High-dimensional Stochastic Interpolation, Optimization and Inversion via Manifold Learning

CASIS 2018- Annual Signal & Image Sciences Workshop

Charanraj A. Thimmisetty\textsuperscript{a}, Xiao Chen\textsuperscript{a}, Roger Ghanem\textsuperscript{b}, Joshua A. White\textsuperscript{a}, Charles H. Tong\textsuperscript{a} and Panagiotis Tsilifis\textsuperscript{b}

\textsuperscript{a} Lawrence Livermore National Laboratory, USA
\textsuperscript{b} University of Southern California, USA

May 23, 2018
High-dimensionality poses challenges in several scientific and engineering applications such as design, model parameter inversion and interpolation.

Common challenges:
- **Data** (measurements)
  - High dimensional
  - Nonlinearly correlated
  - Non-Gaussian
  - Noisy and sparse
  - Expensive to obtain
- **Simulations**
  - Computationally intensive
  - Models have uncertainties

Need:
- Understand patterns in the data and build data driven models
- Produce reliable and faster solutions for timely analysis

**Elasticity inversion**

Target: Invert elasticity field  
Data: Displacement measurements

**Oil well placement**

Goal: Place injection and production wells that maximizes production random permeability field

**Subsurface Characterization**

Goal: Given proxy measurements, obtain wave velocity characterization of the subsurface
Manifold learning can provide us easy-to-search space and true metric

- Manifold: A topological space that resembles Euclidean space near each point
- Intrinsic geometry: Geometry experienced by the inhabitants
- Few intrinsic parameters: Length, Area, Gaussian curvatures
- Most of the dataset require non-linear manifold learning techniques to identify underlying manifold
- Non-linear manifold techniques explored in this work include diffusion maps and kernel principle component analysis (KPCA)

Euclidian distance have different meaning on both geometries even though inhabitants have similar experience.

Example: Left original manifold (left), 1D projection using diffusion maps (middle), 1D projection using PCA (right)
Stochastic intrinsic interpolation: Gaussian process built on the manifold uses true metric and provides probabilistic estimates

- Gaussian Process Regression (GPR)/ Kriging acts as interpolation tool for standard kernels
- Advancements in data acquisition techniques provide high-dimensional proxy datasets
- GPR produces uninformative predictions as the dimension of predictors increases
- Soft computing methods such as neural network and support vector machines by default will not provide a full probabilistic predictions
- We propose a novel intrinsic interpolation method
- We obtain a low dimensional embedding of the data using diffusion maps
- GPR is built on the manifold that takes into account of the distance on manifold (diffusion distance) instead of the Euclidean distance
Example: Characterization of geo-physical parameters using a suite of well-log measurements

Scatter plot of first three diffusion & principle components

DTSM predictions in Original space (left), Principle component space (middle) and diffusion space (right)

- Goal: To build a regression model for the wave velocities
- Direct measurement of wave velocities can be difficult and/or expensive.
- Proxy quantities are available at dense locations while wave velocities are measured at sparse locations
- Out method gives better metric and more data points per dimension while training thus improved interpolation accuracy

**Inversion on the manifolds:** Manifolds can provide us easy-to-search space where the inference is computationally cheaper

ML algorithms can distinguish channelized and unchannelized sub-surfaces, but they do not account for physical law as a constraint for credible data analysis. Our work will couple the nonlinear manifold deep learning with physics models.

- Physics constrained data analysis enforces the physics law on the ML algorithm to improve the extracted information and provides valuable insight into distinct datasets
- Data integration on an easy-to-search feature space extracted by deep learning algorithms reduces the computational complexity of the simulation-driven data integration
- Seamless data analysis and data integration allow us to identify the relevant features for the quantities of interest in distinct datasets
Machine learning can help us to transform the proxy data into prior knowledge and feature space identification. Adjoint models facilitates higher acceptance rate of MCMC chains.

Snapshots generated from proxy data using multipoint statics

KPCA motivating examples; XOR data (top) and arbitrary non-linear data (bottom)
We solve distinct challenges in inverse problems via manifold learning and constructing adjoint PDEs

<table>
<thead>
<tr>
<th>Challenge</th>
<th>Approach</th>
<th>Explanation</th>
</tr>
</thead>
<tbody>
<tr>
<td>High-fidelity gradient computation</td>
<td>Adjoint gradient</td>
<td>Adjoint PDE allows us to compute gradients in the parameter space with two model runs (a forward and adjoint simulation)</td>
</tr>
<tr>
<td>High dimensionality of the parameters</td>
<td>Manifold learning via KPCA/Machine learning</td>
<td>KPCA is used to find a low-dimensional feature space where the solution is not an outlier in the prior probability space</td>
</tr>
<tr>
<td>Sampling non-Gaussian feature random variables</td>
<td>PCE</td>
<td>PCE is used to sample KPCA feature random variables that are uncorrelated but dependent non-Gaussians</td>
</tr>
<tr>
<td>Ill-posedness of the inverse problem</td>
<td>Bayesian inference</td>
<td>Provides a systematic way to address noisy and measurements and provides a probabilistic inverse solution</td>
</tr>
<tr>
<td>Computational intractability of the MCMC</td>
<td>Langevin MCMC</td>
<td>MCMC require $O(n)$ computational capacity while LMCMC need $O(n^{1/3})$ and inference is done in feature space</td>
</tr>
</tbody>
</table>
Example: Elasticity inversion with KPCA and adjoint models

True (left), KPCA projected (middle) and Posterior sample of parameter space.

Prior posterior and true densities of feature random variables in PCA space and KPCA space

MCMC Chains

Langevin MCMC

Optimization on the manifolds: Manifolds are adopted to the quantity of interest

- In general, the physical system acts as a filter (removes the noise in the input) and Quantity of interest (QoI) often lies in a subspace of the input parameter space.
- We obtain QoI adapted random variables ($\eta$) by rotation of the original random variables ($\xi$).
- Rotation does not affect accuracy of the solution if the random variables are Gaussian.
- Rotation can be found via linear or quadratic adaptation.
- We obtain a manifold (subspace) in the rotated space ($\eta$) and build a surrogate in the subspace.

Example: Oil well placement problem

- **Goal**: To maximize the oil production rate (with some confidence, $\alpha$) by placing injection and production wells at optimal locations.

- Uncertainty in permeability ($\text{dim}=20$) implies we need to solve optimization problem under uncertainty.

- For each design point we have to construct a surrogate model to compute statistics.

- Surrogates are built using a manifold adapted to the QOI via basis adaptation (about 100 model runs for each design point).

- Computationally efficient compared to MCMC, variants of MCMC and traditional Polynomial Chaos Expansion (PCE).

Conclusions and future research

- Manifold learning is used improve high-dimensional
  - Stochastic interpolation: Via providing a better metric and more data points per dimension while training thus improved the interpolation accuracy
  - Stochastic inversion: Via doing inversion on easy-to-search feature space extracted by manifold learning reduces the computational complexity of the simulation-driven data integration
  - Optimization under uncertainty: By building quantity of interest aware manifold thus constructing a surrogate with a few model evolutions

- Open Source software (Oct 2018): Data Assimilation for Stochastic Source Inversion (DASSI)

- Developing a novel machine learning (ML) framework for seamless data analysis and data integration to identify the relevant features for the quantities of interest in distinct datasets and apply physics constraints on ML models
Thank You