Orbit Uncertainty Propagation With Separated Representations

Marco Balducci
Brandon Jones
Alireza Doostan

*University of Colorado Boulder
SIAM, APRIL 1-5, 2014
Motivation
Motivation

Space is Becoming More Crowded

- 2013: 20,000
- 2025: 100,000s
- Expected improvements in sensor technology
- SSA analyses must remain accurate over increasingly long time spans

Image credit: NASA Orbital Debris Program Office
Background

Balducci, et al. | University of Colorado Boulder

Problem Introduction

Orbit Uncertainty Propagation With SR

4 of 27
Considered Scenario

Observation Gap of 1.5 Days

- 6 Stochastic dimensions at the least $(x, y, z, \dot{x}, \dot{y}, \dot{z})$
- Additional dimensions in gravitational parameter $\mu$, atmospheric drag, gravity perturbations, solar radiation pressure, etc

<table>
<thead>
<tr>
<th>Element</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>$a$</td>
<td>7172.5 km</td>
</tr>
<tr>
<td>$e$</td>
<td>0.00111</td>
</tr>
<tr>
<td>$i$</td>
<td>98.6°</td>
</tr>
<tr>
<td>$\Omega$</td>
<td>253.7°</td>
</tr>
<tr>
<td>$\omega$</td>
<td>77.2°</td>
</tr>
<tr>
<td>$\nu$</td>
<td>59.6°</td>
</tr>
</tbody>
</table>

Table: Keplerian Elements
Traditional Astrodynamic UQ

Established Techniques

• Monte Carlo
• Linearization and the state transition matrix (STM)
• Unscented Transform

These Methods Have Drawbacks

• Convergence rate of MC is slow
• Linearization and STM, as well as unscented transform, rely on Gaussian distribution assumption

Therefore, more robust methods must be considered
Proposed Astrodyanmic UQ

Methods in Development

- Polynomial Chaos Expansions (PC)
- Gaussian Mixtures
- State Transition Tensors (STT)

Without mitigation techniques, PC and Gaussian Mixtures suffer from the curse of dimensionality

- Computation time increases exponentially with respect to input dimensions $d$
- Resulting in increased computation time or dimension truncation

Image credit: Jones, et al. (2013)

STT Must Solve for Multiple Differential Equations
Non-Adapted Non-Intrusive Approximation Methods

Random Sampling - Monte Carlo

- “No” curse of dimensionality
- Very slow convergence
- \( C_d = O(C_1) \)

Random Sampling - Regression

- Least-squares polynomial (chaos) fit
- \( C_d = O(C_1^d) \)

Deterministic Sampling - Stochastic Collocation

- Sparse grids
- \( C_d = O(C_1 (\log C_1)^{d-1}) \)
Separated Representations
Separated Representations

Separated Representation

**Premise:** Decompose a multi-variate function into a linear combination of the products of uni-variate functions

\[ u(y_1, \ldots, y_d) = \sum_{l=1}^{r} s_l u_1^l(y_1) u_2^l(y_2) \cdots u_d^l(y_d) + \mathcal{O}(\epsilon) \]

- \( r \) is the separation rank
- \( u_i^l(y_i) \) are the unknown uni-variate functions/factors
- Computation cost dominated by relatively few MC propagations

**Extensive Background**

- Chemistry, data mining, imaging, etc
Separated Representations

Connections With SVD

Singular-value decomposition (SVD):

\[
U = s_1 u_1^1 + \cdots + s_r u_2^r + s_1 u_1^r
\]

Separated approx: generalization of matrix SVD to tensors:

\[
U \approx \sum_{l=1}^{r} s_l u_1^l \otimes u_2^l \otimes u_3^l
\]

Functions:

\[
u(y_1, y_2, y_3) = s_1 u_1^1(y_1) + \cdots + s_r u_2^r(y_2) u_3^r(y_3)
\]
Problem Set Up: Given $n$ random samples

\[ u_r(y) = \sum_{l=1}^{r} s_l u_1^l(y_1) u_2^l(y_2) \cdots u_d^l(y_d) \text{ s.t. } \|u - u_r\|_D \leq \epsilon \]

where,

\[ \|u\|_D^2 := \frac{1}{N} \sum_{i=1}^{N} u(y_i)^2 \]
A Non-Intrusive Implementation

Spectral Decomposition of Factors

\[ u_i^l (y_i) = \sum_{j=1}^{M} c_{i,j}^l \psi_j (y_i) \]

where,

\[ \int_{\Gamma_i} \psi_j (y_i) \psi_k (y_i) P_{y_i} (y_i) dy_i = \delta_{jk} \quad \forall \ i, j, k \]

Discrete Approximation

\[ \{ c_{i,j}^l \} = \arg \min_{\{ \hat{c}_{i,j}^l \}} \left\| u (\cdot) - \sum_{l=1}^{r} s_l u_1^l (\cdot) u_2^l (\cdot) \cdots u_d^l (\cdot) \right\|_D \]

with \( \epsilon \) as the stopping criterion.
Separated Representations

Optimization With Alternating Least Squares (ALS)

Initialize $r, u_1^l(y_1), u_2^l(y_2), u_3^l(y_3)$

for $k = 1, \ldots, \text{until convergence}$:

$$\{c_{1,j}^l\} = \arg \min_{\{\hat{c}_{1,j}^l\}} \left\| u(\cdot) - \sum_{l=1}^{r} s_l u_1^l(\cdot) u_2^l(\cdot) u_3^l(\cdot) \right\|_D$$

$$\{c_{2,j}^l\} = \arg \min_{\{\hat{c}_{2,j}^l\}} \left\| u(\cdot) - \sum_{l=1}^{r} s_l u_1^l(\cdot) u_2^l(\cdot) u_3^l(\cdot) \right\|_D$$

$$\{c_{3,j}^l\} = \arg \min_{\{\hat{c}_{3,j}^l\}} \left\| u(\cdot) - \sum_{l=1}^{r} s_l u_1^l(\cdot) u_2^l(\cdot) u_3^l(\cdot) \right\|_D$$

end

For $d = 3$:

$r = r + 1$

no

$\|u - u_r\|_D \leq \varepsilon$

yes

end
Computation Cost

**Linear Scalability: Required Number of Solution Samples**

\[
 u_r(y) = \sum_{l=1}^{r} s_l u_1^l(y_1) u_2^l(y_2) \cdots u_d^l(y_d)
\]

Number of unknowns = \( r \cdot d \cdot M \)

\[ N \sim \mathcal{O}(r \cdot d \cdot M) \]

**Total Computation Time is Quadratic With Respect to \( d \) When ALS is Applied**

\[ C_d \sim \mathcal{O}(K \cdot r^2 \cdot d \cdot M^2 (N + S)) \]

This cost should be small when compared to the number of required MC propagations
Analyses
Distribution Characteristics

- $1\sigma^2$ uncertainty of $1 \text{ km}^2$ in position and $1 \text{ (m/sec)}^2$ in velocity
- Propagated for 36 hr, Dormand-Prince (5)4 Integrator
- $r = 3$
- Hermite polynomials up to 2nd degree
- Estimate Cartesian coordinates
- Results compared to 100,000 MC samples
SR Results \((d = 6)\)

6 random inputs required 250 samples
## SR Results ($d = 6$)

<table>
<thead>
<tr>
<th>Moment</th>
<th>Max</th>
<th>Min</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>8.899e-05</td>
<td>4.9114e-06</td>
</tr>
<tr>
<td>STD</td>
<td>0.00082411</td>
<td>0.00030213</td>
</tr>
<tr>
<td>Skewness</td>
<td>0.030836</td>
<td>0.0056492</td>
</tr>
<tr>
<td>Kurtosis</td>
<td>0.033773</td>
<td>0.0076511</td>
</tr>
</tbody>
</table>

**Table:** Relative Residuals
SR Results \((d = 20)\)

20 random inputs required 900 samples

Additional stochastic dimensions included \(\mu, C_d, A/M\) and gravitational perturbations
SR Results \((d = 20)\)

<table>
<thead>
<tr>
<th>Moment</th>
<th>Max</th>
<th>Min</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>2.2873e-05</td>
<td>4.9506e-07</td>
</tr>
<tr>
<td>STD</td>
<td>0.0012369</td>
<td>0.0311e-05</td>
</tr>
<tr>
<td>Skewness</td>
<td>0.033496</td>
<td>0.0038675</td>
</tr>
<tr>
<td>Kurtosis</td>
<td>0.025802</td>
<td>0.0067912</td>
</tr>
</tbody>
</table>

Table: Relative Residuals
SR vs MC

6 Stochastic Dimensions

![Graph showing STD vs Number of Training Samples for MC and SR methods.](image)
Required Samples (Empirical)
Wrap Up
Considerations

Low-Rank Separated Representation?
• Sparsity in a tensor-product basis

The Approximation is Ill-Posed
• Not always unique
• May over-fit the samples if not regularized
• Tikhonov regularization

\[
\left\{ \hat{c}_{i,j}^l \right\} = \arg \min \left\{ c_{i,j}^l \right\} \left\| u(\cdot) - \sum_{l=1}^{r} s_l u_1^l(\cdot) u_2^l(\cdot) u_3^l(\cdot) \right\|_{D}^2 + \lambda \sum_{l,j} \left( c_{1,j}^l \right)^2
\]

ALS Converges to a Local, Not Necessarily Global, Minimizer
Future Work

Prevent Over Fitting

Cross Validation: Total Cost?

Various Realistic Scenarios with Varying “Difficulty”

Better Estimation Convergence
Wrap Up

Conclusions

Sparse observations lead to non-Gaussian distributions
Conclusions

Sparse observations lead to non-Gaussian distributions

Orbit determination can involve large numbers of stochastic dimensions
Conclusions

Sparse observations lead to non-Gaussian distributions

Orbit determination can involve large numbers of stochastic dimensions

Low rank separated representations use non-intrusive polynomial surrogates
Conclusions

Sparse observations lead to non-Gaussian distributions

Orbit determination can involve large numbers of stochastic dimensions

Low rank separated representations use non-intrusive polynomial surrogates

SR does not suffer from the curse of dimensionality
  • Cost increases quadratically with respect to stochastic dimension count
Conclusions

Sparse observations lead to non-Gaussian distributions

Orbit determination can involve large numbers of stochastic dimensions

Low rank separated representations use non-intrusive polynomial surrogates

SR does not suffer from the curse of dimensionality
  • Cost increases quadratically with respect to stochastic dimension count

Questions?