PCA & Kriging for model reduction

G. Aversano[*,a], A. Parente[a]

[a] *Université Libre de Bruxelles*

Introduction

Simulations of reacting flows for predictive purposes and engineering designs are much often a challenging task. Knowing a system's solution in every point in the parameter-space may be prohibitive. Many techniques exist in order to find a trade-off between solution accuracy and reasonable computational cost. Principal Component Analysis combined with Kriging interpolation can be viewed as one of the most promising methods. The idea is simple: from a *small* data-set, PCA is performed to reduce the dimension (though retaining the physics) of a system, then Kriging is used to predict the system's state at an untried point in the parameter-space.

PCA

Principal Component Analysis (PCA) identifies correlations among the variables defining the state space. As a result, a new coordinate system is identified in the directions of maximal data variance, which allows less important dimensions to be eliminated while maintaining the primary structure of the original data.
The PCA algorithm can be summarized as follows. We take M observations of N variables and make up the data-set $\mathbf{Y}$ ($N \times M$). The data-set is centered and scaled (to its variance): $\tilde{\mathbf{y}}_j = \frac{(\mathbf{y}_j - \bar{\mathbf{y}})}{d_j}$ where $\mathbf{y}_j$ is the j-th column of $\mathbf{Y}$.
The covariance matrix is evaluated. This can be done in *spatial* domain: $\mathbf{S} = \frac{1}{M-1}\tilde{\mathbf{Y}}\tilde{\mathbf{Y}}^T$ or *temporal* domain (Sirovich): $\mathbf{S} = \frac{1}{N}\tilde{\mathbf{Y}}^T\tilde{\mathbf{Y}}$. Then, the eigenvectors of $\mathbf{S}$ are the PCA modes $\boldsymbol{\phi}_i$. Only $q \ll N$ modes are chosen, the ones with the biggest eigenvalues, and stored in the columns of the matrix $\boldsymbol{\Phi}$. The PCA coefficients relative to the j-th observation are evaluated by $\boldsymbol{\alpha}_j = \boldsymbol{\Phi}^T\tilde{\mathbf{y}}_j$. These vectors are stored in the columns of $\mathbf{A}$ ($q \times M$), which can be viewed now as the data-set instead of $\tilde{\mathbf{Y}}$.

The PCA method does not explicitly guarantee the respect of the physical laws involved, such as conservation of mass. To guarantee that, Constrained-PCA is performed: $\boldsymbol{\alpha} = min_{\boldsymbol{\alpha}}\|\mathbf{y} - \mathbf{y}^{pca}\|^2 \quad s.t.\ \mathcal{L}(\boldsymbol{\alpha}) = 0$, where $\mathcal{L}(\boldsymbol{\alpha}) = 0$ represents the constraints adopted. Accuracy may be lower, but no physical laws are infriged, provided that the optimization problem is solved correctly. It is important to stress this out, as the solution of the optimization problem may not be trivial.
Improved performance can be achieved by applying a Local-PCA. Data are optimally divided into clusters, and PCA is performed in each cluster indipendently. The idea is that if the local regions are small enough, the data manifold will not curve much over the extent of the region, and the PCA model will be a good fit.

Kriging

The data-set $\mathbf{A}$ (the values of the PCA coefficients) is used to build a response surface which maps from the parameter-space to $\boldsymbol{\alpha}$. If $\mathbf{p}^*$ is an untried point in the parameter-space, the corresponding $\boldsymbol{\alpha}^*$ is predicted, which in turn leads to $\mathbf{y}^*$, whereas $\boldsymbol{\Phi}$ remains unchanged. This way the physics of the system is stored in $\boldsymbol{\Phi}$ and unaffected by the interpolation. For a scalar target y and a data-set $\mathbf{y}$, the Kriging predictor is: $\tilde{y}(\mathbf{x}) = \mathbf{f}(\mathbf{x})^T\tilde{\boldsymbol{\beta}} + \mathbf{r}(\mathbf{x})^T\mathbf{R}^{-1}(\mathbf{y} - \mathbf{F}\tilde{\boldsymbol{\beta}}) = \mathbf{f}(\mathbf{x})^T\tilde{\boldsymbol{\beta}} + \mathbf{r}(\mathbf{x})^T\boldsymbol{\gamma}$. A target is modeled as a trend function $\mu(\mathbf{x}) = \mathbf{f}(\mathbf{x})^T\tilde{\boldsymbol{\beta}}$ plus a stationary zero-mean stochastic field. The trend function is a weighted sum with weights $\tilde{\boldsymbol{\beta}}$ (determined by a Generalizes Least Square) of some linear or quadratic functions of the inputs $\mathbf{x}$. Differences from the trend are modeled with a white noise. The function describing the correlation matrix $\mathbf{R}$ and cross-correlation vector $\mathbf{r}(\mathbf{x})$ has to be chosen *arbitrarily*, though with an uspecified set of hyperparameters which are determined by a Maximum Likelihood Estimator.

Application

With the equivalence ratio ψ as varying parameter, both PCA-K and CPCA-K have been applied to

*G. Aversano
Email address: Gianmarco.Aversano@ulb.ac.be

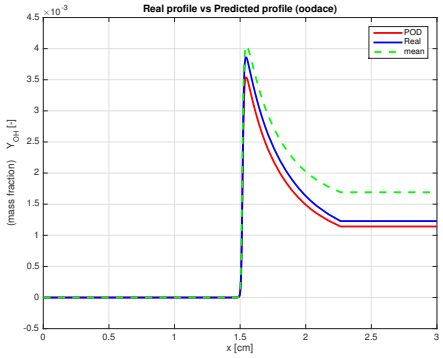

Figure 1: Prediction for OH relative to the case with the largest global error.

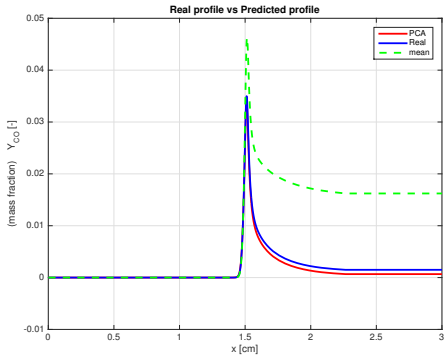

Figure 2: Prediction for CO relative to the case with the largest global error.

1-D laminar flames (GRI 3.0 mechanism) solved in OpenSMOKE. In the range $\psi = 0.8 \div 1.2$, both methods performed effictively with $q = 5$ modes, predicting the system's state at untried points of ψ with a global error below 1% and 4%, respectively. The algorithm consisted in using some samples in $\psi = 0.8 \div 1.2$ to build the PCA-K model, and the remainder for its validation.

In Fig. 3, it is possible to notice how CPCA-K may deliver some inaccuracies in order to, for example, keep the mass fractions positive. Both PCA-K and CPCA-K generally provide very good results and show discrepancies only for few cases. This can be attributed to the range of ψ the model has to deal with and it can be likely improved using Local-PCA.

The Local-PCA formulation combined with Kriging is expected to deliver prosiming results.

The method has also been used to model the non-linear source terms of the system's governing equations. This can be very useful when directly solving the system's equations as it allows to avoid evaluating these highly non-linear terms, thus leading to computational savings.

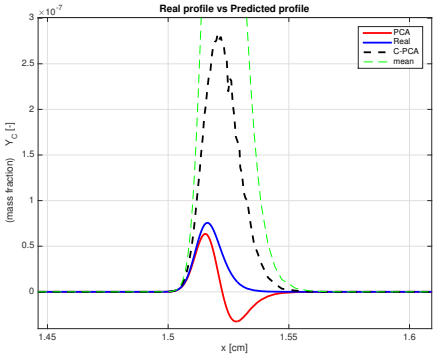

Figure 3: Prediction for C relative to the case with the largest global error. Both PCA-K and CPCA-K predictions are reported.

Conclusion

The results show that the method presented here is promising and versatile, as it can be helpful in different applications and set up according to one's needs. Whether PCA, CPCA or Local-PCA is applied, combined with Kriging, the presented method proved to be a smart choice for model reduction, combining simplicity and efficiency. Further developments and applications will probably confirm what has been asserted so far, with particular attention to application of the method to much higher-dimensional systems, with more than one avarying parameters.

Acknowledgments

This project has received funding from the European Unions Horizon 2020 research and innovation programme under the Marie Sklodowska-Curie grant agreement No 643134.

References

[1] Cuoci A., Frassoldati A., Faravelli T., Ranzi E., *A computational tool for the detailed kinetic modeling of laminar flames: application to C2H4/CH4 coflow flames*, Combustion and Flame (2013), DOI: 10.1016/j.combustflame.2013.01.011

[2] Kambhatla, N., & Leen, T. K. (1997). *Dimension Reduction by Local Principal Component Analysis*. Neural Computation, 9(7), 1493-1516.

[3] Kennedy, M., & O'Hagan, A. (2001). *Bayesian calibration of computer modes*. J. R. Stat. Soc. B (2001), 63, part 3, pp 425-464.

[4] Isaac, B. J., Coussement, A., Gicquel, O., Smith, P. J., Parente, A. (2014). *Reduced-order PCA models for chemical reacting flows.*. Combustion and Flame, 161(11), 2785-2800.