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

Outline

- 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

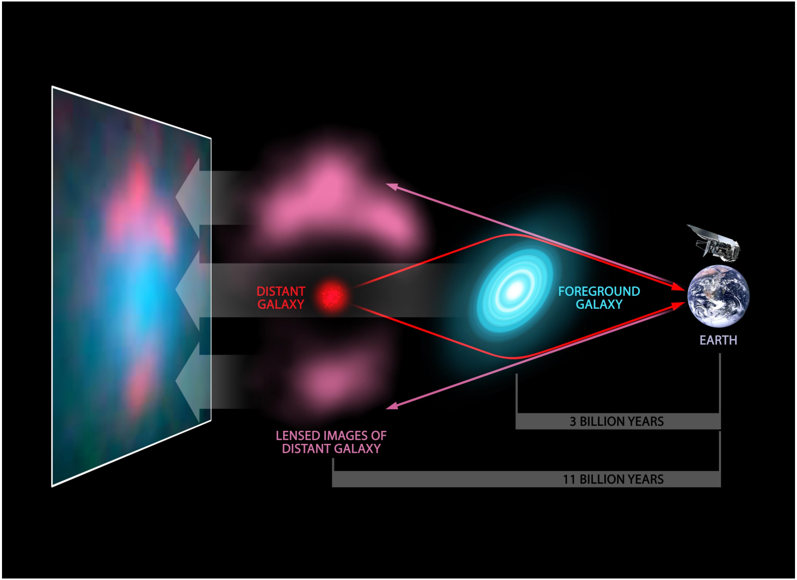
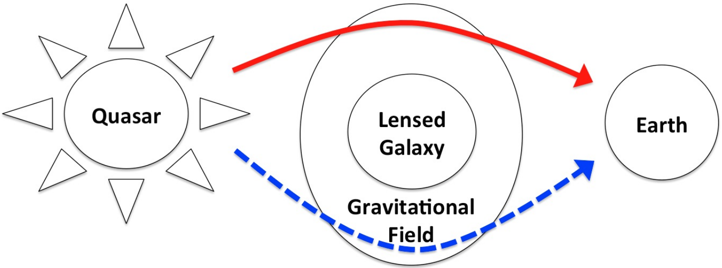


Image Credit: NASA/JPL-Caltech

Introduction



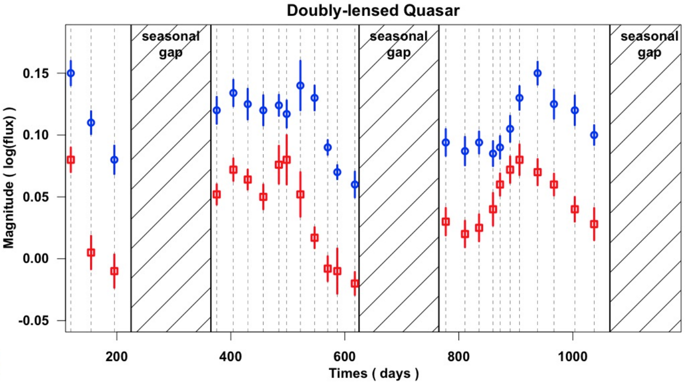
Light rays are bent by a strong gravitational field of a lensed galaxy.

- 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

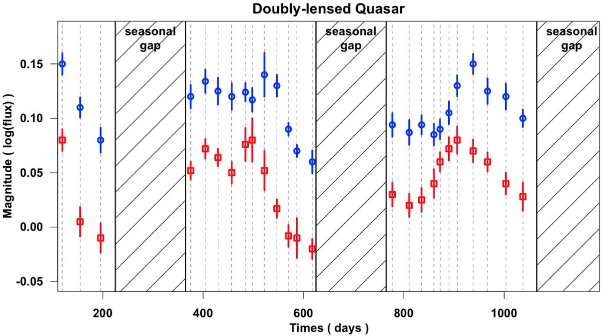
Data and Difficulties



Data comprise of two time series with measurement errors.

- Observation times $\mathbf{t}' \equiv \{t_1, t_2, \ldots, t_n\}$
- Observed magnitudes $\mathbf{x(t)}' \equiv \{x(t_1), x(t_2), \ldots, x(t_n)\}$, and $\mathbf{y(t)}$
- Measurement errors (se) $\boldsymbol{\delta}\mathbf{(t)}' \equiv \{\delta(t_1), \delta(t_2), \ldots, \delta(t_n)\}$ and $\boldsymbol{\eta}\mathbf{(t)}$

Data and Difficulties



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

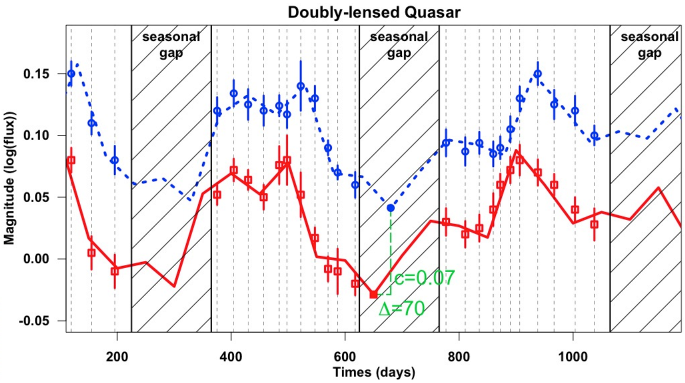
Grid optimization methods are dominating this field!

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

Non-grid-based Bayesian approach

- Principled way of model construction: likelihood-based
- Computational efficiency: simple and fast posterior sampling scheme
- The whole space of Δ

Bayesian Approach: State-space modeling



- Assumption 1: $\exists$ 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 $\mathbf{t}$
- Assumption 2: $\mathbf{Y(t)} = \mathbf{X}(\mathbf{t} - \Delta) + c$ (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 $\mathbf{t}^* = sort(t_1, t_2, \ldots, t_n, t_1 - \Delta, t_2 - \Delta, \ldots, t_n - \Delta)$
- $L(\mathbf{X(t^*)}, \Delta, c) \propto \prod_{j=1}^{n} p(x(t_j)|X(t_j)) \cdot p(y(t_j) - c|X(t_j - \Delta), \Delta, c)$

Bayesian Approach: Prior

- Ornstein-Uhlenbeck process for $\mathbf{X}(\cdot)$ (Kelly et al., 2009)
  - Intrinsic variability of quasar $\rightarrow$ stochastic process in continuous time
  - Easy way to sample true values at irregularly-spaced times
  - $dX(t) = -\frac{1}{\tau}\big(X(t) - \mu\big)dt + \sigma dB(t)$
  - Markovian property
    $X(t_j^*)|X(t_{j-1}^*), \mu, \sigma, \tau \sim \mathcal{N}\big[\text{mean: } \mu + e^{-(t_j^* - t_{j-1}^*)/\tau}\big(X(t_{j-1}^*) - \mu\big),$
    $\text{variance: } \frac{\tau\sigma^2}{2}(1 - e^{-2(t_j^* - t_{j-1}^*)/\tau})\big]$
  - $p(\mathbf{X}(\mathbf{t}^*)|\mu, \sigma, \tau, \Delta) =$
    $p(X(t_1^*)|\mu, \sigma, \tau, \Delta) \cdot \prod_{j=2}^{2n} p(X(t_j^*)|X(t_{j-1}^*), \mu, \sigma, \tau, \Delta)$
- $p(\Delta, c) = p(\Delta)p(c) \propto I_{\{|\Delta| \,\in\, [0,\ (t_n - t_1)]\}}$

Bayesian Approach: Hyper-prior

- μ is a mean parameter of the underlying process
- σ is a scale parameter of the underlying process
- τ 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 μ, InvGamma(0.01, 0.01) on σ^2, and InvGamma(1, 1) on τ

Full Posterior and Sampler based on ASIS

Suppose $\theta_{hyp} \equiv (\mu, \sigma, \tau)$ and $D_{obs} \equiv (\mathbf{x(t)}, \mathbf{y(t)})$

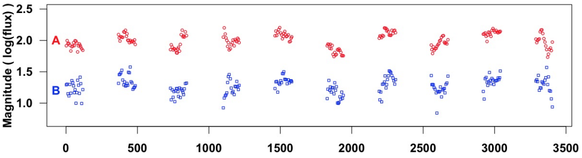
- Full Posterior: $p(\mathbf{X(t^*)}, \Delta, c, \theta | D_{obs})$

$$\propto L(\mathbf{X(t^*)}, \Delta, c) \cdot p(\mathbf{X(t^*)}, \Delta, c | \theta_{hyp}) \cdot p(\theta_{hyp})$$

Likelihood Prior Hyper-prior

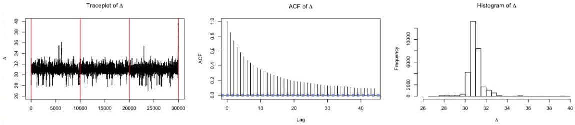
- Metropolis-Hastings within Gibbs
  - $p(\mathbf{X(t^*)}, \Delta | D_{obs}, c, \theta_{hyp})$
  - $p(c | D_{obs}, \mathbf{X(t^*)}, \Delta, \theta_{hyp})$
  - $p(\theta_{hyp} | D_{obs}, c, \mathbf{X(t^*)}, \Delta)$
- Ancillarity-Sufficiency Interweaving Strategy (Yu and Meng, 2011)
  - $p(\mathbf{X(t^*)}, \Delta | D_{obs}, c, \theta_{hyp})$!
  - Interweaving $p(c | D_{obs}, \mathbf{X(t^*)}, \Delta, \theta_{hyp})$ with $p(c | D_{obs}, \mathbf{S(t^*)}, \Delta, \theta_{hyp})$
  - $p(\theta_{hyp} | D_{obs}, c, \mathbf{X(t^*)}, \Delta)$

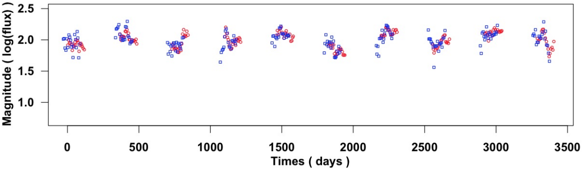
Example 1: Simulated Data



Summary of posterior Δ_{AB} with the blinded truth 30.98

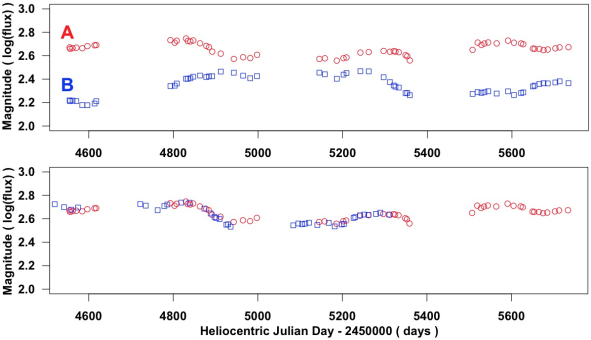
Post. Mean	Post. Median	Post. SD	Half-length of 68% PI
30.94	30.89	0.694	0.423





Example 2: Real Data (Q0957+561)

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



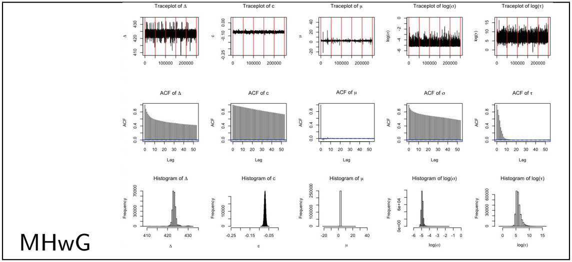
Researchers	Method	$\hat{\Delta}_{AB}$	$SE(\hat{\Delta}_{AB})$
Oscoz et al. (1997)	Discrete cross-correlation & Dispersion	424	3
Serra-Ricart et al. (1999)	Cross-correlation functions	425	4
Tak et al. (?)	Bayesian	423.16	1.22

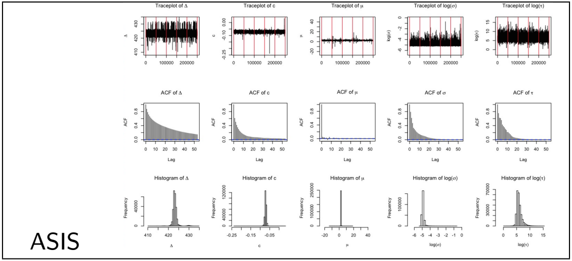
Example 2: Real Data (Q0957+561) (cont.)

Convergence Checks

Each row: Traceplot, ACF, and histogram from the top

Each column: $\Delta, c, \mu, \log(\sigma), \log(\tau)$ from the left





Discussion with Reference

- Prior on Δ
- Quadruply-lensed quasar data
- Microlensing
- Reference

1. L. Hainline, C. Morgan, J. Beach, C. Kochanek, H. Harris, T. Tilleman, R. Fadely, E. Falco, and T. Le (2012) "A new microlensing event in the doubly imaged quasar Q 0957+561" *The Astrophysical Journal*, **744**:104(9pp)
2. M. Harva and S. Raychaudhury (2006) "Bayesian estimation of time delays between unevenly sampled signals" *IEEE Machine Learning for Signal Processing*, ISSN 1551-2541
3. B. Kelly, J. Bechtold, and A. Siemiginowska (2009) "Are the variations in quasar optical flux driven by thermal fluctuation?" *The Astrophysical Journal*, **698**, 895 - 910.
4. A. Oscoz, E. Mediavilla, L. Goicoechea, M. Serra-Ricart, and J. Buitrago (1997) "Time delay of QSO 0957+561 and cosmological implications" *Astronomy and Astrophysics*, **479**, L89-L92
5. M. Serra-Ricart, A. Oscoz, T. Sanchis, E. Mediavilla, L. Goicoechea, J. Licandro, D. Alcalde, and R. Gil-Merino (1999) "BVRI photometry of QSO 0957+561A, B: observations, new reduction method, and time delay" *Astronomy and Astrophysics*, **526**, 40-51