

A BAYESIAN PERSPECTIVE TO THE STUDY OF EARTHQUAKES PROBABILISTIC MODELING OF EARTHQUAKES

ZACHARIE DUPUTEL

JOINT INVERSION IN GEOPHYSICS - 2015

OUTLINE

Introduction & Motivation:

Why Bayesian Analysis ?

Bayes theorem

Monte Carlo

- Hands-on exercice: Toy problem
- Bayesian inference in HD spaces and error model
- Application: Study of earthquakes
 - > 2013 Balochistan earthquake (Pakistan)

> 2014 Pisagua eartquake (Chile)

Explosion of available observations

- Global broad-band seismic network
- Dense seismic networks
- Satellite imagery
- GPS networks
- Tsunami





Explosion of available observations

- Global broad-band seismic network
- Dense seismic networks
- Satellite imagery
- GPS networks
- Tsunami

Higher-dimensional problems



Explosion of available observations

- Global broad-band seismic network
- Dense seismic networks
- Satellite imagery
- GPS networks
- Tsunami

Higher-dimensional problems



NOAA, USGS MOST model

Explosion of available observations

- Global broad-band seismic network
- Dense seismic networks
- Satellite imagery
- GPS networks
- Tsunami

Higher-dimensional problems



WHY BAYESIAN ANALYSIS ?

Geophysics is filled with ill-posed inverse problems

Traditional optimization:

- Regularized inverse picks one solution which fits data and regularization scheme
- Choice of regularization is usually based on convenience not physics
- Choice of regularization can have significant effects on the resulting solution

Izmit earthquake (1999) Yagi and Kikuchi (2001) Mw=7.4 (1.41x10²⁰ N-m) Rupture-speed=3.0 km/s Depth, km Sip. X==EW [km] Sekiguchi and Iwata (2002) Mw=7.4 (1.63x1019 N-m) Rupture-speed=3.0 km/s Slip, Depth, km X==EW [km]



WHY BAYESIAN ANALYSIS ?

Geophysics is filled with ill-posed inverse problems

Bayesian inference:

- Solution includes ensemble of all plausible models
 - ➡ Regularization is not required



Izmit earthquake (1999)



EMBRACE UNCERTAINTY

- When faced with a ill-posed problem, the best answer is to identify the ensemble of all plausible values for your model parameters
- "Plausible" models are those that satisfy the data and which agree with your prior assumptions about the model parameters

BAYESIAN INFERENCE AND ERROR MODEL



Shear-modulus, GPa

-star

Distance between observations (km)

BAYESIAN INFERENCE AND ERROR MODEL



- The "prior" is what you believe the likely values of the model parameters are in the absence of data
 - Can be used to impose physicsbased constraints



BAYESIAN INFERENCE AND ERROR MODEL



HOW TO DO BAYESIAN ANALYSIS

• Two options:

- 1. You can choose whatever priors and data likelihood you want and draw samples of $p(\mathbf{m}|\mathbf{d}_{obs})$ by Monte Carlo simulation
- 2. For certain problems, you can get an analytical expression for p(m|d_{obs}): For example: p(m) & p(d_{obs}|m) Gaussian → p(m|d_{obs}) Gaussian
- Option 1 may require a lot of computer power
- Option 2 requires brain power and possible compromise

APPROACH 1: MONTE CARLO SIMULATION

- Draw samples from the posterior PDF. The samples will have all the properties of the posterior PDF (mean, variance, etc.).
- There are only a few PDFs (e.g., uniform, Gaussian) that you can simulate directly with a random generator.
- Monte Carlo sampling looks like:
 - 1.Draw a random realization of your model from a proposal PDF
 - 2.Decide whether to keep or reject the candidate model (not all samples are accepted... can be very inefficient)
 - 3.Go back to step 1

METROPOLIS-HASTINGS

A memoryless random walk through the parameter space

• Choose:

- a starting vector of model parameters (m⁰)

- a proposal PDF ($q(\mathbf{m})$):

commonly: $q(\mathbf{m}|\mathbf{m}^{(j)}) \sim \mathcal{N}(\mathbf{m}^{(j)}, \boldsymbol{\Sigma})$

• Repeat for j=1,2,...,N

- Generate a candidate \mathbf{m}^* from $q(\mathbf{m}|\mathbf{m}^{(j)})$
- Calculate r = min{1, $p(\mathbf{m}^* | \mathbf{d}_{obs})/p(\mathbf{m}^{(j)} | \mathbf{d}_{obs})$ }
- Generate u from $\mathcal{U}(0,1)$
- If $u \leq r$
 - → Set $\mathbf{m}^{(j+1)} = \mathbf{m}^*$
- Else:
 - → Set $\mathbf{m}^{(j+1)} = \mathbf{m}^{(j)}$

 $m = [p_1, p_2]^T$



• { $\mathbf{m}^{(0)}, \mathbf{m}^{(1)}, \dots, \mathbf{m}^{(N)}$ } is a set of samples of the posterior PDF $p(\mathbf{m}|\mathbf{d}_{obs})$

QUESTIONS? COMING UP: TOY PROBLEM

TOY PROBLEM: LOCKED FAULT MODEL

Infinite strike-slip fault in a half-space

- Vertical fault
- Locked from the surface to depth H
- Deep dislocation: $H \rightarrow \infty$ (constant slip)

• Forward problem:

- $v(x) = S \cdot \tan^{-1}(x/H) / \pi$
- Velocity: v
- Slip rate: S
- Locking depth: H
- Distance from the fault: x
- Inverted parameters: $\mathbf{m} = [S, H]^T$

Posterior distribution:

- $p(\mathbf{m} \mid \mathbf{d}_{\text{obs}}) \propto p(\mathbf{d}_{\text{obs}} \mid \mathbf{m}) p(\mathbf{m})$
- Likelihood: Gaussian, Uncorrelated errors
- Prior: Uniform PDF
- Bayesian sampling using metropolis

Interseismic deformation



BURN-IN PERIOD

• Burn-in

- Refers to the practice of discarding an initial portion the samples so that the effect of initial values on the posterior is minimized



Red dots: burn-in samples

https://faculty.washington.edu/wjs18/Lecture%20PDFs/MCMC.pdf

TOY PROBLEM: LOCKED FAULT MODEL

"Hierarchical" Bayesian inversion

• Forward problem:

- $v(x) = S \cdot tan^{-1}(x/H) / \pi$
- Slip rate: S
- Locking depth: H
- Distance from the fault: x
- Inverted parameters: $\mathbf{m} = [S, H]^T$
- Hyper-parameter : σ_d
- (control the shape of $p(\mathbf{d}_{obs} \mid \mathbf{m})$)

Uncertain data noise:

- Posterior PDF:

- $p(\mathbf{m}, \sigma_{\rm d} | \mathbf{d}_{\rm obs}) \propto p(\mathbf{d}_{\rm obs} | \mathbf{m}, \sigma_{\rm d}) p(\mathbf{m}) p(\sigma_{\rm d})$
- Priors ($p(\mathbf{m})$, $p(\sigma_d)$): Uniform PDF
- Likelihood:

 $p(\mathbf{d}_{obs} | \mathbf{m}, \sigma_d) = [(2\pi)^{N/2} \sigma_d^N]^{-1} \exp[-||\mathbf{d}_{obs} - \mathbf{g}(\mathbf{m})|| / 2 \sigma_d]^2$



TOY PROBLEM: LOCKED FAULT MODEL

"Hierarchical" Bayesian inversion, bias in data/predictions

• Forward problem:

- $v(x) = S \cdot \tan^{-1}(x/H) / \pi$
- Slip rate: S
- Locking depth: H
- Distance from the fault: x
- Inverted parameters: $\mathbf{m} = [S, H]^T$
- Hyper-parameters : σ_d, μ_d
 (control the shape of p(d_{obs} | m))

Uncertain data noise:

- Posterior PDF:
 - $p(\mathbf{m}, \sigma_{d}, \boldsymbol{\mu}_{d} | \mathbf{d}_{obs}) \propto p(\mathbf{d}_{obs} | \mathbf{m}, \sigma_{d}, \boldsymbol{\mu}_{d}) p(\mathbf{m}) p(\sigma_{d}) p(\boldsymbol{\mu}_{d})$
- Priors: $p(\mathbf{m})$, $p(\sigma_d)$, $p(\boldsymbol{\mu}_d)$
- Likelihood:

 $p(\mathbf{d}_{obs} | \mathbf{m}, \sigma_d, \boldsymbol{\mu}_d) = [(2\pi)^{N/2} \sigma_d^N]^{-1} \exp [-\|\mathbf{d}_{obs} - \mathbf{g}(\mathbf{m}) - \boldsymbol{\mu}_d\| / 2 \sigma_d]^2$

e.g., 2nd order polynomial functions to remove SAR offsets due to possible orbital errors



http://funnel.sfsu.edu

QUESTIONS? COMING UP: MONTE CARLO IN H.D. SPACES

MONTE CARLO - CURSE OF DIMENSIONALITY

Monte Carlo simulation requires drawing enough samples to fill the model space

- Huge numbers of samples required for highdimensional problems

- One sample = One forward model evaluation

• Total numeric cost is:

Nsamples x Time(forward model)
For large number of model parameters, we need very fast forward model

A few of the many sampling algorithms in existence:

- Reject method: parallel but inefficient
- Metropolis algorithm: more efficient, but MCMC (e.g., a random walk) is serial
- **Gibbs sampling:** certain restrictions on types of PDFs that can be simulated
- Tempering and transitioning algorithms (e.g., **Parallel tempering**)



Good reference for parallel tempering: Sambridge, M., 2013. A Parallel Tempering algorithm for probabilistic sampling and multimodal optimization. Geophys. J. R. astr. Soc. 196, 357-374. doi:10.1093/gji/ggt342

MONTE CARLO - ALTAR

• Based on the CATMIP algorithm (Minson et al., 2013)

- Developed in collaboration with Caltech, IPGS, Géoazur
- A parallel tempered MCMC algorithm
- Embarrassingly parallel

GPU parallelization

IDEX project BAYESANR project BEBOP (JCJC)

Parallelization scheme



ALTAR benchmark for 288 model parameters



PREDICTION UNCERTAINTY

 $p(\mathbf{m}|\mathbf{d}_{\rm obs}) \propto p(\mathbf{m}) \, p(\mathbf{d}_{\rm obs}|\mathbf{m})$

 $p(\mathbf{m}|\mathbf{d}_{obs}) \propto p(\mathbf{m}) \int_D p(\mathbf{d}_{obs}|\mathbf{d}) p(\mathbf{d}|\mathbf{m}) d\mathbf{d}$

Stochastic model for the measurement process

- p(d*|d) is the probability (density) for getting the measured value d* when the uncertain physical quantity being measured has the value d
- p(d_{obs} | d) comes from substituting d* = d_{obs} in the probability model p(d*|d). It describes the likelihood of having observed d_{obs} if the actual displacement was d

Stochastic forward model for the predictions

- Uncertainties in the forward modeling
- Uncertainties of the theory
- Model discrepancy

PREDICTION UNCERTAINTY



Jolivet et al. (2013) https://profile.usgs.gov/mcnamara

PREDICTION UNCERTAINTY



Calculation of Cp based on the physics of the problem: a perturbation approach

$$\delta \mathbf{p} = \mathbf{K}_{\mu} \cdot \delta \ln \mu \qquad \qquad \mathbf{C}_{p} = \mathbf{K}_{\mu} \cdot \mathbf{C}_{\mu} \cdot \mathbf{K}_{\mu}^{T}$$

Partial derivatives w.r.t. the elastic parameters (sensitivity kernel)

Covariance matrix describing uncertainty in the Earth model parameters

PREDICTION UNCERTAINTY: TOY MODEL

Finite strike-slip fault

- Top of the fault at 0 km
- ▶ South-dipping = 80°

Earth model

Data for a layered half-space



Duputel et al., 2014

PREDICTION UNCERTAINTY: TOY MODEL

Depth, km



PREDICTION UNCERTAINTY: TOY MODEL

Finite strike-slip fault

- Top of the fault at 0 km
- South-dipping = 80°
- Data for a layered half-space





Posterior mean model, No Cp





Posterior mean model, including Cp

Duputel et al., 2014

QUESTIONS? COMING UP: APPLICATION EXAMPLES

Collaborations: R. Jolivet, M. Simons group (Caltech), L. Rivera (IPGS)

• Large earthquakes rarely occur within accretionary prisms

- aseismic deformation
- partitioning of deformation
- high pore pressures weak unconsolidated sediments
- Low seismicity:
 - Séisme de Quetta (1935, Mw~7.5)
 - Séisme de Makran (1945, Mw~8.1)
- Shallow locking depths:
 - Szeliga et al. (2012)
 - smaller than 5km on several faults

Low likelihood for a large earthquake in this region





Listric fault geometry

- Based on W-phase dip angles (δ =70° in the north, δ =50° in the south)
- Structural studies: décollement at 10km depth (e.g., Ellouz-Zimmermann et al., 2007)

East

West

- Depth dependent dip angle

Static Parameters:

- Slip vectors
- Nuisance parameters





Jolivet et al., 2013

Collaborations: M. Simons group (Caltech), L. Rivera (IPGS)

Occurred in the north chilean seismic gap

- Last ruptured in 1877 (Mw~8.8)
- Before 1877: unclear if the region failed in huge single ruptures or in sequences of smaller ruptures

• Foreshock activity:

- Started on 16 March with a Mw=6.7 thrust event
- Followed by thrust faulting aftershocks

Aftershock activity

• Mw=7.7 aftershock on 3 April 2014

<u>Goal</u>: Provide a reliable description of the mainshock source attributes



Observations



Duputel et al., in prep.

Curved fault geometry, 132 subfaults Static parameters

- Slip vectors
- Nuisance parameters

Kinematic Parameters:

- Hypocenter location
- Rupture velocity
- Rise time

20

50

100

Depth. km 200 150

250

300

350



S



Information integration using a cascading approach

Static inversion: p(m_s|d_s) ∝ p(m_s) p(d_s|m_s)
 Data (d_s): GPS, InSAR, Tide gauges, Tsunami
 Parameters (m_s): Slip vectors

² Kinematic inversion: p(m_s, m_k|d_s, d_k) ∝ p(m_s) p(d_s|m_s) p(m_k) p(d_k|m_s, m_k)
 ▶ Data (d_s, d_k): GPS, InSAR, Tide gauges, Tsunami, High-Rate GPS, Strong motion
 ▶ Parameters (m_s, m_k): Slip vectors, Hypocenter loc., Rupture velocity, Rise-Time

Minson et al., 2013



Kinematic slip inversion (+High-Rate GPS, Strong motion)

2

Static slip inversion (GPS, InSAR, Tide gauges, Tsunami)



Predictions



Joint static-kinematic model

- Localized slip that did not rupture up to the trench
- Foreshocks located up-dip of the main slip area



WHAT WE DISCUSSED

Motivation

-Big data

-Build new models for inter-, co- and post-seismic deformation

-e.g., Earthquakes, Slow-slip events

Issues and opportunities

-The Bayesian approach has many advantages over traditional optimisation solutions

-Solve ill-posed problems without smoothing

-Produce the ensemble of acceptable models

-Given sufficient computational ressources, Bayesian solution can be formed by Monte Carlo simulation