Estimation of Information-Theoretic Quantities for Particle Clouds

Shankar Kulumani, Islam I. Hussein, Christopher W. T. Roscoe, Matthew P. Wilkins, and Paul W. Schumacher, Jr.

George Washington University, Washington, DC
Applied Defense Solutions, Columbia, MD
Air Force Research Laboratory, Kihei, HI

September 13-16, 2016

Motivation

- Particle systems represent uncertainty with higher fidelity, especially for long term propagation
- Information theoretic quantities such as entropy and divergence require integration of analytical functions
- Given a particle cloud, how do we estimate information theoretic properties of the underlying unknown probability density function (PDF)?
    - Need to estimate the underlying PDF

Applications

- Entropy: Computing entropy of nonlinear systems for non-linearity detection (DeMars et al., 2013)
- Divergence: Solving track-to-track association for UCT correlation problems (Hussein et al., 2015 and 2016)
- Mutual information: Solving observation-to-observation data association/IOD (Hussein et al., 2014, 2015)

PDF Estimation Approaches

- Histogram: Construct a histogram from the particle system
  - Drawbacks: non-smooth, PDF estimate is highly sensitive to the bin parameters, initractable for higher dimensional systems
- Kernel Methods: Use positive basis functions that integrate to 1 to fit the data
  - Use k-NN algorithm to estimate the PDF
  - Use the EM algorithm to estimate the PDF
  - Implemented the k-NN algorithm in this paper
  - Compare k-NN performance against EM in future work

Information Quantities

- Rényi entropy:

$$
\begin{aligned}
H_\alpha(p) &= \frac{1}{1-\alpha}\log\int p^\alpha(\boldsymbol{x})\mathrm{d}\boldsymbol{x} \\
&= \frac{1}{1-\alpha}\log\left(E[p^{\alpha-1}(\boldsymbol{x})]\right) \\
&\simeq \frac{1}{1-\alpha}\log\left(\frac{1}{N}\sum_{i=1}^{N} p^{\alpha-1}(\boldsymbol{x}_i)\right)
\end{aligned}
$$

- Shannon information entropy

$$
\begin{aligned}
H(p) &= -\int p(\boldsymbol{x})\log(p(\boldsymbol{x}))\mathrm{d}\boldsymbol{x} \\
&= -E[\log(p(\boldsymbol{x}))] \\
&\simeq -\frac{1}{N}\sum_{i=1}^{N}\log(p(\boldsymbol{x}_i))
\end{aligned}
$$

Information Quantities

- Bhattacharyya divergence:

$$\begin{aligned}
D_{\mathrm{B}}(p_1||p_2) &= -\log \mathrm{B}_C(p_1, p_2) = -\log\left[\int \sqrt{p_1(\boldsymbol{x})p_2(\boldsymbol{x})}\mathrm{d}\boldsymbol{x}\right] \\
&= -\log\left[\int p_1(\boldsymbol{x})\sqrt{\frac{p_2(\boldsymbol{x})}{p_1(\boldsymbol{x})}}\mathrm{d}\boldsymbol{x}\right] \\
&= -\log\left[E_{p_1}\left[\sqrt{\frac{p_2(\boldsymbol{x})}{p_1(\boldsymbol{x})}}\right]\right] \\
&\simeq -\log\left[\frac{1}{N}\sum_{i=1}^{N}\sqrt{\frac{p_2(\boldsymbol{x}_i)}{p_1(\boldsymbol{x}_i)}}\right], \ \boldsymbol{x}_i \sim p_1 \\
&\simeq -\log\left[\frac{1}{N}\sum_{i=1}^{N}\sqrt{\frac{p_1(\boldsymbol{x}_i)}{p_2(\boldsymbol{x}_i)}}\right], \ \boldsymbol{x}_i \sim p_2
\end{aligned}$$

The k-NN Algorithm

- A non-parametric method that relies on locally approximating the PDF based on particle concentration in a neighborhood

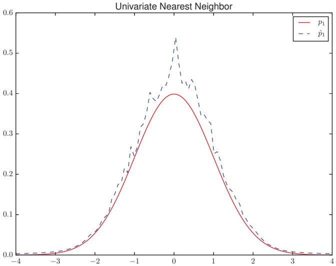
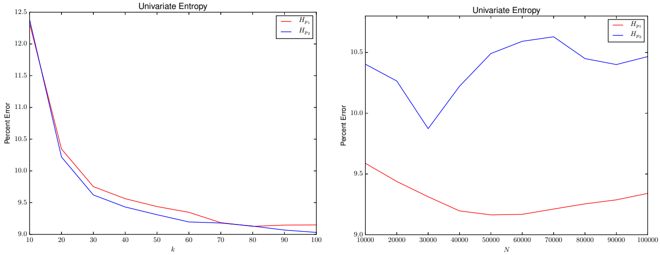


Figure: Univariate nearest neighbor estimate, $N = 20,000$, $k = 200$

The k-NN Algorithm

- Sensitivity to the parameters k and N (1-d)

$$p_1(x) = \mathcal{N}\{1, 1\}$$
$$p_2(y) = \mathcal{N}\{-1, 1\}$$

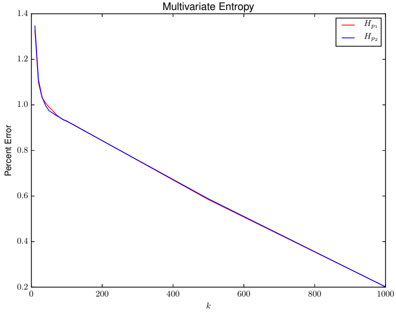


(a) Sensitivity to kth nearest neighbor. (b) Sensitivity to sample size N.

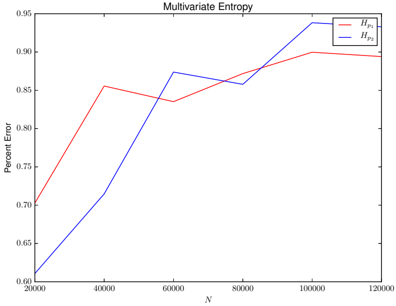
Figure: Univariate Entropy Estimation.

The k-NN Algorithm

- Sensitivity to the parameters k and N (6-d)



(a) Sensitivity to kth nearest neighbor ($N = 50,000$).

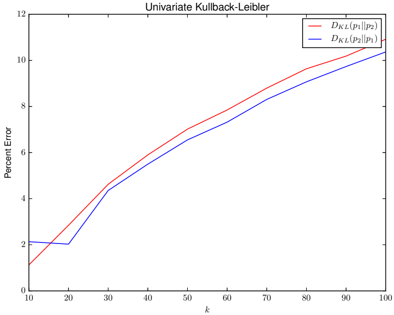


(b) Sensitivity to sample size N ($k = 200$).

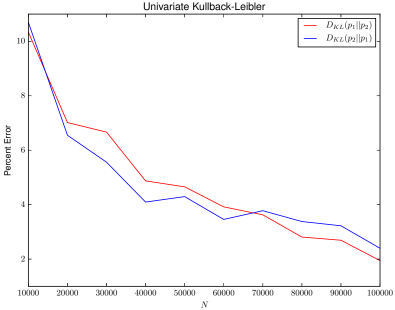
Figure: Multivariate Entropy Estimation.

The k-NN Algorithm

- Sensitivity to the parameters k and N (1-d)



(a) Sensitivity to kth nearest neighbor ($N = 20,000$).

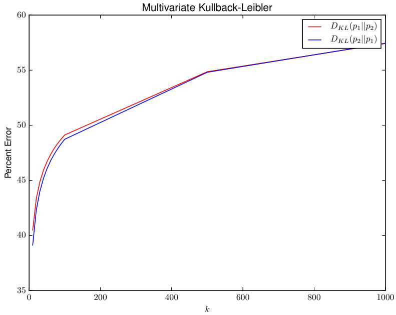


(b) Sensitivity to sample size N ($k = 50$).

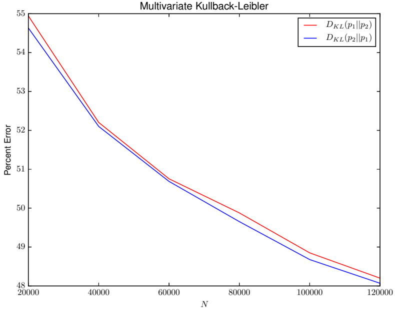
Figure: Univariate Kullback-Leibler Estimation.

The k-NN Algorithm

- Sensitivity to the parameters k and N (6-d)



(a) Sensitivity to kth nearest neighbor ($N = 50,000$).

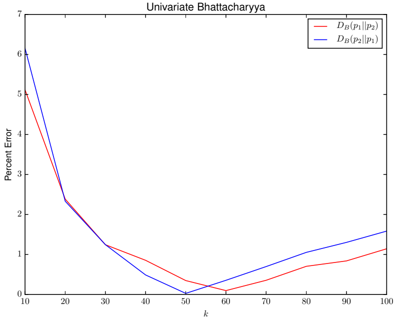


(b) Sensitivity to sample size N ($k = 200$).

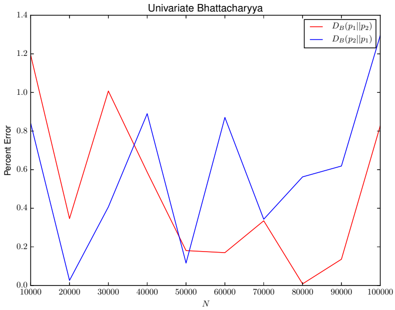
Figure: Multivariate Kullback-Leibler Estimation.

The k-NN Algorithm

- Sensitivity to the parameters k and N (1-d)



(a) Sensitivity to kth nearest neighbor ($N = 20,000$).

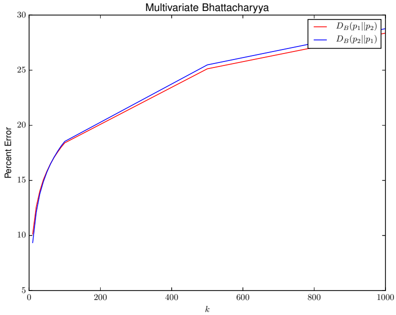


(b) Sensitivity to sample size N ($k = 50$).

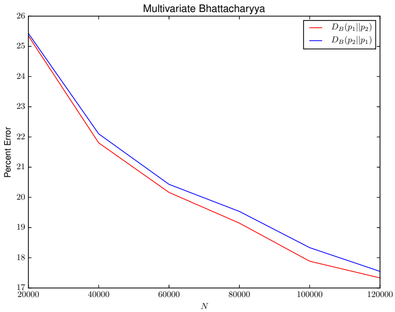
Figure: Univariate Bhattacharyya Estimation.

The k-NN Algorithm

- Sensitivity to the parameters k and N (6-d)



(a) Sensitivity to kth nearest neighbor ($N = 50,000$).

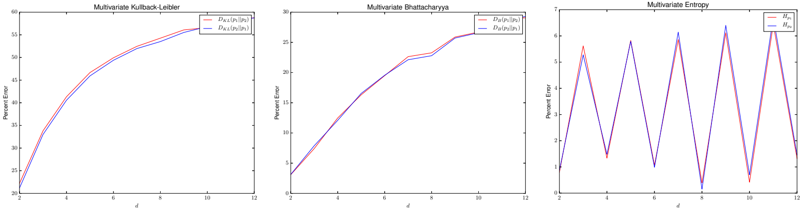


(b) Sensitivity to sample size N ($k = 200$).

Figure: Multivariate Bhattacharyya Estimation.

The k-NN Algorithm

- Sensitivity to the parameters k and N ($k = 50$, $N = 50,000$)



(a) Kullback-Leibler divergence estimation.

(b) Bhattacharyya divergence estimation.

(c) Entropy estimation.

Figure: Sensitivity of estimation to varying dimensions.

Entropy Conservation in Hamiltonian Systems

- Entropy is conserved for a Hamiltonian system when canonical coordinates are used to parametrize the system
- Object in a zero eccentricity, inclined (97.9 degrees) LEO orbit
- Initial uncertainty is Gaussian with 10 m position and 0.1 m/s velocity standard deviation (uniform in all directions)
- 10,000 particles were generated
- Transform the particles' to Delaunay orbital elements, and use to approximate the underlying PDF
- Initial entropy was found to be 44.90855570500575
- After a 30 day orbit propagation, the final entropy was estimated to be 44.90800388940583 (5.5×10^{-2}% error)

Bhattacharyya Divergence for UCT Correlation

- Two objects with same mean orbital elements:
  - Position standard deviation: 5 km to 200 km
  - Velocity standard deviation: 0.5 m/s
- Duration between t_* and t^*: 30 sidereal days
- 1000 MC runs for each position separation (1000 particles)

Table: Parameters of the True Orbit

Parameter	**Value**
Semimajor Axis, km	26 600.0
Eccentricity	0.2
Inclination, deg	55.0
Argument of Perigee, deg	-120.0
Right Ascension of the Ascending Node, deg	207.0
True Anomaly, deg	20.0

Bhattacharyya Divergence for UCT Correlation

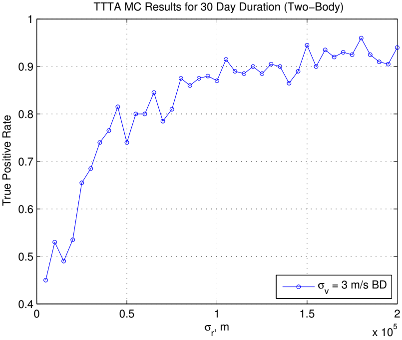


Figure: UCT Correlation results over a 30 day period.

Conclusion

- Introduced the k-NN algorithm for density estimation
- Used k-NN algorithm for information theoretic quantity estimation for particle systems
  - Entropy
  - Divergence
- Applied k-NN to demonstrate conservation of entropy for Hamiltonian systems and for solving UCT correlation problems
- Currently:
  - Implementing EM algorithm for PDF estimation
  - Comparing EM algorithm against k-NN algorithm performance
  - Applying approaches to estimating mutual information (for observation-to-observation association/IOD)