Centroidal Voronoi Tessellation Based Proper Orthogonal Decomposition Analysis

Qiang Du and Max D. Gunzburger

Abstract. Proper orthogonal decompositions (POD) have been used to systematically extract the most energetic modes while centroidal Voronoi tessellations (CVT) have been used to systematically extract best representatives. We combine the ideas of CVT and POD into a hybrid method for model reduction. The optimality of such an approach and various practical implementation strategies are discussed.

1. Introduction

In the study of turbulent and chaotic systems and in the real-time feedback control of complex systems, model reduction plays a very important role. In the former case, there is a need to identify highly persistent spatio-temporal structures using simple approaches. In the latter case, low-dimensional state models are needed so that actuation can be determined quickly from sensed data. As a result, there have been many studies devoted to the development, testing, and use of reduced-order models for complex dynamical systems such as unsteady fluid flows.

Today, perhaps the most popular technique for model reduction is based on *proper orthogonal decomposition* (POD). POD is closely related to the statistical method known as Karhunen-Loève analysis or the method of empirical orthogonal eigenfunctions. POD has become popular due to its potential for extracting empirical information from experimental data or from data obtained from high-fidelity numerical simulations; it has also become popular as a means of building low-dimensional models.

For model reduction in the context of partial differential equations, approximation is effected by solving partial differential equations for long time periods or for various parameter values, then performing the POD analysis on snapshots of the solution, and then using the Galerkin method to project the partial differential equation model onto the reduced POD basis.

1991 *Mathematics Subject Classification.* 65M99, 62H30, 68U99, 65C20, 93A15.
Key words and phrases. proper orthogonal decomposition, centroidal Voronoi tessellation, reduced-order modeling.
This work was supported by the National Science Foundation under grant number CCR-9988303.

There have been many studies devoted to the use of POD for obtaining low-dimensional dynamical system approximations; see, for example [2, 3, 4, 5, 6, 13, 15, 16, 21, 24, 28, 29, 32, 33, 34, 36]. The use of POD analysis in control problems for partial differential equations has been considered in [1, 17, 18, 25, 26, 27, 30, 31, 35].

In this paper, we consider reduced order modeling through the use of a *centroidal Voronoi tessellation* (CVT) of a snapshot set; CVT may be viewed as a clustering technique. Thus, CVT provides an alternative to POD. Moreover, CVT and POD may be combined to define a generalization of POD. The plan of the paper is as follows. In Section 2 we briefly review the POD technique and some of its properties; we do so so that POD and CVT can be compared and contrasted. In Section 3, we introduce the CVT concept. In Section 4, we combine the POD and CVT concepts into a hybrid method (CVOD) which inherits good features of both. Finally, in Section 5, we provide brief remarks about the use of the hybrid method for model reduction. The testing of the usefulness of CVT and CVOD for this purpose through numerical experiments is currrently under way.

2. Proper orthogonal decomposition

In the proper orthogonal decompositions (POD) technique, dominant features from experimental or numerical data are extracted through a set of orthogonal functions which are related to the eigenfunctions of the correlation matrix of the data.

For n snapshots $\widetilde{\mathbf{x}}_j \in \mathbb{R}^N$, $j = 1, \ldots, n$, let

$$\widetilde{\boldsymbol{\mu}} = \frac{1}{n} \sum_{j=1}^{n} \widetilde{\mathbf{x}}_j$$

and

$$\mathbf{x}_j = \widetilde{\mathbf{x}}_j - \widetilde{\boldsymbol{\mu}}\,, \qquad j = 1 \ldots, n,$$

be a set of modified snapshots. Let $d \leq n$. Then, the POD basis $\{\boldsymbol{\phi}_i\}_{i=1}^d$ of cardinality d is found by successively solving, for $i = 1, \ldots, d$,

$$\lambda_i = \max_{\|\boldsymbol{\phi}_i\|=1} \frac{1}{n} \sum_{j=1}^{n} |\boldsymbol{\phi}_i^T \mathbf{x}_j|^2 \qquad \text{and} \qquad \boldsymbol{\phi}_i^T \boldsymbol{\phi}_\ell = 0 \quad \text{for } \ell \leq i-1\,.$$

If $n \geq N$, this decomposition is known as the *direct* method; if $n < N$, then it is known as the *snapshot* method. For the latter case, $\boldsymbol{\phi}_i = \frac{1}{\sqrt{n\lambda_i}} A \boldsymbol{\chi}_i$, where $\boldsymbol{\chi}_i$ with $\|\boldsymbol{\chi}_i\| = 1$ denotes the eigenvector corresponding to the i-th largest eigenvalue λ_i of the $n \times n$ correlation matrix $K = (K_{j\ell})$, where $K_{j\ell} = \frac{1}{n} \mathbf{x}_j^T \mathbf{x}_\ell$. From now on, we will only consider the case $n < N$.

The POD basis is optimal in the following sense [15]. Let $\{\boldsymbol{\psi}_i\}_{i=1}^n$ denote an arbitrary orthonormal basis for the span of the modified snapshot set $\{\mathbf{x}_j\}_{j=1}^n$. Let $P_{\boldsymbol{\psi},d}\mathbf{x}_j$ be the projection of $\mathbf{x}_j$ in the subspace spanned by $\{\boldsymbol{\psi}_i\}_{i=1}^d$ and let the *error* be defined by

$$\mathcal{E} = \sum_{j=1}^{n} \|\mathbf{x}_j - P_{\boldsymbol{\psi},d}\mathbf{x}_j\|^2 \,. \tag{1}$$

Then, the minimum error is obtained when $\boldsymbol{\psi}_i = \boldsymbol{\phi}_i$ for $i = 1, \ldots, d$, i.e., when the $\boldsymbol{\psi}_i$'s are the POD basis vectors.

2.1. POD and the singular value decomposition

We review some of the close connections between the proper orthogonal decomposition of a set of snapshots $\widetilde{\mathbf{x}}_j \in \mathbb{R}^N$, $j = 1, \ldots, n$, and the singular value decomposition (SVD) of the $N \times n$ matrix whose j-th column is the modified snapshot $\mathbf{x}_j = \widetilde{\mathbf{x}}_j - \widetilde{\boldsymbol{\mu}}$. Recall that throughout, $N > n$.

The SVD of an $N \times n$ matrix A is given by [14] $A = U\Sigma V^T$, where U is an $N \times N$ orthonormal matrix, V is an $n \times n$ orthogonal matrix, and Σ is an $N \times n$ diagonal matrix having only non-negative and non-increasing entries on the diagonal. (Here we only treat the real case; everything can be easily generalized to the complex case.) The non-zero entries of Σ are the *singular values* of A, the columns $\{\mathbf{u}_j\}_{j=1}^N$ of U are the *left singular vectors* of A, and the columns $\{\mathbf{v}_i\}_{i=1}^n$ of V are the *right singular vectors* of A.

For the given set of snapshots $\widetilde{\mathbf{x}}_j \in \mathbb{R}^N$, let A denote the matrix whose columns are the modified snapshots $\mathbf{x}_j$, i.e.,

$$A = \big(\widetilde{\mathbf{x}}_1 - \mu(\widetilde{\mathbf{x}})\,,\, \widetilde{\mathbf{x}}_2 - \mu(\widetilde{\mathbf{x}})\,,\, \ldots, \widetilde{\mathbf{x}}_n - \mu(\widetilde{\mathbf{x}})\,\big)\,.$$

Let $A = U\Sigma V^T$ denote the SVD of A. Then, the correlation matrix (or the scaled normal matrix) is the $n \times n$ matrix

$$K = \frac{1}{n}A^T A = \frac{1}{n}V\Sigma^T\Sigma V^T\,.$$

Recall that

$$K\boldsymbol{\chi}_i = \frac{1}{n}A^T A\boldsymbol{\chi}_i = \boldsymbol{\chi}_i\lambda_i \qquad \text{and} \qquad \|\boldsymbol{\chi}_i\| = 1 \qquad\qquad \text{for } i = 1, \ldots, n\,.$$

Then, it is well known that $\boldsymbol{\chi}_i = \mathbf{v}_i$, i.e., the eigenvectors of the matrix $A^T A$ coincide with the right singular vectors $\mathbf{v}_i$ of the snapshot matrix A. Moreover, we have that $\sigma_i^2 = n\lambda_i$, where σ_i and λ_i denote the i-th singular value of A and the i-th largest eigenvalue of K, respectively. It then follows that the POD basis vectors $\{\boldsymbol{\phi}_i\}_{i=1}^n$ may be expressed in terms of the singular values and right singular vectors of A by $\boldsymbol{\phi}_i = \frac{1}{\sigma_i}A\mathbf{v}_i$, for $i = 1, \ldots, n$. It then easily follows that $\boldsymbol{\phi}_i = \mathbf{u}_i$ for $i = 1, \ldots, n$, i.e., the POD basis vectors are the first n left singular vectors of the snapshot matrix A. Note that both $\{\boldsymbol{\chi}_i\}_{i=1}^n$ and $\{\boldsymbol{\phi}_i\}_{i=1}^n$ are orthonormal sets in $\mathbb{R}^n$ and $\mathbb{R}^N$, respectively.

The connection between POD and SVD can be exploited to give a simple proof of the optimality property of the POD basis that was stated earlier. Indeed, we have that

$$\mathcal{E} = \sum_{j=1}^{n} \|\mathbf{x}_j - P_{\boldsymbol{\phi},d}\mathbf{x}_j\|^2 = \sum_{j=1}^{n} \|\mathbf{x}_j\|^2 - \sum_{j=1}^{n} \|P_{\boldsymbol{\phi},d}\mathbf{x}_j\|^2 \,.$$

Now,

$$\sum_{j=1}^{n} \|P_{\boldsymbol{\phi},d}\mathbf{x}_j\|^2 = \sum_{i=1}^{d} \|A^T\boldsymbol{\psi}_i\|^2 = \sum_{i=1}^{d} \|\Sigma'\boldsymbol{\alpha}_i\|^2 \,,$$

where $\boldsymbol{\alpha}_i = U'\psi_i$ and Σ' and U' are the $n \times n$ submatrices defined by the partitionings

$$\Sigma = \begin{pmatrix} \Sigma' \\ 0 \end{pmatrix} \qquad \text{and} \qquad U = (U' \;\; U'') \,.$$

Note that also $\{\boldsymbol{\alpha}_i\}_{i=1}^d$ forms an orthonormal basis of a d-dimensional subspace of $\mathbb{R}^n$. Clearly,

$$\max_{\{\boldsymbol{\alpha}_i\}_{i=1}^d} \sum_{i=1}^{d} \|\Sigma'\boldsymbol{\alpha}_i\|^2$$

is achieved for $\boldsymbol{\alpha}_i = \mathbf{e}_i$, where $\mathbf{e}_i$ is the unit vector with i-th component 1 and the other components zero. Hence, we have that $\mathcal{E}$ is minimized when $\boldsymbol{\psi}_i = U\mathbf{e}_i = \mathbf{u}_i$ for $i = 1, \ldots, d$, i.e., by the POD basis.

The connection between POD and SVD also makes it is easy to show that the "error" of the d-dimensional POD subspace is given by

$$\mathcal{E}_{\text{pod}} = \sum_{j=d+1}^{n} \sigma_j^2 = n \sum_{j=d+1}^{n} \lambda_j \,.$$

Thus, if one wishes for the relative error to be less than a prescribed tolerance δ, i.e., if one wants $\mathcal{E}_{\text{pod}} \le \delta\Sigma_{j=1}^n |\mathbf{x}_j|^2$, one should

$$\begin{array}{r}\text{choose } d \text{ to be the} \\ \text{smallest integer such that}\end{array} \quad \sum_{j=d+1}^{n} \sigma_j^2 = n \sum_{j=d+1}^{n} \lambda_j < \delta\Sigma_{j=1}^n |\mathbf{x}_j|^2 \,. \tag{2}$$

Since we also have that

$$\sum_{j=1}^{n} \|\mathbf{x}_j\|^2 = \sum_{j=1}^{n} \sigma_j^2 = n \sum_{j=1}^{n} \lambda_j \qquad \text{and} \qquad \sum_{j=1}^{n} \|P_{\boldsymbol{\phi},d}\mathbf{x}_j\|^2 = \sum_{j=1}^{d} \sigma_j^2 = n \sum_{j=1}^{d} \lambda_j \,,$$

we may recast (2) in the more familiar form:

$$\begin{array}{r}\text{choose } d \text{ to be the} \\ \text{smallest integer such that}\end{array} \quad \frac{\displaystyle\sum_{j=1}^{d} \sigma_j^2}{\displaystyle\sum_{j=1}^{n} \sigma_j^2} \;=\; \frac{\displaystyle\sum_{j=1}^{d} \lambda_j}{\displaystyle\sum_{j=1}^{n} \lambda_j} \;\ge\; \gamma = 1 - \delta \,.$$

3. Centroidal Voronoi tessellations

A centroidal Voronoi tessellation (CVT) is a Voronoi tessellation of a given set such that the associated generating points are centroids, i.e., the centers of mass with respect to a given density function, of the corresponding Voronoi regions.

Given the *discrete* set of modified snapshots $W = \{\mathbf{x}_j\}_{j=1}^n$ belonging to $\mathbb{R}^N$, a set $\{V_i\}_{i=1}^k$ is a *tessellation* of W if $V_i \subset W$ for $i = 1, \dots, k$, $V_i \cap V_j = \emptyset$ for $i \neq j$, and $\cup_{i=1}^k V_i = W$. Given a set of points $\{\mathbf{z}_i\}_{i=1}^k$ belonging to $\mathbb{R}^N$ (but not necessarily to W), the *Voronoi set* corresponding to the point $\mathbf{z}_i$ is defined by

$$\begin{aligned} \widehat{V}_i = \{\, \mathbf{x} \in W \quad | \quad \|\mathbf{x} - \mathbf{z}_i\| \le \|\mathbf{x} - \mathbf{z}_j\| \quad &\text{for } j = 1, \dots, k,\ j \neq i, \\ &\text{where equality holds only for } i < j \,\}\,. \end{aligned} \tag{3}$$

Other tie-breaking rules for points equidistant to two or more of the $\mathbf{z}_i$'s can also be used. The set $\{\widehat{V}_i\}_{i=1}^k$ is called a *Voronoi tessellation* or *Voronoi diagram* of W.

Given a density function $\rho(\mathbf{y}) \ge 0$, defined for $\mathbf{y} \in W$, the mass centroid $\mathbf{z}^*$ of any subset $V \subset W$ is defined by

$$\sum_{\mathbf{y} \in V} \rho(\mathbf{y}) |\mathbf{y} - \mathbf{z}^*|^2 = \inf_{\mathbf{z} \in V^*} \sum_{\mathbf{y} \in V} \rho(\mathbf{y}) |\mathbf{y} - \mathbf{z}|^2\,, \tag{4}$$

where the sums extend over the points belonging to V and the set V^* can be taken to be V or it can be an even larger set such as $\mathbb{R}^N$. In the latter case, $\mathbf{z}^*$ is the ordinary mean

$$\mathbf{z}^* = \frac{\displaystyle\sum_{\mathbf{y} \in V} \rho(\mathbf{y}) \mathbf{y}}{\displaystyle\sum_{\mathbf{y} \in V} \rho(\mathbf{y})}\,;$$

in this case, $\mathbf{z}^* \notin W$ in general.

If $\mathbf{z}_i = \mathbf{z}_i^*$ for $i = 1, \dots, k$, where $\{\mathbf{z}_i\}_{i=1}^k$ is the set of generating points for the Voronoi tessellation $\{\widehat{V}_i\}_{i=1}^k$ and $\{\mathbf{z}_i^*\}_{i=1}^k$ are the set of mass centroids of the Voronoi regions, we refer to the Voronoi tessellation as a *centroidal Voronoi tessellation.* The concept of CVT's can be extended to more general sets, including regions in Euclidean spaces, and more general metrics. They have a variety of applications including data compression, optimal allocations of resources, cell division, territorial behavior of animals, optimal sensor and actuator location, and numerical analysis including both grid-based and meshfree algorithms for interpolation, multi-dimensional integration, and partial differential equations; see [7, 8, 9, 10, 11, 12, 19, 23].

Given the discrete set of points $W = \{\mathbf{x}_j\}_{j=1}^n$ belonging to $\mathbb{R}^N$, we define the error with respect to a tessellation $\{V_i\}_{i=1}^k$ of W and a set of points $\{\mathbf{z}_i\}_{i=1}^k$ belonging to W or, more generally, belonging to $\mathbb{R}^N$ by

$$\mathcal{F}\big((\mathbf{z}_i, V_i), i = 1, \dots, k\big) = \sum_{i=1}^k \sum_{\mathbf{y} \in V_i} \rho(\mathbf{y}) |\mathbf{y} - \mathbf{z}_i|^2\,. \tag{5}$$

It can be shown that a necessary condition for the error $\mathcal{F}$ to be minimized is that the pair $\{\mathbf{z}_i, V_i\}_{i=1}^k$ form a CVT of W. We note that the error (5) is also often referred to as the *variance, cost, distortion error,* or *mean square error.*

CVT's of discrete sets are closely related to optimal *k-means clusters* and Voronoi regions and centroids are referred to as *clusters* and *cluster centers*, respectively. Clustering analysis provides a selection of a finite collection of templates that well represent, in some sense, a large collection of data. To illustrate the connection between centroidal Voronoi diagrams and optimal k-means clustering, let us consider the case of $\rho(\mathbf{y}) \equiv 1$. Given any subset (cluster) V of W with m points, the cluster is represented by its arithmetic mean

$$\overline{\mathbf{x}}_i = \frac{1}{m} \sum_{\mathbf{x}_j \in V} \mathbf{x}_j$$

which corresponds to the mass centroid of V in the definition (4) with $V^* = \mathbb{R}^N$. The variance is given by

$$Var(V) = \sum_{\mathbf{x}_j \in V} |\mathbf{x}_j - \overline{\mathbf{x}}|^2$$

and, for a k-clustering $\{V_i\}_{i=1}^k$ (a tessellation of W into k disjoint subsets), the total variance is given by

$$Var(W) = \sum_{i=1}^{k} Var(V_i) = \sum_{i=1}^{k} \sum_{\mathbf{x}_j \in V_i} |\mathbf{x}_j - \overline{\mathbf{x}}_i|^2 \,. \tag{6}$$

(Compare (6) with (5).) The optimal k-clustering having the minimum total variance occurs when $\{V_i\}_{i=1}^k$ is the Voronoi partition of W with $\{\overline{\mathbf{x}}_i\}_{i=1}^k$ as the generators; i.e., using the variance-based criteria to define optimality, the optimal k-clustering is a centroidal Voronoi diagram.

3.1. Algorithms for constructing CVT's

There are several algorithms known for constructing centroidal Voronoi tessellations of a given set; see [7, 19, 20, 22]. One representative is *MacQueen's method* [22] (see also [7, 19]), a very elegant probabilistic algorithm which divides sampling points into k sets or clusters by taking means of clusters. A second representative is a deterministic algorithm known in some circles as *Lloyd's method* [20] (see also [7]) and which is the obvious iteration between computing Voronoi diagrams and mass centroids, i.e., for a given set of generators, they are replaced in an iterative process by the mass centroids of the Voronoi regions corresponding to those generators. There are various other methods based on the minimization properties of CVT's. A new probabilistic method has been suggested in [19] which may be viewed as a generalization of both the MacQueen and Lloyd methods; the method of [19] is amenable to efficient parallelization.

3.2. CVT and model reduction

We mentioned previously that CVT's have been used in data compression; one particular application was to image reconstruction; see [7]. Therefore, it is natural to examine CVT's in another data compression setting, namely model reduction. The idea, just as it was in the POD setting, is to extract, from a given set of snapshots $\{\mathbf{x}_j\}_{j=1}^n$ of vectors in $\mathbb{R}^N$, a smaller set of vectors also belonging to $\mathbb{R}^N$. In the POD setting, the reduced set was the d-dimensional set of vectors $\{\boldsymbol{\phi}_j\}_{i=1}^d$ defined in Section 2. In the CVT setting, the reduced set is the k-dimensional set of vectors $\{\mathbf{z}_k\}_{i=1}^k$ that are the generators of a centroidal Voronoi tessellation of the set of modified snapshots. Just as POD produced an optimal reduced basis in the sense that the error $\mathcal{E}$ defined in (1) is minimized, CVT produces an optimal reduced basis in the sense that the error $\mathcal{F}$ defined in (5) is minimized.

4. CVT based POD

We have already mentioned that the concept of centroidal Voronoi tessellations can be extended to more general notions of distance. This allows us to combine POD and CVT to take advantage of the best features of both approaches.

To generalize the CVT concept, there are essentially two ingredients that we need to redefine, namely, the concept of distance (which appears in the definition of and thus serves to define Voronoi tessellations) and the concept of centroid.

First, the square of the distance between the one-dimensional spaces spanned by two vectors $\mathbf{x}$ and $\mathbf{y}$ may be defined by [14]

$$\delta^2(\mathbf{x},\mathbf{y}) = 1 - \frac{|\mathbf{x}^T\mathbf{y}|^2}{\|\mathbf{x}\|^2\,\|\mathbf{y}\|^2}\,.$$

More generally, the square of the distance from a one dimensional subspace spanned by a vector $\mathbf{x}$ to a d-dimensional subspace $\mathcal{Z}$ is given by

$$\delta^2(\mathbf{x},\mathcal{Z}) = 1 - \frac{1}{\|\mathbf{x}\|^2}\sum_{i=1}^d |\mathbf{x}^T\boldsymbol{\theta}_i|^2\,,$$

where $\{\boldsymbol{\theta}_i\}_{i=1}^d$ forms an orthonormal basis for $\mathcal{Z}$. Then, given a set of vectors (e.g., modified snapshots) $W = \{\mathbf{x}_j\}_{j=1}^n$ and a set of d-dimensional subspaces $\{\mathcal{Z}_i\}_{i=1}^k$ (which are called the generators), we define the generalized Voronoi tessellation of W by

$$\mathcal{V}_i = \{\mathbf{x}_j \in W \;\mid\; \delta^2(\mathbf{x}_j,\mathcal{Z}_i) \le \delta^2(\mathbf{x}_j,\mathcal{Z}_\ell) \quad \forall\, \ell \ne i\,\} \qquad \text{for } i=1,\ldots,k\,.$$

A tie-breaking rule may be applied to insure that in the case equality holds in the above definition, each modified snapshot only belongs to one generalized Voronoi region.

Second, given a set of vectors $\mathcal{V} = \{\mathbf{x}_j\}$ that span an m-dimensional subspace of $\mathbb{R}^N$ (e.g., again, for us these are a subset of cardinality m of the modified snapshots), the concept of d-generalized centroid ($d \leq m$) of $\mathcal{V}$ may be defined by an orthonormal basis $\{\boldsymbol{\phi}_i\}_{i=1}^d$ which minimizes

$$\mathcal{D} = \sum_{\mathbf{x}_j \in \mathcal{V}} \|\mathbf{x}_j - P\mathbf{x}_j\|^2 \,,$$

where P denotes the projection operator into the d-dimensional subspace spanned by $\{\boldsymbol{\phi}_i\}_{i=1}^d$. For simplicity, we call such a centroid or basis the *d-g centroid of* $\mathcal{V}$. Note that, by Section 2, the optimal basis $\{\boldsymbol{\phi}_i\}_{i=1}^d$ is the d-dimensional POD basis for the set $\mathcal{V}$.

We are now ready to define CVT based POD. We note that the generators $\{\mathcal{Z}_j\}$ in general may not be required to have the same dimension. Thus, if k denotes the number of generators, we may use a multi-index $\mathbf{d} = \{d_i\}_{i=1}^k$ to replace the scalar index d.

Definition 4.1. *A set of finite subspaces* $\{\mathcal{Z}_j\}_{i=1}^k$ *with dimensions* $\mathbf{d} = \{d_i\}_{i=1}^k$, *respectively, along with the corresponding generalized Voronoi tessellation* $\{\mathcal{V}_j\}_{i=1}^k$ *is called a d-g* CVT *if and only if the* $\mathcal{Z}_i$*'s are themselves the d-g centroids of the* $\mathcal{V}_i$*'s.*

Definition 4.2. *The union of basis vectors corresponding to a d-g CVT is called a* CVT based POD *or a* centroidal Voronoi orthogonal decomposition (CVOD).

To recapitulate, CVOD can be viewed as a generalization of CVT for which the set W of modified snapshots is divided into k clusters or generalized Voronoi regions $\{\mathcal{V}_i\}_{i=1}^k$ and for which the generators are d_i-dimensional spaces each of which is spanned by the d_i-dimensional POD basis for the cluster. CVOD can also be viewed as a generalization of POD for which the set of modified snapshots is divided into k clusters and then a POD basis is separately determined for each cluster. In fact, if $d_i = 1$ for $i = 1, \ldots, k$, then CVT based POD reduces to the standard CVT of Section 3. On the other hand, if $k = 1$, then CVT based POD reduces to the standard POD of Section 2.

Algebraically, one may also interpret CVOD as follows. First, the original correlation matrix for the whole set of snapshots W is replaced by a block diagonal matrix with diagonal blocks being the correlation matrices for the snapshots in individual Voronoi sets $\{V_i\}$; then, the POD analysis is separately performed on each of the blocks. These Voronoi sets form a generalized centroidal Voronoi tessellation of W in the sense given in Definition 4.1. Thus, the role of CVT within CVOD may be viewed as providing, in some sense, an optimal clustering of the modified snapshots; the role of POD is then to provide an optimal reduced basis for each cluster.

There are cases where certain snapshots need to be weighted more heavily; thus, weighted POD's have been defined [5]. In light of the fact that a nonuniform density function can be used in the standard CVT construction, we may also define the weighted CVOD with a prescribed *discrete density* or a set of weights, i.e., we may minimize

$$\sum_{\mathbf{x}_j \in \mathcal{V}} \rho(\mathbf{x}_j)\delta^2(\mathbf{x}_j, \mathcal{Z}_i)$$

over a d_i-dimensional subspace of $\mathcal{V}$ for a given density function ρ.

4.1. Optimization properties of CVT based POD

Similar to the original CVT, the d-g CVT minimizes the error functional

$$\mathcal{G}\Big((\mathcal{Z}_i, \mathcal{V}_i),\, i = 1\ldots,k\Big) = \sum_{i=1}^{k} \sum_{\mathbf{x}_j \in \mathcal{V}_i} \delta^2(\mathbf{x}_j, \mathcal{Z}_i)$$

over all possible subdivisions of the set $\{\mathbf{x}_j\}_{j=1}^n$ of modified snapshots into k clusters $\{\mathcal{V}_i\}_{i=1}^k$ and all possible d_i-dimensional spaces $\mathcal{Z}_i$, $i = 1, \ldots, k$.

More generally, we have for a density function with values $\{\rho_j\}_{j=1}^n$, that the d-g CVT minimizes

$$\mathcal{G}\Big((\mathcal{Z}_i, \mathcal{V}_i),\, i = 1\ldots,k\Big) = \sum_{i=1}^{k} \sum_{\mathbf{x}_j \in \mathcal{V}_i} \rho_j \delta^2(\mathbf{x}_j, \mathcal{Z}_i)\,.$$

This optimization property is one of the key properties of CVT based POD that may make it very useful in practice.

The functional $\mathcal{G}$ also provides a natural error tolerance measure in the sense that

$$\mathcal{G}\Big((\mathcal{Z}_i, \mathcal{V}_i),\, i = 1\ldots,k\Big) = \sum_{i=1}^{k} |\mathcal{V}_i| \sum_{j=d_i+1}^{|\mathcal{V}_i|} \lambda_{i_j}\,,$$

where $|\mathcal{V}_i|$ denotes the cardinality of the Voronoi set or cluster $\mathcal{V}_i$ and the λ_{i_j}'s are the eigenvalues (in decreasing order) of the (weighted) local correlation matrix of the snapshots in the cluster.

In addition, for k large, it has been conjectured [7] that CVT's enjoys the equi-partition of error property; it is natural to extend such a conjecture to CVT based POD. Such an error equi-partition property leads naturally to adaptive strategies to refine the CVOD analysis. Intuitively, one may compare the relative *local error*

$$\frac{|\mathcal{V}_i| \displaystyle\sum_{j=d_i+1}^{|\mathcal{V}_i|} \lambda_{i_j}}{\mathcal{G}\Big((\mathcal{Z}_i, \mathcal{V}_i),\, i = 1\ldots,k\Big)}$$

with a given tolerance. One possible strategy is to enlarge the index d_i if the local error for the corresponding cluster (Voronoi set) $\mathcal{V}_i$ is much bigger than the errors for other clusters. On the other hand, if the error for one cluster is much smaller, then the index may be reduced. If the overall local errors are all very big, then besides enlarging d_i's, a larger value of k may also be desirable, i.e., more clusters may be used. While enlarging d_i may reduce the global error more efficiently, it also increases the computational cost in solving the eigenvalue problem. Thus, a balance needs to be maintained between enlarging k and increasing the d_i's.

4.2. Constrained CVT based POD

Sometimes, the physical system and thus the modified snapshots inherit certain symmetry properties, such as rotational symmetry, which are to be preserved by the selected representations [2]. In other situations, constraints needs to be enforced such as the vectors need to be divergence free or are constrained to a hypersurface, etc. [5]. For CVT, it is easy to modify the basic definition to allow additional constraints to be placed on the centroids [7, 9]. Thus, we can introduce the notion of *constrained CVOD* by extending the definition of generalized mass centroids from Euclidean spaces to other manifolds or more general constrained sets.

4.3. Algorithms for CVT based POD

A natural extension of the Lloyd method for computing standard CVT's is readily available. Let us begin with a given set of d_i-dimensional subspaces $\{\mathcal{Z}_i\}_{i=1}^k$. One may then construct the generalized Voronoi tessellation of the set of modified snapshots and then compute the *d-g* centroid of each generalized Voronoi set; these new centroids replace $\{\mathcal{Z}_i\}_{i=1}^k$ for the next iteration. More precisely, we have the following algorithm.

Algorithm 1. *Generalized Lloyd's method* (a deterministic iteration)

Given a set of modified snapshots $\{\mathbf{x}_j\}_{j=1}^n$ *and a discrete density function* $\{\rho_j\}_{j=1}^n$*, a positive integer* k*, and a multi-index* $\mathbf{d} = \{d_i\}_1^k$*;*

: 0. *choose an initial set of* k *subspaces* $\{\mathcal{Z}_i\}_1^k$ *with dimensions* $\mathbf{d} = \{d_i\}_1^k$*;*

: 1. *determine the generalized Voronoi tessellation*

$$\mathcal{V}_i = \{\mathbf{x}_j \in W \mid \delta^2(\mathbf{x}_j, \mathcal{Z}_i) \le \delta^2(\mathbf{x}_j, \mathcal{Z}_\ell) \quad \forall\, \ell \neq i\,\}$$

for $i = 1, \ldots, k$*, along with a tie-breaking rule;*

: 2. *find the d-g centroids* $\{\mathcal{Z}_i^*\}_{i=1}^k$ *of* $\{\mathcal{V}_i\}_{i=1}^k$*;*

: 3. *set* $\{\mathcal{Z}_i = \mathcal{Z}_i^*\}_{i=1}^k$ *as the new set of generators;*

: 4. *if the new generators meet some convergence criterion, terminate; otherwise, return to step* 1.

Note that the determination of the *d-g* centroids in Step 2 is equivalent to conducting a POD analysis of each of the Voronoi regions. Thus, one may also view Lloyd's method as an iterative procedure that decomposes the whole process of finding a *d-g* CVT into a sequence of POD analyses in sets with a smaller number of modified snapshots. Since the computational complexity of the POD analysis for n snapshots is related to that of solving the eigenproblem for an $n \times n$ matrix, it is very demanding computationally when n is large. The POD analysis of the smaller set of snapshots in Step 2, on the one hand, reduces the dimension of the matrix eigenproblem and thus requires less memory and computation time; on the other hand, it can also be done concurrently for each generalized Voronoi region, thus leaving much room for improvements in efficiency through parallelization.

The above algorithm has the nice feature that the *d-g* CVT error functional

$$\mathcal{G}\Big((\mathcal{Z}_i, \mathcal{V}_i),\, i = 1\ldots,k\Big) = \sum_{i=1}^{k} \sum_{\mathbf{x}_j \in \mathcal{V}_i} \rho_j \delta^2(\mathbf{x}_j, \mathcal{Z}_i)$$

decreases during the iteration. Moreover, as in the case of the original Lloyd iteration for the standard CVT [11], it can be shown that if the local minimizers of $\mathcal{G}$ share the same functional value, then the iteration is globally convergent. For the more general case, we also expect the iteration to converge to local minimizers based on earlier computational experiences, though no rigorous theory is yet available.

Lloyd's iteration can be regarded as a fixed point iteration between the generalized Voronoi generators and the *d-g* centroids. It can also be interpreted as a gradient descent method for the functional $\mathcal{G}$ with a fixed step size. Naturally, one can then define improvements to the Lloyd iteration. Extensions of other algorithms for computing the CVT's to the case of computing CVOD's are currently under investigation.

5. Remarks about model reduction

A reduced basis, be it POD or CVT or CVOD, can be used to define a low-order model in the usual manner. The partial differential equations governing the dynamics of the system are projected over the subspace spanned by a particular basis and a system of ordinary differential equations for the temporal modes is obtained. The projection is effected through the standard Galerkin method. For instance, let $F(t, X, u(X,t)) = 0$ be a system of partial differential equations with suitable boundary and/or initial conditions for the unknown function u. Here, t could be the time variable or some system parameter. Then, the CVOD based model reduction is performed as follows.

Algorithm 2. *CVOD based model reduction procedure*

: 1. *Construct a set of modified snapshots* $\{u_j\}_1^n$ *by solving (perhaps approximately) the system of differential equations for different values of* t.

: 2. *Calculate the CVOD for the set* $\{u_j\}_1^n$ *for some integer* k *and multi-index* $\{d_j\}_{j=1}^k$ *to obtain a set of basis vectors* $\{\phi_m\}_{m=1}^{|\mathbf{d}|}$.

: 3. *Solve the reduced system:*

$$\Big\langle \phi_m, F(t, X, \sum_{l=1}^{|\mathbf{d}|} \beta_l \phi_l) \Big\rangle = 0 \qquad \textit{for } m = 1, 2, \ldots, |\mathbf{d}|\ .$$

Naturally, an important question is why should one use CVOD instead of POD? Although answers have to be substantiated through numerical experiments, heuristically, one can make some arguments.

CVOD naturally introduces the concept of clustering into the decomposition. Imagine situations where intermittency is important or the dynamics are described by somewhat less related modes. By imposing a clustering, each sub-CVOD basis for a specific cluster can be used to capture the dynamics of that cluster.

As we have already mentioned, CVOD also reduces the amount of work relative to the full POD analysis. POD involves the solution of an $n \times n$ eigenproblem, where n is the number of snapshots; CVOD instead requires the solution of several smaller eigenproblems. CVT itself requires no eigenproblem solution.

Another interesting feature of CVOD which has been observed in other contexts, e.g., image processing [7], is that it avoids the over-crowding of the reduced basis into a few dominant modes which is viewed as one of the drawbacks of POD in some practical simulations.

References

[1] E. Arian, M. Fahl, and E. Sachs, *Trust-region proper orthogonal decomposition for flow control*, to appear.

[2] N. Aubry, W. Lian, and E. Titi, *Preserving symmetries in the proper orthogonal decomposition*, SIAM J. Sci. Comp., **14** (1993), 483–505.

[3] G. Berkooz, P. Holmes, and J. Lumley, *The proper orthogonal decomposition in the analysis of turbulent flows*, Ann. Rev. Fluid. Mech., **25** (1993), 539–575.

[4] G. Berkooz and E. Titi, *Galerkin projections and the proper orthogonal decomposition for equivariant equations*, Physics Letters A, **174** (1993), 94–102.

[5] E. Christensen, M. Brons, and J. Sorensen, *Evaluation of proper orthogonal decomposition-based decomposition techniques applied to parameter-dependent nonturbulent flows*, SIAM J. Sci. Comp., **21** (2000), 1419–1434.

[6] A. Deane, I. Kevrekidis, G. Karniadakis, and S. Orszag, *Low-dimensional models for complex geometry flows: applications to grooved channels and circular cylinders*, Phys. Fluids A, **3** (1991), 2337–2354.

[7] Q. Du, V. Faber, and M. Gunzburger, *Centroidal Voronoi tessellations: Applications and algorithms*, SIAM Review, **41** (1999), 637–676.

[8] Q. Du and M. Gunzburger, *Grid generation and optimization based on centroidal Voronoi tessellations*, to appear in Appl. Comp. Math., (2002).

[9] Q. Du, M. Gunzburger, and L.-L. Ju, *Meshfree, probabilistic determination of point sets and support regions for meshless computing*, Comp. Meth. Appl. Mech. Engrg., **191** (2002) 1349–1366.

[10] Q. Du, M. Gunzburger, and L.-L. Ju, *Centroidal Voronoi tessellations on surfaces*, to appear.

[11] Q. Du and T. Wong, *On the Lloyd's algorithm for computing CVTs*, to appear.

[12] A. Faulds and B. King, *Sensor location in feedback control of partial differential equation systems*, in Proc. 2000 IEEE CCA/CACSD, IEEE, Washington, (2000), 536–541.

[13] M. Graham and I. Kevrekidis, *Pattern analysis and model reduction: Some alternative approaches to the Karhunen-Lóeve decomposition*, Comp. Chem. Engng., **20** (1996), 495–506.

[14] G. Golub and C. van Loan, *Matrix Computation*, 2nd edition, John Hopkins, Baltimore, (1991).

[15] P. Holmes, J. Lumley, and G. Berkooz, *Turbulence, Coherent Structures, Dynamical Systems and Symmetry*, Cambridge University Press, Cambridge, (1996).

[16] P. Holmes, J. Lumley, G. Berkooz, J. Mattingly, and R. Wittenberg, *Low-dimensional models of coherent structures in turbulence*, Phys. Rep., **287** (1997), 337–384.

[17] K. Kunisch and S. Volkwein, *Control of Burger's equation by a reduced order approach using proper orthogonal decomposition*, JOTA, **102** (1999), 345–371.

[18] K. Kunisch and S. Volkwein, *Galerkin proper orthogonal decomposition methods for parabolic problems*, Spezialforschungsbereich F003 Optimierung und Kontrolle, Projektbereich Kontinuierliche Optimierung und Kontrolle, Bericht Nr. 153, (1999), Graz.

[19] L.-L. Ju, Q. Du, and M. Gunzburger, *Probabilistic methods for centroidal Voronoi tessellations and their parallel implementations*, to appear in Parallel Comput.

[20] S. Lloyd, *Least squares quantization in PCM*, IEEE Trans. Infor. Theory, **28** (1982), 129–137.

[21] J. Lumley, *Stochastic Tools in Turbulence*, Academic, New York, (1971).

[22] J. MacQueen, *Some methods for classification and analysis of multivariate observations*, Proc. Fifth Berkeley Symposium on Mathematical Statistics and Probability, Vol. I, Ed. by L. Le Cam and J. Neyman, University of California, Berkeley, (1967), 281–297.

[23] A. Okabe, B. Boots, K. Sugihara, and S. Chiu, *Spatial Tessellations: Concepts and Applications of Voronoi Diagrams*, 2nd edition, Wiley, Chichester, (2000).

[24] H. Park and D. Cho, *Low dimensional modeling of flow reactors*, Int. J. Heat Mass Transf., **39** (1996), 3311–3323.

[25] H. Park and O. Kim, *Reduction of modes for the control of viscous flows*, to appear.

[26] H. Park and M. Lee, *Control of Navier-Stokes equations by means of mode reduction*, to appear.

[27] H. Park and W. Lee, *Boundary optimal control of natural convection by means of mode reduction*, to appear.

[28] M. Rathinam and L. Petzold, *A new look at proper orthogonal decomposition*, to appear.

[29] M. Rathinam and L. Petzold, *Dynamic iteration using reduced order models: A method for simulation of large scale modular systems*, to appear.

[30] S. Ravindran, *Proper orthogonal decomposition in optimal control of fluids*, Int. J. Numer. Meth. Fluids, **34** (2000), 425–448.

[31] S. Ravindran, *Reduced-order adaptive controllers for fluid flows using POD*, to appear.

[32] J. Rodríguez and L. Sirovich, *Low-dimensional dynamics for the complex Ginzburg-Landau equations*, Physica D, **43** (1990), 77–86.

[33] L. Sirovich, *Turbulence and the dynamics of coherent structures, I-III*, Quart. Appl. Math., **45** (1987), 561–590.

[34] N. Smaoui and D. Armbruster, *Symmetry and the Karhunen-Loève analysis*, SIAM J. Sci. Comput., **18** (1997), 1526–1532.

[35] S. Volkwein, *Optimal control of a phase field model using the proper orthogonal decomposition*, ZAMM, **81** (2001), 83–97.

[36] S. Volkwein, *Proper orthogonal decompositionand singular value decomposition*, Spezialforschungsbereich F003 Optimierung und Kontrolle, Projektbereich Kontinuierliche Optimierung und Kontrolle, Bericht Nr. 153, (1999), Graz.

Department of Mathematics
Penn State University
University Park PA 16802, USA
E-mail address: qdu@math.psu.edu

Department of Mathematics
Iowa State University
Ames IA 50011-2064, USA
E-mail address: gunzburg@iastate.edu