\mathbf{I}-\mathbf{A})^{-1}\mathbf{V}_{r+1}\right\|_H\left\|\mathbf{e}_r^T\left(s\mathbf{I}-\overline{\mathbf{H}}_r\right)^{-1}\mathbf{e}_1\right\|_H$$

whenever $u \in D_2(\mathbf{R})$.
The proof is finished. ♦

5. Conclusion

In this paper, we have proposed model reduction based on output of matrices in Arnoldi process. This result can be used to estimate error output system for single input single output system over a certain class of input function $u(x)$.

Acknowledgment. This research was partially supported by FRGS and Postgraduate Financial Scheme (SKS) from Universiti Malaysia Terengganu (UMT) 2008.

References

[1]. Antoulas AC., Approximation of large scale dynamical system, SIAM Philidelphia, USA, 2005.

[2]. Biswa Nath Datta, Numerical methods for linear control systems, Elsevier Academic Press, UK, 2004.

[3]. Gallivan K, Vandendorpe A, and Van Dooren P, Model reduction via truncations: an interpolaion point of view, Linear algebra and its applications, 375 (2003), 115-134.

[4]. Gugercin S and K Wilcox, Krylov projection framework for Fourier model reduction, Automatica, 44 (2008), 209-215.

[5]. Gugercin S, An iterative SVD-Krylov based method for model reduction of large scale dynamical system, Linear algebra and its application, 428 (2008), 1964-1986.

[6]. Herng-Jer Lee, Chia-Chi Chu, Wu-shung Feng, An adaptive-order Arnoldi method for model-order eduction of linear time-invariant systems, Linear algebra and its applications, 415 (2006), 235-261.

[7]. Horn, Roger A and Johnson, Charles R, Matrix Analysis, Cambrige University Press, UK, 1985.

[8]. Laub, Ana J, Matrix Analysis for scientists and engineers, SIAM Publishers, Philadelfia, USA, 2005.

[9]. MM Awais, Shafay Shamail, and Nisar Ahmed, Dimensionally reduced Krylov subspace model reduction for large scale systems, Applied Mathematics and Computation, 191 (2007), 21-30.

[10]. Roland W Freud, Krylov subspace associated with higher order linear dynamical system, BIT Numerical Mathematics, 45 (2005), 495-516.

[11]. Thilo Penzl, Algorithms for model reduction of large dynamical system, Linear algebra and its applications, 415 (2006), 322-343.

[12]. Young, N, An Introduction to Hilbert Space, Cambridge University Press, UK, 1988.

[13]. Zhoujun Bai, Krylov subspace techniques for reduced-order modelling of large scale dynamical system, Applied numerical mathematics, 43 (2002), 9-44.

[14]. Zill D G and Shanahan PD, A First Course in Complex Analysis with Applications, John and Bartlett Publisher, USA, 2003.

Received: December, 2008