Parameter Estimation of Gravitational Wave Data

Chia-Jui Chou
National Yang Ming Chiao Tung University, Taiwan

2023/04/24

Detection of Gravitational Waves: Matched Filtering

If we have some knowledge of $h(t)$, by multiplying $s(t)$ and $h(t)$:

$$\frac{1}{T}\int_0^T s(t)h(t)dt = \frac{1}{T}\int_0^T h^2(t)dt + \frac{1}{T}\int_0^T n(t)h(t)dt,$$

where

$$\frac{1}{T} = \int_0^T h^2(t)dt \sim h_0^2, \quad \frac{1}{T}\int_0^T n(t)h(t)dt \sim \left(\frac{\tau_0}{T}\right)^{1/2} n_0 h_0.$$

Optimal Signal-to-Noise Ratio:

$$\left(\frac{S}{N}\right)^2 = 4\int_0^\infty \frac{|\tilde{h}(f)|^2}{S_n(f)}df, \quad \langle|\tilde{n}(f)|^2\rangle = \frac{1}{2}S_n(f)T.$$

Parameter Estimation of Gravitational Waves Signal

- There are models describing gravitational wave form the coalescence binary compact objects: IMRPhenom, SEOBNR, SpinTaylor, IMRSpinPrecEOB...
- From the strain data of an event, we estimate the most probable model and the parameters.
- Intrinsic Parameters: $m_1, m_2, a_1, a_2, \theta_1, \theta_2, \delta\phi, \phi_{jl}$.
- Extrinsic Parameters: $ra, dec, \theta_{jn}, \psi, d_L, \phi_c, t_c$.

Bayes' Theorem

Bayes' Theorem

$$p(\theta_i|d, H) = \frac{p(d|\theta_i, H)p(\theta_i|H)}{p(d|H)}.$$

- H: model,
- θ_i: parameters for the model H,
- d: observed data,
- $p(\theta_i|H)$: prior,
- $p(d|\theta_i, H)$: likelihood,
- $p(\theta_i|d, H)$: posterior,
- $p(d|H)$: evidence.

Bayes' Theorem

Marginalization

$$p(\theta_1|d,H) = \int_{\theta_2^{\min}}^{\theta_2^{\max}} \cdots \int_{\theta_N^{\min}}^{\theta_N^{\max}} p(\theta_1,\cdots,\theta_N|d,H)d\theta_2\cdots d\theta_N.$$

Evidence

$$p(d|H) = \int_{\theta_1^{\min}}^{\theta_1^{\max}} \cdots \int_{\theta_N^{\min}}^{\theta_N^{\max}} p(\theta_1,\cdots,\theta_N|d,H)d\theta_1\cdots d\theta_N.$$

Evaluating Posterior

Bayes' Theorem

$$p(\theta_i|d,H) = \frac{p(d|\theta_i,H)p(\theta_i|H)}{p(d|H)} = \frac{L(\theta_i)\cdot\pi(\theta_i)}{Z}.$$

- We don't know the normalization constant Z.
- Use Markov Chain Monte Carlo algorithms to generate samples from $L(\theta_i)\cdot\pi(\theta_i)$ in the parameter space.

$$L(\theta_i) = \mathcal{N}\exp\left\{-\frac{1}{2}\left(s-h(\theta_i)|s-h(\theta_i)\right)\right\},$$

where

$$(A|B) = 4\mathrm{Re}\int_0^\infty \tilde{A}^*(f)[S_n^{-1}(f)]\tilde{B}(f)df.$$

Markov Chain Monte Carlo Algorithm

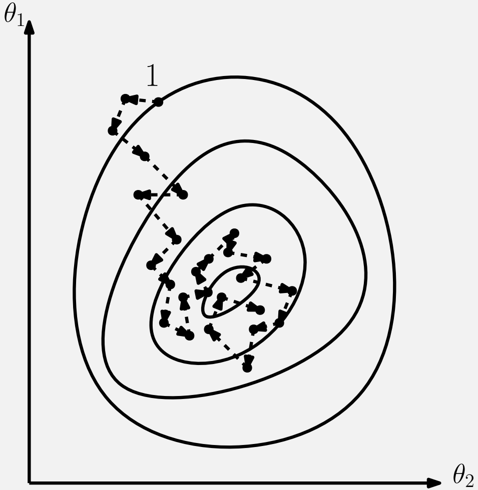
Generate a Markov Chain:

$$\left\{\theta_i^{(0)} \to \theta_i^{(1)} \to \theta_i^{(2)} \to \cdots \to \theta_i^{(N)}\right\},$$

The steps are stochastic and determined by the probabilities $T(\theta_i, \theta_i')$ associated with the trasition $\theta_i \to \theta_i'$.

- $T(\theta_i, \theta_i') \geq 0$.
- The posterior is an invariant distribution of the chain: $p(\theta_i'|d, H) = \int p(\theta_i|d, H) T(\theta_i, \theta_i') d\theta_i$.
- Detailed Balance: $p(\theta_i|d, H) T(\theta_i, \theta_i') = p(\theta_i'|d, H) T(\theta_i', \theta_i)$
- Ergodicity:
  - $\exists n$ such that $T^n(\theta_i, \theta_i') > 0$ for all θ_i, θ_i',
  - $\exists \theta_i$ such that $T(\theta_i, \theta_i) > 0$.

Markov Chain Monte Carlo Algorithm



Metropolis-Hasting Sampling

1. Starting at a random $\theta_i^{(0)}$ in the parameter space.
2. An update proposal θ_i' of $\theta_i^{(k)}$ is generated by sampling form a known proposal distribution $\tilde{p}(\theta_i')$ (e.g. Gaussian distribution centered at $\theta_i^{(k)}$).
3. Acceptance: $A(\theta_i^{(k)}, \theta_i') = \min\left(1, \frac{p(\theta_i')}{p(\theta_i^{(k)})}\frac{\tilde{p}(\theta_i^{(k)})}{\tilde{p}(\theta_i')}\right)$.
4. If $A(\theta_i^{(k)}, \theta_i') \geq 1$ (accepted): record $\theta_i^{(k+1)} = \theta_i'$, else:
   - $\theta_i^{(k+1)} = \theta_i'$ with the probability: $\frac{p(\theta_i')}{p(\theta_i^{(k)})}\frac{\tilde{p}(\theta_i^{(k)})}{\tilde{p}(\theta_i')}$.
   - $\theta_i^{(k+1)} = \theta_i^{(k)}$ with the probability: $1 - \frac{p(\theta_i')}{p(\theta_i^{(k)})}\frac{\tilde{p}(\theta_i^{(k)})}{\tilde{p}(\theta_i')}$.

Metropolis-Hasting Sampling

- Now the Markov chain we get: $\left\{\theta_i^{(0)} \rightarrow \cdots \rightarrow \theta_i^{(N)}\right\}$ can be considered as a correction of the proposal distribution $\tilde{p}(\theta_i)$ to the posterior distribution $p(\theta_i|d, H)$.
- Now Drawing the histogram plot of $\left\{\theta_i^{(0)}, \cdots, \theta_i^{(N)}\right\}$, we can see the distribution proportional to the posterior and find out the most probable parameters.

Nested Sampling

- MCMC methods: Generates samples proportional to the the posterior.
- Nested Sampling: Simultaneously estimates the evidence and the posterior.

Pros of Nested Sampling:

- well-defined stopping criteria for terminating sampling,
- generating a sequence of independent samples,
- flexibility to sample from complex, multi-modal distributions,
- the ability to derive how statistical and sampling uncertainties impact results from a single run,
- being trivially parallelizable.

Nested Sampling

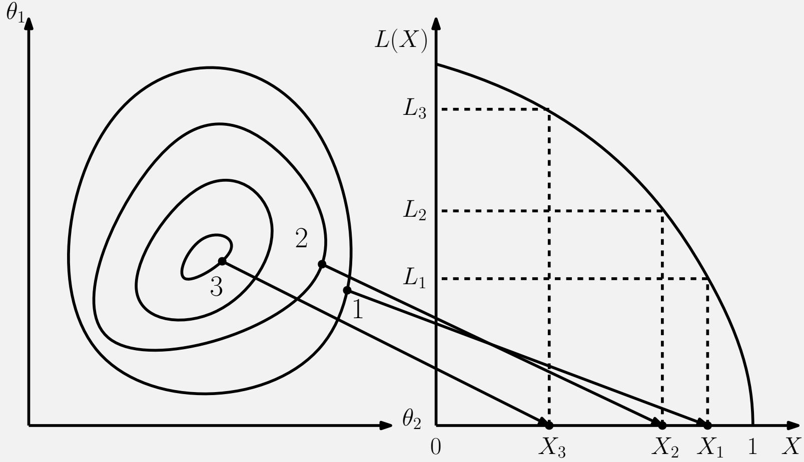
Prior Mass

$$X(\lambda) = \int_{\theta : L(\theta) > \lambda} \pi(\theta) d\theta.$$

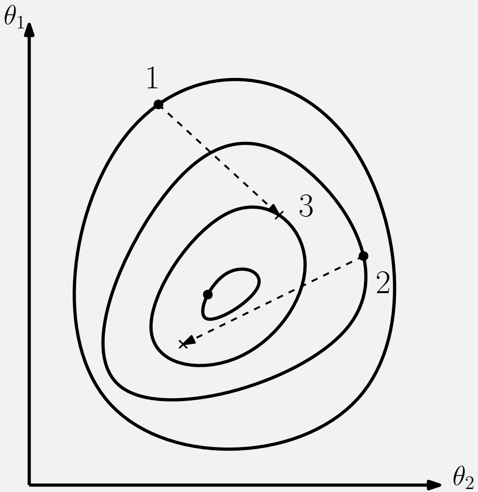
Evidence

$$Z = \int_0^1 L(X) dX, \quad \text{where} \quad L(X(\lambda)) = \lambda.$$

Nested Sampling



Nested Sampling



Nested Sampling

1. Sample N live points $\{\theta^{(1)} \cdots \theta^{(N)}\}$ from prior $\pi(\theta)$.
2. While not termination condition:
   1. record live point (i) with the lowest L_i as L_k,
   2. assign $X_k = t_k X_{k-1}$ where t_k from $P(t_k) = N t_k^{N-1}$,
   3. replace point (i) with sample from $\pi(\theta)$ subject to $L_i > L_k$.
3. Estimate evidence Z by integrating $\{L_k, X_k\}$.

Using Bilby for Parameter Estimation

1. Get strain data,
2. Estimate Power Spectral Density,
3. Determine Waveform model,
4. Set the prior distributions of the estimated parameters,
5. Set likelihood,
6. Set the sampler,
7. Run the sampler!
8. Plot the results.

Thank you