Uncertainty Quantification with High-Dimensional Experimentally Measured Inputs

Ilias Bilionis and Nicholas Zabaras

Sibley School of Mechanical and Aerospace Engineering
Cornell University
Ithaca, NY 14853
nzabaras@gmail.com
http://mpdc.mae.cornell.edu/

SIAM Uncertainty Quantification Conference, Raleigh, NC, April 4, 2012
Problem Definition

Deterministic Solver
Reduction
Density Estimation
Reconstruction
Observed input
Bayesian Training
Surrogate
Model

• Tree construction.
• Experimental design.
• Output correlations.
• HDMR terms.

\[ A = \{ \alpha(s) \}_{s=1}^{S_A} \]

Reduced input space
Output space

Data collection

Bayesian Training

Statistics
PDFs
Error bars
Data-driven Stochastic Input Modeling

\[ R(\alpha) = z \]

\[ \alpha = C(z) \]

**Input reduction:**
KPCA, Isomap, LLE, LEM, HEM, GTM, etc.

**Input reconstruction:**
Optimization problem constrained on physical info.

**Density estimation:**
Polynomial Chaos Representation

\[ z_i = \sum_m \beta_{im} H_m(\xi) \]
Use a non-linear map

\[ \Phi : \mathbb{R}^M \rightarrow F \subset \mathbb{R}^M, \quad \alpha \rightarrow \phi(\alpha) = A, \]

to unwrap the observed data

\[ \{ \alpha^{(s)} \}_{s=1}^{S_A} \rightarrow \{ \Phi(\alpha^{(s)}) \}_{s=1}^{S_A} = \{ A^{(s)} \}_{s=1}^{S_A} \]
on \( F \) (feature space). Then, do PCA (K-L) on the feature space \( F \).

Reduction achieved by keeping just a few terms of PCA.

References: [Schlkopf et al. (1998)], [Ma & Zabaras (2011)].
Fortunately...

Only dot products are needed:

\[ \Phi(\alpha^{(s)}) \cdot \Phi(\alpha^{(s')}) := k(\alpha^{(s)}, \alpha^{(s')}) \]

Choices of kernels:

- \( k(\alpha^{(s)}, \alpha^{(s')}) = (\alpha^{(s)} \cdot \alpha^{(s')}) \) (standard PCA).
- Gaussian kernel:

\[
k(\alpha^{(s)}, \alpha^{(s')}) = \exp \left( -\frac{\|\alpha^{(s)} - \alpha^{(s')}\|^2}{2\sigma^2} \right). \tag{1}
\]

In this work, we use:

\[
\sigma^2 = c \frac{1}{S_A} \sum_{s=1}^{S_A} \min_{j \neq i} \|\alpha^{(s)} - \alpha^{(s')}\|^2, \quad s = 1, \ldots, S_A, \tag{2}
\]

where \( c \) is a user defined.
SGeMS (Stanford Geostatistical Modeling Software).
- 0-1 image large scale image of channelized permeability field.
- Cut it in 1000 45x45 pieces to generate training set.
Training set consists of 1000 samples.
Each sample has 2025 dimensions.
- Performing KPCA reduction.
- We keep 30 eigenvalues of feature space.
- 75% of the field energy.
- Residual variance (in feature space) is 0.003.
Problem Definition

Given a reduced input \( z \), reconstruct a field \( \alpha = C(z) \).

Assuming locally linear reduced manifold:

- Find \( L \) observed nearest neighbors \( z^{(s_l)} \).
- Find the corresponding observed high-dimensional inputs \( \alpha^{(s_l)} \).
- Assume the following form for the reconstruction:

\[
C^L(z) = \sum_{l=1}^{L} d^*_l \alpha^{(s_l)}.
\]

Where:

\[
\{d^*_l\}_{l=1}^{L} = \arg \min \sum_{l=1}^{L} d_l \alpha^{(s_l)} - z \|_2.
\]
Example: Reconstruction

- Reconstruct a test sample:
  - Test sample
  - Reconstruction from Kernel PCA with \( k = 30 \)
  - Reconstructed with Linear PCA: \( k = 30 \)

- Reconstruction of several test samples:
Given the reduced observations \( \{ z^{(s)} \}_{s=1}^{S_A} \), deduce their probability density.

- Expand using a GPC representation:
  \[
  z_i(\xi) = \sum_{m} \beta_{im} H_m(\xi), \quad i = 1, \ldots, K_r.
  \]

- Coefficients may be found via:
  - **Maximum likelihood**: [Descelier et al. (2006)], [Stefanou et al. (2009)].
  - **Rosenblatt transformation**: [Rosenblatt (1952)], [Das et al. (2009)].

- Here, we use Rosenblatt assuming **independence**.
Example: Samples from the Input Model

Sampling the 30 dimensional space and reconstructing...
CUT-HDMR with adaptivity: [Ma & Zabaras (2010)]:

\[ g(\xi) = g_0 + \sum_{1 \leq i \leq K_r} g_i(\xi_i) + \sum_{1 \leq i < j \leq K_r} g_{ij}(\xi_i, \xi_j) + \sum_{1 \leq i < j < k \leq K_r} g_{ijk}(\xi_i, \xi_j, \xi_k) + \ldots \]

Each term is a treed GP’s: [Bilionis & Zabaras (2012)].
In [Bilionis & Zabaras (2012)]:

- Local patches of stationary models.
- Outputs conditionally independent GP’s:

\[ g_i(\cdot) | b_i, \sigma_i, r \sim \mathcal{N}(b_T h(\cdot), \sigma_i c(\cdot, \cdot; r)) \],

- Evidence approximation [Bishop (2006)].
- Analytic error bars for statistics.
In [Bilionis & Zabaras (2012)]:

- Split element if:
  \[ \int_{\Xi} \sigma_i^2(\xi) p(\xi) d\xi > \delta. \]

- Perpendicular to:
  \[ k^* = \arg \max_k l^i_k = \arg \max_k \frac{p^i_k}{r^i_k}. \]

- Informative Data Collection
  [Sacks (1989), MacKay (1992)]:
  \[ \xi^* = \arg \max_{\xi \in \Xi} \sigma_i^2(\xi) p(\xi). \]

- Tree structure depends on \( p(\xi) \).
In [Bilionis & Zabaras (2012)]:

- Split element if:
  \[ \int_{\Xi_i} \sigma^2_g(\xi)p(\xi)\,d\xi > \delta. \]

- Perpendicular to:
  \[ k^* = \arg \max_k l^i_k = \arg \max_k \frac{p^i_k}{r^i_k}. \]

- Informative Data Collection
  [Sacks (1989), MacKay (1992)]:
  \[ \xi^* = \arg \max_{\xi \in \Xi} \sigma^2_g(\xi)p(\xi). \]

- Tree structure depends on \( p(\xi). \)
In [Bilionis & Zabaras (2012)]:

- Split element if:
  \[ \int_{\Xi_i} \sigma_g^2(\xi) p(\xi) d\xi > \delta. \]

- Perpendicular to:
  \[ k^* = \arg \max_k l_i^k = \arg \max_k \frac{p_i^k}{r_i^k}. \]

- Informative Data Collection
  [Sacks (1989), MacKay (1992)]:
  \[ \xi^* = \arg \max_{\xi \in \Xi} \sigma_g^2(\xi) p(\xi). \]

- Tree structure depends on \( p(\xi) \).
In [Bilionis & Zabaras (2012)]:

- Split element if:
  \[ \int_{\Xi_i} \sigma_g^2(\xi)p(\xi)\,d\xi > \delta. \]

- Perpendicular to:
  \[ k^* = \arg \max_k \frac{p_k^i}{r_k^i}. \]

- Informative Data Collection
  [Sacks (1989), MacKay (1992)]:
  \[ \xi^* = \arg \max_{\xi \in \Xi} \sigma_g^2(\xi)p(\xi). \]

- Tree structure depends on \( p(\xi) \).
Examples: Two Phase Flow through Porous Medium

Water and oil, ignore gravity effects and capillary forces, assume porosity is a constant.

\[ \nabla \cdot \mathbf{u} = 0 \]
\[ \mathbf{u} = -K(x, \omega) \nabla p, \forall x \in D, \]
\[ \frac{\partial S(x, t, \omega)}{\partial t} + \mathbf{u} \cdot \nabla S(x, t, \omega) = 0, \forall x \in D, t \in [0, T], \]
\[ p = \bar{p}, \text{ on } \partial D_p, \quad \mathbf{u} \cdot \mathbf{n} = 0, \text{ on } \partial D_u. \]

- Deterministic solver: Mixed FEM on a 45x45 grid.
- Permeability is defined as constant on each element.
- We look at the response at the saturation at 0.2 PVI.
- We don’t need HDMR for 30D input.
Figure: Comparison of standard deviations of the saturation $S$ using MGP with approx. 800 (a) and 6,500 (b) samples with a MC simulation using $10^6$ samples at 0.2 PVI.
Select center point $\bar{\xi}$. Calculate:

$$g_0 = g(\xi_0).$$

Build all first order terms:

$$g_i(\xi_i) = g(\xi_i, \bar{\xi}^j) - g_0.$$

Find important dimensions:

$$\frac{\int g_i(\xi_i)p(\xi_i)d\xi_i}{g_0} > \theta.$$

Evaluate only the 2D terms that come from important dimensions:

$$g_{ij}(\xi_i, \xi_j) = g(\xi_i, \xi_j, \bar{\xi}^{ij}) - g_i(\xi_i) - g_j(\xi_j) - g_0$$

etc...
Single phase porous flow:

\[ \nabla \cdot \mathbf{u} = 0 \]
\[ \mathbf{u} = -K(x, \omega) \nabla p, \forall x \in D, \]
\[ p = 1 - x_1 \text{ on } \partial D. \]

- Deterministic solver: Mixed FEM on a 64x64 grid.
- Response observed on a 10x10 coarse grid.
- \( G(x, \omega) = \log K(x, \omega) \) = exponential random field with:

\[
\text{Cov}[G((x_1, y_1), (x_2, y_2)) = \exp \left( - \frac{|x_1 - x_2|}{L} - \frac{|y_1 - y_2|}{L} \right). \]
**Figure:** [L=0.1, K=100]: Mean of the x-velocity. Adaptivity threshold $\theta = 10^{-3}$. HDMR-1 (terms=100, samples=2001). HDMR-2 (terms=2876, samples=57501). For greater $\theta$ no second order terms are added.
Figure: [L=0.1, K=100]: Std. of the x-velocity. Adaptivity threshold \( \theta = 10^{-3} \). HDMR-1 (terms=100, samples=2001). HDMR-2 (terms=2876, samples=57501). For greater \( \theta \) no second order terms are added.
Example: Porous Flow in High Dimensions

Figure: [L=0.01, K=250]: Mean of the x-velocity. Adaptivity threshold $\theta = 10^{-4}$. HDMR-1 (terms=250, samples=5001). HDMR-2 (terms=4811, samples=96201). For greater $\theta$ no second order terms are added.
Figure: [L=0.01, K=250]: Std. of the x-velocity. Adaptivity threshold $\theta = 10^{-4}$. HDMR-1 (terms=250, samples=5001). HDMR-2 (terms=4811, samples=96201). For greater $\theta$ no second order terms are added.