Bayesian Approach to Time Delay Estimation

Hyungsuk Tak

Stat300

23 Sep 2014

Joint work with Xiao-Li Meng, David van Dyk, Kaisey Mandel, Aneta Siemiginowska, and Vinay Kashyap
Introduction

Data and Difficulties

Bayesian Approach
  ▶ Motivation
  ▶ State-space Modeling
  ▶ Likelihood
  ▶ Prior
  ▶ Hyper-prior
  ▶ Full Posterior and Sampler based on ASIS

Example 1: Simulated data

Example 2: Real Data (Q0957+561)

Discussion with Reference
INTRODUCTION
**Introduction**

Light rays are bent by a strong gravitational field of a lensed galaxy.
- Each route has **different length**.
- **Different arrival times** of light rays

**Why time delay?**
- **Mass structure** of the lens galaxy
- **Cosmological parameters**, e.g., Hubble constant, $H_0$
Data and Difficulties

Data comprise of two time series with measurement errors.

- Observation times \( t' \equiv \{ t_1, t_2, \ldots, t_n \} \)
- Observed magnitudes \( x(t)' \equiv \{ x(t_1), x(t_2), \ldots, x(t_n) \} \), and \( y(t) \)
- Measurement errors (se) \( \delta(t)' \equiv \{ \delta(t_1), \delta(t_2), \ldots, \delta(t_n) \} \) and \( \eta(t) \)
Some difficulties occur in estimating the time delay

1. **Irregular** observation times (∵ weather conditions)
2. **Seasonal gaps** (∵ rotation of the earth)
3. **Magnitude shift** (∵ different gravitational potentials)
4. **Measurement errors**

Our job is to estimate time delay (shift in x-axis) between two time series.
Bayesian Approach: Motivation

Grid optimization methods are dominating this field!
- eg. Cross-correlation method
  1. Shift one light curve by \( \Delta \) in \( x \)-axis
  2. Calculate \( r_\Delta \), sample cross-correlation function
  3. Find \( \Delta \) that maximizes \( r_\Delta \) on the grid of \( \Delta \)
- \( SE(\hat{\Delta}) \) by computationally expensive repeated sampling procedure
- Grid of \( \Delta \) (\( \neq \) the whole space of \( \Delta \))

Non-grid-based Bayesian approach
- Principled way of model construction: likelihood-based
- Computational efficiency: simple and fast posterior sampling scheme
- The whole space of \( \Delta \)
Bayesian Approach: State-space modeling

Assumption 1: ∃ unobserved underlying processes representing the true magnitudes in continuous time (red and blue dashed curves)

$\mathbf{X}(t)' = (X(t_1), X(t_2), \ldots, X(t_n))$ and $\mathbf{Y}(t)$, values on curves at $t$

Assumption 2: $\mathbf{Y}(t) = \mathbf{X}(t - \Delta) + c$ (Harva, 2006)
Bayesian Approach: Likelihood

Independent Gaussian measurement errors

- $x(t_j) | X(t_j) \sim \mathcal{N}[X(t_j), \delta^2(t_j)]$
- $y(t_j) | Y(t_j) \sim \mathcal{N}[Y(t_j), \eta^2(t_j)]$
- $y(t_j) | X(t_j - \Delta) + c, \Delta, c \sim \mathcal{N}[X(t_j - \Delta) + c, \eta^2(t_j)]$.

Likelihood function

- Suppose $t^* = \text{sort}(t_1, t_2, \ldots, t_n, t_1 - \Delta, t_2 - \Delta, \ldots, t_n - \Delta)$
- $L(X(t^*), \Delta, c) \propto \prod_{j=1}^n p(x(t_j) | X(t_j)) \cdot p(y(t_j) - c | X(t_j - \Delta), \Delta, c)$
Bayesian Approach: Prior

- Ornstein-Uhlenbeck process for $X(\cdot)$ (Kelly et al., 2009)
  - Intrinsic variability of quasar → stochastic process in continuous time
  - Easy way to sample true values at irregularly-spaced times
  - $dX(t) = -\frac{1}{\tau} (X(t) - \mu) dt + \sigma dB(t)$
  - Markovian property
    - $X(t^*_j) | X(t^*_{j-1}), \mu, \sigma, \tau \sim \mathcal{N} \left[ \text{mean: } \mu + e^{-(t^*_j - t^*_{j-1})/\tau} (X(t^*_{j-1}) - \mu), \variance: \frac{\tau \sigma^2}{2} (1 - e^{-2(t^*_j - t^*_{j-1})/\tau}) \right]$

- $p(X(t^*) | \mu, \sigma, \tau, \Delta) =$
  - $p(X(t^*_1) | \mu, \sigma, \tau, \Delta) \cdot \prod_{j=2}^{2n} p(X(t^*_j) | X(t^*_{j-1}), \mu, \sigma, \tau, \Delta)$

- $p(\Delta, c) = p(\Delta)p(c) \propto I_{\{ |\Delta| \in [0, (t_n - t_1]) \}}$
Bayesian Approach: Hyper-prior

- $\mu$ is a mean parameter of the underlying process
- $\sigma$ is a scale parameter of the underlying process
- $\tau$ is a relaxation time of the underlying process
- Naively informative hyper-prior distribution:
  \[
p(\mu, \sigma^2, \tau) = p(\mu)p(\sigma^2)p(\tau) \propto \frac{e^{-0.01/\sigma^2}}{(\sigma^2)^{1.01}} \frac{e^{-1/\tau}}{\tau^2}
\]
- Flat on $\mu$, InvGamma(0.01, 0.01) on $\sigma^2$, and InvGamma(1, 1) on $\tau$
Suppose $\theta_{hyp} \equiv (\mu, \sigma, \tau)$ and $D_{obs} \equiv (x(t), y(t))$

- **Full Posterior:** $p(X(t^*), \Delta, c, \theta|D_{obs})$
  \[ \propto L(X(t^*), \Delta, c) \cdot p(X(t^*), \Delta, c|\theta_{hyp}) \cdot p(\theta_{hyp}) \]

  - Likelihood  
  - Prior  
  - Hyper-prior

- **Metropolis-Hastings within Gibbs**
  - $p(X(t^*), \Delta|D_{obs}, c, \theta_{hyp})$
  - $p(c|D_{obs}, X(t^*), \Delta, \theta_{hyp})$
  - $p(\theta_{hyp}|D_{obs}, c, X(t^*), \Delta)$

- **Ancillarity-Sufficiency Interweaving Strategy (Yu and Meng, 2011)**
  - $p(X(t^*), \Delta|D_{obs}, c, \theta_{hyp})$
  - Interweaving $p(c|D_{obs}, X(t^*), \Delta, \theta_{hyp})$ with $p(c|D_{obs}, S(t^*), \Delta, \theta_{hyp})$
  - $p(\theta_{hyp}|D_{obs}, c, X(t^*), \Delta)$
Example 1: Simulated Data

Summary of posterior $\Delta_{AB}$ with the blinded truth 30.98

<table>
<thead>
<tr>
<th></th>
<th>Post. Mean</th>
<th>Post. Median</th>
<th>Post. SD</th>
<th>Half-length of 68% PI</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\Delta_{AB}$</td>
<td>30.94</td>
<td>30.89</td>
<td>0.694</td>
<td>0.423</td>
</tr>
</tbody>
</table>
**Example 2: Real Data (Q0957+561)**

Data observed at the United States Naval Observatory (Hainline et al., 2012)

![Graph showing magnitude vs. heliocentric Julian day](image)

<table>
<thead>
<tr>
<th>Researchers</th>
<th>Method</th>
<th>$\hat{\Delta}_{AB}$</th>
<th>$SE(\hat{\Delta}_{AB})$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Oscoz et al. (1997)</td>
<td>Discrete cross-correlation &amp; Dispersion</td>
<td>424</td>
<td>3</td>
</tr>
<tr>
<td>Serra-Ricart et al. (1999)</td>
<td>Cross-correlation functions</td>
<td>425</td>
<td>4</td>
</tr>
<tr>
<td>Tak et al. (?)</td>
<td>Bayesian</td>
<td>423.16</td>
<td>1.22</td>
</tr>
</tbody>
</table>
Example 2: Real Data (Q0957+561) (cont.)

Convergence Checks
Each row: Traceplot, ACF, and histogram from the top
Each column: $\Delta, c, \mu, \log(\sigma), \log(\tau)$ from the left

MHwG

ASIS
Prior on $\Delta$

Quadruply-lensed quasar data

Microlensing

Reference


