Controllable and Observable Subspaces in Computational Fluid Dynamics

Karen Willcox

Massachusetts Institute of Technology, Cambridge, MA, 02139, USA

Abstract. The concepts of controllability and observability are explored for computational fluid dynamic (CFD) applications. Identification of a balanced subspace which accounts for both controllability (inputs) and observability (outputs) can be performed using the proper orthogonal decomposition (POD) and the concepts of primal and dual state-space systems. Conventional model reduction approaches in CFD consider only controllability and do not take account of engineering outputs of interest. For a two-dimensional, inviscid flow problem, the balanced reduced-order model is shown to have superior performance, implying that the underlying CFD model is unbalanced with respect to the chosen input and output.

1 Introduction

Consider a general linearised CFD model, which can be written as

$$\frac{d\mathbf{x}}{dt} = A\mathbf{x} + B\mathbf{u}, \qquad \mathbf{y} = C\mathbf{x}, \tag{1}$$

where $\mathbf{x}(t)$ is the state vector containing the n unknown flow perturbation quantities at each point in the computational grid. The vectors $\mathbf{u}(t)$ and $\mathbf{y}(t)$ contain the system forcing inputs and outputs respectively. The linearisation matrices A, B and C in (1) are real, and are evaluated at steady-state conditions.

The CFD model typically contains tens of thousands of unknowns, even if just two dimensions are considered. The equations are expensive to time-march, even under the assumption of linearity, and the system is prohibitively large for many applications such as aeroelasticity, optimisation and active flow control. Model reduction is a powerful technique for obtaining low-order, high-fidelity models for complex systems. Reduction techniques have been developed and applied in many contexts, including fluid dynamics, structural dynamics and controls.

The fundamental concept of model reduction is to identify a low-order, dominant subspace. By projecting the high-order system onto this subspace, a low-order model is obtained which accurately captures the dynamics of interest. A number of reduction techniques have been developed and applied to CFD models, including the POD, which uses empirical data to determine a basis which is optimal in the sense that it minimises the error between the exact and projected data. The POD has been used extensively for identifying dominant subspaces in fluid dynamic applications (see [1] for a review). In this paper, it is shown that the POD can be applied using either system inputs (the conventional approach), system outputs or, as shown in [2], a balanced combination of the two.

2 Model Order Reduction

Consider the large n^{th}-order linear system defined by (1). The objective of the reduction procedure is to determine an n_r^{th}-order reduced-space basis onto which the state vector can be projected, that is $\mathbf{x} = V\mathbf{x_r}$, and an orthonormal set $\tilde{V}$, so that $\tilde{V}V = I$. This basis is chosen appropriately so that the reduced-order system

$$\dot{\mathbf{x}}_r = \tilde{V}AV\mathbf{x}_r + \tilde{V}B\mathbf{u} \qquad \mathbf{y}_r = CV\mathbf{x}_r \tag{2}$$

accurately reproduces the desired dynamics of the original system (1) with many fewer states ($n_r << n$). While the basis vectors are determined mathematically, the space they span often has physical significance. While many model reduction techniques have been developed in a controls setting, the challenge has been to find techniques which can be used for the very large systems encountered in CFD applications. One widely used, efficient reduction technique will now be described, and its connection to classical control concepts will be outlined.

2.1 Proper Orthogonal Decomposition

The POD has been widely used to determine efficient bases for dynamic systems. It can be shown that an optimal set of basis vectors, $\mathbf{\Psi}$, are the eigenfunctions of the kernel, K. Sirovich introduced the method of snapshots as a way of determining the eigenfunctions $\mathbf{\Psi}$ without explicitly calculating the kernel [3]. The kernel can be approximated as

$$K(\theta, \theta') = \frac{1}{m}\sum_{i=1}^{m} \mathbf{x}_i(\theta)\mathbf{x}_i^T(\theta'), \tag{3}$$

where $\mathbf{x}_i(\theta)$ is the instantaneous system state or "snapshot" at a time t_i and the number of snapshots m is sufficiently large. It can be shown that the eigenvectors of K are linear combinations of the snapshots, and can be computed by solving an $m \times m$ eigenvalue problem.

Rather than performing a set of time domain simulations to obtain the snapshots $\mathbf{x}_i$, a better approach is to use linearity and evaluate the POD basis vectors in the frequency domain. Any general forcing function can be considered as a superposition of sinusoidally time-varying components each at a frequency ω. Because the system is linear, the component of forcing at frequency ω induces a response which is also harmonic with frequency ω, that is $\mathbf{x}(t) = \mathbf{X}e^{j\omega t}$ and $\mathbf{y} = \mathbf{Y}e^{j\omega t}$. Considering a single temporal harmonic, ω_k, the state-space system (1) can be written in the frequency domain as

$$\mathbf{X}_k = (j\omega_k I - A)^{-1} B\mathbf{U}_k, \qquad \mathbf{Y}_k = C\mathbf{X}_k. \tag{4}$$

The POD snapshots can therefore be obtained by choosing a set of sample frequencies $\{\omega_k\}$ based on the frequency content of problems of interest and solving the frequency domain system (4) to obtain the responses $\{\mathbf{X}_k\}$ [4].

2.2 Controllability

The POD process is often explained in the context of a least squares fit through a set of data obtained from the high-order system; the POD vectors minimize the error between the actual snapshots and their reconstruction in the reduced space. Another insightful interpretation of the reduction process can be gained by considering the concept of controllability, borrowed from state-space theory.

The controllable subspace is that set of states that can be obtained in a finite time with zero initial state and a given input $\mathbf{u}(t)$ (also called the set of reachable states). The controllability grammian is an $n \times n$ matrix whose eigenvectors span the controllable subspace and is defined for the linear system (1) as

$$W_c = \int_0^\infty e^{At} B B^T e^{A^T t} dt. \tag{5}$$

Note that W_c does not depend on the definition of the output. The eigenvalues of W_c are real, non-negative numbers and denote the "reachability" of the corresponding eigenvector, or the amount of control "energy" required to obtain it. For a state $\mathbf{x}_0$, the controllability function is defined as

$$L_c(\mathbf{x}_0) = \min_{\mathbf{u} \in L_2(-\infty,0)_{\mathbf{x}(-\infty)=0, \mathbf{x}(0)=\mathbf{x}_0}} \frac{1}{2} \int_{-\infty}^0 ||\mathbf{u}(t)||^2 dt \tag{6}$$

where $||.||$ denotes the L_2 norm. The lower the value of $L_c(\mathbf{x}_0)$, the more reachable the state $\mathbf{x}_0$. If $\mathbf{x}_0$ is the kth eigenvector of W_c, $L_c(\mathbf{x}_0) = 1/2\lambda_k^c$, where λ_k^c is the kth eigenvalue of W_c. The eigenvectors of W_c, ordered by increasing eigenvalue, therefore describe the most reachable states of the system.

Further insight can be gained by thinking about the problem in the frequency domain. The controllability grammian can also be written as

$$W_c = \frac{1}{2\pi} \int_{-\infty}^\infty (j\omega I - A)^{-1} B B^T \left(-j\omega I - A^T\right)^{-1} d\omega. \tag{7}$$

The term $\mathbf{X}_\omega = (j\omega I - A)^{-1} B$ is the magnitude of the response of the linear system to sinusoidal forcing at a frequency ω (the actual response is given by $\mathbf{x}(t) = \mathbf{X}_\omega e^{i\omega t}$).

Although the kernel is never explicitly computed in the POD frequency domain analysis, by choosing a finite set of discrete frequencies for the snapshots, (3) can be written

$$K = \frac{1}{m} \sum_{i=1}^m (j\omega_i I - A)^{-1} B B^T \left(-j\omega_i I - A^T\right)^{-1}. \tag{8}$$

By comparing equations (7) and (8), we can see that, within a scaling factor, the POD kernel is therefore an approximation to the controllability grammian over a chosen, restricted frequency range. The subspace spanned by the POD basis vectors approximates the most controllable subspace.

2.3 Observability

While controllability considers only the inputs to the state-space system, an analogous concept exists which considers system outputs. The observable subspace comprises those states which as initial conditions could produce a non-zero output with no external input, and the observability grammian is defined as

$$W_o = \frac{1}{2\pi} \int_{-\infty}^{\infty} \left(j\omega I - A^T\right)^{-1} C^T C \left(-j\omega I - A\right)^{-1} d\omega. \tag{9}$$

The observability function L_o is defined as

$$L_o(\mathbf{x}_0) = \frac{1}{2} \int_0^{\infty} ||\mathbf{y}(t)||^2 dt \tag{10}$$

where $\mathbf{y}(t)$ is the output generated with $\mathbf{x}_0$ as an initial condition and no external forcing. The higher the value of $L_o(\mathbf{x}_0)$, the more observable the state $\mathbf{x}_0$. If $\mathbf{x}_0$ is the kth eigenvector of W_o, $L_o(\mathbf{x}_0) = \lambda_k^o/2$, where λ_k^o is the kth eigenvalue of W_o. The eigenvectors of W_o, ordered by increasing eigenvalue, therefore describe the most observable states of the system.

Controllability and observability are dual concepts. Consider the state-space system

$$\frac{d\mathbf{z}}{dt} = A^T \mathbf{z} + C^T \mathbf{u}_d, \qquad \mathbf{y}_d = B^T \mathbf{z}, \tag{11}$$

which is the dual of the primal system defined by (1). Here, $\mathbf{z}$ is the dual state vector, and $\mathbf{u}_d$ and $\mathbf{y}_d$ contain the dual inputs and outputs. Comparing (7) and (9), it can be seen that the controllability grammian of the primal system is equal to the observability grammian of the dual system, and vice versa.

We have seen that the POD basis vectors approximate the most reachable states of the system. Similarly, a set of vectors could be generated which approximate the most observable states of the system by applying the POD technique to the dual state-space system. The question remains: which basis is the better one to use for the purposes of model reduction? If a state is very reachable, then it is likely to play an important role in the system dynamics. However, if this state generates little or no output, then it will not be important for predicting the output of interest accurately. Truly what is of importance is the way in which inputs are transferred to outputs; the state itself is an intermediate mechanism, which is moreover dependent on the arbitrary choice of state coordinates.

This dilemma can be resolved through the concept of a balanced realization, first introduced by Moore [6]. The basic idea is that a balanced representation of the system is one in which the controllability grammians are equal and diagonal. Model reduction can then be performed using the Hankel singular values of the system. These values are independent on the choice of system realization, and represent the importance of a state for transmitting input to output. As described in [2], using classical balanced realization theory, the POD-based grammian approximations can be combined in an efficient way to obtain an approximately balanced subspace for projection. In the following section, results will be presented that compare the three subspaces, controllable, observable and balanced, and the reduced-order models they produce.

3 Results

Results will be presented for a two-dimensional NACA 0012 airfoil operating in unsteady plunging motion with a steady-state Mach number of 0.755. A finite volume CFD formulation for the Euler equations is used, which is described in [5]. The CFD mesh has 3482 grid points, which corresponds to a total of $n = 13928$ unknowns in the linear state-space system. The input to this system is the rigid plunging motion (vertical motion of the airfoil), while the output of interest is the lift force generated.

POD was applied to the primal and dual systems to generate two sets of basis vectors approximating the most controllable and most observable subspaces for the input and output selected. The approximate balancing technique from [2] was also used to create a set of balanced basis vectors. The snapshots of the primal and dual systems were taken over a frequency range from $\omega = 0$ to $\omega = 2$ in steps of $\Delta\omega = 0.1$. From these snapshots, 15 POD basis vectors were computed to perform the balancing. Reduced-order models were constructed using each of the resulting bases. Figure 1 shows the transfer function from plunging motion to lift force. The results from the CFD model are compared to results from each of the reduced-order models with four and ten states.

From Figure 1 it is apparent that the primal POD vectors form a much better basis than the dual POD vectors, although the approximately balanced basis yields the best results. Although a wider range is plotted, the reduced-order models are strictly only valid over the frequency range sampled ($\omega \leq 2$). With just four states, the balanced reduced-order model gives a very accurate representation of the dynamics. The controllability basis shows some error for $\omega > 1$, while the observability basis is very inaccurate for all frequencies. If the size of the model is increased to ten states, results for all three bases improve. The dual POD model gives good results for low frequencies, but still shows significant error at higher frequencies. In order to obtain an accurate representation using this basis, more states would be required.

This result yields interesting insight to the CFD formulation. For this case, choosing vectors from the most controllable subspace yields a much better approximation than choosing vectors from the most observable subspace, which suggests that the flow states capable of generating the most lift force are not typically excited by a plunging motion. The better performance by the balanced basis suggests that the CFD coordinate system is unbalanced with respect to the chosen input and output. An interesting avenue to explore is whether a rescaling of the CFD model could lead to improved convergence properties.

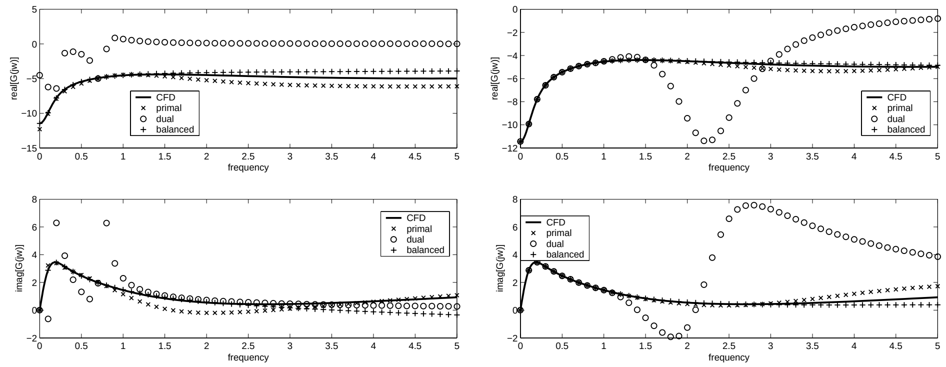

Fig. 1. Transfer function from plunging motion to lift force for subsonic airfoil. Results from CFD model ($n = 13,928$) are compared to reduced-order models derived using POD for primal, dual and balanced models. Left: $k = 4$. Right: $k = 10$.

4 Conclusions

Controllability and observability subspaces have been investigated for a CFD model in the context of model reduction. While the results will vary depending on the specific CFD formulation and the inputs and outputs of interest, the case presented demonstrates some interesting aspects. The problem studied considers the lift force generated on an airfoil due to a plunging motion input. For this case, the most controllable subspace provides a much better approximation to the system dynamics than the most observable subspace. This suggests that the flow states capable of producing the most lift force are not those excited by plunging motion. Conversely, the states reached during a typical plunging motion are important for generating lift. However, the approximately balanced model yields the best results, suggesting that the CFD formulation is unbalanced with respect to the chosen input and output.

References

1. E.H. Dowell and K.C. Hall: 'Modeling of fluid-structure interaction', *Annual Review of Fluid Mechanics*, 33, 2001, pp. 445-90
2. K. Willcox and J. Peraire: 'Balanced Model Reduction via the Proper Orthogonal Decomposition', AIAA Paper 2001-2611, 2001, to appear *AIAA Journal*.
3. L. Sirovich: 'Turbulence and the Dynamics of Coherent Structures. Part 1 : Coherent Structures', *Quarterly of Applied Mathematics*, 45(3), Oct. 1987, pp. 561-571
4. T. Kim: 'Frequency-Domain Karhunen-Loeve Method and Its Application to Linear Dynamic Systems', *AIAA Journal* 36(11), 1998, pp. 2117-2123
5. K. Willcox: *Reduced-Order Aerodynamic Models for Aeroelastic Control of Turbomachines*, PhD Thesis, Dept. of Aeronautics and Astronautics, MIT, Feb. 2000
6. B.C. Moore: 'Principal Component Analysis in Linear Systems: Controllability, Observability, and Model Reduction', *IEEE Transactions on Automatic Control*, AC-26(1), 1981, pp. 17-31