Reduced Order Modeling of PDE's Through Clustering

ISC5316-01: Applied Computational Science II

....

John Burkardt
Department of Scientific Computing
Florida State University

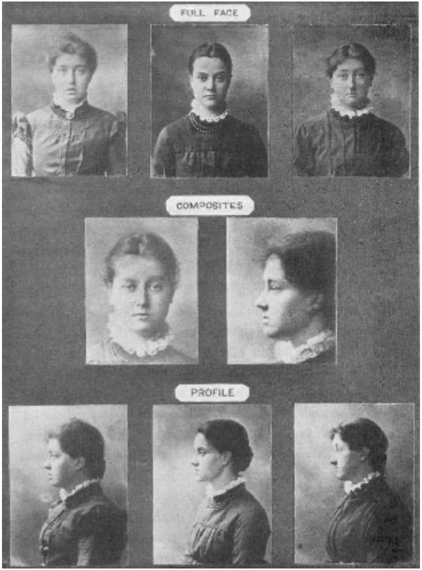
https://people.sc.fsu.edu/$\sim$jburkardt/presentations/...
...rom_2003_bozeman.pdf

....

July 2003

February 1, 2024

Francis Galton's Composite Photographs



Inspiration

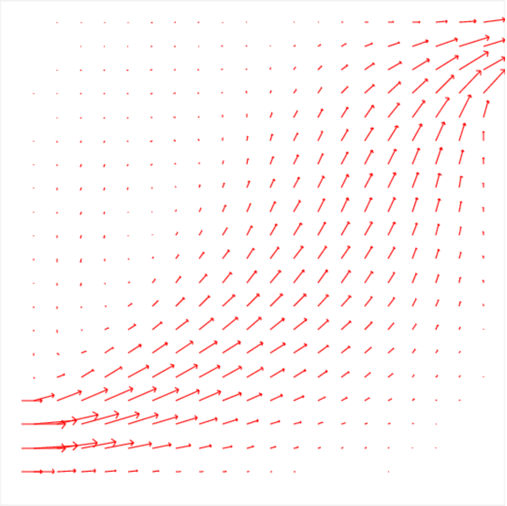
"There comes a time when, for every addition of knowledge, you forget something that you knew before. It is of the highest importance, therefore, not to have useless facts elbowing out the useful ones!"

–Sherlock Holmes in "A Study in Scarlet"

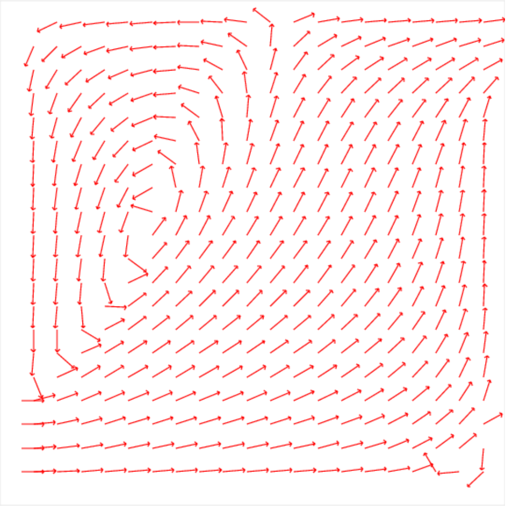
Outline

- The IN/OUT Steady State Fluid Problem
- Case Study: SVD
- Sample Flow Problems
- Generating Snapshots
- POD Analysis of the IN/OUT Problem
- Strengths and Weaknesses of POD
- Clustering by K-Means
- Clustering Analysis of the IN/OUT Problem
- Strengths and Weaknesses of Clustering
- Solving New Reduced Problems
- Conclusion

The IN/OUT Steady State Vector Field



The IN/OUT Steady State Direction Field



The Setting

- Our system is time-dependent, and parameterized;
- We have "snapshots" at various times and parameters;
- (Snapshots from finite element simulation, or windtunnel).
- We need trajectories for many parameter values;
- Generic finite elements too slow or too expensive.

Observations

- Most finite element "degrees of freedom" are not used;
- Solutions inhabit a low dimensional space of "likely behaviors";
- New behaviors emerge with increasing energy;
- Trajectories: energy moves to preferred behavior subspace;
- A basis using "preferred behaviors" would be tiny.

Goals

- **System exploration + experiment design**
- **Data analysis, weighting, compression**
- **Choice of reduced model**
- **Computation of approximate trajectories**

Long range goal:

- *Closed loop control of parameterized flow.*

Goals (Comments)

System exploration + experiment design:

- *may we choose which snapshots to create?*;
- *which parameter and boundary values should be tested?*
- *can we identify and weight "important" behaviors?*

Data analysis, weighting, compression:

- *a behavior is important if it has high energy*;
- *a behavior is important if it is "far away" from others*;
- *a behavior is important if it occurs often*;

Model reduction:

- *reduce existing snapshot data to a smaller representative set*;
- *find a low dimensional basis for fluid behavior*;

Approximate trajectories:

- *approximate points on known trajectory*;
- *calculate points on unknown trajectory using reduced model*;

Case Study: SVD

Given an m by n matrix A, regard the columns as "behaviors".

The energy associated with a direction is the projection of all columns onto that direction.

The SVD produces a factorization:

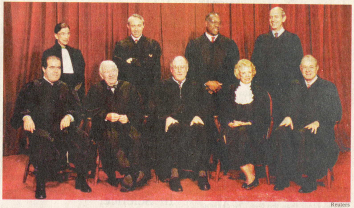
$$A = U \cdot \Sigma \cdot V' \tag{1}$$

- The leading columns of U are the "preferred behaviors";
- The diagonal of Σ is an energy or importance weight;
- A reduced model of A can be constructed;

A 9-D Space with 2-D Subspace

A Mathematician Crunches the Supreme Court's Numbers

United Press International



Reuters

Although the Warren Supreme Court of 1967, left, and current Rehnquist court are ideologically quite different, their dynamics were found similar, with two dominant voting patterns.

By NICHOLAS WADE

The nine justices of the United States Supreme Court do not always issue unanimous decisions, nor do their votes match the pattern that would be shown by nine totally independent thinkers.

For those who have ever wondered where the court's decisions might fall on the spectrum between monolithic unity and total randomness, an answer is now in.

The voting pattern of the Rehnquist court over the last nine years "shows that the court acts as if composed of 4.68 ideal justices," says Dr. Lawrence Sirovich, a mathematician at the Mount Sinai School of Medicine in Manhattan whose day job is figuring out how the visual system works.

By another measure, "the decision space of the Rehnquist court requires only two dimensions for its description," he writes in the issue of The Proceedings of the National Academy of Sciences being published today.

Nine independent thinkers who focus solely on the merits of cases might be expected to vote in all possible combinations over a long enough

But the actual number of voting patterns is very much less, as if generated by a smaller number of wholly independent individuals.

Analyzing nearly 500 opinions issued since 1995 — the court membership has not changed since Justice Stephen G. Breyer joined it in 1994 — Dr. Sirovich calculates, based on information theory, that 4.68 ideal justices would have produced the same diversity of decision making.

By ideal, Dr. Sirovich means a justice whose voting is uncorrelated with any other's. His measure, thus, points up the high degree of correlation in the court's voting pattern.

Looking only at final votes, of course, ignores many nuances about how the court operates in reality, Dr. Sirovich acknowledged. Justices often file concurring opinions or dissents, allowing them to cooperate in the creation of a majority while preserving independence in outlining the reasons behind a vote.

Considering the decisions with another technique known as singular value decomposition, Dr. Sirovich has also found considerably less diversity than might be expected.

It would take nine dimensions for a mathematician to describe the voting patterns of the nine uncorrelated

able as a left-right point of view," said Jeffrey Segal, a political scientist at the State University of New York at Stony Brook. "He comes up with two dimensions but doesn't label them."

Dr. Yochai Benkler of the New York University Law School said the analysis was important, because instead of starting from some theory about the politics of the court, Dr. Sirovich had used a purely mathematical analysis, yet one whose results fit with common sense observation.

"What you see here is not someone trying to prove a point, but someone who has said, 'Beyond the stories, this is the math of how people behave and how they ally,'" Professor Benkler said. "And the interesting thing here is the fit between the mathematical observation and the widely held intuition about the politics of adjudication."

Dr. Sirovich's specialty is in extracting patterns from information. "He is the master of S.V.D.'s," Dr. Mitchell Feigenbaum, a mathematician at the Rockefeller University, said of singular value decomposition.

Dr. Sirovich introduced the technique for use in machine recognition of human faces, and it is now the

resentations of information — in those 100 dimensions represent basic faces from which any other face can be constructed.

The same method applied to the Supreme Court's voting patterns enables vectors in just two dimensions to describe almost all its decisions. One vector is very close to the vector that represents a unanimous decision. The other lies near the vector representing the most common of the court's 5-to-4 voting patterns, including the one that decided the 2000

Two very different courts, one very similar result.

presidential election.

Each justice's vote can be regarded as fixed mixture of those two voting patterns, Dr. Sirovich writes. Only three decisions out of 468 are not fully captured by his two vectors. One, Rogers v. Tennessee, split both sides of the usual 5-to-4 vote, with Chief Justice William H. Rehnquist

pendently," Dr. Sirovich said.

Dr. Sirovich applied his two yardsticks — information theory and single value decomposition — to the Warren court for two periods when its composition was unchanged, from 1959 to 1961 and from 1967 to 1969. The first Warren court was somewhat more diverse than the Rehnquist court, operating as if with 5.16 ideal justices.

But its dynamics were quite similar, with two dominant voting patterns, William O. Douglas being the regular dissenter, as Justice Stevens is in the Rehnquist court, and Tom C. Clark and Potter Stewart playing the role of the swing votes, similar to that of Justices Anthony M. Kennedy and Sandra Day O'Connor in the present court.

The second Warren court was much more liberal than the first one, and some players switched roles. Justices Douglas, Earl Warren and William J. Brennan Jr., often in minority in the first Warren court, became part of the 5-to-4 majority in the second. Justice Thurgood Marshall, in the second court, dissented just once.

The surprising similarity in the voting dynamics of the Warren and

PDE Data Reduction: POD

Each column of an m by n matrix A is a solution or "snapshot".

1. compute the m by m matrix $B = A * A'$,
2. determine eigen-decomposition $B = X * \Lambda * X'$,
3. each eigenvector x_i is a mode;
4. each eigenvalue λ_i is the *energy* of a mode;
5. select eigenvectors by energy to get dominant modes.

In fact, we can carry out the same computation by computing the SVD of A. The $\sigma' s$ are the square roots of the energy, and the corresponding left singular vectors of U are the mode vectors.

An Algorithm of Many Names!

The method has many names, depending in part on the field (meteorology, statistics, biomedical data analysis, mechanical engineering):

- **KL:** the Karhunen Loève analysis;
- **POD:** Principal Orthogonal Direction
- **POD:** Proper Orthogonal Decomposition
- **EOF:** Empirical Orthogonal Functions
- **MDS:** Multidirectional Scaling
- **PCA:** Principal Component Analysis

Sample Flow Problems

- **BACKSTEP** used to model flow separation;
- **CAVITY**, driven cavity;
- **IN/OUT**, flow into box and out again;
- **TCELL**, driven cavity + flow region above it;

Most of our results will concern the **INOUT** problem:

- square 2D box, 1 by 1 units;
- parabolic inflow from the lower right
- inflow amplitude parameter α.
- outflow on the upper right.
- dynamic viscosity $= 1/300$;

An IN/OUT Velocity Field

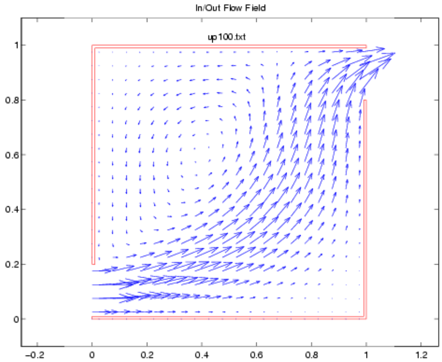


Figure: Velocity field at time step 100 for IN/OUT

The inflow parameter is $\alpha = 5/3$ for this snapshot;
MATLAB's automatic scaling makes the largest vector one cell long. This has been doubled for visibility, and 3/4 of the data has been deleted!

An IN/OUT Direction Field

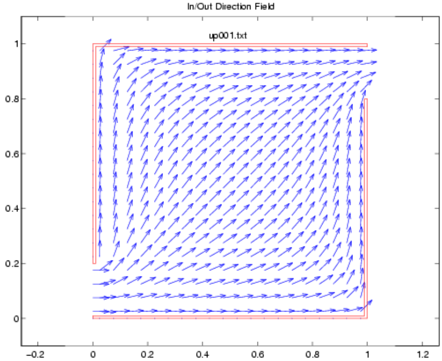


Figure: Direction field at time step 1 for IN/OUT

All vectors are shown with the same length. This allows smaller magnitude portions of the flow to be visible.

Generating Snapshots

Collecting snapshots involves observing the system, if possible, over a range of parameters and times.

- initial condition is steady state ($\alpha = 1/3$);
- time step $\Delta t = 0.01$;
- solve with $\alpha = 5/3$ for 250 time steps.
- continue with $\alpha = 1/3$ for 250 time steps.
- two impulsive changes to α intended to excite many modes;

- 41 by 41 evenly spaced grid of nodes;
- 800 elements (quadratic velocity, linear pressure);

This system has 3,362 degrees of freedom. We are interested in approximating it with 2, 4, 8 or 16 degrees of freedom!

IN/OUT Animation

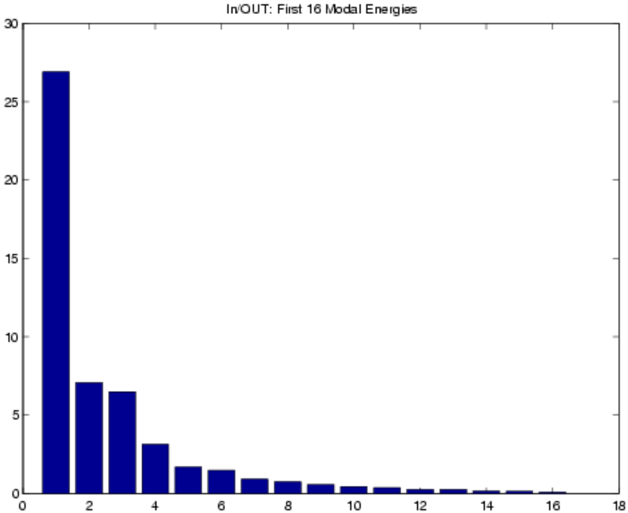
This animation displays the velocity direction field in the IN/OUT flow. The sharp change in the value of α is quite noticeable.

For clarity, only 1/4 of the nodes are shown.

This animation is available as an MPEG-4 file at

http://people.sc.fsu.edu/∼jburkardt...
/datasets/inout_flow/inout_flow_movie.html

The IN/OUT Modal Energies (Figure)



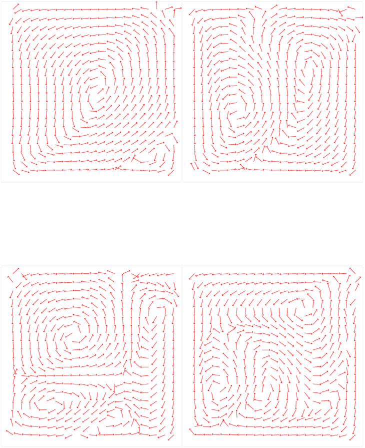
The IN/OUT Modal Energies (Table)

For this problem, the energy is concentrated in a few modes. The 12-th mode has 1/100 the energy of the first. If energy "prefers" to stay in the first few modes, everything is fine, but if changes to the parameters mean the energy moves to other modes, we may be in trouble!

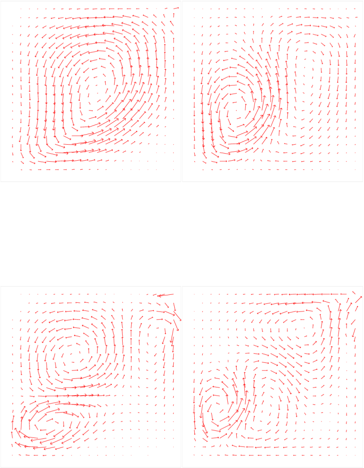
Vector	Singular value	Vector	Singular value
1	26.9107	9	0.5738
2	7.0878	10	0.4570
3	6.5015	11	0.3736
4	3.1420	12	0.2749
5	1.6973	13	0.2707
6	1.4947	14	0.1787
7	0.9253	15	0.1453
8	0.7592	16	0.0994

The IN/OUT Modal Vectors

The modes read left to right, and are plausibly organized by complexity, which we presume reflects their natural energy capacity.



The IN/OUT Modal Vectors



Strengths of POD

POD strengths are largely the strengths of eigen-analysis:

- the data is easy to set up, process and use;
- LAPACK provides standard efficient software;
- parallel LAPACK available;
- basis vectors are orthonormal;
- basis vectors are ordered by energy;
- basis vectors have physical meaning as modes;
- the basis vectors are "nested", the set of 8 is created by adding 1 to the set of 7.
- the weights have physical meaning as energy;
- the energies give a natural way to measure approximation error;

Weaknesses of POD

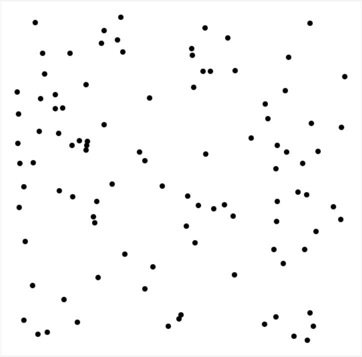
POD weaknesses include:

- requires solution of large eigensystem or SVD;
- adding just one more snapshot requires recomputation;
- high energy modes may not be the only interesting phenomena;
- rapid drop-off in energy may not be be typical;
- limited modeling of boundary conditions;

PDE Data Reduction: Clustering

We can also seek "preferred behaviors" by looking for **clusters** in the data, using **K-means**.
We seek k "generators" G minimizing

$$E(X,G) = \Sigma_{i=1}^{N} ||X(i) - nearestG||^2 \tag{2}$$



Clustering by K-Means

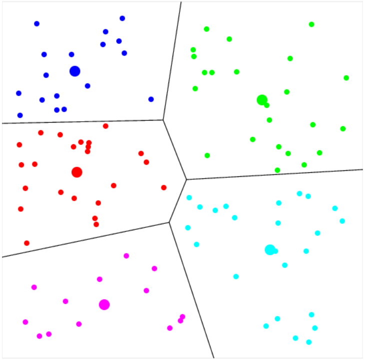


Figure: 100 random points clustered by K-Means

The K-Means Algorithm

```
Given N points X to be assigned to K clusters:

Assign each X to a random cluster;

Do Forever

  For each point X

    For each cluster C
      determine energy change if X -> C

    Move X to its preferred cluster.

  If no point moved, exit;

End Forever
```

Comments on K-Means

- if *any* X switches clusters, affects all points.
- Algorithm terminates at "local minimum" of cluster energy;
- Result depends on initial cluster assignment;
- The weaker H-means algorithm can precondition.

Parallel execution?

- H-means can be parallelized.
- K-means is **not** parallel;
- Want to do several ($\sim$15 to 30) cycles, these can be in parallel.

Questions:

- Should we normalize the snapshot data?

The IN/OUT Clustering Energies

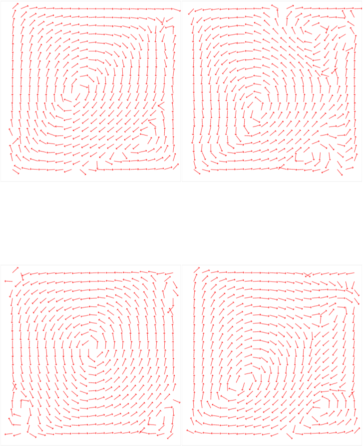
Note that the clusters correspond to time intervals. This makes the clustering plausible. In some runs, the two relatively quiet "tail ends" (here, clusters 4 and 5) are clustered together.

Table: Cluster Energy

Vector	Cluster energy	Population	Extent
1	6.30	15	[1, 15]
2	4.79	27	[37, 63]
3	3.77	30	[267,296]
4	3.28	153	[348,500]
5	4.68	187	[64,250]
6	5.54	21	[16, 36]
7	3.43	51	[297,347]
8	4.75	16	[251,266]
Total	36.58	500	[1,500]

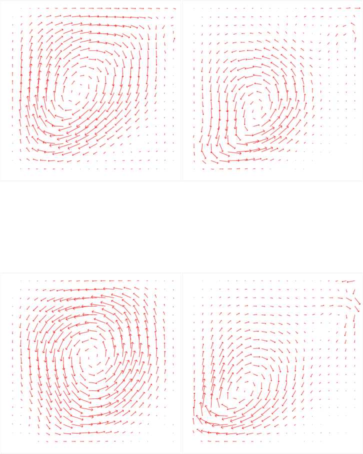
The IN/OUT 8-Cluster Vectors

The first four of the set of 8.
Here, the CVT basis vectors are not orthogonalized, so they share a large component in common.



The IN/OUT 8-Cluster Vectors

The first four of the set of 8.



Clustering Strengths

Clustering strengths include:

- can cluster anything for which you can define a distance;
- a clustering can be updated, rather than being recomputed, when data is added (or removed);

Clustering Weaknesses

Clustering weaknesses include:

- vectors are not orthogonal;
- the cluster energies do not yield the same information as POD energies;
- the K-means computation must be designed by hand;
- possibility of local minima;
- limited parallelizability (H-means versus K-Means)
- the solution of the 8 vector problem is not the 7 vector problem plus 1!

Solving New Reduced Problems

The same procedure for POD or clustering (once cluster vectors are orthonormalized).

Basis vectors Z_i form reduced order finite element basis functions $z_i(\mathrm{x})$.

A typical reduced model solution:

$$u(\mathrm{x}, t) = \beta(t)v(\mathrm{x}) + \Sigma_{i=1}^{d} c_i(t) z_i(\mathrm{x}) \tag{3}$$

where v is a flow solution satisfying the boundary conditions.

- The momentum equations are formally the same;
- The continuity equation is unnecesary;
- The boundary conditions are taken care of by β. (not so easy for complicated boundary cases!)

System was solved using 4th order Runge-Kutta.

Simulation of IN/OUT with Varying Inflow

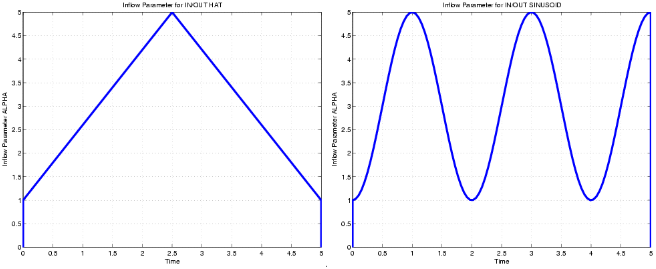


Figure: The time history of α for HAT and SINUSOID problems

This animation shows the simulation of the IN/OUT flow with a sinusoidal inflow parameter α.

Simulation of IN/OUT with Varying Inflow

This flow field is to be approximated by a reduced order model POD or CVT basis set derived from the original set of snapshots in which α took on the values 1/3 and 5/3.

This animation is available as an MPEG-4 file at

http://people.sc.fsu.edu/∼jburkardt/datasets/...
inout_flow/inout_case2_movie.html

Errors of IN/OUT with SINUSOIDAL Inflow

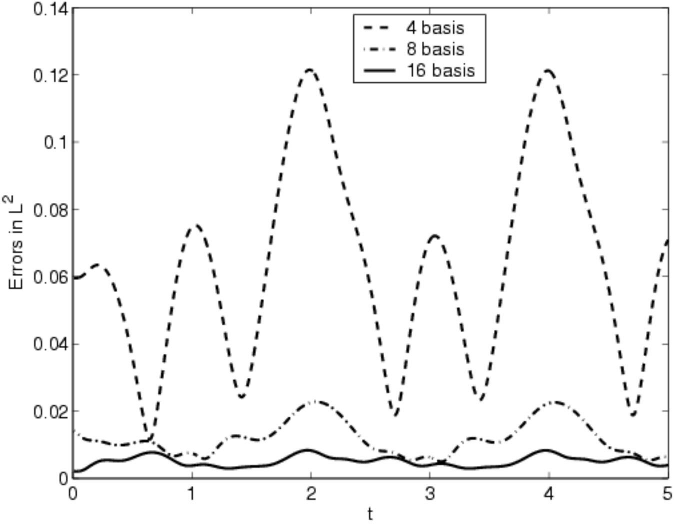


Figure: CVT approximation error, snapshots were NOT normalized

Conclusion

- better methods of exhibiting modes than by impulse?;
- weighting by energy or distance?;
- for clustering, normalization of snapshot data helps and hurts;
- treatment of problems with complex boundary conditions?
- can we adjust basis functions to account for parameter change?

References

Centroidal Voronoi Tessellation-based Reduced-Order Modeling of Complex Systems,
Burkardt, Gunzburger, Lee,
submitted to SIAM Journal of Scientific Computing.

The PDE snapshots available in:

http://people.sc.fsu.edu/∼jburkardt/datasets/...
cavity_flow + inout_flow + tcell_flow