1. Background Information

Purpose

- The purpose of my study is to use various machine learning methods to facilitate the generation of summary statistics in Approximate Bayesian Computation (ABC). Specifically, I will use the following methods
  - Moment estimation and linear regression
  - Radius basis function
  - Neural network

What is an ABC algorithm?

- In Bayesian Estimation, one of the most important steps is to calculate the likelihood function

\[ p(X | \theta = \theta) \]

- However, in many important applications, this is infeasible. ABC algorithm can get posterior estimation without knowing the likelihood function.

\[ \theta \rightarrow \text{ABC algorithm} \rightarrow \text{posterior estimation} \]

- So the central question is to find a good summary statistic using machine learning methods to reduce the dimension.

The following theorem guarantees the validity.

**Theorem 1.** Assume that \( E[\theta] < \infty \). ABC procedure with observed data \( X \), summary statistic \( S \), any finite vector space norm \( \| \cdot \| \), and tolerance \( \varepsilon \) produces a posterior distribution \( p(\theta | X) \). Then we have

\[ \| S(X') - S(X) \| < \varepsilon \]

and

\[ \lim_{n \to \infty} \frac{1}{n} \sum_{i=1}^{n} \| \theta_i - \theta \| = 0 \]

2. Method

- Generate a data set \( D = \{ (\theta_i, X_i) \} \), where \( X_i \) is a random sample from the distribution \( p(X | \theta = \theta_0) \).
- Use the data set \( D \) to train the supervised learning model \( \ell \), to minimize the objective function

\[ \frac{1}{n} \sum_{i=1}^{n} \| \theta_i - \hat{S}(X_i) \| \]

This objective function can be viewed as an approximation of the posterior mean \( E[\ell(X) | X] \) because the solution to an squared error minimization problem will result in such an estimation.
- Run the ABC algorithm with the summary statistic \( S \) to get a posterior distribution. Then check the calculated posterior distribution with the posterior distribution derived from the sufficient statistics to judge whether the machine learning method \( \ell \) performs well.

3. Evaluate with Moving Average Model

- Assuming the underlying model

\[ X_i = \varepsilon_1 + \theta \varepsilon_{i-1} + \gamma \varepsilon_{i-2} \]

- And underlying parameter \( (\theta, \gamma) = (1, -1) \)

- A typical time series with the above parameters

4. Evaluate with Hidden MC model

- The hidden Markov Chain model is as follows

- Automaic summary statistic generation is a good way to do ABC algorithm when we don’t know the exact sufficient statistics.

- More complex models tend to generate more scattered posterior distribution, but the mean and standard deviation seem better.

- In practice, the efficiency of above methods are moment method > linear regression > NN > RBF

5 Conclusions and Insights

- All three methods tend to have similar results in this case, while complex methods have better predictions.

6. Reference